FACHBEREICH MATHEMATIK UND
NATURWISSENSCHAFTEN
BERGISCHE UNIVERSITÄT WUPPERTAL

Estimation of the W+jet background in the analysis of semileptonic top-quark pair events

Dissertation
von
Thorsten T. Boek

March 2013

Introduction

One of the fundamental questions in science is "What is matter made of?". One aim of elementary particle physics is to solve this question. Therefore, the building blocks of matter and their interactions are analyzed in many experiments. The worldwide largest particle accelerator is the Large Hadron Collider (LHC), situated at the European Organization of Nuclear Research (CERN) near Geneva, Switzerland. It is constructed to provide proton-proton collisions at a center-of-mass energy up to $\sqrt{s} = 14$ TeV. Four large experiments have been constructed at the accelerator ring to analyze the remnants of the pp collisions. One of these experiments is the 'A Toroidal Large Aircoil System' (ATLAS) experiment. The ATLAS detector is a general purpose detector, which is designed to cover a broad range of topics in high energy physics. This thesis has been performed within the ATLAS collaboration.

The current theoretical knowledge of the elementary particles and their interactions is collected in the Standard Model (SM) of particle physics. The SM uses a relativistic quantum field theory to describe the interactions of the fundamental particles. Within the SM, matter consists of 12 fundamental particles, divided into 6 leptons and 6 quarks. They interact by the strong, the weak and the electromagnetic interaction. The interaction is described by the exchange of a boson. The $W^{\pm}$ and the Z^0 boson, for instance, are the mediators of the weak interaction. The gravity, with a coupling strength tens of orders of magnitude smaller than the other interactions, is not included in the SM. Another point of interest are the masses of the elementary particles. Even if the SM Higgs mechanism is used to afflict the elementary particles with mass, the cause of the wide spread of their masses is not known. Although the SM is not a complete theory, its predictions have passed many precision tests.

The heaviest quark, discovered at the Fermi National Accelerator Laboratory in proton anti-proton collisions at a center-of-mass energy of $\sqrt{s} = 1.8$ TeV in the year 1995, is the top-quark. With a mass of $173.5 \pm 0.6 \pm 0.8$ GeV [1] it is about forty times heavier than the second heaviest quark, the bottom-quark. At the LHC, the top-quark is produced mainly pairwise via the strong interaction, but also singly via the charged currents of the weak interaction. It decays weakly into a W boson and a bottom-quark. The W boson can decay into a pair of leptons or quarks.

This thesis addresses the top-quark pair production, associated with a semileptonic $t\bar{t}$ decay. The challenge in the analysis of pp collision data is to extract the signal events

from a dataset dominated by a large fraction of background events. Therefore, the selection of events bases on the signature of the signal process. The signal signature of the semileptonic $t\bar{t}$ decay is composed of a high energetic lepton and at least four jets, which are the reconstructed detector objects correspondent to quarks. Although the signal is highly enriched after the event selection, physics processes with a similar event signature remain. Event models are constructed to compare the measurement with the signal and background process expectation. Since some of the background processes are not accurately predicted by perturbation theory, data-driven techniques are applied.

One of the dominant background contributions comes from the W boson production in association with jets, referred to as W+jets. The simulation of W+jets events contains large uncertainties with an increasing number of jets. Because of the presence of a W boson decay, a separation of background and signal processes in data is also challenging. The aspect of this thesis is the improvement of the data-driven W+jets background estimation. An equivalent process to the W+jets production can be observed in the production of a Z-boson in association with jets. Especially the generation of jets in addition the boson production is identical in both processes. In the context of this thesis, two data-driven approaches to estimate the W+jets background are developed, both on the basis of the similarity of W and Z boson events. The results shown in this thesis are based on a dataset recorded at a center-of-mass energy of $\sqrt{s} = 7$ TeV in the year 2011, related to an integrated luminosity of 4.7 fb^{-1}.

The precise validation of $t\bar{t}$ events and their expected backgrounds allows the search for physics processes beyond the SM. Since the top-quark is the heaviest fundamental particle known, it plays an important role in many of these theories. Therefore, new particles with a coupling to the top-quark are predicted in theories beyond the SM. A production mechanism of top-quark pairs, in addition to the strong production process, is expected to be observed in the invariant $t\bar{t}$ mass distribution. The determined W+jets background estimation is evaluated in the analysis of resonant $t\bar{t}$ production in the last chapter of this thesis.

Contents

Chapter 1

Elementary Particle Physics

The Standard Model (SM) of elementary particle physics [2–8] provides the theoretical framework to describe the fundamental particles and their interactions. The precision of the SM was tested in various experiments during the last 40 years. However, no significant deviation from its predictions could be observed. Nevertheless, the SM does not provide a complete theory since the gravitational force is not included in the theoretical framework. The origin of the masses of the elementary particles, connected in the SM with the so-called Higgs mechanism [5], is of current interest. An evidence for this mechanism was discovered in 2012. The full confirmation of the SM Higgs mechanism would be a big success for the SM theory. However, several extensions of the SM are developed to solve the question of the elementary-particle masses. The heaviest elementary particle, the top quark, is supposed to play a special role in many of these theories. The discovery of new particles with a preferred coupling to top quarks is one of the prospects at the LHC.

In the first part of this chapter, the phenomenological structure of the SM is introduced. Furthermore, a brief illustration of the Higgs mechanism is given. A description of the SM top-quark pair production and an overview of expected top-quark-pair production beyond the SM is given in the second part of this chapter. Since this chapter only serves as introduction to the topic of this thesis, it is mainly based on the given sources. A more detailed summary can be found in [9, 10].

1.1 The Standard Model

Quantum Electrodynamics (QED) describes the electromagnetic interaction, combining quantum mechanics and special relativity. The strong force, for instance present in the interactions within the hadron (like the proton or neutron), is described by Quantum Chromodynamics (QCD), a quantum field theory similar to

the QED. A third fundamental interaction, the weak interaction is, among others, responsible for radioactive decays.

The SM of elementary particle physics comprises those three fundamental interactions, described in a quantum field theory. It describes all fundamental constituents of matter and its interaction. As it will be shown later, matter consists of fermions while its interactions are mediated by gauge bosons. Each interaction can be described by a separate theory. In an electroweak unification, described in this chapter, the electromagnetic and the weak interactions are combined in one theory. The gravitational force, with a relative strength that is more than 30 magnitudes weaker than the other interactions, is not included.

1.1.1 Matter Particles

The SM includes twelve point-like fundamental particles, which carry half-integral spin. These fermions are divided into quarks and leptons and grouped in three generations. The particles within the generations differ strongly in their mass. The second and third generation are built up by heavier copies of the first one. Each generation consists of two quarks, an elementary charged and a neutral lepton. The fermions of the SM and their properties are summarized in the first three columns of Figure 1.1.

Each fermion has a corresponding antiparticle with the same mass, spin and lifetime, but opposite signed quantum numbers like charge, baryon and lepton numbers. The fermions of the first generation are the constituents of the visible matter in the universe. The heavier fermions of the other two generations decay into the light and stable particles of the first generation. Whereas the masses of the charged fermions in the first generation are of the order of a few MeV, the heaviest elementary particle, the top-quark of the third generation, has a mass of $(173.5 \pm 0.6 \pm 0.8)$ GeV [1]. Even if a mechanism to generate fermion masses is included in the SM, the cause of the large range of the fermion masses is unknown.

The best-known lepton is the elementary charged Q_e electron e. The muons μ and tauons τ are the leptons of the second and third generations. Three neutrinos (ν_e, ν_μ and ν_τ) are the electrically neutral partners of the charged leptons.

The two quarks in each generation have fractional electrical charges of $+\frac{2}{3}Q_e$ and $-\frac{1}{3}Q_e$. Up (u) and down (d) quarks are the valence quarks of the electrically positive proton (uud) and the neutral neutron (udd). The electrically positive charged quarks, named up-type quarks, in the second and third generation, are the charm (c) and the top (t) quark while the strange (s) and the bottom (b) quarks are their negatively charged down-type quark complements.

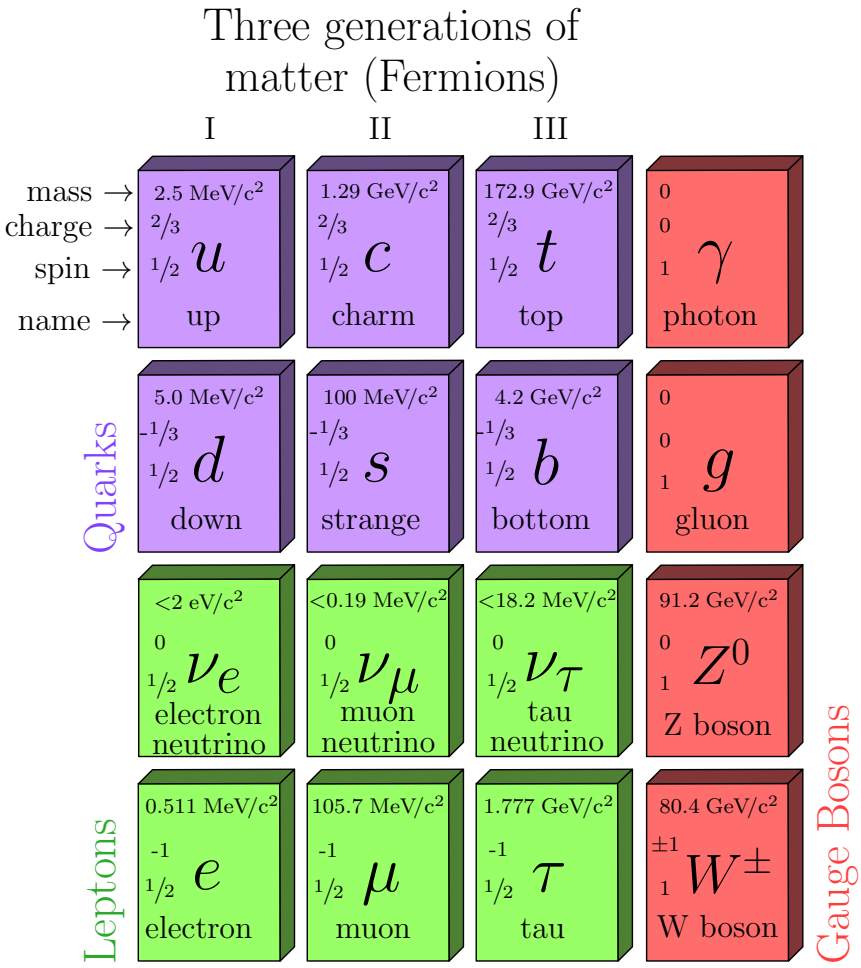


Figure 1.1: *The elementary fermions and the gauge bosons of the Standard Model are shown.*

1.1.2 Interactions

In Hamiltonian mechanics, the Lagrangian was introduced to describe the particle equation of motion in discrete systems. Within quantum-field theory [8], it is expanded to continuously varying systems. The requirement of invariance under local gauge transformations leads to the implementation of symmetry groups $SU(3)_C \times SU(2)_L \times U(1)_Y$ which correspond to the fundamental interactions of the SM. The mediators of the fermion interactions are spin 1 particles. These gauge bosons, depicted in the fourth column of Figure 1.1, are related to the gauge symmetries [9]. The photon γ and the gluons g are the gauge bosons of the electromagnetic and the strong interaction. The weak interaction is associated with an electrically neutral charged Z^0 and two oppositely charged $W^{\pm}$ gauge bosons. The gauge bosons couple to the fermion charges, identified by the conserved quantity of their symmetry. Table 1.1 shows a sketch of the fermion couplings to the fundamental SM interactions.

interactions	neutrinos	charged leptons	quarks
weak	✓	✓	✓
electromagnetic		✓	✓
strong			✓

Table 1.1: *Coupling of the fermions to the SM interactions.*

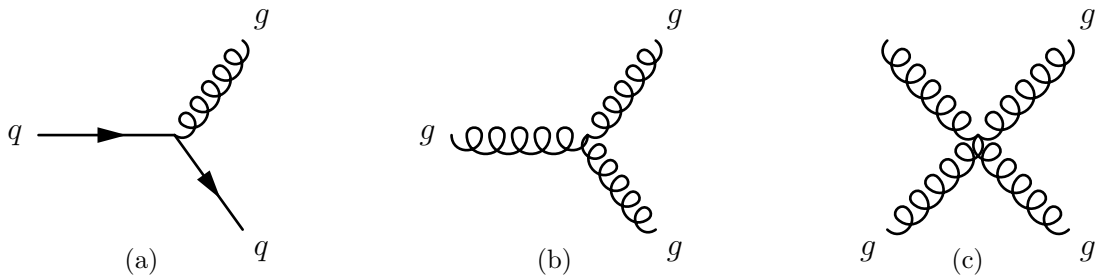


Figure 1.2: *Feynman diagrams of the three fundamental QCD interactions. The gluon radiation from a quark (a) and the gluon self-coupling diagrams with three (b) and four gluons (c) are shown.*

Strong Interaction

The strong interaction is described by the QCD, introducing a $SU(3)_C$ color symmetry group. Three states of color f (red, blue and green) and the corresponding anti-colors $\bar{f}$ are identified as conserved quantity. The mediators of the strong interaction are eight massless gluons, charged with a color doublet $f\bar{f}'$. Since the gluon carries a color charge, gluon-self-coupling is allowed. Quarks are color charged singlets, all equally affected by the strong interaction. The fundamental Feynman diagrams of QCD interactions are shown in Figure 1.2.

In high energetic processes, quarks can be handled as almost free objects. In the low energy, large distance regime, however, free color-charged objects do not exist. They bundle to composed objects, the hadrons. Hadrons are classified into bound states of a quark and an anti-quark or of three quarks, named meson or baryon, respectively. In addition to these valence quarks, the hadrons contain an indefinite number of virtual quarks, referred to as sea quarks, and gluons.

The Electroweak Interaction

The electroweak theory, developed by Glashow, Weinberg and Salam [2,3], describes electromagnetic and weak interactions in a combined theory. Before introducing the unification, the electromagnetic and weak interaction is discussed separately.

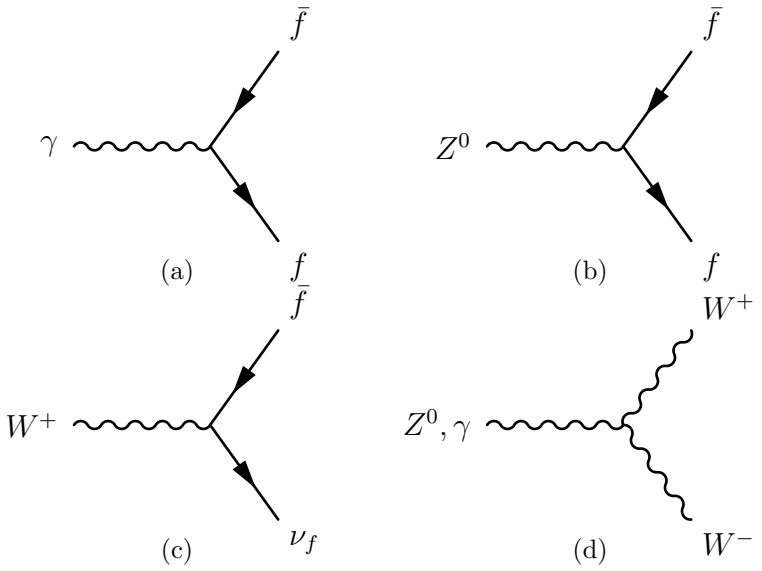


Figure 1.3: *Feynman diagrams of electroweak interactions. The fermion pair production via a photon (a), a Z^0 boson (b) or a W boson (c) are shown. The interaction of γ and Z^0 bosons and the weak charged-current interaction look identical. In addition, Figure (d) shows the vector boson self-coupling.*

Electromagnetic Interaction

The electromagnetic (EM) interaction is described by the QED, introducing a $U(1)_Q$ symmetry group. The photon is the gauge boson of the EM interaction. It couples to electrical charge. The fundamental Feynman diagrams of a QED interaction are shown in Figure 1.3(a). Since the uncharged photon itself is massless, the EM coupling has an infinite range.

Weak Interaction

The weak interaction couples to all fermions. The mediators of the weak interaction are the Z^0 and the $W^{\pm}$ bosons. While the gauge bosons of the electromagnetic and strong interaction are massless, the mediators of the weak interaction are heavy. Hence, the range of the weak interaction is short and the interaction appears weak. The coupling strength of the weak interaction depends on the involved fermions. While the coupling is universal for leptons, it differs regarding the quark flavors. Furthermore, it depends on the particle helicity, defined by the orientation of the spin with respect to the direction of flight. The $W^{\pm}$ bosons interact only with left-handed particles. The Z^0 boson interacts with both chiralities, but with different coupling strengths.

The weak interaction of a $W^{\pm}$ with a quark is the sole possibility in the SM to alter the quark flavor. Hence, in nature, the charged-current weak-interaction is responsible for the radioactive decay of hadrons. The preferential process is the change of an up-type quark into a down-type quark or vice versa without a change of the

quark generation. However, also a change across the generations is allowed. The quark-flavor eigenstates of the weak interaction are not identical to their mass eigenstates. A unitary transformation matrix, the Cabibbo-Kobayashi-Maskawa (CKM) matrix [4], connects flavor and mass eigenstates:

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}$$

The strength of the charged-current interaction between two quarks i and j is proportional to the CKM matrix element $|V_{ij}|^2$. Since these values are not predicted by the theory, they have to be measured in data. Their current numerical values are [11]:

$$V_{CKM} = \begin{pmatrix} 0.97413 \text{ to } 0.97443 & 0.2246 \text{ to } 0.2260 & 0.00330 \text{ to } 0.00358 \\ 0.2245 \text{ to } 0.2259 & 0.97330 \text{ to } 0.97329 & 0.0421 \text{ to } 0.0403 \\ 0.00836 \text{ to } 0.00822 & 0.0414 \text{ to } 0.0396 & 0.999122 \text{ to } 0.999107 \end{pmatrix}$$

The CMK matrix is almost diagonal. Its off-diagonal entries are small. The largest mixing across the generations is observed between the first and the second generation.

Electroweak Unification

The electromagnetic and the weak interaction are successfully described by a unified theory [3], based on a $SU(2)_L \times U(1)_Y$ gauge symmetry. The conserved quantity of the $SU(2)_L$ symmetry group is the weak isospin I. The generator of the $U(1)_Y$ gauge symmetry is the hypercharge Y. It is defined as $Y = 2(Q - I_3)$, connected to the electrical charge generator Q and the third isospin component I_3. The electroweak quantum numbers of the fermions are shown in Table 1.2. The predicted mediators of the $SU(2)_L \times U(1)_Y$ symmetry group are four gauge bosons $W_\mu^{(1)}, W_\mu^{(2)}, W_\mu^{(3)}$ and B_μ, arranged in an isospin triplet and an isospin singlet. The three experimentally observed gauge bosons of the weak interaction W_μ^+, W_μ^- and Z_μ and the photon γ of the QED can be represented as linear combinations of these fundamental gauge bosons [9]:

$$W_\mu^\pm = \frac{1}{2}[W_\mu^{(1)} \pm iW_\mu^{(2)}] \tag{1.1}$$

$$Z_\mu = W_\mu^{(3)} cos\Theta_W - B_\mu sin\Theta_W \tag{1.2}$$

	Generation 1.	2.	3.	I	I_3	Q	Y
Leptons	$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L$	$^1/_2$	$^1/_2$ $-^1/_2$	0 -1	-1
	e_R	μ_R	τ_R	0	0	-1	-2
Quarks	$\begin{pmatrix}u\\ d'\end{pmatrix}_L$	$\begin{pmatrix}c\\ s'\end{pmatrix}_L$	$\begin{pmatrix}t\\ b'\end{pmatrix}_L$	$^1/_2$	$^1/_2$ $-^1/_2$	$^2/_3$ $-^1/_3$	$^1/_3$
	u_R	c_R	t_R	0	0	$^2/_3$	$^4/_3$
	d_R	s_R	b_R	0	0	$-^1/_3$	$-^2/_3$

Table 1.2: *Quantum numbers of the electroweak interaction. The left- (L) and right-handed (R) fermion charges of the weak isospin 'I', its third component 'I_3', the charge 'Q' and the hypercharge 'Y' [10] are shown.*

$$\gamma = W_\mu^{(3)} cos\Theta_W + B_\mu sin\Theta_W \tag{1.3}$$

The angle Θ_W, referred to as Weinberg angle, is the electroweak mixing angle. While the photon is massless, the bosons of the weak interaction are known to be massive. They obtain their masses through the concept of spontaneous-symmetry breaking within the SM Higgs mechanism [5]. The Higgs mechanism introduces a complex scalar isospin doublet [5]:

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix}\Phi^+\\ \Phi^0\end{pmatrix} \tag{1.4}$$

The consistency with the gauge symmetry is achieved with the scalar potential [5]:

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \tag{1.5}$$

The parameters μ and λ, defined as $\mu^2 < 0$ and $\lambda > 0$, are connected to the Higgs boson mass and the Higgs boson self-coupling, respectively. The symmetry of the potential is broken by choosing a particular ground state. This non-zero vacuum expectation value is defined as [5]:

$$\Phi(0) = \frac{1}{\sqrt{2}} \begin{pmatrix}0\\ v\end{pmatrix} \quad \text{with} \quad v^2 = -\frac{\mu^2}{\lambda} \tag{1.6}$$

While the parameter v is experimentally determined to be $v = 246$ GeV, the parameters μ and λ, interdependently from each other, are undetermined. They can

be identified with the discovery and a precise survey of a SM Higgs boson. The fermions obtain their masses in the theory by the introduction of a Yukawa coupling to the Higgs field. The strength of the couplings is associated with the mass of the fermions.

1.1.3 The Higgs Boson

Three of the four components of the Higgs field, shown in Equation 1.4, provide the longitudinal degrees of freedom of $W^{\pm}$ and Z^0 bosons. A fourth component [5]:

$$H = \frac{1}{\sqrt{2}}(\Phi^0 + \bar{\Phi}^0) \tag{1.7}$$

leads to a neutral Higgs scalar. This Higgs boson is predicted to exist as a free particle. Due to the parameter μ, the mass of the Higgs boson is not predicted in the SM. Hence, it has to be determined in experiments. In Figure 1.4 the Feynman diagrams of the favored Higgs-boson-production channels at the LHC are shown. The top quark is the fermion with the by far largest Yukawa coupling and therefore the preferred coupling candidate in the fermion sector. Furthermore, the Higgs boson couples to the massive vector-bosons.

A new boson, discovered in the year 2012 at the LHC, with a mass of approximately 125 GeV, is a good candidate for the SM Higgs boson. The analyzed decay channels are $H \to ZZ^* \to llll$, $H \to WW^* \to e\nu\mu\nu$, $H \to \gamma\gamma$, $H \to b\bar{b}$ and $H \to \tau^+\tau^-$. Figure 1.5(a) shows the invariant di-photon mass for the $H \to \gamma\gamma$ channel, indicating a peak in the spectrum. With a significance of 5.9 σ [12] and 5.8 σ [13] a new boson was detected by the ATLAS and the CMS collaboration, respectively. The discovery of a new boson is achieved with the combination of 2011 and 2012 LHC-collision data. The reconstructed mass of the boson is:

ATLAS [12]: $\mathrm{m_H} = (126.0 \pm 0.4(\mathrm{stat}) \pm 0.4(\mathrm{sys}))$ GeV

CMS [13]: $\mathrm{m_H} = (125.3 \pm 04(\mathrm{stat}) \pm 0.5(\mathrm{sys}))$ GeV

In Figure 1.5(b) the 95% CL upper limit on the signal strength as a function of the Higgs mass in the combination of all channels analyzed in ATLAS is shown. An important aspect is the consistency of the discovered boson with the SM prediction. Hence, further measurements on Higgs properties, such as the spin, the coupling to the fermions and the self coupling strength, have to be determined. These measurements have already started. An update on the ATLAS measurement of the determined Higgs boson mass for instance can be found in [**?**]. Furthermore, an exclusion of a benchmark spin-2 model for the observed particle in the WW decay channels at 95% confidence level [**?**] and the $\gamma\gamma$ decay channels in the gluon-fusion production process at 99% confidence level [**?**] is in favor of the spin 0 prediction of the SM.

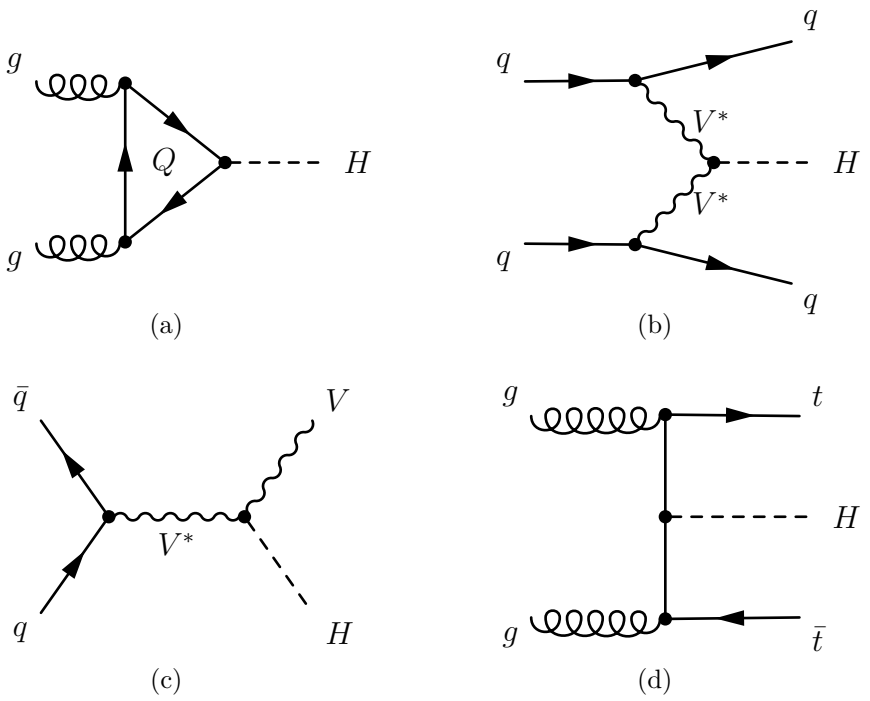


Figure 1.4: *Feynman diagrams for Higgs boson production at the LHC, including gluon-gluon fusion (a), vector boson fusion (b), Higgs bremsstrahlung (c) and top-quark associated production (d).*

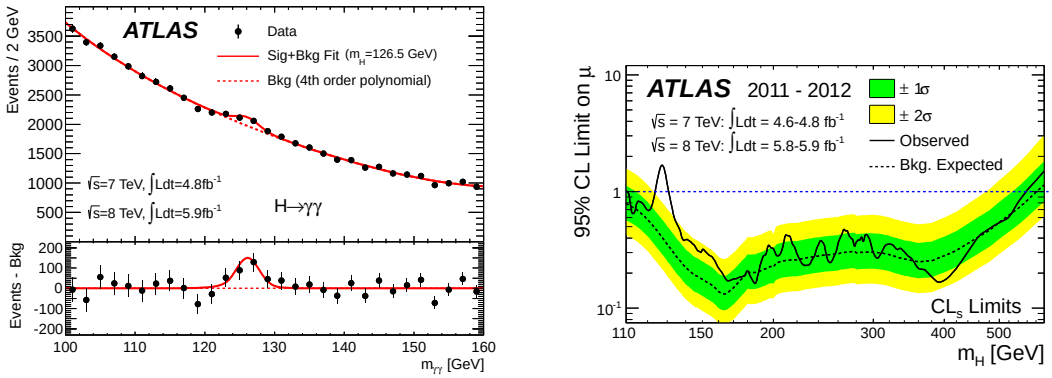


Figure 1.5: *Figure (a) shows the invariant di-photon mass distribution of the Higgs decay channel into two photons, measured in* $\sqrt{s} = 7$ TeV *and* $\sqrt{s} = 8$ TeV *ATLAS data. The dashed line shows the background fit while the solid line shows the combination with a* 126.5 GeV *Higgs signal [12]. Figure (b) shows the observed (solid) 95% CL upper limit on the signal strength as a function of the Higgs mass from combined ATLAS data [12]. The expected limit with an background only hypothesis is shown as dashed line including 1 and 2* σ *uncertainty bands.*

1.2 The Top Quark

The top quark, discovered in 1995 at the Tevatron collider of the Fermi National Accelerator Laboratory near Chicago, USA [14], is the heaviest fermion in the Standard Model of elementary particle physics. Its mass of $m_{top} \approx 173$ GeV/c^2 is in the same order as the electroweak symmetry-breaking scale. Therefore, the top quark plays an important role in many of the physical processes observed at the LHC. One example is the strong fermion coupling to the SM Higgs boson. The verification of the top quark associated Higgs boson production for example, shown in Figure 1.4(a) and 1.4(d), will allow a direct measurement of the top quark Yukawa coupling.

1.2.1 Standard Model Top Physics

The top quark is the weak isospin partner of the bottom quark in the third generation. The top pair production via the strong force is the dominant production channel. Although the top quark decays via the weak interaction, the lifetime of $\tau = 0.5 \times 10^{-24}$ s [1] leads to a decay before the formation process of hadrons starts. Hence, it is the only quark which can be accessed as almost free quark.

Top Quark Production

The production of top quarks in the SM can be divided into top-quark pair production and single-top quark production. The top-quark pair production via strong interaction leads to a total cross-section of $\sigma = 166.8^{+16.5}_{-17.8}$ pb [15] at a centre-of-mass energy of 7 TeV. Top quark pairs are mainly produced by gluon-gluon fusion processes in the proton collisions at the LHC. All tree level LO production mechanism are shown in the Feynman diagrams a, b and c of Figure 1.6.

The production of single top quarks is based on the weak interaction. In the Feynman diagrams d, e and f of Figure 1.6 the production channels are shown. At a centre-of-mass energy of 7 TeV the single-top cross-section is $\sigma = 64.6$ pb^{-1} [16], $\sigma = 4.6$ pb^{-1} [17], $\sigma = 15.7$ pb^{-1} [18] for the t-, s- and Wt- production channels, respectively.

Top Quark Decay

Due to its high mass, the top quark has a very short lifetime. With a probability of almost 100%, related to the CKM matrix element V_{tb}, the top-quark will decay into a b-quark and a W boson before the hadronization process starts. The W boson decays into two quarks or a lepton-neutrino pair, respectively. Hence, the decay of a top-quark pair is divided into three final states, corresponding to the number of leptons generated in the decay of the two W bosons.

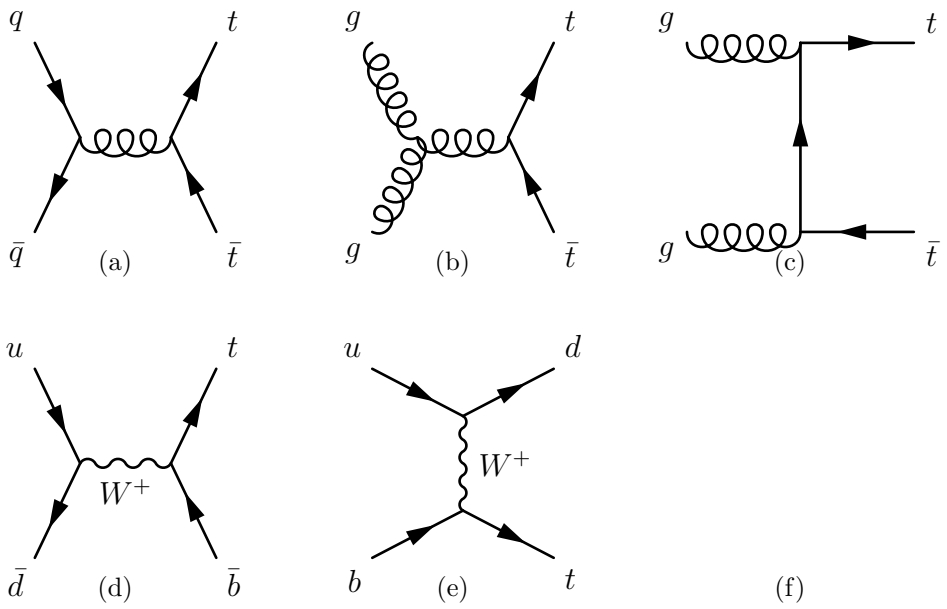


Figure 1.6: *Feynman diagrams of top quark production (LO). In the first row, top quark pair production processes like the quark anti-quark annihilation (a) and the gluon-fusion ((b) and (c)) processes are shown. The second row shows processes of single top quark production in the s-channel (d), the t-channel (e) and the W boson associated (f) production.*

With a probability of 10.3%, the leptonic decay of both W bosons leads to two lepton-neutrino pairs plus two b-quarks in the final state. Both W bosons decay hadronically with a probability of 46.2%. The final state of these decays consists of six quarks, the two b-quarks from the top-quark decay and four quarks from the hadronic W-boson decays.

In 43.5% of the $t\bar{t}$ decays one W boson decays leptonically while the other one decays hadronically. The final state is built up by four quarks, two of them b-quarks from the top decay, and a lepton-neutrino pair. The Feynman diagram of the semi-leptonic top-quark pair decay, analyzed in this thesis, is shown in Figure 1.7.

Top Quark Mass

The top quark mass was precisely measured at the Tevatron collider, colliding protons and anti-protons at a centre-of-mass energy of 1.96 TeV in RunII. First measurements of the top quark mass at the LHC experiments ATLAS and CMS are shown in Figure 1.8. The actual value of the top-quark mass, combining all measurements done at the Tevatron collider, is $(173.5 \pm 0.6 \pm 0.8)$ GeV [1].

Figure 1.7: *Feynman diagram of a semi-leptonic $t\bar{t}$ process. The top-quarks, produced by the hard process, decay into a $W^{\pm}$ boson and a $b(\bar{b})$ quark. The decay of the $W^{\pm}$ bosons into a charged lepton and its (anti-)neutrino and into a quark-pair respectively, leads to a final state of four quarks and a lepton-neutrino pair.*

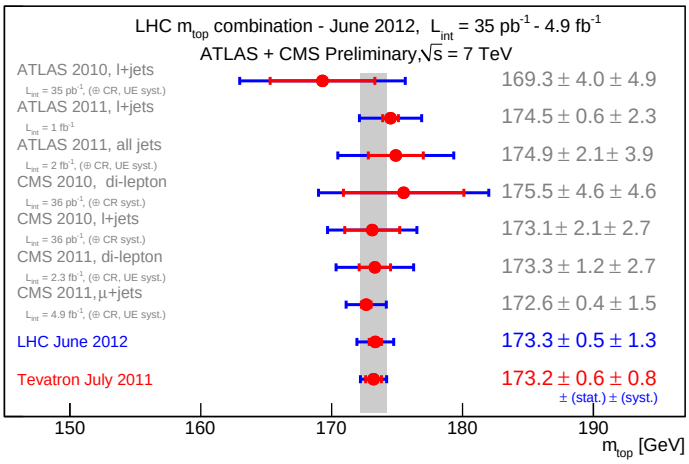


Figure 1.8: *ATLAS and CMS measurements of the top mass based on 4.9 fb^{-1} [19]. The recent Tevatron measurement is given as reference below.*

1.2.2 Extensions of the Standard Model

Even if the SM predictions are precisely verified by a large number of experiments, there are still a couple of issues left [10]. For instance:

- The Planck scale of 10^{19} GeV is the typical gravitational mass scale. Loop-corrections in the quantum-field theory lead to an increase of the effective mass of scalar particles. Hence, the natural value of the SM Higgs-boson mass would be expected in the order of the Planck scale.

- Neutrinos, only existing in one helicity state in the SM, are theoretically assumed to be massless. However, neutrino-flavor oscillations, associated with a non-zero neutrino-mass, are experimentally observed.

- In the SM electroweak-symmetry breaking the fermion masses enter as free coupling parameters. No mechanism in the SM is able to predict the exact masses of the fermions and the huge differences between the fermion masses accordingly.

These are only some of the reasons that motivate a search with focus on physical phenomena beyond the SM (BSM) parametrization. The discovery of the new boson could only be the beginning of new discoveries at the LHC. Therefore, searching for new particles and resonances, predicted by theories extending the SM, is an important challenge.

BSM Top Pair Production

The top quark with its high Yukawa coupling plays a very important role in many physics processes at the LHC. Therefore, the top quark, with a mass of the order of the electroweak-symmetry breaking $m_{top} \approx \frac{v}{\sqrt{2}}$ is a good probe of alternative models beyond the SM. The need of mass generation motivates various theories beyond the SM. New resonances, predicted by some of these theories, are supposed to have a strong coupling to the top quark. Top-quark pair production without strong interaction processes, shown in Figure 1.9, would increase the production cross-section and manifest itself in distributions, such as the invariant-top-pair mass.

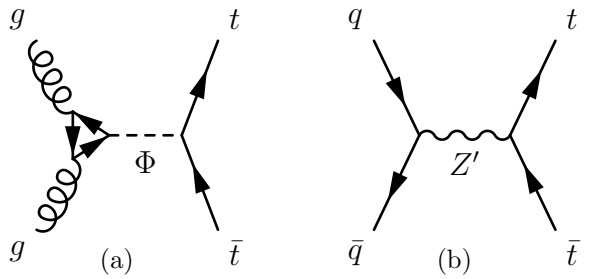


Figure 1.9: *Feynman diagrams of a possible resonance decaying into top pairs. Figure 1.9(a) shows a Spin-0 color-singlet or octet particle while in Figure 1.9(b) a Spin-1 resonance is shown.*

The nature of these resonances, concerning the mass, width and couplings to the SM particles, strongly depends on the theoretical model. The following section will introduce two models that are considered in the subsequent analysis.

Topcolor Assisted Technicolor

The approach of 'technicolor' [20] is based on a dynamically broken theory of weak interaction. As a consequence, new fermions with a new strong interaction at ≈ 1 TeV

appear. Color-octet gauge bosons, referred to as colorons, provide a mechanism of fermion-mass generation. They mediate the interaction between fermions and technifermions. In the model of 'topcolor' [21], the mass of the top quark is arranged by a $t\bar{t}$ condensate of a new strong force. A combination of both theories, established in the 'topcolor assisted technicolor' model [22], provides a natural explanation of the large top quark mass. The symmetry breaking is mainly driven by the concept of technicolor. An additional $U(1)'$ gauge symmetry predicts a neutral spin-1 color-singlet Z' boson, produced via quark anti-quark annihilation. With the assumption of a suppressed coupling to b-quarks, described in [23], a narrow Z' resonance peak in the invariant $t\bar{t}$ mass spectra is predicted.

Kaluza Klein Gluon

The 'Randall Sundrum' model [24] provides a higher-dimensional mechanism solving the hierarchy problem by postulating the existence of a warped extra dimension. The resulting Kaluza-Klein Gluon g_{kk} is a Spin-2 color-octet particle that preferably couples to the top quark. Hence, a strongly produced, wide resonance decaying into a pair of top-antitop quarks [25] is predicted.

The test and validation of top quark physics and its extensions, based on recoded data of proton collisions at the LHC, is an important elementary particle physics topic in the next years.

Chapter 2

The ATLAS Experiment

The European Organisation of Nuclear Research (CERN) was founded in 1954 close to Geneva in Switzerland. During the last 60 years it has become one of the largest international centers of particle physics.

The world-wide largest particle accelerator, the Large Hadron Collider (LHC), started its operation in 2010. Installed in a 27 km long tunnel of the previous accelerator experiment, the Large Electron Positron (LEP) collider, and equipped with superconducting magnets, the LHC, shown in Figure 2.1, is designed to collide protons and heavy ions with a center of mass energy up to $\sqrt{s}$ =14 TeV for proton collisions. Older accelerators at CERN, like the 7 km long Super Proton Synchrotron (SPS), serve as pre-accelerator for the LHC particle bunches.

2.1 The Large Hadron Collider

The LHC is built to reach an instantaneous luminosity of 10^{34} cm^{-2}s^{-1} by colliding proton bunches every 25 ns. These bunches contain up to 10^{11} protons and will reach a minimal distance of 7 m between each other. At full energy, the protons will have a velocity of 99.9999991% of the speed of light. The dataset analyzed in this thesis is recorded in the year 2011 at $\sqrt{s}$ =7 TeV with a maximum number of 1380 bunches in the LHC, up to 1.45×10^{11} protons per bunch and a maximum luminosity of 3.6×10^{33} cm^{-2}s^{-1} [26].

To keep the particles in the accelerator ring, 1232 superconducting dipole magnets with a length of 15 m are placed all around the ring. In addition, 392 superconducting quadrupole magnets with a length of 5 to 7 m are installed to focus the beam. The operation temperature of these magnets is about 2 Kelvin and the generated magnetic fields reach up to 8.4 Tesla.

Collisions take place in four interaction points where the four main LHC experiments ATLAS, CMS, LHCb and ALICE are constructed.

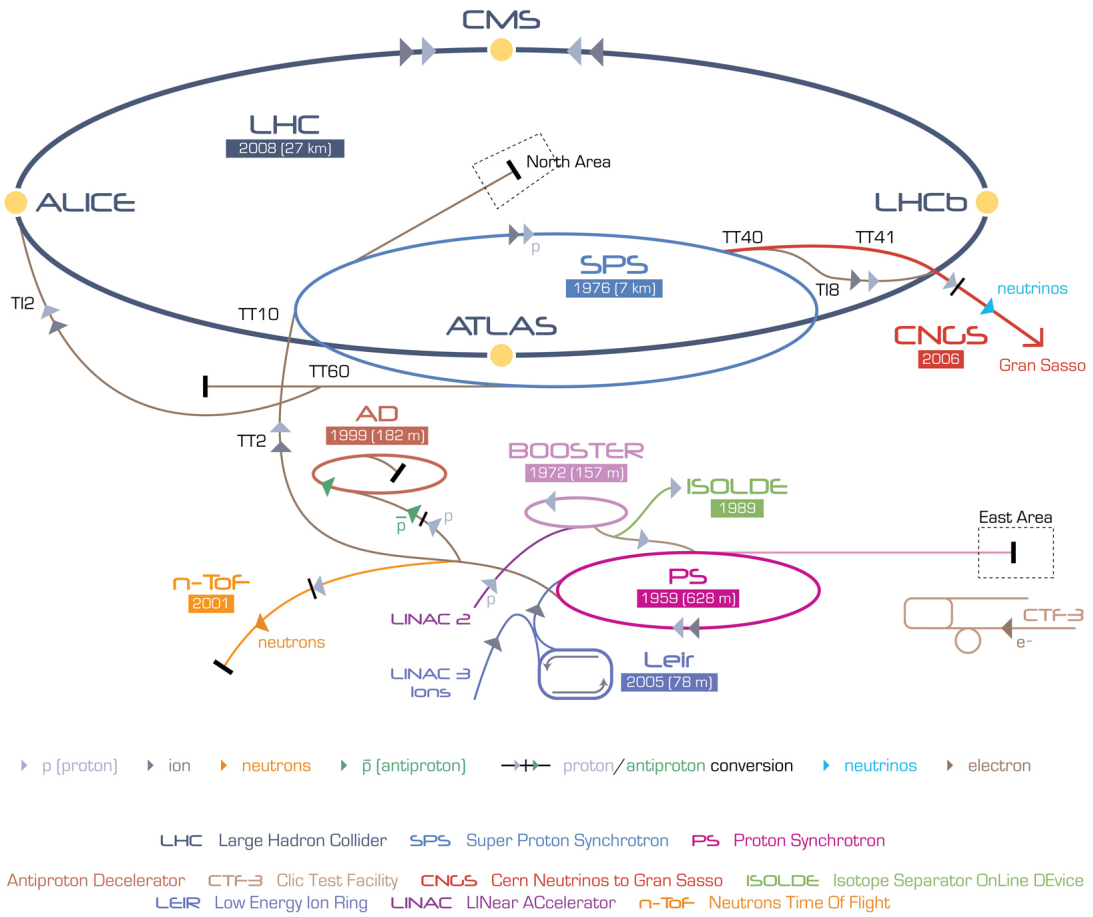


Figure 2.1: *Sketch of the accelerator complex at CERN [27]. The Large Hadron Collider, depicted as blue circle, is designed to provide* $\sqrt{s}$ =14 TeV *proton-proton collisions in four beam interaction points (yellow dots) where the main LHC experiments are placed. The protons for the LHC collisions are produced at LINAC2 and pre-accelerated in various accelerators before the LHC injection.*

ATLAS The "A Toroidal Large Aircoil System" (ATLAS) detector, described in detail in Section 2.2, is one of the two multi-purpose detectors at the LHC. The task of the experiment is to perform more precise measurements of the Standard Model parameters and search for new physics phenomena. To reconstruct all particles produced in the collisions, the detector covers nearly the entire solid angle around the collision region. The design is embossed by the large air coil toroid magnets containing the muon tracking chambers.

CMS The second multi-purpose detector is the "Compact Muon Solenoid" (CMS) detector. Just like the ATLAS detector, it is constructed symmetrically around the point of interaction but in a more compact assembly. Characteristic trait of this experiment is a precise and homogenous calorimeter.

LHCb Searching asymmetries in the B hadron physics, the LHCb detector is interested only in the forward region of the collision events. Therefore, it is the only of the four experiments, which is not symmetrically arranged around the interaction point.

ALICE The experiment with main focus on the heavy ion collisions is the "A Large Ion Collider Experiment" (ALICE). It is specially designed to analyze the produced quark gluon plasma to learn about the processes close to the Big Bang.

2.2 The ATLAS Detector

The ATLAS experiment, described in [28], is a multiple-purpose particle physics detector operating at one of the beam interaction points of the LHC. It is 44 m long, 25 m high, has a weight of 7 kilo tons and covers nearly the entire solid angle around the collision region. As shown in Figure 2.2, it consists of an inner tracking detector surrounded by a thin superconducting solenoid providing a 2 Tesla magnetic field, electromagnetic and hadronic calorimeters, and a muon spectrometer.

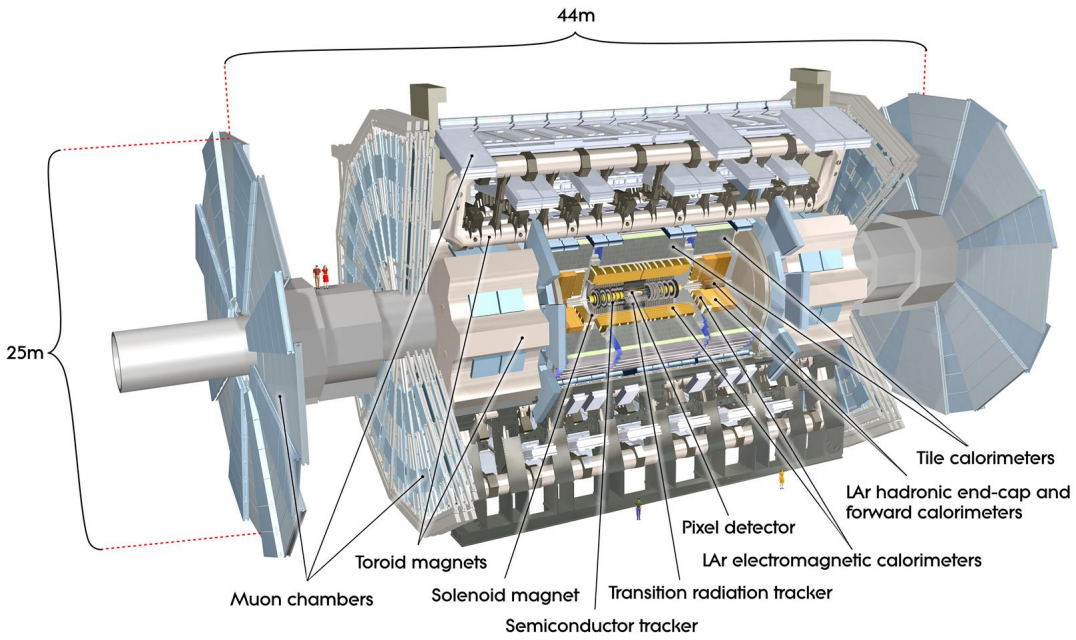


Figure 2.2: *Scheme of the* 44 m *long and* 25 m *high ATLAS detector [28]. The muon spectrometer in the outer part, included into eight big air coil toroidal magnets, embosses the design of the detector.*

This arrangement allows the detection and classification the particles produced in the collisions. Charged fermions leave tracks in the tracking system and deposit their energy in the electromagnetic calorimeter. Hadrons and muons will additionally deposit energy in the hadronic calorimeter. Photons, the only uncharged electromagnetically interacting particles, interact in the electromagnetic calorimeter without a signal in the tracking system. The muon is the sole particle which transverses the calorimeter and leaves an additional signal in the muon chambers. The neutrinos pass the detector without any signal. Since the transverse momentum in the event is conserved, their transverse component can be reconstructed by knowing all other particles and computing the missing transverse energy of the event. A magnet configuration of a thin superconducting solenoid magnet surrounding the inner detector volume and large superconducting toroid magnets in the outer part of the experiment bend the tracks of charged particles to allow a precise momentum reconstruction.

For a common parametrization of the detector a coordinate system, using the nominal interaction point as origin, is defined. The z component is defined to be along the beam pipe of the accelerator. Transverse to the beam pipe the x and y coordinated describe the horizontal and vertical coordinates.
In polar coordinates r is the radial distance to the interaction point, ϕ describes the position in the x,y plane while Θ is the polar angle in the x,z plane. The pseudo-rapidity is defined as $\eta = -log[tan(\frac{\Theta}{2})]$. It becomes a Lorenz transformation invariant parameter, if considered relatively.

A Lorenz transformation invariant parameter pseudo-rapidity is defined as:
$\eta = -log[tan(\frac{\Theta}{2})]$

The ATLAS detector is built including the benchmarks set by various physics constrains. The reconstruction of a Standard Model Higgs boson, for instance, in the various production- and decay channels, described in Chapter 1.1.3, set great demands on all ATLAS detector systems. An accurate photon identification ($H \to \gamma\gamma$), a precise lepton reconstruction (WH, ZH, $H \to ZZ^* \to llll$) and an efficient B-hadron identification inside the hadron showers (ttH, $H \to bb$) are as important as the reconstruction of forward jets to guarantee a good reconstruction sensitivity. The analysis of processes beyond the Standard-Model requires a charge identification of leptons in the TeV range. Other predictions claim additional particles, which pass through the detector without any response. Hence, similar to the reconstruction of neutrinos, a precise measurement of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event is crucial. To fulfill these demands, also high requirements on the detector systems were set. For instance, the material budget in the inner detector has to be low to reduce mis-measurements of particle tracks caused by multiple scattering effects. Also the read out of the detector with a rate of 40 MHz and the high radiation dose demand high standards on the electronic components.

detector component	required resolution	η coverage measurement	η coverage trigger
tracking	$\sigma_{p_T}/p_T = 0.05\% \cdot p_T \oplus 1\%$	± 2.5	-
EM calorimetry	$\sigma_E/E = 10\% \,/\, \sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry			
barrel and end-cap	$\sigma_E/E = 50\% \,/\, \sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward	$\sigma_E/E = 100\% \,/\, \sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7	± 2.4

Table 2.1: *General performance goals of the ATLAS experiment [28]. The units of E and p_T are stated in* GeV.

detector component	measured resolution
ID tracking	$\sigma_{p_T}/p_T = 0.07\% \cdot p_T \oplus 2\%$
Muon spectrometer	$\sigma_{p_T}/p_T = 6\%$ at $p_T = 0.2$ TeV

Table 2.2: *Performance of the ATLAS tracking systems measured in ATLAS data. The resolution of the tracking systems is measured in di-muon Z decays in* 40 pb^{-1} *2011 data [29]. The units of E and p_T are stated in* GeV.

The top quark is produced at the LHC with a rate of a few tens of Hz in design luminosity operation mode. For 2011, a maximal frequency of 0.5 Hz was reached, assuming a top-quark pair production cross-section of approximately 161 pb and a maximal luminosity of 3.6 $10^{33}\text{cm}^{-2}\text{s}^{-1}$. To analyze the semileptonic decay of the top-quark pair, analyzed this thesis, the reconstruction of one isolated lepton, jets with additional flavor information and the E_T^{miss} in the event is needed. Hence, a good performance of all ATLAS detector subsystems has to be required:

- a high-resoluting tracking system to determine displaced vertices from the decay of B-hadrons
- exact track reconstruction of leptons in the inner and outer parts of the detector
- precise momenta measurements of the decay products
- full calorimeter coverage to compute the imbalance of transverse momenta caused by the neutrino escape

Table 2.1 shows the performance goals of the ATLAS detector systems. In comparison, Table 2.2 presents first measurements on the ATLAS performance in 2011 data.

2.2.1 Muon Spectrometer

The muon system, marked in blue in Figure 2.2, is instrumented with trigger and high resolution drift chambers. It has three cylindrical layers and is built into the eight huge air coil magnets, forming a toroidal magnetic field up to 5.5 Tesla, in the barrel region of $0<|\eta|<1.4$. In the forward region of $1.6<|\eta|<2.7$, the drift chambers are assembled in three large wheels per end-cap, enclosed by two end-cap magnets per side.

The muon spectrometer contains four different detection systems. Monitored Drift Tubes (MDT) with $|\eta|<2.7$ (innermost layer $|\eta|<2.0$) and more precise Cathode Strip Chambers (CSC) at high $|\eta|$ ($2.0<|\eta|<2.7$) provide information for the precision tracking. Resistive Plate Chambers (RPC) in the barrel ($|\eta|<1.05$) and Thin Gap Chambers (TGC) in the end-cap regions ($1.05<|\eta|<2.7$ (<2.4 for trigger information)) are responsible for the trigger information, including well defined p_{T} thresholds, and provide additional muon coordinates.

2.2.2 Calorimeter Design

The composition of the ATLAS calorimeter is shown in Figure 2.3, covering the range up to $|\eta| < 4.9$. ATLAS is using various techniques of sampling calorimeters with the goal to completely stop electromagnetically and hadronically interacting particles and to perform a precise measurement of the originating shower. The difference between electromagnetic and hadronic calorimeter is the choice of absorber material and granularity.

Hadronic Calorimeter

The broader granularity of the hadronic calorimeter is a compromise between cost and sufficient precision for jet reconstruction and the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement. In the outer part of Figure 2.3, the hadronic calorimeter is shown. It is divided into a barrel ("Tile barrel" and "Tile extended barrel") and an end-cap region ("LAr hadronic end-cap (HEC)"). The goal of the hadronic calorimeter is to stop all hadronic particles and measure their position and the amount of energy deposition.
The absorber material of the tile barrel and tile extended barrel calorimeter in the central region $|\eta| < 1.7$ is realized by steel planes, discontinued by scintillating tiles as active material.
Hadronically interacting particles shower in the steel planes. The shower particles will activate the scintillator disks proportional to the deposited energy of the hadrons.

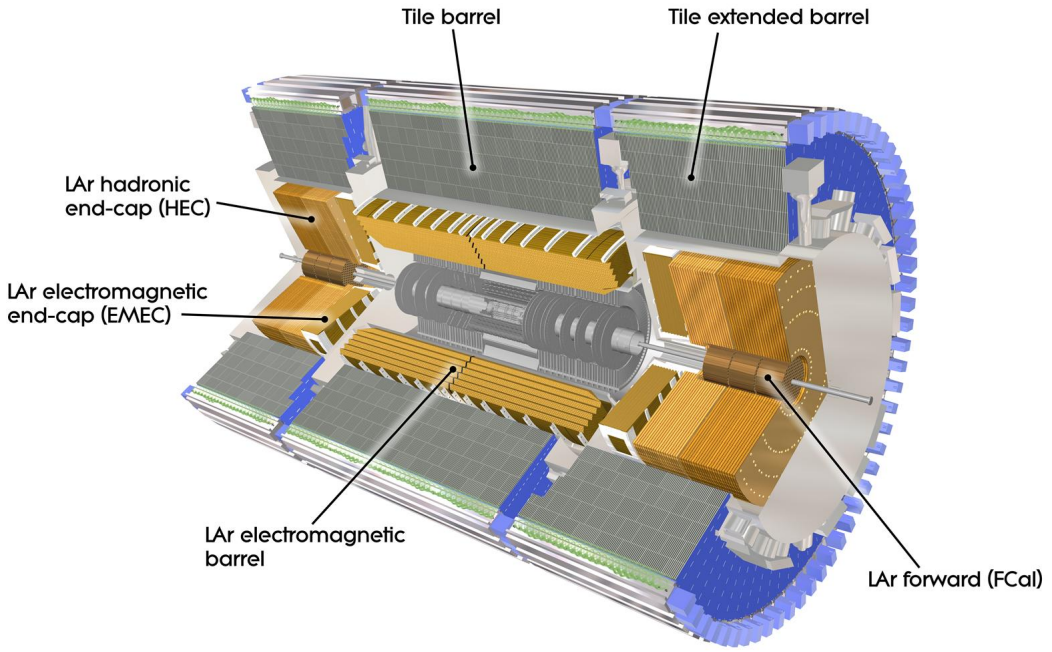


Figure 2.3: *Detailed view on the calorimeter arrangement of the ATLAS detector [28]. Marked in gray the hadronic calorimeter barrel and extended barrel are displayed. Shown in dark orange the forward region of the hadronic calorimeter is shown. The orange section builds the electromagnetic calorimeter barrel and forward region. In the high forward region the combined forward calorimeter, marked in red, is shown.*

The forward region of $|\eta| < 3.2$ is covered on both sides by the LAr Hadronic End Cap (HEC). Here copper is used as absorber. Due to the radioactive environment and limited space, liquid argon is used as detection material as in the bordering electromagnetic end-cap described later on.

Electromagnetic Calorimeter

The electromagnetic liquid argon calorimeter (LAr EM and EMEC in Fig. 2.3) surrounds the inner detector and the solenoid magnet covering a region of $|\eta| < 3.2$. Its accordion structure is filled with liquid argon as active material and lead absorber disks, thus ensuring full coverage in ϕ. A thickness of > 22 radiation length (X_0) in the barrel and > 24 X_0 in the end-cap region and a fine granularity lead to a precision measurement of electrons and photons.

Forward Calorimeter

In the very forward direction, the LAr Forward Calorimeter (FCal) is placed to cover an $|\eta|$ range of up to $|\eta| = 4.9$ and shield the forward muon spectrometer. It consists of one copper and two tungsten absorber layers to serve as electromagnetic and hadronic calorimeter. Main tasks of the FCal are the forward jet reconstruction and the generation of additional $E_{\mathrm{T}}^{\mathrm{miss}}$ balance information.

2.2.3 Inner Detector Tracking System

The inner part of the ATLAS detector consists of a tracking system, shown in Figure 2.4, surrounded by a thin solenoid magnet. The magnetic field of up to 2 Tesla parallel to the beam-pipe bends the tracks and allows to achieve a precise momentum information in the event reconstruction. The task of the inner detector is to detect the trajectory of the particles as precisely as possible with a minimum amount of distortion. To provide a complete picture of the event, it covers a large angular region. The two innermost detector systems can measure charged particles up to a pseudorapidity range $|\eta| < 2.5$.

Figure 2.5 shows the radial positions of the three different detector components. The Transition Radiation Tracker (TRT) with a drift-tube technology forms the outer part of the inner detector. The two inner precision-tracking detectors are based on silicon technology. The Semi Conduction Tracker (SCT) is made of silicon strips while the Pixel detector, to allow a good two particle separation close to the interaction point, is made of silicon pixel cells. The innermost detector layer has a distance of only 50.5 mm from the interaction point. As supporting structure for the silicon modules carbon barrel structures in the central region and end-cap disks for the high $|\eta|$ coverage are designed.

Transition Radiation Tracker

The TRT is built out of straw drift tubes with a diameter of 4 mm, filled with a xenon based gas mixture at overpressure. It consists of 73 axial straw barrel layers and two times 160 radial straw end-cap layers, thus having a total amount of 350.000 channels.

Charged particles ionize the gas in the straw. The ions drift to the anode, the signal amplitudes and arrival times are measured. Ultra relativistic charged particles like electrons produce transition radiation in the radiator material. The absorption of these transition-radiation photons produces a high energy deposition in the straw which allows one to distinguish the signal from a usual ionization. This effect gives an additional handle on the identification of different types of particles leaving signals

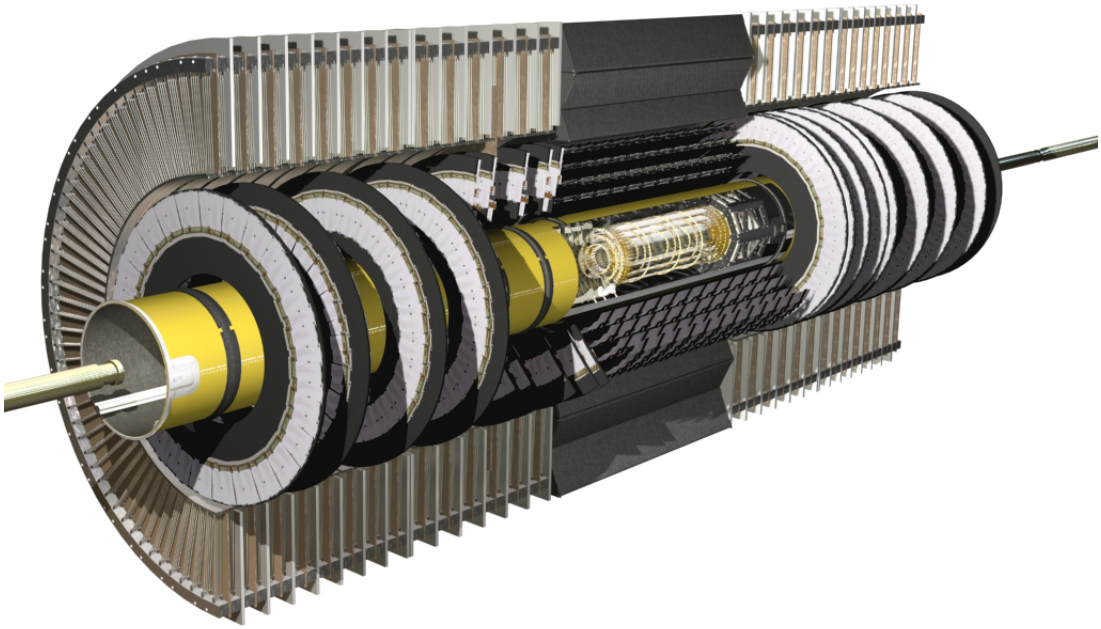


Figure 2.4: *Drawing of the inner detector [28]. Shown is the Transition Radiation Tracker (TRT) in the outer part, the Semi Conduction Tracker (SCT) with its four barrel layers and nine cylindrical disks at each end and the Pixel detector having three layers in the midmost region.*

in the detector. The transition radiation of charged particles and photons can be measured up to a pseudorapidity range $| \eta | < 2.0$.

Whereas its resolution of 130 µm in r-ϕ is worse compared to the other tracking detectors, the TRT provides up to 36 tracking points with almost no disruption of the particles trajectory. Therefore the TRT provides valuable information for the track reconstruction.

Silicon Strip Detector

At a radial distance of 299 mm till 514 mm from the interaction point, the SCT detector is placed. It has four barrel layers and nine end-cap disks on each side with a total amount of 4088 modules and 6.3 million readout channels. With silicon p in n sensor strips of 80 µm width and 6.4 cm length, placed on double sided modules rotated by 40 mrad stereo angle to each other, the hit position is detected. The resolution in the transverse plane is 17 µm in R-ϕ, while it is 580 µm in beam direction.

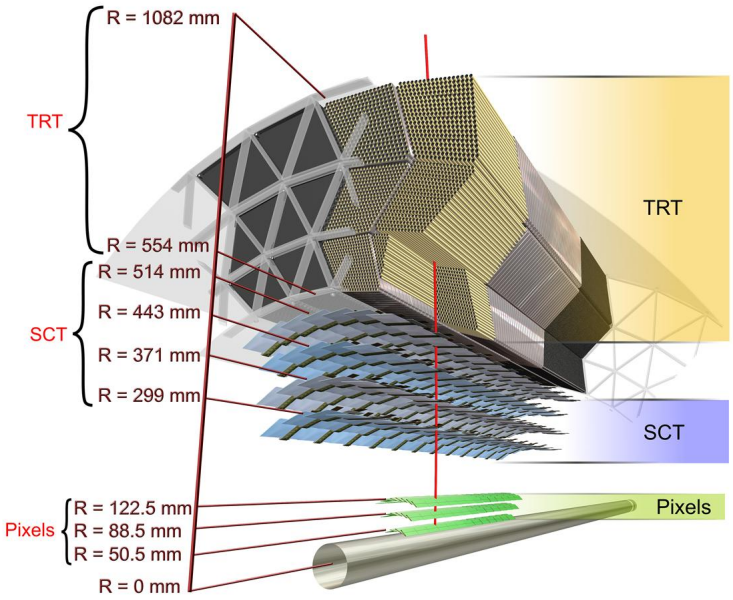


Figure 2.5: *Detailed view of the radial distances of the inner detector component positions with respect to the beam pipe [27]. The Pixel detector has a distance of only* 50.5 mm *while the TRT provides tracking point up to* 1,08 m.

Pixel Detector

The highest granularity is provided by the Pixel detector. At a length of 1.3 m and a diameter of 37.7 cm, the Pixel detector has 80.4 million readout channels which are almost 90% of all channels in the ATLAS experiment. As innermost tracking detector, it is important to reconstruct the collision point of the particles as precisely as possible. The short distance of just 50.5 mm to the interaction point provides a good lever arm for the tracking algorithm. The identification of long lived particles such as B hadrons and τ leptons, decaying in a certain distance to the primary vertex, enters into many physics analyses.

The ATLAS Pixel detector (Fig. 2.6) consists of three barrel layers and three disks at each detector end. The barrel layers are build with carbon staves which hold 13 Pixel-detector modules and provide the cooling. The innermost layer, alias "b-layer", consists of 22 staves. The next two layers have 38 and 52 staves. In the forward regions of the detector, three disks with 48 modules each are in place. A module is 16.4 mm wide and 60.8 mm long and its silicon sensor contains 46080 pixel cells of the size of 50 x 400 μm^2. The 1744 modules on staves and disks overlap each other to reduce non-sensitive areas inside the detector volume. The resolution which can

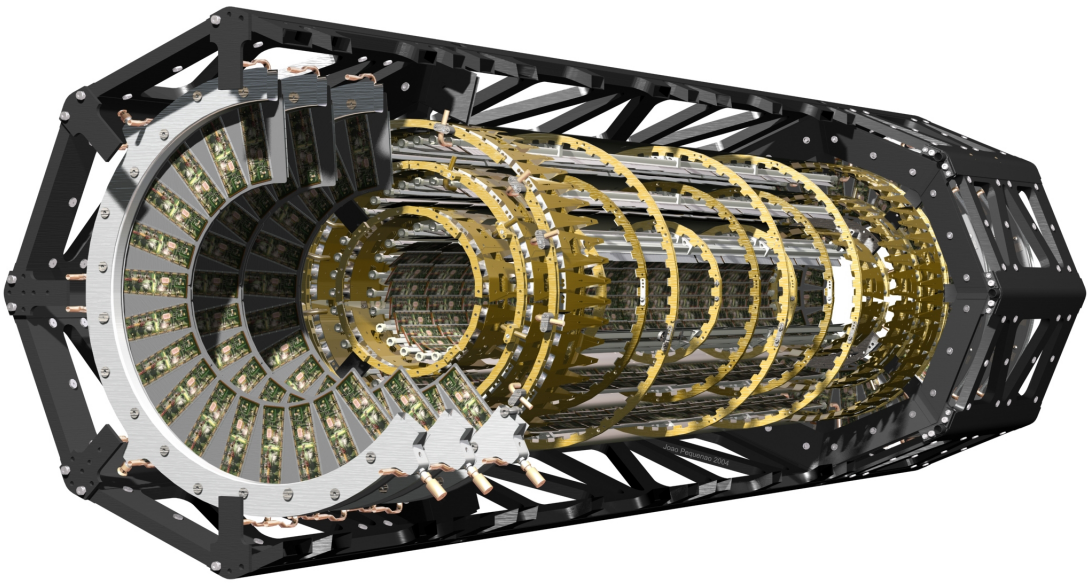


Figure 2.6: *Picture of the Pixel detector and its supporting structure [27]. One can see the stave structure in the barrel region and the disk arrangement containing the Pixel detector modules.*

be reached in the barrel region of the Pixel detector is 10 µm in R-ϕ and 115 µm in beam direction.

2.2.4 Forward Detectors

The measurement of the luminosity is an important issue in a collider experiment. The absolute luminosity is defined as [1]:

$$L = \frac{nfN_1N_2}{A} \tag{2.1}$$

with:
n – number of bunches,
f – colliding frequency,
N_x– number of particles per bunch,
A – interacting area.

The uncertainty of this calculation is about 3% due to the precision of the measured quantities. A more precise measurement of the integrated luminosity can be reached by reconstructing processes with known cross-sections.

$$\mathcal{L} = \frac{N}{\sigma} \tag{2.2}$$

with:
N – number of events,
σ – cross-section.

Therefore, the ATLAS detector measures rates of W $\rightarrow l\nu$ and Z $\rightarrow ll$ production just as exclusive muon pair production via two photon interaction.

In addition to the ATLAS detector, three smaller sets of detectors with good coverage in the very forward region are installed to provide luminosity information. By measuring inelastic p-p scattering and small-angle elastic scattering from Coulomb interaction in the forward direction, a precise luminosity measurement is reached.

LUCID The LUminosity measurement using Cerenkov Integrating Detector (LUCID), placed ±17 m from the interaction point, detects the inelastic p-p scattering in the forward direction. It measures the relative luminosity for the ATLAS experiment with an uncertainty of about 5% and provides online luminosity monitoring.

ALFA At a position of ±240 m, the Absolute Luminosity For ATLAS (ALFA) detectors do a more precise measurement of the luminosity. They use scintillating-fiber trackers inside Roman pots to measure small-angle elastic scattering coming from Coulomb interactions. The optical theorem links the total interaction rate to the imaginary part of the forward elastic scattering amplitude [30].

$$\sigma_{tot} = 4\pi \times Im[f_{el}(0)] \tag{2.3}$$

with:
σ_{tot} – total cross-section,
$f_{el}(0)$ – forward elastic scattering amplitude, extrapolated to zero momentum transfer squared.

For the measurement of Coulomb interaction at extremely small scattering angles (3 µrad), special beam conditions are necessary as the detectors approach 1 nm to the beam which is smaller than the nominal beam divergence. The measured luminosity is taken to calibrate the measurement done by the LUCID detector.

ZDC To determine the centrality of heavy-ion collisions and measure the luminosity, the Zero-Degree Calorimeter (ZDC) is located in a distance of ±140 m from the interaction point just after the split of the beam pipe. Alternating quartz rods and tungsten plates measure neutral particles at high $\mid \eta \mid > 8.2$. The ZDC detector enhances the acceptance of ATLAS forward and central detectors for diffractive processes and provides an additional minimum bias trigger for the ATLAS experiment.

2.2.5 Readout, Data Acquisition and Detector Control

During the detector operation, several systems for detector control and data handling are necessary. To handle the huge amount of data, a trigger system has to slim the number of events, recorded at ATLAS. The system reduces the event rate from 40 MHz to approximately 200 Hz in three steps.

- The level one (LVL1) trigger receives information about high transverse momentum muons from the muon trigger system and electromagnetic and hadronic interacting particles including $E_{\mathrm{T}}^{\mathrm{miss}}$ from the calorimeter system with reduced granularity. To model the bandwidths of passing events in the various trigger channels, pre-scaling factors are defined offline in the trigger menu of the central trigger processor.
- The second level (LVL2) has the whole event data available and is looking into regions of interests, defined by the LVL1 trigger. In addition to inclusive criteria, for example event and object properties over a certain threshold, the LVL2 trigger is sensitive to B-hadron decays. This requires the reconstruction of exclusive decays into low momentum particles.
- A last decision is made by the event filter after a rough offline reconstruction of the complete event. The amount of data is reduced by the trigger system to about 200 MB/s.

The Trigger and Data AcQuisition (TDAQ) system, synchronized to the timing of the trigger system, provides data pipe-line structures and transfers the data to the Read Out Driver (ROD) with respect to the LVL1 trigger decision. From the ROD, the digitized raw data is transferred to the LVL2 trigger. Processed by the event-building system and passing the event filter trigger, the data is written on a permanent storage. The size of a reconstructed ATLAS event is around 1.3 MB.

A safe detector operation is garanteed by the Detector Control System (DCS). It controls, monitors and archives the operation parameters of the ATLAS system and allows automatic and manual actions in case of an abnormal behavior. With a bi-directional communication to the DAQ, the definition of the detector state is synchronized to the offline data.

2.3 LHC Computing Grid

During ATLAS operation, the amount of data is about 2 PB per year. In addition, the simulation of data for physical analyses takes about 15 PB disk space per year. To process and store these huge amounts of data, a world wide LHC computing grid (GRID) is set up for all four experiments. In a hierarchical order, with CERN as

Tier-0, participating institutes provide their computing power in national Tier-1 and regional Tier-2 centers. This distributed computing system, with well defined tasks for the Tier centers, provides the basis of a common computing facility for ATLAS and the ATLAS users.

Chapter 3

Event Reconstruction and Modeling

The data, collected with the ATLAS detector, has to be translated into physical objects to provide a picture of the physics in the event. Various processes initiated by the proton-proton collision lead to the production of different particles emerging from the center of the ATLAS detector. The electrical signals generated by the interactions of these particles with the active detector material are collected, digitized and read out. It is important to deduce the initial physical objects from the signals recorded by the detector.

The second ingredient to measure the predicted properties of the Standard Model, and to search for physical processes beyond, is a precise theoretical description and modeling of the processes generated in the proton collisions.

3.1 Event Reconstruction

In the proton collisions, particles like photons, electrons, muons, hadrons and neutrinos are produced, traversing the detector. The aim of the ATLAS detector is to identify these particles and measure their properties. Figure 3.1 shows the "signatures" of different particle types in the detector sub-systems, introduced in Chapter 2.2.

- Electrons generate hits in the inner detector and are absorbed in the EM calorimeter.
- Photons, being uncharged, appear in the electromagnetic calorimeter without a track in the tracking system.

- Hadrons pass the electromagnetic calorimeter and deposit the remaining energy in the hadronic calorimeter. Additionally, in the case of charged hadrons hits are generated in the tracking system.
- Muons leave signals in all detector systems, including the muon-tracking system outside of the calorimeter, as only a fraction of their energy is deposited in the calorimeter.
- The weakly interacting neutrino leaves the detector without any signal. A nearly full coverage of the detector allows a reconstruction of the missing part of the transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$, assuming momentum conservation in the transverse plane.

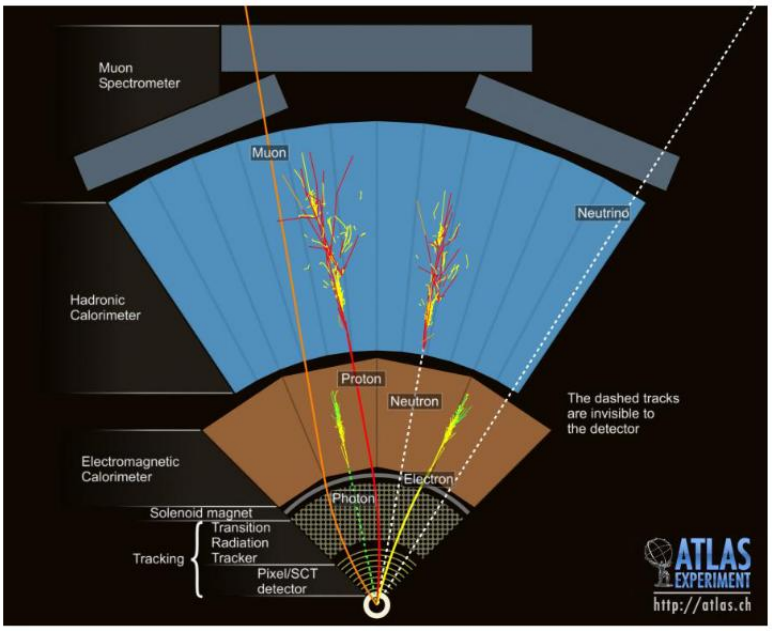


Figure 3.1: *Sketch of the detector response induced by different types of particles in the ATLAS detector [27].*

This section provides an introduction of the reconstruction algorithms included in the ATLAS reconstruction framework. The goal of these algorithms is to identify the type of the particles and to precisely reconstruct their properties, such as energy, momenta, charge and direction of flight. A more detailed description can be found in [28].

3.1.1 Track and Vertex Reconstruction

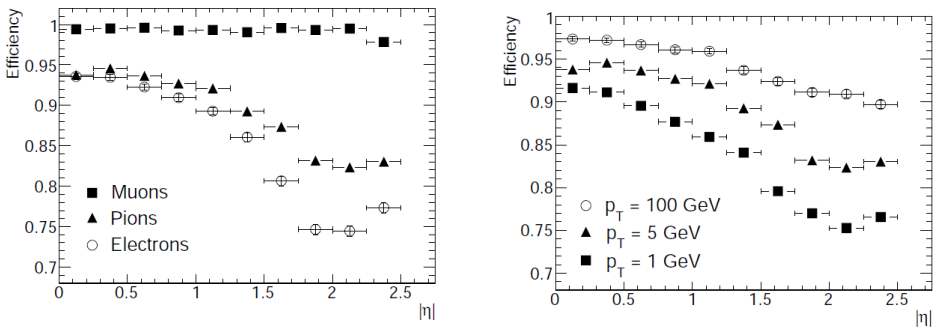
The space points, detected in the inner detector, has to be converted into particle tracks. The generator of these tracks is investigated by the association of the tracks

to their position of origin. The developed reconstruction algorithm are discussed in this section.

Track reconstruction

The tracking algorithms reconstruct the tracks of charged particles in the Inner Detector with a transverse momentum of $p_{\mathrm{T}} > 0.4$ GeV inside a pseudo-rapidity region of $\mid \eta \mid < 2.5$. All hits detected with the Pixel and SCT detector are transformed into three dimensional space points in a pre-processing stage. A track finding procedure, starting from the high granularity of the Pixel detector, detects track-seeds formed by combined space points. These seeds are extrapolated into the SCT to build track candidates. The candidates are fitted, outlying clusters removed and ambiguities in the cluster-to-track association solved. After the extrapolation into the TRT, the final track is refitted with the full information. An accurate description of the active and passive material is important in order to allow for material corrections.

Unused track segments of the TRT are extrapolated backwards into the SCT and the Pixel detector to increase the reconstruction efficiency of secondary tracks originating from photon conversions or long-lived particle decays. Figure 3.2 shows the reconstruction efficiency as a function of $\mid \eta \mid$ for different particle types and momenta. Inefficiencies of electron and pion tracks with increasing values of $\mid \eta \mid$ reflect the amount of material in the Inner Detector [31].



(a) (b)

Figure 3.2: *Track reconstruction efficiency as a function of* $\mid \eta \mid$ *[28]. Figure (a) shows muon, pion and electron tracks. In Figure (b), pions with different track momenta are shown.*

Vertex Reconstruction

Vertex reconstruction tools, described in detail in [32], are important to assign tracks to a common point of origin. The primary vertex is defined as the vertex of the hard interaction in the proton-proton collision. Additional less energetic collisions of the proton bunches, so called pile-up, can produce additional vertices. The reconstruction of the primary vertex in presence of pile-up is as crucial as the reconstruction of secondary-vertices, for example from heavy-flavor decays, photon conversions or decays of long lived hadrons. Figure 3.3 shows a sketch of the different vertex candidates:

- The primary vertex reconstruction algorithm calculates a vertex position, including all tracks of the event into a vertex fit. Tracks which are considered to be outliers are used to create iterative pile-up vertex candidates.
- Vertices in hadron showers, such as B-hadron decays, are reconstructed with the tracks inside the hadron shower region, which is defined by the hadronic-calorimeter response. Since the tracks of secondary vertices are not expected to arise from the primary vertex, the selection of tracks can be reduced requiring a track impact-parameter, which is defined as the closest approach of the the trajectory to the reconstructed primary vertex. The kinematic properties of the B-hadron decay, like the B $\rightarrow$ C-hadron decay chain, can be used as an additional constraint to reduce the mis-identification rate.
- Inside the Inner Detector, about 40% of all photons convert into e^+e^- pairs, caused by the material contingent of the inner detector. The vertices of photon conversions, marked as (γ) in Figure 3.3, are expected to be reconstructed in positions close to detector material, such as detector layers and mechanical and electrical supports. Whereas, th reconstruction of the decay of long lived hadrons (V^0) is a more challenging task, since there is no emphasized detector region defined. The decay of long lived hadrons is randomly distributed in the detector volume.

Constrained vertex fitting is CPU-time consuming, therefore a careful pre-selection of tracks and track pair candidates, based on the vertex type, is necessary. One of the tasks was the tuning of the pre-selection in the reconstruction of photon conversions and neutral-particle decays. These vertices appear in the whole detector space, which leads to a high amount of two-track vertex candidates in the event passed to the vertex fit.

Electron identification capabilities provided by the TRT detector and geometrical constructions of the two-track vertex candidates, such as the opening angle or a pre-calculated invariant two-track mass, can be used to increase the purity of two-track candidates. The applied modifications lead to a significant improvement of the required computing time of the vertex reconstruction algorithms.

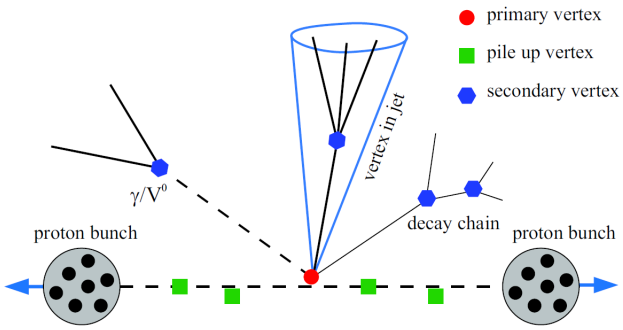


Figure 3.3: *Expected vertex topologies in the proton collision. Depicted are the primary vertex, pile-up vertices, vertices from conversions or long-living particles, B-hadron decay vertices in jets and decay chains [32].*

3.1.2 Inner Detector Survey

Important measurements in the early operation phase of the ATLAS detector are the parametrization of the detector-module positions with respect to each other and the implemented material budget. The position of the modules, defined by the detector construction plan, can be misaligned by distortions of single modules, detector layers or complete sub-detectors. The assignment of the hits to a track depends on the alignment of the tracking detector. Furthermore, the knowledge of the material budget implemented in the Inner Detector is important to parametrize the influence of multi-scattering correctly. Both effects will have an influence on the track reconstruction efficiency and momentum resolution [28].

Detector Alignment

The first step in aligning the tracking-system geometry was the minimization of track residuals, using high-momenta tracks of the cosmic-ray muons passing the Inner Detector. These tracks were collected in operation phases without beam presence. As these tracks were mostly vertical, they improve the alignment in the barrel region mainly. Horizontal tracks were generated by particles coming from beam-gas collisions during beam presence. They were used to improve the end-cap alignment.

To reach the required precision of 10 µm on the silicon-module position [33], tracks from the interaction point, generated in proton-proton collisions, have to be considered as well. Systematic correlations between the misalignment of different detector elements can be measured using well known physics processes. The processes $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$, for example, constrain the alignment constructing the invariant dilepton mass or the rate of positive and negative tracks as a function of p_{T} .

Detector Material

Material in front of the calorimeter, such as detector components, mechanical support structures and electrical services, leads to multiple scattering of particles and photon conversions. The amount of converting photons at a certain position, reconstructed in the Inner Detector volume, allows the visualization of the material budget. To validate the parametrization of the material budget in the ATLAS software, the decay of particles with well known mass, such as the K_s^0 decaying into two charged pions, is measured.

One of the focusses in course of my research was the study of K_s^0 decay vertices in data, collected at a center of mass energy of $\sqrt{s} = 900$ GeV in December 2009. Within these so called *minimum bias* events only $soft$ interactions take place where only a small momentum is transferred. The event selection and the track reconstruction used for the kinematics constrained vertex fit is described in [34]. In Figure 3.4, distributions of collected track and two-track V^0 vertex candidates generated by a Pythia [35] minimum-bias simulation are compared to data collected in the proton collisions. The simulated distribution of the track p_T 3.4(a) agrees well with the recorded to data. The dip at 500 MeV is induced by the transition from a special low-p_T into the standard-tracking algorithm. The track η distribution 3.4(b) confirms a good description by the Pythia simulation. The distribution of the radial position of the V^0 vertex candidates, shown in Figure 3.4(c), can be explained with the expected exponential decrease of V^0 decays as a function of the distance in combination with variations in the vertex reconstruction efficiency. At a radial position of 34.3 mm for example, where the LHC beam-pipe starts, the reconstruction efficiency of V^0 vertices is reduced due to multiple-scattering effects. In the region of 50.5 mm and 88.5 mm the Pixel b-layer and the second Pixel barrel layer is placed. For a particle decay in a short distance to a detector layer the reconstruction efficiency is increased. In addition, the rate of V^0 vertices is enhanced due to the misidentification of photon conversions at the position of the detector layers. The observed radial position and the transverse momenta of the vertex candidates, shown in 3.4(d), are well described by the simulation.

In Figure 3.5, the Podalanski-Armenteros plot [36] is shown. It is based on a kinematic construction. The transverse vertex momentum is shown on the y-axis. The relative difference of the longitudinal momenta of the two decay products is shown on the x-axis. Therefore, a K_s^0 decay vertex with the decay to two pions is symmetrically arranged on the large ellipse. Decays of Λ baryons ($\Lambda \to p\pi^-$) have larger relative differences in the longitudinal track momenta and reduced transverse vertex p_T . In the Podalanski-Armenteros plot they are placed in the small ellipses on the bottom edges. Photons converted into two electrons have a vertex without transverse momenta, therefore they are placed at the bottom line.

Differences in the parametrization of the material budget in a certain detector region would lead to a difference in the modeling of the pion momenta compared to data.

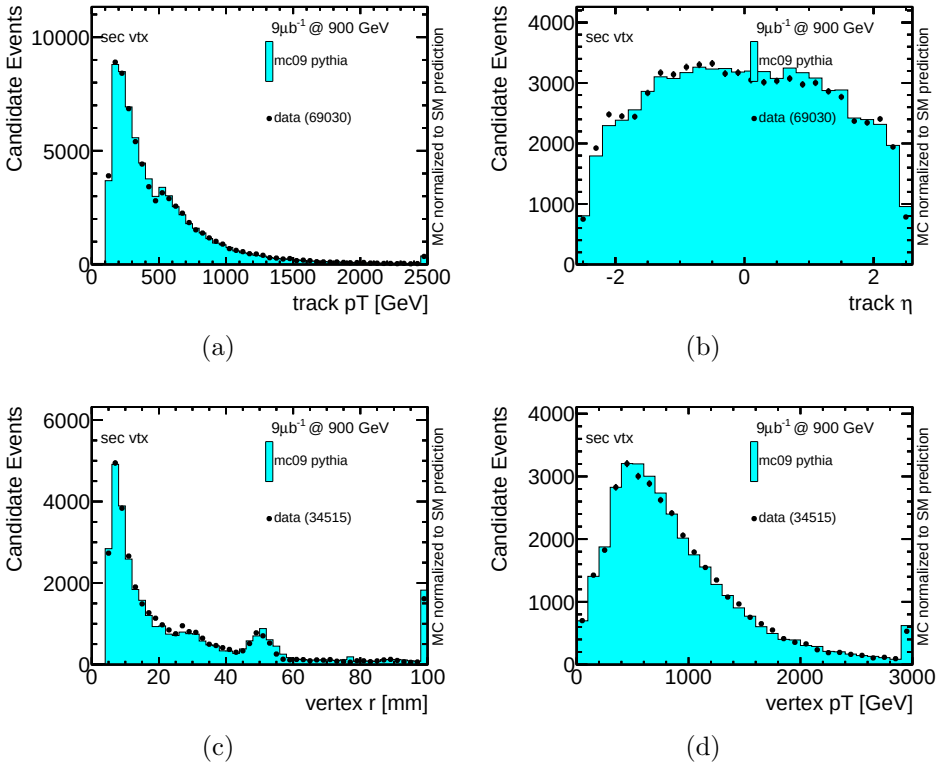


Figure 3.4: *Comparison of the reconstructed secondary vertex candidates selected in* $\sim 9\ \mu b^{-1}$ *minimum-bias data to a* PYTHIA *minimum-bias event simulation. The* p_T *and* η *distributions of the selected track candidates, the vertex momenta and the radial vertex position are depicted.*

Additional material, not parametrized in the detector volume, leads to a reduced reconstructed track p_T , since the energy loss is underestimated in the reconstruction algorithm. An enhancement of multiple scattering broadens the track p_T resolution. Both effects are indirectly visible in the invariant K^0_s mass distribution. The pions, originating from the K^0_s decay vertex, have an average momentum of ∼400 MeV. They are strongly bent in the magnetic field which leads to an increased freedom in the track fit. Hence, misalignment effects of the detector geometry are expected to be small and dominated by the effects of multiple scattering.

Figure 3.6 shows the invariant vertex mass, based on the assumption that both tracks originate from a pion. A clear peak at the K^0_s mass is visible. The distribution of background vertices is almost flat in the K^0_s mass region. An observed disagreement of the K^0_s rate comparing data to the PYTHIA simulation, which is without influence in the presented study, is not further investigated.

Figure 3.7 shows the reconstructed K^0_s mass, estimated by a Gaussian fit of the mass spectra in the signal region, relative to the nominal mass of (497.614 ± 0.024) MeV [1] as a function of the decay radius and the vertex position in η. A difference of 10%

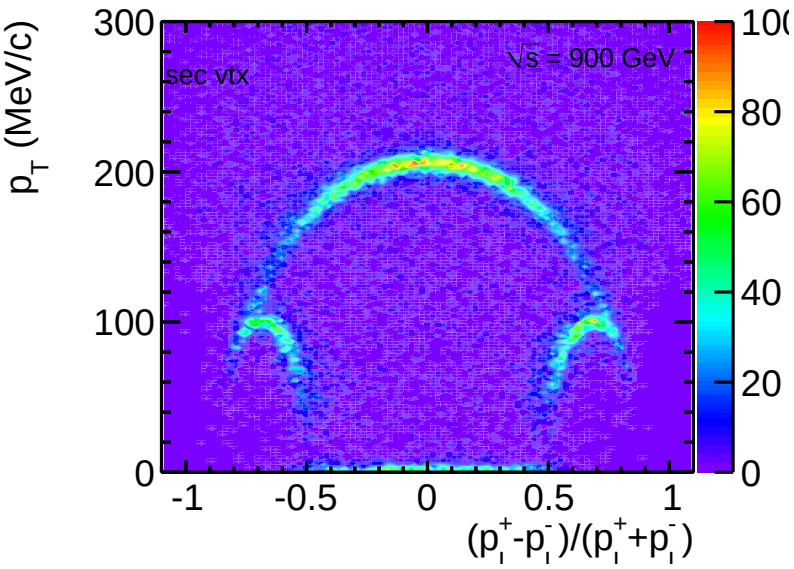


Figure 3.5: *The Podalanski-Armenteros plot shows the dependence of the transverse vertex momentum and the longitudinal track momenta. It has the power to separate* Λ *and* K^0_s *decays. Depicted are vertex candidates reconstructed in* $\sim 9\ \mu b^{-1}$ *collision events with* $\sqrt{s} = 900$ GeV. *The fraction of* K^0_s *decay vertices is arranged on the large ellipse while the fraction of* Λ *decays is steeled in the small ellipses on the bottom left (*$\bar{\Lambda}$*-decay) and bottom right (*Λ*-decay) corner. Converted photons appear at the bottom of the plot.*

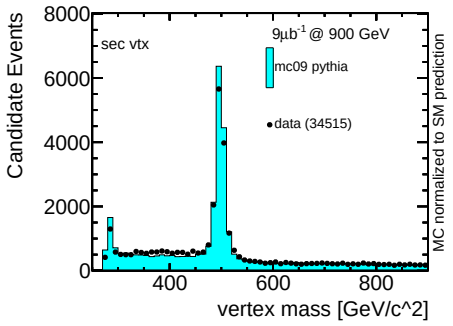
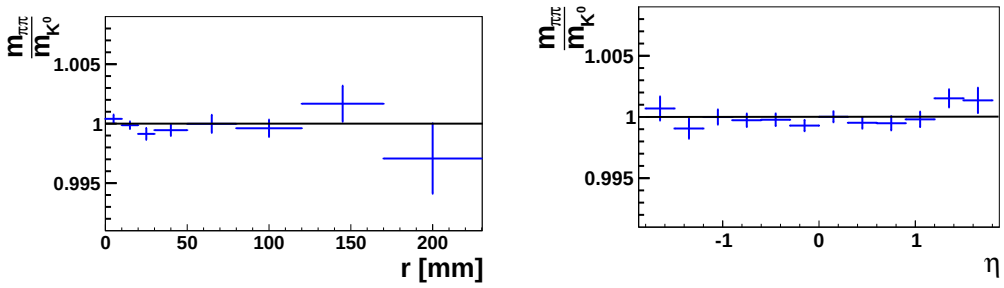


Figure 3.6: *Comparison of the reconstructed invariant vertex mass, based on a pion assumption, selected in* $\sim 9\ \mu b^{-1}$ *minimum-bias data to a* Pythia *minimum-bias event simulation.*

in the parametrization of the material in the inner detector would lead to a shift of 2% in the reconstructed K^0_s mass [37]. In accordance with the measurements shown in [37], based on a different vertex reconstruction algorithm, this measurement shows a good agreement of the offline material parametrization compared to the built-in detector material.



(a) (b)

Figure 3.7: *Deviation from the K^0_s mass as a function of the decay position inside the tracking volume. Figure (a) shows the radial distribution, Figure (b) the η dependence. The results are based on $\sim 9\ \mu b^{-1}$ data at a center of mass energy of $\sqrt{s} = 900$ GeV.*

3.1.3 Electron Reconstruction and Identification

The electron is detected as a charged particle track in the Inner Detector tracking system. It is supposed to be stopped in the electromagnetic calorimeter, radiating photons via Bremsstrahlung which then convert to electron-positron pairs. Similar electromagnetic showers can be produced by photons and hadrons as well. An accurate identification of electrons is a difficult task in the LHC environment with a ratio of electrons to jets in the order of 10^{-5}, considering a transverse momentum of 40 GeV. In addition, the material contingent in front of the calorimeter leads to electron energy losses and a large fraction of secondary electrons produced by converting photons.

Electron Reconstruction

The reconstruction of electrons based on different approaches is in detail described in [38]. The standard method of electron reconstruction starts with a shower seed, established by a sliding-window algorithm. With a window size of 5×5 cells, the energy deposit in the middle layer of the electromagnetic calorimeter is analyzed. A cluster with the fixed size of 3×7 cells (0.025×0.025 in units of $\eta \times \phi$ space) is reconstructed around this seed. These choices optimize the collection of all shower energy and simultaneously minimizes the contribution of noise and pile-up. The requirement of an associated track to the EM seed cluster, not associated to a photon conversion, separates electrons and photons reasonably.

A second choice of reconstruction algorithms starts with a good-quality track and constructs a cluster around the extrapolated position of impact into the EM calorimeter. This algorithm relies more on the electron identification capabilities of the Inner Detector and has been developed to improve the reconstruction of low p_T electrons.

The cluster energy itself is calculated as the sum of the expected energy deposited in the material in front of the calorimeter, the measured energy in the cluster, the estimated external deposit outside the cluster and the estimated energy deposited beyond the EM calorimeter.

The four momenta of a central electron in the region $| \eta | < 2.5$ is computed from track and cluster information. The energy of the electron is taken from the EM calorimeter while η and ϕ are calculated using the precise tracking information.

Cut Based Optimization of the Electron Identification

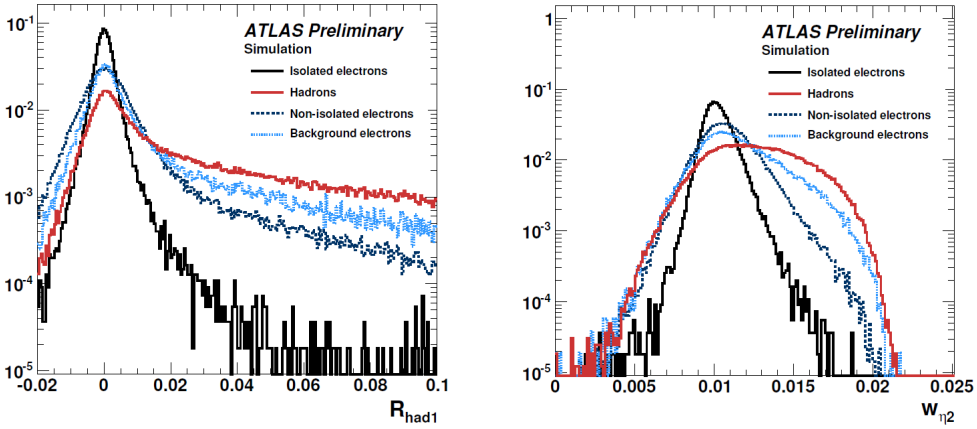
Hadronically interacting particles, such as protons or charged pions, can fake electromagnetic objects. To provide an additional separation, Shower-shape variables[1)] take advantage of the fine granularity of the EM calorimeter and combined reconstructed properties [39], such as the ratio of energy in the calorimeter to the momentum in the Inner Detector.

- Electrons and photons produce a narrow shower with a core of high energy density surrounded by a halo of small activity. In comparison, the deposition of energy in the EM calorimeter by hadron showers is much broader.
- The ratio of energy in the EM calorimeter divided by the momentum of the inner detector track is higher for electrons than for protons.
- The fraction of neutral pions decaying into two photons and, therefore, faking an electron candidate can be reduced by the requirement of one center of energy in the calorimeter shower.
- Hadrons, like protons and neutrons, only deposit a small amount of their energy in the EM calorimeter before they are stopped in the hadronic calorimeter. The ratio of electromagnetic energy to hadronic energy deposition, called "hadronic leakage" provides a good separation.

ATLAS has defined several working points in the electron selection, ordered by the power of rejection as "loose", "medium" and "tight" [38] with a background rejection factor of about 500, 5000 and 50000.

The "loose" electron object selection, reconstructed in a region $| \eta | < 2.47$ (excluding the crack region $1.37 < | \eta | < 1.52$ between barrel and forward calorimeter), is based on the shower shape and the hadronic leakage. An example of observables, used for the separation of fake electrons, is shown in Figure 3.8. The quality parameters are defined as:

[1)] lateral and longitudinal shower profiles, etc.



(a) (b)

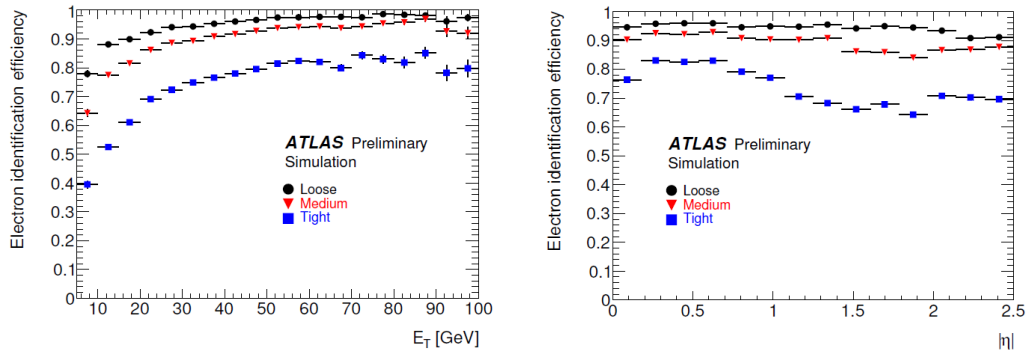
Figure 3.8: *Examples of observables to separate signal electrons and fake candidates. Figure (a) shows the ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster. In Figure (b), the shower width in the second compartment of the electromagnetic calorimeter for isolated electrons and the main backgrounds is shown [38].*

- Ratio in η of cell energies in 3×7 versus 7×7 cells
- Lateral width of the EM shower ($W_{\eta 2}$, shwon in Figure 3.8(b))
- Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (R_{had1}, shown in Figure 3.8(a))
- Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster in the barrel-forward crack region of the detector ($| \eta | > 0.8$ and $| \eta | < 1.37$)

The "medium" definition rejects hadrons by additional requirements on the shower width and the ratio of the largest and second largest energy deposit in the EM calorimeter over the sum of energy deposit. The track requirements are tightened, including a requirement of ≥ 1 hit in the Pixel detector and ≥ 7 hits in the SCT detector. The demand of a transverse impact parameter < 5 mm and an Inner Detector track, matching the second EM cluster layer within $\Delta\eta < 0.01$, further reduces the number of fake electrons.

Object criteria, defined as "tight" further reject charged hadrons and converted photons. Here the $\Delta\phi$ between the cluster and the track is required to be less than 0.01, the track pointing to the cluster has to match within $\Delta\eta < 0.005$ and the requirement on the high cluster energy over track momentum E/p ratio is tightened. In addition, one b-layer hit in the Pixel detector and a transverse impact parameter < 1 mm of the track are required. The difference between expected and measured numbers of TRT hits has to be smaller than 15 and electron candidates matched to photon conversions are rejected.

Figure 3.9 shows the efficiency of the three electron object definitions as a function of E_T and $|\eta|$. The further optimization of the fake rejection without a big loss in efficiency is done by a special tuning, described in [40], providing "loose++", "medium++" and "tight++" electron objects [39].



(a) (b)

Figure 3.9: *Electron reconstruction efficiency for different object definition in dependence on E_T in Figure (a) and $|\eta|$ in Figure (b) [38].*

3.1.4 Muon Reconstruction

The reconstruction of muons takes advantage of the multiple sub-detector configuration, which provides a standalone tracking system in the muon chambers combined with the track information given by the Inner Detector.

Muons are detected in a momentum range of approximately 3 GeV up to 3 TeV. For high momentum tracks with $p_{\mathrm{T}} > 100$ GeV, the standalone reconstruction provides a high precision. For low momenta tracks, the matching of the track in the muon system with the Inner Detector track improves in the muon reconstruction. Therefore, the track is propagated to the interaction point considering corrections on the expected energy loss in the calorimeter. In addition, the combination of the two tracking systems provides an important improvement in terms of background rejection. Muons, coming from pion or kaon decays, will be suppressed by the requirement of hits in the first tracking layers of the Inner Detector.

The object quality is defined as "loose", "medium" and "tight", as it is for electrons [41]. A "loose" muon is reconstructed only in the muon chambers. To fulfill the "medium" definition, the standalone track is required to be in a range of $|\eta| < 0.2$ or has to have a track extension into the Inner Detector. The "tight" muon candidate, used in this analysis, is required to have a matched track in the Inner Detector.

In addition to the object requirements, criteria on the track are defined in the top-quark analysis [68]. The muon track is required to have one hit in the b-layer, at least 6 SCT hits and less than 3 missed hits in the Pixel and SCT detectors. The number of hits in the TRT is demanded to be greater than 5 and the ratio of hits in the TRT divided by the number of all hits is required to be less than 0.9. The impact parameter z_0 component has to be smaller than 2 mm.

3.1.5 Jet Reconstruction

A color charged object, originating from the hard interaction, initial- and final-state radiation or the underlying event, will produce a shower, composed of a bundle of hadrons produced in the fragmentation process. The reconstruction of these hadron showers, visible as jets in the calorimeter, plays an important role in many physical processes at the LHC. The fine granularity of the of ATLAS hadronic calorimeter allows for a precise jet reconstruction.

However, the jet reconstruction is influenced by several effects. Background activity in the calorimeter, generated by pile-up events and calorimeter-noise, will influence the measurement. Low p_{T} particles of the hadron shower can be lost due to deflection of the magnetic field in the inner-detector or material in front of the calorimeter. The hadron shower itself contains hadronic and electromagnetic fractions, which differ in their calorimeter response. A hadronic signal in a non-compensating calorimeter leads to low signal densities in the calorimeter cells in comparison to an electromagnetic signal. Therefore, a compensation weight of the order of the electron-pion signal ratio e/π has to be applied. Since the fractions are unknown, this leads to an uncertainty in the energy measurement. Furthermore, due to nuclear breakup, a fraction of the hadronic component is not included in the visible energy constructed from the calorimeter cell activity. All these effects have to be taken into account to make a precise statement on the properties of the initial parton. The ATLAS jet reconstruction, described in detail in [43], consists of several steps, specified in the next sections.

Calorimeter Cell Clustering

Hadrons, entering the calorimeter, deposit their energy in various calorimeter cells. These cells are grouped together to signal objects, referring to the energy deposition of the incoming hadron. In the ATLAS reconstruction framework, there are two algorithms implemented to cluster the calorimeter cells, based on two-dimensional and three-dimensional cluster objects [**?**].

Calorimeter Towers Calorimeter towers are two-dimensional objects, summing up calorimeter cells in a fixed grid of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. The longitudinal

separation of the calorimeter is not used in this approach. By building the sum of the cell energy in a fixed area, no suppression of noise is applied and the correction and calibration of the cluster energy is complex.

Topological Calorimeter Cluster Topological calorimeter cluster are three-dimensional "energy blobs" taking in account the individual shower development of each particle entering the calorimeter. Hence, a better background suppression and energy calibration are reached. Starting from a seed cell with a significant amount of energy $\mid E \mid_{cell} > 4\sigma_{cell}$, neighbouring cells are collected. If one of these cells has an energy $\mid E \mid_{cell} > 2\sigma_{cell}$, it is handled as second seed, including its nearest neighbours as well. With this approach, a ring of guard cells is added to each topological cluster. If more than one local maximum is detected, the topological cluster is split into multiple cluster objects.

Jet Reconstruction Algorithm

Jet algorithms [43] are among the main tools used to analyze data from hadron collisions. They bundle the calorimeter clusters to jet objects, referring to the momenta and direction of flight of the initial parton. As the jet algorithms are able to handle reconstructed tracks and generated particles from MC as well, an interface is provided to compare calorimeter jet objects to jets formed by tracks of the Inner Detector (track jets) and jets formed by the truth particles of the simulation (truth jets).

The jet algorithm has to be physically well defined. Therefore it has to be infra-red (IR) and collinear (CL) save [43].

IR: soft particles, not coming from the fragmentation process of the hard scattered proton, should not affect the number of jets produced.

CL: a certain amount of transverse momenta carried by one particle or a disjunct two-collinear particle system should lead to the same reconstructed jet.

The anti-k_t [43] and the cone jet algorithm [43] are the two main approaches of jet reconstruction, implemented in the ATLAS framework. The choice of a specific algorithm depends on the specific analysis.

Cone Jet Algorithm In the cone jet algorithm, a geometrical approach is used to cluster objects, close-by in the $(\eta - \phi)$ plane. The algorithm starts with a seed cluster of $p_T > 1$ GeV. All energy contributions inside a cone of a predefined cone size R are summed up. The cone axis is placed iteratively by recollecting the clusters until a stable jet axis is found. Jets with $E_T < 7$ GeV are removed. In the case

of overlapping jet clusters, the energy in the overlap region is shared between both jets.

The cone jet algorithm is fast and provides plain jet shapes, shown on the bottom left of Figure 3.10. Cone shaped objects provide an easy application of jet energy calibrations. However, the collection of all energy contributions in the cone, partly generated by the underlying event, influences the energy measurement. In addition, supplementary energy of the parton shower outside the jet cone is not taken into account at all. With the default construction, the algorithm is not infra-red and collinear save.

k_T Jet-Clustering Algorithm The k_T jet-clustering algorithm uses a sequential recombining approach of the calorimeter cluster. Object pairs (particles, tracks or detector objects) are analysed with respect to their relative distance $d_{i,j}$, defined as:

$$d_{i,j} = min(k_{T,i}^{2p}, k_{\mathrm{T},j}^{2p}) \frac{(\Delta)_{ij}^2}{R^2} \tag{3.1}$$

with:

$k_{\mathrm{T},i}$	–	transverse momentum of cluster i,
$(\Delta)_{ij}^2$	–	distance of cluster i and j in the $\eta - \phi$ -plane and
R	–	scale parameter at which two jets are resolved from each other.

In case of the k_T jet-clustering algorithm the parameter p is one. As in the cone algorithm, a radius parameter R defines the maximum distance of the objects. Furthermore, the momentum of one object relative to the beam position $d_{i,B}$ is calculated as:

$$d_{i,B} = k_{\mathrm{T},i}^{2p} \tag{3.2}$$

If the distance between two objects $d_{i,j}$ is smaller than $d_{i,B}$, they are combined into a new object. Otherwise $d_{i,B}$ is called a jet and removed from the object list. With this approach, neighboring objects with low energy deposition are clustered before they are assigned to a jet. In the end, all calorimeter clusters are assigned to jets or are defined as jets themselves. Shared objects do not exist.

The extensions of the $\mathrm{k_t}$ algorithm consist in the definition of the distance. By varying the parameter p, the relative power of energy versus geometrical (Δ_{ij}) scales is shifted. With $p = 1$ the inclusive $\mathrm{k_t}$ algorithm is selected. Figure 3.10 shows the collection of jets reconstructed with four different jet algorithms. The Cambridge/Aachen algorithm is defined by $p = 0$ while for the anti-$\mathrm{k_t}$ algorithm $p = -1$ is set. The shape of the jets differs a lot, depending on the choice of the p parameter.

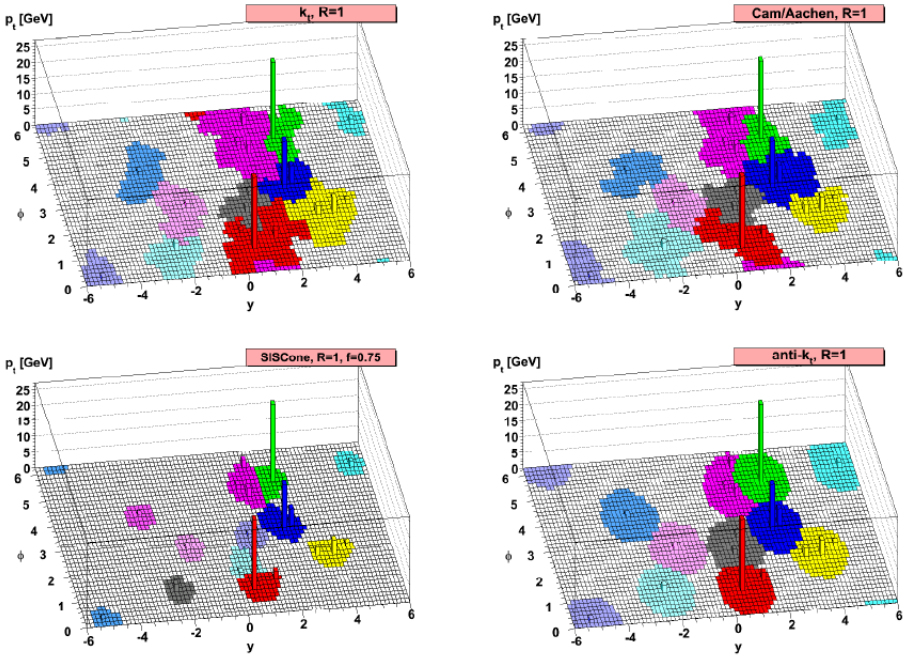


Figure 3.10: *Association of calorimeter cluster to jet candidates in a parton-level event with many random soft clusters shown for different jet reconstruction algorithms [44]. Presented are the* k_t *(top-left), Cambridge/Aachen (top-right), SISCone (bottom-left) and anti-*k_t *(bottom-right) algorithm.*

The anti-k_t jet-clustering algorithm performs best in terms of graduation of energy in case of overlapping jets. In addition, the anti-k_t jets have a more circular jet shape compared to the other algorithm. Soft particles, having a low transverse momentum, do not modify the shape of the anti-k_t jet. They will cluster with hard, high transverse momentum particles before they cluster among themselves. If a hard particle $k_{T,1}$ has no neighbor within a distance of $2R$, it will accumulate all soft particles within a circle of radius R, resulting in a conical jet.

If there is another hard particle $k_{T,2}$ within $R < \Delta_{12} < 2R$, the result is that there are two hard jets, both not perfectly conical. If $k_{T,1} >> k_{T,2}$ the jet, leading in p_T , will be partly conical as the overlapping part is missing. For $k_{T,1} = k_{T,2}$ the jets are equally divided.

For $\Delta_{12} < R$, the hard particles $k_{T,1}$ and $k_{T,2}$ are clustered together, forming a more complex shape consistent of the unification of two small cones around the hard particles and one cone with radius R, centered around the final jet.

In sum, the anti-k_t algorithm constructs jets which are similar to jets of an idealized cone algorithm. Jets with only soft fractions are absolutely conical. Since the anti-k_t algorithm is in good infra-red and collinear safe conditions, it became the standard algorithm in ATLAS.

Calorimeter Jet Calibration

The measured jet energy and the reconstructed jet position do not necessarily match the parton properties. As mentioned above, detector effects such as noise in the calorimeter, dead material in front of the calorimeter, particles bend by the magnetic field, crack regions of the calorimeter and the fraction of electro-magnetic content in the hadronic shower, have an influence on the reconstructed jet.

Global calibrations are applied to compensate these effects. The performance of the calibration in terms of linearity and resolution is determined in various processes, such as QCD di-jet events. In addition, corrections of jet algorithm effects and mis-calibrations are applied. More details on the calibration of calorimeter jets can be found in [43].

B Hadron Identification

The knowledge of the presence of a b-quark in a jet, labeled as b-jet, plays a big role in many physical analyses. The long lifetime of the b-quark in the order of pico seconds leads to a displacement of the B-hadron decay vertex of several millimeters in dependence on the B-hadron momentum. The B-hadron decays into lighter particles. Tracks of these particles stick out due to a non-zero impact-parameter, which is the closest approach of the trajectory to the reconstructed primary-vertex in the $R-\phi$ plane. The high resolution of the ATLAS tracking detector, with its innermost layer being placed only 50.5 mm away from the interaction point, allows one to resolve these vertices in the vicinity of the primary vertex.

Several algorithms, called b-tagging algorithms, are implemented in the ATLAS framework and described in [68]. The purpose of these algorithms is the separation of b-jets from light-flavour jets using a track-based, a vertex-based or a combined approach. For the analysis of the b-quark presence, all tracks selected within the jet are taken into account.

The IP3D b-tagging algorithm calculates the signed transverse-impact-parameter significance $\mathrm{d}_0/\sigma_{d_0}$ and the longitudinal-impact-parameter significance $\mathrm{z}_0/\sigma_{z_0}$ for each track, taking advantage of the correlation between the two variables. The sign of the impact parameter discriminates the tracks from B-hadron decays from tracks originating from the primary vertex, since the point of the B-hadron decay must lie along its direction of flight, defined by the jet axis.

The SV1 tagger uses a secondary vertex approach, starting with the search for two track pairs displaced from the primary vertex. To reduce the number of fake vertices, objects compatible to V^0 decays or photon conversions are rejected. A tagging discrimination weight is calculated using likelihood ratio techniques to combine the invariant mass of all tracks assigned to the vertex, the sum of energy of the vertex

assigned tracks with respect to the energy sum of all tracks, and the number of two-track vertices.

The JetFitter exploits the topology of the B- and the C-hadron decay inside the jet by assuming a common line between the primary vertex, the b- and c-hadron decay vertex. A discrimination of b-,c- and light jets is based on the likelihood, using the masses, momenta, flight length significances and track multiplicities of the reconstructed vertices as inputs.

Using a neural network, the MV1 algorithm determines a single discriminant combining the output weight of all three b-tagging algorithms (JetFitter, IP3D and SV1) and the η and p_{T} of the jet. Figure 3.11 shows the light jet rejection rate for different b-tagging algorithms as a function of tagging efficiency. The best performance is obtained by the MV1 b-tagging algorithm.

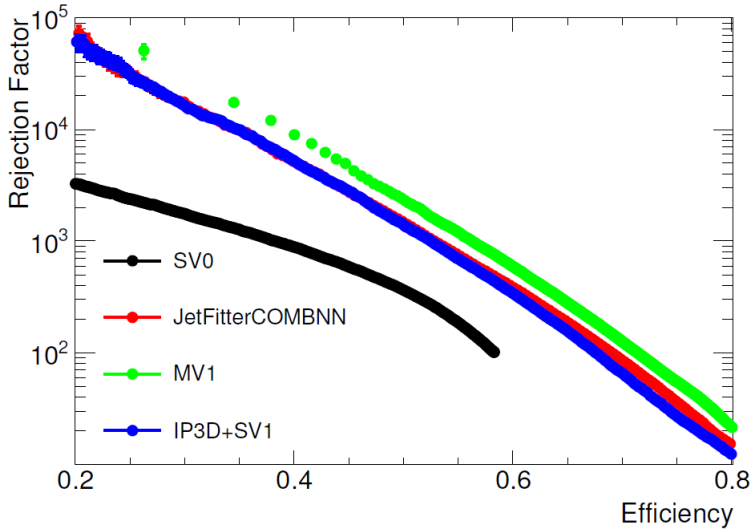


Figure 3.11: *Light jet rejection rate as a function of the tagging efficiency for different b-jet tagging algorithms. The MV1 tagger shows the best rejection power of light jets [38].*

3.1.6 Missing Transverse Energy

The neutrino is the only SM particle which escapes undetected, as it interacts only by the weak interaction with an extremely low interaction probability. The identification and the measurement of the particle properties therefore have to be done indirectly. As the momenta of the colliding partons point into the longitudinal direction, the conservation of momentum predicts the total transverse energy of an event to be zero. Losses in the total p_{T} balance, so-called $E_{\mathrm{T}}^{\mathrm{miss}}$, point to the presence of neutrinos.

In general this value can be computed by an E_{T} vector sum of all calorimeter cells plus the muon energy. With a good azimuthal and η coverage and a high spatial resolution, the ATLAS calorimeter precisely determines the neutrino direction of flight in the $R-\phi$ plane. Because energy losses of particles bias the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, their interactions with material, such as the cryostats between the Inner Detector and the EM and hadronic calorimeter barrel, are taken into account. To reduce the contribution of fake muons in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation, an extrapolation of the track, reconstructed by the muon-system into the Inner Detector tracking system, is required.

The construction of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the analysis described in this thesis, is performed object based. Equation 3.4 shows the exact calculation, based on jets calibrated to the electromagnetic scale [71]. The calorimeter cells are calibrated according to the associated object with the highest p_{T}. The order in Equation 3.4 refers to the order of the cells associated with the objects. SoftJets are defined as jets with a low transverse momentum (7 GeV $< p_{\mathrm{T}} <$ 20 GeV). Remaining cells are summed up in a CellOut therm, calibrated to the EM scale .

$$E_{\mathrm{T}}^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \tag{3.3}$$

where

$$E_{x,y}^{miss} = E_{x,y}^{RefElec} + E_{x,y}^{RefJet} + E_{x,y}^{RefSoftJet} + E_{x,y}^{RefMuon} + E_{x,y}^{CellOut} \tag{3.4}$$

The resolution of $E_{\mathrm{T}}^{\mathrm{miss}}$ in simulation, analyzed using a simulated semileptonic $t\bar{t}$ sample, is shown in Figure 3.12. By taking the truth information of the neutrinos into account, the resolution can be calculated with a gaussian fit on the $E_{x,y}^{miss} - E_{x,y}^{NonInt}$ residual distribution. As expected, the resolution of $E_{\mathrm{T}}^{\mathrm{miss}}$ depends on the scalar sum of the transverse energy of the cells ($\sum E_{\mathrm{T}}$) in the calorimeter.

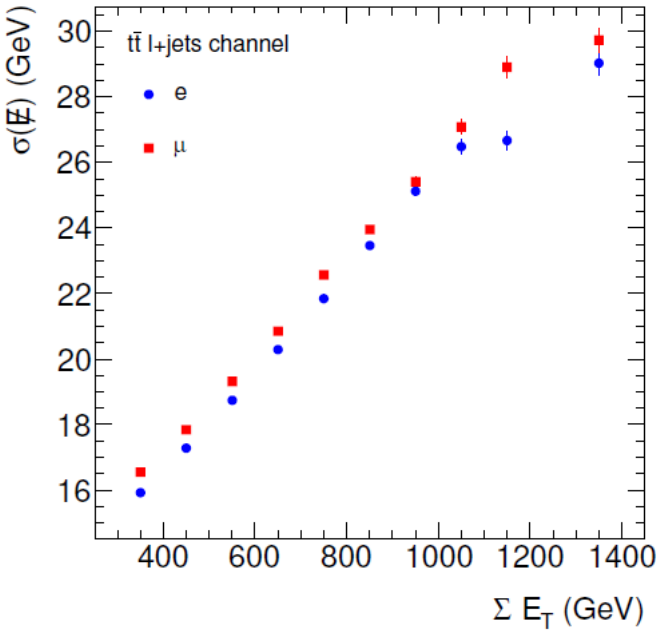


Figure 3.12: *Resolution of the reconstructed E_{T} in bins of $\sum E_{\mathrm{T}}$ for the semileptonic $t\bar{t}$ selection for the electron and muon channel [68].*

3.2 Modeling of Simulated Events

The interpretation of data, collected at high-energy particle colliders, requires a theoretical prediction and modeling of the collisions. Collision events recorded with the ATLAS detector have to be compared to hypothetical events, simulated according to the frequency and signatures of expected processes, to measure SM processes and observe phenomena beyond the SM.

The aim of the event modeling is not only to provide a complete picture of the proton interaction final state but also to describe the acceptance and response of the detector. The final-state particles, simulated according to our knowledge of particle physics, pass through a detector simulation which simulates interactions with the detector material and effects of the detector electronics to provide a realistic estimation of the detector response. In a final reconstruction step, the same reconstruction algorithms as used for collision data are applied, allowing a direct comparison of simulated events and observed data.

3.2.1 Event Generation

The event simulation describes high energy collisions and their final states, using numerical Monte Carlo (MC) techniques. The fundamental physical concept to separate independent phases of the entire collision, referring to their energy scales, is denoted by factorization.

Event Structure

Perturbation theory can be applied in the QCD+EW theory after the separation and decoupling of processes at different scales. With a momentum exchange of the order of $Q \gg \Lambda_{QCD}$, the hard process interacts at a short distance. Hence, it can be handled as interaction among fundamental, freely moving constituents. Long distance functions, such as the fragmentation functions, are decoupled and can be treated separately.

Based on this assumption, the event generation can be split into a PDF, a hard process, a shower evolution, the hadronization, the underlying event and higher-oder corrections. The general structure of a proton collision is shown in Figure 3.13. It is split into different sub-processes:

The Hard Interaction The hard interaction of two of the proton constituents ($qg \rightarrow qg$ in Figure 3.13), as well as the decay of the generated particles, can be described using matrix elements.

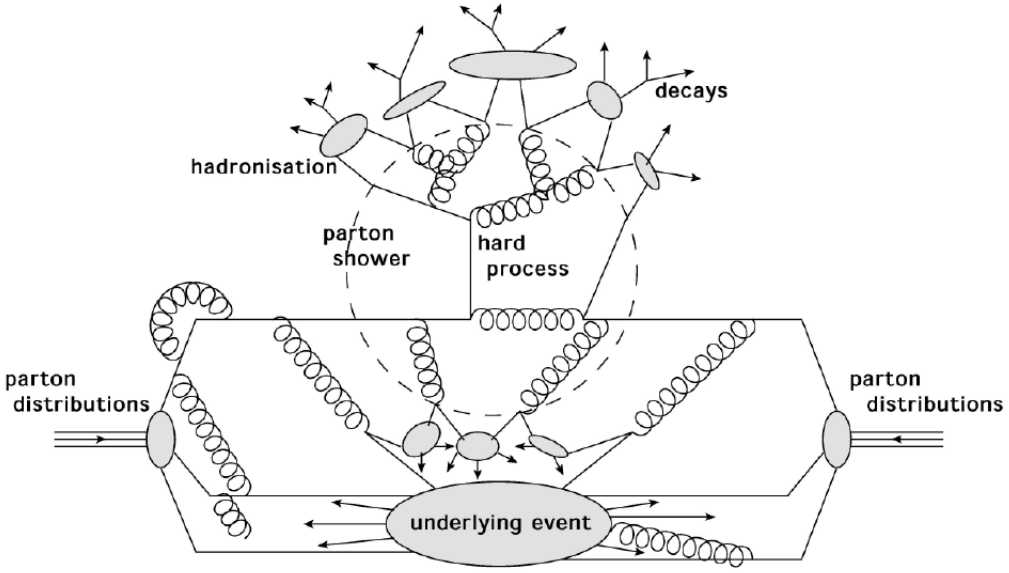


Figure 3.13: *Structure of a hard proton-proton collision, including a hard quark-gluon interaction, initial and final state radiation and contributions from the underlying event [46].*

Initial State Radiation (ISR) ISR is generated by emitted gluons (splitting into a $q\bar{q}$ pair in Figure 3.13) in the perturbation theory evolution of the initial-state. Therefore it depends on the energy scale Q, which is the momentum consumed by the hard interaction. It is not completely independent of the hard interaction as the overall color neutrality and conservation of the baryon number has to be fulfilled.

Final State Radiation (FSR) FSR describes the radiation of the remaining constituents of the hard interaction. Hence, it also depends on the scale Q of the process, responsible for the amount and energy of the radiated gluons.

Parton Shower The parton-shower is produced in the fragmentation process by strongly interacting partons radiating gluon bremsstrahlung and gluon splitting processes. In the generation of events this process continues until a certain threshold, defined by a virtuality scale, the transverse momentum or a radiation angle, is reached (represented by a dashed line in Figure 3.13).

Hadronization Hadronization is achieved when nearby partons merge into color-singlet clusters (depicted as gray blobs in Figure 3.13), which then decay into physical hadrons.

Underlying Event The underlying event models the additional activity beyond the hard process, for instance produced by additional color exchange between particles, not involved in the hard interaction.

To predict the rate of various partons, a Parton Distribution Function (PDF), measured in deep-inelastic scattering processes, describes the momentum fraction of the proton constituents. Figure 3.14 shows the parton distribution $xf(x, Q^2)$ correlated to the momentum fraction x for one set of PDFs and two different values of Q^2. The most relevant PDFs at the LHC are the ones from the gluons followed by the PDFs of the up and down valence quarks.

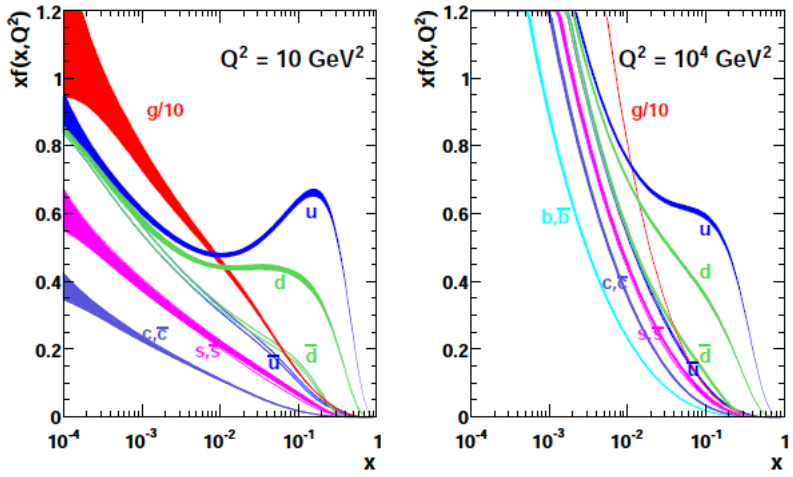


Figure 3.14: *MSTW2008 NNLO Parton Distribution Function at* $Q^2 = 10\ \mathrm{GeV}^2$ *and* $Q^2 = 10^4\ \mathrm{GeV}^2$*, including a one-sigma confidence level uncertainty band [47].*

Event Generators

The precision of the perturbative calculation depends on the order of perturbation theory taken into account. Leading order processes are based on calculations of the hard process without loop corrections, so called 'at tree-level'. More complex Next-to-Leading Order (NLO) calculations are necessary to take loop corrections into account.

The pool of event generators, described in detail in [48], includes multipurpose generators which handle all phases of the event simulation and specialized generators which cover only single steps of the event generation. Matrix element generators reach a good description of the hard process but need the application of showering and hadronization algorithms. The application of the different generator concepts depends on the physical process of interest. A physical analysis at hadron-colliders often obtains the most accurate theoretical prediction by combining components of different simulation programs.

The Monte Carlo (MC) simulation of a final state with a large number of jets is a challenging problem. In the generation parton-shower models in addition to the matrix-element calculation are included. Therefore, the radiation of a gluon in the initial- or final-state of the event can lead to the same n-parton configuration as an additional gluon radiation in the hard process. Hence, the problem of double counting n-parton-final-state configurations has to be covered by special "matching scheme" algorithms.

Matrix Element Event Generators

The parton-level Monte Carlo event generators simulate final states consisting of leptons, bare quarks and gluons. Reliable predictions can be provided using perturbative expansion because colliding partons can be considered as almost free particles at smallest scales.

AlpGEN AlpGEN is designed for the generation of SM processes in hadron collisions, using a leading-order evaluation [49]. It puts emphasis on large parton multiplicities in the final state and is used to generate W+jet and Z+jet samples in this analysis. The shower evolution and hadronization can be interfaced to Herwig [50] or Pythia [35].

MC@NLO MC@NLO includes the full next-to-leading order QCD corrections in the computation of the hard subprocess [51]. The next-to-leading order calculations include virtual and real emissions of partons. The result is a set of leading order and next-to-leading order parton configurations given to the parton-shower generator Herwig. The number of processes of the MC@NLO generator is limited, but includes $t\bar{t}$ production, which is of special interest in this thesis.

Multi Purpose Event Generators

Multi-purpose event generators are standalone tools to generate a whole picture of a collision event. They simulate matrix-element calculations as well as parton showers and hadronization processes which have the advantage of a consistent matching scheme implementation. By providing interfaces, it is possible to adopt their subcomponents, such as the parton-shower implementation, to a ME calculated by a different generator.

HERWIG Herwig is a general-purpose event generator for the simulation of lepton-lepton, lepton-hadron and hadron-hadron collisions [50]. It includes a large range of hard scattering processes for the Standard Model and supersymmetric extensions. All constituent parts of the event, such as initial- and final-state radiation,

an angular-ordered parton shower, hadronization and hadron decays, and underlying event simulation, are modeled. To simulate multiple parton interactions, HERWIG makes use of JIMMY [52].

For additional QCD radiation in the HERWIG parton shower, a coherent branching algorithm for both initial- and final-state particles is used. The full phase-space is restricted to an angular-ordered region in order to treat both leading soft and collinear singularities.

PYTHIA PYTHIA is, like HERWIG, a general-purpose event generator for hadronic events in pp , ep and e^+e^- collisions [35]. It contains around 240 different $2 \to n$ subprocesses, which are all at leading order calculation. The parton shower matches first "NLO" matrix elements for gluon emissions in $1 \to 2$ resonance decays. For hadronization and decays the Lund-String-Model [53] is used, based on a linear confinement, where color (anti-)triples, such as (anti-)quarks, are located at the end of the string, while gluons are energy and momentum carrying kinks.

SHERPA SHERPA consists of a tree-level matrix-element generator for the calculation of hard scattering processes within the Standard Model and various new physical models [54]. The emission of additional QCD partons coming from the initial- and final-states is described through a parton-shower model. To consistently combine multi-parton matrix elements with the QCD parton cascades, the approach of Catani, Krauss, Kuhn and Webber (CKKW) [55] is employed. It leads to a separation of matrix-elements and parton showers at a given jet resolution value y_{ini}, defined according to the resolution y_{cut} of the k_t -algorithm for the jet clustering. The fragmentation of partons into primary hadrons is described using a phenomenological cluster-hadronisation model.

3.2.2 Detector Simulation

The simulated events consist of particles with a certain energy and direction of flight. The ATLAS detector simulation translates these observables into a detector response, taking the interactions with the detector material into account.

The standard ATLAS detector simulation, described in detail in [56], is based on the GEANT4 [57] simulation tool-kit. It contains physical models to parametrize the particle transit through the detector material, including multiple scattering and energy deposition. After the simulation of the physical interactions, a digitization of the energy deposited in the sensitive detector regions, such as silicon modules, straw tubes chambers and scintillating tiles, into a detector response is applied. The final output of the detector simulation is identical to the digital signals of the ATLAS detector readout. To allow a most precise comparison of collected data

and simulated events, the detector simulation is validated and tuned with test-beam studies, cosmic-ray and collision data.

Chapter 4

Event Selection and Background Modeling

The analysis of proton-proton-collision data with regard to a specific process requires the separation of the signal from the various background processes. Figure 4.1 shows the cross-sections of hard-interaction processes as a function of the center-of-mass energy. The analyzed collisions in this thesis are recorded at a center-of-mass energy of $\sqrt{s} = 7$ TeV. At the LHC, the cross-section of top quark production (σ_t) is suppressed by eight orders of magnitude compared to the total cross-section (σ_{tot}). Therefore, the analyzed signal process has to be extracted from a huge amount of background events. The selection of events is based on reconstructed physics objects and event-kinematic requirements with regard to their difference in signal and background processes.

The decay of a top quark pair can be subdivided into three decay channels, according to the number of leptons in the final state. Thus, resulting decay-channels are linked to the decay of the W-bosons, discussed in section 1.2.1. A different potential in terms of signal to background ratio, frequency of occurrence and reconstruction issues can be observed. With the highest probability, both W-bosons decay hadronically. While this decay channel has a high frequency, the resulting event signature is hard to distinguish from multi-jet background events ($\sigma_{\mathrm{QCD}} \approx \sigma_{\mathrm{tot}} \approx 10^8$ nb). The case that both W bosons decay leptonically is less likely. While the two resulting leptons allow a good separation from background processes, the disadvantage of this channel is a very low branching ratio compared to the other decay-channels. In addition, this process contains two neutrinos in the final state which makes an independent reconstruction of all decay products impossible.

The semileptonic decay of a top quark pair, drafted in Figure 1.7 on page 12, is the topic of this thesis. An isolated electron or muon in the final state, originating from the leptonic W-boson decay, is used to trigger candidate events. In addition to the W-boson decay into an $e\nu_e$ or $\mu\nu_\mu$ pair, the decay $W \to \tau\nu_\tau$, with the τ decaying

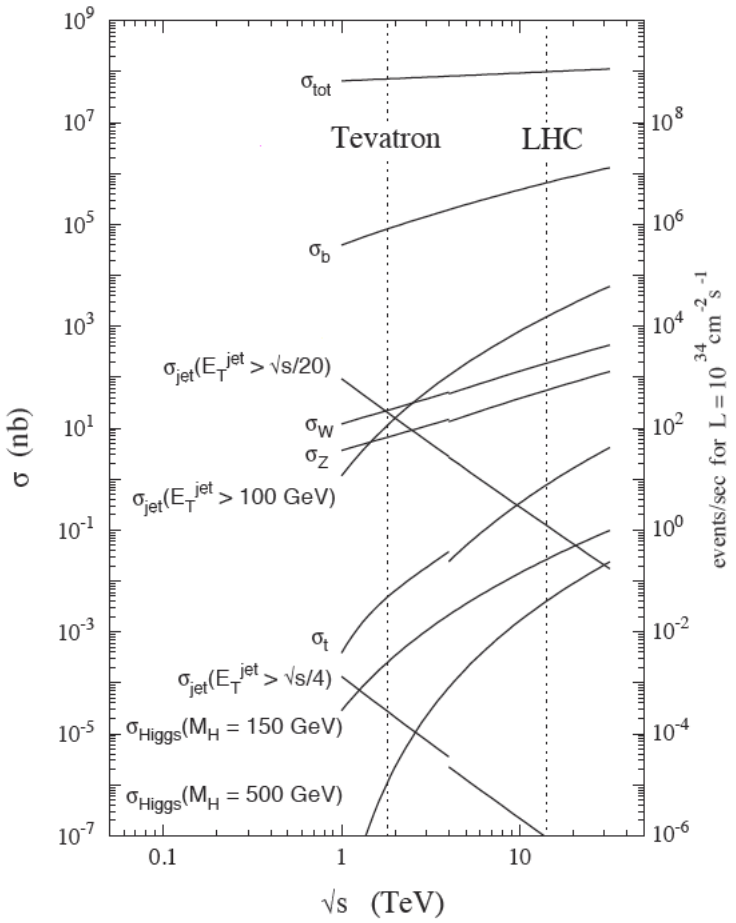


Figure 4.1: *Cross-section of various processes in the proton collisions as a function of the center-of-mass energy assuming a luminosity of* $1 \times 10^{34} cm^{-2} s^{-1}$ *[58].*

into an isolated electron or muon, is also selected. The second W-boson is required to decay hadronically. Hence, the event signature of a semileptonic top quark pair decay consists of one isolated high energetic electron or muon, missing transverse energy originating from the neutrino, two jets from the hadronically decaying W-boson and two jets including B-hadrons, originating from the top quark decays. Therefore, this $t\bar{t}$ decay is referred to as lepton plus jets top quark decay. The lepton plus jets signature allows a sufficient background suppression and achieves a high statistics of events. This chapter introduces the reconstruction of top-quark pair event candidates and discusses the considered background processes. Furthermore, the systematic uncertainties on the predictions are discussed. The descriptions in this chapter base on the prescriptions of top-quark reconstruction defined by the ATLAS collaboration.

4.1 Selection of Semileptonic $t\bar{t}$ Events

This section introduces the selection of semileptonic top-quark pair event candidates, described in detail in [68]. The event signature of a semileptonic top-quark-pair decay consists of one high energetic lepton (electron or muon), missing transverse energy and four jets. The lepton originating from the W boson decay is isolated from the jets and serves as seed to trigger the event. Even if the event signature is quite significant, also other processes with event properties similar to the signal process can lead to an identical event signature. These events have to be considered as background events in the analysis. Within the selection of events, specific requirements on the reconstructed physics objects, such as electrons, muons and jets, are defined to reduce the influence of mis-identified objects in background processes. The selection of events is subdivided into the selection of events with an electron or a muon in the final state, referring to the charged lepton of the leptonic W boson decay.

4.1.1 Electron Selection

The selection of top-quark events associated with an isolated electron starts with the requirement of an electron-event trigger. An electromagnetic-cluster energy of 20 GeV to 22 GeV is demanded for the trigger objects. The electron-trigger efficiency, measured in decays of $Z \to ee$ and $W \to e\nu$, is almost constant and above 98%, requiring objects with a transverse momentum larger than 25 GeV [59].

The electron object in the event reconstruction, matched with the event-trigger object, has to fulfill the tight++ identification criteria. Since the momentum of the lepton is high in case of the top-quark decay, a transverse momentum of $\mathrm{p_T} > 25$ GeV is required. With the demand of an acceptance $|\eta| < 2.47$ the electron track is detected within the two innermost detector systems and its energy can be measured in the calorimeter volume. The energy deposit measured in the calorimeter cluster is calibrated based on the electromagnetic energy scale [77] The transition region of the electromagnetic calorimeter $1.37 <| \eta |< 1.52$ is excluded from the sensitive region to reduce uncertainties in the energy calibration.

Thus the lepton of the W boson decay is isolated from the remaining top-quark decay products, an isolation criteria can be used to reduce the contribution of misidentified-electron objects. An electron very close to a jet would indicate to a heavy-flavor decay inside the jet. If there is another jet within a cone of $\Delta R = 0.4$ around the electron and with a $\mathrm{p_T} > 20$ GeV after the jet-electron overlap removal, discussed later on, the electron is discarded. To suppress the contribution of QCD multijet background events with mis-identified electrons, tight isolation cuts are imposed on the electron. A measure for extra energy in the calorimeter and additional track momenta is given by a cone with the radius 0.2 and 0.3 around the electron track.

These cone variables are called E_T^{cone} and ${p_\mathrm{T}}^{cone}$ [68]. A relative isolation criterion at an efficiency of 90% is demanded with requirement on the E_T^{cone20} and ${p_\mathrm{T}}^{cone30}$ variable [68].

4.1.2 Muon Selection

Event candidates associated with a muon require a muon-event trigger-object with a p_T larger than 18 GeV [**?**]. The muon selection requires an isolated muon, reconstructed by the combined MuidMuon algorithm in the region of $|\eta| < 2.5$. The transverse momentum is defined to be larger than 20 GeV. Since the muon is reconstructed in the inner detector and in the muon system of the ATLAS detector, the selection of muon objects has a minor fake contributions compared to the electron object selection. Hence, the muon object is defined by an absolute isolation $E_T^{cone20} < 4.0$ and ${p_\mathrm{T}}^{cone30} < 2.5$ [68]. Non-prompt muons from heavy-flavor decays within jets are rejected by a cut on muons overlapping with a jet axis within $\Delta R < 0.4$, considering jets with a p_T larger than 20 GeV.

4.1.3 Jet Selection

The jets, included in this analysis, are reconstructed from topological clusters by the ATLAS anti-k_T algorithm with a cone-width parameter of 0.4. The topological clusters are calibrated by the electromagnetic scale. The energy of the jets has to be corrected by the hadronic energy scale, which corrects the effects of charged and neutral particles in the hadron shower. Also detector effects originated by dead material in front of the calorimeter and calorimeter leakage are taken into account. Therefore, the jet-energy scale is calibrated by comparing the energy of reconstructed jets with the energy of simulated truth jets in bins of η and p_T [77].

The reconstructed jets have to pass a transverse momentum requirement of at least 25 GeV. Since the jet algorithm also classifies electromagnetic objects as jets, jets with the axis within $\Delta R = 0.2$ from the electron direction are removed from the event. This procedure is called jet-electron overlap removal. The Jet-Vertex Fraction (JVF) observable defines the fraction of tracks in the jet which originate in the primary vertex. To remove the contribution of pile-up jets from secondary proton interactions, a minimum threshold on the JVF of 0.75 is applied [68].

The measurement of the jet energy can be corrupted by problems in the calorimeter hardware, beam-gas interactions or cosmic-ray induced showers. The 'bad-jet' quality-criterion [**?**] refers to a jet with an energy contribution from one of these effects. Such a jet in the event would lead to a mis-measurement of the $E_\mathrm{T}^\mathrm{miss}$. Since the precise $E_\mathrm{T}^\mathrm{miss}$ measurement is important in the reconstruction of the leptonic W boson decay, a veto on events containing a bad-jet with a transverse momentum larger than 20 GeV is applied.

4.1.4 Event Requirements

Since most of the recorded processes are not of interest, the collected set of recorded proton-collision data has to be reduced. Events with a semileptonic top-quark decay can be triggered on account of their lepton. Furthermore, a primary vertex has to be reconstructed with at least five associated tracks and a position consistent to the $x - y$ profile of the beam spot. The exact reconstruction of a primary vertex is important, since the information is passed to the track and jet reconstruction and especially to the b-tagging algorithm. The event is required to contain exactly one high p_{T} isolated lepton after the application of the object requirements. This lepton has to match the event-trigger object. If an electron track is overlapping with a muon candidate, the event is rejected.

A further separation of background events is facilitated by applying requirements on the W boson decay. The neutrino, involved in the leptonic W-boson decay, can only be quantified by the amount of missing transverse energy in the event. Therefore, observables characteristic for a leptonic decay of a real W boson such as the missing transverse momentum and the W boson mass in the transverse plane are applied. The missing transverse momentum is calculated using the MetRefFinal algorithm described in Section 3.1.6. The mis-reconstruction of a jet as an electron or a weak hadron decay inside a jet can also lead to a certain amount of misidentified $E_{\mathrm{T}}^{\mathrm{miss}}$. To reject the fraction of these multi-jet background in the signal selection, the kinematic construction of a transverse W boson mass ($\mathrm{m_T}(W)$) is calculated. The transverse W boson mass is defined as:

$$m_T(W) = \sqrt{(2p_{\mathrm{T}}E_{\mathrm{T}}^{\mathrm{miss}}(1 - cos(\Delta\phi))} \tag{4.1}$$

where p_{T} is the transverse momentum of the charged lepton and $\Delta\phi$ the azimuthal angle between the charged lepton and the missing transverse momentum. In the electron selection channel, both the missing transverse energy and the transverse W boson mass are required to be larger than 30 GeV. In the muon selection channel, the requirements are $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV and the sum $\mathrm{E_T^{miss}}+m_{\mathrm{T}}(W) > 60$ GeV.

Events containing less than four selected jets are rejected, based on the expectation, that two b-quark jets originate from the top-quark decay and two jets produced by the hadronic W boson decay are present. To suppress the fraction of multi-jet events, the requirement that at least one of the jets in the event is identified as b-jet can be applied additionally. According to this requirement, the selection is divided into two samples, referred to as 'pre-tag' or 'tagged' sample, respectively. The b-jet identification in the top-quark selection is based on the MV1 b-tagging algorithm. A working-point on the identification is set corresponding to a required tagging efficiency of 70% [73]. This leads to a light-quark rejection factor of 140 for jets with a p_{T} larger than 20 GeV [75]. A summary of the object requirements and the selection criteria on the event candidates is given in Table 4.1.

Cut	Value
Electron Object	
trigger	EF_e20_medium, EF_e22_medium, EF_e22vh_medium1
transverse momentum	$p_{\mathrm{T}} > 25$ GeV
acceptance	$\mid \eta \mid < 2.47$, $1.37 < \mid \eta \mid < 1.52$ excluded
quality and isolation	tight++, Etcone20@90, ptcone30@90
jet overlap	$\Delta R(e, jet) > 0.4$
Muon Object	
trigger	EF_mu18_medium
transverse momentum	$p_{\mathrm{T}} > 20$ GeV
acceptance	$\mid \eta \mid < 2.5$
quality and isolation	tight, Etcone20 < 4 GeV, ptcone30 < 2.5 GeV
jet overlap	$\Delta R(\mu, jet) > 0.4$
Jet Object	
transverse momentum	$p_{\mathrm{T}} > 25$ GeV
acceptance	$\mid \eta \mid < 2.5$
quality	no bad-jet with $p_{\mathrm{T}} > 20$ GeV
Event Level Cuts	
primary vertex	at least 1 good vertex with $N_{\mathrm{tracks}} \geq 4$
lepton	one isolated electron or muon
$E_{\mathrm{T}}^{\mathrm{miss}}$	$\mathrm{E_T^{miss}} > 30$ GeV(electron)
	$\mathrm{E_T^{miss}} > 20$ GeV(muon)
transv. W mass	$\mathrm{m_T}(W) > 30$ GeV(electron)
	$\mathrm{m_T}(W) > 60\ GeV$ - $E_{\mathrm{T}}^{\mathrm{miss}}$ (muon)
jets	at least four jets
b-tag	one jet tagged by MV1 at 70% efficiency

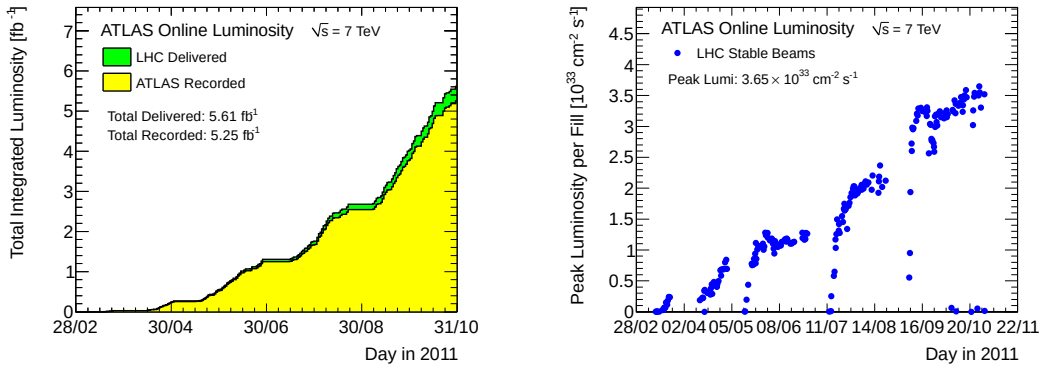
Table 4.1: *Selection cuts for the reconstructed objects and event candidates.*

4.2 Data Samples and Simulated Events

The analysis of the physics in *pp* collisions is based on the comparison of reconstructed events with events generated by MC simulations. This section introduces the recorded dataset of *pp* collisions and the simulated samples of signal and background processes.

4.2.1 Recorded Data

In the proton-proton collision operation phase of the year 2011 with a center-of-mass energy of $\sqrt{s} = 7$ TeV, the LHC delivered 5.61 fb^{-1} of integrated luminosity to the ATLAS experiment. Vast improvements of the beam conditions and the operation of the ATLAS experiment are visible in the distribution of the delivered and recorded integrated luminosity and the peak instantaneous luminosity per fill, shown with respect to the operation time in Figure 4.2.



(a) (b)

Figure 4.2: *Histogram of the integrated luminosity (a) and the instantaneous peak luminosity per fill (b) of* $\sqrt{s} = 7$ TeV *proton collisions as a function of days, delivered to and recorded by the ATLAS experiment during the period of collecting data in the year 2011 [60].*

The analysis presented in this thesis is based on the complete set of proton-proton collision data collected in the year 2011. Since the event selection requires a good functionality of all detector sub-systems, not all recorded data is suitable for this analysis. The Pixel and the SCT detector, for example, with their short distance to the beam pipe, are not powered during the beam injection phase. Large signal amplitudes, induced by unstable beam conditions, would lead to high electrical currents damaging the detector modules. Hence, collisions before a 'stable-beam' signal was given by the LHC operation, are rejected. In addition, run-periods with a restricted detector operation or restrictions in the data quality are not considered. The exact list of good quality run-periods is defined in a so called *Good Run List*. The analyzed data corresponds to an integrated luminosity of 4.7 fb^{-1}.

4.2.2 Simulated Events

Connected with the collection of data, various physical processes are generated in simulations with identical beam and detector conditions. Hence, the recorded data

can be directly compared to simulated events. In this section, the samples of simulated signal and background processes are introduced. A list of all samples, their cross-sections, normalization-factors and the number of generated events can be found in Table 6.5 in the appendix.

Signal Process

Apart from Chapter 6, where resonant top-quark production is discussed, the signal process is defined as SM top-quark-pair production. It is generated with the MC@NLO event-generator, providing a next-to-leading order calculation of the $t\bar{t}$ matrix-element. The parton-shower and the underlying event are added using the HERWIG and JIMMY algorithm. Both MC@NLO and HERWIG use the $CTEQ6L1$ [61] PDF set. To reduce the size of the dataset, events in which both W bosons decay hadronically are excluded. In total, around 15 million events are generated to model the distributions of top-quark pair events.

Background Processes

The selection of the signal events is contaminated by a certain fraction of background events. An accurate description of these processes is necessary to achieve a precise measurement of the signal. The main background contribution in the lepton plus jets selection stems from events with a real W boson decay. In addition, mis-reconstructed objects of physical processes, suppressed by event selection criteria, can lead to an event signature identical to the signal process.

- The SM process, with a final state nearly identical to the signal process, is the production of a real W boson in the leptonic decay mode in association with more than three jets. The decay of the W boson is not distinguishable in both kinds of events. If the W boson production is associated with b-quarks or a jet is mis-identified as b-jet, the event signature is nearly identical to the signal process.

- The production of Z bosons decaying into leptons in association with jets can also fake the signal signature. The loss or mis-identification of one of the two leptons leads to a similar event signature. The amount of $E_{\mathrm{T}}^{\mathrm{miss}}$, necessary to pass the event selection-requirements, can be reached by a corrupted $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement, for instance due to pile-up energy in the calorimeter or a missed lepton.

- A small background contribution derives from di-boson production, such as WW, WZ and ZZ events, associated with jets. The dominant input to this background fraction is expected to come from the WW production if one W boson decays leptonically while the other one decays hadronically.

- The production of single top-quark events has a low cross-section and is expected to take place in the small jet multiplicities. In particular the single top-quark production with a W boson in the final state (Wt channel) can lead to a background contribution in the four jet selection.
- QCD multi-jet events can pass the event selection if a lepton is reconstructed. Non-promt leptons, originating from a heavy-flavor decay inside a jet, can fake the signal leptons. In addition, a misidentified jet in the reconstruction can fake an electron. The $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold in those 'lepton+jets' events can be reached by mis-measurements in the calorimeter.

The production of leptonically decaying vector bosons, such as $Z/\gamma^*(\rightarrow ee$, $\mu\mu$ and $\tau\tau)$ and $W(\rightarrow e\nu$, $\mu\nu$ and $\tau\nu)$ associated with jets, is simulated by ALPGEN, interfaced to HERWIG and JIMMY. The ALPGEN parameters, controlling the minimal transverse momentum and angular separation of light quarks in the hadronization process, are set to $\mathrm{p}_{\mathrm{T,j}}^{\mathrm{min}} = 15$ GeV and $\Delta R_j^{min} = 0.7$. An overlap of heavy-flavor production can be caused by a double counting of matrix-element and parton-shower heavy-flavor production. Hence, the heavy-flavor overlap-removal tool [62] corrects this double counting.

Di-boson events, such as WW, WZ and ZZ, are produced using HERWIG in association with JIMMY. For both ALPGEN and HERWIG the matrix element calculations and the parton-shower evolution are based on the $CTEQ6.1$ [61] PDF. The cross-sections for the vector-boson production are normalized to next-to-leading-order cross-section predictions, presented in [63]. To provide a complementary sample of vector-boson production, events generated by SHERPA are used in Chapter 5.3.

The single-top-quark production in the s-channel, t-channel and Wt production is modeled by ACERMC, interfaced to PYTHIA to parametrize the hadronisation process.

A simulation of the multi-jet background mis-identified as lepton plus jets events is challenging. The mis-identification of the lepton in the object reconstruction is associated with instrumental detector effects. Also, the mis-measurement of the $E_{\mathrm{T}}^{\mathrm{miss}}$ depends on the calorimeter-system performance. Since these effects are hard to parametrize in the simulation, the multi-jet background is derived from a data-driven model, that is explained in the next section.

Corrections on the Simulation

Since the simulated collision processes and the simulated detector response have not necessarily to be identical to the physical processes in nature, deviating effects have to be adjusted. Therefore, differences between data and simulated events are corrected by the application of scale factors, derived from data. For instance, the number of collisions in each bunch crossing varies during the LHC operation phase.

To model the number of interactions correctly, the simulated events were re-weighted according to their run-period. Also, the jet-vertex fraction cut in the selection of jets shows differences in the selection efficiency comparing data and simulation. Using a tag and probe method, which analyzes hadronically decaying Z+jets events, scale-factors to compensate this difference are determined. Furthermore, event-weight corrections for the trigger efficiencies are applied. A detailed description of all corrections can be found in [68].

4.3 Multi-Jet Background Modeling

Lepton plus jets events are identified by an isolated high p_{T} lepton in the final state, passing the trigger, object and analysis requirements. In addition to the prompt electrons of the signal process, an additional fraction of non-prompt or mis-identified leptons passes the object requirements. Main sources for fake leptons are:

- semileptonic heavy-flavor decays inside a jet
- the weak decay of long lived particles like the $\pi^{\pm}$ or K meson
- mis-identification of a π^0 shower as an electron
- reconstruction of an electron which stems from a conversion or a direct photon

Events with gluon exchanges in the hard interaction, also referred to as multi-jet events, have the highest cross-section compared to all other processes at the LHC. Even though the electron reconstruction and selection requirements are very tight, the large amount of QCD multi-jet events leads to a significant fraction of events, where a lepton is wrongful reconstructed.

The shape of the QCD multi-jet background strongly depends on the object reconstruction and the performance of the calorimeter system responsible for the mis-measurement of the $E_{\mathrm{T}}^{\mathrm{miss}}$. These instrumental effects are difficult to include in the detector simulation. Thus, in order to simulate QCD multi-jet background-events, data-driven models are applied. Different approaches for estimating the multi-jet background are presented in this section.

4.3.1 Matrix Method

The Matrix Method [64] studies the difference in lepton-identification related properties based on prompt leptons and non-prompt or mis-identified leptons. For this purpose, two samples with different lepton-identification requirements are defined, after the final kinematic selection criteria are applied: A 'tight' sample, typically

identical to the signal identification criteria, and a 'loose' sample with more basic requirements. This method is based on the assumption that both samples can be expressed as a linear combination of the number of real and fake lepton events.

$$N^{loose} = N^{loose}_{real} + N^{loose}_{fake} \tag{4.2}$$

$$N^{tight} = \epsilon_{real} N^{loose}_{real} + \epsilon_{fake} N^{loose}_{fake} \tag{4.3}$$

The factor $\epsilon_{real}(\epsilon_{fake})$ is parametrizing the probability that a real(fake) lepton that satisfies the loose criteria, also passes the tight ones. These factors can be estimated from data in control regions. The efficiency for a real lepton ϵ_{real} is derived from Z-boson events by a tag and probe method. The determination of the fake efficiency ϵ_{fake}, which is measured in a QCD enriched data-sample, is more challenging. In terms of 'tight' and 'loose' event numbers, these efficiencies are defined as:

$$\epsilon_{real} = \frac{N^{tight}_{real}}{N^{loose}_{real}} \tag{4.4}$$

$$\epsilon_{fake} = \frac{N^{tight}_{fake}}{N^{loose}_{fake}} \tag{4.5}$$

Solving the system of equations, the number of estimated events with a fake lepton in the tight selection is defined as:

$$N^{tight}_{fake} = \frac{\epsilon_{fake}}{\epsilon_{real} - \epsilon_{fake}} (N^{loose} \epsilon_{real} - N^{tight}) \tag{4.6}$$

Equation 4.6 allows one to determine the normalization of the multi-jet background from data. Bin-wise applied, it is used to derive the shape of the fake-lepton event distributions in any observable of interest. However, the precision of the method depends on the definition of the control region. If the efficiencies are not sufficiently different, the statistical precision is compromised. An independence of the efficiencies from the event topology is also necessary.

4.3.2 Jet-Electron Model

The Jet-Electron mis-identification model [64] is able to directly derive the shape of the multi-jet background from data. Therefore, multi-jet events are selected from collision data. Jets with a high electromagnetic energy fraction are used to model the mis-identified electron. The selection of event candidates is done by a set of

Process	Constrain
$t\bar{t}$ and single top	5%
W+jets	10%
Z+jets	10%
di-boson	5%

Table 4.2: *Normalization constrains of simulated SM processes in the fit estimating the multi-jet fraction.*

pre-scaled jet triggers, starting at a jet p_{T} threshold of 10 GeV. Events containing exactly one jet with a high electromagnetic energy fraction $0.8 < f_{EM} < 0.95$ are requested. To remove the signal contamination from prompt electrons, an electron veto is applied. Also, the background contribution of converted photons is reduced with the requirement of at least four tracks in the electron fake candidate. After the transition of the jet into an electron, the events have to pass the default event selection criteria. For the recalculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$, the jet energy is re-calibrated to the electromagnetic energy-scale. Thus, a template of the QCD multi-jet background-events with properties identical to the multi-jet events is produced. Therefore, the Jet-Electron model provides a robust modeling of the multi-jet background. The major disadvantage is the prescale of the jet trigger, which leads to a relatively low statistics of events.

An example of distributions generated by the JetElectron template is shown in Figure 4.3. The shape is compared to the distribution of signal and background processes after the event selection except for the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement. The fake lepton events indicate smaller values of missing transverse energy (a) and transverse W boson mass (b), since the origin of the $E_{\mathrm{T}}^{\mathrm{miss}}$ has instrumental reasons. The p_{T} of a fake lepton (c) is typically softer than the p_{T} of a prompt lepton originating from a W boson decay. The same is true for the jet with the highest momentum, which is softer in case of a QCD multi-jet interaction.

In general, the JetElectron model is designed to describe the shape of the QCD multi-jet background in the electron channel. Since events including non-prompt leptons from heavy-flavor decays and events with a misidentified jet nearly have the same kinematics, it is justified to also apply the JetElectron model for the multi-jet background modeling in the muon channel [64]. To obtain the normalization of the QCD-multi-jet background, a binned maximum-likelihood fit is used. The fit is performed separately in the electron and muon channel. To enrich the fraction of multi-jet background and increase the discriminating power, the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut is not applied at that stage. An additional freedom is provided by leaving the rate of expected SM processes floating. The constraints, shown in Table 4.2, are defined within the expected uncertainties on the expected cross-sections.

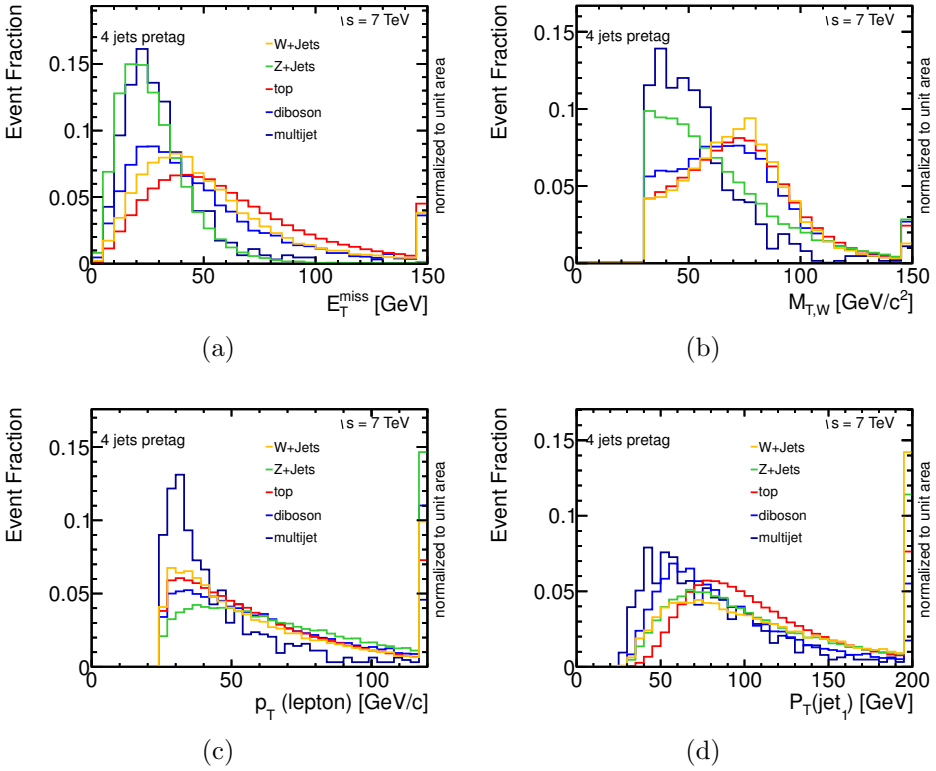


Figure 4.3: *Examples of QCD multi-jet background distributions derived with the JetElectron model. The missing transverse energy (a), the transverse W boson mass (b), the p_{T} of the lepton (c) and the p_{T} of the leading jet (d) are shown. All event selection cuts are applied, except for the b-tag requirement and cut on the $E_{\mathrm{T}}^{\mathrm{miss}}$. The SM expectation of the other processes of interest is plotted as reference.*

The observables with the best performance in terms of the fit result are identified as the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the transverse W boson mass distribution. Since the separation power of the multi-jet background in both observables is similar, as shown in Figure 4.3, the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution is chosen for the template fit. Ideally the fit would be performed in the low $E_{\mathrm{T}}^{\mathrm{miss}}$ side-band region, independent from the signal region. Because the fit in the full $E_{\mathrm{T}}^{\mathrm{miss}}$ range provides more stable results, especially for higher jet multiplicities, this possibility is rejected in the analysis later on. The estimation of the QCD fraction in the 'pre-tag' four jet multiplicity is shown in Figure 5.5 for the electron and in Figure 5.6 and for the muon selection.

Additional kinematic variables, shown in Figures 4.4 and 4.5 for the electron and the muon channel, validate the accurate normalization of the multi-jet background fraction. Except for low values of muon p_{T} , a good agreement of the data and the expected processes can be observed in the control distributions. The slight disagreement of the p_{T} description in the muon channel can be explained by the fact, that the muon in multi-jet events is mainly generated by semileptonic heavy-

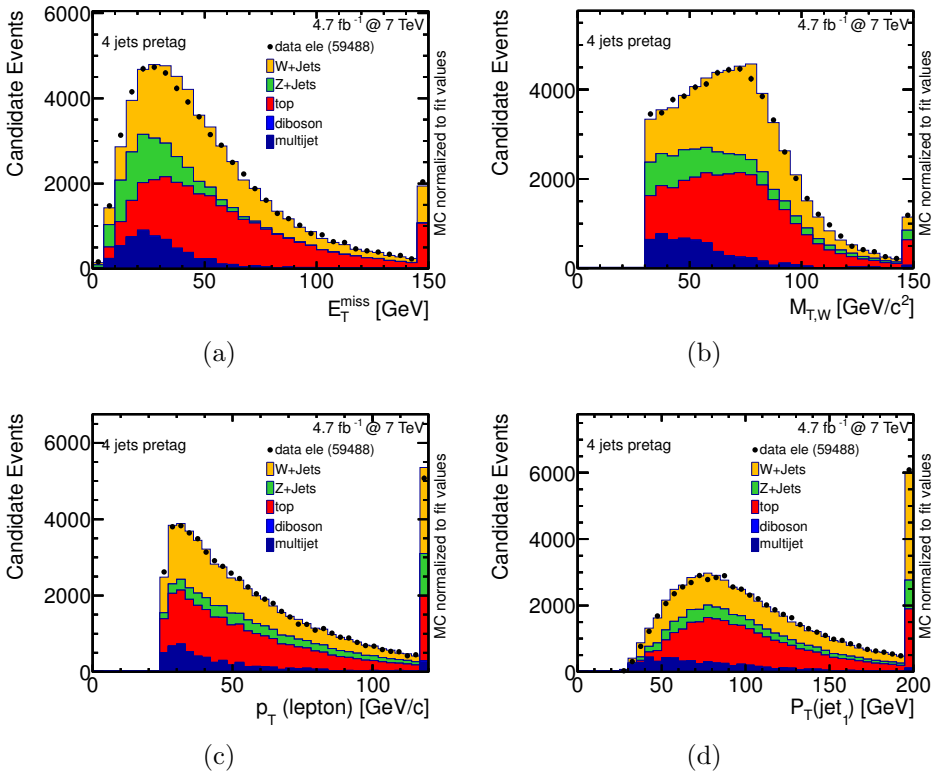


Figure 4.4: *Estimation of the QCD multi-jet fraction in the 'pre-tag' four jet multiplicity for the electron channel using the JetElectron model. Figure (a) shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, used to estimate the fraction of multi-jet background. Distributions like the transverse W boson mass (b), the p_{T} of the lepton (c) and the p_{T} of the leading jet (d) validate the fit result. All selection cuts, except for the cut on the $E_{\mathrm{T}}^{\mathrm{miss}}$, are applied.*

flavor decays inside the jet. Hence, these muons are expected to have slightly less p_{T} than the mis-identified electron candidates.

The uncertainty of the estimated rate is conservatively set to 50% [64]. These uncertainty is motivated by a comparison with the Matrix Method. In addition variations of the fit result due to a different choice of observables in the estimation of the multi-jet background content were studied. Furthermore, the dependence on the phase space is investigated, for instance, analyzing the fraction of multi-jet events with a high pile-up contribution.

4.4 W+jet Background Estimation

The production of W bosons in association with at least four jets is the dominant background in the analysis of semileptonic $t\bar{t}$ decays. The final-state signature of both processes, consisting of one isolated lepton, missing transverse energy and

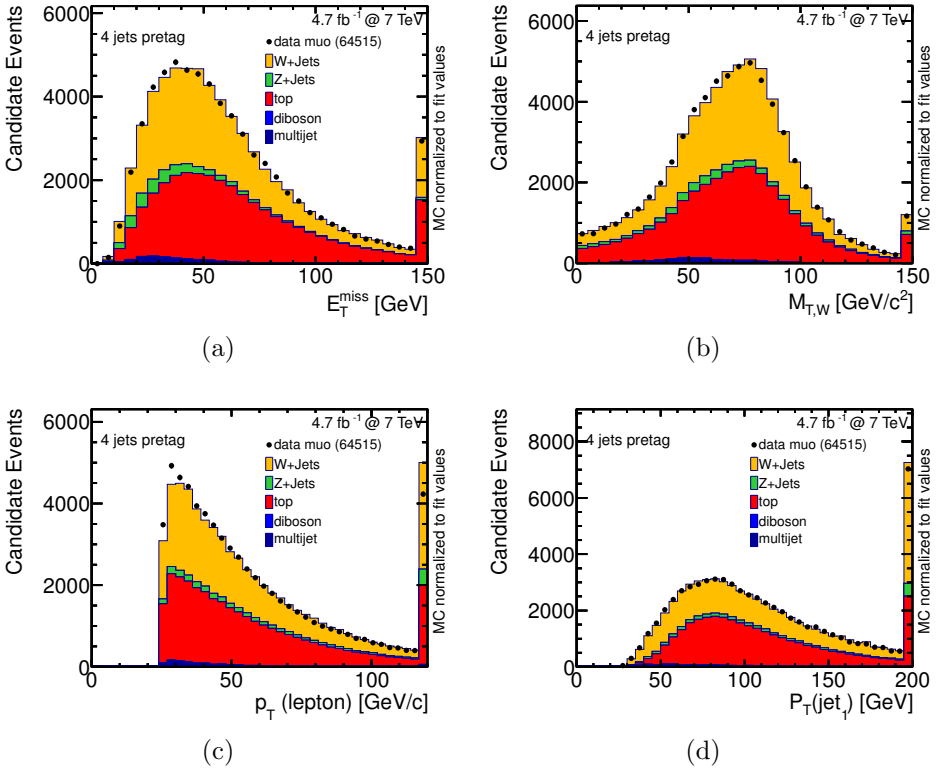


Figure 4.5: *Estimation of the QCD multi-jet fraction in the four jet multiplicity for the muon channel using the JetElectron model. Figure (a) shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, used to estimate the fraction of multi-jet background. Distributions like the transverse W boson mass (b), the p_{T} of the lepton (c) and the p_{T} of the leading jet (d) validate the fit result. All selection cuts, except the cut on the $E_{\mathrm{T}}^{\mathrm{miss}}$, are applied.*

at least four jets, is almost identical. Hence, a separation of the signal process from this background process is challenging. In the analysis presented, the production of W bosons in association with jets is simulated by ALPGEN and normalized to NNLO-calculations of the W boson production cross-section [65]. Even if the W+jets normalization is precisely calculated for small jet multiplicities, the perturbation calculation of QCD effects lead to considerable uncertainties for larger jet multiplicities [66]. Also the composition of W+jets events, produced in association with heavy-flavor quarks, is not accurately known from perturbation theory. Hence, the W+4 jets normalization and the W+jets heavy-flavor composition is not accurately known from theory. Therefore, data-driven techniques to estimate the W+jets background are required. This section introduces the recommended approaches of the ATLAS top group analysis to estimate the W+jets normalization and the composition of the W+jets heavy-flavor fraction from data.

4.4.1 Determination of the W+jet Normalization

The determination of the overall W+jets normalization is based on the asymmetry in the number of charged W bosons, generated in pp collisions. The asymmetry is a result of different quark and anti-quark distributions in the PDFs. Positively charged W bosons are produced in parton-level processes such as $u\bar{d} \to W^+$ while negatively charged W bosons are produced in processes such as $d\bar{u} \to W^-$. The ratio of the cross-sections $r \equiv \frac{\sigma(pp\to W^+)}{\sigma(pp\to W^-)}$ is theoretically understood within a few percent and nearly independent on the number of additionally produced partons [47, 67]. The generation of additional jets in association with the W+jets production rests on QCD processes and, therefore, it has to be extracted in perturbation theory calculations. In comparison with the charge-asymmetry prediction, the W boson production cross-section in association with more than three jets is known less precisely [65,68]. Therefore, the theoretical prediction for the ratio r is used to estimate the overall W+jets background normalization.

This assumption is implemented in the Charge-Asymmetry (CA) method [68]. The production of W+jets leads to the dominant contribution of the asymmetry observable in data. Therefore, an estimation of the W+jets normalization can be achieved. It is based on the following equation:

$$N_{W^+} + N_{W^-} = \frac{N^{MC}_{W^+} + N^{MC}_{W^-}}{N^{MC}_{W^+} - N^{MC}_{W^-}}(D^+ - D^-) = \frac{r_{MC}+1}{r_{MC}-1}(D^+ - D^-) \tag{4.7}$$

where $D^+(D^-)$ is the number of events with a positively and negatively charged lepton reconstructed in data and r_{MC} the ratio of positively and negatively charged W boson events predicted by the simulation. The formula leads to a good approximation, since the signal process and all other background processes, except for the single top-quark production, are assumed to be symmetric with regard to their lepton charge. The contribution of charge asymmetric background processes in the selected data are subtracted according to their MC predictions. Since the CA method depends on the W+jets heavy-flavor composition, data-driven heavy-flavor scaling factors, introduced in the next section, are applied in the estimation of r_{MC}. The normalization of the W+jets prediction is corrected by the application of scale factors (SF) in the analysis. They are calculated from the ratio of the number of W+jets events determined in data and predicted by the simulation.

The number of W+jets events, containing at least one jet identified as b-jet, is estimated from the estimated number of pre-tag W+jets events multiplied by W+jets tagging efficiencies f^{2j}_{tag} and $f^{2\to4}_{tag}$:

$$W^{4-jet}_{\geq 1tag} = W^{4-jet}_{pretag} \times f^{2j}_{tag} \times f^{2\to4}_{tag} \tag{4.8}$$

The tagging efficiency of W+jets events in the two jet multiplicity f^{2j}_{tag} is determined in data. The extrapolation of the tagging efficiency into the four jet multiplicity

$f_{tag}^{2\to4}$ is based on simulations. Since the heavy-flavor content of W+jets production is determined data-based in the two jet bin, as discussed in the next section, the tagging efficiency f_{tag}^{2j} in data and simulation is identical by construction. Thus, in the calculation of the SF in the b-Hadron tagged selection, both tagging efficiencies f cancel each other out [68]. Therefore, the pre-tag and tagged W+jets SF have the same nominal value, though they have different systematic uncertainties. Table 4.3 shows the SF for the electron and the muon channel requiring exactly four jets in the event.

	Electron	Muon
pre-tag	$0.83^{+0.08}_{-0.08}$	$0.87^{+0.07}_{-0.07}$
tagged	$0.83^{+0.14}_{-0.13}$	$0.87^{+0.11}_{-0.11}$

Table 4.3: *Scaling factors to correct the W+jets overall normalization, derived from the charge-asymmetry method in events with at least four jets [68].*

4.4.2 Estimation of the W+jets Heavy Flavor Content

The composition of W+jets events, produced in association with heavy-flavor quarks, is not precisely known from theoretical calculations. Figure 4.6 shows the Feynman diagrams of the main production processes. The splitting processes of the gluon, for instance $g \to bb$ in Figure 4.6(c), is not predictable in perturbation calculation.

Heavy-flavor scale factors F_{bb}, F_{cc} and F_c are derived from data to correct the composition of W_{bb}, W_{cc} and W_c production in the MC simulation [68]. For each jet bin, the relation of the pre-tag and tagged number of events is:

$$N_{W^\pm}^{tag} = N_{W^\pm}^{pre-tag}(F_{bb}P_{bb} + F_{cc}P_{cc} + F_cP_c + F_{light}P_{light}) \tag{4.9}$$

where P_x is the b-tagging probability for each W+jets heavy flavor type x. To keep the overall W+jets normalization constant, the sum of all scale factors is required to be one:

$$F_{bb} + F_{cc} + F_c + F_{light} = (1 + k_{cc2bb})F_{bb} + F_c + F_{light} = 1 \tag{4.10}$$

Since the production of W_{bb} and W_{cc} is based on the same mechanism, the ratio $k_{cc2bb} = \frac{F_{bb}}{F_{cc}}$ is constant.

For the final calculation of the W+jets heavy-flavor scaling factors, a system of equations based on Equation 4.9, separated into the number of positively and negatively charged lepton events, and Equation 4.10 has to be solved. The b-tagging probabilities P_x and the ratio of W_{bb} and W_{cc} production k_{cc2bb} are taken from simulation. The number of pre-tag and tagged W+jets events, required in Equation

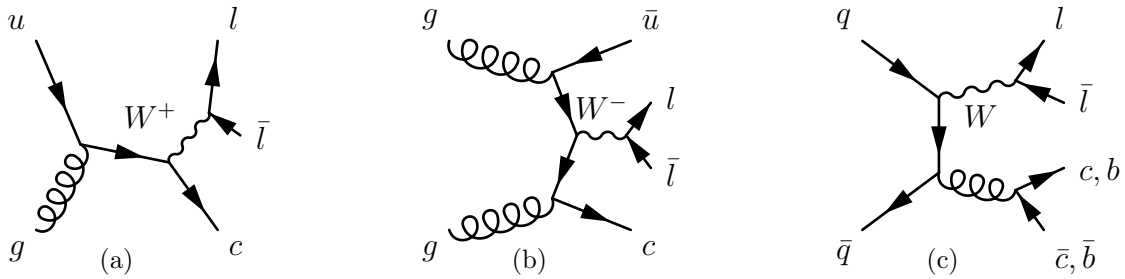


Figure 4.6: *Feynman diagrams of W+jets production processes associated with heavy-flavor jets.*

Channel	Electron	Muon
K_{bb}	1.22 ± 0.29	$1.36^{+0.30}_{-0.29}$
K_{cc}	0.95 ± 0.23	$0.71^{+0.30}_{-0.31}$
K_{light}	0.98 ± 0.07	$1.01^{+0.08}_{-0.08}$

Table 4.4: *W+jets heavy-flavor scaling factors including statistical and systematical uncertainties. The estimation is done in the exclusive two-jet multiplicity [68].*

4.9, is estimated with the CA method. Since the method depends on the W+jets heavy-flavor fraction, both the W+jets heavy-flavor scaling factors and the overall W+jets normalization are calculated iteratively until the best accordance is reached.

The exclusive two-jet multiplicity was evaluated to show the best performance in the estimation of the heavy-flavor scaling factors [68]. Hence, the scaling factors for the analysis are estimated in a two-jet selection. The uncertainty of the extrapolation of the W+jets flavor scale-factors from the two jet into the four jet multiplicity is estimated to be less than 25% [68], based on the variation of ALPGEN parameters. Finally, scale factors K_{xx} are calculated from the determined W+jets heavy-flavor fractions. Equivalent to the W+jets normalization scale factors, they are used to correct the predicted W+jets heavy-flavor fractions in the simulated samples. The W+jets heavy-flavor scaling factors are shown in Table 4.4.

4.5 Systematic Uncertainties

Simulated events are required to describe the physics of the event generation, the parton shower and the response of the ATLAS detector. Their calculations are based on models which can only approximate the processes occurring in nature and, therefore, leading to systematic uncertainties in the prediction of simulated events. They can be roughly divided into detector, and signal and background modeling effects. In this section, the different sources of systematic uncertainties considered in the analysis of top-quark events are shown [68]:

Jet Energy Scale
As discussed in Section 3.1.5, the fraction of hadronic and electromagnetic content in a hadron shower is not known precisely. Therefore, the energy measurement of the calorimeter cells has to be calibrated to the hadronic energy scale. The Jet Energy Scale (JES) is derived from test-beam data, LHC collision data and simulated events. Since an accurate hadron shower simulation is challenging, an uncertainty on the order of 2% to 6% in the central detector region, depending on the jet p_{T} and η, is applied. An additional uncertainty due to pile-up can be as large as 3% in the central detector region [69]. In the analysis, the uncertainty is included by an up- and down scaling of the reconstructed jet energy. Since the jet energy is correlated with the measurement of the missing transverse energy, an adjustment also has to be propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation.

Jet Energy Resolution
The Jet-Energy Resolution (JER) can be overestimated in the MC simulations compared to the measured resolution in collision data. Therefore, the JER is measured in di-jet events in data. It agrees with the predictions of simulated events at the level of 10% [70]. The uncertainty on the JER is evaluated by smearing jets according to the systematic uncertainty on the resolution measurement.

Jet Reconstruction Efficiency
The calorimeter Jet Reconstruction Efficiency (JRE) can be measured in minimum bias and QCD di-jet events. In the comparison of collision data and simulated events an uncertainty of 2% on the reconstruction efficiency on jets with a $\mathrm{p_T}$ <30 GeV is found. For jets with a larger p_{T} the effect is negligible [70]. The impact of jet reconstruction efficiency is evaluated in the analysis by randomly dropping jets from events.

$\boldsymbol{E}_{\mathbf{T}}^{\mathbf{miss}}$
The systematic uncertainties from the energy scale and energy resolution of leptons and jets are already propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction. Calorimeter cells not associated to any jet, so called "cell-out" terms, are considered in the $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty. Furthermore, the presence of soft jets, defined by a p_{T} larger than 7 GeV and smaller than 20 GeV, contribute to an uncertainty of the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. The uncertainty is calculated by variations of the soft jets and cell-out terms in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation [71]. The cell-out and soft-jet uncertainties are treated as fully correlated.

$\boldsymbol{E}_{\mathbf{T}}^{\mathbf{miss}}$ Pile-Up
The modeling of the pile-up contribution in the simulation leads to additional uncer-

tainties in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. A flat 10% uncertainty on the soft-jet and cell-out terms in the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is applied to take this effect into account [68].

Jet Vertex Fraction Scale Factor
A cut on the jet vertex fraction (JVF) is applied in the selection of the jets to reduce pile-up influences. JVF scale factors correct for induced disagreements, comparing simulated events with the reconstructed data. The uncertainty on these scaling factors leads to a systematic uncertainty in the shape of the simulated distributions [68].

Lepton Reconstruction Efficiency
Errors in the modeling of lepton trigger, reconstruction and selection efficiencies are parametrized by scale factors, determined by efficiency measurements in data. The uncertainty on these scale factors is on the order of 1% for both lepton flavors [39,41]. It is applied by a reweighting of the selected events according to the uncertainty of the scale factors.

Lepton Resolution and Energy Scale
Scale factors are applied to model the lepton resolution and the lepton-energy scale in the simulation compared to the observed data. These scale factors are determined from measurement of $Z \rightarrow ll$ processes [39, 72]. To model the energy scale uncertainty, the transverse lepton momentum is varied up and down by 1σ and re-applying the event selection. The lepton energy resolution is parametrized by a smearing of the lepton energy with a random number from a Gaussian distribution.

b-Tag Efficiency
Different b-tagging efficiencies, observed in data and simulated events, are corrected by b-tagging scale factors. These scale factors have an uncertainty of 10% to 20% for b-jets and 20% to 40% for mistagged light and c-jets. The uncertainty depends on the p_{T} and the η of the jet [73–75]. To evaluate the uncertainties in the analysis, the scaling factors of each jet are varied within their uncertainties.

Multi-Jet Background Normalization
The multi-jet background is estimated data-driven. A systematic uncertainty of 50% on the normalization is assigned based on the comparison of different estimation methods and the comparison of different pile-up conditions.

Luminosity
The ATLAS detector has different detectors to determine the luminosity, described in Section 2.2.4. The measurement of luminosity is based on dedicated "van der Meer"

scans in combination with calorimeter-based techniques during the data-taking period for 2011 data. Therefore, the determination of the integrated luminosity has an uncertainty of 3.7% [76]. Since all simulated processes are scaled to the integrated luminosity, this uncertainty affects both, the normalization of signal and background contributions.

MC Generator
The uncertainty on the generator is calculated for the $t\bar{t}$ signal process. Different MC generators are compared studying the acceptance and the impact on the predicted shape of the process. The $t\bar{t}$ events, generated with MC@NLO+Herwig, are compared to Powheg+Herwig. In addition to the uncertainty of the top-pair production cross-section, determined in NNLO calculations [15], an 7% uncertainty on the $t\bar{t}$ rate is found [68].

Parton Shower and Fragmentation
The correct parametrization of the parton shower is crucial for the generation of simulated pp collision events. The effect of the parton shower modeling is tested by comparing the predictions of two different models used for the parton shower. The $t\bar{t}$ sample generated by Powheg+Pythia is compared to Powheg+Herwig. Furthermore, the parameter sensitive to the parton shower activity is varied in a set of AcerMC+Pythia $t\bar{t}$ samples to obtain samples with more and less initial and final state radiation [77].

Parton Distribution Functions
The parametrization of the PDF leads also to an uncertainty in the simulated processes. It is estimated by taking the envelope of different PDF sets following the recommendations described in [78]. Technically, the simulated events are reweighted according to an alternative PDF parametrisation. The largest deviation from the default PDF is defined as PDF uncertainty.

***W*+jets Shape**
A variation in the W+jets shape is observed by the variation of parameters in the Alpgen event generator. The "iqopt3" parameter varies the functional form of the factorization and renormalization scale. The minimum p_{T} of the parton is changed to 10 GeV, compared to the nominal value of 15 GeV, by the "ptjmin10" variation. Since the W+jets background is normalized data-driven, for both effects only the shape variation is considered [68].

The influence of the systematic uncertainties can be seen in the next chapters, comparing the distributions of observables in recorded data and simulated events. An

improvement in the estimation of this uncertainties will allow one to rise the sensitivity of the physics analysis.

Chapter 5

W+Jets Background Estimation using Z+Jets Events

The contribution from W boson events associated with the production of jets is a major background in the analysis of semileptonically decaying top-quark pairs. The decay of a real W boson in combination with additional jets, some of them originating from heavy-flavor quarks, produces an event signature which is almost indistinguishable from the signal event signature. Since the fraction of W+jets background events in the top-quark selection is substantial, a accurate description of the W+jets background in the quantity to be measured of interest is indispensable. A data-driven estimation of W+jets events is difficult, since $t\bar{t}$ production is W+4 jets background and vice versa. As already discussed in Chapter 4.4 also the simulation of W+4 jets events has large uncertainties. The jets generated in addition to the W boson production process originate from initial and final state radiation, which is estimated in perturbation theory calculations. Therefore, the W+jets event properties and especially the W+jets cross-section are only known with limited precision for the W boson production with additional jets.

Chapter 4.4 presents the current ATLAS recommendation to determine the W+4 jets background data-driven. This chapter introduces two alternative approaches to estimate the W+4 jets background. Since the production of Z+jets events has many similarities compared to the W+jets production, in both methods the Z-boson is involved. The commonalities of W and Z boson events are addressed in the first section of this chapter. A conversion method, introduced in the second section, allows a generation of a data-driven template of W+jets events by means of Z+jets events. The last section introduces a method to determine the overall W+4 jets normalization from data, based on the ratio of reconstructed W and Z boson events as a function of the jet multiplicity.

5.1 Comparison of W and Z Boson Events

The vector bosons of the weak interaction mainly differ in terms of their production cross-sections, masses and the particles involved in their productions and decays. However, the production processes of W+jets and Z+jets events have many similarities which are addressed in this section. In particular, the production of associated jets is almost identical.

The Z-boson is an interesting candidate to analyze the W+jets production. In contrast to the W+jets production, the two charged leptons of the leptonic Z-boson decay lead to a distinct signature in the event. Requiring two oppositely charged leptons with an invariant mass near the Z-boson mass are hard kinematic constraints on the event. Hence, Z+jets events can be selected with a high purity. This section motivates the adoption of Z+jets events to investigate the W+jets background in the top-quark analysis.

Both W and Z bosons are produced in Drell-Yan processes. The production processes differ in terms of the input particles. While the Z-boson is produced via $q\bar{q}$ annihilation of same flavour, the production of the W boson requires an up-type and a down-type quark/anti-quark combination. Furthermore, for the production of Z bosons, higher energies are necessary since the mass of the Z-boson is about 10 GeV [1] higher. Apart from that, the production processes of both vector bosons are very similar. In Figure 5.1, Feynman diagrams of weak vector-boson production with zero, one, and two additional partons are shown. The vector boson can be characterized by a W or Z boson, depending on input and output particles. Any higher parton multiplicity is reached by additional gluon radiation and gluon splitting processes. These processes are identical in the production of both W and Z bosons.

Vector Boson Couplings As shown in Figure 5.1, W and Z bosons are produced via $q\bar{q}$ annihilation in Drell-Yan processes. The leading order cross-section for on-shell W and Z boson production, neglecting the gauge boson decay width, is shown in Equations 5.1 and 5.2 [79], respectively.

$$\sigma^{q\bar{q}'\to W} = \frac{\pi}{3}\sqrt{2}G_F M_W^2 |V_{q\bar{q}'}|^2 \delta(s - M_W^2) \tag{5.1}$$

$$\sigma^{q\bar{q}\to Z} = \frac{\pi}{3}\sqrt{2}G_F M_Z^2 (c_V^2 + c_A^2)\delta(s - M_Z^2) \tag{5.2}$$

The similar formulas for the production cross-sections imply that the production mechanisms are related. Only the coupling strengths to the fermions are different. While the cross-section of W boson production depends on the squared CKM matrix element $|V_{q\bar{q}'}|$, the production cross-section of Z bosons depends on a composition of

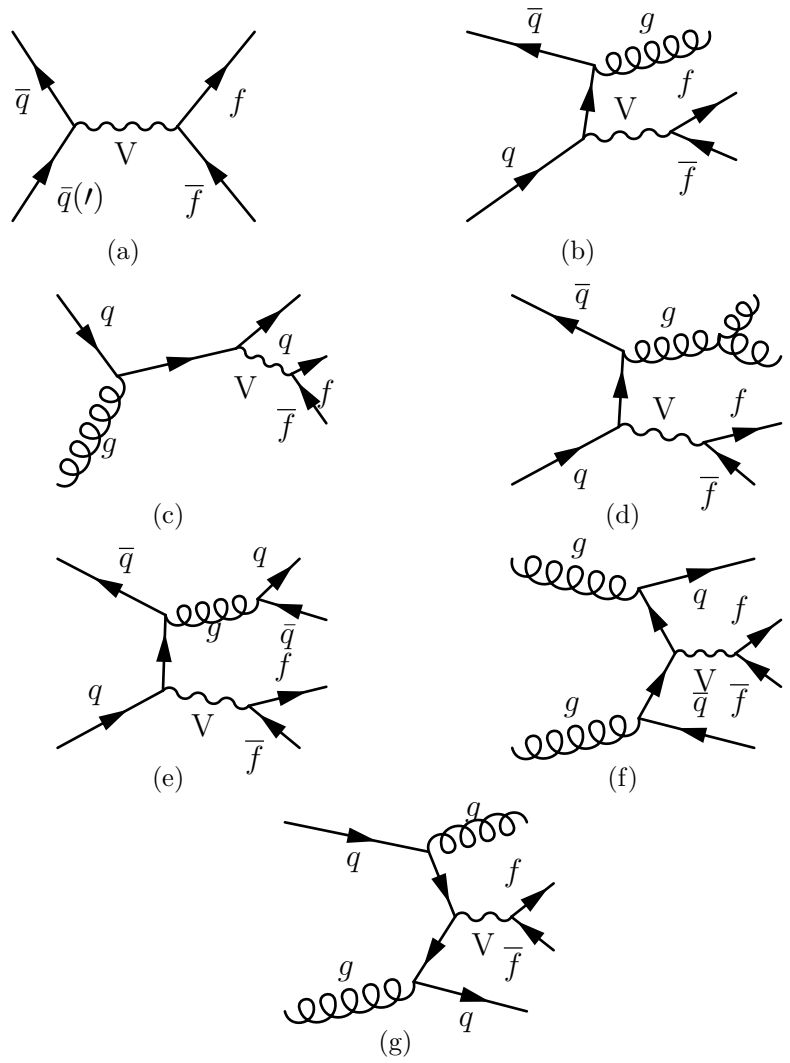


Figure 5.1: *Feynman diagrams of W/Z-boson production processes. The Drell-Yan process is shown in Figure (a). Figures (b), (c) and (d), (e), (f), (g) show the production with one and two additional partons in the final state, respectively.*

vector c_V and axial-vector c_A fermion couplings. The fermion couplings are defined as [79]:

$$c_V = I_{3L}^f + I_{3R}^f - 2e_f sin^2\Theta_W \tag{5.3}$$

$$c_A = I_{3L}^f - I_{3R}^f \tag{5.4}$$

The factor $I_{3L}^f(I_{3R}^f)$ describes the left (and right) handed third iso-spin component of the fermion, e_f is the fermion charge and Θ_W the electroweak mixing angle. In addition to a different cross-section, the small variations of the couplings lead

Process	Production Cross-Section [pb]
W^+	6198
W^-	4216
Z^0	974

Table 5.1: *Cross-sections of the W- and Z-boson production for LHC proton-proton collisions at a center-of-mass energy of* 7 TeV *calculated by ZWPROD [65].*

to slightly different angular distributions of the particles in the production and decay processes [80]. Table 5.1 shows the cross-sections for the W and Z boson production. Due to the different couplings, different particle accessibility in the production process, and a slightly higher Z-boson mass, the production cross-section of Z-boson events in pp collisions is decreased by a factor of ten.

Initial and Final State Objects Z bosons in pp collision are mainly produced by $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, $c\bar{c}$, $u\bar{s}$ and $d\bar{c}$ quark pairs. The W^+ is mainly produced from $u\bar{d}$ and the W^- from $d\bar{u}$. The PDFs for these quarks, shown in 3.14 on page 51, are not identical. While the PDFs for the sea quarks are comparable, the valence quark PDFs differ. Hence, different momentum fractions of the quarks in the protons affect the rapidity distributions of the bosons [80]. Furthermore, the different quark content in the W and Z boson production leads to a disparity in the remaining parton flavor content. For instance, a boson production in association with two bottom-quarks in the final state, possible with the Feynman diagram shown in Figure 5.1(f), only exists in association with a Z boson. While this difference in the heavy-flavor content is negligible in an analysis without a b-jet requirement, it has to be investigated in the context of a b-tag requirement in the event selection.

The difference of the boson mass leads to a larger phase space of the Z-boson decay products. The associated partons in the final state have, on average, a higher transverse momentum. Therefore, the relative number of events without additional hard parton radiation ($p_T > 25$ GeV) is reduced for W boson production compared to Z-boson production. Hence, the average number of selected jets in Z-boson events is higher. Apart from that, the parton related differences in W and Z boson events are compensated almost by QCD effects. As soon as additional hard partons exist, the QCD effects show almost no difference between W and Z boson events. Analyzing the ratio of W and Z boson production with at least one additional jet, one observes a constant ratio as a function of the jet multiplicity [80].

Both W and Z bosons decay into a pair of fermions. The branching ratio for the W and Z bosons into the different decay channels is shown in Table 5.2. Within the leptonic decay of W and Z bosons analyzed, both W and Z bosons decay into a pair of leptons. The sole difference is the detector reconstruction efficiency of the decay products, which differs strongly. While in case of the Z-boson decay both

leptons are charged, the W boson decay contains one neutrino. The reconstruction of the charged leptons originating in the W and Z boson decay is identical for both. In contrast, the reconstruction of the neutrino can only be done by means of the missing transverse energy in the event. The resulting differences in the efficiency in the reconstruction of the decay products have to be considered. As shown in this chapter, these differences can be modeled and, therefore, nearly abrogated.

Comparison of Event Properties

The expected differences in the angular distribution of the vector bosons and their decay products are shown in Figure 5.2. In Figures 5.2(a) and 5.2(b) the η and p_{T} of the reconstructed[1)] W and Z boson bosons are shown, based on MC truth information. The η distributions of W and Z boson bosons show only small differences on generator level, shown in [82]. The disagreement observed in Figure 5.2(a) is mainly linked to the reconstruction efficiencies of the objects which are included in the distributions shown. While leptons, for instance, have to be reconstructed in the inner detector acceptance, such a demand is not made for the neutrino in this selection. Therefore, the pseudo-rapidity distribution is more central for Z+jets events compared to W boson events. Due to the slightly smaller W boson mass, the transverse momentum of the boson, shown in Figure 5.2(b), is lower in W+jets events compared to Z-boson events.

The remaining graphs of Figure 5.2 show distributions of jet-related event properties. The p_{T} and the η of the jet with the highes transverse momentum are depicted. In addition, the separation and the angular difference of the two most energetic jets are shown. All jet-related variables show almost no difference in the comparison of W and Z boson events.

[1)] Detailed information on the reconstruction of W and Z boson events can be found in Chapter 5.3.

Decay Channel			Branching Ratio
$W^{\pm}$	$\rightarrow$	$e^{\pm}\nu_e$	$(10.75 \pm 0.13)\%$
$W^{\pm}$	$\rightarrow$	$\mu^{\pm}\nu_{\mu}$	$(10.57 \pm 0.15)\%$
$W^{\pm}$	$\rightarrow$	$\tau^{\pm}\nu_{\tau}$	$(11.25 \pm 0.20)\%$
$W^{\pm}$	$\rightarrow$	$q\bar{q}'$	$(67.60 \pm 0.27)\%$
Z^0	$\rightarrow$	e^+e^-	$(3.363 \pm 0.004)\%$
Z^0	$\rightarrow$	$\mu^+\mu^-$	$(3.366 \pm 0.007)\%$
Z^0	$\rightarrow$	$\tau^+\tau^-$	$(3.370 \pm 0.008)\%$
Z^0	$\rightarrow$	$q\bar{q}$	$(69.91 \pm 0.06)\%$
Z^0	$\rightarrow$	invisible	$(20.00 \pm 0.06)\%$

Table 5.2: *Branching ratios of the W- and Z-boson [1].*

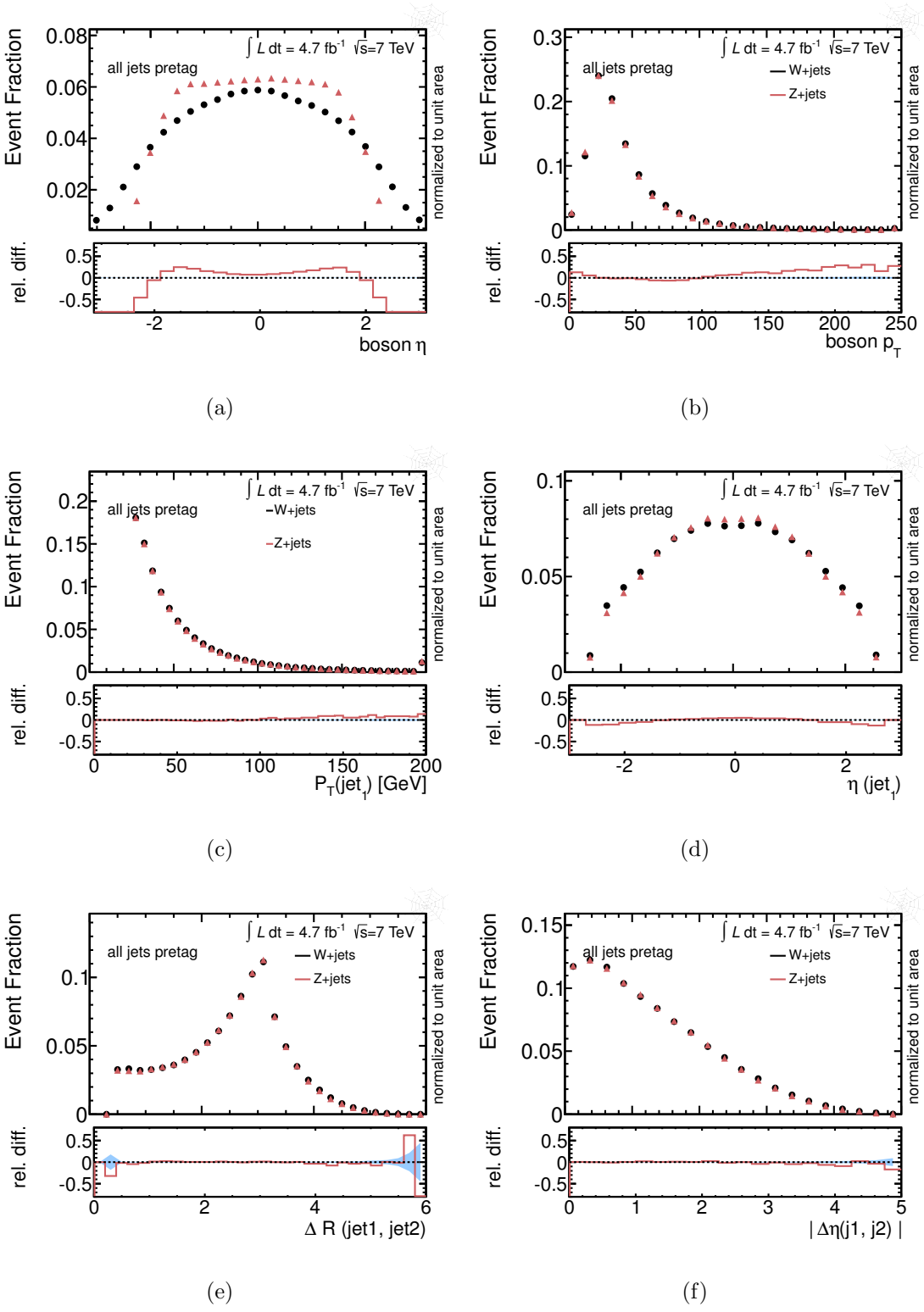


Figure 5.2: *Angular distributions of W and Z boson events derived from MC truth information. The the pseudo-rapidity (a) and the p_{T} distribution (b) of the vector bosons, determined in MC truth information, are shown. Furthermore, the p_{T} (c) and the η (d) of the highest energetic jet and the angular separations in ΔR (e) and $\Delta\eta$ (f) of the two most energetic jets are depicted. In addition to the histogram shape, also the relative difference of Z- divided by W-events is displayed.*

Commonalities and differences of the W and Z boson production are discussed in more detail in [80]. As shown in the next sections, all W and Z boson differences can either be corrected or are negligibly small in the application of Z+jets events to analyze W+jets events. In summary, it can be stated that Z+jets events are well suited to investigate W+jets events.

5.2 W+jets Modeling using Z+jets Events

The production of a Z-boson associated with jets is a process with many similar properties compared to the W+jets production. In this section, a 'conversion' algorithm is introduced. It allows one to model W+jets events by means of Z+jets events. Since Z-boson events can be extracted from data in a very pure sample, a data-driven template of W+jets events based on Z+jets events can be obtained. Thus, the theoretical modeling of W+jets events is replaced by the event modeling from data. In particular, systematic uncertainties on the event generation and detector effects are largely eliminated.

5.2.1 Conversion of Z+jets Events

Since we know that the production of jets is very similar in W+jets and Z+jets events, the conversion algorithm addresses only the decay products of the weak vector boson.

- The difference in the boson mass is corrected. This is achieved by the momenta of both leptons from the Z-boson decay are scaled by the relative mass difference:

$$\vec{p}_{l_{1,2}}^{\,scaled} = \vec{p}_{l_{1,2}} \cdot \frac{m_W}{m_Z} \tag{5.5}$$

  The energy by which the leptons are reduced is removed from the event.

- Both W and Z bosons decay into two leptons. In the case of the W boson decay, one of the leptons is a neutrino. Therefore, the difference in the detector response is corrected. The algorithm chooses one of the leptons l_i randomly as fake neutrino candidate. The random choice prevents a bias in the event properties, which would be present e.g. with a p_{T} ordered choice. The transverse momentum of lepton l_i is added to the missing transverse energy of the event:

$$\vec{E}_T^{\,\mathrm{miss}\prime} = \vec{E}_{\mathrm{T}}^{\,\mathrm{miss}} + \begin{pmatrix} p_{l_i,x}^{\mathrm{scaled}} \cdot w_x \\ p_{l_i,y}^{\mathrm{scaled}} \cdot w_y \\ 0 \end{pmatrix} \tag{5.6}$$

The factors $w_{x,y}$ refer to the contribution of the lepton l_i in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. Since the $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated from the reconstructed objects, described in 3.1.6, each object has a $E_{\mathrm{T}}^{\mathrm{miss}}$ weight with regarding to its object type and its location in the detector.

- Finally, the lepton is discarded from the event.

Apart from the adoption of the whole jet distribution of the Z-boson event, also pile-up effects in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation are automatically included. In the case of a Z-boson decay, the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event is expected to be very small or zero. The amount of measured $E_{\mathrm{T}}^{\mathrm{miss}}$ mainly results from mismeasurements in the calorimeter system. It is identical to the instrumental fraction of $E_{\mathrm{T}}^{\mathrm{miss}}$ present in the W+jets events. Therefore, the instrumental fraction of $E_{\mathrm{T}}^{\mathrm{miss}}$ is also modeled from data using this approach.

To provide an accurate W+jets background description, the event selection of Z+jets candidates is based upon the same object definitions and nearly identical event-selection cuts as used in the top-quark selection, introduced in Chapter 4.1. The requirement for missing transverse energy in the event and a consistency of the transverse W boson mass is replaced by the demand of two oppositely charged leptons with a reconstructed invariant mass $81 < \mathrm{m_{ll}} < 101$ GeV.

After the conversion algorithm is applied, the events have to pass the $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\mathrm{m_T}(W)$ requirements of the top-quark selection. Hence, the kinematics of the W+jets and converted Z+jets events are equally defined. In Table 5.3, the number of resulting and converted Z+jets events with four reconstructed jets is shown. The efficiency of the conversion algorithm including the additional cuts on $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\mathrm{m_T}(W)$ is roughly 75%. The expected number of W+jets background events in the semileptonic $t\bar{t}$ analysis for instance is approximately twenty times higher. However, it is not possible to increase the statistics of the Z+jets sample by a reduction of the object requirements for the second lepton. Even if it would be possible in terms of the purity of the Z-boson sample, the conversion method would be biased, since the remaining lepton in the modeled W+jets event has to fulfill the tight object requirements. Therefore, the lepton which is selected with reduced object requirements would have to be used as fake neutrino candidate. It has been shown that soften object requirements are often related to a reduced transverse object momentum [68]. Hence, the conversion method would be biased.

Figure 5.3 shows a set of control distributions in which event properties of W+jets and converted Z-boson events are compared. The distributions are chosen such, that the modeling of jets and the W boson decay can be investigated. The properties of the W+jets events are based on simulation. The properties of converted Z-boson events, referred to as $Z to W$ events, are based on Z+jets events from data and simulation, respectively. The event shape observable aplanarity [81] measures the

Jet Bin	Selected Z Boson Events	Converted Z Boson Events
4	768 ± 28	615 ± 25

Table 5.3: *Number of selected Z+jets events in data shown before and after the conversion into W+jets events. Additional kinematic cuts on $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\mathrm{m_T}(W)$ on the converted events are applied. The numbers refer to the electron selection and an integrated luminosity of* $4.7\ \mathrm{fb}^{-1}$.

transverse momentum component out of the event plane. It is defined as:

$$A = \frac{3}{2}\lambda_3 \tag{5.7}$$

where λ_3 is the third largest eigenvalue of the sphericity tensor [81]. The sphericity tensor is defined as:

$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i \|p_i\|^2} \tag{5.8}$$

where α and β correspond to the x, y and z components and p_i is the momentum of object i in the event.

An aplanarity value of zero refers to a planar event while a value of $1/2$ corresponds to an isotropic event. Apart from the statistical fluctuations in the tail region (> 0.12), the aplanarity in converted Z+jets events from simulation and data is compatible to the one of the W+jets prediction within 10%. As expected, also the distributions of the jet kinematics in W+jets events, represented in Figure 5.3(b) by the p_{T} of the third leading jet, are well described by Z+jets events.

The distribution of the transverse momentum of the lepton for converted Z-boson events, shown in Figure 5.3(c), depends on the conversion algorithm, since it is scaled by the relative boson mass difference. Within the statistical uncertainties, the distributions show a reasonable agreement when compared to the W+jets simulation. Also, the missing transverse energy and the transverse W boson mass distributions depend on the conversion algorithm for $ZtoW$ events. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution of converted Z-boson events indicates a slight shift and, therefore, a disagreement which is observable especially for low $E_{\mathrm{T}}^{\mathrm{miss}}$ values. This shift, expected to be correlated with the different W and Z boson couplings, could be compensated by an additional angular correction of the Z-boson decay products in the conversion algorithm. Apart from that, both distributions are modeled suitable within an uncertainty of 20%. Even more complex observables are in good agreement. The distribution of the invariant $t\bar{t}$ mass in the W+jets background estimation, for instance, is of particular interest in the top-quark analysis. It is composed of the four leading jets, the lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$. Therefore, it is dependent on many event properties. Also, in this particular case, the modeling with converted Z-boson events agrees very well with simulated W boson events.

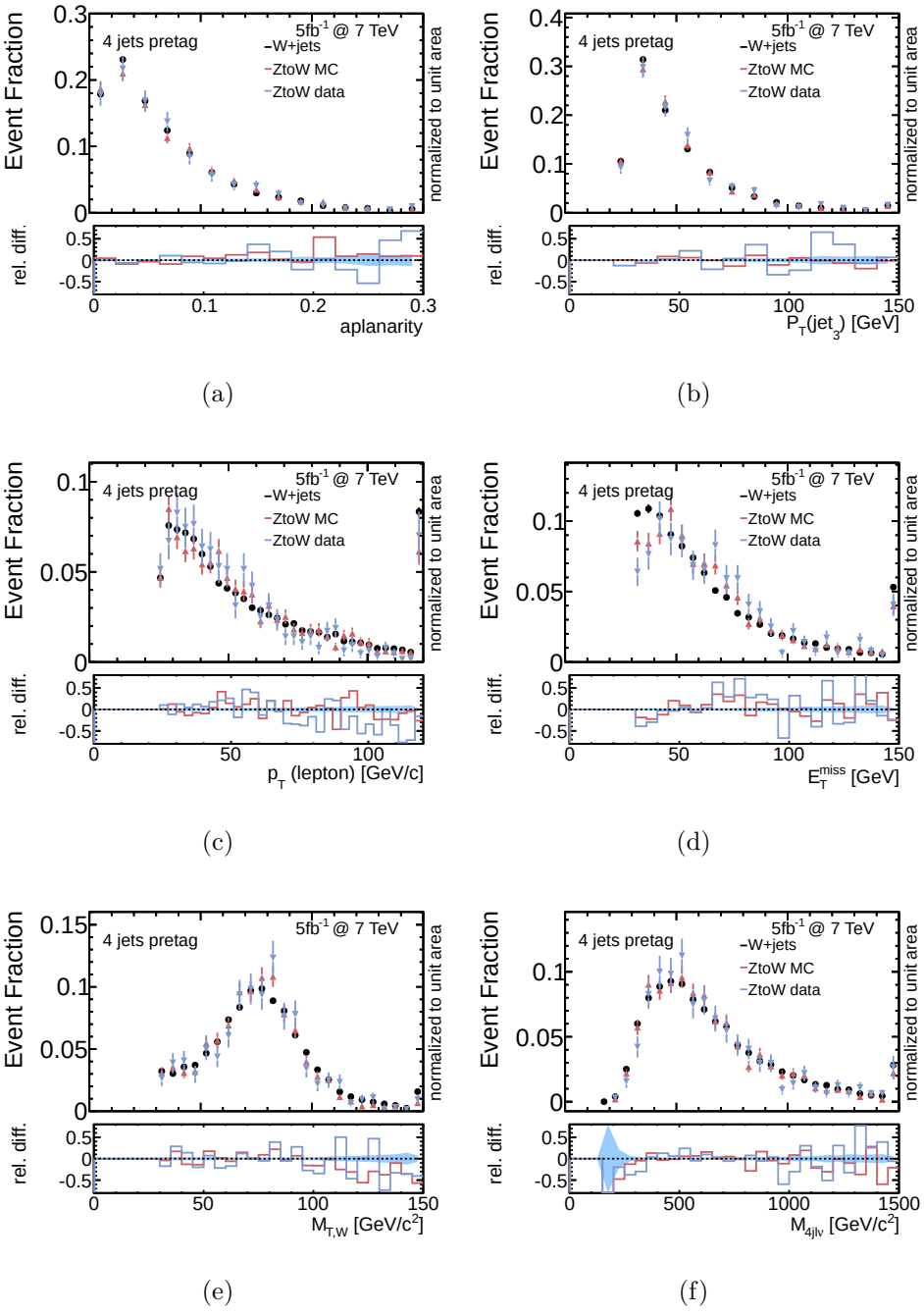


Figure 5.3: *Comparison of simulated W+jets events to converted Z+jets events from simulation and data for the electron channel. Shown are the aplanarity of the event (a), the p_T of the third leading jet (b), the p_T of the lepton (c), the missing transverse energy (d), the transverse W boson mass (e) and the invariant mass (f) of the composition of the four leading jets, the lepton and the E_T^{miss}. At the bottom of each histogram, the relative difference, compared to the W+jets simulation, is shown.*

A validation of the conversion algorithm with a reduced statistical uncertainty is shown for the two-jet multiplicity in Figure 6.7 in the appendix. Since the agreement in the muon channel is compatible, the conversion of Z-boson events with a muon in the final-state is not explicitly shown. A more detailed validation of the W+jets template derived from Z-boson events, including additional observables and studies on the different heavy-flavor content, can be found in [82].

5.2.2 Conclusion

In this section, a conversion method to produce a W+jets template from Z+jets events has been developed. The agreement of converted Z-boson events compared to simulated W+jets events is very good, even if the conversion algorithm is kept as simple as possible. Since the method is applied on Z+jets events from data, the resulting W+jets template is not conditioned of systematic uncertainties on the detector-response modeling.

Further corrections to model the small differences observed in W and Z boson events might lead to an even better description. The differences caused by the varying heavy-flavor content of W and Z boson events are avoided by performing the analysis without applying a b-tag. However, it has been shown that these differences are very small [82]. An estimate of the W+jets normalization, based on a fit of the introduces W+jets template, is shown in [82]. The low statistics of Z+jets events leads to a statistical uncertainty in the analyzed set of data. Since the event properties do not depend on the lepton flavor, it is possible to increase the statistic by a factor of two, merging the electron and the muon samples. Furthermore, this uncertainty will be reduced with the analysis of the ATLAS data collected in the year 2012.

5.3 W+jets Normalization using the Ratio of W and Z+Njet Events

The theoretical calculation of the normalization of W boson production in association with at least four jets, discussed in Chapter 4.4, has a large uncertainty. Hence, a data-driven estimation of the W+jets background normalization in the lepton plus jets selection is necessary. An application of data-driven methods is difficult, since both $t\bar{t}$ and W+jets events have an almost identical event signature.

The recommended method to estimate the W+jets normalization in the top-quark analysis, introduced in Chapter 4.4.1, is based on the charge asymmetry in the W boson production. It compares the reconstructed charge asymmetry in data with the predicted one from MC simulations. The uncertainties on the theoretical calculations are predicted to be within 3% [67]. Applied on data, additional uncertainties, for

instance on the PDFs, lepton-charge identification and the estimation of the W+jets heavy-flavor content, have to be considered.

In this section an alternative data-driven approach to estimate the number of W+4 jets background events in the semileptonic $t\bar{t}$ analysis is introduced. The presented N_{jet} ratio method is based on the ratio of W and Z boson events, evaluated in terms of their dependence on the number of jets. Similar to the estimation of the W+jets template, this method benefits from the similarity of the production processes in W and Z boson events.

5.3.1 Motivation

As discussed in 5.1, in particular the production mechanism of jets associated with vector-boson production, is identical for W and Z boson events. The fraction of events with a given number of jets

$$f_n(V) = \frac{\sigma(V + n \text{ jets})}{\sum_m \sigma(V + m \text{ jets})}, \tag{5.9}$$

where V is either a Z or W boson, shows a similar behavior for W and Z boson production in the theory prediction [83]. Therefore, the cross-section ratio

$$r = \frac{\sigma(W + n \text{ partons})}{\sigma(Z + n \text{ partons})} \tag{5.10}$$

is expected to be independent of the parton multiplicity apart from $\mathcal{O}(2\%)$ deviations. A detailed analysis can be found in [80].

Based on this assumption, also the ratio of selected W and Z boson events

$$R_n = \frac{W + n \text{ jets}}{Z + n \text{ jets}} \tag{5.11}$$

is expected to be flat as a function of the jets multiplicity, if at least one jet is selected. The ratio of events with zero selected jets is excluded, since a difference compared to the other jet multiplicities is expected, discussed in Chapter 5.1.

The ratio R_n can be measured in data. While the number of Z+jets events can be determined with a high purity in all jet multiplicities, the dominant background process in the W+jets selection, requiring a larger number of jets, comes from $t\bar{t}$ production. To be almost independent from the $t\bar{t}$ background, and based on the linear hypothesis, the ratio is measured in data for $1 \leq n \leq 3$ jets and extrapolated to the four-jet ratio R_4^{extrap}. The number of W+4 jets events is calculated by the equation:

$$N^{\text{estimation}}_{W+4 \text{ jet}} = R_4^{\text{extrap}} \cdot N^{\text{data}}_{Z+4 \text{ jet}} \tag{5.12}$$

where N^{data}_{Z+4jet} is the number of recorded Z+4 jets events in data.

Both terms R_4^{extrap} and N^{data}_{Z+4jet} largely base on data. Systematic uncertainties arise from the background contributions in the W and Z boson selection, which are evaluated from MC simulations and a data-driven approach. With the construction of the W and Z boson ratio some of the systematic effects cancel out each other. For both reasons, it is expected that the systematic uncertainty on the W+jets normalization can be reduced compared to the uncertainty of the currently used data-based prediction. Furthermore, with the exclusion of events with four selected jets in the determination of R_4^{extrap}, the analysis is performed exclusively in the side-band region of the top-quark analysis. The result is independent from the dataset, used in the analysis of top-quark pair production.

5.3.2 Estimation of the W and Z Boson Ratio

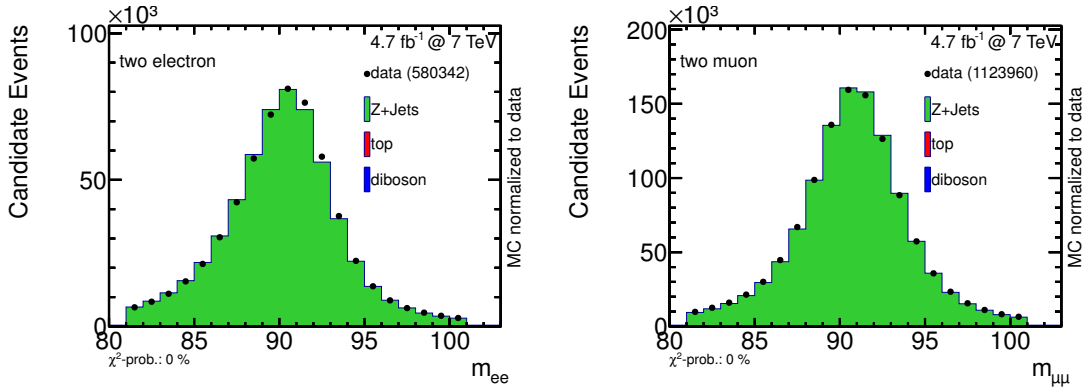
In this section, the event selection and background estimation to determine a data-driven ratio R_n are presented. The goal is to achieve a consistent prediction of the normalization for the W+jets background in the $t\bar{t}$ selection. Therefore, the required objects and event-selection cuts follow the criteria of the top-quark analysis, defined in Chapter 4.

Selection of Events

The selection of W+jets events is identical to the top-quark selection. In the case of the Z+jets selection, the same object definitions and almost identical event-selection cuts are applied. This requirement enables a reduction of systematic effects, since many systematic uncertainties cancel out each other in the constructed ratio.

Within the Z+jets selection the cut on the missing transverse energy and the transverse W boson mass is replaced by a requirement of two oppositely charged identically flavored leptons in a Z-boson mass window. This definition is identical to the requirement in Section 5.2. Hence, it leads to an almost background-free Z-boson selection. Figure 5.4 shows the distribution of the invariant di-lepton mass of events passing the Z-boson event selection in the electron and muon channels. As expected, an almost background-free Z+jets sample is obtained.

In both selections, events with at least one selected jet are chosen. Since differences in the behavior of the W and Z boson ratio are expected applying a b-tag requirement, no b-tag is requested in both selection schemes. The event-selection requirements for both selection schemes are depicted in Table 5.4. The transverse momentum of the reconstructed jets is required to be larger than 25 GeV for both selections.



(a) (b)

Figure 5.4: *Invariant di-lepton mass for events passing the Z+jets selection, shown for two electrons (a) and two muons (b) normalized to data.*

	W Selection	Z Selection
	1 good lepton	2 good leptons
$\mathrm{E}^{\mathrm{T}}_{\mathrm{miss}}$	>30 GeV (e) >20 GeV (μ)	81< m_{ll} <101 GeV
$\mathrm{m_T}(W)$	>30 GeV (e) $\geq$60 - $\mathrm{E}^{\mathrm{T}}_{\mathrm{miss}}$ (μ)	opp. charged leptons

Table 5.4: *Comparison of W and Z boson selection cuts.*

Background Estimation

While the selection of Z+jets events can be performed nearly background-free, the W+jets selection is contaminated by background events. The relevant background processes are QCD multi-jet, top quark, Z+jets and di-boson production. In the subsequent discussion, the 'top-quark background' refers to a combination of $t\bar{t}$ and single top-quark processes. Background processes investigated in the Z+jets event selection are top-quark and di-boson production. The number of Z+jets and W+jets events is determined by subtracting the number of expected background events from the amount of selected events in data. While the amount of most of the background processes is taken from simulation, the contribution of the multi-jet background in the W+jets selection is determined from a data-driven template.

Estimation of the Multi Jet Background

The huge number of QCD multi-jet events in the pp collisions leads to a certain fraction of events, where a jet is misidentified as an electron or the identified lepton

originates from a heavy-flavor decay inside a jet. The fraction of multi-jet background in the di-lepton selection of Z+jets events is less than 1% [84] and, therefore, neglected in this analysis. In the case of the W+jets selection, the multi-jet background contribution is substantial.

Since the simulation of QCD multi-jet events has large uncertainties, the multi-jet background is determined using a data-driven method. A template of the multi-jet background is produced by the JetElectron model, which is described in Section 4.3.2. To obtain the fraction of QCD multi-jet background in each jet multiplicity, the multi-jet background template, in combination with the simulated signal and background processes, is fitted to data in the missing transverse energy distribution. The result of the fit, applied as described in Section 4.3.2, is shown in Figure 5.5 for the electron and Figure 5.6 for the muon channel, individually for the different jet multiplicities each. The left column shows the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution used for the multi-jet estimation. In the right column of Figure 5.5 (5.6), the $\mathrm{m_T}(W)$ distribution is shown. It is displayed as a control variable and indicate a good agreement of background expectations and data. The disagreement in the low $\mathrm{m_T}(W)$ region is associated with a slightly mismodeled lepton p_{T} and vanishes after the cut on the $E_{\mathrm{T}}^{\mathrm{miss}}$. The estimated multi-jet event fraction in the event selection is shown in Table 5.5 for the different numbers of jets.

N_{Jets}	Electron	Muon
1	$4.2 \pm 2.1\%$	$3.9 \pm 2.0\%$
2	$9.1 \pm 4.6\%$	$4.8 \pm 2.4\%$
3	$8.3 \pm 4.2\%$	$3.6 \pm 1.8\%$

Table 5.5: *Estimated fraction of the multi-jet background per jet multiplicity in the W+jets selection, including the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut. The uncertainty of the estimated background fraction is appreciated to be 50% [68].*

Event Yields

The analysis is based on the same samples as required by the top-quark analysis, presented in Section 4.2. The yields of simulated and reconstructed events passing the W+jets event selection can be found in Table 5.6 for the electron channel and Table 5.7 for the muon channel. The amount of multi-jet background events in the tables is scaled to the estimated fraction determined in the fit. Since no W+jets scaling factors are applied at this stage, the amount of W+jets events shown in Tables 5.6 and 5.7 is also scaled to the corresponding fit result. The adjustment of the W+jets contribution in the event yield tables is necessary to achieve a reasonable agreement in the number of recorded and simulated events. This adjustment is not applied in the further analysis and especially not included in the calculation of the

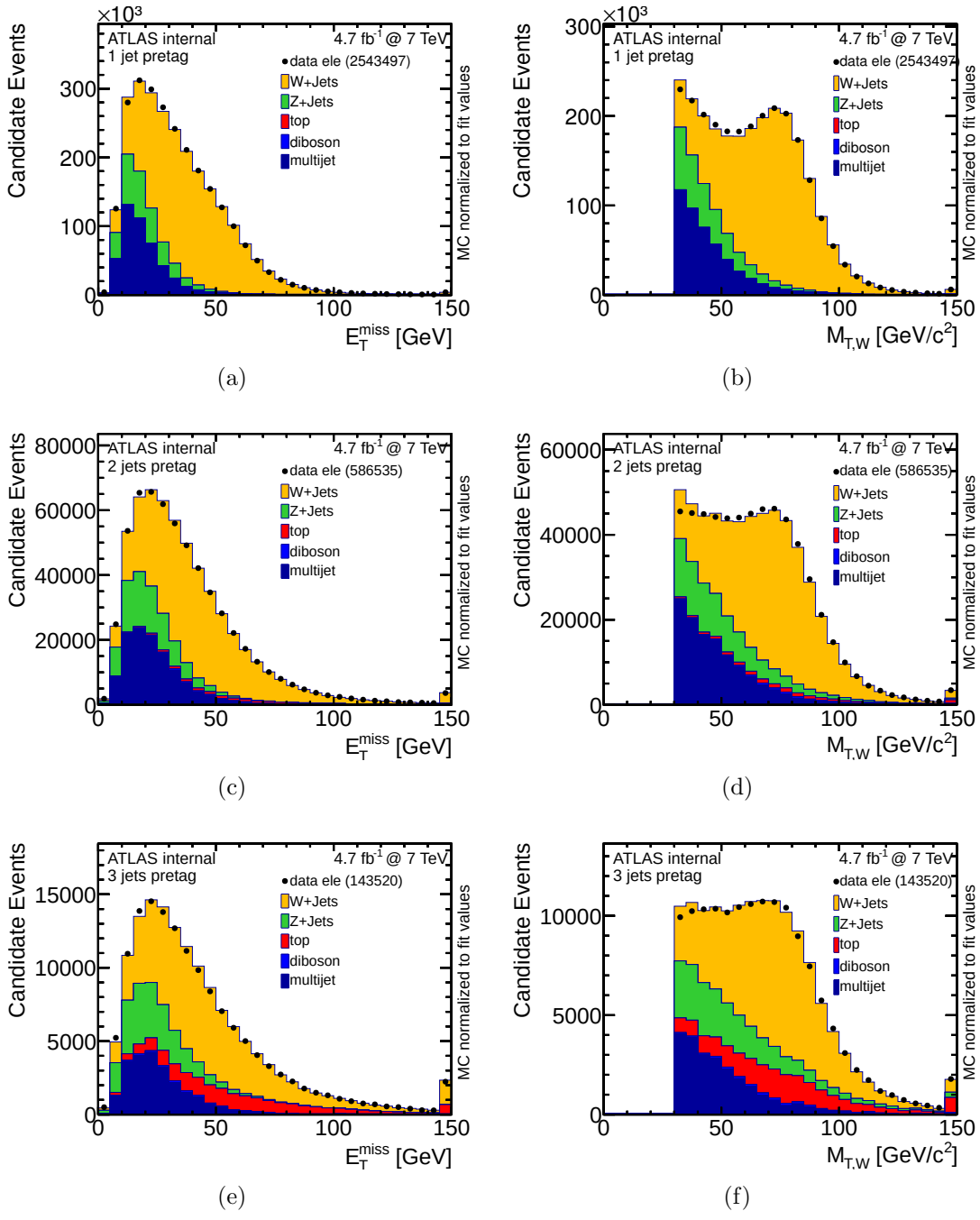


Figure 5.5: *Multi-jet background estimation of the electron channel. In the left column are the fit results in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in the one (d), two (e) and three (f) jet multiplicity are shown. The column on the right shows the data agreement of the fit result in the $\mathrm{m_T}(W)$ distribution.*

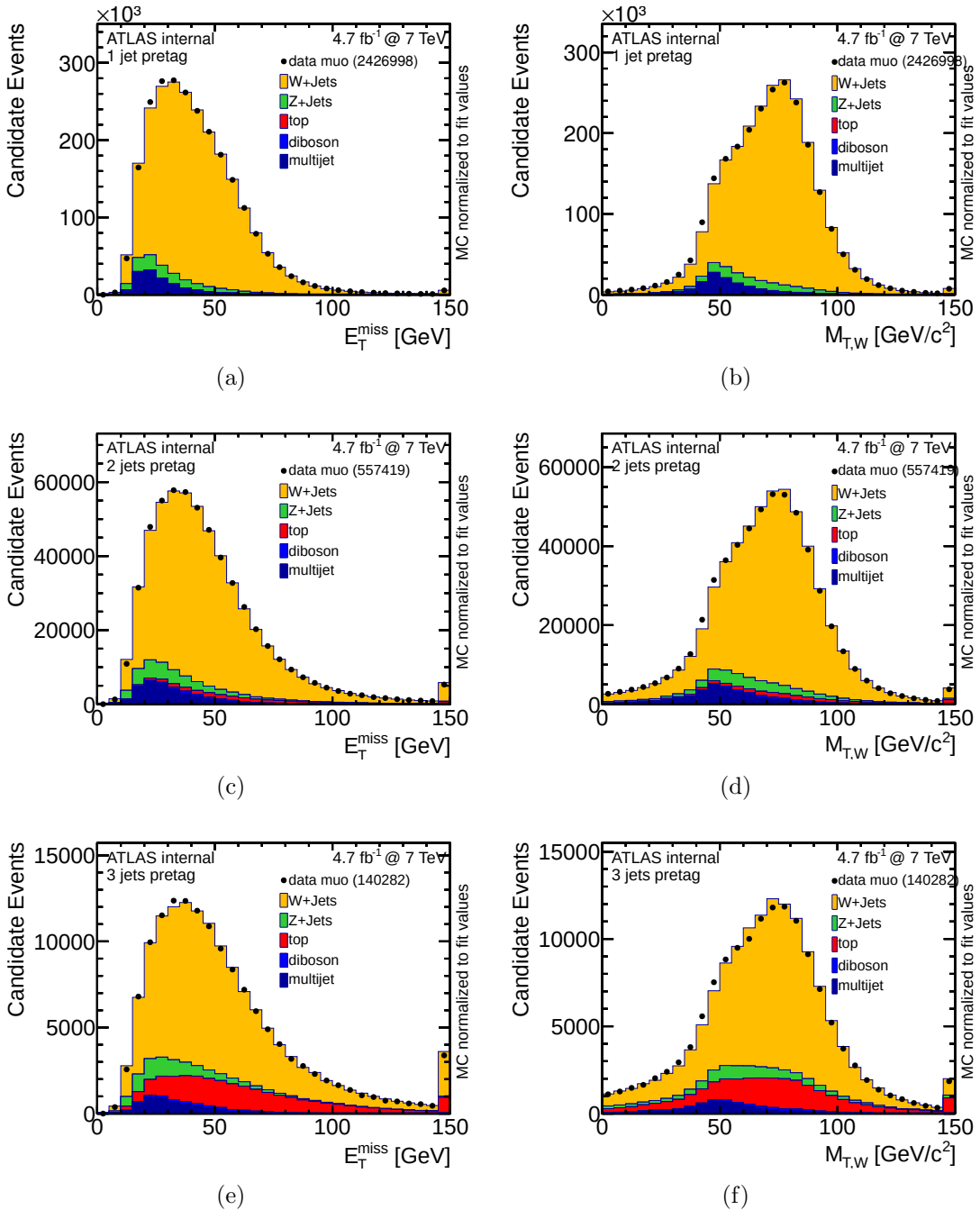


Figure 5.6: *Multi-jet background estimation of the muon channel. In the left column are the fit results in the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in the one (d), two (e) and three (f) jet multiplicity are shown. The column on the right shows the data agreement of the fit result in the $\mathrm{m_T}(W)$ distribution.*

ratio R_n, shown in this Chapter. The number of events passing the Z+jets selection cuts is shown in Tables 5.8 and 5.9. Since an adjustment of Z+jets normalization was not present, the disagreement of reconstructed and simulated events is larger as compared to the W+jets selection. As with the W+jets adjustment, the number of simulated W and Z boson events is not included in the further analysis.

Figure 5.7 shows the N-jet distribution for the W+jets event selection in the top and the Z+jets event selection in the bottom area. The electron channel is shown on the left, while the muon channel is shown on the right. It can be observed that the background processes in selected W+jets events are mainly coming from $t\bar{t}$, multi-jet, Z+jets and di-boson production. The fraction of $t\bar{t}$ background in the three-jet bin is already about 30%. For jet multiplicities n $\geq$ 4 it becomes the dominant W+jets background process. In comparison to the W+jets selection, the top-quark background in the Z+jets selection is much smaller and nearly constant, apart from a slight increase in the two-jet multiplicity.

The predicted fraction of Z+jets background in the W+jets selection is in the order of 5% to 10%, increasing with the number of jets. Since the Z+jets production is also a signal process in the calculation of the W and Z boson ratio and, therefore, determined in data, migration effects of misidentified Z+jets events in the W+jets selection have to be taken into account. In this analysis the number of misidentified Z+jets events in the W+jets selection is taken from MC simulations. To minimize the uncertainty of the migration effect, a data-driven estimation of the misidentified Z+jets contribution is desirable.

Process	1 Jet	2 Jet	3 Jet
top	4374 ± 66	10953 ± 105	15371 ± 124
W+jets	1148693 ± 1072	249951 ± 500	55845 ± 236
Z+jets	46080 ± 215	20253 ± 142	6791 ± 82
di-boson	4732 ± 69	3532 ± 59	986 ± 31
total MC	1203881 ± 1097	284689 ± 534	78995 ± 281
QCD	51985 ± 228	28559 ± 168	7003 ± 83
total exp.	1255866 ± 1121	313249 ± 560	85999 ± 293
data observed	1250173 ± 1118	313223 ± 560	84642 ± 291

Table 5.6: *Expected number and the statistical uncertainty of selected events per jet bin, passing the W+jets event selection, in the electron channel. The numbers correspond to an integrated luminosity of 4.7 fb^{-1}. The uncertainties correspond to the statistical uncertainty.*

The final numbers of W+jets and Z+jets events for each jet-multiplicity derived from data are summarized in Table 5.10. They are determined by the subtraction of the number of expected background processes from the total number of W and

Process	1 Jet	2 Jet	3 Jet
top	6287 ± 79	15537 ± 125	22300 ± 149
W+jets	2007989 ± 1417	438052 ± 662	95290 ± 309
Z+jets	97651 ± 312	27081 ± 165	7205 ± 85
di-boson	7461 ± 86	5645 ± 75	1521 ± 39
total MC	2119390 ± 1456	486316 ± 697	126318 ± 355
QCD	85215 ± 292	24885 ± 158	4647 ± 68
total exp.	2204606 ± 1485	511201 ± 715	130966 ± 362
data observed	2212427 ± 1487	513756 ± 717	130536 ± 361

Table 5.7: *Expected number and the statistical uncertainty of selected events per jet bin, passing the W+jets event selection, in the muon channel. The numbers correspond to an integrated luminosity of 4.7 fb^{-1}. The uncertainties correspond to the statistical uncertainty.*

Process	1 Jet	2 Jet	3 Jet	4 Jet
Z+jets	70654 ± 266	15866 ± 126	3463 ± 59	768 ± 28
di-boson	289 ± 17	222 ± 15	58 ± 8	10 ± 3
top	59 ± 8	136 ± 12	71 ± 8	22 ± 5
total MC	71005 ± 266	16225 ± 127	3594 ± 60	801 ± 28
data observed	76000 ± 276	16236 ± 127	3422 ± 58	718 ± 27

Table 5.8: *Expected number and the statistical uncertainty of selected events per jet bin, passing the Z+jets event selection, in the electron channel. The numbers correspond to an integrated luminosity of 4.7 fb^{-1}. The uncertainties correspond to the statistical uncertainty.*

Process	1 Jet	2 Jet	3 Jet	4 Jet
Z+jets	132901 ± 365	29517 ± 172	6391 ± 80	1378 ± 37
di-boson	491 ± 22	371 ± 19	98 ± 10	18 ± 4
top	102 ± 10	227 ± 15	118 ± 11	38 ± 6
total MC	133497 ± 365	30117 ± 174	6608 ± 81	1436 ± 38
data observed	144104 ± 380	30666 ± 175	6648 ± 82	1511 ± 39

Table 5.9: *Expected number and the statistical uncertainty of selected events per jet bin, passing the Z+jets event selection, in the muon channel. The uncertainties correspond to the statistical uncertainty.*

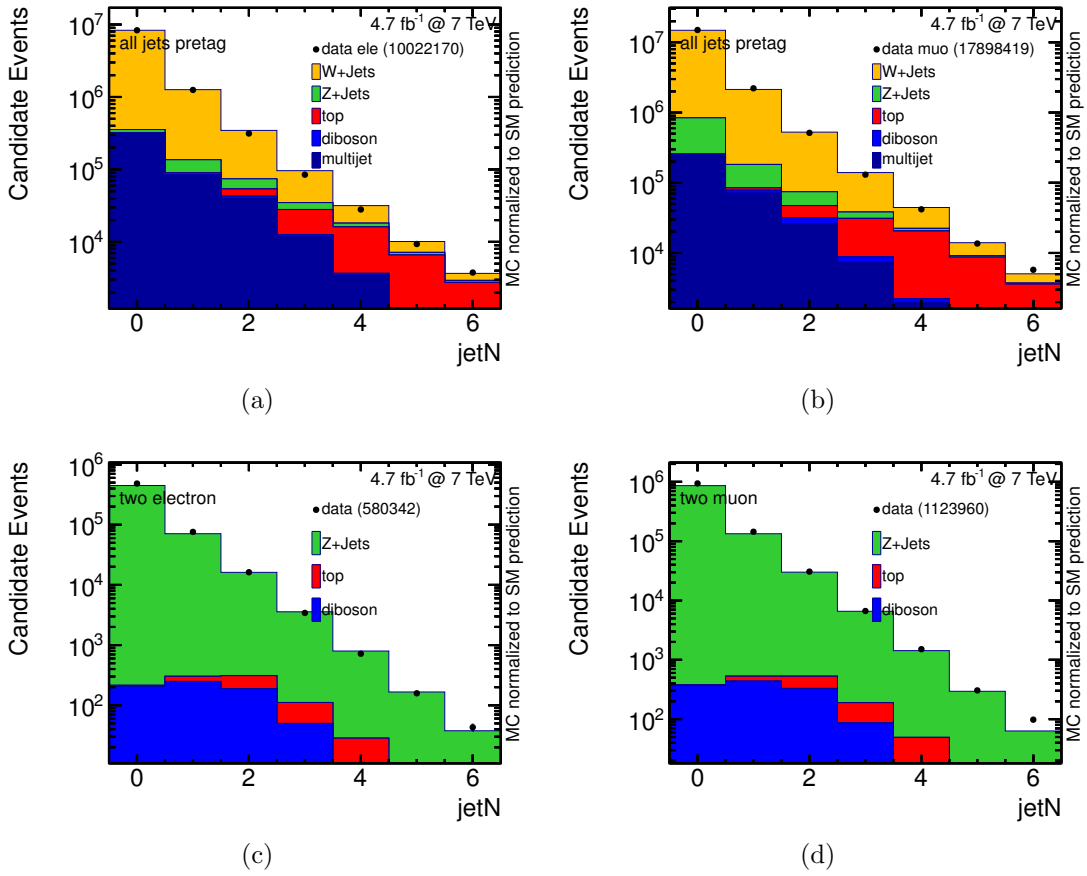


Figure 5.7: *Number of events per jet bin in the electron and muon selection, plotted with logarithmic y-axis. Figures (a) and (b) show the W+jet selection while Figures (c) and (d) show the Z+jet selection. Both, W and Z boson selections are shown for the electron and muon channel, respectively.*

Z boson event candidates selected in data. The calculation is based on the event yields shown in Tables 5.6 to 5.9. In the last column of Table 5.10, the calculated ratio R_n is presented.

5.3.3 Extrapolation of R_n

The basic idea of the N_{jet} ratio method is the extrapolation of R_n to the number of four selected jets. Therefore, this section presents the examination of the ratio characteristics based on simulated events. Afterward, the extrapolation of the data-driven R_n in the four-jet multiplicity is discussed. Furthermore, the influences of systematic effects are investigated.

N_{Jet}	W Selection	Z Selection	$\frac{W+n_{jets}}{Z+n_{jets}}$
		Electron	
1	1143000 ± 1118	75649 ± 276	15.1 ± 0.1
2	249925 ± 560	15876 ± 127	15.7 ± 0.1
3	54488 ± 291	3292 ± 58	16.6 ± 0.3
		Muon	
1	2015810 ± 1487	143508 ± 380	14.0 ± 0.1
2	440606 ± 717	30066 ± 175	14.7 ± 0.1
3	94860 ± 361	6431 ± 82	14.8 ± 0.2

Table 5.10: *Expected number of events for W+jet and Z+jet production per jet bin derived from data. The uncertainty shown is the statistical uncertainty.*

Examination of R_n using Simulated Events

Theory predicts a almost constant ratio of the W and Z boson cross-section dependent on the number of generated partons. Since R_n is constructed from reconstructed events, one observes a linear dependence on the number of jets. Therefore, influences of selection cuts in the W and Z boson event selection have to be analyzed. Simulated W and Z boson events allow an investigation of the ratio R_n. In this section, the behavior of the ratio is analyzed by comparing different kinematic regions, detector efficiencies and MC generators.

Influence of Event Kinematics The W and Z boson event kinematics is shaped by the event-selection requirements defined by the top-quark analysis. While the number of selected events in the different jet multiplicities is not expected to depend on the event-selection cuts in the Z-boson event selection, a dependency is awaited for the W+jets event-selection cuts. The reconstructed missing transverse energy of the event, important in the W+jets selection, has a clear dependency on the number of selected jets. Therefore, the event-selection cuts on the $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\mathrm{m_T}(W)$ lead to a variation in the behavior of R_n. Figure 5.8 shows the ratio R_n, based on ALPGEN, with different W+jets selection cuts applied. At the top, the dependency in the electron selection-channel is shown. On the left hand side, the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is alternated while the one on $\mathrm{m_T}(W)$ is held constant. The distribution on the right shows the variation of the $\mathrm{m_T}(W)$ requirement with a constant $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold. At the bottom, the dependency on the triangular cut in the muon channel is shown. In all histograms, the default R_n distribution based on the top selection cuts is marked in black.

All distributions indicate a linear dependence on R_n up to a number of three selected jets. The slope of R_n depends on the selection requirements of $E_{\mathrm{T}}^{\mathrm{miss}}$ and $\mathrm{m_T}(W)$,

while the general R_n progression remains linear. A deviation from the linearity is observed with the requirement of four selected jets. The $E_{\mathrm{T}}^{\mathrm{miss}}$ variation, shown in Figure 5.8(a), indicates that the deviation has a dependency on the required $E_{\mathrm{T}}^{\mathrm{miss}}$. Since the extrapolation in the four-jet bin is the crucial approach in the presented method, this deviation in the case of four selected jets is analyzed in the following sections.

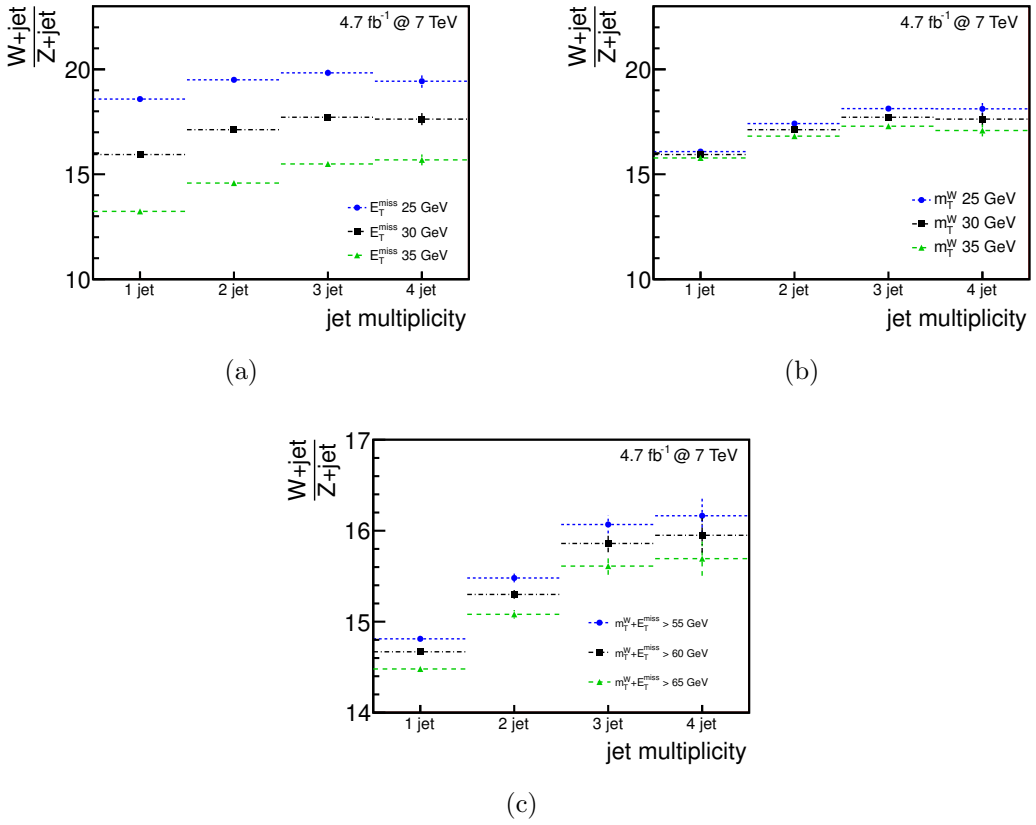
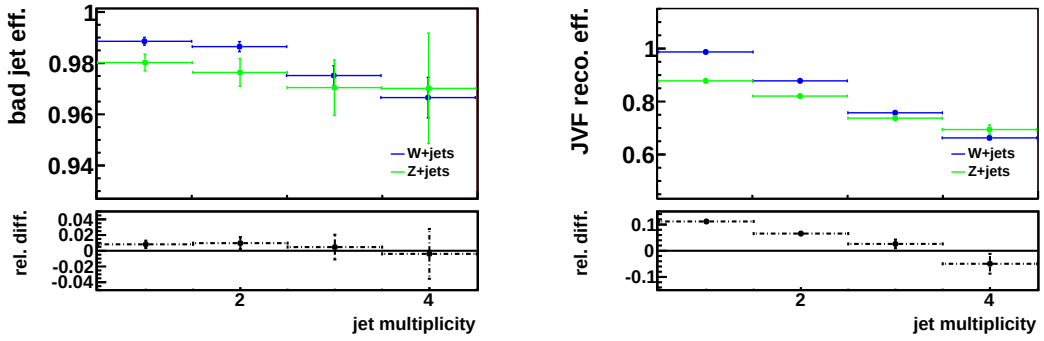


Figure 5.8: *R_n dependence on the W boson selection cuts. Figures (a) and (b) show the variation of the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut ($\mathrm{m_T(W)} > 30$ GeV) and the $\mathrm{m_T}(W)$ cut ($E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV) for the electron channel. Figure (c) shows the dependence of the triangular cut in the muon selection. The W and Z boson events in the calculation of R_n are generated with* ALPGEN.

Validation of the Reconstruction Efficiency The observed deviation from a linear progression, induced by the reconstruction of events with four selected jets, could be caused by differences in the reconstruction efficiencies. Since the reconstruction requirements of W+jets events are inherited from the top-quark analysis, a special tuning of the parameters in the four-jet multiplicity has to be ruled out. Therefore, all event-selection criteria and object definitions are tested in terms of their influence on R_n. For instance, Figure 5.9 shows the reconstruction efficiency

of the W and Z boson selection for the bad-jet veto on the left, and the cut on the JVF in the jet object definition on the right. In the case of a bad-jet event-selection veto, the whole event is rejected. As shown, the induced difference in the W and Z boson selection is of the order of one percent. The JVF requirement in the jet object definition leads to the rejection of jet candidates. Since it is tuned in the four-jet selection, it is a good candidate for a bias in the behavior of R_n. The distribution indicates a difference up to 10% in the one jet bin. Since the relative difference shows a flat dependence on the jet multiplicity, the linearity of R_n is not disturbed, while the slope of R_n is affected. In summary, it can be concluded that the influences of the selection requirements on R_n are either independent or linearly dependent on the number of jets. Therefore, the deviations from the linearity in the case of four selected jets have been found to be uncorrelated with the reconstruction efficiency.



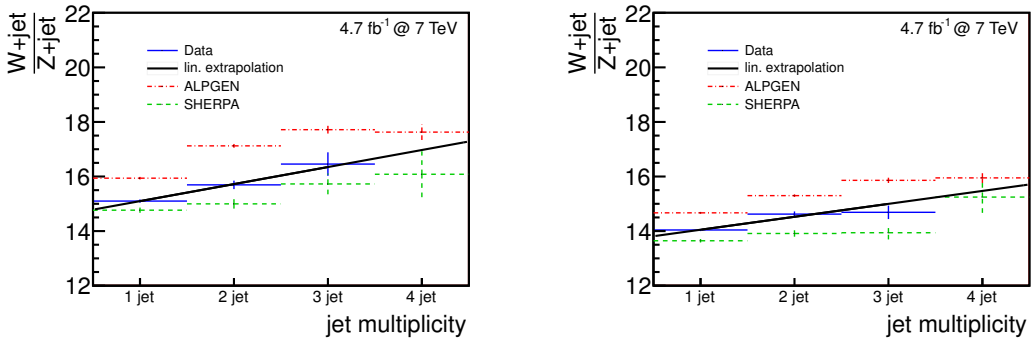
(a) (b)

Figure 5.9: *Validation of the reconstruction efficiency influence on R_n. The bad-jet event-selection veto (a) and cut on the JVF in the jet-object definition (b), both for the electron channel, are shown. At the bottom of both distributions, the relative difference is shown.*

Comparison of Different Monte Carlo Generators Different event generators have different approaches for the simulation of the hard interaction and the generation of parton shower. Therefore, there are differences in the final-state properties of the events. The dependency of R_n on the choice of the event generator is examined in this section. The default W and Z boson samples are produced using Alpgen. These samples are compared with the Sherpa predictions. Figure 5.10 shows the N-jet ratio for the Alpgen and Sherpa predictions, based on the same event-selection requirements. The ratio determined from data, corresponding to Table 5.10, is shown in blue with its linear extrapolation as a reference. Comparing Alpgen and Sherpa predictions, one observes an overall shift of the ratio R_n. The shift is caused by differences in the predicted cross-sections and difference acceptances of W and Z boson events. In the comparison of the predicted W and Z boson ratio the overall shift can be neglected. Also, the slope of R_n shows slight

differences. This can be explained by differences in the momentum of the generated partons. Apart from that, both event generators predict a similar progression of the ratio $R_n$in the first three jet multiplicities.

However, deviations from the linearity in the case of four selected jets can be observed. The drop of R_4 predicted by ALPGEN in the electron channel is not observed for SHERPA. In contrast, the muon channel shows an increase in the SHERPA prediction of R_4 while ALPGEN agrees with the assumed linear dependency. In Table 5.11 the percentage deviations of the predicted ratio R_4, compared to the linear extrapolation of the ratio R_n from simulation, are shown.



(a) (b)

Figure 5.10: *Comparison of the predicted ratio of* ALPGEN *and* SHERPA *to the data-driven result, shown for the electron (a) and the muon (b) channel.*

In general, there are no differences expected in the production of W and Z boson events with an electron or a muon in the final state. In the comparison of the ALPGEN and SHERPA predictions of R_n for the electron and muon channel, a trend of the deviations is not obvious. While Figure 5.8(a) indicates a dependency on the required $E_{\mathrm{T}}^{\mathrm{miss}}$, the comparison of the two generators quantifies this effect. In the electron channel a maximal deviation of minus 2 σ is observed for ALPGEN while the muon channel deviates maximal by plus 2 σ for PYTHIA. Therefore, the deviations are not significant. A systematic uncertainty is added in the determination of R_n to be conservative. It is expected that the fluctuations disappear with higher statistics of simulated W and Z boson events, redoing the analysis for the dataset, recorded in 2012 at a center-of-mass energy of 8 TeV.

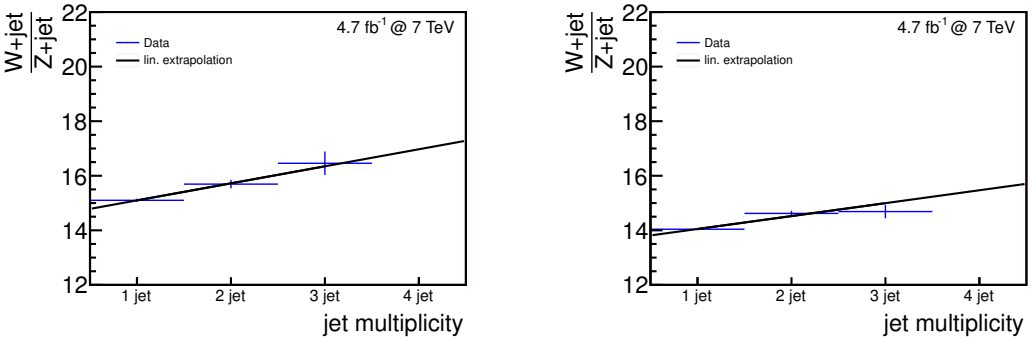
5.3.4 Data Driven Determination of R_4^{extrap}

The data-driven determination of the W and Z boson ratio in the four-jet multiplicity is the basis of the $\mathrm{N_{jet}}$ ratio method, referring to Equation 5.12. With reference to the validation of a linear progression of R_n, analyzed in the previous section, the

	Electron	Muon
ALPGEN	$(-7.6 \pm 4.3)\%$	$(-3.4 \pm 3.1)\%$
SHERPA	$(+1.5 \pm 2.4)\%$	$(+6.9 \pm 3,4)\%$

Table 5.11: *Percentage deviation of R_4 predicted by* ALPGEN *and* SHERPA *simulation from the linear extrapolation of R_n using* ALPGEN *and* SHERPA *predictions.*

data-driven ratio R_n can be extrapolated from the one-, two- and three-jet multiplicity to the four-jet bin. In Figure 5.11, the distribution of R_n as a function of the jet multiplicity is shown for the electron and muon selection. The histogram refers to the numbers shown in Table 5.10. The error bars reflect the statistical uncertainty of R_n . A different slope in the linear dependence, comparing the electron and the muon channel, is associated with different event-selection cuts on the transverse W boson mass. The value of the extrapolated ratio $R_4^{\text{extrap}} = 17.0 \pm 0.3$ for the electron and $R_4^{\text{extrap}} = 15.4 \pm 0.2$ for the muon channel is estimated by a fit of a straight line. The quality of the fit, defined by $\frac{\chi^2}{NDF}$, is 0.17 and 3.79, respectively.



(a) (b)

Figure 5.11: *Ratio of the number of W+n_{jet} divided by Z+n_{jet} events selected in data, shown for the electron (a) and the muon (b) channel.*

The extrapolation to the four-jet multiplicity allows for a prediction of the exclusive four-jet ratio R_4^{extrap}. In the top-quark analysis the ≥ 4 jet multiplicity is analyzed. Therefore, a scaling factor from the exclusive to the inclusive four-jet ratio has to be determined. The number of selected events decreases drastically with the requirement of additional jets. Hence, the difference of the inclusive ratio compared to the exclusive one is expected to be small. The determination of the inclusive scaling factor is achieved on the basis of MC simulations. A relative increase of 1% for the electron and the muon channel is observed.

5.3.5 Treatment of Systematic Uncertainties

The exact knowledge of the ratio of W and Z boson events with four jets is the crucial point of the analysis. The systematic uncertainty on R_4^{extrap} are expected to be small, since the numbers of W and Z boson events in the calculated ratio R_n are determined mainly on the basis of data. The systematic effects in the determination of R_4^{extrap} can be divided into an uncertainty on the considered background processes and an uncertainty on the extrapolation of R_n.

Within this determination of the number of W and Z boson events in data, the amount of expected background events is subtracted according to the MC prediction. Therefore, systematic uncertainties on the simulated events are applied in the same way as described in Section 4.5. Since the systematic effects lead to a modified number of background events in the W and Z boson selection, the number of determined W and Z boson events in data and, therefore, the calculated ratio R_n changes. In Figures 6.8 and 6.9 in the appendix, the deviations on R_n caused by the different systematic effects are shown. The values of R_4^{extrap} are determined for each systematic effect individually. The result is compared to the central value of R_4^{extrap} to estimate the influence of the systematic effect.

One exception is the calculation of the influence of the uncertainty on the multi-jet background fraction. In this particular case the correlations among the jet bins are not known. Hence, the influence of the systematic uncertainty on the multi-jet background in the determination of R_4^{extrap} is estimated by the mean variation of R_n. In addition to the uncertainties on the background events considered, also uncertainties from the assumptions arise, which are made in the determination of R_4^{extrap}. The N_{jet} ratio method suffers from two sources of systematic uncertainties, both related to the extrapolation of R_n:

- The extrapolation of R_n is done by using the one-, two- and three-jet bins as basis. The three jet bin is already afflicted by a large fraction of $t\bar{t}$ background. Since the dependence on the jet multiplicity is linear, a fit based only on the one- and two-jet multiplicity is technically possible, although it has a larger uncertainty. The different fit basis investigated in the extrapolation leads to a slightly different fit outcome of 2% in both lepton selection channels. This systematic effect is referred to as 'number of jet bins'.

- Furthermore, a systematic influence arises with the assumption that R_n has a linear progression as a function of the jet multiplicity. In the previous section it could be shown that the progression of R_n is linear as a function of the jet multiplicity. However, the predictions of different MC generators show a discrepancy from the linear progression in the case of four selected jets. Although the predicted discrepancies are not significant, a systematic uncertainty, labeled as 'extrapolation', is estimated by the deviation of ALPGEN and SHERPA predictions from the linear progression, shown in Table 5.11.

The influence of all investigated systematic uncertainties on R_4^{extrap} is presented in Table 5.12. Shown are the relative uncertainties of R_4^{extrap} for the individual systematic effects in percentage terms. The systematic uncertainty of the method (*total method*), the background modeling (*total other*) and the total uncertainty (*total*) are specified separately for both the electron and the muon channel. Since the object definition and selection cuts of W+jets and Z+jets events are kept similar, the background modeling uncertainties cancel each other out largely in the construction of the ratio R_n.

	Electron	Muon
number of jet bins 2/3	$^{+2.0\%}_{-2.0\%}$	$^{+2.0\%}_{-2.0\%}$
extrapolation	$^{+7.6\%}_{-1.5\%}$	$^{+3.4\%}_{-6.9\%}$
total method	$^{+8\%}_{-3\%}$	$^{+4\%}_{-7\%}$
Luminosity	$^{+1.5\%}_{-1.5\%}$	$^{+1.1\%}_{-1.1\%}$
top cross-section	$^{+2.2\%}_{-2.2\%}$	$^{+1.9\%}_{-1.9\%}$
QCD	$^{+6.3\%}_{-6.5\%}$	$^{+3.4\%}_{-3.4\%}$
ISR/FSR	$^{+0.7\%}_{-1.6\%}$	$^{+0.3\%}_{-0.8\%}$
JES	$^{+1.6\%}_{-0.0\%}$	$^{+0.5\%}_{-0.1\%}$
JER	$^{+0.4\%}_{-3.7\%}$	$^{+1.2\%}_{-0.1\%}$
JRE	$^{+0.5\%}_{-0.5\%}$	$^{+0.1\%}_{-0.1\%}$
JVF SF	$^{+0.6\%}_{-0.5\%}$	$^{+0.2\%}_{+1.0\%}$
E_T^{miss}	$^{+0.2\%}_{-0.2\%}$	$^{+0.1\%}_{-0.1\%}$
E_T^{miss} pile-up	$^{+0.1\%}_{-0.1\%}$	$^{+0.6\%}_{-0.1\%}$
Lepton identification	$^{+0.4\%}_{+0.9\%}$	$^{+1.3\%}_{+1.1\%}$
Lepton resolution	$^{+0.1\%}_{-0.3\%}$	$^{+0.1\%}_{-0.1\%}$
Lepton energy scale	$^{+0.2\%}_{-0.1\%}$	$^{+0.1\%}_{-0.1\%}$
total other	$^{+7\%}_{-8\%}$	$^{+5\%}_{-4\%}$
total	$^{+11\%}_{-9\%}$	$^{+6\%}_{-8\%}$

Table 5.12: *Systematic uncertainties on R_4^{extrap}, divided into uncertainties on the method and uncertainties on background estimation and simulation*

The systematic uncertainties on R_4^{extrap}, shown in Table 5.12, indicate a small influence from the uncertainties on the MC simulation. One of the main systematic

uncertainty arises from the uncertainty of the data-driven multi-jet background estimation. The multi-jet background is equally present in all jet multiplicities in the W boson selection. With an uncertainty of 50% it leads to one of the largest uncertainties of R_n. Although, the determination of R_4^{extrap} is constructed with the attempt to be independent of $t\bar{t}$ production, the top cross-section uncertainty leads to the second largest uncertainty considering the background modeling uncertainties.

With an uncertainty of up to 7.6% in the electron and up to 6.9% in the muon channel, the extrapolation uncertainty has a large influence in the precision of the determined W and Z boson ratio in the four-jet multiplicity. Since the uncertainty is motivated by the fluctuation of the ratio determined in MC simulations, it is expected that this uncertainty has a reduced influence in the analysis of 8 TeV collision data recorded by ATLAS in the year 2012. To provide accurate analyses based on this dataset, the MC statistics is increased by a factor of approximately one hundred.

5.3.6 W+4jet Normalization based on Data

In the previous section, the ratio of W and Z boson events was estimated from data and extrapolated to the number of four selected jets. The final number of W+4 jets events results, as already shown in Equation 5.12, from the product of R_4^{extrap} and the number of Z+4 jets events derived from data.

The selection of W+jets events is identical to the top-quark selection and, therefore, the resulting number of estimated W+4 jets events refers directly to a W+jets background estimation in the lepton plus jets $t\bar{t}$ analysis. Since the selection is performed without the b-jet requirement, the determined number of W+4 jets events corresponds to the W+jets background in the pre-tag $t\bar{t}$ analysis. The central value of R_4^{extrap}, its statistical uncertainty and the combination of all systematic uncertainties is shown in Table 5.13. The number of Z+4 jets events determined in data is shown in Table 5.14. Table 5.15 shows the number of identified W+jets events in the inclusive and exclusive four-jet multiplicity.

	Electron	Muon
exclusive	$17.0 \pm 0.03^{+1.89}_{-1.48}$	$15.4 \pm 0.02^{+0.92}_{-1.31}$
inclusive	$17.2 \pm 0.03^{+1.91}_{-1.49}$	$15.6 \pm 0.02^{+0.93}_{-1.33}$

Table 5.13: *Data-driven estimated value of R_4^{extrap} and its statistical and systematic uncertainty. The value is denoted for the exclusive and the inclusive four-jet multiplicity for the electron and muon selection separately.*

In the top-quark analysis, the W+jets rate predicted by the MC simulation is corrected for the data-driven results by the application of scaling factors. Therefore,

	Electron	Muon
exclusive	685±26	1453±38
inclusive	879±30	1841±43

Table 5.14: *Number and statistical uncertainty of Z+jets events in the four-jet multiplicity selected in data. The numbers are specified for the exclusive and the inclusive four-jet multiplicity for the electron and muon selection separately.*

	Electron	
	number of events	scaling factor
exclusive	$11638 \pm 493(stat.)^{+1297}_{-1014}(sys.)$	$0.86 \pm 0.04(stat.)^{+0.10}_{-0.07}(sys.)$
inclusive	$15084 \pm 577(stat.)^{+1681}_{-1314}(sys.)$	$0.87 \pm 0.03(stat.)^{+0.10}_{-0.08}(sys.)$
	Muon	
	number of events	scaling factor
exclusive	$22405 \pm 656(stat.)^{+1342}_{-1909}(sys.)$	$1.02 \pm 0.03(stat.)^{+0.06}_{-0.09}(sys.)$
inclusive	$28672 \pm 763(stat.)^{+1717}_{-2443}(sys.)$	$1.02 \pm 0.03(stat.)^{+0.06}_{-0.09}(sys.)$

Table 5.15: *Estimated number of W+4 jets events and their relative difference to the MC prediction.*

in the right column of Table 5.15, the estimated W+jets scaling-factors are shown. The presented scaling factors are equivalent to the scaling factors estimated by the charge-asymmetry method, introduced in Section 4.4.1. Hence, the determined scaling factors of both methods can be compared directly. The scaling factors calculated with the charge-asymmetry method in the pre-tag selection are shown in Table 5.16, together with their total uncertainty. The comparison of the scaling factors derived from the two different methods shows that the N_{jet} ratio method leads to slightly higher values in the predicted W+jets background rate. The uncertainties of both methods are compatible. Within these uncertainties, the estimated scaling factors are consistent.

	Electron	Muon
exclusive	$0.82^{+0.09}_{-0.08}$	$0.89^{+0.08}_{-0.07}$
inclusive	$0.83^{+0.08}_{-0.08}$	$0.87^{+0.07}_{-0.07}$

Table 5.16: *W+jets scaling factors for the pre-tag selection [68], determined with the charge-asymmetry method in events with at least four jets.*

The idea of the presented N_{jet} ratio method is to improve the estimation of the number of W+jets background events in the analysis of the semileptonic $t\bar{t}$ decay. Since the number of W and Z boson events in the calculated ratio R_n is determined mainly on the basis of data, an improvement in the systematic uncertainty, compared to the uncertainty of the charge-asymmetry method, was expected. Table 5.17 shows a disposition of the most important systematic effects investigated. The systematic effects in the table are arranged in common and independent effects comparing the W+jets normalization determination of the charge-asymmetry and the N_{jet} ratio method.

The charge-asymmetry method combines data-driven techniques with simulated properties. Hence, several systematic uncertainties have to be considered. This includes the uncertainty due to the MC-generator choice, charge misidentification, jet energy scale, PDF variations and the estimation of the W+jets heavy-flavor fractions. Since the charge-asymmetry method requires the measurement of the W+jets heavy-flavor fractions in the pre-tag selection, the pre-tagged result depends also on the uncertainty of the b-tagging efficiency.

The N_{jet} ratio method is affected primarily by uncertainties on R_4^{extrap}, already discussed in Section 5.3.5. Since R_n is estimated mainly from data and the systematic effects in the considered background samples partly cancel each other out, most of the systematic uncertainties are reduced compared to the systematic uncertainties on charge-asymmetry method. In addition, many of the systematic effects, such as the PDF uncertainty, considered in the result of the charge-asymmetry method and referred to as 'only CA method', don't have significant influences on the N_{jet} ratio method result.

The uncertainty on the assumptions made in the N_{jet} ratio method , referred to as 'only N_{jet} ratio method', is dominated by the uncertainty on the extrapolation, caused by the deviation of the MC prediction in the case of four selected jets. It is expected that this effect disappears with an increased MC statistic, which is available with the analysis of the 8 TeV dataset. A more detailed study on the MC generator effects, not performed within this thesis, could reduce the uncertainty as well. Without the consideration of this systematic effect, the total uncertainty of the N_{jet} ratio method result would be significantly improved compared to the total uncertainty of the charge-asymmetry method.

However, the presented result of the N_{jet} ratio method with a comparable uncertainty compared to the charge-asymmetry method allows also an improvement of the W+jets prediction. Since both methods start from two completely independent approaches, most of the systematic effects are uncorrelated. A combination of both methods is assumed to reduce the systematic uncertainty of the W+4 jets normalization in the top-quark analysis significantly.

Systematic Source	charge asymmetry		N_{jet} ratio	
	Electron	Muon	Electron	Muon
top cross-section up/down	1.1/-1.1	0.9/-0.8	2.2/-2.2	1.9/-1.9
QCD up/down	4.6/-4.7	3.5/-3.5	6.3/-6.5	3.4/-3.4
ISR/FSR up/down	0.4/-0.4	1.0/-1.0	0.7/-1.6	0.3/-0.8
JES up/down	-0.6/-2.8	0.0/-3.4	1.6/-0.0	0.5/-0.1
JER up/down	-1.1/1.1	-0.37/0.37	0.4/-3.7	1.2/-0.1
JRE up/down	0.45/-0.45	0.21/-0.21	0.5/-0.5	0.1/-0.1
JVF SF up/down	0.29/-0.34	0.35/-0.33	0.6/-0.5	0.2/1.0
$E_{\mathrm{T}}^{\mathrm{miss}}$ up/down	-0.61/-0.76	-0.14/0.83	0.2/-0.2	0.1/-0.1
$E_{\mathrm{T}}^{\mathrm{miss}}$ pile-up up/down	-0.5/0.5	0.039/0.31	0.1/-0.1	0.6/-0.1
lepton identification up/down	1.28/-1.43	-0.14/0.14	0.4/0.9	1.3/1.1
lepton resolution up/down	0.01/0.56	0.5/0.5	0.1/-0.3	0.1/-0.1
lepton energy scale up/down	0.16/0.56	0.5/0.5	0.2/-0.1	0.1/-0.1
common total	5.0/-6.0	3.9/-5.2	6.9/-8.0	4.1/-4.0
b-tag efficiency	2.6/-2.0	3.4/-2.5		
W heavy-flavor up/down	-2.4/2.6	-3.2/3.4		
PDF up/down	-8.5/9.8	-6.7/8.9		
charge misidentification up/down	-0.7/0.7	-0.7/0.7		
W shape up/down	0.65/-0.65	0.23/-0.23		
W modeling ($p_{\mathrm{T}}{}^{j}$ min) up/down	0.61/-0.61	0.24/-0.24		
W shower up/down	-2.2/2.2	-1,4/1.4		
only CA method (total)	-9.5/10.6	-8.3/9.9	x	x
number of jet bins 2/3			2.0/-2.0	2.0/-2.0
extrapolation			7.6/-1.5	3.4/-6.9
luminosity up/down			1.5/-1.5	1.1/-1.1
only N_j*et*-ratio method (total)	x	x	8.0/-3.0	4.1/-7.3
total	11.7/-11.2	10.6/-9.8	10.6/-8.5	5.8/-8.3

Table 5.17: *Comparison of the relative systematic uncertainties in the two data-driven W+jets normalization estimations for electron and muon events with at least four jets. The uncertainties on the CA method are taken from [85].*

Estimation of the Normalization with b-jet Requirement

The final analysis of semileptonic $t\bar{t}$ events is done with the requirement of at least one b-tagged jet in the event. As with the W+jets normalization determination with the charge asymmetry method, the N_{jet} ratio method employs an event selec-

tion without requiring a b-jet. Hence, the difference in the W+jets normalization compared to the tagged selection has to be calculated.

In the case of the charge-asymmetry method, the W+jets normalization in the tagged selection is determined from the pre-tag result by means of W+jets tagging efficiencies. The detailed calculation is shown in Equation 4.8 in Chapter 4.4.1. It can be shown that the central value of the W+jets normalization is independent of the b-jet requirement while the systematic uncertainties differ substantially [68].

The W+jets normalization predicted by the N_{jet} ratio method can be determined in the tagged selection with the same approach. Table 5.18 shows a itemization of all systematic uncertainties investigated in both W+jets normalization methods for the tagged selection. In comparison with Table 5.17 it becomes apparent that many systematic uncertainties of the charge-asymmetry method are strongly correlated with the W+jets tagging fractions. In contrast, the systematic uncertainties of the N_{jet} ratio method stay the same. Two additional uncertainties, linked with the uncertainties on the W+jets tagging efficiencies, introduced in Chapter 4.4.1, are included. The W+jets tagging efficiency f_{tag}^{2j} is derived from data [68] while the W+jets tagging efficiency $f_{tag}^{2\to4}$ is estimated based on W+jets MC predictions. In the estimation of both W+jets tagging efficiencies the uncertainty on the b-tag efficiency and the uncertainty on the W+jets heavy-flavor content is relevant.

A tiny modification of the N_{jet} ratio method could also provide a measurement in the tagged selection independent of the W+jets tagging efficiencies. It is based on the idea that a W and Z boson ratio (R_n^{tagged}) can be determined directly in a selection with a b-tag requirement. In contrast to the presented N_{jet} ratio method this approach is more challenging. A higher $t\bar{t}$ background fraction in the selection of W and Z boson events is expected to cause a stronger $t\bar{t}$ cross-section dependence. Furthermore, the difference in the production processes of W and Z boson events associated with heavy-flavor jets, discussed in 5.1, might affect the progression of R_n^{tagged}. In addition, the requirement of at least one b-tagged jet in the event reduces the number of selected events and, therefore, increases the statistical uncertainty of R_n^{tagged}. The numbers of expected events in the pre-tag and tagged selection are presented in Table 5.19. The problem of a limited statistics could be avoided with a modification of the event-selection criteria for W and Z boson events. For instance, the removal of the b-tag requirement in the Z+jets selection would increase the number of selected Z+jets events. Detailed studies have to prove that the progression of R_n^{tagged} is independent from such a approach.

The final scaling factors of the W+jets background normalization for the tagged semileptonic $t\bar{t}$ selection are shown in Table 5.20. Within this table, the predicted W+jets normalizations of the N_{jet} ratio method and the charge-asymmetry method are compared. The given uncertainty is the statistical and systematic uncertainty, summed in quadrature. It can be shown that the determination of the W+jets normalization with the N_{jet} ratio method leads to an improved uncertainty of up

Systematic Source	Charge Asymmetry		N_{jet} ratio	
	Electron	Muon	Electron	Muon
top cross-section up/down	2.3/-2.1	1.9/-1.9	2.2/-2.2	1.9/-1.9
QCD up/down	9.1/-9.1	5.8/-5.8	6.3/-6.5	3.4/-3.4
ISR/FSR up/down	1.2/-1.2	1/-1	0.7/-1.6	0.3/-0.8
JES up/down	0.7/-1.9	-2.2/4.3	1.6/-0.0	0.5/-0.1
JER up/down	-1.3/1.3	1.2/-1.2	0.4/-3.7	1.2/-0.1
JRE up/down	0.05/-0.05	0.06/-0.06	0.5/-0.5	0.1/-0.1
JVF SF up/down	0.29/-0.34	0.35/-0.33	0.6/-0.5	0.2/1.0
$E_{\mathrm{T}}^{\mathrm{miss}}$ up/down	0.63/-0.01	-0.03/0.00	0.2/-0.2	0.1/-0.1
$E_{\mathrm{T}}^{\mathrm{miss}}$ pile-up up/down	0.78/0.93	0.03/0.28	0.1/-0.1	0.6/-0.1
lepton identification up/down	1.28/-1.43	0.54/-0.54	0.4/0.9	1.3/1.1
lepton resolution up/down	1.3/0.9	-0.005/0.001	0.1/-0.3	0.1/-0.1
lepton energy scale up/down	1.16/2.51	-0.019/0.019	0.2/-0.1	0.1/-0.1
b-tag efficiency	5.3/-5.3	4.1/-4.1	1.0/-1.0	0.9/-0.9
W heavy-flavor up/down	-13/13	-12/12	5.3/-5.3	5.1/-5.1
common total	11.2/-11.2	7.8/-8.5	8.3/-9.6	6.8/-6.5
PDF up/down	-6.9/6.9	-4.0/4.0		
charge misidentification up/down	-0.7/0.7	-0.7/0.7		
W shape up/down	0.05/-0.05	0.09/-0.09		
W modeling ($p_{\mathrm{T}}{}^{j}$ min) up/down	0.09/-0.09	0.14/-0.14		
W shower up/down	-0.96/0.96	-1,64/1.64		
only CA method (total)	-7.0/7.0	-4.4/4.4	x	x
number of jet bins 2/3			2/-2	2/-2
extrapolation			7.6/-1.5	3.4/-6.9
luminosity up/down			1.5/-1.5	1.1/-1.1
only N_jet-ratio method (total)	x	x	8.0/-3.0	4.1/-7.3
total	13.2/-13.2	9.0/-9.6	11.5/-10.1	7.9/-9.8

Table 5.18: *Comparison of the relative systematic uncertainties in the two data-driven W+jets normalization estimations for electron and muon events with at least four jets and one B-hadron identified jet. The uncertainties on the CA method are taken from [68, 85].*

to 4% in both electron and muon channel. Furthermore, a combination of both results would reduce the systematic uncertainty on the W+jets normalization more significantly, since it is expected that the systematic effects are mainly uncorrelated.

Jet	W+jets		Z+jets	
Bin	pre-tag	tagged	pre-tag	tagged
1	1112137	40835	70654	1969
2	261287	20840	15866	940
3	57776	7157	3463	310
4	12475	2210	768	90

Table 5.19: *Number of W and Z boson events in the pre-tag and tagged selection predicted by MC simulation.*

Charge Asymmetry		N_{jet} ratio	
Electron	Muon	Electron	Muon
$0.83^{+0.14}_{-0.13}$	$0.87^{+0.11}_{-0.11}$	$0.87^{+0.11}_{-0.10}$	$1.02^{+0.09}_{-0.11}$

Table 5.20: *W+jets scaling factors for the tagged inclusive four-jet selection. The prediction of the W/Z-boson ratio estimation and the charge-asymmetry method [68] are compared. The given uncertainty is the statistical and systematic uncertainty, summed in quadrature.*

5.3.7 Conclusion

In this chapter a method was presented to determine the number of W+jets background events in the four-jet multiplicity. A result was achieved for the event selection with and without a b-jet requirement. The results shown are in good agreement with the predictions of the charge asymmetry method, which is currently used as standard method in the $t\bar{t}$ analysis. The N_{jet} ratio method provides an independent measurement of the W+4 jets background fraction in the $t\bar{t}$ analysis, complementary to the charge-asymmetry method. Since both methods have independent approaches and, therefore, can be assumed to be almost uncorrelated, a combination of the results will lead to a reduction of the uncertainty in the W+4 jets normalization.

One of the dominant systematic uncertainties is related to the extrapolation of the W and Z boson ratio. While differences in the predictions of Alpgen and Sherpa lead to an significant uncertainty in the presented result, it is expected that the Alpgen and Sherpa predictions with an increased MC statistics will confirm the linear progression of R_n. Furthermore, a possible extension of the N_{jet} ratio method to estimate the tagged W+jets normalization independently from the simulated W+jets tagging efficiencies could further reduce the systematic uncertainties on the N_{jet} ratio method. Therefore, the influence on the progression of R_n^{tagged} and on the systematic and statistic uncertainties has to be validated.

5.4 Summary

In this chapter, W and Z boson events are compared with respect to their event properties. The main causes for differences of W and Z boson events are induced by a different vector-boson mass, different couplings and a different heavy-flavour content. No differences are expected for the QCD production of jets in association with the vector-boson production. W+jets events are difficult to separate in the higher jet multiplicities since they are contaminated with top-quark events. It is shown that Z+jets events can be used as a tool to model W+jets events. The conversion of one lepton from the Z-boson decay into the "neutrino signal" allows a modeling of all crucial observables in W+jets events.

Since the jet production is identical in both W and Z boson events, the calculation of the W and Z boson ratio allows an estimation of the number of W+jets background events in the semileptonic $t\bar{t}$ selection. The presented results, published in [86], are in good agreement with the calculations of the charge-asymmetry method. The N_{jet} ratio method provides a W+jets normalization determination which is independent from the W+jets charge asymmetry. Therefore, it can be applied also in analyses of charge asymmetric processes in the four-jet multiplicity. A combination of the two methods and additional studies on the N_{jet} ratio method promise further improvements in the estimation of the W+4 jets normalization.

Both Z+jets based methods presented in this section show that Z+jets events are well suited to be used as a tool to investigate W+jets events. The application of both methods would lead to an estimation of the W+jets event properties and the W+jets normalization from data simultaneously. Therefore, it would allow one to estimate the W+jets background independently from the simulation of W+jets events.

Chapter 6

Analysis of Top Quark Resonances

The study of top-quark pair production in the semileptonic decay channel is described in the previous chapters. This chapter briefly discusses the analysis of beyond SM resonant production of particles, which are expected to decay to top-quark pairs. Since both processes lead to the same final state signature, the $t\bar{t}$ continuum produced by SM $t\bar{t}$ production is the major background in this analysis. The resonant $t\bar{t}$ production is expected to be visible as peak in the invariant mass-spectra of the $t\bar{t}$ system. Searches for new heavy particles in top-quark events have already been performed at the Tevatron at the Fermi National Accelerator Laboratory in proton anti-proton collisions at a center-of-mass energy of $\sqrt{s} = 1.8$ TeV. Within these measurements, a lower mass limit of around 820 GeV at 95% C.L. for a leptophobic Z' boson [23] was achieved [87, 88].

The analysis that will be presented in this chapter, published in [89], is based on data with the integrated luminosity of 4.7 fb^{-1} recorded in $\sqrt{s} = 7$ TeV pp collisions at the LHC in the year 2011. Since the LHC with its higher center-of-mass energy is able to extend the mass range of $t\bar{t}$ resonance searches, also the reconstruction techniques have to be adapted to the higher energies. In the previous chapters, the resolved signature of the $t\bar{t}$ decay, in which all decay products can be reconstructed individually, is discussed. At higher energy and larger momentum of the top-quarks, the top-quark decay products can overlap. In this case, resolving the individual decay products in the event reconstruction is no longer possible. The first section of this chapter introduces the reconstruction of the $t\bar{t}$ event candidates. Two reconstruction methods are adapted to reconstruct $t\bar{t}$ events at low and high invariant masses in an optimal way. The analysis of the events is subdivided in a *resolved* and a *boosted* reconstruction scheme. While the reconstruction in the resolved scheme is similar to the one described in Chapter 4.1, in the boosted event reconstruction a new approach is necessary. Since no evidence for a $t\bar{t}$ resonance is observed, an exclusion limit on the cross-section times branching-ratio is calculated, presented in the second section. In this calculation two benchmark models, partitioned into a narrow and a broad resonance with respect to the experimental mass resolution, are considered.

In Chapter 5.3, a new method to estimate the W+jets background normalization was introduced. To evaluate the efficiency of this N_{jet} ratio method, the analysis of $t\bar{t}$ resonances is chosen. Therefore, the calculation of cross-section limits, based on the estimated W+jets normalization, is presented in the section of statistical analysis. Finally, the exclusion limits on the investigated benchmark models, determined in [89], are presented.

6.1 Reconstruction of Event Candidates

In the presented analysis, only semileptonic decays of top-quark pairs are considered. The boost of the top-quark, due to a higher mass of the resonance, leads to a smaller angular separation of the top-quark decay-products. The separation of the decay products in a two-body decay is approximately

$$\Delta R \approx \frac{2M}{p_{\mathrm{T}}} \tag{6.1}$$

where M and p_{T} is the mass and the momentum of the decaying particle respectively. The decay of a top quark with a transverse momentum of 400 GeV, for instance, would produce a W boson and a b-quark with a separation of $\Delta R \approx 0.8$. Therefore, a separation of a hadronically decaying W boson and the b-jet with the default $\Delta R = 0.4$ jet reconstruction algorithm is no longer possible. Hence, the different event topologies in the low and high $m_{t\bar{t}}$ region motivate the separation of the analysis in a resolved and boosted selection scheme. The selection schemes are chosen in a way that the analysis shows a good performance in both $m_{t\bar{t}}$ regions. Most of the object definitions and event-selection requirements are kept identical to the analysis of SM $t\bar{t}$ production, described in the previous chapters. Only the differences in the event selection compared to the analysis of SM $t\bar{t}$ production are discussed in the chapter.

First, the lepton isolation criterion which is common in the resolved and boosted selection scheme is discussed. Afterwards, the event selection and the $m_{t\bar{t}}$ reconstruction is described for both methods.

6.1.1 Mini Isolation

The default isolation criterion used in the $t\bar{t}$ analysis is defined in Chapter 4.1. Since the distance of lepton from the leptonically decaying top quark to the b-jet is expected to be reduced in this analysis, the default isolation requirement would filter a large amount of events out. Thus, a new isolation criterion, referred to as 'mini-isolation', is defined [90]:

$$I^{l}_{mini} = \sum_{tracks} {p_{\mathrm{T}}}^{track} \tag{6.2}$$

Tracks with a transverse momentum ${p_T}^{\text{track}} > 1$ GeV, which pass a set of quality cuts and fulfill the ΔR requirement $\Delta R(l, \text{track}) < {K_T}/{{p_T}^l}$, are considered. Here K_T is an empirical scale parameter defined as 10 GeV. In the analysis, a cut of ${I^l_{mini}}/{{p_T}^l} < 0.05$ is used, corresponding to a selection efficiency of 98% in case of the electron and 95% in case of the muon selection channel. Since the isolation requirement also shows a good performance in the low $m_{t\bar{t}}$ region, this requirement is applied in both selection schemes.

6.1.2 Resolved Analysis

The analysis in the resolved selection scheme is very similar to the analysis of SM $t\bar{t}$ production, since it is optimized for the reconstruction of $t\bar{t}$ decays with a low invariant $t\bar{t}$ mass. Events that pass a single-lepton trigger are selected. Exactly one isolated lepton with a p_T >25 GeV is required. The only difference in the object definition is the isolation criterion of the lepton, which changed from a relative isolation to the mini-isolation ${I^l_{mini}}/{{p_T}^l} < 0.05$. Apart from that, all object definitions are applied identically to the recommendations for top-quark analyses defined by the ATLAS collaboration, introduced in Chapter 4.1.

In contrast to the event selection in the SM $t\bar{t}$ analysis, the resolved analysis requires events with at least three instead of four anti-k_t jets ($R = 0.4$) with a p_T larger than 25 GeV. One of these jets is required to have a mass above 60 GeV [91]. This jet is assumed to contain the two quarks of the hadronic W boson decay or one of these quarks and the b-quark from the top quark decay. Therefore, also the resolved analysis shows indications for the expectation of a high top quark boost. If such a jet is not present in the event, events with less than four jets are discarded. Apart from that, all the event-selection criteria, for instance the cut on the E_T^{miss} or the b-tag requirement, are applied identically to the recommendations introduced in Chapter 4.1.

Background Estimation

While most of the background processes are estimated using MC simulation, two of the backgrounds and their uncertainties are determined using data-driven methods.

The multi-jet background is estimated from data in a control region using the matrix method, introduced in Chapter 4.3.1. The defined control regions are discussed in detail in [89]. The JetElectron template method is used to evaluate the systematic uncertainty of the background shape in the $m_{t\bar{t}}$ distribution.

The treatment of W+jets background is identical to the description in Chapter 4.4. While the shape of the W+jets background is estimated using simulated events, the W+jets normalization and the W+jets heavy-flavor content are determined data-driven. The scaling factors in the resolved analysis are identical to the ones

calculated for the analysis of SM $t\bar{t}$ production [68]. This approach is tested to be valid [89], since the differences in the event selection, compared to the event selection used for the analysis of SM $t\bar{t}$ production, are negligible in terms of the W+jets normalization. The differences are a modified lepton isolation and a minimal jet p_{T} requirement on the most energetic jet. In case of the altered isolation definition, the objects originated from the W boson decay in W+jets background events are well separated and mostly pass both isolation requirements. Therefore, the different lepton isolation has a negligibly small influence in the determination of the W+jets normalization. Since the W+jets normalization is assumed to be independent of the minimal jet momentum requirement [89], also the requirement that at least one jet has a $\mathrm{p_T} > 60$ GeV in case of three selected jets has a negligible influence in the calculation of the W+jets normalization.

$m_{t\bar{t}}$ Reconstruction

The resonant signal is expected to be visible as a peak in the invariant mass of the $t\bar{t}$ system. Within the $m_{t\bar{t}}$ determination, all top-quark decay products have to be reconstructed. Since one of the W bosons decays leptonically, the produced neutrino leads to an additional degree of freedom. Only the transverse components of the neutrino momentum can be reconstructed. The missing longitudinal component p_z is reconstructed in the lepton plus $E_{\mathrm{T}}^{\mathrm{miss}}$ system, applying an on-shell W boson mass constraint. The missing momentum component p_z is obtained from a quadratic equation [92]. In 60% of all cases, the generated neutrino is better reproduced choosing the solution with smaller p_z in case of a real solution [89]. With the presence of pile-up in the event or due to an imperfect $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement, the solution can also be complex. In this case the $E_{\mathrm{T}}^{\mathrm{miss}}$ is modified in a way that the discriminant is zero. Therefore, the value of $\mathrm{m_T}(W)$ is set to the on-shell W-mass in such cases.

Additional jets in the event can be misidentified as top quark decay products. An inaccurate assignment in the calculation of $m_{t\bar{t}}$, for instance distorted by the choice of a jet from the initial or final state radiation, leads to non-Gaussian tails in the mass resolution. A χ^2 method is used to obtain the best jet assignment [89].

$$\begin{aligned}\chi^2 = &\left[\frac{m_{jj} - m_W}{\sigma_W}\right]^2 \\ &+ \left[\frac{m_{jjb} - m_{jj} - m_{t_h - W}}{\sigma_{t_h - W}}\right]^2 + \left[\frac{m_{jl\nu} - m_{t_l}}{\sigma_{t_l}}\right]^2 \\ &+ \left[\frac{(p_{T,jjb} - p_{T,jl\nu}) - (p_{T,t_h} - p_{T,t_l})}{\sigma_{p_{T,diff}}}\right]^2\end{aligned} \tag{6.3}$$

In the calculation of the χ^2, the reconstructed top-quark and W boson mass are used as constraints. The parameter t_h and t_l denote the hadronically and leptonically

decaying top-quarks, respectively. The parameters m_W, m_{t_h-W}, m_{t_l}, σ_W, σ_{t_h-W}, σ_{t_l}, $p_{T,t_h} - p_{T,t_l}$ and $\sigma_{p_{T,diff}}$ are taken from simulation. The permutation of all selected jets with the lowest χ^2 is used to calculate $m_{t\bar{t}}$. In the reconstruction of $t\bar{t}$ events, the correct assignment is achieved in approximately 65% of the events [89]. For events containing jets with p_T >60 GeV the χ^2 is slightly modified, since this heavy jet can be either assigned to two quarks from the W boson decay or to one of these quarks and the corresponding b-quark from the top-quark decay [89].

The χ^2 method replaces the established ΔR_{min} approach, used in the resonance analysis based on 2 fb^{-1} proton collision data at 7 TeV [93]. The ΔR_{min} method reduces the contribution of wrong associated jets by excluding jets from the $m_{t\bar{t}}$ calculation, if their distance to the lepton or the closest jet is larger than $\Delta R_{min} > 2.5 - 0.015 \cdot m_j$. Therefore, it starts from the four most energetic jets and is iterated until no fake jet candidate is found or four jets remain. The jet of the leptonically decaying top-quark is defined as the closest jet to the lepton. The three remaining jets are assigned to the hadronically decaying top-quark. While the ΔR_{min} approach performs better in a dataset without the separation in boosted and resolved $t\bar{t}$ events, the χ^2 shows a better performance in the resolved analysis [89]. Within the analysis of the 2 fb^{-1} dataset, one of my contributions to the resonance analysis was the estimation of the multi-jet background using the JetElectron template model. Since the topology of the selected events differs from the one presented in this chapter, the results are not shown explicitly within this thesis.

Figure 6.1 shows distributions of some variables for the resolved analysis. The transverse momentum of the most energetic jet and the reconstructed invariant $t\bar{t}$ mass spectra are shown. The distributions, shown for the electron and muon selection-channel individually, indicate a good agreement of data with the expected background processes. A deficit in data can be observed for both selection channels in the $m_{t\bar{t}}$ region of 1 TeV to 2 TeV. The $m_{t\bar{t}}$ spectra in the electron selection channel, shown in Figure 6.1(c), indicates a deviation above one σ in this region. The cause of this effect is associated with a problem in the simulation of the $t\bar{t}$ continuum at high mass. Investigations of this effect are ongoing.

6.1.3 Boosted Analysis

The analysis of top-quark pair production with a high $m_{t\bar{t}}$ differs from the resolved analysis scheme described above. Since the top quarks have a high momentum, their decay products are close together. With the assumption that all jets of the hadronic top quark decay overlap, the boosted analysis is optimized for the reconstruction of $t\bar{t}$ events with a high invariant mass. To assemble all constituents in the resulting jet, a new jet object, referred to as 'fat-jet', is constructed with a higher jet-radius parameter in the jet-reconstruction algorithm.

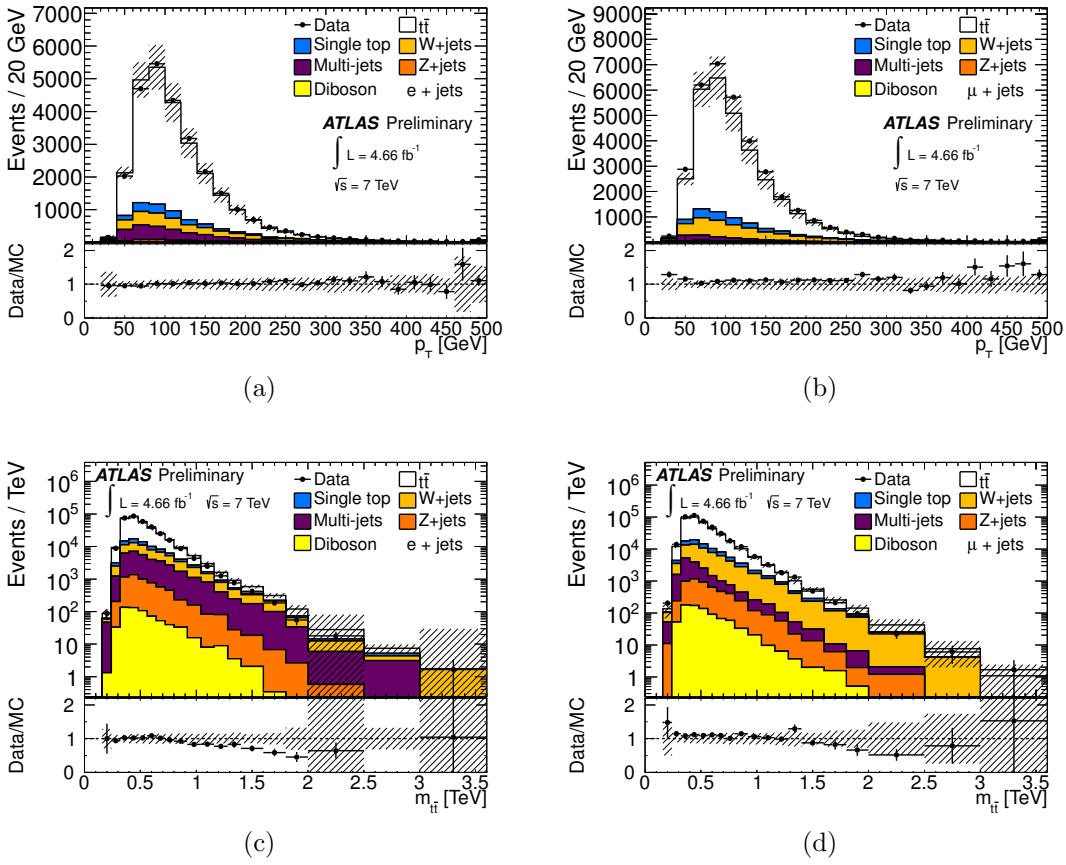


Figure 6.1: *Transverse momentum of the leading jet and the reconstructed invariant $t\bar{t}$ mass for the resolved analysis. The left side shows the distributions for events containing an electron and on the right events containing a muon are shown. Events, also passing the boosted reconstruction scheme are excluded. The shaded areas indicate the total systematic uncertainty. [89].*

Fat Jet Reconstruction

The fat-jet reconstruction also uses the anti-k_T algorithm. The radius parameter is enlarged to $R = 1.0$. While the topological clusters for narrow $R = 0.4$ jets are calibrated at the EM scale, the broad $R = 1.0$ jets use locally calibrated topological clusters.

The momentum of the fat-jets, reconstructed within a region of $|\eta| < 2.0$, is required to be larger than 350 GeV. This value is chosen, since the $R = 1.0$ jet-trigger, used in the boosted selection, is 100% efficient within its uncertainty above this threshold. In addition, the invariant mass of the fat-jet is required to be larger than 100 GeV [91]. Fake jets can be further reduced by including information on the sub-structure of the fat-jet. Therefore, the jet is split up into sub-jets including the main energy

depositions. The identification of relatively symmetric splittings in a jet points to the decay of a heavy particle. In this analysis, the first splitting scale of the k_T-algorithm $\sqrt{d_{12}}$ [94] is used. It is calculated from the momentum and the distance of the two remaining jet constituents before the final k_T-algorithm merging step. This sub-structure variable is approximately half the mass of the initial particle in case of a heavy particle decay. The largely asymmetric splittings of quarks and gluons in misidentified $R = 1.0$ jets tend to have smaller values $\sqrt{\mathrm{d}_{12}} \approx 10$ GeV with extensions also to high $\sqrt{d_{12}}$ values. A cut is applied at $\sqrt{\mathrm{d}_{12}} > 40$ GeV.

The b-jet identification inside the fat-jet is more complicated, compared to the identification in narrow jets. A larger number of tracks originated from more than one final-state particle reduces the efficiency of the default tagging algorithms. Since there is no b-tagging algorithm available yet, calibrated for the b-jet identification using fat-jets as b-tag seeds, the b-tagging in the boosted selection relies on reconstructed $R = 0.4$ jets with a p_{T} larger than 25 GeV. Therefore, at least one $R = 0.4$ jet in the event is required to be b-tagged, using the $MV1$ b-tagging algorithm, identical to the resolved selection scheme.

Event Selection

The event signature of a boosted top-quark decay consists of one high-momentum fat-jet, at least one $R = 0.4$ jet and a high p_{T} lepton. The event selection starts with events that pass a $R = 1.0$ jet trigger. As in the resolved analysis, exactly one isolated lepton with a $\mathrm{p_T}$ >25 GeV with an isolation criterion of $I^l_{mini}/{p_{\mathrm{T}}}^l < 0.05$ is required.

The $R = 0.4$ jet, assigned to the leptonically decaying top-quark, has to fulfill the same selection criteria as used in the resolved reconstruction. Since the jet is assigned to the leptonically decaying top-quark, the ΔR separation to the lepton has to be smaller than $\Delta R = 1.5$. The decay products from the two top-quarks are required to be in different hemispheres. Therefore, a cut on the angle $\Delta\phi$ between the lepton and the fat-jet $\Delta\phi > 2.3$ is applied. In addition, a cut on the ΔR between the two jets $\Delta R > 1.5$ guarantees that there is no energy-overlap among the two jets.

Apart from these differences, all event-selection requirements are equal to the resolved selection scheme described above.

Background Estimation

The background estimation in the boosted selection is done analogously to the estimation in the resolved selection scheme. Most of the background processes are determined from MC simulation. The multi-jet background and the W+jets background contribution are determined data-driven.

The fraction of multi-jet background is estimated from data using the matrix method. The defined control regions are discussed in detail in [89].

The W+jets background normalization and heavy-flavor content are estimated with the methods shown in Chapter 4.4. Since the number of events passing the event selection is very low,the event-selection requirements have to be modified. To increase the size of the sample, the b-tagging requirement, the jet mass cut and the $\sqrt{d_{12}}$ requirement are not applied in the calculation of the W+jets normalization. In addition, also the p_{T} cut on the fat-jet is reduced to $\mathrm{p_T} > 300$ GeV.

$m_{t\bar{t}}$ Reconstruction

The detection of the objects in case of the boosted selection has less ambiguities compared to the assignment in resolved events. Here, the fat-jet is assigned to the hadronically decaying top-quark, while the $R = 0.4$ jet with the smallest distance to the lepton is associated to the leptonically decaying top-quark.

Figure 6.2 shows distributions of some variables for the boosted analysis. The distributions show the transverse momentum of the $R = 1.0$ jet, which is assigned to the hadronic top-quark decay, and the reconstructed invariant $t\bar{t}$ mass spectra. Both distributions, shown for the electron and the muon selection-channel individually, indicate a good agreement between data and the SM expectation within the systematic uncertainty.

6.1.4 Systematic Uncertainties

The final observables of the analysis are the invariant $t\bar{t}$ mass spectra. This observable is afflicted with two categories of systematic uncertainties. There are systematic uncertainties on the reconstructed objects and systematic uncertainties on the modeling of the signal and background contributions. The systematic uncertainties, described in detail in [89], can either affect the shape or the normalization of the spectrum.

This section will address the systematic uncertainties in the resolved selection, which will be of particular interest in the second part of this chapter. Most of the uncertainties on the reconstructed objects are introduced in Section 4.5. The only difference originates from the different choices of the lepton isolation [90].

The uncertainty on the $m_{t\bar{t}}$ modeling of the signal and background contributions has different sources. The largest effect in the resolved selection is associated with $t\bar{t}$ cross-section uncertainty [89], already discussed in Section 4.5. The impact on the event yield of the total background is 8.2% [89]. The uncertainty on the PDF leads to the second largest effect with 4.7% [89].

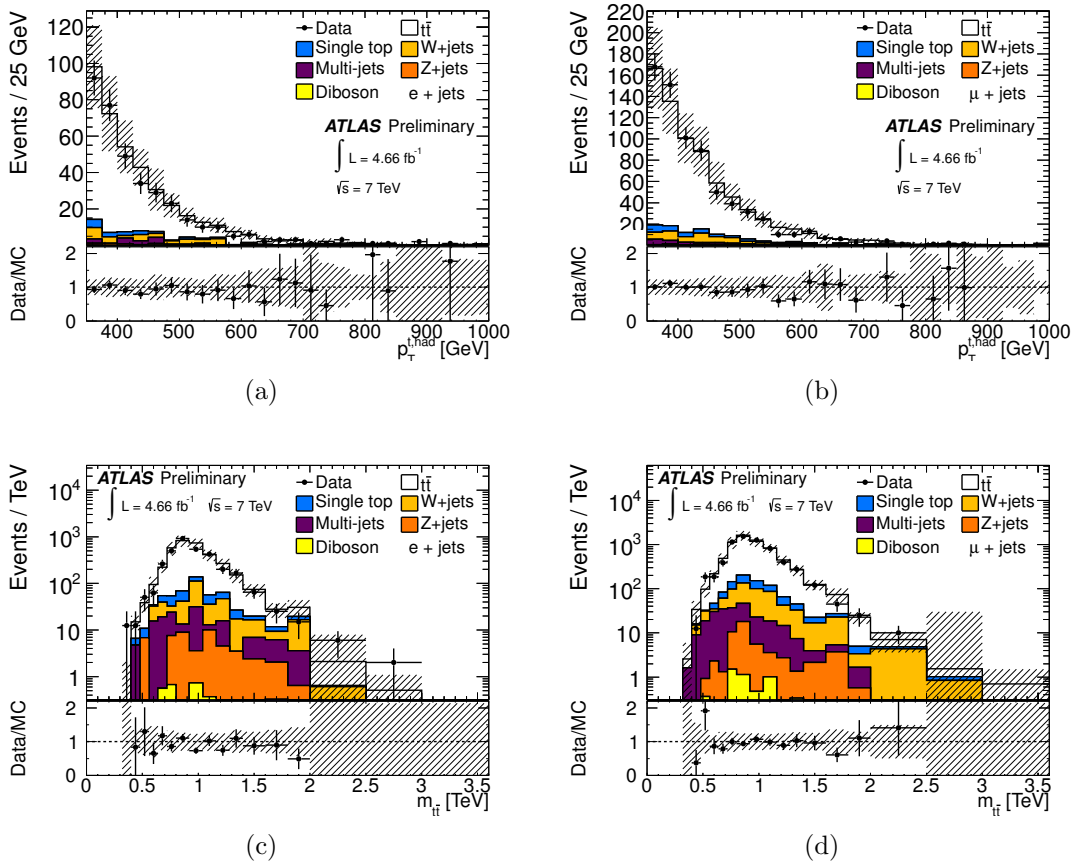


Figure 6.2: *Transverse momentum of the $R = 1.0$ jet, assigned to the hadronically decaying top, and the reconstructed invariant $t\bar{t}$ mass for the boosted analysis. The left side shows the distributions for events containing an electron and on the right, events containing a muon are shown. The shaded areas indicate the total systematic uncertainty. [89].*

The W+jets normalization is of particular interest in this thesis. Within the analysis of the systematic uncertainties, the normalization uncertainty of the W+jets background has an impact of 1.3% and, therefore, it is placed at the tenth position of the most significant systematic effects for both, the resolved and the boosted selection scheme [89].

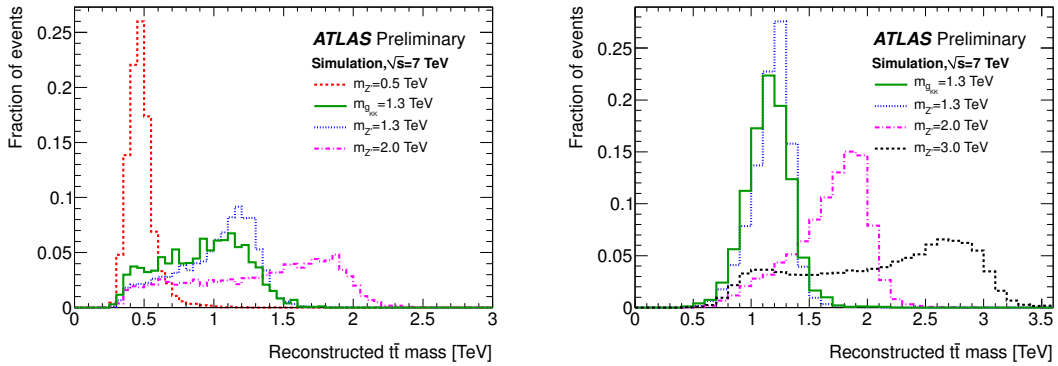
6.1.5 Signal and Background Samples

The continuous background in the $m_{t\bar{t}}$ spectra is modeled by the same samples as used in the SM top-quark analysis and described in Chapter 4.2.2 for both selection schemes. In addition, samples for the considered benchmark models are simulated. The samples of Z' boson production are generated by Pythia, including all top-quark decay modes. The Kaluza-Klein gluon samples are generated with Mad-

GRAPH [95]. The parton shower is computed with PYTHIA. Both generators use the $CTEQ6L1$ PDFs [61]. The generated resonance masses and their corresponding cross-sections are shown in Table 6.1. Figure 6.3 shows the invariant mass distribution of the simulated samples in the resolved and the boosted selection scheme.

Mass [GeV]	$Z' \to t\bar{t}$ $\sigma \times BR$ [pb]	$g_{KK} \to t\bar{t}$ $\sigma \times BR$ [pb]
500	19.6	81.3
1000	1.2	4.1
1500	0.13	0.50
2000	0.019	0.095
2500	0.0030	0.026
3000	0.00097	0.0097

Table 6.1: *Production cross-section times branching fraction for the resonant signal process $pp \to Z' \to t\bar{t}$ in the topcolor model and $pp \to g_{KK} \to t\bar{t}$ for the Kaluza Klein gluon in Randall-Sundrum models with warped extra dimensions [96].*



(a) (b)

Figure 6.3: *Reconstructed $t\bar{t}$ invariant mass using the (a) resolved and (b) boosted event selection. Distributions are shown for different Z' masses and a broad Kaluza-Klein gluon resonance at a mass of* 1.3 TeV *[89].*

6.1.6 Event Yields

Since the combination of the boosted and resolved analyses is intended, the analyzed dataset is split into two orthogonal datasets. The main interest is set to events with a high invariant $t\bar{t}$ mass. Therefore, events that pass the boosted event selection are preferred. Only these events that fail the boosted event selection but pass the resolved event selection are considered in the resolved analysis. The number of

selected events in the resolved and the boosted selection schemes after the removal of overlapping events are listed in Table 6.2 along with the systematical uncertainties. A reasonable agreement between data and the expected background processes is observed. The uncertainties on the number of events in the $t\bar{t}$ continuum are in the order of 10% for both selections. Within these uncertainties, an indication of additional $t\bar{t}$ production beyond the SM is not expected to be visible in the observed event yields.

Type	Resolved selection	Boosted selection
$t\bar{t}$	44,000 ± 4,700	950 ± 100
Single top	3,250 ± 250	49 ± 4
Multi-jets e	2,500 ± 1500	12 ± 7
Multi-jets μ	1,010 ± 610	20 ± 12
W+jets	6,940 ± 730	82 ± 15
Z+jets	840 ± 410	11 ± 5
Di-boson	124 ± 43	0.9 ± 0.3
Total	58,700 ± 5,300	1,120 ± 100
Data	61,954	1,079

Table 6.2: *Selected data and expected background event yields after the resolved and boosted selection [89]. The overlap of events in the resolved reconstruction, which also pass the boosted reconstruction scheme, is removed.*

6.2 Statistical Analysis

Since the signal of the investigated resonance is assumed to appear as a bump in the reconstructed invariant $t\bar{t}$ mass spectra, data and expected background distributions are compared in this observable. Therefore, the reconstructed invariant mass spectra of the four independent selections, corresponding to the resolved and boosted selections in the electron and muon decay channels, are added. The combined $m_{t\bar{t}}$ spectrum is shown in Figure 6.4. The study of $t\bar{t}$ events produced in processes beyond the SM is evaluated in two steps. In the first step, the compatibility of the data with the SM-only hypothesis is evaluated. In the absence of deviations, an upper cross-section limit for the two signal benchmark models is computed in a second step.

6.2.1 Compatibility with the Null Hypothesis

The *BumpHunter* [97] is a hypothesis-testing tool that searches for local excesses or deficits in data compared to the expected backgrounds. The $m_{t\bar{t}}$ spectra are analyzed

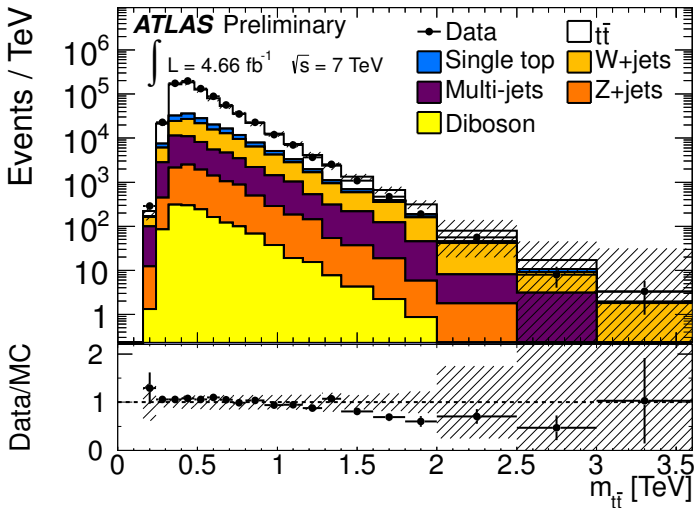


Figure 6.4: *Reconstructed invariant $t\bar{t}$ mass, adding the two channels and selection methods. The shaded areas indicate the total systematic uncertainty. [89].*

by a sliding window of variable size with a minimum width of two bins. The most prominent deviation corresponding to the smallest Poisson probability is analyzed. Pseudo experiments are used to validate the observed deviation. In addition to the analysis of the combined spectra, the two spectra of the two lepton flavors for the resolved and boosted analysis are tested individually. Since it is expected, that a deviation is observed in all four spectra at the same position, a comparison of the observed most significant window interval gives additional significance. With this approach, the look-elsewhere effect is taken into account. The most significant deviation was found in the mass range of $400 - 560$ GeV in the boosted μ+jets spectra, shown in Figure 6.1(d), with a significance of 0.88 σ [89]. Since no significant deviation in the analyzed $m_{t\bar{t}}$ mass spectra could be found, there is no indication of $t\bar{t}$ resonances in the dataset analyzed.

6.2.2 Determination of an Upper Cross Section Limit

The calculation of an upper cross-section limit on a specific theoretical model using a given sample of data provides a point of intersection between theory predictions and the experimental data. Given that no significant excess is observed, upper limits on the cross-section times branching ratio are set for the considered Z' boson and Kaluza-Klein benchmark models. Since these benchmark model are chosen to characterize a narrow and a broad resonance with respect to the experimental mass resolution, the cross-section limit is nearly independent from the consulted model. This section gives a short introduction of the applied limit setting while a more detailed description can be found in [98].

A tool developed by the $D\emptyset$ collaboration [99], based on a Bayesian technique, is used to calculate the exclusion limits from the binned signal and background distributions. Therefore, a likelihood function L_v, depending on the signal cross-section σ_v for a particular resonance with the mass v, is defined as [89]:

$$L_v(D|\sigma_v, a_v, b) = \prod_{i=1}^{N} \frac{e^{-(a_{v,i}\sigma_v + b_i)}(a_{v,i}\sigma_v + b_i)^{D_i}}{\Gamma(D_i + 1)} \tag{6.4}$$

where D is the number of data events, b is the sum of all expected backgrounds and a_v is the acceptance times luminosity for the signal process. The index i addresses all bins N of the spectra. The gamma function $\Gamma(D_i+1)$ reduces to the faculty $D_i!$ if D_i is an integer. To retrieve a likelihood value for the whole distribution, the likelihoods of the single bins are multiplied. Using Bayes theorem with a positive flat prior in the signal cross-section, the posterior probability density $p(\sigma, a, b|D)$ is calculated [99]. Systematic uncertainties are incorporated by smearing the parameters of the Poisson distribution in each bin. A Gaussian prior controls the the probability for a given deviation of the parameter from the nominal value for each systematic effect. The 95% confidence level upper limit is calculated by integrating the posterior probability up to 95%. The one- and two-sigma bands around the expected limit are determined from pseudo-experiments. For each pseudo-experiment, an expected upper limit on the signal cross-section is calculated to estimate the sensitivity of the limit setting. The median of this distribution is taken as the central value and the 16% (84%) and 2.5% (97.5%) quantils as one and two σ band.

6.2.3 Application of the W+jets Estimation based on the R_N Method

The predictive power of the limit setting is not only dependent on the available data statistics. The good knowledge of the background processes is crucial in the determination of an exclusion limit. For example, a reduction of the systematic uncertainties in the background expectations improves the significance of the analysis. The second-most dominant background process to be considered in the analysis is the W boson production in association with jets. The numbers of selected events, shown in Table 6.2, indicate a W+jets background fraction of about 11% in the resolved and about 8% in the boosted analysis. The relative impact of the systematic uncertainty related to the W+jets background yield is 1.3% and 1.5%, respectively [89].

In Chapter 5.3, a new method developed to determine the W+jets normalization is described. By means of this method a set of W+jets normalization scaling factors is determined. These scaling factors also can be applied in the analysis of beyond SM $t\bar{t}$ production. An application of the N_{jet} ratio method in the boosted event topology is not yet investigated. Therefore, this section focuses only on the resolved selection-scheme of the analysis.

Determination of Cross-Section Limits

The W+jets normalization in the resolved analysis scheme is based on the scaling factors, recommended to be applied in the analysis of SM $t\bar{t}$ production by the ATLAS collaboration. Since the W+jets normalization, determined with the N_{jet} ratio method, is based on the same requirements, both W+jets normalization methods can be compared in terms of the determined exclusion limits. The scaling factors determined by the two methods and their uncertainties are shown in Table 5.20 on page 110.

Within the estimation of an exclusion limit based on the new W+jets normalization, the scaling factor of the W+jets background for events with at least four selected jets is exchanged with the one from the N_{jet} ratio method . The systematic uncertainties on the W+jets normalization are adapted. Based on the new background expectation, the limit on the cross-section times branching ratio is recalculated. The upper cross-section limits are given with systematic and statistical uncertainties for both determined W+jets normalizations in Figure 6.5. The results based on the recommended W+jets normalization are shown on the left while the result of the N_{jet} ratio method is applied in the histograms on the right.

In the column of Figure 6.5, the W+jets background contribution is normalized with the charge-asymmetry method, while on the right the W+jets scaling factors of the N_{jet} ratio method are applied. The expected cross-section of the two benchmark models is drawn as reference. In Tables 6.3 and 6.4, the numerical values of the upper cross-section limits are shown. In the comparison of both W+jets normalization methods in the limit setting, one obtains a nearly identical observed and expected limit. Within the uncertainties no difference could be found. The progression of the cross-section limit is rather flat towards higher masses because only the resolved spectrum without the boosted events is used.

The results confirm the good agreement of the determined scaling factors, calculated with the two independent methods. As already discussed, the R_n method predicts a slightly higher W+jets normalization, especially for the muon-selection channel. In addition the systematic uncertainty on the W+jets normalization is slightly reduced. Thus, the expected limits at low masses are slightly higher. The observed limits are closer to the expected limits. The intended combination of both methods will lead to a significantly reduced systematic uncertainty on the W+jets estimate. Since the resulting improvement in the cross-section limit is mainly expected in the lower $m_{t\bar{t}}$ region, a significantly improvement of the determined mass limits on the benchmark models is not expected.

	CA normalization				N_{jet} ratio normalization			
Mass (TeV)	Obs.	Exp.	-1σ	$+1\sigma$	Obs.	Exp.	-1σ	$+1\sigma$
			(pb)				(pb)	
500	4.28	6.72	3.68	11.18	4.01	6.86	3.65	11.04
600	5.48	2.84	1.66	4.47	5.80	2.82	1.61	4.53
700	3.12	1.86	1.07	2.99	3.45	1.92	1.11	3.04
800	2.00	1.63	0.95	2.64	1.57	1.64	0.94	2.66
1000	1.78	1.60	0.93	2.57	1.80	1.60	0.93	2.60
1300	2.04	1.44	0.83	2.32	1.86	1.48	0.85	2.37
1600	1.21	1.32	0.75	2.12	1.27	1.34	0.76	2.15
2000	0.55	1.01	0.57	1.60	0.52	1.03	0.57	1.62
3000	0.36	0.75	0.45	1.23	0.40	0.76	0.46	1.27

Table 6.3: *Upper 95% CL cross-section limit times branching ratio on a leptophobic topcolor Z' boson decaying to a $t\bar{t}$ pair calculated in the resolved selection scheme. The left and the right columns show the limit according to a W+jets normalization based on the charge asymmetry and the N_{jet} ratio method. The observed and expected limits for each mass point are given, as well as the $\pm 1\sigma$ variation*

	CA normalization				N_{jet} ratio normalization			
Mass (TeV)	Obs.	Exp.	-1σ	$+1\sigma$	Obs.	Exp.	-1σ	$+1\sigma$
			(pb)				(pb)	
700	5.18	2.89	1.67	4.67	4.66	2.89	1.66	4.56
800	2.45	2.25	1.28	3.56	2.34	2.25	1.27	3.53
900	3.45	2.36	1.33	3.86	3.39	2.36	1.34	3.81
1000	3.47	2.26	1.28	3.72	3.22	2.26	1.28	3.73
1150	3.72	2.30	1.33	3.76	2.38	2.31	1.31	3.78
1300	3.45	2.20	1.26	3.54	3.26	2.18	1.23	3.55
1600	2.39	2.20	1.24	3.61	2.39	2.20	1.21	3.59
1800	1.50	1.99	1.12	3.18	1.53	2.04	1.14	3.26
2000	1.23	2.14	1.17	3.52	1.17	2.15	1.18	3.46
2250	1.16	1.86	1.06	3.08	1.37	1.85	1.06	3.04

Table 6.4: *Upper 95% CL cross-section limit times branching ratio on a Kaluza-Klein gluon decaying to a $t\bar{t}$ pair calculated in the resolved selection scheme. The left and the right columns show the limit according to a W+jets normalization based on the charge asymmetry and the N_{jet} ratio method. The observed and expected limits for each mass point are given, as well as the $\pm 1\sigma$ variation*

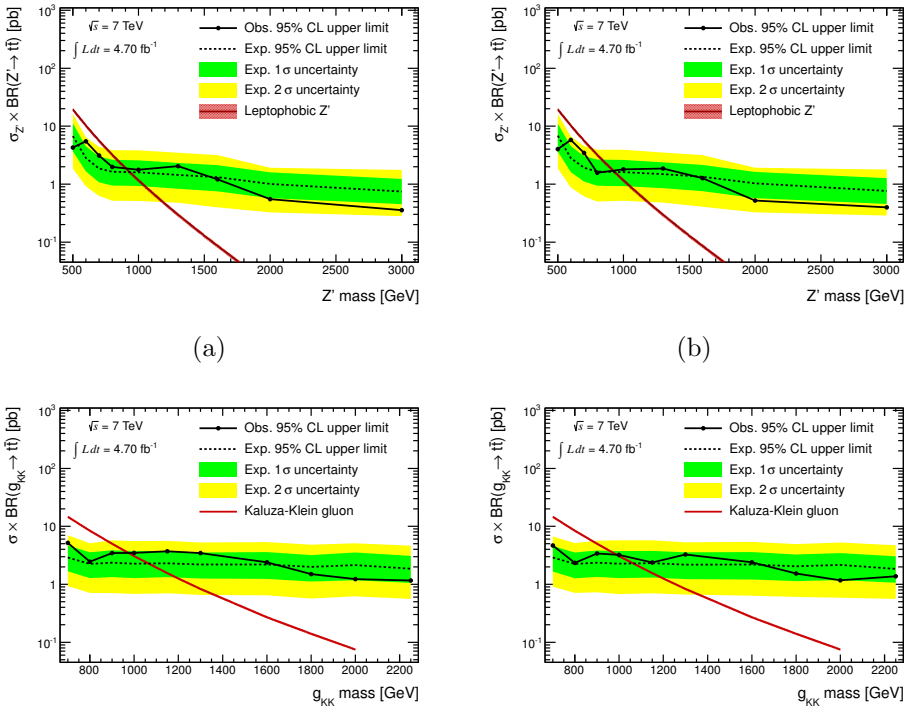
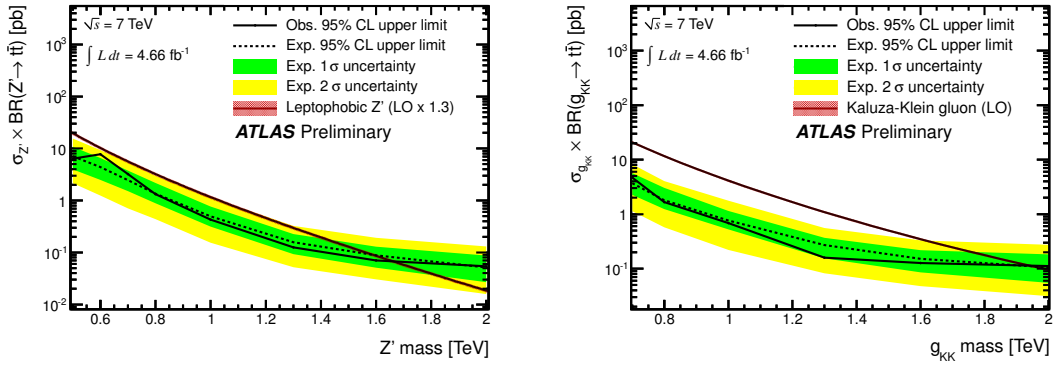


Figure 6.5: *Expected and observed upper cross-section times branching ratio limits based on the Z' and Kaluza-Klein benchmark model. The estimation of the limit is done only in the resolved selection. Figurex (a) and (c) show the estimated limit with the recommended W+jets normalization while Figures (b) and (d) show the estimated limit with the W+jets normalization based on the N_{jet} ratio method . Both systematic and statistic uncertainty are included.*

6.2.4 Upper Cross Section Limits on $t\bar{t}$ Resonances

This section summarizes the final results based on the combination of both selection schemes, presented in [89]. The expected cross-section times branching ratio for the two benchmark models is shown in Table 6.1 in dependence on the mass. For each of the models investigated, a 95% upper limit is set. The limits are calculated for the two selection schemes separately and combined. In Figure 6.6, the expected and observed upper cross-section limits are shown for the two benchmark models combining both selection schemes. At a 95% confidence level, the leptophobic topcolor Z' boson with a mass between 0.5 and 1.7 TeV and a Kaluza Klein gluon in a mass range of 0.6 to 1.9 TeV are excluded [89].



(a) (b)

Figure 6.6: *Expected and observed upper cross-section times branching ratio limits on the (a) Z' and (b) KK gluon benchmark models. The resolved and boosted selections have been combined in the estimation of the limit. Both systematic and statistic uncertainties are included. The shaded area on the theoretical cross-section corresponds to the PDF uncertainty on the prediction [89].*

6.3 Summary

This chapter introduces a search for $t\bar{t}$ resonances in the lepton plus jets decay-channel. Due to different event topologies at high and low $m_{t\bar{t}}$, two methods are applied. For the resolved selection, the hadronic top-quark decay is reconstructed in two or three $R = 0.4$ jets while in the boosted selection one $R = 1.0$ jet is reconstructed. The analyzed observable, expected to be sensitive on $t\bar{t}$ production beyond the SM, is the invariant $t\bar{t}$ mass. Since no excess of events produced in addition to the SM expectation is observed, an upper limit on the cross-section times branching ratio is set on two benchmark models. With regard to the determined cross-section limits, the existence of a leptophobic Z' in the mass range of 0.5 TeV to 1.7 TeV and a Kaluza-Klein gluon for masses between 0.6 TeV to 1.9 TeV are excluded at 95% CL.

It has been shown that the W+jets background normalization, determined in Chapter 5.3, leads to the same results in the estimation of the upper cross-section limit. An improvement of the cross-section limits of $t\bar{t}$ resonances is expected with the combination of both W+jets normalization results.

Conclusion and Outlook

In the year 2011, the ATLAS experiment has recorded pp collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV. The generated dataset with an integrated luminosity of $5.3\,\mathrm{fb}^{-1}$ is the basis of many physical analyses. The top quark with its high mass plays an important role in many of the processes at the LHC. The considered semileptonic $t\bar{t}$ decay, with an event signature of four jets, an isolated high p_{T} lepton and missing transverse momentum, combines an adequate branching ratio with a characteristic event signature.

The signal and most of the background processes are modeled using simulated events. One of the dominant background contributions arises from W+jets production in association with jets. Large uncertainties in the simulation of the W+jets background in association with at least four jets motivate the application of data-driven methods. The methods, currently recommended by the ATLAS top-quark analysis group, scale the predicted W+jets rate and W+jets heavy-flavor content data-driven while the kinematic shape of the W+jets processes is taken from MC simulations. Therefore, the estimation of the W+jets background strongly relies on properties extracted from simulated events.

Within this thesis, a W+jets background estimation directly from data was presented. The basic idea is that W+jets and Z+jets events have many similarities. It was shown in this thesis that both the production and the decay mechanism of W and Z bosons are very similar. Particularly the generation of jets associated to the boson production is identical. Differences in the distribution of the final state leptons are well understood and can be modeled. A template modeling W+jets events was constructed by the conversion of Z+jets events. It was shown that the kinematic properties in W+jets and converted Z+jets events agree very well. Finally, a W+jets template determined from Z+jets data events and, therefore, independent of modeling and reconstruction uncertainties, could be obtained.

Furthermore, a new approach to estimate the W+jets background rate in the semileptonic $t\bar{t}$ decay was presented within this thesis. It bases upon the fact that the production mechanism of additional jets for both W and Z boson events are identical, which leads to a W and Z boson ratio constant as a function of the jet multiplicity. The W and Z boson ratio in the low-jet multiplicities can be determined from data. The extrapolation of the ratio to the number of four selected jets

allows a determination of the W and Z boson ratio in the four-jet bin, almost independent of the $t\bar{t}$ signal process. The number of W+4 jets events was determined by the calculation of the product of the extrapolated ratio R_4^{extrap} and the number of Z+4 jets events, determined in data. The determined rate of W+jets events is in good agreement with the official numbers of the top-quark analysis group. While the systematic uncertainties in the pre-tagged determination are similar, an 4% improvement of the W+jets normalization uncertainty was reached by a selection with a b-tag requirement applied. It is the first estimation of a W+4 jets background normalization by means of Z+jets events performed at a hadron collider ever.

The result of the W+jets background rate is tested in the analysis of $t\bar{t}$ production beyond the SM. Within the comparison of both W+jets normalization methods, the results of the novel method show the same performance in terms of an exclusion limit on the cross-section of resonant $t\bar{t}$ production. Since both methods are based on independent approaches, their uncertainties are mainly uncorrelated. Therefore, a combination of both methods would lead to a further improvement of the uncertainty on the W+jets background normalization in the $t\bar{t}$ analysis. This improvement is expected to show benefits in the analysis of resonant $t\bar{t}$ production. Furthermore, an improved W+jets normalization would allow a more precise study of SM $t\bar{t}$ production. A possible example of an application is the measurement of the $t\bar{t}$ cross-section.

Within this thesis two methods were developed to estimate the W+jets background by means of Z+jets events. Since one method produces a template of the W+jets kinematic shape while the other one predicts the W+jets rate, the application of both methods would lead to a W+jets background description completely independent from the simulation of W+jets events. Since the number of Z+jets events in data is limited due to a smaller production cross-section, compared to the cross-section of W+jets production, the statistical uncertainty of the shown results is not negligible in both methods. A dataset corresponding to an integrated luminosity of 21.7 fb^{-1} and with an increased center-of-mass energy of $\sqrt{\text{s}} = 8$ TeV was already recorded in the year 2012. The application of the developed methods in a analysis based on this dataset has already started.

Danksagung

Ich möchte mich für die Unterstützung im Laufe meiner Doktorarbeit bei meinem Doktorvater Prof. Dr. Mättig, Prof. Dr. Wolfgang Wagner, meiner Arbeitsgruppe und meinen Kollegen am CERN ganz herzlich bedanken. Die Eindrücke, Erfahrungen und Freundschaften dieser Zeit möchte ich nicht missen!

Ein ganz besonderer Dank für die Aufmunterung, gerade in schwierigen Zeiten der Doktorarbeit, geht an meine Familie und meine Freunde. Dabei muss ich die tolle Unterstützung meiner Frau und die hervorragende Ablenkung meiner Kinder gesondert herausheben. Nur so war es mir möglich, auch das wirklich Wichtige im Leben nicht aus den Augen zu verlieren.

Bibliography

[1] P. D. Group, *Extracted from the Review of Particle Physics*, Phys. Rev. D 86, 010001 (2012) .

[2] S. L. Glashow., *Partial Symmetries of Weak Interactions*, Nucl. Phys., 22:579-588 (1961) .

[3] S. W. Jeffrey Goldstone, Abdus Salam, *Broken Symmetries*, Phys. Rev., 127:965-970 (1962) .

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

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

[6] H. Georgi and S. L. Glashow, *Unified Weak and Electromagnetic Interactions without Neutral Currents*, Phys. Rev. Lett. 28, 1494 (1972) .

[7] H. L. H. Fritzsch, M. Gell-Mann, *Advantages of the Color Octet Gluon Picture*, Phys. Lett. B47 (1973) 365–368.

[8] S. Weinberg, *What is Quantum Field Theory, and What Did We Think It Is?*, `arXiv:hep-th/9702027`.

[9] F. Halzen and A. D. Martin, *Quarks and Leptons: An introductory course in Modern Particle Physics*, QC793.5.Q2522H34 (1984) .

[10] D. H. Perkins, *Introduction to High Energy Physics*, QC93.2.P47, 4th edition (1999) .

[11] Particle Data Group Collaboration, K. Nakamura, *Review of Particle Physics*, Journal of Physics G: Nuclear and Particle Physics 37, no. 7A (2010) 075021.

[12] The ATLAS Collaboration, *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 CERN-PH-EP-2012-218, `arXiv:1207.7214 [hep-ex]`.

[13] CMS Collaboration Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys.Lett. **B716** (2012) 30–61 CMS-HIG-12-028, CERN-PH-EP-2012-220, arXiv:1207.7235 [hep-ex].

[14] CDF Collaboration Collaboration, F. Abe et al., *Observation of top quark production in $p\bar{p}$ collisions*, Phys.Rev.Lett. **74** (1995) 2626–2631 FERMILAB-PUB-95-022-E, CDF-PUB-TOP-PUBLIC-3040, ANL-HEP-PR-95-44, arXiv:hep-ex/9503002 [hep-ex].

[15] M. A. et al., *HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR*, Comput. Phys. Commun., 182:1034-1046 (2011) .

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

[17] N. Kidonakis, *NNLL resummation for s-channel single top quark production*, Phys. Rev., D81:054028 (2010) .

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

[19] CMS Collaboration Collaboration, *LHC Combination: Top mass*, CMS-PAS-TOP-12-001.

[20] E. Eichten and K. Lane, *Dynamical breaking of weak interaction symmetries*, Physics Letters **B90** (11 February 1980) 125–130.

[21] C. T. Hill and S. J. Parke, *Top production: Sensitivity to new physics*, Phys.Rev. **D49** (1994) 4454–4462 FERMILAB-PUB-93-397-T, arXiv:hep-ph/9312324 [hep-ph].

[22] C. T. Hill, *Topcolor assisted technicolor*, Phys.Lett. **B345** (1995) 483–489 FERMILAB-PUB-94-395-T, arXiv:hep-ph/9411426 [hep-ph].

[23] R. M. Harris and S. Jain, *Cross Sections for Leptophobic Topcolor Z' Decaying to Top-Antitop*, Eur.Phys.J. **C72** (2012) 2072 FERMILAB-FN-0940-CMS-PPD, arXiv:1112.4928 [hep-ph].

[24] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, Phys.Rev.Lett. **83** (1999) 3370–3373 MIT-CTP-2860, PUPT-1860, BUHEP-99-9, arXiv:hep-ph/9905221 [hep-ph].

[25] K. Agashe, A. Belyaev, T. Krupovnickas, G. Perez, and J. Virzi, *LHC Signals from Warped Extra Dimensions*, Phys.Rev. **D77** (2008) 015003 BNL-HET-06-13, MSUHEP-060915, SU-4252-838, YITP-SB-06-43, arXiv:hep-ph/0612015 [hep-ph].

[26] A. F. R. e. a. Wenninger J, *Operation of the LHC at High luminosity and High stored energy*, CERN-ATS-2012-101 (2012) .

[27] The ATLAS Collaboration, *ATLAS images*, atlas.ch (2012) .

[28] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST (August 2008) .

[29] The ATLAS Collaboration, *ATLAS Muon Momentum Resolution in the First Pass Reconstruction of the 2010 p-p Collision Data at sqrts* $= 7$ TeV, ATLAS-CONF-2011-046 (2011) .

[30] The ATLAS Collaboration, *ATLAS Forward Detectors for Luminosity Measurement and Monitoring*, CERN/LHCC/2004-010 **LHCC I-014** (2004) .

[31] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, and A. Salzburger, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, Tech. Rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar, 2007.

[32] The ATLAS Collaboration, *Vertex Reconstruction in the ATLAS Experiment at the LHC*, ATLAS-CONF-2009-001 (2009) .

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

[34] The ATLAS Collaboration, *Charged-particle multiplicities in pp interactions at* $\sqrt{s}$ *= 900 GeV measured with the ATLAS detector at the LHC*, Phys.Lett.B688:21-42,2010 (2010) , `arXiv:1003.3124`.

[35] S. T.Sjostrand and P. Skands, *A brief introduction to Pythia 8.1. Comput. Phys. Commun.*, JHEP **178:852-867** (2008) .

[36] J. P. . R. Armenteros, *Analysis of V-events*, Philosophical Magazine Series 7 (1954) .

[37] The ATLAS Collaboration, *Study of the Material Budget in the ATLAS Inner Detector with* K^0_s *decays in collision data at* $\sqrt{s} = 900$ GeV, ATLAS-CONF-2010-019 (2010) .

[38] The ATLAS Collaboration, *Expected electron performance in the ATLAS experiment*, ATLAS NOTE **ATL-PHYS-PUB-2011-006** (2011) .

[39] The ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, `arXiv:1110.3174`.

[40] The ATLAS Collaboration, *Object selection and calibration, background estimations and MC samples for the Winter 2012 Top Quark analyses with 2011 data*, ATLAS NOTE **ATL-COM-PHYS-2012-224** (2012) .

[41] The ATLAS Collaboration, *Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector*, .

[42] The ATLAS Collaboration, *Object selection and calibration, background estimations and MC samples for the Summer 2012 Top Quark analyses with 2011 data*, ATLAS Note **ATL-COM-PHYS-2012-499** (2012) .

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

[44] The ATLAS Collaboration, M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP 0804:063,2008 (2008) , arXiv:0802.1189.

[45] The ATLAS Collaboration, *Missing et in rel17*, https://twiki.cern.ch/twiki/bin/view/Main/Rel17MissingEt **ATLAS TWiki** (2012) .

[46] T. S. M.L. Mangano, *Tools for the simulation of hard hadronic collisions*, Annu. Rev. Nucl. Part. Sci. 55 **10.1146/annurev.nucl.55.090704.151505** (2005) .

[47] 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].

[48] M. A. Dobbs, S. Frixione, E. Laenen, K. Tollefson, H. Baer, E. Boos, B. Cox, R. Engel, W. Giele, J. Huston, S. Ilyin, B. Kersevan, F. Krauss, Y. Kurihara, L. Lonnblad, F. Maltoni, M. Mangano, S. Odaka, P. Richardson, A. Ryd, T. Sjostrand, P. Skands, Z. Was, B. R. Webber, and D. Zeppenfeld, *Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics*, arXiv:hep-ph/0403045.

[49] M. M. et al., *Alpgen, a generator for hard multiparton processes in hadronic collisions.*, JHEP **0307:001** (2003) .

[50] G. C. et al., *Herwig 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **0101:010** (2001) .

[51] S. Frixione and B. Webber, *Matching nlo qcd computations and parton shower simulations*, JHEP **0206:029** (2002) .

[52] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton Interactions in Photoproduction at HERA*, arXiv:hep-ph/9601371.

[53] G. I. B. Andersson, G. Gustafson and T. Sjostrand, *Parton Fragmentation and String Dynamics*, Phys. Rept. 97 **31** (1983) .

[54] T. G. et al., *Event generation with sherpa 1.1.*, JHEP **0902:007** (2009) .

[55] R. K. S. Catani, F. Krauss and B. R. Webber, *QCD Matrix Elements + Parton Showers*, arXiv:hep-ph/0109231.

[56] The ATLAS Collaboration, "", *The ATLAS Simulation Infrastructure.*, Eur.Phys.J.C70:823-874,2010 (2010) , arXiv:1005.4568.

[57] S. A. et al., *Geant4-a simulation toolkit.*, NIM A 506 (2003) .

[58] The ATLAS Collaboration, P. Jenni, M. Nessi, M. Nordberg, and K. Smith, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*, .

[59] The ATLAS Collaboration, *Lepton trigger and identification for the Winter 2011 top quark analyses*, Tech. Rep. ATL-COM-PHYS-2011-123, CERN, Geneva, Feb, 2011. Supporting document for Winter 2011 top physics measurements.

[60] The ATLAS Collaboration, *Luminosity Public Results*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults (2011) .

[61] P. M. Nadolsky, H. L. Lai, Q. H. Cao, J. Huston, J. Pumplin, D. Stump, W. K. Tung, and C. P. Yuan, *Implications of CTEQ global analysis for collider observables*, Phys.rev.d78:013004,2008, 2008. arXiv:0802.0007.

[62] The ATLAS Collaboration, *Twiki page: Heavy Flavor Overlap Removal Tool*, https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HforTool .

[63] The ATLAS Collaboration, *Single Boson and Diboson Production Cross Sections in pp Collisions at* $\sqrt{s}$ *= 7 TeV*, ATL-COM-PHYS-2010-695 (2010) .

[64] The ATLAS Collaboration, *Mis-identified lepton backgrounds in top quark pair production studies for EPS 2011 analyses*, ATLAS Note **ATL-COM-PHYS-2011-768** (2011) .

[65] J. Butterworth, E. Dobson, U. Klein, B. Mellado Garcia, T. Nunnemann, J. Qian, D. Rebuzzi, and R. Tanaka, *Single Boson and Diboson Production Cross Sections in pp Collisions at sqrts=7 TeV*, Tech. Rep. ATL-COM-PHYS-2010-695, CERN, Geneva, Aug, 2010.

[66] C. F. Berger, Z. Bern, L. J. Dixon, F. F. Cordero, D. Forde, T. Gleisberg, H. Ita, D. A. Kosower, and D. Maitre, *Precise Predictions for W + 4 Jet Production at the Large Hadron Collider*, Phys.rev.lett.106:092001,2011, 2010. arXiv:1009.2338.

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

[68] The ATLAS Collaboration, B. A. et al., *Object selection and calibration, background estimations and MC samples for the Autumn 2012 Top Quark analyses with 2011 data*, Tech. Rep. **ATL-COM-PHYS-2012-1197** (Aug, 2012) .

[69] The ATLAS Collaboration, *Jet energy scale and its systematic uncertainty in proton-proton collisions at sqrt(s)=7 TeV with ATLAS 2011 data*, ATLAS-CONF-2013-004.

[70] The ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at sqrt(s) = 7 TeV recorded in 2010 with the ATLAS detector*, arXiv:1210.6210.

[71] The ATLAS Collaboration, *Performance of missing transverse momentum reconstruction in proton-proton collisions at 7 TeV with ATLAS*, arXiv:1108.5602.

[72] The ATLAS Collaboration, *Muon Momentum Resolution in First Pass Reconstruction of pp Collision Data Recorded by ATLAS in 2010*, .

[73] The 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*, .

[74] The ATLAS Collaboration, *b-jet tagging calibration on c-jets containing D^{*+} mesons*, .

[75] The ATLAS Collaboration, *Measurement of the Mistag Rate with 5 fb^{-1} of Data Collected by the ATLAS Detector*, .

[76] The ATLAS Collaboration, *Luminosity Determination in pp Collisions at sqrt(s) = 7 TeV using the ATLAS Detector in 2011*, ATLAS-CONF-2011-116, ATLAS-COM-CONF-2011-130.

[77] The ATLAS Collaboration, *Measurement of ttbar production with a veto on additional central jet activity in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector*, arXiv:1203.5015.

[78] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, S. Forte, A. Glazov, J. Huston, R. McNulty, T. Sjostrand, and R. Thorne, *The PDF4LHC Working Group Interim Recommendations*, arXiv:1101.0538.

[79] J. M. Campbell, J. W. Huston, and W. J. Stirling, *Hard Interactions of Quarks and Gluons: a Primer for LHC Physics*, Rept.prog.phys.70:89,2007, 2006. arXiv:hep-ph/0611148.

[80] M. Sandhoff, *Prospects of a cross section ratio measurement of W and Z events as a function of the jet multiplicity with early ATLAS data*, urn:nbn:de::hbz **468-20100341** (2010) .

[81] J. Bjorken and S. Brodsky Phys. Rev., D1:1416 (1970) .

[82] A. Bannoura, *Studies on a New Method to Estimate W+jets Background in $t\bar{t}$ Events at the ATLAS Experiment*, Master Thesis (2012) .

[83] F. A. Berends, W. Giele, H. Kuijf, R. Kleiss, and W. J. Stirling, *Multi-jet production in W, Z events at $p\bar{p}$ colliders*, Phys.Lett. **B224** (1989) 237.

[84] The ATLAS Collaboration, *Measurement of the production cross section of jets in association with a Z boson in pp collisions at $\sqrt{s}$ =7 TeV with the ATLAS detector*, ATL-COM-PHYS-2012-1639.

[85] M. Vreeswik, *List of all systematic uncertainties present in the Charge Asymmetry method*, private communication (2012) .

[86] The ATLAS Collaboration, *Data driven W+jets background estimation for semi-leptonic $t\bar{t}$ decays in pp collisions at $\sqrt{7}$ TeV using the W/Z-boson N-jet ratio*, ATL-COM-PHYS-2012-1792.

[87] D0 Collaboration, V. Abazov, *Search for ttbar resonances in the lepton plus jets final state in ppbar collisions at sqrts=1.96TeV*, Phys.lett.b668:98-104,2008, 2008. arXiv:0804.3664.

[88] D. Collaboration, *Search for $t\bar{t}$ Resonances in the Lepton+Jets Final State in $p\bar{p}$ Collisions at sqrts=1.96TeV*, D0 note 5882-conf, 2009.

[89] The ATLAS Collaboration, *A search for ttbar resonances in the lepton plus jets final state using 5 fb-1 of pp collisions at sqrt(s) = 7 TeV*, .

[90] K. Rehermann and B. Tweedie, *Efficient Identification of Boosted Semileptonic Top Quarks at the LHC*, Jhep 1103:059,2011, 2010. arXiv:1007.2221.

[91] The ATLAS Collaboration, *Jet mass and substructure of inclusive jets in sqrt(s) = 7 TeV pp collisions with the ATLAS experiment*, arXiv:1203.4606.

[92] T. Chwalek, *Measurement of the W boson helicity in top-antitop quark events with the CDF II experiment*, FERMILAB-MASTERS-2006-04 (2006) .

[93] The ATLAS Collaboration, *A Search for $t\bar{t}$ Resonances in the Lepton Plus Jets Channel using* 2.05 fb^{-1} *of pp Collisions at* $\sqrt{s} = 7$ *TeV*, ATLAS-CONF-2012-029, ATLAS-COM-CONF-2012-022.

[94] The ATLAS Collaboration, *Jet mass and substructure of inclusive jets in* $\sqrt{7}$ TeV *pp collisions with the ATLAS experiment*, arXiv:1203.4606.

[95] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, F. Maltoni, T. Plehn, D. L. Rainwater, and T. Stelzer, *MadGraph/MadEvent v4: The New Web Generation*, Jhep0709:028,2007, 2007. arXiv:0706.2334.

[96] The ATLAS Collaboration, *A search for ttbar resonances in the lepton plus jets final state using* 4.66 fb^{-1} *of pp collisions at* $\sqrt{7}$ TeV, ATLAS-CONF-2012-1753.

[97] G. Choudalakis, *On hypothesis testing, trials factor, hypertests and the BumpHunter*, 2011. arXiv:1101.0390.

[98] T. Lenz, *Prospects for ttbar resonance searches at ATLAS*, .

[99] D0 Collaboration, I. Bertram et al., *A Recipe for the construction of confidence limits*, FERMILAB-TM-2104, D0-NOTE-3476, D0-NOTE-2775-A.

List of Figures

List of Tables

Appendices

Monte Carlo Samples

Table 6.5: *Overview of the Process, matrix element, parton shower, cross section, k factor and number of generated events*

process	ME	parton shower	FE * x-sec [pb]	k-f.	events
$W \to e\nu$ 0p	ALPGEN	HERWIGJIMMY	6921.6	1.2	6952874
$W \to e\nu$ 1p	ALPGEN	HERWIGJIMMY	1304.3	1.2	4998487
$W \to e\nu$ 2p	ALPGEN	HERWIGJIMMY	378.29	1.2	3768632
$W \to e\nu$ 3p	ALPGEN	HERWIGJIMMY	101.43	1.2	1008947
$W \to e\nu$ 4p	ALPGEN	HERWIGJIMMY	25.87	1.2	250000
$W \to e\nu$ 5p	ALPGEN	HERWIGJIMMY	7.0	1.2	69999
$Z \to ee$ 0p	ALPGEN	HERWIGJIMMY	668.32	1.2	6618284
$Z \to ee$ 1p	ALPGEN	HERWIGJIMMY	134.36	1.2	1334897
$Z \to ee$ 2p	ALPGEN	HERWIGJIMMY	40.54	1.2	2004195
$Z \to ee$ 3p	ALPGEN	HERWIGJIMMY	11.16	1.2	549949
$Z \to ee$ 4p	ALPGEN	HERWIGJIMMY	2.88	1.2	149948
$Z \to ee$ 5p	ALPGEN	HERWIGJIMMY	0.83	1.2	50000
$W \to \mu\nu$ 0p	ALPGEN	HERWIGJIMMY	6919.6	1.2	6962239
$W \to \mu\nu$ 1p	ALPGEN	HERWIGJIMMY	1304.2	1.2	4988236
$W \to \mu\nu$ 2p	ALPGEN	HERWIGJIMMY	377.83	1.2	3768737
$W \to \mu\nu$ 3p	ALPGEN	HERWIGJIMMY	101.88	1.2	1008446
$W \to \mu\nu$ 4p	ALPGEN	HERWIGJIMMY	25.75	1.2	254950
$W \to \mu\nu$ 5p	ALPGEN	HERWIGJIMMY	6.92	1.2	70000
$Z \to \mu\mu$ 0p	ALPGEN	HERWIGJIMMY	668.68	1.2	6615230
$Z \to \mu\mu$ 1p	ALPGEN	HERWIGJIMMY	134.14	1.2	1334296
$Z \to \mu\mu$ 2p	ALPGEN	HERWIGJIMMY	40.33	1.2	1999941
$Z \to \mu\mu$ 3p	ALPGEN	HERWIGJIMMY	11.19	1.2	549896
$Z \to \mu\mu$ 4p	ALPGEN	HERWIGJIMMY	2.75	1.2	150000
$Z \to \mu\mu$ 5p	ALPGEN	HERWIGJIMMY	0.77	1.2	50000
$W \to \tau\nu$ 0p	ALPGEN	HERWIGJIMMY	6918.6	1.2	3418296
$W \to \tau\nu$ 1p	ALPGEN	HERWIGJIMMY	1303.2	1.2	2499194
$W \to \tau\nu$ 2p	ALPGEN	HERWIGJIMMY	378.18	1.2	3750986
$W \to \tau\nu$ 3p	ALPGEN	HERWIGJIMMY	101.51	1.2	1009946
$W \to \tau\nu$ 4p	ALPGEN	HERWIGJIMMY	25.64	1.2	249998
$W \to \tau\nu$ 5p	ALPGEN	HERWIGJIMMY	7.04	1.2	65000
$Z \to \tau\tau$ 0p	ALPGEN	HERWIGJIMMY	668.4	1.2	10613179
$Z \to \tau\tau$ 1p	ALPGEN	HERWIGJIMMY	134.81	1.2	3334137
$Z \to \tau\tau$ 2p	ALPGEN	HERWIGJIMMY	40.36	1.2	1004847
$Z \to \tau\tau$ 3p	ALPGEN	HERWIGJIMMY	11.25	1.2	509847
$Z \to \tau\tau$ 4p	ALPGEN	HERWIGJIMMY	2.79	1.2	144999
$Z \to \tau\tau$ 5p	ALPGEN	HERWIGJIMMY	0.77	1.2	45000
$W \to e\nu$	Sherpa	PHYTHIA	8755.50	1.19	1699846

Table 6.5 – continued from previous page

$W \to mu\nu$	Sherpa	Phythia	8752.80	1.20	1699694
$Z \to ee$	Sherpa	Phythia	925.29	1.16	1999196
$Z \to \mu\mu$	Sherpa	Phythia	925.00	1.16	1999294
Wc 0p	Alpgen	Herwig	644.4	1.20	6427837
Wc 1p	Alpgen	Herwig	205.0	1.20	2069646
Wc 2p	Alpgen	Herwig	50.8	1.20	519998
Wc 3p	Alpgen	Herwig	11.4	1.20	115000
Wc 4p	Alpgen	Herwig	2.8	1.20	30000
Wcc 0p	Alpgen	HerwigJimmy	127.53	1.20	1274846
Wcc 1p	Alpgen	HerwigJimmy	104.68	1.20	1049847
Wcc 2p	Alpgen	HerwigJimmy	52.08	1.20	524947
Wcc 3p	Alpgen	HerwigJimmy	16.96	1.20	170000
Wbb 0p	Alpgen	HerwigJimmy	47.35	1.20	474997
Wbb 1p	Alpgen	HerwigJimmy	35.76	1.20	205000
Wbb 2p	Alpgen	HerwigJimmy	17.33	1.20	174499
Wbb 3p	Alpgen	HerwigJimmy	7.61	1.20	69999
WW	Herwig	Jimmy	1.48	11.50	2490000
ZZ	Herwig	Jimmy	1.60	3.46	250000
WZ	Herwig	Jimmy	1.30	0.97	1000000
single top (s-ch.) $e\nu$	AcerMC	Phythia	0.498	1.	199985
single top (s-ch.) $\mu\nu$	AcerMC	Phythia	0.489	1.	199678
single top (s-ch.) $\tau\nu$	AcerMC	Phythia	0.521	1.	489452
single top (t-ch.) $e\nu$	AcerMC	Phythia	6.941	1.	843357
single top (t-ch.) $\mu\nu$	AcerMC	Phythia	6.825	1.	844530
single top (t-ch.) $\tau\nu$	AcerMC	Phythia	7.264	1.	845118
single top (Wt)	AcerMC	Phythia	15.74	1.	994972
$t\bar{t}$	MCNLO	HerwigJimmy	79.01	1.146	14983835

Conversion of Z+2jets event

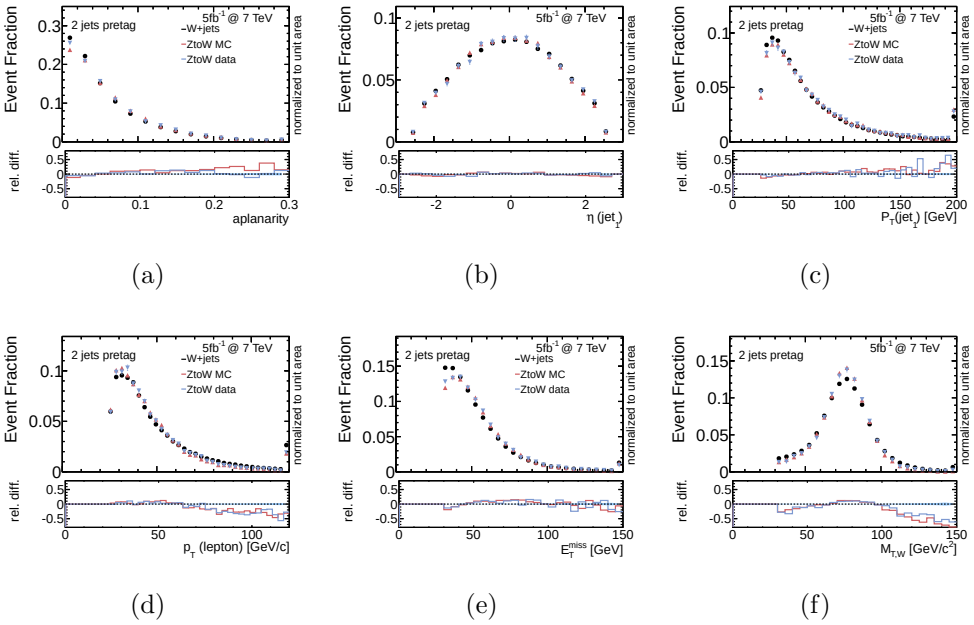


Figure 6.7: *Comparison of simulated W+jets event with converted Z+jets events from simulation and data for the electron channel in the two jet bin. Shown are the aplanarity (a) of the event, the η (b) and the p_{T} (c) of the leading jet, the p_{T} of the lepton (d) , the missing transverse energy (e) and the transverse W boson mass (f). At the bottom of each histogram, the relative difference, compared to W+jets simulation, is shown.*

Systematic uncertainties on the NJet ratio

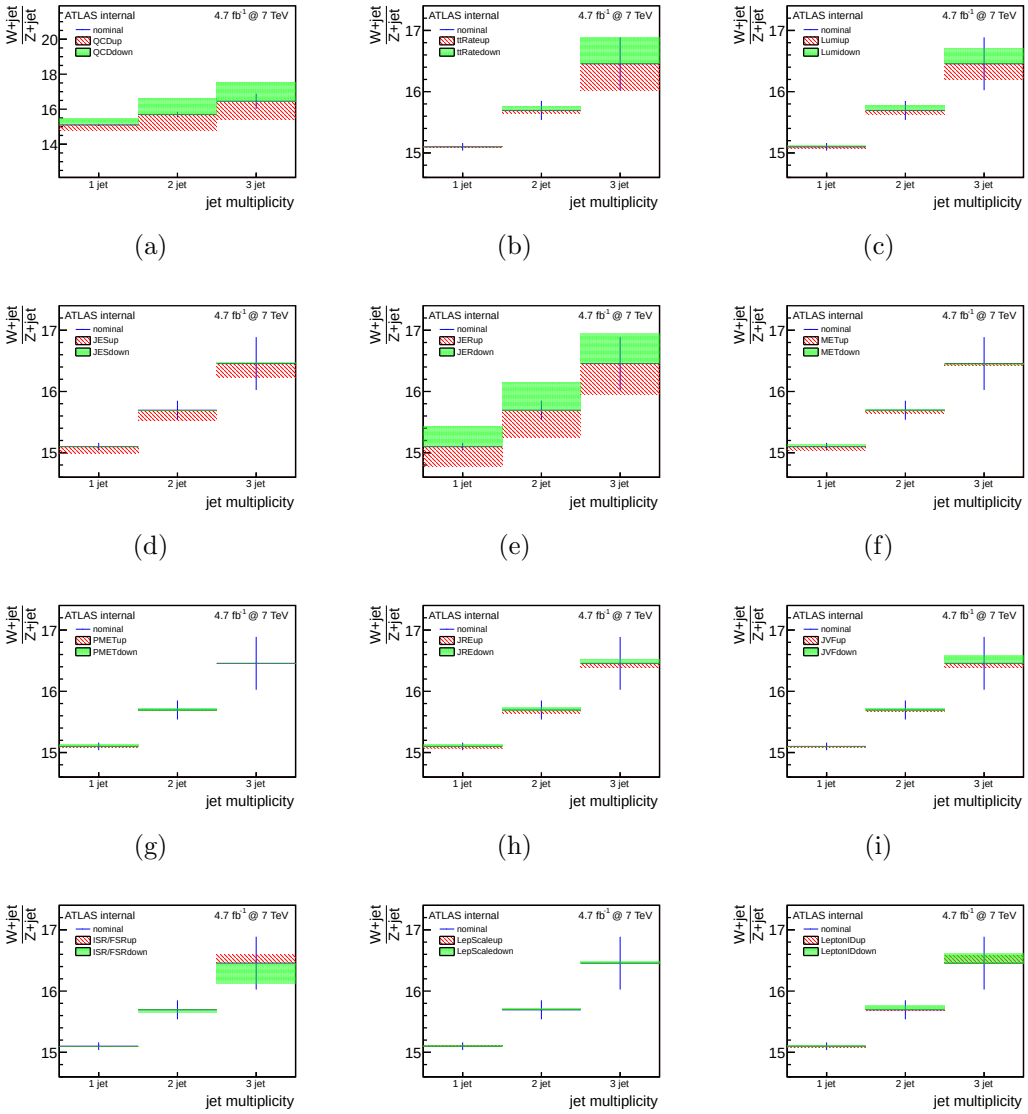


Figure 6.8: *Systematic uncertainties of the electron NJet ratio. Shown are the uncertainty on (a) the multi-jet background estimation, (b) the $t\bar{t}$ -rate, (c) the luminosity, (d) the jet-energy scale, (e) the jet-energy resolution, (f) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (g) the pile-up $E_{\mathrm{T}}^{\mathrm{miss}}$; (h) the jet-reconstruction efficiency, (i) the jet-vertex-fraction scale factors, (j) the ISR/FSR, (k) the lepton-energy scale and (l) the lepton identification.*

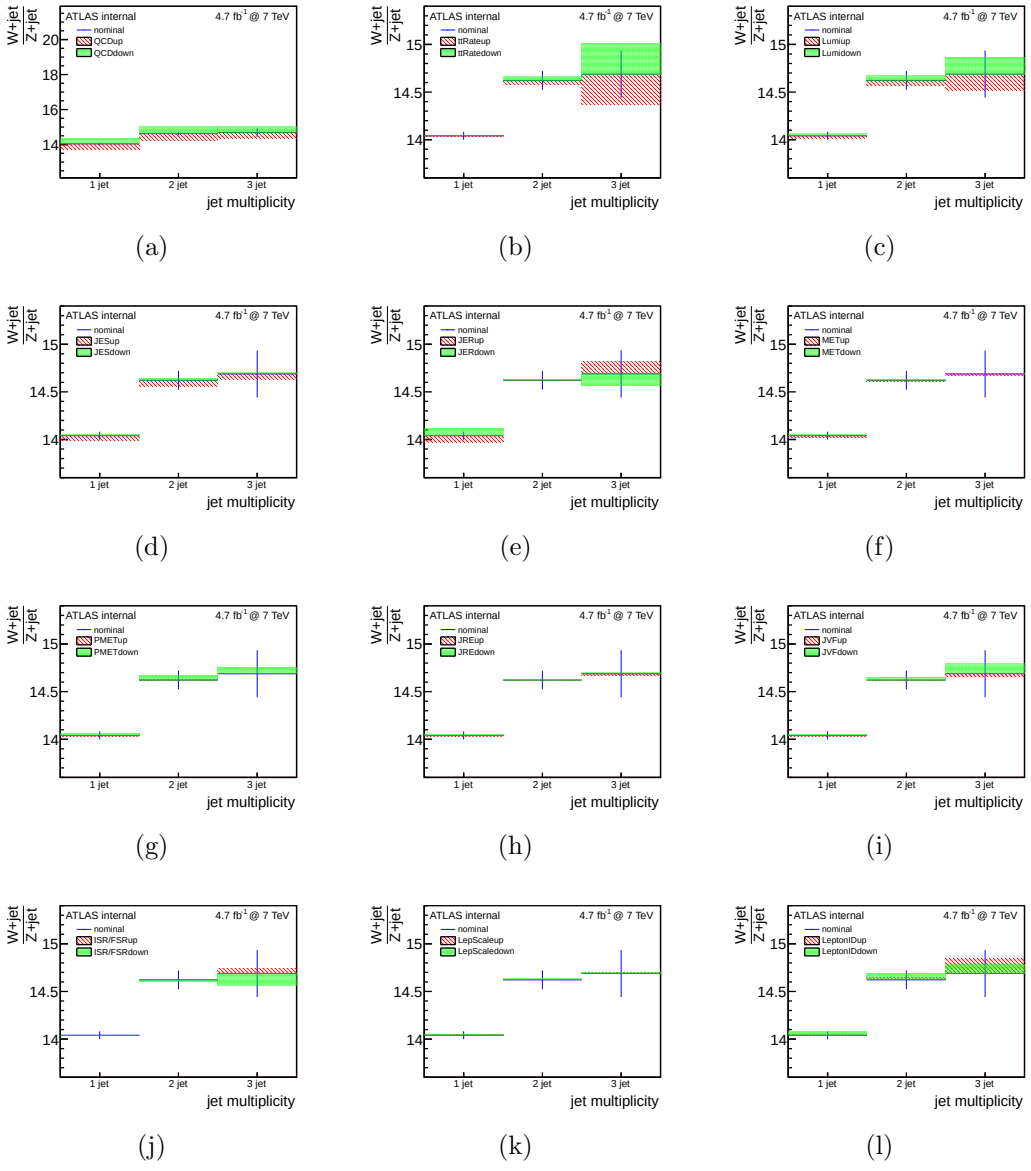


Figure 6.9: *Systematic uncertainties of the muon NJet ratio. Shown are the uncertainty on (a) the multi-jet background estimation, (b) the $t\bar{t}$ -rate, (c) the luminosity, (d) the jet-energy scale, (e) the jet-energy resolution, (f) the $E_{\mathrm{T}}^{\mathrm{miss}}$, (g) the pile-up $E_{\mathrm{T}}^{\mathrm{miss}}$; (h) the jet-reconstruction efficiency, (i) the jet-vertex-fraction scale factors, (j) the ISR/FSR, (k) the lepton-energy scale, (l) the lepton identification*