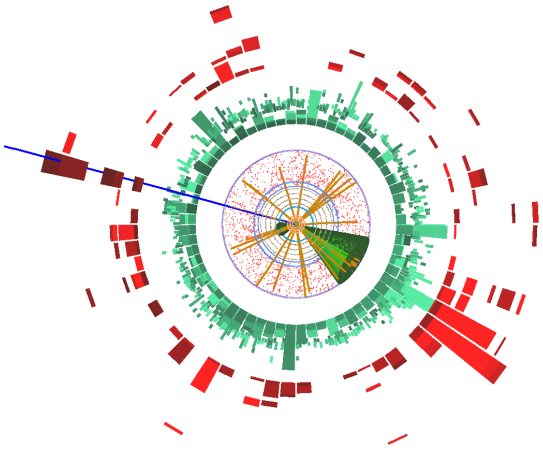
Measurement of the top quark mass in topologies enhanced with single top-quarks produced in the t-channel using flavour-tagging and a neural network with ATLAS data at $\sqrt{s} = 8\,\text{TeV}$

Dissertation

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



vorgelegt von:

Dipl. Phys. Hendrik Esch

Lehrstuhl für Experimentelle Physik IV
Fakultät Physik

technische universität
dortmund

Dortmund, Januar 2015

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

Measurement of the top quark mass in topologies enhanced with single top-quarks produced in the t-channel using flavour-tagging and a neural network with ATLAS data at $\sqrt{s} = 8$ TeV

vorgelegt von:

Dipl. Phys. Hendrik Esch

Lehrstuhl für Experimentelle Physik IV

Fakultät Physik

Technische Universität Dortmund

Dortmund, Januar 2015

Gutachter:	PD Dr. Reiner Klingenberg
Zweitgutachter:	Prof. Dr. Wolfgang Wagner
Leitung der Prüfungskommission:	Prof. Dr. Shaukat Khan
wiss. Mitarbeiter in der Prüfungskommission:	Dr. Christian Sternemann

Abstract

In this thesis a measurement of the top quark mass in topologies that have been enhanced with single-top quark decays in the t-channel produced via weak interactions is presented. The dataset consists of proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 8$ TeV collected with the ATLAS detector at the LHC with a total integrated luminosity of $\mathcal{L}_{\text{int}} = 20.3\,\text{fb}^{-1}$.

Selected events contain exactly one charged lepton – which can be either an electron or a muon –, missing transverse energy and two jets with exactly one of the two being b-tagged. The techniques of b-tagging used to identify jets induced by heavy quarks is explained further. In addition, the signal is enhanced using a neural network based discriminant that combines the ability to discriminate between signal and background of several correlated variables.

To determine the mass of the top quark a template method is used in combination with the mass sensitive variable, $m(\ell b)$, which is the invariant mass of the lepton and the b-tagged jet.

The top quark mass is measured as $m_{\text{top}} = [172.2 \pm 0.7\,(\text{stat.}) \pm 1.9\,(\text{syst.})]$ GeV.

Zusammenfassung

In dieser Dissertation wird die Messung der Masse des Top-Quarks in Topologien, die mit Zerfällen einzeln produzierter Top-Quarks angereichert wurden, gezeigt. Dabei handelt es sich um die Produktion im t-Kanal, die durch die schwache Wechselwirkung vermittelt wird. Der verwendete Datensatz enthält Proton-Proton-Kollisionen, die bei einer Schwerpunktsenergie von $\sqrt{s} = 8$ TeV mit einer integrierten Luminosität von $\mathcal{L}_{\text{int}} = 20.3\,\text{fb}^{-1}$ mit dem ATLAS Detektor am LHC aufgezeichnet wurden.

Die selektierten Ereignisse enthalten genau ein geladenes Lepton – entweder Elektron oder Myon –, fehlende transversale Energie und exakt zwei Jets, von denen genau einer b-getaggt sein muss. Diese b-tagging Techniken, mit denen Jets aus Zerfällen schwerer Quarks identifiziert werden können, werden näher erläutert. Darüber hinaus wird der Signalprozess mit Hilfe einer Diskriminante, die auf einem künstlichen neuronalen Netz basiert, angereichert. Dieses Netz kombiniert unterschiedlichste korrelierte Variablen zur Trennung von Signal und Untergrund.

Die Top-Quark-Masse wird mit Hilfe der Template-Methode bestimmt und nutzt die massenabhängige Variable $m(\ell b)$. Dabei handelt es sich um die invariante Masse des geladenen Leptons und des b-getaggten Jets.

Die gemessene Top-Quark-Masse ist $m_{\text{top}} = [172.2 \pm 0.7\,(\text{stat.}) \pm 1.9\,(\text{syst.})]$ GeV.

Contents

CHAPTER 1

Introduction

The top quark distinguishes itself from other elementary particles via its large mass, m_{top}, which is a fundamental parameter of the Standard Model. Since its discovery in 1995 at the Tevatron [1,2] various properties of the top quark have been measured, with one of the most precise quantities being the top quark mass m_{top}.

The recent world combination of measurements performed by the CDF and DØ experiments at the Tevatron collider and the ATLAS and CMS experiments at the Large Hadron Collider (LHC) yields $m_{\text{top}} = (173.34 \pm 0.27\,(\text{stat.}) \pm 0.71\,(\text{syst.}))$ GeV [3]. This result is based on an integrated luminosity of up to $8.7\,\text{fb}^{-1}$ of proton-antiproton collisions from Run II of the Tevatron at a centre-of-mass energy of 1.96 TeV and LHC data corresponding to up to $4.9\,\text{fb}^{-1}$ of proton-proton collisions from the run at a centre-of-mass energy of 7 TeV.

The most precise measurements of m_{top} have been performed with the production of a top-antitop quark pair, $t\bar{t}$, in the lepton+jets decay channel. In this final state one of the $W^{\pm}$-bosons from the two top quark decays is decaying into a charged lepton and a neutrino. The other $W^{\pm}$-boson decays into a quark-antiquark pair. Other significant contributions to the combination arise from the $t\bar{t} \to$ dilepton, $t\bar{t} \to$ all jets and $t\bar{t} \to \text{E}_{\text{T}}^{\text{miss}} +$ jets final states that distinguish themselves from the $t\bar{t} \to$ lepton + jets decay channel by different decay modes of the $W^{\pm}$-bosons. All decay channels have in common that a $t\bar{t}$ pair has been produced via the strong interaction by either gluon fusion or $q\bar{q}$ annihilation.

The LHC at CERN is providing high luminosities at high energies making it possible to study single top-quarks produced via weak, charged-current interactions. The dominant single top-quark production process, that is the main signal process under study in this thesis, is the t-channel exchange of a virtual $W^{\pm}$-boson. The predicted cross-section at a centre-of-mass energy of $\sqrt{s} = 8$ TeV for the sum of t and $\bar{t}$ at ap-

proximate next-to-next-to-leading order (NNLO) is $\sigma_t = (87.8^{+3.4}_{-1.9})\,\text{pb}$ assuming a top quark mass of $m_{\text{top}} = 172.5$ GeV [4]. The two sub-leading single-top quark production processes are associated production of a $W^{\pm}$-boson and a top quark and the s-channel production. These processes have a predicted cross-section of $\sigma_{Wt} = (22.4 \pm 1.5)\,\text{pb}$ [5] and $\sigma_s = (5.6 \pm 0.2)\,\text{pb}$ [6], respectively, assuming a top quark mass of $m_{\text{top}} = 172.5$ GeV at a centre-of-mass energy of $\sqrt{s} = 8$ TeV.

The analysis presented in this thesis makes use of topologies that have been enhanced with single top-quarks produced in the t-channel via the weak interaction to measure the mass of the top quark. This approach, which has never been applied before, uses a neural network discriminant to enhance the selected dataset with signal events and to reject non-reducible background with the same final-state signature.

In this analysis the leptonic decay channel of the $W^{\pm}$-boson is used ($W \to \ell\nu$). Thus, the complete process targeted to measure the top quark mass is $qb \to qt \to qW(\to \ell\nu)\,b$. Events are characterised by one isolated charged lepton (electron or muon), missing transverse momentum from the non-detectable neutrino and exactly two jets. One of the two jets is produced by the hadronisation of the b-quark from the top quark decay and another light-flavour jet arises from the t-channel production process. Also events from cascade decays of the $W^{\pm}$-boson are included, $W \to \tau\nu_\tau \to \text{e}(\mu)\nu_\tau\nu_{\text{e}(\mu)}$, leading to an electron or muon in the final state.

Apart from the t-channel production itself also the top-antitop quark pair production gives a significant contribution to the measurement while the Wt- and s-channel production processes only give a minor contribution.

The $qb \to qt \to qW(\to \ell\nu)\,b$ process used to measure m_{top} is characterised by three main differences compared to $t\bar{t}$ pair production used in previous measurements.

Firstly, the top quark is produced via the weak interaction instead of the $t\bar{t}$ production which is always mediated by the strong interaction. Because of the different colour structure of the two final states the analysis is sensitive to different sources of systematic uncertainties.

Secondly, there is no ambiguity in the assignment of the reconstructed jets to the initial partons reducing combinatorial background. Together with the expected presence of exactly one neutrino it allows the final state to be fully reconstructed with higher precision than in a phase-space dominated by $t\bar{t}$. This leads to a better overall mass resolution.

Finally, there is more irreducible background. This makes it necessary to apply multivariate techniques on top of a simple cut-based event selection. In contrast to this drawback there is no statistical correlation between the selected dataset and the datasets used

for the other measurements. This leads to very good prospects for a future world combination that includes measurement of the top quark mass in t-channel single top-quark topologies.

The variable sensitive to m_{top} is the invariant mass of the charged lepton and the b-tagged jet denoted as $m(\ell b)$. The technique applied is known as the template method [7] that uses templates of $m(\ell b)$ depending on the top quark mass. These templates estimated from simulation are fitted to data yielding the value of m_{top} that best describes the data.

This thesis is organised in the following way. In Chapter 2 the Standard Model of particle physics is described. It is the theoretical basis to describe the fundamental structure of matter and interacting forces within our Universe. The top quark itself is part of the Standard Model but explained in Chapter 3 focusing on its properties and previous measurements in more detail. In Chapter 4 the LHC, giving the possibility to produce top quarks, and the ATLAS detector, making it possible to measure the properties of the top quark in well-defined experimental conditions, are described. The way how different objects, which appear in a particle collision, are reconstructed with the ATLAS detector is described in Chapter 5. In Chapter 6 the identification of jets induced by b-quarks, referred to as flavour-tagging, is explained in detail. This divides into two parts. At first, different algorithms that are able to identify b-quark induced jets, called b-tagging algorithms, are described. Secondly, different measurements are presented to measure the performance of these b-tagging algorithms. The selected dataset and simulated samples used for modelling the different signal and background processes are described in Chapter 7. In Chapter 8 the cut-based event selection and background estimation are discussed, while the classification of events into signal- and background-like events is explained in Chapter 9. Here, the multivariate neural network technique is introduced that combines the separation power between signal and background that is distributed over a set of correlated variables into one powerful discriminant. In Chapter 10 the measurement of the top quark mass in the selected phase-space is presented. This includes an explanation of the measurement technique and a discussion of different sources of systematic uncertainties. The result of the measurement, a summary and a conclusion are given in Chapter 11.

1.1 Definitions of common variables

In this thesis some kinematic quantities will be used that are common when working with events collected by the ATLAS detector. They are motivated by the cylindrical

geometry of the detector which will be described in detail in Section 4.2. The goal of the event reconstruction is to estimate the full momentum 4-vector of physics objects produced in a collision. This 4-vector is defined as:

$$p^{\mu} = \begin{pmatrix} E = \sqrt{m^2 + p_x^2 + p_y^2 + p_z^2} \\ p_x \\ p_y \\ p_z \end{pmatrix}.$$

It consists of the three Euclidean momentum components p_x, p_y and p_z based on a right-handed coordinate system where the x-axis points to the centre of the LHC accelerator. The positive y-axis points upwards perpendicular to the plane of the LHC ring. The z-axis runs along the beam axis pointing towards the centre of the city of Geneva.

The coordinate system with the important quantities is shown in Figure 1.1. Together

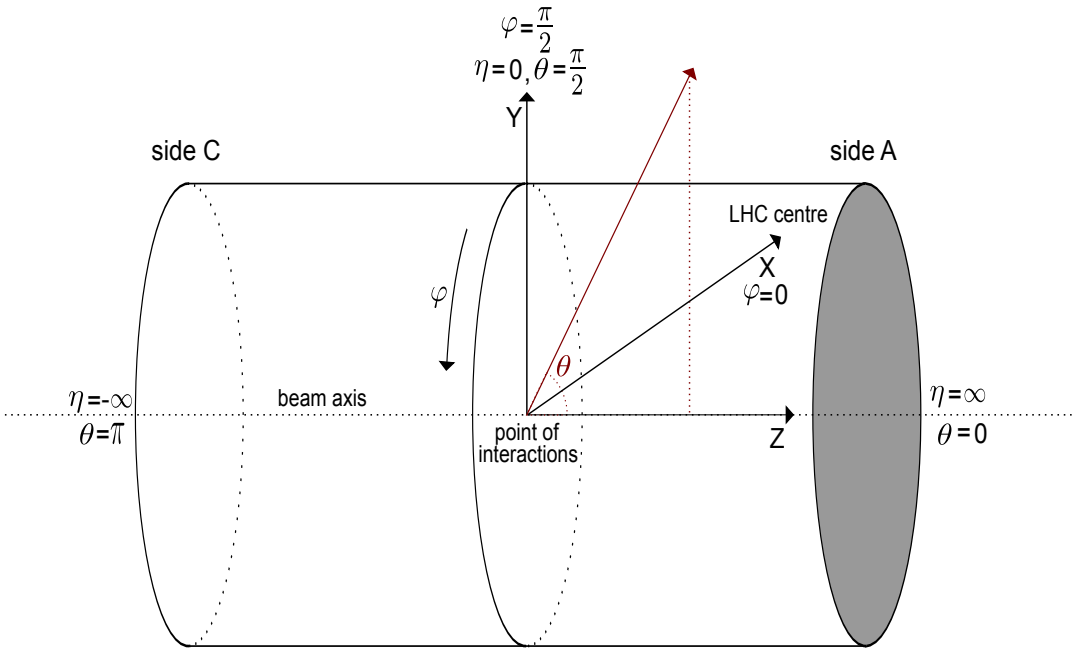


Figure 1.1: Right-handed coordinate system of the ATLAS detector.

with the rest mass m of the reconstructed particle object its energy is related using the relation that is known from special relativity $m^2 = p^{\mu}p_{\mu} = E^2 - p_x^2 - p_y^2 - p_z^2$. Out of

these 4-vector components angular relations can be constructed that will be useful in the later described physics analysis. These are:

$$\begin{aligned}
\text{transverse momentum:}\quad & p_T = \sqrt{p_x^2 + p_y^2} \\
\text{polar angle:}\quad & \theta = \arctan\left(\frac{p_T}{p_z}\right) \in [0, \pi] \\
\text{azimuthal angle:}\quad & \varphi = \arctan\left(\frac{p_x}{p_y}\right) \in (-\pi, \pi] \\
\text{pseudo-rapidity:}\quad & \eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \in (-\infty, \infty) \\
\text{rapidity :}\quad & y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \in (-\infty, \infty) \\
\text{distance in the } \eta\text{-}\varphi\text{-plane:}\quad & \Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2}
\end{aligned}$$

In addition, natural units are used in this thesis which means that formally $\hbar = c = 1$. This has the consequence that mass, energy and momentum are given in units of energy, i.e. $1\,\text{eV} = 1.602 \cdot 10^{-19}\,\text{J}$.

All Feynman diagrams shown follow the convention that the x-axis corresponds to the time axis. Anti-particles are travelling backwards in time indicated by the drawn arrows while explicit overbars are avoided.

CHAPTER 2

The Standard Model of Particle Physics

2.1 Overview over the Standard Model

The theoretical basis of this thesis is described by the Standard Model of particle physics which is also referred to as Standard Model. Since the 1930s thousands of theoretical and experimental physicists were trying to find a model that describes the fundamental structure of matter in our Universe. It was found that only a few building blocks of fundamental particles and four different forces are necessary to explain matter and how it interacts. The Standard Model relates these particles and three of the forces. It is able to explain almost all experimental results and make precise predictions about a huge variety of phenomena.

The components of the Standard Model are summarised in Figure 2.1 and will be explained in this chapter. They can be grouped into different parts. The fermionic quarks and leptons as matter particles with spin $\frac{1}{2}$ will be described in Section 2.2. Forces and their gauge bosons with spin 1 will be explained in detail in Section 2.3. In Section 2.4 the unification of the weak and electromagnetic force as two of the forces will be explained. The scalar Higgs particle is related to the origin of mass and will be described in Section 2.5. Section 2.6 shortly describes the structure of hadrons that are composite objects of two or three quarks bound together by the strong interaction.

2.2 Matter particles: Quarks and leptons

Matter appearing around us can be built out of two different building blocks that are called quarks and leptons. Each block has six different fundamental particles organised in three generations of paired particles. The first generation contains the lightest and therefore stable particles and builds up the matter we observe in everyday life. The second

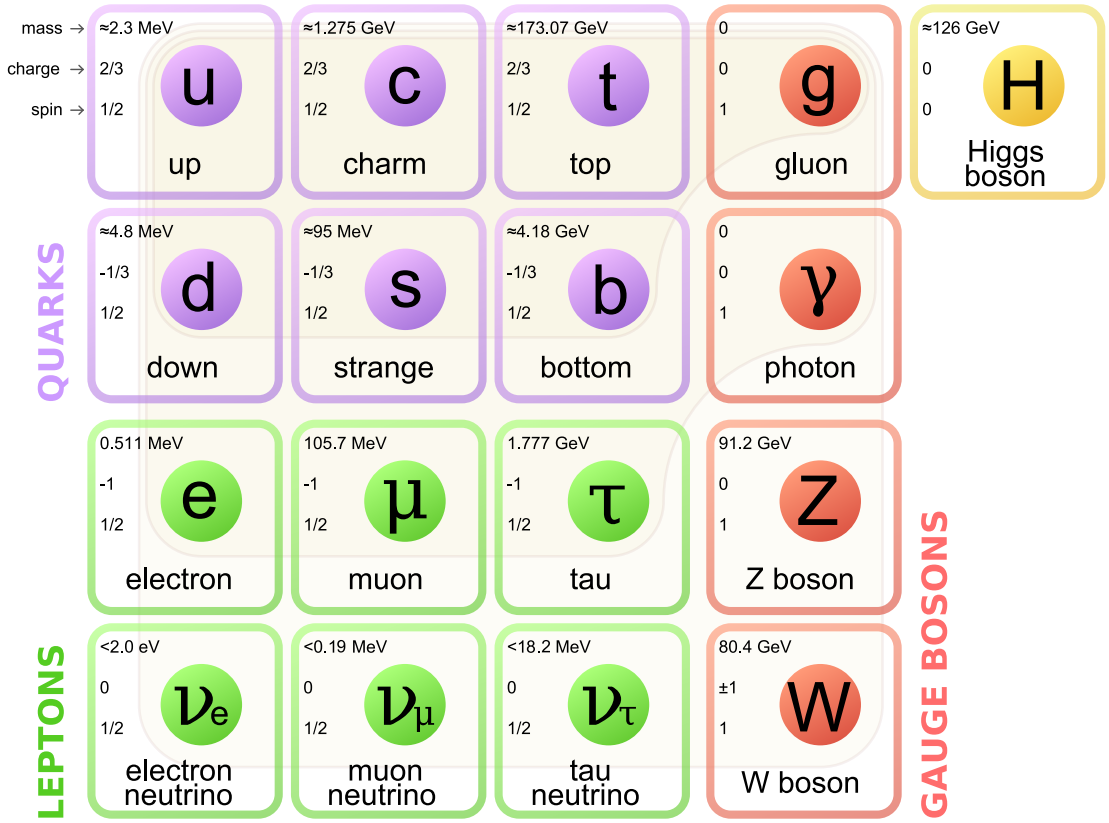


Figure 2.1: Overview of the Standard Model of particle physics [8, 9].

and third generation contain particles that are not stable and decay to the particles of the first generation.

2.2.1 Quarks

There are six different quarks called the up quark and the down quark in the first generation which build up protons, neutrons and atomic nuclei. Then, there is the strange quark and the charm quark in the second generation and bottom quark and top quark in the third generation that are produced for instance in high energy collisions at the LHC.

The difference of the quarks compared to the leptons is that quarks appear in three different colours - red (R), green (G) and blue (B) - following the $SU(3)_C$ symmetry group. Due to the strong interaction quarks are building colourless bound states that are singlets remaining unchanged by rotations in the $SU(3)_C$ space. Only three bound particle states of colourless singlets are known. Those states that need to be symmetrised

and normalised[1] are mesons, $R\bar{R} + G\bar{G} + B\bar{B}$, baryons, RGB, or antibaryons, $\bar{R}\bar{G}\bar{B}$. Here, $\bar{R}$, $\bar{G}$ and $\bar{B}$ correspond to the anticolours of the corresponding antiquarks that are explained later in this section.

Quarks all have a weak isospin of $I = \frac{1}{2}$ and the third component $I_3 = \pm\frac{1}{2}$ either has a positive or negative sign. So called down-type quarks have a negative sign while up-type quarks have a positive sign. Each pair in a generation therefore forms a so called weak isospin doublet. The third important quantity is the electric charge which causes quarks also to interact via the electromagnetic force. For up-type quarks this charge is $Q = +\frac{2}{3}\,e$ and for down-type quarks it is $Q = -\frac{1}{3}\,e$. All these important quantities of quarks are summarised together with the quark masses, which are growing with the generation, in Table 2.1. They are shown in Figure 2.1 in blue.

Table 2.1: Overview on the properties of the quark particles [9].

generation	1	2	3
name	up quark u	charm quark c	top quark t
mass m	$2.3^{+0.7}_{-0.5}$ MeV	1.275 ± 0.025 GeV	$173.07 \pm 0.52 \pm 0.72$ GeV
weak isospin (I, I_3)	$\left(\frac{1}{2}, +\frac{1}{2}\right)$	$\left(\frac{1}{2}, +\frac{1}{2}\right)$	$\left(\frac{1}{2}, +\frac{1}{2}\right)$
electric charge Q	$+\frac{2}{3}\,e$	$+\frac{2}{3}\,e$	$+\frac{2}{3}\,e$
name	down quark d	strange quark s	bottom quark b
mass m	$4.8^{+0.5}_{-0.3}$ MeV	95 ± 5 MeV	4.18 ± 0.03 GeV
weak isospin (I, I_3)	$\left(\frac{1}{2}, -\frac{1}{2}\right)$	$\left(\frac{1}{2}, -\frac{1}{2}\right)$	$\left(\frac{1}{2}, -\frac{1}{2}\right)$
electric charge Q	$-\frac{1}{3}\,e$	$-\frac{1}{3}\,e$	$-\frac{1}{3}\,e$

The previously explained quark model was developed over decades starting with the up, down and strange quark in the early 1960s. First of all bound states of quarks called hadrons were ordered in isospin multiplets by Murray Gell-Mann [11] and Yuval Ne'eman [12]. This classification scheme was called the Eightfold Way and followed a SU(3) symmetry but the underlying structure remained unknown. Gell-Mann [13] and Zweig [14, 15] then independently proposed the quark model starting with the three quarks in 1964. Finally, in 1968 at the Stanford Linear Accelerator Centre direct evidence could be achieved [16, 17]. It was shown using deep inelastic scattering that protons do have a substructure and a set of three fundamental particles could explain the data.

The charm quark had already been predicted by many authors like Bjorken and Glashow in 1964 [18] and then in the context of the so called GIM mechanism by

[1]The correctly symmetrised and normalised states are $\frac{1}{\sqrt{3}}(R\bar{R} + G\bar{G} + B\bar{B})$, $\frac{1}{\sqrt{6}}(RGB - RBG + BRG - BGR + GBR - GRB)$ and $\frac{1}{\sqrt{6}}(\bar{R}\bar{G}\bar{B} - \bar{R}\bar{B}\bar{G} + \bar{B}\bar{R}\bar{G} - \bar{B}\bar{G}\bar{R} + \bar{G}\bar{B}\bar{R} - \bar{G}\bar{R}\bar{B})$ [10].

Glashow, Iliopoulos and Maiani in 1970 [19]. It was finally discovered independently by experimentalists at the Stanford Linear Accelerator Centre [20] and the Brookhaven National Laboratory [21] in 1974 as a bound charm-anticharm state called J/Ψ[2].

The bottom quark was firstly mentioned in theory in 1973 by Kobayashi and Maskawa with the intention to explain the origin of CP-violation [22]. It was then discovered in 1977 by the E288 experiment located at Fermilab where bottom-antibottom quark pairs were produced [23, 24].

The existence of the bottom quark as a third generation quark already implied that the top quark would exist but it took another 18 years until it was finally discovered. The reason for this was that the top quark is very heavy reaching a mass that is comparable with the mass of a gold atom. Accelerators that were built to discover it were not able to reach these energies. It was then finally discovered by the two experiments DØ and CDF at the Tevatron in 1995 [1, 2] where top-antitop quark pairs have been produced. First evidence that top quarks appear also in single-top quark production was found in December 2006 [25]. The measurement of the mass of the top quark in single-top quark production in the so called t-channel has been performed for the first time within the scope of this thesis. This will be explained in detail in the later chapters.

2.2.2 Leptons

The second block of matter particles contains the leptons shown in green in Figure 2.1. As well as the quarks they can be organised as weak isospin doublets in three different generations, but leptons do not carry colour charge and do not interact via the strong interaction. In this case each doublet is built out of one lepton with $I_3 = -\frac{1}{2}$ that carries an electric charge of $Q = -1\,e$ and an associated electrically neutral neutrino with $I_3 = +\frac{1}{2}$ that only interacts by the weak interaction.

The leptons are called the electron e^- and electron-neutrino ν_e in the first generation, the muon μ^- and its muon-neutrino ν_μ in the second generation followed by the τ-lepton τ^- together with the τ-neutrino ν_τ. The charged leptons do have a sizeable mass whereas the mass of the neutrinos is very small. It was long thought that neutrinos would even be massless, but with direct observations of oscillations between the different neutrino flavours it was indirectly proven that neutrinos must have a mass different from zero [26]. Table 2.2 summarises the explained quantities of the leptons and their masses which are as for the quarks growing with the generation.

[2]The J/Ψ particle is the only particle with a two-letter name. The reason is the almost simultaneous discovery by the two groups that first intended to name the particle either J or Ψ.

The electron that appears in our daily life related to electric currents was already discovered by J.J.Thomson in 1897 [27] and its associated electron-neutrino was proposed by Pauli in 1930 to solve the problem of preserved energy, momentum and angular momentum conservation in beta decays [28]. The electron-neutrino was then discovered by Cowan and Reines in 1956 using a nuclear reactor as the source of the neutrino flux [29].

Table 2.2: Overview on the properties of the lepton particles [9].

generation	1	2	3
name	electron-neutrino ν_e	muon-neutrino ν_μ	τ-neutrino ν_τ
mass m in MeV	$< 2 \cdot 10^{-6}$	< 0.19(CL = 90%)	< 18.2(CL = 90%)
weak isospin (I, I_3)	$\left(\frac{1}{2}, +\frac{1}{2}\right)$	$\left(\frac{1}{2}, +\frac{1}{2}\right)$	$\left(\frac{1}{2}, +\frac{1}{2}\right)$
electric charge Q	0	0	0
name	electron e^-	muon μ^-	τ-lepton τ^-
mass m in MeV	$0.511 \pm 1.1 \cdot 10^{-8}$	$105.7 \pm 3.5 \cdot 10^{-6}$	1776.82 ± 0.16
weak isospin (I, I_3)	$\left(\frac{1}{2}, -\frac{1}{2}\right)$	$\left(\frac{1}{2}, -\frac{1}{2}\right)$	$\left(\frac{1}{2}, -\frac{1}{2}\right)$
electric charge Q	$-1\,e$	$-1\,e$	$-1\,e$

The muon was discovered in 1936 by Anderson using showers of cosmic rays and was first misidentified as a bound state of a quark-antiquark pair [30]. Later it was found that the muon does not interact via the strong force which was the time where the group of leptons was firstly created with the three known particles. In 1962 it was shown by Lederman, Schwartz and Steinberger at the Alternating Gradient Synchrotron that more than one flavour of neutrinos must exist corresponding to the muon-neutrino [31].

Between 1974 and 1977 various experiments were made by the SLAC LBL group led by Perl and evidence for the existence of the τ-lepton was found in 1975 in collisions at the SPEAR ring [32]. At that time an associated τ-neutrino was already expected and as the last part of the matter particles this τ-neutrino was then found in 2000 by the DONUT collaboration at Fermilab [33].

Both quarks and leptons are fermions with spin $\frac{1}{2}$ and can be described by the relativistic Dirac equations. The solution of this equation yields that there must be a second set of particles with the same mass but negative energies. These can also be interpreted as anti particles with positive energies moving backwards in time and therefore fermions get an associated anti particle[3]. It has a flipped sign in all additive

[3]For the particle, p, the associated anti particle, $\overline{p}$, is commonly denoted with an overbar.

quantum numbers like the electric charge, Q, or third component of the weak isospin, I_3, i.e. charge-like quantum numbers.

2.3 Forces and gauge bosons

Four different forces are known to interact in our Universe. They are known as the gravitational force, the electromagnetic force, the weak force and the strong force. Except for the gravitational force all forces are part of the Standard Model. They are mediated by gauge bosons with spin 1 that carry discrete amounts of energy between fundamental particles. In Section 2.3.1 the fundamental interactions will be described while Section 2.3.2 will deal with the mediating gauge bosons.

2.3.1 Fundamental interactions

The most familiar force from our daily life but also weakest is the gravitational force which has an infinite range and is always attractive. This is why it becomes important on macroscopic scales when large amounts of matter accumulate. However, on very small scales like particle physics the gravitational force can be neglected. It is about 36 orders of magnitude smaller than for instance the electromagnetic force. The electromagnetic force also has an infinite range but can be attractive as well as repulsive. Both the weak and the strong force only have a short range and are only important on subatomic scales.

The gauge boson mediating the electromagnetic force is the photon - also denoted γ - that couples to electrically charged particles. The gauge bosons of the weak force are the two either positively or negatively charged $W^{\pm}$-bosons and the Z^0-boson which is electrically neutral. The $W^{\pm}$- and Z^0-boson couple to particles carrying weak isospin.

The strong force has eight different so called gluons that couple to all particles carrying colour charge.

A hypothetical boson called Graviton is a potential candidate often mentioned as the gauge boson of the gravitational force. However, efforts to combine the general theory of relativity that is important on macroscopic scale like the gravitational force and quantum mechanics are challenging and even lead to inconsistencies between these theories. That is why the gravitational force could not be integrated in the Standard Model. As the gravitational force is so weak, the Standard Model is still able to describe all phenomena observed in particle physics.

The strength of each interaction is represented by a coupling constant which is the fine structure constant $\alpha \approx \frac{1}{137}$ in the case of the electromagnetic force. The weak force

appears to have a smaller coupling compared to the electromagnetic force because of the large mass of its mediating gauge bosons. However, as will be explained in Section 2.4 its coupling is actually stronger. The coupling strength of the strong interaction can be of the order of $O(1)$ for low energies. It is the only force that increases with growing distance which leads to the Confinement that binds quarks to bound states, as mentioned in Section 2.2.1. It also exists an asymptotic freedom at very short distances of colour charged particles where also quarks are seen as free particles [34]. Table 2.3 summarises the four different forces and their properties.

Table 2.3: Overview of the four different forces

Force	Gauge-Boson	couples to	range [m]
gravitational	graviton	mass m	∞
electromagnetic	photon γ	electric charge Q	∞
weak	$W^{\pm}$, Z^0	weak isospin I	O(10^{-18})
strong	gluon g	colour charge Q_C	O(10^{-15})

2.3.2 Gauge bosons

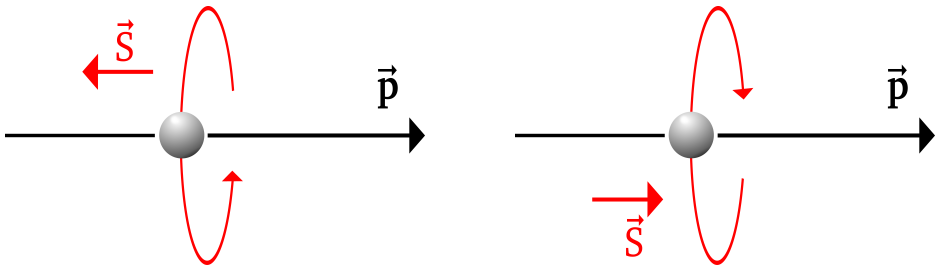
The photon itself is a massless particle that neither carries electric or colour charge. It also has a weak isospin of 0 and therefore cannot interact itself via the three fundamental interactions. As the particle we see as light it was studied by a large variety of physicists starting with Maxwell [35] within the theory of electrodynamics, Planck in the context of black-body radiation [36] or Einstein who finally explained the photoelectric effect [37]. He added the energetic quantisation as a property of the photon radiation itself. This was not part of Maxwell's theory who interpreted the photon as a light wave. The quantisation was leading to the photon as a particle that is able to move around carrying quantised amounts of energy. This was then confirmed experimentally by Compton in 1923 [38] and is now very well described by the theory of quantum electrodynamics [39] as a part of the Standard Model.

The $W^{\pm}$- and Z^0-bosons as the gauge bosons of the weak interaction all have a weak isospin of $I = 1$ while the third component is different for the three particles. The electrically charged $W^{\pm}$-bosons have $I_3 = \pm 1$ while the Z^0-boson has $I_3 = 0$. They do not carry colour charge and therefore do not interact via the strong force. But as the only gauge bosons they are not massless. The large masses of $m_{W^{\pm}} = 80.4\,\mathrm{GeV}$ [9] and

$m_{Z^0} = 91.2\,\text{GeV}$ [9] are responsible for the limited range of the weak interaction. It can be estimated using the uncertainty principle $\Delta E \Delta t \geq \frac{\hbar}{2}$ [40] following Equation (2.1):

$$\Delta s = c \cdot \Delta t = \frac{c\hbar}{2\Delta E} \approx 10^{-18}\,\text{m}. \tag{2.1}$$

Another speciality of the weak force and its gauge bosons is that they only couple to particles with left-handed helicity and antiparticles with right-handed helicity[4]. The definition of helicity is illustrated in Figure 2.2. Right-handed means that the spin is pointing in the same direction as the momentum of the particle and left-handed vice versa respectively. This so called V-A or left-handed coupling[5] directly implies the violation of



(a) Left-handed particle (b) Right-handed particle

Figure 2.2: Left-handed (a) and right-handed (b) particles where $\vec{p}$ is intended to represent the momentum and $\vec{S}$ the spin orientation of the particle.

the parity symmetry which was thought to be a fundamental symmetry until the 1950s. The parity violation was proven to exist by Wu in 1957 [41]. After this discovery it was expected that the compound symmetry of charge C and parity P called CP would be conserved but in 1964 it was proven in kaon decays by Cronin and Fitch that also CP is violated by the weak interaction [42]. The existence of both the $W^{\pm}$- and Z^0-bosons was proven by the UA1 and UA2 experiments at the Super Proton Synchrotron located at CERN in 1983 [43].

Gluons as the gauge bosons of the strong force are massless, electrically neutral and have a weak isospin of 0 which means that they do not interact via the electromagnetic

[4]For particles with mass it is possible to boost the particle in another reference frame which makes it possible to flip the helicity. This is taken into account in the lorentz-invariant definition of chirality. So for massless particles the chirality and helicity are the same.

[5]Another consequence is that right-handed neutrinos do not exist. The isospin doublets, introduced in Section 2.2.2, can therefore be divided in left-handed doublets and a right-handed singlets without the neutrinos.

or weak interaction. They are carrying a combination of colour and anticolour and are organised in a $SU(3)_C$ octet with the eight gluon states being $R\bar{G}$, $R\bar{B}$, $G\bar{R}$, $G\bar{B}$, $B\bar{R}$, $B\bar{G}$, $\frac{1}{\sqrt{2}}(R\bar{R} - G\bar{G})$ and $\frac{1}{\sqrt{6}}(R\bar{R} + G\bar{G} - 2B\bar{B})$. Since gluons are not colourless they not only mediate the strong force between quarks, but are also able to interact between themselves[6]. First evidence of their existence was reported in 1978 by the PLUTO experiment in $e^+e^- \to \Upsilon(9.46) \to 3g$ decays [44] and in 1979 at PETRA using $e^+e^- \to q\bar{q}g$ events [45].

The gauge bosons are shown in Figure 2.1 in red and the coupling to the different fundamental particles is indicated with the yellow shaded areas.

2.3.3 Quark mixing and the CKM mixing matrix

Another anomaly related to quarks is that they carry weak isospin but exist as eigenstates of the strong interaction in space-time. This equivalently means that the mass eigenstates of quarks, $|q\rangle$, are not the same as the eigenstates of the weak interaction, $|q'\rangle$. To relate these eigenstates Cabbibo, Kobayashi and Maskawa introduced the CKM quark mixing matrix where the mass eigenstates of down-type quarks exist as a mixture of their weak eigenstates while the up-type quarks remain unchanged [22, 46]. This relationship can be expressed as follows:

$$\begin{pmatrix} |d'\rangle \\ |s'\rangle \\ |b'\rangle \end{pmatrix} = V_{CKM} \cdot \begin{pmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{pmatrix}. \tag{2.2}$$

The entries of the CKM matrix act as free parameters of the Standard Model and have mostly been measured by various experiments. One can also exploit that the matrix is an unitary matrix with complex elements. All together this yields for the absolute values [9]:

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97425 \pm 0.00022 & 0.2252 \pm 0.0009 & (4.15 \pm 0.49)\,10^{-3} \\ 0.230 \pm 0.011 & 1.006 \pm 0.023 & (40.9 \pm 1.1)\,10^{-3} \\ (8.4 \pm 0.6)\,10^{-3} & (42.9 \pm 2.6)\,10^{-3} & 0.89 \pm 0.07 \end{pmatrix}.$$

The matrix is strongly diagonal dominant but has non vanishing elements on the off-diagonal. This has far-reaching consequences for particle physics since on the one hand it introduces the possibility of transitions between different quark generations. On the other hand a transition between two quarks $q \to q'$ is suppressed by the respective

[6] A hypothetical colourless singlet gluon would be unconfined and act like a strong interacting photon. This would lead to an infinite range of the strong force which is not observed in nature [10].

matrix element $|V_{qq'}|^2$. This principle can for instance be exploited to identify particle jets induced by b-quarks which will be explained in more detail in Chapter 6. It also plays an important role in decays of the top quark which will be explained in Chapter 3.

In addition, the CKM quark mixing matrix has a complex phase. This phase includes the possibility of CP-violation of the weak interaction[7], mentioned in Sections 2.2.1 and 2.3.2, but it does not predict its magnitude.

2.4 The electroweak unification

In experiments with decays of the Z^0-boson it was found that the coupling to either quarks or leptons seems to be different. This was surprising since they both have the same weak isospin. Furthermore, it turned out that the coupling depends on the strength of the electric charge of the quark or lepton. This inconsistency could be solved with the unification of the electromagnetic and the weak force that was already proposed by Glashow, Salam and Weinberg in the 1970s [47, 48].

The gauge bosons of the electroweak force are the massless W_1, W_2, W_3 and B^0 that couple to the combination of weak isospin and electric charge. This is called weak hypercharge and is defined as $Y = 2(Q - I_3)$. What can be observed are the already explained $W^{\pm}$-, Z^0-bosons and the photon. Those are formed as a mixture of the gauge bosons of the electroweak theory caused by the electroweak symmetry breaking induced by the Higgs mechanism:

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos\theta_w & \sin\theta_w \\ -\sin\theta_w & \cos\theta_w \end{pmatrix} \begin{pmatrix} B^0 \\ W_3 \end{pmatrix} \tag{2.3}$$

$$\begin{pmatrix} W^+ \\ W^- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \end{pmatrix} \tag{2.4}$$

The amount of mixing is set by the electroweak mixing angle θ_w. It is directly related to the mass difference of the $W^{\pm}$- and Z^0-boson caused by the Higgs mechanism [49–51] that will be explained in Section 2.5. The masses of the weak gauge bosons have already been measured with high precision as explained in Section 2.3.2. With the relationship $m_{W^{\pm}}/m_{Z^0} = \cos\theta_w$ the mixing angle is known to be $\theta_w \approx 28.74°$ [9]. From this angle

[7] This indirectly implies the existence of a third generation of quarks since at least three dimensions are needed to introduce such a complex phase.

also the relationship between the coupling[8] of the electromagnetic and weak force can be derived with $e = g \sin^2 \theta_w$.

2.5 Higgs-Mechanism and the origin of particle mass

The unification of the weak and electromagnetic force explained in the previous section was a major success but still one caveat was remaining. The masses of the gauge bosons are necessary for the proposed unification mechanism and to keep the Standard Model renormalisable. Also the experimental observation yields that these gauge bosons should be massive. The problem was solved by Brout, Englert and Higgs in 1964 who introduced a spontaneously broken Higgs field that pervades the Universe.

Spontaneous symmetry breaking means that a system has a global symmetry which is broken in the lowest energy state. This is the case for the complex Higgs field ϕ with the spontaneously broken potential given in Equation (2.5) that is also displayed in Figure 2.3:

$$V(\phi) = -\mu\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \tag{2.5}$$

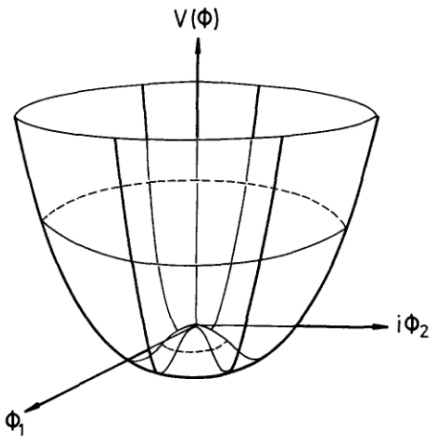


Figure 2.3: Potential of the complex Higgs field given by Equation (2.5) [52].

In quantum field theories in general the existence of a field compellingly implies the existence of an associated particle that can be interpreted as excitation of the field. The

[8]The coupling constants are $\alpha_w \propto g^2$ and $\alpha \propto e^2$ showing that the coupling of the weak force is actually stronger than for the electromagnetic force and only appears to be smaller due to the large mass of its gauge bosons.

scalar Higgs boson associated to the Higgs field is predicted for almost half a century since the mechanism was proposed. It was finally discovered in 2012 by the ATLAS and CMS experiments at the LHC [53, 54] at a mass of $m_H = (125.9 \pm 0.4)$ GeV [9]. The Higgs boson is shown as a part of the Standard Model in Figure 2.1 in yellow.

The masses of the gauge bosons are generated directly by introducing the spontaneous symmetry breaking due to the Higgs field. The masses of the fermions are generated via a Yukawa coupling with a coupling strength $\tilde{g}_f$. The mass m_f is depending on the coupling strength and the vacuum expectation value of the Higgs field v with the relation $\tilde{g}_f = \frac{m_f\sqrt{2}}{v}$ [52].

2.6 Structure of hadrons

As described in Section 2.2.1 quarks do only exist as bound states called hadrons. Hadrons exist in two types which are mesons as a quark-antiquark pairs or baryons which are either three quarks or three antiquarks bound together. One example of a baryon is the proton which is a bound state of two up quarks and one down quark. However, what can be seen is that the sum of the masses of the three so called valence quarks of the proton does not match up with the proton mass of $m_p = 938\,\text{MeV}$ observed in data.

In experiments with electron-nucleon scattering [16, 17] it was found that hadrons also contain so called sea quarks that are quark-antiquark pairs of the same flavour that are produced and annihilate each other in fluctuations of the vacuum. All together a hadron is a complex object made out of the valence quarks, that already determine all quantum numbers, sea quarks and additional gluons. This object cannot be described by perturbative theories. The aggregate of valence quarks, sea quarks and gluons within a hadron is also referred to as partons.

In particle collisions using hadrons like proton-proton collisions at the LHC it is important that the initial state can be described in simulation. Because of that parton distribution functions $f_i(x, Q)$, in short called PDF, are used. A PDF is a probability density function giving the probability of a certain parton i to have the momentum fraction x of the total momentum of the hadron at a certain energy scale Q. In a proton-proton collision the actual scattering is not occurring between the incoming protons but between the incoming partons. This is resulting in a usually unknown and much lower collision energy than beam energy which is the main caveat when using hadrons in particle colliders.

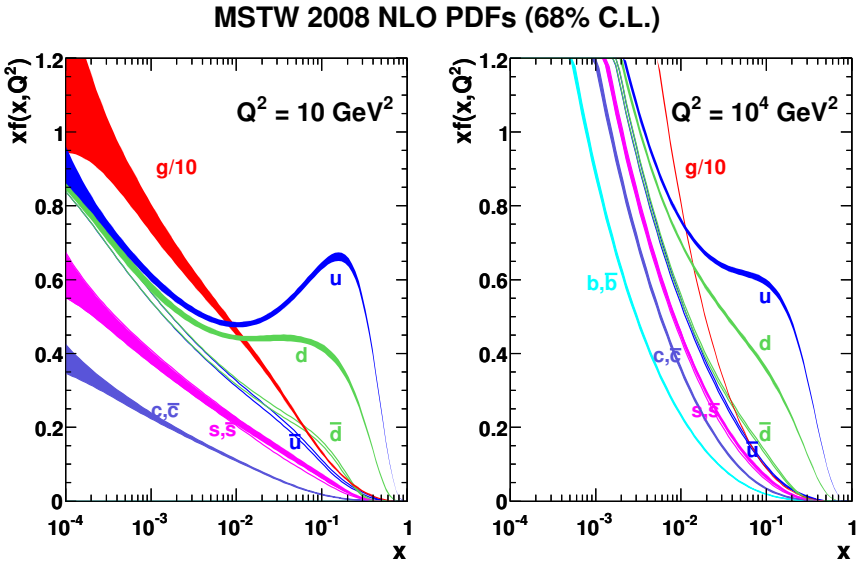


Figure 2.4: Parton distribution functions provided by the MSTW group at $Q^2 = 10$ GeV^2 (left) and $Q^2 = 10^4$ GeV^2 (right). The contribution of gluons is suppressed by a factor of 10 for a better visibility. Uncertainties of one standard deviation are shown as shaded areas [55].

Since PDFs can only be described non-perturbatively they are parametrised and then fitted with data mostly from the HERA accelerator at DESY [56] and the Tevatron at Fermilab [57,58]. The most common PDF set used by the ATLAS collaboration is given by the CTEQ group [59]. CTEQ provides as well a certain value of the momentum fraction, x, depending on the energy scale and parton flavour and the uncertainty in the form of a set of uncorrelated eigenvectors. PDF sets derived from other parametrisations are also provided by the MSTW [55] group or NNPDF [60]. All three are taken into account when calculating systematic uncertainties due to the parton distribution function as it is described in Section 10.4.2 for the analysis that is the topic of this thesis. As an example, the PDF set of the MSTW group is shown at two different energy scales in Figure 2.4.

CHAPTER 3

The top quark

The top quark, already introduced in Section 2.2, is the heaviest known elementary particle and forms a weak-isospin doublet together with the bottom quark. Due to its very high mass the top quark has special properties distinguishing it from the other known quarks.

As the topic of this thesis is a measurement of the mass of the top quark in a special production mode, an overview of the production, decay and the different properties of the top quark is given. Also the methods that have been used in the past to measure the mass of the top quark will be explained focusing on the differences compared to the analysis presented in this thesis.

3.1 Production of the top quark

The production of the top quark can be separated into two main classes. Firstly, top quarks can be produced as top-antitop quark pairs due to the strong interaction and, secondly, single top-quark production can occur in three different channels. This single-top quark production can only occur via the weak interaction due to the conservation of the flavour quantum numbers by the other interactions. Since the weak coupling is suppressed with respect to the strong interaction the dominant process occurring in hadron collisions like at the LHC is $t\bar{t}$ pair production. That is why most properties of the top quark have been measured in $t\bar{t}$ events so far. However, the single-top production has some advantages that will be explained in more detail later.

3.1.1 Top-antitop quark pair production

A top-antitop quark pair, $t\bar{t}$, can be produced in two different subprocesses which are displayed in Figure 3.1. The dominant process at the LHC is the gluon fusion $gg \to t\bar{t}$

production shown in (a) to (c) while at the Tevatron the $q\bar{q}$ annihilation $q\bar{q} \to t\bar{t}$ shown in (d) was the dominant production mode[9].

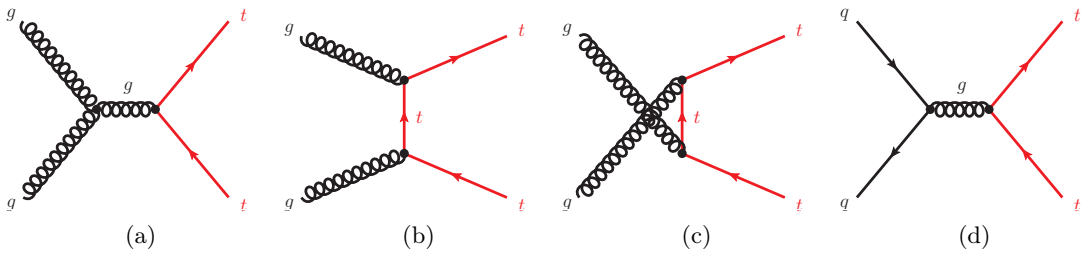
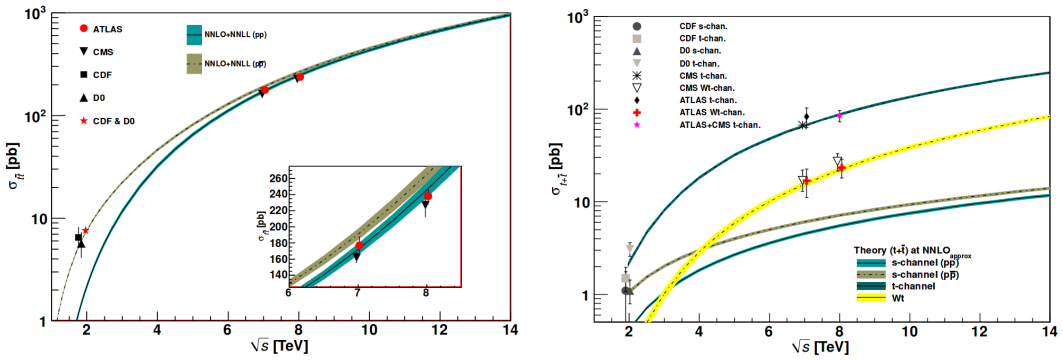


Figure 3.1: Feynman diagrams of $t\bar{t}$ production processes: (a) to (c) gluon fusion and (d) $q\bar{q}$ annihilation.

The total cross-section[10] of the top-antitop quark pair production is predicted to be $\sigma_{t\bar{t}} = (253^{+13}_{-15})$ pb [61–66] at a centre-of-mass energy of $\sqrt{s} = 8$ TeV at an assumed top quark mass of $m_{\text{top}} = 172.5$ GeV. A large variety of measurements of the cross-section have been performed by different experiments at the Tevatron and the LHC. All of them are in very good agreement with the predicted values. The data points in Figure 3.2(a) summarise the measurements in different $t\bar{t}$ decay channels which will be explained in Section 3.2.



(a) Top-antitop quark pair production

(b) Single-top quark production

Figure 3.2: Measured and predicted $t\bar{t}$ (a) and different single-top (b) quark production cross-sections for Tevatron energies in $p\bar{p}$ collisions and for the LHC in pp collisions [9].

[9] In pp collisions at the LHC at $\sqrt{s} = 8$ TeV the fraction of gluon fusion production is about $80 - 90\%$ while in $p\bar{p}$ collisions at the Tevatron at $\sqrt{s} = 1.96$ TeV the fraction of $q\bar{q}$ annihilation is about 85% [9]

[10] The cross-section σ is a measure giving the probability of a process to happen. Details can e.g. be found in Ref. [39].

3.1.2 Single-top quark production

The single-top quark production via the weak interaction can be separated into three different production modes. The dominant process $qb \to q't$ is mediated by a virtual $W^{\pm}$-boson in the t-channel as shown in Figure 3.3(a). The two sub-leading processes are the associated Wt production $bg \to Wt$ displayed in Figure 3.3 ((b), (c)) and the s-channel production $q\bar{q}' \to t\bar{b}$ shown in Figure 3.3(d).

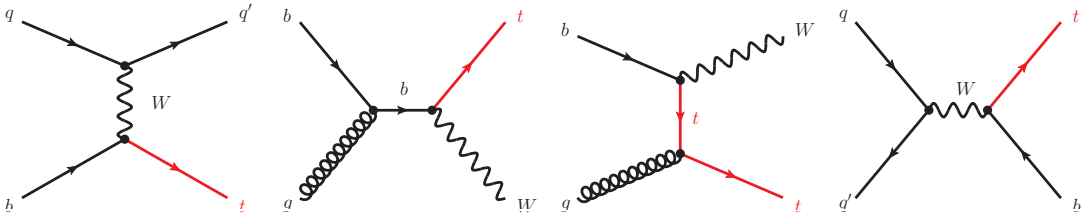


Figure 3.3: Feynman diagrams of single-top quark production processes: (a) t-channel production, (b) and (c) associated Wt production, (d) s-channel production.

In contrast to the symmetric production at the Tevatron the t-channel and s-channel production cross-sections are different for the t and $\bar{t}$ production at the LHC[11]. The cross-section of the combination of the two is predicted to be $\sigma_t = (87.8^{+3.4}_{-1.9})\,\text{pb}$ [4] and $\sigma_s = (5.6 \pm 0.2)\,\text{pb}$ [6]. For the Wt-channel the predicted cross-section is $\sigma_{Wt} = (22.4 \pm 1.5)\,\text{pb}$ [5] with an equal portion of t and $\bar{t}$. All the predicted cross-sections assume a top quark mass of $m_{\text{top}} = 172.5$ GeV.

A first observation of the single-top quark production at all was made in 2009 in the combined t-channel and s-channel production at the Tevatron [67]. The cross-section of the Wt-channel production was too small for a direct observation at the Tevatron but first evidence was reported by ATLAS and CMS at the LHC [68, 69]. The cross-section could be measured in addition to the separate t-channel production at different centre-of-mass energies.

First direct evidence of the s-channel corresponding to 3.7 standard deviations was found by DØ [70] and a first direct observation of the s-channel is expected with more LHC data that will be collected by the experiments in the future. Within the same analysis from DØ the first observation of the t-channel only was reported with a significance of 7.7 standard deviations [70]. The results of all measurements performed so far at the

[11] The fraction of top quarks is about 65% in the t-channel and 69% in the s-channel [9].

Tevatron and the LHC are summarised and compared to the theoretical prediction in Figure 3.2(b).

3.1.3 Measurement of the fiducial t-channel cross-section

In March 2014 the t-channel cross-section was measured for the first time in a so called fiducial volume within the detector acceptance [71]. This measurement relies on the definition of a reconstructed phase-space which in this case corresponds exactly to the signal region defined in Section 8.1.2. The fiducial volume is defined on particle-level in simulation and chosen such that it is as close as possible to the reconstructed phase space. The value of the fiducial cross-section, σ_{fid}, is defined as:

$$\sigma_{\text{fid}} = \frac{\varepsilon_{\text{corr,sel}}}{\varepsilon_{\text{corr,fid}}} \cdot \frac{\hat{\nu}}{\mathcal{L}}. \tag{3.1}$$

Here, $\varepsilon_{\text{corr,sel}}$ is the fraction of events that are selected by the offline selection at usual reconstruction level to be within the fiducial volume; $\varepsilon_{\text{corr,fid}}$ corresponds to the fraction of events selected within the fiducial volume to be selected by the offline selection:

$$\varepsilon_{\text{corr,sel}} = \frac{N_{\text{sel+fid}}}{N_{\text{sel}}} \text{ and } \varepsilon_{\text{corr,fid}} = \frac{N_{\text{sel+fid}}}{N_{\text{fid}}}. \tag{3.2}$$

$\hat{\nu}$ is the number of expected single-top t-channel events that is being obtained from data with the integrated luminosity $\mathcal{L}$.

To estimate $\hat{\nu}$ a binned maximum likelihood fit to the data distribution of a neural network discriminant similar to the one defined in Chapter 9 is performed. In summary one obtains a value for σ_{fid} that is independent from the choice of the MC generator [71]:

$$\sigma_{\text{fid}} = (3.37 \pm 0.05(\text{stat.}) \pm 0.47(\text{syst.}) \pm 0.09(\text{lumi.}))\,\text{pb}. \tag{3.3}$$

With the value of σ_{fid} one can now extrapolate to the full inclusive phase-space by applying the selection efficiency ε_{fid} of a certain generator with the simple equation:

$$\sigma_t = \frac{1}{\varepsilon_{\text{fid}}} \cdot \sigma_{\text{fid}} \tag{3.4}$$

The estimated inclusive cross-section, σ_t, is shown in Figure 3.4 for different MC generators taking into account the branching ratio of leptonic top quark decays. These decay modes will be explained in Section 3.2. It is compared to the expected theoretical cross-section, see Section 3.1.2, and a good agreement is found.

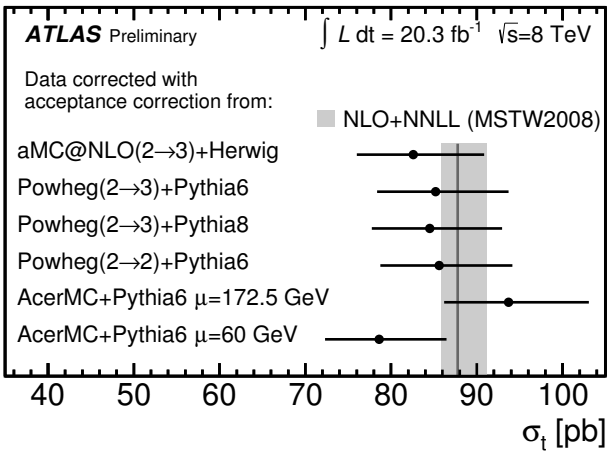


Figure 3.4: Inclusive total single top-quark t-channel cross-section obtained by extrapolation from the fiducial cross-section to the full space for different MC generators. The vertical line shows the expected theoretical cross-section calculated at NLO+NNLL [71].

The advantage of measuring the cross-section via the fiducial volume is that systematic uncertainties due to the event generation affecting the signal acceptance only appear as differences within the fiducial volume. This reduces the total systematic uncertainty of σ_t from about 17% to 14%.

3.2 Decay of the top quark

As explained in Section 2.3.3 the probability of the decay of the top quark is proportional to the square of the respective matrix elements of the CKM quark mixing matrix. Since $|V_{tb}| \gg |V_{td}|$, $|V_{ts}|$ the by far dominating decay of the top quark is into a bottom quark and a $W^{\pm}$-boson, $t \to W^{\pm}b$. The predicted total decay width is [9][12]:

$$\Gamma_{\text{top}} = \frac{G_F m_{\text{top}}^3}{8\pi\sqrt{2}} \left(1 - \frac{m_W^2}{m_{\text{top}}^2}\right)^2 \left(1 + 2\frac{m_W^2}{m_{\text{top}}^2}\right) \left[1 - \frac{2\alpha_s}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right]. \tag{3.5}$$

With G_F being the Fermi coupling constant, the top quark mass $m_{\text{top}} = 173.3$ GeV, the mass of the $W^{\pm}$-boson $m_W = 80.4$ GeV and the strength of the strong coupling at the scale of the mass of the Z^0-boson $\alpha_s(M_Z) = 0.118$ the value of the total decay width is $\Gamma_{\text{top}} = 1.35$ GeV. Thus, the top quark has a very short lifetime of about $\tau_{\text{top}} = \frac{\hbar}{\Gamma_{\text{top}}} \approx 0.5 \cdot 10^{-24}$ s. It almost exclusively decays before hadrons including top quarks or

[12]Terms of order $\frac{m_b^2}{m_{\text{top}}^2}$, α_s^2 and $\left(\frac{\alpha_s}{\pi}\right)\frac{m_W^2}{m_{\text{top}}^2}$ have been neglected.

$t\bar{t}$-bound states can be formed. This means that the top quark is the only quark that can be handled as a quasi-free quark state.

3.2.1 Decay channels in $t\bar{t}$ production

As explained, in the $t\bar{t}$ production a top-antitop quark pair is produced. Both quarks are decaying almost instantaneously due to their very short lifetime via $t \to W^{\pm}b$. Therefore, two b-quarks and two $W^{\pm}$-bosons are produced. While the b-quarks build hadrons and then manifest themselves as b-jets, the two $W^{\pm}$-bosons can decay either hadronically in a quark-antiquark pair or leptonically in a charged lepton and its corresponding neutrino. The final state of a $t\bar{t}$ decay can, therefore, be ordered in three different classes depending on the decay of the two $W^{\pm}$-bosons:

- **fully-hadronic (all-jets):** $t\bar{t} \to W^{+}b\,W^{-}\bar{b} \to q\bar{q}'\,b\,q''\bar{q}'''\,\bar{b}$
- **semi-leptonic (lepton+jets):** $t\bar{t} \to W^{+}b\,W^{-}\bar{b} \to q\bar{q}'\,b\,\ell^{-}\bar{\nu}_{\ell}\,\bar{b}$
  $t\bar{t} \to W^{+}b\,W^{-}\bar{b} \to \ell^{+}\nu_{\ell}\,b\,q''\bar{q}'''\,\bar{b}$
- **dileptonic (lepton+lepton):** $t\bar{t} \to W^{+}b\,W^{-}\bar{b} \to \bar{\ell}\nu_{l}\,b\,\ell'\bar{\nu}_{\ell'}\,\bar{b}$

The different quarks q from the decay can be reconstructed as jets and ℓ can be either an electron, muon or τ-lepton that are reconstructed with the respective sub-detectors.

Since quarks exist in three different colour states, see Section 2.2.1, the fully-hadronic decay channel is the dominant channel. It is followed by the semi-leptonic channel and the dilepton channel with the smallest branching fraction. The relative fraction, including hadronic corrections and lepton universality, are given in Table 3.1 together with the objects that have to be reconstructed.

τ-leptons only have a very short lifetime of $\tau_{\tau} = (2.906 \pm 0.010) \cdot 10^{-13}\,\mathrm{s}$ [9] and can decay either into electrons, muons or hadronically into quarks seen as jets. That is why they are difficult to reconstruct and only the electron and the muon channel is looked at in most experimental analyses.

The most precise measurements of top quark properties have been made in the semi-leptonic electron or muon channel. The one created charged lepton can be well triggered and reconstructed. The presence of only one neutrino makes it possible to fully reconstruct its 4-vector by using the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ and a constraint on the mass of the $W^{\pm}$-boson. The only ambiguity remaining is the association of the four reconstructed jets to the two b-jets directly coming from the top quark decay and to the two jets from the decay of the $W^{\pm}$-boson.

The fully-hadronic channel is suffering even more from the jet combinatorics and large non-reducible background contributions due to QCD-multijet production. The dilepton channel has less statistics than the other two due to the low branching ratio and the presence of two neutrinos makes it more difficult to fully reconstruct the final state.

Table 3.1: Summary of the different $t\bar{t}$ decay channels [9].

	number of jets (b-jets)	number of leptons (e,μ,τ)	number of neutrinos	relative fraction
fully-hadronic	6(2)	0	0	45.7%
semi-leptonic (e, μ)	4(2)	1	1	29.2%
semi-leptonic (τ)	4 or 5(2)	1	1	14.6%
dileptonic (e, μ)	2(2)	2	2	7.0%
dileptonic (τ)	2, 3 or 4(2)	2	2	3.5%

3.2.2 Decay channels in single-top quark production

In the single-top quark production no further classification of the decay channels is applied since there is only one top quark decaying via $t \to W^{\pm}b$.

In this thesis t-channel topologies where the $W^{\pm}$-boson decays leptonically are being studied. Therefore, the final state is characterised by the b-jet from the top quark decay, one charged lepton and missing transverse energy from the $W^{\pm}$-boson decay and an additional jet[13] from the t-channel production itself, see Figure 3.3(a). Also here final states with τ-leptons are not explicitly reconstructed, but some acceptance is present when the τ-lepton decays leptonically and either the electron or the muon is reconstructed.

The final state signature of the Wt-channel characterises by either more leptons or a higher jet multiplicity. This is caused by the presence of an additional $W^{\pm}$-boson in the final state as can be seen in Figures 3.3 ((b), (c)).

The final state topology of the single-top s-channel is similar compared to the t-channel but with a b-quark induced jet as the additional jet, see Figure 3.3(d).

[13]It is typical for the t-channel production that this additional jet is in the forward region of the detector.

3.3 The mass of the top quark

The property of the top quark that has been studied with the highest emphasis is the mass of the top quark. Over the last two decades physicists worked on estimating the value of m_{top} starting with indirect measurements with fits to other electroweak parameters. In 1995 the first direct observation including a first measurement of m_{top} at the Tevatron was done. In Figure 3.5 the time series of the determination of the top quark is shown ending with the first world combination done in March 2014 [3].

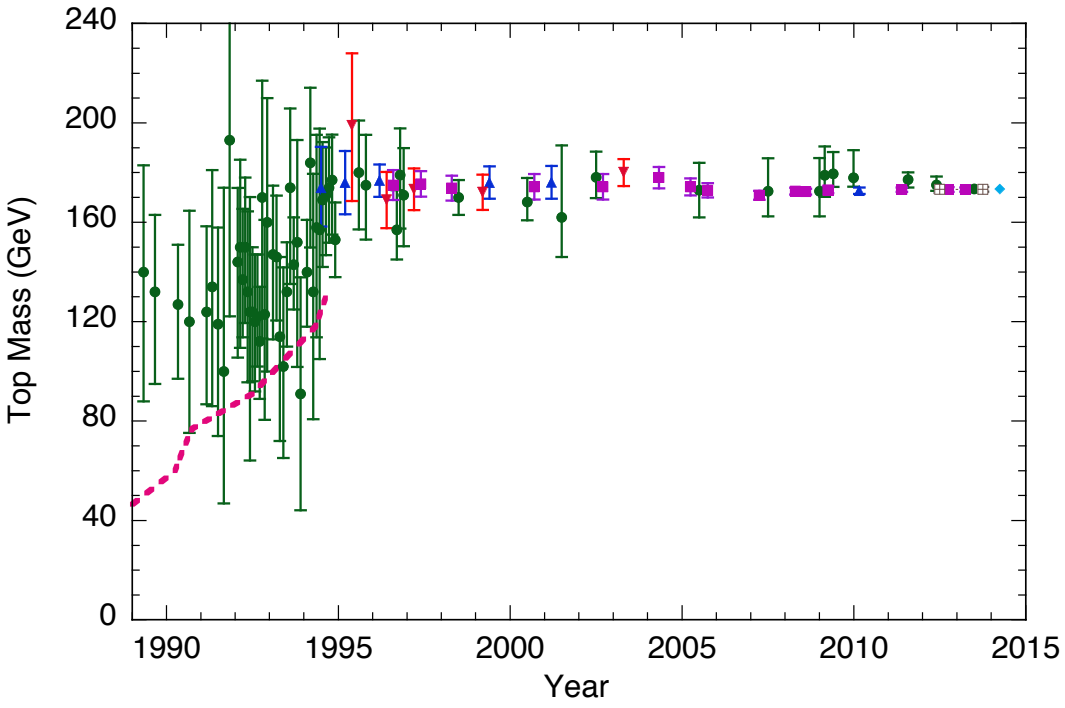


Figure 3.5: Time series of the determination of the mass of the top quark. First indirect determinations from fits to electroweak observables are shown as green circles ● and 95% confidence-level lower bounds of first direct searches as a broken line. Measurements made at the Tevatron by CDF are shown as blue triangles ▲ and by DØ as inverted red triangles ▼ both at the time of initial evidence of the top quark, discovery claim and thereafter. The Tevatron average is shown as magenta squares ■ and the first LHC averages as a crossed box ⊞. The newest World Average from March 2014 is shown as a cyan diamond ♦ [72].

The world combination includes the most precise measurements in different $t\bar{t}$ decay channels that have been done by the four different experiments. They are summarised in Figure 3.6. All these measurements have been done with top quarks that have been

produced in $t\bar{t}$ pairs as described in Section 3.1. In this thesis a first measurement of m_{top} in a dataset that is dominated by top quarks from single-top t-channel production is presented. This decay channel has very good prospects to be included in a future world combination and significantly increase the overall precision of the top quark mass estimation.

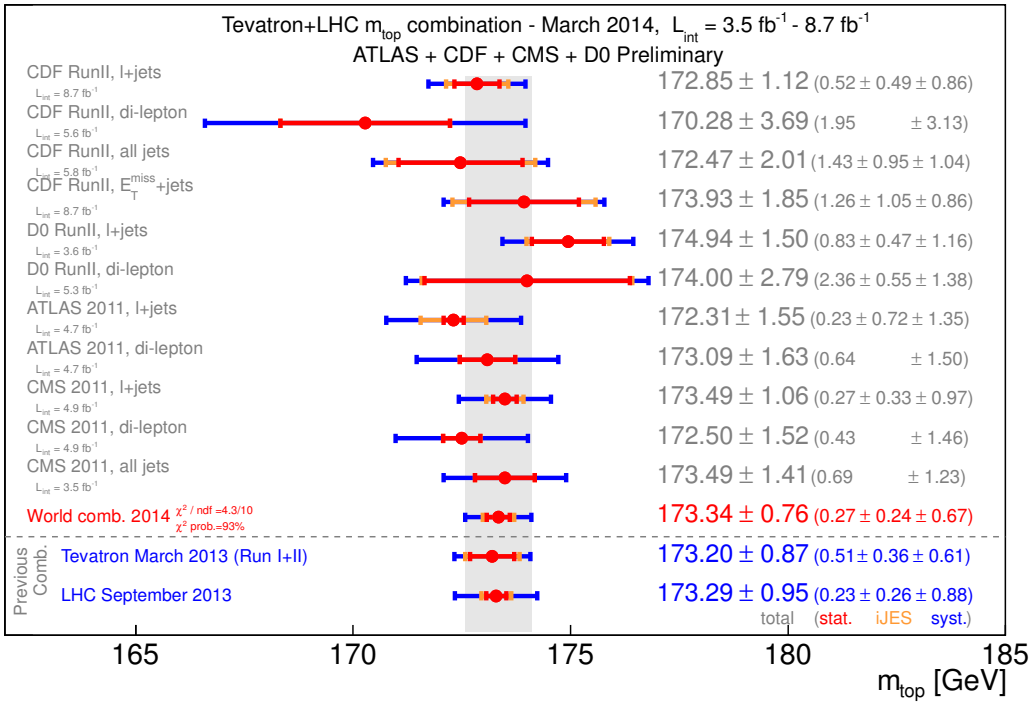


Figure 3.6: Most precise measurements performed by different experiments in the different decay channels used for the March 2014 world combination [3].

3.3.1 Theoretical aspects about the top quark mass definition

The definition of the top quark mass in quantum field theory is a complicated task and still under theoretical investigation. It is impossible to mention all aspects related to this topic but at least an overview shall be given.

A possible definition of the top quark mass within a perturbative quantum field theory is the pole mass, $m_{\text{top}}^{\text{pole}}$. This mass definition corresponds to the appearance of a pole in the propagator representing a two-point function. The strong interaction between quarks, introduced in Section 2.3, is theoretically described by the quantum field theory

of QCD which is not a fully perturbative theory. Thus, two kind of caveats arise with the pole mass definition:

- At high energies divergent terms occur that have to be renormalised with a running mass $m(\mu)$. It depends on the renormalisation scale μ that is controlling the absorption of ultraviolet fluctuations into the mass.

- In the quark-propagator of a non-perturbative theory like the QCD there is no pole appearing. This is caused by the confinement of quarks, mentioned in Section 2.3, and so the pole mass is in general ill-defined.
  A special feature of the top quark is its large width and very short lifetime as described in Section 3.2. Since the top quark decays before it hadronises it can be handled as a free nearly deconfined quark. It was hoped that the large width would prevent the top quark from effects due to non-perturbative QCD leading to a well-defined top quark pole mass. However, it was shown that $m_{\text{top}}^{\text{pole}}$ remains ambiguous by an amount proportional to the strong interaction scale $\Lambda_{\text{QCD}} \approx 200$ MeV [73][14].

The mass of the top quark can be defined in other, so called short-distance, renormalisation schemes. In these schemes no ambiguity is present and the measurement can in principle be done with arbitrary accuracy. These schemes always depend on a scale R controlling the absorption of infrared fluctuations. A common representation of these short-distance schemes is the $\overline{MS}$ scheme with the intrinsic definition of $R = \overline{m}(\mu)$. More details about the definition of different renormalisation schemes can be found in Ref. [74].

In summary, the top quark mass in a quantum field theory is described by the long-distance behaviour, given by the pole mass scheme, and the short-distance behaviour given by e.g. the $\overline{MS}$ scheme. The Differences between the two schemes can be evaluated as a perturbative series in α_s. This was done at the three-loop level yielding [73, 75, 76]:

$$\begin{aligned} m_{\text{top}}^{\text{pole}} = m_{\text{top}}^{\overline{\text{MS}}}(m_{\text{top}}^{\overline{\text{MS}}}) \Bigg[1 &+ \frac{4}{3}\frac{\overline{\alpha}_s(m_{\text{top}}^{\overline{\text{MS}}})}{\pi} \\ &+ (-1.0414 N_L + 13.4434)\left(\frac{\overline{\alpha}_s(m_{\text{top}}^{\overline{\text{MS}}})}{\pi}\right)^2 \\ &+ (0.6527 N_L^2 + 26.655 N_L + 190.595)\left(\frac{\overline{\alpha}_s(m_{\text{top}}^{\overline{\text{MS}}})}{\pi}\right)^3 \Bigg], \end{aligned} \tag{3.6}$$

[14] This is known as the pole mass $O(\Lambda_{\text{QCD}})$ renormalon problem.

with $N_L = 5$ being the number of other quark flavours and $\overline{\alpha}_s(m_{\text{top}}^{\overline{\text{MS}}})$ being the strong coupling in the $\overline{MS}$ scheme taken from Ref. [77][15].

MC simulations, with a top quark mass referred to as $m_{\text{top}}^{\text{MC}}$, are currently based on the calculation of the matrix element at leading (LO) or next-to-leading order (NLO) precision. All higher order effects are then handled by the parton shower at modified leading logarithm (LL) level. This means that still there is no coherent analytic framework relating perturbative and non-perturbative effects in a consistent way. It is impossible to directly relate $m_{\text{top}}^{\text{MC}}$ with the pole or any other mass scheme without the parton shower calculated with at least next-to-leading logarithms (NLL) accuracy [78].

Currently the definition of $m_{\text{top}}^{\text{MC}}$ is given by the implementation of the parton shower of the top quark and the chosen shower cutoff scale [79]. Details about this definition for the different event generators can be found in Ref [80]. It is predicted that $m_{\text{top}}^{\text{MC}}$ is close to $m_{\text{top}}^{\text{pole}}$ with a possible difference being of the order of 1 GeV [78, 79].

3.3.2 Techniques to measure the mass of the top quark

Different techniques have been applied to measure m_{top} and some of them are briefly presented in this subsection. One has to be aware of the fact that the mass estimated in all direct measurements is the mass used in the calculation of the MC generators in simulation for the reasons given in Section 3.3.1. This value of $m_{\text{top}} \stackrel{\triangle}{=} m_{\text{top}}^{\text{MC}}$ is not equivalent to the mass that appears in the theoretical calculations of quantum field theories for the reasons given in Section 3.3.1.

Indirect measurements Indirect estimations of m_{top} using theoretical calculations have been performed in many different ways. As the variety of these calculations can not be summarised shortly only one example is given here.

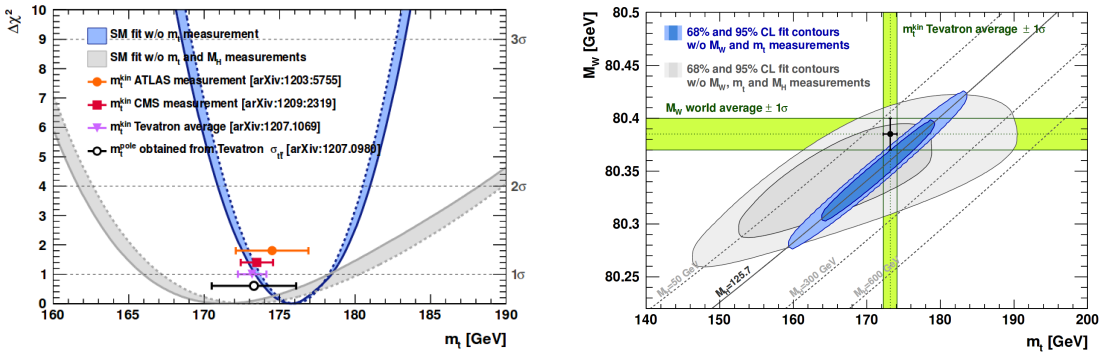
Effects of new physics can occur by vacuum polarisation and lead to differences in electroweak precision observables. To constrain these models of new physics three oblique gauge self-energy parameters S, T and U have been introduced in 1992 by Peskin and Takeuchi [81].

The T parameter is measuring a possible occurrence of isospin violation[16]. As the bottom quark is the isospin partner of the top quark the large mass splitting between the two is an evidence of a broken isospin symmetry. The mass of the bottom quark is

[15] $\overline{\alpha}_s(m_{\text{top}}^{\text{pole}})$ is iteratively transformed to $\overline{\alpha}_s(m_{\text{top}}^{\overline{\text{MS}}})$ with Equation (3.6).

[16] It is sensitive to differences between the loop corrections to the vacuum polarisation function to the Z^0-boson and the $W^{\pm}$-boson. This is equivalent to the differences of contributions from new physics to neutral and charged currents at low energies.

known with high precision and with the T parameter the mass of the top quark can be constrained. In Figure 3.7(a) the result of a fit to electroweak precision data is shown estimating the value of m_{top}. Also included are the directly measured values that are in good agreement with the indirect estimate [82].



(a) Indirect measurement of m_{top} using a fit to electroweak precision data.

(b) Fit result relating the three important electroweak parameters m_{top}, m_W and m_H

Figure 3.7: (a) $\Delta\chi^2$ profiles as a function of m_{top} of the fit to electroweak precision observables. The blue band includes the measured mass of the Higgs boson m_H while the grey band shows the fit with m_H being excluded. (b) Confidence level contours of the fit to indirectly estimate m_{top}, m_W and m_H compared to direct measurement [82].

Figure 3.7(b) shows confidence level contours of the fits to electroweak precision data relating the three masses of the $W^{\pm}$-boson, the Higgs-Boson and the top quark. The grey areas show the fit results where m_{top}, m_W and m_H have been used as free parameters. The fit result represented by the blue contour area was made with only m_{top} and m_W as free parameters. For both fits very good agreement with the direct measurement shown as green bands is found [82].

With more precise measurement of the parameters of the Standard Model, like m_{top}, it is possible to gain a better knowledge about other yet unmeasured parameters. The relationship of m_{top}, m_W and m_H shown in Figure 3.7 is a good example of this principle.

The matrix element method An approach to measure m_{top} directly is the matrix element method only done so far at the Tevatron. In this method a probability P_{event} is calculated for each event depending on m_{top}. The probability includes an in-situ mea-

surement of the jet energy scale represented by k_{JES}. The event-based probability is calculated depending on the fraction of $t\bar{t}$ signal events f_{sig}, m_{top} and k_{JES}:

$$P_{\text{event}}(m_{\text{top}}, k_{\text{JES}}) = A(x)\left[f_{\text{sig}} \cdot P_{\text{sig}}(x; m_{\text{top}}, k_{\text{JES}}) + (1 - f_{\text{sig}}) \cdot P_{\text{bkg}}(x; k_{\text{JES}})\right]. \quad (3.7)$$

x represents all measured quantities of the event like jet or lepton energies and angles. $A(x)$ accounts for acceptance and efficiencies of the detector used to reconstruct the quantities x. As the signal here are $t\bar{t}$ events the signal probability density function $P_{\text{sig}}(x; m_{\text{top}}, k_{\text{JES}})$ is given by the probability density of the $t\bar{t}$ pair production $P_{t\bar{t}}(x; m_{\text{top}}, k_{\text{JES}})$.

The calculation of $P_{t\bar{t}}(x; m_{\text{top}}, k_{\text{JES}})$ is the heart of the method. The two initial partons carry momenta denoted as q_1 and q_2 meaning that one has to sum over all possible initial-state parton flavours and integrate over all possible momenta q_i. To do so one uses the parton density functions $f(q_i)$. These PDFs have already been explained in Section 2.6 and are an estimate of the probability to find a parton with a certain flavour and momentum fraction q_i. They are convoluted with the partonic differential cross-section of the $t\bar{t}$ production $\mathrm{d}\sigma(y; m_{\text{top}})$ and an additional transfer function $W(x, y; k_{\text{JES}})$. This is necessary to include the detector resolution and relate the measured quantities x with the partonic quantities y of the original final-state. In total $P_{t\bar{t}}(x; m_{\text{top}}, k_{\text{JES}})$ is given by:

$$P_{t\bar{t}}(x; m_{\text{top}}, k_{\text{JES}}) = \frac{1}{\sigma_{t\bar{t}}(m_{\text{top}})} \int \sum_{\text{flavours}} \mathrm{d}\sigma(y; m_{\text{top}}) \mathrm{d}q_1 \mathrm{d}q_2 f(q_1) f(q_2) W(x, y; k_{\text{JES}}). \quad (3.8)$$

The event-based weights are then combined in a likelihood to estimate the value of f_{sig} and a 2-dimensional likelihood to measure m_{top} and k_{JES}. The definitions of the likelihood, the background probability density $P_{\text{bkg}}(x; k_{\text{JES}})$ and the transfer functions are explained in much more detail in Ref. [83, 84].

The template method The template method is a complementary method used to measure m_{top}. It is based on simulated templates of an observable sensitive to m_{top} which is then fitted to the data distribution of that observable. This can be extended to more than one observable and more than one physics parameter to reduce the influence of systematic uncertainties.

An example is the most precise measurement of m_{top} in the lepton+jets channel done by ATLAS using a 3-dimensional template likelihood fit [85]. In that analysis three

observables are calculated for each selected event being the reconstructed top quark mass obtained in a likelihood fit, $m_{\mathrm{top}}^{\mathrm{reco}}$, the invariant mass of the hadronically decaying $W^{\pm}$-boson, $m_{\mathrm{W}}^{\mathrm{reco}}$, and the $R_{\mathrm{lb}}^{\mathrm{reco}}$ variable[17]. This variable is defined according to the number of tagged b-jets in the event following Equation (3.9):

$$R_{\mathrm{lb}}^{\mathrm{reco,2b}} = \frac{p_T^{b_{\mathrm{had}}} + p_T^{b_{\mathrm{lep}}}}{p_T^{W_{\mathrm{jet}_1}} + p_T^{W_{\mathrm{jet}_2}}} \quad \text{or} \quad R_{\mathrm{lb}}^{\mathrm{reco,1b}} = \frac{p_T^{b_{\mathrm{tag}}}}{\left(p_T^{W_{\mathrm{jet}_1}} + p_T^{W_{\mathrm{jet}_2}}\right)/2}. \tag{3.9}$$

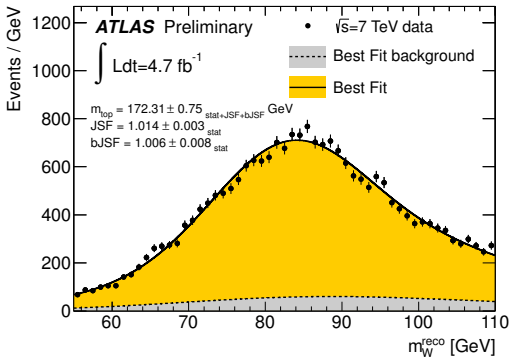
The fit of the 3-dimensional templates to the data is shown in Figure 3.8 for all three variables. It is used to obtain the value of m_{top}, a global jet energy scale factor (JSF) and a relative b-jet to light-flavour-jet scale factor (bJSF). In this way the dependence on the jet energy scale, which is one of the dominant uncertainties in most measurements of m_{top}, is significantly reduced yielding a very high precision measurement.

Fitting p_T^{lep} and L_{xy} As the jet energy scale is the dominant systematic uncertainty in most measurements of the top quark mass different methods have been developed that rely almost exclusively on tracking instead of calorimetry. These complementary approaches rely on the fact that the information about m_{top} is passed to the decay products of the top quark via a kinematic boost. This can be exploited by using either the transverse decay length of the b-hadron, L_{xy}, or the p_T of the lepton from the decay of the $W^{\pm}$-boson, p_T^{lep}. The spectrum of both variables is expected to be harder with larger values of m_{top}.

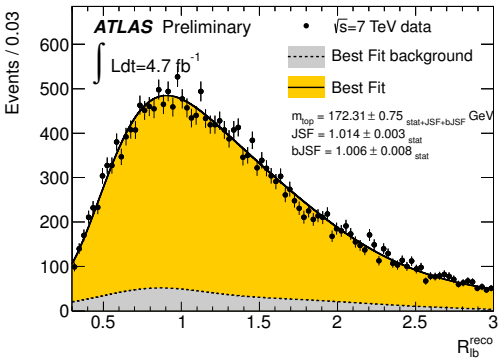
Both variables have been used individually already at the Tevatron [86, 87] and another analysis combining both in a 2-dimensional likelihood is currently prepared by ATLAS and documented in Ref. [88]. In that analysis the mass of the top quark as well as the amount of final state gluon radiation (FSR) is measured simultaneously. This is sensible due to the fact that L_{xy} has a large sensitivity to the amount of FSR leading to a large systematic uncertainty. This can be avoided by the in-situ measurement.

In summary the total uncertainty of the measurement using these tracking variables as mass estimators is slightly larger than in the standard approaches. On the other hand the uncertainty due to the jet energy scale is very small. This leads to very good prospects for combinations with other measurements.

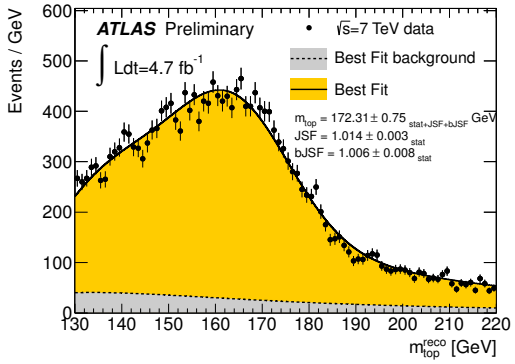
[17] With the kinematic fit the jets are assigned the their originating partons of the $t\bar{t}$ decay. This means that the measured jets are related to the quark decay products of the $W^{\pm}$-boson and to the b-quarks produced in the top quark decays. Thus, invariant masses of the top quarks and the $W^{\pm}$-bosons can be calculated by adding the reconstructed charged lepton and neutrino using the missing transverse energy.



(a) Reconstructed mass of the hadronic $W^{\pm}$-boson



(b) $R_{\mathrm{lb}}^{\mathrm{reco}}$, defined in Equation (3.9)



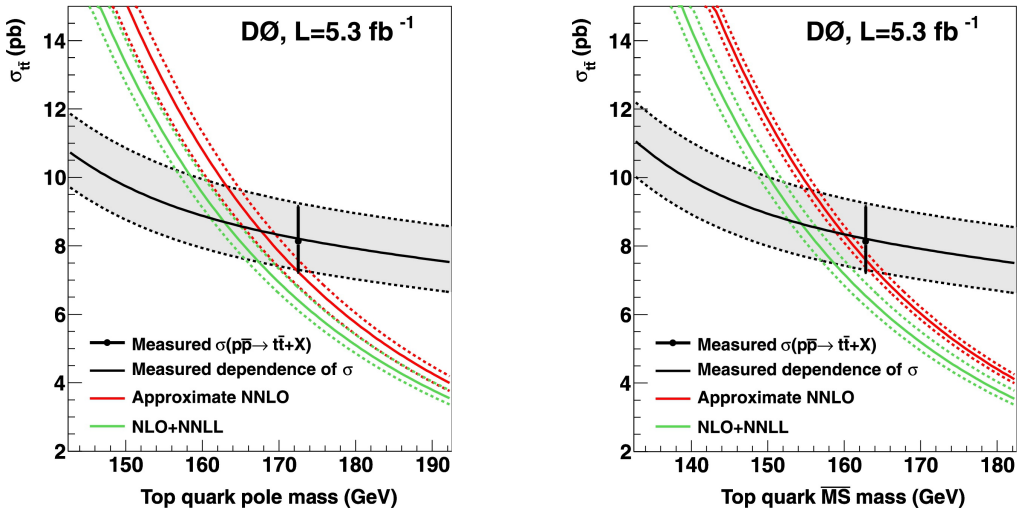
(c) Reconstructed mass of the leptonic top quark

Figure 3.8: Fitted distributions in the 3d-template method performed in 4.7fb^{-1} of ATLAS data. Shown is the mass of the $W^{\pm}$-boson from the top decaying hadronically via $W \to jj$ (a), the $R_{\mathrm{lb}}^{\mathrm{reco}}$ variable that is used for the in-situ estimation of the b-jet to light-flavour-jet scale factor (b) and the mass of the reconstructed top quark (c) [85].

Mass estimation from the $t\bar{t}$ production cross-section The mass of the top quark is directly related to the $t\bar{t}$ production cross-section and can, therefore, be estimated when the cross-section itself is measured.

The DØ experiment at the Tevatron has measured the value of $\sigma_{t\bar{t}}$ and estimated the cross-section depending on the top quark mass used in simulation $\sigma_{t\bar{t}}\left(m_{\mathrm{top}}^{\mathrm{MC}}\right)$ [78]. This parametrisation is compared to different higher-order QCD calculations. The extraction of m_{top} is done by applying a normalised joint-likelihood function including experimental and theoretical uncertainties.

The theoretical calculation of $\sigma_{t\bar{t}}$ involves an unambiguous definition of the top quark mass either in the pole mass scheme or the $\overline{MS}$ scheme. Consequently, this means that



(a) $\sigma_{t\bar{t}}$ as a function of $m_{\text{top}}^{\text{pole}}$

(b) $\sigma_{t\bar{t}}$ as a function of $m_{\text{top}}^{\overline{\text{MS}}}$

Figure 3.9: Comparison of the measured cross-section $\sigma_{t\bar{t}}(m_{\text{top}}^{\text{pole}} = m_{\text{top}}^{\text{MC}})$ with theoretical calculations of $\sigma_{t\bar{t}}$ as a function of $m_{\text{top}}^{\text{pole}}$ (a) and as function of $m_{\text{top}}^{\overline{\text{MS}}}$ (b). The coloured dashed lines represent the theoretical uncertainties on the choice of the PDF and the renormalisation and factorisation scales. The black point shows the measured value of $\sigma_{t\bar{t}}$ for $m_{\text{top}}^{\text{MC}} = 172.5$ GeV and the grey band corresponds to the total experimental uncertainty [78].

by comparing $\sigma_{t\bar{t}}\left(m_{\text{top}}^{\text{MC}}\right)$ to either the calculation in the pole mass scheme or the $\overline{MS}$ scheme provides a measurement of $m_{\text{top}}^{\text{pole}}$ or $m_{\text{top}}^{\overline{\text{MS}}}$, respectively.

In Figure 3.9 the parametrisation is shown as a black line with grey uncertainty band under the assumption of $m_{\text{top}}^{\text{MC}} = m_{\text{top}}^{\text{pole}}$. It is compared with the theoretical calculations as a function of $m_{\text{top}}^{\text{pole}}$ (a) or $m_{\text{top}}^{\overline{\text{MS}}}$ (b). The obtained values of the top quark mass based on the theoretical calculation at approximate NNLO are [78]:

$$m_{\text{top}}^{\text{pole}} = 167.5^{+5.4}_{-4.9} \text{ GeV} \quad \text{and} \quad m_{\text{top}}^{\overline{\text{MS}}} = 160.0^{+5.1}_{-4.5} \text{ GeV}.$$

The uncertainties involve theoretical uncertainties from the choice of the PDF and the renormalisation and factorisation scales. Also included is a large systematic uncertainty related to the assumption about the top quark mass in simulation of $m_{\text{top}}^{\text{MC}} = m_{\text{top}}^{\text{pole}}$. This was tested by building a new likelihood where m_{top} is transformed using Equation (3.6) and taking half of the observed difference as a symmetric systematic uncertainty.

Although the uncertainties are still large, a first measurement of the top quark mass was provided in a well-defined mass scheme. It shows that the obtained value of $m_{\text{top}}^{\text{pole}}$

is closer to the currently most precise direct measurement of $m_{\text{top}} = 173.07 \pm 0.52 \pm 0.72$ GeV [9] based on $m_{\text{top}}^{\text{MC}}$. It agrees within one standard deviation while the obtained value of $m_{\text{top}}^{\overline{\text{MS}}}$ agrees within two standard deviations.

3.4 Other properties of the top quark

The top quark has many other interesting properties that are not directly studied in this thesis but shall be mentioned briefly.

3.4.1 Electrical charge

The electric charge of quarks is accessible by measuring the ratio $R = \frac{\sigma(e^+e^- \to \text{Hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)}$ in e^+e^--collisions. However, only if a quark-antiquark pair can be produced directly its respective flavour contributes to the ratio R. Until now no e^+e^--collider with a centre-of-mass energy large enough to produce $t\bar{t}$ pairs has been built. Thus, the top quark is the only quark whose electric charge has not been measured yet with high precision.

Exotic scenarios with the charge of the top quark being $q_t = -\frac{4}{3}$ were still compatible with electroweak measurements but in 2013 ATLAS performed a first measurement of the charge of the top quark yielding $q_{\text{top}} = (0.64 \pm 0.02(\text{stat.}) \pm 0.08(\text{syst.}))e$. This excludes the exotic scenario by more than 8 standard deviations [89].

Many more searches for exotic charges or analyses to extract the exact electromagnetic coupling of the top quark have been performed. However, the exclusion of various exotic coupling scenarios will only be possible with more experimental data.

3.4.2 Spin correlations and $W^{\pm}$-boson helicity

As described in Section 3.2 the top quark only has a very short lifetime meaning that it can be interpreted as a free quark. This has the consequence that its spin is conserved and passed on to its decay products. Therefore, the polarisation of the top quark is measurable via angular distributions of the decay products. In the production of a $t\bar{t}$ pair, see Section 3.1.1, the top and antitop quarks are usually produced unpolarised, however, their spins are correlated. This hypothesis was tested in several analyses and the scenario of no spin correlations was excluded at 5.1 standard deviations [9].

In the Standard Model it is expected that the top quark decays into a $W^{\pm}$-boson that is longitudinally polarised with a fraction of about 70%. In the remaining 30% the $W^{\pm}$-boson is expected to be left-handed while the right-handed component is strongly suppressed. This is a direct consequence of the V-A structure of the weak coupling that

was already mentioned in Section 2.3.2. The Standard Model hypothesis was measured with different analyses all confirming the expectation. Most of the direct measurement use the cosine of the helicity angle, θ^*, between the lepton and the b-quark in the rest frame of the $W^{\pm}$-boson from the decay of the top quark [9].

CHAPTER 4

The ATLAS detector at the LHC

As explained in Section 2.2 matter we observe in everyday life is build out of quarks and leptons from the first generation. Other fermions, the Higgs boson or the gauge bosons mediating fundamental forces are unstable. They have to be produced in high energy particle collisions at particle accelerators.

In Section 4.1 the LHC accelerator complex is explained and in Section 4.2 details are given about the experimental setup of the ATLAS detector. This particle detector is used to reconstruct known or yet unobserved fundamental particles related to new physics. In Section 4.3 the Tevatron as another accelerator complex with other particles detectors is shortly introduced.

4.1 The Large Hadron Collider

The **L**arge **H**adron **C**ollider (LHC) [90] is a circular synchrotron particle accelerator located at the European Organisation for Nuclear Research, CERN[18], in Geneva at the Swiss-French border. The accelerator is built in the former tunnel of the **L**arge **E**lectron **P**ositron (LEP) collider with a circumference of approximately 27 km about 100 m underground. The LHC was built to accelerate bunches containing about $1.15 \cdot 10^{11}$ protons to a beam energy of 4 TeV. This is leading to a centre of mass energy of $\sqrt{s} =$ 8 TeV in the proton-proton collision[19].

The main goal of the whole accelerator complex is to answer a variety of questions about the human understanding of physical laws. For this purpose the collider tunnel

[18] **C**onseil **E**uropéen pour la **R**echerche **N**ucléaire

[19] In 2011 at first the LHC was operated at a centre of mass energy of $\sqrt{s} = 7$ TeV and the $\sqrt{s} = 8$ TeV data was taken during 2012. After a technical stop in the years 2013-2015 the LHC will be operating with a centre of mass energy of $\sqrt{s} = 13$ TeV and then finally reach its design centre of mass energy of $\sqrt{s} = 14$ TeV.

includes two parallel beam pipes where the protons are accelerated in opposite directions and the two beams are brought to collisions in four interaction points around the tunnel. At these interaction points four large experiments are located trying to reconstruct the particles being produced in a collision. The four large experiments that are targeting different fields of particle physics are:

- **ATLAS** (**A** **T**oroidal **L**HC **A**pparatu**s**): ATLAS is a general purpose detector searching for different kinds of new physics like the origin of mass of elementary particles, dark matter or extra dimensions. The dataset used in this thesis was recorded by the ATLAS detector [91].

- **CMS** (**C**ompact **M**uon **S**olenoid): As well as ATLAS the CMS detector is a multi-purpose detector searching for new physics [92].

- **LHCb**: As an asymmetric detector with very high granularity in the forward direction, LHCb is able to measure parameters of the Standard Model with high precision. Deviations from the Standard Model at high energy scales that cannot be measured directly are accessible indirectly by precision measurements with b-hadrons [93].

- **ALICE** (**A** **L**arge **I**on **C**ollider **E**xperiment): ALICE is studying the quark-gluon plasma that is produced in Pb-Pb collisions[20]. A quark-gluon plasma is characterised by a very high temperature and energy density where quarks and gluons are deconfined [94].

In addition to the four larger detectors there are three smaller experiments located around the tunnel which are:

- **LHCf**: The LHCf experiment is located about 140m from the interaction point of ATLAS and is measuring the energy and number of neutral pions (π^0) in very forward direction. The physics goal is to gain a better understanding of the origin of ultra-high-energy cosmic rays [95].

- **TOTEM** (**TOT**al **E**lastic and diffractive cross-section **M**easurement): TOTEM is located close to the interaction point of CMS and targets to measure the total cross-sections of elastic scattering and diffractive processes [96].

[20] Apart from the mode providing proton-proton collisions the LHC is also able to provide collisions of lead ions at an energy of 2.76 TeV per nucleon.

- **MoEDAL**: The MoEDAL experiment is sharing the cavern with the LHCb experiment and searches for not yet observed magnetic monopoles [97].

The accelerator complex with its underground tunnel and the four experiments and their access points is shown in Figure 4.1. Also shown is the SPS (**S**uper **P**roton **S**ynchrotron) accelerator that acts as a pre-accelerator and injection ring for the LHC.

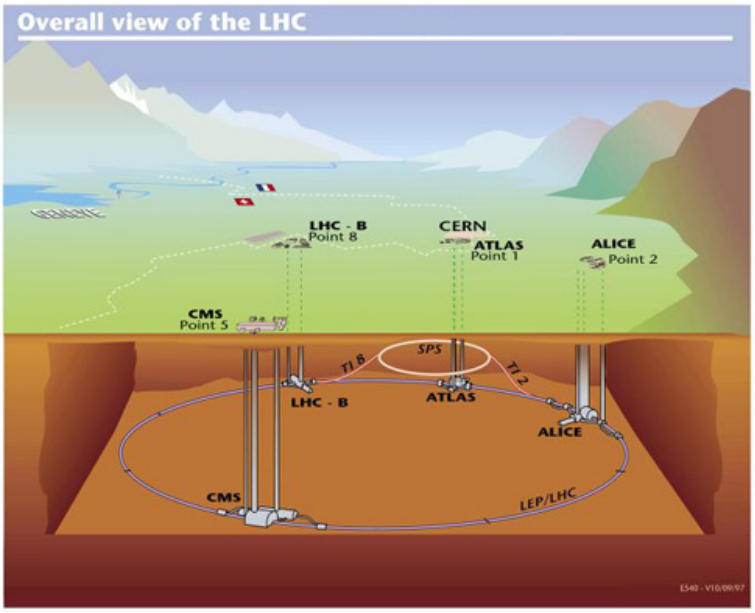
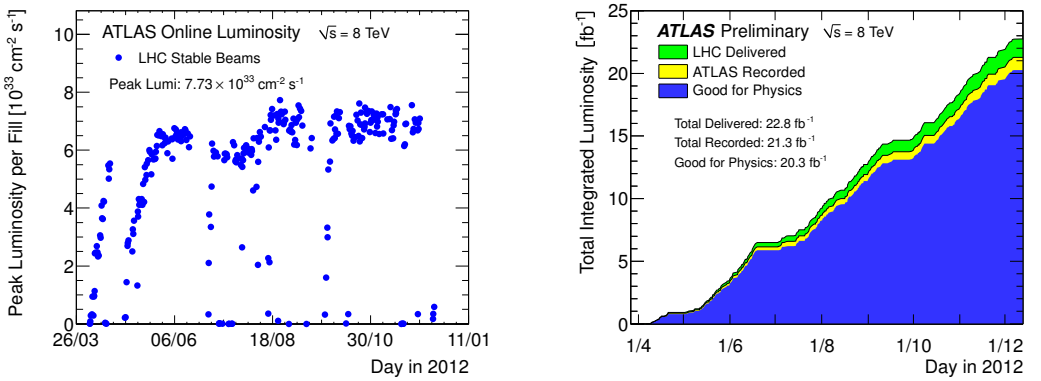


Figure 4.1: The four experiments ATLAS, CMS, LHCb and ALICE located in the four interaction points of the LHC. Not shown are the three smaller experiments LHCf, TOTEM and MoEDAL [98].

Processes at subatomic scales are explained using probabilities based on quantum field theories described in Chapter 2. This means that the experiment has to be repeated many times to access processes that are very unlikely to happen or particles that are very unlikely to be produced. This makes it very important that high event rates, corresponding to high luminosities, are provided by the LHC. The design luminosities of the LHC of $L = 10^{34}\,\frac{1}{\mathrm{cm}^2\mathrm{s}}$ was almost reached during data taking in 2012 as can be seen in Figure 4.2(a). In Figure 4.2(b) the integrated luminosity recorded by ATLAS is shown versus time. A total amount of $\mathcal{L} = \int L\mathrm{d}t = 20.3\,\mathrm{fb}^{-1}$ of data was taken and is used in this thesis[21].

[21] The shown luminosities have been measured with the same techniques explained in Ref. [99] in 7 TeV collision data taken in 2011.

The process under study in this thesis is the top quark production in the t-channel via weak interaction. Its cross-section in proton-proton collisions with a centre-of-mass energy of $\sqrt{s} = 8$ TeV is predicted to be $\sigma_t = (87.8^{+3.4}_{-1.9})$ pb assuming a top quark mass of $m_{\text{top}} = 172.5$ GeV [4]. This means that $N_{\text{t-channel}} = \mathcal{L} \times \sigma_t \approx 1.7 \cdot 10^6$ events are expected to have been produced during the data taking period in 2012[22].



(a) Maximum instantaneous luminosity versus day

(b) Cumulative luminosity versus time

Figure 4.2: (a) The maximum instantaneous luminosity versus day delivered to ATLAS during stable beam periods. (b) Cumulative luminosity versus time delivered by the LHC (green), recorded by ATLAS (yellow) and also fulfilling good quality criteria (blue) during stable beams in 2012 [100].

4.2 The ATLAS detector

Since ATLAS is a multi-purpose detector it has to fulfil several criteria to fully reconstruct the particles produced during a collision. Shown in Figure 4.3 is the experiment consisting of different sub-detectors. The detector is providing almost full angular coverage around the interaction point. In total the detector has a cylindrical form with a length of about 44 m and 25 m in diameter with a total weight of about 7000 t.

The closest sub-detector to the collision point is a three-stage tracking system. Its main purpose is the reconstruction of charged particle tracks as well as the reconstruction of primary and secondary vertices. Around the inner-detector a thin superconducting solenoid is located following an electromagnetic and hadronic calorimeter used to measure the energy of produced particles like charged leptons or hadrons. Around

[22] For simplicity this calculation does not account for the limited reconstruction efficiency, acceptance and branching ratios.

the calorimeter, the muon spectrometer is located together with three superconducting toroidal magnets making it possible to measure the momentum and charge of muons with high accuracy.

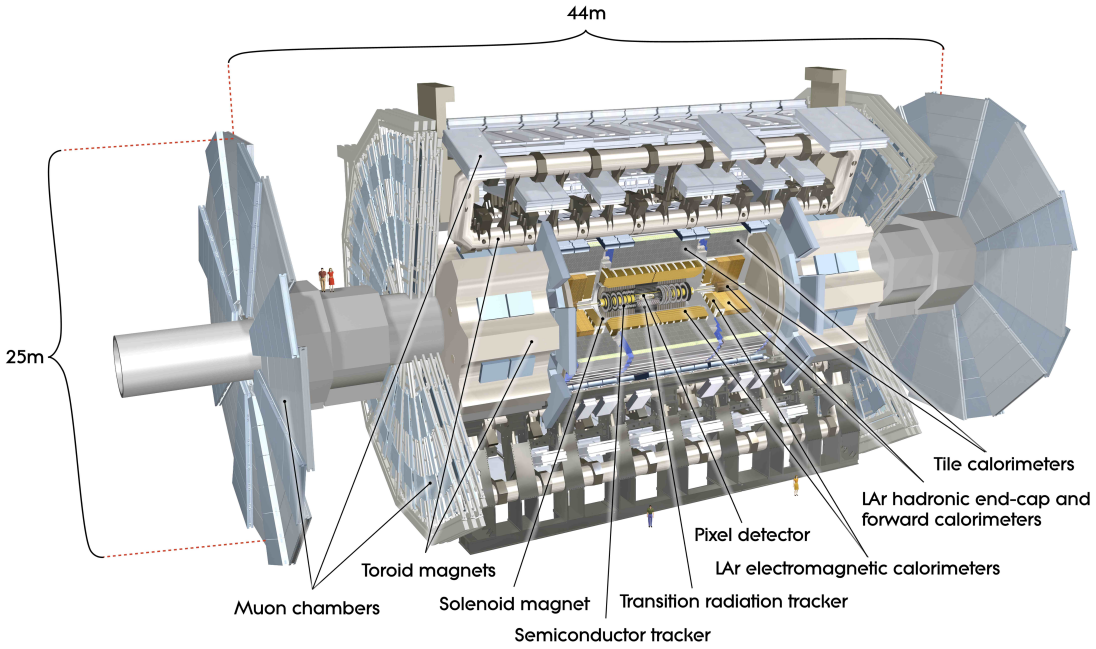


Figure 4.3: The ATLAS detector with its main components centred around the interaction point [101].

In Appendix A.3 two proton-proton collision events recorded with the ATLAS detector during 2012 are shown[23]. They have a high probability to contain a single top-quark produced in the t-channel which is the main signal process in this thesis.

4.2.1 The inner tracking detector

The **I**nner **D**etector (ID) is used to reconstruct the trajectories of charged particles as well as for vertexing. Due to the high resolution of the tracking detector it is also possible to reconstruct secondary vertices that are displaced from the primary vertex where the hard collision has happened. This is very important for instance in the context of flavour tagging where jets originating from b-hadrons can be identified. This will be explained in more detail in Section 5.1 and in Chapter 6.

Also the inner detector consists of three subsystems with increasing granularity, if one gets closer to the interaction point and the beam pipe. The innermost **pixel** detector

[23]The event displays have been created using the VP1 [102] and the Atlantis [103] software packages.

covering the vertex region is followed by a silicon micro-strip tracker, **S**emi**c**onductor **T**racker (SCT), which is then followed by the **T**ransition **R**adiation **T**racker (TRT). All three sub-detectors are divided in a barrel and end-cap region and in total an η range of $|\eta| < 2.5$ is covered by the ID shown in Figure 4.4.

The main parameters like the resolution and extension of the inner detector depend on the sub-detectors and are summarised in Table 4.1. The Pixel barrel consists of three cylindrical layers with the closest being at a radial distance of 50.5 mm from the interaction point. In addition, on each side three end-cap discs enhance the total coverage in forward direction up to $|\eta| < 2.5$.

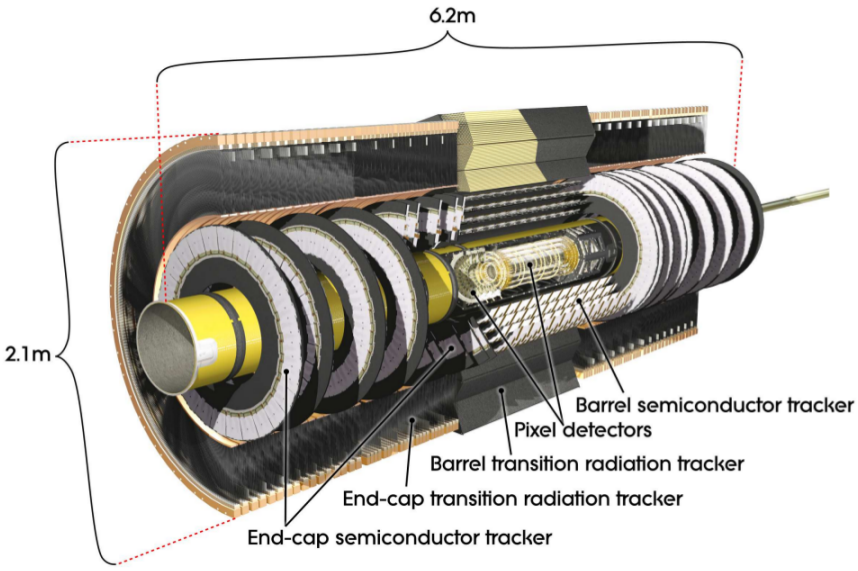


Figure 4.4: The inner tracking detector of the ATLAS experiment [91]

The SCT is built out of 4 cylindrical double-strip layers with the closest having a radial distance of 299 mm from the interaction point in the barrel. 9 additional end-cap discs on each side are covering the same η range as the pixel detector. The outermost TRT at a radial distance of $554 - 1082$ mm from the interaction point covers a range of $|\eta| < 2.0$ using straw-tubes.

Each drift (straw) tube of the TRT has a diameter of 4 mm and is filled with a xenon-based gas mixture[24]. The gas is ionised if a charged particle passes the tube. Typically each particle is detected by more than 30 tubes and therefore the overall momentum

[24] 70% Xe, 27% CO_2 and 3% O_2 with $5 - 10$ mbar over-pressure [91]

resolution of the inner detector is improved. The TRT also provides information to distinguish electrons from other heavier charged particles like pions.

In addition the whole inner-detector is embedded in a 2 T axial magnetic field bending the charged particle tracks to measure momentum and charge of the particle [91].

Table 4.1: Main parameters of the ATLAS inner detector. The actual resolution depends on $|\eta|$ and therefore only typical values are given [91].

	Radial extension [mm]	Length [mm]	Resolution σ [μm]
Overall ID	$0 < R < 1150$	$0 < \|z\| < 3512$	—
Beam-pipe	$29 < R < 36$	—	—
Pixel (envelope)	$45.5 < R < 242$	$0 < \|z\| < 3092$	—
Pixel (barrel)	$50.5 < R < 122.5$	$0 < \|z\| < 400.5$	$10(R-\varphi)\,115(z)$
Pixel (end-cap)	$88.8 < R < 149.6$	$495 < \|z\| < 650$	$10(R-\varphi)\,115(R)$
SCT (envelope)	$255 < R < 549$ (ba.)	$0 < \|z\| < 805$	—
	$251 < R < 610$ (e.-c.)	$810 < \|z\| < 2797$	—
SCT (barrel)	$299 < R < 514$	$0 < \|z\| < 749$	$17(R-\varphi)\,580(z)$
SCT (end-cap)	$275 < R < 560$	$839 < \|z\| < 2735$	$17(R-\varphi)\,580(R)$
TRT (envelope)	$554 < R < 1082$ (ba.)	$0 < \|z\| < 780$	130 (per straw)
	$617 < R < 1106$ (e.-c.)	$827 < \|z\| < 2744$	130 (per straw)
TRT (barrel)	$563 < R < 1066$	$0 < \|z\| < 712$	130 (per straw)
TRT (end-cap)	$644 < R < 1004$	$848 < \|z\| < 2710$	130 (per straw)

4.2.2 The electromagnetic and hadronic calorimeter

The calorimeter system of the ATLAS detector consists of two subsystems, the inner electromagnetic (EM) and the outer hadronic calorimeter which are both shown in Figure 4.5. The calorimeter system in total is covering a range of $|\eta| < 4.9$. The electromagnetic calorimeter is located in an η region that is matched with the inner detector and provides a fine granularity that can be used to separate electrons from photons. The hadronic calorimeter with coarser granularity is mainly used to reconstruct jets. The calorimeter has a total thickness of up to 11 radiation length[25], λ, at $\eta = 0$ which ensures a sufficient reduction of punch-through into the muon spectrometer [91].

A range up to $|\eta| < 3.2$ is covered by the three parts of the electromagnetic calorimeter. The barrel is in the region up to $|\eta| < 1.475$ and the two end-cap parts are covering $1.375 < |\eta| < 3.2$. The transition region between the barrel and the end-cap is located

[25] λ is defined as the mean path length necessary to reduce the numbers of relativistic charged particles by the factor $\frac{1}{e}$ when passing through a certain material.

between $1.37 < |\eta| < 1.52$. It is challenging to model this region in simulation and therefore it is excluded in most analyses. The EM calorimeter consists of lead absorber plates and accordion-shaped kapton electrodes using liquid Argon (LAr) as active material to measure the energy of particles showered by the absorber plates.

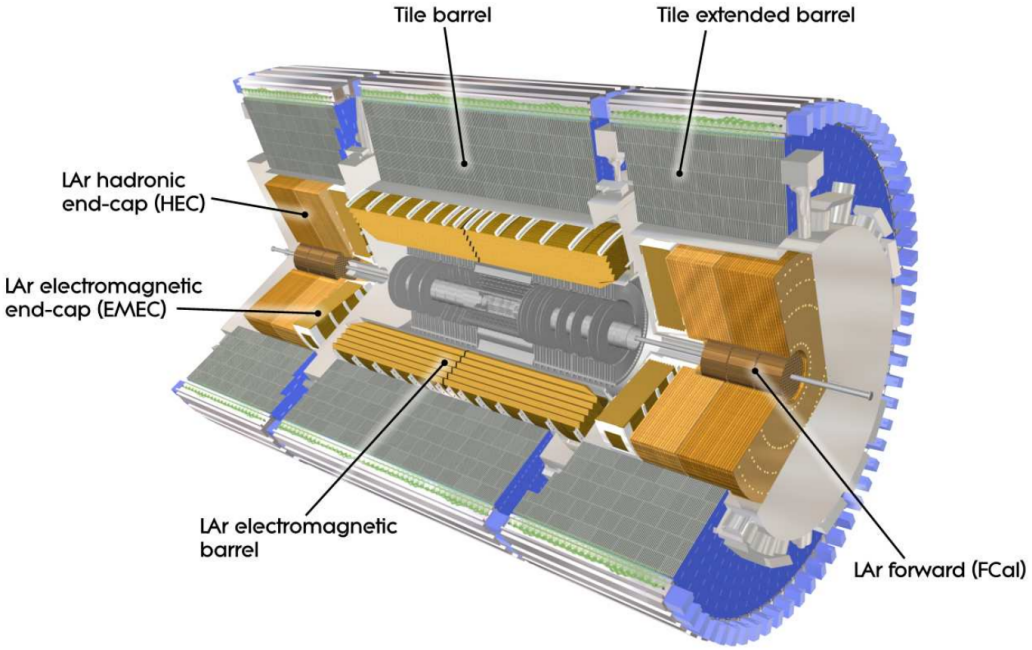


Figure 4.5: Electromagnetic and the hadronic calorimeter of the ATLAS experiment [91]

The hadronic calorimeter is also divided into barrel region and a forward end-cap region. The tile barrel is matched to the EM calorimeter barrel up to $|\eta| < 1$ and is expended by the tile extended barrel in $0.8 < |\eta| < 1.7$. Both barrel parts use steel as an absorber and scintillating tiles as active material. The total thickness of the tile-instrumented calorimeter is only 9.7λ at $\eta = 0$. In the end-cap region between $1.5 < |\eta| < 3.2$ another part of the hadronic calorimeter is located using LAr as active material and copper as absorber instead of lead. This LAr hadronic end-cap calorimeter is overlapping with the tile calorimeter and the forward calorimeter to reduce the drop in material density in the transition region. The last part of the calorimeter is the three-stage LAr forward calorimeter with a thickness of about 10λ in total. Three modules on each side with the first one using copper as absorber and the other two being made with tungsten as the absorber contain tubes integrated in a block of the absorber material. The spaces in between are filled with liquid argon. In this way the full η range can be covered [91].

4.2.3 The muon spectrometer

Muons are minimal ionising and are the only particles[26] that can get through the dense material of the calorimeter system. They can be detected and reconstructed with high precision by the muon system. This spectrometer is the outermost part of the whole detector with two main components. Over the range of $|\eta| < 1.4$ a large barrel toroid and two additional end-cap toroids in $1.6 < |\eta| < 2.7$ are bending the muon tracks to measure the momentum and the charge of the muon candidate. In the central region of the detector the bending power of the magnetic field can be summed up to 8 Tm. The complex system of muon chambers is built out of separate trigger and high-precision tracking chambers. They are arranged in a way that each muon track is detected by at least three different chambers [91].

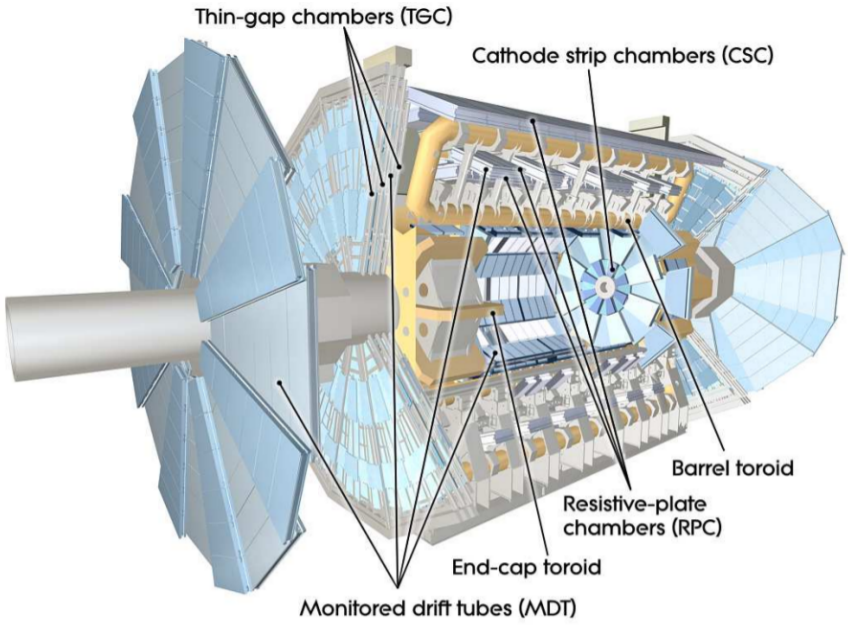


Figure 4.6: The muon system of the ATLAS experiment [91]

Four different kinds of drift chambers are used in the muon system that are optimised for different purposes. Two kinds of drift chambers mainly serve as high-precision tracking chambers. The other two are used for triggering muon candidates and provide a measurement of track coordinates that are not provided by the high-precision chambers. They are all included in Figure 4.6 where the whole muon spectrometer is displayed [91].

[26] Apart from only weakly interacting neutrinos that cannot be detected at all by the detector.

- **MDT** (**M**onitoring **d**rift **t**ubes): 1150 chambers of this kind cover almost the total η range and are used for high-precision tracking. The MDTs are drift tubes measuring the distance of a muon track to a wire using a time measurement. It is only possible to measure the track coordinate in the principal direction of the magnetic field and not along the drift tube.

- **CSC** (**C**athode **s**trip **c**hambers): 32 CSCs are located in the forward direction in $2.0 < |\eta| < 2.7$ and have a higher granularity than MDTs. They are optimised to also withstand high event rates that are present in this region of the detector. A CSC contains four layers of two cathode strips on opposite sides with anode wires in between that are perpendicular to the cathodes. In this way high-precision tracking is possible in all three directions of space.

- **RPC** (**R**esistive **p**late **c**hambers): 606 RPCs are located in front and behind the MDT or CSC, respectively, in the barrel region up to $|\eta| < 1.05$. These chambers serve as a trigger chamber by providing well-defined p_T thresholds. They also measure the coordinate of the muon track orthogonal to the component measured by the high-precision MDT or CSC. They contain two isolated perpendicular electrode-plates with ionising gas[27] in between. A muon is ionising the gas and the induced charges are measured by the electrodes.

- **TGC** (**T**hin **g**ap **c**hambers): 3588 TGCs are working in a similar way as the RPCs with the same purpose but in the forward region of the detector in $1.05 < |\eta| < 2.7$ (2.4 for triggering). They are multi-wire proportional chambers with a small wire-to-wire distance that leads to a very good time resolution. That is necessary to provide a sufficient sharp cut-off momentum for the muon that fires the trigger.

In Table 4.2 the resolution and number of measurements for each track is summarised for each drift chamber type.

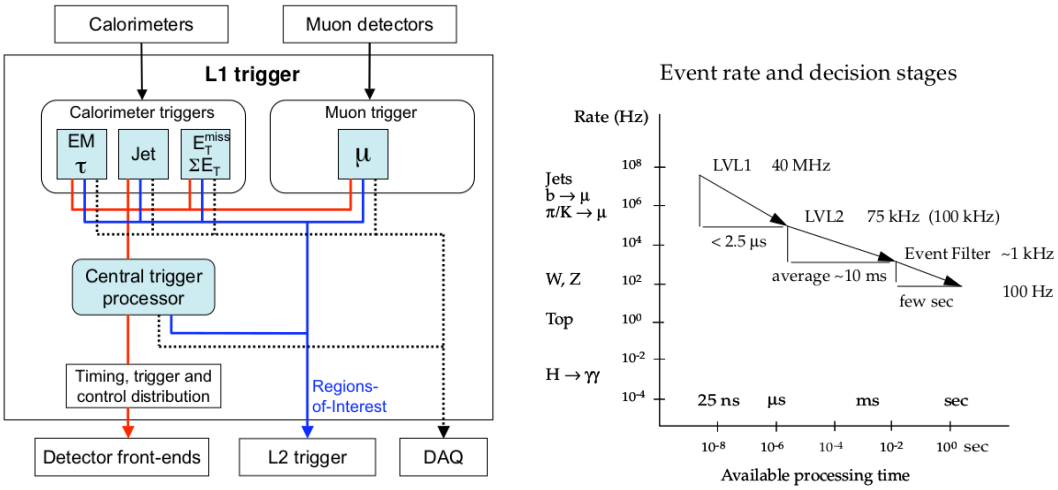
4.2.4 The trigger and data acquisition system

With an event rate of 40 MHz that can be delivered by the LHC and a data volume of approximately $O(1.3\,\text{MB})$ for each event one would end up with a total amount of data of $O(100\,\text{TB})$ that had to be stored per second. This cannot be handled by any data acquisition system and therefore events with interesting physics signatures have to be selected before they are stored [91].

[27] The used gas is a mixture of $C_2H_2F_4$/Iso-C_4H_{10}/SF_6 (94.7 %/5.0 %/0.3 %) [91]

Table 4.2: Main parameters of the ATLAS muon spectrometer. The actual resolution also depends on alignment, signal-propagation and electronics which is not included in these numbers [91].

		Resolution σ in			Measurement / track	
Type	Function	z, R	φ	time	barrel	end-cap
MDT	tracking	$35\,\mu\text{m}(z)$	—	—	20	20
CSC	tracking	$40\,\mu\text{m}(R)$	5 mm	7 ns	—	4
RPC	triggering	10 mm(z)	10 mm	1.5 ns	6	—
TGC	triggering	$2-6$ mm(R)	$3-7$ mm	4 ns	—	9



(a) (b)

Figure 4.7: (a) Block diagram of the L1 trigger level [104]. (b) Orders of magnitude of the expected event rates and processing times at the three trigger levels. An initial event rate of approximately 40 MHz is reduced stepwise to 100 Hz [91].

ATLAS is using a three-step **T**rigger and **D**ata **Aq**uisition System (TDAQ) that is gradually reducing the event rate and in the end $100-200$ events per second are being stored.

The Level-1 (L1) trigger uses custom-made electronics and is directly integrated in the hardware of the detector. Its flow is shown in Figure 4.7(a). The trigger level is able to handle an output rate of approximately 75 kHz and has to reach its decision within about 2.5 μs. High-p_T muons, electrons, photons, jets or τ-leptons that decayed hadronically are signatures that are searched for independently together with events that have large total transverse energy or large missing transverse energy. Used are the trigger chambers

of the muon system (RPC and TGC) and both EM and hadronic calorimeter, while the inner detector is not part of the L1 trigger.

The L1 trigger is followed by a two-step software-based trigger which is also called **H**igh-**L**evel **T**rigger (HLT). The first step of the HLT is the Level-2 (L2) trigger that uses Regions-of-Interest (RoI's) as seeds. The RoI's are given by the L1 trigger in the η-φ-plane. They are analysed further by the L2 also using the information from the inner detector and the full granularity of the muon system and calorimeter. The L2 makes a decision whether an event is kept or not within approximately 10 ms and reduces the event rate to about 1 kHz.

The event filter (EF) as last trigger step uses fully reconstructed events based on offline analysis algorithms. This reduces the event rate to about 200 Hz within a few seconds. The event rates and their reduction are shown in Figure 4.7(b).

An important property of each trigger is the so called prescale. The cross-section of some processes is several orders of magnitudes larger than the cross-section of other processes that are of particular interest within the ATLAS physics program. E.g. the rate of QCD-multijet production is about ten orders of magnitude larger than $H \rightarrow \gamma\gamma$ production as indicated on the y-axis in Figure 4.7(b). A trigger selecting events containing jets only would make use of a large fraction of the full trigger bandwidth. Thus not much capacity would be left for processes with smaller cross-sections. The prescale is giving the possibility to steer the trigger rate which is often used e.g. for calibration analyses. A prescale value of 1000 would cause the trigger to only fire each thousandth time although its requirements are fulfilled[28].

The data transfer between the different trigger levels is handled by the data acquisition system (DAQ). This system buffers the event information at the L1 trigger rate and transmits for instance the information about the RoI's requested by the L2 trigger. It also performs the event-building of the events fulfilling the L2 criteria and moves the information to the EF or permanent storage if also the event filter criteria are met [91].

4.3 Other particle detectors at the Tevatron

Another particle collider where many important measurements related to the top quark could be made is the Tevatron [105]. It is located at the Fermilab near Chicago and is a synchrotron with a circumference of 6.86 km. In contrast to the LHC protons and anti-protons are accelerated and brought to collision at a beam energy of up to 0.98 TeV.

[28] In case of the p_T^{rel} analysis, explained in Section 6.2.1, the prescales of triggers are adjusted such that a constant trigger rate of 1 Hz is achieved.

At the Tevatron two large experiments were built to reconstruct the collisions, called CDF and DØ. Both detectors act as multi-purpose detectors like ATLAS and CMS with its goal to identify new particles or measure particle properties with higher precision.

The Tevatron was finally shut down in 2011 because of the completion of the LHC that provides higher energies and luminosities. However, due to the different $p\bar{p}$ initial state and the performance of the detectors being tuned for years, the data taken by the two experiments still yields the most precise measurements of many parameters of the Standard Model.

CHAPTER 5

Reconstruction of physics objects

Every analysis done with data taken by the ATLAS detector relies on the reconstruction of physics objects and the experimental signature in the phase-space under study. The final state of the signal from single-top t-channel production in this thesis is characterised by one b-jet, one charged lepton and one neutrino from the decay of a real $W^{\pm}$-boson and an additional light-flavour-jet.

The lepton can be either an electron or a muon and the non-interacting neutrino can be reconstructed using the missing transverse energy. In this chapter the reconstruction of the different objects is explained and an explanation of extra cuts applied to the objects is given. In addition, a short description of tracking and the reconstruction of vertices is given that are the basis of flavour-tagging.

5.1 Tracking and vertexing

The primary vertex is the point where the hard scattering of the two protons occurred. Due to the presence of pile-up, denoting the occurrence of more than one proton-proton interaction during one bunch-crossing, usually more than one primary vertex candidate is reconstructed from the charged particle tracks. In Figure 5.1 the number of interactions per bunch crossing for the data that was taken during 2012 is shown. On average about 20 interactions took place in parallel with a maximum number of about 35 interaction.

The chosen primary vertex is defined as the one that maximises the sum of the p_T^2 of all tracks that are associated to the particular primary vertex. The selected primary vertex always has to contain at least two associated tracks.

In order to reconstruct secondary vertices within jets or use the track information for flavour-tagging, the tracks are associated to jets with a spatial matching in ΔR. As

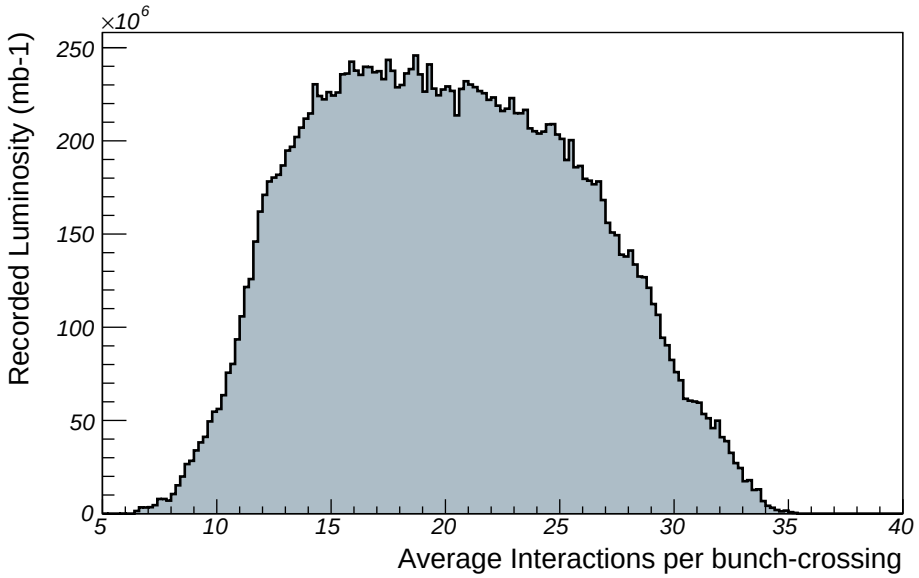


Figure 5.1: Average interactions per bunch-crossing in the dataset collected in 2012.

jets with higher momentum are more collimated, the cone size R varies as a function of the jet p_T^{jet} according to [106]:

$$R = 0.239 + \mathrm{e}^{-1.22-1.64\cdot10^{-5}\,\text{MeV}^{-1}\cdot p_T^{\text{jet}}}. \tag{5.1}$$

Associated tracks have to fulfil specific requirements to ensure that all tracks are well-measured. For the usual track selection the requirements are defined as:

$$\begin{aligned} N_{\text{PIXEL}}^{\text{hits}} + N_{\text{SCT}}^{\text{hits}} > 6, \qquad & p_T^{\text{track}} > 1\ \text{GeV}, \\ N_{\text{PIXEL}}^{\text{hits}} > 1, \qquad & N_{\text{B-Layer}}^{\text{hits}} > 0, \\ |d_0| < 1\,\text{mm}, \qquad & |z_0| \sin\theta^{\text{track}} < 1.5\,\text{mm}. \end{aligned}$$

$N_{\text{PIXEL}}^{\text{hits}}$ and $N_{\text{SCT}}^{\text{hits}}$ correspond to the number of hits in the pixel and SCT sub-detector, introduced in Section 4.2. $N_{\text{B-Layer}}^{\text{hits}}$ corresponds to the number of hits in the first layer of the pixel sub-detector denoted as the B-layer. The impact parameters, d_0 and z_0, are displayed in Figure 6.1 and defined in the transverse or longitudinal plane with respect to the beam axis. The transverse impact parameter, d_0, is the distance of closest approach of the track to the position of the selected primary vertex in the r-φ-projection. The longitudinal impact parameter, z_0, is the difference between the z-position of the reconstructed primary vertex and the position of the track at its point of closest approach

in the r-φ-plane. All associated tracks are then used for flavour-tagging as described in Section 6.1.2.

To reconstruct secondary vertices the track requirements are slightly loosened in order to optimise the reconstruction efficiency. The changed cuts correspond to:

$$N_{\text{PIXEL}}^{\text{hits}} + N_{\text{SCT}}^{\text{hits}} > 6, \qquad p_T^{\text{track}} > 400 \text{ MeV},$$
$$N_{\text{PIXEL}}^{\text{hits}} > 0, \qquad |d_0| < 3.5\,\text{mm}.$$

Building an inclusive secondary vertex within a jet starts with building all two-track pairs of the remaining associated tracks if they form a good vertex. Tracks are considered if they are significantly displaced from the chosen primary vertex. From these vertices the ones compatible with long-lived particles or material interaction are rejected[29]. Details about this procedure can be found in Ref. [107].

All tracks that are associated to a remaining two-track pair are fitted to a single inclusive secondary vertex. The algorithm includes fitting the vertex and estimating the χ^2 for each track to the fitted vertex. It then removes iteratively the track with the largest χ^2 contribution until three criteria are met:

- The fit probability of the secondary vertex must be larger than 0.001.
- The invariant mass of the vertex, denoted as SV0 mass, is calculated from its associated tracks and has to be smaller than 6 GeV.
- The χ^2 of each track is required to be smaller than 7.

After the secondary vertex is formed tracks that have been rejected during the creation of the two-track vertices are re-incorporated if possible. Furthermore, the jet direction is corrected to match the line joining the selected primary and reconstructed inclusive secondary vertex [106, 107].

5.2 Jets

In this section the reconstruction of jets from the information collected by the electromagnetic and hadronic calorimeter is described. This includes the reconstruction based on topological clusters that have been calibrated with the local cluster weighting (LCW) method [108] and the anti-k_t clustering algorithm [109] using a width parameter of

[29] They are rejected if the mass is compatible with a K_s^0-meson, a Λ^0-baryon or photon conversion. Additionally, they are rejected if the vertex is located close to one of the three layers of the Pixel sub-detector.

$R = 0.4$. In addition, cuts are applied that ensure a good quality of the reconstructed objects.

5.2.1 Jet reconstruction

The reconstruction starts with building topological clusters by clustering calorimeter cells that are calibrated at the electromagnetic (em) scale[30] if the signal within the cells is significant compared to noise [111]. The calibration with the LCW method then starts with a classification based on the location and the shape of the formed cluster in the three categories: electromagnetic, hadronic or noise. After this the cells of hadronically classified clusters are reweighted depending on the cluster location, energy and signal density. That corrects for the lower response of the calorimeter to hadronic energy deposits. In this step electromagnetic clusters are kept unchanged. Afterwards two corrections to all clusters are applied being a correction for the energy loss due to dead materials close or within the cluster (DM correction) and a correction for the clustering itself (out-of-cluster correction)[31]. The clusters now calibrated at the local hadronic energy scale can be used as the input for the anti-k_t algorithm that forms jets feasible to be used in the later analysis.

To retrieve the final physics jets calibrated to interaction level three more correction steps are being applied. Firstly, a pile-up correction from in-situ measurements is applied. The correction subtracts the average additional energy due to additional pile-up interactions from the energy measured by the calorimeter. Secondly, a vertex correction makes sure that the jet direction is pointing to the location of the primary vertex instead of the geometrical centre of the detector. In a last step, the jet energy and direction are corrected based on comparisons between reconstructed observables with truth jets in simulation [108]. After the final calibration described in Ref. [113] the jets are referred to as calibrated with the LCW+JES scheme.

5.2.2 Additional jet requirements

It is possible that due to pathological noise bursts in the calorimeter a jet can mistakenly be reconstructed from noisy cells. This kind of jets has been studied in detail in Ref. [114] using data collected in 2010 corresponding to an integrated luminosity of $0.3\,\mathrm{nb}^{-1}$. Cri-

[30] Electromagnetic scale means the raw signal from the calorimeter only including basic corrections derived from electron signals. Not included are corrections for high-precision electron or photon reconstruction as it is described in Ref. [110].

[31] One should note that no information of the later jet structure is used for the calibration which is why the procedure is called local. The corrections applied are described in much more detail in Ref. [112]

teria have been developed to flag these jets as "bad". To ensure that only good quality jets are being used events that contain at least one "bad" jet with $p_T > 20$ GeV are being rejected.

If a jet and an electron are overlapping in the η-φ-plane within $\Delta R < 0.2$ the jet is removed since most likely the electron and the jet are originating from the same physical object.

To reject jets originating from pile-up the jet vertex fraction (JVF) criterion is applied to all jets with $|\eta| < 2.4$ and $p_T < 50$ GeV. The jet vertex fraction is defined as the sum of the transverse momenta of all tracks of the jet that are also coming from the primary vertex over the sum of transverse momenta from all tracks associated to the jet. The jet is accepted if $|\text{JVF}| > 0.5$ is fulfilled.

Finally, kinematic cuts on the jet candidates are applied. A jet must have a transverse momentum of $p_T > 30$ GeV and $|\eta| < 4.5$. This includes forward jets that are only reconstructed from the calorimeter. No information from the inner tracking detector is available above $|\eta| < 2.5$ as explained in Section 4.2.1. To remove some mismodelling in the transition region between central and forward calorimeter the threshold on the transverse momentum is raised to $p_T > 35$ GeV within $2.7 < |\eta| < 3.5$.

5.3 Muons

The reconstruction of muons can be divided in two parts. The event under study has to be selected by a dedicated muon trigger. Then object related and kinematic cuts are being applied. In summary a muon is accepted if it has a transverse momentum of $p_T > 25$ GeV and is in the pseudo-rapidity region of $|\eta| < 2.5$. In addition, the muon candidate has to fulfil the criteria described below, namely being trigger matched, well reconstructed and isolated.

5.3.1 Muon trigger

Two different triggers using the three-step trigger system of ATLAS, explained in Section 4.2.4, are used to collect events with a muon. These events are then referred to as part of the muon channel within the analysis.

Both triggers are starting with a L1 muon trigger chamber track with a threshold of $p_T > 15$ GeV which is matched with a muon that is reconstructed on the EF trigger level. One of the two triggers requires the muon to be isolated[32] and have a transverse

[32]The isolation requirement in this case is based on reconstructed tracks. The isolation variable is defined as the sum of the transverse momenta with $p_T > 1$ GeV found in the inner detector within a

momentum of $p_T > 24$ GeV. The second trigger uses a raised transverse momentum threshold at $p_T > 36$ GeV but has no isolation requirement [115].

5.3.2 Muon reconstruction

Muon tracks are being reconstructed independently with the inner tracking detector (ID) using the ID track requirements explained in Section 5.3.2.1 and the muon spectrometer (MS). The two tracks are then combined to use the complete track information accounting for material effects of the ATLAS detector structure. There are two different approaches to combine the two tracks. The first combines the two using a global refit of the two tracks and the second performs a search for segments and tracks in the muon spectrometer using the reconstructed ID track as seed. Since with the second approach it is possible that a muon without matched ID and MS track is accepted, the result of the global fit is always used as final muon track under the circumstance that the mentioned segments and tracks search was also successful [116].

5.3.2.1 Muon ID track requirements

A number of cuts are applied to ensure that the pure ID muon track is of good quality and not coming e.g. from pile-up interactions. These cuts are based on the three subsystems of the inner detector, explained in Section 4.2.1, and are defined as:

$$
\begin{aligned}
N^{\text{hits}}_{\text{PIXEL}} + N^{\text{dead sensors}}_{\text{PIXEL}} &> 0, \\
N^{\text{hits}}_{\text{SCT}} + N^{\text{dead sensors}}_{\text{SCT}} &\geq 5, \\
N^{\text{holes}}_{\text{PIXEL}} + N^{\text{holes}}_{\text{SCT}} &< 3, \\
N^{\text{hits}}_{\text{TRT}} + N^{\text{outliers}}_{\text{TRT}} &> 5 \qquad (\text{if } 0.1 < |\eta^{\text{track}}_{\text{muon}}| < 1.9), \\
0.9 \cdot (N^{\text{hits}}_{\text{TRT}} + N^{\text{outliers}}_{\text{TRT}}) &> N^{\text{outliers}}_{\text{TRT}} \qquad (\text{if } 0.1 < |\eta^{\text{track}}_{\text{muon}}| < 1.9).
\end{aligned}
$$

Here, N_i^{hits} refers to the number of hits and $N_i^{\text{dead sensors}}$ is the number of crossed dead sensors in the corresponding inner detector subsystem i. N_i^{holes} corresponds to the number of holes in the PIXEL or SCT subsystem, see Section 4.2.1. In case of the TRT a successful TRT extension is required within its acceptance where $N_{\text{TRT}}^{\text{outliers}}$ corresponds to the number of TRT outliers.

To suppress tracks from pile-up events the longitudinal impact parameter z_0, defined in Section 5.1, of the muon track has to be $z_0 \leq 2\,\text{mm}$.

cone of $\Delta R < 0.2$ around the muon candidate. This does not include the muon itself and the muon is called isolated if $\sum p_T/p_T(\mu) < 0.12$.

5.3.3 Muon isolation

In order to suppress the contribution of muons from heavy-flavour decays the so-called mini-isolation requirement [117] was developed. In the mini-isolation approach the cone radius is allowed to vary as a function of the muon p_T. This improves especially the pile-up robustness and performance in boosted top quark topologies. The mini-isolation variable $\text{MiniIso10}_\text{R}$ is defined as the scalar sum of the transverse momenta with $p_T > 1$ GeV of all tracks within a cone with radius $R_\text{iso} = \frac{10 \text{ GeV}}{p_T(\mu)}$. Not included is the reconstructed muon track itself. The muon is being accepted as isolated if

$$\frac{\text{MiniIso10}_\text{R}}{p_T(\mu)} < 0.05. \tag{5.2}$$

In addition, every muon is removed that is overlapping within $\Delta R < 0.4$ with a jet with $p_T > 25$ GeV fulfilling the JVF requirement, see Section 5.2.2.

5.3.4 Performance of the muon reconstruction

The muon reconstruction efficiency and muon momentum resolution and scale have been measured based on the reconstruction of $Z^0 \to \mu^+\mu^-$, $J/\Psi \to \mu^+\mu^-$ and $\Upsilon \to \mu^+\mu^-$ decays in simulated events and experimental data. It was found that a reconstruction efficiency of more than 0.98 can be achieved that is independent from pile-up and the geometry of the detector. Scale factors close to unity are obtained to correct for small differences between data and simulation [116, 118].

The mass spectrum of $Z \to \mu^+\mu^-$ decays provides a possibility to estimate the muon momentum scale and resolution. The mass resolution of the di-muon pair ranges from $1.5 - 3$ GeV and smearing corrections of the order of 0.1% are applied to correct for small differences between simulation and collision data [116, 118].

The uncertainties on the three sources are propagated to the final measurement of the top quark mass and are further explained in Section 10.4.

5.4 Electrons

The reconstruction of electrons also can be divided into two parts. Events are selected by dedicated single electron triggers and electrons are reconstructed from inner detector tracks and the electromagnetic calorimeter. The electron candidate has to fulfil offline reconstruction criteria that will be explained in the following. As for the muon it is required that an electron candidate has a transverse momentum of $p_T > 25$ GeV and is

in the pseudo-rapidity region of $|\eta_{\text{clus}}| < 2.47$ to be accepted. η_{clus} indicates the position of the electromagnetic cluster. In addition, the selected electron has to be isolated and matched to the trigger object.

5.4.1 Electron trigger

The trigger selection is done in a similar way as in the muon channel described in Section 5.3.1. Events that are selected by a single electron trigger are referred to as part of the electron channel.

At the L1 trigger level an electromagnetic energy deposit with $E_T > 30$ GeV or $E_T > 18$ GeV when the electron is also isolated is required[33]. The cluster at L1 then has to match a track at EF trigger level. The EF trigger object must have a transverse momentum of $E_T > 60$ GeV or $E_T > 24$ GeV plus the isolation criterion. The triggered electron candidates must meet the requirements defined as "medium1". These selection cuts correspond to the "medium++" definition used in the offline reconstruction of electrons and is explained in Ref. [119].

5.4.2 Electron reconstruction

To reconstruct electrons two different algorithms are used. About half of the electrons are reconstructed by an algorithm that uses calorimeter information as seed and the second half is reconstructed using calorimeter as well as tracking information [120, 121]. The electron candidates have to fulfil the so-called "tight++" definition that is explained in detail in Ref. [122]. This selection makes use of information from the first and second layer calorimeter, from the two silicon detectors and the TRT. Additionally, the η and φ position of the reconstructed object is matched between tracking and the calorimeter.

The overlap region between the barrel and end-cap of the calorimeter $1.37 < |\eta_{\text{clus}}| < 1.52$ is vetoed because of the limited calorimeter instrumentation. Events with bad quality clusters or fake clusters originating e.g. from noise bursts in the LAr calorimeter are explicitly rejected to ensure a good quality of the reconstructed electrons. To reject electrons from pile-up tracks the longitudinal impact parameter of the electron track has to be $z_0 \leq 2\,\text{mm}$.

[33] The isolation criterion is fulfilled if the total transverse momentum within a cone of $\Delta R < 0.2$ around the electron is less than 10% of the transverse energy of the electron cluster.

5.4.3 Electron isolation

Multiple sources are able to fake an electron like hadronic jets, electrons from heavy flavour decays or photon conversion. These sources can be suppressed by an isolation criterion since signal electrons from $W^{\pm}$-boson decays are usually isolated from other calorimeter activity. The isolation criterion uses a minimal amount of information from the calorimeter and only few tracks in a cone around the electron in the η-φ-plane. The cuts are tuned in a way that a uniform isolation efficiency across η_{clus} and the transverse momentum $E_T = \frac{E}{\cosh(\eta)}$ is achieved[34]. For the track and the EM calorimeter two cuts and corresponding working points are defined:

- **EtCone20@90:** Cell based isolation at 90% efficiency
- **PtCone30@90:** Track isolation at 90% efficiency

The track isolation variable is computed by summing up the transverse momentum of all ID tracks within a cone of $\Delta R < 0.3$ where the track requirements are slightly different compared to the ones explained in Sections 5.1 or 5.3.

The cell based isolation variable is calculated as the sum of calorimeter cell energies at the electromagnetic scale within a cone of $\Delta R < 0.2$ around the centre of the cluster. A grid of 5x7 cells in the centre of the cluster is explicitly excluded. Details about the isolation criteria can be found in Ref. [123, 124].

5.4.4 Performance of the electron reconstruction

The reconstruction and identification efficiency of electrons was measured based on a tag-and-probe ansatz using $Z \to e^+e^-$, $J/\Psi \to e^+e^-$ decays. Scale factors are used to correct for differences in the efficiencies between data and simulation. They are close to unity with deviations of a few percent in regions of low E_T or high η_{clus} [120, 125].

The electron energy scales and resolution have been estimated using different resonances ($Z \to e^+e^-, J/\Psi \to e^+e^-$) or E/p studies with isolated electrons from $W^{\pm}$-boson decays. The techniques used are explained in Ref. [121]

The electron cluster energy is smeared in the simulation to correct for differences between data and simulation. All of these sources are being assigned a dedicated systematic uncertainty which is explained in Section 10.4.

[34] In this case E is the energy of the cluster deposits and η corresponds to the associated track.

5.5 Missing transverse momentum

Non-interacting particles like neutrinos can be reconstructed via the missing transverse momentum. $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the absolute value of the vectorial sum of all topological clusters in the event. This is taking into account the corrections mentioned in the previous sections due to the reconstruction of electrons, muons or jets. $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated by summing up the individual components:

$$E_{x(y)}^{\mathrm{miss}} = E_{x(y)}^{\mathrm{miss},e} + E_{x(y)}^{\mathrm{miss,jets}} + E_{x(y)}^{\mathrm{miss},\mu} + E_{x(y)}^{\mathrm{miss,SoftTerm}}. \tag{5.3}$$

Here, the individual terms are calculated as the negative sum of the calibrated reconstructed objects which are projected onto either the x or y direction. Noise contributions from the calorimeter are suppressed by using only energy deposits from topological clusters referred to as $E_{x(y)}^{\mathrm{miss,SoftTerm}}$ that are not associated to another reconstructed jet or electron. To avoid double counting of energy in case of the muons the parametrised energy loss of the muon in the calorimeter is subtracted [126]. Finally, the missing transverse momentum is calculated from the x and y components: $E_T^{\mathrm{miss}} = \sqrt{\left(E_x^{\mathrm{miss}}\right)^2 + \left(E_y^{\mathrm{miss}}\right)^2}$.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ is then referred to as MET_RefFinal_AntiKt4LCTopoJets_tightpp which corresponds to the reconstruction used for the leptons and jets. Systematic uncertainties have been determined for the $E_{x(y)}^{\mathrm{miss,SoftTerm}}$ term [127] and are propagated to the measurement. This is explained in more detail in Section 10.4.

5.6 Identification of b-quark jets

Different algorithms exploit special characteristics of jets induced by b-quarks. The procedures range from simple impact parameter or secondary vertexing algorithms to sophisticated decay chain reconstruction and versatile multivariate techniques combining individual inputs. These algorithms and their performance is explained in much more detail in Chapter 6.

CHAPTER 6

Flavour tagging in ATLAS

One important part of the reconstruction of various final states is the identification of jets that have been induced by a b-quark. This is especially the case for all processes containing top quarks because the top quark decays almost exclusively to bottom quarks, see Section 3.2. In general background processes that contain mostly light-flavour-jets can be suppressed very effectively using b-tagging techniques. In the first part of this chapter the concept behind different b-tagging algorithms is explained and a description of their performance is given. In the second part the calibration of these b-tagging algorithms will be explained which is of great importance to ensure a reliable performance when applying the algorithms to collision data.

6.1 B-tagging algorithms

B-hadrons have a relatively long lifetime of the order of $\tau \approx 1.5\,\text{ps}$. This means that the b-hadron is able to travel a measurable distance before its decay. This significant flight path decay length $\langle l \rangle = \beta\gamma c\tau$ is about 3 mm in average in the transverse plane. Two different approaches are used to exploit this topology:

- When the b-hadron decays after travelling the distance l a typical secondary vertex can be seen. It is displaced from the primary vertex, PV, where the hard scattering collision has happened. The secondary vertex can explicitly be reconstructed using the inner tracking detector of ATLAS with the algorithm explained in Section 5.1.

- Even if no secondary vertex is reconstructed due to the limited detector resolution, charged particle tracks from the b-hadron have large associated track impact parameters. This quantity was explained in detail in Section 5.1.

The track impact parameter is displayed in Figure 6.1 within the cone of the so-called reconstructed b-jet together with the primary vertex and the displaced secondary vertex.

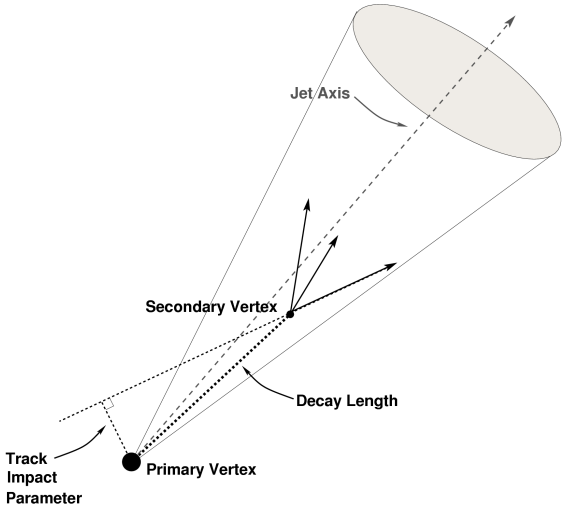


Figure 6.1: A secondary vertex with a significant decay length indicates the presence of a long-lived particle in the jet. The secondary vertex is reconstructed from tracks with a large impact parameter significance with respect to the primary vertex [107].

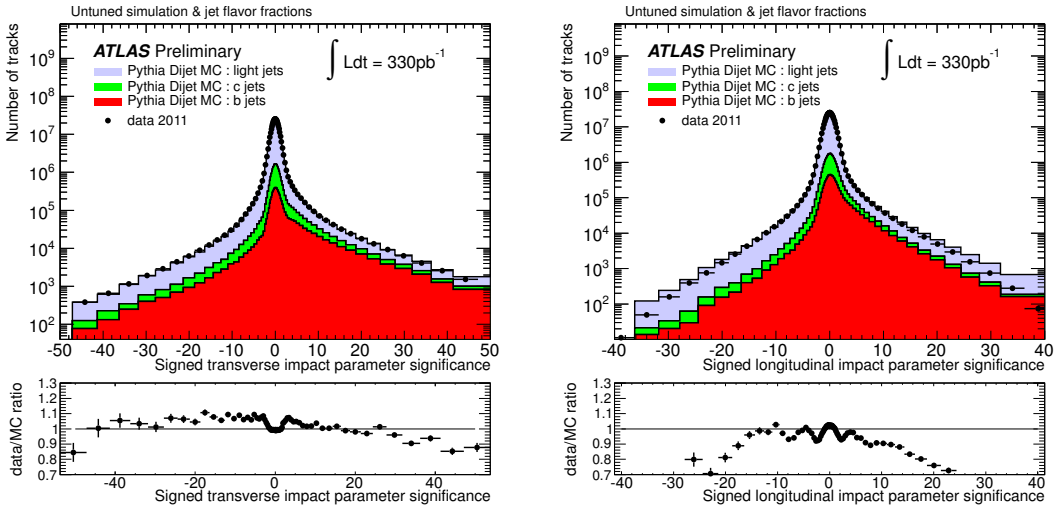
Each tagging algorithm is evaluating a weight, called b-tagging weight, for each jet. It is designed such that the higher the weight the higher the probability that the jet is a b-quark induced jet.

6.1.1 Flavour labelling in simulation

In simulation the information about the origin of a jet is known from truth information. The labelling of the flavour of a jet in simulation is derived by looking at the true partons found within a cone with size $\Delta R < 0.3$ around the jet direction axis. If a b-quark is found the jet is labelled as a b-jet. If no match is found it is checked for a c-quark within the same cone and a τ-lepton afterwards. The jet is labelled a c-jet or a τ-jet, respectively, if a match is found. A jet without any association is labelled as a light-flavour-jet.

6.1.2 Impact parameter based algorithms

Figure 6.2 shows the transverse (a) and the longitudinal (b) signed impact parameter significance for data at $\sqrt{s} = 7$ TeV. These two parameters are the basis of the algorithms that are explained in the following.



(a) Transverse impact parameter significance (b) Longitudinal impact parameter significance

Figure 6.2: Distributions of S_{d_0} (a) and S_{z_0} (b) with respect to the primary vertex for tracks of b-tagging quality associated to jets for collision data (solid black points) and for simulated data (The filled histograms show the different flavours based on the labelling procedure explained in Section 6.1.1.) [128].

JetProb algorithm A basic algorithm based on the impact parameter of all reconstructed tracks associated to a jet is the JetProb algorithm [106]. It uses the signed transverse impact parameter significance $S_{d_0} = \frac{d_0}{\sigma_{d_0}}$ to calculate the probability that the track is associated to the primary vertex. The information for each individual track is then combined in a likelihood used to estimate the b-tagging weight for each jet.

IP3D algorithm The more advanced IP3D algorithm [106, 110] is using both the longitudinal impact parameter significance $S_{z_0} = \frac{z_0}{\sigma_{z_0}}$ and S_{d_0} taking advantage of the correlations between the two. A log-likelihood ratio technique is applied to compare the 2-dimensional (S_{d_0}, S_{z_0}) input variable distribution for each track to distributions obtained in simulation for both the b-jet and light-flavour-jet hypothesis. The ratio of the probabilities is then giving a weight for each track. The b-tagging weight of the jet is obtained as the sum of the logarithms of the individual track weights.

6.1.3 Vertex based algorithms

The reconstruction of secondary vertices within jets is described in Section 5.1. The following algorithms are exploiting this approach.

SV0 algorithm A simple algorithm that is discriminating between b-jets and light-flavour-jets is called SV0. The discriminating variable of the algorithm is the signed decay length significance $S_l = \frac{l}{\sigma_l}$. It is the distance, shown in Figure 6.1, between the selected primary and the reconstructed secondary vertex divided by the measurement's uncertainty. A general limitation of this algorithm is the reconstruction efficiency of the secondary vertex of about 70%.

SV1 algorithm A more advanced algorithm that is using the same log-likelihood ratio technique as the IP3D algorithm, explained in Section 6.1.2, is the SV1 algorithm. It uses the same secondary vertex reconstruction as the SV0 algorithm and combines three properties of the vertex found in a likelihood. These properties are:

- The invariant mass of all tracks used to reconstruct the secondary vertex
- The ratio of the sum of the energies of these tracks to the sum of the energies of all tracks in the jet
- The number of two-track vertices

In addition, the information of the ΔR between the jet direction and the direction of the line between the reconstructed primary and secondary vertex is added to the likelihood. To also take advantage of the correlations between the first two properties of the secondary vertex those two are combined in a 2-dimensional distribution.

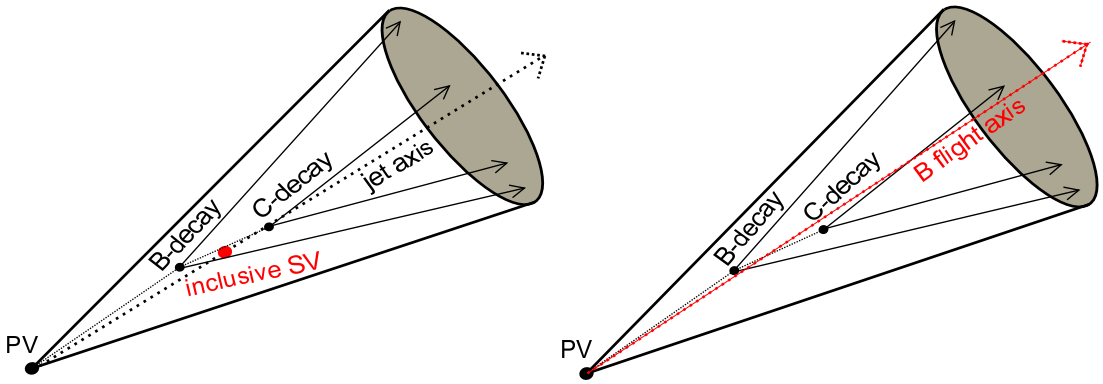
JetFitter algorithm Another algorithm using even more information than the SV1 algorithm is called JetFitter, explained in more detail in Ref. [110, 129].

The default reconstruction of secondary vertices builds an inclusive vertex also in the case of topologies of weak b- and c-hadron decay chains inside the jet. This is displayed in Figure 6.3(a). The JetFitter algorithm exploits these topologies by fitting a common line between the PV and the SVs of the b- and c-quark decays using a Kalman filter [130], displayed in Figure 6.3(b).

With this approach one can reconstruct an approximated b-hadron flight axis. Consequently the different vertices are not necessarily merged even if they are formed only by single associated charged tracks. The reconstructed decay chain topology is in the end saved in six variables which are:

- Number of vertices with at least two tracks
- Number of tracks at the reconstructed vertices

- Number of single track vertices at the reconstructed b-hadron flight axis
- Invariant mass of all charged particle tracks attached to the decay chain
- Energy of all charged particle tracks attached to the decay chain divided by the sum of the energies of all charged particles associated to the jet
- Weighted average vertex position divided by its uncertainty



(a) Inclusive secondary vertex reconstruction (b) Multi-vertex fit reconstruction

Figure 6.3: Default reconstruction of an inclusive secondary vertex (a) and multi-vertex fit (b) using the b-hadron flight direction constraint done by the JetFitter algorithm.

The variables are combined in an artificial neural network with eight input nodes also including the p_T and the η of the jet[35]. The neural network output has three different output nodes corresponding to the three hypotheses of a b-, c- or light-flavour-jet, called P_b, P_c and P_l. The b-tagging weight is then given by the relation $w_{\text{JetFitter}} = \ln \frac{P_b}{P_l}$ which is optimised to reject especially light-flavour-jets.

6.1.4 Combined algorithms

Different combinations of the explained algorithms are possible to achieve an even better separation between b-jets and light-flavour-jets. They have been studied in detail in Ref. [128].

[35]Before the training a 2-dimensional reweighing is used to ensure that the distribution of p_T as well as the η distribution of the jets are flat for the different jet flavours.

IP3D+SV1 The IP3D as well as the SV1 algorithm are using the same log-likelihood ratio technique and can easily be combined by summing up the individual b-tagging weights. This yields for the combined b-tagging weight: $w_{\mathrm{IP3D+SV1}} = w_{\mathrm{IP3D}} + w_{\mathrm{SV1}}$

IP3D+JetFitter: JetFitterCOMBNN and JetFitterCOMBNNc The combination of the IP3D and JetFitter algorithm is done by adding another input node to the artificial neural network of JetFitter explained in Section 6.1.3. The neural network again has three output nodes P_b, P_c and P_l and the b-tagging weight corresponding to the algorithm called JetFitterCOMBNN is calculated by using $w_{\mathrm{JetFitterCOMBNN}} = \ln \frac{P_b}{P_l}$.

Since the JetFitter algorithm is also providing an individual weight for c-jets another specific tuning is created called JetFitterCOMBNNc. The b-tagging weight of JetFitterCOMBNNc is defined as $w_{\mathrm{JetFitterCOMBNNc}} = \ln \frac{P_b}{P_c}$. It provides a better discrimination between b-jets and c-jets with the caveat that the separation between b-jets and light-flavour-jets is worsened.

For instance the main background in the case for single-top t-channel production is the production of a real $W^{\pm}$-boson in combination with an additional c-jet. This large background contribution can be significantly reduced by using a b-tagging algorithm tuned to reject c-quark induced jets.

MV1 and MV1c The tagging algorithm that, presently, is mostly used in ATLAS analyses is the MV1 algorithm which provides a final combination of the explained individual or already combined algorithms. The input variables are the weights of the SV1, IP3D and JetFitterCOMBNN algorithms. Furthermore, an input variable is used to classify each jet in a two-dimensional grid in (p_t, η) of the jet. This is necessary since the underlying training samples are simulated b-jets (signal) and simulated light-flavour-jets (background). The kinematic spectra of the two samples are different and therefore weights are applied to each jet depending on its (p_t, η)-category to avoid any kinematic bias.

The neural network used is implemented in the TMVA framework [131] and consists of the four mentioned input nodes, two hidden layers with three and two nodes, respectively, and one output node which gives the b-tagging weight of MV1.

By replacing the input of the JetFitterCOMBNN weight with the weight calculated by JetFitterCOMBNNc again a specific instance of the MV1 algorithm is created that is optimised to discriminate between b-jets and c-jets. This is referred to as the MV1c algorithm and is used as the main b-tagging algorithm to measure the top quark mass in this thesis, see Section 8.1.2.

6.1.5 Comparison of b-tagging algorithms

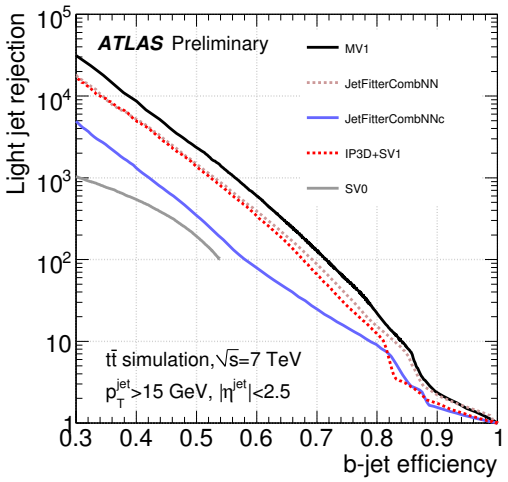
The performance of each algorithm has to be studied to enable them for reliable usage in physics analyses. The most important quantity is the tagging efficiency which is defined as

$$\varepsilon_b = \frac{\text{number of true } b\text{-jets that are tagged by the algorithm}}{\text{number of all true } b\text{-jets}} \tag{6.1}$$

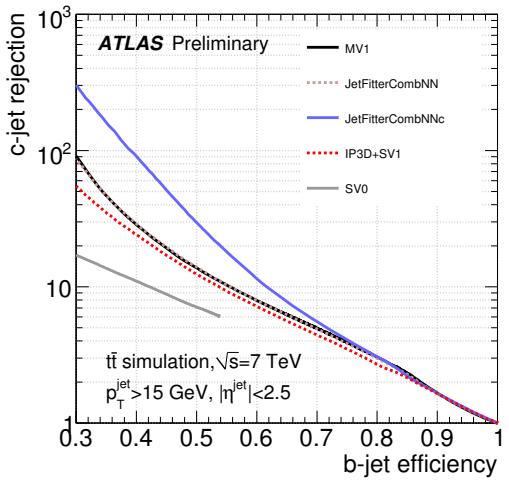
in the case of b-jets. In the same way the tagging efficiency can be defined for c-jets, ε_c, or light-flavour-jets, ε_l. The rejection is defined as

$$r_{light} = \frac{\text{number of all true light-flavour-jets}}{\text{number of true light-flavour-jets that are tagged by the algorithm}} \tag{6.2}$$

for light-flavour-jets and analogue for c-jets, r_c. In Figure 6.4(a) the light-flavour-jet



(a) Light-flavour-jet rejection as a function of the b-jet tagging efficiency



(b) c-jet rejection as a function of the b-jet tagging efficiency

Figure 6.4: Rejection of light-flavour-jet (a) and rejection of c-jets (b) as a function of the b-jet tagging efficiency for different b-tagging algorithms based on simulated $t\bar{t}$ events [132].

rejection is shown as a function of the b-tagging efficiency and Figure 6.4(b) shows the c-jet rejection against the b-tagging efficiency for different tagging algorithms. Both plots are made by varying the cut value on the output b-tagging weight distribution of the algorithm. To ensure the comparability of the different algorithms the jets used are

always stemming from the identical simulated $t\bar{t}$ sample. The jets must fulfil the selection criteria of $p_T > 15$ GeV and $|\eta| < 2.5$ and are counted as b-tagged if the b-tagging weight is larger than the respective cut value. It is clearly visible in Figure 6.4 that a higher rejection can be achieved with the more advanced algorithms while the b-jet tagging efficiency remains unchanged. Also the higher rejection of the algorithms optimised to reject c-jets is clearly visible in Figure 6.4(b).

6.2 Performance measurements of b-tagging algorithms

In Section 6.1.5 the important properties of b-tagging being the tagging efficiency and the rejection have been explained. In addition, the mistag rate is defined as the reciprocal of the rejection[36]. To ensure that no artificial bias to a physics analysis is introduced due to a b-tagging selection the properties have to be measured in data and be compared to the expectation from simulation. This is done independently for b-jets, c-jets and light-flavour-jets by various performance measurements. All of these b-tagging calibration analyses are using a sample that is dominated by the specific flavour.

In this section a basic measurement of the b-jet tagging efficiency in collision data using dijet events containing jets with muons will be explained in detail in Section 6.2.1. A brief description of other measurements based on $t\bar{t}$ events is following and in the end an overview of measurements of the c-jet tagging efficiency and the light-flavour-jet mistag rate is given.

6.2.1 Measuring the b-jet tagging efficiency with p_T^{rel}

With the p_T^{rel}-method the b-jet tagging efficiency can be measured in a sample of inclusive jets that contain a muon within the reconstructed jet cone. It is a very robust method used especially for an early calibration. During data taking it can be updated regularly since it uses a very simple event selection. The calibration results are given in the form of jet p_T and η-depending scale factors

$$\kappa_{\varepsilon_b} = \kappa_{\varepsilon_b}^{\mathrm{data/sim}} = \frac{\varepsilon_b^{\mathrm{data}}}{\varepsilon_b^{\mathrm{sim}}}, \tag{6.3}$$

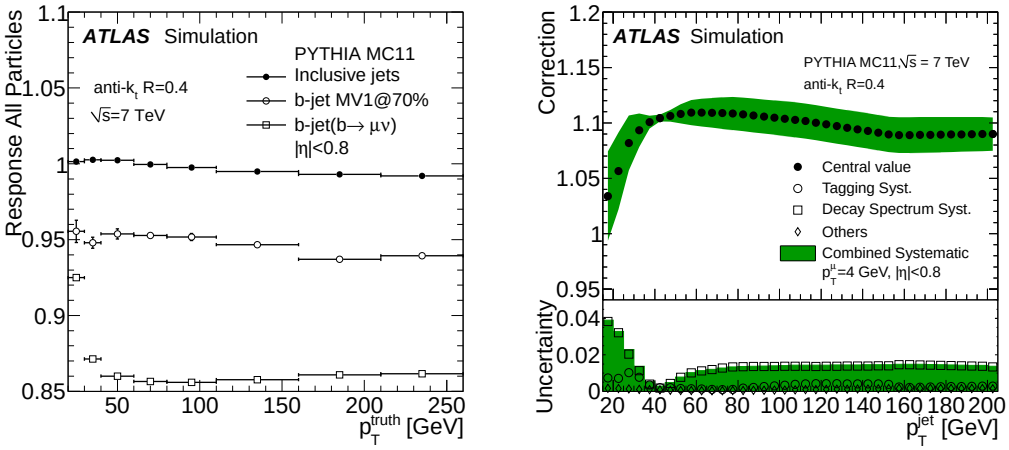
that can be used to correct the efficiency in simulation, $\varepsilon_b^{\mathrm{sim}}$, to match the one observed in data called $\varepsilon_b^{\mathrm{data}}$.

[36] The mistag rate is often used for light-flavour-jets instead of the efficiency, ε_l, which is actually equivalent.

6.2.1.1 Semileptonic jet correction

Before the selection the transverse momentum and the direction of the jets are corrected for the component of a reconstructed muon and the neutrino assigned to the jet [113]. In this way the kinematic quantities of the initial parton are restored also in the case of a semileptonic decay.

The correction is done in two steps as follows. Firstly, the jet is corrected for the muon contribution. This is done by subtracting the average energy deposition of a muon in the calorimeter [133] in the flight direction of the reconstructed muon from the jet and adding the full 4-vector of the reconstructed muon. Secondly, the contribution of the assumed neutrino to the transverse momentum of the jet is added. This contribution been derived using the all-particle response $\mathcal{R} = \frac{p_T^{\text{jet+muon}}}{p_T^{\text{truth,all}}}$ with $p_T^{\text{jet+muon}}$ being the momentum of the reconstructed jet and the reconstructed muon. $p_T^{\text{truth,all}}$ is being built from all final states with a lifetime of $\tau > 10\,\text{ps}$ on truth level. The all-particle response is shown in three



(a) Average jet response as a function of true transverse momentum of jets built using all stable particle in three different samples.

(b) Semileptonic correction, as a function of calorimeter jet p_T, used to transform the p_T of a jet in the semileptonic sample to the p_T of a jet in an inclusive sample of b-jets

Figure 6.5: (a) Average jet response as a function of true transverse momentum of jets built using all stable particles. The different points correspond to a sample of inclusive jets, a sample of b-jets tagged by the MV1 tagging algorithm and a sample of b-jets decaying semileptonically. (b) The semileptonic correction, as a function of calorimeter jet p_T. Systematic uncertainties are shown as coloured bands [113].

different samples in Figure 6.5(a) and the resulting correction is shown in Figure 6.5(b) together with its associated systematic uncertainties. It can be seen that the semileptonic

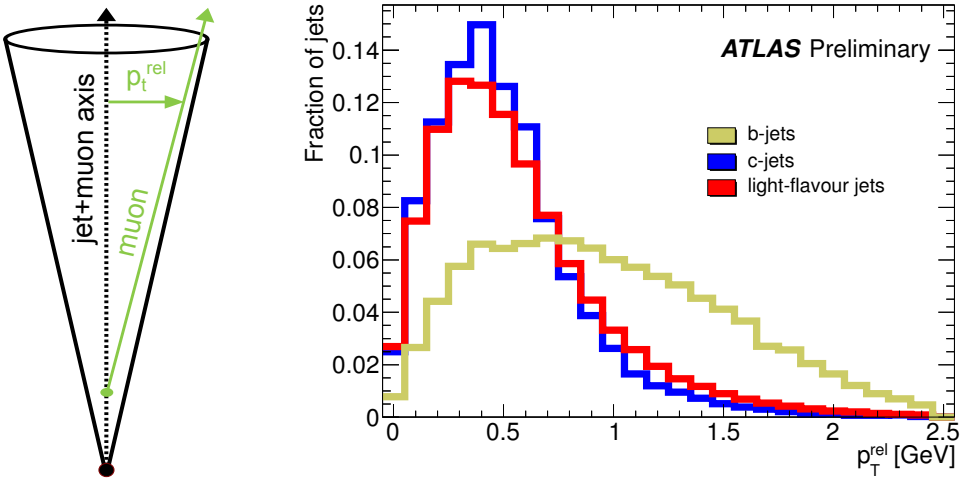
correction of the transverse momentum of the jet ranges up to about 10 % in a wide p_T^{jet} range.

6.2.1.2 Key ingredients: The p_T^{rel} variable

The p_T^{rel}-method uses the transverse momentum component with respect to the combined jet+muon axis of a soft muon reconstructed within the jet cone. The variable illustrated in Figure 6.6(a) is called p_T^{rel}. It is defined in Equation (6.4) based on the momentum axis vectors of the jet after the semileptonic correction[37], $\vec{p}_{j+\mu}$, and of the muon, $\vec{p_\mu}$ [134]:

$$p_T^{\text{rel}} = \sqrt{\vec{p_\mu}^{\,2} - \frac{(\vec{p}_{j+\mu} \cdot \vec{p_\mu})^2}{\vec{p}_{j+\mu}^{\,2}}}. \tag{6.4}$$

Due to the higher mass, muons from b-hadron decays tend to have a harder p_T^{rel} spectrum than muons in c- and light-flavour-jets. This makes the p_T^{rel} variable sensitive to the amount of jets induced by b-quarks in a sample. The number of b-jets can be estimated by fitting the p_T^{rel} distribution to templates that have been derived separately for b-, c- and light-flavour-jets. These templates are displayed in Figure 6.6(b), clearly showing the harder spectrum of p_T^{rel} for b-jets.



(a) p_T^{rel} variable

(b) p_T^{rel} templates for b-, c- and light-flavour-jets [135].

Figure 6.6: Sketch visualising the definition of the p_T^{rel} variable (a) and p_T^{rel} templates (b).

[37] One should note that the correction due to the neutrino only changes the transverse momentum of the jet and not the direction of the jet axis.

6.2.1.3 Selection of events

The presented measurement with the p_T^{rel}-method was done using a subset of the proton-proton collision data recorded at a centre-of-mass energy of $\sqrt{s} = 8$ TeV during 2012. The data sample corresponds to an integrated luminosity of $\mathcal{L} = 13.0\,\text{fb}^{-1}$. However, since the triggers used to select events are highly prescaled the amount of data usable for the analysis is much lower, corresponding to the values given in Table 6.1. This will be detailed in the following.

Trigger selection As explained in Section 6.2.1.2 the measurement of the b-jet tagging efficiency with p_T^{rel} is based on reconstructed jets that contain a muon within the jet cone from semileptonic decays. This experimental signature of jets being produced in particle collisions has a huge cross-section and is usually wanted to be suppressed in order to use the full trigger bandwidth for signatures of new physics or other rare processes. This means that usual triggers selecting jets are suffering from very large prescales and low selection efficiencies, $\varepsilon_{\text{prescale}}$, leading to only a limited amount of selected events for the final measurement[38].

To select a suitable amount of data a set of different triggers has been designed with each trigger being optimised in a different kinematic region of the jet p_T. Each trigger relies on a muon that is reconstructed from hits in the muon spectrometer with a threshold at $p_T^{\mu} > 4$ GeV. This muon has to be matched geometrically with a reconstructed calorimeter jet. The measurement is done in 9 different bins for which five different triggers have been designed. The respective jet p_T thresholds are listed in Table 6.1.

Table 6.1: Overview of the used muon-in-jet triggers to select a data sample usable for the p_T^{rel}-method.

kinematic region	jet threshold	muon threshold	$\mathcal{L} \cdot \varepsilon_{\text{prescale}}$
$20\text{ GeV} < p_T^{\text{jet}} < 40\text{ GeV}$	$p_T^{\text{jet,trig}} > 15\text{ GeV}$	$p_T^{\mu,trig} > 4\text{ GeV}$	$1.3\,\text{pb}^{-1}$
$40\text{ GeV} < p_T^{\text{jet}} < 50\text{ GeV}$	$p_T^{\text{jet,trig}} > 25\text{ GeV}$	$p_T^{\mu,trig} > 4\text{ GeV}$	$4.2\,\text{pb}^{-1}$
$50\text{ GeV} < p_T^{\text{jet}} < 75\text{ GeV}$	$p_T^{\text{jet,trig}} > 35\text{ GeV}$	$p_T^{\mu,trig} > 4\text{ GeV}$	$11.7\,\text{pb}^{-1}$
$75\text{ GeV} < p_T^{\text{jet}} < 110\text{ GeV}$	$p_T^{\text{jet,trig}} > 55\text{ GeV}$	$p_T^{\mu,trig} > 4\text{ GeV}$	$50.2\,\text{pb}^{-1}$
$110\text{ GeV} < p_T^{\text{jet}} < 200\text{ GeV}$	$p_T^{\text{jet,trig}} > 80\text{ GeV}$	$p_T^{\mu,trig} > 4\text{ GeV}$	$188.7\,\text{pb}^{-1}$

[38] This is the case especially for jets with high transverse momentum due to the exponentially decreasing jet p_T spectrum.

Different aspects which trigger is used in which jet p_T bin have to be considered:

- The trigger has to be fully efficient in the respective jet p_T bin to ensure that no bias due to different jet p_T spectra of b-, $c-$ and light-flavour-jets is introduced.
- To accumulate as much data as possible the trigger with the highest possible jet p_T threshold is chosen.
- The triggered jets are being reconstructed as in the usual offline reconstruction explained in Section 5.2.1 from topological clusters. They are calibrated at the hadronic scale, explicitly not including the momentum of the muon or the neutrino from semileptonic decays. Offline jets used in the later analysis are corrected for the muon as well as the neutrino component as explained in Section 6.2.1.1. This difference of about 10 % has to be taken into account when assigning the trigger to a jet p_T bin. E.g. the trigger with a threshold at $p_T^{\text{jet,trig}} > 25$ GeV will not yet be fully efficient at $p_T^{\text{jet}} \approx 30$ GeV and should not yet be used in the $30 - 40$ GeV jet p_T bin.

Each trigger is configured to have a rate of 1 Hz which is ensured by prescales being changed during data taking depending on the current run conditions. The production of high-p_T jets has a much lower cross-section than the production of lower-p_T jets. This means that the accumulated dataset in terms of integrated luminosity is much larger when using a higher jet p_T threshold, but the number of selected events is roughly the same due to the constant trigger rate[39].

Offline event selection To ensure a good data quality only data is used for which the inner detector, the muon system and the calorimeter all are fully operational. Only data collected during stable beam periods in which the silicon systems operated at full depletion voltage are used.

The p_T^{rel} method relies mainly on three types of objects: calorimeter jets, muons and tracks, where the latter two are required to be associated to the calorimeter jets as explained in Section 5.1. Jets are being reconstructed from topological clusters as it is described in Section 5.2.1 and are required to have $p_T > 20$ GeV and $|\eta| < 2.5$.

The usability of the method is limited to jets with p_T that is not too high. This is because high p_T jets are very collimated, and the muon track becomes almost collinear with the jet axis. Due to the finite resolution of the jet direction measurements the p_T^{rel} templates for b-, c- and light-flavour-jets all become dominated by resolution effects and

[39] The high prescales also cause that the selected datasets only have a very small statistical overlap.

start to look very similar. It is, therefore, not possible to distinguish a b-quark jet from a light-flavour-jet based on the p_T^{rel} of the muon and the method breaks down. Because of that only jets with $p_T < 200$ GeV are used.

The reconstructed muons are mainly required to fulfil the similar criteria as described in Section 5.3 apart from the trigger and isolation requirements. While in Section 5.3 high-p_T muons that are isolated with $p_T^{\mu} > 25$ GeV are reconstructed, in the p_T^{rel}-analysis one is explicitly interested in soft non-isolated muons. Therefore, the requirement on the transverse momentum is much lower and depending on the transverse momentum of the associated reconstructed jet. This dynamic muon cut is:

$$\begin{aligned} p_T^{\mu} &> 4 \text{ GeV} \qquad (\text{if } 20 \text{ GeV} < \mathrm{p_T^{jet}} < 60 \text{ GeV}), \\ p_T^{\mu} &> 6 \text{ GeV} \qquad (\text{if } 60 \text{ GeV} < \mathrm{p_T^{jet}} < 90 \text{ GeV}), \\ p_T^{\mu} &> 8 \text{ GeV} \qquad (\text{if } 90 \text{ GeV} < \mathrm{p_T^{jet}} < 200 \text{ GeV}). \end{aligned}$$

To increase the heavy-flavour fraction in the selected dataset a tag-and-probe ansatz is applied. This makes use of the fact that b-jets are very often produced in pairs. So in each event at least one jet has to be tagged by the simple SV0 tagger, see 6.1.3, using an operating point with a b-jet tagging efficiency of about $\varepsilon_b = 50\,\%$. In particular this means that at least one jet must contain a reconstructed secondary vertex and fulfil $\frac{l}{\sigma_l} > 1.0$. To not bias the sample towards larger $\frac{l}{\sigma_l}$ this jet (tag-jet) is discarded and only the other jets are used for the later analysis (probe jets)[40].

The explained selection is applied as well on collision data and simulated events to obtain the template shape of p_T^{rel} for b-jets and c-jets separately.

Selection of the light-flavour-template Contributions from light-flavour-jets are naturally suppressed by the usual event selection. The existence of a muon associated to the jet and the tag-and-probe requirement are preferring heavy-flavour jets. This means that a very large number of simulated events would be necessary in order to compensate for the low light-flavour-jet acceptance to obtain a light-flavour-jet template with decent statistics[41].

The low light-flavour-jet acceptance is bypassed by modifying the selection and estimating the light-flavour-template in a data-driven way. The tag-and-probe requirement is dropped and events from collision data are used if they do not contain a jet that has

[40] If more jets are fulfilling the tag criterion one of them is randomly dropped preferring jets that include a muon to retain the maximum number of jets with a muon for the measurement.

[41] In case of the tagged template the light-flavour-jet acceptance is even lower.

been tagged by the IP3D+SV1 tagging algorithm, see Section 6.1.2, at an efficiency of $\varepsilon_b = 80\,\%$.

The selected sample is dominated by light-flavour-jets with a contribution of about 20 % from c-jets. This does not bias the measurement since the p_T^{rel} is not sensitive to the difference between c-jets and light-flavour-jets as can be seen in Figure 6.6(b). The contamination of the sample with b-jets is $2 - 6\,\%$ depending on the jet p_T bin and is taken into account as a correction and a systematic uncertainty is applied.

6.2.1.4 Simulated dataset

The simulated samples that are used to measure the b-jet tagging efficiency are listed in Table 6.2. These samples have been produced using the event generation strategy that is described in Section 7.1. All samples contain dijet events generated by Pythia8 (v8.165) [136] using the parton density function from CT10 [59] together with the ATLAS AUET2 tune [137]. All b-decays are handled by a dedicated program with the newest implementation of b-hadron decay tables and decay modes called EvtGen [138]. The detector and trigger simulation of all samples is done using the full detector simulation that is also briefly described in Section 7.1.

Table 6.2: Simulated samples used for the p_T^{rel}-method.

Sample	# events	σ (nb)	Comment p_T in GeV
J0 dijet	2000000	$7.17 \cdot 10^7$	$0 < p_T^{\text{jet,truth}} < 20$
J1 dijet	9000000	$9.39 \cdot 10^3$	$20 < p_T^{\text{jet,truth}} < 80$
J2 dijet	3000000	$1.05 \cdot 10^2$	$80 < p_T^{\text{jet,truth}} < 200$
J3 dijet	2000000	$6.67 \cdot 10^{-1}$	$200 < p_T^{\text{jet,truth}} < 500$
J0 muon-filtered dijet	4000000	$2.78 \cdot 10^4$	$0 < p_T^{\text{jet,truth}} < 20$, $p_T^{\mu,\text{gen.}} > 3$
J1 muon-filtered dijet	8000000	$1.39 \cdot 10^2$	$20 < p_T^{\text{jet,truth}} < 80$, $p_T^{\mu,\text{gen.}} > 3$
J2 muon-filtered dijet	10000000	$4.00 \cdot 10^0$	$80 < p_T^{\text{jet,truth}} < 200$, $p_T^{\mu,\text{gen.}} > 3$
J3 muon-filtered dijet	4000000	$4.05 \cdot 10^{-2}$	$200 < p_T^{\text{jet,truth}} < 500$, $p_T^{\mu,\text{gen}} > 3$

The J0-J3 muon-filtered dijet samples are created such that each individual sample covers a different range in truth jet p_T of the leading jet. The samples are summed according to their relative cross-section σ to form an inclusive sample. This is referred to as the JXμ dijet sample.

Since the p_T^{rel}-method uses only jets that contain a muon, these samples have been enriched with events containing muons by applying a filter at generator level. Only events that contain at least one muon with $p_T > 3$ GeV during event generation are saved and

all others are directly discarded. However, the JXμ dijet sample contains muons from b- and c-decays, but has too few muons from in-flight decays[42]. Therefore, the fraction of light-flavour-jets with pions or kaons decaying in-flight is underestimated in the JXμ dijet sample. Although the analysis is targeting only to measure the fraction of b-jets at some points also the correct fraction of light-flavour-jets is needed[43]. This is why the JX dijet sample was created. It has the same configuration as the JXμ dijet sample but without the muon generator filter applied.

6.2.1.5 p_T^{rel} measurement

The b-jet tagging efficiency is defined as the fraction of b-jets that are being tagged by an algorithm. This means that this quantity can be calculated from the number of tagged b-jets and the number of b-jets that have not been tagged by the algorithm. So the selected dataset is split up into two disjoint datasets called the tagged and the untagged sample according to the tagging requirement under study.

As the p_T^{rel} variable is sensitive to the amount of b-jets in any sample the fractions of b-jets in the two samples, $f_{\text{tagged}}^{b-\text{jets}}$ and $f_{\text{untagged}}^{b-\text{jets}}$, can be estimated using a binned template likelihood fit. In this fitting technique each bin of the three different templates, estimated as explained in Section 6.2.1.3, is treated as an independent Poisson variable [139]. With the total number of jets in the tagged sample, $N_{\text{tagged}}^{\text{all jets}}$, and the untagged sample, $N_{\text{untagged}}^{\text{all jets}}$, the efficiency ε_b can be calculated as:

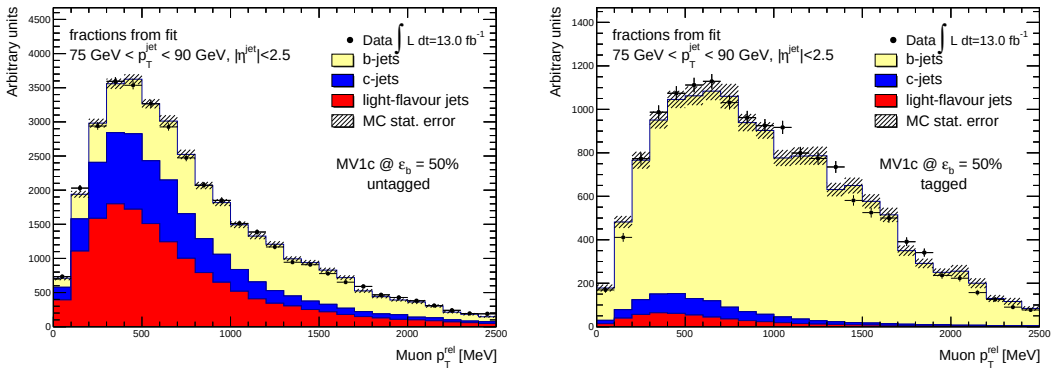
$$\varepsilon_b = \frac{f_{\text{tagged}}^{b-\text{jets}} \cdot N_{\text{tagged}}^{\text{all jets}}}{f_{\text{tagged}}^{b-\text{jets}} \cdot N_{\text{tagged}}^{\text{all jets}} + f_{\text{untagged}}^{b-\text{jets}} \cdot N_{\text{untagged}}^{\text{all jets}}}. \tag{6.5}$$

In Figure 6.7 the fit result is shown for the MV1c tagging algorithm at an operating point of $\varepsilon_b^{t\bar{t}} = 50\%$ for jets with a transverse momentum of 75 GeV $< p_T <$ 90 GeV. Shown are the selected untagged (a) and tagged (b) jets in data (solid black points) and the b-, c- and light-flavour-templates scaled to the respective fit result. Since the p_T^{rel} variable is only sensitive to the amount of b-jets and does not separate well between c-jets and light-flavour-jets only two templates are used in the fit. These are the b-template itself and a combined (c+light)-template. The light-to-c ratio is constrained to the expectation obtained in simulation using the JX dijet sample.

[42] Pions and kaons are treated as stable particles on generator level. The decay of these particles is added at the level of detector simulation.

[43] For instance, the correct fraction is needed when estimating the ratio of c-jets to light-flavour-jets, explained in Section 6.2.1.5.

The fitting procedure is repeated for every algorithm and operating point in the same way. Since the b-tagging performance strongly depends on the jet momentum, p_T^{jet}, and pseudo-rapidity, η^{jet}, the fits are performed in bins of these two quantities.



(a) Untagged sample fitted with p_T^{rel}

(b) Tagged sample fitted with p_T^{rel}

Figure 6.7: p_T^{rel} fit to the untagged (a) and tagged (b) sample using the tagging requirements of the MV1c algorithm at an operating point of $\varepsilon_b^{t\bar{t}} = 50\%$ for jets with a transverse momentum of 75 GeV $< p_T <$ 90 GeV.

Two different corrections are applied to the scale factor κ_{ε_b}. Those two corrections arise from the jet direction resolution that was found to be lower in simulation than in data [140]. Secondly, the contamination of b-jets within the light-flavour-template, see Section 6.2.1.3 is corrected for.

Correction due to the jet direction resolution The direction of the jet has a direct influence on the p_T^{rel} variable. Therefore, a poor resolution of the jet direction would cause the b-, c- and light-flavour-template to look more alike which has a direct influence on the fit result.

To study any possible influence an independent jet axis was formed from the p_T of all tracks that were associated to the jet. This track-jet based axis was compared to the axis of the usual calorimeter jet in both data and simulation with an uncertainty of 6 mrad in η and 4 mrad in φ [140]. According to the uncertainty the jet direction is smeared and the smeared and unsmeared scenario are compared. The difference might be caused by a poor modelling of tracks in jets[44] and not by the jet direction resolution.

[44] The intention of the measurement is exactly to correct these kind of deviations between data and simulation with the scale factors κ_{ε_b}.

Thus, half the difference is applied as the correction. A relative uncertainty of 100% on the correction is taken as a systematic uncertainty on κ_{ε_b}.

Correction due to the b-jet contamination in the light-flavour-template As explained in Section 6.2.1.3 the light-flavour-template is estimated in a data-driven way and a contamination with b-jets as well as c-jets is expected. Both is obtained in simulation and varied by 100%. The difference between the nominal measured $\varepsilon_b^{\text{data,nom}}$ and $\varepsilon_b^{\text{data,var}}$ including the variation is very small. However, it is applied as a correction on the scale factor, κ_{ε_b}, and is used as a systematic uncertainty.

6.2.1.6 Systematic uncertainties

A variety of systematic uncertainties have been considered and tested for a possible bias on the measurement of ε_b. They are summarised in Table 6.3 for the MV1c algorithm at the operating point of $\varepsilon_b^{t\bar{t}} = 50\%$. An analogue table is created for every tagging algorithm and operating point leading to a large number of systematic uncertainties. To obtain the total systematic uncertainty the individual systematic uncertainties in each jet p_T bin are added in quadrature.

The individual systematic uncertainties are explained in the following. If not stated otherwise they are calculated by repeating the fit to the same dataset with modified templates including the systematic variation. If possible an upwards and downwards variation by $\pm 1\sigma$ is done and half the difference is taken as the systematic according to Equation (6.6):

$$\Delta\varepsilon_b = 0.5 \cdot \left(\varepsilon_b^{+1\sigma} - \varepsilon_b^{-1\sigma}\right). \tag{6.6}$$

For a single variation the value of ε_b is compared directly to the nominal fit result according to Equation (6.7):

$$\Delta\varepsilon_b = \varepsilon_b^{1\sigma} - \varepsilon_b^{\text{nominal}}. \tag{6.7}$$

The uncertainty on the estimated efficiency is then propagated yielding the uncertainty of the scale factor given in Table 6.3. In case of systematic uncertainties related to the reconstruction of energies[45] also the efficiency in simulation is recalculated including the systematic variation.

[45] This means the systematic uncertainties due to the jet energy scale and resolution and the semileptonic correction.

Table 6.3: Table of the systematic uncertainties for the MV1c tagging algorithm at $\varepsilon_b^{t\bar{t}} = 50\%$ offline b-tagging efficiency.

MV1c @ $\varepsilon_b^{t\bar{t}} = 50\%$	p_T^{jet} [GeV]								
Source of uncertainty	20-30	30-40	40-50	50-60	60-75	75-90	90-110	110-140	140-200
	Systematic uncertainty [%]								
modelling of $g \to b\bar{b}$	0.1	0.4	0.0	0.0	0.1	0.0	0.4	0.3	2.1
modelling of $g \to c\bar{c}$	0.4	0.8	1.3	2.0	2.6	4.6	5.7	11.5	10.5
b-quark fragmentation fraction	0.3	0.7	0.1	0.0	0.4	0.2	0.3	0.0	0.7
b-quark fragmentation function	0.0	0.3	0.2	0.1	0.1	0.0	0.4	0.5	1.3
b-decay branching fractions	0.0	0.1	0.0	0.0	0.1	0.1	0.4	0.3	1.1
b-decay p^* spectrum	1.9	0.9	0.7	1.4	0.7	0.5	2.9	1.9	3.0
fake muons in b-jets	0.0	0.5	0.1	0.1	0.3	0.1	0.1	0.3	1.4
jet direction resolution	0.2	0.3	0.5	0.2	0.2	0.6	0.6	1.6	2.4
jet energy resolution	1.6	0.4	0.5	0.5	0.2	0.3	0.3	0.0	1.1
jet energy scale	0.9	1.1	0.6	0.6	0.5	0.4	0.0	0.9	0.8
semileptonic correction	0.3	0.4	0.2	0.3	0.4	0.1	0.0	0.3	0.2
muon p_T spectrum	0.9	2.0	1.7	1.8	2.4	0.7	4.2	4.5	7.2
pile-up $\langle\mu\rangle$ reweighing	0.5	0.6	0.0	0.4	0.2	0.1	0.3	0.5	0.3
light-template contamination	0.5	0.4	0.4	0.5	0.5	0.4	0.2	0.3	0.1
charm-light ratio	1.9	1.7	1.1	0.4	0.0	0.5	0.0	2.4	5.3
scale factor for inclusive b-jets	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
simulation tagging efficiency	1.3	0.3	0.3	0.4	0.4	0.4	0.3	0.2	0.2
template statistics	2.5	0.6	0.5	0.5	0.7	0.7	0.9	1.1	1.8
total systematic uncertainty	6.2	5.4	5.0	5.4	5.7	6.4	8.8	13.5	15.0
statistical uncertainty	2.4	2.1	1.5	1.4	1.5	1.5	1.5	1.8	2.3

Modelling of $g \to b\bar{b}$ and $g \to c\bar{c}$ It is possible that a b-jet is actually a gluon jet with the gluon decaying in a $b\bar{b}$-pair where both quarks end up within the same jet. Since jets with two b-quarks have a higher probability to be tagged this would have a direct influence on the b-jet tagging efficiency. The systematic uncertainty due to these double-b-jets is calculated by varying the ratio of double-b-jets to single-b-jets by 100%. In practice this is done by counting the number of true b-quarks within $\Delta R < 0.4$ around the jet+muon axis and giving each truth-labelled b-jet with more than one b-quark a weight of 0 or 2. Half the difference between the two variations is assigned the systematic uncertainty.

The same procedure is done for truth-labelled c-jets where more than one c-quark within $\Delta R < 0.4$ around the jet+muon axis is found.

b-quark fragmentation Two different sources of systematic uncertainties are being considered for the b-quark fragmentation.

X_b corresponds to the fraction of b-quark energy that is carried by the b-hadron. The resulting scale factor obtained with templates where the b-quark fragmentation function was reweighted such that the average X_b was changed up and down by 5%. Half the difference in ε_b is referred to as the systematic of the b-quark fragmentation function.

Secondly, the production fractions of different b-hadrons have been measured at LEP as well as the Tevatron showing a $2\,\sigma$ disagreement [9, 141]. Events are being reweighted so that the distribution of the different hadrons matches with the Tevatron results. The difference to the nominal sample which is more compatible with the measurement from LEP is the systematic uncertainty called b-quark fragmentation fraction.

b-hadron decay Two sources of systematic uncertainties are considered for the decay of the b-hadron to a muon. The spectrum of the muon momentum in the rest frame of the b-hadron is called p^* and two different kind of decays have to be considered.

The b-hadron can either decay directly into a muon via $b \to \mu + X$ or as a cascade decay with an intermediate c-quark via $b \to c/\overline{c} \to \mu + X$. The branching fractions $BF(b \to lX) = (10.69 \pm 0.22)\%$ as well as $BF(b \to c/\overline{c} \to l + X) = (9.62 \pm 0.53)\%$ with l being either an electron or a muon and the p^* spectrum have been measured [9, 142].

For the systematic uncertainty called b-decay branching fractions the ratio of $BF(b \to lX)/BF(b \to c/\overline{c} \to l + X)$ was varied by one standard deviation. For the b-decay p^* spectrum systematic uncertainty the p^* spectrum was reweighting such that it matches the measurement.

Fake muons in b-jets If a reconstructed muon does not match with a truth muon in simulation it is defined as a fake muon[46]. To estimate the influence of these fake muons on the measurement the amount of fake muons was doubled. It is done by giving the corresponding jet a weight of 2 if a fake muon is found. The fit result with the variation is compared to the nominal fit result and the full difference is called the systematic uncertainty due to fake muons.

[46] To be more precise, only about half of the muons identified as fakes in this way are true fake muons. A second contribution is coming from decay-in-flight muons since e.g. kaons or pions decaying within the calorimeter are treated as stable particles. In those cases the track is not labelled as a muon track on truth level.

Jet direction resolution The correction applied to compensate for a different jet direction resolution between data and simulation was explained in Section 6.2.1.5 and the full size of the correction is taken as the systematic uncertainty on the scale factor.

Jet energy resolution The energy of each jet in simulation is smeared by a Gaussian function such that the width of the resulting Gaussian distribution corresponds to the one including the uncertainty on the jet energy resolution [143]. The difference between the smeared and unsmeared fit result is taken as the systematic uncertainty according to Equation (6.6).

Jet energy scale The p_T of each jet in simulation was varied up and down by $\pm 1\,\sigma$ according to an inclusive uncertainty on the jet energy scale. The techniques used to measure the uncertainty are the same as they are described for 7 TeV data in Ref. [113].

Semileptonic correction A correction is applied to include the contribution from the reconstructed muon and the neutrino matched to the reconstructed jet. The semileptonic correction, explained in Section 6.2.1.1, has various systematic uncertainties which are partly correlated with the systematic uncertainties applied in the p_T^{rel}-method [113]. In these cases the uncertainties on the semileptonic correction are varied at the same time as the p_T^{rel} uncertainties. All other components are summarised in one inclusive up and down variation of the p_T of all jets in simulation. Half the difference between this upwards and downwards variation is taken as the systematic uncertainty due to the semileptonic correction.

Muon p_T spectrum Slight deviations in the muon p_T spectrum are observed between simulation and collision data. The spectrum in simulation is reweighted to match the distribution in data and the obtained result compared to the nominal result is assigned a systematic uncertainty.

Pile-up $\langle\mu\rangle$ reweighing The distribution of the average interactions per bunch-cross, called $\langle\mu\rangle$, is shown in Figure 5.1 for the data collected during 2012. A good description of the minimum bias vertex multiplicity was found in simulation when scaling $\langle\mu\rangle$ by 1.11 ± 0.088. The uncertainty was estimated by checking the level of agreement in other variables sensitive to pile-up like the number of primary vertices [144]. Afterwards in simulation the distribution is reweighted to agree with the data distribution. The differ-

ence, when varying the scaling of $\langle\mu\rangle$ by one standard deviation, in the scale factor is the systematic uncertainty.

p_T^{rel} light-flavour-template contamination As explained in Section 6.2.1.5 a correction is applied because of the heavy-flavour contamination in the p_T^{rel} light-flavour-template. The full correction is also assigned as systematic uncertainty.

Charm-light ratio As explained in Section 6.2.1.5 the ratio of c-jets to light-flavour-jets is constrained to the prediction from simulation. The combined (c+light)-template is varied twice by doubling and halving the amount of light-flavour-jets in the sample. The difference between the two fit results is the systematic uncertainty calculated following Equation (6.6).

Scale factor for inclusive b-jets One caveat of the p_T^{rel}-method is that only jets that contain a soft muon contribute to the measurement of ε_b. Since semileptonic jets will always contain a well-measured muon track it is expected that ε_b is higher in a sample only containing those jets than in an inclusive sample. However, the calibration results are given in the form of a data-to-simulation scale factor κ_{ε_b} that is applied to the inclusive sample. As long as the relative difference between semileptonic and inclusive jets is modelled well in simulation the difference should cancel.

The ratio of the scale factor $\frac{\kappa_b^{\mathrm{incl}}}{\kappa_b^{\mathrm{semi}}}$ for jets with muons and all jets was measured separately in a $t\bar{t}$ dilepton sample [145] and was found to be consistent with one with an uncertainty of 4%. This uncertainty is assigned as a constant uncertainty for all tagging algorithms, operating points and kinematic bins. This uncertainty currently is the limiting factor of the overall precision of the method.

Simulation tagging efficiency The statistical uncertainty on the value of $\varepsilon_b^{\mathrm{sim}}$ is shown separately and called the simulation tagging efficiency.

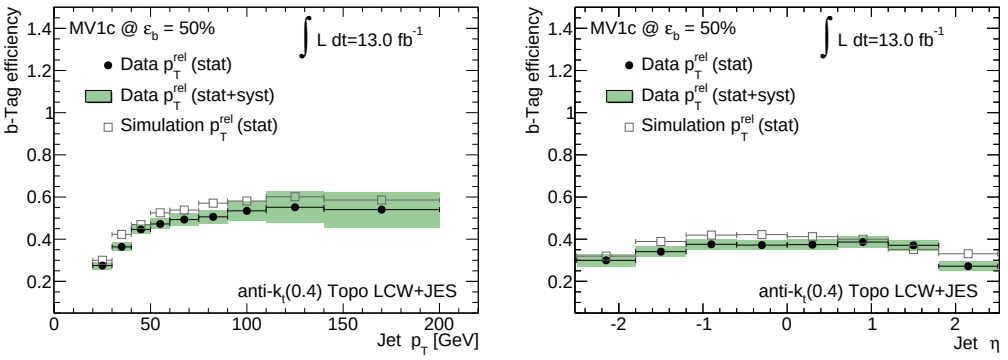
Template statistics The influence of the limited template statistics was checked by using pseudo-experiments. To create a pseudo-dataset a fluctuation was added to each bin of the b-, c- and light-flavour-template according to their statistical uncertainty. This was done by letting each bin content vary around its central value according to a Gaussian distribution with the width set to the statistical uncertainty in each bin. The fit is repeated 1000 times yielding 1000 values of ε_b that are filled into a histogram. The RMS of this histogram is taken as the systematic uncertainty.

Two additional effects are covered at once with this procedure. Firstly, to cover a possible systematic bias due to the usage of the data-driven light-flavour-template the statistical uncertainty in case of the light-flavour-jets is taken from the JX simulation sample.

Secondly, the templates used as well for the tagged and the untagged fit are not using any b-tagging requirement. This is justified by the fact that the p_T^{rel} variable is uncorrelated to the b-tagging weights and as well the tagged, untagged and pretagged distributions of p_T^{rel} agree within their statistical uncertainty for the three different flavours. To cover a possible difference the statistical uncertainty used to create the pseudo-datasets is based on the tagged and untagged templates instead of the pretagged ones used in the nominal fit.

6.2.1.7 Calibration results

The estimated efficiencies for the MV1c algorithm at an operating point of $\varepsilon_b^{t\bar{t}} = 50\%$ are shown in Figure 6.8 in bins of p_T^{jet} (a) and η^{jet} (b). Also κ_{ε_b} is shown in in Figure 6.9 in bins of p_T^{jet} (a) and η^{jet} (b). The green band is showing the total uncertainty which is defined as the quadratic sum of all systematic uncertainties and the statistical uncertainty.

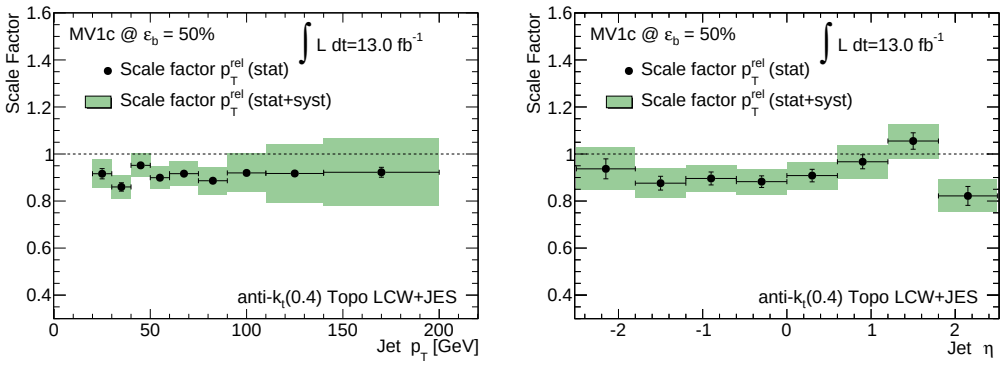


(a) ε_b measured in bins of p_T^{jet} with p_T^{rel}

(b) ε_b measured in bins of η^{jet} with p_T^{rel}

Figure 6.8: The b-jet tagging efficiency ε_b in data and simulation measured with the p_T^{rel}-method for the MV1c tagging algorithm at the operating point of $\varepsilon_b^{t\bar{t}} = 50\%$.

The efficiency in data has been measured using the explained fit and the efficiency in simulation was calculated directly from the number of tagged and untagged jets that are labelled as a b-jet. The efficiency measured in data is about 10% lower than the efficiency in simulation depending on the kinematic bin. This can be corrected using the obtained data-to-simulation scale factors κ_{ε_b}. Several reasons could cause the difference between



(a) κ_{ε_b} measured in bins of p_T^{jet} with p_T^{rel}

(b) κ_{ε_b} measured in bins of η^{jet} with p_T^{rel}

Figure 6.9: The scale factor κ_{ε_b} in data and simulation measured with the p_T^{rel}-method for the MV1c tagging algorithm at the operating point of $\varepsilon_b^{t\bar{t}} = 50\%$.

data and simulation like general differences in the tracking and vertexing performance that are described e.g. in Ref. [128]. In general one can say that κ_{ε_b} is closer to unity if the tagging requirement is loosened which corresponds to using an operating point with a higher tagging efficiency but lower purity.

6.2.2 Measuring the b-jet tagging efficiency with System8

The system8-method is using the same semileptonic jets with a non-isolated soft muon within the jet cone as the p_T^{rel}-method. The sample is divided into 8 disjoint subsamples according to three different criteria:

- The b-jet efficiency operating point under study
- A soft-muon tagging requirement implemented as a cut on the p_T^{rel} variable itself
- A second jet in the event was tagged by another tagging algorithm

The last criterion is made such that it corresponds exactly to the tag-and-probe criterion

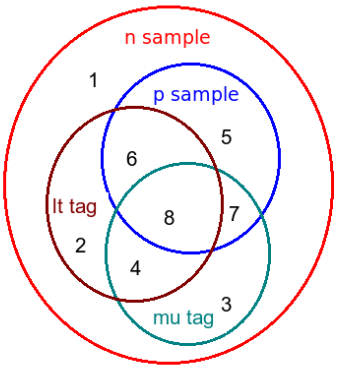


Figure 6.10: Venn diagram showing the relations between different subsamples used in system8 [146].

used in the p_T^{rel} sample. The selected subsample after this requirement is called the p-sample and is the same as the p_T^{rel} pretagged sample. A set of eight equations can be written down that is correlating the different flavours within the subsamples and the tagging efficiencies:

$$\begin{array}{lcrcr}
n & = & n_b & + & n_{cl} \\
p & = & p_b & + & p_{cl} \\
n^{\mathrm{LT}} & = & \epsilon_b^{\mathrm{LT}} n_b & + & \epsilon_{cl}^{\mathrm{LT}} n_{cl} \\
p^{\mathrm{LT}} & = & \alpha_6 \epsilon_b^{\mathrm{LT}} p_b & + & \alpha_4 \epsilon_{cl}^{\mathrm{LT}} p_{cl} \\
n^{\mathrm{MT}} & = & \epsilon_b^{\mathrm{MT}} n_b & + & \epsilon_{cl}^{\mathrm{MT}} n_{cl} \\
p^{\mathrm{MT}} & = & \alpha_5 \epsilon_b^{\mathrm{MT}} p_b & + & \alpha_3 \epsilon_{cl}^{\mathrm{MT}} p_{cl} \\
n^{\mathrm{LT,MT}} & = & \alpha_1 \epsilon_b^{\mathrm{LT}} \epsilon_b^{\mathrm{MT}} n_b & + & \alpha_2 \epsilon_{cl}^{\mathrm{LT}} \epsilon_{cl}^{\mathrm{MT}} n_{cl} \\
p^{\mathrm{LT,MT}} & = & \alpha_7 \alpha_6 \alpha_5 \epsilon_b^{\mathrm{LT}} \epsilon_b^{\mathrm{MT}} p_b & + & \alpha_8 \alpha_4 \alpha_3 \epsilon_{cl}^{\mathrm{LT}} \epsilon_{cl}^{\mathrm{MT}} p_{cl}
\end{array} \tag{6.8}$$

Here, LT stands for "life-time tagging" requirement meaning the criterion of the operating point under study. MT stands for "muon tagging" requirement meaning the cut on the value of p_T^{rel}. The eight values of α_i stand for the correlation between the different subsamples. The set of equations can be solved numerically yielding a measurement of the efficiency. This method is also described in detail in Ref. [132, 146].

6.2.3 Measuring the b-jet tagging efficiency in $t\bar{t}$ events

Since the top quark decays almost exclusively to b-quarks, see Section 3.2, $t\bar{t}$ events can provide a sample of almost pure b-jets well suited for the calibration. Here, the semileptonic and the dilepton $t\bar{t}$ decay channel can be used that are statistically uncorrelated. Both have the advantage that ε_b is measured in an inclusive jet sample instead of the sample using only semileptonic jets as in the p_T^{rel}- or system8-method.

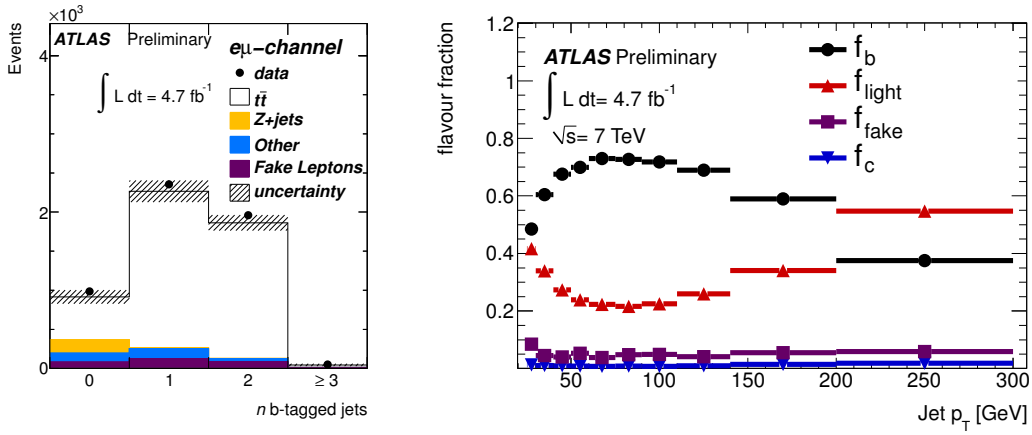
Four different approaches have been used for calibrations in ATLAS all exploiting the favourable flavour fraction in $t\bar{t}$ events. They will shortly be explained in the following, and more details can be found in in Ref. [145, 147].

Tag-counting method In the tag-counting method the number of tagged jets is fitted in $t\bar{t}$ candidate events. Two b-jets are expected as well in the semileptonic and the dilepton final state and the number of events with one or two b-tagged jets would then be directly related to ε_b via $N_{2tags} = \varepsilon_b^2 N_{t\bar{t}}$ and $N_{1tag} = \varepsilon_b \cdot (1 - \varepsilon_b) 2 N_{t\bar{t}}$.

However, if e.g. one of the expected b-jets lies outside of the detector acceptance or additional b-jets are produced due to gluon radiation and $g \to b\bar{b}$ decays the number

of expected b-jets can be different. In addition, tagged c- or light-flavour-jets contribute and have to be taken into account.

All these effects are included by fitting the expected fractions, F_{ijk}, of events with i b-jets, j c-jets and k light-flavour-jets estimated in simulation. All contributions of F_{ijk} can be summed up to estimate the number of expected events with n b-tagged jets that is again directly related to ε_b. In Figure 6.11(a) the number of b-tagged jets is shown in the $e + \mu$ dilepton $t\bar{t}$ decay channel. It can be seen that a large fraction of $t\bar{t}$ signal events either have one or two b-tagged jets.



(a) Number of b-tagged jets

(b) Jet flavour composition in dilepton $t\bar{t}$ events

Figure 6.11: (a) Number of b-tagged jets in the $e+\mu$-channel and (b) expected jet flavour composition of the two leading jets in a selected dilepton $t\bar{t}$ sample as a function of jet p_T [145].

Kinematic selection method In the kinematic selection method the b-tagging rate of the leading jet is measured. This fraction is directly related to the b-jet tagging efficiency by the equation:

$$\begin{aligned} f_{\mathrm{b-tag}} &= \varepsilon_b f_{\mathrm{b-jets}} + \varepsilon_c f_{\mathrm{c-jets}} + \varepsilon_l f_{\mathrm{light-flavour-jets}} + \varepsilon_{\mathrm{fake}} f_{\mathrm{fake}} \\ \Leftrightarrow \varepsilon_b &= \frac{1}{f_{\mathrm{b-jets}}} \cdot \left(f_{\mathrm{b-tag}} - \varepsilon_c f_{\mathrm{c-jets}} - \varepsilon_l f_{\mathrm{light-flavour-jets}} - \varepsilon_{\mathrm{fake}} f_{\mathrm{fake}}\right). \end{aligned} \tag{6.9}$$

Other contributions from c-jets and light-flavour-jets as well as fake lepton contributions in the dilepton channel and QCD-multijet production in the semileptonic channel have to be taken into account. The contributions can be taken from simulation using the

correction factor estimated e.g. with the p_T^{rel}-method explained in Section 6.2.1 or they are estimated from collision data.

In Figure 6.11(b) the estimated flavour composition of the two leading jets in a dilepton $t\bar{t}$ sample are displayed with a high fraction of b-jets in a wide range of jet p_T.

Kinematic fit method In this method a kinematic fit is used to obtain a sample that is highly purified with b-jets using the $t\bar{t}$ event topology. The fit is only done in the semileptonic decay channel and provides a mapping of the reconstructed jets and leptons to the decay products of the $t\bar{t}$ decay. This yields the information which jet is either a b-jet directly from the decay of a top-quark or a jet stemming from the hadronic decay of the $W^{\pm}$-boson produced in the top decay.

Of course the correct assignment is not made in all cases and also background processes have to be taken into account. This is done by using a statistical background subtraction based on two subsamples. The signal sample is constructed in a way that it contains a high fraction of correct assignments while the background sample contains a large fraction of incorrect assignments. Using a truth matching in both samples the background contribution can be estimated in the background sample. This is extrapolated and subtracted in the signal sample.

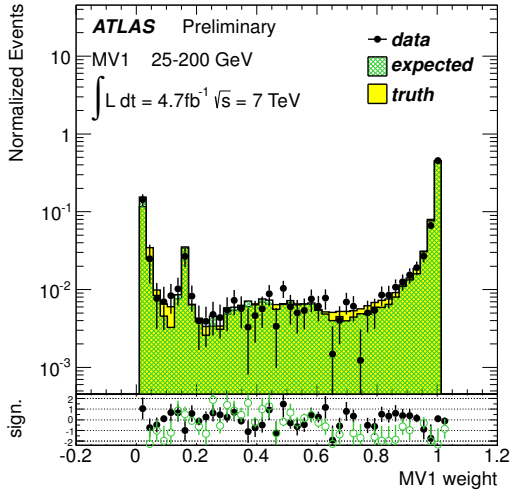
The b-jet tagging efficiency is extracted using the jet that has been assigned as the b-jet on the leptonic side of the $t\bar{t}$ decay. Its background-subtracted b-tagging weight distribution is fully reconstructed allowing not only to obtain the efficiency but also a continuous calibration of this distribution. In Figure 6.12(a) the background-subtracted b-tagging weight distribution for the MV1 tagging algorithm is shown. Figure 6.12(b) shows the obtained efficiency for the MV1 algorithm in data together with the expected efficiency using simulated data and the true distribution from true b-jets in simulation.

Combinatorial likelihood approach (PDF method) In general the information whether a second jet is tagged or not can be used to extend Equation (6.9) as follows:

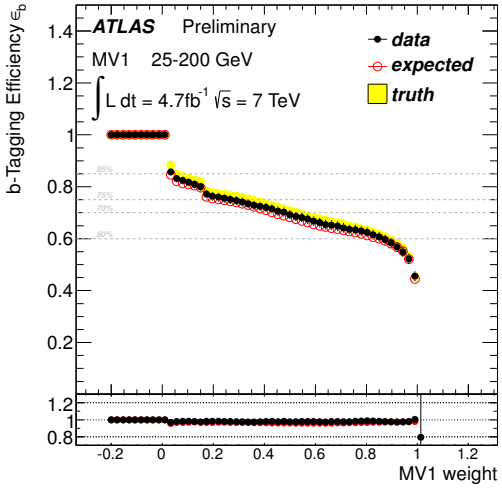
$$f_{2\ tags} = f_{bb}\varepsilon_b^2 + f_{bl}\varepsilon_l\varepsilon_b + (1 - f_{bb} - f_{bl})\varepsilon_l^2, \tag{6.10}$$

$$f_{1\ tag} = 2f_{bb}\varepsilon_b(1-\varepsilon_b) + f_{bl}[\varepsilon_l(1-\varepsilon_b) + (1-\varepsilon_l)\varepsilon_b] + (1 - f_{bb} - f_{bl})2\varepsilon_l(1-\varepsilon_l). \tag{6.11}$$

A problem arises if a binning in any kinematic variable like the jet p_T is needed due to possible correlations. For N kinematic bins N^2 combinations for two jets would be possible. Thus, a complex system of $2N^2$ non-linear equations had to be solved.



(a) b-tagging weight distribution for the MV1 tagging algorithm



(b) Efficiency for the MV1 tagging algorithm

Figure 6.12: Background-subtracted b-tagging weight distribution for the MV1 tagging algorithm (a) and obtained efficiency for the MV1 tagging algorithm (b). Both are extracted using the jet assigned as the b-jet on the leptonic side by the kinematic fit in semileptonic $t\bar{t}$ events [145].

An alternative, explained in detail in Ref. [147], is to model the system using a powerful likelihood function $\mathcal{L}$ and solving it by minimising $\mathcal{L}$ explicitly including the correlations:

$$\begin{aligned}\mathcal{L}(p_{T,1}, p_{T,2}, w_1, w_2) = [&f_{bb} \cdot \mathrm{PDF}_{bb}(p_{T,1}, p_{T,2}) \cdot \mathrm{PDF}_b(w_1|p_{T,1}) \cdot \mathrm{PDF}_b(w_2|p_{T,2}) \\ &+ f_{b\ell} \cdot \mathrm{PDF}_{b\ell}(p_{T,1}, p_{T,2}) \cdot \mathrm{PDF}_b(w_1|p_{T,1}) \cdot \mathrm{PDF}_\ell(w_2|p_{T,2}) \\ &+ f_{\ell\ell} \cdot \mathrm{PDF}_{\ell\ell}(p_{T,1}, p_{T,2}) \cdot \mathrm{PDF}_\ell(w_1|p_{T,1}) \cdot \mathrm{PDF}_\ell(w_2|p_{T,2}) \\ &+ 1 \leftrightarrow 2]/2.\end{aligned}$$

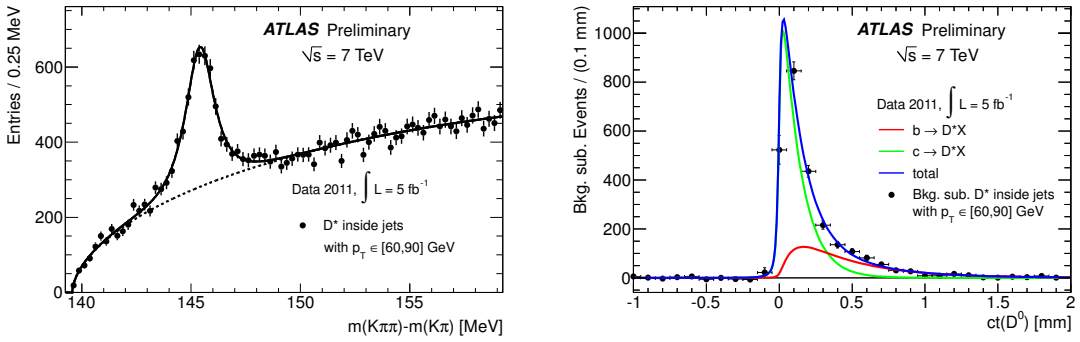
Firstly, the two-dimensional probability density functions $\mathrm{PDF}_{f_1 f_2}(p_{T,1}, p_{T,2})$ within the interval $[p_{T,1}, p_{T,2}]$ for the flavour combination $[f_1, f_2]$ have to be estimated in simulation. Secondly, in $\mathcal{L}$ the PDFs for the b-tagging weight for a jet of flavour f, dependent on the jet p_T, are included as $\mathrm{PDF}_f(w|p)$. Those are being used to finally extract the efficiency depending on the jet p_T or any other kinematic variable in an analogue way.

The method using the combinatorial likelihood approach applied to dileptonic $t\bar{t}$ events is currently providing the most precise calibration of ε_b in ATLAS. It reaches a

total uncertainty of about 2 % for jets with transverse momenta around 100 GeV and provides a calibration of ε_b for jets with transverse momenta in the range $20 - 300$ GeV.

6.2.4 Measuring the c-jet tagging efficiency with D^* mesons

The c-jet tagging efficiency can be measured using jets associated with D^* mesons that decayed via the decay chain $D^{*+} \to D^0(K^-\pi^+)\pi^+$. A sample highly enriched with c-jets is created by fitting the $\Delta m = m(K^-\pi^+\pi^+) - m(K^-\pi^+)$ distribution built out of the three reconstructed charged particle tracks. The distribution is shown in Figure 6.13(a) leading to the wanted sample by applying a background-subtraction technique, that will be explained in the following.



(a) Δm distribution

(b) Fitted D^0 pseudo proper time distribution

Figure 6.13: Δm distribution of D^{*+} candidates (a) and fitted D^0 pseudo proper time distribution on a background-subtracted D^{*+} sample (b) for jets with a transverse momentum between $60 - 90$ GeV [148].

The signal region is defined as the region within $3\,\sigma$ of the fitted Gaussian Δm peak centre and the background region is defined by requiring $\Delta m > 150$ MeV. To obtain a background-subtracted variable its distribution of events from the background region, normalised to the fitted background fraction in the signal region, is subtracted from the data distribution in the signal region.

In this way the pseudo proper time of the D^0 candidate, which is defined in Ref. [148, 149], can be obtained in the background-subtracted sample. This quantity is discriminating between b-jets and c-jets and is fitted to extract the flavour composition in the background-subtracted sample. The fit result is shown in Figure 6.13(b).

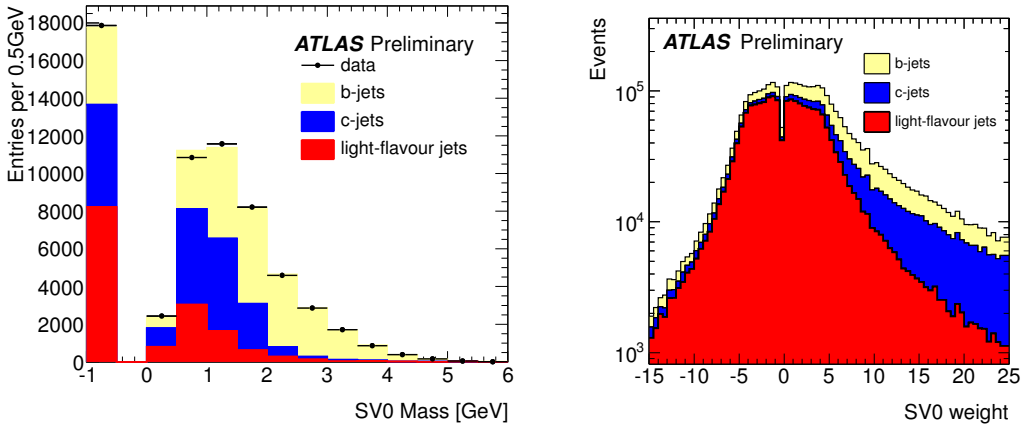
The c-jet tagging efficiency is extracted by a combined fit to the Δm distribution before and after applying the requirement of the b-tagging operating point under study.

With this information and the information about the b-jet contamination from the pseudo proper time distribution fit, ε_c can be obtained. The measurement is explained in detail in Ref. [148, 149].

6.2.5 Measuring the mistag rate

Two complementary approaches are used to extract the mistag rate which corresponds to the light-flavour-jet tagging efficiency ε_l. Both methods are using an inclusive jet sample and are described in more detail in Ref. [149, 150].

SV0 mass fits A variable that is discriminating between light-flavour-jets and heavy-flavour jets is the SV0 mass already introduced in Section 5.1. To extract ε_l templates for the different flavours of the SV0 mass are estimated in simulation. They are fitted to the data distribution before and after the tagging requirement is applied. This yields the number of all light-flavour-jets, N_l, in the inclusive jet sample and the number of light-flavour-jets, N_l^{tag}, that have been selected by the tagging algorithm at the operating point under study. The mistag rate is then: $\varepsilon_l = \frac{N_l^{\text{tag}}}{N_l}$. A fit to the SV0 mass distribution in the tagged sample using the b-tagging algorithm MV1 at an operating point of $\varepsilon_b^{t\bar{t}} = 70\%$ is shown as an example in Figure 6.14(a).



(a) SV0 mass template fit

(b) SV0 tag weight distribution

Figure 6.14: (a) SV0 mass template fit to the jet sample tagged by the MV1 algorithm at an operating point of $\varepsilon_b^{t\bar{t}} = 70\%$. The first bin includes jets without a reconstructed secondary vertex. (b) SV0 tag weight distribution for b-, c- and light-flavour-jets in simulation [150].

Negative tags The impact parameter and the decay length distributions, that are the key ingredients of every b-tagging algorithm introduced in Section 6.1, are expected to be symmetric for light-flavour-jets. This is displayed in Figure 6.14(b) for the signed decay length significance. This means that by counting the number of jets that are negatively tagged and comparing to the total number of jets within the sample, it is possible to measure $\varepsilon_l^{\text{neg}}$.

To obtain the correct mistag rate two correction factors, k_{hf} and k_{ll}, have to be taken from simulation. k_{hf} is accounting for the contamination of heavy-flavour jets on the negative side of the b-tagging weight distribution. k_{ll} is used to correct for contributions from long-lived particles like K_s^0. Those contributions introduce an asymmetry in the b-tagging weight distribution for light-flavour-jets towards the positive side[47].

[47]The reason is that decaying long-lived particles are creating real secondary vertices.

CHAPTER 7

Simulated samples

In every analysis an excellent knowledge of the processes that are contributing after applying an event selection is of great importance. In case of the basic selection with two jets, one charged high-p_T lepton and missing transverse energy as it is used in this thesis the contributing processes are:

- Single-top t-channel, Wt-channel and s-channel processes
- Top-antitop quark pair production: $t\bar{t}$
- Production of a $W^{\pm}$-boson with additional jets: W+jets
- Production of a Z^0-boson with additional jets: Z+jets
- Di-boson production: WW, WZ and ZZ
- QCD-multijet production

Different approaches are used to predict the contribution of the different processes which will be detailed in the following.

All Feynman diagrams shown are meant as examples of the different process categories. The theoretical calculations of all MC generators are performed including all possible Feynman diagrams of the respective order in perturbation theory.

7.1 General strategy to create simulated events

The creation of simulated proton-proton collision events is done stepwise by dedicated programs. First of all the hard process of an event is calculated based on the matrix elements. This calculation is performed by so called event generators at a certain order

of perturbation theory in the coupling constants. The incoming initial-state protons are described by PDF sets that have already been described in Section 2.6.

On top of the calculated hard process, parton shower models are taking care of the evolution of the QCD. This connects the hard scale of the coloured partons with the hadronic scale where colourless hadrons are formed. This can be done with an accuracy of the order of leading logarithm in perturbation theory. The following hadronisation is done using phenomenological models only where different parameters have been fitted to experimental data. The formed hadrons then decay further to stable particles. Furthermore, in hadron collision events additional secondary interactions can occur. This underlying event is also described by purely phenomenological models.

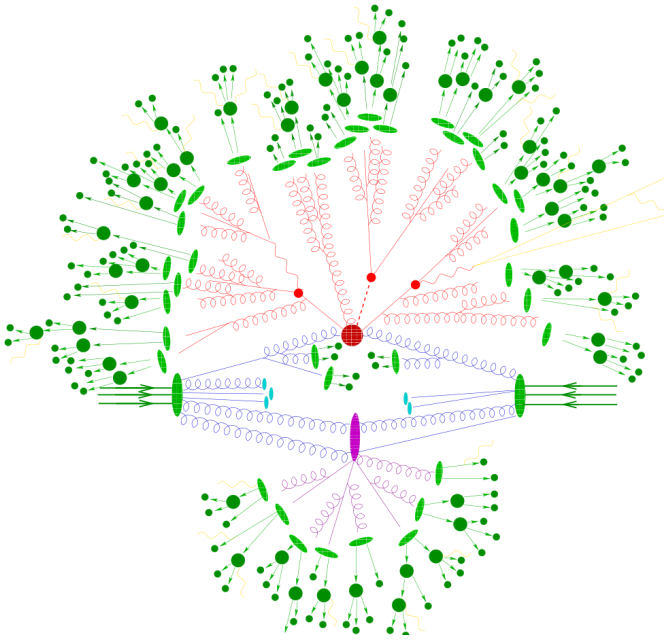


Figure 7.1: Representation of a simulated $t\bar{t}H$ event: Initial state partons (blue), hard process (big red blob), top quark and Higgs decays (small red blobs), QCD radiation (red), underlying event (purple), hadronisation (light green blobs), decaying hadrons (dark green blobs) and photon radiation (yellow) [151].

Apart from the described processes also photon radiation can occur at all stages of the event simulation. In Figure 7.1 the exemplary generation of a $t\bar{t}H$ event is shown with the incoming initial-state partons shown in blue, the hard process as big red blob and the decay of the top quarks and Higgs boson as small red blobs. QCD radiation is shown in red and the hadronisation as light green blobs with decays of the hadrons

displayed as dark green blobs. The underlying event is displayed in purple and photon radiation in yellow [151].

The simulation of the ATLAS detector in all samples described in this chapter is done using the GEANT4 framework [152][48]. The simulation can be done in two different modes being the full simulation of the detector or a fast simulation mode. In the latter case, called ATLFAST-II, presimulated showers stored in memory are used. In this way the very time consuming simulation of particles travelling through the ATLAS calorimeter and electromagnetic particles can be avoided [101].

7.2 Processes including top quarks

The processes including top quarks have already been explained in more detail in Chapter 3. The single-top t-channel, Wt-channel and s-channel processes as well as the $t\bar{t}$ pair production are of course carrying the information about the mass of the top quark and are treated as signal processes.

All nominal samples have been generated using an assumed top quark mass of $m_{\text{top}} = 172.5$ GeV and the decay of the top quark is fixed to $t \to Wb$.

7.2.1 Electroweak single-top t-channel production

To generate t-channel events the AcerMC generator [153] was used with the corresponding leading-order parton density function from CTEQ6L1 [154]. AcerMC is a leading-order generator that is calculating the matrix-element of the two processes $qb \to q't$ and $qg \to q't\bar{b}$ which are combined to one consistent sample with the ACOT method [155]. Generated events are interfaced and passed to Pythia6 (v6.426) [156] which is simulating the parton shower, underlying event and the hadronisation. The parameters used by Pythia6 correspond to the Perugia 2011C tune [157]. Details about the predicted cross-section and the number of events that have been generated can be found in Table 7.1.

7.2.2 Single-top Wt-channel and s-channel production

Both the single-top quark processes from the associated production of an on-shell $W^{\pm}$-boson with a top quark and the s-channel have a smaller cross-section than the single-top t-channel. However, both of them still are giving a significant contribution to the signal and have to be taken into account.

[48] The geometry used for the ATLAS detector is ATLAS-GEO-20-00-01.

Table 7.1: Top quark event MC samples used for this analysis. The cross-section σ column includes filter efficiencies, k-factors and branching ratios [4–6, 61–66].

	σ [pb]	Generator	N_{MC}	dataset ID
t-channel (lepton+jets)	28	AcerMC + Pythia6	9,000,000	110101
$t\bar{t}$ no fully hadronic	137	Powheg + Pythia6	15,000,000	117050
s-channel (lepton+jets)	1.8	Powheg + Pythia6	1,200,000	110119
Wt all decays (DR)	22	Powheg + Pythia6	1,000,000	110140

Both processes are calculated using the Powheg generator [158] which calculates the matrix elements with NLO precision together with the CT10 next-to-leading order PDF sets [59]. AcerMC as well as the Powheg generator is only able to calculate the hard matrix element and again Pythia6 with the Perugia 2011C tune and the leading-order CTEQ6L1 PDF set is used to simulate the parton shower, underlying event and the hadronisation.

In the case of the Wt-channel some overlap has to be removed which occurs at NLO between $t\bar{t}$ production and the Wt-channel. This is done using the diagram removal scheme (DR) where all diagrams that include a $t\bar{t}$ pair are removed from the calculation of the matrix element in case of the Wt-channel calculation [5]. The two generated samples and the corresponding predicted cross-sections are listed with their main parameters in Table 7.1.

7.2.3 Top-antitop quark pair production: $t\bar{t}$

The $t\bar{t}$ pair production has a larger cross-section than the single-top t-channel process. Although the event selection suppresses the contribution from $t\bar{t}$ production two different sources are contributing.

Firstly, dilepton $t\bar{t}$ events are expected to contain two jets, two charged high-p_T leptons and missing transverse momentum. This means if one of the two high-p_T leptons is not reconstructed, the dilepton $t\bar{t}$ events are fulfilling the basic event selection. Secondly, for semileptonic $t\bar{t}$ events four jets, one charged high-p_T lepton and missing transverse momentum is expected. Though, if only two of the four jets are identified the event is selected. Since the branching ratio of semileptonic $t\bar{t}$ events is larger than for dilepton $t\bar{t}$ events the contribution from both decay channels is of a similar amount.

The sample was generated using the same strategy as for the Wt-channel and s-channel with Powheg + Pythia6 with the same PDF sets and tunes. It is also listed

with the predicted cross-section in Table 7.1. Not included is the fully hadronic $t\bar{t}$ decay channel which is expected to contain six jets, no charged high-p_t lepton and no missing transverse momentum and, thus, is effectively suppressed by the event selection.

Table 7.2: Top quark event MC samples used for this analysis with different top quark masses used during event generation. The cross-section σ column includes filter efficiencies, k-factors and branching ratios [4–6, 61–66].

	m_{top} [GeV]	σ [pb]	Generator	N_{MC}	dataset ID
t-channel	165.0	30.539	AcerMC + Pythia6	1,500,000	110113
t-channel	167.5	29.819	AcerMC + Pythia6	1,500,000	110114
t-channel	170.0	29.116	AcerMC + Pythia6	1,500,000	110115
t-channel	175.0	27.777	AcerMC + Pythia6	1,500,000	110116
t-channel	177.5	27.138	AcerMC + Pythia6	1,500,000	110117
t-channel	180.0	26.552	AcerMC + Pythia6	1,500,000	110118
Wt all decays	165.0	25.491	Powheg + Pythia6	3,000,000	110124
Wt all decays	167.5	24.401	Powheg + Pythia6	3,000,000	110126
Wt all decays	170.0	23.361	Powheg + Pythia6	3,000,000	110128
Wt all decays	175.0	21.429	Powheg + Pythia6	3,000,000	110130
Wt all decays	177.5	20.540	Powheg + Pythia6	3,000,000	110132
Wt all decays	180.0	19.689	Powheg + Pythia6	3,000,000	110134
s-channel	165.0	2.1675	Powheg + Pythia6	500,000	110123
s-channel	167.5	2.0411	Powheg + Pythia6	500,000	110125
s-channel	170.0	1.9246	Powheg + Pythia6	500,000	110127
s-channel	175.0	1.7172	Powheg + Pythia6	500,000	110129
s-channel	177.5	1.6264	Powheg + Pythia6	500,000	110131
s-channel	180.0	1.5389	Powheg + Pythia6	500,000	110133
$t\bar{t}$ no fully hadronic	165.0	161.83	Powheg + Pythia6	6,000,000	117836
$t\bar{t}$ no fully hadronic	167.5	150.01	Powheg + Pythia6	6,000,000	117838
$t\bar{t}$ no fully hadronic	170.0	139.18	Powheg + Pythia6	6,000,000	117840
$t\bar{t}$ no fully hadronic	175.0	120.16	Powheg + Pythia6	6,000,000	117842
$t\bar{t}$ no fully hadronic	177.5	111.77	Powheg + Pythia6	6,000,000	117844
$t\bar{t}$ no fully hadronic	180.0	104.07	Powheg + Pythia6	6,000,000	117846

7.3 Processes including top quarks with mass variation

In order to measure the top quark mass in single-top t-channel topologies all of the processes containing top quarks have been generated assuming different values for m_{top}. Six additional samples have been made for all four processes with m_{top} at intervals of 2.5 GeV between 165 GeV and 180 GeV. The MC generators, PDF sets and generator tunes are the same as for the central samples with $m_{\text{top}} = 172.5$ GeV that have been explained in Section 7.2. All samples are summarised with the corresponding cross-sections that have also been calculated in the same way as the central sample in Table 7.2.

7.4 Samples used to estimate systematic uncertainties

To estimate systematic influences of the chosen MC generator on the measurement of m_{top} different configurations are used to generate alternative signal samples. All samples used to estimate systematic variations are listed in Tables 7.3 and 7.4.

7.4.1 t-channel systematic variation samples

In case of the single-top t-channel a sample was generated using the POWHEG generator using the four-flavour scheme [159] with the fixed four-flavour PDF set called CT104f [59]. Here, the simulation of the parton shower and hadronisation remains the same as in the AcerMC + Pythia6 case.

To cover any possible systematic difference in the parton shower and hadronisation, events are generated as well with the POWHEG generator but passed to HERWIG instead of Pythia6. HERWIG (v6.520) [160] works together with JIMMY (v4.31) [161] as an alternative parton shower generator with the same purpose as Pythia6 but with a different modelling approach. In the Pythia6 model the shower is created in a p_T-ordered way meaning that the hardest emission is coming first. The HERWIG approach creates an angular-ordered shower, i.e. the shower is ordered by the emission angle. In this case HERWIG uses the ATLAS AUET2 tune [162] and LO** PDF set [163].

Modelling uncertainties related to the non-perturbative QCD like colour reconnection or the underlying event can be addressed by comparing three POWHEG + Pythia6 samples containing exactly the same events at the level of the hard process but different parton shower tunes. The default tune here is the Perugia2012 tune which can be compared to the Perugia2012loCR and Perugia2012mpiHi tune. All three tunes are based

on the Perugia2011C, Perugia2011CmpiHI and Perugia2011CnoCR tunes documented in Ref. [157][49].

7.4.2 Wt-channel and s-channel systematic variation samples

To estimate any systematic influence of the choice of the generator in case of the Wt-channel and the s-channel, both processes are being generated with the NLO generator called MC@NLO [165] interfaced with Herwig for the parton shower, hadronisation and underlying event. Here, also the ATLAS AUET2 tune and LO** PDF sets are used and the samples with their predicted cross-section are listed in Table 7.3.

For the Wt-channel the same sample as the nominal Powheg + Pythia6 sample is made using the diagram subtraction scheme (DS) instead of the default DR scheme. In this scheme resonant $t\bar{t}$ contributions are subtracted locally from the cross-section in the Wt-channel calculation [166].

7.4.3 $t\bar{t}$ systematic variation samples

To study a systematic influence of the hadronisation of the $t\bar{t}$ process Powheg can be interfaced as well to Herwig or Pythia6 [167, 168]. Two dedicated samples have been created with large statistics using fast-simulation and are listed in Table 7.4.

Furthermore, a $t\bar{t}$ sample generated with MC@NLO + Herwig [169] is created giving the possibility to separate the influence of the matrix element calculation and the parton shower. All systematic variation samples for the $t\bar{t}$ process are using the CT10 PDF sets and the ATLAS AUET2 tune in case of Herwig and the Perugia 2011C tune in the case of Pythia6.

Analogue to what was explained for the t-channel in Section 7.4.1 three Powheg + Pythia6 samples using three different Perugia2012 tunes are used to estimate the uncertainties due to non-perturbative QCD effects in $t\bar{t}$ production.

An additional possible systematic influence is stemming from the amount of Initial and Final State Radiation (ISR/FSR). Two samples are simulated with AcerMC + Pythia6. In both samples the generated events based on the calculated matrix element with AcerMC are exactly the same, but parameters steering the strength of the parton shower and the hadronisation in Pythia6 are modified. This variation is done such that differences in observed distributions of sensitive variables are covered [170, 171]. The varied Pythia6 parameters and their default values are:

[49] All samples described in this section use MadSpin [164] for the decay of the top quark to preserve all spin correlations.

Table 7.3: Top quark event MC samples used for systematic uncertainties in this analysis. The cross-section σ column includes filter efficiencies, k-factors and branching ratios [4–6].

	σ [pb]	Generator	N_{MC}	dataset ID
t-channel (t)	18.39	POWHEG + HERWIG	5,000,000	110086
t-channel ($\bar{t}$)	9.97	POWHEG + HERWIG	5,000,000	110087
t-channel (t)	18.39	POWHEG + PYTHIA6	3,000,000	110090
t-channel ($\bar{t}$)	9.97	POWHEG + PYTHIA6	2,000,000	110091
t-channel (t, P2012)	18.39	POWHEG + PYTHIA6	5,000,000	110070
t-channel ($\bar{t}$, P2012)	9.97	POWHEG + PYTHIA6	5,000,000	110071
t-channel (t, mpiHi)	18.39	POWHEG + PYTHIA6	5,000,000	110072
t-channel ($\bar{t}$, mpiHi)	9.97	POWHEG + PYTHIA6	5,000,000	110073
t-channel (t, P2012, loCR)	18.39	POWHEG + PYTHIA6	5,000,000	110074
t-channel ($\bar{t}$, P2012, loCR)	9.97	POWHEG + PYTHIA6	5,000,000	110075
t-channel	28	aMC@NLO + HERWIG	1,000,000	110095
s-channel ($W \to e\nu_e$)	0.6	MC@NLO + HERWIG	200,000	108343
s-channel ($W \to \mu\nu_\mu$)	0.6	MC@NLO + HERWIG	200,000	108344
s-channel ($W \to \tau\nu_\tau$)	0.6	MC@NLO + HERWIG	200,000	108345
Wt all decays (DS)	22	POWHEG + PYTHIA6	1,000,000	110142
Wt all decays	22	MC@NLO + HERWIG	2,000,000	108346

- ISR: `PARP(67) = 1` and `PARP(64) = 2`
- FSR: `PARP(72) = 0.260` GeV

The `PARP(67)` parameter is proportional to $\frac{1}{\left(\Lambda^{\mathrm{ISR}}_{QCD}\right)^2}$ and the strong coupling α_s^{ISR}. Therefore, it directly steers the amount of simulated ISR. `PARP(64)` steers the coherence imposed by the first emission in the space-like parton shower and `PARP(72)` directly changes the value of the $\Lambda^{\mathrm{FSR}}_{QCD}$ scale and is also proportional to α_s^{FSR}. The variations that are done around the default values are[50]:

- **more PS**: `PARP(67) = 1.40`, `PARP(64) = 0.90`
  `PARP(72) = 0.370` GeV

[50] A more detailed description of the parameters can be found in the PYTHIA6 manual, Ref. [156].

- **less PS**: `PARP(67) = 0.60`, `PARP(64) = 3.50`
  `PARP(72) = 0.110` GeV

All these systematic samples are listed in Table 7.4.

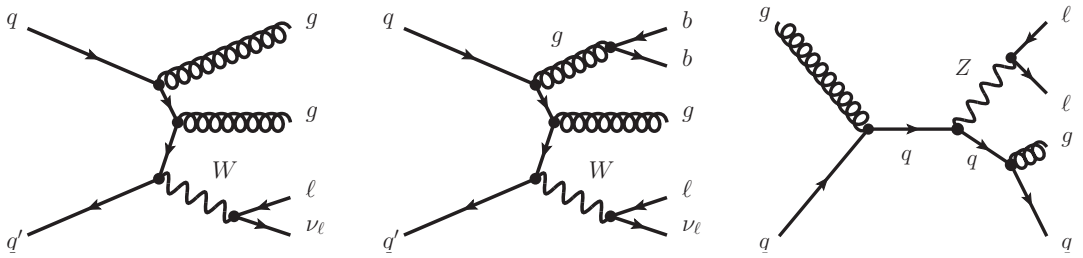
Table 7.4: Top quark event MC samples used for systematic uncertainties in this analysis. The cross-section σ column includes filter efficiencies, k-factors and branching ratios [61–66].

	σ [pb]	Generator	N_{MC}	dataset ID
$t\bar{t}$ no full hadr.	137	MC@NLO + Herwig	15,000,000	105200
$t\bar{t}$ no full hadr. (AFII)	137	Powheg + Herwig	30,000,000	105860
$t\bar{t}$ no full hadr. (AFII)	137	Powheg + Pythia6	75,000,000	117050
$t\bar{t}$ no full hadr. (P2012)	137	Powheg + Pythia6	15,000,000	117428
$t\bar{t}$ no full hadr. (mpiHi)	137	Powheg + Pythia6	15,000,000	117426
$t\bar{t}$ no full hadr. (loCR)	137	Powheg + Pythia6	15,000,000	117429
$t\bar{t}$ no full hadr. less PS	137	AcerMC + Pythia6	15,000,000	117209
$t\bar{t}$ no full hadr. more PS	137	AcerMC + Pythia6	15,000,000	117210

7.5 W+jets background

The most important background is stemming from the production of a $W^{\pm}$-boson with additional jets. Due to the possibility of a leptonic decay of the $W^{\pm}$-boson the charged high-p_T lepton as well as missing transverse momentum is expected for these events. With additional jets e.g. due to gluon radiation the event selection can be passed although no true top quark is present. With the b-tagging requirement the contribution from W+jets can be suppressed, however, the production of a $W^{\pm}$-boson in association with heavy flavour jets remains as a non-reducible background. In Figure 7.2 Feynman diagrams of the production of a $W^{\pm}$-boson with jets (a) and additional heavy flavour induced jets (b) are shown as an example.

To simulate the W+jets background a different approach is used compared to the top quark processes explained in Section 7.2. Multi-leg generators are used that are generating subsamples for processes with a different number of additional partons in addition to the $W^{\pm}$-boson. These subsamples are then merged together to obtain a sample at the same order in perturbation theory for the different jet multiplicities. The



(a) W+jets production (b) W+HF+jets production (c) Z+jets production

Figure 7.2: Feynman diagrams of $V^{\pm,0}$+jets production processes: (a) Production of a $W^{\pm}$-boson and additional jets, (b) production of a $W^{\pm}$-boson with jets and two heavy flavour jets, (c) production of a Z^0-boson with additional jets.

generator used is Sherpa (v1.4.1) [151] with the CT10 PDF sets and the inherent Sherpa tunes. Sherpa includes not only the calculation of the matrix element but also the parton shower, hadronisation and underlying event. By simply merging the different subsamples, overlaps between the $W + n$ and the $W + (n + 1)$ subsample are occurring due to the mixing of the matrix element calculated with n additional partons and the parton shower. To avoid this double-counting the CKKW method [172] is applied by Sherpa.

In Table 7.5 all used subsamples are listed with the expected cross-sections. To avoid any possible double counting between the inclusive $W + n$ parton subsamples and the subsamples with additional heavy flavour pair production, massive c- and b-quarks have been used in the shower. Generator filters are applied to separate the different contributions. The b-filtered subsamples contain only events with any b-hadron within $|\eta| < 4.0$. The c-filtered events must contain a jet associated to a c-hadron with $p_T > 15$ GeV within $|\eta| < 3.0$. In the W+light subsamples both filters are vetoed to avoid any overlap.

7.6 Z+jets background

Another background contribution is coming from the production of a Z^0-boson in association with additional jets. An example of a Feynman diagram for these processes is shown in Figure 7.2(c). The events are effectively suppressed by the event selection but still give a significant contribution due to their relatively large cross-section. The strategy to simulate the Z+jets background is following the same approach as for the W+jets background explained in Section 7.5. The Sherpa generator is used again, with

the same b- and c-filters to avoid any overlap. Details about the different subsamples and the predicted cross-sections are given in Table 7.5.

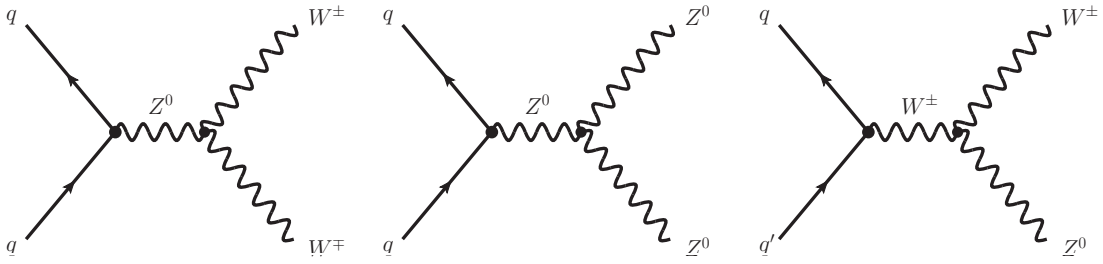
Table 7.5: Background MC samples used for the presented analysis. The cross-section σ column includes filter efficiencies, k-factors and branching ratios [173].

	σ [pb]	Generator	N_{MC}	dataset ID
$W \to e\nu_e$ b-, c-vetoed	11324.5	SHERPA	50,000,000	167742
$W \to \mu\nu_\mu$ b-, c-vetoed	11324.5	SHERPA	50,000,000	167745
$W \to \tau\nu_\tau$ b-, c-vetoed	11324.5	SHERPA	50,000,000	167748
$W \to e\nu_e$ b-filtered	154.0	SHERPA	15,000,000	167740
$W \to \mu\nu_\mu$ b-filtered	154.0	SHERPA	15,000,000	167743
$W \to \tau\nu_\tau$ b-filtered	154.0	SHERPA	15,000,000	167746
$W \to e\nu_e$ c-filtered	591.8	SHERPA	10,000,000	167741
$W \to \mu\nu_\mu$ c-filtered	591.8	SHERPA	10,000,000	167744
$W \to \tau\nu_\tau$ c-filtered	591.8	SHERPA	10,000,000	167747
$Z \to ee$ b-, c-vetoed	855.68	SHERPA	5,000,000	167751
$Z \to \mu\mu$ b-, c-vetoed	855.68	SHERPA	5,000,000	167754
$Z \to \tau\tau$ b-, c-vetoed	855.68	SHERPA	5,000,000	167757
$Z \to ee$ b-filtered	34.72	SHERPA	4,000,000	167749
$Z \to \mu\mu$ b-filtered	34.72	SHERPA	4,000,000	167752
$Z \to \tau\tau$ b-filtered	34.72	SHERPA	4,000,000	167755
$Z \to ee$ c-filtered	351.68	SHERPA	3,000,000	167750
$Z \to \mu\mu$ c-filtered	351.68	SHERPA	3,000,000	167753
$Z \to \tau\tau$ c-filtered	351.68	SHERPA	3,000,000	167756
WW di-boson	35.28	HERWIG	2,500,000	105985
WZ di-boson	11.40	HERWIG	1,000,000	105987
ZZ di-boson	2.325	HERWIG	245,000	105986
JF17 dijet	93052540	PYTHIA8	111,000,000	129160

7.7 Di-boson production

The smallest background contribution is coming from the production of two vector-bosons. Three exemplary Feynman diagrams of these processes are shown in Figure 7.3.

The three combinations are simulated using the generators HERWIG and JIMMY only with the ATLAS AUET2 tune and CTEQ6L1 PDF sets. They are listed in Table 7.5 where the given cross-sections have been corrected using an NLO prediction [174]. The three samples are filtered to always contain one lepton with $p_T > 10$ GeV and $|\eta| < 2.8$.

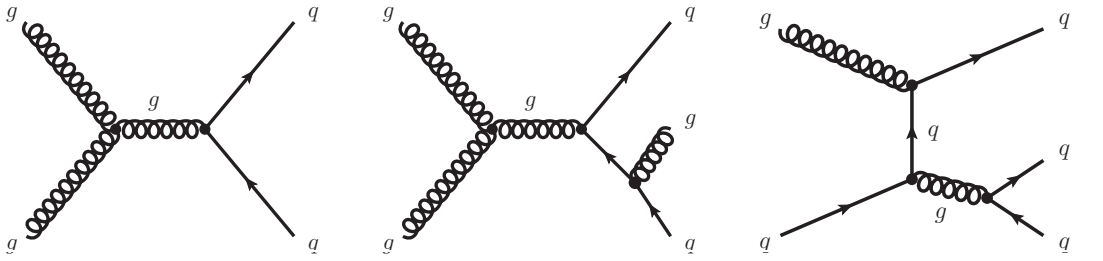


(a) $W^\pm W^\mp$ production (b) $Z^0 Z^0$ production (c) $W^\pm Z^0$ production

Figure 7.3: Feynman diagrams of di-boson production processes with two vector-bosons.

7.8 QCD-multijet production

The probability that an event from pure QCD-multijet production passes the event selection is small since most selection criteria are suppressing this kind of background. However, the cross-section of QCD-multijet events is several orders of magnitude larger than all other contributing processes. In the end, this leads to a significant amount of selected events that have to be taken into account. In Figure 7.4 the Feynman diagrams of three possible QCD-multijet production processes are shown for exemplarity.



(a) QCD-dijet production (b) QCD-multijet production (c) QCD-multijet production

Figure 7.4: Feynman diagrams of QCD-multijet production processes: (a) Dijet production with two quarks, (b) multijet production with two quarks and one gluon, (c) multijet production with three quarks.

To obtain a multijet sample with feasible statistics an enormous amount of computing as well as storage capacity would be necessary making it expensive to obtain the sample from simulation. That is why estimating the background contribution from QCD-multijet events is a combination of data-driven techniques and simulation. The applied fitting technique is explained in more detail in Section 8.2.1.

The only simulated dataset used to obtain a template for the jet-lepton model, see 8.2.2, was created using PYTHIA8 (v8.160) [136] with the AU2 generator tune and CTEQ6L1 leading-order PDF sets. To enhance the available statistics, the dijet sample denoted as JF17 dijet, was filtered in a way that every event must contain at least one jet with $E_T(jet) > 17$ GeV and $|\eta(jet)| < 2.7$ at generator level. The sample is listed in Table 7.5 with its predicted cross-section. Since the simulated dataset is only used to obtain a template shape and the normalisation is taken from a fit to data the cross-section is not used in the further analysis.

CHAPTER 8

Basic event selection and background estimation

In this chapter the selection of events based on basic reconstructed objects is explained. The selection in three kinematic regions is detailed in Section 8.1. The second part in Section 8.2 is giving a detailed explanation of the strategy to estimate the expected background. A comparison of the expected and observed event yield and kinematic modelling is presented in Sections 8.3 and 8.4.

8.1 Event selection

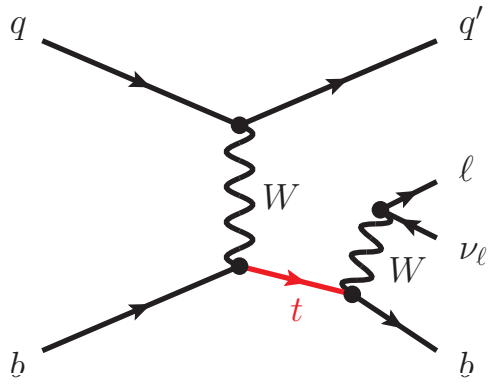


Figure 8.1: Feynman diagram of single-top t-channel production and a possible decay channel of the top quark

In Figure 8.1 the Feynman diagram of the single-top quark production in the t-channel is shown. The basic event selection is following this final state signature with the main criteria being:

- exactly one electron or muon
- $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV, representing ν_ℓ
- exactly two jets
- one of the two jets being b-tagged

All objects used in the event selection have to be well reconstructed with the algorithms and selection cuts detailed in Chapter 5. To further reduce contributions from

QCD-multijet events two cuts calculated from the reconstructed lepton, jets and missing transverse energy are being used which will be explained in the following.

8.1.1 Multijet veto: triangular and $m_T(W)$ cut

If an event contains a fake electron that did not arise from $W^{\pm}$-boson decay the event tends to have lower missing transverse energy than events with a real $W^{\pm}$-boson due to the missing neutrino contributions. Therefore, the cut on $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV already is a good way to reduce background contributions from QCD-multijet events. To exploit the expected existence of a real $W^{\pm}$-boson further, the transverse $W^{\pm}$-boson mass can be reconstructed from the lepton and the missing transverse energy:

$$m_{\mathrm{T}}(W) \;=\; \sqrt{2\left[p_{\mathrm{T}}(\ell)E_{\mathrm{T}}^{\mathrm{miss}} - \vec{p}_{\mathrm{T}}(\ell)\cdot\vec{E}_{\mathrm{T}}^{\mathrm{miss}}\right]}. \tag{8.1}$$

Here, $\vec{p}_{\mathrm{T}}(\ell)$ denotes the transverse momentum of the lepton with $p_{\mathrm{T}}(\ell) = |\vec{p}_{\mathrm{T}}(\ell)|$. Since $m_{\mathrm{T}}(W)$ is expected to be lower for QCD-multijet events, all events are required to satisfy the requirement of $m_{\mathrm{T}}(W) > 50$ GeV.

In addition an isolation criterion is applied for low p_T leptons depending on the p_T of the lepton itself. The cut is using the reconstructed lepton and the leading jet, j_1, which is defined as the jet with the highest transverse momentum present in the event[51]:

$$p_{\mathrm{T}}(\ell) > 40 \text{ GeV}\left(1 - \frac{\pi - |\Delta\phi(j_1,\ell)|}{\pi - 1}\right). \tag{8.2}$$

In Figure 8.2 the lepton p_{T} vs. $\Delta\phi(j_1,\ell)$ plane is shown to illustrate the reduction of the QCD-multijet background due to the cut only. Shown is a histogram containing collision data at $\sqrt{s} = 8$ TeV with histograms from all expected processes apart from the QCD-multijet contribution being subtracted[52]. This is leading to a histogram that contains the contribution from QCD-multijet events only. These are concentrated in the two triangular regions at high values of $|\Delta\phi(j_1,\ell)|$ and low lepton p_{T}, which is effectively removed by the cut given by Equation (8.2). In Appendix A.1 the two-dimensional $(\Delta\phi(j_1,\ell), p_{\mathrm{T}})$ distributions can be found for the data and the different contributing processes separately. One can see that the cut effectively suppresses the QCD-multijet background, while most of the events from signal processes remain.

[51] The physics motivation of this cut is to reduce e.g. dijet events with two jets produced back-to-back. If one of the two jets is reconstructed as a fake lepton and another jet, e.g. from gluon radiation, is reconstructed these kind of multijet events might be able to pass the event selection.

[52] To subtract the other processes, simulated events, explained in Chapter 7, have been used.

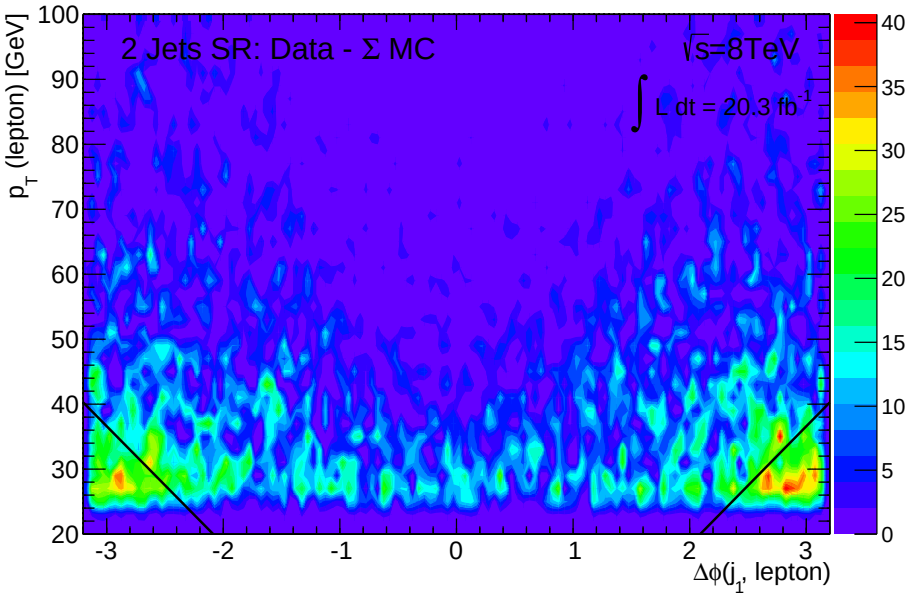


Figure 8.2: The lepton p_{T} depending on the $\Delta\phi\,(j_1, \ell)$ is shown for the data where all MC predictions for background processes are subtracted. Thus, only the contribution from QCD-multijet production remains.

8.1.2 Definition of the signal and control regions

Three different not overlapping regions are defined and used in this thesis. The signal region is used to finally measure the mass of the top quark. Two different control regions are constructed in a way that they are dominated by a certain background process. This makes it possible to check the kinematic modelling of the dedicated background. All three regions are using the same selection of leptons, $E_{\mathrm{T}}^{\mathrm{miss}}$, jets, and the cuts to reduce QCD-multijet contribution, explained in Section 8.1.1, are applied. They differ in the choice of either exactly one or two b-tagged jets:

Signal region (SR) In the signal region (SR) exactly one of the two jets is required to be b-tagged with a tight tagging requirement. Since one of the largest background contributions of the single-top t-channel production is the production of a $W^{\pm}$-boson with additional c-jets the MV1c tagging algorithm is chosen. This algorithm is optimised to also reject c-quark induced jets and is explained in detail in Section 6.1.4. Using MV1c a jet is b-tagged if the MV1c tagger weight is $w_{\mathrm{MV1c}} > 0.9195$. This cut value corresponds to a b-tagging efficiency of 50% in a $t\bar{t}$ control sample defined in Ref. [128]

that was already mentioned in Section 6.2. In the same control sample the mistag rate is 3.9% in the case of c-jets and 0.07% for light-flavour-jets.

The number of single-top t-channel events and the amount of events from other top processes have been measured using a binned template likelihood fit [71]. All distributions in the signal region are normalised to the fit result using correction factors. A systematic uncertainty on the measurement of m_{top} is arising from this source which is explained in more detail in Section 10.4.3.

$W^{\pm}$-boson control region ($W^{\pm}$ CR) The dominant background not containing top quarks arises from the production of a $W^{\pm}$-boson with additional jets. Therefore, the modelling of the background has to be checked carefully. The main difference of this background compared to the single-top t-channel signal process is that there is always a b-hadron present from the decay of the top quark in case of the signal. Because of that the $W^{\pm}$-boson control region is defined by using a less stringent b-tagging requirement in two steps. To avoid any overlap with the signal region all events containing exactly one jet with $w_{\text{MV1c}} > 0.9195$ are vetoed. However, exactly one jet is required to have $w_{\text{MV1}} > 0.3511$. This cut value of the MV1 tagging algorithm corresponds to a b-tagging efficiency of 80% in the same $t\bar{t}$ control sample as used for the MV1c algorithm. As the tagging requirement is not as tight as in the signal region the mistag rates are higher. They are 32.5% in the case of c-jets and 4.0% for light-flavour-jets.

The explained selection ensures that the kinematics and flavour fractions are similar to the signal region. Also the choice of the b-jet is well defined due to the tagging requirement.

$t\bar{t}$ control region ($t\bar{t}$ CR) A second control region is defined to also check the modelling of the $t\bar{t}$ production process which also gives a significant signal contribution in the top quark mass measurement. The main difference compared to the single-top t-channel is the presence of two b-hadrons from the decays of the two top quarks. So the $t\bar{t}$ control sample is obtained by requiring both jets to be tagged with the same tagging requirement as in the signal region. This means that both jets must fulfil $w_{\text{MV1c}} > 0.9195$.

In the later analysis observables will be defined which act as input variables of a neural network based discriminant. To define some of the observables it is necessary to assign the two reconstructed jets as one "b-jet" and one "light-flavour-jet"[53]. Since both of the two jets have to fulfil the same tagging requirement another criterion for the assignment is needed. A true t-channel event is often characterised by a light-flavour-jet

[53] A detailed description of the observables can be found in Section 9.2.

in the forward region[54], which is why the jet with the highest value of $|\eta|$ is treated as the "light-flavour-jet" in the $t\bar{t}$ control region.

8.2 Background estimation

To estimate the different background contributions simulated events, introduced in Chapter 7, and data-driven techniques are used. These techniques will be explained in the following.

8.2.1 QCD-multijet events

Various processes related to the strong interaction can lead to events passing the signal event selection. Some examples for different contributions are:

- **Semileptonic b-quark decays:** A significant amount of b-quarks is decaying to either c-quarks or light-flavour-quarks and a $W^{\pm}$-boson that can decay leptonically. These leptons are usually non-isolated leptons but in some cases can appear as isolated leptons fulfilling all reconstruction requirements. A neutrino is present and can be reconstructed as $E_{\mathrm{T}}^{\mathrm{miss}}$. Therefore, semileptonic b-quark decays are a significant QCD-multijet background contribution.

- **Long-lived weakly decaying particles ($\pi^{\pm}$, K mesons):** Long-lived particles can decay when flying through the detector and so they can be mis-identified as a lepton.

- **Mis-identified electrons:** Various sources can lead to a mis-identified reconstructed electron. π^0 particles are decaying with a very high fraction to two photons seen as showers in the electromagnetic calorimeter. These are possibly mis-identified if randomly associated with a track of for example a $\pi^{\pm}$. Another source could arise from electrons from photon conversion or from prompt photons reconstructed in the EM calorimeter.

All of these sources are summarised as the QCD-multijet background. The corresponding cross-section is huge compared to all other processes. Because of the very low mis-identification efficiency this would lead to an enormous amount of computing resources necessary to fully simulate this background contribution. That is why different data-driven techniques have been developed to be used in the electron and muon channel,

[54] Further details about the light-flavour-jet in the forward region can also be found in Section 9.2.1.

respectively. The shape of QCD-multijet events is obtained using the jet-lepton model in the electron-channel, see Section 8.2.2, and the anti-muon model in the muon channel, see Section 8.2.3. The assignment of the two models to the two channels has been made to ensure the best possible modelling of the control variables. Both approaches have to be normalised which is done using a binned maximum-likelihood fit in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution explained in Section 8.2.4.

8.2.2 The jet-lepton model

As explained in the previous section the main difference between the QCD-multijet background and the signal is the presence of a reconstructed lepton. The idea behind the jet-lepton model is to choose a jet with similar kinematics compared to a real reconstructed lepton and to use this jet in place of the lepton [175, 176]. With this approach an event from the QCD-multijet background is able to pass the selection with similar kinematics as the signal sample.

In principle it would be possible to take this so called jet-lepton sample from collision data but to select a feasible sample the usage of jet triggers with low jet multiplicities would be necessary. Those triggers are highly prescaled leading to a selected sample that would suffer from statistical fluctuations. Therefore, the jet-lepton model is fully based on simulation and since it is only used as a background model no dedicated trigger selection is required. The model is using a simulated JF17 dijet sample listed in Table 7.5 explained in Section 7.8.

Table 8.1 lists the cuts that the selected "jet-lepton" has to fulfil. The cuts make sure that the same coverage in transverse energy and η is taken as it is the case for a real lepton and the fraction of energy measured in the electromagnetic calorimeter is relatively large. Additionally, the number of reconstructed tracks associated to the "jet-lepton" should be larger than three to avoid converted photons or so called trident electrons. These are originating from electrons with bremsstrahlung and subsequent photon conversion[55].

To ensure a high purity of QCD-multijet events and only low contamination from W+jets background all events are vetoed that contain another lepton. In case of electrons they are reconstructed with the same criteria as described in Section 5.4.2 but with the "loose++" identification [122] instead of the "tight++" definition. Also no isolation requirement is used for these electrons. The veto against reconstructed muons uses the same definition as given in Section 5.3.

[55] An example for this could be $e^- \rightarrow e^- + \gamma(\rightarrow e^+e^-)$ when the photon is interacting with the ID.

Table 8.1: Applied cuts to define a jet-lepton sample. The CRACK region is the region between the barrel and the end-cap region of the electromagnetic calorimeter and is defined as $1.37 \leq \eta \geq 1.52$.

Variable	Cut
Transverse energy of jet	$E_{\mathrm{T}} > 25$ GeV
η of jet	$\lvert\eta\rvert < 2.47$, no CRACK region
EM fraction	$0.8 < f_{\mathrm{EM}} < 0.95$
Number of tracks within the jet	$n_{\mathrm{tracks}} > 3$

In summary the selected events must contain exactly one "jet-lepton", no other reconstructed lepton, exactly two jets and the QCD-multijet veto explained in Section 8.1.1 must be satisfied. Since the model is only constructed to obtain a shape a looser b-tagging is required which remains the same in all kinematic regions defined in Section 8.1.2. Exactly one of the two jets must fulfil $w_{\mathrm{MV1}} > 0.3511$ which corresponds to the same b-tagging efficiency as in the $W^{\pm}$-boson control region. This ensures a well-defined b-jet and a similar flavour fraction in all selected samples.

8.2.3 The anti-muon model

A second data-driven ansatz used to obtain a sample that is highly enriched with muons from QCD-multijet events is the anti-muon model. By inverting or changing cuts applied in the usual muon selection, it is possible that an event containing a fake muon or a non-isolated muon passes the standard event selection. An anti-muon has to fulfil the same requirements, given in Section 5.3.2, as the well reconstructed muons apart from the inverted or changed cuts. The list of cuts that are changed is given in Table 8.2. eTcone20 means the transverse energy measured in a cone within $\Delta R < 0.2$ around the

Table 8.2: Cuts that are different from the nominal cuts in the anti-muon sample.

Variable	
impact parameter z_0	z_0 can be any value
isolation	$\mathrm{eTcone20}/p_T^{\mathrm{lepton}} > 0.03, \mathrm{MiniIso10_R}/p_T^{\mathrm{lepton}} < 0.1$
energy loss type	not isolated (energyLossType = 1)
energy loss	energyLoss < 6 GeV

reconstructed muon track not including the muon track itself. $\mathrm{MiniIso10_R}$ is explained in Section 5.3.3 and p_T^{lepton} means the transverse momentum of the reconstructed muon.

The impact parameter, z_0, is explained in Section 5.1. More information about the energy loss of the muon when travelling through the calorimeter can be found in Ref. [133].

All cuts are favouring non-isolated muons from QCD-multijet decays and by requiring energyLoss < 6 GeV fake jets from muons losing large amounts of energy in the calorimeter can be avoided.

Apart from the mentioned changes in the muon reconstruction the event selection remains completely unchanged. Since the selected sample is obtained from collision data it still contains a small amount of real signal muons mainly coming from decays of real $W^{\pm}$- and Z^0-bosons.

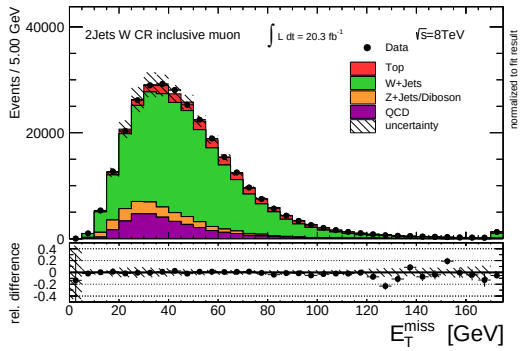
8.2.4 Estimation of the QCD-multijet background

Both the jet-lepton model in the electron channel and the anti-muon model in the muon channel can only be used to obtain a shape for the QCD-multijet background but the normalisation has to be estimated separately. This is done using a binned maximum-likelihood fit [139] on the distribution of the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. The obtained QCD-multijet background estimates obtained with the fitting procedure explained in the following are shown in Table 8.3. In Figure 8.3 the fitted $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions are

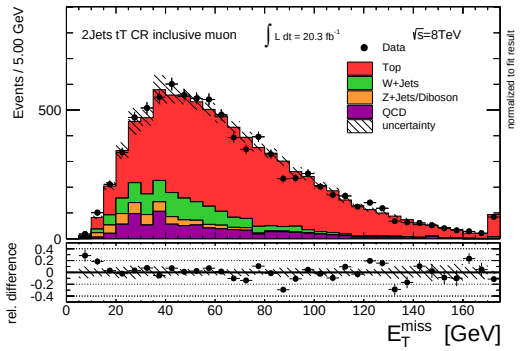
Table 8.3: Estimates of the QCD-multijet background for the signal region and the two control regions using the binned maximum-likelihood fit in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. The quoted numbers are the expected number of events in each region. The uncertainties given reflect the uncertainty of the QCD-multijet normalisation of 50 %.

	$W^{\pm}$-boson control region		signal region	
Channel	events	fraction	events	fraction
electrons forward (end-cap)	13300 ± 6700	25.1 %	3300 ± 1600	20.7 %
electrons central (barrel)	11500 ± 5800	10.2 %	2300 ± 1100	5.0 %
electrons combined	24800 ± 12400	15.0 %	5500 ± 2800	9.0 %
muon	23200 ± 11600	10.3 %	6000 ± 3000	7.2 %
	$t\bar{t}$ control region			
Channel	events	fraction		
electrons forward (end-cap)	140 ± 70	13.0 %		
electrons central (barrel)	250 ± 130	5.4 %		
electrons combined	390 ± 200	6.9 %		
muon	670 ± 340	9.2 %		

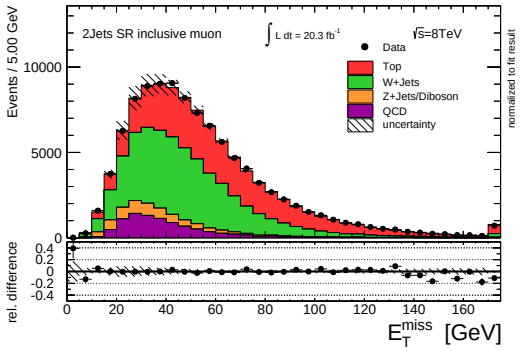
shown for the muon channel using the template from the anti-muon model for the shape of the QCD-multijet background. In Figure 8.4 the electron channel is shown using the jet-lepton template.



(a) Muon channel in the $W^{\pm}$-boson control region



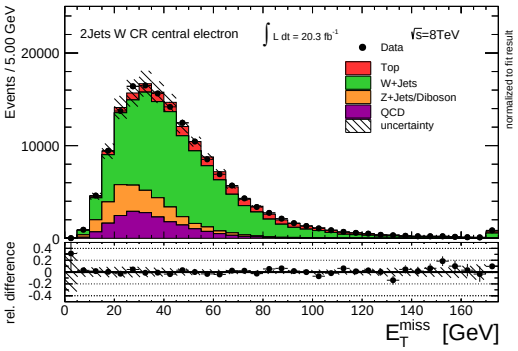
(b) Muon channel in the $t\bar{t}$ control region



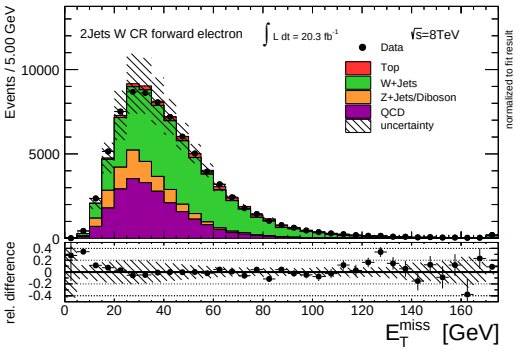
(c) Muon channel in the signal region

Figure 8.3: Result of the binned maximum-likelihood fit to the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions for the signal region (c), for the $W^{\pm}$-boson control region (a) and the $t\bar{t}$ control region (b) for the muon channel. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the QCD-multijet normalisation.

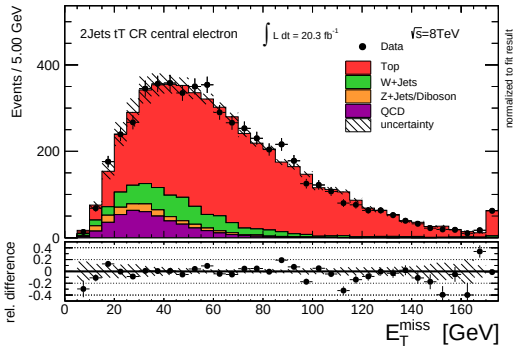
The $E_{\mathrm{T}}^{\mathrm{miss}}$ in QCD-multijet events is expected to be smaller compared to other contributing processes and is therefore sensitive to the amount of the corresponding background. The fit is done after all selection cuts explained in Section 8.1 but without the cut of $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV to increase the sensitivity. Since the expected background fraction is different in the three kinematic regions the fit has to be done separately in the signal region and the two control regions. In the electron channel to ratio of events between the barrel and end-cap region of the EM calorimeter is not well described by the jet-



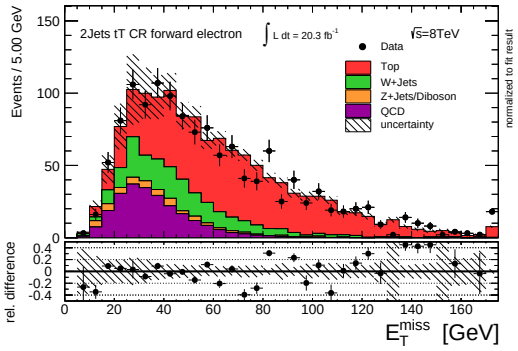
(a) Central electron channel in the $W^{\pm}$-boson control region



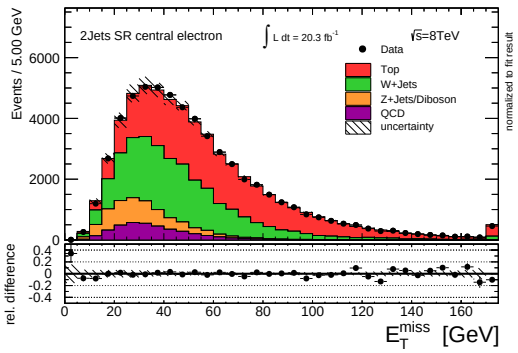
(b) Forward electron channel in the $W^{\pm}$-boson control region



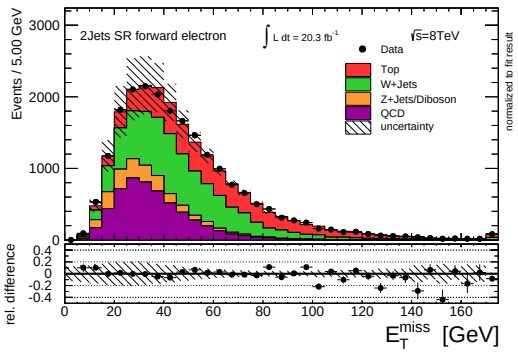
(c) Central electron channel in the $t\bar{t}$ control region



(d) Forward electron channel in the $t\bar{t}$ control region



(e) Central electron channel in the signal region



(f) Forward electron channel in the signal region

Figure 8.4: Result of the binned maximum-likelihood fit to the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions for the signal region ((e),(f)), the $W^{\pm}$-boson control region ((a),(b)), and the $t\bar{t}$ control region ((c),(d)) for electrons in the central (left) and forward (right) regions, respectively. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the QCD-multijet normalisation.

lepton model. Therefore, the fit is split-up at the transition region and done separately for electrons in the barrel ($|\eta_{\text{electron}}| < 1.5$) and the end-cap ($|\eta_{\text{electron}}| > 1.5$).

The other contributing processes which are top quark, W+jets, Z+jets and diboson processes are taken from simulation and their rate uncertainties are taken into account as constrained nuisance parameters. That means that these processes are first of all normalised to their theoretical cross-section given in Chapter 7. The constrained correction factors are then fitted simultaneously during the fit of the $E_{\text{T}}^{\text{miss}}$ distribution. The contribution from Z+jets and diboson production is very small and has been fixed to the theoretical prediction. This increases the sensitivity to distinguish the dominant processes. Also the top quark processes containing the dominant t-channel, $t\bar{t}$, the Wt-channel and the s-channel have been grouped together in one template. The correction factors obtained by the fit are shown in Table 8.4. In all regions the $E_{\text{T}}^{\text{miss}}$ distribution is well described by the fit result.

Table 8.4: Scale factors for the combined contributions from W+jets and from $t\bar{t}$ and single-top quark production in the signal region and the two control regions as obtained from the simultaneous binned maximum-likelihood fit of the $E_{\text{T}}^{\text{miss}}$ distribution.

	signal region		
Channel	electrons		muons
	forward	central	inclusive
$t\bar{t}$ / single-top	0.86 ± 0.05	0.91 ± 0.02	0.90 ± 0.02
W+jets	0.96 ± 0.07	1.10 ± 0.07	1.15 ± 0.03
Z+jets / Diboson	fixed	fixed	fixed
	$W^{\pm}$-boson control region		
Channel	electrons		muons
	forward	central	inclusive
$t\bar{t}$ / single-top	fixed	fixed	fixed
W+jets	0.89 ± 0.01	0.96 ± 0.01	1.00 ± 0.01
Z+jets / Diboson	fixed	fixed	fixed
	$t\bar{t}$ control region		
Channel	electrons		muons
	forward	central	inclusive
$t\bar{t}$ / single-top	0.98 ± 0.04	1.00 ± 0.02	1.01 ± 0.03
W+jets	fixed	fixed	fixed
Z+jets / Diboson	fixed	fixed	fixed

8.3 Event yield and modelling in the control regions

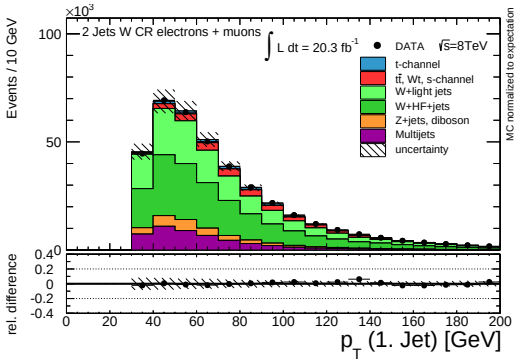
In Figures 8.5 and 8.6 the modelling of important kinematic variables in the $W^{\pm}$-boson and the $t\bar{t}$ control region, respectively, is shown. The number of expected events in these two regions is split up into the contributing processes and listed in Tables 8.5 and 8.6. All tables and figures are using the normalisation obtained using the fits of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution described in Section 8.2.4. Good agreement with the number of observed events is seen and also all important observables are well described by the simulation and the normalised QCD-multijet background model. The uncertainty on the QCD-multijet normalisation is fixed to 50% due to the reasons formulated in Section 10.4.3.

Table 8.5: Number of observed and expected events in the $W^{\pm}$-boson control region. The uncertainties shown are derived using the statistical uncertainty. The uncertainty on the QCD-multijets is fixed to 50%.

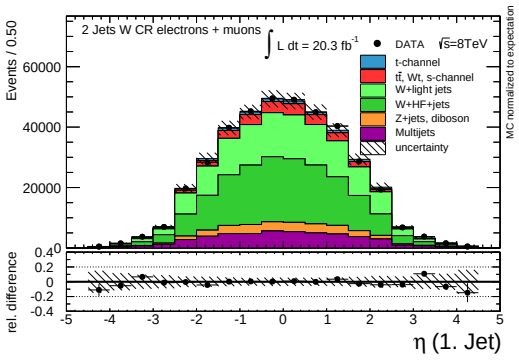
	electron channel			muon channel
Process	barrel	end-cap	total	total
t-channel	3403 ± 21	777 ± 10	4180 ± 24	5410 ± 28
$t\bar{t}$	7706 ± 48	1821 ± 23	9527 ± 53	11923 ± 61
s-channel	123 ± 2	53 ± 2	177 ± 3	247 ± 4
Wt-channel	1361 ± 31	289 ± 15	1650 ± 34	2018 ± 39
W+jets (b)	6124 ± 40	3224 ± 29	9348 ± 49	14540 ± 66
W+jets (c)	41460 ± 240	13820 ± 140	55280 ± 270	82300 ± 340
W+jets (light)	31900 ± 350	15540 ± 240	47440 ± 420	73240 ± 560
Z+jets/Diboson	9200 ± 240	4300 ± 190	13510 ± 310	11820 ± 230
QCD-multijets	11500 ± 5800	13300 ± 6700	24800 ± 12400	23200 ± 11600
Total expected	112800 ± 5800	53100 ± 6700	166900 ± 12400	224700 ± 11600
data	112362	52643	165005	224914

In Figure 8.7(a) the fractions of the different processes in the two control regions are shown. It can be seen that the $W^{\pm}$-boson control region is strongly dominated by W+jets events with the amount of W+HF+jets being a little larger than the amount of W+light jets. Also there is a sizeable contribution from QCD-multijet events and a smaller fraction from Z+jets and diboson events. Processes including top quarks also contribute but do have no overlap with the events selected in the signal region.

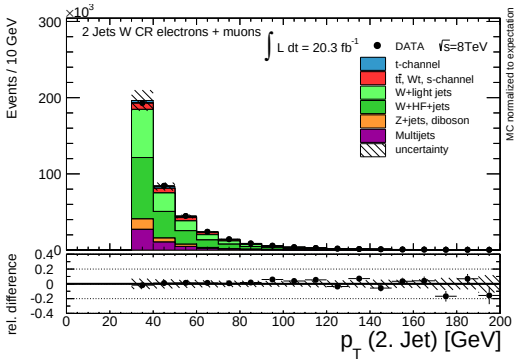
The $t\bar{t}$ control region is clearly dominated by top quark processes shown in red in Figure 8.7(b) not including the single-top t-channel. As can be seen in Table 8.6 these



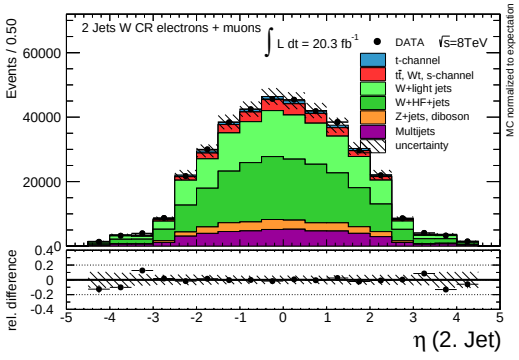
(a) Transverse momentum of the leading jet



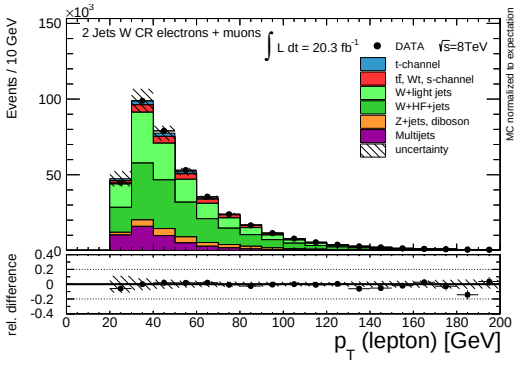
(b) Pseudorapidity of the leading jet



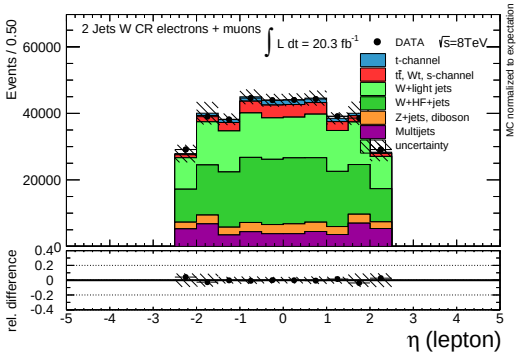
(c) Transverse momentum of the second leading jet



(d) Pseudorapidity of the second leading jet

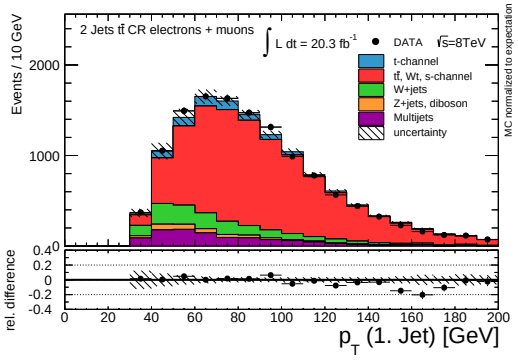


(e) Transverse momentum of the lepton

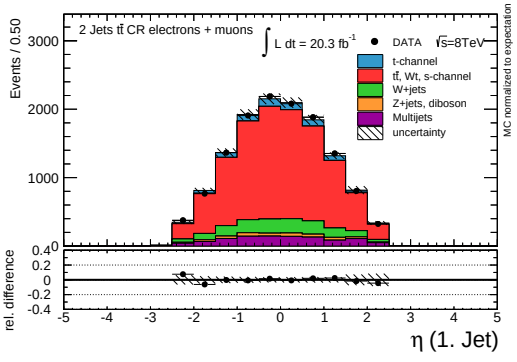


(f) Pseudorapidity of the lepton

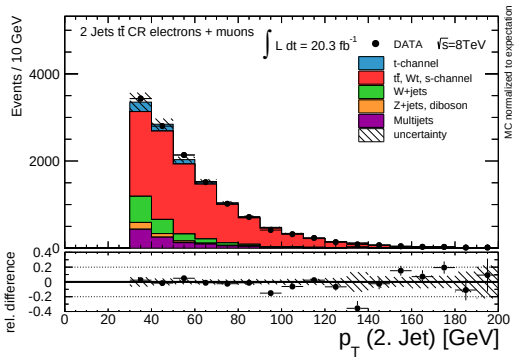
Figure 8.5: Basic kinematic distributions of the combination of muon and electron channels in the $W^{\pm}$-boson control region. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the QCD-multijet normalisation.



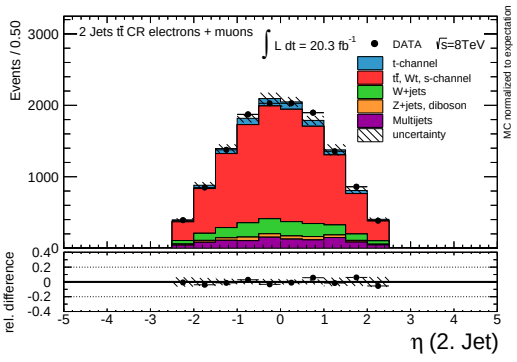
(a) Transverse momentum of the leading jet



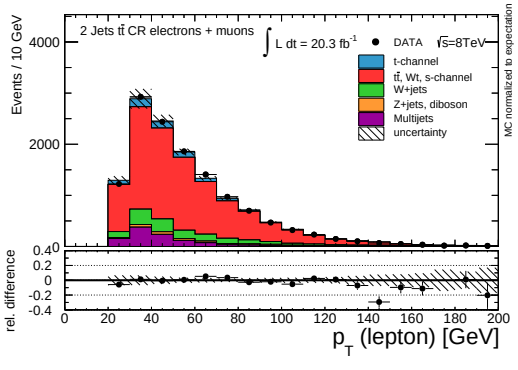
(b) Pseudorapidity of the leading jet



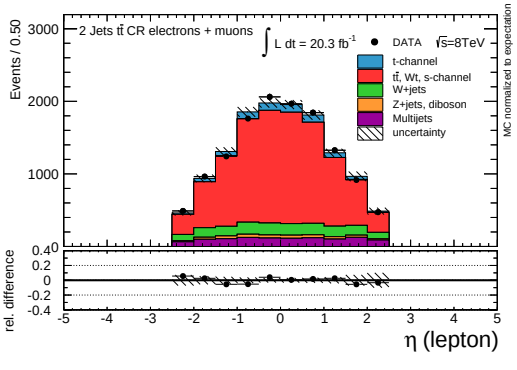
(c) Transverse momentum of the second leading jet



(d) Pseudorapidity of the second leading jet



(e) Transverse momentum of the lepton



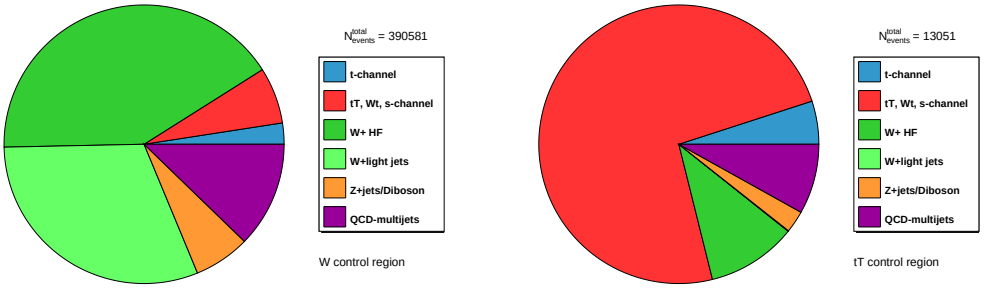
(f) Pseudorapidity of the lepton

Figure 8.6: Basic kinematic distributions of the combination of muon and electron channels in the $t\bar{t}$ control region. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the QCD-multijet normalisation.

Table 8.6: Number of observed and expected events in the $t\bar{t}$ control region. The uncertainties shown are derived using the statistical uncertainty. The uncertainty on the QCD-multijets is fixed to 50%.

	electron channel			muon channel
Process	barrel	end-cap	total	total
t-channel	241 ± 6	42 ± 3	283 ± 7	371 ± 8
$t\bar{t}$	3381 ± 31	631 ± 13	4011 ± 34	4954 ± 39
s-channel	148 ± 3	36 ± 1	183 ± 3	244 ± 3
Wt-channel	93 ± 8	26 ± 5	119 ± 9	128 ± 10
W+jets (b)	396 ± 10	168 ± 6	564 ± 12	733 ± 13
W+jets (c)	31 ± 8	3 ± 2	34 ± 8	25 ± 6
W+jets (light)	4 ± 4	0 ± 0	4 ± 4	5 ± 5
Z+jets/Diboson	88 ± 4	28 ± 2	116 ± 5	217 ± 7
QCD-multijets	250 ± 130	139 ± 70	390 ± 200	670 ± 340
Total expected	4631 ± 130	1073 ± 72	5700 ± 200	7350 ± 340
data	4642	1061	5703	7338

events are mostly $t\bar{t}$ events. Again events from QCD-multijet production, V+jets[56] and diboson are included and a small fraction of t-channel events are selected without any overlap with the signal region.



(a) Fractions in the $W^{\pm}$-boson control region

(b) Fractions in the $t\bar{t}$ control region

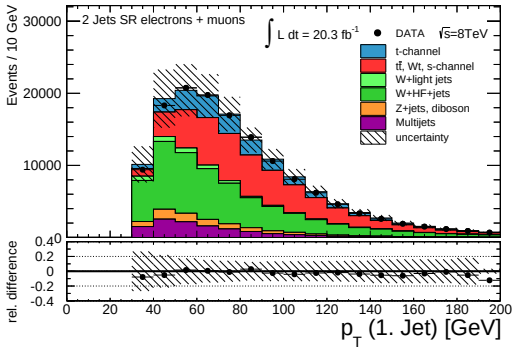
Figure 8.7: Event fractions of the different processes in the $W^{\pm}$-boson control region (a) and the $t\bar{t}$ control region (b).

[56] This means the aggregate of W+jets and Z+jets.

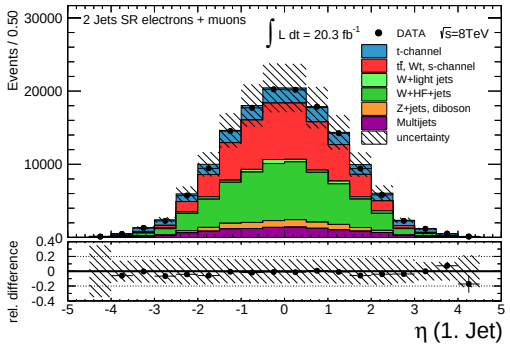
8.4 Event yield and modelling in the signal region

The number of observed and expected events in the signal region is shown in Table 8.7. Also here good agreement is found in all kinematic observables shown in Figures 8.8 and 8.9.

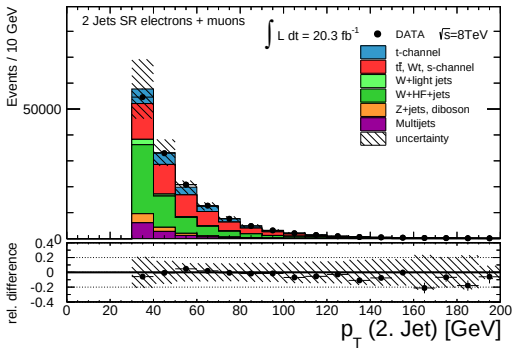
The selected sample has dominant contributions from $t\bar{t}$ production and the W+jets background and will be further enhanced with events from t-channel production by a neural network discriminant. This will be explained in Chapter 9.



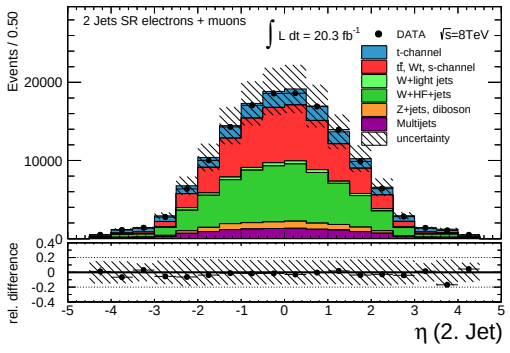
(a) Transverse momentum of the leading jet



(b) Pseudorapidity of the leading jet

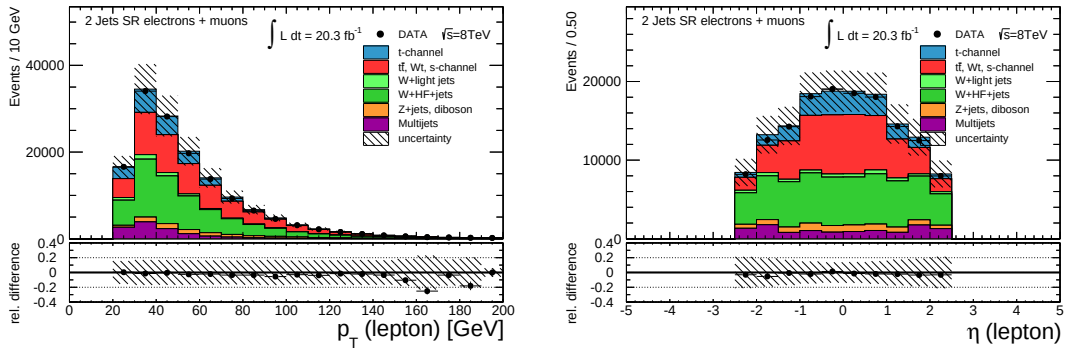


(c) Transverse momentum of the second leading jet



(d) Pseudorapidity of the second leading jet

Figure 8.8: Basic kinematic distributions of the combination of muon and electron channels in the signal region. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the W+jets normalisation.



(a) Transverse momentum of the lepton

(b) Pseudorapidity of the lepton

Figure 8.9: Basic kinematic distributions of the combination of muon and electron channels in the signal region. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the W+jets normalisation.

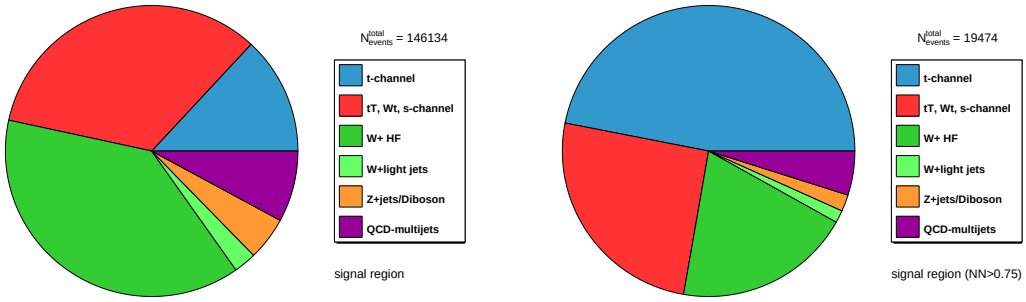
Table 8.7: Number of observed and expected events in the signal region. The data correction factors estimated in Ref. [71] are applied for all the top processes. The uncertainties shown are derived using the statistical uncertainty. The uncertainty on the QCD-multijets is fixed to 50%.

	electron channel			muon channel
Process	barrel	end-cap	total	total
t-channel	6966 ± 31	1440 ± 15	8406 ± 34	10751 ± 40
$t\bar{t}$	15414 ± 63	3341 ± 30	18755 ± 70	23324 ± 80
s-channel	355 ± 4	120 ± 2	475 ± 5	630 ± 5
Wt-channel	2179 ± 37	435 ± 17	2614 ± 41	3133 ± 46
W+jets (b)	8050 ± 47	3815 ± 30	11865 ± 56	20233 ± 81
W+jets (c)	7070 ± 110	2021 ± 56	9090 ± 130	14617 ± 170
W+jets (light)	1058 ± 87	419 ± 50	1480 ± 100	2170 ± 140
Z+jets/Diboson	2669 ± 79	936 ± 31	3605 ± 85	3430 ± 48
QCD-multijets	2300 ± 1100	3300 ± 1600	5500 ± 2800	6000 ± 3000
Total expected	46000 ± 1200	15800 ± 1600	61800 ± 2800	84300 ± 3000
data	45279	15467	60746	82632
bkgd. frac. r_{MC} [%]	45.9 ± 2.2	66.2 ± 9.1	51.1 ± 3.9	55.1 ± 3.1

CHAPTER 9

Neural network based event selection

In Figure 9.1(a) the fractions of the different processes in the signal region are shown. The largest contribution is W+heavy flavour production. As there are no real top quarks included in these processes these events cannot contribute to the measurement of m_{top} and a way to reduce this amount of background has to be found. Also the contribution from $t\bar{t}$ events is still dominating over the t-channel production.



(a) Signal region

(b) Signal region + $NN > 0.75$

Figure 9.1: Fractions of the different processes in the signal region before (a) and after (b) cutting on the neural network output distribution for the combined electron+muon channel.

To enhance the fraction of t-channel events multivariate techniques based on the Neurobayes package [177, 178] are applied, selecting only a subset of events out of the signal region. The resulting fractions of the different processes after this selection are shown in Figure 9.1(b) and the applied cut will be explained and justified in this chapter. As will be shown, it is possible to select a sample with almost 75% of the events from processes including top quarks with about two thirds of those events stemming from single-top

t-channel production. This is the reason for calling the selected sample enhanced with single top-quarks produced in the t-channel.

9.1 Reconstruction of the $W^{\pm}$-boson and the top quark

The $W^{\pm}$-boson can be reconstructed from its decay products being the charged lepton and the neutrino. The top quark can then be reconstructed by combining the $W^{\pm}$-boson with the b-tagged jet that is treated as the b-jet from the top decay.

The charged lepton, either an electron or muon, can be reconstructed with very high precision as it is described in Sections 5.3 and 5.4. For the neutrino only the x- and y-components of the four-momentum are known by the measurement of the missing transverse momentum. However, it is possible to exploit that the neutrino was created together with the charged lepton in the $W^{\pm}$-boson decay and the invariant mass of the combined four-momentum of the two should correspond to the mass of the $W^{\pm}$-boson. This means that by solving the following quadratic equation the z-component of the neutrino momentum $P_{z,\nu}$ can be reconstructed [179]:

$$P_{z,\nu}^2 - 2 \cdot \frac{\mu \cdot P_{z,\ell}}{E_\ell^2 - P_{z,\ell}^2} \cdot P_{z,\nu} + \frac{E_\ell^2 \cdot P_{T,\nu}^2 - \mu^2}{E_\ell^2 - P_{z,\ell}^2} = 0,$$
$$\text{with} \quad \mu = \frac{m_W^2}{2} + \cos\Delta\Phi\left(\ell, E_{\mathrm{T}}^{\mathrm{miss}}\right) \cdot P_{T,\ell} \cdot P_{T,\nu}.$$

The reconstruction algorithm was already developed at the Tevatron and also applied by CMS. It is explained in more detail in Ref. [179]. What has to be taken into account is that either due to the limited detector resolution or no presence of a real $W^{\pm}$-boson in the event the solution of $P_{z,\nu}$ can become imaginary. This translates into a transverse mass of the $W^{\pm}$-boson[57] that becomes larger than the pole mass $m_W = 80.4$ GeV. To avoid this unphysical behaviour a fit is applied that modifies also the x- and y-component of the neutrino momentum. The fit ensures that a real solution of $P_{z,\nu}$ is obtained. In addition, the difference between the transverse momentum of the reconstructed neutrino $P_{T,\nu} = \sqrt{P_{x,\nu}^2 + P_{y,\nu}^2}$ and the measured missing transverse momentum $E_{\mathrm{T}}^{\mathrm{miss}}$ is minimised [179, 180].

With the reconstructed neutrino the four-momentum of the $W^{\pm}$-boson can be calculated as $P_W^\mu = P_\nu^\mu + P_\ell^\mu$. The four-momentum of the top quark is obtained by also

[57] The definition of $m_{\mathrm{T}}(W)$ is given in Section 8.1.1.

adding the reconstructed four-momentum of the b-tagged jet. This means that P^{μ}_{top} is given by $P^{\mu}_{\text{top}} = P^{\mu}_{W} + P^{\mu}_{b} = P^{\mu}_{\nu} + P^{\mu}_{\ell} + P^{\mu}_{b}$.

After this step the event topology is fully reconstructed, making it possible to calculate various variables to discriminate the signal from background processes.

9.2 Discriminating input variables

The full reconstruction of the event topology is described in the previous section and a variety of variables can be derived from the reconstructed objects. The different kind of variables will be explained in this section. In the end these variables will be combined in one powerful discriminant that is based on a neural network.

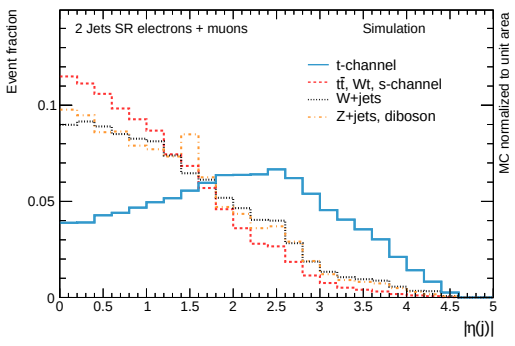
The goal is to select a dataset that is dominated by single-top quark production in the t-channel to measure the mass of the top quark. However, events including the production of top quarks like $t\bar{t}$ production and single-top quark production in the Wt-channel and s-channel of course also contain information about the top quarks mass. Therefore, they are treated as signal in the final top quark mass measurement described in Chapter 10.

In the neural network training only events from t-channel production are treated as signal and W+jets, Z+jets and diboson processes are considered as background. To enhance the fraction of top quark processes in the signal region, the $t\bar{t}$ process is not included in the training. The reason for that is that some of the input variables considered do discriminate well between the t-channel production and other processes, but not among $t\bar{t}$, Wt-channel and s-channel production and processes not containing top quarks like W+jets production. Also, only the MC simulated events are used, thus, the QCD-multijet model is not included in the training.

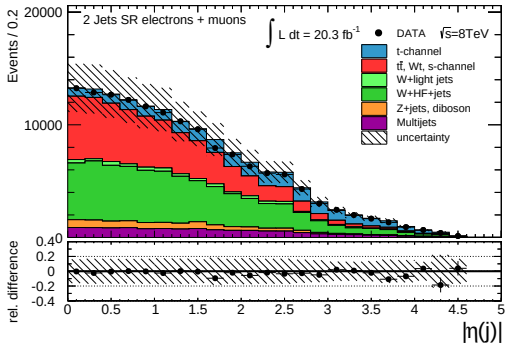
9.2.1 Kinematic variables

A basic kinematic variable separating well between the t-channel signal and background processes is the pseudorapidity of the light untagged jet, $|\eta(j)|$. This jet being reconstructed in the forward direction of the detector is typical for single top-quarks produced in the t-channel and separates well from other processes as can be seen in Figure 9.2(a).

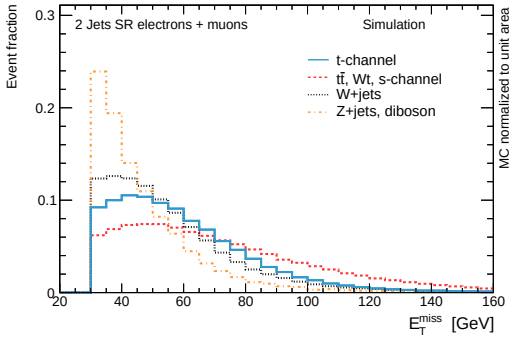
Events not containing neutrinos like Z+jets events have less missing transverse energy $E_{\text{T}}^{\text{miss}}$ than events where neutrinos are expected. Also W+jets processes typically have lower $E_{\text{T}}^{\text{miss}}$ than events containing top quarks due to the larger mass of the top quark compared to the $W^{\pm}$-boson, see Figure 9.2(c).



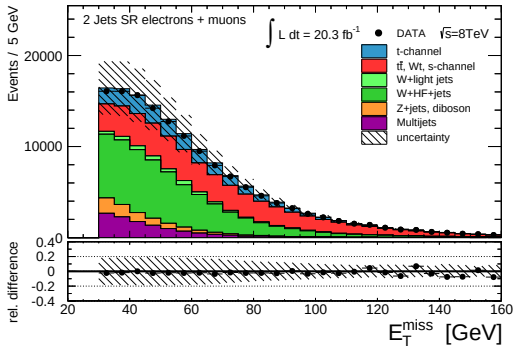
(a) Pseudorapidity of the reconstructed light jet



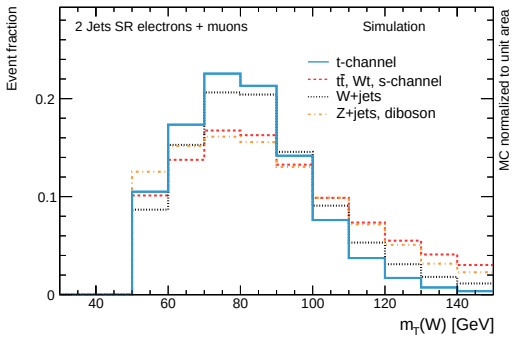
(b) Pseudorapidity of the reconstructed light jet



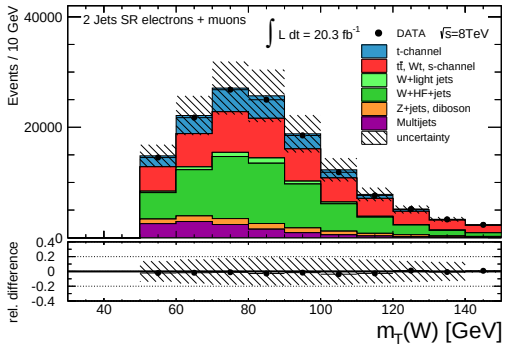
(c) Missing transverse momentum



(d) Missing transverse momentum

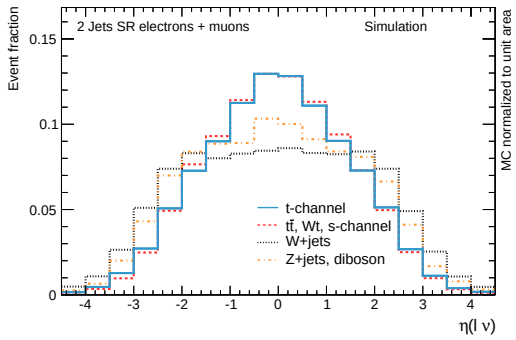


(e) Transverse $W^{\pm}$-boson mass

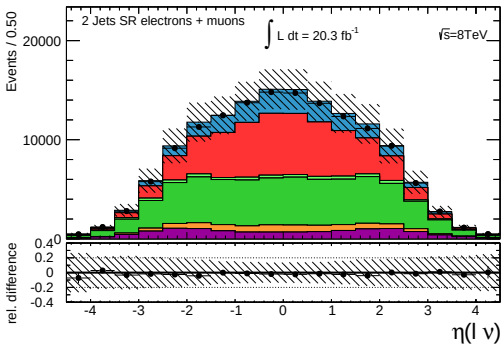


(f) Transverse $W^{\pm}$-boson mass

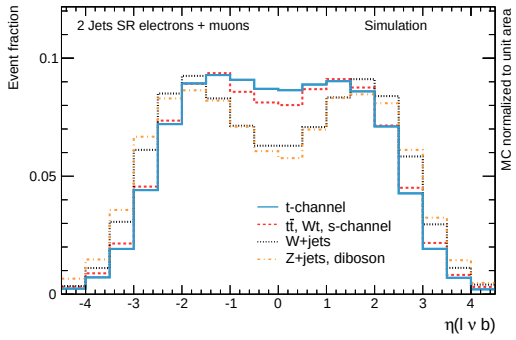
Figure 9.2: Distributions of the input variables used in the neural network selection normalised to unit area (left). Simulated and observed distributions of the input variables used in the neural network selection in the combined electron+muon channel in the signal region. The normalisation given in Table 8.7 and a top quark mass of $m_{\text{top}} = 172.5$ GeV in simulation is used. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the W+jets normalisation.



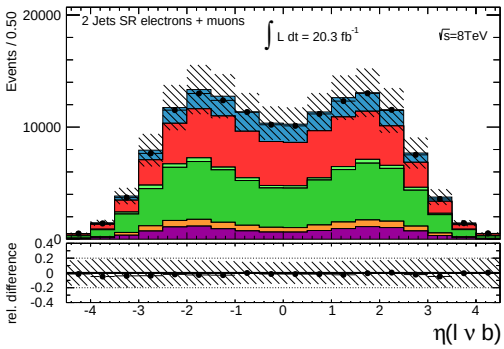
(a) Pseudorapidity of the reconstructed $W^{\pm}$-boson



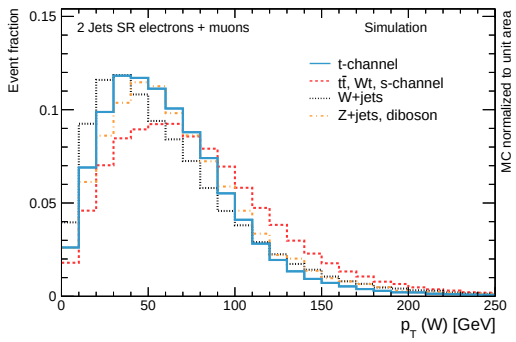
(b) Pseudorapidity of the reconstructed $W^{\pm}$-boson



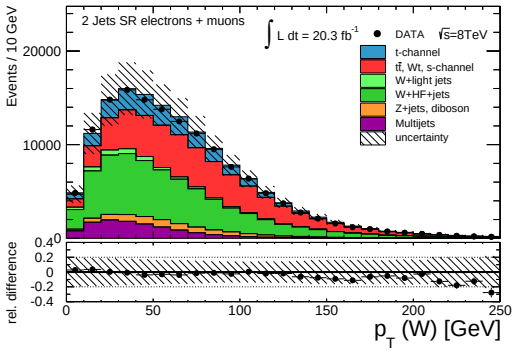
(c) Pseudorapidity of reconstructed top quark



(d) Pseudorapidity of reconstructed top quark



(e) Transverse momentum of the $W^{\pm}$-boson



(f) Transverse momentum of the $W^{\pm}$-boson

Figure 9.3: Distributions of the input variables used in the neural network selection normalised to unit area (left). Simulated and observed distributions of the input variables used in the neural network selection in the combined electron+muon channel in the signal region. The normalisation given in Table 8.7 and a top quark mass of $m_{\text{top}} = 172.5$ GeV in simulation is used. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the W+jets normalisation.

The transverse mass of the $W^{\pm}$-boson, $m_{\mathrm{T}}(W)$ plays a minor role in the multivariate separation since a cut at $m_{\mathrm{T}}(W) > 50$ GeV is already applied in the basic event selection explained in Section 8.1.1. However, processes not containing real $W^{\pm}$-bosons like Z+jets processes have a broader $m_{\mathrm{T}}(W)$ distribution and can be separated from the signal in this way. This is clearly visible in Figure 9.2(e).

More information to discriminate between the signal and background can be taken from the pseudorapidity of the reconstructed $W^{\pm}$-boson, $\eta(\ell\nu)$, shown in Figure 9.3(a) and the reconstructed top quark, $\eta(\ell\nu b)$, shown in Figure 9.3(c). While the $W^{\pm}$-boson is reconstructed more often in the central region of the detector in processes containing top quarks the distribution of $\eta(\ell\nu)$ is more flat for the other processes. This behaviour is propagated to the $\eta(\ell\nu b)$ variable where the pseudorapidity of the b-jet adds discriminating information[58].

The transverse momentum of the reconstructed $W^{\pm}$-boson, $p_T(W)$, is displayed in Figure 9.3(e). One can see that again due to the high mass of the top quark processes containing top quarks peak at larger values.
The distributions of the mentioned variables in the signal region using the normalisation given in Table 8.7 are shown in Figures 9.2((b), (d), (f)) and Figures 9.3((b), (d), (f)). Excellent agreement between the data and simulation can be seen.

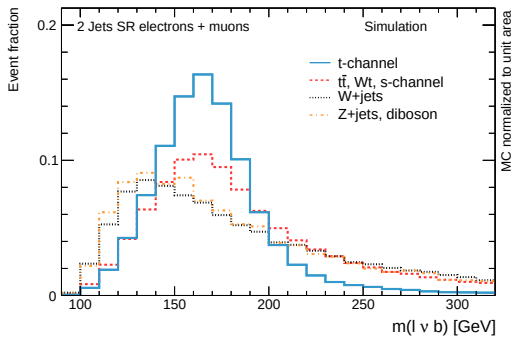
9.2.2 Invariant mass and event topology variables

The variables related to reconstructed invariant masses or the event topology are providing the best separation between the signal and the background processes.

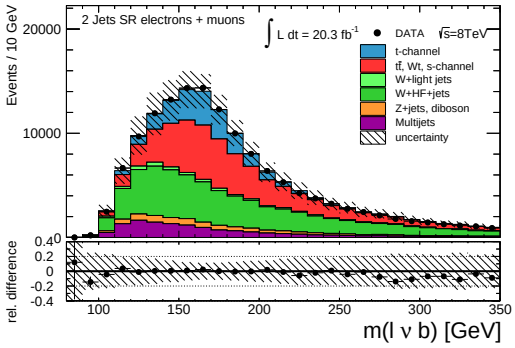
The invariant mass of the reconstructed top quark, $m(\ell\nu b)$, is shown in Figure 9.4(a). A peak around the true mass used in simulation at $m_{\mathrm{top}} = 172.5$ GeV can be seen for the t-channel while the distribution for the other processes including top quarks is broader with a large tail towards higher values. This is a direct consequence of the better resolution due to the unambiguous jet assignment to the final state partons in the t-channel. For processes without true top quarks the distribution also peaks at lower values with a large tail.

The $m(\ell b)$ variable, being the invariant mass of the reconstructed top quark without the neutrino component, behaves similar as $m(\ell\nu b)$. Again a sharp peak is seen in Figure 9.4(c) for the t-channel, while the other dominant processes have a worse resolution.

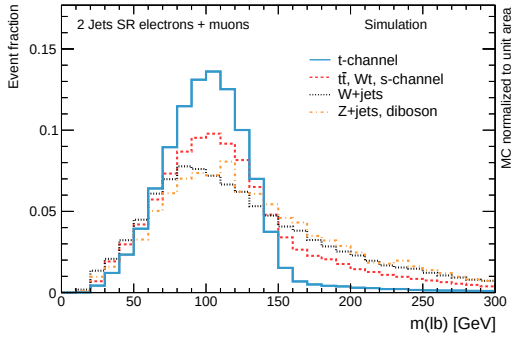
[58] In this case the discriminating information from the pseudorapidity of the b-jet could also be added to the neural network with $\eta(b)$ instead of $\eta(\ell\nu b)$. The overall performance does not vary much if using either the one or the other variable.



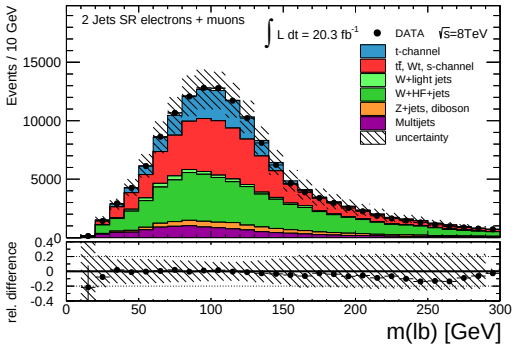
(a) Invariant mass of the reconstructed top quark



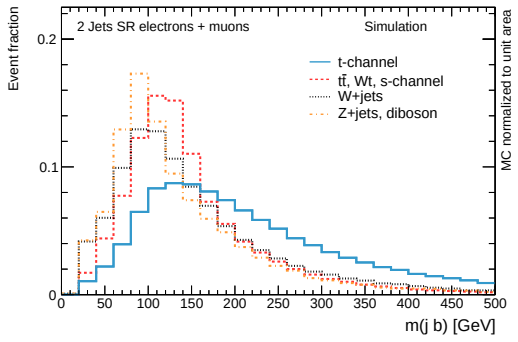
(b) Invariant mass of the reconstructed top quark



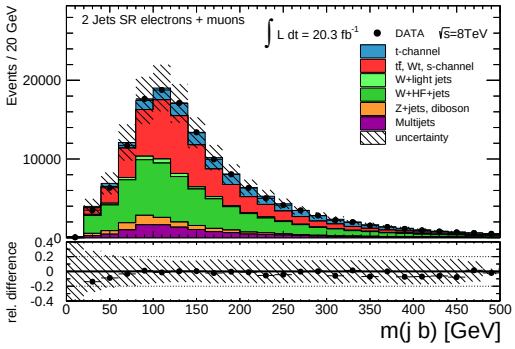
(c) Invariant mass of the lepton and b-tagged jet



(d) Invariant mass of the lepton and b-tagged jet

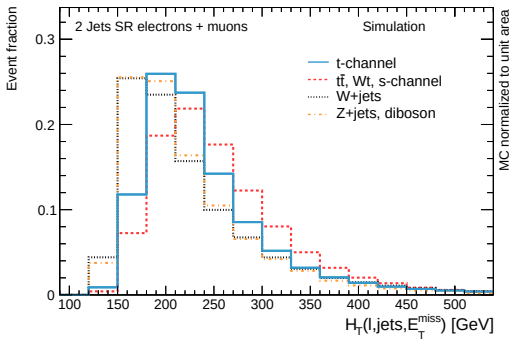


(e) Invariant mass of the two reconstructed jets

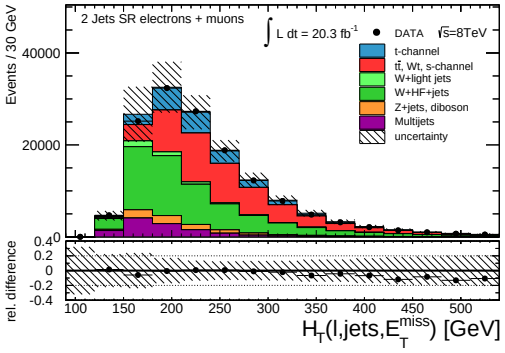


(f) Invariant mass of the two reconstructed jets

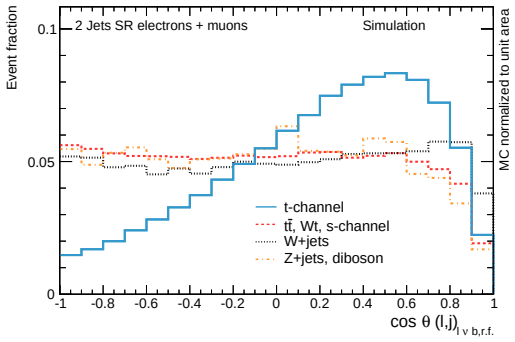
Figure 9.4: Distributions of the input variables used in the neural network selection normalised to unit area (left). Simulated and observed distributions of the input variables used in the neural network selection in the combined electron+muon channel in the signal region. The normalisation given in Table 8.7 and a top quark mass of $m_{\text{top}} = 172.5$ GeV in simulation is used. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the W+jets normalisation.



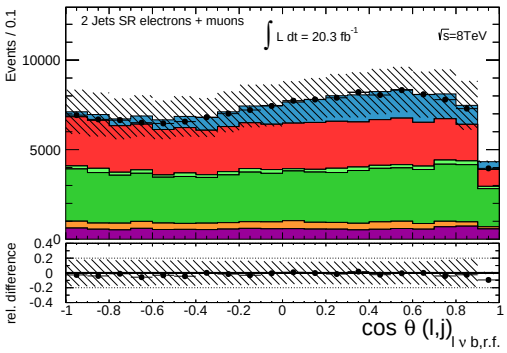
(a) Summed transverse momentum of all reconstructed objects in the event



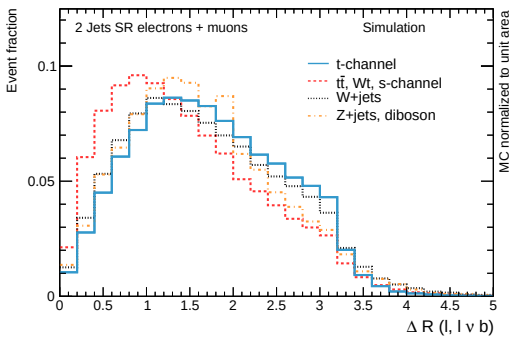
(b) Summed transverse momentum of all reconstructed objects in the event



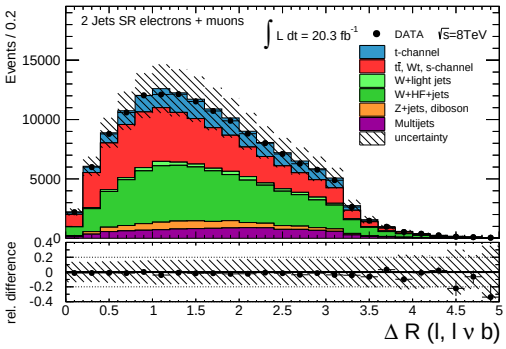
(c) Polarisation angle of the top quark



(d) Polarisation angle of the top quark



(e) ΔR between the top quark and the lepton



(f) ΔR between the top quark and the lepton

Figure 9.5: Distributions of the input variables used in the neural network selection normalised to unit area (left). Simulated and observed distributions of the input variables used in the neural network selection in the combined electron+muon channel in the signal region. The normalisation given in Table 8.7 and a top quark mass of $m_{\text{top}} = 172.5$ GeV in simulation is used. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the W+jets normalisation.

Another sensitive variable is the invariant mass of the two reconstructed jets called $m(jb)$ shown in Figure 9.4(e). For the t-channel the two jets are not directly related to e.g. a decay of a particle. Instead, as explained in Section 9.2.1, the light-flavour-jet is often reconstructed in the forward region of the detector, while the b-tagged jet is reconstructed more central. This separates well from W+jets events where the two jets are often produced by the decay of a gluon from initial state or final state gluon radiation.

$H_{\mathrm{T}}(\ell, \mathrm{jets}, E_{\mathrm{T}}^{\mathrm{miss}})$ is the sum of the transverse momenta of all reconstructed objects within the event and can be interpreted as the total transverse energy of the collision process. To produce a top quark in the t-channel one needs at least the energy to produce the top quark itself plus the light-flavour-jet from the t-channel production mode, corresponding to a momentum of about 200 GeV. Analogue for the W+jets background one needs the energy to produce a real $W^{\pm}$-boson plus two additional jets corresponding to at least 140 GeV while in the $t\bar{t}$ pair production the needed energy is higher. This is clearly visible in Figure 9.5(a).

In the Figures 9.4((b), (d), (f)) and 9.5(b) the observed and expected distributions in the signal region are displayed showing very good agreement between data and simulation.

9.2.3 Angular correlation variables

The variable denoted as $\cos\Theta(\ell, j)_{\ell\nu b, r.f.}$ is the cosine of the angle between the charged lepton and the untagged jet in the rest frame of the reconstructed top quark. This variable illustrates the polarisation of the top quark. In the single-top t-channel production the top quark is expected to be produced polarised due to V-A structure of the weak interaction [181]. Thus, this variable shown in Figure 9.5(c) can be used as a separating variable.

$\Delta R(\ell, \ell\nu b)$ is defined as the distance between the reconstructed top quark and the charged lepton from the top quark decay in the η-φ-plane. The variable is sensitive to the boost of the top quark meaning that the larger the boost the more collinear the two reconstructed objects are. This is clearly visible in Figure 9.5(e) when comparing the more boosted top quarks in $t\bar{t}$ with top quarks in the t-channel. Processes not containing real top quarks are in between the two scenarios.

A comparison of the observed and well-modelled expected distributions in the signal region is shown in Figure 9.5((d), (f)).

9.3 Variable preprocessing

The training of the neural network, detailed in Section 9.5, involves a multi-dimensional minimisation of high complexity. The Neurobayes package uses a complex stepwise preprocessing of the input variables. With this procedure it obtains a better starting point for the minimisation and possible issues related to input variables are avoided. Two kinds of preprocessing are applied. The global preprocessing is defined once for all input variables and the individual preprocessing can be defined independently for each input variable. Both preprocessings are implemented within Neurobayes. The neural network used in this thesis is configured to use the following robust methods automatically [177].

9.3.1 Global preprocessing

Global preprocessing steps are performed for every input variable and defined globally before the training.

Flattening The input variables are transformed such that the cumulative signal and background distribution becomes a flat distribution. In practice this is done by filling a histogram with variable bin widths. This significantly reduces the influence of possible outliers that could otherwise saturate the output of an individual node or the network output.

Gaussian transformation The flattened distributions are transformed to a standard Gaussian. This has two different advantages. Firstly, all input variables have the same dynamic range leading to a better numerical performance during the later training process. Secondly, it ensures optimal learning conditions from the very beginning of the training.[59]

Ranking The performance of the neural network of course depends very much on the choice of the input variables. It is possible that the overall performance is worsened if variables are added that do not add significant information to discriminate between signal and background. To avoid this, the input variables are automatically ranked according to their significance with the following procedure.

[59]The training process starts with random weights also distributed according to a Gaussian centred around zero and width $\sigma = 1/\sqrt{n_{in}}$, with n_{in} being the number of incoming weights of a neuron. Together with the Gaussian input variables also the output of the hidden layer nodes will be Gaussian before they are passed through the activation function. Thus, the whole network will be in a well-defined condition. In addition, values not close to zero would imply large eigenvalues of the Hessian in Equation (9.7) limiting the initially allowed learning rate [182].

Firstly, the $N+1$-dimensional correlation matrix is calculated for the N input variables and the target variable[60]. Secondly, the total correlation of the input variables to the target is calculated [183]. Then a variable is removed and the correlation matrix plus the remaining total correlation to the target is recalculated. This is done independently for each variable and the full list of input variables is reordered according to the loss of total correlation to the target caused by the removal of the respective variable. From this new list the least significant variable is removed leading to a subset of $N-1$ input variables and an N-dimensional correlation matrix. The procedure is repeated iteratively until only the one most significant variable remains.

Different quantities can be derived to rank the correlation of the input variables.

- The loss of correlation estimated with the iterative procedure causing the variable to be removed is called ρ^{iter}.
- The correlation of only the one variable under study to the target is called ρ^{only}. This quantity is not influenced by any other variable.
- The loss in total correlation of the input variables to the target caused by the removal of the one variable is defined as ρ^{loss}.
- From the N-dimensional correlation matrix of the N input variables only the correlation of each variable to all the other variables is calculated. This quantity is denoted as κ.

The correlation coefficients, ρ_i, are directly related to the significance via $\sigma_i = \rho_i \cdot \sqrt{n}$ where n corresponds to the sample size used in the training. Variables with $\sigma^{\text{iter}} < 3$ are automatically removed and a minimal subset of input variables is kept for the neural network.

Decorrelation A neural network is able to take the correlations between input variables into account. However, a better performance can be achieved if the input variables would be uncorrelated from the beginning. The linear decorrelation is done after estimating the ranking by calculating the covariance matrix of the Gaussian input variables. It is then diagonalised with iterative Jacobi rotations [184]. After that the correlation matrix of the transformed input variables finally used in the neural network is close to a unit matrix, i.e. variables are mostly uncorrelated.

[60]The binary target variable holds the information whether an event is signal or background.

9.3.2 Individual preprocessing

It is possible that different kinds of input variables are preprocessed in a different way before the ranking step of the global preprocessing[61]. In practice, all used input variables are of the same kind and the individual preprocessing step is the same for all variables.

From the flattened distribution of variable i in each bin j the probability $P(x_{i,j}|\text{signal})$ of an event being either signal or background is calculated. This probability corresponds to the purity or mean value of the target defined in Equation (9.1):

$$P(x_{i,j}|\text{signal}) = \frac{n_{i,j}^{\text{signal}}}{n_{i,j}^{\text{signal}} + n_{i,j}^{\text{background}}}. \tag{9.1}$$

$n_{i,j}^{\text{signal}}$ corresponds to the number of signal events of variable i in bin j while $n_{i,j}^{\text{background}}$ is the same for the background events. To reduce statistical fluctuations the resulting distribution is fitted with a spline [184].

Instead of using the initial variable a new variable is defined. This contains the information of the correlation to the target given by the regularised spline fit. More details about the complex procedure to transform the input variables can be found in Ref. [177, 178].

9.4 Neural network architecture

An artificial neural network is inspired by the human brain and consists of neurons (nodes) that are arranged in layers. Neurons in one of the layers are connected by edges (synapses) to the neurons of the adjacent layer.

In this special case the neural network is a feed-forward network with three layers. This means that neurons do only influence the neurons of the following layer. The first layer represents the input variables and its output creates the input of the neurons in the hidden layer. Furthermore, the output of the hidden layer neurons is used to calculate the value of the neurons of the third layer[62].

The first layer consists of 13 nodes for the 12 input variables plus one bias node with a constant value of 1. The hidden layer consists of 15 nodes, while only one node is used to obtain the output of the network in the output layer. The strength of the connections

[61] E.g. a continuous variable like the jet momentum should be treated differently compared to a discrete variable like the number of jets.

[62] One could think about adding a second hidden layer. However, it was shown that one hidden layer in combination with sigmoid activation functions is enough to approximate any multivariate function with arbitrary precision [185].

between neuron i and j corresponds to the weight w_{ij}. The described neural network architecture is shown schematically together with the variable preprocessing, explained in Section 9.3, in Figure 9.6.

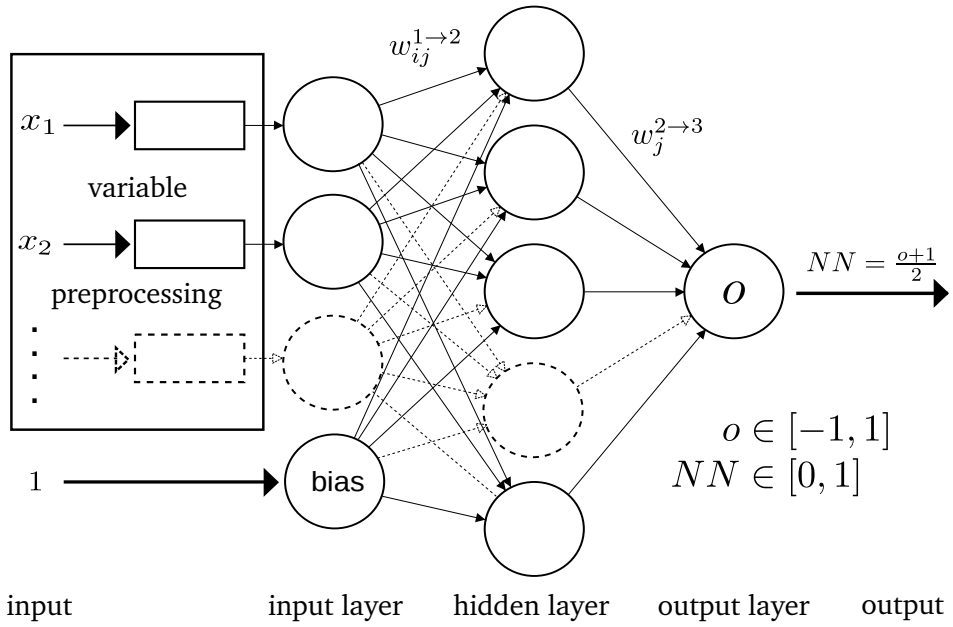


Figure 9.6: Sketch of the feed-forward neural network with three layers. The input variables are preprocessed, as described in Section 9.3, and passed to the input layer neurons. The dashed lines represent additional neurons and edges that are integrated in the neural network analogical to the other shown neurons and edges.

The value of a neuron j of the hidden layer is calculated as the weighted sum of all input variables x_i following Equation (9.2):

$$a_j(\vec{x}) = \sum_{i=1}^{13} w_{ij}^{1\to2} x_i + \mu_j. \tag{9.2}$$

μ_j is evaluated from the strength of the connection with the bias node and can be used to shift the weighted sum as visualised in Figure 9.8. As the range $a_j(\vec{x})$ is not restricted it is passed to the activation function of the node which is mapping the output of the

node from $[-\infty, \infty]$ to $[-1, 1]$. The activation function used in Neurobayes is the shifted sigmoid function $S(a_j)$ defined in Equation (9.3):

$$S(a_j) = \frac{2}{1 + \mathrm{e}^{-a_j}} - 1. \tag{9.3}$$

This function is sensitive for values around zero and saturates for large positive or negative values. That is why the shift induced by the bias node and μ_j is necessary.

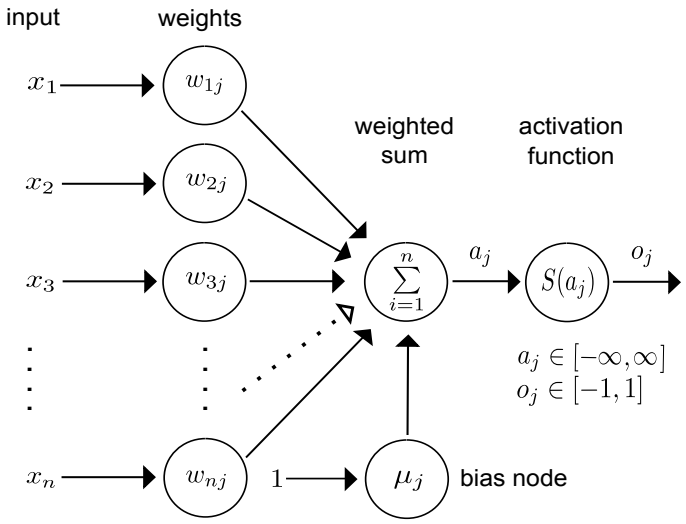


Figure 9.7: Data processing at neuron j. In the special case of j being the one output node of the neural network it would be $o_j = o$.

The data flow at one neuron j is shown in Figure 9.7 and the sigmoid activation function is depicted in Figure 9.8. The function becomes steeper in its sensitive region for larger input weights and is shifted to its sensitive region by the weight connected to the bias node.

In the end the output o of the neural network is again the weighted sum of the output of the hidden nodes passed through a sigmoid activation function:

$$o = S\left[\sum_{j=1}^{15} w_j^{2\to 3} S\left(\sum_{i=1}^{13} w_{ij}^{1\to 2} x_i + \mu_j\right)\right]. \tag{9.4}$$

To interpret the output of the neural network as a measure of a probability, o is transformed with Equation (9.5) to be within the range $[0, 1]$:

$$NN = \frac{1+o}{2}. \tag{9.5}$$

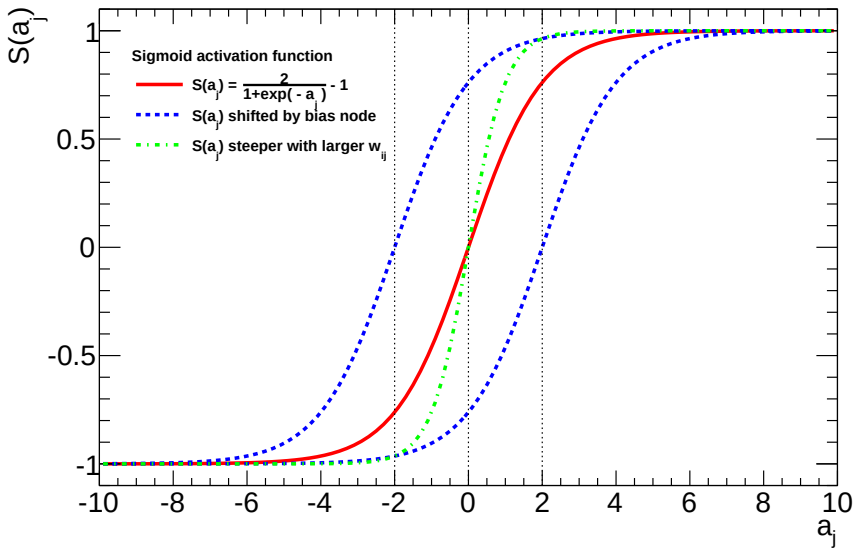


Figure 9.8: Sigmoid activation function defined in Equation (9.3). The function is shifted by the bias node and steeper in its sensitive region if the input weights are larger.

9.5 Training of the neural network

The training of the neural network relies on a training sample from simulation where the target value is known. The goal of the training procedure is to minimise the difference between the predicted output of the network and the true target value. In Neurobayes this is quantified by the entropy loss function[63] which is defined based on the sum of n events in the training sample:

$$E_D = -\sum_{k=1}^{n} \log\left[\frac{1}{2}\left(1 + T_K \cdot o_K + \varepsilon\right)\right]. \tag{9.6}$$

[63] One advantage of the entropy loss function is that completely wrong classifications like $T_K = 1$ and $o_K = -1$ have an infinite strong influence on the error and are learned by the network very fast. This behaviour is not present e.g. if using a χ^2-function.

In this function $T_K \in [-1,1]$ is the target value of each event while $o_K \in [-1,1]$ is the network output. ε is a constant used for regularisation decreasing with each training iteration. It avoids saturation effects at early stages of the training where the weights are initialised randomly.

In principle the training process corresponds to a minimisation problem of E_D. However, in practice this is a highly non-trivial task with caveats like finding local minima instead of the global minimum. Therefore, the training is done by applying the back-propagation algorithm [186]. In this algorithm the weights of the synapses are modified by propagating E_D backwards from the output layer to the input layer. The change of the i^{th} input weight of neuron j is Δw_{ij} and depends on the influence of the weight on E_D. This is quantified by the partial derivative:

$$\begin{aligned}\Delta w_{ij} = -\eta \frac{\partial E_D}{\partial w_{ij}} &= -\eta \frac{\partial E_D}{\partial o_j}\frac{\partial o_j}{\partial a_j}\frac{\partial a_j}{\partial w_{ij}} \\ &= -\eta \frac{\partial E_D}{\partial o_j} S'(a_j) x_i = -\eta \delta_j x_i.\end{aligned} \tag{9.7}$$

η is a free parameter usable to steer the speed of the training. With a higher value of η the training is performed faster but with lower precision. The neurons' error signal δ_j distinguishes two different cases for neurons being either in the hidden or output layer:

$$\delta_j = S'(a_j) \cdot \frac{\partial E_D}{\partial o_j} = S'(a_j) \cdot \frac{-T_j}{1 + T_j o_j + \varepsilon} \quad \text{(if } j \text{ is output neuron)}, \tag{9.8}$$

$$\delta_j = S'(a_j) \cdot \sum_k \delta_k w_{jk} \quad \text{(if } j \text{ is hidden neuron)}. \tag{9.9}$$

In case of neurons of the output layer the influence on E_D is taken into account directly. For a hidden layer neuron the signal error of its k connected nodes in the following layer is propagated. Thus, the influence of the neuron is taken indirectly.

The modification of the weights can be done in different ways and in this particular case the weight update is done every 100 events during the training with $w_{ij}^{\text{new}} = w_{ij}^{\text{old}} + \Delta w_{ij}$.

9.5.1 Training with momentum correction

Equation (9.7) can be modified leading to Equation (9.10) that includes the amount of weight change in the previous iteration:

$$\Delta w_{ij} = -\eta \delta_j x_i + \alpha \Delta w_{ij}^{\text{previous}}. \tag{9.10}$$

The second, so called momentum term, can be useful to avoid getting stuck in either steep valleys or flat plateaus of the multi-dimensional plane.

It was found that the momentum term only has a minor influence on the neural network used in this thesis. Therefore, the momentum parameter α is set to zero.

9.5.2 Bayesian regularisation and significance control

One issue that can appear when using a neural network is overtraining. This means that the network has learned to identify statistical fluctuations instead of the general underlying structure. In Neurobayes complex techniques of Bayesian regularisation are used that ensure a good generalisation ability of the network. During the training not only E_D but also the weights themselves are minimised in parallel. Insignificant synapses or neurons are removed automatically leading to a minimal set of nodes and connections after a successful training. Details about Bayesian regularisation techniques are given in Ref. [187].

9.5.3 Intrinsic overtraining test

To check for a possible overtraining the training sample is split up in the actual training sample and a test sample with lower statistics. 20% of the events are used as a test

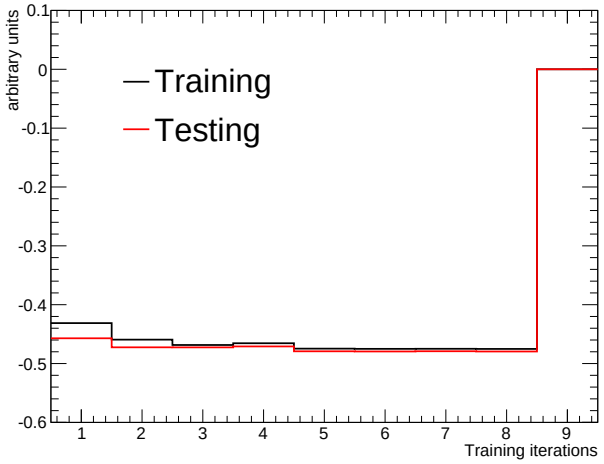


Figure 9.9: Entropy loss error function during the training of the neural network. No overtraining is present and the training finishes after 8 iterations indicated by the error fixed to a value of 0.

sample while the 80% remain to perform the training. For both samples the entropy loss error function is calculated after each iteration of the training. This is shown in Figure 9.9. A possible overtraining would be visible by an increase of the error on the test sample after an iteration. In this case the training finishes after 8 iterations and no indication for an overtraining is seen.

9.6 Final neural network selection

The final neural network used to enhance the signal region with topologies of single-top t-channel events uses the input variables that have been explained in Section 9.2. It is expected that the shape of some input variables depends on the mass of the top quark. It was validated that no bias in the mass estimation is introduced due to the neural network based event selection. This is explained in Section 10.3.3.

In general, all input variables should be modelled well in the control regions. Appendix A.2 contains comparisons between the expected and observed distributions of the used variables. In both control regions very good agreement is found in the electron as well as the muon channel separately. Apart from a good modelling only variables that add significant discrimination between signal and background are considered. Their ranking estimated with the procedure explained in Section 9.3 is shown in Table 9.1. As one can see all of them add significant information to distinguish between signal and

Table 9.1: The 12 variables which are used in the training of the neural network ordered by their importance. The definitions of ρ_i, σ_i and κ are given in Section 9.3.

Variable	ρ^{iter} [%]	σ^{iter}	ρ^{only} [%]	σ^{only}	ρ^{loss} [%]	σ^{loss}	κ [%]
$m(\ell\nu b)$	37.89	134.12	37.89	134.12	14.78	52.29	67.8
$m(jb)$	30.82	109.06	31.37	111.00	12.18	43.10	69.6
$m(\ell b)$	17.50	61.91	36.26	128.31	13.41	47.47	68.7
$\lvert\eta(j)\rvert$	13.88	49.13	32.41	114.69	15.02	53.17	52.8
$\eta(l\nu)$	13.26	46.93	18.71	66.22	9.39	33.24	67.7
$H_T(l, jets, E_T^{\text{miss}})$	9.92	35.09	23.20	82.11	8.79	31.10	58.7
E_T^{miss}	7.37	26.09	12.78	45.22	8.00	28.23	35.8
$m_T(W)$	6.92	24.48	12.47	44.15	6.48	22.93	31.7
$\cos\theta(\ell, j)_{\ell\nu b, r.f.}$	6.24	22.09	21.91	77.55	5.59	19.79	48.4
$p_T(W)$	2.45	8.67	10.08	35.69	2.79	9.86	48.7
$\eta(l\nu b)$	1.58	5.58	12.74	45.10	1.41	4.98	76.6
$\Delta R(\ell, \ell\nu b)$	1.41	4.98	6.48	22.92	1.34	4.73	38.4

background. The correlation matrix calculated for all input variables initially used in the training and the target variable can be found in Appendix A.4. The setup of the neural network corresponds to the setup of the feed-forward network with three layers explained in Section 9.4.

The training was done with the procedure that is detailed in Section 9.5. Events selected in the signal region according to Table 8.7 apart from the QCD-multijet model are used as the training sample. The admixture of signal to background events in the training was chosen to be 50% signal and 50% background, where the different background processes are weighted according to their number of expected events.

In the training only events from t-channel production are treated as signal and W+jets, Z+jets and diboson processes are considered as background. Other top quark processes will be treated as signal in the later top quark mass measurement but are excluded from the training due to the reasons formulated in Section 9.2.

The resulting output distribution normalised to unit area is shown in Figure 9.10 for all processes. It is clearly visible that the single-top t-channel events are accumulating at higher values of the distribution.

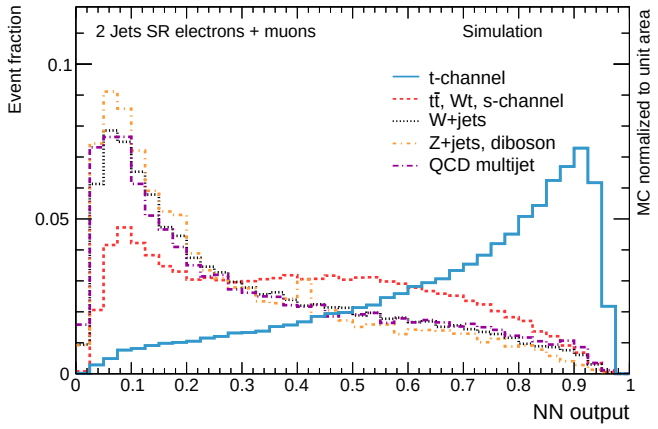
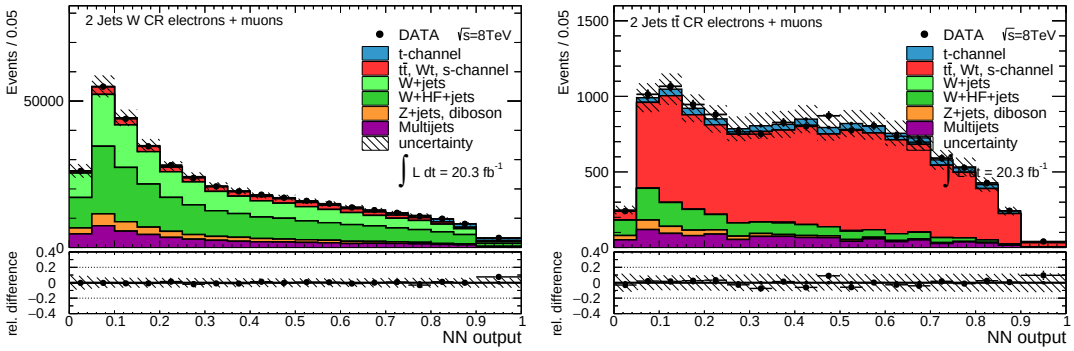


Figure 9.10: Distributions of the neural network output normalised to unit area showing all background processes.

Since the training of the neural network is done with simulated events, one has to check the modelling of the output distribution with observed events. This is tested in the two control regions defined in Section 8.1.2. In Figure 9.11(a) the neural network output distribution is shown in the $W^{\pm}$-boson control region and in Figure 9.11(b) for the $t\bar{t}$ control region. In both regions the normalisation obtained in Section 8.2 is used

and good agreement is seen. The neural network output distribution in the signal region is shown in Figure 9.12.



(a) $W^{\pm}$-boson control region

(b) $t\bar{t}$ control region

Figure 9.11: Comparison of the expected and the observed neural network output distributions in the two control regions. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the systematic uncertainties on the QCD-multijet normalisation and relative top normalisation. In simulation a top quark mass of $m_{\text{top}} = 172.5$ GeV is used.

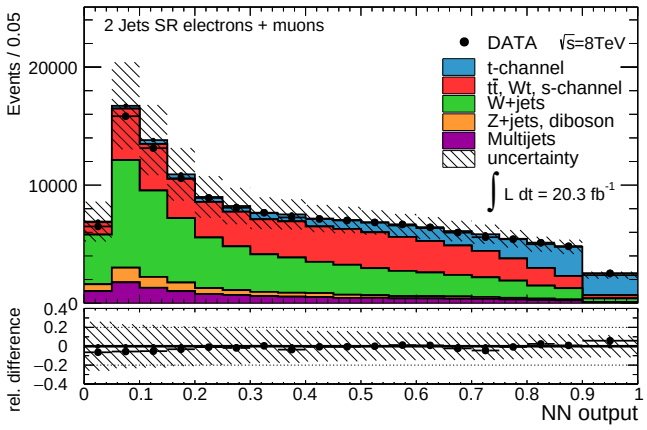
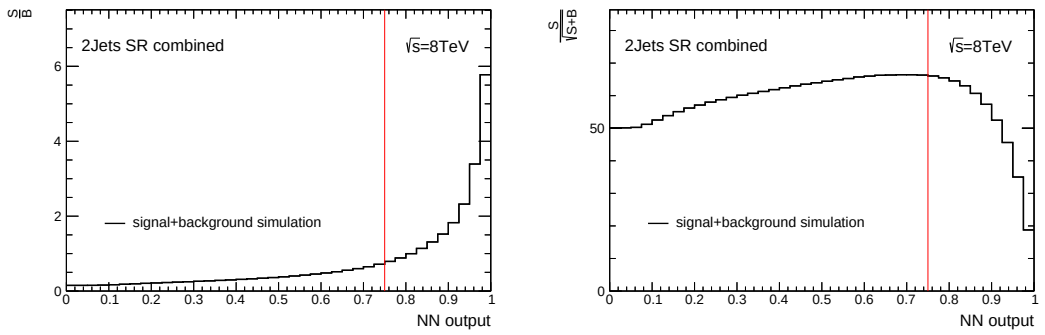


Figure 9.12: Simulated and observed output distributions of the neural network in the combined electron+muon channel in the signal region. The normalisation given in Table 8.7 and a top quark mass of $m_{\text{top}} = 172.5$ GeV in simulation is used. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the systematic uncertainties on the W+jets normalisation and relative top normalisation.

9.6.1 Cutting on the neural network output

A cut value of $NN > 0.75$ was chosen to enhance the signal purity for the later top quark mass analysis. The choice of the cut on the neural network was chosen to be the best possible compromise between a high signal purity and signal efficiency leading to the highest possible statistical precision and sensitivity.

The neural network is constructed in a way that the signal-to-background ratio is monotonically increasing as a function of the cut value. Therefore, the value of $\frac{S}{\sqrt{S+B}}$ is more useful as a classifier to optimise the cut value because it is a direct measure of the statistical precision[64]. In Figure 9.13 the ratios $\frac{S}{B}$ (a) and $\frac{S}{\sqrt{S+B}}$ (b) are shown as a function of the neural network cut showing that the chosen cut value maximises the latter ratio[65].



(a) $\frac{S}{B}$ vs neural network output cut

(b) $\frac{S}{\sqrt{S+B}}$ vs neural network output cut

Figure 9.13: Signal-to-background ratios as a function of the neural network cut showing that the chosen cut value maximises the respective ratio.

Studies with pseudo-experiments have shown that if the cut is raised from $NN > 0.75$ to e.g. $NN > 0.8$ the statistical uncertainty on the top quark mass measurement increases by about 100 MeV. This is mostly driven by low statistics that remains for the background sample described in Section 10.2.2. If the cut is lowered to e.g. $NN > 0.7$ the statistical precision is only slightly better but the non-top background fraction increases. The cut value at $NN > 0.75$ was chosen as the best balance between the two as it also

[64] $\frac{S}{\sqrt{S+B}}$ is a measure optimising the product of signal efficiency $\varepsilon_S = \frac{S}{X_S}$ and signal purity $p_S = \frac{S}{S+B}$. S stands for the number of selected events from processes including top quarks while B corresponds to the number of background processes without top quarks. X_S corresponds to the total number of events from processes containing top quarks. It is constant and does not influence $\frac{S}{\sqrt{S+B}}$ for different cut values.

[65] Actually the maximum is slightly lower than at 0.75 but still the observed performance was better with the little higher purity.

provides the highest possible value of $\frac{S}{\sqrt{S+B}}$. Systematic uncertainties only play a minor role since the dominant uncertainties from the jet energy scale are mostly independent from the cut on the neural network.

9.6.2 Event yields and kinematic modelling after cutting on the neural network output distribution

In Table 9.2 the respective event yields in the signal region after cutting at $NN > 0.75$ on the neural network output is summarised. The event yields are calculated using the acceptance from MC samples normalised to their respective theoretical cross-sections including the (N)NLO k factors and the correction factors obtained in Ref. [71]. The event numbers for QCD-multijets reflect the yields as derived from the binned maximum-likelihood fit in the $E_{\mathrm{T}}^{\mathrm{miss}}$-distribution as it is described in Section 8.2.

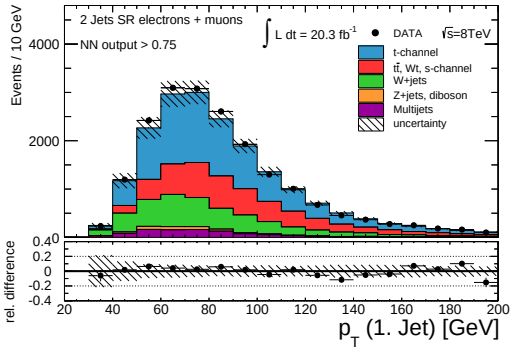
Table 9.2: Number of observed and expected events in the signal region after cutting at $NN > 0.75$ on the neural network output. The uncertainties shown are derived using the statistical uncertainty and the uncertainty on the theoretical cross-section. The uncertainty on the QCD-multijets is fixed to 50%.

	electron channel			muon channel
Process	barrel	end-cap	total	total
t-channel	3587 ± 22	495 ± 8	4082 ± 23	5055 ± 27
$t\bar{t}$	1811 ± 22	171 ± 7	1983 ± 23	2299 ± 25
s-channel	56 ± 2	14 ± 1	70 ± 2	92 ± 2
Wt-channel	222 ± 12	16 ± 3	238 ± 12	254 ± 13
W+jets (b)	632 ± 14	113 ± 5	745 ± 15	1108 ± 19
W+jets (c)	749 ± 38	58 ± 10	806 ± 39	1161 ± 49
W+jets (light)	117 ± 27	23 ± 12	140 ± 30	129 ± 26
Z+jets/Diboson	139 ± 7	26 ± 3	165 ± 8	194 ± 16
QCD-multijets	260 ± 130	132 ± 66	390 ± 190	570 ± 280
Total expected	7570 ± 140	1048 ± 69	8620 ± 200	10860 ± 290
data	7796	999	8795	11038
bkgd. fraction r_{MC} [%]	25.0 ± 1.6	33.5 ± 5.7	26.0 ± 2.1	29.1 ± 2.4

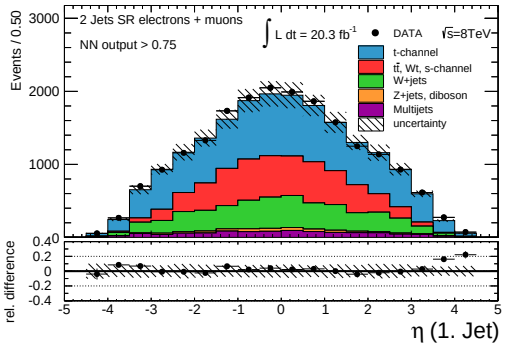
Good agreement between the expectation and the data can be seen. By applying the neural network selection it was possible to reduce the predicted background fraction from $r_{\mathrm{MC}}^{\mathrm{SR}} \approx 53\%$, see Table 8.7, to $r_{\mathrm{MC}}^{\mathrm{SR},NN>0.75} \approx 28\%$.

In Figures 9.14 and 9.15 the distributions of kinematic observables are shown in the signal region after cutting at $NN > 0.75$ normalised to the number of expected

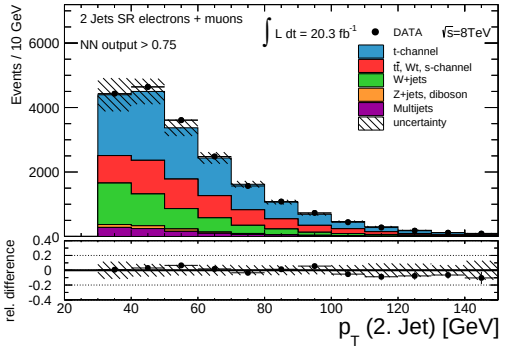
events according to Table 9.2. Overall a very good agreement between the simulated distributions and the observed distributions in data is seen.



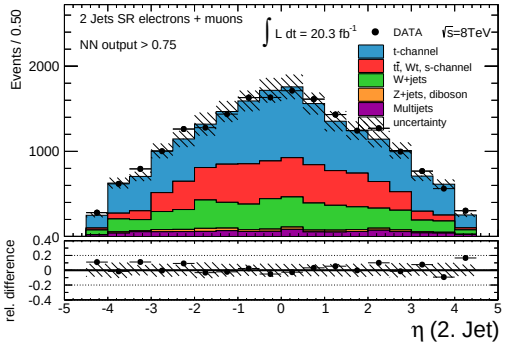
(a) Transverse momentum of the leading jet



(b) Pseudorapidity of the leading jet

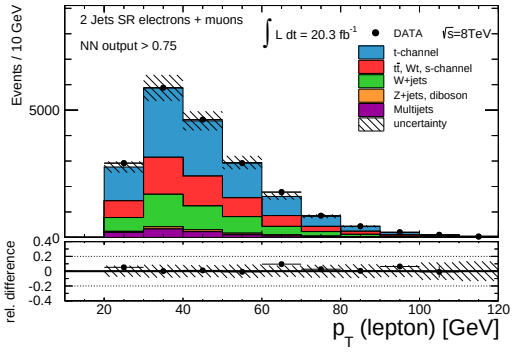


(c) Transverse momentum of the second leading jet

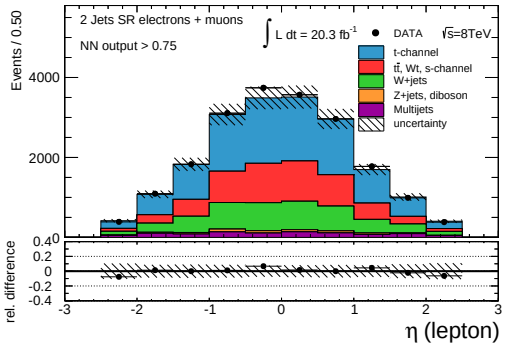


(d) Pseudorapidity of the second leading jet

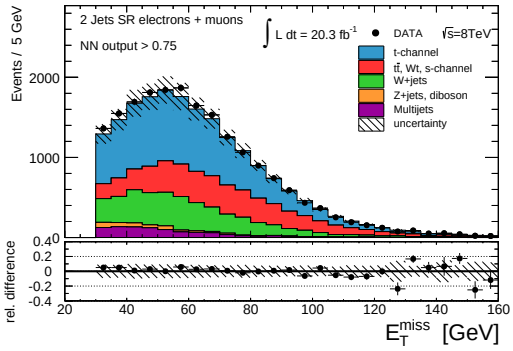
Figure 9.14: Basic kinematic distributions of the combination of muon and electron channels in the signal region after cutting at $NN > 0.75$. The normalisation given in Table 9.2 and a top quark mass of $m_{\text{top}} = 172.5$ GeV in simulation is used. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the W+jets normalisation.



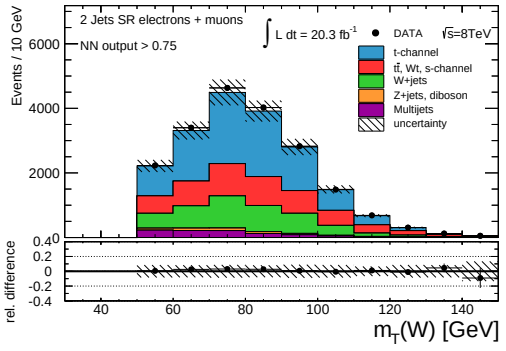
(a) Transverse momentum of the lepton



(b) Pseudorapidity of the lepton



(c) Missing transverse momentum



(d) Transverse W-boson mass

Figure 9.15: Basic kinematic distributions of the combination of muon and electron channels in the signal region after cutting at $NN > 0.75$. The normalisation given in Table 9.2 and a top quark mass of $m_{\text{top}} = 172.5$ GeV in simulation is used. The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the W+jets normalisation.

CHAPTER 10

Measurement of m_{top} in t-channel topologies

In this chapter the measurement of m_{top} in a sample selected with the neural network based selection explained in Chapter 9 will be discussed. In Section 10.2 the analysis method is described. It is using the $m(\ell b)$ variable which will be introduced in Section 10.1. The validation of the procedure is discussed in Section 10.3 and the discussion about systematic uncertainties is included in Section 10.4.

The measurement is done in the electron, muon and combined electron+muon channel. The combination is done by adding together the two disjoint contributions and repeating all analysis steps in the same way as it is done for the separate channels.

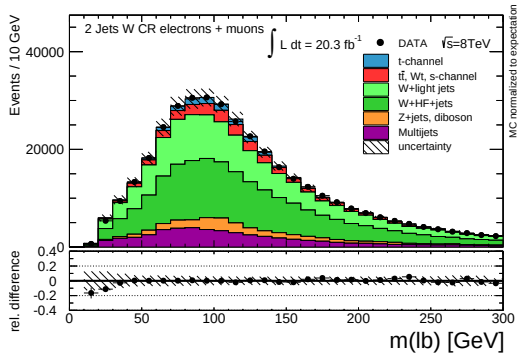
10.1 The $m(\ell b)$ observable

As the observable that is sensitive to m_{top} the invariant mass of the reconstructed lepton and b-jet in the event, $m(\ell b)$, is chosen. This corresponds to the reconstructed top quark mass without the neutrino component. The full 4-vector of the charged lepton can always be reconstructed with good accuracy. Additionally, the requirement of exactly two jets with one of them being tagged by the MV1c algorithm at an efficiency of $\varepsilon_b = 50\%$ yields a well defined b-jet. With this procedure the assignment of the b-jet to the correct parton can be done with high accuracy and the assignment is correct in about 95% of the cases.

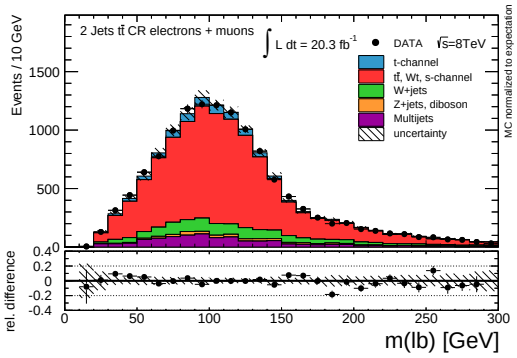
Selected events contain exactly one charged lepton and one b-tagged jet which make the assignment unambiguous. Therefore, the 4-vector can be built by summing up the vectors of the two decay products $x^{\mu}_{\ell\text{b}} = x^{\mu}_{\ell} + x^{\mu}_{\text{b-jet}}$ and the invariant mass $m(\ell b)$ can be calculated from the 4-vector components.

In Figure 10.1 ((a),(b)) the observed and expected distribution of $m(\ell b)$ is shown in the two control regions while in Figure 10.1(c) the signal region is shown. Figure 10.1(d)

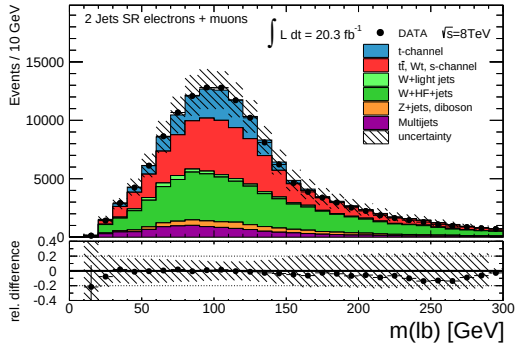
shows the predicted and expected distribution of $m(\ell b)$ after the final event selection including the $NN > 0.75$ cut. All plots assume a top quark mass of $m_{\mathrm{top}} = 172.5$ GeV in simulation and good agreement is seen.



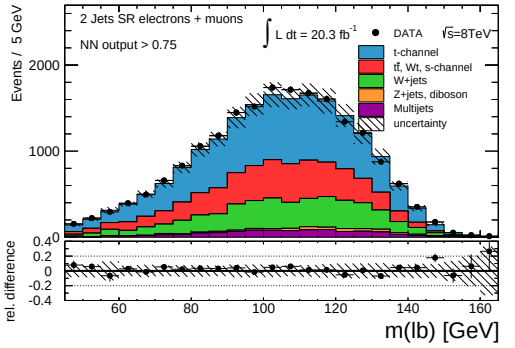
(a) $W^{\pm}$-boson control region



(b) $t\bar{t}$ control region



(c) Signal region

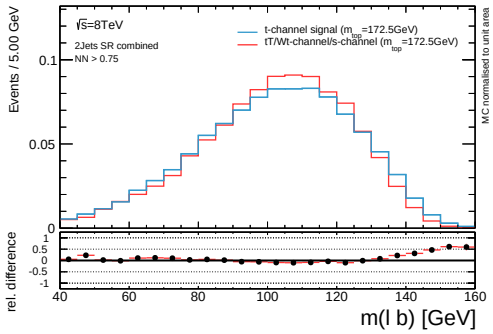


(d) Signal region including the cut on the neural network output distribution

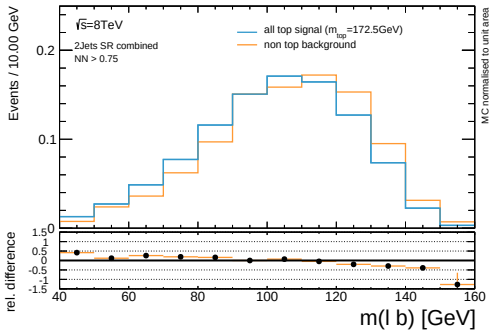
Figure 10.1: Expected and observed distribution of $m(\ell b)$ assuming $m_{\mathrm{top}} = 172.5$ GeV in the two control regions, in the signal region and in the signal region after the cut on the neural network output distribution. The hatched bands indicate the statistical uncertainty and the uncertainty on the QCD-multijet normalisation in the control regions and on the W+jets normalisation in the signal region.

In Figure 10.2 the shapes of the different signal and background components of the $m(\ell b)$ observable normalised to unit area are shown. In Figure 10.2(a) one can see that the shape of $m(\ell b)$ in the peak region is similar for the t-channel process compared to the other top quark mass dependent processes. $t\bar{t}$ events have a larger jet (and b-jet) multiplicity leading to the possibility of a wrong assignment of the second b-jet from the decay of the second top quark. This can be seen as a tail at high values of $m(\ell b)$ in

Figure 9.4(c) in the signal region before the neural network selection is applied. After cutting on the neural network in Figure 10.2(a) deviations at higher $m(\ell b)$ values are still present but much less pronounced when comparing the t-channel with the other top quark processes. The interpretation of this is that the neural network often selects the correctly assigned b-jet and lepton stemming from the same top quark in $t\bar{t}$ events.



(a) t-channel compared to other top quark processes



(b) All top quark processes compared to mass independent background

Figure 10.2: (a) Shape of the invariant mass $m(\ell b)$ after all cuts in the signal region for the t-channel signal and other top quark mass dependent processes. (b) Shape with all signal and background processes included. $t\bar{t}$, single-top s-channel and Wt-channel production are treated as signal and the remaining background is completely independent from the top quark mass. All plots are normalised to unit area and a top quark mass of $m_{\text{top}} = 172.5$ GeV is assumed (See Appendix A.9 for other top quark masses.).

To be able to use a combined (t-channel + $t\bar{t}$ + Wt-channel + s-channel) signal distribution it is necessary that it is described by the template parametrisation, which will be introduced in Section 10.2.1 in the sensitive region of m_{top}. That this is the case can be seen in Figure 10.3 for the central mass point with $m_{\text{top}} = 172.5$ GeV. All other mass points are also well described and can be found in Appendix A.5. Additionally, in Section 10.3 it is shown that no bias is present in any mass region when using a combined signal template.

Thus, for the $m(\ell b)$ observable the background is completely independent from the top quark mass. In Figure 10.2(b) the shapes of the final signal and background processes are shown including contributions from W+jets, Z+Jets, diboson and QCD-multijet production. The relative fraction of signal and background will be included as a parameter in the fit to measure the top quarks mass. Because of that it is important that the shapes of the two show differences yielding sensitivity to discriminate between the signal and the background.

10.2 The Template Method

The template method has already been explained in Section 3.3.2 and is a commonly used method in particle physics. Simulated distributions are constructed for $m(\ell b)$ using a number of discrete values of m_{top}. The fitted parameters are then interpolated between different values of m_{top}. Using the interpolations it is possible to predict the parameters and, thus, predict a distribution of $m(\ell b)$ for all values of m_{top} within a certain mass region. These distributions of $m(\ell b)$ depending on the mass of the top quark are called templates giving the method its name.

In the final step a likelihood fit to the observed data distribution is used to obtain the value of m_{top} that best describes the data. In this procedure, the distributions are constructed such that they are unbiased estimators of the top quark mass used in the simulated sample. That this is the case will be validated in Section 10.3. Consequently, the top quark mass measured in this way in data corresponds to the mass definition that has been used in simulation. This definition is different from e.g. the pole mass definition in quantum field theory as explained in Section 3.3.1.

10.2.1 Signal probability density function

The signal distributions of $m(\ell b)$ are constructed for top quark masses in the range of $165 - 180$ GeV, using the separate simulated samples for each of the $N = 7$ mass points. These samples are all listed in Tables 7.1 and 7.2.

As a first step the seven signal distributions of $m(\ell b)$ are fitted separately with the sum of a Landau and a Gaussian function[66] corresponding to Equation (10.1):

$$s\left(p_0 \dots p_6, x = m(\ell b)\right) = p_0 \cdot \left((1 - p_1) \cdot \mathrm{G}(x|p_5, p_6) + p_1 \cdot \mathrm{L}(x|p_2, p_3, p_4)\right). \qquad (10.1)$$

This functional form depends on $n = 7$ parameters $p_0 \dots p_{n-1}$ yielding $N \times n = 49$ fit parameters. The individual fit parameters are the mean, p_5, and the width, p_6, of the Gaussian, the shift, p_2, centre, p_3, and scale, p_4, of the Landau, one parameter, p_1, of the relative Gaussian and Landau fraction and one parameter, p_0, as the overall normalisation. Since all parameters are correlated only the parameters p_3 and p_5 are left as free parameters for the mass measurement. The parameter p_0 is also a free parameter

[66] The definition of the used flipped and shifted Landau function is $\mathrm{L}(x|p_2, p_3, p_4) = \mathrm{L}(-x + p_2|p_3, p_4)$. The changed argument $x \to -x + p_2$ flips the x-axis and shifts the distribution to the sensitive region of $m(\ell b)$. The landau function $\mathrm{L}(x|\mu, \sigma)$ is defined in Ref. [188] with the most probable value $\mu = p_3$ and scale $\sigma = p_4$. The definition of the Gaussian function is analogue to the definition given in Equation (10.7) with $\mu = p_5$ and scale $\sigma = p_6$.

in each individual fit but is not used in the later mass estimation. This makes sure that there is no influence of the cross-section on the mass estimation. The parameters p_1, p_2, p_4 and p_6 are fixed to the prediction from the central mass fit with $m_{\text{top}} = 172.5$ GeV.

In Figure 10.3 the fit is shown for the central mass point at $m_{\text{top}} = 172.5$ GeV for the combined electron+muon channel. In addition, all fits of the different mass points are shown in Appendix A.5. One can see that the functional parametrisation describes the distribution of $m(\ell b)$ very well for all mass points with a very good χ^2. In Figure 10.4 the

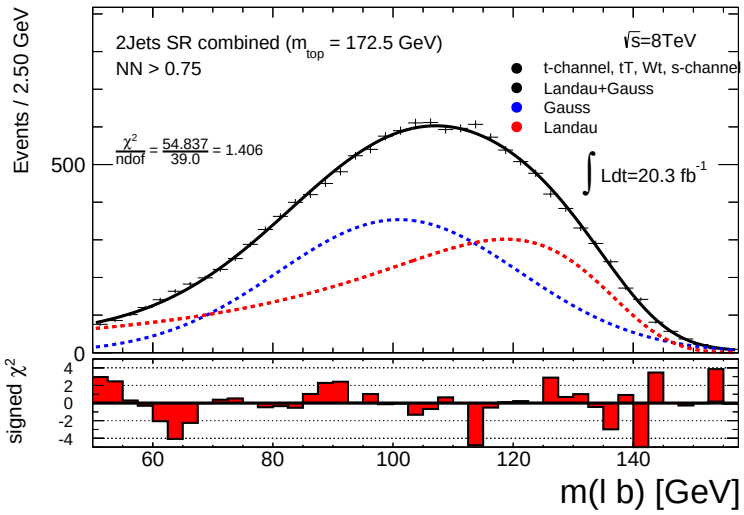


Figure 10.3: Signal distribution of the $m(\ell b)$ observable scaled to the number of expected events fitted with the function defined as the sum of a Landau and a Gaussian distribution defined in Equation (10.1). In simulation a top quark mass of $m_{\text{top}} = 172.5$ GeV is used.

sensitivity of the $m(\ell b)$ observable to the input value of m_{top} is shown as a superposition of three different mass points together with the fitted parametrisation.

By verifying and imposing a linear dependence of the parameters on m_{top} they can be expressed as

$$\tilde{p_i}(m_{\text{top}}) = a_i \cdot m_{\text{top}} + b_i \quad \text{with}\, i \in \{0, \ldots, n-1\}, \tag{10.2}$$

which reduces the number of free parameters[67] to $n \times 2$. The linear approximation of the fit parameter dependence on m_{top} is shown in Figure 10.5 for the mass-dependent

[67] To be more precise, since p_0 is only used for the normalisation and p_1, p_2, p_4 and p_6 are fixed the actual number of free parameters used to determine m_{top} would be $(n \times 2) - 10 = 4$ in each channel.

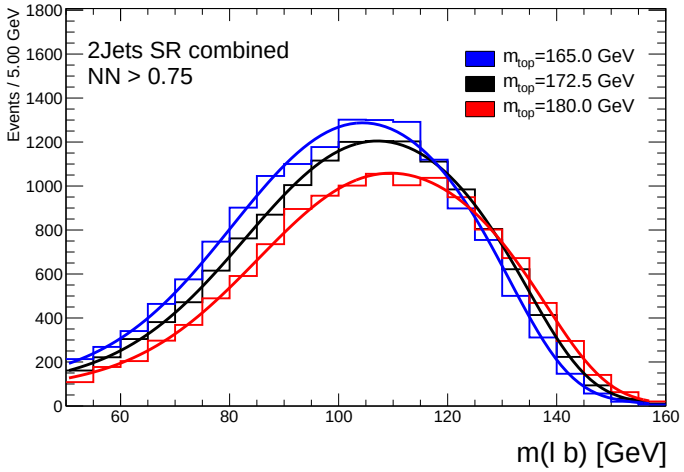
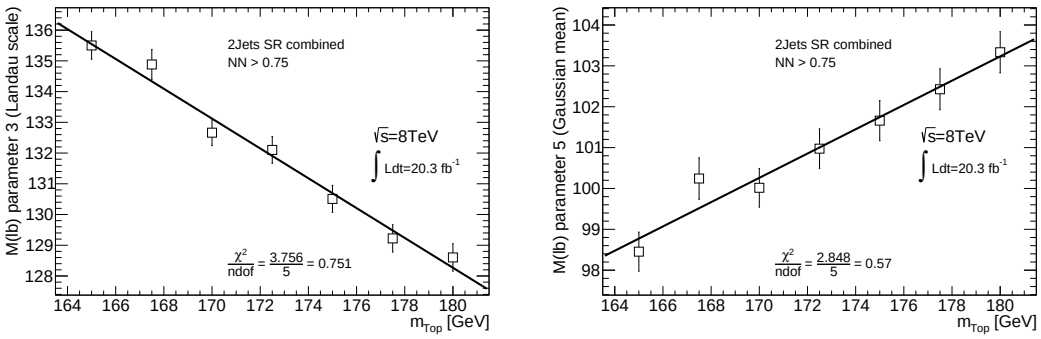


Figure 10.4: Dependence of the $m(\ell b)$ distribution on m_{top} for simulated samples generated with different input top quark masses, together with the signal probability density functions obtained from the parametrisation described in Section 10.2.1.

parameters p_3 (a) and p_5 (b). The optimal values of a_i and b_i are estimated using a fit of Equation (10.2) to the N mass points for each parameter. The calibration curves of all parameters for the electron, muon and combined electron+muon channel are shown in Appendix A.6.



(a) p_3: scale of the Landau distribution

(b) p_5: mean of the Gaussian distribution

Figure 10.5: Parameters fitted with Equation (10.2) that are used for the determination of m_{top} are shown for the two mass dependent parameters. (a) The scale, p_3, of the Landau distribution and (b) the mean, p_5, of the Gaussian distribution. The error bars indicate the one standard deviation uncertainty of the fitted parameters.

Therefore, a probability density function for the $m(\ell b)$ observable depending on m_{top} for the signal is built using the Equations (10.1) and (10.2):

$$P_{\text{sig}}(m(\ell b)|m_{\text{top}}) = s\left(\tilde{p_i}(m_{\text{top}}), m(\ell b)\right). \tag{10.3}$$

10.2.2 Background probability density function

The background events are included in the analysis also using a functional parametrisation. To estimate the background template only one fit that is independent from the top quark mass is necessary. This means that the background template is kept constant for all different mass points. The same parametrisation of a Landau and a Gaussian distribution, given by Equation (10.1), is used for the background. The seven fit parameters are referred to as $\hat{p_i}$ and in Figure 10.6 the fit to the background only $m(\ell b)$ distribution is shown.

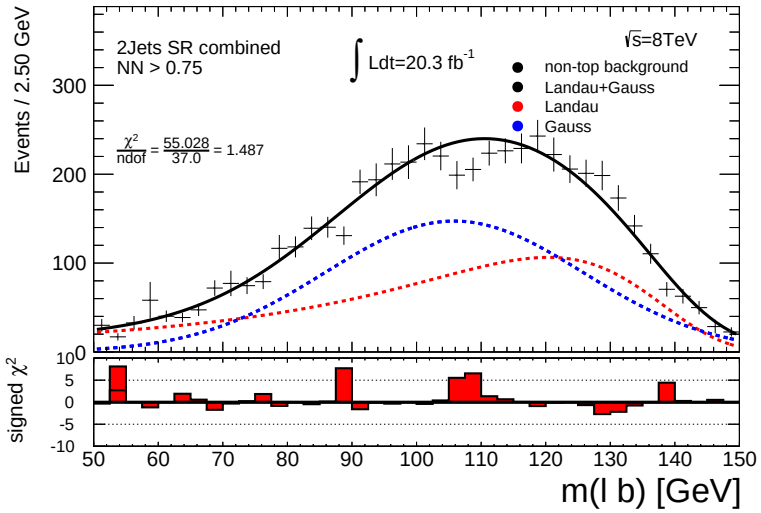


Figure 10.6: Background distribution of the $m(\ell b)$ observable scaled to the number of expected events fitted with the function defined as the sum of a Landau and a Gaussian distribution in Equation (10.1).

One caveat of the very effective reduction of the background due to the neural network selection is the low statistics remaining to estimate a decent background template. Although the total simulated background sample consists of about $350 - 400$ million events in total, compare Table 7.5, statistical fluctuations can be seen in the distributions in Figure 10.6. This is a limiting factor of the method's precision which can be

avoided by using very large datasets in simulation. That is the reason for the enormous need in terms of available computing resources.

The probability density function of the background based on the one set of seven fitted parameters, $\hat{p_i}$, and Equation (10.1) is given by:

$$P_{\mathrm{bkgd}}(m(\ell b)) = s\left(\hat{p_i}, m(\ell b)\right). \tag{10.4}$$

10.2.3 Template likelihood fit to estimate m_{top}

Finally the signal and background templates are used as the input of a binned template likelihood fit [139] to the $m(\ell b)$-distribution in data. This is done for $m(\ell b)$ in the range between 50.0 GeV and 157.5 GeV using a bin width of 2.5 GeV.

The full likelihood $\mathcal{L} = \mathcal{L}_{Model} \cdot \mathcal{L}_r$ consists of two terms where the first term, $\mathcal{L}_{Model}$, includes the signal and background templates as well as the data distribution. The second term, $\mathcal{L}_r$, constraints one of the fit parameters, f_{back}, which is the relative background fraction. This will be explained in more detail later in this section. The full logarithmic likelihood function can be written down as:

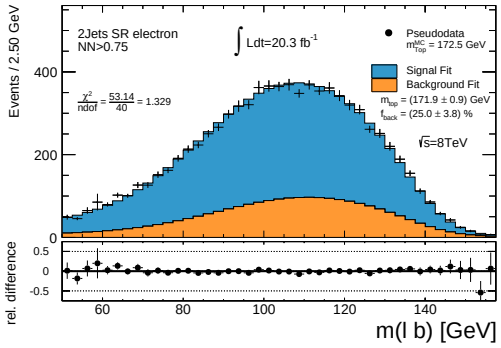
$$\begin{aligned}\ln\left(\mathcal{L}\right) = &\ln\left[\mathcal{L}_{Model}(A, f_{\mathrm{back}}, m_{\mathrm{top}} | m(\ell b)^{\mathrm{data}}_{\mathrm{bin}}, P_{\mathrm{sig}}(m(\ell b)|m_{\mathrm{top}}), P_{\mathrm{bkgd}}(m(\ell b)))\right] \\ &+ \ln\left[\mathcal{L}_r(f_{\mathrm{back}}|r_{\mathrm{MC}}, \sigma_{r_{\mathrm{MC}}})\right].\end{aligned} \tag{10.5}$$

The function has three parameters: the top quark mass, m_{top}, the relative background fraction, f_{back} and the overall normalisation, A.

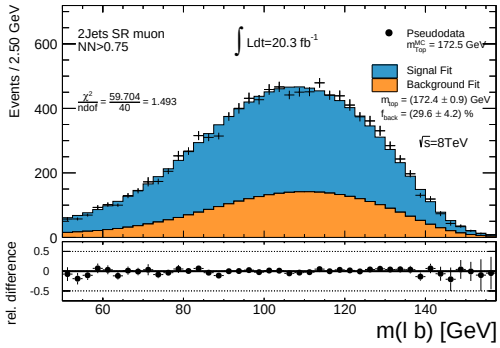
The likelihood of the model, $\mathcal{L}_{Model}$, uses the signal and background templates, $P_{\mathrm{sig}}(m(\ell b)|m_{\mathrm{top}})$ and $P_{\mathrm{bkgd}}(m(\ell b)))$, as well as the distribution of $m(\ell b)$ in data, referred to as $m(\ell b)^{\mathrm{data}}_{\mathrm{bin}}$. In a binned template likelihood fit it is defined as the product of probability functions comparing data to simulation in bins of $m(\ell b)$. The probability in each bin is taken from Poisson distributions with mean, λ_{bin}, from simulation and values, k_{bin}, from data. Thus, the likelihood of the model reads:

$$\begin{aligned}\mathcal{L}_{Model} = \prod_{\mathrm{bins}} \mathrm{Pois}_{\lambda_{\mathrm{bin}}}(k_{\mathrm{bin}}) \quad \text{with} \quad k_{\mathrm{bin}} = m(\ell b)^{\mathrm{data}}_{\mathrm{bin}} \\ \text{and} \quad \lambda_{\mathrm{bin}} = A\left[(1 - f_{\mathrm{back}}) \cdot P_{\mathrm{sig}}(m(\ell b)_{\mathrm{bin}}|m_{\mathrm{top}}) + f_{\mathrm{back}} \cdot P_{\mathrm{bkgd}}(m(\ell b)_{\mathrm{bin}})\right].\end{aligned} \tag{10.6}$$

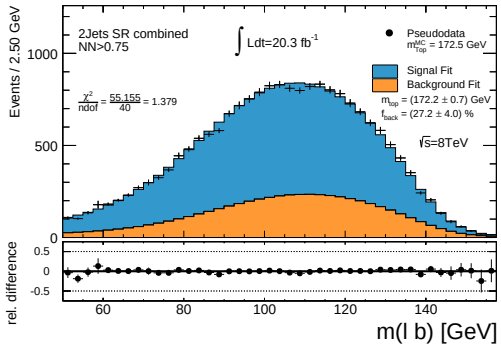
In Figure 10.7 an example fit with the signal and background probability density functions is shown using pseudo-data with $m_{\mathrm{top}} = 172.5$ GeV.



(a) Electron channel



(b) Muon channel



(c) Combined electron+muon channel

Figure 10.7: Fitted signal and background probability density functions to pseudo-data with $m_{\text{top}} = 172.5$ GeV scaled to the number of expected events.

f_{back} is constrained by a Gaussian distribution centred around the prediction from simulation r_{MC}. The width of the Gaussian $\sigma_{r_{\text{MC}}}$ is reflecting the theoretical uncertainty on the background fraction. Both values, r_{MC} and $\sigma_{r_{\text{MC}}}$, are given in Table 9.2 separately for the electron and the muon channel. The constraint is included in the likelihood as mentioned above and is given by:

$$\mathcal{L}_r = G(f_{\text{back}}|r_{\text{MC}}, \sigma_{r_{\text{MC}}}) = \frac{1}{\sqrt{2\pi}\sigma_{r_{\text{MC}}}} \cdot \exp\left(-\frac{(f_{\text{back}} - r_{\text{MC}})^2}{2\sigma^2_{r_{\text{MC}}}}\right) \tag{10.7}$$

The strength of the constraint was varied to make sure that the mass estimation is not biased in any way. Details can be found in Appendix A.7. The background fraction given by the fit is $(25 \pm 4)\,\%$ in the electron channel and $(30 \pm 4)\,\%$ in the muon channel. This is in good agreement with the predicted fractions from Table 9.2 of $(26 \pm 2)\,\%$ in the

electron channel and $(29 \pm 2)\,\%$ in the muon channel, respectively. Also the obtained top quark mass is in good agreement with the mass used in simulation.

10.3 Validation of the method

Following the procedure described in Section 10.2 the templates are estimated for measuring the mass of the top quark. For the $m(\ell b)$ observable a sum of a Gaussian and a Landau function was chosen where the Landau was flipped to account for the flatter right flank. This function provided the best performance in terms of χ^2 per degrees of freedom and stability compared to other investigated functional parametrisations. The background is also parametrised by the same function but since all top quark mass dependent processes are treated as signal there is no dependence left in the background fit.

In this section it will be shown that the estimation of the top quark mass and the statistical uncertainty is unbiased using pseudo-experiments.

10.3.1 Drawing pseudo-data and oversampling correction

It is often useful to create ensembles of pseudo-experiments generated from a distribution like the $m(\ell b)$-distribution. With this technique the analysis method can e.g. be tested for a possible bias or statistical fluctuations can be averaged out to isolate systematic uncertainties.

One pseudo-dataset is created by randomly drawing k events from the underlying $m(\ell b)$-distribution where k is following a Poisson distribution, Pois(λ), with λ being the number of expected events. This procedure can be repeated N times in order to generate N ensembles of pseudo-data.

In principle it would be possible to reduce the statistical uncertainty of the underlying simulated dataset to zero by drawing $N \to \infty$ pseudo-datasets. However, the pseudo-datasets are drawn from a finite amount of simulated events. In this case these are the events contributing to the distribution all ensembles of pseudo-data are drawn from. This incorrect calculation of the statistical uncertainty of the simulated sample is referred to as oversampling. The Barlow correction [189] used to correct for oversampling is defined in Equation (10.8):

$$\sigma_{\mathrm{stat}}(N, \rho) \propto \sqrt{\frac{1}{N} + \rho}, \tag{10.8}$$

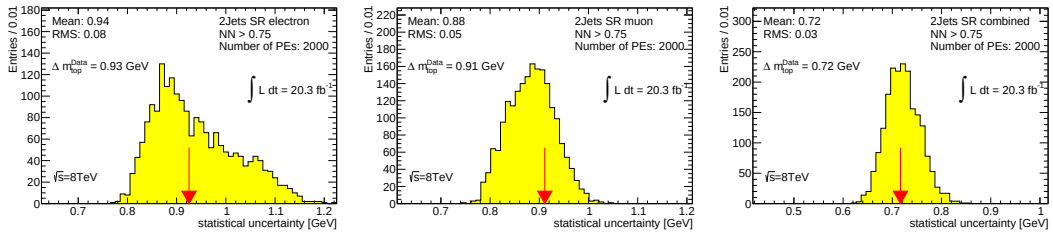
with N being the number of pseudo-experiments and ρ the correlation between the pseudo-data samples. For drawing of events without replacement the value of ρ is defined as $\rho = \frac{n}{m}$. n corresponds to the number of events in the underlying simulated sample and m is the number of events drawn for the pseudo-dataset.

All uncertainty bars in Sections 10.3.2 and 10.3.3 are corrected for oversampling using the Barlow correction depending on the number of pseudo-experiments drawn. This makes sure that the correct statistical uncertainty corresponding to the true uncertainty of the underlying simulated dataset is obtained.

10.3.2 Expected sensitivity and fit bias

To validate the method 2000 ensembles of pseudo-experiments per mass variation sample were performed and the results are presented in this section later on.

The expected statistical uncertainty distributions for a measured mass of $m_{\text{top}} = 172.5$ GeV are shown in Figure 10.8 in comparison to the obtained statistical uncertainty that is estimated in the template likelihood fit to the data, which will be detailed in Chapter 11. Good agreement to the predicted precision which is of the order of 0.7 GeV in the combined electron+muon channel can be seen.



(a) Electron channel (b) Muon channel (c) Electron+muon channel

Figure 10.8: Expected statistical uncertainty assuming $m_{\text{top}} = 172.5$ GeV evaluated with pseudo-experiments in the electron (a), muon channel (b) separately and in the combined electron+muon channel (c). Also shown is the observed statistical uncertainty that is estimated in the template likelihood fit to the data, which will be detailed in Chapter 11, indicated by a red arrow.

In Figure 10.9 ((a), (c), (e)) the mean estimated mass $\langle m_{\text{top}}^{\text{out}} \rangle$ given by the template fit and in ((b), (d), (f)) the resulting mass differences $\Delta m_{\text{top}} = m_{\text{top}}^{\text{in}} - \langle m_{\text{top}}^{\text{out}} \rangle$ depending on $m_{\text{top}}^{\text{in}}$ is displayed. It is clearly visible that the mass used to draw the pseudo-experiments, $m_{\text{top}}^{\text{in}}$, is obtained reliably and no bias is present in the whole sensitive mass region.

The so called *pull* is defined as:

$$pull = \frac{m_{\mathrm{top}}^{\mathrm{in}} - m_{\mathrm{top}}^{\mathrm{out}}}{\sigma_{\mathrm{top}}^{\mathrm{out}}}, \tag{10.9}$$

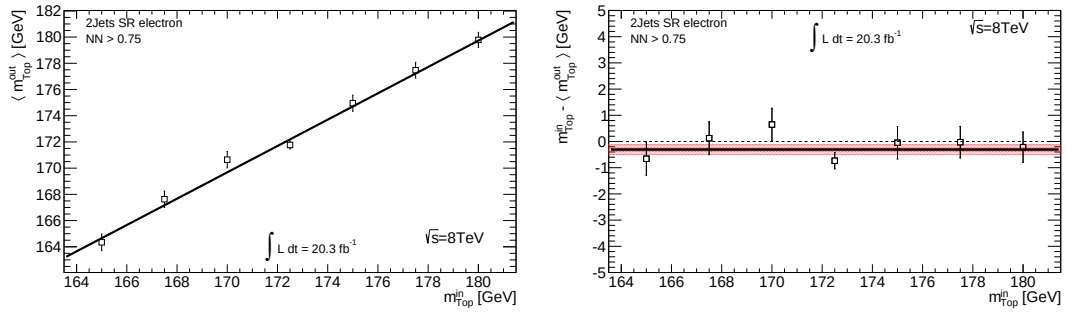
with $m_{\mathrm{top}}^{\mathrm{in}}$ being the expected value of m_{top} and $m_{\mathrm{top}}^{\mathrm{out}}$ being the fitted value of m_{top} with its statistical uncertainty $\sigma_{\mathrm{top}}^{\mathrm{out}}$. If the statistical uncertainty is correctly estimated and the mass estimation is unbiased, the pull is expected to be distributed as a standard Gaussian with mean zero and unit width.

The pull width values, shown in Figure 10.10 ((b), (d), (f)), are deviating slightly from one. However, as these deviations are in opposite direction in the electron and the muon channel it is unlikely that a general problem with the template fit is present. Also a hypothetical correction would be far below the statistical precision of the method and, thus, the final result of the measured value of m_{top} would not change at all. This proves together with the plots in Figure 10.9 that the method is unbiased. Also the expected statistical uncertainty depending on m_{top} is shown in Figure 10.10 ((a), (c), (e)). The uncertainty bars in all figures indicate the remaining uncertainty on the result of the pseudo-experiments. This corresponds to the true uncertainty of the underlying simulated dataset that is corrected for oversampling following the procedure given in Section 10.3.1. In Appendix A.8 the pull-distributions, distributions of the fitted mass and distributions of the statistical uncertainties are shown for all mass points.

10.3.3 Influence of m_{top} on neural network training

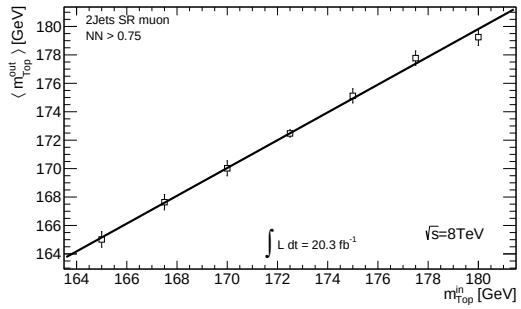
Some of the input variables of the neural network, introduced in Section 9.2, are clearly expected to have a dependence on the mass of the top quark. Thus, a possible bias due neural network based event selection can not be excluded a priori and has to be checked.

Three independent trainings of the neural network have been performed using a mass sample with $m_{\mathrm{top}} = 165.0$ GeV, $m_{\mathrm{top}} = 172.5$ GeV (default) or $m_{\mathrm{top}} = 180.0$ GeV as signal sample. Figure 10.11(a) shows the $m(\ell b)$ distribution of all processes including top quarks after the selection with the three neural networks normalised to the number of events obtained for the central mass sample as given in Table 9.2. The three simulated samples all use a top quark mass of $m_{\mathrm{top}} = 172.5$ GeV and are selected with the full event selection including the cut on the neural network at $NN > 0.75$. It is clearly visible that the distribution is shifted to higher (lower) values if the neural network was trained with a higher (lower) top quark mass. This is the expected behaviour which is

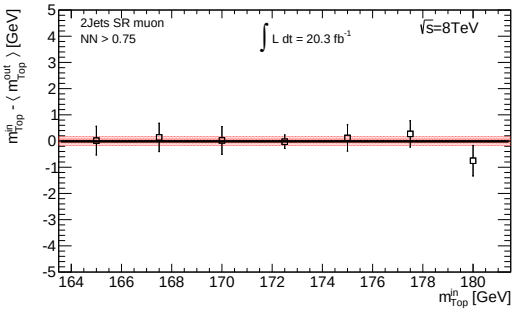


(a) Electron channel

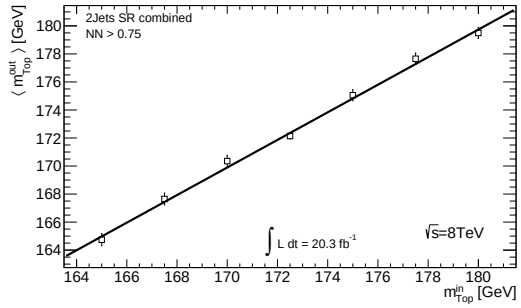
(b) Electron channel



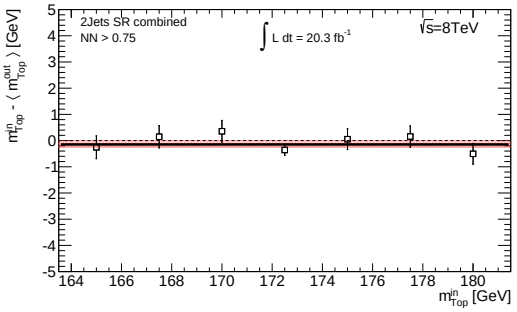
(c) Muon channel



(d) Muon channel

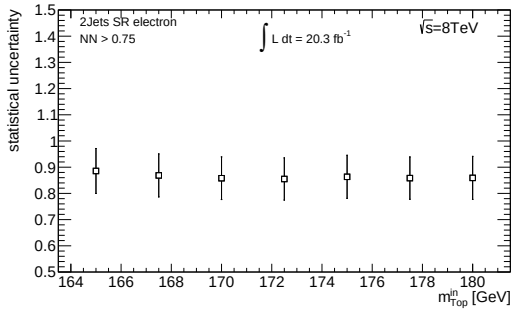


(e) Electron+muon channel

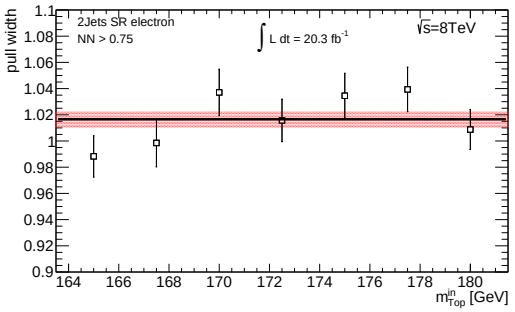


(f) Electron+muon channel

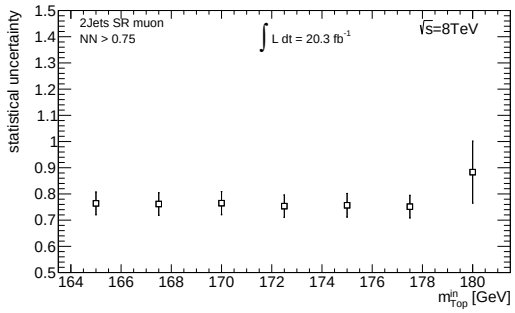
Figure 10.9: Plots to check for a possible bias in the mass determination. The x-axis refers to the top quark mass used as a parameter during simulation. The y-axis is the mean fitted mass $\langle m_{\text{top}}^{\text{out}} \rangle$ (left) and the difference of the mean mass to the mass used in the pseudo-data sample $\Delta m_{\text{top}} = m_{\text{top}}^{\text{in}} - \langle m_{\text{top}}^{\text{out}} \rangle$ (right).



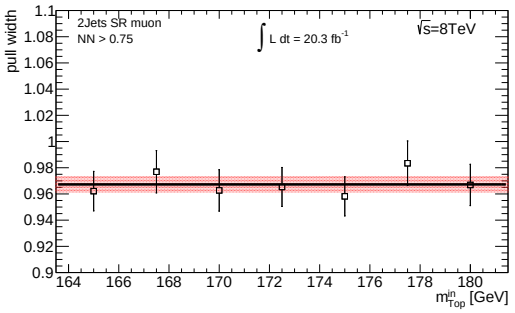
(a) Electron channel



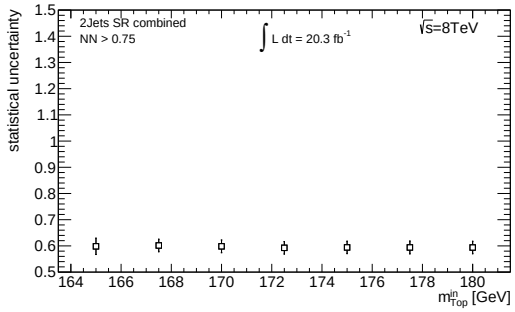
(b) Electron channel



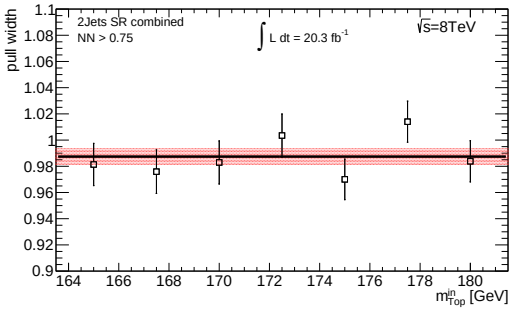
(c) Muon channel



(d) Muon channel

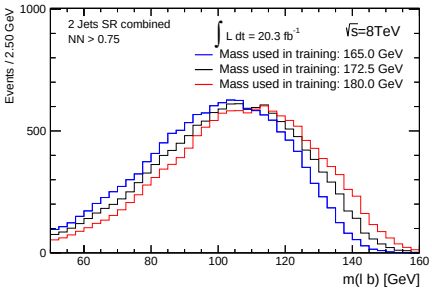


(e) Electron+muon channel

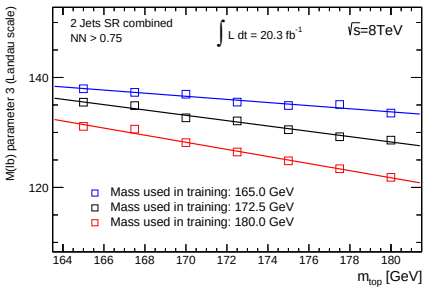


(f) Electron+muon channel

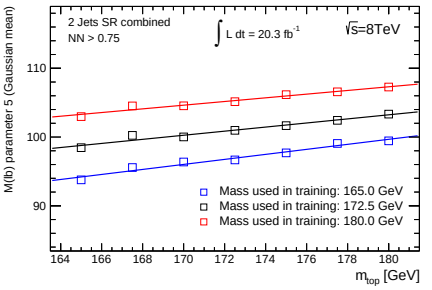
Figure 10.10: Plots to check the consistency of the estimation of the statistical uncertainty in the mass determination. The x-axis refers to the top quark mass used as a parameter during simulation. The y-axis is the expected statistical uncertainty (left) and the width of the pull-distribution (right) showing that the estimated statistical uncertainty is correctly estimated.



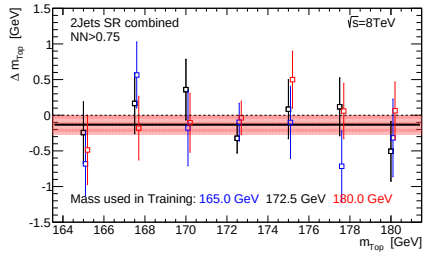
(a) $m(\ell b)$-distribution for all mass dependent processes after the full event selection



(b) p_3: scale of the Landau distribution



(c) p_5: mean of the Gaussian distribution



(d) Mass difference $\Delta m_{\text{top}} = m_{\text{top}}^{\text{in}} - \langle m_{\text{top}}^{\text{out}} \rangle$ depending on $m_{\text{top}}^{\text{in}}$

Figure 10.11: Plots to check if any possible bias is introduced due to the mass dependence of the neural network input variables. Shown is the combined electron+muon channel where the training of the neural network has been performed using a mass sample with $m_{\text{top}} = 165.0$ GeV (blue), $m_{\text{top}} = 172.5$ GeV (default,black) or $m_{\text{top}} = 180.0$ GeV (red) as signal sample. In (d) it can be seen that no bias is observed.

also seen in Figure 10.4 where the same neural network selection is applied to samples with different top quark masses.

The estimation of the signal and background probability density functions, detailed in Sections 10.2.1 and 10.2.2, is now repeated for the three scenarios. Figures 10.11(b) and 10.11(c) show the parameter calibration curves, corresponding to Equation (10.2), for the two mass dependent parameters of the signal parametrisation. It can be seen that both parameters are shifted which absorbs the mass dependence introduced by the selection.

To ensure that no bias remains pseudo-experiments have been performed. The procedure to check if any bias is present is done analogue as described in Section 10.3.2. It

is repeated three times with the three neural networks trained with different top quark mass samples and the corresponding signal and background probability density functions. In Figure 10.11(d) the mass difference $\Delta m_{\mathrm{top}} = m_{\mathrm{top}}^{\mathrm{in}} - \langle m_{\mathrm{top}}^{\mathrm{out}} \rangle$ is shown for the three selections depending on $m_{\mathrm{top}}^{\mathrm{in}}$. For all mass points $m_{\mathrm{top}}^{\mathrm{in}}$ is obtained by the fit within uncertainties. Thus, no indication for a bias introduced by the top quark mass influence on the neural network training is observed.

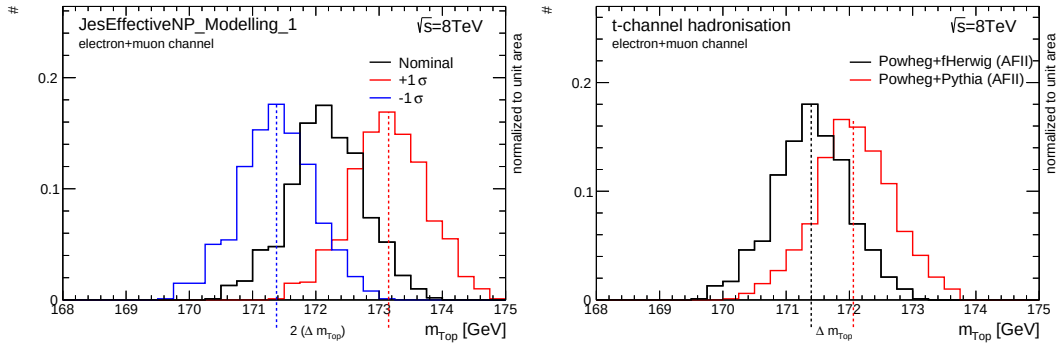
10.4 Systematic uncertainties

The measurement of the top quark mass suffers from a variety of systematic uncertainties. These are calculated by performing pseudo-experiments with the technique described in Section 10.3. Each systematic uncertainty is estimated by varying the respective uncertainty source and determining the impact on the top quark mass measurement[68]. Whenever it is possible the uncertainty source is varied by one standard deviation ($\pm 1\sigma$) with respect to the default value. The average value of m_{top} obtained with the pseudo-experiments including the $+1\sigma$ systematic variation, $\langle m_{\mathrm{top}}^{+1\sigma} \rangle$, is compared to the corresponding value of the -1σ systematic variation, $\langle m_{\mathrm{top}}^{-1\sigma} \rangle$. The difference is used to determine the systematic uncertainty. This is depicted in Figure 10.12(a) as an example for one of the systematic uncertainties that are discussed later.

For systematic sources where only one variation is available the difference between $\langle m_{\mathrm{top}}^{1\sigma} \rangle$ and $\langle m_{\mathrm{top}}^{\mathrm{nominal}} \rangle$ is used to determine the systematic uncertainty. Furthermore some uncertainties involve a comparison of two scenarios that are directly compared. An example is shown in Figure 10.12(b) for the single-top t-channel hadronisation where the simulated sample representing the single-top t-channel has been replaced. Some other special systematic uncertainties involve different procedures and are explained in the respective section.

For most of the uncertainty sources the same simulated events are used such that the observed values of m_{top} for the different sample are highly statistically correlated. In all cases the actual observed difference is quoted as the systematic uncertainty on the corresponding source, even if it is smaller than the statistical precision of the difference, following the recommendation given in Ref. [190]. The statistical precision of the systematic uncertainties was tested using pseudo-experiments and found to be about 0.3 GeV. Thus, statistical fluctuations between the electron and muon channel can appear if the

[68] Usually this means that the simulated sample with the nominal selection is replaced by a simulated sample including systematic variations or changes in the selection.



(a) ($\pm 1\sigma$) systematic variation (b) Systematic uncertainty based on two scenarios

Figure 10.12: Examples how systematic uncertainties are calculated for a ($\pm 1\sigma$) variation (a) and a systematic uncertainty comparing two scenarios (b).

corresponding systematic uncertainty itself is small. In all cases these fluctuations are below the statistical precision of the measurement.

The total uncertainty is calculated as the quadratic sum of the individual contributions. This approach neglects possible correlations between different systematic uncertainties, which in any case are expected to be small.

10.4.1 Object energy scale/resolution and efficiencies

Systematic uncertainties due to the residual differences between data and simulation on reconstructed objects after calibration are propagated through the whole analysis. This involves the reconstruction of all objects explained in Chapter 5 like jets, leptons and the missing transverse momentum.

Lepton energy scale and resolution The underlying lepton momentum correction factors and associated uncertainties were derived by the ATLAS combined performance groups as briefly explained in Sections 5.3.4 and 5.4.4. To evaluate the effect of the lepton momentum scale uncertainties the event selection is redone with the lepton momentum varied by one standard deviation ($\pm 1\sigma$). The systematic uncertainties on the lepton energy scales are then given by half the difference between the two average fitted top quark masses following Equation (10.10):

$$\Delta m_{\text{top}} = 0.5 \cdot \left(\langle m_{\text{top}}^{+1\sigma} \rangle - \langle m_{\text{top}}^{-1\sigma} \rangle \right) . \tag{10.10}$$

This procedure corresponds to the example displayed in Figure 10.12(a).

In case of the lepton momentum resolution the event selection is redone with the lepton momentum smeared by one standard deviation ($\pm 1\sigma$). Since muons are reconstructed using the inner tracking detector and the muon spectrometer two variations are needed. This is accounting for the different energy resolution of the two sub-detectors including correlations among them. The final systematic uncertainty due to the muon energy resolution is half the difference between the maximum and minimum value of the four estimated average fitted top quark masses:

$$\begin{aligned}\Delta m_{\mathrm{top}} = 0.5 \cdot [\, &\max(\langle m_{\mathrm{top}}^{+1\sigma,\mathrm{ID}}\rangle, \langle m_{\mathrm{top}}^{-1\sigma,\mathrm{ID}}\rangle, \langle m_{\mathrm{top}}^{+1\sigma,\mathrm{MS}}\rangle, \langle m_{\mathrm{top}}^{-1\sigma,\mathrm{MS}}\rangle) \\ &- \min(\langle m_{\mathrm{top}}^{+1\sigma,\mathrm{ID}}\rangle, \langle m_{\mathrm{top}}^{-1\sigma,\mathrm{ID}}\rangle, \langle m_{\mathrm{top}}^{+1\sigma,\mathrm{MS}}\rangle, \langle m_{\mathrm{top}}^{-1\sigma,\mathrm{MS}}\rangle)]. \end{aligned} \tag{10.11}$$

The electron energy resolution uncertainty is given by half the difference analogue to the lepton energy scale uncertainties following Equation (10.10).

Lepton reconstruction The mis-modelling of lepton trigger, reconstruction and selection efficiencies in simulation are corrected by scale factors depending on the lepton kinematics. These have been derived with uncertainties by the combined performance groups of ATLAS as briefly explained in Sections 5.3.4 and 5.4.4. The systematic uncertainties on the measurement of m_{top} were evaluated by varying the respective scale factors by one standard deviation ($\pm 1\sigma$). The systematic sources are called lepton identification, lepton reconstruction and lepton trigger efficiency in Table 10.7. They have been derived following the same procedure as explained in the previous paragraph using Equation (10.10).

Jet energy scale (JES) The jet energy scale uncertainty is accounting for the limited ability to reconstruct kinematic quantities of jets with the ATLAS calorimeter. It was derived by the ATLAS JetEtMiss combined performance group using information from test-beam data, LHC collision data and simulation.

The total jet energy scale is split up into 21 components, which are considered uncorrelated. These components vary as a function of jet p_T and η and are evaluated in different categories: detector, statistical, physics modelling, mixed detector and modelling, η intercalibration, pile-up, single particle and flavour. A detailed description of the different categories can be found in Ref. [113]. Additional contributions, namely pile-up (p_T term and ρ topology), arise due to the large pile-up effects in 2012 data. Another component called b-jet energy scale (bJES) ranges up to 2.5%, depending on

the p_{T} of the jet. It is applied for b-quark induced jets only due to differences between light-flavour- and gluon-jets as opposed to jets containing b-hadrons.

For each of the 21 components the kinematic quantities[69] of the reconstructed jets are scaled by one standard deviation ($\pm 1\sigma$) and the full event selection and analysis are repeated. Thus, every component of the systematic uncertainty is estimated following Equation (10.10). The individual components are given in Table 10.1 and the total jet energy scale uncertainty also listed in Table 10.7 is given by the quadratic sum.

Table 10.1: Systematic uncertainties due to the jet energy scale on the measurement of the top quark mass for variations at the level of $\pm 1\sigma$.

Source of uncertainty	$\lvert\Delta m_{\mathrm{top}}\rvert$ [GeV] electron	$\lvert\Delta m_{\mathrm{top}}\rvert$ [GeV] muon	$\lvert\Delta m_{\mathrm{top}}\rvert$ [GeV] electron+muon
EffectiveNP detector 1	0.30	0.41	0.37
EffectiveNP detector 2	0.03	0.07	0.03
EffectiveNP detector 3	0.01	0.13	0.08
η intercalibration (statistical)	0.15	0.32	0.24
EffectiveNP statistical 1	0.05	0.15	0.10
EffectiveNP statistical 2	0.00	0.11	0.06
EffectiveNP statistical 3	0.10	0.02	0.06
η intercalibration (modelling)	0.72	1.06	0.90
EffectiveNP modelling 1	0.85	0.91	0.91
EffectiveNP modelling 2	0.06	0.07	0.07
EffectiveNP modelling 3	0.10	0.23	0.16
EffectiveNP modelling 4	0.01	0.10	0.06
EffectiveNP mixed 1	0.00	0.06	0.02
EffectiveNP mixed 2	0.03	0.02	0.02
Pile-up (ρ topology)	0.14	0.18	0.18
Pile-up (p_T term)	0.06	0.04	0.00
Pile-up offset (μ term)	0.13	0.22	0.18
Pile-up offset (NPV term)	0.03	0.29	0.18
Single particle high p_T	0.01	0.06	0.02
Flavour composition	0.18	0.16	0.20
Flavour response	0.07	0.01	0.05
b-jet energy scale	0.46	0.39	0.43
total jet energy scale	1.29	1.64	1.49

[69] This mainly concerns energy and transverse momentum of the jets.

Jet reconstruction efficiency The jet reconstruction efficiency was measured using minimum bias events and simulated events containing dijets. It was found that the jet reconstruction is fully efficient for jets with $p_T > 30$ GeV [191]. Only jets above this threshold contribute to the analysis and no dedicated systematic uncertainty for the reconstruction efficiency of jets is needed.

Jet energy resolution The impact of the jet energy resolution is evaluated by smearing the jet energy in simulation. Before the event selection is performed the energy of each jet is smeared by a Gaussian function such that the width of the resulting Gaussian distribution corresponds to the one which includes the measured jet energy resolution uncertainty [143]. The difference among the average fitted top quark mass for the smeared and unsmeared scenario is taken as the systematic uncertainty:

$$\Delta m_{\mathrm{top}} = \langle m_{\mathrm{top}}^{\mathrm{nominal}} \rangle - \langle m_{\mathrm{top}}^{1\sigma} \rangle. \tag{10.12}$$

Missing transverse momentum The missing transverse momentum is reconstructed from the reconstructed leptons and jets in the event as described in Section 5.5. The uncertainties from the energy scale and resolution of these objects explained previously are propagated into the calculation of the missing transverse momentum. This means that the systematic uncertainties arising from these sources modifying the missing transverse momentum are already covered.

An inherent uncertainty of the $E_{\mathrm{T}}^{\mathrm{miss}}$-resolution remains that arises from the contributions of calorimeter cells not associated to any jets (cellout term) and soft jets ($7\ \mathrm{GeV} < p_T^{\mathrm{jet}} < 20\ \mathrm{GeV}$). These components are again varied by one standard deviation ($\pm 1\sigma$) and the systematic uncertainties calculated with Equation (10.10) are listed in Table 10.7.

Flavour-tagging efficiency Since the analysis makes use of b-tagging, the uncertainties on the b- and c-tagging efficiencies and the mistag rate are taken into account. Methods to evaluate correction factors from collision data in dijet or $t\bar{t}$ events have been explained in detail in Section 6.2. These correction factors are applied to all simulated events used in the analysis to match the flavour-tagging performance measured in data.

The uncertainties on the correction factors are propagated independently for b-jets, c-jets and light-flavour-jets. In the case of the b-tagging efficiency the eigenvector approach is chosen which splits up the inclusive uncertainty into a set of nine uncorrelated eigenvector components. Each component is varied by one standard deviation ($\pm 1\sigma$) and

systematic uncertainties are calculated following Equation (10.10). The detailed components are listed in Table 10.2, while the quadratic sum of all components is taken as the systematic uncertainty called b-tagging efficiency.

In the case of the tagging efficiency of c-jets and the mis-tagging efficiency of light-flavour-jets only an inclusive uncertainty is estimated neglecting possible correlations.

Table 10.2: Systematic uncertainties due to the flavour-tagging on the measurement of the top quark mass for variations at the level of $\pm 1\sigma$.

Source of uncertainty	$\lvert\Delta m_{\text{top}}\rvert$ [GeV] electron	$\lvert\Delta m_{\text{top}}\rvert$ [GeV] muon	$\lvert\Delta m_{\text{top}}\rvert$ [GeV] electron+muon
b-tagging efficiency 1	0.02	0.02	0.02
b-tagging efficiency 2	0.00	0.01	0.01
b-tagging efficiency 3	0.04	0.04	0.04
b-tagging efficiency 4	0.03	0.04	0.03
b-tagging efficiency 5	0.04	0.04	0.04
b-tagging efficiency 6	0.00	0.00	0.00
b-tagging efficiency 7	0.06	0.07	0.06
b-tagging efficiency 8	0.07	0.08	0.08
b-tagging efficiency 9	0.05	0.06	0.05
b-tagging efficiency	0.13	0.14	0.13
c-tagging mistag rate	0.03	0.22	0.14
light-jet tagging mistag rate	0.23	0.25	0.22
flavour tagging efficiency	0.26	0.36	0.30

10.4.2 Monte Carlo generators and parton densities

The analysis strongly depends on simulated samples and systematic effects from the MC modelling have to be taken into account. This is done by comparing different generators directly or varying parameters used during the event generation. In most cases the scenario to calculate the uncertainty corresponds to the example shown in Figure 10.12(b). The simulated samples used are all explained in Section 7.4. Additionally, another source of theoretical uncertainty arises from the PDF of the initial-state protons. The individual components of the different systematic uncertainties are all listed in Table 10.4.

Proton PDF The initial-state protons are complex composite objects and their internal structure is explained in Section 2.6. In simulation the PDFs are taken from

different PDF sets that contain a central value and a dedicated error set of uncertainty eigenvectors.

To evaluate the systematic uncertainty due to the PDF, events are reweighted according to three different central values and their respective uncertainty eigenvectors. Used here are the most common CT10 [59] PDF set, the MSTW2008nlo68cl [55] PDF set, and the NNPDF2.3 [192] set.

The reweighting procedure taking into account all top quark mass dependent samples is done by using the formula given in Equation (43) of Ref. [193][70].

Different strategies are combined to evaluate the final uncertainty due to the proton PDF. The split-up of the systematic uncertainties and the envelope [194] are shown in Table 10.3. For the components indicated with "inter" events are reweighted to switch from the central value of one PDF set to another PDF set. This means that the uncertainty is given by the difference between the two. Components labelled as "intra" use a reweighting according to the uncertainty eigenvectors of the respective PDF set itself. Each PDF set has a different set of uncertainty eigenvectors and different prescription how to derive the uncertainty. CT10 has 52 uncertainty eigenvectors and the uncertainty is the symmetric Hessian of the 52 average fitted top quark masses. For the MSTW2008nlo68cl PDF set 42 uncertainty eigenvectors and the asymmetric Hessian are used while for NNPDF2.3 the uncertainty corresponds to the standard deviation of 100 uncertainty eigenvectors.

The final systematic uncertainty due to the PDF of the initial-state protons is calculated as the envelope of the individual components, following the PDF4LHC recommendation. The whole procedure is explained in more detail in Ref. [194].

$t\bar{t}$ MC generator Systematic effects from the MC modelling of the $t\bar{t}$ process are estimated by comparing different simulated samples that are listed in Table 7.4.

The full difference of the average fitted top quark mass, $\langle m_{\mathrm{top}}^{\mathrm{out}} \rangle$, is estimated if either using the MC@NLO + Herwig (dataset ID: 105200) or Powheg + Pythia6 (dataset ID: 117050) sample for $t\bar{t}$. To also cover a potential systematic uncertainty due to the parton showering the full difference is also calculated if either using the Powheg + Pythia6 (dataset ID: 117050) or Powheg +Herwig (dataset ID: 105860) sample with fast-simulation. The largest of the two differences is taken as a systematic uncertainty denoted as $t\bar{t}$ MC generator in Table 10.7.

[70] For the Powheg event generator the information necessary to perform the reweighting is only available for the leading-order process. This can lead to an underestimated systematic uncertainty. That is why for $t\bar{t}$ production and s-channel and Wt-channel production the sample generated with MC@NLO listed in Table 7.3 and 7.4 is used.

Table 10.3: Systematic uncertainties due to the PDFs on the measurement of the top quark mass for variations at the level of $\pm 1\sigma$.

Source of uncertainty	$\lvert\Delta m_{\text{top}}\rvert$ [GeV] electron	$\lvert\Delta m_{\text{top}}\rvert$ [GeV] muon	$\lvert\Delta m_{\text{top}}\rvert$ [GeV] electron+muon
Proton PDF, inter MSTW/CTEQ	0.00	0.00	0.00
Proton PDF, inter MSTW/NNPDF	0.05	0.07	0.07
Proton PDF, inter NNPDF/CTEQ	0.05	0.07	0.06
Proton PDF, intra CTEQ	0.06	0.07	0.05
Proton PDF, intra MSTW	0.14	0.00	0.00
Proton PDF, intra NNPDF	0.00	0.00	0.01
Proton PDF (envelope)	0.14	0.07	0.05

Initial and final state radiation ($t\bar{t}$) The systematic uncertainty due to initial-state radiation (ISR) and final-state radiation (FSR) was evaluated by comparing the two samples generated with AcerMC + Pythia6. The samples with less PS[71] (dataset ID: 117209) and more PS (dataset ID: 117210) are listed in Table 7.4 and explained in Section 7.4.3. Half the difference of the average fitted top quark mass when using the two samples is taken as the systematic uncertainty:

$$\Delta m_{\text{top}} = 0.5 \cdot \left| \langle m_{\text{top}}^{\text{lessPS}} \rangle - \langle m_{\text{top}}^{\text{morePS}} \rangle \right| . \tag{10.13}$$

$t\bar{t}$ colour reconnection / underlying event A possible systematic uncertainty is conceivable due to non-perturbative QCD effects like the modelling of colour reconnection and the underlying event. To estimate the effect on the average fitted top quark mass, $\langle m_{\text{top}}^{\text{out}} \rangle$, three samples listed in Table 7.4 have been generated using Powheg + Pythia6 with three different tunes. The nominal sample uses the Perugia2012 tune (dataset ID: 117428) which is compared to the Perugia2012loCR tune (dataset ID: 117429) in case of colour reconnection and Perugia2012mpiHi tune (dataset ID: 117426) in case of the underlying event. The full difference is taken as the respective systematic uncertainty.

Single-top t-channel MC generator The systematic uncertainty due to the MC generator for the single-top t-channel process is tested by replacing the AcerMC + Pythia6 sample in Table 7.1. To cover the pure difference due to the event generation the default scenario is compared to the average fitted top quark mass, $\langle m_{\text{top}}^{\text{out}} \rangle$, estimated if using a sample generated with Powheg + Pythia6 (dataset ID: 110090,110091).

[71] The variation of the parton shower (PS) was explained in Section 7.4.3.

A possible difference due to the parton showering and hadronisation is estimated by comparing this POWHEG + PYTHIA6 sample with a POWHEG + HERWIG (dataset ID: 110086,110087) sample. The details of the samples used are given in Table 7.3.

To keep the procedure close to the estimation of the $t\bar{t}$ MC generator uncertainty, the largest of the two differences is taken as the final uncertainty. It can be found as t-channel MC generator in Table 10.7.

Single-top t-channel colour reconnection / underlying event The three samples generated with POWHEG + PYTHIA6 using different tunes are all listed in Table 7.3. The nominal sample with the Perugia2012 tune (dataset ID: 110070,110071) is compared to either the sample with Perugia2012loCR tune (dataset ID: 110074,110075) or Perugia2012mpiHi tune (dataset ID: 110072,110073) in the same way as it was explained for the $t\bar{t}$ production previously.

Single-top Wt-channel MC generator The three samples used to estimate the influence on the choice of the single-top Wt-channel generator are listed in Tables 7.1 and 7.3. The full difference in the average fitted top quark mass, $\langle m_{\mathrm{top}}^{\mathrm{out}} \rangle$, between a POWHEG + PYTHIA6 (dataset ID: 110140) and MC@NLO + HERWIG (dataset ID: 108346) sample is calculated. Additionally, the POWHEG + PYTHIA6 samples with DR and DS separation scheme (dataset ID: 110140 and 110142) are compared to each other. The largest of the two differences is taken as the systematic uncertainty denoted as Wt-channel MC generator in Table 10.7.

Single-top s-channel MC generator The difference of the average fitted top quark mass, $\langle m_{\mathrm{top}}^{\mathrm{out}} \rangle$, is estimated when using the POWHEG + PYTHIA6 (dataset ID: 110119) or MC@NLO + HERWIG (dataset ID: 108343-108345) sample in Tables 7.1 and 7.3. The full difference is taken as the systematic uncertainty due to the single-top s-channel MC generator.

10.4.3 Background

Different systematic uncertainties related to the modelling of the background processes have to be taken into account. They can be divided into two classes. One being the overall normalisation and one concerning the shape of the $m(\ell b)$-distribution.

If the background is varied, the Gaussian constraint on the background fraction defined in Equation (10.5) is kept constant at the nominal value. This is conservative and will cover possible differences between the predicted and the true background fraction.

Table 10.4: Systematic uncertainties due to the simulation on the measurement of the top quark mass.

Source of uncertainty	$\lvert\Delta m_{\text{top}}\rvert$ [GeV] electron	$\lvert\Delta m_{\text{top}}\rvert$ [GeV] muon	$\lvert\Delta m_{\text{top}}\rvert$ [GeV] electron+muon
Proton PDF	0.14	0.07	0.05
t-channel AcerMC vs. Powheg	0.21	0.32	0.08
t-channel Pythia6 vs. Herwig	0.71	0.65	0.66
t-channel colour reconnection	0.12	0.48	0.31
t-channel underlying event	0.10	0.15	0.04
$t\bar{t}$ ISR/FSR	0.18	0.16	0.16
$t\bar{t}$ Powheg vs. MC@NLO	0.20	0.28	0.21
$t\bar{t}$ Pythia6 vs. Herwig	0.11	0.05	0.08
$t\bar{t}$ colour reconnection	0.07	0.36	0.22
$t\bar{t}$ underlying event	0.11	0.12	0.11
Wt-channel Powheg DR vs. DS	0.26	0.16	0.00
Wt-channel Powheg vs. MC@NLO	0.17	0.25	0.08
s-channel Powheg vs. MC@NLO	0.23	0.00	0.08
Simulation modelling	0.92	1.06	0.84

In Table 10.6 the systematic uncertainties related to the background are shown together with their quadratic sum.

QCD-multijet background normalisation The normalisation of the QCD-multijet background is done by fitting the $E_{\text{T}}^{\text{miss}}$-distribution. This is explained in detail in Section 8.2. A systematic uncertainty of $\sigma_{\text{QCD}} = 50\%$ is assigned based on alternative methods to estimate the QCD-multijet normalisation. What was done is e.g. using the $m_{\text{T}}(W)$ variable instead of the $E_{\text{T}}^{\text{miss}}$ or replacing the default multijet model. The result of the different estimations is given in Table 10.5. The differences are fully covered by the σ_{QCD} variation in the signal region[72].

The systematic uncertainty is calculated by shifting the QCD-multijet background contribution by one standard deviation ($\pm 1\sigma_{\text{QCD}}$) and applying Equation (10.10).

W+jets background normalisation The theoretical uncertainty for inclusive $W^{\pm}$-boson production is 4%. This is extended by an additional uncertainty per additional jet of 24% to be added in quadrature. This leads to an overall uncertainty of $\sigma_{\text{W+jets}} = 40\%$ on the W+jets background normalisation [196, 197]. To estimate the impact on the

[72] A description of the data-driven matrix method can be found in Ref. [195].

Table 10.5: Estimate of the multijet background in the signal and control region using a binned maximum-likelihood fit. The quoted numbers are the expected number of events in each region and the relative difference is calculated with respect to the default method.

Method	$W^{\pm}$-boson control region			signal region		
	events	fraction	rel. diff	events	fraction	rel. diff
	electron channel					
Jet-lepton model (def.)	24800	15.0 %	-	5500	9.0 %	-
Matrix method	13000	7.7 %	−47.8 %	6400	10.5 %	15.5 %
$m_{\mathrm{T}}(W)$	18300	24.8 %	−26.4 %	7900	15.6 %	42.7 %
	muon channel					
Anti-muons (def.)	23200	10.3 %	-	6000	7.2 %	-
Matrix method	4800	2.1 %	−79.5 %	4500	5.4 %	−26.1 %
Jet-lepton model	11500	5.1 %	−50.2 %	4400	5.3 %	−27.6 %
$m_{\mathrm{T}}(W)$	10500	4.7 %	−54.8 %	5200	6.3 %	−13.5 %

top quark mass measurement the W+jets background contribution is shifted by on standard deviation ($\pm 1\sigma_{\mathrm{W+jets}}$) and the difference based on Equation (10.10) is assigned the systematic uncertainty.

Z+jets/Diboson background normalisation The same variation as explained for the W+jets background is applied on the distribution of the combined Z+jets and diboson background. The systematic uncertainty is also calculated using Equation (10.10).

W+jets background shape The shape of $m(\ell b)$ for the most important background, that is W+jets production, is estimated using the Sherpa MC generator as explained in Section 7.5. To estimate a possible bias due to the shape a reweighting of the parton distribution function analogue to the method explained in Section 10.4.2 is performed. All contributing simulated samples have been reweighted to the central value of the PDF sets CT10, MSTW2008nlo68cl and NNPDF2.3. The maximum difference among the three scenarios is taken as the systematic uncertainty:

$$\begin{aligned}\Delta m_{\mathrm{top}} = \max(&|\langle m_{\mathrm{top}}^{\mathrm{CT10}}\rangle - \langle m_{\mathrm{top}}^{\mathrm{NNPDF}}\rangle|,\\ &|\langle m_{\mathrm{top}}^{\mathrm{CT10}}\rangle - \langle m_{\mathrm{top}}^{\mathrm{MSTW}}\rangle|,\\ &|\langle m_{\mathrm{top}}^{\mathrm{NNPDF}}\rangle - \langle m_{\mathrm{top}}^{\mathrm{MSTW}}\rangle|).\end{aligned} \tag{10.14}$$

Relative top normalisation In Section 10.1 it was explained that the signal is defined as the combination of $t\bar{t}$ and single top-quark production processes. In Figure 10.2(a) the differences in the shape of the $m(\ell b)$ distribution is shown. This could result in a bias in the estimation of the top quark mass if the ratio of t-channel production to the other processes is not estimated correctly. Both contributions have been measured and correction factors were obtained with uncertainties [71]. The nominal correction factors, one factor for the t-channel production and one for the other processes including top quarks, are used in this analysis, as explained in Section 8.1.2.

To estimate the effect of a possible bias, the correction factors are shifted by one standard deviation ($\pm 1\sigma$). This means that four combinations of correlated and anti-correlated variations are possible. The systematic uncertainty due to the relative top normalisation is given by the maximum difference of the four variations with respect to the nominal scenario:

$$\begin{aligned}\Delta m_{\text{top}} = \max(&|\langle m_{\text{top}}^{+1\sigma,+1\sigma}\rangle - \langle m_{\text{top}}^{\text{nominal}}\rangle|, |\langle m_{\text{top}}^{-1\sigma,-1\sigma}\rangle - \langle m_{\text{top}}^{\text{nominal}}\rangle|,\\ &|\langle m_{\text{top}}^{+1\sigma,-1\sigma}\rangle - \langle m_{\text{top}}^{\text{nominal}}\rangle|, |\langle m_{\text{top}}^{-1\sigma,+1\sigma}\rangle - \langle m_{\text{top}}^{\text{nominal}}\rangle|).\end{aligned} \tag{10.15}$$

Table 10.6: Systematic uncertainties due to the background on the measurement of the top quark mass for variations at the level of $\pm 1\sigma$.

Systematic	Δm_{top}[GeV] electron	Δm_{top}[GeV] muon	Δm_{top}[GeV] electron+muon
W+jets normalisation	0.20	0.54	0.40
W+jets shape	0.07	0.12	0.06
Z+jets normalisation	0.26	0.23	0.25
QCD-multijet normalisation	0.37	0.11	0.22
rel. top normalisation	0.19	0.22	0.22
Background	0.53	0.65	0.56

10.4.4 Simulation statistics

Pseudo-experiments are performed in which the bin content of each bin of the $m(\ell b)$-distribution is fluctuated according to its true statistical uncertainty. The width of the resulting distribution of m_{top} is giving an uncertainty due to the limited size of the simulated sample. The procedure to estimate the systematic uncertainty is visualised in Figure 10.13 for the combined electron+muon channel.

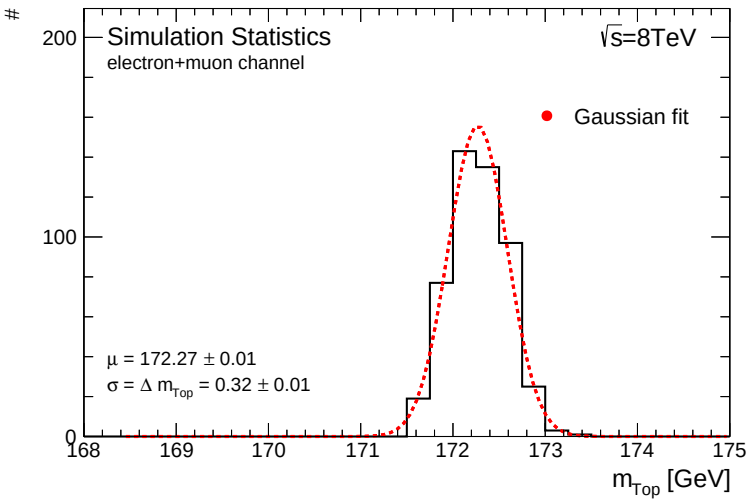


Figure 10.13: Systematic uncertainty due to the limited size of the simulated sample. The uncertainty is given by the width of the fitted Gaussian to the distribution of m_{top}, which is estimated using pseudo-experiments.

10.4.5 Summary of systematic uncertainties

In Table 10.7 the resulting systematic uncertainties on the measurement of the top quark mass for the combined electron+muon channel and each channel separately are summarised. The total uncertainty is dominated by the jet energy scale and t-channel modelling in simulation. Other important contributions are coming from the background normalisation.

Table 10.7: Summary of the systematic uncertainties on the measurement of the top quark mass. If possible the variation corresponds to a variations at the level of $\pm 1\sigma$.

Source of uncertainty	$\lvert\Delta m_{\text{top}}\rvert$ [GeV] electron	$\lvert\Delta m_{\text{top}}\rvert$ [GeV] muon	$\lvert\Delta m_{\text{top}}\rvert$ [GeV] electron+muon
$E_{\text{T}}^{\text{miss}}$ cellout + softjet resolution	0.03	0.09	0.04
$E_{\text{T}}^{\text{miss}}$ cellout + softjet scale	0.15	0.28	0.22
Total jet energy scale	1.29	1.64	1.49
Jet energy resolution	0.21	0.12	0.03
Jet vertex fraction	0.08	0.11	0.08
Lepton ID efficiency	0.00	0.00	0.00
Lepton reconstruction efficiency	0.00	0.00	0.00
Lepton trigger efficiency	0.06	0.00	0.02
Electron energy resolution	0.02	0.00	0.01
Electron energy scale	0.68	0.00	0.31
Muon energy resolution	0.00	0.15	0.08
Muon energy scale	0.00	0.14	0.07
Flavour tagging efficiency	0.26	0.36	0.30
Background	0.53	0.65	0.56
t-channel colour reconnection	0.12	0.48	0.31
t-channel underlying event	0.10	0.15	0.04
t-channel MC generator	0.71	0.65	0.66
$t\bar{t}$ MC generator	0.20	0.28	0.21
$t\bar{t}$ ISR/FSR	0.18	0.16	0.16
$t\bar{t}$ colour reconnection	0.07	0.36	0.22
$t\bar{t}$ underlying event	0.11	0.12	0.11
Wt-channel MC generator	0.26	0.25	0.08
s-channel MC generator	0.23	0.00	0.08
Proton PDF	0.14	0.07	0.05
Simulation statistics	0.44	0.45	0.32
Total systematic uncertainty	1.87	2.14	1.89

CHAPTER 11

Results and conclusion

The distribution of $m(\ell b)$ in the full dataset together with the corresponding fitted probability density functions for the signal and background is shown in Figure 11.1. The inlet shows the corresponding $-2\ln(L)$ profile as a function of the top quark mass with the horizontal line at $-2\ln(L) = 1$ corresponding to $\pm 1\sigma$ of the statistical uncertainty.

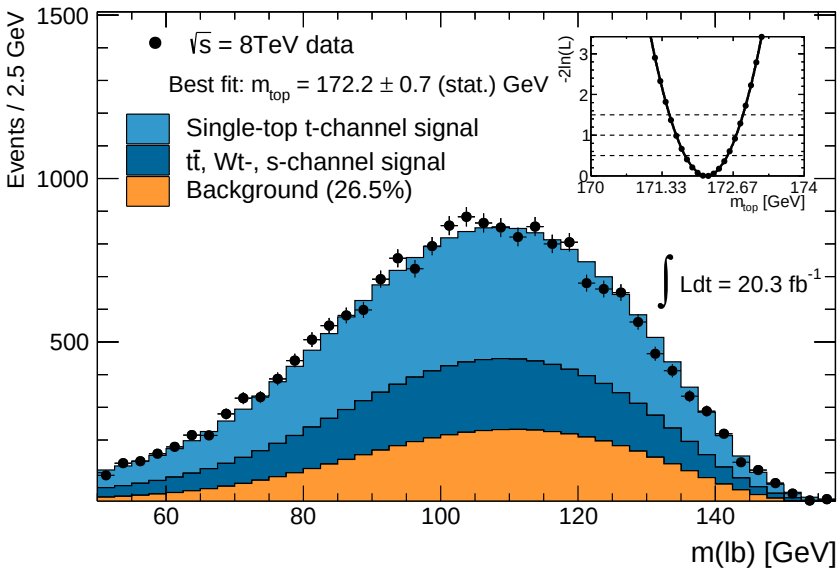


Figure 11.1: Fitted $m(\ell b)$ distribution in data to estimate m_{top} in the combined electron+muon channel. The inlet shows the $-2\ln\mathcal{L}$ profile as a function of the top quark mass. The relative mixture of the dominant single top t-channel production process (light blue) and the other top processes dominated by $t\bar{t}$ (dark blue) corresponds to what is given in Table 9.2.

The relative mixture of the dominant single top t-channel production process (light blue) and the other top processes dominated by $t\bar{t}$ (dark blue) corresponds to what is given in Table 9.2.

The result of the fit to 2012 ATLAS data in topologies enhanced with single top-quarks produced in the t-channel is:

$$\boxed{m_{\text{top}} = [172.2 \pm 0.7\,(\text{stat.}) \pm 1.9\,(\text{syst.})]\ \text{GeV}}$$

The result is obtained by adding the two disjoint contributions from the electron and muon channel and building a combined parametrisation of the signal and background templates. Those are used in the maximum likelihood fit to data to estimate the value of m_{top} following the procedure explained in Chapter 10.

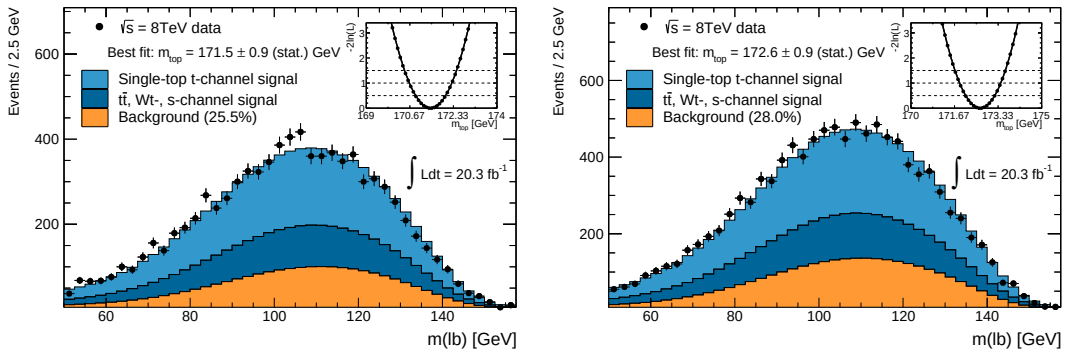
The given value of the total systematic uncertainty corresponds to the quadratic sum of the components that are summarised in Table 11.1. The uncertainties for the different sources are grouped together and listed with the fitted top quark mass and its statistical uncertainty.

Table 11.1: Measured value of m_{top} and uncertainties on the measurement for the systematic variations explained in Section 10.4

	Value [GeV]
Measured value m_{top}	172.2
Statistical uncertainty	0.7
Background	0.56
Electron uncertainties	0.31
Flavour tagging efficiency	0.30
Jet energy scale	1.49
Jet energy resolution	0.03
Jet vertex fraction	0.08
Missing transverse momentum	0.22
Muon uncertainties	0.11
Simulation modelling	0.84
Simulation statistics	0.32
Total systematic uncertainty	1.89
Total uncertainty	2.02

In Figure 11.2 the distributions of $m(\ell b)$ and the fitted values of m_{top} are shown separately in the electron (a) and muon channel (b). The measured value of m_{top} in the electron channel is slightly lower compared to the one in the muon channel. Also

the fitted background is lower in the electron channel which is in agreement with the prediction given in Table 9.2. Within the statistical uncertainty both channels agree well with each other and are contributing in a similar extent to the final result in the combined electron+muon channel.



(a) Electron channel

(b) Muon channel

Figure 11.2: Fitted $m(\ell b)$ distribution in data to estimate m_{top} in the electron channel (a) and the muon channel (b) separately. The inlet is showing the $-2\ln\mathcal{L}$ profile as a function of the top quark mass.

The result is dominated by the systematic uncertainties, with the largest contributions coming from the jet energy scale uncertainty and the simulation of the t-channel signal process. Good prospects to further reduce the total uncertainty are expected. This is the case because the jet energy calibration currently used for the 2012 dataset is still preliminary. It is expected that the contribution of the jet energy scale uncertainty will decrease. Particularly the component due to the η-intercalibration method is expected to be significantly reduced for jets in the forward region of the detector.

Furthermore, in the future it will be possible to use a NLO generator POWHEG or aMC@NLO instead of the $2 \to 2$ and $2 \to 3$ ACOT-matched ACERMC generator for the single-top t-channel simulation. This has potential to lead to a smaller systematic uncertainty due to the t-channel MC generator.

The measurement presented in this thesis is sensitive to different sources of systematic uncertainties compared to the other channels. The dominant single-top t-channel process is a process induced by the weak interaction. Therefore, the analysis is less depending on colour flow in the final state leading to smaller systematic uncertainties. Additionally, the selection relies on exactly one tagged plus untagged jet and only one neutrino that has to be reconstructed. This leads to a better mass resolution compared to final states where

ambiguities occur. However, larger contributions from irreducible background mainly from W+jets and QCD-multijet processes have to be taken into account.

The statistical uncertainty of 0.7 GeV is similar compared to the ATLAS m_{top} measurement in the dileptonic $t\bar{t}$ decay channel at 7 TeV [198]. It still does not reach the precision of the measurement done in the $t\bar{t}$ l+jets channel at 7 TeV [85] but outdoes already the current ATLAS m_{top} measurement in the $t\bar{t}$ all hadronic decay channel at 7 TeV [199]. Regarding future combinations very good prospects are given as well. Due to the selection with exactly one lepton and exactly two jets there is no statistical correlation between the selected datasets and the datasets of the other three measurements. A comparison between the measured mass of the top quark in t-channel topologies and the three measurements in the different decay channels from ATLAS is shown in Figure 11.3. Also included are the most recent combinations which, as well as the single measurements, show very good agreement with the measured value of m_{top} presented in this thesis.

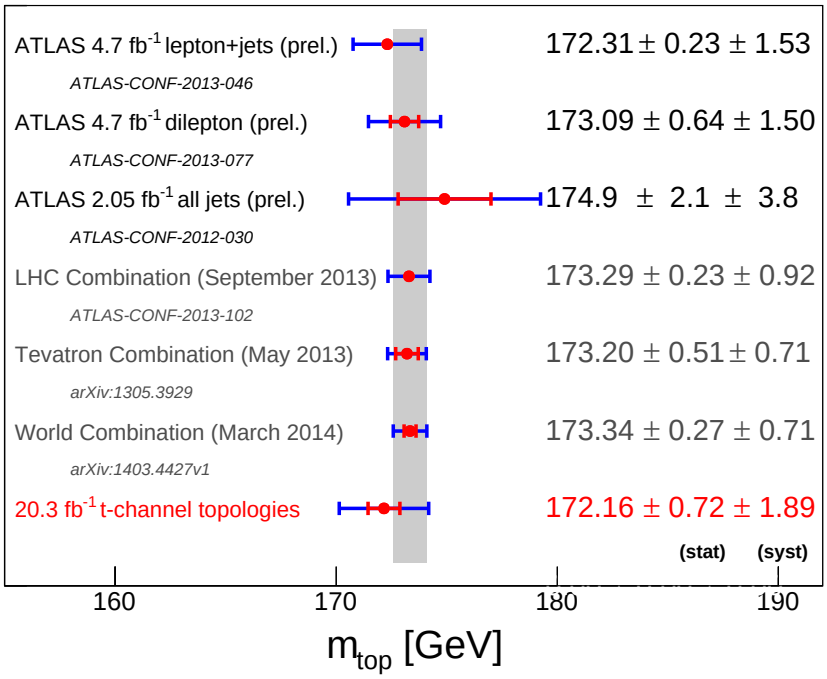
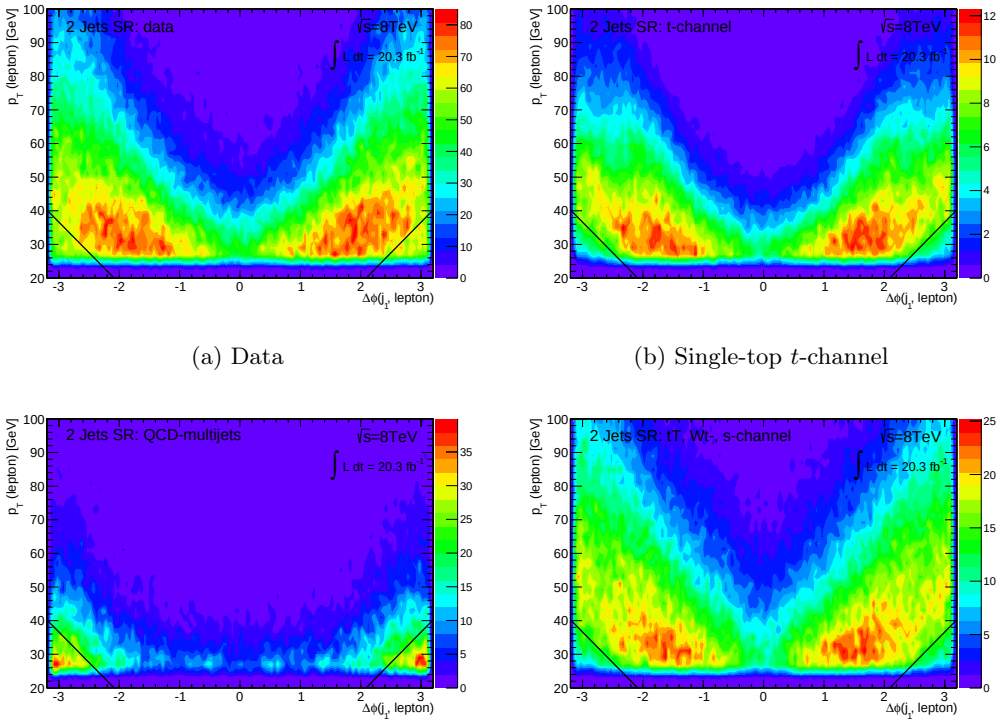


Figure 11.3: Measured top quark mass compared to the most recent ATLAS measurements in the different $t\bar{t}$ decay channels. Also shown are the newest combinations of measurements from the Tevatron, the LHC and both which is referred to as the world combinations. The latter combination is also shown as the grey band.

APPENDIX A

Appendix

A.1 Multijet veto for different processes



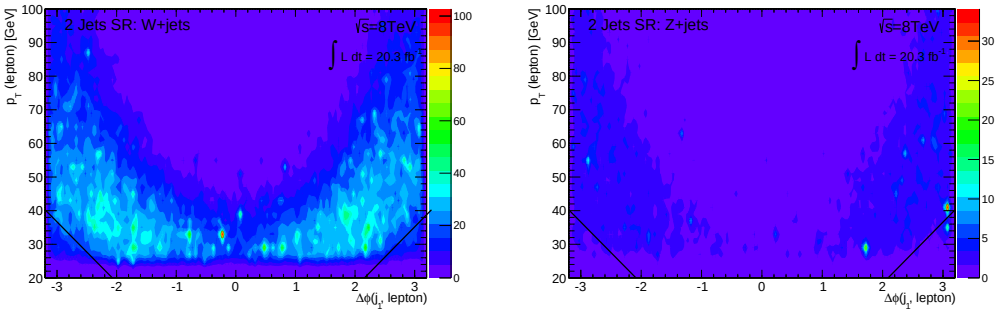
(a) Data

(b) Single-top t-channel

(c) QCD-multijets (jet-lepton and anti-muon)

(d) $t\bar{t}$, single-top Wt- and s-channel

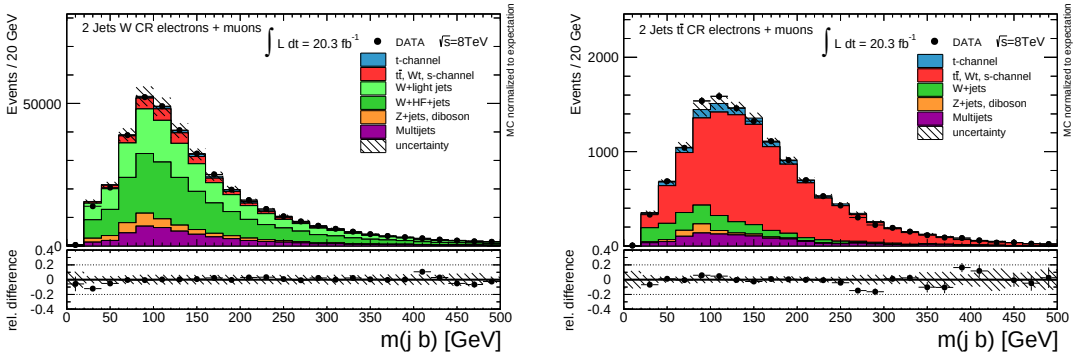
Figure A.1: The lepton p_{T} in dependence of the $\Delta\phi\,(j_1, \ell)$ is shown for the different contributing processes containing top quarks and the other background processes. The combined electron+muon channel in the signal region is shown.



(a) W+jets (b) Z+jets/diboson

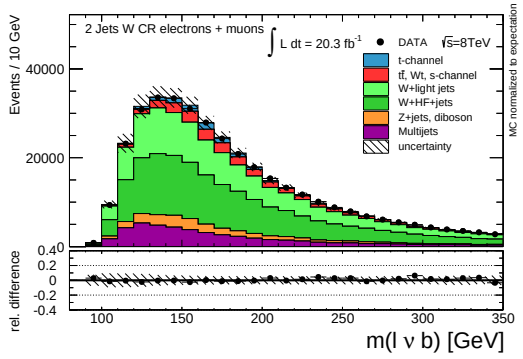
Figure A.2: The lepton p_{T} in dependence of the $\Delta\phi\,(j_1, \ell)$ is shown for the data and the single-top t-channel process. The combined electron+muon channel in the signal region is shown.

A.2 Modelling of the NN input variables

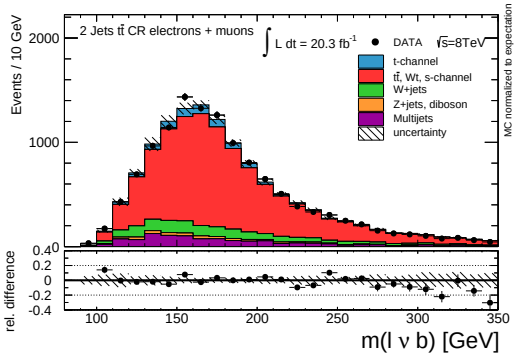


(a) Invariant mass of the two reconstructed jets (b) Invariant mass of the two reconstructed jets

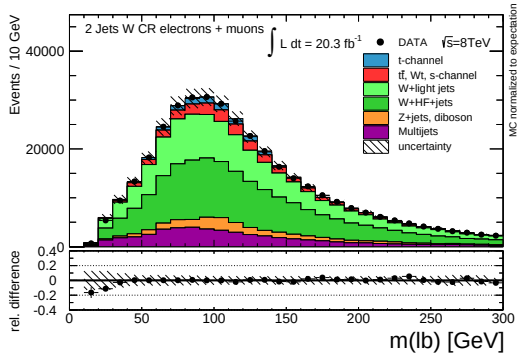
Figure A.3: Expected and observed distributions in the combined electron+muon channel of the input variables used in the neural network selection in the $W^{\pm}$-boson control region (left) and the $t\bar{t}$ control region (right). The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the QCD-multijet normalisation. In simulation a top quark mass of $m_{\mathrm{top}} = 172.5$ GeV is used.



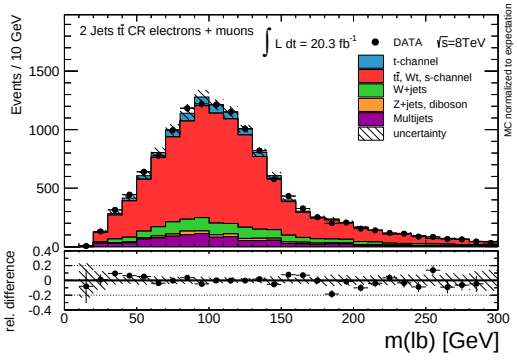
(a) Invariant mass of the reconstructed top quark



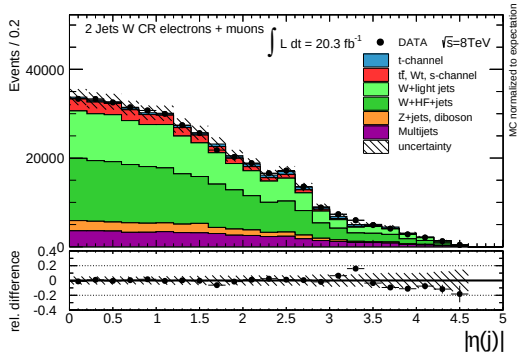
(b) Invariant mass of the reconstructed top quark



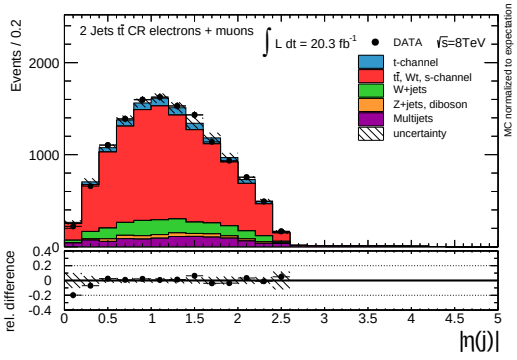
(c) Invariant mass of the lepton and b-tagged jet



(d) Invariant mass of the lepton and b-tagged jet

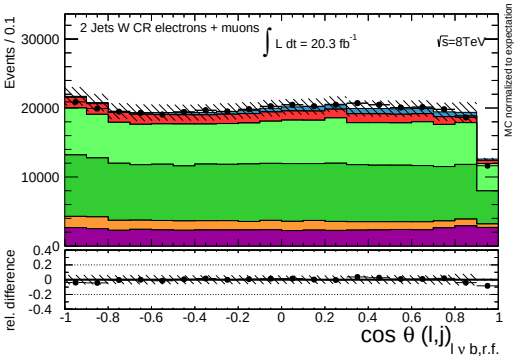


(e) Pseudorapidity of the reconstructed light jet

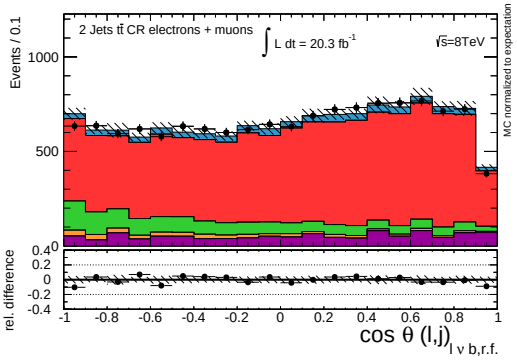


(f) Pseudorapidity of the reconstructed light jet

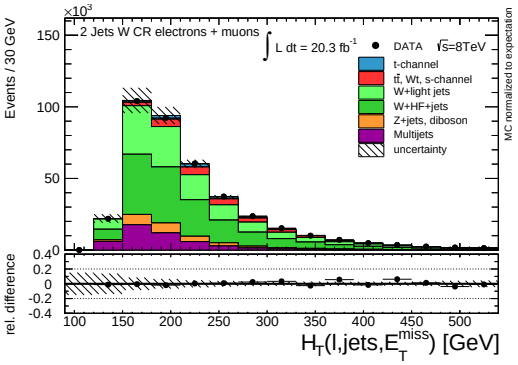
Figure A.4: Expected and observed distributions in the combined electron+muon channel of the input variables used in the neural network selection in the $W^{\pm}$-boson control region (left) and the $t\bar{t}$ control region (right). The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the QCD-multijet normalisation. In simulation a top quark mass of $m_{\text{top}} = 172.5$ GeV is used.



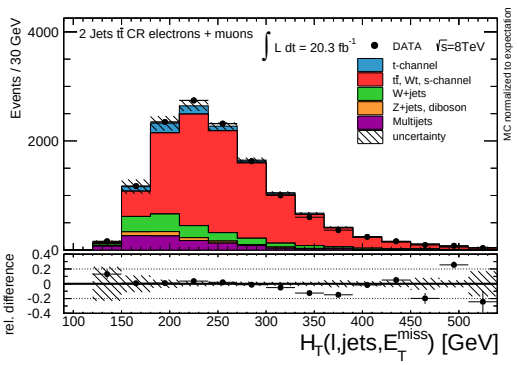
(a) Polarisation angle of the top quark



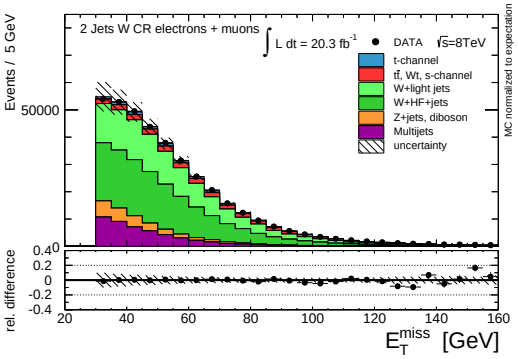
(b) Polarisation angle of the top quark



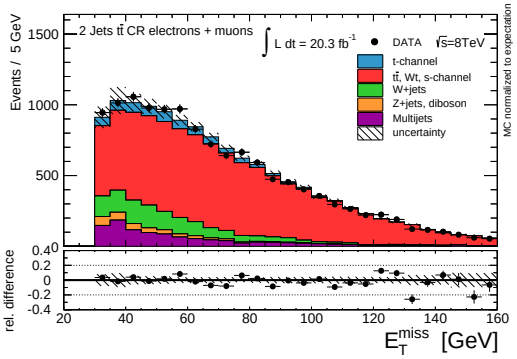
(c) Summed transverse momentum of all reconstructed objects in the event



(d) Summed transverse momentum of all reconstructed objects in the event

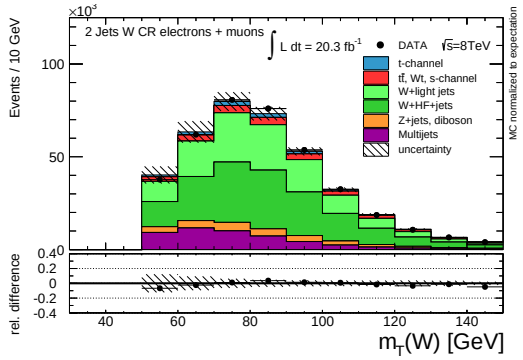


(e) Missing transverse momentum

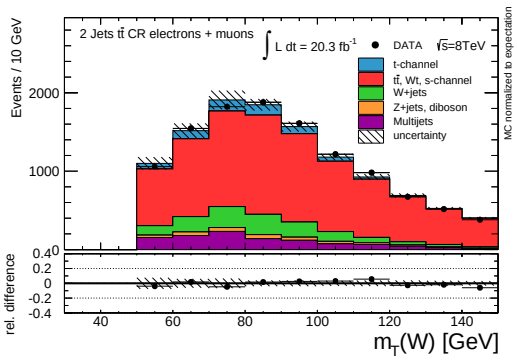


(f) Missing transverse momentum

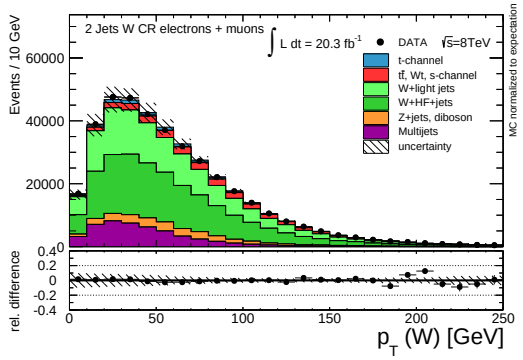
Figure A.5: Expected and observed distributions in the combined electron+muon channel of the input variables used in the neural network selection in the $W^{\pm}$-boson control region (left) and the $t\bar{t}$ control region (right). The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the QCD-multijet normalisation. In simulation a top quark mass of $m_{\text{top}} = 172.5$ GeV is used.



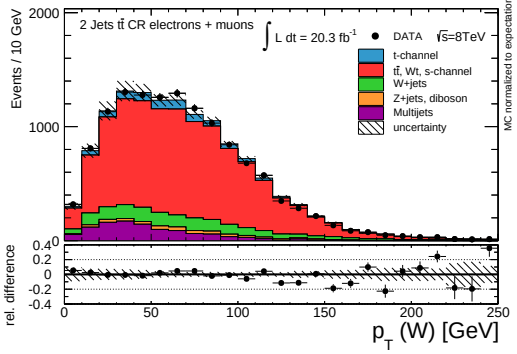
(a) Transverse $W^{\pm}$-boson mass



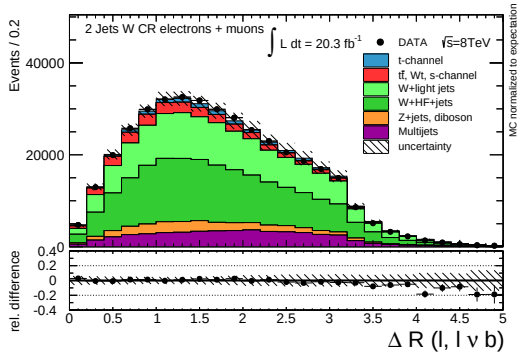
(b) Transverse $W^{\pm}$-boson mass



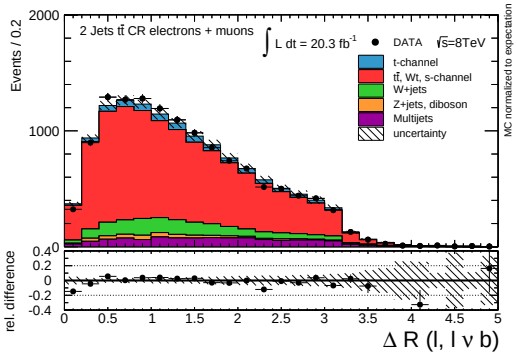
(c) Transverse momentum of the $W^{\pm}$-boson



(d) Transverse momentum of the $W^{\pm}$-boson

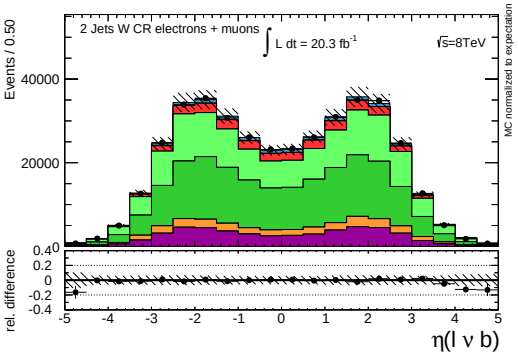


(e) ΔR between the top quark and the lepton

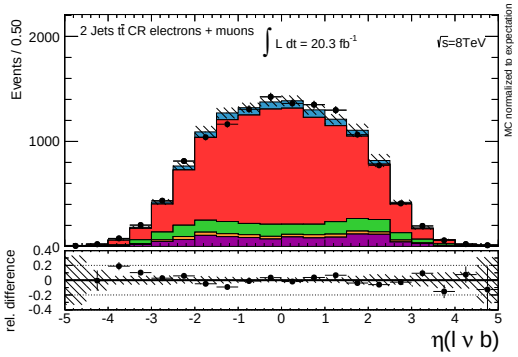


(f) ΔR between the top quark and the lepton

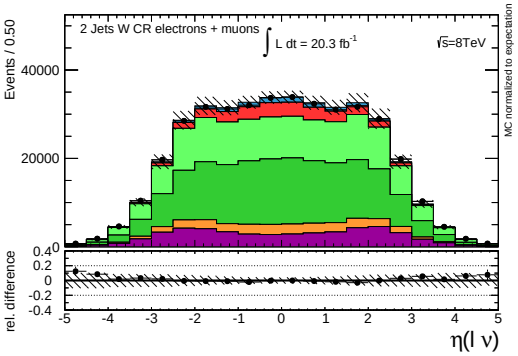
Figure A.6: Expected and observed distributions in the combined electron+muon channel of the input variables used in the neural network selection in the $W^{\pm}$-boson control region (left) and the $t\bar{t}$ control region (right). The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the QCD-multijet normalisation. In simulation a top quark mass of $m_{\text{top}} = 172.5$ GeV is used.



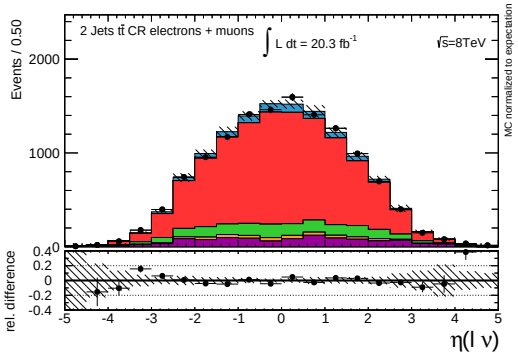
(a) Pseudorapidity of reconstructed top quark



(b) Pseudorapidity of reconstructed top quark



(c) Pseudorapidity of the reconstructed $W^{\pm}$-boson



(d) Pseudorapidity of the reconstructed $W^{\pm}$-boson

Figure A.7: Expected and observed distributions in the combined electron+muon channel of the input variables used in the neural network selection in the $W^{\pm}$-boson control region (left) and the $t\bar{t}$ control region (right). The hatched bands indicate the size of the statistical uncertainty of the simulated sample and the uncertainty on the QCD-multijet normalisation. In simulation a top quark mass of $m_{\text{top}} = 172.5$ GeV is used.

A.3 Event displays of t-channel candidate events

The following pages are showing two proton-proton collision events taken during 2012 in the $\sqrt{s} = 8$ TeV run. Both events passed the event selection of the signal region that was explained in Chapter 8. This means that the events contain two jets with exactly one of them being b-tagged, $E_{\mathrm{T}}^{\mathrm{miss}}$, and exactly one reconstructed lepton. In Figures A.8 and A.9 an event with an electron and in Figures A.10 and A.11 an event with a muon can be seen.

For both events the output value of the neural network, explained in Chapter 9, is larger than 0.95. This means that events with a high probability to contain a single top-quark produced in the t-channel are shown. This is the main signal process in this thesis. In detail the event displays are showing the following:

- **Figures A.8 and A.9:** The event shows two jets with the untagged jet reconstructed in the forward region of the detector. The b-tagged jet is located more central in the barrel region. This b-jet was reconstructed by as well the electromagnetic and hadronic calorimeter clearly visible as coloured calorimeter cells in the r-φ-plane in both figures. The untagged jet was reconstructed mostly from one cell in the forward calorimeter nicely visible in Figure A.9.
  The electron is measured almost exclusively in the electromagnetic calorimeter visible in the bottom right of Figure A.9 and the inner detector. The blue track is the reconstructed electron track also shown in the zoom together with the Pixel and SCT hits in the bottom left of Figure A.9. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is shown in the r-φ-plane of Figure A.8 as a white dotted line.

- **Figures A.10 and A.11:** Two jets are reconstructed with one of them in forward direction. The b-tagged jet is seen in the central detector region from cells of the electromagnetic and hadronic calorimeter while the untagged jet is measured by the forward calorimeter.
  In Figure A.10 the muon track is shown in blue that was reconstructed as well with the inner detector and the muon spectrometer. In Figure A.11 the same muon track is shown in red in the four different views. In the bottom left also the muon chambers that have been hit are drawn (MDT: green, TGC: dark red). In the same plot at the top right the SCT hits and TRT drift circles are drawn together with the muon track and the jets. The top left shows the Pixel and SCT hits in the barrel region together with the muon track and the two jets and associated tracks.

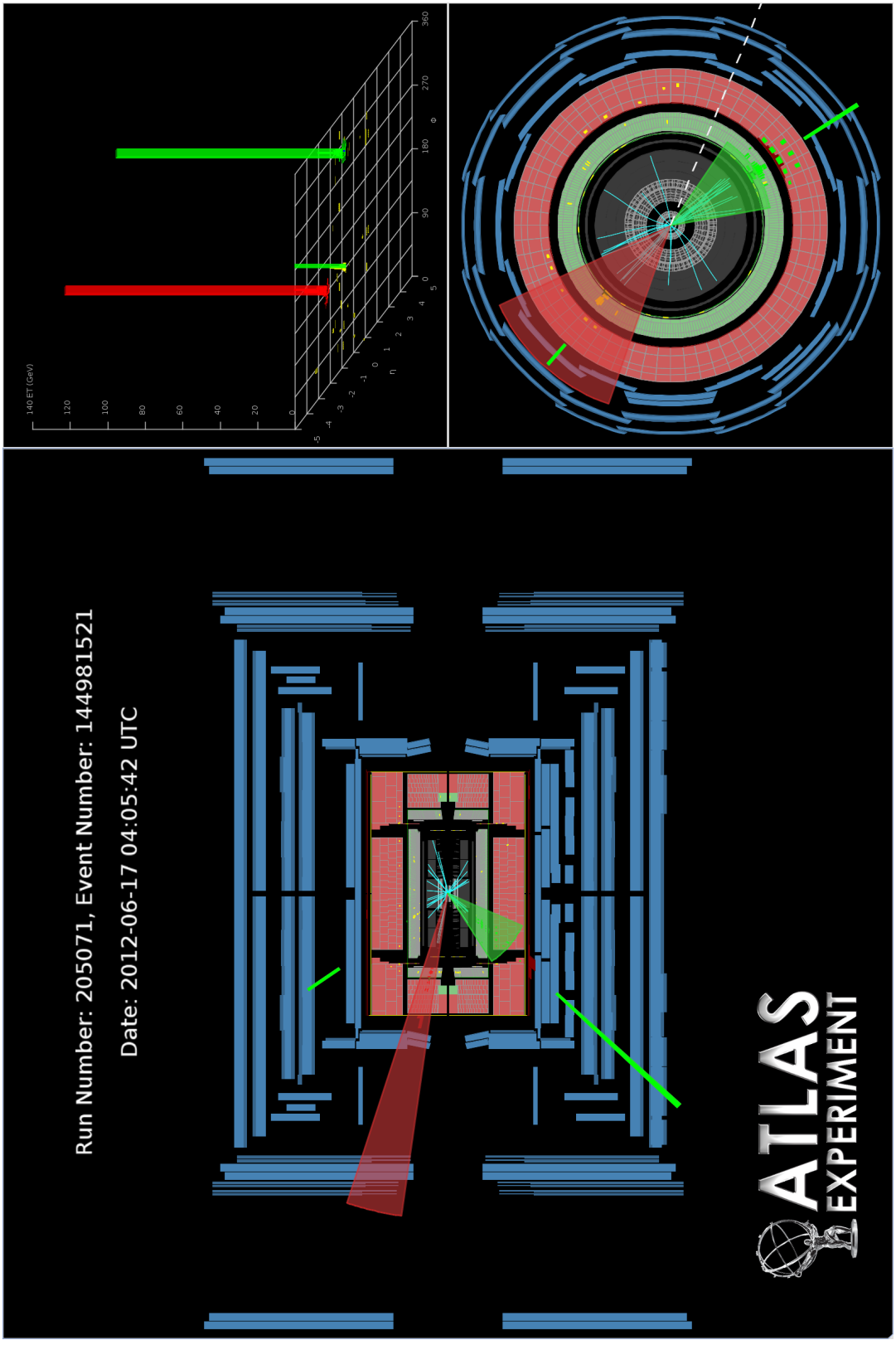


Figure A.8: Event display of a single-top t-channel candidate event in the electron channel. Shown is the same event as in Figure A.9.

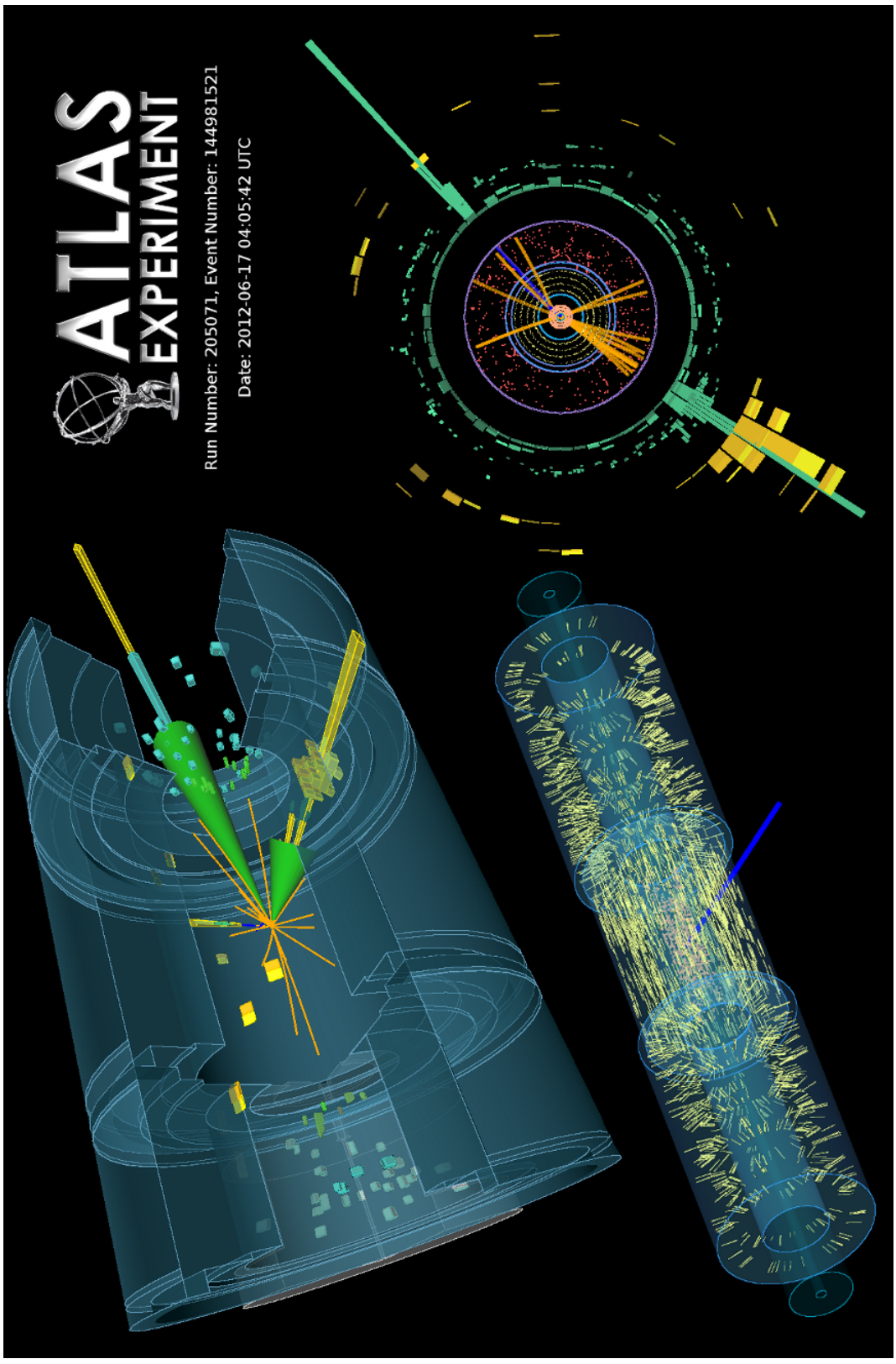


Figure A.9: Different views of a single-top t-channel candidate event in the electron channel. Shown is the same event as in Figure A.8.

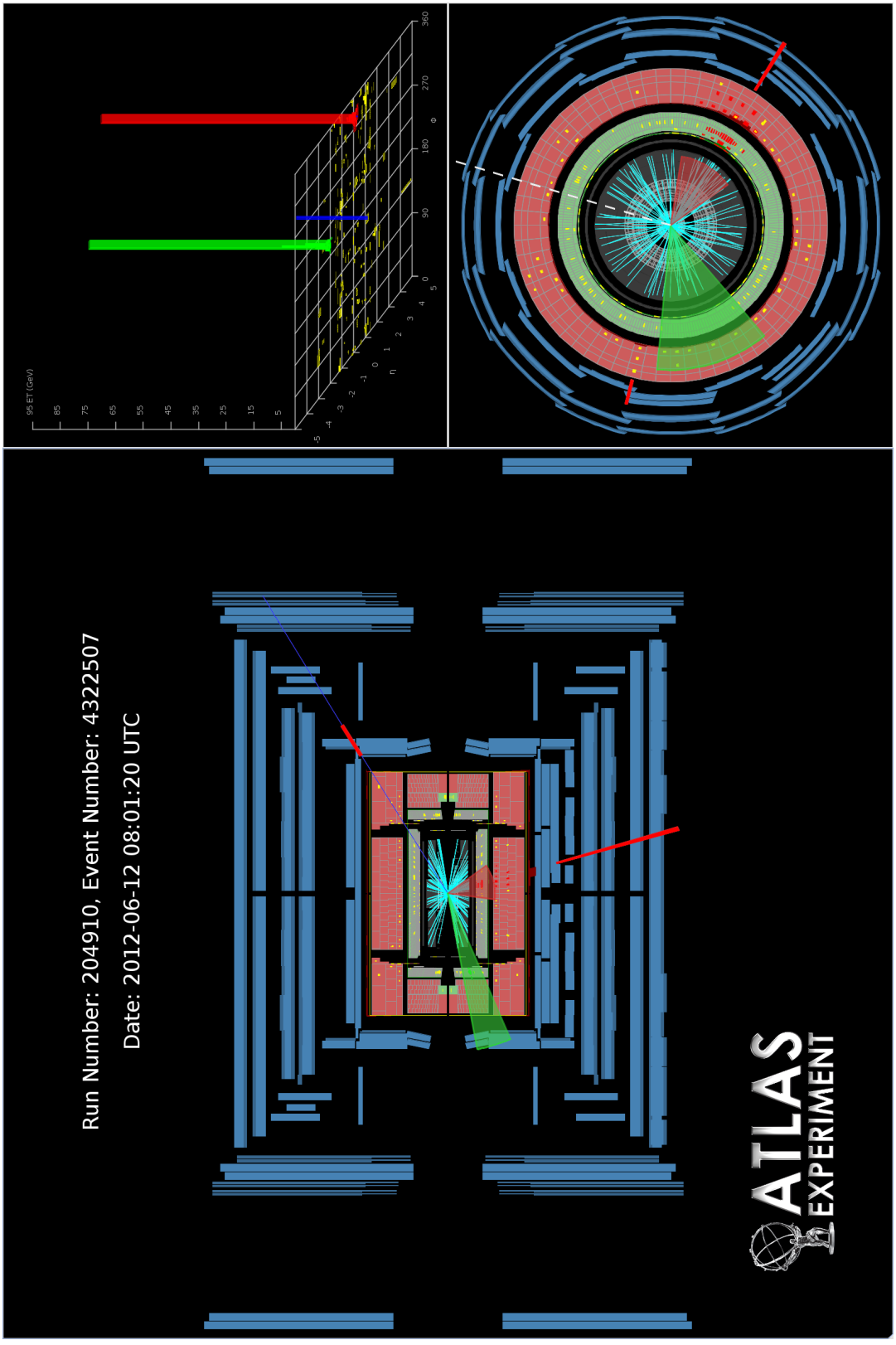


Figure A.10: Event display of a single-top t-channel candidate event in the muon channel. Shown is the same event as in Figure A.11.

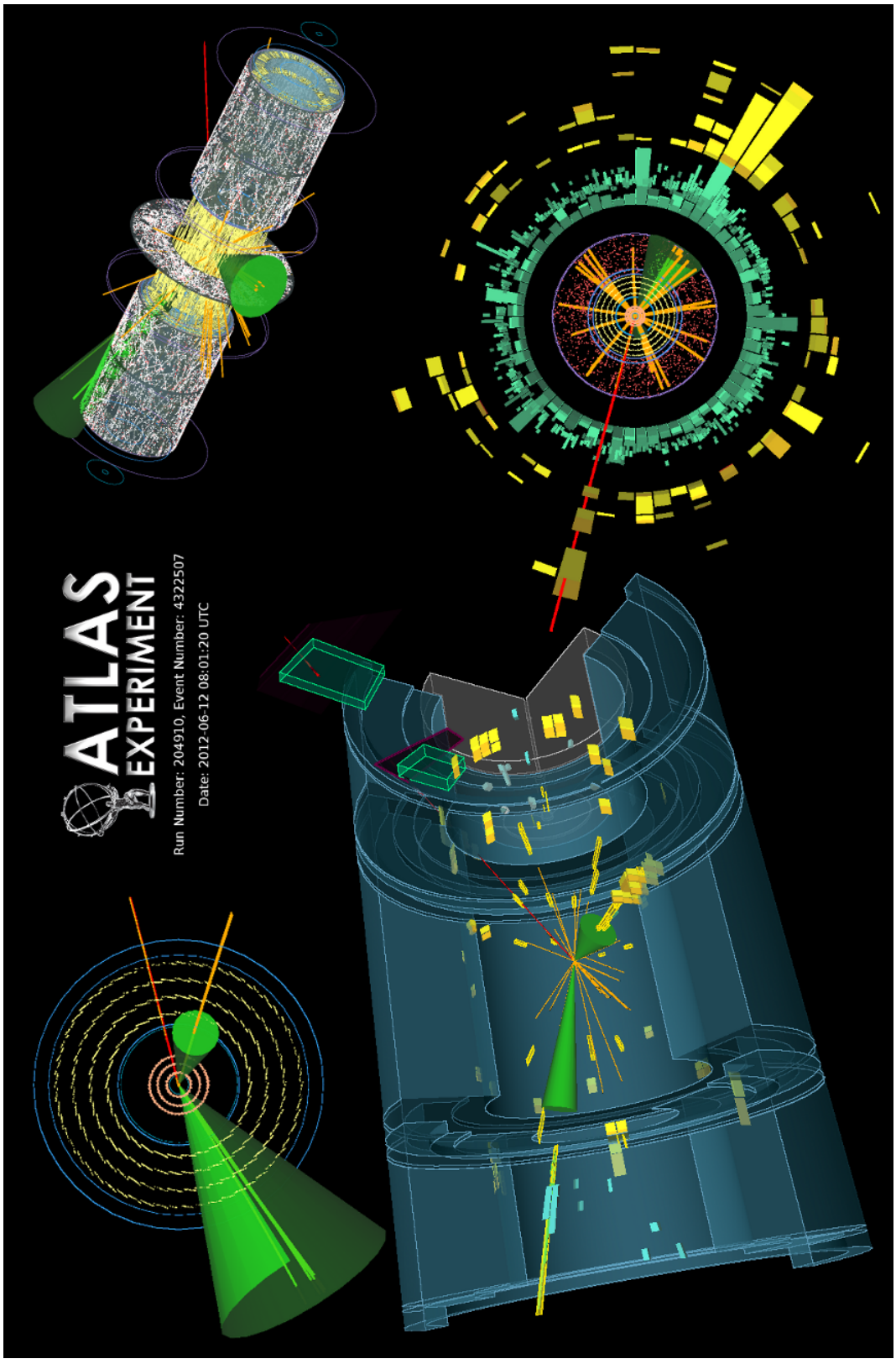


Figure A.11: Different views of a single-top t-channel candidate event in the muon channel. Shown is the same event as in Figure A.10.

A.4 Correlation matrix of the NN input variables

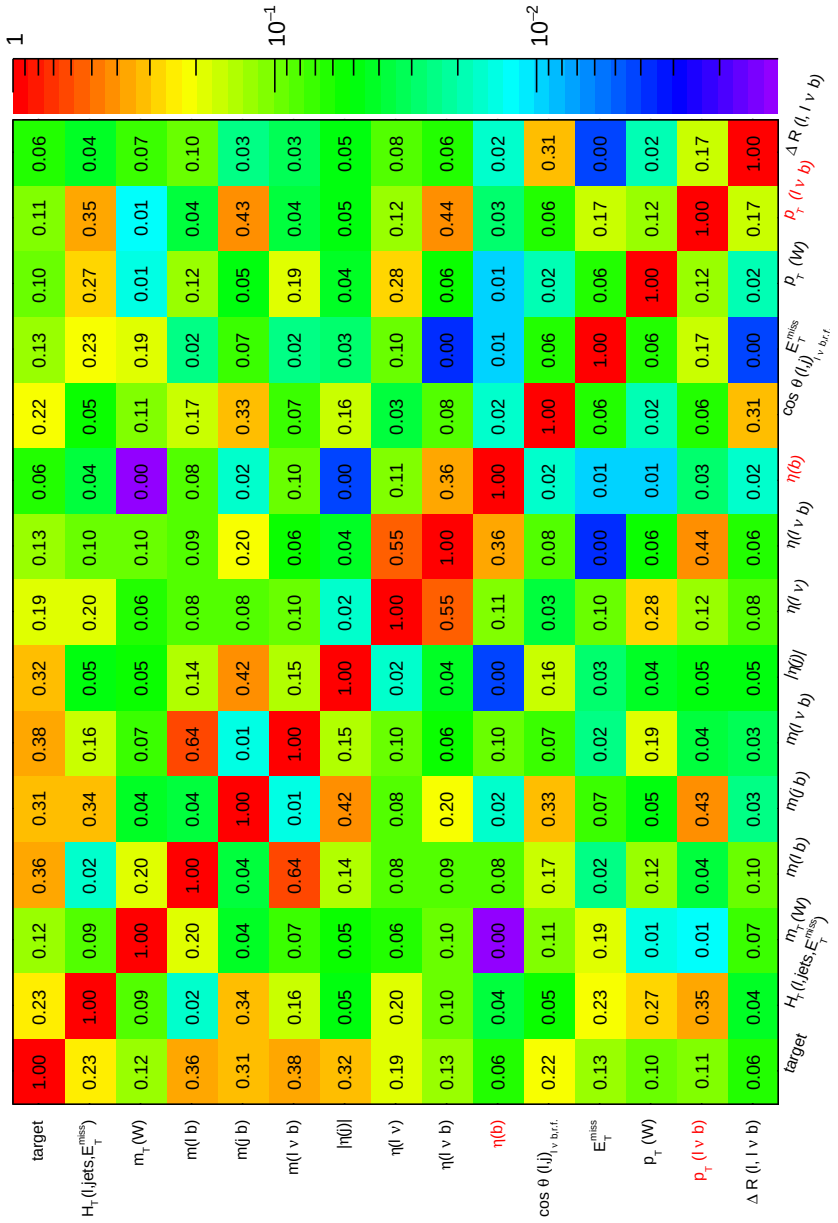
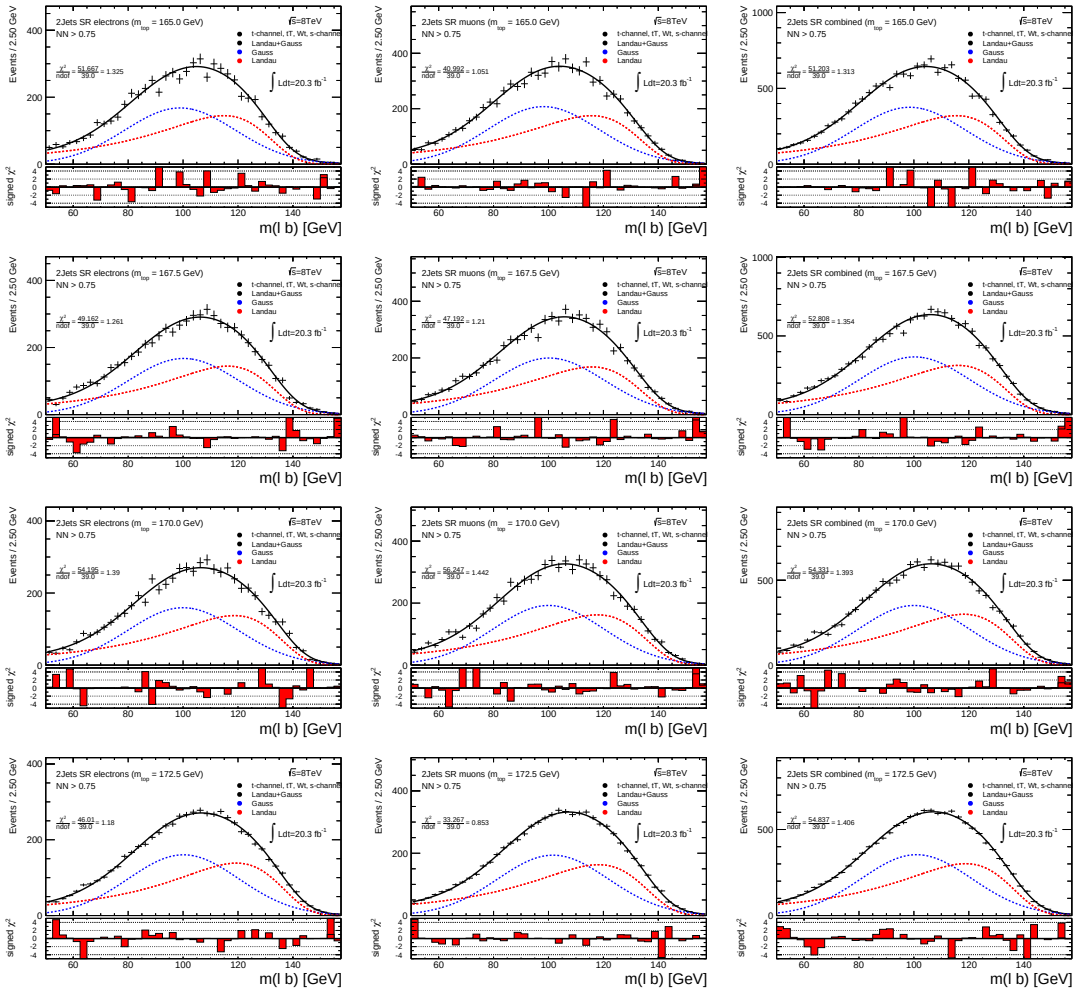


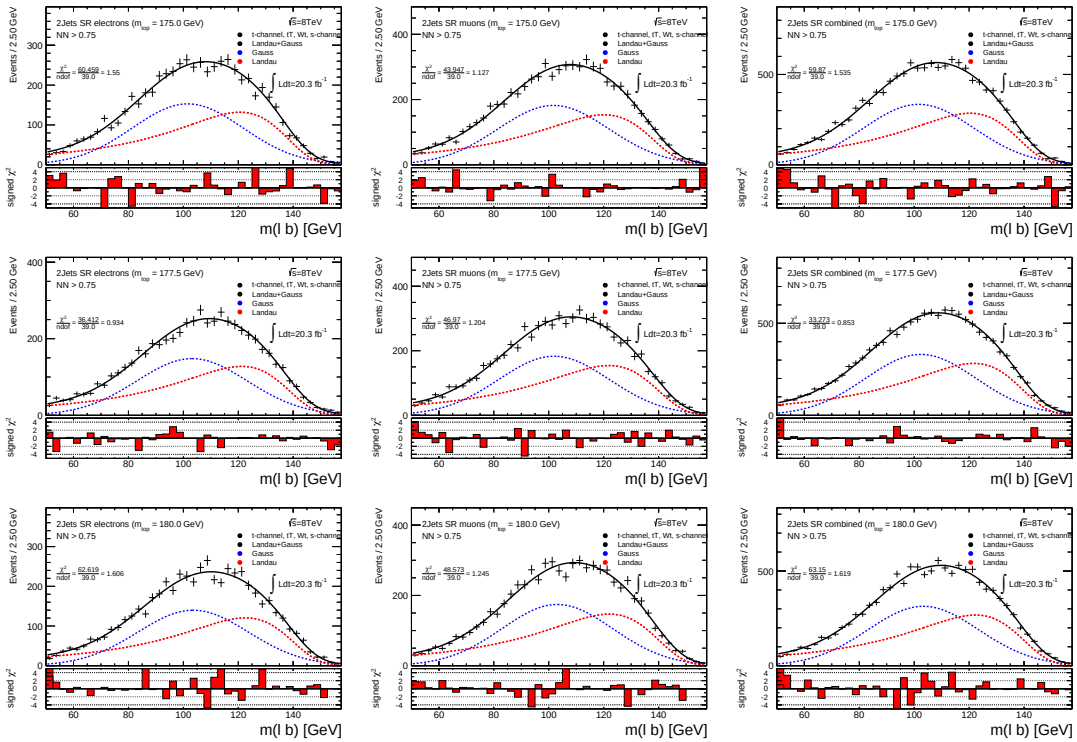
Figure A.12: Matrix showing the absolute correlation between the neural network input variables and the target variable. Variables with $\sigma^{\text{iter}} < 3$ estimated with the procedure explained in Section 9.3.1 are marked in red and automatically removed. They are not used in the final neural network.

A.5 Fitted $m(\ell b)$ distributions using different mass points



a) b) c)

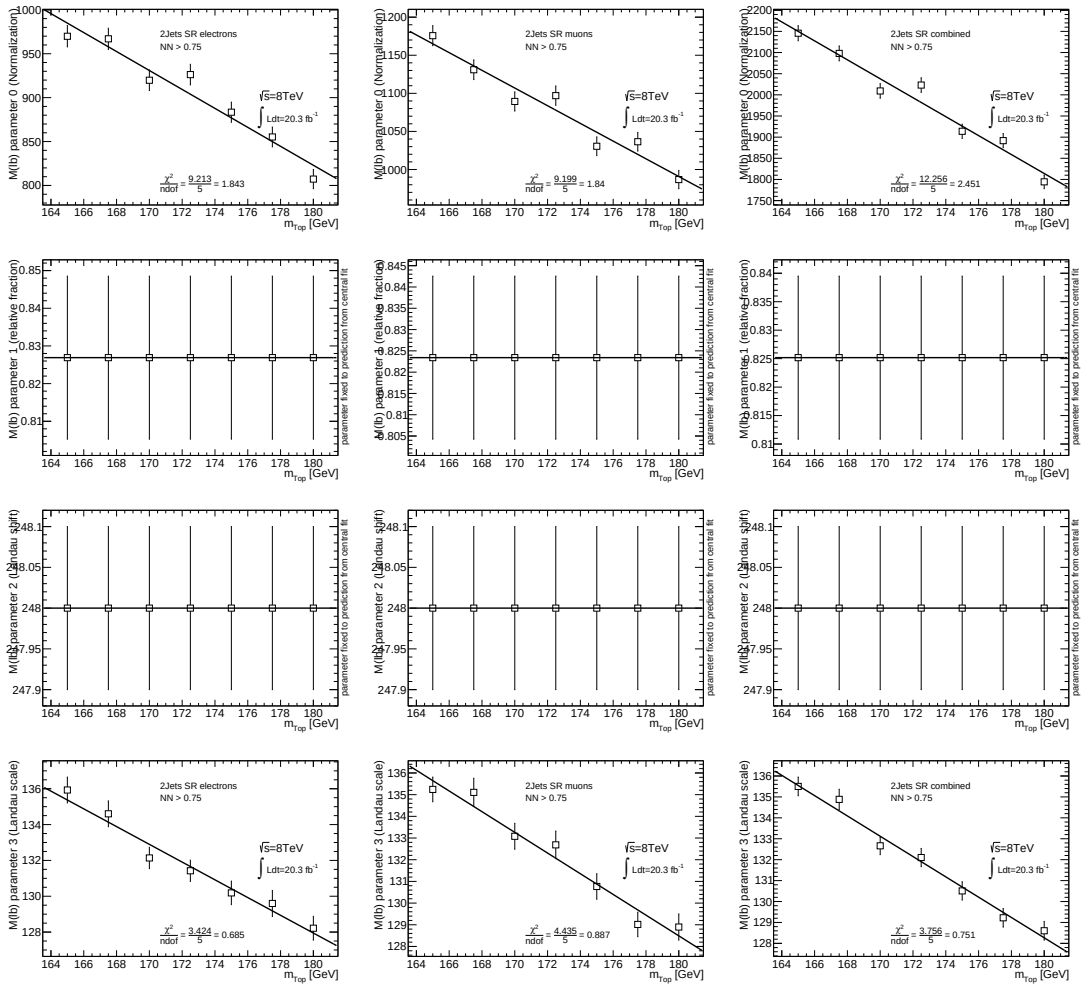
Figure A.13: Signal $m(\ell b)$ distribution fitted with a sum of Landau and Gaussian distribution defined in Equation (10.1) for the different top quark mass points (top: $m_{\text{top}} = 165.0$ GeV to bottom: $m_{\text{top}} = 172.5$ GeV). The calibration curves displayed in Figures A.15 and A.16 are based on these fits. The electron a), muon b) and combined channel c) are shown.



a) b) c)

Figure A.14: Signal $m(\ell b)$ distribution fitted with a sum of Landau and Gaussian distribution defined in Equation (10.1) for the different top quark mass points (top: $m_{\text{top}} = 175.0$ GeV to bottom: $m_{\text{top}} = 180.0$ GeV). The calibration curves displayed in Figures A.15 and A.16 are based on these fits. The electron a), muon b) and combined channel c) are shown.

A.6 Signal fit parameters depending on m_{top}



a) b) c)

Figure A.15: Parameters $p_0 - p_3$ (top to bottom) from Equation (10.2) that are used for the determination of m_{top} are shown for the electron a), muon b) and combined channel c). The error bars indicate the one standard deviation uncertainty of the fitted values. Parameter p_0 is used for the overall normalisation only and the parameters p_1, p_2 are fixed to the prediction from the central mass point at $m_{\mathrm{top}} = 172.5$ GeV.

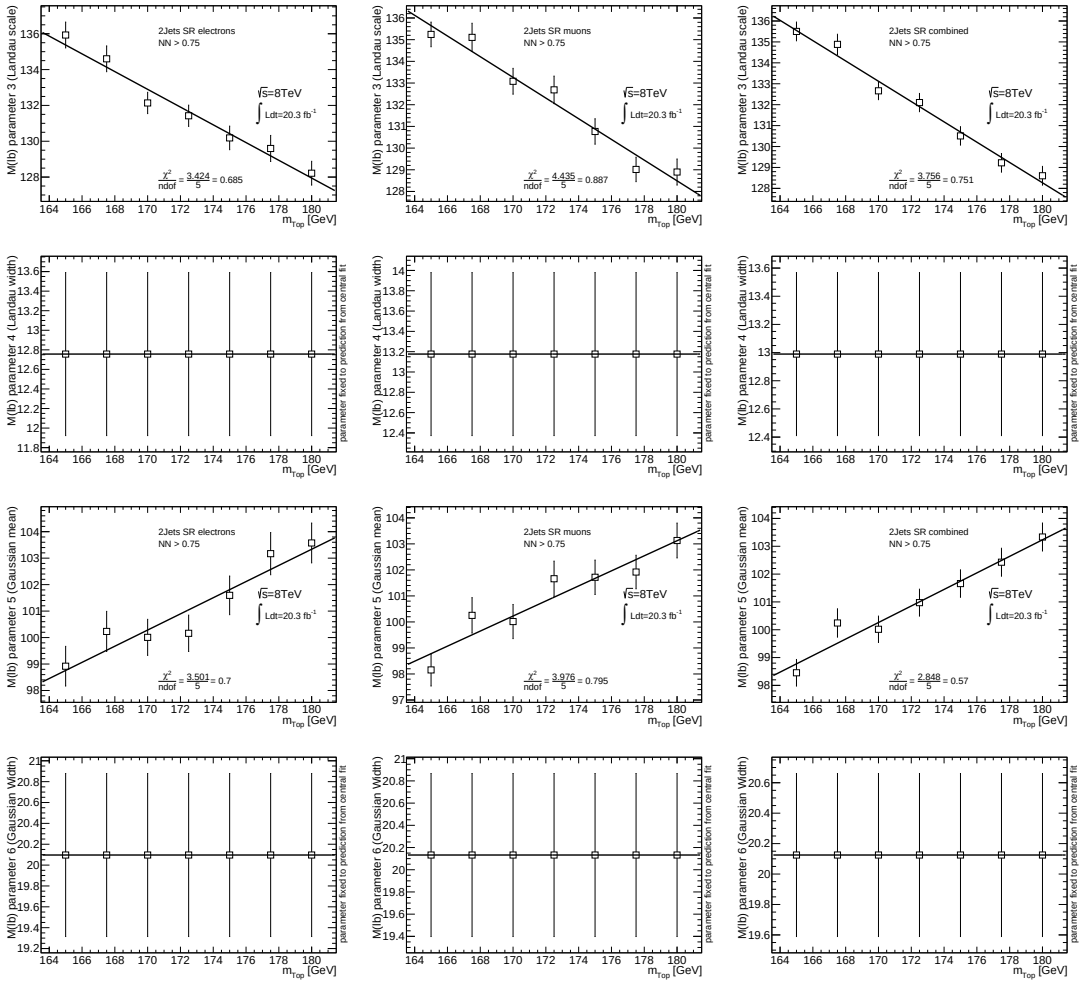
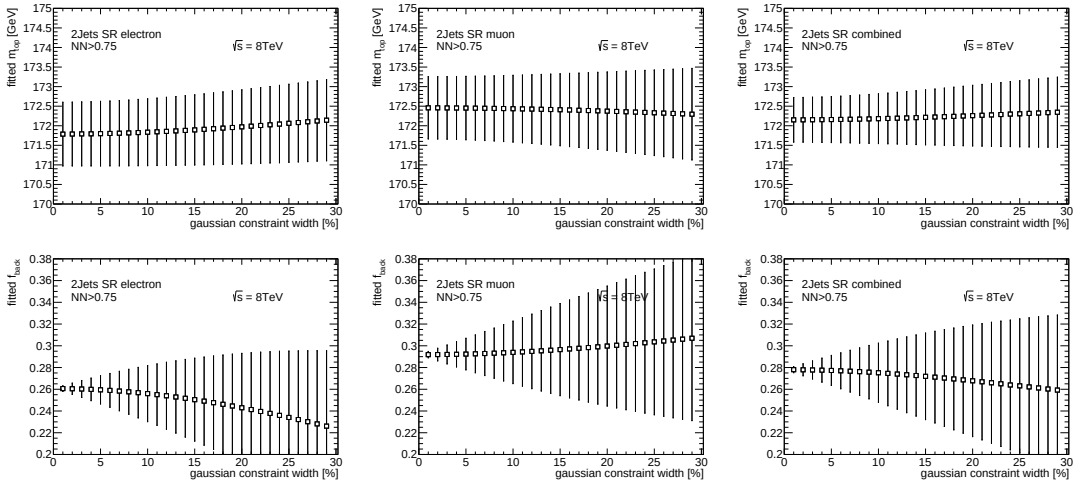


Figure A.16: Parameters $p_4 - p_6$ (top to bottom) from Equation (10.2) that are used for the determination of m_{top} are shown for the electron a), muon b) and combined channel c). The error bars indicate the one standard deviation uncertainty of the fitted values. The parameters p_4 and p_6 are fixed to the prediction from the central mass point at $m_{\text{top}} = 172.5$ GeV.

A.7 Background fraction constraint

In Figure A.17 a variation of the "strength" of the background fraction constraint, explained in Section 10.2.3, is shown. This strength corresponds to the width that is used for the Gaussian distribution. The theoretical uncertainty on the background fraction is roughly 10%-15% (relative uncertainty), see Table 9.2, and this corresponds to the value of 10-15% on the x-axis. The plot shows that either a width of 0 (which corresponds to fixing the background fraction) or a width of 30% does not influence the mass measurement significantly. With a stronger constraint the statistical uncertainty of the estimated background fraction decreases which is the reasonable behaviour.

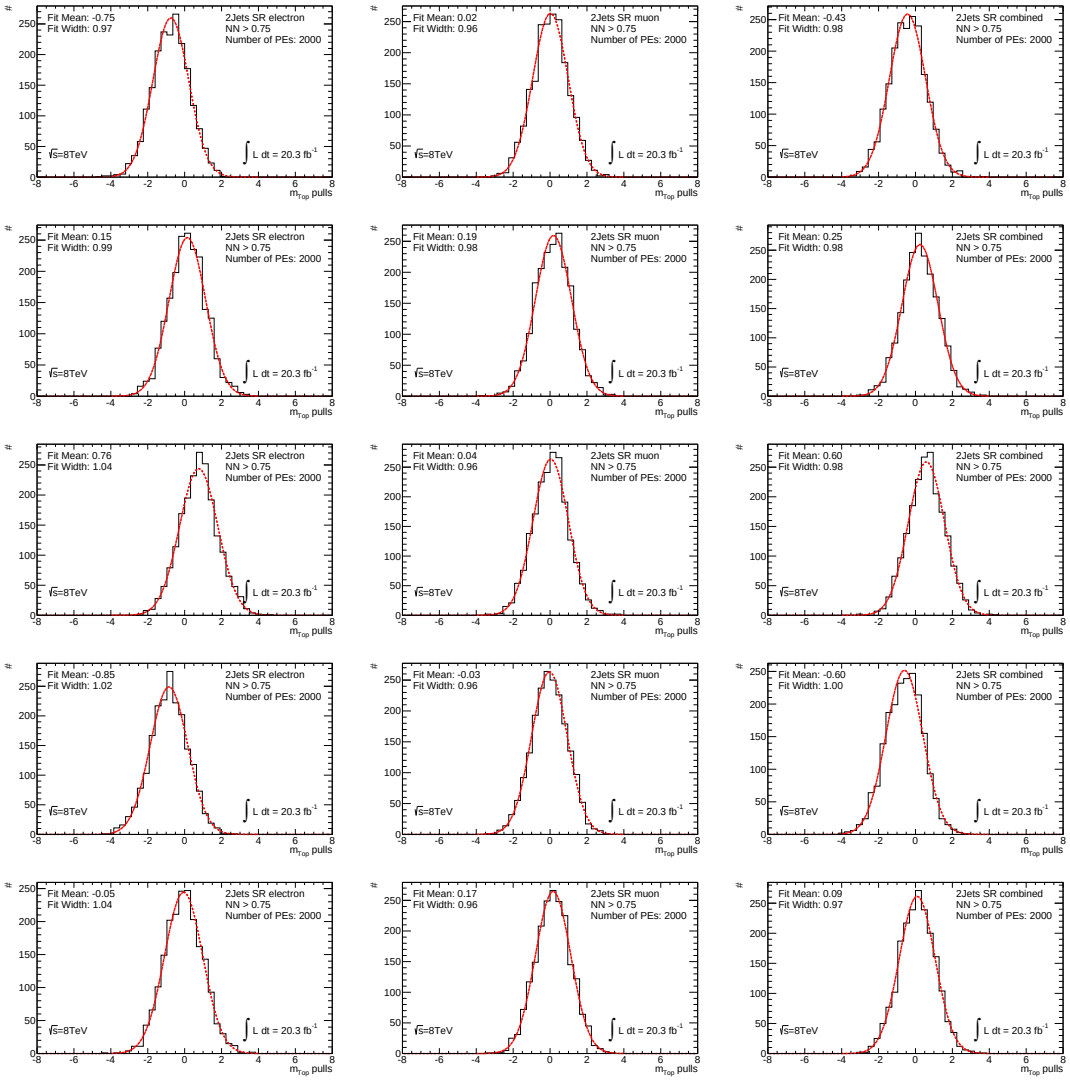


a) electron b) muon c) combined

Figure A.17: Variation of strength of Gaussian constraint on predicted background fraction used in the template fit explained in Section 10.2.3. The x-axis in all plots shows the chosen width of the Gaussian constraint and on the y-axis the fitted top quark mass (top row) and fitted background fraction (bottom row) is shown. No indication for a bias of the top quark mass measurement induced by the constraint is seen.

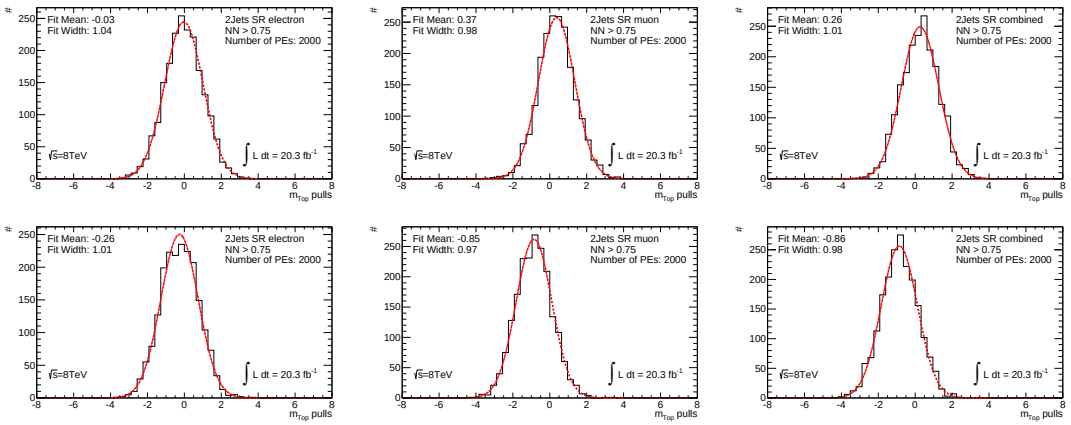
A.8 Validation plots

A.8.1 Pull distributions



a) electron b) muon c) combined

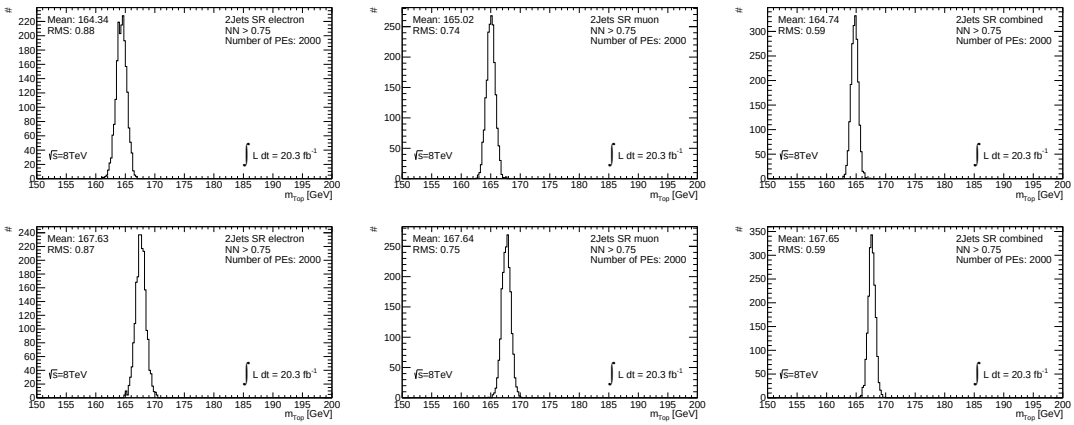
Figure A.18: Pull distributions in the electron (left), muon (middle) and combined channel (right) for the different top quark mass points (top: $m_{\text{top}} = 165.0$ GeV to bottom: $m_{\text{top}} = 175.0$ GeV) based on pseudo-experiments explained in Section 10.3.



a) electron b) muon c) combined

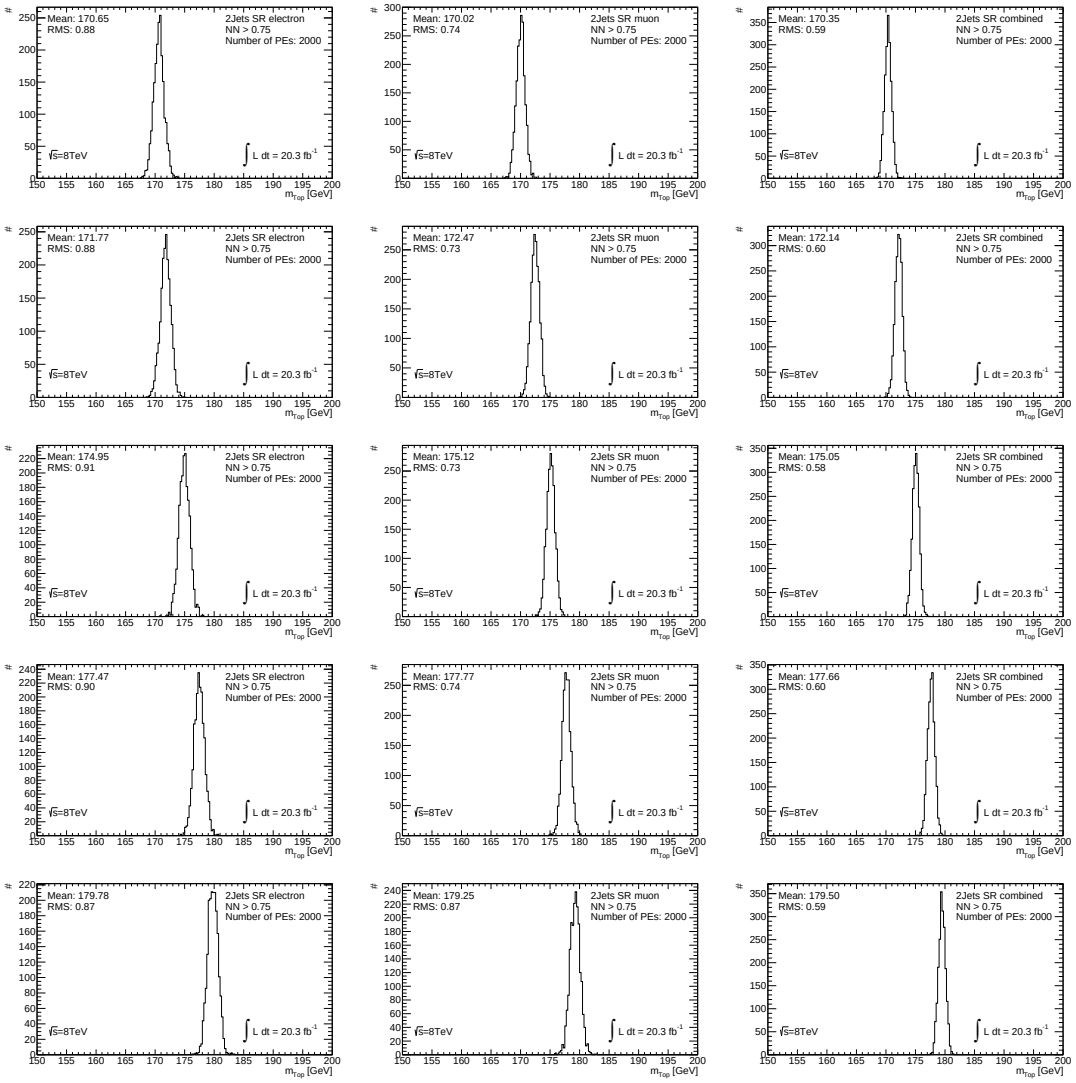
Figure A.19: Pull distributions in the electron (left), muon (middle) and combined channel (right) for the different top quark mass points (top: $m_{\text{top}} = 177.5$ GeV to bottom: $m_{\text{top}} = 180.0$ GeV) based on pseudo-experiments explained in Section 10.3.

A.8.2 Top quark mass distributions



a) electron b) muon c) combined

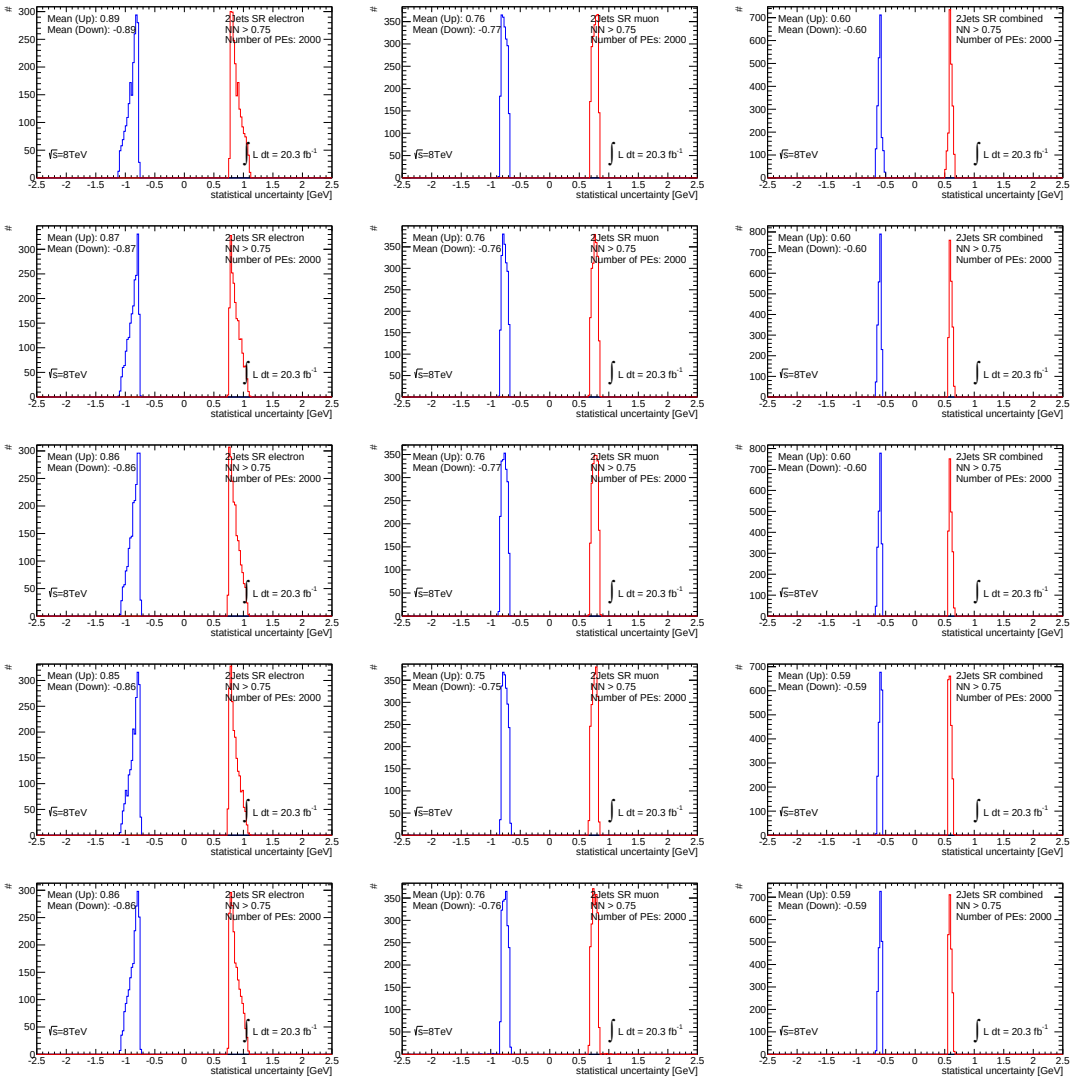
Figure A.20: Mass distributions in the electron (left), muon (middle) and combined channel (right) for the different top quark mass points (top: $m_{\text{top}} = 165.0$ GeV to bottom: $m_{\text{top}} = 167.5$ GeV) based on pseudo-experiments explained in Section 10.3.



a) electron b) muon c) combined

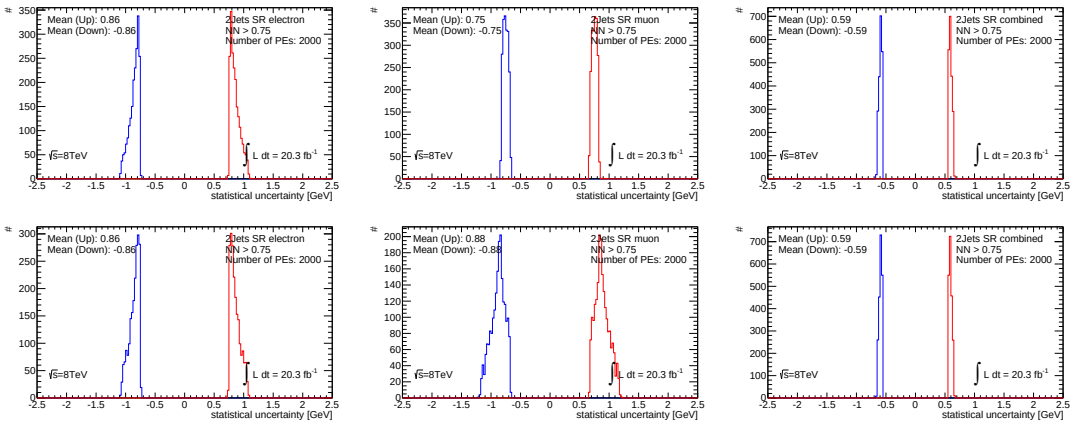
Figure A.21: Mass distributions in the electron (left), muon (middle) and combined channel (right) for the different top quark mass points (top: $m_{\text{top}} = 170.0$ GeV to bottom: $m_{\text{top}} = 180.0$ GeV) based on pseudo-experiments explained in Section 10.3.

A.8.3 Top quark mass uncertainty distributions



a) electron b) muon c) combined

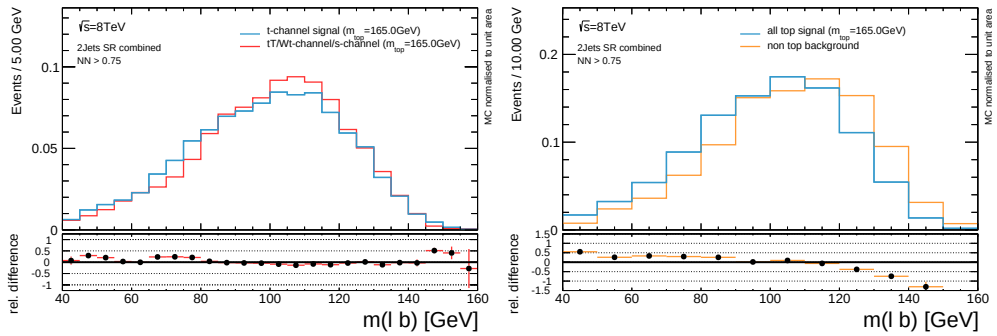
Figure A.22: Asymmetric mass error distributions in the electron (left), muon (middle) and combined channel (right) for the different top quark mass points (top: $m_{\text{top}} = 165.0$ GeV to bottom: $m_{\text{top}} = 175.0$ GeV) based on pseudo-experiments explained in Section 10.3.



a) electron b) muon c) combined

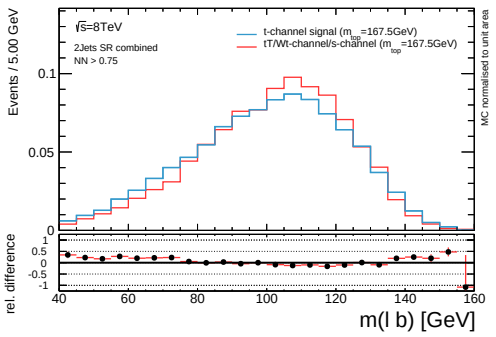
Figure A.23: Asymmetric mass error distributions in the electron (left), muon (middle) and combined channel (right) for the different top quark mass points (top: $m_{\text{top}} = 177.5$ GeV to bottom: $m_{\text{top}} = 180.0$ GeV) based on pseudo-experiments explained in Section 10.3.

A.9 Shape comparisons for different top quark masses

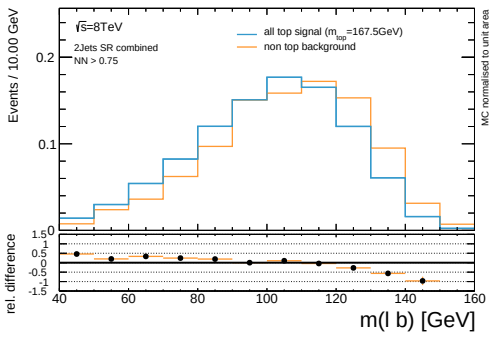


(a) $m_{\text{top}} = 165.0$ GeV (b) $m_{\text{top}} = 165.0$ GeV

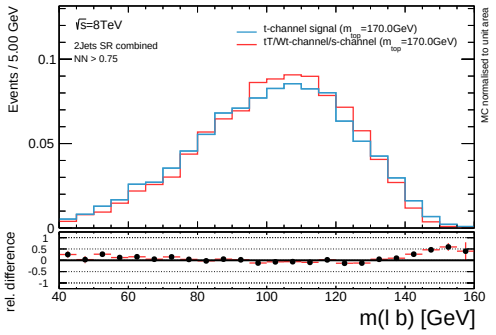
Figure A.24: Shapes of the invariant mass $m(\ell b)$ after all cuts for the t-channel signal and the top quark mass dependent background components (left). Shapes with all signal and background processes included (right). $t\bar{t}$, single-top s-channel and Wt-channel production are treated as signal. The combined electron + muon channel normalised to unit area is shown.



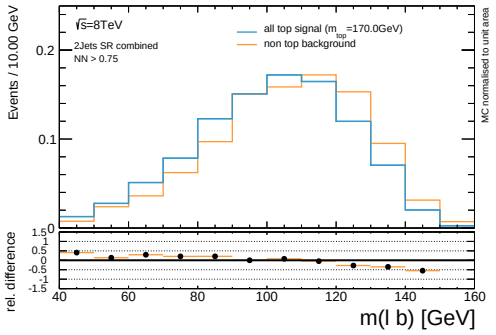
(a) $m_{\text{top}} = 167.5$ GeV



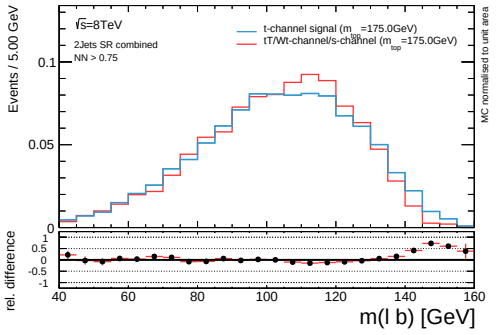
(b) $m_{\text{top}} = 167.5$ GeV



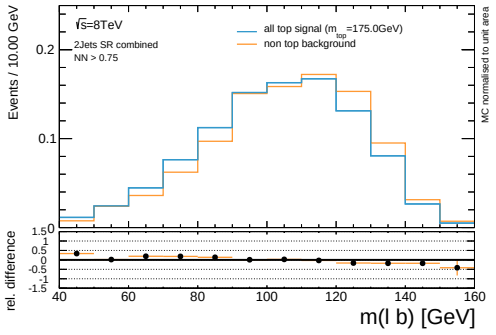
(c) $m_{\text{top}} = 170.0$ GeV



(d) $m_{\text{top}} = 170.0$ GeV

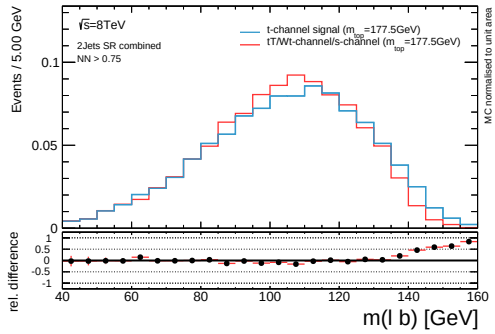


(e) $m_{\text{top}} = 175.0$ GeV

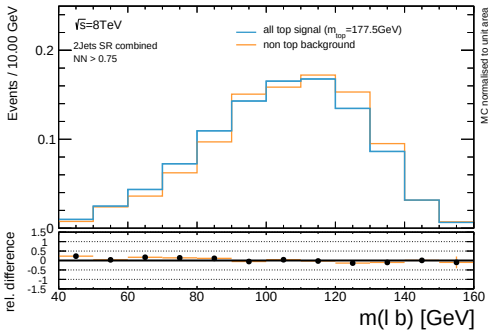


(f) $m_{\text{top}} = 175.0$ GeV

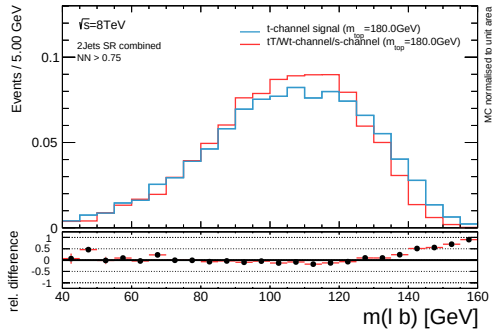
Figure A.25: Shapes of the invariant mass $m(\ell b)$ after all cuts for the t-channel signal and the top quark mass dependent background components (left). Shapes with all signal and background processes included (right). $t\bar{t}$, single-top s-channel and Wt-channel production are treated as signal. The combined electron + muon channel normalised to unit area is shown.



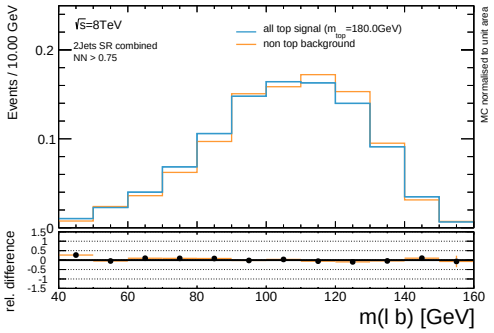
(a) $m_{\text{top}} = 177.5$ GeV



(b) $m_{\text{top}} = 177.5$ GeV



(c) $m_{\text{top}} = 180.0$ GeV



(d) $m_{\text{top}} = 180.0$ GeV

Figure A.26: Shapes of the invariant mass $m(\ell b)$ after all cuts for the t-channel signal and the top quark mass dependent background components (left). Shapes with all signal and background processes included (right). $t\bar{t}$, single-top s-channel and Wt-channel production are treated as signal. The combined electron + muon channel normalised to unit area is shown.

APPENDIX B

Publications

Conference posters

Hendrik Esch on behalf of the ATLAS collaboration,
b-Jet Tagging Performance Calibration in ATLAS
26th International Symposium on LEPTON PHOTON INTERACTIONS at HIGH ENERGIES June 2013, San Francisco (CA, USA)
http://cds.cern.ch/record/1559311/

Hendrik Esch on behalf of the ATLAS collaboration,
Top quark mass in topologies enhanced with single top quarks produced in the t-channel at $\sqrt{s} = 8$ TeV with ATLAS
7th International Workshop on Top-Quark Physics
September/October 2014, Cannes-Mandelieu (FR)
https://cds.cern.ch/record/1953711/

Conference talks

Hendrik Esch, Measurement of the top quark mass in topologies enhanced with single top quarks produced in the t-channel at $\sqrt{s} = 8$ TeV using the ATLAS experiment
8th Annual Helmholtz Alliance Workshop on "Physics at the Terascale"
December 2014, Hamburg (Germany)
https://indico.desy.de/contributionDisplay.py?contribId=50&confId=9821

Hendrik Esch, Kalibrierung der *b*-Jet Tagging Effizienz mit p_T^{rel} am ATLAS Experiment
Particle Physics DPG Spring Meeting 2012 March 2012, Göttingen (Germany)

Hendrik Esch, Systematic uncertainty due to gluon radiation among the determination of the $t\bar{t}$ cross section with the ATLAS experiment
Particle Physics DPG Spring Meeting 2011 April 2011, Dresden (Germany)

Articles

ATLAS Collaboration, Measurement of the b-tag Efficiency in a Sample of Jets Containing Muons with $5\,\text{fb}^{-1}$ of Data from the ATLAS Detector
ATLAS-CONF-2012-043, CERN, Geneva, Mar, 2012.
http://cds.cern.ch/record/1435197

ATLAS Collaboration, Measurement of the Inclusive and Fiducial Cross-Section of Single Top-Quark t-Channel Events in pp Collisions at $\sqrt{s} = 8$ TeV
ATLAS-CONF-2014-007, CERN, Geneva, Mar, 2014
http://cds.cern.ch/record/1668960

ATLAS Collaboration, Measurement of the top quark mass in topologies enhanced with single top-quarks produced in the t-channel in $\sqrt{s} = 8$ TeV ATLAS data
ATLAS-CONF-2014-055, CERN, Geneva, Sep, 2014
https://cds.cern.ch/record/1951323
ATL-COM-PHYS-2013-1665: https://cds.cern.ch/record/1639555 (internal doc.)

ATLAS Collaboration, Performance of b-Jet Identification in the ATLAS Experiment
paper in preparation, to be submitted to JINST
ATL-COM-PHYS-2012-824: https://cds.cern.ch/record/1455862 (paper draft)

Conference proceedings

Hendrik Esch for the ATLAS Collaboration, Measurement of the top quark mass in topologies enhanced with single top quarks produced in the t-channel at $\sqrt{s} = 8$ TeV using the ATLAS experiment
ATL-PHYS-PROC-2014-252 Proceedings of Top 2014, Cannes, 28th Sept - 3rd Oct 2014
https://cds.cern.ch/record/1970430 preprint arXiv:1411.3879 [hep-ex]

Bibliography

[1] CDF Collaboration, F. Abe et al., *Observation of Top Quark Production in $p\overline{p}$ 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. 1, 2.2.1

[2] D0 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. 1, 2.2.1

[3] ATLAS Collaboration, CDF Collaboration, CMS Collaboration, D0 Collaboration, *First combination of Tevatron and LHC measurements of the top-quark mass*, Tech. Rep. ATLAS-CONF-2014-008. CDF-NOTE-11071. CMS-PAS-TOP-13-014. D0-NOTE-6416, CERN, Geneva, Mar, 2014. 1, 3.3, 3.6

[4] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. **D83** (May, 2011) 091503, arXiv:1103.2792 [hep-ph]. 1, 3.1.2, 4.1, 7.1, 7.2, 7.3

[5] 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]. 1, 3.1.2, 7.1, 7.2.2, 7.2, 7.3

[6] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev. **D81** (2010) 054028, arXiv:1001.5034 [hep-ph]. 1, 3.1.2, 7.1, 7.2, 7.3

[7] ATLAS Collaboration, G. Aad et al., *Measurement of the top quark mass with the template method in the $t\bar{t} \to$ lepton + jets channel using ATLAS data*, Eur. Phys. J. **C72** (2012) 2046, arXiv:1203.5755 [hep-ex]. 1

[8] Wikimedia Commons: MissMJ, *Standard Model of Elementary Particles*, Website, 2006. Available online at http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg; visited on May 06th 2014. This file is licensed under the Creative Commons Attribution 3.0 Unported license. 2.1

[9] Particle Data Group Collaboration, J. Beringer et al., *Review of Particle Physics (RPP)*, Phys. Rev. **D86, (2012) and 2013 partial update for the 2014 edition** 010001. http://journals.aps.org/prd/abstract/10.1103/PhysRevD.86.010001. 2.1, 2.1, 2.2, 2.3.2, 2.3.3, 2.4, 2.5, 3.2, 9, 11, 3.2, 3.2.1, 3.1, 3.3.2, 3.4.2, 6.2.1.6

[10] F. Halzen and A. D. Martin, *QUARKS AND LEPTONS: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, Inc., 1984. 1, 6

[11] M. Gell-Mann, *The Eightfold Way: A Theory of strong interaction symmetry*, Tech. Rep. CTSL-20, TID-12608, 1961. 2.2.1

[12] Y. Ne'eman, *Derivation of strong interactions from a gauge invariance*, Nucl. Phys. **26** (1961) 222–229. 2.2.1

[13] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964) 214–215. 2.2.1

[14] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking; Part I*, Tech. Rep. CERN-TH-401, CERN, Geneva, Jan, 1964. 2.2.1

[15] G. Zweig, *An SU(3) model for strong interaction symmetry and its breaking; Part II*, Tech. Rep. CERN-TH-412, CERN, Geneva, Feb, 1964. 2.2.1

[16] E. D. Bloom et al., *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. 2.2.1, 2.6

[17] M. Breidenbach et al., *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. 2.2.1, 2.6

[18] J. Bjorken and S. Glashow, *Elementary Particles and SU(4)*, Phys. Lett. **11** (1964) 255–257. 2.2.1

[19] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2** (Oct, 1970) 1285–1292. http://link.aps.org/doi/10.1103/PhysRevD.2.1285. 2.2.1

[20] J. E. Augustin et al., *Discovery of a Narrow Resonance in e^+e^- Annihilation*, Phys. Rev. Lett. **33** (Dec, 1974) 1406–1408. http://link.aps.org/doi/10.1103/PhysRevLett.33.1406. 2.2.1

[21] J. J. Aubert et al., *Experimental Observation of a Heavy Particle J*, Phys. Rev. Lett. **33** (Dec, 1974) 1404–1406. http://link.aps.org/doi/10.1103/PhysRevLett.33.1404. 2.2.1

[22] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. **49** (1973) 652–657. 2.2.1, 2.3.3

[23] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (Aug, 1977) 252–255. http://link.aps.org/doi/10.1103/PhysRevLett.39.252. 2.2.1

[24] L. M. Lederman, *Logbook: Bottom Quark*, Symmetry Magazine **2 (8)** (2005) . http://www.symmetrymagazine.org/sites/default/files/legacy/pdfs/200510/logbook.pdf. 2.2.1

[25] D0 Collaboration, V. Abazov et al., *Evidence for production of single top quarks and first direct measurement of $|Vtb|$*, Phys. Rev. Lett. **98** (2007) 181802, arXiv:hep-ex/0612052 [hep-ex]. 2.2.1

[26] Super-Kamiokande Collaboration, Y. Fukuda et al., *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. **81** (Aug, 1998) 1562–1567. http://link.aps.org/doi/10.1103/PhysRevLett.81.1562. 2.2.2

[27] I. Falconer, *Corpuscles, Electrons and Cathode Rays: J.J.Thomson and the 'Discovery of the Electron'*, Brit. J. Hist. Sci. **20** (1987) 241–276. 2.2.2

[28] L. M. Brown, *The idea of the neutrino*, Phys. Today **31** (1978) 23. 2.2.2

[29] C. Cowan, F. Reines, F. Harrison, H. Kruse, and A. McGuire, *Detection of the free neutrino: A Confirmation*, Science **124** (1956) 103–104. 2.2.2

[30] C. D. Anderson and S. H. Neddermeyer, *Cloud Chamber Observations of Cosmic Rays at 4300 Meters Elevation and Near Sea-Level*, Phys. Rev. **50** (Aug, 1936) 263–271. http://link.aps.org/doi/10.1103/PhysRev.50.263. 2.2.2

[31] G. Danby et al., *Observation of High-Energy Neutrino Reactions and the Existence of Two Kinds of Neutrinos*, Phys. Rev. Lett. **9** (Jul, 1962) 36–44. http://link.aps.org/doi/10.1103/PhysRevLett.9.36. 2.2.2

[32] M. L. Perl et al., *Evidence for Anomalous Lepton Production in e^+- e^- Annihilation*, Phys. Rev. Lett. **35** (Dec, 1975) 1489–1492. http://link.aps.org/doi/10.1103/PhysRevLett.35.1489. 2.2.2

[33] DONUT Collaboration, K. Kodama et al., *Observation of tau neutrino interactions*, Phys. Lett. **B504** (2001) 218–224, arXiv:hep-ex/0012035 [hep-ex]. 2.2.2

[34] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Non-Abelian Gauge Theories*, Phys. Rev. Lett. **30** (Jun, 1973) 1343–1346. http://link.aps.org/doi/10.1103/PhysRevLett.30.1343. 2.3.1

[35] J. Maxwell, *A Dynamical Theory of the Electromagnetic Field*, Phil. Trans. R. Soc. London **155** (Jan, 1865) no. 1, 459–513. http://rstl.royalsocietypublishing.org/content/155/459. 2.3.2

[36] M. Planck, *On the Law of Distribution of Energy in the Normal Spectrum*, Ann. Phys. **4** (1901) 553. 2.3.2

[37] A. Einstein, *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt*, Ann. Phys. **322** (1905) 132–148. 2.3.2

[38] A. H. Compton, *A Quantum Theory of the Scattering of X-rays by Light Elements*, Phys. Rev. **21** (May, 1923) 483–502. http://link.aps.org/doi/10.1103/PhysRev.21.483. 2.3.2

[39] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Westview Press, 1995. Reading, USA: Addison-Wesley (1995) 842 p. 2.3.2, 10

[40] W. Heisenberg, *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*, Z. Phys. **A43** (1927) 172–198. http://dx.doi.org/10.1007/BF01397280. 2.3.2

[41] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (Feb, 1957) 1413–1415. http://link.aps.org/doi/10.1103/PhysRev.105.1413. 2.3.2

[42] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay, *Evidence for the* 2π *Decay of the* K_2^0 *Meson*, Phys. Rev. Lett. **13** (Jul, 1964) 138–140. http://link.aps.org/doi/10.1103/PhysRevLett.13.138. 2.3.2

[43] L. Di Lella, *Proton-Antiproton Collider Physics: Experimental Aspects*, in *Particle Physics*, vol. 173 of *NATO ASI Series*, pp. 533–596. Springer US, Jan, 1988. http://dx.doi.org/10.1007/978-1-4613-0977-2_14. 2.3.2

[44] B. R. Stella and H.-J. Meyer, *Υ(9.46 GeV) and the gluon discovery (a critical recollection of PLUTO results)*, Eur. Phys. J. **H36** (2011) 203–243, arXiv:1008.1869 [hep-ex]. 2.3.2

[45] P. Söding, *On the discovery of the gluon*, EPJ **H35** (2010) no. 1, 3–28. http://dx.doi.org/10.1140/epjh/e2010-00002-5. 2.3.2

[46] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (Jun, 1963) 531–533. http://link.aps.org/doi/10.1103/PhysRevLett.10.531. 2.3.3

[47] S. Bais, *The Equations: Icons of knowledge*. Harvard University Press, 2005. 2.4

[48] Nobelprize.org. Nobel Media AB 2013. Web. 6 Nov 2013, *The Nobel Prize in Physics 1979*, . http://www.nobelprize.org/nobel_prizes/physics/laureates/1979/. 2.4

[49] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (Aug, 1964) 321–323. http://link.aps.org/doi/10.1103/PhysRevLett.13.321. 2.4

[50] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (Oct, 1964) 508–509. http://link.aps.org/doi/10.1103/PhysRevLett.13.508. 2.4

[51] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (Nov, 1964) 585–587. http://link.aps.org/doi/10.1103/PhysRevLett.13.585. 2.4

[52] P. Schmüser, *Feynman-Graphen und Eichtheorien für Experimentalphysiker*. Springer-Verlag GmbH, 1995. 2.3, 2.5

[53] ATLAS Collaboration, 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*, Phys. Lett. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex]. 2.5

[54] CMS Collaboration, S. Chatrchyan et al., *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex]. 2.5

[55] A. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285, arXiv:0901.0002 [hep-ph]. 2.4, 2.6, 10.4.2

[56] H1 and ZEUS Collaboration, F. Aaron et al., *Combined Measurement and QCD Analysis of the Inclusive $e^{\pm}$ p Scattering Cross Sections at HERA*, JHEP **1001** (2010) 109, arXiv:0911.0884 [hep-ex]. 2.6

[57] D0 Collaboration, V. M. Abazov et al., *Measurement of the Inclusive Jet Cross Section in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. Lett. **101** (Aug, 2008) 062001. http://link.aps.org/doi/10.1103/PhysRevLett.101.062001. 2.6

[58] CDF Collaboration, T. Aaltonen et al., *Measurement of the Inclusive Jet Cross Section at the Fermilab Tevatron $p\bar{p}$ Collider Using a Cone-Based Jet Algorithm*, Phys. Rev. **D78** (2008) 052006, arXiv:0807.2204 [hep-ex]. 2.6

[59] H.-L. Lai et al., *New parton distributions for collider physics*, Phys. Rev. **D82** (2010) 074024, arXiv:1007.2241 [hep-ph]. 2.6, 6.2.1.4, 7.2.2, 7.4.1, 10.4.2

[60] F. Demartin et al., *The impact of PDF and alphas uncertainties on Higgs Production in gluon fusion at hadron colliders*, Phys. Rev. **D82** (2010) 014002, arXiv:1004.0962 [hep-ph]. 2.6

[61] M. Cacciari et al., *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, Phys. Lett. B **710** (2012) 612–622, arXiv:1111.5869 [hep-ph]. 3.1.1, 7.1, 7.2, 7.4

[62] 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]. 3.1.1, 7.1, 7.2, 7.4

[63] 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]. 3.1.1, 7.1, 7.2, 7.4

[64] 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]. 3.1.1, 7.1, 7.2, 7.4

[65] M. Czakon, P. Fiedler, and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$*, Phys. Rev. Lett. **110** (2013) no. 25, 252004, arXiv:1303.6254 [hep-ph]. 3.1.1, 7.1, 7.2, 7.4

[66] 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]. 3.1.1, 7.1, 7.2, 7.4

[67] A. Heinson and T. R. Junk, *Observation of Single Top Quark Production*, Ann. Rev. Nucl. Part. Sci. **61** (2011) 171–196, arXiv:1101.1275 [hep-ex]. 3.1.2

[68] ATLAS Collaboration, G. Aad et al., *Evidence for the associated production of a W boson and a top quark in ATLAS at* $\sqrt{s} = 7$ *TeV*, Phys. Lett. **B716** (2012) 142–159, arXiv:1205.5764 [hep-ex]. 3.1.2

[69] CMS Collaboration, S. Chatrchyan et al., *Evidence for associated production of a single top quark and W boson in pp collisions at* $\sqrt{s} = 7$ *TeV*, Phys. Rev. Lett. **110** (2013) 022003, arXiv:1209.3489 [hep-ex]. 3.1.2

[70] D0 Collaboration, V. M. Abazov et al., *Evidence for s-channel single top quark production in* $p\bar{p}$ *collisions at* $\sqrt{s} = 1.96$ *TeV*, Phys. Lett. **B726** (2013) 656–664, arXiv:1307.0731 [hep-ex]. 3.1.2

[71] ATLAS Collaboration, *Measurement of the Inclusive and Fiducial Cross-Section of Single Top-Quark t-Channel Events in pp Collisions at* $\sqrt{s} = 8$ *TeV*, Tech. Rep. ATLAS-CONF-2014-007, CERN, Geneva, Mar, 2014. 3.1.3, 3.1.3, 3.4, 8.1.2, 8.7, 9.6.2, 10.4.3

[72] C. Quigg, *Top-ology*, Phys. Today **50N5** (1997 (March 2014 update to the top-mass time series originally published as Figure 1)) 20–26, arXiv:hep-ph/9704332 [hep-ph]. http://lutece.fnal.gov/TTS/. 3.5

[73] M. C. Smith and S. S. Willenbrock, *Top quark pole mass*, Phys. Rev. Lett. **79** (1997) 3825–3828, arXiv:hep-ph/9612329 [hep-ph]. 3.3.1

[74] A. H. Hoang, A. Jain, I. Scimemi, and I. W. Stewart, *Infrared Renormalization Group Flow for Heavy Quark Masses*, Phys. Rev. Lett. **101** (2008) 151602, arXiv:0803.4214 [hep-ph]. 3.3.1

[75] N. Gray, D. J. Broadhurst, W. Grafe, and K. Schilcher, *Three Loop Relation of Quark (Modified) Ms and Pole Masses*, Z. Phys. **C48** (1990) 673–680. http://link.springer.com/article/10.1007%2FBF01614703. 3.3.1

[76] K. Melnikov and T. v. Ritbergen, *The Three loop relation between the MS-bar and the pole quark masses*, Phys. Lett. **B482** (2000) 99–108, arXiv:hep-ph/9912391 [hep-ph]. 3.3.1

[77] S. Bethke, *The 2009 World Average of alpha(s)*, Eur. Phys. J. **C64** (2009) 689–703, arXiv:0908.1135 [hep-ph]. 3.3.1

[78] D0 Collaboration, V. M. Abazov et al., *Determination of the pole and MSbar masses of the top quark from the $t\bar{t}$ cross section*, Phys. Lett. **B703** (2011) 422–427, arXiv:1104.2887 [hep-ex]. 3.3.1, 3.3.2, 3.9, 3.3.2

[79] A. H. Hoang and I. W. Stewart, *Top Mass Measurements from Jets and the Tevatron Top-Quark Mass*, Nucl. Phys. Proc. Suppl. **185** (2008) 220–226, arXiv:0808.0222 [hep-ph]. 3.3.1

[80] A. Buckley et al., *General-purpose event generators for LHC physics*, Phys. Rept. **504** (2011) 145–233, arXiv:1101.2599 [hep-ph]. 3.3.1

[81] M. E. Peskin and T. Takeuchi, *Estimation of oblique electroweak corrections*, Phys. Rev. **D46** (1992) 381–409. 3.3.2

[82] 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]. 3.3.2, 3.7, 3.3.2

[83] 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]. 3.3.2

[84] D0 Collaboration, V. Abazov et al., *A precision measurement of the mass of the top quark*, Nature **429** (2004) 638–642, arXiv:hep-ex/0406031 [hep-ex]. 3.3.2

[85] ATLAS collaboration, *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. 3.3.2, 3.8, 11

[86] CMS collaboration, *Measurement of the top quark mass using the B-hadron lifetime technique*, Tech. Rep. CMS-PAS-TOP-12-030, CERN, Geneva, 2013. 3.3.2

[87] CDF Collaboration, T. Aaltonen et al., *Measurement of the Top Quark Mass in the Lepton+Jets Channel Using the Lepton Transverse Momentum*, Phys. Lett. **B698** (2011) 371–379, arXiv:1101.4926 [hep-ex]. 3.3.2

[88] C. A. Jung, *Measurement of the mass of the top quark using the transverse decay length and lepton transverse momentum techniques.* PhD thesis, Technische Universität Dortmund, February, 2014. http://hdl.handle.net/2003/33096. 3.3.2

[89] ATLAS Collaboration, G. Aad et al., *Measurement of the top quark charge in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **1311** (2013) 031, arXiv:1307.4568 [hep-ex]. 3.4.1

[90] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) no. 08, S08001. http://stacks.iop.org/1748-0221/3/i=08/a=S08001. 4.1

[91] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) no. 08, S08003. http://stacks.iop.org/1748-0221/3/i=08/a=S08003. 4.1, 4.4, 24, 4.2.1, 4.1, 4.2.2, 4.5, 4.2.2, 4.2.3, 4.6, 4.2.3, 4.2.4, 27, 4.2, 4.7, 4.2.4

[92] CMS Collaboration, S. Chatrchyan et al., *The CMS experiment at the CERN LHC*, JINST **3** (2008) no. 08, S08004. http://stacks.iop.org/1748-0221/3/i=08/a=S08004. 4.1

[93] LHCb Collaboration, A. Augusto Alves Jr et al., *The LHCb Detector at the LHC*, JINST **3** (2008) no. 08, S08005. http://stacks.iop.org/1748-0221/3/i=08/a=S08005. 4.1

[94] ALICE Collaboration, K. Aamodt et al., *The ALICE experiment at the CERN LHC*, JINST **3** (2008) no. 08, S08002. http://stacks.iop.org/1748-0221/3/i=08/a=S08002. 4.1

[95] LHCf Collaboration, O. Adriani et al., *The LHCf detector at the CERN Large Hadron Collider*, JINST **3** (2008) S08006. http://stacks.iop.org/1748-0221/3/i=08/a=S08006. 4.1

[96] TOTEM Collaboration, G. Anelli et al., *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007. http://stacks.iop.org/1748-0221/3/i=08/a=S08007. 4.1

[97] J. Pinfold, *CERN Courier: MoEDAL becomes the LHC's magnificent seventh*, Website, 05, 2010. Available online at http://cerncourier.com/cws/article/cern/42329; visited on April 28th 2014. 4.1

[98] AC Team, *The four main LHC experiments*, Available online at `http://cdsweb.cern.ch/record/40525`; visited on March 10th 2014., Jun, 1999. 4.1

[99] 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]`. 21

[100] *Luminosity Public Results*, Website. Available online at `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults`; visited on April 28th 2014. 4.2

[101] ATLAS Collaboration, G. Aad et al., *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823–874, `arXiv:1005.4568 [physics.ins-det]`. 4.3, 7.1

[102] T. Kittelmann et al., *The Virtual Point 1 event display for the ATLAS experiment*, J. Phys.: Conf. Ser. **219** (2010) no. 3, 032012. `http://iopscience.iop.org/1742-6596/219/3/032012/`. 23

[103] *Atlantis, Event display for ATLAS*, Website. Available online at `http://atlantis.web.cern.ch/atlantis/`; visited on July 09th 2014. 23

[104] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1999. 4.7

[105] L. M. Lederman, *The Tevatron*, Sci. Am. **264** (Mar, 1991) no. 3, 48–55. `http://www.nature.com/scientificamerican/journal/v264/n3/pdf/scientificamerican0391-48.pdf`. 4.3

[106] ATLAS Collaboration, *Performance of Impact Parameter-Based b-tagging Algorithms with the ATLAS Detector using Proton-Proton Collisions at $\sqrt{s} = 7$* TeV, Tech. Rep. ATLAS-CONF-2010-091, CERN, Geneva, Oct, 2010. 5.1, 5.1, 6.1.2

[107] ATLAS Collaboration, *Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data*, Tech. Rep. ATLAS-CONF-2010-042, CERN, Geneva, Jul, 2010. 5.1, 6.1

[108] ATLAS Collaboration, G. Aad et al., *Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s}$ = 7 TeV*, Eur. Phys. J. **C73** (2013) 2304, `arXiv:1112.6426 [hep-ex]`. 5.2, 5.2.1

[109] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **0804** (2008) 063, `arXiv:0802.1189 [hep-ph]`. http://fastjet.fr/. 5.2

[110] ATLAS Collaboration, G. Aad et al., *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, SLAC-R-980, CERN-OPEN-2008-020, arXiv:0901.0512 [hep-ex]. 30, 6.1.2, 6.1.3

[111] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008. 5.2.1

[112] G. Pospelov and the Atlas Hadronic Calibration Group, *The overview of the ATLAS local hadronic calibration*, J. Phys.: Conf. Ser. **160** (2009) no. 1, 012079. http://stacks.iop.org/1742-6596/160/i=1/a=012079. 31

[113] ATLAS Collaboration, G. Aad et al., *Jet energy measurement and its systematic uncertainty in proton–proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, arXiv:1406.0076 [hep-ex]. submitted to EPJC. 5.2.1, 6.2.1.1, 6.5, 6.2.1.6, 10.4.1

[114] ATLAS Collaboration, *Data-Quality Requirements and Event Cleaning for Jets and Missing Transverse Energy Reconstruction with the ATLAS Detector in Proton-Proton Collisions at a Center-of-Mass Energy of sqrts = 7 TeV*, Tech. Rep. ATLAS-CONF-2010-038, CERN, Geneva, Jul, 2010. 5.2.2

[115] ATLAS Collaboration, *Performance of the ATLAS muon trigger in 2011*, Tech. Rep. ATLAS-CONF-2012-099, CERN, Geneva, Jul, 2012. 5.3.1

[116] ATLAS Collaboration, *Preliminary results on the muon reconstruction efficiency, momentum resolution, and momentum scale in ATLAS 2012 pp collision data*, Tech. Rep. ATLAS-CONF-2013-088, CERN, Geneva, Aug, 2013. 5.3.2, 5.3.4

[117] K. Rehermann and B. Tweedie, *Efficient identification of boosted semileptonic top quarks at the LHC*, JHEP **2011** (2011) 1–27, arXiv:1007.2221 [hep-ph]. 5.3.3

[118] ATLAS Collaboration, G. Aad et al., *Muon reconstruction efficiency and momentum resolution of the ATLAS experiment in proton-proton collisions at $\sqrt{s}$=7 TeV in 2010*, arXiv:1404.4562 [hep-ex]. submitted to EPJC. 5.3.4

[119] A. Collaboration, *Performance of the ATLAS Electron and Photon Trigger in p-p Collisions at $\sqrt{s} = 7$ TeV in 2011*, Tech. Rep. ATLAS-CONF-2012-048, CERN, Geneva, May, 2012. 5.4.1

[120] ATLAS Collaboration, *Expected electron performance in the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2011-006, CERN, Geneva, Apr, 2011. 5.4.2, 5.4.4

[121] 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]. 5.4.2, 5.4.4

[122] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, Tech. Rep. ATLAS-CONF-2012-047, CERN, Geneva, May, 2012. 5.4.2, 8.2.2

[123] M. Hance, D. Olivito, and H. Williams, *Performance Studies for e/gamma Calorimeter Isolation*, Tech. Rep. ATL-COM-PHYS-2011-1186, CERN, Geneva, Sep, 2011. 5.4.3

[124] S. Laplace and J. de Vivie, *Calorimeter isolation and pile-up*, Tech. Rep. ATL-COM-PHYS-2012-467, CERN, Geneva, May, 2012. 5.4.3

[125] ATLAS Collaboration, ATLAS collaboration, *Electron efficiency measurements with the ATLAS detector using the 2012 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2014-032, CERN, Geneva, Jun, 2014. 5.4.4

[126] ATLAS Collaboration, G. Aad et al., *Performance of missing transverse momentum reconstruction in proton-proton collisions at* $\sqrt{s} = 7$ *TeV with ATLAS*, Eur. Phys. J. **C72** (Jan, 2012) no. 1, 1844. http://dx.doi.org/10.1140/epjc/s10052-011-1844-6. 5.5

[127] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS studied in Proton-Proton Collisions recorded in 2012 at 8 TeV*, Tech. Rep. ATLAS-CONF-2013-082, CERN, Geneva, Aug, 2013. 5.5

[128] ATLAS Collaboration, *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. 6.2, 6.1.4, 6.2.1.7, 8.1.2

[129] G. Piacquadio and C. Weiser, *A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS*, J. Phys.: Conf. Ser. **119** (2008) no. 3, 032032. 6.1.3

[130] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987) 444–450. 6.1.3

[131] A. Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, et al., *TMVA - Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007) 040, arXiv:physics/0703039 [physics.data-an]. 6.1.4

[132] ATLAS Collaboration, *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. 6.4, 6.2.2

[133] B. Lenzi, R. Nicolaidou, and S. Hassani, *TrackInCaloTools: A package for measuring muon energy loss and calorimetric isolation in ATLAS*, J. Phys.: Conf. Ser. **219** (2010) no. 3, 032049. 6.2.1.1, 8.2.3

[134] F. Hirsch, *B-Tagging calibration at the ATLAS experiment.* PhD thesis, Technische Universität Dortmund, August, 2011. `http://hdl.handle.net/2003/28957`. 6.2.1.2

[135] ATLAS Collaboration, *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. 6.6(b)

[136] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867, `arXiv:0710.3820`. 6.2.1.4, 7.8

[137] ATLAS collaboration, *Further ATLAS tunes of PYTHIA6 and Pythia 8*, Tech. Rep. ATL-PHYS-PUB-2011-014, CERN, Geneva, Nov, 2011. 6.2.1.4

[138] D. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. **A462** (2001) 152–155. `http://www.sciencedirect.com/science/article/pii/S0168900201000894`. 6.2.1.4

[139] R. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) no. 2, 219 – 228. `http://www.sciencedirect.com/science/article/pii/001046559390005W`. 6.2.1.5, 8.2.4, 10.2.3

[140] ATLAS Collaboration, *Measurement of the b-jet production cross section using muons in jets with ATLAS in pp Collisions at $\sqrt{s} = 7$ TeV*, Tech. Rep. ATLAS-CONF-2011-057, CERN, Geneva, Apr, 2011. 6.2.1.5

[141] CDF Collaboration, T. Aaltonen et al., *Measurement of Ratios of Fragmentation Fractions for Bottom Hadrons in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$-TeV*, Phys. Rev. **D77** (2008) 072003, `arXiv:0801.4375 [hep-ex]`. 6.2.1.6

[142] DELPHI Collaboration, J. Abdallah et al., *Determination of heavy quark non-perturbative parameters from spectral moments in semileptonic B decays*, Eur. Phys. J. **C45** (2006) 35–59, `arXiv:hep-ex/0510024 [hep-ex]`. 6.2.1.6

[143] 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]. 6.2.1.6, 10.4.1

[144] *Inner Detector Tracking Performance Guidelines*, Website. Available online at https://twiki.cern.ch/twiki/bin/view/AtlasProtected/InDetTrackingPerformanceGuidelines; visited on June 6th 2014. 6.2.1.6

[145] ATLAS Collaboration, *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. 6.2.1.6, 6.2.3, 6.11, 6.12

[146] ATLAS Collaboration, *b-Jet Tagging Efficiency Calibration using the System8 Method*, Tech. Rep. ATLAS-CONF-2011-143, CERN, Geneva, Oct, 2011. 6.10, 6.2.2

[147] ATLAS Collaboration, *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2014-004, CERN, Geneva, Feb, 2014. 6.2.3, 6.2.3

[148] ATLAS Collaboration, *b-jet tagging calibration on c-jets containing D^{*+} mesons*, Tech. Rep. ATLAS-CONF-2012-039, CERN, Geneva, Mar, 2012. 6.13, 6.2.4

[149] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, Tech. Rep. ATLAS-CONF-2014-046, CERN, Geneva, Jul, 2014. 6.2.4, 6.2.5

[150] ATLAS Collaboration, *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. 6.2.5, 6.14

[151] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **0902** (2009) 007, arXiv:0811.4622 [hep-ph]. 7.1, 7.5

[152] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A Simulation Toolkit*, Nucl. Instrum. Meth. **A 506** (2003) 250–303. http://www.sciencedirect.com/science/article/pii/S0168900203013688. 7.1

[153] 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 ARI-*

ADNE 4.1, Comput. Phys. Commun. **184** (2013) no. 3, 919–985. http://www.sciencedirect.com/science/article/pii/S001046551200375X. 7.2.1

[154] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0207** (2002) 012, arXiv:hep-ph/0201195 [hep-ph]. 7.2.1

[155] M. Aivazis et al., *Leptoproduction of heavy quarks. 2. A Unified QCD formulation of charged and neutral current processes from fixed target to collider energies*, Phys. Rev. **D50** (1994) 3102–3118, arXiv:hep-ph/9312319 [hep-ph]. 7.2.1

[156] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **0605** (2006) 026, arXiv:hep-ph/0603175 [hep-ph]. 7.2.1, 50

[157] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. **D82** (2010) 074018, arXiv:1005.3457. 7.2.1, 7.4.1

[158] 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]. 7.2.2

[159] R. Frederix, E. Re, and P. Torrielli, *Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*, JHEP **09** (Jul, 2012) 130. 18 p, arXiv:1207.5391 [hep-ph]. 7.4.1

[160] G. Corcella et al., *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **0101** (2001) 010, arXiv:hep-ph/0011363 [hep-ph]. 7.4.1

[161] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. **C72** (1996) 637–646, arXiv:hep-ph/9601371 [hep-ph]. 7.4.1

[162] ATLAS Collaboration, *New ATLAS event generator tunes to 2010 data*, Tech. Rep. ATL-PHYS-PUB-2011-008, CERN, Geneva, Apr, 2011. 7.4.1

[163] A. Sherstnev and R. S. Thorne, *Different PDF approximations useful for LO Monte Carlo generators*, arXiv:0807.2132 [hep-ph]. 7.4.1

[164] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **1303** (2013) 015, arXiv:1212.3460 [hep-ph]. 49

[165] S. Frixione, E. Laenen, P. Motylinski, and B. R. Webber, *Single-Top Production in MC@NLO*, JHEP **03** (2006) 092, arXiv:hep-ph/0512250 [hep-ph]. 7.4.2

[166] S. Frixione et al., *Single-top hadroproduction in association with a W boson*, JHEP **0807** (2008) 029, arXiv:0805.3067 [hep-ph]. 7.4.2

[167] S. Frixione, P. Nason, et al., *Positive weight next-to-leading-order Monte Carlo*, JHEP 11(2007)-126 and JHEP 09(2007)111, arXiv:07092092 [hep-ph] and arXiv:07073088 [hep-ph] (2007) . 7.4.3

[168] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **0411** (2004) 040, arXiv:0409146 [hep-ph]. 7.4.3

[169] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **0206** (2002) 029, arXiv:hep-ph/0204244 [hep-ph]. 7.4.3

[170] ATLAS Collaboration, *Monte Carlo generator comparisons to ATLAS measurements constraining QCD radiation in top anti-top final states*, Tech. Rep. ATL-PHYS-PUB-2013-005, CERN, Geneva, Mar, 2013. 7.4.3

[171] 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]. 7.4.3

[172] S. Hoeche et al., *QCD matrix elements and truncated showers*, JHEP **0905** (2009) 053, arXiv:0903.1219 [hep-ph]. 7.5

[173] *Top group's MC12 Samples For 2012 Data Analyses*, Website. Available online at https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopMC12; visited on June 20th 2014. 7.5

[174] J. Butterworth et al., *Single Boson and Diboson Production Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV*, Tech. Rep. ATL-COM-PHYS-2010-695, CERN, Geneva, Aug, 2010. 7.7

[175] ATLAS Collaboration, *Measurement of the t-channel single top-quark and top-antiquark production cross-sections and their ratio in pp collisions at $\sqrt{s} = 7\ TeV$*, Tech. Rep. ATLAS-CONF-2012-056, CERN, Geneva, Jun, 2012. 8.2.2

[176] B. Abi et al., *Mis-identified lepton backgrounds to top quark pair production*, Tech. Rep. ATL-PHYS-INT-2010-139, CERN, Geneva, Dec, 2010. 8.2.2

[177] M. Feindt and U. Kerzel, *The NeuroBayes Neural Network Package*, Nucl. Instrum. Meth. **A559** (2006) 190–194. 9, 9.3, 9.3.2

[178] M. Feindt, *A Neural Bayesian Estimator for Conditional Probability Densities*, 2004 , arXiv:physics/0402093 [physics.data-an]. 9, 9.3.2

[179] T. Chwalek, T. Mueller, and W. Wagner, *Measurement of the W boson helicity in top-antitop quark events with the CDF II experiment*, Master's thesis, Karlsruhe U., Karlsruhe, 2006. http://inspirehep.net/record/733740. 9.1

[180] P. Sturm, T. Mueller, and W. Wagner, *Studies for the Measurement of Single-Top-Quark-Events with the CMS-Experiment*, Master's thesis, Karlsruhe U., Karlsruhe, 2008. http://cds.cern.ch/record/1311220?ln=de. Presented on 17 Mar 2008. 9.1

[181] M. Merola and the CMS Collaboration, *Measurement of the single top quark t -channel cross section with a template fit analysis*, J. Phys.: Conf. Ser. **452** (2013) no. 1, 012018. http://stacks.iop.org/1742-6596/452/i=1/a=012018. 9.2.3

[182] Y. LeCun, L. Bottou, G. Orr, and K.-R. Müller, *Efficient BackProp*, in *Neural Networks: Tricks of the Trade*, vol. 1524 of *Lecture Notes in Computer Science*, pp. 9–50. Springer Berlin Heidelberg, 1998. http://dx.doi.org/10.1007/3-540-49430-8_2. 59

[183] J. Cohen, P. Cohen, S. G. West, and L. S. Aiken, *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, 3rd Edition.* Hillsdale, NJ: Lawrence Erlbaum Associates., 2003. http://www.worldcat.org/isbn/0805822232. 9.3.1

[184] R. J. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences.* John Wiley & Sons, Inc., 1989. 9.3.1, 9.3.2

[185] G. Cybenko, *Approximation by superpositions of a sigmoidal function*, Math. Control Signal **2** (1989) no. 4, 303–314. http://dx.doi.org/10.1007/BF02551274. 62

[186] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning Representations by Back-propagating Errors*, in *Neurocomputing: Foundations of Research*, J. A. Anderson and E. Rosenfeld, eds., pp. 696–699. MIT Press, Cambridge, MA, USA, 1988. http://dl.acm.org/citation.cfm?id=65669.104451. 9.5

[187] D. J. C. MacKay, *A Practical Bayesian Framework for Backpropagation Networks*, Neural Comput. **4** (May, 1992) no. 3, 448–472. http://dx.doi.org/10.1162/neco.1992.4.3.448. 9.5.2

[188] R. Brun and F. Rademakers, *ROOT - An object oriented data analysis framework*, Nucl. Inst. & Meth. in Phys. Res. **A389** (1997) no. 12, 81 – 86. http://www.sciencedirect.com/science/article/pii/S016890029700048X. See also http://root.cern.ch/. 66

[189] R. Barlow, *Application of the Bootstrap resampling technique to Particle Physics experiments*, Tech. Rep. MAN/HEP/99/4, 1999. http://www.hep.man.ac.uk/preprints/1999.html. 10.3.1

[190] R. Barlow, *Systematic errors: Facts and fictions*, Tech. Rep. MAN-HEP-02-01, 2002. arXiv:hep-ex/0207026 [hep-ex]. 10.4

[191] ATLAS Collaboration, ATLAS collaboration, *Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector Using Proton-Proton Collisions at a Center of Mass Energy* $\sqrt{s} = 7$ TeV, Tech. Rep. ATLAS-CONF-2010-054, CERN, Geneva, Jul, 2010. 10.4.1

[192] R. D. Ball et al., *Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology*, Nucl. Phys. **B849** (2011) 296–363, arXiv:1101.1300 [hep-ph]. 10.4.2

[193] J. M. Campbell, J. Huston, and W. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, Rept. Prog. Phys. **70** (2007) 89, arXiv:hep-ph/0611148 [hep-ph]. 10.4.2

[194] PDF4LHC, *Recommendation for LHC cross section calculations*, Website. Available online at http://www.hep.ucl.ac.uk/pdf4lhc/; visited on May 09th 2014. 10.4.2

[195] B. Acharya et al., *Object selection and calibration, background estimations and MC samples for the Winter 2013 Top Quark analyses with 2012 data*, Tech. Rep. ATL-COM-PHYS-2013-088, CERN, Geneva, Jan, 2013. 72

[196] ATLAS Collaboration, G. Aad et al., *Measurement of the production cross section for W-bosons in association with jets in pp collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*, Phys. Lett. B **698** (2011) 325–345, arXiv:1012.5382 [hep-ex]. 10.4.3

[197] ATLAS Collaboration, B. Acharya et al., *Estimation of the W+Jets Background for Top Quark Re-Discovery in the Single Lepton+Jets Channel.*, Tech. Rep. ATL-COM-PHYS-2010-834, CERN, Geneva, Oct, 2010. 10.4.3

[198] ATLAS Collaboration, ATLAS collaboration, *Measurement of the Top Quark Mass in Dileptonic Top Quark Pair Decays with* $\sqrt{s} = 7$ *TeV ATLAS Data*, Tech. Rep. ATLAS-CONF-2013-077, CERN, Geneva, Jul, 2013. 11

[199] ATLAS Collaboration, ATLAS collaboration, *Determination of the Top Quark Mass with a Template Method in the All Hadronic Decay Channel using* $2.04\,fb^{-1}$ *of ATLAS DATA*, Tech. Rep. ATLAS-CONF-2012-030, CERN, Geneva, Mar, 2012. 11

Acknowledgement

This thesis would not have been possible without the support and help of many people that deserve the concluding words of this document.

First of all I want to thank my thesis supervisor PD Dr. Reiner Klingenberg for his support in so many fields. This includes the support of our IT infrastructure (also on weekends and during holidays!) which is indispensable when running any kind of analysis in particle physics. I especially admire his ability to ask questions that first of all sound simple but then turn out to really hit the core of a problem. Many times the resulting discussion lead to a significant improvement of a topic.

I want to thank Prof. Dr. Wolfgang Wagner for being the second examiner of my thesis and especially for enabling the close collaboration between our group in Dortmund and the Wuppertal group. In this context I want to thank the people from Wuppertal, namely Phillipp Tepel, Dominic Hirschbühl and Kathrin Becker for many productive discussions and providing me with their knowledge about the single top t-channel and Neurobayes. Also I want to thank Prof. Dr. Shaukat Khan, Dr. Christian Sternemann and Prof. Dr. Ian C. Brock for being part of the thesis defence, reading and assessing this admittedly lengthy thesis.

A very special thanks goes to Christian Jung for answering all my questions related to physics, programming, technical details, statistics, sports, geography, history et cetera sometimes before I even knew that I had these questions. Thanks for teaming up from the very beginning of studying physics in Dortmund until today.

A big thanks deserve the members of the chair "Experimentelle Physik 4" lead by Prof. Dr. C. Gößling for creating the possibility to work in this great relaxed atmosphere. In particular I want to thank all members of the ATLAS analysis group in Dortmund I was able to team up with over the last years. This involves Ingo Burmeister I shared the office with and never (almost never ;)) regretted it and Michael Homann who will continue working on the t-channel analysis to raise the methods precision into unimaginable heights. I also want to thank all diploma, bachelor and master students Sonja Bartkowski, Lennart Behrendt, Gregor Geßner, Timea Krones, Tobias Kupfer and Is-

abel Nitsche for their enthusiasm when starting to work in the field of an analysis with ATLAS. Thanks also to Mona Abt for bringing the spirit of the ATLAS upgrade into our office and providing me with helpful feedback regarding the thesis.

Apart from Reiner Klingenberg I also want to thank Markus Alex and Oliver Schulz for their all time quick support fixing issues related to the IT infrastructure and our local cluster (again also on weekends and during holidays!). The analysis results I was able to obtain would not at all be conceivable with a lot of CPU power and disk storage that was provided by them.

I learned a lot of how to run an analysis in ATLAS during my participation in the flavour tagging combined performance group. Thus, I want to thank Florian Hirsch who started the first efforts of the p_T^{rel}-method in ATLAS. Furthermore I want to thank Georges Aad, Dominik Duda, Frank Filthaut, James Mueller, Fabrizio Parodi, Kevin Sapp, Sara Strandberg, Gordon Watts, Christian Weiser and all the others who provided feedback to improve the b-tagging calibration results and finally publish the results to the particle physics community.

The same that applies to the members of the flavour tagging group also applies to many people working in the field of top quark physics and mass measurements within the ATLAS collaboration. Without the comments and suggestions from Stefanie Adomeit, Barbara Alvarez Gonzalez, Duc Bao Ta, Oleg Brandt, Tancredi Carli, Giorgio Cortiana, Carlos Escobar, Jose Enrique García Navarro, Alison Lister, Kevin Kröninger and again many more people I would not have been able to perform the analysis as it is described. In addition, thanks to the members of the editorial board Dominique Pallin, Lily Asquith and Johannes Erdmann for providing me with comments about the analysis and how to explain what is done in a written form.

I very much want to thank the people of Dorph, a small group of people that once started studying physics in 2006 (at least most of them ;)). By now we are distributed over the country but the group still persists. People like them made it very easy for me to feel at home when I moved to Dortmund in 2006. The same applies to the participants and especially the organisers Stefan, Thorsten, Julian and Nico of our Physikerfußball group.

I thank my parents Margret and Laurenz and my sister Carolin for being such a great family that gave me unreserved support and protection against risks in all phases of life. Last but not least I want to thank my wonderful wife Karin for so many things. It is invaluable if there is somebody listening (or pretending to be listening ;)) to the concerns and needs of a particle physicist. Thank you for all your patience, support and love I enjoy every day.