Momentum Bias Determination in the Tracker Alignment and First Differential $t\bar{t}$ Cross Section Measurement at CMS

Dissertation

zur Erlangung des Doktorgrades

des Department Physik

der Universität Hamburg

vorgelegt von

Holger Enderle

aus Lübeck

Hamburg

2012

Gutachter der Dissertation:	Prof. Dr. Peter Schleper Prof. Dr. Johannes Haller
Gutachter der Disputation:	Prof. Dr. Dieter Horns PD Dr. Thomas Schörner-Sadenius
Datum der Disputation:	09.01.2012
Vorsitzender des Prüfungsausschusses:	PD Dr. Michael Martins
Vorsitzender des Promotionsausschusses:	Prof. Dr. Peter Hauschildt
Leiterin des Fachbereichs Physik:	Prof. Dr. Daniela Pfannkuche
Dekan der Fakultät für Mathematik, Informatik und Naturwissenschaften:	Prof. Dr. Heinrich Graener

Abstract

This thesis is prepared within the framework of the CMS experiment at the Large Hadron Collider. It is divided into a technical topic and an analysis.

In the technical part, a method is developed to validate the alignment of the tracker geometry concerning biases in the momentum measurement. The method is based on the comparison of the measured momentum of isolated tracks and the corresponding energy deposited in the calorimeter. Comparing positively and negatively charged hadrons, the twist of the tracker is constrained with a precision of $\frac{\Delta\phi}{\Delta z} = 12\,\frac{\mu\mathrm{rad}}{\mathrm{m}}$.

The analysis deals with cross section measurements in events containing an isolated muon and jets. The complete dataset of proton-proton collisions at a centre-of-mass energy of 7 TeV taken in 2010 is investigated. This corresponds to an integrated luminosity of $35.9\,\mathrm{pb}^{-1}$. Cross sections including different physics processes with an isolated muon and jets in the final state are measured for different jet multiplicities ($N_{\mathrm{jets}} \geq 1, 2, 3, 4$). With increasing jet multiplicity, the transition from a $W \to l\nu$ dominated to a strongly $t\bar{t}$ enriched phase space becomes evident. The inclusive cross section for $t\bar{t}$ production derived from the four jet sample is measured to be $\sigma = 172 \pm 15(\mathrm{stat.}) \pm 41(\mathrm{syst.}) \pm 7(\mathrm{lumi.})\,\mathrm{pb}$. Cross sections differentially in kinematic quantities of the muon, $\frac{d\sigma}{dp_T}$, $\frac{d\sigma}{d\eta}$, are measured as well and compared to theoretical predictions.

Kurzfassung

Diese Arbeit wurde im Rahmen des CMS-Experiments am Large Hadron Collider erstellt. Sie gliedert sich in einen technischen Abschnitt und eine Analyse.

Im technischen Teil wird eine Methode entwickelt, um das Alignment der Spurdetektorgeometrie bezüglich der Verfälschung von Impulsmessungen zu validieren. Die Methode basiert auf dem Vergleich zwischen gemessenem Impuls isolierter Spuren und der zugehörigen im Kalorimeter deponierten Energie. Durch den Vergleich von positiv und negativ geladenen Hadronen können Verdrillungen des Spurdetektors mit einer Präzision von $\frac{\Delta\phi}{\Delta z} = 12\,\frac{\mu\mathrm{rad}}{\mathrm{m}}$ ermittelt werden.

Die Analyse behandelt Wirkungsquerschnittsmessungen in Ereignissen, die ein isoliertes Myon und Jets enthalten. Hierbei wurde der gesamte im Jahr 2010 aufgezeichnete Datensatz von Protonenkollisionen bei einer Schwerpunktsenergie von 7 TeV verwendet. Er entspricht einer integrierten Luminosität von $35.9\,\mathrm{pb}^{-1}$. Es wurden für verschiedene physikalische Prozesse mit einem isolierten Myon und Jets im Endzustand kombinierte Wirkungsquerschnitte für verschiedene Jetmultiplizitäten ($N_{\mathrm{Jets}} \geq 1, 2, 3, 4$) untersucht. Mit zunehmender Jetmultiplizität zeigt sich ein Übergang von einem Phasenraum, der durch $W \to l\nu$ Ereignisse dominiert ist, hin zu einem Phasenraum, der stark mit $t\bar{t}$ Produktion angereichert ist. Aus den Ereignissen mit mindestens vier Jets wurde der inklusive Wirkungsquerschnitt der $t\bar{t}$ Produktion zu $\sigma = 172 \pm 15(\mathrm{stat.}) \pm 41(\mathrm{syst.}) \pm 7(\mathrm{lumi.})\,\mathrm{pb}$ bestimmt. Die Messung wurde ebenfalls differentiell in kinematischen Größen des Myons, $\frac{d\sigma}{dp_T}$, $\frac{d\sigma}{d\eta}$, durchgeführt und mit theoretischen Vorhersagen verglichen.

Contents

List of Figures

List of Tables

1 Introduction

The aim of particle physics is to identify the basic building blocks the universe is made of and to understand their interactions.

According to our current conception of nature, matter consists of twelve elementary matter particles, namely six leptons and six quarks and their corresponding antiparticles. They are all sorted into three families each consisting of one neutral and one charged lepton and one up- and one down-type quark plus corresponding antiparticles.

Four fundamental forces act between these particles: gravitational interactions between all known particles, weak interactions between all quarks and leptons, electromagnetic interactions between all electrically charged particles, and strong interactions between all colour-charged particles. In particle physics, the so-called 'Standard Model' (SM) [1] sets up the theoretical and mathematical framework comprising the present knowledge on the particles and their interactions, except for gravity, which is decoupled from the SM and described by the general theory of relativity [2].

The 'SM forces' are explained by gauge theories and thus by the exchange of gauge bosons. These are the Z boson and the two W bosons being responsible for weak interactions, the photon for electromagnetism and the gluons, existing in eight different colour states, for strong interactions. Weak interactions and electromagnetism are unified in the theory of electroweak interactions [3, 4], whereas the strong interactions are described within the theory of Quantum Chromodynamics (QCD) [5].

The existence of all these particles has been verified by different experiments in the past. However, the Higgs boson [6], an essential building block of the SM, has not been observed yet [7, 8]. It is supposed to cause the masses of all fundamental particles. A compilation of the observed SM particles and their masses[1] is given in table 1.1.

The masses of the fundamental particles range over many orders of magnitude. In principle, all these particles can be produced at particle colliders, given that the centre-of-mass energy is high enough to generate the particle mass. Therefore, the heaviest particle, the top quark, was only discovered about 20 years after its prediction [9] within the framework of the SM. It was finally discovered in 1995 at the proton-antiproton collider Tevatron being operated at a centre-of-mass energy of nearly 2 TeV [10, 11].

In the following, the focus will be on this heaviest particle of the SM. A physical consequence of the large top-quark mass exceeding the sum of the masses of the W boson and the bottom (b) quark is the rapid decay of the top quark. It decays after a very short mean lifetime of the order of 10^{-25} s via the weak interaction into a W boson and a bottom quark.

[1] Throughout the thesis, the 'natural units' of particle physics are used, i. e. $\hbar = c = 1$. In this notation, it is easier to compare energy, momentum, and mass as they are all given in the same unit 'eV'.

		family			gauge bosons		
		1	2	3	mass [MeV]		
leptons	neutral (0e)	ν_e	ν_μ	ν_τ			
	mass [MeV]	$< 2{\cdot}10^{-6}$	< 0.19	< 18.2			
	charged (-1e)	e	μ	τ		photon	Z boson
	mass [MeV]	0.511	105.66	1776.8		$<10^{-24}$	$91188^{+2.1}_{-2.1}$
quarks	up type ($+\frac{2}{3}$e)	u	c	t	gluon		W boson
	mass [MeV]	$2.5^{+0.6}_{-0.8}$	1290^{+50}_{-110}	173200^{+900}_{-900}			80399^{+23}_{-23}
	down type ($-\frac{1}{3}$e)	d	s	b	0		
	mass [MeV]	$5.0^{+0.7}_{-0.9}$	100^{+30}_{-20}	4190^{+180}_{-60}			

Table 1.1: Standard Model particles and their measured masses [12, 13]. The given upper limits for neutrino masses are determined from direct measurements. The stated mass of the gluon is a theoretical value. The electric charge of the different types of leptons and quarks is given in elementary charges. The corresponding antiparticles not being listed here have the same mass but opposite charge. The braces indicate to which particles the gauge bosons couple.

Due to the structure of QCD, the quarks being colour charged are confined to colourless bound states called hadrons. Generally, a bunch of hadrons emerges from each high-momentum quark produced in a collision. This bunch of hadrons is measured as a so-called 'jet' in the detector and only allows indirect measurements of the quark properties. The timescale of hadronisation is of the order of 10^{-23} s. Therefore, the top quark already decays before it could hadronise. This provides the unique possibility to measure properties of the bare quark leading for instance to a very low relative uncertainty in the mass measurement of below 1 % at the Tevatron [13].

Since March 2010, the Large Hadron Collider provides proton-proton collisions at a centre-of-mass energy of 7 TeV being the second collider in the world producing top quarks. The production cross section for top quark pairs at this centre-of-mass energy is more than a factor 20 higher than at the Tevatron. Due to the large cross section and a high collision rate, the Large Hadron Collider already produced more top quarks during its first two years of operation than the Tevatron during its complete operation period, which lasted for about 25 years and ended in September 2011.

This high rate of top-quark production will soon lead to an amount of statistics that particularly enables differential measurements of kinematic top quark quantities with unprecedented precision. As many models for physics beyond the SM differ from the SM in the prediction of these distributions, an accurate test of the SM in this sector serves as a promising search for new phenomena. Furthermore, due to the similar event topologies, the top quark production constitutes one of the most important background processes for other search methods.

In this thesis, the very first data of proton-proton collisions at a centre-of-mass energy of 7 TeV taken in 2010 by the Compact Muon Solenoid detector at the Large Hadron Collider was used. This corresponds to an integrated luminosity of $35.9\,\mathrm{pb}^{-1}$. At this early stage of the experiment, the understanding of the detector plays an important role for a proper object reconstruction to ensure robust measurements. Muons are proved to be reconstructed and identified reliably already from the very beginning of the data taking [14].

A vital contribution to the reliability of the muon reconstruction is made by the track-based alignment of the inner tracking system, i. e. the determination of the tracker geometry. A very high alignment precision has been achieved in 2010 using the tracks of secondary muons from cosmic rays as well as tracks originating from the collisions [15]. However, there are certain correlated tracker distortions that the track-based alignment is insensitive to. Therefore, in this thesis, a method is developed to validate tracker geometries concerning biases in the momentum measurement. It is based on a comparison between the energy measurement in the calorimeter and the corresponding momentum measurement in the inner tracking system.

Taking advantage of the accurate muon reconstruction of the CMS detector, events containing an isolated muon in the final state are investigated in the analysis part of this thesis. As the first step towards a measurement of the top-quark pair production cross section, a measurement of a combined muon+jets cross section is established. This combination basically comprises all main physics processes resulting in a final state containing a muon originating from a W boson and additional hard jets. Performing this measurement as a function of the jet multiplicity allows for the study of the transition from background dominated towards top-quark dominated kinematic regions and helps to estimate the main background processes for the top quark pair production. The top quark pair production cross section measurement is performed inclusively and also differentially in kinematic quantities of the muon to test the underlying kinematics.

The structure of this thesis is as follows. An overview of the experimental framework needed for this thesis is given in chapter 2. After a short introduction to track reconstruction and track-based alignment, the momentum bias study is explained in chapter 3. Phenomenological aspects, which are important for the top-quark analysis, are discussed in chapter 4. Chapter 5 briefly presents the event simulation based on Monte-Carlo generators required for the analysis. Chapter 6 defines the selection of a sample enriched with events containing isolated muons and jets. The derivation of the muon+jets and $t\bar{t}$ cross sections including background estimation and the calculation of uncertainties is described in chapter 7 and results are shown in chapter 8. Finally, chapter 9 summarises the thesis and and gives an outlook on possible extensions of the analysis.

2 Experimental Setup

The European Organization for Nuclear Research CERN[1] was founded in 1954 on the Swiss-French border near Geneva. According to the name, the first experiments were dedicated to the study of atomic nuclei. But soon the main focus moved to particle physics and thus larger and larger accelerators providing higher and higher energies were built, see figure 2.1. For now, this evolution of energy culminated in the construction of the Large Hadron Collider (LHC) designed to collide proton beams with a centre-of-mass energy of 14 TeV [16]. There are four collision points along the beamline each equipped with a detector. One of these detectors is the Compact Muon Solenoid (CMS) [17, 18], which provides the data used in this thesis.

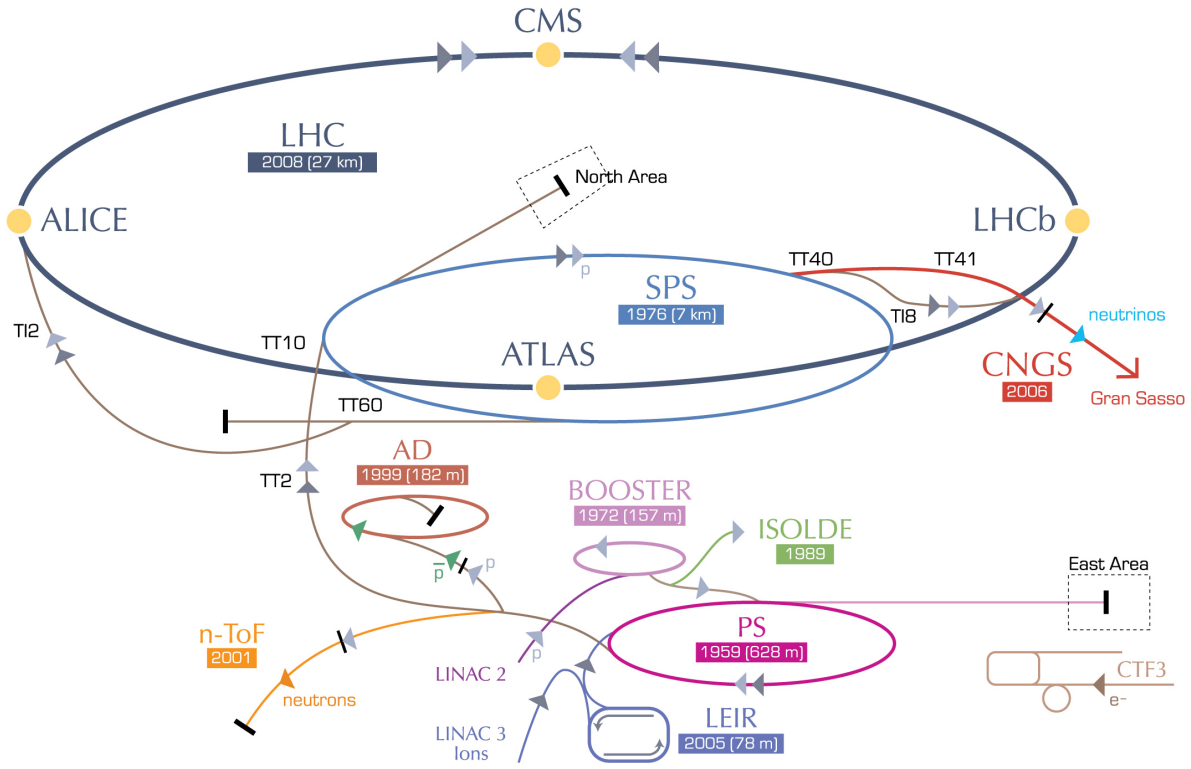

Figure 2.1: Sketch of the CERN accelerator complex. Below each accelerator name, the year of completion and, for ring accelerators, the circumference are given. [19]

[1]The acronym CERN educes from the French *Conseil Européen pour la Recherche Nucléaire* and was not changed when it was renamed to *Organisation Européenne pour la Recherche Nucléaire*.

2.1 Large Hadron Collider

The LHC is a collider ring with a circumference of 26.7 km located in a tunnel which lies between 45 m and 170 m below the surface. The main purpose is to collide protons counter-rotating in two beam pipes. The additional ability for colliding heavy ions such as lead will not be discussed here. The protons are ramped up by several pre-accelerators to an energy of 450 GeV before they are injected in bunches into the LHC. The first collisions with stable proton beams were achieved in November 2009 at injection energy. Then the LHC successively accelerated the protons and, since March 2010, provides collisions at a centre-of-mass energy of 7 TeV. After the experiments will have taken an adequate amount of data at this energy, which is of the order of 10 fb^{-1}, the beam energies will be raised until the design centre-of-mass energy of 14 TeV is reached. An important parameter, which characterises the performance of the accelerator, is the luminosity $\mathcal{L}$. It is a measure for the rate of particle collisions and directly relates the cross section σ of a certain process to the event rate $\mathcal{N}$ associated with that process:

$$\mathcal{N} = \sigma \cdot \mathcal{L}. \tag{2.1}$$

In order to obtain a statistically relevant number of events of rare processes a high luminosity is needed. The LHC is aimed at reaching up to $\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. There are

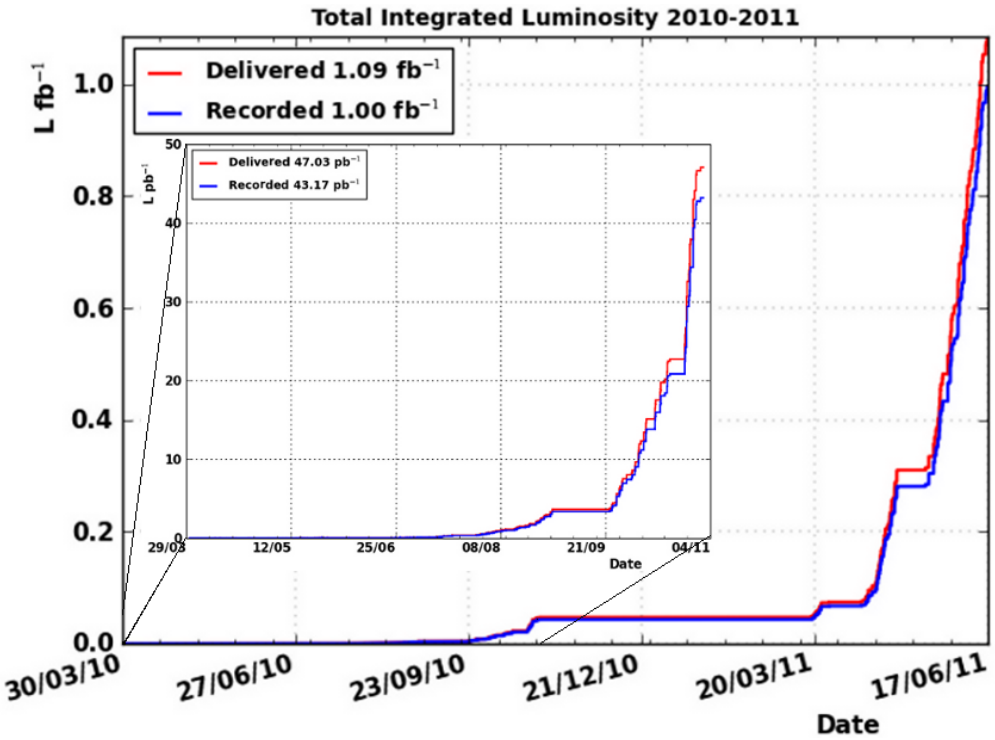

Figure 2.2: Total integrated luminosity in 2010 and 2011 as a function of time, delivered by LHC (red line) and recorded by CMS (blue line). The inlay is a zoom into 2010 data.

several beam parameters which can be tuned to increase the luminosity. The functional relation is

$$\mathcal{L} = \frac{N_p^2 N_b f}{A}. \tag{2.2}$$

N_p is the number of protons per bunch, N_b is the number of colliding bunches, f is the revolution frequency, and A is the beam profile, which depends on the bunch size and the crossing angle. The revolution frequency is about 11 kHz assuming protons travelling almost at speed of light, the other parameters were modified during the 2010/2011 data taking. The resulting increase in luminosity can be seen from the rise in the integrated luminosity $L = \int \mathcal{L}(t)dt$ as a function of time shown in figure 2.2. In 2010, most of the integrated luminosity was collected in the last few weeks before the winter shut down. In 2011, the integrated luminosity of the first three month of running exceeded the complete 2010 luminosity by almost a factor of 30.

2.2 Compact Muon Solenoid

CMS is a barrel-shaped multipurpose detector with a total weight of 12500 t. As the first part of the name implies, the architecture of CMS is very compact leading to a length of 21.6 m and a diameter of 14.6 m, covering almost $4\,\pi$ of the solid angle. It is situated in a cavern 100 m underground between Lake Geneva and the Jura mountains. The geometry of the detector is defined in a right-handed coordinate system with its

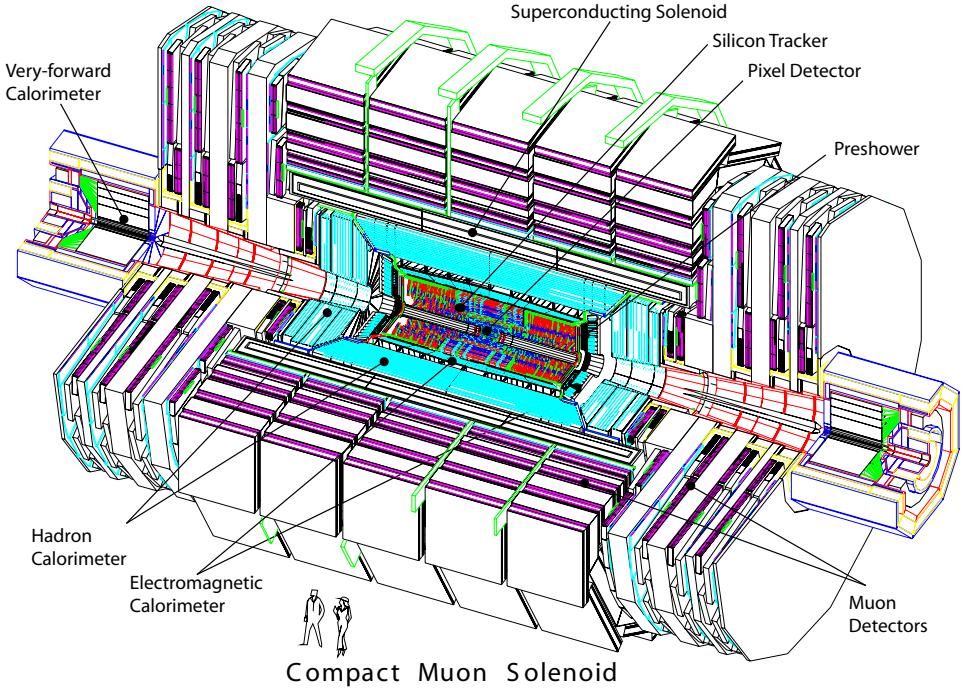

Figure 2.3: A perspective view of the CMS detector showing its main components. [18]

origin in the nominal interaction point. The positive z-direction is oriented along the beam direction towards the Jura mountains and the x axis points towards the centre of the LHC. In the polar coordinate system (r, θ, ϕ), the azimuthal angle ϕ is defined with respect to the x axis in the x-y plane and the polar angle θ is defined with respect to the z axis. In particle physics, a commonly used alternative to the polar angle is the pseudorapidity $\eta = -\ln(\tan\frac{\theta}{2})$ mostly because the particle flux is nearly constant in η. The CMS detector with its main components is shown in figure 2.3. The general concept is driven by the choice of the magnetic field configuration, which is needed to bend the tracks of charged particles enabling a momentum measurement. A 3.8 T magnetic field is provided by a superconducting solenoid giving CMS the last part of its name. With the length of 13 m and an inner diameter of 6 m it is large enough to accommodate the inner tracking system and the main calorimeter parts. The large muon system, causing the experiment's middle name, is embedded in 1.5 m of iron around the solenoid divided into five wheels, which are serving as a return yoke for the magnetic flux.

2.2.1 Inner Tracking System

A schematic layout of the inner tracking system can be seen in figure 2.4. The tracker has to cope with very high charged particle fluxes and to deliver precise measurements of particle trajectories at the same time. For this reason, the tracker is completely based on silicon detector technology. Due to the reduction of the particle flux with increasing radius, the tracker is divided into different regions from the inside to the outside.

- **Pixel:** The innermost part is the pixel detector. It covers radii from 4.4 to 10.2 cm where the charged particle flux is the highest. It will reach up to 10^7-$10^8\,\mathrm{cm^{-2}s^{-1}}$ at design luminosity. The pixel cell size is $100 \times 150\,\mu\mathrm{m}^2$ leading to a pixel occupancy of the order of 10^{-4} per LHC bunch crossing. It consists of three barrel layers with 768 modules and two endcap disks with 672 modules. The layers and disks are arranged in a way that the pixel detector provides three tracking points over almost the full η range ($|\eta| < 2.5$). A spatial resolution in the range of 15-20 μm can be achieved in $r\phi$ and z. Hence, a good secondary vertex reconstruction is possible with the help of the pixel detector.

- **Strip tracker:** Radii from 20 to 116 cm are covered by the strip tracker. It is composed of 15148 strip modules and is divided into further parts.

  The Tracker Inner Barrel and Disks (TIB/TID) cover radii up to 55 cm. The particle flux in this region is reduced by one order of magnitude with respect to the pixel detector allowing the use of micro-strip modules with a minimum cell size of $10\,\mathrm{cm} \times 80\,\mu\mathrm{m}$. This results in a strip occupancy of up to 2-3 % per LHC bunch crossing. The TIB comprises 4 layers and is enclosed by the three disks of the TID at both ends. The modules of the TID are arranged radially and grouped in rings. The two inner layers of the TIB and the two inner rings of the TID have double modules mounted back-to-back. The additional module is oriented with a stereo angle of 100 mrad with respect to the other module providing a measurement of a

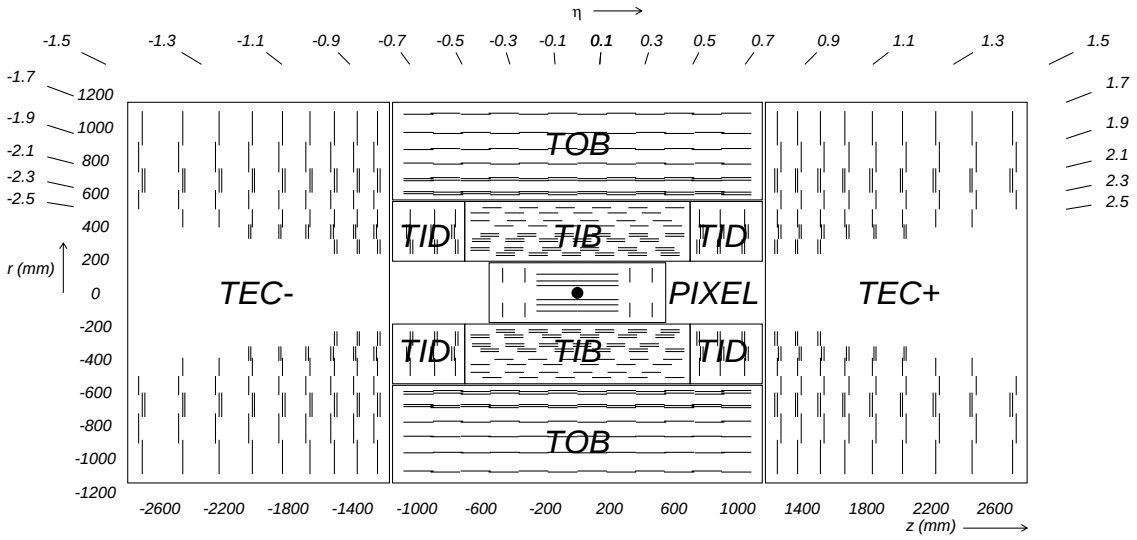

Figure 2.4: Layout of the inner tracking system in the $r-z$ view. The lines indicate the positions of the tracker modules. [18]

second coordinate in addition to $r\phi$, z for the TIB and r for the TID. The single point resolution in the TIB ranges from 23 to 34 μm in $r\phi$ and is about 230 μm in z.

The Tracker Outer Barrel and EndCaps (TOB/TEC) surround TIB and TID. Due to a further reduction of the particle flux by one order of magnitude, the strip size can be increased to a maximum of 25 cm $\times$ 180 μm in order to limit the number of read-out channels while keeping the strip occupancy at about 1 % per LHC bunch crossing. The TOB consists of six layers, out of which the two inner are equipped with double modules. Nine disks enclosing TID and TOB at each end form the TEC. Here, the rings 1, 2 and 5 are equipped with additional stereo modules. The TOB provides a single point resolution of 35 to 52 μm in $r\phi$ and 530 μm in z.

2.2.2 Calorimeter

The calorimeter consists of the electromagnetic calorimeter (ECAL) and the surrounding hadronic calorimeter (HCAL). The design of the calorimeter is strongly influenced by the environment in CMS. There is not much space as the main parts are placed inside the solenoid, and in addition, the calorimeter has to manage high particle fluxes in short time intervals.

- **ECAL:** The material chosen for the ECAL is lead tungstate ($PbWO_4$). Crystals made from $PbWO_4$ have a short radiation length[2] of $X_0 = 0.89$ cm enabling them to

[2]The radiation length is the average distance that an electron travels through the material until its energy is decreased to a fraction of $\frac{1}{e}$ due to bremsstrahlung.

meet the requirement of the available space. Moreover, they emit 80% of the light within 25 ns, which is the minimal bunch spacing of LHC, and they are radiation hard resisting up to 100kGy. With test beams the energy resolution for single electrons was measured to be $\sigma_E/E = 2.8\%/\sqrt{E[\mathrm{GeV}]} \otimes 12.4\%/E[\mathrm{GeV}] \otimes 0.26\%$. The ECAL divides into a barrel section and two endcaps.

The region up to $|\eta| < 1.479$ is covered by the ECAL barrel (EB). Here, the calorimeter has its finest granularity with a crystal size of 0.0174×0.0174 in (η,ϕ), which corresponds to a front face cross section of roughly $22 \times 22\,\mathrm{mm}^2$. In length the crystals cover $25.8\,X_0$. They are mounted in a quasi-projective way, which means that they are slightly tilted with respect to the projection from the nominal interaction point. This reduces the number of particles passing through a crack between the crystals.

The ECAL endcaps (EE) add to the barrel at both sides extending the pseudorapidity coverage to $|\eta| < 3.0$. The crystals have a size of $28.6 \times 28.6\,\mathrm{mm}^2$ at their front face and cover $24.7\,X_0$ in length. The EE is supplemented by a preshower detector in front of the crystals, which is a sampling calorimeter composed of two layers of lead absorber each followed by silicon strip sensors. The main purpose of this preshower device is the discrimination between a single photon and a π^0 decaying into two closely spaced photons.

- **HCAL:** The HCAL had to be built in a way that the number of interaction lengths inside the solenoid is maximised. Thus, the choice was made in favour of a sampling calorimeter using brass as a non-magnetic absorber material with a reasonably short interaction length[3]. Thin plastic scintillator layers are used as active material. The HCAL falls into four subsystems, the hadronic barrel (HB), endcap (HE), outer (HO), and forward (HF) calorimeters.

  The central region is instrumented inside the solenoid with the HB extending to a pseudorapidity of $|\eta| < 1.392$ and outside the solenoid the region $|\eta| < 1.3$ is additionally covered by the HO, which ensures an adequate sampling depth. HB and HO are arranged in towers pointing in the direction of the nominal interaction point. The front faces of these towers have a size of 0.087×0.087 in (η,ϕ), which exactly corresponds to a cluster of 5×5 ECAL crystals. The length of the towers in interaction lengths is about $5.4\,\lambda_I$ in transverse direction, which adds up to about $10.3\,\lambda_I$ in the most forward tower due to its angle. The added HO extends the depth to a minimum of $11.8\,\lambda_I$ across its whole η range. The HE directly follows the HO in η and ranges up to $|\eta| < 3.0$. The HE towers have cross sections varying from 0.087×0.087 at low $|\eta|$ to 0.35×0.175 in (η,ϕ) at high $|\eta|$.

  11.2 m from the interaction point in z direction, the HF is placed in order to extend the η coverage further to $|\eta| < 5.2$. At these high $|\eta|$ values the particle flux is much higher than in all other detector regions. Thus, this part of the HCAL primarily needs to be radiation hard. Steel was chosen as absorber material and quartz

[3]The interaction length is the average distance a particle can traverse the material until it interacts.

fibres as active medium. The fibres are bundled to towers with front faces of the size 0.175×0.175 in (η,ϕ). The missing ECAL at this η region is compensated by starting half of the fibres at a depth of 22 cm, where most of the electromagnetic showers are already died off.

The difference in the efficiency of converting electromagnetic energy depositions on the one hand and hadronic energy depositions on the other hand into electrical signals is described by their ratio $\frac{e}{h}$. The different materials which are used for ECAL and HCAL result in a different $\frac{e}{h}$ for the two calorimeter parts. This leads to different particle responses in ECAL and HCAL, which can be improved but not fully recovered by corrections to the energy reconstruction [20].

2.2.3 Muon System

The muon system combines three different types of gaseous detectors, which together are responsible for muon triggering, muon identification, and the improvement of the muon momentum measurement taking advantage of the large lever arm. The construction of the different parts is shown in figure 2.5. The pseudorapidity range $|\eta| < 1.2$ is covered by drift tubes (DT). They are organised into four concentric cylinders, of which the inner three provide a z measurement in addition to the measurement in the r-ϕ plane. The

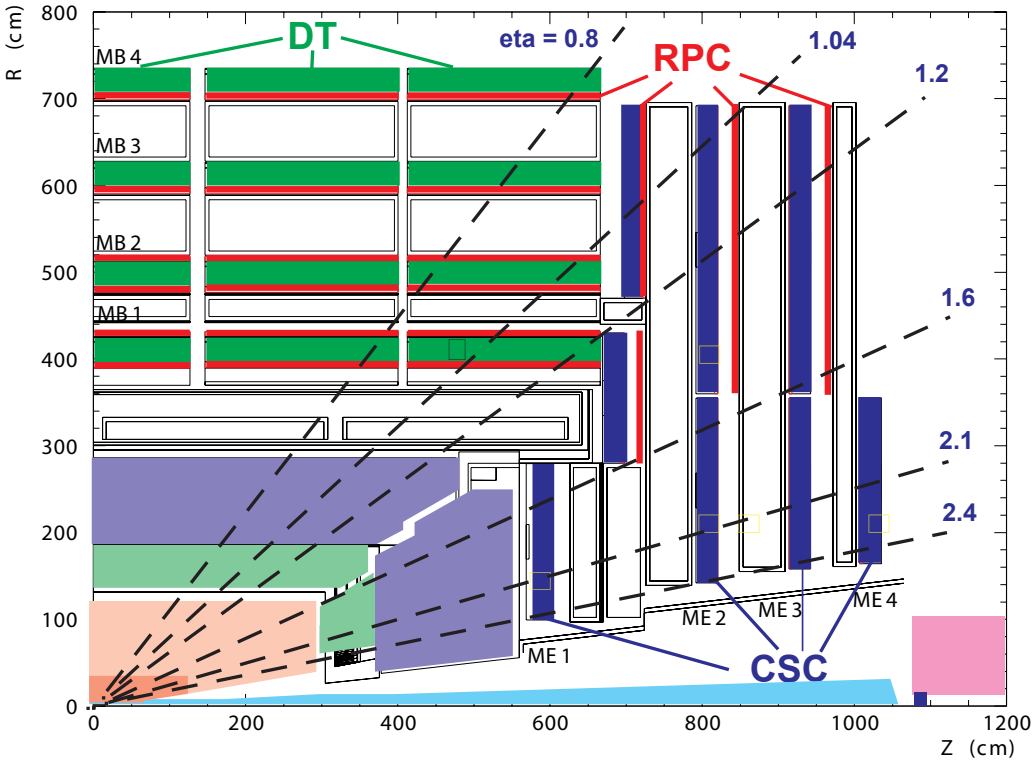

Figure 2.5: Scheme of a quarter of the CMS detector in the $r - z$ view. The different muon chambers are shown and labelled accordingly. [17]

cathode strip chambers (CSC) enclose the DT at both ends and cover $0.9 < |\eta| < 2.4$. The CSC endcaps are arranged in four discs perpendicular to the beam line. DT and parts of CSC ($|\eta| < 1.6$) are interspersed with resistive plate chambers (RPC). Due to the very fast response of the RPC, it provides a good bunch crossing assignment and improves the trigger efficiency.

2.2.4 Trigger System

If the LHC runs at design luminosity, the time between two bunch crossings is 25 ns. That is a frequency of 40 MHz, which is impossible to be managed by the data storage system. At maximum, collision events at a frequency of about 100 Hz can be stored and processed offline. Therefore, the event rate has to be reduced already online with a trigger while preserving a high efficiency for interesting events. The CMS trigger works with a two-step concept. In the first step, the hardware-based level-1 (L1) trigger reduces the rate to about 100 kHz, which, in the second step, can be handled and further decreased to the final event rate of 100 Hz by the software-based high-level trigger (HLT).

- **L1:** A frequency of 40 MHz is an enormous challenge even for a hardware-based trigger. In order to gain in speed, the L1 trigger uses calorimeter and muon system information in a reduced granularity and completely omits information of the inner tracking system. From this information, the presence of simple particle candidates is checked and global energy sums are calculated. Still, the latency of the L1 trigger amounts to 3.2 μs. Thus, the data is stored in pipelined buffers until a L1 trigger decision is made to ensure a quasi-deadtime-free data taking.

- **HLT:** The HLT gets events at a reduced rate and hence can take advantage of all subdetectors in full granularity. The calculations done by the HLT are nearly as complex as the final offline reconstruction. They are performed in parallel by several hundred processors. The thresholds of the trigger decisions have to be adapted to the instantaneous luminosity in order not to exceed the final output rate of about 100 events per second. If for some reason a trigger with a certain threshold should not be abandoned, there is also the possibility of assigning a prescale n to the trigger. This means that only every n-th event which fulfils the requirement fires the trigger.

2.2.5 Luminosity Measurement

For the translation of a measured event rate into a dedicated cross section, the exact knowledge of the luminosity is needed, what can be seen from equation (2.1). The same relation can be used to determine the luminosity provided to an experiment if the rate of a process with a precisely known cross section is measured. At CMS, this is currently done via the total event rate of inelastic proton-proton scattering [21]. There are online methods for a fast feedback and offline methods, which have a latency of about 24 h but provide a better background rejection. Finally, the luminosity is calibrated in order to achieve a reliable absolute normalisation.

- **Online monitoring:** The HF calorimeter is used in two ways for the online luminosity determination, either by zero counting, which means estimating the event rate from the average fraction of empty towers, or by using the linear relation between the average deposited transverse energy per tower and the luminosity. Noise in the HF is treated by subtracting an average pedestal estimated from nominal bunch crossings without any proton bunches. Nominal bunch crossings with only one bunch filled with protons can be used to estimate the effect of beam-gas interactions and beam scraping. The latter occurs when the beam scratches the aperture, most likely a collimator.

- **Offline monitoring:** Either the HF or the inner tracking system is used offline. The HF method requires energy depositions which sum up to $\sum E_T \geq 1\,\mathrm{GeV}$ coinciding within a time frame of 8 ns in forward and backward HF. The tracking method counts events with at least two tracks coming from the same vertex which does not lie too far away from the nominal interaction point longitudinally ($|z| \leq 15\,\mathrm{cm}$). However, in a significant fraction of bunch crossings, more than one interaction happens especially for high instantaneous luminosities ($\mathcal{L} > 10^{29}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$). As the offline methods are only sensitive to the rate of bunch crossings with colliding protons, the average number of collisions per bunch crossing is estimated via zero counting.

- **Absolute calibration:** Initially, a temporary absolute normalisation of the monitoring is estimated by comparing the measurements with Monte Carlo simulations. The outcome can be used as a cross check. But in order to be independent of unconfirmed theoretical inputs to the simulation, a different method is used to obtain the final absolute calibration of the luminosity measurement. Van der Meer scans are performed to determine the beam profile and calculate the luminosity via equation (2.2). The basic concept of these scans is to monitor the relative interaction rate while the transverse beam separation is varied.

3 Tracker Alignment

The tracker is one of the most essential parts of the CMS detector. For this thesis, the most important tasks utilising the tracker are the precise determination of muon momenta and the reconstruction of the tracks of all charged particles for the so-called 'particle flow algorithm', which is described in section 6.6.1. The intrinsic resolution of the tracker modules depends on the part of the tracker where the module is placed. Modules that are close to the interaction point need better resolutions. Thus, the pixel modules are designed to reach almost $10\,\mu$m (see section 2.2.1). In order to be able to utilise this resolution for an accurate reconstruction of track parameters, a proper knowledge of the geometrical position (alignment) of the individual modules is needed. This is done via a χ^2-minimisation of the residuals between hit measurements and track predictions. However, for certain track topologies, there are different solutions that result in almost identical χ^2 values. These solutions are related to systematic shifts of neighbouring modules, which cause a bias in the track parameters. Some of these shifts, e. g. a rotation of different detector layers in the azimuth ϕ relative to each other, affect the measured curvature and therefore the transverse momentum measurement of oppositely charged particles in different directions. As the calorimeter should give a charge independent measurement at high energies, the ratio of calorimeter energy and track momentum can help to discover these weak modes. The aim of this study is therefore to correlate isolated charged particles with calorimeter measurements to constrain the weak modes of the tracker alignment.

3.1 Track Reconstruction

The standard track reconstruction algorithm used at CMS is the combinatorial track finder (CTF) [22]. The track reconstruction can be decomposed into three steps, seed finding, pattern recognition and track fitting. The latter two are performed by the CTF using the capacity of the Kalman Filter [23].

3.1.1 Seed Finding

Three points are needed for the calculation of the starting trajectory parameters. For collision tracks, this seed consists either of a pair of hits in the inner layers with an additional beam spot constraint or of a hit triplet in the inner layers. Preferably, pixel hits are used, but in case of missing hits on pixel layers, the seed can also consist of hits on double-module layers of the strip tracker. The starting trajectory is a helix with a maximum curvature, which translates to a minimum transverse momentum of the track.

For cosmic tracks, some adaptations had to be made. There are also seeds from outer layers but a beam spot constraint can not be used at all in this case. Modifications are needed for the case that the magnetic field is off since then a helix is not appropriate.

3.1.2 Pattern Recognition

The next step is assigning further hits to the track, which is the computationally most intensive part of the reconstruction procedure. Starting from each seed, the following layers are scanned. The width of the search window is defined by the uncertainty of the track parameters. Found hits are consecutively added to the candidate trajectory and each time the track parameters are updated by performing a track fit (see section 3.1.3). In the case of two consecutive layers without any hits which can be assigned to the track, the trajectory is rejected and not propagated any further.

3.1.3 Track Fitting

A Kalman Filter is a progressive Least Squares Method, which increases iteratively the χ^2 of the track based on the difference between the hit under consideration and the predicted trajectory state on this detector surface. It is starting with an initial guess of the track parameters and then adding step by step the hits provided by the pattern recognition. The trajectory state on each surface is calculated as a trade off between the position of the current hit and the state which is predicted by the propagation from the previous surface. Only the trajectory state on the last added surface holds the information of all preceding hits so with each step the precision of the state improves. When all hits are processed, the fit is rerun in the opposite direction and the final trajectory states are given by the combination of both fits in order to benefit from the information of all hits on each surface.

3.2 Track-Based Alignment

For a reliable track reconstruction with best possible resolution, the position of every tracker module has to be known to a precision comparable to the intrinsic resolution of the modules, which is of the order of $10\,\mu$m. Within this section, *position* represents both the location and the orientation of the module. The determination of the module positions has been redone several times over the last years with different samples of input tracks gaining comprehensive knowledge of track-based alignment.

3.2.1 Alignment Algorithm

A very simple one-dimensional example of the position determination with the help of track-based alignment is illustrated in figure 3.1. In this example, a particle traverses the tracker in the arbitrarily chosen $x-z$ plane. One of the hit modules is displaced in x with respect to the design geometry. Before the alignment, the track reconstruction assumes

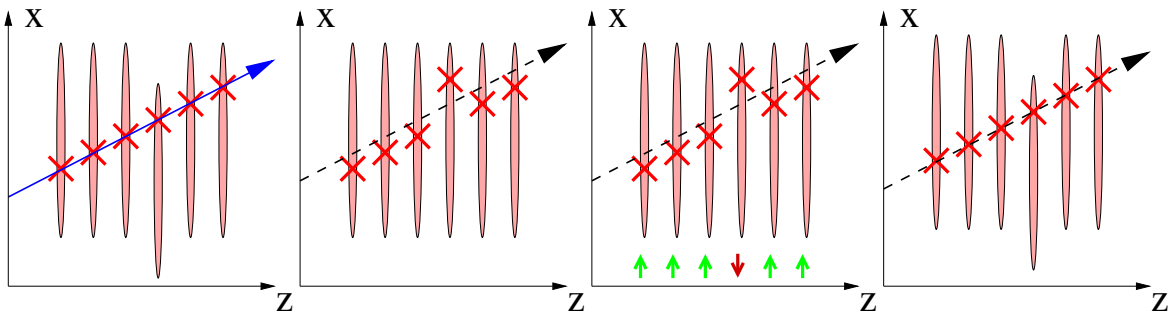

Figure 3.1: Simplified example for the track-based alignment. The blue solid arrow shows the real trajectory of the particle leaving hits indicated by red crosses on the modules shown as salmon ellipses. The dashed black arrow indicates the reconstructed track. The small arrows are shifts suggested by the alignment.

the tracker to be in its design geometry, which results in a worsened track parameter resolution. Based on the distance between the measured hits and the reconstructed track, the alignment suggest a shift of the modules. Repeating the reconstruction of the track with the geometry determined in the alignment procedure improves the obtained track parameter resolution.

In general, algorithms used for aligning the modules minimise the objective function

$$\chi^2(\mathbf{p},\mathbf{q}) = \sum_{j}^{\text{tracks}} \sum_{i}^{\text{hits}} \mathbf{r}_{ij}^T(\mathbf{p},\mathbf{q}_j)\mathbf{V}_{ij}^{-1}\mathbf{r}_{ij}(\mathbf{p},\mathbf{q}_j) \tag{3.1}$$

where $\mathbf{p}$ is the vector of alignment parameters, i. e. corrections to the module positions, and $\mathbf{q}_j$ the vector of track parameters of track j. The track residuals $\mathbf{r}_{ij}$ are defined as the difference between the measured hit position $\mathbf{m}_{ij}$ and the hit position predicted by the trajectory $\mathbf{f}_{ij}(\mathbf{p},\mathbf{q}_j)$. The hit uncertainties are given by $\mathbf{V}_{ij}$, which are the two-dimensional covariance matrices in case of the pixel modules and the squared position uncertainties for the strip modules.

Mainly, two complementary statistical alignment methods have been used for the alignment in the past, a global method and a local method. The "Millepede II" [24] algorithm, which is the global method, is used in the latest alignment approach. It determines the alignment and the track parameters simultaneously in one go by minimising the χ^2 given in equation (3.1). Possible correlations between the alignment parameters of different modules induced by tracks connecting them are thus taken into account. It assumes uncorrelated hits $\mathbf{m}_{ij}$ by using a diagonal covariance matrix $\mathbf{V}_{ij}$. The residuals are linearised in $\mathbf{p}$ and $\mathbf{q}_j$ around the values obtained with the starting geometry. This is valid since changes of the parameters should be small. Larger changes can be handled by iterations of the complete alignment procedure. In order to allow for material effects, such as multiple scattering and energy loss, trajectories based on general broken lines are used [25, 26].

3.2.2 Alignment with Different Datasets

The first alignment of the whole tracker was done using about 3.2 million cosmic tracks, i. e. tracks of charged particles produced by cosmic rays. These tracks were collected in 2008 during a Cosmic Run at Almost Four Tesla (CRAFT). The exact alignment strategy is described in reference [27]. In 2009 and 2010, the CRAFT exercise was repeated with similar strategies. The first collision data, which was taken at reduced centre-of-mass energy in late 2009, was not sufficient for a full tracker alignment. But in 2010, the alignment could be performed combining cosmic tracks and collision tracks [15]. The inclusion of collision tracks especially improves the alignment of the endcaps, which are only weakly illuminated by cosmic tracks. In addition, the combination of two different types of tracks helped to suppress weak modes.

3.3 Systematic Geometry Distortions

There are systematic deviations from the design geometry of the tracker which leave the χ^2 (eq. (3.1)) untouched. Therefore, the χ^2-minimising alignment approach is not able to cure these deviations. Furthermore, they can even be introduced by the alignment procedure itself. So certain constraints have to be used in order to circumvent these distortions.

3.3.1 Trivial Distortions

Trivial examples of these distortions are global translations and rotations around the z axis of the tracker as a whole. The existence of these modes does not affect the internal alignment of the tracker, and as other detector components are aligned with respect to the tracker, it does not harm. The creation of these trivial distortions by the alignment procedure, which can vary in time and thus would be a nuisance to the interalignment of different detector components, can easily be avoided. For instance, global translations are prevented by restricting the centre of gravity of all modules to coincide with the design position.

A tilt of the tracker around the x or y axis does not belong to these kind of distortions as it raises the χ^2 due to the appearance of transverse magnetic field components.

3.3.2 Weak Modes

Non-trivial systematic distortions of the tracker which keep the χ^2 unchanged, however, bias the track parameters and therefore are more worrisome. These not or only weakly χ^2-dependent distortions are called weak modes. In figure 3.2, the principle is illustrated using the twist of the tracker, which is a rotation of the modules in ϕ depending on the z position. A particle that traverses an ideally aligned tracker in the $z - r\phi$ plane, is reconstructed with a different curvature if a twisted geometry is assumed in the track reconstruction. This means, the χ^2 remains virtually unchanged but a wrong momentum is assigned to the particle.

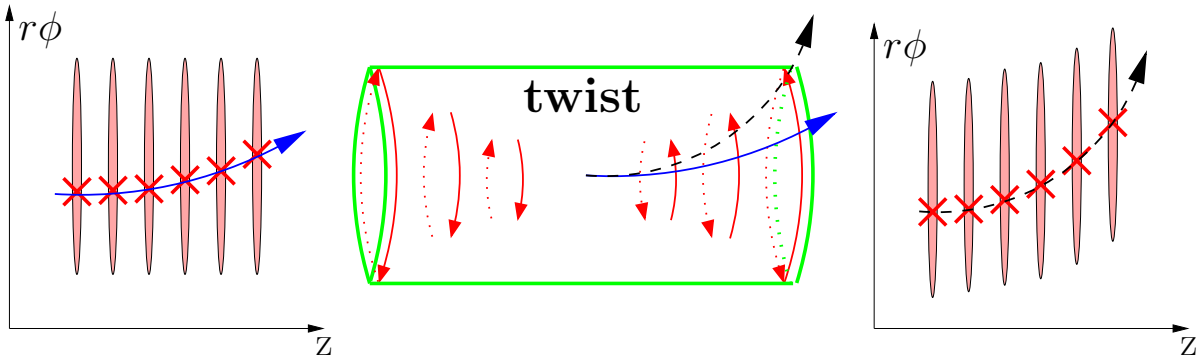

Figure 3.2: Basic principle of weak modes shown for the twist. The left and the right plot are twist-free and twisted tracker, respectively. For details, see caption of figure 3.1. In the middle, a twist of the tracker is sketched. The lengths of the small red arrows symbolise the amplitude of the module shifts. The blue solid arrow shows the real trajectory of the particle and the dashed black arrow indicates the reconstructed track.

The nine basic systematic distortions in a system with cylindrical geometry and multiple layers are biases in the cylindrical coordinates (Δr, Δz, and $\Delta\phi$) each as a function of r, z, or ϕ, listed in table 3.1. Which of these distortions actually are preserving the

	Δr	Δz	$\Delta\phi$
vs. r	radial	telescope	layer rotation
vs. z	bowing	z-expansion	twist
vs. ϕ	elliptical	skew	sagitta

Table 3.1: Basic systematic distortions for the CMS tracker.

χ^2 strongly depends on the type of tracks used for the alignment. In a former study, which was performed in connection with the first CRAFT exercise [27], it was tested which modes are weak modes using only cosmic tracks for the alignment following the approach described in reference [28]. The study was completely based on Monte-Carlo simulation.

Exemplary, the outcome is shown for the layer rotation and the twist in figure 3.3. By the introduction of the layer rotation, the normalised χ^2 becomes on average much larger. This reflects the fact that this mode can be cured by the alignment with cosmic tracks. The module positions after the alignment have a slight spread in r$\Delta\phi$ with respect to the design position but are distributed around 0 μm. The simulated twist, however, is not fully recovered by the alignment. The spread in r$\Delta\phi$ of the module positions after the alignment is larger than for the layer rotation and a z dependency of the module shifts remains. This is expected because the twist changes the χ^2 only slightly.

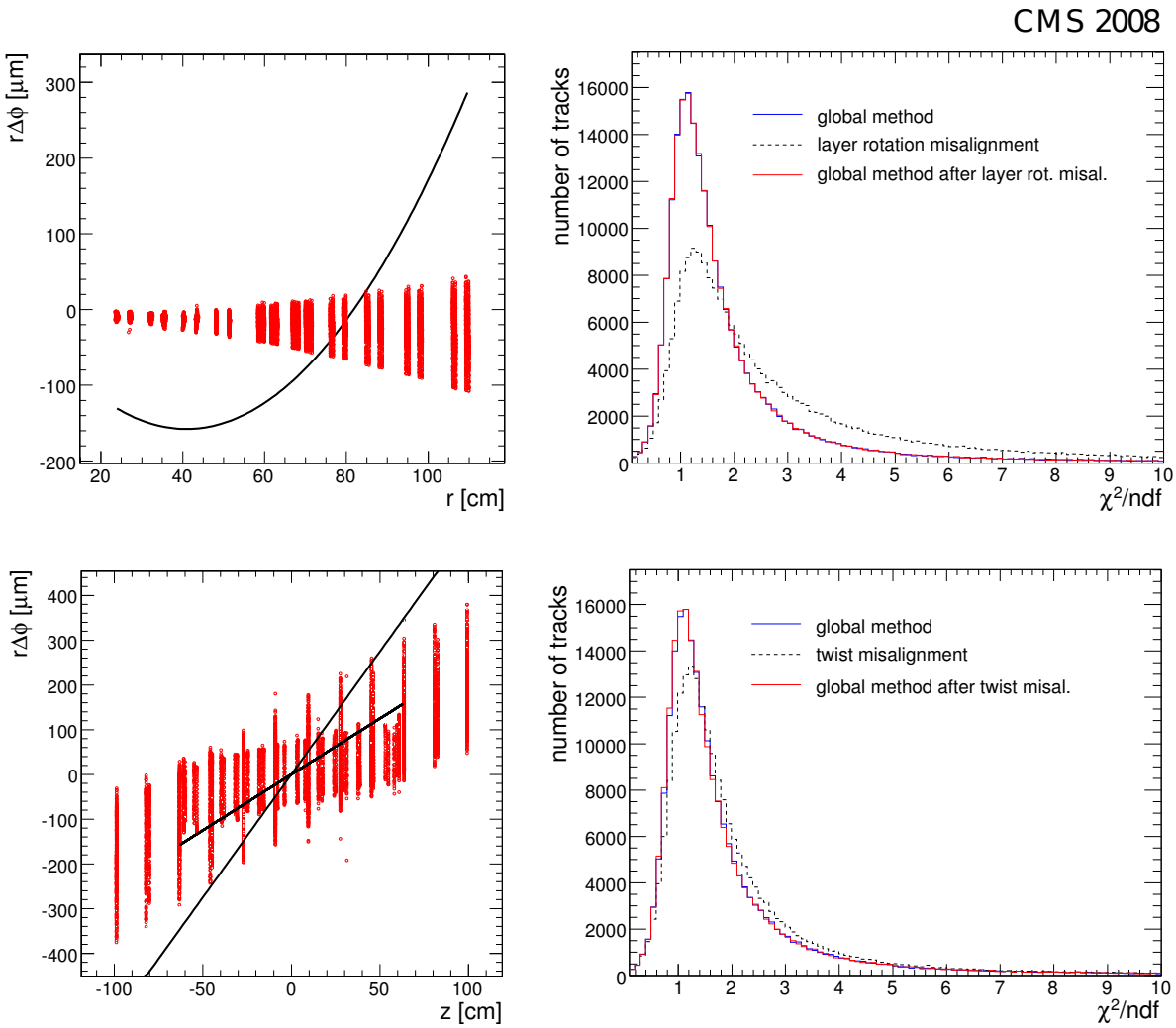

Figure 3.3: Alignment results for layer rotation (upper row) and twist (lower row). On the left, module shifts with respect to design geometry are shown. The black lines show the average position for TIB and TOB modules before alignment and the red dots each module position after alignment. On the right, the normalised χ^2 is shown for distorted geometry (black dashed line), realigned geometry (red solid line), and aligned geometry without previous systematic distortion (blue solid line). [27]

3.3.3 Constraints against Weak Modes

There are different complementary possibilities to avoid certain weak modes. The basic idea is to find additional requirements for the geometry that are not fulfilled by the distorted geometry.

One possibility, which was already mentioned shortly in the previous section, is the usage of different track topologies, e. g. collision tracks and cosmic tracks. An example for the benefit against layer rotation is illustrated in figure 3.4. If the geometry used

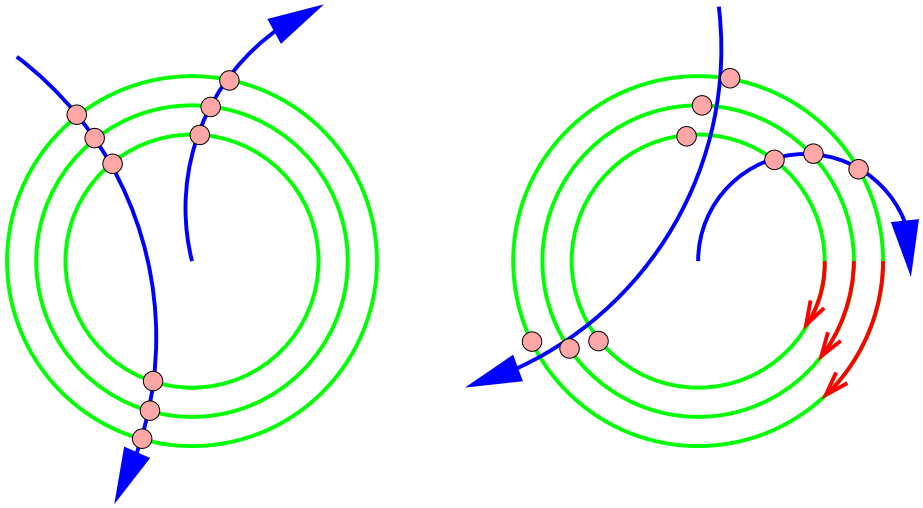

Figure 3.4: Effect of layer rotation on cosmic (long arrow) and collision (short arrow) tracks. The green rings are layers of the tracker in the $x - y$ plane and the salmon circles symbolise positions of modules hit by the particles. Modules on the left are in ideal position, on the right rotated along the red arrows.

for track reconstruction has a layer rotation with respect to the real geometry, collision tracks are reconstructed with a changed curvature preventing a change in χ^2 whereas cosmic tracks can not be properly reconstructed resulting in a large χ^2. Thus, it is obvious that layer rotations, which are weak modes for collision tracks, are constrained by the use of cosmic tracks. This combination of collision and cosmic tracks was used for the alignment derived from the complete 2010 dataset (*GR10_v3*), which was used in the beginning of the data taking of 2011.

In addition, constraints from the masses of resonances can be used. The mass of the Z boson is known to a very high precision, $m_Z = 91.1876 \pm 0.0021\,\text{GeV}$, and has a small width of $2.4952 \pm 0.0023\,\text{GeV}$ [12]. Thus, a Z boson decaying into two muons is an ideal standard candle for testing an alignment. By reconstructing the Z-boson mass in $Z \to \mu^+\mu^-$ events, a twist was found in the alignment obtained with the 2010 data (*GR10_v3*). Hereupon, the twist was removed by hand in order to provide a twist-free geometry (*GR10_v4*). The Z-boson mass measurements for these two geometries are shown in figure 3.5. A twist appears as a dependence of the Z-boson mass on the pseudorapidity of the muons. For 2011, the Z-boson mass constraint is already implemented in the common alignment approach and a new geometry (*GR10_v5*) is produced.

A third possibility is to get reference measurements from external systems. So far, this is not used in the alignment, but it constitutes a valuable tool for the monitoring of the alignment. The study described in the next section, for instance, uses calorimeter information as an independent reference. After implementing the Z-boson mass constraint into the common alignment approach, it remains the only developed validation tool for certain weak modes.

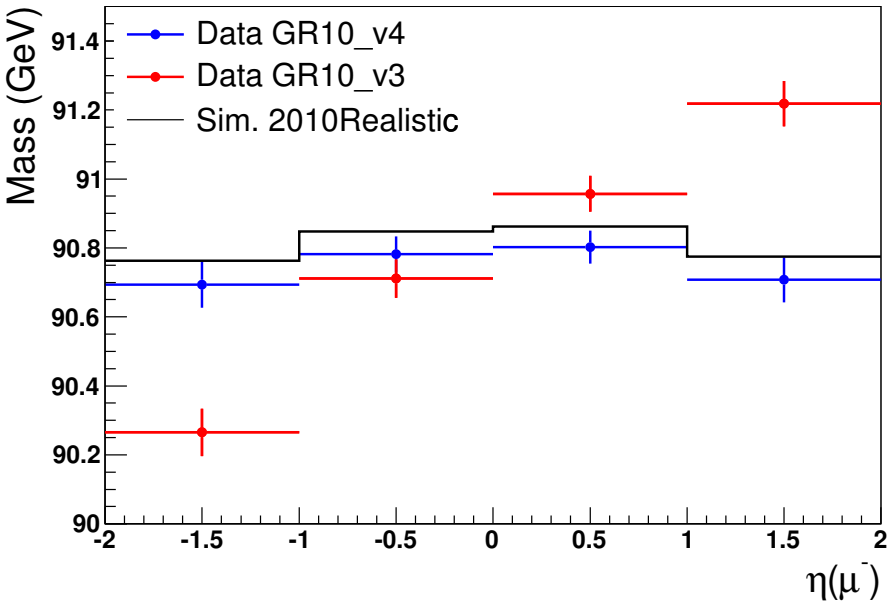

Figure 3.5: The invariant mass of the two muons from a Z-boson decay is shown as a function of the pseudorapidity of the negatively charged muon. The black line shows the result for a twist-free simulation. The red points represent the alignment derived from the complete 2010 dataset (*GR10_v3*). For the blue points (*GR10_v4*), the twist was removed by hand from the alignment. [15]

3.4 Momentum Bias Study

Since momentum changing weak modes are supposed to have the largest impact on physics, a study is set up to validate tracker alignments in the context of $\Delta\phi$ modes, which directly bias the reconstructed curvature of tracks with an opposite effect for positively and negatively charged particles. With the help of test beams, it has been shown that the calorimeter has a charge independent response for pions with energies of $E > 5\,\text{GeV}$ [20]. Hence, the energy deposited by isolated tracks in the HCAL represents a charge- and tracker-independent reference. This can be used for the determination of the bias by comparing the ratio $\frac{E}{p}$ for positively and negatively charged particles. The principle is first tested with simulated single pions on tracker geometries which are constructed to have weak modes. Finally, those tracker geometries are tested which are obtained from the official alignments that are most up to date at the time of writing.

3.4.1 Concept

The concept can be split into three parts. First of all, the relation between a misalignment $\Delta\phi$ and the resulting bias in the track momentum has to be found. Although the layer rotation is not a weak mode in the current tracker alignment using cosmic tracks, it is used to demonstrate this part since it is the most simple $\Delta\phi$ weak mode for collision

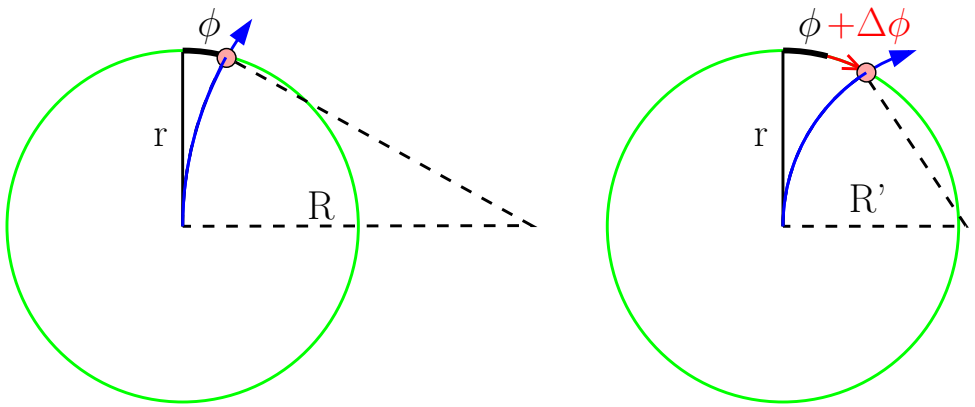

Figure 3.6: Effect of misalignment on track curvatures. On the left, a charged particle traverses the tracker and hits a module located at the radius r and the angle ϕ relative to the impact of a hypothetical straight track. The dashed lines represent the radius of curvature R. On the right, the hit module is shifted by $\Delta\phi$ resulting in a reconstructed radius of curvature of R'.

tracks. Then the calorimeter energy is introduced as the required independent reference. Finally, the calculation for the layer rotation is translated to the twist case.

- **Relation:** The relation between the radius of curvature R of a track and its relative azimuthal angle ϕ with respect to a hypothetical straight track starting in the same direction is given by

$$\phi + \Delta\phi = \arcsin\left(\frac{r}{2 \cdot R}\right) \tag{3.2}$$

where r is the radial position of the hit module and $\Delta\phi$ a possible shift of this module in ϕ (see figure 3.6). Possible signs of ϕ and R denote the direction of the curvature. Without loss of generality, tracks of positively charged particles are defined to have a positive curvature. When adding equation (3.2) for two oppositely charged particles with a real curvature of identical absolute value, i. e. $\phi^+ = -\phi^-$, ϕ can be eliminated leading to

$$\Delta\phi = \frac{1}{2}\left(\arcsin\left(\frac{r}{2 \cdot R^+}\right) + \arcsin\left(\frac{r}{2 \cdot R^-}\right)\right) \tag{3.3}$$

where R^+ and R^- are the radii of curvature of the positively and negatively charged particle, respectively, with $R^- = -R^+$. Via the balance of Lorentz force and centripetal force, each of these radii can be translated to the corresponding transverse momentum

$$R = \frac{p_T}{q \cdot B_z} \tag{3.4}$$

where q is the charge of the particle and B_z the magnetic field in z direction. Using particles with one elementary charge ($q = \pm 1\,\mathrm{e}$) and the magnetic field as in the

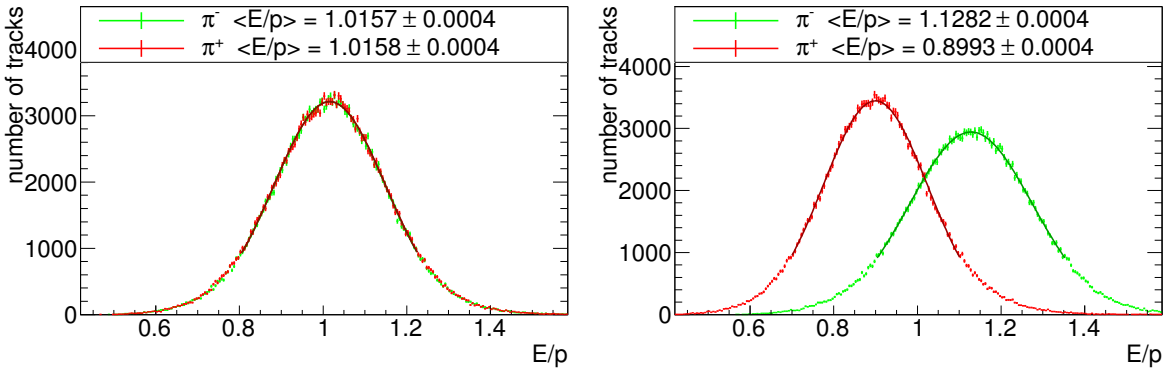

Figure 3.7: Simulated energy over momentum distributions for ideal geometry (left) and layer rotation (right). Gaussians are fitted to the distributions of the negatively (green crosses) and positively (red crosses) charged pions.

CMS tracker gives

$$R[\mathrm{m}] = \frac{p_T[\mathrm{GeV}]}{0.3 \cdot q[\mathrm{e}] \cdot B_z[\mathrm{T}]} \overset{B_z=3.8\,\mathrm{T}}{\longrightarrow} \mathrm{sign}(q) \cdot \frac{p_T[\mathrm{GeV}]}{1.14}. \tag{3.5}$$

Combining equations (3.3) and (3.5) results in

$$\Delta\phi = \frac{1}{2}\left(\arcsin\left(\frac{0.57 \cdot r[\mathrm{m}]}{p_T^+[\mathrm{GeV}]}\right) - \arcsin\left(\frac{0.57 \cdot r[\mathrm{m}]}{p_T^-[\mathrm{GeV}]}\right)\right). \tag{3.6}$$

Following the derivation, p_T^+ and p_T^- are the reconstructed transverse momenta of two particles with exactly the same real transverse momentum.

- **Reference:** At this point, the calorimeter comes into play as a charge- and tracker-independent reference. The tracks are propagated to the calorimeter to assign the corresponding energy to the track. Some details of the assignment are given in the next section. There are basically two ways of employing the assigned calorimeter energy.

  One would be to literally use it as a reference. If the energy measurement would be perfect, the measured energy could be used to select tracks of positively and negatively charged particles with identical real transverse momentum. With the help of the reconstructed transverse momenta of these tracks, the misalignment could be calculated via equation (3.6). However, for most tracks ($p_T \lesssim 100\mathrm{GeV}$), the energy resolution effects in the calorimeter are larger than possible momentum biases in the tracker, and only the average transverse energy $\langle E_T \rangle$ of many particles can be used as an indicator for the real transverse energy. Therefore, the misalignment can not be determined event by event but from the average inverse transverse momentum $\left\langle \frac{1}{p_T} \right\rangle$ of these particles.

  The second approach is to evaluate the ratio energy over momentum separately

for positively and negatively charged particles and to determine the mean of a Gaussian fitted to this distribution $\left\langle \frac{E}{p} \right\rangle$ as a direct measure for the momentum bias as shown in figure 3.7. The fit is iterated and each time the range of the fit is restricted to plus and minus 1.5 times the width of the previous fit in order to exclude the non-Gaussian tails, which are caused among other things by fake tracks, energy mismeasurements and wrong energy assignments. Therefore, this method is more robust against these effects than the first method and will be used in the following. According to this approach, the substitution,

$$\frac{1}{p_T} \to \frac{1}{\langle E_T \rangle} \cdot \left\langle \frac{E}{p} \right\rangle, \tag{3.7}$$

is made where $\langle E_T \rangle$ is the average transverse energy of those particles that are falling into the fitted core region of the energy over momentum distribution. The substitution (3.7) in equation (3.6) results in

$$\Delta\phi = \frac{1}{2}\left(\arcsin\left(\frac{0.57 \cdot r[\mathrm{m}]}{\langle E_T \rangle^{+}\,[\mathrm{GeV}]}\left\langle \frac{E}{p} \right\rangle^{+}\right) - \arcsin\left(\frac{0.57 \cdot r[\mathrm{m}]}{\langle E_T \rangle^{-}\,[\mathrm{GeV}]}\left\langle \frac{E}{p} \right\rangle^{-}\right)\right). \tag{3.8}$$

- **Translation:** In order to determine a twist, the bias is measured in different intervals of z and a function $\Delta\phi(z)$ is calculated. But before that, possible functional relations have to be figured out that would constitute weak modes for the alignment with collision tracks. Using relation (3.5), equation (3.2) is approximated for small angles ϕ and $\Delta\phi$,

$$p_T[\mathrm{GeV}] \approx \frac{0.57 \cdot r[\mathrm{m}]}{\phi + \Delta\phi}. \tag{3.9}$$

This is valid for particles with transverse momenta above 10 GeV propagating in the magnetic field of CMS, where ϕ is smaller than 0.1 rad at a radius of $r = 1\,\mathrm{m}$. From geometry comparisons, the maximum expected misalignment $\Delta\phi$ is of the order of 1 mrad. Neglecting energy loss effects, which are small for the considered energies, the curvature of a track does not change on its way through the tracker. This naturally holds for unbiased tracks, but it also holds for tracks reconstructed with geometries biased by weak modes since otherwise the χ^2 of the track would be raised as a change of curvature is not foreseen in the track reconstruction. So in the given approximation, $\Delta\phi$ is proportional to r for the layer rotation. This translates via $r = z \cdot \tan\theta$ to

$$\Delta\phi \propto z \tag{3.10}$$

for the twist because the polar angle θ of a particle hardly changes during the propagation through the tracker. Hence, for layer rotation and twist, only the linear biases are weak modes for the alignment with collision tracks. There can still be a combination of both, which is not linear in each contribution. But as the layer rotation is not a weak mode for the alignment with cosmic tracks as shown in section 3.3.3, a possible twist is expected to be approximately linear in z.

3.4.2 Track Selection

For a reasonable measurement of the ratio energy over momentum, tracks are needed which are well reconstructed and isolated. The exact requirements are described in the following.

- **Trigger:** In order to select the required tracks during data taking, an appropriate trigger is needed. There are two High-Level Triggers which are fired by isolated tracks, the *HLT_IsoTrackHB* for the barrel region ($|\eta| < 1.15$) and the *HLT_IsoTrackHE* for the endcap region ($1.1 < |\eta| < 2.2$). As high-p_T QCD events frequently contain isolated tracks, the events are preselected on trigger Level 1 to have a jet with an energy threshold depending on the instantaneous luminosity. To avoid a bias due to this energy threshold in the later correlation of the isolated track with the energy measurement, only tracks are considered that differ in the transverse momentum from the highest energetic Level-1 jet ($\Delta p_T > 4\,\text{GeV}$). The track isolation on trigger level only accounts for charged particles. It is defined by extrapolating all tracks to the ECAL surface. No track with a transverse momentum of $p_T > 2\,\text{GeV}$ is allowed to be in a circle with a radius of 40 cm around the impact point of the considered track. In addition, the tracks have to fulfil a momentum requirement to keep the trigger rate low. For low instantaneous luminosities in the beginning of the data taking, the requirement started at $p > 10\,\text{GeV}$ and was then raised to a final threshold of

  $$p > 38\,\text{GeV}, \tag{3.11}$$

  for later stages of the data taking.

- **Number of hits:** Long tracks ensure a small uncertainty on the momentum measurement and on the extrapolation to the calorimeter.

  So only tracks with many valid hits,

  $$\#\text{hits}_{\text{valid}} \geq 13, \tag{3.12}$$

  are included. The track of a particle has an invalid hit if the particle traverses a module without leaving a hit that can be matched to its track. Tracks with invalid hits, also called 'lost hits', are not selected,

  $$\#\text{hits}_{\text{lost}} = 0. \tag{3.13}$$

  The distributions of the number of valid and lost hits are shown in the upper plots of figure 3.8.

  In addition, the outermost hit of a track needs to be at a radius of

  $$R_{\text{outer}} > 99\,\text{cm}, \tag{3.14}$$

  which explicitly excludes tracks with the outermost hit on the second to last barrel

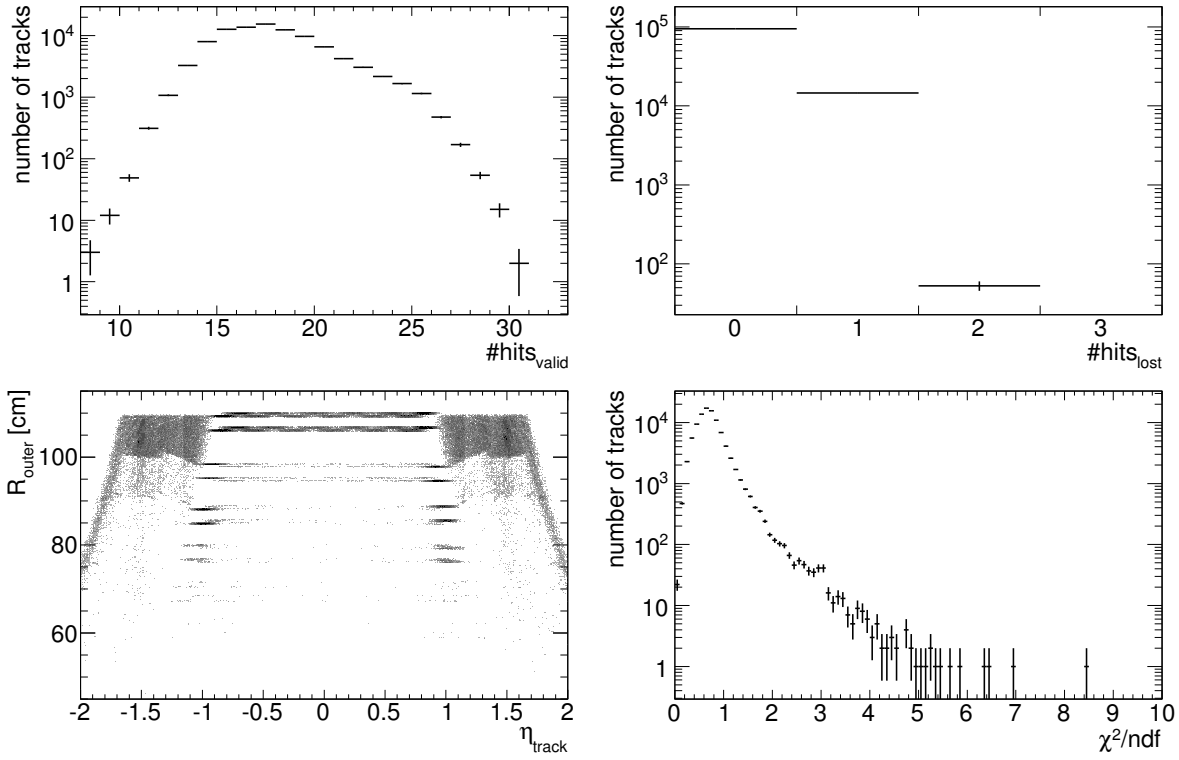

Figure 3.8: Distributions of track quality variables after the cuts on all other variables have been applied. The upper plots show the number of valid (left) and lost (right) hits per track. The lower left plot shows the two-dimensional distribution of the η of the track and the radius of its outermost hit illustrating the layer structure of the tracker in the barrel region. Each layer has two radii because the tracker modules are staggered to provide hermetic coverage in ϕ. The lower right plot shows the normalised χ^2 distribution.

layer as can be seen from the lower left plot of figure 3.8. This constraint renders a limitation of the momentum uncertainty unnecessary.

- **Normalised χ^2:** Fake tracks are reduced by cutting on

$$\frac{\chi^2}{ndf} < 5. \tag{3.15}$$

This only rejects a few outliers as shown in the lower right plot of figure 3.8.

- **Late showering particles:** In consequence of the difference in response of ECAL and HCAL (see section 2.2.2), the measured energy of a particle depends on the position in the detector where the showering starts. The ECAL energy in a cylinder of 9 cm radius around the track impact point on the ECAL surface $E_{9\,\mathrm{cm\ cylinder}}$ is assigned to the track. The distribution of this energy is plotted on the left of figure 3.9. Even if particles do not shower inside the ECAL, they deposit energy via

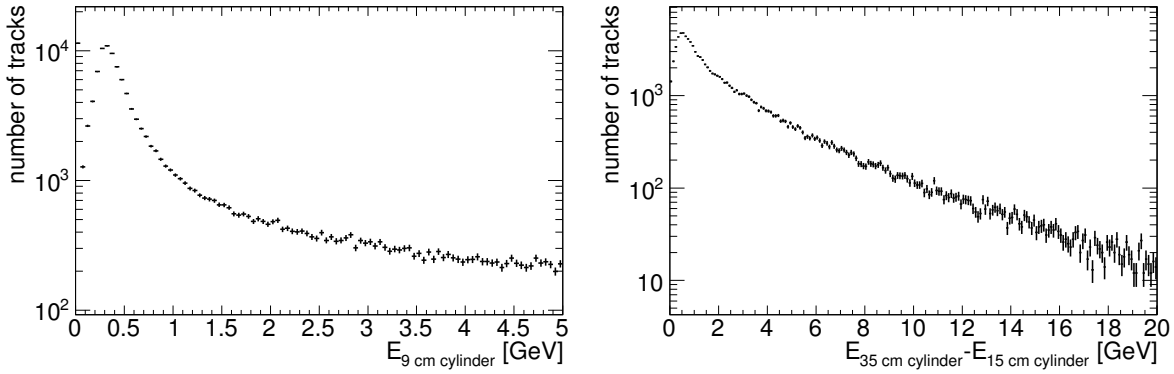

Figure 3.9: Distributions of the energy around the tracks after the cuts on all other variables have been applied. The left plot shows the energy in a cylinder of 9 cm radius around the track impact point on the ECAL surface. The right plot shows the difference between two concentric cylinders of 35 cm and 15 cm radius, respectively.

ionisation and bremsstrahlung, which leads to the peak slightly above 0.3 GeV. The large number of tracks in the lowest energy interval is due to the noise suppression threshold in the ECAL. The particles showering in the ECAL populate the long tail to higher energies.

By requiring

$$E_{9\,\text{cm cylinder}} < 1\,\text{GeV}, \tag{3.16}$$

a clean hadron sample is selected containing mainly particles that do not shower until they enter the HCAL.

- **Isolation:** As the trigger only takes care of charged particles in the vicinity of the considered track, an isolation requirement against neutral particles has to be added. For this purpose, the ECAL energy between two concentric cylinders around the track impact point on the ECAL surface with a radius of 35 cm and 15 cm, respectively, has been studied as shown in the right plot of figure 3.9 and is restricted to

$$E_{35\,\text{cm cylinder}} - E_{15\,\text{cm cylinder}} < 8\,\text{GeV}. \tag{3.17}$$

By using the ECAL energy for the isolation, no bias is fed into the energy measurement in the HCAL.

- **Energy selection:** As only late showering particles are used, the energy measurement is solely based on the HCAL, namely by taking the energy in a 3×3 HCAL tower cluster ($E_{\text{HCAL}(3\times3)}$) around the extrapolated impact point of the track. The lowest acceptable energy needs to be well above the momentum cut of the trigger in order not to bias the measurement of the ratio energy over momentum. The width of the distribution of this ratio is dominated by the energy resolution. Therefore, the average energy response and resolution in the HCAL of

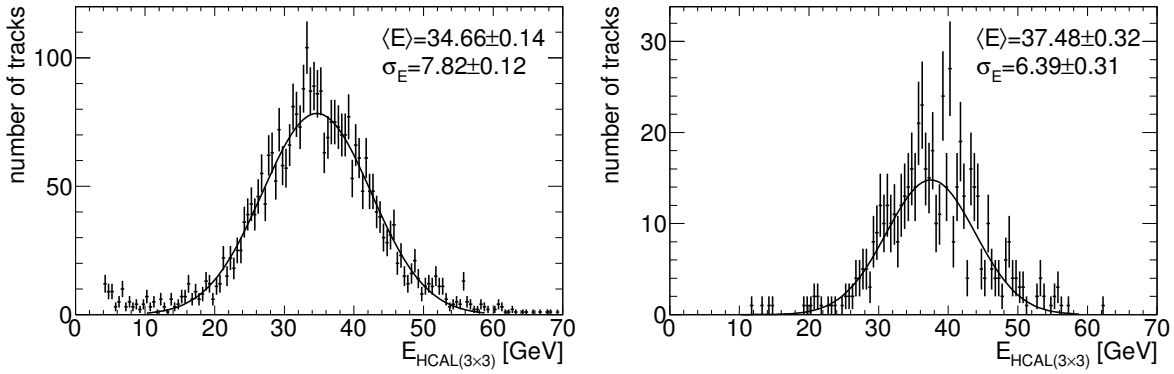

Figure 3.10: HCAL energy distribution of data (left) and simulated single pions (right) for tracks with a reconstructed momentum of $38\,\text{GeV} < p < 38.1\,\text{GeV}$ fulfilling the selection criteria. A Gaussian is fit around the mean of the distribution. Mean $\langle E \rangle$ and width σ_E of the fit are given in the upper right.

particles with a momentum of 38 GeV, which is the current trigger threshold, is determined as shown in figure 3.10. The response is found to be $34.7 \pm 0.1\,\text{GeV}$ with a resolution of $7.8 \pm 0.1\,\text{GeV}$ for data and $37.5 \pm 0.3\,\text{GeV}$ with a resolution of $6.4 \pm 0.3\,\text{GeV}$ for simulated single pions. The difference in response between data and simulation is expected because the tracks in data have a steeply falling momentum spectrum while the single pions are simulated flat in momentum up to 100 GeV (see figure 3.11). Therefore, the response is biased due to migration only in data. However, as the response does not influence the method as long as it is equal for positively and negatively charged particles, the response bias in data does not harm. In addition to energy response and resolution, a possible bias in the momentum reconstruction of the trigger due to a $\Delta\phi$ misalignment has to be considered for the energy cut. Assuming a maximum misalignment of the order of $\Delta\phi = 1\,\text{mrad}$ at a radius of $r = 1\,\text{m}$, the bias in the momentum of these tracks is at most about 3 GeV (see eq. (3.9)). The Gaussian fit to the $\left\langle \frac{E}{p} \right\rangle$ distribution is done in a range of $\pm 1.5\,\sigma$ around the mean. Taking the response corrected for the possible bias and adding 1.5 times the energy resolution leads to an energy cut of

$$E_{\text{HCAL}(3\times3)} > 50\,\text{GeV}, \tag{3.18}$$

which is safe for data and simulation. At lower trigger thresholds, significantly less events are taken so there is no cut dedicated to these run periods.

In addition, intervals in $E_{\text{HCAL}(3\times3)}$ need to be introduced as the ratio energy over momentum is energy dependent due to non-linearities in the calorimeter and possibly due to a bias in the momentum measurement. The results of all intervals are combined as a weighted mean. On the one hand, these intervals should not be too large because the intention is to pick tracks with preferably similar true momenta. On the other hand, an interval which is too small leads to a lack of

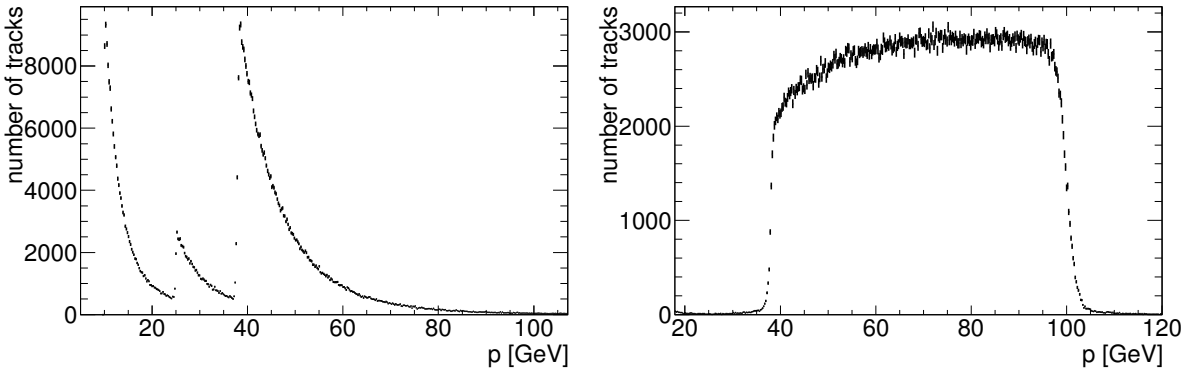

Figure 3.11: Momentum spectrum of data showing the different trigger thresholds (left) and of simulated single pions (right) for tracks fulfilling the selection criteria.

statistics. The actual interval limits depend on the sample which is used and are specified in the corresponding sections.

3.4.3 Proof of Principle

A valuable test of this method can be done with the help of simulated events containing single pions. The tracks in these events are refitted with three different geometries, the first is the design geometry, the second has an additional layer rotation, and the third a simulated twist. The trigger was already required before the refitting, the rest of the track selection was applied to the refitted tracks. As the momentum spectrum of the tracks was simulated flat between 1 and 100 GeV, equidistant energy intervals with the limits $\{50, 55, 60, 65, 70, 75, 80\}$ can be used. The upper limit is chosen to be well below the corresponding cut off in the artificial momentum spectrum to avoid a bias. A total of about 360 000 tracks is selected where the exact number depends on the geometry used for the track reconstruction. For each geometry, the misalignment is measured in 18 intervals of η translated to the location in z where the track reaches a radius of $r = 1\,\mathrm{m}$. Since the underlying z distribution is not flat, a bin centre correction is done to take this into account. According to the proportionality (3.10), a linear function in z is fitted to the outcome. A layer rotation leads to an offset in this function and a twist can be read off from the slope. The measured misalignments and the fit results for the three geometries can be seen in figure 3.12. All fit results agree with the simulated geometries.

3.4.4 Results

After the validity of the method has been proven, it can be used to check aligned geometries. This is done with tracks from the complete 2010 dataset and the 2011 data collected until the technical stop in the beginning of July, which corresponds to a recorded luminosity of about $1.2\,\mathrm{fb}^{-1}$. This amounts to about 85 000 isolated tracks, which can be split into two energy intervals with the limits $\{50, 58, 80\}$. The second

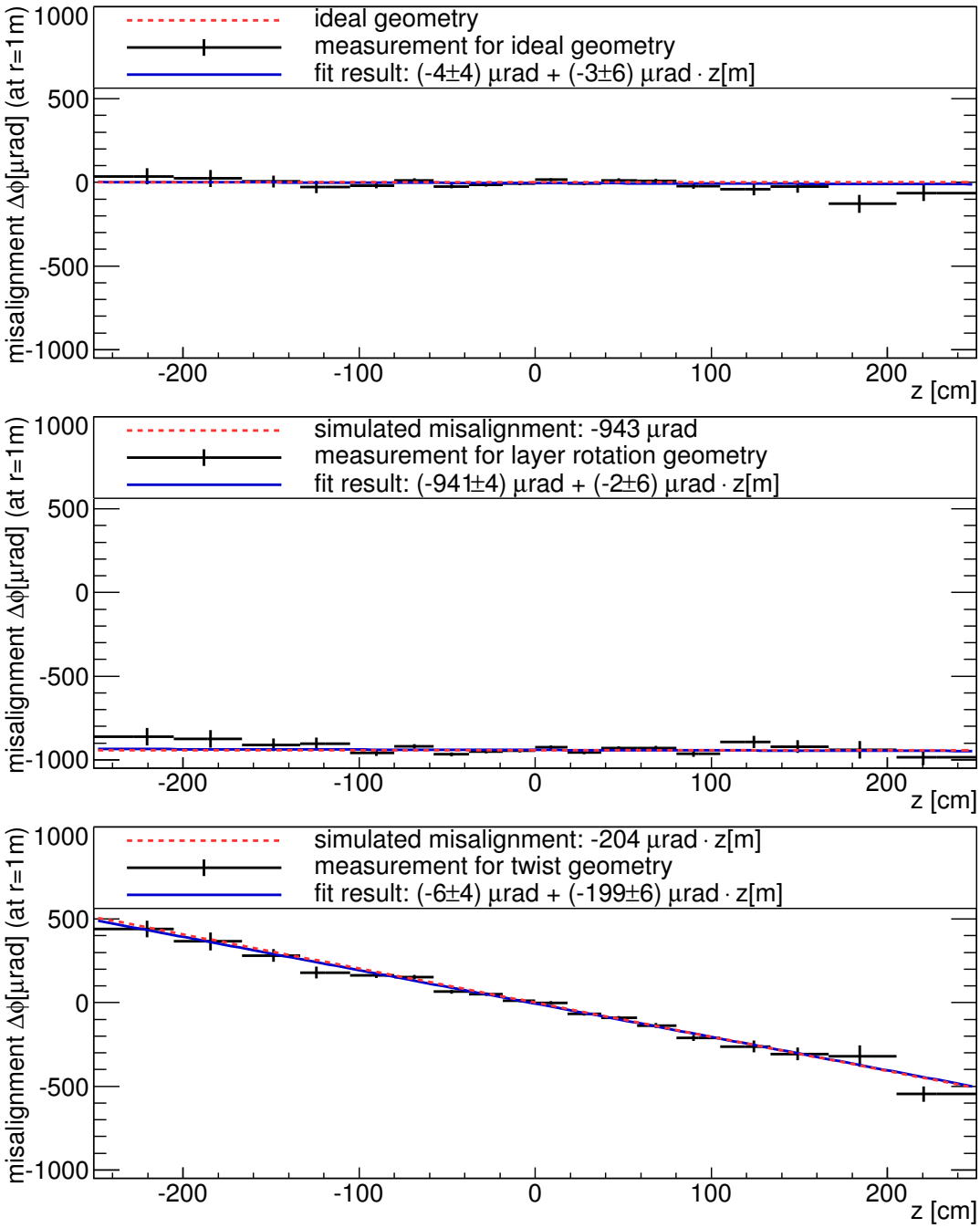

Figure 3.12: Misalignment $\Delta\phi$ as a function of z in MC for design geometry (upper plot), layer rotation (middle plot), and a twist (lower plot). The points represent the measurement and the solid line the corresponding linear fit. Vertical error bars are statistical errors whereas horizontal error bars simply illustrate the interval width. The dashed line shows the simulated misalignment.

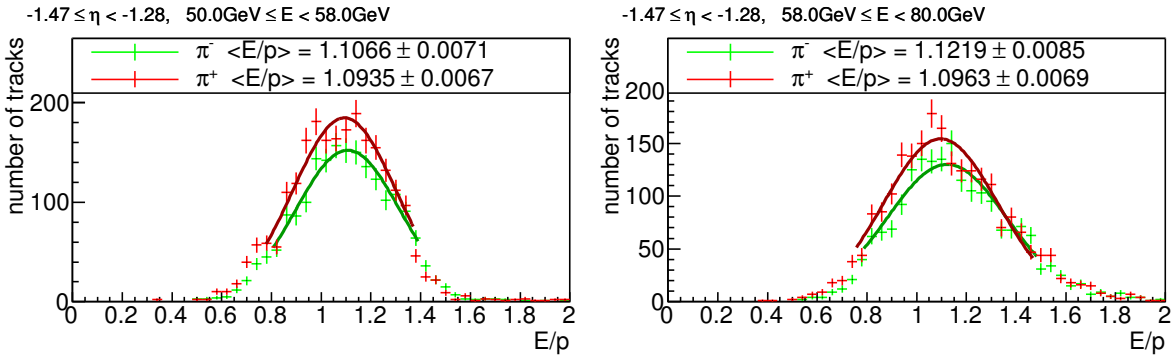

Figure 3.13: Energy over momentum distribution for measured particles reconstructed with the *GR10_v3* geometry for both energy intervals at $-1.65 \leq \eta < -1.47$.

interval is chosen much larger to compensate for the steeply falling momentum spectrum. At the moment, the interval widths for data are larger than for simulation but will be decreased in future with rising number of events. The large intervals do not seem to cause problems since using the same intervals for simulation gives consistent results. The division in η is done in the same way as for the simulation. The validation is performed for the three geometries presented in section 3.3.3. The energy over momentum distributions for all intervals and all three geometries are shown in appendix A and an example is given in figure 3.13. It shows the difference in Energy over momentum between negatively and positively charged particles with a pseudorapidity of $-1.65 \leq \eta < -1.47$ reconstructed with the *GR10_v3* geometry. In figure 3.14, the validation result for the three geometries is presented. This method not only confirms the findings of the η dependent Z-boson mass measurement but also quantifies the twist. In the *GR10_v3* geometry, it is found to be

$$\frac{\Delta\phi}{\Delta z} = 133 \pm 12 \, \frac{\mu\text{rad}}{\text{m}}. \tag{3.19}$$

In the *GR10_v4* geometry, the absolute value is strongly reduced, but the twist seems to be slightly overcorrected to

$$\frac{\Delta\phi}{\Delta z} = -25 \pm 12 \, \frac{\mu\text{rad}}{\text{m}}. \tag{3.20}$$

In the *GR10_v5* geometry, the best result,

$$\frac{\Delta\phi}{\Delta z} = -9 \pm 12 \, \frac{\mu\text{rad}}{\text{m}}, \tag{3.21}$$

is achieved with no significant twist remaining.

For a particle with a transverse momentum of 100 GeV, a twist of $\frac{\Delta\phi}{\Delta z} = 12 \, \frac{\mu\text{rad}}{\text{m}}$, which corresponds to the current statistical uncertainty, would lead to a relative transverse-momentum bias of 0.85 % for $\eta = 2.1$. This is smaller than the expected transverse-momentum resolution of 1-2 % [18].

Concluding, one can say that the measurement of $\left\langle \frac{E}{p} \right\rangle$ serves as a valuable tool for the validation of geometries concerning momentum biases.

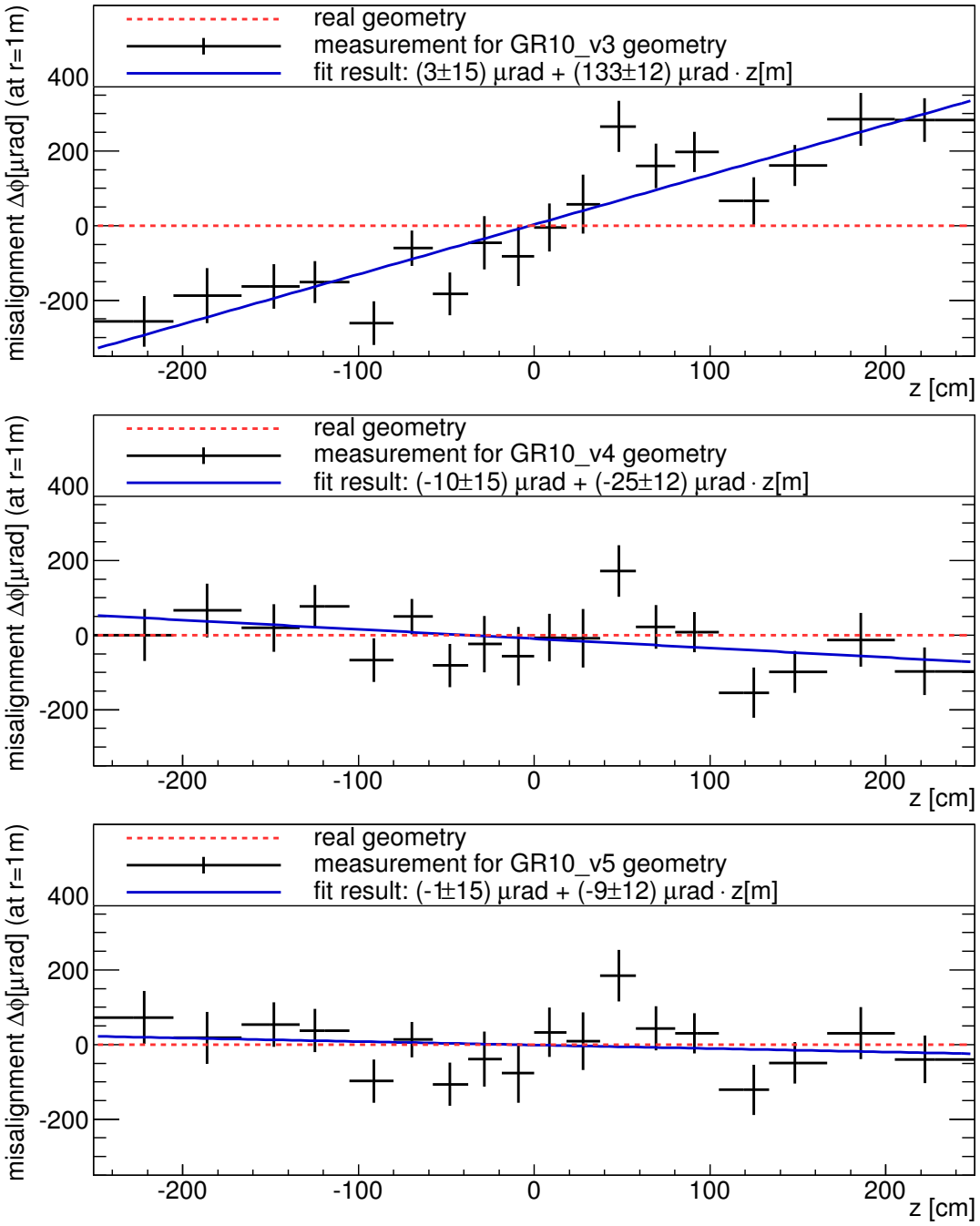

Figure 3.14: Misalignment $\Delta\phi$ as a function of z in data for the geometry of the full 2010 tracker alignment (*GR10_v3*) in the upper plot, the geometry with subsequently removed twist (*GR10_v4*) in the middle plot, and the geometry obtained using the Z-boson mass constraint (*GR10_v5*) in the lower plot. The dashed line shows the real geometry of the tracker, which is to be found by the alignment. For further details, see caption of figure 3.12.

4 Phenomenology of Top Quarks

Top quarks are described in the theoretical framework of the Standard Model of particle physics. The mass of the top quark cannot be calculated from basic principles in the Standard Model but is defined via a parameter which is to be measured. However, the production and decay of top quarks, provided that the mass is known, are precisely predicted and thus can be measured to test the Standard Model. A comprehensive review of the phenomenology of top-quark physics is given in reference [29].

4.1 Mass Definition

The mass is one of the most precisely measured top-quark properties. The current value, which is an average of different mass measurements from Tevatron, is $173.2 \pm 0.9\,\text{GeV}$ [13]. The high measurement precision necessitates a careful theoretical definition of the top-quark mass to correctly interpret the measured value. There are basically two different theoretical definitions, the on-shell mass also referred to as pole mass and the $\overline{\text{MS}}$ mass.

- **On-shell mass scheme:** The invariant mass reconstructed from the decay products of a non-virtual[1] particle that does not form bound states is its on-shell mass. As a top quark decays before its colour charge leads to hadronisation, the on-shell mass seems to be its natural mass definition. However, the colour charge of the top quark is conserved by the b quarks arising from the decay. The b quarks then form bound states together with other quarks, which do not belong to the decay products of the top quark. This leads to long-range effects, which cannot be calculated in perturbative QCD leading to an intrinsic ambiguity of the on-shell mass [30, 31].

- **$\overline{\text{MS}}$ mass scheme:** In the modified minimal subtraction ($\overline{\text{MS}}$) scheme, non-perturbative ambiguities are cured. Divergences in perturbative QCD are removed via renormalisation introducing an arbitrary renormalisation scale. The relation between on-shell and $\overline{\text{MS}}$ mass is calculated to the order α_s^3 [32]. Choosing the $\overline{\text{MS}}$ mass itself as the renormalisation scale leads to a roughly 10 GeV lower $\overline{\text{MS}}$ mass compared to the on-shell mass.

Experimentally, Monte-Carlo generators are used for corrections in mass measurements. Therefore, the measured mass is in principle the mass parameter of the Monte-Carlo generator which is by construction close to the on-shell mass.

[1] Virtual particles serve as a mathematical construct to calculate observable processes without being observable themselves. They are not fixed to their mass shell, i. e. $E^2 - p^2 \neq m^2$.

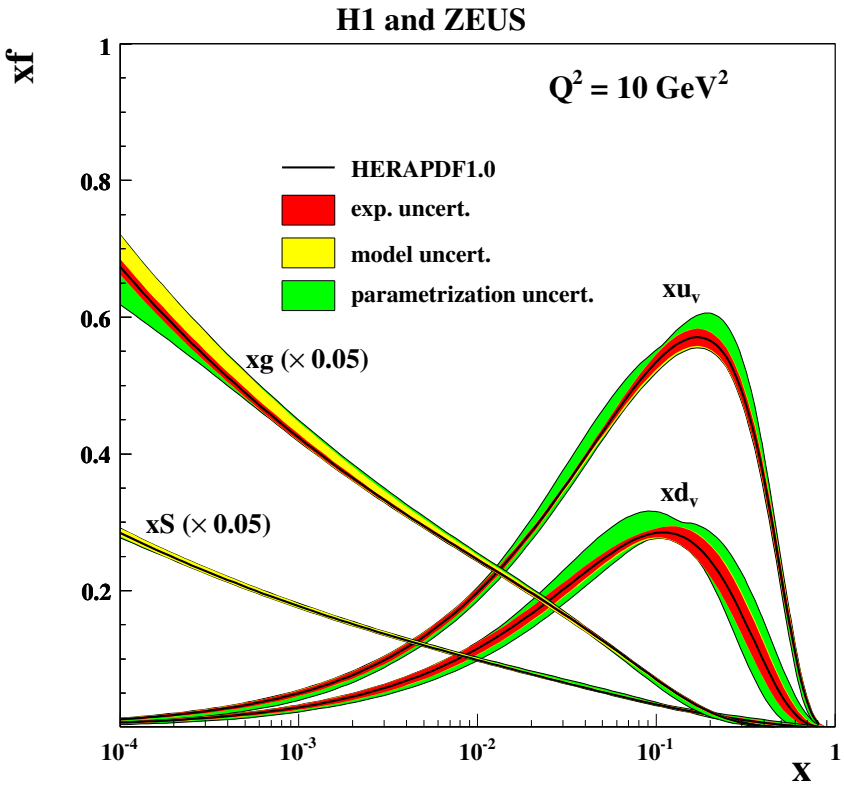

Figure 4.1: Parton distribution functions of the proton as a function of the proton momentum fraction x. The gluon and sea quark distributions are scaled down by a factor 0.05. [33]

4.2 Production Mechanisms

For the production of top quarks in proton-proton collisions, partons, i. e. quarks and gluons, inside the protons interact. Therefore, the centre-of-mass energy provided by a proton-proton collider is only partly used for the production because the interacting partons only carry fractions of the proton momentum. These fractions are given by the parton distribution functions (PDFs) of the proton, which are very precisely measured by HERA experiments [33] as shown in figure 4.1. At high momentum fractions the so-called 'valence quarks' consisting of up (u) and down (d) quarks dominate. For lower momentum fractions, the probability for finding gluons (g) rises as well as for finding so-called 'sea quarks'. Sea quarks are quarks of all flavours that are steadily produced and annihilated inside the proton.

The probability for a certain production mechanism is given by its cross section. In proton-proton collisions, the production cross section arises from the convolution of the proton PDFs and the cross section of the parton interaction. The latter results from the kinematics of the interacting partons and the matrix element of the participating interaction.

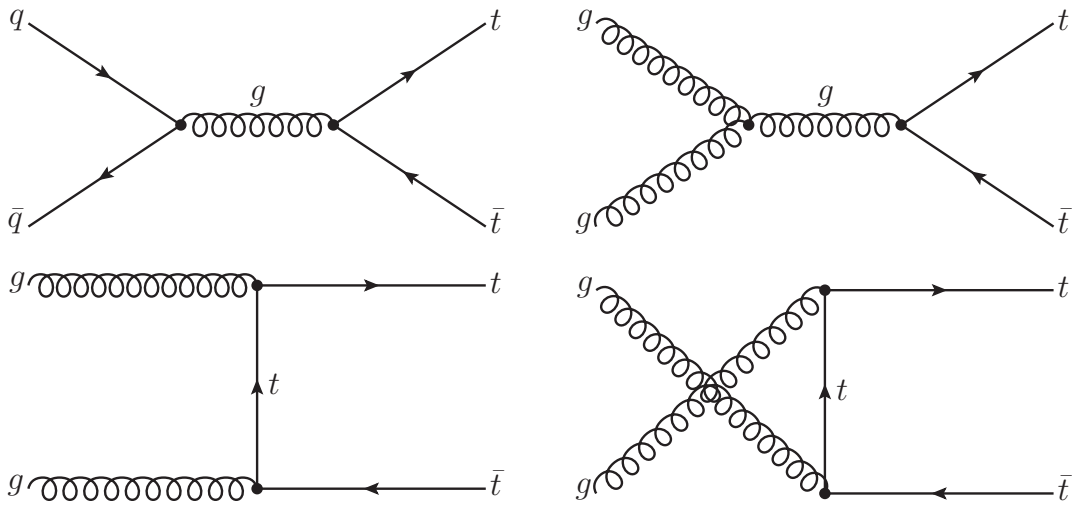

Figure 4.2: Feynman graphs for top-quark pair production. The upper left graph shows $q\bar{q}$ annihilation. The other graphs show the gluon-gluon fusion in s channel (upper right), t channel (lower left), and u channel (lower right).

In the naive expectation, the production of a pair of top quarks should be less probable than the production of a single top quark because about twice of the energy is needed to create the mass of the outgoing particles. However, the top-quark production via the weak interaction is suppressed owing to the large mass of the W boson being the mediator of the weak interaction. The production of single top quarks is only possible via the weak interaction, whereas top-quark pair production is also possible via the strong interaction. Therefore, the cross section of the top-quark pair production is expected to be significantly larger than the single-top cross section.

The main production mechanisms of top-quark pairs are the gluon-gluon fusion and the quark-antiquark annihilation, where the top quarks are produced via the strong interaction indicated by the participating gluons. The corresponding Feynman graphs are shown in figure 4.2. In the LHC proton-proton collisions at a centre-of-mass energy of 7 TeV, a momentum fraction of about 0.05 for both partons is sufficient to generate twice the top mass. As gluons dominate the proton PDFs at this momentum fraction, the gluon-gluon fusion constitutes about 85 % of the total top-quark pair production cross section.

The dominant single-top production mechanisms are illustrated in figure 4.3 showing the corresponding leading-order Feynman graphs. In each graph, a gauge boson of the weak interaction, the W boson, is participating. To obtain the corresponding Feynman graphs for antitop-quark production t and $\bar{t}$ as well as b and $\bar{b}$ have to be interchanged. One important difference between single top- and antitop-quark production in proton-proton colliders is due to the participating quark q in the initial state of the s and t channel. For top-quark production, this is an up-type quark, for antitop-quark production, this is a down-type quark. The required momentum fraction is 0.025 assuming both partons have the same momentum. Due to the significant contribution of the valence

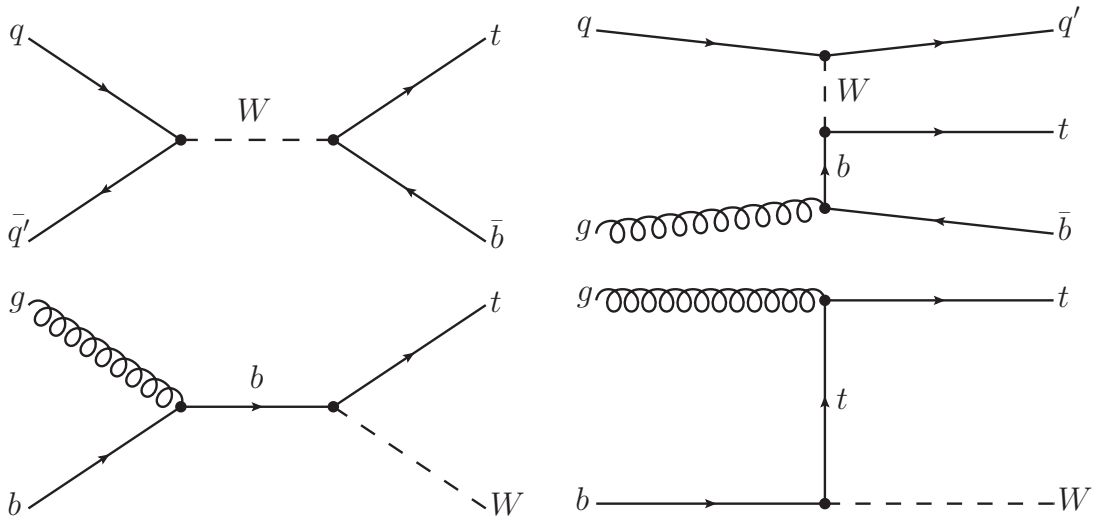

Figure 4.3: Feynman graphs for single-top production. The upper graphs show the s-channel (left) and the t-channel (right) production. The lower graphs show the associated tW production in s channel (left) and t channel (right).

quarks for momentum fractions above this threshold, the probability to find an up-type quark is nearly twice as high as the probability to find an down-type quark as can be seen from the proton PDFs in figure 4.1. Therefore, the production probability for top quarks is almost twice as high as for antitop quarks in these two single-top production channels.

4.3 Decay Channels

Due to the mass of the top quark, which is even larger than the mass of the W boson, it always decays via the weak interaction into an on-shell W boson and a down-type quark. The Cabibbo-Kobayashi-Maskawa matrix gives the probabilities for the different down-type quark flavours to occur in such a decay. A recent Standard Model fit [34] results in $|V_{tb}| = 0.999135^{+0.000057}_{-0.000018}$. $|V_{tb}|^2$ being the probability for the decay of a top quark into a bottom quark amounts to nearly 100 %.

Therefore, the different decay channels of the top quark are characterised by the decay branch of the W boson emerging from the top-quark decay. The W boson can decay either into one of the three charged leptons and the corresponding neutrino or into a pair of quarks of the first or second family. The third family is excluded as the sum of top- and bottom-quark masses is larger than the mass of the W boson. Taking into account the three possible colour states of the two quark pairs, there are six quark final states. The quarks directly form jets of colourless hadrons. Therefore, the colour cannot be measured and the flavours are hardly distinguishable. At leading order, this results in a branching ratio of $\frac{1}{9}$ for each of the three leptonic decays and $\frac{6}{9}$ for the combined hadronic decays. Higher order corrections only lead to small deviations from this assumption.

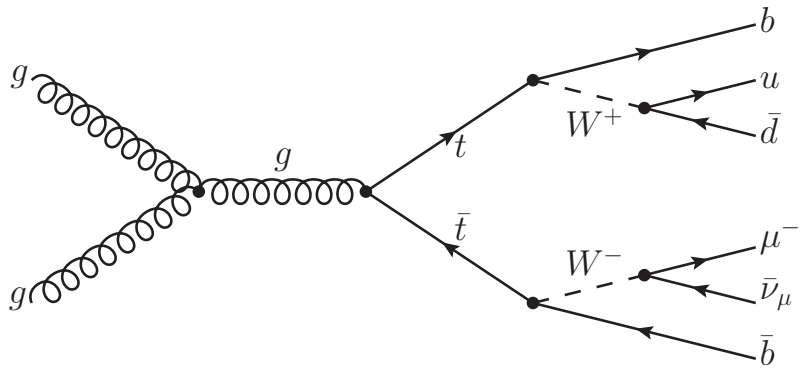

Figure 4.4: An example of top- antitop-quark pair production with subsequent decay via the muon+jets decay channel.

Due to the two W bosons in top-quark pair decays, the decay channels are classified in three types, namely the fully hadronic decay channel, where both W bosons decay into quarks, the so-called lepton+jets decay channels, where one of the W bosons decays leptonically and the other hadronically, and the dileptonic decay channels with both W bosons decaying leptonically. The probabilities for the different decay channels according to the latest measurements of W-boson decays are given in table 4.1.

The advantage of measurements in the fully hadronic decay channel due to the large branching ratio of 45.7 % is compensated by the overwhelming background in hadron colliders because the final state exclusively consists of jets. Similar final states are produced in many different processes involving the strong interaction subsumed under the label 'QCD multijet production'. Furthermore, a discrimination between t and $\bar{t}$ is experimentally hardly possible in this channel.

The same flavour and different flavour dileptonic decay channels represent the other extreme each having branching ratios of about 1 % and 2 %, respectively. They have the smallest branching ratio but are extremely clean, i. e. SM processes with a confusingly similar final state are rare. The final state consists of two jets originating from b quarks and two isolated leptons, which allow for tagging the event and identifying t and $\bar{t}$

		W^+ boson			
		$u\bar{d}$, $c\bar{s}$	$e^+\nu_e$	$\mu^+\nu_\mu$	$\tau^+\nu_\tau$
W^- boson	$\bar{u}d$, $\bar{c}s$	45.7 %	7.27 %	7.15 %	7.61 %
	$e^-\bar{\nu}_e$	7.27 %	1.16 %	1.14 %	1.21 %
	$\mu^-\bar{\nu}_\mu$	7.15 %	1.14 %	1.12 %	1.19 %
	$\tau^-\bar{\nu}_\tau$	7.61 %	1.21 %	1.19 %	1.27 %

Table 4.1: Decay channel probabilities of top-antitop-quark pairs according to the W branching ratios quoted in reference [12].

via the charge. In addition, there are two neutrinos escaping the detector undetected. The resulting so-called 'missing transverse energy' (MET) helps to tag the event, but it complicates the proper reconstruction of the top quarks because only a combined MET for both neutrinos can be determined.

The lepton+jets decay channels with an electron or a muon in the final state combine a relatively large branching ratio of about 15 % each and a signature distinguishable from background processes due to the lepton. Moreover, the hadronic branch does not involve a neutrino from the W-boson decay and thus can be used for a proper reconstruction of the top quark. The tau+jets decay channel is an exception because the tau decays after a mean lifetime of about $2.9 \cdot 10^{-13}$ s. The rapid decay makes them difficult to identify. In about $\frac{2}{3}$ of the cases, it decays hadronically and the decays into the lighter leptons have a branching ratio of about $\frac{1}{6}$ each.

Therefore, the lepton+jets decay channel with a muon in the final state as illustrated in figure 4.4 is chosen in this thesis. The muon is preferred to the electron because it is very easy to identify a muon taking advantage of the muon system of the CMS detector.

4.4 Top Quarks as a Test of the Standard Model

In the past, many precise measurements of SM parameters have been performed especially in the electroweak sector. A fit to these high precision measurements shows a remarkably good agreement with SM predictions provided the Higgs boson exists. [35]

Nevertheless, some observations in nature cannot be explained within the framework of the SM. One of the most popular examples is the so-called 'Dark Matter'. There is a strong evidence that the known luminous matter constitutes only a small part of the matter in the universe [36]. For instance, the measured velocity of stars as a function of their distance to the centre of the galaxy is different from the expectation as obtained from the luminous matter in the galaxy. This discrepancy indicates the existence of a large amount of non-luminous matter, the Dark Matter, which is not predicted by the SM.

Observations of this kind initiate the development of new theories beyond the SM. In many cases, these theories imply new heavy particles that are expected to be produced at the LHC. $t\bar{t}$ decays often are an important background in searches for these particles due to their similar decay signature involving leptons, jets and MET. Moreover, particles decaying into a $t\bar{t}$ pair directly become evident in certain kinematic distributions of $t\bar{t}$ pairs as described in the following.

- **Invariant $t\bar{t}$ mass:** The most striking distribution is the invariant mass of the $t\bar{t}$ pair. A resonance in this distribution directly yields the mass of the decaying particle. Up to now, no significant deviation from SM expectations has been found. Therefore, model independent limits are set on the production of narrow resonances [37]. Broad resonances would also appear in the invariant-mass distribution as a deviation from SM expectations, but for this kind of resonances, the charge asymmetry is more sensitive.

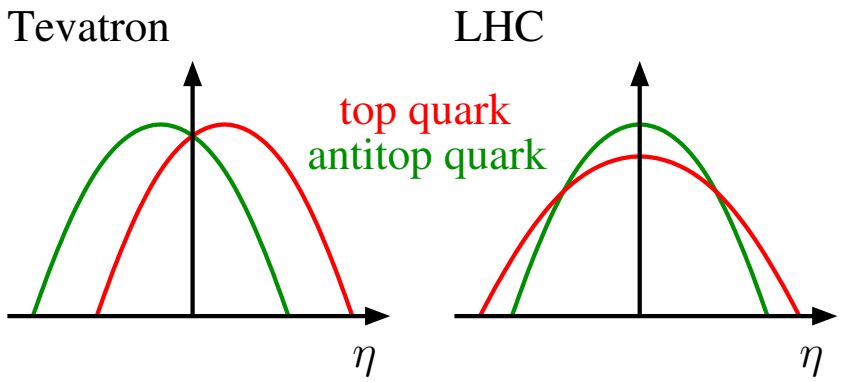

Figure 4.5: Illustration of the charge asymmetry in the pseudorapidity distribution of $t\bar{t}$ pairs produced at the Tevatron (left) and at the LHC (right).

- **Charge asymmetry:** For the $t\bar{t}$ production, the SM predicts a charge asymmetry in the pseudorapidity distribution [38] as illustrated in figure 4.5. The asymmetry arises from higher order QCD corrections to the quark-antiquark annihilation, which induce a correlation in the momentum between incoming quark and top quark and between incoming antiquark and antitop quark.

  At the Tevatron, which is a proton-antiproton collider, this leads to a forward-backward asymmetry, where the top quarks are on average boosted in the direction of the proton and the antitop quarks in the direction of the antiproton. This asymmetry is measured by two experiments at the Tevatron and independently found to be significantly higher than predicted by the SM [39, 40].

  Due to the symmetric proton-proton initial state at the LHC, there is no forward-backward asymmetry in the $t\bar{t}$ production. At the LHC, the charge asymmetry manifests in an asymmetry in the centrality arising from the fact that in protons, antiquarks only occur as sea quarks, whereas quarks also exist as valence quarks carrying a larger momentum fraction on average. Therefore, top quarks experience a stronger boost along the beam direction than the antitop quarks and thus have a broader angular distribution. In contrast to the Tevatron experiments, the experiments at the LHC do not measure a significant deviation from SM predictions [41, 42]. However, the predicted charge asymmetry is very small because the charge symmetric gluon-gluon fusion is the main production mechanism, and the measurements still reveal uncertainties being larger than the predicted asymmetry itself. In the near future, the uncertainties are expected to decrease allowing for more conclusive answers.

5 Event Simulation

A very important aim of experimental physics is to improve or reject theoretical models in order to test our understanding of the underlying physics.

Monte-Carlo (MC) generators make an essential tool for this in high energy physics. They simulate pp collisions according to certain QCD models up to the final state as measured in the detector. If the simulation sufficiently describes basic measured distributions, the simulated events can be used to calculate reconstruction and selection efficiencies. In addition, it helps to estimate the amount of background events which remain in the selected data sample. Thus, not only the signal process but all processes which can significantly contribute to the selected number of events have to be simulated.

Another valuable tool is the fixed-order QCD calculation of cross sections based on perturbation theory. It includes the matrix element of the hard scattering, i. e. the main process, with all corrections up to the given order of the strong coupling constant α_s. However, no detailed simulation of the final state as measured in the detector to compare with data is possible, but it can be used to normalise the background predictions if they are not derived from data. Furthermore, it provides the reference value of the signal cross section to compare with the measured value.

5.1 Monte-Carlo Event Generation

Typically, event simulation is factorised into multiple steps. First, the cross section of the hard scattering process is calculated on matrix element level. Depending on the actual process and the MC generator, hard scattering processes with up to n particles in the final state ($2 \to n$) are calculated. In this analysis, two different generators, *MadGraph* [43] and *Pythia* [44], have been used for this task. *MadGraph* calculates up to $2 \to 9$ processes. Hence, for $t\bar{t}$ production and decay, at most three hard partons are generated in addition to the six decay products of the top and the anti-top quark. *Pythia*, in contrast, calculates the hard scattering processes without any additional partons.

Higher-order QCD effects coming from initial- or final-state radiation (ISR/FSR) are taken into account simulating parton showers. As this step is performed by *Pythia*, *MadGraph* has an interface to pass the generated hard process on to *Pythia*. However, the combination of matrix element calculation and parton showering necessitates a matching between the hard partons from the matrix element level and the partons from the parton shower to prevent double counting of partons due to radiation in overlapping phase space regions. This matching is done via the MLM algorithm [45] by introducing a phase space cutoff and discarding the event if the parton shower generates partons harder than this matching threshold. A subsequent hadronisation merges the partons to colourless objects

called hadrons. The output of the simulation before the hadronisation is referred to as 'parton level' and after the hadronisation as 'hadron level'.

In addition to the hard process, softer interactions of secondary partons happen inside the same pp collision. This, in combination with proton remnants ending up in the detector, is called 'underlying event'. Because perturbation theory is not applicable at this energy regime, a correct description of the underlying event is a challenging task and many different *Pythia* tunes have been studied in this context with the first LHC data in 2010 [46]. The tune D6T was chosen to be the default *Pythia* tune in this analysis.

In the last step of the event generation, all final-state particles are passed through a full simulation of the CMS detector including the trigger chain based on GEANT4 [47]. The output of the detector simulation is referred to as 'reconstruction level'. To ensure the comparability between simulated and measured events, the same reconstruction algorithms are used in both cases.

5.2 Simulated Processes

The processes considered in this analysis are listed in table 5.1 together with the corresponding matrix-element generator and the number of simulated events. The signal process ($t\bar{t}$ production), the production of single top quarks (s, t, and tW channel), and the production of one vector boson (W and Z) in association with hard final state partons is generated by *MadGraph*. $t\bar{t}$ events are generated with up to three additional

process	generator	N_{events}	σ [pb]	order	weight ($35.9\,\text{pb}^{-1}$)
$t\bar{t}$	*MadGraph*	1,306,182	158^{+23}_{-24}	NLO	$4.34 \cdot 10^{-3}$
single top (s)	*MadGraph*	494,967	$4.59^{+0.20}_{-0.18}$	NNLL	$1.08 \cdot 10^{-4}$
single top (t)	*MadGraph*	484,060	$64.6^{+3.4}_{-3.2}$	NLO	$1.55 \cdot 10^{-3}$
single top (tW)	*MadGraph*	494,961	10.6 ± 0.8	NLO	$7.69 \cdot 10^{-4}$
$W \to l\nu$	*MadGraph*	14,805,546	31300 ± 1600	NNLO	$7.59 \cdot 10^{-2}$
$Z/\gamma^* \to l^+l^-$	*MadGraph*	2,543,727	3050 ± 130	NNLO	$4.30 \cdot 10^{-2}$
QCD multijet	*Pythia*	29,504,866	~ 296600000	LO	$1.03 \cdot 10^{-1}$
diboson (WW)	*Pythia*	2,061,760	43 ± 1.5	NLO	$7.49 \cdot 10^{-4}$
diboson (WZ)	*Pythia*	2,194,752	18.2 ± 0.7	NLO	$2.98 \cdot 10^{-4}$
diboson (ZZ)	*Pythia*	2,113,368	5.9 ± 0.15	NLO	$1.00 \cdot 10^{-4}$

Table 5.1: Compilation of the simulated processes used in this analysis. The Monte-Carlo generators and the number of generated events are given. In addition, the corresponding theoretical cross sections σ and the order of α_s to which they are calculated are listed, and the weight which has to be applied to the simulated events in order to normalise them to the integrated luminosity of $35.9\,\text{pb}^{-1}$ is stated.

hard partons, whereas the other processes are generated with up to four additional hard partons. QCD-multijet production and diboson production (WW, WZ, or ZZ) are exclusively done with *Pythia*. The default underlying event tune D6T was not available for single-top, QCD-multijet, and diboson production so the tune Z2 was used instead for these processes. Although this leads to an inconsistency, its effect on the result of this analysis is assumed to be small and in any case covered by the systematic uncertainty on the background.

For plots that are shown in this thesis, the $t\bar{t}$ sample is often split into the muon+jets decay channel, which is labelled as '$t\bar{t}$ signal (μ prompt)', and all other decay channels '$t\bar{t}$ other'. The label 'μ prompt' distinguishes between the muon+jets channel and the tau+jets channel with a subsequent decay of the tau lepton into a muon and neutrinos, which is also considered as signal. For technical reasons, in plots shown at reconstruction level, the latter is included in $t\bar{t}$ other but in plots corrected to parton level, it is separated and labelled as '$t\bar{t}$ signal ($\tau \to \mu$)'. The three diboson samples and the three single-top samples are combined to one diboson and one single-top sample, respectively, for all plots.

5.3 Normalisation from Fixed-Order Calculations

The cross sections of the processes considered here, the order of α_s up to which they are computed using perturbative QCD, and the resultant event weights, which are applied to normalise the simulated events to an integrated luminosity of $35.9\,pb^{-1}$, are given in table 5.1. For $t\bar{t}$, single-top, and diboson production, the most precise calculations available at the time of writing are of next-to-leading order (NLO). The calculations are carried out with the parton-level event integrator MCFM [48]. For the single-top production in the s-channel, there is even an approximate next-to-next-to-leading order (NNLO) calculation available, which is achieved via next-to-next-to-leading-logarithm (NNLL) resummation [49]. The cross sections for W-boson production with subsequent decay into lepton and neutrino ($W \to l\nu$) and for Z-boson production decaying into two leptons ($Z/\gamma^* \to l^+l^-$) are computed with FEWZ [50], which provides full NNLO precision. For QCD-multijet production, the plain LO cross section given by *Pythia* is used.

Some of the processes are filtered during their generation concerning their final state on parton level whereas the related fixed-order calculation refers to the inclusive cross section. This is taken into account for the determination of the event weight by correcting for the efficiency of the corresponding filter. The events with single top quarks from the s- and t-channel production are filtered by *MadGraph* to have a lepton in the final state. The efficiency of this filter ($\epsilon = 0.324$) purely results from the branching ratio of the W boson. The cross section of the QCD-multijet production is too large to permit a timely production of the full phase space. Therefore, at least one muon with a transverse momentum of $p_\mathrm{T} \geq 15\,\mathrm{GeV}$, which in most cases evolves from a parton shower of a heavy quark, is required in the final state leading to an efficiency of $\epsilon = 0.0002855$.

6 Event Selection

The basis for a proper cross section determination is a clean sample of well reconstructed events. In addition to the trigger selecting events online, a number of offline cuts are needed, which are applied on data and on reconstruction level of MC. As a step towards the phase space that is relevant for the $t\bar{t}$ cross section in the muon+jets decay channel, an inclusive muon sample is prepared to investigate event topologies for different numbers of additional hard jets. The entire event selection follows a common reference selection recommended by the Top Physics Analysis Group *(TopPAG)* as described in reference [51].

6.1 Data Sample

All the data from proton-proton collisions at $\sqrt{s} = 7\,\text{TeV}$ recorded in 2010 with the CMS detector are considered in this analysis. The complete dataset is divided into three run periods listed in table 6.1. The split into Run A and Run B is due to a technical stop in

run period	run range	luminosity	trigger threshold	$\epsilon^{\text{data}}_{\text{trigger}}$ [%]
Run A	136035 - 144114	$3.2\,\text{pb}^{-1}$	9 GeV	87.7 ± 0.8
Run B1	146428 - 147116	$5.0\,\text{pb}^{-1}$	9 GeV	92.0 ± 0.5
Run B2	147196 - 149294	$27.7\,\text{pb}^{-1}$	15 GeV	92.8 ± 0.2
Sum		$35.9\,\text{pb}^{-1}$		92.2 ± 0.2

Table 6.1: List of run periods with associated run range and luminosity. In addition, the threshold of the applied single-muon HLT and the corresponding efficiency measured via the tag-and-probe method in data is given.

the beginning of September 2010 implying major changes in the data taking conditions, e. g. the introduction of the cascade algorithm in the trigger software. The split of Run B into B1 and B2 is due to smaller changes in the trigger configuration, nevertheless being of relevance for this analysis. Generally, the data taking at CMS is divided into runs with stable conditions for beam and detector and further into so-called 'lumi sections' defined to be 2^{18} beam orbits each, which corresponds to about 23 seconds. Based on the status of the individual detector components, the decision is made which lumi sections can be used for further studies. From these lumi sections, the well reconstructed and interesting events have to be selected.

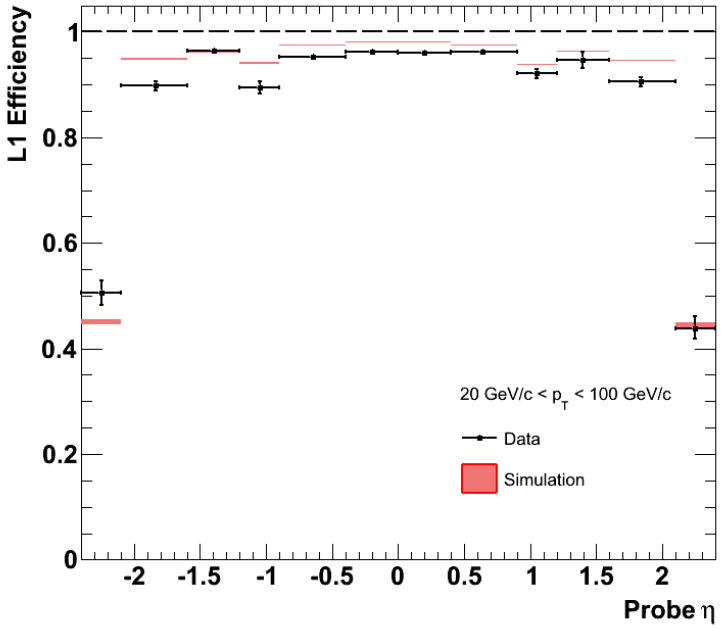

Figure 6.1: Efficiency of the Level-1 Single-muon trigger *L1_SingleMu7* as a function of the muon η. [52]

6.2 Trigger

In 2010, the luminosity was sufficiently low to obtain manageable event rates using unprescaled single-muon triggers without isolation requirement. As these triggers do not require any hadronic activity in the detector, they are ideally suited to collect the inclusive muon sample. High-level muon triggers usually take Level-1 trigger objects as seed for the reconstruction of the track in the muon system, which is then combined via a global fit with the corresponding track reconstructed in the tracker. Two high-level triggers, *HLT_Mu9* and *HLT_Mu15* are applied, which are both based on the Level-1 trigger *L1_SingleMu7*. The two HLTs have a different threshold for the transverse momentum of the muon. The threshold had to be raised from 9 to 15 GeV in the beginning of October due to the continuous raise of the instantaneous luminosity. As L1 single-muon triggers use the muon system, they have a geometrical acceptance of $|\eta| < 2.4$. But in the region $2.1 < |\eta| < 2.4$, the cabling of the trigger electronics leads to ambiguities in the read-out and hence compromises the L1 trigger efficiency for muons beyond $|\eta| = 2.1$ as shown in figure 6.1.

6.3 Data Cleaning

Two cuts are recommended for data cleaning. They both are applied only on data to remove non-collision events and events which can not be reconstructed correctly. One of them filters out beam scraping events by requesting 25 % of all tracks in the events to be of high quality. The second cut removes events with too much noise in the HCAL.

6.4 Vertex

The tracks for the primary vertex reconstruction [53] are selected based on the number of hits in the inner tracking system, the normalized track χ^2, and the transverse impact parameter. Tracks that are close in longitudinal direction ($z < 1\,\text{cm}$) at their point of closest approach to the beamline are clustered into groups to form the vertex candidates. Selected events are required to have at least one good primary vertex. A primary vertex is declared as 'good' if it lies within a longitudinal distance of $|z| < 24\,\text{cm}$ and a transverse distance of $\rho < 2\,\text{cm}$ with respect to the nominal interaction point and has at least four degrees of freedom (ndof > 4). The quantity ndof corresponds to the weighted sum of tracks. The weight of each track refers to its compatibility with the common vertex derived with an adaptive vertex fit [54].

6.5 Inclusive Muon Sample

One important reason for choosing the muon+jets decay channel for the determination of the $t\bar{t}$ cross section is the muon in the final state. Its identification is almost unambiguous and the reconstruction of its kinematics is very accurate. Thus, it provides the possibility to select a well defined kinematic region. Furthermore, the muon is valuable for the discrimination of certain backgrounds. The muon efficiencies, which are needed for the correction of the cross section later, are obtained in a data-driven way.

6.5.1 Reconstruction of Muons

Tracks from the inner tracking system (tracker tracks) and tracks from the muon system (standalone-muon tracks) are independently reconstructed and form the basis for the muon reconstruction [52]. There are two main reconstruction methods.

- **Tracker-muon reconstruction:** This method is an inside-out approach starting with the tracker tracks and extrapolating them to the muon system including energy-loss and multiple-scattering effects. As it only requires one muon segment to be matched to the tracker track, it is very efficient also for low momentum muons. The muons reconstructed with this method are called tracker muons.

- **Global-muon reconstruction:** In contrast, a global muon is reconstructed in an outside-in approach matching standalone-muon tracks to tracker tracks. A global fit is combining the hits of both tracks to the track of the global muon. Due to the large lever arm of the muon system, the momentum resolution is improved in comparison to the tracker muons especially for large momenta ($p_T \geq 200\,\text{GeV}$).

If both previous methods fail, an alternative method is the usage of the standalone-muon track as a standalone muon. But this is only used for cosmic-ray muons, which often do not pass the inner tracking system.

Two different types of muons are defined in the following, loose and tight muons. For both, only muons that can be reconstructed as a global muon are considered to benefit

from the low fake rate of this method and from the improved resolution of the global fit. Loose muons do not need to fulfil any further criteria on the track quality whereas tight muons are additionally required to be reconstructed as a tracker muon from at least eleven hits with one of them in the pixel detector. Furthermore, the tight muons have to fulfil some basic quality criteria. Their global track has to contain hits from at least two muon stations and the normalized χ^2 of the global fit needs to be below ten. In addition, there is a requirement to the longitudinal distance between the muon vertex and the primary vertex of $\Delta z < 1\,\mathrm{cm}$.

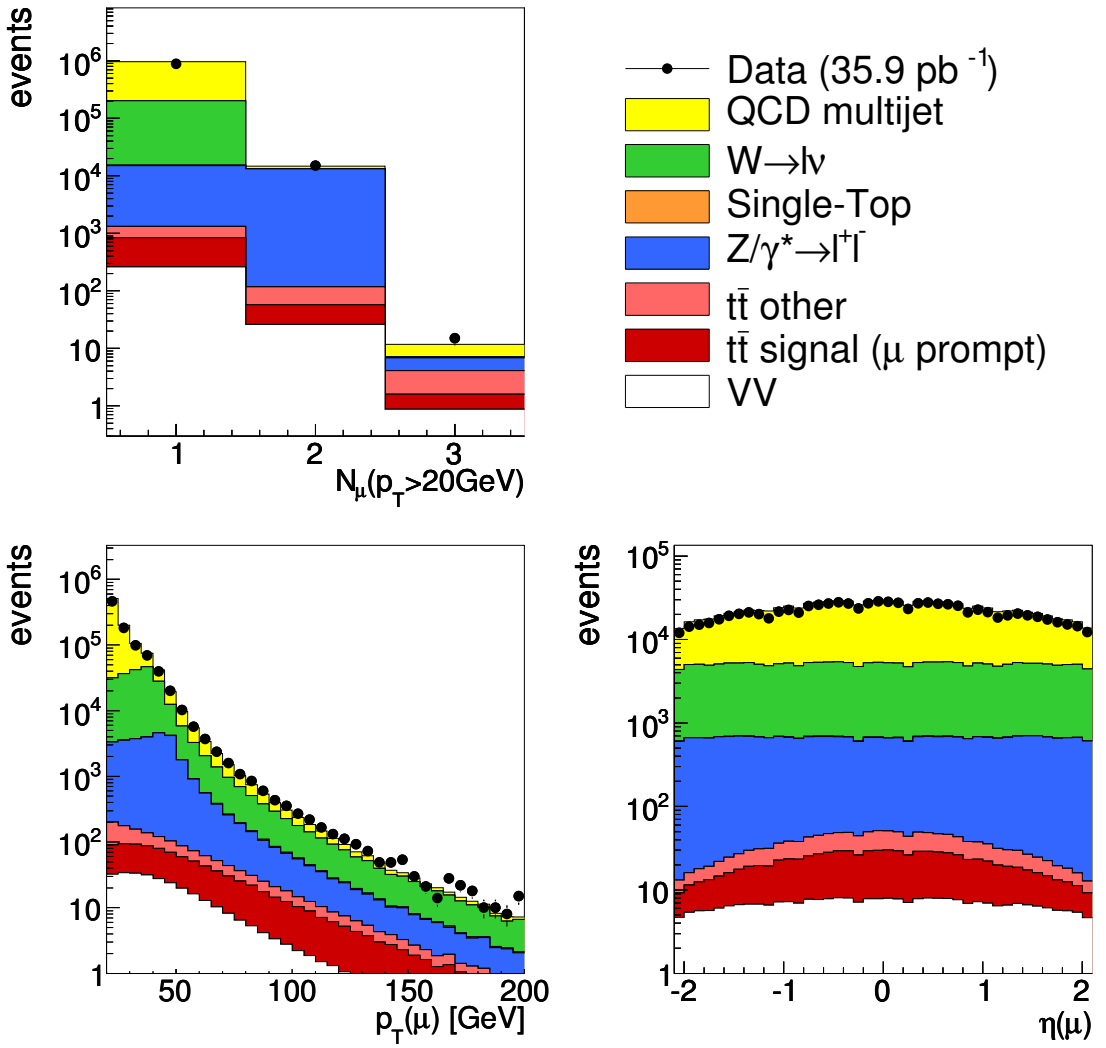

Figure 6.2: Muons in the analysed phase space. The upper left plot shows the multiplicity and the distributions of p_T and η of the leading muon are shown in the lower left and right plot, respectively. Black points denote data, coloured areas different simulated processes.

6.5.2 Muon Kinematics

Muons that are relevant for this analysis are defined by the following kinematic region. Only tight muons are considered.

- **Pseudorapidity:** The efficiency of the muon trigger is steeply falling for very forward muons, so the muons are restricted to the central region of the detector,

$$|\eta(\mu)| < 2.1. \tag{6.1}$$

  Due to the large mass of the top quarks, they are mostly produced nearly at rest and do not have a large boost in z direction. Hence, a large fraction of their decay products traverse the central part of the detector.

- **Transverse momentum:** As the W boson on the one hand and its decay products, muon and muon neutrino, on the other hand have a large mass difference, central muons tend to have a large transverse momentum on average. Thus, cutting on

$$p_T(\mu) > 20\,\text{GeV} \tag{6.2}$$

  does not reject too many signal events. Furthermore, this cut ensures high reconstruction and trigger efficiencies.

For selected events, the distributions of kinematic variables for the highest-p_T muon (leading muon) that fulfils all requirements mentioned so far, are shown in figure 6.2. Data and MC prediction are in a good agreement. Muons from $t\bar{t}$ decays seem to have a harder p_T spectrum and they are more central than other muons. Especially, QCD multijet events seem to have primarily low-p_T muon candidates and muons from $W \to l\nu$ and $Z/\gamma^* \to l^+l^-$ events are distributed flat in η. The majority of $t\bar{t}$ events in the muon+jets decay channel has only one of the considered muons. In contrast, $Z/\gamma^* \to l^+l^-$ events often have two muons in this kinematic region.

6.5.3 Background Reduction

For the selection of the inclusive muon sample, those events are considered as background that do not contain exactly one muon coming from the decay of a W boson. Thus, at this stage, the background processes can be categorised into two groups, namely on the one hand, the fully hadronic events with no muon from a Vectorboson decay such as QCD-multijet events and fully hadronically decaying $t\bar{t}$ events, and on the other hand, the dileptonic events such as $Z/\gamma^* \to l^+l^-$ events and dileptonic $t\bar{t}$ events.

- **fully hadronic events:** Muons fulfilling the tight criteria mentioned in section 6.5.1 that are evolving in fully hadronic events, originate in most cases from the decay of hadrons containing heavy quarks such as D and B mesons because lighter hadrons mostly escape the tracker before they decay. Since the hadrons are contained in jets, the same is true for the muons from their decay. Thus, two cuts are used to reject these muons selecting only isolated ones.

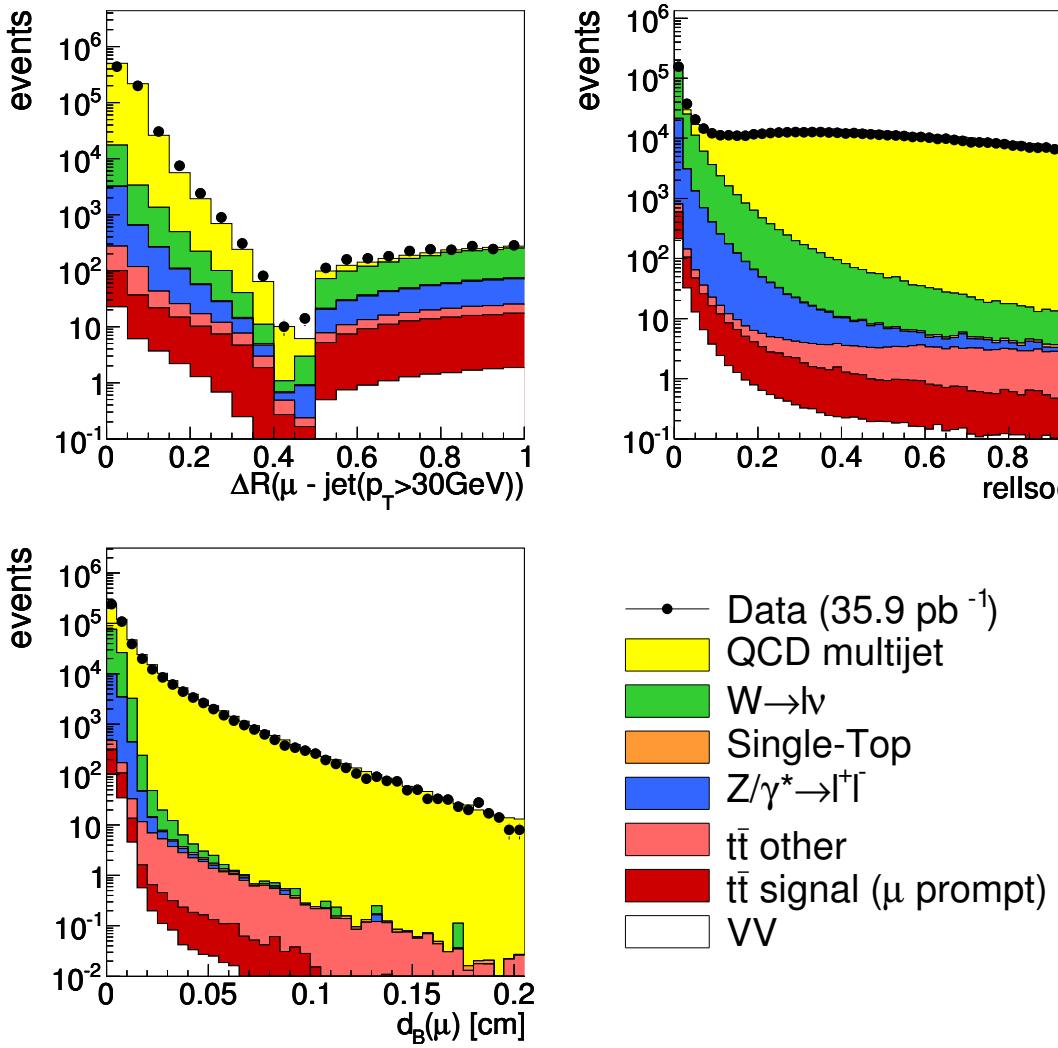

Figure 6.3: Distributions of cut variables for the leading muon in the analysed phase space. The upper plots show the isolation variables, $\Delta R(\mu - \text{jet})$ (left) and relIso (right) and the lower left plot shows the transverse impact parameter. Black points denote data, coloured areas different simulated processes.

The first cut rejects muons which are in the vicinity of a jet, where the distance $\Delta R(\mu, jet)$ between the muon and a jet is defined in the (η, ϕ) plane and has to be larger than

$$\Delta R(\mu, jet) \equiv \sqrt{\Delta\eta^2 + \Delta\phi^2} > 0.3 \tag{6.3}$$

with respect to all jets fulfilling the criteria defined in section 6.6.

For the second cut, variables indicating the isolation of the muon in the tracker (trkIso) and in the calorimeter (calIso) are introduced. trkIso is defined as the sum of track momenta in a cone of $\Delta R < 0.3$ around the muon excluding the muon momentum itself. calIso is defined as the sum of energy depositions in this cone

excluding the calorimeter energy in a smaller cone of $\Delta R < 0.007$. The combined relative isolation, relIso, is defined as the sum of trkIso and calIso divided by the transverse momentum of the muon. Muons with

$$\text{relIso} \equiv \frac{\text{trkIso} + \text{calIso}}{p_T(\mu)} < 0.05 \tag{6.4}$$

are considered as isolated.

A third cut takes advantage of the distance of flight before the decay, which is for heavy hadrons typically a few $100\,\mu$m. Depending on the angle of emission, this often leads to large transverse impact parameters d_B with respect to the beam line. Hence, cutting on

$$d_B < 0.02\,\text{cm} \tag{6.5}$$

further reduces fully hadronic background events.

The background-reduction capability of these three cuts can be seen from the corresponding plots in figure 6.3. They show the distributions of these variables for muons after the kinematic cuts are applied. The vast majority of QCD-multijet events can be rejected by these cuts. In the same manner, these cuts are appropriate to reject fully hadronically decaying $t\bar{t}$ events, which are a subset of the other $t\bar{t}$ events shown in these plots.

- **dileptonic events:** The dileptonic events are suppressed by applying vetoes against electrons and additional muons. In order to make these lepton vetoes efficient, very loose definitions for muons and electrons are introduced. The track criteria of the loose muons are defined in section 6.5.1 and the electrons are reconstructed using tracks reconstructed in the tracker and clusters of energy in the ECAL [55]. Muons as well as electrons are considered for the lepton veto if they are loosely isolated (relIso < 0.2).

  Events are allowed to have one muon but no electron in the acceptance region of the inner tracking system ($|\eta| < 2.5$) with a transverse momentum above the lepton dependent threshold.

$$N_\mu(p_T > 10\,\text{GeV}) = 1 \tag{6.6}$$

$$N_e(p_T > 15\,\text{GeV}) = 0 \tag{6.7}$$

  The multiplicities of the leptons are shown in figure 6.4. The veto-lepton multiplicities are shown for events containing exactly one tight muon. Especially, the second-muon veto reduces a large fraction of $Z/\gamma^* \to l^+l^-$events. In addition, large parts of the other $t\bar{t}$ events are removed by the lepton vetoes. These events are mostly $t\bar{t}$ events with two muons or one electron and one muon in the final state.

This selection provides a sample enriched with events containing a single well reconstructed and isolated muon mainly coming directly from the decay of a W boson. Events

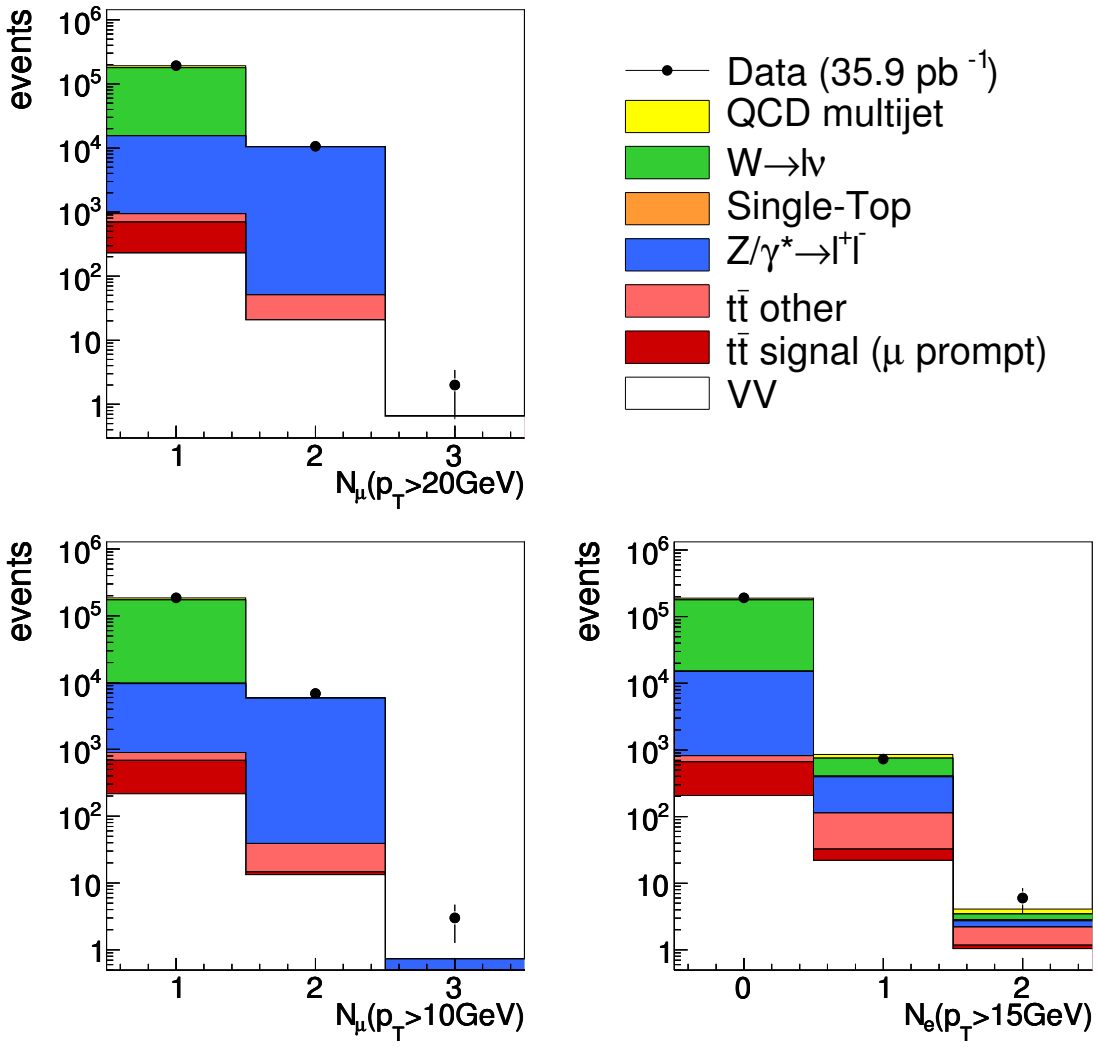

Figure 6.4: Multiplicity of leptons after the full muon selection is applied. In the upper left plot, the tight muons are shown. In the lower plots the loose leptons used for the vetoes, muons (left) and electrons (right), are shown for events containing exactly one tight muon. Black points denote data, coloured areas different simulated processes.

with a W boson decaying into a tau lepton with a subsequent decay into a muon and neutrinos have a very similar topology and hence will be considered as signal in the following. However, as mentioned in the previous chapter, they are treated separately in the simulation.

6.5.4 Muon Efficiencies

There are many steps in the compilation of the muon sample, which are not fully efficient. The main sources of inefficiency are the trigger, the reconstruction, the identification,

and the isolation of the muon. A correction for these inefficiencies is mandatory as they lower the measured event yields, which are used for the cross section determination. The overall efficiency, which is used for the correction in the end, is calculated from simulation as the total number of reconstructed events that are selected divided by the total number of events that are generated in the proper phase space,

$$\epsilon = \frac{N_{\text{reco}}^{\text{selected}}}{N_{\text{gen}}^{\text{phase space}}}. \tag{6.8}$$

Efficiencies can be factorised corresponding to the different sources of inefficiency involved if they are independent of each other. For the muon efficiency, this results in

$$\epsilon_\mu = \epsilon_{\text{reco}} \cdot \epsilon_{\text{ID [reco]}} \cdot \epsilon_{\text{iso [ID*reco]}} \cdot \epsilon_{\text{trig [iso*ID*reco]}}, \tag{6.9}$$

where each factor is the efficiency for the criterion given in the index relative to a preselection given in the brackets so each step refers to all previous steps.

- $\boldsymbol{\epsilon_{\text{reco}}}$**:** efficiency for reconstructing a muon as a global and as a tracker muon.
- $\boldsymbol{\epsilon_{\text{ID}}}$**:** efficiency of all tight muon criteria besides reconstruction and isolation.
- $\boldsymbol{\epsilon_{\text{iso}}}$**:** efficiency for the isolation, ΔR and relIso.
- $\boldsymbol{\epsilon_{\text{trig}}}$**:** efficiency of the applied muon trigger.

A priori, it is not clear if all parts are modelled correctly by the simulation. The reconstruction efficiency was studied [56] and found to be described sufficiently precise by the simulation, i. e. the ratio between the efficiency determined from data and from simulation is consistent with unity. For all other efficiencies, a tag-and-probe (T&P) method is used to check the accuracy in a data-driven way. The two muons of $Z \to \mu^+\mu^-$ events are very well suited as input for this method. One of them, which has to fulfil all tight muon criteria, is used to tag the event. In the tagged events, muons are selected that survive the preselection, which is indicated in brackets in eq. (6.9). Out of these muons, a muon is accepted as probe muon if the invariant mass of the tag and the probe muon is reconstructed in a window around the Z boson mass of $(91 \pm 15)\,\text{GeV}$. The fraction of probe muons that fulfil the test criterion is then the efficiency of the test criterion relative to the preselection.

The efficiencies obtained for simulation and data are listed in table 6.2. The trigger efficiencies are separately determined for the different run periods (see tab. 6.1) and combined afterwards. At the end all efficiencies are combined to an overall muon efficiency. In addition, the scale factors, which need to be applied to the simulated events to correct for the difference between data and Monte Carlo, are given in this table. A scale factor that is close to unity is an evidence of an accurate simulation. This is the case for all scale factors beside the one for the trigger efficiency. Thus, some trigger inefficiencies seem to be underestimated in the simulation.

source	$\epsilon^{\text{sim}}_{\text{T\&P}}$ [%]	$\epsilon^{\text{data}}_{\text{T\&P}}$ [%]	scale factor
trigger	95.21 ± 0.03	92.2 ± 0.2	0.969 ± 0.002
identification	98.36 ± 0.02	97.7 ± 0.1	0.993 ± 0.001
isolation	89.88 ± 0.04	90.0 ± 0.2	1.002 ± 0.002
combination	84.17 ± 0.05	81.1 ± 0.3	0.964 ± 0.003

Table 6.2: Muon efficiencies and scale factors for different selections in simulation and data obtained from the T&P method using $Z \to \mu^+\mu^-$ events. The quoted uncertainties are statistical only.

6.6 Muon+Jets Samples

As the inclusive muon sample still includes all background processes with a W boson decaying into muon and muon neutrino, further measures are required to suppress them. A particular characteristic of $t\bar{t}$ events in the muon+jets channel compared to the remaining background processes is the relatively large number of hard jets evolving from the quarks in the final state. By investigating the event topologies of processes containing exactly one muon in the final state in dependence of the jet multiplicity, the transition from background to signal events can be studied.

6.6.1 Reconstruction of Jets

In inelastic proton-proton collisions, the density of particles spread into the detector is very high. Many of these particles originate from a parton, quark or gluon, which immediately forms a jet due to the confinement of QCD. Jets are narrow bunches of almost collinear particles produced by radiation, pair production, and decay. First, all particles in the event are reconstructed taking advantage of the full particle-identification capacity of the CMS detector. Subsequently, the reconstructed particles are fed into a clustering algorithm to form the jets, which help to trace back to the underlying partons.

- **Particle reconstruction:** The aim is to reconstruct and identify all stable particles in the event. The particle flow algorithm [57], which is used for this task, attempts to distinguish between muons, electrons, photons, charged hadrons, and neutral hadrons. The CMS detector with its silicon tracker in the strong magnetic field and the large muon system is well suited for the reconstruction and identification of muons. The high granularity and the brilliant energy resolution of the ECAL allow an accurate reconstruction of electrons and photons. The latter two are distinguished by the tracker track, which is only present for electrons. The HCAL, which is responsible for the hadron measurement, provides a more coarse resolution, which does not allow for separating charged and neutral hadrons. Thus, the information from HCAL and tracker are combined to identify charged hadrons. A remaining energy excess in the HCAL after subtracting the charged component is considered as neutral hadrons.

- **Jet clustering algorithm:** The complete list of reconstructed particles is given to the anti-k_t algorithm [58] to cluster them to jets. This algorithm is infrared and collinear safe and fast enough to cope with the high particle multiplicity per event in a timely manner. Infrared and collinear safety mean the invariance of the final result under emission of very soft and collinear particles, respectively. From the theoretical point of view, this is needed because the probability of radiating a gluon increases with falling energy of the radiated gluon and decreasing angular distance between quark and gluon leading to IR divergences for very low energetic or collinear gluons. These divergences are of course not physical, but they refer to the breakdown of the perturbative approach in QCD. Experimentally, it renders the separation of close-by produced particles and the efficient reconstruction of very low energetic particles unnecessary.

  As a first step, the anti-k_t algorithm calculates a weighted distance between all possible pairs of particles (i, j) in the list,

$$d_{i,j} = \min\left(\frac{1}{k_{T,i}^2}, \frac{1}{k_{T,j}^2}\right) \cdot [(y_i - y_j)^2 + (\phi_i - \phi_j)^2], \tag{6.10}$$

  where k_T is the transverse momentum, y the rapidity, and ϕ the azimuth of the corresponding particle. Starting with the smallest weighted distance, all pairs are merged to a new particle until the smallest $d_{i,j}$ is larger than the maximum allowed weighted distance d_i

$$\min(d_{i,j}) > d_i = \left(\frac{1}{k_{T,i}^2}\right) \cdot R_0^2, \tag{6.11}$$

  where the parameter R_0 is set to 0.5 in this analysis. Then particle i is added to the final list of jets and not further taken into account. R_0 can be interpreted as the opening angle of the jet cones. This procedure is repeated until all particles are clustered to jets.

The jets are supposed to carry roughly the same four-momentum as the underlying parton. However, due to instrumental effects, large angle radiation, pile up, and neutrinos, which can evolve in jets and escape undetected, the measured energy of the jets has to be corrected. At CMS, this is done in a factorised approach, where the η and p_T dependent correction factors are derived from simulation [59]. These Corrections also include an offset correction subtracting energy caused by pile-up interactions, i. e. additional generally softer interactions occurring in the same bunch crossing as the interaction under study. Small differences between data and MC are resolved by residual corrections applied to data. The residual corrections are basically derived in two steps. First, events are selected containing at least two jets. This selection provides enough statistics to correct the response relative in η by balancing the p_T of the two jets. In the second step, a sample is chosen, which contains a jet and a well measured photon as the reference object for the p_T-balancing. Photons have a much better energy resolution than jets due to the excellent ECAL. Therefore, this sample is used to fix the absolute energy scale.

6.6.2 Jet Selection

In order to select hadronic jets avoiding fake jets from instrumental noise, the jets have to fulfil a few loose requirements concerning their composition. They need to consist of at least two particles, out of which at least one particle is charged. Furthermore, the jets are required to have a non-vanishing fraction of charged hadronic energy, whereas the neutral hadronic and the charged and neutral electromagnetic energy fraction each have to be below 99 %. The requirements are found to retain more than 99 % of the genuine jets. Jets that are closer to a selected muon than 0.1 in the (η, ϕ) plane are regarded as not being a genuine jet but being mainly induced by the muon. These jets are removed from the jet collection.

The allowed kinematic region for these jets is defined as follows.

- **Pseudorapidity:** Jets are restricted to the central region of the detector,

$$|\eta(\text{jet})| < 2.4, \tag{6.12}$$

  to ensure that the jet cones are completely contained within the ECAL and HCAL.

- **Transverse momentum:** Due to large uncertainties for low jet energy values, hard jets with

$$p_T(\text{jet}) > 30\,\text{GeV} \tag{6.13}$$

  are required to enable a reliable jet energy reconstruction.

6.6.3 Jet Multiplicity

Remaining background events for the top-quark pair production in the muon sample are mainly due to the direct production of W bosons. As the event topology of the production of single top quarks is very similar, they are also considered here. The final state of $t\bar{t}$ events in the muon+jets decay channel comprises four jets from the hard process, out of which two are formed by the b quarks from the decay of the top quarks and two are formed by the light quarks from the hadronically decaying W boson. As the W bosons from single top events passing the muon selection are mostly decaying leptonically, the only jet evolving in the decay of these single top quarks is coming from the b quark. However, there is often an additional jet arising from the light quark involved in the production process because it can be scattered into the acceptance region of the detector. In both cases, further jets can emerge from QCD radiation, which is the only production mechanism for jets in $W \to l\nu$ events. Each QCD radiation increases the order of α_s in the process by one, which reduces the probability of the process to happen. This leads to the difference in jet multiplicity for the different processes shown in figure 6.5 with the signal process peaking at the highest value compared to the other processes. Jet multiplicities smaller than the one expected from the number of partons produced in hard interaction, e. g. $N_{\text{jets}} < 4$ for $t\bar{t}$ signal events, are possible due to reconstruction and selection inefficiencies.

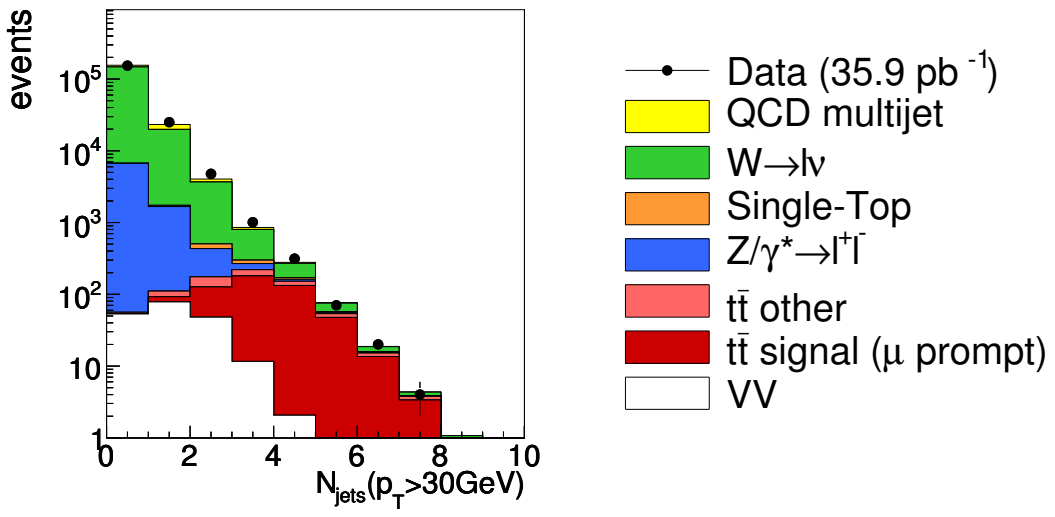

Figure 6.5: Multiplicity of selected jets after the full muon selection. Black points denote data, coloured areas different simulated processes.

Four samples are compiled with exactly one isolated muon and at least one, two, three, and four selected jets, respectively.

6.7 Selected Events

The event yields measured in 35.9 pb^{-1} of data, the simulated event yields scaled to the same integrated luminosity, and the event composition expected from the simulation for different inclusive jet multiplicities are shown in table 6.3. The measured number of events ranges from 184 845 events in the inclusive muon sample, where no jet require-

		muon sample	≥ 1 Jet	≥ 2 Jets	≥ 3 Jets	≥ 4 Jets
	N(data)	184 845	31 145	6 195	1 419	408
	N(simulation)	178 100 $\pm$ 100	27 590 $\pm$ 50	5 090 $\pm$ 20	1 188 $\pm$ 8	365 $\pm$ 4
simulation	$t\bar{t}$ signal	$< 1\,\%$	2 %	8 %	29 %	52 %
	$t\bar{t}$ other	$< 1\,\%$	$< 1\,\%$	2 %	5 %	7 %
	single top	$< 1\,\%$	1 %	2 %	4 %	3 %
	$W \to l\nu$	88 %	77 %	72 %	51 %	32 %
	$Z/\gamma^* \to l^+l^-$	5 %	7 %	6 %	5 %	3 %
	QCD multijet	6 %	13 %	9 %	5 %	2 %
	diboson	$< 1\,\%$	$< 1\,\%$	1 %	1 %	1 %

Table 6.3: Event yields in data and simulation scaled to 35.9 pb^{-1} and the expected composition for different selection steps. The combined scale factor as determined in section 6.5.4 has been applied to the numbers derived from simulation.

ments are imposed, to 408 events in the inclusive four jet sample. The simulated event yields are systematically below the measured numbers. The event fractions given in the table are derived from the simulation without those correction factors that are determined in the next chapter to cure the underestimation in simulation. Nevertheless, they are supposed to give a good estimate on the real event composition. The inclusive muon sample is with 88 % completely dominated by the process $W \to l\nu$. With increasing number of jets this dominance becomes weaker. Finally, in the inclusive four jet sample, the majority of events stems from $t\bar{t}$ signal events. However, even in the inclusive four jet sample, about one third of all events are $W \to l\nu$ events being produced in association with hard jets.

In order to validate the final selection, cut variables are shown for certain selection steps leaving out the cut on the variable itself. Small differences between data and MC in muon distributions are resolved by the muon efficiency correction presented in section 6.5.4.

Vertex and muon cut variables are shown after applying the full vertex and muon selection except for the cut on the respective variable. Figure 6.6 shows the vertex and muon quality variables. The z distribution of the primary vertex is described reasonably well and the cut ($|z(\mathrm{PV})| < 24\,\mathrm{cm}$) is set to about $\pm 4\sigma$ of a Gaussian fitted to the distribution giving a cut efficiency above 99.99 %. The ndof distribution of the primary vertex reveals an underestimation in the simulation at high values ($\mathrm{ndof(PV)} \gtrsim 150$). This could be related to the known difficulties in the simulation of the underlying event and the associated charged particle multiplicity [60]. However, this does not harm because the cut is set to a very low value ($\mathrm{ndof(PV)} > 4$), where the description is good.

The number of hits per track $N_{\mathrm{Hits\ trk}}$ of the muon is on average slightly overestimated by the simulation, i. e. the simulated distribution is slightly shifted to higher values, whereas the normalized χ^2 is on average slightly underestimated, i. e. the simulated distribution is slightly shifted to smaller values. The effect of these deviations is small because the cuts $N_{\mathrm{Hits\ trk}} \geq 11$ and $\chi^2/\mathrm{ndof} < 10$ both have an efficiency of nearly 99.9 % in data and MC.

The distance in z between primary vertex and muon vertex $\Delta z(\mu - PV)$ is only described sufficiently for distances smaller than 0.1 cm. At larger distances, the data has a distinct tail, which is not present in simulation. This small contribution is due to additional pile-up vertices, which are not simulated in the default MC. Simulations including pile up also feature this tail. The reason is that two vertices which are closer than 1 cm in z are merged to one vertex [53]. The resulting combined vertex often is close to the single vertex with the larger track multiplicity. Therefore, the primary vertex associated with the muon in some cases might be merged with a pile-up vertex. The combined vertex could than be up to 1 cm shifted with respect to the muon-related primary vertex. The efficiency of the cut ($\Delta z(\mu - PV) < 1\,\mathrm{cm}$) is above 99.9 % even in data. Therefore, the deviation does not have a significant influence.

Figure 6.7 shows the muon isolation and kinematic variables. The distance in the (η, ϕ) plane between the muon and the closest jet with $p_T > 30\,\mathrm{GeV}$ only starts at 0.1 because jets closer to the muon are removed from the jet collection. In principle, the distribution should rise because the area around the muon rises with the radius enhancing

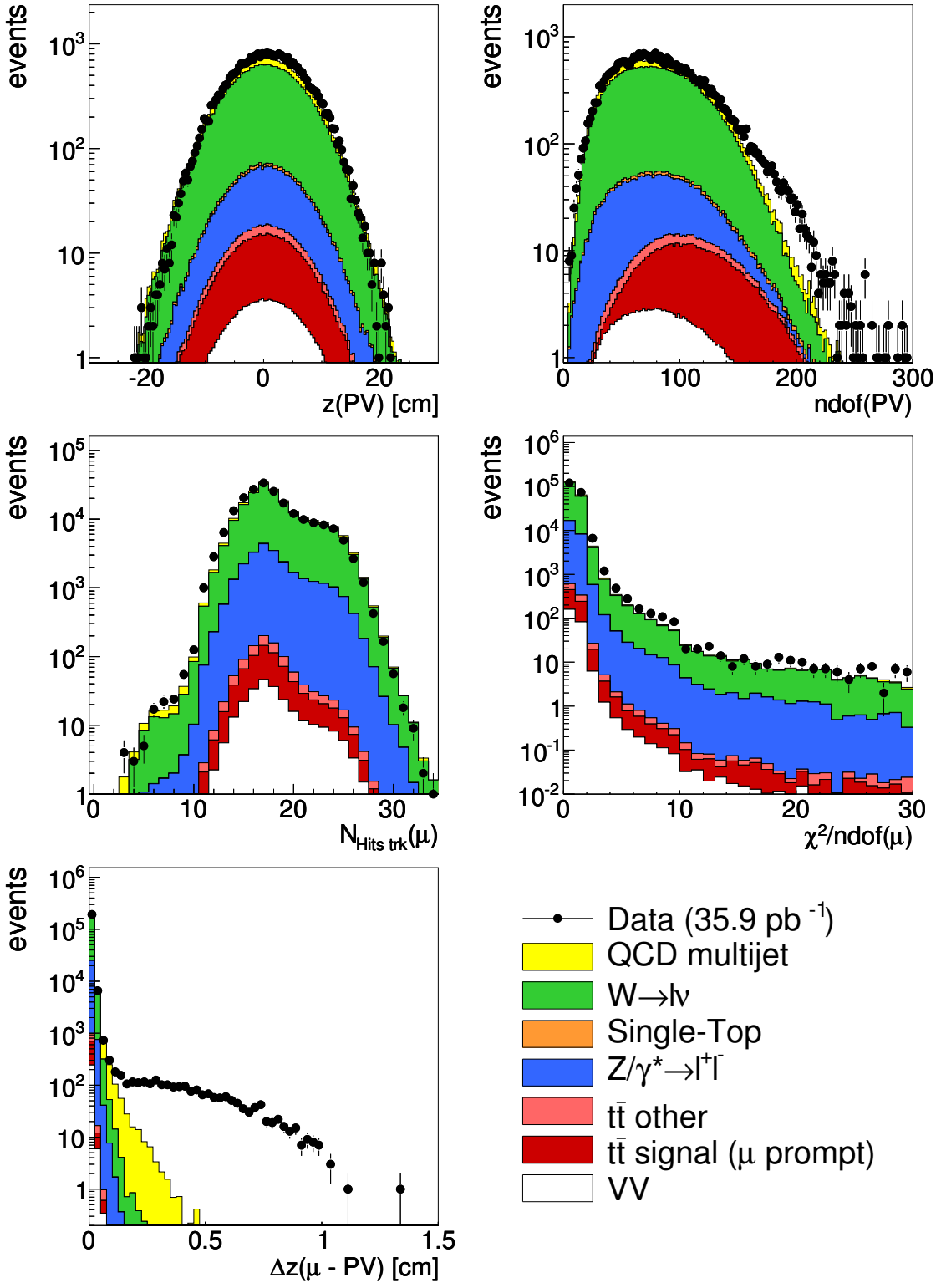

Figure 6.6: Vertex and muon quality variables after the full muon selection is applied. Black points denote data, coloured areas different simulated processes.

the probability to find a jet. This can be seen in the plot starting at $\Delta R = 0.5$, which is the cone size of the jets. Below this radius muons are clustered into the jet and therefore pull the jet towards the muon, which leads to the steeply falling distribution at smaller radii. The cut ($\Delta R(\mu - \text{jet}) > 0.3$) removes most of the biased jets. Meanwhile, it is technically possible to remove the muons from the jet clustering to completely avoid this small bias. The relative isolation and the transverse impact parameter of the muon are well simulated. The cuts (relIso < 0.05 and $d_B < 0.02\,\text{cm}$) are very well suited to remove the QCD-multijet background.

The simulated kinematics, p_T and η, of the muon show a good agreement with the measurement in the analysed phase space ($p_T > 20\,\text{GeV}$ and $|\eta| < 2.1$).

Figure 6.8 shows the transverse momentum of the four leading jets each before the p_T requirement on the respective jet is applied. In addition, the p_T of the first leading jet is shown in figure 6.9 after requiring at least two, three, or four jets, the p_T of the second leading jet is plotted in figure 6.10 after requiring at least three or four jets, and figure 6.11 shows the p_T of the third leading jet after requiring at least four jets. All jet-p_T distributions are reasonably well simulated, albeit some differences are visible, which are related to the underestimation in simulation addressed in section 7.2.

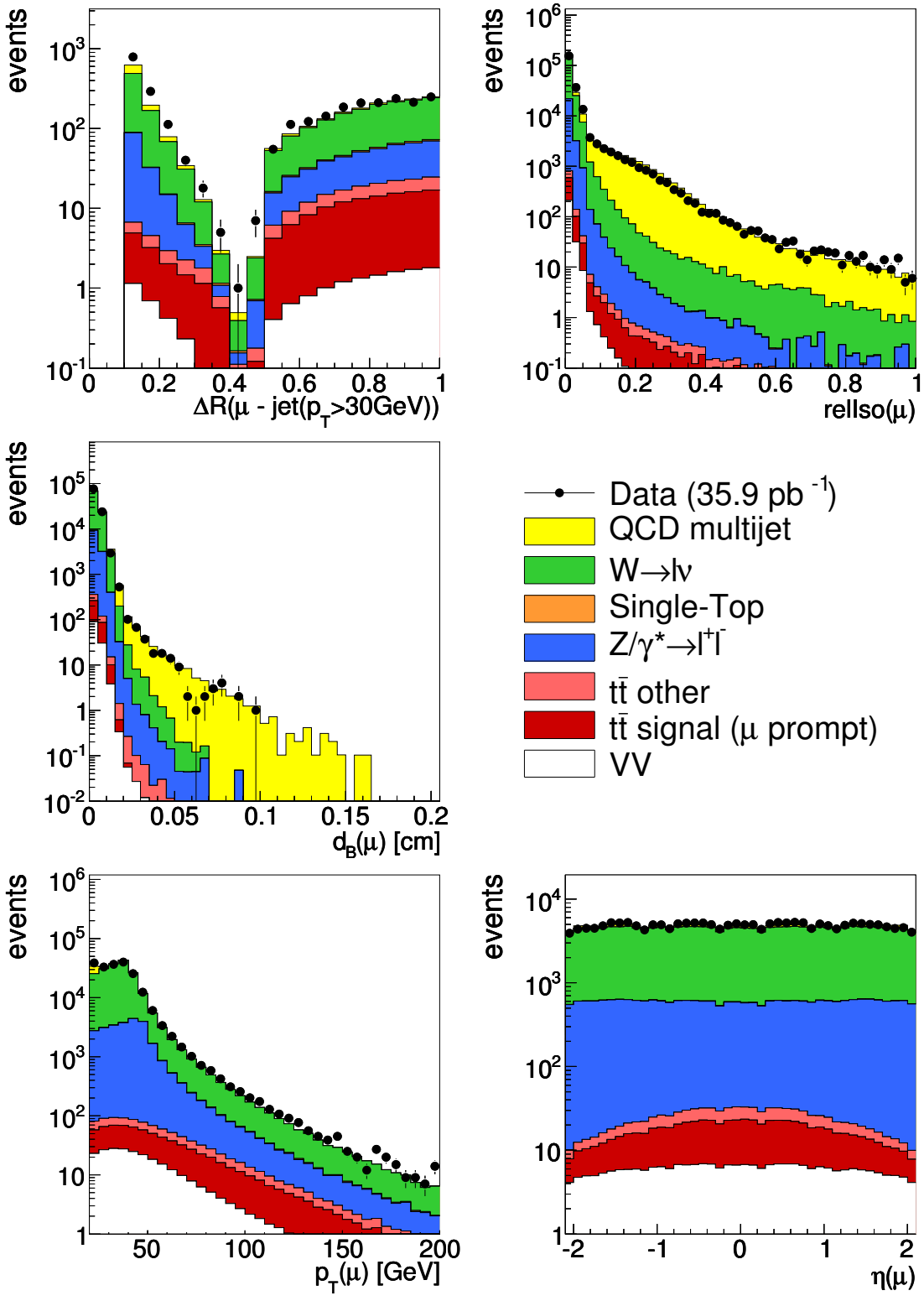

Figure 6.7: Muon isolation and kinematic variables after the full muon selection is applied. Black points denote data, coloured areas different simulated processes.

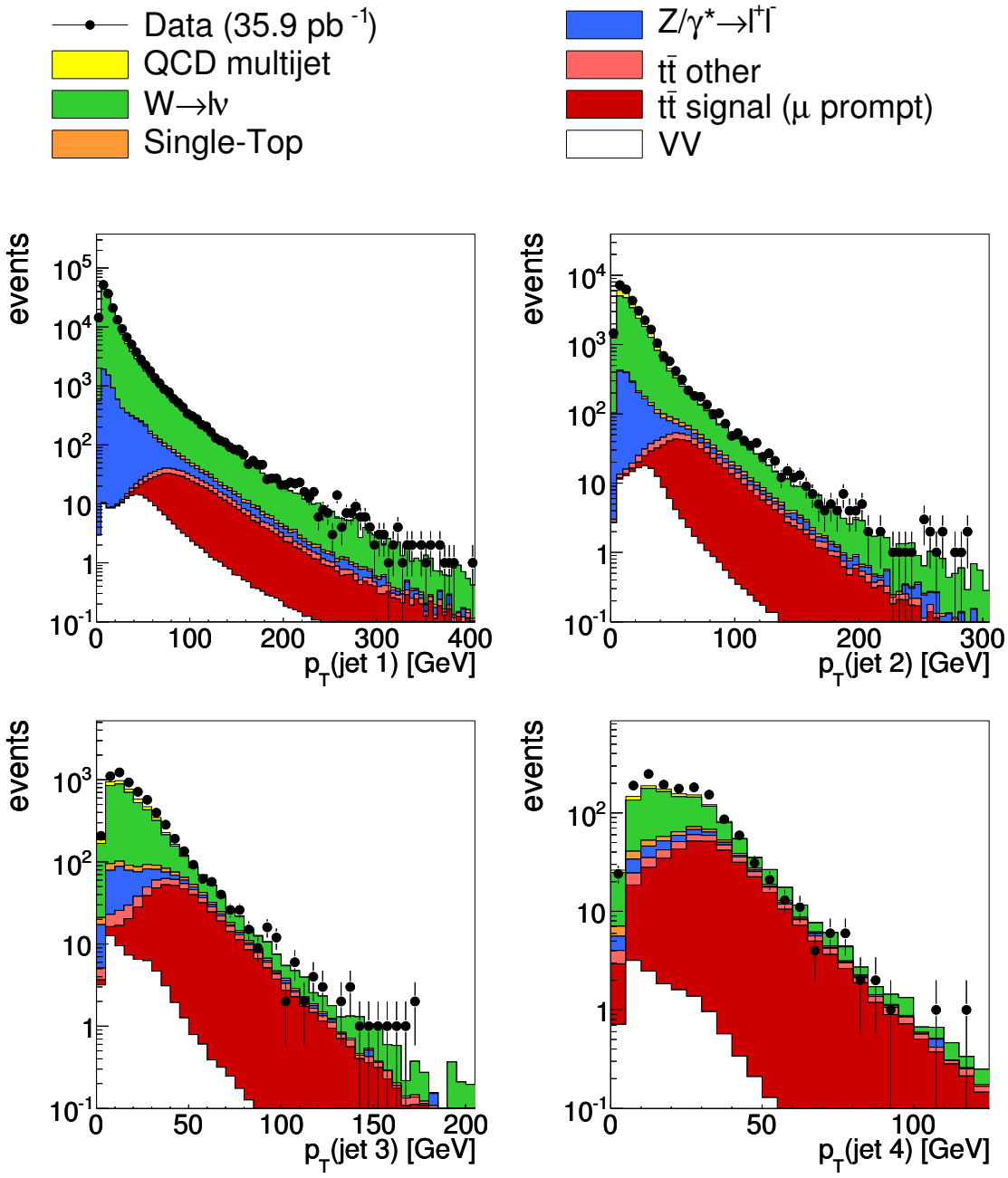

Figure 6.8: Transverse momentum of the first leading jet (upper left), second leading jet (upper right), third leading jet (lower left), and fourth leading jet (lower right) before applying the p_T cut on the particular jet. Black points denote data, coloured areas different simulated processes.

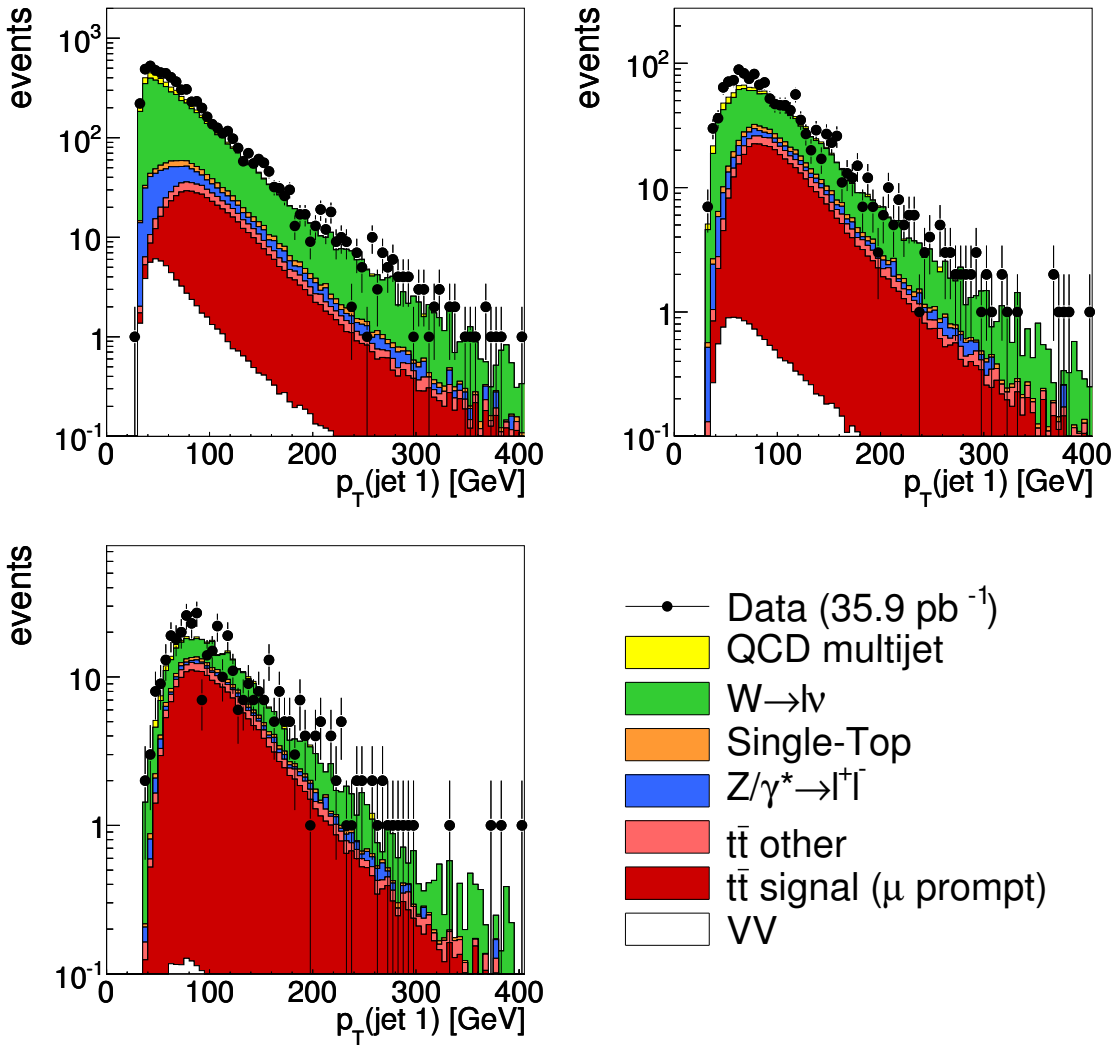

Figure 6.9: Transverse momentum of the first leading jet for at least two (upper left), three (upper right), or four (lower left) jets in the event. Black points denote data, coloured areas different simulated processes.

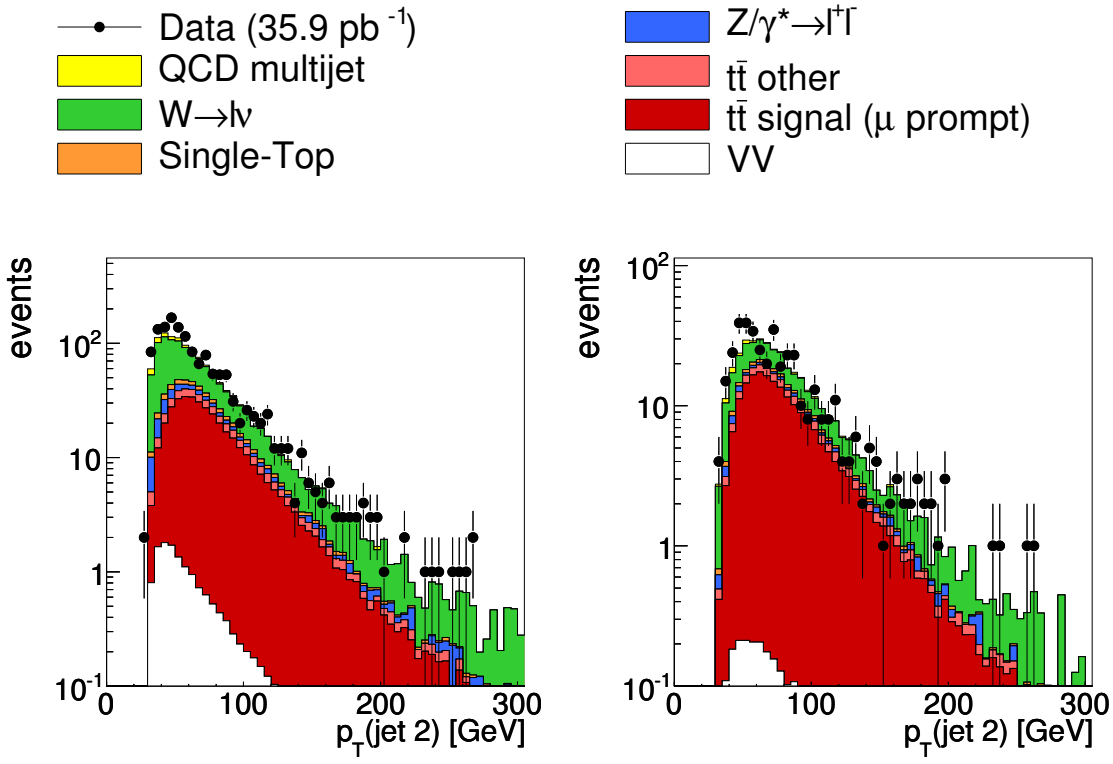

Figure 6.10: Transverse momentum of the second leading jet for at least three (left) or four (right) jets in the event. Black points denote data, coloured areas different simulated processes.

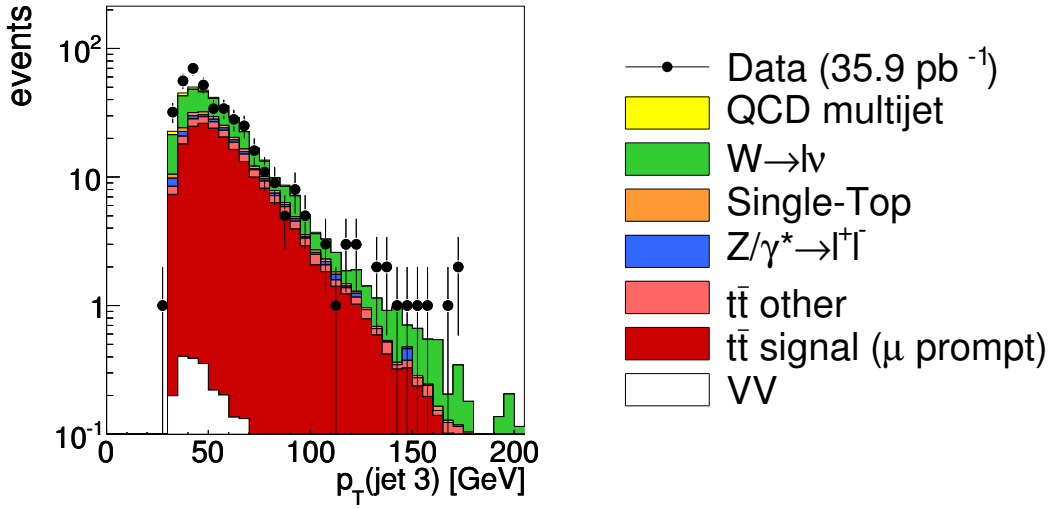

Figure 6.11: Transverse momentum of the third leading jet for at least four jets in the event. Black points denote data, coloured areas different simulated processes.

7 Derivation of Cross Sections

In this thesis, cross sections are derived for the production of exactly one isolated muon associated with at least one, two, three or four additional hard jets ($\sigma(\mu + \text{jets})$) and for the top-quark pair production ($\sigma(t\bar{t})$). All cross sections are determined inclusively, i. e. integrated over the analysed phase space, and moreover, differentially in kinematic quantities X of the muon, namely p_T and η, via

$$\left.\frac{d\sigma}{dX}\right|_i = \frac{N_i^{\text{sig}} - N_i^{\text{bkg}}}{\Delta X_i \cdot \epsilon_i \cdot L}, \tag{7.1}$$

where the index i indicates a distinct interval in p_T or η. N_i^{sig} is the number of measured events in the signal region, N_i^{bkg} the number of expected background events in this region. ΔX_i represents the width of the interval. ϵ_i covers all efficiencies and L is the integrated luminosity.

The formula (7.1) basically gives the line of approach for the derivation of the cross sections. First of all, the signal for which the cross section should be derived has to be defined by delimiting the phase space. Then the number of background events in the signal region is estimated. Migration effects have to be taken into account for the arrangement of the intervals. Reconstruction and selection efficiencies are estimated for the correction of the cross section. Finally, statistical and systematic uncertainties are estimated.

7.1 Phase Space

A well defined phase space is needed for the comparison of measured and calculated cross sections, which is performed not only in order to test the underlying theory but also in order to unfold the data, i. e. correcting for instrumental effects. The phase space is defined by the kinematic requirements to the muon and the jets as specified in the previous chapter. For the muon, these are

$$|\eta(\mu)| < 2.1 \tag{7.2}$$
$$p_T(\mu) > 20\,\text{GeV} \tag{7.3}$$

and for the jets,

$$|\eta(\text{jet})| < 2.4 \tag{7.4}$$
$$p_T(\text{jet}) > 30\,\text{GeV}. \tag{7.5}$$

σ	signal process
μ + jets	$t\bar{t} \to bWbW \to b\mu\nu_\mu bqq'$
	$t\bar{t} \to bWbW \to b\tau\nu_\tau bqq' \to b\mu\nu_\mu\nu_\tau\nu_\tau bqq'$
	$t \to bW \to b\mu\nu_\mu$
	$t \to bW \to b\tau\nu_\tau \to b\mu\nu_\mu\nu_\tau\nu_\tau$
	$W \to \mu\nu_\mu$
	$W \to \tau\nu_\tau \to \mu\nu_\mu\nu_\tau\nu_\tau$
$t\bar{t}$	$t\bar{t} \to bWbW \to b\mu\nu_\mu bqq'$
	$t\bar{t} \to bWbW \to b\tau\nu_\tau bqq' \to b\mu\nu_\mu\nu_\tau\nu_\tau bqq'$

Table 7.1: Signal processes for the μ + jets and the $t\bar{t}$ cross sections. No differentiation is made between particles and their corresponding antiparticles as they are coherently interchangeable. Additional partons from radiation processes are allowed but not stated here.

The signal processes for the different cross sections are compiled in table 7.1. For the muon+jets cross sections, the production of top quarks (in pairs and single) and the production of W bosons are considered as signal processes if exactly one W boson in the event is decaying leptonically. Furthermore, the final state has to consist of exactly one muon in the given phase space originating from this decay either directly or via an intermediate tau lepton and of at least one, two, three and four jets, respectively, obeying the kinematic criteria.

The signal process for the $t\bar{t}$ pair production cross section is exclusively the $t\bar{t}$ production. Here, the final state has to consist of exactly one muon in the given phase space and of at least four jets that fulfil the kinematic criteria. The origin of the muon is constrained like above.

Due to lepton universality, similar numbers of muons and tau leptons are produced in W-boson decays. However, the $\tau \to \mu\nu_\mu\nu_\tau$ branching ratio only amounts to $(17.36 \pm 0.05)\,\%$ [12]. Moreover, caused by the additionally produced neutrinos, the p_T spectrum of muons from tau decays is softer than that of prompt muons. Therefore, it is predicted from simulations that their fraction yields only less than 10 % of all selected signal events.

7.2 Background Estimation

The backgrounds are treated differently depending on their occurrence and their expected simulation accuracy. As the amount of diboson events is almost negligible and the amount of single top events is small and expected to be well modelled, their yields are both simply taken from the simulation. The $t\bar{t}$ pairs not decaying into muon+jets are treated together with the $t\bar{t}$ signal via the efficiency correction as described in section 7.4 since they are simulated identically. The emergence of isolated muons in QCD-multijet events is complicated to simulate so the estimation of its normalisation is completely data driven. The production mechanisms of the vector-boson samples, $Z/\gamma^* \to l^+l^-$

and $W \to l\nu$, are both well understood. They are handled together using simulation supported by data.

7.2.1 QCD Multijet Production

Especially for low jet multiplicities, QCD-multijet events are expected to make a substantial fraction of the selected data. As the modelling of isolated muons coming from hadron decays might lead to large uncertainties, the estimation of the MC simulation is corrected using data-driven estimates by means of a template fit, which basically is a fit of simulated histograms to a measured distribution.

For this method, a variable is needed that distinguishes between QCD-multijet events and events with one isolated muon and additional jets. As it is one of the aims of this analysis to measure cross sections differentially in kinematic quantities of the muon, these quantities should not be used for the template fit to not bias the measurement. A characteristic of events containing an isolated muon from the decay of a W boson is the missing transverse energy (MET) due to the undetected neutrino. QCD multijet events only have MET due to acceptance effects, finite detector resolution and mismeasurement, so the MET distribution is expected to be very discriminating.

To avoid relying on properly modelling the instrumental effects, the MET distribution in QCD-multijet events is taken from data in a control region, which is constructed by inverting the cut on the relative isolation (eq. (6.4)) of the muon and raising the cut value to 0.5. This provides a region that is completely dominated by QCD multijet events with an event fraction between 99.9 % in the inclusive one jet bin and 97.5 % in the inclusive four jet bin. Additionally, an upper cut of 1.0 is introduced on the relIso variable in order to reduce effects of a possible correlation between MET and relIso. The resulting relIso region of

$$0.5 < \text{relIso} < 1.0, \tag{7.6}$$

is called 'QCD region' in the following. Possible biases due to the choice of the cuts on relIso for the definition of the QCD region are included in the systematic uncertainties as described in the end of this section.

The kinematics of the remaining muon candidates, however, is well modelled as can be seen from figure 7.1. It shows a comparison of the p_T distribution of muon candidates in the QCD region between data and simulation, which agree very well. Due to the dominance of QCD-multijet events in this region, this is a test of the modelling of the muon kinematics in QCD-multijet events. Therefore, the result of the fit performed in a certain $p_T(\mu)$ region can be used for the whole $p_T(\mu)$ range. Restricting the transverse momentum of the muons to

$$20\,\text{GeV} < p_T(\mu) < 35\,\text{GeV}. \tag{7.7}$$

selects a kinematic region where a substantial fraction of events are QCD-multijet events even for events with isolated muons. This will make the following fit more stable.

The area of the resulting MET distribution is then scaled to the predicted event yield of the QCD-multijet simulation in the signal region and used as the QCD MET

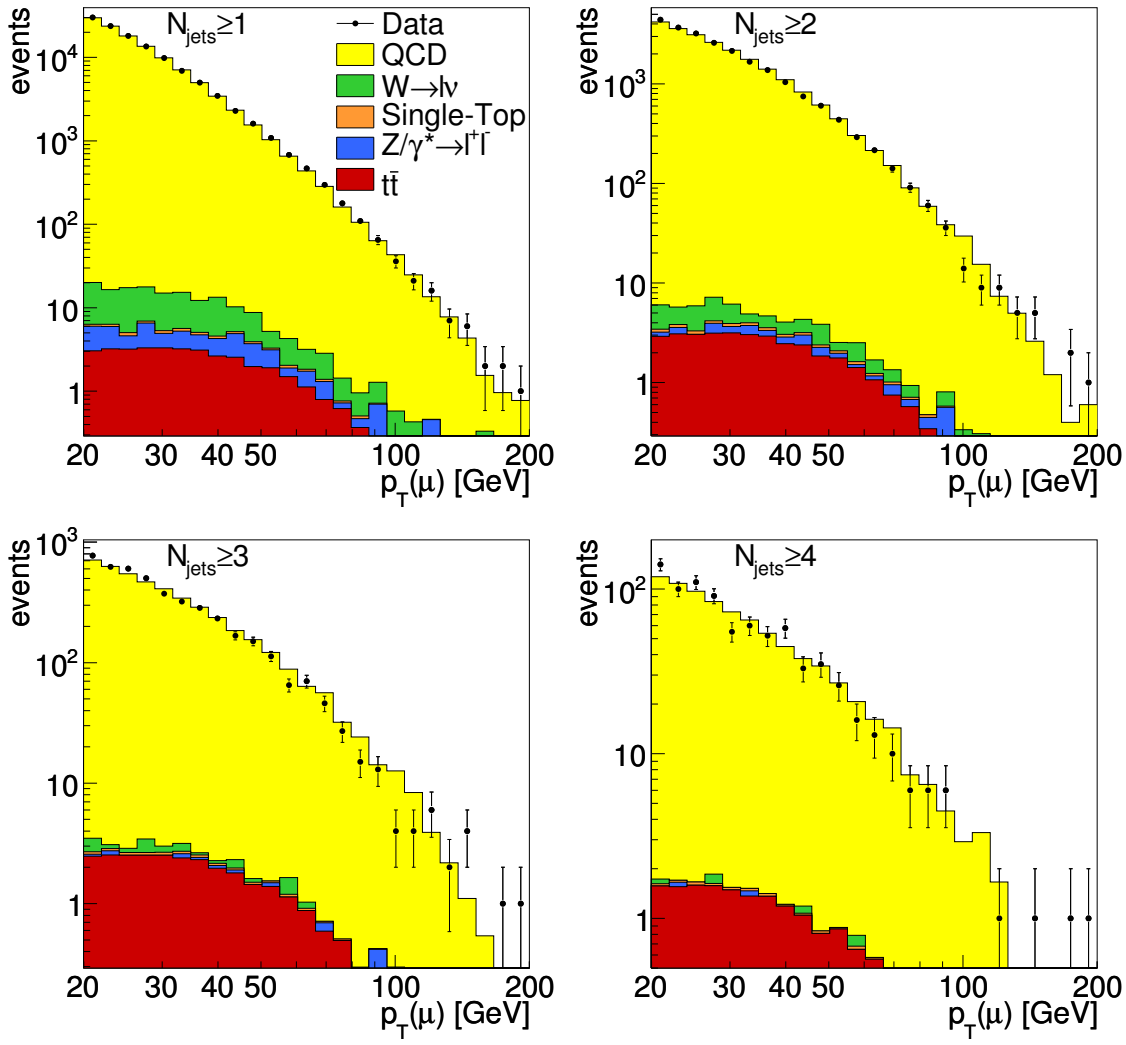

Figure 7.1: Transverse momentum of non-isolated muons (QCD region) for different jet multiplicities. Black points denote data, coloured areas different simulated processes.

template. The other MET templates are taken from simulation in the same p_T region but in the signal relIso region. The diboson samples are ignored due to their negligible contribution. Figure 7.2 shows the MET templates for the different jet multiplicities in comparison to data from the signal region without having performed a fit. It already hints at an underestimation of the QCD-multijet simulation.

The template scales are subject to a log likelihood[1] fit to match the MET distribution as observed in data in the signal region. As MET in $Z/\gamma^* \to l^+l^-$ and QCD-multijet events is likewise generated, the two MET templates are very similar. To dissolve the

[1]The likelihood is based on a Poisson probability density function for each bin.

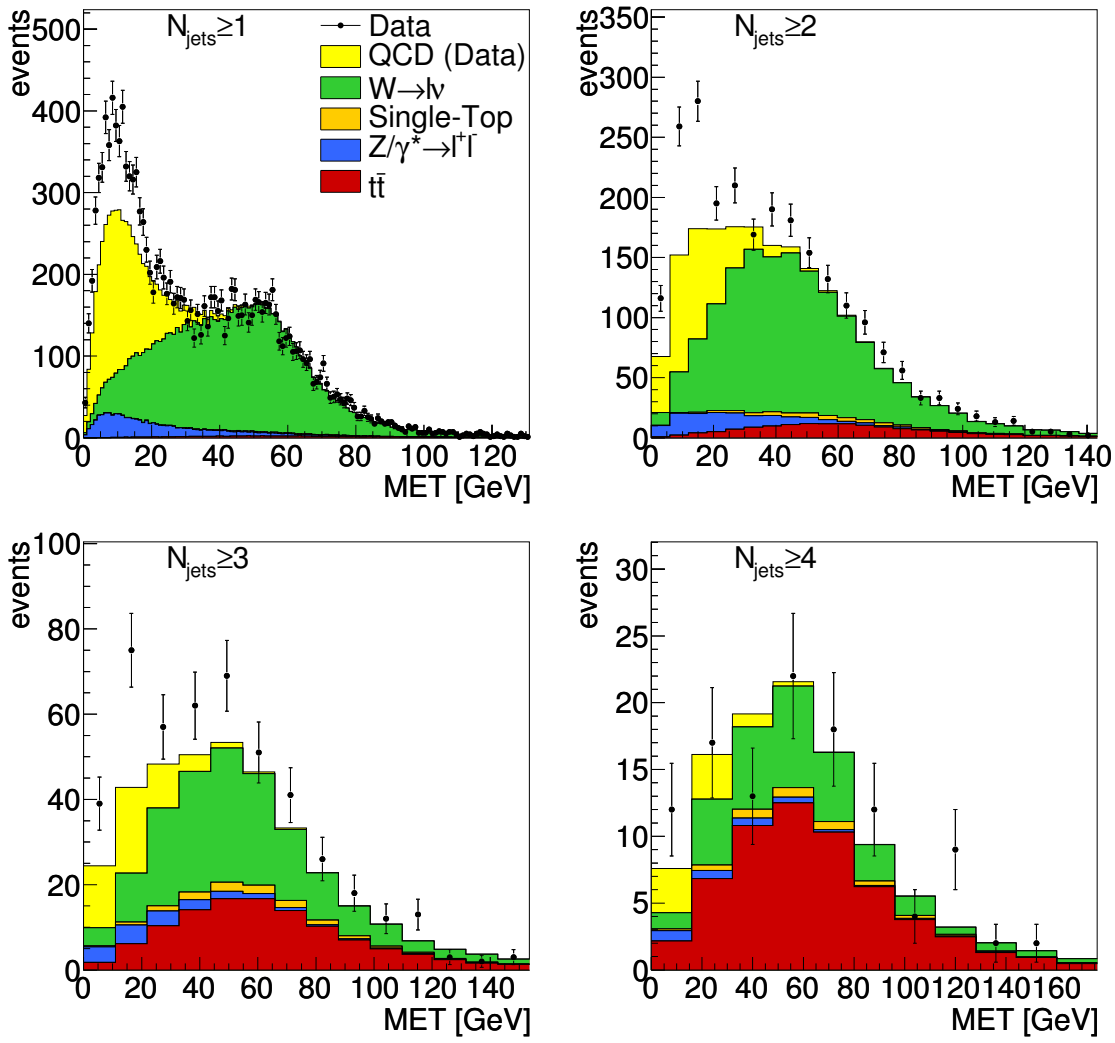

Figure 7.2: MET distributions for different jet multiplicities for events containing a muon with 20 GeV $< p_T <$ 35 GeV before the template fit. The QCD template is derived from data in the QCD region, the other templates from simulation.

similarity of the shapes in the fit, the scale of the $Z/\gamma^* \to l^+l^-$ template is fixed to the scale of the $W \to l\nu$ template. Physically, this is well justified because the production mechanisms of both vector bosons are equally understood so that the ratio of W/Z production is expected to be reasonably modelled. Furthermore, the shape of the $t\bar{t}$ and the single-top templates are very similar to the shape of the $W \to l\nu$ template so they are both fixed to the NLO prediction at this stage. The fit is performed for each jet multiplicity separately since this quantity is not expected to be correctly modelled by *Pythia* as it only calculates $2 \to 2$ processes on matrix-element level.

The MET distribution for the different jet multiplicities after performing the template fit are shown in figure 7.3. A much better agreement between data points and templates

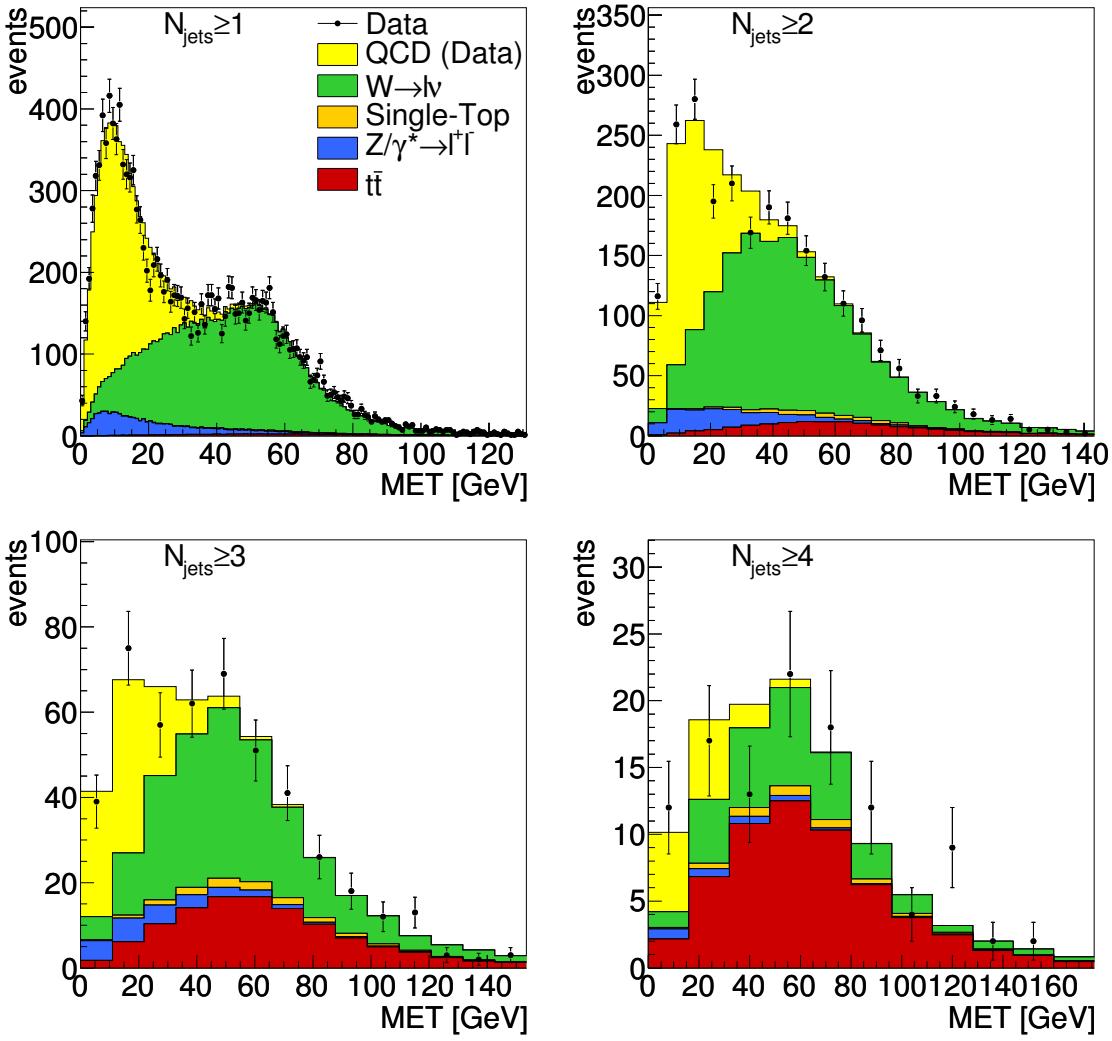

Figure 7.3: MET distributions for different jet multiplicities for events containing a muon with 20 GeV $< p_T <$ 35 GeV after the template fit. The QCD template is derived from data in the QCD region, the other templates from simulation.

is achieved by means of the fit. The resulting scale factors with statistical and systematic uncertainties for each jet multiplicity are given in table 7.2. It shows that the simulation underestimates the fraction of QCD for all jet multiplicities by factors of about 1.5 to 2.0. However, the estimation is suffering from statistics especially in the inclusive four jet bin, where the statistical uncertainty is considerably larger than the systematic uncertainties.

The systematic uncertainties quoted in the table are composed of three contributions.

- **variation of $t\bar{t}$:** The fixed scale of the $t\bar{t}$ template is varied up and down by a factor of 2 and 0.5, respectively.

jet multiplicity	QCD scale factor
$N_{\text{jets}} \geq 1$	$1.52 \pm 0.03(\text{stat.})^{+0.10}_{-0.05}(\text{syst.})$
$N_{\text{jets}} \geq 2$	$1.89 \pm 0.10(\text{stat.})^{+0.13}_{-0.08}(\text{syst.})$
$N_{\text{jets}} \geq 3$	$2.02 \pm 0.31(\text{stat.})^{+0.39}_{-0.19}(\text{syst.})$
$N_{\text{jets}} \geq 4$	$1.79 \pm 0.82(\text{stat.})^{+0.08}_{-0.37}(\text{syst.})$

Table 7.2: QCD scale factors for the different jet multiplicities with corresponding statistical and systematic uncertainty.

- **shift of relIso region:** The relIso region used for obtaining the QCD template is shifted once to lower values ($0.3 <$ relIso < 0.5) and once to higher values ($1.0 <$ relIso < 2.0).
- **different QCD templates:** The difference between the signal region and the QCD region is estimated by taking the QCD template from simulation separately in both regions and comparing the results.

The variation of the top scale is the dominant contribution in all but the inclusive one jet sample. The shift of the relIso region is meant to estimate the influence of signal contamination in the QCD region, which is larger for smaller values of relIso, and it is the dominant contribution in the inclusive one jet sample. The use of the different QCD templates from simulation gives an estimate of the correlation between relIso and MET. This is only a small contribution for all jet multiplicities.

The transverse mass m_T of the leptonically decaying W boson is used as control variable to check the validity of the scale factors. It is defined as

$$m_T = \sqrt{(p_T(\nu) + p_T(l))^2 - ((p_x(\nu) + p_x(l))^2 + (p_y(\nu) + p_y(l))^2)}, \tag{7.8}$$

where the momenta of the neutrino (ν) are taken from the missing transverse energy and the lepton (l) is the selected muon. The transverse mass is shown before (figure 7.4) and after (figure 7.5) applying the scale factors derived from the MET-template fit. The W mass peak seems to be slightly broader in data than in simulation, but apart from that the shape of the m_T distribution in data is well described by the templates after the scaling. Especially for lower jet multiplicities, the scaling of the templates cures the underestimation at low transverse masses, which is dominated by the QCD-multijet production.

7.2.2 Vector Boson + Jets

The production of W bosons associated with hard jets is by far the largest background for the $t\bar{t}$ cross section determination. Its fraction in the inclusive four jet bin is 32 %. Therefore, the prediction from simulation is improved by a data-driven approach taking advantage of the muon+jets cross section measurement for different jet multiplicities.

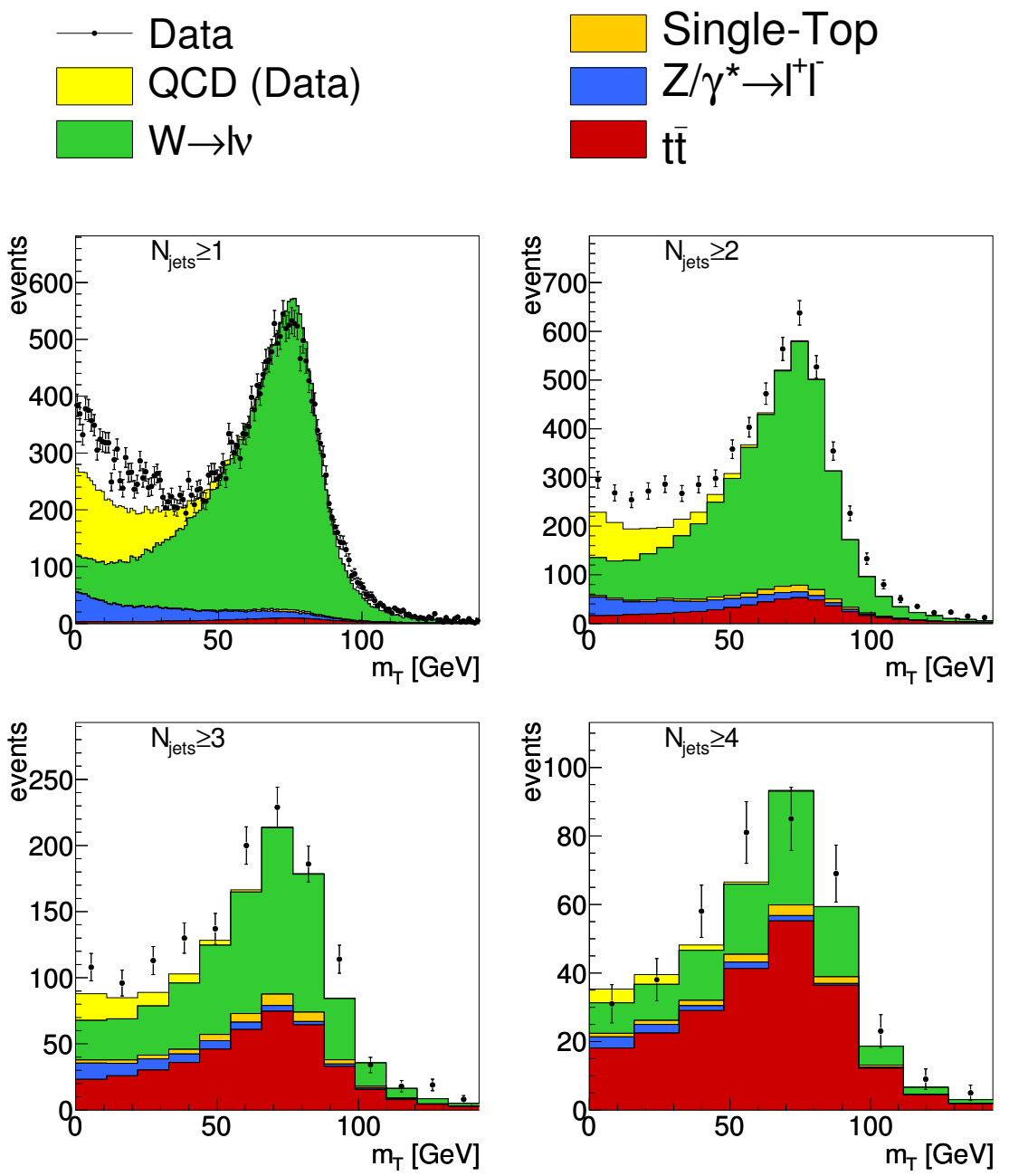

Figure 7.4: Distributions of the transverse W mass for different jet multiplicities for all selected events before the template fit. The QCD template is derived from data in the QCD region, the other templates from simulation.

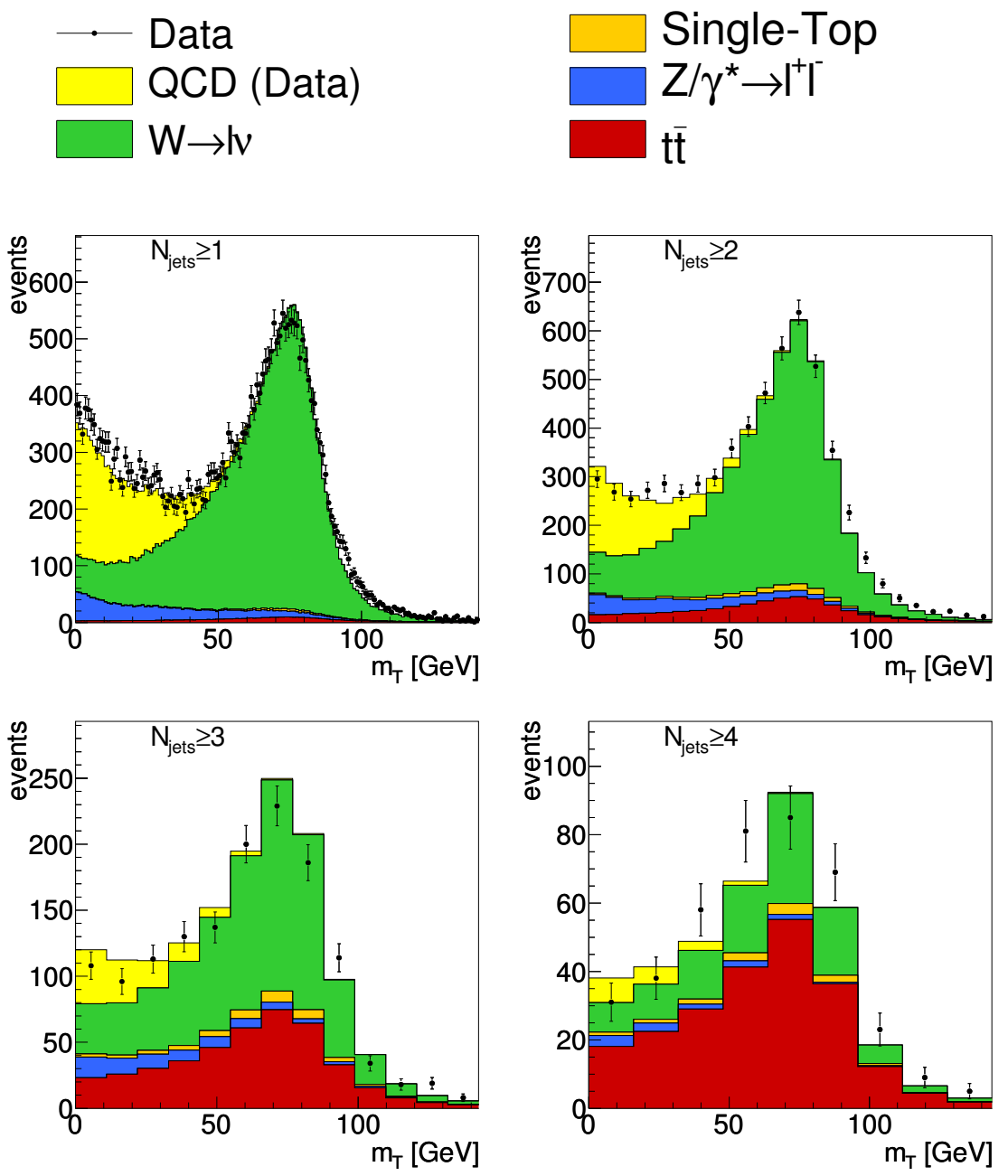

Figure 7.5: Distributions of the transverse W mass for different jet multiplicities for all selected events after the template fit. The QCD template is derived from data in the QCD region, the other templates from simulation.

process	including QCD	excluding QCD
$t\bar{t}$ signal	2 %	2 %
$t\bar{t}$ other	1 %	1 %
single top	2 %	2 %
$W \to l\nu$	78 %	86 %
$Z/\gamma^* \to l^+l^-$	6 %	7 %
QCD multijet	9 %	—
diboson	1 %	1 %

Table 7.3: Event composition for events with exactly one isolated muon and exactly two jets. The fraction is given for all events before and after excluding QCD-multijet production.

The event topologies of $W \to l\nu$ and $t\bar{t}$ decaying in the lepton+jets channel are very similar. Thus, the idea is to find a phase space that is preferably signal like but contains only a small fraction of $t\bar{t}$ events and to determine a normalisation for the predicted $W \to l\nu$ yield in this phase space such that it matches the data. This trade-off is found in events containing exactly one isolated muon and exactly two additional hard jets. The event composition for these events is shown in table 7.3. The only significant contribution besides $W \to l\nu$ and $Z/\gamma^* \to l^+l^-$ is QCD-multijet production, where the yield is estimated as described in section 7.2.1 and subtracted from data. All other processes are found to be insignificant because they contribute only $1 - 2\,\%$ each even after subtracting the QCD-multijet events. They are subtracted from data with the normalisation predicted from the NLO calculations. The yields of $W \to l\nu$ and $Z/\gamma^* \to l^+l^-$ events are then simultaneously normalized to match the data. A common normalisation factor of

$$1.16 \pm 0.02(\text{stat.}) \pm 0.04(\text{syst.}) \tag{7.9}$$

is obtained from this method. As *MadGraph* calculates up to four additional partons on matrix-element level for $W \to l\nu$ and $Z/\gamma^* \to l^+l^-$ events, it is assumed that the jet multiplicity is sufficiently modelled for up to four jets. Therefore, this normalisation factor is used equally for all jet multiplicities. It is applied to $Z/\gamma^* \to l^+l^-$ events for the determination of the muon+jets cross sections and to $W \to l\nu$ and $Z/\gamma^* \to l^+l^-$ events for the determination of the $t\bar{t}$ cross sections.

The systematic uncertainty of this method is derived by varying the predicted normalisation of the $t\bar{t}$ events by $\pm 100\,\%$. This is an estimate for the bias due to the signal contribution, which actually is to be measured. In order to cross-check the estimation of the normalisation factor, the normalisation of $W \to l\nu$ and $Z/\gamma^* \to l^+l^-$ events obtained from the template fit described in section 7.2.1 is taken from the inclusive two jet sample because the fit was not performed in the exclusive two jet sample. It amounts to 1.08 ± 0.04, where the error indicates the statistical uncertainty of the fit. Within the uncertainties, these values are compatible.

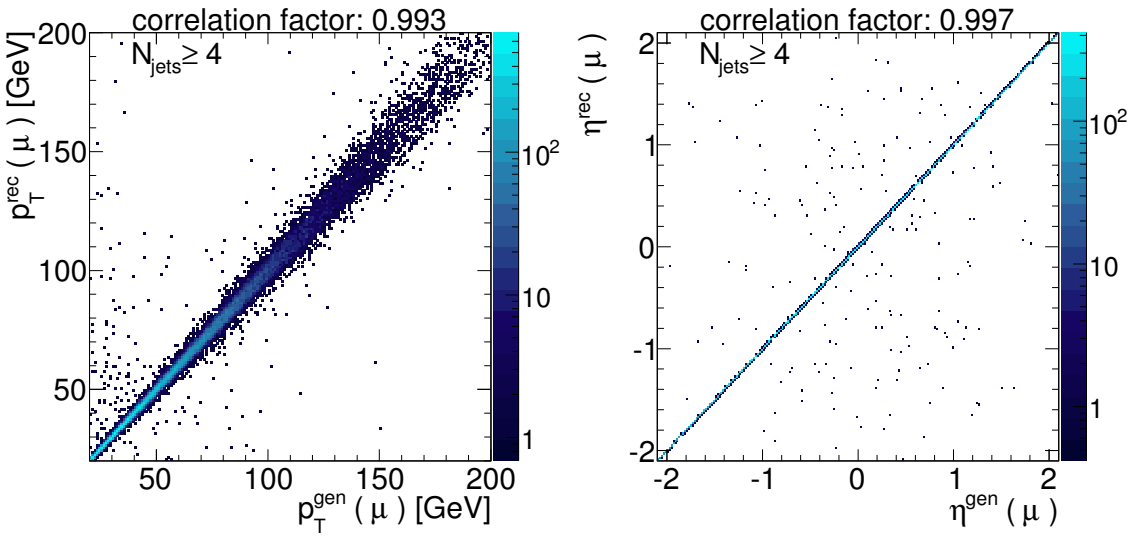

Figure 7.6: Correlation between generated and reconstructed kinematic variables of the muon. The left and the right plot show the transverse momentum and the pseudorapidity, respectively.

7.3 Migrations

If a quantity of an event, e. g. the p_T of the final-state muon, is not properly reconstructed, the event can end up in a wrong bin for a given binning in that quantity. This effect is called 'migration'. It can be studied using a MC simulation because it contains both the true information X^{gen} and the reconstructed information X^{rec}. Figure 7.6 shows the correlation between X^{gen} and X^{rec} for the transverse momentum and the pseudorapidity of the isolated muon in $t\bar{t}$ events after the full event selection. Generally, a matching between generated and reconstructed particles is done to ensure comparing the same object on different levels. However, only events containing exactly one muon in the final state are selected so mismatching is assumed to be rare. For both quantities, nice correlations with correlation factors close to unity can be seen, which is graphically expressed by the arrangement of the events on a diagonal line. The correlation of the transverse momentum is smeared for higher values because the momentum resolution of the tracker is getting worse for high-p_T tracks, which are becoming more and more straight. The correlation of the pseudorapidity is nearly perfect attesting an excellent spatial resolution to the inner tracking system. The few off-diagonal entries, which only amount to a fraction of about 0.3 %, are most likely due to events where the reconstructed and generated information originate from different muons.

For the determination of differential cross sections an adequate binning has to be chosen for each variable. In general, the binning is chosen in a way that a balance is found between preserving differential information, i. e. many small bins, and obtaining reasonably moderate migrations and small statistical uncertainties, i. e. only a few large bins. Due to the excellent reconstruction accuracy of muons in the CMS experiment,

as can be seen from the correlations in figure 7.6, migrations do not seem to play a relevant role for the cross sections differential in muon kinematics. The bin widths are therefore set to ensure a minimum bin content of the order of 100 events to limit statistical uncertainties. The resulting bin ranges are given in table 7.4.

To quantify the strength of migrations two measures, purity and stability, are introduced.

The purity for each bin i is defined as

$$p_i = \frac{N_i^{\text{gen\&rec}}}{N_i^{\text{rec}}}, \tag{7.10}$$

where $N_i^{\text{gen\&rec}}$ is the number of events that are generated and reconstructed in the same bin, and N_i^{rec} is the number of events reconstructed in that bin. Only those events that are generated within the phase space defined in section 7.1 are considered. The purity is a measure of how many events are migrating into a certain bin and therefore impurify the bin.

The stability for each bin i is defined as

$$s_i = \frac{N_i^{\text{gen\&rec}}}{N_i^{\text{gen}}}, \tag{7.11}$$

where $N_i^{\text{gen\&rec}}$ is defined as above, and N_i^{gen} is the number of events that are generated in bin i. Here, only those events that are selected on reconstruction level are considered. The stability is a measure for the fraction of events that are migrating out of the bin testifying its instability.

The outcome for purity and stability given the binning defined in table 7.4 is shown in figure 7.7. For the transverse momentum, purity and stability are well above 90 % and for the pseudorapidity they are even well above 99 % again reflecting the highly correlated distributions shown in figure 7.6.

In addition, figure 7.7 shows a migration correction, which is calculated as the ratio of purity and stability,

$$C_i^{\text{mig}} = \frac{p_i}{s_i} = \frac{N_i^{\text{gen}}}{N_i^{\text{rec}}}. \tag{7.12}$$

Using this correction factor is a simple but robust way of unfolding. According to [61], it is appropriate if the correction factors are of order unity and the bin widths are much larger than the resolution in the considered variable. Both conditions are fulfilled in the present analysis.

	1	2	3	4	5
p_T [GeV]	[20 , 35[	[35 , 50[	[50 , 70[	[70 , 95[	[95 , 200[
η	[-2.1 , -0.75[	[-0.75 , 0.0[	[0.0 , 0.75[	[0.75 , 2.1[	—

Table 7.4: Bin ranges for the differential cross sections in p_T and η of the isolated muon.

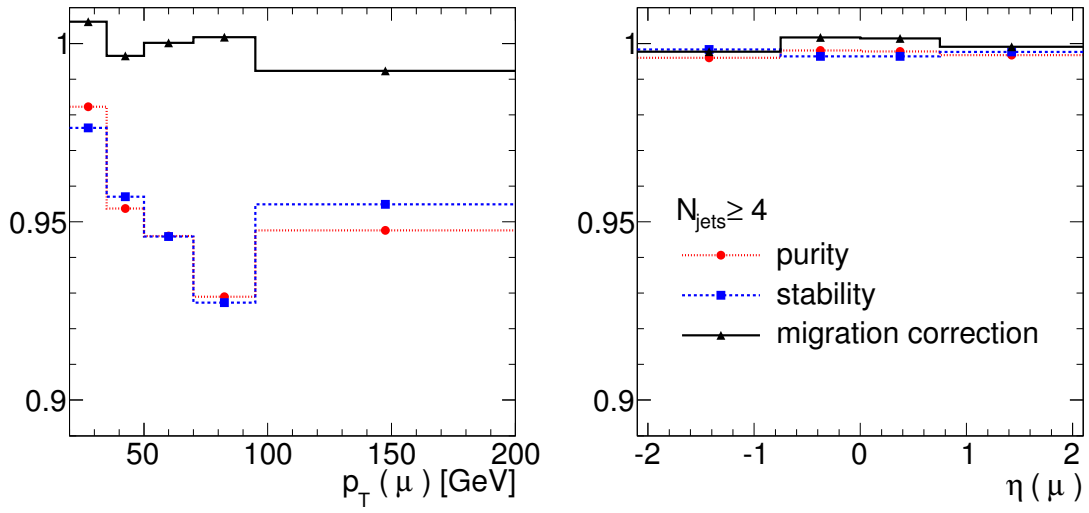

Figure 7.7: Migrations between the bins of the differential cross sections are shown for the transverse momentum on the left and for the pseudorapidity on the right. The dotted line indicates the purity, the dashed line the stability, and the solid line the resultant migration correction.

7.4 Efficiency Corrections

Events are associated to a certain process by means of the measured final-state particles. For several reasons, it is impossible to reconstruct the complete list of particles for all events. First of all, a real detector has a certain acceptance, i. e. it can not cover the full solid angle of 4π. Accelerator equipment and beam pipe have to fit in the detector structure. Furthermore, the instrumented regions contain so-called 'dead material', e. g. the supporting structure, which leads to inefficiencies in the reconstruction of particles. Then, concerning event selection, online and offline criteria are optimized to find a trade-off between high selection efficiencies for signal processes and an efficient background reduction. This also leads to a rejection of signal events.

For a proper cross section determination, the measured number of events has to be corrected for these inefficiencies. In this analysis, the efficiency is calculated from simulation via

$$\epsilon = \frac{N^{\mathrm{rec}}}{N^{\mathrm{gen}}}, \tag{7.13}$$

where N^{rec} is the total number of reconstructed and selected events originating from one of the according signal processes given in table 7.1. N^{gen} is the total number of events that are generated within the chosen phase space. As described in section 6.5.4, the efficiency is partly corrected for differences between data and simulation by applying a scale factor derived from data. As the scale factor is related to muon efficiencies only, it is the same for all jet multiplicities. The determined efficiencies after the scale factor is applied are shown in table 7.5 for the different inclusive cross sections.

σ	jet multiplicity	efficiency [%]
μ + jets	$N_{\text{jets}} \geq 1$	83.2 ± 0.1
	$N_{\text{jets}} \geq 2$	77.6 ± 0.3
	$N_{\text{jets}} \geq 3$	74.3 ± 0.5
	$N_{\text{jets}} \geq 4$	68.9 ± 0.6
$t\bar{t}$	$N_{\text{jets}} \geq 4$	69.1 ± 0.3

Table 7.5: Efficiencies for the inclusive μ + jets and $t\bar{t}$ cross sections corrected for known differences between data and simulation. The quoted uncertainties are statistical only.

For the μ + jets cross sections, this is a combined efficiency for the different signal processes. The obtained value is valid only if either the efficiencies of the different processes are similar or the event composition is correctly predicted. As for the inclusive four jet multiplicity, the combined efficiency for μ + jets and the pure $t\bar{t}$ efficiency agree within the uncertainties, the combined efficiency correction is applicable.

The few $t\bar{t}$ events that do not decay in the muon+jets channel but nevertheless survive the complete event selection, are included in the number of reconstructed events in equation (7.13). Hence, this background is removed by applying the efficiency correction. Therefore, a measured $t\bar{t}$ cross section that deviates from the predicted cross section is automatically considered for the removal of background events originating from other $t\bar{t}$ decays.

In order to obtain the inclusive $t\bar{t}$ production cross section, the calculated cross section is extrapolated to the whole phase space including all $t\bar{t}$ decay channels. The extrapolation factor $f_{\text{extrapolation}}$ is derived by comparing the total predicted number of events in the whole phase space, $N^{t\bar{t}}_{\text{total}}$, with the number of events generated by *MadGraph* restricted to the chosen phase space for the given decay channel, $N^{t\bar{t}\to\mu+\text{jets}}_{\text{phase space}}$, and amounts to

$$f_{\text{extrapolation}} = \frac{N^{t\bar{t}}_{\text{total}}}{N^{t\bar{t}\to\mu+\text{jets}}_{\text{phase space}}} = 17.85. \tag{7.14}$$

MadGraph uses the simplified assumption for the W branching ratios of $\frac{1}{9}$ for each possible decay channel. This is corrected for in the calculation of the extrapolation factor using the up-to-date knowledge of the W-boson decay (see section 4.3).

For differential cross sections the efficiencies are determined separately for each bin. The resulting efficiency corrections then per se include a correction for migrations between the bins as described in section 7.3. The efficiencies as a function of p_T and η of muons in the analysed phase space are shown in figure 7.8. For $p_T(\mu)$, the efficiencies rise with increasing p_T from about 80 % to 90 % for $N_{\text{jets}} \geq 1$ and from about 60 % to 80 % for $N_{\text{jets}} \geq 4$. The difference in shape for different jet multiplicities is probably due to the requirement on the relative isolation of the muon, which is p_T dependent. In $\eta(\mu)$, the efficiencies are flat slightly above 80 % for $N_{\text{jets}} \geq 1$ and at about 70 % for $N_{\text{jets}} \geq 4$.

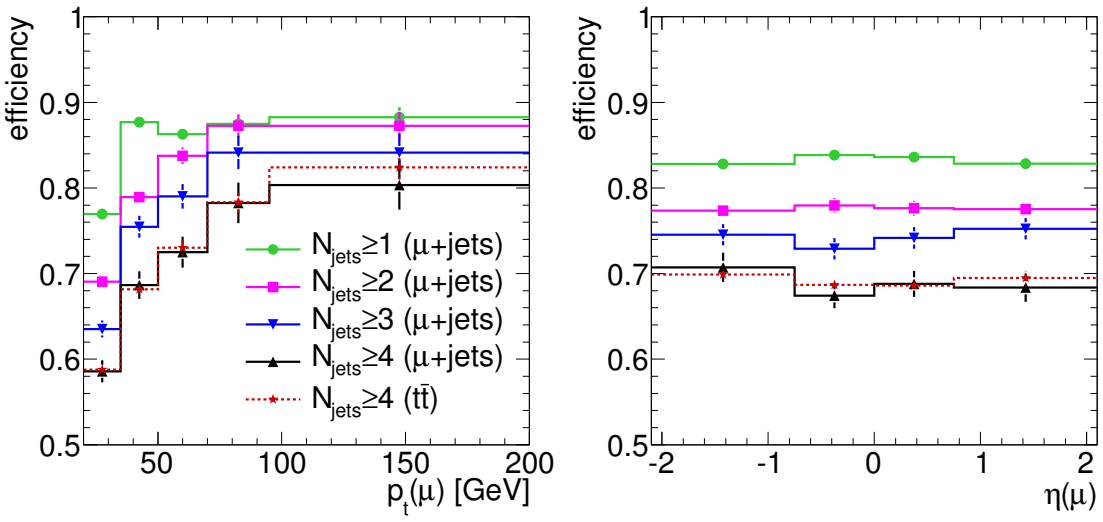

Figure 7.8: Efficiencies for the μ + jets (solid lines) and $t\bar{t}$ (dashed line) cross sections differential in p_T (left) and η (right) of muons in the analysed phase space.

7.5 Uncertainties

After correcting the measurement for all known inefficiencies, there are still possible sources of inaccuracy that can not easily be eliminated. They often originate from biases in the measurement and uncertainties in theoretical models. Also, limited statistics contributes to these inaccuracies, either in the measurement itself or in the derivation of corrections. In order to be able to correctly judge about possible deviations between measurement and prediction, statistical as well as systematic uncertainties are determined.

7.5.1 Statistical Uncertainty

The statistical uncertainty on the cross sections is obtained via Gaussian error propagation of equation (7.1), where only the variance of the number of measured events, $\sqrt{N_i^{\text{sig}}}$, is considered.

$$\delta\left(\left.\frac{d\sigma}{dX}\right|_i\right) = \frac{\sqrt{N_i^{\text{sig}}}}{\Delta X_i \cdot \epsilon_i \cdot L}, \tag{7.15}$$

Uncertainties due to limited statistics in the determination of corrections and in the estimation of background normalisations are included in the systematic uncertainties.

7.5.2 Systematic Uncertainty

There are several different experimental and theoretical sources of systematic uncertainty for the cross section measurement. Table 7.6 contains all known relevant sources and

Source	$\sigma_{\text{incl}}(\mu + \text{jets})$				$\sigma_{\text{incl}}(t\bar{t})$
	$N_{\text{jets}} \geq 1$	$N_{\text{jets}} \geq 2$	$N_{\text{jets}} \geq 3$	$N_{\text{jets}} \geq 4$	$N_{\text{jets}} \geq 4$
JES	8%	11%	11%	13%	15%
JER	1%	$< 1\%$	$< 1\%$	$< 1\%$	1%
QCD-multijet estimate	12%	8%	4%	2%	2%
$Z/\gamma^* \to l^+l^-$ normalisation	3%	2%	2%	1%	$< 1\%$
Single-top normalisation	-	-	-	-	1%
Diboson normalisation	$< 1\%$	$< 1\%$	$< 1\%$	$< 1\%$	$< 1\%$
$t\bar{t}$ in W/Z+Jets estimate	-	-	-	-	2%
Q^2 scale ($t\bar{t}$)	$< 1\%$	$< 1\%$	$< 1\%$	$< 1\%$	7%
Q^2 scale (W/Z)	3%	$< 1\%$	2%	1%	15%
Matching threshold ($t\bar{t}$)	$< 1\%$	$< 1\%$	$< 1\%$	$< 1\%$	2%
Matching threshold (W/Z)	$< 1\%$	$< 1\%$	1%	1%	6%
ISR/FSR ($t\bar{t}$)	$< 1\%$	$< 1\%$	1%	2%	2%
Underlying-event tune ($t\bar{t}$)	$< 1\%$	$< 1\%$	$< 1\%$	$< 1\%$	2%
Muon efficiencies	3%	3%	3%	3%	3%
Pile up	2%	4%	3%	2%	$< 1\%$
Systematic	15%	14%	13%	14%	24%
Statistical	1%	2%	3%	5%	9%
Luminosity	4%	4%	4%	4%	4%
Combined	16%	15%	14%	16%	26%

Table 7.6: Uncertainties for the muon+jets and $t\bar{t}$ cross sections. There is no uncertainty on the estimate of single top events for the muon+jets measurement as single top is part of the signal. For the top quark cross section measurements the uncertainty of the normalization of W/Z boson events is absorbed into other uncertainties as described in the text. The uncertainty on the $t\bar{t}$ signal fraction in the exclusive two jet bin for the estimate of the W/Z+jets estimate is not applicable for the muon+jets measurement.

their impact on the different inclusive cross sections. The size of the variations is chosen according to the recommendations of the *TopPAG* [51]. The single contributions are added in quadrature to obtain the combined uncertainty. A detailed description of the different sources is given in the following.

- **Jet Energy:** One of the largest uncertainties stems from the energy measurement of jets affecting their selection. In order to emulate this effect, the jet-energy scale (JES) in simulation is shifted depending on η and p_T of the jets [59]. In the analysed phase space, these shifts range from about 2.5 % to 6 % of the measured jet energy. Furthermore, the following analysis- and dataset-specific uncertainties

are added in quadrature to the recommended shift. Pile-up effects remaining after the jet-energy offset correction are taken into account by applying an uncertainty of $35.2\,\%/p_T(\text{jet})$, which is a jet-energy variation of at most about 1.2 % for the lowest energetic jets in this analysis ($p_T = 30\,\text{GeV}$). The difference in the energy scale for light-quark and b-quark induced jets amounts to an uncertainty in the energy of b-quark induced jets of 2 % for central ($\eta < 2.0$) jets with $50\,\text{GeV} < p_T < 200\,\text{GeV}$ and 3 % for all other jets. An additional jet-energy uncertainty of 1.5 % is assigned to possible discrepancies due to the old simulation version that was used to derive the jet-energy correction factors which were available at the time of finishing this analysis.

In addition, the jet-energy resolution (JER) was found to be larger in data than in simulation by about 10 %. For the measurement, this was taken into account by applying a smearing, which is performed by adding 10 % of the difference between the jet energy on reconstruction level and on hadron level to the jet energy on reconstruction level. The matching of the jets between the two levels is performed via the distance in the (η, ϕ) plane, which must not be above 0.4. The difference in the cross section between applying no smearing and increasing the resolution by 20 % applying a larger smearing was taken as an additional uncertainty.

- **Background:** Another source of uncertainties is the normalisation of background yields. For low jet multiplicities, the uncertainty on the QCD normalisation is dominating. Although uncertainties on the QCD scaling factors are derived in the course of QCD-multijet background estimation as quoted in table 7.2, a conservative approach of taking a larger uncertainty of systematically 50 % was chosen in order to be consistent with other analyses by CMS.

  In general, the $Z/\gamma^* \to l^+l^-$, single-top, and diboson predictions are varied by ±30 %. As described in section 7.2.2, a combined vector-boson normalisation is derived from data. Hence, the variation of the $Z/\gamma^* \to l^+l^-$ prediction is basically a variation of the $Z/\gamma^* \to l^+l^-$ fraction in the combined vector-boson sample ($Z/\gamma^* \to l^+l^-$ and $W \to l\nu$). This is not relevant for the $t\bar{t}$ cross section because both are subtracted. The single-top variation is only relevant for the $t\bar{t}$ cross section as it is considered as signal for the muon+jets cross sections. The diboson variation is almost negligible in either case.

  As already mentioned in section 7.2.2, the $t\bar{t}$ prediction is varied by ±100 % for the estimation of the common vector-boson normalisation. This variation is taken into account only for the $W \to l\nu$ estimate in the $t\bar{t}$ cross section determination. Larger variations of the $W \to l\nu$ estimation are due to uncertainties in theoretical models used for the simulation.

- **Theory:** There are many theoretical uncertainties in the modelling of processes. In perturbative QCD, there are two relevant scales that are chosen arbitrarily, the renormalisation scale μ_R and the factorisation scale μ_F. μ_R is introduced to modify the strong coupling constant α_s such that it absorbs certain divergencies

that occur in calculations of higher order corrections, and μ_F basically is the energy scale below that processes are 'factorised' into the proton PDF. The choice of these scales does not have an influence on hypothetical calculations done to an infinite order of α_s. However, *MadGraph* calculates processes at leading order of α_s, so the choice of the scales actually affects the calculations. The scales are parametrised in *MadGraph* via

$$\mu_F^2 = \mu_R^2 = Q^2 \left(m_{FS}^2 + \sum_j p_T^2 \right), \tag{7.16}$$

where Q is a scale factor, which is set to unity as a default, and m_{FS} is the mass of the heaviest final state particle. The sum accounts for the transverse momenta of all particles that are generated by *MadGraph*, e. g. the top and the antitop quark and the additional partons in case of the $t\bar{t}$ sample. In order to account for a possible bias due to this arbitrary choice of μ_F and μ_R, Q is varied up and down by a factor of 2 and 0.5, respectively. As $W \to l\nu$ and $Z/\gamma^* \to l^+l^-$ have a very similar production mechanism, the uncertainties on their scales are assumed to be correlated and thus, the variation is performed simultaneously, whereas it is done separately for $t\bar{t}$.

The matching threshold for parton showers mentioned in section 5.1 is varied up and down by a factor of 2 and 0.5, respectively. Assuming the same correlation as above, this variation is also done simultaneously for $W \to l\nu$ and $Z/\gamma^* \to l^+l^-$ and separately for $t\bar{t}$.

Different settings for initial- and final-state QCD radiation (ISR/FSR) are used for the simulation of $t\bar{t}$ events to estimate the effect of the uncertainty on this radiation. As the dedicated systematic samples suffer from poor statistics, a significant contribution to the determined uncertainties might be due to statistical fluctuations in the samples. However, the statistics is adequate to conclude that it does not belong to the dominating uncertainties.

Uncertainties on the underlying event are determined by using the tune Z2 mentioned in section 5.2 instead of the default tune D6T for the simulation of $t\bar{t}$ events.

The above mentioned variations in the simulation of processes do not have a large effect on the muon+jets cross section determination as the simulation of signal processes are not strongly involved in the cross section determination. The variation of the Q^2 scale has the largest effect as it influences the estimation of the $Z/\gamma^* \to l^+l^-$ normalisation.

For the determination of the $t\bar{t}$ cross section, variations for $W \to l\nu$ and $Z/\gamma^* \to l^+l^-$ have a large influence as $W \to l\nu$ constitutes the largest background. Especially, the jet multiplicity that is important for the estimation of the normalisation, is strongly affected by these variations. Variations for the simulation of $t\bar{t}$ events have a negligible effect on the determination of the $t\bar{t}$ cross section in the chosen phase space, but they strongly influence the extrapolation to the whole phase space.

- **Muon Efficiency:** A further uncertainty arises from the determination of the muon efficiencies as described in section 6.5.4. The scale factor is varied by $\pm 3\,\%$ to account for differences between $t\bar{t}$ and $Z/\gamma^* \to l^+l^-$ events, from which the scale factor is derived.

- **Pile Up:** The default simulation was done without any pile up. In data, differences in the jet energy due to the existing pile up are corrected via the jet-energy offset correction and remaining systematics are discussed above. However, not only jets are affected by pile up but also muons due to the isolation requirement. Therefore, simulated samples of all processes with pile up are used to estimate the influence of pile up.

- **Luminosity:** The uncertainty on the luminosity measurement was determined to be $4\,\%$ [62]. This uncertainty is separately stated in the following as it is not specifically related to this analysis.

8 Results

The results of the cross section measurements are summarised in this chapter. The muon+jets cross sections are determined inclusively and differentially in p_T and η of the final state muon for at least one, two, three, and four jets in the final state.[1] The inclusive $t\bar{t}$ cross section is measured in the chosen phase space but also extrapolated to the whole phase space. Finally, the $t\bar{t}$ cross section is presented differentially in p_T and η of the final state muon.

8.1 Muon+Jets Cross Sections

The inclusive muon+jets cross section measurements for four exclusive and inclusive jet multiplicities are shown in figure 8.1. Comparing the expected event composition for the different bins with each other exhibits the transition from a completely $W \to l\nu$ dominated ($N_{\text{jets}} \geq 1$) to a strongly $t\bar{t}$ enriched ($N_{\text{jets}} \geq 4$) phase space. There is a reasonable agreement between the measurement and the prediction from simulation within the uncertainties of about 20 %. Nevertheless, the prediction is systematically below the data points. This reflects the fact that the $W \to l\nu$ prediction has to be scaled up for the $t\bar{t}$ cross section measurement as discussed in chapter 7.

The muon+jets cross sections differential in the transverse momentum of the final state muon for the different jet multiplicities are shown in figure 8.2. Again, a clear transition from dominating $W \to l\nu$ events in the inclusive one jet bin to a large fraction of $t\bar{t}$ events in the inclusive four jet bin is visible. The predicted cross sections are systematically below the measurement. Additionally, a tendency to a larger discrepancy for higher transverse momenta of the muon at least for low jet multiplicities can be observed. The large systematic uncertainty on the lowest $p_T(\mu)$ bin for low jet multiplicities is due to the conservatively large uncertainty on the QCD estimation of 50 % because QCD makes a large fraction of the event yield in this bin before background subtraction.

The muon+jets cross sections differential in the pseudorapidity of the final state muon for the different jet multiplicities are shown in figure 8.3. Here, the transition from $N_{\text{jets}} \geq 1$ to $N_{\text{jets}} \geq 4$ also exhibits a difference in the shape of the distributions. In the inclusive one jet bin, the $\eta(\mu)$ distribution is approximately flat only showing a weakly pronounced enhancement of central muons. With increasing jet multiplicity, the centrality of the muons rises and central muons obviously dominate for high jet multiplicities. This is due to the different involved masses and production mechanisms

[1] The inclusive muon+jets cross sections for the different jet multiplicities that are presented in this thesis have been published in the European Physical Journal C (EPJC) [51].

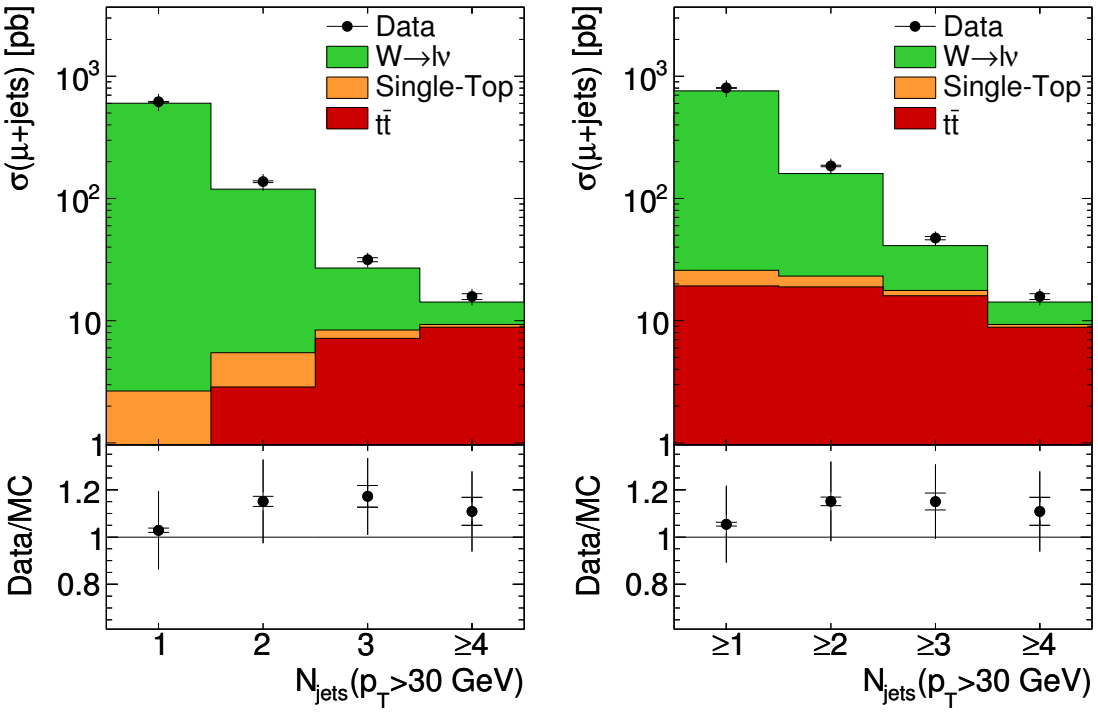

Figure 8.1: Muon+jets cross section as a function of the exclusive (left) and inclusive (right) jet multiplicity. The inner error bars correspond to the statistical uncertainties while for the full error bars, systematic uncertainties are added in quadrature. The data points are compared to the prediction obtained by the simulation with the *MadGraph* Monte Carlo event generator.

for $W \to l\nu$ and $t\bar{t}$ events, which are dominating different jet multiplicities. The centre-of-mass energy of the collision is used for creating the masses and the momenta of the outgoing particles. The mass of one W boson of about 80.4 GeV is considerably smaller than the mass of two top quarks of about 173 GeV each. Therefore, the produced W boson has on average a larger momentum than the top quarks. Moreover, the main production mechanism of single W^+ and W^- bosons at the LHC is through $u\bar{d}$ and $\bar{u}d$ annihilation, respectively. In a proton-proton collider like the LHC, this implies that the antiquark is a sea quark, mostly carrying only a small fraction of the proton momentum, whereas the quark often is a valence quark with a larger fraction of the proton momentum. Thus, the produced W boson often has a boost in the flight direction of the incoming quark, which is transferred to the muon from the decay. In contrast, the main production mechanism of $t\bar{t}$ pairs is mediated via gluon-gluon fusion. Due to the large mass of the $t\bar{t}$ pair, both of the gluons have to carry a considerable fraction of the proton momentum to provide the required invariant mass. Therefore, the $t\bar{t}$ pair is often produced with little kinetic energy and thus without a significant boost in z direction. Within the uncertainties, the shape of the measurement is nicely described by the prediction indicating that the assumed event composition is reasonable.

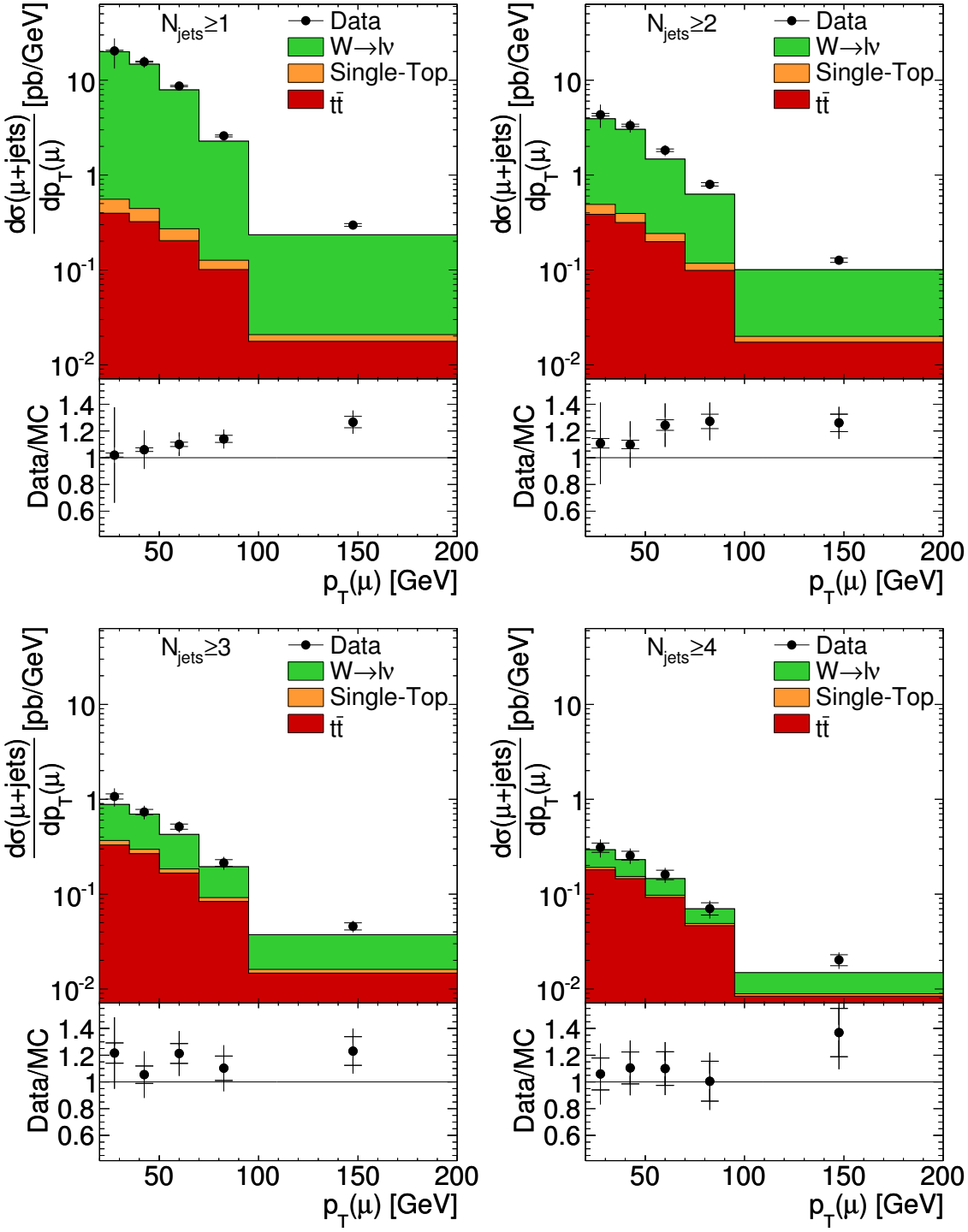

Figure 8.2: Differential muon+jets cross sections as a function of the p_T of the final state muon for different jet multiplicities. The inner error bars correspond to the statistical uncertainties while for the full error bars, systematic uncertainties are added in quadrature. The data points are compared to the prediction obtained by the simulation with the *MadGraph* Monte Carlo event generator.

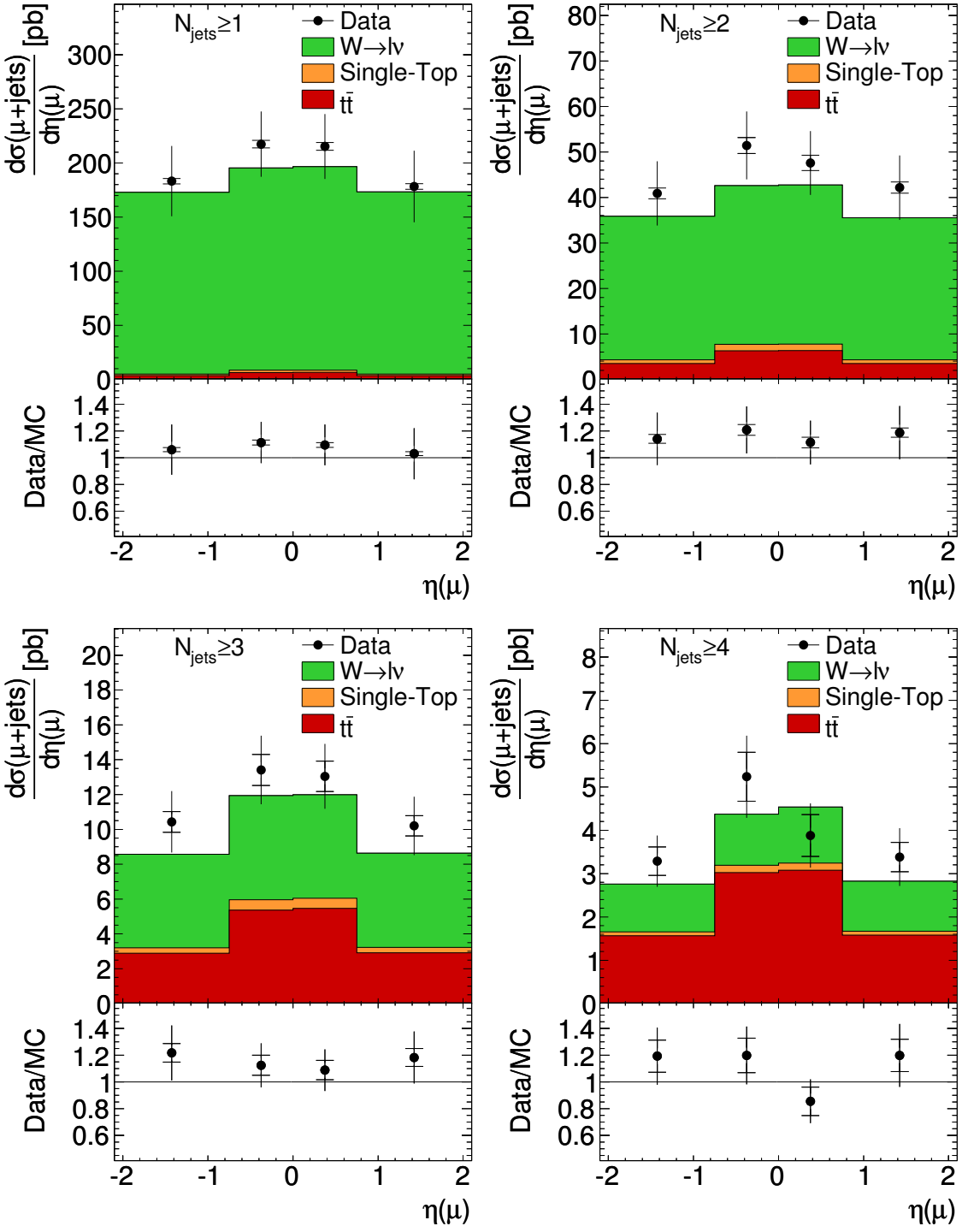

Figure 8.3: Differential muon+jets cross sections as a function of the η of the final state muon for different jet multiplicities. The inner error bars correspond to the statistical uncertainties while for the full error bars, systematic uncertainties are added in quadrature. The data points are compared to the prediction obtained by the simulation with the *MadGraph* Monte Carlo event generator.

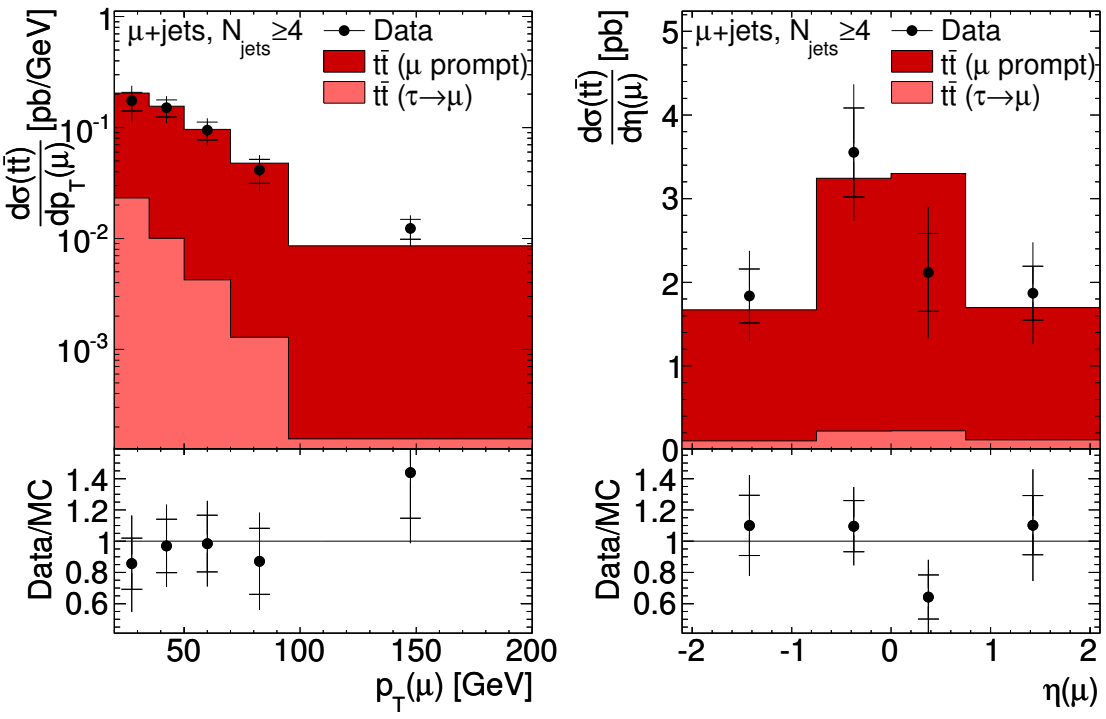

Figure 8.4: Differential $t\bar{t}$ cross section as a function of p_T and η of the final state muon in the muon+jets decay channel for the chosen phase space with at least four hard jets. The inner error bars correspond to the statistical uncertainties while for the full error bars, systematic uncertainties are added in quadrature. The data points are compared to the prediction obtained by the simulation with the *MadGraph* Monte Carlo event generator.

8.2 Top Pair Production Cross Section

The inclusive measurement for the $t\bar{t}$ production is performed in the muon+jets decay channel with a single muon and at least four jets in the phase space described in chapter 7. This cross section measurement yields

$$\sigma^{\mu+\text{jets}}_{\text{phase space}}(t\bar{t}) = 9.6 \pm 0.8(\text{stat.}) \pm 2.2(\text{syst.}) \pm 0.4(\text{lumi.})\,\text{pb}. \tag{8.1}$$

Based on the simulation using the *MadGraph* Monte Carlo event generator, this cross section is extrapolated to the whole phase space including all $t\bar{t}$ decay channels. The extrapolation results in

$$\sigma_{\text{total}}(t\bar{t}) = 172 \pm 15(\text{stat.}) \pm 41(\text{syst.}) \pm 7(\text{lumi.})\,\text{pb}. \tag{8.2}$$

Being slightly below but well within the uncertainties, the approximate NNLO prediction of 163^{+11}_{-10} pb [63] is in good agreement with the measured cross section value.

The differential cross section for the $t\bar{t}$ production in the muon+jets decay channel

with a single muon and at least four jets in the chosen phase space as a function of p_T and η of the final state muon are presented in figure 8.4. The comparison between data and simulation shows a good agreement within the uncertainties of about 30 %. This indicates a sound understanding of the dynamics and kinematics of the production process as it is implemented in the event generation by the *MadGraph* Monte Carlo event generator.

9 Summary and Outlook

This thesis was prepared during the start-up phase of the data taking at the LHC. In this phase, measurements in the top-quark sector are an interesting and important subject because they require a good understanding of all main detector parts and they are vital for the search for new phenomena beyond the SM.

The technical part of this thesis contributes to the understanding of the geometry of the tracker, which is one of the most essential parts of the CMS detector. The tracker geometry is determined via track-based alignment at CMS. However, the alignment algorithm is insensitive to certain systematic misalignment modes, called weak modes. In this thesis, a method is developed to determine momentum changing weak modes in the tracker geometry. This is done by comparing the momentum of an isolated hadron measured in the tracker with the corresponding energy measured in the calorimeter. Especially, the sensitivity to a twist of the tracker turned out to be valuable because the track-based alignment of the tracker seems to introduce such modes. By implementing the Z-boson mass constraint into the track-based alignment, the presented method currently remains the only independent validation method concerning momentum changing weak modes and is therefore very important for the ongoing alignment of the CMS tracker. It shows that the geometry obtained by the current alignment is twist free.

The analysis presented in this thesis is divided into two parts, namely the measurement of muon+jets cross sections and the measurement of $t\bar{t}$ cross sections. They are both determined inclusively and differentially in p_T and η of the muon in the final state. The phase space is defined by the kinematics of the muon ($|\eta| < 2.1$, $p_T > 20\,\text{GeV}$) and the kinematics of the jets ($|\eta| < 2.4$, $p_T > 30\,\text{GeV}$) in the final state. The analysis is performed with the complete dataset taken at a centre-of-mass energy of 7 TeV by the CMS detector at the LHC in the year 2010. These data correspond to an integrated luminosity of $35.9\,\text{pb}^{-1}$.

The muon+jets cross sections comprise $W \to l\nu$, $t\bar{t}$, and single-top events. They are measured for at least one, two, three, and four jets in the final state. With increasing jet multiplicity, the transition from a $W \to l\nu$ dominated to a strongly $t\bar{t}$ enriched phase space becomes evident. The dominating background in this measurement stems from QCD-multijet production, which is estimated from the data via a template fit to the distribution of the missing transverse energy. The largest systematic uncertainties originate from the estimation of the QCD-multijet event yield especially for low jet multiplicities and from the energy measurement of the jets. They are typically of the order of 10 % each. Within the uncertainties, the prediction obtained by the simulation with the *MadGraph* Monte Carlo event generator is consistent with the measurements. Nevertheless, the prediction is systematically below the data for all jet multiplicities indicating a slight underestimation of $W \to l\nu$ events. This measurement offers valuable

information about the W-boson production in association with jets and helps to estimate its contribution to the background in $t\bar{t}$ cross section measurements. It is published in reference [51] consolidating the top cross section determination therein.

The muon+jets measurement as a function of the jet multiplicity was dedicated to the early low luminosity phase of the LHC. It is hard to perform in future because single-muon triggers with a reasonably low momentum threshold have a rate that is too high to cope with already in the data taking of 2011. Instead, so-called 'cross triggers' will be used that already require at least three jets in addition to the muon, which contradicts the idea of measuring quantities as a function of the jet multiplicity.

The $t\bar{t}$ cross sections are measured in the lepton+jets decay channel with a muon in the final state. Additionally, the measured inclusive cross section is extrapolated to the whole phase space including all $t\bar{t}$ decay channels. The jet energy measurement is the most significant experimental uncertainty also in this measurement. A second significant source of uncertainty originates from the underlying theory used for the simulation of events as the definition of the renormalisation and factorisation scale yields a large uncertainty in the $W \to l\nu$ background modelling and in the extrapolation to the whole phase space. These two main sources yield about 15 % relative uncertainty each. The prediction for the inclusive $t\bar{t}$ production cross section obtained by NLO calculations is within the uncertainties in good agreement with the measurement. Also, the differential $t\bar{t}$ cross sections are well described by the prediction obtained by the simulation with the *MadGraph* Monte Carlo event generator indicating a good understanding of the kinematics and dynamics in $t\bar{t}$ events.

The differential cross section measurements presented here are a first step towards a variety of possible $t\bar{t}$ cross section measurements differential in kinematic quantities of the top quarks. These can serve to further test the production of $t\bar{t}$ events within the frame of perturbative QCD and hence the SM, and therefore are vital also for the search for new phenomena beyond the SM. These studies can be performed with the data collected in 2011 corresponding to an integrated luminosity of almost $5\,\mathrm{fb}^{-1}$. The higher statistics allow for applying tighter cuts to further reduce the background. The QCD-multijet background can be strongly suppressed by a higher momentum threshold for muons. A requirement for one or two jets being identified as originating from b quarks via their distance of flight before the decay would efficiently reduce $Z/\gamma^* \to l^+l^-$ and $W \to l\nu$ events. The measurement differential in η of the muon can then be used to estimate the remaining $W \to l\nu$ background due to the different shapes for $W \to l\nu$ and $t\bar{t}$ events.

Cross section measurements differential in kinematic properties of the top quarks are a natural further step but require more data than collected in 2010.

A E/p Distributions in Data

A.1 GR10_v3 Geometry

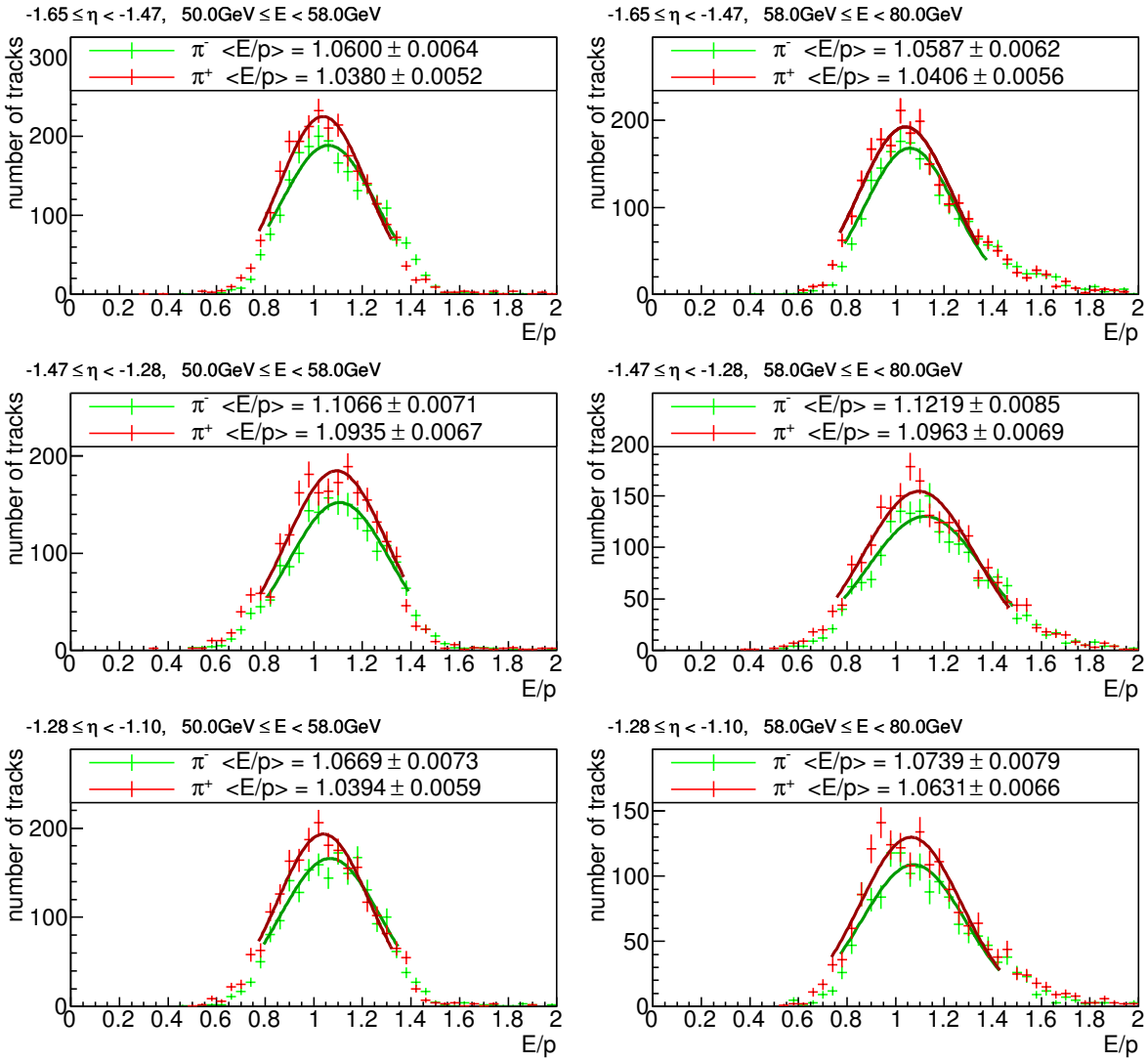

Figure A.1: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v3* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

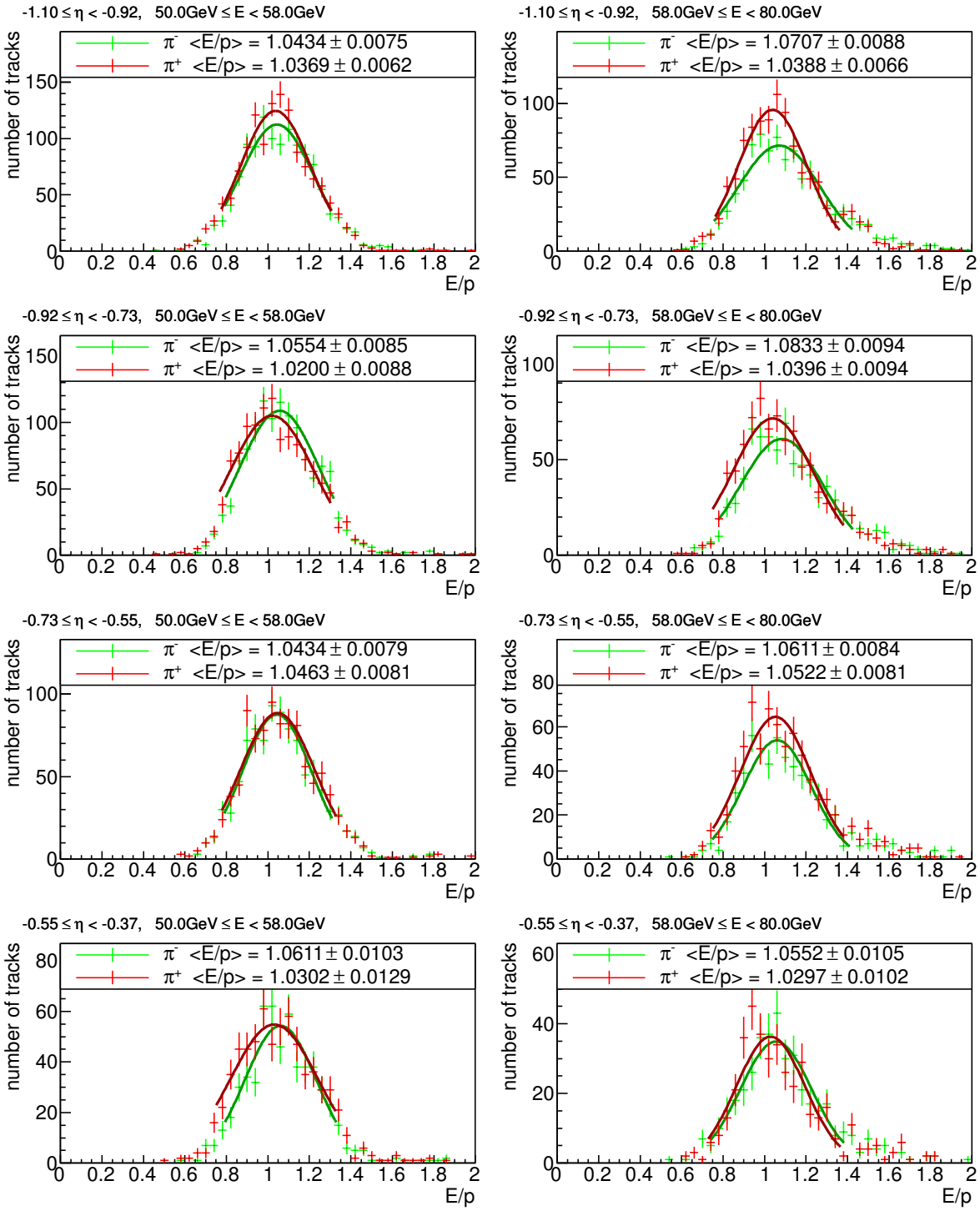

Figure A.2: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v3* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

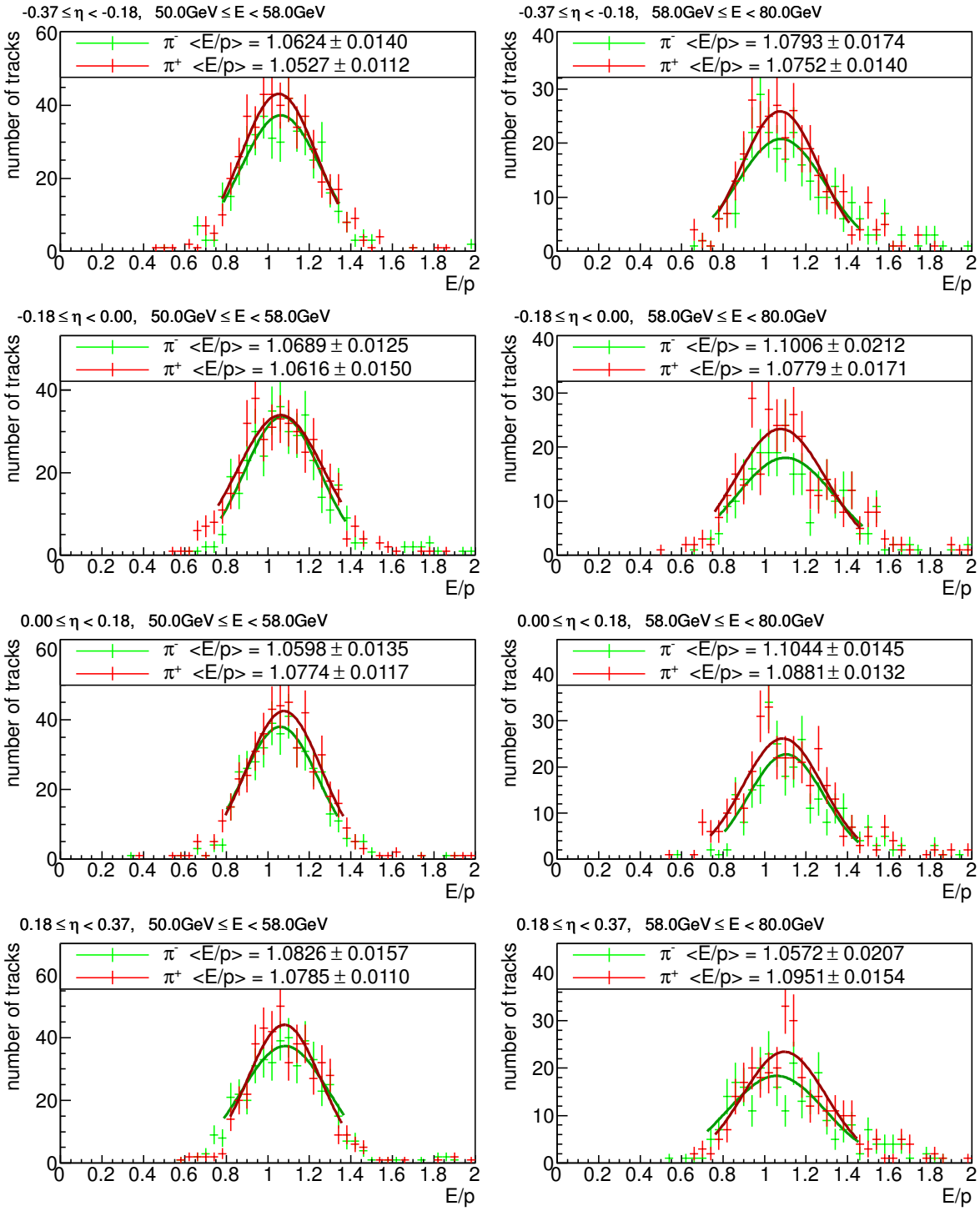

Figure A.3: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v3* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

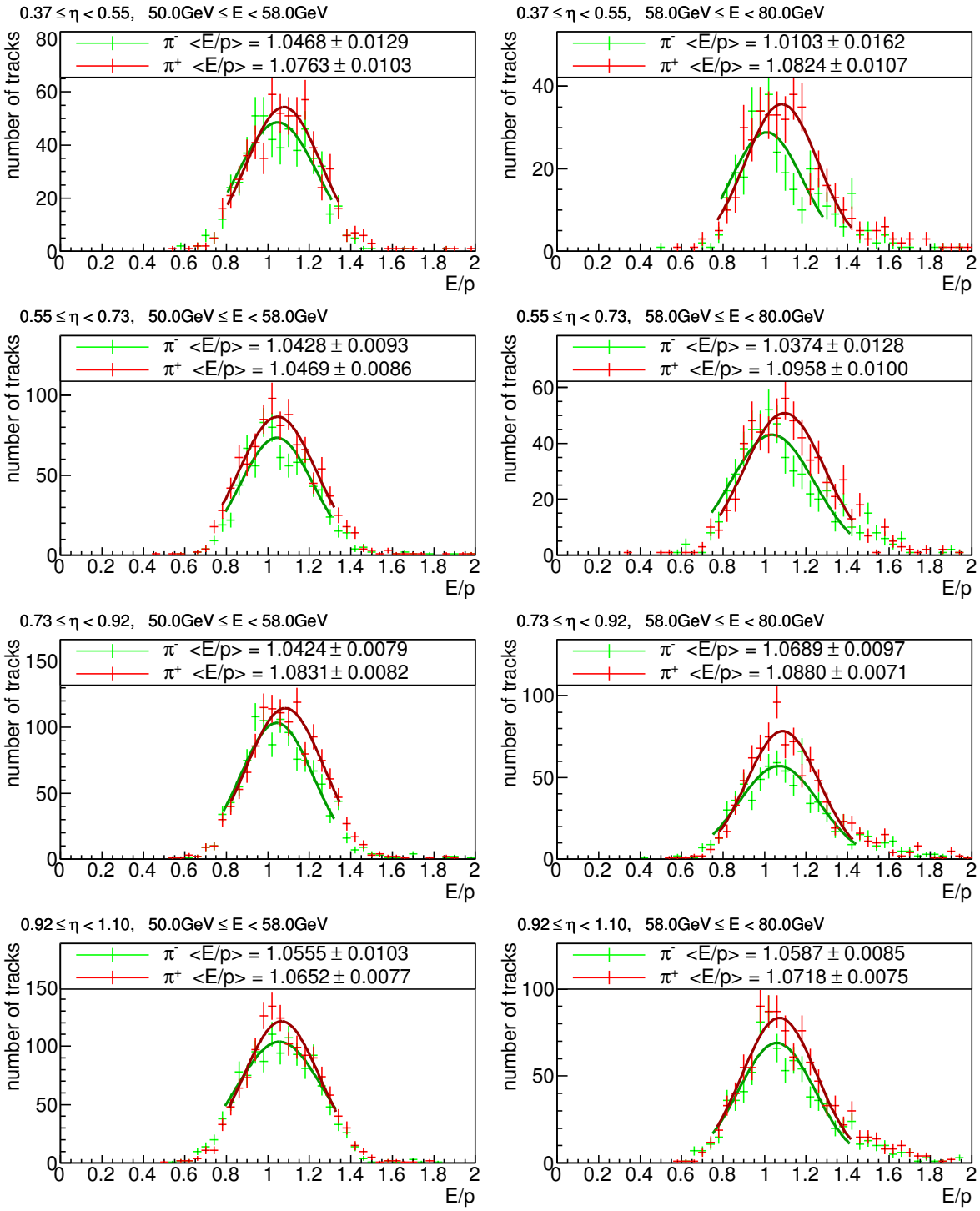

Figure A.4: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v3* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

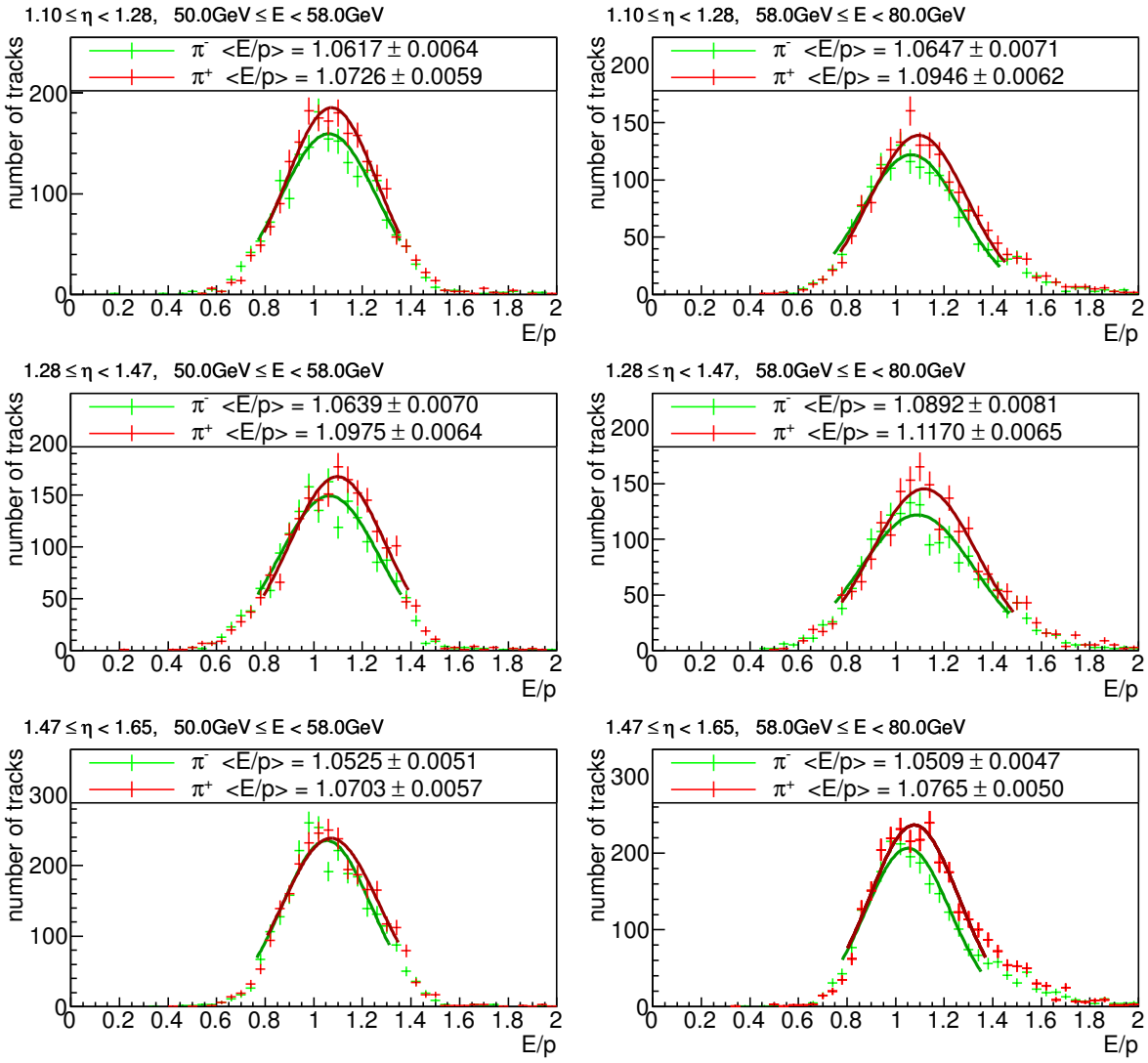

Figure A.5: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v3* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

A.2 GR10_v4 Geometry

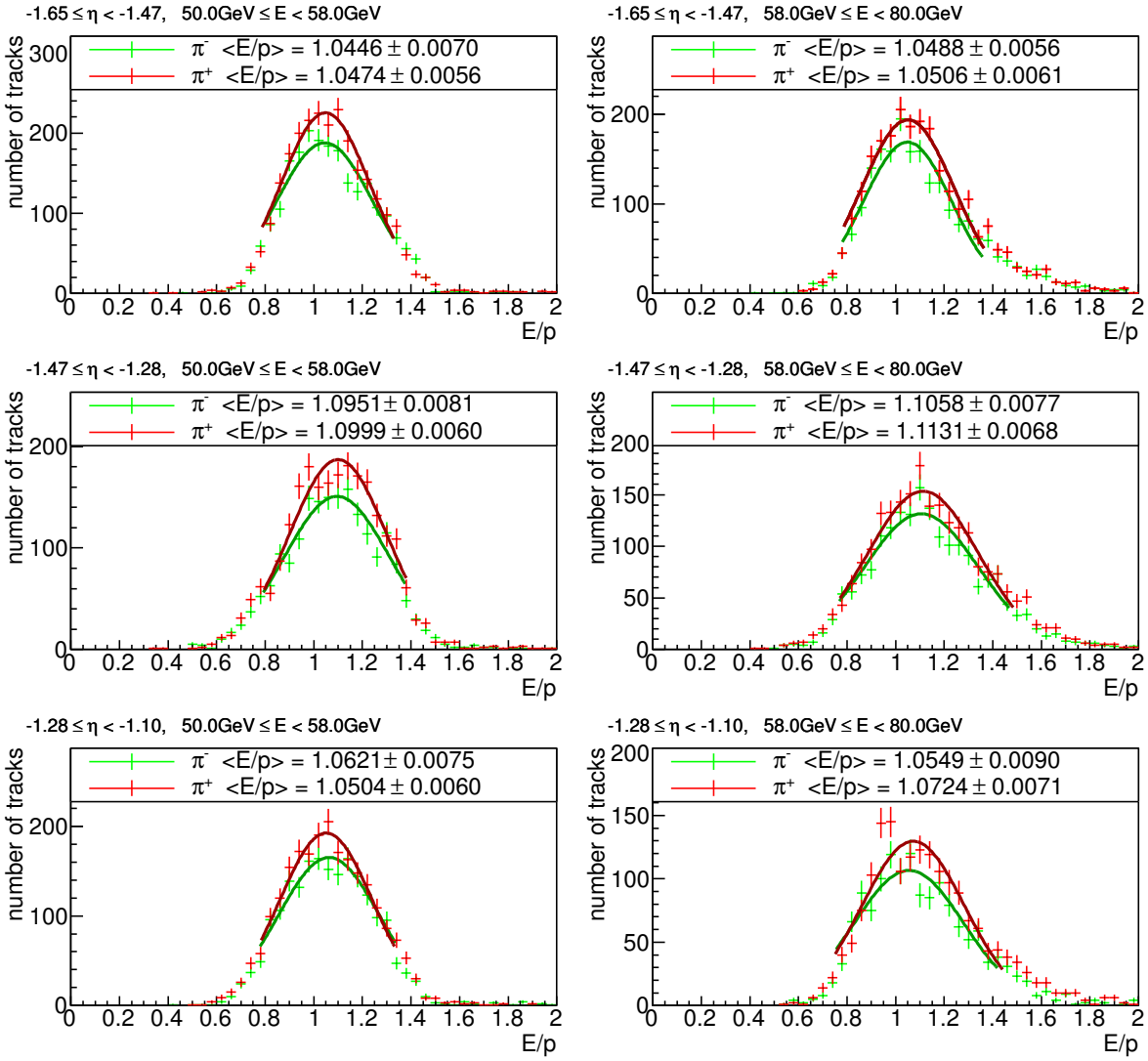

Figure A.6: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v4* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

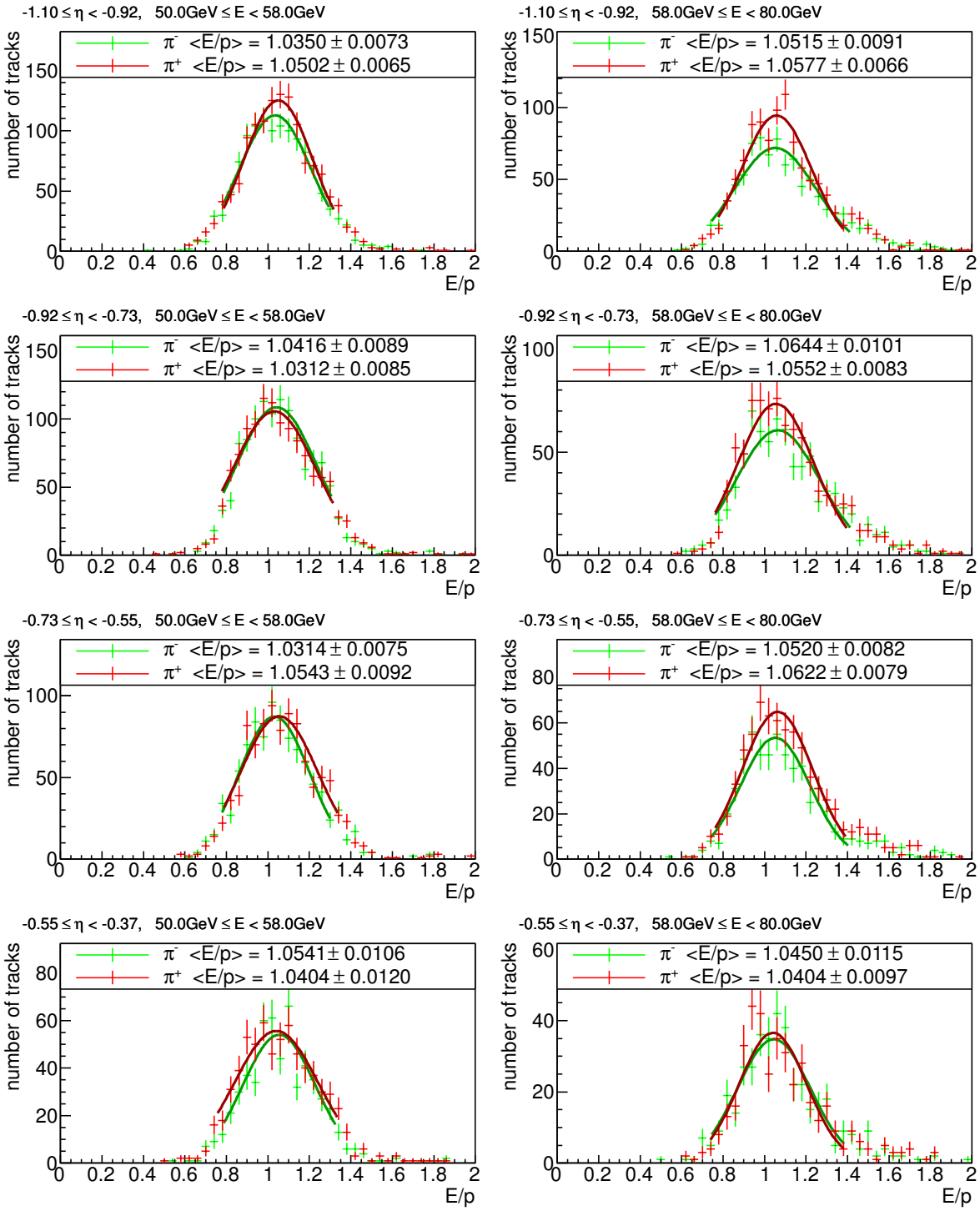

Figure A.7: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v4* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

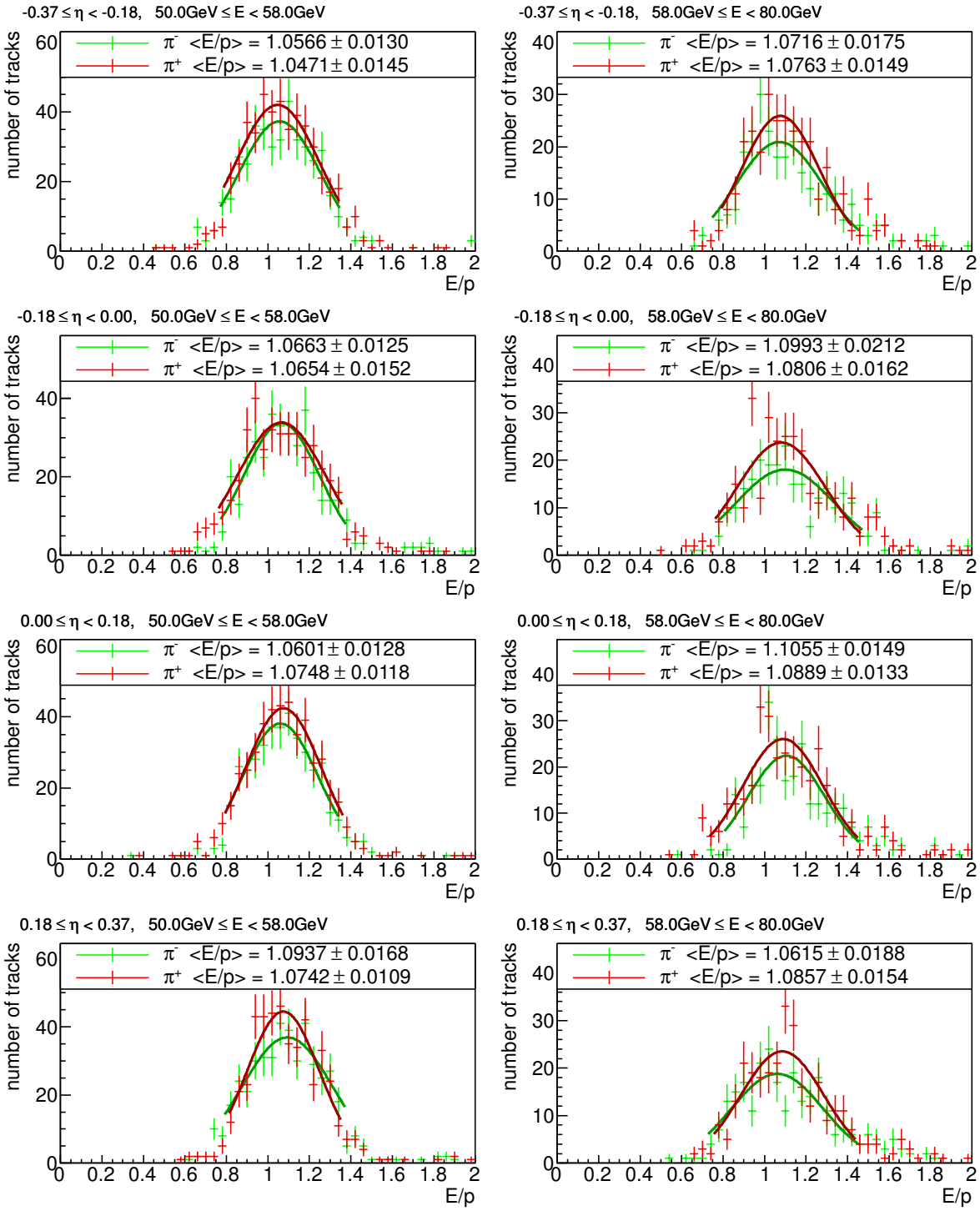

Figure A.8: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v4* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

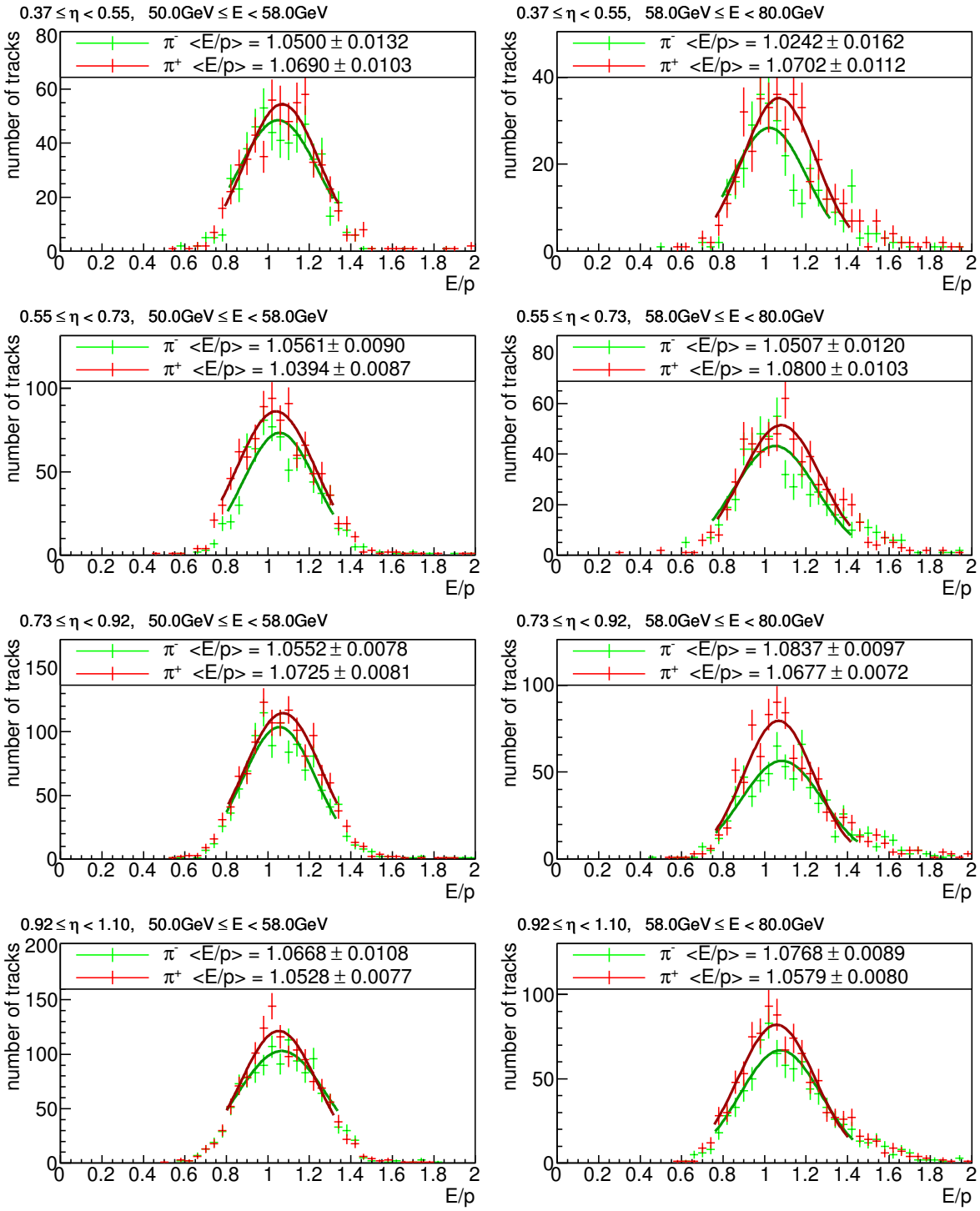

Figure A.9: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v4* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

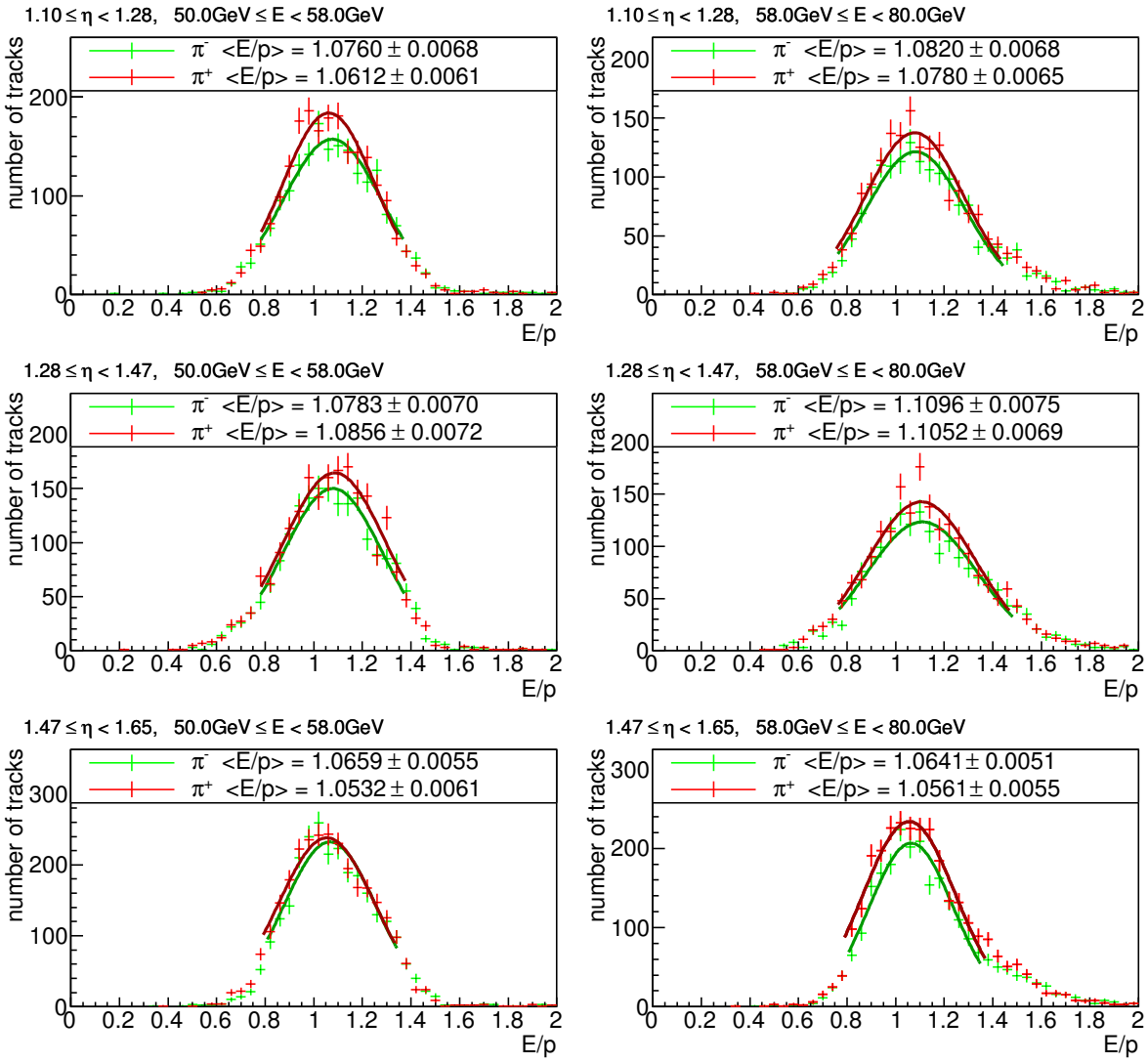

Figure A.10: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v4* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

A.3 GR10_v5 Geometry

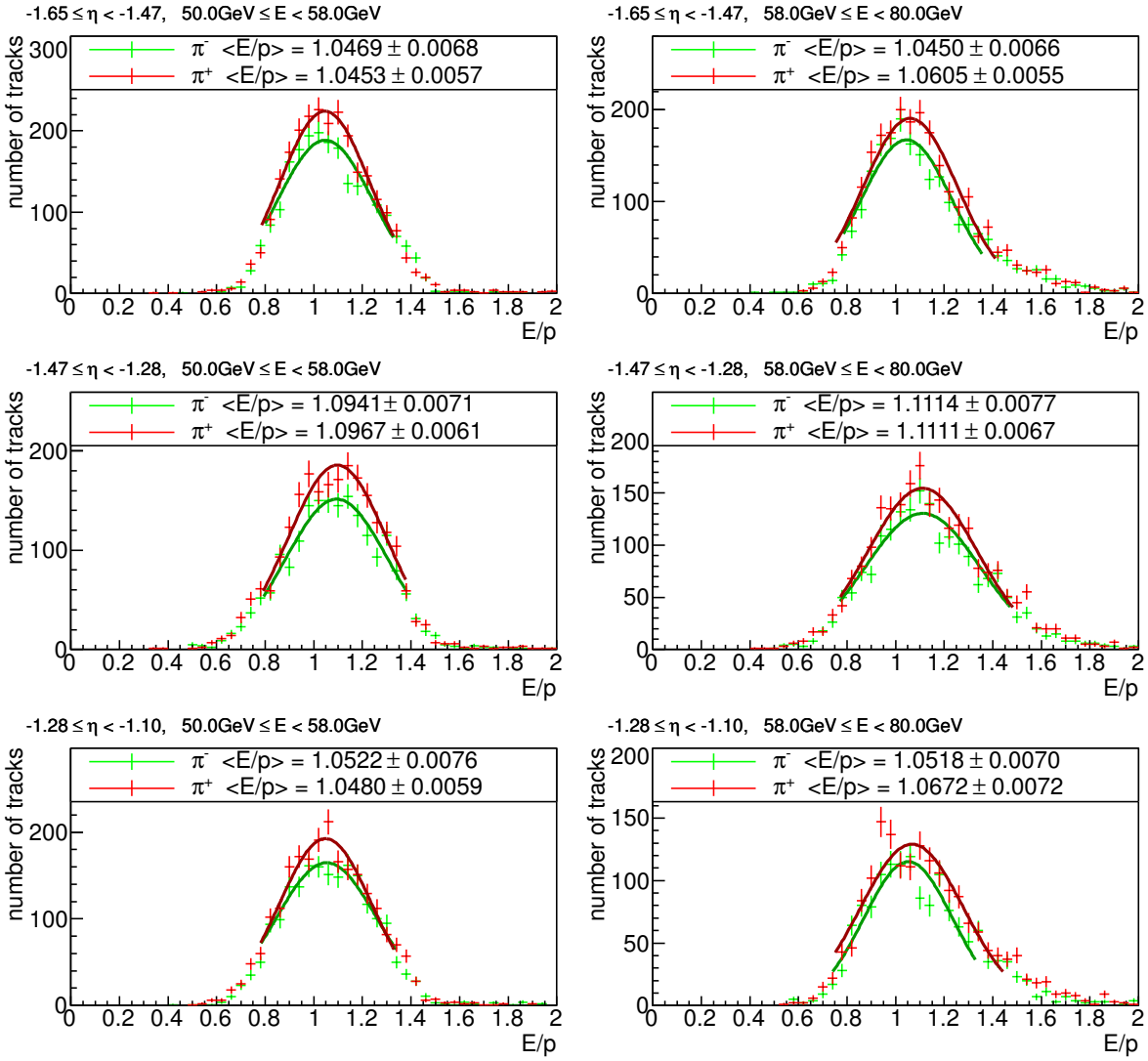

Figure A.11: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v5* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

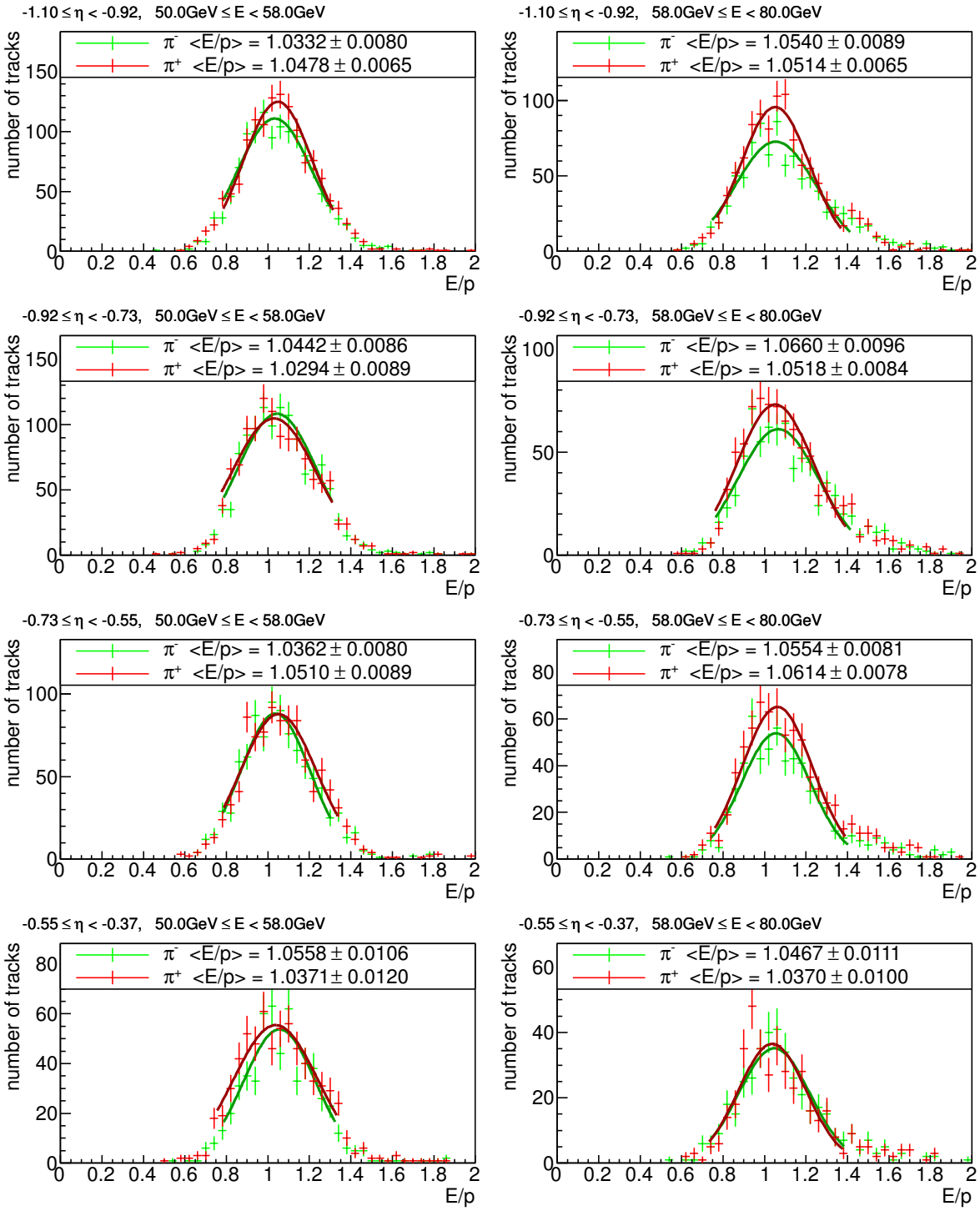

Figure A.12: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v5* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

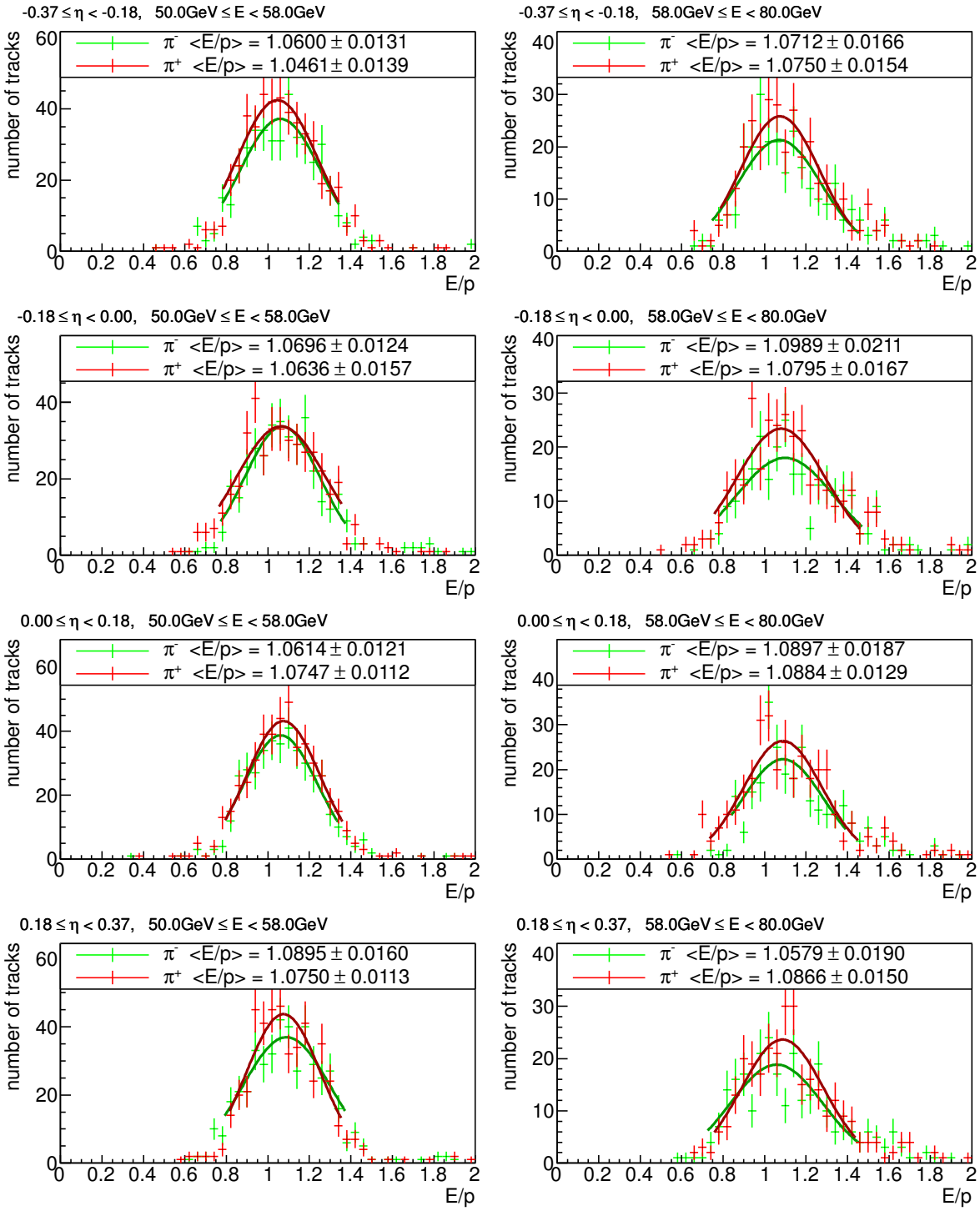

Figure A.13: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v5* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

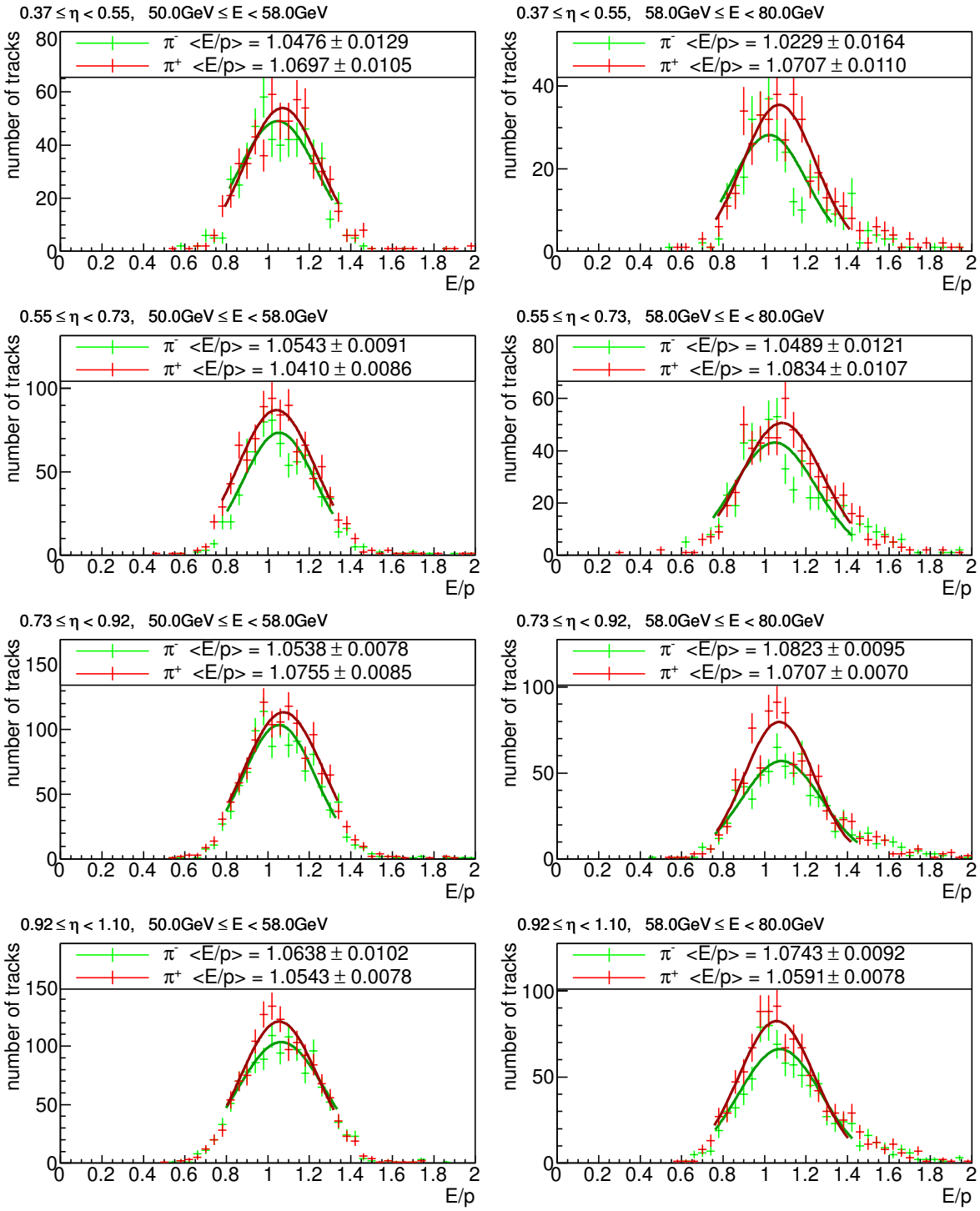

Figure A.14: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v5* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

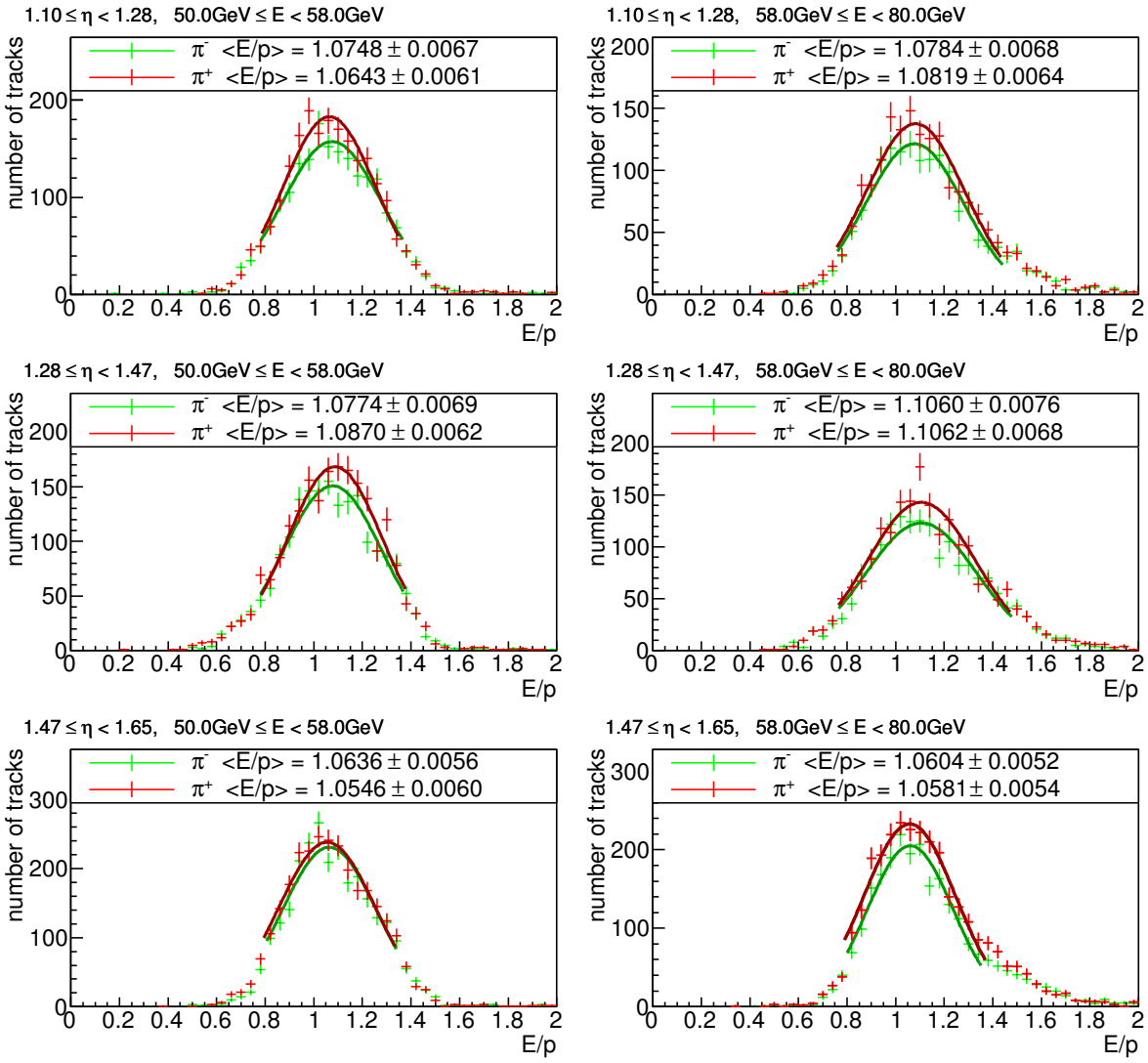

Figure A.15: Energy over momentum distributions of negatively (green points) and positively (red points) charged particles in data for the *GR10_v5* geometry. The mean of a Gaussian fit to each distribution is specified in the legend. Columns are for different energy intervals and rows for different pseudorapidity intervals as given in the upper left corner.

Bibliography

[1] H. Spiesberger, M. Spira, and P. Zerwas, "The Standard model: Physical basis and scattering experiments", arXiv:hep-ph/0011255.

[2] A. Einstein, "The Foundation of the General Theory of Relativity", *Annalen Phys.* **49** (1916) 769–822. doi:10.1002/andp.200590044, 10.1002/andp.200590044.

[3] S. Glashow, "Partial Symmetries of Weak Interactions", *Nucl.Phys.* **22** (1961) 579–588. doi:10.1016/0029-5582(61)90469-2.

[4] S. Weinberg, "A Model of Leptons", *Phys.Rev.Lett.* **19** (1967) 1264–1266. doi:10.1103/PhysRevLett.19.1264.

[5] D. Gross and F. Wilczek, "Asymptotically Free Gauge Theories. 1", *Phys.Rev.* **D8** (1973) 3633–3652. doi:10.1103/PhysRevD.8.3633.

[6] P. W. Higgs, "Broken symmetries, massless particles and gauge fields", *Phys.Lett.* **12** (1964) 132–133. doi:10.1016/0031-9163(64)91136-9.

[7] CMS Collaboration, "Search for standard model Higgs boson in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ and integrated luminosity up to $1.7\,\text{fb}^{-1}$", *CMS Physics Analysis Summary* **CMS-PAS-HIG-11-022** (2011).

[8] J. Vossebeld, "Standard Model Higgs Boson Searches in ATLAS", Technical Report ATL-PHYS-PROC-2011-102, CERN, Geneva, (Sep, 2011).

[9] M. Kobayashi and T. Maskawa, "CP-Violation in the Renormalizable Theory of Weak Interaction", *Progress of Theoretical Physics* **49** (1973), no. 2, 652–657. doi:10.1143/PTP.49.652.

[10] CDF Collaboration, "Observation of top quark production in $\bar{p}p$ collisions", *Phys.Rev.Lett.* **74** (1995) 2626–2631, arXiv:hep-ex/9503002. doi:10.1103/PhysRevLett.74.2626.

[11] D0 Collaboration, "Observation of the top quark", *Phys.Rev.Lett.* **74** (1995) 2632–2637, arXiv:hep-ex/9503003. doi:10.1103/PhysRevLett.74.2632.

[12] Particle Data Group Collaboration, "Review of particle physics", *J. Phys.* **G37** (2010) 075021. doi:10.1088/0954-3899/37/7A/075021.

[13] Tevatron Electroweak Working Group Collaboration, "Combination of CDF and D0 results on the mass of the top quark using up to 5.8 fb-1 of data", arXiv:1107.5255.

[14] CMS Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s} = 7\,\text{TeV}$", *CMS Physics Analysis Summary* **CMS-PAS-MUO-10-002** (2010).

[15] J. Draeger, "Track Based Alignment of the CMS Silicon Tracker and its Implication on Physics Performance". PhD thesis, University of Hamburg, Germany, 2011.

[16] L. Evans, (ed.) and P. Bryant, (ed.), "LHC Machine", *JINST* **3** (2008) S08001. doi:10.1088/1748-0221/3/08/S08001.

[17] CMS Collaboration, "CMS physics: Technical design report",. CERN-LHCC-2006-001.

[18] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004. doi:10.1088/1748-0221/3/08/S08004.

[19] C. Lefèvre, "The CERN accelerator complex. Complexe des accélérateurs du CERN", (Dec, 2008).

[20] USCMS Collaboration, "The CMS barrel calorimeter response to particle beams from 2-GeV/c to 350-GeV/c", *Eur. Phys. J.* **C60** (2009) 359–373. doi:10.1140/epjc/s10052-009-0959-5.

[21] CMS Collaboration, "Measurement of CMS luminosity", *CMS Physics Analysis Summary* **CMS-PAS-EWK-10-004** (2010).

[22] P. Lenzi, C. Genta, and B. Mangano, "Track reconstruction of real cosmic muon events with CMS tracker detector", *J. Phys. Conf. Ser.* **119** (2008) 032030. doi:10.1088/1742-6596/119/3/032030.

[23] R. Fruhwirth, "Application of Kalman filtering to track and vertex fitting", *Nucl. Instrum. Meth.* **A262** (1987) 444–450. doi:10.1016/0168-9002(87)90887-4.

[24] V. Blobel, "Software alignment for tracking detectors", *Nucl. Instrum. Meth.* **A566** (2006) 5–13. doi:10.1016/j.nima.2006.05.157.

[25] V. Blobel, "A new fast track-fit algorithm based on broken lines", *Nucl. Instrum. Meth.* **A566** (2006) 14–17. doi:10.1016/j.nima.2006.05.156.

[26] V. Blobel, C. Kleinwort, and F. Meier, "Fast alignment of a complex tracking detector using advanced track models", *Comput. Phys. Commun.* **182** (2011) 1760–1763, arXiv:1103.3909. doi:10.1016/j.cpc.2011.03.017.

[27] CMS Collaboration, "Alignment of the CMS Silicon Tracker during Commissioning with Cosmic Rays", *JINST* **5** (2010) T03009, arXiv:0910.2505. doi:10.1088/1748-0221/5/03/T03009.

[28] D. N. Brown, A. V. Gritsan, Z. J. Guo et al., "Local Alignment of the BABAR Silicon Vertex Tracking Detector", *Nucl. Instrum. Meth.* **A603** (2009) 467–484, arXiv:0809.3823. doi:10.1016/j.nima.2009.02.001.

[29] W. Bernreuther, "Top quark physics at the LHC", *J. Phys.* **G35** (2008) 083001, arXiv:0805.1333. doi:10.1088/0954-3899/35/8/083001.

[30] M. Beneke and V. M. Braun, "Heavy quark effective theory beyond perturbation theory: Renormalons, the pole mass and the residual mass term", *Nucl. Phys.* **B426** (1994) 301–343, arXiv:hep-ph/9402364. doi:10.1016/0550-3213(94)90314-X.

[31] I. I. Y. Bigi, M. A. Shifman, N. G. Uraltsev et al., "The Pole mass of the heavy quark. Perturbation theory and beyond", *Phys. Rev.* **D50** (1994) 2234–2246, arXiv:hep-ph/9402360. doi:10.1103/PhysRevD.50.2234.

[32] K. G. Chetyrkin and M. Steinhauser, "The relation between the MS-bar and the on-shell quark mass at order alpha(s)**3", *Nucl. Phys.* **B573** (2000) 617–651, arXiv:hep-ph/9911434. doi:10.1016/S0550-3213(99)00784-1.

[33] H1 and ZEUS Collaboration, "Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA", *JHEP* **1001** (2010) 109, arXiv:0911.0884. doi:10.1007/JHEP01(2010)109.

[34] A. Lenz et al., "Anatomy of New Physics in B-Bbar mixing", *Phys. Rev.* **D83** (2011) 036004, arXiv:1008.1593. doi:10.1103/PhysRevD.83.036004.

[35] ALEPH Collaboration, "Precision Electroweak Measurements and Constraints on the Standard Model", arXiv:1012.2367.

[36] V. Zacek, "Dark Matter", arXiv:0707.0472. Lake Louise Winter Institute 2007.

[37] CMS Collaboration, "Search for Resonances in Semi-leptonic Top-pair Decays Close to Production Threshold", *CMS Physics Analysis Summary* **CMS-PAS-TOP-10-007** (2011).

[38] J. H. Kuhn and G. Rodrigo, "Charge asymmetry of heavy quarks at hadron colliders", *Phys.Rev.* **D59** (1999) 054017, arXiv:hep-ph/9807420. doi:10.1103/PhysRevD.59.054017.

[39] CDF Collaboration Collaboration, "Evidence for a Mass Dependent Forward-Backward Asymmetry in Top Quark Pair Production", *Phys.Rev.* **D83** (2011) 112003, arXiv:1101.0034. doi:10.1103/PhysRevD.83.112003.

[40] D0 Collaboration Collaboration, "Forward-backward asymmetry in top quark-antiquark production", arXiv:1107.4995.

[41] CMS Collaboration, "Measurement of the Charge Asymmetry in Top Quark Pair Production", *CMS Physics Analysis Summary* **CMS-PAS-TOP-11-014** (2011).

[42] ATLAS Collaboration, "Measurement of the charge asymmetry in top quark pair production in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ using the ATLAS detector", Technical Report ATLAS-CONF-2011-106, CERN, Geneva, (Aug, 2011).

[43] J. Alwall et al., "MadGraph/MadEvent v4: The New Web Generation", *JHEP* **09** (2007) 028, arXiv:0706.2334. doi:10.1088/1126-6708/2007/09/028.

[44] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual", *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175. doi:10.1088/1126-6708/2006/05/026.

[45] M. L. Mangano, M. Moretti, F. Piccinini et al., "Matching matrix elements and shower evolution for top- quark production in hadronic collisions", *JHEP* **01** (2007) 013, arXiv:hep-ph/0611129. doi:10.1088/1126-6708/2007/01/013.

[46] R. Field, "Early LHC Underlying Event Data - Findings and Surprises", arXiv:1010.3558.

[47] J. Allison et al., "Geant4 developments and applications", *IEEE Trans. Nucl. Sci.* **53** (2006) 270. doi:10.1109/TNS.2006.869826.

[48] J. M. Campbell and R. K. Ellis, "MCFM for the Tevatron and the LHC", *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 10–15, arXiv:1007.3492. doi:10.1016/j.nuclphysbps.2010.08.011.

[49] N. Kidonakis, "NNLL resummation for s-channel single top quark production", *Phys. Rev.* **D81** (2010) 054028, arXiv:1001.5034. doi:10.1103/PhysRevD.81.054028.

[50] K. Melnikov and F. Petriello, "Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_s^2)$", *Phys. Rev.* **D74** (2006) 114017, arXiv:hep-ph/0609070. doi:10.1103/PhysRevD.74.114017.

[51] CMS Collaboration, "Measurement of the $t\bar{t}$ Production Cross Section in pp Collisions at $\sqrt{s} = 7\,\text{TeV}$ using the Kinematic Properties of Events with Leptons and Jets", *Eur. Phys. J.* **C71** (2011) 1721, arXiv:1106.0902. doi:10.1140/epjc/s10052-011-1721-3.

[52] CMS Collaboration, "Performance of muon reconstruction and identification in pp collisions at $\sqrt{s} = 7\,\text{TeV}$", *CMS Physics Analysis Summary* **CMS-PAS-MUO-10-004** (2011).

[53] CMS Collaboration, "Tracking and Primary Vertex Results in First 7 TeV Collisions", *CMS Physics Analysis Summary* **CMS-PAS-TRK-10-005** (2010).

[54] R. Fruhwirth, W. Waltenberger, and P. Vanlaer, "Adaptive vertex fitting", *J. Phys.* **G34** (2007) N343. doi:10.1088/0954-3899/34/12/N01.

[55] S. Baffioni et al., "Electron reconstruction in CMS", *Eur. Phys. J.* **C49** (2007) 1099–1116. doi:10.1140/epjc/s10052-006-0175-5.

[56] CMS Collaboration, "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7\,\text{TeV}$", *JHEP* **01** (2011) 080, arXiv:1012.2466. doi:10.1007/JHEP01(2011)080.

[57] CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET", *CMS Physics Analysis Summary* **CMS-PAS-PFT-09-001** (Apr, 2009).

[58] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, arXiv:0802.1189. doi:10.1088/1126-6708/2008/04/063.

[59] CMS Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS", *JINST* **6** (2011) P11002, arXiv:1107.4277. doi:10.1088/1748-0221/6/11/P11002.

[60] CMS Collaboration, "Charged particle multiplicities in pp interactions at $\sqrt{s} = 0.9$, 2.36, and $7\,\text{TeV}$", *JHEP* **1101** (2011) 079, arXiv:1011.5531. doi:10.1007/JHEP01(2011)079.

[61] G. Cowan, "A survey of unfolding methods for particle physics", 2002. Prepared for Conference on Advanced Statistical Techniques in Particle Physics, Durham, England, 18-22 Mar, www.ippp.dur.ac.uk/Workshops/02/statistics/papers/cowan_gdc_talk.ps.gz.

[62] CMS Collaboration, "Absolute luminosity normalization", *CMS Detector Performance Summary* **CMS-DP-2011-002** (Mar, 2011).

[63] N. Kidonakis, "Next-to-next-to-leading soft-gluon corrections for the top quark cross section and transverse momentum distribution", *Phys.Rev.* **D82** (2010) 114030, arXiv:1009.4935. doi:10.1103/PhysRevD.82.114030.

Danksagung

Die Arbeit ist geschrieben und ich blicke zurück auf eine tolle Zeit. Es gibt Viele, bei denen ich mich bedanken möchte, weil sie dazu ihren Beitrag geleistet haben.

Prof. Peter Schleper danke ich dafür, dass er mit seinem ansteckenden Enthusiasmus schon im fünften Semester maßgeblich daran beteiligt war, mein Interesse an der Teilchenphysik zu wecken und mich später in seine Arbeitsgruppe aufgenommen hat.

Prof. Johannes Haller danke ich für das Erstellen des Zweitgutachtens der Dissertation und Prof. Dieter Horns und PD Dr. Thomas Schörner-Sadenius dafür, dass sie sich bereit erklärt haben, die Disputation zu begutachten.

Ein großer Dank geht an meine beiden Betreuer, Dr. Gero Flucke bei der Alignmentstudie und Dr. Roger Wolf bei der Topquarkanalyse. Ich habe viel von euch gelernt und ihr wart auch noch für mich da, als ihr die Gruppe schon verlassen hattet.

Ich danke auch allen Postdocs, die mir Kommentare zu einzelnen Kapiteln meiner Arbeit gegeben haben, Dr. Gordon Kaussen, Dr. Martijn Gosselink und ganz speziell Dr. Thomas Hermanns, der die meisten Kapitel gegengelesen und schnell und konstruktiv geantwortet hat.

Allen Kollegen aus der Arbeitsgruppe danke ich für die angenehme Arbeitsatmosphäre und den freundschaftlichen Umgang auch in der Freizeit. Die Gruppe ist zu groß, um alle namentlich zu nennen. Daher hebe ich einige stellvertretend hervor: Friederike Nowak und Matthias Schröder wegen der vielen anregenden Gespräche im Büro, die nicht nur meinen physikalischen Horizont erweitert haben. Martin Görner und Jörn Lange für den „Approval“-Prozess, den wir gemeinsam mit Roger erfolgreich durchgestanden haben; wie groß der Streß zeitweise auch war, es war immer eine Freude, mit euch zusammen zu arbeiten. Dr. Sebastian Naumann-Emme und Eike Schlieckau für die aufschlussreichen Diskussionen über Topquarkphysik. Niklas Pietsch und Matthias Stein für den gemeinsamen Sport und andere außerdienstliche Aktivitäten.

Außerdem bin ich dankbar für die Ablenkung von der Arbeit, wann immer es notwendig war, um wieder den Kopf frei zu bekommen. Dieser Dank gilt meinen Freunden und in letzter Zeit vor allem Jette.

Zu guter Letzt danke ich meiner Familie für die bedingungslose Unterstützung. Es ist gut zu wissen, dass es immer jemanden gibt, auf den man sich verlassen kann.

Danke!