Search for new physics with top quark pairs in the ℓ+jets channel at $\sqrt{s}$=13 TeV with the ATLAS experiment at the LHC

Dissertation
zur Erlangung des Grades

Doktor der Naturwissenschaften

am Fachbereich Physik, Mathematik und Informatik
der Johannes Gutenberg-Universität in Mainz

Sabrina Groh
geboren in Kirn

Mainz, 14.April 2020

1. Berichterstatter: [In der elektronischen Fassung aus Datenschutzgründen entfernt]
2. Berichterstatter: [In der elektronischen Fassung aus Datenschutzgründen entfernt]
Datum der mündlichen Prüfung: 30.09.2020
Dissertation an der Johannes Gutenberg-Universität Mainz (D77)

"Faithless is he that says farewell when the road darkens.", said Gimly.
"Maybe", said Elrond," but let him not vow to walk in the dark, who has not seen the nightfall."
The Fellowship of the Ring by J.R.R Tolkien

Kurzfassung

Das Standardmodell der Teilchenphysik verkörpert das gesammelte Wissen über die Zusammensetzung und Wechselwirkung gewöhnlicher Materie. Jedoch verbleiben Fragen unbeantwortet, die eine Erweiterung des Modells um neue Teilchen motivieren. Mittels direkter und indirekter Suchen sollen diese Lücken geschlossen werden, wobei das top Quark als schwerstes bekanntes Elementarteilchen eine zentrale Rolle spielt. In dieser Arbeit werden Daten des ATLAS Detektors am Teilchenbeschleuniger LHC analysiert, die 2015 und 2016 bei einer Schwerpunktsenergie von 13 TeV aufgezeichnet wurden. Anhand von konkreten Beispielen wird sowohl der direkte, als auch der indirekte Ansatz zur Suche nach Hinweisen auf neue physikalische Phänomene im ℓ+Jets Zerfallskanal der top-antitop Paarproduktion vorgestellt.

Im ersten Teil der vorliegenden Arbeit wird in einer direkten Suche nach Hinweisen auf neue Physik das Massenspektrum von top-antitop Paaren auf Resonanzen von theoretisch vorhergesagten schweren Teilchen untersucht. Der Schwerpunkt der Analyse ist hierbei auf die Abschätzung des W+Jets Untergrundes gelegt, der in einer externen Kontrollregion bestimmt und basierend auf der Jetflavor abhängigen und daher unterschiedlich ausgeprägten W^+/W^- Asymmetrie in den Wirkungsquerschnitten aus den Daten abgeschätzt wird. Da keine signifikanten Abweichungen von den Standardmodell Erwartungen innerhalb des top-antitop Massenspektrums beobachtet werden, können Ausschlussgrenzen im Bereich von 0.5 bis 3.8 TeV auf ausgewählte theoretische Modelle gesetzt werden.

Im zweiten Teil wird eine indirekte Suche anhand der Präzisionsmessung der Ladungsasymmetrie in der top-antitop-Paarproduktion vorgestellt. Neben der inklusiven Messung wird auch eine differentielle Messung bezüglich der invarianten top-antitop Masse durchgeführt, die gerade bezogen auf geboostete top-Zerfälle im Bereich hoher Massen als besonders sensitiv eingeschätzt wird. Die zentrale Frage der Messung gilt jedoch dem Optimierungspotential der Analyse bezüglich der systematischen Unsicherheit durch die Verwendung des zuvor bestimmten W+Jets Untergrundes. Im Rahmen der Unsicherheiten stimmen die Ergebnisse der Ladungsasymmetriemessungen mit den Erwartungen gemäß des Standardmodells überein.

Abstract

The Standard Model of particle physics embodies the accumulated knowledge of the structure and the interaction of ordinary matter. However, questions remain unanswered, which motivate the extension of the model in order to include new particles. These gaps are to be closed in direct and indirect searches, with the top quark, the heaviest known elementary particle, playing a central role. In the thesis at hand data from the ATLAS detector at the particle accelerator LHC, which were recorded in 2015 and 2016 at a center of mass energy of 13 TeV, are analyzed. Using actual examples, both approaches to the search for evidence of new physical phenomena in the ℓ+jets decay channel of top-antitop quark pair production are presented.

In the first part of this thesis the mass spectrum of top-antitop pairs is analyzed in a direct search for hints of new physics via resonances of theoretically predicted heavy particles. The main focus of the analysis is the estimation of the W+jets background, which is determined in an external control region and estimated from data based on the jet flavor dependent W^+/W^- asymmetry in the cross-section. Since no significant deviations from the standard model expectations are observed within the top-antitop mass spectrum, exclusion limits in the range of 0.5 to 3.8 TeV can be set for selected theoretical models.

In the second part, an indirect search is presented by performing a precision measurement of the charge asymmetry in the top-antitop pair production. In addition to the inclusive measurement, a differential measurement as a function of the invariant top-antitop mass is accomplished, which is considered to be particularly sensitive to boosted top decays in the high mass range. The central question of the measurement, however, is the optimization potential of the analysis regarding the systematic uncertainty by using the previously determined W+jets background. Within the scope of the uncertainties, the results of the charge asymmetry measurements are in line with the expectations according to the Standard Model.

Contents

1 Introduction

> Now, Introduction to Physics. What is physics? Physics comes from the ancient Greek word physiké[1]. It's at this point that you'll want to start taking notes. Physiké means the science of natural things. And it is there, in ancient Greece, that our story begins.
> Dr. Sheldon Lee Cooper [2]

Following the words of the unquestionably famous television-physicist Ph.D., Sc.D. Sheldon Lee Cooper leads us to the beginnings of particle physics, i.e. the discovery of unbreakable small particles, the so called *atomos*[2]. Since then it has been a long way until the description of nature with the Standard Model containing the smallest unbreakable particles of our time. Nevertheless mankind is still trying to "break" the smallest particles for the next step and/or find new particles to increase the understanding of our surroundings and the universe. Unexplainable phenomena, such as dark matter and dark energy, motivate scientists all over the world to explore new theory models and search with particle detectors for physics beyond the current knowledge.

The latest boom of discoveries, in which many new particles have been discovered in a short period of time, happened at the beginning of quark discovery, with explicit searches for the direct detection of particles. With the Higgs particle being the last unbound particle and missing piece of the SM, a plateau has been reached and no new particles have been discovered since.
The search for new physics can be split into two strategies: direct and indirect searches. Direct searches for new phenomena are in general model-independent, but currently have not lead to a discovery of new particles within the explored energy regions at the present experiments. Nevertheless exclusion limits can be set for theoretic models within the direct searches, since although no new particles have been discovered, new information have been gained. Indirect searches on the other hand contribute to the search for new physics by performing precise measurements of Standard Model properties. These type of searches evaluate their measurements regarding evidence of new phenomena, visible as deviation from the Standard Model expectation. The focus lies on increasing the precision of the measurement with every iteration.

The two parts in the thesis at hand represent the different approaches for searches for new physics involving top quarks. The first one is an example of a direct search for heavy resonances decaying into a $t\bar{t}$ quark pair and is published in [3]. The second analysis represents an indirect search for new physics by measuring the charge asymmetry

1 physiké, romanized spelling from the ancient greek word $\varphi\nu\sigma\iota\zeta$[1].
2 Not to be confused with atoms in current physics speech.

in top-pair production. The study has provided valuable input regarding strategy decisions for the publication [4]. Both studies analyze events with the signature of a top-antitop ($t\bar{t}$) decay in the lepton+jets channels with $\int \mathcal{L}$=36.1 fb^{-1} of data[1] recorded at the ATLAS detector during 2015 and 2016 at a center of mass energy of $\sqrt{s}$=13 TeV.

The thesis starts with a brief introduction into particle physics in chapter 2, providing the theoretical foundation and predictions needed for this work. The second chapter, chapter 3, contains a short overview over the particle accelerator complex LHC and the ATLAS detector, where the analyzed data is collected. In chapter 4 the reconstruction and identification of physics objects at the ATLAS detector is described. The dataset and the simulated events for comparison are documented in chapter 5, while the collected data is analyzed with respect to the applied object and selection criteria defined in chapter 6. Special emphasis is paid to the estimation of the W+jets background in chapter 7 in preparation for the search for resonances in the $t\bar{t}$ decay mass spectrum. A short overview over the systematic uncertainties affecting this direct search is presented in chapter 8. The final results of the direct search are presented in chapter 9, including the setting of exclusion limits. In chapter 10 the indirect search via the measurement of the top charge asymmetry in top-antitop pair production is described. The study concludes with a summary in chapter 11.

1 The amount of data is described in the unit *barn* [b] as an alternative unit for cross-sections with 1 b $=10^{-24}$cm^2.

2 Theory foundation

The foundation of particle physics is provided by the Standard Model of Particle Physics (subsequently often abbreviated to SM), which describes the collection of knowledge of fundamental particles and their interactions. A brief introduction into the theoretical aspects of the SM is provided in the first section of this chapter, section 2.1. The second section, section 2.2, focuses on the top quark, the field of attention of the thesis at hand, and discusses in detail its production and decay. A short outlook into selected theoretical models beyond the SM in section 2.3 motivates the search for new physics with top quarks, while the theory foundation is concluded with an introduction into the W+jets asymmetry in section 2.4. The last section lays the foundation of understanding and estimating a major source of uncertainty, the W+jets background, within the top analysis presented further on.

2.1 Standard Model of Particle Physics

The Standard Model of Particle Physics was developed during the late 20th century as a theoretical model, see section 2.1.1, to describe the fundamental interactions of the combined electroweak interaction and the strong interaction along with the fundamental particles and the Higgs mechanisms, see 2.1.2. Murray Gell-Mann and Juval Ne'eman proposed individually in 1961 a scheme to order the previously discovered hadrons, similar to the periodic table of elements [5, 6]. In 1969 Gell-Mann was awarded the Nobel prize in physics "for his contribution and discoveries concerning the classification of elementary particles and their interactions."[7], referring to his work introducing a symmetry scheme, which he called the "Eightfold way"[1]. In 1964 Gell-Mann and Georg Zweig developed the quark model independently[8, 9].

With particle accelerators as the Large Hadron Collider at CERN, see section 3.1, an ideal environment for probing the boundaries of the SM is provided. Center of mass energies in the TeV scale provide input for the search for new particle in continuously increasing mass regions, while precision measurements profit from high luminosities, resulting in high statistics. Physicists all over the world utilized the collected data to search for deviations from the SM, which hints to physics beyond the currently established model.

2.1.1 Mathematical description of the Standard Model

The SM is a relativistic quantum field gauge theory, which describes the particle interactions by a quantum field theory, consistent with both quantum mechanics and special relativity.

1 The "Eightfold Way" refers to the "Noble Eigthfold Path", a path of Buddhist practices, since the number 8 plays a dominant role in the symmetry scheme.

Thus, combining the electroweak theory and quantum chromodynamics into a structure denoted by the gauge symmetry group

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{2.1}$$

the interactions are a consequence of the local gauge invariance under this gauge group. The structure describes the corresponding gauge field of the colour charge C for the strong interaction, of the weak isospin L for the weak interaction and of the hypercharge Y, for the electromagnetic interaction. The underlying gauge theory is non-Abelian due to the non-commutative nature of the $SU(3)$ and $SU(2)$ field strength tensors. With the inability to incorporate general relativity and failing at energies or distances, where the graviton is expected to emerge, the SM can be interpreted as an effective field theory[1].

2.1.2 Interactions and particles

Based on the theoretical introduction in the previous section 2.1.1, a closer look at the field content of the SM is the focus of this section. The first sector presented here is the gauge sector, containing the elementary particles representing the gauge group generators with integer spin, mediating the interactions as force carriers. Next the fermionic sector is introduced, which consists of twelve half-integer spin matter fields, representing the fundamental particles of matter. Finally the Higgs sector is presented with its doublet of Higgs scalar fields, providing mass to the fermions and massive bosons via spontaneous electroweak symmetry breaking.

Gauge sector

In our current understanding gauge bosons with an integer spin of 1 mediate the fundamental forces between the elementary matter building particles. In table 2.1 an overview is provided over the forces included in the Standard Model. A representative of the gravitational force is absent, since a satisfactory quantum theory of gravity to introduce the hypothetical graviton has yet to be worked out. Due to the fact that the strength of gravitation is negligible at the subatomic scale, the missing integration into the Standard Model is not necessary to describe the relevant interactions at high energy.

The gauge sector consists of the electroweak interaction, i.e. the mathematical combination of the electromagnetic and the weak force, and the strong interaction, which will be introduced next.

The oldest and simplest form of these dynamical theories is the Quantum Electrodynamics (QED), which describes the electromagnetic interaction between electrically charged particles via its gauge boson, the photon (γ). The photon is massless and electrically neutral, resulting in an infinite range for the electromagnetic force.

1 An effective field theory is a quantum field theory, that only describes nature at energies below a certain scale or cutoff.

Interaction	Boson	Mass [GeV]	corresponding charge
electromagnetic	photon (γ)	0	electric charge
weak	$W^{\pm}$	80.379±0.012	weak isospin
	Z	91.1876±0.0021	
strong	gluon (g)	0	color charge

Table 2.1: Overview over the interactions within the Standard model, listing the gauge bosons, their recent mass measurements[10] and the corresponding charge, determining the coupling to fermions .

The weak interaction however is mediated by massive gauge bosons, either electrically charged (W^+ and W^-) or neutral (Z). These gauge bosons couple to particles with weak isospin I_3 and have a limited range due to their mass[1].

The strongest force, the hence so called strong interaction, is described by Quantum Chromodynamics (QCD) with gluons as gauge bosons, mediating the interaction between particles, which carry a color charge, a quantum number described as pseudo-color Red (r), Green (g) or Blue (b). The gluons are electrically neutral, massless, carry two color charges, a combination of color and anti-color, leading to an octet of different gluons ($3\otimes3=8\otimes1$) and a singlet. The singlet though is non-existent, due to the fact that this color-free gluon would be expected to exist as a free particle and give rise to a long range force - which is a contradiction to the observations and the current understanding of the strong interaction. Since the gluons carry a color charge themselves, they are the only bosons capable of self-interaction. These gluon-gluon interactions constrain the range of the strong force and confine gluons and color charged elementary particles within composite particles, since the coupling strength increases with the distance. Only color-neutral composite particles are allowed, for example with three colors or anti-colors or a combination of color and anti-color.

Fermionic sector

With the gauge bosons introduced, the focus shifts to the twelve particles with spin 1/2 of the fermionic sector, which represent the building blocks of matter. These so called fermions are categorized into leptons and quarks, each group divided into three generations. The elementary particles throughout the generations have identical quantum numbers, but increase in mass with the generation. While ordinary matter is built up from the first generation of leptons and quarks, the second and third generations are instable and decay via weak interaction into particles of the first generation. An overview over the field content of the fermionic sector is provided in table 2.2.

1 Natural units are used throughout this thesis, setting c=$\hbar$=1 and omitting c and $\hbar$ from corresponding expressions.

Leptons						
Gen.	**Name**	**Symbol**	**Color**	**I_3**	**Q/e**	**Mass**
1.	electron neutrino	ν_e	none	+1/2	0	<1.1 eV*
	electron	e^-	none	-1/2	-1	0.511 MeV
2.	muon neutrino	ν_μ	none	+1/2	0	<0.19 MeV*
	muon	μ^-	none	-1/2	-1	105.66 MeV
3.	tauon neutrino	ν_τ	none	+1/2	0	<18.2 MeV**
	tauon	τ^-	none	-1/2	-1	1776.86±0.12 MeV

Quarks						
Gen.	**Name**	**Symbol**	**Color**	**I_3**	**Q/e**	**Mass**
1.	up	u	yes	+1/2	2/3	$2.16^{+0.49}_{-0.26}$ MeV
	down	d	yes	-1/2	-1/3	$4.67^{+0.48}_{-0.17}$ MeV
2.	charm	c	yes	+1/2	2/3	1.27±0.02 GeV
	strange	s	yes	-1/2	-1/3	95^{+11}_{-5} MeV
3.	top	t	yes	+1/2	2/3	$173.34^{+0.27}_{-0.71}$ GeV
	bottom	b	yes	-1/2	-1/3	$4.18^{+0.03}_{-0.02}$ GeV

Table 2.2: The three generations of fermions of the Standard Model, separated into leptons and quarks. Some quantum numbers are presented and either the mass or mass limits, provided at (*) 90% or (**) 95% level of confidence[10–12]. If no explicit uncertainty on a mass measurement is provided, it is smaller than the given precision. Antiparticles are not listed explicitly. The third component of the isospin I_3 is given for the left-handed fermions, while the right handed fermions have $I_3 = 0$.

The SM differentiates between six types or flavors of quarks, i.e. up (u), down (d), charm (c), strange (s), top (t) and bottom (b), with an antiquark $\bar{q}$ for each quark q, distinguishable by the opposite electric charge. The three lightest quarks (u,d,s) can be grouped together and are referred to as *light quarks* within the thesis at hand. The six quarks are arranged as pairs, consisting of a quark with a positive electric charge of $+2/3e$ and quark with negative electric charge of -1/3e. Besides the electric charge quarks carry a weak and a color charge and are therefore capable of interacting with all forces. Composite states (hadrons) can be built out of either two (mesons) or three (baryons) quarks under the condition, that the overall color-charge is neutral. States with four or five (anti-)quarks have also been observed[13], so-called tetra- and pentaquarks.
Leptons form the second part of the fermionic sector and carry either an electrical charge of -1e or are neutral and called neutrinos. There are three different types of leptons, i.e. electron (e), muon (μ) and tauon (τ), arranged in three generations, which differ in mass. While the charged leptons interact with all electroweak gauge bosons, neutrinos only participate in weak interactions and are therefore difficult to detect. The mass of

neutrinos has not been measured yet, but due to neutrino flavor oscillations, measured from atmospheric and solar neutrinos[14, 15], massless neutrinos can be excluded[16] and upper limits of the mass measurements are experimentally set. The latest individual result for the electron neutrino provides a limit of $m_{\nu_e} < 1.1$ eV at a 90% confidence level[12].

Higgs sector

Since the electroweak $SU(2)_L \otimes U(1)_Y$ symmetry forbids masses for its force mediators W and Z, as well as for elementary matter particles, which is in conflict with observations of $m_W \approx 80$ GeV and $m_Z \approx 90$ GeV, the symmetry needs to be broken. Within the SM the ground state of the electroweak symmetry corresponds to $U(1)_{EM}$ of the electromagnetism, resulting in a massless photon, but generating massive gauge bosons $m_W \approx m_Z$, which are of the order of the electroweak scale M_{EW}[1]. Symmetry breaking can be achieved by adding a $SU(2)_L$ doublet scalar Higgs,

$$H = \begin{pmatrix} H^+ \\ H^0 \end{pmatrix} \tag{2.2}$$

representing the Higgs field with a "Mexican-hat" potential V(H), assuming a negative mass squared parameter $\mu^2 < 0$ and a positive and non-vanishing Higgs field self-coupling $\lambda > 0$, see equation 2.3.

$$V(H) = \mu^2 H^\dagger H + \lambda (H^\dagger H)^2, \text{ with } \mu^2 < 0,\ \lambda > 0 \tag{2.3}$$

With the non-vanishing vacuum expectation value H^0,

$$H^0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix}, \nu = m_H / \sqrt{\lambda} \tag{2.4}$$

with m_H representing the Higgs mass, the gauge group of the Standard Model is broken spontaneously, see equation 2.5.

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \Rightarrow SU(3)_C \otimes U(1)_{EM} \tag{2.5}$$

The masses of the weak gauge bosons are acquired by their coupling to the Higgs field

$$m_W = \frac{\nu g}{2} \text{ and } m_Z = \frac{\nu \sqrt{g^2 + g'^2}}{2}. \tag{2.6}$$

with the gauge coupling g to $\mathrm{SU}(2)_L$ and g' to $\mathrm{U}(1)_Y$. The masses of the fermions can be estimated by calculating the coupling between the vacuum Higgs field H^0 and them, the so called Higgs-Yukawa coupling

1 The electroweak scale, also known as Fermi scale, is the energy scale around 246 GeV, a typical energy of processes described by the electroweak theory.

The mass term of an additional boson, the Higgs boson,

$$m_H^2 = 2\nu^2\lambda \tag{2.7}$$

is a free parameter within the Standard Model and its existence has been confirmed by the ATLAS and CMS experiment in 2012 at the LHC [17, 18]. Latest measurements yield a Higgs mass of $m_{H^0} = (125.10 \pm 0.14)$ GeV[10].

2.2 Top quark physics

Within the Standard Model and its elementary particles, the top quark is the heaviest with a mass of $m_t = 173.34^{+0.27}_{-0.71}$ GeV[11], comparable to the mass of a gold atom. It has been discovered in 1995 with the detectors CDF and D0 at the proton-antiproton ($p\bar{p}$) collider Tevatron at the Fermi National Accelerator Laboratory (Fermilab)[19, 20], next to last of all the currently known elementary particles of the SM, with only the Higgs boson being younger with respect to its discovery. Due to the short lifetime of 10^{-24}s, the experimental measurement of the top properties is only possible via the decay products, since no composite states with top quarks are known.
In the following section 2.2.1, the production mechanism of top quark pairs[1] at proton-proton colliders is introduced, followed by an overview over the top decay channels in section 2.2.2, with focus on the relevant decay channel for the thesis at hand.

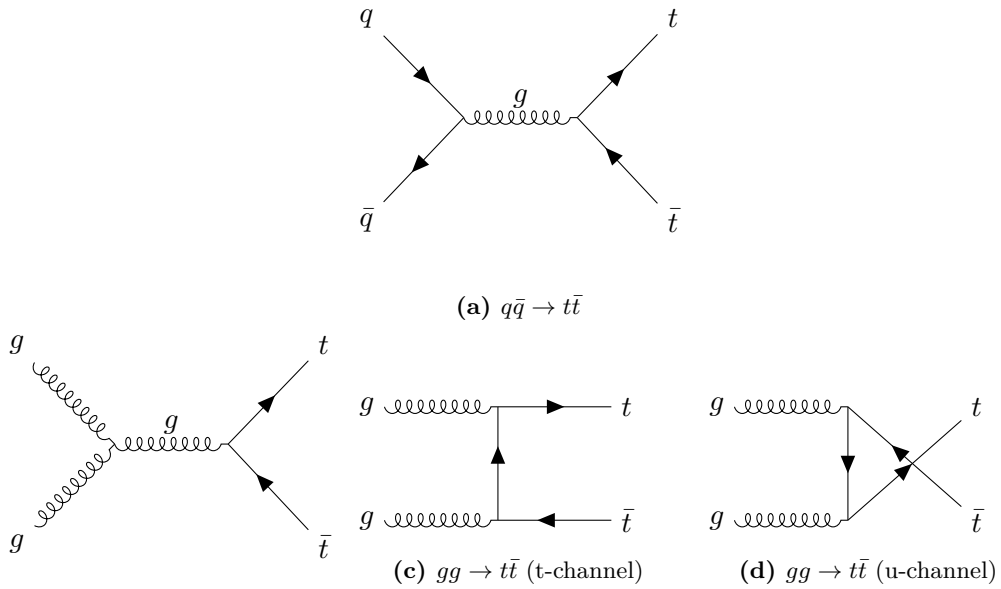
2.2.1 Top quark production at LHC

At proton-proton(pp) colliders top-antitop pairs are mainly produced via strong interaction described by perturbative QCD. The constituents of the colliding protons interact via hard scattering process and produce a top quark pair in the final state. At Born level approximation the dominant production processes are quark-antiquark annihilation, $q\bar{q} \to t\bar{t}$ and gluon-gluon fusion, $gg \to t\bar{t}$. The relevant leading order Feynman diagrams for the contributing processes to top pair production are shown in figure 2.1.
At pp-colliders with a center of mass energy of $\sqrt{s} = 13$ TeV gluon fusion is the dominant process, providing 90% of the top quark pair production, due to the high gluon density within the proton at low fractions χ of the proton momentum, see figure 2.2.

Since the fraction χ of the initial hadron momentum is not known for the parton interaction, the total top cross-section can be described by separating the partonic reaction into a short distance and a long distance contribution

$$\sigma_{t\bar{t}}(\sqrt{s},m_t) = \sum_{i,j} \int\int d\chi_i d\chi_j f_i(\chi_i,Q^2) f_j(\chi_j,Q^2) \times \sigma_{ij\to t\bar{t}}(\rho, m_t^2,\chi_i,\chi_j,\alpha_s(Q^2),Q^2) \tag{2.8}$$

1 Single top production represents a minor background for the analysis at hand and the production processes are briefly discussed in section 5.4.3.



(a) $q\bar{q} \to t\bar{t}$

(b) $gg \to t\bar{t}$ (s-channel)

(c) $gg \to t\bar{t}$ (t-channel)

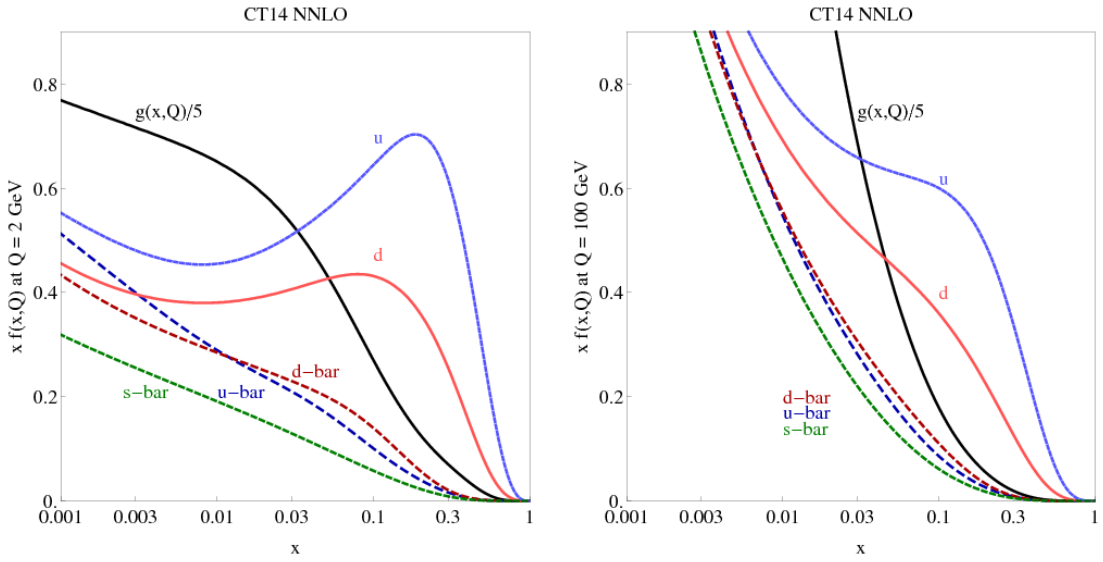
(d) $gg \to t\bar{t}$ (u-channel)

Figure 2.1: Feynman diagrams of lowest order contributing to top pair production at pp-colliders, either in quark-antiquark annihilation (a), or via gluon-gluon fusion (b),(c),(d). Charge conjugation can be applied.

for a given center-of-mass energy $\sqrt{s}$ and a top mass parameter m_t. The summation is performed over all permutations of $i, j \in \{q,\bar{q}, g\}$. The long distance part can be factorized into longitudinal parton momentum distribution functions (PDFs) $f_i(\chi_i, Q^2)$, with the momentum transfer Q. The PDFs represent the probability distribution of observing a parton of type i with a longitudinal parton momentum fraction χ_i and the squared transferred momentum Q^2. Since these probabilities cannot be universally derived from QCD, they have to be provided from experimental studies of the proton structure, mostly from deep inelastic lepton-proton scattering experiments.
The short distance term arises from the hard scattering process of the respective partons, denoted by the partonic cross-section σ_{ij} for partons i and j. This contribution is characterized by high momentum transfer. Hence, it is not dependent on the incoming hadron type or the respective wave functions and can be described by perturbative QCD, see Feynman diagrams in figure 2.1.the parameter ρ is given by equation 2.9,

$$\rho = \frac{4m_t^2}{\sqrt{\chi_1 \chi_j s}} = \frac{4m_t^2}{\sqrt{\hat{s}}} \tag{2.9}$$

where $\chi_i \chi_j s \equiv \hat{s}$ denotes the effective center of mass energy in the partonic reaction. The probability of a parton i to be carrying a momentum fraction of χ_i decreases significantly with rising χ_i, as can be seen in figure 2.2.



(a) $Q = 2$ GeV **(b)** $Q = 100$ GeV

Figure 2.2: Parton distribution functions χf for quarks and gluons inside a proton as a function of the fraction χ of the proton momentum at a given energy scale Q, evaluated for low ($Q = 2$ GeV), see (a), and high ($Q = 100$ GeV), see (b), energy scales inside the proton[21].

Latest next-to-next-to leading order (NNLO) calculations, including next-to-next-to-leading-log soft gluon resummation, predict a cross-section of $\sigma_{t\bar{t}} = 831.8^{+19.8+35.1}_{-29.2-35.1}$ pb[11, 22] at 13 TeV center of mass energy, assuming a top quark mass of 172.5 GeV. Listed are the uncertainties from scale dependence and parton distribution functions.

2.2.2 Top quark decay

The weak decay of quarks within the three quark generations is described by their coupling strength and summarized in the Cabibbo-Kobayashi-Maskawa (CKM) matrix (V_{CKM})[23, 24]. Each matrix element V_{ij} provides the coupling strength between quarks of flavor i and j, while the coupling strength within one quark generation, represented by the diagonal entries in V_{CKM}, is close to unity, see equation 2.10 with the latest results[25].

$$|V_{CKM}| = \begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97446(10) & 0.22452(44) & 0.00365(12) \\ 0.22438(44) & 0.97359^{+(10)}_{-(11)} & 0.04214(76) \\ 0.00896^{+(24)}_{-(23)} & 0.04133(74) & 0.999105(32) \end{pmatrix} \quad (2.10)$$

For top quarks the decay is expected to be dominated by the two-body decay $t \to b + W$, since the top mass is above the Wb threshold and the fraction of decays into a bottom quark is predicted by the branching ratio BR$(t \to b + W) \approx |V_{tb}|^2 = 99.8\%$.
The top decay width Γ_t is 1.35 GeV [11], assuming a top mass of $m_t = 173.3$ GeV. Due to

the short lifetime of $\tau_t = 1/\Gamma_t \approx 0.5 \times 10^{-24}$ s the top quark is expected to decay before composite states with tops quarks can be built.
The final states for the top quark pair decay are defined by the decays of the two W bosons, which individually decay either hadronically into two quarks, i.e. one up- and one down-type quark, or leptonically into a lepton and its neutrino, see Feynman diagram in figure 2.3 illustrating a hadronically and a leptonically decaying W boson within the top quark pair decay.

For the hadronic decay of the W boson a combination of a top quark and an arbitrary down-type quark is excluded due to the low mass of the W in comparison to the top quark mass. The remaining six possible combinations of up-type and down-type quarks and a color factor of three lead to an expected branching ratio of the hadronic W boson decay of $\Gamma_{W,had.}/\Gamma_W$ =(67.41±0.27)% [10]. The leptonic side is almost equally distributed with about 11% branching ratio per lepton flavor, see table 2.3, resulting in an overall branching ratio for the leptonic W decay of $\Gamma_{W,lep.}/\Gamma_W \approx 33\%$.

Decay mode	**Branching ratio in %**
$W^+ \to q\bar{q}'$	67.41±0.27
$W^+ \to e^+\nu_e$	10.71±0.16
$W^+ \to \mu^+\nu_\mu$	10.63±0.15
$W^+ \to \tau^+\nu_\tau$	11.38±0.21

Table 2.3: Decay modes of a W^+ boson, either hadronic or leptonic, with the respective measured branching ratio[10]. W^- modes are charge conjugates of the listed W^+ modes above.

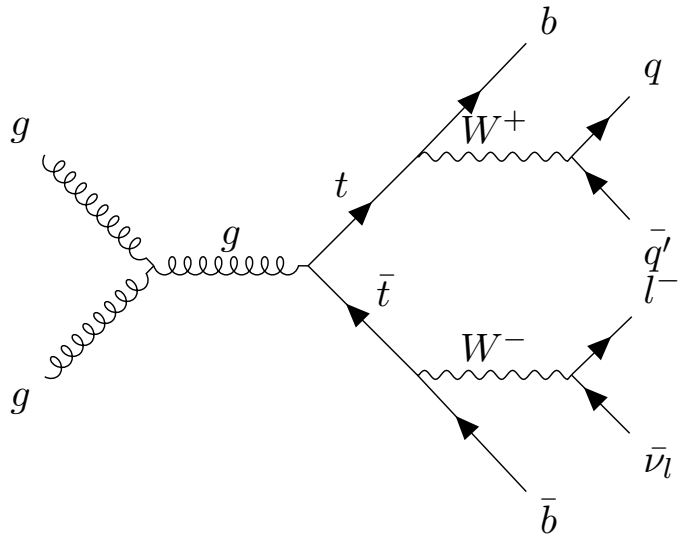


Figure 2.3: Feynman diagram for $t\bar{t}$ decay process with a hadronically decaying W^- and a leptonically decaying W^+, resulting in top quark pair decay of the category ℓ+jets.

For top quark pair decays three final signatures are defined, depending on the decay mode of the two W-bosons, see table 2.4 and the following list:

- **Full hadronic final state (all-hadronic)**
  Both W bosons from the top quark pair decay further into quarks, which hadronize and form a parton shower each, a so-called jet, leading to a total amount of six jets, including the b quarks from the initial top quark decay.
  $\mathrm{BR}(t\bar{t}_{\,all-had.}) = 45.7\%$
- **Semileptonic final state (lepton+jets)**
  One W boson decays hadronically, while the other decays leptonically, leading to a total amount of four jets, including the b quarks from the initial top quark decay, one charged lepton and its corresponding (anti-)neutrino.
  $\mathrm{BR}(t\bar{t}_{\,\ell+jets}) = 43.8\%$
- **Dileptonic final state (dilepton)**
  Both W-bosons from the top quark pair decay further into a charged lepton and the corresponding (anti-)neutrino, respectively. In addition to these two leptons and neutrinos, two jets from the b quarks of the initial top decay are produced.
  $\mathrm{BR}(t\bar{t}_{\,\ell\ell}) = 10.5\%$

Decay mode	BR in %	Process	BR of process in %
all-hadronic	45.7	$t\bar{t} \to qqb + qqb$	45.7
lepton+jets	43.8	$t\bar{t} \to qqb + e\nu_e b$	14.4
		$t\bar{t} \to qqb + \mu\nu_\mu b$	14.3
		$t\bar{t} \to qqb + \tau\nu_\tau b$	15.3
dilepton	10.5	$t\bar{t} \to e\nu_e b + e\nu_e b$	1.2
		$t\bar{t} \to e\nu_e b + \mu\nu_\mu b$	2.3
		$t\bar{t} \to e\nu_e b + \tau\nu_\tau b$	2.4
		$t\bar{t} \to \mu\nu_\mu b + \mu\nu_\mu b$	1.1
		$t\bar{t} \to \mu\nu_\mu b + \tau\nu_\tau b$	2.4
		$t\bar{t} \to \tau\nu_\tau b + \tau\nu_\tau b$	1.3

Table 2.4: Overview of branching fractions of the top pair decay modes, distinguished between all-hadronic, lepton+jets and dilepton decay mode. The branching ratios are extracted from [10, 11]. Anti-/matter identities are not provided explicitly and depend on the charge of the W-boson respectively.

In the all-hadronic channel the correct assignment of the decay products, i.e. the six jets, to the two top quarks as well as a large background from multijet processes are challenging. Within the dilepton channel two neutrinos have to be reconstructed by splitting the missing transverse energy.
The analysis at hand uses the lepton+jets decay channel, which provides a statistics as large as in the all-hadronic channel, but without the dominating multijet background and the non-trivial assignment of decay products to the initial top quark pair. Compared to the

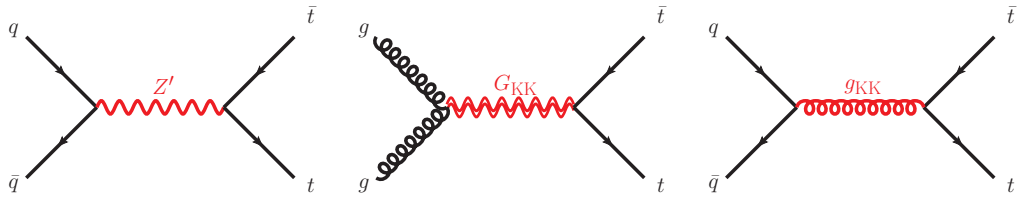
dilepton channel the advantage of a satisfying signal to background ratio is also fulfilled in the lepton+ jets decay mode, but only the reconstruction of a single neutrino is required. The Feynman diagram of a top quark pair decaying in lepton+jets mode is provided in figure 2.3.

2.3 Theoretical models beyond the Standard Model

The Standard Model in its current form is not capable of explaining the open questions of our time regarding for example the building blocks of Dark Matter (DM) or providing an unified theory, where all symmetry groups are contained within a single one, the so called Grand Unification Theory (GUT). With a mass close to the electroweak symmetry breaking scale M_{EW}, see section 2.1.2, the top quark is a candidate to probe physics beyond the Standard Model (subsequently often abbreviated to BSM), since many theories expect new particles to decay to (or couple to) top quarks, due to the large Yukawa-coupling. In addition to that the continuously high production rate of top quarks at the LHC provides a rich testing ground for theoretical models of BSM physics.
Within this section hypothetical new particles are presented, which are assumed to decay into top quark pairs. These particles are distinguished by spin and type. In section 2.3.1 an additional neutral gauge boson Z' is introduced, followed by a graviton field excitation in section 2.3.2 and a gluon field excitation in section 2.3.3.

2.3.1 Spin-1 color singlet Z'

The existence of an additional neutral heavy particle, a so called Z' (Z-prime) boson, is predicted by many theories, to provide an extension of the Standard Model. Within the search for resonances in the top quark pair mass spectrum, two representative theories are probed. The primary production mode for both is $q\bar{q}$ annihilation as shown in figure 2.4(a). Technicolor theories are an example for BSM theories and describe the electroweak symmetry breaking mechanism not by introducing an elementary Higgs boson but with new gauge interactions coupled to new massless fermions. Since the Higgs boson is discovered, topcolor-assisted technicolor theories[26] assume, that the Higgs boson is a composite state of a top and an antitop quark, while the original models without a Higgs boson (elementary



(a) $Z' \to t\bar{t}$ **(b)** $G_{KK} \to t\bar{t}$ **(c)** $g_{KK} \to t\bar{t}$

Figure 2.4: Leading-order Feynman diagrams for the signal processes studies in the search for resonances in the $t\bar{t}$ mass spectrum. The Z' (a) and Kaluza-Klein gluons (g_{KK})(b) have spin 1, while the Kaluza-Klein graviton (G_{KK}) has spin 2(c) [3].Charge conjugation can be applied.

or composite) are ruled out. They also provide an explanation for the large top mass through top quark condensation[27].
The topcolor-assisted technicolor TCA2[28] benchmark model predicts a Z' boson, denoted Z'_{TC2}, which is leptophobic and couples only to the first- and third-generation of quarks. The properties of the hypothetical heavy Z'_{TC2} boson are defined by the topcolor tilting parameter, $\cot\theta_H$, which controls the width and the production cross-section, and $f_1 = 1$ and $f_2 = 0$, which maximize the fraction of Z'_{TC2} bosons that decay into $t\bar{t}$. The tilting parameter is tuned to achieve a resonance width of 1% of its mass[29].
Previous searches by ATLAS [30] and CMS [31] collaboration set lower limits of $m(Z'_{TC2}) >$ 1.8 TeV and $m(Z'_{TC2}) > 2.5$ TeV, respectively, on the allowed mass for the hypothetical Z'_{TC2} boson at 95% confidence level (CL).

With astrophysical and cosmological evidence for dark matter, simplified DM models predict the existence of a messenger particle, which provides a coupling to both regular and dark sector and therefore may be produced in high energy collisions at the LHC. Since the production via quark annihilation is assumed to be likely, the mediator could decay back into two quarks, especially into heavy top quarks.
An axial-vector model with equal couplings to quarks and leptons is considered to be probed within the direct search in the thesis at hand, predicting a weakly coupled TeV-scale axial-vector mediator and a vector mediator, denoted as $Z'_{DM,ax}$ and $Z'_{DM,vec}$, proposed within a framework of simplified models by the LHC Dark Matter Working Group[32]. The five free parameters of the mediators are chosen as follows to serve as comparable signal for the analysis at hand:

- g_q: The coupling to quarks is set to 0.25, to represent a weak coupling strength
- g_l: The coupling to leptons is set to 0, resulting in a leptophobic Z'
- g_{DM}: The coupling to Dark Matter is set to 1, representing a strong coupling strength
- m_{DM} The Dark Matter mass is set to 10 GeV, following the benchmarks A1 and V1 defined in Ref.[32].
- m_{DM}: The Dark Matter mass is varied between 0.5 TeV an 5 TeV

The width of the two Z'_{DM} mediators is 5.6% of their masses, with the $Z'_{DM,ax}$ width kinematically limited to 5.3% at 0.5 TeV.
This model has not been tested so far in searches for top quark pair resonances in the lepton+jets decay channel.

2.3.2 Spin-2 color singlet G_{KK}

The large discrepancy of the Planck and weak scale is denoted as Planck-weak hierarchy problem of the SM and one of the unsolved riddles in particle physics. Many theories exist to propose solutions. One of them is the here presented model based on the Randall-Sundrum (RS) framework with a warped extra dimension [33]. The postulation of an extra dimension of space leads to the Kaluza-Klein excitation of the graviton and produces a spin-2 color singlet boson, denoted as G_{KK} (Kaluza-Klein graviton). The considered

graviton G_{KK} is referred to as a Bulk RS graviton, since the SM fields in the warped bulk and the fermions are localized appropriately[34, 35]. Characterized by a dimensionless coupling constant $k/\overline{M}_{Pl} \propto 1$, where k is the curvature of the warped extra dimension and $\overline{M}_{Pl}$=$M_{Pl}/\sqrt{8\pi}$ is the reduced Planck mass, the decay of G_{KK} into light fermions is suppressed, since the KK gravitons are localized near the TeV brane, whereas the light fermions (1st/2nd generation) are localized near the Planck brane. The primary production mode is gluon-gluon fusion as shown in figure 2.4(b).
The width of the gravitons, searched for within this thesis, varies from 3% to 6% in the mass range 0.4-3 TeV. The branching ratio of G_{KK}, decaying into a top and antitop quark pair, increases rapidly from 18% to 50% for masses between 400 and 600 GeV, reaching a plateau at 68 % for masses larger than 1 TeV.
A search for G_{KK} has been performed previously by the ATLAS and CMS collaboration in the $G_{KK} \to ZZ/WW$ decay channel, resulting in an exclusion of Bulk RS G_{KK} with masses less than 1.3 TeV at 95% CL [36–38].

2.3.3 Spin-1 color octet g_{KK}

Similar to the spin-2 color singlet G_{KK} in section 2.3.2, the spin-1 color octet bosons are produced in RS models with a single warped extra dimension in space leading to Kaluza-Klein excitations of the gluons[39]. The primary production mode is $q\bar{q}$ annihilation as shown in figure 2.4(c).
The Bulk RS Kaluza-Klein gluons, g_{KK}, considered for the direct search, have a width varying between 10% and 40% of the g_{KK} mass. Characteristic for the chosen setting is the strong coupling of the hypothetical gluons to light quarks with $g_q = -0.2g_s$, with the SM gluon coupling g_s. In addition to that the left-handed coupling to top quarks is fixed to $g_{L(t)} = g_s$ and the right-handed coupling to top quarks, $g_{R(t)}$, is varied to achieve the desired width.
Previous searches of the ATLAS Collaboration exclude a g_{KK} (15% width) with a mass below 2.1 TeV[40]. Using slightly different benchmark models the CMS Collaboration sets exclusion limits for g_{kk} masses of less than 3.3 TeV at 95% CL[41].

2.4 W^+/W^- boson production ratio

Production of a W boson with associated jets is the largest background in the top quark related analyses in this thesis. The theoretical prediction of the W+jets cross-section is not as precise as the ratio of the W^+ to W^- cross-sections, since the cross-section ratio benefits from significant cancellations of correlated uncertainties, e.g. PDF uncertainties[42]. In chapter 7 the ratio of the asymmetric W^+/W^- production is utilized for a data-driven background estimation. The theoretical background of the asymmetric production of W^+ and W^- bosons is introduced in this section.

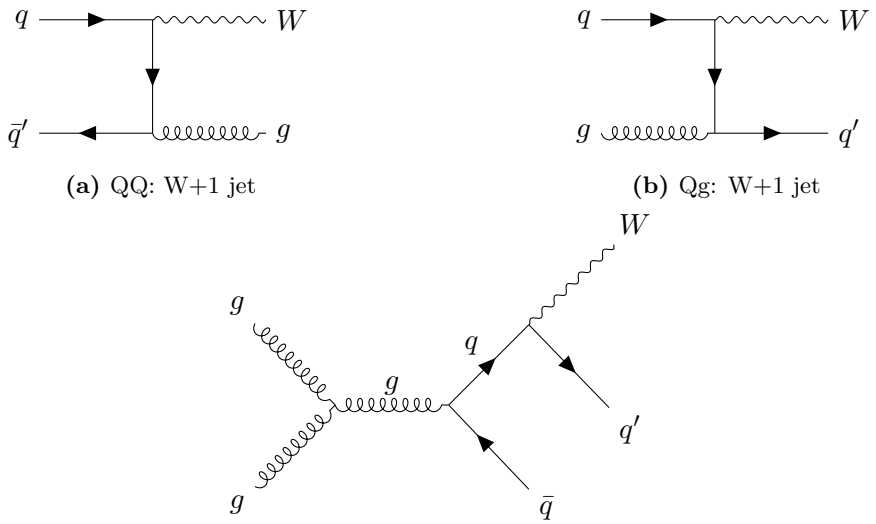
The predominant production mechanism of W bosons without associated jets at proton-proton colliders is the quark-antiquark annihilation, see figure 2.5, where W^+ bosons are produced primarily via $u\bar{d} \to W^+$ or $c\bar{s} \to W^+$ and W^- bosons in the corresponding

charge inverted processes. The production rate of W^+ is significantly higher than for W^-, since the W^+/W^- ratio reflects the u/d ratio of PDFs, due to the proton being a composite state of (uud)[42]. The production process with c and s quark is symmetric, since only sea quarks are available.

(a) W^+ **(b)** $W-$

Figure 2.5: Feynman diagrams for leading order $W^{\pm}$ boson production via quark-antiquark annihilation.

The production of a W boson with associated jets consists of three subprocesses, with the quark-gluon production (Qg) as the dominant process, see table 2.5 and figure 2.6.



(a) QQ: W+1 jet **(b)** Qg: W+1 jet

(c) gg: W+2 jets

Figure 2.6: Exemplary Feynman diagrams for W+jet production at LO via quark-antiquark annihilation (a), quark gluon scattering (b) and gluon-fusion (c). The W-boson charge depends on the quark type. Additional jets may result from initial and final state radiation.

The expected ratios $R^{\pm}$ for the W^+/W^- production without and with up to 4 jets exclusive is summarized in table 2.6. The ratio drops from $n = 0$ to $n = 1$ and increases steadily with n afterwards. This demonstrates the impact of the different subprocesses, which are present at associated jet production only. The leading QQ and Qg subprocesses are dominated by $u\bar{d}$ and gu as well as $d\bar{u}$ and gd, respectively, for W^+ and W^-, following the u/d ratio of the proton PDF. For $n = 1$ a non negligible contribution of $Q + g \rightarrow W^+$

n	QQ in [%]	Qg [%]	gg [%]
0	100	0	0
1	18	82	0
2	21	73	6
3	23	70	7
4	25	67	8

Table 2.5: W+jets production subprocess breakdown (in percent) of leading order $W^{\pm}$ +n jet production at the LHC ($\sqrt{s} = 14$ TeV) and $Q = q$ or $\bar{q}$ and $q \in \{u,d,c,s\}$[43]

$(Q' + g \rightarrow W^-)$with $Q \in \{\bar{d},\bar{c},s\}$ $(Q' \in \{\bar{u},\bar{s},c\})$ arises and decreases the ratio by its symmetric production of W^+ and W^- due to the initial quarks being sea quarks.

n	$\mathbf{R}^{\pm}$
0	1.31 ± 0.01 (scl) ± 0.01 (pdf)
1	1.27 ± 0.01 (scl) ± 0.01 (pdf)
2	1.33 ± 0.02 (scl) ± 0.01 (pdf)
3	1.45 ± 0.03 (scl) ± 0.02 (pdf)
4	1.55 ± 0.04 (scl) ± 0.02 (pdf)

Table 2.6: Prediction for the W^+/W^- production ratio $R^{\pm}$ at the LHC. The ratios are calculated at NLO for n=0,1,2 and estimated at LO for n=3,4. The scale (scl) and PDF uncertainties are displayed. [43]

For the preparation of the W+jets background estimation method in chapter 7 the production subprocesses are distinguished by the flavor of the quark in the final state. This splitting is necessary to estimate the flavor fractions from data, since the theoretical predictions for heavy flavor fractions suffer from large uncertainties. The following four final states are considered:

- $W + b\bar{b}$: The final state contains a $b\bar{b}$ quark pair besides the W boson.
- $W + c\bar{c}$: The final state contains a $c\bar{c}$ quark pair besides the W boson and no $b\bar{b}$ quark pair.
- $W + c$: The final state contains a single c or $\bar{c}$ quark besides the W boson, and no $b\bar{b}$ quark pair or additional c quarks.
- W+light quarks ($W + lf$): The final state contains only light quarks, with as light defined flavors, besides the W boson.

Additional light flavor quarks in the final states are allowed for all four defined subprocesses.

W+c production

The main production channel for W+c production are the processes $sg \to W^- + c$ and $\bar{s}g \to W^- + \bar{c}$, since $\bar{d}g$ and dg are Cabibbo suppressed[1]. Overall the $W^- + c$ yield is expected to be slightly larger than the $W^+ + c$ yield, due to the d valence quark in the proton.

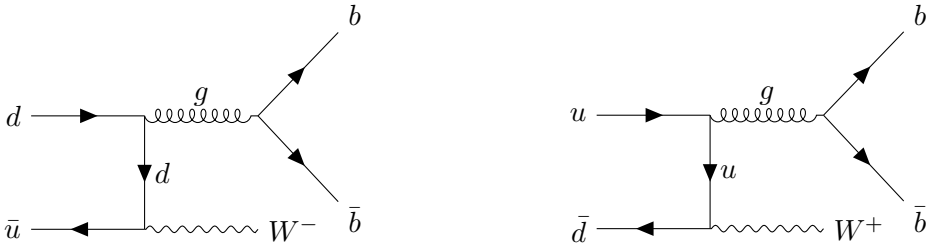
(a) W^-+c **(b)** W^++c

Figure 2.7: Feynman diagrams for $W + c$ production for $W^- + c$ (a) and $W^+ + c$ (b). Additional jets may result from initial and final state gluon radiation.

Of all the different final states, only for the $W + c$ production channel dedicated measurements of the W^+/W^- ratio have been conducted. Previous publications at $\sqrt{s} = 7$ TeV at ATLAS[44] and CMS[45] measured a value of $R^{\pm}$ in agreement with the theoretical prediction of $R^{\pm}$=$0.953^{+0.009}_{-0.007}$ for NLO[46]

W+bb production

The $b\bar{b}$ pair within $W + b\bar{b}$ production is the result of gluon splitting and according to the W^+/W^- ratio identical to $W + c\bar{c}$, where a $c\bar{c}$ quark pair is produced similar to the $b\bar{b}$ quark pair.



(a) W^-+bb **(b)** W^++bb

Figure 2.8: Exemplary Feynman diagrams for $W + b\bar{b}$ production for $W^- + b\bar{b}$ (a) and $W^+ + b\bar{b}$ (b). Additional jets may result from initial and final state gluon radiation.

1 Transitions within one quark generation are favored and transitions across generations suppressed, see CKM matrix in section 2.10.

3 Experimental apparatus

The experimental setup of the analyses described in this thesis is provided by the ATLAS detector[47] at the Large Hadron Collider (LHC)[48] situated at CERN in Geneva, Switzerland. The ATLAS detector is one of the four experiments at collision points of the proton beams, accelerated by the LHC.
The collaboration running the ATLAS detector comprises of about 3000 scientific authors, representing 38 countries. CERN itself was founded in 1953 by 12 founding states. In 2020 23 member states are providing contributions to capital and operational costs and build the CERN Council, responsible for steering the operation of current experimental setups and future plans.

A general overview over the LHC accelerator complex is provided in section 3.1, followed by a short introduction into the ATLAS detector and its components in section 3.2.

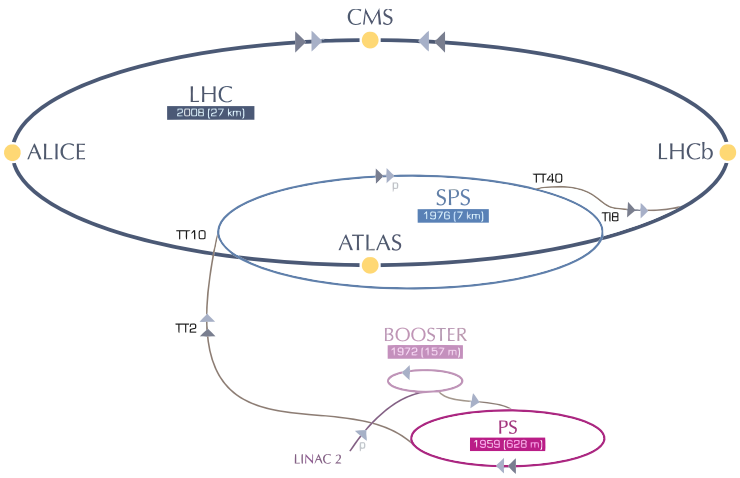
3.1 The Large Hadron Collider

The historic origins of the Large Hadron Collider (LHC) go back to 1994, when the construction of a new accelerator was approved by the CERN Council. Build within the circular tunnel of its predecessor LEP[1] the LHC has a circumference of 27 km and is located on average 100 m below ground level near Geneva in the border region of Switzerland and France.

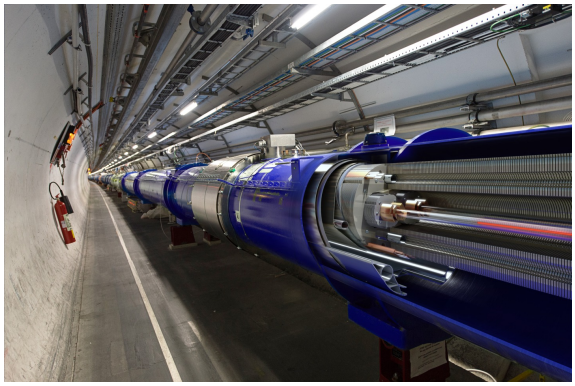
Before high energetic proton-beams collide in the ATLAS detector, protons are extracted from hydrogen and accelerated within pre-accelerators, before reaching their final energies within the LHC, see figure 3.1(a).

The accelerator chain starts with LINAC2, a linear accelerator, in which the protons reach an energy of 50 MeV at a total accelerator length of 33 m. Before entering the LHC the energy is increased in steps up to 450 GeV by a chain of circular accelerators, i.e. the BOOSTER, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS). These accelerators were upgraded to meet the very stringent needs of the LHC: many high intensity proton bunches (up to 2.808 per beam) with small transverse and well defined longitudinal emittance. Within the LHC the protons reach their maximum energy of 7 TeV, resulting in a center of mass energy of 14 TeV. The magnetic field of the 1232 superconducting dipole magnets keeps the protons on their circular trajectory in the course of a length of 27 km. The dipoles are operated at a temperature of 1.9 K, colder than outer space, and provide a magnetic field of up to 8.3 T. The protons circle the LHC in two beam pipes in opposite directions, are focused by 392 quadrupole magnets and

1 The Large Electron Positron Collider (LEP) was running at CERN from 1989 until 2000.



(a) Overview of the layout of the CERN accelerator complex



(b) LHC dipole

Figure 3.1: Overview of the CERN accelerator complex in (a), showing the pre-accelerators, providing proton beams with step-by-step increasing beam energy up to 7 TeV within the Large Hadron Collider for the four collision points. Simplified version based on [49]. In (b) a 3D cut of the LHC dipole is presented[50].

brought into collision at four distinct interaction points. At these four interaction points the experiments ATLAS[47], CMS[51], LHCb[52] and ALICE[53] are located, fulfilling their purpose in lifting the veil in particle physics and beyond.

3.2 The ATLAS experiment

ATLAS, originally an acronym from A Toroidal LHC ApparatuS now considered as proper name, is a multi-purpose experiment and together with its independent partner/rival, the CMS detector (Compact Muon Solenoid), it is designated for the search for new physics

and the precise measurement of the Standard Model of Particle Physics. At 46 m long, 25 m high and 25 m wide, the 7.000-tonne ATLAS detector is the largest volume particle detector ever constructed. It is situated in a cavern 100 m below ground near the main CERN site, at one of the interaction points of the proton beams within the LHC.

The detector consists of an inner detector (Pixel Detector, SCT Tracker and TRT Tracker) for tracking and vertexing, an electromagnetic and hadronic calorimeter system and last a muon spectrometer, as shown in figure 3.2. A solenoid magnetic field of 2 T in the inner detector enables to measure the transverse momentum of charged particles, besides the measurement of vertices. A more detailed description of the inner detector can be found in section 3.2.2. The electromagnetic calorimeter is realized as a Liquid Argon (LAr) calorimeter, while the hadronic calorimeter is based on the principle of a Tile calorimeter, see section 3.2.3. The outer layer of the detector is formed by the muon spectrometer, including a toroidal magnet system, separated into a barrel and two end-caps, see section 3.2.4 for details. To reduce the rate of proton-proton collisions from 40 MHz to a manageable rate of 1 kHz a trigger system is deployed, see section 3.2.6, before the collision data can be stored and analyzed all over the world via GRID-computing, see section 3.2.7.

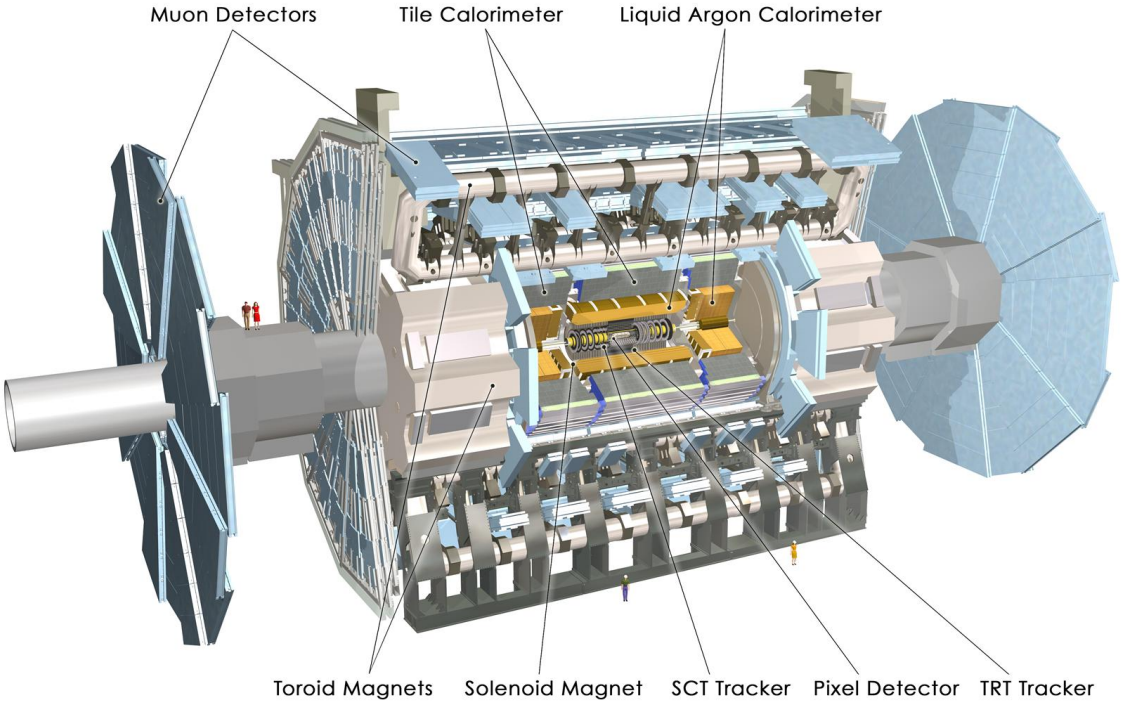


Figure 3.2: Cut-away view of the ATLAS detector and its components[54]

The geometry and the detector components as well as the trigger and data acquisition system described in the following sections are based on the technical design report for ATLAS [55, 56] and [47].

3.2.1 ATLAS Coordinate System

Throughout this thesis the standard ATLAS definition of coordinates is used, a right-handed system with the z-axis defined by the beam axis within the detector, the x-axis pointing towards the center of the LHC ring and the y-axis pointing towards the surface. As origin of the system the primary interaction vertex is chosen. The part of the detector described by the positive z-axis is called A-side, while the opposite side is defined as C-side. The polar angle θ is defined relative to the z-axis and the azimuthal angle ϕ is defined in the $x-y$ plane around the beam axis. The pseudorapidity η is defined via the θ as

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

and in case of massive objects the rapidity $y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right)$ is used, where E is the total energy and p_z the component of the momentum in z direction. Angular separation is defined as $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$. Furthermore the transverse momentum p_T, the transverse energy E_T and the missing transverse energy E_T^{miss} defined by

$$p_T = \sqrt{p_x^2 + p_y^2}, \tag{3.2}$$

$$E_T = \sqrt{p_T^2 + m^2} \tag{3.3}$$

and

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \tag{3.4}$$

are used to describe an event based on its kinematical properties.

3.2.2 Inner Detector

The inner detector is the first part of the ATLAS detector to register the decay products of the collision, therefore it is compact and close to the beam line, as well as highly sensitive by simultaneously being resistant to high radiation. It consists of three different systems of sensors covering the range $|\eta| \leq 2.5$. The inner detector is immersed in a magnetic field parallel to the beam axis, generated by the solenoid magnet, see figure 3.2 and 3.3. With the capability to identify primary and secondary vertices due to delayed decay processes, the inner detector provides necessary information to identify jets from bottom quarks within the top quark decay process.

Of the three subsystems the Pixel detector with its modules builds the innermost component of the inner detector, with the Insertable B-Layer (IBL)[58] closest to the beamline at a radius of 3.3 cm. The IBL is surrounded by three concentric layers of Pixel modules at a radius of 4 cm, surrounding the beam axis and further. This innermost layer is called the B-layer and it is, combined with the IBL, vital for good vertexing, since it contributes to the reconstruction of the secondary vertices of B-hadrons. The concentric pixel modules are complemented by three disks at either side of the inner detector. In total the pixel modules are build of about 80 million pixels (readout channels), covering a total area of 1.7

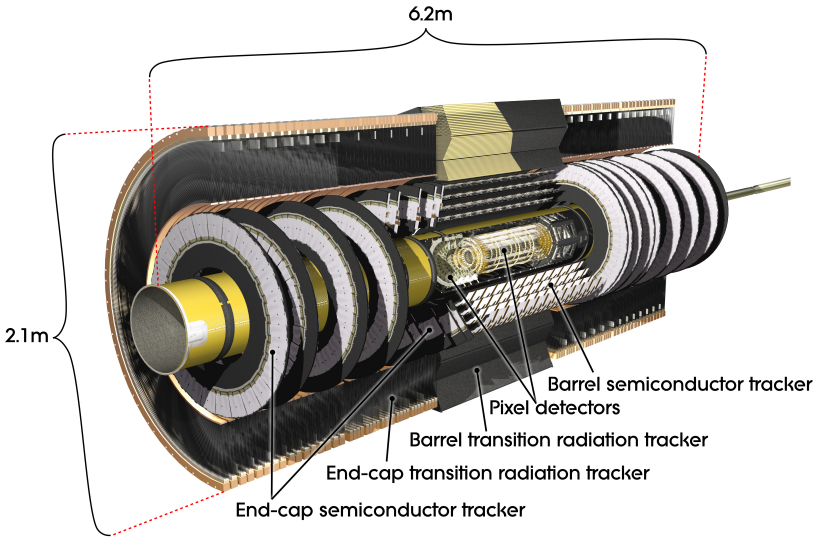


Figure 3.3: Cut-away view of the ATLAS inner detector[57].

m^2 with a power consumption of 15 kW. A single pixel is of the size 50×400 μm^2, providing a high resolution of 14×115 μm^2. Since the high instantaneous luminosity of the LHC leads to multiple interactions per bunch crossing, a high spatial resolution for vertex and track reconstruction is necessary to ensure a sufficient reconstruction and identification of physics objects, see chapter 4.

The Semi-Conductor Tracker (SCT), surrounding the Pixel detector, is arranged in four double-layers of silicon microstrip detectors, each strip of the size 80 $\mu m\times12$ cm. In total the detector component consists of over 4.000 silicon microstrip detectors, covering 60 m^2, and over 6 million readout strips, distributed over 4 cylindrical barrel layers and 18 planar endcap discs. The SCT represents the main part of the ATLAS tracking system and provides an accuracy of 17 μm per layer in the direction transverse to the strips.

The outermost subsystem of the inner detector is the Transition Radiation Tracker (TRT), consisting of densely packed straw tubes filled with a Xenon based gas mixture and containing polypropylene fibers (in the barrel region) or foils (in the endcap region), which serve as transition radiation material. In case of highly relativistic charged particles traversing the straw tubes, transition radiation is emitted proportional to the Lorentz factor $\gamma = E/m$. Electrons can be identified by their transition radiation being above a certain threshold compared to the lower transition radiation generated by heavy objects like hadrons. More than 350.000 read-out channels are utilized for the 50.000 straws in the barrel and 250.000 straws in both end-caps, resulting in an accuracy of 130 μm in $R-\phi$ per straw. In contrast to Pixel detector and SCT, which cover a range up to $|\eta|<2.5$, the TRT is limited to $|\eta|<2.0$.

3.2.3 Calorimeter System

The calorimeter system within ATLAS measures the energy loss of particles passing the detector. With its multiple layers of passive material, i.e. high density material to absorb energy, and active material, such as liquid argon, it is a sampling calorimeter, designated to absorb most of the particles from the collision. Traversing particles, such as electrons, photons or hadrons, interact with the passive material, causing particle showers. The energy of a particle shower can be measured in the active layers as a so called cluster of energy, providing information about the energy of the incident particle.
The ATLAS calorimeter consists of two subcomponents, the electromagnetic (EMCal) and the hadronic calorimeter (HCal), specialized in different particles. The coverage of the calorimeter system is up to 4.9 in $|\eta|$, facing a wide range of different radiation conditions. To accommodate these individual requirements the detector is split into different sections, each adapted to the η dependent conditions. The central region covers the inner detector and has the finest granularity within the electromagnetic calorimeter to measure electrons and photons with high precision. The remaining parts are supporting the reconstruction of jets and missing E_T. A schematic view of the ATLAS calorimeter is provided in figure 3.4. During the reconstruction of top decay products the calorimeter system provides the energy information for electrons and jets as well as the input for the neutrino reconstruction, see chapter 4.

The electromagnetic calorimeter surrounding the inner detector and the solenoid magnet is optimized to measure the energy of electrons or photons as they interact with the detector layers electromagnetically. Electrons and positrons radiate photons via bremsstrahlung, which further produce electron-positron pairs via pair production until the ionization stops the cascade of particles. Photons produce electron-positron pairs first via pair-production.

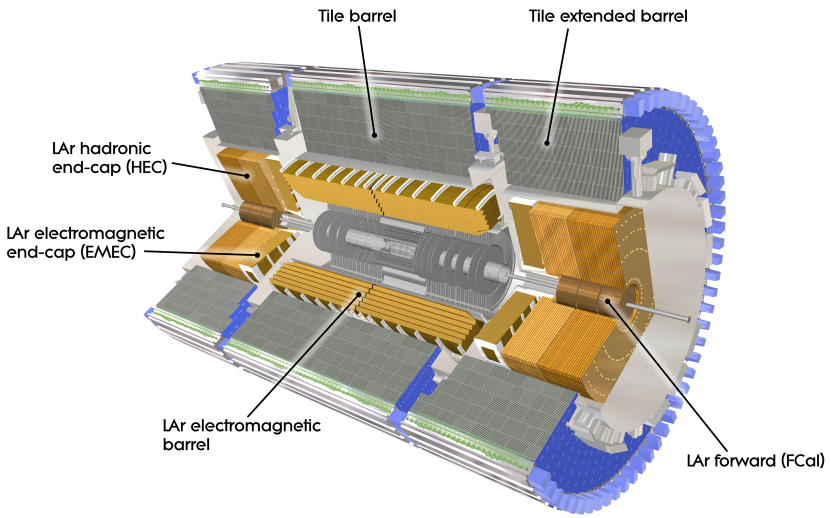


Figure 3.4: Cut-away view of the ATLAS calorimeter system[59]

The EMCal consists of lead layers as passive material, arranged in an accordion like structure, and liquid argon as an active material, providing a homogeneous response, that allows the determination of the energy of the incident particles. The accordion geometry provides complete ϕ symmetry without azimuthal cracks. Within the η range of the EMCal, the thickness of the lead absorber plates varies as a function of η to achieve optimal performance results regarding the energy solution. The calorimeter itself is divided into a barrel part, $|\eta| < 1.475$, and two end-caps (EMEC), covering $1.375 < |\eta| < 3.2$.

The hadronic calorimeter surrounds the electromagnetic calorimeter and is designated to measure the energy deposits of hadrons. The hadronization of quarks and gluons via strong interaction produces hadronic parton showers, which are absorbed within the HCal. Covering a pseudorapidity range of up to 4.9 in $|\eta|$, different techniques are used to accommodate the various requirements. The tile calorimeter over the range of $|\eta| < 1.7$ is a sampling calorimeter using iron as passive material and scintillating tiles as active material. It consists of the tile barrel and two tile extended barrel sections. The next subsystem within the hadronic calorimeter is the LAr hadronic end-cap calorimeter along the z-axis at both sides of the barrel, build with copper plates as absorber material and liquid argon as active material. The end-caps cover an $|\eta|$ range up to 3.2, overlapping the forward calorimeter thereby. The forward calorimeter (FCal), also using LAr as active material, faces high level of radiation due to its position about 4.7 m from the interaction point along the beamline. It consists of three sections, varying in the choice of the passive material: the first is build from copper, while the other two are made out of tungsten.

3.2.4 Muon Spectrometer

Muons pass the described calorimeter system in the previous section 3.2.3 nearly without loss, since their energy loss by bremsstrahlung or ionization of the calorimeter material during transition is minimal[1]. The muons track and energy is measured within the muon spectrometer surrounding the hadronic calorimeter. The muon spectrometer consists of precision-tracking chambers and trigger chambers, see figure 3.5.
Tracking chambers measure and reconstruct the muon momentum via the curvature of the track, bend by the toroidal magnetic field of the superconducting air core toroid magnets, arranged in separate coils. In the range $|\eta| < 1.4$ the magnetic field is provided by the barrel toroid, while in the region $1.6 < |\eta| < 2.7$ smaller end-cap magnets generate the field. In the transition region $1.4 < |\eta| < 1.6$ the combined magnetic field of the two systems provide the deflection of the muon tracks. The magnetic field is inhomogeneous with an average of 0.5 T. As precision-tracking chambers Monitored Drift Tubes (MDT) are arranged in cylindrical shells around the hadronic calorimeter in the central region, $|\eta| < 2.0$, while Cathode Strip Chambers (CSC) are installed in the forward and backward region, $2.0< |\eta| <2.7$. The MDTs measure the curves of the muon tracks and are built of about 1200 chambers with a total of 350k tubes, each tube 3 cm in diameter and between

1 The energy loss per transition depth of swift charged particles is described by the Bethe formula, estimating the energy deposition per flight distance within a given material.

0.85 m and 6.5 m long, providing a tube resolution of 80 μm. The CSCs consist of about 70k channels with a resolution of 60 μm and provide a precise measurement of the particle transition point at the end of the muon spectrometer.
The second component of the muon spectrometer, the trigger chambers, are utilized to provide a fast signal, about 15-25 ns after the passage of a particle, which allows to identify and tag a beam crossing. Besides that the trigger chambers, in form of Resistive Plate Chambers (RPC) in the barrel region ($|\eta| < 1.05$) and Thin Gap Chambers (TGC) in the end-cap region ($1.05 < |\eta| < 2.4$), measure the coordinates of the tracks in η and ϕ. The RPCs have a resolution of 10 mm in η and ϕ direction, while the TGCs provide a resolution of 2-6 mm in η and 3-7 mm in ϕ.
Based on the track reconstruction provided by the inner detector and the information of the muon spectrometer, an efficient muon identification and reconstruction can take place as described in section 4.2.2.

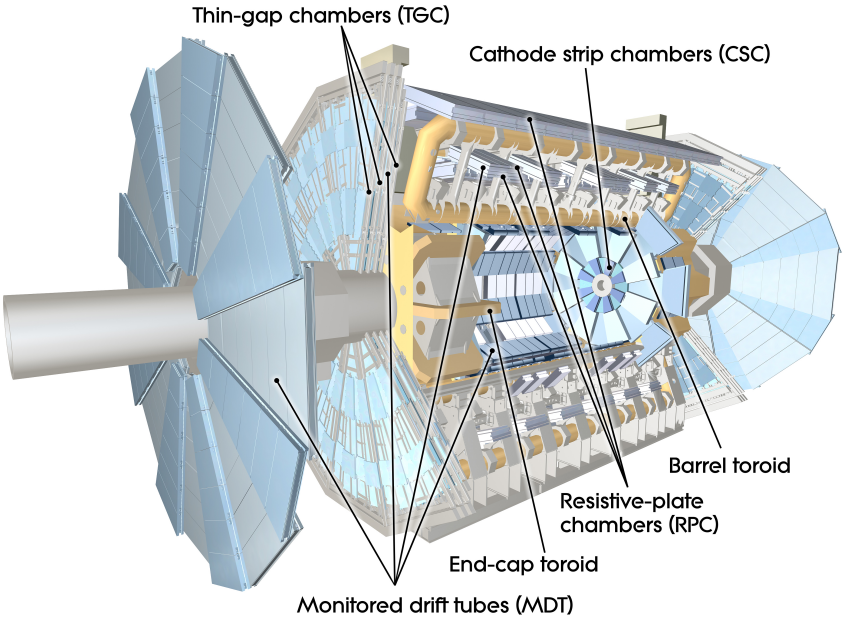


Figure 3.5: Cut-away view of the ATLAS muon subsystems[60]

3.2.5 Luminosity Detectors

The estimation of the integrated luminosity is a source of uncertainty in every analysis, highly important in cross-section measurements but not negligible in the analysis at hand either, see section 8.1. The luminosity measurement for ATLAS is based on an absolute calibration of the primary luminosity-sensitive detectors, i.e. LUCID2[1] [61] and BCM[2][62]

1 LUCID2: LUminosity measurement using Cherenkov Integrating Detector, further referred to as LUCID
2 BCM: Beam Condition Monitor

for ATLAS, in low luminosity runs with special tailored LHC conditions using the van der Meer (vdM) method[1][63]. The calibration settings obtained by the vdM scans once per year are extrapolated to the physics data-taking regime at high luminosities while relative comparisons to identify possible luminosity changes are measured trough the year by LUCID and BCM.

For the ATLAS detector the primary bunch-by-bunch luminosity measurement is provided by the LUCID Cherenkov detector, complemented by bunch-by-bunch measurements from the ATLAS beam condition monitors (BCM) diamond detectors.
LUCID contains 16 photomultipliers tubes (PMTs) on each side of the ATLAS detector along the z-axis, placed approximately 17 m from the interaction point. Cherenkov light is produced in the quartz window of the PMTs, which are coated with ^{207}Bi radioactive sources, which provide a calibration signal. The read-out electronic is designated to measure the luminosity counts for all 3564 nominal LHC bunch slots, where a colliding bunch pair could be present. The raw hit count is converted into a visible interaction rate per bunch crossing, which is proportional to the instantaneous luminosity.
The BCM detector consists of two stations, i.e. forward and backward in direction of z-axis, with four modules each. These modules are build of radiation-hard diamond sensors located around the beam pipe on each side of the ATLAS detector, placed approximately at $z = \pm$ 184 cm and $|\eta| = 4.2$. They provide a measurement of bunch-by-bunch luminosity by counting in-time and out-of-time collisions. Besides the luminosity measurement the BCM is designated to measure the beam quality in real time. In case of anomalies caused e.g. by beam loss BCM triggers a beam dump to protect the ATLAS detector from radiation damage.

3.2.6 Trigger System

The Trigger and Data Acquisition system (TDAQ)[64] is responsible for deciding whether or not an event from a given bunch crossing is stored for later analysis. Due to the high instant luminosity of $L = 1.2 \cdot 10^{34}$ cm^{-2}s^{-1} an efficient and reliable trigger system is required to cope with the high occupancy in the detector, caused by a mean number of $< \mu > \approx 35$ interactions per bunch crossing with a bunch spacing of 25 ns. The aim of this trigger is to reduce the flow of data to manageable levels by selecting only events with distinguishable characteristics, interesting for the physics program.
The ATLAS trigger system is build of a two level online event selection system, the so called Level-1 hardware trigger (L1) and the High Level Trigger (HLT). The first stage retrieves information from the calorimeter and muon systems, identified as Regions of Interest, and reduces the incoming LHC interaction rate of 40 MHz to approximately 100 kHz. Within a time window of 2.5 μs L1 decides to trigger or not, while in the mean time the detector information is stored in a temporary pipeline memory.
During the second stage within the HLT, offline like reconstruction algorithms run in large computer farms and perform decisions within typically 300 ms. The total event rate is

1 In a vdM scan the two beams are shifted relative to each other step by step and the reaction is measured to determine information about the beam dimension $\sigma_{x,y}$.

reduced from 100 kHz to 1 kHz by processing the previously identified Regions of Interest or even the full detector data.
A well functioning and dedicated trigger system is mandatory to identify interesting event signatures at an early stage, e.g. a potential top pair decay with a single lepton, to identify the relevant data for a specific analysis by dismissing the events without significance for any analysis, which would otherwise fill up the data storage.

3.2.7 Data Acquisition and Grid Computing

Following a positive trigger decision events will be recorded and fully reconstructed by the Data Acquisition system (DAQ) at the Tier-0 center, i.e. the computer cluster forming the base of the ATLAS computing system, part of the worldwide LHC Computing Grid (LCG)[65]. Data is stored corresponding to individual LHC runs, where each run is assigned a unique run number. Runs are further divided into luminosity blocks to minimize data loss in case of detector failure. The luminosity blocks are flagged according to the detector performance, where the overall status of a luminosity block can be considered as good for physics and summarized in good run lists (GRL) for use in physics analyses.
The LCG consists of different clusters of computing centers, organized in 4 levels or tiers. After the reconstruction of the data within the Tier-0 computing center, copies of the raw and reconstructed events are transferred to several large computing centers worldwide, the so called Tier-1 cluster. These centers have the prime responsibility of managing the storage of the raw data, as backup to Tier-0, further processed data, created at Tier-1 with its computational capacity designated for reprocessing and analyzing, and simulated Monte Carlo events, generated at Tier-2 centers. Besides the generation of events, the Tier-2 centers process the analyses, mathematical calculations and the final selections of data for the specialized physics groups, such as the Exotics or Top group. Computing facilities at universities and research centers build the Tier-3 centers, where scientists can access the GRID and request data or process calculations.

4 Reconstruction of physics objects

While the previous chapter 3 describes in detail the individual components of the ATLAS detector and their role, the current chapter focuses on the identification of physical objects for the later reconstruction of the top-antitop quark pair. The objects in question, the products of the top quark pair decay, see section 2.2.2, are:

- leptons: muon or electron
- jets: reconstructed from tracks or energy deposits in the calorimeter
- b-tagged jets: jets fulfilling b-tagging requirements
- neutrinos: estimated based on missing transverse energy

As the reconstruction of physics objects is of major concern for every analysis, the methods and algorithms used in the process need to maintain a robust and optimal performance under the challenging working conditions. The first section in this chapter, section 4.1, describes the process of track and vertex reconstruction, the second section deals with the individual physics objects required for the analysis, see section 4.2. This chapter follows the substantial descriptions in [66], updated studies are indicated at the corresponding sections. The ATLAS recommendations regarding common objects for top reconstruction, see [67], are followed for the presented analysis.

4.1 Charged particle reconstruction

The basic constituents of an event and clearly visible in an event display are tracks, primary vertices and the energy deposits in the calorimeter, which form the basic concepts for higher level physic object reconstruction. With these information the origin and the charge of the particles created during the collision can be estimated as well as the direction and the momentum of the particles.

4.1.1 Track reconstruction

Tracks are the reconstructed trajectories of charged particles flying through the inner detector and their traces are reconstructed with information from the subsystems Pixel, SCT and TRT, see section 3.2.2. For muons the information from the designated spectrometer is considered additionally and discussed in section 4.2.2. Tracks provide the basic input for many higher level objects, e.g. electrons and muons, and for the identification of jets, origination from a bottom quark.

Applying local and global pattern recognition algorithms the measurements from the same charged particle are identified and track parameter are estimated using a track fit.

The inner detector track reconstruction consists of several sequences with different track strategies, where the pattern finding and track fitting modules are not clearly distinguishable in modern procedures[68].
The main sequence is referred to as *inside-out* track finding as it starts with forming seeds from the silicon detector measurements (SCT and Pixel), followed by a window search in the direction determined by the seed. The resulting very high number of track candidates needs to be cleared out of fake or incomplete tracks and overlapping track segments with shared hits before extending the reconstruction into the outer TRT. Scoring the tracks in a reward/penalty schema with respect to another achieves the best track collection. Based on the surviving track candidates extensions in the TRT are identified, fitting to the original tracks in the silicon module. The reconstructed tracks are selected by applying track quality criteria cuts, where the so called Tight Primary selection, used in the hand at work, is build on the Loose quality selection. The cuts rely on p_T, η and the number of hits in the inner detector.

Loose:

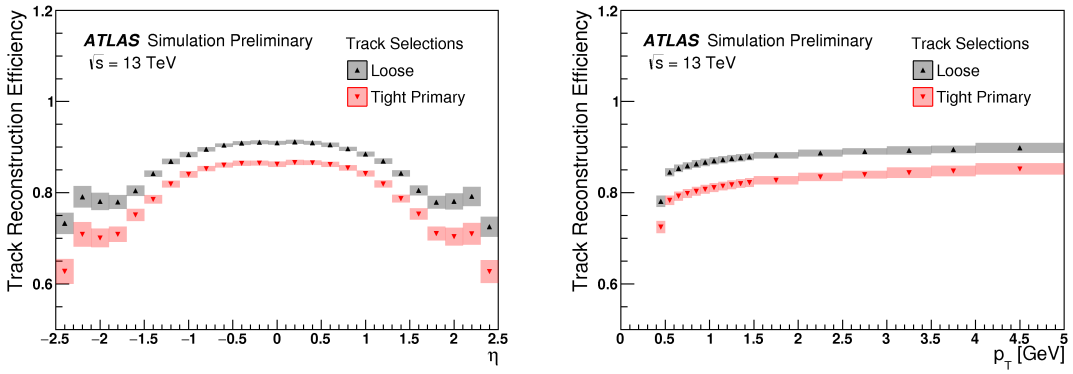
- p_T>400 MeV
- $|\eta|$<2.5
- number of silicon hits[1] $\geq$ 7
- maximum of 1 shared hit[2]
- maximum of 2 silicon holes, i.e. missing hits in SCT and Pixel
- maximum of 1 hole in the pixel detector

Tight Primary (in addition to Loose selection criteria):

- number of silicon hits $\geq$ 9 if $|\eta| \leq 1.65$
- number of silicon hits $\geq$ 11 if $|\eta| \geq 1.65$
- at least one hit on one of the two innermost pixel layers[3]
- no pixel holes

Shared hits are defined as hits used by more than one track and holes are intersections of the reconstructed trajectory with a sensitive detector element lacking a hit. Inactive modules or problematic regions, such as edge areas, are excluded from the hole definition. The tracking efficiency for the Loose and Tight Primary track selection is shown in figure 4.1.

1 hits in Pixel and SCT combined
2 1 shared pixel hit or 2 shared SCT hits
3 IBL and B-layer



(a) Track efficiency as a function of η

(b) Track efficiency as a function of p_T

Figure 4.1: Track reconstruction efficiency evaluated by using minimum bias events, as a function of η (a) and p_T (b) for Loose and Tight Primary track selection. Bands indicate total syst. uncertainty [69].

4.1.2 Vertex reconstruction

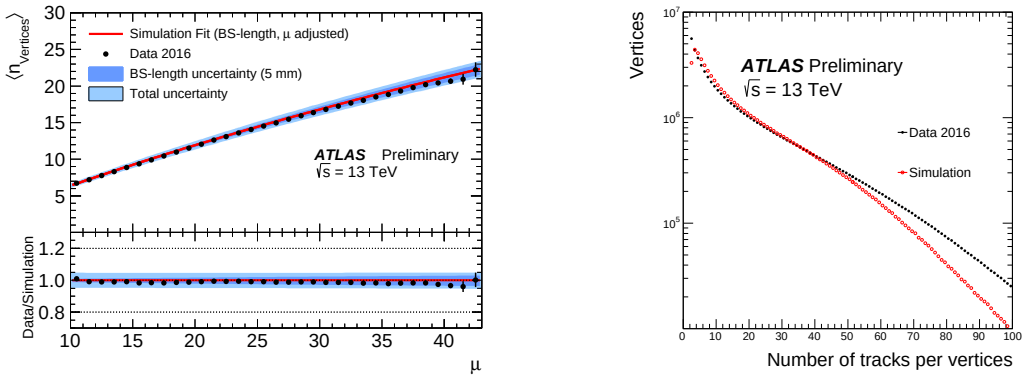
The interaction point in particle collisions is defined as vertex, with primary vertex referring to the reconstructed location of an individual collision. Pile-up[1] complicates the identification of the most valuable hard scattering vertex per event for analysis and complex vertex finding algorithms are necessary to separate the particles coming from a hard scattering from those from pile-up collisions.
The vertexing process consists of the track selection and the primary vertex reconstruction. For the track selection a sub-set of the reconstructed tracks with Tight Primary quality criteria, see section 4.1.1, is used by applying the following additional impact parameter cuts

- $|d_0|<$ 4mm
- $\sigma(d_0)<$ 5mm
- $\sigma(z_0)<$ 10mm
- number of SCT hits $\geq$4
- maximum of 1 SCT hole if $|\eta| \geq$1.65

to reduce the contamination from tracks origination from secondary interactions. The parameters d_0 and z_0 describe the transverse and longitudinal impact parameters of the tracks with respect to the center of the luminous region and $\sigma(d_0)$, $\sigma(z_0)$ represent the according uncertainties.

1 The superposition of multiple inelastic proton-proton interactions reconstructed as a single physical event with many primary vertices is called *pile-up*. They build the dominant part of the total cross section and are usually soft-QCD interactions[70]

From this track selection a seed position for the first vertex is estimated based on the z_0 of all tracks using an iterative vertex finding approach, see [71] for details, to find the most likely value. The position of a vertex is afterwards estimated with an iterative fit procedure, where in each turn less compatible tracks are down-weighed and the vertex position is recomputed. Once the vertex position is determined, incompatible tracks are removed and are available for other vertices. The procedure is repeated until all tracks are assigned to a primary vertex.



(a) number of reconstructed vertices vs. μ **(b)** number of tracks per primary vertex

Figure 4.2: Comparison of 2016 data and simulation for the distribution of the average number of reconstructed vertices as a function of the number of interactions μ [72] (a) and the number of tracks per reconstructed primary vertex[73] (b).

4.2 Object reconstruction

Build on the reconstructed tracks and additional information from the measured energy deposited in the electromagnetic calorimeter and muon spectrometer, more complex physical objects are feasible. At first the reconstruction of two lepton types, i.e. electron and muon, is introduced in section 4.2.1 and 4.2.2, as they are a selection criterion for the in this analysis considered ℓ+jets top quark pair decay channel, in contrast to the third lepton flavor, i.e. tauons. The cone shaped calorimeter showers originating from strongly interacting particles, the so called jets, and the calculation of missing transverse energy to seize the otherwise for the detector invisible neutrinos, are completing this chapter in section 4.2.3 and 4.2.4.

4.2.1 Electrons

For the reconstruction of electrons[1] with transverse energies of the order of a few GeV up to 5 TeV in the central part of the detector, i.e. $|\eta| < 2.47$, energy deposits from the EM calorimeter are matched to reconstructed tracks of charged particles in the inner detector. A brief overview is provided in the following section, for details see [74–77].

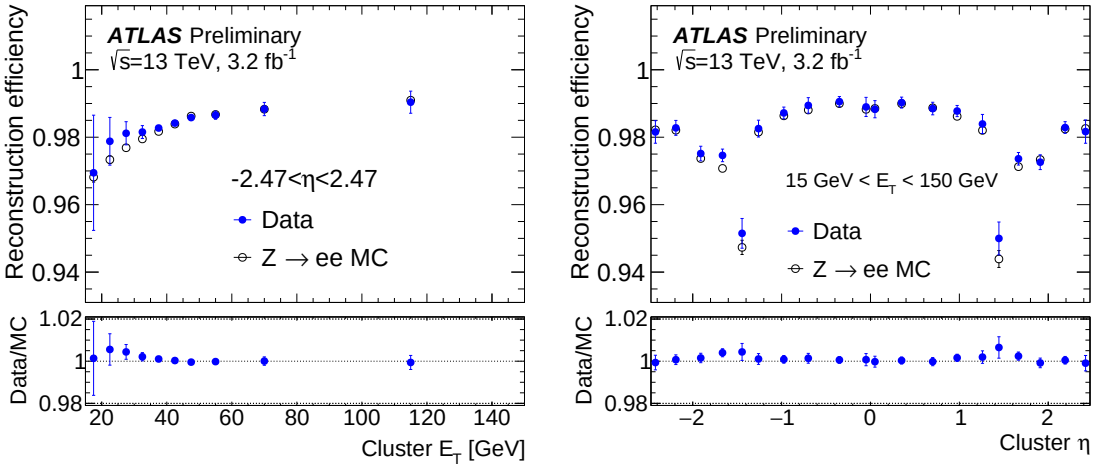
Electron reconstruction

The electron reconstruction algorithm is based on two consecutive steps, the building of cluster in the EM calorimeter and the association of a reconstructed track to a particular cluster.
From the energy deposits in the η-ϕ-plane of the electromagnetic calorimeter towers of the fixed size[2] $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ are built by summing up the energies of the calorimeter cells of all longitudinal layers[3]. For the clustering the *sliding-window algorithm*[78] is used to combine towers within a fixed-size window[4] around a central tower, here $N_\eta^{window} \times N_\phi^{window} = 3 \times 5$. Within this algorithm the central tower is defined as a seed in case of fulfilling the condition, that the position of the window secures a local maximum concerning the contained transverse energy and the sum is above the threshold of $E_T^{thresh} > 2.5$ GeV[76]. Based on a seed, a cluster is formed using a clustering algorithm[78], which includes a duplicate seed removal. The kinematics of the cluster are reconstructed using an extended window depending on the position of the cluster in the calorimeter. The efficiency of this cluster search ranges from 95% at $E_T = 7$ GeV to above 99% for $E_T \geq 15$ GeV[76]. To form electron candidates the tracks are loosely matched to the identified clusters using the distance in η and ϕ between the position of the track, extrapolated into the calorimeter middle layer, and the barycenter of the cluster. Details on the matching procedure, including a refit of tracks associated to electron clusters, can be found in [79]. The efficiency to reconstruct an electron in Data and Monte Carlo simulated events as a function of E_T and η is shown in figure 4.3.

Electron identification

To achieve a large efficiency for the electron reconstruction at the previous step a significant contamination, e.g. with converted photons, is accepted. Algorithms for electron identification (ID) are applied to distinguish between signal like objects and background. Besides quantities related to EM clusters and tracks, information from the TRT and variables measuring bremsstrahlung effects are used. Table 4.1 presents the complete list of variables

1 Electrons and their positive counterpart, the positrons, are treated equally as they differ only in the sign of their electric charge. Therefore *electron* can be replaced by *electron or positron* throughout this text if not explicitly provided.
2 The provided sizes correspond to the granularity of the EM calorimeter middle layer.
3 The energy of a cell spanning several towers is distributed according to the fraction of the area belonging to the different towers
4 The size of the window is provided in units of tower size.



(a) Reconstruction efficiency vs. E_T **(b)** Reconstruction efficiency vs. η

Figure 4.3: Measured reconstruction efficiency for electrons as a function of E_T integrated over the full η range (a) and as a function of η for 15 GeV $<E_T<$10 GeV(b) for the 2015 dataset[76].

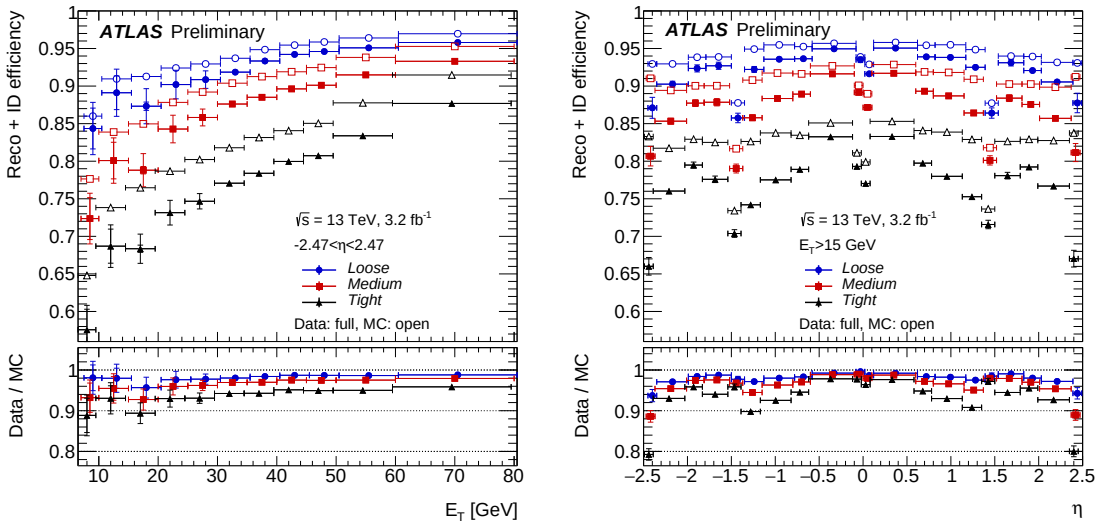
with discrimination power.
The baseline ID algorithm is the likelihood-based (LH) approach[76], i.e. a multivariate analysis technique that simultaneously evaluates all input variables when making a decision. For the LH method signal and background probability density functions of the discriminating variables are used and combined into a single discriminant. Simple selection cuts on the number of hits in the track complete the identification algorithm.
Three levels of identification operating points are provided with increasing background rejection power and including the previous selection criteria:

- LooseLH
- MediumLH
- TightLH

The variables E/p, w_{stot} and $\Delta\phi_2$ are only used as rectangular cuts in the Tight operating point for high E_T and are not included as discriminating variables in the LH algorithm. Figure 4.4 shows the combined reconstruction and identification efficiencies for the LooseLH, MediumLH and TightLH workingpoint. In the work at hand MediumLH and TightLH electrons are utilized for different requirements, see section 6.1.1.

4.2.2 Muons

The reconstruction of muons is based on the independently reconstructed hit and track information from the inner detector, the muon spectrometer and in exceptional cases from



(a) Reconstruction efficiency vs. E_T **(b)** Reconstruction efficiency vs. η

Figure 4.4: Combined electron reconstruction and identification efficiency as a function of E_T integrated over the full η range (a) and as a function of η over the full E_T range (b) for the 2015 dataset[76].

the electromagnetic calorimeter. Within the muon spectrometer a pseudorapidity range up to $|\eta|$ <2.7 can be reached and a momentum measurement with a relative resolution above 3% over a wide p_T range and up to 10% at $p_T \approx 1$ TeV provided[80]. After the reconstruction the identification process is used to suppress background, i.e. mainly pion and kaon decays[1].

Muon reconstruction

Muon track candidates in the muon spectrometer are built using a segment-seeded combinatorial search beginning with seeds from the middle layer and extending afterwards to include the outer and inner layers. For the track measurement hits in at least two matching segments, except in the barrel/end-cap transition region where a single hit in the high-quality segment is sufficient, are required. A global χ^2 fit is applied to relate the hits to the track candidates.
Different combined reconstruction approaches are in application, varying in the choice of the detector components providing the cluster and track information, see [80] for details.

- **Standalone/Extrapolated (ME) muon:**
  The muon track is reconstructed within the muon spectrometer only and fulfills loose requirements for originating from the interaction point. A general criterion is the traversing of at least two layer of muons spectrometer chambers, three in the

1 Pion and Kaon describe two groups of light mesons, with similar attributes within each group

Type	Description	Name
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $\|\eta\|<0.8$ or $\|\eta\|<1.37$)	R_{had1}
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over range $0.8<\|\eta\|<1.37$)	R_{had}
Back layer of EM calorimeter	Ratio of the energy in the back layer to the total energy in the EM accordion calorimeter[5]	f_3
Middle layer of EM calorimeter	Lateral shower width $\sqrt{(\sum E_i\eta_i^2)/(\sum E_i) - ((\sum E_i\eta_i))^2}$, energy E_i and pseudorapidity η_i of cell i and the sum is calculated within a window of 3 $\times$5 cells	w_{η^2}
	Ratio of the energy in 3×3 cells over the energy in 3×7 cells centered at the electron cluster position	R_ϕ
	Ratio of the energy in 3×7 cells over the energy in 7×7 cells centered at the electron cluster position	R_η
Strip layer of EM calorimeter	Shower width, $\sqrt{\sum E_i(i-i_{max})^2)/(\sum E_i)}$, where i runs over all strips in a window of $\Delta\eta \times \Delta\phi \approx 0.0625\times0.2$, corresponding typically to 20 strips in η, and i_{max} is the index of the highest-energy strip	w_{stot}
	Ratio of the energy difference between largest and second largest energy deposits in the cluster over sum of energies	E_{ratio}
	Ratio of the energy in the strip layer to the total energy in the EM accordion calorimeter	f_1
Track conditions	Number of hits in the innermost pixel layer; discriminates against photon conversion	n_{Blayer}
	Number of hits in the pixel detector	n_{Pixel}
	Number of total hits in the pixel and SCT detectors	n_{Si}
	Transverse impact parameter with respect to the beam-line	d_0
	Significance of d_0 defined as the ratio of d_0 and ist uncertainty	d_0/σ_{d_0}
	Momentum lost by the track between the perigee and the last measurement point divided by the original momentum	$\Delta p/p$
TRT	Likelihood probability based on transition radiation in TRT	eProbabil-ityHT
Track-cluster matching	η between cluster position in strip layer and extrapolation track	$\Delta\eta_1$
	$\Delta\phi$ between cluster position in middle layer and track extrapolated from the perigee	$\Delta\phi_2$
	Defined as $\Delta\phi_2$, but track momentum is rescaled to the cluster energy before extrapolating the track from the perigee to the midddle layer of the calorimeter	$\Delta\phi_{res}$
	Ratio of the cluster energy to the track momentum	E/p

Table 4.1: Definition of electron discriminating variables[76]

forward region, to provide a track measurement. Standalone muons are mainly used to extend the acceptance into the region $2.5<|\eta|<2.7$, which is not covered by the inner detector.

- **Combined (CB) muon:**
  The tracks found in the inner detector and the muon spectrometer are combined via a global fit to improve the quality. An outside-in pattern recognition, starting with a reconstructed muon in the muon spectrometer followed by an inward extrapolation, or the complementary inside-out approach, starting with inner detector tracks followed by an extrapolation of these to the muon spectrometer hits outside, are possible.
- **Calorimeter-Tagged (CT) muon:**
  A track in the inner detector can be identified as muon, in case it can be matched to an energy deposit in the electromagnetic calorimeter, which is compatible with a minimum-ionizing particle. This reconstruction algorithm comes with a low purity, but it recovers acceptance in regions, where the muon spectrometer is only partially instrumented ($|\eta|<0.1$). CT muons are optimized for that region and a p_T range of $15<p_T<100$ GeV.
- **Segment-Tagged (ST) muon:**
  Muons leaving a track in the inner detector are defined as ST muons, if the extrapolated track into the muon spectrometer can be matched to at least one local track segment in the MDT or CSC chambers, see section 3.2.4. Covering muons when they only cross one layer of these chambers due to low p_T or regions of reduced acceptance is the main application of the ST muon definition.

Overlap between the different muons types is resolved by giving preferences in case of shared inner detector tracks, listed from high to low: CB, ST and CT muon.
While the ME and CB muons are used for the signal region within the work at hand, the control regions to estimate the multijet background are filled with all four types of reconstructed muons.

Muon identification

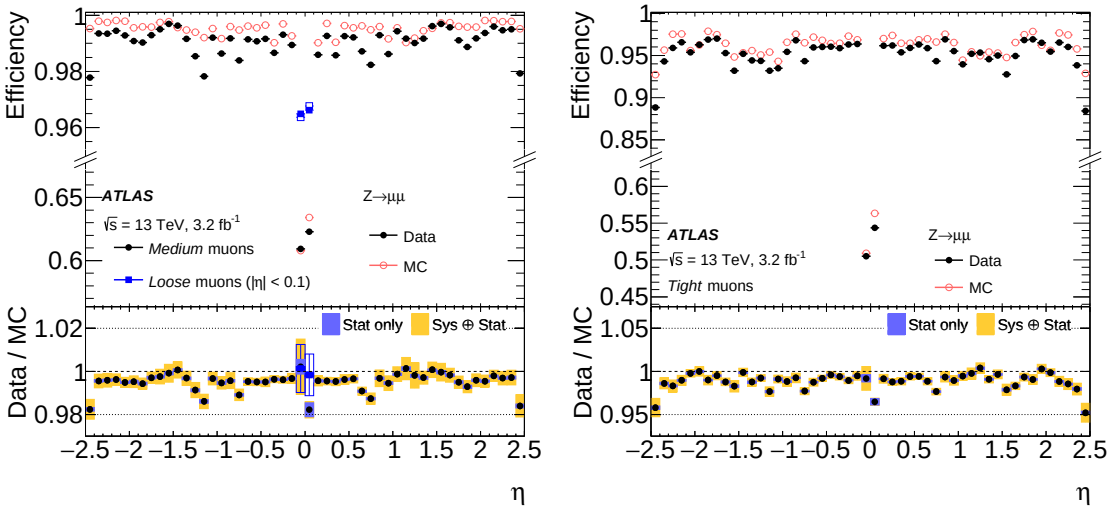
Muon identification quality cuts are applied to suppress background while aiming for a high efficiency for selecting prompt muons and a robust momentum measurement. General selection criteria on the number of hits and holes in the inner detector are used in combination with track specific variables for the combined muons, listed below.

- q/p significance:
  the absolute value of the difference between the ratio of the charge and momentum of the muons measured in the inner detector and the muon spectrometer divided by the sum in quadrature of the corresponding uncertainties
- ρ': the absolute value of the difference between the transverse momentum measurements in the inner detector and the muon spectrometer divided by the p_T of the combined track

- normalized χ^2 of the combined track fit

Three inclusive level of quality criteria are introduced to meet the requirements of different analysis, ranging from low to high background rejection power: Loose, Medium and Tight, while the Loose and Medium muons are selected for the thesis fulfilling different requirements, see section 6.1.2.

- Medium:
  CB or ME muons fulfilling specific hit criteria and with loose selection cuts applied to the compatibility between inner detector and muon spectrometer momentum measurements. The latter suppresses the contamination by hadrons by requiring q/p *significance* <7. The selection is optimized for minimal systematic uncertainties associated with muon reconstruction and calibration.
- Loose:
  To maximize the reconstruction efficiency all muons types are used, combining *Medium* CB and ME muons with CT and ST muons restricted to the $|\eta|$<0.1 region.
- Tight:
  To achieve a high purity at the cost of efficiency, only CB muons passing the medium selection criterion with a modified q/p significance cut <8 are considered. In addition to that a two-dimensional cut in ρ' and q/p significance is applied as a function of the muon p_T to achieve a stronger background rejection for muon momenta below 20 GeV.



(a) Loose and Medium muons **(b)** Tight muons

Figure 4.5: Muon reconstruction efficiency as a function of η measured in $Z \to \mu\mu$ for muons with p_T>10 GeV for the Medium and the Loose muon selection in (a) and for the Tight selection in (b). The error bars indicate the statistical uncertainty[80]

The reconstruction efficiency for muons of the Medium and Tight selections are displayed in figure 4.5.

4.2.3 Jets

Jets are the representatives of quarks and gluons, as the former produce hadronization induced cone-shaped particle showers mostly in the hadronic calorimeter. In addition to that hits in the inner detector can be used for the reconstruction as well as energy deposits registered in the electromagnetic calorimeter. Jets are assigned a radius parameter and a total energy, which needs to be calibrated before further usage. Jets originating from b-quarks exhibit significant properties used for b-tagging. With increasing center of mass energy heavy particle decays tend to be boosted, requiring reconstruction in large jets with substructure analyzing techniques to fulfill the aggravated conditions. All these individual aspects are discussed in the following part of the section, starting with the foundation of the jet reconstruction process, the topocluster.

Topocluster

Before any jet finding algorithms can start its work, the individual hits from the calorimeter system are unified with a topological clustering algorithm[78] into so-called topoclusters. The key idea is to group neighboring cells into a cluster if they fulfill a specific criterion. Contrary to the sliding-window algorithm, topoclusters do not have a fixed size, but they start from seed cells as well. The topological cluster algorithm consists of two steps, the cluster maker and the cluster splitter.
Accountable for the finding of seeds is the cluster maker. Here seeds are defined by having a signal to noise ratio above a certain (high) threshold, t_{seed}, and noise is estimated as the expected RMS of the electronics noise for the current gain and conditions plus additional pile-up contributions. For each seed, starting with the highest signal-to-noise ratio and continuing in descending order, a proto-cluster is formed by considering all neighboring cells and evaluating their signal-to-noise ratio in comparison to a new threshold, $t_{neighbor}$, lower than the seed threshold. Passing this threshold, the neighboring cell is added to the proto-cluster and the seed list. This iterative procedure ends if the tested cells do not pass the threshold $t_{neighbor}$, but pass a third even lower threshold, t_{cell}, which results in adding the cell to the proto-cluster, but dismissing its neighbors as potential candidates. Failing the third threshold t_{cell} also ends the procedure for the tested neighbouring cell. This approach results in forming clusters of individual size and shape in the process. In case of a cell belonging to more than one proto-cluster, the corresponding clusters are merged. The remaining proto-clusters are converted to clusters, sorted in descending order of E_T and dismissed below a certain threshold.

As individual jets are usually not isolated, single cluster can cover large areas at this stage, especially in the end-caps and forward calorimeter region, the cluster splitter searches for local maxima within a cluster, originating from individual particles. A local maximum is defined by having a total energy above a certain threshold and above the energy of

any neighboring cell in combination with the number of neighboring cells within the parent cluster being above a threshold. The number of local maxima found defines the final number of clusters, as clusters with several maxima are being split into the same number of individual clusters and cluster without a local maximum stay untouched. At this point a topological cluster represents energy deposits in three dimensions throughout the calorimeter. The number of topocluster within a jet in a specific p_T range is shown in figure 4.6. The energy of the topocluster is either reconstructed in electromagnetic (EM) scale or LC scale, i.e. the hadronic scale, using the local cluster weighting method to calibrate the topocluster, see [81] for details.

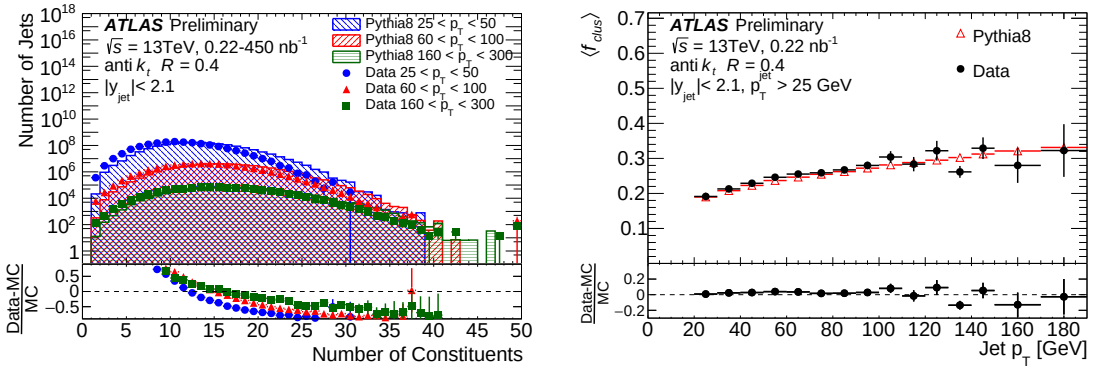
Jet reconstruction

The main jet algorithms used within ATLAS are sequential clustering algorithms. Following the theoretical guidelines regarding infrared safety, collinear safety and order independence[83] these algorithms introduce distances d_{ij} between input objects i, j and d_{iB} between the beamline and the object i as described in equations 4.1 and 4.2, see [84] for details.

$$d_{ij} = min(p_{T,i}^{2p}, p_{T,j}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \qquad \text{with} \qquad \Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 \tag{4.1}$$

$$d_{iB} = p_{T,i}^{2p} \tag{4.2}$$

In an iterative calculation over a list of input objects the minimal distance d_{min} of all d_{ij} and d_{iB} is found and if d_{min} is the distance between two objects i and j, the corresponding objects are replaced by their combination. In the second case, the minimal distance is between an object i and the beamline, i is considered as a jet and removed from the list.



(a) Topocluster multiplicity **(b)** mean F_{clus} vs. p_t^{jet}

Figure 4.6: Topocluster multiplicity for jets in three p_T^{jet} ranges (a) and dependence of fraction of energy carried by the leading cluster in each jet, f_{clus} on p_T^{jet} (b). Are distributions are scaled to 6.4 pb^{-1} and error bars reflect statistical uncertainty[82].

The algorithm is repeated until all elements from the input list are removed. The size of a jet is defined by the radius parameter R, with $R = 0.4$ and $R = 1.0$ for small and large jets, respectively.
The main difference of the available algorithms lies within the parameter p, see table 4.2. Within the anti-k$_t$ algorithms input objects with large momentum are combined first, they are combined last in the k$_t$ algorithm. While the anti-k$_t$ algorithm is the main algorithm for

parameter p	algorithm
1	k$_t$ algorithm
-1	anti-k$_t$ algorithm

Table 4.2: Differentiation of jet algorithm by choice of parameter p in 4.1

the reconstruction of small and large radius jets, k_t is used for trimming and top tagging, so are in this thesis different input objects for different requirements used, see figure 4.7: tracks for b-tagging, topocluster at EM scale for small radius jets and at LC scale for large radius jets.

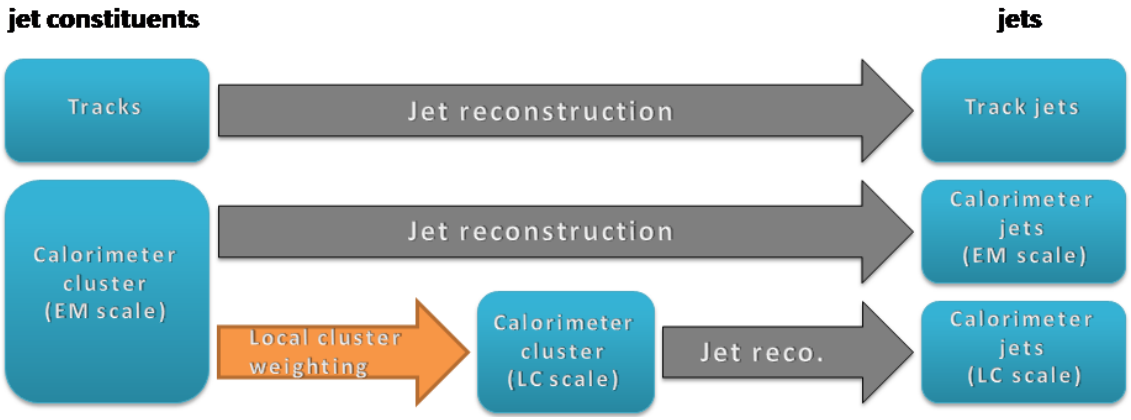


Figure 4.7: Overview of jet reconstruction. The jet four momentum is defined as the four momentum sum of its constituents (based on [81])

Jet calibration

For the calorimeter jets with topoclusters as input, the following correction and calibration is applied, see [85, 86] for details:

- **Origin correction:**
  corrects the jet direction to point to the selected primary vertex for the analysis
- **Pile-up correction:**
  corrects for the energy offset introduced by pile-up using an area-based subtraction procedure and residual corrections. The latter is a function of the number of primary vertices, N_{PV}, and the average number of interactions per bunch crossing $< \mu >$,

and captures residual in-time and out-of-time effects[1]. Derived from Monte Carlo simulated events.

- **Energy and η calibration:**
  calibrates the jet energy by applying p_T and η-dependent corrections derived from Monte Carlo simulated events. Biases in η due to different calorimeter technologies in different regions are corrected for, too.
- **Global sequential corrections:**
  reduce the dependence of the jet energy on the longitudinal and transverse structure of the jet and additionally correct for jets that are only partially contained in the calorimeter. Derived from data events.
- **Residual in situ corrections:** calibrating remaining (residual) issues, derived in situ, i.e. from data events

Jet trimming

Jet trimming is a grooming method applied to large radius jets to reduce the impact of contained energy from underlying events and pile-up contamination. The constituents of the jet are reclustered into subjets with the k_t algorithm[2] and the radius parameter R_{sub} is the first step. In a second step all subjets within a jet with p_T of less than a fraction f_{cut} of the ungroomed jet are discarded[87]. The trimming parameters used for the presented analysis are R_{sub}=0.2 and f_{cut}=5%.

B-tagging

The identification of jets originating from b-quarks within an event is an important tool to distinguish $t\bar{t}$ signal events from background events with light-flavor jets. Several algorithms to tag jets containing a b-hadron as b-jets exist and use the significant decay length resulting in a secondary vertex and/or the significance of the transverse and longitudinal impact parameters of the charged particle tracks. In ATLAS the multivariate-based algorithm (MV2) is the the main algorithm for b-jet identification and it combines the results from impact parameter based algorithms (IP2D and IP3D), a secondary vertex finding algorithm (SV1) and a decay chain multi-vertex algorithm (JetFitter) on jets reconstructed from tracks within the inner detector. A detailed description of the method can be found in [88]. The application of a boosted decision tree achieves a better discrimination than any basic b-jet finding algorithm individually.

1 In-time effects are caused by multiple interactions within one events, while out-of-time effects are effects from preceding/subsequent interactions

2 The k_t algorithm is a sequential clustering algorithm for jet reconstruction, same as anti-k$_t$, but with a difference in the distance calculation $d_{ij} = min(p_{T,i}^{-2}, p_{T,j}^{-2}) \frac{\Delta R_{ij}^2}{R^2}$, compare to equation 4.1

Top-tagging

The identification of heavy particle decays, in particular top decays, discriminates signal from background and can be achieved via jet substructure techniques. For the thesis a so called Smooth Top Tagger, see [89], is used, which applies cuts on one or two substructure observables only. As the latter are in general depending on the jet momentum an important algorithm criterion is its stability concerning the signal efficiency for a wide range of jet p_T. A flexible procedure has been developed to find the optimal two variable combination with smoothed jet p_T dependent cuts at fixed signal efficiency working points. The following substructure variables are used for the Smooth Top Tagger:

- **k_t splitting scales:**
  k_t splitting scales[90] are determined by re-clustering a large R-jet using the k_t algorithm, which tends to cluster high p_T constituents with large distances last. They are defined as the distance $\sqrt{d_{ij}}$ 4.3 between the subjets i and j in the k_t history starting with its last step.

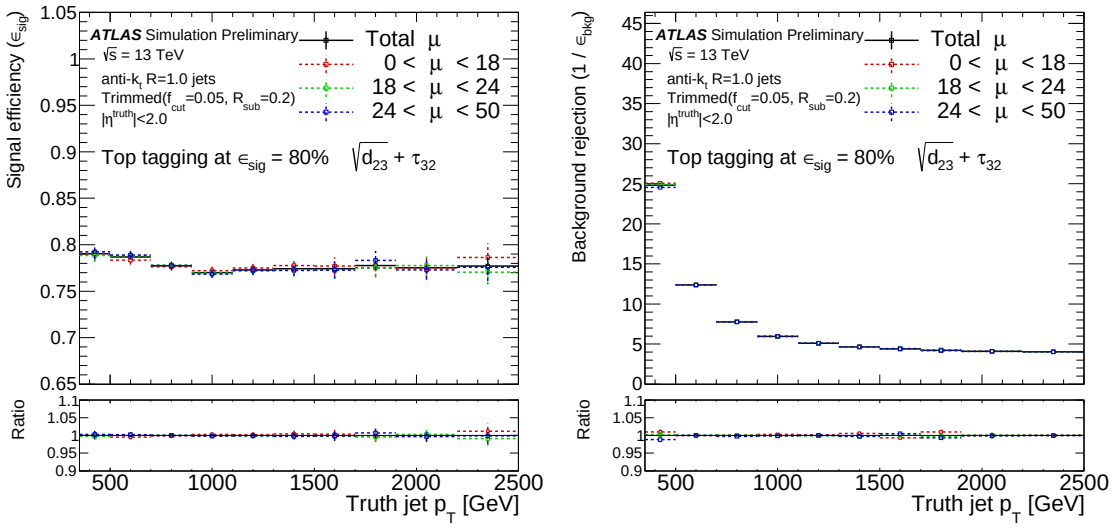
  $$\sqrt{d_{ij}} = min(p_{T,i},p_{T,j}) \times \Delta R_{ij} \tag{4.3}$$

  The splitting scale $\sqrt{d_{23}}$ is determined from the second-to-last clustering step, expected to cover subjets originating from the W boson decay with $\sqrt{d_{23}} \approx \frac{m_W}{2}$ ($\sqrt{d_{12}} \approx \frac{m_t}{2}$)
- **N-subjetiness:**
  The N-subjetiness variables τ_N [91, 92] are an indicator of the subjet multiplicity of a large-R jet. To calculate τ_N the k_t algorithm is applied to a jet, merging the constituents until N subjets are formed. τ_N is defined as the sum of all p_T weighted distances between the constituent k and the closest subjet axis and can be interpreted as a measure of how well the substructure of the jet is described with N subjets. The N-subjetiness ratio $\tau_{32} = \frac{\tau_3}{\tau_2}$ is used to discriminate top events against soft gluons, light quark jets or W bosons.

The signal efficiency and the background rejection regarding dijet events for the Smooth Top Tagger used in the presented analysis with τ_{32} and $\sqrt{d_{23}}$ as discriminating substructure variables is presented in figure 4.8.

4.2.4 Missing transverse energy

Not all particles created in proton-proton collisions interact with the detector material to deposit energy or leave a track hit. For these undetected particles the missing transverse energy (E_T^{miss}) serves as an experimental proxy to estimate the transverse momentum. The reconstruction is challenging as it involves all detector subsystems for the precise reconstruction of all detectable particles of the hard interaction of interest. As the sum of all transverse momentum of all particles involved is zero, due do momentum conservation, a deviation points to a contribution from undetectable particles. Successful strategies have been developed by ATLAS for a high-quality E_T^{miss} estimation, focusing on minimizing the disturbing contribution of pile-up[94].
The reconstruction procedure of E_T^{miss} is based on two contributions, see equation 4.4:



(a) Signal efficiency **(b)** Background rejection

Figure 4.8: Comparison of signal efficiency (a) and background rejection regarding dijet events (b) as a function of the truth jet p_t of the 80% signal efficiency Smooth Top Tagger for different pile-up conditions by using μ bins[93]

- **hard-event signals:**
  fully reconstructed, calibrated particles and jets, so called hard objects.
- **soft-event signals:**
  reconstructed but unused charged particle tracks (soft signals) within the inner detector, associated with the hard scatter vertex assigned to the event.

The basic input for the estimation of the E_T^{miss} are its components E_x^{miss} and E_y^{miss} derived from the hard-event and soft-event signals, 4.4.

$$E_{x(y)}^{miss} = \underbrace{E_{x(y)}^{miss,e} + E_{x(y)}^{miss,\gamma} + E_{x(y)}^{miss,\tau_{had}} + E_{x(y)}^{miss,\mu} + E_{x(y)}^{miss,jet}}_{\text{hard term}} + \underbrace{E_{x(y)}^{miss,soft}}_{\text{soft term}} \tag{4.4}$$

From the components the magnitude of E_T^{miss} can be calculated via 4.5.

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \tag{4.5}$$

5 Data and Monte Carlo Samples

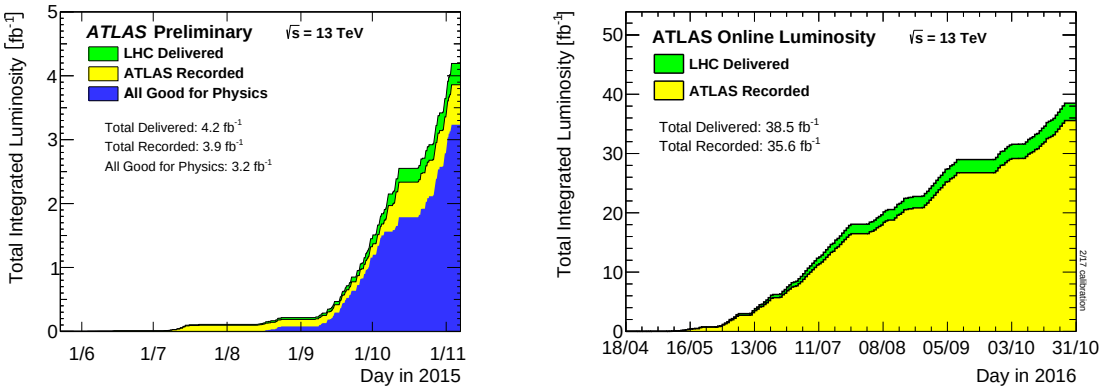
Datasets of particle collisions, collected with the ATLAS detector, are evaluated by comparison with theoretical expectations within the thesis at hand. Based on the known and well established Standard Model or by modeling new theory predictions, simulated events are the foundation of the theoretical expectation. These simulated events are first produced using Monte Carlo (MC) generators, with the transition to physical objects taking place within a full simulation of the ATLAS detector, based on GEANT4 [95]. Pile-up, an occurring effect of several collisions per event in the recorded data, is reproduced within the simulated collisions by overlaying individually generated events.
The choice of simulated signal and background samples is consistent with the published analysis in [3], complemented by the artificial charge asymmetric $t\bar{t}$ quark pair samples for the top charge asymmetry measurement in chapter 10, related to [4]. The explicit listing of Monte Carlo samples and the triggers for the collected data are attached in appendix A. If no reference is provided the Feynman diagrams are created with TikZ[96].

The recorded datasets, evaluated for the direct and indirect search for new physics, are described in section 5.1, while the simulated samples for the signal prediction are presented in section 5.2 (indirect search) and 5.3 (direct search), followed by the simulated background samples in section 5.4 and the multijet estimation in section 5.5.

5.1 Dataset

The direct and indirect search for new physics, presented within the thesis at hand, uses the full data sample collected by the ATLAS detector during the proton-proton collisions in 2015 and 2016 at a center of mass energy of 13 TeV. In total an integrated luminosity of 42.7 fb^{-1} has been delivered with an average of $< \mu >= 23.7$ interactions per bunch crossing, see figure 5.2. Quality criteria are applied to the collected collision data in order to remove the fraction of recorded data, which is unusable due to downtime of individual detector subsystems or identified malfunctions. In 2015 87.1% of the delivered data passed the quality criteria check of the individual subsystems providing $\int \mathcal{L}$=3.2 fb^{-1} of data tagged as *Good for Physics*, while the downtime of the IBL, see section 3.2.2, caused a decrease of 6%, see figure 5.1(a). In 2016 an integrated luminosity of $\int \mathcal{L}$=32.9 fb^{-1} compared to 38.5 fb^{-1} was defined as Good for Physics, see section 3.2.7, resulting in a combined total integrated luminosity of 36.1 fb^{-1}.

Interesting events for the considered analysis at hand are selected from the collision data tagged as Good for Physics by requesting a specific triggers or combinations thereof, see section 3.2.6 for the ATLAS trigger system. Since the signature characteristic of the $t\bar{t}$ quark pair decay in the lepton+jets channel is a single lepton, a trigger to distinguish



(a) Integrated luminosity over time for 2015 **(b)** Integrated luminosity over time for 2016

Figure 5.1: Integrated luminosity delivered to (green) and recorded by (yellow) ATLAS during stable beams for proton-proton collisions at a center-of-mass energy of 13 TeV in 2015 in (a) and 2016 in (b). In blue the data considered as Good for Physics is visualized for 2015 only [97].

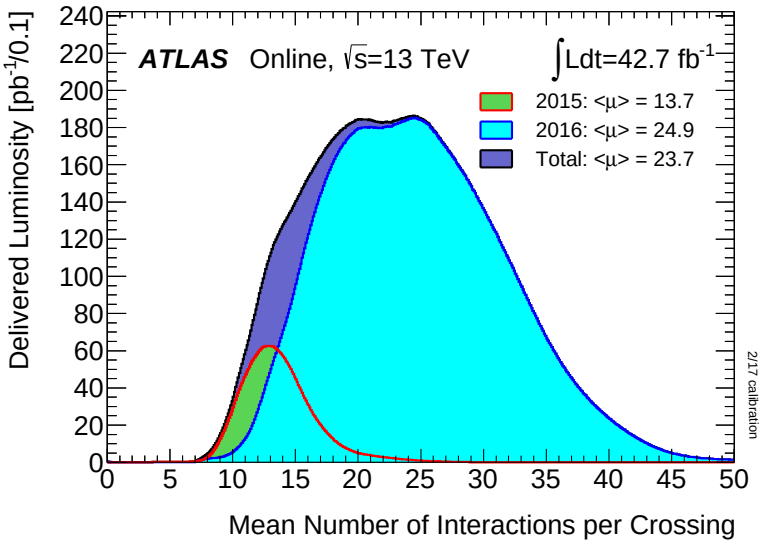


Figure 5.2: Shown is the weighted luminosity distribution of the mean number of interactions per crossing for the combined 13 TeV proton-proton collision data. All data delivered to ATLAS during stable beams is visualized and the integrated luminosity and the mean μ value are provided in the figure [97].

between events with and without a lepton following the basic requirements can be used to decrease the sample size by removing non-relevant collision events. There is no advantage for this analysis in additional triggers, for e.g. jet triggers, since the single-lepton trigger provides the best efficiency for the here considered decay channel. All the variations of single-lepton triggers, utilized for this analysis, are high level triggers, based on Level 1

seeds.
The individual single-lepton triggers are listed in appendix A.1 and are based on individually combined selection criteria on the lepton. Every trigger applies a cut on the lepton p_t. In addition to that some single-lepton triggers apply a criterion based on the lepton identification process and/or on the isolation, and/or on the track information.
For example the single-electron trigger HLT_e24_lhmedium_L1EM20VH applies a cut on the electron p_t of 24 GeV, requests the ID selection MediumLH, see section 4.2.1, and uses the so-called EM20VH seed, corresponding to an electromagnetic trigger with p_t>20 GeV (EM20) and applied hadronic core isolation (H) criterion and a p_t threshold, which varies with η to account for energy loss (V).

5.2 Signal simulation for indirect search for new physics

Monte Carlo samples of the $t\bar{t}$ process are generated to estimate the expected signal events for the validation of the analysis setup as well as to model the uncertainties driven by the choice of the Monte Carlo generator. The Feynman diagram for the corresponding process of $t\bar{t}$ production is displayed in figure 2.1. For the top quark mass a value of m_{top}=172.5 GeV is assumed for all $t\bar{t}$ samples, except for the mass variation samples.

5.2.1 Standard $t\bar{t}$ samples

The Standard model process is simulated by the Powheg+Pythia6 generators. The processing of the hard scattering is described by the Powheg v2[98–100] generator using the CT10[101, 102] parton distribution function (PDF) set. Parton shower, fragmentation and underlying events are simulated using Pythia v6.428[103] with the CTEQ6L1[104] PDF set and the corresponding Perugia 2012[105] set of tune parameters. Tuning is done with the h_{damp} parameter[1] being set to the reference top quark mass of 172.5 GeV.
Mass slices are added orthogonally to the standard $t\bar{t}$ sample to increase statistics in the following five ranges, beginning at a truth $m_{t\bar{t}}$ above 1.1 TeV:

- 1.1 TeV $< m_{t\bar{t}} \leq$ 1.3 TeV
- 1.3 TeV $< m_{t\bar{t}} \leq$ 1.5 TeV
- 1.5 TeV $< m_{t\bar{t}} \leq$ 1.7 TeV
- 1.7 TeV $< m_{t\bar{t}} \leq$ 2 TeV
- 2 TeV $< m_{t\bar{t}} \leq$ 14 TeV

An electroweak correction is applied to account for higher order electroweak processes, see [106] for details.

1 h_{damp} is the parameter used in modeling to control matrix element(ME)/parton shower matching in POWHEG and effectively regulates the high-p_t radiation by controlling the transverse momentum of the first additional emission beyond the Born configuration.

5.2.2 $t\bar{t}$ samples for modeling studies

With the standard $t\bar{t}$ sample defined in 5.2.1, the following sections describe the exploration of different MC generators, considered for the semi-leptonic $t\bar{t}$ event production. The impact of individual generator choices is estimated with respect to the selected standard sample. The resulting modeling uncertainty, see section 8.3.1, provides a dominant contribution to the total uncertainty in the direct and indirect search for new physics, see chapter 9 and 10.

Variation of parton shower generator

To estimate the impact of the parton shower simulation, the standard $t\bar{t}$ sample of this analysis is compared to a $t\bar{t}$ sample with a different generator for the parton showering. In agreement with the standard sample, the hard scattering is processed with the Powheg v2 generator using CT10, while the parton shower, fragmentation and underlying events are simulated using Herwig++ v2.7.1[107] with the same CTEQ6L1 PDF set, but the corresponding CTEQ6L1-UE-EE-5[108] set of tune parameters.

Variation of hard scattering generator

The uncertainty provided by the choice of the hard scattering generator is estimated by comparing the MadGraph5_aMC@NLO v2.2.3 [109] hard scattering generator with the Powheg generator, used within the standard sample. The parton showering for the comparison is in both samples simulated by Herwig++ v2.7.1.

Variation of initial/final state radiation

The variation of the initial and final state radiation is evaluated by selecting different Perugia2012[105] sets of tune parameters, i.e. Perugia2012radLo for a decreased radiation and Perugia2012radHi for an increased radiation, for a $t\bar{t}$ sample otherwise identical to the standard $t\bar{t}$ sample in section 5.2.1. Low radiation is defined by the renormalization (R) and factorization (F) scale $\mu_{R,F}$ varied by a factor of 2.0 and $h_{damp} = m_{top}$, while the scale for high radiation is altered by a factor of 2 and h_{damp}=$2m_{top}$.

5.3 Signal simulation for direct search for new physics

For the direct approach to search for physics beyond the Standard Model the previously in section 5.2.1 defined $t\bar{t}$ sample is considered as background while theoretical BSM models are declared signal and tested for explaining potential deficits/excesses in the data and SM expectation comparison. Interference between the new signal models and the SM $t\bar{t}$ production are not considered, since no strongly interference between the dominant component of the SM $t\bar{t}$ background and the potential signal models is to be expected. For further information see [3, 110].

5.3.1 Spin-1 color singlet - Z'

For the production of a BSM spin-1 color singlet particle decaying into $t\bar{t}$ the $Z' \rightarrow t\bar{t}$ process is simulated with Pythia v8.165[111] with the NNPDF23LO[112] PDF set and the A14[113] set of tune parameters.
For the topcolor-assisted technicolor Z'_{TC2} the samples were normalized to cross-section calculations performed at NLO in QCD[114] using PDF4LHC2015[115].
The hypothetical dark matter mediator particles $Z'_{DM,vec}$ and $Z'_{DM,ax}$ are estimated from the same sample, reweighted to the appropriate resonance width as simulated in MadGraph5_aMC@NLO[109]. The leading order process is displayed in figure 2.4(a).

5.3.2 Spin-2 color singlet - G_{KK}

The production of a BSM spin-2 color singlet particle was simulated using Madgraph5_-aMC@NLO with the NNPDF2.3 LO PDF set for the process $G_{KK} \rightarrow t\bar{t}$. The parton showering was performed by Pythia v8.165 with the A14 set of tune parameters. The leading order process is displayed in figure 2.4(b).

5.3.3 Spin-1 color octet g_{KK}

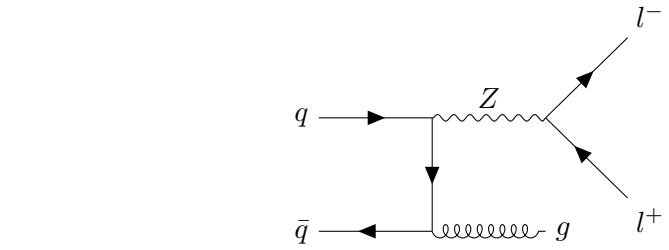
The decay process of the spin-1 color octet particle g_{KK} into $t\bar{t}$ is modeled by Pythia v8.165 at leading order with the NPDF2.3 LO PDF set and the A14 set of tune parameters. The leading order process is displayed in figure 2.4(c).

5.4 Simulated background samples

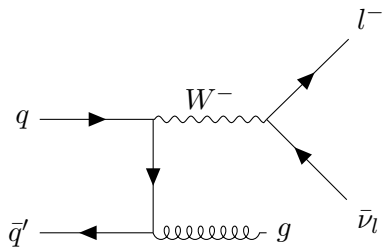
The characteristic decay signature for $t\bar{t}$ processes, i.e. a single lepton, a neutrino and four jets, can be mimicked by different non-top quark related particles or with single top quark events. Therefore the dataset collected by ATLAS is corresponding to a combination of expected signal and background samples generated by Monte Carlo detectors. The relevant backgrounds regarding the signal signature are described in the following subsections.

5.4.1 V+jets and diboson production

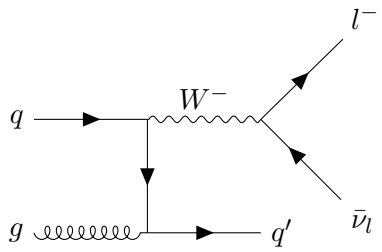
The background consisting of a boson V$\in\{W,Z\}$ with additional jets or two bosons (WW, WZ or ZZ) is generated with Sherpa 2.2.1 [116], using the NNPDF 3.0 PDF set at NNLO. Decay processes similar to the signal with at least one leptonically decaying boson per event are considered as contributing to the background, including all three lepton flavors. For the diboson production WW, WZ or ZZ, with different decay channels ($\ell\nu q\bar{q}$, $\ell\ell q\bar{q}$) orthogonal cuts for the lepton p_t within the individual samples and the heavy flavor of the additional jets are applied by default to guarantee enough statistics for the analysis in all variations.
The leading order Feynman diagrams for the V+jets and diboson processes are visualized in figure 5.3 and figure 5.4 respectively.



(a) Z+jets

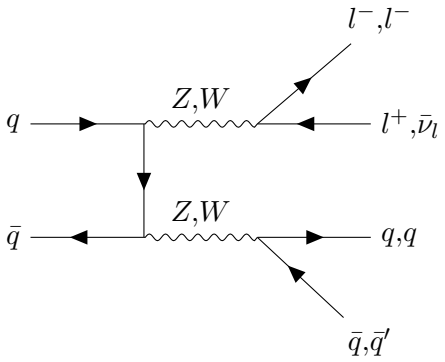


(b) W+jets ($q\bar{q}$)

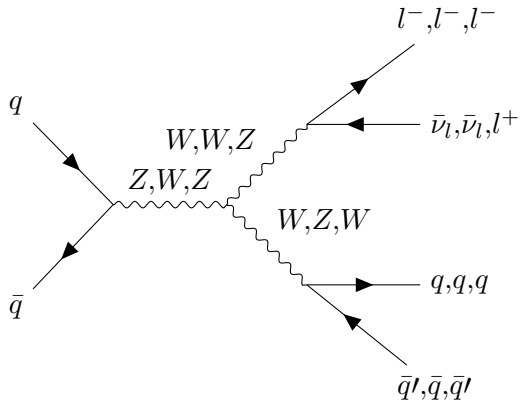


(c) W+jets (qg)

Figure 5.3: Feynman diagrams for V+jets processes for Z+jets (a) and W+jets production (b),(c). The main production channels are visualized, charge conjugation can be applied. Additional jets may result from initial and final state radiation.



(a) VV t-channel



(b) VV s-channel

Figure 5.4: Feynman diagrams for VV processes with V=W,Z in t-channel (a) and s-channel (b). The main production channels are visualized, charge conjugation can be applied.

W+jets flavor filter

Filters to the jet flavor, i.e. the flavor of the quark producing the jet during hadronization, are applied in the production process of the samples. These heavy flavor filters are a combinations of c quark-veto and b quark-veto filters and not suitable to distinguish between W+c and W+cc events, which is necessary for the estimation of the W+jets background

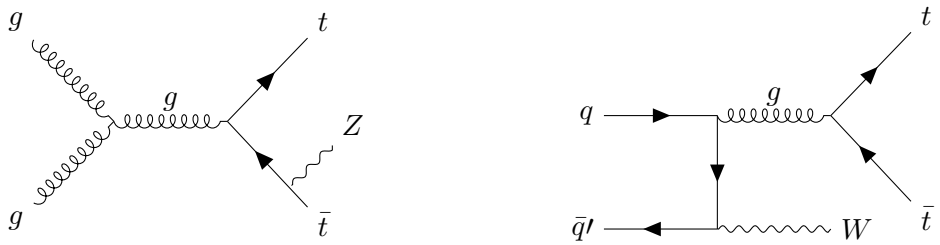
based on the W^+/W^- cross-section ratio, see chapter 7. Therefore a second filter has been implemented and included into the processing of the generated W+jets samples within the Top Analysis framework, to tag the event with the corresponding category, based on the jet flavor.
The four categories deduced from theory, see section 2.4, are

- W+bb
- W+cc
- W+c
- W+lf

The new filter has been developed within the thesis to accommodate the need first for the publication of the direct search and second for the W+jets estimation comparison in the top charge asymmetry measurement.

5.4.2 $t\bar{t}$ +V

The production of a $t\bar{t}$ quark pair with an associated boson V $\in\{Z,W\}$ with up to two additional jets is generated using the MadGraph5_MC@NLO v2.2.2, interfaced with Pythia8.186. The A14 set of tune parameters and the NNPDF23LO PDF set is used. The leading order Feynman diagrams are visualized in figure 5.5

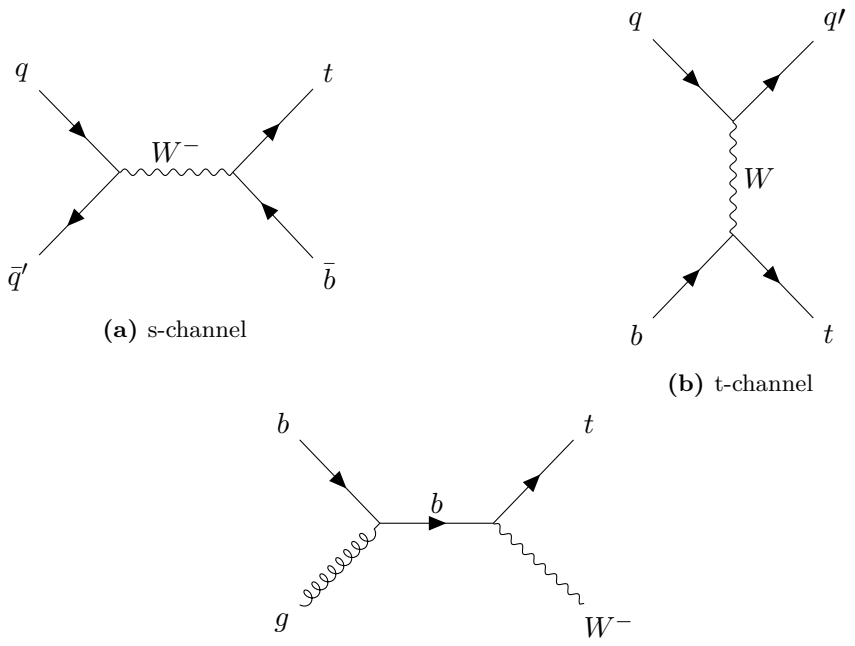


(a) $t\bar{t}$ +Z **(b)** $t\bar{t}$ +W

Figure 5.5: Feynman diagrams for $t\bar{t}$ +V processes for V=Z (a) and V=W production (b). The main production channels are visualized, charge conjugation can be applied.

5.4.3 Single top

The hard scattering processes of single top events in the Wt-[117] and s-channel[118] are simulated with the Powheg v2 generator with CT10 PDF set, similar to $t\bar{t}$ production. Overlap between $t\bar{t}$ and Wt-channel is treated within the diagram removal scheme[119]. The third single top production channel, t-channel, is generated with Powheg-Box v1[120], using the four-flavor scheme for NLO matrix element calculations together with the four flavor PDF set CT10f4. While for Wt- and s-channel the parton showering is simulated with Pythia v6.428 and the CTEQ6L1 PDF set with the Perugia 2012 set of tune parameters, t-channel events are processed with Madspin[121], preserving all spin correlations. The leading order Feynman diagrams for single top processes are provided in figure 5.6.



(a) s-channel

(b) t-channel

(c) Wt-channel

Figure 5.6: Feynman diagrams for single top processes in s-channel (a), t-channel (b) and Wt-channel (c). The main production channels are visualized, charge conjugation can be applied.

5.5 Data driven multijet background estimation

For the background arising from non-prompt leptons[1], mainly from QCD multijet production, the normalization, kinematic distributions and the statistical and systematic uncertainties are estimated from data with a technique called matrix method. The particular matrix method used in the thesis at hand is a variation of the one used in previous ATLAS analyses, described in detail here [122].
The idea of the matrix method is to use the lepton misidentification probability and lepton identification efficiency to estimate the contribution of non-prompt leptons in the signal selection. The efficiency ε is defined as the probability, that a prompt lepton, e.g. from a W or Z boson, fulfills the loose lepton identification criteria, see sections 6.1.1 (el) and 6.1.2 (μ), as well as the tight identification criteria. The efficiency f, the so called *fake rate*, is defined as the probability, that a non-prompt lepton fulfills those two categories as well. With f and ε derived from or validated with data within dedicated control regions, see [110], the estimation of f and ε was provided externally for the publication, see [3], and the thesis at hand.

1 A *non-prompt* lepton is either a lepton not originating from the main collision, e.g. produced within jets or from heavy particles, or a jet, mis-identified as a lepton

The number of events with leptons fulfilling the loose identification criterion , see section 4.2.1 and 4.2.1 respectively, N_{loose}, is defined in equation 5.1 with N_{prompt} and N_{QCD} equal the number of events satisfying those criteria with prompt and non-prompt leptons from QCD multijet events respectively.

$$N_{loose} = N_{prompt} + N_{QCD} \tag{5.1}$$

The number of events fulfilling the tight lepton identification criteria, N_{tight}, is defined as

$$N_{tight} = \varepsilon \cdot N_{prompt} + f \cdot N_{QCD}. \tag{5.2}$$

The number of anti-tight events $N_{anti-tight}$, where leptons passing the loose identification criterion fail the tight lepton selection, can be defined similarly via

$$N_{anti-tight} = (1 - \varepsilon) \cdot N_{prompt} + (1 - f) \cdot N_{QCD}. \tag{5.3}$$

Solving the two equations 5.2 and 5.3 for N_{prompt} and N_{QCD} in equation 5.4 enables the assessment of the QCD contribution in the signal region. The QCD multijet background is estimated from data events, that pass all the signal selection criteria, except that the loose lepton definition is used instead of the tight one. Data contains events from prompt-lepton sources and QCD multijet events both. The multijet contribution in the signal region can be estimated as

$$f \cdot N_{QCD} = \frac{\varepsilon - 1}{\varepsilon - f} N_{tight} + \frac{\varepsilon \times f}{\varepsilon - f} N_{anti-tight}. \tag{5.4}$$

In addition to the overall yield estimation kinematic distributions, necessary for the mass spectrum and the W+jet background estimation method, can be provided for the multijet background by applying a weight $w(N_{tight},N_{anti-tight})$ to each event. Equation 5.4 provides the weight $w(N_{tight},N_{anti-tight})$, where the tuple $(N_{tight},N_{anti-tight})$ is either (1,0) or (0,1) depending on whether the data event passes the tight or fails the tight selection, but passes the loose selection.

A good modeling of the shape of the kinematic distributions, in the signal region for the analysis as well as in the the control region within the W+jets background estimation is mandatory. The best modeling with focus on the search for resonances in the $t\bar{t}$ mass spectrum is achieved by parameterizing the efficiencies f and ε as a function of relevant kinematic variables. For the electron channel a two-dimensional function of the lepton p_t and a calorimeter based isolation variable provides the best result, while for the muon channel an additional parameter, the angular separation between the lepton and the closest jet is necessary to achieve satisfying result, see [110] for details.

The modeling has been validated externally in separate dedicated regions, where one of the multijet excluding selection criteria of the signal selection cuts, see section 6.2 , i.e. E_T^{miss} or $E_T^{miss} + m_T^W$, is inverted. For further information see [3]. For the thesis at hand the provided f and ε from [3] have been used to estimate the QCD multijet background for the pre-tag control region within the W+jet background estimation as well as for the signal region for the direct and indirect search for BSM physics.

6 Analysis

Focusing on the lepton+jets topology, this chapter describes the core aspects of the analysis: the object selection in section 6.1, the event selection and reconstruction for the top quark pair candidates in section 6.2. The event yields are presented in section 6.2.3. The presented cuts and selections are applied to the direct and indirect search for new physics, described in this thesis. The analysis in the context of the search for $t\bar{t}$ resonances has been published in [3], with auxiliary material in [110].

6.1 Object selection

Before reconstructing the top and antitop candidates, the necessary building blocks, the physical objects, see chapter 4, are selected requiring specific quality criteria. This object selection is defined by the setup of the top-antitop resonance search presented in [3] and follows the recommendations of the ATLAS TopReconstruction group[67]. These prescriptions are used for a wide range of analyses involving $t\bar{t}$ decays in order to ensure a common standard and compatibility, besides gaining an advantage from the optimized selection based on previous experience.

6.1.1 Electrons

Besides a general selection, two electron quality criteria are introduced in this section: Loose, for multijet background estimation only, and Tight otherwise. The electrons passing the corresponding object selection will be called *selected electrons* further on.
General quality cuts, see table 6.1, are applied to ensure the basic kinematic coverage in the central region of the detector regarding the transverse momentum, p_T, and the pseudo rapidity, η. The association of the electron candidate track to the collision vertex is handled via a requirements to the significance of the transversal, d_0, and restrictions to the longitudinal, z_0, impact parameter. To ensure a fully functional detector environment electrons with clusters affected by dead front end boards in the liquid argon calorimeter or dead high-voltage (HV) regions, so called *BADCLUSELECTRONS*, are removed via bitwise comparison to an object quality map.
The loose and tight electrons differ from each other in the electron identification, see section 4.2.1, and the isolation criterion. For the tight selection isolated electrons are required, where the total transverse momentum of all tracks, excluding the track belonging to the electron, within a variable-cone size with radius ΔR

$$\Delta R = \min\left(\frac{10\text{GeV}}{p_T}, R_{max}\right), \tag{6.1}$$

is below a certain threshold, defined via a fraction of transverse momentum of the electron. For the tight electron selection criteria a reconstruction efficiency of 90% at p_T=25 GeV and 99% at $p_T t$=60 GeV is reached [4].

General selection	
Transverse momentum	p_T>25 GeV
Pseudo rapidity	$\|\eta\|$<2.47, excluding calorimeter transition region: 1.37<$\|\eta\|$<1.52
Impact parameters	$\|d_0\|/\sigma(d_o)$<5 $\|z_0 \sin\theta\|$<0.5 mm
Dead calorimeter regions	BADCLUSELECTRON removal

Loose selection	
Identification	MediumLH

Tight selection	
Identification	TightLH
Isolation	$\sum p_T$ in ΔR of 0.2/p_T<0.06

Table 6.1: Electron quality requirements applied to the electron selection to identify suitable electrons for the multijet estimation (Loose) and the analysis (Tight)

6.1.2 Muons

The muon selection follows a similar procedure as for electrons described in the previous section 6.1.1, with the specific cuts listed in table 6.2. The general selection requires the Combined Muon reconstruction criteria, see section 4.2.2, and suppresses muons with a high probability to have a bad momentum resolution, by flagging those as "bad". The Loose and Tight muons are distinguished via the identification and the isolation criterion. For the tight selection the average identification efficiency for muons is 98% [4].

6.1.3 Jets

Looking at top pair decays with a resolved topology, small-R jets are a requirement to reconstruct the at least four jets in total from the hadronic W decay, and the two b-quark jets directly from the top decay, see section 2.2.2. With a center-of-mass energy of 13 TeV, the $t\bar{t}$ system can be boosted, i.e. the decay products of the hadronically decaying top collimate in one large-R jet and can not be reconstructed individually anymore. Identifying jets originating from the hadronization process of the b-quark improves the signal clarity

General selection	
Reconstruction	combined
Transverse momentum	p_T>25 GeV
Pseudorapidity	$\|\eta\|$<2.7,
Impact parameters	$\|d_0\|/\sigma(d_0)$<3 $\|z_0 \sin\theta\|$<0.5 mm
Veto	isBadMuon

Loose selection	
Identification	Loose

Tight selection	
Identification	Medium
Isolation	$\sum pT$ in ΔR of 0.3)/p_T<0.06

Table 6.2: Quality requirements applied to the muon selection to identify suitable muon for the multijet QCD estimation (Loose) and the analysis (Tight)

and acquires a better track resolution by using track instead of calorimeter jets. Therefore three different jet collections are defined within the analysis:

- small-R jets, R=0.4
- large-R jets, R=1.0
- small-R track jets, R=0.4

All jet collections are clustered via the anti-k$_t$ algorithm with the applied kinematic cuts listed in table 6.3. For the small radius jets the Jet Vertex Tagger (JVT)[123] is a supplementary quality criterion to suppress pile-up applied to jets within the kinematic range of p_T<60 GeV and $|\eta|$<2.4. The chosen medium working point has an average efficiency of 92% and a rejection rate for pile-up and noise of 98%. For the large-R jets the reduction of pile-up effects is achieved with the trimming procedure described in section 4.2.3. The anti-k$_t$ and k$_t$ algorithms have been implemented using FastJet[124].

To the small-R track jets b-tagging requirements are applied, see section 4.2.3, using the MV2 algorithm at a working point with 77% efficiency for the tagged jet being a b-jet. A purity of 95% and a rejection[1] of c- (light-) jets of 6.2 (134.3) respectively is achieved[125, 126].

1 The rejection is the inverse of the probability, that a given non-x jet is tagged as x, $x \in$(c, light). A rejection of 100 means that one in 100 jets will be mistagged.

Small radius jets	
Clustering algorithm	anti-k$_t$ on EM topocluster
Jet radius	R=0.4
Transverse momentum	25 GeV
Pseudorapidity	2.5
JVT$_{min}$	JVT>0.59

Small radius track jets	
Clustering algorithm	anti-k$_t$ on tracks
Jet radius	R=0.4
Transverse momentum	10 GeV
Pseudorapidity	2.5

Large radius jet	
Clustering algorithm	anti-k$_t$ on LC topocluster
Jet radius	1
Transverse momentum	100 GeV
Pseudorapidity	2.0
Trimming	R_{sub}=0.2, f_{cut}=0.05

Table 6.3: Quality requirements for the three jet containers used during the analysis

6.1.4 Overlap removal

With a reconstruction procedure based on tracks and energy deposits in the calorimeter, electrons and muons are also reconstructed as jets or might share the same information in their reconstruction process. To discard duplicates of the same physical object and to isolate electrons and muons from jets, an overlap removal procedure is applied.
In case of shared tracks between selected muons and electrons, the electron is rejected as it is most likely a bremstrahlung photon. For the duplicate removal between electrons and jets, the closest jet within ΔR<0.2 of the electron is removed and in a consecutive step electrons within a distance of ΔR <0.4 to a jet are discarded.
For the muon-jet overlap removal the distance parameter is depending on the muon's transverse momentum with ΔR<0.04+10 GeV/p_T^{μ} to target the removal of jets faked by muons through bremsstrahlung. If the nearest jet to the muon within this range has fewer than three associated tracks, the jet is removed, otherwise the muon is discarded instead, because it is likely to be non-prompt from hadron decays in a real jet. This

procedure avoids inefficiencies for high energetic muons with a significant energy loss in the calorimeter, registered as jets[3].

6.2 Event selection and reconstruction

After the description of a common event selection in section 6.2.1 to achieve a high data quality and utilizing fundamental cuts to ensure a $t\bar{t}$ event signature the events are further categorized into a resolved and a boosted decay topology. In section 6.2.2 the reconstruction mechanism for the neutrino is described, as it is independent of the event categorization. The necessary cuts implying a boosted decay are described in subsection 6.2.3, the second category, the resolved decay topology, is detailed in subsection 6.2.4.
For the search for new particles decaying into top-pairs and for the measurement of the charge asymmetry in top-pair decays, a precise reconstruction of the decay products is fundamental for the analysis. The top-quark-pair reconstruction method is presented combined with the selection cuts in the sections for resolved and boosted accordingly.

6.2.1 Common event selection

Following the initial pre-selection applied by the choice of single-lepton triggers in section 5.1, the event selection applied to all events consists of the following steps.

Event quality

At this selection stage only events recorded while all parts of the detector operated correctly are allowed. A centrally provided Good Runs List flags events with e.g. incomplete information, turned off detector components or corrupted calorimeter cells to be excluded from any analysis and the remaining events are flagged "good for physics", see section 3.2.7. Only events without any noise bursts detected in the EM or magnetic calorimeter are considered for the analysis, ensured by a so-called "GOODCALO" flag[110]. In addition to that a primary vertex, i.e. the vertex with the highest sum of squared transverse momentum of the tracks assigned to it and at least two tracks, is required.

Charged Lepton selection

With the focus on the lepton+jets decay mode, a single lepton (electron or muon) is used as a signal criterion to identify the corresponding events next. At first the lepton has to have a transverse momentum of at least 30 GeV, second no additional electrons or muons with $p_T \geq 25$ GeV are allowed. With the event passing the single-lepton trigger, the selected lepton has to to match the trigger decision to fully pass the charged lepton selection.

Leptonic W selection

Due to the required leptonic decay of one of the W-bosons, the event is expected to have a charged lepton and missing transverse momentum. Additionally a jet cleaning event veto, see [127] for details, is applied, preventing jets defined as bad from affecting the E_T^{miss}

calculation. This is not only affecting the jets in the chosen three jets containers, but all jets considered for the E_T^{miss} calculation.
The event selection based on the leptonic W kinematics is achieved by the following two criteria

- $E_T^{miss} > 20$ GeV
- E_T^{miss}+m_T^W> 60 GeV

with the transverse mass of the W, m_T^W, being defined as

$$m_T^W = \sqrt{2p_T^\ell E_T^{miss}(1 - \cos \Delta\phi(\ell, E_T^{miss}))} \tag{6.2}$$

using the transverse momentum of the lepton,p_T^ℓ and the angular distance in ϕ between lepton and the missing transverse momentum.

b-tagging

The b-tagging requirement is vital to separate signal events from background events with only light flavor jets, multijet background and Z+jets,. Although the $t\bar{t}$ decay contains two b-jets, the condition is loosened to at least 1 b-tagged jet to take into account the comparatively low efficiency of the method of 77%[125].

Classification into Boosted or Resolved selection

Fulfilling the later selection criteria the event is afterwards tested against the boosted selection and - in case of failure - against the resolved selection. Double counting of events is avoided with the preference set to boosted. The specific selection criteria are described in the following sections 6.2.3 and 6.2.4.

6.2.2 Common reconstruction: Neutrino

Neutrinos can not be detected directly within the ATLAS detector, because the are unlikely to interact with the detector material. The transverse momentum of the single neutrino within the $t\bar{t}$ decay can be estimated from the missing transverse momentum of the event. With the x- and y-component of the neutrino momentum acquired from E_T^{miss}, a solution for the z-component has to be found.
Under the assumption that neutrino and lepton originate from an on-shell W-boson decay and including a W-mass constraint, an equation system can be formed to estimate the z-component of the neutrino. Are both solutions real, the choice depends on the topology, see the corresponding choice for the boosted and resolved topology in subsections 6.2.3 and 6.2.4. In case of no real solution a mis-measurement of E_T^{miss} is assumed and E_T^{miss} is minimally rescaled and rotated until a real solution is found. For a detailed description of the procedure see [128].

6.2.3 Boosted selection and reconstruction

Boosted $t\bar{t}$ decays are characterized by the decay products of the hadronically decaying top being collimated into a single large jet, while the leptonically decaying partner is fully resolved and under ideal circumstances back to back with the boosted jet. This topology is more likely with a high mass of the $t\bar{t}$ system.

Leptonic-top b-jet:

To pass the boosted selection, the event is required to contain at least one small-R jet with $\Delta R(\text{jet},l)<1.5$. In case of more than one jet matching this criterion, the jet highest in p_T is chosen and further on defined as *selected jet*. This jet represents the b-jet from the leptonically decaying top, although no b-tagging requirement is enforced.

Hadronic-top jet:

The boosted jet candidate representing the hadronically decaying top quark is a large-R jet, which passes the following criteria:

- $p_T>300$ GeV
- $\Delta\phi(\text{jet}_{R=1.0},\ell)>2.3$
- $\Delta R(\text{jet}_{R=1.0}, \text{selected jet})>1.5$

Additionally top-tagging is applied using p_T and $|\eta|$ dependent cuts on the subjetiness τ_{32} and the splitting scale $\sqrt{d_{32}}$, see section 4.2.3, at a working point with an expected efficiency of 80%.

Boosted top candidates reconstruction

In the boosted topology the large-R jet represents the hadronic top, while the selected jet, lepton and neutrino form the leptonic top. They are distinctly identified during the event selection. In case of more than one large-R jet passing the top-tagging requirements, the highest transverse momentum large-R jet is taken. With a possibility of having two real solutions for the z-component of the neutrino, the smallest absolute $|p_z|$ is chosen. The reconstruction of the top quark candidates is achieved by combining the four-vectors of the reconstructed objects respectively, while the top-pair system is formed by combining the two top-candidates, providing the invariant mass $m_{t\bar{t}}$.

6.2.4 Resolved selection and reconstruction

If the boosted selection fails, the event is tested against the criteria for the resolved selection. The event must contain at least four small-R jets and the χ^2 algorithm for the reconstruction of the $t\bar{t}$ system, see section 6.2.4, has to yield a value of $\log_{10}(\chi^2)<0.9$, to pass the resolved selection.

Resolved top candidates reconstruction

The resolved topology consists of a lepton, a neutrino and a small-R jet, representing the leptonic top decay identical to the boosted topology. In addition to that three small-R jets, two of them originating from the W-boson and one from the b-quark of the top decay, form the hadronic top candidate In contrast to the boosted topology these small-R jets are expected to be resolved individually. To assign the at least four small-R jet candidates from the resolved event selection to the respective top quark candidate a minimal χ^2 is constructed to find the best combination, see

$$\chi^2 = \underbrace{\left[\frac{m_{jj} - m_W}{\sigma_W}\right]^2}_{\text{hadronic W}} + \underbrace{\left[\frac{m_{jjb} - m_{jj} - m_{t_h-W}}{\sigma_{t_h-W}}\right]^2}_{\text{hadronic top}} + \underbrace{\left[\frac{m_{j\ell\nu} - m_{t_\ell}}{\sigma_{t_l}}\right]^2}_{\text{leptonic top}} + \underbrace{\left[\frac{(p_{T,jjb} - p_{T,j\ell\nu}) - (p_{T,t_h} - p_{T,t_\ell})}{\sigma_{\Delta p_T}}\right]^2}_{\text{top quark transverse momentum}}. \tag{6.3}$$

The first term represents the constraint from the hadronically decaying W-boson, while the second one represents the hadronic top decay, with the W-boson being subtracted to decouple it from the first term. Next is the constraining term formed by the leptonic top quark followed by the constraint of the top quark candidates to have similar transverse momenta. For the b-jet candidate only jets passing the b-tag requirement are considered. The parameter values for the central-values m_W, m_{t_h-W}, m_{t_ℓ}, m_{t_ℓ} and the width values σ_W, σ_{t_h-W}, σ_{t_ℓ} and $\sigma_{\Delta p_T}$ are estimated from Gaussian fits to the distributions of relevant reconstructed variables, using MC events for which the reconstructed objects can be matched to partons with the MC truth information[1], see[3].

6.3 Yields and comparison

With the events selection, see section 6.2, applied, table 6.4 presents the event yields for the resolved topology and table 6.5 for the boosted topology. Here only the statistical uncertainties are displayed for the base of the in-situ calibration of the backgrounds used in the direct and indirect search presented in the work at hand.
In figures 6.1 and 6.2 exemplary MC-Data comparisons are presented for relevant distributions for the $t\bar{t}$ resonance and top charge asymmetry analysis including the total uncertainties. The so called total uncertainties here represent the dominant uncertainties for the top charge asymmetry measurement and combine the statistical uncertainty and the systematic uncertainties of the cross-sections, $t\bar{t}$ modeling and W+jets scale factors, see chapter 8. Further uncertainties, which are identified as significant in the search for resonances in the $t\bar{t}$ mass spectrum, see table 9.2, are not included. The distributions for

1 The parameter values are: $m_W = 80.51$ GeV, $m_{t_h-W} = 85.17$ GeV, $m_{t_\ell} = 167.36$ GeV, $\sigma_W = 12.07$ GeV, $\sigma_{t_h-W} = 16.05$ GeV, $\sigma_{t_\ell} = 25.41$ GeV, $(p_{T,t_h} - p_{T,t_\ell}) = 0.23$ GeV and $\sigma_{\Delta p_T} = 18.85$ GeV[110]

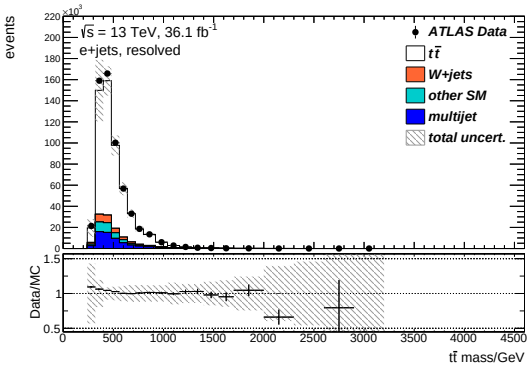
observed data and Standard Model expectations agree within the covered uncertainties for both channels.

Source	e+jets			μ+jets			combined		
$t\bar{t}$	443600	±	400	439200	±	400	882800	±	600
single top	21530	±	90	21340	±	90	42870	±	130
W + bb,cc	19640	±	290	23000	±	400	42600	±	500
W + c	4880	±	140	4840	±	130	9720	±	190
W + lf	2060	±	130	1990	±	140	4050	±	190
Z+jets	7470	±	100	3680	±	60	11150	±	120
diboson	1480	±	40	1450	±	40	2930	±	50
$t\bar{t}$ +V	1010	±	10	990	±	10	2000	±	10
multijet	58200	±	800	24500	±	700	82700	±	1000
Total	559900	±	1000	521000	±	900	1080900	±	1300
Data	580194			569977			1150171		

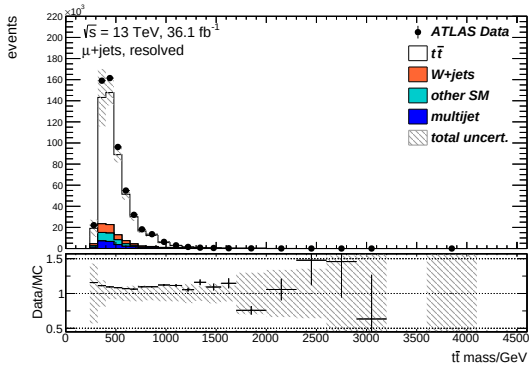
Table 6.4: Event yields for data and Standard Model expectation after the application of the event selection for the resolved topology, including only stat. uncertainties.

Source	e+jets			μ+jets			combined		
$t\bar{t}$	35750	±	120	32740	±	110	68490	±	170
single top	2220	±	40	1910	±	30	4120	±	40
W + bb,cc	1750	±	20	2000	±	20	3750	±	30
W + c	620	±	10	600	±	10	1220	±	20
W + lf	250	±	10	230	±	10	480	±	10
Z+jets	370	±	10	250	±	10	620	±	10
diboson	460	±	20	370	±	20	830	±	20
$t\bar{t}$ $t\bar{t}$ +V	320	±	10	300	±	10	620	±	10
multijet	4680	±	290	1400	±	160	6070	±	330
Total	46410	±	320	39790	±	200	86200	±	370
Data	40649			35701			76350		

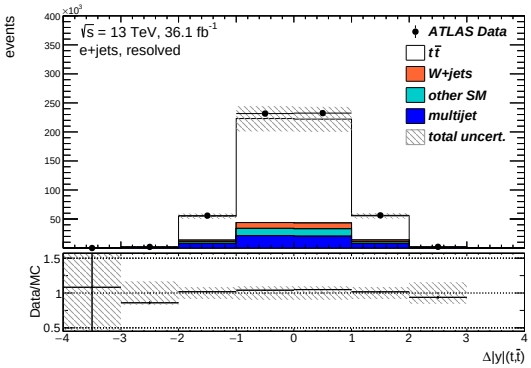
Table 6.5: Event yields for data and Standard Model expectation after the application of the event selection in the boosted topology, including only stat. uncertainties.



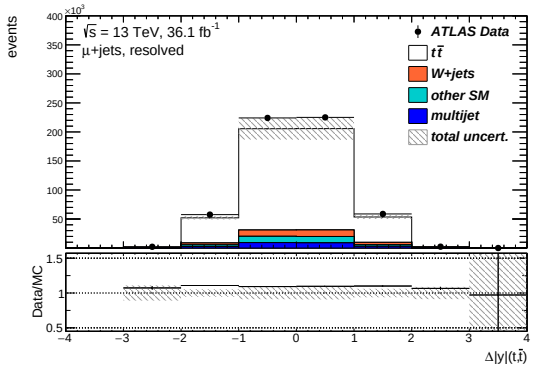
(a) invariant $t\bar{t}$ mass distribution in e+jets channel



(b) invariant $t\bar{t}$ mass distribution in μ+jets channel

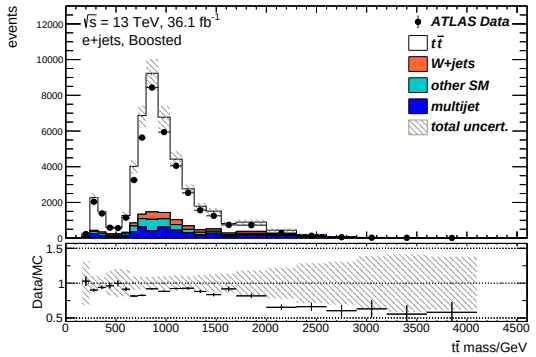


(c) $\Delta|y|$ distribution in e+jets channel

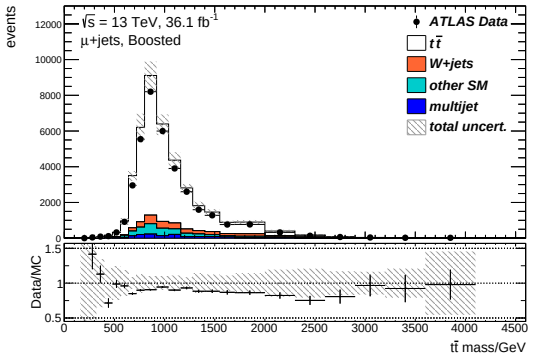


(d) $\Delta|y|$ distribution in μ+jets channel

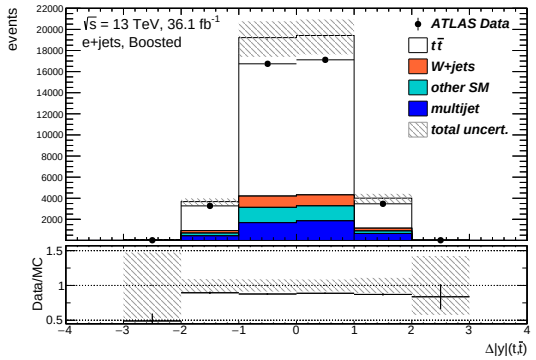
Figure 6.1: Comparison of data and Standard Model expectation for the reconstructed invariant mass of the $t\bar{t}$ system and the $\Delta|y|$ distribution in the e+jets channel ,(a) and (c), and the μ+jets channel ,(b) and (d). Total uncertainties are the combination of statistical uncert. and systematic uncert.



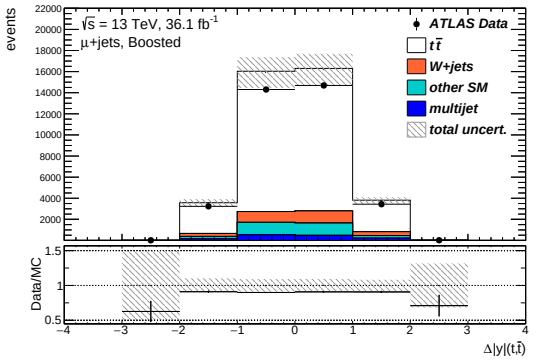
(a) invariant $t\bar{t}$ mass distribution e+jets channel



(b) invariant $t\bar{t}$ mass distribution μ+jets channel



(c) $\Delta|y|$ distribution in e+jets channel



(d) $\Delta|y|$ distribution in μ+jets channel

Figure 6.2: Comparison of data and Standard Model expectation for reconstructed invariant mass of the $t\bar{t}$ system and $\Delta|y|$ distribution in the e+jets channel, (a) and (c), and the μ+jets channel, (b) and (d). The small peak at low masses in the $t\bar{t}$ mass spectrum in the e+jets channel originates in jet miss-assignment or miss-reconstruction and is well modeled in the simulation. A further investigation is evaluated as redundant.

7 Data driven W+jets background estimation with flavor scale factors

Since the total amount of W-bosons produced with associated jets, W+jets, is not precisely predicted in Monte Carlo, a data-driven approach is used to estimate this background. The chosen method is based on an overall normalization factor and an individual heavy flavor scale factor for the different flavors of the additional jets. First descriptions of the method used for $\sqrt{s} = 8$ TeV can be found in [129], an improved method for the search for $t\bar{t}$ resonances is published in [130] and [131].

In this study the estimation method is adapted to the conditions of data taken with $\sqrt{s}$ = 13 TeV. The final scale factors represent the main contribution to the published analysis for the direct search for new particles [3] and are further used in the study of additional use cases within the $t\bar{t}$ charge asymmetry measurement in chapter 10.

The overall normalization of the W+jets background is based on the theoretically well understood ratio of the cross sections for W^+ and W^-, see section 2.4. A short description of the normalization procedure via charge asymmetry is presented in section 7.1. In section 7.2 the extraction of the individual flavor scale factors from a control region within the resolved event selection is described and the resulting combined scale factors are presented in 7.3. W+jets specific uncertainties and the treatment of general uncertainties are discussed in 7.4. The final section 7.5 is dedicated to closure tests for the individual scaling factors.

7.1 W+jets normalization

The data driven charge asymmetry normalization deduces the estimated total amount of W+jets in data, $N_{W,Data}$ and corrects the amount of simulated W+jets events, $N_{W,MC}$ accordingly. This is achieved by comparing the asymmetry in the corrected data yields, $D_{asym,pm}$, which remain after the subtraction of asymmetric non-W+jets MC backgrounds, to the expected asymmetry of W+jets in MC, r_{MC}, see equation 7.1. While a contribution to the asymmetry is expected from the MC samples $t\bar{t}$ +V, single top (s-,t-channel) and diboson, the remaining backgrounds, $t\bar{t}$, single top (Wt-channel), multi-jet and Z+jets, are treated as charge symmetric.

$$N_{W,Data} = \left(\frac{r_{MC}+1}{r_{MC}-1}\right)(D_{asym,+} - D_{asym,-}) \tag{7.1}$$
$$\text{with } r_{MC} = \frac{N_{W^+,MC}}{N_{W^-,MC}} = \frac{\sigma(pp \to W^+)}{\sigma(pp \to W^-)}$$

As a consequence of this a factor for the so called charge asymmetry normalization, CA, can be extracted as

$$CA = \frac{N_{W,Data}}{N_{W,MC}}. \tag{7.2}$$

Separating the MC samples into symmetric and asymmetric with respect to the lepton charge is based on the asymmetry value, A_c[1], compatible with zero, see table 7.1. Contributions of the symmetric MC processes to the asymmetry in data are considered as an uncertainty, see section 7.4.

excl. 4-jets $^-$-channel	$\mathbf{A_c}$	$\mathbf{\Delta A_c}$	asymmetric
$W + bb$	13.7%	0.3%	✓
$W + cc$	12.4%	0.8%	✓
$W + c$	5.8%	0.7%	✓
$W + lf$	15.6%	0.7%	✓
single top (s,t)	22.7%	0.5%	✓
$t\bar{t}$ +V	5.4%	0.3%	✓
diboson	1.9%	0.7%	✓
QCD	0.4%	0.9%	
single top (Wt)	0.1%	0.3%	
$t\bar{t}$	0.0%	0.1%	
Z+jets	0.0%	0.7%	

Table 7.1: The asymmetry A_c of all MC samples is listed for events with exactly four jets in the *mu*-channel in the pretag region, including the statistical uncertainty.

The W+jets normalization is obtained from a sample of events selected without b-tagging requirement, so called *pretag*, and before the application of topology specific selection cuts, see sections 6.2.3 and 6.2.4. Three different control regions are defined, based on the jet multiplicity: 2 exclusive, 4 exclusive and 5 inclusive jets. For these the events are split according to their lepton charge.

The total input for the charge asymmetry normalization procedure of the W+jets MC samples is listed in table 7.2 (e) and table 7.3 (μ).

CA normalization control region e+jets			
jet flavor	**2ex**	**4ex**	**5in**
W+bb	738100 ± 2000	91600 ± 400	46010 ± 170
W+cc	507400 ± 3200	90400 ± 600	50150 ± 300
W+c	1624000 ± 7000	140000 ± 1000	53600 ± 400
W+lf	4870000 ± 40000	326900 ± 2800	122200 ± 800

Table 7.2: Event yields in the e+jets channel of the W+jets MC sample in the charge asymmetry (CA) normalization control region, statistical uncertainties only.

1 The asymmetry A_c is based on event counts of events with positive, N^{l+}, or negative lepton charge, N^{l-}, with $A_c = \frac{N^{l+}-N^{l-}}{N^{l+}+N^{l-}}$.

CA normalization control region μ+jets			
jet flavor	**2ex**	**4ex**	**5in**
W+bb	722800 ± 2100	94640 ± 330	49700 ± 170
W+cc	550000 ± 4000	97500 ± 800	54430 ± 300
W+c	1770000 ± 9000	149600 ± 1100	57500 ± 400
W+lf	5570000 ± 40000	380600 ± 2500	139200 ± 800

Table 7.3: Event yields in the μ+jets channel of the W+jets MC sample in the charge asymmetry (CA) normalization control region, muon channel, statistical uncertainties only

7.2 Flavor scale factors

For the different jets flavors in the W+jets samples individual flavor scale factors are estimated in a b-tagged (tag), lepton charge separated control region and extrapolated to the signal region. The control region for the flavor scale factors the same as for the charge asymmetry normalization, except for the additional b-tag requirement.
Due to a large contamination with multijet background in the flavor control region with an absolute uncertainty (50%) larger than the total W+c jets contribution, a correction factor for the QCD background is used in the 2 exclusive tag control region only. While the charge contribution difference can be used to quantify the $W+bb$, $W+cc$ and $W+lf$ contribution, the $W+c$ jets and the multijet contribution are both (almost) charge symmetric and therefore indistinguishable. This procedure can not estimate the $W+c$ jets contribution while simultaneously fitting the QCD contribution. As a consequence of this the $W+c$ flavor scale factor K_c is confined to 1.00 during the fit. The bb and cc flavored W+jets contributions are combined and treated with a single flavor scale factor, $K_{bb,cc}$, because their asymmetric behavior is similar, due to identical production mechanisms, see section 2.4, and are estimated together with K_{lf} and K_{QCD} within this method.
The flavor scale factors and the QCD correction factor are estimated by solving the equation system described in equation7.3. There the number of W+jets MC events for the individual subsamples, $N^{\ell\ charge}_{MC,<flav.>}$, scaled with the charge asymmetry normalization factor CA and the flavor scale factors K_i, is compared to the corrected data yields D_{corr}, i.e. the data yields minus all non-W MC samples, but including QCD. In addition the sum of the flavor fractions $f_{<flav.>}$, with $flav.(flavor) \in \{bbcc,c,lf\}$, estimated from W+jets MC yields in the 2 excl. pretag region, is forced to be equal 1 after scaling with K_i, so that it does not change the total W+jets yields, as this is the purpose of the normalization factor. The third line in equation 7.3 is dedicated to the $W+c$ flavor scale factor, which is fixed to 1.00 in the control region. After 10 iterations of solving the equation system, including the recalculation of the CA normalization factor, the results have proven to be stable within the uncertainties.

$$\begin{pmatrix} CA \cdot N^{-}_{W_{bb,cc}} & CA \cdot N^{-}_{W_c} & CA \cdot N^{-}_{W_{lf}} & N^{-}_{QCD} \\ f_{bb+cc} & f_c & f_{lf} & 0 \\ 0 & 1 & 0 & 0 \\ CA \cdot N^{+}_{W_{bb,cc}} & CA \cdot N^{+}_{W_c} & CA \cdot N^{+}_{W_{lf}} & N^{+}_{QCD} \end{pmatrix} \begin{pmatrix} K_{bb,cc} \\ K_c \\ K_{lf} \\ K_{QCD} \end{pmatrix} = \begin{pmatrix} D^{-}_{corr} \\ 1.0 \\ 1.0 \\ D^{+}_{corr} \end{pmatrix} \tag{7.3}$$

After the estimation of the flavor scale factors $K_{<flav>}$ for the 2 excl. region the new flavor fractions can be calculated via

$$f^{2ex,final}_{<flav>} = f^{2ex}_{<flav>} \cdot K^{<2ex}_{<flav>}. \tag{7.4}$$

The flavor scale factors are extrapolated to the 4 exclusive and 5 inclusive jet region, which are representing the common signal region for analysis with top-pair decays. For the resulting flavor scale factors $K^{<\#jets>}_{<flav>}$ the flavor fractions $f^{final,<\#jets>}_{<flav>}$ need to be normalized within each jet multiplicity, which can be achieved via the following equation

$$f^{final,<\#jets>}_{<flav>} = f^{<\#jets>}_{<flav>} \cdot \frac{f^{final,2excl}_{<flav>}}{f^{2excl}_{<flav>}} \qquad \text{and} \qquad K^{<\#jets>}_{<flav>} = \frac{f^{norm,<\#jets>}_{<flav>}}{f^{<\#jets>}_{<flav>}}. \tag{7.5}$$

The final set of heavy flavor scale factors are listed in section 7.3 with statistical uncertainties listed in 7.4. The correcting factors for the multijet background only used within the control region for the flavor scale factor estimation are presented in table 7.4.

channel	$\mathbf{K_{QCD}}$	**stat. uncertainties**
e	1.05	$\pm$ 0.03 (Data) $\pm$ 0.10 (MC W) $\pm$ 0.01 (MC asym)
μ	1.07	$\pm$ 0.06 (Data) $\pm$ 0.21 (MC W) $\pm$ 0.1 (MC asym)

Table 7.4: Correcting factors for the multijet background in the resolved control region (2 exclusive jets) for heavy flavor scale factor extraction, including statistical uncertainties only.

7.3 Final scale factors for W+jets and application

In this section the final normalization and heavy flavor scale factors are presented. To visualize the optimization concerning the agreement with data, the control regions for the heavy flavor scale factors are presented before and after the application of the scale factors, including the QCD correction factor for demonstration purpose.
For the boosted region in the $t\bar{t}$ resonance search analysis, see chapter 9, the flavor scale factors from the resolved region are used in combination with the charge asymmetry normalization estimated in the resolved region, since a lack of statistics prevents individual control regions for the boosted topology.

Two sets of heavy flavor scale factors are obtained for the final analysis, depending on the jet multiplicity: 4 exclusive jets and 5 inclusive jets, which cover the signal region with a 4 inclusive jet requirement. For reference the 2 exclusive jet bin results, the bin in which

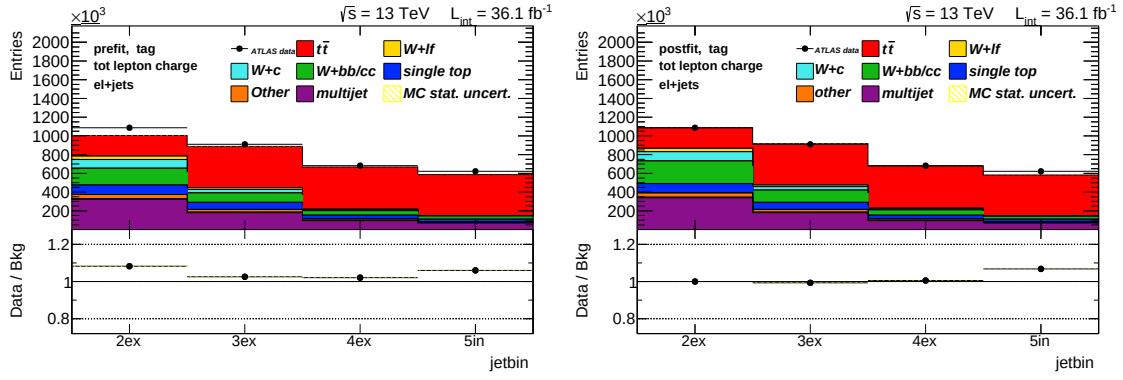
the flavor scale factors are estimated from, are listed additionally. The extrapolation from control to signal region is done via equation 7.5.

#jets	K_{bb}, K_{cc}			K_c			K_{lf}			CA		
2	1.28	±	0.18	1.00	±	0.30	0.93	±	0.05	1.06	±	0.06
4	1.22	±	0.17	0.96	±	0.29	0.89	±	0.04	1.02	±	0.05
≥5	1.20	±	0.17	0.94	±	0.28	0.87	±	0.04	0.78	±	0.04

Table 7.5: W+Jets scale factors for each W+jets flavor and the charge asymmetry normalization in different jet multiplicity bins for the e+jets channel in resolved topology. Combined statistical uncertainties are listed.

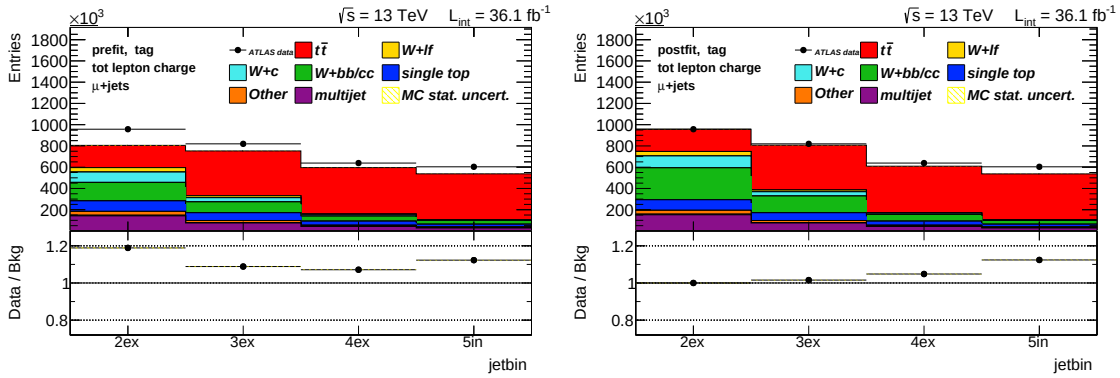
#jets	K_{bb}, K_{cc}			K_c			K_{lf}			CA		
2	1.51	±	0.17	1.00	±	0.30	0.88	±	0.04	1.15	±	0.05
4	1.41	±	0.16	0.93	±	0.28	0.82	±	0.04	0.95	±	0.03
≥5	1.35	±	0.15	0.89	±	0.27	0.79	±	0.03	0.80	±	0.03

Table 7.6: W+jets scale factors for each W+jets flavor and the charge asymmetry normalization in different jet multiplicity bins for the μ+jets channel in the resolved topology. Combined statistical uncertainties are listed.



(a) prefit **(b)** postfit

Figure 7.1: Jet multiplicity in the resolved electron channel, before application of the scale factors (a), and afterwards (b), including QCD correction factor in 2 exclusive bin only.



(a) prefit **(b)** postfit

Figure 7.2: Jet multiplicity in the resolved muon channel, before application of the scale factors (a), and afterwards (b), including QCD correction factor in 2 excl. bin only.

7.4 Uncertainties

Concerning the W+jets scale factors one can distinguish between two different kind of uncertainties:

- uncertainties affecting the estimation of the scale factors, resulting in uncertainties on the scale factors
- effect of other systematic uncertainties on the scale factors, resulting in respective scale factors

7.4.1 Uncertainties affecting the estimation of W+jets scale factors

Statistical uncertainties

In order to estimate the influence of the statistic of the collected data sample as well as the size of the Monte Carlos samples, 10.000 sets of scale factors are calculated with varied data yields in the pretag and tag control region under the assumption of a Poisson distribution. For the W+jets yields and the asymmetric Monte Carlo samples, the size of the variation is based on the square root of the quadratic sum of the event weights. The impact of the symmetric Monte Carlos samples is neglected due to its size being smaller than the one from the asymmetric Monte Carlo samples.

Method uncertainties

As an uncertainty of the method itself, the as symmetric defined samples have been checked thoroughly and it has been exposed, that their asymmetry value, see section 7.1, is varying within different jet multiplicity bins, see appendix B.

An impact on the charge asymmetry normalization could be traced back during the closure tests, see section 7.5, to the single top s- and t-channel as well as the Z+jets

sample, depending on the lepton channel. To allow for this, a set of scale factors has been estimated, where all MC samples are considered as asymmetric and extracted from data before the charge asymmetry normalization is calculated. This is treated as a separate uncertainty in the $t\bar{t}$ resonance search analysis, with a minor impact only.

For the W+jets with c flavor, the uncertainty can not be estimated, as the flavor scale factor is set to 1.0 in the control region. Therefore a conservative uncertainty of 30% has been chosen, based on the uncertainty on the cross-section measurement for $W + c$ with 2 additional jets.

7.4.2 Affect of other systematic uncertainties on the W+jets scale factors

For every other systematic uncertainty, listed in chapter 8, an individual set of scale factors is estimated and used for the overall uncertainty in the final result in the direct search.

7.5 Closure tests for W+jets estimation

The W+jets estimation has been tested with different pseudo data sets, consisting of the multijet estimate and the expected Monte Carlo background with varied fractions of the W+jets components $W + bb$, $W + cc$, $W + c$ and W+light flavor jets ($W + lf$).

The variations tested are based on the W+jets flavor and normalization scale factors estimated for the analysis:

- W+jets total yields + 10%
- W+bb, W+cc fraction + 40%
- W+bb, W+cc fraction - 40%
- W+lf fraction + 15%
- W+lf fraction - 15%

The $W + c$ scale factor is not included in the closure tests, since the $W + c$ fraction can not be estimated in parallel to the multijet correcting factor in the control region of exclusive 2 jets and without b-tag jet requirement.

In a so called Asimov data sample pseudo data equals the sum of Monte Carlo samples and the multijet background. The extracted W+jets scale factors differ from the expected scale factors, 1.00 for all flavors and CA normalization, see Table 7.7, but are within the statistical uncertainties[1]. This can be explained by the sensitivity of the normalization to the charge asymmetry in the Asimov data set. The contributions from $t\bar{t}$, Z+jets and single top Wt-channel are expected to be symmetric, but their asymmetries in Monte Carlo

1 For the analysis the statistical uncertainties of data and Monte Carlo W+jets and charge asymmetric Monte Carlo backgrounds are combined, with the uncertainties on W+jets dominating.

are partially above statistical fluctuations. For the analysis this is covered by the lepton charge mis-identification uncertainty in the electron channel and by introducing a new uncertainty, which corresponds to a CA normalization calculated with all Monte Carlo samples subtracted from the data sample.

2ex bin	$\mathbf{K}_{bb}$,$\mathbf{K}_{cc}$	$\mathbf{K}_{c}$	$\mathbf{K}_{lf}$	**QCD**	**CA**
expected	1.0	1.00	1.00	1.00	1.00
HFSF e-channel	0.96±0.06	1.00	1.01±0.01	0.99	1.03±0.01
HFSF μ-channel	0.95±0.05	1.00	1.01±0.01	0.97	1.04±0.01

Table 7.7: Charge asymmetry normalization CA and heavy flavor scale factors K_i, i $\in$ [bb,cc,c,lf] for an Asimov pseudo-dataset, scaled to the integrated luminosity of 36.1 fb^{-1}, statistical uncertainties on data only, expected and estimated heavy flavor scale factors with QCD scale factor and CA normalization in control region.

For the closure tests an ideal environment is defined, with symmetric yields for $t\bar{t}$, Z+jets, single top (Wt-channel) and the multijet background, resulting in a perfect agreement for the Asimov data set.

Within the closure test several variations, i.e. variations of the total W+jets yields and individual flavor fraction variations, have been checked and they all agree within expectations, se appendix C.

8 Systematic uncertainties

Within the analysis, different sources of uncertainty affect the significance of the measurement. In this section the uncertainties, that are considered as relevant for the search for a resonance within the $t\bar{t}$ mass spectrum[3], are detailed.
The personal contribution to these uncertainties regarding the W+jets estimation within the thesis at hand is described in 7.4.2. The dominant systematic uncertainties for the measurement of the top charge asymmetry are presented in section 10.7.
An overview over the final uncertainties for the direct search for new physics is presented combined with the final results in section 9.2.

At first the uncertainty on the luminosity and the pile-up is presented in section 8.1, followed by the uncertainties divided into affecting reconstructed objects, see section 8.2, and backgrounds in section 8.3.
For details and further information regarding the systematic uncertainties evaluated within the analysis at hand see [3, 110].

8.1 Systematic uncertainty on integrated luminosity and pile-up

The uncertainty in the combined integrated luminosity for data taken in 2015 and 2016 is 2.1 % [132], obtained using the LUCID-2 detector [61] for the primary luminosity measurements, see section 3.2.5. The uncertainty is applied as a constant shift to each simulated background.
In addition to that an uncertainty for the pile-up in collision events is estimated and applied by reweighting the MC samples in order to achieve the same luminosity profile as in data. The uncertainty is extracted from the official tool [133]. The luminosity induced uncertainty on the event yields in the $t\bar{t}$ resonance measurement varies between 1.9 and 2.1% in the different signal regions.

8.2 Systematic uncertainty on reconstructed objects

Uncertainties originating from the identification and reconstruction methods of physics objects affect the event yields and measured $t\bar{t}$ mass in the analysis at hand, since the reconstruction of the top quark pair from physics objects provides the input for the mass spectrum, i.e. the figure of merit for the search for resonances within the spectrum.

8.2.1 Electron and muon

For the electron and the muon the following sources for systematic uncertainties are considered:

- trigger efficiencies
- identification efficiencies
- energy scale
- energy resolution

The modeling of these sources is studied by using leptonic Z boson decays in data and simulation at $\sqrt{s}$=13 TeV. The identified corrections for the simulation, to better model the performance seen in data, have associated uncertainties that are propagated to the estimated signal and background yields. The combined systematic uncertainties associated to electron or muon are smaller than 2% in the event yields within the individual signal regions.

8.2.2 Small-R jets

Within the different uncertainties regarding small-R jets, the jet energy scale (JES) is estimated using a combination of simulations, test-beam data and in situ measurements. The jet energy resolution (JER) is derived with an in situ measurement of the jet response in di-jet events. With additional contributions from jet flavor composition, calorimeter response to different jet flavors, pile-up and more, a set of 19 eigenvectors is defined to describe the systematic uncertainty subcomponents [134]. The JES and JER uncertainties for small-R jets are one of the dominant uncertainties within the presented direct search for new physics within the $t\bar{t}$ mass spectrum, with the largest uncertainty from JES with 5.8% in the resolved channel, see figure 9.2.

8.2.3 Large-R jets

The systematic uncertainties on the large-R jet energy, mass and τ_{32} scales are derived from Run-1 in situ uncertainties for $\sqrt{s}$= 8 TeV, extrapolated to $\sqrt{s}$=13 Tev and complemented by additional Run-2 uncertainties[89]. The method used for estimating the uncertainties in Run-1 is the calorimeter-to-tracker response double-ratio method, or R_{trk} method, see [135] for details.
The considered uncertainties also have a minor impact on the resolved selection, since the uncertainty on the large-R jet can have an impact on the decision, whether the event passes the boosted or the resolved selection. JES and JER are the dominant large-R jet uncertainties with a significant impact on the final yields of up to 6.0% for the JES uncertainty in the boosted channel, see figure 9.2.

8.2.4 b-tagging on the track jets

To compensate for differences between data and simulation within the b-tagging efficiency for b-, c- and light jets [88, 136] correction factors are applied. The correction factors are derived from flavor enriched data samples, i.e. $t\bar{t}$ events with di-leptonic final states. An additional uncertainty term is considered to extrapolate the measured uncertainty for various variables affecting the b-tagging performance, such as track multiplicity per jet

and impact parameter resolution, to the high p_T regions of interest. Within the b-tagging uncertainties and overall, the b-tagging efficiency and extrapolation uncertainty are within the dominant uncertainties regarding the analysis with up to 3.2% for the b-tagging efficiency in the resolved channel, see figure 9.2.

8.2.5 Missing transverse momentum

For the uncertainty on the E_T^{miss} two different sources are considered: variations due to the systematic uncertainty of the E_T^{miss} inputs, such as small-R jets or leptons, and the uncertainties assigned to the soft term of the E_T^{miss} calculation, see section 4.2.4, recommended by the JetEtMiss group[137]. The E_T^{miss} uncertainty is of minor relevance for the analysis compared to the dominant systematics, since it is below 2%.

8.3 Systematic uncertainties on the background

Uncertainties can be specific to individual processes, such as $t\bar{t}$ or W+jets. The following subsections address these subcategory of uncertainties.

8.3.1 Uncertainties affecting only $t\bar{t}$

Since the search for new physics takes place within the mass spectrum of $t\bar{t}$ events, a deep understanding of the uncertainties in the $t\bar{t}$ process is of utter importance.

Cross-section

One sources of uncertainty for $t\bar{t}$ processes is the uncertainty on the cross-section. At a center-of-mass energy of $\sqrt{s}$= 13 TeV the cross section is estimated as $\sigma_{t\bar{t}}$=832^{+46}_{-52} pb for a top mass of 172.5 GeV. It has been calculated at next-to-next-to-leading order (NNLO) in QCD including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms with top++2.0 [22, 138–143].

Modeling

Since different $t\bar{t}$ models are available, distinguished by the choice of the individual generators for simulating hard scattering and parton shower as well as the degree of radiation, uncertainties are estimated to cover those variations. To account for those variations the nominal Powheg+Pythia $t\bar{t}$ sample are compared to Powheg+Herwig to estimate the effect of different parton showering generators, while MC@NLO+Herwig and Powheg+Herwig are considered for an estimation of the hard scattering process simulation. For the radiation uncertainty two additional Powheg+Pythia samples with increased (decreased) radiation are compared to the nominal $t\bar{t}$ sample, see section 5.2.2 for a short sample description.

Parton distribution functions (PDF)

The choice of the parton distribution functions is evaluated using the PDF4LHC15 PDF prescription, providing 30 separate uncertainty eigenvectors[115], with the MSTW2008 68% CL NNLO [144, 145], CT10 NNLO [101, 102] and NNPDF2.3 5f FFN [112] PDF prescriptions, added in quadrature to the scale uncertainty. This uncertainty is taken as a systematic uncertainty in the nominal Powheg+Pythia 6 $t\bar{t}$ sample.

Electroweak corrections

Electroweak corrections to $t\bar{t}$ production add an uncertainty of 10% of their deviation from unity[3]. The corrections for higher order electroweak effects are applied to the nominal $t\bar{t}$ sample as a weight. The weight depends on the flavor and energy of the initial partons in the center of mass frame, and on the angle of the top quark in the same frame. The value of the correction decreases with the invariant mass of the $t\bar{t}$ system.

Next-to-next-to-leading order QCD corrections

NNLO QCD uncertainties are derived from theoretical calculations [146] and applied as a function of the top quark p_T and the transverse momentum of the $t\bar{t}$ system, as recommended in reference [146]. The effect of this uncertainty in the $m_{t\bar{t}}$ mass spectrum is very small in the low mass region, but increases up to 7% for a mass of 2 TeV and higher in the resolved and up to 20% above 3 TeV in the boosted selection.

8.3.2 Uncertainties affecting only single top

The cross-section uncertainty for the single top background is estimated to be $\pm 5.3\%$, corresponding to the theoretical uncertainty in the dominant Wt-channel contribution at NNLO in QCD [147–149]. Differences between the predictions from diagram removal and diagram subtraction approaches [119] are represented by an according uncertainty, since both methods lead to different results with respect to the interference between tW and $t\bar{t}$ production.

8.3.3 Uncertainties affecting only W+jets

The uncertainties for W+jets are described within the estimation procedure in section 7.4.

8.3.4 Uncertainties affecting only QCD multijet

Various definitions of multijet control regions are utilized to estimate the systematic uncertainty for the multijet background, resulting in slightly different values for the efficiency rate f. Including systematic uncertainties associated with object reconstruction and MC simulation, a total normalization uncertainty of 50% is assigned to the QCD multijet background, lepton flavor independent.

9 Search for resonances in $t\bar{t}$ decay mass spectrum - a direct search for BSM physics

A search for new heavy particles decaying into top quark pairs has been performed using 36.1 fb^{-1} of data at $\sqrt{s}$= 13 TeV by the ATLAS detector at the LHC. Since no excesses have been found in the $t\bar{t}$ mass spectrum beyond background expectation exclusion limits on cross-section times branching ratio for different signal models have been set. The tested signal models are the hypothetical Z' boson, Kaluza-Klein gluons (g_{KK}) and Kaluza Klein gravitons (G_{KK}).
This chapter presents the results of the publication [3] of the $t\bar{t}$ resonance search group, while the providing of the W+jets estimate with heavy flavor scale factors, explained in detail in chapter 7, represents the main contribution to the group effort.

With the analysis procedure described in chapter 6, the following section 9.1 focuses on the adaption regarding the event selection and the acceptance of the tested models, followed by the considered systematic uncertainties in section 9.2, the results in section 9.3 and a short discussion in section 9.4.

9.1 Event categorization and acceptance

To improve the sensitivity, see [110], of the $t\bar{t}$ resonance search analysis with respect to the limit results, a splitting into four different b-tag categories is introduced after the event selection presented in chapter 6. For the categorization, the reconstructed hadronic and leptonic top, as described in section 6.2, as well as the b-tagged jet(s) are considered.
The events are sorted into one of the following four b-tag categories:

Category 0: ∄ b-tagged jet matching the hadronic or leptonic top candidate

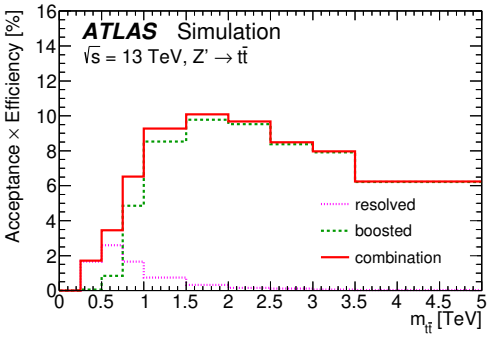
Category 1: only the leptonic top candidate has a matching b-tagged jet

Category 2: only the hadronic top candidate has a matching b-tagged jet

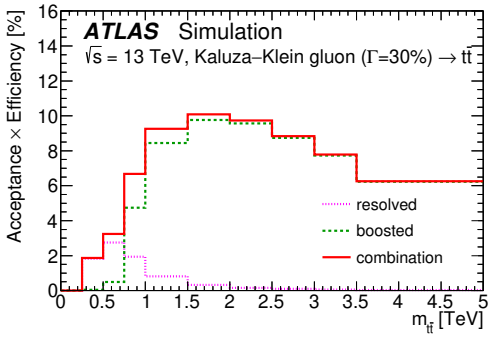
Category 3: both top candidate have a matching b-tagged jet

A b-tagged jet is considered as a match, if it is reconstructed within $\Delta R = 0.4$ to the small-R jet assigned to the leptonic top candidate or within $\Delta R = 1.0$ of the large-R jet of the hadronic top candidate. The resolved equivalent of the large-R jet is the one small-R jet assigned as a b quark by the χ^2 algorithm, described in section 6.2.4. All events within category 0 are rejected for the signal regions.

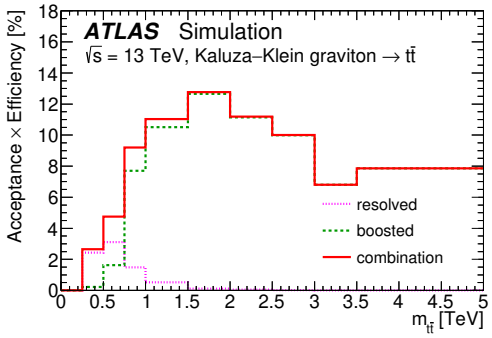
For these final signal regions the acceptance[1] times efficiency[2] ($A \times \varepsilon$) is shown in figure 9.1, including the branching ratio for simulated BSM particles decaying into $t\bar{t}$, described in section 2.3, in the detector acceptance, since other decay channels than $t\bar{t}$ are possible. For the separated acceptances in the individual lepton channels and b-tag category split into 1,2 and 3 see [110].



(a) Z'→$t\bar{t}$



(b) g_{KK} →$t\bar{t}$



(c) G_{KK} →$t\bar{t}$

Figure 9.1: Acceptance times efficiency ($A \times \varepsilon$), including the branching ratio for MC simulated BSM particles decaying into $t\bar{t}$, as a function of the $t\bar{t}$ invariant mass $m_{t\bar{t}}$ for simulated signal events. The e and μ channel efficiencies are combined[3]

9.2 Resulting systematic uncertainties

For the analysis the effect of the systematic uncertainties described in 8 is estimated and the largest contributions are shown in figure 9.2. The $t\bar{t}$ modeling and jet energy uncertainty provide the largest contribution to the overall uncertainties.

1 Acceptance: The fraction of the signal events, that is visible to and recorded by the detector systems.
2 Efficiency: The fraction of signal events, that pass the selection.

Systematic Uncertainty	Background [%]		Z'_{TC2}, 2 TeV [%]		Z'_{TC2}, 3 TeV [%]	
	resolved	boosted	resolved	boosted	resolved	boosted
$t\bar{t}$ extra QCD radiation	4.0	2.4	–	–	–	–
$t\bar{t}$ QCD NNLO	0.8	7.4	–	–	–	–
$t\bar{t}$ cross-section	5.2	–	–	–	–	–
$t\bar{t}$ generator	1.7	3.8	–	–	–	–
$t\bar{t}$ parton shower	0.6	3.2	–	–	–	–
Multi-jet	2.6	2.7	–	–	–	–
Anti-k_t $R = 0.4$ JER	1.1	0.2	3.2	0.2	1.2	0.2
Anti-k_t $R = 0.4$ JES	5.8	0.9	7.0	0.7	3.6	0.6
Anti-k_t $R = 1.0$ JER	0.1	4.0	5.3	3.7	2.0	4.2
Anti-k_t $R = 1.0$ JES	0.3	6.0	3.7	4.7	2.8	6.0
b-tagging efficiency	3.2	1.8	1.8	1.9	2.3	2.7
b-tagging extrapolation	2.4	2.3	2.0	0.6	1.2	1.8
Luminosity	1.9	1.9	2.1	2.1	2.1	2.1
Pile-up	4.4	0.5	4.4	0.8	3.9	0.5
Total	11.6	12.8	11.7	7.1	7.6	8.7

Figure 9.2: The systematic uncertainties in the yields in the background, as well as in the 2 TeV and 3 TeV Z'$_{TC2}$ signal models, in percentage. Only rows with at least one column with an uncertainty larger than 2% are shown individually.

9.3 Results

After the reconstruction of the $t\bar{t}$ mass spectra, a model independent search for excesses in form of bumps or deficits is performed, described in section 9.3.1. Since no deviations beyond expected fluctuations were found, model dependent upper limits have been set in section 9.3.2 for each of the tested BSM models.

9.3.1 Compatibility with the SM-only hypothesis

The here presented search for new particles is performed using the $t\bar{t}$ mass spectra distributions of data and expected SM background with BumpHunter[150], a hypothesis testing tool sensitive to local excesses of data. This tool also takes the look-elsewhere effect (LEE)[151] over the full mass range into account, requiring a significant bump at roughly the same mass in all spectra, to avoid a bias by random fluctuation. The detailed SM-only hypothesis test procedure performed by BumpHunter is described in the following section, based on [110].

Common to all kind of hypothesis tests is the consistency check between some data D and a hypothesis, typically "Null" or in this case "SM-only", represented by the MC background following SM predictions. Three steps summarize the estimation of potential excesses or deficits in data:

1. quantification of the discrepancy between data and background by a single number, the test statistic t
2. generation of pseudo-data based on the SM-only hypothesis and repetition of step 1 between pseudo-data and background

3. consistency check of data and SM-only hypothesis with the p-value

The quantification of the discrepancy between data and background in a $t\bar{t}$ mass spectrum is performed in sliding windows of variable size, with a minimum width of two bins, through the full mass spectrum in boosted (480 GeV to 6 TeV) and resolved (390 GeV to 2 TeV) topology. In order to estimate the test statistic t for each window the Poisson probability $P(d_i,b_i)$, defined in equation 9.1 with data yields d_i and background yields b_i, is estimated while searching for an excess[1].

$$P(d_i,b_i) = \begin{cases} \Gamma(d_i,b_i) = \sum_{n=d_i}^{\infty} \frac{b_i^n}{n!} e^{-b_i} & \text{if } d_i \geq b_i \\ 1 - \Gamma(d_i+1,b_i) & \text{if } d_i < b_i \end{cases} \tag{9.1}$$

The smallest $P(d_i,b_i)$ value of all windows within one mass spectrum, defined as P_i^{min}, hereby represents the window containing the most interesting candidate for an excess. Including the total systematic uncertainty θ_i for window i, equation 9.1 is redefined as 9.2 with $\lambda_i \in \{-8,8\}$ being the real number that maximizes $\frac{b_i^n}{n!} e^{-b_i} e^{-\frac{\lambda_i^2}{2}}$.

$$P(d_i,b_i) \rightarrow (d_i,b_i + \lambda\theta_i) e^{-\frac{\lambda_i^2}{2}} \tag{9.2}$$

From P_i^{min} the BumpHunter test statistic t can be calculated as

$$t = \begin{cases} 0 & \text{if} d_i \leq b_i \\ -\log P_i^{min} & \text{otherwise} \end{cases} \tag{9.3}$$

After the quantification of the differences in the mass spectra of data and SM background expectation, at least $N = 10.000$ pseudo experiments have been generated by Poisson fluctuation of the expected background.
The figure of merit to decide, whether the "SM-only" hypothesis can be rejected, is the p-value, defined in equation 9.4, where $f(t)$ is the distribution of the test statistic values derived from the pseudo-data and t_{obs} is the test statistic obtained from data.

$$p - \text{value} = \frac{\int_{t_{obs}}^{\infty} f(t)}{\int_0^{\infty} f(t)} \tag{9.4}$$

The p-value lies within 0 and 1 and can be considered a false-discovery probability, with a value of 0 representing a deviation in data of a larger size than any corresponding excess in the pseudo-data.
All 12 distinct channels, divided by lepton flavor (e,μ), topology (resolved, boosted) and the three b-tagging categories (1,2,3), have been evaluated with BumpHunter for excesses and deficits in the $t\bar{t}$ mass spectrum. Without finding any significant deviations the "SM-only"

1 For the search for deficits the inequality signs are reversed

hypothesis could not be rejected. Therefore no hints for the existence of a BSM particle decaying into a top-antitop quark pair could be identified.

9.3.2 Upper limit setting on cross-sections

The search for excesses or deficits in the $t\bar{t}$ mass spectra was accomplished in a model-independent way. In order to exclude mass ranges for hypothetical particles benchmark models of specific processes need to be defined. Therefore upper limits are set on the cross-section times branching ratio for each tested signal model, defined in section 2.3, using a profile likelihood-ratio test. Within this test confidences are derived from the pdfs of $-2\ln Q$, with Q representing the ratio of likelihoods for the two hypothesis of interest, see equation 9.5, i.e. b for background (SM-only) and $s+b$ for signal+background.

$$Q = \frac{L(s+b)}{L(b)} \tag{9.5}$$

A frequentist approach with the CL_s prescription[152] is used to derive one-sided 95% confidence level (CL) cross-section limits on the hypothesis, that data agrees with the expected SM background and signal samples. The CL_s procedure provides conservative limits on the signal hypothesis and is described by

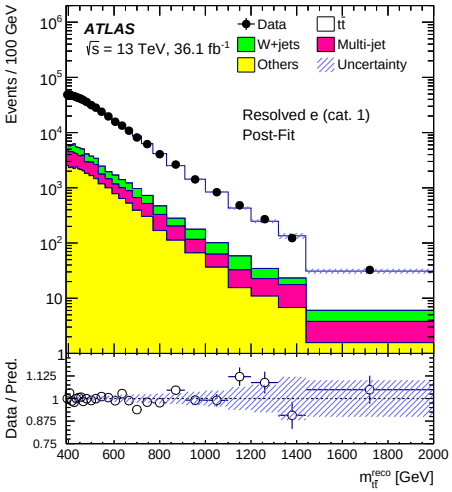
$$CL_s = \frac{CL_{s+b}}{CL_b} \tag{9.6}$$

Statistical and systematic uncertainties in the expected distributions are included as nuisance parameters, constrained by Gaussian pdfs in the likelihood fits. For details see [3]. Exemplary the mass distributions after the likelihood fit under the background-only hypothesis for the resolved e+jets and the boosted μ+jets channel are displayed in figure 9.4.

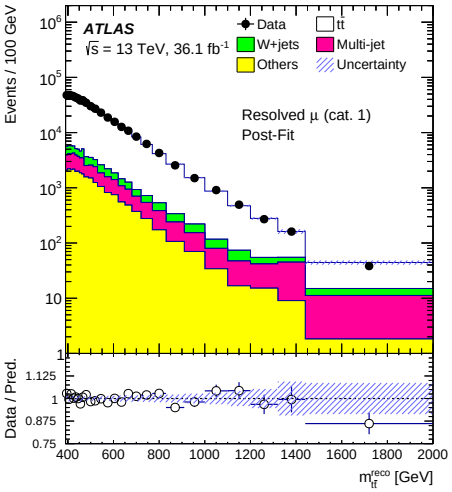
The expected and observed limits on the cross-section for the different tested BSM models versus the particle mass are shown in figure 9.5, 9.6, 9.7 and 9.8. In these figures the regions between the extracted limits for the individual mass points are interpolated with straight lines. The summary of the excluded mass limits for the BSM particles can be found in figure 9.3.

Summary of 95 % Confidence Level mass exclusion ranges on benchmark models		
Model	Observed excluded mass [TeV]	Expected excluded mass [TeV]
Z'_{TC2} (1% width)	< 3.0	< 2.6
$Z'_{\mathrm{DM,ax}}$	< 1.2	< 1.4
$Z'_{\mathrm{DM,vec}}$	< 1.4	< 1.6
G_{KK}	[0.45, 0.65]	[0.45, 0.65]
g_{KK} (15% width)	< 3.8	< 3.5
g_{KK} (30% width)	< 3.7	< 3.2

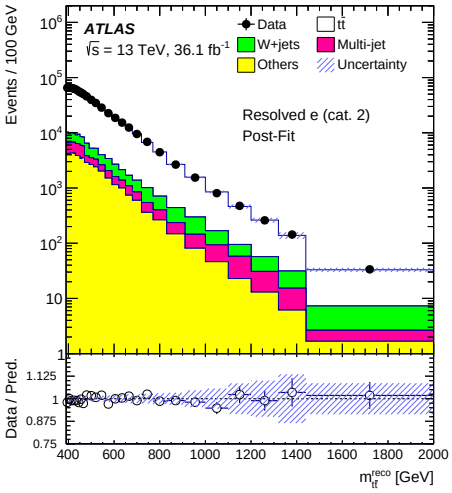
Figure 9.3: Summary of excluded mass ranges for the signals studied in the analysis[3].



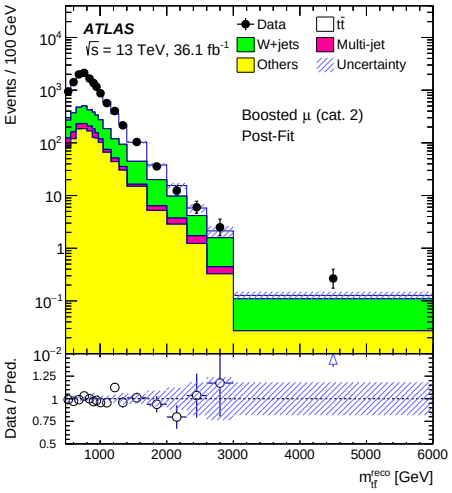
(a) resolved e+jets, cat. 1



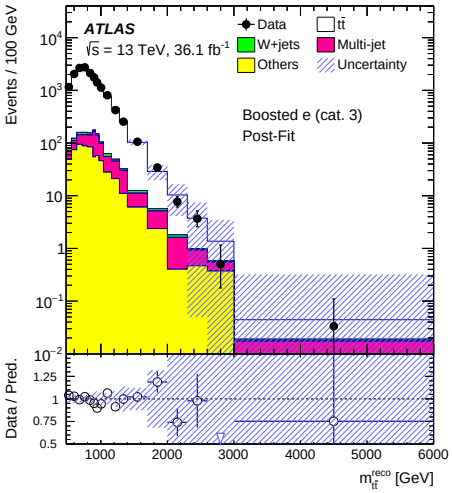
(b) resolved e+jets, cat. 2



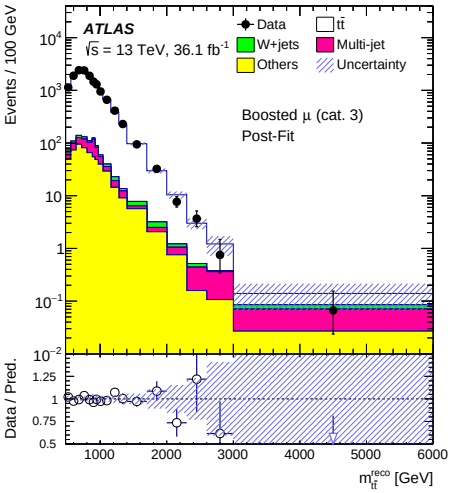
(c) resolved e+jets, cat. 3



(d) boosted μ+jets, cat. 1



(e) boosted μ+jets, cat. 2



(f) boosted μ+jets, cat. 3

Figure 9.4: The $m_{t\bar{t}}^{reco}$ distributions for the resolved e, (a),(b),(c), and boosted μ, (d),(e), (f). The shaded areas indicate the total systematic uncertainties, The ratio of the data to the final fitted expectations is shown in the lower panel, open triangles indicate that the ratio point would appear outside the panel[3].

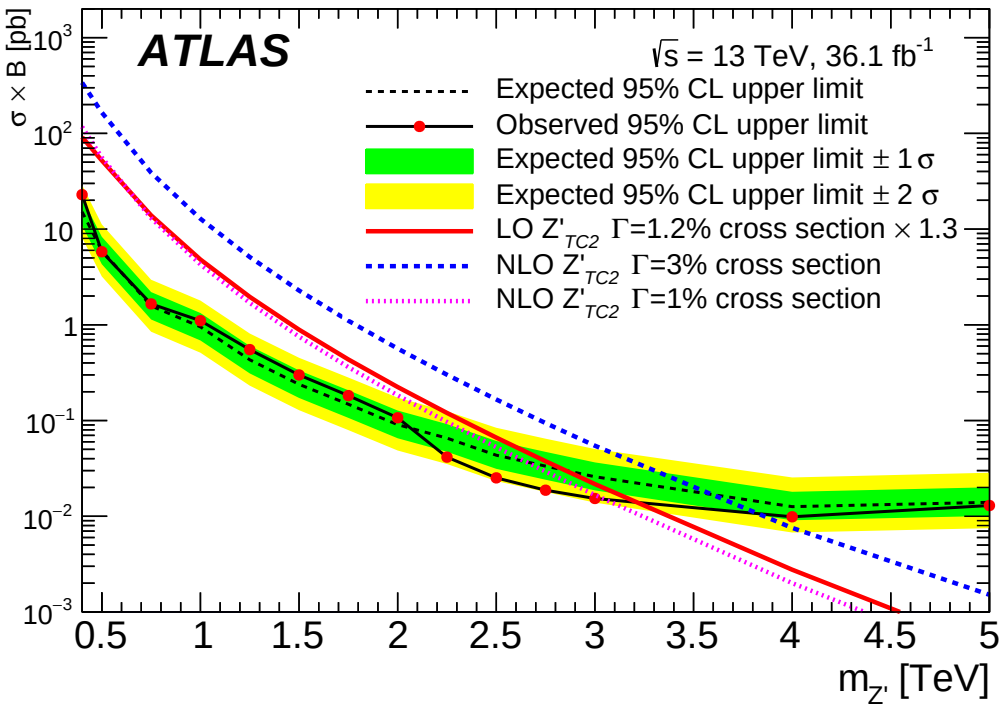
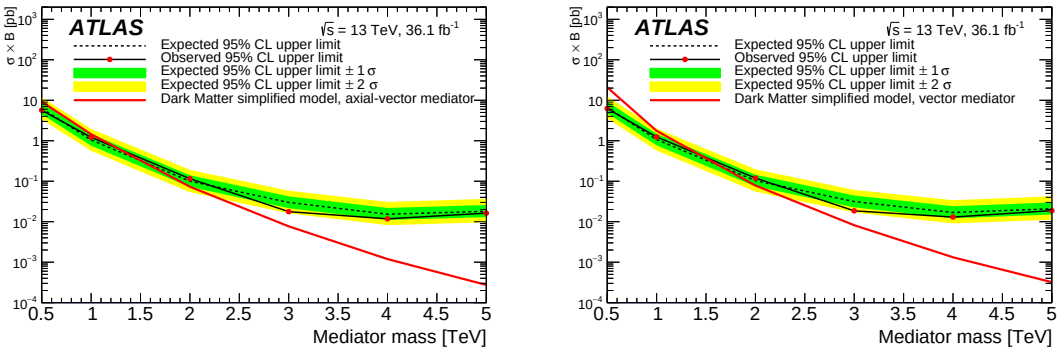


Figure 9.5: The observed and expected cross-section 95% CL upper limits on the Z'$_{TC2}$ signal. The theoretical predictions for the production cross-section times branching ratio of Z'$_{TC2}$ $\rightarrow t\bar{t}$ at the corresponding masses are also shown[3].



(a) Axial-vector mediator **(b)** Vector mediator

Figure 9.6: The observed and expected cross-section 95% CL upper limits on the (a) Z'$_{DM,ax}$ and (b) Z'$_{DM,vec}$ signals. The theoretical predictions for the cross-section times branching ratio of Z'$_{DM}$ $\rightarrow t\bar{t}$ at the corresponding masses are also shown[3].

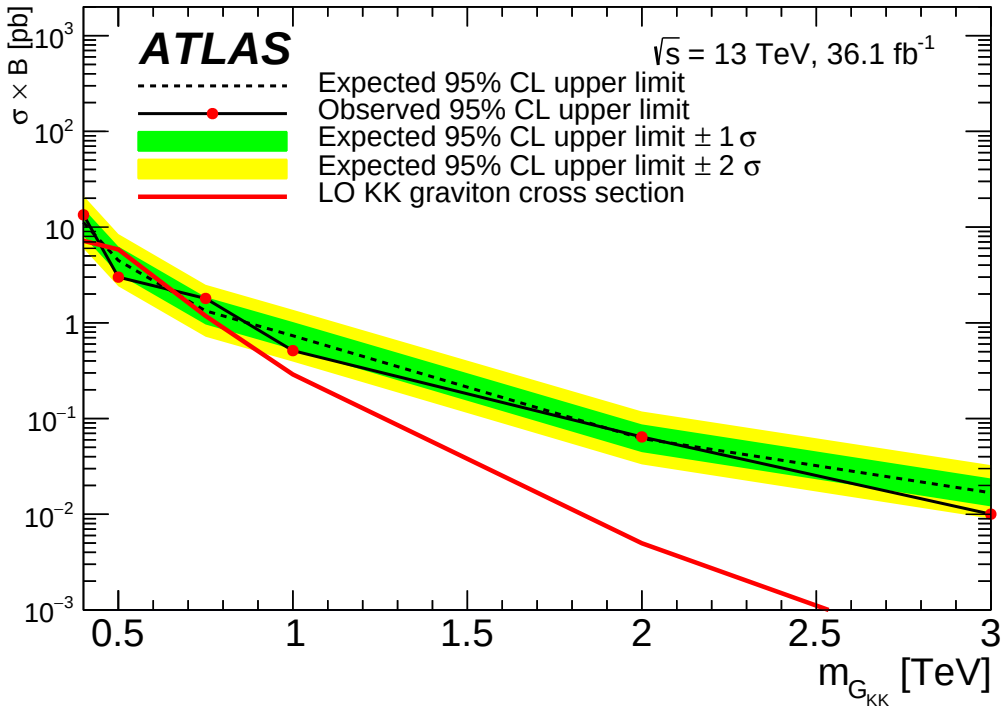
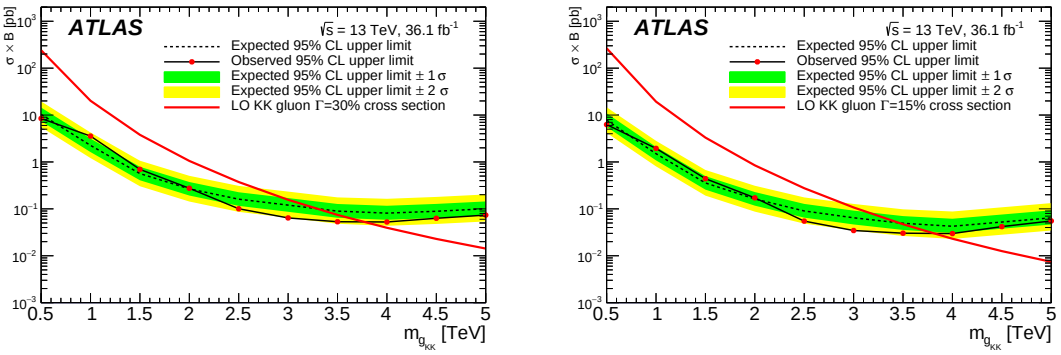


Figure 9.7: The observed and expected cross-section 95% CL upper limits on the G_{KK} signal. The theoretical predictions for the production cross-section times branching ratio of $G_{KK} \rightarrow t\bar{t}$ at the corresponding masses are also shown[3].



(a) 30% width **(b)** 15% width

Figure 9.8: The observed and expected cross-section 95% CL upper limits on the g_{KK} signal for resonances with widths of 30% (a) and 15% (b). The theoretical predictions for the cross-section times branching ratio of $g_{KK} \rightarrow t\bar{t}$ at the corresponding masses are also shown[3].

9.4 Discussion

Since no evidence has been observed for the existence of new particles decaying into top quarks for the here considered models for Z', g_{KK} and G_{KK} at masses up to 3.8 TeV, the frontier needs to be pushed further. This can be achieved on one hand by decreasing systematic uncertainties affecting only the SM background, especially $t\bar{t}$ modeling and multi-jet estimation. On the other hand the impact on the overall uncertainty for signal and background is dominated by JES and JER uncertainties as well as b-tagging related uncertainties and the effect of uncertainties by measuring the total luminosity and estimating the pile-up. Since the analysis is not in particular sensitive to statistics, using the full 2017-2018 dataset is not expected to improve the analysis in a significant way. Contrary the increase in the center of mass energy would gain statistics in the higher $t\bar{t}$ mass region and extend the range for the search.
A representative summary over the current effort regarding the search for new physics is provided in figure 9.9, containing the limits estimated in the here presented analysis [3]. Supersymmetry for example is not included in this summary, since the focus lies on comparable models to the considered examples in the $t\bar{t}$ resonance search.

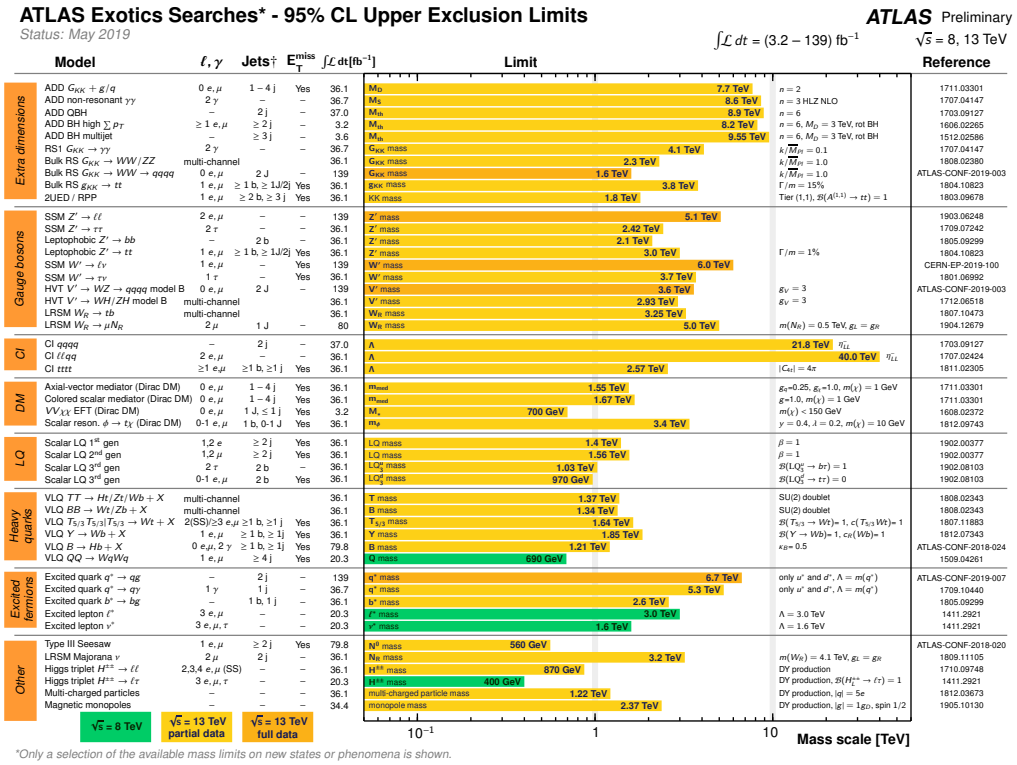


Figure 9.9: Reach of ATLAS searches for new phenomena. Only a representative selection is shown. Green bands indicate 8 TeV data results; yellow (orange) bands indicate 13 TeV data results with partial (full) dataset [153].

10 Measurement of charge asymmetry in top pair production - an indirect search for BSM physics

In the previous chapter, chapter 9, the direct search for physics beyond the standard model in form of new particles decaying into top quark pairs has been presented, resulting in new exclusion limits. Within the following section a different approach to search for new physics, representing a so called indirect search, is performed.
Measuring the properties of SM particles with increasing precision provides not only a validation of the modeled processes, but is also sensitive to new particles. In case of hidden new particles measurable interferences are expected resulting in derivations from the SM prediction, hence this approach is considered as 'indirect' search.

The indirect search performed within this thesis is the precision measurement of the charge asymmetry, A_C, within the pair production of top quarks in the lepton+jets channel. The corresponding theoretical background is outlined in chapter 10.1. The challenge of this analysis is the size of the dominant uncertainties. The main source of uncertainty, the $t\bar{t}$ modeling uncertainty described in section 5.2.2, is sensitive to the statistic of the modeled events and restricted to the available samples. Within the background uncertainties the focus here is on the W+jets estimation. Since a complete rejection of the W+jets background is impossible, due to the conformity with the $t\bar{t}$ decay signal, a realistic modeling of the estimated W+jets events is necessary to decrease the impact of the asymmetric production of W^+ and W^- on the top charge asymmetry measurement as good as possible. As standard procedure for the W+jets background estimation an in-situ approach is considered, where the amount of simulated W+jets events is adapted during the measurement in the signal region. In contrast to that a second approach is evaluated for potential improvements.
Physically and technically the analysis at hand follows the recommendations and experience of the 8 TeV top charge asymmetry analysis, see [154], wheres the event selection is based on the recommendations from the $t\bar{t}$ resonance search group and the 13 TeV publication [3], presented in chapter 9. Under these circumstances it is possible to reutilize and investigate the potential of the already obtained heavy flavor scaling factors, estimated for the $t\bar{t}$ resonance publication, see chapter 7.

Evaluating these two different W+jets estimation approaches regarding the potential to decrease the corresponding uncertainty is the main purpose of this complementary study. During the preparation and estimation of the dominant uncertainties, a closer look has been applied to the largest impact on the total uncertainties for now, the $t\bar{t}$ modeling uncertainties.
The analysis presented here has provided input to the public results in [4], but is based on a different set of data and simulation samples, hence yielding results not directly comparable

with the public ones.

An outlook of the here presented analysis with respect to the published top charge asymmetry measurement in [4] with the full 13 TeV dataset is provided in section 10.10.

In this chapter advantage of the full 2015 and 2016 statistic is taken within the inclusive measurement of the top charge asymmetry for resolved topologies. A differential approach regarding the $t\bar{t}$ mass is presented, motivated by the sensitivity of the A_c measurement to various theoretical models, see section 10.1.

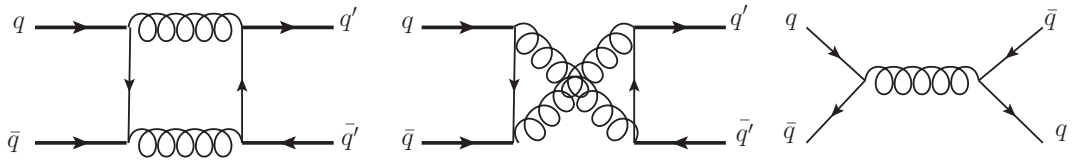
The first section 10.1 is dedicated to the theoretical introduction into the phenomenology of the charge asymmetry, followed by a description of the analysis setup in section 10.2. The introduction into the analysis methods contains a short description of the Bayesian Unfolding techniques and the configuration thereof for the measurement in 10.3. The systematic uncertainty of the unfolding procedure is evaluated in the following section 10.4 regarding its linearity between reconstructed A_c values from artificially asymmetric $t\bar{t}$ pseudo-datasets with the intrinsic truth values. The general configuration is followed by a validation of the method in section 10.5. The comparison of the two approaches for assessing the W+jets background in discussed in section 10.6. In section 10.7 the estimation of the considered dominant uncertainties is described. The analysis results for the inclusive charge asymmetry measurement are presented in section 10.8, the results of the differential measurement in section 10.9. A discussion of potential improvements of the presented analysis with respect to the publication with the full dataset and an updated set of Monte Carlo samples in section 10.10 finalizes this chapter.

10.1 Theoretical description of the top quark pair decay charge asymmetry

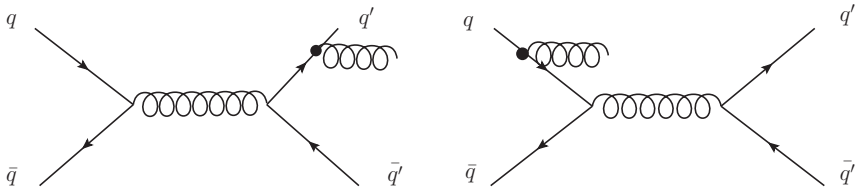
The production process of $t\bar{t}$ is described within the perturbative quantum field theory (QFT) in the SM. Within the theory introduction in section 2.2.1 only the production process in leading order Born approximation for gluon fusion and quark-antiquark annihilation is described, which is symmetric under charge conjugation. For next-to-leading order an additional production mechanism is considered, the quark-gluon scattering $qg \to t\bar{t}g$, also known as flavor excitation.

The origin of the asymmetry in top and antitop production is based on two different reactions [155], i.e. radiative corrections to $q\bar{q}$ fusion, see figure 10.1, and interference of higher order amplitudes affecting qg-scattering, see figure 10.2. Gluon fusion does not provide a contribution to the charge asymmetry, since it is symmetric for all orders. The relative QED contribution to the charge asymmetry is about 13%, while weak interactions are negligible [156].

The interference terms displayed as Feynman diagrams in figure 10.1(a) correspond to virtual radiation, the complement Feynman diagrams in figure 10.1(b) to real emission. They contribute with opposite signs, with the virtual radiation contribution always larger than the real emission contribution, see [157].



(a) Interference between box/crossed box diagram and LO Born diagram



(b) Interference of final state and initial state with gluon bremsstrahlung

Figure 10.1: Visualization of NLO $q\bar{q}$ annihilation, providing a contribution to charge asymmetry in $t\bar{t}$ production

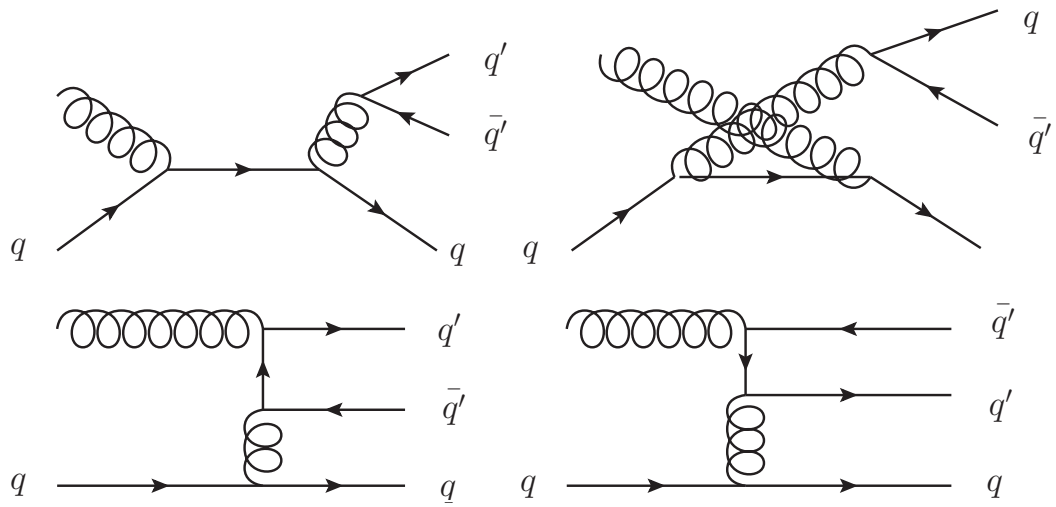
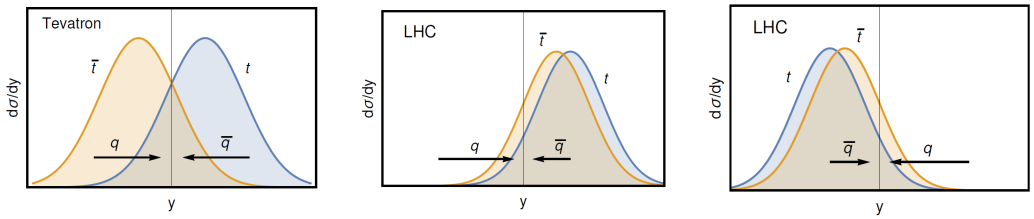


Figure 10.2: Asymmetry in $t\bar{t}$ production due to quark-gluon-scattering (NLO)

The charge asymmetry expresses itself in the tendency of the top quark to be more likely produced in the direction of the incoming quark in a $q\bar{q}$ production process and respectively the antitop in the direction of the antiquark. This results in a $p\bar{p}$ collider like the Tevatron[1] in a forward-backward asymmetry, since the direction of quark and antiquark are defined most likely by the direction of the proton and antiproton beam, see illustration 10.3(a). Experiments conducted with the ATLAS detector at the proton-proton

1 Tevatron was a particle accelerator used for proton-antiproton collisions running from 1983 to 2011 at Fermi National Accelerator Laboratory, Chicago.

collider LHC, such as the work at hand, do not result in an observable forward-backward asymmetry, due to the symmetric pp initial state, but in a forward-central asymmetry. With the proton consisting of valence quarks (uud) with relative large proton momentum x compared to sea-quarks and gluons, the $q\bar{q}$ production process of the $t\bar{t}$ pair is dominated by initial quarks with a large x and antiquarks with small x. Thus we expect a broader rapidity distribution for the top quark than for the antitop quark, since the boost in the laboratory frame increases the probability of the top quark to be produced in the forward and backward direction, while the antitop quark is produced more likely in the central region, see illustration 10.3(b) and 10.3(c). As a conclusion the charge asymmetry leads to a sizable difference in the differential production cross sections of top and antitop quarks, first predicted in [157]. The thesis at hand focuses only on the forward-central asymmetry at the LHC.



(a) $p\bar{p}$ collider **(b)** pp collider I **(c)** pp collider II

Figure 10.3: Not to scale partonic rapidity distributions of top and antitop quark at $p\bar{p}$ colliders like Tevatron (a) and $p\bar{p}$ colliders like ATLAS (b),(c) to illustrate the measurable charge asymmetry within top-antitop production. Sketches are taken from [158]

Considering BSM physics new particles, e.g. axigluons[159], heavy Z'[160] or colored Kaluza-Klein gluon excitations [161] are expected to interfere with the SM process due to additional amplitudes, resulting in a modified asymmetry.

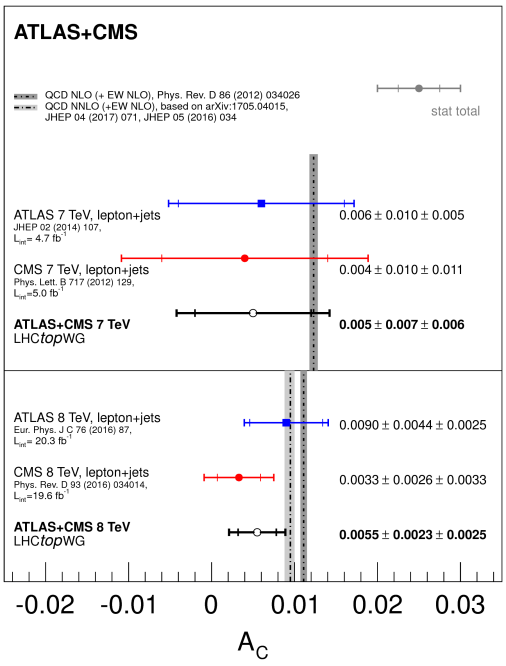
The charge asymmetry of $t\bar{t}$ pair decays at a proton-proton collider can be quantitatively expressed by A_C through the rapidity of the top and antitop quark as

$$A_C^{t\bar{t}}(y) = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)} \qquad \text{with} \quad \Delta|y| = |y_t| - |y_{\bar{t}}|. \tag{10.1}$$

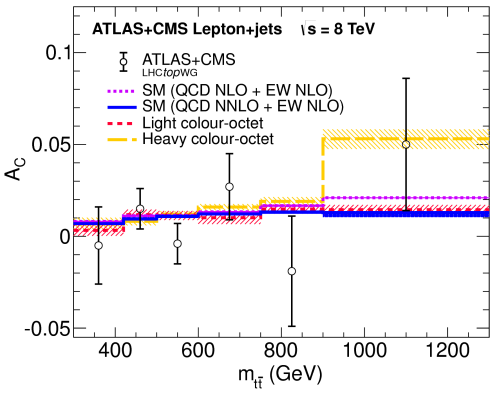
Previous measurements with the full statistics of data at 7 and 8 TeV at ATLAS and CMS have been performed of the top-antitop charge asymmetry in the lepton+jets channel individually [162–165] and combined [166], both inclusive and differential, see summary of results in figure 10.4. Additionally at the ATLAS detector the asymmetry in highly boosted top-quark pairs at $\sqrt{s} = 8$ TeV has been measured and published in [167]. Recently ATLAS presented new measurements for the lepton+jets channel for the full statistics of Run II data at 13 TeV [4].

All of the performed measurements by ATLAS and CMS are consistent with SM predictions. Since the impact of the studied BSM scenarios increases with the center of mass energy, while within higher mass regions the significance of the measurement is limited

by uncertainties, measurements with high statistics reaching for higher mass regions at the order of 1 TeV and beyond are necessary to break down the borders of our current understanding of physics.



(a) Summary of single inclusive measurements and the LHC combinations at 7 and 8 TeV compared to theory predictions (calculated at (N)NLO in QCD, incl. NLO electroweak corrections). The inner bars indicate the statistical uncertainty, the outer bars the total uncertainty.



(b) The combined ATLAS+CMS charge asymmetry in six bins of the invariant mass of the $t\bar{t}$ system in comparison with theoretical predictions for the SM and two versions of a colour-octet model. The last bin includes the overflow. The uncertainties, indicated by the shaded areas, reported for the SM predictions are dominated by scale variations and are small, while the uncertainties reported for the colour-octet model are statistical uncertainties in the simulation.

Figure 10.4: Combination of inclusive and differential $t\bar{t}$ charge asymmetry measurements using ATLAS and CMS data at $\sqrt{s}$=7 TeV and 8 TeV [166]

10.2 Analysis description for top charge asymmetry measurement

The motivation of the following study is the question, whether the top charge asymmetry measurement gains an advantage by using the elaborate estimation of the W+jets background with the heavy flavor scale factors in an external control regions. Since the scale factors have already been provided for the $t\bar{t}$ resonance analysis, see chapter 9, a charge asymmetry estimation with similar analysis conditions has been set up.
In chapter 4 the reconstruction of the physical objects is described for both analyses. The recorded data, as well as the considered Monte Carlo samples, are described in chapter 5. Additionally to the samples for the $t\bar{t}$ resonances analysis specific $t\bar{t}$ samples with enhanced charge asymmetry are considered to estimate the modeling uncertainties, see section 5.2.

For the selection of the physical objects required to reconstruct the $t\bar{t}$ quark pair decay and the reconstruction itself an identical setup has been used for both analyses.

The yields for the inclusive measurement are presented in table 6.4 (resolved) and table 6.5 (boosted) with the statistical uncertainties. The total yields separated into the four different mass bins for the differential measurement are listed in table 10.1 (resolved) and 10.2 (boosted).

source	**events in M1**	**events in M2**	**events in M3**	**events in M4**
$t\bar{t}$	517470	31510	310830	22990
single top	32640	1510	8290	430
W + bb,cc	36500	2840	3120	180
W + c	8630	7620	250	8
W + lf	3720	310	20	6
Z+jets	9630	560	920	40
diboson	2490	170	260	10
$t\bar{t}$ +V	1130	80	730	50
multijet	60800	5510	15080	1420
Total	673000	50110	339500	25150
Data	702744	44146	375609	27672

Table 10.1: Events yields in $t\bar{t}$ mass bins (M1-M4) for the differential measurement are listed for the ℓ+jets channel for resolved $t\bar{t}$ decays. The statistical uncertainties are small for all MC channels, see table 6.4, while the uncertainty of the multijet background is provided by the prior of the respective NP.

Source	**events in M1-M3**	**events in M4**
$t\bar{t}$	31900	36600
single top	2500	1620
W + bb,cc	2850	900
W + c	1140	90
W + lf	470	10
Z+jets	300	120
diboson	640	190
$t\bar{t}$ +V	270	350
multijet	4240	1860
Total	44500	41800
Data	37183	39167

Table 10.2: Events yields in the first three (M1-M3) and fourth (M4) $t\bar{t}$ mass bin for the differential measurement are listed for the ℓ+jets channel for boosted $t\bar{t}$ decays. The statistical uncertainties are small for all MC channels, see table 6.5.

10.3 Unfolding - theoretical introduction and application

Every measurement suffers from disturbing influences, with statistical fluctuations and additional random effects due to finite resolution, limited acceptance and restricted efficiency of the reconstruction being the dominant factors. *Detector level* expresses an interpretation of the experimental results including experiment specific biases due to reconstruction efficiencies, i.e. the reconstructed jets and leptons. In addition to that within the so called *parton-level* deficiencies in the migrations through QCD and the event reconstruction are subtracted to allow the comparison of the experimental result and theory predictions or with other independent measurements. Therefore the results of the analysis at hand are provided at parton-level.
The transformation from detector- to parton-level is accomplished in two steps: a bin-by-bin correction for the acceptance and an unfolding of the observed distribution of the figure of merit, $\Delta|y|$, to correct for resolution effects. For the unfolding procedure the PyFBU [168] implementation of the Fully Bayesian Unfolding (FBU) [169] algorithm is used, as it has been sucessfully applied in the previous analysis with 8 TeV [154]. In case of Poisson distributed data events, such as the recorded collision data for the presented approach, the result of the unfolding algorithm is even valid in regions of low statistics, such as the high $t\bar{t}$ mass regions in the differential charge asymmetry measurement, where Poisson probability is not approximated well by a Gaussian.

In this section the procedure is described to raise the detector-level results to parton-level. After a short introduction to the theoretical background of the FBU the applied unfolding procedure for the top charge asymmetry measurement is described. The mathematical description of the unfolding procedure is based on [169–171].

Theoretical introduction to Unfolding

Given a measured spectrum $g(s)$ and a migration model, unfolding targets the estimation of the actual truth spectrum $f(t)$.

$$f(t) \rightarrow g(s)\text{: Folding} g(s) \rightarrow f(t)\text{: Unfolding} \tag{10.2}$$

The Friedholm integral equation of the first kind, equation 10.3, describes the relationship between the distribution $f(t)$ of the truth variable, and the distribution of the measured value $g(s)$, which is related to the truth distribution by migration, distortions and transformations, $K(s,t)$ and additional background contributions, $b(s)$.

$$\int_{\Omega} K(s,t) f(t) dt + b(s) = g(s) \tag{10.3}$$

The Fully Bayesian unfolding offers a complete solution for this problem, which can be expressed analytically through the Bayes theorem, see equation 10.4. There the conditional probability $P(A|B)$ is described for the likelihood of event A occurring, given that B is true, through the conditional probability P(B|A) and the probabilities of observing event

A, $P(A)$, and event B, $P(B)$, independently.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{10.4}$$

In particle physics the common question is:
"*Given the data D and the corresponding migrations model $\mathcal{M}$, what is the truth-level spectrum T of the measured property?*"

With $\Delta|y|$[1] being reconstructed in N_r bins and originating from N_t truth bins, the observed spectrum is defined as $D \in \mathbb{N}^{N_r}$, and the migration matrix as $\mathcal{M} \in \mathbb{R}^{N_r} \times \mathbb{R}^{N_t}$.

Utilizing the Bayes theorem the posterior probability for the truth level spectrum $P(T|D,\mathcal{M})$ can be derived given a prior probability $\pi(T,\mathcal{M})$ and a likelihood function $\mathcal{L}(D|T,\mathcal{M})$, derived from a statistical model for the observed data. The posterior probability is expressed as

$$P(T|D,\mathcal{M}) = \mathcal{L}(D|T,\mathcal{M}) \times \frac{\pi(T,\mathcal{M})}{\text{Norm.Const.}}. \tag{10.5}$$

The former denominator $P(B)$ in equation 10.4 defines a normalization factor, constant for all possible hypothesis, and can be neglected due to our interest in the relative difference of the hypothesis. In combination with omitting the migration model $\mathcal{M}$, since it represents a constant factor, equation 10.5 can be simplified to

$$P(T|D) \approx \mathcal{L}(D|T) \times \pi(T). \tag{10.6}$$

Applied unfolding procedure and ingredients

The transformation of the detector-level results to parton level distinguishes between acceptance and resolution induced deviations and corrects them within the unfolding process. During the unfolding procedure different sampling points are chosen, for which the posterior probability, eq. 10.6, is evaluated through the Likelihood term $\mathcal{L}(D|T)$ and the prior $\pi(T)$. The estimation of uncertainties is introduced via nuisance parameters for background related variations. In this analysis the reconstructed $\Delta|y|$ distributions are initially separated by lepton charge and number of b-tags, and unfolded in parallel.

Concerning the correction of resolution effects with the FBU algorithm the choices for the likelihood $\mathcal{L}(D|T)$ and the prior $\pi(T)$ are described next , followed by the description of the sampling procedure , the introduction of nuisance parameters and the description of the channel combination .

1 The FBU is valid for any measured observable. As this thesis is about the usage of the FBU in context of the charge asymmetry measurement, the according figure of merit is used to describe the principles.

Likelihood

Under the assumption that the spectrum of reconstructed data is conform with Poisson distribution, the likelihood $\mathcal{L}$ can be estimated from the following two pieces of information:

- the transfer matrix $\mathcal{M}$, whose element m_{ij} is the probability $P(r_i|t_j)$ of an event produced in the true bin t_j to be observed in the reconstructed bin r_i.
- the efficiency ε_{t_j} for an event produced in the true bin t_j to be reconstructed in any reconstructed bin r. Calculated from the ratio of events for each of the N_t truth bins of the $\Delta|y|$ distribution before and after the selection, an unskimmed $t\bar{t}$ sample is taken into account to estimate the number of events before.

For the likelihood the observed spectrum D, with d_i events in bin i is compared to the expected one for this bin, r_i, based on the truth distribution T and the migration matrix $\mathcal{M}$ plus background contribution b_i as

$$\mathcal{L}(D|T,\mathcal{M},B) = \prod_{i}^{N_r} Poisson(d_i, r_i(T,\mathcal{M}) + b_i). \tag{10.7}$$

The predicted reconstructed distribution to estimate r_i is described as

$$r_i(T,\mathcal{M}) = \sum_{j=0}^{N_r} \varepsilon_{t_j} \cdot m_{ij} \cdot t_j. \tag{10.8}$$

As the probability of observing d_i events for Poisson distributed cases with an average value of r_i in bin i is given by $P(d_i) = \frac{r_i^{d_i}}{d_i!} e^{-r_i}$ the likelihood for an observed spectrum D can be described with equation 10.9.

$$\mathcal{L}(D|T) = \prod_{i=1}^{N_r} \frac{r_i^{d_i}}{d_i!} e^{-r_i} \tag{10.9}$$

Prior

The simplest prior $\pi(T)$ is an uninformative prior, where it is assumed that all points are equally probable, described by an N_t dimensional step-function as

$$\pi(T) \propto \begin{cases} 1 \text{ if } t_j \in [T^{\ulcorner}, T^{\urcorner}], \forall j \in [1, N_t] \\ 0 \text{ otherwise.} \end{cases} \tag{10.10}$$

The form chosen for the prior is depending on the information known beforehand of the truth distribution T and based on the experience of previous analysis with 8 TeV [171] an uninformative prior is chosen for the presented analysis within the limits $T^{\ulcorner} = \tilde{T}/5$ and $T^{\urcorner} = 2 \cdot \tilde{T}$, where $\tilde{T}$ is the true spectrum in simulation.

Sampling procedure

For the production of the testing points along the iterative unfolding procedure a Markov chain Monte Carlo (MCMC)[172] sampling algorithm[1] is chosen: a variation of the Metropolis-Hastings algorithm [173]. This algorithm is an adaption of a random walk with an acceptance/rejection rule to converge to the specific target distribution [174].
The first sample of equation 10.6 is taken with the MC truth distribution $\tilde{T}$, the origin of the MC reconstructed distribution. The next sampling point is taken randomly, following a uniform distribution within a hyper-box centered at the last position and tested against the acceptance condition as follows

$$(L(D|T)\cdot\pi(T))^{\text{new point}} > (L(D|T)\cdot\pi(T))^{\text{previous point}}. \tag{10.11}$$

In case of

- **acceptance:** the new sample point is kept
- **rejection:** the ratio between $(L(D|T)\cdot\pi(T))^{\text{new point}}$ and $(L(D|T)\cdot\pi(T))^{\text{previous point}}$ is calculated and taken as the probability to accept the new point.

As a result of this, the sampling algorithm tends to drift towards the most likely region in the T-space and samples it.

Marginalization

With marginal likelihood the systematic uncertainties can be estimated with all correlations between uncertainties taken properly into account, as well as different analysis channels combined, such as different b-tag categories, different charge and different topologies. The explicit marginalization is described in the following.

Estimation of systematic uncertainties

Systematic uncertainties, such as variations of Monte Carlo sample cross-sections within their uncertainties, can be included in the unfolding process by adding nuisance parameter terms to $\mathcal{L}(D|T)$, leading to the marginalized likelihood defined in equation 10.12, where θ are the nuisance parameters with their assumed priors $\pi(\theta)$.

$$\mathcal{L}(D|T) = \int \mathcal{L}(D|T,\theta)\pi(\theta)d\theta \tag{10.12}$$

In the work at hand the priors are assumed to be Gaussian distributions G^2 with $\mu = 0$, $\sigma = 1$ and a lower bound to prevent negative background normalizations, with the exception

1 MCMC is a general method based on drawing values θ from approximate distributions and then correcting those draws to better approximate the target posterior distribution $p(\theta,y)$. The sampling is done sequentially, with the distribution of the sampled draws depending on the last value drawn; hence the Markov chain.

2 A Gaussian distribution $G(x|\mu,\sigma)$ is of the form $G(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}$

of the nuisance parameters for the individual multijet bins. In this case uncut Gaussian distributions have been used, to vary the shape of the multijet background distributions binwise.
A nuisance parameter is in general associated with each of the uncertainty sources described in chapter 8, resulting in two categories of nuisances: normalization of background processes (θ_b) and the uncertainties related to object identification, reconstruction and calibration (θ_s). The first ones only affect the background prediction and are referred to as $B(\theta_s,\theta_b)$, while the others influence background and signal and are referred to as $R(T;\theta_s)$.

For the analysis nuisance parameters for the cross-section uncertainties of the individual Monte Carlo samples are considered, as well as to calibrate the estimated multijet background. The W+jets in situ calibration procedure consists of fitting the calibration factors $K_{b\bar{b}/c\bar{c}}$,K_c and K_{lf} for scaling the flavor components of the W+jets background with different charge asymmetries, assuming uniform prior probabilities π during the posterior probability estimation defined in equation 10.13. The b-jet multiplicity provides information about the heavy- and light-flavor composition of the W+jets back-ground, while the lepton charge asymmetry is used to determine the normalization of each component. Figure 10.5 shows the individual W+jets contributions for the different b-jet multiplicities and lepton charges in an exemplary channel. In addition to the expected number of $t\bar{t}$ events for each bin in T, the W+jets calibration factors are free parameters in the likelihood. With these additional information the marginal likelihood can be described as

$$\mathcal{L}(D|T) = \int \mathcal{L}(D|R(T;\theta_s),B(K_{b\bar{b}/c\bar{c}},K_c,K_{ll};\theta_s,\theta_b)) \cdot G(\theta_s) \cdot G(\theta_b) d\theta_s d\theta_b. \tag{10.13}$$

As the integral is in practice too complex to integrate, the evaluation is determined with a Monte Carlo method, where the nuisance parameter is sampled randomly according to its prior distribution. The corresponding pdf term is estimated for each sampling point.

Combination of channels

To estimate especially the W+jets background a combination of orthogonal channels with different background contributions, such as 0-b-tagged jets, 1 exclusive b-tagged jets and 2 inclusive b-tagged jets is considered. As described in [171] the marginalization approach provides a natural framework to treat simultaneously unfolding and background estimation using multiple data regions. For the distributions D_i measured in N_{ch} independent channels, the extension of the likelihood function defined in 10.12 is the product of likelihoods of each channel as shown in equation 10.14, with the nuisance parameters being common to all analysis channels.

$$\mathcal{L}(\{D_1,\cdots D_{N_{ch}}\}|T) = \int \prod_{i=1}^{N_{ch}} \mathcal{L}(D_i|T;\theta) \cdot \pi(\theta) d\theta \tag{10.14}$$

The inclusive measurements are performed using a combination of 18 channels based on the lepton charge (positive or negative), three different categories based on the number of

b-tagged jets (0,1,≥2), and a splitting into four $\Delta|y|$ bins in the signal region, and one bin for the total yield of the 0-b-tagged region, see figure 10.5.

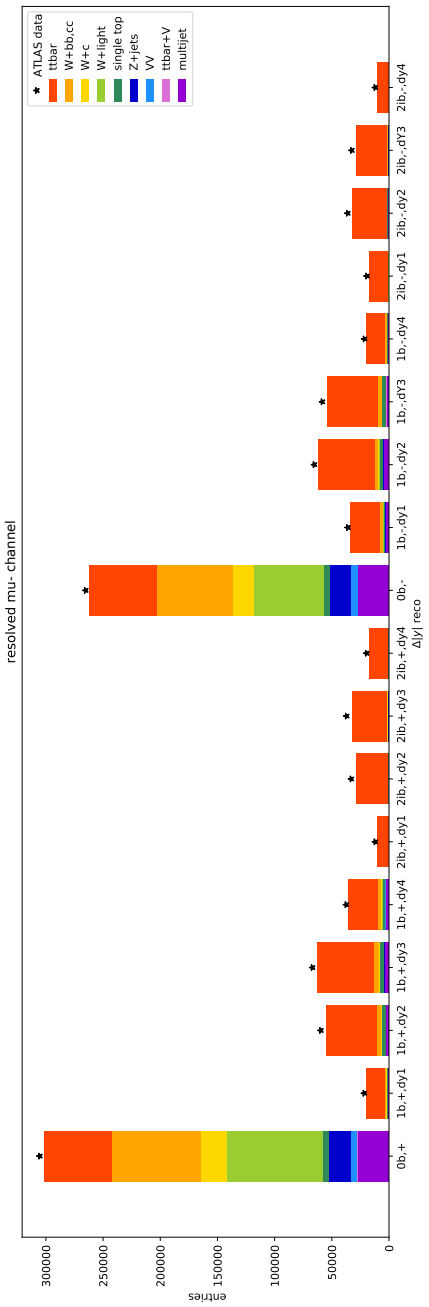


Figure 10.5: Visualization of channels for the resolved μ+jets channel. The breakdown into four bins in $\Delta|y|$ is applied to the signal region with 1 and 2 inclusive b-tagged jets, not in the control region with 0 b-tagged jets.The entries correspond to $\sqrt{s}$=13 TeV and L_{int}=36.1 fb^{-1}.

Two-dimensional unfolding

In addition to the inclusive measurement of the charge asymmetry A_c a differential measurement in mass bins of the reconstructed $t\bar{t}$ system is performed. Hereby two dimensional distributions are unfolded by transforming the two-dimensional histograms into one-dimensional histograms via wrapping the second dimension as consecutive sub ranges of the first dimension. This is displayed in figure 10.6, where a segment of the response matrix is shown to exemplary visualize the wrapping of the two dimensions for the differential $\Delta|y|$, restricted to the resolved channel with exactly 1 b-tagged jet and positive lepton charge.

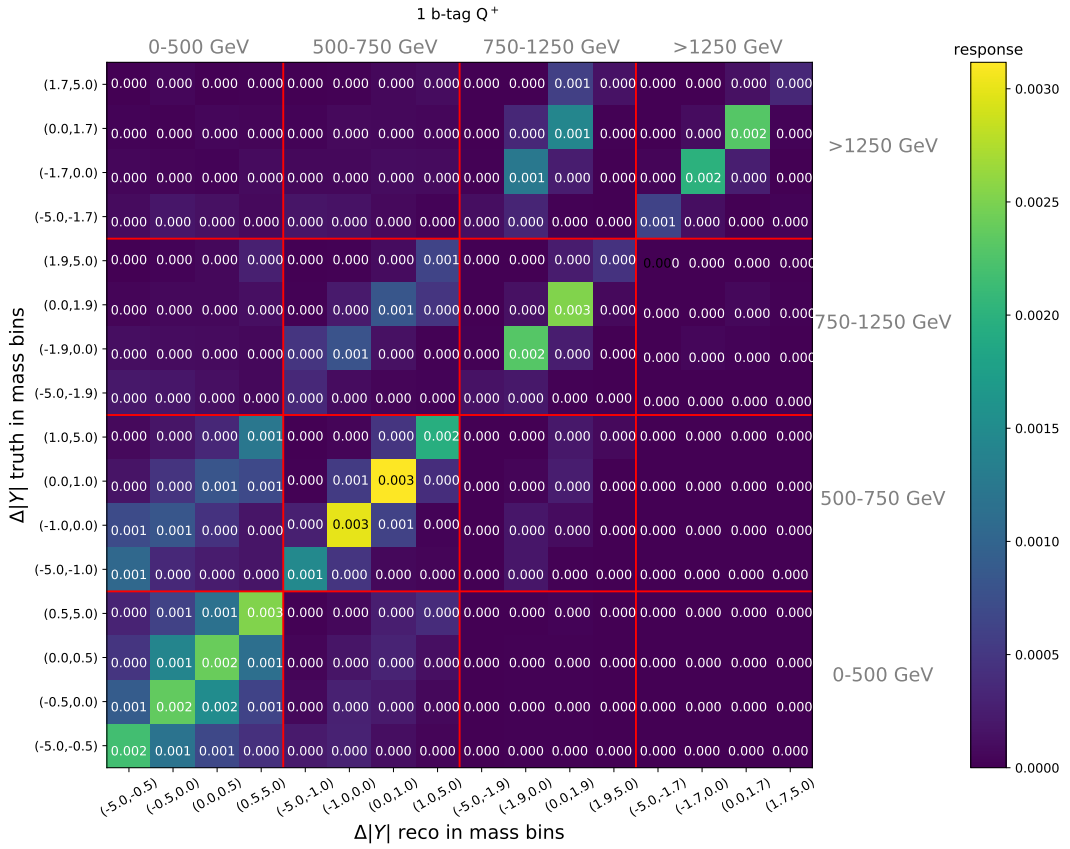


Figure 10.6: Segment of the response matrix for the differential charge asymmetry measurement. It displays the binning for the $\Delta|y|$ measurement in four bins by simultaneously differentiation into four mass bins for the $t\bar{t}$ system. The response matrix is taken from the resolved ℓ+jets channel, restricted to exactly 1 b-tagged jet and positive lepton charge.

10.4 Linearity

With the unfolding approach being defined and set up, a validation has been performed by testing the linearity of the A_c value based on $t\bar{t}$ pseudo data. Here the A_c value estimated

from reconstructed $t\bar{t}$ events is compared to the A_C value directly from the simulated top and antitop particles in Monte Carlo simulation, so called *truth* top and antitop. Different values of charge asymmetry are simulated by re-weighting the pseudo $t\bar{t}$ sample according to a linear function of $\Delta|y|$, defined by a weight $\omega = 1 + K\Delta|y|$. The parameter K is varied between -0.3 and 0.3 in steps of 0.1, introducing asymmetries between approximately ±30%, far beyond the expected charge asymmetry in the standard model. The reweighted samples are illustrated with their respective truth $\Delta|y|$ distribution in figure 10.7.

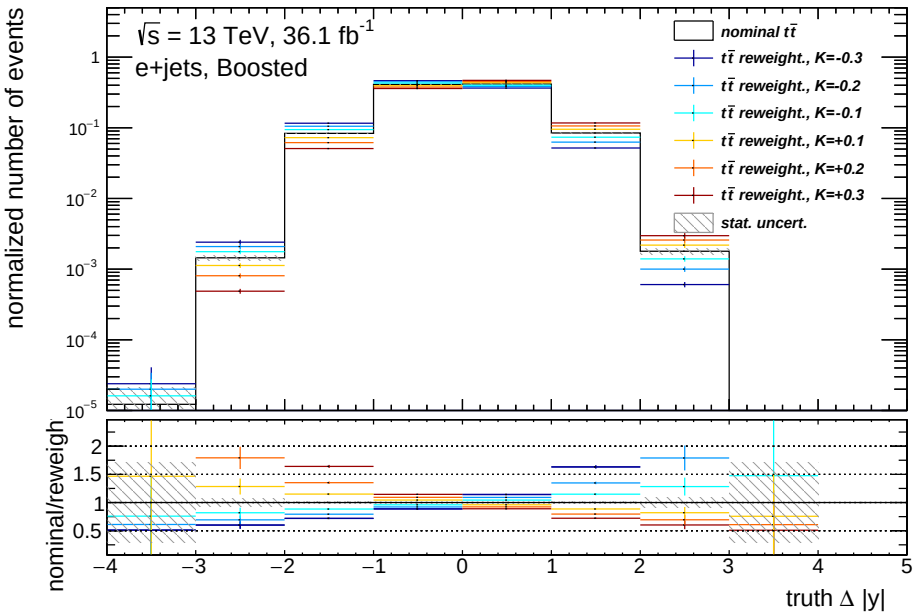


Figure 10.7: Comparison of truth $\Delta|y|$ distribution for nominal $t\bar{t}$ and reweighted $t\bar{t}$ samples with $K = \pm 0.3$, ± 0.2 and ± 0.1 in the boosted e+jets channel.

Based on the linearity, i.e. the slope and the offset of the linear fit of the six tested pseudo-sample variations, the binning of $\Delta|y|$ has been chosen. The linearity test is considered successful, if the slope (offset) of the calibration curve is compatible with 1 (0) within its statistic uncertainty. With an originally planned estimation of a combined resolved and boosted A_c value in mind, an emphasis has been put on the binning choice being acceptable for resolved, boosted and their combination in parallel[1]. For the differential measurement in four $t\bar{t}$ mass bins :[0, 500, 750, 1250, ∞] GeV an individual $\Delta|y|$ binning has been estimated for each mass bin.
For the charge asymmetry measurement the following $\Delta|y|$ binning is set:
Inclusive measurement

- Binning in $\Delta|y|$: [-5.0, -0.8, 0.0, 0.8, 5.0]

1 For the 750-1250 GeV mass bin the slope of the linearity curve differs 2σ from 1.0 for resolved and boosted, but the combination is in an acceptable range within 1σ

Differential measurement in $m_{t\bar{t}}$ bins

- $\leq$**500 GeV** $m_{t\bar{t}}$: [-5.0, -0.5, 0.0, 0.5, 5.0]
- **500-750 GeV** $m_{t\bar{t}}$: [-5.0, -1.0, 0.0, 1.0, 5.0]
- **750-1250 GeV** $m_{t\bar{t}}$: [-5.0, -1.9, 0.0, 1.9, 5.0]
- $\geq$**1250 GeV** $m_{t\bar{t}}$: [-5.0, -1.7, 0.0, 1.7, 5.0]

The calibration curve for the inclusive measurement is presented for the resolved ℓ+jets channel in figure 10.8, while the differential measurement in $t\bar{t}$ mass bins can be found in appendix D.

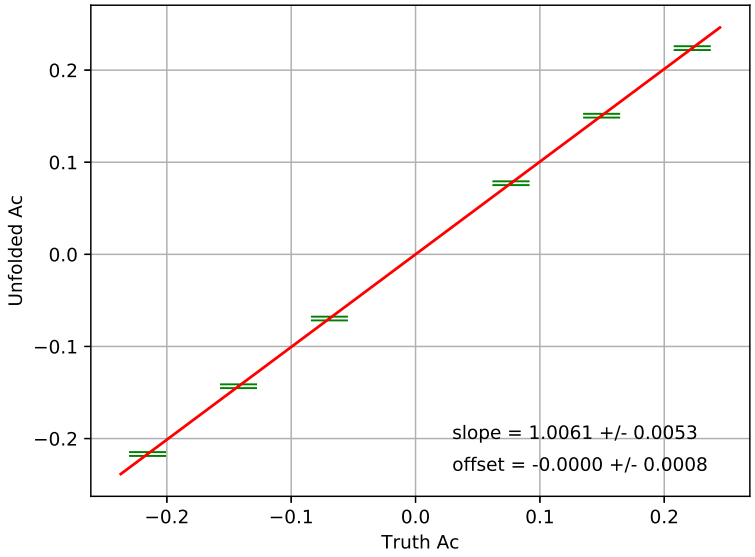


Figure 10.8: Linearity of reconstructed and truth A_c values for the inclusive measurement of $\Delta|y|$ in the resolved ℓ+jets channel.

A satisfying linearity is achieved with the presented choice of $\Delta|y|$ binning for the inclusive and differential measurement of the top charge asymmetry with the results of the linearity checks being summarized in the appendix D.1. Further improvements can be expected with a binning optimization including the mass bin ranges.

10.5 Method validation

The procedure to calibrate the W+jets backgrounds with respect to the cross-section uncertainty, described in section 10.3, is evaluated with well known pseudo-datasets in this section.
At first a validation is performed with an Asimov pseudo-dataset in section 10.5.1, followed by a test to reproduce an artificial modification of the W+jets background contribution in section E.

10.5.1 Validation with Asimov pseudo-dataset

To evaluate the in-situ estimation of the MC background, with focus on the W+jets background estimation, a closure test has been conducted, with Monte Carlo events being compared to an Asimov pseudo-dataset. The pseudo-dataset is identical to the expectation.

For the in-situ marginalization of the QCD estimation individual nuisance parameters are used to estimate the total event yields and the amount of events with a positive or negative lepton. The asymmetry in the distribution shall not be adjustable, since an asymmetry in the y distribution is not physically expected. The following choice of nuisance parameters are complied with these guidelines.

Within the Δy histogram in the inclusive measurement there are in total 18 bins:

- Bin 1 (bin 10) corresponds to the region with 0 b-tagged jets, but with a positive (negative) charged lepton
- Bin 2-9 (11-18) correspond to the region with 1 ($\geq$2) b-tagged jets with four Δy bins for positive/negative lepton charge each

The total event yields are calibrated with a nuisance parameter per b-tag region per lepton flavor, in total 6 individual parameters. To enable the fluctuation regarding the event yields with differently charged leptons, but prohibiting an artificial asymmetry for QCD in the y distribution, the two inner and the two outer bins of the y-distribution are combined for each lepton charge, resulting in four combined bins for the region with 1 b-tagged jet and 2 b-taged jets each. The bin combination with the lowest uncertainty has been removed for each tag region due to the correlation to the previous defined nuisance parameters, which calibrate the total amount. The calibration of the QCD background is based on the statistical uncertainties, with large effects on regions with low statistics, such as the differential measurement with higher $t\bar{t}$ masses.

An overview over the resulting nuisance parameters with an Asimov pseudo-dataset in the combined lepton channel for the resolved topology is presented in figure 10.9. A good agreement with the expectation can be seen, with minor fluctuation in the nuisance parameters within the posterior width.

In table 10.3 the mean value of the individual nuisance parameters is listed, the posterior width with respect to the prior and the prior itself.

10.5.2 Validation with modified W+jets contribution in pseudo-dataset

To evaluate the in-situ approach for the W+jets background in detail, a second test has been conducted, with Monte Carlo events being set up against a pseudo-dataset. This pseudo-dataset differs from the previous Asimov pseudo-dataset in the W+jets background, since the heavy flavor scale factors and the charge asymmetry normalization factors are applied to the W+jets MC samples, see tables 7.5 (el) and 7.6 (μ) for the explicit scale factors.

The figures of merit of the closure test are the estimated scale factors from the nuisance parameters of the W+jets background in comparison to the applied scale factors to the

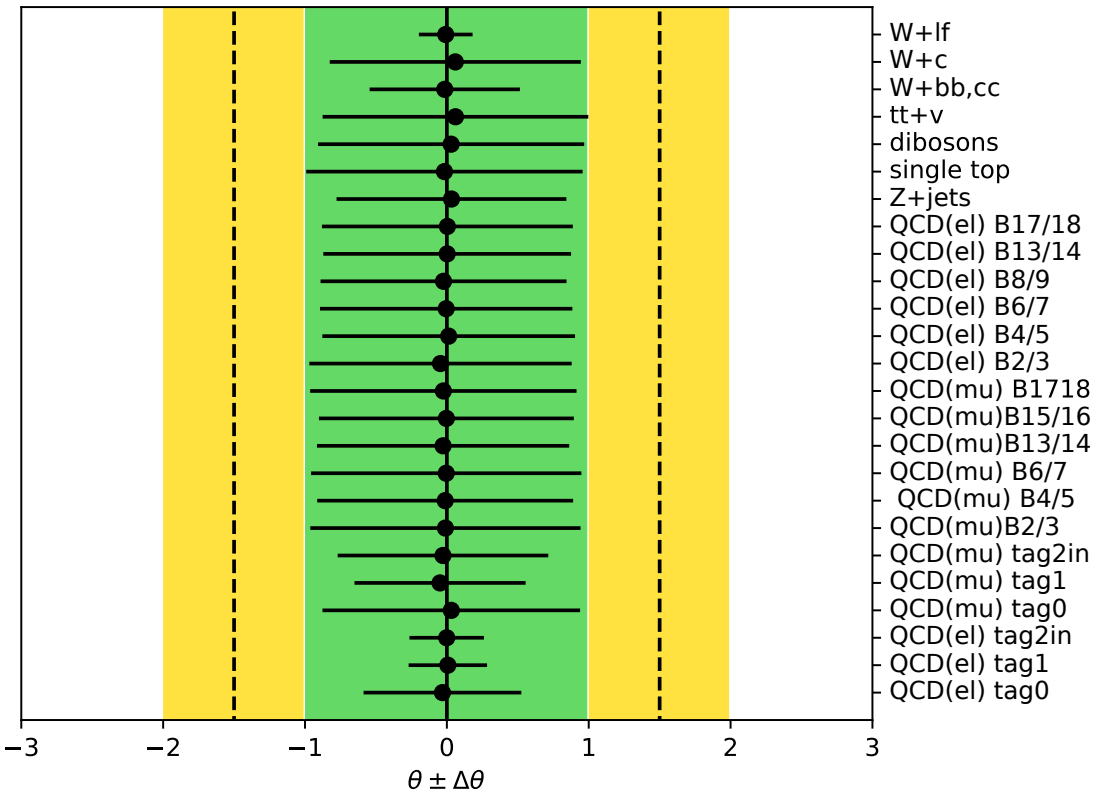


Figure 10.9: Overview nuisance parameters for inclusive measurement in resolved combined lepton+jets channel with Asimov pseudo-dataset. The W+jets background is estimated in-situ. All nuisance parameters agree within the uncertainties with the expectation.

pseudo-dataset. Since a direct comparison with the applied scale factors is not possible, due to their breakdown into different jet-multiplicities, general scale factors have been estimated based on the yields for the corresponding regions covered by the nuisance parameters.

For the in-situ marginalization of the unscaled W+jets background, the nuisance parameters for the three W+jets subsamples are estimated with a Gaussian prior and a variation of 50% is considered, reflecting the large uncertainty on the cross-section for W-bosons with the additional production of jets [175].
A detailed evaluation of the inclusive resolved e+jets channel is presented in this section, while the evaluation of resolved μ+jets and the combined channel for the inclusive measurement can be found in appendix E.

In the considered channel, e+jets in the resolved topology, the scale factors for the marginalization of the W+jets background extracted from the nuisance parameters, SF_{NP} are compared to the equivalent scale factors SF_{HFSF} applied to the pseudo-dataset.

Nuisance parameter	prior	NP	posterior width
QCD el, tag0	0.5	-0.03	0.56
QCD el, tag1	0.5	0.01	0.28
QCD el, tag2 incl.	0.5	0.00*	0.26
QCD mu, tag0	0.5	0.03	0.92
QCD mu, tag1	0.5	-0.05	0.60
QCD mu, tag2 incl.	0.5	-0.03	0.74
Z+jets, prior	0.5	0.03	0.81
single top	0.05	-0.02	0.97
VV	0.5	0.03	0.94
$t\bar{t}$ +V	0.5	0.06	0.94
W+jets bb/cc	0.5	-0.01	0.53
W+jets c	0.5	0.06	0.89
W+jets lf	0.5	-0.01	0.19
QCD mu, bin 2+3	0.16	-0.01	0.95
QCD mu, bin 4+5	0.10	-0.01	0.90
QCD mu, bin 6+7	1.00	0.0*	0.95
QCD mu, bin 13+14	0.17	-0.03	0.90
QCD mu, bin 15+16	0.29	-0.0*	0.90
QCD mu, bin 17+18	1.16	-0.02	0.94
QCD el, bin 2+3	0.10	-0.04	0.92
QCD el, bin 4+5	0.05	0.01	0.89
QCD el, bin 6+7	0.36	0.0*	0.89
QCD el, bin 8+9	0.10	-0.02	0.87
QCD el, bin 13+14	0.09	0.0*	0.87
QCD el, bin 17+18	0.27	0.0*	0.88

Table 10.3: Overview of nuisance parameters for the inclusive measurement in the resolved combined ℓ+jets channel with Asimov pseudo-dataset. The mean value of the nuisance parameters is presented together with the posterior width with respect to the prior. The W+jets background is marginalized in-situ. For the QCD sample the priors per bin-pair are calculated from the individual relative statistical uncertainties per bin and combined via quadratic sum. The entries 0.0* refer to absolute values below 0.01. All priors are Gaussian.

For the closure test the nuisance parameters are estimated in three different scenarios:

1. For all MC generated background samples plus the multijet background uncertainties are considered by application of individual nuisance parameters.
2. Only for W+jets (W+cc/bb, W+c and W+light flavor) and the multijet back-

ground contributions are marginalized in-situ during the unfolding procedure. The uncertainties for non-W backgrounds are neglected.

3. Only for the W+jets subsamples individual nuisance parameters are applied.

The first scenario is representing the final in-situ approach. The results are presented in table 10.4 and with restricting the variation of the non-W MC backgrounds, scenario 2, an agreement of the SF_{HFSF} and SF_{NP} within their uncertainties is reached. For the full background variation in the first setup the differences between SF_{HFSF} and SF_{NP} within a single flavor are assumed to be based on the correlations of W+jets with other MC backgrounds.

	SF_{HFSF}	Δ	SF_{NP} *scen. 1*	Δ	SF_{NP} *scen. 2*	Δ	SF_{NP} *scen. 3*	Δ
W+bb,cc	1.13	0.12	1.04	0.20	1.11	0.17	1.13	0.13
W+c	0.87	0.30	0.98	0.41	0.88	0.38	0.87	0.38
W+lf	0.83	0.05	0.82	0.05	0.84	0.05	0.84	0.04

Table 10.4: Closure test for resolved e+jets channel: The uncertainties of the estimated SF_{HFSF} are the combined uncertainties of the original scale factors for different jet bins. The uncertainties of SF_{NP} are the posterior width uncertainty respecting a prior of 0.5. Agreement between SF_{HFSF} and SF_{NP} is reached within the uncertainties after reduction of background calibration to W+jets and multijet background.

As a conclusion the applied scale factors to the W+jets background can be extracted with the in-situ approach, but they are instable due to correlations with other backgrounds.

10.6 Comparison of W+jets background estimation methods

While in chapter 7 the external estimation of the simulated W+jets background is described, the utilization of nuisance parameters for individual background samples provides a second tool to optimize the Monte Carlo generated asymmetric main background for the charge asymmetry measurement in-situ, described in subsection 10.3.
For previous $t\bar{t}$ charge asymmetry measurements in 8 Tev, see publication [154], the in-situ method has been used. Since at this stage different approaches are available, a validation of the two methods for W+jets background estimation has been performed.

In this section a comparison of the W+jets estimation within the two methods, focusing on the expected uncertainties, is presented. The findings in the inclusive channel for resolved are considered representative for the full analysis, since the events in this channel dominate the measurement in the differential measurement in three of four mass bins. For the fourth bin, the statistics in the boosted channel are too low for a significant decision on an estimation method preference.

For a direct comparison of the two methods of W+jets estimation, first the respective uncertainties are evaluated. For the approach with heavy flavor scale factors the scaled

W+jets sample is used and the MC samples plus the multijet background are compared to the corresponding Asimov pseudo-dataset, representing an integrated luminosity of 36.1 fb^{-1}. Second the uncertainties of the in-situ marginalization are estimated by comparing the unscaled W+jets sample and the MC samples plus multijet background to the corresponding Asimov pseudo-dataset sample. During the unfolding process, the background samples are varied within the given priors, which depend on the uncertainty of the W+jets sample. For the HFSF approach with scaled W+jets MC samples, the uncertainty is estimated from the individual uncertainties of the combined scale factors. For the in-situ approach, the uncertainty of the cross-section of W bosons with additional jet production, 50%, is considered. The applied marginalization factors for the W+jets background after unfolding are provided in table 10.5 for both approaches with the respective uncertainty for the ℓ+jets channel in the resolved topology for the inclusive measurement with Asimov pseudo-datasets.

	HFSF applied			**w/o HFSF**		
	prior	$\mathbf{SF_{NP}}$	$\mathbf{\Delta\ SF_{NP}}$	**prior**	$\mathbf{SF_{NP.}}$	$\mathbf{\Delta\ SF_{NP}}$
$\mathbf{W + bb,cc}$	0.11	1.00	0.09	0.50	0.99	0.27
$\mathbf{W + c}$	0.35	1.00	0.25	0.50	1.03	0.44
$\mathbf{W + lf}$	0.06	1.00	0.05	0.50	1.00	0.09

Table 10.5: Overview over the scale factors, applied to the MC W+jets sample within the in-situ marginalization with and without HFSF application to the W+jets MC sample. MC samples compared to Asimov pseudo-dataset in the resolved ℓ+jets channel

Considering the size of the uncertainties of the resulting scale factors for the μ+jets channel in table 10.5, the approach with the HFSF results in smaller uncertainties. The according comparison for the individual electron and muon channel can be found in appendix F.

As conclusion the usage of the heavy flavor scale factors is slightly preferred over the full in-situ approach. From now on within this chapter the scaled W+jet background is considered as input for the unfolding procedure with the priors listed in table 10.5.

10.7 Dominant uncertainties for measurement

For the here presented charge asymmetry measurement only the dominant uncertainties of those introduced in chapter 8 for the full $t\bar{t}$ resonance measurement are estimated, since the purpose of this study is to evaluate the advantages of a different W+jet background estimation approach regarding the estimated uncertainties.
The systematic uncertainties considered for the analysis are:

- modeling uncertainty, see section 10.7.1
  derived with $t\bar{t}$ samples simulated by different generators and settings
- cross-section uncertainties of background, see section 10.7.2
  covered within the in-situ calibration during unfolding procedure

- uncertainty of the estimated multi-jet background, see section 10.7.3
  covered within the in-situ marginalization during unfolding procedure and additional nuisance parameters for scale and Δy distribution
- uncertainty of the W+jets estimation, see section 10.7.4
  combination of uncertainty of the external heavy flavor scale factor estimation and an additional in-situ approach within these uncertainties
- the uncertainty of the unfolding method, see section 10.7.5
  estimated with artificial asymmetric pseudo-datasets for calibration
- numerical uncertainty, see section 10.7.6
  estimated but negligible, due to the procedure of combining 10 individual measurements to increase precision

10.7.1 Modeling uncertainties

Since the general idea of $t\bar{t}$ modeling systematics is explained in section 8.3.1, here only the procedure for the estimation of the uncertainties for the charge asymmetry measurement is described.
The signal modeling uncertainties are estimated using the same response matrix, obtained from nominal Powheg+Pythia $t\bar{t}$ signal sample, and Asimov pseudo-datasets corresponding to alternative signal models. No background samples are included.

The systematic uncertainty for a given signal model is estimated as

$$\Delta_{syst.} = (A_c^{unf.,model} - A_c^{truth.,model}) - (A_c^{unf.,nom.} - A_c^{truth,nom.}) \tag{10.15}$$

based on the unfolded and truth A_C values of the nominal $t\bar{t}$ and the alternative $t\bar{t}$ sample. A statistical uncertainty based on the number of generated events is assigned to the unfolded asymmetry, $\Delta A_c^{unf.}$, and the truth asymmetry, ΔA_c^{truth}, for each signal sample. The obtained Δ_{syst} are evaluated against statistical fluctuations, since the amount of generated events for the nominal sample is about two to three times the size of the alternative signal model sample. Therefore the statistical uncertainty $\Delta\Delta_{syst}$) is estimated by rescaling the standard deviation of the unfolded results with an individual scale factor SF to the equivalent amount of Poisson-like data events via

$$\Delta\Delta_{syst} = \sqrt{\left(\frac{\partial \Delta_{syst}}{\partial A_c^{x,y}} \Delta A_c^{x,y}\right)^2} = \sqrt{\sum_{x,y} (\Delta A_c^{x,y})^2} \tag{10.16}$$

While $\Delta A_c^{truth,y}$ is directly estimated, $\Delta A_c^{unf.,model}$ is provided by

$$\Delta A_c^{unf.,y} = \frac{RMS_y^{unf.}}{\sqrt{SF_y}}, y \in \{model,nominal\}. \tag{10.17}$$

These scale factors, SF_y, see equation 10.18, are estimated for each $t\bar{t}$ sample, i.e. nominal

and alternative models, with the event weights w_i and the effective entries N_{eff}, corresponding to the size of a Poisson distributed dataset with equal statistical uncertainties[1].

$$SF = \frac{N_{eff}}{\sum_i w_i} \tag{10.18}$$

For the final modeling systematic uncertainty the maximum of $\Delta_{syst.}$ and $\Delta\Delta_{syst.}$ is taken to represent the systematic uncertainty of the individual modeling uncertainty, e.g radiation, parton showering and hard scattering. The estimation of the radiation uncertainty is based on the mean value of high and low radiation in comparison to the nominal $t\bar{t}$ sample. In addition to that the hard scattering uncertainty is defined by the comparison of MC samples produced with the generator combination MC NLO+Herwig and Powheg+Herwig. The results are listed for the inclusive charge asymmetry measurement in table 10.6 and for the differential approach in table 10.7 and 10.8. In appendix G the values considered as input for the evaluation of the systematic uncertainty are provided.

model	**uncertainty**
radiation	0.003
parton showering	0.002
hard scattering	0.004
TOTAL	0.005

Table 10.6: $t\bar{t}$ modeling uncertainties for inclusive top charge asymmetry measurement in resolved ℓ+jets channel. For the systematic uncertainty the maximum of $\Delta_{syst.}$ and $\Delta\Delta_{syst}$ is taken.

	differential in $t\bar{t}$ mass bins, resolved			
model	**<500 GeV**	**500-750 GeV**	**750-1250 GeV**	**>1250 GeV**
radiation	0.009	0.004	0.126	0.045
parton showering	0.008	0.005	-0.020	0.089
hard scattering	0.008	0.008	0.025	0.076
TOTAL	0.014	0.010	0.034	0.125

Table 10.7: $t\bar{t}$ modeling uncertainties for differential top charge asymmetry measurement in resolved ℓ+jets channel. For the systematic uncertainty the maximum of $\Delta_{syst.}$ and $\Delta\Delta_{syst}$ is taken.

1 The statistical uncertainty for MC samples is defined as $\Delta_{stat.,MC} = \frac{\sqrt{\sum_i wi^2}}{\sum_i w_i}$ and for Poisson distributed data as $\Delta_{stat.,D} = \frac{\sqrt{N}}{N}$. The resulting effective entries N_{eff} are calculated as $N_{eff} = \frac{\left(\sum_i w_i\right)^2}{\sum_i wi^2}$.

	differential in $t\bar{t}$ mass bins, boosted			
model	**<500 GeV**	**500-750 GeV**	**750-1250 GeV**	**>1250 GeV**
radiation	0.225	0.085	0.457	0.025
parton showering	0.236	0.072	0.048	0.027
hard scattering	0.328	0.090	-0.096	-0.056
TOTAL	0.463	0.143	0.117	0.067

Table 10.8: $t\bar{t}$ modeling uncertainties for differential top charge asymmetry measurement in boosted ℓ+jets channel. For the systematic uncertainty the maximum of $\Delta_{syst.}$ and $\Delta\Delta_{syst}$ is taken.

10.7.2 Cross-section uncertainties

For the nuisance parameters covering the Monte Carlo yield variation due to cross-section uncertainties, Gaussian distributed priors are used. The following values are applied as variation prior for the nuisance parameters

- Z+jets: 50%
- single top: 5%
- diboson: 50%
- $t\bar{t}$ +V: 50%

While the cross-section uncertainty for the single top production is estimated as 5%, due to the fact, that the main contribution comes from the Wt channel, see subsection 8.3.2, for the other backgrounds a conservative cross-section normalization uncertainty of 50% is applied, representing the large uncertainty in cross-sections with $\geq$4 jets.
The cross-section uncertainty on $t\bar{t}$ is not directly considered for the charge asymmetry measurement, since the amount of $t\bar{t}$ events is varied for each sampling point with a flat prior.

10.7.3 Multijet background uncertainty

For the mulitjet background a total uncertainty concerning the total yields is assumed to be 50%, based on the conservative estimation in [3] and represented by six nuisance parameters, one for each lepton channel per measurement and per tag region (0 b-tagged jets, 1 b-tagged jets and at least 2 b-tagged jets). Gaussian priors are considered. Additionally individual nuisance parameters for combinations of Δy bins, in detail described in subsection 10.5.1 are used. The nuisance parameters of the Δy bins are varied during the unfolding with an uncut Gaussian distribution and the combined statistical uncertainty of the combined bins.

10.7.4 W+jets

The externally estimated W+jets background has a systematic uncertainty for each subsample (W+b/cc, W+c and W+light flavor) from the estimation of the heavy flavor scale factors and the charge asymmetry normalization factor. During the unfolding process Gaussian distributed nuisance parameter of the size of the external systematic uncertainties are utilized to adapt the W+jets background.

W-fraction	**uncertainty**
W+bb,cc	0.11
W+c	0.35
W+lf	0.06

Table 10.9: Uncertainty of W+jets estimation with HFSF approach, used with an Gaussian distribution to adapt with in-situ approach during unfolding procedure. A conservative estimation of the uncertainty of the scale factors for normalization and flavor scaling is considered and used for both lepton channels and the combination.

10.7.5 Unfolding uncertainty

The unfolding uncertainty represents the not exactly linear correlation between reconstructed and truth A_c values, see section 10.4. Here the uncertainty of the offset is dominant with respect to the uncertainty of the slope, since the value of the charge asymmetry is of the same order of magnitude, $\mathcal{O}(10^{-3})$. Table 10.10 shows the results for the inclusive measurement, the differential overview is presented in the appendix D.

resolved	**slope**	**offset**
e+jets	1.005 $\pm$0.007	$\pm$0.001
μ+jets	1.009$\pm$0.008	$\pm$0.001
combined ℓ+jets	1.006$\pm$0.005	$\pm$0.001

Table 10.10: Overview of linearity of truth to reconstructed charge asymmetry in the inclusive measurement for the resolved topology.

10.7.6 Numerical uncertainty

To investigate the numerical uncertainty of this statistical approach, the final results for each channel are merged from 10 individual repetitions of the FBU calculation. For the inclusive and differential measurement the uncertainty is of the order of $\mathcal{O}(10^{-5})$, and further neglected due to its size in comparison to the other uncertainties $\mathcal{O}(10^{-3})$.

10.8 Results of inclusive measurement

The inclusive measurement has been performed for the resolved topology only, since this channel is dominating the overall statics. The decay channels of the $t\bar{t}$ pair e+jets and

μ+jets are combined into one channel, ℓ+jets.
The validation of the method with an Asimov pseudo-dataset has been conducted first. An overview over the resulting nuisance parameters - all within the posterior width, is presented in figure 10.9 in section 10.5.1.

The measurement has been conducted with the full 2015 and 2016 dataset and a total integrated luminosity of 36.1 fb^{-1}. Figure 10.10 displays the resulting mean values of the nuisance parameters. All nuisance parameters for the background samples generated from Monte Carlo are within one σ, the posterior width. For QCD shifts close to two σ are observed. Since the estimation of QCD is not ideal, because of large statistic uncertainties in some bins, these fluctuations can be understood.

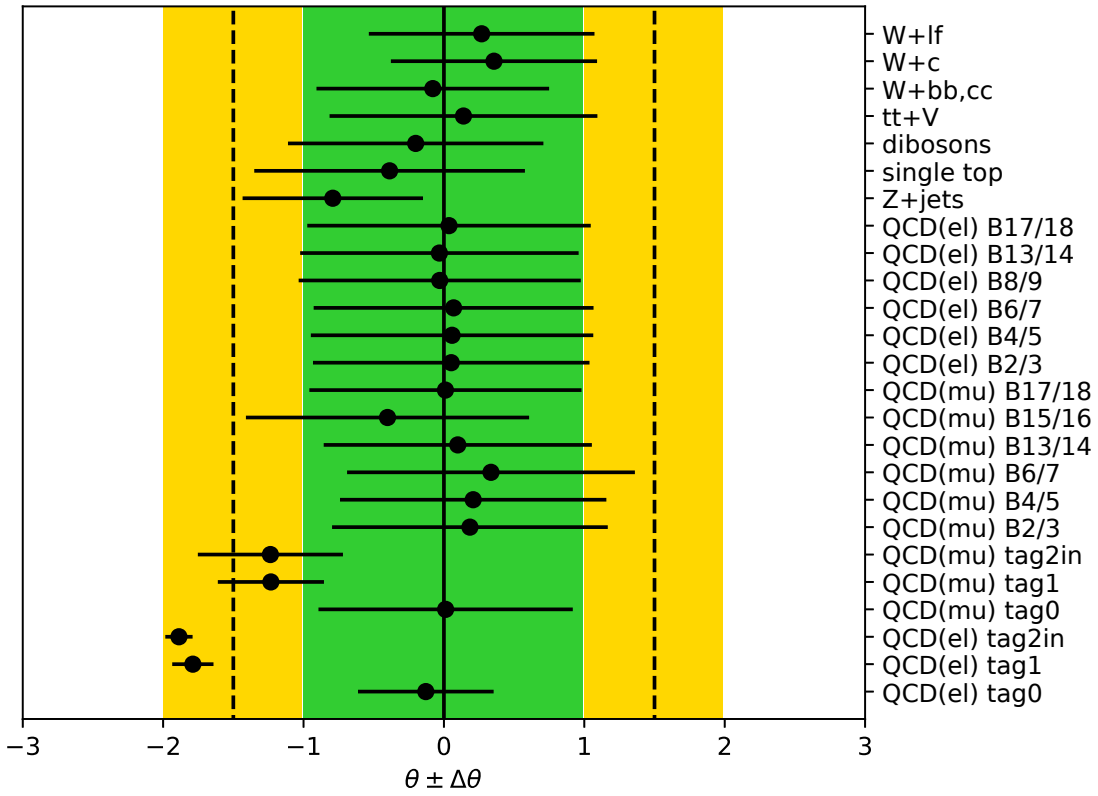


Figure 10.10: Overview of nuisance parameters for inclusive measurement in resolved combined lepton+jets channel with 36.1 fb^{-1} of collected data. The W+jets background is scaled with heavy flavor scale factors and within their uncertainties estimated in-situ. All nuisance parameters agree within their uncertainties with the expectation.

The final result of the inclusive measurement of the top-antitop production charge asymmetry in the resolved topology for the combined lepton+jets channel is presented in table 10.11.

	ATLAS Data 36.1 fb^{-1}					
	A_c	**Stat.+Syst.**	**Model**	**Unfold.**	**Total unc.**	**SM prediction**
resolved	0.0069	0.0021	0.0053	0.0008	0.0058	$0.0064^{+0.0005}_{-0.0006}$

Table 10.11: Results for the inclusive measurement of the ℓ+jets channel is presented for resolved topology with the combined statistical and systematical uncertainty (including the MC cross-section uncertainty and the W+jets uncertainty), $t\bar{t}$ modeling uncertainty and the uncertainty due to the unfolding bias. The total uncertainty is the quadratic sum of the listed individual uncertainties. The Standard Model predictions are calculated at NNLO in QCD and NLO in electroweak theory[4].

10.9 Results of differential measurement

The differential measurement has been performed in the ℓ+jets channel, the combined analysis for the electron and muon decay channel. For the first three mass bins, up until a $t\bar{t}$ mass of 1250 GeV, only the resolved topology is considered, while the last mass bin, above 1250 GeV, is evaluated for the boosted topology. This split has shown the best results regarding agreement with expectations for the Asimov pseudo-dataset. In addition to that, the nuisance parameters in the last mass bin for W+jets have been split into two sets with respect to the lepton flavor, instead of one combined set, to accommodate the different pulls for the W+jets background from electron and muon channel in this low statistic region. Under these modified conditions all nuisance parameters agree within their uncertainties to the expectations for the method evaluation with an Asimov pseudo-dataset, see section H.1 in the appendix.

In figure 10.11 the resulting mean values of the nuisance parameters for data are displayed for the 4th bin in the boosted topology. It is noticeable, that subsets from the W+jets and QCD background are marginalized with a nuisance parameter of 1.5σ and more beyond the expectation for a well modeled background. In table 10.12 the corresponding nuisance parameters are listed in detail, in table H.2 in the appendix all nuisance parameters are documented. The largest discrepancy between before and after marginalization within the unfolding procedure is visible in figure 10.11 for the W+jets background, with the jets originating from a pair of b- or c-quarks and leptonically decaying with a muon. Considering the effectively applied marginalization the expected $W + b\bar{b}$,$c\bar{c}$ background contribution of 2% (approx. 500 events in the μ+jets channel) with respect to the total expected yields before calibration is reduced by 50±20 %. The additional visible large fluctuations are corresponding to individual QCD bins (e or μ channel) and the total QCD yields for the 1 tagged b-jet region in the e-channel. These corrections to a not well modeled background, struggling with low statistics in the higher mass region during estimation and here in application, are expected and supported by the allocation in bin-wise and overall nuisance parameters for QCD.

The final results of the differential measurement of the top-antitop production charge asymmetry in four bins of $m_{t\bar{t}}$ in the combined lepton channel are presented in table 10.13.

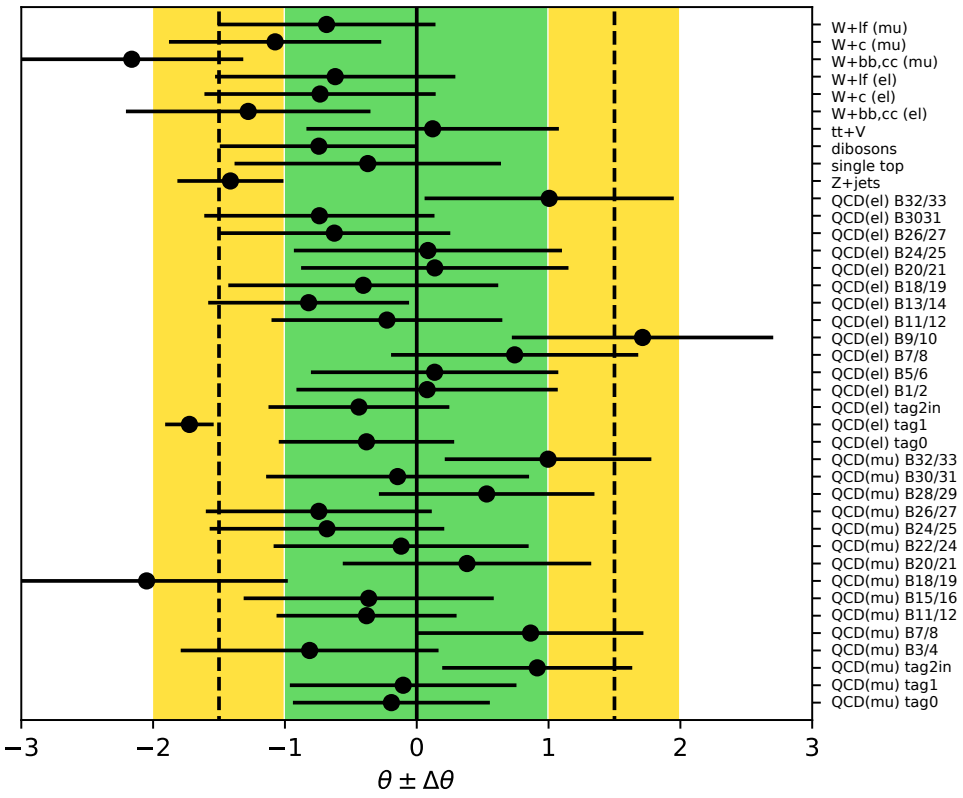


Figure 10.11: Overview of nuisance parameters for differential measurement in the combined lepton+jets channel with 36.1 fb^{-1} collected data. The displayed nuisance parameters and their posterior width represent the fourth mass bin, $m_{t\bar{t}}$>1250 GeV, measured within the boosted topology. The W+jets background is scaled with heavy flavor scale factors and within their uncertainties calibrated in-situ.

background sample	prior	NP	posterior width
W+bb/cc (μ)	0.23	-2.2	0.8
QCD, bin 9+10 (e)	0.5	1.7	1.0
QCD, tag1 (e)	0.5	-1.7	0.2

Table 10.12: The largest nuisance parameter values for the fourth mass bin are listed with the posterior width with respect to the prior and the corresponding prior. The nuisance parameter for QCD muon bin 18+19 is not further discussed here, since it references to a $m_{t\bar{t}}$ bin covered by the resolved channel. All priors are Gaussian.

Standard Model predictions for the asymmetry in top pair production currently exist for NNLO calculations at QCD level and NLO for electroweak, see [176] for details. Due to the large $t\bar{t}$ modeling uncertainties the processing of the calculations for the here presented

	ATLAS Data 36.1 fb^{-1}					
	A_c	**Stat.+Syst.**	**Model**	**Unfold.**	**Total unc.**	**SM prediction**
res. M1	0.0155	0.0048	0.0142	0.0020	0.0151	0.0033±0.0001
res. M2	0.0003	0.0043	0.0098	0.0016	0.0108	0.0043± 0.0001
res. M3	0.0228	0.0132	0.0343	0.0048	0.0371	0.0050±0.0003
boost. M4	0.0152	0.0495	0.0673	0.0099	0.0841	0.0065±0.0013

Table 10.13: Results for the differential measurement in the ℓ+jets channel is presented in four mass bins, exclusively measured in the resolved and boosted topology respectively. The combined statistical and systematical uncertainty (including the MC cross-section uncertainty and the W+jets uncertainty), $t\bar{t}$ modeling uncertainty and the uncertainty due to the unfolding bias are listed. The total uncertainty is the quadratic sum of the individual uncertainties. The Standard Model predictions are extracted from the Powheg+Pythia 6 $t\bar{t}$ sample at NLO.

mass bins has been deemed redundant. Reference values for related mass bins can be found in [4]. Instead the values provided by the MC sample with NLO precision are used for the comparison.

10.10 Discussion

The study at hand was performed in order to choose the strategy for the W+jets estimation in the publication of the top charge asymmetry measurement [4]. As a conclusion an improvement is obtained with the heavy flavor scale factors estimated in an external control region, see section 10.6. In order to utilize this approach for the publication, including data collected in 2017, new scaling factors have to be estimated. In total the improvement was not sufficient to justify the effort of deriving these factors for the 2017 dataset. Therefore the in-situ strategy was used for the publication with data collected from 2015 to 2017.

The publication [4] represents the latest results for a measurement of the top charge asymmetry at $\sqrt{s}$=13 TeV with ATLAS and is performed in the ℓ+jets channel. The CMS Collaborations published in [177] a similar measurement in the ℓ+jets channel. The latest joined publication of the ATLAS and CMS collaboration contains the combination of the inclusive and differential $t\bar{t}$ charge asymmetry measurement using data collected at ATLAS and CMS at $\sqrt{s}$=7 TeV and 8 TeV, see [166] for details.

11 Summary and outlook

In the last years since the discovery of the Higgs boson in 2012, all eyes have been on the LHC as the laergest accelerator of this time, eager to discover the next new particle to solve the open mysteries in particle physics. But currently no new elementary particles have been discovered, neither at multi-purpose experiments like ATLAS and CMS, nor at individual experiments with different approaches, specialized in for e.g. direct dark matter searches. Therefore the focus has shifted to high precision measurements to provide insight into current theoretical models, including the until this day well established standard model of particle physics. In the work at hand an exemplary measurement for each approach has been conducted with the ATLAS detector using top quarks. For the search for new particles the mass spectrum of the decay particles of a top quark pair has been evaluated to find a new particle, decaying into top and antitop, as predicted for the theoretical Z' boson or others. Within the measurement no new particles could be discovered but new exclusion limits for the specific considered theoretical models could be set, e.g. the mass of a g_{KK} with a 15% width is excluded below 3.8 TeV with 90% Cl. In the second part of the thesis a precise measurement of the charge asymmetry in the production of top quark pairs was performed, resulting e.g. in A_C=0.0069±0.0058 for the inclusive measurement in the ℓ+jets channel for the resolved topology. Within the uncertainties all results agree with the expectation from the standard model of particle physics, but due to the large statistical and modeling uncertainties especially for the higher mass region in the differential approach, the potential of this kind of asymmetry measurement is not yet exhausted. Future plans for the exploration of the top-antitop charge asymmetry include the top-antitop energy asymmetry measurement as described in [178, 179], and the continuation of a reduction of systematic uncertainties.
Currently[1] the LHC is shut down for a technical break to get ready for the next phase of collecting data and also to prepare for the upgrade of the LHC to the High-Luminosity LHC to be run from 2026. Only time and the data collected within the ATLAS detector and the other experiments already existing or in preparation at CERN and all over the world will show, what lies ahead and will provide us with the answers to the final questions of life, the universe and everything ... but answering the questions about dark matter and the speculations about the grand unification theory would be fine for the beginning.

1 Since December 2018 the LHC is shut down for approximately two years.

12 Bibliography

[1] H. G. Liddel and R. Scott.
'A Greek-English Lexicon at the Perseus Project'.
φνσιζ.
2020, accessed 11.04.2020.
URL: http://www.perseus.tufts.edu.

[2] C. Lorre and B. Prady.
The Big Bang Theory: The Gorilla Experiment, season 3, episode 10.
Television Comedy.
2007.

[3] M. Aaboud et al.
'Search for heavy particles decaying into top-quark pairs using lepton-plus-jets events in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector'.
In: *The European Physical Journal C* 78.7 (2018).
DOI: 10.1140/epjc/s10052-018-5995-6.
arXiv: 1804.10823 [hep-ex].

[4] The ATLAS Collaboration.
'Inclusive and differential measurement of the charge asymmetry in $t\bar{t}$ events at 13 TeV with the ATLAS detector'.
In: *CERN Document Server* ATLAS-CONF-2019-026 (July 2019).
URL: http://cds.cern.ch/record/2682109.

[5] M. Gell-Mann.
'Symmetries of Baryons and Mesons'.
In: *Phys. Rev.* 125 (3 Feb. 1962), pp. 1067–1084.
DOI: 10.1103/PhysRev.125.1067.

[6] Y. Ne'eman.
The application of Lie groups to invariance principles in physics.
Thesis(PhD)–University of London (Imperial College of Science and Technology), 1961-62. Index to theses, 12-977.
1961.
URL: http://catalogue.libraries.london.ac.uk/record=b2997936~S16.

[7] Nobel Foundation.
The Nobel Prize in Physics 1969.
Nobel Media AB 2020. Fri. 28 Feb 2020.
1969.
URL: https://www.nobelprize.org/prizes/physics/1969/summary/.

[8] M. Gell-Mann.

'The Symmetry Group of the Vector and Axial Vector Currents'.
In: *Science* 144.3618 (1964), pp. 563–563.
DOI: 10.1126/science.144.3618.563.

[9] G Zweig.
An SU_3 model for strong interaction symmetry and its breaking; Version 1.
Tech. rep. CERN-TH-401.
Geneva: CERN, Jan. 1964.
URL: http://cds.cern.ch/record/352337.

[10] M. Tanabashi et al.
'Review of Particle Physics - Summary tables'.
In: *Phys. Rev. D* 98 (3 Aug. 2018 and 2019 update).
DOI: 10.1103/PhysRevD.98.030001.

[11] M. Tanabashi et al.
'Review of Particle Physics -Top quark'.
In: *Phys. Rev. D* 98 (3 Aug. 2018 and 2019 update).
DOI: 10.1103/PhysRevD.98.030001.

[12] M. Aker et al.
'Improved Upper Limit on the Neutrino Mass from a Direct Kinematic Method by KATRIN'.
In: *Physical Review Letters* 123.22 (Nov. 2019).
DOI: 10.1103/physrevlett.123.221802.

[13] R. Aaij et al.
'Observation of $J/\Psi p$ Resonances Consistent with Pentaquark States in $\lambda_b^0 \to J/\Psi K^- p$Decays'.
In: *Physical Review Letters* 115.7 (Aug. 2015).
DOI: 10.1103/physrevlett.115.072001.

[14] Y. Fukuda et al.
'Evidence for Oscillation of Atmospheric Neutrinos'.
In: *Physical Review Letters* 81.8 (Aug. 1998), pp. 1562–1567.
DOI: 10.1103/physrevlett.81.1562.

[15] Q. R. Ahmad et al.
'Direct Evidence for Neutrino Flavor Transformation from Neutral-Current Interactions in the Sudbury Neutrino Observatory'.
In: *Physical Review Letters* 89.1 (June 2002).
DOI: 10.1103/physrevlett.89.011301.

[16] I. Esteban et al.
'Global analysis of three-flavour neutrino oscillations: synergies and tensions in the determination of θ23, δCP, and the mass ordering'.
In: *Journal of High Energy Physics* 2019.1 (Jan. 2019).
DOI: 10.1007/jhep01(2019)106.

[17] G. Aad et al.

'Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC'.
In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29.
DOI: 10.1016/j.physletb.2012.08.020.

[18] S. Chatrchyan et al.
'Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC'.
In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61.
DOI: 10.1016/j.physletb.2012.08.021.

[19] F. Abe et al.
'Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab'.
In: *Phys. Rev. Lett.* 74 (14 Apr. 1995), pp. 2626–2631.
DOI: 10.1103/PhysRevLett.74.2626.

[20] S. Abachi et al.
'Observation of the Top Quark'.
In: *Phys. Rev. Lett.* 74 (14 Apr. 1995), pp. 2632–2637.
DOI: 10.1103/PhysRevLett.74.2632.

[21] S. Dulat et al.
'New parton distribution functions from a global analysis of quantum chromodynamics'.
In: *Physical Review D* 93.3 (Feb. 2016).
DOI: 10.1103/physrevd.93.033006.

[22] M. Czakon, P. Fiedler, and A. Mitov.
'Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$'.
In: *Physical Review Letters* 110.25 (June 2013).
DOI: 10.1103/physrevlett.110.252004.

[23] N. Cabibbo.
'Unitary Symmetry and Leptonic Decays'.
In: *Phys. Rev. Lett.* 10 (12 June 1963), pp. 531–533.
DOI: 10.1103/PhysRevLett.10.531.

[24] M. Kobayashi and T. Maskawa.
'CP-Violation in the Renormalizable Theory of Weak Interaction'.
In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657.
DOI: 10.1143/PTP.49.652.
eprint: http://oup.prod.sis.lan/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf.

[25] M. Tanabashi et al.
'Review of Particle Physics - CKM matrix'.
In: *Phys. Rev. D* 98 (3 Aug. 2018 and 2019 update), p. 030001.

DOI: 10.1103/PhysRevD.98.030001.

[26] C. T. Hill.
'Topcolor assisted technicolor'.
In: *Phys. Lett.* B345 (1995), pp. 483–489.
DOI: 10.1016/0370-2693(94)01660-5.
arXiv: hep-ph/9411426 [hep-ph].

[27] V.A. Miransky, M. Tanabashi, and K. Yamawaki.
'Dynamical electroweak symmetry breaking with large anomalous dimension and t quark condensate'.
In: *Physics Letters B* 221.2 (1989), pp. 177–183.
DOI: https://doi.org/10.1016/0370-2693(89)91494-9.

[28] K. Lane and E. Eichten.
'Natural topcolor-assisted technicolor'.
In: *Physics Letters B* 352.3-4 (June 1995), pp. 382–387.
DOI: 10.1016/0370-2693(95)00482-z.

[29] R. M. Harris, C. T. Hill, and S. J. Parke.
'Cross-Section for Topcolor Z'_t Decaying to $t\bar{t}$'.
In: *FERMILAB-FN-0687* (1999).
arXiv: hep-ph/9911288 [hep-ph].

[30] G. Aad et al.
'A search for $t\bar{t}$ resonances using lepton-plus-jets events in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector'.
In: *Journal of High Energy Physics* 2015.8 (Aug. 2015).
DOI: 10.1007/jhep08(2015)148.

[31] A. M. Sirunyan et al.
'Search for $t\bar{t}$ resonances in highly boosted lepton+jets and fully hadronic final states in proton-proton collisions at $\sqrt{s} = 13$ TeV'.
In: *Journal of High Energy Physics* 2017.7 (July 2017).
DOI: 10.1007/jhep07(2017)001.

[32] A. Albert et al.
'Recommendations of the LHC Dark Matter Working Group: Comparing LHC searches for heavy mediators of dark matter production in visible and invisible decay channels'.
In: *CERN-LPCC-2017-01* (2017).
arXiv: 1703.05703 [hep-ex].

[33] L. Randall and R. Sundrum.
'Large Mass Hierarchy from a Small Extra Dimension'.
In: *Physical Review Letters* 83.17 (Oct. 1999), pp. 3370–3373.
DOI: 10.1103/physrevlett.83.3370.

[34] L. Fitzpatrick et al.
'Searching for the Kaluza-Klein graviton in bulk RS models'.

In: *Journal of High Energy Physics* 2007.09 (Sept. 2007), pp. 013–013.
DOI: 10.1088/1126-6708/2007/09/013.

[35] K. Agashe et al.
'Warped gravitons at the CERN LHC and beyond'.
In: *Physical Review D* 76.3 (Aug. 2007).
DOI: 10.1103/physrevd.76.036006.

[36] G. Aad et al.
'Search for resonant diboson production in the $\ell\ell q\bar{q}$ final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector'.
In: *The European Physical Journal C* 75.2 (Feb. 2015).
DOI: 10.1140/epjc/s10052-015-3261-8.

[37] V. Khachatryan et al.
'Search for massive resonances decaying into pairs of boosted bosons in semi-leptonic final states at s $\sqrt{s} = 8$ TeV'.
In: *Journal of High Energy Physics* 2014.8 (Aug. 2014).
DOI: 10.1007/jhep08(2014)174.

[38] A. M. Sirunyan et al.
'Search for ZZ resonances in the $2\ell 2\nu$ final state in proton-proton collisions at 13 TeV'.
In: *Journal of High Energy Physics* 2018.3 (Mar. 2018).
DOI: 10.1007/jhep03(2018)003.

[39] B. Lillie, L. Randall, and L. T. Wang.
'The Bulk RS KK-gluon at the LHC'.
In: *Journal of High Energy Physics* 2007.09 (Sept. 2007), pp. 074–074.
DOI: 10.1088/1126-6708/2007/09/074.

[40] G. Aad et al.
'Search for $t\bar{t}$ resonances in the lepton plus jets final state with ATLAS using 4.7 fb^{-1}of pp collisions at $\sqrt{s} = 7$ TeV'.
In: *Physical Review D* 88.1 (July 2013).
DOI: 10.1103/physrevd.88.012004.

[41] A. M. Sirunyan et al.
'Search for $t\bar{t}$ resonances in highly boosted lepton+jets and fully hadronic final states in proton-proton collisions at $\sqrt{s} = 13$ TeV'.
In: *Journal of High Energy Physics* 2017.7 (July 2017).
DOI: 10.1007/jhep07(2017)001.

[42] G. Aad et al.
'Measurement of W± and Z-boson production cross sections in pp collisions at s=13 TeV with the ATLAS detector'.
In: *Physics Letters B* 759 (2016), pp. 601–621.
DOI: https://doi.org/10.1016/j.physletb.2016.06.023.

[43] C. H. Kom and W. J. Stirling.

'Charge asymmetry in W + jets production at the LHC'.
In: *The European Physical Journal C* 69.1-2 (June 2010), pp. 67–73.
DOI: 10.1140/epjc/s10052-010-1353-z.

[44] G. Aad et al.
'Measurement of the production of a W boson in association with a charm quark in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector'.
In: *Journal of High Energy Physics* 2014.5 (May 2014).
DOI: 10.1007/jhep05(2014)068.

[45] S. Chatrchyan et al.
'Measurement of associated W + charm production in pp collisions at $\sqrt{s} = 7$ TeV'.
In: *Journal of High Energy Physics* 2014.2 (Feb. 2014).
DOI: 10.1007/jhep02(2014)013.

[46] W. J. Stirling and E. Vryonidou.
'Charm Production in Association with an Electroweak Gauge Boson at the LHC'.
In: *Physical Review Letters* 109.8 (Aug. 2012).
DOI: 10.1103/physrevlett.109.082002.

[47] The ATLAS Collaboration.
'The ATLAS Experiment at the CERN Large Hadron Collider'.
In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08003–S08003.
DOI: 10.1088/1748-0221/3/08/s08003.

[48] L. Evans and P. Bryant.
'LHC Machine'.
In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001–S08001.
DOI: 10.1088/1748-0221/3/08/s08001.

[49] C. Lefèvre.
'The CERN accelerator complex'.
Schema of LHC complex.
Dec. 2008.
URL: http://cds.cern.ch/record/1260465.

[50] D. Dominguez.
'3D cut of the LHC dipole'.
Accessed: 2020-01-19.
July 2014.
URL: https://home.cern/resources/image/accelerators/lhc-images-gallery.

[51] The CMS Collaboration.
'The CMS experiment at the CERN LHC'.
In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08004–S08004.
DOI: 10.1088/1748-0221/3/08/s08004.

[52] The LHCb Collaboration.
'The LHCb Detector at the LHC'.

In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08005–S08005.
DOI: 10.1088/1748-0221/3/08/s08005.

[53] The ALICE Collaboration.
'The ALICE experiment at the CERN LHC'.
In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08002–S08002.
DOI: 10.1088/1748-0221/3/08/s08002.

[54] J. Pequenao.
'Computer generated image of the whole ATLAS detector'.
Mar. 2008.
URL: http://cds.cern.ch/record/1095924.

[55] A. Airapetian et al.
ATLAS detector and physics performance: Technical Design Report, 1.
Technical Design Report ATLAS.
Geneva: CERN, 1999.
URL: http://cds.cern.ch/record/391176.

[56] A. Airapetian et al.
ATLAS detector and physics performance: Technical Design Report, 2.
Technical Design Report ATLAS.
Geneva: CERN, 1999.
URL: http://cds.cern.ch/record/391177.

[57] J. Pequenao.
'Computer generated image of the ATLAS inner detector'.
Mar. 2008.
URL: http://cds.cern.ch/record/1095926.

[58] A. La Rosa.
The ATLAS Insertable B-Layer: from construction to operation.
2016.
arXiv: 1610.01994 [physics.ins-det].

[59] J. Pequenao.
'Computer Generated image of the ATLAS calorimeter'.
Mar. 2008.
URL: http://cds.cern.ch/record/1095927.

[60] J. Pequenao.
'Computer generated image of the ATLAS Muons subsystem'.
Mar. 2008.
URL: https://cds.cern.ch/record/1095929.

[61] G. Avoni et al.
'The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS'.
In: *JINST* 13.07 (2018), P07017.
DOI: 10.1088/1748-0221/13/07/P07017.

[62] V Cindro et.al.
'The ATLAS Beam Conditions Monitor'.
In: *Journal of Instrumentation* 3.02 (Feb. 2008), P02004–P02004.
DOI: 10.1088/1748-0221/3/02/p02004.

[63] S. van der Meer.
Calibration of the effective beam height in the ISR.
Tech. rep. CERN-ISR-PO-68-31. ISR-PO-68-31.
Geneva: CERN, 1968.
URL: http://cds.cern.ch/record/296752.

[64] G. Aad et al.
Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System.
Tech. rep. CERN-LHCC-2013-018. ATLAS-TDR-023.
Final version presented to December 2013 LHCC.
Sept. 2013.
URL: https://cds.cern.ch/record/1602235.

[65] K. Bos et al.
LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005).
Technical Design Report LCG.
Geneva: CERN, 2005.
URL: https://cds.cern.ch/record/840543.

[66] G. Aad et al.
'Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics'.
In: *CERN-OPEN-2008-020* (2008).
arXiv: 0901.0512 [hep-ex].

[67] ATLAS TopReconstructionGroup.
TopCommonObjects.
ATLAS internal.
2017.

[68] T. Cornelissen et al.
'The new ATLAS track reconstruction (NEWT)'.
In: *Journal of Physics: Conference Series* 119.3 (July 2008), p. 032014.
DOI: 10.1088/1742-6596/119/3/032014.

[69] The ATLAS Collaboration.
Early Inner Detector Tracking Performance in the 2015 data at $\sqrt{s}$ = 13 TeV.
Tech. rep. ATL-PHYS-PUB-2015-051.
Geneva: CERN, Dec. 2015.
URL: http://cds.cern.ch/record/2110140.

[70] S. Boutle et al.
'Primary vertex reconstruction at the ATLAS experiment'.
In: *J. Phys. Conf. Ser.* 898.4 (2017), p. 042056.
DOI: 10.1088/1742-6596/898/4/042056.

[71] The ATLAS Collaboration.
Performance of primary vertex reconstruction in proton-proton collisions at $\sqrt{s}$ =7 TeV in the ATLAS experiment.
Tech. rep. ATLAS-CONF-2010-069.
Geneva: CERN, July 2010.
URL: https://cds.cern.ch/record/1281344.

[72] The ATLAS Collaboration.
Number of reconstructed vertices versus number of interactions in 2016 data and simulation.
Public Results.
URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2016-013/.

[73] The ATLAS Collaboration.
2016 Primary Vertex Performance Plots.
Public Results.
URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2016-007/.

[74] A. A. J. Lesage.
'Lepton and photon performance at ATLAS and CMS'.
In: *Prcoceedings of the 5. Annual LHCP* (2017).
URL: https://arxiv.org/abs/1709.02598.

[75] J. Mitrevski.
'Electron and Photon Reconstruction with the ATLAS Detector'.
In: *Nuclear and Particle Physics Proceedings* 273-275 (2016). 37th International Conference on High Energy Physics (ICHEP), pp. 2539–2541.
DOI: https://doi.org/10.1016/j.nuclphysbps.2015.09.452.

[76] The ATLAS Collaboration.
Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data.
Tech. rep. ATLAS-CONF-2016-024.
Geneva: CERN, June 2016.
URL: http://cds.cern.ch/record/2157687.

[77] The ATLAS Collaboration.
Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s}$ = 13 TeV.
Tech. rep. ATL-PHYS-PUB-2016-015.
Geneva: CERN, Aug. 2016.
URL: http://cds.cern.ch/record/2203514.

[78] W. Lampl et al.
Calorimeter Clustering Algorithms: Description and Performance.
Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003.
Geneva: CERN, Apr. 2008.

URL: https://cds.cern.ch/record/1099735.

[79] The ATLAS Collaboration.
Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung.
Tech. rep. ATLAS-CONF-2012-047.
Geneva: CERN, May 2012.
URL: http://cds.cern.ch/record/1449796.

[80] G. Aad et al.
'Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$=13 TeV'.
In: *Eur. Phys. J. C (2016)* 76 (292 Mar. 2016). arXiv:1603.05598 [hep-ex].
DOI: 10.1140/epjc/s10052-016-4120-y.

[81] The ATLAS Collaboration.
Jet energy scale and its systematic uncertainty in proton-proton collisions at sqrt(s)=7 TeV with ATLAS 2011 data.
Tech. rep. ATLAS-CONF-2013-004.
Geneva: CERN, Jan. 2013.
URL: https://cds.cern.ch/record/1509552.

[82] The ATLAS Collaboration.
Properties of Jets and Inputs to Jet Reconstruction and Calibration with the ATLAS Detector Using Proton-Proton Collisions at $\sqrt{s} = 13$ *TeV.*
Tech. rep. ATL-PHYS-PUB-2015-036.
Geneva: CERN, Aug. 2015.
URL: http://cds.cern.ch/record/2044564.

[83] G. C. Blazey et al.
'Run II Jet Physics: Proceedings of the Run II QCD and Weak Boson Physics Workshop'.
In: *High Energy Physics -Experiment (hep-ex)* (2000).
URL: https://arxiv.org/abs/hep-ex/0005012.

[84] M. Cacciari, G. P. Salam, and G. Soyez.
'The anti-k_t jet clustering algorithm'.
In: *JHEP 0804:063,2008* (2008).
URL: https://arxiv.org/abs/0802.1189.

[85] M. Aaboud et al.
'Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector'.
In: *Phys. Rev.* D96.7 (2017), p. 072002.
DOI: 10.1103/PhysRevD.96.072002.
arXiv: 1703.09665 [hep-ex].

[86] The ATLAS collaboration.

Monte Carlo Calibration and Combination of In-situ Measurements of Jet Energy Scale, Jet Energy Resolution and Jet Mass in ATLAS.
Tech. rep. ATLAS-CONF-2015-037.
Geneva: CERN, Aug. 2015.
URL: https://cds.cern.ch/record/2044941.

[87] ATLAS Collaboration.
Impact of Alternative Inputs and Grooming Methods on Large-R Jet Reconstruction in ATLAS.
Tech. rep. ATL-PHYS-PUB-2017-020.
Geneva: CERN, Dec. 2017.
URL: https://cds.cern.ch/record/2297485.

[88] The ATLAS Collaboration.
Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run.
Tech. rep. ATL-PHYS-PUB-2016-012.
Geneva: CERN, June 2016.
URL: http://cds.cern.ch/record/2160731.

[89] The ATLAS Collaboration.
Boosted hadronic top identification at ATLAS for early 13 TeV data.
Tech. rep. ATL-PHYS-PUB-2015-053.
Geneva: CERN, Dec. 2015.
URL: http://cds.cern.ch/record/2116351.

[90] J. M. Butterworth, B. E. Cox, and J. R. Forshaw.
'WW scattering at the CERN LHC'.
In: *Phys. Rev. D* 65 (9 May 2002), p. 096014.
DOI: 10.1103/PhysRevD.65.096014.

[91] J. Thaler and K. Van Tilburg.
'Maximizing boosted top identification by minimizing N-subjettiness'.
In: *Journal of High Energy Physics* 2012.2 (Feb. 2012), p. 93.
DOI: 10.1007/JHEP02(2012)093.
arXiv: 1108.2701v2 [hep-ph].

[92] Jesse Thaler and Ken Van Tilburg.
'Identifying boosted objects with N-subjettiness'.
In: *Journal of High Energy Physics* 2011.3 (Mar. 2011), p. 15.
DOI: 10.1007/JHEP03(2011)015.
arXiv: 1011.2268v3 [hep-ph].

[93] The ATLAS Collaboration.
Performance of Top Quark and W Boson Tagging in Run 2 with ATLAS.
Tech. rep. ATLAS-CONF-2017-064.
Geneva: CERN, Aug. 2017.
URL: https://cds.cern.ch/record/2281054.

[94] M. Aaboud et al.

In: *The European Physical Journal C* 78.11 (Nov. 2018).
DOI: 10.1140/epjc/s10052-018-6288-9.
arXiv: 1802.08168 [hep-ex].

[95] GEANT4 collaboration.
'Geant4—a simulation toolkit'.
In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303.
DOI: https://doi.org/10.1016/S0168-9002(03)01368-8.

[96] J. P. Ellis.
'TikZ-Feynman: Feynman diagrams with TikZ'.
In: *Computer Physics Communications* 210 (Jan. 2017), pp. 103–123.
DOI: 10.1016/j.cpc.2016.08.019.

[97] The ATLAS Collaboration.
ATLAS (2015+2016).
Public results of mutiple year collision plots.
URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[98] P. Nason.
'A New method for combining NLO QCD with shower Monte Carlo algorithms'.
In: *JHEP* 11 (2004), p. 040.
DOI: 10.1088/1126-6708/2004/11/040.
arXiv: hep-ph/0409146 [hep-ph].

[99] S. Frixione, P. Nason, and C. Oleari.
'Matching NLO QCD computations with Parton Shower simulations: the POWHEG method'.
In: *JHEP* 11 (2007), p. 070.
DOI: 10.1088/1126-6708/2007/11/070.
arXiv: 0709.2092 [hep-ph].

[100] S. Alioli et al.
'A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX'.
In: *JHEP* 06 (2010), p. 043.
DOI: 10.1007/JHEP06(2010)043.
arXiv: 1002.2581 [hep-ph].

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

[102] J. Gao et al.
'CT10 next-to-next-to-leading order global analysis of QCD'.
In: *Phys. Rev.* D89.3 (2014), p. 033009.

DOI: 10.1103/PhysRevD.89.033009.
arXiv: 1302.6246 [hep-ph].

[103] T. Sjostrand, S. Mrenna, and P. Z. Skands.
'PYTHIA 6.4 Physics and Manual'.
In: *JHEP* 05 (2006), p. 026.
DOI: 10.1088/1126-6708/2006/05/026.
arXiv: hep-ph/0603175 [hep-ph].

[104] J. Pumplin et al.
'New generation of parton distributions with uncertainties from global QCD analysis'.
In: *JHEP* 07 (2002), p. 012.
DOI: 10.1088/1126-6708/2002/07/012.
arXiv: hep-ph/0201195 [hep-ph].

[105] P. Z. Skands.
'Tuning Monte Carlo Generators: The Perugia Tunes'.
In: *Phys. Rev.* D82 (2010), p. 074018.
DOI: 10.1103/PhysRevD.82.074018.
arXiv: 1005.3457 [hep-ph].

[106] J. H. Kühn, A. Scharf, and P. Uwer.
'Weak Interactions in Top-Quark Pair Production at Hadron Colliders: An Update'.
In: *Phys. Rev.* D91.1 (2015), p. 014020.
DOI: 10.1103/PhysRevD.91.014020.
arXiv: 1305.5773 [hep-ph].

[107] G. Corcella et al.
'HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)'.
In: *Journal of High Energy Physics* 2001.01 (Jan. 2001), pp. 010–010.
DOI: 10.1088/1126-6708/2001/01/010.

[108] S. Gieseke, C. Röhr, and A. Siódmok.
'Colour reconnections in Herwig++'.
In: *The European Physical Journal C* 72.11 (Nov. 2012).
DOI: 10.1140/epjc/s10052-012-2225-5.

[109] J. Alwall et al.
'The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations'.
In: *Journal of High Energy Physics* 2014.7 (July 2014).
DOI: 10.1007/jhep07(2014)079.

[110] J. K. Behr and ohters.
A search for top-antitop resonances in the lepton+jets final state using 36.1 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV.
Tech. rep. ATL-COM-PHYS-2016-1036.
Geneva: CERN, July 2016.

URL: https://cds.cern.ch/record/2201651.

[111] T. Sjostrand, S. Mrenna, and P. Z. Skands.
'A Brief Introduction to PYTHIA 8.1'.
In: *Comput. Phys. Commun.* 178 (2008), pp. 852–867.
DOI: 10.1016/j.cpc.2008.01.036.
arXiv: 0710.3820 [hep-ph].

[112] R. D. Ball et al.
'Parton distributions with LHC data'.
In: *Nucl. Phys.* B867 (2013), pp. 244–289.
DOI: 10.1016/j.nuclphysb.2012.10.003.
arXiv: 1207.1303 [hep-ph].

[113] The ATLAS Collaboration.
ATLAS Run 1 Pythia8 tunes.
Tech. rep. ATL-PHYS-PUB-2014-021.
Geneva: CERN, Nov. 2014.
URL: https://cds.cern.ch/record/1966419.

[114] R. Bonciani et al.
'Electroweak top-quark pair production at the LHC with Z' bosons to NLO QCD in POWHEG'.
In: *JHEP* 02 (2016), p. 141.
DOI: 10.1007/JHEP02(2016)141.
arXiv: 1511.08185 [hep-ph].

[115] J. Butterworth et al.
'PDF4LHC recommendations for LHC Run II'.
In: *J. Phys.* G43 (2016), p. 023001.
DOI: 10.1088/0954-3899/43/2/023001.
arXiv: 1510.03865 [hep-ph].

[116] E. Bothmann et al.
'Event generation with Sherpa 2.2'.
In: *SciPost Physics* 7.3 (Sept. 2019).
DOI: 10.21468/scipostphys.7.3.034.

[117] E. Re.
'Single-top Wt-channel production matched with parton showers using the POWHEG method'.
In: *Eur. Phys. J.* C71 (2011), p. 1547.
DOI: 10.1140/epjc/s10052-011-1547-z.
arXiv: 1009.2450 [hep-ph].

[118] S. Alioli et al.
'NLO single-top production matched with shower in POWHEG: s- and t-channel contributions'.
In: *JHEP* 09 (2009). [Erratum: JHEP02,011(2010)], p. 111.

DOI: 10.1007/JHEP02(2010)011,10.1088/1126-6708/2009/09/111.
arXiv: 0907.4076 [hep-ph].

[119] S. Frixione and others.
'Single-top hadroproduction in association with a W boson'.
In: *JHEP* 07 (2008), p. 029.
DOI: 10.1088/1126-6708/2008/07/029.
arXiv: 0805.3067 [hep-ph].

[120] R. Frederix, E. Re, and P. Torrielli.
'Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO'.
In: *JHEP* 09 (2012), p. 130.
DOI: 10.1007/JHEP09(2012)130.
arXiv: 1207.5391 [hep-ph].

[121] P. Artoisenet et al.
'Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations'.
In: *JHEP* 03 (2013), p. 015.
DOI: 10.1007/JHEP03(2013)015.
arXiv: 1212.3460 [hep-ph].

[122] G. Aad et al.
'Measurement of the top quark pair production cross-section with ATLAS in the single lepton channel'.
In: *Physics Letters B* 711.3-4 (May 2012), pp. 244–263.
DOI: 10.1016/j.physletb.2012.03.083.

[123] The ATLAS Collaboration.
Tagging and suppression of pileup jets with the ATLAS detector.
Tech. rep. ATLAS-CONF-2014-018.
Geneva: CERN, May 2014.
URL: https://cds.cern.ch/record/1700870.

[124] M. Cacciari, Gavin Salam, and Gregory Soyez.
'FastJet user manual'.
In: *The European Physical Journal C* 72.3 (Mar. 2012), p. 1896.
DOI: 10.1140/epjc/s10052-012-1896-2.
URL: https://doi.org/10.1140/epjc/s10052-012-1896-2.

[125] The ATLAS Collaboration.
Expected performance of the ATLAS b-tagging algorithms in Run-2.
Tech. rep. ATL-PHYS-PUB-2015-022.
Geneva: CERN, July 2015.
URL: https://cds.cern.ch/record/2037697.

[126] ATLAS Flavour Tagging Group.
B-tagging benchmarks for mc15c samples, MV2c10 tagger.
ATLAS internal.

[127] The ATLAS Collaboration.
Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector.
Tech. rep. ATLAS-CONF-2015-029.
Geneva: CERN, July 2015.
URL: https://cds.cern.ch/record/2037702.

[128] T. Chwalek.
'Messung der W-Boson-Helizitätsanteile in Top-Quark-Zerfällen mit dem CDF II Experiment und Studien zu einer frühen Messung des ttbar-Wirkungsquerschnitts mit dem CMS Experiment'.
Presented 12 Feb 2010.
PhD thesis. 2010.
URL: https://cds.cern.ch/record/1416031.

[129] G. Aad et al.
'Measurement of the differential cross-section of highly boosted top quarks as a function of their transverse momentum in $\sqrt{s} = 8$ TeV proton-proton collisions using the ATLAS detector'.
In: *Phys. Rev. D* 93 (3 Feb. 2016), p. 032009.
DOI: 10.1103/PhysRevD.93.032009.

[130] G. Aad et al.
'A search for top-antitop resonances in the lepton+jets final state using 36.1 fb^{-1} of proton-proton collisions at $\sqrt{s} = 8$ TeV'.
Supporting note ATL-COM-PHYS-2016-1036 (internal).

[131] T. Heck.
'Search for $t\bar{t}$ resonances in the lepton+jets channel in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector'.
Presented 08 Nov 2016.
PhD thesis. Apr. 2016.
URL: http://cds.cern.ch/record/2235253.

[132] The ATLAS Collaboration.
Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC.
Tech. rep. ATLAS-CONF-2019-021.
Geneva: CERN, June 2019.
URL: https://cds.cern.ch/record/2677054.

[133] W. Buttinger.
Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data.
Tech. rep. ATL-COM-SOFT-2015-119.
Geneva: CERN, May 2015.
URL: https://cds.cern.ch/record/2014726.

[134] M. Aaboud et al.

'Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector'.
In: *Phys. Rev. D* 96 (7 Oct. 2017), p. 072002.
DOI: 10.1103/PhysRevD.96.072002.

[135] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, D. K. Abhayasinghe, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, and et al.
'In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector'.
In: *The European Physical Journal C* 79.2 (Feb. 2019).
DOI: 10.1140/epjc/s10052-019-6632-8.
URL: http://dx.doi.org/10.1140/epjc/s10052-019-6632-8.

[136] The ATLAS Collaboration.
'Performance of b-jet identification in the ATLAS experiment'.
In: *Journal of Instrumentation* 11.04 (Apr. 2016), P04008–P04008.
DOI: 10.1088/1748-0221/11/04/p04008.

[137] JetEtMiss group within the ATLAS Collaboration.
Jet EtMiss recommendation for ATLAS (2015+2016), internal.

[138] M. Cacciari et al.
'Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation'.
In: *Physics Letters B* 710.4-5 (Apr. 2012), pp. 612–622.
DOI: 10.1016/j.physletb.2012.03.013.

[139] M. Beneke et al.
'Hadronic top-quark pair production with NNLL threshold resummation'.
In: *Nuclear Physics B* 855.3 (Feb. 2012), pp. 695–741.
DOI: 10.1016/j.nuclphysb.2011.10.021.

[140] P. Bärnreuther, M. Czakon, and A. Mitov.
'Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t} + X$'.
In: *Physical Review Letters* 109.13 (Sept. 2012).
DOI: 10.1103/physrevlett.109.132001.

[141] M. Czakon and A. Mitov.
'NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels'.
In: *Journal of High Energy Physics* 2012.12 (Dec. 2012).
DOI: 10.1007/jhep12(2012)054.

[142] M. Czakon and A. Mitov.
'NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction'.
In: *Journal of High Energy Physics* 2013.1 (Jan. 2013).
DOI: 10.1007/jhep01(2013)080.

[143] M. Czakon and A. Mitov.
'Top++: A program for the calculation of the top-pair cross-section at hadron colliders'.
In: *Computer Physics Communications* 185.11 (Nov. 2014), pp. 2930–2938.
DOI: 10.1016/j.cpc.2014.06.021.

[144] A. D. Martin et al.
'Parton distributions for the LHC'.
In: *The European Physical Journal C* 63.2 (July 2009), pp. 189–285.
DOI: 10.1140/epjc/s10052-009-1072-5.

[145] A. D. Martin et al.
'Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections'.
In: *The European Physical Journal C* 64.4 (Oct. 2009), pp. 653–680.
DOI: 10.1140/epjc/s10052-009-1164-2.

[146] M. Czakon, D. Heymes, and A. Mitov.
'Dynamical scales for multi-TeV top-pair production at the LHC'.
In: *Journal of High Energy Physics* 2017.4 (Apr. 2017).
DOI: 10.1007/jhep04(2017)071.

[147] N. Kidonakis.
'Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-'.
In: *Physical Review D* 82.5 (Sept. 2010).
DOI: 10.1103/physrevd.82.054018.

[148] N. Kidonakis.
'Next-to-next-to-leading logarithm resummation fors-channel single top quark production'.
In: *Physical Review D* 81.5 (Mar. 2010).
DOI: 10.1103/physrevd.81.054028.

[149] N. Kidonakis.
'Next-to-next-to-leading-order collinear and soft gluon corrections fort-channel single top quark production'.
In: *Physical Review D* 83.9 (May 2011).
DOI: 10.1103/physrevd.83.091503.

[150] G. Choudalakis.
'On hypothesis testing, trials factor, hypertests and the BumpHunter'.
In: *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011.*
2011.
arXiv: 1101.0390 [physics.data-an].

[151] L. Lyons.

'Open statistical issues in particle physics'.
In: *ArXiv e-prints* (Nov. 2008).
arXiv: 0811.1663 [stat.AP].

[152] A. L. Read.
'Presentation of search results: the CL s technique'.
In: *Journal of Physics G: Nuclear and Particle Physics* 28.10 (2002), p. 2693.
URL: http://stacks.iop.org/0954-3899/28/i=10/a=313.

[153] The ATLAS Collaboration.
Summary plots from the ATLAS Exotic phsysics group.
Public results of the ATLAS Exotics group.
URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/EXOTICS/.

[154] The ATLAS Collaboration.
'Measurement of the charge asymmetry in top-quark pair production in the lepton-plus-jets final state in pp collision data at $\sqrt{s}$= 8 TeV with the ATLAS detector'.
In: *The European Physical Journal C* 76.2 (Feb. 2016), p. 87.
DOI: 10.1140/epjc/s10052-016-3910-6.

[155] J. H. Kühn and G. Rodrigo.
'Charge asymmetry of heavy quarks at hadron colliders'.
In: *Phys. Rev. D* 59 (5 Feb. 1999), p. 054017.
DOI: 10.1103/PhysRevD.59.054017.

[156] J. H. Kühn and G. Rodrigo.
'Charge asymmetries of top quarks at hadron colliders revisited'.
In: *Journal of High Energy Physics* 2012.1 (Jan. 2012), p. 63.
DOI: 10.1007/JHEP01(2012)063.

[157] J. H. Kühn and G. Rodrigo.
'Charge Asymmetry in Hadroproduction of Heavy Quarks'.
In: *Phys. Rev. Lett.* 81 (1 July 1998), pp. 49–52.
DOI: 10.1103/PhysRevLett.81.49.

[158] J. H. Kuehn and G. Rodrigo.
'Forward-backward and charge asymmetries at Tevatron and the LHC'.
In: *8th International Workshop on the CKM Unitarity Triangle (CKM 2014) Vienna, Austria, September 8-12, 2014.*
2014.
arXiv: 1411.4675 [hep-ph].

[159] O. Antuñano, J. H. Kühn, and G. Rodrigo.
'Top quarks, axigluons, and charge asymmetries at hadron colliders'.
In: *Phys. Rev. D* 77 (1 Jan. 2008), p. 014003.
DOI: 10.1103/PhysRevD.77.014003.

[160] J. L. Rosner.
'Prominent decay modes of a leptophobic Z''.

In: *Physics Letters B* 387.1 (1996), pp. 113–117.
DOI: https://doi.org/10.1016/0370-2693(96)01022-2.

[161] P. Ferrario and G. Rodrigo.
'Massive color-octet bosons and the charge asymmetries of top quarks at hadron colliders'.
In: *Phys. Rev. D* 78 (9 Nov. 2008), p. 094018.
DOI: 10.1103/PhysRevD.78.094018.

[162] The ATLAS Collaboration.
'Measurement of the top quark pair production charge asymmetry in proton-proton collisions at 7 TeV using the ATLAS detector'.
In: *arXiv e-prints* (Nov. 2013), arXiv:1311.6724.
DOI: 10.1007/JHEP02(2014)107.
arXiv: 1311.6724 [hep-ex].

[163] The ATLAS Collaboration.
'Measurement of the charge asymmetry in top-quark pair production in the lepton-plus-jets final state in pp collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector'.
In: *The European Physical Journal C* 76.2 (Feb. 2016), p. 87.
DOI: 10.1140/epjc/s10052-016-3910-6.

[164] The CMS Collaboration.
'Inclusive and differential measurements of the tt^{-} charge asymmetry in proton–proton collisions at s=7 TeV'.
In: *Physics Letters B* 717.1 (2012), pp. 129–150.
DOI: https://doi.org/10.1016/j.physletb.2012.09.028.

[165] The CMS Collaboration.
'Inclusive and differential measurements of the $t\bar{t}$ charge asymmetry in pp collisions at 8 TeV'.
In: *Physics Letters B* 757 (2016), pp. 154–179.
DOI: https://doi.org/10.1016/j.physletb.2016.03.060.

[166] The ATLAS Collaboration and The CMS Collaboration.
'Combination of inclusive and differential $t\bar{t}$ charge asymmetry measurements using ATLAS and CMS data at $\sqrt{s}$ =7 TeV and 8 TeV'.
In: *Journal of High Energy Physics* 2018.4 (Apr. 2018), p. 33.
DOI: 10.1007/JHEP04(2018)033.

[167] 'Measurement of the charge asymmetry in highly boosted top-quark pair production in 8 TeV pp collision data collected by the ATLAS experiment'.
In: *Physics Letters B* 756 (2016), pp. 52–71.
DOI: https://doi.org/10.1016/j.physletb.2016.02.055.
URL: http://www.sciencedirect.com/science/article/pii/S0370269316001519.

[168] F. Rubbo, C. Helsens, and D. Gerbaudo.
PyFBU.
GitHub directory of the implementation of the Fully Bayesian Unfolding Algorithm.

URL: https://pypi.org/project/fbu/.

[169] G. Choudalakis.
'Fully Baeysian Unfolding'.
In: *ArXiv e-prints* (Jan. 2012).
eprint: 1201.4612.

[170] O. Behnke et al.
Data Analysis in High Energy Physics : a practical guide to statistical methods.
Wiley-VCH Verlag GmbH & Co. KGaA, 2013.

[171] S. Gonzalez de la Hoz et al.
'Measurement of the charge asymmetry in boosted top quark pair production on 8 TeV *pp* collision data collected by the ATLAS experiment'.
internal note.
Geneva, Mar. 2015.

[172] P. Diaconis.
'The Markov Chain Monte Carlo Revolution'.
In: *Bull. Amer. Math. Soc. (N.S.)* 46.2 (Apr. 2009), pp. 179–205.

[173] N. Metropolis et al.
'Equation of State Calculations by Fast Computing Machines'.
In: *J. Chem. Phys.* 21 (1953).
DOI: https://doi.org/10.1063/1.1699114.

[174] A. Gelman et al.
Bayesian Data Analysis.
Chapman and Hall/CRC press, 2013.

[175] M. Aaboud et al.
'Measurement of differential cross sections and W^+/W^- cross-section ratios for W boson production in association with jets at $\sqrt{s}$=8 TeV with the ATLAS detector'.
In: *Journal of High Energy Physics* 2018.5 (May 2018).
DOI: 10.1007/jhep05(2018)077.

[176] Centre for Precision Studies in Particle Physics.
URL: http://www.precision.hep.phy.cam.ac.uk/results/.

[177] CMS Collaboration.
Measurement of the top quark forward-backward production asymmetry and the anomalous chromoelectric and chromomagnetic moments in pp collisions at $\sqrt{s} =$ 13 TeV.
2019.
arXiv: 1912.09540 [hep-ex].

[178] A. Basan et al.
Measuring the top energy asymmetry at the LHC: QCD and SMEFT interpretations.
2020.
arXiv: 2001.07225 [hep-ph].

[179] S. Berge and S. Westhoff.
'Observing the top energy asymmetry at the LHC'.
In: *Phys. Rev. D* 95 (1 Jan. 2017), p. 014035.
DOI: 10.1103/PhysRevD.95.014035.

List of Figures

List of Tables

A General: Datasets and Monte Carlo samples

A.1 Datasets

The single lepton triggers used for the individual lepton streams and the 2015 and 2016 datasets each are presented in table A.1 and A.2.

Dataset (year)	topology	single-electron trigger
2015	resolved	HLT_e24_lhmedium_L1EM20VH HLT_e60_lhmedium HLT_e120_lhloose
	boosted	HLT_e24_lhmedium_L1EM20VH HLT_e60_lhmedium, HLT_e120_lhloose
2016	resolved	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0
	boosted	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0

Table A.1: The individual single-electron triggers used on the datasets for 2015 and 2016 are listed,for the resolved and boosted topology.

Dataset (year)	topology	single-muon trigger
2015	resolved	HLT_mu20_iloose_L1MU15 HLT_mu50
	boosted	HLT_mu20_iloose_L1MU15 HLT_mu50
2016	resolved	HLT_mu26_ivarmedium HLT_mu50
	boosted	HLT_mu26_ivarmedium HLT_mu50

Table A.2: The individual single-muon triggers used on the datasets for 2015 and 2016 are listed,for the resolved and boosted topology.

A.2 Monte Carlo samples

This section contains a list of the Monte Carlo simulation samples used for the analysis for the Standard Model background expectation and the $t\bar{t}$ signal samples from various generators, including those considered for modeling uncertainty studies.

$t\bar{t}$ (Powheg+Pythia) + systematic samples

-Standard analysis-

mc15_13TeV.410000.PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad.merge.DAOD_EXOT4.e3698_s2608_s2183_r7725_r7676_p2949

mc15_13TeV.3015[28-32].PowhegPythiaEvtGen_P2012_ttbar_hdamp172p5_nonallhad_mtt_[1-5].merge.DAOD_EXOT4.e3954_s2608_r7772_r7676_p2952

-Generator comparison for parton showering (Powheg+Herwig)-

mc15_13TeV.410004.PowhegHerwigppEvtGen_UEEE5_ttbar_hdamp172p5_nonallhad.merge.DAOD_EXOT4.e3836_s2726_r7772_r7676_p2949

-Generator comparison for hard scattering (MC@NLO+Herwig) (+ PDF uncertainty)-

mc15_13TeV.410003.aMcAtNloHerwigppEvtGen_ttbar_nonallhad.merge.DAOD_EXOT4.e4441_s2726_r7772_r7676_p2949

-Initial/Final state radiation study (Powheg+Pythia)-

mc15_13TeV.410002.PowhegPythiaEvtGen_P2012radLo_ttbar_hdamp172_up_nonallhad.merge.DAOD_EXOT4.e3783_s2608_r7725_r7676_p2949

mc15_13TeV.410001.PowhegPythiaEvtGen_P2012radHi_ttbar_hdamp345_down_nonallhad.merge.DAOD_EXOT4.e3783_s2608_r7725_r7676_p2949

$t\bar{t}$ +V (MadGraph+Pythia)

mc15_13TeV.41006[6-8].MadGraphPythia8EvtGen_A14NNPDF23LO_ttW_Np[0-2].merge.DAOD_EXOT4.e4111_s2608_s2183_r7725_r7676_p2949

mc15_13TeV.41007[3-5].MadGraphPythia8EvtGen_A14NNPDF23LO_ttZnnqq_Np[0-2].merge.DAOD_EXOT4.e4631_s2726_r7725_r7676_p2949

W+jets (Sherpa)

[X]:=[CVetoBVeto,CfilterBVeto,BFilter]

mc15_13TeV.36417[0-2].Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_[X].merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.36417[3-5]Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_[X].merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.36417[6-8].Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_[X].merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.3641[79-81].Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_[X].merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364182.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV500_1000.merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364183.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV1000_E_CMS.merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.36415[6-8].Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_[X].merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.3641[59-61].Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_[X].merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.36416[2-4].Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_[X].merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.36416[5-7].Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_[X].merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364168.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV500_1000.merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364169.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV1000_E_CMS.merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.36418[4-6].Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_[X].merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.36418[7-9].Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_[X].merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.36419[0-2].Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_[X].merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.36419[3-5].Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_[X].merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364196.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV500_1000.merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

mc15_13TeV.364197.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV1000_E_CMS.merge.DAOD_EXOT4.e5340_s2726_r7772_r7676_p2949

Z+jets (Sherpa)

[X]:=[CVetoBVeto,CfilterBVeto,BFilter]

mc15_13TeV.36410[0-2].Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_[X].merge.DAOD_EXOT4.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.36410[3-5].Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_[X].merge.DAOD_EXOT4.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.36410[6-8].Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_[X].merge.DAOD_EXOT4.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.3641[09-11].Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_[X].merge.DAOD_EXOT4.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364112.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV500_1000.merge.DAOD_EXOT4.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.364113.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV1000_E_CMS.merge.DAOD_EXOT4.e5271_s2726_r7772_r7676_p2949

mc15_13TeV.36411[4-6].Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_[X].merge.DAOD_EXOT4.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.36411[7-9].Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_[X].merge.DAOD_EXOT4.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.36412[0-2].Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_[X.merge.DAOD_EXOT4.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.36412[3-5].Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_[X].merge.DAOD_EXOT4.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364126.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV500_1000.merge.DAOD_EXOT4.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.364127.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV1000_E_CMS.merge.DAOD_EXOT4.e5299_s2726_r7772_r7676_p2949

mc15_13TeV.3641[28-30].Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_[X].merge.DAOD_EXOT4.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.36413[1-3].Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_[X].merge.DAOD_EXOT4.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.36413[4-6].Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_[X].merge.DAOD_EXOT4.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.36413[7-9].Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_[X].merge.DAOD_EXOT4.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364140.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV500_1000.merge.DAOD_EXOT4.e5307_s2726_r7772_r7676_p2949

mc15_13TeV.364141.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV1000_E_CMS.merge.DAOD_EXOT4.e5307_s2726_r7772_r7676_p2949

Single top (Powheg+Pythia)

mc15_13TeV.410011.PowhegPythiaEvtGen_P2012_singletop_tchan_lept_top.merge.DAOD_EXOT4.e3824_s2608_s2183_r7725_r7676_p2952

mc15_13TeV.410012.PowhegPythiaEvtGen_P2012_singletop_tchan_lept_antitop.merge.DAOD_EXOT4.e3824_s2608_s2183_r7725_r7676_p2952

mc15_13TeV.410013.PowhegPythiaEvtGen_P2012_Wt_inclusive_top.merge.DAOD_EXOT4.e3753_s2608_s2183_r7725_r7676_p2952

mc15_13TeV.410014.PowhegPythiaEvtGen_P2012_Wt_inclusive_antitop.merge.DAOD_EXOT4.e3753_s2608_s2183_r7725_r7676_p2952

mc15_13TeV.410025.PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_top.merge.DAOD_EXOT4.e3998_s2608_s2183_r7725_r7676_p2952

mc15_13TeV.410026.PowhegPythiaEvtGen_P2012_SingleTopSchan_noAllHad_antitop.merge.DAOD_EXOT4.e3998_s2608_s2183_r7725_r7676_p2952

Diboson (Sherpa)

mc15_13TeV.361081.Sherpa_CT10_WplvWmqq.merge.DAOD_EXOT4.e3836_s2608_s2183_r7772_r7676_p2952

mc15_13TeV.361082.Sherpa_CT10_WpqqWmlv.merge.DAOD_EXOT4.e3836_s2608_s2183_r7772_r7676_p2952

mc15_13TeV.361083.Sherpa_CT10_WlvZqq.merge.DAOD_EXOT4.e3836_s2608_s2183_r7772_r7676_p2952

mc15_13TeV.361084.Sherpa_CT10_WqqZll.merge.DAOD_EXOT4.e3836_s2608_s2183_r7772_r7676_p2952

mc15_13TeV.361086.Sherpa_CT10_ZqqZll.merge.DAOD_EXOT4.e3926_s2608_s2183_r7772_r7676_p2952

B General: Definition of asymmetric and symmetric MC samples for W+jets charge normalization factor

In this appendix asymmetry value A_C is estimated for each background sample, as additional material for section 7.4.1.

B.1 μ+jets channel

2ex μ+jets channel	$\mathbf{A_c}$	$\mathbf{\Delta A_c}$	**defined as asymmetric**
$W + bb$	11.5%	0.3%	✓
$W + cc$	10.3%	0.7%	✓
$W + c$	5.9%	0.5%	✓
W+lf	10.8%	0.7%	✓
single top (s,t)	21.7%	0.2%	✓
$t\bar{t}$ +V	5.3%	0.6%	✓
Diboson	3.6%	0.4%	✓
QCD	1.5%	0.3%	
single top (Wt)	0.2%	0.3%	
$t\bar{t}$	-0.3%	0.1%	
Z+jets	2.5%	0.7%	

Table B.1: The asymmetry A_c of all MC samples is listed for events with exactly 2 jets in the μ+jets channel in the pretag region, including the statistical uncertainty.

B.2 e+jets channel

2ex e+jets channel	$\mathbf{A_c}$	$\mathbf{\Delta A_c}$	**defined as asymmetric**
$W + bb$	10.1%	0.3%	✓
$W + cc$	8.8%	0.6%	✓
$W + c$	4.7%	0.4%	✓
W+lf	10.1%	0.7%	✓
single top (s,t)	20.6%	0.2%	✓
$t\bar{t}$ +V	4.7%	0.6%	✓
Diboson	3.1%	1.7%	✓
QCD	0.5%	0.1%	
single top (Wt)	-0.3%	0.3%	
$t\bar{t}$	-0.1%	0.1%	
Z+jets	0.3%	0.4%	

Table B.2: The asymmetry A_c of all MC samples is listed for events with exactly 2 jets in the e+jets channel in the pretag region, including the statistical uncertainty.

4ex e+jets channel	$\mathbf{A_c}$	$\mathbf{\Delta A_c}$	**defined as asymmetric**
$W + bb$	11.0%	0.4%	✓
$W + cc$	10.5%	0.7%	✓
$W + c$	5.1%	0.7%	✓
W+lf	12.3%	0.9%	✓
single top (s,t)	23.1%	0.5%	✓
$t\bar{t}$ +V	5.2%	0.3%	✓
Diboson	-0.3%	0.7%	✓
QCD	1.1%	0.4%	
single top (Wt)	0.4%	0.3%	
$t\bar{t}$	0.0%	0.1%	
Z+jets	-0.1%	0.6%	

Table B.3: The asymmetry A_c of all MC samples is listed for events with exactly 4 jets in the e+jets channel in the pretag region, including the statistical uncertainty.

C General: Closure test for W+jets estimation

For the closure tests an ideal environment is defined, with symmetric yields for $t\bar{t}$, Z+jets, single top (Wt-channel) and the multijet background, resulting in a perfect agreement for the Asimov pseudo-dataset, see Table C.1.

2ex bin	$\mathbf{K}_{bb}$,$\mathbf{K}_{cc}$	$\mathbf{K}_{c}$	$\mathbf{K}_{lf}$	**QCD**	**CA**
expected	1.0	1.00	1.00	1.00	1.00
HFSF e-channel	1.00±5.6%	1.00	1.00±1.4%	1.00	1.00±0.5%
HFSF μ-channel	1.00±4.6%	1.00	1.00±1.0%	1.00	1.00±0.4%

Table C.1: Charge asymmetry normalization CA and heavy flavor scale factors K_i, i ∈ [bb,cc,c,lf] for an Asimov pseudo-dataset with symmetrized QCD, Z+jets, $t\bar{t}$ and single top (Wt-channel) contributions, scaled to the integrated Luminosity of 36.1 fb^{-1}, statistical uncertainties on data only, expected and heavy flavor scale factors with QCD scale factor in control region and CA normalization.

As the charge asymmetry normalization factor CA is accounting for the normalization between the pseudo data set and the expected yields, employing an increase of 10% to the total amount of W+jets in the pseudo data is expected to be corrected by a normalization factor of $CA = 1.10$ for each jet multiplicity bin, constant heavy flavor scale factors with K_i=1.0, i∈ [bb,cc,c,lf] and a constant multijet estimate, see Table C.2.

2ex bin	$\mathbf{K}_{bb}$,$\mathbf{K}_{cc}$	$\mathbf{K}_{c}$	$\mathbf{K}_{lf}$	**QCD**	**CA**
expected	1.0	1.00	1.00	1.00	1.10
HFSF e-channel	1.00±5.2%	1.00	1.00±1.3%	1.00	1.10±0.5%
HFSF μ-channel	1.00±4.2%	1.00	1.00±1.0%	1.00	1.10±0.4%

Table C.2: Charge asymmetry normalization CA and heavy flavor scale factors K_i, i ∈ [bb,cc,c,lf] for a pseudo data set with 10% increase in the total W+jets yields, symmetric expected backgrounds are symmetrized, scaled to the integrated luminosity of 36.1 fb^{-1}, statistical uncertainties on data only

The variation of the W+jets flavor fractions utilizes a normalization factor for each jet multiplicity bin and each W+jets flavor, see Table C.3 and C.4, to obtain a constant total amount of W+jets in the pretag region. This approach allows to measure the reliability of the method to extract the heavy flavor scale factors by expecting the overall normalization to be 1.0.

As the flavor scale factor for $W + c$ is fixed to 1.0 in the control region, the reproduction of the initial scale factors with the W+jets estimation method for the variation of one of

e+jets	**W+bb,cc (+40%)**		**W+bb,cc (-40%)**		**W+lf (+15%)**		**W+lf (-15%)**	
bin	$\boldsymbol{F_{bb}}$**,**$\boldsymbol{F_{cc}}$	F_c,F_{lf}	$\boldsymbol{F_{bb}}$**,**$\boldsymbol{F_{cc}}$	F_c,F_{lf}	$\boldsymbol{F_{lf}}$	F_{bb},F_{cc},F_c	$\boldsymbol{F_{lf}}$	F_{bb},F_{cc},F_c
2ex	1.32	0.94	0.64	1.07	1.05	0.91	0.94	1.10
4ex	1.26	0.90	0.68	1.13	1.07	0.93	0.92	1.08
5in	1.23	0.88	0.70	1.17	1.08	0.94	0.91	1.07

Table C.3: Initial scale factors for W+jets components used for the creation of the different pseudo datasets for the e-channel, including the chosen variation of one fraction (***highlighted***) and the normalization of the total pretag yields applied to all W+jets contributions, for pretag and tag region

μ+jets	**W+bb,cc+40%**		**W+bb,cc-40%**		**W+lf+15%**		**W+lf-15%**	
bin	$\boldsymbol{F_{bb}}$**,**$\boldsymbol{F_{cc}}$	F_c,F_{lf}	$\boldsymbol{F_{bb}}$**,**$\boldsymbol{F_{cc}}$	F_c,F_{lf}	$\boldsymbol{F_{lf}}$	F_{bb},F_{cc},F_c	$\boldsymbol{F_{lf}}$	F_{bb},F_{cc},F_c
2ex	1.32	0.94	0.64	1.06	1.05	0.91	0.94	1.11
4ex	1.27	0.90	0.67	1.12	1.07	0.93	0.92	1.09
5in	1.23	0.88	0.70	1.16	1.08	0.94	0.91	1.07

Table C.4: Initial scale factors for W+jets components used for the creation of the different pseudo data sets for μ-channel, including the chosen variation of one fraction (***highlighted***) and the normalization of the total pretag yields applied to all W+jets contributions, for pretag and tag region

the W+jets contributions $W + bb$ and $W + cc$ or W+lf is expected to deviate regarding the flavor scale factor for $W + c$. In the control region the difference of the amount of $W + c$ in pseudo data compared to the Monte Carlo sample is compensated with the multijet estimate and a variation of the total W+jets yields, see Table C.5 for a 40% up variation of $W + bb$ and $W + cc$. Taking only the statistical uncertainties of the pseudo data set into account the differences are covered for the 40% variation of $W + bb$ and $W + cc$. The results are shown for the electron channel only, since the procedure is identical for the muon channel

	W+bb , W+cc		**W+c**		**W+lf**		**QCD**	**CA**
	exp.	**est.**	**exp.**	**est.**	**exp.**	**est.**	**est.**	**est.**
2ex	1.32	1.32±4.4%	0.94	1.01	0.94	0.93±1.7%	0.98	1.01±0.5%
4ex	1.26	1.26±4.4%	0.90	0.96	0.90	0.89±2.6%	-	1.01±2.6%
5in	1.23	1.23±4.4%	0.88	0.94	0.88	0.86±2.8%	-	1.01±2.8%

Table C.5: Closure test for 40% up variation of $W + bb$,$W + cc$ with expected, F_i, and estimated scale factors, $K_i \times$ CA, i ∈ [bb,cc,c,lf], for the e+jets-channel; multijet and CA are expected to be 1.0, statistical uncertainty on data only, symmetric expected backgrounds are symmetrized.

To confirm the precise extraction of the applied correction factors in the varied pseudo

data sets for all W+jets contributions, the fixed value of the flavor scale factor for $W + c$, K_c, in the 2 exclusive jet bin is modified to match the expected scale factor. As a result a perfect closure can be reached in all jet multiplicity bins for all W+jets components. The comparison of expected scale factor F_i and the estimated scale factor, which is the product of the flavor scale factor K_i, i $\in$ [bb,cc,c,lf] and the overall normalization CA, for the different W+jets contributions is given for the 2 exclusive jet bin for the individual flavor variations in the tables C.6, C.7, C.8 and C.9. The extracted scale factors are presented for the nominal approach and the flavor scale factor K_c fixed to the expected value instead of 1.0 in the control region.

$W + bb$,cc (+40%)	$\mathbf{F_{bb}}$,$\mathbf{F_{cc}}$	$\mathbf{F_c}$	$\mathbf{F_{lf}}$	**QCD**	**CA**
expected	1.32	0.94	0.94	1.00	1.00
$\mathbf{K_c := 0.94}$	1.32 ±4.4%	0.94	0.94 ±1.7%	1.00	1.00±0.5%
$\mathbf{K_c := 1.00}$	1.32 ±4.4%	1.01	0.93 ±1.7%	0.98	1.01±0.5%

Table C.6: $W + bb$, $W + cc$ (+40%): Comparison of expected W+jets scale factors and the obtained scale factors F_i with $F_i = K_i \times CA$, including the CA normalization scale factor, and the multijet correction factor for different fixed values of F_c; the pseuda data set contains an up variation of the $W + bb$ and $W + cc$ contribution by 40%, statistical uncertainty on data only, symmetric expected backgrounds are symmetrized

$W + bb$,cc (-40%)	$\mathbf{F_{bb}}$,$\mathbf{F_{cc}}$	$\mathbf{F_c}$	$\mathbf{F_{lf}}$	**QCD**	**CA**
expected	0.64	1.07	1.07	1.00	1.00
$\mathbf{K_c := 1.07}$	0.64±8.6%	1.07	1.07±1.4%	1.00	1.00±0.5%
$\mathbf{K_c := 1.00}$	0.64±8.7%	0.99	1.08±1.4%	1.02	0.99±0.5%

Table C.7: $W + bb$, $W + cc$ (-40%): Comparison of expected W+jets scale factors and the obtained scale factors F_i with $F_i = K_i \times CA$, including the CA normalization scale factor, and the multijet correction factor for different fixed values of F_c; the pseudo- ataset contains a down variation of the $W + bb$ and $W + cc$ contribution by 40%, statistical uncertainty on data only, symmetric expected backgrounds are symmetrized

W+lf (+15%)	$\mathbf{F_{bb}}$,$\mathbf{F_{cc}}$	$\mathbf{F_c}$	$\mathbf{F_{lf}}$	**QCD**	**CA**
expected	0.91	0.91	1.05	1.00	1.00
$\mathbf{F_c := 0.91}$	0.91±6.2%	0.91	1.05±1.5%	1.00	1.00±0.5%
$\mathbf{F_c := 1.00}$	0.92±6.1%	1.01	1.05±1.5%	0.97	1.01±0.5%

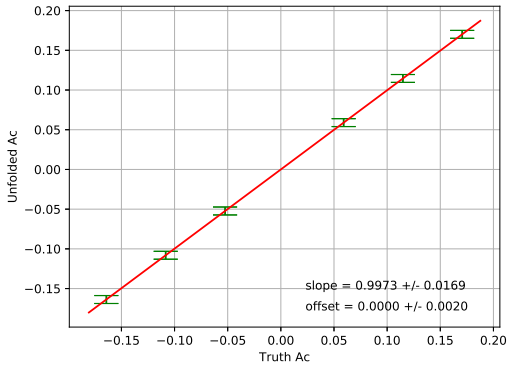
Table C.8: W+lf (+15%): Comparison of expected W+jets scale factors and the obtained scale factors F_i with $F_i = K_i \times CA$, including the CA normalization scale factor, and the multijet correction factor for different fixed values of F_c; the pseudo-dataset contains an up variation of the W+lf contribution by 15%, statistical uncertainty on data only, symmetric expected backgrounds are symmetrized

$W+lf$ (-15%)	$\mathbf{F_{bb}}$,$\mathbf{F_{cc}}$	$\mathbf{F_c}$	$\mathbf{F_{lf}}$	**QCD**	**CA**
expected	1.10	1.10	0.94	1.00	1.00
$\mathbf{F_c := 1.10}$	1.10±5.2%	1.10	0.94±1.6%	1.00	1.00±0.5%
$\mathbf{F_c := 1.00}$	1.10±5.2%	0.99	0.96±1.6%	1.04	0.99±0.5%

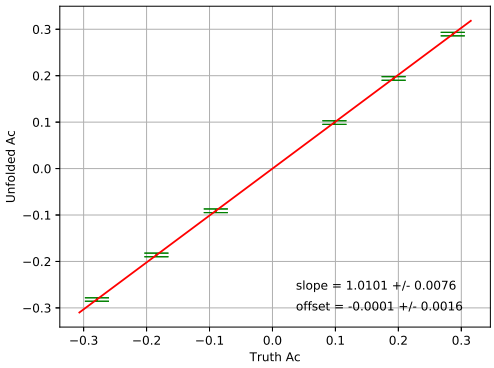
Table C.9: W+lf (-15%): Comparison of expected W+jets scale factors and the obtained scale factors F_i with $F_i = K_i \times CA$, including the CA normalization scale factor, and the multijet correction factor for different fixed values of F_c; the pseudo-dataset contains a down variation of the W+lf contribution by 15%, statistical uncertainty on data only, symmetric expected backgrounds are symmetrized

D Charge asymmetry: Linearity tests

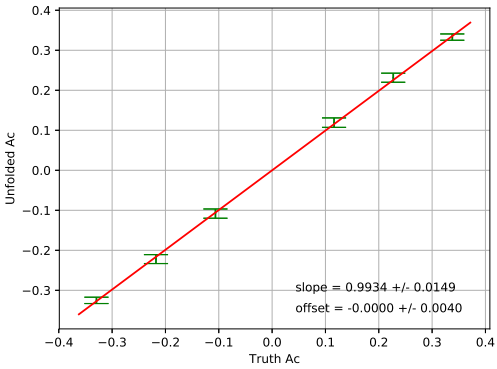
The linearity plots of truth and reconstructed A_c for the inclusive ℓ+jets channel are presented in figure D.2 and for the combined resolved+boosted topology for the ℓ+jets channel for the individual mass bins in figure D.1.
The results for the slope and offset parameters for the differential setup in resolved and boosted topology separately, as well as for the individual lepton flavor selections are listed in table D.1.



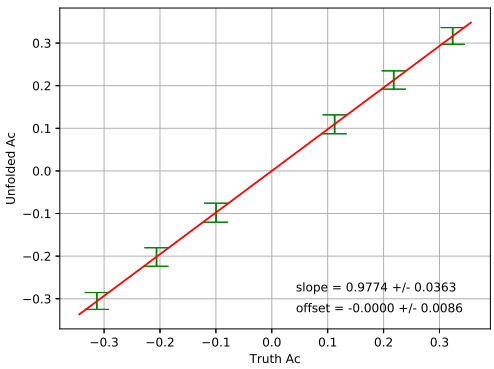
(a) $m_{t\bar{t}}$< 500 GeV



(b) 500 GeV <$m_{t\bar{t}}$< 750 GeV



(c) 750 GeV <$m_{t\bar{t}}$< 1250 GeV

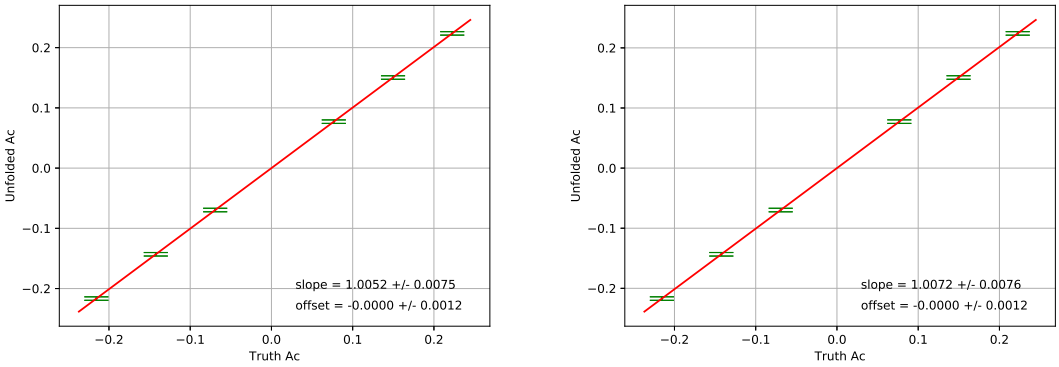


(d) 1250 GeV <$m_{t\bar{t}}$

Figure D.1: Linearity of reconstructed and truth A_C values for the differential measurement of $\Delta|y|$ in $t\bar{t}$ mass bins for the combined resolved+boosted signal region in ℓ+jets

	$m_{t\bar{t}}$<500 GeV		500 GeV<$m_{t\bar{t}}$<750 GeV		750 GeV<$m_{t\bar{t}}$<1250 GeV		1250 GeV<$m_{t\bar{t}}$	
resolved	**slope**	**offset**	**slope**	**offset**	**slope**	**offset**	**slope**	**offset**
e+jets	1.002±0.024	±0.003	0.995±0.011	±0.002	1.053±0.024	±0.006	0.919±0.102	0.004±0.024
μ+jets	1.001±0.024	±0.003	1.002±0.011	±0.002	1.051±0.024	±0.006	0.896±0.104	0.002±0.025
ℓ+jets	1.001±0.017	±0.002	0.998±0.008	±0.002	1.059±0.017	±0.005	0.940±0.074	0.001±0.018
boosted	**slope**	**offset**	**slope**	**offset**	**slope**	**offset**	**slope**	**offset**
e+jets	0.945±0.776	-0.011±0.095	1.051±0.156	-0.001±0.035	0.720±0.083	0.002±0.021	0.997±0.060	-0.001±0.015
μ+jets	0.732±0.847	0.006±0.102	1.000±0.177	-0.002±0.039	0.786±0.070	0.003±0.018	0.999±0.053	-0.001±0.013
ℓ+jets	0.780±0.822	-0.002±0.100	1.138±0.125	-0.004±0.028	0.806±0.060	±0.016	1.004±0.041	±0.010
combined	**slope**	**offset**	**slope**	**offset**	**slope**	**offset**	**slope**	**offset**
e+jets	0.997±0.024	±0.003	1.009±0.011	±0.002	0.981±0.021	0.001±0.006	0.971±0.052	±0.013
μ+jets	0.997±0.024	±0.003	1.013±0.011	±0.002	0.987±0.020	±0.005	0.974±0.047	±0.011
ℓ+jets	0.997±0.017	±0.002	1.010±0.008	±0.002	0.993±0.015	±0.004	0.977±0.036	±40.009

Table D.1: Overview of results of linearity check for differential topology for e+jets, μ+jets and the combined channel. Not listed offsets are below 0.001.



(a) e+jets **(b)** μ+jets

Figure D.2: Linearity of reconstructed and truth A_C values for the inclusive measurement of $\Delta|y|$ in the resolved e+jets (a) and μ+jets (b) channel.

E Charge asymmetry: Closure test for W+jets estimation in full Bayesian unfolding

The following appendix provides additional material for section . A continuation of the evaluation is presented for the resolved μ+jets and the resolved ℓ+jets channel.

E.1 Resolved μ+jets channel

In the considered channel, μ+jets in the resolved topology, the scale factors for the marginalization of the W+jets background extracted from the nuisance parameters, SF_{NP} are compared to the equivalent scale factors SF_{HFSF} applied to the pseudo-dataset. The results are presented in table E.1 and with restricting the variation of the non-W MC backgrounds, scenario 2, an agreement of the SF_{HFSF} and SF_{NP} within their uncertainties is reached. For the full background variation in the first setup the differences between SF_{HFSF} and SF_{NP} within a single flavor are assumed to be based on the correlations of W+jets with other MC backgrounds.

	SF_{HFSF}	Δ in \$	SF_{NP} *scen. 1*	Δ in %	SF_{NP} *scen. 2*	Δ in %	SF_{NP} *scen. 3*	Δ in %
W+bb,cc	1.24	0.14	1.13	0.05	1.26	0.17	1.25	0.11
W+c	0.81	0.28	0.94	0.40	0.76	0.32	0.77	0.32
W+lf	0.73	0.04	0.73	0.07	0.74	0.07	0.74	0.04

Table E.1: Closure test for resolved μ+jets channel: uncertainties of SF_{HFSF} are combined uncertainties on original scale factors, uncertainties of SF_{NP} are the RMS of the NP times σ=0.5. Agreement between SF_{HFSF} and SF_{NP} is reached within Δ_{NP} after reduction of background variation to W+jets and multi-jet background.

E.2 Resolved ℓ+jets

While the closure test achieves the expected results for the individual resolved lepton channels, for the combination of those a significant different set of scale factors is preferred within the nuisance parameter estimation to achieve an agreement between the Monte Carlo samples and the scaled pseudo data set.
Since the externally estimated scale factors are individual for the e+jets and μ+jets distribution in $\Delta|y|$, an overall scale factor is not expected to extract the combined scale factors from the external approach based on the total yields.

	SF_{HFSF}	Δ	SF_{NP} *scen. 1*	Δ	SF_{NP} *scen. 2*	Δ	SF_{NP} *scen. 3*	Δ
W+bb,cc	1.19	0.13	0.94	0.15	1.15	0.12	1.40	0.09
W+c	0.84	0.29	1.09	0.33	0.67	0.28	0.46	0.25
W+lf	0.78	0.05	0.77	0.04	0.80	0.04	0.75	0.03

Table E.2: Closure test for resolved ℓ+jets channel: uncertainties of SF_{HFSF} are combined uncertainties on original scale factors, uncertainties of SF_{NP} are the RMS of the NP times σ=0.5. Agreement between SF_{HFSF} and SF_{NP} is not expected.

F Charge asymmetry: Comparison of W+jets background estimation

The following appendix provides additional material to section 10.6. The equivalent of table 10.5 ($\ell + jets$ channel) for the μ+jets channel, see table F.1, and the e+jets, see table F.2, are presented.

	HFSF applied			**w/o HFSF applied**		
	prior	$\mathbf{SF_{NP}}$	$\mathbf{\Delta\ SF_{NP}}$	**prior**	$\mathbf{SF_{NP}}$	$\mathbf{\Delta\ SF_{NP}}$
W + bb,cc	0.11	1.00	0.09	0.50	1.02	0.28
W + c	0.35	1.00	0.26	0.50	1.01	0.43
W + lf	0.06	1.00	0.05	0.50	0.99	0.10

Table F.1: Overview over the scale factors, applied to the MC W+jets sample within the in-situ marginalization with and without HFSF application to the W+jets MC sample. MC samples compared to Asimov pseudo-dataset in the resolved μ+jets channel

	HFSF applied			**w/o HFSF applied**		
	prior	$\mathbf{SF_{marg.}}$	$\mathbf{\Delta\ SF_{marg.}}$	**prior**	$\mathbf{SF_{marg.}}$	$\mathbf{\Delta\ SF_{marg.}}$
W + bb,cc	0.11	1.00	0.10	0.50	1.00	0.30
W + c	0.35	1.00	0.26	0.50	1.03	0.43
W + lf	0.06	1.00	0.05	0.50	0.99	0.12

Table F.2: Overview over the scale factors, applied to the MC W+jets sample within the in-situ marginalization with and without HFSF application to the W+jets MC sample. MC samples compared to Asimov pseudo-dataset in the resolved e+jets channel

G Charge asymmetry: Estimation of top modeling uncertainties

The full input for the estimation of the modeling uncertainties is listed here. For the radiation uncertainty the median of radiation low and radiation high for (unfolded - truth) is used for the comparison to the nominal sample. Parton showering compares nominal (Powheg+Pythia) with (Powheg+Herwig) to estimate the effect of the choice of the parton showering generator. While the hard scattering uncertainty is estimated not with the nominal sample, but between Powheg+Herwig and MC@NLO+Herwig, to derive the uncertainty based on the choice of the hard scattering generator.
The uncertainty Δ_{syst} is calculated via equation 10.15, $\Delta\Delta_{syst}$ via equation 10.16.

	inclusive: resolved ℓ+jets							
	truth			**unfolded**			**syst. uncert.**	
model	$\boldsymbol{A_c}$	$\boldsymbol{\Delta A_c}$	$\boldsymbol{SF}$	$\boldsymbol{A_c}$	$\boldsymbol{\Delta A_c}$	$\boldsymbol{SF}$	$\boldsymbol{\Delta_{syst}}$	$\boldsymbol{\Delta\Delta_{syst.}}$
Nominal	0.0038	0.0001	2.91	0.0038	0.0021	2.54	-	-
Rad. low	0.0029	0.0003	0.97	0.0065	0.0020	0.79	0.0033	0.0020
Rad. high	0.0036	0.0003	1.22	0.0066	0.0021	1.00		
Parton shower	0.0032	0.0002	1.03	0.0048	0.0021	1.47	0.0015	0.0021
Hard scattering	0.0045	0.0003	0.60	0.0071	0.0021	0.48	0.0010	0.0035

Table G.1: Input for $t\bar{t}$ modeling uncertainty estimation for inclusive A_c measurement in the resolved ℓ+jets channel

	differential for $m_{t\bar{t}}$>500GeV: resolved ℓ+jets							
	truth			unfolded			syst. uncert.	
model	A_c	ΔA_c	SF	A_c	ΔA_c	SF	Δ_{syst}	$\Delta\Delta_{syst.}$
Nominal	0.0033	0.0001	2.91	0.0033	0.0050	2.54	-	-
Rad. low	0.0023	0.0004	0.97	0.0114	0.0049	0.79	0.0087	0.0049
Rad. high	0.0028	0.0004	1.22	0.0111	0.0050	1.00		
Parton shower	0.0030	0.0004	0.60	0.0144	0.0047	0.48	0.0081	0.0049
Hard scattering	0.0018	0.0003	1.03	0.0098	0.0045	1.47	0.0033	0.0078

	differential for 500 GeV<$m_{t\bar{t}}$<750GeV: resolved ℓ+jets							
	truth			unfolded			syst. uncert.	
model	A_c	ΔA_c	SF	A_c	ΔA_c	SF	Δ_{syst}	$\Delta\Delta_{syst.}$
Nominal	0.0043	0.0001	2.91	0.0043	0.0040	2.54	-	-
Rad. low	0.0037	0.0005	0.97	0.0035	0.0038	0.79	0.0014	0.0040
Rad. high	0.0047	0.0005	1.22	0.0077	0.0043	1.00		
Parton shower	0.0059	0.0006	0.60	0.0115	0.0046	0.48	0.0020	0.0046
Hard scattering	0.0050	0.0004	1.03	0.0070	0.0047	1.47	0.0036	0.0077

	differential for 750 GeV<$m_{t\bar{t}}$<1250GeV: resolved ℓ+jets							
	truth			unfolded			syst. uncert.	
model	A_c	ΔA_c	SF	A_c	ΔA_c	SF	Δ_{syst}	$\Delta\Delta_{syst.}$
Nominal	0.0050	0.0003	2.91	0.0049	0.0129	2.54	-	-
Rad. low	0.0029	0.0011	0.97	0.0235	0.0125	0.79	-0.0005	0.0126
Rad. high	0.0045	0.0009	1.22	-0.0173	0.0134	1.00		
Parton shower	0.0088	0.0011	0.60	0.0005	0.0148	0.48	-0.0203	0.0144
Hard scattering	0.0063	0.0008	1.03	-0.0141	0.0145	1.47	0.0120	0.0246

	differential for $m_{t\bar{t}}$>1250GeV: resolved ℓ+jets							
	truth			unfolded			syst. uncert.	
model	A_c	ΔA_c	SF	A_c	ΔA_c	SF	Δ_{syst}	$\Delta\Delta_{syst.}$
Nominal	0.0065	0.0013	2.91	0.0021	0.0472	2.54	-	-
Rad. low	0.0047	0.0034	0.97	-0.0185	0.0390	0.79	0.0379	0.0453
Rad. high	0.0058	0.0029	1.22	0.0960	0.0526	1.00		
Parton shower	0.0122	0.0036	0.60	0.0983	0.0522	0.48	0.0888	0.0297
Hard scattering	0.0092	0.0026	1.03	0.0980	0.0000	1.47	-0.0027	0.0758

Table G.2: Input for $t\bar{t}$ modeling uncertainty estimation for differential A_c measurement in four $t\bar{t}$ mass bins in the resolved ℓ+jets channel

	differential for $m_{t\bar{t}}$>500GeV: boosted ℓ+jets							
	truth			unfolded			syst. uncert.	
model	A_c	ΔA_c	**SF**	A_c	ΔA_c	**SF**	Δ_{syst}	$\Delta\Delta_{syst.}$
Nominal	0.0033	0.0001	2.91	-0.0165	0.2502	2.54	-	-
Rad. low	0.0023	0.0004	0.97	0.2109	0.2539	0.79	0.1301	0.2249
Rad. high	0.0028	0.0004	1.22	0.0147	0.1503	1.00		
Parton shower	0.0030	0.0004	0.60	0.3674	0.1854	0.48	0.0559	0.2361
Hard scattering	0.0018	0.0003	1.03	0.0379	0.2140	1.47	0.3283	0.3214

	differential for 500 GeV<$m_{t\bar{t}}$<750GeV: boosted ℓ+jets							
	truth			unfolded			syst. uncert.	
model	A_c	ΔA_c	**SF**	A_c	ΔA_c	**SF**	Δ_{syst}	$\Delta\Delta_{syst.}$
Nominal	0.0043	0.0001	2.91	0.0060	0.0919	2.54	-	-
Rad. low	0.0037	0.0005	0.97	0.0108	0.0852	0.79	-0.0439	0.0847
Rad. high	0.0047	0.0005	1.22	-0.0868	0.0793	1.00		
Parton shower	0.0059	0.0006	0.60	0.0663	0.0549	0.48	-0.0048	0.0718
Hard scattering	0.0050	0.0004	1.03	0.0020	0.0520	1.47	0.0635	0.0904

	differential for 750 GeV<$m_{t\bar{t}}$<1250GeV: boosted ℓ+jets							
	truth			unfolded			syst. uncert.	
model	A_c	ΔA_c	**SF**	A_c	ΔA_c	**SF**	Δ_{syst}	$\Delta\Delta_{syst.}$
Nominal	0.0050	0.0003	2.91	0.0049	0.0501	2.54	-	-
Rad. low	0.0029	0.0011	0.97	-0.0155	0.0452	0.79	0.0032	0.0457
Rad. high	0.0045	0.0009	1.22	0.0291	0.0428	1.00		
Parton shower	0.0088	0.0011	0.60	-0.0417	0.0500	0.48	0.0456	0.0483
Hard scattering	0.0063	0.0008	1.03	0.0517	0.0445	1.47	-0.0960	0.0813

	differential TCA for $m_{t\bar{t}}$>1250GeV : boosted ℓ+jets							
	truth			unfolded			syst. uncert.	
model	A_c	ΔA_c	**SF**	A_c	ΔA_c	**SF**	Δ_{syst}	$\Delta\Delta_{syst.}$
Nominal	0.0065	0.0013	2.91	0.0063	0.0261	2.54	-	-
Rad. low	0.0047	0.0034	0.97	0.0095	0.0245	0.79	0.0143	0.0248
Rad. high	0.0058	0.0029	1.22	0.0289	0.0246	1.00		
Parton shower	0.0122	0.0036	0.60	-0.0174	0.0236	0.48	0.0271	0.0268
Hard scattering	0.0092	0.0026	1.03	0.0360	0.0254	1.47	-0.0564	0.0403

Table G.3: Input for $t\bar{t}$ modeling uncertainty estimation for differential A_c measurement in four $t\bar{t}$ mass bins in the boosted ℓ+jets channel

H Charge asymmetry: NP of differential measurement

In this appendix additional material for section 10.9 is provided.

H.1 Evaluation with an Asimov pseudo-dataset

The summary of nuisance parameters for the top anti-top production charge asymmetry measurement method check with an Asimov pseudo-dataset is presented in figure H.1 (resolved) and H.2 (boosted) and in table H.1.

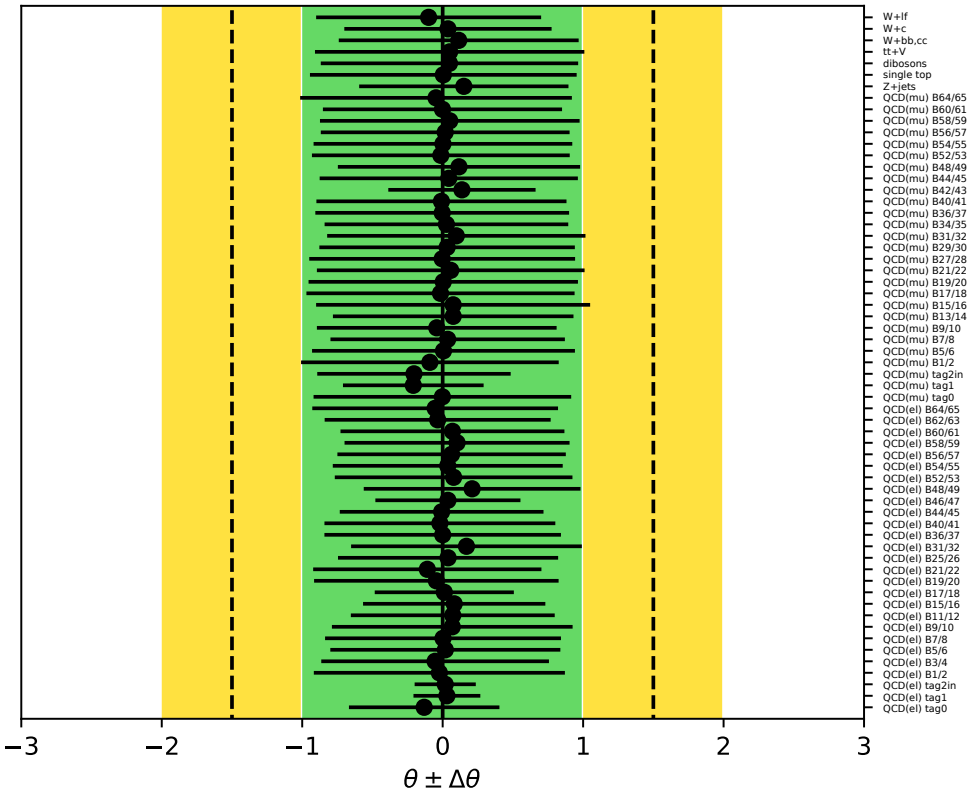


Figure H.1: Overview of nuisance parameters for differential measurement in the combined lepton+jets channel with an Asimov pseudo-dataset. The first three mass bins, up until $m_{t\bar{t}}$ ≥1250 GeV, are estimated within the here presented resolved topology. The W+jets background is scaled with heavy flavor scale factors and within their uncertainties estimated in-situ.

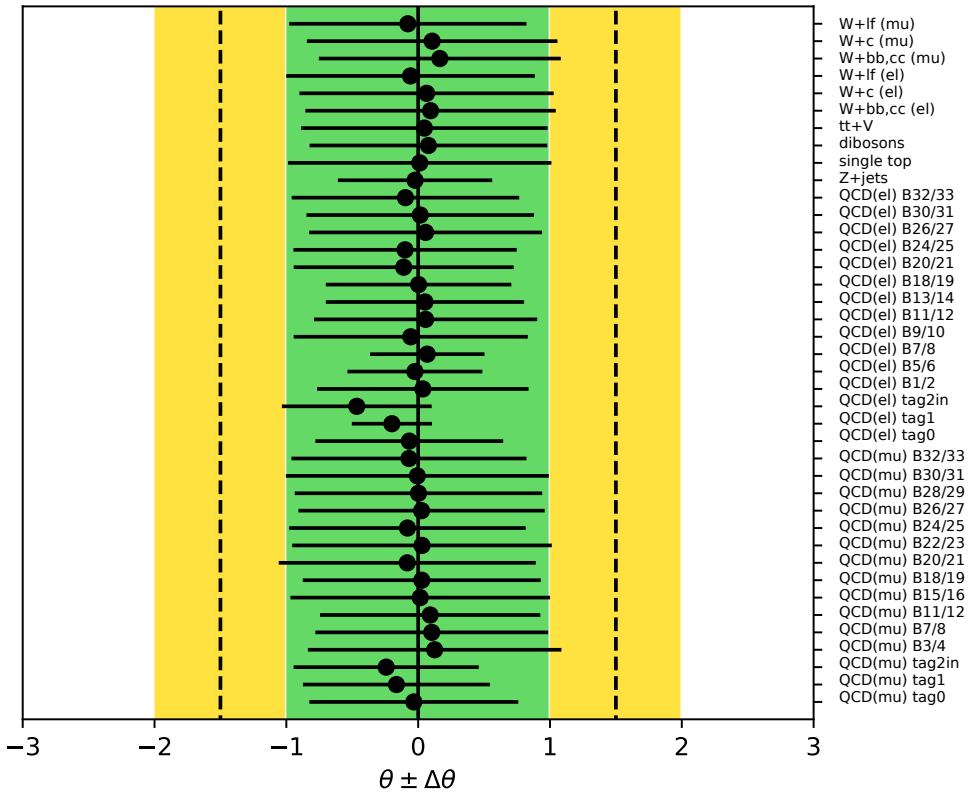


Figure H.2: Overview of nuisance parameters for differential measurement in the combined lepton+jets channel with an Asimov pseudo-dataset. The fourth mass bin, for $m_{t\bar{t}}$>1250 GeV, is measured within the boosted topology presented in this figure. The W+jets background is scaled with heavy flavor scale factors and within their uncertainties estimated in-situ.

		resolved (M1-M3)		boosted (M4)	
source	**prior**	**NP**	**posterior width**	**NP**	**posterior width**
QCD (e), 0-btag	0.5	-0.13	0.54	-0.07	0.71
QCD (e), 1-btag	0.5	0.03	0.24	-0.20	0.30
QCD (e), 2-btags	0.5	0.02	0.22	-0.47	0.57
QCD (μ), 0-btags	0.5	0.0*	0.92	-0.03	0.79
QCD (μ), 1-btag	0.5	-0.21	0.50	-0.16	0.71
QCD (μ), 2-btags	0.5	-0.20	0.69	-0.24	0.70
Z+jets	0.5	0.15	0.75	-0.02	0.59
single top	0.05	0.0*	0.95	0.01	1.00
VV	0.5	0.05	0.92	0.08	0.90
$t\bar{t}$ +V	0.5	0.05	0.96	0.05	0.93
$W + bb$,cc (ℓ)	0.23	0.11	0.85	-	-
$W + c$ (ℓ)	0.40	0.04	0.74	-	-
$W + lf$ (ℓ)	0.21	-0.10	0.80	-	-
$W + bb$,cc (e)	0.23	-	-	0.09	0.95
$W + c$ (e)	0.40	-	-	0.06	0.96
$W + lf$ (e)	0.21	-	-	-0.06	0.94
$W + bb$,cc (μ)	0.23	-	-	0.16	0.92
$W + c$ (μ)	0.40	-	-	0.11	0.95
$W + lf$ (μ)	0.21	-	-	-0.08	0.90

Table H.1: Overview of nuisance parameters for measurement with an Asimov pseudo-dataset. The mean value of the nuisance parameters is presented together with its standard derivation. The W+jets background is calibrated in-situ. For the QCD sample the nuisance parameters per bin-pair are not displayed, since the selected bins are different for resolved and boosted channel. The entries 0.0* refer to absolute values below 0.01. All priors are Gaussian. The posterior with is provided with respect to the prior.

H.2 Comparison with data

The summary of nuisance parameters for the top anti-top production charge asymmetry measurement method check with an Asimov pseudo-dataset is presented in table H.1. The nuisance parameter are discussed in section 10.9.

		resolved (M1-M3)		boosted (M4)	
source	**prior**	**NP mean**	**std mean**	**NP mean**	**std mean**
QCD (e), 0-btag	0.5	-0.10	0.47	-0.38	0.67
QCD (e), 1-btag	0.5	-1.84	0.12	-1.72	0.18
QCD (e), 2-btags	0.5	-1.85	0.13	-0.44	0.69
QCD (μ), 0-btags	0.5	0.03	0.88	-0.19	0.75
QCD (μ), 1-btag	0.5	0.03	0.88	-0.10	0.86
QCD (μ), 2-btags	0.5	-0.59	0.62	0.91	0.72
Z+jets	0.5	-0.20	0.64	-1.41	0.40
single top	0.05	-0.26	0.99	-0.37	1.01
VV	0.5	-0.30	0.87	-0.74	0.75
$t\bar{t}$ +V	0.5	0.04	0.95	0.12	0.96
W+bb,cc (ℓ)	0.23	-0.13	0.79	-	-
$W + c$ (ℓ)	0.40	0.34	0.70	-	-
$W + lf$ (ℓ)	0.21	0.24	0.79	-	-
$W + bb$,cc (e)	0.23	-	-	-1.28	0.93
$W + c$ (e)	0.40	-	-	-0.73	0.88
$W + lf$ (e)	0.21	-	-	-0.62	0.91
$W + bb$,cc (μ)	0.23	-	-	-2.16	0.85
$W + c$ (μ)	0.40	-	-	-1.07	0.81
$W + lf$ (μ)	0.21	-	-	-0.68	0.83

Table H.2: Overview of nuisance parameters for measurement in ℓ+jets channel with data corresponding to 36.1 fb^{-1} integrated luminosity. The mean value of the nuisance parameters is presented together with its posterior width. The W+jets background is marginalized in-situ, with HFSF applied. For the multijet sample the nuisance parameters per bin-pair are not displayed, since the selected bins are different for resolved and boosted channel. All priors are Gaussian.