Measurement of Top-Quark-Pair Differential Cross Sections in Proton-Proton Collisions at $\sqrt{s} = 8\,\text{TeV}$ with the CMS Experiment

Dissertation with the aim of achieving a doctoral degree
at the Faculty of Mathematics, Informatics and Natural Sciences
Department of Physics
of Universität Hamburg

submitted by

Ivan Asin Cruz

2014 in Hamburg

Day of oral defense: 2014/12/16
The following evaluators recommended the admission of the dissertation:

PD Dr. A. B. Meyer

Prof. Dr. J. Haller

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

Hamburg, den 2014/11/01 Unterschrift

Dedicado
a mis padres, a mi hermano y a mi familia
por todo vuestro apoyo en este largo viaje.
Que nunca falte.

To Annika
for your patience.

Muchas gracias.
Vielen Dank.

Abstract

The normalised differential cross sections for the production of top-quark pairs are measured using proton-proton collisions at centre-of-mass energy $\sqrt{s} = 8\,\text{TeV}$. Events are delivered at the Large Hadron Collider and recorded with the Compact Muon Solenoid experiment in 2012, corresponding to a total integrated luminosity of $L = 19.7\,\text{fb}^{-1}$. The analysis is performed in the dilepton (e^+e^-, $\text{e}^\pm\mu^\mp$ and $\mu^+\mu^-$) top-quark pair ($\text{t}\bar{\text{t}}$) decay channels. Events are selected based on topological requirements of the final state particles and a kinematic reconstruction is performed to determine the individual top-quark 4-momenta. The differential cross sections are measured as a function of kinematic variables of the top quark, the $\text{t}\bar{\text{t}}$ system and their decay products (leptons, lepton pairs, b-jets and b-jet pairs). The reconstructed distributions are corrected for detector effects by use of a regularised unfolding technique. The results are normalised to the total production cross section measured in the same kinematic range. Therefore, systematic uncertainties correlated among the bins cancel and a typical precision of 3–5% is achieved. The results are compared to Standard Model (SM) predictions from various Monte Carlo event generators and several QCD perturbation theory calculations up to approximate next-to-next-to-leading order accuracy. In general, the results agree with the SM expectations and the achieved precision is able to discriminate between certain theory models.

Zusammenfassung

Normierte differentielle Wirkungsquerschnitte von Top-Quark-Paarproduktion werden in Proton-Proton-Kollisionen des Large Hadron Colliders aus dem Jahre 2012 bei einer Schwerpunktsenergie von $\sqrt{s} = 8\,\text{TeV}$ gemessen, welche mit dem CMS-Experiment aufgezeichnet wurden und einer Gesamtluminosität von $L = 19.7\,\text{fb}^{-1}$ entsprechen. Analysiert werden die di-leptonischen (e^+e^-, $\text{e}^\pm\mu^\mp$ und $\mu^+\mu^-$) Zerfallskanäle des Top-Quark-Paares ($\text{t}\bar{\text{t}}$), basierend auf einer der Topologie des Endzustandes entsprechenden Ereignisselektion, unter Verwendung einer kinematischen Rekonstruktion der beiden Top-Quark-Vierervektoren. Die Messungen erfolgen als Funktion kinematischer Observablen des Top-Quarks, des $\text{t}\bar{\text{t}}$-Systems und der Zerfallsprodukte (Leptonen, Lepton-Paare, b-Jets und b-Jet-Paare). Die Verteilungen der gemessenen Größen werden bezüglich Detektoreffekten mit Hilfe einer regularisierten Entfaltungsmethode korrigiert. Durch Normierung der Resultate auf den totalen Produktionswirkungsquerschnitt – gemessen im selben kinematischen Bereich – kürzen sich zwischen Einzelintervallen korrelierte Unsicherheiten, wodurch eine typische Unsicherheit von 3–5% erreicht wird. Die Resultate werden mit Vorhersagen des Standardmodells (SM) diverser Monte-Carlo-Simulationen, und perturbativen QCD-Rechnungen in approximierter dritter Ordnung verglichen. Im Allgemeinen stimmen die Ergebnisse mit den SM-Erwartungen überein, ferner ermöglicht die erreichte Präzision die Unterscheidung zwischen verschiedenen Theoriemodellen.

Contents

Chapter 1

Introduction

The Standard Model (SM) of particle physics is one of the most successful theories ever developed. It has undergone and survived a vast set of experimental tests for almost 50 years. The latest great accomplishment of the SM was the discovery of a boson compatible with the properties of the Higgs boson at the Large Hadron Collider (LHC) in 2012 [1,2]. Nevertheless, there are experimental observations with no explanation within the SM such as dark matter, dark energy and matter-antimatter imbalance. Due to the enormous collision energy and high collision rate delivered by the LHC collider, the experiments located at it can now test and evaluate physics processes and parameters of the SM with unprecedented precision and energy and search for signals of physics beyond the SM (BSM).

A prominent constituent of the SM is the top quark, the heaviest fundamental particle ever measured. The high value of the top-quark mass and its closeness to the electroweak scale makes the top-quark analysis sensitive to possible new physics effects. Similarly, results of top-quark analyses, such as differential and inclusive cross sections, can be used to constrain theory parameters, e.g.: the parton density functions [3], the strong coupling constant [4], etc. The top-quark physics programme has been a very active and productive research field since it was discovered in 1995 [5,6]. In the year 2012 the LHC produced approximately five million top-quark pairs ($t\bar{t}$), making possible a detailed study of the top quark and its properties, confirming the validity of the perturbative QCD sector of the SM and setting more stringent limits to the presence of BSM physics processes [7].

In this analysis the normalised differential cross section for the production of $t\bar{t}$-pairs in proton-proton (pp) collisions at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ is measured. The collision events were delivered by the LHC accelerator during the year 2012 and collected by the Compact Muon Solenoid (CMS) experiment corresponding to approximately $20\,\text{fb}^{-1}$ of total integrated data luminosity. The analysis is performed in the dileptonic $t\bar{t}$ decay channel characterised by the presence of two high-p_{T} oppositely-charged leptons, two b-quarks and two neutrinos which escape direct detection. The collected dataset is filtered with the application of several kinematic requirements based on the $t\bar{t}$ dilepton event topologies such that a large signal-to-background fraction of events is selected. The individual top-quark 4-momentum components are derived from a kinematic reconstruction algorithm applied to the remaining events after imposing the complete event selection criteria. Differential cross section distributions are measured as a function of 19 kinematic variables of the top quark, top-quark pair and $t\bar{t}$ decay products and are presented in two alternative phase spaces defined by the detector acceptance. The results are compared to the predictions obtained from event simulation tools and higher-order theory calculations, wherever available.

This thesis is organised as follows: A brief introduction on the theory framework of the

SM of particle physics, the properties of the top quark and the relevance of the top quark in the SM and theories beyond are outlined in Chapter 2.

Chapter 3 is devoted to the description of the experimental apparatus: the LHC collider and the CMS experiment. The certification process of the quality of the collected data used in physics analysis is also introduced in this chapter.

A complete pp collision cannot be analytically described by the current theories, and thus event simulations based on theory modelling assumptions are used to provide a representation of the pp collision. Event simulations, introduced in Chapter 4, are important analysis tools which are also used to estimate the analysis efficiency and to calibrate the detector response to the different particles produced in the events under investigation.

Details of event reconstruction and selection including the top-quark momentum reconstruction algorithm are described in Chapter 5.

In Chapter 6 the differential cross section is defined. The unfolding technique used to correct the bin-to-bin migrations caused by the detector is described in the same chapter. The total $t\bar{t}$ production cross section is measured and used as consistency check of the analysis methodology with dedicated experimental measurements and with the most precise theory predictions.

A precision measurement can only be obtained with a detailed study of the systematic uncertainties of the analysis. Additional details are provided in Chapter 7.

In Chapter 8 the normalised $t\bar{t}$ differential cross section results are derived, separately for kinematic distributions of the top quarks and the top-quark pairs and the top-quark decay products, namely leptons, b-jets, lepton pairs and b-jet pairs.

Chapter 2

Theory

The evolution and development of any field of science is an interplay of the theoretical hypotheses and experimental observations and measurements. Precise experimental results are used to improve the concepts described by the theory. And vice versa, the proposed theory models can only be validated or refuted thanks to new and better experimental measurements. The current understanding of the high-energy and particle physics is based on the so-called Standard Model (SM) of particle physics which comprises the description of the particles and their interactions.

In this Chapter a brief description of the SM theory is given followed by an overview of the top quark and its properties. A more extensive discussion can be found for example in [8].

2.1 Standard Model of Particle Physics

The SM of particle physics is a theory which describes the current understanding of the particles and their interactions as a result of the evolution of several concepts and experiments started in the early 20th century. It is based on a relativistic quantum field theory and gauge symmetry conservation described by the Ward–Takahashi identity [9, 10]. The interactions act on half-integer spin fundamental particles (fermions) exchanging integer-spin mediators (bosons). Four forces or interactions are known in nature: the electromagnetic, the weak, the strong interaction and the gravitational force. The former three are accounted for in the SM framework, while the latter cannot be described by the SM since it is not a complete quantum field theory.

Fermions

The SM contains 12 particles with spin 1/2, thus fermions, with masses[1] in the range of few MeV to hundreds of GeV organised in two main categories depending on the type of interaction they undergo. The fermions can be organised in three families which share common properties. Each particle has an antiparticle[2] with the same mass and quantum properties, and opposite electric charge.

A diagram with a summary of the properties and classification of the fermions is shown in Figure 2.1. The gauge bosons and their properties, discussed in the following section, are

[1] In this analysis "natural units" are used $\hbar = c = 1$. Transverse momenta, masses and energies are expressed in units of electron-volts (eV): MeV, GeV, etc.

[2] In the following, unless stated otherwise, particle and antiparticle terms are used indifferently. Thus "fermion" will refer to the fermion and antifermion, "top quark" will indicate the top quark and antiquark, etc.

also presented for completeness.

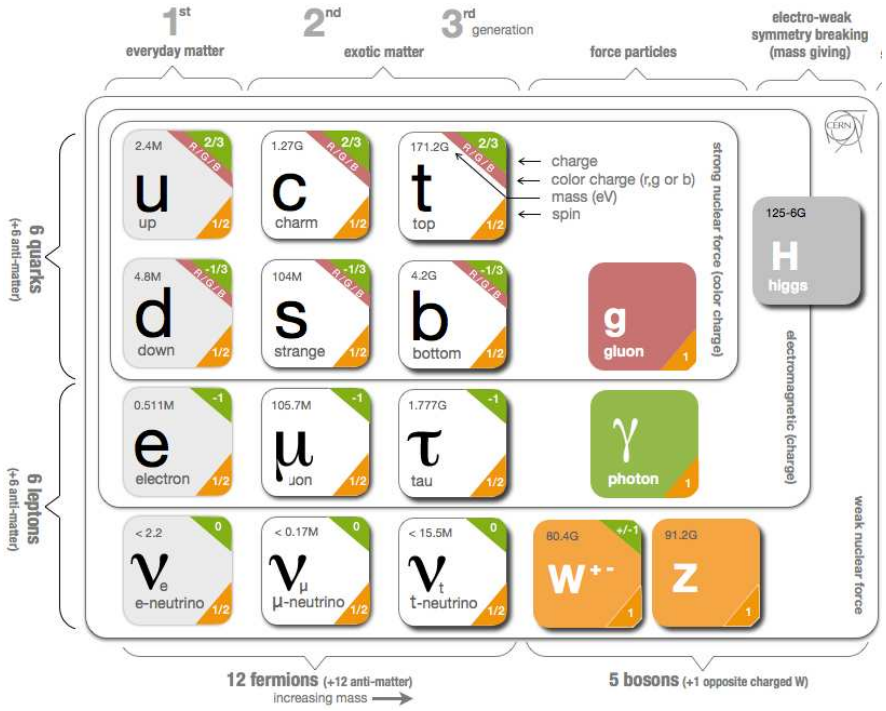


Figure 2.1: List of fundamental fermions and gauge boson mediators of the electroweak and strong interactions. The electroweak symmetry breaking boson is also indicated. The name, spin, mass and electric- and colour-charge are indicated for each of the particles. Figure (modified) taken from [11].

Interactions

The electromagnetic interaction between particles with non-zero electric charge is described by the quantum electro-dynamics (QED) theory. The QED presents a U(1) group symmetry properties with the electrically-neutral photon, γ, as the gauge mediator. The coupling constant associated to the electromagnetic interaction is defined as

$$\alpha_{\mathrm{EM}} \simeq \frac{e^2}{4\pi} \ll 1, \text{ at low energies.} \tag{2.1}$$

The QED coupling remains small, $\alpha_{\mathrm{EM}} \ll 1$, up to high energies and therefore the QED theory can be perturbatively expanded in series of powers of α_{EM}.

The weak interaction is described by a SU(2) symmetry group. The unification of the electromagnetic and weak interaction implies the presence of a neutral and two charged force mediators, the Z^0 and the $\mathrm{W}^{\pm}$ bosons[3], respectively. The W boson presents a non-zero electric charge and thus it couples at the same interaction vertex to a lepton and the corresponding neutrino or to a quark-antiquark pair. The coupling occurs within same family fermions in the case of the leptons while quarks from different generations can be coupled to the W boson. This effect is known as *flavour mixing* and is described by the Cabibbo–Kobayashi–Maskawa

[3]In the following and wherever is possible the charge superscript is omitted.

(CKM) matrix elements V_{CKM} [12], see Equation 2.2.

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.973 & 0.041 \\ 0.009 & 0.041 & 0.999 \end{pmatrix} \tag{2.2}$$

The strong interactions are described by the Quantum Chromodynamics (QCD) theory based on a SU(3) symmetry group. It presents three *colour* charge eigenstates, namely red, green and blue. The mediator of the QCD is the gluon, a colour-singlet state, which and thus interacts with other gluons or with the coloured quarks. The QCD-coupling parameter, α_s, is known to *run*, i.e. its value varies with the momentum transfer Q^2 of the interaction (see Figure 2.2). The running of the strong coupling presents two important consequences: the *asymptotic freedom* at short distances and the *colour confinement* at large distances. At high energies the strong coupling is small and therefore the quarks behave as free particles. The colour confinement instead occurs at small energies where $\alpha_s \simeq 1$. This causes the formation of colourless bound states of quarks, called hadrons. The lack of experimental evidence of colour singlet particles provides a confirmation of the validity of the QCD model.

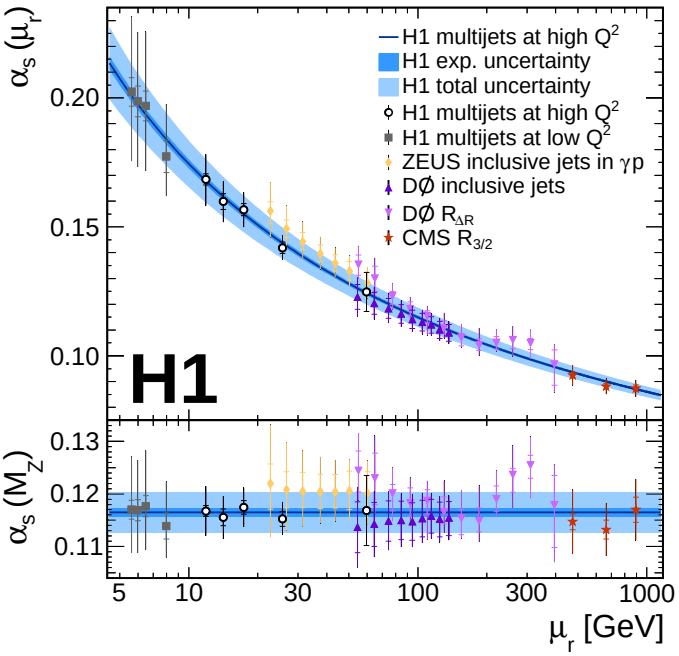


Figure 2.2: Distribution of the values of the strong coupling α_s determined in processes involving different momentum transfer scales μ_r measured in different experiments [13]. The bottom panel shows the same distribution normalised to the α_s value evaluated at the Z-boson mass scale $\alpha_s(\mu_r = m_Z)$.

Electroweak Symmetry Breaking

The simplest group notation of the SM is SU(3) × SU(2) × U(1) which predicts massless fermion and bosons. This consequence is in contradiction to the experimental results where

the W- and Z-boson masses are experimentally measured with very high precision [12]:

$$\begin{aligned} m(W) &= 80.385 \pm 0.015\,\text{GeV} \\ m(Z) &= 91.1876 \pm 0.0021\,\text{GeV} \end{aligned} \tag{2.3}$$

The presence of massive particles is introduced in the SM formalism by the *Higgs formalism*. It adds an additional scalar doublet field and the corresponding potential to the SM. The gauge invariance requirement produces a set of mathematically equivalent minimum vacuum eigenstates of the potential. The symmetry is *spontaneously broken* and predicts a vacuum expectation value to $v = 246\,\text{GeV}$. The physical manifestation of the spontaneous symmetry breaking model is the Higgs boson [14]. Experimentally, a particle resonance with spin-parity and decay branching fractions properties compatible with a Higgs boson was measured by the CMS and ATLAS Collaborations with a mass $m_H \simeq 125\,\text{GeV}$ [1,2].

The W and Z bosons gain mass by a gauge interaction with the Higgs field and the SM fermions through Yukawa interactions with a coupling constant Y_f.

$$m_f = \frac{Y_f \cdot v}{\sqrt{2}} \tag{2.4}$$

2.2 Top Quark Physics

The hadron collision event, such as the production of a $\mathrm{t\bar{t}}$ pair in a collider such as the LHC can be described by factorising the global event in sub-processes occurring at different energy scales separated by the *factorisation* scale μ_F.

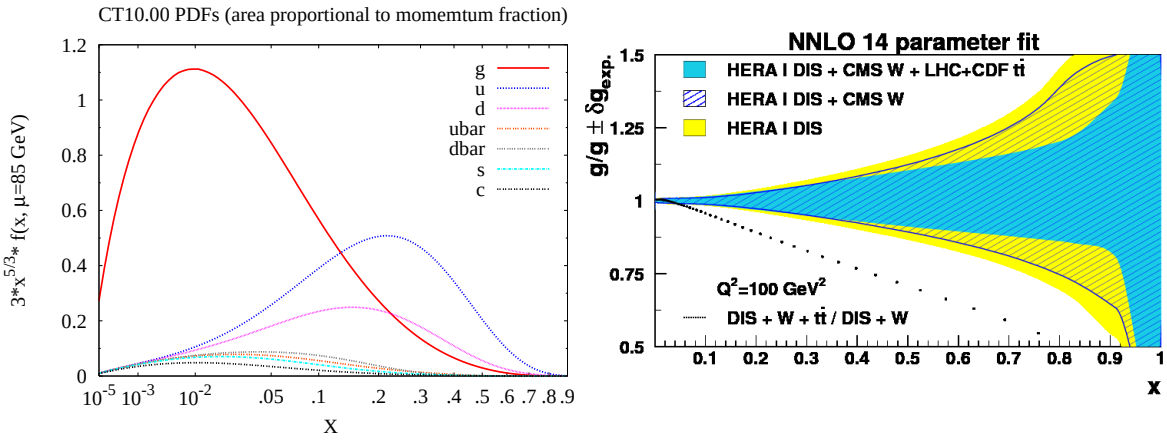
$$\sigma(\mathrm{pp} \to \mathrm{t\bar{t}}) = \sum_{i,j} \int_{x_i} \int_{x_j} \mathrm{d}x_i \mathrm{d}x_j f_{p|i}(\mu_r, x_i) f_{p|j}(\mu_r, x_j) \hat{\sigma}_{ij \to \mathrm{t\bar{t}}}(\mu_r, \mu_F). \tag{2.5}$$

The large distance confinement of coloured particles into hadrons is modelled by the parton density function (PDF) $f_{p|i}(\mu_r, x_i)$ which indicates the probability density to find a parton i inside the hadron p with a certain longitudinal hadron momentum fraction x_i at a renormalisation scale μ_r. The PDFs are universal distributions for a certain hadron type. They are obtained from fits to experimental results at a given energy scale which depends on the process. They can be derived at a different energy scale using the DGLAP evolution equations [15–17]. An example of the proton PDF obtained by the CTEQ collaboration is shown in Figure 2.3a for a $\mu_r = 85\,\text{GeV}$ scale. Top quark analyses can be used to constrain the PDF and to reduce the PDF uncertainties. As an example, Figure 2.3b shows the gluon PDF uncertainty obtained by the HERAFitter collaboration [18] using three different sets of experimental data [3], including top-quark measurements.

The partonic cross section is represented by $\hat{\sigma}_{ij \to \mathrm{t\bar{t}}}(\mu_r, \mu_F)$ and indicates the actual interaction between fundamental partons and presents a dependency on the factorisation and renormalisation scales μ_F and μ_r, respectively. Scattering processes between partons occur at energy scales where the coupling constant is small $\alpha \ll 1$ and thus a perturbative expansion of $\hat{\sigma}_{ij \to \mathrm{t\bar{t}}}$ in powers of α can be calculated.

Production

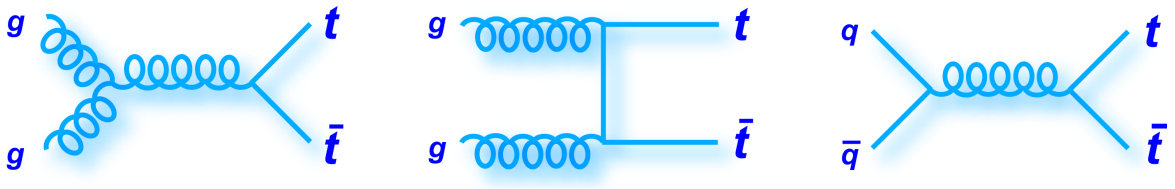
The top quark, as a coloured particle, can be produced via the strong interaction in pairs ($\mathrm{t\bar{t}}$ production) or via the electro-weak interaction individually (single-top production).



(a) (b)

Figure 2.3: (a): Result of the CT10 [19] PDF model of the proton as a function of the parton momentum fraction x at 85 GeV probe scale. (b): Uncertainties of the gluon PDF at a scale $Q^2 = 100\,\mathrm{GeV}^2$ as a function of the parton momentum fraction x as obtained in [3] using inclusive DIS measurements (light shaded band), DIS and W lepton charge asymmetry data (hatched band), and DIS, W lepton charge asymmetry and the $\mathrm{t\bar{t}}$ measurements (dark shaded band).

Top-quark-pairs are dominantly produced in gluon-gluon fusion ($gg \rightarrow \mathrm{t\bar{t}}$) or quark-antiquark annihilation ($q\bar{q} \rightarrow \mathrm{t\bar{t}}$) events at tree level in perturbation theory, as can be seen in Figure 2.4. The production via quark-gluon scattering processes ($qg/\bar{q}g \rightarrow \mathrm{t\bar{t}}$) only appear in next-to-leading (NLO) terms in the QCD perturbative expansion. The total $\mathrm{t\bar{t}}$ production



(a) (b) (c)

Figure 2.4: Feynman diagrams for $\mathrm{t\bar{t}}$ production at leading-order in QCD perturbative expansion: gluon-gluon fusion (a, b) and quark-antiquark annihilation (c).

cross section in pp collisions at $\sqrt{s} = 8\,\mathrm{TeV}$ has been calculated at next-to-next-to-leading order (NNLO) accuracy [20]: $\sigma_{\mathrm{t\bar{t}}}(\mathrm{NNLO}) = 245.8^{+8.8}_{-10.6}\,\mathrm{pb}$, for a top quark mass of 173.3 GeV. However, differential production cross sections are so far only calculated up to approximate NNLO accuracy level by resumming the soft-gluon emission near the partonic production threshold [3, 21].

In addition to the production of top-quark pairs via the strong interaction, single-top quarks can be produced via weak interactions. Three production channels can be distinguished at LO the t- and s-channels mediated by a virtual W boson and the associated production of a top quark and a W boson (tW channel). The Figure 2.5 shows the LO single-top production diagrams. The single-top production cross sections are calculated up to approximate NNLO

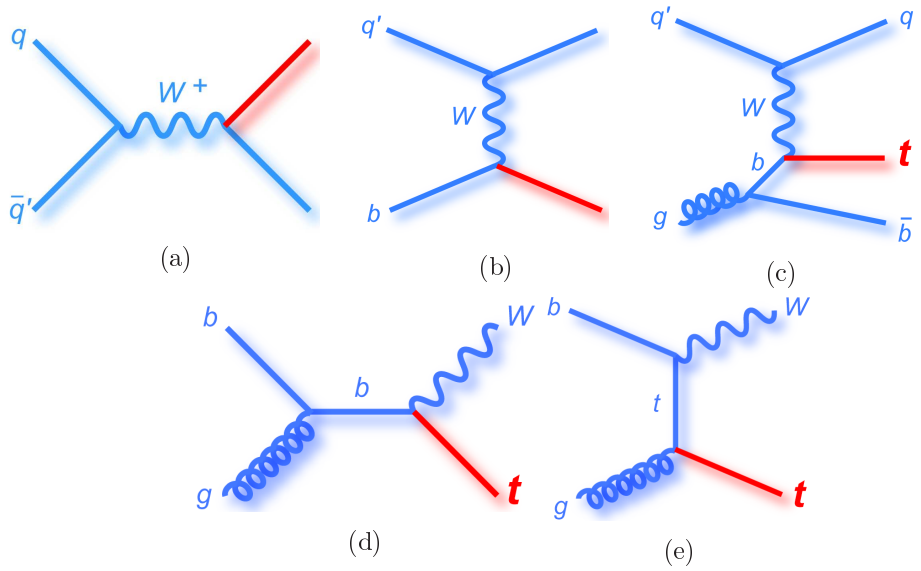


Figure 2.5: Leading-order single-top electroweak production diagrams: s-channel (a), t-channel (b and c), and tW-associated production in (d and e).

accuracy [21]. The results for the LHC $\sqrt{s} = 8\,\text{TeV}$ collision energy are shown in Table 2.1. The dominant production mechanism is the t-channel followed by the production in association with a W boson. The s-channel production at the LHC is suppressed due to low quark and antiquark fractions in the proton.

Channel	Cross Section [pb]
t-channel	$87.2^{+3.4}_{-2.4}$
s-channel	5.55 ± 0.22
tW-channel	22.2 ± 1.5

Table 2.1: Electroweak single-top production cross section expectations in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ [21]. Results are provided separately for the t- and s-channel and tW-associated production. The uncertainties account for renormalisation and factorisation scales and the PDF uncertainties on the calculation.

Decay

The top quark is the heaviest fundamental particle ever discovered with a mass of $m_t = 173.21 \pm 0.51 \pm 0.71\,\text{GeV}$ [12]. The large mass corresponds to a short lifetime, $\tau_t \propto m_t^3 \simeq 0.5 \times 10^{-24}\,\text{s}$ which is approximately an order of magnitude smaller than the QCD scale, $\Lambda_{\text{QCD}} \simeq 3 \times 10^{-24}\,\text{s}$. Therefore the top quark decays before the strong force can act on it. Top quark properties are transferred to its decay products which are not affected by the colour confinement effects. Such properties are for example the $t\bar{t}$ spin correlations. Different angular distributions of the lepton-pairs are expected depending on the correlation level between the top quark and antiquark (see Figure 2.6).

The top quark decays almost exclusively via the electroweak interaction into a W-boson and a b-quark as the CKM matrix element is expected to be $|V_{tb}| = 0.999$, as shown in

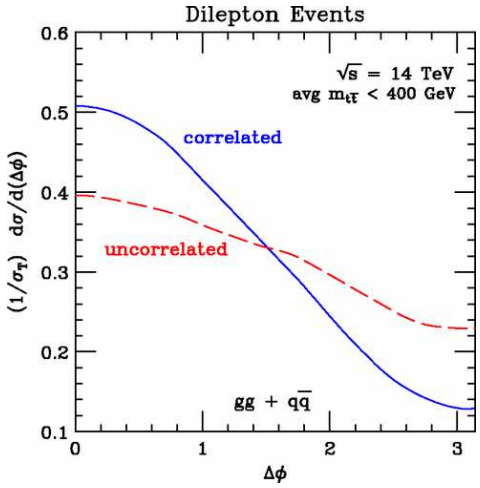


Figure 2.6: Normalised differential cross section distribution as a function of the azimuthal angle difference between the leptons produced from $t\bar{t}$ decays. Different results arise depending on the correlation level between the top quark and antiquark [22]. The results are shown for the particular configuration of pp collisions at $\sqrt{s} = 14\,\text{TeV}$ with $m^{t\bar{t}} \leq 400\,\text{GeV}$.

BR [%]		W^+ Decay			
		$u\bar{d}, c\bar{s}$	$e^+\nu_e$	$\mu^+\nu_\mu$	$\tau^+\nu_\tau$
W^- Decay	$\bar{u}d, \bar{c}s$	45.7	7.3	7.3	7.3
	$e^-\bar{\nu}_e$	7.3	1.2	1.2	1.2
	$\mu^-\bar{\nu}_\mu$	7.3	1.2	1.2	1.2
	$\tau^-\bar{\nu}_\tau$	7.3	1.2	1.2	1.2

Table 2.2: Branching ratios (BR), in %, of all possible $t\bar{t}$ decay channels.

Equation (2.2). The W-boson can further decay into a quark-antiquark pair or into a lepton-neutrino pair. The $t\bar{t}$ decay channel is defined by the decay of the W-bosons, see Table 2.2.

The fully-hadronic decay channel is characterised by the decay of both W-bosons into quark-antiquark pairs. This channel presents the largest $t\bar{t}$ decay branching ratio (BR), 45.7%, but suffers from a very high contamination of background events e.g. multijet production.

The semi-leptonic decay channel occurs when a W-boson decays into a lepton-neutrino pair and the remaining W-boson into a $q\bar{q}$ pair. This channel presents a 43.8% BR. The presence of a lepton, which usually has a high momentum, is used to discriminate $t\bar{t}$ events from background processes.

The dilepton channel accounts for approximately 10.8% of the $t\bar{t}$ decays and only for 4.8% when the $t\bar{t}$ decays via τ-leptons are excluded. Despite having the lowest BR of all $t\bar{t}$ channels, the presence of two high-energetic opposite-charged leptons is used to separate $t\bar{t}$ events from background processes.

The analysis presented in this thesis focuses in the dilepton $t\bar{t}$ decay channel where the W-bosons decay into electrons or muons, and the corresponding neutrinos.

$$\text{pp} \rightarrow \text{t}\bar{\text{t}} \rightarrow \bar{l}\,\nu_l\,b\,l\,\bar{\nu}_l\,\bar{b}, \text{ where l = e, } \mu. \tag{2.6}$$

The dilepton channel events with τ-leptons, which may decay further into electron or muon and the corresponding neutrinos, are considered background process. This choice is motivated by the presence of additional neutrinos from the τ decay which complicates the detection and reconstruction of the τ-kinematics.

Implications of the Top Quark in the Standard Model and Beyond

Top-quark production and decay involve the strong and electroweak interactions. A top-quark physics analysis provides an important test of various aspects of SM predictions.

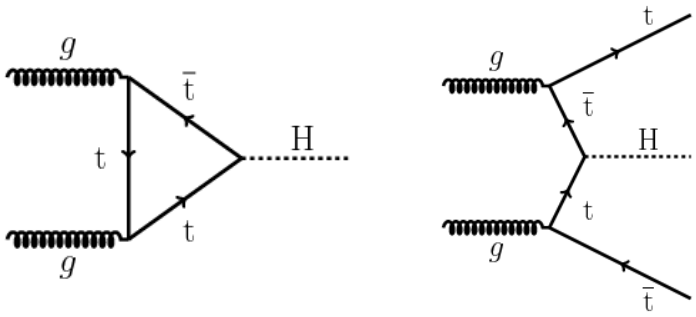
Higher-order effects in perturbative QCD can be tested by measuring the $t\bar{t}$ charge asymmetry, which only appears from the interference between the LO amplitudes of $q\bar{q}$-annihilation process and the one-loop corrections (NLO accuracy terms) in perturbative series [23].

The spin correlation between the top quark and antiquark can be studied in the distribution of the difference in azimuthal angles, $\Delta\phi$, between the lepton and antilepton produced in a $t\bar{t}$ dileptonic decay.

The precise measurement of the $t\bar{t}$ production cross section, both inclusive and differential, can be used to improve the accuracy of the QCD parameters such as α_s [4] and the PDF [3,18].

The unitarity of the CKM matrix can be tested measuring, or setting limits to, the V_{td}, V_{ts} and V_{tb} CKM matrix elements. These measurements require precise results of the top-quark-decay ratio [24] and single-top production cross sections [25].

The Higgs mechanism provides an explanation to the presence of massive gauge mediator bosons and fermions as explained in Section 2.1. Since scales of the top-quark mass and of the electroweak symmetry breaking are similar, the Yukawa coupling of the Higgs boson and top quark is dominant and close to unity ($Y_t \simeq 1$), see Equation (2.4). The main Higgs production mechanism at the LHC collision energy is the gluon-fusion process via a fermion loop. The loop involves predominantly top quarks due to the large Yukawa coupling (see Figure 2.7a). The Higgs boson can also be produced in association with a $t\bar{t}$ pair, denoted as $t\bar{t}$H-production, as represented in Figure 2.7b. A possible method to validate the Higgs mechanism is to measure the Yukawa coupling using the $t\bar{t}$H production cross section and comparing the results to the theoretical predictions.



(a) (b)

Figure 2.7: Feynman diagrams showing the gluon fusion production of a Higgs boson through a top-quark loop (a) and the production of a Higgs boson in association with a top-quark pair (b). These diagrams are representative of SM processes with sensitivity to the coupling between the top quark and the Higgs boson.

An accurate modelling of the SM $t\bar{t}$ production is required by Higgs-boson analyses aiming

to measure the Yukawa coupling via the $t\bar{t}H$ production cross section analyses. For a measured Higgs mass value of approximately 125 GeV the main decay channel is $H \rightarrow b\bar{b}$ (see Figure 2.8). Therefore $t\bar{t}H$ and the production of $t\bar{t}$-pairs accompanied by a pair of b-quarks will present identical final state topologies.

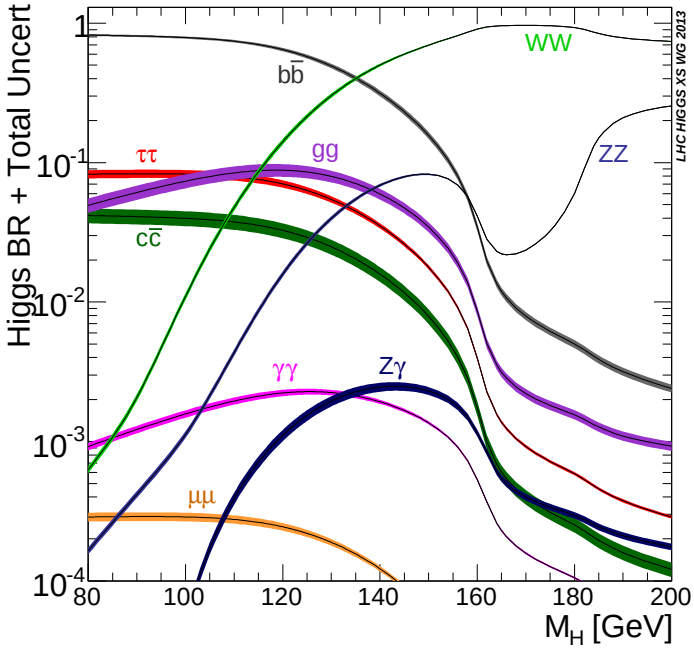


Figure 2.8: Standard Model Higgs-boson decay branching ratio as a function of the Higgs-boson mass [26]. The hatched area corresponds to the total uncertainty on the calculation.

Despite its big predictive success, the SM contains several limitations such as the hierarchy problem or the existence of dark matter. Different BSM models propose solutions to these limitations by introducing new particles, such as heavy boson resonances [27], or new symmetries as a consequence new particles, such as SUSY models [28]. Many of these SM extensions predict the decay of the new particles into a top quark or top-quark pairs. BSM searches rely on an excellent SM description of the $t\bar{t}$ production as signatures of the new processes might appear as excesses or deviations from the SM predictions.

Chapter 3

Experimental Setup

3.1 Large Hadron Collider

The Large Hadron Collider (LHC) [29–31] is a superconducting accelerator complex operated by the European Organization for Nuclear Research (CERN) located in the vicinity of Geneva (Switzerland). The LHC occupies the 26.7 km long tunnel constructed and used for the former Large Electron-Positron (LEP) collider. The LHC can accelerate beams of protons or lead ions in two opposite directions. Superconducting dipole and quadrupole magnets are used to guide and focus the beams and radio-frequency cavities are used to accelerate the particle bunches. This thesis is based on pp collision data and in the following only this operation mode is discussed.

Protons are accelerated in the CERN accelerator complex prior to injection in the LHC ring. A scheme of the location of the different accelerators and experiments in the CERN complex is shown in Figure 3.1. The pre-acceleration chain starts by creating the protons from hydrogen atoms and boosting them up to 50 MeV energy in the linear accelerator (LINAC2). The protons are then transferred via the BOOSTER into the PS (proton-synchrotron) where they achieve an energy of 25 GeV. From the PS, protons are sent to the Super Proton Synchrotron (SPS) from which they are injected in the LHC at an energy of 450 GeV, before being accelerated within the LHC to the desired collision energy. During the year 2012 a 4 TeV [33] per proton-beam operation mode was used.

Four interaction points exist along the LHC ring where the particles are brought to collision. In these points, four high-energy physics (HEP) experiments are located: A Large Ion Collider Experiment (ALICE) [34], A Toroidal LHC Apparatus (ATLAS) [35], the Compact Muon Solenoid (CMS) [36] and the Large Hadron Collider beauty (LHCb) [37]. The ALICE detector is designed to study heavy-ion collisions and LHCb is mainly devoted to study B physics. The ATLAS and CMS experiments are multi-purpose high-luminosity experiments with a broad HEP research programme: detailed study of the SM in a new energy regime, precision measurement of top-quark properties, Higgs-boson searches and property measurements, search for new physics beyond the SM. The Total Elastic and Diffractive Cross-Section Measurement (TOTEM) [38] and Large Hadron Collider forward (LHCf) [39] experiments share the interaction points with the CMS and ATLAS experiments. Their main focus is the measurement of the total inelastic pp cross section, forward physics analyses and the measurement of the proton size. This thesis is carried out in the context of the CMS experiment. In Section 3.2 further details are given.

The most relevant parameters of a collider dedicated to particle physics experiments are the centre-of-mass energy $\sqrt{s}$ and the instantaneous luminosity $\mathcal{L}$ since both are related to

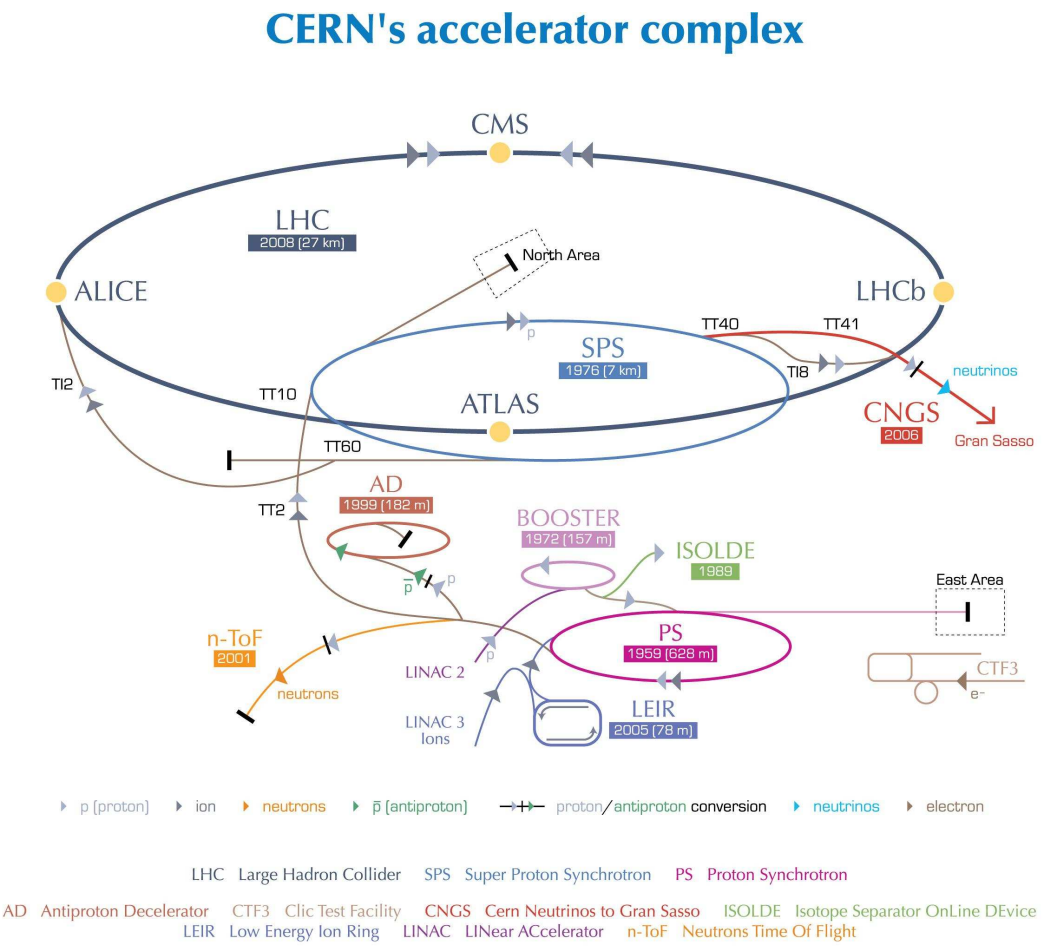


Figure 3.1: The CERN accelerator complex. Figure (modified) taken from [32].

the rate of events $\mathrm{d}N/\mathrm{d}t$ produced for a particular process.

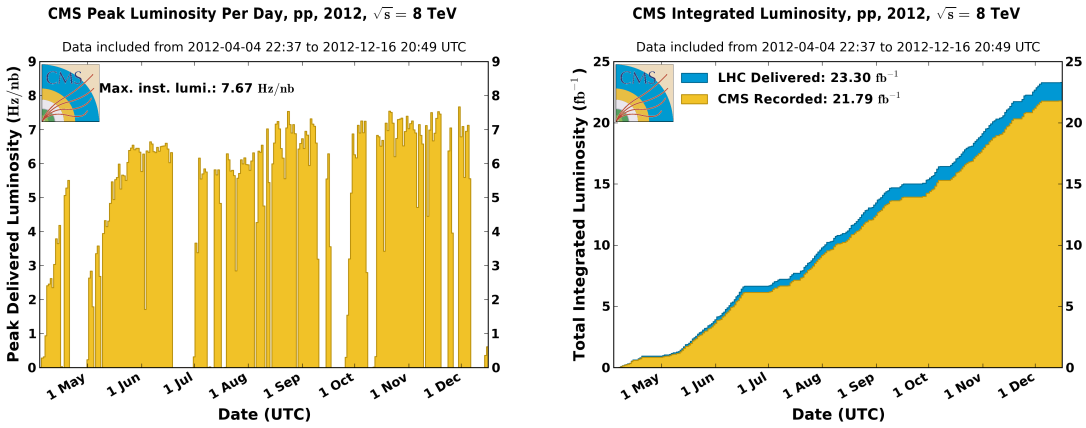
$$\frac{\mathrm{d}N}{\mathrm{d}t} = \mathcal{L} \cdot \sigma \tag{3.1}$$

where σ denotes the production cross section for a given process and depends on the collision centre-of-mass energy $\sqrt{s}$.

The instantaneous luminosity can be obtained from the running conditions of the accelerator by measuring the number of bunches in each beam, N_b, the number of protons in the bunch, N_p, the revolution frequency of the bunches, f, and the geometrical form of the beam in the interaction point, A:

$$\mathcal{L} = \frac{N_b \cdot N_p \cdot f}{A} \tag{3.2}$$

The LHC design parameters are $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ instantaneous luminosity, $1.15 \cdot 10^{11}$ protons per bunch and 25 ns bunch-spacing. During the year 2012, in the pp collision mode the running conditions evolved up to maximum values of $\mathcal{L} = 7.67 \cdot 10^{-33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and 1380 bunches containing $1.6 \cdot 10^{11}$ protons each separated by 50 ns [40]. Figure 3.2 shows the evolution of the peak instantaneous luminosity per day recorded by the CMS experiment and the cumulative total integrated luminosity delivered by the LHC and recorded by the CMS Collaboration during the year 2012.



(a) (b)

Figure 3.2: Peak instantaneous luminosity (a) recorded by the CMS experiment and cumulative integrated luminosity (b) delivered by the LHC (in blue) and recorded by the CMS experiment (in blue) as a function of days. The results correspond to pp collisions during the year 2012. Constant luminosity periods correspond to periods of no data taking. Figures taken from [41].

3.2 The Compact Muon Solenoid Experiment

This analysis is based on pp collision data collected by the Compact Muon Solenoid (CMS) detector installed at the LHC interaction point 5 (Point 5), near the village of Cessy (France). The CMS detector is a multi-purpose experiment with the usual onion-like design of collider-based particle-physics experiments. It is cylindrically symmetric around the beam axis and symmetric in both directions from the nominal interaction point around the axis defined by the beam line. The detector is 28.7 m long and 15.0 m wide with an approximate 14000 tonnes weight[1].

The main feature of the CMS detector is the very strong and homogeneous magnetic field of 3.8 T produced by a 12.5 m long and approximately 6 m wide superconducting solenoid coil made of NbTi. The silicon tracker and both electromagnetic and hadronic calorimeters are situated inside the solenoid in increasing radial distance from the axis defined by the LHC beam. The magnetic flux is returned through the iron yoke in which the muon spectrometer chambers are embedded. The high magnetic field bends the trajectories of charged particles produced in a transverse direction with respect to the beam axis making possible a precise momentum measurement from the curvature radius. A sketch of the CMS detector is presented in Figure 3.3. In the following the inner tracker, the electromagnetic calorimeter, the hadronic calorimeter and the muon system of the CMS experiment are described.

[1]The CMS collaboration adopted a right-handed reference system with the origin in the nominal interaction point, the x-axis pointed towards the centre of the LHC ring, the y-axis pointing vertically upward and the z-axis along the beam direction. The azimuthal angle ϕ is measured from the x-axis in the xy-plane, and the polar angle θ is defined from the z-axis. The pseudorapidity η defined as $\eta = -\ln(\tan(\theta/2))$ is preferred since it presents Lorentz invariance properties while the polar angle does not. Similarly the transverse momentum, p_{T}, of a particle is defined in plane transverse to the beam-axis: $p_{\mathrm{T}} = |\vec{p}| \cdot \cos\theta$. The angular separation is defined in terms of the $\Delta\mathrm{R} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ variable.

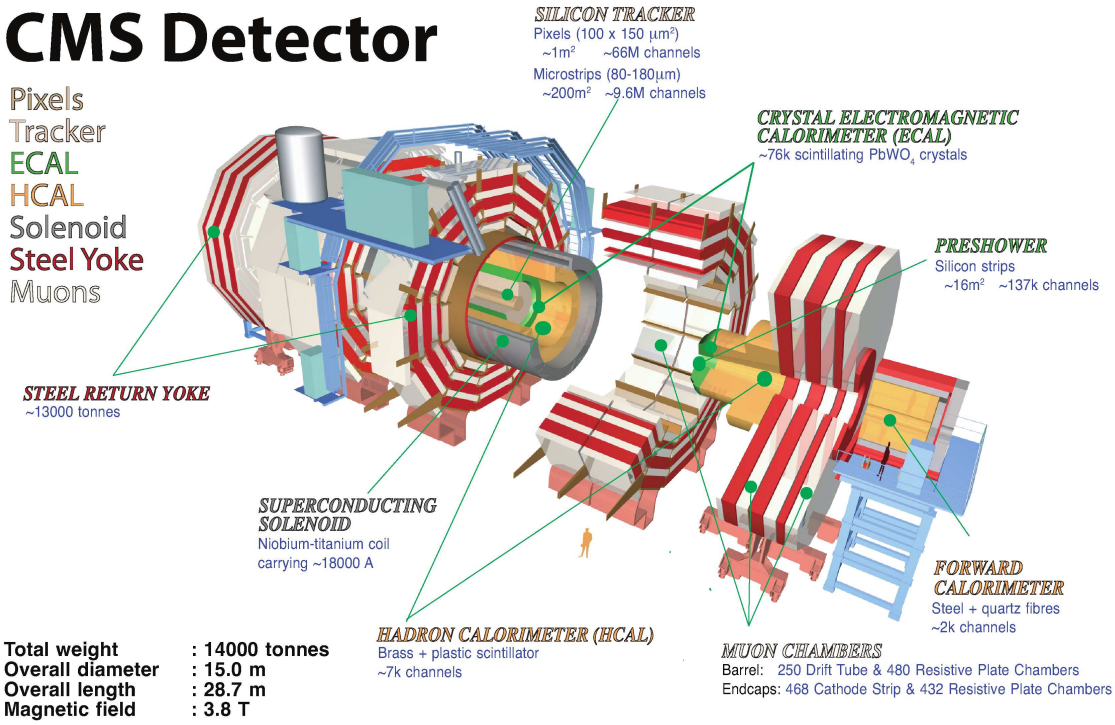


Figure 3.3: The CMS detector.

Inner Tracking Detector

The inner tracking system [42], also referred to as tracker, is the CMS subdetector located closest to the LHC beam axis and to the pp interaction point. It is designed to measure the trajectories of charged particles with high accuracy and to reconstruct the position of the collision and the decay vertices precisely. The inner tracking system is composed of a pixel detector and a strip tracker made of semiconductor silicon covering a total length of 5.8 m and 2.5 m in diameter. A schematic display of the position of the tracker subsystems is shown in Figure 3.4.

The **pixel detector** consists of three co-axial barrels located at distances between 4.4 cm and 10.2 cm in the central detector region and two forward discs at ± 34.5 cm and ± 46.5 cm in each z direction covering a pseudorapidity range $-2.5 < \eta < 2.5$. It contains 66 million pixels each $100 \times 150\,\mu\text{m}^2$ in size and it covers a total area of about $1\,\text{m}^2$.

The **strip detector** occupies the radial region between 20 cm and 116 cm with 15148 modules containing a total of 9.3 million silicon strips. It is organised in four subsystems: the tracker inner barrel (TIB), the outer barrel (TOB), the inner disks (TID) and the endcaps (TEC). The modules in the TIB and TOB cover the central pseudorapidity region of the tracker with their strips parallel to the beam axis. While the orientation of the strips in the TID and TEC is perpendicular to the beam. The average strip pitch ranges between 80 μm and 500 μm.

The tracker achieves a reconstruction efficiency between 85% in high η regions increasing up to essentially 100% in the barrel. The resolution of the transverse momentum is mea-

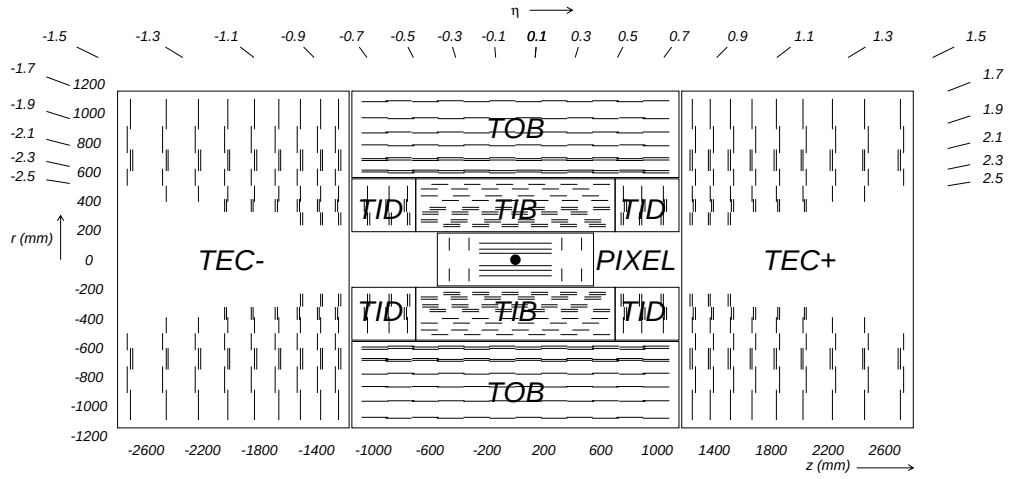


Figure 3.4: Schematic cross section through the CMS tracker in the $r - z$ plane.

sured [43] to be 1.8% for low momentum ($1 < p_{\mathrm{T}} < 10\,\mathrm{GeV}$) charged particles and increases up to 2.8% for muons with $p_{\mathrm{T}} = 100\,\mathrm{GeV}$.

Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) [44] is an hermetic and homogeneous detector composed of approximately 67000 lead tungstate ($PbWO_4$) crystals. Scintillation and short radiation-length properties of the lead tungstate are exploited to detect and measure the energy and position of electrons and photons by absorbing the full electromagnetic shower energy in the ECAL detector. The ECAL is divided into the barrel (EB) and the two endcaps (EE) covering the pseudorapidity regions $|\eta| \leq 1.479$ and $1.479 \leq |\eta| \leq 3.0$, respectively. Additionally a preshower detector is placed in front of the EE to identify neutral pions decaying into photon pairs. The layout of the ECAL is shown in Figure 3.5.

The ECAL barrel is composed of 61200 crystals organised in submodules with 5 crystal pairs each. Each crystal is 230 mm long and is constructed in a tapered shape starting with $22 \times 22\,\mathrm{mm}^2$ cross section in the front area, closest to the beam axis, increasing up to $26 \times 26\,\mathrm{mm}^2$ in the outermost rear area. The length of the crystals corresponds to 25.8 electromagnetic radiation lengths (X_0). The submodules are arranged in a quasi-projective geometry with a small tilt of their axis with respect to the direction to the nominal interaction point such that the number of particle trajectories going along cracks between the crystals is minimised.

Each ECAL endcap contains 156 groups of crystals arranged in units of 5×5 crystals. Individual crystals are 220 mm long with an cross section increasing from approximately $28 \times 28\,\mathrm{mm}^2$ to $30 \times 30\,\mathrm{mm}^2$. The spatial dimensions of the EE correspond to $24.7\,\mathrm{X}_0$.

The preshower is composed of two layers of lead each followed by a silicon strip detector located in front of the EE. The preshower improves the position determination of electrons and photons and enhances the identification of neutral pion decays.

Prior to the installation into the CMS detector, the energy resolution of the ECAL was studied using electrons with energies in the 20 GeV to 250 GeV range in a test beam [45]. The ECAL energy resolution, $\sigma(E)$, is parametrised as a function of a stochastic term (S), a

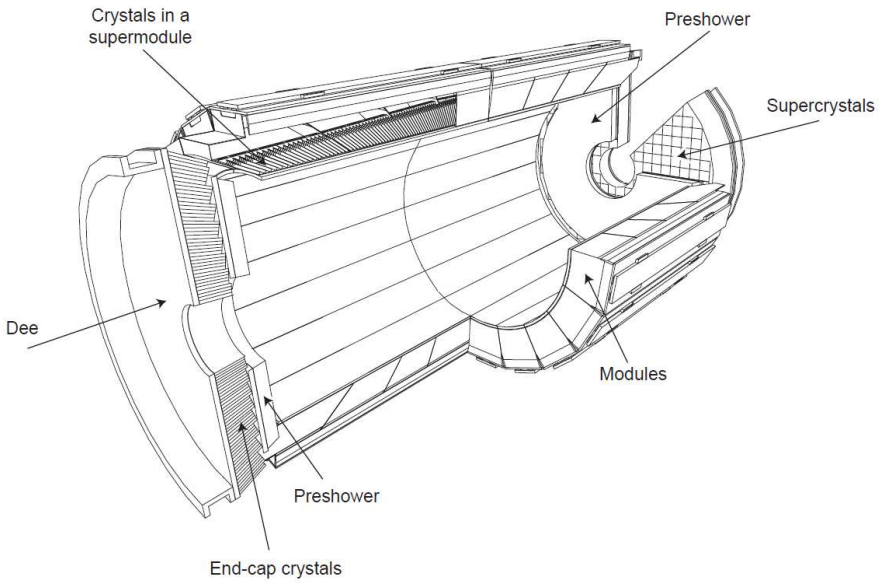


Figure 3.5: Layout of the CMS electromagnetic calorimeter. Figure (modified) taken from [36].

detector noise factor (N) and a term to account for intercalibration and energy leakage effects (I):

$$\begin{aligned} \left(\frac{\sigma(E)}{E}\right)^2 &= \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + I^2 \\ &= \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.3\%)^2 \end{aligned} \tag{3.3}$$

where the energy E is measured in GeV.

After the start of the LHC, the energy resolution of the ECAL was determined using pp collisions from LHC operation in the years 2010 and 2011 at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ [46]. The electron energy resolution was found to be better than 2% in the barrel region ($|\eta| \leq 0.8$) and between 2% and 5% elsewhere for electrons with transverse energies of approximately 45 GeV. The energy resolution of photons with transverse energies of approximately 60 GeV varied between 1.1% and 2.6% in the EB and from 2.2% to 5% in the EE.

Hadronic Calorimeter

The calorimetric system of the CMS detector is completed with a hadronic sampling calorimeter (HCAL) [36, 47]. The HCAL design is mainly driven by the decision to place it within the solenoid, and aims to precisely measure the energy and direction of hadron showers. High hermeticity is required to the HCAL in order to be sensitive to the detection of missing transverse energy due to weakly interacting particles such as neutrinos or exotic particles. The HCAL is composed of four subdetectors: the barrel HCAL (HB), the HCAL endcaps (HE), the outer barrel HCAL (HO) and the forward calorimeter (HF). An schematic view of the HCAL and its subcomponents is shown in Figure 3.6.

The HB is a sampling calorimeter constructed out of flat brass absorber plates aligned parallel to the beam axis interleaved with scintillation plastics. The thickness of the HB

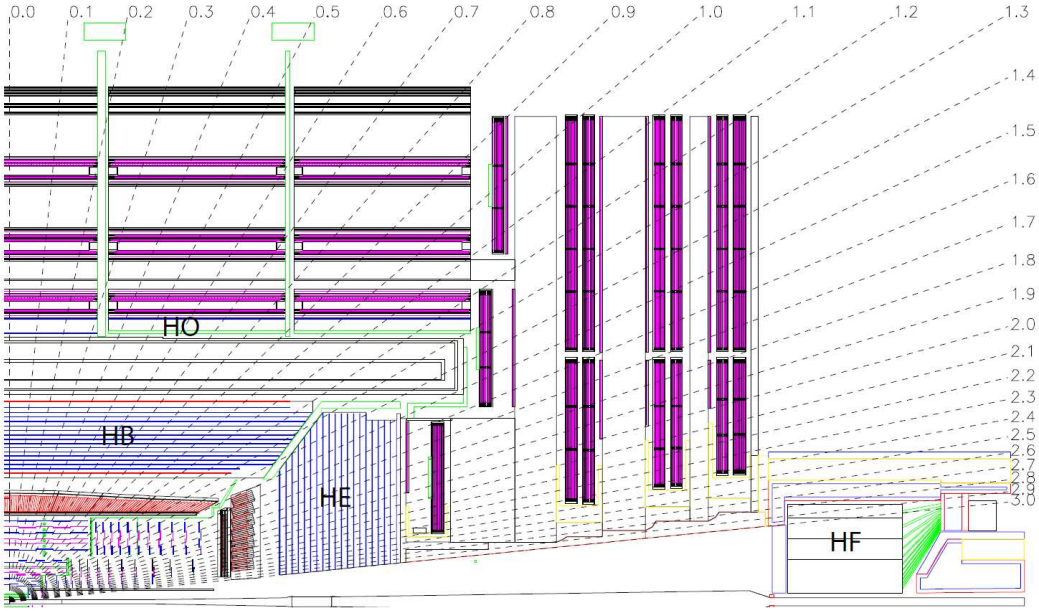


Figure 3.6: Longitudinal view of the CMS detector showing the locations of the hadron calorimeters subsystems: barrel (HB), endcap (HE), outer (HO) and forward (HF) hadronic calorimeters.

corresponds to a nuclear interaction length (λ_I) varying from 5.8 to approximately 10 when moving from $|\eta| = 0$ to 1.3 in pseudorapidity.

The HO ensures the full detection of the hadronic shower energies in the central detector regions, where the material budget is low. A pair of scintillator layers in between the solenoid and the muon detector increases the "thickness" of the central HCAL detector to more than $10\,\lambda_I$.

The HE complements the hermetic coverage of the HB up to $|\eta| < 3.0$. The granularity of the HE varies from $(\Delta\eta, \Delta\phi) = (0.087, 0.087)$ in the geometrical region $|\eta| < 1.6$ increasing to $(\Delta\eta, \Delta\phi) = (0.17, 0.17)$ in the remaining $1.6 < |\eta| < 3.0$ regions. The material budget of the HE reaches a similar interaction length as the combination of the HB and HO.

The forward subsystem complements the hadron-shower detection in the $3 < |\eta| < 5.2$ region. Iron and quartz fibres are used as active materials which resist the unprecedented particle fluxes expected at such high-pseudorapidity regions.

The energy resolution, σ, of the HCAL was measured using test beam data obtaining a result of

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{115^2}{E}\right) + 5.5^2 \tag{3.4}$$

for energies E measured in GeV [48]. This result is consistent with the HCAL design values.

Muon Detector

Muon detection is a powerful tool for the recognition of signatures of interesting physics processes amongst the very high background rate expected at the LHC, e.g. new physics or BSM processes. The large amount of material between the interaction point and the muon

chambers absorbs practically all other particles. Thus, only muons and neutrinos can reach the muon stations which are the outermost group of subdetectors of the CMS experiment.

The CMS muon system [36,49] has three functions: muon identification, muon momentum reconstruction and triggering. The CMS experiment uses three types of muon subsystems (see Figure 3.7): drift tubes (DT) in the $|\eta| < 1.2$ barrel region, cathode strip chambers (CSC) in the endcaps $0.9 < |\eta| < 2.4$ and resistive plate chambers (RPC).

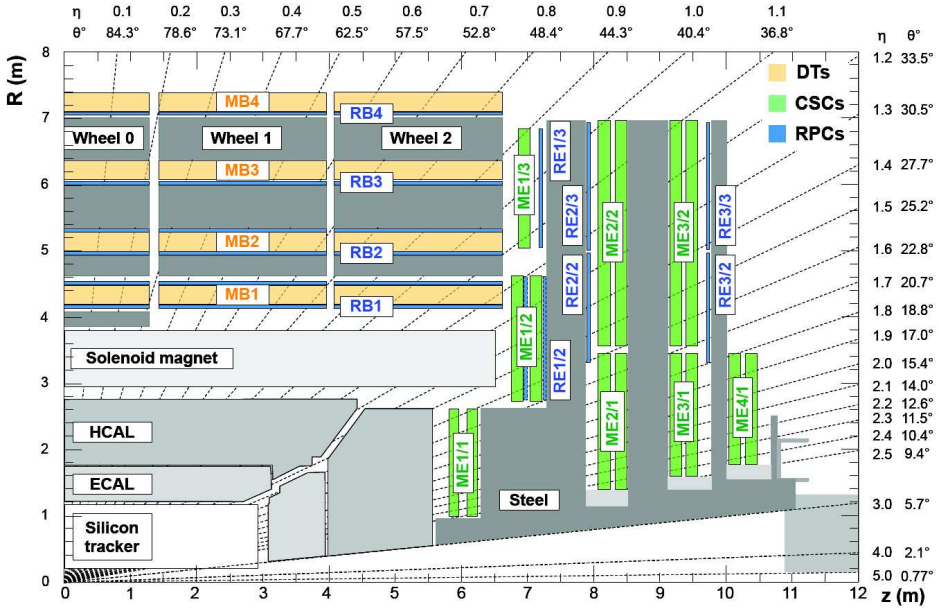


Figure 3.7: Longitudinal layout of one quadrant of the CMS detector. The four DT stations in the barrel (green), the four CSC stations in the endcap (blue), and the RPC stations (red) are shown.

The **drift tubes** are organised in a four stations interspersed among the layers of the flux return plates and provide a complete (r, ϕ, z) hit position. A spatial resolution of 100 μm in the $r - \phi$ plane and 150 μm in the $r - \theta$ plane is achieved in each DT station.

A total of 468 **cathode strip chamber** modules are installed in the CMS endcap region. This subsystem achieves a spatial resolution of the muon track position of ~ 5 mm in the radial direction and better than 150 μm longitudinal direction.

Due to the very fast response and excellent time resolution the **resistive plate chambers** are generally used for triggering. The barrel RPCs are located in the inner face of the DTs with two additional layers in the back face of the two innermost DT chambers and four additional RPC layers complete the coverage in the endcap region.

Proton-proton collision events at $\sqrt{s} = 7$ TeV were used to measure the average muon p_T resolution [50]. The transverse momentum resolution of muons with $p_T < 200$ GeV is below 9%, while at high momenta up to 1 TeV the p_T resolution of muons is 15–40% depending on the pseudorapidity region.

Trigger System

The 2012 LHC running conditions used proton bunches separated by 50 ns producing of the order of $3 \cdot 10^8$ collisions per second. Technical limitations allow a maximum data storage at a

rate of a few hundred crossings per second, therefore a drastic reduction of the rate is needed. This task is performed by the CMS trigger system in a two-step process.

The hardware-based Level-1 trigger (L1T) [51] system uses only calorimeter and muon system signals to produce a global trigger decision based on information related to p_{T} and η variables of the particle-candidate or particle-identification criteria.

Based on the L1T decision, a software-based high-level trigger (HLT) [52] produces the final rate of events which are stored for physics analysis. The HLT algorithm uses a *partial event reconstruction* approach such that only those particle candidates and/or regions of the detector that are actually needed for reconstruction are evaluated.

This analysis uses dilepton based HLT paths [53] which requires the presence of pairs of electrons, pairs of muons or an electron and a muon with p_{T} above a given threshold.

- Double Muon HLT: requires two muons with at least 17 and 8 GeV transverse momentum reconstructed using combined information from the tracker and muon detector systems. A second HLT path requires the second muon to be reconstructed in the inner tracker only.
- Double Electron HLT: the presence of two electrons with at least 17 and 8 GeV transverse momentum is demanded using a track from the inner detector matched to a energy cluster from the calorimeter. Additionally minimal identification and isolation criteria are required to the triggering electrons.
- Electron-Muon HLT: events containing electron-muon pairs with $p_{\mathrm{T}} > 17\,\mathrm{GeV}$ and 8 GeV are kept for further processing. The muon is reconstructed as a combination of inner detector and muon system tracks, while the electron makes use of the tracker and the calorimeter information.

The specific HLT paths used in this analysis are listed in Section A.2.

3.3 Data Quality Monitoring

The high quality of the data used in physics analyses is guaranteed by the Data Quality Monitoring (DQM) process [36,54] which evaluates the detector conditions during the collision data taking, the software used for reconstructing and processing the recorded collisions [55], the generation of simulation samples, etc.

The CMS experiment has set up a two-step process with the capability of monitoring the status of the detector subsystems in real-time, *online* DQM, and the certification of the reconstruction and calibration of the recorded data, *offline* DQM. The DQM applications produce automatically control distributions in histogram format which are then collected and stored for inspection in a web-based graphical user interface (GUI) [56]. The DQM GUI is the central tool used for data certification since it comprises in a compact format the status histograms of all subdetectors and physics objects. A representative example is shown in Figure 3.8. The collision data offline DQM process (in the following generically referred to as DQM) is briefly described in the following.

The DQM files are produced in parallel to the reconstruction of the data used for analysis. A harvesting process collects, merges and stores the results into the offline DQM GUI. The DQM file harvesting is a semi-automated process running periodically every 12 hours and requires human intervention only in case of a request with non-standard harvesting conditions.

The actual certification occurs in shift operation with a 24-hour daily coverage shared between CERN (Switzerland), DESY (Germany) and the Fermilab National Laboratory (USA)

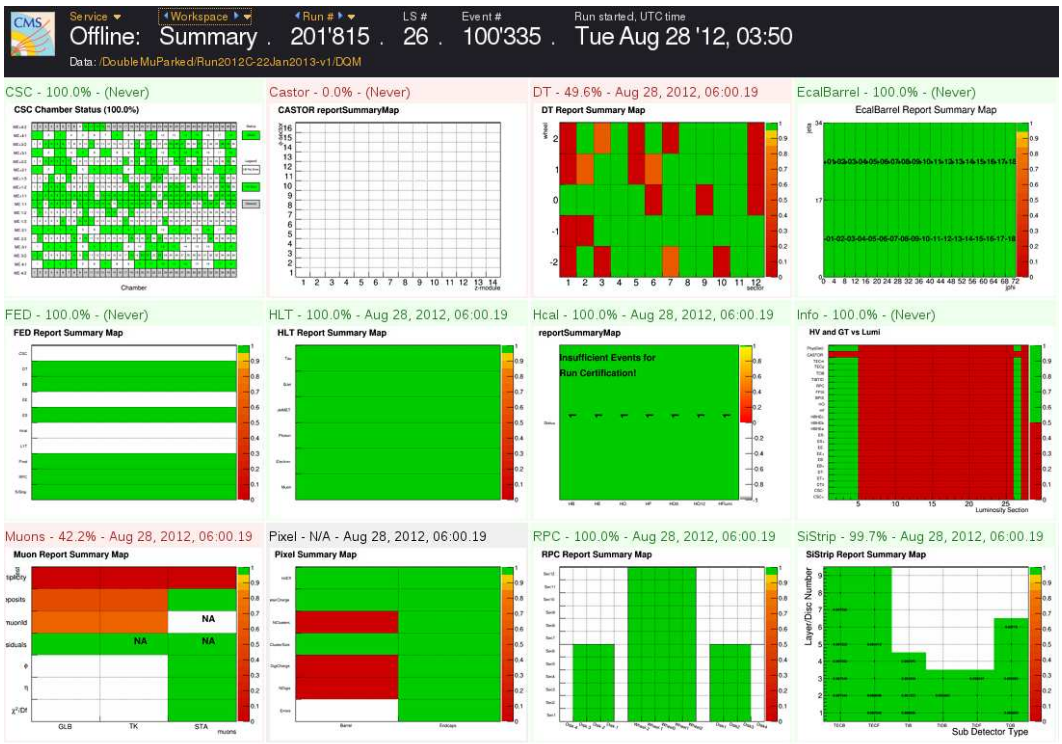


Figure 3.8: Offline DQM GUI example showing different histograms representing the status of a selection of CMS subdetectors.

research centres. The continuous monitoring allows a fast and early detection of incidences in the data taking or in the reconstruction process. The different subdetector and physics object experts define evaluation procedures and reference distributions based on selected data collisions. The offline DQM validation is done by a visual comparison of the collected distributions and the reference distribution. A goodness status flag is assigned to each system under evaluation following expert instructions. A summary of each certified data collection period is stored in a central database for a later confirmation and sign-off by the detector and physics object groups, and a final verdict is delivered to the experiment as a common list of runs and luminosity sections[2] of good data-taking time-period for all CMS subdetector and physics objects.

The CMS experiment collected approximately $L = 21.8\,\text{fb}^{-1}$ of pp collisions during the 2012 LHC operation time, but only about 90% of this dataset was certified as valid for physics analysis. The remaining 10% corresponds to pp collision events collected and recorded during unstable LHC beam conditions, detector or subdetector malfunctioning periods.

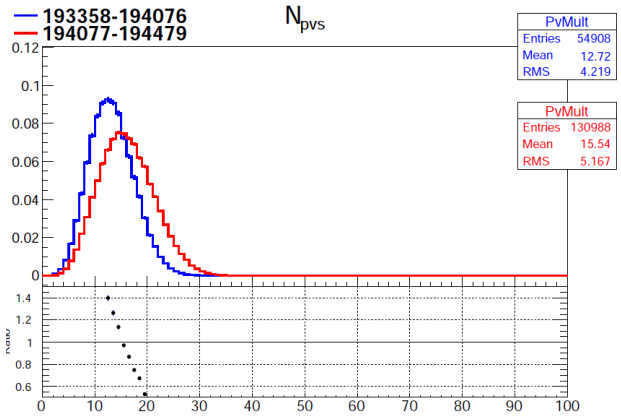
Data Quality Monitoring for Top Quark Physics

The large amount of collected events allowed for the monitoring of top-quark-like events, *top DQM*. During the 2012 data taking period, the top DQM tools and procedures were updated and used to monitor variables sensitive to $\mathrm{t\bar{t}}$ events and periodic reports were provided to

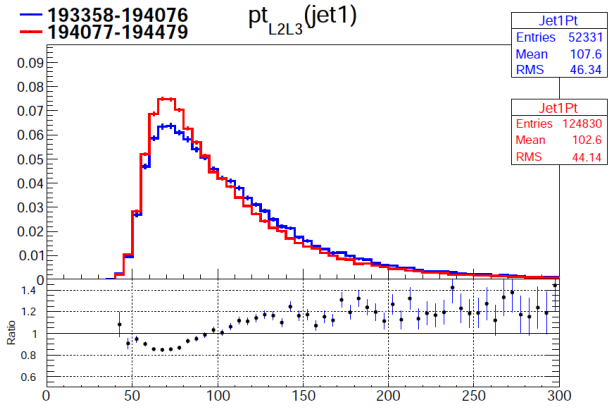
[2] A *luminosity section* is defined as the time period corresponding to 2^{18} LHC orbits of data taking, equivalently $t = 23.31$ seconds [57]. An uninterrupted set of luminosity sections with approximately constant detector conditions is defined as *run*.

the TOP Physics Analysis Group [58]. The top DQM procedure kept track of the evolution of the LHC running conditions by comparing pp collision events collected in two consecutive weeks. Top-quark-pair-like event signatures were monitored based on the standard ℓ + jets and dilepton $t\bar{t}$ decay channel selection: collision events were filtered with constraints in the number of leptons and jets and in the p_T and η of the particles. In addition to "standard" distributions such as lepton p_T or jet multiplicity, more complex variables sensitive to the top quark properties were implemented, e.g. the trijet invariant mass (M3) variable which represents a rough mass value of the hadronically decaying top-quark. A representative example is shown in Figure 3.9, where the distributions of the number of primary vertices and the transverse momentum of the leading jet are shown for events selected with a topology of the e + jets decay channel containing at least a high-p_T electron and at least four hard jets.

No unexpected observations were made and any differences in the monitored distributions were understood to be produced due to the varying LHC running conditions.



(a)



(b)

Figure 3.9: Distributions of the number of primary vertices N_{pvs} (a) and leading jet transverse momentum, in GeV, $pt_{L2L3}(jet1)$ (b) observed in the top DQM process. The results, normalised to unit area, are obtained in events with at least one electron with $p_T > 20\,GeV$ and $|\eta| < 2.4$, and four jets with $p_T > 25\,GeV$ and $|\eta| < 2.5$ in pp collisions collected in the consecutive 2014/5/4-13 (blue) and 2014/5/14-19 (red) week-periods. The differences in the distributions are caused by the increasing instantaneous luminosity by a factor ~ 2 with the evolution of the data taking.

Chapter 4

Simulation of Events

The understanding of a pp collision requires complicated theoretical calculations often analytically unresolvable. Additionally, the analyses require a detailed understanding of the complex detector and its interactions with the particles produced in the collisions. This limitations are usually solved using Monte Carlo (MC) simulation programs which produce expectation yields and kinematic distributions which are then compared to the experimental results. These programs are also used to calibrate the detector, to estimate experimental efficiencies and corrections, to quantify the fraction of background events, etc. Furthermore, event simulation tools play an important role in the viability studies and design of future experiments.

Separated phases of the pp collision occur at different energy scales. The QCD factorisation theorem [59] separates the full event description into more specialised subprocesses. The colliding proton components are described by the parton distribution functions (PDF). The PDFs are derived from fits of a variety of experimental measurements (see Section 2.2). The hard scattering generators describe the collision between the partons producing a desired final state event, e.g. $gg \to \mathrm{t\bar{t}} \to l\bar{\nu}b\ \bar{l}\nu\bar{b}$. Details on the hard scattering calculators are given in Section 4.1. Many of the generated final-state particles contain electric- or colour-charge, thus they evolve emitting electromagnetic and/or QCD radiation. This simulation step is described by the parton showering tools (Section 4.2). The radiation process finishes with a colour confinement and hadron formation process (Section 4.3). The possible interaction between proton remnants are described by the underlying event models as explained in Section 4.4. Figure 4.1 presents a graphical summary of a pp collision event simulation process.

In this chapter a brief description of the individual sub-processes of the Monte Carlo simulation chain are discussed. Special emphasis is given to the tools used in this analysis to simulate $\mathrm{t\bar{t}}$ events. The complete list of generated event-samples is presented at the end of the chapter.

4.1 Hard Scattering

The production of $\mathrm{t\bar{t}}$-pairs in a hadron collider occurs via the scattering or annihilation of gluons or quarks as presented in Section 2.2. The parton interaction occurs in an energy regime where the QCD-coupling constant is small, $\alpha_s(Q^2) \ll 1$, and thus the partonic cross section $\hat{\sigma}$ can be derived from perturbative theory calculations.

The $\mathrm{t\bar{t}}$ production model implemented in MadGraph [61] calculates the hard scattering process at leading-order (LO) QCD accuracy with up to nine final state particles. In the case of dileptonic $\mathrm{t\bar{t}}$ events six particles arise from the $\mathrm{t\bar{t}}$ decay (two leptons, two neutrinos and two

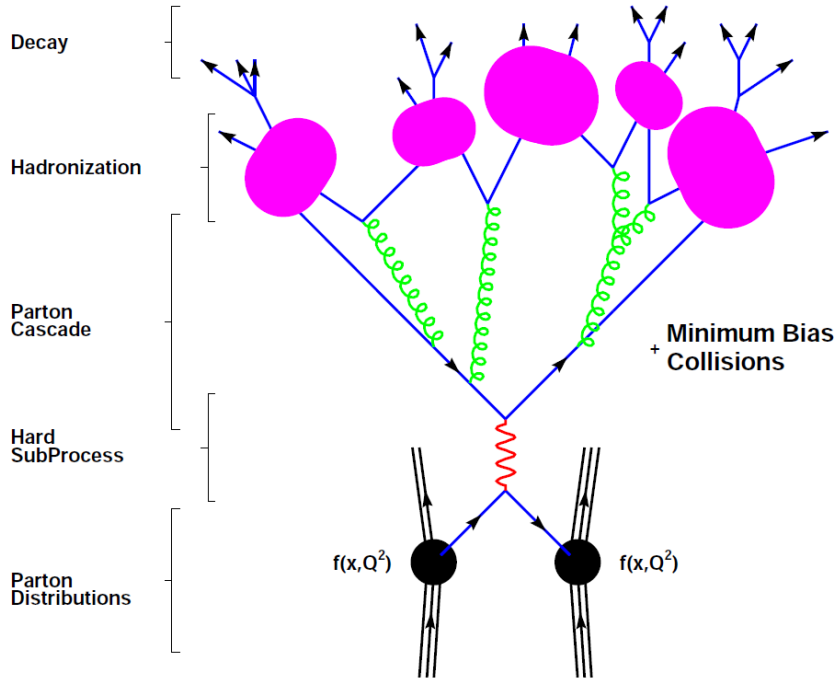


Figure 4.1: Visualisation of a pp collision indicating the different evolution steps. Figure (modified) taken from [60].

b-quarks) and the remaining three are additional particles. As discussed in Section 2.2 the spin-correlation effects of the $t\bar{t}$ pair are transferred to its decay products before being diluted by the QCD interactions. The MadSpin [62] software package, attached to the MadGraph simulation, accounts for such effects in the event simulation process.

Alternative generator models, such as Powheg [63] or MC@NLO [64], simulate the production of $t\bar{t}$ events at next-to-leading order (NLO) accuracy, $\mathcal{O}(\alpha_s^3)$, by including virtual and real-emission corrections to the LO terms of the calculation (see Figure 4.2).

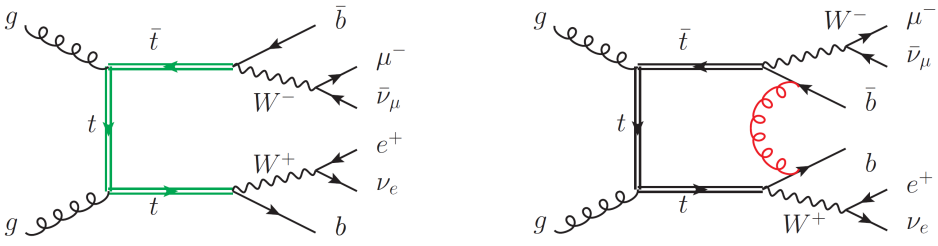


Figure 4.2: Examples of Feynman diagrams for $t\bar{t}$ production in the dilepton final state at LO (a) and NLO accuracy (b) including the presence of a virtual gluon (in red). Figures (modified) obtained from [65].

Other simulation programs instead include a complete event simulation from the hard scattering of partons to the final hadron formation process at the expense of lower accuracy in the hard scattering model. This is the case for Pythia [66] which is used to simulate background processes such as multijet production at LO QCD accuracy.

4.2 Parton Showering

The $t\bar{t}$ initial and final states contain coloured particles which can radiate gluons. The gluons can further split into a quark-antiquark or gluon pair. This process continues iteratively as long as the system contains sufficient energy, e.g. the $g \to b\bar{b}$ process requires a minimum gluon-energy of $2 \cdot m^{b}$ in order to be energetically allowed to occur. This radiation emission processes is known as *parton showering* and occurs in an energy regime where $\alpha_s \ll 1$ and perturbative QCD calculations are still valid.

The parton-shower models iteratively produce parton splittings reducing the energy scale of the event. The process starts at the maximum energy scale Q_0 determined by the hard scattering and continues until a cut-off Q_{min} minimum value is reached. At the Q_{min}, typically $\mathcal{O}(\text{ GeV})$, the emitted partons are not experimentally distinguishable because they are either soft (low energetic) or emitted in a collinear direction with respect to the parent parton. The specific Q_0 and Q_{min} values are user-defined parameters in the simulation program which regulate the amount of radiation produced by the showering process.

The *evolution variable* parametrises the radiation emission process. Thus, different parametrisations of the evolution variable generate different shower shapes.

- The Pythia [66] model orders the parton emissions in decreasing momentum by selecting a $Q^2 = p_{\mathrm{T}}^2$ dependency, where p_{T} represents the transverse momentum of the radiating parton.
- The Herwig [67] tool instead implements an angular-ordered shower with a $Q^2 = 2E_a^2(1-\cos\theta)$ relation, where E_a is the parent parton energy and θ is the angle between the parent and the emitted parton.

The final state particle showers, described above, are also suitable for initial state parton emission by inverting the logic of the process. The parton participating in the hard scattering is assumed to be produced with a virtuality Q^2 (fixed by the hard scattering) as the endpoint of a showering process from partons with lower scales.

The matching of a hard scattering generator to a parton shower tool can produce a double counting effect since the final state configuration of a hard parton (p) emitted by the parton shower in addition to the LO $t\bar{t}$ hard scattering is the same as the $t\bar{t}+1p$ produced by the hard scattering. Different techniques overcome this issue with an appropriate scale separation.

The MC@NLO [64] method introduces two parton shower terms which start the emission at different scales. One term starts the evolution at energy scales of the event if no additional parton is generated from the matrix element, while the second term starts the showering at a maximum energy of the 1^{st} additional parton produced by the matrix element.

The Powheg [68] approach reweights the parton shower model according to the matrix element term such that it acts only on the phase space configuration defined by the leading order expansion terms.

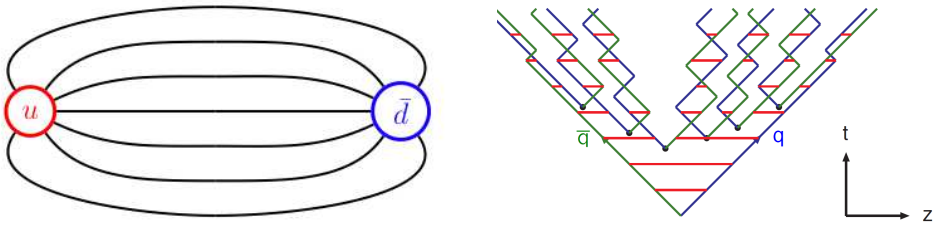
The MLM matching [69] uses a veto on events which contain emissions from the parton shower already covered by the matrix element. This strategy is used to remove double-counting terms when the MadGraph matrix element generator is interfaced with the Pythia parton shower.

4.3 Hadronisation Models

The evolution of the parton showers finishes at an energy regime of the order of few GeV where $\alpha_s \simeq 1$ and thus the strong interaction becomes indeed strong and forces the formation of

colourless bound states, the hadrons. The process is modelled by the hadronisation algorithms.

The string or Lund hadronisation model [70], used in PYTHIA [66], is based on long-distance colour-confinement approach. It assumes a uniform colour potential which depends on the spatial separation between a quark-antiquark ($q\bar{q}$) pair. The $q\bar{q}$ pair evolves in time by increasing their mutual distance and therefore the energy of the colour potential. When the potential energy is of the order of the hadron masses, it is energetically favourable for the string to break at some point along its length creating a new quark-antiquark pair, $q'\bar{q}'$. The new antiquark is created at the end of the string segment connected to the original quark q and the new quark to the original antiquark $\bar{q}$. The string evolution stops when the string potential energy cannot produce more hadrons, i.e. when the separation between the quark and antiquark produces a potential energy smaller than the typical hadron masses. Figure 4.3 shows a representation of the colour potential and string fragmentation evolution.



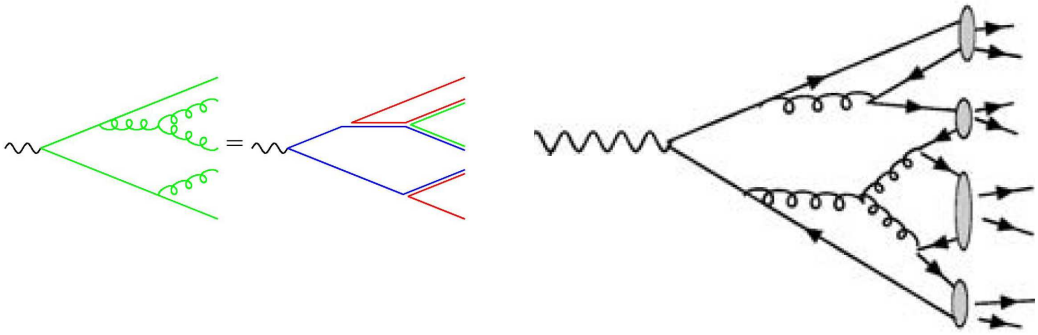
(a)

Figure 4.3: (a): Graphical representation of the colour potential between a u-quark and a $\bar{d}$-quark. (a): Sketch of the time-space evolution of the string hadronisation model. Figures (modified) obtained from [71].

An alternative hadronisation technique is based on the colour pre-confinement assumption [72], i.e. colour-connected partons are produced in a collinear direction. A graphical representation is shown in Figure 4.4a. The HERWIG [67] program implements this hadronisation model. It produces colourless clusters of partons with a cluster-invariant-mass distribution independent of the process [71]. The clusters are considered as proto-hadrons which are forced to decay into hadrons via a quasi-two-body decay. In this model, gluons are identified as colour-anticolour-pair lines and are forced to produce a $q\bar{q}$ pair before being assigned to a cluster. A conceptual representation of the cluster hadronisation model is shown in Figure 4.4b.

4.4 Underlying Event

In addition to the hard scattering of the partons, secondary scattering or diffraction processes between the proton remnants can occur in a pp collision. These secondary effects are collectively classified as underlying events (UE). The UE usually involves low energy processes and therefore the models are tuned to experimental data. The Z2* model was derived using data collected by the CMS experiment from proton-proton collisions at $\sqrt{s} = 0.9$ and 7 TeV [73]. This UE model is conventionally used in Monte Carlo samples generated with PYTHIA. Simulation samples using the HERWIG showering and hadronisation model use two different UE parametrisations: the HERWIG model itself and the AUET2 model [74]. The latter is derived by the ATLAS Collaboration using LHC collision data.



(a) (b)

Figure 4.4: (a): Schematic representation of the colour preconfinement assumption. The gluons are represented by a colour-anticolour line pair connected at the vertices and quarks by single-colour lines. (b): Representation of the cluster hadronisation model where the grey circles indicate the clusters. Gluon are forced to split into $q\bar{q}$ pairs before being assigned to a cluster which later produces final state hadrons by a quasi-two-body decay. Figures (modified) obtained from [71].

4.5 List of Simulated Samples

Several Monte Carlo simulation samples are used in this analysis to model both $t\bar{t}$ events and background processes. They all correspond to the *Summer12* Monte Carlo production campaign produced and validated centrally by the CMS Collaboration. All generated particles are processed through a full CMS detector simulation implemented in the GEANT4 software [75] such that real collision data and simulation events are reconstructed with the same algorithms. Additional details on the event reconstruction are given in Chapter 5.

Various SM processes can produce final states that can be identified as produced from $t\bar{t}$ events due to particle mis-reconstruction or mis-identification. MADGRAPH hard scattering model interfaced to PYTHIA, for showering and hadronisation effects, is used to simulated the production of W-bosons accompanied by additional partons, the Drell-Yan process $Z/\gamma* \to l\bar{l}$, and the production of top-quark pairs accompanied by Z, W or γ bosons. Different regimes of the Drell-Yan production are described with two distinct samples. The samples are produced filtering the generated events according to the lepton-pair invariant mass value: $m^{l\bar{l}} \in [10\,\text{GeV}, 50\,\text{GeV}]$ and $m^{l\bar{l}} \geq 50\,\text{GeV}$. The single-top event production is modelled by the POWHEG hard scattering tool interfaced to PYTHIA. The PYTHIA generator is used in the production of diboson (WW, WZ and ZZ) and multijet events (referred to hereafter as QCD). The simulation of QCD events is separated in different samples preselecting events with partons and muons above a given p_T threshold or requiring the presence of an electron stemming from the decay of a B- or a C-hadron. All background processes are normalised to the total integrated data luminosity using the most precise theory predictions for the cross sections. A complete list of background process samples, including their cross sections and an approximate numbers of simulated events, is given in Appendix A.1.

The main sample of simulated $t\bar{t}$ events is produced combining the MADGRAPH and PYTHIA generators to model the hard scattering and the showering and hadronisation effects, respectively. This sample is produced with the CTEQ6L1 PDF set and a top-quark-mass value of $m^t = 172.5\,\text{GeV}$. Additional samples are used to evaluate the systematic uncertainties on the $t\bar{t}$ modelling (c.f. Chapter 7) by varying a simulation parameter. The effect of the top-

quark mass is evaluated using MadGraph+Pythia samples in which the top quark mass is varied $m^{\mathrm{t}} = 172.5 \pm 1\,\mathrm{GeV}$. The matrix element-parton showering matching scale and the Q-scale are varied by conventional factors 0.5 and 2 with respect to the default 20 GeV and $Q_0^2 = m_t^2 + \sum_{\text{add. partons}} p_{\mathrm{T}}^2$ settings in the MadGraph+Pythia simulation. These variations are used to evaluate the matching and Q^2-scale uncertainties.

Alternative $\mathrm{t\bar{t}}$ event samples are produced interfacing the Powheg matrix element generator with Pythia and Herwig showering and hadronisation models, and another sample using MC@NLO interfaced with Herwig. These samples are used to estimate the uncertainties related to the matrix element and parton showering and hadronisation models. They are also used as theory comparison of the results. The different $\mathrm{t\bar{t}}$ simulations are normalised to the in situ measured cross section value (Section 6.5.2). A summary of the different $\mathrm{t\bar{t}}$ simulation samples is shown in Table 4.1.

Description	Hard Scattering	PS & Hadronisation	PDF Set	UE Model	Additional Choices
Default	MadGraph (v5.1.5.11)	Pythia (v.6.424)	CTEQ6L1	Z2*	$m_t = 172.5\,\text{GeV}$ $Q_0^2 = m_t^2 + \sum_{\text{add. partons}} p_{\text{T}}^2$ ME-PS scale 20 GeV
Q^2-Scale Up	MadGraph (v5.1.5.11)	Pythia (v.6.424)	CTEQ6L1	Z2*	$Q^2 = 4 \cdot Q_0^2$
Q^2-Scale Down	MadGraph (v5.1.5.11)	Pythia (v.6.424)	CTEQ6L1	Z2*	$Q^2 = 0.25 \cdot Q_0^2$
ME-PS Up	MadGraph (v5.1.5.11)	Pythia (v.6.424)	CTEQ6L1	Z2*	ME-PS scale 40 GeV
ME-PS Down	MadGraph (v5.1.5.11)	Pythia (v.6.424)	CTEQ6L1	Z2*	ME-PS scale 10 GeV
Mass Up	MadGraph (v5.1.5.11)	Pythia (v.6.424)	CTEQ6L1	Z2*	$m_t = 173.5\,\text{GeV}$
Mass Down	MadGraph (v5.1.5.11)	Pythia (v.6.424)	CTEQ6L1	Z2*	$m_t = 171.5\,\text{GeV}$
PDF variation	MadGraph (v5.1.5.11)	Pythia (v.6.424)	CT10	Z2*	PDF variation
Alternative Model	Powheg (v.1.0)	Pythia (v.6.424)	CT10	Z2*	–
Alternative Model	Powheg (v.1.0)	Herwig (v.6.520)	CT10	Herwig	–
Alternative Model	MC@NLO (v.3.41)	Herwig (v.6.520)	CTEQ6M	AUET2	–

Table 4.1: Summary of configurations of the different $\text{t}\bar{\text{t}}$ simulation samples used in this analysis. The version of the hard scattering, parton showering (PS) and hadronisation tools is displayed in parentheses. The different PDF sets and UE models are listed separately. Additional simulation parameter choices is indicated in the last column. The default parameters of the simulation tools are used, unless explicitly stated.

Chapter 5

Event Reconstruction and Selection

The top-quark-pair-production differential cross section is measured in the dilepton decay channel, $t\bar{t} \to \bar{b}l\bar{\nu}\, b l\nu$. The event final state topology is composed of two neutrinos detected as missing transverse energy ($\vec{\not{E}}_T$), two leptons, electron or muon, and two jets produced in the hadronisation of the b-quarks. The full-leptonic $t\bar{t}$ decay channel has the lowest branching ratio BR $\simeq$ 4.8% of all $t\bar{t}$ decay channels. The presence of an oppositely-sign lepton-pair facilitates the selection of a clean event sample with a low contamination from background processes.

The measurement is based on the full 2012 dataset collected with the CMS detector in pp collisions at $\sqrt{s} = 8\,\text{TeV}$. The dataset corresponds to a total integrated luminosity of $\mathcal{L} = 19.7 \pm 0.5\,\text{pb}$ of data certified for physics analysis by the DQM process (c.f. Section 3.3). Several SM processes produce event final states similar to the dileptonic decay of top-quark pairs. These are the production of W bosons accompanied by additional partons, Drell-Yan process (in the following referred to as DY), tW-channel single-top production, diboson production (WW, WZ and ZZ), multijet events (hereafter referred to as QCD) and the production of top-quark-pairs in channels other than the dilepton or $t\bar{t}$-pairs accompanied by Z, W or γ bosons.

The reconstruction and selection of events follows the recommendations of the CMS Top-Quark-Physics-Analysis Group [76] based on the particle-flow (PF) [77, 78] algorithm. The PF method combines information from each CMS subdetector in order to reconstruct the trajectory, the momentum and the energy of each particle-candidate in the event: muons, electrons, photons, and neutral and charged hadrons. The PF-particle-candidates are later used to build up more complex objects such as jets (clusters of particles) or missing transverse energy.

The analysis is optimised to select a large fraction of $t\bar{t}$ dilepton events rejecting contamination from background processes based on kinematic requirements applied to the leptons, the jets and the missing transverse energy present in the event. The individual top-quark 4-momentum components cannot be measured directly in the detector, they are only accessible via a kinematic reconstruction algorithm.

In the following the event and particle reconstruction and selection criteria are described following the logic of the analysis. The chapter concludes with the event yield results and a set of control distributions.

5.1 Trigger

In Section 3.2 the trigger system of the CMS experiment was introduced and the need of such a system was motivated by the need to collect only those pp collisions that are interesting for physics analysis. The events are selected based on HLT paths that require the presence of at least two leptons, electron or muon, with transverse momenta larger than 17 and 8 GeV. The muon candidates are reconstructed using the CMS muon and inner tracker subsystem information without isolation requirements. Electrons candidates, reconstructed combining silicon tracker and calorimeter information, are required to fulfil very loose identification and isolation conditions. The detailed HLT paths are listed in the Appendix A.2.

The dilepton trigger efficiencies ϵ in data and simulation are estimated using an ensemble of events selected with a set of orthogonal HLT paths based on $\not\!\!E_{\mathrm{T}}$ requirements. The fraction of these events that additionally fire the dilepton triggers define the trigger efficiency. Small differences in data and simulation efficiencies are corrected with a data-to-simulation factor (SF). The SF is derived double-differentially as a function of pseudorapidity of the two triggering leptons as the ratio of the data efficiency over the simulation efficiency. The SFs are found to be stable with respect to the vertex multiplicity, and in general they are above 90% in the full phase space for all HLT paths. The results are shown in Figure 5.1.

5.2 Collision Vertex

The presence of several hadron collisions in each bunch crossing is generally called as *pileup* (PU) and requires a careful treatment because particles produced in secondary collisions might add spurious energy to the primary interaction. The collision vertices are reconstructed assigning tracks to the vertex candidates based on the track-longitudinal position at the closest distance with respect to the beam axis. An adaptive vertex fitter [79] assigns to each track a weight which represents the likelihood that the track arises from the associated vertex.

The selected events must contain at least one good vertex by requiring the sum of the associated-track weights to be larger than 4. In addition the vertex has to be in the central detector region, i.e. $|z| < 24\,\mathrm{cm}$ with respect to the nominal interaction point and less than 2 cm in the transverse plane around the beam axis. The vertex with the highest scalar sum of the transverse momentum squared (p_{T}^2) of the tracks associated to the vertex is assigned to the primary pp interaction and the rest to PU vertices. The charge hadron subtraction (CHS) algorithm removes from the event the tracks of charged particles produced in secondary vertices. Contributions from neutral particles are corrected at the jet calibration step (see Section 5.5).

The average vertex multiplicity distribution is measured in data based on the total pp inelastic cross section [80] and the instantaneous luminosity of the events. Simulation events are generated with a distribution of the average number of vertices based on a priori assumptions of the LHC running conditions. The distribution of the event number of vertices in data and simulation differs, as shown in Figure 5.2a. Therefore events from simulation are corrected event-by-event to match the measured distribution in data. In Figure 5.2b the vertex multiplicity distribution is shown after applying the correction to the simulation events. A good agreement between the data and the simulation can be observed.

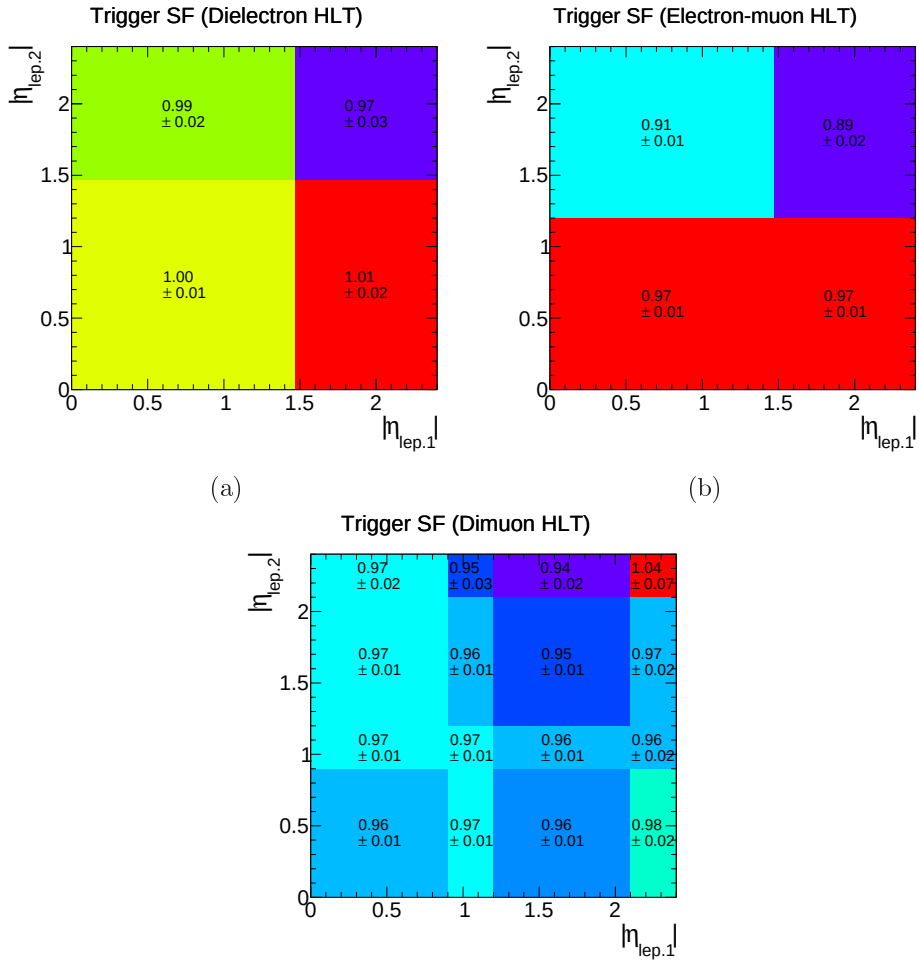


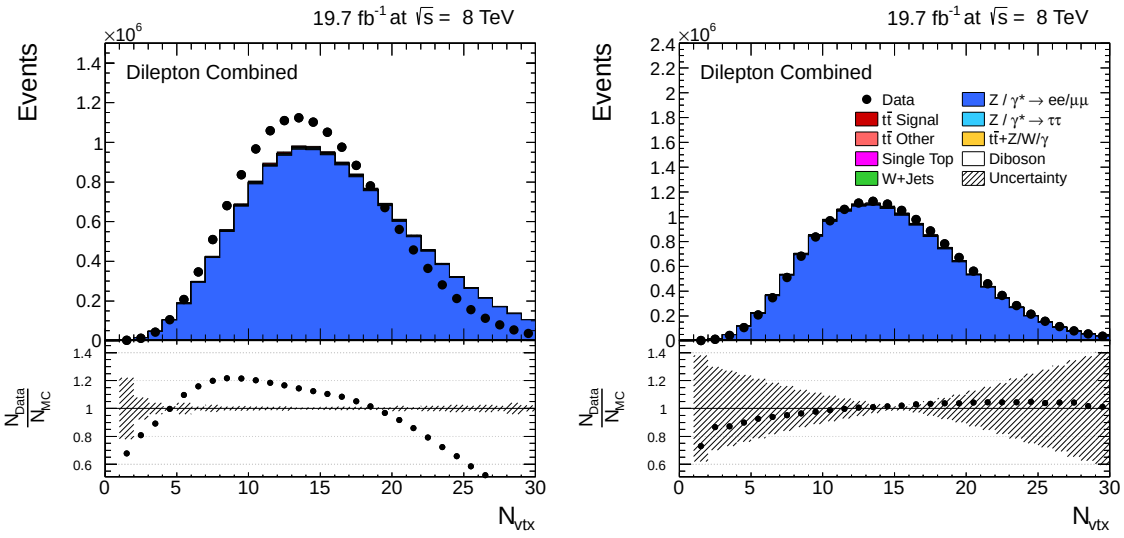
Figure 5.1: Distributions of the data-to-simulation efficiency ratio (SF) for the ee (a), eμ (b) and $\mu\mu$ (c) HLT trigger paths. The results are presented as a function of pseudorapidity of the triggering leptons ($|\eta_{\text{lep.},1}|, |\eta_{\text{lep.},2}|$) and include the total uncertainty on the determination method (c.f. Section 7.2).

5.3 Lepton Reconstruction and Selection

As described in Section 2.2 the dileptonic $t\bar{t}$ decay is characterised by the presence of two oppositely-charged high-p_{T} leptons. In this analysis, only prompt electrons or muons are considered, and generically referred to as leptons unless stated otherwise.

The **muon** reconstruction combines information from the inner tracker and muon subdetector to produce a trajectory along the full detector [81]. Three muon categories can be distinguished:

- *Standalone muons* are based on tracks reconstructed using the muon subsystem only.
- *Tracker muons* are reconstructed as inner detector tracks matched to at least one hit in the muon subdetector.
- *Global muons* are reconstructed as a combined fit of a track from the muon system and



(a) (b)

Figure 5.2: Distribution of the average number of reconstructed vertices in data and simulation after trigger and dilepton selection criteria (see Section 5.3 for further details). The distributions are shown before (a) and after (b) correcting the vertex distribution in simulation. The hatched band corresponds to all shape uncertainties on the signal $t\bar{t}$ events (c.f. Chapter 7). The data-over-simulation ratio is displayed in the bottom panel.

a track from the inner tracker.

Only global or tracker muons are selected for further analysis.

Tracks from the inner detector and energy clusters deposited in the ECAL are combined to reconstruct **electron** candidates [82]. Initially the energy deposits of the ECAL crystals are added into a so-called *supercluster* taking into account the EM shower properties. The complete electron track is obtained after fitting the supercluster position to a track from the silicon detector [83]. The fitting algorithm takes into account possible bremsstrahlung-radiation emissions produced by the transition of the electron through the detector material.

The presence of hits in the inner tracker detector associated to the electron-candidate track, and a maximum track distance of 0.04 cm in the transverse plane with respect to the primary vertex are minimum identification criteria applied to the candidate electron. Additionally, a multivariate (MVA) discriminant increases the separation power between prompt electrons and mis-reconstructed candidates. The discriminant is trained on shower-shape, track-quality and particle-kinematic variables. A minimum discriminant value of 0.5 is required to the selected electron candidates.

The events with at least two oppositely-charged, well reconstructed and identified leptons are considered in the analysis if the leptons are detected in the $|\eta| \leq 2.4$ central detector region with a minimum transverse momentum of 20 GeV.

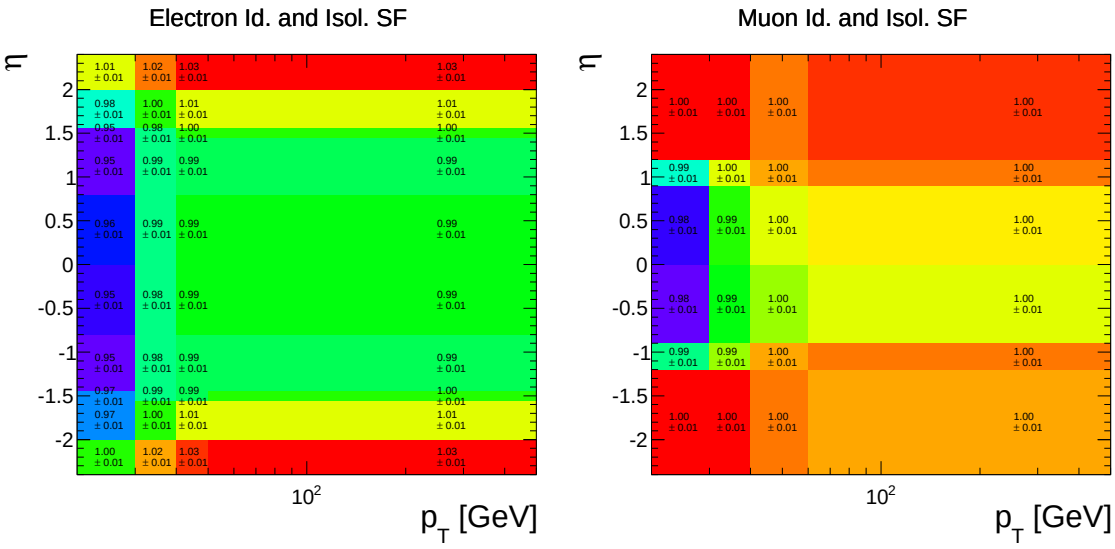
In order to reject leptons from semi-leptonic decays of hadrons isolation criteria are imposed. The relative isolation variable, $I_{\text{rel.}}$, is an estimator of the additional energy around the track of a given particle. It is defined as the sum of the photon, the charged and the neutral hadron energy relative to the transverse momentum p_{T} of the lepton inside a cone of

ΔR size around the track of the lepton:

$$I_{rel.} = \frac{\sum_{\text{photons}} E + \sum_{\text{charged had.}} E + \sum_{\text{neutral had.}} E}{p_{\text{T}}} \tag{5.1}$$

Electrons and muons are only considered if they are isolated, i.e. $I_{\text{rel.}}\,(\Delta\text{R} = 0.3) \leq 0.15$.

The lepton identification (Id.) and isolation (Isol.) efficiencies are estimated using a so-called *tag and probe* (T&P) method. The T&P algorithm exploits Z-boson event topologies and the $m^{\ell\bar{\ell}} \in [76, 106]\,\text{GeV}$ constraint imposed on the oppositely-charged pair of leptons produced in the Z-boson decay. One of the leptons (tag) is required to pass tight selection conditions in order to ensure a clean event sample, and the second lepton (probe) is used to determine the efficiency. The fraction of probe leptons passing in addition the lepton selection criteria used in the analysis defines the efficiency. Efficiencies are derived in data and $\text{Z}/\gamma^* \to \text{ee}/\mu\mu$ simulation and the data-to-simulation ratio (SF) is used to correct the simulated events for small differences observed between the data and the simulation efficiencies. The SFs are obtained double-differentially as a function of the p_{T} and η observables of the electron and muon separately. In general, the lepton Id. and isolation SF are larger than 95% in the phase space of interest for the analysis (see Figure 5.3).



(a) (b)

Figure 5.3: Distribution of the lepton Id. and Isol. correction factors as a function of lepton p_{T} and η. The results are shown for electrons (a) and muons (b) separately.

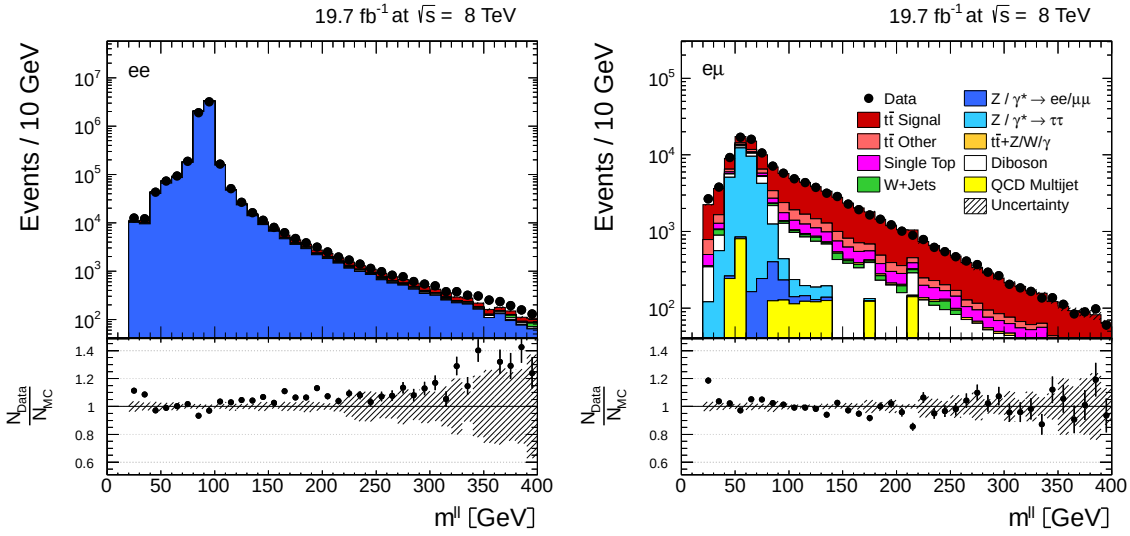
5.4 Lepton-Pair Selection

In every selected event, the oppositely-charged lepton pair maximising the transverse momentum $p_{\text{T}}^{\ell\bar{\ell}}$ is selected as the pair deriving from the $\text{t}\bar{\text{t}}$ decay. This criterion assigns uniquely the event to a particular dilepton $\text{t}\bar{\text{t}}$ decay channel (ee, eμ or $\mu\mu$) and removes ambiguities in events with three or more selected leptons.

The fraction of selected events from SM processes other than $\text{t}\bar{\text{t}}$ can be reduced with a constraint on the selected lepton-pair invariant mass. Low Drell-Yan resonances and QCD-multijet events are rejected by imposing a minimum dilepton mass of 20 GeV. Figure 5.4

shows the dilepton invariant mass distribution of events after the trigger, lepton and low $m^{l\bar{l}}$ requirements.

The eμ channel composition is dominated by t$\bar{\text{t}}$ signal events, $\sim 45\%$ of the selected events, while the ee channel is dominated by $Z/\gamma^* \rightarrow ee/\mu\mu$ events with approximately 98% of the selected events. A similar conclusion can be drawn for the $\mu\mu$ channel. In order to reduce the contamination from $Z/\gamma^* \rightarrow ee/\mu\mu$ events an additional veto on the dilepton invariant mass is applied to the same-flavour channels: ee and $\mu\mu$. Events with lepton-pair masses compatible with the Z-boson resonance, $76\,\text{GeV} \leq m^{l\bar{l}} \leq 106\,\text{GeV}$, are rejected.



(a) (b)

Figure 5.4: Dilepton invariant mass distributions for the ee (a) and eμ (b) final state channels obtained in events selected after trigger and lepton requirements. The contributions from different processes are presented according to the legend. The efficiency corrections due to PU and lepton identification and isolation are applied to the events from simulation. The hatched band corresponds to all shape uncertainties on the signal t$\bar{\text{t}}$ events (c.f. Chapter 7). The data-over-simulation ratio is displayed in the bottom panel.

5.5 Jets: Reconstruction, Calibration and Selection

The hard scattering in a hadron collider includes in general coloured particles, globally denoted as partons. Due to colour confinement partons fragment producing colour-neutral hadrons. The radiated hadrons are usually detected as a collimated set of particles along the direction of the original parton known as *jets*.

In order to estimate the properties of the original parton producing a jet, the anti-k$_t$ clustering algorithm is used with a $R = 0.5$ radius parameter [84]. The anti-k$_t$ algorithm merges the particle i into the object j (a particle or a proto-jet) if their distance mutual d_{ij} is smaller than $p_{\text{T},i}^{-2}$. The distance between two particles is defined as:

$$d_{ij} = \min\left(p_{\text{T},i}^{-2}, p_{\text{T},j}^{-2}\right)\frac{\Delta\text{R}^2}{R^2} \tag{5.2}$$

where ΔR is the angular separation between two particles as defined in Section 3.2, $p_{\mathrm{T},i}$ is the transverse momentum of the particle i and R is the cone-size parameter of the algorithm. The algorithm is iteratively applied to all particles and finishes when they are all clustered into jets.

The anti-k$_t$ algorithm produces circular-shaped jets which are infrared and collinear safe, i.e. the jet energy and momentum is invariant under the addition of infinitesimally low momentum particles or when the energy of a particle is distributed between two collinear particles. These properties are relevant since they permit the use of the same clustering algorithm in experimental observations and in theory predictions, thus facilitating direct comparisons.

As all reconstructed particles can be clustered in a jet, possible double counting effects occur when a lepton is both identified and included within a jet. An algorithm, referred to in CMS as *top projection* [85], solves these effects by removing from the clustering candidate list all isolated leptons as defined in Section 5.3.

A factorised jet calibration method [86] corrects sequentially the jet energy component for non-linear and non-uniform responses of the calorimeter relating, on average, the measured jet energy to its true value. A sketch of the different calibration steps applied to the jets in data and in simulation events is shown in Figure 5.5. Energy excess due to neutral hadrons

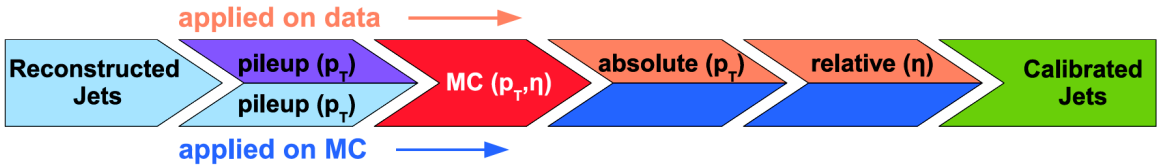


Figure 5.5: Diagram of the jet calibration steps applied on data and simulations.

originating from pileup collisions are removed first with a data-driven correction derived from minimum bias events. Additional corrections make jet response flat as a function of η and p_{T}, and correct the absolute scale. Angular dependent corrections are derived in dijet data events exploiting the energy-momentum balance in the transverse plane of the detector. The p_{T} response corrections are based on Drell-Yan events accompanied by jet emission and benefit from the excellent resolution of the lepton energy and momentum. Additional residual corrections are applied on data as a function of the jet pseudorapidity to correct small differences with respect to simulation.

The total energy correction is provided double differentially in several jet p_{T} and η ranges [87]. The corrections are generally smaller than 5% in the kinematic region of interest for this analysis.

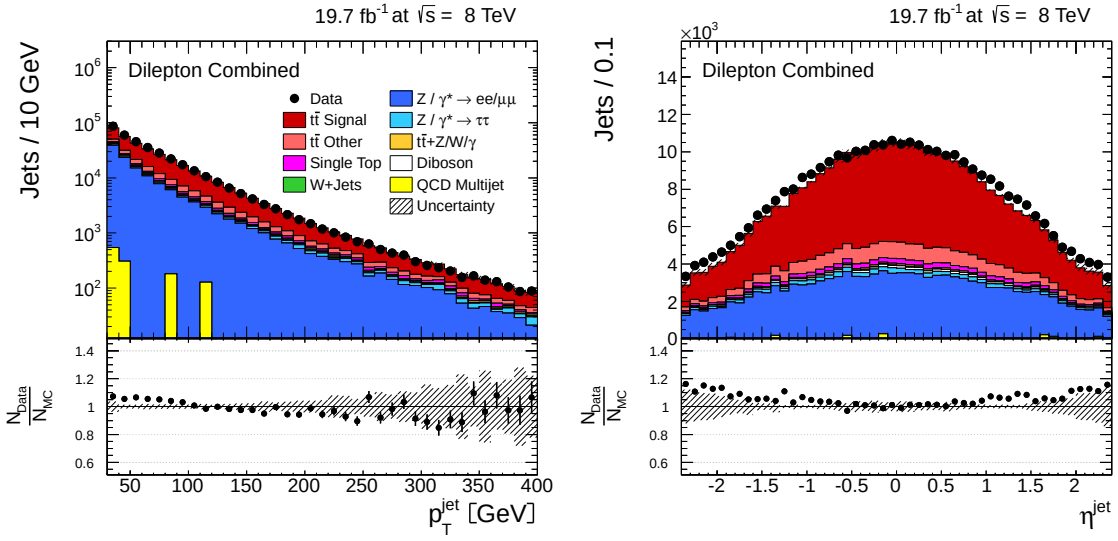
The differences in jet energy resolution between data and simulation are corrected with $|\eta|$-dependent factors applied to the simulated events. The correction factors are obtained in [87] by exploiting the momentum balance of dijet events in the transverse plane of the detector. Results are shown in Table 5.1.

$\lvert\eta\rvert$ Range	$[0, 0.5]$	$[0.5, 1.1]$	$[1.1, 1.7]$	$[1.7, 2.3]$	$[2.3, 2.8]$	$[2.8, 3.2]$	$[3.2, 5.0]$
JER Factor	1.079	1.099	1.121	1.208	1.254	1.395	1.056
$\pm1\sigma$ Uncertainty	$\pm$0.026	$\pm$0.028	$\pm$0.029	$\pm$0.046	$\pm$0.062	$\pm$0.063	$\pm$0.191

Table 5.1: Jet energy resolution (JER) correction factors, including the total uncertainty, in different reconstructed jet $|\eta|$ ranges [87].

The analysis selects events with at least two well reconstructed and calibrated jets, mea-

sured in the $|\eta| \leq 2.4$ detector region with a minimum jet transverse momentum of 30 GeV. These requirements enhance the fraction of selected $t\bar{t}$ signal events to $\sim 45\%$, with a similar amount of $Z/\gamma^* \to ee/\mu\mu$ background events. Figure 5.6 shows the jet p_T and η distributions after selecting events containing high transverse momentum jets.



(a) (b)

Figure 5.6: Distributions of the p_T (a) and η (b) variables of the jets selected after requiring the presence of at least two hard jets in the event. The efficiency corrections due to PU and lepton identification and isolation are applied to the events from simulation. The hatched band corresponds to all shape uncertainties on the signal $t\bar{t}$ events (c.f. Chapter 7). The data-over-simulation ratio is displayed in the bottom panel.

5.6 Missing Transverse Energy

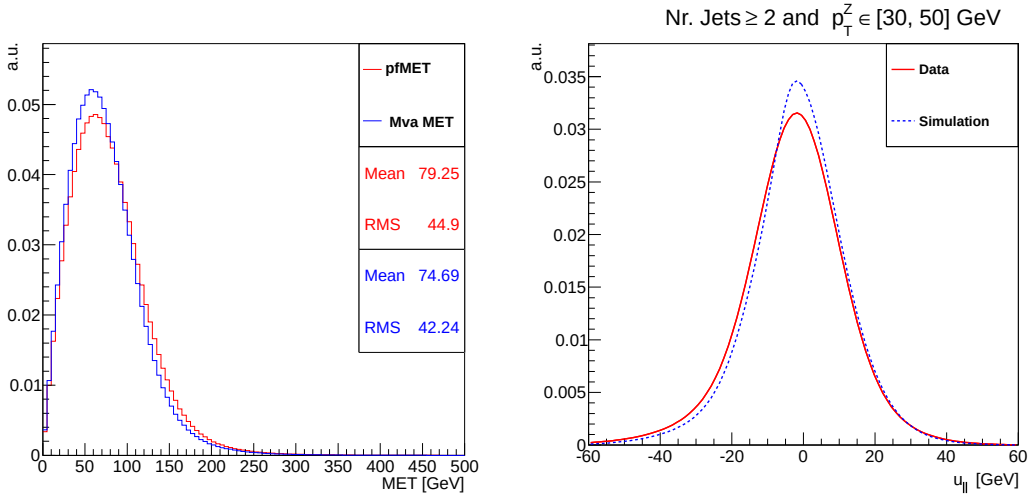
Weakly interacting particles, such as neutrinos, present in the final state of the event escape direct detection. Thus, their presence can only be inferred from energy and momentum imbalance of the measured particles in the detector transverse plane. The missing transverse momentum, $\vec{\not{E}}_T$, is defined as the negative vectorial sum of the transverse momenta of the leptons and jets present in the event. Together with its magnitude, called missing transverse energy and denoted by $\not{E}_T$, the undetected particle contribution to the event is determined.

$$\not{E}_T = \sqrt{\not{E}_{T,x}^2 + \not{E}_{T,y}^2} = \left| - \sum_{\text{leptons}} \vec{p}_T - \sum_{\text{jets}} \vec{p}_T \right| \tag{5.3}$$

Because of the sensitivity of the missing transverse energy to detector malfunctioning and to mismeasurements or misidentification of the detected particles, a precise and accurate reconstruction of the particles in the event is needed.

The resolution of the $\not{E}_T$ depends on the measured lepton and jet energy resolution, and thus is affected by contributions from PU vertices. The reconstruction of the $\not{E}_T$ is based on a MVA algorithm, described in detail in [88]. It uses jet-, PU- and event-related variables to

enhance the discriminating power against spurious contamination. The algorithm is trained in $Z/\gamma^* \to \mu^+\mu^-$ simulated events where the only contribution to the $\not\!\!E_T$ is caused by detector resolution effects. The MVA $\not\!\!E_T$ has an approximately 6% better resolution than particle-flow MET (pfMET) reconstruction methods used by the CMS Collaboration (see Figure 5.7a). The pfMET reconstruction uses all particles reconstructed by the PF-algorithm with the charged hadron subtraction method [88].



(a) (b)

Figure 5.7: (a): Missing transverse energy distributions obtained from the pfMET and from the MVA algorithm derived from dilepton $t\bar{t}$ simulation events. (b): Fit of the measured hadronic recoil $u_{\parallel}$ component obtained in data and $Z/\gamma^* \to e^+e^-$ simulation. The results are derived in events with two high-p_T oppositely-charged electrons with $m^{l\bar{l}} \in [60\,\text{GeV}, 120\,\text{GeV}]$, $30\,\text{GeV} \leq p_T^{l\bar{l}} \leq 50\,\text{GeV}$ and at least two jets with $|\eta| \geq 2.4$ and $p_T \geq 30\,\text{GeV}$. Figure taken from [89].

Data and simulation differences in the $\not\!\!E_T$ resolution are corrected with a *recoil* correction proposed in [89] and briefly described in the following. Dileptonic decays of the Z-boson do not contain intrinsic missing transverse energy source besides detector resolution effects or particle mismeasurements. Therefore, they can be used to accurately calibrate the $\not\!\!E_T$ response. The precise reconstruction of the Z-boson momentum allows an accurate determination of the hadronic-recoil momentum-components, $\vec{u} = -\vec{p}_T^{\,Z} - \vec{\not\!\!E}_T$. A Drell-Yan dominated sample of events is chosen in data and in simulation selecting events with two oppositely-charged electrons in the range $m^{ee} \in [60, 120]$ GeV and classified according to the jet multiplicity N_{jets} and to the Z-boson p_T. In each (N_{jet}, p_T^Z)-bin the recoil parallel ($u_{\parallel}$) and perpendicular ($u_{\perp}$) components, with respect to the Z-boson boost, are fitted in simulation and in data where the backgrounds have been subtracted. As an illustration, the result of the fit f obtained in data and simulation event with $\geq$2 jets and $30\,\text{GeV} \leq p_T^Z \leq 50\,\text{GeV}$ is shown in Figure 5.7b. Additional results can be found in [89].

The hadronic recoil components in DY simulation events are corrected ($u^{corr.}$) based on an isomorphic mapping of the data and simulation fits f. The mapping is such that the area of the data-fit equals that of the simulation for a measured $u^{uncorr.}$-value (see Equation (5.4)).

$$\int_{-\infty}^{u^{corr.}} f_{\text{data}} = \int_{-\infty}^{u^{uncorr.}} f_{\text{sim.}} \tag{5.4}$$

The isomorphic mapping applies identically for the parallel $u_{\parallel}$ and orthogonal $u_{\perp}$ components of the hadronic recoil.

The corrected recoil is then used to re-estimate the missing transverse energy: $\not\!\!E_{\mathrm{T}} = \left|-\vec{p}_{\mathrm{T}}^{Z} - \vec{u}^{corr.}\right|$. A significant improvement over the $\not\!\!E_{\mathrm{T}}$ distribution is achieved with the recoil corrections described above. The Figure 5.8 shows the $\not\!\!E_{\mathrm{T}}$ distribution measured in ee-channel events before and after the recoil corrections are applied, a similar improvement is found in the $\mu\mu$-channel.

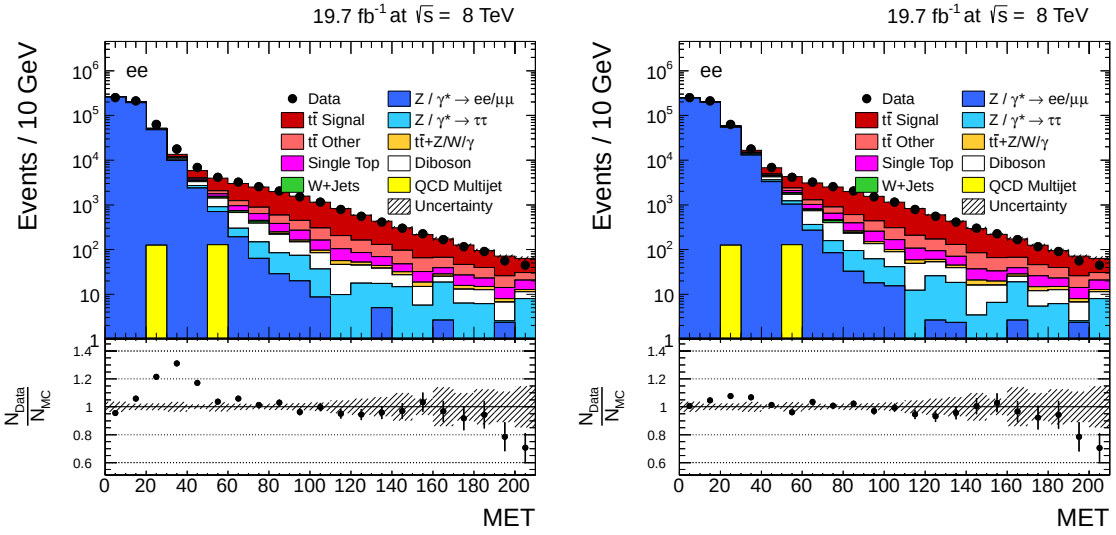


Figure 5.8: Missing transverse energy distribution of events selected in the ee channel after requiring the trigger, lepton, Z-veto and jet selection criteria. The distributions are shown before (a) and after (b) applying the recoil correction to $\mathrm{Z}/\gamma^{*} \to \mathrm{ee}/\mu\mu$ Monte Carlo events. The hatched band corresponds to all shape uncertainties on the signal $\mathrm{t\bar{t}}$ events (c.f. Chapter 7). The data-over-simulation ratio is displayed in the bottom panel.

Data and simulation events selected in the ee and $\mu\mu$ channels are required to have a minimum $\not\!\!E_{\mathrm{T}}$ of 40 GeV. This criteria rejects events from background processes with not real missing transverse energy.

The normalisation of the $\mathrm{Z}/\gamma^{*} \to \mathrm{ee}/\mu\mu$ simulation is derived from data and used to remove remaining differences in the $\not\!\!E_{\mathrm{T}}$ distribution [90–92]. The data-based $R_{in/out}$ method scales the $\mathrm{Z}/\gamma^{*} \to \mathrm{ee}/\mu\mu$ event yield in ee and $\mu\mu$ channels to the Drell-Yan event yield from data in the $76 \leq m^{\mathrm{l\bar{l}}} \leq 106\,\mathrm{GeV}$ control region. The scaling factor is then used to normalise the $\mathrm{Z}/\gamma^{*} \to \mathrm{ee}/\mu\mu$ simulation in the analysis region, $m^{\mathrm{l\bar{l}}} \notin [76\,\mathrm{GeV}, 106\,\mathrm{GeV}]$. The fraction of simulation events in the control region of the eμ channel provides an estimate of background events which are subtracted from data in order to estimate the fraction of Drell-Yan events. The scaling factor results are 0.89 in the ee, 0.96 in the $\mu\mu$ and 0.92 in the eμ channel, respectively. The latter is obtained as the geometric mean of the same-flavour channel normalisation factors as proposed in [92].

5.7 Identification and Selection of b-Quark Jets

Hadrons containing b-quarks have a lifetime of about 1.6 ps and thus travel $\sim 500\,\mu$m before they decay. The relatively large travel distance can be measured in the inner tracker making possible an identification of original parton-flavour of the jet.

The *combined secondary vertex* (CSV) b-tagging algorithm combines kinematic and geometric properties of B-hadron decays into a multivariate discriminant which distinguishes jets originating from b-quarks and from other quarks [93]. Three working points (WP) are defined as a function of the light-jet misidentification efficiency which decreases by approximately an order of magnitude between two consecutive WPs. Table 5.2 summarises the CSV discriminator values for each WP together with the light-jet misidentification and the b-jet tagging efficiency values obtained from simulation [93]. In this analysis a b-jet is considered b-tagged

Name	Threshold	Light-Jet Efficiency [%]	b-Jet Efficiency [%]
Loose (CSVL)	≥ 0.244	~ 10	~ 80
Medium (CSVM)	≥ 0.679	~ 1	~ 65
Tight (CSVT)	≥ 0.898	~ 0.1	~ 50

Table 5.2: CSV b-tagging discriminant threshold for the different working points of the algorithm [93]. The approximate light-jet misidentification and b-jet identification efficiencies are also indicated.

if the discriminator value exceeds the 0.244 threshold. This corresponds to the loose working point.

The efficiency of the CSV algorithm is estimated in data and simulated QCD-multijet events [93] and shown in Figure 5.9a. In order to correct for small differences observed in the data and simulation efficiencies, an efficiency ratio (SF) is derived as a function of the b-tagging algorithm working point, the jet flavour and jet kinematics (η, p_{T}). The SF results for the CSVL algorithm for heavy-flavoured jets are larger than 96% in the full rapidity region increasing up to approximately 100% with lower momenta, while light-jet results present different functional dependencies as a function of the η regions. The SF for the light-jet mis-reconstruction efficiency range between 1.0% and 1.4%. Figure 5.9b shows the simulation correction factor for heavy-flavoured b-tagged jets as a function of the jet transverse momentum in the $0 \leq |\eta| \leq 2.4$ range. The light-jet SFs are shown in Figure 5.9c as a function of the jet p_{T}, separately for different $|\eta|$-regions.

The b-tagging efficiency of the analysis, ε_{MC}, is used to correct the efficiency differences of the CSV algorithm between data and simulation. ε_{MC} is defined as the fraction of selected jets which are in addition tagged by the CSV algorithm. It is derived from a $\mathrm{t\bar{t}}$ simulation sample and is parametrised as a function of the p_{T}, pseudorapidity and flavour of the jets as shown in Figure 5.10. The efficiencies are in general larger than 75% for b-flavoured jets in the selected kinematic region, while the c- and light-flavoured jet misidentification efficiency varies between 5% and 50% depending on the specific flavour and jet kinematic properties. Similar results are obtained in ee and $\mu\mu$ channels.

In order to match the efficiency measured in simulation to the results from data, the individual jet b-tagging status is updated. A fraction $1 - SF$ of a priori b-tagged jets is downgraded to untagged status if the correction factor is $SF < 1$ and a fraction $(1-SF)/(1-1/\varepsilon_{MC})$ of untagged jets is promoted to a b-tagged status only if $SF > 1$.

After the tagging status of the jets has been updated according to the method described above, the analysis proceeds by selecting events in data and simulation if they contain at least

one CSVL b-tagged jet. Figure 5.11 shows the b-tagged jet multiplicity and selected jet p_{T} distribution before and after selecting events with at least one b-tagged jet, respectively.

5.8 Reconstruction of the Top-Quark-Kinematics

The analysis presented in this thesis measures the differential $\mathrm{t\bar{t}}$ production cross section as a function of several kinematic variables of the $\mathrm{t\bar{t}}$ system, individual top quarks and decay products. The top quark decays before hadronisation, as described in Chapter 2. Therefore, a direct experimental detection of its kinematics is not possible and must be reconstructed from the final state lepton, b-jet and neutrino 4-momenta under the assumption that $t \to b + W(\to l\bar{\nu})$:

$$p_i(t) = p_i(b) + p_i(l) + p_i(\bar{\nu}), \text{ where } i = x, y, z, t \tag{5.5}$$

A similar conditions applies to the top antiquark.

The complex $\mathrm{t\bar{t}}$ dileptonic final state presents several limitations. The neutrinos escape the detector and their presence is only inferred as energy-momentum imbalance in the transverse plane of the detector, $\not{E}_{\mathrm{T}}$. Furthermore, the individual neutrino contribution to the $\not{E}_{\mathrm{T}}$ cannot be disentangled. The selected leptons can be assigned to the top quark or antiquark decay branch according to their electric charge information. The electric charge of the b-jet cannot be used to associate the jet to the separate decay branches because this information is diluted in the hadronisation process due to the presence of oppositely-charged muons from semi-leptonic decays of D mesons in the charged B-hadron decay products.

The 4-momenta of the top quarks and antiquarks are derived using a reconstruction algorithm similar to the Analytical Matrix Weighting Technique (AMWT) [94]. It is based on the 4-momenta of the selected leptons, jets and $\not{E}_{\mathrm{T}}$ present in the set of events surviving all selection criteria described above. The kinematic reconstruction is adapted to select the $\mathrm{t\bar{t}}$ final state which fits the fully leptonic decay hypothesis best. It is required, furthermore, that the algorithm assigns the reconstructed jets to the b-quarks from the $\mathrm{t\bar{t}}$ decay.

Mathematical Formulation

The kinematics of $\mathrm{t\bar{t}}$ dilepton events can be expressed as an under-constrained system of equations. One pair of equations relate the projection of the $\vec{\not{E}}_{\mathrm{T}}$ components to the sum of the neutrino and antineutrino momentum components belonging to the same projection. It is assumed that the neutrinos produced in dilepton $\mathrm{t\bar{t}}$ decay are the only contribution to the $\vec{\not{E}}_{\mathrm{T}}$. Two additional equations represent the relativistic energy-momentum neutrino and antineutrino relations. The remaining equations describe the W-boson and top-quark mass constraints by relating the invariant masses to the energy and momenta of their decay particles via relativistic 4-vector arithmetic. Mathematically the $\mathrm{t\bar{t}}$ event kinematics are expressed as

$$\begin{aligned}
\vec{\not{E}}_{\mathrm{T}} &= \vec{p}(\nu) + \vec{p}(\bar{\nu}) \\
E^2(\nu) &= m^2(\nu) + |\vec{p}(\nu)|^2 \\
E^2(\bar{\nu}) &= m^2(\bar{\nu}) + |\vec{p}(\bar{\nu})|^2 \\
m^2(\mathrm{W}^+) &= \left[E(\mathrm{l}^+) + E(\nu)\right]^2 - \left|\vec{p}(\mathrm{l}^+) + \vec{p}(\nu)\right|^2 \\
m^2(\mathrm{W}^-) &= \left[E(\mathrm{l}^-) + E(\bar{\nu})\right]^2 - \left|\vec{p}(\mathrm{l}^-) + \vec{p}(\bar{\nu})\right|^2 \\
m^2(\mathrm{t}) &= \left[E(\mathrm{b}) + E(\mathrm{l}^+) + E(\nu)\right]^2 - \left|\vec{p}(\mathrm{b}) + \vec{p}(\mathrm{l}^+) + \vec{p}(\nu)\right|^2 \\
m^2(\bar{\mathrm{t}}) &= \left[E(\bar{\mathrm{b}}) + E(\mathrm{l}^-) + E(\bar{\nu})\right]^2 - \left|\vec{p}(\bar{\mathrm{b}}) + \vec{p}(\mathrm{l}^-) + \vec{p}(\bar{\nu})\right|^2
\end{aligned} \tag{5.6}$$

with the unknowns being the neutrino 4-momentum components and the W-boson and top-quark masses.

The number of degrees of freedom of the system of equations (5.6) is reduced by imposing equality between the $W^{\pm}$ boson masses, $m(W^+) = m(W^-)$, and between the top quark and antiquark masses, $m(t) = m(\bar{t})$. A second assumption considers the leptons and neutrinos massless because their momenta in $t\bar{t}$ events are in general much larger than their invariant masses $p_T^l \gtrsim 20\,\text{GeV} \gg 100\,\text{MeV} \gtrsim m(l) \gg m(\nu) \simeq 0\,\text{GeV}$. The invariant masses of the b-quarks are set to $m^b = m^{\bar{b}} = 4.8\,\text{GeV}$, the values used in the $t\bar{t}$ simulation.

With these assumptions the kinematic equations of the $t\bar{t}$ events can be expressed without loss of generality as a 4^{th}-order polynomial as a function of the x-component of the antineutrino 4-momentum:

$$0 = \sum_{i=0}^{4} c_i(m^t, p_{l^+}, p_{l^-}, p_b, p_{\bar{b}})\, p_x(\bar{\nu})^i \tag{5.7}$$

The energy and momentum resolution of the measured leptons and jets is not perfect and can yield solutions of Equation (5.7) with no physical meaning, e.g. negative energy values or imaginary momentum values. This unphysical situations can be reduced by smearing the energy, the scalar momentum and the direction of measured leptons and jets 100 times according to the resolution distributions derived from $t\bar{t}$ signal simulation. Additionally, the reconstructed W-boson mass is also smeared with the resolution distribution obtained in $t\bar{t}$ simulation around the nominal mass value of 80.4 GeV [12]. The solution of Equation (5.7) is searched for every smeared event configuration recovering event solutions a priori non-valid.

Resolving the Ambiguity Problem

For each smearing of the event up to four mathematically equivalent solutions of the Equation (5.7) can exist. This ambiguity is disentangled by selecting the solution yielding the lowest $t\bar{t}$ invariant mass value [95]. This assumption is based on the proportionality between $m^{t\bar{t}}$ and the Bjorken-x variable. The Bjorken-x distribution presents a very rapidly decreasing probability distribution and therefore large values are less likely to occur. The top-quark (analogously for the top-antiquark) 3-momentum components are obtained as the weighted average of the selected solution in each of the smeared event configuration $\vec{p}_{t,i}$.

$$\langle \vec{p}_t \rangle = \frac{1}{w} \sum_{i=1}^{100} w_i \cdot \vec{p}_{t,i} \tag{5.8}$$

$$w = \sum_{i=1}^{100} w_i \tag{5.9}$$

where the weight w_i is obtained from the true lepton-b-jet invariant mass probability distribution. The last component of the top-quark 4-momentum vector, the energy, is obtained from the relativistic energy-momentum relations and the constraint $m^t = 172.5\,\text{GeV}$. Possible biases due to the fixed top quark mass constraint used in the kinematic reconstruction algorithm are discussed in Chapter 7.

Selection of the Lepton-Jet-Pair

All well reconstructed and selected jets (see Section 5.5) in the event are considered in the kinematic reconstruction algorithm. Thus several combination of jets could yield a valid

Channel	ee	eμ	$\mu\mu$
$\varepsilon_{\text{data}}$ [%]	89.62	92.87	90.11
$\varepsilon_{\text{simu.}}$[%]	90.75	93.62	90.63
SF	0.987 ± 0.004	0.992 ± 0.001	0.994 ± 0.004

Table 5.3: Data and simulation efficiency values of the kinematic reconstruction algorithm in the ee, eμ and $\mu\mu$ channel. The data-to-simulation efficiency ratio value, together with the statistical uncertainty, is indicated in the separate channels.

result of Equation (5.7). In cases with two oppositely-charged leptons (l_1 and l_2) and only two selected jets (j_1 and j_2) two possible lepton-jet assignments exist:

- l_1–j_1 and l_2–j_2
- l_1–j_2 and l_2–j_1

The number of possible combinations increases with the number of jets present in the event.

The kinematic reconstruction algorithm is tested for every possible combination of lepton-jet. The preferred configurations contain two b-tagged jets. Only if no valid solution is found, lepton-jet combinations with only one b-tagged jet are used. Combinations with no b-tagged jets are discarded.

Only events with a valid solution of the kinematic reconstruction algorithm are selected. The performance of the reconstruction algorithm is studied as a function of several event and particle kinematic distributions. The kinematic reconstruction method used here yields on average a reconstruction efficiency of about 90% in data and simulation. In general, the data and simulation efficiencies are found to be larger than 80% up to very large jet multiplicities and $\not\!\!E_{\text{T}}$ values, while the results at very high lepton-p_{T} values drop to $\sim 65\%$ with a statistical uncertainty of the order of 20–30% (see Figure 5.12). Small differences in the efficiency between simulation and data are corrected in simulation events with a global efficiency ratio (SF). Numerical values and statistical uncertainties are presented in Table 5.3.

5.9 Event Yields and Control Distributions

In the following a summary of the selection criteria is presented:

- Events are selected based on dilepton triggers which require the presence of an electron-pair, muon-pair or electron-muon-pair with lepton transverse momenta larger than 17 and 8 GeV, respectively.
- The event must contain at least one well reconstructed vertex in the vicinity of the nominal pp interaction point.
- Events must have at least two oppositely-charged, isolated and high-p_{T} leptons (electron or muon).
- A minimum lepton-pair invariant-mass value of 20 GeV rejects events from low mass QCD-multijet and Drell-Yan processes. Additionally, a veto on the Z-boson mass window is imposed on the ee and $\mu\mu$ channels.

Sample	$\mu\mu$	eμ	ee
Data	14403	39640	10678
Total Simulation	13939.1	38574.3	10221.5
Fraction of Events (%)			
$t\bar{t}$ signal	78.0	79.5	77.2
$t\bar{t}$ other	12.9	13.5	13.0
Single top	3.4	3.5	3.2
$Z/\gamma^* \to ee/\mu\mu$	3.4	0.1	2.4
$Z/\gamma^* \to \tau\tau$	0.7	1.2	1.1
$t\bar{t}$+W/Z/γ	1.1	1.0	1.1
Diboson	0.4	0.4	0.7
W+jets	0.1	0.2	0.3
QCD-multijet	0.0	0.6	1.0

Table 5.4: Number of selected data events and relative composition of the different simulation processes after applying the full event selection criteria including the top quark kinematic reconstruction. The results are shown separately for the ee, eμ and $\mu\mu$ channels. The events from simulation are normalised to the data luminosity $L = 19.7\,\text{fb}^{-1}$ and corrected for efficiency differences between data and simulation.

- The events must have at least two well reconstructed and calibrated jets in the region $|\eta| \leq 2.4$ with a minimum transverse momentum of at least 30 GeV.
- A minimum $\not\!\!E_{\text{T}}$ of 40 GeV is required on ee and $\mu\mu$ channels.
- The presence of at least one CSVL b-tagged jet is imposed to select events.
- Selected events must have a valid solution of the kinematic reconstruction algorithm.

A total of ~ 62700 data events are selected, where $\sim 78\%$ of the sample corresponds to $t\bar{t}$ production in the dilepton channel ($t\bar{t}$ signal). All other $t\bar{t}$ events ($t\bar{t}$ other), specifically those originating from decays via τ leptons, are considered as background and amount to approximately 10% of the final event sample. Semi-leptonic and fully-hadronic $t\bar{t}$ decay channels represent a negligible contribution to the final event sample due to the lepton selection criteria. Remaining background processes are due to electroweak single-top production $\sim 3.4\%$, $Z/\gamma^* \to ee/\mu\mu$, $t\bar{t}$ production in association with photons, W and Z bosons ($t\bar{t}$+W/Z/γ) $\sim 1\%$ each, Z-boson decays via τ leptons ($Z/\gamma^* \to \tau\tau$), diboson events (WW, WZ and ZZ), W boson produced in association with jets and QCD-multijet production. In Table 5.4, the selected number of data events and the expected relative contributions of the signal and background processes are presented in the $\mu\mu$, eμ and ee channels separately. The expected yields agree well with the data.

In the following, control distributions are presented to evaluate the description of data by the simulation. All efficiency correction factors described previously are applied to the simulated events which are additionally scaled to the data luminosity, $L = 19.7\,\text{fb}^{-1}$.

In Figure 5.13 the p_{T} and η distributions of the two selected leptons are shown. The spectra are shown in the combined channel defined as the sum of events reconstructed and selected in ee, eμ and $\mu\mu$ channels. The lepton p_{T} shows a good agreement in the region $\lesssim 90$ GeV, while at higher momenta data presents a softer spectrum than the prediction with

differences of $\sim 20\%$. The pseudorapidity distribution of the selected leptons is well described by the simulation within the systematic uncertainties on the signal shape, indicated by the hatched band (see Chapter 7).

The $\not\!\!E_{\mathrm{T}}$ data distribution is well described by the simulations up to 150 GeV, where a slight excess of Monte Carlo events appears. Differences between data and simulation are covered by the $\mathrm{t\bar{t}}$ signal systematic uncertainty as shown in Figure 5.14.

In the case of the kinematic variables of the lepton-pair, the simulation describes the data distributions of the mass and transverse momentum well. Any small fluctuation is contained within the statistical uncertainty of the data. The transverse momentum of the lepton pair is shown in Figure 5.15a and 5.15b, separately for the ee and eμ channels. The lepton-pair invariant mass distribution is presented in Figures 5.15c and 5.15d for ee and eμ channel, respectively. The $\mu\mu$ channel, not shown, presents a similar behaviour as the ee channel.

The transverse momentum distribution of the two selected b-jets shows a shape difference between data and simulated events at values $p_{\mathrm{T}}^{\mathrm{b}} \geq 120\,\mathrm{GeV}$, while below this threshold the agreement is good (see Figure 5.16a). This observation is similar to the lepton p_{T} result discussed previously. Simulation predicts more central b-jet pseudorapidity values in comparison to data (see Figure 5.16b). Small differences are contained within the $\mathrm{t\bar{t}}$ signal uncertainties.

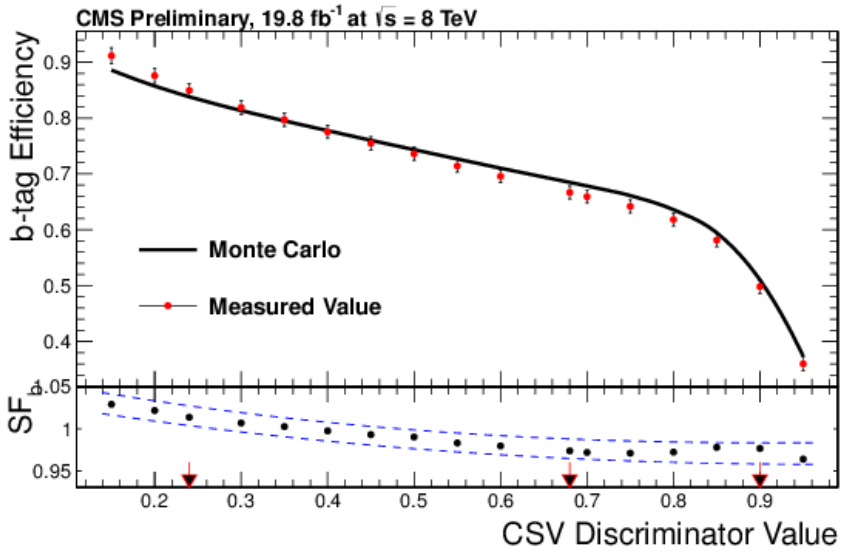
The transverse momentum $p_{\mathrm{T}}^{\mathrm{b\bar{b}}}$ and invariant mass $m^{\mathrm{b\bar{b}}}$ distributions of the system of the b-jet pairs are shown in Figure 5.17. A generally good description of the measured distributions is observed.

The top-quark-pair transverse momentum $p_{\mathrm{T}}^{\mathrm{t\bar{t}}}$, the rapidity $y^{\mathrm{t\bar{t}}}$, and the invariant mass $m^{\mathrm{t\bar{t}}}$, are shown in Figures 5.18a, 5.18b and 5.18c, respectively. Distributions from simulation are in good agreement with those observed in data and the differences are covered by the shape uncertainties on the signal. Data tends to slightly higher values of top quark rapidity, y^{t}, but differences of the simulation with respect to data are contained within the signal uncertainties. Figure 5.18d presents the y^{t} distribution.

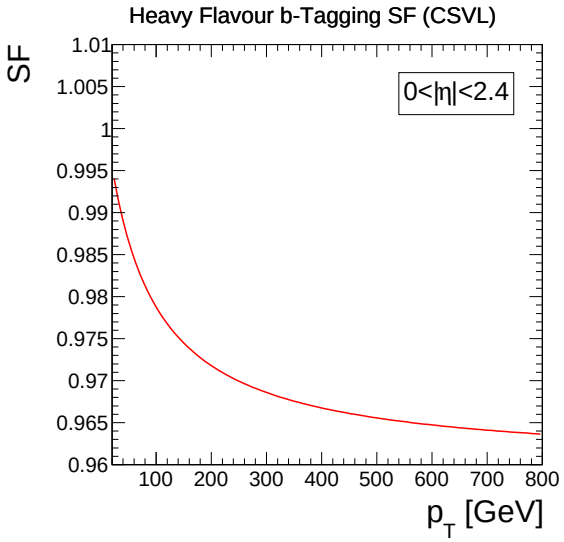
Distributions of the transverse momentum of the top quarks, $p_{\mathrm{T}}^{\mathrm{t}}$, and separately of the p_{T}-ordered leading, $p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{lead.\,t})$, and trailing $p_{\mathrm{T}}^{\mathrm{t}}(2^{\mathrm{nd}}\,\mathrm{lead.\,t})$ top quark or antiquark are shown in Figure 5.19. Data shows a softer spectrum than the predictions. A similar observation is found if the top quark p_{T} is measured in the $\mathrm{t\bar{t}}$ reference system (see Figure 5.19b). Similar results have been observed in several CMS analyses [91,96–98], where measurements are performed using different datasets, alternative event selection criteria and distinct reconstruction algorithms of the top quark 4-momenta.

Distributions of angular differences between the lepton and antilepton or top quark and antiquark are well described by the event simulation as displayed in Figure 5.20.

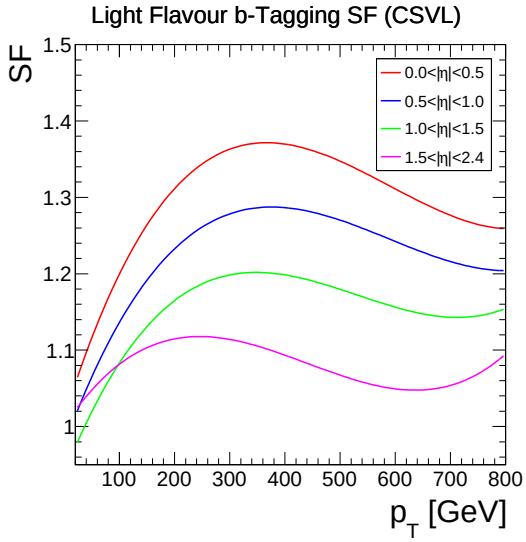
In order to evaluate the possible sources of the differences in the p_{T} distributions of the leptons, b-jets and top quarks, the analysis is repeated in three exclusive PU regions: $1 \leq PU \leq 10$, $10 < PU \leq 20$ and $PU > 20$. The results are shown in Appendix D.1. Similar results are observed in all three regions, thus no bias due to the PU is found.



(a)

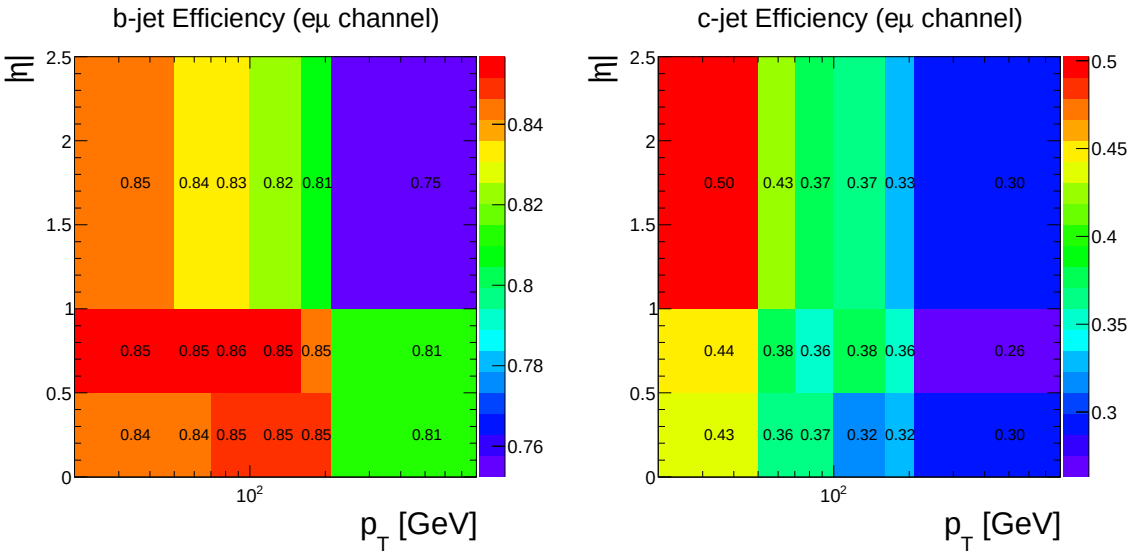


(b)



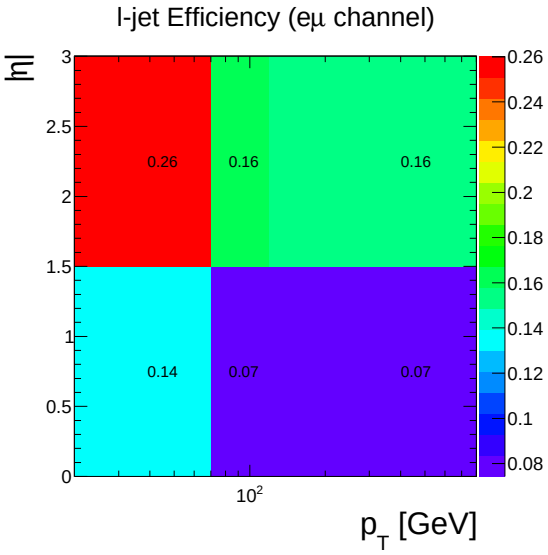
(c)

Figure 5.9: (a): b-jet tagging efficiency as a function of the discriminator threshold of the CSV algorithm measured in data and simulation. In the bottom panel the efficiency ratio is shown together with the total uncertainty indicated by the dashed lines. The three WPs of the CSV algorithm are indicated by the vertical arrows. Figure (modified) taken from [93]. Distribution of the simulation correction factor for the CSVL b-tagging algorithm as a function of the jet p_{T} for heavy (b) and light flavour (c) jets respectively [93]. The pseudorapidity region of validity is indicated in the legend.



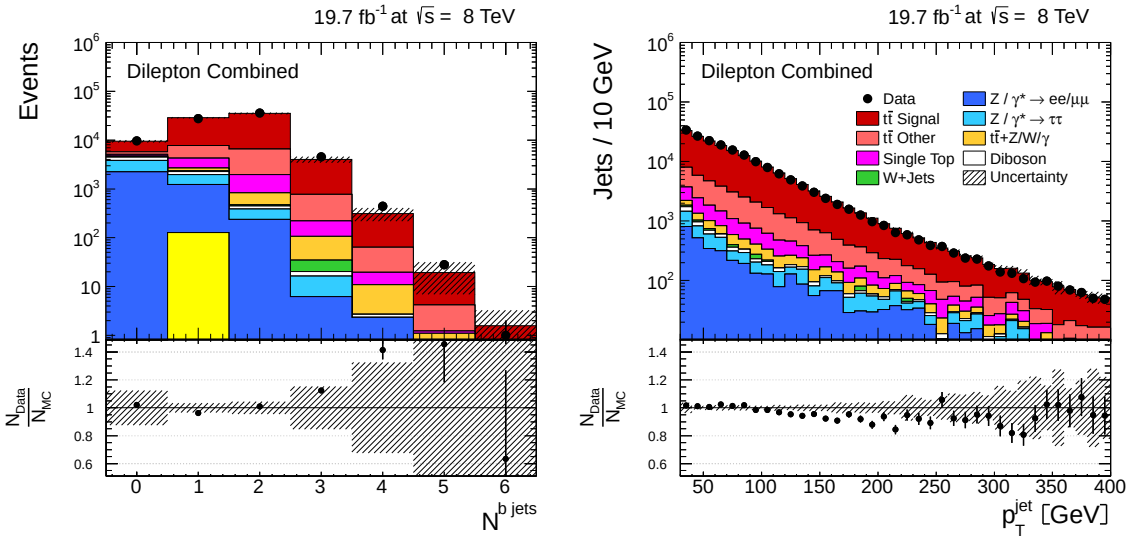
(a)

(b)



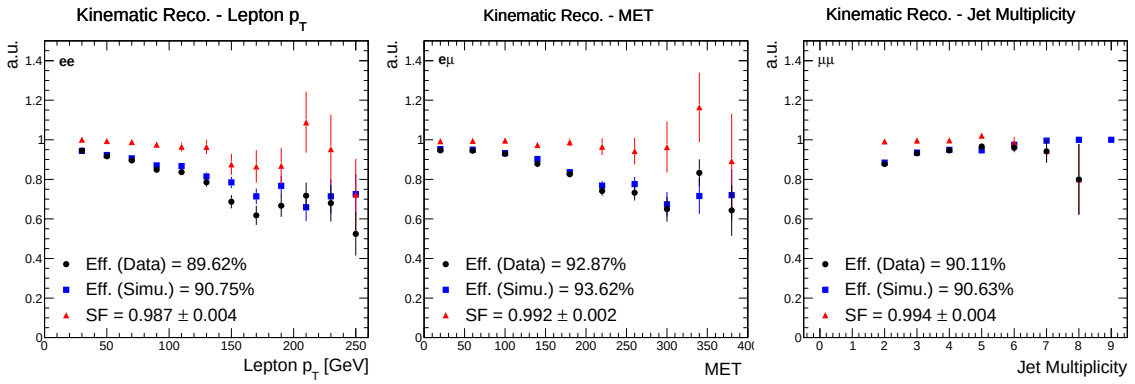
(c)

Figure 5.10: b-tagging efficiency distribution of the b-jets (a), c-jets (b) and light-jets (c) as a function of the jet p_{T} and η. The results are derived from $\mathrm{t\bar{t}}$ simulation events selected the eμ channel after dilepton, jet and $\not\!\!E_{\mathrm{T}}$ requirements.



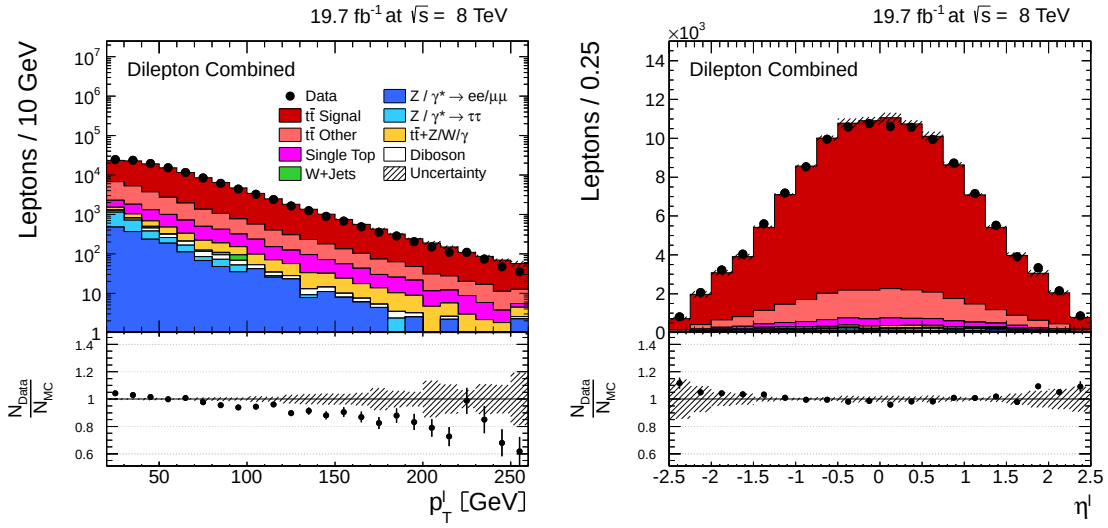
(a) (b)

Figure 5.11: (a): b-tagged jet multiplicity distribution of events after lepton, jet and $\not\!\!E_{\mathrm{T}}$ selection. (b): Transverse momentum distribution of jets in events with at least one CSVL b-tagged jet. The efficiency corrections due to PU and lepton identification and isolation are applied to the events from simulation. The hatched band corresponds to all shape uncertainties on the signal $\mathrm{t\bar{t}}$ events (c.f. Chapter 7). The data-over-simulation ratio is displayed in the bottom panel.



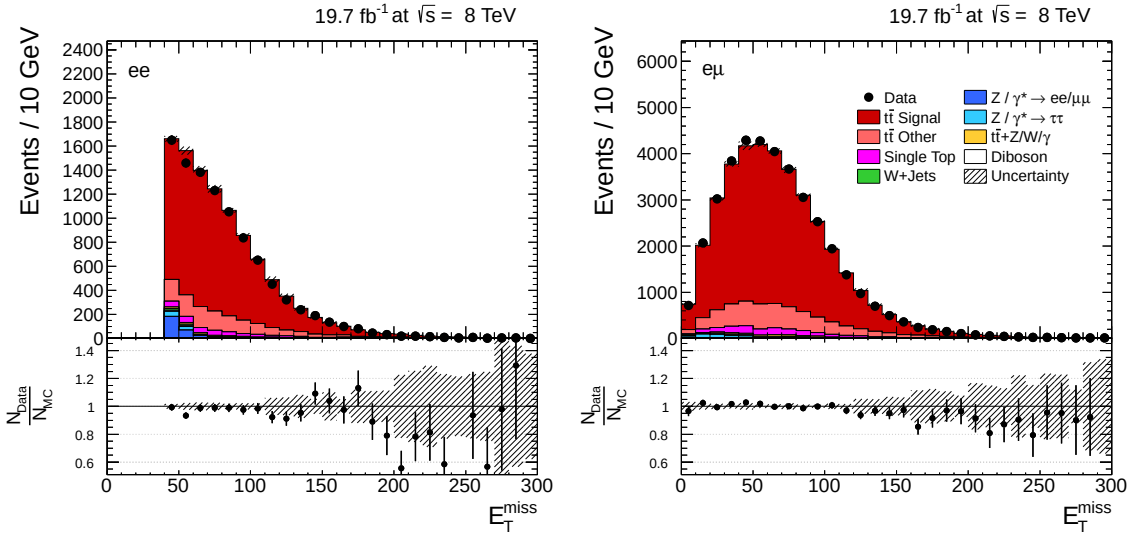
(a) (b) (c)

Figure 5.12: Distribution of the efficiency of the kinematic reconstruction for data (circle) and simulation (square), and data-to-simulation efficiency ratio, SF, (triangle). Results are shown as a function of the leading lepton p_{T} in the ee channel (a), the $\not\!\!E_{\mathrm{T}}$ in the emu channel (b) and the jet multiplicity in the $\mu\mu$ channel (c). The error bars represent only the statistical uncertainty on the result. The total efficiency in data and simulation events and the total efficiency ratio values are indicated in the legend of the figures.



(a) (b)

Figure 5.13: Distributions of the transverse momentum $p_{\mathrm{T}}^{\mathrm{l}}$ (a) and pseudorapidity η^{l} (b) of the two selected leptons. Results are presented after the full event selection. The error bars in the data points represent the statistical uncertainty and the hatched band corresponds to all shape uncertainties on the $\mathrm{t\bar{t}}$ simulation events (c.f. Chapter 7). The bottom panels present the data-to-simulation ratio.



(a) (b)

Figure 5.14: Distribution of the $\not{E}_{\mathrm{T}}$ measured in the ee (a) and eμ (b) channels separately. Results are presented after the full event selection. The error bars in the data points represent the statistical uncertainty and the hatched band corresponds to all shape uncertainties on the $\mathrm{t\bar{t}}$ simulation events (c.f. Chapter 7). The bottom panels present the data-to-simulation ratio.

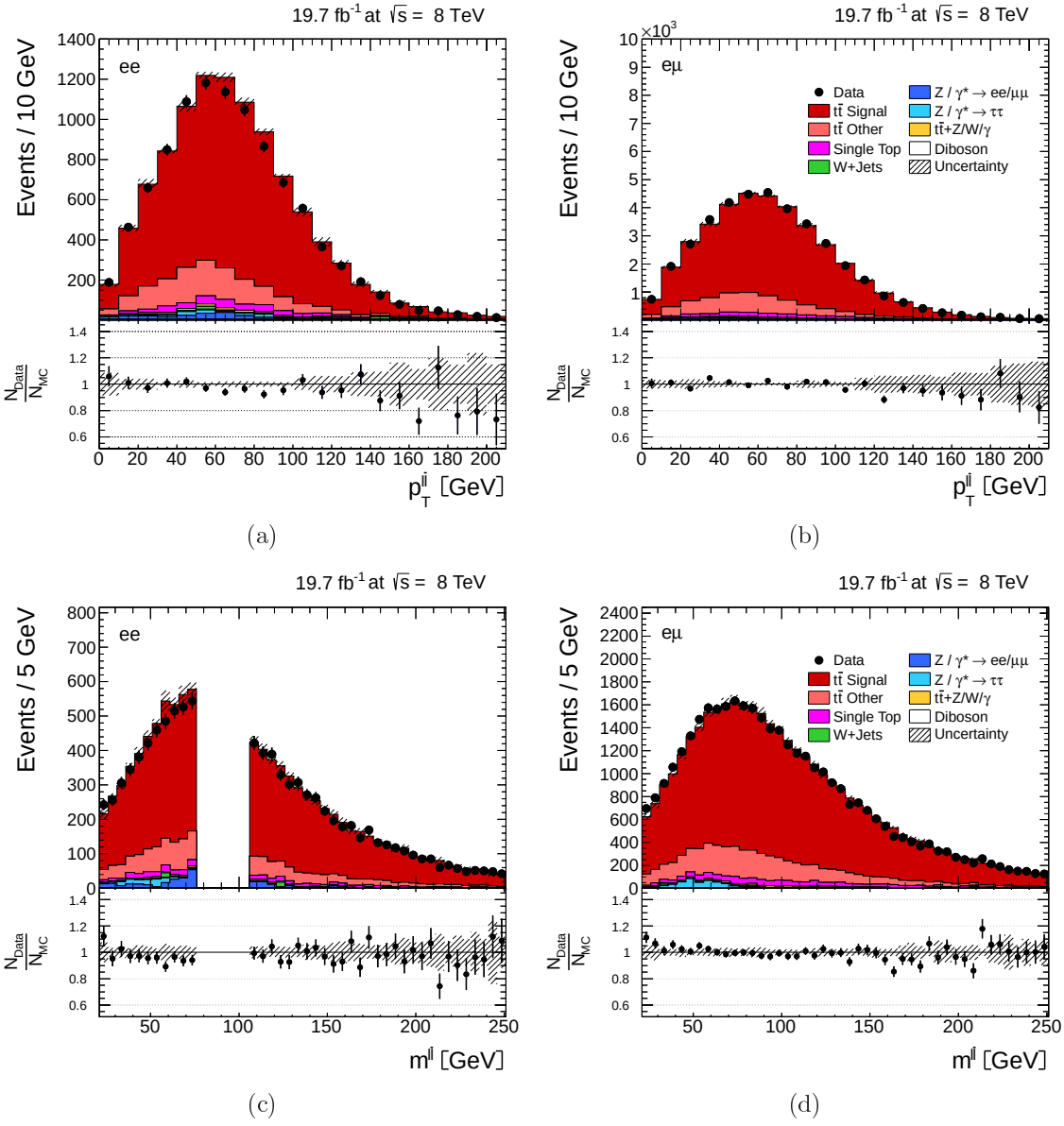
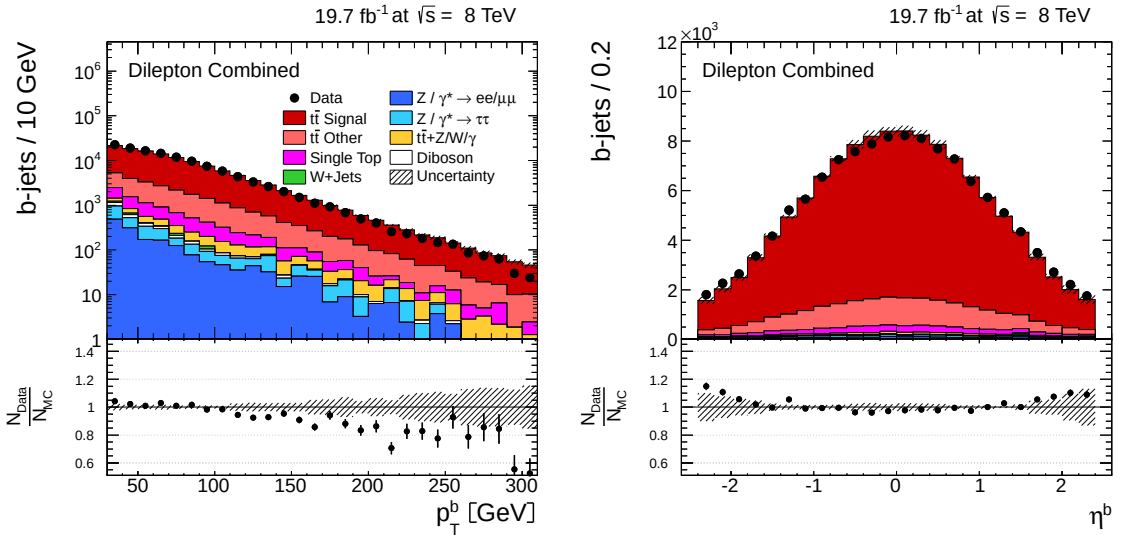
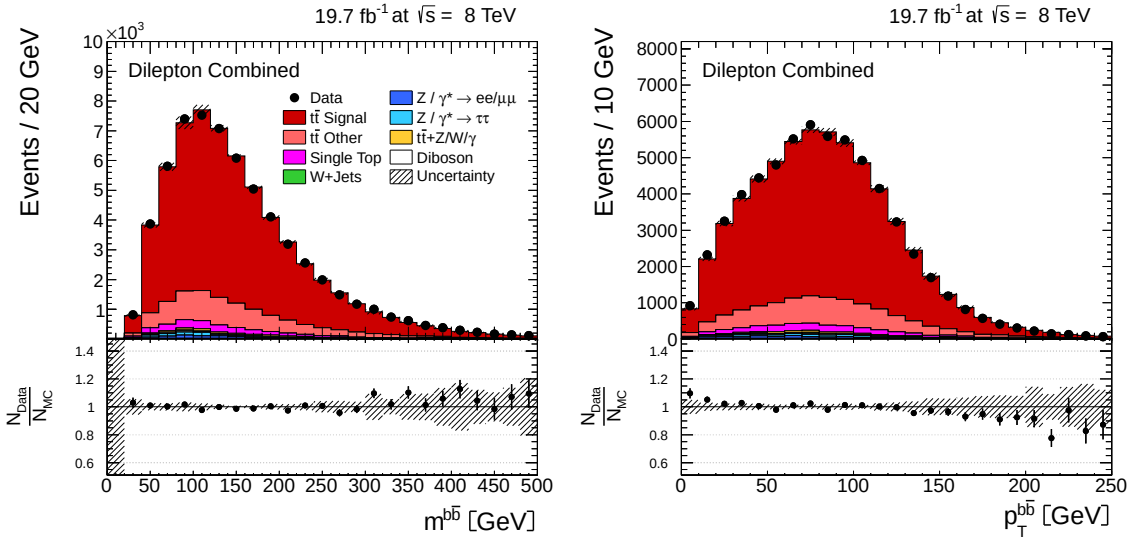


Figure 5.15: Distributions of the lepton-pair transverse momentum $p_{\mathrm{T}}^{\mathrm{l\bar{l}}}$, and invariant mass $m^{\mathrm{l\bar{l}}}$, reconstructed in the ee (a and c) and eμ (b and d) channels separately. Results are presented after the full event selection. The error bars in the data points represent the statistical uncertainty and the hatched band corresponds to all shape uncertainties on the $\mathrm{t\bar{t}}$ simulation events (c.f. Chapter 7). The bottom panels present the data-to-simulation ratio.



(a) (b)

Figure 5.16: Distributions of the transverse momentum $p_{\mathrm{T}}^{\mathrm{b}}$ (a) and pseudorapidity η^{b} (b) of the two selected jets. Results are presented after the full event selection. The error bars in the data points represent the statistical uncertainty and the hatched band corresponds to all shape uncertainties on the $\mathrm{t\bar{t}}$ simulation events (c.f. Chapter 7). The bottom panels present the data-to-simulation ratio.



(a) (b)

Figure 5.17: Distributions of the invariant mass $m^{\mathrm{b\bar{b}}}$ (a) and transverse momentum $p_{\mathrm{T}}^{\mathrm{b\bar{b}}}$ (b) of the b-jet pair. Results are presented after the full event selection. The error bars in the data points represent the statistical uncertainty and the hatched band corresponds to all shape uncertainties on the $\mathrm{t\bar{t}}$ simulation events (c.f. Chapter 7). The bottom panels present the data-to-simulation ratio.

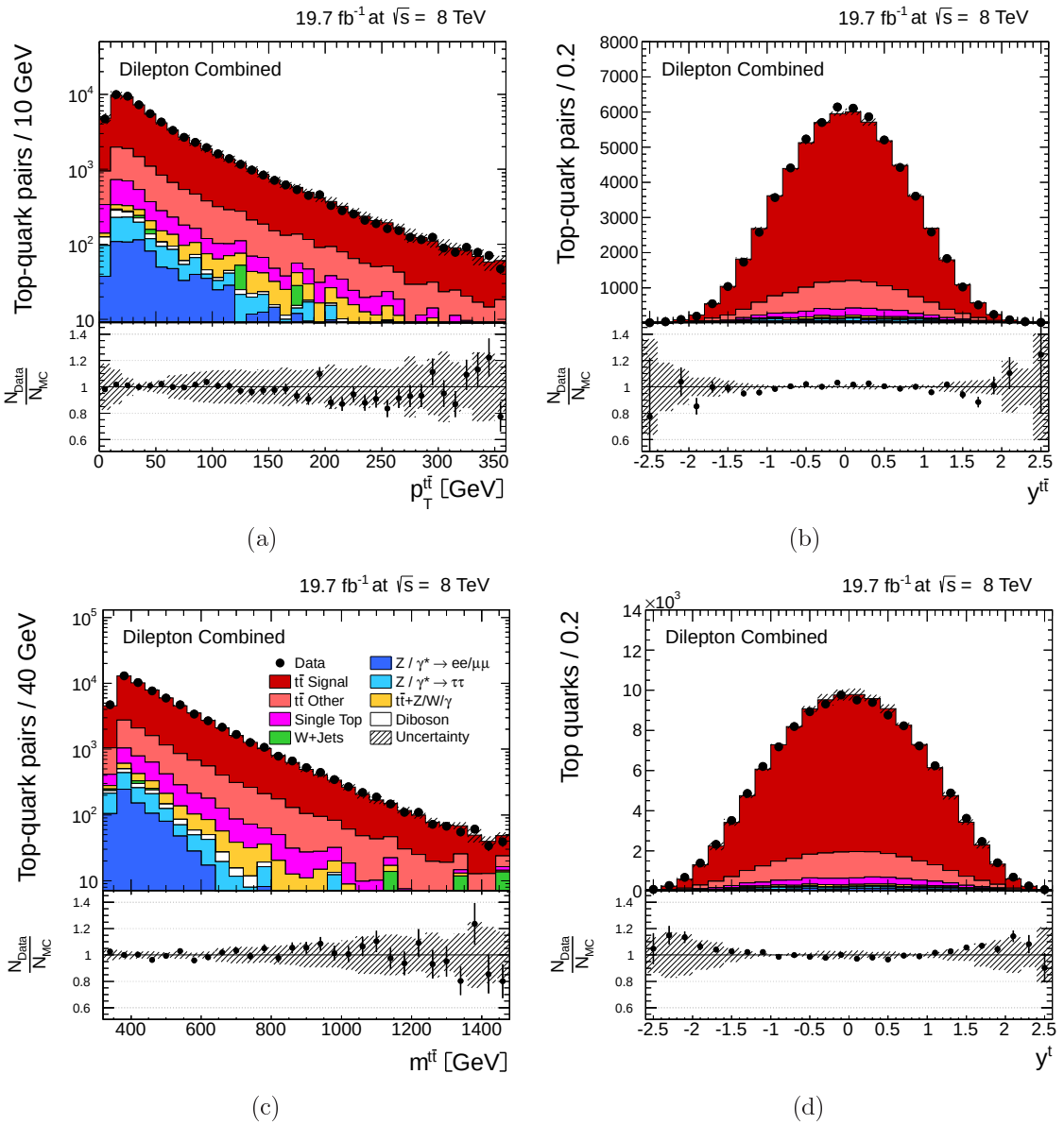


Figure 5.18: Distributions of the transverse momentum $p_T^{t\bar{t}}$ (a), rapidity $y^{t\bar{t}}$ (b) and invariant mass $m^{t\bar{t}}$ (c) of the top-quark-pair system, and top-quark rapidity y^t (d) distribution as obtained from the kinematic reconstruction algorithm. The error bars in the data points represent the statistical uncertainty and the hatched band corresponds to all shape uncertainties on the $t\bar{t}$ simulation events (c.f. Chapter 7). The bottom panels present the data-to-simulation ratio.

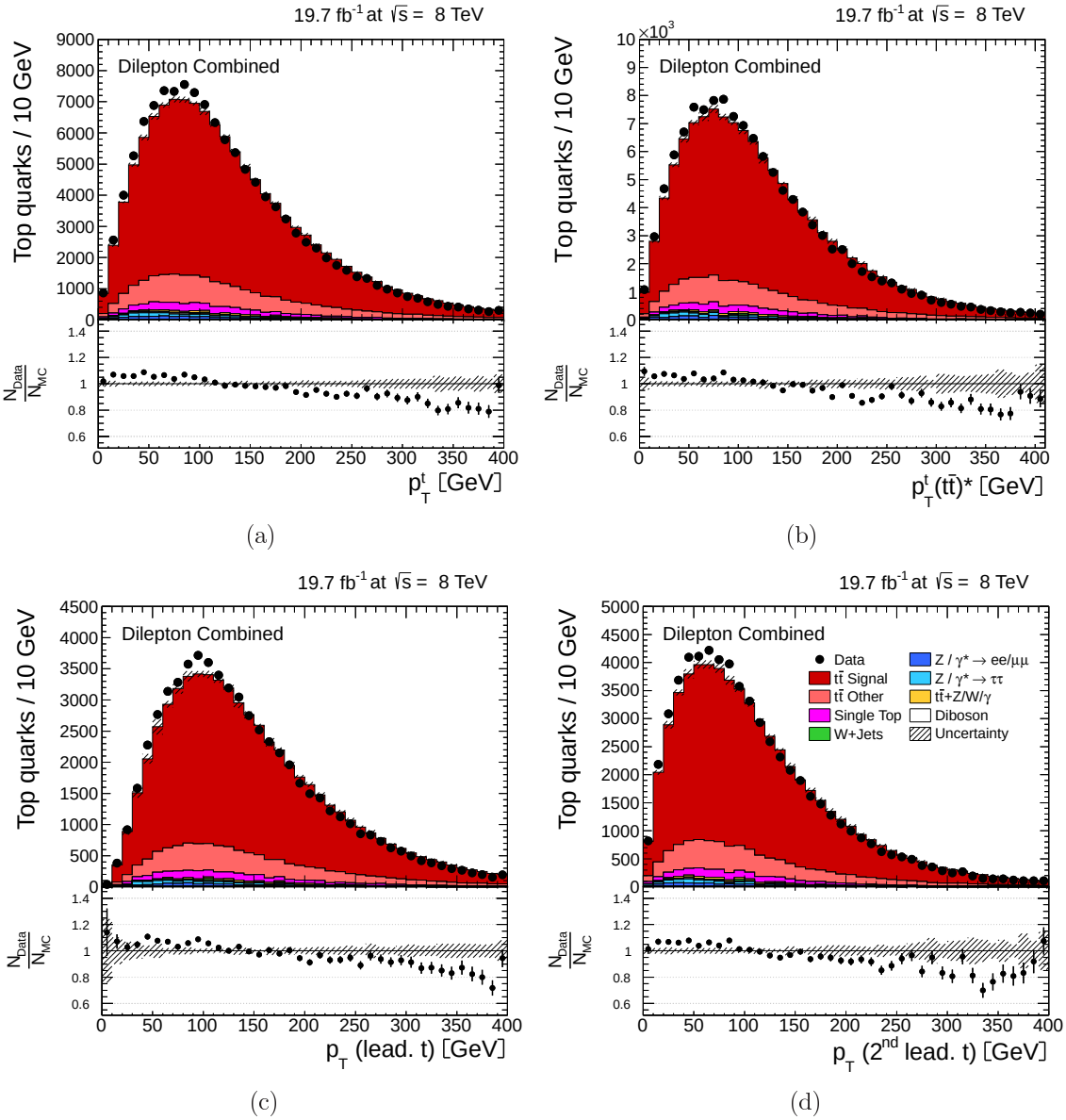


Figure 5.19: Distributions of the transverse momentum of the top quarks as obtained from the kinematic reconstruction algorithm. The distributions measured in the laboratory and $t\bar{t}$ rest frame are shown in (a) and (b), respectively. The transverse momentum distributions of the p_T ordered top quarks are shown in (c) for the leading top quark and in (d) for the trailing top quark. The error bars in the data points represent the statistical uncertainty and the hatched band corresponds to all shape uncertainties on the $t\bar{t}$ simulation events (c.f. Chapter 7). The bottom panels present the data-to-simulation ratio.

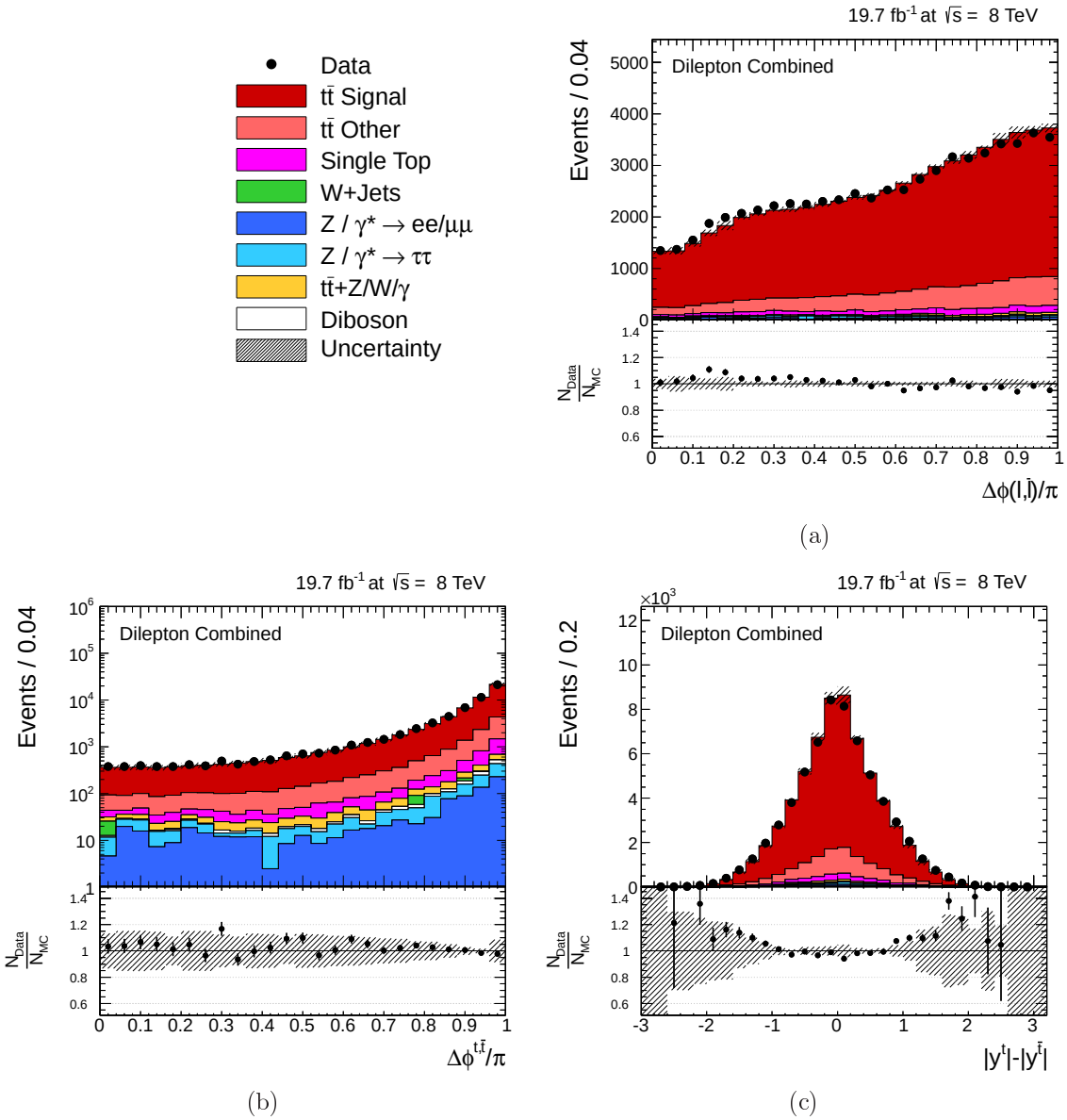


Figure 5.20: Distributions of the azimuthal difference between the lepton and antilepton $\Delta\phi^{l\bar{l}}$ (a) and between the top quark and antiquark $\Delta\phi\,(t,\bar{t})$ (b), and difference in the absolute value of the rapidity of the top quark and antiquark $|y^t| - |y^{\bar{t}}|$ (c). Results are presented after the full event selection. The error bars in the data points represent the statistical uncertainty and the hatched band corresponds to all shape uncertainties on the $t\bar{t}$ simulation events (c.f. Chapter 7). The bottom panels present the data-to-simulation ratio.

Chapter 6

Cross Sections: Definition, Binning and Migrations

The event sample selected by the criteria detailed in Chapter 5 is used to measure the $t\bar{t}$ production cross sections differentially as a function of kinematic distributions of the top quark, the top-quark pair and the decay products, namely the leptons and the b-jets. The cross section in every bin i of an observable X is extracted from the number of measured signal events in data x, corrected for detector efficiencies, acceptances, and migrations and divided by bin-width ΔX, branching-ratio BR and luminosity L. The differential cross section is normalised by dividing it by the total cross section σ, measured in the same kinematic range, such that only a shape measurement is performed. Mathematically it can be written as

$$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}X^i} = \frac{1}{\sigma}\frac{x^i}{BR \cdot L \cdot \Delta X^i} \tag{6.1}$$

The cross section results are measured in the visible phase space defined by the detector acceptance or in the extrapolated full phase space. Phase space definition are detailed in Section 6.1. The correction for background contamination is explained in Section 6.2 followed by the description of the choice of binning (Section 6.3). Bin-to-bin migrations caused by limited detector resolution are corrected using an unfolding technique presented in Section 6.4. The differential and inclusive $t\bar{t}$ production cross sections are discussed in Section 6.5.

6.1 Phase Space Definition

The normalised differential cross section is measured as a function of several kinematic variables of different particles:

- Variables: transverse momentum (p_{T}), pseudorapidity (η), rapidity (y), invariant mass (m), difference in azimuthal angle ($\Delta\phi$) and difference in rapidity.
- Particles and systems of particle-pairs: lepton (l), lepton pair ($l\bar{l}$), b-jets (b), b-jet pair ($b\bar{b}$), top quark (t) and top-quark pair ($t\bar{t}$).

Lepton and jets originating from b-quark decays present experimentally detectable signatures, while top-quark and top-quark-pair kinematic properties can only be inferred from their decay products. Additionally, higher-order theory calculations are provided in unconstraint kinematic phase spaces. Two separate kinematic regions are used to measure the differential cross section results: the visible and the full phase space.

Visible Phase Space

Differential cross sections as a function of lepton and b-jet observables are measured in the *visible* phase space defined as the kinematic region with detector coverage. This selection avoids uncertainties or model dependencies due to extrapolation effects to kinematic regions that are experimentally inaccessible.

Particle level leptons are defined as the electrons or muons produced from a W-boson decay which comes from the decay of a top quark or antiquark.

Particle level jets are constructed by clustering all stable particles after the hadronisation simulation using the anti-k$_t$ algorithm with a $\Delta R = 0.5$ cone size. The list of particles which can be clustered includes the so-called *ghost* B-hadrons. They correspond to the last B-hadrons, before decay, in the simulation-evolution chain with their 4-momenta scaled by a very small factor, typically 10^{-20}. The momentum scaling allows to identify which jet contains the B-hadron within its constituents without modifying the jet kinematic properties.

In order to associate a ghost B-hadron to the b-quark produced in the top-quark decay, the simulation-true information is exploited. The particle (*mother particle*) from which the B-Hadron is produced is searched. The process continues iteratively by evaluating the mother particle type of the B-hadron-mother until a b-quark from a top-quark decay or a parton from the hard scattering is found. This method determines uniquely the original parton with an efficiency higher than 99% [99].

The b-jets are those two jets containing within its constituents the ghost B-hadrons associated to the b-quarks from the top-quark decay.

The visible phase space is defined by the presence of two leptons and two b-jets in the kinematic region indicated in Table 6.1.

	p_{T}	$\lvert\eta\rvert$
Lepton	$\geq 20\,\mathrm{GeV}$	≤ 2.4
b-Jets	$\geq 30\,\mathrm{GeV}$	≤ 2.4

Table 6.1: Kinematic requirements on the lepton and b-jets defining the visible phase space.

Full Phase Space

The kinematic region without any kinematic constraints is referred to as *full* phase space. This definition includes areas where no detector coverage exist and no experimental detection of particles is possible. A measurement in the full phase space is only possible with an extrapolation to the unknown region based on a particular theory model. In this analysis the extrapolation is based on MadGraph+Pythia top-quark-pair event simulation.

The cross sections measured as a function of $t\bar{t}$ and top-quark kinematic variables are presented in the full phase space. This allows for comparison not only with Monte Carlo simulations but also with the most precise QCD theory calculations of the $t\bar{t}$ production [3, 21,100,101]. These predictions use threshold resummation techniques and include corrections due to soft-gluon emission at approximate next-to-next-to-leading order (approx. NNLO) accuracy. Instead, Monte Carlo models describing the $t\bar{t}$ production include QCD expansion terms up to next-to-leading (NLO) order.

The differential cross section results are extrapolated to the parton level top-quark which corresponds to the last stable top quark after radiation and before decay in the Monte Carlo simulation process. The level of constrain imposed by the phase space definition is quantified

by the *acceptance* A. It is defined as the ratio of events generated in the full phase space $\mathrm{N}^{gen.}$(full PS) which also lies in the visible phase space $\mathrm{N}^{gen.}$(visible PS):

$$A = \frac{\mathrm{N}^{gen.}(\text{visible PS})}{\mathrm{N}^{gen.}(\text{full PS})} \tag{6.2}$$

The performance of the detector can be measured as a function of the detector efficiency $\varepsilon_i^{\text{detector}}$, which is defined as the ratio of the number of reconstructed events over the number of generated events in the visible phase space. The $\varepsilon_i^{\text{detector}}$ together with the acceptance A defines the full-analysis efficiency $\varepsilon_i^{\text{analysis}}$.

$$\varepsilon_i^{\text{analysis}} = A_i \cdot \varepsilon_i^{\text{detector}} = \underbrace{\frac{\mathrm{N}_i^{\text{gen.}}(\text{visible PS})}{\mathrm{N}_i^{\text{gen.}}(\text{full PS})}}_{A_i} \cdot \underbrace{\frac{\mathrm{N}_i^{\text{rec.}}(\text{visible PS})}{\mathrm{N}_i^{\text{gen.}}(\text{visible PS})}}_{\varepsilon_i} \tag{6.3}$$

In the case of distributions measured in the visible phase space, no acceptance correction is applied and therefore the analysis efficiency reduces to the detector efficiency. The distributions of the analysis efficiency and acceptance are shown for illustration in Figure 6.1 as a function of top-quark transverse momentum and $\mathrm{t\bar{t}}$ invariant mass. Additional distributions for all measured quantities are displayed in Appendix B.1. For the presented binning it can be observed that more than 60% of the events lie in the phase space with detector coverage, and only a small extrapolation to the full phase space is performed.

6.2 Background Subtraction

The final sample of selected data events N_{Data} contains a large fraction of $\mathrm{t\bar{t}}$ dilepton events, approximately 80% of the selected events. The remaining fraction correspond to contamination due to background processes mainly produced from $\mathrm{t\bar{t}}$ decays other than the dilepton decay channel and single-top production (see Table 5.4).

The number of signal events in the selected data sample is obtained by subtracting the non-$\mathrm{t\bar{t}}$ background events $\mathrm{N}^{\text{MC}}_{\text{non-t}\bar{\text{t}}\text{ BG.}}$ and correcting by the fraction of $\mathrm{t\bar{t}}$ decays other than the dilepton $N^{\text{MC}}_{\text{t}\bar{\text{t}}\text{ other}}$, i.e.: dilepton decays via intermediate τ leptons, semi-leptonic and fully-hadronic $\mathrm{t\bar{t}}$ decays. The use of the signal fraction factor, $f_{sig.}$, in the background correction process removes the dependency of the expected $\mathrm{t\bar{t}}$ events on the normalisation cross section value. Expression (6.4) summarises the number of measured $\mathrm{t\bar{t}}$ signal events in the selected data sample.

$$\begin{aligned} \mathrm{N}^{\text{sel.}}_{\text{t}\bar{\text{t}}\text{ signal}} &= f_{sig.}\left(\mathrm{N}_{\text{Data}} - \mathrm{N}_{\text{non-t}\bar{\text{t}}\text{ BG.}}\right) \\ &= \underbrace{\frac{N^{\text{MC}}_{\text{t}\bar{\text{t}}\text{ signal}}}{N^{\text{MC}}_{\text{t}\bar{\text{t}}\text{ signal}} + N^{\text{MC}}_{\text{t}\bar{\text{t}}\text{ other}}}}_{f_{sig.}} \underbrace{\left(\mathrm{N}_{\text{Data}} - \mathrm{N}^{\text{MC}}_{\text{non-t}\bar{\text{t}}\text{ BG.}}\right)}_{\text{all t}\bar{\text{t}}\text{ events in data}} \end{aligned} \tag{6.4}$$

6.3 Migrations and Selection of Binning

Limitations in the detector resolution and the requirement of finite bin boundaries causes *migrations* of selected events, i.e.: an event generated in a bin i of a given distribution is reconstructed in a different bin j. Such effects have to be removed in order to produce an unbiased and detector independent measurement.

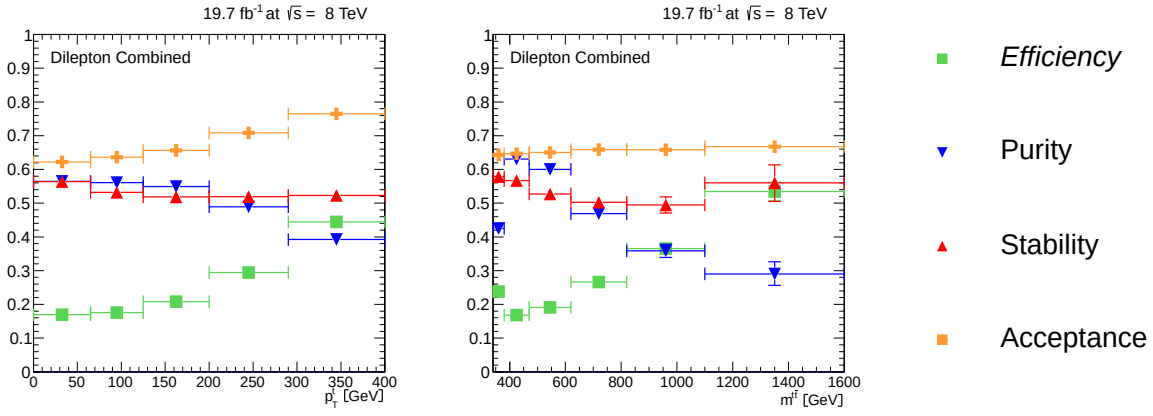
The migrations are quantified by the *purity* and *stability* variables derived from $t\bar{t}$ simulation events. The purity p_i, defined as the fraction of reconstructed ($N_{\text{rec.}}$) events which are also generated in the same bin ($N_{\text{gen. and rec.}}$), estimates the migrations out of a certain bin. The stability s_i instead quantifies migrations into a bin by counting the fraction of events generated ($N_{\text{gen.}}$) in a bin i, that are also correctly reconstructed in the same bin.

$$p_i = \frac{\mathrm{N}_i^{\text{gen. and rec.}}}{\mathrm{N}_i^{\text{rec.}}} \tag{6.5}$$

$$s_i = \frac{\mathrm{N}_i^{\text{gen. and rec.}}}{\mathrm{N}_i^{\text{gen.}}} \tag{6.6}$$

It is desired a measurement in a small bin-size, in the ideal case infinitesimal bin-size, such that the result is sensitive to possible shape differences. With the reduction of the bin size the amount of migrations and statistical uncertainty increase, rendering the cross section less precise. The choice of binning must balance these limitations. The number and size of the bins of the measurement are chosen such that the stability is approximately constant and larger than 50% along all bins of the measured distribution. Thus at least 50% of the events are reconstructed in the same bin where they were originated.

Figure 6.1 shows the purity and stability distributions as a function of the top-quark transverse momentum and $t\bar{t}$ invariant mass as an illustrative example. Corresponding distributions for all other measured quantities can be found in Appendix B.1.



(a) (b)

Figure 6.1: Distributions of the stability, purity, acceptance and analysis efficiency as a function of the reconstructed top-quark transverse momentum (a) and top-quark-pair invariant mass (b). The vertical error bars indicate the statistical uncertainty of the distribution, and the horizontal bars the size of the bin.

6.4 Unfolding of Signal Events

The analysis aims to measure the true spectrum of $t\bar{t}$ dilepton events, corrected for detector effects, from the experimentally measured number of events. The detector introduces bin-to-bin migration effects which are limited by an appropriate bin-size choice. Information on the smearing effects caused by the detector are studied with signal event simulation and

codified into a detector response matrix $\mathcal{A}$. This matrix provides a relation between the true distribution x and the smeared measured distribution $\mathrm{N}^{\mathrm{sel.}}_{\mathrm{t\bar{t}\ signal}}$. This relation is described by Equation (6.7).

$$\mathrm{N}^{\mathrm{sel.}}_{\mathrm{t\bar{t}\ signal}} = \mathcal{A} \cdot x \tag{6.7}$$

Obtaining the true distribution x from a direct inversion of Equation (6.7) is an ill-posed problem known as *unfolding*. In addition, the direct inversion introduces non-physical anti-correlations between neighbouring bins [102]. This leads to rapidly oscillating variations between consecutive bins represented by the opposite sign in consecutive elements of the inverse response matrix $\mathcal{A}^{-1}$. An illustrative example is shown in Equation (6.8).

$$\mathcal{A} = \begin{pmatrix} 0.7 & 0.4 \\ 0.3 & 0.6 \end{pmatrix} \rightarrow \mathcal{A}^{-1} = \begin{pmatrix} 2 & -4/3 \\ -1 & 7/3 \end{pmatrix} \tag{6.8}$$

These limitations are overcome by introducing a *regularisation* condition $\mathcal{C}$ based on an a priori knowledge of the solution. It requires a continuous and smooth transition between neighbouring bins of the measured distribution.

The original unfolding problem of finding the solution x can be expressed as χ^2-minimisation problem as a function of the floating parameter x. The Equation (6.9) summarises the unfolding problem and includes the regularisation condition $\mathcal{C}$ modulated by the regularisation strength parameter τ.

$$\chi^2(x) = \underbrace{(\mathrm{N}^{\mathrm{sel.}}_{\mathrm{t\bar{t}\ signal}} - \mathcal{A} \cdot x)^{\top} \cdot COV^{-1}_{\mathrm{N}^{\mathrm{sel.}}_{\mathrm{t\bar{t}\ signal}}} \cdot (\mathrm{N}^{\mathrm{sel.}}_{\mathrm{t\bar{t}\ signal}} - \mathcal{A} \cdot x)}_{\chi^2_{\mathrm{Unfolding}}} + \underbrace{\tau\,(\mathcal{C} \cdot x)^{\top}\,(\mathcal{C} \cdot x)}_{\chi^2_{\mathrm{Reg.}}} = \mathrm{min.} \tag{6.9}$$

The $COV^{-1}_{\mathrm{N}^{\mathrm{sel.}}_{\mathrm{t\bar{t}\ signal}}}$ represents the covariance matrix of the measured signal events, it is a diagonal matrix containing the statistical uncertainty of measured signal events.

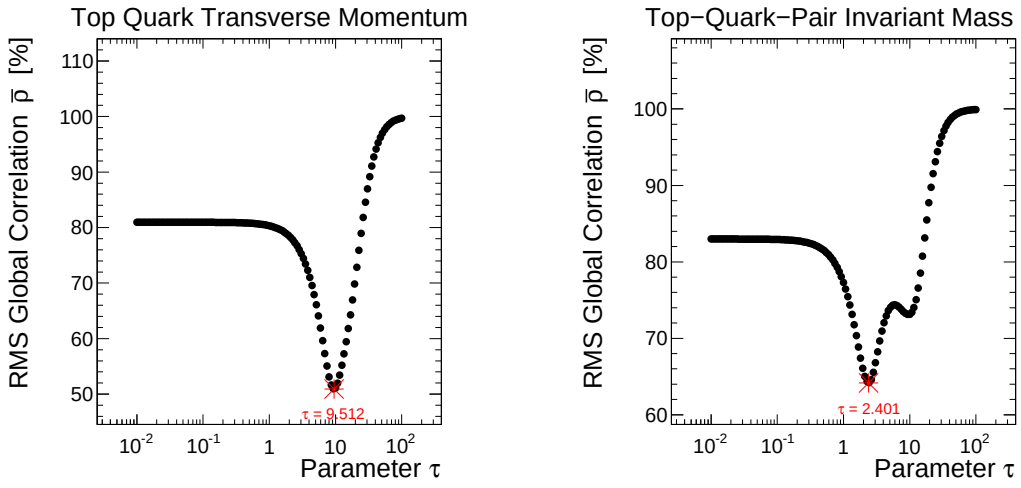
The Equation (6.9) is solved based on the Single Value Decomposition (SVD) [103] matrix diagonalisation method as implemented in the TSVDUnfold [104] package. The regularisation condition $\mathcal{C}$ allows only correlations between direct neighbours, and is technically implemented as the second derivative of a discrete distribution.

$$\mathcal{C} = \begin{pmatrix} -1 & 1 & 0 & & \\ 1 & -2 & 1 & 0 & \\ & \ddots & \ddots & \ddots & \\ & 0 & 1 & -2 & 1 \\ & & 0 & 1 & -1 \end{pmatrix} \tag{6.10}$$

The appropriate determination of the regularisation strength τ is important since:

- Large τ-values, $\tau \rightarrow \infty$, cause over-smoothed solutions: the regularisation condition is dominant over the unfolding problem $\chi^2_{\mathrm{Unfolding}}$ and therefore correlations between largely-separated bins are artificially removed.
- Under-smoothed solutions are obtained when $\tau \rightarrow 0$ and therefore the oscillating results between neighbouring bins are weakly suppressed.

The determination of the regularisation strength differs from the original SVD [103] problem and it is based on the minimisation of the average global correlation $\bar{\rho}$ of the result [105]. Figure 6.2 shows the $\bar{\rho}$ value as a function of the regularisation strength for the top-quark p_{T} and $\mathrm{t\bar{t}}$ invariant mass results. Additional distributions for all other measured quantities are presented in Appendix B.2.



(a) (b)

Figure 6.2: Average global correlation $\bar{\rho}$ distribution as a function of the regularisation strength τ for the cross section result obtained as a function the top quark p_{T} (a) and $\mathrm{t\bar{t}}$ invariant mass (b).

6.4.1 Consistency Tests of the Unfolding

The unfolding method relies on a good description of the detector effects by the simulation and on some a priori assumptions on the smoothness of the result between two consecutive bins. The stability of the result against variations in the unfolding set up is tested in multiple modes:

- Variation of the regularisation strength parameter.
- Unfolding of pseudo-data obtained from $\mathrm{t\bar{t}}$ signal simulation with and without shape variations on the original distribution.

A detailed discussion on these studies is provided in the following.

Regularization Strength

As described previously a regularisation term is added to the original unfolding problem to remove rapidly and largely oscillating solutions imposing a continuous and a smooth solution between consecutive bins of the measurement. The τ parameter indicates the relevance or the *strength* of the regularisation term in comparison to the unfolding term. It is selected such that the average global correlation $\bar{\rho}$ is minimal.

The possible biases due the τ-selection are evaluated by unfolding the selected data events with varied regularisation-strength-parameter values with respect to the default value $\tau_{\mathrm{opt.}}$: $\tau = 0$, $10^{-2} \cdot \tau_{\mathrm{opt.}}$, $10^{-1} \cdot \tau_{\mathrm{opt.}}$, $10 \cdot \tau_{\mathrm{opt.}}$, $50 \cdot \tau_{\mathrm{opt.}}$ and $10^{2} \cdot \tau_{\mathrm{opt.}}$. Results are shown in Appendix B.3. The data markers are horizontally displaced from the bin centre for graphical purposes only. In the same figures the Monte Carlo model used to derive the migration matrix which describe the detector effects is shown, i.e. the original and unmodified MadGraph+Pythia prediction.

The results are found to be largely insensitive to the regularisation parameter with no significant difference in the measurement when varying the τ-value in the range $[0, 10{\cdot}\tau_{\text{opt.}}]$. As expected, the use of very *strong* regularisation values ($50 \cdot \tau_{\text{opt.}}$ and $100 \cdot \tau_{\text{opt.}}$) biases the result towards the simulation model used to derive the migration matrix [103, 105, 106]. Increasing the regularisation term value $\tau \to \infty$ implies that the curvature term on the unfolding equation allows only migrations between neighbouring bins ignoring correlations between bins largely separated. This situation is equivalent to a bin-wise efficiency correction, also called bin-by-bin unfolding which is known to bias the result towards the simulation model used to derive the detector efficiency corrections [105].

Pseudo-data

The recorded data sample from collision events only provides insights into the post-detector stage since previous stages of the collision are experimentally inaccessible. In contrast, Monte Carlo simulations provide full access to kinematic properties of an event before, during and after the detector simulation and reconstruction process.

In order to test the unfolding framework and the assumptions made, the analysis is repeated by replacing the real collision data with a sample of simulated $\mathrm{t\bar{t}}$ dilepton events obtained from the MadGraph+Pythia simulation. The pseudo-data sample is scaled to the integrated luminosity of the real data, $L = 19.7\,\text{fb}^{-1}$. The analysis follows in an identical manner by subtracting the background events, correcting for the signal fraction and unfolding the signal events with the same migration matrix and regularisation parameters as for real data.

The results of the pseudo-data are shown in Appendix B.3. The figures show also the original pseudo-data distribution (*Simu. Reweighted*) and the shape of the theory prediction used to model the detector effects (MadGraph+Pythia). For reference a higher-order QCD theory calculation is also shown, wherever available.

The results derived from pseudo-data are in good agreement with the original distribution. No bias due to the unfolding procedure is observed.

In order to evaluate the impact of a possible biases due to the shape of the original pseudo-data a distortion is introduced on the p_{T}^{t} or y^{t} distribution of the pseudo-data events. The distortion is linear in p_{T} and y around the values $p_{\text{T}}^{\text{t}} = 100\,\text{GeV}$ and $y^{\text{t}} = 1$ and is modulated by a *slope* factor s. A weight w, derived as a product of top quark and antiquark variations, is applied to the pseudo-data distribution. The mathematical expressions summarising the shape-variations are given in Equations (6.11) and (6.12) for p_{T}^{t} and y^{t} dependent changes, respectively.

$$w = \max\left(0.1, \left(1 + \left(p_{\text{T}}^{t} - 100\right) \cdot s\right) \cdot \left(1 + \left(p_{\text{T}}^{\bar{t}} - 100\right) \cdot s\right)\right) \tag{6.11}$$

$$w = \max\left(0.1, \left(1 + \left|y^{t} - 1\right| \cdot s\right) \cdot \left(1 + \left|y^{\bar{t}} - 1\right| \cdot s\right)\right) \tag{6.12}$$

The parameter s regulates the size of the variation introduced to the original distribution and is limited to $s = \pm 0.002, \pm 0.004$ and $s = \pm 0.08, \pm 0.16$ for p_{T}^{t} and y^{t} dependent variations, respectively. Larger variations are not evaluated since they would not represent realistic physical situations and would be significantly larger than the uncertainties of the measurement.

The modified pseudo-data events are unfolded using the standard unfolding set up including the migration matrix obtained from the original and unmodified MadGraph+Pythia signal simulation. The results are found to be described by the modified pseudo-data dis-

tribution (*Simu. Reweighted*), and any difference is within the statistical uncertainty of the pseudo-data (see Appendix B.3). No bias is found.

6.5 Cross Section Definition

6.5.1 Differential Cross Section

Once the reconstructed signal events are corrected for migration effects and the binning is defined, all relevant values are in place to calculate the normalised differential cross section $(1/\sigma)(\mathrm{d}\sigma/\mathrm{d}X)$ as a function of the variable X. The calculation proceeds by dividing, in each bin of the measurement, the number of $\mathrm{t\bar{t}}$ events in data corrected for detector efficiency and migration events x^i by the bin-width ΔX^i. The shape measurement is finally obtained by normalising to unit area in the range where the result is presented. The normalisation factor is obtained by summing up the cross section values in each bin.

$$\frac{1}{\sigma}\left(\frac{\mathrm{d}\sigma}{\mathrm{d}X}\right)^i = \frac{1}{\sum_j \left(\Delta X^j \frac{x^j}{\Delta X^j \cdot \cancel{BR} \cdot \cancel{L}}\right)} \frac{x^i}{\Delta X^i \cdot \cancel{BR} \cdot \cancel{L}} \tag{6.13}$$

Constant terms along all bins of the measurement, such as the branching ratio and the luminosity, cancel due to the normalisation of the result. This leads to a reduction of the systematic uncertainties, as discussed in Chapter 7. The differential cross section results are presented and discussed in Chapter 8.

6.5.2 Total Production Cross Section

The total $\mathrm{t\bar{t}}$ production cross section is measured counting the number of signal events contained on the selected dataset after the complete event selection described in Chapter 5, including the requirement of a valid solution of the top-quark kinematic reconstruction algorithm. The number of $\mathrm{t\bar{t}}$ dilepton events on the selected dataset is obtained by subtracting the background contamination and correcting the $\mathrm{t\bar{t}}$ signal fraction (see Section 6.2). Finally the signal events are corrected for detector efficiency effects $\varepsilon^{\mathrm{detector}}$, visible phase space acceptance A, total integrated luminosity L and branching ratio of decay channel BR. Mathematically it is expressed by the Equation (6.14). The total $\mathrm{t\bar{t}}$ cross section result is presented in Section 8.6.

$$\sigma = \frac{f_{\mathrm{sig.}} \cdot (N_{\mathrm{Data}} - N_{\mathrm{non\ t\bar{t}\ BG.}})}{\varepsilon^{\mathrm{detector}} \cdot A \cdot BR \cdot L} \tag{6.14}$$

6.5.3 Combination of Results

The differential cross sections are measured in the ee, eμ and $\mu\mu$ channels from three statistically independent data samples. The results are combined by minimising a negative-log-likelihood function under the assumption of three statistically independent Gaussian-distributed results. Under this hypothesis the problem transforms into the so-called *weighted average* [107]. The normalised differential cross section and statistical uncertainty $\delta_{\mathrm{Combined}}$ for an arbitrary bin of the combined channel are expressed as

$$\left(\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}X}\right)_{\mathrm{Combined}} = \frac{1}{\delta^2_{\mathrm{Combined}}} \sum_i \left(\left(\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}X}\right)_i / \delta_i^2\right) \tag{6.15}$$

$$\delta_{\mathrm{Combined}} = \frac{1}{\sqrt{\sum_i \frac{1}{\delta_i^2}}} \tag{6.16}$$

where the index i runs on the three statistically independent channels: ee, eμ and $\mu\mu$.

6.6 Covariance Matrix and Statistical Uncertainty

In the unfolding procedure to correct for detector effects the special relevance of the correlations between the bins was introduced. The extension of the statistical uncertainty of a particular given bin, which dilutes any correlation information, is a covariance matrix generally defined as:

$$COV_{ij}(y) = \langle (y_i - \langle y_i \rangle) \cdot (y_j - \langle y_j \rangle) \rangle \tag{6.17}$$

where $\langle y_i \rangle$ represents the expectation value of the differential cross section result y_i in the i-th bin.

In this analysis the covariance matrix is obtained via error propagation. The measured $\mathrm{t\bar{t}}$ signal yield N is overlaid with random noise proportional to the statistical uncertainty, $\sqrt{N}$. For each varied configuration the unfolding procedure is repeated and a new result is obtained: y'. A new covariance matrix is calculated from the difference with respect to the original result: $COV(y')_{ij} = (y'_i - y_i) \cdot (y'_j - y_j)$. The average result of 1000 pseudo-experiments represents the final covariance matrix of the cross section result.

The statistical uncertainty of the results is obtained from the diagonal elements of the results covariance matrix, $COV(y)_{ii}$.

The maximum correlation of a bin k with all other bins from a particular measurement can be retrieved from the covariance matrix [108] using the expression (6.18). Averaging this expression the average global correlation $\bar{\rho}$ of the bins of the measurement is obtained, see Equation (6.19).

$$\rho_i = \sqrt{1 - \frac{1}{COV(y)_{ii} \cdot COV^{-1}(y)_{ii}}} \tag{6.18}$$

$$\bar{\rho} = \sqrt{\frac{1}{\mathrm{N_{bins}}} \sum_{i=1}^{\mathrm{N_{bins}}} \rho_i^2} \tag{6.19}$$

Chapter 7

Systematic Uncertainties

Systematic uncertainties arising from limited understanding of the detector efficiencies and resolutions, as well as theory modelling uncertainties limit the precision of the measurement. A detailed study and evaluation of the uncertainties is at the core of this analysis.

Detector efficiencies are known with a certain finite accuracy and any deviation from the true value will cause a shift of the number of selected signal and background events and as a consequence of the differential cross sections.

Similarly, the simulation of signal events includes some parameters in the Monte Carlo model such as top-quark mass or hadronisation model. These are derived from experimental results, with a certain accuracy, or they are based on some a priori assumptions. Different parameters in the simulation cause different acceptances and as a consequence different cross-section results.

In this chapter the method to evaluate the systematic uncertainty is described first, in Section 7.1. It is accompanied by a discussion of the individual sources of the uncertainties. The detector modelling sources are explained in Section 7.2 followed by the description of the uncertainties on the theory modelling of the $t\bar{t}$ events in Section 7.3. In Section 7.4 the typical uncertainty values are given separately for results measured in the visible phase space and in the extrapolated full phase space. A method to estimate the correlation between the systematic uncertainties is described in Section 7.5.

7.1 Systematic Uncertainty Determination Procedure

The effect of a particular source of systematic uncertainty on the result is assessed by repeating the full analysis procedure and re-obtaining a new measurement. The difference of the varied and the nominal measurement is then quoted as the systematic uncertainty.

Detector modelling effects are inserted in the analysis as efficiency correction factors on the selected events from simulation. The uncertainties are obtained by varying these factors within their uncertainties and re-evaluating the cross section result.

The evaluation of the $t\bar{t}$ theory model is addressed by replacing the nominal simulation sample by an alternative sample generated with a variation of a certain parameter. The cross section results are derived using the alternative sample to estimate the signal fraction and the visible phase space acceptance corrections. A summary table with the simulation configurations is given in Table 7.1 while further details on the Monte Carlo simulation are given in Chapter 4.

In general, a given source of uncertainty has an *up* and a *down* variation from its nominal

Uncertainty	ME, PS and Hadron. Model	Comments
Nominal	MadGraph+Pythia	$m_t = 172.5\,\text{GeV}$ $Q_0^2 = m_t^2 + \sum_{\text{add.partons}} p_{\text{T}}^2$ ME-PS scale 20 GeV CTEQ6L1 PDF set
Q^2-scale up	MadGraph+Pythia	$Q^2 = 4 \cdot Q_0^2$
Q^2-scale down	MadGraph+Pythia	$Q^2 = 0.25 \cdot Q_0^2$
Matching-scale up	MadGraph+Pythia	ME-PS scale 40 GeV
Matching-scale down	MadGraph+Pythia	ME-PS scale 10 GeV
Top-quark mass up	MadGraph+Pythia	$m_t = 173.5\,\text{GeV}$
Top-quark mass down	MadGraph+Pythia	$m_t = 171.5\,\text{GeV}$
PDF set	MadGraph+Pythia	CT10 PDF set
Hard scattering model	Powheg+Pythia	–
Parton shower and Hadronisation model	MC@NLO+Herwig	–

Table 7.1: Summary of the different $\text{t}\bar{\text{t}}$ simulation configurations used for the estimation of the signal modelling uncertainties. The hard scattering (ME), parton shower (PS) and hadronisation model together with the different simulation parameters are given for each of the evaluated uncertainties.

value which consequently causes two different cross section results. In this analysis, the total uncertainty due to a particular source is obtained as the symmetric average of the up and down differences with respect to the nominal result:

$$\delta_s = \frac{1}{2}\left(|\text{result}_{up} - \text{result}_{\text{nominal}}| + |\text{result}_{down} - \text{result}_{\text{nominal}}|\right) . \tag{7.1}$$

On the other hand, the direct comparison between two models, e.g. MadGraph vs Powheg or Pythia vs Herwig, produces only a single difference: $\text{result}_{\text{Model 1}} - \text{result}_{\text{Model 2}}$. In this case a conservative approach is taken by symmetrically applying the difference as an up and down uncertainty.

All sources of uncertainty are assumed to be uncorrelated with each other and their contributions are summed in quadrature to obtain the total systematic uncertainty $\delta_{\text{tot.}}$

$$\delta_{\text{total}} = \sqrt{\sum_s \delta_s^2}\,, \tag{7.2}$$

where the sum runs on all systematic sources s considered.

The systematic uncertainties are calculated separately in each bin of every measured distribution. For illustration, the median[1] value of the individual uncertainty sources in all bins of every distribution is quoted as *typical* size of each systematic uncertainty. Typical uncertainties are separated for results measured in the visible phase space defined by the detector acceptance and in the extrapolated full phase space. Additional details on the phase space definition are given in Section 6.1.

[1]The median of an ordered set is defined as the numerical value dividing the set in two equally populated subsets.

Because the differential cross sections are normalised, only those sources of uncertainties that lead to changes in the shape of the result contribute to the total uncertainty. Normalisation uncertainty sources such as the branching-ratio or the data luminosity largely cancel in the measurement of the normalised differential cross section.

7.2 Experimental Uncertainties

The experimental efficiency differences in data and simulation are corrected by applying efficiency correction factors to the events in simulation. These correction factors, also referred to as data-to-simulation factor or scale factors (SF), are measured with finite precision and a variation within their uncertainty leads to a shift of the result of this measurement.

Trigger

The trigger data-to-simulation factors are measured as a function of the lepton pseudorapidity with a typical precision of 1%, as described in Section 5.1. A shift of the result by typically 0.1% is observed when varying the correction factors within their uncertainties.

Pileup

The distribution of the number of additional pp collisions in data is derived as a function of the measured luminosity and the pp cross section (see Section 5.2). Following the recommendation from [80], the pp cross section is varied by 5% in order to conservatively account for the luminosity and pp-cross-section uncertainties which amount to 2.6% and 2.5%, respectively. The pileup distribution variation produces typically a 0.1% change of the $t\bar{t}$ differential cross sections.

Lepton Identification and Isolation

The lepton identification and isolation efficiency differences are corrected based on a tag-and-probe technique (Section 5.3). The method has a systematic uncertainty of the order of 0.3%, but to cover possible differences between the Drell-Yan and $t\bar{t}$ event topologies, a conservative 1% uncertainty is added [109]. The typical shift on the result is in general below 0.1%.

Jet Energy Scale

The calibration of the energy of the jets, described in Section 5.5, is based on a factorised process and is implemented as a set of double-differential correction factors [86]. The maximum uncertainty in the kinematic region of interest for the analysis is below 5%.

The differences in the $t\bar{t}$ differential cross section obtained from the propagation of the jet-energy scale uncertainties are typically 0.4% and 1.0% for distributions measured in the visible and full phase space, respectively.

Jet Energy Resolution

The jet energy resolution (JER) in simulated events is corrected using $|\eta|$-dependent factors measured with a precision of about 6% (see Section 5.5).

The uncertainty on the $t\bar{t}$ differential cross sections due to JER is typically 0.1%.

Normalisation of Simulation Events

The measured Drell-Yan event yield is normalised using a method based on data (Section 5.6), while the normalisation of other background contributions is based on the most precise theory calculations (Section 4.5).

The uncertainty of the background normalisation, including Drell-Yan process, is conservatively assumed to be 30%, accounting for theory uncertainties and shape differences. The typical systematic uncertainties of the differential cross section are 0.2% due to Drell-Yan process normalisation and 0.3% due to the normalisation of other backgrounds.

Missing Transverse Energy

Any lepton or jet uncertainty is propagated to the $\vec{\not{E}}_{\mathrm{T}}$ distribution by recomputing $\vec{\not{E}}_{\mathrm{T}}$-observable after varying the kinematic properties of the particles in the event. Therefore no specific uncertainty associated to the missing transverse energy is evaluated, as recommended by the experts [88].

Uncertainty on the b-Tagging Techniques

The systematic uncertainty due to the b-jet identification is due to the simulation correction factors. These are derived as a function of the jet flavour and kinematic properties [93] with uncertainties up to 4% due to modelling of b-fragmentation and gluon splitting.

In this analysis, the b-tagging uncertainty is assessed by varying the correction factor as a function of $|\eta|$ and p_{T}. Two regions, separated by the median point, are varied in opposite directions such that the shape of the distribution is maximally modified. The median points correspond to $|\eta| = 0.75$ for all jet flavours. The median points of the p_{T} distribution are $p_{\mathrm{T}} = 65$ and 45 GeV for heavy- (b- and c-jets) and light-jets, respectively. The p_{T} and η variations can be considered uncorrelated and their contributions to the systematic uncertainty are added in quadrature. The uncertainties on heavy-flavour jets are considered fully correlated and therefore varied simultaneously, while light-flavoured jets are uncorrelated to the heavy-flavour components and are varied separately and summed in quadrature to the heavy-flavour component of the uncertainty.

The shape variations on the b-tagging correction factors produce an uncertainty of 0.1% (0.2%) on the cross sections measured in the visible (full) phase space.

Uncertainty on the Reconstruction Algorithm of Top Quark Kinematics

As detailed in Section 5.8, only events with a solution of the top-quark kinematic reconstruction are considered in the measurement. Events from data and simulation have very similar efficiencies and a small correction factor is applied to correct for the remaining differences. The shift on the results obtained by varying these correction factors within their uncertainties is found to be negligible (below 0.1%).

Branching Ratio

The uncertainties on the $\mathrm{t\bar{t}}$ dilepton final state are propagated from the individual branching ratios of W bosons decaying into leptons [12]. This uncertainty source largely cancels in the normalised differential cross section measurement.

Luminosity

The uncertainty on the integrated luminosity is obtained based on a pixel cluster counting method and is found to be 2.6% [57]. The effect of the luminosity largely cancels in the measurement of the differential cross section while it has a substantial effect on the total inclusive $t\bar{t}$ cross section result.

7.3 Signal Model Uncertainties

The $t\bar{t}$ signal model uncertainties are obtained by replacing the nominal simulation with alternative samples generated with varied simulation parameters. The nominal $t\bar{t}$ signal sample is produced interfacing the MadGraph matrix element generator and Pythia for showering and hadronisation effects, a top-quark mass of 172.5 GeV and the CTEQ6L1 PDF set. Additional details on the simulation of events and parameters used in the signal simulation are given in Chapter 4.

Parton Distribution Function

The uncertainty due to the PDF modelling is assessed, following the recommendations from [19], by reweighting the simulated $t\bar{t}$ events according to the 53 error eigenvalues of the CT10 PDF set evaluated at 90% C.L. The envelope with the maximum difference between the results evaluated with the central and the uncertainty eigenvalues is considered as the PDF uncertainty on the result. Typically, the uncertainty due to the PDF is in the range of 0.3–0.5% depending on the phase space in which the results are measured.

Hard Scattering Model

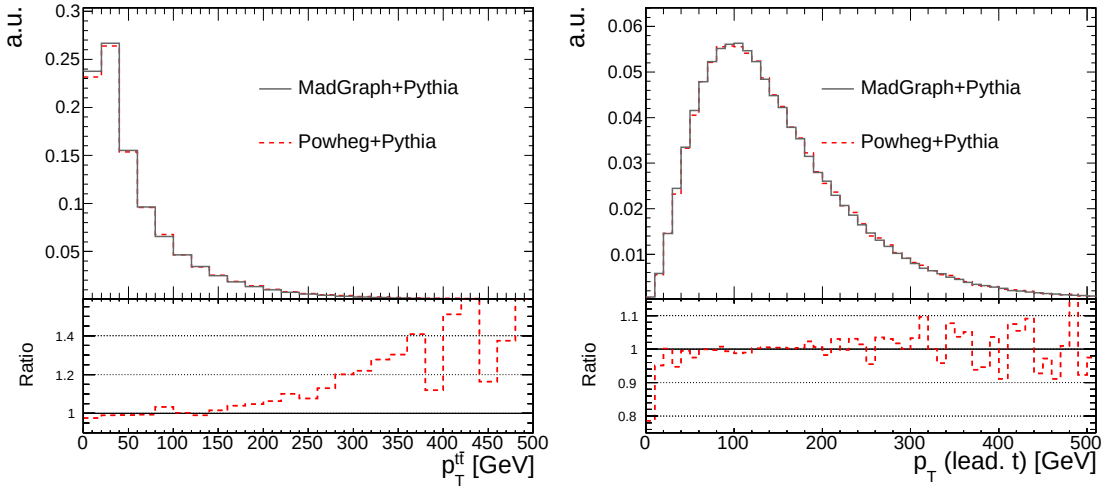
The uncertainty due to the hard scattering model is obtained as the difference of the results measured with MadGraph and Powheg matrix element generators. The former generates $t\bar{t}$ events at LO with up to 9 final state particles, while the latter generates $t\bar{t}$ events with NLO accuracy in perturbative QCD. Both models, MadGraph and Powheg, are interfaced to Pythia for showering and hadronisation effects. It is assumed that Pythia performs identically in both samples. Figure 7.1 shows the generator-level distribution of the transverse momentum of the $t\bar{t}$ system and of the leading top quark generated with MadGraph+Pythia and Powheg+Pythia simulations. The observed differences are assumed to be produced by differences in the hard scattering models only.

A typical generator model uncertainty of 0.9 (1.1)% is determined for results in visible (full) phase space, respectively.

Parton Showering and Hadronisation Model

In Chapter 4 the different showering and hadronisation models implemented in Pythia and Herwig were discussed. The parton showering and hadronisation uncertainties are evaluated simultaneously by comparing the results obtained using the $t\bar{t}$ simulations produced by Powheg+Pythia and MC@NLO+Herwig. Both Powheg and MC@NLO generate $t\bar{t}$ events up to NLO accuracy at matrix element level, and therefore it is assumed that differences in the generated events are only due to the showering and hadronisation models.

The showering and hadronisation uncertainty is typically in the 1.2–1.5% range.



(a) (b)

Figure 7.1: Generator distributions of the top-quark-pair (a) and leading top quark (b) transverse momentum produced by MadGraph+Pythia and Powheg+Pythia models. The distributions are normalised to unit area. In the bottom panel the ratio with respect to MadGraph+Pythia model is shown.

Top-Quark Mass

The nominal $t\bar{t}$ simulation is generated with a top-quark mass value of $m^{t}_{MC} = 172.5\,GeV$, while the experimentally measured value is $m^{t}_{exp.} = 173.4 \pm 0.74\,GeV$ [110]. Two alternative samples with $m^{t}_{MC} = 171.5\,GeV$ and $173.5\,GeV$ are used to account for the experimental uncertainty arising from the top-quark mass. The transverse momentum spectra of the top quarks as produced by the simulations with different top-quark masses are presented in Figure 7.2. The top-quark-mass uncertainty is found to be 0.5–0.9% depending on the phase space of the measurement.

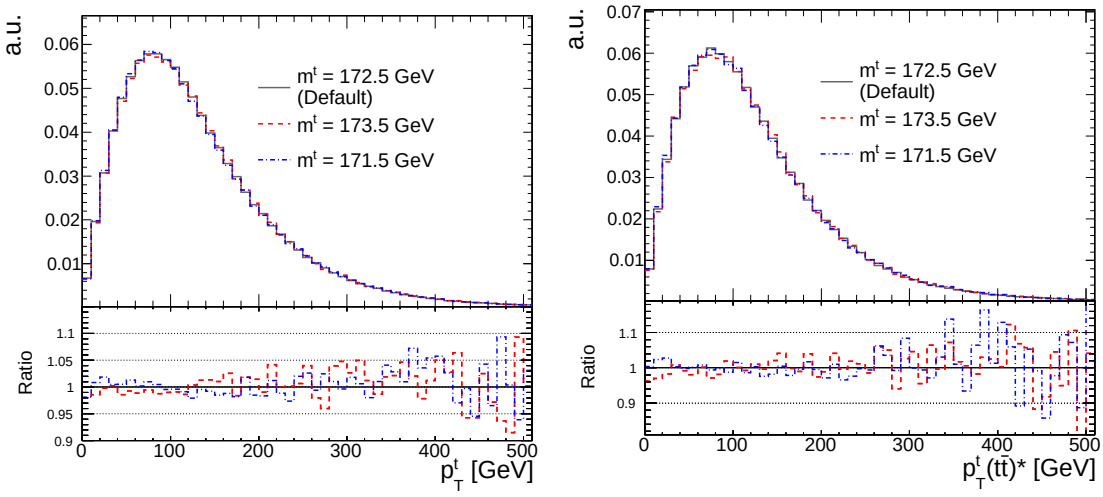
A kinematic reconstruction algorithm is used to reconstruct the 4-momentum components of the individual top quarks (Section 5.8). Possible biases due to the fixed top-quark mass constraint used in the kinematic reconstruction algorithm were studied by varying $m^{t} = 172.5 \pm 1\,GeV$. The effect on the differential cross sections is found to be negligible in comparison to the statistical uncertainty of the result.

Matching and Hard Scattering Scales

The matching of the matrix element and parton showering was discussed in Section 4.2. The systematic uncertainty associated to the matching scale is evaluated by comparing the nominal results for the cross section and the results obtained with two alternative Monte Carlo simulation samples where the matching scale is varied by a factors 2 and 0.5 with respect to the default value of 20 GeV.

The uncertainty on the hard scattering scale Q comprises the renormalisation and factorisation scale. Dedicated $t\bar{t}$ samples are generated varying the Q-scale by a factor 2 and 0.5 with respect to the nominal value, following the convention adopted by the CMS experiment.

The uncertainty on the cross section due to the matching scale is typically 0.7%. The



(a)

Figure 7.2: Generator distributions of the top-quark transverse momentum measured in the laboratory reference system () and in the $t\bar{t}$ rest frame (a) produced by the MadGraph+Pythia model with m^t=171.5 (blue), 172.5 (black) and 173.5 GeV (red). The distributions are normalised to unit area. In the bottom panel the ratio with respect to MadGraph+Pythia model generated with m^t=172.5 GeV is shown.

result uncertainty due to the Q-scale ranges between 0.7% and 1.1%, depending on the phase space of the measurement.

7.4 Summary of Typical Uncertainties

A summary of the typical uncertainties on the measured differential cross section is listed in Table 7.2, separately for results measured in the visible phase space and in the extrapolated full phase space. Tables containing the statistical, total systematic and total error in each bin for all measured variables can be found in Appendix C.1. The main systematic uncertainty of the result is due to the parton showering and hadronisation modelling which amounts to typically 1.2%–1.5%.

The inclusive $t\bar{t}$ production cross section is measured with a total systematic error of 5.2% and with an additional 2.6% uncertainty due to the luminosity measurement [57]. The main uncertainty is due to the integrated luminosity with dominant systematic uncertainty sources arising from the jet energy scale and lepton identification and isolation efficiencies, approximately 2.2% each. Among the theory uncertainties, the precision of the result is limited by the hard scattering and parton shower and hadronisation models. Table 7.3 summarises the measured uncertainty in the inclusive $t\bar{t}$ production cross section separately for the individual channels and uncertainty sources.

7.5 Correlation Matrices of the Total Uncertainty

The determination of SM parameters from experimental results benefits from the knowledge of the correlation between the bins [111], e.g. the PDF fit using $t\bar{t}$ differential cross sections.

Uncertainty Source	**Uncertainty [%]**	
	Visible Phase Space	Full Phase Space
Experimental		
Trigger Efficiency	0.1	0.1
Lepton Id. and Isolation	0.1	0.1
Drell-Yan Normalisation	0.2	0.2
Background Normalisation	0.3	0.3
Jet Energy Scale	0.4	1.0
Jet Energy Resolution	0.1	0.1
Pileup	0.1	0.1
b-tagging	0.1	0.2
Kinematic Reconstruction	< 0.1	< 0.1
Theory Model		
Hard Scattering Model	0.9	1.1
Showering - Hadronisation Model	1.5	1.2
Top quark mass	0.5	0.9
Q^2 scale	0.7	1.1
ME/PS matching	0.7	0.7
PDF	0.3	0.5
Total Systematic	2.2	2.5

Table 7.2: Summary of the typical value of the systematic errors on the normalised differential cross sections separately for variables measured in the visible phase space and full phase space. The uncertainty due to the normalisation of background processes is separated into two contributions: $Z/\gamma^* \to ee/\mu\mu$ events (Drell-Yan) and other backgrounds.

The correlation information is parametrised in the *correlation matrix*, $corr_{ij}$. It is derived using the method proposed in [112] and briefly described in the following.

The correlation matrix is derived from the statistical covariance matrix cov_{ij} of the systematic uncertainties. The latter is constructed from the differences of the results for each of the uncertainty sources.

The absolute difference d_i in each bin i of the measurement due to a systematic source is

$$d_i^{up/down} = \left(\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}X}\right)_i^{up/down} - \left(\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}X}\right)_i^{nominal} \tag{7.3}$$

for every *up* and *down* variation. Each variation is then averaged and applied symmetrically in both directions of the result:

$$\delta_i = \mathrm{sign}(d_i^{up}) \cdot \frac{1}{2}\left(|d_i^{up}| + \left|d_i^{down}\right|\right). \tag{7.4}$$

In order to recover the sign information lost while symmetrising, the convention to consider the sign of the *up* variation is selected [112]. Covariance matrices for each systematic source s are constructed via

$$cov_{ij}(s) = \delta_i(s) \cdot \delta_j(s). \tag{7.5}$$

All uncertainty sources are assumed uncorrelated and therefore are summed up in quadrature, as explained in Section 7.1. The diagonal elements of the covariance matrix contain the

Uncertainty Source	**Uncertainty [%]**			
Channel	ee	eμ	$\mu\mu$	Combined
Experimental				
Trigger Efficiency	1.4	1.3	1.3	1.3
Lepton Id. and Isolation	2.3	2.2	2.3	2.2
Drell-Yan Normalisation	0.8	< 0.1	1.1	0.4
Background Normalisation	2.1	2.1	1.8	2.0
Jet Energy Scale	2.3	2.2	2.2	2.2
Jet Energy Resolution	< 0.1	< 0.1	< 0.1	< 0.1
Pileup	0.2	0.4	0.1	0.3
b-tagging	0.6	0.6	0.6	0.3
Kinematic Reconstruction	0.5	0.2	0.4	0.1
Branching Ratio	1.5	1.5	1.5	1.5
Theory Model				
Hard Scattering Model	1.9	1.7	1.1	1.6
Showering - Hadronisation Model	2.2	2.0	1.1	1.8
Top quark mass	0.9	0.7	1.1	0.8
Q^2 scale	1.8	1.1	2.0	1.4
ME/PS matching	0.8	0.4	0.7	0.5
PDF	0.3	0.4	0.2	0.3
Total Systematic	5.8	5.2	5.2	5.2
Luminosity	2.6	2.6	2.6	2.6

Table 7.3: Systematic uncertainty values on the total inclusive $\mathrm{t\bar{t}}$ production cross section.

squared values of the systematic uncertainty. Thus, the sum of the covariance matrices is assumed equivalent to the quadratic sum of the systematic uncertainties [112].

The final systematic correlation matrix is derived after normalising the covariance matrix elements by the corresponding diagonal components: $corr_{ij} = cov_{ij}/\sqrt{cov_{ii} \cdot cov_{jj}}$. The correlation matrices for all measured quantities are presented in Appendix E.

Chapter 8

Results

In this chapter the results of the $\mathrm{t\bar{t}}$ normalised differential cross section measurement are presented. Results are measured as a function of 19 kinematic variables of the top quarks, the $\mathrm{t\bar{t}}$ system and the top-quark-decay products: leptons, lepton pair, b-jets and b-jet pair. The results, shown in Figures 8.1–8.9, are in preparation for publication by the CMS Collaboration.

The cross sections are measured as a function of the top-quark-pair invariant mass $m^{\mathrm{t\bar{t}}}$, the rapidity $y^{\mathrm{t\bar{t}}}$ and the transverse momentum $p_{\mathrm{T}}^{\mathrm{t\bar{t}}}$. The measurements as a function of angular distributions of the top quarks comprise the rapidity y^{t}, the absolute-rapidity difference $|y^{\mathrm{t}}| - |y^{\bar{\mathrm{t}}}|$ and the difference in azimuthal angle of the two top quarks $\Delta\phi\,(\mathrm{t},\bar{\mathrm{t}})$. Additionally, the transverse momentum of the top quarks is measured in the laboratory, $p_{\mathrm{T}}^{\mathrm{t}}$, and in the $\mathrm{t\bar{t}}$ rest frame, $p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{t\bar{t}})^{*}$, and separately for the p_{T}-ordered leading $p_{\mathrm{T}}^{\mathrm{t}}(\text{lead. t})$ and the trailing $p_{\mathrm{T}}^{\mathrm{t}}(2^{\mathrm{nd}}\,\text{lead. t})$ top quark. Top quark and $\mathrm{t\bar{t}}$ variables are measured in the extrapolated full phase space as defined in Section 6.1.

The differential cross sections obtained as a function of the lepton and b-jet kinematic variables are measured in the visible phase space as defined by the detector acceptance (see Section 6.1). These results comprise distributions of the lepton transverse momentum $p_{\mathrm{T}}^{\mathrm{l}}$, the pseudorapidity η^{l} and the azimuthal-angle difference $\Delta\phi^{\mathrm{l\bar{l}}}$, the lepton-pair invariant mass $m^{\mathrm{l\bar{l}}}$ and the transverse momentum $p_{\mathrm{T}}^{\mathrm{l\bar{l}}}$, the transverse momentum $p_{\mathrm{T}}^{\mathrm{b}}$ and the pseudorapidity η^{b} of the b-jets, and the invariant mass $m^{\mathrm{b\bar{b}}}$ and the transverse momentum $p_{\mathrm{T}}^{\mathrm{b\bar{b}}}$ of the system formed by the pair of b-jets.

In the following, the distributions corresponding to the combination of the individual ee, eμ and $\mu\mu$ channels are shown and discussed. Summary tables with the results of the normalised differential cross sections including the statistical, systematic and total uncertainty components in each bin of the different variables are presented in Appendix C.1. The differential cross sections measured separately in the ee, eμ and $\mu\mu$ channels are shown in Appendix C.2.

The figures include data points with an inner error bar indicating the statistical component of the uncertainty and an outer error bar indicating the total uncertainty obtained as the quadratic sum of the statistical and systematic uncertainties. The same components of uncertainties on the data are represented by the error bands in the theory-to-data ratios shown in the bottom panels of the figures.

All results are compared to four Monte Carlo predictions: MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig. Additionally, the top-quark-pair invariant mass and transverse momentum results are compared to a QCD calculation up to NLO+NNLL accuracy [100, 101], and the top-quark p_{T} and rapidity results are compared to two approximate NNLO calculations [3,21]. The calculation from [3] includes a theory uncer-

tainty band due to the choice of the PDF set, the strong-coupling constant α_s, the top-quark mass, the renormalisation and the factorisation scales. All theory predictions are presented in the same binning as the data for direct visual comparison.

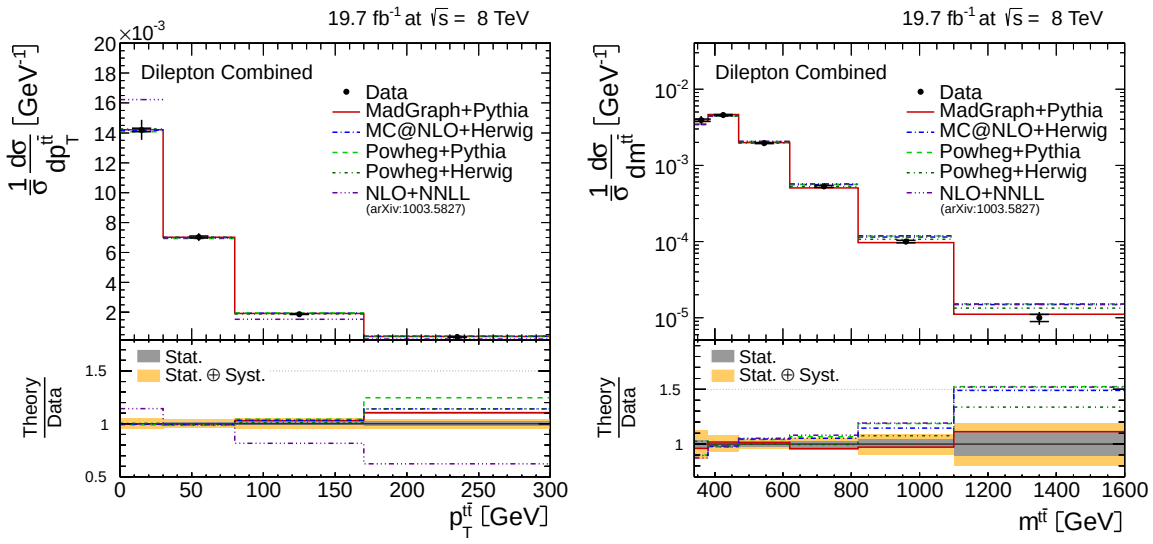
The differential cross sections of the top quark and top-quark pair are discussed in Section 8.1, followed by the discussion of the results measured in the visible phase space in Section 8.2. A χ^2 statistical test is performed to evaluate the consistency of the measurement with the different predictions (see Section 8.3). Several checks are performed to ensure the stability of the results as discussed in Section 8.4. The results are compared in Section 8.5 to the measurements obtained in other analyses. The chapter finishes with a discussion of the total $\mathrm{t\bar{t}}$ production cross section (see Section 8.6).

8.1 Top-Quark-Pair and Top Quark Differential Cross Sections

The top-quark-pair transverse momentum ($p_{\mathrm{T}}^{\mathrm{t\bar{t}}}$) is shown in Figure 8.1a. This variable is sensitive to higher-order terms of the QCD perturbative expansion because terms beyond LO in the perturbative expansion include the emission of radiation which balances the transverse momentum of the $\mathrm{t\bar{t}}$ system. The differential cross section is measured up to 300 GeV with an uncertainty smaller than 5.1% limited by the parton showering and hadronisation modelling differences in Pythia and Herwig. In general, the Monte Carlo predictions describe the data well, while the NLO+NNLL theory calculation [101] is steeper than the data and fails to describe the observed shape. In Figure 8.1b the top-quark-pair invariant-mass ($m^{\mathrm{t\bar{t}}}$) result is shown. This distribution is sensitive to possible resonances due to BSM models [27, 28, 113]. The $m^{\mathrm{t\bar{t}}}$ result is measured with an uncertainty ranging between 4.2% at small $m^{\mathrm{t\bar{t}}}$ values increasing up to 19% at high $m^{\mathrm{t\bar{t}}}$. The latter is dominated by the statistical uncertainty of the data 10% in the last bin of the measurement, $m^{\mathrm{t\bar{t}}} \in [1100\,\mathrm{GeV}, 1600\,\mathrm{GeV}]$. The result is well described by the different predictions. In the last bin, $m^{\mathrm{t\bar{t}}} \gtrsim 1100\,\mathrm{GeV}$, the expectations from MC@NLO+Herwig, Powheg+Pythia and the NLO+NNLL calculation [100] predict higher cross section values than the measurement. The χ^2-statistical-test result (Table 8.1) indicates that the best description is given by the MadGraph+Pythia prediction.

The Figure 8.2a shows the normalised differential cross section measured as a function of the top-quark-pair rapidity ($y^{\mathrm{t\bar{t}}}$). This result presents a maximum uncertainty of $\sim 13\%$. This uncertainty is dominated by the different hard scattering models implemented in MadGraph and Powheg generators. Within the current uncertainties all predictions provide a good description of the data. The differential cross section as a function of the top-quark rapidity (y^{t}) is well described by the SM predictions (see Figure 8.2b). The spectra simulated by Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig models are very similar. In contrast, MadGraph+ Pythia expects a more central distribution. The results are measured with an uncertainty smaller than 3% in $|y^{\mathrm{t}}| \leq 1.6$ increasing up to approximately 6.0% in the outermost bins.

The $\mathrm{t\bar{t}}$ differential cross sections as a function of the difference of the angular variables $\Delta\phi\,(\mathrm{t}, \bar{\mathrm{t}})$ and $|y^{\mathrm{t}}| - |y^{\bar{\mathrm{t}}}|$ between the top quark and antiquark quark are shown in Figure 8.3. In absence of additional radiation the $\Delta\phi\,(\mathrm{t}, \bar{\mathrm{t}})$ distribution is expected to be distributed in the vicinity of $\Delta\phi\,(\mathrm{t}, \bar{\mathrm{t}}) = \pi$ because the top quarks are produced back-to-back in the transverse plane of the detector. The result is very well described by all Monte Carlo predictions within the precision of the measurement which achieves a systematic uncertainty of 3.7–8.8%, depending on the bin. The top-quark charge asymmetry can be measured using a variable proportional to $|y^{\mathrm{t}}| - |y^{\bar{\mathrm{t}}}|$, as explained in Chapter 2. The $|y^{\mathrm{t}}| - |y^{\bar{\mathrm{t}}}|$ differential cross section is found to be well described by the Powheg+Pythia, Powheg+Herwig and



(a) (b)

Figure 8.1: Normalised differential cross sections as a function of the transverse momentum (a) and invariant mass (b) of the top-quark-pair system. The inner (outer) error bar in the data points indicate the statistical (total) uncertainties. The ratios of the predictions over data are shown in the bottom panels, with the uncertainties of the measurement indicated by the coloured bands.

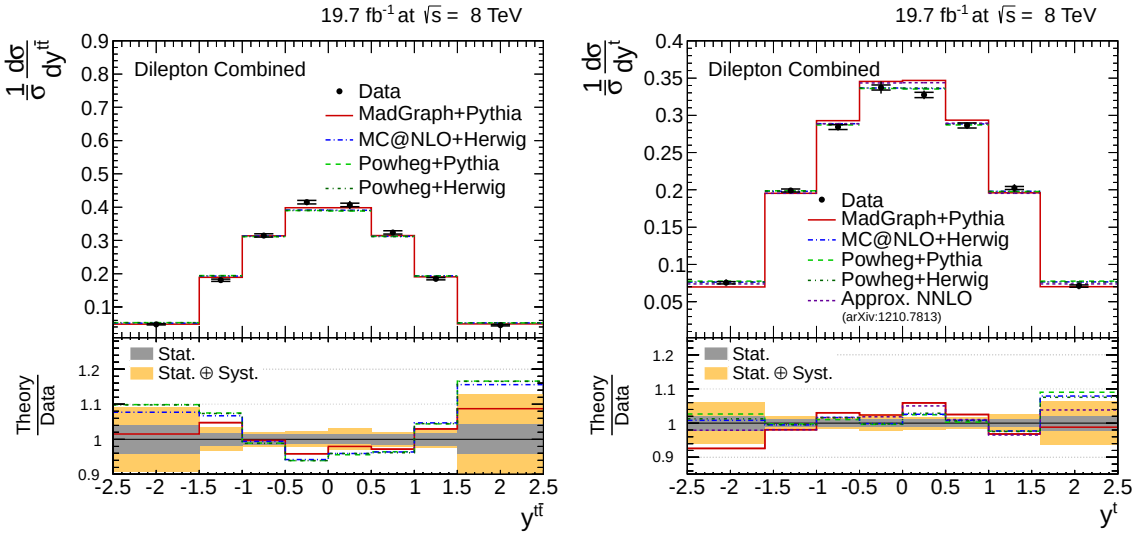
MC@NLO+Herwig predictions, while the MadGraph+Pythia model predicts a more central result. The precision of the result is limited by the uncertainties on the Q^2 scale and hard scattering model, typically 1.6% and 1.4% respectively.

The top-quark transverse momentum distribution measured in the laboratory system ($p_{\mathrm{T}}^{\mathrm{t}}$) is presented in Figure 8.4. The result has a maximum uncertainty of 8.4% in the $p_{\mathrm{T}} \in [290\,\mathrm{GeV}, 400\,\mathrm{GeV}]$ bin, due to the top-quark mass and hadronisation model uncertainties. The result, within the current uncertainties, is well described by the Powheg+Herwig simulation and by the theory calculations up to approximate NNLO accuracy. Distributions generated by the MadGraph+Pythia, Powheg+Pythia and MC@NLO+Herwig models predict top-quark transverse momentum spectra that are harder than the observed in data. This observation is corroborated by the χ^2-test results summarized in Table 8.1.

Similar conclusions can be drawn from distributions of the top quark transverse momentum measured in the $\mathrm{t\bar{t}}$ rest frame ($p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{t\bar{t}})^*$), and for the transverse momenta of the leading and trailing top quark $p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{lead.\,t})$ and $p_{\mathrm{T}}^{\mathrm{t}}(2^{\mathrm{nd}}\,\mathrm{lead.\,t})$, respectively. These results are shown in Figure 8.5.

The $p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{t\bar{t}})^*$ distribution is well described by the Powheg+Herwig prediction, while the alternative MadGraph+Pythia, Powheg+Pythia and MC@NLO+Herwig simulations predict in general harder spectra in comparison to the data. The precision of this variable is limited by the uncertainties due to the hard scattering and parton showering and hadronisation models which produce uncertainties up to 5.4% at very high transverse momentum values, $p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{t\bar{t}})^* \in [400\,\mathrm{GeV}, 500\,\mathrm{GeV}]$.

The $p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{lead.\,t})$ and $p_{\mathrm{T}}^{\mathrm{t}}(2^{\mathrm{nd}}\,\mathrm{lead.\,t})$ results reveal softer spectra than the Monte Carlo predictions. MC@NLO+Herwig, MadGraph+Pythia and Powheg+Pythia present differences up to 40% with respect to data at very high p_{T} values, while Powheg+Herwig

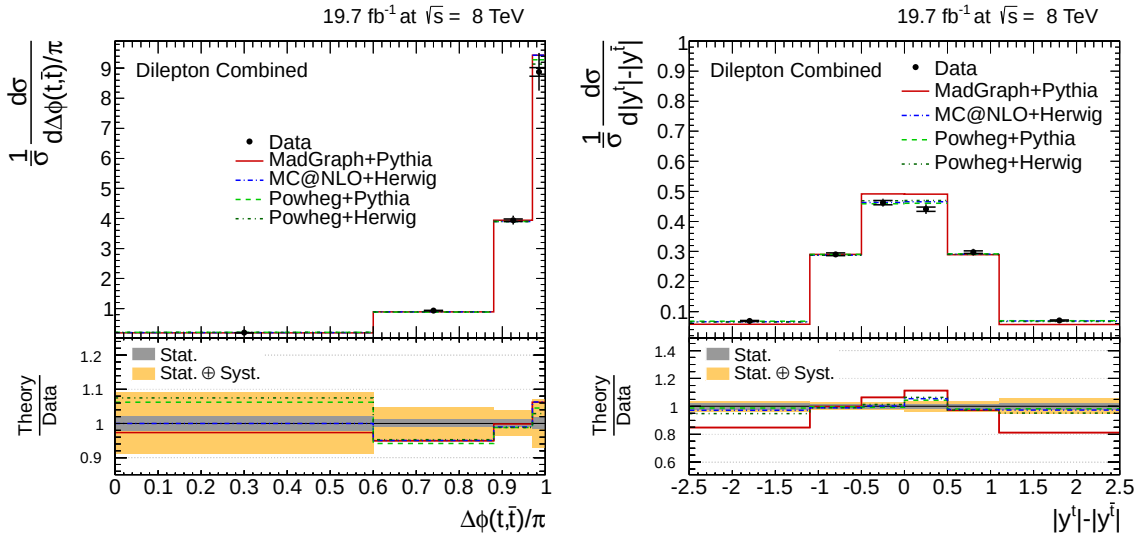


(a) (b)

Figure 8.2: Normalised differential cross sections as a function of the rapidity distribution of the $t\bar{t}$ system (a) and top quarks (b). The inner (outer) error bar in the data points indicate the statistical (total) uncertainties. The ratios of the predictions over data are shown in the bottom panels, with the uncertainties of the measurement indicated by the coloured bands.

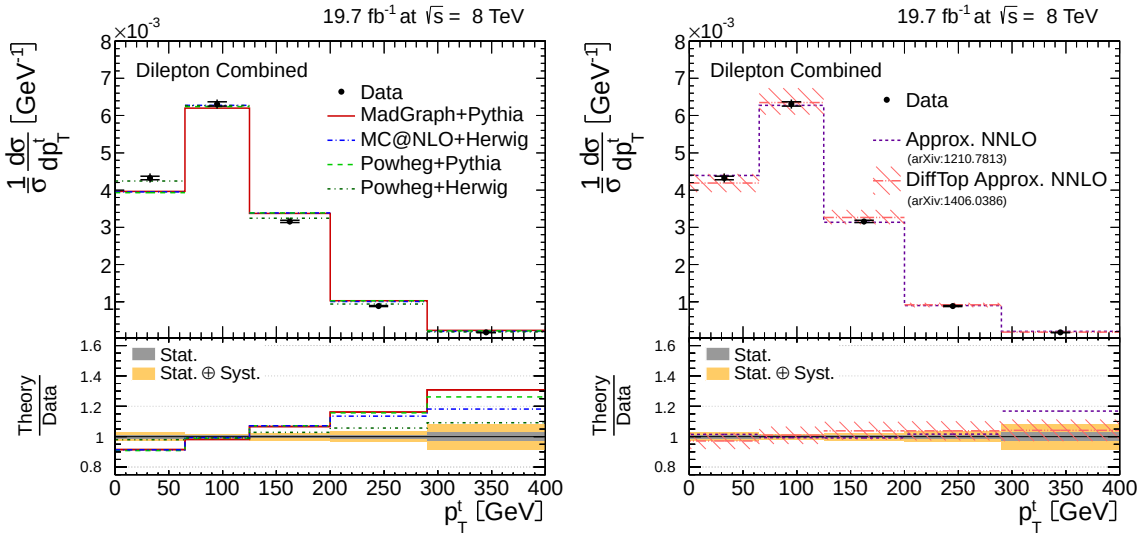
produces a good description of the data in the measured range. The precision of the measurements is limited by the parton showering and hadronisation model uncertainties which range between 0.5% and 5.5% with increasing momentum.

Differences in the data and predicted spectra are observed for the top-quark transverse momentum. The differences are also seen when ordering the individual top quark or antiquark according to their p_{T}, and when removing possible effects due to an initial boost of the $t\bar{t}$ system. Furthermore, disparate predictions arise from POWHEG+PYTHIA and POWHEG+HERWIG models. Recently it has been discovered that these differences are due to the different treatments of the initial state radiation and the technical implementation of the energy-momentum conservation in the HERWIG model with respect to PYTHIA [114]. Further investigations and technical developments are under evaluation.



(a) (b)

Figure 8.3: Normalised differential cross sections as a function of the difference in the azimuthal angle of the top quarks (a) and the difference of the absolute rapidities of the top quark and antiquark (b). The inner (outer) error bar in the data points indicate the statistical (total) uncertainties. The ratios of the predictions over data are shown in the bottom panels, with the uncertainties of the measurement indicated by the coloured bands.



(a) (b)

Figure 8.4: Normalised differential cross section as a function of the top-quark p_{T} measured in the laboratory system. The results are compared to predictions from Monte Carlo simulations (a) and to QCD perturbative calculations up to approximate NNLO accuracy (b), including the theory uncertainties due to renormalisation and factorisation scales, PDF, top mass and α_s [3]. The inner (outer) error bar in the data points indicate the statistical (total) uncertainties. The ratios of the predictions over data are shown in the bottom panels, with the uncertainties of the measurement indicated by the coloured bands.

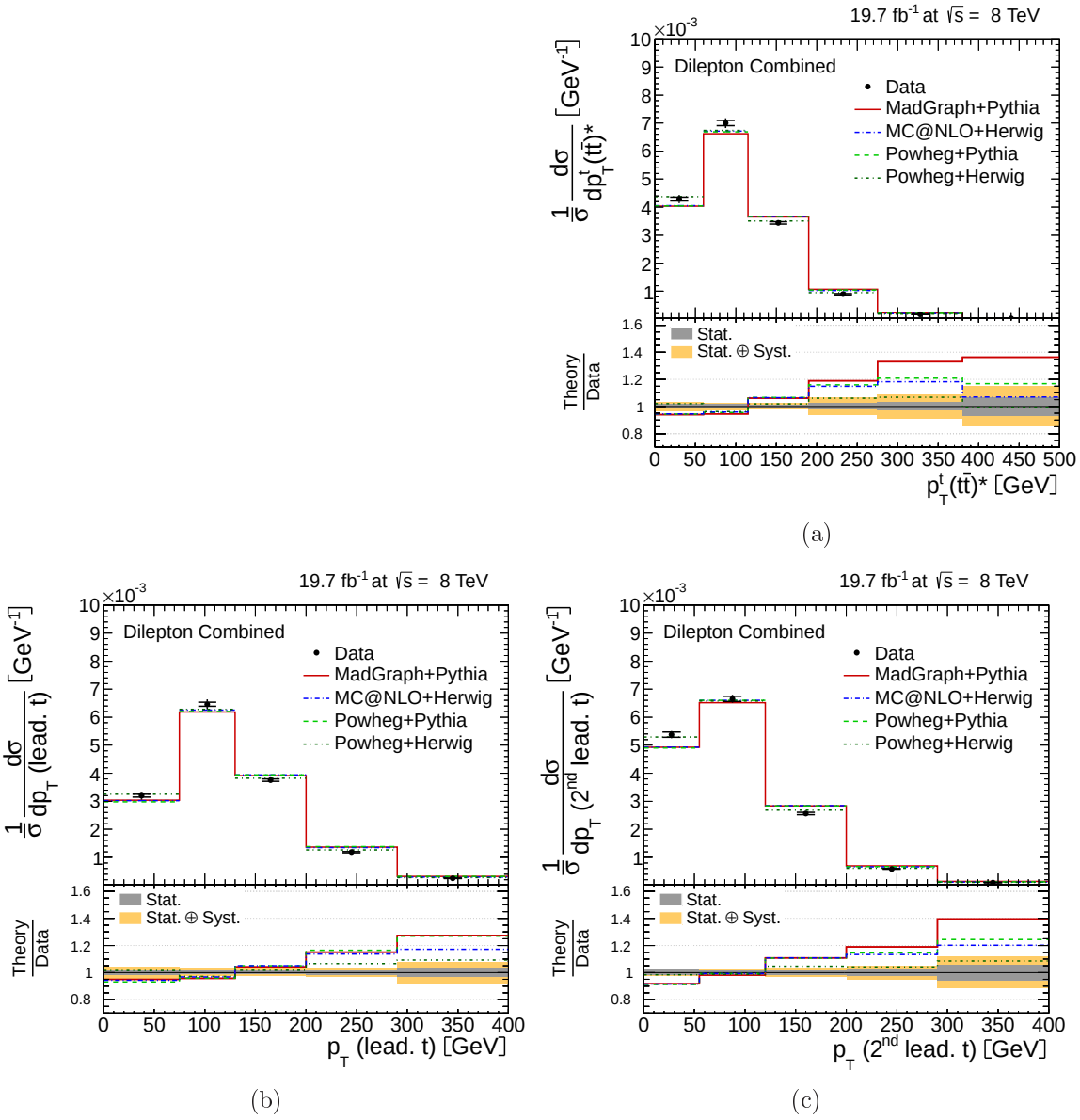


Figure 8.5: (a): Normalised differential cross section as a function of the top-quark transverse momentum measured in the $t\bar{t}$ rest frame $p_T^t(t\bar{t})^*$. The results as a function of the p_T distribution of the leading, p_T^t(lead. t), and trailing, $p_T^t(2^{nd}$ lead. t), top quark are shown in (b) and (c), respectively. The inner (outer) error bar in the data points indicate the statistical (total) uncertainties. The ratios of the predictions over data are shown in the bottom panels, with the uncertainties of the measurement indicated by the coloured bands.

8.2 Lepton and b-Jet Differential Cross Sections

The normalised differential cross section measured as a function of the leptons and b-jets are presented in the visible phase space as defined in Section 6.1.

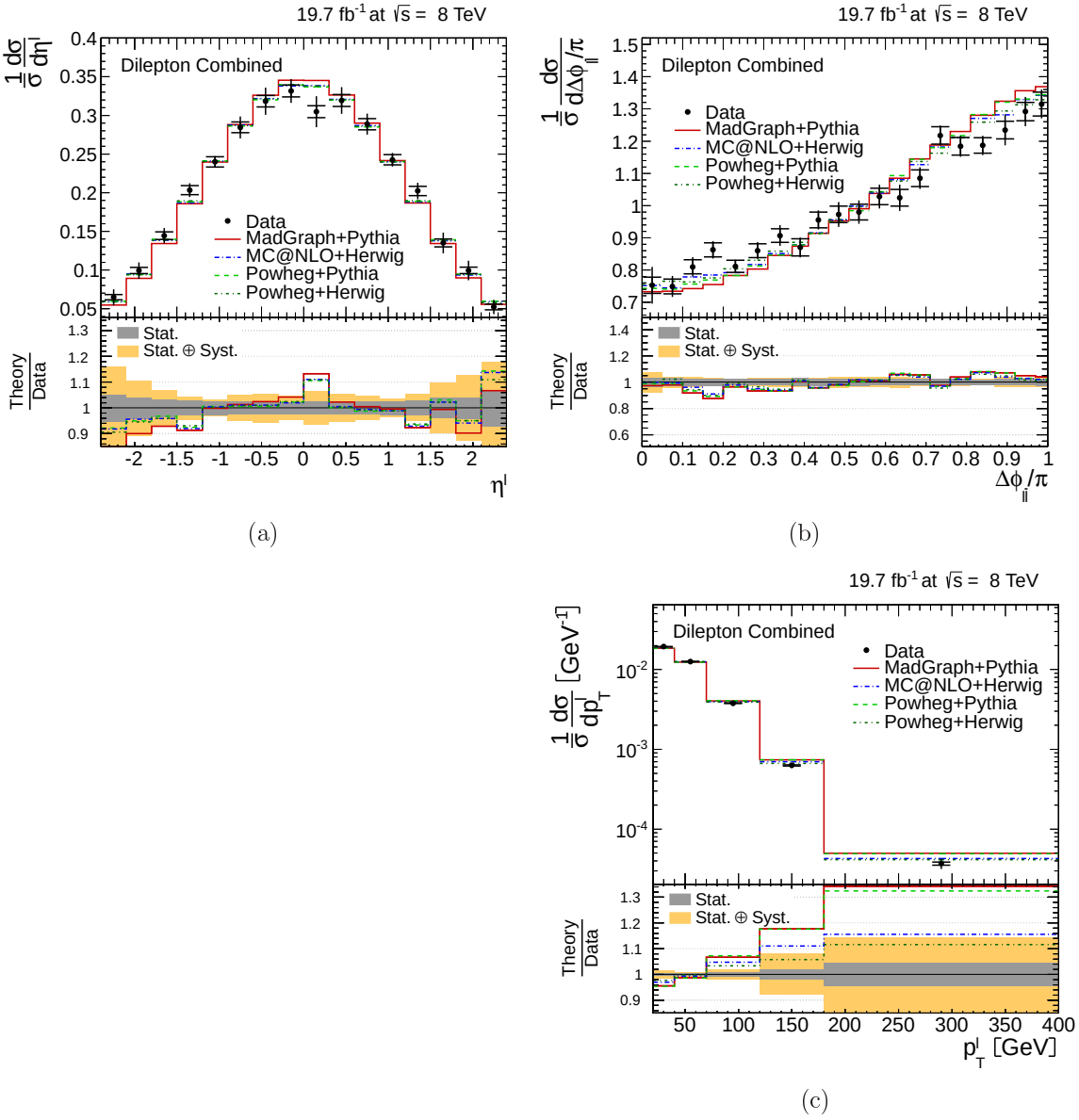
The lepton pseudorapidity (η^{l}) result is shown in Figure 8.6a. The measurement has a typical uncertainty of up to 15% in the outermost bins ($1.8 < |\eta^{l}| < 2.4$). Within this accuracy all predictions agree with the data. Due to the very short top-quark lifetime the spin correlations between the top quark and antiquark are transferred to the final state leptons, as discussed in Chapter 2. The $\Delta\phi^{l\bar{l}}$ variable is sensitive to the spin correlations. The results is shown in Figure 8.6b, and within the uncertainty of the data all Monte Carlo predictions describe the data well.

The p_{T} distribution of the leptons, shown in Figure 8.6c, shows a similar behaviour as that observed for the top-quark p_{T}, i.e. the measured differential cross section has a more rapidly falling spectrum than the predictions. The best description is obtained by the POWHEG+HERWIG prediction, while MADGRAPH and POWHEG interfaced to PYTHIA generate leptons with higher transverse momenta similarly as the MC@NLO+HERWIG simulation. The differential cross section as a function of the b-jet p_{T} is shown in Figure 8.7a. The data spectrum shows a softer spectrum than the expectations. The difference between data and prediction in the lepton and b-jet transverse momentum spectra is due to the difference in the top-quark spectrum, which is transferred to the decay products: leptons and b-jets.

The b-jet pseudorapidity (η^{b}) distribution presents a broader distribution in data than the predictions from Monte Carlo simulations (see Figure 8.7b). Data and Monte Carlo predictions agree within the total uncertainty, which ranges between 2.0% and 6.0%, mainly driven by the jet energy and Q^{2}-scale uncertainties.

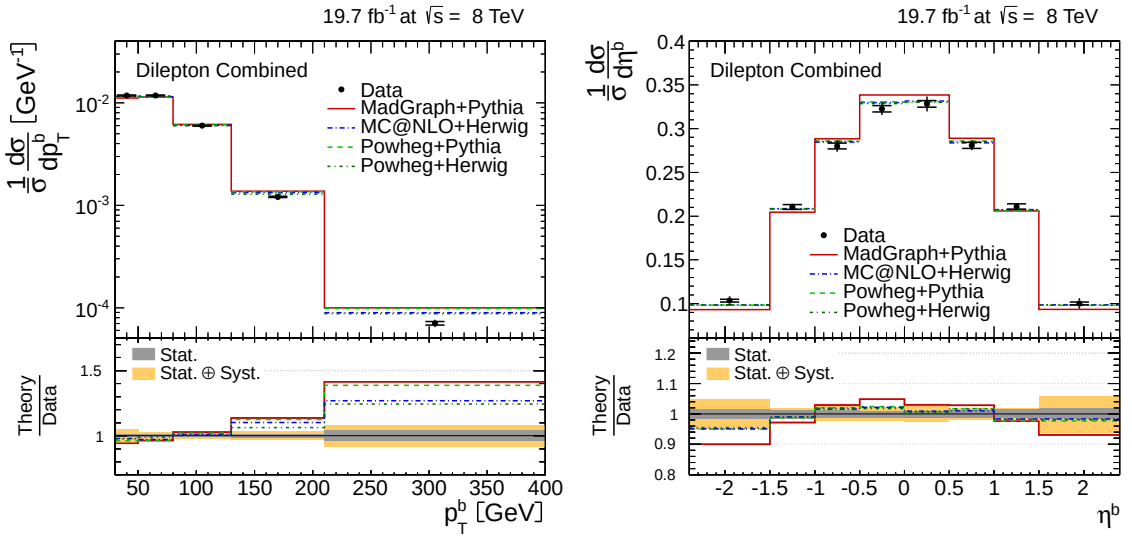
The results as a function of the transverse momentum of the lepton-pair ($p_{T}^{l\bar{l}}$) and b-jet-pair ($p_{T}^{b\bar{b}}$) are shown in Figure 8.8. These variables are sensitive to the spin correlations of the $t\bar{t}$ system, as explained in Section 2.2. Generally the various Monte Carlo predictions describe the data well, as confirmed by the χ^{2}-test results listed in Table 8.2. The best description is given by POWHEG model interfaced to HERWIG. In contrast, POWHEG interfaced to PYTHIA predicts a higher cross section at large lepton-pair transverse momenta. The main uncertainties are due to the different models used to describe the parton showering and hadronisation, which are typically in the range 1.5–6.0%, depending on the bin and on the distribution.

The differential cross section as a function of the invariant masses of the lepton pair ($m^{l\bar{l}}$) and of the b-jet pair ($m^{b\bar{b}}$) are presented in Figure 8.9. The $m^{l\bar{l}}$ distribution is measured with a precision better than 5% mainly limited by the parton shower and hadronisation modelling uncertainty. All Monte Carlo expectations predict $t\bar{t}$ cross sections larger than those observed in the data. The $m^{b\bar{b}}$ spectrum, shown in Figure 8.9b, is well described by all predictions. The dominant uncertainty is due to the parton showering and hadronisation model described by PYTHIA and HERWIG.



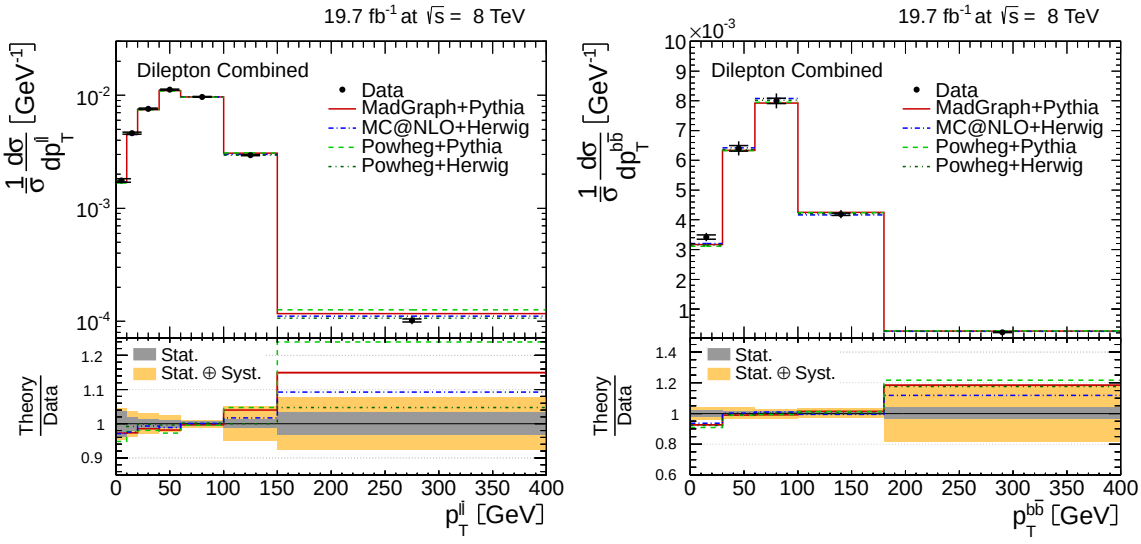
(a) (b) (c)

Figure 8.6: Normalised differential cross sections as a function of the lepton pseudorapidity (a), the difference in azimuthal angles of the leptons (b) and the lepton transverse momentum (c). The inner (outer) error bar in the data points indicate the statistical (total) uncertainties. The ratios of the predictions over data are shown in the bottom panels, with the uncertainties of the measurement indicated by the coloured bands. Results are obtained in the visible phase space (see Section 6.1).



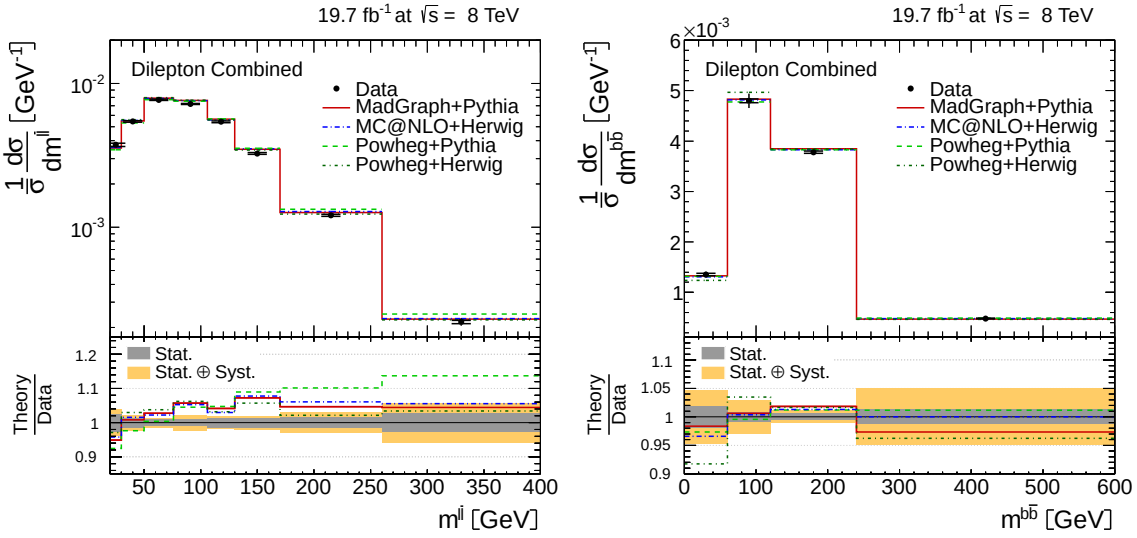
(a) (b)

Figure 8.7: Normalised differential cross sections as a function of the p_{T} (a) and pseudorapidity (b) of the b-jets. The inner (outer) error bar in the data points indicate the statistical (total) uncertainties. The ratios of the predictions over data are shown in the bottom panels, with the uncertainties of the measurement indicated by the coloured bands. Results are obtained in the visible phase space (see Section 6.1).



(a) (b)

Figure 8.8: Normalised differential cross sections as a function of the transverse momentum transverse momentum of the lepton pair (a) and of the b-jet pair (b). The inner (outer) error bar in the data points indicate the statistical (total) uncertainties. The ratios of the predictions over data are shown in the bottom panels, with the uncertainties of the measurement indicated by the coloured bands. Results are obtained in the visible phase space (see Section 6.1).



(a) (b)

Figure 8.9: Normalised differential cross sections as a function of the invariant mass of the lepton pair (a) and of the b-jet pair (b). The inner (outer) error bar in the data points indicate the statistical (total) uncertainties. The ratios of the predictions over data are shown in the bottom panels, with the uncertainties of the measurement indicated by the coloured bands. Results are obtained in the visible phase space (see Section 6.1).

8.3 Compatibility Tests of the Results

A numerical method to evaluate the agreement or disagreement of a certain theory model with the data is based on normalised-χ^2 test. This test quantifies the deviation of a prediction from the measured data relative to the total data uncertainty averaged over all independent bins of the measurement. The χ^2 is defined as

$$\chi^2/\text{ndof} = \frac{1}{\text{ndof}} \sum_{\text{N}_{\text{bins}}} \frac{(\text{result}_i - \text{theory}_i)^2}{\text{uncertainty}_i^2} \tag{8.1}$$

where *result* and *theory* indicate the i^{th}-bin differential cross section value obtained in data and in the prediction under evaluation, and *uncertainty* refers to the total error on the measurement. The number of degrees of freedom (ndof) corresponds to the number of bins, reduced by one unit due to the normalisation of the result, i.e. a normalised differential cross section measured in 6 bins will contain only 5 degrees of freedom since the result in the 6^{th} bin is already constrained by the normalisation condition.

The measured value of χ^2 is transformed into a probability value, p-value [107]. The p-value indicates the probability that the χ^2 is larger than or equal the observed value, assuming a valid theory. A large p-value, $p \to 1$, indicates that the theory assumption provides a good data description, while small p-values, $p \to 0$, represents a bad description of the data by the theory.

The results of the normalised-χ^2 and corresponding p-values are presented in the Tables 8.1 and 8.2. In general, the best description of the data is given by the Powheg+Herwig prediction, and by the higher-order QCD calculations, depending on the observable. All cross section results as a function of the rapidity or pseudorapidity of the particles are described well by all predictions. The $|y^{\text{t}}| - |y^{\bar{\text{t}}}|$ result is not so well described by the MadGraph+Pythia in comparison to Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig.

	χ^2/ndof (p-value)				
Distribution	MadGraph +Pythia	Powheg +Pythia	Powheg +Herwig	MC@NLO +Herwig	QCD Calculation
p_{T}^{t}	12.9 (0.01)	12.2 (0.02)	1.4 (0.84)	9.2 (0.06)	1.2 (0.88) [21] 1.1 (0.90) [3]
y^{t}	3.2 (0.87)	0.8 (1.00)	0.9 (1.00)	0.9 (1.00)	1.9 (0.96) [21]
$p_{\text{T}}^{\text{t}}(\text{t}\bar{\text{t}})^*$	8.8 (0.12)	5.5 (0.36)	1.0 (0.97)	4.7 (0.46)	–
$p_{\text{T}}^{\text{t}}(\text{lead. t})$	10.5 (0.03)	12.0 (0.02)	1.9 (0.76)	7.8 (0.10)	–
$p_{\text{T}}^{\text{t}}(2^{\text{nd}}\,\text{lead. t})$	13.2 (0.01)	10.1 (0.04)	1.2 (0.89)	9.2 (0.06)	–
$\Delta\phi\,(\text{t}, \bar{\text{t}})$	0.6 (0.89)	0.8 (0.85)	0.6 (0.89)	0.6 (0.89)	–
$\lvert y^{\text{t}}\rvert - \lvert y^{\bar{\text{t}}}\rvert$	8.7 (0.12)	0.4 (1.00)	1.3 (0.93)	0.7 (0.98)	–
$p_{\text{T}}^{\text{t}\bar{\text{t}}}$	1.9 (0.60)	9.6 (0.02)	3.1 (0.37)	3.2 (0.36)	29.24 (0.00) [101]
$m^{\text{t}\bar{\text{t}}}$	0.3 (1.00)	3.0 (0.70)	0.8 (0.98)	2.4 (0.79)	3.38 (0.64) [100]
$y^{\text{t}\bar{\text{t}}}$	1.9 (0.96)	4.3 (0.74)	4.3 (0.74)	3.9 (0.79)	–

Table 8.1: Summary of the normalised-χ^2 statistical test and corresponding p-value (in parentheses) for the differential cross sections measured in the full phase space. Results are given separately for the different Monte Carlo predictions and perturbative QCD calculations.

Distribution	χ^2/ndof (p-value)			
	MadGraph +Pythia	Powheg +Pythia	Powheg +Herwig	MC@NLO +Herwig
$p_{\mathrm{T}}^{\mathrm{l}}$	8.0 (0.09)	8.6 (0.07)	1.5 (0.82)	3.3 (0.51)
η^{l}	1.0 (1.00)	0.6 (1.00)	0.6 (1.00)	0.7 (1.00)
$\Delta\phi^{\mathrm{l\bar{l}}}$	3.3 (1.00)	2.8 (1.00)	1.9 (1.00)	2.0 (1.00)
$p_{\mathrm{T}}^{\mathrm{l\bar{l}}}$	1.0 (0.98)	2.3 (0.89)	0.1 (1.00)	0.4 (1.00)
$m^{\mathrm{l\bar{l}}}$	4.9 (0.67)	7.7 (0.36)	4.3 (0.75)	4.8 (0.69)
$p_{\mathrm{T}}^{\mathrm{b}}$	11.8 (0.02)	9.9 (0.04)	3.2 (0.52)	5.4 (0.25)
η^{b}	2.8 (0.90)	0.9 (1.00)	0.6 (1.00)	0.6 (1.00)
$p_{\mathrm{T}}^{\mathrm{b\bar{b}}}$	1.2 (0.89)	1.7 (0.79)	1.2 (0.88)	0.8 (0.94)
$m^{\mathrm{b\bar{b}}}$	1.1 (0.77)	0.6 (0.92)	2.4 (0.49)	0.7 (0.88)

Table 8.2: Summary of the normalised-χ^2 statistical test and corresponding p-value (in parentheses) for the differential cross sections measured in the visible phase space. Results are given separately for the different Monte Carlo predictions.

8.4 Reweighting the Signal Simulation to Data

The analysis presented so far showed differences in the observed and predicted $p_{\mathrm{T}}^{\mathrm{t}}$ spectrum both in the control distributions and in the differential cross section results (see Figures 5.19a and 8.4a). A similar observation was found in the lepton and b-jet transverse momentum distributions.

To evaluate a possible bias due to the shape difference between data and simulation the analysis is repeated after scaling the Monte Carlo signal yields to the data event-by-event using a so-called p_T^t-*reweighting* [115]. This technique uses the detector-independent $\mathrm{t\bar{t}}$ differential cross section measurement to derive event-correction factors. These factors are obtained from the ℓ + jets and dilepton preliminary results of the CMS Collaboration [96, 97] using pp collisions recorded at $\sqrt{s} = 8\,\mathrm{TeV}$ centre-of-mass energy and corresponding to an integrated luminosity of $L = 12.1\,\mathrm{fb}^{-1}$.

The CMS preliminary analysis and the one presented here are based on MadGraph+ Pythia $\mathrm{t\bar{t}}$ simulation model, and therefore the data ratio to this model is used to correct the expected yields. In order to remove possible dependencies on the choice of binning the ratio is fitted with an exponential function in the measured $p_{\mathrm{T}}^{\mathrm{t}}$ ranges after applying a bin-centre-correction [1] (BCC) [116] to the ratio values based on the approximate NNLO prediction [117]. It is checked that alternative predictions introduce negligible differences in the BCCs. The fit result, $f(p_{\mathrm{T}}^{\mathrm{t}}) = \exp\left(0.156 - p_{\mathrm{T}}^{\mathrm{t}} \cdot 1.37 \cdot 10^{-3}\right)$, is then used to correct the events from $\mathrm{t\bar{t}}$ simulation based on an event weight: $w = \sqrt{f(p_{\mathrm{T}}^{\mathrm{t}}) \cdot f(p_{\mathrm{T}}^{\bar{\mathrm{t}}})}$ where $p_{\mathrm{T}}^{\mathrm{t}}$ ($p_{\mathrm{T}}^{\bar{\mathrm{t}}}$) is the generated top-quark (antiquark) transverse momentum. The result of the fit is shown in Figure 8.10. The figure shows also the fit result obtained from $\mathrm{t\bar{t}}$ differential cross section measurements derived in the ℓ+jets and dilepton decay channels by the CMS Collaboration using $\sqrt{s} = 7\,\mathrm{TeV}$ pp collision data [91]. The results obtained at the same collision centre-of-mass energies are compatible with each other, and similarly results from the same $\mathrm{t\bar{t}}$ decay channel are

[1] The horizontal position of the data point in each bin is that where the bin average measured cross section equals the cross section of an unbinned continuous prediction.

compatible with each other despite the difference in the centre-of-mass energy of the collisions.

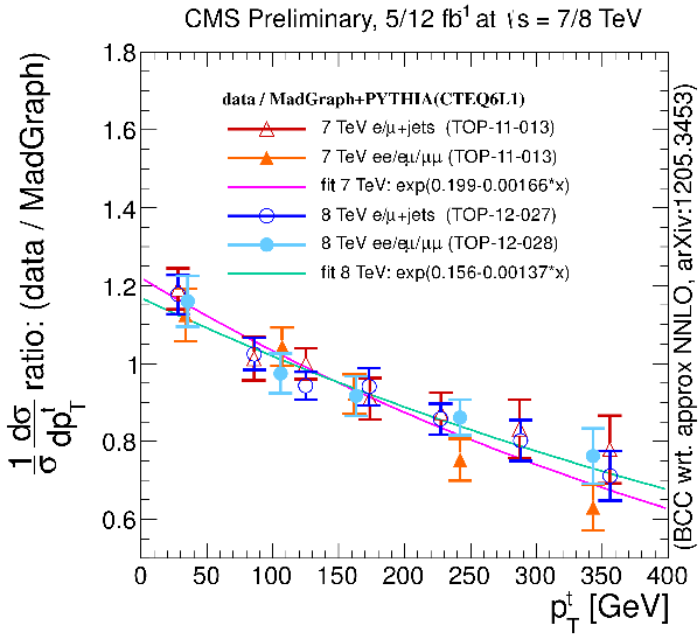


Figure 8.10: Ratio of the differential cross section as a function of the top quark transverse momentum as obtained from data and from MadGraph+Pythia prediction. Results are obtained in the ℓ + jets and dilepton $t\bar{t}$ decay channel using pp collisions recorded at $\sqrt{s}$ = 7 TeV and 8 TeV by the CMS Collaboration [91, 96, 97]. A bin-centre-correction with respect to an approx. NNLO prediction [117] is applied to all data points.

The data yield is very well described by the simulation after applying the aforementioned correction as can be seen in Figures 8.11a and 8.11c. Figures 8.11b and 8.11d show a comparison of the default results and those obtained after correcting the $t\bar{t}$ simulation at an event-by-event basis. The MadGraph+Pythia simulation model with and without the corrections are also presented for reference. The difference between measurements of the normalised differential cross section with and without the reweighting is found to be negligible proving that the results are independent of the shape assumed in the simulation. Additional distributions for all the quantities measured in the analysis are presented in Appendix D.2.

8.5 Comparison to Other Analyses

The CMS Collaboration has presented a preliminary result of the $t\bar{t}$ normalised differential production cross section in the ℓ + jets decay channel in pp collisions at $\sqrt{s}$ = 8 TeV, corresponding to an integrated luminosity of $L = 12.1\,\text{fb}^{-1}$ [96]. The analysis includes results as a function of variables of the top quark, the top-quark pair, the lepton and the b-jets derived from events with one high-p_T isolated electron or muon, and at least four jets from which at least two are required to be identified as b-jets. The kinematic properties of the top-quark pair are determined from the four-momenta of all final-state particles by means of a kinematic fitting algorithm.

The preliminary result of the CMS Collaboration has been extended in [112] by measuring additional distributions and using a total integrated luminosity of $L = 19.7\,\text{fb}^{-1}$ of pp collisions at $\sqrt{s} = 8\,\text{TeV}$. The selection of the events is similar to the official CMS analysis [96] but in this case an iterative kinematic fitting procedure is used to determine the top quark and antiquark 4-momentum components. Additional details can be found in [112].

The semi-leptonic $\text{t}\bar{\text{t}}$ analyses [96, 112] use a visible phase-space definition which differs from the definition used in this analysis (see Section 6.1). A direct comparison of the lepton and b-jet results is therefore not possible. Only the top quark and $\text{t}\bar{\text{t}}$ results can be compared because all these distributions are measured in the extrapolated full phase space.

For the comparison of the three analyses, the data-to-prediction ratio is used to emphasise the differences between the measurements and the predictions. The MadGraph+Pythia model is selected as reference theory since it is the main simulation model in all the analyses. In order to facilitate a direct visual comparison of the results a BCC is applied based on the MadGraph+Pythia model to determine the correct placement of the data result in the figures.

The comparison of the results obtained in the three aforementioned analyses is shown in Figures 8.12 and 8.13. The transverse momentum of the top quark measured in the laboratory reference system and in the $\text{t}\bar{\text{t}}$ rest frame, and separately for the leading and trailing top quark are shown in Figure 8.12. In Figure 8.13 the top-quark rapidity, and the top-quark-pair rapidity, p_{T} and invariant mass are presented. A good consistency is found between the results of this thesis and the semi-leptonic analyses.

8.5.1 Differential Cross Section Results at $\sqrt{s} = 7\,\text{TeV}$

The ATLAS and CMS Collaborations measured the $\text{t}\bar{\text{t}}$ differential cross section in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ in the dilepton and ℓ + jets final state [91, 118]. The analysis from ATLAS include measurements as a function of the top-quark p_{T}, and $\text{t}\bar{\text{t}}$-system transverse momentum, rapidity and invariant mass. The CMS experiment presents additionally results for the top-quark rapidity, and lepton and b-jet observables.

The results of the ATLAS experiment are compared to predictions from Alpgen+ Herwig, Powheg+ Pythia, Powheg+ Herwig and MC@NLO+ Herwig Monte Carlo simulations, and wherever available to higher order perturbative QCD calculations from [100, 117]. Similarly the CMS results are compared to the same QCD calculations an to Monte Carlo predictions from MadGraph+ Pythia, Powheg+ Pythia, Powheg+ Herwig and MC@NLO+ Herwig. The *NLO+NNLL* prediction used in the p_{T}^{t} result of ATLAS corresponds to the *Approx. NNLO* prediction used in CMS, despite the different naming.

The results measured by both collaborations are shown in Figures 8.14–8.17. The results are in good agreement with each other and they observe similar behaviours of the data. In particular, the top-quark p_{T} data spectrum is softer than most of the Monte Carlo predictions, as it is observed in the analysis presented in this thesis. The best description of the ATLAS result is given by the Powheg+Herwig and by the higher-order calculations in the case of CMS. In both $\sqrt{s} = 7\,\text{TeV}$ analyses, the $\text{t}\bar{\text{t}}$ observables are well described by the SM predictions.

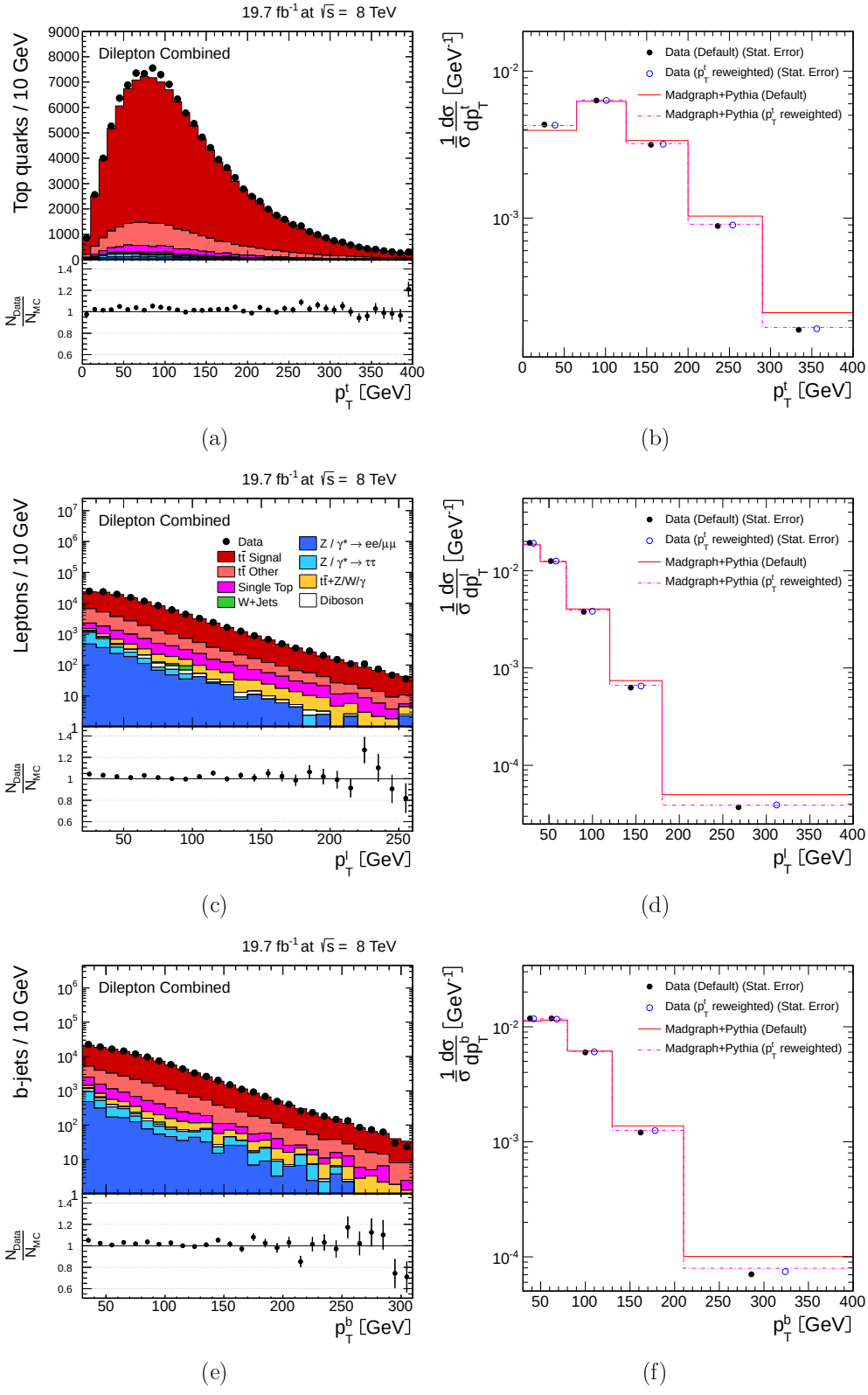


Figure 8.11: Comparison of the differential cross sections (right) with and without the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. The results include only the statistical uncertainty and are compared to the unmodified and $p_{\mathrm{T}}^{\mathrm{t}}$-corrected MadGraph+Pythia prediction. Event yields (left) are obtained after the full selection and include the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. Distributions are shown as a function of the transverse momentum of the top quarks (a and b), leptons (c and d) and b-jets (e and f).

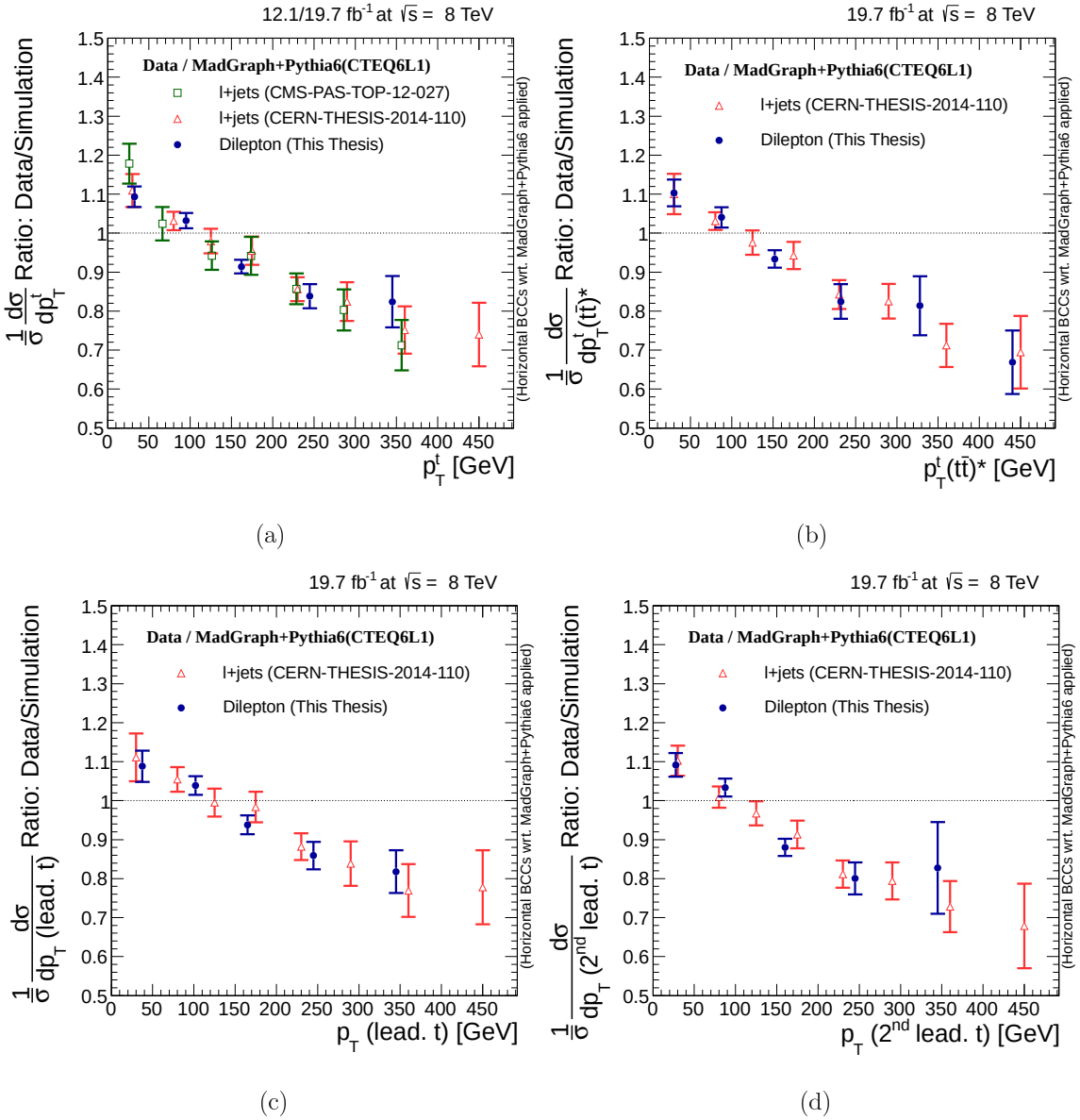


Figure 8.12: Data-to-simulation ratio comparison of the differential cross section result as a function of the top quark transverse momentum in the laboratory rest frame (a) and in the $t\bar{t}$ rest frame (b), transverse momentum of the leading (c) and trailing (d) top quark. The errors bars correspond to the total error on the measurements. The ratio and the BCCs are obtained with respect to the MadGraph+Pythia $t\bar{t}$ simulation.

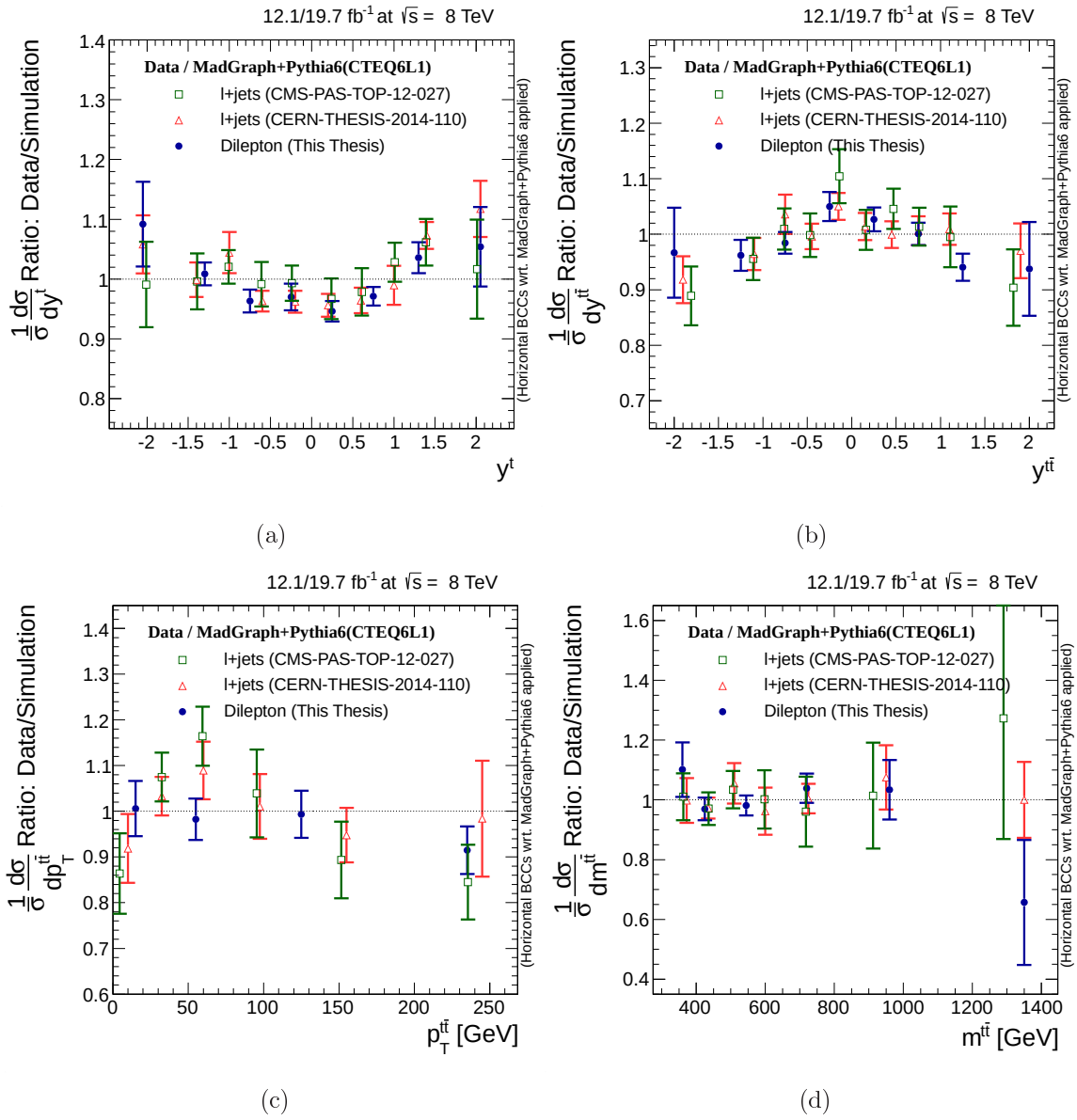


Figure 8.13: Data-to-simulation ratio comparison of the differential cross section result as a function of the top-quark rapidity (a), and of the top-quark-pair rapidity (b), transverse momentum (c) and invariant mass (d). The errors bars correspond to the total error on the measurements. The ratio and the BCCs are obtained with respect to the MadGraph+Pythia $t\bar{t}$ simulation.

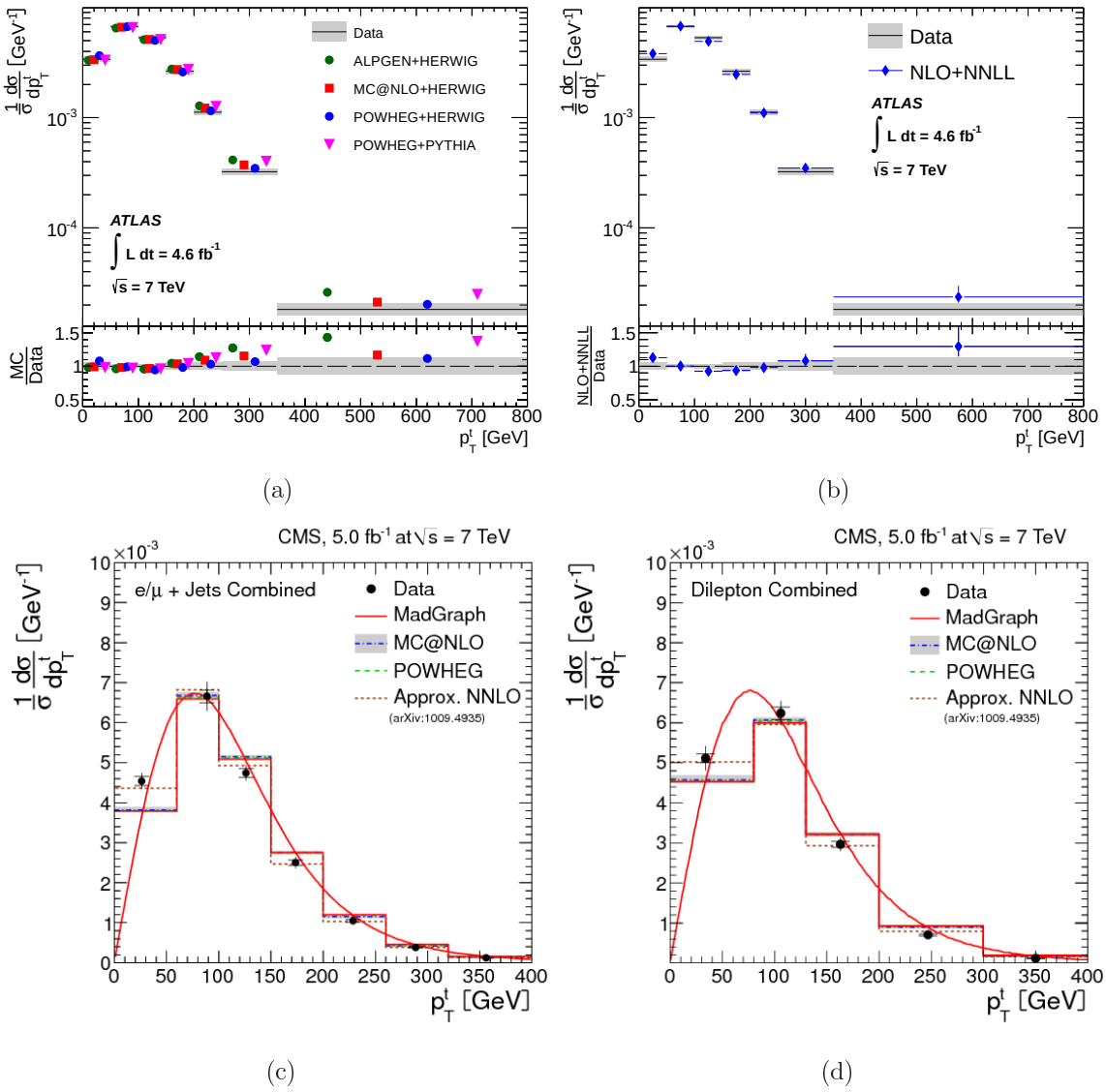


Figure 8.14: Differential cross section results as a function of the top-quark p_{T} obtained by the ATLAS (a and b) and CMS (c) Collaborations in the $\ell+$jets final state, and by the CMS Collaboration in the dilepton decay channel (d). It has to be noticed that both Collaborations denoted differently to the same higher-order QCD perturbation theory [117]: *NLO+NNLL* (in ATLAS) and *Approx. NNLO* (in CMS). All results were derived in pp collisions at a centre-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV}$ [91, 118].

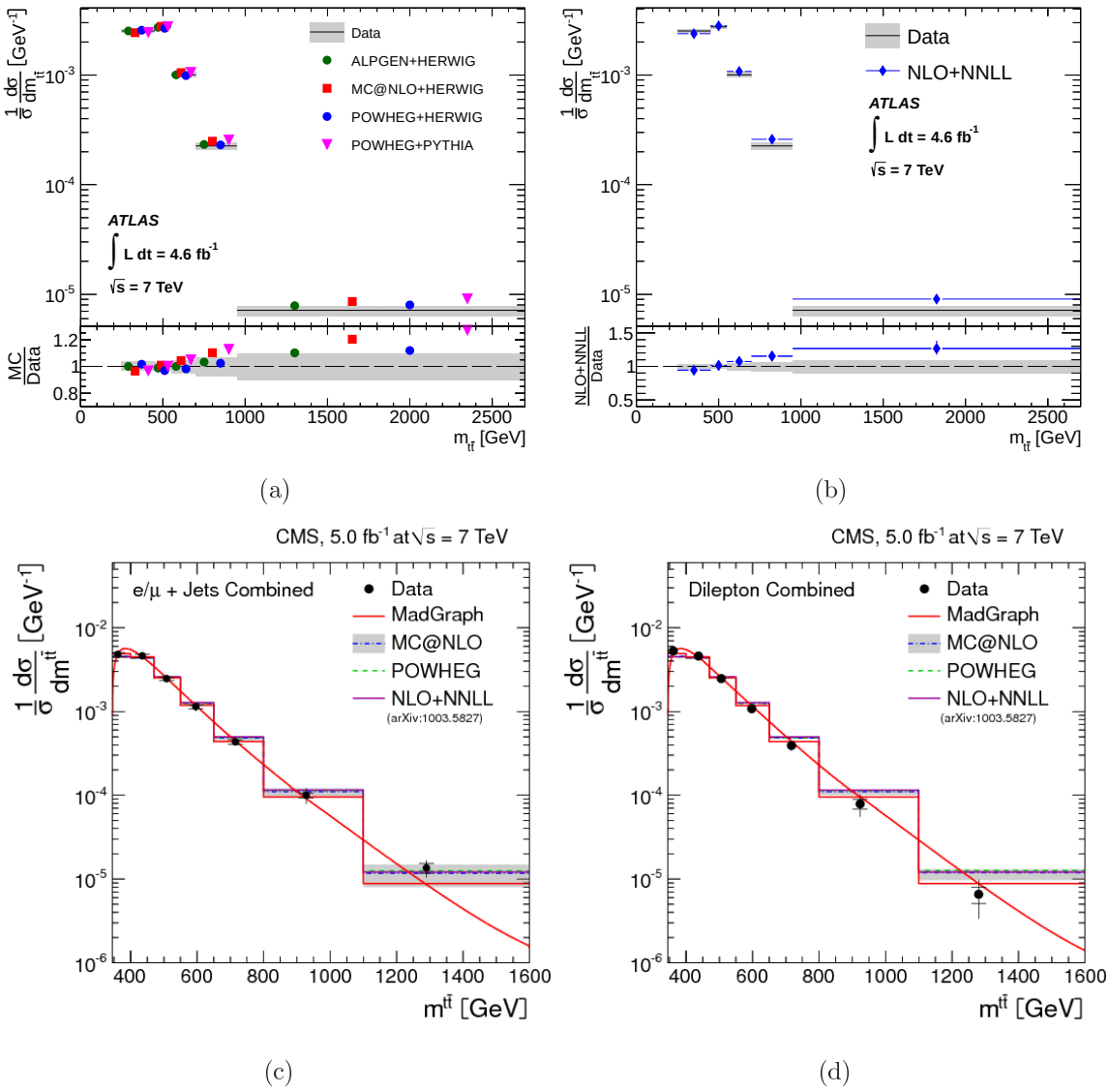
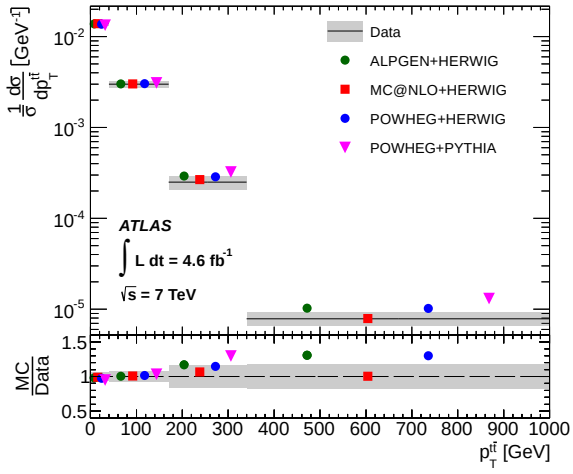
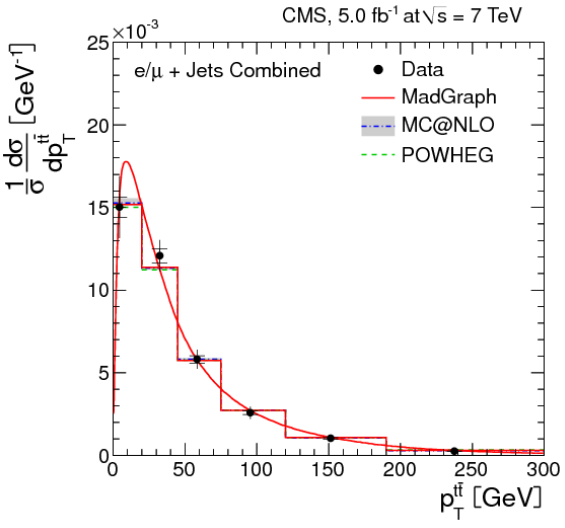


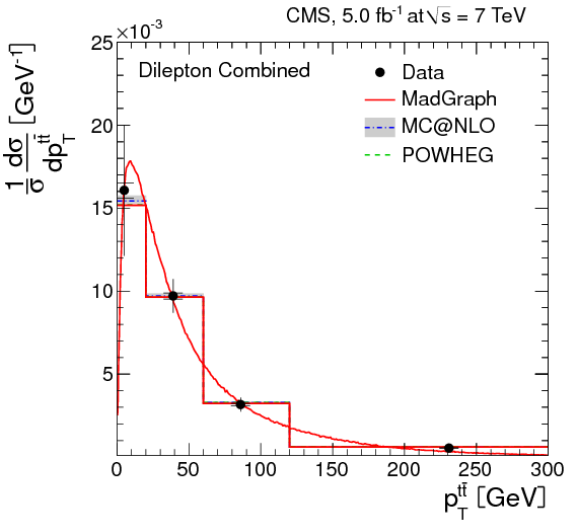
Figure 8.15: Differential cross section results as a function of the top-quark-pair m obtained by the ATLAS (a and b) and CMS (c) Collaborations in the ℓ + jets final state, and by the CMS Collaboration in the dilepton decay channel (d). All results were derived in pp collisions at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ [91, 118].



(a)

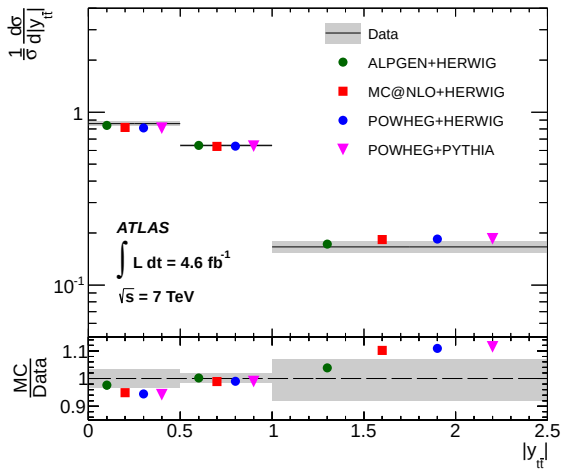


(b)

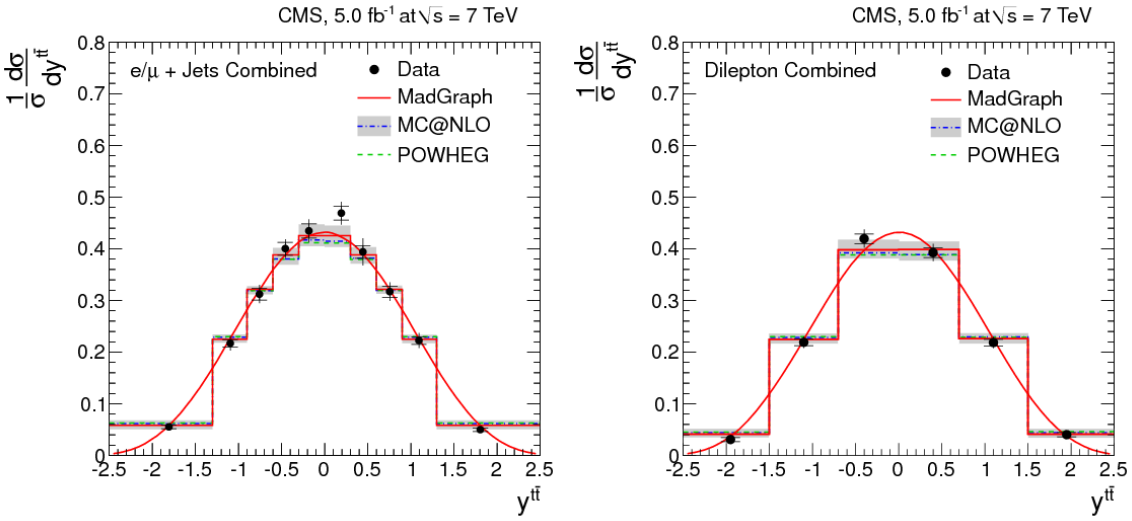


(c)

Figure 8.16: Differential cross section results as a function of the top-quark-pair p_{T} obtained by the ATLAS (a) and CMS (b) Collaborations in the ℓ + jets final state, and by the CMS Collaboration in the dilepton decay channel (c). All results were derived in pp collisions at a centre-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV}$ [91, 118].



(a)



(b) (c)

Figure 8.17: Differential cross section results as a function of the top-quark-pair y obtained by the ATLAS (a) and CMS (b) Collaborations in the ℓ + jets final state, and by the CMS Collaboration in the dilepton decay channel (c). All results were derived in pp collisions at a centre-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ [91, 118].

Channel	ee	eμ	$\mu\mu$	Combined
Data	9964	38404	14353	62721
Non-$t\bar{t}$ Bg.	836.8	2555.8	1265.1	4669.6
$f_{sig.}$ [%]	86.0	85.6	85.7	85.7
Efficiency [%]	33.4	63.7	45.1	51.5
Acceptances [%]	42.6	42.8	42.7	42.8
Branching ratio [%]	1.166	1.166	2.332	4.664
Cross Section [pb]	238.9	244.7	253.7	245.6
Stat. Unc. [pb]	2.6	1.3	2.3	1.1
Syst. Unc. [pb]	14.5	12.8	13.1	13.0
Lumi. Unc. [pb]	6.2	6.4	6.6	6.4
Total Unc. [pb]	16.0	14.4	14.9	14.5

Table 8.3: Number of data and non-$t\bar{t}$ background events, signal fraction, event selection efficiency, acceptance and branching ratio [12] results for the ee, eμ, $\mu\mu$ channels and the combination. The inclusive $t\bar{t}$ cross sections together with the statistical, systematic, luminosity and total uncertainties are also given separately for the individual channels and their combination.

8.6 Total Inclusive Cross Section

The total $t\bar{t}$ inclusive cross section is measured by counting the fraction of $t\bar{t}$ signal events in the selected data sample and correcting them for detector efficiencies and acceptances. Details are given in Section 6.5. The number of selected data events, the expected background events, the signal fraction, the efficiency and the acceptance values are listed in Table 8.3 together with the total cross section results. The uncertainties on the measurement are given separately for the statistical, systematic and luminosity components. The total uncertainty, obtained as the quadratic sum of the mentioned components, is indicated. Further details on the systematic uncertainty calculations are given in Chapter 7.

The total $t\bar{t}$ cross section measurement can be compared to theory predictions and to dedicated experimental results. The most precise theory result is available at NNLO accuracy in QCD perturbation theory and includes the resummation of NNLL soft-gluon terms. It is calculated with Top++2.0 [119] for a top-quark mass of 172.5 GeV, $\mu_r = \mu_F = m^{\mathrm{t}}$ renormalisation and factorisation scale choice and using the MSTW2008 NNLO PDF set [120, 121]. The result is $\sigma_{\mathrm{t\bar{t}}}(m^{\mathrm{t}} = 172.5\,\mathrm{GeV}) = 252.9^{+6.4}_{-8.6}(\mathrm{scales}) \pm 11.7(\mathrm{PDF} + \alpha_s)\,\mathrm{pb}$ [122].

Recently, the first LHC combination of public $t\bar{t}$ cross section measurements in the dilepton channel was obtained [122] using the most precise ATLAS [123] and CMS [92] Collaboration measurements based on $\sqrt{s} = 8\,\mathrm{TeV}$ pp collisions. The combined result is $\sigma_{\mathrm{t\bar{t}}}(\mathrm{LHC}) = 241.5 \pm 1.4\,(\mathrm{stat.}) \pm 5.7\,(\mathrm{syst.}) \pm 6.2\,(\mathrm{lumi.})\,\mathrm{pb}$ assuming a top quark mass of 172.5 GeV.

The result from ATLAS [123] uses a dataset corresponding to an integrated luminosity of $L = 20.3\,\mathrm{fb}^{-1}$. The inclusive cross section and the efficiency to reconstruct and identify a b-jet from a top-quark decay are measured simultaneously. The results is $\sigma_{\mathrm{t\bar{t}}}(\mathrm{ATLAS}) = 242.4 \pm 1.7\,(\mathrm{stat.}) \pm 5.5\,(\mathrm{syst.}) \pm 7.5\,(\mathrm{lumi.})\,\mathrm{pb}$.

The CMS Collaboration uses a $L = 5.3\,\mathrm{fb}^{-1}$ subset of the full pp dataset collected in 2012. The $t\bar{t}$ production cross section is measured with a 5.7% accuracy: $\sigma_{\mathrm{t\bar{t}}}(\mathrm{CMS}) = 239.0 \pm 2.6\,(\mathrm{stat.}) \pm 11.9\,(\mathrm{syst.}) \pm 6.2\,(\mathrm{lumi.})\,\mathrm{pb}$.

Figure 8.18 presents a comparison of the $t\bar{t}$ cross section result obtained in this analysis and

the different experimental public results obtained at $\sqrt{s} = 8\,\text{TeV}$. The comparison includes the individual CMS and ATLAS results described above as well as measurements from the semi-leptonic decay channel and the τ+lepton channel. It can be seen that the measurement obtained in this thesis is in good agreement with all experimental results from the ATLAS and the CMS Collaborations, with the LHC combination and with the most precise theory calculation.

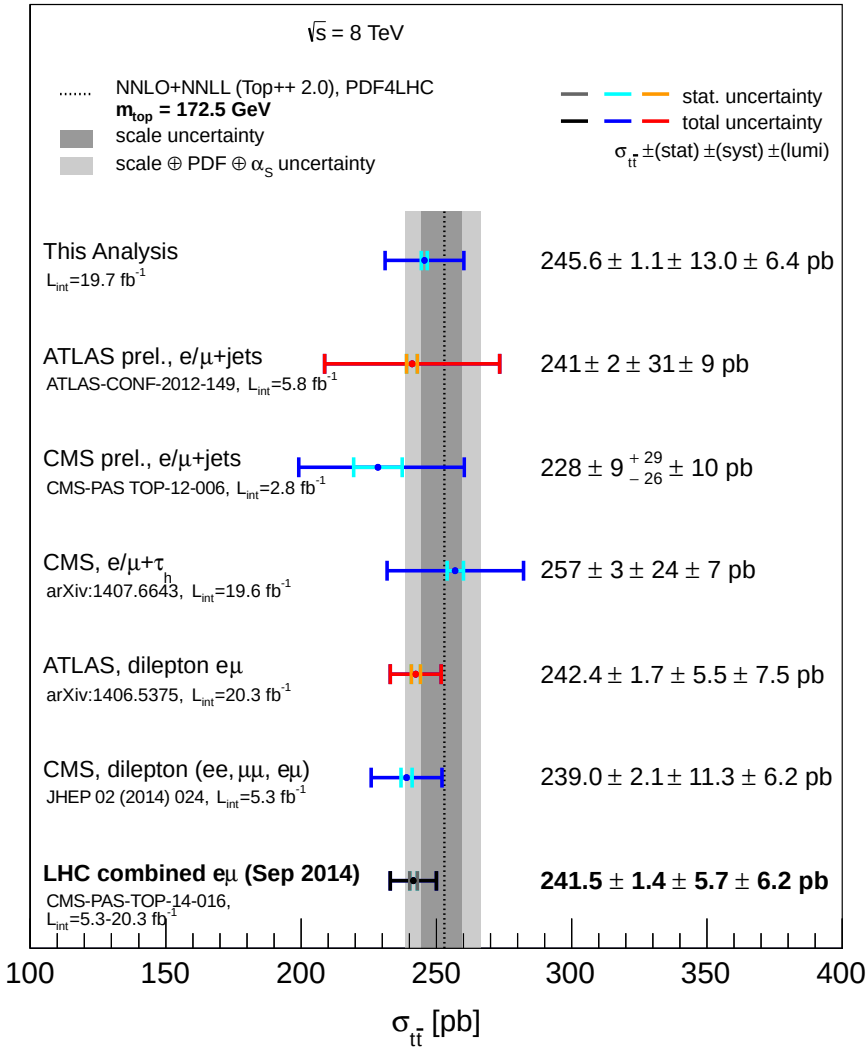


Figure 8.18: Comparison of the $t\bar{t}$ inclusive cross section results derived in this analysis, with the public measurements obtained by the ATLAS and CMS Collaborations, and with the result of the LHC combination. The experimental results indicate separately the statistical, systematic and luminosity uncertainty components. The vertical grey band corresponds to the NNLO+NNLL QCD calculation with Top++2.0 [119] of $\sigma_{t\bar{t}} = 252.9^{+6.4}_{-8.6}(\text{scale}) \pm 11.7(\text{PDF} + \alpha_s)$ pb for a top-quark mass of 172.5 GeV [122].

Chapter 9

Conclusions

The research work presented in this thesis is the first comprehensive measurement of the normalised differential cross section for the $t\bar{t}$ production in the dilepton final state channel. The analysis is based on a proton-proton-collision dataset collected by the CMS experiment at a centre-of-mass energy of $\sqrt{s} = 8\,\text{TeV}$ corresponding to an integrated luminosity of $L = 19.7\,\text{fb}^{-1}$. The excellent quality of the data used in the analysis was ensured by an extensive and meticulous Data Quality Monitoring process which allowed for early detection of incidents during data-taking and in the event reconstruction procedure. A specific quality monitoring process of top-quark events was also implemented as a routine procedure in the CMS Top-Quark-Analysis group.

The $t\bar{t}$ cross section has been measured differentially as a function of 19 kinematic distributions of the top quark, the top-quark pair and the dilepton $t\bar{t}$ decay products, namely the leptons, the lepton pairs, the b-jets and the b-jet pairs. The analysis provides a thorough test of perturbative QCD predictions of the SM. The results have been compared to SM predictions from Monte Carlo event generators which include up to next-to-leading order in perturbative QCD theory. These models are MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig, and MC@NLO+Herwig. Higher order perturbative QCD calculations up to approximate next-to-next-to-leading order accuracy are used to compare the p_{T} and invariant mass of the $t\bar{t}$ system, and the rapidity and p_{T} of the top quarks.

The precision reached in this result allows discriminating between the different theory models. The level of agreement between the results and the theory predictions is quantified by a χ^2-test and corresponding probability value. From these tests, it can be concluded that, in general, the best description of the data is provided by the Powheg+Herwig Monte Carlo prediction, and by the higher-order calculations, depending on the observable.

In general, the angular distribution results are well described by the different models. Small differences between data and the MadGraph+Pythia model can be observed in the $|y^{\text{t}}| - |y^{\bar{\text{t}}}|$ result. They can be associated to the lower accuracy in the perturbative expansion of the theory prediction.

The top-quark-pair transverse momentum result is well described by all Monte Carlo predictions evaluated. The $p_{\text{T}}^{\text{t}\bar{\text{t}}}$ spectrum is not well described by the NLO+NNLL calculation [101] which predicts a distribution softer than data.

The top-quark transverse momentum spectrum is found to be softer in data than most of the predictions. Within the uncertainties, the result is best described by Powheg+Herwig prediction and by the two approximate NNLO calculations [3, 21]. Similar conclusions apply to the top-quark p_{T} measured in the $t\bar{t}$ rest frame or the transverse momentum distributions separately for the highest-p_{T} and second-highest-p_{T} top quark: data is well de-

scribed by the POWHEG+HERWIG model, while the predictions from MADGRAPH+PYTHIA, POWHEG+PYTHIA, and MC@NLO+HERWIG are harder than data.

The differences on data and predictions in the top-quark p_{T} spectrum are consistently measured in the top-quark decay-product distributions. The transverse momentum spectra of the leptons and b-jets in data is generally softer than the predictions from MADGRAPH+PYTHIA, POWHEG+PYTHIA and MC@NLO+HERWIG. Instead, data is well described by the POWHEG+HERWIG model.

Detailed studies have demonstrated that the observed differences in the p_{T}-spectra between data and the predictions are not due to the unfolding model or the assumptions used to correct the migrations caused by the detector neither from experimental sources such as PU effects. The differences are already observed before the event kinematic reconstruction is performed.

The results are found to be consistent with those obtained by the CMS collaboration in the single-lepton $\mathrm{t\bar{t}}$ decay channel at $\sqrt{s} = 8\,\mathrm{TeV}$ [96, 112]. A joint publication of the results from this thesis and from the ℓ + jets top-quark-pair decay channel is in preparation by the CMS Collaboration [124].

Previous differential cross section analysis performed by the CMS and ATLAS collaborations with pp collisions at a centre-of-mass energy $\sqrt{s} = 7\,\mathrm{TeV}$ [91, 118] observed similar differences in the top-quark p_{T} as those presented in this analysis. In particular, the conclusions in the $p_{\mathrm{T}}^{\mathrm{t}} \geq 200\,\mathrm{GeV}$ are similar for all measurements, while slight differences are observed in $p_{\mathrm{T}}^{\mathrm{t}} \leq 200\,\mathrm{GeV}$.

The observed differences between data and predictions, and between the POWHEG+PYTHIA and POWHEG+HERWIG predictions has served as starting point for very fruitful discussions between the experimental analysts of the CMS and the ATLAS collaborations and top-quark theory experts within the Top Quark LHC Working Group [125]. First studies have found possible sources to the differences in the POWHEG+PYTHIA and POWHEG+HERWIG predictions [114].

Theory calculations at NNLO+NNLL order accuracy of the $\mathrm{t\bar{t}}$ production differential cross section as a function of the top-quark p_{T} might provide an explanation for the observed data spectrum. First results have been obtained already as a function of the top-charge asymmetry [126], only. Experimentally, a double differential cross section measurement with a detailed evaluation of the correlations between various event kinematic variables may be able to provide further insights on the origin of the model differences.

Measurements of the Higgs boson properties and beyond SM searches rely on an excellent $\mathrm{t\bar{t}}$ production description. The differential cross section measured as a function of $p_{\mathrm{T}}^{\mathrm{t}}$ is a valuable input for the determination $\mathrm{t\bar{t}}$-modelling corrections which are widely used in the CMS Collaboration [127–129].

In the forthcoming LHC Run-II the $\mathrm{t\bar{t}}$ production rate will be enhanced due to increasing collision centre-of-mass energy. The the top-quark physics programme and particularly the differential measurements will become of major interest for the ultimate precision test of the SM and the Higgs mechanism, as well as searches beyond the SM.

Appendix A

Datasets and Triggers

A.1 Data and Simulation Sets

In the following the different data and simulation sets used in the determination of the normalised differential cross section are detailed.

The proton-proton collision data used in the analysis was collected by the CMS experiment during the year 2012 in four separated time periods according to the LHC running conditions: Run2012A, Run2012B, Run2012C and Run2012D. The collision datasets correspond to the `23Jan2012` re-reconstruction campaign where improved detector calibration constants were used to reconstruct the recorded events and particles in them. The complete list of datasets is given in Table A.1 separated by the HLT trigger used to pre-select the events: double-electron (ee), double-muon ($\mu\mu$) and electron-muon (eμ).

Channel	Dataset Name	Run range	Data Period	Nr. Events
ee	`/DoubleElectron/Run2012A-22Jan2013-v1/AOD`	190456–193621	2012/Apr./05-May/08	13×10^6
	`/DoubleElectron/Run2012B-22Jan2013-v1/AOD`	193834–196531	2012/May/10-Jun./18	23.5×10^6
	`/DoubleElectron/Run2012C-22Jan2013-v1/AOD`	198022–203742	2012/Jul./01-Sep./27	39×10^6
	`/DoubleElectron/Run2012D-22Jan2013-v1/AOD`	203777–208686	2012/Sep./28-Dec./06	34.5×10^6
$\mu\mu$	`/DoubleMu/Run2012A-22Jan2013-v1/AOD`	190456–193621	2012/Apr./05-May/08	5.6×10^6
	`/DoubleMuParked/Run2012B-22Jan2013-v1/AOD`	193834–196531	2012/May/10-Jun./18	29×10^6
	`/DoubleMuParked/Run2012C-22Jan2013-v1/AOD`	198022–203742	2012/Jul./01-Sep./27	37×10^6
	`/DoubleMuParked/Run2012D-22Jan2013-v1/AOD`	203777–208686	2012/Sep./28-Dec./06	38×10^6
eμ	`/MuEG/Run2012A-22Jan2013-v1/AOD`	190456–193621	2012/Apr./05-May/08	2.5×10^6
	`/MuEG/Run2012B-22Jan2013-v1/AOD`	193834–196531	2012/May/10-Jun./18	15×10^6
	`/MuEG/Run2012C-22Jan2013-v1/AOD`	198022–203742	2012/Jul./01-Sep./27	21×10^6
	`/MuEG/Run2012D-22Jan2013-v1/AOD`	203777–208686	2012/Sep./28-Dec./06	22×10^6

Table A.1: List of proton-proton collision dataset names used in this analysis separated by channel. The LHC run numbers, data taking periods and approximate number of events contained in each sample are also indicated.

The event simulation process is detailed in Chapter 4. A list of $\mathrm{t\bar{t}}$ simulation samples is given in Table 4.1 and includes details on the parameters used in the simulation. Several SM background processes are considered and briefly described in Section 4.5 and a summary list with the dataset names is given in Table A.2.

Process	Generator	Sample	Nr. Evts.	σ [pb]
Single top	POWHEG	/Tbar_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola	$0.5 \cdot 10^6$	11.1 (NLO+NNLL) [21]
Single top	POWHEG	/T_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola	$0.5 \cdot 10^6$	11.1 (NLO+NNLL) [21]
$t\bar{t}W$	MADGRAPH	/TTWJets_8TeV-madgraph	$0.2 \cdot 10^6$	0.232 (NLO) [130]
$t\bar{t}Z$	MADGRAPH	/TTZJets_8TeV-madgraph_v2	$0.2 \cdot 10^6$	0.2057 (NLO) [131]
$t\bar{t}\gamma$	MADGRAPH	/TTGJets_8TeV-madgraph	$0.1 \cdot 10^6$	1.8 (LO) [132, 133]
W+jets	MADGRAPH	/WJetsToLNu_TuneZ2Star_8TeV-madgraph-tarball	$18.3 \cdot 10^6$	36257.2 (NNLO) [92]
Drell-Yan	MADGRAPH	/DYJetsToLL_M-10To50filter_8TeV-madgraph	$7.1 \cdot 10^6$	860.5 (NNLO) [92]
Drell-Yan	MADGRAPH	/DYJetsToLL_M-50_TuneZ2Star_8TeV-madgraph-tarball	$30.5 \cdot 10^6$	3503.71 (NNLO) [92]
Diboson	PYTHIA	/WW_TuneZ2star_8TeV_pythia6_tauola	$10.0 \cdot 10^6$	54.838 (NLO) [134]
Diboson	PYTHIA	/WZ_TuneZ2star_8TeV_pythia6_tauola	$10.0 \cdot 10^6$	33.21 (NLO) [134]
Diboson	PYTHIA	/ZZ_TuneZ2star_8TeV_pythia6_tauola	$9.8 \cdot 10^6$	17.654 (NLO) [134]
QCD Multijet	PYTHIA	/QCD_Pt_20_MuEnrichedPt_15_TuneZ2star_8TeV_pythia6	$7.5 \cdot 10^6$	134680.0 (LO) [66]
QCD Multijet	PYTHIA	/QCD_Pt_20_30_EMEnriched_TuneZ2star_8TeV_pythia6	$35.0 \cdot 10^6$	2914860.0 (LO) [66]
QCD Multijet	PYTHIA	/QCD_Pt_30_80_EMEnriched_TuneZ2star_8TeV_pythia6	$33.1 \cdot 10^6$	4615893.0 (LO) [66]
QCD Multijet	PYTHIA	/QCD_Pt_80_170_EMEnriched_TuneZ2star_8TeV_pythia6	$34.5 \cdot 10^6$	183294.9 (LO) [66]
QCD Multijet	PYTHIA	/QCD_Pt_20_30_BCtoE_TuneZ2star_8TeV_pythia6	$1.7 \cdot 10^6$	167388.0 (LO) [66]
QCD Multijet	PYTHIA	/QCD_Pt_30_80_BCtoE_TuneZ2star_8TeV_pythia6	$2.0 \cdot 10^6$	167040.0 (LO) [66]
QCD Multijet	PYTHIA	/QCD_Pt_80_170_BCToE_TuneZ2star_8TeV_pythia6	$1.9 \cdot 10^6$	12981.9 (LO) [66]

Table A.2: Summary of background simulation samples in this analysis. All simulation samples correspond to the *Summer12_DR53X_PU_S10_START53_*/AODSIM* production campaign. All samples use PYTHIA parton showering and hadronisation model interfaced to the hard scattering generator indicated in the table. The approximate number of generated events, theory cross section value and accuracy level used to normalise each simulated sample are also indicated for reference.

A.2 HLT Trigger Paths

In Chapter 5 the complete event reconstruction and selection process is detailed including the trigger decision requirement. An OR decision is used to select the events based on double-lepton based HLT trigger paths which require the presence of an electron-pair, muon-pair or electron-muon-pair. The specific HLT paths are listed in Table A.3 for the different channels and separated for collision data and Monte Carlo simulation events.

Channel	Sample	HLT Path
$\mu\mu$	Data	HLT_Mu17_Mu8_v*
		HLT_Mu17_TkMu8_v*
	Simulation	HLT_Mu17_Mu8_v17
		HLT_Mu17_TkMu8_v10
ee	Data	HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_ Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*
	Simulation	HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_ Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v17
eμ	Data	HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*
		HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v*
	Simulation	HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v7
		HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_v7

Table A.3: Summary of the HLT triggers used in data and simulation events for the three dilepton channels (ee, eμ and $\mu\mu$). The asterisk (*) wildcard is used to represent any version of the HLT menu.

Appendix B

Unfolding: Additional Distributions

B.1 Migrations

The acceptance of the visible phase space defined by the detector was introduced in Section 6.1 as the fraction of total generated signal events in the region with detector coverage. In Section 6.3 the bin-to-bin migrations caused by the detector were quantified in terms of the stability and purity. The differential cross section results are measured using a number of bins and a bin size such that the stability is approximately constant and larger or equal to 50% along the full distribution.

In Figures B.1–B.4 the stability, purity, acceptance (wherever possible) and full analysis *efficiency* distributions are presented for all observables measured in the analysis.

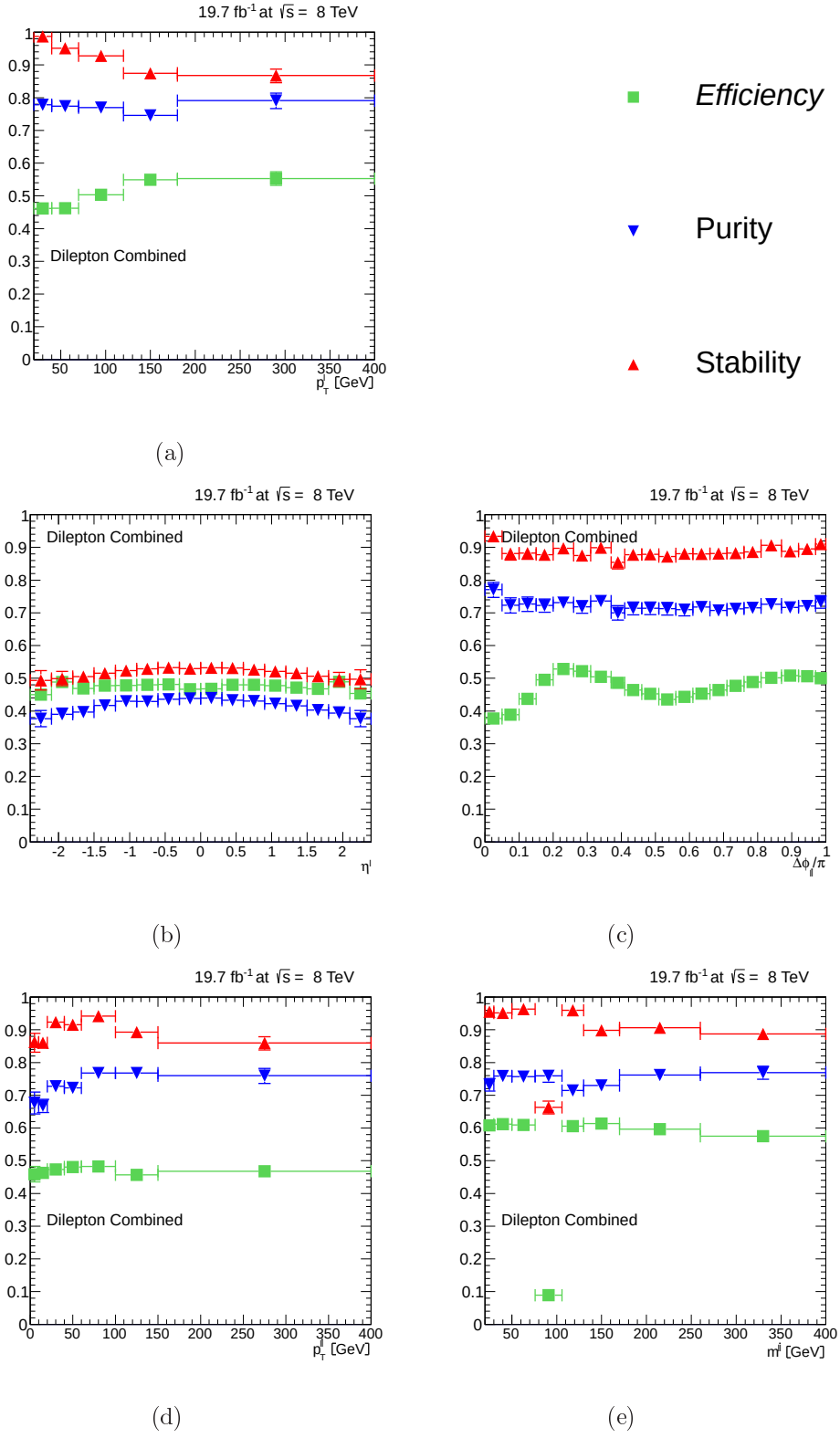
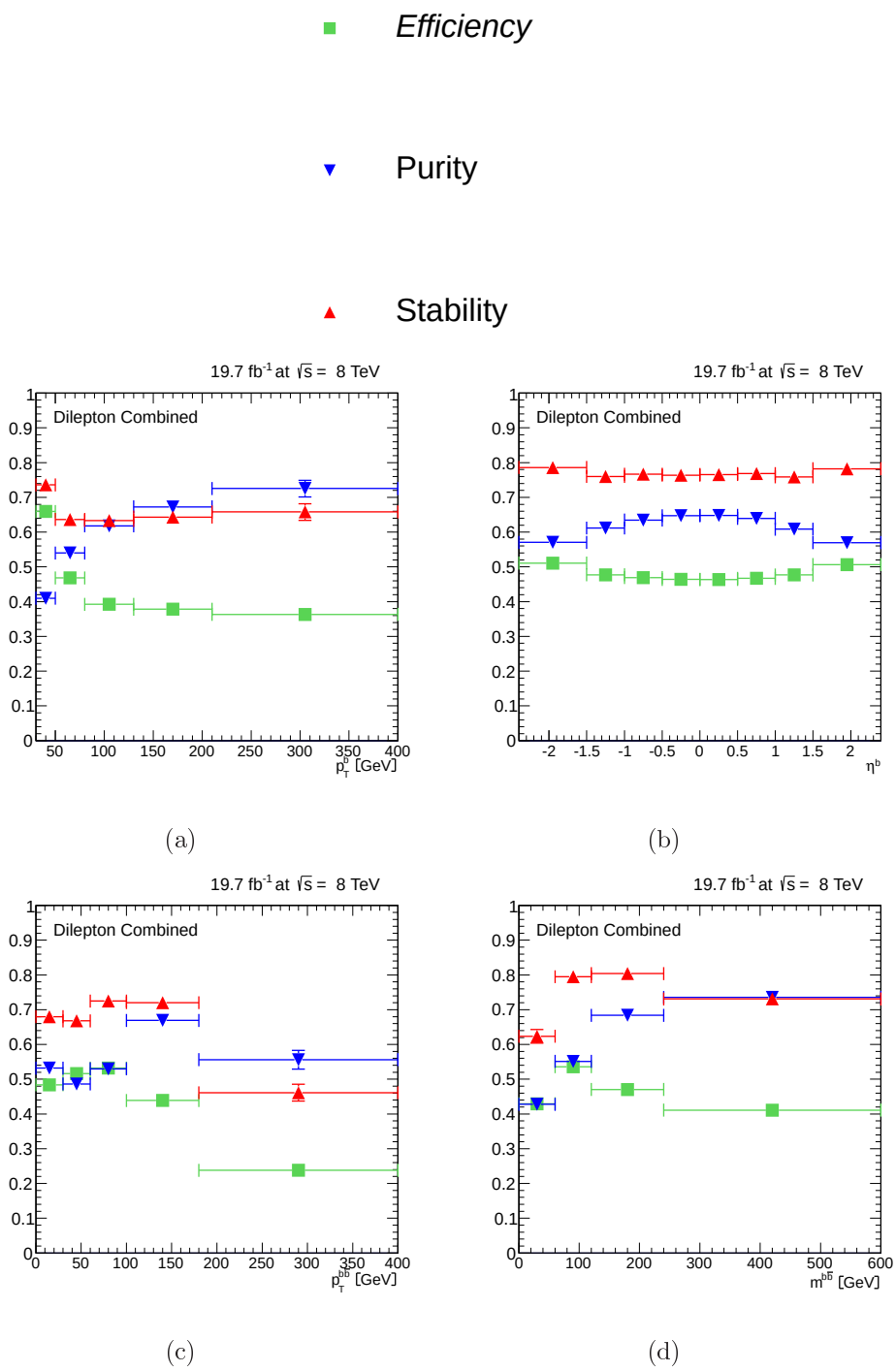


Figure B.1: Stability, purity and efficiency distributions of the lepton transverse momentum (a), pseudorapidity (b), azimuthal angle difference between the leptons (c), lepton-pair transverse momentum (d) and lepton-pair invariant mass (e). The vertical error bars represent the statistical uncertainty of the distributions, and the horizontal bars the bin range.



(a) (b) (c) (d)

Figure B.2: Stability, purity and efficiency distributions of the b-jet transverse momentum (a), pseudorapidity (b), b-jet-pair transverse momentum (c) and invariant mass (d). The vertical error bars represent the statistical uncertainty of the distributions, and the horizontal bars the bin range.

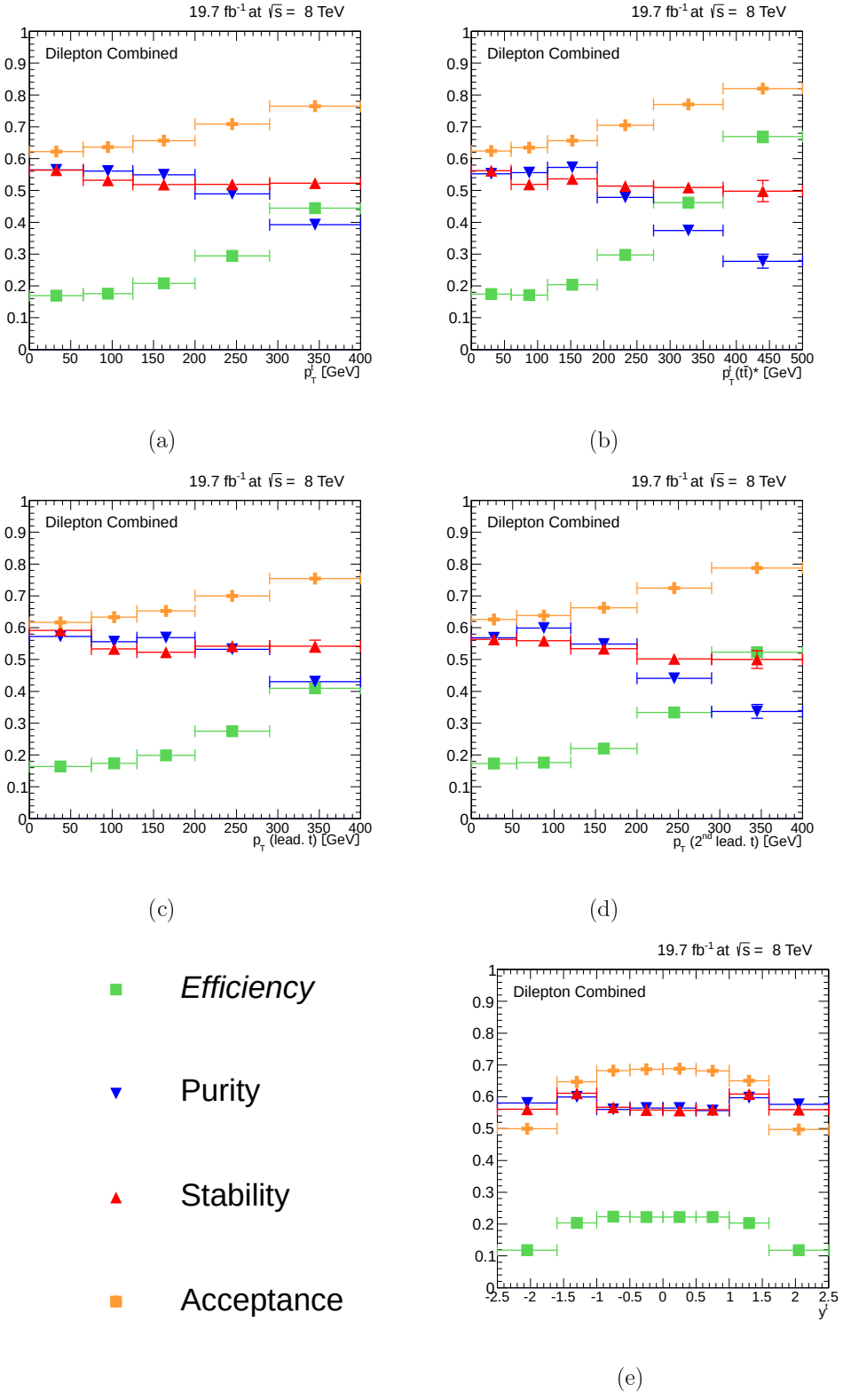


Figure B.3: Stability, purity, acceptance and efficiency distributions of the top quark transverse momentum measured in the laboratory (a) and in the $t\bar{t}$ rest frame (b) reference system, leading and trailing top quark transverse momentum (c) and (d) and top quark rapidity (e). The vertical error bars represent the statistical uncertainty of the distributions, and the horizontal bars the bin range.

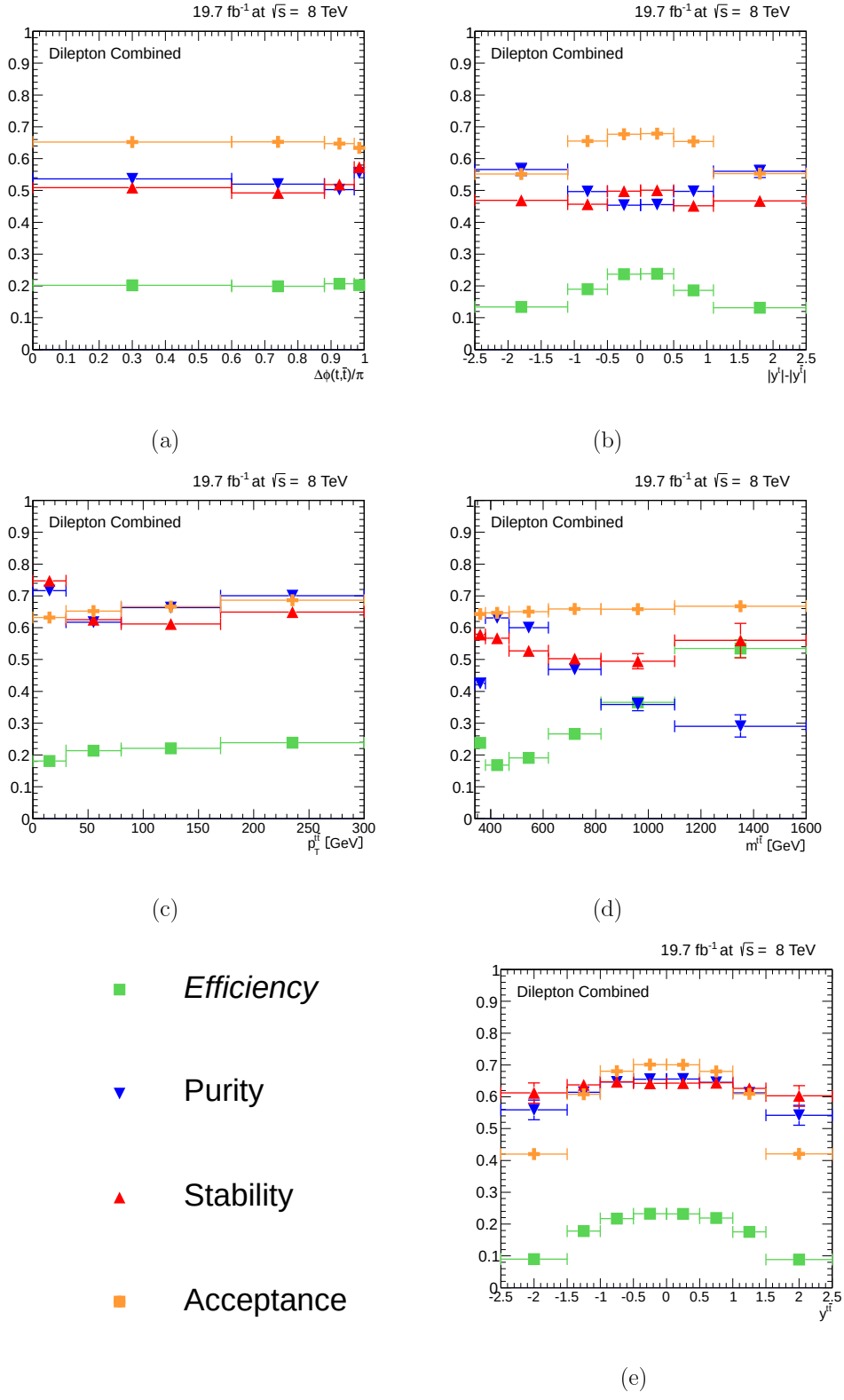


Figure B.4: Stability, purity, acceptance and efficiency distributions of the azimuthal angle difference between the top quarks (a) and rapidity difference between the top quarks (b), top-quark-pair system transverse momentum (c), invariant mass (d) and rapidity (e). The vertical error bars represent the statistical uncertainty of the distributions, and the horizontal bars the bin range.

B.2 Regularisation Strength

The bin-to-bin migration correction is based on a SVD unfolding algorithm which includes a regularisation condition based on a priori knowledge of the solution (see Section 6.4). The relevance of the regularisation term is modulated by a strength parameter τ selected such that the average global correlation $\bar{\rho}$ is minimised. The $\bar{\rho}$ distribution as a function of different regularisation terms is shown for all measured variables in Figures B.5–B.8. The optimal τ-value is indicated numerically and graphically with a red marker.

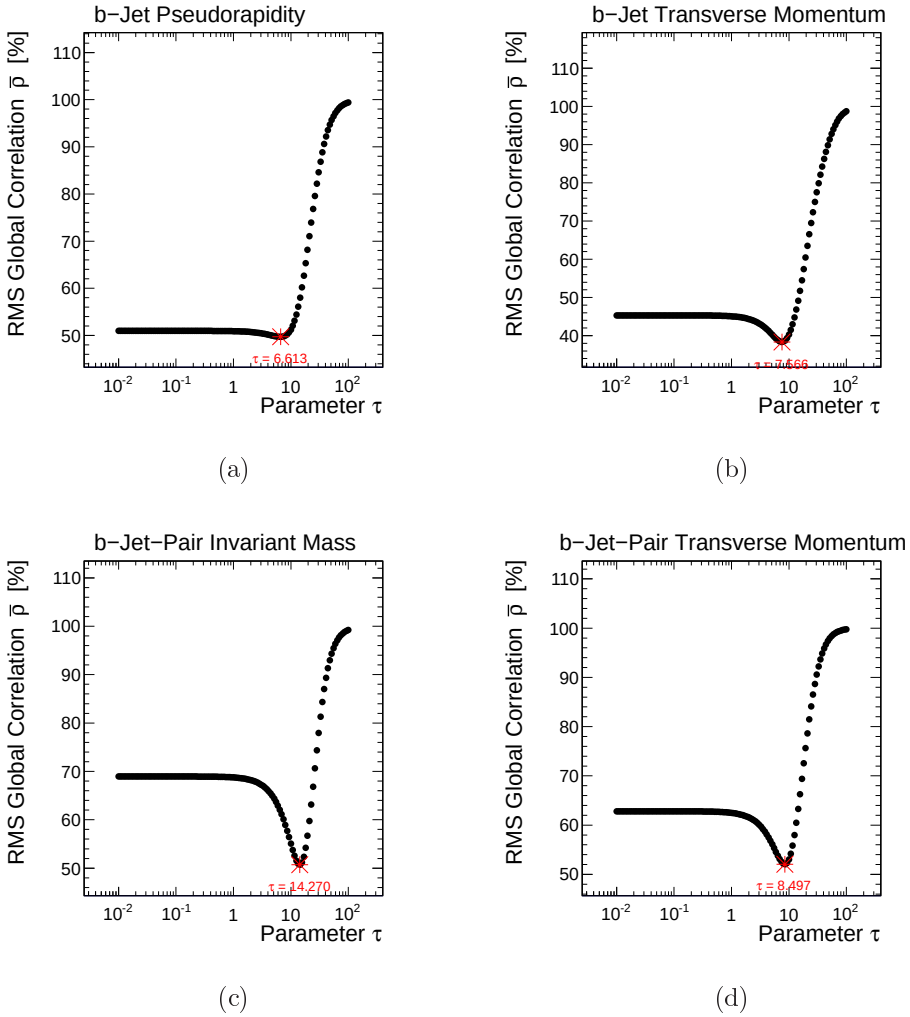
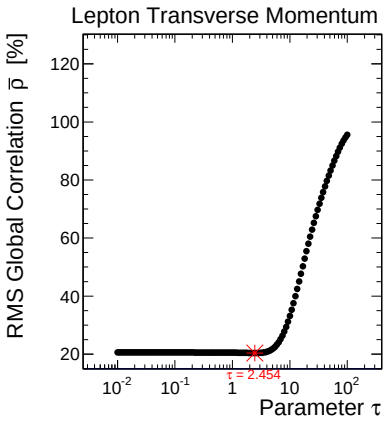
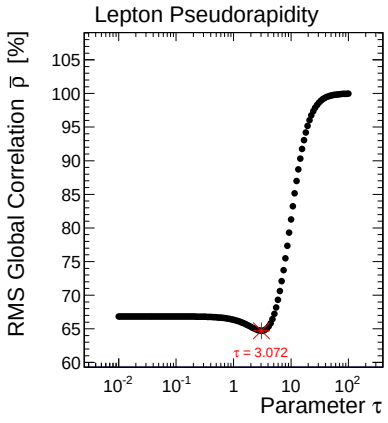


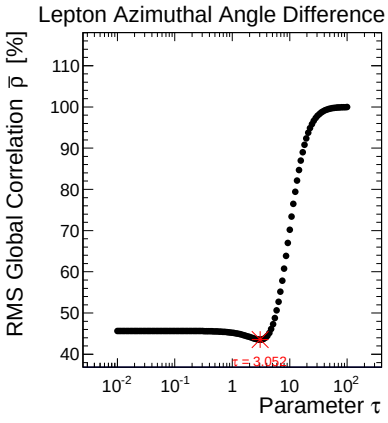
Figure B.5: Average global correlation as a function of the regularisation strength τ for the b-jet transverse momentum (a) and pseudorapidity (b), b-jet-pair transverse momentum (c) and invariant mass (d). The optimal τ-value is indicated in red.



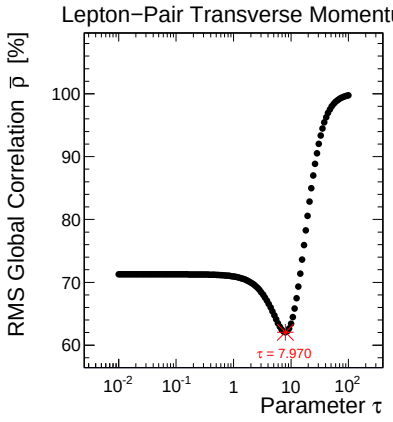
(a)



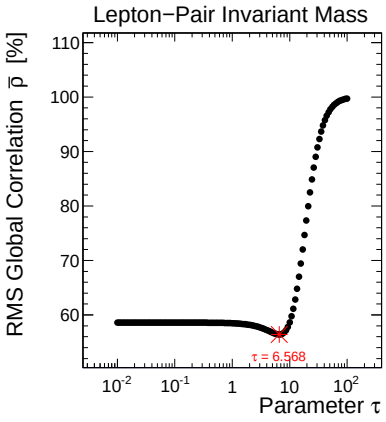
(b)



(c)

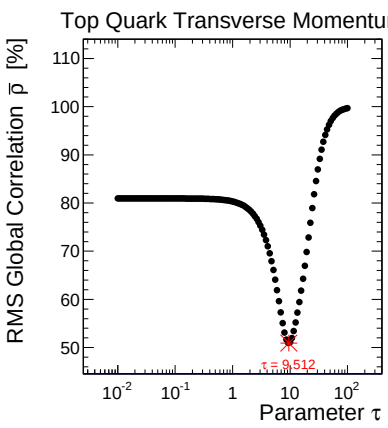


(d)

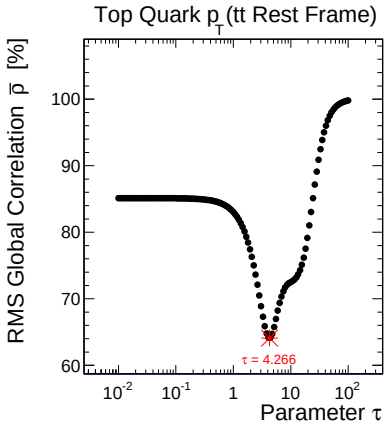


(e)

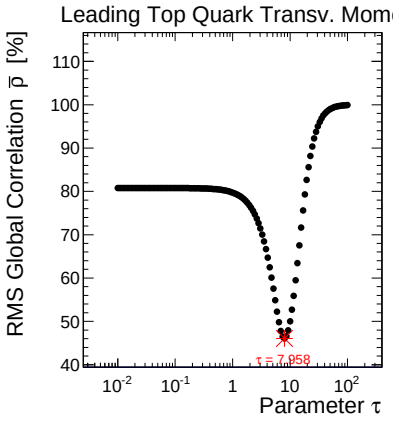
Figure B.6: Average global correlation as a function of the regularisation strength τ for the lepton transverse momentum (a), pseudorapidity (b) and azimuthal angle difference (c), lepton-pair transverse momentum (d) and invariant mass (e). The optimal τ-value is indicated in red.



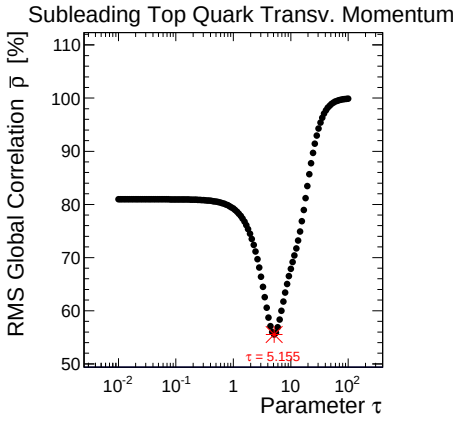
(a)



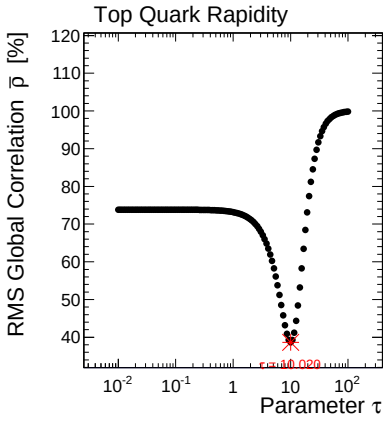
(b)



(c)



(d)



(e)

Figure B.7: . Average global correlation as a function of the regularisation strength τ for the transverse momentum of the top quark measured in the laboratory reference system (a) and in the $t\bar{t}$ rest frame (b), transverse momentum of the leading top quark (c) and trailing top quark (d), and top quark rapidity (e). The optimal τ-value is indicated in red.

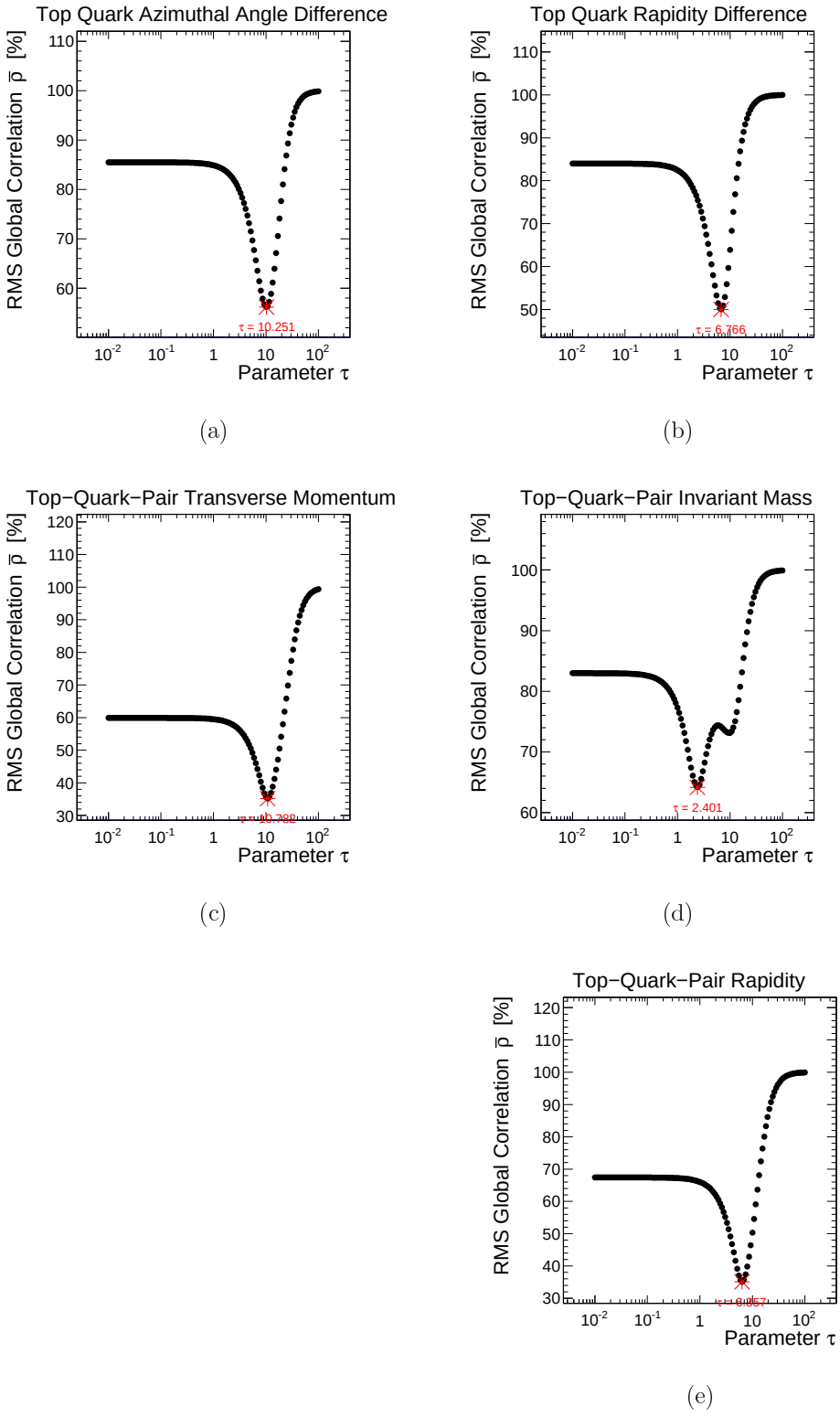


Figure B.8: . Average global correlation as a function of the regularisation strength τ for the difference of the azimuthal angles between the two top quarks (a) and rapidity difference of the top quarks (b), top-quark-pair transverse momentum (c), invariant mass (d) and rapidity (e). The optimal τ-value is indicated in red.

B.3 Consistency Checks

Regularization Strength

Figures B.9–B.12 show the comparison of the differential cross sections obtained with different regularisation strength parameters, indicated in the legend of the figures. $\tau_{\mathrm{Opt.}}$ indicates the optimal regularisation strength used to derive the main result presented in Chapter 8, specific values can be found in Appendix B.2. The theory model used to model the detector effects is indicated by MadGraph+Pythia. No bias due to the unfolding tools is observed as discussed in Section 6.4.1.

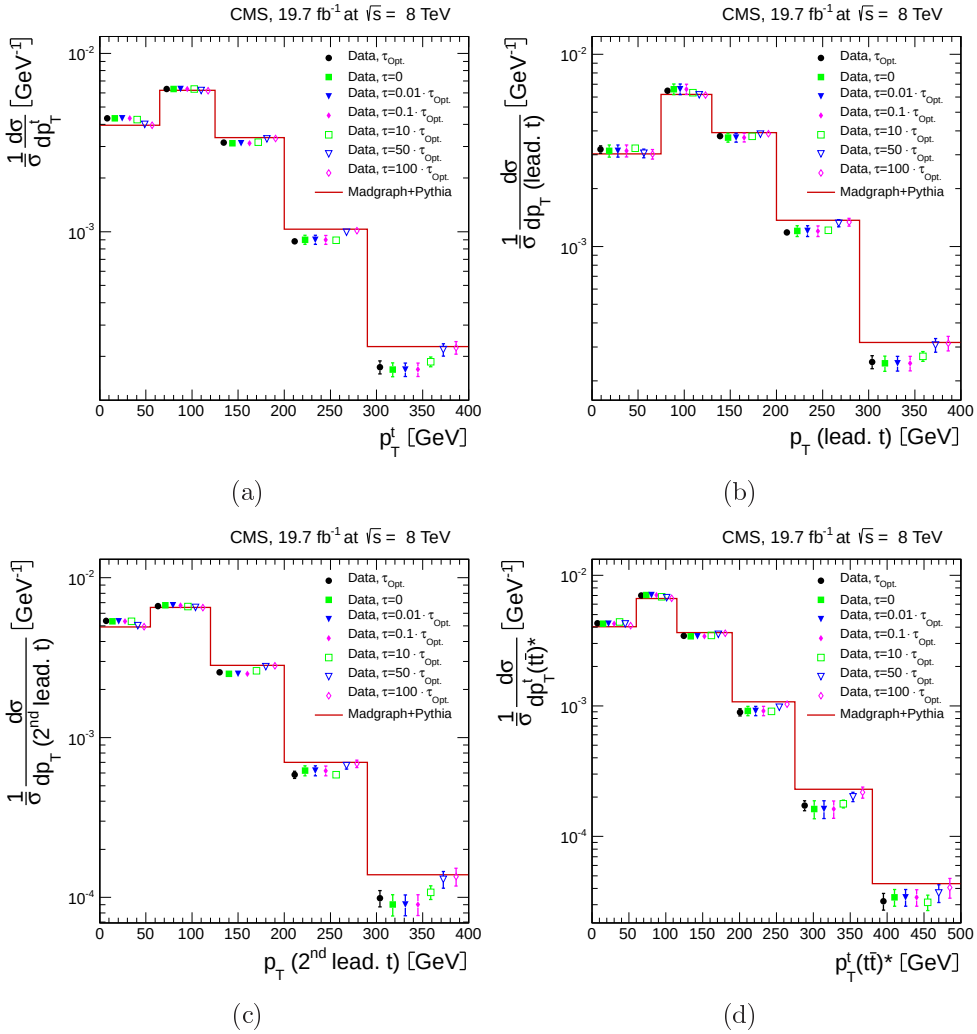
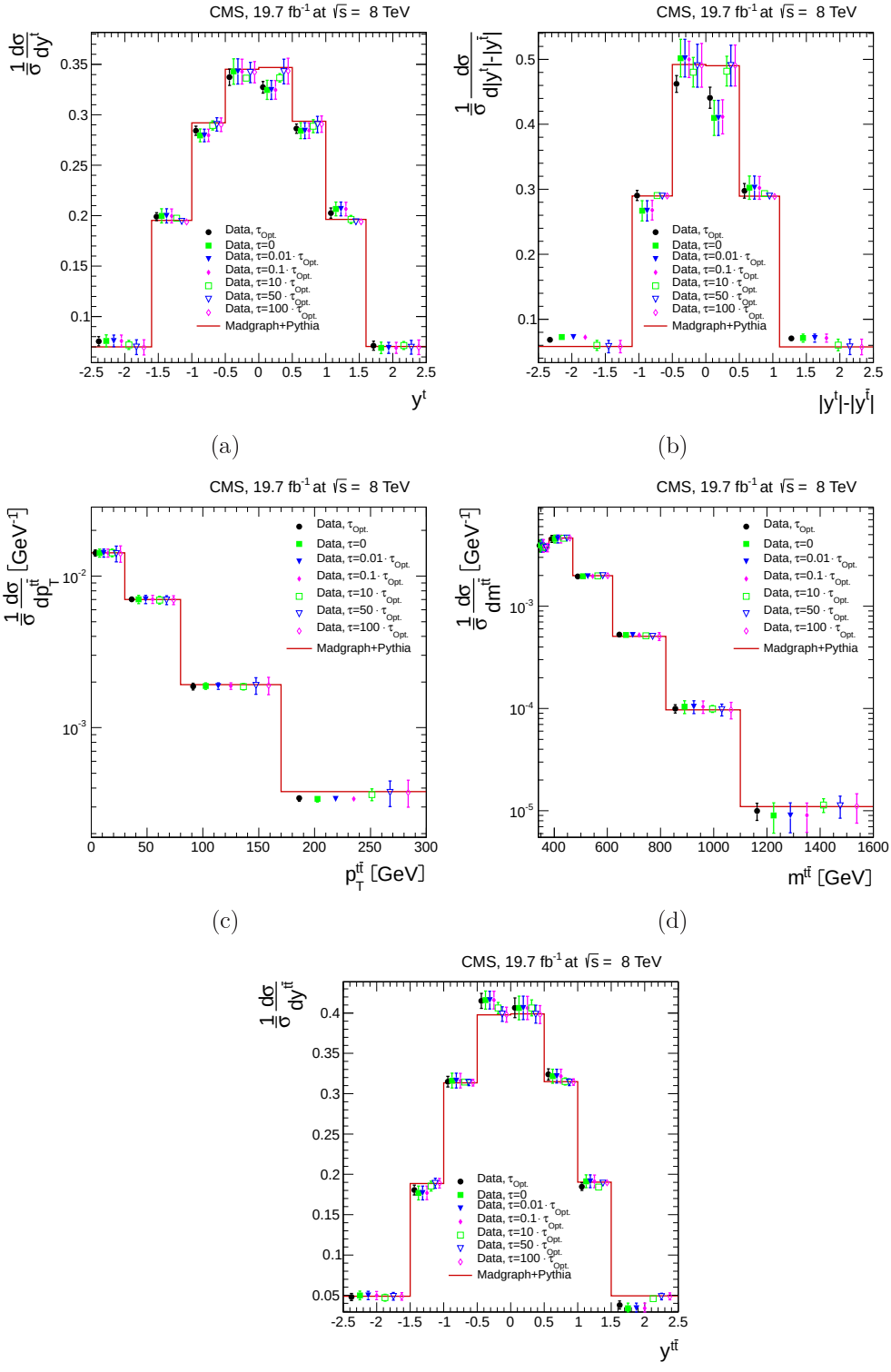
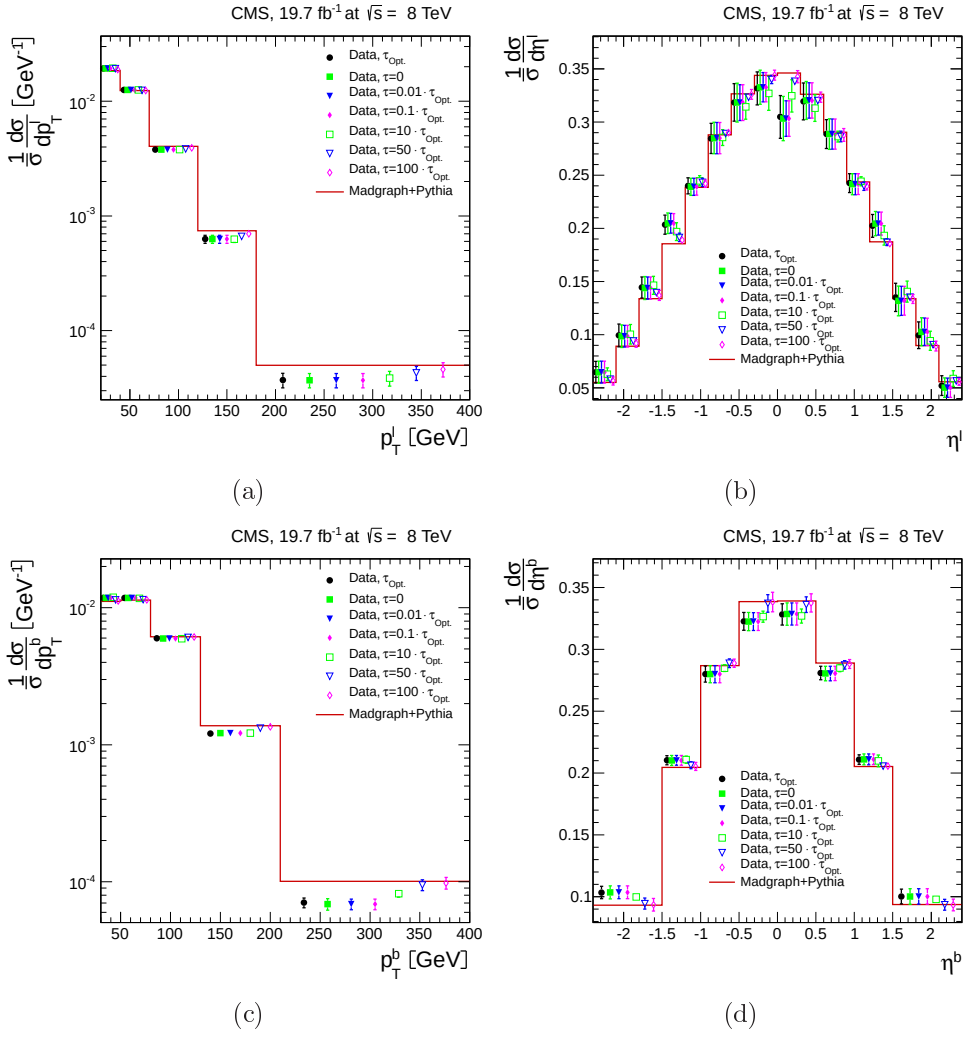


Figure B.9: Differential cross section results as a function of $p_{\mathrm{T}}^{\mathrm{t}}$ (a), $p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{lead.\,t})$ (b), $p_{\mathrm{T}}^{\mathrm{t}}(2^{\mathrm{nd}}\,\mathrm{lead.\,t})$ (c) and $p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{t\bar{t}})^{*}$ (d) for different regularisation strength values indicated in the legend. The Monte Carlo model used to derived the migration matrix describing the detector effects, and used in the unfolding, is indicated by MadGraph+Pythia. The data points are horizontally displaced from the bin centre for graphical purposes only. The vertical error bars on the data markers correspond to the total uncertainty.



(a) (b) (c) (d) (e)

Figure B.10: Differential cross section results as a function of y^{t} (a), $|y^{t}| - |y^{\bar{t}}|$ (b), $p_{T}^{t\bar{t}}$ (c), $m^{t\bar{t}}$ (d) and $y^{t\bar{t}}$ (e) for different regularisation strength values indicated in the legend. The Monte Carlo model used to derived the migration matrix describing the detector effects, and used in the unfolding, is indicated by MadGraph+Pythia. The data points are horizontally displaced from the bin centre for graphical purposes only. The vertical error bars on the data markers correspond to the total uncertainty.



(a) (b) (c) (d)

Figure B.11: Differential cross section results as a function of $p_{\mathrm{T}}^{\mathrm{l}}$ (a), η^{l} (b), $p_{\mathrm{T}}^{\mathrm{b}}$ (d) and η^{b} (d) for different regularisation strength values indicated in the legend. The Monte Carlo model used to derived the migration matrix describing the detector effects, and used in the unfolding, is indicated by MadGraph+Pythia. The data points are horizontally displaced from the bin centre for graphical purposes only. The vertical error bars on the data markers correspond to the total uncertainty.

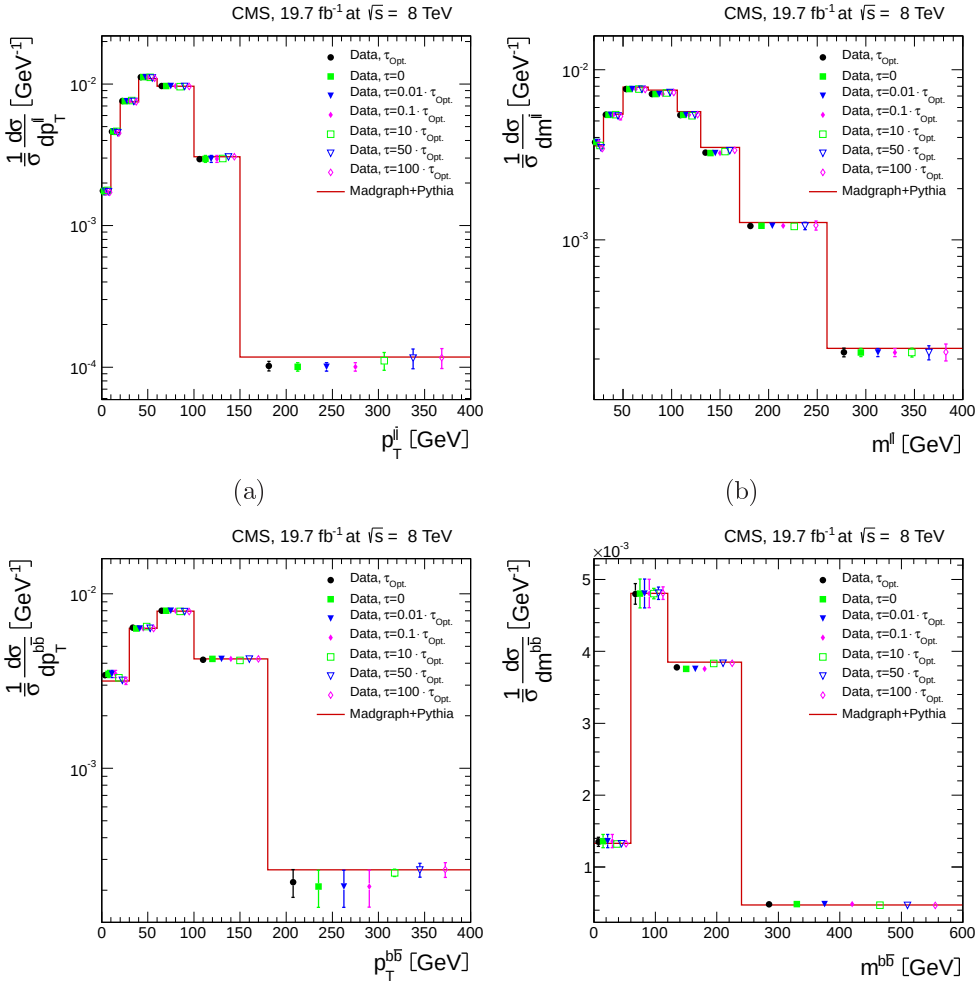


Figure B.12: Differential cross section results as a function of $p_{\mathrm{T}}^{\mathrm{l\bar{l}}}$ (a), $m^{\mathrm{l\bar{l}}}$ (b), $p_{\mathrm{T}}^{\mathrm{b\bar{b}}}$ (c) and $m^{\mathrm{b\bar{b}}}$ (d) for different regularisation strength values indicated in the legend. The Monte Carlo model used to derived the migration matrix describing the detector effects, and used in the unfolding, is indicated by MadGraph+Pythia. The data points are horizontally displaced from the bin centre for graphical purposes only. The vertical error bars on the data markers correspond to the total uncertainty.

Pseudo-data

The cross section results derived from pseudo-data as a function of the top-quark transverse momentum and rapidity are shown in Figures B.13a and B.13b, respectively. The figures show also the original pseudo-data distribution (*Simu. Reweighted*) and the shape of the theory prediction used to model the detector effects (MadGraph+Pythia). For reference a higher-order QCD theory calculation is also shown. The results derived from pseudo-data are in good agreement with the original distribution.

Differential cross sections derived from pseudo-data with varied $p_{\mathrm{T}}^{\mathrm{t}}$ and y^{t} distributions are presented in Figures B.13c, B.13d, B.13e and B.13f. The original pseudo-data shape is indicated by *Simu. Reweighted*. The distribution MadGraph+Pythia indicates the theory prediction used to model the detector effects. Higher-order theory prediction is indicated for reference, wherever available.

No bias due to the unfolding procedure is observed.

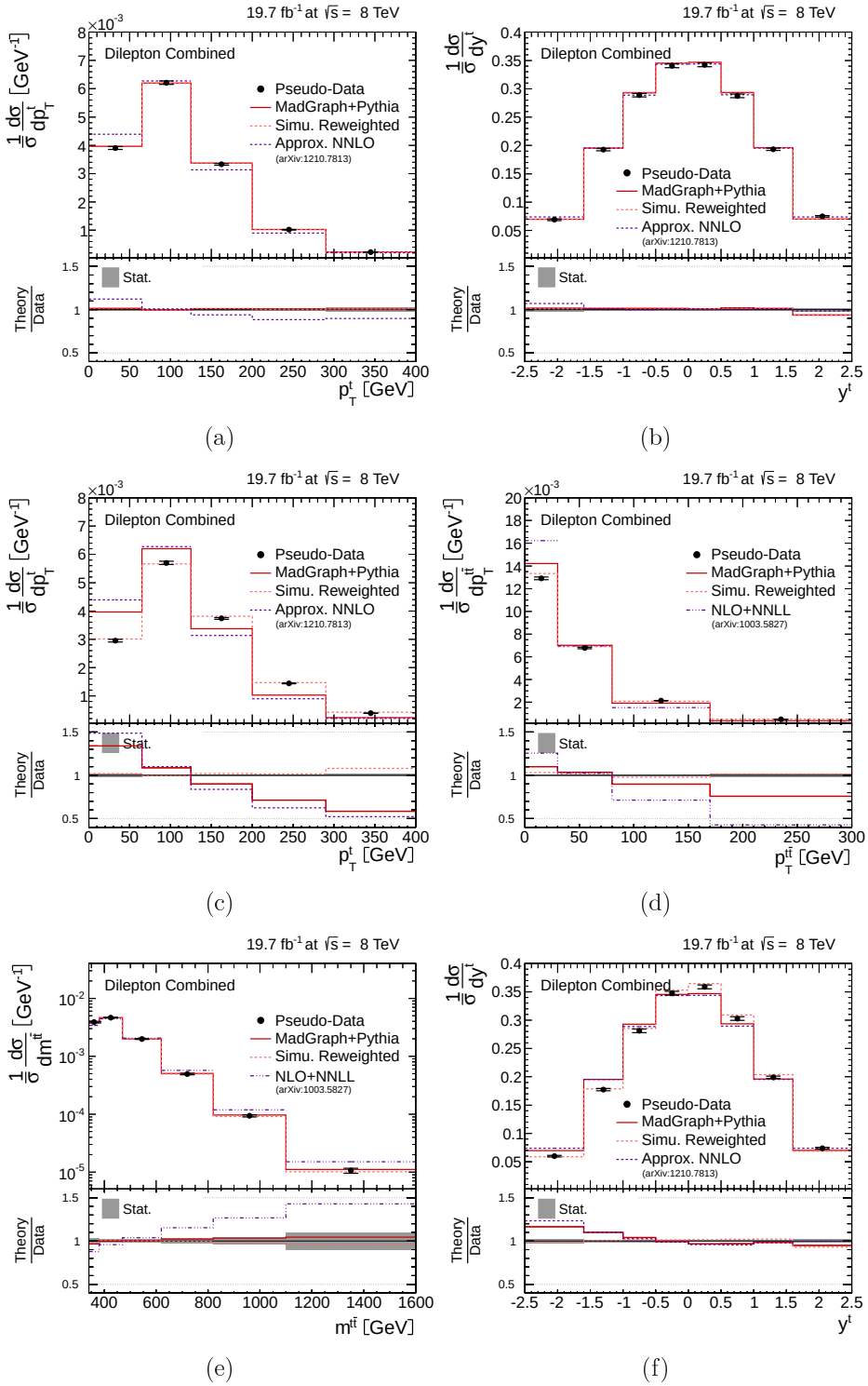


Figure B.13: Differential cross section results as a function of the $p_{\mathrm{T}}^{\mathrm{t}}$, y^{t}, $p_{\mathrm{T}}^{\mathrm{t\bar{t}}}$ and $m^{\mathrm{t\bar{t}}}$ obtained from pseudo-data events derived from MadGraph+Pythia $\mathrm{t\bar{t}}$ simulation without (a and b) and with variations on the shape of the distribution based on $p_{\mathrm{T}}^{\mathrm{t}}(s = 0.002)$ (c), $p_{\mathrm{T}}^{\mathrm{t}}(s = 0.004)$ (d), $y^{\mathrm{t}}(s = -0.160)$ (e) and $y^{\mathrm{t}}(s = -0.080)$ (f). The distribution used to derived the unfolding migration matrix is indicated by MadGraph+Pythia and the original shape of the pseudo-data is indicated by *Simu. Reweighted*. Additionally higher-order QCD theory calculation are shown for reference: NLO+NNLL [100, 101] and Approx. NNLO [21].

Appendix C

Results: Tables and Indidivual Channels

C.1 Result Tables

In the following the numerical values of the cross section results presented in Chapter 8 are given. The Tables C.1–C.9 include the bin range, measured normalised differential cross section in the bin, relative statistical, systematic and total uncertainty measured in the bin indicated. In Chapter 7 the sources and the methodology to obtain the systematic uncertainties are detailed.

	Results	Uncertainties [%]		
Bin range [GeV]	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{b}}}$ [GeV^{-1}]	Stat.	Syst.	Total
30 to 50	1.18e-02	0.9	5.0	5.1
50 to 80	1.18e-02	0.9	2.7	2.8
80 to 130	5.99e-03	0.9	1.9	2.1
130 to 210	1.21e-03	1.4	3.0	3.3
210 to 400	7.05e-05	4.0	7.1	8.2
Bin range	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}\eta^{\mathrm{b}}}$	Stat.	Syst.	Total
−2.40 to −1.50	1.03e-01	1.6	4.6	4.9
−1.50 to −1.00	2.10e-01	1.4	1.1	1.8
−1.00 to −0.50	2.80e-01	1.2	2.0	2.3
−0.50 to 0.00	3.23e-01	1.1	1.9	2.2
0.00 to 0.50	3.28e-01	1.1	2.4	2.6
0.50 to 1.00	2.81e-01	1.2	1.5	1.9
1.00 to 1.50	2.11e-01	1.4	1.2	1.8
1.50 to 2.40	1.00e-01	1.7	5.7	6.0

Table C.1: Normalised differential cross sections and uncertainties for the dilepton channel in the visible phase space as a function of the b-jet transverse momentum ($p_{\mathrm{T}}^{\mathrm{b}}$) and pseudo-rapidity ($\eta^{\mathrm{b}}$). The uncertainties are given separated for the statistical and total systematic component as well as the quadratic sum of them (total).

	Results	Uncertainties [%]		
Bin range [GeV]	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{b\bar{b}}}}$ [GeV^{-1}]	Stat.	Syst.	Total
0 to 30	3.42e-03	2.1	3.4	4.0
30 to 60	6.40e-03	1.4	3.5	3.7
60 to 100	8.00e-03	1.1	2.9	3.1
100 to 180	4.19e-03	0.9	3.1	3.2
180 to 400	2.22e-04	3.9	17.7	18.1
Bin range [GeV]	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}m^{\mathrm{b\bar{b}}}}$[GeV^{-1}]	Stat.	Syst.	Total
0 to 60	1.35e-03	1.8	4.3	4.7
60 to 120	4.80e-03	0.8	2.9	3.0
120 to 240	3.78e-03	0.6	0.9	1.1
240 to 600	4.81e-04	1.3	4.8	5.0

Table C.2: Normalised differential cross sections and uncertainties for the dilepton channel in the visible phase space as a function of the b-jet-pair transverse momentum ($p_{\mathrm{T}}^{\mathrm{b}}$) and invariant mass ($m^{\mathrm{b\bar{b}}}$). The uncertainties are given separated for the statistical and total systematic component as well as the quadratic sum of them (total).

	Results	Uncertainties [%]		
Bin range [GeV]	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{l}}}$ [GeV^{-1}]	Stat.	Syst.	Total
20 to 40	1.94e-02	0.5	1.6	1.6
40 to 70	1.26e-02	0.5	0.7	0.9
70 to 120	3.78e-03	0.8	1.8	1.9
120 to 180	6.28e-04	1.9	7.7	8.0
180 to 400	3.70e-05	4.5	13.8	14.5

Table C.3: Normalised differential cross sections and uncertainties for the dilepton channel in the visible phase space as a function of the lepton transverse momentum ($p_{\mathrm{T}}^{\mathrm{l}}$). The uncertainties are given separated for the statistical and total systematic component as well as the quadratic sum of them (total).

	Results	Uncertainties [%]		
Bin range	$\frac{1}{\sigma}\frac{d\sigma}{d\eta^l}$	Stat.	Syst.	Total
−2.40 to −2.10	6.48e-02	5.2	14.9	15.8
−2.10 to −1.80	9.93e-02	4.0	10.0	10.8
−1.80 to −1.50	1.44e-01	3.4	5.9	6.8
−1.50 to −1.20	2.03e-01	2.9	3.3	4.4
−1.20 to −0.90	2.40e-01	2.7	1.6	3.2
−0.90 to −0.60	2.84e-01	2.4	4.4	5.0
−0.60 to −0.30	3.18e-01	2.3	5.0	5.5
−0.30 to 0.00	3.32e-01	2.3	4.1	4.7
0.00 to 0.30	3.05e-01	2.5	6.1	6.6
0.30 to 0.60	3.19e-01	2.4	4.9	5.4
0.60 to 0.90	2.89e-01	2.5	4.0	4.7
0.90 to 1.20	2.43e-01	2.7	2.4	3.6
1.20 to 1.50	2.02e-01	2.9	4.4	5.3
1.50 to 1.80	1.35e-01	3.8	9.1	9.9
1.80 to 2.10	9.95e-02	4.2	12.0	12.7
2.10 to 2.40	5.22e-02	7.0	16.3	17.7

Table C.4: Normalised differential cross sections and uncertainties for the dilepton channel in the visible phase space as a function of the lepton pseudorapidity (η^l). The uncertainties are given separated for the statistical and total systematic component as well as the quadratic sum of them (total).

	Results	Uncertainties [%]		
Bin range/π	$\frac{1}{\sigma}\frac{d\sigma}{d\Delta\phi^{l\bar{l}}/\pi}$	Stat.	Syst.	Total
0.00 to 0.05	7.53e-01	3.4	6.8	7.6
0.05 to 0.10	7.49e-01	3.1	3.1	4.3
0.10 to 0.15	8.10e-01	2.7	2.5	3.7
0.15 to 0.20	8.63e-01	2.5	1.8	3.1
0.20 to 0.26	8.11e-01	2.3	1.0	2.5
0.26 to 0.31	8.60e-01	2.5	1.8	3.1
0.31 to 0.37	9.07e-01	2.3	2.7	3.6
0.37 to 0.41	8.70e-01	3.0	1.8	3.5
0.41 to 0.46	9.56e-01	2.5	1.8	3.1
0.46 to 0.51	9.73e-01	2.6	3.0	4.0
0.51 to 0.56	9.80e-01	2.5	2.7	3.7
0.56 to 0.61	1.03	2.5	2.6	3.6
0.61 to 0.66	1.02	2.5	3.1	4.0
0.66 to 0.71	1.09	2.4	2.0	3.1
0.71 to 0.76	1.22	2.2	1.2	2.5
0.76 to 0.81	1.18	2.3	1.5	2.7
0.81 to 0.87	1.19	2.0	1.5	2.5
0.87 to 0.92	1.23	2.2	3.2	3.9
0.92 to 0.97	1.29	2.2	2.7	3.5
0.97 to 1.00	1.32	2.8	1.9	3.4

Table C.5: Normalised differential cross sections and uncertainties for the dilepton channel in the visible phase space as a function of the azimuthal angle difference of the leptons ($\Delta\phi^{l\bar{l}}/\pi$). The uncertainties are given separated for the statistical and total systematic component as well as the quadratic sum of them (total).

	Results	Uncertainties [%]		
Bin range [GeV]	$\frac{1}{\sigma}\frac{d\sigma}{dp_T^{l\bar{l}}}$ [GeV^{-1}]	Stat.	Syst.	Total
0 to 10	1.76e-03	3.7	2.8	4.6
10 to 20	4.62e-03	2.0	3.2	3.8
20 to 40	7.59e-03	1.2	2.8	3.1
40 to 60	1.12e-02	1.0	2.1	2.3
60 to 100	9.67e-03	0.7	0.8	1.1
100 to 150	2.95e-03	1.2	4.9	5.1
150 to 400	1.02e-04	3.3	7.0	7.7
Bin range [GeV]	$\frac{1}{\sigma}\frac{d\sigma}{dm^{l\bar{l}}}$[GeV^{-1}]	Stat.	Syst.	Total
20 to 30	3.75e-03	2.5	2.9	3.8
30 to 50	5.44e-03	1.4	1.9	2.3
50 to 76	7.69e-03	1.0	0.9	1.4
76 to 106	7.20e-03	1.0	2.0	2.2
106 to 130	5.41e-03	1.3	1.3	1.9
130 to 170	3.25e-03	1.4	1.3	1.9
170 to 260	1.21e-03	1.5	2.6	3.0
260 to 400	2.19e-04	2.6	5.2	5.8

Table C.6: Normalised differential cross sections and uncertainties for the dilepton channel in the visible phase space as a function of the lepton-pair transverse momentum ($p_T^{l\bar{l}}$) and invariant mass ($m^{l\bar{l}}$). The uncertainties are given separated for the statistical and total systematic component as well as the quadratic sum of them (total).

	Results	Uncertainties [%]		
Bin range [GeV]	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{t}}}$ [GeV^{-1}]	Stat.	Syst.	Total
0 to 65	4.32e-03	1.1	2.6	2.8
65 to 125	6.31e-03	0.9	1.7	1.9
125 to 200	3.15e-03	1.0	2.4	2.6
200 to 290	8.84e-04	1.3	3.3	3.5
290 to 400	1.74e-04	2.9	7.8	8.4
Bin range [GeV]	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{t\bar{t}})^{*}}$ [GeV^{-1}]	Stat.	Syst.	Total
0 to 60	4.29e-03	1.6	2.9	3.3
60 to 115	7.00e-03	1.3	2.1	2.5
115 to 190	3.44e-03	1.3	2.0	2.4
190 to 275	8.95e-04	2.1	5.9	6.3
275 to 380	1.73e-04	2.9	8.3	8.8
380 to 500	3.19e-05	6.9	12.8	14.6
Bin range [GeV]	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{lead.\,t})}$ [GeV^{-1}]	Stat.	Syst.	Total
0 to 75	3.20e-03	1.6	4.1	4.4
75 to 130	6.46e-03	1.1	2.4	2.7
130 to 200	3.75e-03	1.2	2.3	2.6
200 to 290	1.18e-03	1.5	2.8	3.1
290 to 400	2.51e-04	3.3	7.0	7.7
Bin range [GeV]	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{t}}(2^{\mathrm{nd}}\,\mathrm{lead.\,t})}$ [GeV^{-1}]	Stat.	Syst.	Total
0 to 55	5.38e-03	1.7	1.6	2.4
55 to 120	6.65e-03	1.4	1.6	2.1
120 to 200	2.56e-03	1.7	2.4	2.9
200 to 290	5.86e-04	2.3	4.5	5.1
290 to 400	9.88e-05	5.8	10.1	11.7

Table C.7: Normalised differential cross section values and uncertainties for the dilepton channel in the full phase space as a function of the top quark transverse momentum measured in the laboratory ($p_{\mathrm{T}}^{\mathrm{t}}$) and in the $\mathrm{t\bar{t}}$ rest frame ($p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{t\bar{t}})^{*}$), leading ($p_{\mathrm{T}}^{\mathrm{t}}(\mathrm{lead.\,t})$) and second leading top-quark transverse momentum ($p_{\mathrm{T}}^{\mathrm{t}}(2^{\mathrm{nd}}\,\mathrm{lead.\,t})$). The uncertainties are given separated for the statistical and total systematic component as well as the quadratic sum of them (total).

	Results	Uncertainties [%]		
Bin range	$\frac{1}{\sigma}\frac{d\sigma}{dy^t}$	Stat.	Syst.	Total
−2.50 to −1.60	7.55e-02	2.1	5.7	6.0
−1.60 to −1.00	1.99e-01	1.0	1.7	2.0
−1.00 to −0.50	2.84e-01	1.1	1.1	1.6
−0.50 to 0.00	3.37e-01	1.0	2.2	2.4
0.00 to 0.50	3.27e-01	1.1	1.4	1.8
0.50 to 1.00	2.86e-01	1.1	1.1	1.6
1.00 to 1.60	2.03e-01	1.1	2.3	2.5
1.60 to 2.50	7.11e-02	2.1	5.9	6.3
Bin range/π	$\frac{1}{\sigma}\frac{d\sigma}{d\Delta\phi(t,\bar{t})/\pi}$	Stat.	Syst.	Total
0.00 to 0.60	1.96e-01	2.1	8.8	9.0
0.60 to 0.88	9.34e-01	1.1	4.6	4.8
0.88 to 0.97	3.94e+00	1.1	3.6	3.8
0.97 to 1.00	8.88e+00	1.6	6.9	7.1
Bin range	$\frac{1}{\sigma}\frac{d\sigma}{d\lvert y^t\rvert-\lvert y^{\bar{t}}\rvert}$	Stat.	Syst.	Total
−2.50 to −1.10	6.84e-02	2.5	2.8	3.7
−1.10 to −0.50	2.90e-01	1.5	2.2	2.6
−0.50 to 0.00	4.62e-01	1.5	2.3	2.8
0.00 to 0.50	4.41e-01	1.6	3.3	3.7
0.50 to 1.10	2.98e-01	1.5	3.5	3.8
1.10 to 2.50	7.04e-02	2.4	5.0	5.5

Table C.8: Normalised differential cross section results and uncertainties for the dilepton channel in the full phase space as a function of the top quark rapidity (y^t), azimuthal angle difference ($\Delta\phi\left(t,\bar{t}\right)/\pi$) and rapidity difference between the top quarks ($|y^t| - |y^{\bar{t}}|$). The uncertainties are given separated for the statistical and total systematic component as well as the quadratic sum of them (total).

	Results	Uncertainties [%]		
Bin range [GeV]	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}p_{\mathrm{T}}^{\mathrm{t\bar{t}}}}$ [GeV^{-1}]	Stat.	Syst.	Total
0 to 30	1.42e-02	0.9	4.6	4.7
30 to 80	7.03e-03	1.0	3.6	3.8
80 to 170	1.87e-03	1.1	5.0	5.1
170 to 300	3.43e-04	2.2	4.1	4.6
Bin range [GeV]	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}m^{\mathrm{t\bar{t}}}}$[GeV^{-1}]	Stat.	Syst.	Total
340 to 380	3.91e-03	3.2	12.8	13.2
380 to 470	4.56e-03	1.7	7.6	7.8
470 to 620	1.96e-03	1.8	3.8	4.2
620 to 820	5.28e-04	2.8	4.1	5.0
820 to 1100	9.96e-05	3.8	8.5	9.3
1100 to 1600	9.97e-06	10.6	16.0	19.2
Bin range	$\frac{1}{\sigma}\frac{\mathrm{d}\sigma}{\mathrm{d}y^{\mathrm{t\bar{t}}}}$	Stat.	Syst.	Total
−2.50 to −1.50	4.80e-02	4.1	8.2	9.2
−1.50 to −1.00	1.81e-01	1.8	2.8	3.3
−1.00 to −0.50	3.15e-01	1.5	1.5	2.1
−0.50 to 0.00	4.15e-01	1.3	1.8	2.2
0.00 to 0.50	4.07e-01	1.3	2.7	3.0
0.50 to 1.00	3.24e-01	1.5	1.5	2.1
1.00 to 1.50	1.85e-01	1.8	1.8	2.5
1.50 to 2.50	4.50e-02	4.2	12.2	12.9

Table C.9: Normalised differential cross sections and uncertainties for the dilepton channel in the full phase space as a function of the of top-quark-pair transverse momentum ($p_{\mathrm{T}}^{\mathrm{t\bar{t}}}$), invariant mass ($m^{\mathrm{t\bar{t}}}$) and rapidity ($y^{\mathrm{t\bar{t}}}$) distributions. The uncertainties are given separated for the statistical and total systematic component as well as the quadratic sum of them (total).

C.2 Results in Individual Channels

In Chapter 8 the $t\bar{t}$ differential cross section results are presented in the combined channel, and the tables summarising the results are given in Appendix C.1. As explained in Section 6.5.3, the combined channel is obtained as the statistical combination of the individual ee, eμ and $\mu\mu$ $t\bar{t}$ decay channels. The combined result is driven by lower statistical uncertainty of the eμ channel. The $t\bar{t}$ normalised differential cross section results of the individual ee, eμ and $\mu\mu$ channels are presented in Figures C.1–C.11 for all measured quantities discussed in Chapter 8. The results are compared to the predictions obtained from MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig Monte Carlo simulations, and wherever available to higher-order perturbative QCD theory calculations [3, 21, 100, 101] up to approximate NNLO accuracy. The results measured in the individual channels present similar results among them. In general, the best description of the measured distributions is achieved by the Powheg+Herwig prediction or by the approximate NNLO calculation, depending on the observable.

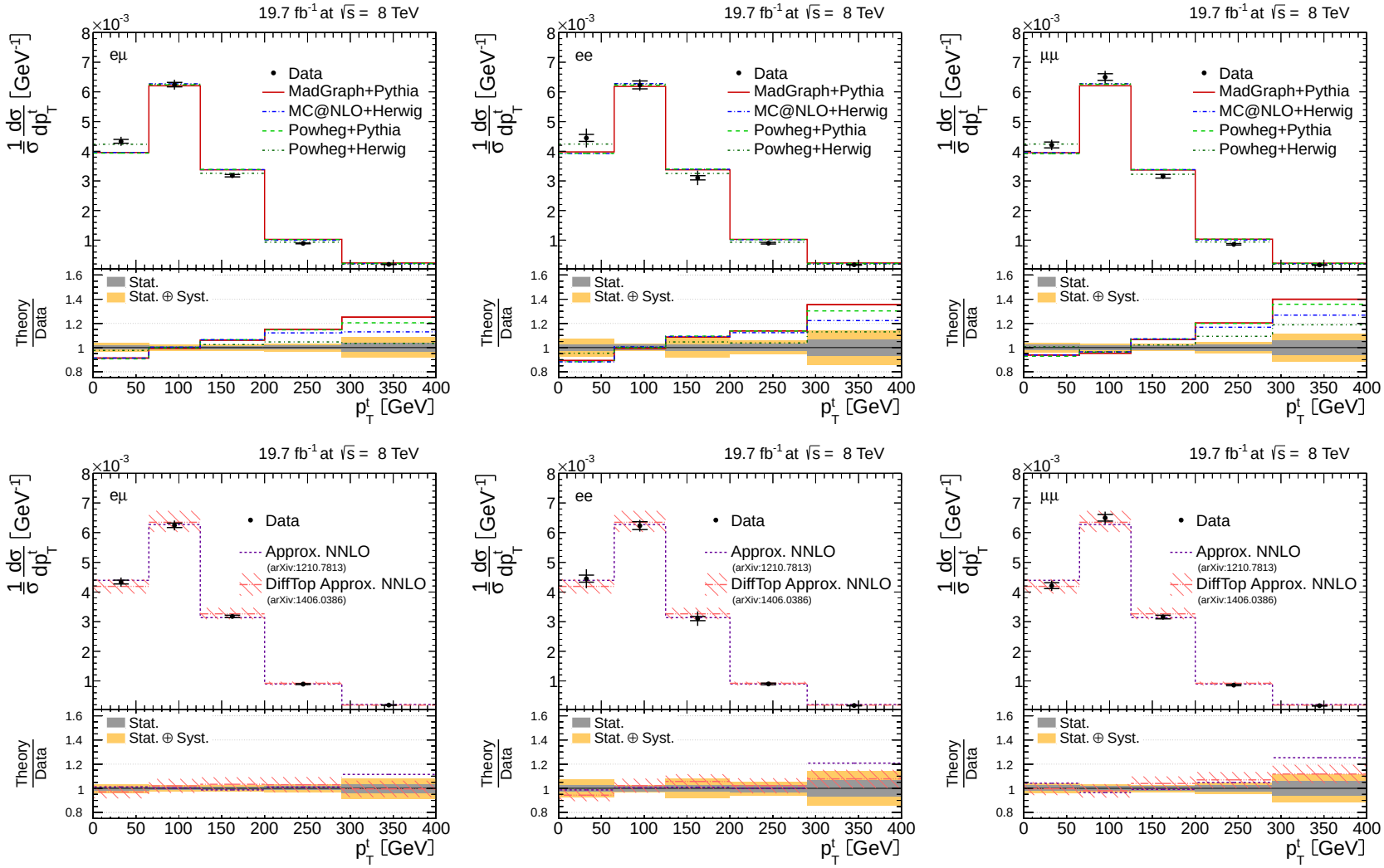


Figure C.1: Normalised $t\bar{t}$ differential cross section as a function of the top transverse momentum for the eμ (left), ee (centre) and $\mu\mu$ (right) channels. The inner (outer) error bars indicate the statistical (total) uncertainty of the measurement. A theory-to-data ratio plot is shown in the bottom panel, with the grey (yellow) error band indicating the statistical (total) data uncertainty.
Top: Results are compared to predictions from MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig. Bottom: Results are compared to perturbative QCD theory calculations up to approximate NNLO accuracy [3,21]. The error band on the DiffTop prediction includes the renormalisation and factorisation scales, α_s, PDF set and top-quark mass uncertainty.

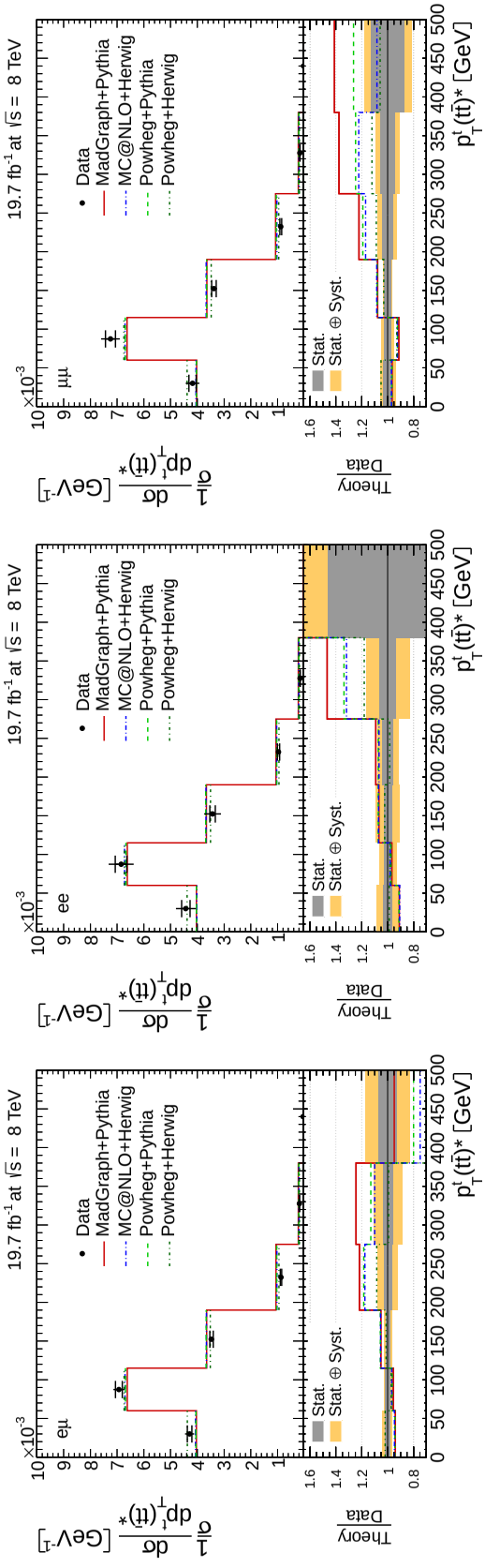


Figure C.2: Normalised $t\bar{t}$ differential cross section as a function of the top transverse momentum, measured in the $t\bar{t}$ rest frame reference system, for the $e\mu$ (left), ee (centre) and $\mu\mu$ (right) channels. The inner (outer) error bars indicate the statistical (total) uncertainty of the measurement. A theory-to-data ratio plot is shown in the bottom panel, with the grey (yellow) error band indicating the statistical (total) data uncertainty. Results are compared to predictions from MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig.

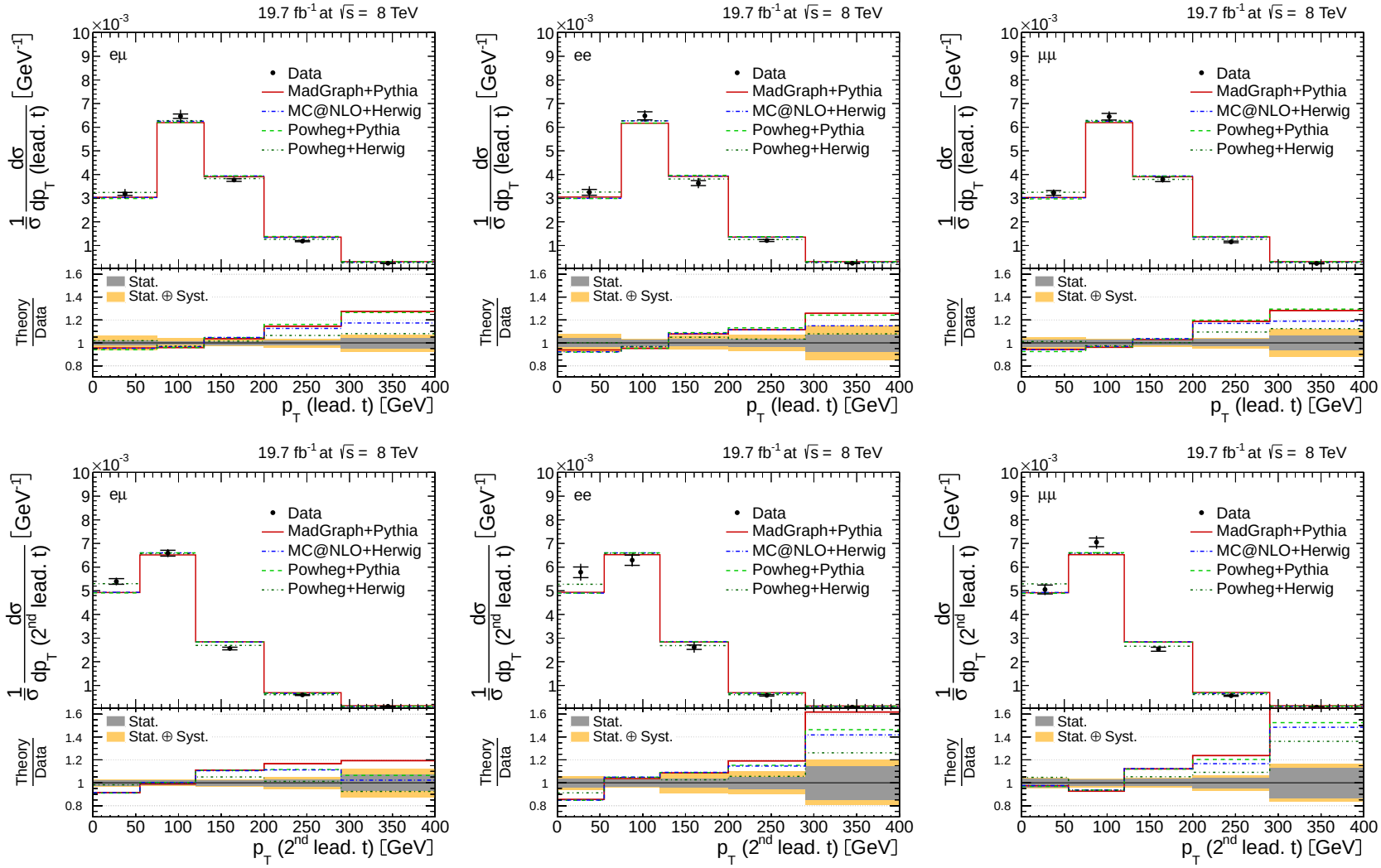


Figure C.3: Normalised $t\bar{t}$ differential cross section as a function of the transverse momentum of the leading top quark (top row) and trailing top quark (bottom row). Results are presented for the eμ (left), ee (centre) and $\mu\mu$ (right) channels separately. The inner (outer) error bars indicate the statistical (total) uncertainty of the measurement. A theory-to-data ratio plot is shown in the bottom panel, with the grey (yellow) error band indicating the statistical (total) data uncertainty. Results are compared to predictions from MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig.

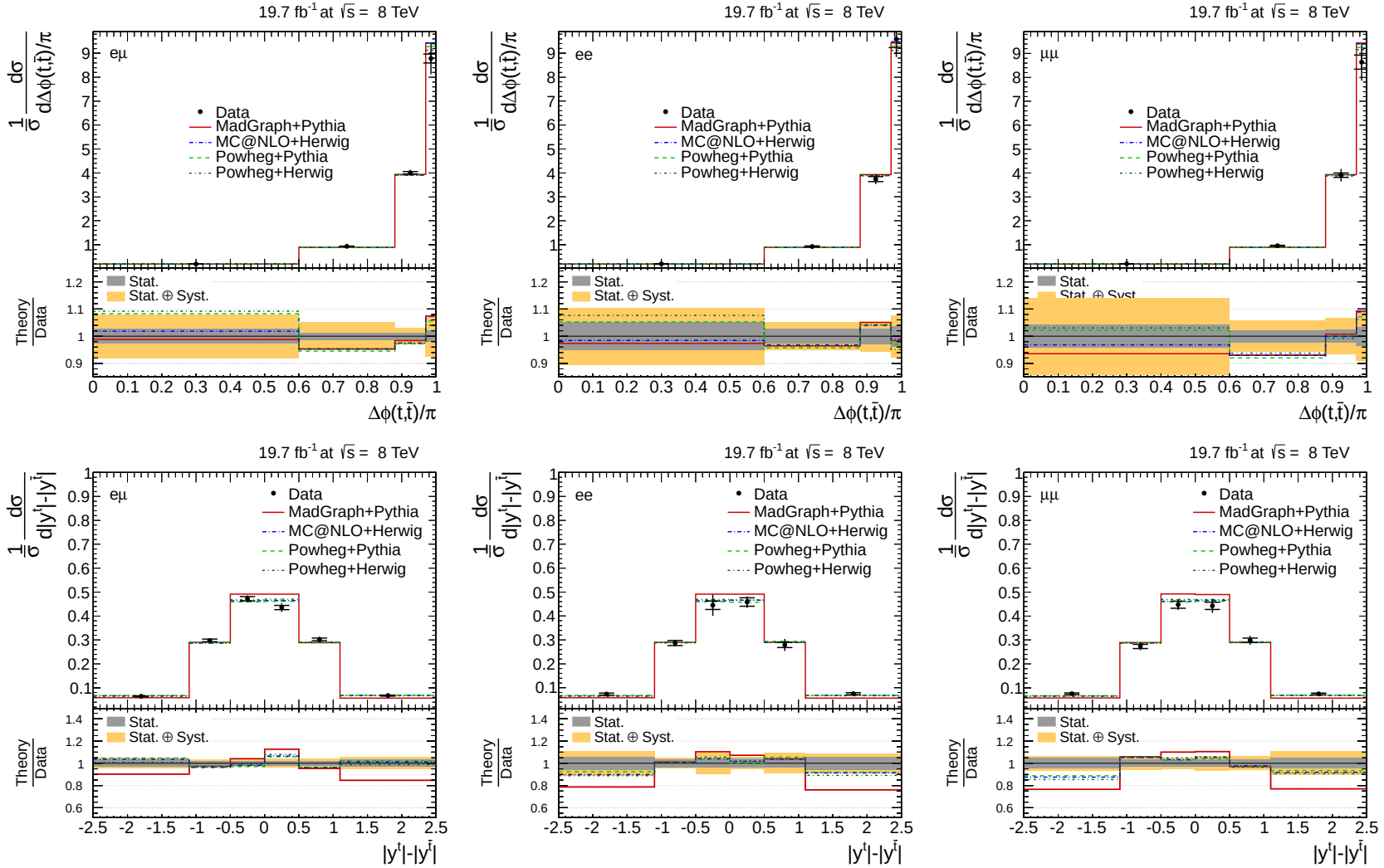


Figure C.4: Normalised $t\bar{t}$ differential cross section as a function of the difference in azimuthal angle (top row) and in absolute rapidity (bottom row) between the top quark and antiquark. Results are presented for the eμ (left), ee (centre) and $\mu\mu$ (right) channels separately. The inner (outer) error bars indicate the statistical (total) uncertainty of the measurement. A theory-to-data ratio plot is shown in the bottom panel, with the grey (yellow) error band indicating the statistical (total) data uncertainty. Results are compared to MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig Monte Carlo predictions.

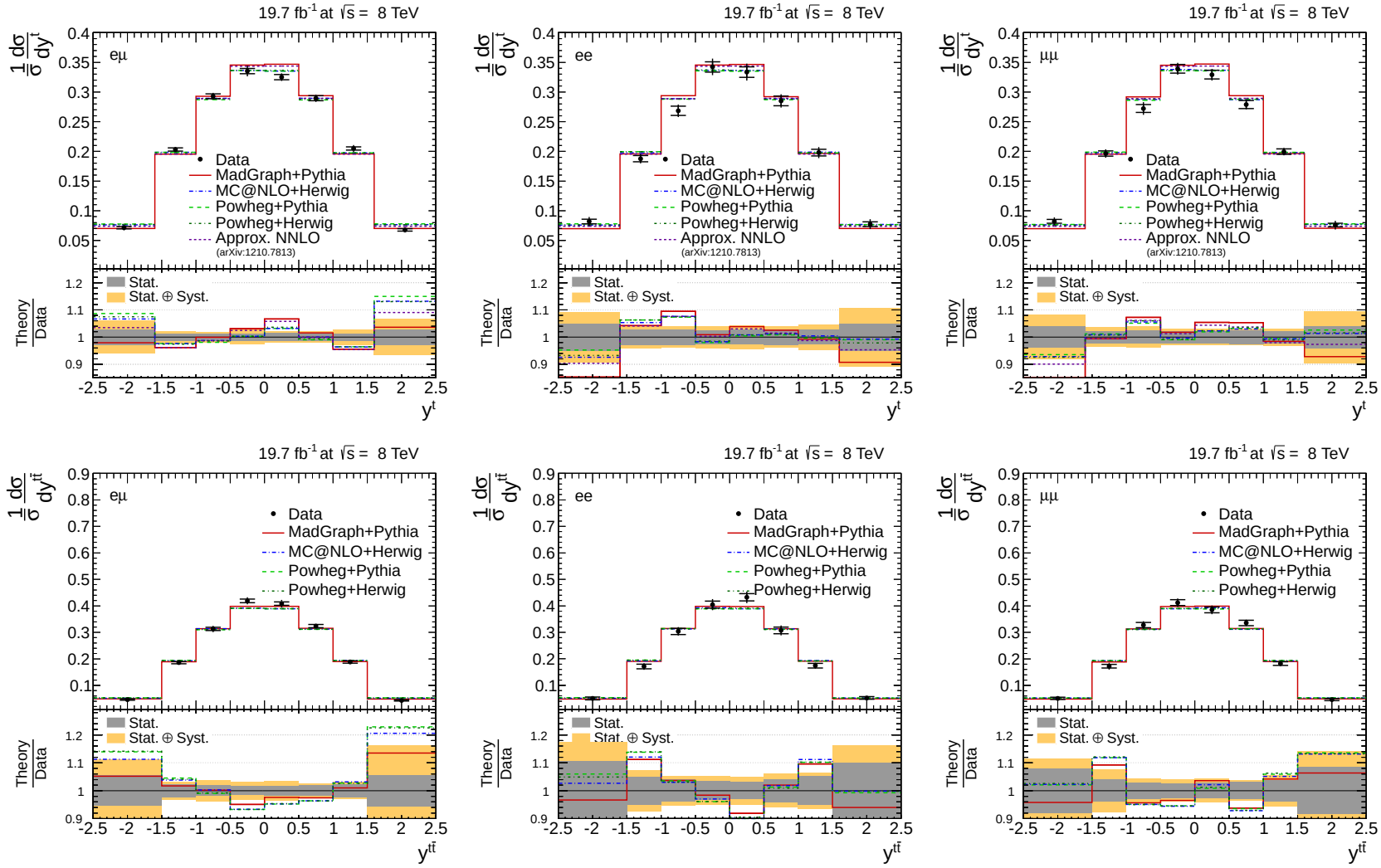


Figure C.5: Normalised $t\bar{t}$ differential cross section as a function of the rapidity of the top quark (top row) and $t\bar{t}$ system (bottom row). Results are presented for the eμ (left), ee (centre) and $\mu\mu$ (right) channels separately. The inner (outer) error bars indicate the statistical (total) uncertainty of the measurement. A theory-to-data ratio plot is shown in the bottom panel, with the grey (yellow) error band indicating the statistical (total) data uncertainty. Results are compared to MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig Monte Carlo predictions and to QCD theory calculation up to approximate NNLO accuracy [21]. The top quark rapidity result is also compared to a QCD theory calculation up to approximate NNLO accuracy [21].

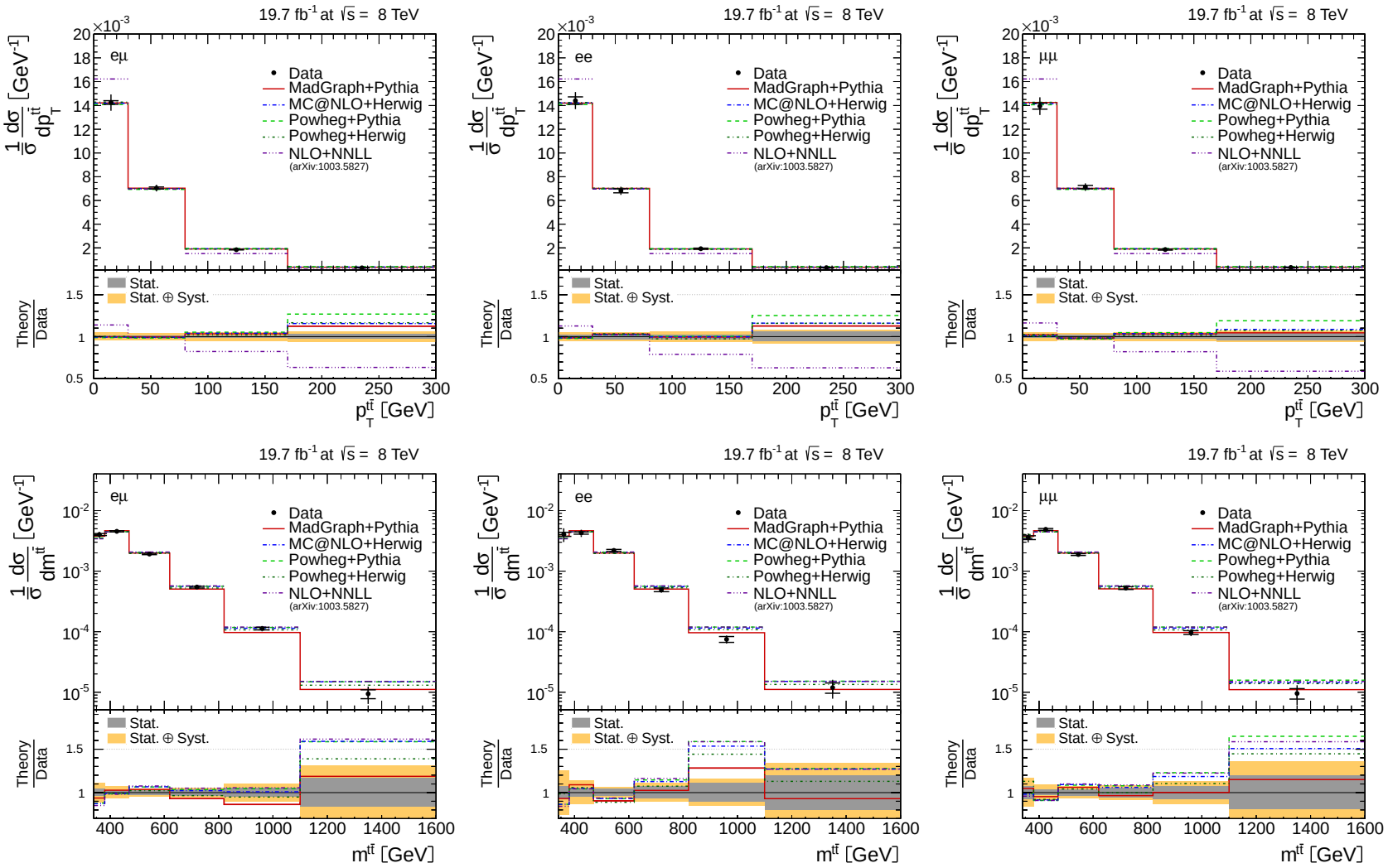


Figure C.6: Normalised $t\bar{t}$ differential cross section as a function of the $t\bar{t}$ transverse momentum (top row) and invariant mass (bottom row). Results are presented for the eμ (left), ee (centre) and $\mu\mu$ (right) channels separately. The inner (outer) error bars indicate the statistical (total) uncertainty of the measurement. A theory-to-data ratio plot is shown in the bottom panel, with the grey (yellow) error band indicating the statistical (total) data uncertainty. Results are compared to predictions from MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig and to a QCD theory calculation up to NLO+NNLL accuracy [100, 101]

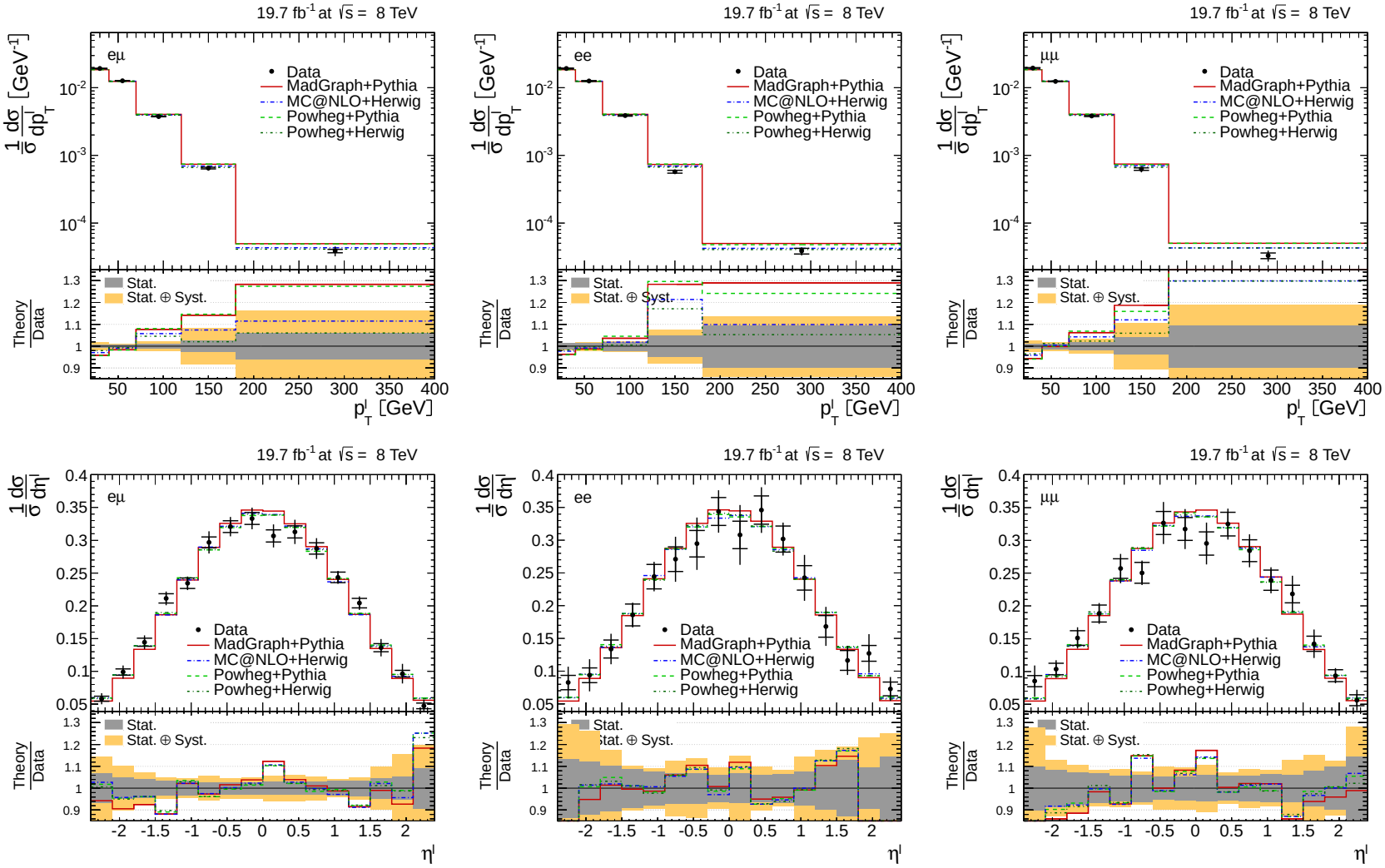


Figure C.7: Normalised $t\bar{t}$ differential cross section as a function of the lepton transverse momentum (top row) and pseudorapidity (bottom row). Results are presented for the eμ (left), ee (centre) and $\mu\mu$ (right) channels separately. The inner (outer) error bars indicate the statistical (total) uncertainty of the measurement. A theory-to-data ratio plot is shown in the bottom panel, with the grey (yellow) error band indicating the statistical (total) data uncertainty. Results are compared to predictions from MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig.

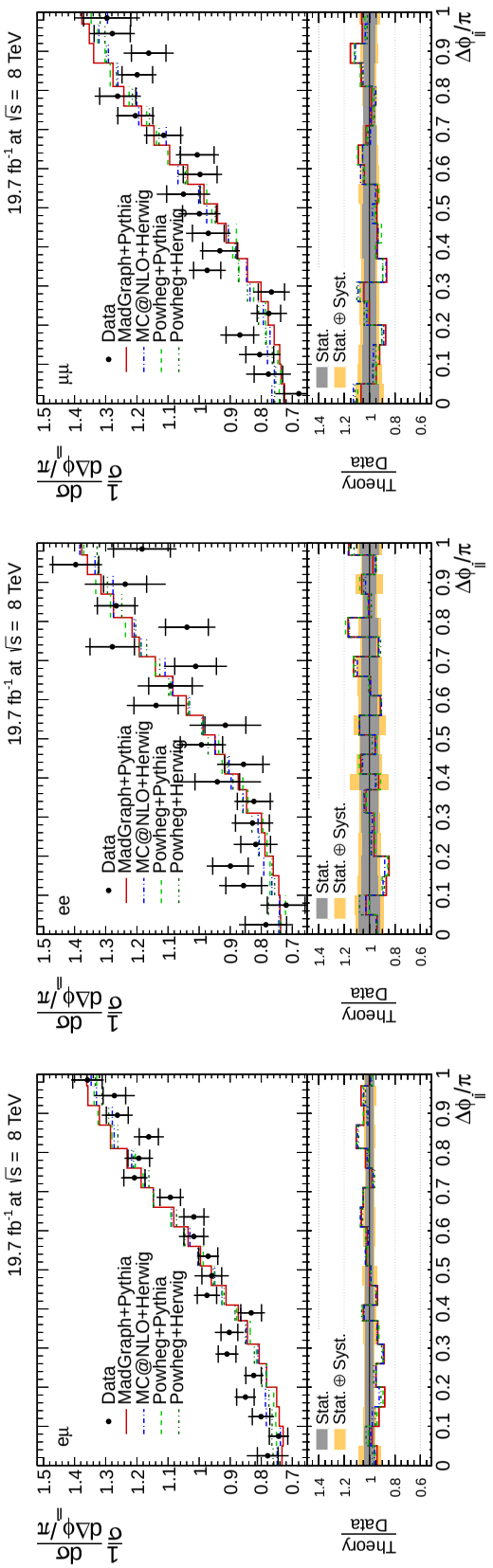


Figure C.8: Normalised $t\bar{t}$ differential cross section as a function of the difference in azimuthal angle between the lepton and antilepton. Results are presented for the eμ (left), ee (centre) and $\mu\mu$ (right) channels separately. The inner (outer) error bars indicate the statistical (total) uncertainty of the measurement. A theory-to-data ratio plot is shown in the bottom panel, with the grey (yellow) error band indicating the statistical (total) data uncertainty. Results are compared to predictions from MadGraph+Pythia, Powheg+Herwig and MC@NLO+Herwig.

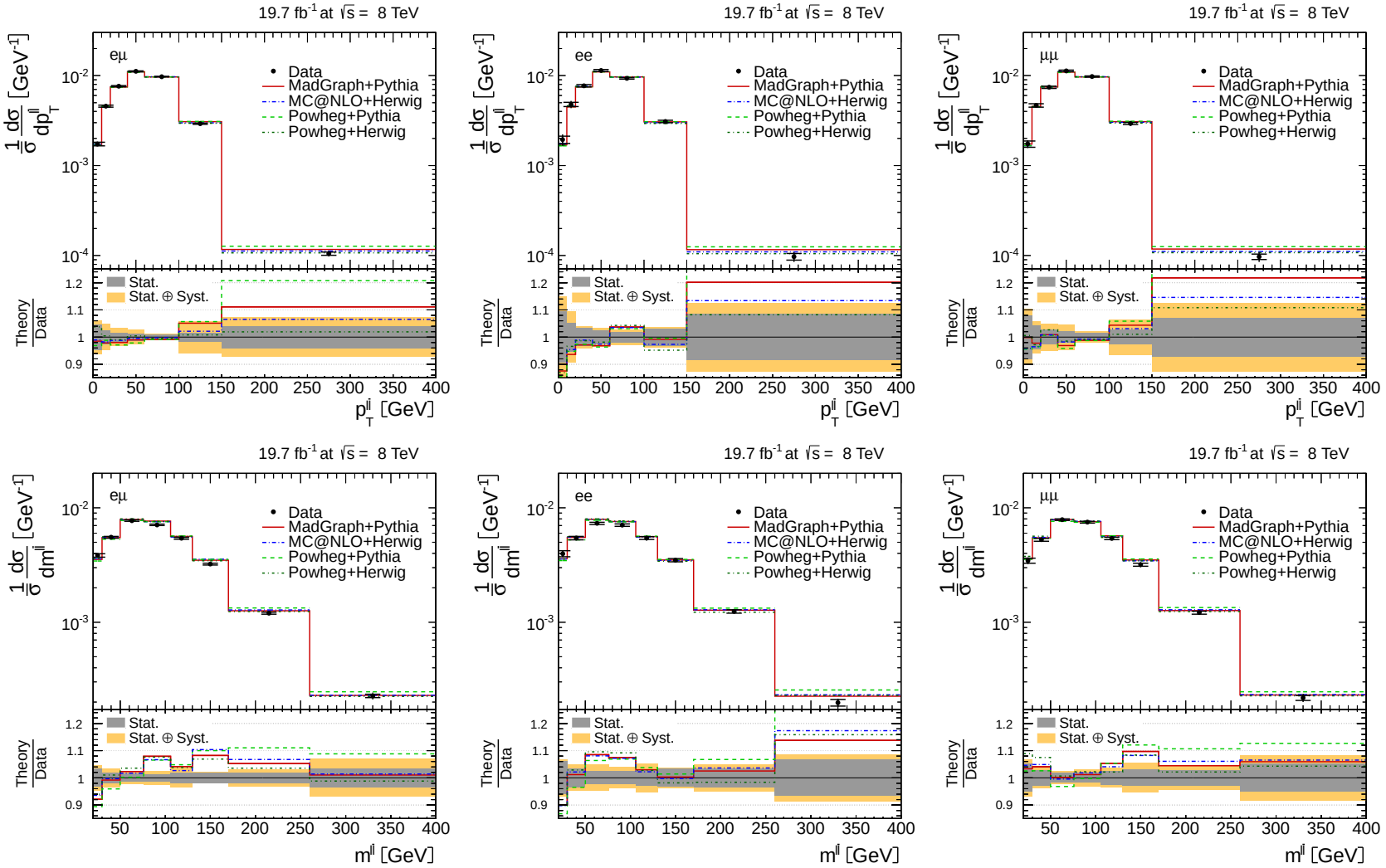


Figure C.9: Normalised $t\bar{t}$ differential cross section as a function of the lepton-pair transverse momentum (top row) and invariant mass (bottom row). Results are presented for the eμ (left), ee (centre) and $\mu\mu$ (right) channels separately. The inner (outer) error bars indicate the statistical (total) uncertainty of the measurement. A theory-to-data ratio plot is shown in the bottom panel, with the grey (yellow) error band indicating the statistical (total) data uncertainty. Results are compared to predictions from MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig.

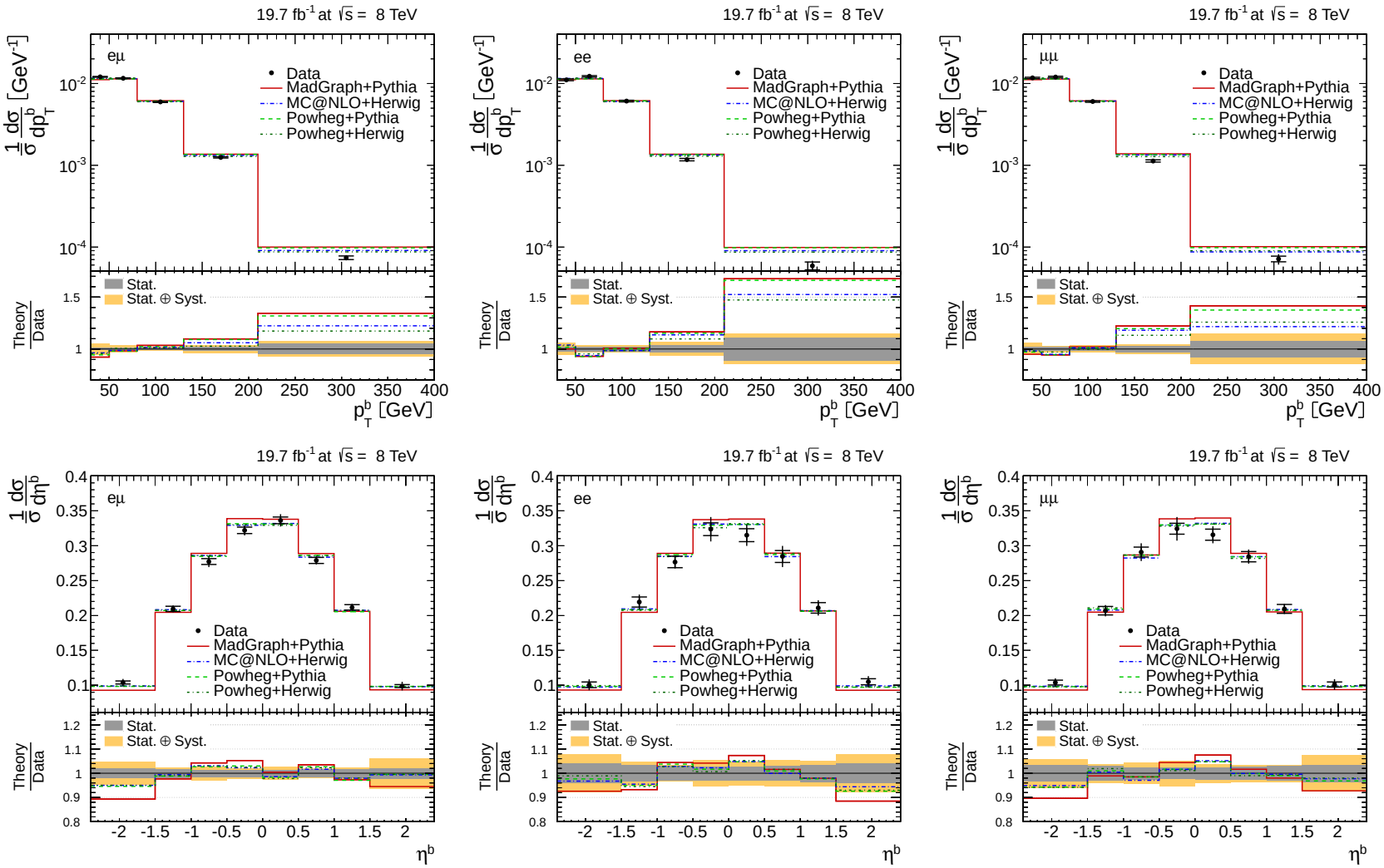


Figure C.10: Normalised $t\bar{t}$ differential cross section as a function of the b-jet transverse momentum (top row) and pseudorapidity (bottom row). Results are presented for the eμ (left), ee (centre) and $\mu\mu$ (right) channels separately. The inner (outer) error bars indicate the statistical (total) uncertainty of the measurement. A theory-to-data ratio plot is shown in the bottom panel, with the grey (yellow) error band indicating the statistical (total) data uncertainty. Results are compared to predictions from MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig.

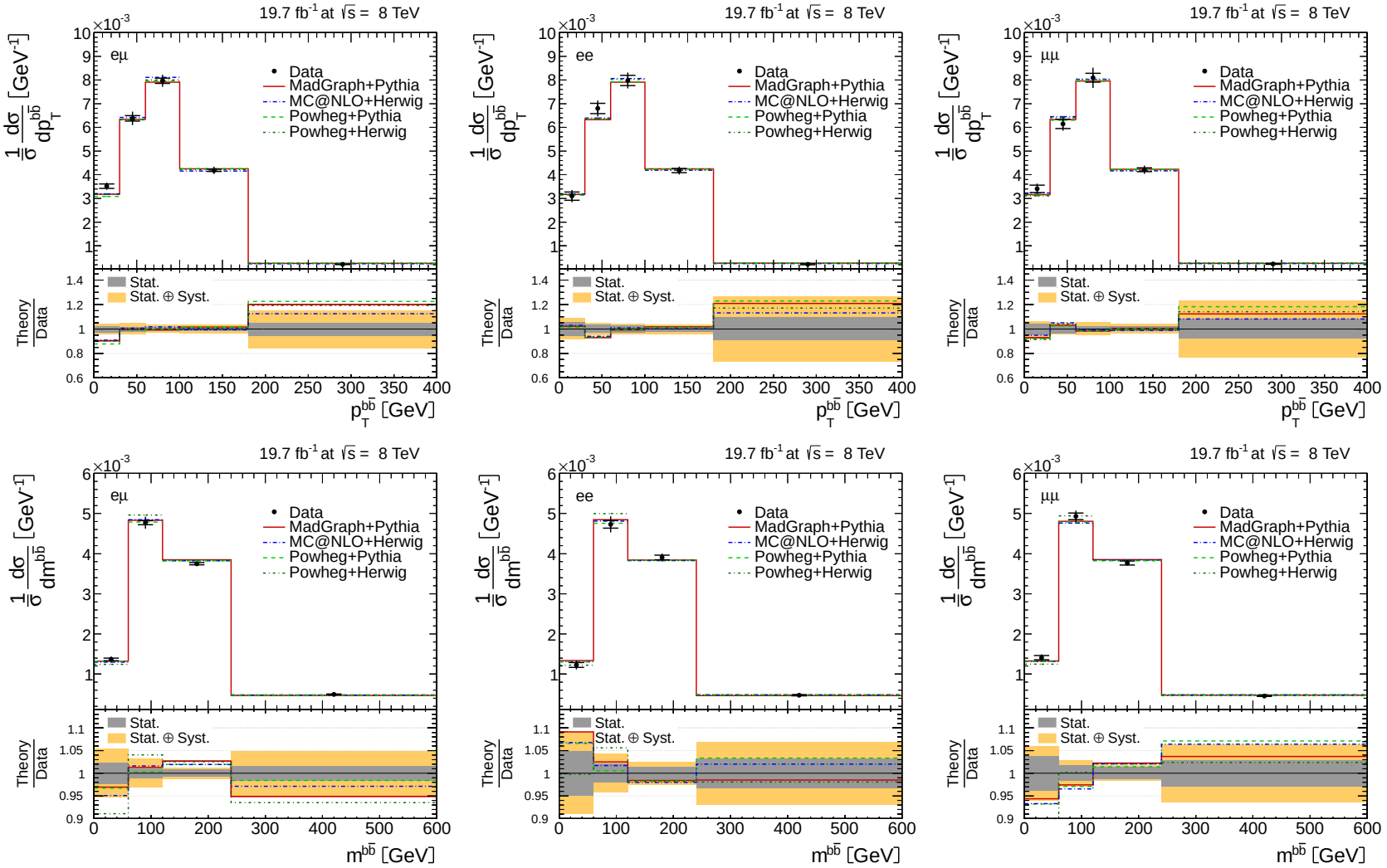


Figure C.11: Normalised $t\bar{t}$ differential cross section as a function of the b-jet-pair transverse momentum (top row) and invariant mass (bottom row). Results are presented for the eμ (left), ee (centre) and $\mu\mu$ (right) channels separately. The inner (outer) error bars indicate the statistical (total) uncertainty of the measurement. A theory-to-data ratio plot is shown in the bottom panel, with the grey (yellow) error band indicating the statistical (total) data uncertainty. Results are compared to predictions from MadGraph+Pythia, Powheg+Pythia, Powheg+Herwig and MC@NLO+Herwig.

Appendix D

Stability Checks: Distributions

D.1 Dependency on Pileup Conditions

The default analysis requires the presence of at least one well reconstructed vertex on the selected events as discussed in Section 5.2 and shown in Figure D.1a. The analysis is repeated by selecting events in three exclusive PU regions:

- *Low* PU: events with at least 1 and less than 10 well reconstructed vertices.
- *Medium* PU: events with at least 10 and less than 20 vertices.
- *High* PU: events with more than 20 vertices.

The calculation of the results proceeds by applying all event selection criteria detailed in Chapter 5, with the alternative vertex multiplicity condition. The individual top-quark 4-momentum components are determined using the same kinematic reconstruction algorithm described in Section 5.8. The measured signal yields are corrected for detector and migration effects by means of unfolding techniques (see Section 6.4). Finally, the results are obtained in the same binning as the default analysis such that the results can be compared bin-by-bin.

Figure D.1b presents the top-quark p_{T} differential cross section obtained selecting events in the three aforementioned exclusive PU regions. For comparison, the default result (PU ≥ 1) and the default MadGraph+Pythia prediction are shown. The data points are horizontally displaced for graphical purposes only, but the meaning of each data point remains unaltered: the normalised cross section value in the range of the bin where it is presented.

It can be observed that the top-quark p_{T} distribution remains softer than the spectrum predicted by the MadGraph+Pythia simulation, independently of the number of vertices selected. The same conclusion is found for the results measured as a function of the transverse momentum distribution of the leptons and the b-jets. No bias due to the PU is found.

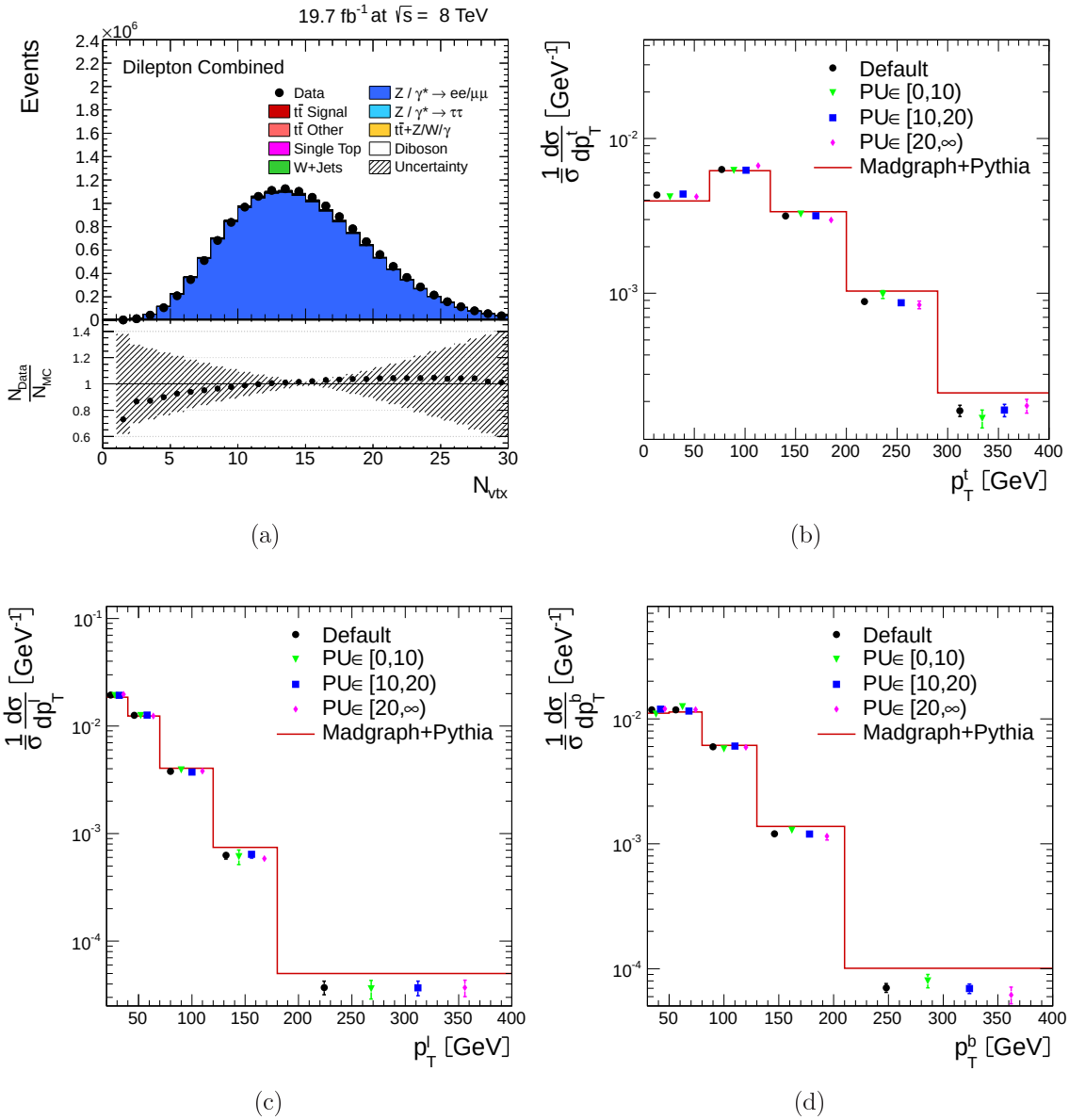


Figure D.1: (a): Distribution of the average number of reconstructed vertices in data and simulation events selected after trigger and dilepton selection criteria (c.f. Section 5.3). The hatched area corresponds to all shape uncertainties on the signal $t\bar{t}$ events (c.f. Chapter 7). Differential cross section as a function of the transverse momentum of the top-quark (b), lepton (c) and b-jet (d) obtained in events selected in four different PU regions, the error bars indicate the total uncertainty on the result. The data points are horizontally displaced from the centre of the bin only for graphical purposes. The default MadGraph+Pythia prediction is shown as reference.

D.2 Reweighting the Signal Simulation

As described in Section 8.4 the stability of the analysis is tested by correcting the $t\bar{t}$ simulation such that it describes the data. In Figures D.2–D.7 the control distribution and the differential cross section results for all measured observables are presented. The differential cross section figures include the results obtained without any correction factor for reference. Additionally, the theory spectra obtained from the MadGraph+Pythia Monte Carlo prediction with and without the event correction factor are displayed.

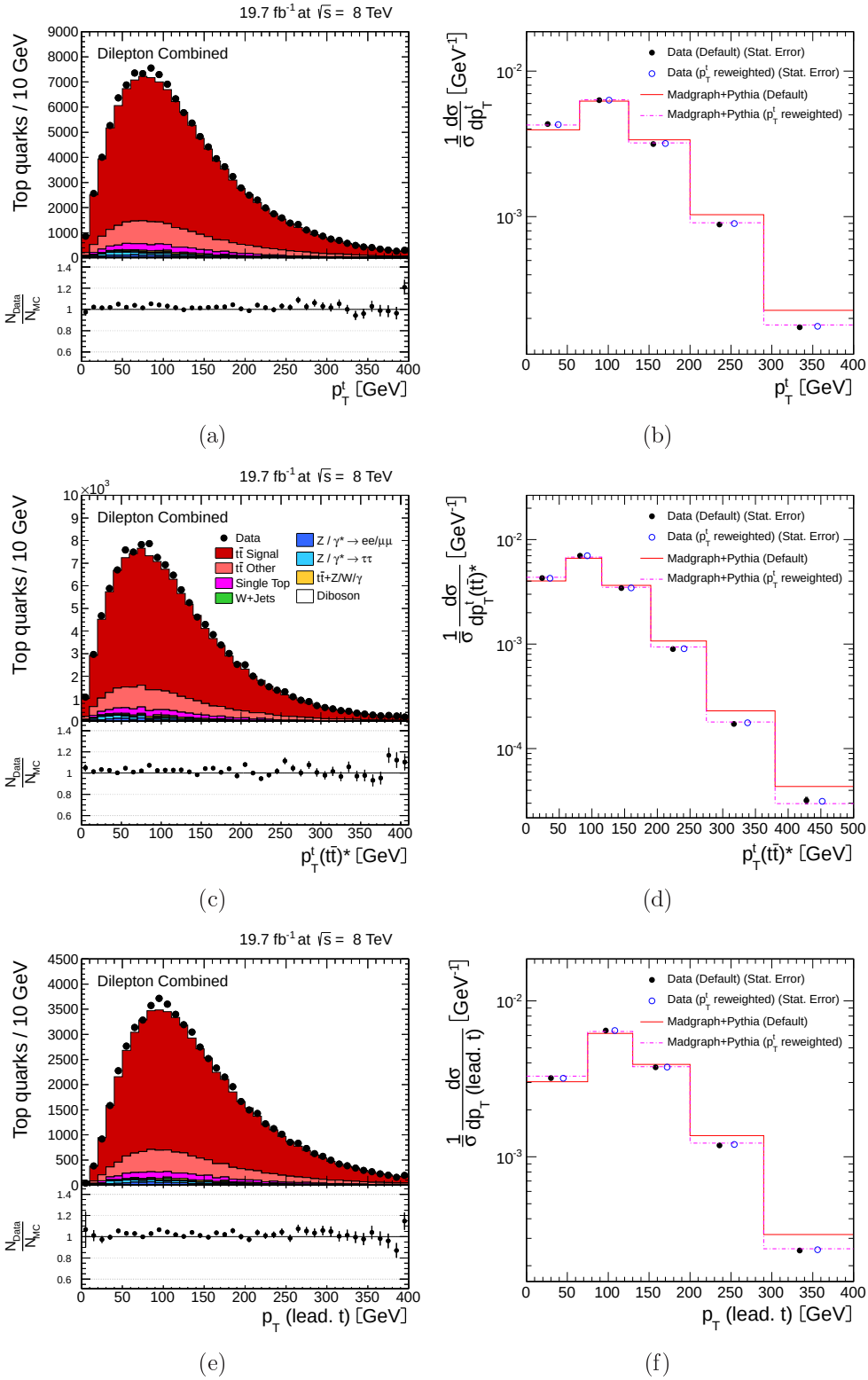


Figure D.2: Comparison of the differential cross sections (right) with and without the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. The results include only the statistical uncertainty and are compared to the unmodified and $p_{\mathrm{T}}^{\mathrm{t}}$-corrected MadGraph+Pythia prediction. Event yields (left) are obtained after the full event selection and include the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. Distributions are shown as a function of the top-quark transverse momentum in the laboratory (a and b), in the $\mathrm{t\bar{t}}$ rest frame (c and d), and leading top-quark p_{T} (e and f).

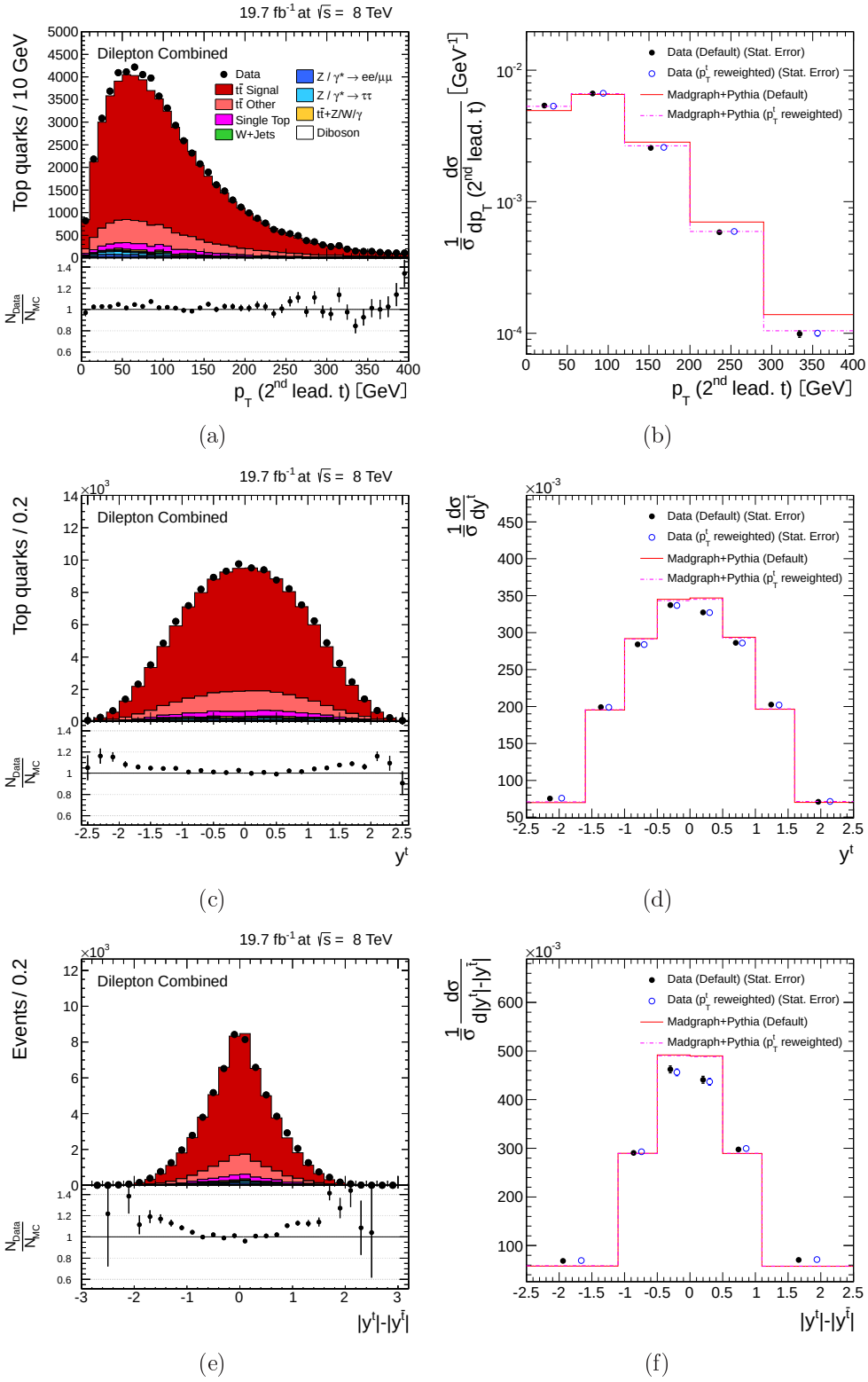


Figure D.3: Comparison of the differential cross sections (right) with and without the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. The results include only the statistical uncertainty and are compared to the unmodified and $p_{\mathrm{T}}^{\mathrm{t}}$-corrected MadGraph+Pythia prediction. Event yields (left) are obtained after the full event selection and include the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. Distributions are shown as a function of the subleading top-quark p_{T} (a and b), top quark rapidity (c and d), and difference in the absolute rapidities of the top quark and antiquark (e and f).

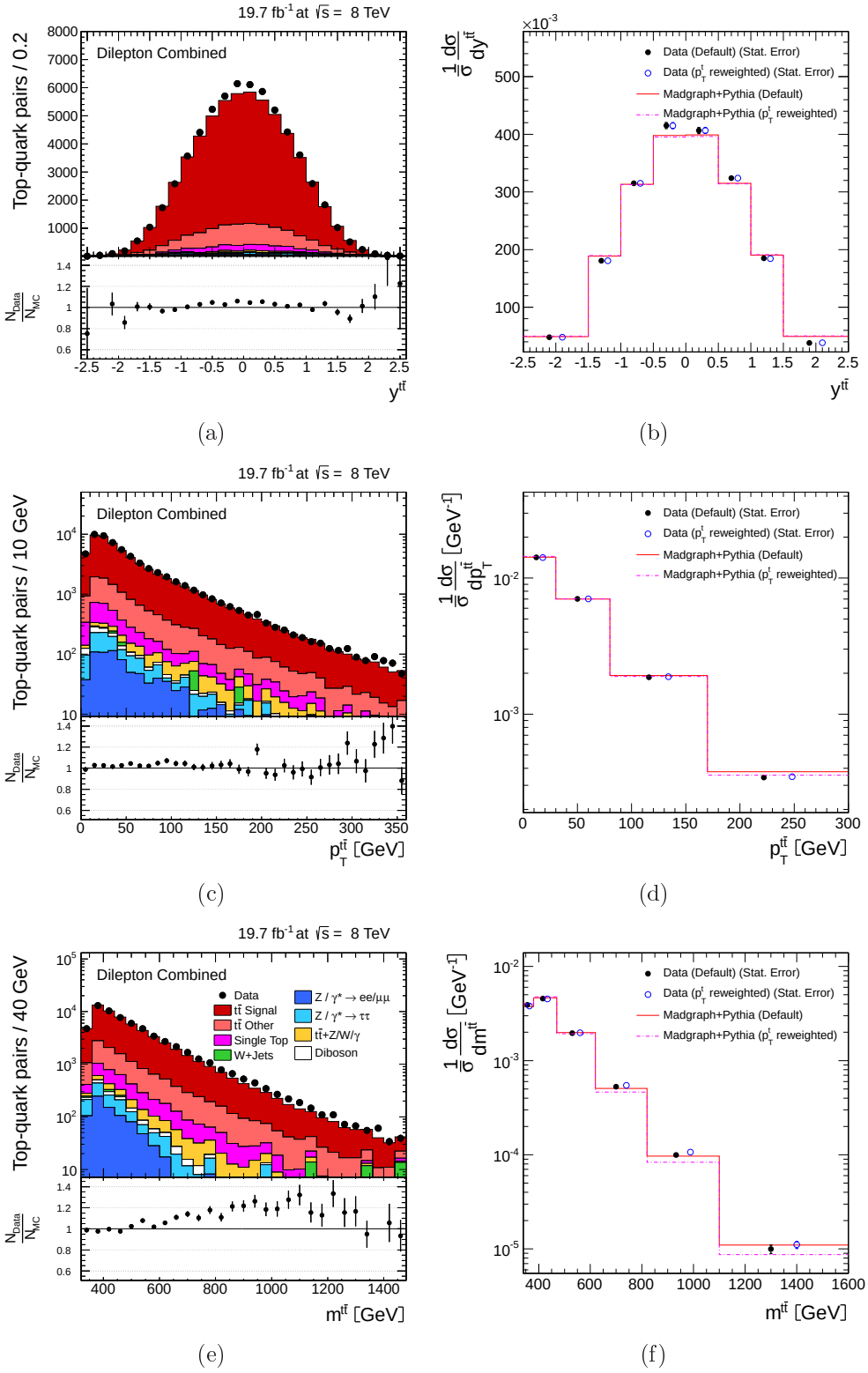


Figure D.4: Comparison of the differential cross sections (right) with and without the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. The results include only the statistical uncertainty and are compared to the unmodified and $p_{\mathrm{T}}^{\mathrm{t}}$-corrected MadGraph+Pythia prediction. Event yields (left) are obtained after the full event selection and include the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. Distributions are shown as a function of the $\mathrm{t\bar{t}}$ rapidity (a and b), p_{T} (c and d), and invariant mass (e and f).

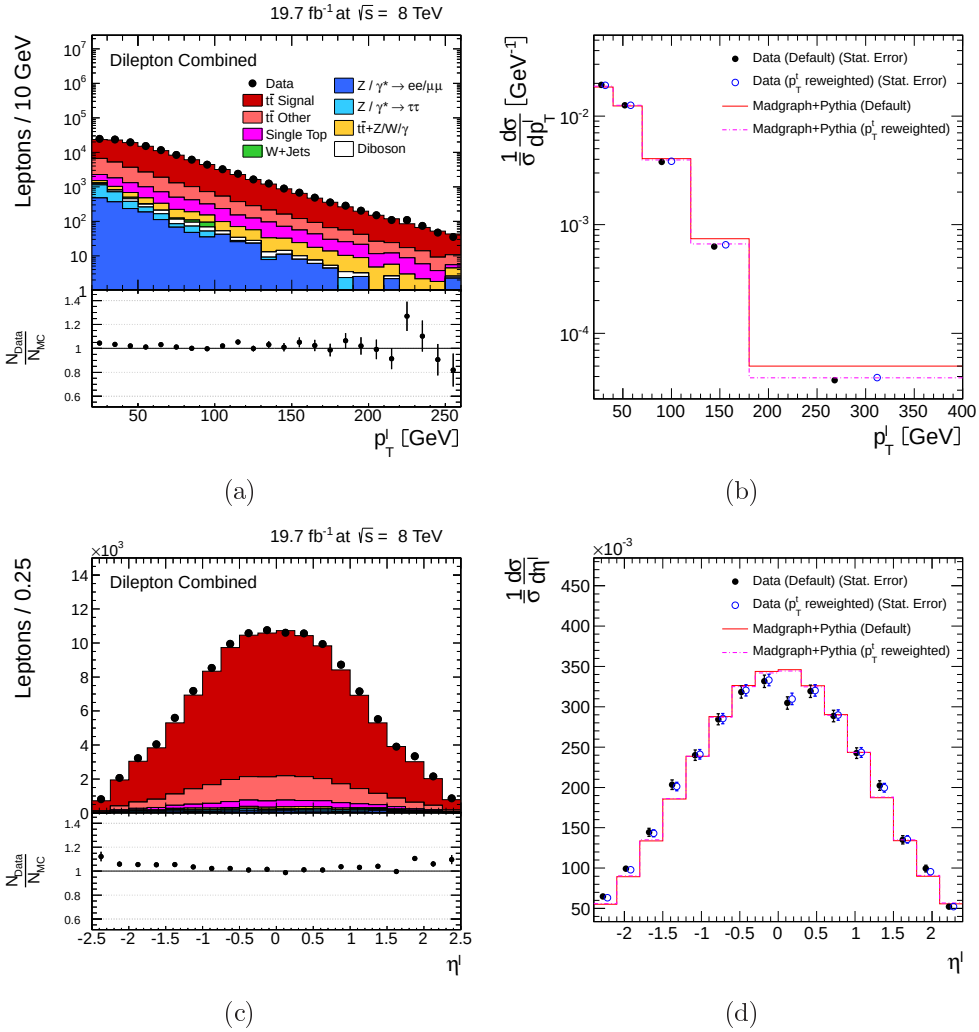


Figure D.5: Comparison of the differential cross sections (right) with and without the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. The results include only the statistical uncertainty and are compared to the unmodified and $p_{\mathrm{T}}^{\mathrm{t}}$-corrected MadGraph+Pythia prediction. Event yields (left) are obtained after the full event selection and include the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. Distributions are shown as a function of the lepton p_{T} (a and b) and pseudorapidity (c and d).

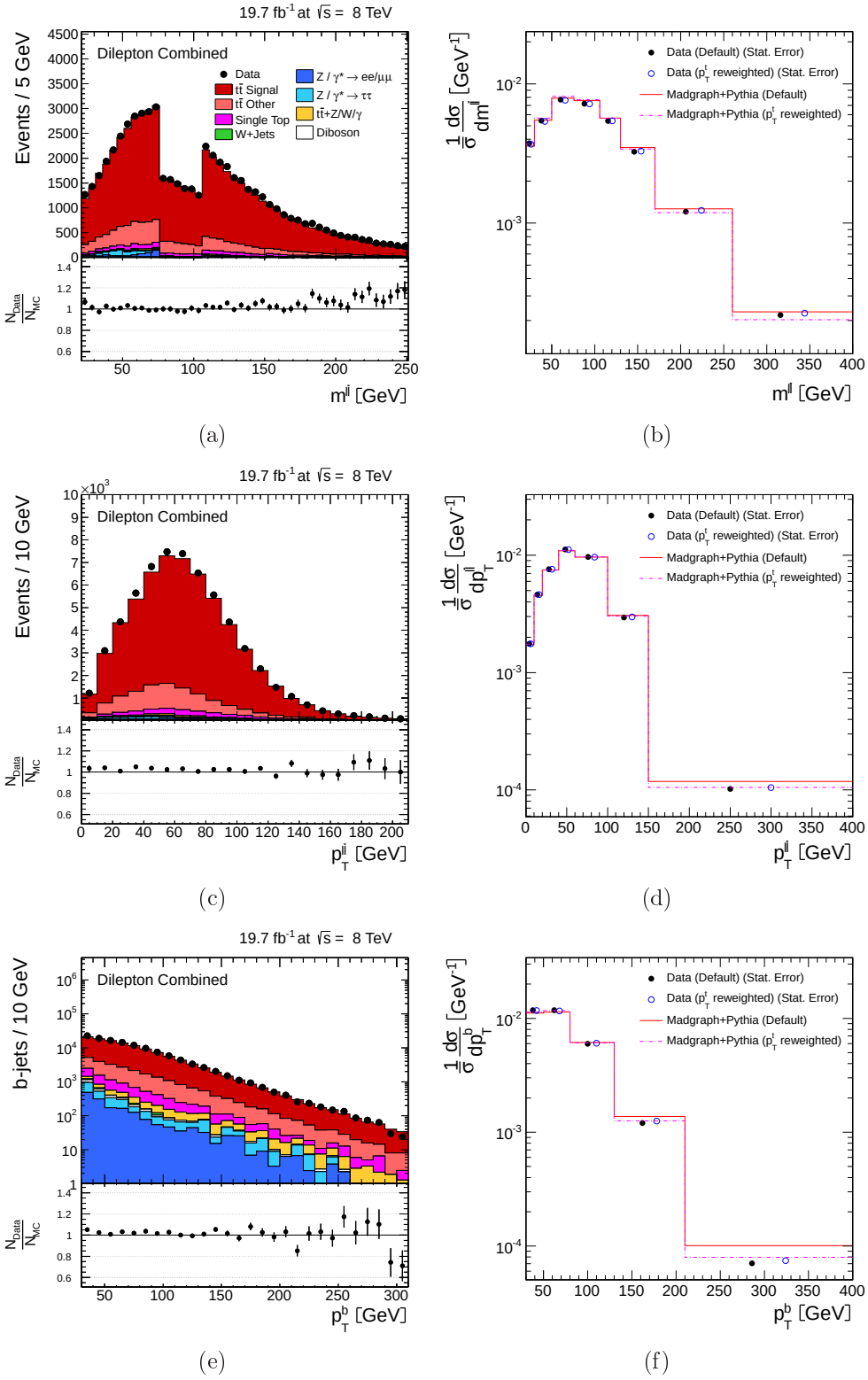


Figure D.6: Comparison of the differential cross sections (right) with and without the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. The results include only the statistical uncertainty and are compared to the unmodified and $p_{\mathrm{T}}^{\mathrm{t}}$-corrected MadGraph+Pythia prediction. Event yields (left) are obtained after the full event selection and include the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. Distributions are shown as a function of the lepton-pair invariant mass (a and b) and p_{T} (c and d), and b-jet p_{T} (e and f).

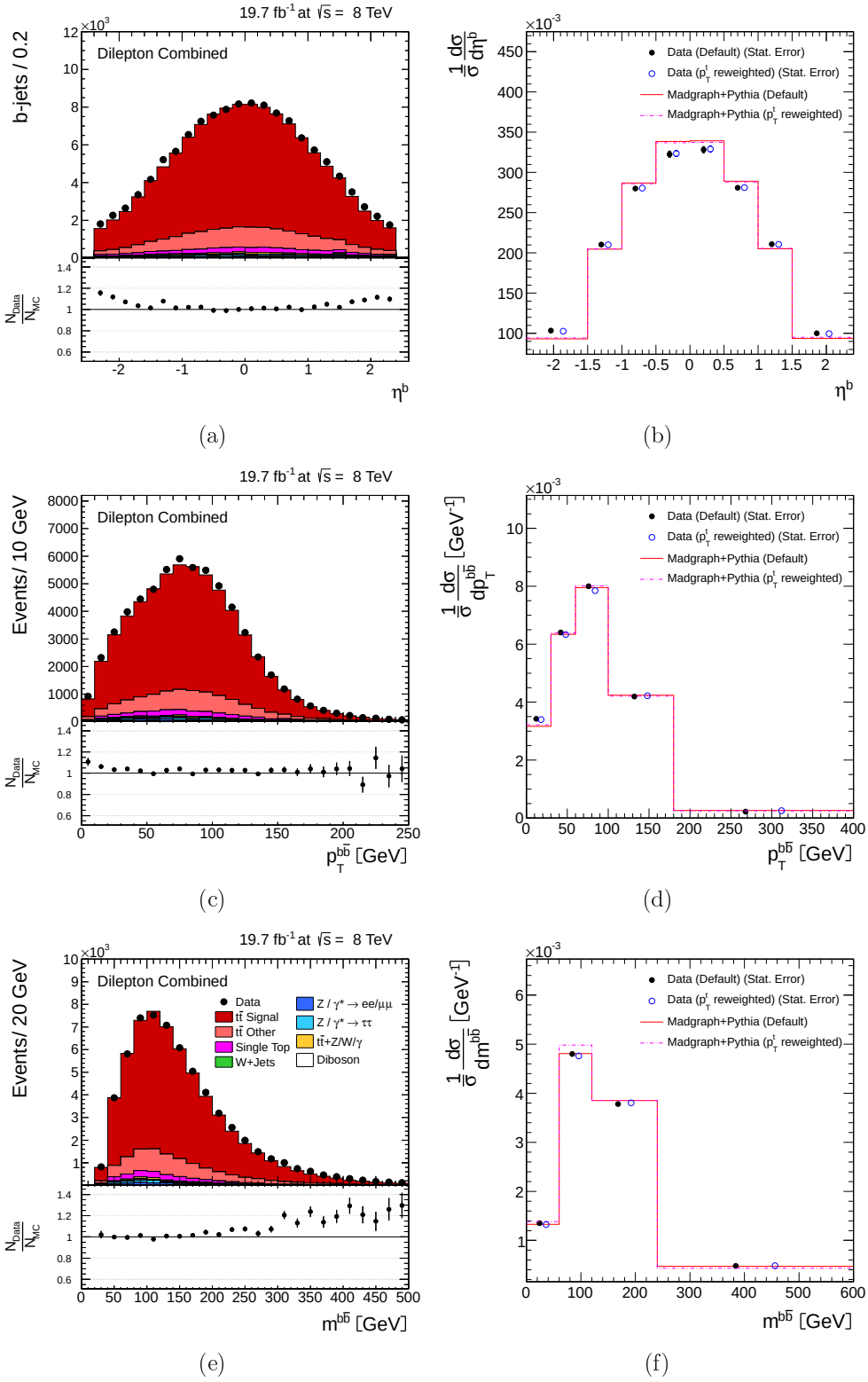
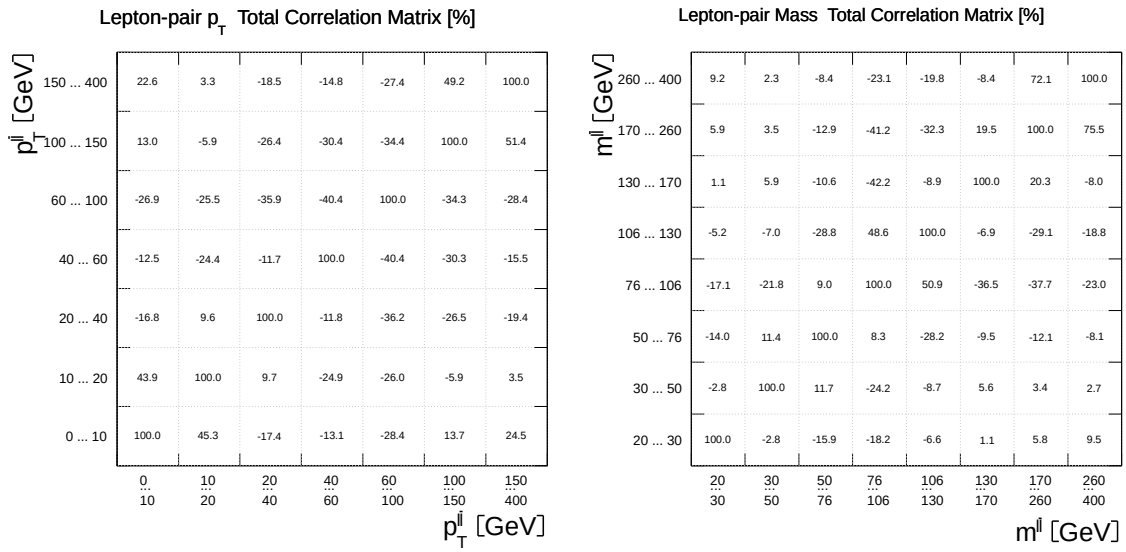


Figure D.7: Comparison of the differential cross sections (right) with and without the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. The results include only the statistical uncertainty and are compared to the unmodified and $p_{\mathrm{T}}^{\mathrm{t}}$-corrected MadGraph+Pythia prediction. Event yields (left) are obtained after the full event selection and include the $p_{\mathrm{T}}^{\mathrm{t}}$-correction. Distributions are shown as a function of the b-jet pseudorapidity (a and b) and b-jet-pair p_{T} (c and d), and invariant mass (e and f).

Appendix E

Covariance Matrices

In Figures E.1–E.6 the correlation matrix of the full uncertainty are displayed as a function of the kinematic distributions measured in this analysis. The method to derived the correlation matrix is explained in Section 7.5.



(a) (b)

Figure E.1: Covariance matrix of the total uncertainty of the differential cross section result measured as a function of the lepton-pair transverse momentum (a) and invariant mass (b).

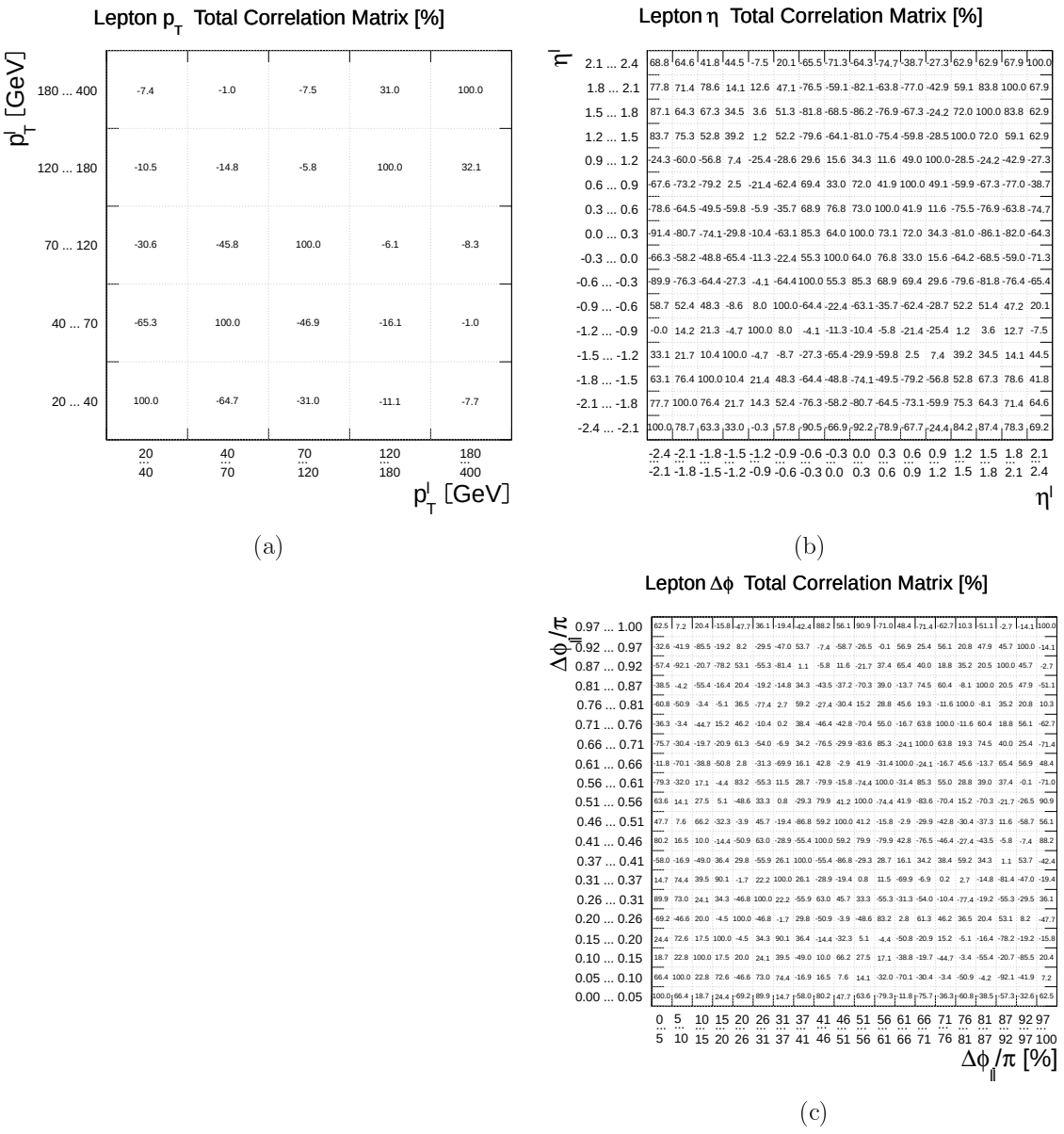


Figure E.2: Covariance matrix of the total uncertainty of the differential cross section result measured as a function of the lepton transverse momentum (a), pseudorapidity (b) and azimuthal angle difference (c).

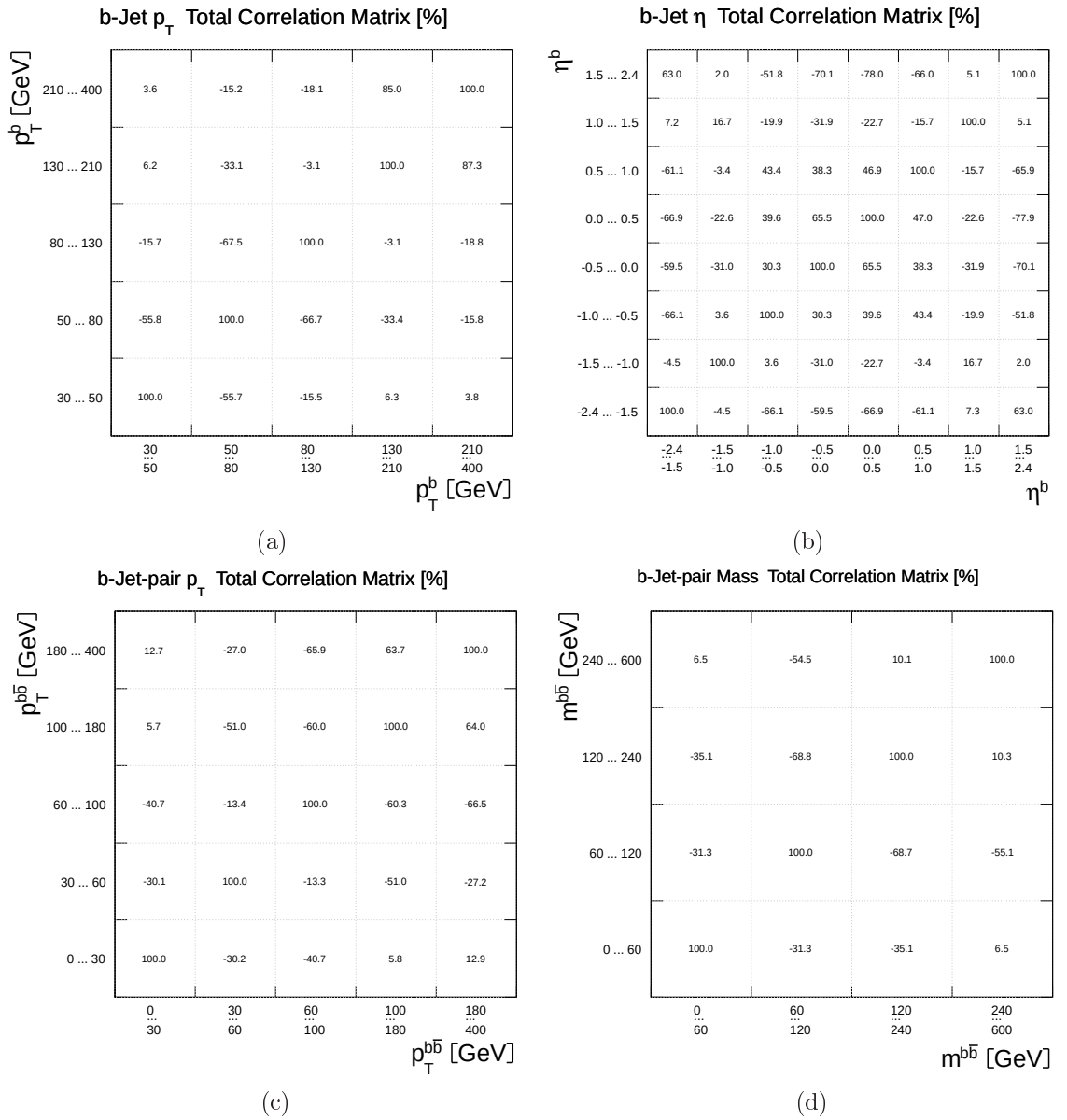


Figure E.3: Covariance matrix of the total uncertainty of the differential cross section result measured as a function of the b-jet transverse momentum (a) and pseudorapidity (b), and b-jet-pair transverse momentum (c) and invariant mass (d).

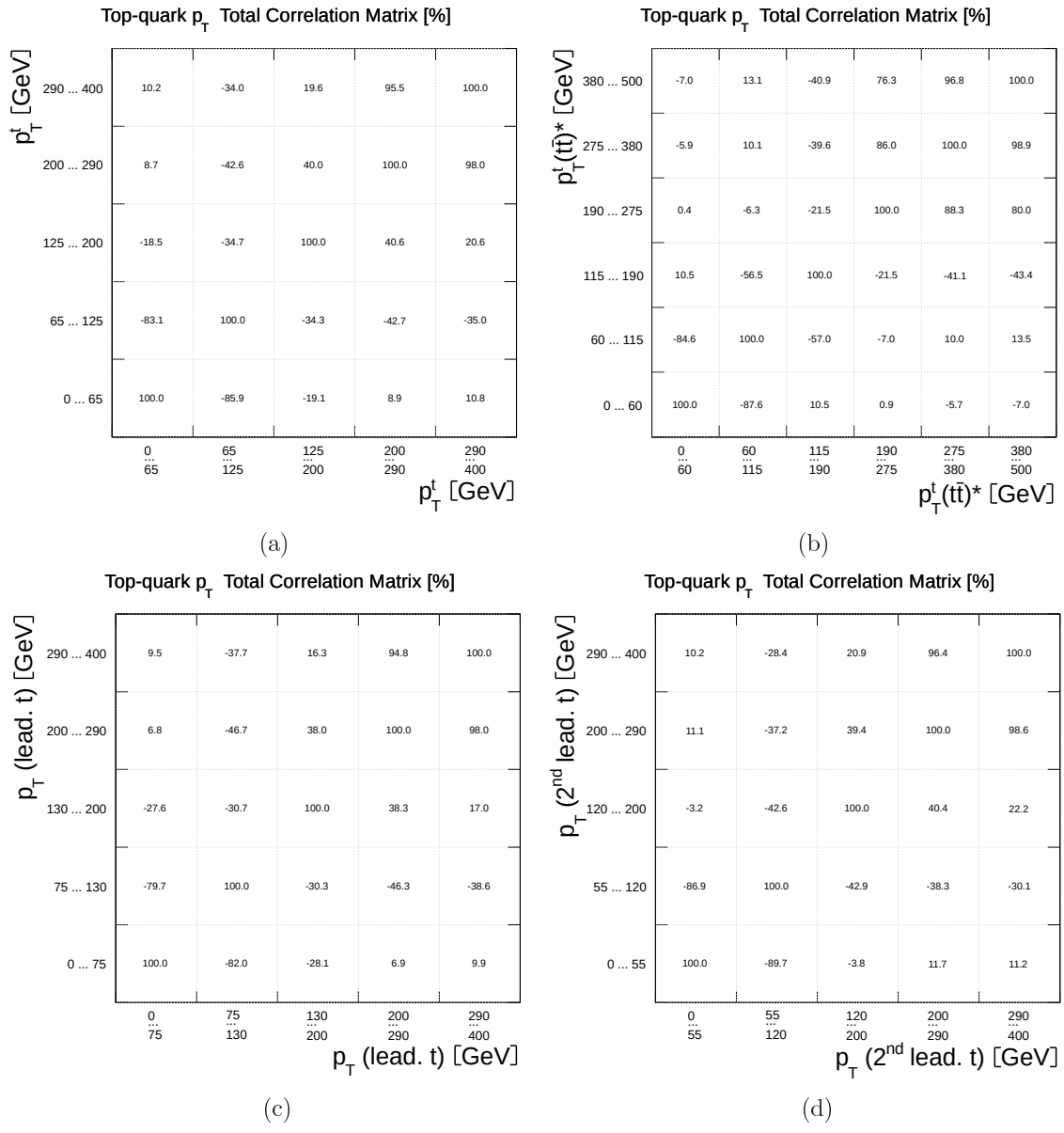


Figure E.4: Covariance matrix of the total uncertainty of the differential cross section results measured as a function of the transverse momentum of the top quark in the laboratory (a) and in the $t\bar{t}$ rest frame reference system (b), of the leading top quark (c) and trailing top quark (d).

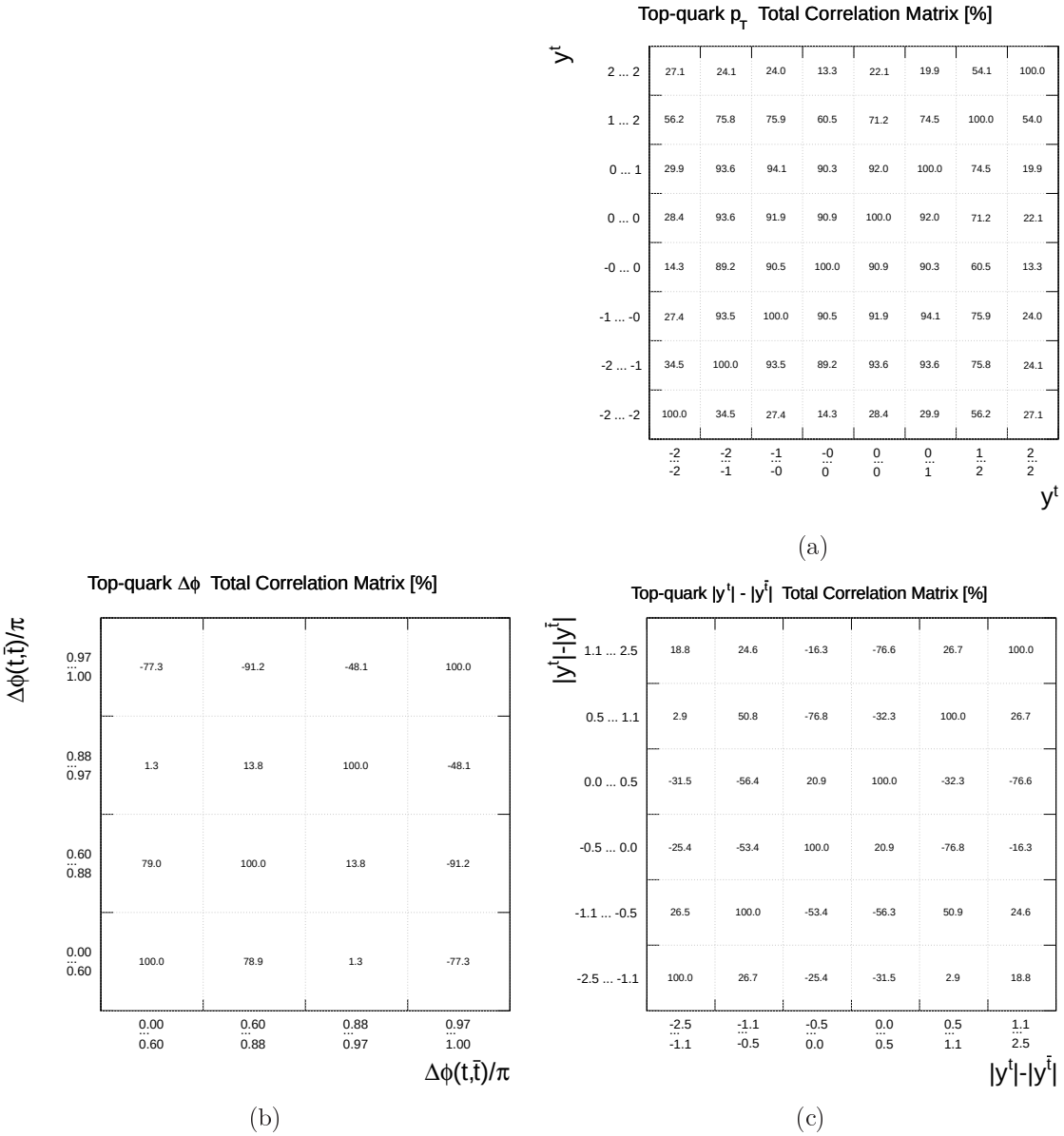


Figure E.5: Covariance matrix of the total uncertainty of the differential cross section results as a function of the top quark rapidity (a), azimuthal angle difference (b) and absolute rapidity difference (c).

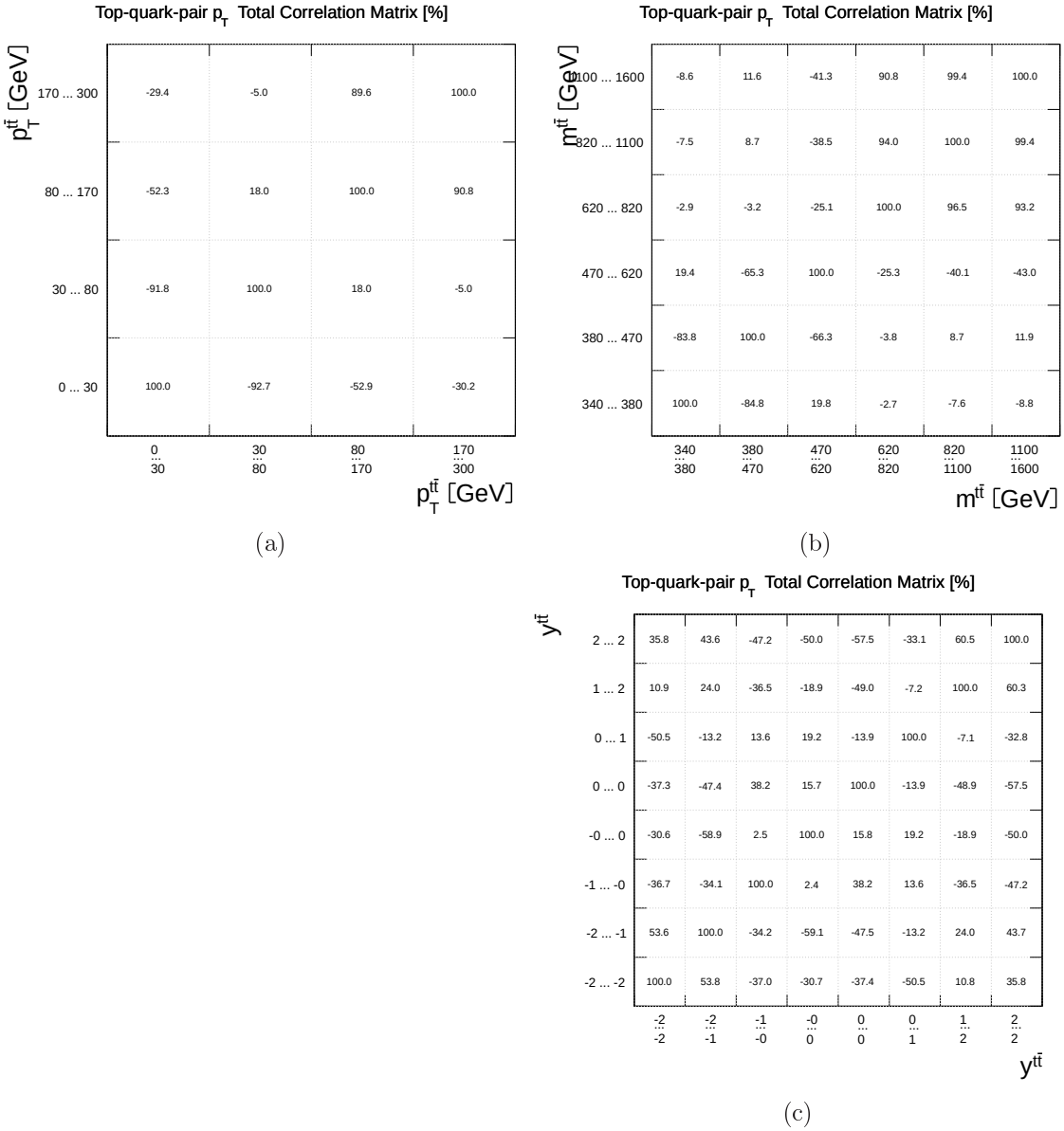


Figure E.6: Covariance matrix of the total uncertainty of the differential cross section results as a function of the top-quark-pair transverse momentum (a), invariant mass ((b) and rapidity (c).

Bibliography

[1] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.

[2] The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[3] M. Guzzi, K. Lipka, and S.-O. Moch. Top-quark pair production at hadron colliders: differential cross section and phenomenological applications with DiffTop. *arxiv:1406.0386*, 2014.

[4] The CMS Collaboration. Determination of the top-quark pole mass and strong coupling constant from the $t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 7\,\text{TeV}$. *Phys. Lett.*, B728:496–517, 2014.

[5] The CDF Collaboration. Observation of top quark production in $p\bar{p}$ collisions. *Phys. Rev. Lett.*, 74:2626–2631, 1995.

[6] The D0 Collaboration. Observation of the top quark. *Phys. Rev. Lett.*, 74:2632–2637, 1995.

[7] The CMS Collaboration. Exclusion limits on gluino and top-squark pair production in natural SUSY scenarios with inclusive razor and exclusive single-lepton searches at 8 TeV. 2014. CMS-PAS-SUS-14-011.

[8] M. E. Peskin and D. V. Schroeder. *An introduction to quantum field theory*. Advanced book classics. Addison-Wesley Publishing Company, 1995.

[9] Y. Takahashi. On the generalized Ward identity. *Nuovo Cim.*, 6:371, 1957.

[10] J. C. Ward. An identity in quantum electrodynamics. *Phys. Rev.*, 78:182, 1950.

[11] A. Purcell. Go on a particle quest at the first CERN webfest. (BUL-NA-2012-269. 35/2012):10, 2012. CERN Bulletin.

[12] J. Beringer et al. Review of particle physics (RPP). *Phys. Rev.*, D86:010001, 2012.

[13] V. Andreev et al. Measurement of multijet production in ep collisions at high Q^2 and determination of the strong coupling α_s. *arxiv:1406.4709*, 2014.

[14] P. W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.*, 145:1156–1163, 1966.

[15] V. N. Gribov and L. N. Lipatov. Deep inelastic ep scattering in perturbation theory. *Sov. J. Nucl. Phys.*, 15:438–450, 1972.

[16] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nucl. Phys.*, B126:298, 1977.

[17] Y. L. Dokshitzer. Calculation of the structure functions for deep inelastic scattering and e+e- annihilation by perturbation theory in quantum chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977.

[18] S. Alekhin, O. Behnke, P. Belov, S. Borroni, M. Botje, et al. HERAFitter, open source QCD fit project. *arxiv:1410.4412*, 2014.

[19] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al. New parton distributions for collider physics. *Phys. Rev.*, D82:074024, 2010.

[20] M. Czakon, P. Fiedler, and A. Mitov. Total top-quark pair-production cross section at hadron colliders through $O(\alpha_s^4)$. *Phys. Rev. Lett.*, 110:252004, 2013.

[21] N. Kidonakis. NNLL threshold resummation for top-pair and single-top production. *Phys. Part. Nucl.*, 45(4):714–722, 2014.

[22] G. Mahlon and S. J. Parke. Spin correlation effects in top-quark-pair production at the LHC. *Phys. Rev.*, D81:074024, 2010.

[23] J. H. Kuhn and G. Rodrigo. Charge asymmetries of top quarks at hadron colliders revisited. *JHEP*, 1201:063, 2012.

[24] The CMS Collaboration. Measurement of the ratio $B(t \to Wb)/B(t \to Wq)$ in pp collisions at $\sqrt{s} = 8\,\text{TeV}$. *Phys. Lett.*, B736:33, 2014.

[25] The CMS Collaboration. Measurement of the t-channel single-top-quark production cross section and of the $\mid V_{tb} \mid$ CKM matrix element in pp collisions at $\sqrt{s} = 8\,\text{TeV}$. *JHEP*, 1406:090, 2014.

[26] LHC Higgs Working Group. LHC Higgs cross section WG picture gallery. `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CrossSectionsFigures`. [Online, accessed 29 September 2014].

[27] G. Altarelli, B. Mele, and M. Ruiz-Altaba. Searching for new heavy vector bosons in $p\bar{p}$ colliders. *Z. Phys.*, C45:109, 1989.

[28] I. Aitchison. *Supersymmetry in particle physics: An elementary introduction*. Cambridge University Press, 2007.

[29] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, et al. LHC design report. 1. The LHC main ring. volume 1 of *LHC Design Report*. 2004. CERN-2004-003-V-1.

[30] O. Buning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, et al. LHC design report. 2. The LHC infrastructure and general services. volume 2 of *LHC Design Report*. 2004. CERN-2004-003-V-2.

[31] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl. LHC design report. 3. The LHC injector chain. volume 3 of *LHC Design Report*. 2004. CERN-2004-003-V-3.

[32] C. Lefèvre. The CERN accelerator complex. Complexe des accélérateurs du CERN. 2008.

[33] J. Wenninger. Energy Calibration of the LHC beams at 4 TeV. 2013. CERN-ATS-2013-040.

[34] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008.

[35] The ATLAS Collaboration. The ATLAS experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[36] The CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[37] A. A. Jr. Alves et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.

[38] G. Anelli et al. The TOTEM experiment at the CERN Large Hadron Collider. *JINST*, 3:S08007, 2008.

[39] The LHCf Collaboration. *LHCf experiment: Technical Design Report*. Technical Design Report LHCf. 2006. CERN-LHCC-2006-004.

[40] M. Lamont. Status of the LHC. In *Proceedings, Workshop on Discovery Physics at the LHC (Kruger 2012)*, volume 455, page 012001, 2013.

[41] The CMS Collaboration. Public CMS luminosity information. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`. [Online, accessed 01 November 2014].

[42] V. Karimäki et al. The CMS tracker system project: Technical Design Report. Technical Design Report CMS. 1997. CERN-LHCC-98-006.

[43] The CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9:P10009, 2014.

[44] The CMS Collaboration. The CMS electromagnetic calorimeter project: Technical Design Report. Technical Design Report CMS. 1997. CERN-LHCC-97-033.

[45] P. Adzic, R. Alemany-Fernandez, C. B. Almeida, N. M. Almeida, G. Anagnostou, et al. Energy resolution of the barrel of the CMS electromagnetic calorimeter. *JINST*, 2:P04004, 2007.

[46] The CMS Collaboration. Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7\,\text{TeV}$. *JINST*, 8:P09009, 2013.

[47] The CMS Collaboration. The CMS hadron calorimeter project: Technical Design Report. Technical Design Report CMS. 1997. CERN-LHCC-97-031.

[48] V. D. Elvira. Measurement of the pion Energy response and resolution in the CMS HCAL test beam 2002 experiment. 2004. CMS-NOTE-2004-020.

[49] The CMS Collaboration. The CMS muon project: Technical Design Report. Technical Design Report CMS. 1997. CERN-LHCC-97-032.

[50] The CMS Collaboration. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7\,\text{TeV}$. *JINST*, 7:P10002, 2012.

[51] The CMS Collaboration. CMS The TriDAS project: Technical Design Report, Volume 1: The Trigger Systems. volume 1 of *Technical Design Report CMS*. CERN-LHCC-2000-038.

[52] The CMS Collaboration. CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project. volume 2 of *Technical Design Report CMS*. CERN-LHCC-2002-026.

[53] The CMS Collaboration. The CMS Trigger. 2014. [In preparation].

[54] F. De Guio. The CMS data quality monitoring software: experience and future prospects. In *Proceedings, 20th International Conference on Computing in High Energy and Nuclear Physics (CHEP 2013)*, volume 513, page 032024, 2014.

[55] O. Gutsche. Validation of software releases for CMS. In *Proceedings, 17th International Conference Computing in high Energy and Nuclear Physics (CHEP 2009)*, volume 219, page 042040, 2010.

[56] Physics Performance and Dataset Group. CMS DQM graphical user interface. `https://cmsweb.cern.ch/dqm/offline`. [Unpublished. Online, accessed 7 Oct. 2014].

[57] The CMS Collaboration. CMS luminosity based on pixel cluster counting - Summer 2013 update. 2013. CMS-PAS-LUM-13-001.

[58] Top Quark Physics Object Group. Top commissioning: Operations. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/TWikiTopQuarkComWG4Operations`. [Unpublished. Online, accessed 4 Oct. 2014].

[59] S. D. Drell and T.-M. Yan. Massive Lepton-pair production in hadron-hadron collisions at high-energies. *Phys. Rev. Lett.*, 25:316–320, 1970.

[60] M. Dobbs and J. B. Hansen. The HepMC C++ Monte Carlo event record for high energy physics. *Comput. Phys. Commun.*, 134:41–46, 2001.

[61] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer. MadGraph 5: Going beyond. *JHEP*, 1106:128, 2011.

[62] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk. Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations. *JHEP*, 1303:015, 2013.

[63] S. Alioli, P. Nason, C. Oleari, and E. Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 1006:043, 2010.

[64] S. Frixione and B. R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 0206:029, 2002.

[65] G. Bevilacqua, M. Czakon, A. van Hameren, C. G. Papadopoulos, and M. Worek. Complete off-shell effects in top quark pair hadroproduction with leptonic decay at next-to-leading order. *JHEP*, 1102:083, 2011.

[66] T. Sjostrand, S. Mrenna, and P. Z. Skands. PYTHIA 6.4 physics and manual. *JHEP*, 0605:026, 2006.

[67] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al. HERWIG 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 0101:010, 2001.

[68] S. Frixione, P. Nason, and C. Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 0711:070, 2007.

[69] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani. Matching matrix elements and shower evolution for top-quark production in hadronic collisions. *JHEP*, 0701:013, 2007.

[70] B. Andersson. *The Lund model*. Cambridge University Press, 1998. Cambridge Books Online.

[71] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Hoche, et al. General-purpose event generators for LHC physics. *Phys. Rept.*, 504:145–233, 2011.

[72] D. Amati and G. Veneziano. Preconfinement as a property of perturbative QCD. *Phys. Lett.*, B83:87, 1979.

[73] The CMS Collaboration. Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} = 7$ TeV and comparison with $\sqrt{s} = 0.9$ TeV. *JHEP*, 1109:109, 2011.

[74] The ATLAS Collaboration. ATLAS tunes of PYTHIA 6 and Pythia 8 for MC11. 2011. ATL-PHYS-PUB-2011-009.

[75] S. Agostinelli et al. GEANT4: A simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[76] Top Physics Analysis Group. Top PAG reference selections. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/TWikiTopRefEventSel`. [Unpublished. Online, accessed 21 Mar. 2014].

[77] The CMS Collaboration. Particle-flow event reconstruction in CMS and performance for jets, taus, and MET. 2009. CMS-PAS-PFT-09-001.

[78] The CMS Collaboration. Commissioning of the particle-flow reconstruction in minimum-bias and jet events from pp collisions at 7 TeV. 2010. CMS-PAS-PFT-10-002.

[79] R. Fruhwirth, W. Waltenberger, and P. Vanlaer. Adaptive vertex fitting. *J. Phys.*, G34:N343, 2007.

[80] CMS Physics Validation Group. Estimating systematic errors due to pileup modeling. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupSystematicErrors`. [Unpublished. Online, accessed 25 June 2014].

[81] The CMS Collaboration. The performance of the CMS muon detector in proton-proton collisions at sqrt(s) = 7 TeV at the LHC. *JINST*, 8:P11002, 2013.

[82] The CMS Collaboration. Electron reconstruction and identification at $\sqrt{s} = 7$ TeV. 2010. CMS-PAS-EGM-10-004.

[83] W. Adam, R. Fruhwirth, A. Strandlie, and T. Todorov. Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC. In *Proceedings, 13th International Conference on Computing in High-Enery and Nuclear Physics (CHEP 2003)*, volume C0303241, page TULT009, 2003.

[84] M. Cacciari, G. P. Salam, and G. Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 0804:063, 2008.

[85] The Top Quark Physics Object Group. Top projections, or avoiding double-counting. `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuidePF2PAT`. [Unpublished. Online, accessed 06 May 2014].

[86] The CMS Collaboration. Determination of jet energy calibration and transverse momentum resolution in CMS. *JINST*, 6:P11002, 2011.

[87] K. Goebel, J. Haller, J. Ott, and H. Stadie. Jet transverse momentum resolution measurement using dijet events at $\sqrt{s} = 8\,\text{TeV}$. 2014. [Unpublished].

[88] The CMS Collaboration. Performance of missing transverse momentum reconstruction algorithms in proton-proton collisions at $\sqrt{s} = 8\,\text{TeV}$ with the CMS detector. 2012. CMS-PAS-JME-12-002.

[89] J. M. H. Salfeld-Nebgen. *Search for the Higgs boson decaying into τ-leptons in the di-electron channel*. PhD thesis, University of Hamburg. DESY-THESIS-2014-019.

[90] The CMS Collaboration. Measurement of the $\text{t}\bar{\text{t}}$ production cross section and the top quark mass in the dilepton channel in pp collisions at $\sqrt{s} = 7\,\text{TeV}$. *JHEP*, 1107:049, 2011.

[91] The CMS Collaboration. Measurement of differential top-quark pair production cross sections in pp colisions at $\sqrt{s} = 7\,\text{TeV}$. *Eur. Phys. J.*, C73:2339, 2013.

[92] The CMS Collaboration. Measurement of the $\text{t}\bar{\text{t}}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8\,\text{TeV}$. *JHEP*, 1402:024, 2014.

[93] The CMS Collaboration. Performance of b tagging at sqrt(s)=8 TeV in multijet, ttbar and boosted topology events. 2013. CMS-PAS-BTV-13-001.

[94] L. Sonnenschein. Analytical solution of ttbar dilepton equations. *Phys. Rev.*, D73:054015, 2006.

[95] The CDF Collaboration. Measurement of the top quark mass using template methods on dilepton events in proton antiproton collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev.*, D73:112006, 2006.

[96] The CMS Collaboration. Measurement of differential top-quark pair production cross sections in the lepton+jets channel in pp collisions at 8 TeV. 2013. CMS-PAS-TOP-12-027.

[97] The CMS Collaboration. Measurement of the differential top-quark pair production cross section in the dilepton channel in pp collisions at 8 TeV. 2013. CMS-PAS-TOP-12-028.

[98] The CMS Collaboration. Measurement of the $t\bar{t}$ charge asymmetry with lepton+jets events at 8 TeV. 2013. CMS-PAS-TOP-12-033.

[99] N. Bartosik. [Private communication].

[100] A. Ferroglia, B. D. Pecjak, and L. L. Yang. Top-quark pair production at high invariant mass: an NNLO soft plus virtual approximation. *JHEP*, 1309:032, 2013.

[101] H. T. Li, C. S. Li, D. Y. Shao, L. L. Yang, and H. X. Zhu. Top quark pair production at small transverse momentum in hadronic collisions. *Phys. Rev.*, D88:074004, 2013.

[102] V. Blobel. A survey of unfolding methods for particle physics. `https://www.wiki.terascale.de/images/a/a8/Blobel_Unfold_Note_20100524.pdf`, 2010. Terascale Statistics Tools School Spring 2010. [Online, accessed 5 Sep. 2014].

[103] A. Hocker and V. Kartvelishvili. SVD approach to data unfolding. *Nucl. Instrum. Meth.*, A372:469–481, 1996.

[104] K. Tackmann, A. Hoecker, and H. Lacker. TSVDUnfold software package. `http://root.cern.ch/root/html/TSVDUnfold.html`. [Online, accessed 5 Sep. 2014].

[105] D.-J. Fischer. *Inclusive neutral current ep cross sections with HERA II and two-dimensional unfolding*. PhD thesis, University of Hamburg. DESY-THESIS-2011-020.

[106] V. Blobel. An unfolding method for high-energy physics experiments. In *Advanced statistical techniques in particle physics*, pages 258–267, 2002.

[107] O. Behnke, K. Kröninger, G. Schott, and T. Schörner-Sadenius. *Data analysis in high energy physics: A practical guide to statistical methods*. Wiley, 2013.

[108] V. Blobel. Some comments on χ^2 minimization applications. In *Statistical problems in particle physics, astrophysics and cosmology. Proceedings, PHYSTAT 2003*, volume C030908, page MOET002, 2003.

[109] C. Diez Pardos and J. Kieseler. Dilepton trigger and lepton identification efficiencies for the top quark pair production cross section measurements at 8 TeV in the dilepton decay channel. 2013. CMS AN-2012/389.

[110] ATLAS Collaboration, CDF Collaboration, CMS Collaboration and D0 Collaboration. First combination of Tevatron and LHC measurements of the top-quark mass. *arxiv:1403.4427*, 2014.

[111] J. Rojo. Top quark production at the LHC as a gluon luminometer. `https://indico.cern.ch/event/217721/session/1/contribution/12`. TOP LHC WG Meeting. [Online, accessed 20 Oct. 2014].

[112] M. Görner. *Differential cross sections for top-quark-pair production in the e/μ+jets final state at $\sqrt{s} = 8$ TeV in CMS*. PhD thesis, University of Hamburg. DESY-THESIS-2014-023.

[113] The CMS Collaboration. Search for electroweak production of higgsinos in channels with two Higgs bosons decaying to b quarks in pp collisions at 8 TeV. 2014. CMS-PAS-SUS-13-022.

[114] P. Nason. Issues on the top pt distribution in ME+PS and NLO+PS interfaced to Herwig. `https://indico.cern.ch/event/301787/session/10/contribution/23`. Top LHC Working Group meeting. [Online, accessed 10 Oct. 2014].

[115] Top Physics Analysis Group. pt(top-quark) based reweighting of $t\bar{t}$ MC. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopPtReweighting`. [Unpublished. Online, accessed 08 Aug. 2014].

[116] G. D. Lafferty and T. R. Wyatt. Where to stick your data points: The treatment of measurements within wide bins. *Nucl. Instrum. Meth.*, A355:541–547, 1995.

[117] N. Kidonakis. Differential and total cross sections for top pair and single top production. In *Proceedings, 20th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2012)*, pages 831–834, 2012.

[118] The ATLAS Collaboration. Measurements of normalized differential cross-sections for $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ using the ATLAS detector. *Phys. Rev.*, D90:072004, 2014.

[119] M. Czakon and A. Mitov. Top++: A program for the calculation of the top-pair cross-section at hadron colliders. *Comput.Phys.Commun.*, 185:2930, 2014.

[120] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63:189–285, 2009.

[121] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Uncertainties on α_s in global PDF analyses and implications for predicted hadronic cross sections. *Eur. Phys. J.*, C64:653–680, 2009.

[122] The CMS and ATLAS Collaborations. Combination of ATLAS and CMS top quark pair cross section measurements in the eμ final state using proton-proton collisions at 8 TeV. 2014. CMS-PAS-TOP-14-016, ATLAS-CONF-2014-054.

[123] The ATLAS Collaboration. Measurement of the $t\bar{t}$ production cross-section using eμ events with b-tagged jets in pp collisions at $\sqrt{s} = 7\,\text{TeV}$ and 8 TeV with the ATLAS detector. *arxiv:1406.5375*, 2014.

[124] The CMS Collaboration. Measurement of the differential top-quark pair production cross sections in pp collisions at $\sqrt{s} = 8\,\text{TeV}$. In preparation.

[125] Top Physics LHC Working Group (TOPLHCWG). TOPLHCWG - Top physics LHC working group. `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TopLHCWG`. [Online, accessed 30 September 2014].

[126] M. Czakon. Latest developments in generators and NNLO computations. `https://indico.in2p3.fr/event/10114/session/2/contribution/2`. Top 2014: 7^{th} International Workshop on Top-Quark Physics. [Online, accessed 10 Oct. 2014].

[127] The CMS Collaboration. Search for Higgs boson production in association with a top-quark pair and decaying to bottom quarks or tau leptons. CMS-PAS-HIG-13-019.

[128] The CMS Collaboration. Search for the production of dark matter in association with top quark pairs in the di-lepton final state in pp collisions at $\sqrt{s} = 8\,\text{TeV}$. CMS-PAS-B2G-13-004.

[129] The CMS Collaboration. Search for third generation scalar leptoquarks decaying to top quark - tau lepton pairs in pp collisions. CMS-PAS-EXO-13-010.

[130] J. M. Campbell and R. K. Ellis. $t\bar{t}$ $W^{\pm}$ production and decay at NLO. *JHEP*, 1207:052, 2012.

[131] M. V. Garzelli, A. Kardos, C. G. Papadopoulos, and Z. Trocsanyi. $t\bar{t}$ $W^{\pm}$ and $t\bar{t}\,Z$ hadroproduction at NLO accuracy in QCD with parton shower and hadronization effects. *JHEP*, 1211:056, 2012.

[132] J. M. Campbell and R. K. Ellis. An update on vector boson pair production at hadron colliders. *Phys. Rev.*, D60:113006, 1999.

[133] K. Melnikov, M. Schulze, and A. Scharf. QCD corrections to top quark pair production in association with a photon at hadron colliders. *Phys. Rev.*, D83:074013, 2011.

[134] J. M. Campbell, R. K. Ellis, and C. Williams. Vector boson pair production at the LHC. *JHEP*, 1107:018, 2011.