Karlsruher Institut für Technologie

IEKP-KA/2014-08

Differentielle Messung der Ladungsasymmetrie in Top-Quark-Paarereignissen im sichtbaren Phasenraum mit dem CMS-Experiment

Christian Buntin

Masterarbeit

An der Fakultät für Physik
Institut für Experimentelle Kernphysik

Referent: Prof. Dr. Th. Müller
Korreferent: Prof. Dr. G. Quast

13. Mai 2014

Zusammenfassung

2012 war ein äußerst spannendes Jahr für die Teilchenphysik. Am Large Hadron Colliders (LHC) wurden Proton-Proton-Kollisionen mit einer unter Laborbedingungen bisher unerreichten Schwerpunktsenergie und Luminosität erzeugt. In diesen Kollisionen wurden schwere Teilchen produziert, die in alltäglicher Materie nicht vorkommen. Um diese Teilchen zu messen und zu rekonstruieren wurden große Teilchendetektoren wie das Compact-Muon-Solenoid-Experiment (CMS) um die Wechselwirkungspunkte herum errichtet. Die so im Jahr 2012 gemessenen Daten ermöglichen es, Theorievorhersagen des Standardmodells der Teilchenphysik auf höheren Energieskalen zu prüfen, als es zuvor möglich war.

Dieses Standardmodell der Teilchenphysik wurde in den 1960er Jahren entwickelt und ist bis heute die allgemein akzeptierte Theorie der Elementarteilchen und der Wechselwirkungen zwischen ihnen. Es beschreibt die drei fundamentalen Kräfte, die starke, die schwache und die elektromagnetische Kraft, welche durch entsprechende Eichbosonen vermittelt werden. Weiterhin beschreibt das Standardmodell zwölf Fermionen, von denen sechs Leptonen sind, die über die schwache und elektromagnetische Wechselwirkung interagieren. Die anderen sechs Teilchen sind Quarks, welche über alle drei fundamentalen Kräfte wechselwirken. Das schwerste bekannte Teilchen im Standardmodell ist das Top-Quark, welches im Jahr 1995 am Tevatron-Beschleuniger entdeckt wurde [1, 2]. Aufgrund seiner großen Masse von etwa 173 $\mathrm{GeV/c^2}$ zerfällt es, bevor es gebundene Zustände bilden kann. Dies ermöglicht es, die Eigenschaften eines quasi-freien Quarks zu studieren, um die Vorhersagen des Standardmodells bei großen Energien zu prüfen.

Top-Quarks können als Quark-Antiquark-Paare über die starke Wechselwirkung in Gluon-Gluon-Fusionsprozessen oder in Quark-Antiquark-Annihilationsprozessen erzeugt werden. Im asymmetrischen Prozess der Quark-Antiquark-Annihilation tritt das Phänomen der $\mathrm{t\bar{t}}$-Ladungsasymmetrie [3–5] auf, die aus Interferenzen zwischen verschiedenen Feynmandiagrammen der $\mathrm{t\bar{t}}$-Produktion resultiert: Bei einer positiven Asymmetrie werden Top-Quarks bevorzugt in die Richtung des einlaufenden Quarks und Top-Antiquarks in die Richtung des einlaufenden Antiquarks emittiert. Am Tevatron-Beschleuniger äußerte sich dies in einem Überschuss von Top-Quarks in Vorwärtsrichtung und Top-Antiquarks in Rückwärtsrichtung. Am LHC ist der Anfangszustand dagegen symmetrisch. Da Antiquarks in Protonen jedoch nur als See-Quarks mit einem geringeren Impulsanteil existieren, äußert sich die Ladungsasymmetrie am LHC durch einen Überschuss von Top-Antiquarks im zentralen Bereich des Detektors, wohingegen bei Top-Quarks ein Überschuss in der Vorwärts- und Rückwärtsrichtung auftritt. Dies führt zu unterschiedlichen Breiten der Rapiditätsverteilungen von Top-Quarks und Top-Antiquarks. Um diesen Effekt zu bestimmen, wird die Rapidität y als Maß für die Flugrichtung eines Teilchens eingeführt. Diese ist durch

$$y := \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right)$$

gegeben, wobei E die Energie und p_z den Impuls des Teilchens entlang der Strahlachse bezeichnen.

Zur Quantifizierung des Breitenunterschieds der Rapiditätsverteilungen von Top-Quarks und Top-Antiquarks wird die sensitive Variable

$$\Delta|y| := |y_{\mathrm{t}}| - |y_{\bar{\mathrm{t}}}|$$

eingeführt. Damit kann die Ladungsasymmetrie in Top-Quark-Paar-Ereignissen als Asymmetrie dieser sensitiven Variable gemessen werden:

$$A_C := \frac{N^+ - N^-}{N^+ + N^-},$$

wobei N^+ und N^- die Anzahl der Ereignisse mit positivem beziehungsweise negativem Vorzeichen von $\Delta|y|$ sind.

Das Ziel dieser Arbeit ist die differentielle Messung der $\mathrm{t\bar{t}}$-Ladungsasymmetrie nicht nur im vollen Phasenraum, sondern auch in einem eingeschränkten sichtbaren Phasenraum, um Modellabhängigkeiten zu reduzieren. Dieser sichtbare Phasenraum ist dabei eine auf Generatorgrößen klar definierte Region im vollen Phasenraum, die vergleichbar mit der Akzeptanzregion des Detektors ist. Als Teil dieser Masterarbeit wurden dazu zwei verschiedene Definitionen von sichtbaren Phasenräumen erarbeitet. Der *sichtbare Phasenraum definiert auf Teilchenniveau* basiert dabei auf Selektionskriterien angewendet auf Generatorjets. Da die Simulation solcher Generatorjets aber nur phänomenologisch möglich ist, sind für Messungen in diesem Phasenraum keine expliziten Theorievorhersagen vorhanden. Daher wurde der *sichtbare Phasenraum definiert auf Partonniveau* eingeführt. Dieser basiert auf Selektionskriterien, welche auf geclusterte Partonen vor der Hadronisierung angewendet wurden, was damit explizite Theorievorhersagen ermöglicht. Die Messung der $\mathrm{t\bar{t}}$-Ladungsasymmetrie wurde anschließend für den vollen Phasenraum und für diese beiden Phasenräume durchgeführt.

Die Ladungsasymmetrie wird dabei differentiell in der invarianten Masse $m_{\mathrm{t\bar{t}}}$, im transversalen Impuls $p_{\mathrm{T,t\bar{t}}}$ und in der Rapidität $|y_{\mathrm{t\bar{t}}}|$ des Top-Quark-Paares gemessen, um so genauere Einblicke in die Ladungsasymmetrie zu erhalten.

Hierfür wurde der im Jahr 2012 vom CMS-Experiment aufgezeichnete Datensatz von Proton-Proton-Kollisionen bei einer Schwerpunktsenergie von 8 TeV verwendet. Diese Daten entsprechen einer integrierten Luminosität von 19.7 fb^{-1}. Durch zahlreiche Selektionsschritte wurden Ereignisse selektiert, die der Signatur des Lepton+Jets Zerfallskanals von Top-Quark-Paaren entsprechen. Dabei zerfallen beide Top-Quarks jeweils in ein W-Boson und in ein Bottom-Quark. Eines der W-Bosonen zerfällt daraufhin hadronisch in zwei leichte Quarks, während das andere W-Boson in diesem Kanal in ein geladenes Lepton und das entsprechende Neutrino zerfällt. Daher werden in der Ereignisselektion mindestens vier Jets, sowie ein einziges isoliertes geladenes Lepton gefordert. Dabei werden nur Elektronen und Myonen berücksichtigt. Um die Reinheit des Signals weiter zu erhöhen, wurde ein b-Tagger eingesetzt. Dabei wird das Vorhandensein von mindestens einem Jet, der durch die Hadronisierung eines Bottom-Quarks entstanden ist, gefordert.

Trotz dieser Selektion bleiben immer noch Ereignisse übrig, welche nicht dem Zerfall eines Top-Quark-Paares entsprechen. Um die verbleibenden Beiträge der verschiedenen Untergrundprozesse zu bestimmen, wurden Computersimulationen eingesetzt. Diese sogenannten Monte-Carlo-Simulationen ermöglichen die Berechnung aller möglichen Prozesse in Proton-Proton-Kollisionen. Die oben beschriebene Selektion wurde auch auf die simulierten Untergrundprozesse angewendet. Anschließend wurde mittels eines Binned-Likelihood-Fits der simulierten Prozesse an die gemessenen Daten die Größe der einzelne Beiträge abgeschätzt.

Für den Multijet-Untergrundprozess wurden keine Simulationen eingesetzt, stattdessen wurde eine datenbasierte Verteilung aus einem Seitenband verwendet. Als Teil dieser Masterarbeit wurde dabei die Selektion für die Verteilung des Myon+Jets Zerfallskanals weiter verbessert.

Um die sensitive Variable und die sekundären Variablen zu bestimmen, ist eine Rekonstruktion der Vierervektoren beider Top-Quarks erforderlich. Dazu werden die Vierervektoren der zugehörigen Zerfallsprodukte entsprechend kombiniert. Allerdings ist die Zuordnung der gemessenen Zerfallsprodukte zu den Zerfallsprodukten des Top-Quark-Paars nicht eindeutig. Daher wurde ein Likelihood-Kriterium basierend auf simulierten Wahrscheinlichkeitsdichten der rekonstruierten invarianten Massen und der b-Tagger-Ausgaben angewendet.

Nach der Rekonstruktion der Top-Quarks werden die Beiträge der Untergrundprozesse den Fit-Ergebnissen entsprechend abgezogen. Die dadurch erhaltenen Verteilungen der sensitiven Variable beinhalten allerdings noch Verzerrungen durch die Selektion und Ungenauigkeiten in der Rekonstruktion. Um diese zu korrigieren wird eine regularisierte Entfaltungsmethode auf Grundlage einer generalisierten Matrixinversion angewendet. Die dabei durchgeführte Extrapolation der Messung in den vollen Phasenraum wurde für diese Arbeit auf Extrapolationen in die sichtbaren Phasenräume ergänzt. Dazu wurde die Selektionseffizienz für die Akzeptanzkorrektur entsprechend der sichtbaren Phasenräume angepasst und in der Entfaltung angewendet.

Auf der Basis von verschiedenen Pseudoexperimenten wurden zusätzliche Tests durchgeführt, um die Konsistenz der Entfaltungsmethode zu verifizieren.

Über die Entfaltung der rekonstruierten Verteilungen der sensitiven Variable konnte so schließlich die Ladungsasymmetrie für die Messung im vollen Phasenraum zu

$$A_{\mathrm{C}}^{\Delta|y|} = \quad 0.005 \pm 0.007 \text{ (stat.)} \pm 0.006 \text{ (syst.)}$$

bestimmt werden, sowie

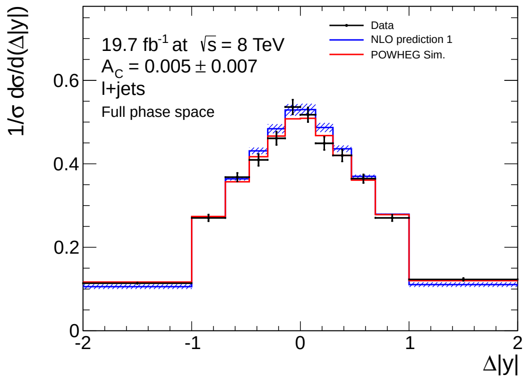
$$A_{\mathrm{C}}^{\Delta|y|,\mathrm{sichtbar,Teilchen}} = -0.001 \pm 0.008 \text{ (stat.)} \pm 0.006 \text{ (syst.)}$$

und

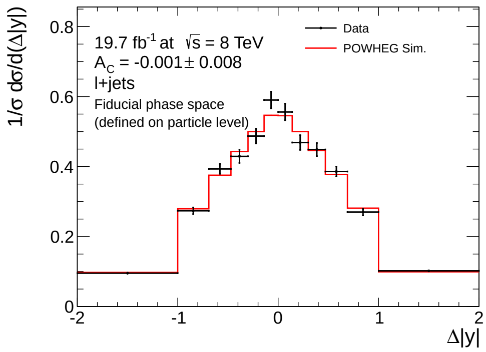
$$A_{\mathrm{C}}^{\Delta|y|,\mathrm{sichtbar,Parton}} = \quad 0.001 \pm 0.008 \text{ (stat.)} \pm 0.006 \text{ (syst.)}$$

für die Messungen im sichtbaren Phasenraum definiert auf Teilchen- beziehungsweise Parton-Niveau.

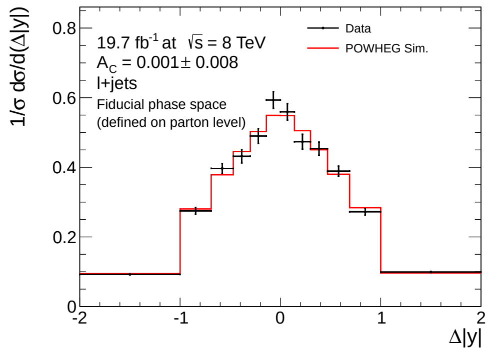
Die Messung im vollen Phasenraum stimmt gut mit dem vorhergesagten Wert im Standardmodell von $A_{C}^{\Delta|y|,\mathrm{SM}} = 0.0111 \pm 0.0004$ [6] überein. Die Messungen in den



(a)



(b)



(c)

Abbildung 0.1.: Entfaltete inklusive Verteilungen der sensitiven Variable $\Delta|y|$ im vollen Phasenraum (a) und im sichtbaren Phasenraum definiert auf Teilchen-Niveau (b) beziehungsweise auf Parton-Niveau (c). Die Verteilungen werden mit den Vorhersagen einer Powheg-Simulation [7–9] verglichen. Für die Verteilung im vollen Phasenraum ist zusätzlich eine Vorhersage [10], die in nächstführender Ordnung im Standardmodell bestimmt wurde, eingezeichnet.

beiden sichtbaren Phasenräumen sind vergleichbar miteinander und zeigen im Mittel etwas kleinere systematische Unsicherheiten mit leicht erhöhten statistischen Unsicherheiten, deren Ursache aber in dieser Arbeit untersucht und verstanden wurde. Die resultierenden Verteilungen der sensitiven Variable, aus denen die Ladungsasymmetrie bestimmt wurde, sind in Abbildung 0.1 dargestellt.

Die in dieser Arbeit durchgeführten Messungen ermöglichen durch Extrapolationen in den sichtbaren Phasenraum statt in den vollen Phasenraum neue Einblicke in die Ladungsasymmetrie. Weitere Erkenntnisse werden im Jahr 2015 erwartet, wenn der LHC bei einer Schwerpunktsenergie von 13 TeV seine Arbeit aufnehmen wird. Zwar geht bei höheren Schwerpunktsenergien die inklusive Ladungsasymmetrie wegen der Dominanz des symmetrischen Gluon-Gluon-Fusionsprozesses zurück, allerdings ermöglichen die größeren Wirkungsquerschnitte und höheren Luminositäten noch detailliertere Messungen der Ladungsasymmetrie, vor allem in den Phasenräumen der differentiellen Messungen und mit neuen sensitiven Variablen, um somit entscheidende Kenntnisse über die Existenz von Physik jenseits des Standardmodells zu liefern.

IEKP-KA/2014-08

Fiducial Measurement of the Charge Asymmetry in Top Quark Pair Production at the CMS-Detector

Master's thesis
of
Christian Buntin

At the Department of Physics
Institut für Experimentelle Kernphysik

Advisor:	Prof. Dr. Th. Müller
Second advisor:	Prof. Dr. G. Quast

May 13, 2014

Introduction

The year 2012 has been a very exciting year for particle physics. The outstanding performance of the Large Hadron Collider (LHC) allowed proton-proton collisions at a center-of-mass energy and with a luminosity that had not been reached before in laboratory conditions. In these collisions, massive particles were produced, that are not present in ordinary matter. These then decay into known lighter particles, which can be detected. To measure and reconstruct this collision products, large particle detectors like the Compact Muon Solenoid (CMS) experiment have been constructed around the interaction points. The data taken in 2012 allows to test theoretical predictions of the Standard Model at higher energies than it was possible before.

The Standard Model of particle physics was developed in the 1960s and is today's universally accepted theory of the elementary particles and the interactions between them. It describes the three fundamental forces, the strong, the weak and the electromagnetic forces, which are mediated by the corresponding gauge bosons. The matter in the Standard Model consists of twelve fermions, divided into a group of six leptons interacting via the weak and electromagnetic force and into a group of six quarks interacting via all three fundamental forces. The heaviest known particle in the Standard Model is the top quark, which was discovered at the Tevatron collider in 1995 [1, 2]. Due to its high mass of about 173 $\mathrm{GeV/c^2}$ it instantly decays before forming bound states. This allows to study the properties of a quasi-free quark to probe the Standard Model at high energies.

Top quarks are mostly produced as quark-antiquark pairs via the strong interaction by gluon-gluon fusion or quark-antiquark annihilation processes. The asymmetric quark-antiquark annihilation induces the specific behavior of the $\mathrm{t\bar{t}}$ charge asymmetry [3–5]: For a positive asymmetry the top quark is emitted preferentially in the direction of the incoming quark and the top antiquark in the direction of the incoming antiquark. At the Tevatron collider this behavior leads to an excess of top quarks in the forward direction and an excess of top antiquarks in the backward direction. Since antiquarks only occur as sea quarks in protons, their average momentum fraction is much smaller than of valence quarks. Therefore the charge asymmetry appears at the LHC as an excess of top antiquarks in the central region of the detector and of top quarks in the forward and backward regions of the detector.

In this thesis, inclusive and differential measurements of the charge asymmetry in top-quark pair-production are presented, using the lepton+jets decay channel. The secondary variables of the differential measurements are the invariant mass, the transverse momentum and the rapidity of the top-quark pair. The dataset with an integrated luminosity of 19.7 $\mathrm{fb^{-1}}$, which was recorded by the CMS experiment in 2012, is analyzed to gain further insights into the nature of the charge asymmetry.

The measurement of the charge asymmetry involves an extrapolation of the measured result to the full phase space. To reduce the model dependence of this extrapolation procedure, the charge asymmetry can be measured in fiducial phase spaces. These are well defined regions in the full phase space, which are comparable to the detector's acceptance region. In this thesis two different fiducial phase spaces are analyzed and additional measurements of the charge asymmetry are performed in these phase spaces.

The first chapter of this thesis gives a short theoretical introduction of the Standard Model and the charge asymmetry in top-quark pair-production. The functionality of the LHC and the individual components of the CMS experiment are described in Chapter 2. In Chapter 3 the simulation of collision events based on Monte Carlo methods as well as the reconstruction of events from raw detector information are explained. The event selection procedure and the estimation of the remaining background contributions are described in Chapter 4. Finally in Chapter 5 the reconstruction of the top quarks' four-momenta, the background subtraction and the correction for selection and reconstruction effects using a regularized unfolding procedure, as well as the estimation of systematic uncertainties, are described.

Contents

1. Theoretical Introduction

The Standard Model of Particle Physics [11–20] was developed between 1960 and 1970 and is today's universally accepted theory of the elementary particles and their properties and interactions. It did not only explain the experimental observations in the time of its development, it also predicted lots of new particles, which were discovered in the following years. The charm quark was discovered in 1974 [21, 22] and the bottom quark in 1977 [23]. With the discovery of the top quark in 1995 [1,2] and the tau neutrino in 2000 [24] the third generation of fermions was complete. Finally in 2012 the discovery of the long before predicted Higgs boson [25, 26] completed the Standard Model as known today.

Despite all this, the Standard Model is not a complete theory of nature. For example the gravitation as the fourth fundamental force can be explained by Einstein's theory of General Relativity [27], but not by the Standard Model. Furthermore, the cosmological problems about the nature of dark matter and dark energy [28] along with the reasons for the baryon asymmetry of the universe are not covered yet. The explanation of these and other missing parts is attempted by many models that extend the theory of the Standard Model.

In the following sections a short overview of the Standard Model is given. The pair-production and decay of top quarks as well as the $\mathrm{t\bar{t}}$ charge asymmetry and its measurement are explained in more detail.

1.1. The Standard Model of Particle Physics

The Standard Model of particle physics is a relativistic quantum field theory based on quantum mechanics and the theory of special relativity. Particles and their interactions are described by a Lagrangian that is invariant under certain local gauge transformations. These transformations correspond to the symmetries $SU(3) \times SU(2) \times U(1)$, which represent the three fundamental forces and their gauge bosons: The *strong force*, with gluons as exchange particles, the *weak force*, exchanged via W and Z bosons, and the *electromagnetic force*, mediated via photons.

According to Noether's theorem [29], which demands a corresponding conservation law for any continuous symmetry of a physical system, all these symmetries can be linked to three different charges in the Standard Model: The *color charge* for the strong force, the *electrical charge* for the electromagnetic force and the *weak isospin* for the weak interaction.

Additionally each particle has a quantum number called *spin*, which describes a particle's intrinsic angular momentum. Particles with an integer spin are called *bosons* and mediate the fundamental forces. Particles with a half-integer spin are called *fermions* and are usually associated with matter.

One extra feature of the Standard Model is the fact that each particle has its corresponding antiparticle with the same mass but with oppositely signed electrical charge. Because neutrinos do not have an electrical charge, they can be theorized

Force	Conserved Quantity	Mediating Bosons	Mass [GeV/c^2]
strong	color charge	gluons (g)	0
electromagnetic	electric charge	photons (γ)	$\leq 1 \cdot 10^{-27}$
weak	weak isospin	W boson ($W^{\pm}$)	80.385 ± 0.015
		Z boson (Z^0)	91.1876 ± 0.0021

Table 1.1.: List of the three forces and the associated conserved quantities and gauge bosons of the Standard Model along with their symbols and masses. The gluon mass is noted as zero according to the theory prediction. All other values are taken from [39].

as Majorana fermions [30], which would be their own antiparticles.

1.1.1. Gauge Bosons and Fermions

The gauge bosons with a spin quantum number of one are the mediators of the three fundamental forces.

The strong interaction has its theoretical description in Quantum Chromodynamics (QCD) [19,20]. It couples to the color charge, which appears in three states named red, green and blue and their associated anti-colors anti-red, anti-green and anti-blue. In composite particles either these colors and anti-colors cancel each other out or they all add up to white, so only color neutral particles exist. The strong force is mediated by gluons, which themselves carry color charges, namely one color charge and one anti-color charge. This results in eight linearly independent types of gluons which also can couple to each other. This force holds together the protons and neutrons of atomic nuclei.

The electromagnetic force is described theoretically in Quantum Electrodynamics (QED) [31–38]. Its mediating particles are photons, which couple to electrically charged particles. They have no mass, no electric charge and no color charge, therefore they can not couple to other photons and their range is unlimited. This force makes it possible for negatively charged electrons and positively charged nuclei to form atoms.

The weakest of the three fundamental forces of the Standard Model is the weak force. It is exchanged by massive gauge bosons, the electrically neutral Z^0 boson with a mass of 91.2 GeV/c^2 and the electrically charged W^+ and W^- bosons with a mass of 80.4 GeV/c^2. These large masses of the mediating particles result in a range on only sub-nuclear scales, making the force weak. In atoms this force is responsible for the beta decay.

The three fundamental forces and their corresponding quantities and gauge bosons are summarized in table 1.1.

Like the unification of the electric and the magnetic force in the electromagnetic theory, the electromagnetic force and the weak force are unified in the electroweak theory [11, 12, 14, 16]. A problem in this unification was the need to introduce mass terms for the bosons into the Lagrangian. This would cause a violation of local gauge symmetry, which was solved by the electroweak symmetry breaking, introduced by the Higgs mechanism.

Besides the bosons the Standard Model also consists of particles with a half-integer spin quantum number: the fermions. They can be divided into color-charged quarks and color-neutral leptons, which are all arranged in three generations. The

Gen.	Name	EC [e]	CC	Mass [MeV/c^2]
I	up quark (u)	$+\frac{2}{3}$	r, g or b	$2.3^{+0.7}_{-0.5}$
	down quark (d)	$-\frac{1}{3}$	r, g or b	$4.8^{+0.5}_{-0.3}$
	electron neutrino (ν_e)	0	0	$< 2 \cdot 10^{-6}$
	electron (e)	-1	0	$510.998928 \cdot 10^{-3} \pm 11 \cdot 10^{-9}$
II	charm quark (c)	$+\frac{2}{3}$	r, g or b	$(1.275 \pm 0.025) \cdot 10^3$
	strange quark (s)	$-\frac{1}{3}$	r, g or b	95 ± 5
	muon neutrino (ν_μ)	0	0	< 0.19
	muon (μ)	-1	0	$105.6583715 \pm 3.5 \cdot 10^{-6}$
III	top quark (t)	$+\frac{2}{3}$	r, g or b	$(173.34 \pm 0.27 \pm 0.71) \cdot 10^3$
	bottom quark (b)	$-\frac{1}{3}$	r, g or b	$(4.18 \pm 0.03) \cdot 10^3$
	tau neutrino (ν_τ)	0	0	< 18.2
	tau (τ)	-1	0	1776.82 ± 0.16

Table 1.2.: List of all fermions of the Standard Model, ordered by generation (Gen.). The electric charge (EC) is given in units of the elementary charge e. The possible values of the color charge (CC) are listed with the abbreviations r, g, b for red, green and blue and 0 for colorless particles. The value of the top quark mass is taken from [40], all other values are from [39].

only difference between these generations is the higher mass of particles of higher generations. Each generation includes an up-type quark, a down-type quark, an electrically charged lepton and a neutrino. A summary of all the fermions in the Standard Model is shown in table 1.2.

The charged leptons consist of electrons, muons and taus and interact via the weak and the electromagnetic forces. Each of them has a weak isospin partner, one of the electrically neutral neutrinos. They only interact via the weak force and for a long time they were considered to be massless. But direct measurements of neutrino oscillations [41] can only be explained by finite mass differences between the neutrino generations. To introduce the required mass generation mechanism for the neutrinos, theorists have proposed various extensions of the Standard Model [42].

The quarks carry the charges of all three forces and therefore can interact via all three of them, but typically the interactions are dominated by the strong force. This force is characterized by the *asymptotic freedom* and the *confinement*: At high energies and small distances the force is only weak, allowing the quarks to act like quasi-free particles (asymptotic freedom). But for large distances or small energies the color charge of the gluons leads to a divergent strength of the strong force. Therefore color-charged particles like quarks can only exist in bound colorless states, meaning they are confined. These colorless compound states called hadrons can be achieved in two different ways: A quark and an antiquark charged with a color and the corresponding anti-color form color-neutral bound states called mesons, while three quarks with different color charges form colorless baryons. If a quark in a bound state gets enough energy to escape its hadron, the gluon field becomes more and more energetic until it is converted into a new quark-antiquark pair, which restores color neutrality and allows new hadrons to form.

1.1.2. Spontaneous Symmetry Breaking and Higgs Mechanism

As mentioned before, to explain the experimentally observed masses of the W and Z bosons, a mass term can be introduced into the Lagrangian. However this violates the local gauge symmetry, so a different mechanism is required. Therefore an additional four-component scalar field with a non-zero vacuum expectation value that forms a complex doublet of the weak isospin SU(2) symmetry, called the *Higgs field*, is introduced. By interactions with the weak gauge fields it causes spontaneous symmetry breaking.

Such a broken symmetry results in the generation of pseudo-Goldstone bosons, as demanded by the Goldstone theorem [43, 44]. These Goldstone bosons are then absorbed by the fields of the W and Z bosons. This results in three out of the four degrees of freedom turning into the mass terms of these bosons. The remaining degree of freedom leads to the manifestation of a massive boson, which can be identified with the Higgs boson as described by the Higgs mechanism [45–47].

Via the Yukawa coupling to fermions, the Higgs field also generates masses of quarks and charged leptons. But the masses of neutrinos cannot be explained by the Standard Model alone.

The recent experimental discovery of a Higgs boson by ATLAS [25] and CMS [26, 48] with a mass of about 126 GeV/c^2 finally confirms the existing of this last missing particle of the Standard Model.

1.1.3. Processes in the Standard Model

Because the Standard Model is based on quantum mechanics it can only predict the probability for a given process to occur. This probability for a transition from an initial state $|i\rangle$ to a final state $|f\rangle$ is given by Fermi's Golden Rule [49] and is proportional to the transition amplitude $|\mathcal{M}_{fi}|^2$ of the given process. The cross section σ of this process then results from an integration of its transition amplitude over all possible initial and final states and is given in units of barn, with $1\ \text{b} = 10^{-28}\ \text{m}^2$.

Richard Feynman developed a set of rules by which such transition amplitudes can be calculated from simple visualizations, the so-called *Feynman diagrams* [36]. Examples for Feynman diagrams of the fundamental interactions in the Standard Model are shown in figure 1.1.

1.2. Top Quark Physics

With the discovery of the top quark in 1995 by the CDF [1] and DØ [2] collaborations at the Tevatron, the last missing quark of the Standard Model has been confirmed. It was already predicted in 1973 together with the bottom quark, its weak isospin partner by the prediction of a third generation of fermions. That resulted from the introduction of the CKM matrix as a unitary matrix with a complex phase that contains information on the strength of flavour-changing interactions of the weak force to explain the observed CP-violation in kaon decays [50].

With its mass of $173.34 \pm 0.27 \pm 0.71$ GeV/c^2, given by the combination [40] of the CDF and DØ results with the latest ATLAS and CMS results, the top quark

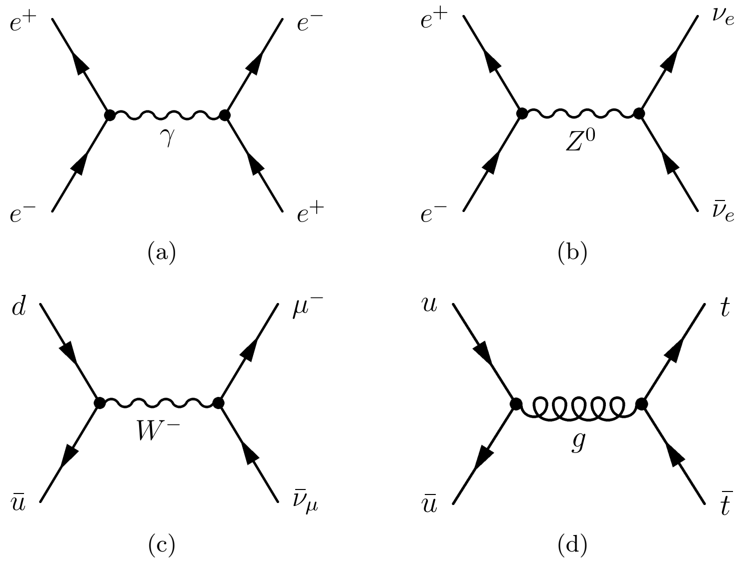


Figure 1.1.: Examples of leading order Feynman diagrams for the fundamental interactions. At the top the electron-positron annihilation via the electromagnetic (a) and weak interaction (b) is shown. An example for the other type of weak interaction is the annihilation of two quarks via a charged W boson. This is shown in (c) at the bottom together with the the annihilation of two quarks via the strong force in (d). Time is flowing from left to right.

is the heaviest particle in the Standard Model. This makes this particle special in two different ways:

First, the high mass of the top quark leads to a large decay width and therefore a very short lifetime of $\tau \approx 10^{-25}$ s [39]. This is shorter than the typical time scale of QCD interactions of $1/\Lambda_{\mathrm{QCD}} \approx 10^{-23}$ s by about two orders of magnitude and makes the top quark decay before hadronization. Because in this process it passes all of its spin information to its decay products, it can be studied as a quasi-free quark.

Secondly, it is linked to the electroweak symmetry breaking with the vacuum expectation value of the Higgs field of $v = 246$ GeV: The top quark Yukawa coupling y_{t} has a value of

$$y_{\mathrm{t}} = \sqrt{2}\,\frac{m_{\mathrm{t}}}{v} \approx 1, \tag{1.1}$$

which is quite surprising because there is no obvious reason for this. Therefore this might be a hint for the important role of the top quark in the understanding of the Higgs mechanism.

A more detailed overview over top quark physics can be found in [51].

1.2.1. Production of Top Quarks

At hadron colliders top quarks are generated by two different mechanisms:

On the one hand, single top quarks can be produced by processes of the weak interaction, which were first observed at the Tevatron in 2009 [52, 53]. These production processes allow the determination of the $|V_{\mathrm{tb}}|$ element of the CKM matrix, because all of them contain a vertex with a top and a bottom quark and a W boson.

But on the other hand the more common process is the production of pairs

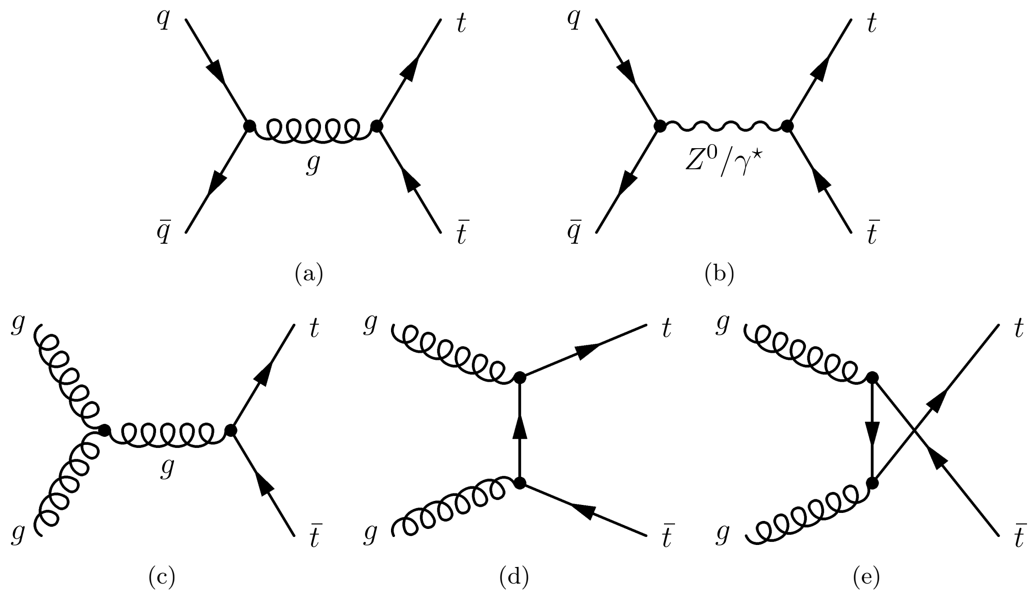


Figure 1.2.: Leading order Feynman diagrams of the production of top quark pairs in hadron collisions: At the top the quark-antiquark annihilation processes via the strong interaction (a) and via the electroweak interaction (b) are shown, at the bottom gluon-gluon fusion processes via strong interactions in the s-channel (c), the t-channel (d) and the u-channel(e) are shown.

of top quarks and top anti-quarks. Due to the high center-of-mass energies of current particle accelerators, the dominating strong force also is the main force involved in top-quark pair-production. The leading order Feynman diagrams for the production of top quark pairs, which consist of quark-antiquark annihilation and gluon fusion processes, are shown in figure 1.2.

At the Tevatron with its proton-antiproton collisions the quark-antiquark annihilation was the dominant process. At the LHC with its symmetric initial state in proton-proton collisions, gluon fusion becomes more important instead, which becomes even more prominent when going to larger center-of-mass energies. These effects are due to the different distributions of the partons in the proton, which are given by the so-called *parton distribution functions* (PDFs) shown in figure 1.3.

The latest approximate next-to-next-to-leading order calculation of the cross-section of top-quark pair-production in proton-proton collisions at a center-of-mass energy of 8 TeV [55] is

$$\sigma(\mathrm{pp} \to \mathrm{t\bar{t}}) = 245.8^{+6.2+6.2}_{-8.4-6.4}\ \mathrm{pb}, \tag{1.2}$$

with uncertainties arising from scale variations and PDF uncertainties [56].

1.2.2. Decay of the Top Quark

In nearly all cases the top quark decays into a bottom quark and a W boson, because the corresponding element of the CKM-matrix V_{tb} is close to one. The W boson either decays leptonically into a charged lepton and the corresponding neutrino or hadronically into a quark-antiquark pair, as shown in figure 1.4. The coupling of the W boson to these two different kinds of weak isospin doublets is the

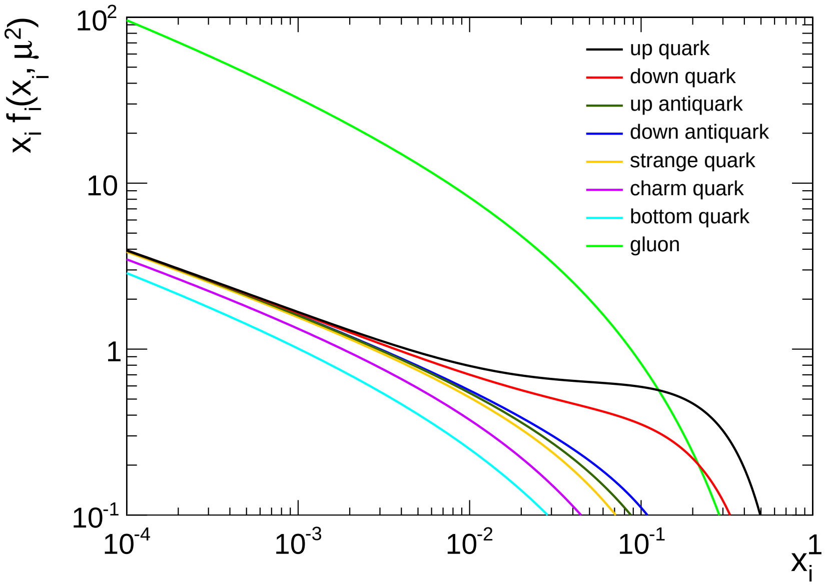


Figure 1.3.: The CT10 [54] proton PDF for gluons and quarks at a scale of $\mu^2 = (172.5\ \mathrm{GeV/c^2})^2$. This value for μ^2 is commonly used for the simulation of top-quark events.

same, but because quarks appear in three different color charges, their branching ratio is three times higher than for leptons.

In the decay of a top-quark pair three different channels can be observed, depending on the decays of the top quarks:

- In the full-hadronic channel both W bosons decay hadronically into a quark-antiquark pair. This results in a total of six jets in an event, with two jets originating from bottom quarks.

- The di-leptonic channel requires both W bosons to decay into charged leptons and neutrinos, resulting in two jets from the bottom quarks, two charged leptons and missing energy from the two neutrinos.

- In the semi-leptonic channel one W boson decays leptonically and one hadronically. This leads to two jets from the W boson and two from the bottom quarks, one charged lepton and missing energy from one neutrino. Thus this decay channel is also called the *lepton+jets channel*.

Because of its clear signature and the still good branching ratio of about 30%, the semi-leptonic mode is considered as the most powerful one for many analyses.

1.3. The Charge Asymmetry of Top-Quark Pair-Production

One phenomenon in top-quark pair-production is the charge asymmetry [3,4]. This describes an excess of top quarks over top antiquarks in certain kinematic regions and vice versa. In the Standard Model it occurs from higher order corrections in QCD calculations of the annihilation of quark and antiquark pairs to top-quark pairs. The interference of the Born diagram with the box diagram yields a positive

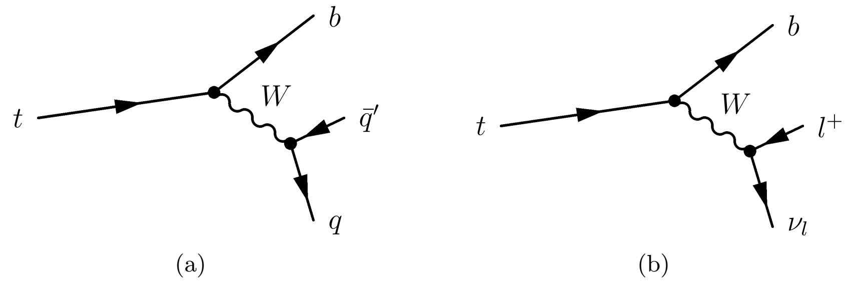


Figure 1.4.: The top-quark decay into a bottom quark and a W boson. The W boson decays either hadronically into a quark q and an antiquark $\bar{\mathrm{q}}'$ (a) or leptonically into a charged lepton and the corresponding neutrino (b).

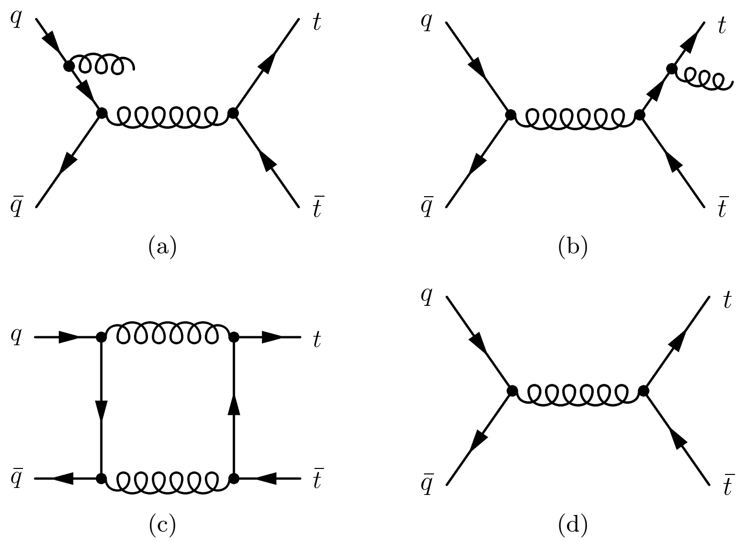


Figure 1.5.: The origin of the charge asymmetry in the Standard Model: The interference of initial state radiation (a) and final state radiation (b) shown at the top yields a negative contribution, the interference between box diagram (c) and Born diagram (d) shown at the bottom yields a positive one. Only representative diagrams for each process are displayed.

contribution to the charge asymmetry, the interference between initial state radiation (ISR) and final state radiation (FSR) yields a negative one. These different processes are shown in figure 1.5. Additionally a small contribution to the charge asymmetry is also given by the electroweak annihilation of quark-antiquark pairs.

The sum of these effects results in top quarks being emitted preferentially in the direction of the incoming quark and top antiquarks being emitted preferentially in the direction of the incoming antiquark, yielding a slightly positive charge asymmetry.

The top quark also plays a special role in theories beyond the Standard Model (BSM), because the proposed exchange particles often couple preferentially to heavy quarks like the top quark. Then top-quark pair-production is possible not only via the gauge bosons, but also via these new exchange particles. Examples for these are Z' bosons [57], axigluons [5, 58] and Kaluza Klein excitations of gluons [59, 60].

By different couplings to top quarks and top antiquarks, for example, these new exchange particles could affect the value of the charge asymmetry. An overview of theoretical models which could result in deviations from the charge asymmetry predicted by the Standard Model can be found in [61–63].

1.3.1. Measurement of the Charge Asymmetry

The measurement of the charge asymmetry is generally performed by the measurement of the pseudorapidities or the rapidities of the top quarks and antiquarks.

The rapidity y of a particle depends on a preferred direction given by the z axis, usually chosen to be along the direction of the beam pipe. It is given by

$$y := \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right), \tag{1.3}$$

where E is the particles energy and p_z its momentum component in the chosen z direction. This quantity has the advantage of being invariant under Lorentz boosts along the z axis.

The pseudorapidity η is related to the rapidity, but easier to measure experimentally. It solely depends on the angle θ between the particle's momentum vector and the z axis and is given by

$$\eta := -\ln\left(\tan\frac{\theta}{2}\right). \tag{1.4}$$

For massless particles the rapidity and the pseudorapidity are equal and therefore both invariant under Lorentz boosts along the z axis. For massive particles this does not hold for the pseudorapidity. To actually measure the charge asymmetry, the sensitive variable has to be constructed in consideration of the experimental setup, like proton-antiproton collisions at the Tevatron or proton-proton collisions at the LHC.

At the Tevatron with its colliding protons and antiprotons, both the quarks and the antiquarks appear as valence quarks. As these particles are coming from different directions, it is also possible to define a general forward direction. Therefore a positive charge asymmetry leads to an excess of top quarks in the forward direction and of top antiquarks in the backward direction. This turns the charge asymmetry into a forward-backward asymmetry, which can be probed by the sensitive variable Δy, which is the difference of the rapidities of top quark and top antiquark:

$$\Delta y := y_{\mathrm{t}} - y_{\bar{\mathrm{t}}}. \tag{1.5}$$

In contrast to the Tevatron the initial state at the LHC with its proton-proton collisions is a symmetric one. Therefore no general forward direction can be defined and a new approach is needed. As there are no antiprotons present, all antiquarks occurring in initial states are sea antiquarks, which have on average a smaller momentum fraction than valence quarks. In the presence of a charge asymmetry, these momentum fractions of the initial state particles are transferred to the final state top quarks and antiquarks. As top antiquarks are related to the less energetic sea antiquarks of the initial state, a positive charge asymmetry leads to top antiquarks being produced more centrally than top quarks. This makes the charge asymmetry result in a central-peripheral asymmetry, which affects the width of the rapidity distributions of top quarks and antiquarks. Therefore a new sensitive variable $\Delta|y|$, which is the difference of the absolute values of the rapidities of the top quark and the top antiquark can be constructed:

$$\Delta|y| := |y_{\mathrm{t}}| - |y_{\bar{\mathrm{t}}}|. \tag{1.6}$$

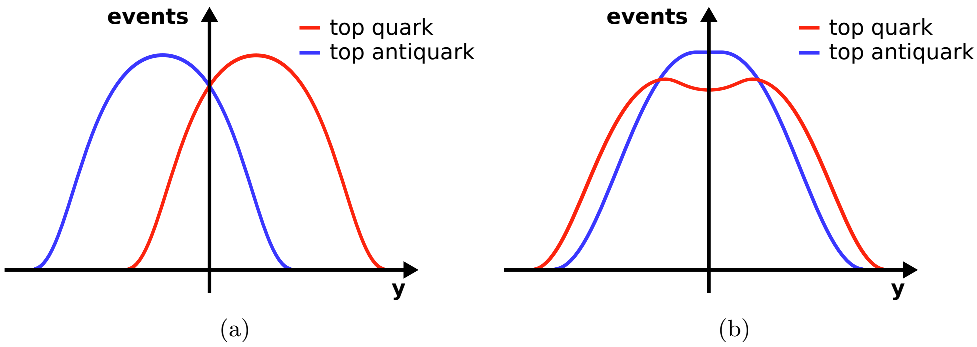


Figure 1.6.: The effect of the charge asymmetry on the y distribution, taken from [64]. In (a) the effect in proton-antiproton collisions, like at the Tevatron, is shown. With the top quark flying in the direction of the incoming proton and the top antiquark flying in the direction of the incoming antiproton, the charge asymmetry appears as a forward-backward asymmetry. In proton-proton collisions instead, like at the LHC, the same effect results in the charge asymmetry appearing as a central-peripheral asymmetry in the rapidity distributions, as shown in (b).

The qualitative differences caused by the charge asymmetry in the rapidity distributions of top quarks and antiquarks at the Tevatron and at the LHC are shown in figure 1.6.

With the so defined sensitive variables for the Tevatron and the LHC, the asymmetry itself can then be defined as

$$A_C := \frac{N^+ - N^-}{N^+ + N^-}, \tag{1.7}$$

with N^+ and N^- as the numbers of events with positive and negative signs of the sensitive variable.

The Standard Model predicts an asymmetry of $A_C^{\text{Tev,SM}} = 0.087 \pm 0.010$ [65] for the measurement at the Tevatron. For the measurement at the LHC with a center-of-mass energy of 8 TeV an asymmetry of $A_C^{\Delta|y|,\text{SM}} = 0.0102 \pm 0.0005$ [65] or $A_C^{\Delta|y|,\text{SM}} = 0.0111 \pm 0.0004$ [6] is predicted. These results have all been calculated at NLO precision including electroweak corrections. The uncertainties arise from scale variations and PDF uncertainties. The different predictions for the LHC result from different considerations of electroweak corrections. The reason for the much smaller asymmetry at the LHC is the much smaller fraction of the asymmetric quark-antiquark annihilation process and the use of a different sensitive variable.

The first measurement of the charge asymmetry in hadron collisions has been performed by the CDF collaboration in 2007 [66] with a result of $A_C^{\text{Tev}} = 0.24 \pm 0.14$. The most recent measurements in the lepton+jets channels of the Tevatron detectors using all collected data of its lifetime yield $A_C^{\text{Tev}} = 0.164 \pm 0.047$ for the CDF measurement [67] and $A_C^{\text{Tev}} = 0.106 \pm 0.030$ measured by DØ [68].

At the LHC the first measurement of the charge asymmetry was done by the CMS collaboration at a center-of-mass energy of 7 TeV using $\Delta|\eta|$ as the sensitive variable. Its result of $A_C^{\Delta|\eta|} = 0.06 \pm 0.14$ [69, 70] is compatible with the Standard Model prediction of $A_C^{\Delta|\eta|,\,\text{SM}} = 0.0136 \pm 0.0008$ [65] within its uncertainties. The latest 7 TeV analyses in the semi-leptonic decay channel at ATLAS and CMS used $\Delta|y|$ as the sensitive variable and measure an asymmetry of $A_C^{\Delta|y|} = 0.006 \pm 0.010$

for ATLAS [71] and $A_C^{\Delta|y|} = 0.004 \pm 0.010 \pm 0.011$ for CMS [72]. The combination of these two results yields an overall charge asymmetry at the LHC based on 7 TeV data of $A_C^{\Delta|y|} = 0.005 \pm 0.007 \pm 0.006$ [73], which is within its uncertainties also compatible with the corresponding Standard Model prediction of $A_C^{\Delta|y|,\mathrm{SM}} = 0.0115 \pm 0.0006$ [65].

All measured results at the LHC are compatible with a zero asymmetry as well as with the predicted asymmetry by the Standard Model. But in contrast to the large positive Tevatron results, the LHC results all show smaller values of the charge asymmetry than predicted.

1.3.2. Differential Measurements

By measuring the charge asymmetry as a function of other appropriate secondary variables, the sensitivity to different physics scenarios can be enhanced [61–63, 65, 74], improving the search for BSM physics and allowing a more exact measurement of the Standard Model asymmetry.

Because top-quark pair-production modes via new heavy particles with different couplings to top quarks and antiquarks would become apparent only at high energies, a measurement differential in the invariant mass of the top-quark pair $m_{\mathrm{t\bar{t}}}$ allows a more promising search for them. The asymmetry of the Standard Model also rises with higher values of $m_{\mathrm{t\bar{t}}}$ because of an enrichment of the asymmetric $\mathrm{q\bar{q}}$ initial state.

The charge asymmetry in the Standard Model consists of positive and negative contributions. The negative contributions arises from the interference of initial state radiation and final state radiation Feynman diagrams, which both show an additional gluon in the final state. But the transverse momentum of this extra gluon has to be balanced by the top-quark pair also having a higher transverse momentum $p_{\mathrm{T,t\bar{t}}}$. Thus a differential measurement in $p_{\mathrm{T,t\bar{t}}}$ allows a separation of the two contributions, enhancing the positive contribution for low values of $p_{\mathrm{T,t\bar{t}}}$.

As the energies of the initial state particles are symmetric, the final state particles of gluon fusion processes can be found predominantly in the central region of the detector. In contrast to that, the annihilation of quarks and antiquarks leads to a boost of the final state system in the direction of the incoming quark, because initial state quarks mostly appear as valence quarks with a larger momentum fraction than sea quarks. Therefore a measurement differential in the absolute value of the top-quark pair's rapidity $|y_{\mathrm{t\bar{t}}}|$ is able to distinguish between the different production processes, enriching the asymmetry for high rapidities.

A differential measurement of the charge asymmetry as a function of these variables has been performed by CMS [72] for 7 TeV data, which is in good agreement with the Standard Model predictions.

1.3.3. Fiducial Measurement

All the measurements described above refer to the charge asymmetry in the full phase space. But for its determination usually an extrapolation from the selected events in the acceptance region of the used particle detector to the full phase space is needed. To reduce this amount of extrapolation, a so-called *fiducial measurement* is performed in this thesis. This kind of measurement only refers to a well defined region in phase space, the *visible phase space*. It should be in agreement with the

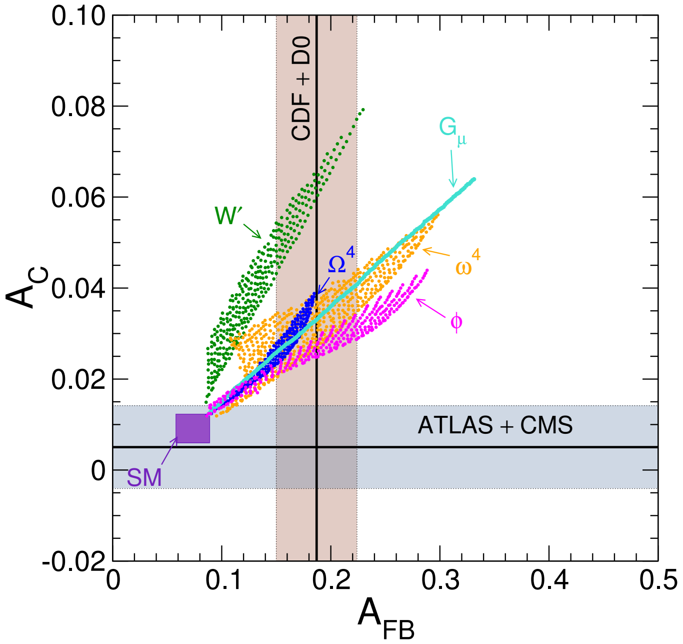


Figure 1.7.: Comparison of BSM theory predictions for the charge asymmetry with measurements performed at CMS and ATLAS at 7 TeV and Tevatron at 1.96 TeV, adapted from [63]. The horizontal axis shows the Tevatron charge asymmetry, while the vertical axis shows the charge asymmetry for the LHC. The central values of the measurements are indicated by a solid line with dashed lines for the corresponding uncertainties.

detector's acceptance region or the analysis' event selection and be defined in a way to allow predictions from the Standard Model.

1.3.4. Theories Beyond the Standard Model

Several models exist to explain the asymmetry measurements of the Tevatron with new physics [61–63]. But these theories, which generally introduce a new extra particle, have to be consistent with the precise $t\bar{t}$ cross section measurements and the measured charge asymmetry at the LHC. Examples for these new possible particles are light neutral color-octet vector bosons G_μ, a color-singlet scalar doublet ϕ with hypercharge $-1/2$ exchanged in the t channel, and a scalar particle with an electric charge of $4/3\,e$ exchanged in the u channel that can be either a color sextet (Ω^4) or a color triplet (ω^4). Because the color-singlet vector bosons Z′ and W′ predict large inclusive asymmetries at the LHC, they have been disfavored by the recent measurements. A comparison of these different models and their predicted asymmetries with the Tevatron and 7 TeV CMS measurements is shown in figure 1.7.

2. Experimental Setup

All the matter we have around us only consists of particles of the first generation of the Standard Model, like electrons or up and down quarks, which form protons and neutrons. Even particles generated in the Earth's atmosphere in interactions with high energetic cosmic rays mostly decay into members of the first generation until they reach the Earth's surface.

But after Einstein's famous equation $E = mc^2$ [75], new particles can be created if enough energy is concentrated in one point. Therefore, to study heavy leptons, bosons or quarks like the top quark, huge machines are needed to produce them and to analyze them.

In particle accelerators many kinds of them can be produced by collisions of hadrons or leptons with a very high center-of-mass energy. But because all sorts of particles will be produced there, these all have to be understood in detail to analyze the desired ones.

The most powerful accelerator today is the Large Hadron Collider (LHC) [76] with its many particle detectors at CERN, where protons collide with a center-of-mass energy of $\sqrt{s} = 8$ TeV.

2.1. The Large Hadron Collider

The LHC has a circumference of 27 km and is located in the tunnel where the Large Electron Positron Collider (LEP) was located before, about 100 m underground. To keep each of the proton beams with an energy of 4 TeV on a circular track, 1232 superconducting magnets with a magnetic field of 4.8 T are used.

The LHC is divided into eight sections, each of them containing one possible collision point P1 to P8. In four of these points there are particle detectors: The ATLAS detector [77] at P1 and the Compact Muon Solenoid (CMS) detector [78,79] at P5 are the two general purpose particle detectors at the LHC. Their aim was the discovery of the Higgs boson and is the search for physics beyond the Standard Model. At P2 the ALICE detector [80] is analyzing lead-ion collisions, in which a quark-gluon plasma like right after the big bang is produced. At P7 the LHCb experiment [81] is analyzing the decays of bottom quarks, trying to observe CP-violating decays.

The other four sectors contain the infrastructure to keep the LHC running. The RF system at P4 is needed to accelerate the particles up to the desired beam energy and to compensate energy loss due do synchrotron radiation. The beam dump at P6 is used to safely get rid of the beam after a run or when it becomes unstable. Additional cleaning structures are at P3 and P7 to get a compact focused beam and avoid quenching. The general layout of the sections of the LHC together with the chain of all the preaccelerators is shown in figure 2.1

The operation of the LHC started in 2008, in 2010 a center-of-mass energy of 7 TeV was reached. In 2012 the energy was increased to a center-of-mass energy of 8 TeV until the first long shutdown from 2013 to 2015. After that the LHC is

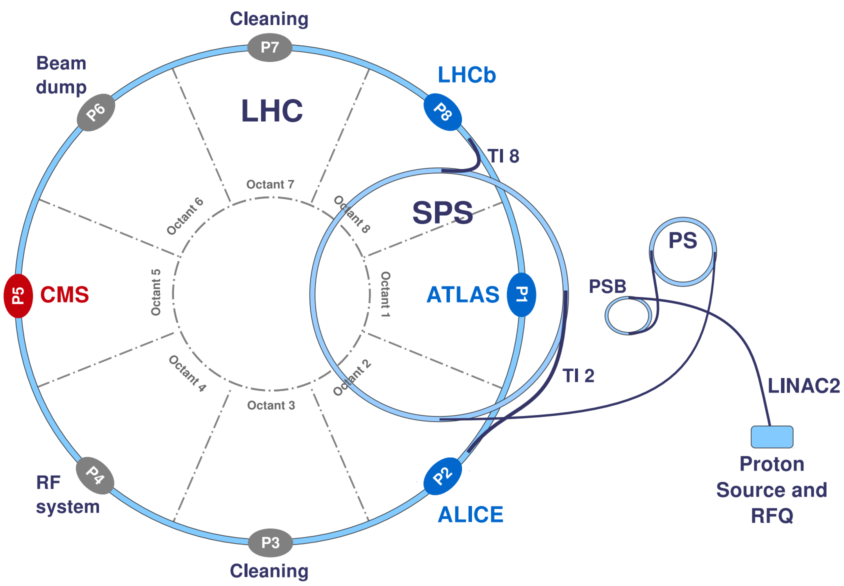


Figure 2.1.: Schematic overview over the preaccelerators and the subdivisions of the main accelerator ring at CERN, taken from [82]. The protons of the proton source are preaccelerated in the LINAC2, the PSB, the PS and the SPS. From the SPS the bunches are injected into the main ring in opposite directions through the transfer lines Tl 2 and Tl 8, resulting in a opposite direction of revolving. The main ring of the LHC is divided into eight octants, each one containing one possible collision point P1 to P8. At four of these collision points particle detectors have been installed.

planned to operate on a center-of-mass energy of 13 TeV.

The protons for the LHC originate from a duoplasmatron, where hydrogen gas is injected and the electrons and protons are separated from each other [83]. These protons are accelerated in the linear accelerator LINAC2 up to an energy of 50 MeV [84] and injected into the Proton Synchrotron Booster (PSB), to reach an energy of 1.4 GeV [85]. Then the protons continue their way in the accelerator chain to the Proton Synchrotron (PS) [86], which brings them to an energy of 26 GeV, followed by the Super Proton Synchrotron (SPS), which accelerates the protons up to an energy of 450 GeV to be ready to be injected into the main ring of the LHC [87]. This injection into the LHC happens in two directions to get two beams running in opposite directions. At last the superconducting dipole magnets of the LHC are ramped up to reach the final beam energy of 4 TeV, so that the protons can be brought to collisions at the collision points of the four particle detectors.

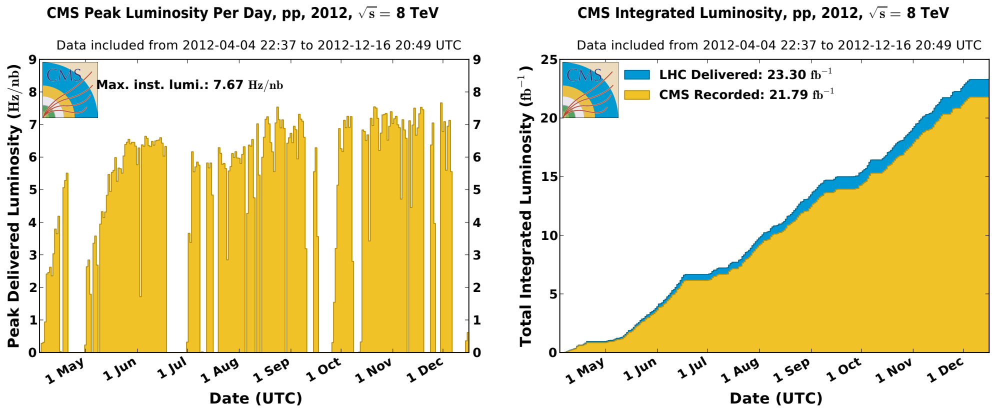
The relationship between a measured event rate $\dot{N}$ and the cross section σ of a specific process is given by

$$\dot{N} = \mathcal{L}\sigma, \tag{2.1}$$

where $\mathcal{L}$ is the luminosity, which is a measure for the amount of particle collisions. For an accelerator with a revolution frequency f, the number of bunches per beam n, the numbers of particles in one bunch of beam A or B, N_A and N_B, and the gaussian height and width of a beam's profile of σ_x and σ_y, the luminosity is given by

$$\mathcal{L} = fn\frac{N_\mathrm{A}N_\mathrm{B}}{4\pi\sigma_x\sigma_y}. \tag{2.2}$$

The maximum instantaneous luminosity per day and the development of the integrated luminosity over time of the LHC operation in 2012 are shown in figure 2.2.



(a) (b)

Figure 2.2.: Luminosity profile in the 2012 proton operation at the LHC during stable beams at a center-of-mass energy of 8 TeV [88]. Shown in (a) is the maximum instantaneous luminosity per day delivered to CMS. In (b), the integrated luminosity over time is shown. The blue shape represents the integrated luminosity delivered by the LHC, while the yellow shape indicates the luminosity that was recorded by the CMS detector.

2.2. The Compact Muon Solenoid Detector

The Compact Muon Solenoid (CMS) detector [78, 79] is – beside ATLAS – the second general purpose detector at the LHC, located at the interaction point 5.

Although the CMS detector has a diameter of 15 m and a length of 22 m, compared to ATLAS it still can be called "compact".

Starting from the beam pipe it consists of the tracking system with silicon strip and pixel detectors, the calorimetry system with the electromagnetic and the hadronic calorimeter, enclosed by the superconducting magnet and the muon system embedded in the iron return yoke. All these detector systems completely enclose the interaction point to be able to detect all particles originating from there. Despite the compact construction and due to the iron return yoke the detector has a total weight of 14 000 t. A schematic overview over CMS is shown in figure 2.3.

The coordinate system used in CMS is a right-handed one with the origin at the nominal interaction point. The x axis is pointing to the center of the LHC, the y axis is pointing up and the z axis is pointing along the counterclockwise beam direction. The azimuthal angle ϕ is measured in the x-y plane and the polar angle θ is measured from the positive z axis. Instead of this angle usually the pseudorapidity $\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$ is used.

2.2.1. Tracking System

The tracking system of CMS [90, 91] consists of silicon pixel detectors located directly around the beam pipe, enclosed by silicon strip detectors. It allows the detection and reconstruction of the track of charged high energy particles. Because of the deflection of the particles in the magnetic field of 3.8 T the momenta and the charges of these particles can be determined. The track measurements are accurate to 10 µm, which allows an extrapolation of the particle tracks back to the interaction point to search for secondary vertices, which can originate from possible decays of relatively long-living bottom quarks.

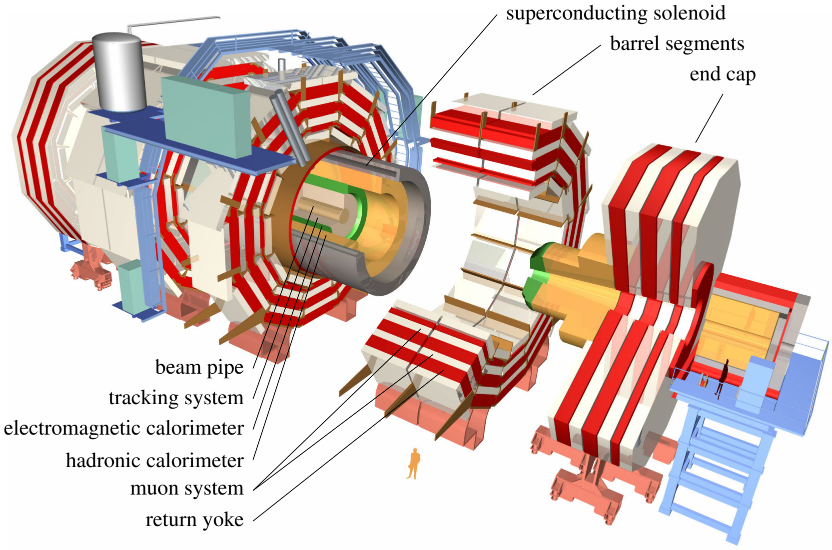


Figure 2.3.: Schematic overview of the CMS detector and its components, adopted from [82, 89]. The beam pipe is encased by the tracking system, the electromagnetic and hadron calorimeters, the superconducting solenoid and the muon system, which is embedded in the iron return yoke. The different subdetectors are layered around the interaction points like the layers of an onion. Also the segmented structure of CMS can be seen, which allowed construction on the surface and makes easy access for maintenance possible.

Silicon Pixel Detector

The silicon pixel detector is the detector component closest to the interaction point: The first cylindrical layer is at 4 cm, the second at 7 cm and the third at 11 cm. Each layer consists of silicon pixel sensors each of which measures 100 µm by 150 µm. Overall there are 65 million pixels, which results in an equal number of channels that have to be read out.

Silicon Strip Detector

The silicon strip detector consists of ten layers surrounding the pixel detector. These layers are arranged in barrels and endcaps as shown in figure 2.4, to cover all possible directions of outgoing particles. All the layers contain 15 200 modules with a total of 10 million channels. Some of these modules are stereo modules, which consist of two layers of silicon strips with a slightly different angle of orientation. Despite only having information of the silicon strips in two dimensions, the overlap of these strips makes the measurement of all three coordinates of specific points on a particle's track possible.

2.2.2. Calorimetry System

The calorimetry system of CMS consists of two different hermetic components to absorb and measure the energy of produced particles in all directions as shown in figure 2.5: An electromagnetic calorimeter (ECAL) [92, 93] for electromagnetically interacting particles like electrons, positrons and photons, followed by a hadron calorimeter (HCAL) [94] for particles interacting via the strong force like charged and neutral hadrons.

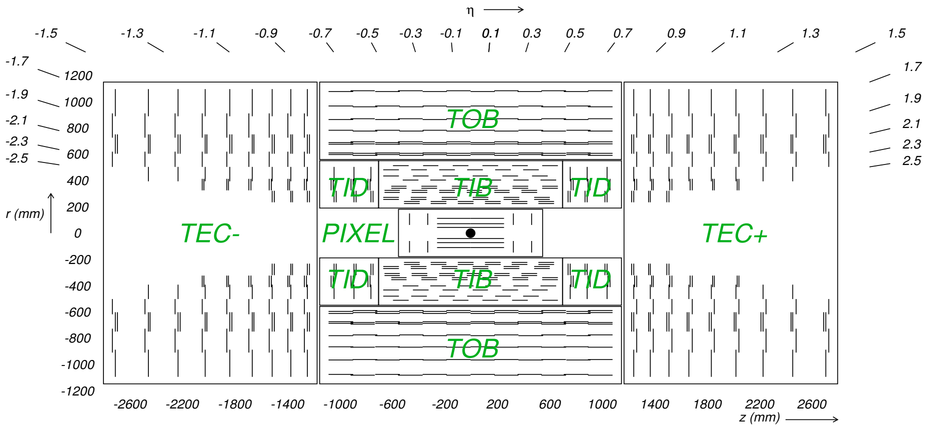


Figure 2.4.: Schematic longitudinal overview over the CMS tracking system with the CMS coordinate system [78]. The interaction point, shown as black dot, is enclosed by pixel and strip detectors. The strip detector is divided into barrels, with the Tracker Inner Barrel (TIB) and the Tracker Outer Barrel (TOB) in the central region, and endcaps, with the Tracker Inner Discs (TID) and the Tracker End Caps (TEC) on each side. The detector modules are shown as lines, stereo modules as double lines.

The exact measurement of the deposited energies in the different directions of ϕ makes it possible to vectorially sum up these transverse energies. Because of momentum and energy conservation, the resulting vector would be zero if the energies of all produced particles were considered. But this vector often does not vanish due to particles not being detected in the calorimeters, like neutrinos or possible dark matter candidates. So with this vector, the vector of the so-called missing transverse energy can be constructed, pointing in the opposite direction of the vectorial sum of the transverse energies.

The absorption of electrons and photons in the calorimeter material can be described as a particle shower consisting of alternating sequences of pair-production and bremsstrahlung. For a particle with the starting energy E_0 this leads to an exponential decrease of the shower energy $E(x)$ of

$$E(x) = E_0 \cdot \exp\left(-\frac{x}{X_0}\right), \tag{2.3}$$

where x is the depth of the shower in the absorber material. The parameter X_0 is called radiation length and only depends on the absorber material. Because an absorber usually has a depth of several radiation lengths, this quantity is used to measure these depths. For hadrons instead the nuclear interaction length λ_{I} is used.

Electromagnetic Calorimeter

The electromagnetic calorimeter of CMS is made up of lead tungstate crystals ($PbWO_4$). With a density of 8.3 g/cm^3, resulting in a radiation length of $X_0 =$ 0.89 cm, it is a good absorber for electromagnetic particle showers. Because it also is transparent and a scintillating material, it is an excellent choice for a homogeneous electromagnetic calorimeter.

The cylindrical barrel in the central region of the electromagnetic calorimeter consists of 61 200 of these crystals, each with a length of 25.8 radiation lengths X_0.

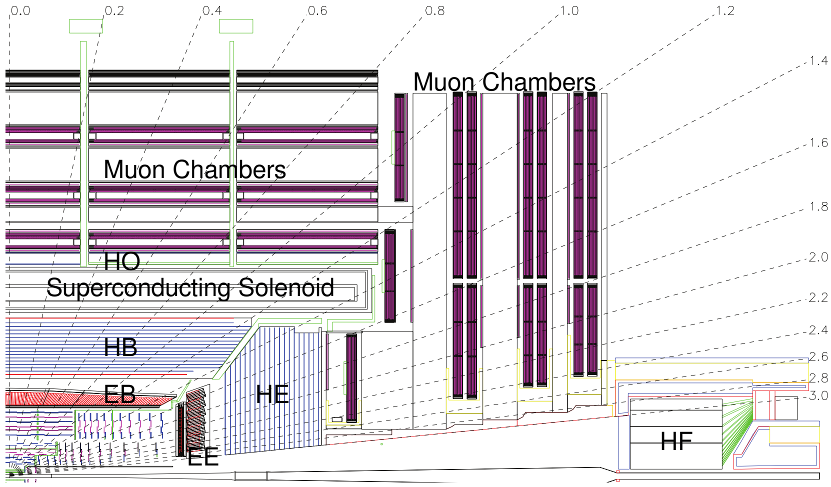


Figure 2.5.: Overview of the calorimeters and the muon system of the CMS detector [79]. The calorimeters, like most of the components, are separated into a barrel part and an endcap part. The CMS tracker (shown in figure 2.4) is encased by the Electromagnetic Barrel (EB) and the Hadron Barrel (HB) calorimeters in the center region and by the Electromagnetic Endcap (EE) and the Hadron Endcap (HE) calorimeters on each side. These calorimeter systems are enclosed by the superconducting solenoid, followed by the Hadron Outer (HO) calorimeter. All these detector layers are surrounded by the muon system, which is embedded in the iron return yoke for the magnetic field. Close to the beam pipe and far from the interaction point the Hadron Forward (HF) calorimeter is placed, which measures the energies of hadrons with large pseudorapidities. Different values of these pseudorapidities are indicated by dashed lines.

The endcaps at both ends are made up of 14 648 further crystals with a length of 24.7 X_0.

Hadron Calorimeter

The hadron calorimeter is a sampling calorimeter consisting of brass absorbing layers and plastic scintillator layers in between. In the brass layers, particles generate hadronic showers with a nuclear interaction length of $\lambda_\mathrm{I} = 16.42$ cm. This shower produces light in the plastic scintillator layers, which is read out as a signal proportional to the deposited energy.

The barrel part of the hadron calorimeter has a depth of 5.82 nuclear radiation lengths λ_I, the endcaps roughly 10 λ_I.

2.2.3. Muon System

The muon system is one of the most important parts in CMS because muons are an important signature for many physics analyses. Therefore the detection and accurate measurement of muons had a high priority in the detector design. Because muons are minimal ionizing particles that can pass through much material without further interactions, as a first approximation everything detected outside the calorimeters can be considered as muons.

The muon system itself [95] consists of three different technologies, all embedded in the iron return yoke: drift tubes (DTs) around the solenoid, cathode strip chambers (CSCs) in the endcaps, and resistive plate chambers (RPCs) in both regions.

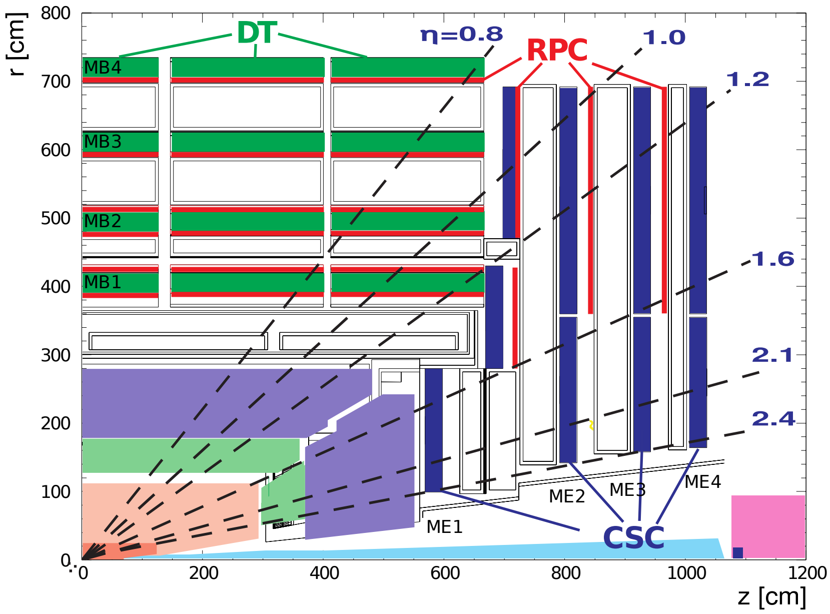


Figure 2.6.: Overview of the arrangement of the muon system of the CMS detector [78]. In the four stations MB 1 to MB 4 of the barrel region drift tubes (DT) and resistive plate chambers (RPC) are used for muon detection, while in the four endcap muon stations ME 1 to ME 4 RPCs and cathode strip chambers (CSC) are used. Different values of η are indicated as dashed lines.

An overview of the arrangement of the muon system is shown in figure 2.6.

2.2.4. Trigger System

When CMS is performing at its peak, proton-proton interactions will happen at a rate of 40 MHz. Additionally, there are about 20 simultaneous collisions expected per interaction. The huge amount of data produced during this operation of all subdetectors can neither be read out nor be stored on disks at this rate. Therefore multiple levels of triggers select only events which seem to look interesting and bring down the event rate to a more reasonable value.

The level one trigger L1 [96] is implemented in programmable hardware and looks for simple signs of interesting physics, reducing the event rate to 0.1 MHz. Then the high-level trigger (HLT) [97], which is implemented completely in software on a computing farm, uses more detailed information of the events to look for specific signatures. This finally reduces the event rate to 100 Hz, in which the events are handed over to the computing infrastructure.

2.2.5. Computing Infrastructure

Despite the reduced event rate due to the trigger system there is still a lot of data being recorded. To cope with that a tiered computing infrastructure was established, the so-called LHC computing grid [98]. The architecture of these tiers is shown in figure 2.8.

The Tier-0 computing site is located at the CERN research facility. There the first event reconstruction of the so-called RAW data from the trigger system is

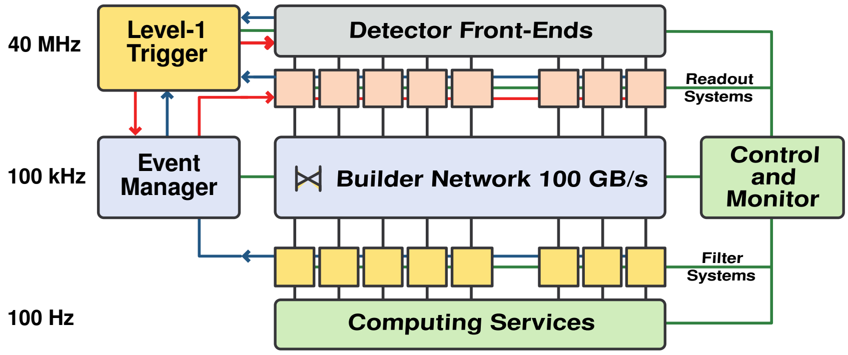


Figure 2.7.: Architecture of the CMS data acquisition (DAQ) system [79]. The level one trigger L1 reduces the event rate from 40 MHz to 100 kHz. Only these events are read out and passed to the builder network, where they are processed in parallel and forwarded to the HLT filter system. There the decision is made whether to keep them or not, which reduces the event rate to 100 Hz.

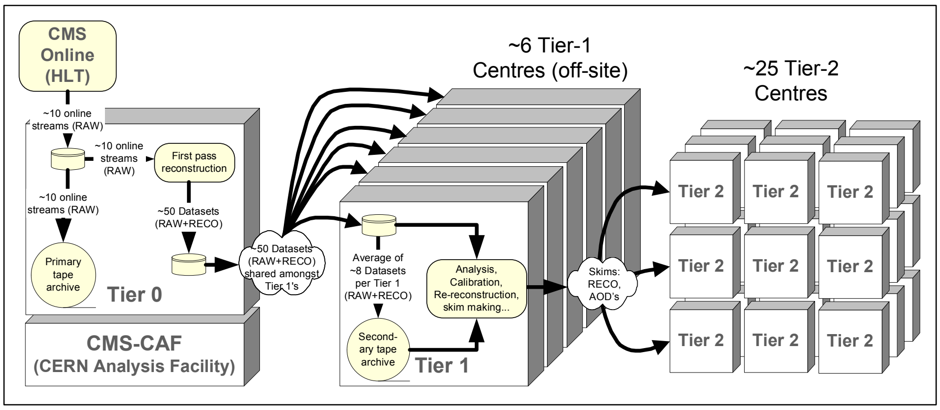


Figure 2.8.: Architecture of the CMS workflow in the LHC computing grid [79]. The flow of recorded events starts at the single Tier-0 site at CERN, where the RAW data from CMS is stored and used to produce RECO datasets. The RAW and RECO datasets are then transferred to and permanently stored at Tier-1 sites, which also hold all available AOD datasets. The Tier-2 centers, only storing AOD datasets, provide CPU and storage resources for the individual analyses.

performed to gain the RECO data set, which already contains the reconstructed high-level physics objects and the detector hits used to reconstruct them.

These datasets are archived on Tier-0 and transfered to the Tier-1 computing centers. There the Analysis Object Data (AOD) are produced from the RECO datasets and distributed between them. These AOD datasets contain the high-level physics objects plus a summary of other RECO information needed for specific analysis actions. They are small enough to be shared with the many Tier-2 sites which finally provide the computing infrastructure and storage capacities to the different analysis groups.

3. Generation, Simulation and Reconstruction of Collision Events

In particle colliders lots of secondary particles are produced per event, which result in even more electric signals in all the subdetectors, which have to be read out. These raw signals have to be translated back to information about the physical objects to determine their tracks and energies by reconstruction algorithms linking all the information of the subdetectors.

But the reconstructed particles cannot be compared to theory directly. Instead, theory predictions are implemented as so-called *Monte Carlo* (MC) simulations. These allow approximations for the intricate integrals in theory calculations by drawing large amounts of randomly generated numbers which follow the calculated probability distributions. Afterwards a simulation of the CMS detector is applied to model its response to the generated particles. So both the data and the MC simulations can undergo the same analysis procedure resulting in comparable distributions with the advantage of the available truth information for simulated events.

In this chapter an overview of the process of Monte Carlo event generation is given, including the used production programs. Furthermore the different steps of the reconstruction procedure are described.

3.1. Generation of Events

The generation of MC events requires the proton-proton collision events to be factorized into smaller subprocesses. At first the hard process is generated, then gluon radiation is added. In the following hadronization process color-neutral particles are formed and the decays of remaining unstable particles are simulated. The sequence of these processes is shown in figure 3.1.

Hard Scattering Process

The hard scattering process can be described by Feynman diagrams, from which the probability of the interaction can be estimated by perturbative calculations with the matrix element method. This is possible because the energy scale is large enough for the coupling α_s of the strong interaction to become small and allow perturbative theory calculations. The initial state particles are given by the PDFs of the incoming protons, shown in figure 1.3. The decays of particles with a short lifetime – like top quarks or W bosons – are also included already.

Parton Shower

In the parton showering process the radiation of accelerated color charges is calculated, which results in initial state radiation (ISR) and final state radiation (FSR). Radiations with high momentum transfers correspond to low values of α_s which

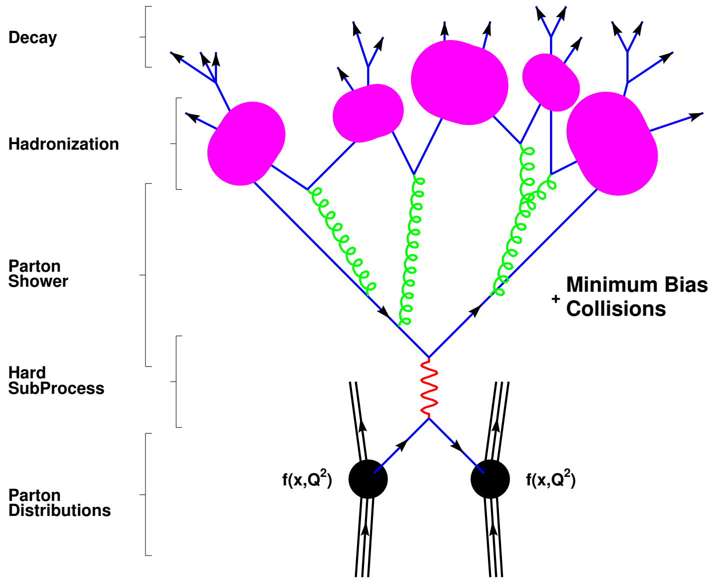


Figure 3.1.: Overview of the different subprocesses of Monte Carlo event generation, taken from [99]. In the hard process the interactions of the partons within the protons are simulated, with their fractions given by the PDFs of the protons. Afterwards additional parton showers are calculated. In the hadronization process colorless particles are formed and the decay of unstable particles is simulated.

allow perturbative calculations. For radiations with lower momentum transfers the Parton Shower (PS) approach is used, which uses simplified models for the kinematics of the interactions. There the branchings of the partons are calculated using the DGLAP QCD evolution equations [100–102], where the probability of a gluon radiation is calculated using the Altarelli-Parisi splitting functions.

Hadronization

At lower energy scales and larger distances between the particles the coupling α_s of the strong force increases, which makes calculations based on perturbative theory become invalid. Therefore the hadronization process is based on pure phenomenological models. These models describe the formation of colorless particles, with their parameters tuned to match the observed data. One established model is the *Lund string model* [103]. It describes the strong force as color-flux strings spanned between two color-singlet particles. With increasing distances of the particles the tension of the strings increases until their energy is high enough for the color-flux strings to break up and create a quark-antiquark pair. This procedure continues until the energy is too low to create new pairs and only colorless particles are left. Afterwards the decays of remaining unstable baryons or mesons are simulated according to their known branching ratios.

Underlying Event and Pile-Up

In the measurement of resulting particles from a specific parton interaction, also the detection of particles from additional parton interactions has to be considered. To keep the simulated events comparable to measured data, an additional simulation of these effects is performed, which uses phenomenological models adjusted to measured data.

Additional interactions of other protons are called *in-time pile-up* when originating from protons of the same bunch crossing, or *out-of-time pile-up* when originating from earlier or later bunch crossings. At a center-of-mass energy of 8 TeV, an average of 21 interactions per event was observed by the CMS detector.

The hadronization of the color-charged remnants of the colliding protons as well as multiple interactions of the partons are called *underlying event*.

3.1.1. Monte Carlo Event Generators

As described before the generation of simulated events consists of several steps based on different algorithms. For these processes several Monte Carlo generators exist, with some of them focused on specific generation steps and others being general purpose generators.

For the simulation of the hard process based on the perturbative matrix element, leading order generators like MadGraph/MadEvent [104–106] and Alpgen [107] exist, which also simulate ISR and FSR by leading order matrix element calculations. Generators including next-to-leading-order (NLO) corrections are for example Powheg [7] and MC@NLO [108, 109], but they only generate higher order radiations resulting from perturbative calculations.

To generate the subsequent parton shower and the hadronization processes parton shower generators like Pythia [110] and Herwig [111] are used. Although they are general purpose event generators, they only provide leading order matrix element calculations for $2 \to 2$ processes and are therefore only used for shower generation purposes.

These two different kinds of generators have to be interfaced using matching schemes like CKKW [112,113] and MLM [114] for leading order generators and the Powheg method [8,9] or the MC@NLO matching scheme [108, 109] for next-to-leading order generators to achieve a complete event simulation.

In the following paragraphs the MC generators involved in the generation of the samples which are used in this thesis are described.

Powheg

The Powheg Box [7], or just Powheg, is a matrix element generator providing NLO QCD calculations. It can be interfaced to Pythia or Herwig as shower generators, for example, and uses the Powheg method [8, 9] matching scheme. The main point of this scheme is that the hardest radiation of each event is always generated by Powheg. Therefore subsequent shower generators are required to simulate only softer radiations than generated by Powheg. In contrast to MC@NLO, only positive event weights are produced with this method. Also the properties of the partons before the showering process are usable for physics analyses, as long as the analyses refer to the NLO picture only.

MadGraph and MadEvent

MadGraph [104, 106] is a matrix element generator which calculates the amplitudes for a given process and its subprocesses providing additional information for phase space integration. With this information the tree-level generator MadEvent [105] then generates the simulated events, which are stored in the Les Houches event format [115]. These events can then be processed by subsequent parton shower generators with the MLM matching scheme [114] applied.

MC@NLO

The MC@NLO package includes a standalone matrix element generator, which provides NLO QCD calculations, as well as the MC@NLO matching scheme [108, 109]. In this scheme the contributions of the calculations which are also simulated by the following shower generator are subtracted in advance. Therefore the matrix element generator explicitly needs to support the matching to a given shower generator, which is the case for the generators of the Herwig family. Due to the described correction method, MC@NLO applies negative weights to a small amount of events, which can cause unphysical results in low-statistics regions. Because this correction also modifies the properties of the produced partons of the hard process, only the particles and jets after showering and not at parton level are suitable for the analysis presented in this thesis.

Herwig

Herwig [116] is a general-purpose event generator with special emphasis on the accurate simulation of QCD radiation. It is able to simulate the elementary hard subprocess, initial- and final-state parton showers, the decay of heavy objects like the top quark, multiple scattering as the dominant component of the underlying event as well as the hadronization process, followed by hadron decays. In this thesis it is only used as a shower generator.

Pythia

Pythia [110] also is a full event generator. It simulates the hard scattering process and has many implementations for the calculation of soft interactions. Also ISR and FSR, hadronization processes, the decay of unstable particles and the underlying event can be simulated. Despite Pythia being able to handle the complete event generation process, it is more often used as a shower generator following after the hard scattering process generated by MadGraph/MadEvent or Powheg.

Tauola

An accurate simulation of τ-lepton decays including spin correlations is performed by Tauola [117], which is used as the last step in event generation for all used MC samples.

3.1.2. Detector Simulation

These previously described Monte Carlo event generators all perform only a simulation of the processes as they happen in a vacuum. In reality, all the generated particles pass through the different materials of the CMS detector and are affected by the magnetic field of the superconducting magnet. To simulate the effects resulting from interactions with the detector material like multiple scattering, ionization, bremsstrahlung and electromagnetic and hadronic showering, a detailed model of the CMS detector is needed. That simulation also includes the processes in the subdetectors, which result in measurable signals comparable to real raw data taken from the experiment.

The GEANT4 [118] toolkit is able to perform such a full detector simulation with a high precision. But to reduce the needed computing power by about a factor of 100 there is also a fast simulation [119] available. This simulation uses more simplified models and algorithms and is used if a huge number of generated events is needed for a specific process.

3.2. Reconstruction of Events

To reconstruct all the resulting physical objects of a hard scattering event, the raw electric signals of the CMS detector have to be interpreted by various algorithms. These algorithms are implemented in the CMS SoftWare (CMSSW) [78] framework.

In CMS, the powerful Particle Flow [120–122] algorithm is used for reconstruction. In this algorithm the identification and reconstruction of all stable particles is done individually by the combination of the information of all detector systems. At first, the trajectories of charged particles are gained from the tracking system, which also yields the coordinates of the interaction point. Then these tracks are linked to energy deposits in the calorimeters and to hits in the muon system. After further specific procedures depending on the types of physical objects are applied, the algorithm provides a complete list of all stable particles and their precisely measured energies and momenta.

In the following sections the reconstruction algorithms are described which are applied to observed data and to generated Monte Carlo events.

3.2.1. Reconstruction of Tracks

The trajectories of charged particles traversing a magnetic field can be described as helices. By interacting with the detector material of the tracking system they generate energy deposits called *hits* in the different layers of the tracker along their way.

For the reconstruction of the trajectories from these hits the Combined Track Finder [123] is used, which iteratively applies the Kalman Filter (KF) track finder [124] performing local fits. The reconstruction first starts by looking for a seed consisting of multiple hits close to the beam pipe combined with a beam spot constraint. Then the Kalman filter extrapolates this track candidate by adding more hits from further layers and adjusting the approximation. In this procedure the magnetic field as well as effects from multiple scattering and energy loss are taken into account. All the trajectory candidates are fitted to the hits in parallel and are compared to each other to avoid overlaps.

After the local fits of the Kalman Fitter, a global fit is performed for the appropriate hits to improve the precision of the measurement of the track parameters.

3.2.2. Reconstruction of Vertices

The information about the exact locations of the interaction vertices is important in two ways: First to provide the number of pile-up interactions in an event, and secondly to allow a more precise track fitting by also taking the primary vertex into account.

The vertex finding process itself consists of two steps: At first tracks are grouped together to form possible vertex candidates, then the best vertex parameters for given tracks are determined by an appropriate χ^2 minimization. The vertex with the highest sum of its tracks squared transverse momenta p_{T}^2 is then named the primary vertex.

Instead of the described procedure, which is used in offline analysis, the HLT uses a simpler algorithm, where the x and y axes are fixed to the beam pipe and only the z axis is determined.

3.2.3. Reconstruction of Electron Candidates

Being charged particles, electrons generate hits in the tracking system and deposit their energy in the electromagnetic calorimeter. Because of their large electric-charge-over-mass ratio they lose much energy through bremsstrahlung. These radiated photons are spread along the ϕ direction due to the helical track in the magnetic field, leading to energy deposits in the ECAL. Taking this into account, superclusters of energies larger than 1 GeV are formed from these signals, which act as starting points for the track fitting procedure for electron candidates [125]. Due to the nonlinear effects caused by bremsstrahlung, a nonlinear Gaussian sum filter [126, 127] is used to fit the electron tracks. Resulting tracks matching a supercluster that are in $|\eta| < 2.5$ with $p_{\mathrm{T}} \geq 5$ GeV/c correspond to primary electron candidates. Misidentification of charged hadrons or other objects as electrons is avoided by applying additional quality criteria [128] concerning the bremsstrahlung and electromagnetic showering behavior of electrons.

3.2.4. Reconstruction of Muon Candidates

Muon candidates in CMS consist of hits in the muon system and in the tracking system. Depending on whether the reconstruction focuses on the muon system, on the tracking system or on both, three different classes of muon candidates [78, 129] can be distinguished:

- For *stand-alone muons* the hits in the DT, CSC and RPC of the muon system are used for a reconstruction of their tracks with the Kalman Filter, corresponding to the reconstruction in the tracking system. After the found track reaches the edge of the muon system an additional KF fit is applied backwards to the interaction point to accurately measure the track parameters.

- *Tracker muons* are reconstructed from trajectories found in the tracking system, which are extrapolated to the muon system considering energy loss and the magnetic field and matched to corresponding hits there. This allows the reconstruction of muons with only low values of transverse momentum, which produce not enough hits in the muon system to be recognized as stand-alone muons.

- The reconstruction of *global muons* starts with stand-alone muons, whose tracks are extrapolated to the tracking system. There they are combined with track candidates of the tracking system, and global fits including all hits of the muon candidate are performed to yield the most accurate reconstruction.

To avoid misidentification, similar to the electron reconstruction additional quality criteria are applied to the tracks of the reconstructed muons. For example, hadrons with a large transverse momentum which reach the muon system, so-called *punch-through hadrons*, can be misidentified as muons. But they also deposit a large amount of energy in the ECAL and the HCAL. Therefore the total deposited energy in a cone around the track, the χ^2 value of the fit and the number of invalid hits are used as discriminators against misidentification.

3.2.5. Reconstruction of Photons and Hadrons

After the final reconstruction of electron and muon candidates their corresponding tracks are removed. The remaining tracks are then linked to energy deposits in the ECAL and the HCAL along their paths. Depending on the discrepancy of the momentum determined by the particle's track and the energy deposited in the ECAL or HCAL, different approaches are chosen.

If the measurements of the calorimeter and the tracker are compatible, a charged hadron is reconstructed and the calorimeter information is included in the fit for the track, improving the accuracy. When the calorimeters measure a lower energy than expected by the tracker's momentum measurement, the particle is assumed to be a pion and its energy and momentum are reconstructed directly from the track.

An excess of the measured energy in the calorimeters compared to the momentum measurement leads to the reconstruction of neutral hadrons and photons. If this excess is smaller than the deposited energy in the ECAL, only a single photon is reconstructed. Otherwise, if the excess is larger than the ECAL energy deposit, also a neutral hadron is reconstructed in addition to the photon. This is justified by the observation, that 3% of a jet's energy deposit in the ECAL are caused by neutral hadrons, whereas 25% are caused by photons.

Finally the remaining clusters in the ECAL and in the HCAL are reconstructed as further photons and neutral hadrons, respectively.

3.2.6. Reconstruction of Jets

Single quarks and gluons are color-charged particles, so their direct observation is impossible due to the QCD confinement. Instead, they fragment into colorless hadrons, which form narrow cones called *jets*. The evolution of such a jet is shown in figure 3.2.

The reconstruction of a jet aims to identify its origin and reconstruct the four-vector of the initial particle as precisely as possible. This is achieved by various clustering algorithms, which use a specific type of objects like reconstructed particle tracks or calorimeter clusters as an input.

These clustering algorithms need to fulfill two requirements. Additional soft radiations should not affect the number of reconstructed jets, which is called *infrared safety*. These soft radiations can originate from the underlying event and from gluon emissions or they can just be a result of detector noise and should therefore not influence the number of jets. Also the splitting of a particle into two particles at a small angle should not affect the jet output, which is called *collinear safety*. This is essential because gluon splitting shows this behavior, which is common in showering processes.

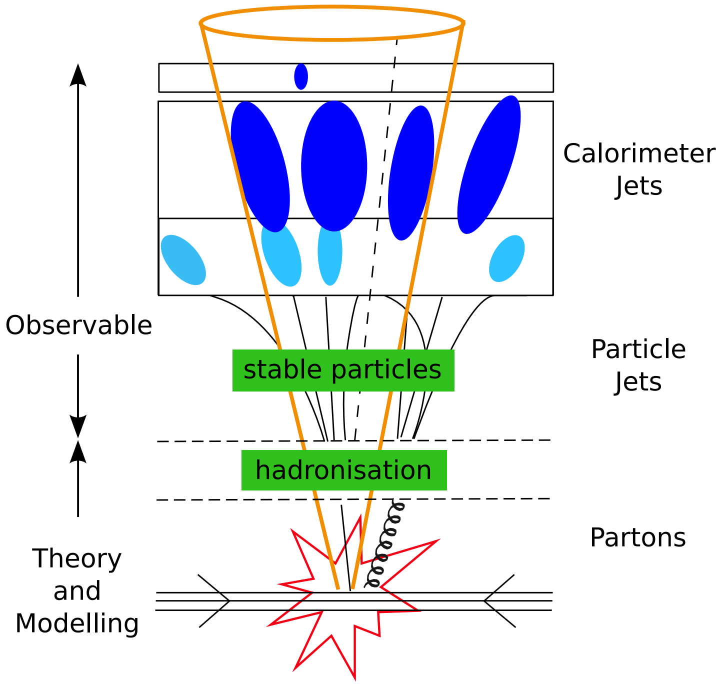


Figure 3.2.: The evolution of a jet, taken from [82]. After hadronization colorless stable particles are formed in jets, which deposit their energy in the ECAL, marked in light blue, and in the HCAL, marked in dark blue. These energy deposits can be used as input for the jet reconstruction algorithms.

Clustering Algorithms

For the clustering of the four-momenta of physical objects two important approaches exist. Cone-type algorithms make up one category, the other category consists of sequential clustering algorithms.

Cone-type algorithms cluster all objects within a given cone with a fixed radius R in the η-ϕ plane. The iterative cone (ICONE) [78] algorithm selects the particle with the highest transverse energy E_{T} as a seed. The transverse energy is given by

$$E_{\mathrm{T}} = E \cdot \sin\Theta, \tag{3.1}$$

with the particle's energy E and the azimuthal angle Θ of its momentum vector. Then a proto-jet is formed by the particles within a cone with radius R around the seed. In the next step this proto-jet is used as the new seed and this procedure is repeated until a stable jet is found, which is then removed from the list. After that the whole algorithm is repeated until no particles are left over. Because which jet is chosen as the seed depends on the particles energy, this algorithm is not collinear safe. Therefore the seedless infrared-safe cone (SISCone) algorithm [130] was developed, which uses a split and merge method and finds all stable cones in a sufficient amount of time.

Sequential clustering algorithms instead are not limited to geometrical shapes and are infrared and collinear safe by construction. Also they are easier to implement for theory predictions and have a low computational cost. For these algorithms, the two distances $d_{i,j}$ and $d_{i,\text{beam}}$ are defined based on the objects transverse momenta p_T by

$$d_{i,j} = \min\left(p_{\text{T},i}^{2n}, p_{\text{T},j}^{2n}\right) \cdot \frac{\Delta_{i,j}^2}{R^2} \qquad \text{and} \tag{3.2}$$

$$d_{i,\text{beam}} = p_{\text{T},i}^{2n} \ , \tag{3.3}$$

with $\Delta_{i,j}$ as the distance of object i to object j in the η-ϕ plane. The resolution parameter R defines the size of the resulting jets and the parameter n represents different approaches of the distance calculation. In an iterative procedure all possible distances $d_{i,j}$ and $d_{i,\text{beam}}$ from the list of input objects are evaluated and the smallest distance is determined. If this distance is of type $d_{i,j}$, then the corresponding objects i and j are merged into one new object in the input list with the original objects i and j being removed. But if a distance of type $d_{i,\text{beam}}$ is the smallest one, then the corresponding object i is removed from the list of input objects and declared as a final jet. This whole procedure is repeated until no input object is left over, resulting in a list of jets which all have a minimum distance $\Delta_{i,j}$ from each other larger than R.

Depending on the chosen value for the parameter n in equations 3.2 and 3.3, three different approaches are commonly used. For the k_t algorithm [131] $n = 1$ is chosen, whereas the Cambridge-Aachen algorithm [132, 133] uses $n = 0$ and the anti-k_t algorithm [134] uses $n = -1$, which means that the distances are weighted with the inverse of the larger of the two p_T values.

Depending on the type of the input objects, several types of jets can be distinguished. When the stable particles after the hadronization process in MC event generation are used, the reconstructed jets are generator jets, which can be used as a reference for resolution comparisons. The clustering algorithms can also be applied on the generated partons before hadronization, resulting in parton jets, which have the advantage that they can be used in calculations for theory predictions.

Calorimeter jets are reconstructed from the unweighted sums of the deposited energies in the ECAL and the HCAL, which has the disadvantage that the higher precision of the ECAL is lost in this sum. Therefore the jet plus tracks (JPT) algorithm [135] improves the calorimeter jets by matching the jets to the information of the tracking system.

Because CMS makes use of the Particle Flow algorithm, the list of reconstructed objects based on the combination of all detector components can also be used as an input for the clustering algorithms, resulting in the Particle Flow jets (PF jets). These jets have the best transverse momentum resolution of the described jet types. Therefore the PF jets clustered by the anti-k_t algorithm are the commonly used jets at CMS and are also used in this thesis.

Jet Energy Corrections

The reconstruction of jets enables the measurement of the four-vectors of the original particles. But this measurement is influenced by many different effects, why multiple correction steps for these effects have to be applied to data and MC. These steps are called Level 1 to Level 7 corrections, where the first three Levels correct for instrumental effects in the energy measurement. The Levels 4 to 7 are optional corrections to improve the measurement of the jet energy and the four-vector of the original particle.

Level 1: At first, noise of the detector's electronics and pile-up contributions are corrected for. For this, the corresponding energy density is determined, which is then subtracted from the energy of the jet.

Level 2: This correction accounts for the η-dependent jet response because of uninstrumented regions and non-compensating behavior of the detector. This is achieved by correcting the jets in such a way that the jet response in the whole detector equals the one in $|\eta| < 1.3$.

Level 3: The calorimeter's response also depends on the transverse momentum of the jet. This Level corrects for this, resulting in a flat jet response in p_{T}.

Level 4: Because the jet response also depends on the fraction of the energy deposited in the electromagnetic calorimeter, this Level corrects for a uniform response.

Level 5: The goal of this Level is to correct for the jet flavor dependence by assuming that the jet originated from a specific parton flavor.

Level 6: This correction fixes energy offsets arising from the underlying event.

Level 7: The last Level optionally corrects the jets to parton level, which means that the corrected transverse momentum of the jet is equal to the transverse momentum of the originating parton on average. This correction is calculated by comparing the transverse momenta of a generator jet and the corresponding parton.

Additional residual corrections are applied because differences have been observed between MC and data in Level 2 and Level 3 corrections. These are applied to data, because the MC truth calibration is already good enough.

3.2.7. b Tagging

The identification of b jets, which originate from the hadronization of bottom quarks, is called *b tagging*. A discrimination between these jets and jets from lighter quarks or gluons is important for many analyses. Especially for the analysis of top quarks, which decay to bottom quarks and W bosons in nearly 100% of all cases, the use of b tagging results in a higher purity of the selected dataset.

Bottom quarks produced in collision events form B hadrons, which mostly decay via the weak interaction, leading to a relatively long lifetime. Therefore their decay takes place at a distance of the order of $c\tau_{\mathrm{B}} \approx 480\ \mu\mathrm{m}$ away from the primary vertex, which results in a displacement of the tracks of the decay products. By reconstructing these tracks and their displacement parameter d_0, the secondary vertex can be determined, as shown in figure 3.3.

Several algorithms for b tagging exist in CMSSW [137], which can be characterized by their tagging efficiency of bottom quarks and the rate of wrongly tagged jets from light quarks, called mistag rate. Because these two quantities are related to each other, an optimal working point needs to be chosen depending on the analysis' requirements.

In this analysis the combined secondary vertex (CSV) algorithm [138] is used at the medium working point. This working point is characterized by a mistag rate of 1%, which is achieved by a cut on the CSV discriminator at 0.679.

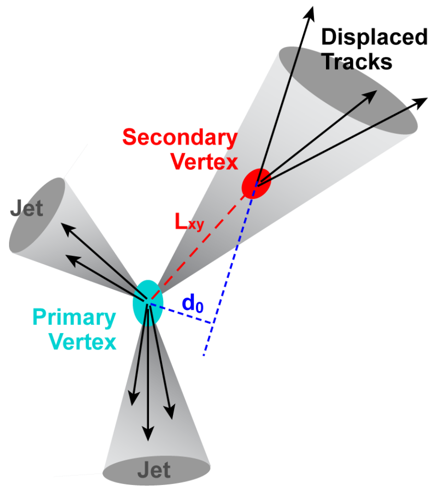


Figure 3.3.: Illustration of the reconstruction of a b jet, taken from [136]. The produced bottom quark forms a B hadron, which has a relatively long lifetime of about $\tau \approx 1.6$ ps due to the small corresponding CKM matrix element $V_{\mathrm{cb}} \approx 0.041$. The decay of this B hadron then results in jets whose tracks do not point to the primary vertex but are displaced by a distance d_0. The displaced tracks can be used to reconstruct a secondary vertex at a distance L_{xy} from the primary vertex of the hard interaction.

3.2.8. Missing Transverse Energy

The colliding particles at the LHC only carry a momentum along the beam axis. As no initial momentum transverse to the beam axis exists, it also has to vanish in the final state due to momentum conservation. Therefore the vectorial sum of the transverse momenta of all particles of an event adds up to zero. Particles in an event that are not recognized by the detector, like the weakly interacting neutrinos or some hypothesized particles in extensions to the Standard Model, result in an imbalance of the total transverse momentum. The measurement of this effect is achieved by introducing the vector of the so-called *missing transverse energy* (MET, $E_{\mathrm{T}}^{\mathrm{miss}}$ or $\not{E}_{\mathrm{T}}$) [120–122]. In the Particle Flow approach it is defined as the negative vectorial sum of the transverse momenta of all reconstructed particles:

$$\vec{\not{E}}_{\mathrm{T}} = -\sum_i \vec{p}_{\mathrm{T},i} \,. \tag{3.4}$$

The jet energy corrections described in section 3.2.6 are propagated to the missing transverse energy as Type-I corrections. Therein the vector sum of the transverse momenta of particles which can be clustered as jets is replaced by the vector sum of the transverse momenta of the jets to which the jet energy corrections are applied. If other input objects are used, different additional corrections are needed.

4. Selection of Events

An enormously large number of events has been collected by the CMS detector in 2012, but most of them do not contain processes of interest. An overview of the cross sections for different processes depending on the center-of-mass energy $\sqrt{s}$ is shown in figure 4.1. Usually analyses are only interested in specific processes, therefore an event selection procedure has to be applied to enrich the data in signal events by suppressing events originating from different background processes.

In this chapter the modeling of the signal and the various background processes is described, followed by the applied selection criteria and the estimation of the background contributions in this selection.

4.1. Modeling of Signal and Background Events

4.1.1. The Lepton+Jets Channel

The lepton+jets decay channel of top-quark pairs is characterized by one top quark decaying leptonically and one decaying hadronically. This provides a clean signature in combination with a still high branching ratio.

An example of a leading order Feynman diagram for this process is shown in figure 4.2, where the typical signature of this decay can be seen. Each of the two top quarks decays into a bottom quark and a W boson. One W boson then decays into a quark-antiquark pair, the other into a charged lepton and the corresponding neutrino. This results in an overall number of four jets, with two of them originating from bottom quarks, and an isolated charged lepton with a high transverse momentum in the final state. Additionally, missing transverse energy due to the final state neutrino can be measured. Despite this channel being called lepton+jets, only muons and electrons are considered, because tau leptons almost instantly decay into lighter leptons or quark-antiquark pairs, which makes their selection and reconstruction more complicated.

The $t\bar{t}$ signal and the main background events were generated by two different event generators. MadGraph/MadEvent is able to simulate the radiation of up to four jets through matrix element based calculations. This leads to a more accurate simulation of the jet kinematics, but as different jet multiplicities are considered as different processes, the interference which causes the charge asymmetry is not simulated. Therefore MadGraph is only used for the simulation of the background processes. For the signal process, the matrix element generator Powheg is used, which includes next-to-leading-order calculations that result in the charge asymmetry. But the simulated asymmetry is smaller than predicted by theory by a factor of about 1.5 [65]. A factor of 1.2 originates from the inclusion of QED effects in the prediction, and another factor of about 1.3 originates from normalizing with respect to the Born cross section instead of the NLO result. Because the asymmetric part of the cross section is known to leading order only, the normalization to this leading order cross section was considered to be more plausible in the theory prediction.

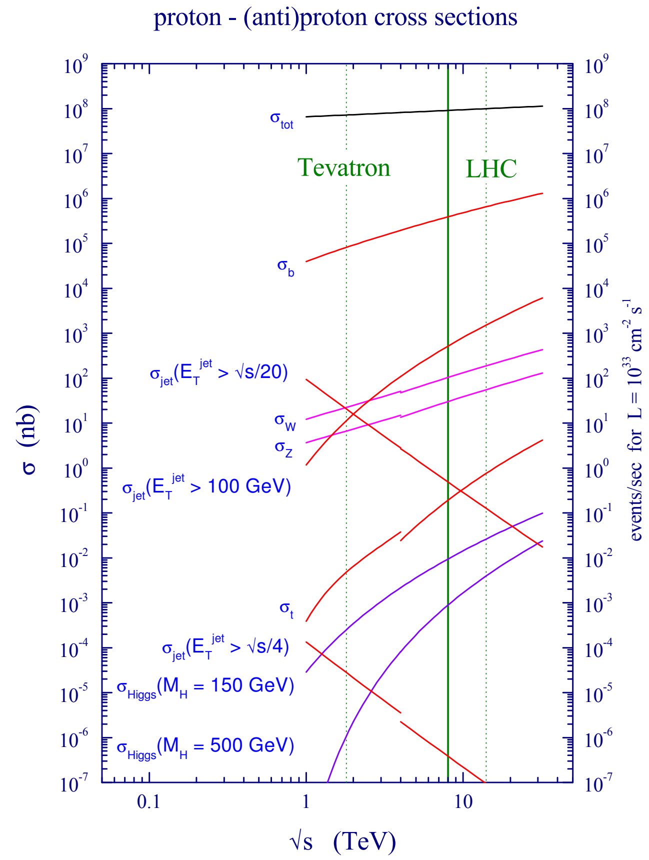


Figure 4.1.: Predicted cross sections of different physics processes depending on the center-of-mass energy $\sqrt{s}$, taken from [139]. For a center-of-mass energy $\sqrt{s}$ below 4 TeV the cross sections in proton-antiproton collisions are shown, for $\sqrt{s}$ larger than 4 TeV those for proton-proton collisions. The center-of-mass energy of LHC's 2012 operation at $\sqrt{s} = 8$ TeV is marked by a solid green line. The total inelastic proton-proton cross section (black) is about nine orders of magnitude larger than the cross section of top-quark pair-production (red) at this energy. Additionally, the center-of-mass energy of the Tevatron ($\sqrt{s} = 1.96$ TeV) and the design value of the LHC ($\sqrt{s} = 14$ TeV) are indicated by dashed green lines.

The mass of the top quark is set to $m_t = 172.5$ GeV/c^2 for both generators. For MadGraph the PDF of the protons is described by the CTEQ6L1 parametrization, for Powheg by the CT10 parametrization. These two generators are both interfaced to Pythia to simulate additional radiation, the hadronization process and the decay of unstable particles. For all these simulations with Pythia, the Z2* tune [140] was used.

4.1.2. Background Processes

Several background processes exist that have to be considered. These processes can have a signature similar to the $t\bar{t}$ signal, or their signatures can look the same due to an incomplete or wrong reconstruction of physical objects.

A signature with a highly energetic lepton can be generated by leptons originating directly from the decay of W bosons. Also the electroweak production of single top quarks as well as Drell-Yan processes [141] via Z bosons or virtual hard photons γ^* can produce such leptons. When additional hard radiation leads to multiple jets,

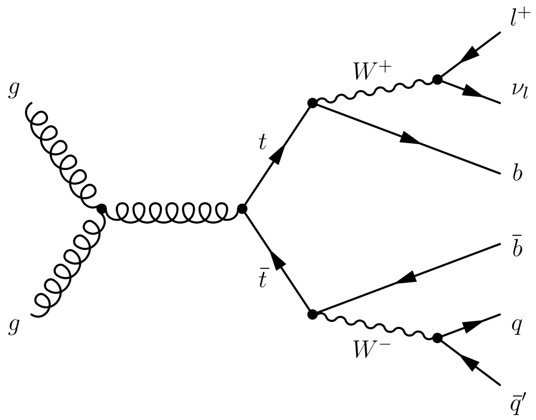


Figure 4.2.: Feynman diagram of a $t\bar{t}$ event in the lepton+jets decay channel. Each top quark decays into a bottom quark and a W boson. One of the W bosons further decays into a charged lepton and a neutrino, the other one decays into a pair of quarks.

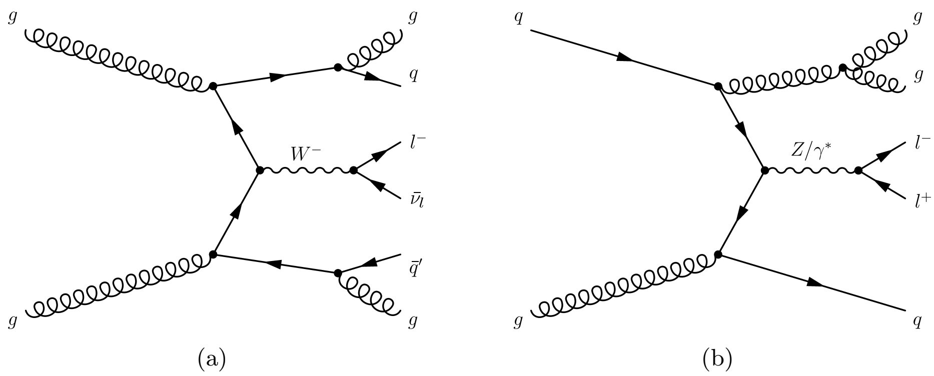


Figure 4.3.: Examples of leading order Feynman diagrams of $W^\pm$ (a) and Drell-Yan (b) production. The signature of the shown W+jets production mode is similar to the signature of the $t\bar{t}$ lepton+jets decay channel. The signature of the shown Drell-Yan process can look like a signal event if one of the leptons is misidentified as a jet.

these processes can mimic the $t\bar{t}$ signature.

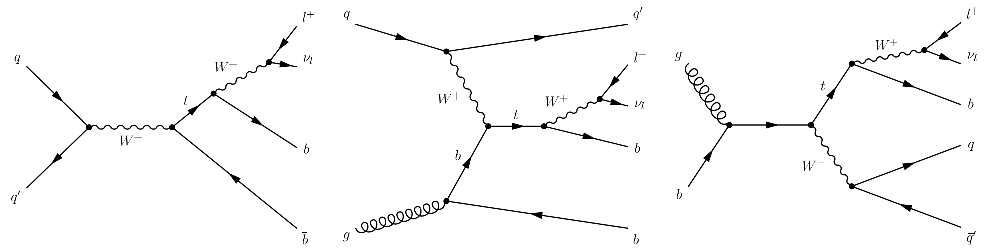
The jet requirement of the $t\bar{t}$ signature is easily fulfilled by QCD events. If an additional highly energetic lepton is produced in a shower or if jet components are misidentified as a lepton, QCD processes can also have a signal-like signature.

W+Jets and Z/γ^*+Jets

The leptonic decay of a produced W boson in association with additional radiations has the same signature as the $t\bar{t}$ signal, as shown in figure 4.3(a). Due to the still large cross section this process is the dominant background of the lepton+jets channel.

Drell-Yan processes with a leptonically decaying Z boson or virtual photon can also mimic a $t\bar{t}$ event if additional radiated partons exist, as it is shown in figure 4.3(b). This background can be highly suppressed by a veto on the second lepton, but due to possible non-detected or misidentified leptons this process still can have a signal-like signature.

The Monte Carlo samples of these processes were generated by MadGraph



(a) (b) (c)

Figure 4.4.: Examples of Feynman diagrams for the electroweak production of single top quarks: (a) the *s*-channel, (b) the *t*-channel and (c) the tW-channel. These processes can look like signal events when the top quark decays leptonically like it is shown and additional partons are radiated.

interfaced to Pythia. For the Drell-Yan process only events with an invariant mass of the lepton-antilepton pair larger than 50 GeV/c^2 were considered.

Single-Top-Quark Production

The electroweak production of single top quarks, shown in figure 4.4, can have the same signature as the $\mathrm{t\bar{t}}$ signal if the top quark decays leptonically and additional radiation is present. But the cross section for this process is rather small, which reduces the impact of this background. The MC events were generated by Powheg interfaced to Pythia in the t and tW-channel, with the s-channel being negligible.

QCD Multijet Processes

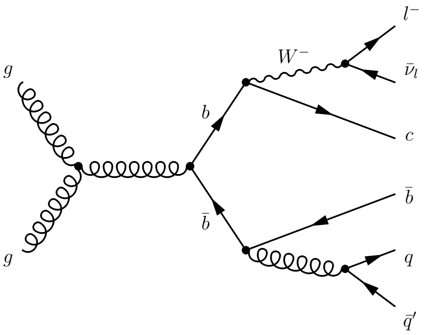
Due to their high cross section multijet events produced via the strong force play an important role in the lepton+jets channel. Usually the final state of these events only consist of jets, but if hadrons with bottom or charm quarks are produced, these quarks can decay into W bosons which decay semileptonically. It is also possible that parts of a jet are misidentified as a lepton, leading to a signal like signature. These processes are shown in figure 4.5.

For the actual analysis in this thesis no simulated samples of multijet processes were used because the available number of events in the Monte Carlo samples for the selected phase space region is not sufficient. Instead the shape was taken from a sideband region in data enriched with QCD multijet events to approximate the QCD background.

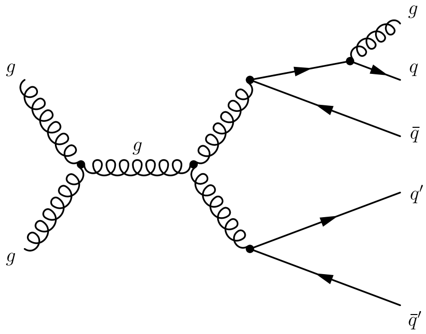
4.1.3. Used Monte Carlo Samples

The samples used in this analysis have been produced in the “Summer12” production campaign of the CMS collaboration. An overview over the used samples is given in table 4.1.3, including the effective cross sections and the numbers of produced events. Here the effective cross section corresponds to the cross section of a specific simulated process considering the restrictions on decay channels, additional radiations and event kinematics.

The cross section of the signal $\mathrm{t\bar{t}}$ process has been determined at next-to-next-to-leading order (NNLO) [55] to be $245.8^{+6.2}_{-8.4}$ pb. For the single-top-quark cross sections, approximate NNLO calculations yield $\sigma_{\mathrm{tW}} = 11.1 \pm 0.3 \pm 0.7$ pb [142,143]



(a)



(b)

Figure 4.5.: Examples of Feynman diagrams of QCD multijet processes. In (a) the production of a $\mathrm{b\bar{b}}$ pair via the strong interaction is shown. The bottom quark further decays via the weak interaction into a charm quark, a charged lepton and an antineutrino. The bottom antiquark instead radiates a gluon that splits into a pair of quarks, resulting in two jets. In (b) an event containing only quarks and gluons in the final state is depicted. If one of the particles of the produced jets is misidentified as a charged lepton, these events can mimic signal events.

Process			σ_{eff} [pb]	N_{prod}
signal	$\mathrm{t\bar{t}}$	inclusive	245.8	21 675 970
	W+2jets	$\mathrm{W} \to \mathrm{l}\nu$	1 750.0	34 044 921
electroweak	W+3jets	$\mathrm{W} \to \mathrm{l}\nu$	519.0	15 519 503
background	W+4jets	$\mathrm{W} \to \mathrm{l}\nu$	214.0	13 382 803
	Z/γ^*+jets	$\mathrm{Z}/\gamma^* \to \mathrm{ll}$, $m_{\mathrm{ll}} > 50$ GeV/c^2	3 503.7	30 439 503
	t-channel, t	inclusive	56.4	3 758 227
single top	t-channel, $\bar{\mathrm{t}}$	inclusive	30.7	1 935 072
background	tW-channel, t	inclusive	11.1	497 658
	tW-channel, $\bar{\mathrm{t}}$	inclusive	11.1	493 460

Table 4.1.: Overview of the used Monte Carlo samples with the number of generated events N_{prod} and the effective cross section σ_{eff}.

for the single-top-quark or top-antiquark associated production (tW-channel). The cross section of the t-channel of single top quark production has been calculated separately for the top quark to $\sigma_{t,\mathrm{t}} = 56.4^{+2.1}_{-0.3} \pm 1.1$ pb and for the top antiquark to $\sigma_{t,\bar{\mathrm{t}}} = 30.7 \pm 0.7^{+0.9}_{-1.1}$ pb [143, 144]. Compared to the other single-top-quark production processes the s-channel has a much smaller cross section and was therefore neglected for this analysis. For the leptonically decaying W bosons the cross section was determined at NNLO to $\sigma_{W\to l\nu} = 37\,509$ pb using the FEWZ [145] framework. The samples actually used in the analysis have been generated separately for the different jet multiplicities and have been combined according to their individual leading order cross sections. With the same framework the NNLO cross section of Drell-Yan processes, where two leptons with an invariant mass of at least 50 GeV$/c^2$ are produced, has been calculated to be $\sigma_{\mathrm{Z}/\gamma^*\to\mathrm{ll}}(m_{\mathrm{ll}} > 50\ \mathrm{GeV/c^2}) = 3\,503.71$ pb.

4.1.4. Used Data

In this thesis the collision data recorded in 2012 by the CMS detector in proton-proton collisions at a center-of-mass energy of 8 TeV is used.

Due to the high instantaneous luminosity of the LHC shown in figure 2.2(a) the recorded events are required to fulfill certain criteria of the HLT system to be recorded. These criteria are based on the existence of physical objects with certain kinematic quantities and result in different corresponding primary datasets. In this analysis the `SingleElectron` and `SingleMu` datasets are used, which require the presence of one electon or muon.

But only those events of these datasets are considered that fulfill the requirements of dedicated single lepton triggers, which require one highly energetic electron or muon. For the electron channel the `HLT_Ele27_WP80` trigger is used, which requires an electron with a transverse energy of more than 27 GeV passing certain identification requirements. In the muon channel the `HLT_IsoMu24_eta2p1` trigger is used, which requires a muon with at least 24 GeV/c of transverse momentum that passes an isolation criterion and lies in the central region of $|\eta| < 2.1$. These triggers are also simulated in MC production of events, therefore no corrections for this had to be applied on the simulated samples.

The data is separated into different runs, which indicate specific time periods in which the detector has been recording without interruption. These runs are further divided into luminosity sections, in which the instantaneous luminosity is considered to be constant. The used data in this analysis covers the runs 190 456 to 208 686. But only the data from those runs and luminosity sections can be used in which all detector components have been working correctly. Therefore the Data Quality Monitoring System (DQM) [146] determines the runs and luminosity sections which can be used for analyses by documenting its results in configuration files in the JSON format. With these files unvalidated events are filtered out in the first step of the analysis procedure. This thesis uses a certified JSON file corresponding to an integrated luminosity of 19.8 fb^{-1}.

After reconstruction several filters regarding noise cleaning [147] have to be applied on data. Furthermore events with more than ten tracks on the whole, but less than 25% high-purity tracks, which are called *beam scraping events*, are discarded.

4.1.5. Corrections on Simulated Events

Not all of the effects in particle collisions can be simulated as precisely as necessary by the event generators. Therefore additional corrections are applied on the simulated events.

Pile-Up Reweighting

The pile-up distributions of the produced MC samples do not match the pile-up conditions observed in data. Therefore the simulated events are corrected by the so-called *pile-up reweighting* [148], where the events of a sample get reweighted in such a way that the pile-up distributions of the MC samples match the observed distributions in data.

b Tag Scale Factors

To achieve equal probabilities for tagging scenarios in simulations and in data, b-tag scale factors [149] are applied to the MC samples. For this the probabilities of all possible tagging scenarios that result in a given event to be selected are determined. Then this event is reweighted by the ratio of the amalgamate probabilities for these scenarios of data and simulation.

Reweighting of Top-Quark p_T

The p_{T} spectrum of top quarks in simulated $\mathrm{t\bar{t}}$ events is harder than the result of differential cross section measurements [150] and theory predictions. To correct for this effect, scale factors SF for the individual generated top quarks in each event are determined [151] by

$$SF = \exp\Big(0.156 - 0.00137 \cdot p_{\mathrm{T}}^{\mathrm{top}}/(\mathrm{GeV}/c)\Big)\,. \tag{4.1}$$

The event is then reweighted by the geometric mean $\sqrt{SF(\mathrm{top}) \cdot SF(\text{anti-top})}$ of the scale factors of top quark and antiquark.

Jet Resolution

Measurements of the imbalance of the transverse momenta in di-jet events show that the jet energy resolution (JER) in data is worse than in simulations [152]. To correct the simulation for this, selected jets are matched to generated jets. The difference between the transverse momenta of reconstructed and generated jets is then multiplied by the corresponding $|\eta|$-dependent scale factor given in table 4.2 and is propagated to the jet four-vectors.

$\lvert\eta\rvert$ range	$\sigma(\mathrm{Data})/\sigma(\mathrm{MC})$
0.0 – 0.5	1.052
0.5 – 1.1	1.057
1.1 – 1.7	1.096
1.7 – 2.3	1.134
2.3 – 5.0	1.288

Table 4.2.: Ratios of the jet p_{T} resolutions in data and MC simulation, taken from [152].

4.2. Selection Criteria

After the reconstruction of the measured data and the simulated samples a selection is applied on the events. In the following section the requirements on the physical objects considered in this analysis are described and the consecutive selection steps are explained.

4.2.1. Definitions of Physical Objects

In the semileptonic decay of a top-quark pair a single charged lepton is produced. To distinguish this isolated lepton from other lepton candidates, which result from the decay of hadrons in background processes, the *corrected relative isolation* is considered. It is a measure for the amount of energy deposited in the hadron calorimeter around the track of the lepton, corrected for the effective area of the lepton to reduce pile-up contributions, and enables a discrimination of signal events against background. The relative isolation is defined as

$$I^{\ell}_{\text{Rel, cor}} = \frac{E^{\ell}_{\text{CH}} + \max\Big(0, E^{\ell}_{\gamma+\text{NH}} - \rho \cdot A_{\text{eff}}(E^{\ell}_{\gamma+\text{NH}}, \eta)\Big)}{p^{\ell}_{\text{T}} \cdot c} \,, \tag{4.2}$$

with the energy E^{ℓ}_{CH} deposited by charged hadrons in a cone with radius $\Delta R = 0.3$ in the η-ϕ-plane around the lepton track, the energy $E^{\ell}_{\gamma+\text{NH}}$ of neutral hadrons and photons, the transverse momentum density ρ of the event as determined using k_{T} jets with a distance parameter of 0.6, the effective area A_{eff} of the lepton as determined from data, and the angular position η of the lepton's supercluster. Low values of the relative isolation parameter indicate a well isolated lepton and vice versa.

Electron Definition

Electron candidates are required to lie within $|\eta| < 2.5$ and to have a transverse energy larger than 30 GeV. Candidates within the ECAL endcap–barrel transition region defined by $1.4442 < |\eta_{\text{sc}}| < 1.5660$ are rejected, where η_{sc} is the pseudorapidity of the electron's supercluster. They are also required to be isolated by a cut on the relative isolation of $I^{\text{e}}_{\text{Rel, cor}} < 0.1$. Furthermore, candidates have to fulfill the criteria of a multivariate identification [153], which takes various variables related to calorimetry and tracking parameters as well as the momentum and η of the electron into account. The discriminant of this identification method is required to be larger than 0.9.

To improve the purity of the electron selection, the possible conversion of a highly energetic photon into a pair of electrons is taken into account by a conversion rejection procedure [154]. Therein missed hits of the electron track in inner tracker layers as well as the location of a reconstructed conversion vertex and its fit probability are considered.

Muon Definition

Muon candidates are PF muons, which are required to be reconstructed as "global" muons, to have a transverse momentum larger than 26 GeV/c and to lie within $|\eta| < 2.1$. The χ^2 value of the global fit has to be smaller than 10 and the number of hits in the tracker has to be larger than 5. The longitudinal position of the muon track at its closest approach to the beam line is required to lie within a distance of 0.5 cm to the longitudinal position of the primary vertex. The global muon track fit needs to contain at least one muon chamber hit, there must be muon segments in at least two muon stations, and the track must contain at least one pixel hit. Muons also must be isolated by requiring a relative isolation of $I^{\mu}_{\text{Rel, cor}} < 0.12$.

Jet Definition

The jets used in this thesis were reconstructed using the anti-k_{T} jet algorithm with a distance parameter of $R = 0.5$ and Particle Flow objects as input objects.

Contributions of charged hadrons, which were identified as originating from pile-up vertices, have been subtracted. Jets are required to lie within $|\eta| < 2.5$ and to have a transverse momentum larger than 30 GeV/c.

Additional criteria have to be fulfilled to reduce the possibility of misidentification: Jets must have at least two constituents as well as charged hadron contributions. If a jet has a neutral hadron energy fraction, a neutral electromagnetic energy fraction or a charged electromagnetic energy fraction larger than 0.99 it is rejected.

The energy of the jets is corrected by applying the L1, L2 and L3 corrections described in section 3.2.6, and for the data events also the L2L3Residual corrections are applied.

4.2.2. Selection Steps

The following selection steps are applied consecutively on the events of the simulated and the measured data samples.

Primary Vertex Criteria

For each event one good primary vertex is required. For this, the number of degrees of freedom n_{dof} of the vertex is required to be larger than 4, where n_{dof} corresponds to the weighted sum of the number of tracks used for its construction. Additionally the primary vertex has to lie in the central detector region with $|z| < 24$ cm and $r < 2$ cm.

Lepton Cuts

For the semileptonic decay channel of the top-quark pair, which is used in this analysis, exactly one isolated lepton is required. Therefore events containing additional leptons are rejected, even if these leptons do not satisfy the original definitions. These veto leptons are defined more loosely compared to the isolated leptons and are therefore called loose electrons and loose muons. Loose electrons are required to have a transverse energy of only at least 20 GeV with a corrected relative isolation smaller than 0.15 and a value larger than 0.5 for the multivariate discriminant. For loose muons a transverse momentum of at least 10 GeV/c with a corrected relative isolation smaller than 0.2 is required. Both types of leptons must lie within $|\eta| < 2.5$. This reduces the Drell-Yan background as well as the dileptonic decay channel of the top-quark pair drastically.

Jet Cut

For this analysis at least four jets as defined above are required. This reduces all other backgrounds by a large amount, enriching the top-quark pair signal.

b Tagging Requirements

In the decay of a top-quark pair two jets are expected to originate from bottom quarks. To account for this the selection requires at least one b tagged jet. This b tagging is performed by the combined secondary vertex (CSV) algorithm [138] at the medium working point, corresponding to a cut on the CSV discriminator at 0.679. Simulations show that the gain of purity by an additional requirement of a

process	electron+jets	muon+jets	lepton+jets
$t\bar{t}$	132 218	154 675	286 893
single top (t)	1 409	2 397	3 806
single top (tW)	5 549	6 425	11 974
W+jets	12 496	14 480	26 976
Drell-Yan+jets	2 527	2 086	4 613
observed data	176 835	198 290	375 125

Table 4.3.: Event yields of the selected dataset corresponding to an integrated luminosity of 19.7 fb^{-1}. The predicted numbers for the different processes based on simulated samples are shown together with the observed number of events in data.

second b tagged jet is not large enough to compensate the huge loss in the selection efficiency.

4.2.3. Selection Results

After the triggers and the event selection cuts are applied, 375 125 events remain with 176 835 events in the electron+jets and 198 290 events in the muon+jets channel of the data sample.

The total number of predicted events for a process before the event selection is given by the product of the integrated luminosity $\mathcal{L}_{\mathrm{int}}$ of the dataset and the calculated effective cross section σ_{eff} of the process. By multiplying the result with the efficiency ϵ_{sel} of the selection procedure the number of predicted events N_{pred} corresponding to the given dataset can be determined by

$$N_{\mathrm{pred}} = \epsilon_{\mathrm{sel}} \cdot \mathcal{L}_{\mathrm{int}} \cdot \sigma_{\mathrm{eff}}, \tag{4.3}$$

$$\text{with } \epsilon_{\mathrm{sel}} = \frac{N_{\mathrm{sel}}}{N_{\mathrm{prod}}}. \tag{4.4}$$

Here the efficiency of the event selection procedure is calculated as the quotient of the numbers of selected events N_{sel} and the numbers of generated events N_{prod} in the simulated MC samples for each process. In table 4.3 the summary of the predicted and measured event numbers is shown.

4.3. Data-Driven Modeling of QCD Multijet Production Processes

The overall cross section for multijet production is several orders of magnitude larger than that of any other process. But after applying the event selection on simulated QCD multijet samples, only a limited number of events remain, even if the samples were enriched with processes expected to fulfill the selection. This limited number of events leads to big fluctuations in distributions created from such samples. Therefore this specific background is modeled directly from a sideband region of the actual dataset enriched in QCD multijet events, which is possible due to the large multijet cross section.

To avoid too large deviations, the sideband selection criteria have to stay close to the selection criteria of the signal region. Because the largest suppression of QCD

Primary Dataset	Run Range	HLT Trigger
SingleElectron	190 456 − 208 686	HLT_Ele27_WP80
SingleMu	190 456 − 193 621	HLT_Mu24_eta2p1_CentralPFJet30_CentralPFJet25
	193 834 − 203 742	HLT_Mu24_CentralPFJet30_CentralPFJet25
	203 777 − 208 686	HLT_Mu18_CentralPFJet30_CentralPFJet25

Table 4.4.: Summary of the HLT paths used for the QCD multijet template.

multijet events in the default event selection is achieved by the requirement of an isolated lepton, the selection of the sideband region is defined by an inversion of this corrected relative isolation of the lepton. Thus, the sideband selection requires an electron with a relative isolation of $0.2 < I^{\mathrm{e}}_{\mathrm{Rel,\ cor}} < 0.5$ instead of $I^{\mathrm{e}}_{\mathrm{Rel,\ cor}} < 0.1$ for the electron+jets channel and $0.2 < I^{\mu}_{\mathrm{Rel,\ cor}} < 0.5$ instead of $I^{\mu}_{\mathrm{Rel,\ cor}} < 0.12$ for the muon+jets channel.

Because the trigger for single muons used in the default selection includes an isolation criterion, a set of different triggers is selected. These triggers for the multijet sideband are listed in table 4.4 together with their run ranges to make sure they do not overlap. They have different p_{T} and $|\eta|$ requirements as well as additional jet requirements, but the values of these requirements are looser than the requirements of the default event selection. For the multijet template in the electron channel the same trigger as in the default selection can be used.

The contributions of the various processes to the multijet sideband region were examined with Monte Carlo studies. These contributions are shown in figure 4.6 for three important variables: the sensitive variable $\Delta|y|$, the transverse mass $m_{\mathrm{T,W}}$ of the leptonically decaying W boson and a variable called M3, which is the invariant mass of those three jets in an event with the largest vectorially summed transverse momentum. The resulting multijet event purity in the sideband region is 90% for the electron+jets channel and 91% for the muon+jets channel. The determined contributions from non-QCD processes were subtracted from the multijet sideband region resulting in a data driven template of only QCD events.

4.4. Background Estimation

The selection procedure described in section 4.2 does not yield a 100% pure $\mathrm{t\bar{t}}$ sample. But for a measurement of the charge asymmetry the contributing backgrounds have to be determined precisely. Theory predictions based on Monte Carlo samples as shown in section 4.2.3 are not viable for the analysis. Reasons for this are possible mismodeling issues and the lack of sufficient QCD multijet simulations in the selected phase space. Therefore the background estimation is performed by a binned likelihood fit of templates to the distributions of discriminating variables in data. For this procedure, the THETA framework [155] is used and the templates for the different contributions are taken from MC simulations, except for the QCD multijet template, which is taken from a sideband region in data as described in section 4.3.

The variables used for the background estimation are the transverse mass $m_{\mathrm{T,W}}$ of the leptonically decaying W boson and the M3 variable. Their discriminating power can be seen in figure 4.7. The $m_{\mathrm{T,W}}$ distribution is able to discriminate

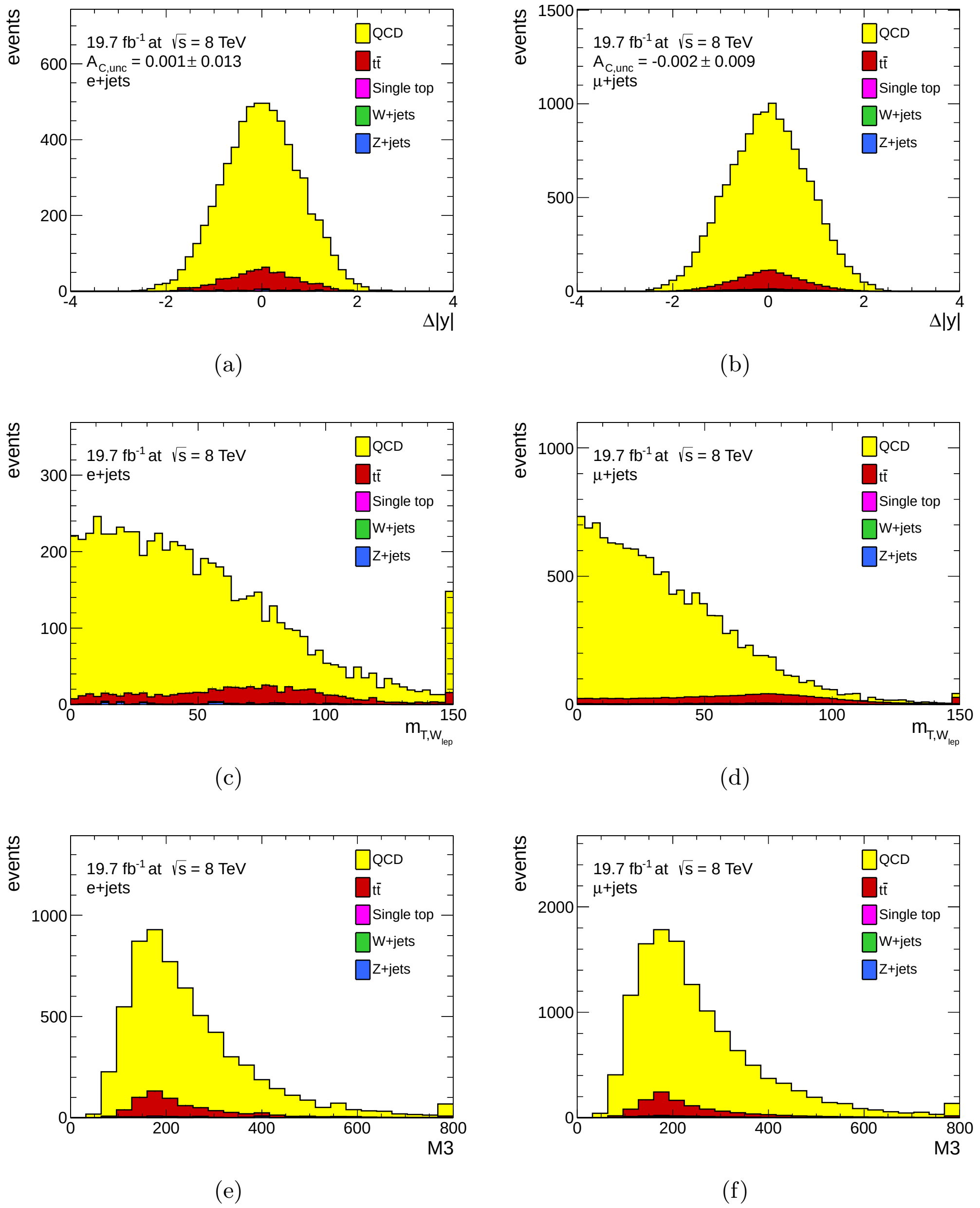


Figure 4.6.: Contributions to the data-driven QCD multijet templates in the sensitive variable $\Delta|y|$ (top) in $m_{\mathrm{T,W}}$ (middle) and in M3 (bottom). On the left in (a), (c) and (e) the distributions in the electron+jets channel and on the right in (b), (d) and (f) the distributions in the muon+jets channel are shown.

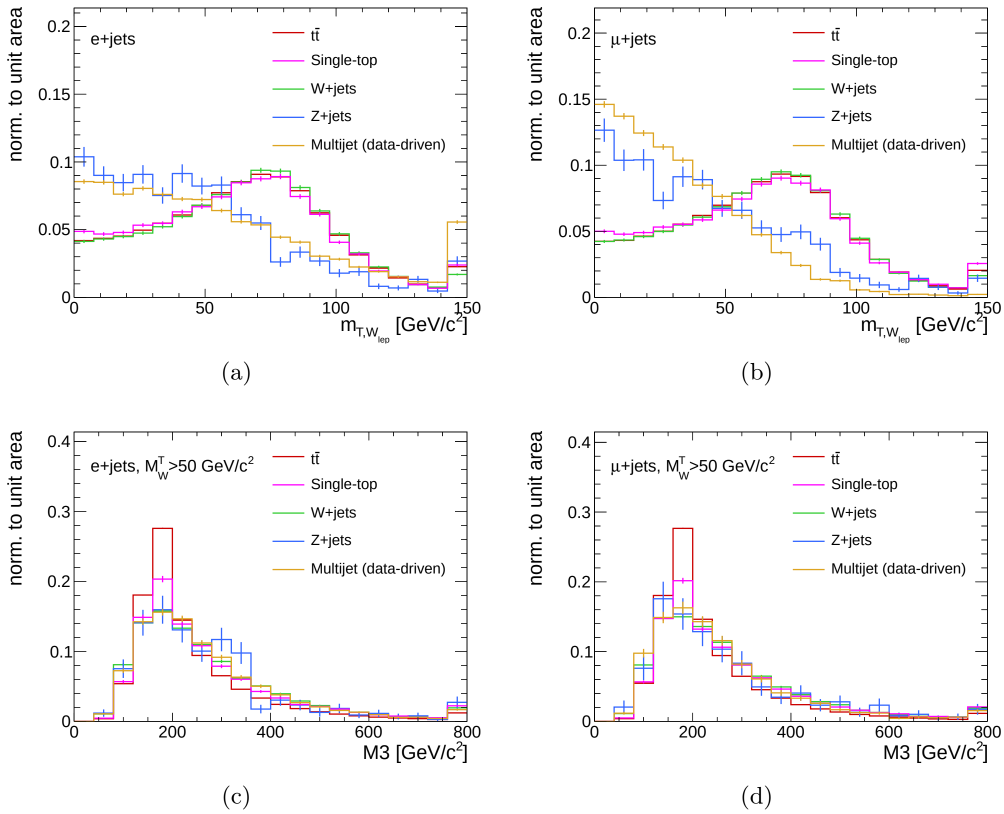


Figure 4.7.: Shape comparisons of the $m_{\mathrm{T,W}}$ (top) and M3 (bottom) distributions for signal and background processes. In (a) and (c) on the left the electron+jets channel and in (b) and (d) on the right the muon+jets channel is shown. For the M3 distributions only events with an $m_{\mathrm{T,W}}$ larger than 50 GeV/c^2 are used for the fitting procedure. Connected to that, the remaining events with an $m_{\mathrm{T,W}}$ below 50 GeV/c^2 enter into the template fit in the $m_{\mathrm{T,W}}$ distribution.

between processes with and without the production of a real W boson. Processes including a real W boson show a peak at the mass of the W boson, processes without a real W boson are enriched at low $m_{\mathrm{T,W}}$ masses. Therefore a fit to this distribution is able to determine the Drell-Yan and the multijet contributions. The M3 mass can be seen as a simple approach of reconstructing the mass of the hadronically decaying top quark. Therefore the events including a hadronically decaying top quark, like the signal process, show a peak at the top-quark mass, in contrast to the other background processes, which have a smoother distribution. Therefore a fit to this distribution yields the signal fraction of the selected sample.

In total, ten different templates are used in the fitting procedure: the $\mathrm{t\bar{t}}$ signal and the W+jets, Drell-Yan+jets, single-top-quark production and QCD multijet background processes, with separate templates for the electron+jets and the muon+jets decay channels. The separation of electron and muon channels is done to account for differences in the selection efficiencies and in the multijet contributions of these channels. For a correct determination of uncertainties each event must only appear once in the fitting procedure, so the two distributions of the fitting variables have to be orthogonal. Therefore the samples are divided into events with $m_{\mathrm{T,W}} < 50$ GeV/c^2 used for the fit to the $m_{\mathrm{T,W}}$ distributions, and into events

process	electron+jets	muon+jets	total
single top (t + tW)	$6\,804 \pm 690$	$8\,395 \pm 868$	$15\,199 \pm 1\,109$
W+jets	$20\,344 \pm 1\,557$	$18\,401 \pm 1\,450$	$38\,745 \pm 2\,128$
Drell-Yan+jets	$1\,761 \pm 805$	$1\,771 \pm 632$	$3\,532 \pm 1\,024$
QCD multijet	$11\,053 \pm 1\,517$	$5\,491 \pm 678$	$16\,544 \pm 1\,662$
total background	$39\,963 \pm 2\,419$	$34\,057 \pm 1\,928$	$74\,020 \pm 3\,093$
$\mathrm{t\bar{t}}$ signal	$136\,886 \pm 1\,486$	$164\,239 \pm 1\,483$	$301\,124 \pm 2\,099$
observed data	176 835	198 290	375 125

Table 4.5.: Fit results with statistical uncertainties together with the numbers of observed events in data. These statistical uncertainties of the fit are partially correlated, which can be seen in figure 4.9. The samples are fitted separately in the electron+jets and muon+jets channel and then combined for the background estimation of the total lepton+jets channel, given in the last column. The resulting signal purity is 77% in the electron+jets channel and 83% in the muon+jets channel.

with $m_{\mathrm{T,W}} > 50$ GeV/c^2 for the fit to the M3 distributions.

These templates are normalized to the predicted numbers of events for their specific processes, obtained from MC studies and shown in table 4.3. For the data-driven QCD template the event yield of available MC samples is used as an approximation for the predictions, but this parameter is left free in the fitting procedure anyway. The actual parameters determined by the fitting procedure are the β factors of each template, which are the ratio of measured and predicted numbers of events

$$\beta_k = \frac{N_k^{\mathrm{meas}}}{N_k^{\mathrm{pred}}} . \tag{4.5}$$

The contribution of Drell-Yan+jets events to the selected data is difficult to discriminate from the multijet background and the single-top-quark production is difficult to discriminate from the $\mathrm{t\bar{t}}$ signal process. But both Drell-Yan+jets and single-top-quark production processes are well understood theoretically and their contributions are expected to be small. Therefore the numbers of events in the fit are constrained to the values predicted by MC simulations with an uncertainty of 30% for the Drell-Yan+jets process and of 10% for the single-top-quark production process assigned through Gaussian functions in the likelihood. All other parameters are left free in the fitting procedure.

The resulting event yields, which are shown in table 4.5, are given by the product of the predicted event numbers and the corresponding β factors. The background estimation shows signal purities of 77% in the electron+jets channel and of 83% in the muon+jets channel. A comparison of the templates normalized to the fit result and the measured distributions in the fit variables is shown in figure 4.8. A good agreement can be found between simulation and data.

In figure 4.9 the correlation matrix of the fit parameters is shown. Large negative correlations between the Drell-Yan+jets and QCD multijet parameters as well as between the W+jets and the $\mathrm{t\bar{t}}$ parameters can be seen. This can be explained by the fact, that these processes are hard to distinguish from each other, respectively. Because these correlations cannot be neglected, this correlation matrix is used in the subsequent steps of the analysis to estimate the statistical errors on the measurement results correctly.

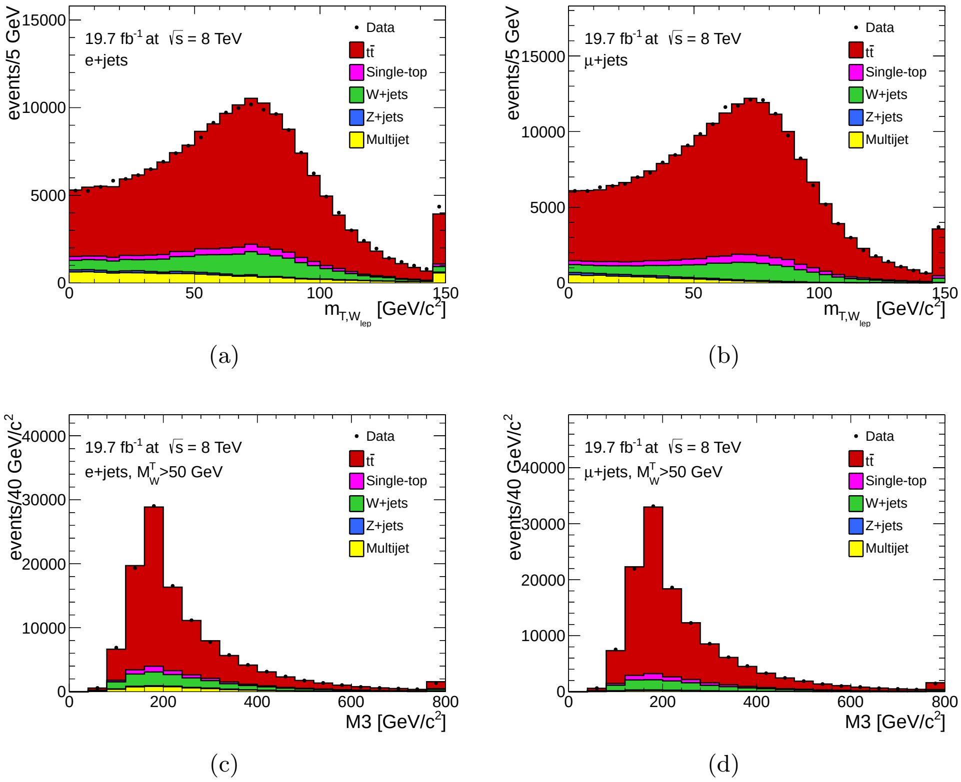


Figure 4.8.: Comparison of data and normalized templates in the variables $m_{\mathrm{T,W}}$ (top) and M3 (bottom) in the electron+jets channel on the left ((a) and (c)) and in the muon+jets channel on the right ((b) and (d)). All templates are normalized to the fit result shown in table 4.5.

	tt̄, e	tt̄, μ	W+jets, e	W+jets, μ	DY+jets, e	ST, e	DY+jets, μ	ST, μ	QCD, e	QCD, μ
QCD, μ	0.00	-0.15	0.00	-0.00	0.00	0.00	-0.63	-0.03	-0.00	1.00
QCD, e	-0.16	0.05	-0.43	-0.05	-0.68	-0.05	-0.00	0.00	1.00	-0.00
ST, μ	0.00	-0.23	-0.00	-0.35	0.00	0.00	0.00	1.00	0.00	-0.03
DY+jets, μ	-0.00	-0.03	0.00	-0.11	-0.00	-0.00	1.00	0.00	-0.00	-0.63
ST, e	-0.21	0.00	-0.19	-0.00	-0.01	1.00	-0.00	0.00	-0.05	0.00
DY+jets, e	-0.05	-0.04	0.19	0.04	1.00	-0.01	-0.00	0.00	-0.68	0.00
W+jets, μ	-0.02	-0.70	0.05	1.00	0.04	-0.00	-0.11	-0.35	-0.05	-0.00
W+jets, e	-0.61	-0.05	1.00	0.05	0.19	-0.19	0.00	-0.00	-0.43	0.00
tt̄, μ	0.02	1.00	-0.05	-0.70	-0.04	0.00	-0.03	-0.23	0.05	-0.15
tt̄, e	1.00	0.02	-0.61	-0.02	-0.05	-0.21	-0.00	0.00	-0.16	0.00

Figure 4.9.: Correlation matrix of the fit parameters of the background estimation. Large negative correlations can be seen between the Drell-Yan+jets and QCD multijet parameters as well as between the W+jets and the $t\bar{t}$ parameters.

5. Measurement of the $t\bar{t}$ Charge Asymmetry

The analysis documented in this thesis is a measurement of the charge asymmetry in top-quark pair-production in full and fiducial phase spaces. For this purpose, the sensitive variable $\Delta|y|$ introduced in section 1.3 is used. But this variable is not directly accessible in the reconstructed and selected datasets described in the previous chapters. Instead, an additional reconstruction procedure is necessary to determine the properties of the two top quarks by combining the observed decay products.

This selected and reconstructed data still contains background events without a top-quark pair. These background contributions are subtracted according to the results of the background estimation.

Additionally, the measured distributions are corrected for distortions from the selection and reconstruction procedures. For this, a regularized unfolding procedure is applied. The correction for selection effects is applied to correct into the full phase space and to correct into smaller fiducial phase spaces to reduce uncertainties.

In the following sections the mentioned analysis steps are documented. At first the full reconstruction of the $t\bar{t}$ event kinematics is described, followed by the unfolding procedure including the choice of the binning, the definition of fiducial phase spaces and consistency checks. Then the various systematic uncertainties are determined, after which finally the results of the measurements are presented.

5.1. Full Reconstruction of $t\bar{t}$ Events

To calculate the sensitive variable, the four-vectors of the two top quarks of the signal events need to be determined. For this, the reconstructed physical objects of an event are individually assigned to the final state particles expected in the semileptonic decay of the top-quark pair. For this assignment, the reconstructed jets, charged leptons, and the missing transverse energy are used. But the assignment is ambiguous, which allows multiple reconstruction hypotheses. Therefore it is necessary to find the best hypothesis by defining an appropriate criterion.

In the following sections the hadronically decaying top quark is labeled t_{had} and its decay products, the W boson and the bottom quark, are labeled W_{had} and b_{had}. The decay products of the leptonically decaying top quark t_{lep} are labeled W_{lep} and b_{lep}, respectively. For calculations natural units with $c = 1$ are used.

5.1.1. Reconstruction of all Possible Hypotheses

The charged lepton is the only object which can directly be assigned without any ambiguities, therefore it is the starting point for the reconstruction. The charge of this lepton as a decay product of t_{lep} indicates the nature of the leptonically decaying top quark. A positively charged lepton determines t_{lep} to be the top quark and $\mathrm{t_{had}}$ to be the top antiquark, or vice versa for a negatively charged lepton.

The missing transverse energy is assumed to be caused by the non-detected neutrino, therefore the x and y components of $\vec{\not{E}}_{\mathrm{T}}$ can directly be assigned to the $p_{x,\nu}$ and $p_{y,\nu}$ components of the neutrino's four-momentum p_ν. The remaining z component can be calculated by taking the energy and momentum conservation of the $\mathrm{W_{lep}}$ decay into account, given by

$$p_{\mathrm{W_{lep}}} = p_{\mathrm{l}} + p_\nu \,, \tag{5.1}$$

with p_{l} and $p_{\mathrm{W_{lep}}}$ as the four-momenta of the charged lepton and the leptonically decaying W boson. By setting the neutrino's rest mass to zero, as well as by assuming the rest mass of the lepton to be negligible and the W boson to be produced on-shell, a quadratic equation for the z component of the neutrino's four-vector can be derived [156] from equation 5.1:

$$m_{\mathrm{W}}^2 = \left(E_{\mathrm{l}} + \sqrt{\vec{\not{E}}_{\mathrm{T}}^{\,2} + p_{z,\nu}^2} \right)^2 - \left(\vec{p}_{\mathrm{T,l}} + \vec{\not{E}}_{\mathrm{T}} \right)^2 - \left(p_{z,\mathrm{l}} + p_{z,\nu} \right)^2 . \tag{5.2}$$

Here, E_{l} and $\vec{p}_{\mathrm{T,l}}$ are the energy and the transverse momentum of the charged lepton, while $p_{z,\mathrm{l}}$ and $p_{z,\nu}$ are the corresponding z components of the charged lepton's or neutrino's momentum. m_{W} is the mass of the W boson given by $m_{\mathrm{W}} = 80.4\ \mathrm{GeV/c^2}$. The solution of equation 5.2 is then given by

$$p_{z,\nu}^{\pm} = \frac{\mu\, p_{z,\mathrm{l}}}{p_{\mathrm{T,l}}^2} \pm \sqrt{\frac{\mu^2 p_{z,\mathrm{l}}^2}{p_{\mathrm{T,l}}^4} - \frac{E_{\mathrm{l}}^2 p_{\mathrm{T},\nu}^2 - \mu^2}{p_{\mathrm{T,l}}^2}} \,, \tag{5.3}$$

where $p_{\mathrm{T},\nu}$ is the transverse momentum of the neutrino. The parameter μ is defined by

$$\mu = \frac{m_{\mathrm{W}}^2}{2} + p_{\mathrm{T,l}} \cdot p_{\mathrm{T},\nu} \cdot \cos\Delta\phi \,, \tag{5.4}$$

with $\Delta\phi$ as the azimuthal angle between the charged lepton and the neutrino.

If the radicand in this solution is positive, then two possible solutions exist, which both are taken into account. But if the radicand is negative, only complex solutions exist. Further studies [156] showed that this occurs when the measured transverse mass $m_{\mathrm{T,W}}$ of the W boson is larger than the nominal W boson mass m_{W}, which can happen due to the finite resolution of the missing transverse energy measurement. Therefore the $p_{x,\nu}$ and $p_{y,\nu}$ components of the neutrino are minimally adjusted in an iterative procedure to achieve a single real solution for $p_{z,\nu}$. In this way the entire four-vector of the neutrino can be reconstructed.

By adding up the four-vectors of the neutrino and the charged lepton, the four-vector of the leptonically decaying W boson $\mathrm{W_{lep}}$ is determined.

The remaining jets in an event have to be matched to the two bottom quarks $\mathrm{b_{had}}$ and $\mathrm{b_{lep}}$ originating from the decay of the two top quarks $\mathrm{t_{had}}$ and $\mathrm{t_{lep}}$, as well as to two light quarks labeled as $\mathrm{q_1}$ and $\mathrm{q_2}$, which originate from the hadronically decaying W boson $\mathrm{W_{had}}$.

Because all possible jet combinations are taken into account, this results in a total of $N \cdot (N-1) \cdot (N-2) \cdot (N-3)$ possible assignments in an event with N jets. Taking into account that the two light quarks $\mathrm{q_1}$ and $\mathrm{q_2}$ can be interchanged without any effect, the number of assignments is halved. With N_{sol}^{ν} as the number of possible solutions for the z component of the neutrino's momentum in equation 5.3, the total number of reconstruction hypotheses is given by

$$N_{\text{hyp}} = \frac{1}{2} \cdot N_{\text{sol}}^{\nu} \cdot N \cdot (N-1) \cdot (N-2) \cdot (N-3)\,. \tag{5.5}$$

A typical event passing the event selection has four or five jets, which results in 24 or 120 possible reconstruction hypotheses, respectively, if the more common case of two solutions for the neutrino is assumed.

After the jet assignment, the four-vector of the hadronically decaying W boson $\mathrm{W_{had}}$ is determined by combining the two light quarks $\mathrm{q_1}$ and $\mathrm{q_2}$. Then the two top quarks $\mathrm{t_{had}}$ and $\mathrm{t_{lep}}$ are reconstructed by adding up the four-vectors of the corresponding W boson and bottom quark, $\mathrm{W_{had}}$ and $\mathrm{b_{had}}$ or $\mathrm{W_{lep}}$ and $\mathrm{b_{lep}}$. The reconstruction of these particles is done for each jet assignment to select the best possible hypothesis for each event in the following step.

5.1.2. Selection of One Reconstruction Hypothesis per Event

For each event one hypothesis has to be selected. Therefore a criterion has to be chosen that is able to distinguish good from bad hypotheses by relying only on reconstructed information.

At first a discriminator d is introduced as a measure for the difference of the event kinematics of simulated and reconstructed events, which is defined as

$$d = \Delta R\left(\mathrm{t_{lep}^{rec}}, \mathrm{t_{lep}^{gen}}\right) + \Delta R\left(\mathrm{t_{had}^{rec}}, \mathrm{t_{had}^{gen}}\right) + \Delta R\left(\mathrm{W_{lep}^{rec}}, \mathrm{W_{lep}^{gen}}\right) + \Delta R\left(\mathrm{W_{had}^{rec}}, \mathrm{W_{had}^{gen}}\right). \tag{5.6}$$

Here, the ΔR function is meant as the distance in the η–ϕ plane between the momentum vectors of the two given objects. This discriminator d describes the sum of the differences between the reconstructed and generated top quarks and W bosons. Lower values of d indicate a better agreement between reconstructed and generated quantities, therefore the hypothesis with the lowest value of d in an event is considered to be the best possible hypothesis.

To select the best hypothesis by referring only to reconstructed objects, another criterion is needed. For this the ψ discriminator [64] is used, which is the product of several likelihood terms depending on different variables. The probability distributions for these likelihood terms are obtained from simulated samples, where the truth information is available and the discriminator d can be determined.

The first three terms in the likelihood discriminant determine the probability of a hypothesis to be the best one based on the masses of the two top quarks and the hadronically decaying W boson. Because the hadronically decaying top quark $\mathrm{t_{had}}$ is reconstructed from the hadronically decaying W boson $\mathrm{W_{had}}$, their masses are strongly correlated. Therefore a linear decorrelation procedure is applied on these three masses, which results in three new masses m_1, m_2 and m_3 with much smaller correlations. These new masses are then given by a linear combination of the reconstructed particles' masses by

$$\begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = \begin{pmatrix} 1.00 & -0.06 & 0.01 \\ 0.06 & 0.93 & 0.37 \\ -0.02 & -0.37 & 0.93 \end{pmatrix} \begin{pmatrix} m_{t,\text{lep}} \\ m_{t,\text{had}} \\ m_{W,\text{had}} \end{pmatrix}. \tag{5.7}$$

The mentioned correlations between the masses of the hadronically decaying top quark and W boson can clearly be seen in this coefficient matrix.

The other terms in the likelihood discriminant are based on the output of the CSV b tagger. This gives additional information about the correct jet assignment by considering the probabilities $P_\mathrm{b}(x)$ of jets with a b tag output of x to be b jets in the best possible hypothesis.

With all these terms, the likelihood ψ is defined as

$$\psi = \mathcal{L}(m_1)\mathcal{L}(m_2)\mathcal{L}(m_3)P_\mathrm{b}(x_{\mathrm{b,lep}})P_\mathrm{b}(x_{\mathrm{b,had}})(1 - P_\mathrm{b}(x_{\mathrm{q1}}))(1 - P_\mathrm{b}(x_{\mathrm{q2}}))\,, \tag{5.8}$$

with $x_{\mathrm{b,had}}$ and $x_{\mathrm{b,lep}}$ as the b tag output of the assigned b jets and x_{q1} and x_{q2} as the b tag output of the assigned light jets.

The hypothesis with the smallest value of ψ is then chosen for each event. With this, the best possible hypothesis is chosen in 29% of all events and in about 73% of all cases the value of $\Delta|y|$ is reconstructed with the correct sign.

In figures 5.1 and 5.2 comparisons of simulated and data samples in various kinematic quantities of the reconstructed objects $\mathrm{t_{had}}$, $\mathrm{t_{lep}}$, $\mathrm{W_{had}}$ and $\mathrm{W_{lep}}$ are shown for the electron+jets and the muon+jets channels. Both channels show a good agreement. Additionally the distributions of the rapidities of the top quarks and the top antiquarks, which are especially important for this analysis, are shown in figure 5.3.

The complete reconstruction of the top-quark pair allows the determination of the distributions of the important variables for this analysis: the sensitive variable $\Delta|y|$, as well as the three secondary variables given by the mass $m_{\mathrm{t\bar{t}}}$, the transverse momentum $p_{\mathrm{T,t\bar{t}}}$ and the absolute value of the rapidity $|y_{\mathrm{t\bar{t}}}|$ of the top-quark pair. The distributions of these variables are shown in figures 5.4 and 5.5 for the reconstructed data samples along with the simulated and data-driven templates normalized to the fit result of the background estimation. From the distribution of the sensitive variable in the combined lepton+jets channel already a first uncorrected value of the charge asymmetry can be calculated, yielding $A_{C,\mathrm{unc}}^{\Delta|y|} = 0.003 \pm 0.002$. But this result is distorted by several effects and can therefore not be compared directly to theory predictions.

The distributions of the rapidity $y_{\mathrm{t\bar{t}}}$ of the $\mathrm{t\bar{t}}$ system in figure 5.5 (e) and (f) show a small excess of events in the central region compared to the fit result of the background estimation. Similarly to the reweighting of the top quark's p_T described in section 4.1.5 an additional reweighting of the simulated $\mathrm{t\bar{t}}$ events according to the results of cross section measurements differential in $y_{\mathrm{t\bar{t}}}$ [150] could be applied. But despite this has been studied to show a nearly perfectly described shape of the $y_{\mathrm{t\bar{t}}}$ distribution, it has a negligible influence on the $\Delta|y|$ shape and on the asymmetry.

5.1.3. Cross Checks

As a cross check of the selection and reconstruction algorithms, the selection and reconstruction steps were applied to an orthogonal dataset from a sideband region. For this purpose, the number of b tagged jets is required to be exactly zero instead of at least one, with all other selection criteria remaining unchanged. This phase space with jets without any b tag is named the *zero-tag* region. The selected events of this region undergo the fitting procedure of the background estimation followed by the reconstruction algorithm. The resulting distributions of the sensitive variable, of the variables used for the background estimation and of the secondary variables are shown in figure 5.6. They all show a good agreement.

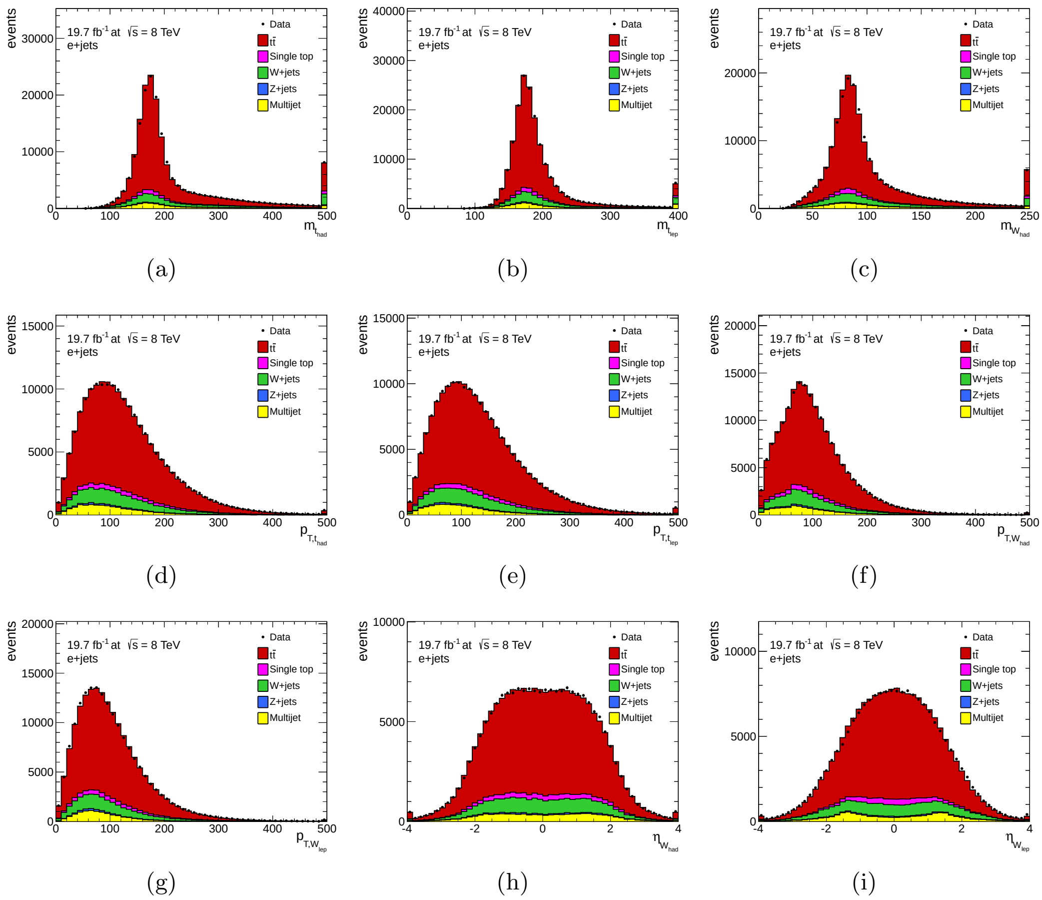


Figure 5.1.: Data to MC comparison after reconstruction for the **electron+jets** channel. The masses of the hadronically decaying top quark (a), of the leptonically decaying top quark (b) and of the hadronically decaying W boson (c) are shown. Further, the transverse momenta of the hadronically and leptonically decaying top quarks ((d) and (e)) and W bosons ((f) and (g)) are depicted. Finally, the pseudorapidities of the hadronically (h) and leptonically (i) decaying W bosons are illustrated. All simulated distributions are normalized to the fit results.

From this sideband a data-driven template of W+jets events was derived by subtracting the contributions of $t\bar{t}$, QCD and Z+jets events. This template is later used to estimate the systematic uncertainties due to mismodeling of the simulated W+jets samples.

5.2. Background Subtraction

The background contributions as determined in the background estimation in section 4.4 make up about 20% of all selected events. These contributions are subtracted from the measured contributions according to the fit results of the background estimation. The shapes of the different background templates are shown in figure 5.7.

As it was shown in the correlation matrix of the background estimation in figure 4.9, some of the fit parameters of the background estimation are correlated. To account for that, the templates $\vec{b}_i$ are decorrelated before subtraction. For this,

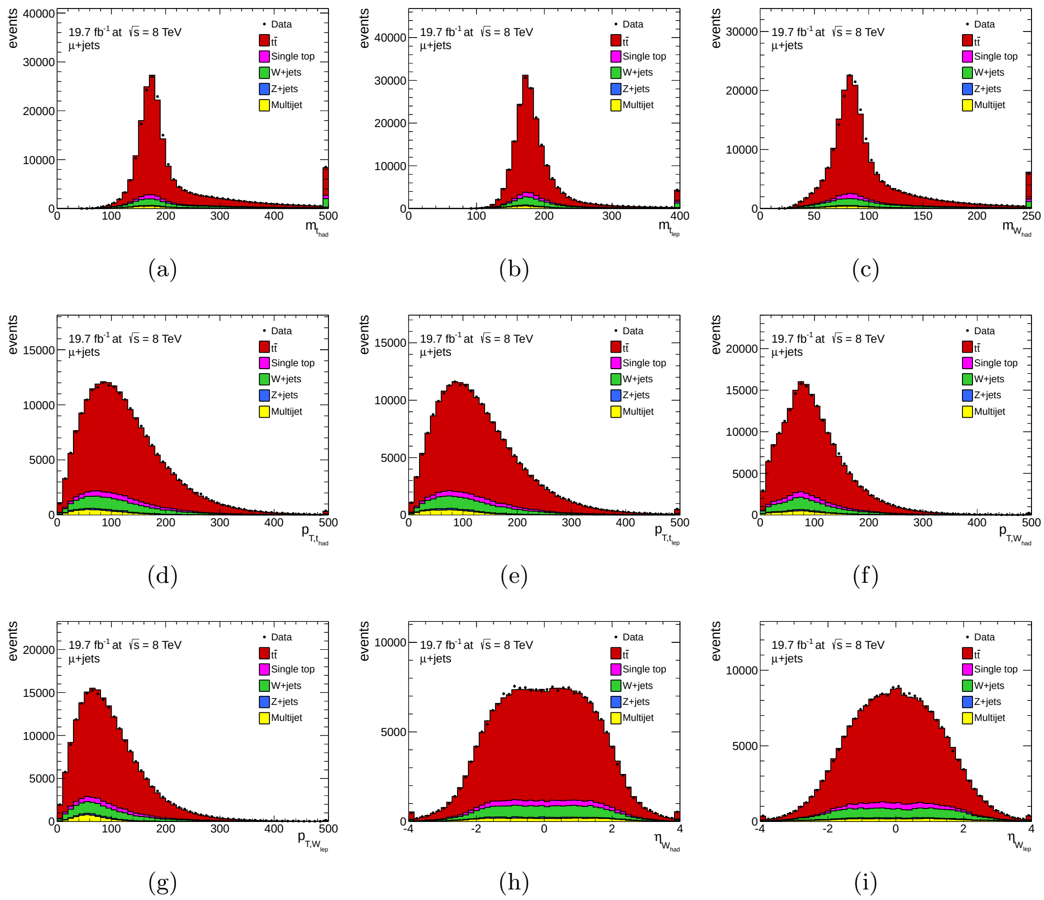
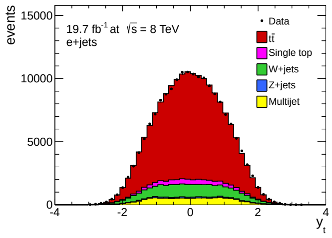


Figure 5.2.: Data to MC comparison after reconstruction for the **muon+jets** channel. The masses of the hadronically decaying top quark (a), of the leptonically decaying top quark (b) and of the hadronically decaying W boson (c) are shown. Further, the transverse momenta of the hadronically and leptonically decaying top quarks ((d) and (e)) and W bosons ((f) and (g)) are depicted. Finally, the pseudorapidities of the hadronically (h) and leptonically (i) decaying W bosons are illustrated. All simulated distributions are normalized to the fit results.

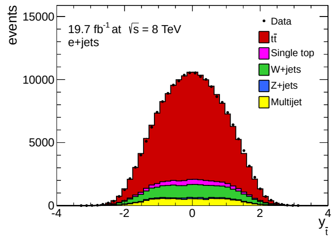
the eigenvalues and eigenvectors $\vec{v}_i$ of the covariance matrix $\mathbf{M}$ of the fit results are determined. These eigenvectors then make up the transformation matrix $\mathbf{V}$, so that $\mathbf{V}^{-1}\mathbf{M}\mathbf{V}$ is diagonal. The decorrelated background contributions are then given by $\vec{b}_{\text{decorr}} = \mathbf{V}\vec{b}$. To keep the normalizations of the different contributions constant, the eigenvectors $\vec{v}_i$ are normalized individually by scale factors $\vec{s}_i$. The array $\vec{s}$ of these scale factors is determined by $\vec{s} = \mathbf{V}^{-1}\vec{\mathbb{1}}$, with $\vec{\mathbb{1}}$ being a vector of ones, so that each row in $\mathbf{V}$ sums up to one. The statistical uncertainties of the background templates' normalizations are then given by the square roots of the corresponding eigenvalues. After the subtraction of the backgrounds an inclusive asymmetry of $A_{C,\text{bgsub}}^{\Delta|y|} = 0.002 \pm 0.002$ is measured in data.

5.3. Unfolding

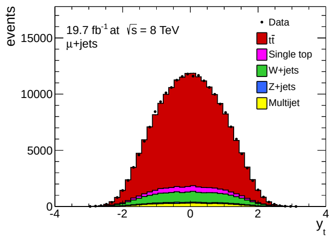
After the subtraction of the background contributions it can be assumed that the resulting distributions correspond to pure signal events. But still these distributions



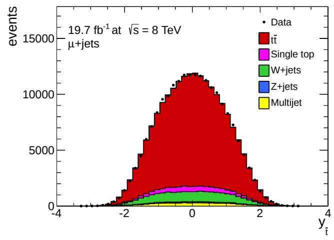
(a)



(b)



(c)



(d)

Figure 5.3.: Comparison of the rapidity distributions of simulated and data samples for the reconstructed top quarks (left) and antiquarks (right), which are important for this analysis. The electron+jets channel is shown in (a) and (b) at the top, the muon+jets channel in (c) and (d) at the bottom. All simulated distributions are normalized to the fit results.

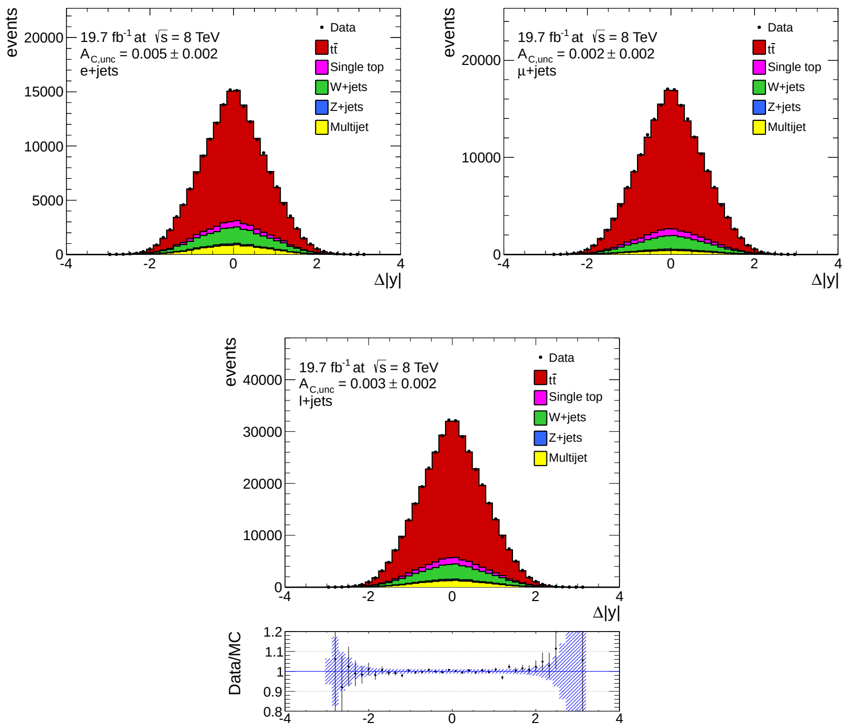


Figure 5.4.: Distributions of the sensitive variable $\Delta|y|$ for the electron+jets, muon+jets, and combined lepton+jets channels. The simulated samples are normalized to the fit results. Additionally, the uncorrected asymmetries are shown for each distribution. For the combined lepton+jets channel also the ratio between data and simulated samples is shown including the statistical uncertainties of the fit result as a hatched band.

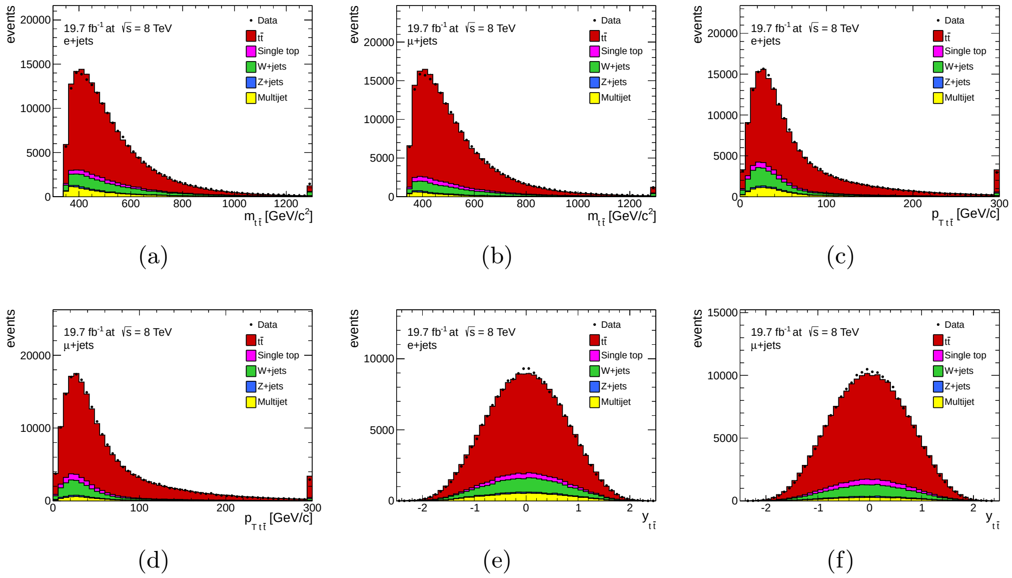


Figure 5.5.: Distributions of the secondary variables $m_{\mathrm{t\bar{t}}}$ ((a) and (b)), $p_{\mathrm{T,t\bar{t}}}$ ((c) and (d)) and $|y_{\mathrm{t\bar{t}}}|$ ((e) and (f)) for the electron+jets and muon+jets channels. The simulated samples are normalized to the fit results.

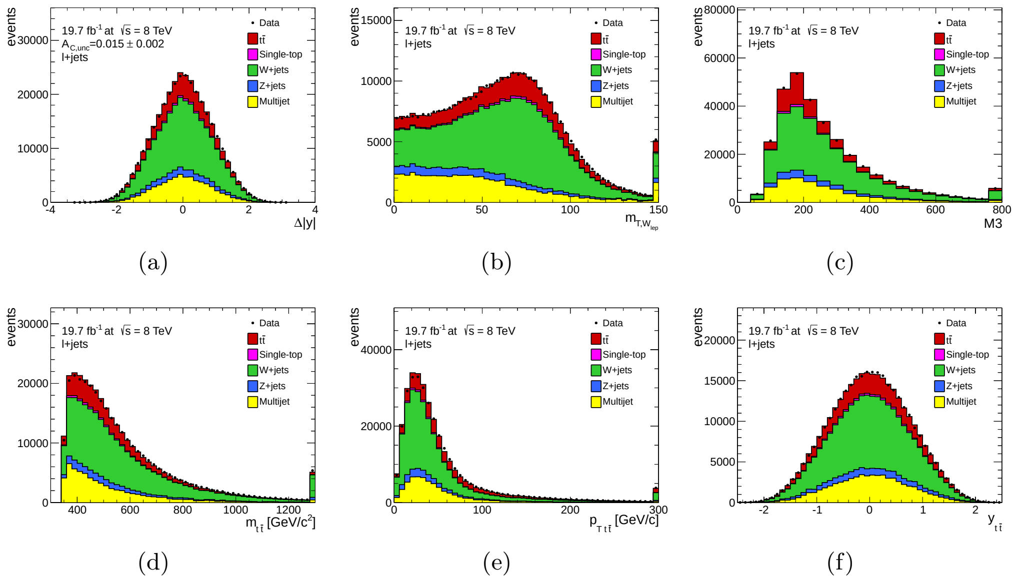


Figure 5.6.: Data to MC comparison after reconstruction for the lepton+jets channel of the zero-tag sideband. The sensitive variable (a), the transverse mass of the leptonically decaying W boson (b) and the M3 mass (c) are shown. In addition the distributions of the three secondary variables $m_{\mathrm{t\bar{t}}}$ (d), $p_{\mathrm{T,t\bar{t}}}$ (e) and $|y_{\mathrm{t\bar{t}}}|$ (f) are displayed. All simulated distributions are normalized to the fit results.

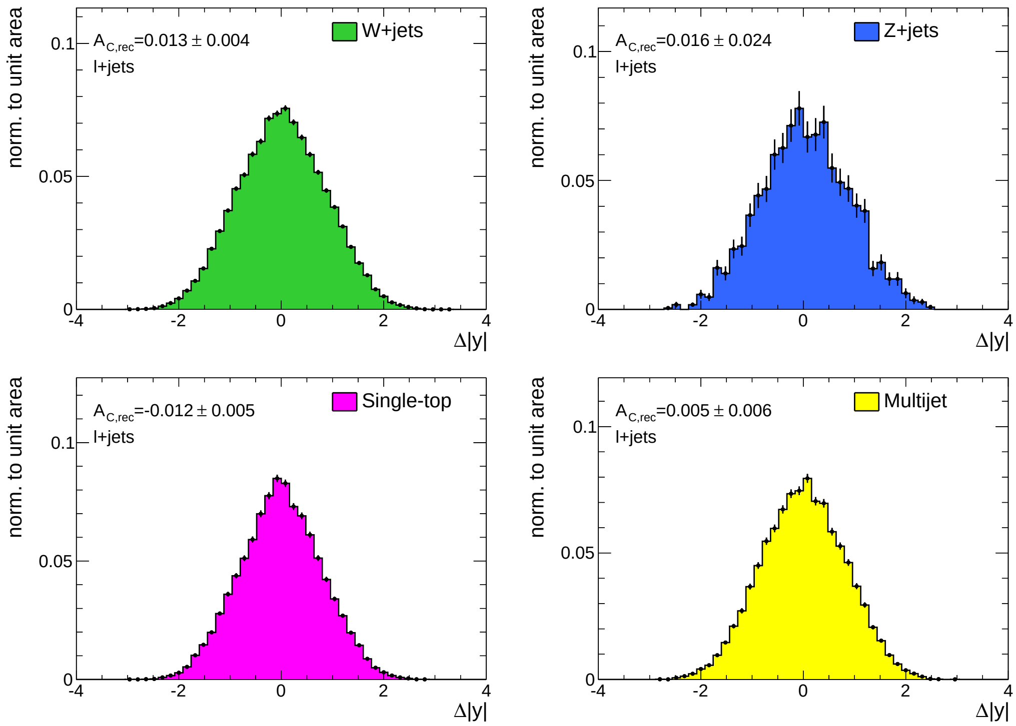


Figure 5.7.: Reconstructed distributions of the sensitive variable $\Delta|y|$ in the templates used for the background subtraction. For each template the individual asymmetry is shown.

cannot be compared to theory predictions directly because several distortions are affecting the measurement.

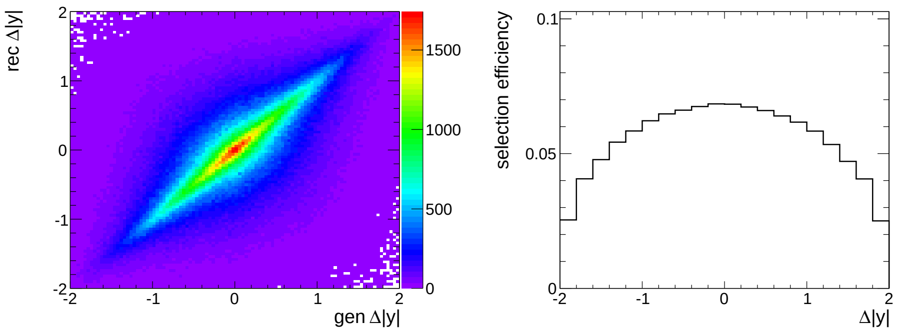
On the one hand, distortions are caused by imperfect reconstruction steps, resulting in migration effects between the bins of the measured distribution. Multiple reasons for these imperfections exist, from uncertainties in the reconstruction of jet energies to the selection of non-optimal hypotheses in the top-quark pair reconstruction. These effects can be summarized in migration matrices that show the reconstructed values over the generated truth values, generated using simulated samples like in figure 5.8(a).

On the other hand, the measurement itself only takes place in a small part of the phase space, which is defined by the event selection steps. To compare the result with theory calculations, an extrapolation to the full phase space or to a fiducial phase space defined on generated quantities has to be applied. The selection efficiency needed for such an extrapolation to the full phase space is shown in figure 5.8(b).

To correct for these effects a regularized unfolding procedure is applied, which is described in detail in section 5.3.3.

Not only an inclusive measurement of the charge asymmetry is performed, but also measurements differential in $m_{\mathrm{t\bar{t}}}$, $p_{\mathrm{T,t\bar{t}}}$ and $y_{\mathrm{t\bar{t}}}$. For all these measurements, a special binning has been chosen, so that the bins in each individual measurement have a comparable number of events. This binning is described in detail in section 5.3.1.

Additionally, the model dependence of the extrapolation can be reduced by only extrapolating to a well defined phase space comparable to the detector's acceptance region instead of extrapolating to the full phase space. The definition of such *visible*



(a) (b)

Figure 5.8.: Illustration of migration effects (a) and selection efficiencies (b). For the migration effects, the reconstructed value of $\Delta|y|$ is shown over the generated value for the selected events. The selection efficiency is calculated by the number of selected events divided by the total number of events in the full phase space. For these studies the simulated $\mathrm{t\bar{t}}$ sample is used.

phase spaces or *fiducial phase spaces* and the implications for the measurement, which is then called *fiducial measurement*, is explained in section 5.3.2 in more detail.

To check the unfolding procedure for correct behavior concerning the values and uncertainties of the unfolded results, consistency checks and linearity tests are performed. These are described in sections 5.3.4 and 5.3.5, respectively.

5.3.1. Choice of Binning

To reduce potential biases a proper binning has to be chosen. Due to the performed unfolding method, a separate binning is needed for the reconstructed and for the unfolded distributions. For the inclusive measurement, the binning is characterized by the number of bins and their edges in the sensitive variable. For the differential measurements the number of bins and their edges in the secondary variables are also part of the binning, which results in two dimensional distributions. For the decision about the binning several considerations are taken into account.

Generally it is advised to use twice as many bins for the reconstructed distribution as for the true one [157]. For the differential measurement this results in twice as many bins for each measured variable. The exact numbers of bins have to be chosen carefully. A larger number of bins leads to a higher accuracy of the estimations of the selection efficiencies and the migration effects, but with too many bins the resulting precision is limited by the number of simulated signal events.

Taking these considerations into account, twenty-four reconstructed and twelve truth bins were chosen for the inclusive measurement. For the differential measurements, sixteen reconstructed and eight truth bins for the sensitive variable and six measured and three truth bins for each of the secondary variables were chosen.

The placement of the bin edges also influences the uncertainties introduced by the unfolding method. So the edges were chosen to achieve a comparable number of events in the reconstructed distribution and in the unfolded distribution before correcting for the selection efficiency, for the inclusive and for each differential measurement. This results in a more stable unfolding procedure.

The final binnings were determined on simulated signal samples and symmetrized around zero in the sensitive variable to keep the method itself free from asymmetries. For the differential measurements for each bin of the secondary variable a separate binning of the sensitive variable has been determined. All these binnings are listed in table 5.1 for the reconstructed distributions and in table 5.2 for the unfolded distributions.

5.3.2. Fiducial Phase Space

For the definition of the fiducial phase space criteria similar to the ones in the default event selection are used: exactly one lepton and four jets. In contrast to the nominal event selection, which is based on reconstructed quantities, these criteria for the fiducial regions are applied to generated quantities.

The single lepton is either an electron with $p_{\mathrm{T}} > 30$ GeV/c and $|\eta| < 2.5$ or a muon with $p_{\mathrm{T}} > 26$ GeV/c and $|\eta| < 2.1$. No additional electrons with $p_{\mathrm{T}} > 20$ GeV/c and $|\eta| < 2.5$ or muons with $p_{\mathrm{T}} > 10$ GeV/c and $|\eta| < 2.5$ should be present.

The requirement of four jets with $p_{\mathrm{T}} > 30$ GeV/c and $|\eta| < 2.5$ can be applied on the jets generated in the showering procedure, called *GenJets*. This means the fiducial region is defined on *particle level*. A measurement in the fiducial region defined on particle level seems appropriate to do from an experimental point of view. But due to the definition based on objects generated in the showering process, no explicit theory predictions are available for the charge asymmetry in this phase space. Still predictions from Monte Carlo simulations do exist, but these yield different asymmetries than theory predictions, as explained in section 4.1.1.

For a comparison with theory predictions, the definition of the fiducial region has to be applied on partons before the showering process, which is then called a fiducial region defined on *parton level*. For a definition that still is close to the default event selection and that allows theory predictions by being based on partons, parton jets are used. These are formed by clustering the quarks and gluons in a simulated event via the anti-k_t algorithm with a distance parameter of $R = 0.5$, as explained in section 3.2.6. On these jets then the requirement of four jets with $p_{\mathrm{T}} > 30$ GeV/c and $|\eta| < 2.5$ is applied.

As an additional criterion for all jets the requirement of the distance ΔR between a jet and a lepton being larger than 0.4 is applied. This enlarges the agreement with the events of the default selection procedure by mimicking the isolation criteria used there.

Because a measurement in a phase space defined on particle level is considered to be more appropriate, but a fiducial region defined on parton level enables comparisons with theory predictions, the measurement was performed in both fiducial regions, in addition to the measurement in the full phase space. The relative amount of events and the overlap of the fiducial regions as well as the generated spectra of the sensitive variable can be seen in figure 5.9.

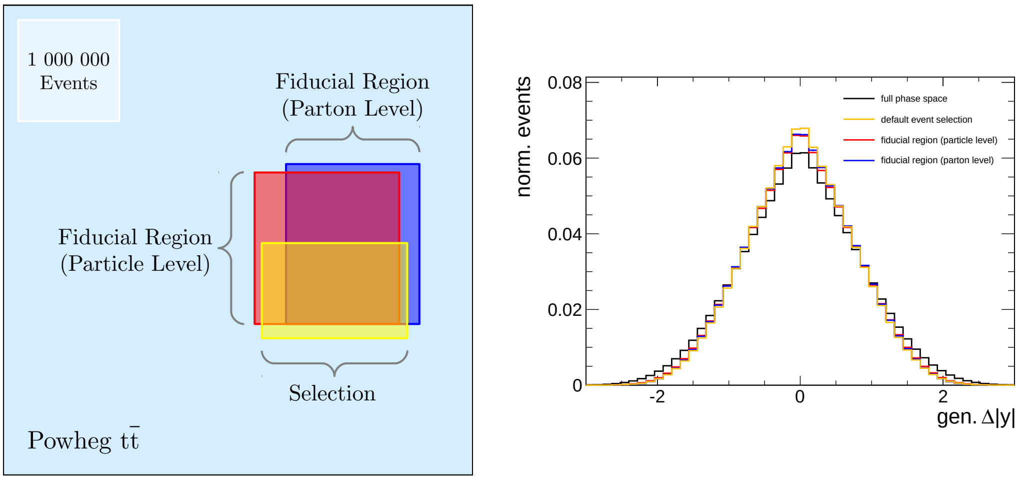
Compared to the distribution of $\Delta|y|$ in the full phase space, the distributions are more narrow for the fiducial phase spaces as well as for the selected events, due to the $|\eta|$ requirements of the jets and the leptons. The two fiducial phase spaces show a large overlap with each other, but they are still larger than the phase space of the event selection. The reason for this lies in the manifold criteria for the event selection procedure. There, the presence of a positive trigger decision is demanded, the criteria for selected leptons are stricter and at least one b tagged jet is required. Also the different definitions of jets for the event selection procedure and for the

		Inclusive Measurement					
$\Delta\|y\|$	Bin 1	$-\infty, -1.27$					
	Bin 2	$-1.27, -1.01$					
	Bin 3	$-1.01, -0.83$					
	Bin 4	$-0.83, -0.70$					
	Bin 5	$-0.70, -0.58$					
	Bin 6	$-0.58, -0.48$					
	Bin 7	$-0.48, -0.38$					
	Bin 8	$-0.38, -0.30$					
	Bin 9	$-0.30, -0.22$					
	Bin 10	$-0.22, -0.14$					
	Bin 11	$-0.14, -0.07$					
	Bin 12	$-0.07, 0.00$					
	Bin 13	$0.00, 0.07$					
	Bin 14	$0.07, 0.14$					
	Bin 15	$0.14, 0.22$					
	Bin 16	$0.22, 0.30$					
	Bin 17	$0.30, 0.38$					
	Bin 18	$0.38, 0.48$					
	Bin 19	$0.48, 0.58$					
	Bin 20	$0.58, 0.70$					
	Bin 21	$0.70, 0.83$					
	Bin 22	$0.83, 1.01$					
	Bin 23	$1.01, 1.27$					
	Bin 24	$1.27, \infty$					
		Differential in $m_{t\bar{t}}$ [GeV/c^2]					
		Bin 1 $0-395$	Bin 2 $395-435$	Bin 3 $435-481$	Bin 4 $481-542$	Bin 5 $542-647$	Bin 6 $647-\infty$
$\Delta\|y\|$	Bin 1	$-\infty, -0.65$	$-\infty, -0.95$	$-\infty, -1.14$	$-\infty, -1.28$	$-\infty, -1.40$	$-\infty, -1.49$
	Bin 2	$-0.65, -0.50$	$-0.95, -0.76$	$-1.14, -0.90$	$-1.28, -1.01$	$-1.40, -1.10$	$-1.49, -1.17$
	Bin 3	$-0.50, -0.39$	$-0.76, -0.60$	$-0.90, -0.71$	$-1.01, -0.79$	$-1.10, -0.86$	$-1.17, -0.92$
	Bin 4	$-0.39, -0.29$	$-0.60, -0.46$	$-0.71, -0.54$	$-0.79, -0.61$	$-0.86, -0.66$	$-0.92, -0.70$
	Bin 5	$-0.29, -0.21$	$-0.46, -0.33$	$-0.54, -0.39$	$-0.61, -0.44$	$-0.66, -0.48$	$-0.70, -0.51$
	Bin 6	$-0.21, -0.14$	$-0.33, -0.21$	$-0.39, -0.25$	$-0.44, -0.28$	$-0.48, -0.31$	$-0.51, -0.33$
	Bin 7	$-0.14, -0.07$	$-0.21, -0.10$	$-0.25, -0.12$	$-0.28, -0.14$	$-0.31, -0.15$	$-0.33, -0.16$
	Bin 8	$-0.07, 0.00$	$-0.10, 0.00$	$-0.12, 0.00$	$-0.14, 0.00$	$-0.15, 0.00$	$-0.16, 0.00$
	Bin 9	$0.00, 0.07$	$0.00, 0.10$	$0.00, 0.12$	$0.00, 0.14$	$0.00, 0.15$	$0.00, 0.16$
	Bin 10	$0.07, 0.14$	$0.10, 0.21$	$0.12, 0.25$	$0.14, 0.28$	$0.15, 0.31$	$0.16, 0.33$
	Bin 11	$0.14, 0.21$	$0.21, 0.33$	$0.25, 0.39$	$0.28, 0.44$	$0.31, 0.48$	$0.33, 0.51$
	Bin 12	$0.21, 0.29$	$0.33, 0.46$	$0.39, 0.54$	$0.44, 0.61$	$0.48, 0.66$	$0.51, 0.70$
	Bin 13	$0.29, 0.39$	$0.46, 0.60$	$0.54, 0.71$	$0.61, 0.79$	$0.66, 0.86$	$0.70, 0.92$
	Bin 14	$0.39, 0.50$	$0.60, 0.76$	$0.71, 0.90$	$0.79, 1.01$	$0.86, 1.10$	$0.92, 1.17$
	Bin 15	$0.50, 0.65$	$0.76, 0.95$	$0.90, 1.14$	$1.01, 1.28$	$1.10, 1.40$	$1.17, 1.49$
	Bin 16	$0.65, \infty$	$0.95, \infty$	$1.14, \infty$	$1.28, \infty$	$1.40, \infty$	$1.49, \infty$
		Differential in $p_{T,t\bar{t}}$ [GeV/c]					
		Bin 1 $0-20.5$	Bin 2 $20.5-32.7$	Bin 3 $32.7-46.8$	Bin 4 $46.8-68.8$	Bin 5 $68.8-117.2$	Bin 6 $117.2-\infty$
$\Delta\|y\|$	Bin 1	$-\infty, -1.18$	$-\infty, -1.18$	$-\infty, -1.20$	$-\infty, -1.20$	$-\infty, -1.19$	$-\infty, -1.16$
	Bin 2	$-1.18, -0.89$	$-1.18, -0.89$	$-1.20, -0.90$	$-1.20, -0.90$	$-1.19, -0.90$	$-1.16, -0.88$
	Bin 3	$-0.89, -0.68$	$-0.89, -0.68$	$-0.90, -0.69$	$-0.90, -0.69$	$-0.90, -0.69$	$-0.88, -0.67$
	Bin 4	$-0.68, -0.51$	$-0.68, -0.51$	$-0.69, -0.52$	$-0.69, -0.52$	$-0.69, -0.52$	$-0.67, -0.50$
	Bin 5	$-0.51, -0.36$	$-0.51, -0.37$	$-0.52, -0.37$	$-0.52, -0.37$	$-0.52, -0.37$	$-0.50, -0.36$
	Bin 6	$-0.36, -0.23$	$-0.37, -0.23$	$-0.37, -0.24$	$-0.37, -0.24$	$-0.37, -0.24$	$-0.36, -0.23$
	Bin 7	$-0.23, -0.11$	$-0.23, -0.11$	$-0.24, -0.12$	$-0.24, -0.11$	$-0.24, -0.11$	$-0.23, -0.11$
	Bin 8	$-0.11, 0.00$	$-0.11, 0.00$	$-0.12, 0.00$	$-0.11, 0.00$	$-0.11, 0.00$	$-0.11, 0.00$
	Bin 9	$0.00, 0.11$	$0.00, 0.11$	$0.00, 0.12$	$0.00, 0.11$	$0.00, 0.11$	$0.00, 0.11$
	Bin 10	$0.11, 0.23$	$0.11, 0.23$	$0.12, 0.24$	$0.11, 0.24$	$0.11, 0.24$	$0.11, 0.23$
	Bin 11	$0.23, 0.36$	$0.23, 0.37$	$0.24, 0.37$	$0.24, 0.37$	$0.24, 0.37$	$0.23, 0.36$
	Bin 12	$0.36, 0.51$	$0.37, 0.51$	$0.37, 0.52$	$0.37, 0.52$	$0.37, 0.52$	$0.36, 0.50$
	Bin 13	$0.51, 0.68$	$0.51, 0.68$	$0.52, 0.69$	$0.52, 0.69$	$0.52, 0.69$	$0.50, 0.67$
	Bin 14	$0.68, 0.89$	$0.68, 0.89$	$0.69, 0.90$	$0.69, 0.90$	$0.69, 0.90$	$0.67, 0.88$
	Bin 15	$0.89, 1.18$	$0.89, 1.18$	$0.90, 1.20$	$0.90, 1.20$	$0.90, 1.19$	$0.88, 1.16$
	Bin 16	$1.18, \infty$	$1.18, \infty$	$1.20, \infty$	$1.20, \infty$	$1.19, \infty$	$1.16, \infty$
		Differential in $\|y_{t\bar{t}}\|$					
		Bin 1 $0-0.16$	Bin 2 $0.16-0.33$	Bin 3 $0.33-0.52$	Bin 4 $0.52-0.73$	Bin 5 $0.73-1.02$	Bin 6 $1.02-\infty$
$\Delta\|y\|$	Bin 1	$-\infty, -0.30$	$-\infty, -0.62$	$-\infty, -0.95$	$-\infty, -1.28$	$-\infty, -1.57$	$-\infty, -1.45$
	Bin 2	$-0.30, -0.24$	$-0.62, -0.55$	$-0.95, -0.86$	$-1.28, -1.15$	$-1.57, -1.32$	$-1.45, -1.10$
	Bin 3	$-0.24, -0.19$	$-0.55, -0.49$	$-0.86, -0.78$	$-1.15, -1.05$	$-1.32, -1.04$	$-1.10, -0.85$
	Bin 4	$-0.19, -0.15$	$-0.49, -0.44$	$-0.78, -0.71$	$-1.05, -0.87$	$-1.04, -0.81$	$-0.85, -0.65$
	Bin 5	$-0.15, -0.11$	$-0.44, -0.39$	$-0.71, -0.62$	$-0.87, -0.64$	$-0.81, -0.59$	$-0.65, -0.47$
	Bin 6	$-0.11, -0.07$	$-0.39, -0.33$	$-0.62, -0.43$	$-0.64, -0.42$	$-0.59, -0.39$	$-0.47, -0.31$
	Bin 7	$-0.07, -0.04$	$-0.33, -0.20$	$-0.43, -0.22$	$-0.42, -0.21$	$-0.39, -0.19$	$-0.31, -0.16$
	Bin 8	$-0.04, 0.00$	$-0.20, 0.00$	$-0.22, 0.00$	$-0.21, 0.00$	$-0.19, 0.00$	$-0.16, 0.00$
	Bin 9	$0.00, 0.04$	$0.00, 0.20$	$0.00, 0.22$	$0.00, 0.21$	$0.00, 0.19$	$0.00, 0.16$
	Bin 10	$0.04, 0.07$	$0.20, 0.33$	$0.22, 0.43$	$0.21, 0.42$	$0.19, 0.39$	$0.16, 0.31$
	Bin 11	$0.07, 0.11$	$0.33, 0.39$	$0.43, 0.62$	$0.42, 0.64$	$0.39, 0.59$	$0.31, 0.47$
	Bin 12	$0.11, 0.15$	$0.39, 0.44$	$0.62, 0.71$	$0.64, 0.87$	$0.59, 0.81$	$0.47, 0.65$
	Bin 13	$0.15, 0.19$	$0.44, 0.49$	$0.71, 0.78$	$0.87, 1.05$	$0.81, 1.04$	$0.65, 0.85$
	Bin 14	$0.19, 0.24$	$0.49, 0.55$	$0.78, 0.86$	$1.05, 1.15$	$1.04, 1.32$	$0.85, 1.10$
	Bin 15	$0.24, 0.30$	$0.55, 0.62$	$0.86, 0.95$	$1.15, 1.28$	$1.32, 1.57$	$1.10, 1.45$
	Bin 16	$0.30, \infty$	$0.62, \infty$	$0.95, \infty$	$1.28, \infty$	$1.57, \infty$	$1.45, \infty$

Table 5.1.: The bin ranges for the reconstructed $\Delta|y|$ distributions in the inclusive measurement as well as in the individual bins of the three secondary variables $m_{t\bar{t}}$ [GeV/c^2], $p_{T,t\bar{t}}$ [GeV/c] and $|y_{t\bar{t}}|$ of the differential measurements.

		Inclusive Measurement		
$\Delta\|y\|$	Bin 1	$-\infty, -1.00$		
	Bin 2	$-1.00, -0.69$		
	Bin 3	$-0.69, -0.47$		
	Bin 4	$-0.47, -0.30$		
	Bin 5	$-0.30, -0.14$		
	Bin 6	$-0.14, 0.00$		
	Bin 7	$0.00, 0.14$		
	Bin 8	$0.14, 0.30$		
	Bin 9	$0.30, 0.47$		
	Bin 10	$0.47, 0.69$		
	Bin 11	$0.69, 1.00$		
	Bin 12	$1.00, \infty$		
		Differential in $m_{\mathrm{t\bar{t}}}$ [GeV/c^2]		
		Bin 1	Bin 2	Bin 3
		$0 - 430$	$430 - 530$	$530 - \infty$
$\Delta\|y\|$	Bin 1	$-\infty, -0.61$	$-\infty, -0.93$	$-\infty, -1.12$
	Bin 2	$-0.61, -0.36$	$-0.93, -0.56$	$-1.12, -0.67$
	Bin 3	$-0.36, -0.16$	$-0.56, -0.26$	$-0.67, -0.31$
	Bin 4	$-0.16, 0.00$	$-0.26, 0.00$	$-0.31, 0.00$
	Bin 5	$0.00, 0.16$	$0.00, 0.26$	$0.00, 0.31$
	Bin 6	$0.16, 0.36$	$0.26, 0.56$	$0.31, 0.67$
	Bin 7	$0.36, 0.61$	$0.56, 0.93$	$0.67, 1.12$
	Bin 8	$0.61, \infty$	$0.93, \infty$	$1.12, \infty$
		Differential in $p_{\mathrm{T,t\bar{t}}}$ [GeV/c]		
		Bin 1	Bin 2	Bin 3
		$0 - 41$	$41 - 92$	$92 - \infty$
$\Delta\|y\|$	Bin 1	$-\infty, -0.86$	$-\infty, -0.88$	$-\infty, -0.88$
	Bin 2	$-0.86, -0.50$	$-0.88, -0.51$	$-0.88, -0.50$
	Bin 3	$-0.50, -0.23$	$-0.51, -0.23$	$-0.50, -0.23$
	Bin 4	$-0.23, 0.00$	$-0.23, 0.00$	$-0.23, 0.00$
	Bin 5	$0.00, 0.23$	$0.00, 0.23$	$0.00, 0.23$
	Bin 6	$0.23, 0.50$	$0.23, 0.51$	$0.23, 0.50$
	Bin 7	$0.50, 0.86$	$0.51, 0.88$	$0.50, 0.88$
	Bin 8	$0.86, \infty$	$0.88, \infty$	$0.88, \infty$
		Differential in $\|y_{\mathrm{t\bar{t}}}\|$		
		Bin 1	Bin 2	Bin 3
		$0 - 0.34$	$0.34 - 0.75$	$0.75 - \infty$
$\Delta\|y\|$	Bin 1	$-\infty, -0.44$	$-\infty, -0.98$	$-\infty, -1.16$
	Bin 2	$-0.44, -0.27$	$-0.98, -0.73$	$-1.16, -0.70$
	Bin 3	$-0.27, -0.13$	$-0.73, -0.40$	$-0.70, -0.34$
	Bin 4	$-0.13, 0.00$	$-0.40, 0.00$	$-0.34, 0.00$
	Bin 5	$0.00, 0.13$	$0.00, 0.40$	$0.00, 0.34$
	Bin 6	$0.13, 0.27$	$0.40, 0.73$	$0.34, 0.70$
	Bin 7	$0.27, 0.44$	$0.73, 0.98$	$0.70, 1.16$
	Bin 8	$0.44, \infty$	$0.98, \infty$	$1.16, \infty$

Table 5.2.: The bin ranges for the unfolded $\Delta|y|$ distributions in the inclusive measurement as well as in the individual bins of the three secondary variables $m_{\mathrm{t\bar{t}}}$ [GeV/c^2], $p_{\mathrm{T,t\bar{t}}}$ [GeV/c] and $|y_{\mathrm{t\bar{t}}}|$ of the differential measurements.



(a) (b)

Figure 5.9.: An illustration of the overlap of the different fiducial phase spaces with the default event selection is shown in (a). The area of the fiducial region defined on particle level is shown in red, of the fiducial region defined on parton level in blue and of the default event selection in yellow. The full phase space is indicated by the whole light blue square. The light gray square in the top left corner corresponds to an amount of one million events, all areas including the overlap regions are drawn to scale. The distributions of the sensitive variable in these phase spaces is shown in (b). The distribution in the full phase space is drawn in black, the other phase spaces are indicated by the same colors as in (a). The overlaps and distributions have been determined on the simulated $t\bar{t}$ sample.

fiducial phase spaces have an influence on whether four jets are present in an event or not. But still the differences between the selected and the fiducial phase spaces are rather quantitative than qualitative and the amount of extrapolation from the selected events to the fiducial phase spaces is about a factor of ten smaller than the extrapolation to the full phase space.

In additional studies different values for the criteria of the fiducial regions were examined to enlarge the overlap with the phase space of the selected events. But the small observed improvements were not sufficient to outweigh the overall loss of events in the fiducial region.

5.3.3. Regularized Unfolding Procedure

To correct for the mentioned migrations and the selection efficiencies, a regularized unfolding procedure [157] is applied. The used method is a generalized matrix inversion method, which is described in the following sections.

Matrix Inversion Method

Generally the distortion of a true spectrum $\vec{x}$ can be described by a transition matrix $\mathbf{A}$. Applying this transition matrix to the true spectrum then results in the measured distribution $\vec{w}$:

$$\vec{w} = \mathbf{A}\vec{x}. \tag{5.9}$$

Here the components of the vectors $\vec{x}$ and $\vec{w}$ represent the contents of the individual bins of the involved histograms and the transition matrix $\mathbf{A}$ is the product of the

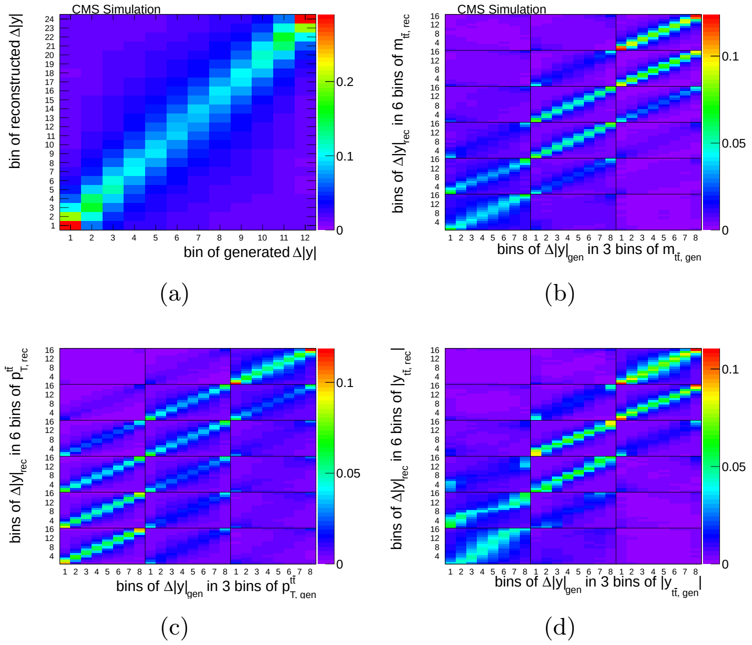


Figure 5.10.: Migration matrices for the inclusive (a) and for the three differential measurements (b), (c) and (d). For this illustration the columns of each matrix are normalized to unity, so that the individual entries correspond to the probability that a selected $\mathrm{t\bar{t}}$ event with given true values of the sensitive and secondary variables is found with the specific reconstructed values of these variables.

migration matrix and the diagonal selection efficiency matrix. To derive the true spectrum $\vec{x}$ from the measured distribution $\vec{w}$, equation 5.9 has to be solved. Taking the covariance matrix $\mathbf{V}_w$ of the measured distribution $\vec{w}$ into account, this problem can be transformed into a least-squares (LS) problem

$$F_{\mathrm{LS}}(\vec{x}) = (\mathbf{A}\vec{x} - \vec{w})^{\mathrm{T}} \, \mathbf{V}_w^{-1} \, (\mathbf{A}\vec{x} - \vec{w}) \,, \tag{5.10}$$

where the solution is obtained by minimizing F_{LS}. By introducing a generalized inverse matrix $\mathbf{A}^{\#}$

$$\mathbf{A}^{\#} = \left(\mathbf{A}^{\mathrm{T}}\, \mathbf{V}_w^{-1}\, \mathbf{A}\right)^{-1} \, \mathbf{A}^{\mathrm{T}}\, \mathbf{V}_w^{-1}, \tag{5.11}$$

a naive solution can be determined by

$$\vec{x}_{\mathrm{LS}} = \mathbf{A}^{\#}\vec{w}, \tag{5.12}$$

but a more advanced method is used to solve this LS problem, which is described later in this chapter.

Migration Matrices

The matrices describing the migration effects between the different bins in the sensitive and the secondary variables are shown in figure 5.10. They are based on simulated events and implemented for the binning described in section 5.3.1.

These corrections for migration effects are applied to the selected events before any acceptance correction. Therefore this step is the same for all extrapolations into separate phase spaces.

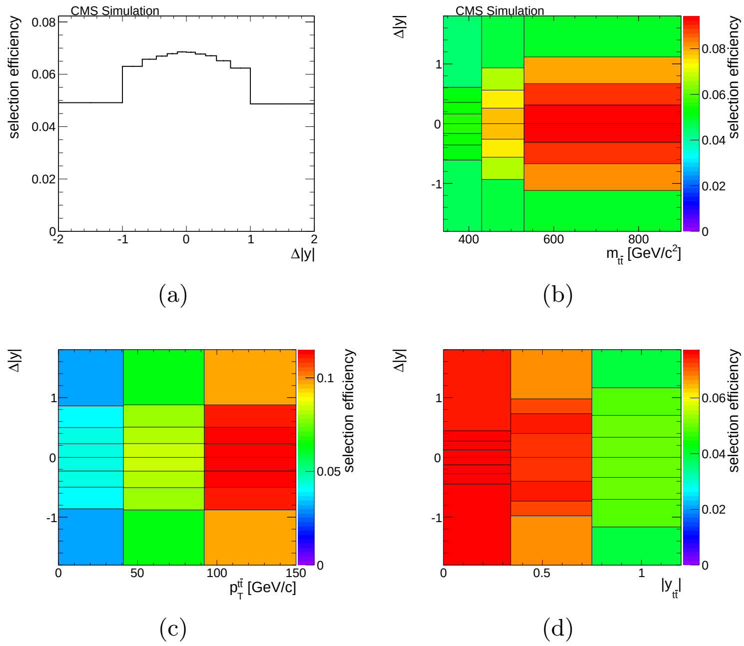


Figure 5.11.: The $\mathrm{t\bar{t}}$ event selection efficiency for the extrapolation to the **full phase space**. The selection efficiencies are shown for the sensitive variable in the inclusive measurement in (a) as well as for the sensitive variable and one of the secondary variables $m_{\mathrm{t\bar{t}}}$, $p_{\mathrm{T,t\bar{t}}}$ and $|y_{\mathrm{t\bar{t}}}|$ of the differential measurements in (b), (c) and (d). The binning was chosen to optimize the unfolding performance.

Selection Efficiencies

After the correction of migration effects, the measured result is extrapolated into different phase spaces: The full phase space and the fiducial phase spaces defined on parton or particle level. The selection efficiencies for this procedure are based on simulated samples. They are shown for the special binning used in this analysis in figure 5.11 for the full phase space and in figures 5.12 and 5.13 for the fiducial phase spaces.

These selection efficiencies are determined by the ratio of the number of selected events to the number of total events. Therefore they also correct the measured distributions for the small amount of selected events, which are not part of any of both fiducial regions.

It can be clearly seen that the total selection efficiency is larger by a factor of ten in the fiducial regions. Also the selection efficiencies for the fiducial regions are much more flat compared to the selection efficiency for the full phase space. This results in all bins being almost equally important for the resulting asymmetry in the fiducial regions. In contrast to that, in the full phase space measurement the outer bins undergo a much larger correction than the inner bins, as can be seen in figure 5.11(a), making these inner bins less important for the result than the outer ones.

Regularization

The transition matrix **A** described above has singular values of different orders of magnitude. Therefore the generalized inversion of **A** will be dominated by the

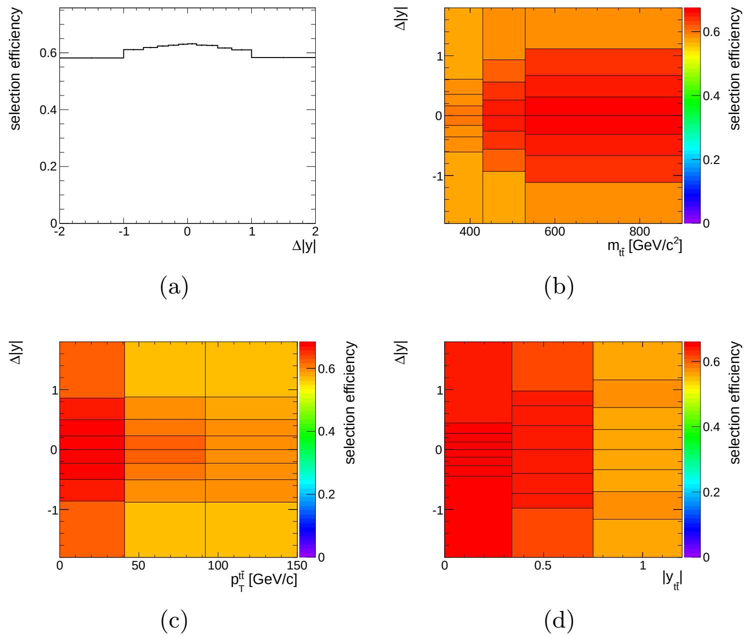


Figure 5.12.: The $t\bar{t}$ event selection efficiency for the extrapolation to the **fiducial phase space defined on particle level**. The selection efficiencies are shown for the sensitive variable in the inclusive measurement in (a) as well as for the sensitive variable and one of the secondary variables $m_{t\bar{t}}$, $p_{T,t\bar{t}}$ and $|y_{t\bar{t}}|$ of the differential measurements in (b), (c) and (d). The binning was chosen to optimize the unfolding performance.

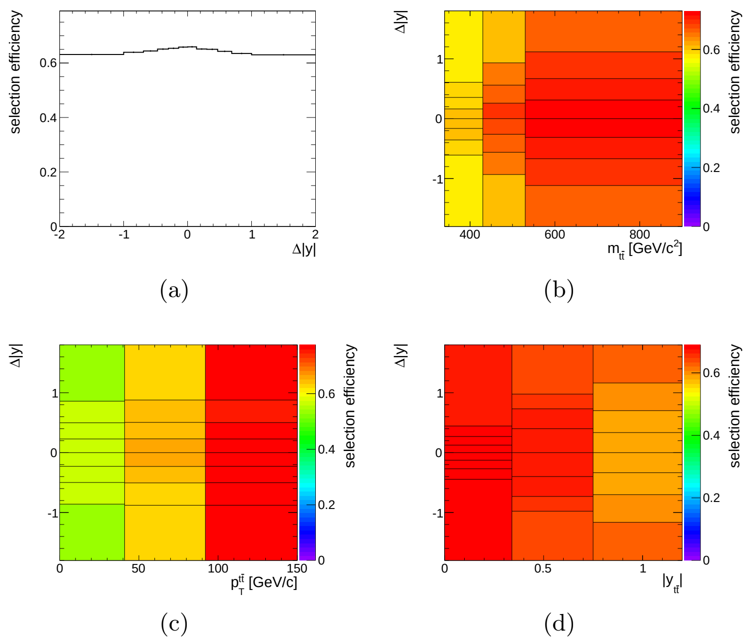


Figure 5.13.: The $t\bar{t}$ event selection efficiency for the extrapolation to the **fiducial phase space defined on parton level**. The selection efficiencies are shown for the sensitive variable in the inclusive measurement in (a) as well as for the sensitive variable and one of the secondary variables $m_{t\bar{t}}$, $p_{T,t\bar{t}}$ and $|y_{t\bar{t}}|$ of the differential measurements in (b), (c) and (d). The binning was chosen to optimize the unfolding performance.

smallest singular values, which correspond to highly fluctuating eigenmodes of $\vec{w}$. This results in unstable and heavily fluctuating solutions even for small changes in $\vec{w}$.

Therefore a more advanced unfolding method is used by introducing two additional terms into the LS formula 5.10 to regularize the solution and avoid unphysical fluctuations [158, 159]:

$$F(\vec{x}, \kappa) = F_{\mathrm{LS}}(\vec{x}) + \kappa \left(N_{\mathrm{obs}} - \sum_{i=1}^{n} (\mathbf{A}\vec{x})_i \right)^2 + \tau \|\mathbf{L}(\vec{x} - \vec{x}_{\mathrm{bias}})\|^2. \tag{5.13}$$

The first new term in this advanced LS problem is proportional to the Lagrangian multiplier κ and sets the norm of the solution to the observed number of events N_{obs}. This is important when the assumption of Gaussian uncertainties instead of Poisson uncertainties for $\vec{w}$ does not hold, which is the case for bins with a small number of observed events [160].

The second term, proportional to the regularization strength τ, introduces a new matrix $\mathbf{L}$. It is defined in such a way that $||\mathbf{L}(\vec{x} - \vec{x}_{\mathrm{bias}})||^2$ is a measure of the summed absolute values of the second derivatives of the distribution $\vec{x} - \vec{x}_{\mathrm{bias}}$.

Because large bin-by-bin fluctuations result in a large curvature, a term proportional to the second derivatives allows to suppress such solutions with large fluctuations. From a Bayesian point of view it can be seen as prior information on the smoothness of the true distribution.

To take the expected shape of the measured distribution into account, a bias distribution $\vec{x}_{\mathrm{bias}}$ is introduced, which is generated from the simulated signal sample. This bias distribution is normalized to the observed number of events with respect to the overall selection efficiency.

The matrix $\mathbf{L}$ introduces terms of the form

$$x_i - 2x_j + x_k \tag{5.14}$$

for the neighboring bins i to k, which can be seen as an approximation of the second derivative of a function f:

$$f'' \approx \frac{f(x + \Delta x) - 2f(x) + f(x - \Delta x)}{\Delta x^2}. \tag{5.15}$$

In two-dimensional distributions sequent bins are considered, which all lie in the same column or row. But the special binning used in this analysis allows a different amount of overlap between horizontally neighboring bins. This is taken into account by considering all possible combinations of horizontally neighboring bins and weighting the curvature from equation 5.14 with an additional factor

$$w = \frac{o_i}{h_j} \cdot \frac{o_k}{h_j}. \tag{5.16}$$

Here, o_i and o_k are the vertical overlaps of the outer bins with the middle bin, which has the height h_j. Considering all combinations these weights w all add up to one.

A proper choice of the regularization parameter τ is important for the regularization method. If it is too small, the regularization will not affect the result and unphysical fluctuations will occur. Otherwise, if τ is too large, the LS problem will be dominated by the regularization term. To find the optimal value of τ the method of minimizing the global correlations [161] is used. Therein the value of τ is chosen which results in a minimum of the mean value of the global correlation coefficients of the result vector. This global correlation coefficient is the total amount of correlation between an element i of the result vector $\vec{x}$ and all other elements. With the covariance matrix $\mathbf{V}_x$ of the result it is defined as

$$\rho_i = \sqrt{1 - \frac{1}{(\mathbf{V}_x^{-1})_{ii} \cdot (\mathbf{V}_x)_{ii}}}. \tag{5.17}$$

Positive correlations are a hint of too much regularization and negative correlations can result from heavily fluctuating distributions. Therefore this approach of minimizing the correlations was chosen.

The actual τ parameters used in the analysis are determined by so-called *pseudo experiments*, which are explained in section 5.3.4. For each inclusive and differential measurement 100 pseudo experiments are performed. In each of these pseudo experiments the best value of τ between $\tau_{\text{min}} = 10^{-6}$ and $\tau_{\text{max}} = 10^{-2}$ is evaluated iteratively and their mean is used for the analysis. Due to different bias distributions in the different phase spaces, τ is also determined independently for the individual phase spaces of the unfolding procedures. The resulting regularization parameters are given in table 5.3.

Measurement in	Inclusive	Differential in		
		$m_{\text{t}\bar{\text{t}}}$	$p_{\text{T},\text{t}\bar{\text{t}}}$	$y_{\text{t}\bar{\text{t}}}$
full phase space	−4.86	−4.61	−4.78	−4.59
fiducial phase space (defined on particle level)	−4.057	−3.717	−3.827	−3.699
fiducial phase space (defined on parton level)	−4.041	−3.696	−3.814	−3.681

Table 5.3.: Base-10 logarithms of the regularization parameters τ used in the individual measurements.

Statistical Uncertainties

The full covariance matrix $\mathbf{V}_x$ of the result allows the correct calculation of the statistical uncertainty of the resulting asymmetry:

$$\sigma^2_{A_\text{C}} = \begin{pmatrix} \frac{\partial A_\text{C}}{\partial N_1} & \dots & \frac{\partial A_\text{C}}{\partial N_n} \end{pmatrix} \mathbf{V}_x \begin{pmatrix} \frac{\partial A_\text{C}}{\partial N_1} \\ \vdots \\ \frac{\partial A_\text{C}}{\partial N_n} \end{pmatrix}, \tag{5.18}$$

with the unfolded bin contents N_i. Applying the partial derivatives on the definition of the asymmetry in equation 1.7, they can be calculated to

$$\frac{\partial A_\text{C}}{\partial N_{i,\text{pos}}} = \frac{2N^-}{(N^+ + N^-)^2} \quad \text{and} \quad \frac{\partial A_\text{C}}{\partial N_{i,\text{neg}}} = -\frac{2N^+}{(N^+ + N^-)^2}, \tag{5.19}$$

for bins corresponding to positive or negative values of the sensitive variable.

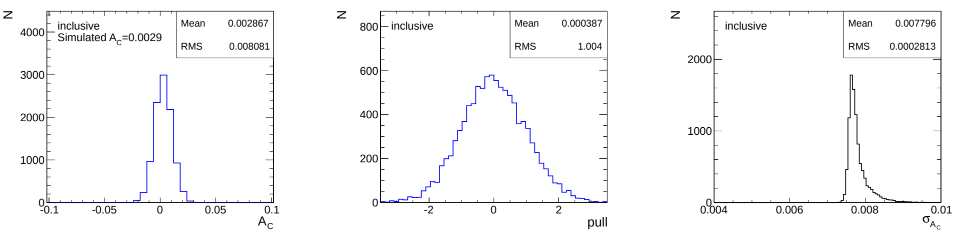


Figure 5.14.: Distribution of the measured asymmetries (left), the pull distributions of the measured asymmetries (middle), and the distribution of the statistical uncertainties (right) for the measured A_C in the inclusive measurement in the fiducial phase space defined on particle level, using pseudo experiments.

5.3.4. Consistency Checks

To check for the consistency of the unfolding method, pseudo experiments are performed. In each pseudo experiment a pseudo dataset is obtained by drawing signal and background events from the MC and data-driven templates according to the background estimation. The number of events of each process is drawn from a Poisson distribution, whose expected number of events is drawn from a Gaussian distribution with mean value and width given by the fit result of the process. The rotation of the backgrounds with correlated fit uncertainties is performed as described in section 5.2 and the fit result of the signal is set to have no uncertainty. These distributions, drawn randomly from simulated or data-driven samples, are then treated like measured data distributions by performing the background subtraction and applying the unfolding method.

For each of the measurement scenarios, the inclusive and the differential ones, extrapolated to the full phase space and to the fiducial phase spaces, 10 000 pseudo experiments are drawn. The unfolded asymmetry values are then filled into a histogram. The mean of this histogram should be identical to the true generated asymmetry of the signal sample. Also the pull P is calculated, which is the difference of true and unfolded asymmetries $A_\mathrm{C}^\mathrm{true}$ and $A_\mathrm{C}^\mathrm{unf}$ divided by the calculated statistical uncertainty σ:

$$P = \frac{A_\mathrm{C}^\mathrm{true} - A_\mathrm{C}^\mathrm{unf}}{\sigma}. \tag{5.20}$$

The resulting distribution of the pulls should have a mean value of zero and a width close to one, which indicates that the calculation of the statistical uncertainty is performed correctly. Additionally, the calculated statistical uncertainties of each measurement are filled into histograms to analyze the variations of the statistical uncertainties.

These distributions are shown for the inclusive measurement in figure 5.14 and for the three differential measurements in figures 5.15 to 5.17 for the fiducial phase space defined on particle level. Both the mean values of the unfolded asymmetries and the widths of the pull distributions are very close to the ideal values, which shows the consistency of this method. The same holds for the consistency checks in the full phase space and the fiducial phase space defined on parton level, which are not shown explicitly here.

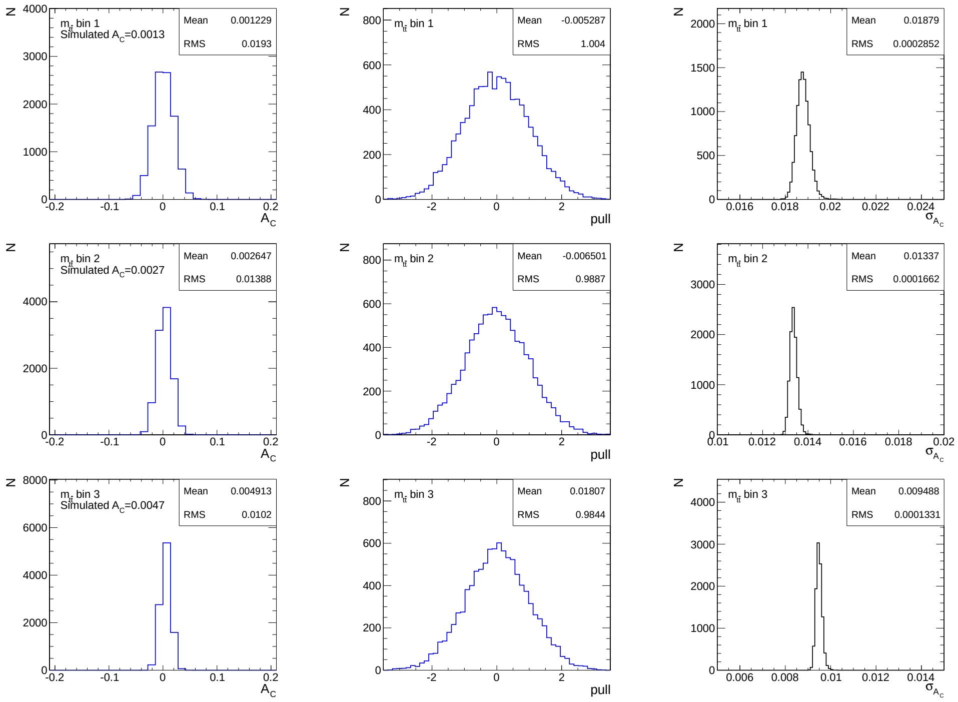


Figure 5.15.: Distribution of the measured asymmetries (left), the pull distributions of the measured asymmetries (middle), and the distribution of the statistical uncertainties (right) for the measured A_C in three bins of $\boldsymbol{m_{t\bar{t}}}$ in the fiducial phase space defined on particle level, using pseudo experiments.

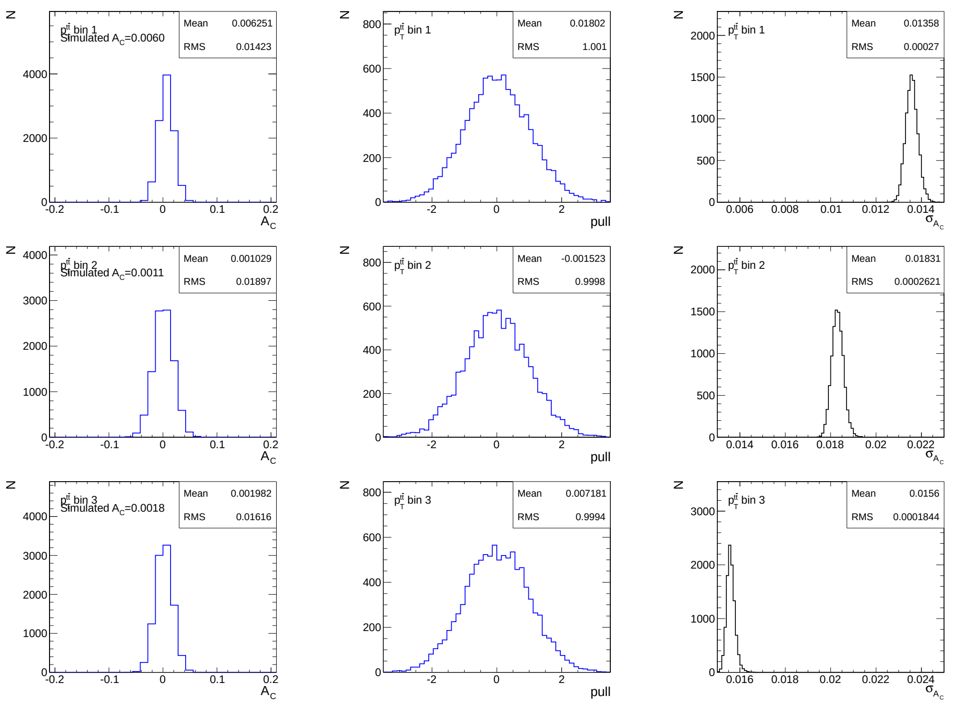


Figure 5.16.: Distribution of the measured asymmetries (left), the pull distributions of the measured asymmetries (middle), and the distribution of the statistical uncertainties (right) for the measured A_C in three bins of $\boldsymbol{p_{T,t\bar{t}}}$ in the fiducial phase space defined on particle level, using pseudo experiments.

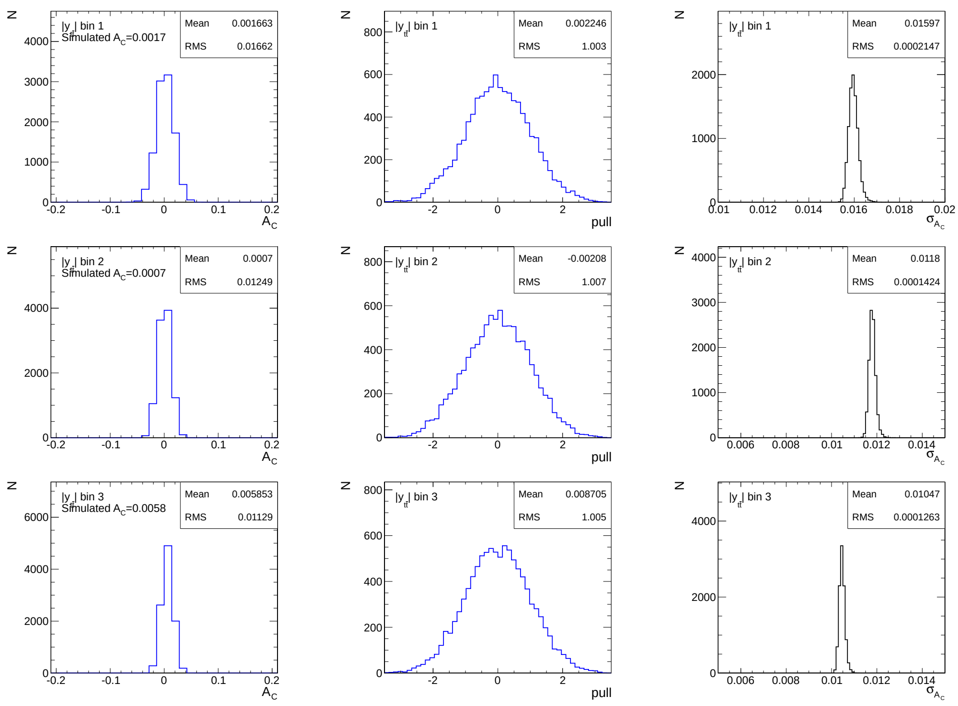


Figure 5.17.: Distribution of the measured asymmetries (left), the pull distributions of the measured asymmetries (middle), and the distribution of the statistical uncertainties (right) for the measured A_C in three bins of $|y_{\mathrm{t\bar{t}}}|$ in the fiducial phase space defined on particle level, using pseudo experiments.

5.3.5. Linearity Tests

For all consistency checks described above the default $\mathrm{t\bar{t}}$ sample with an inclusive asymmetry of $A_\mathrm{C}^\mathrm{sim} = 0.65\%$ is used. To check whether the unfolding method is also sensitive to other values of the charge asymmetry, tests are performed with different values of generated asymmetries. For this the events of the signal sample are reweighted with a factor w according to the value of the sensitive variable, defined by

$$w = k \cdot \Delta|y| + 1. \tag{5.21}$$

The factor k is varied between -0.25 and 0.25, which results in different generated asymmetries. For each value of k, 600 pseudo experiments are performed and the mean value of the unfolded asymmetries is determined, which ideally is identical to the generated reweighted asymmetry.

This linearity test is shown for the inclusive and the differential measurements in the fiducial phase space defined on particle level in figures 5.18 and 5.19. These figures show the unfolded asymmetries over the generated ones. A linear regression of the unfolded asymmetries is shown in red, which can be compared to the bisector in blue, which represents a perfect unfolding procedure. The statistical uncertainties of the individual measurements are represented by the error bars. In most cases they are larger than the deviations between the red and blue lines. Remaining small deviations are covered by the unfolding systematic. The same holds for the linearity tests in the full phase space and the fiducial phase space defined on parton level, which are not shown explicitly here.

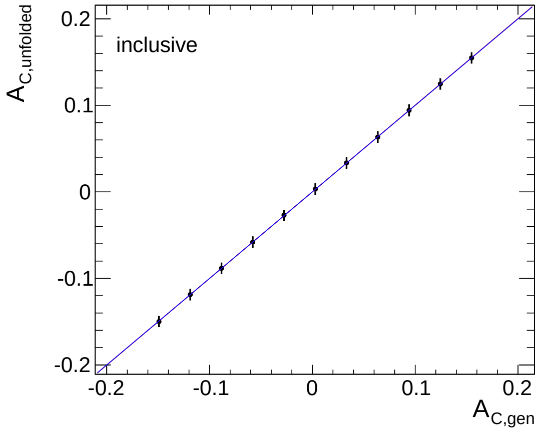


Figure 5.18.: Linearity checks for $\Delta|y|$ unfolded to the fiducial phase space defined on particle level. The mean values of unfolded A_C over the true values of A_C are shown for the inclusive measurement. The error bars indicate the statistical uncertainty of a single measurement. The result of a linear regression is drawn in red and the bisector is shown in blue.

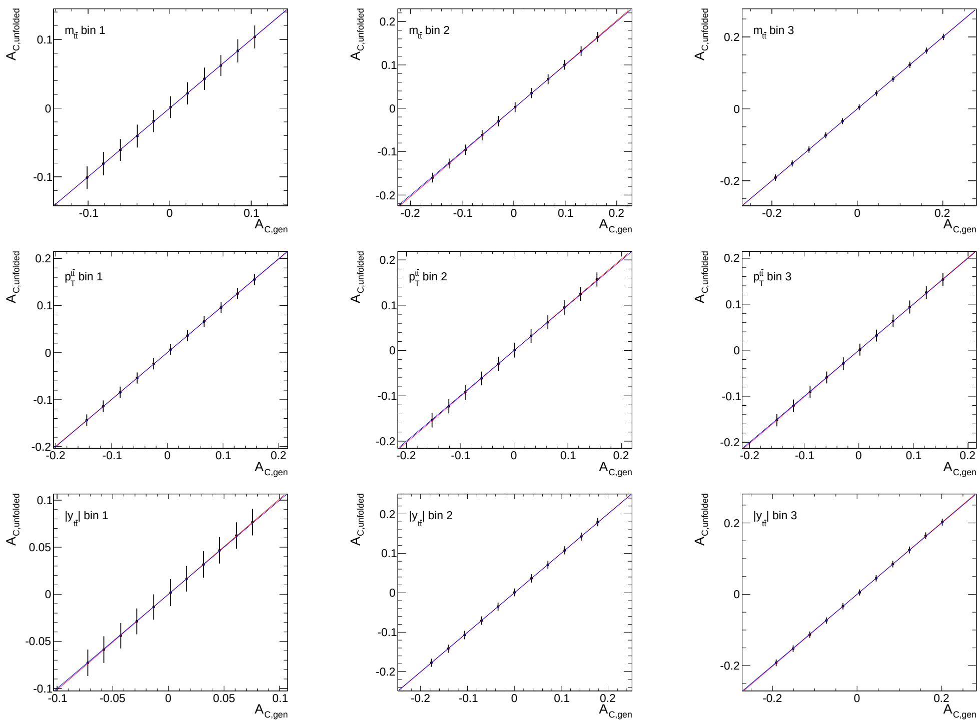


Figure 5.19.: Linearity checks for $\Delta|y|$ unfolded to the fiducial phase space defined on particle level. The mean values of unfolded A_C over the true values of A_C are shown in the three bins (left, center and right) of the sensitive variables $m_{\mathrm{t\bar{t}}}$ (top), $p_{\mathrm{T,t\bar{t}}}$ (middle) and $|y_{\mathrm{t\bar{t}}}|$ (bottom). The error bars indicate the statistical uncertainty of a single measurement. The result of a linear regression is drawn in red and the bisector is shown in blue.

5.4. Systematic Uncertainties

Besides the statistical uncertainties, the measurement is also affected by a number of sources of systematic uncertainties. Due to the measurement of an asymmetry and not of a rate, quantities like the uncertainty on the luminosity or the overall $\mathrm{t\bar{t}}$ selection efficiency and acceptance have no influence here. They would only change the normalization of the $\Delta|y|$ distribution and not the shape. But other sources of systematic uncertainties that affect the shape of the $\Delta|y|$ distribution do exist, which influences the asymmetry.

To determine these uncertainties, a new background estimation followed by the measurement of the asymmetry is performed for each source of systematic uncertainty. Unless described otherwise, the largest observed deviation from the nominal measurement is then taken to be the symmetric systematic uncertainty for the corresponding source.

Jet Energy Scale (JES)

The systematic uncertainty arising from the imperfect knowledge of the jet energy scale is determined by varying the four-momenta of all jets of the simulated samples at the same time by either $+1\sigma$ or -1σ in their η- and p_{T}-dependent uncertainties. The resulting shift in the energy is also propagated to $\not{E}_{\mathrm{T}}$.

Jet Energy Resolution (JER)

Measurements of jet asymmetries suggest that jet p_{T} resolutions are, depending on the $|\eta|$ value of the jet, worse in data compared to simulations. Therefore the difference of the transverse momenta of matched reconstructed and generated jets is scaled by $|\eta|$-depending scale factors, as explained in section 4.1.5. These measurements have themselves an uncertainty of about 6% to 20% [152], depending on the jet-$|\eta|$. To account for this uncertainty, the measurement is performed with up and down variations of the $|\eta|$-depending scale factors corresponding to their uncertainties.

Pile-Up Reweighting Uncertainty

To account for uncertainties in the measured pileup distributions in data events, the simulated samples are reweighted to match systematically shifted versions of the data pileup distribution, as described in [148].

Event Generator

The used unfolding procedure relies to some degree on the correctness of the used MC samples for the $\mathrm{t\bar{t}}$ signal. To account for possible mismodeling effects, the measurement is performed with the selection efficiencies, migration matrices and bias distributions generated from MC@NLO $\mathrm{t\bar{t}}$ signal samples instead of Pythia samples. Because the available MC@NLO samples are interfaced to Herwig as a shower generator, the results are compared to Powheg samples also showered with Herwig instead of Pythia.

The fiducial phase space defined on parton level, as described in section 5.3.2, refers to partons before the showering process. Because of MC@NLO's matching scheme, described in section 3.1.1, these partons are already prepared for the subsequent shower generator, which makes them not suitable for this phase space definition. Therefore no systematic uncertainty is determined for the event generator for this specific measurement in the fiducial phase space defined on parton level.

Model-Dependence of Unfolding Method

The unfolding uncertainty has to be evaluated on pseudo experiments to be able to compare with the true distributions. To estimate the effect of the unfolding procedure, the simulated $\mathrm{t\bar{t}}$ events are reweighted in three scenarios to individually reproduce the asymmetries measured differentially in $m_{\mathrm{t\bar{t}}}$, $p_{\mathrm{T,t\bar{t}}}$ and $|y_{\mathrm{t\bar{t}}}|$. For each of these scenarios pseudo experiments are performed and the unfolding procedures for the inclusive and the differential measurements are applied. The maximum deviation between the reweighted simulated asymmetry and the unfolded asymmetry for a specific bin in the three scenarios is then taken as the uncertainty of this bin.

The three secondary variables together are a good description of the whole $\mathrm{t\bar{t}}$ system, but the three individual variables all describe different properties. So the different reweighting scenarios will produce different effects, whereas the result in data can be seen as a mixture of these three reweighting scenarios. Therefore taking the largest deviation of these scenarios is a good estimate for the uncertainty of the unfolding procedure.

W+Jets Modeling

For the determination of the uncertainty resulting from possible W+jets mismodeling, a data-driven W+jets template taken from the zero-tag sideband region as described in section 5.1.3 is used. In this template the $\mathrm{t\bar{t}}$, Z+jets and QCD contributions are subtracted. Because the fraction of heavy quarks is very different in this sideband sample, this approach can be assumed to estimate the uncertainty in a conservative way.

Multijet Modeling

The data-driven modeling of the QCD process, as described in section 4.3, is biased towards non-isolated leptons. This influences the modeling of the angles between leptons and jets, which affects the asymmetry. To determine the uncertainty due to this effect, the maximum deviation out of three scenarios is taken as a conservative estimation of the uncertainty: Replacing the multijet contribution with the simulated $\mathrm{t\bar{t}}$ or W+jets samples, or inverting the sign of the sensitive variable in the multijet template itself.

b Tagging Uncertainty

The b tagging uncertainty is determined in eight different scenarios using the scale factors and uncertainties given in [149]. Because an overall scale factor of the b tagging efficiency has only a negligible influence on the measured asymmetry, the scale factors are varied separately for light and heavy quarks as well as for jets within and outside of $|\eta| < 0.8$.

The 1σ up- and down variations, separately for light and heavy quarks, make up the first four scenarios. In the other four scenarios the scale factors for jets within $|\eta| < 0.8$ are varied up and for jets with $|\eta| > 0.8$ are varied down or vice versa, again separately for light and heavy quarks. The resulting maximal deviations of these variations are then added in quadrature.

Lepton Identification and Selection Efficiency

Since the nature of a top quark to be a quark or an antiquark is determined by the charge of the lepton, their identification and selection efficiency has an influence on the resulting charge asymmetry.

For the determination of the uncertainty the simulated events are reweighted based on the uncertainties on the lepton scale factors as determined in [162, 163]. This reweighting is performed in such a way that maximally different efficiencies for differently charged leptons within the overall uncertainties are achieved. These variations are done separately for electrons and muons, with the largest observed deviations added in quadrature as the resulting uncertainty.

Factorization and Renormalization Scale Q^2

The statistical uncertainty resulting from the renormalization and factorization scale is determined for the signal sample using dedicated samples generated at scales which are shifted systematically by factors of 2. For the W+jets sample this uncertainty is covered by the data-driven W+jets modeling systematic, which is described above.

An alternative approach of this systematic is to use the nominal sample with events reweighted accordingly to their individual Q^2. Studies for this approach can be found in appendix A.

Parton Shower (PS) and Hadronization

To determine the influence of the showering procedure on the result, the measurement is performed with selection efficiencies, migration matrices and bias distributions generated from Powheg $\mathrm{t\bar{t}}$ signal samples interfaced to Herwig instead of Pythia as in the default sample.

Parton Distribution Functions (PDF)

Uncertainties of the parton distribution functions affect the rate and the shape of the signal and background processes. To account for these effects the selected events are reweighted in 52 different scenarios corresponding to the up and down variations of the 26 eigenvectors of the used CT10 PDF, resulting in 52 alternative templates.

Reweighting of Top-Quark p_T

The p_{T} spectrum of top quarks in simulated $\mathrm{t\bar{t}}$ events is harder than the result of differential cross section measurements [150] and theory predictions. To correct for this effect the generated events are reweighted according to scale factors derived from these measurements. To measure the uncertainty resulting from this reweighting, the measurement is performed again with non-reweighted samples and with samples that have been reweighted twice.

Summary

The shifts of the unfolded asymmetries due to the systematic uncertainties described above are summarized in table 5.4 for the inclusive measurements in all phase spaces and in tables 5.5, 5.6 and 5.7 for the differential measurements in the full and the fiducial phase spaces. The total uncertainty for each bin is obtained by summing up the individual uncertainties in quadrature.

An illustration of the variances corresponding to the systematic uncertainties is shown in figure 5.20, where the variances of the individual systematic sources are stacked on top of each other. There it can be seen that the dominating systematic uncertainties arise from the Q^2 variations and the hadronization procedure, as well as from the generator uncertainty in the full phase space and in the fiducial phase space defined on particle level. As described earlier, for the fiducial phase space on parton level no generator systematic was determined. These three systematic uncertainties have in common that they are all gathered from dedicated samples instead of reweighted versions of the nominal samples. Therefore the large deviations from the nominal measurements are likely to arise from statistical fluctuations, because the statistical uncertainties of the dedicated samples are checked to be of the same magnitude. In the differential $p_{\mathrm{T},\mathrm{t}\bar{\mathrm{t}}}$ bins the multijet systematic also shows large contributions to the total systematical uncertainty. These mostly arise from the inversion of the sign of the sensitive variable in the data-driven QCD template.

	inclusive A_C in phase space		
Systematic uncertainty	full	fiducial (particle level)	fiducial (parton level)
JES	0.001	0.002	0.002
JER	0.001	0.001	0.001
Pileup	0.001	0.001	0.001
Generator	0.003	0.001	n. a.
Unfolding	0.002	0.001	0.001
W+jets	0.002	0.001	0.001
Multijet	0.001	0.001	0.001
b tagging	0.000	0.000	0.000
Lepton ID/sel. efficiency	0.002	0.001	0.001
Q^2 scale	0.003	0.005	0.005
PS + Hadronization	0.000	0.002	0.001
PDF	0.001	0.001	0.001
p_T weighting	0.001	0.000	0.000
Total	0.006	0.006	0.006

Table 5.4.: Systematic uncertainties of the inclusive measurement of the charge asymmetry in the full phase space and in the fiducial phase spaces defined on particle and parton level. Listed are the shifts induced by systematic uncertainties in the measurement on data.

Systematic uncertainty	A_C $m_{\mathrm{t\bar{t}}}$ bin 1	A_C $m_{\mathrm{t\bar{t}}}$ bin 2	A_C $m_{\mathrm{t\bar{t}}}$ bin 3
JES	0.002	0.002	0.004
JER	0.004	0.003	0.001
Pileup	0.003	0.000	0.000
Generator	0.003	0.012	0.001
Unfolding	0.003	0.002	0.004
W+jets	0.007	0.003	0.001
Multijet	0.004	0.005	0.005
b tagging	0.001	0.002	0.001
Lepton ID/sel. efficiency	0.001	0.002	0.002
Q^2 scale	0.009	0.004	0.000
PS + Hadronization	0.008	0.016	0.007
PDF	0.002	0.001	0.002
p_{T} weighting	0.000	0.000	0.001
Total	0.016	0.022	0.011

Systematic uncertainty	A_C $p_{\mathrm{T,t\bar{t}}}$ bin 1	A_C $p_{\mathrm{T,t\bar{t}}}$ bin 2	A_C $p_{\mathrm{T,t\bar{t}}}$ bin 3
JES	0.002	0.004	0.003
JER	0.001	0.002	0.003
Pileup	0.002	0.002	0.002
Generator	0.010	0.010	0.002
Unfolding	0.002	0.002	0.002
W+jets	0.001	0.007	0.001
Multijet	0.005	0.009	0.009
b tagging	0.003	0.002	0.002
Lepton ID/sel. efficiency	0.002	0.002	0.002
Q^2 scale	0.004	0.001	0.004
PS + Hadronization	0.001	0.001	0.000
PDF	0.001	0.001	0.003
p_{T} weighting	0.000	0.002	0.001
Total	0.013	0.017	0.012

Systematic uncertainty	A_C $\lvert y_{\mathrm{t\bar{t}}}\rvert$ bin 1	A_C $\lvert y_{\mathrm{t\bar{t}}}\rvert$ bin 2	A_C $\lvert y_{\mathrm{t\bar{t}}}\rvert$ bin 3
JES	0.005	0.002	0.001
JER	0.005	0.001	0.001
Pileup	0.001	0.002	0.003
Generator	0.009	0.015	0.007
Unfolding	0.001	0.002	0.004
W+jets	0.002	0.004	0.005
Multijet	0.003	0.002	0.002
b tagging	0.001	0.001	0.001
Lepton ID/sel. efficiency	0.001	0.002	0.003
Q^2 scale	0.008	0.008	0.005
PS + Hadronization	0.014	0.013	0.001
PDF	0.001	0.002	0.001
p_{T} weighting	0.001	0.001	0.000
Total	0.020	0.022	0.012

Table 5.5.: Systematic uncertainties of the differential measurements of the charge asymmetry in three bins of the differentiating variables in the **full phase space**. Listed are the shifts induced by systematic uncertainties in the measurement on data.

Systematic uncertainty	A_C $m_{\mathrm{t\bar{t}}}$ bin 1	A_C $m_{\mathrm{t\bar{t}}}$ bin 2	A_C $m_{\mathrm{t\bar{t}}}$ bin 3
JES	0.002	0.002	0.003
JER	0.005	0.003	0.001
Pileup	0.003	0.000	0.001
Generator	0.000	0.007	0.003
Unfolding	0.002	0.002	0.001
W+jets	0.006	0.003	0.003
Multijet	0.004	0.005	0.004
b tagging	0.000	0.000	0.001
Lepton ID/sel. efficiency	0.001	0.001	0.002
Q^2 scale	0.012	0.008	0.002
PS + Hadronization	0.010	0.016	0.007
PDF	0.003	0.002	0.002
p_{T} weighting	0.000	0.000	0.000
Total	0.018	0.021	0.011

Systematic uncertainty	A_C $p_{\mathrm{T,t\bar{t}}}$ bin 1	A_C $p_{\mathrm{T,t\bar{t}}}$ bin 2	A_C $p_{\mathrm{T,t\bar{t}}}$ bin 3
JES	0.005	0.004	0.001
JER	0.001	0.002	0.002
Pileup	0.002	0.001	0.002
Generator	0.001	0.011	0.003
Unfolding	0.001	0.000	0.001
W+jets	0.002	0.003	0.000
Multijet	0.005	0.009	0.008
b tagging	0.006	0.002	0.001
Lepton ID/sel. efficiency	0.002	0.002	0.002
Q^2 scale	0.007	0.002	0.003
PS + Hadronization	0.003	0.001	0.003
PDF	0.002	0.001	0.002
p_{T} weighting	0.001	0.002	0.001
Total	0.013	0.016	0.011

Systematic uncertainty	A_C $\lvert y_{\mathrm{t\bar{t}}}\rvert$ bin 1	A_C $\lvert y_{\mathrm{t\bar{t}}}\rvert$ bin 2	A_C $\lvert y_{\mathrm{t\bar{t}}}\rvert$ bin 3
JES	0.005	0.002	0.001
JER	0.005	0.001	0.001
Pileup	0.001	0.002	0.003
Generator	0.010	0.012	0.004
Unfolding	0.001	0.001	0.002
W+jets	0.002	0.004	0.005
Multijet	0.003	0.002	0.003
b tagging	0.001	0.000	0.001
Lepton ID/sel. efficiency	0.001	0.001	0.003
Q^2 scale	0.006	0.009	0.007
PS + Hadronization	0.015	0.012	0.001
PDF	0.001	0.003	0.001
p_{T} weighting	0.000	0.001	0.000
Total	0.020	0.020	0.011

Table 5.6.: Systematic uncertainties of the differential measurements of the charge asymmetry in three bins of the differentiating variables in the **fiducial phase space defined on particle level**. Listed are the shifts induced by systematic uncertainties in the measurement on data.

Systematic uncertainty	A_C $m_{\mathrm{t\bar{t}}}$ bin 1	A_C $m_{\mathrm{t\bar{t}}}$ bin 2	A_C $m_{\mathrm{t\bar{t}}}$ bin 3
JES	0.002	0.002	0.003
JER	0.005	0.003	0.001
Pileup	0.003	0.000	0.001
Generator	n. a.	n. a.	n. a.
Unfolding	0.002	0.001	0.001
W+jets	0.006	0.003	0.003
Multijet	0.004	0.005	0.004
b tagging	0.000	0.000	0.001
Lepton ID/sel. efficiency	0.001	0.001	0.002
Q^2 scale	0.012	0.008	0.001
PS + Hadronization	0.007	0.017	0.007
PDF	0.002	0.002	0.002
p_{T} weighting	0.000	0.000	0.000
Total	0.017	0.020	0.010

Systematic uncertainty	A_C $p_{\mathrm{T,t\bar{t}}}$ bin 1	A_C $p_{\mathrm{T,t\bar{t}}}$ bin 2	A_C $p_{\mathrm{T,t\bar{t}}}$ bin 3
JES	0.004	0.004	0.002
JER	0.001	0.002	0.003
Pileup	0.002	0.001	0.002
Generator	n. a.	n. a.	n. a.
Unfolding	0.003	0.003	0.001
W+jets	0.002	0.003	0.000
Multijet	0.005	0.009	0.008
b tagging	0.006	0.002	0.001
Lepton ID/sel. efficiency	0.002	0.002	0.002
Q^2 scale	0.007	0.002	0.005
PS + Hadronization	0.002	0.001	0.003
PDF	0.002	0.001	0.002
p_{T} weighting	0.001	0.002	0.000
Total	0.013	0.012	0.011

Systematic uncertainty	A_C $\lvert y_{\mathrm{t\bar{t}}}\rvert$ bin 1	A_C $\lvert y_{\mathrm{t\bar{t}}}\rvert$ bin 2	A_C $\lvert y_{\mathrm{t\bar{t}}}\rvert$ bin 3
JES	0.005	0.002	0.001
JER	0.005	0.001	0.001
Pileup	0.001	0.002	0.003
Generator	n. a.	n. a.	n. a.
Unfolding	0.002	0.001	0.001
W+jets	0.002	0.004	0.006
Multijet	0.003	0.002	0.003
b tagging	0.001	0.000	0.001
Lepton ID/sel. efficiency	0.001	0.001	0.002
Q^2 scale	0.006	0.010	0.007
PS + Hadronization	0.013	0.013	0.002
PDF	0.001	0.002	0.001
p_{T} weighting	0.001	0.001	0.000
Total	0.016	0.018	0.011

Table 5.7.: Systematic uncertainties of the differential measurements of the charge asymmetry in three bins of the differentiating variables in the **fiducial phase space defined on parton level**. Listed are the shifts induced by systematic uncertainties in the measurement on data.

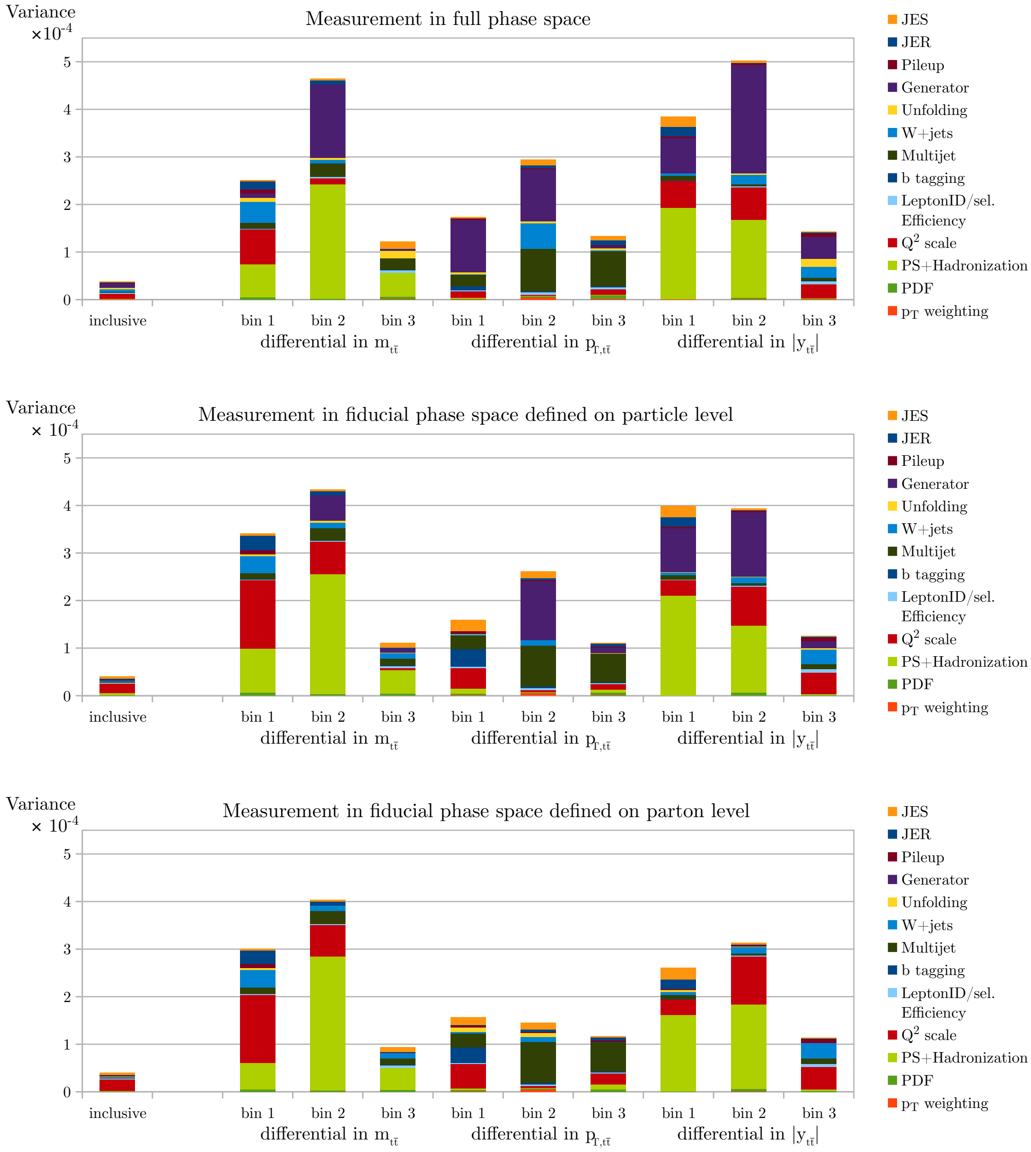


Figure 5.20.: Illustration of the variances for the full and the fiducial phase spaces. The variances of the various systematic uncertainties are stacked for each bin, like they are added up for the total variance. This visualization gives an insight into which systematic uncertainties are dominating in the total uncertainty.

5.5. Results

In table 5.8 the values of the raw uncorrected asymmetry and the asymmetry after background subtraction are shown. The final unfolded asymmetry in the inclusive measurement in the full phase space, along with theory predictions, is also given there, together with the unfolded inclusive asymmetries of the fiducial phase spaces. The unfolded inclusive asymmetry in the full phase space has been calculated from the unfolded $\Delta|y|$ distribution, shown in figure 5.21(a), yielding

$$A_{\mathrm{C}}^{\Delta|y|} = 0.005 \pm 0.007 \text{ (stat.)} \pm 0.006 \text{ (syst.)}. \tag{5.22}$$

It is well compatible with the Standard Model prediction of $A_C^{\Delta|y|,\mathrm{SM}} = 0.0111 \pm 0.0004$ [6]. The unfolded inclusive asymmetries in the fiducial phase spaces defined on particle or parton level have been calculated from the unfolded $\Delta|y|$ distributions shown in figures 5.22(a) and 5.23(a) and yield

$$A_{\mathrm{C}}^{\Delta|y|,\mathrm{fid.particle}} = -0.001 \pm 0.008 \text{ (stat.)} \pm 0.006 \text{ (syst.)} \tag{5.23}$$

and

$$A_{\mathrm{C}}^{\Delta|y|,\mathrm{fid.parton}} = 0.001 \pm 0.008 \text{ (stat.)} \pm 0.006 \text{ (syst.)}. \tag{5.24}$$

The results of the asymmetry in the three differential measurements are given in table 5.9 and figures 5.21(b) to (d) for the measurement in the full phase space. When possible, the measured asymmetries are compared to predictions from Standard Model calculations [6, 10, 65, 164] and to calculations from an effective field theory [165–167], which is able to explain the results of the CDF experiment by introducing an effective axial-vector coupling of the gluon.

The results of the differential fiducial measurements of the charge asymmetry are given in table 5.10 as well as in figures 5.22(b) to (d) for the fiducial phase space defined on particle level and in figures 5.23(b) to (d) for the fiducial phase space defined on parton level, compared to the result of the Powheg simulation. But as explained in section 4.1.1, the simulation implements not all effects that cause the charge asymmetry.

All measured values in the full phase space are consistent with the values predicted by the Standard Model. The individual measurements in the visible phase spaces are comparable with each other. Their systematic uncertainties have slightly decreased in average, compared to the measurement in the full phase space.

But the statistical uncertainties have increased a little. This is due to the different shapes of the selection efficiencies, as described in section 5.3.3, which make the inner bins of the $\Delta|y|$ distribution more important compared to the measurement in the full phase space. Migrations of events in these inner bins are more likely to cause a change of the sign of $\Delta|y|$, resulting in a larger statistical uncertainty if these bins gain more importance.

This has been checked by performing the unfolding procedure with only two or four bins in the $\Delta|y|$ distribution. With only two bins, only one bin exists for each sign of the sensitive variable and no inner or outer bins exist. For both, the fiducial and the full phase spaces, a statistical uncertainty of 0.0048 was observed in the two bin measurement. The measurement in four bins of $\Delta|y|$ is the simplest approach of inner and outer bins. Here, a statistical uncertainty of 0.0071 was observed for the fiducial measurements, compared to 0.0065 in the full phase space measurement, which is the result of the different shapes of the selection efficiencies

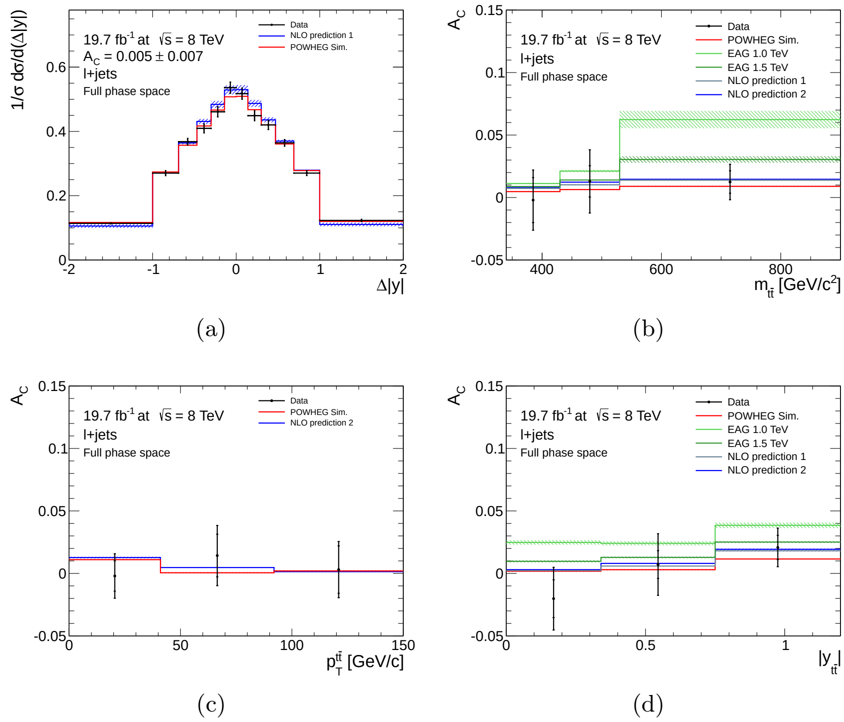


Figure 5.21.: Unfolded inclusive $\Delta|y|$ distribution (a) and the unfolded asymmetry as a function of $m_{t\bar{t}}$ (b), $p_{T,t\bar{t}}$ (c) and $|y_{t\bar{t}}|$ (d) in the **full phase space**. The measured values are compared to the POWHEG simulation as well as to two NLO calculations for the Standard Model (1: [10, 65], 2: [6, 164]) and to predictions of a model featuring an effective axial-vector coupling of the gluon (EAG) at different new physics scales [165–167] where they are available. This model is capable of explaining the results of the CDF experiment [67] at a new physics scale of about 1.3 TeV.

as described. Of course, these uncertainties are only valid for this comparison, because the selection efficiencies are assumed to be equal for large ranges of $\Delta|y|$. These observed differences in the systematical uncertainties are consistent with the observed statistical uncertainties in the measurements in the full and the fiducial phase spaces and thus can explain them.

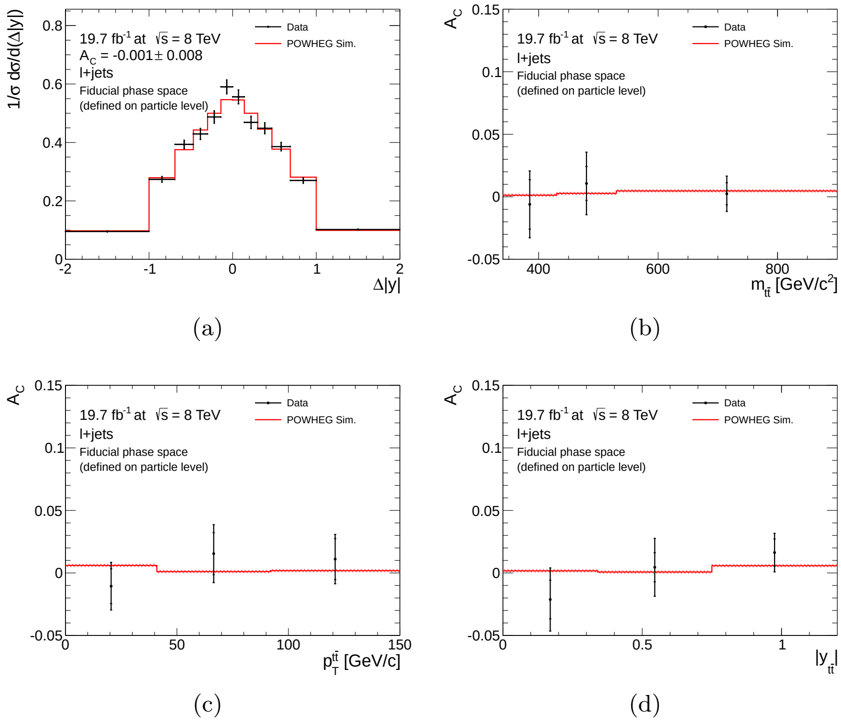


Figure 5.22.: Unfolded inclusive $\Delta|y|$ distribution (a) and the unfolded asymmetry as a function of $m_{\mathrm{t\bar{t}}}$ (b), $p_{\mathrm{T,t\bar{t}}}$ (c) and $|y_{\mathrm{t\bar{t}}}|$ (d) in the **fiducial phase space defined on particle level**. The measured values are compared to the POWHEG simulation.

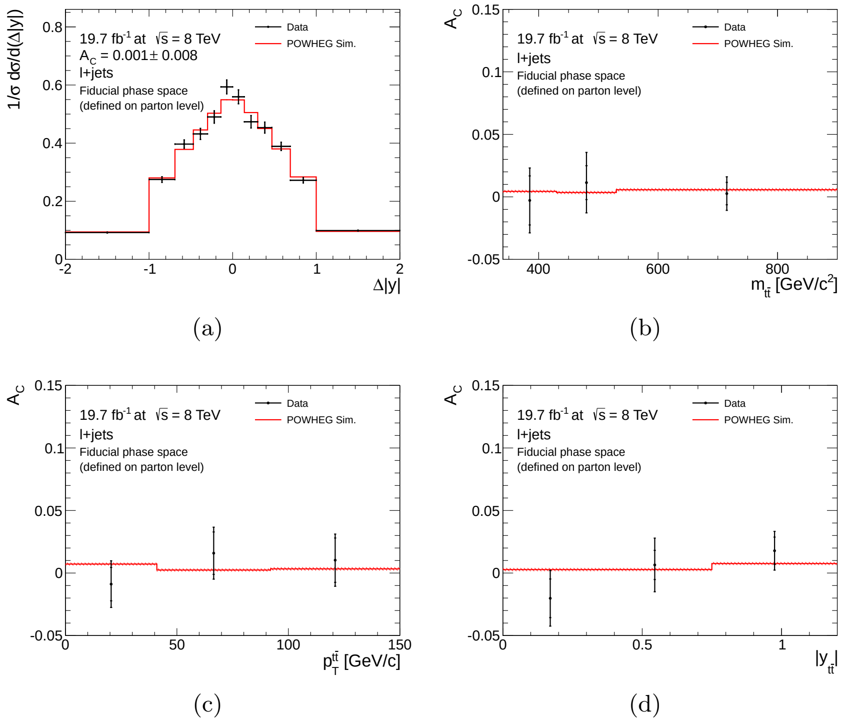


Figure 5.23.: Unfolded inclusive $\Delta|y|$ distribution (a) and the unfolded asymmetry as a function of $m_{\mathrm{t\bar{t}}}$ (b), $p_{\mathrm{T,t\bar{t}}}$ (c) and $|y_{\mathrm{t\bar{t}}}|$ (d) in the **fiducial phase space defined on parton level**. The measured values are compared to the POWHEG simulation.

Asymmetry	A_C
Uncorrected	0.003 ± 0.002 (stat.)
BG-subtracted	0.002 ± 0.002 (stat.)
Final unfolded (full phase space)	0.005 ± 0.007 (stat.) ± 0.006 (syst.)
SM NLO [Kühn, Rodrigo]	0.0102 ± 0.0005
SM NLO [Bernreuther, Si]	0.0111 ± 0.0004
Measured in fiducial phase space (defined on particle level)	-0.001 ± 0.008 (stat.) ± 0.006 (syst.)
Measured in fiducial phase space (defined on parton level)	0.001 ± 0.008 (stat.) ± 0.006 (syst.)

Table 5.8.: The measured **inclusive asymmetry** at the different stages of the analysis and the corresponding theory predictions from the Standard Model [6,10,65,164] for the measurement in the full phase space. Additionally the measured inclusive asymmetries in the fiducial phase spaces are shown.

	A_C in bin 1	A_C in bin 2	A_C in bin 3
$m_{t\bar{t}}$	$-0.002 \pm 0.018 \pm 0.016$	$0.013 \pm 0.013 \pm 0.022$	$0.012 \pm 0.009 \pm 0.011$
NLO prediction 1	0.0073 ± 0.0003	0.0102 ± 0.0004	0.0139 ± 0.0005
NLO prediction 2	0.0082 ± 0.0004	0.0123 ± 0.0003	0.0146 ± 0.0003
$p_{T,t\bar{t}}$	$-0.002 \pm 0.012 \pm 0.013$	$0.014 \pm 0.017 \pm 0.017$	$0.003 \pm 0.019 \pm 0.012$
NLO prediction 2	0.0127 ± 0.0006	0.0047 ± 0.0003	0.0014 ± 0.0002
$\|y_{t\bar{t}}\|$	$-0.020 \pm 0.015 \pm 0.020$	$0.007 \pm 0.011 \pm 0.022$	$0.021 \pm 0.010 \pm 0.012$
NLO prediction 1	0.0023 ± 0.0002	0.0059 ± 0.0003	0.0181 ± 0.0006
NLO prediction 2	0.0030 ± 0.0002	0.0080 ± 0.0003	0.0193 ± 0.0005

Table 5.9.: The unfolded asymmetry values in three bins of the secondary kinematic variables $m_{t\bar{t}}$, $p_{T,t\bar{t}}$ and $|y_{t\bar{t}}|$ along with the Standard Model theory predictions (1: [10,65], 2: [6,164]) in the **full phase space**.

	Measurement in fiducial phase space defined on particle level		
	A_C in bin 1	A_C in bin 2	A_C in bin 3
$m_{t\bar{t}}$	$0.006 \pm 0.020 \pm 0.018$	$0.011 \pm 0.014 \pm 0.021$	$0.002 \pm 0.009 \pm 0.011$
$p_{T,t\bar{t}}$	$-0.011 \pm 0.014 \pm 0.013$	$0.015 \pm 0.017 \pm 0.016$	$0.011 \pm 0.016 \pm 0.011$
$\|y_{t\bar{t}}\|$	$-0.021 \pm 0.015 \pm 0.020$	$0.005 \pm 0.012 \pm 0.020$	$0.016 \pm 0.011 \pm 0.011$
	Measurement in fiducial phase space defined on parton level		
	A_C in bin 1	A_C in bin 2	A_C in bin 3
$m_{t\bar{t}}$	$-0.003 \pm 0.020 \pm 0.017$	$0.011 \pm 0.014 \pm 0.020$	$0.003 \pm 0.009 \pm 0.010$
$p_{T,t\bar{t}}$	$-0.009 \pm 0.013 \pm 0.013$	$0.016 \pm 0.017 \pm 0.012$	$0.010 \pm 0.018 \pm 0.011$
$\|y_{t\bar{t}}\|$	$-0.020 \pm 0.015 \pm 0.016$	$0.006 \pm 0.012 \pm 0.018$	$0.018 \pm 0.011 \pm 0.011$

Table 5.10.: The unfolded asymmetry values in three bins of the secondary kinematic variables $m_{t\bar{t}}$, $p_{T,t\bar{t}}$ and $|y_{t\bar{t}}|$ in the **fiducial phase spaces**.

Summary and Outlook

To measure the charge asymmetry, events from the semileptonic decay channel of the top-quark pair have been selected, where one top quark decays into a bottom quark and two light quarks and the other top quark decays into a bottom quark, a charged lepton and a neutrino. The remaining contributions of backgrounds have been determined via a binned likelihood fit and have then been subtracted accordingly. Further the four-vectors of the top quarks have been reconstructed by using a likelihood criterion to determine the correct assignment of the individual components of an event. To correct for distortions introduced by the event selection and by imperfections of the reconstruction, a regularized unfolding method has been applied. This includes extrapolations of the measured asymmetry to the full phase space and to fiducial phase spaces defined on particle and parton level. The unfolding method has been validated in several consistency and linearity checks.

Finally, the charge asymmetry in top-quark pair-production at the LHC has been measured in the unfolded distribution of $\Delta|y|$, yielding

$$A_{\mathrm{C}}^{\Delta|y|} = \ \ 0.005 \pm 0.007\ (\mathrm{stat.}) \pm 0.006\ (\mathrm{syst.}) \tag{5.25}$$

in the full phase space measurement and

$$A_{\mathrm{C}}^{\Delta|y|,\mathrm{fid.particle}} = -0.001 \pm 0.008\ (\mathrm{stat.}) \pm 0.006\ (\mathrm{syst.}) \tag{5.26}$$

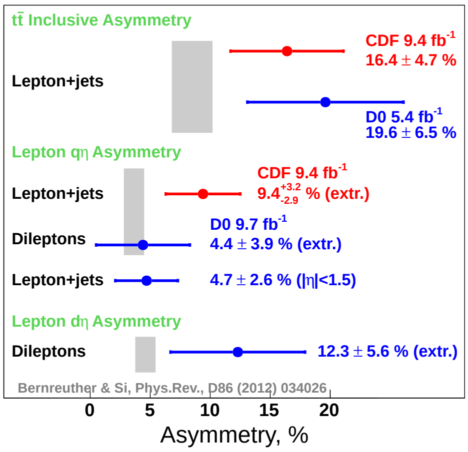
and

$$A_{\mathrm{C}}^{\Delta|y|,\mathrm{fid.parton}} = \ \ 0.001 \pm 0.008\ (\mathrm{stat.}) \pm 0.006\ (\mathrm{syst.}) \tag{5.27}$$

in the fiducial phase spaces defined on particle level and parton level. Additionally, the measurement has been performed differentially in the invariant mass, the transverse momentum and the absolute rapidity of the top-quark pair for the full and the fiducial phase spaces.

All results in the full phase space measurement are in good agreement with Standard Model predictions. The measurements in the fiducial phase spaces are comparable with each other and show in average slightly decreased systematic uncertainties in combination with a small increase of the statistical uncertainties, which is understood. The result of the fiducial measurement defined on parton levels enables future comparisons with theory calculations for the agreement with the Standard Model or with theories of physics beyond the Standard Model. The current results of inclusive charge asymmetry measurements both at the Tevatron and the LHC are shown in 5.24.

To expand the picture of the charge asymmetry, the asymmetry of the produced leptons can be measured, which allows a reduction of the uncertainties. Also the $t\bar{t}$ charge asymmetry can be measured in the dileptonic and full-hadronic decay channels. Further improvements are possible through the next run of the LHC, which is planned for 2015 at a center-of-mass energy of 13 TeV. However, the overwhelming domination of gluon fusion processes at this energy reduces the contributions of quark-antiquark annihilation, where the charge asymmetry occurs. But the huge increase in luminosity still allows further insights of the charge asymmetry by measuring in additional phase spaces or in different sensitive variables. Candidates for this observables are the incline asymmetry, based on the inclination between the planes of initial- and final-state momenta, and the energy asymmetry, based on the energy difference between top quarks and antiquarks [168]. Therefore the topic of the $t\bar{t}$ charge asymmetry will continue to be exciting.



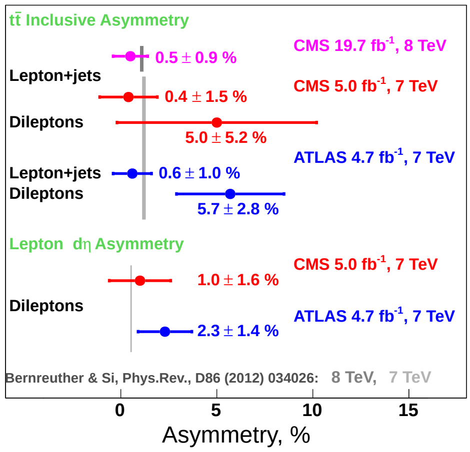


Figure 5.24.: Summary of current results of inclusive charge asymmetry measurements at the Tevatron (left) and the LHC (right), taken from [169], together with theory predictions [6] in gray. The Tevatron results show rather large positive deviations from the theory predictions for the $t\bar{t}$ asymmetry, while the results from ATLAS and CMS in the lepton+jets decay channel represent small negative deviations.

Bibliography

[1] F. Abe et al., “Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab”, Phys. Rev. Lett. 74, 2626–2631 (1995).

[2] S. Abachi et al., “Observation of the Top Quark”, Phys. Rev. Lett. 74, 2632–2637 (1995).

[3] J. H. Kühn and G. Rodrigo, “Charge Asymmetry in Hadroproduction of Heavy Quarks”, Phys. Rev. Lett. 81, 49–52 (1998).

[4] J. H. Kühn and G. Rodrigo, “Charge Asymmetry of Heavy Quarks at Hadron Colliders”, Phys. Rev. D 59, 054017 (1999).

[5] O. Antuñano, J. H. Kühn, and G. Rodrigo, “Top Quarks, Axigluons, and Charge Asymmetries at Hadron Colliders”, Phys. Rev. D 77, 014003 (2008).

[6] W. Bernreuther and Z.-G. Si, “Top Quark and Leptonic Charge Asymmetries for the Tevatron and LHC”, Phys.Rev. D86, 034026 (2012).

[7] S. Alioli, P. Nason, C. Oleari, and E. Re, “A General Framework for Implementing NLO Calculations in Shower Monte Carlo Programs: The PowhegBox”, JHEP 1006, 043 (2010).

[8] P. Nason, “A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms”, JHEP 11, 040 (2004).

[9] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD Computations with Parton Shower Simulations: The Powheg Method”, JHEP 0711, 070 (2007).

[10] G. Rodrigo, Private Communication (2012).

[11] S. L. Glashow, “Partial Symmetries of Weak Interactions”, Nuclear Physics 22, 579 – 588 (1961).

[12] A. Salam and J. Ward, “Electromagnetic and Weak Interactions”, Physics Letters 13, 168 – 171 (1964).

[13] M. Gell-Mann, “A Schematic Model of Baryons and Mesons”, Physics Letters 8, 214 – 215 (1964).

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

[15] S. L. Glashow, J. Iliopoulos, and L. Maiani, “Weak Interactions with Lepton-Hadron Symmetry”, Phys. Rev. D 2, 1285–1292 (1970).

[16] H. Georgi and S. L. Glashow, “Unified Weak and Electromagnetic Interactions without Neutral Currents”, Phys. Rev. Lett. 28, 1494–1497 (1972).

[17] G. 't Hooft, "Renormalizable Lagrangians for Massive Yang-Mills Fields", Nuclear Physics B 35, 167 – 188 (1971).

[18] G. 't Hooft and M. Veltman, "Regularization and Renormalization of Gauge Fields", Nuclear Physics B 44, 189 – 213 (1972).

[19] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Non-Abelian Gauge Theories", Phys. Rev. Lett. 30, 1343–1346 (1973).

[20] H. D. Politzer, "Reliable Perturbative Results for Strong Interactions?", Phys. Rev. Lett. 30, 1346–1349 (1973).

[21] J. J. Aubert et al., "Experimental Observation of a Heavy Particle J", Phys. Rev. Lett. 33, 1404–1406 (1974).

[22] J. E. Augustin et al., "Discovery of a Narrow Resonance in e+ e- Annihilation", Phys. Rev. Lett. 33, 1406–1408 (1974).

[23] S. W. Herb et al., "Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions", Phys. Rev. Lett. 39, 252–255 (1977).

[24] K. Kodama et al., "Observation of Tau Neutrino Interactions", Physics Letters B 504, 218 – 224 (2001).

[25] G. Aad et al., "Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC", Phys.Lett. B716, 1–29 (2012).

[26] S. Chatrchyan et al., "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC", Phys.Lett. B716, 30–61 (2012).

[27] A. Einstein, "Die Grundlage der Allgemeinen Relativitätstheorie", Annalen der Physik 354, 769–822 (1916).

[28] N. Jarosik, C. Bennett, J. Dunkley, B. Gold, M. Greason, et al., "Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Sky Maps, Systematic Errors, and Basic Results", Astrophys.J.Suppl. 192, 14 (2011).

[29] E. Noether, "Invariante Variationsprobleme", Nachr. d. Königl. Ges. d. Wiss. zu Göttingen Math–Phys. Klasse 235–257 (1918).

[30] A. Majorana, "Teoria simmetrica dell'elettrone e del positrone", Nuovo Cimento 14 (1937).

[31] S. Tomonaga, "On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields", Progress of Theoretical Physics 1, 27–42 (1946).

[32] J. Schwinger, "Quantum Electrodynamics. I. A Covariant Formulation", Phys. Rev. 74, 1439–1461 (1948).

[33] J. Schwinger, "Quantum Electrodynamics. II. Vacuum Polarization and Self-Energy", Phys. Rev. 75, 651–679 (1949).

[34] J. Schwinger, "Quantum Electrodynamics. III. The Electromagnetic Properties of the Electron-Radiative Corrections to Scattering", Phys. Rev. 76, 790–817 (1949).

[35] R. P. Feynman, “The Theory of Positrons”, Phys. Rev. 76, 749–759 (1949).

[36] R. P. Feynman, “Space-Time Approach to Quantum Electrodynamics”, Phys. Rev. 76, 769–789 (1949).

[37] R. P. Feynman, “Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction”, Phys. Rev. 80, 440–457 (1950).

[38] J. S. Schwinger, “Renormalisation Theory of Quantum Electrodynamics: An Individual View”, The Birth of Particle Physics 329–353 (1983).

[39] J. Beringer et al., “Review of Particle Physics”, Phys. Rev. D 86, 010001 (2012), and 2013 partial update for the 2014 edition.

[40] The ATLAS, CDF, CMS, D0 Collaborations, “First Combination of Tevatron and LHC Measurements of the Top-Quark Mass”, `arXiv:1403.4427` (2014).

[41] S. Khalil and E. Torrente-Lujan, “Neutrino Mass and Oscillation as Probes of Physics Beyond the Standard Model”, J. Egyptian Math. Soc. 9, 91–141 (2001).

[42] J. W. F. Valle, “Neutrino Physics Overview”, Journal of Physics: Conference Series 53, 473 (2006).

[43] J. Goldstone, “Field Theories with Superconductor Solutions”, Il Nuovo Cimento 19, 154–164 (1961), 10.1007/BF02812722.

[44] J. Goldstone, A. Salam, and S. Weinberg, “Broken Symmetries”, Phys. Rev. 127, 965–970 (1962).

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

[46] P. W. Higgs, “Broken Symmetries and the Masses of Gauge Bosons”, Phys. Rev. Lett. 13, 508–509 (1964).

[47] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles”, Phys. Rev. Lett. 13, 585–587 (1964).

[48] S. Chatrchyan et al., “Study of the Mass and Spin-Parity of the Higgs Boson Candidate Via Its Decays to Z Boson Pairs”, Phys.Rev.Lett. 110, 081803 (2013).

[49] P. A. M. Dirac, “The Quantum Theory of the Emission and Absorption of Radiation”, Proc. R. Soc. Lond. A 114, 243–265 (1927).

[50] M. Kobayashi and T. Maskawa, “CP Violation in the Renormalizable Theory of Weak Interaction”, Prog. Theor. Phys. 49, 652–657 (1973).

[51] J. R. Incandela, A. Quadt, W. Wagner, and D. Wicke, “Status and Prospects of Top-Quark Physics”, Prog.Part.Nucl.Phys. 63, 239–292 (2009).

[52] V. M. Abazov et al., “Observation of Single Top-Quark Production”, Phys. Rev. Lett. 103, 092001 (2009).

[53] T. Aaltonen et al., “Observation of Electroweak Single Top-Quark Production”, Phys. Rev. Lett. 103, 092002 (2009).

[54] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., “New Parton Distributions for Collider Physics”, Phys.Rev. D82, 074024 (2010).

[55] M. Czakon, P. Fiedler, and A. Mitov, “The Total Top Quark Pair Production Cross-Section at Hadron Colliders through O(α_S^4)”, Phys.Rev.Lett. 110, 252004 (2013).

[56] A. Martin, W. Stirling, R. Thorne, and G. Watt, “Parton Distributions for the LHC”, The European Physical Journal C - Particles and Fields 63, 189–285 (2009).

[57] J. L. Rosner, “Prominent Decay Modes of a Leptophobic Z'”, Phys. Lett. B387, 113–117 (1996).

[58] P. H. Frampton, J. Shu, and K. Wang, “Axigluon as Possible Explanation for $\mathrm{p\bar{p}} \to \mathrm{t\bar{t}}$ Forward-Backward Asymmetry”, Phys. Lett. B683, 294–297 (2010).

[59] P. Ferrario and G. Rodrigo, “Massive Color-Octet Bosons and the Charge Asymmetries of Top Quarks at Hadron Colliders”, Phys. Rev. D78, 094018 (2008).

[60] P. Ferrario and G. Rodrigo, “Heavy Colored Resonances in Top-Antitop + Jet at the LHC”, JHEP 02, 051 (2010).

[61] J. Aguilar-Saavedra, “Overview of Models for the $\mathrm{t\bar{t}}$ Asymmetry”, Nuovo Cim. C035N3, 167–172 (2012).

[62] J. Aguilar-Saavedra and M. Perez-Victoria, “Simple Models for the Top Asymmetry: Constraints and Predictions”, JHEP 1109, 097 (2011).

[63] J. Aguilar-Saavedra and M. Perez-Victoria, “$\mathrm{t\bar{t}}$ Charge Asymmetry, Family and Friends”, J.Phys.Conf.Ser. 447, 012015 (2013).

[64] F. Roscher, “Differential Measurement of the Charge Asymmetry in Top Quark Pair Production at the CMS-Detector”, Diploma Thesis, Karlsruhe Institute of Technology IEKP-KA/2012-6 (2012).

[65] J. H. Kühn and G. Rodrigo, “Charge Asymmetries of Top Quarks at Hadron Colliders Revisited”, JHEP 01, 63 (2012).

[66] T. Aaltonen et al., “Forward-Backward Asymmetry in Top-Quark Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV”, Phys. Rev. Lett. 101, 202001 (2008).

[67] T. Aaltonen et al., “Measurement of the Top Quark Forward-Backward Production Asymmetry and its Dependence on Event Kinematic Properties”, Phys.Rev. D87, 092002 (2013).

[68] DØ Collaboration, “Measurement of the forward-backward asymmetry in $p\bar{p} \to t\bar{t}$ production in the l+jets channel”, DØ Note 6425-CONF (2014).

[69] T. Peiffer, “First Measurement of the Charge Asymmetry and Search for Heavy Resonances in Top Quark Pair Production with the CMS Experiment”, PhD Thesis, Karlsruhe Institute of Technology, CERN-THESIS-2009-208 (2011).

[70] CMS Collaboration, “Measurement of the Charge Asymmetry in Top Quark Pair Production with the CMS Experiment”, CMS Physics Analysis Summary TOP-10-010 (2011).

[71] G. Aad et al., “Measurement of the Top Quark Pair Production Charge Asymmetry in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector”, JHEP 1402, 107 (2014).

[72] S. Chatrchyan et al., “Inclusive and Differential Measurements of the $\mathrm{t\bar{t}}$ Charge Asymmetry in Proton-Proton Collisions at 7 TeV”, Phys.Lett. B717, 129–150 (2012).

[73] CMS Collaboration, “Combination of ATLAS and CMS $\mathrm{t\bar{t}}$ Charge Asymmetry Measurements using LHC Proton-Proton Collisions at $\sqrt{s} = 7$ TeV”, CMS Physics Analysis Summary TOP-14-006 (2014).

[74] E. Alvarez, “Improving Top Quark Induced Charge Asymmetries at the LHC using $\mathrm{t\bar{t}}$ Transverse Momentum”, Phys.Rev. D85, 094026 (2012).

[75] A. Einstein, “Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?”, Annalen der Physik 18, 639–643 (1905).

[76] L. Evans and P. Bryant, “LHC Machine”, JINST 3, S08001 (2008).

[77] G. Aad et al., “The ATLAS Experiment at the LHC”, JINST 3, S08003 (2008).

[78] CMS Collaboration, “CMS Physics Technical Design Report Volume I: Detector Performance and Software”, CERN-LHCC-2006-01 (2006).

[79] CMS Collaboration, “The CMS Experiment at the CERN LHC”, JINST 3, S08004 (2008).

[80] K. Aamodt et al., “The ALICE Experiment at the LHC”, JINST 3, S08002 (2008).

[81] A. A. Alves et al., “The LHCb Detector at the LHC”, JINST 3, S08005 (2008).

[82] J. Gruschke, “Observation of Top Quarks and First Measurement of the $\mathrm{t\bar{t}}$ Production Cross Section at a Centre-Of-Mass Energy of 7 TeV with the CMS Experiment at the LHC”, PhD Thesis, Karlsruhe Institute of Technology, CERN-THESIS-2011-030 (2011).

[83] C. E. Hill and M. O’Neill, “The High Voltage System for the High Intensity CERN Proton Source”, CERN-PS-98-035-HP 4 p (1998).

[84] C. E. Hill, A. M. Lombardi, E. Tanke, and M. Vretenar, “Present Performance of the CERN Proton Linac”, CERN-PS-98-045-HP 4 p (1998).

[85] K. Schindl, "The PS Booster as Pre-Injector for LHC", Part. Accel. 58, 63–78 (1997).

[86] R. Cappi, "The PS in the LHC Injector Chain", Part. Accel. 58, 79–89 (1997).

[87] T. Linnecar, "Preparing the SPS for LHC", Part. Accel. 58, 91–101 (1997).

[88] CMS Collaboration, "CMS Luminosity Collision Data" `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults` (February 2014).

[89] CMS Collaboration, "How Was CMS Designed?" `http://cms.web.cern.ch/cms/Detector/Designed/index.html` (November 2011).

[90] CMS Collaboration, "The CMS Tracker System Project: Technical Design Report", CERN-LHCC-98-06 (1997).

[91] CMS Collaboration, "The CMS Tracker: Addendum to the Technical Design Report", CERN-LHCC-2000-016 (2000).

[92] CMS Collaboration, "The CMS Electromagnetic Calorimeter Project: Technical Design Report", CERN-LHCC-97-33 (1997).

[93] CMS Collaboration, "Changes to the CMS ECAL Electronics: Addendum to the Technical Design Report", CERN-LHCC-2002-027 (2002).

[94] CMS Collaboration, "The CMS Hadronic Calorimeter Project: Technical Design Report", CERN-LHCC-97-31 (1997).

[95] CMS Collaboration, "The CMS Muon Project: Technical Design Report", CERN-LHCC-97-32 (1997).

[96] CMS Collaboration, "The TriDAS Project Technical Design Report, Volume 1: The Trigger Systems", CERN-LHCC-2000-038 (2000).

[97] CMS Collaboration, "The TriDAS Project Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger", CERN-LHCC-2002-026 (2002).

[98] WLCG Collaboration, "LHC Computing Grid: Technical Design Report", LCG TDR, CERN, Geneva, CERN-LHCC-2002-024 (2005).

[99] M. A. Dobbs et al., "Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics", Nuclear Physics A 49 (2004).

[100] V. N. Gribov and L. N. Lipatov, "Deep Inelastic ep Scattering in Perturbation Theory", Sov. J. Nucl. Phys. 15, 438–450 (1972).

[101] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language", Nucl. Phys. B 126, 298 (1977).

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

[103] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, “Parton Fragmentation and String Dynamics”, Phys. Rept. 97, 31–145 (1983).

[104] T. Stelzer and W. F. Long, “Automatic Generation of Tree Level Helicity Amplitudes”, Comput. Phys. Commun. 81, 357–371 (1994).

[105] F. Maltoni and T. Stelzer, “MadEvent: Automatic Event Generation with MadGraph”, JHEP 02, 027 (2003).

[106] J. Alwall et al., “MadGraph/MadEvent v4: The New Web Generation”, JHEP 09, 028 (2007).

[107] M. L. Mangano et al., “ALPGEN, a Generator for Hard Multiparton Processes in Hadronic Collisions”, JHEP 07, 001 (2003).

[108] S. Frixione and B. R. Webber, “Matching NLO QCD Computations and Parton Shower Simulations”, JHEP 06, 029 (2002).

[109] S. Frixione, P. Nason, and B. R. Webber, “Matching NLO QCD and Parton Showers in Heavy Flavour Production”, JHEP 08, 007 (2003).

[110] T. Sjöstrand et al., “High Energy Physics Event Generation with Pythia 6.1”, Comput. Phys. Commun. 135, 238–259 (2001).

[111] G. Corcella et al., “Herwig 6 : An Event Generator for Hadron Emission Reactions with Interfering Gluons (Including Supersymmetric Processes)”, JHEP 1, 10 (2001).

[112] S. Catani, F. Krauss, R. Kühn, and B. R. Webber, “QCD Matrix Elements + Parton Showers”, JHEP 11, 063 (2001).

[113] F. Krauss, “Matrix Elements and Parton Showers in Hadronic Interactions”, JHEP 08, 015 (2002).

[114] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, “Matching Matrix Elements and Shower Evolution for Top-Quark Production in Hadronic Collisions”, JHEP 01, 013 (2007).

[115] J. Alwall et al., “A Standard Format for Les Houches Event Files”, Comput. Phys. Commun. 176, 300–304 (2007).

[116] M. Bahr, S. Gieseke, M. Gigg, D. Grellscheid, K. Hamilton, et al., “Herwig++ Physics and Manual”, Eur.Phys.J. C58, 639–707 (2008).

[117] S. Jadach, J. H. Kühn, and Z. Was, “TAUOLA: A Library of Monte Carlo Programs to Simulate Decays of Polarized Tau Leptons”, Comput. Phys. Commun. 64, 275–299 (1990).

[118] S. Agostinelli et al., “Geant4: A Simulation Toolkit”, Nucl. Instrum. Meth. A506, 250–303 (2003).

[119] CMS Collaboration, “Fast Simulation of the CMS Detector at the LHC”, CERN-CMS-CR-2010-297 (2010).

[120] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, $\not\!\!E_{\mathrm{T}}$, and Taus", CMS Physics Analysis Summary PFT-09-001 (2009).

[121] CMS Collaboration, "Commissioning of the Particle-Flow Event Reconstruction with the First LHC collisions Recorded in the CMS Detector", CMS Physics Analysis Summary PFT-10-001 (2010).

[122] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", CMS Physics Analysis Summary PFT-10-002 (2010).

[123] T. Speer et al., "Track Reconstruction in the CMS Tracker", Nucl. Instrum. Meth. A559, 143–147 (2006).

[124] R. Frühwirth, "Application of Kalman Filtering to Track and Vertex Fitting", Nucl. Instrum. Meth. A262, 444–450 (1987).

[125] S. Baffioni et al., "Electron Reconstruction in CMS", Eur. Phys. J. C49, 1099–1116 (2007).

[126] R. Frühwirth, "Track Fitting with Non-Gaussian Noise", Comput. Phys. Commun. 100, 1–16 (1997).

[127] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, "Reconstruction of Electrons with the Gaussian Sum Filter in the CMS Tracker at LHC", ECONF C0303241, TULT009 (2003).

[128] R. Salerno, "Electron Reconstruction and Identification in CMS at LHC", Nucl. Phys. Proc. Suppl. 197, 230–232 (2009).

[129] CMS Collaboration, "Performance of Muon Identification in *pp* Collisions at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary MUO-10-002 (2010).

[130] G. P. Salam and G. Soyez, "A Practical Seedless Infrared-Safe Cone Jet Algorithm", JHEP 05, 086 (2007).

[131] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally Invariant k_{T} Clustering Algorithms for Hadron Hadron Collisions", Nucl. Phys. B406, 187–224 (1993).

[132] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better Jet Clustering Algorithms", JHEP 08, 001 (1997).

[133] M. Wobisch and T. Wengler, "Hadronization Corrections to Jet Cross-Sections in Deep Inelastic Scattering", `arXiv:hep-ph/9907280` (1998).

[134] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k_{T} Jet Clustering Algorithm", JHEP 04, 063 (2008).

[135] CMS Collaboration, "Jet Plus Tracks Algorithm for Calorimeter Jet Energy Corrections in CMS", CMS Physics Analysis Summary JME-09-002 (2009).

[136] DØ Collaboration, "B Jet Identification" `http://www-d0.fnal.gov/Run2Physics/top/singletop_observation/b_tagging_graphic.png` (November 2011).

[137] CMS Collaboration, "Performance of b tagging at $\sqrt{s} = 8$ TeV in multijet, ttbar and boosted topology events", CMS Physics Analysis Summary BTV-13-001 (2013).

[138] S. Chatrchyan et al., "Identification of b-Quark Jets with the CMS Experiment", JINST 8, P04013 (2013).

[139] J. M. Campbell, J. W. Huston, and W. J. Stirling, "Hard Interactions of Quarks and Gluons: A Primer for LHC Physics", Reports on Progress in Physics 70, 89 (2007).

[140] R. Field, "Early LHC Underlying Event Data - Findings and Surprises", `arXiv:1010.3558` (2010).

[141] S. D. Drell and T.-M. Yan, "Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies", Phys. Rev. Lett. 25, 316–320 (1970).

[142] N. Kidonakis, "Two-Loop Soft Anomalous Dimensions for Single Top Quark Associated Production with a W- or H-", Phys. Rev. D82, 054018 (2010).

[143] N. Kidonakis, "Differential and Total Cross Sections for Top Pair and Single Top Production", `arXiv:1205.3453` (2012).

[144] N. Kidonakis, "Next-to-Next-to-Leading-Order Collinear and Soft Gluon Corrections for t-Channel Single Top Quark Production", Phys. Rev. D83, 091503 (2011).

[145] K. Melnikov and F. Petriello, "Electroweak Gauge Boson Production at Hadron Colliders through O(α_S^2)", Phys. Rev. D74, 114017 (2006).

[146] L. Tuura, A. Meyer, I. Segoni, and G. Della Ricca, "CMS Data Quality Monitoring: Systems and Experiences", J. Phys. Conf. Ser. 219, 072020 (2010).

[147] CMS Collaboration, "MET Optional Filters" `https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFilters` (internal), (July 2013).

[148] CMS Collaboration, "Pileup Studies" `https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupInformation` (internal), (July 2013).

[149] CMS Collaboration, "EPS13 Prescription" `https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagPOG` (internal), (July 2013).

[150] CMS Collaboration, "Measurement of Differential Top-Quark Pair Production Cross Sections in the Lepton+Jets Channel in pp Collisions at $\sqrt{s} = 8$ TeV", CMS Physics Analysis Summary TOP-12-027 (2013).

[151] CMS Collaboration, "p_{t}(Top-Quark) based Reweighting of $\mathrm{t\bar{t}}$ MC" https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopPtReweighting (internal), (July 2013).

[152] C. Autermann, C. Sander, P. Schleper, M. Schroder, and H. Stadie, "Measurement of the Full Jet p_T Response Function in QCD Di-Jet Events", CMS Analysis Note AN 2011/330 (2011).

[153] CMS Collaboration, "MVA Based Electron Identification" https://twiki.cern.ch/twiki/bin/view/CMS/MultivariateElectronIdentification (internal), (July 2013).

[154] CMS Collaboration, "Conversion Rejection" https://twiki.cern.ch/twiki/bin/viewauth/CMS/ConversionTools (internal), (July 2013).

[155] J. Ott, "Theta – A Framework for Template-Based Modeling and Inference" http://www.theta-framework.org.

[156] T. Chwalek, "Measurement of the W-Boson Helicity-Fractions in Top-Quark Decays with the CDF II Experiment and Prospects for an Early $t\bar{t}$ Cross-Section Measurement with the CMS Experiment", PhD Thesis, Karlsruhe Institute of Technology, CERN-THESIS-2010-255 (2010).

[157] V. Blobel, "An Unfolding Method for High Energy Physics Experiments", arXiv:hep-ex/0208022 (2002).

[158] A. Tikhonov, "Solution of Incorrectly Formulated Problems and the Regularization Method", Soviet Mathematics Doklady 4, 1035–1038 (1963).

[159] D. L. Phillips, "A Technique for the Numerical Solution of Certain Integral Equations of the First Kind", J. ACM 9, 84–97 (1962).

[160] S. Baker and R. D. Cousins, "Clarification of the Use of χ^2 and Likelihood Functions in Fits to Histograms", Nucl. Instrum. Meth. 221, 437–442 (1984).

[161] V. Blobel, "Unfolding – Linear Inverse Problems –", Notes for the Terrascale Workshop on Unfolding and Data Correction at DESY (2010).

[162] J. Goh, Y. Jo, S. Khalil, and T. J. Kim, "Electron Efficiency Measurement for Top Quark Physics at $\sqrt{s} = 8$ TeV" CMS Note AN-12-429 (2013).

[163] CMS Collaboration, "Reference muon id and isolation efficiencies" https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs (internal) (July 2013).

[164] W. Bernreuther and Z. Si, Private Communication (2013).

[165] E. Gabrielli, M. Raidal, and A. Racioppi, "Implications of the Effective Axial-Vector Coupling of Gluon on Top-Quark Charge Asymmetry at the LHC", Phys.Rev. D85, 074021 (2012).

[166] E. Gabrielli and M. Raidal, "Effective Axial-Vector Coupling of Gluon as an Explanation to the Top Quark Asymmetry", Phys. Rev. D 84, 054017 (2011).

[167] E. Gabrielli, Private Communication (2013).

[168] S. Berge and S. Westhoff, "Top-Quark Charge Asymmetry Goes Forward: Two New Observables for Hadron Colliders", JHEP 1307, 179 (2013).

[169] V. Sharyy, "Experimental Status of Top Charge Asymmetry Measurements", `arXiv:1312.0383` (2013).

[170] T. Peiffer, "Q^2 Systematic MC Samples Re-Weighting", CMS Internal Talk (February 2012).

A. Q^2 Reweighting

The choice of the factorization and renormalization scale Q^2 has an non-negligible influence on the measured result. To quantify this, the measurement is repeated using dedicated samples, which are generated with shifted values of Q^2. But the statistical power of these samples is insufficient in the selected phase space. This results in the statistical uncertainties of the shifted samples being of the same magnitude than the deviations to the nominal measurement.

Therefore the Q^2 reweighting method [170] has been used in previous analyses. For this method, the Q^2 scale for events of the nominal $\mathrm{t\bar{t}}$ sample has been shifted relatively by factors of 2 and 0.5 for the up and down variations. Then α_s and the PDFs have been recalculated at this scale and each event is reweighted accordingly.

A comparison of these reweighted samples with the nominal sample and the dedicated samples generated with shifted Q^2 is shown in figure A.1.

There large deviations can be seen between the dedicated systematic samples and the reweighted nominal samples. Also the asymmetry measurement with the reweighted nominal samples show no deviations from the non-shifted nominal measurement. Therefore it has been decided not to use this approach for this analysis.

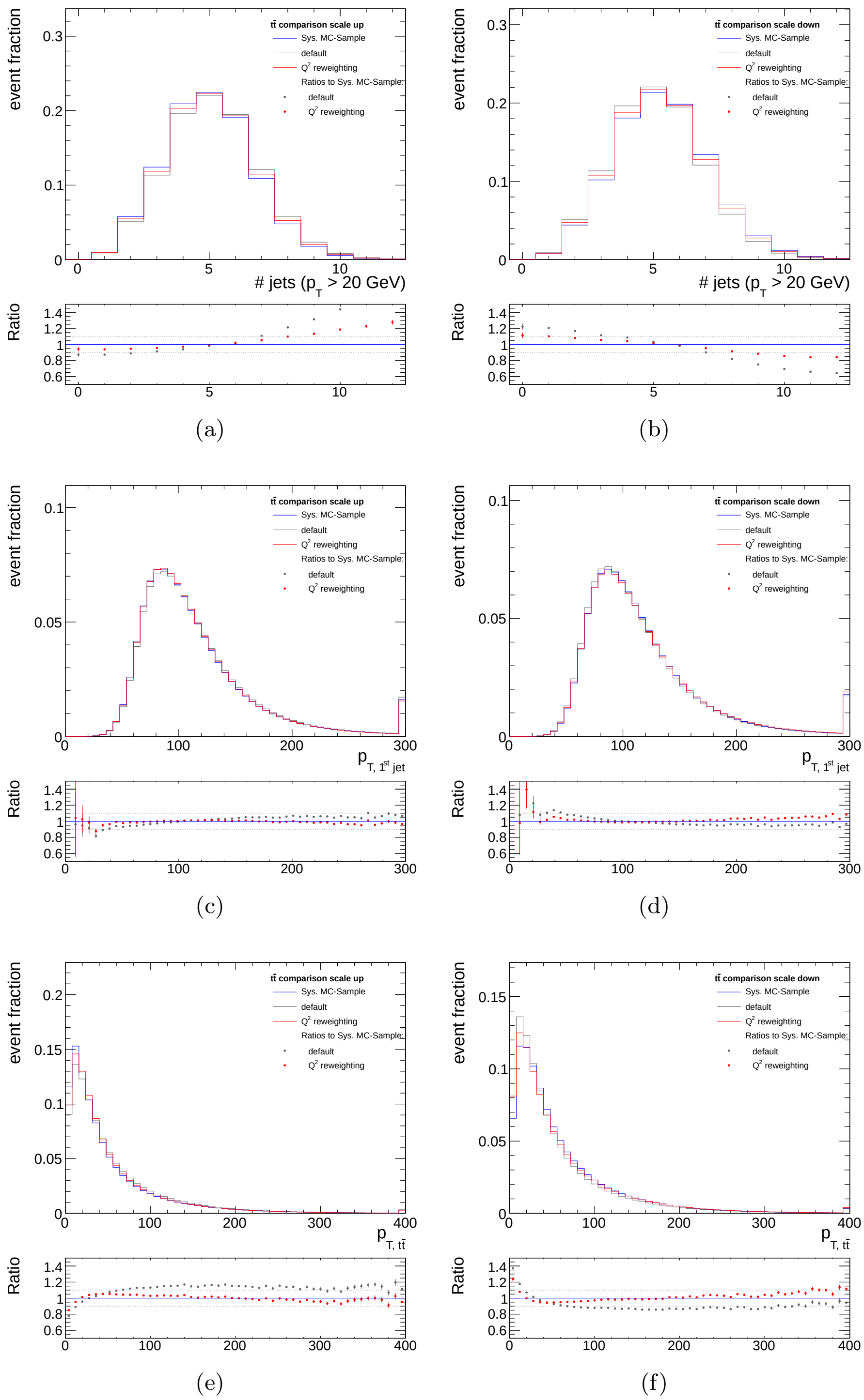


Figure A.1.: Comparison of the nominal $t\bar{t}$ sample (gray), the dedicated samples generated with shifted values of Q^2 (blue), and the reweighted samples with Q^2 shifted by hand (red). The up variations are shown on the left, the down variations on the right. The values of the number of jets with p_{T} larger than 20 GeV/c ((a) and (b)), the transverse momentum of the leading jet ((c) and (d)), and the transverse momentum of the top-quark pair ((e) and (f)) are shown. Additionally the ratios of the distributions to the dedicated systematic samples are depicted.

Danksagung

Zunächst möchte ich mich bei Herrn Prof. Dr. Th. Müller bedanken, dass er mich in seine Top-Physik-Arbeitsgruppe aufgenommen hat und mir die Möglichkeit gegeben hat, an einem so spannenden Thema zu forschen.

Herrn Prof. Dr. G. Quast danke ich für die Übernahme des Korreferats.

Ein großer Dank gilt meinem Betreuer Dr. Thorsten Chwalek für die ausdauernde Betreuung und für die großartige wissenschaftliche, moralische und organisatorische Unterstützung, sowie Dr. Jeannine Wagner-Kuhr für die Motivation und Begeisterung für die Top-Physik.

Ein besonderer Dank gilt Frank Roscher für seine exzellente Unterstützung und der großen Mitarbeit an dieser Analyse, sowie für das Beantworten all meiner aufkommenden Fragen, vor allem auch in technischer Hinsicht. Ihm und Dr. Thorsten Chwalek danke ich auch für das intensive Korrekturlesen dieser Arbeit.

Dr. Matthias Mozer, Christian Böser, Nils Faltermann, Simon Fink, Wajid Ali Khan, Benedikt Maier, Steffen Röcker und Ivan Shetsov danke ich für die hervorragende Atmosphäre in unserer Arbeitsgruppe. Auch bei den Mitgliedern der Arbeitsgruppe von Prof. Dr. M. Feindt bedanke ich mich für die vielen interessanten und unterhaltsamen Gespräche in den Mittagspausen.

Ich danke allen Mitgliedern des EKP, insbesondere dem Admin-Team und Frau Bräunling, für das Ermöglichen eines reibungslosen Arbeitens am Institut.

Meinen Eltern danke ich für die moralische und finanzielle Unterstützung während meines gesamten Studiums, ohne die ein erfolgreicher Abschluss nicht möglich gewesen wäre. Genauso bedanke ich mich bei meinen Geschwistern, meiner Familie und meinen Freunden für deren steten Rückhalt.

Der größte Dank gebührt meiner zukünftigen Ehefrau Corina Friedrich für die Motivation und die Kraft, die sie mir gibt.

Hiermit versichere ich, dass ich diese Arbeit selbstständig verfasst habe, dass ich keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass ich die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht habe, sowie dass ich die Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis beachtet habe.

Christian Buntin

Karlsruhe, Mai 2014