BERGISCHE
UNIVERSITÄT
WUPPERTAL

Measurements of top-quark properties at 8 TeV using the ATLAS detector at the LHC

Dissertation zur Erlangung des Grades
Doktor der Naturwissenschaften (Dr. rer. nat)

vorgelegt der
Bergischen Universität Wuppertal
Fakultät für Mathematik und Naturwissenschaften

von Roger Naranjo

Wuppertal, 2017

Die Dissertation kann wie folgt zitiert werden:

urn:nbn:de:hbz:468-20170329-152229-3
[http://nbn-resolving.de/urn/resolver.pl?urn=urn%3Anbn%3Ade%3Ahbz%3A468-20170329-152229-3]

Abstract

This work presents the measurements of the charge asymmetry and top quark spin observables in $t\bar{t}$ events based on 20.3 fb^{-1} of $\sqrt{s}$ = 8 TeV proton–proton collisions recorded with the ATLAS detector at the LHC. These analyses are performed in the dilepton final state, characterized by the presence of two isolated leptons (electrons or muons). Two different reconstruction algorithms are studied and applied to reconstruct the $t\bar{t}$ system. For the measurements of the charge asymmetry, inclusive and differential measurements are performed using top-quark and lepton based observables. The differential measurements are performed as a function of the transverse momentum, mass, and boost of the $t\bar{t}$ system. For the top quark spin measurements, 15 observables, each sensitive to a different coefficient of the spin density matrix of $t\bar{t}$ production, are measured independently. Ten of these observables are measured for the first time. These measurements represent a full characterization of the spin density matrix. The measurements are corrected to particle and parton level by using the Fully Bayesian Unfolding method. At particle level, all measurements are performed in a fiducial region, while at parton level they are performed in the full phase space. No significant deviation from the Standard Model is observed for any of the measurements performed.

Die vorliegende Arbeit präsentiert Messungen der Ladungsasymmetrie und der Spin Observable des Top Quarks in $t\bar{t}$ Ereignissen auf der Basis von 20.3 fb^{-1} proton–proton Kollisionen, die bei $\sqrt{s}$ = 8 TeV mit dem ATLAS Detektor am LHC aufgezeichnet wurden. Die präsentierten Analysen wurden mit Ereignissen mit zwei isolierten Leptonen im Endzustand durchgeführt (Elektronen oder Muonen). Zwei verschiedene Algorithmen wurden untersucht, um das $t\bar{t}$ System zu rekonstruieren. Die Messung der Ladungsasymmetrie beinhaltet inklusive und differentielle Messungen unter Verwendung von Top Quark und Lepton basierenden Observablen. Die differentiellen Messungen wurden als Funktion des transversalen Momentums, der Masse und des Boosts des $t\bar{t}$ Systems durchgeführt. Für die Messung des Top Quark Spins wurden 15 Observable, die sensitiv auf verschiedene Koeffizienten der Spin-Dichtematrix der $t\bar{t}$ Erzeugung sind, einzeln analysiert. Zehn dieser Observablen wurden das erste Mal gemessen und die Messungen repräsentieren eine vollständige Charakterisierung der Spin-Dichtematrix. Die Messungen wurden auf Teilchen-, und Partonebene durch eine Fully Bayesian Unfloding Methode korrigiert. Auf Teilchenebene wurden alle Messungen im Referenzbereich durchgeführt, während die Messungen auf Partonebene im kompletten Phasenraum durchgeführt wurden. Es wurden keine signifikanten Abweichungen von den Vorhersagen des Standardmodels beobachtet.

Acknowledgments

Firstly, I would like to thanks my wife Grecia, mi amor, for standing beside me throughout my career. I can write a book with all your effort bella, and therefore I can say that you are also getting this degree. Thanks to my kids Roger Francisco and Felipe Alejandro who are the light of my life. This work is dedicated to them.

I am very grateful to the many staff and students at DESY, in particular, I would like to thanks Dr. Reinhild Fatima Yvonne Peters for her supervision, daily morning coffee, chats, funny moments, salary bottles, laughs, and of course giving me this opportunity. Many thanks to Dr. Cecile Francoise Deterre for her daily discussions, moral support, rides, second hand clothes, advice through all this work, etc, etc, and to Dr. James William Howarth (a.k.a Jimmy Billy) for all his help, advice, a full set of empty bottles, playstation, and more! without them this work would not be possible. Thanks to Ralph and Abby, my officemates/teammates/friends, it was great to work with them in such a nice office. It was wonderful to work with this team, great friendships were forged. Again, muchas gracias chicos.

Many thanks to Prof. Dr. Peter Mättig for his helpful discussions and giving me also the chance to be part of this great university.

Gracias a mis padres y hermanas, quienes me apoyaron y continúan haciéndolo en cada instante de mi vida. Los amo mucho.

Contents

Introduction

Particle physics is a branch of science that tries to answer a question which has been captivating mankind since its existence: "what is the universe made of?". It has been developed by brilliant physicists for many years, converging to one of the most powerful models to-date: the Standard Model (SM). The SM describes the elementary particles and their interactions. The properties of the particles can be used as tools to test its veracity. In particular, the heaviest particle of the SM, known as the top quark, offers a versatile way of testing the SM due to its short lifetime. The properties of the top quark can be studied experimentally through its decay products. Properties, such as spin, polarization, and charge asymmetry are excellent tools to test the SM, as well as to probe its possible extensions since new physics could affect their expected behavior. The Large Hadron Collider at CERN is considered as a top quark factory, providing large amount of top quark events at rates incredibility high with respect to previous experiments in particle physics. The amount of data allows for precision measurements and the exploration of properties that were not possible to characterize before.

The measurements of the charge asymmetry offer a precision test of the SM. Measurements performed at the Tevatron Collider at Fermilab showed large discrepancies with respect to the expected values, which led to the development of possible extensions of the SM. These discrepancies are not as large as before, due to more accurate SM predictions recently produced, however, the uncertainties on the measurements are still large. These large uncertainties on the measurements make the results compatible with both SM and its possible extensions. Precise measurements of the charge asymmetry are needed in order to constrain those extensions of the SM as well as to detect any deviation from the expected value.

The measurement of the polarization of the top quark and the spin correlation between pair-produced top and antitop quarks probes the full production and decay chain of the process, since together they characterize the spin density matrix of the process. Measurements of the polarization and spin correlation have been performed at the Tevatron and the LHC. However, these measurements only partially characterized the spin density matrix, since the observables that were used were proportional to linear combinations of the different elements of the matrix.

This work presents measurements of the charge asymmetry, as well as the first full characterization of the spin density matrix. Inclusive and differential measurements of the charge asymmetry are performed, while for the characterization for the spin density matrix the measurement of 15 different observables is performed. The measurements are performed in top and antitop quark pair events using the full 2012 dataset collected by the ATLAS detector. Both sets of measurements are performed in a fiducial region close to the detector acceptance and the full phase space. The dileptonic decay channel of the top pair system is used to perform the measurements. It offers a good balance between purity and the amount of statistics. However, reconstruction of the final state using this channel requires special consideration due to the presence of the two neutrinos. The quality of the results strongly depends on the reconstruction method used. In this work, two different reconstruction methods are studied and optimized for each of the two sets of measurements performed.

This work is divided in the following chapters. In Chapter 1 a brief theoretical overview of the SM, top-pair production, and the measured properties are presented. Chapter 2 describes the experimental apparatus: the LHC machine and the ATLAS detector, including the data used to performed the measurements. In Chapter 3, the main aspect of the simulation of hadron collisions is described. Chapter 4 describes the object definition and event selection performed in order to enrich the data sample with dileptonic events. In Chapter 5, the quality of two different reconstruction procedures are studied. Chapter 6 presents the method used to correct for the effects introduced by the ATLAS detector: Fully Bayesian Unfolding (FBU). Chapter 7 describes the systematic uncertainties affecting all the measurements. This is followed by the measurements of the charge asymmetry in Chapter 8 and the spin observables in Chapter 9, where the optimization of FBU, estimation of systematic uncertainties, and results are presented. Finally, conclusions are presented.

1 Theoretical Framework

In this chapter, the theoretical framework used to define the observables measured in this work is presented. First, the basis of the Standard Model is described, followed by the description of the particle of interest for the measurements presented: the top quark. Finally, the observables used to perform the measurements are presented.

1.1 The Standard Model

The Standard Model (SM) is a quantum field theory that describes the elementary particles and their interactions. It is one of the most successful scientific theories developed so far, with excellent precision and predictive power. The fundamental forces described by the SM are the strong, weak, and electromagnetic interactions. The particles in the SM are categorized in two different types: fermions and bosons. Fermion particles have half-integer spin, whereas bosons are particles with integer spin. The fermions are divided into three families and describe the observable matter in our universe. The fermions consist of three charged leptons, three neutrinos, and six quarks. For each of these particles, there exists an antiparticle with the same mass and opposite quantum numbers. This gives a total of 24 fermionic particles. The bosons, on the other hand, mediate the different fundamental interactions between fermions. The strong force is mediated by the gluon (g), the weak force by three weak bosons ($W^{\pm}$, Z), and the electromagnetic force by the photon (γ). The SM is structured according to the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The three groups represent the fundamental symmetries that correspond to the forces describing the interaction of particles: the *color* symmetry (C), described by the non-Abelian group $SU(3)_C$, and the symmetry with respect to the *weak isospin* (L) and *hypercharge* (Y) gauge groups, described by $SU(2)_L \times U(1)_Y$. All fermions carry a weak charge and interact with the weak force carriers $W^{\pm}$ and Z. All matter particles, except for the neutrinos, are electrically charged and interact with the electromagnetic force, mediated by the photon. Only the quarks have a color charge and undergo strong-force interactions mediated by gluons.

The SM may be considered as the combination of two theories: Quantum Chromodynamics (QCD) [1], which describes the strong interactions from the $SU(3)$ color symmetry and the electroweak (EW) theory [2–4], which describes the electroweak interactions corresponding to the

isospin and *hypercharge* symmetries. These theories are described in the following.

1.1.1 Quantum Chromodynamics

The $SU(3)_C$ symmetry group corresponds to the strong interaction, governed by the theory of QCD. The color charge occurs in three flavors: red, green and blue, and the corresponding anti-colors: anti-red, anti-green and anti-blue. QCD interactions are described by the QCD Lagrangian, defined as:

$$\mathcal{L}_{QCD} = -\frac{1}{4}F_a^{\mu\nu}F^a_{\mu\nu} + \sum_i \bar{\psi}_j(i\nabla - gA^j_\mu - m_j)\psi_j \quad \text{with} \quad F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - ig[A^a_\mu, A^a_\nu], \tag{1.1}$$

where A_μ is the gauge field (corresponding to gluons), ψ are the quarks, and g is the gauge coupling parameter, which is related to the strong coupling constant α_S as $g^2 = 4\pi\alpha_S$. The commutator $[A^a_\mu, A^a_\nu]$ is defined under the $SU(3)_C$ transformations and is responsible for the self-interaction of the gluons. The index a and j run over the color charges and the quark flavors, respectively. The parameter α_S in the one-loop approximation depends on the energy scale, Q^2, as:

$$\alpha_S(Q^2) \sim \frac{1}{\beta_0 \ln(Q^2/\Lambda^2)}, \tag{1.2}$$

where $\beta_0 = \frac{1}{12\pi}(33 - 2N_f)$ for N_f quark flavors and Λ is the QCD scale at which α_S diverges ($\Lambda \sim 200$ MeV). This results in a characteristic property of QCD: at high energies the coupling is small and perturbation theory may be used as a good approximation, but at low energies the coupling is so strong that colored objects are only found in color-neutral composite objects and perturbation theory is no longer a good approximation. This behavior (called *confinement*) implies that all quarks are always found in bound states of colorless hadrons (except the top quark), i.e, color-singlets. Two kinds of color-singlets are observed: baryons (composed of three quarks) and mesons (composed of a quark and an antiquark pair).

1.1.2 Electroweak interaction

The $SU(2)_L \times U(1)_Y$ symmetry group represents the electroweak theory (EW), which combines the $U(1)$ electromagnetic interactions with the $SU(2)$ weak interactions. Here Y refers to the weak hypercharge ($Y/2 = Q + I_3$ with Q the electric charge and I_3 the third component of the weak isospin) and L to the chirality (handedness) of the weak interaction. The electroweak Lagrangian is defined as:

$$\mathcal{L}_{QED} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}(iD_\mu - m)\psi \quad \text{with} \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu], \tag{1.3}$$

where $D_\mu\psi = \partial_\mu\psi + ieA_\mu\psi$ is the covariant derivative applied to the fermion ψ and A_μ corresponds the gauge fields of the theory. The commutator $[A_\mu, A_\nu]$ is defined under the $SU(2)_L \times U(1)_Y$ transformations and it is related to the self-interaction of the gauge fields. There are four gauge bosons associated with the $SU(2)_L \times U(1)_Y$ group: W^a_μ ($a = 1, 2, 3$) and B_μ. The non-Abelian nature of $SU(2)$ implies that the gauge bosons W^a_μ interact with each other. Linear combinations of these four gauge bosons leads to the observed $\gamma(A)$, Z, $W^\pm$ bosons:

$$A = \sin\theta_W W^3 + \cos\theta_W B, \tag{1.4}$$

$$Z = \cos\theta_W W^3 - \sin\theta_W B, \tag{1.5}$$

where θ_W is the weak mixing angle related to the coupling constants g_W of the $SU(2)_L$ and g'_W of the $U(1)_Y$ groups through:

$$\sin\theta_W = \frac{g'_W}{\sqrt{g_W^2 g'^2_W}}. \tag{1.6}$$

The bosons $W^\pm$ are defined by the linear combination

$$W^\pm = \frac{1}{\sqrt{2}}(W^1 \mp iW^2). \tag{1.7}$$

The weak interaction is experimentally observed to violate parity [5] and the theory is constructed such that the weak interaction only acts on left-handed fermions. Left-handed fermions are formed into weak isospin doublets ($SU(2)$) whereas right-handed fermions form singlets ($U(1)$). The eigenstates that interact through the weak interaction, known as weak eigenstates, are different from the physically observed mass eigenstates. These states are related by the Cabbibo-Kobayashi-Maskawa (CKM) matrix. The CKM matrix is parametrized in terms of three mixing angles and a complex phase that causes CP-violation in the SM. The element $|V_{ab}|^2$ of the CKM matrix is proportional to the probability of a transition from a quark a to a quark b. One possible parametrization of the CKM matrix is shown in Eq. 1.8, as well as the elements that are proportional to the quark transitions. In the expression, c_{ij} and s_{ij} are the cosines and sines of the Euler angles (θ_{12}, θ_{23}, θ_{13}) used in the parametrization, and δ_{13} is the complex phase.

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix} \tag{1.8}$$

1.1.3 Spontaneous symmetry breaking

In order to keep the EW theory invariant under transformation of the gauge group (gauge invariance), the gauge bosons must be massless. However, observations show that the Z and W bosons are massive. A spontaneous symmetry breaking, known as the Higgs mechanism, was introduced in order to solve this problem [6–8]. The Higgs mechanism introduces a complex scalar field ϕ that follows the $SU(2)$ symmetry. The Lagrangian describing its kinematics is defined as:

$$\mathcal{L}_{Higgs} = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi), \tag{1.9}$$

where $V(\phi) = \lambda(\phi^\dagger\phi)^2 - \mu^2\phi^\dagger\phi$ is called the Higgs potential and is invariant under $SU(2)$ gauge transformation. If $\mu^2 < 0$, the state of minimum energy will be that with $< \phi >= 0$ and the potential will preserve the symmetries of the Lagrangian. In this case, the theory is simply a scalar theory with a massless photon and a scalar field ϕ with mass μ. However, if $\mu^2 > 0$, the field ϕ will acquire the value of minimum energy of the potential, known as the vacuum expectation value (called *vev*), of $< \phi >= v/\sqrt{2} = \sqrt{\frac{\mu^2}{2\lambda}}$. When the field is situated at a minimum, the $SU(2)\times U(1)$ symmetry is broken to form the $U(1)$ symmetry, providing masses to the W and Z bosons and giving a non-complex field h which corresponds to the Higgs boson. This boson was discovered in 2012 by the ATLAS and CMS Collaborations at the Large Hadron Collider (LHC) [9, 10].

The mass of the fermions is given by the interactions between the fermions of different chiralities and the Higgs field. This term is known as the Yukawa coupling and is given by:

$$\mathcal{L}_{Yukawa} = g_f(\bar{\Psi}_L \phi \Psi_R + \bar{\Psi}_R \phi \Psi_L). \tag{1.10}$$

Each fermion has a different coupling constant (g_f) which is related with the mass of the fermions as $m = g_f v\sqrt{2}$.

1.1.4 Limitations of the SM

The SM is a powerful theory that describes an enormous variety of physics processes. However, some limitations of the SM indicates that it needs to be extended or embedded into broader theories. These theories could unify the strong interaction, weak interaction and gravity (which is not described by the SM). Also, astronomical observations show gravitational effects of non-observable matter (called Dark Matter) that cannot be described by the matter particles contained within the SM [11].

Another problem occurs when the fundamental value of some physical parameter, such as a coupling constant or a mass, is very different from the value that is measured in an experiment (effective value). This occurs because the effective value is related to the fundamental value by its renormalization, which applies corrections to it. The renormalized values of most parameters are close to their fundamental values. In some parameters, it appears that there are large cancellation between the fundamental quantity and the quantum corrections, for example the quantum corrections that the mass of the Higgs boson receives due to the presence of quantum correction from the top quark. The corrections depend on the interplay between the positive contributions of the Higgs boson and negative contribution from the top quark (which is larger than the Higgs mass) leading to the need to "fine tune" the scale Λ at which the SM is valid, which is an intrinsic part of the renormalization procedure. This problem i s commonly referred to as the *hierarchy problem*.

Beyond the SM theories (BSM) have been proposed to address these limitations. For instance, in one of the models, all the particles from the SM get a super partner particle that would share the same mass and internal quantum numbers besides spin with the SM particles. This assumption provides a solution to the hierarchy problem, since it gives extra contributions to the renormalized values, as well as a candidate for dark matter. This model is known as SUSY (Super Symmetry) [12]. Grand unification theories (GUT) have been also proposed, where the three gauge interactions of the SM which define the electromagnetic, weak, and strong interactions, are merged into one single force at very high energies. These theories predict new gauge mediators (such as Z' and W') and also predict the masses of the neutrinos[1] due to the generalization of the group theory in which it is realized [13, 14]. Other more specific models, are focused on providing explanations for deviations observed in SM measurements. Some of these models are studied in this work.

[1] In the SM the neutrinos are assumed to be massless.

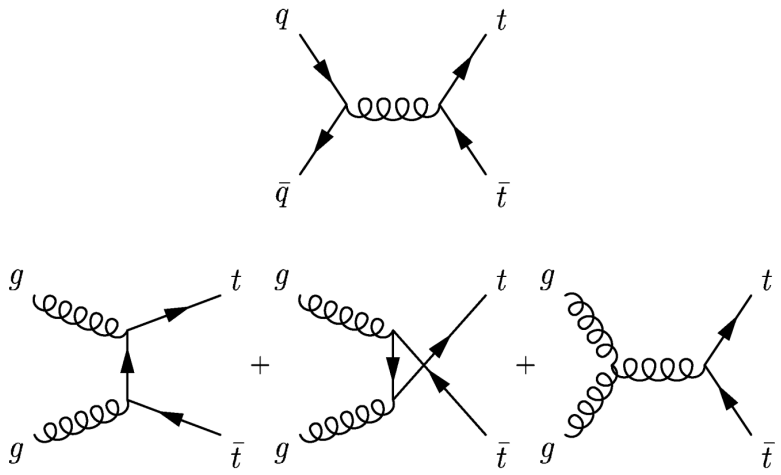


Figure 1.1 – *Leading Feyman diagrams for the $t\bar{t}$ production.*

1.2 The Top Quark

In 1973, Kobayashi and Maskawa proposed a new family of quarks in the SM in order to explain the experimentally measured CP violation in electroweak decays. Before this, only two families were considered. The new family of quarks introduced the bottom and the top quark. After the discovery of the b-quark in 1977 in dimuon resonances [15], intense searches of the top quark started to take place in the colliders available at the time. After many years of setting limits on the top mass in several experiments, it was finally discovered in 1995 at the Tevatron by the CDF and D0 Collaborations [16, 17]. With a mass of $\sim$170 GeV, it is the heaviest elementary particle observed so far.

1.2.1 Production at hadron colliders

At hadron colliders, top quarks are mainly produced in pairs ($t\bar{t}$) via the strong interaction. They can also be produced singly (single top) via the electroweak interaction at about half of the $t\bar{t}$ production rate. The leading order (LO) Feyman diagrams of the $t\bar{t}$ production are shown in Fig. 1.1.

The strong interaction between colliding hadrons (protons or antiprotons) is described by perturbative QCD. A hard scattering process between some constituents of the hadrons, called partons (quarks or gluons), occurs in the collision. Each of the partons carries a fraction x of the mometum of their parent hadrons. The hadron collision can be described by separating the short distance (hard scattering) and long distance interaction. The separation makes use of theorem known as the *factorization theorem*, which allows to write the cross-section of $\sigma_{pp\to X}$ as:

$$\sigma_{pp\to X} = \sum_{a,b} \int dx_a dx_b f_a(x_a, \mu_F^2) f_b(x_b, \mu_F^2) \times \hat{\sigma}_{ab\to X}(x_a p_a, x_b p_b, \mu_F^2 \mu_R^2), \tag{1.11}$$

where $f_i(x_i, \mu_F^2)$ $(i = a, b)$, corresponds to the probability of finding a parton i within the proton carrying a fraction x_i of the momentum of the proton. It is known as the parton distribution

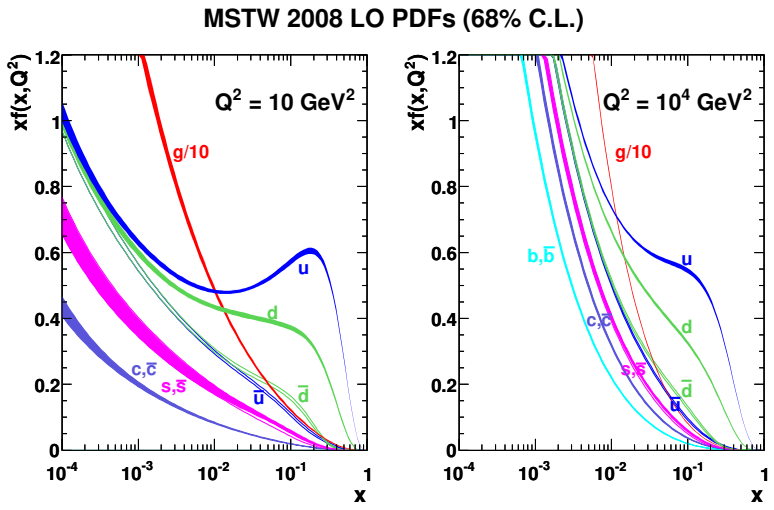


Figure 1.2 – *MSTW parton distribution functions for gluons and quarks at* $Q^2 = 10$ GeV^2 *(left) and* $Q^2 = 10^4$ GeV^2 *(right) [18, 19].*

function (PDF). The PDFs do not depend on a particular process, they are estimated by combining information from deep-inelastic scattering experiments and hadron colliders. The term $\hat{\sigma}_{ab\to X}(x_a p_a, x_b p_b, \mu_F^2, \mu_R^2)$ is the cross section for the partonic process $ab \to X$. The *factorization scale* μ_F^2 describes the scale of the interaction. An additional *renormalization scale* μ_R^2 accounts for higher-order corrections. For calculations and simulations, both scales are set to the typical transferred momentum Q^2 of the process studied. In the case of $t\bar{t}$ production the scale is often chosen to be equal to the top quark mass, $\mu_F^2 = \mu_R^2 = m_t^2$. In $t\bar{t}$ production, the fraction of momentum x carried by each of the colliding partons follows the relationship:

$$\sqrt{x_1 x_2 s} \geq 2m_t, \tag{1.12}$$

and by assuming that both partons carry a similar fraction of the energy $x_1 \approx x_2 = x$, the relation becomes:

$$x \geq \frac{2m_t}{\sqrt{s}}. \tag{1.13}$$

At the LHC for a center-of-mass energy $\sqrt{s} = 8$ TeV, this value is $x \geq 0.04$. Figure 1.2 shows the MSTW parton distribution functions for gluons and quarks within the proton [18, 19]. The probability of gluon collisions is significantly larger than for any other parton in this region. This leads to the fact that the production of top quark pairs at the LHC is dominated by the gluon fusion processes by approximately 85%. This is not the case at $p\bar{p}$ collisions such as the Tevatron, where the same calculation yields $x \approx 0.18$. In this region, quark-antiquark annihilation is the dominant production process and both quarks are expected to be valence quarks.

The total $t\bar{t}$ production cross section at the LHC is calculated at next-to-next-to leading order in QCD to be $\sigma_{t\bar{t}} = 177.3^{+10.1}_{-10.8}$ pb^{-1} at $\sqrt{s} = 7$ TeV, $\sigma_{t\bar{t}} = 252.89^{+13.30}_{-14.52}$ pb^{-1} at $\sqrt{s} = 8$ TeV, and $\sigma_{t\bar{t}} = 832^{+40}_{-46}$ pb^{-1} at $\sqrt{s} = 13$ TeV, for a top quark mass of 172.5 GeV [20]. Measurements have been performed by both ATLAS and CMS collaborations, yielding a combined value of $\sigma_{t\bar{t}} = 173.3 \pm 10.1$ pb^{-1} at $\sqrt{s} = 7$ TeV [21]. Measurements at $\sqrt{s} = 8$ TeV yield $\sigma_{t\bar{t}} = 242.9 \pm 8.8$ pb^{-1}

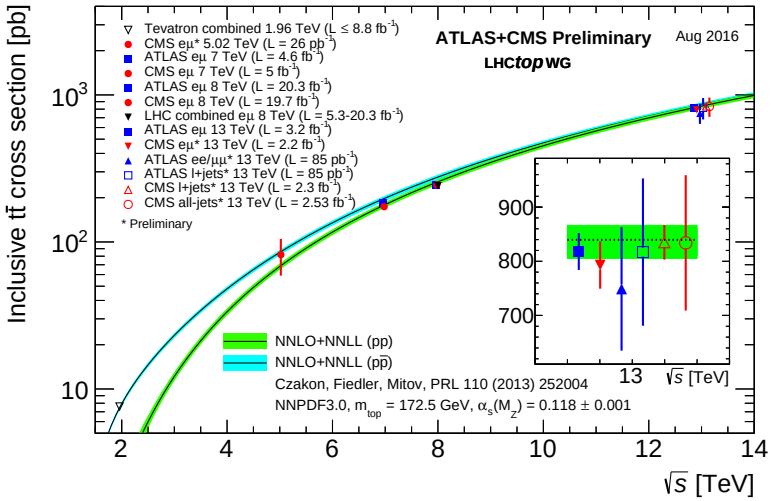


Figure 1.3 – *Summary of LHC and Tevatron measurements of the top-pair production cross-section as a function of the center-of-mass energy compared to the NNLO QCD calculation. The measurements and the theory calculations are quoted considering a mass of the top 172.5 GeV. Measurements made at the same center-of-mass energy are slightly offset for clarity [26].*

(ATLAS) [22] and $\sigma_{t\bar{t}} = 239 \pm 13\,\text{pb}^{-1}$ (CMS) [23], while at 13 TeV the measurement yield $\sigma_{t\bar{t}} = 818 \pm 36\,\text{pb}^{-1}$ (ATLAS) [24] and $\sigma_{t\bar{t}} = 792 \pm 42\,\text{pb}^{-1}$ (CMS) [25]. All measurements are in agreement with the SM predictions. Measurements and predictions for Tevatron and LHC are summarized in Fig. 1.3 as a function of the center-of-mass energy $\sqrt{s}$.

1.2.2 Decay

Since the lifetime of the top quark is 5×10^{-25} s, which is shorter than the hadronization scale, it decays before the hadronization can occur. The top quark decays in almost all cases via electroweak interaction into a b quark and a W boson and the CKM matrix element $|V_{tb}|$ is 0.99914 ± 0.00005 [27]. The W boson decays either leptonically (to one lepton and one neutrino) or hadronically (to one quark and one antiquark). Therefore, the pair decays are categorized based on the decay of the W boson as follows:

- **Fully hadronic**. Both W decay hadronically, leading to a final state with six quarks.
- **Semileptonic**. One W decays hadronically and one leptonically, leading to a final state of three quarks, one charged lepton and one neutrino.
- **Dileptonic**. Both W decay leptonically, leading to a final state with two quarks, two charged leptons, and two neutrinos.

The probability of having a fully hadronic decay is approximately 45%, while is it 44% for a semileptonic decay, and 10% for dileptonic decay. The fully hadronic and semileptonic channels offer large amount of events, at expenses of large background contributions. The purity[2] is higher

[2]The purity of the sample is quantified by the signal over background ratio

for the semileptonic channel than for the fully hadronic channel, due the charged lepton present in the process that helps to have a clear signature of the event. The dileptonic channel offers a pure sample since the background contribution is smaller than in the other channels. The leptonic τ decay in this channel is usually considered as part of the electron or muon channel. In the analyses presented in this work, the dileptonic decay channel is used.

SM backgrounds for dileptonic decays

Other SM processes can have two charged leptons, two jets and two neutrinos in the final states or can imitate the signatures of dileptonic events in the detector. For instance, the quarks or gluons can be identified as the b-quarks coming from the $t\bar{t}$ decay. Another example is the case of the neutrinos which are not detected but they are inferred using the transverse energy of other particles. In dileptons decays, the major background contributions arise from the Drell–Yan process, diboson production and single top production in the Wt channel. This will be discussed in detail in Chapter 4.

1.2.3 Properties

The most important characteristic of the top quark is that, due to its large mass, the lifetime is a order of magnitude shorter than the time scale of the strong interaction. This implies that the top quark decays before hadronization occurs. Thus, the quantum numbers of the top quark are transferred to its decay products and not diluted by hadronization effects. This is the only quark that posses such characteristic and allows for a deep study of its properties.

Since its discovery, the top quark's properties have been thoroughly studied. For example, the top mass has been measured with a precision well below 1 GeV by combining the measurements of the Tevatron and the LHC [28].

The top quark allows for the study of the Wtb vertex through the measurement of the W helicity [29–31].[3] This measurement represents a good test of the SM since W bosons with positive helicity are forbidden in top-quark decays. The measurements performed to-date are consistent with the SM expectations. The coupling of the top quark with the Higgs boson, i.e the Yukawa coupling, is of particular interest in BSM models since it is the highest quantum contribution that the self-coupling constant of the Higgs boson receives. The Yukawa coupling has been studied very recently at the LHC using top pair production in association with the Higgs boson. Currently, there are limits and the first measurement is expected soon with data collected during the period between 2015 and 2018 (called Run 2) [32, 33].

The charge asymmetry in top quark pair production has been also studied for several years at different experiments at the Tevatron and at the LHC, such as D0, CDF, ATLAS, and CMS [34–40]. Precise measurements of this quantity are needed in order to clarify discrepancies between the different results. The measurement of this quantity is the first objective of this work and it will be more thoroughly described in Sec. 1.3.

The spin correlation in top quark pairs and the top quark polarization have been measured by different experiments at the LHC and at the Tevatron [41–45]. The measurements were found

[3] The helicity of a particle is the projection of its spin into an axis defined by its momentum direction.

to be consistent with the SM model predictions. However, the spin correlations and polarization have been measured with a small set of observables that only partially describes the spin density matrix (described in Sec. 1.4). The full characterization of the spin density matrix is the second objective of this work and it will be thoroughly described in Sec. 1.4.

1.3 Charge asymmetry in top pair production

Most of the measurements in the top quark sector have yielded results compatible with SM predictions. The only exception to this was the measurement of the *forward-backward* (FB) asymmetry in $t\bar{t}$ production at the Tevatron in 2011, where a significant discrepancy with respect with the SM prediction was found [46] at that time. The equivalent of the forward-backward asymmetry in symmetric collision such as pp collision at the LHC, is called the *charge asymmetry*.

1.3.1 Generalities

The charge asymmetry describes an excess of top quarks over antitop quarks in certain kinematic regions. In the SM, it occurs from higher-order corrections in QCD calculations of $t\bar{t}$ production through quark-antiquark annihilation. Specifically, it originates from the interference between the Born amplitude for $q\bar{q} \to t\bar{t}$ and its one-loop box correction, and from the interference between initial- and final-state radiation (ISR, FSR) in $q\bar{q} \to t\bar{t}g$ (see Fig. 1.4). The total asymmetry depends on the contributions that these two interference provides individually. The Born-box interference provides a positive contribution to the asymmetry while the ISR/FSR interference provides a negative contribution. The sum of these effects results in top quarks being emitted preferentially in the direction of the incoming quark and top antiquarks being emitted preferentially in the direction of the incoming antiquark. The amount of the ISR and FSR is correlated with the transverse momentum of the $t\bar{t}$ system, since the amount of radiation increases with the transverse momentum. At low transverse momentum ($p_{\mathrm{T},t\bar{t}} < 25$ GeV), the asymmetry is positive since the Born-box contribution dominates, while it is negative at high transverse momentum ($p_{\mathrm{T},t\bar{t}} > 25$ GeV) since the ISR/FSR interference dominates. The overall asymmetry is expected to be positive in the SM. The asymmetry also gets contributions from QED, when the diagrams analogous to Fig. 1.4 (by changing the gluons by photons) interfere with the QCD diagrams. This contribution is approximately 20% for the forward-backward asymmetry at the Tevatron and approximately 10% for the charge asymmetry at the LHC. These contributions depend on the relative amount of the $u\bar{u}$ and $d\bar{d}$ initial-states in $t\bar{t}$ production. Electroweak contributions can also be estimated using the same diagrams as in Fig. 1.4 but changing gluons to Z bosons. This contribution is found to be of the order of 1% at the Tevatron and negligible at the LHC due to the small weak coupling.

The forward-backward charge asymmetry can be defined in top pair production as the asymmetry in the distribution of the angle θ_t between the outgoing top quark and the incoming quark in the $t\bar{t}$ rest-frame:

$$A_{FB} = \frac{N(\cos\theta_t > 0) - N(\cos\theta_t < 0)}{N(\cos\theta_t > 0) + N(\cos\theta_t < 0)}, \tag{1.14}$$

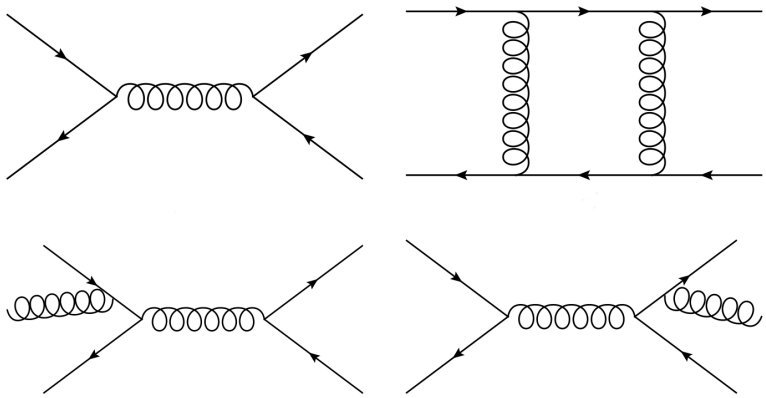


Figure 1.4 – *Main sources of the QCD charge asymmetry in $t\bar{t}$ productions: interference of Born and box diagrams (top- left and right, respectively) and interference of initial state and final state radiation diagrams (bottom- left and right, respectively).*

where N corresponds to the number of the events with $\cos\theta_t > 0$ or $\cos\theta_t < 0$. The charge asymmetry only appears in $q\bar{q}$ processes. However, it can be measured in both $p\bar{p}$ and pp collisions. The measurement can be naturally performed in $p\bar{p}$ collision, as is the case for the Tevatron, due to the asymmetric initial state and the fact the $q\bar{q}$ annihilation is the dominant process for $t\bar{t}$ production. In pp collisions, which is the case for the LHC, the measurement is slightly more complicated to perform due to the symmetric nature of the collision and the fact that $q\bar{q}$ annihilation is not the dominant process to generate $t\bar{t}$ events.

1.3.2 Observables at the Tevatron

The $\cos\theta_t$ angle is used in the definition of asymmetry, but in some cases it is not possible to know a priori the direction of the incoming quark. At the Tevatron, with its colliding protons and antiprotons, both quarks and antiquarks appear as valence quarks. As these particles are coming from different directions, it is possible to redefine the asymmetry in terms of the rapidity y[4] of the top and antitop quarks with respect to the interacting beams as:

$$A_{FB}^{t\bar{t}} = \frac{N(\Delta y > 0) - N(\Delta y > 0)}{N(\Delta y > 0) + N(\Delta y < 0)} \text{ with } \Delta y = y_{\text{top}} - y_{\text{antitop}}. \tag{1.15}$$

In the dileptonic channel, it is also possible to use the two leptons to define a lepton-based asymmetry (called leptonic asymmetry). Since the momentum of the top quark is transferred to its decay product, the leptons are likely to have a similar momentum direction. This is not always the case and leads to a dilution of the asymmetry compared to the $t\bar{t}$ asymmetry. The leptonic asymmetry, instead of using the rapidity, uses an approximation of the rapidity for massless particles, known as pseudorapidity[5] η. This quantity transforms in an additive way under boosts along the

[4] The rapidity is defined as $y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z}$, where E is the energy of the particle and p_z the momentum along the z-axis, defined along the collision axis.

[5] The pseudorapidity is defined as $\eta = -ln\left(\tan\frac{\theta}{2}\right)$, where θ is the angle between the momentum of the particle and the z-axis.

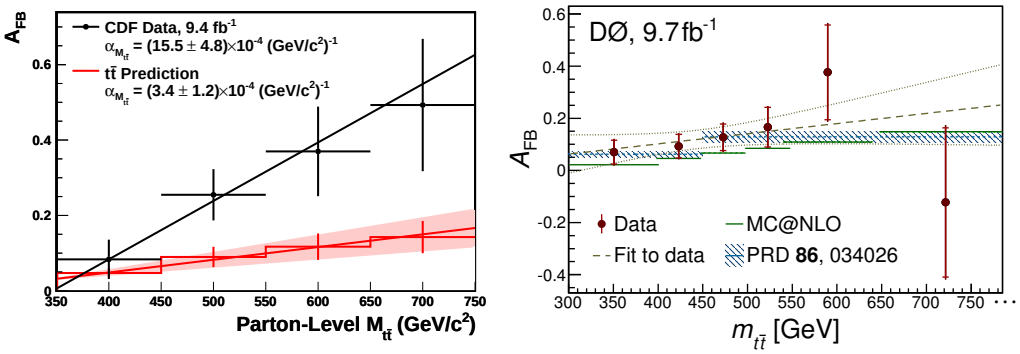


Figure 1.5 – *Measured $A_{FB}^{t\bar{t}}$ as a function of the mass of the $t\bar{t}$ system. The asymmetry value is expected to increase with respect to the energy of the system. Small deviations from the SM predictions are observed [34, 35].*

z axis and its difference is thus Lorentz invariant since the leptons are assumed to be massless. The leptonic asymmetry is constructed using the difference of the pseudorapidities of the leptons as:

$$A_{FB}^{\ell\ell} = \frac{N(\Delta\eta > 0) - N(\Delta\eta < 0)}{N(\Delta\eta > 0) + N(\Delta\eta > 0)} \text{ with } \Delta\eta = \eta_{\ell} - \eta_{\bar{\ell}}. \tag{1.16}$$

The measurements of $A_{FB}^{t\bar{t}}$, using half of the Tevatron Run 2 dataset, showed discrepancies between results and expectations. The measurement performed by CDF showed a deviation of more than 3σ with respect to the SM predictions [46]. However, the measurement performed by D0 was compatible with the SM prediction and the measurements performed by CDF. [47]. The same behavior was observed in the leptonic asymmetry, however it was of the order of 2σ. Different BSM models were proposed to explain the discrepancy. At the present, the tensions have been reduced in the final results using the full Tevatron dataset. CDF reports $A_{FB}^{t\bar{t}} = 0.164 \pm 0.047$, corresponding to a 1.7σ excess over the SM prediction of $A_{FB}^{t\bar{t}} = 0.088 \pm 0.006$ at NLO+EW [48], whereas D0 finds agreement within 1σ with $A_{FB}^{t\bar{t}} = 0.106 \pm 0.030$ [49]. In the case of the leptonic asymmetry, CDF is reporting a value of $A_{FB}^{\ell\ell} = 0.072 \pm 0.060$ [50], while D0 is reporting $A_{FB}^{\ell\ell} = 0.123 \pm 0.056$ [35], both compatible with the SM prediction of $A_{FB}^{\ell\ell} = 0.040 \pm 0.004$. The asymmetry as a function of the mass of the $t\bar{t}$ system in the measurements performed by CDF also showed a discrepancy with respect to the theoretical predictions, as shown in Fig. 1.5. These deviations are not fully covered by the experimental and theoretical uncertainties. CDF reports a deviation with respect to the SM, which is not observed by D0. However, both measurements are compatible between each other. This deviation puzzled the physics community for several years. Recently, the theoretical predictions estimated at next-to-next-to-leading order (NNLO) [51] have lowered the tension, making the predictions compatible with the differential results reported by D0 and about $\sim 2\sigma$ below that of the values reported by CDF.

1.3.3 Observables at the LHC

At the LHC it is not possible to define a forward-backward asymmetry due to the symmetric collision (pp collision). However, due to the fact that the valence quarks carries a larger fraction of the proton's energy than the sea quarks, and no valence antiquark are present in the collision,

the direction of the $t\bar{t}$ system is expected to be the same as the valence quark in the laboratory frame. In this scenario, a forward top quark in the $t\bar{t}$ rest frame has on average a larger absolute rapidity $|y|$, in the laboratory frame, than the backward antiquark. This makes it possible to define a forward-central asymmetry, called charge asymmetry, using the difference of the rapidities of the top and antitop in absolute value as:

$$A_C^{t\bar{t}} = \frac{N(\Delta|y|>0) - N(\Delta|y|<0)}{N(\Delta|y|>0) + N(\Delta|y|<0)} \text{ with } \Delta|y| = |y_{\text{top}}| - |y_{\text{antitop}}|. \tag{1.17}$$

As in the case of the Tevatron, the asymmetry can also be measured in dileptonic events by using the leptons coming from the top quarks. The leptonic asymmetry at the LHC is constructed using the difference in the pseudorapidities of the leptons in absolute value as:

$$A_C^{\ell\ell} = \frac{N(\Delta\eta>0) - N(\Delta\eta<0)}{N(\Delta\eta>0) + N(\Delta\eta>0)} \text{ with } \Delta|\eta| = |\eta_\ell| - |\eta_{\bar{\ell}}|. \tag{1.18}$$

The charge asymmetry has been measured by ATLAS and CMS at 7 and 8 TeV, with no significant deviation from the SM predictions [36–40]. At 7 and 8 TeV, the $t\bar{t}$ asymmetry has been measured by ATLAS and CMS in the semileptonic and dileptonic channel [36, 37, 39]. Using dileptonic events the leptonic asymmetry have been also measured. The measurements performed by ATLAS at 8 TeV using dileptonic events, correspond to the work presented in this work. Differential measurements at the LHC are intended to enhance the asymmetry value by discrimanting $t\bar{t}$ events produced via $q\bar{q}$ annihilation over those produced by gg fusion and have been performed by ATLAS and CMS [36, 37].

1.3.4 Charge asymmetry in BSM models

Possible explanations of the discrepancies that were observed between the experimental results and the SM predictions, is that new physics were affecting the asymmetry value. Various extensions of the SM were proposed to explain the excess of the measured asymmetry at the Tevatron. In these models, the interference of the LO new physics diagram with the SM diagrams is the mechanism that contributes to the charge asymmetry value. The parameters of the models are, in most of the cases, chosen to be compatible with the $t\bar{t}$ cross section measurements and constraints from flavor physics and other electroweak precision tests. However, a relatively wide range of asymmetries can be predicted. Despite the fact that the measurement performed at D0 and CDF are compatible with the recent NNLO calculations, the uncertainty on the measurement is very large, which makes them compatible with both SM and BSM predictions. Figure 1.6 shows the prediction of the A_{FB} and A_C for different BSM models, as well as the measurements performed at the LHC and at the Tevatron. Some of the models are compatible with the current measurements while other models are disfavored. A particular model studied in this work is the color-octet vector G model (showed in cyan). In this model, a new particle G is exchanged in the s channel via flavor-diagonal couplings. It gives an amplitude that interferes with the SM $q\bar{q} \to t\bar{t}$ Born diagram. The corresponding contribution to the charge asymmetry is proportional to the product of axial couplings with light and top quarks $g_A^{u,d} g_A^t$. Depending on the relative sign of the couplings and on the mass M_G, the contribution to the asymmetry can be positive or negative.

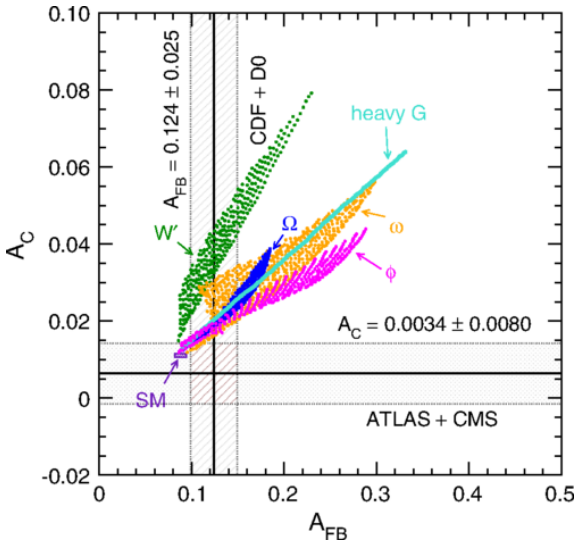


Figure 1.6 – *Comparison of predictions for the inclusive asymmetries A_{FB} and A_C for several simple models, together with the experimental measurements obtained at the Tevatron and LHC [52].*

1.4 Polarization and spin correlation

The degree of polarization of a particle, P, is the expectation value of the helicity of the particle, given by:

$$P = < \vec{S} \cdot \hat{p} >, \tag{1.19}$$

where $\vec{S}$ is the spin of the particle and $\hat{p}$ is a unit vector in the direction of the momentum of the particle. This polarization is commonly referred to as the longitudinal polarization. A fermion which is maximally polarized has a projection of its spin onto its momentum direction of $\pm\hbar/2$; the positive (negative) projection is called parallel (antiparallel). Under parity transformation $\mathcal{P}$, the degree of longitudinal polarization transform as $\mathcal{P}(P) = -P$, therefore, the longitudinal polarization is parity-odd and any mechanism that generates a longitudinally polarized particle must violate parity symmetry. Due to parity conservation in the strong production of $t\bar{t}$, the SM predicts a very small longitudinal polarization of the top quark.

Top quarks produced via QCD are expected to be almost unpolarized but their spins are expected to be correlated. The degree to which the spin of the top quark is aligned with the spin of the antitop quark, in $t\bar{t}$ production, is known as the *spin correlation*. The spin correlation (C) can be expressed as the ratio of the difference of spin-aligned pairs and spin anti-aligned pairs in a given frame of reference:

$$C = \frac{N(\uparrow\uparrow) + N(\uparrow\uparrow) - N(\uparrow\downarrow) - N(\downarrow\uparrow)}{N(\uparrow\uparrow) + N(\uparrow\uparrow) + N(\uparrow\downarrow) + N(\downarrow\uparrow)}. \tag{1.20}$$

This parameter depends on which reference is used to quantify the top and the antitop quark spin direction, commonly referred to as the *spin-quantization axis*.

During decay, the spin information of the top quark is transferred to the W boson (and to its decay product) and the b quark. The degree to which these particles retain the spin information of top quarks is known as *spin-analyzing power* (α) and is defined between the values of -1 and

1. The α parameter relates the measured values of the polarizations and spin correlation with its true values. The spin-analyzing power is known at NLO [53]. For charge leptons, $\alpha = 0.998$, which is the value closest to 1 across the different particles involved in the $t\bar{t}$ decay. This means that by using charge leptons a precise estimation of the spin of their parents top quarks can be obtained.

Some BSM models used to explain the anomalous forward-backward asymmetry include a parity violating chiral coupling to the top quarks that predict a larger longitudinal polarization of the top quark in $t\bar{t}$ production than expected in the SM. They can also alter the spin correlation of the top and antitop quark by modifying the production mechanism of the $t\bar{t}$ pair, as well as the $t\bar{t}$ decay by which the spin information is accessed. The spin correlation measurement has for example been used to exclude masses of the SUSY partner of the top quark in a minimal SUSY model [54].

A precise measurement of the polarization and spin correlation offers a way to validate the SM predictions. It also offers a way to distinguish between the various BSM models if for instance a non-negligible polarization is measured or a spin correlation significantly larger than the SM is observed.

1.4.1 Generalities

The squared spin density matrix of $t\bar{t}$ production can be decomposed in the following way [55]:

$$|M|^2 \propto \tilde{A} + \mathbf{B}^+ \cdot \mathbf{s}_1 + \mathbf{B}^- \cdot \mathbf{s}_2 + C_{ij} s_{1i} s_{2j}. \tag{1.21}$$

The coefficient $\tilde{A}$ is completely spin independent and fixes the cross section of $t\bar{t}$ production. $\mathbf{B}^{\pm}$ corresponds to the polarization vectors for the top and antitop quark and C_{ij} is the spin correlation matrix. The vectors $\mathbf{s}_{1/2}$ denote the spin vectors of the top/antitop quark. Since the leptons are an excellent tool to access the spin information, the Eq. 1.21 can be rewritten in term of their angular distributions with respect to the production angles of the leptons that come from the top and antitop quark as:

$$\frac{1}{\sigma}\frac{d^2\sigma}{d\cos\theta_a d\cos\theta_b} = \frac{1}{4}(1 + B_+^a \cos\theta_+^a + B_-^b \cos\theta_-^b - C(a,b)\cos\theta_+^a \cos\theta_-^b), \tag{1.22}$$

where B^a, B^b and $C(a,b)$ are the polarization and spin correlation along the spin quantization axes a and b. The angles θ^a and θ^b are defined as the angles between the momentum direction of a top quark decay particle in its parent top quark's rest frame and the axis a or b. The subscript $+(-)$ refers to the top (antitop) quark. From Eq. 1.22 it is possible to retrieve the following relation for the spin correlation between the axes a and b:

$$C(a,b) = \alpha_+ \alpha_- C_0 = -9 < \cos\theta_+^a \cos\theta_-^b >, \tag{1.23}$$

where α_+ (α_-) are the spin-analyzing power for the positive (negative) lepton and C_0 is the spin correlation of the top quarks. By integrating out one of the angles in Eq. 1.22, the differential cross section, which only depends linearly on the polarization of the top quark, is obtained:

$$\frac{1}{\sigma}\frac{d\sigma}{d\cos\theta_\pm^a} = \frac{1}{2}(1 + \alpha B_0^a \cos\theta_\pm^a) = \frac{1}{2}(1 + B^a \cos\theta_\pm^a). \tag{1.24}$$

where B_0 represents the true value of the polarization of the top quark. This means the differential cross section has a linear dependence on the polarization B^a, from which also follows:

$$B^a = 3 < \cos\theta^a > . \tag{1.25}$$

1.4.2 Choice of basis

By choosing a specific basis $\{\hat{k}, \hat{n}, \hat{r}\}$, the coefficients from Eq. 1.22 can be further decomposed into:

$$\begin{aligned} B_i^{\pm} &= b_k^{\pm}\hat{k}_i + b_n^{\pm}\hat{n}_i + b_r^{\pm}\hat{r}_i, & (1.26)\\ C_{ij} &= c_{nn}\hat{n}_i\hat{n}_j + c_{rr}\hat{r}_i\hat{r}_j + c_{kk}\hat{k}_i\hat{k}_j \\ &+ c_{rk}(\hat{r}_i\hat{k}_j + \hat{k}_i\hat{r}_j) + c_{kn}(\hat{k}_i\hat{n}_j + \hat{n}_i\hat{k}_j) + c_{rn}(\hat{r}_i\hat{n}_j + \hat{n}_i\hat{r}_j) \\ &+ \epsilon_{ijl}(c_r^{\mathrm{CP}}\hat{r}_l + c_k^{\mathrm{CP}}\hat{k}_l + c_n^{\mathrm{CP}}\hat{n}_l), & (1.27) \end{aligned}$$

where $b_l (l = 1, 2, 3)$ are the polarization coefficients with respect to their corresponding axis and c_{lm} are the spin correlation coefficients. The superscripts $+(-)$ indicate the top (antitop) quark.

In order to perform a measurement of all the coefficients of the spin density matrix, a set of spin-quantization axes must be defined. They can be defined[6] as follows [55]: a helicity axis ($\hat{\mathbf{k}}$), which is the momentum direction of the top quark in the $t\bar{t}$ rest frame, an axis transverse to the production plane ($\hat{\mathbf{n}}$), which is spanned by the direction of the top quark and the beam axis in the laboratory frame, and an orthonormal quantization axis ($\hat{\mathbf{r}}$), chosen to be orthogonal to the other two. They are mathematically defined as:

$$\hat{\mathbf{n}} = \frac{1}{r}\left(\hat{\mathbf{p}} \times \hat{\mathbf{k}}\right), \quad \hat{\mathbf{r}} = \frac{1}{r}\left(\hat{\mathbf{p}} - y\hat{\mathbf{k}}\right), \tag{1.28}$$

$$y = \hat{\mathbf{k}} \cdot \hat{\mathbf{p}}, \quad r = \sqrt{1 - y^2}, \tag{1.29}$$

where the axis $\hat{\mathbf{p}}$ defines the direction of one of the proton beams.

A set of observables, proposed in Ref. [55], can be used to measure the different coefficients of the spin density matrix directly. Each of the observables is sensitive to one coefficient of the spin density matrix. These observables are obtained from angular distributions defined by the double-differential cross section for $t\bar{t}$ production (see Eq. 1.22). The observables consist of single and double $\cos\theta$ distributions and are summarized in Tab. 1.1. The table describes the names used for the different observables, their definition, as well as a shorter notation to which coefficient of the spin density matrix they are connected. The observables consisting of sums and differences of two $\cos\theta^a \cos\theta^b$ terms will be referred to as cross-correlations. The expectation values and their SM predictions at NLO are shown in Tab. 1.2.

All $\cos\theta$ observables are calculated by taking the dot product of the positive/negative lepton momentum direction, indicated by a subscript +/−, and one of the spin quantization axes, indicated by a superscript for the angle θ. For instance, $\hat{\mathbf{k}} \cdot \ell_+ = cos\theta_+^k$ describes the $\cos\theta$ observable of the positive lepton and the helicity axis. For the calculation of $\cos\theta$ with negative leptons, each of the axes is multiplied by a minus sign. The necessity for this additional factor is intuitive in the helicity axis, as $-\hat{\mathbf{k}}$ is the direction of the antitop quark in the $t\bar{t}$ rest frame. The multiplication

[6] This is an arbitrary choice, different spin-quantization axes lead to different observables.

Name	Expectation value	Observable	Proportional to
Helicity polarization	B^k	$\cos\theta^k_\pm$	$b_1^\pm$
Transverse polarization	B^n	$\cos\theta^n_\pm$	$b_2^\pm$
R polarization	B^r	$\cos\theta^r_\pm$	$b_3^\pm$
Helicity correlation	$C(k,k)$	$\cos\theta^k_+\cos\theta^k_-$	c_{kk}
Transverse correlation	$C(n,n)$	$\cos\theta^n_+\cos\theta^n_-$	c_{nn}
R correlation	$C(r,r)$	$\cos\theta^r_+\cos\theta^r_-$	c_{rr}
R-Hel Sum	$C(r,k)+C(k,r)$	$\cos\theta^r_+\cos\theta^k_- + \cos\theta^k_+\cos\theta^r_-$	c_{nn}
R-Hel Diff	$C(r,k)-C(k,r)$	$\cos\theta^r_+\cos\theta^k_- - \cos\theta^k_+\cos\theta^r_-$	c_n^{CP}
Trans-Hel Sum	$C(n,k)+C(k,n)$	$\cos\theta^n_+\cos\theta^k_- + \cos\theta^k_+\cos\theta^n_-$	c_{rk}
Trans-Hel Diff	$C(n,k)-C(k,n)$	$\cos\theta^n_+\cos\theta^k_- - \cos\theta^k_+\cos\theta^n_-$	$-c_r^{CP}$
Trans-R Sum	$C(n,r)+C(r,n)$	$\cos\theta^n_+\cos\theta^r_- + \cos\theta^r_+\cos\theta^n_-$	c_{rk}
Trans-R Diff	$C(n,r)-C(r,n)$	$\cos\theta^n_+\cos\theta^r_- - \cos\theta^k_+\cos\theta^r_-$	c_k^{CP}

Table 1.1 – *The top spin observables, their definitions, and to which coefficient of the spin density matrix they are proportional are listed in this table.*

with the sign of the scattering angle y for the transverse and r axis is necessary as the initial state is symmetric due to the two proton beams, but in the definition of the axes a fixed beam direction is used. Without applying this factor, the corresponding $\cos\theta$ distributions would always be a flat line. Table 1.3 summarizes the corrections to the axes, that are used for the $\cos\theta$ calculation for positive and negative leptons.

In the previous measurements of ATLAS [43, 44] and CMS [45] from the LHC, and CDF [41] and D0 [42] from the Tevatron, only a subset of these coefficients has been measured using different techniques and approaches. All results measured by the different experiments are in agreement with the SM. The polarization was measured along the helicity and transverse axes, which provides two coefficients b_i [42, 43, 45, 56]. The polarization along the transverse axis has been only measured at the Tevatron by D0 [42]. For the measurements performed at 7 TeV at the LHC, ATLAS and CMS used a template fit to extract the polarization value. The template fit shows a good statistical uncertainty, but a large model dependency. At 8 TeV, CMS used a different approach to extract the polarization value, known as unfolding [45]. The unfolding procedure is intended to reduce the model dependency, but with an impact on the statistical uncertainty. D0 also measured the polarization values by using a template fit. In the case of the spin correlation, the measurements performed by ATLAS and CMS at 7 and 8 TeV were sensitive to linear combinations of different correlation coefficients, making it impossible to characterize precisely the spin density matrix [44, 57]. An example of the observables used in these measurements is $\Delta\phi$. This observable is defined as the difference of the azimuthal angle between the two leptons in the laboratory frame and it is sensitive only to the linear combination of the c_{nn} and c_{kk} parameters.

In this work, the set of 15 observables that characterizes the full spin density matrix are measured. Out of all the observables considered for the measurement, only the polarization and the

Expectation values	NLO predictions
B_+^k	0.0030 ± 0.0010
B_-^k	0.0034 ± 0.0010
B_+^n	0.0035 ± 0.0004
B_-^n	0.0035 ± 0.0004
B_+^r	0.0013 ± 0.0010
B_-^r	0.0015 ± 0.0010
$C(k,k)$	0.318 ± 0.003
$C(n,n)$	0.332 ± 0.002
$C(r,r)$	0.055 ± 0.009
$C(n,k) + C(k,n)$	0.0023
$C(n,k) - C(k,n)$	0
$C(n,r) + C(r,n)$	0.0010
$C(n,r) - C(r,n)$	0
$C(r,k) + C(k,r)$	-0.226 ± 0.004
$C(r,k) - C(k,r)$	0

Table 1.2 – *List of the observables and corresponding expectation values measured in this analysis. The SM predictions at NLO are also shown [55]; expectation values predicted to be 0 at NLO are exactly 0 due to term cancellations. The expectation values can be obtained from the corresponding observables using the relations from Eq. (1.23) and (1.25). The uncertainties on the predictions refer to scale uncertainties only; values below 10^{-4} are not quoted. The first six observables correspond to the polarization of the top and antitop quarks along the various axes, the other nine to the spin correlations. In order to distinguish between the correlation observables, the correlations using only one axis are referred to as spin correlations and the last six as cross correlations.*

Axis name	Positive lepton	Negative lepton
Helicity	$\hat{\mathbf{k}}$	$-\hat{\mathbf{k}}$
Transverse	$\text{sign}(y)\,\hat{\mathbf{n}}$	$-\text{sign}(y)\,\hat{\mathbf{n}}$
R	$\text{sign}(y)\,\hat{\mathbf{r}}$	$-\text{sign}(y)\,\hat{\mathbf{r}}$

Table 1.3 – *Axes taken for the $\cos\theta$ calculation by taking the dot product with the momentum direction of the positive (+) or negative (-) lepton.*

spin correlation along the helicity axis have been measured at 7 TeV by ATLAS and CMS and at 8 TeV only by CMS. D0, apart from measuring the polarization along the hecility axis, has recently measured the polarization in the transverse axis ($\hat{n}$), i.e. the B_+^n and the B_-^n observable, and also along the beam axis, with no deviation from the SM predictions. As result, 10 observables are measured for the first time in this work.

2

Experimental Apparatus

In this chapter the experimental apparatus is presented. First, the Large Hadron Collider complex is described. This is followed by the description of the ATLAS detector which collected the data analyzed in this work. The description of the dataset used to performed the analyses is also presented.

2.1 The LHC

The Large Hadron Collider (LHC) at CERN is the largest accelerator ever built. It is a proton–proton (pp) collider, located in the tunnel where the Large Electron-Positron collider (LEP) used to operate between the years 1989 and 2000. The LHC tunnel has a circumference of 27 km and is situated at the Swiss-French border. The LHC is designed to reach a nominal center-of-mass energy of 14 TeV. It consists of two superconducting adjacent rings and four collision points where the main detectors called ATLAS, CMS, LHCb, and ALICE are located. The protons travel through the rings in opposite directions and cross each other at the collision points. The protons travel through 1232 dipoles to maintain a circular path and 392 quadripoles to keep the proton-beam focused, which increases the likelihood of interactions between them at the collision points.

The particle acceleration process involves several steps and starts which the proton source, which is a bottle of hydrogen gas. An electric field is used to strip hydrogen atoms of their electrons to yield protons. The protons are injected in groups (bunches) to a linear accelerator called LINAC2, that accelerates them up to 50 MeV. The resulting proton-beam is then transferred into the first circular accelerator, called BOOSTER (Proton Synchrotron Booster), that increases the energy up to 1.4 GeV in 1.2 seconds. Subsequently, the beam is transferred to the PS (Proton Synchrotron) and the SPS (Super Proton Synchrotron). Both are circular accelerators that increase the energy of the beam up to 26 GeV and 450 GeV, respectively. Finally, the beam is injected to the LHC where it reaches the collision energy. The LHC not only accelerates bunches of protons, but also heavy ions, which are used to produce special environments to be analyzed by different experiments.

Figure 2.1 shows the different components and collision points of the LHC complex. There are four collision points, where the protons interacts and the data is collected by the different

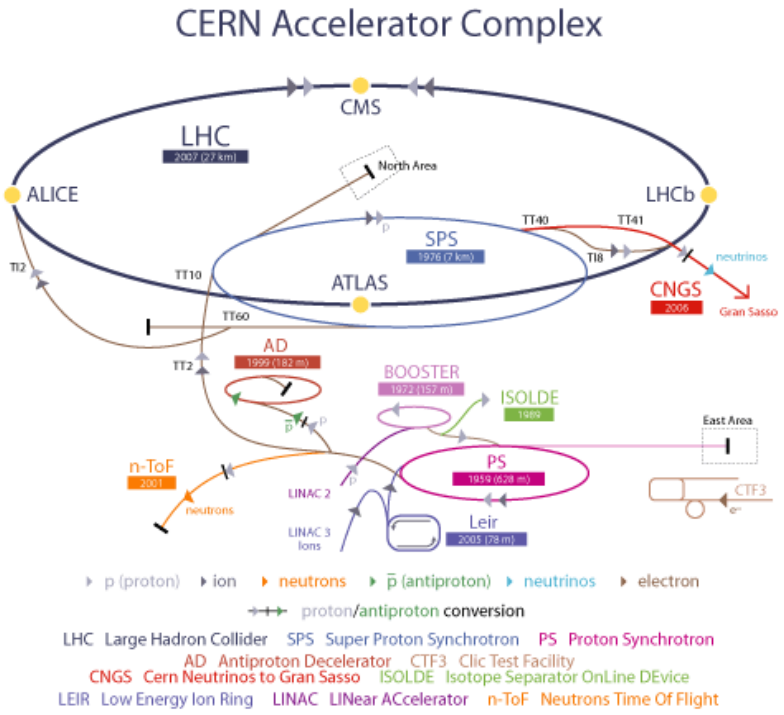


Figure 2.1 – *Schematic diagram of the CERN accelerator complex. The different components and interaction points are shown [58].*

detectors:

- ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid): they are general propose detectors, designed to study the Standard Model of particle physics, its possible violations and explore its frontiers.
- ALICE (A Large Ion Collider Experiment): this is an experiment for heavy ions. The main purpose of this experiment is to study the quark-gluon plasma.
- LHCb (Large Hadron Collider beauty): this is an experiment specialized in b-quark physics. The main purpose of this detector is to study b-hadron production, CP violation, and decays.

Besides the high center-of-mass energy of the LHC, it also outperforms previous accelerators. The integrated luminosity (L) is related with the number of events per second (N), for a given process, as follows:

$$N = L \times \sigma_{int}, \tag{2.1}$$

where σ_{int} is the cross-section of the given process. The integrated luminosity depends on the instantaneous luminosity ($\mathcal{L}$), as:

$$L = \int \mathcal{L} dt, \tag{2.2}$$

where $\mathcal{L}$ only depends on the parameters of the beam and can be estimated through the following equation:

$$\mathcal{L} = \frac{n_b f_r n_1 n_2}{2\pi\sigma_x\sigma_y}, \tag{2.3}$$

where n_1 and n_2 are the bunch populations (protons per bunch) in beams 1 and 2 respectively, f_r is the revolution frequency of the LHC, n_b is the number of bunch pairs colliding in each revolution and σ_x and σ_y characterize the horizontal and vertical beam widths.

Timeline of the LHC and ATLAS data taking

The LHC was designed to operate at 14 TeV and after finishing construction in 2008 the accelerator was expected to be operating at a center-of-mass energy of 14 TeV by the end of 2008. Unfortunately, on September 19th, 2008 a fault occurred in the electrical connection between a dipole and a quadrupole, resulting in mechanical damage and release of approximately 6 tonnes of liquid helium breaking the vacuum conditions of the beam pipe. This event shaped the evolution of the LHC data taking during the following years, making it possible to increase the center-of-mass energy only up to 8 TeV until the first long shutdown in 2013.

The damages were repaired in 2009 and the LHC started to operate with a collision energy of 900 GeV, which allowed ATLAS to obtain approximately 9 μb^{-1} for calibration purposes using minimum bias measurements. In 2010, the energy of the proton beam was gradually increased up to 7 TeV. The number of bunches and the bunch intensity[1] of the machine were also increased up to 368 bunches and 1.2×10^{11} protons per bunch, where an integrated luminosity of 45 pb^{-1} was recorded by the ATLAS detector. The same energy was maintained until 2011, while the number of bunches was increased to 1380 and the bunch intensity to 1.45×10^{11} protons per bunch, making it possible to deliver an integrated luminosity of 5.46 fb^{-1}, of which 5.05 fb^{-1} was recorded by the ATLAS detector.

In 2012, the collision energy was increased to 8 TeV, maintaining the number of bunches but increasing the bunch intensity to 1.7×10^{11} protons per bunch. A total integrated luminosity of 22.6 fb^{-1} was delivered by the LHC while 21.3 fb^{-1} was recorded by the ATLAS detector.

A long shutdown took place during 2013 and 2014, where several repairs took place and allowed the machine to reach the center-of-mass energy of 13 and 14 TeV. This new period of data taking is known as Run 2. The first collision at 13 TeV was delivered in summer of 2015, recording during that year a total integrated luminosity of 3.9 fb^{-1}. In 2016, the data taking continued to take place until the end of the year and more than 40 fb^{-1} were collected. Figure 2.2 shows the delivered luminosity as a function of the time for the 2011, 2012, 2015, and 2016 data-taking periods [59].

2.2 The ATLAS detector

The ATLAS detector is the largest detector at the CERN complex. The detector, which is shown in Fig. 2.3, has a cylindrical shape with a length of 44 m and a radius of 11 m and weighs approximately 7000 tons [60]. The high energy and luminosity of the LHC makes it possible to study a

[1] Number of protons per bunch.

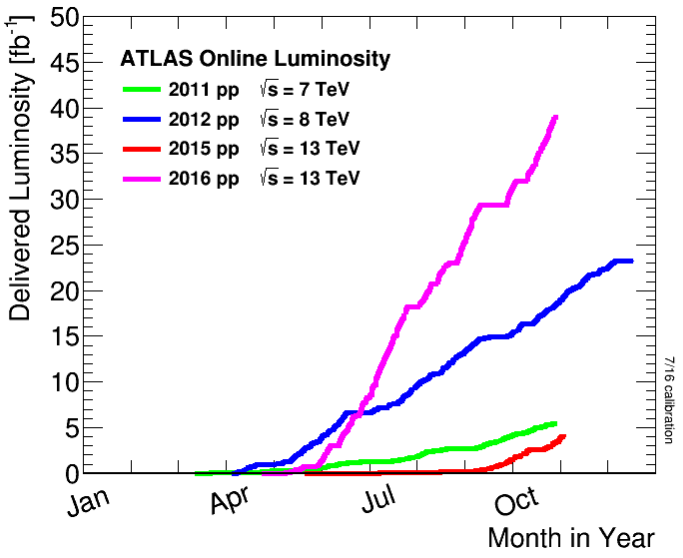


Figure 2.2 – *Delivered luminosity by the LHC as a function of time [59].*

wide range of processes, which includes searches for new physics as well as high precision measurements of QCD and EW processes. These studies present a serious experimental difficulty, as it implies that every candidate event for new physics will on average be accompanied by 23 inelastic events per bunch-crossing at the designed center-of-mass energy and luminosoty. A good identification of final objects, such as the charged particle momentum, missing transverse energy (E_T^{miss}) and secondary vertices, as well as a good spatial coverage is very important.

The detector is built in several layers of sub-detectors, each devoted to the measurement of different properties for different types of particles and physical signatures. The sub-detectors are grouped into three main systems: the Inner Detector (ID), followed by a calorimeter system (electromagnetic and hadronic) and finally by the muon spectrometer. In the following sections these components of the ATLAS detector are described.

2.2.1 Coordinate system

The ATLAS coordinate system is a cartesian right-handed system, with its origin at the nominal collision point. The proton beam defines the z-axis and the $x-y$ plane is defined transverse to the beam direction. The x-axis is defined pointing to the center of the LHC ring and the y-axis pointing to the surface. The azimuthal angle (ϕ) is measured in the $x-y$ plane and the polar angle (θ) in the $y-z$ plane. The radial coordinate r is measured in the $x-y$ plane.

Fundamental quantities, such as the particles momenta, are measured with respect to the $x-y$ plane due to initial momentum in that plane. In the $x-y$ plane, the momentum of the system is zero before the collision, and it will remain zero after the collision. This gives a natural constraint on the initial conditions which is not possible to impose along the z-axis since the initial momentum of the system is unknown. The transverse momentum (p_{T}) is defined as,

$$p_{\text{T}} = p\sin(\theta). \tag{2.4}$$

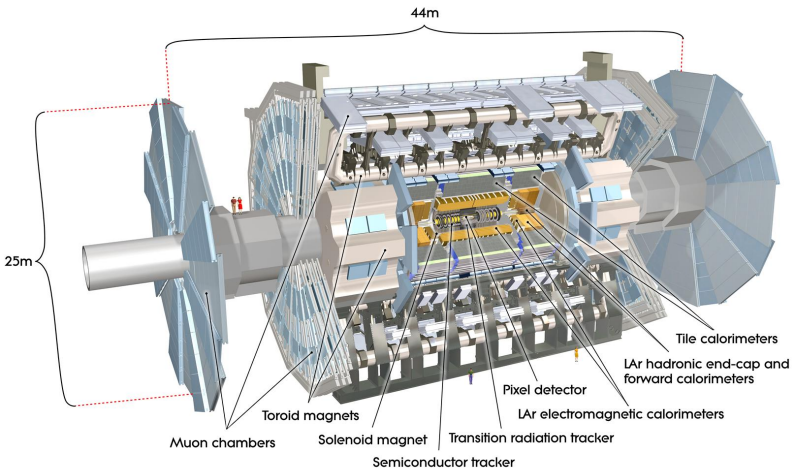


Figure 2.3 – *Schematic diagram of the ATLAS detector. All the different components are specified [60].*

The distances are measured in the $\eta - \phi$ plane (with η is the pseudorapidity) using:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \tag{2.5}$$

2.2.2 Inner detector

The Inner Detector (ID) is the closest detector to the collision point and covers a range up to $|\eta| < 2.5$. It is responsible for measuring particle trajectories (known as tracks) and transverse momenta [61, 62]. The ID is designed to provide a high momentum resolution $\sigma p_T/p_T = 0.05\% \cdot p_T \oplus 1\%$, which gives a resolution of about 2% (3%) for a 25 GeV (50 GeV) charged particle [60]. It consists of three independent sub-detectors: the Pixel Detector, the Semi Conductor Tracker (SCT) and the Transition Radiation Tracker (TRT). Its layout is shown in Fig. 2.4.

The inner and the external part of the ID are realized with different technologies. The inner part consists of high-resolution semiconductor pixel and micro-strip sensors (Pixel and SCT). When a charged particle traverses the silicon layer, it generates a number of free electron-hole pairs along its track, which are proportional to the energy loss of the incident particle. An electric field is applied to the sensors, causing the electrons and holes to drift towards the electrodes in opposite directions. The collected charges induce a signal in the pixel and micro-strips electrodes, which allows for the determination of the position of the particle passing through the detector. The difference between pixels and micro-strips sensors consists mainly in their geometry: pixels sensors are closely spaced pads, capable of good 2-dimensional reconstruction, while micro-strips sensors give a good resolution along one privileged coordinate that depends on the position of the sensor.

The outer part of the detector (TRT) consists of straw-tube tracking detectors. The straw tubes are parallel to the beam direction in the barrel and arranged radially in wheels in the end-cap region. They are interleaved with a transition radiation material (radiator), which is characterized by having a widely varied index of refraction. Charged particles traversing the radiator produce

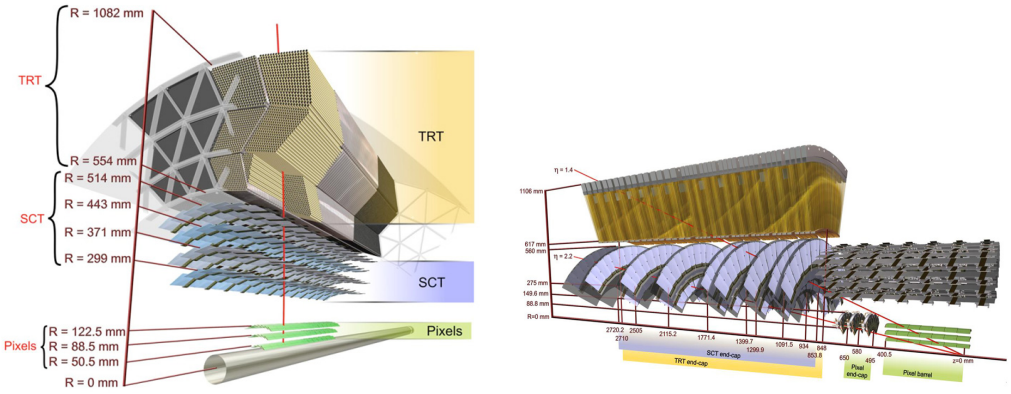


Figure 2.4 – *Schematic diagram of the Inner Detector (ID) barrel (left) and end-cap (right) of the ATLAS detector. All the different components are specified [61, 62].*

X-ray photons, known as transition radiation. The transition radiation photons interact with the molecules in the gas, ionizing the gas and creating free electrons, which then move towards an anode, where the current is measured. The intensity of the transition radiation is proportional to the Lorentz γ factor of the incoming particle. Therefore, the radiation pattern can be used for particle identification, providing a separation of electrons from hadrons.

The ID is immersed in a magnetic field of 2 T along the z-direction provided by the barrel solenoid, which allows for the measurement of charged particles p_{T} using the curvature produced by the magnetic field. The solenoid consist of a single layer of aluminum conductor, cooled down using Helium at a temperature of $\sim 4.5K$. It is located between the inner detector and the calorimeter. In the following, the different components of the ID are detailed.

Pixel detector

The nearest detector to the collision interaction point is the Pixel Detector, which allows for the measurement of the particles impact parameters and to reconstruct any secondary vertices,[2] due to the decay of short living particles like B-hadrons [63]. The Inner Tracker has three layers of silicon pixels, placed at 5, 9 and 12 cm from the center of the detector, and five rings on each side with an inner radius of 11 cm and an outer radius of 30 cm, to complete the angular coverage. It covers the region $|\eta| < 2.5$. The pixel detector has an intrinsic accuracy of 10 μm in the $r-\phi$ plane and 115 μm along the z-axis in the barrel region. In the disks placed end-cap regions perpendicular to the beam axis, the intrinsic accuracy is 10μm ($r-\phi$) and 115 μm (r). The silicon sensors are arranged in three layers and are designed to provide the highest granularity around the interaction point. The pixel detector has approximately 80.4 million readout channels.

[2] The secondary vertices are defined as the vertices produced by a secondary decay of the particles produced after the hard-scattering interaction.

SCT

The Semi Conductor Tracker system is designed to provide precision measurements of momentum, impact parameter, and vertex position in the intermediate radial range [64]. It covers the same η-region as the pixel detector and consists of eight strip layers that are normally crossed by each track. The SCT barrel consists of four layers of micro-strips situated at 299, 371, 443, and 514 mm from the beam. The micro-strips sensors are placed back-to-back with a small stereo angle of 40 mrad between them (stereo strips), that allows for the measurement of the azimuthal coordinate. In the end-caps, the SCT consist of 9 disks of stereo strips arranged radially, covering a region up $|\eta| < 2.5$. The intrinsic accuracy per module are 17 μm ($r - \phi$) and 580 μm (z) in the barrel and 17 μm (r $- \phi$) and 580 μm (r) in the disks, making it possible to distinguish tracks which are separated by at least $\sim 200\,\mu$m. The total number of readout channels in the SCT is approximately 6.3 million.

TRT

The Transition Radiation Tracker is the combination of a drift tube chamber tracking detector with a transition radiation detector for electron/hadron discrimination [65, 66]. It provides a large numbers of hits (typically 36 per track) using 4 mm diameter straw tubes filled with a gas mixture (70 % Xe, 27 % CO_2 and 3 % O_2) and covers the range up to $|\eta| = 2.0$. The TRT only provides ($r - \phi$) information, for which it has an intrinsic accuracy of 130 μm per straw. In the barrel region, the straws are parallel to the beam while in the end-cap region they are arranged radially in wheels. The total number of TRT readout channels is 351000. The combination of precision trackers at small radii with the TRT at a larger radius gives very robust pattern recognition and high precision in both $R - \phi$ and z coordinates. The straw hits on the outer radius contribute significantly to the momentum measurement, since the lower precision per point compared to the silicon is compensated by the large number of measurements and longer measured track length. As mentioned previously, the TRT provides an electron/hadron discrimination, taking advantage of the transition radiation generated by the incoming particles. The pion misidentification probability has been measured and found to be 5% for the majority of the detector and as low as 1 – 2% in the best performing detector regions, for a typical electron identification efficiency of 90% [67].

Material

In order to avoid the influence of the ID on energy measurement in the calorimeters, the ID is designed to minimize the amount of material. Figure 2.5 shows the amount of material in electromagnetic radiation lengths[3] (X_0) as a function of the absolute pseudorapidity η. The total amount of material of the ID corresponds to roughly 0.5 X_0 in the range $|\eta| < 0.6$. In the regions $0.6 < |\eta| < 1.37$ and $1.52 < |\eta| < 2.5$, the amount of material reaches 1.5 X_0. In the transition between the barrel and the end-caps at $1.37 < |\eta| < 1.52$, the amount of material is even larger (up

[3]The radiation length is defined as the amount of material traversed by an electron after which it has lost $1/e$ of its original energy by bremsstrahlung.

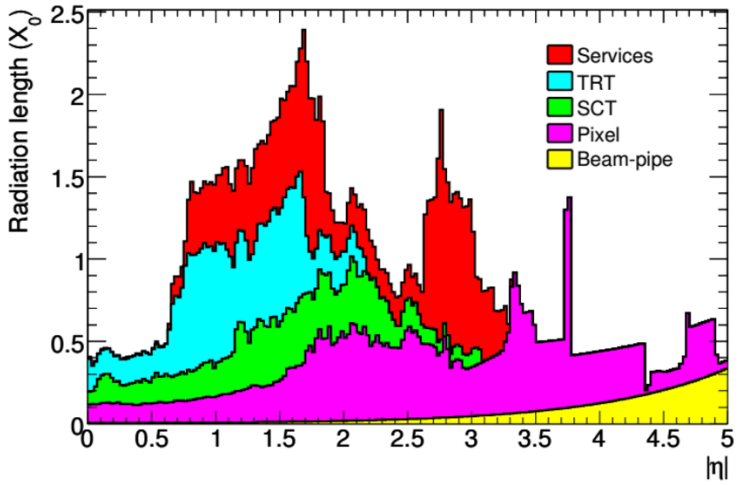


Figure 2.5 – *Amount of material in radiation length (X_0) as a function of η for the different components of the inner detector and services [60].*

to 2.5 X_0), with a significant degradation of the energy measurement in the calorimeters. These regions are usually not included in physics analysis.

2.2.3 Calorimetry

Calorimeters measure the particle energies through their electromagnetic or strong interactions with the detector. The particles that enter the calorimeter produce a cascade of secondary particles as a result of the interaction between the particles and the material. These cascades are commonly known as electromagnetic or hadronic showers (depending on the nature of the interaction). The energy deposited by the shower allows for the measurement of particle energy. A layout of the ATLAS calorimeters is shown in Fig. 2.6. The calorimeter system is divided into four main subsystems, the Liquid Argon (LAr) electromagnetic calorimeter, the Tile calorimeter, the LAr hadronic end-cap calorimeter (HEC) and the LAr forward calorimeter (FCal) [68–70]. They are sampling calorimeters, characterized by alternating layers of an absorber, a dense material used to absorb the energy of the incident particle, and an active medium that provides the detectable signal. These calorimeters cover the range $|\eta| < 4.9$, using different techniques suited to cover a variety of different physics processes over this large η-range. Over the η region matched to the inner detector, the fine granularity of the EM calorimeter is ideally suited for precision measurements of electrons and photons. The total thickness of the EM calorimeter is $\sim$22 radiation length in the barrel and $\sim$24 radiation length in the end-caps. In the following, the different subdetectors that belong to the ATLAS calorimeter system are detailed.

LAr calorimeter

The main purpose of the LAr calorimeter is to measure electron and photon energies. It is divided into a barrel part ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$). It was constructed with accordion shape absorbers (made of lead) and electrodes (made of kapton). The accordion

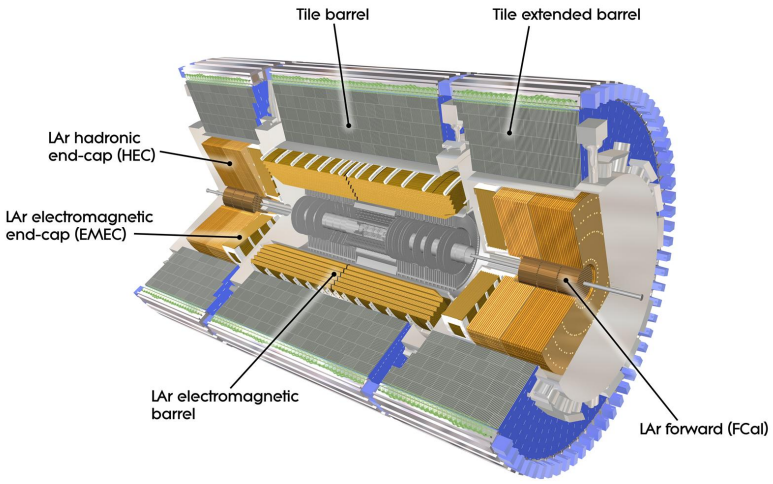


Figure 2.6 – *Schematic diagram of the calorimeter system of the ATLAS detector. All the different components are specified [60].*

shape provides full ϕ symmetry around the detector without any azimuthal cracks, and provides a very uniform performance in terms of linearity and resolution as a function of ϕ. A schematic view with the different ϕ and η segmentations is shown in Fig. 2.7. In the region corresponding to $|\eta| < 2.5$, the calorimeter is segmented into three sections in depth. For the end-cap inner wheel, the calorimeter is segmented into two sections in depth. In the region of $|\eta| < 1.8$, a presampler detector is used to correct for the energy lost by electrons and photons upstream of the calorimeter. The presampler consists of an active LAr layer of thickness 1.1 cm (0.5 cm) in the barrel (end-cap) region. The designed resolution of the EM calorimeter is $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$ (with E in GeV), giving a resolution of about 2.1% for a 25 GeV electron and 1.6% for a 50 GeV electron [60].

Tile calorimeter

The Tile calorimeter is placed around the EM calorimeter. Its main purpose is to reconstruct hadronic jets and to measure their energy. The Tile calorimeter barrel cover the region $|\eta| < 1.0$, and its two extended barrels covers the range $0.8 < |\eta| < 1.7$. It uses steel as the absorber and scintillating tiles as the active material. The barrel and extended barrels are divided azimuthally into 64 modules and segmented in depth into three layers. The total detector thickness at the outer edge of the tile-instrumented region is 9.7 interaction lengths[4] (λ) at $\eta = 0$. The energy resolution achievied is $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$, giving a resolution of about 10% (8%) for a 25 GeV(50 GeV) jet [60].

[4]The interaction length is the mean path length required to reduce the numbers of relativistic charged particles by the factor 1/e, or 0.368, as they pass through matter.

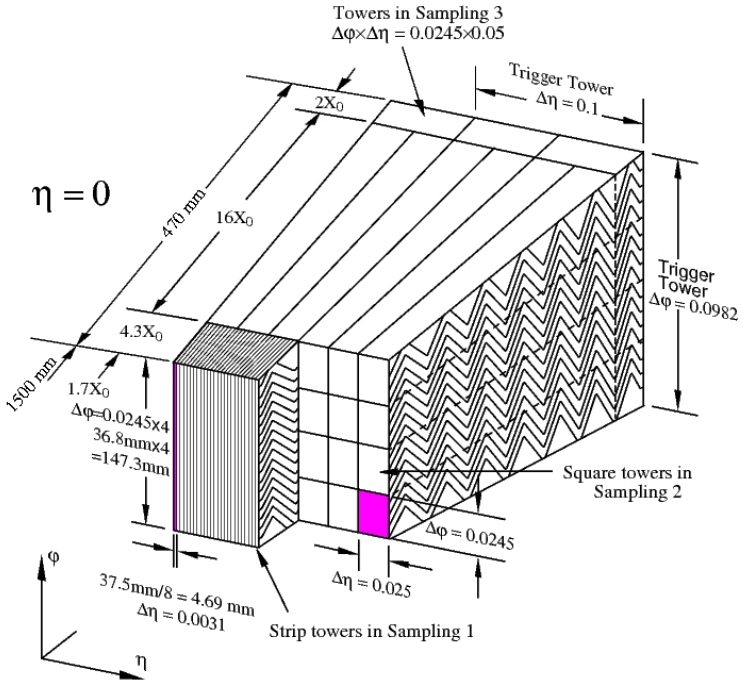


Figure 2.7 – *Schematic diagram of the LAr EM calorimeter All the different components are specified [60].*

HEC

The Hadronic End-cap Calorimeter (HEC) consists of two independent wheels per end-cap, located directly behind the end-cap electromagnetic calorimeter. It extends out to $|\eta| = 3.2$ overlapping with both the Tile and the FCal calorimeters in order to reduce the drop in material density at the transition. Each wheel is built from 32 identical wedge-shaped modules and divided into two segments in depth, for a total of four layers per end-cap. The energy resolution is the same as for the Tile calorimeter.

FCal

The Forward Calorimeter (FCal) provides clear benefits to the calorimeter system in terms of uniformity of the calorimetric coverage as well as reduced radiation background levels in the muon spectrometer. The FCal is approximately 10 interaction lengths deep, and consists of three modules in each end-cap. One module is made of copper and two modules are made of tungsten. The module made of copper is optimized for electromagnetic measurements, while the modules made of tungsten, measure predominantly the energy of hadronic interactions. The energy resolution is $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$, which give a resolution of about 22% and 17% for a 25 GeV and a 50 GeV jet [60].

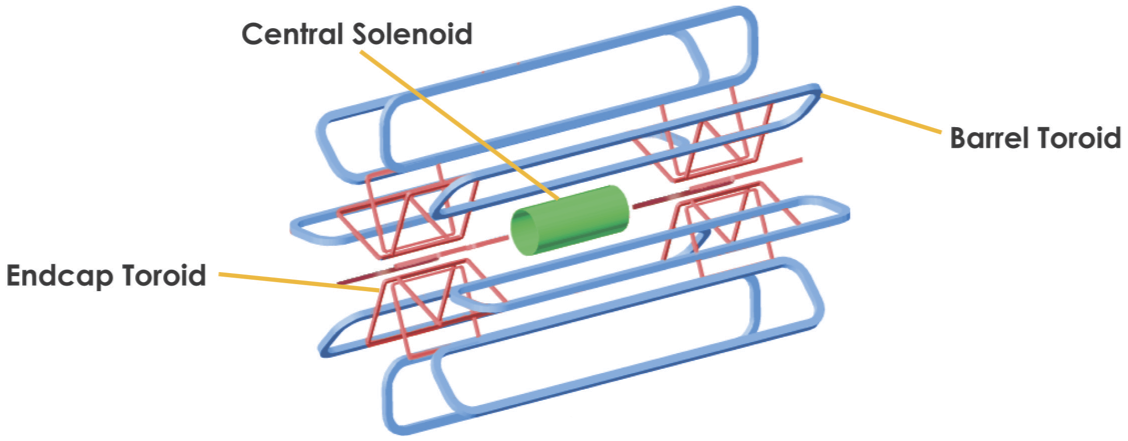


Figure 2.8 – *Schematic diagram of the magnet system of the ATLAS detector [60].*

2.2.4 Magnet system

In order to bend the path of the charged particles interacting with the detector, a magnet system is used by ATLAS. The system consist of one central superconducting solenoid, a barrel toroid, and two end-cap toroids. Figure 2.8 shows a diagram of the system. The central solenoid provides a magnetic field of approximately 2 T. It is built around the ID and it is optimized to reduce the amount of material in front of the calorimeters (which is $\sim 0.5\ \chi_0$). The barrel and end-cap toroids produce a toroidal magnetic field of approximately 0.5 T and 1 T for the muon detectors in the central and end-cap regions, respectively. They consist of eight coils placed in aluminum housings. The toroid magnets are made of aluminum stabilized niobium/titanium/copper conductor cooled with liquid helium ($\sim$4.5 K). The toroidal fields contain non-uniformities which need to be known to high precision to allow for an accurate measurement of muon momenta. The measurement of the magnetic field is performed with Hall sensors placed around the muon spectrometer.

2.2.5 Muon System

The muon spectrometer is the largest subdetector and it is shown in Fig. 2.9. The muon spectrometer together with the ATLAS magnets system provides an independent measurement of the momentum of the muons. It is located in the outer region of the ATLAS detector. The principle of the muon spectrometer is based on the magnetic deflection of muon tracks in the large superconducting air-core toroid magnets, instrumented with separate trigger and high-precision tracking chambers [71]. The muon spectrometer has a resolution of $\sigma p_{\mathrm{T}}/p_{\mathrm{T}} = 10\%$ at $p_{\mathrm{T}} = 1$ TeV [60]. Over the range $|\eta| < 1.4$, the magnetic bending is provided by the large barrel toroid. For $1.6 < |\eta| < 2.7$, muon tracks are bent by the two smaller end-cap magnets inserted into both ends of the barrel toroid. Over $1.4 < |\eta| < 1.6$, the magnetic deflection is provided by a combination of barrel and end-cap fields. With this configuration the magnetic field is mostly orthogonal to the muon trajectories.

The tracks in the barrel are measured in chambers arranged in three cylindrical layers around

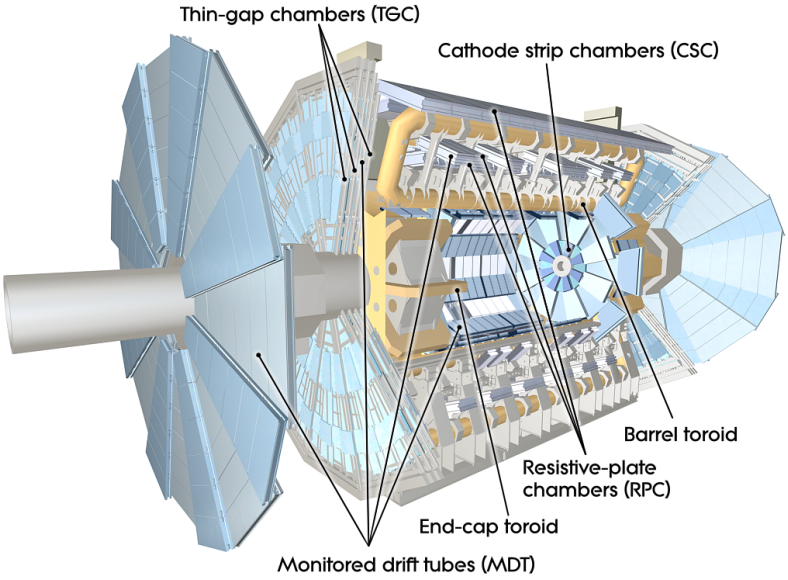


Figure 2.9 – *Schematic diagram of the muon spectrometer of the ATLAS detector [60].*

the beam axis, while in the end-caps the three layers of chambers are installed in planes perpendicular to the beam axis. There are two types of chambers that compose the tracking of the system: the Monitored Drift Tubes (MDT) and the Cathode-Strip chambers (CSC).

MDT

The MDT are responsible for the muon tracking and momentum measurement. They provide coverage in the region $|\eta| < 2.7$ ($|\eta| < 2.0$ for the innermost layer). The basic element of the MDT is a pressurized drift tube with a diameter of 30 mm, operating with Ar/CO_2 gas (with respective fraction of 93/7 %) at 3 bars of pressure. When a muon passes through the MDT chambers it ionizes the gas inside the tubes and causes the electrons to drift towards the anode wires made of tungsten-rhenium. The drift time is used to measure the distance of the particle to the wire, allowing to determine the coordinates of the muon. The chambers are rectangular in the barrel and trapezoidal in the end-cap. Their shapes and dimensions were chosen to optimize solid angle coverage, support structures and access ducts. The direction of the tubes in the barrel and end-caps is along the ϕ coordinate. The MDT has an average spacial resolution of 80 μm per tube or about 35 μm per chamber.

CSC

The CSC are placed in the innermost end-cap wheel ($2.0 < |\eta| < 2.7$) and are designed to cope with the high particle rate in this region, where the MDT are less effective. They are multiwire chambers with segmented cathodes, one with the strips perpendicular to the wires and the other parallel to the wires, providing the measurement of both coordinates. The chambers are filled with a 80% Ar, 20% CO_2 gas mixture. The amount of the gas is reduced in comparison with the MDT, in order to minimize the drift time and the response time. The spatial resolution of a

chamber is 40 μm in the bending plane and about 5 mm in the transverse plane.

Muon Trigger System

The high-precision tracking chambers are complemented with a dedicated trigger system, given the different conditions present in the barrel and end-cap regions during data taking. The trigger system consists of Resistive Plate Chambers (RPC), used in the barrel region ($|\eta| < 1.05$) and Thin Gap Chambers (TGC) used in the end-cap ($1.05 < |\eta| < 2.4$). The RPC consist of gaseous parallel electrode-plates and are placed in three concentric cylinders around the beam pipe. A chamber comprises two detector layers, each layer containing plastic plates separated by an insulating spacer. The TGC uses multiwire chamber technology, similar to the CSC but with a larger distance between anode wires than the distance from the anode wire to the cathode strip, enabling faster signals for triggering. With RPCs and TGCs, a time resolution of 15-25 ns can be achieved, which is sufficient for fast trigger decisions and a good association of tracks to bunch crossings.

2.2.6 Forward Detectors

There are three smaller detector systems in the ATLAS forward region: LUCID (LUminosity measurement using Cerenkov Integrating Detector), ALFA (Absolute Luminosity For ATLAS), and ZDC (Zero-Degree Calorimeter). LUCID and ALFA are situated at ± 17 m and ± 240 m from the interaction point, respectively [72, 73]. Their main purpose is to estimate the luminosity delivered to ATLAS. LUCID detects inelastic scattering in the forward direction, and is the main online luminosity monitor for ATLAS. ALFA consists of scintillating fiber trackers located inside roman pots which are designed to approach as close as 1 mm to the beam. The third system is the Zero-Degree Calorimeter (ZDC), which plays a key role in determining the centrality of heavy-ion collisions and it is located at ± 140 m from the interaction point. The ZDC modules consist of layers of alternating quartz rods and tungsten plates which measure neutral particles at $|\eta| > 8.2$.

2.2.7 Trigger System

The interaction rate (with a luminosity of $10^{34} cm^2 s^{-1}$) is approximately 1 GHz, however, the current technology in ATLAS allows only to store information with a rate of around 400 Hz. This represents an enormous challenge for the ATLAS detector and makes necessary the implementation of a trigger system that efficiently selects events. For this purpose, ATLAS uses a three-levels trigger system which is described in the following [74, 75]:

- **Level 1 (L1) trigger** is a hardware-based trigger with a decision time of 2 μs per event. It searches for high transverse-momentum muons, electrons, photons, jets, and τ- leptons decaying into hadrons, as well as large missing and total transverse energy. It uses reduced granularity information from a subset of detectors: calorimeters and the muon system. This trigger also defines one or more Regions-of-Interest (RoI's). The RoI's are the coordinates in ϕ and η of those regions within the detector where the selection process has identified interesting features.

- **Level 2 (L2) trigger** is a software-based trigger which analyzes the RoI's defined by the L1 trigger. L2 selections use, with full granularity and precision, all the available detector data within the RoI's. This trigger reduces the event rate to 3.5 kHz and has an average processing time of 40 ms per event.
- **Event Filter (EF)** is the final stage of the trigger system which reduces the event rate to roughly 400 Hz. The event selections are implemented using more advanced algorithms than L2, similar to those used in offline analyses, within an average event processing time of the order of 4 seconds. The events that are selected by this trigger are stored and are available for offline analysis.

2.3 Full 8 TeV data sample

The analyses presented in this work are performed using the pp data with a center-of-mass energy of $\sqrt{s} = 8$ TeV collected in 2012. The data correspond to a total integrated luminosity of $20.3 \pm 0.4\,\text{fb}^{-1}$. Not all collision events recorded by ATLAS are used for data analysis. Each subdetector maintains a record of its performance across the run. Only the data collected with subdetectors meeting quality requirements are considered in this work. These quality requirements involve stable beam collisions during which all detector subsystems were fully operational and no data integrity errors were observed in the LAr calorimeter. For each dataset, Good Runs Lists (GRL) are compiled where the quality requirements are satisfied. The data quality efficiency is around 95.5% for the whole Run 1 period.

2.4 Data monitoring in the SCT

Monitoring the performance of the detector is extremely important to evaluate the quality of collected data. ATLAS possesses for each subdetector a monitoring system, which is used to notice any deviation from the expected behavior and take the proper actions to correct them. In particular, the continuous monitoring of the SCT data is essential to ensure good-quality data for physics analysis. Data quality monitoring is performed both online and offline. The online monitoring provides immediate feedback on the conditions of the SCT, allowing quick diagnosis of issues that require intervention during a data-taking period. These may include the recovery of a module that is not returning data, or a more serious problem which requires the early termination of a data-taking period and restart. Offline monitoring allows for the data quality to be checked in more details, and address any problem that could affect its use in physics analyses. A subsample of events is reconstructed promptly by the ATLAS computer farm. Monitoring plots are produced as part of this reconstruction, and used to assess data quality. In the following, one of the steps in the data monitoring chain is described.

2.4.1 Prompt Calibration Loop

The reconstruction of the data proceeds in two stages. First, a fraction of the data from each run is reconstructed immediately, to allow for detailed checks on data quality and detector perfor-

mance. Second, the full dataset is reconstructed. This step is usually performed within 24 to 48 hours after the end of the run. This delay allows to update the detector calibrations. This second step is known as the *prompt calibration loop* [64]. For the SCT, no offline calibrations are performed during this prompt calibration loop, but the run period is used to obtain condition data, i.e. data corresponding to the status of the different detectors. In particular, strips that have become noisy since the last online calibration period are found and excluded from the subsequent bulk reconstruction. Other condition data, such as dead strips or chips, are obtained to monitor the SCT performance.

During the first shutdown of the LHC (between 2012 and 2015), there were several changes in the detector before increasing the collision energy to 13 TeV. Improvements were made to the ID, the muon detectors and calorimeters, as well as to the entire basic infrastructure, including the electrical power supply, the cooling systems, and the monitoring systems. One of the upgrades to the monitoring systems is intended to study the radiation damage in the silicon sensors. This is crucial for Run 2, since the collision energy has increased as well as the radiation that all sensors receive. The radiation affects the silicon lattice by displacing the atoms and creating vacancies that could be filled with impurities. Radiation damage causes leakage current, changes in the depletion voltage and charge collection, which affect the performance of sensors.

One possibility to study the radiation damage is to monitor the Lorentz angle of the silicon sensors. During Run 1, the Lorentz angle was measured every few months. One aspect of this work was to implement Lorentz angle monitoring in real time. For that purpose, it was added to the prompt calibration loop. With this addition, the radiation damage can be monitored run by run. The Lorentz angle is defined in the following, as well as the measurements performed in the past by ATLAS.

Lorentz Angle

Charge carriers in silicon detectors are subject to the electric field, E, generated by the bias voltage and oriented perpendicularly to the module plane, and to the magnetic field from the solenoid, B. The charge carriers, created by the charged particle passing through the sensor, drift along the electric field (Fig. 2.10a). This is not case in the presence of a magnetic field, since charge carriers drift with an angle due to the Lorentz force (Fig. 2.10b). This angle is called the *Lorentz angle*.

In the SCT end-cap modules the effect of the electric and magnetic fields are nearly parallel and the charge carriers drift directly towards the electrodes. In the barrel modules, the effect of these fields are perpendicular and the charge carriers drift along the Lorentz angle, θ_L, perpendicularly to the sensor plane. The value of the Lorentz angle is given by:

$$\tan\theta_L = \mu_H B = \gamma_H \mu_d B \tag{2.6}$$

where μ_H is the Hall mobility, the product of the charge-carrier mobility in silicon μ_d and the Hall factor γ_H. The charge-carrier mobility depends on the bias voltage, the thickness of the depleted region and the temperature.

The Lorentz angle is measured from the dependence of the cluster size on the incident angle of the particle. When the incident angle is equal to the Lorentz angle, all charge carriers generated

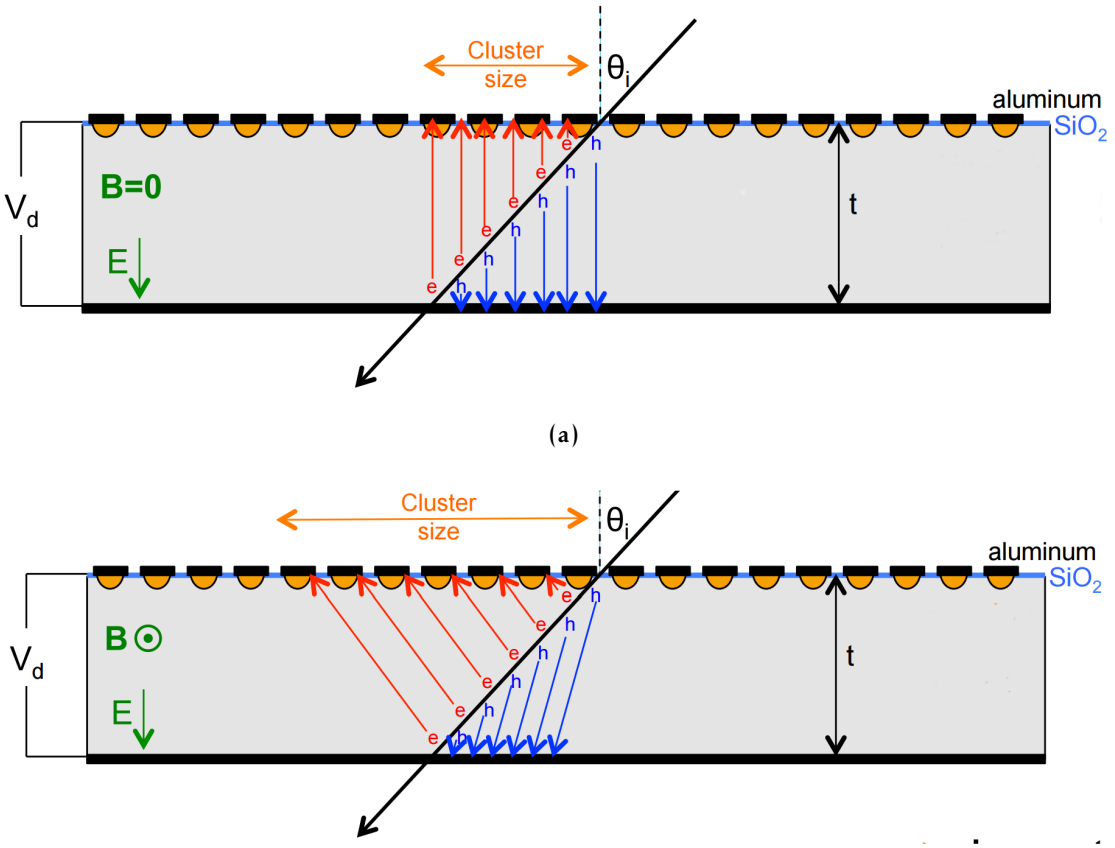


Figure 2.10 – *Drift of charge carriers in the silicon sensor in absence (a) and presence (b) of a magnetic field B.*

by the particle drift along the particle direction and, apart from charge diffusion, are collected at the same point on the sensor surface, giving a minimum cluster size. The tilt of the barrel modules with respect to the radial direction as well as the transverse momentum of particles leads to a range of possible incident angles for positive and negative particles.

The Lorentz angle has been measured using collision data between 2011 and 2012 [64]. For these measurements, the reconstructed tracks are required to have $p_{\mathrm{T}} > 400$ MeV. The measurement is performed separately for two different types of sensor, <100> and <111> sensors. The name of these sensors correspond to the crystal orientation used in the silicon layer. The dependence of the cluster size on the incident angle ϕ_{local} is shown in Fig. 2.11 for data from each barrel corresponding to a luminosity of 0.3 fb^{-1} at 8 TeV. The minimum of the distribution corresponds to the Lorentz angle. The figure on the left and on the right show the results for the two types of sensors (<100> and <111>). Data are fitted using a convolution of the function:

$$f(\phi_{\mathrm{local}}) = a|\tan\phi_{local} - \tan\theta_L| + b \tag{2.7}$$

with a Gaussian distribution. The fitted parameters are the Lorentz angle, θ_L, the shape parameters, a and b, and the width of the Gaussian. The electric field in the silicon, and thus the local Lorentz angle, varies with distance from the electrodes. Figure 2.12 shows a summary of the

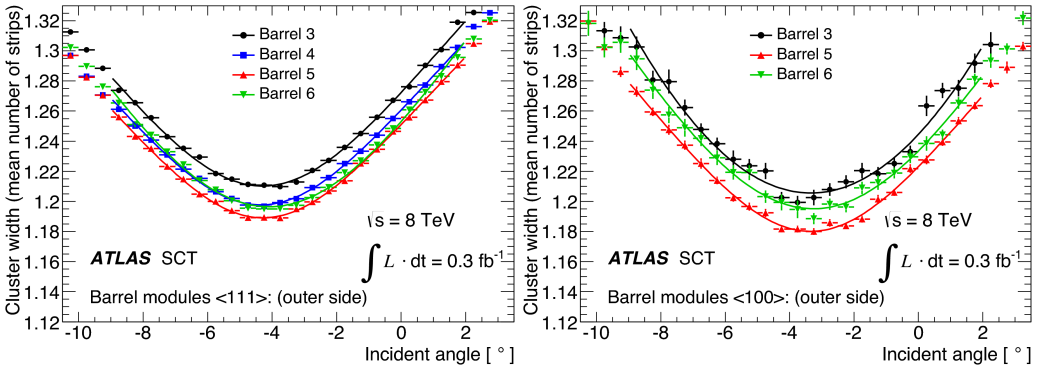


Figure 2.11 – *Cluster-size dependence on the particle incident angle for each SCT barrel for (left) <111> and (right) <100> sensors. The displacement of the minimum from zero is a measurement of the Lorentz angle θ_L [64].*

measurements performed between 2011 and 2012. Over this period, the largest variation in the Lorentz angle is below 0.1.

The track requirements as well as the fit technique used for the measurements previously described were implemented in the prompt calibration loop, in order to have the Lorentz angle estimation after each run. The Lorentz angle monitoring in the prompt calibration loop allows for a more detailed study in shorter periods of time, which is helpful if fast actions related with the performance of the SCT need to be taken.

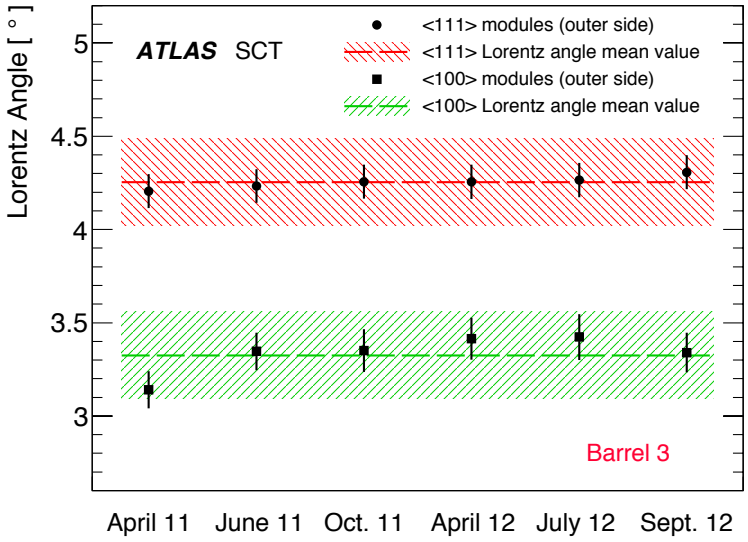


Figure 2.12 – *Average fitted values of the Lorentz angle for <111> and <100> sensors in the innermost barrel for stable-beam running periods in 2011 and 2012. The error bars correspond to the statistical and systematic uncertainties [64].*

3

Monte Carlo simulations

An accurate simulation of the physics processes and the interaction of particles with the detector is necessary to model the impact of the analysis procedure on the measured quantities, and to estimate the background composition expected in data. Computer programs known as Monte Carlo (MC) event generators simulate the physics processes. In this chapter, the production chain of Monte Carlo simulation in pp collisions is briefly described. This is followed by the description of the Monte Carlo generator programs used in this work and a description of the ATLAS detector simulation.

3.1 Simulations of hadron collisions

Monte Carlo (MC) generators can be used to perform simulations to study the response of the detector for a large variety of physical processes as well as the physical process itself. The simulations are performed in several steps. The first step involves the full simulation of the hadron collision, which starts with the hard interaction of the incoming partons, followed by the simulation of radiated partons, hadronization and pile-up. The next step in the process is the simulation of the detector, including the physical processes due to the interaction between the particles and the detector. The final step consists of the digitalization process where the electronic signals, similar to the ones obtained with real data, are simulated. The output of the simulations has a similar format as the collected data, which allows for the usage of the same trigger selection and reconstruction algorithm for real data and simulated events.

3.1.1 Hard interaction

In pp collisions, the full physical event involves the modeling of different sub-processes, illustrated in Figure 3.1. The two colliding protons provoke a deep inelastic interaction with a large momentum transfer. The interacting partons generate the process of interest by hard interaction (between them). The computation of the hard interaction process is performed by taking advantage of the factorization theorem of QCD (explained in Sec. 1.2.1). The term $\hat{\sigma}_{ab\to X}(x_a p_a, x_b p_b, \mu_F^2 \mu_R^2)$ in the theorem is the cross section for the partonic process $ab \to X$ computed at fixed order in per-

turbation theory (LO[1] or NLO[2]). This step is commonly referred to as the Matrix Element (ME) calculation, since it involves the calculation of the scattering matrix relating the initial and the final states particles of the process.

As stated in Chapter 1, the PDFs do not depend on a particular process. They are estimated using phenomenological models (combining theory with experimental results). In this work, CTEQ6L1 [76], CT10 [77], MSTW2008 [18, 19], and NNPDF2.3 [78] PDF sets are used with different MC generators and processes, described in Sec. 3.2. The CTEQ6L1 set uses data from deep-inelastic ep scattering from the H1 and ZEUS experiments at HERA, and data of the inclusive jet production as a function of rapidity and transverse energy from D0 at the Tevatron. It is based on the minimization of a global χ^2 function (the sum of χ^2's over all datasets) with respect to the model parameters. The CT10 set is based on a similar global χ^2 function minimization, but also includes results of the rapidity distribution of Z^0 production, measured by CDF and D0. The MSTW2008 set is determined from global analysis of hard-scattering data and uses data from CCFR/NuTeV (at Fermilab) dimuon cross section and Tevatron Run 2 data on inclusive jet production, the lepton charge asymmetry from W decays, and the Z rapidity distribution. The NNPDF2.3 set makes use of neural networks combining LHC data from ATLAS and LHCb W and Z rapidity measurements from the 2010 run, CMS W electron asymmetry data from the 2011 run, and ATLAS inclusive jet cross sections from the 2010 run.

3.1.2 Parton shower

The hard subprocess, by definition, involves large momentum transfers and therefore the partons involved in it are strongly accelerated. Just as accelerated electric charges emit QED radiation (photons), the accelerated colored partons will emit QCD radiation in the form of gluons. Unlike the uncharged photons, the gluons themselves carry color charges and can therefore emit further radiation, leading to a cascade of secondary partons known as a parton shower (PS). In principle, the showers represent higher-order corrections to the hard subprocess. However, it is not feasible to calculate these corrections exactly, since radiative corrections at a fixed perturbative order are divergent at low energies (infrared divergence) or small angles (collinear divergence). Since a calculation is not possible an approximation scheme (called leading-log) is used where only the dominant contributions are considered. There are three possible processes for QCD emission (splitting): $q \to qg$, $g \to gg$, and $g \to \bar{q}q$. The differential cross section $2 \to n$ with one extra emission $2 \to n+1$ can be expressed as the product of the $2 \to n$ partonic cross section and a factor accounting for the splitting probability of one of the partons as follows:

$$d\sigma_{2\to n+1} \sim d\sigma_{2\to n} \frac{\alpha_S}{2\pi} \frac{d\theta^2}{\theta^2} dz d\phi P_{ij}(z,\phi), \tag{3.1}$$

where θ and ϕ are the opening and azimuthal angles of the splitting, and P_{ij} is the splitting function which describes the distribution of the fraction z of the energy of i carried by j. The parton shower is developed by applying sequentially Eq. 3.1 to the simulation and generating the values of z, θ and ϕ for each splitting of the partons involved in the hard process. The

[1] Leading Order
[2] Next-to-Leading Order

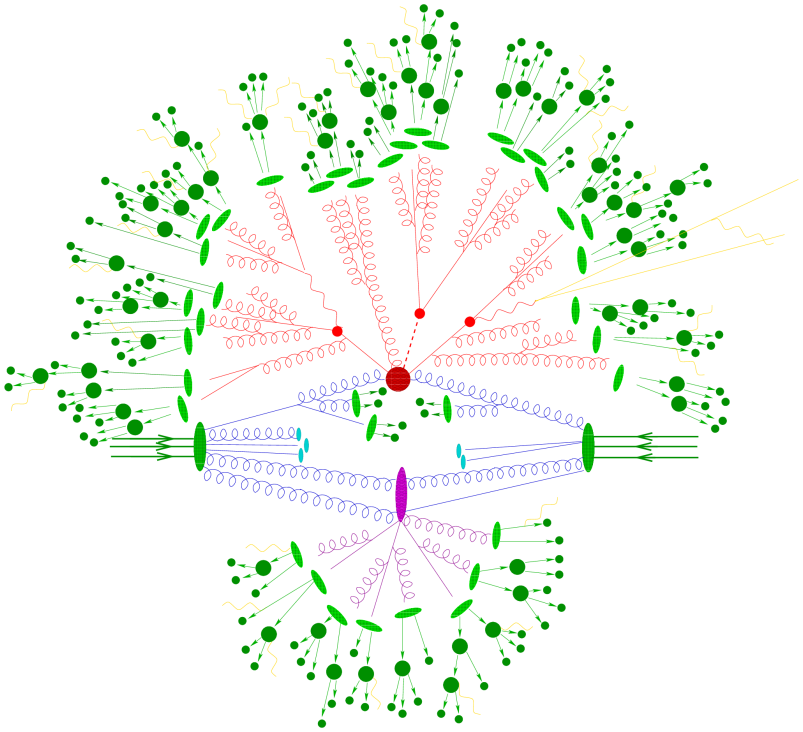


Figure 3.1 – *Representation of a hadron-hadron collision as simulated by a Monte-Carlo event generator. The red central circle represents the hard collision, surrounded by a tree-like structure representing radiations as simulated by parton showers. The purple circle indicates a secondary hard scattering event. Parton-to-hadron transitions are represented by light green ellipses, dark green circles indicate hadron decays, while yellow lines are signal of soft photon radiation [79].*

evolution of the shower is performed by using a parameter called *virtuality*, that can be taken as the virtual mass-squared of the produced partons[3]. The initial virtuality is required to be smaller than the momentum transfer of the hard process, and the shower is terminated when the virtuality is below the hadronization scale ($Q^2 \sim 1$ GeV^2). In order to take into account the real parton emission at each order in perturbation theory during the showering, and also the virtual effects (quantum loops) of the same order, the probability of not splitting during evolution from scale q_1^2 to q_2^2 is considered. The probabiliy is known as the Sudakov form factor, defined as:

$$\Delta_i(q_1^2, q_2^2) = \exp\left[-\int_{q_1^2}^{q_2^2} \frac{dq^2}{q^2} \frac{\alpha_S}{2\pi} \int_{\frac{Q_0^2}{q^2}}^{1-\frac{Q_0^2}{q^2}} dz \int_0^{2\pi} d\phi P_i(z,\phi) \right]. \tag{3.2}$$

The generation of a parton shower thus proceeds as follows. Given the initial scale Q^2, the equation $\Delta_i(Q^2, q_1^2) = R_1$ is solved, where R_1 is a random number uniform on the interval $[0,1]$, for the scale q_1^2 of the first splitting. If $q_1^2 < Q^2$, then the splitting is unresolvable and the showering of that parton is terminated. Otherwise, the process $i \to j + k$ is resolvable and $\Delta_i(q_1^2, q_2^2) = R_2$ is computed again for parton j and then for parton k, until all attempted splittings have fallen below the hadronization scale.

Initial- and Final-State Radiation

A final-state radiation (FSR) is developed from an outgoing parton of the hard process. The procedure to simulate it is the same as the parton showering, which starts from a high energy Q^2 and reaches progressively the hadronization scale. The initial-state radiation (ISR) is developed from the incoming partons of the hard process. There is an important difference in the shower evolution of ISR compared to FSR, as the final energy of the showering is set by the hard interaction energy scale. In this case, the subsequent partons from each of the incoming ones start at a high energy and low virtuality and evolve to a higher virtuality by radiating partons and losing energy. The showering of these partons terminates when they reach the energy of the initial hard process, which sets the scale that limits the endpoint virtualities of the showers. MC generators implement a mechanism, called backward evolution, that first sets the correct parton momentum fractions for the hard scatter, and then develops the showers backward, with the intermediate partons gaining energy at each emission. The Sudakov form factor in this case is slightly modified by taking into account the ratio of the PDFs.

ME and Parton shower matching

The ME calculations at the simplest fixed order LO are often not sufficient for a full description of the final state. In this case, additional radiations or NLO calculations can be used by applying an infrared cut-off in order to avoid divergences from soft and collinear emissions. Procedures, known as ME-Parton shower matching, avoid the overlap between the hard and large angle emission, described by the ME, and the soft and collinear emissions, described by the parton shower. The most widely-used matching schemes are the Catani-Krauss-Kuhn-Webber (CKKW [80]) and

[3] Several parameters can be taken as virtuality. This depends on the MC generator.

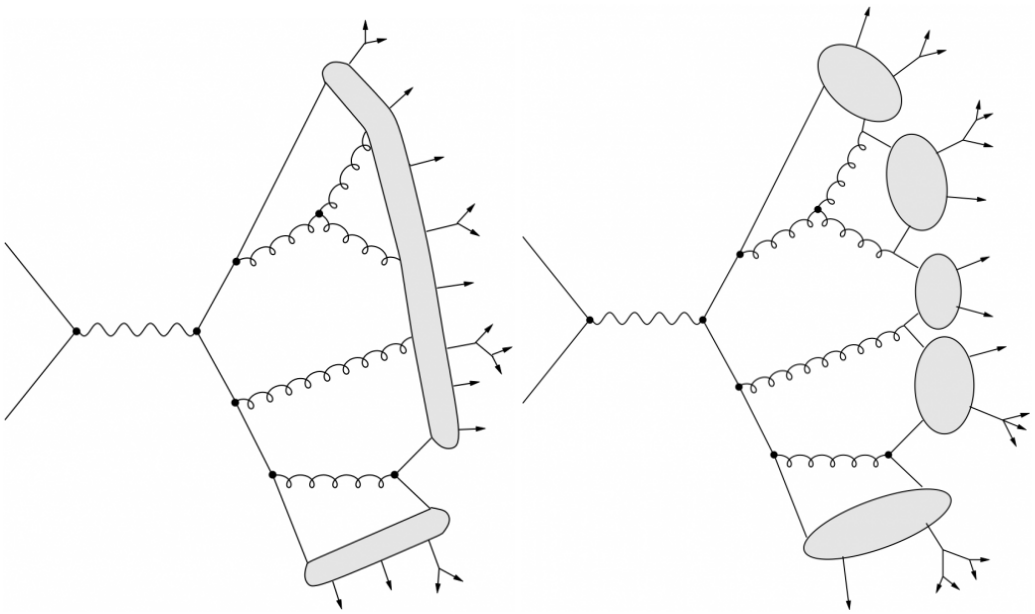


Figure 3.2 – *Schematic diagram for the Lund Model (left) and Cluster Model (right) [82].*

the Michelangelo L. Mangano (MLM [81]) algorithms. The CKKW scheme relies on the ME reweighting using the values of α_S in every vertex of the branching, and the Sudakov factor from every line between the vertices. The MLM scheme separates the events in exclusive samples of n partons in the final state, on which the parton shower is added. The parton configuration after the showering is then processed with a cone jet algorithm, with a radius R_{jet} and ΔR matching between the jets and the considered partons are used for validation.

3.1.3 Hadronization

When the shower evolution brings the parton virtuality q^2 below the hadronization scale $Q_0^2 \sim 1$ GeV^2, the parton enters a non-perturbative phase, which leads to the formation of the final-state colorless hadrons. This hadronization process cannot be calculated in perturbative QCD, and therefore event generators have to rely on phenomenological models based on general features of QCD. The most commonly used phenomenological models are the Lund [83] and Cluster [84] models. The Lund model (or String model) treats all but the highest-energy gluons (which will further radiate) as field lines, which are attracted to each other due to the gluon self-interaction and form a narrow tube (or string) of strong color field. The Cluster model is based on the so-called *preconfinement* property of QCD. At energies much lower than the hard-scattering process, the partons in a shower are clustered in colorless groups with an invariant mass distribution that is independent of the nature and scale of the hard subprocess, depending only on q and the fundamental QCD scale. The clusters are then decaying isotropically into two hadron final states. In Fig. 3.2 a schematic diagram of two models is shown.

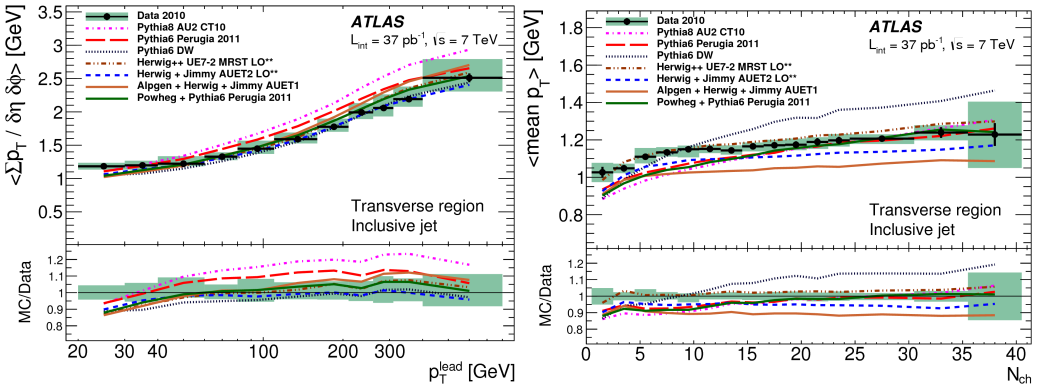


Figure 3.3 – *Charged-particle p_T (left) and multiplicity (right) for data different MC simulations. These variables are used for the study of underlying interactions in different MC simulations.*

3.1.4 The underlying event

Collisions between the partons in the incoming hadrons that do not directly participate in the hard subprocess can interact between each other with lower energy. This kind of interactions are referred to as the underlying event (UE) and are described with phenomenological models, where sets of tuneable parameters of the simulations, such as the charged-particle p_T and multiplicity (Figure 3.3), are fit to experimental data [85]. The dominant subprocess of the UE is gluon-gluon scattering, with a cross section larger than the total pp scattering cross section. The generic soft scattering of partons is referred as multiple parton interactions (MPI) and is modeled in MC generators and tuned to observed data. The color connection with the beam remnants is also simulated with phenomenological models.

3.1.5 Pile-up

The pile-up interactions are classified into two different categories and are simulated as two components: in-time pile-up events and out-of-time pile-up. The in-time pile-up events correspond to the scattering of protons in the same bunch of the hadron generating the hard process of interest. They are simulated in a similar way as the UE. The out-of-time pile-up events correspond to the additional pp collisions occurring in bunch-crossings just before and after the collision of interest. They are modeled as the in-time pile-up but the time response of the readout electronics is also taken into account.

3.2 Monte Carlo Generators

The MC simulation is performed using dedicated software packages referred to as MC generators. The MC generators can be classified as either multipurpose generators, capable of performing the full simulation chain described above, or as specialized generators, optimized for an accurate simulation of specific aspects. The following sections summarize the characteristics of the MC generators used in this work.

3.2.1 General purpose generators

- Pythia is a multipurpose MC generator using LO calculations for $2 \to n$ ($n < 4$) processes and parton shower development [86]. The Lund string model is used for hadronization, and UE simulation is included. The Pythia version 6 is used in most of the cases. Despite the fact that Pythia can simulate UE, Jimmy [87], which is an standalone software at LO, is the most common choice for performing this task.

- Herwig is a multipurpose MC generator using LO calculations for $2 \to 2$ processes and the parton shower [88]. The cluster model is used for hadronization and for the UE description. The generator Jimmy [87] typically simulates UE as well as the multi-parton interactions.

3.2.2 Specialized generators

- Alpgen is a MC generator using LO calculations of $2 \to n (n < 10)$ processes [89]. It is interfaced with either Pythia, Herwig or Jimmy for parton shower development, hadronization and UE. ME-PS matching is applied with the MLM method.

- MadGraph is a MC generator specialized in the computation at LO of $2 \to n (n < 7)$ processes [90]. It is interfaced with either Pythia or Herwig for parton shower evolution and MLM matching is applied to avoid overlapping.

3.2.3 NLO generators

- MC@NLO is a MC generator using NLO calculations [91]. The full NLO ME provides precise cross section estimates, but higher-multiplicity parton emissions are simulated via Herwig PS. The ME and PS matching is performed by a built-in CKKW-like subtraction procedure. Hadronization and UE are simulated through Herwig or Jimmy.

- Powheg-**hvq** is an event generator computing the ME at NLO at pQCD [92]. It uses a different ME-PS matching scheme with respect to MC@NLO, which leads to a better modeling of the jet multiplicity in the event.

3.2.4 Tuning

The baseline MC simulation is sometimes not sufficient to provide an accurate description of some physical processes such as the UE and MPI. All generators include a set of parameters that can be modified in order to provide a better description of data. These parameters can regulate the multiple-parton scattering, the ISR/FSR contribution, UE, beam-remnant, color reconnection and others. In this work, some specific sets of parameters are used to tune different MC generators and processes. These tunes are the *Perugia tunes 2011/2012* [93] for Pythia, and the ATLAS Underlying Event Tune (*AUET2 tunes*) for HERWIG [94]. The Perugia tunes are derived from data collected at LEP, Tevatron, and LHC, while the AUET2 tunes use only data collected at the LHC.

Another important parameter, specific to the POWHEG-hvq generator, is the h_{damp} parameter. The h_{damp} parameter is the resummation scale that is used in the damping function, which is designed to limit the resummation of higher-order effects at large transverse momentum without spoiling the NLO accuracy of the cross section. This parameter is known to improve the agreement of the number of jets in the final state between data and MC when is set to equal to the mass of the top quark in $t\bar{t}$ simulated events.

3.2.5 Parton and Particle level

When the full chain of the MC simulation process is performed, the particles at different stages of the generation are saved in what is called the MC record. These stages are known as *levels*. The first level is defined right after the hard interaction and it is known as *parton level*. At this level the particles do not suffer from any radiation or showering. The behavior of the particles at parton level depends strongly on the MC generator used to simulate the hard interaction. The observable built at this level, however, are more compatible with their theoretical predictions, since this level corresponds to the output of the ME computation. A different level can be considered and corresponds to the final output of the MC generator: a set of four-vectors of all stable particles produced in the event. This level is known as *particle level*. The stable particles correspond to the resulting particles after all decays, radiations and hadronization. The particle level is commonly used to perform the study of physics events due to the fact that they are less dependent on the MC generator. However, the observables built at this level need a special treatment when a prediction is performed, which is not necessarily equivalent to the predictions at parton level.

3.3 Simulation of the ATLAS detector

In order to compare the output of the MC generator with the data recorded by the ATLAS detector, the MC has to pass through a detector simulation that simulates the reconstruction of the particles after they have interacted with the detector. This level is referred to as *reconstructed level*. The ATLAS detector software, created with GEANT4 [95],[4] reproduces the interaction of the particles with the detector. The GEANT4 parameters are tuned using test-beam and pp collision data. The accuracy of the detector simulation is based on the information from two databases: one contains the description of the detector volumes in terms of dimensions, geometry, position and material composition, while the second database provides the information on the detector real-time conditions such as dead channels, and misalignments. Since conditions of the detector vary in time (run by run), it is important that the detector simulation reproduces as closely as possible the real status of ATLAS during a particular data-taking period. For this purpose, the simulations are reprocessed for each data release. A less refined simulation, known as ATLFAST-II [96], is also available. This reduces the CPU time necessary to process an event, since it relies on a parametrized description of the particle showers in the calorimeters. The final output of this simulation is equivalent to the output of the real ATLAS detector when data is collected.

[4] A toolkit for the simulation of the passage of particles through matter.

4

Object Definition and Selection

In this chapter, the definition of the objects used in this work are presented, as well as the criteria used to select dileptonic $t\bar{t}$ events. First, the definitions of objects such as tracks, electrons, muons, jets, and missing transverse momentum are presented. Second, the event selection requirements specific to dileptonic $t\bar{t}$ events are presented. The studies related to the background modeling affecting the measurements are also discussed. Finally, the definition of the particle level and the fiducial region are presented.

4.1 Object definition

4.1.1 Tracks

When an electrically charged particle passes through the solenoidal magnetic field in the ID, it moves along a helicoidal trajectory with a curvature inversely proportional to its momentum. The tracks are the reconstruction of these trajectories from the electric signals (hits) induced in the detectors arising from the ionization produced by the interacting particle. The tracks are used to identify charged particles and to measure their momenta. In addition, the extrapolation of the trajectories allows the identification of the interaction point where the hard scattering takes place.

The parameters describing a track are shown in Fig. 4.1. The angle θ is measured with respect to the z axis in the rz plane measured from the point of the track closest to the origin (known as perigee); ϕ_0 is the angle with respect to the x axis in the xy plane measured from the perigee; d_0 is the impact parameter with respect to the z axis in the xy plane; and z_0 is the z component of the track at the perigee.

There are several steps in the track reconstruction. First, all hits found in the ID are collected and these hits are converted into three-dimensional space points. Then, an algorithm is applied using as initial points at least three aligned points (seed hits) in the pixel detector or the SCT [97], called *inside-out* algorithm. The inside-out algorithm is a reconstruction strategy that iteratively builds a track by combining space points one by one, starting from the hits collected by the silicon detectors and moving towards the points collected by the TRT. For each new point considered by the inside-out algorithm, the compatibility between the track and the new point is checked

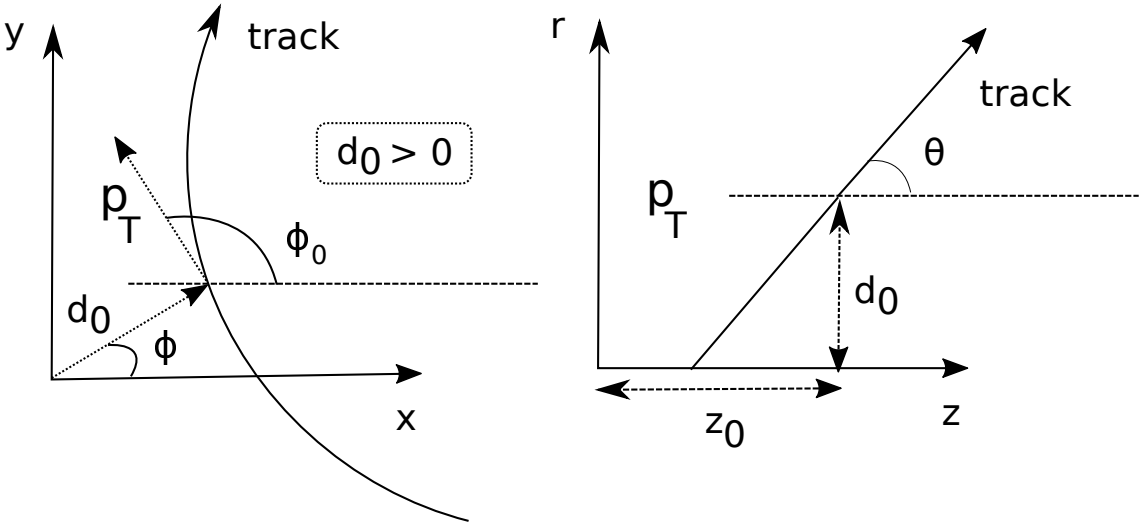


Figure 4.1 – *Track parameters in the xy (left) and Rz (right) planes where the origin is the beam spot, i.e. the region where the protons collide.*

using an algorithm called *Kalman filter algorithm* [98]. Finally, the track quality is improved by including the signals obtained from the TRT and also the particle interaction with the detector material.

When the signals in the TRT are not associated with any track from the silicon detectors, a different algorithm is applied in order to reconstruct tracks from secondary charged particles [97]. The algorithm starts with the seed hits collected at the TRT and performs an extrapolation back to the SCT and pixel detector. This reconstruction strategy is called *outside-inside* algorithm.

4.1.2 Primary Vertices

The primary vertices (PV) correspond to the interaction points where the hard scattering takes place. It is thus quite important to identify them in order to reconstruct the physics objects properly. The reconstruction of primary vertices is divided in two steps which are performed iteratively.

The first step of the PV reconstruction is to find the maximum of impact parameter distribution of the reconstructed tracks. This maximum corresponds to the initial PV. Then, a χ^2 based fitting algorithm is used, which deals with outlying track measurements by down-weighting their contribution to the overall vertex. The tracks that are incompatible with the vertex by more than approximately 7σ are used to seed a new vertex and the procedure is repeated until no unassociated tracks are left in the event or no additional vertex can be found. Once the PV candidates are defined, the one associated with the hard scattering process is assumed to be the one with the highest sum of squared transverse momenta of the tracks. The others PVs are considered to be pile-up interactions. Vertices incompatible with the beam collision region are considered to be secondary vertices.

4.1.3 Electrons

Electrons are reconstructed in the central detector region ($|\eta| < 2.47$) from energy deposits (called clusters) in the calorimeter that are matched to tracks in the ID [99]. The clustering is performed by using a *sliding window* algorithm, described as follows. The EM calorimeter is divided into a grid of elements composed of $N_\eta \times N_\phi$ cells in $\eta - \phi$ space of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. Towers are formed by summing the energy of the cells in all longitudinal layers within each element. A window of a fixed size of 3×5 in units of 0.025×0.025 in $\eta - \phi$ space is moved across each element of the tower grid defined above. If the transverse energy (defined as the sum of the transverse energy of the towers contained in the window) is a local maximum and is above a threshold of $E_T > 2.5$ GeV, a precluster is formed. The precluster is defined as the energy deposit of the electron candidate. The preclusters matched to an ID track with $p_T > 500$ MeV are classified as electrons or converted photons, otherwise they are classified as unconverted photons. Electrons are distinguished from converted photons by investigating the presence of pairs of close-by tracks originating from a vertex displaced from the interaction point and by verifying the location of the first hits along the path of the single tracks. When the track and precluster are matched, a cluster is formed and is optimized to take into account the overall energy distributions in the different regions of the calorimeter. In the barrel region, the energy of the electron cluster is collected by enlarging the size of the cells to 3×7, while in the endcaps the size is increased to 5×5. The total reconstructed electron energy is obtained from the sum of four contributions: the energy of the cluster, the energy deposit in the material in front of the EM calorimeter, the estimated energy deposit outside the cluster (lateral leakage), and the estimated energy deposit beyond the EM calorimeter (longitudinal leakage).

Not all the objects obtained by the reconstruction algorithm are signal electrons (i.e electrons produced during the hard scattering process). Misidentified electrons or electrons coming from secondary decays, Dalitz decays and semileptonic heavy flavor hadron decays can also be obtained. A criteria based on multivariate techniques (called the likelihood technique) is used in order to reject as much of these backgrounds as possible while keeping the efficiency for signal electrons high. This technique uses the information of the shape of the clusters, known as shower shapes variables, to discriminate the signal and background electrons. There are different selections that can be used, based on which shower shapes variables are used and if an isolation criteria is applied. They are called *looseLH*, *mediumLH*, and *VeryTightLH*. In this context, the likelihood technique is referred to as the electron likelihood identification and it makes use of signal and background probability density functions of the discriminating variables. Based on these functions, an overall probability is calculated for the object to be either signal or background. The likelihood-based identification provides a higher rejection of fake electrons for the same identification efficiency (described below) compared to a cut-based approach, as shown in Fig 4.2 [100]. The identification efficiency is above 80% for electron with $E_T > 25$. This energy threshold is commonly used in top-quark physics analyses to reach a background efficiency of less than 1%. In the analyses which are presented here, the *mediumLH* electron likelihood identification is used.

The reconstruction efficiency for electrons is defined as the ratio of the number of signal electrons reconstructed as a cluster matched to a track passing the track quality criteria (numerator)

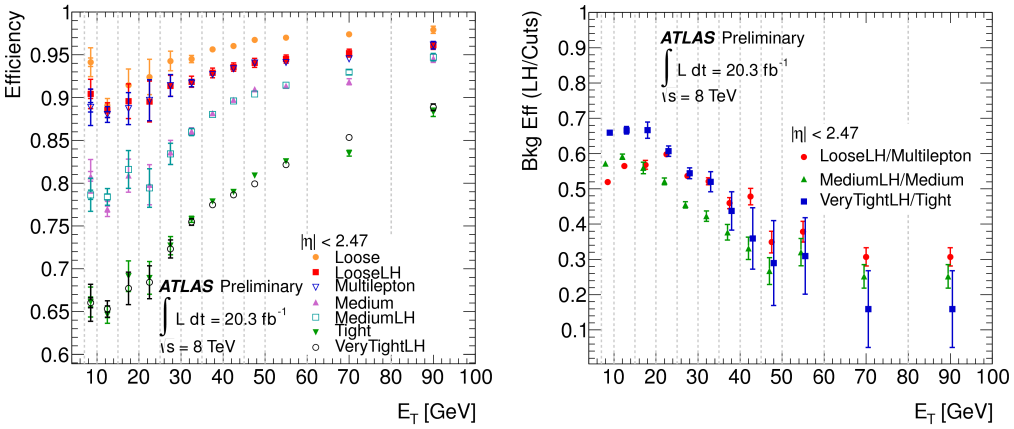


Figure 4.2 – *Electron identification efficiency (left) and background efficiency ratio (right) between cut-based and likelihood-based selection criteria as a function of E_T [100].*

to the number of clusters with or without a matching track (denominator). It is measured using a *tag and probe* method. The tag and probe method is based on the knowledge of well known resonances such as $Z \to \ell^+\ell^-$ or J/Ψ decays. The electrons coming from these are known to be signal electrons and therefore can be used to check if the electrons pass or fail certain identification criteria. In this method, a "tag" electron is selected with a tight identification criteria and a "probe" electron is selected with a looser identification criteria. The invariant mass of both electron are expected to be within the Z or J/Ψ mass peak. The fraction of probe electrons, which pass the selection under study gives an estimate of the corresponding efficiency. The measured efficiency in data varies from 95% to 99% between the endcap and barrel regions for low E_T electrons ($E_T < 20$ GeV). For very high E_T electrons ($E_T > 80$ GeV) the efficiency is of $\sim$99% over the whole η range [100].

The identification efficiency is defined as the ratio of the number of electrons passing a certain identification selection (numerator) to the number of electrons with a matching track passing the track quality requirements (denominator). It is measured in data using $Z \to e^+e^-$, $Z \to e^+e^-\gamma$ and $J/\Psi \to e^+e^-$ events with the tag and probe technique. The measurement of the combined reconstruction and identification efficiency as a function of E_T and η is shown in Fig. 4.3. For the *mediumLH* electron likelihood identification, the efficiency is above 80% (on average) for electrons with $E_T > 25$ GeV.

The electron energy scale[1] and resolution[2] are measured by reconstructing the invariant mass of the $Z \to e^+e^-$, $Z \to e^+e^-\gamma$ and $J/\Psi \to e^+e^-$ events. The electron energy scale is known with 0.5 – 1% accuracy and the energy resolution is known by approximately 2% as shown in Fig. 4.4 [101]. There is a notable increment of the energy scale factor in the region $1 < |\eta| < 1.82$. These differences are related to residual uncertainties in the detector material description. The resolution curve shows a non-flat behavior at $E_T < 40$. This is expected since the energy resolution is predicted to be inversely proportional to $\sqrt{E_T}$.

[1] The energy scale is the absolute energy that relates the energy measured by the detector and their true energies.
[2] The resolution corresponds to the relative uncertainty on the energy measurement as a function of the energy.

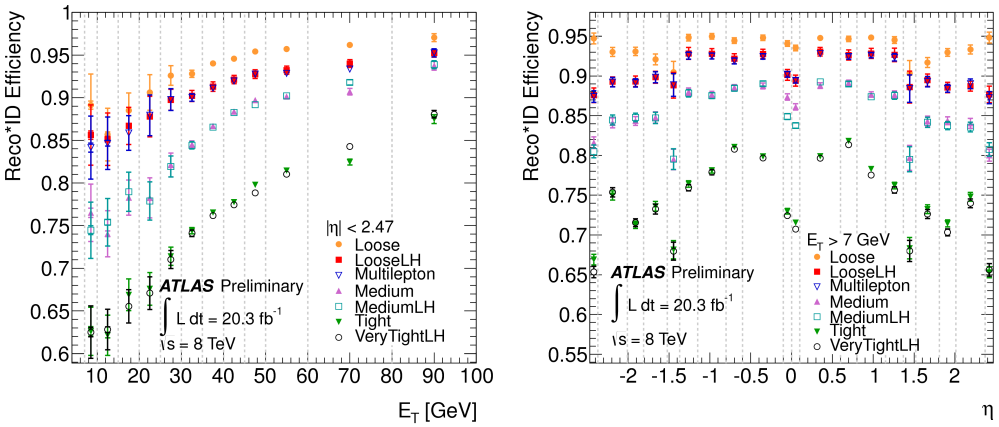


Figure 4.3 – *Measured combined reconstruction and identification efficiency for the various cut-based and likelihood selections as a function of E_T (left) and η (right) for electrons. The data efficiency is derived from the measured data-to-MC efficiency ratios and the MC prediction from $Z \to ee$ decays [100].*

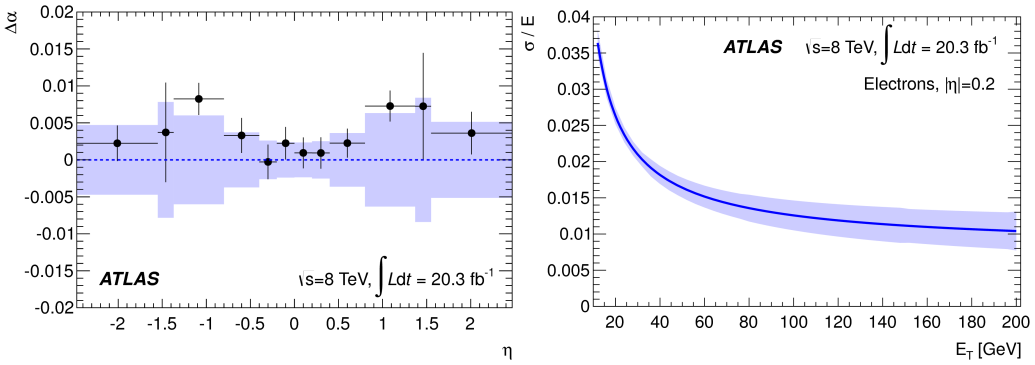


Figure 4.4 – *Energy scale factors obtained after Z- based calibration from the J/Ψ sample, as a function of the electron pseudorapidity (left). Resolution curve and its uncertainty as a function of E_T for electrons with $|\eta| = 0.2$. The selection criteria called multilepton, is optimized for the low energy electrons in the $H \to ZZ^* \to \ell\ell$ analysis [101].*

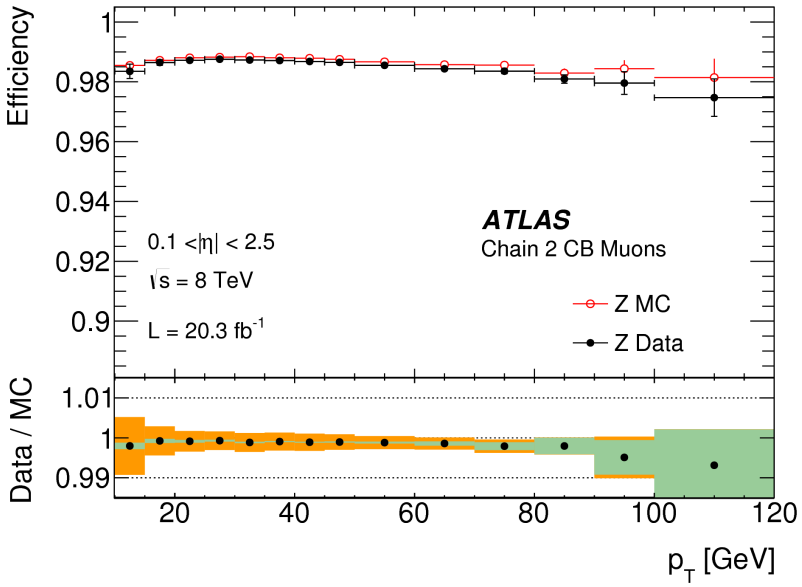


Figure 4.5 – *Reconstruction efficiency as a function of pt for Chain-2 CB muons, derived using $Z \to \mu^+\mu^-$ events [102].*

4.1.4 Muons

Muons can be reconstructed using different parts of the detector that define the muon *type* and using different strategies, known as "Chains" [102]. In this work, muons are reconstructed from tracks formed by combining information from the MS with the ID. They are known as Combined (CB) muons. CB muons are reconstructed independently in the ID and MS, and a combined track is formed if a successful combination of a MS track with an ID track is found. The strategy to reconstruct the muons is known as the *Muid algorithm* (or Chain 2). The algorithm searches for track segments in the RPC and TGC in $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$ regions where the trigger fired. The MS track is built with a least-square fitting method, and the trajectory is extrapolated back to the interaction point. Then, it is combined with the ID track that provides the best match, based on a χ^2 test.

The reconstruction and identification efficiency for muons reconstructed from a combination of inner detector and muon spectrometer tracks is above 99% for muons with $p_T < 80$ GeV and decreases to 98% above that threshold in the region $0.1 < |\eta| < 2.5$, as shown in Fig. 4.5. The decreasing behavior is expected since high energetic muons a more difficult to measure precisely in the muon spectectomer, due to the fact that they are less bended by the magnetic field. The reconstruction and identification efficiency were obtained applying the same *tag and probe* technique used for the electrons, using $Z \to \mu^+\mu^-$, $J/\Psi \to \mu^+\mu^-$ and $Y \to \mu^+\mu^-$ samples as explained in Ref. [102].

In the muon transverse momentum range of $6 < p_T < 100$ GeV, the muon momentum scale is known with 0.04% precision in the barrel region, decreasing to 0.2% in the region with $|\eta| > 2$ [102]. However, most of top-quark physics analyses use a cut at $\eta = 2.5$, which is the region covered by the ID. This is translated into a good muon precision for top-quark physics analyses.

4.1.5 Jets

Jets are showers of particles from the hadronization of quarks or gluons produced after the collision. The resulting stable particles leave tracks in the ID, if charged, and clusters of energy deposits in the calorimeters. Reconstructing the total energy and position of a cluster allows for an estimation of the four-momentum of the jet from which the shower is originating.

Jets are typically reconstructed and clustered from calorimeter energy deposits in the form of three-dimensional topological clusters (topo-clusters) of calorimeter cells in the electromagnetic and hadronic calorimeters [103]. The topological clustering algorithm is based on the significance of the energy deposits in the calorimeter cells E_{cell} with respect to their noise level σ_{cell}, where the significance is defined as $E_{\text{cell}} > 4\sigma_{\text{cell}}$. The noise level σ_{cell} is the sum in quadrature of the standard deviations of the distributions of electronic and estimated pile-up noise. The cell with the most significant signal is identified as the seed cell of the topocluster. All adjacent secondary cells with $E_{\text{cell}} > 2\sigma_{\text{cell}}$ are iteratively added to the topo-cluster until no new adjacent cell satisfies this requirement. This procedure is performed in three-dimensional space and allows topo-clusters to form between the different calorimeters. The topo-cluster is completed by adding adjacent cells with $E_{\text{cell}} > 0$. The entire procedure is applied for each seed cell until all are formed into topo-clusters.

The topological clusters are reconstructed at the so-called electromagnetic energy scale (EM), which corresponds to the energy deposited by particles in an electromagnetic shower in the calorimeter. The EM scale correctly reconstructs the energy deposited by particles in an electromagnetic shower in the calorimeter. This energy scale is established using test-beam measurements [104–107]. The topological clusters are used to form jets, through a process referred to as jet reconstruction or jet finding. The most-used jet algorithm in ATLAS is the anti-k_{T} jet algorithm [108] which combines topological clusters iteratively, based on a distance parameter criterium. The distance parameter is defined as:

$$d_{ij} = min(\frac{1}{p_{T_i}^2}, \frac{1}{p_{T_j}^2})\frac{\Delta R_{ij}^2}{R^2} \tag{4.1}$$

where p_{T_i} is the transverse momentum of topo-cluster i, $\Delta R_{ij} = \sqrt{(\Delta\eta_{ij})^2 + (\Delta\phi_{ij})^2}$ is the distance between topo-clusters i and j, and R is a parameter of the algorithm that controls the size of the jet and is chosen to be 0.4. The algorithm computes the distance between two topo-clusters i and j (d_{ij}) and the distance between the input i and the beam axis (d_{iB}). If d_{iB} is the smallest in the computation, the topo-cluster i is considered a jet and removed from the list and then the algorithm repeats the procedure with the remaining input objects. Otherwise, the i and j topoclusters corresponding to the smallest distance d_{ij} are combined, and the list is updated for a new iteration. The procedure is repeated until the list is empty. These jets then undergo several steps of calibration that account for calorimeter non-compensation,[3] dead material, shower leakage, and pile-up are used to calibrate the jet energy scale (JES) [109, 110]. The calibration scheme used in the dissertation is called the Local Cluster Weighting (LCW) calibration. In this calibra-

[3]The calorimeter non-compensation is defined as the difference in scales of the energy measured from hadronic and electromagnetic showers.

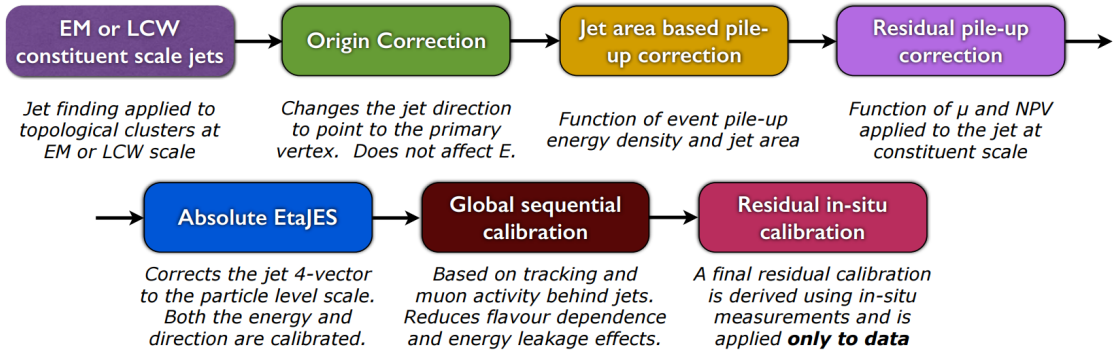


Figure 4.6 – *The different step of the calibration of EM and LCW jets [109].*

tion scheme the cluster shape variables are used to separate the topological clusters into hadronic and electromagnetic ones. Correction factors are applied to hadronic clusters in order to correct for non-compensation effects. Further corrections are applied to both type of clusters in order to correct for the energy lost outside of the cluster as well as in dead material. The calibration steps are shown in Fig. 4.6 and are described in the following.

The first step of the jet calibration is the *origin correction* which changes the direction of the jet so that it originates from the primary vertex instead of the geometrical center of the ATLAS detector. This correction improves the η (and ϕ) resolution of the jet without affecting the jet energy.

The second step is the *pile-up correction* which is intended to mitigate the effect of pile-up interactions. The contribution to the jet energy from pile-up is estimated event-by-event from the product of the area of the jet and the transverse momentum density, in the $\eta \times \phi$ plane. Also, the dependency of the jet transverse momentum on the number of primary vertices and the average number of interactions per bunch crossing are considered in the correction. After the pile-up subtraction, a MC-based correction is applied. The reconstructed jet energy is corrected to the energy of the simulated jet (MC jet) to which it corresponds to. The reconstructed jet is considered to be matched to a MC jet if the distance between them is $\Delta R < 0.3$.

Additional corrections are then applied based on the topology of energy deposits in the calorimeter, properties of the tracks associated with the jet, and any associated energy deposits in the muon spectrometer. These corrections are referred to as the global sequential calibration (GSC) and are applied sequentially such that the mean of the jet energy is unchanged. The information from the muon spectrometer is relevant for jets with very high p_T that reaches the muon detector system (called *punch-through*), which improves the jet energy resolution [111].

The final step of the jet calibration chain consists of a number of data-driven (in-situ) measurements, which are designed to validate the MC-based corrections described above and derive residual corrections to the JES that are applied to jets in data. These corrections are based on the comparison of the p_{T} of reference objects (photons, Z bosons or other jets) and the jets to be calibrated using data and Monte Carlo simulation. The first of these corrections is referred to as the dijet η-intercalibration [112], in which jets in the forward region ($0.8 < |\eta| < 4.5$) of the ATLAS detector are calibrated relative to jets in the central region ($|\eta| < 0.8$). The η-intercalibration

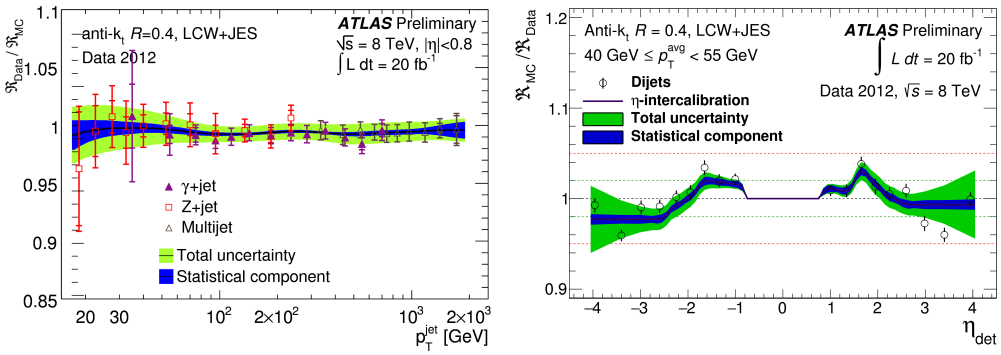


Figure 4.7 – *Ratio of response measured in data to response measured in MC for Z+jet, γ+jet and multi-jet balance in-situ analyses [109].*

correction factors are generally below 2%.

The jet performance is quantified by the precision of the determination of the jet energy scale and the energy resolution (JER). Figure 4.7 shows the JES as a function of p_T and η for 40 GeV < p_T < 55 GeV. The precision of the JES is better than 3% for jets with $p_T > 50$ GeV, while it gets worse for jets with lower p_T, especially in the forward regions [109]. However, the usual cut performed in top-quark physics analyses is of $|\eta| < 2.5$, where the precision of the resolution is significantly better. The JER is measured in multi-jet events by probing the jet imbalance on the transverse plane, as described in Ref. [113]. The JER is of the order of 10% for jets with $p_T > 100$ GeV, growing to 20% for jets with $p_T \sim 30$ GeV.

b-tagging

Several algorithms can be used to identify jets containing the fragmentation products of b-quarks. These algorithms take advantage of the fact that b-hadrons have a relatively high lifetime of $\sim$1.5 ps, which leads to long decay lengths. These can be reconstructed as displaced secondary vertices, which are used in *secondary vertex* (SV) b-tagging algorithms [114]. The displaced decays of b-hadrons produce tracks with large impact parameters with respect to the primary vertex (Fig. 4.1). These tracks are used as input to *impact parameter* (IP) b-tagging algorithms [115]. The semi-leptonic branching ratio of b-hadrons is $\sim$11% for each lepton flavor. This branching ratio is as high as $\sim$20% when $b \to c$ cascade decays are taken into account. The cascade produces another displaced vertex along the flight path of the b-hadron. The topology of the b-hadron decay chain is exploited by the *JetFitter* algorithm [116].
The SV, IP, and JetFitter algorithms are combined into a neural network, known as the MV1 tagger [117], which is used in the analyses presented in this work. The efficiency of the MV1 algorithm to correctly identify b-jets or for mis-identifying charm (c) or light flavor (LF) jets as b-jets, has been measured in samples of $t\bar{t}$, D-mesons and dijet events respectively [118]. The corresponding efficiencies for the operating point corresponding to an inclusive b-jet efficiency of 70% is shown in Fig. 4.8 as a function of the transverse momentum and η of the jet. The mis-identification probability is of around 0.8% for light jets, and 20% for c-quark jets.

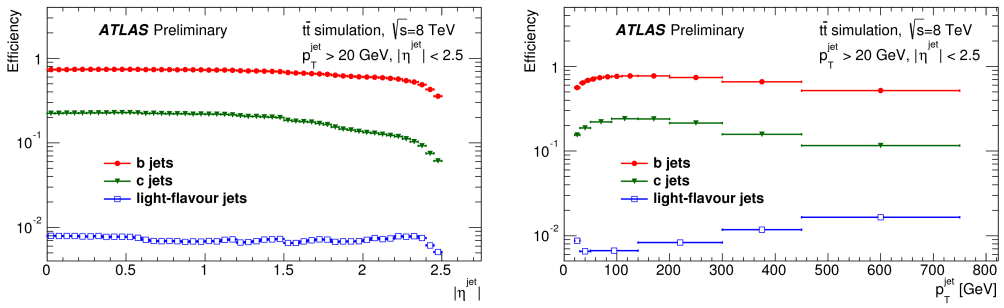


Figure 4.8 – *b-jet, c-jet and light jets identification efficiency as a function η and p_T of the jet using the MV1 tagger algorithm [117].*

4.1.6 Missing Transverse Momentum

The Missing Tranverse Momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$) is computed by combining all the topo-clusters found in the calorimeters. The topo-clusters are associated with reconstructed objects: electrons, photons, jets and muons. Such topo-clusters are then calibrated according to the correction factors applied to the reconstructed physics objects to which they are associated. The contribution of these objects to the $E_{\mathrm{T}}^{\mathrm{miss}}$ computation is known as the *hard terms*. Topo-clusters that are not associated with any physics object are also considered and a dedicated calibration is applied. The contribution of these objects to the $E_{\mathrm{T}}^{\mathrm{miss}}$ computation is known as the *soft terms*. The $E_{\mathrm{T}}^{\mathrm{miss}}$ and its components are defined as:

$$E_{\mathrm{T}}^{\mathrm{miss}} = |-\Sigma\vec{p_{\mathrm{T}}}| = \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2} \tag{4.2}$$

$$E_x^{\mathrm{miss}} = -\Sigma\vec{p_x} \tag{4.3}$$

$$E_y^{\mathrm{miss}} = -\Sigma\vec{p_y} \tag{4.4}$$

where E_x^{miss} and E_y^{miss} correspond to projection of the $E_{\mathrm{T}}^{\mathrm{miss}}$ on the $x-$ and $y-$axis, respectively. The resolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ is estimated typically in $Z \to \ell\ell$ events since no real $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected. The resolution is estimated from the width of the combined distribution of E_x^{miss} and E_y^{miss} as a function of the total transverse energy in the detector ΣE_{T}. The distributions are fitted, for each ΣE_{T} bin, with a Gaussian. The fitted width, σ, is studied as a function of ΣE_{T}. The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution as a function of ΣE_{T} follows a behavior that can be approximated with the function $\sigma = k \cdot \sqrt{\Sigma E_{\mathrm{T}}}$, but deviations from this simple law are expected as shown in Ref. [119]. The constant factor k is found to be 0.97 $\mathrm{GeV}^{1/2}$ for collisions at 8 TeV [120].

4.2 Selection and background estimation

In order to enrich the data sample in dileptonic $t\bar{t}$ events, requirements are imposed on reconstructed charged leptons (electrons and muons), jets, and the missing transverse momentum. Three different final states are considered in the analyses presented : events with two electrons in the final state (ee), with one electron and one muon ($e\mu$), and with two muons ($\mu\mu$). The leptons coming from τ decays are incorporated to the respective channels. In the following, the selection criteria on the objects are presented, as well as studies performed on the background events.

4.2.1 Object selection

Electron candidates passing the likelihood-based *medium* identification requirements are required to have a transverse momentum $p_T > 25$ GeV and must also lie in the region $|\eta_{\mathrm{cl}}| < 2.47$, where η_{cl} is the pseudorapidity of the calorimeter energy cluster associated with the electron, excluding the transition region between the calorimeter barrel and endcaps $1.37 < |\eta_{\mathrm{cl}}| < 1.52$. Moreover, electrons are required to be isolated from surrounding activity in the inner detector. The scalar sum of the track p_{T} present in the event within a cone of $\Delta R = 0.3$ (excluding the track of the electron itself) divided by the electron p_{T} is required to have certain values as a function of η_{cl} and p_{T} in order to achieved an efficiency of 90%.

Muon candidates reconstructed using the Muid algorithm are required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. In addition, they are required to satisfy track-based p_{T}-dependent isolation criteria. The scalar sum of the track p_{T} within a cone of size $\Delta R = 10\ \mathrm{GeV}/p_{\mathrm{T}}^{\mu}$ around the muon (excluding the muon track itself) must be less than 5% of the muon p_{T} (p_{T}^{μ}).

Both the electrons and muons have to be consistent with the primary vertex by requiring the absolute value of the longitudinal impact parameter to be less than 2 mm.

Jets are reconstructed from clustered energy deposits in the electromagnetic and hadronic calorimeters, using the anti-k_t algorithm with a radius parameter $R = 0.4$. After applying the jet calibration and energy correction, described in Sec. 4.1.5, the jets are required to have $p_{\mathrm{T}} > 25$ GeV, to be in the pseudorapidity range $|\eta| < 2.5$, and to have a jet vertex fraction[4] $|\mathrm{JVF}| > 0.5$ [121] if $p_{\mathrm{T}} < 50$ GeV, for pile-up mitigation. The jet that is the closest to a selected electron is removed from the event if their separation is $\Delta R < 0.2$. After this jet overlap removal, electrons and muons that are within a cone of $\Delta R = 0.4$ around the closest jet are removed.

4.2.2 Event selection

Events are required to fire a single electron or single muon trigger [122, 123]. For electrons, the logical OR of two single electron triggers is used: one corresponding to a threshold of 24 GeV for the electron cluster E_{T} (EF_e24vhi_medium1) and the other corresponding to a threshold of 60 GeV (EF_e60_medium1). For the low threshold trigger, an isolation requirement is applied, by requiring the sum of the transverse momenta of the tracks within a cone of radius $R = 0.2$ around the electron-candidate track (excluding the electron track) to be less than 10% of the trigger electron's transverse energy. This isolation cut is applied in order to reduce the high rate of electrons produced in hadron decays. The isolation requirement becomes inefficient at $E_{\mathrm{T}} > 60$ GeV, since high-p_T electrons produce more bremsstrahlung with intrinsic e^-e^+ pair creation, and therefore have a higher probability of failing the isolation cut. For muons, the OR of two unprescaled single muon triggers is used: one corresponding to a threshold of 24 GeV with isolation requirements (EF_mu24i_tight) and the other corresponding to a threshold of 36 GeV with no isolation requirements (EF_mu36_tight). The muon isolation requirement at the trigger level is applied by requiring the sum of the tracks in a cone of radius $R = 0.2$ around the muon candidate track (excluding the muon track itself) to be less than 12% of the muon track's p_{T}.

[4] The jet vertex fraction is defined as the fraction of the scalar p_{T} sum of tracks associated with the jet and the primary vertex and the scalar p_{T} sum of tracks associated with the jet and any vertex.

Charge asymmetry measurement

In the three final states, exactly two isolated leptons with opposite charge and an invariant mass $m_{\ell\ell} > 15$ GeV are required, together with at least two jets. The first requirement is applied in order to reject $Z/\gamma^* \to \ell\ell$ events and meson production such as J/Ψ that could enter the selection by decaying to two opposite sign leptons, while the second removes most of the non-$t\bar{t}$ SM events. In the same-flavor channels (ee and $\mu\mu$), the invariant mass of the two charged leptons is required to be outside of the Z boson mass window, such that $|m_{\ell\ell} - m_Z| > 10$ GeV, which reduces the amount of Drell–Yan processes drastically in the ee and $\mu\mu$ channels. Furthermore, it is required that $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV and at least one of the jets must be b-tagged. The $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is intended to suppress Drell–Yan production of $Z/\gamma^* \to \ell\ell$ (with $\ell = e/\mu$) and also suppress diboson backgrounds. In Drell–Yan processes no real $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected since there are no neutrinos from the hard scatter that result in missing energy. In the case of dibosons events, few or no neutrinos are expected in the final state. In the $e\mu$ channel, the background contamination from the different processes is much smaller and the background suppression is achieved by requiring the scalar sum of the p_{T} of the two leading jets and leptons (H_{T}) to be larger than 130 GeV. This requirement rejects background events with low energy such as $Z \to \tau\tau$ + jets and it is found to be a very loose requirement. No $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is applied to this channel since the events have a real contribution from $Z \to \tau\tau$ decays. The event selection requirements for the charge asymmetry measurements are summarized in Table 4.2.

Table 4.1 – *Summary of the event selection requirements applied in the different channels for the charge asymmetry measurements.*

Requirements	$ee/\mu\mu$	$e\mu$
Leptons	== 2	== 2
Jets	≥ 2	≥ 2
b-tagged jets	≥ 1	–
$m_{\ell\ell}$	> 15 GeV	> 15 GeV
$\lvert m_{\ell\ell} - m_Z \rvert$	> 10 GeV	–
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 30 GeV	–
H_{T}	–	> 130 GeV

Top spin measurement

The same event selection as in the charge asymmetry analysis is used, but a small modification is applied. In this selection, at least one b-tagged jet is required in all three decay channels. This was not the case for the charge asymmetry event selection where it was only applied in the same flavor channel. By applying this requirement, a purer sample is obtained without reducing dramatically the amount of statistics. A reduction of the background in the spin observables is important since their shapes affect the value of the observables more than in the charge asymmetry. Furthermore, the charge asymmetry measurements are more dominated by the statistical uncertainty than in the spin observables measurements. It is thus important to keep as much statistics as possible.

The event selection requirements for the top spin measurements are summarized in Table 4.2.

Table 4.2 – *Summary of the event selection requirements applied in the different channels for the top spin measurements.*

Requirements	$ee/\mu\mu$	$e\mu$
Leptons	== 2	== 2
Jets	≥ 2	≥ 2
b-tagged jets	≥ 1	≥ 1
$m_{\ell\ell}$	> 15 GeV	> 15 GeV
$\lvert m_{\ell\ell} - m_Z \rvert$	> 10 GeV	–
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 30 GeV	–
H_{T}	–	> 130 GeV

4.2.3 Signal modeling

The nominal $t\bar{t}$ signal sample is generated at NLO in QCD using Powheg-hvq and the CT10 PDF set for the hard scattering interaction. The h_{damp} parameter is set to the top-quark mass of 172.5 GeV. The parton shower, hadronization, and underlying event are simulated using Pythia6 +Jimmy with the CTEQ6L1 PDF and the corresponding set of tunable parameters Perugia 2011C [93] intended to be used with this PDF. The $t\bar{t}$ cross section for pp collisions at a center-of-mass energy of 8 TeV is rescaled to $\sigma_{t\bar{t}} = 253^{+13}_{-15}$ pb, calculated at NNLO in QCD including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms with top++2.0 [20, 124–129].

4.2.4 Background estimation

The main background contribution in the $t\bar{t} \to \nu\bar{\nu}\ell^+\ell^- b\bar{b}$ system comes from Drell–Yan production of $Z/\gamma^* \to \ell\ell$ in the same flavor channel, which is estimated by a combination of simulated samples modified with corrections derived from data. These modifications scale the background expectations in order to make them agree with the observations. Contributions from diboson (WW, ZZ, and WZ) and single-top-quark (Wt channel) production are evaluated purely via MC simulations.

Further background contributions can arise from events including a jet or a lepton from a semileptonic hadron decay that is misidentified as an isolated charged lepton as well as leptons from photon conversions, together referred to as *fake leptons*. This contribution is estimated using simulated samples, modified with corrections derived from data. The samples mentioned above, together with simulated samples of $t\bar{t} + W/Z$, t-channel of single-top-quark production, W+jets, and $W+\gamma$+jets are used for the estimation.

Single-top-quark production in the Wt channel is simulated using Powheg-hvq with Pythia6 and the CT10 (NLO) PDF set. The cross section of 22.3 ± 1.5 pb is estimated at approximate NNLO in QCD including resummation of NNLL terms [130]. The parton shower, hadronization, and underlying event are simulated by Pythia6 +Jimmy using the Perugia 2011C tune. The Drell–Yan process is modeled using Alpgen interfaced with Pythia6 with the CTEQ6L1 PDF set using the

MLM matching scheme. Its heavy-flavor component is included in the matrix element calculations to model the $Z/\gamma^* + b\bar{b}$ and $Z/\gamma^* + c\bar{c}$ processes. Diboson processes (WW, ZZ, and WZ) are simulated using ALPGEN interfaced with HERWIG +JIMMY with the CTEQ6L1 PDF set for parton fragmentation. The only exceptions are the same-charge $W^{+(-)}W^{+(-)}$ samples, which are simulated using MADGRAPH interfaced with PYTHIA8. The samples are normalized to the reference NLO QCD prediction, obtained using the MCFM generator [131]. The associated production of a $t\bar{t}$ pair with a vector boson ($t\bar{t}Z$ and $t\bar{t}W$) is simulated with MADGRAPH, interfaced with PYTHIA8, and normalized to NLO cross-section calculations [132, 133]. The W+jets events are simulated using ALPGEN, interfaced with PYTHIA6, and the $W+\gamma$+jets process is simulated using ALPGEN interfaced with JIMMY.

Z+jets modeling

The modeling of Drell–Yan events in the same-flavor channels with $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV may not be accurate in simulation due to a mismodeling of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution. The Drell–Yan background does not contain any real $E_{\mathrm{T}}^{\mathrm{miss}}$. Non-negligible $E_{\mathrm{T}}^{\mathrm{miss}}$ can appear in a fraction of events with misreconstructed objects, which are difficult to model. Moreover, after applying the b-tagging requirement, a large contribution to the background comes from the associated production of Z bosons with heavy-flavor jets, which is not well predicted by MC simulation. These problems are shown in Fig. 4.9. Before applying b-tagging there is an overestimation of MC events in the Z peak, while after applying b-tagging an underestimation is observed.

The $E_{\mathrm{T}}^{\mathrm{miss}}$ mismodeling depends on the reconstructed objects and the corrections to be applied (scale factors) are therefore expected to be different between the ee and $\mu\mu$ channel. For the heavy flavor normalization in Z+jets events, the correction is expected to be the same in the ee, $\mu\mu$, and $e\mu$ channel since the modeling problem does not depend on the topology of the event. The normalization of the inclusive and heavy-flavor components of the Drell–Yan background in the same-flavor channels is computed simultaneously using data in three control regions with three scale factors. Two scale factors are applied to Drell–Yan events to take into account the mismodeling from the $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement (one in the ee and one in the $\mu\mu$ channel) while another one is applied only to Z+heavy-flavor events. Since real $E_{\mathrm{T}}^{\mathrm{miss}}$ is present in $Z \to \tau\tau$ events, the scale factors related with the $E_{\mathrm{T}}^{\mathrm{miss}}$ mismodeling are not applied to these samples. The normalization correction for heavy flavor processes are applied on top of the previously defined scale factors. The calculation is performed in control regions defined by the same cuts as the signal region, but inverting the Z mass windows. The control regions are defined using the selection criteria with b-tagging requirements on the selected jets (b-tagged region) and without b-tagging requirements on the selected jet (pre-tagged region) as follows:

- ee channel, no b-tagging requirement
- $\mu\mu$ channel, no b-tagging requirement
- Combined channels, at least one b-tagged jet

All the constraints previously described lead to the following system of equations:

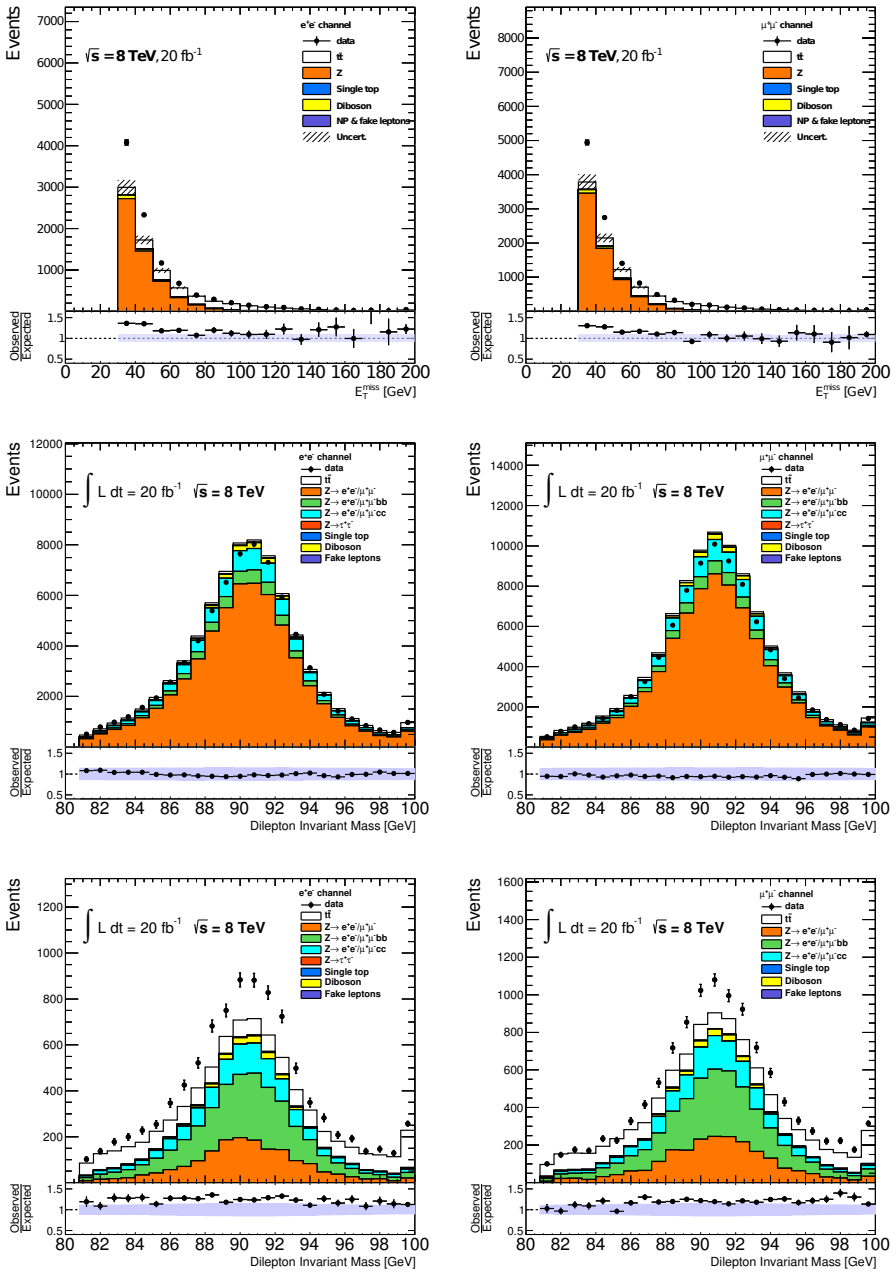


Figure 4.9 – *E_T^{miss} for ee (top left) and μμ (top-right) in the Z mass window after b-tagging. The dilepton invariant mass in the pretag region for ee (middle left) and μμ (middle right) and in the b-tag region for ee (bottom left) and μμ (bottom right) is also shown. The data/MC ratio is shown in the bottom panel of each figure. The shaded area shows the uncertainty on the signal and background normalization.*

$$\begin{cases} C_{ee} \cdot (N_{Z+light}^{ee,pretag} + C_{HF} \cdot N_{Z+HF}^{ee,pretag}) + C_{HF} \cdot N_{Z\to\tau\tau+HF}^{ee,pretag} + N_{Z\to\tau\tau}^{ee,pretag} + N_{others}^{ee,pretag} = N_{data}^{ee,pretag} \\ C_{\mu\mu} \cdot (N_{Z+light}^{\mu\mu,pretag} + C_{HF} \cdot N_{Z+HF}^{\mu\mu,pretag}) + C_{HF} \cdot N_{Z\to\tau\tau+HF}^{\mu\mu,pretag} + N_{Z\to\tau\tau}^{\mu\mu,pretag} + N_{others}^{\mu\mu,pretag} = N_{data}^{\mu\mu,pretag} \\ C_{ee} \cdot (N_{Z+light}^{ee,btag} + C_{HF} \cdot N_{Z+HF}^{ee,btag}) + C_{\mu\mu} \cdot (N_{Z+light}^{\mu\mu,btag} + C_{HF} \cdot N_{Z+HF}^{\mu\mu,btag}) + C_{HF} \cdot N_{Z\to\tau\tau+HF}^{all,btag} + N_{Z\to\tau\tau}^{all,btag} + N_{others}^{all,btag} = N_{data}^{all,btag} \end{cases} \quad (4.5)$$

where,

- C_{ee} is the scale factor applied to $Z \to ee$ events.
- $C_{\mu\mu}$ is the scale factor applied to $Z \to \mu\mu$ events.
- C_{HF} is the scale factor applied to $Z + HF$ events.
- N_{Z+X} is the number of events for the $Z + X$ processes, where X correspond to either *light* of *HF*.
- N_{data} is number of observed events.
- N_{others} is the number of events of all the other processes

The system of equations is solved numerically by minimizing the likelihood of the system. The scale factors that correct for $E_{\mathrm{T}}^{\mathrm{miss}}$ mismodeling are 0.927 ± 0.005 for the *ee* channel and 0.891 ± 0.004 for the $\mu\mu$ channel, while the normalization factor for heavy flavour processes is 1.70 ± 0.03. The uncertainties correspond to statistical uncertainty. In Fig 4.10 the Z mass windows after applying the derived scale factors as well as the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution are shown. The agreement in the normalization after applying the scale factors is significantly improved.

Fakes leptons

The background arising from fakes leptons is determined using both MC simulation and data. The dominant sources of these fake leptons are semileptonic b-hadron decays, long-lived weakly decaying states (such as $\pi^{\pm}$ or $K^{\pm}$ mesons), π^0 showers, photons reconstructed as electrons, and electrons from photon conversions. W+jets, $W+\gamma$+jets, $t\bar{t}$, $t\bar{t}Z$, $t\bar{t}W$, Drell–Yan, single-top-quark, and diboson production are considered when estimating this background. Multijet events do not contribute significantly to this background since the probability of having two jets misidentified as isolated leptons is very small. The shapes of the kinematic distributions are taken from simulated events where at least one of two selected leptons is required not to be matched with a MC generator-level lepton. Scale factors are derived from data in order to adjust the normalization.

The control regions used for the scale factor estimation are defined by applying the same selection criteria as in the signal region but by requiring the two leptons to have the same electric charge. The number of events in data and MC in the control regions used for the estimation are shown in Tab. 4.3. There is a general underestimation of the MC simulation in the three channels, being the more prominent case the $\mu\mu$ channel. A scale factor is needed in order to correct for these effects. The scale factors derived for each dilepton channel are calculated as

$$SF_{\mathrm{fakes}} = \frac{N_{\mathrm{data}} - \sum_{i=\mathrm{other}} N_i}{N_{\mathrm{fakes}}}.$$

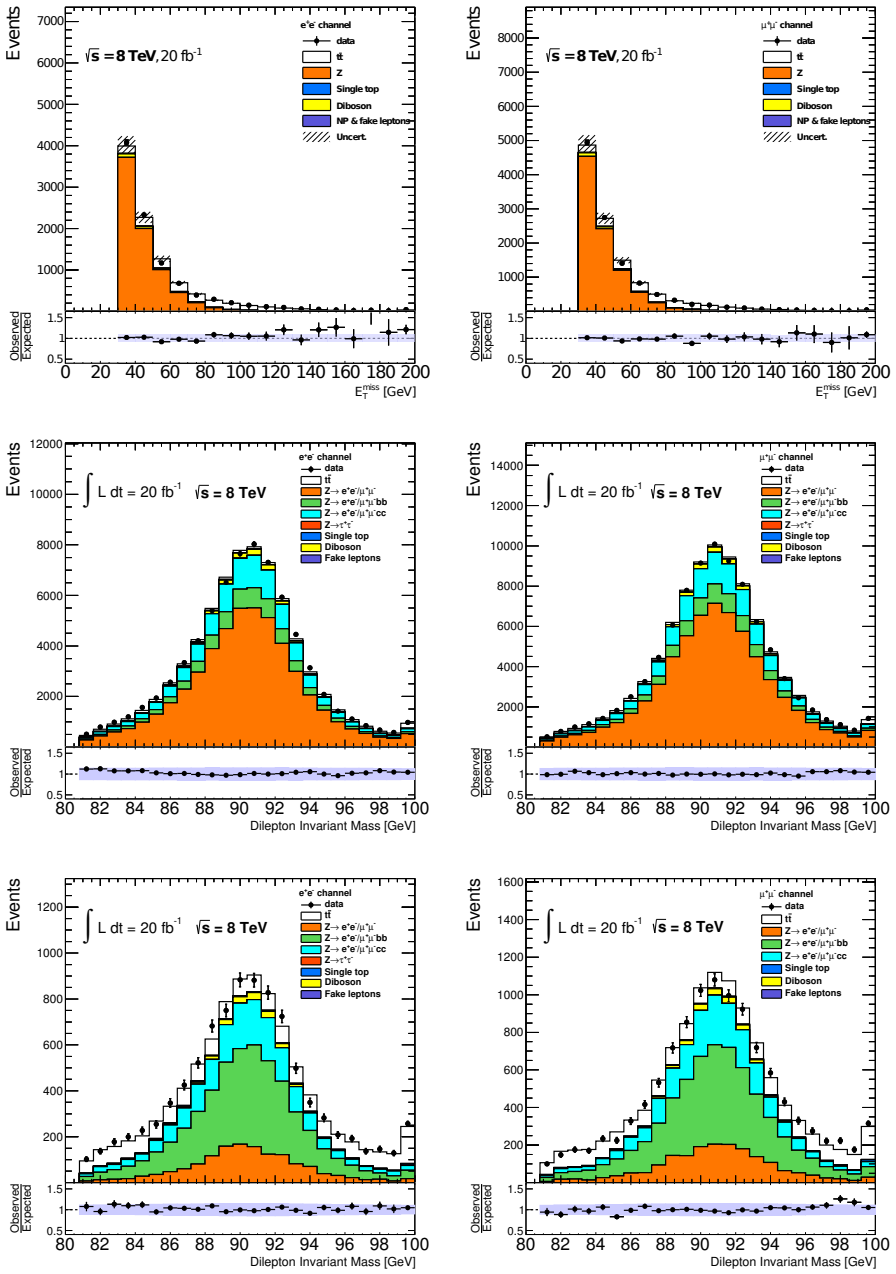


Figure 4.10 – *E_T^{miss} for ee (top left) and μμ (top right) in the Z mass region after b-tagging. The dilepton invariant mass in the pretag region for ee (middle left) and μμ (middle right) and in the b-tag region for ee (bottom left) and μμ (bottom right) after applying the derived scale factors is also shown. The data/MC ratio is also shown. The shaded area shows the uncertainty on the signal and background normalization.*

Table 4.3 – *Numbers of events in data and MC in the fakes control region. Only statistical uncertainties are shown.*

Channel	ee	$\mu\mu$	$e\mu$ - no b-tag	$e\mu$ - b-tag
$t\bar{t}$	127.0 ± 2.8	0.40 ± 0.16	210 ± 4.8	173.0 ± 4.3
$t\bar{t}$ + W/Z	12.47 ± 0.29	13.00 ± 0.31	41.1 ± 0.54	33.03 ± 0.48
Fakes	50.8 ± 4.4	10.9 ± 1.0	351 ± 32	121 ± 20
Dibosons	6.02 ± 0.90	7.9 ± 1.5	272 ± 6.3	17.0 ± 1.6
Single Top	7.22 ± 0.71	< 0.001	11.2 ± 0.90	7.98 ± 0.73
$Z + jets$	19.8 ± 2.5	0.060 ± 0.040	8.6 ± 1.7	1.4 ± 1.2
Total Expected (MC)	223.3 ± 6.0	32.2 ± 1.9	888 ± 33	353 ± 21
Data	232	62	945	406
MC/Data	0.96	0.52	0.95	0.87
SF	1.20 ± 0.3	3.70 ± 0.8	1.10 ± 0.2	1.44 ± 0.3

where N_i is the number of MC events without the fakes contribution, N_{fakes} is the number of fakes events, and N_{data} is the number of data events.

The differences between the scale factors originate from the sources of misidentified electrons and muons, which seem to be modeled better in MC simulation for the electrons. However, the shapes of the distributions of the relevant kinematic variables in the $\mu\mu$ channel, such as the jet multiplicity, jet p_T, missing E_T, and lepton p_T in the $\mu\mu$ channel, are cross-checked with a purely data-driven method, known as the *matrix method* (MM) [134]. The matrix method is based on the selection of two categories of events: events that satisfy loose lepton selection requirements (N^{loose}), and events that satisfy tight lepton selection requirements (N^{tight}). Together with the efficiency $\epsilon = N^{\text{loose}}/N^{\text{tight}}$, which is estimated using data, it is possible to compute the number of fakes events that passes the tight selection criteria. The tight muon selection criteria corresponds to the selection criteria mentioned in Sec. 4.2.1. The comparison between the MC-driven fakes and the MM-driven fakes for different kinematic variables is shown in Fig. 4.12. The MC-driven fakes are normalized to the MM-driven fakes in order to evaluate the agreement in the shapes of the distributions. Both fake estimations are in agreement with each other. The comparison between data and the MC prediction in the same-sign control region after applying the fakes scale factors SF_{fakes} is shown in Fig. 4.11.

4.2.5 Comparison between Data and MC

Charge asymmetry measurements

The number of events for both expectation and data after applying the selection criteria corresponding to the charge asymmetry measurements are shown in Table 4.4 for the three final states. The uncertainties shown correspond to the total uncertainty, which includes: the statistical uncertainties coming from the limited size of the MC samples, the normalization uncertainties, from the uncertainty on the theoretical cross-section predictions, and the systematic

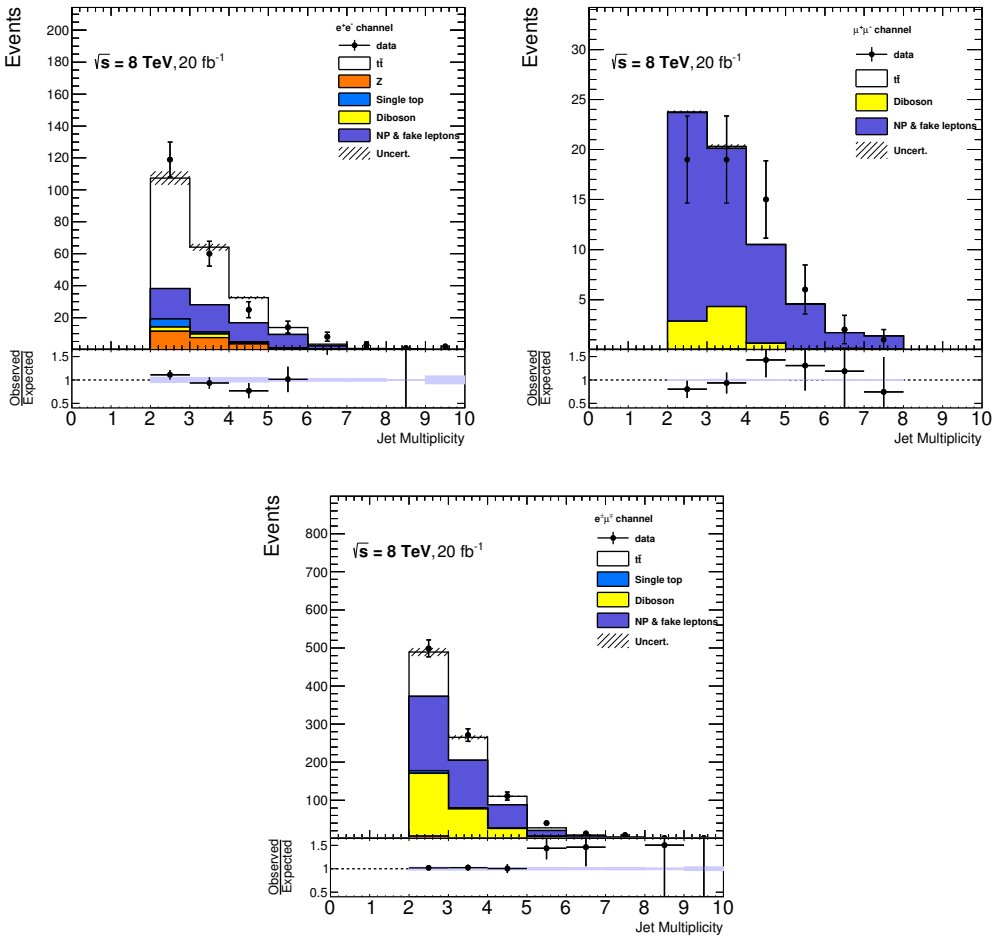


Figure 4.11 – *The jet multiplicity in the same-sign control region in the ee (top-left), μμ (top-right), and eμ (bottom) channel. Here, the comparison is done only in the control regions corresponding to the signal regions, i.e. in the b-tag same-sign region for ee and μμ channels and in the pretag same-sign region for eμ channel. The grey area in the ratio plot corresponds to the normalization uncertainty.*

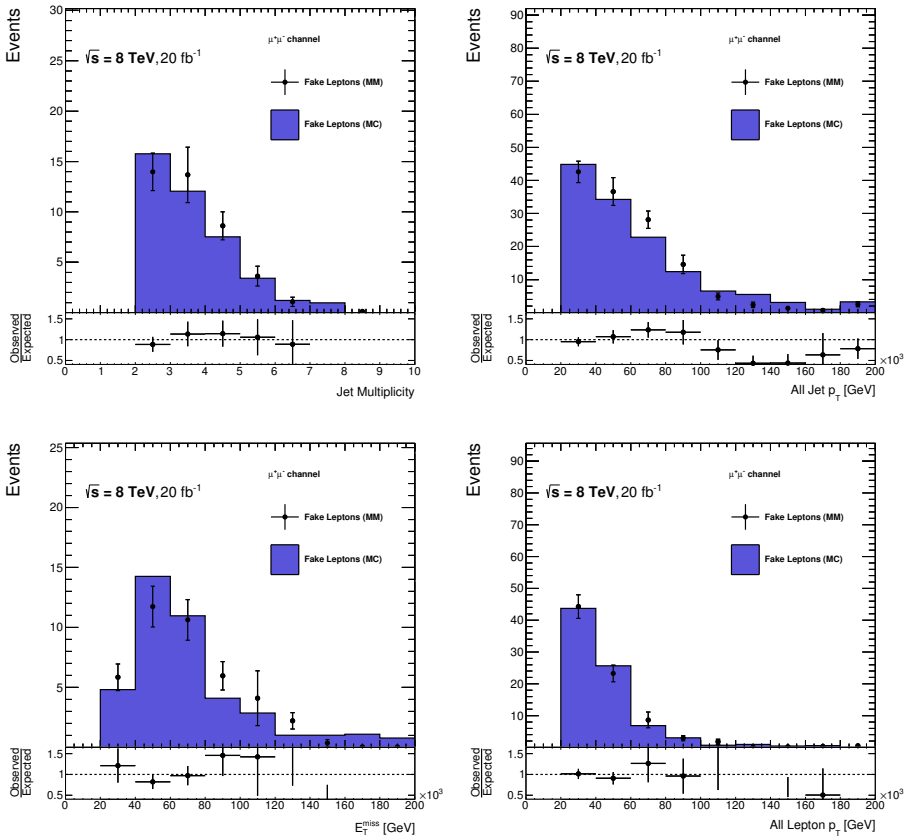


Figure 4.12 – *The comparison between the fakes estimated from MC and by MM method in the SS control region. The jet multiplicity, jet p_T, missing E_T, and lepton p_T are shown in the $\mu\mu$ channel. The MC estimation is normalized to the MM estimation.*

uncertainties which are described in Chapter 7. In the case of the uncertaitnies related with NP & fakes leptons, a Bayesian procedure procedure assuming constant a priori probability for a non-negative number of events was assumed. The $e\mu$ channel contributes with the largest number of events, followed by $\mu\mu$ and ee. The jet multiplicity, the lepton η, and lepton p_{T} distributions are shown in Fig. 4.14 for the three channels combined. There is about 5% more data than MC, however, the overall agreement is acceptable within the uncertainties. A slope observed in the data/MC ratio of the transverse momentum of the leptons is a known mismodeling effect and it is consistent with the differential cross section measurement performed by ATLAS [30].

Table 4.4 – *Observed numbers of data events compared to the expected signal and background contributions in the three decay channels and their combination for the charge asymmetry measurements. The uncertainty corresponds to the total uncertainty on the given process. Data-driven scale factors are applied to the Z+jets and the non-prompt (NP) & fake leptons contributions. The $Z \to \tau\tau$ process in the $e\mu$ channel is estimated using MC simulation only. All the uncertainties are assumed to be uncorrelated.*

Channel	ee	$\mu\mu$	$e\mu$	$\ell\ell$
$t\bar{t}$	10200 ± 800	12100 ± 800	36000 ± 2400	58300 ± 2700
Single-top	510 ± 50	590 ± 50	1980 ± 170	3080 ± 180
Diboson	31 ± 5	40 ± 6	1320 ± 100	1400 ± 100
$Z \to ee$	1200 ± 260	–	–	1200 ± 260
$Z \to \mu\mu$	–	1520 ± 300	–	1520 ± 300
$Z \to \tau\tau$	31 ± 15	58 ± 25	1120 ± 430	1210 ± 430
NP & fake leptons	62^{+119}_{-29}	45^{+36}_{-24}	480^{+240}_{-220}	590^{+270}_{-220}
Total Expected	12000 ± 900	14400 ± 800	40900 ± 2500	67300 ± 2800
Data	12785	14453	42363	69601

Top spin measurements

The number of events for both expectation and data after applying the selection criteria corresponding to the top spin measurements are shown in Table 4.5 for the three final states. The uncertainties shown correspond to the total uncertainty, which includes: the statistical uncertainties coming from the limited size of the MC samples, the normalization uncertainties, from the uncertainty on the theoretical cross-section predictions, and the systematic uncertainties which are described in Chapter 7. The same behavior is observed as for the charge asymmetry measurements: the $e\mu$ channel contributes with the largest number of events, followed by $\mu\mu$ and ee. The contribution of the $e\mu$ channel is $\sim$20% smaller than in the selection used for the charge asymmetry. The jet multiplicity, the lepton η, and lepton p_{T} distributions are shown in Fig. 4.13 for the three channels combined. As before, there is about 5% more data than MC, however, the overall agreement is acceptable within the uncertainties. A slope is also observed in the data/MC ratio of the transverse momentum of the leptons.

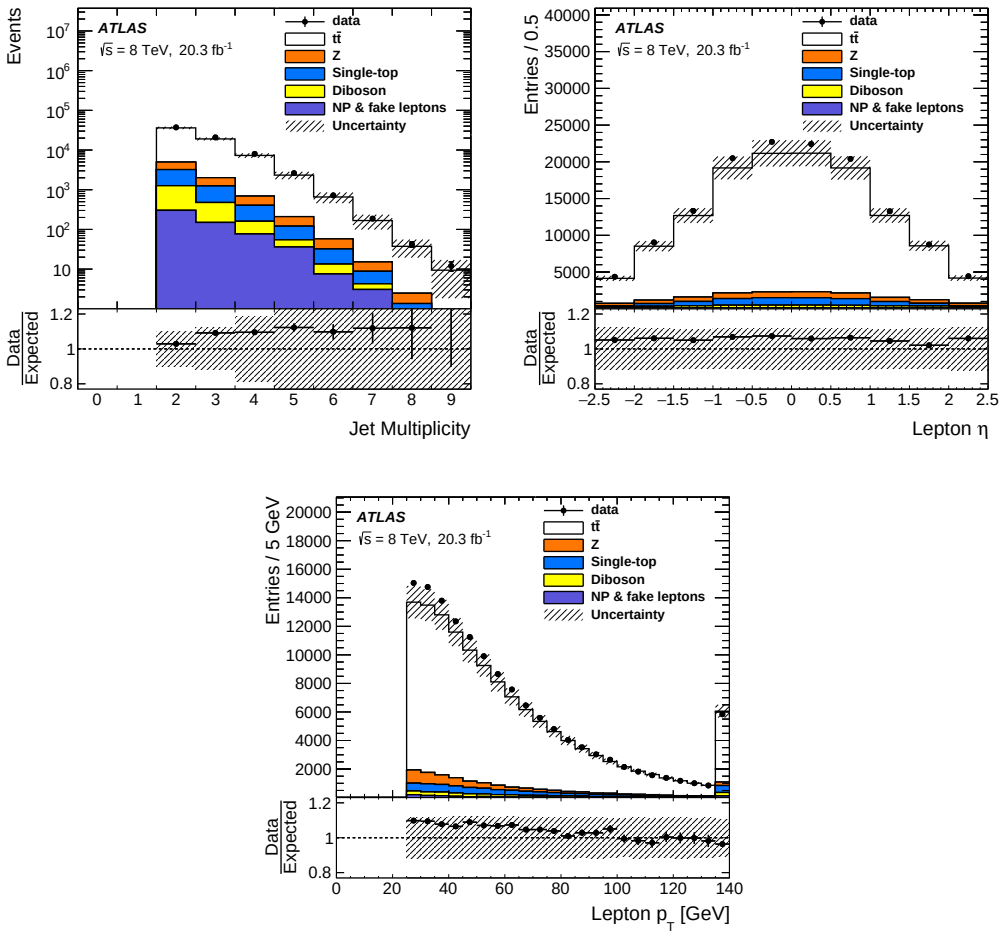


Figure 4.13 – *Comparison between data and expectations for the jet multiplicity, lepton η, and lepton p_{T} after the event selection used in the charge asymmetry analysis. The three channel are combined and the data/expectations ratio is shown in the bottom panel. The shaded area shows the total uncertainty.*

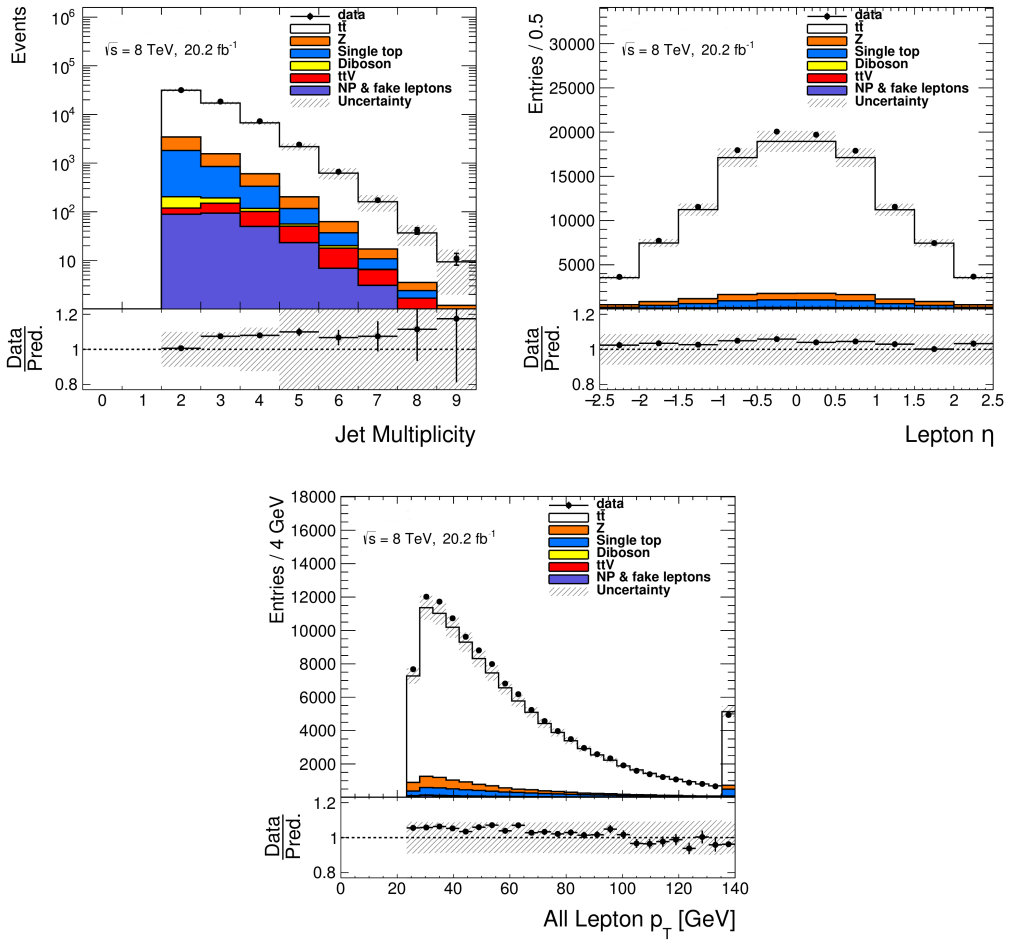


Figure 4.14 – *Comparison between data and expectations for the jet multiplicity, lepton η, and lepton p_{T} after event selection used in the top spin observables analysis. The three channel are combined and the data/expectations ratio is shown in the bottom panel. The shaded area shows the total uncertainty.*

Table 4.5 – *Observed numbers of data events compared to the expected signal and background contributions in the three decay channels and their combination for the top spin measurements. The uncertainty corresponds to the total uncertainty in the given process. Data-driven scale factors are applied to the Z+jets and the NP & fake leptons contributions. The $Z \to \tau\tau$ process in the $e\mu$ channel is estimated using MC simulation only.*

Channel	ee	$\mu\mu$	$e\mu$	$\ell\ell$
$t\bar{t}$	10200 ± 800	12100 ± 800	30300 ± 1800	52600 ± 2100
Single-top	510 ± 50	590 ± 50	1470 ± 100	2570 ± 120
Diboson	31 ± 5	40 ± 6	81 ± 8	152 ± 11
$t\bar{t}V$	39 ± 4	44 ± 4	100 ± 10	183 ± 11
$Z \to ee$	1200 ± 260	–	–	1200 ± 260
$Z \to \mu\mu$	–	1520 ± 300	–	1520 ± 300
$Z \to \tau\tau$	31 ± 15	58 ± 25	120 ± 9	210 ± 20
NP & fake leptons	$62 \pm {}^{+119}_{-29}$	$45 \pm {}^{+36}_{-24}$	$170 \pm {}^{+240}_{-220}$	$277 \pm {}^{+270}_{-230}$
Total Expected	12000 ± 900	14000 ± 800	32000 ± 1800	58000 ± 2000
Data	12785	14453	33482	60720

4.3 Fiducial Region and Particle Level

A fiducial region is defined in order to closely match the phase space region accessible with the ATLAS detector and the requirements made on the reconstructed objects. A fiducial measurement usually allows for MC generator dependencies to be reduced, since it avoids extrapolation to the full phase space. In the fiducial region, only objects defined at particle level are used. Particle-level objects are closer to the reconstructed objects than parton level object, thus, the corrections to be applied are better defined. In the following, the definition of the particle-level objects is presented.

The considered charged leptons (electrons and muons) are required not to originate from hadrons. Photons within $\Delta R = 0.1$ around the charged lepton are included in the four-momentum calculation. This procedures is known as *lepton dressing*. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated as the summed four-momenta of neutrinos from the W/Z boson decays, including those from intermediate τ decays. Jets are reconstructed by clustering all the stable particles in the event, excluding the neutrinos, the selected leptons and the photons used to dress the leptons. The clustering is performed using the anti-k_t algorithm with a radius parameter $R = 0.4$. The identification of jets originating from b-quarks is achieved using a procedure which consists in scaling the b-hadron's four-momentum by a small number (10^{-6}), and including it in the jet clustering. In this case, the MC generator-level b-hadrons are clustered into the particle-level jets. The four-momentum of the particle must be made small enough in order to avoid the alteration of the result of the jet clustering. This procedure is known as *ghost matching* [135].

The fiducial volume is defined by requiring at least two particle-level jets and at least two leptons in the event, with $p_T > 25$ GeV and $|\eta| < 2.5$ for both objects. Events where leptons and jets overlap within a ΔR cone of 0.4 are rejected. The particle-level jets are not required to be b-jets since this requirement is not shared between the three channels in the selection.

Using these objects, the reconstruction of top quarks (known as pseudotops [136]) can be performed. The assignment of the proper jet-lepton-neutrino permutation is chosen first by minimizing the difference between the masses computed from each lepton-neutrino combination and the W boson mass value used in the MC simulation. Then, the difference between the mass of each combination of the chosen lepton-neutrino pairs with a jet and the top quark mass value, used in the MC simulation, is minimized. If b-jets are present in the event, they are used for the jet-lepton-neutrino assignment. The correlation coefficient between top rapidity at the parton and particle levels is found to be 85%, while for the lepton η it is 95% as shown in Fig. 4.15.

The measurements of the asymmetry, polarization and correlations in the fiducial volume require the treatment of an additional background contribution. This background contribution corresponds to signal events that are outside of the fiducial region at generated level but have migrated into the detector acceptance due to resolution effects. The events out of the fiducial region correspond to a subset of the signal events. The estimation of this background, called nonfiducial background, is defined bin-by-bin as follows:

$$b_i^{nonfiducial} = (d_i - \sum_{j=other\,bckg} b_i^j) \cdot f_i, \qquad f_i = \frac{N_i^{non-fiducial}}{N_i^{signal}} \tag{4.6}$$

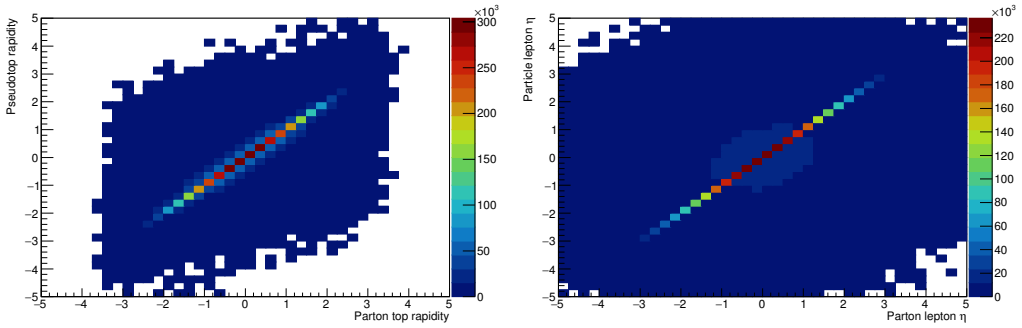


Figure 4.15 – *Correlation between particle and parton level definitions for top rapidity (left) and lepton η (right). The correlation factor are 0.85 and 0.95 for the top rapidity and lepton pseudorapidity, respectively.*

where $N_i^{nonfiducial}$ is the number of signal events in bin i without the generated level information within the fiducial region and N_i^{signal} is the number of signal events in bin i.

The nonfiducial background constitutes about 8% of the expected $t\bar{t}$ events after selection in the charge asymmetry measurements and about 6% in the top spin measurements. It was found to be independent of the charge asymmetry value of the simulated sample, as shown in Fig. 4.16. In the figure, the distributions of the $\Delta|y|$ and $\Delta|\eta|$ observables with the expected SM asymmetry and an asymmetry value more than six times larger (6% asymmetry) are presented at reconstructed level. The distribution corresponding to 6% asymmetry is obtained by reweighting the nominal $t\bar{t}$ signal sample at parton level to the considered value and applying the same weights at reconstructed level. The nonfiducial background was also found to be independent of spin correlation and polarization values, as shown Fig. 4.17.

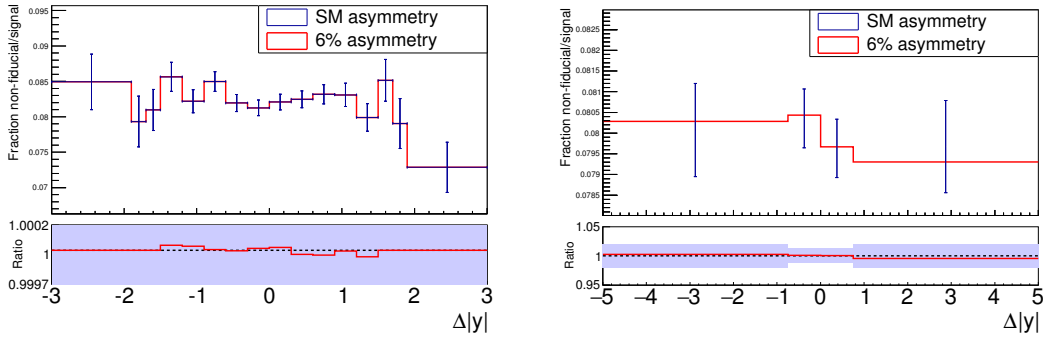


Figure 4.16 – *Fraction of nonfiducial background over all the signal events as a function of $\Delta|\eta|$ (left) and $\Delta|y|$ (right) for the SM and for a distribution that leads to a 6% asymmetry value.*

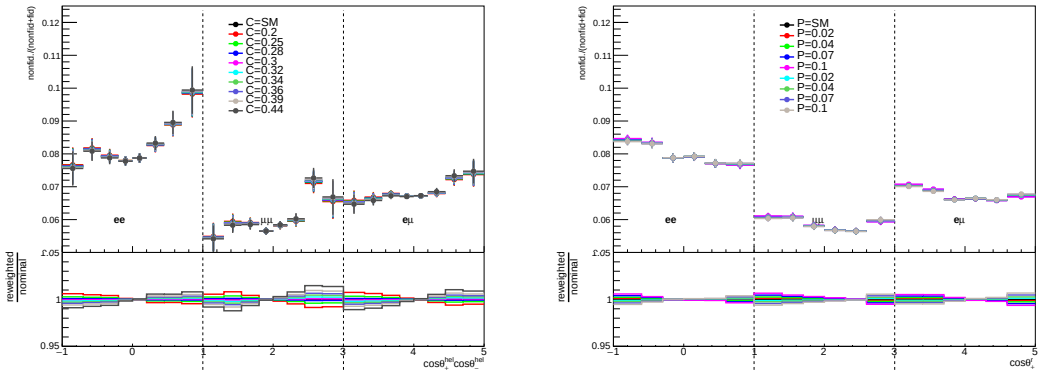


Figure 4.17 – *Fraction of nonfiducial background over all the signal events as a function of* $\cos\theta_+^k \cos\theta_-^k$ *(left) and* B_+^r *(right) for the different values of correlation (C) and polarizations (P).*

5

Kinematic Reconstruction

The top and antitop quark four-momenta need to be reconstructed from the final-state particles in order to compute any observable based on their kinematics (such as $\Delta|y|$ for the asymmetry estimation). The final state consists of two leptons of opposite charge, two neutrinos, and two b-jets. There are four decay vertices: two vertices corresponding to the process $t\bar{t} \to W^+bW^-\bar{b}$ and two vertices corresponding to $W^+ \to l^+\nu$ and $W^- \to l^-\bar{\nu}$. The reconstruction of the $t\bar{t}$ systems is performed by using momentum conservation at each of the decay vertices. The W bosons are reconstructed by summing the four-momenta of the leptons and neutrinos, while the top and antitop quarks are reconstructed by summing the four-momenta of the W bosons and the b-jets present in the event. Thus, assuming the process $pp \to t\bar{t} \to W^+bW^-\bar{b} \to l^+\nu bl^-\bar{\nu}\bar{b}$, the following kinematic constraints can be set on the particles momenta:

$$\begin{aligned} p_b + p_{W^+} &= p_t \\ p_{\bar{b}} + p_{W^-} &= p_{\bar{t}} \\ p_{l^+} + p_\nu &= p_{W^+} \\ p_{l^-} + p_{\bar{\nu}} &= p_{W^-} \end{aligned} \tag{5.1}$$

Considering each component of the four-momenta, there are 16 equations with forty unknowns in this system of equations. The momenta of the leptons and jets are measured, which reduces the number of unknowns to 28. The system of equation thus remains under-constrained. Other assumptions can be used to solve the system, such as the masses of the various particles, which are here assumed to be:

$$\begin{aligned} m_b &= m_{\bar{b}} &&= 4.5 \text{ GeV}, \\ m_{W^+} &= m_{W^-} &&= 80.4 \text{ GeV}, \\ m_t &= m_{\bar{t}} &&= 172.5 \text{ GeV}, \\ m_{l^+} &= m_{l^-} &&= 0.0 \text{ GeV}, \\ m_\nu &= m_{\bar{\nu}} &&= 0.0 \text{ GeV}. \end{aligned} \tag{5.2}$$

The system of equations has now 18 unknowns and 16 equations, implying that two more constraints are needed to solve the system. Different approaches can be used to establish these constraints, which depend on the reconstruction method. Once these constraints are established, there are still several ambiguities to solve. One of the ambiguities is related to the fact that at

least two jets are expected in the final state (coming from the b and $\bar{b}$). This leads to a combinatoric problem between the jets and the leptons, since it is not possible to know a priori which jet is from the top and which jet is from the antitop decay (the charge of the jets are not known). Another ambiguity is related with the multiple solutions that can be obtained when solving the system, since four-degree equations are expected with multiple possible roots, both real and complex. The way of treating these ambiguities depends on the reconstruction method. In this dissertation two methods are discussed: the Neutrino Weighting and the Kinematic Method (KIN method). These methods have been used in several analyses [137–140] but their internal parameters are optimized for the measurements performed in this work.

5.1 Neutrino Weighting

The Neutrino Weighting method makes use of assumptions of the η value of the neutrinos. This introduces two more constraints to the system of equations and makes it possible to solve it, since the numbers of equations and unknowns are the same. By making this assumption, it is possible to rearrange Eq. 5.1 and derive two polynomial equations (one for each neutrino) that describe the neutrinos four-momenta. The neutrino η assumptions are made based on MC simulations, which follow a Gaussian distribution as shown in Fig. 5.1.

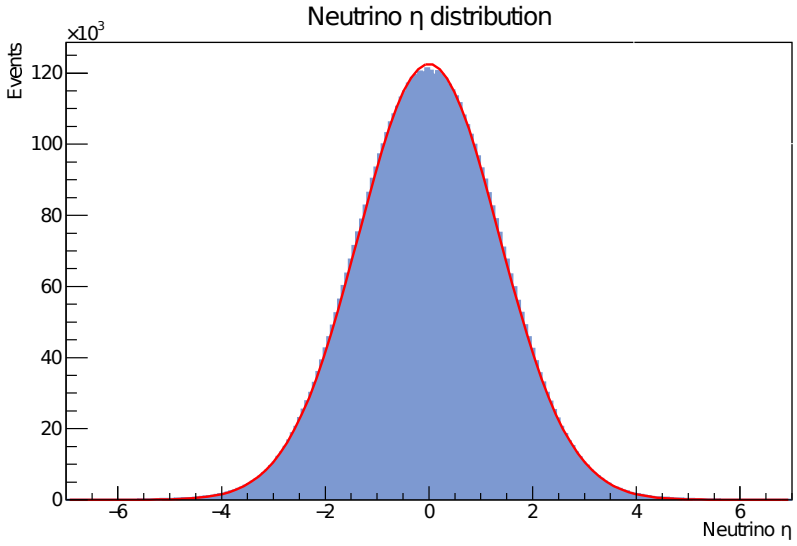


Figure 5.1 – *Neutrino η distribution in dileptonic $t\bar{t}$ events from MC simulation.*

In the Neutrino Weighting method, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is not directly used to solve the system of equations. Instead, it is used to evaluate the quality of the solutions. This evaluation is essential to discriminate solutions, since for a given number n of neutrino pairs assumptions (scanned points), it is possible to obtain more than n^2 solutions per event. The quality of each solution is

evaluated with a weight that estimates the compatibility of the measured $E_{\mathrm{T}}^{\mathrm{miss}}$ with the sum of the transverse momentum of the two assumed neutrinos. The weight is defined as:

$$\omega = \exp\left(-\frac{\left(E_{\mathrm{x}}^{\mathrm{miss}} - \not{E}_{\mathrm{x}}\right)^2}{2\sigma_{\mathrm{x}}^2}\right) \times \exp\left(-\frac{\left(E_{\mathrm{y}}^{\mathrm{miss}} - \not{E}_{\mathrm{y}}\right)^2}{2\sigma_{y}^2}\right), \tag{5.3}$$

where $E_{\mathrm{x}}^{\mathrm{miss}}$ ($E_{\mathrm{y}}^{\mathrm{miss}}$) is the x (y) component of missing transverse energy, σ_x (σ_y) is the resolution of the x (y) component of the missing transverse energy, and $\not{E}_x$ ($\not{E}_y$) is the sum of the momentum of the neutrinos along the x (y) direction. The resolution follows the equation $\sigma = 0.97 \cdot \sqrt{\Sigma E_{\mathrm{T}}}$ described in Sec. 4.1.6. The weight is maximum when the x (y) component of the momentum of the neutrinos is equal to the component $E_{\mathrm{x}}^{\mathrm{miss}}$ ($E_{\mathrm{y}}^{\mathrm{miss}}$) of the missing transverse energy. Two different approaches can be used to select the solutions. One can either select the solution with the highest weight or use all the solutions and take their weighted average. The neutrino-η pair that leads to the highest weight is used for the top and antitop quark four-momentum reconstruction in the presented analysis. The proper lepton-jet assignment is also evaluated through the weight; an incorrect assignment will generally lead to a smaller weight. All possible combinations between the jets and the leptons present in the event are tested.

Neutrino η Scanning

The scanning of the neutrino η is one of the most important steps for the reconstruction, since the quality of the solution with the highest weight is based on the assumptions made. The scanning is an arbitrary choice of the method. Two different criteria need to be taken into account in order to perform an optimal scanning. First, the choice of the sampling method, which can be either random (η points are randomly taken based on a probabilistic distribution) or methodical (η points are scanned following a specific pattern). Second, the shape of the distribution used to perform the sampling, which is very important since for instance the distribution of the neutrino η in boosted and nonboosted topologies are expected to be different. Several sampling procedures were tested in this study: a random sampling that follows Gaussian or flat distributions and a methodical sampling of variable η steps.

When a random sampling that follows a Gaussian distribution for both neutrinos is used (as in Fig. 5.1), the solutions obtained are biased towards the central η values, since the η values are more often scanned in the central region. This not optimal for the boosted $t\bar{t}$ topologies, where the neutrino η distribution is shifted in η, as shown in Fig. 5.2. The boosted topologies are important for asymmetry measurements, since for high boost, the asymmetry value is expected to change for different BSM models. When a random sampling using a flat distribution[1] is performed, the bias problem affecting boosted topologies is not present. However, the number of sampled points needs to be increased in order to reach a performance similar to the previous scanning procedure, which increases the computation time significantly.

The combination of these two approaches was also tested: one neutrino η sampled using a Gaussian distribution centered at zero and the other neutrino η using a flat distribution. However,

[1] In this case all η values have the same probability of being scanned

the performance was not optimal in comparison with a methodical η sampling of variable size (called *grid* sampling).

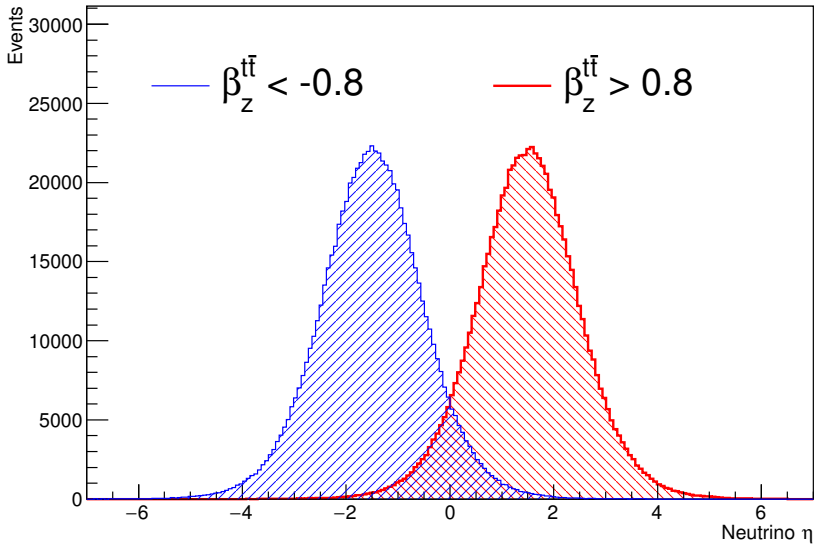


Figure 5.2 – *Neutrino η distributions from MC simulation for boosted topologies.*

In the grid sampling, the neutrino-η values are scanned in the range $[-5, 5]$ in steps of variable size. In the η ranges $[-5, -2]$ and $[2, 5]$, the steps are of the size of 0.05, and in the range $[-2, 2]$ the steps are reduced to 0.025. The finer steps in the central region are due to the high probability of finding neutrinos in this region. Figure 5.3 shows the distribution of the weights for the scanned neutrinos η points in one event of the MC simulation. The highest weight is presented in yellow for the η of the neutrinos and the true neutrino η is represented as the red mark.

Jet Smearing

In 40% of the cases, when the Neutrino Weighting method solves the $t\bar{t}$ system, it is not possible to obtain non-complex solutions. This is due to the finite number of η points that are tested and due to the resolution of the reconstructed objects. In order to reduce this fraction, the jets are smeared according to their resolution taken using MC simulations. The resolution functions are estimated by fitting the distribution $(p_{\mathrm{T}}^{\mathrm{truth}} - p_{\mathrm{T}}^{\mathrm{reco}})/p_{\mathrm{T}}^{\mathrm{truth}}$, in p_{T} bins,[2] using the convolution of two Gaussian functions. The typical resolution across the different p_{T} bins is between 10% and 20%. Figure 5.4 shows the resolution function used for the p_{T} ranges of $[20, 34]$, $[75, 87]$, and $[184, 500]$ GeV. The smearing is performed by randomly sampling the resolution function. In principle, the same procedure could be performed for the leptons. However, the lepton resolution is two

[2] The p_{T} bins used for the resolutions are $[20, 34, 42, 50, 57, , 65, 75, 87, 103, 130, 156, 184, 500]$ GeV

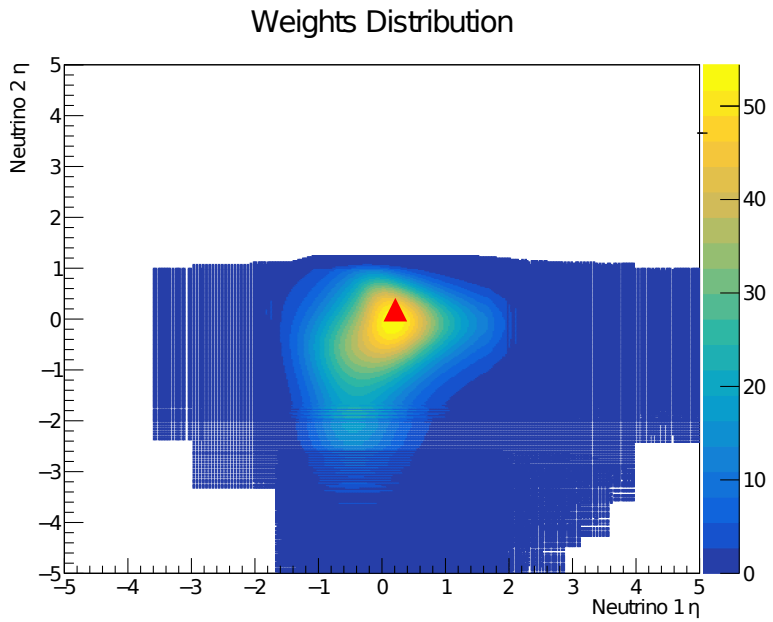


Figure 5.3 – *Distribution of the weights for the scanned neutrinos η pairs using the Grid sampling. The highest weight is presented in bright yellow. The true neutrinos η pair is shown with the red mark.*

orders of magnitude better than the jet resolution and therefore the smearing on the leptons is not considered. Appendix B shows all the resolution functions for the different energy bins on the different objects. By smearing the jets, the number of unsolvable events drops from 40% to around 10%. The jets are smeared 50 times for each event. The computation time increases significantly when more smearings are performed. The chosen configuration gives a good balance between computation time and reconstruction efficiency. The optimization will be further described in Sec. 5.3.

Comparison between data and MC

In Fig. 5.5 the p_{T}, invariant mass, and boost along the z-axis of the $t\bar{t}$ system after applying the Neutrino Weighting method are shown for the three combined decay channels. The purity of the sample is increased around 5% after the reconstruction, since it is more likely to find solutions for signal events than for background events. The expectation/data ratio remains unaffected (it is of 0.96). More distributions can be found in Appendix A.

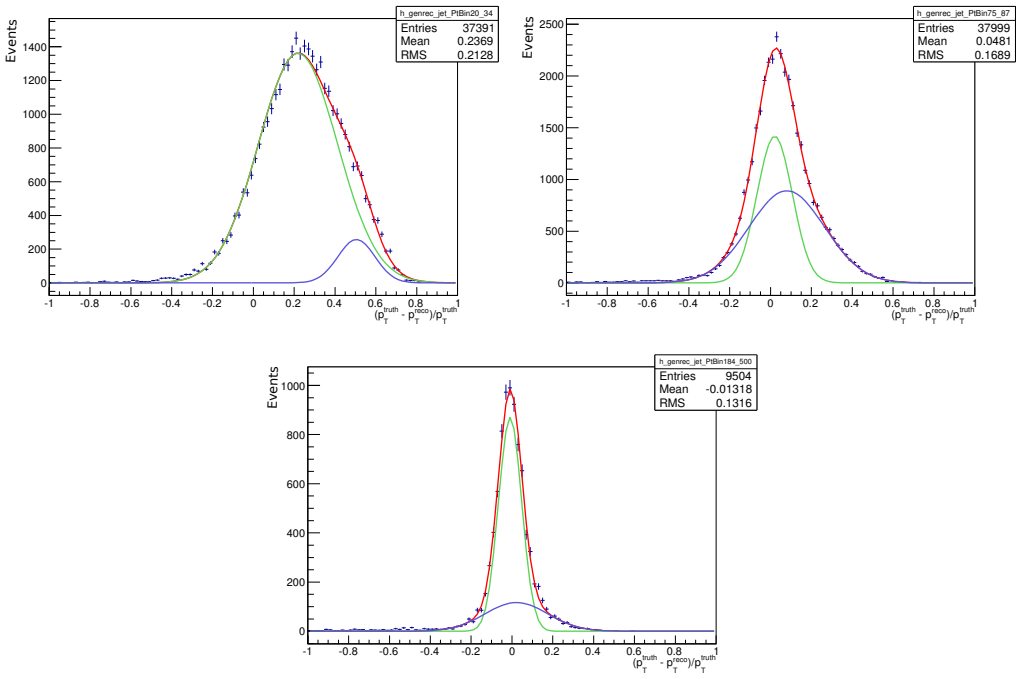


Figure 5.4 – *Jet resolution function for* p_{T} *in the ranges of* $[20, 34]$ *(left),* $[75, 87]$ *(right), and* $[184, 500]$ *GeV (bottom). The distributions are fit using a double Gaussian functions.*

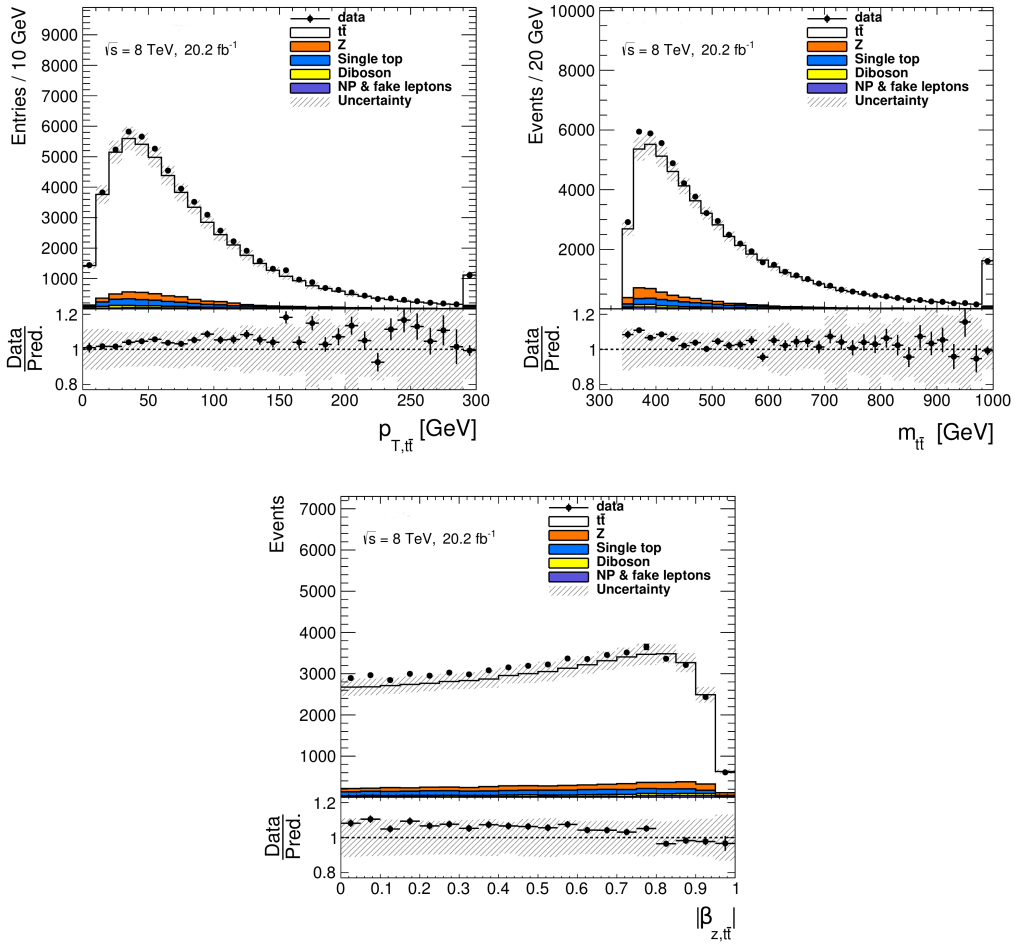


Figure 5.5 – *Comparison between data and expectations after applying the Neutrino Weighting reconstruction for the p_T (left), invariant mass mass (right), and boost along the z-axis (bottom) of the $t\bar{t}$ system. The total uncertainty corresponds to the quadratic sum of the statistical, normalization and systematic uncertainties.*

5.2 KIN Method

The KIN method starts by adding two constraints to the system of equations of sixteen equations and eighteen unknowns. These constraints relate the $E_{\mathrm{T}}^{\mathrm{miss}}$ of the event with the neutrinos, as:

$$\begin{aligned} p_{\nu_x} + p_{\bar{\nu}_x} &= E_x^{miss}, \\ p_{\nu_y} + p_{\bar{\nu}_y} &= E_y^{miss}. \end{aligned} \tag{5.4}$$

Since two more equations are added, the system can be solved numerically. The Newton-Rhapson method is used to solve the system. If no real solution is found, the jet energy and $E_{\mathrm{T}}^{\mathrm{miss}}$ are smeared within their resolutions and the system is solved again until real solutions are found. In contrast to the Neutrino Weighting, the KIN method uses the measured $E_{\mathrm{T}}^{\mathrm{miss}}$ during the computation. By rearranging the system, it is possible to obtain a four degree polynomial for each neutrino. When at least two real solutions are found, the method stops due to computation time efficiency, since the probability of finding four real solutions is around 1%. The smearing of the objects (sampling) is performed up to N_{smear} times. The chosen solution is either the solution found using the nominal jet energies and measured $E_{\mathrm{T}}^{\mathrm{miss}}$, if available, or the first solution found during the sampling. The kinematic reconstruction fails for a given event if no solution is found in any of the N_{smear} sampled points.

If there are more than two reconstructed jets in a given event, the two jets with the highest b-tagging weights (as determined by the MV1 b-tagging algorithm) are used. This improves the probability of choosing the correct jets, compared to just choosing the two jets with the highest p_T, from about 54% to about 69% in the nominal $t\bar{t}$ sample.

For each sampling point, several solutions can be obtained. The KIN method chooses the solution that leads to the lowest reconstructed mass of the $t\bar{t}$ system. The reason for this is that the $t\bar{t}$ cross section is a decreasing function of the partonic center-of-mass energy $\sqrt{\hat{s}} \simeq m_{t\bar{t}}$, so events with smaller $m_{t\bar{t}}$ are more likely. The twofold ambiguity in the lepton and b-jet assignment is solved by taking the permutation that has been reconstructed more often. In other words, the permutation that maximizes $N_{\mathrm{smear}}^{\mathrm{reco}}/N_{\mathrm{smear}}$, where $N_{\mathrm{smear}}^{\mathrm{reco}}$ corresponds to the number of successful reconstructions and N_{smear} the number of trials (sampling).

Jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ Smearing

If the KIN method fails to find real solutions, the jet energy and the $E_{\mathrm{T}}^{\mathrm{miss}}$ are smeared within their resolutions. The jet resolution functions are derived in the same way as for the Neutrino Weighting. The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution function are derived, in a similar way as the jet resolutions, in bins of the total transverse energy in the event ($\sum E_{\mathrm{T}}$). The distribution $(\sum_{\nu} p_{\mathrm{T}}^{x/y}) - E_{\mathrm{T}}^{\mathrm{miss,x/y}}$, where $(\sum_{\nu} p_{\mathrm{T}}^{x/y})$ is the x/y component of the total transverse momentum of the neutrinos, is fit using a Gaussian. The function obtained is used by the KIN method to perform the random sampling during the smearing procedure. In Fig. 5.6 the resolution functions for $\sum E_{\mathrm{T}}$ bins in the ranges [0,375], [590,640], and [1170,1600] GeV are shown. The typical resolution across the different bins is between 20 GeV and 30 GeV. The parameter N_{smear} is set to 500. It was optimized based on the computation time and reconstruction efficiency (described in Sec 5.3).

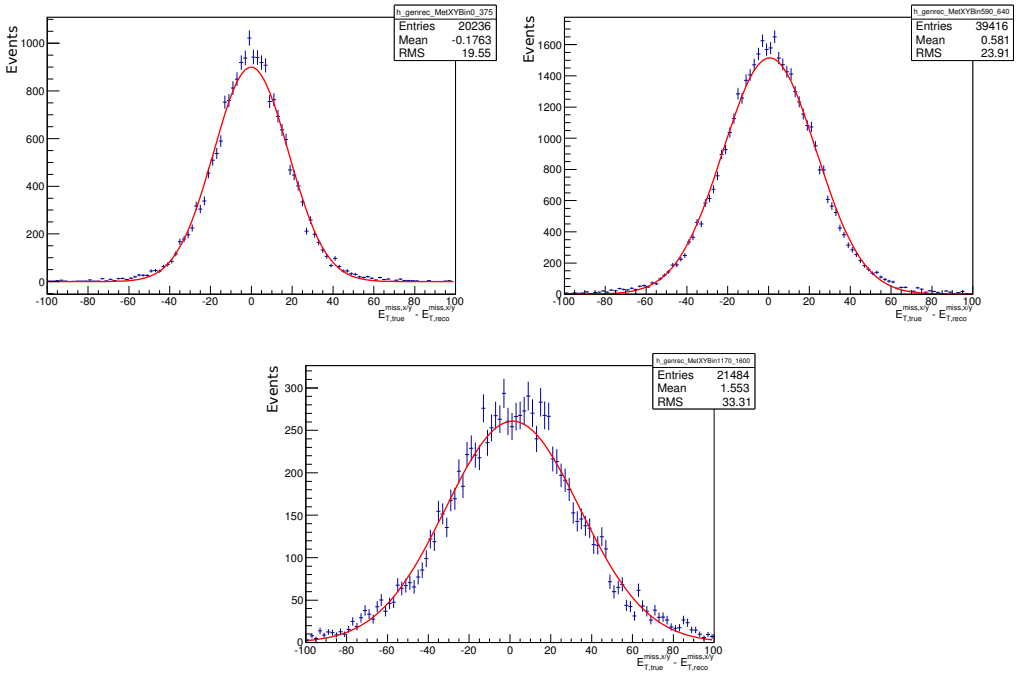


Figure 5.6 – E_T^{miss} *resolution function for energies in the ranges of* $[0,375]$ *(left),* $[590,640]$ *(right), and* $[184,500]$ *GeV(bottom). The distributions are fit using a Gaussian function.*

Comparison between data and MC

In Fig. 5.7 the p_T, invariant mass, and boost along the z-axis of the $t\bar{t}$ system after applying the KIN method is shown for the three channels combined. As in case of the Neutrino Weighting method, the purity of the sample is increased 4% after the reconstruction, since it is more likely to find solutions for signal events than for background events. The MC/data ratio remains unaffected (0.96). The distributions are compatible with the ones obtained using Neutrino Weighting. More distributions can be found in Appendix A.

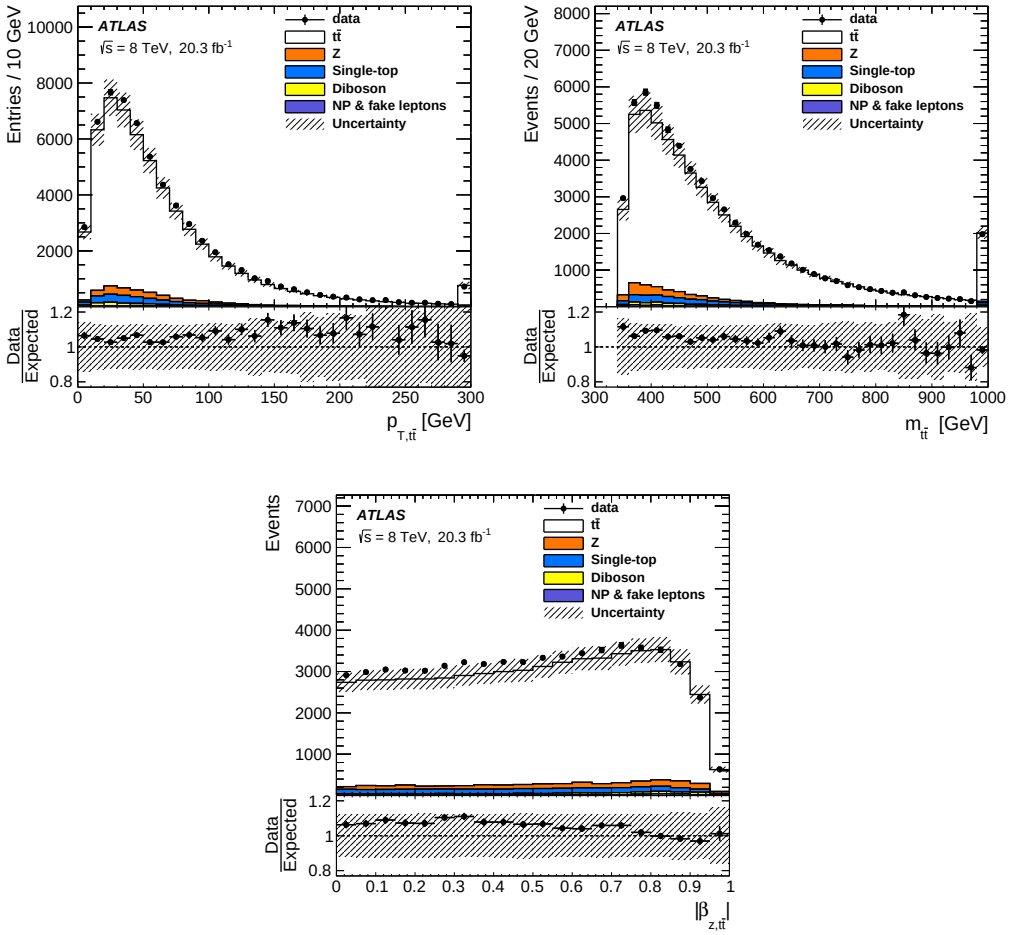


Figure 5.7 – *Comparison between data and expectations after the reconstruction performed using the KIN method for the p_T (left), mass (right), and boost along the z-axis (bottom) of the $t\bar{t}$ system. The uncertainty corresponds to the quadratic sum of the statistical, normalization and systematic uncertainties.*

5.3 Performance of the methods

5.3.1 Efficiencies and Resolutions

The performance of the methods previously described is evaluated using several estimators. One essential estimator is the reconstruction efficiency, defined as the fraction of events entering the reconstruction procedure which have at least one solution. An efficiency of 100% means that at least one solution is found for all selected events. The resolution of the kinematic variables of the top quarks and the $t\bar{t}$ system are also studied. The resolutions are defined as the standard deviation of the difference between the reconstructed and generated variables associated with the kinematics of the top quarks and $t\bar{t}$ system. The p_{T}, η, y, ϕ of the top quarks are studied, while the p_{T}, η, y, ϕ, mass, and boost along the z-axis are studied for the $t\bar{t}$ system. The study of the p_{T}, mass and boost along the z-axis are of particular interest for the differential measurements of the charge asymmetry, since a better resolution on these variables gives a more precise differential measurement. In addition, the $\Delta|y|$ resolution, the correct sign of the $\Delta|y|$ observable, as well as the computing time per event are considered. Figure 5.8 shows the distributions of the difference between the reconstructed and generated p_{T} of the top and mass of the $t\bar{t}$ system using both reconstruction methods. The standard deviation of these distributions are used to evaluate the performance of the methods. In these examples, the KIN method is offering a slightly smaller standard deviation. The mean of the distributions, which quantify the bias of the method is slighly better for Neutrino Weighting, however, the difference present in both reconstruction methods is not significant and it will be corrected during the unfolding procedure described in Chapter 6. In Tab. 5.1 the performance estimators are shown for both reconstruction methods using the full $t\bar{t}$ sample. The reconstruction efficiency is found to be 89% for Neutrino Weighting and 91% for KIN. Despite the fact that the average computation time is better for Neutrino Weighting, KIN offers a slightly better resolution in all the kinematic variables. Also, by using KIN, the correct sign of the $\Delta|y|$ observable is better than by using Neutrino Weighting by 2 %. The migration between the generated and reconstructed $\Delta|y|$ observable due to resolution and acceptance effects is shown for both methods in Fig. 5.9. The migration matrix is slightly more diagonal for KIN than for Neutrino Weighting, especially in the central regions.

The charge asymmetry is expected to change as a function of the boost of the $t\bar{t}$ system as stated in Chapter 1. The performance of the reconstruction methods as a function of the boost along the z-axis of the $t\bar{t}$ system is evaluated. The performance is evaluated in two boost regimes: high boost ($\beta_z^{t\bar{t}} > 0.6$) and low boost ($\beta_z^{t\bar{t}} < 0.6$). The performance estimators are presented in Tab. 5.2 for the low and high boost regimes. In general, the KIN method offers better resolutions across the different variables in both regimes. Given this fact, the KIN method was selected as the reconstruction method for the charge asymmetry measurements.

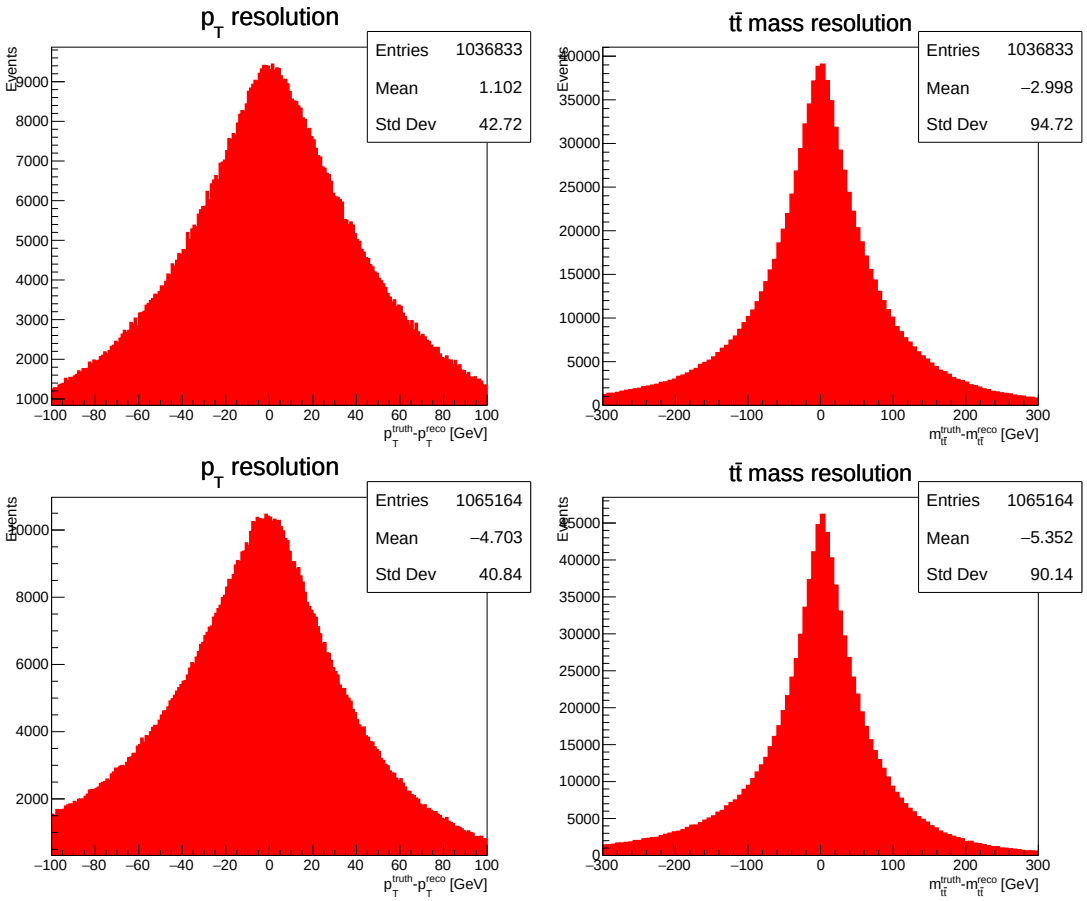


Figure 5.8 – *Distribution of the difference between generated and reconstructed p_T (left) and mass (right) of the $t\bar{t}$ system for Neutrino Weighting (top) and KIN (bottom) method. The standard deviation of these distributions is used as a performance estimator.*

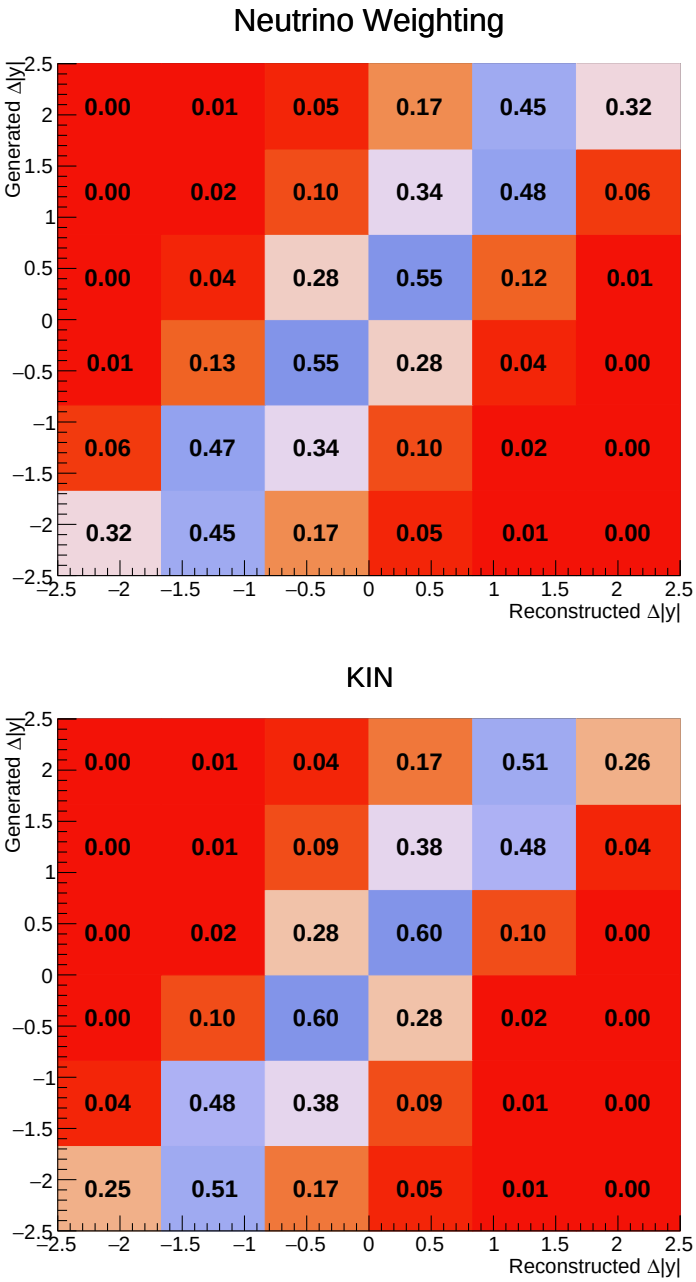


Figure 5.9 – *Response matrices for the Δ|y| observable, obtained using Neutrino Weighting (top) and KIN Method (bottom). The matrices are normalized by row in order to clearly see the migration between the different bins.*

Table 5.1 – *The performances and resolutions for the nominal $t\bar{t}$ sample. The statistical uncertainties on the numbers are negligible. The column named NW corresponds to the Neutrino Weighting performance.*

	NW	KIN
Reconstruction efficiency [%]	89 %	91%
$\Delta\|y\|$ resolution	0.63	0.56
Correct $\Delta\|y\|$ sign	74 %	76 %
Average CPU time per event [s]	3.2	4.1
Resolutions - top & antitop		
p_T [GeV]	42.7	40.8
y	0.49	0.44
ϕ	1.37	1.19
η	0.95	0.84
Resolutions - $t\bar{t}$ system		
p_T [GeV]	37.4	30.5
η	1.45	1.34
y	0.33	0.31
$m_{t\bar{t}}$ [GeV]	94.7	90.1
$\beta_z^{t\bar{t}}$	0.24	0.22

5.3.2 Additional uncertainty on the KIN method

In the charge asymmetry analysis a systematic uncertainty affecting the KIN method was found. This uncertainty is due to the randomness in the smearing procedure. If the smearing starts from a different point, it could lead to a different solution due to the stopping criteria (method stops when the first solution is found). The uncertainty from this effect is computed by performing pseudoexperiments on MC events. For each event, the $t\bar{t}$ system is reconstructed multiple times varying the starting point of the smearing procedure. Then, for each variation the measurement is repeated and the standard deviation of the asymmetries obtained is taken as the uncertainty. This represents one of the major systematic uncertainties for the charge asymmetry measurements, but it is still only half of the statistical uncertainty for most of them. This uncertainty will be explained in more details in Chapter 8. Due to this extra uncertainty and the fact that in the top spin measurements the size of the sample after reconstruction is smaller than in the charge asymmetry analysis, the reconstruction method was changed to Neutrino Weighting for the top spin measurement.

Table 5.2 – *The performances and resolutions for the nominal $t\bar{t}$ sample in the low and high boost regimes using Neutrino Weighting (NW) and KIN Method for the kinematic reconstruction. The statistical uncertainties on the numbers are negligible.*

	Low boost		High boost	
	NW	KIN	NW	KIN
$\Delta\|y\|$ resolution	0.56	0.50	0.70	0.62
Correct $\Delta\|y\|$ sign	71 %	73 %	78 %	79 %
Resolutions - top & antitop				
p_T [GeV]	42.7	40.7	42.8	41.1
η	0.97	0.86	0.91	0.83
ϕ	1.35	1.16	1.40	1.24
y	0.48	0.43	0.51	0.47
Resolutions - $t\bar{t}$ system				
p_T [GeV]	37.5	30.2	37.4	30.8
η	1.55	1.48	1.28	1.10
y	0.31	0.30	0.35	0.34
$m_{t\bar{t}}$ [GeV]	96.8	91.6	91.6	88.0
$\beta_z^{t\bar{t}}$	0.27	0.25	0.19	0.18

6 Unfolding

In order to allow for a direct comparison between the experimental results and the theoretical predictions, the parton-level and particle-level asymmetries, polarizations, and spin correlation distributions need to be estimated. After selection and reconstruction, the measurements are affected by the resolution and acceptance of the ATLAS detector. These effects can shift or dilute the measured value of a given observable compared to its true value. Figure 6.1 shows an example of the reconstructed charge asymmetry values as a function of the true value of the charge asymmetry in MC simulations. A linear fit is performed in order to evaluate the degradation of the asymmetry value due to detector effects (blue line). The red dashed line corresponds to the $y = x$ curve. The reconstructed asymmetry values are shifted by around 30% with respect to the true values. The nominal $t\bar{t}$ sample is used to map the distortions, efficiency, and migrations (caused by acceptance and resolution effects) in the distribution of the observable. Using this information, combined with the distribution observed in data, it is possible to correct for detector and acceptance effects and bring the measurement back to parton and particle level. This procedure is known as unfolding. It can be applied to a single distribution (one dimensional unfolding) or to multiple distributions simultaneously (multi-dimensional unfolding). In this chapter, the unfolding problem in particle physics is described, followed by the description of the unfolding procedure used in the analyses.

6.1 The unfolding problem

When a measurement x of an observable y is performed, it is distorted by the limited resolution and acceptance of the detector. This implies that each observation i is characterized by a measured value x_i, corresponding to a true value y_i, and the distributions of x and y are related by the relation:

$$f_{\text{measured}}(x) = \int dy M(x,y) f_{\text{true}}(y), \tag{6.1}$$

where $f_{\text{true}}(y)$ is the distribution of the variable y and $M(x,y)$ is the response function, describing the probability for the true value y to be measured as x. The unfolding consists of the determination of the function f_{true} by de-convoluting Eq. 6.1. The problem of de-convoluting resolutions

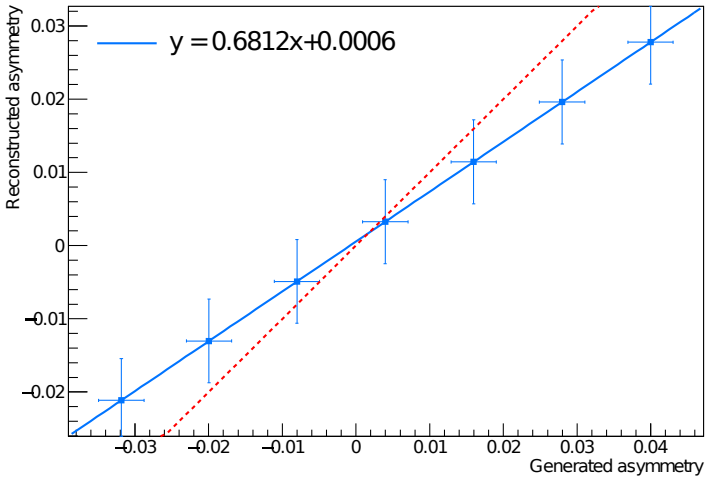


Figure 6.1 – *Reconstructed asymmetry values as a function of the true asymmetry generated by MC. The red dashed line correspond to the curve $y = x$ while the blue line represents the linear fit performed on the points.*

and detector effects from the distribution of true observable has been studied for a long time. The first time that this was discussed in the context of particle physics was in 1984 [141].

The function $f_{\text{true}}(y)$ is usually represented in particle physics by histograms of N_t bins defined as $\boldsymbol{T} = \{t_1, t_2, ..t_{N_t}\}$. In the same way, the function $f_{\text{measured}}(x)$ is represented by the reconstructed histogram $\boldsymbol{R} = \{r_1, r_2, ...r_{N_r}\}$ with N_r bins, and the response function M becomes the $N_t \times N_r$ response matrix $\mathcal{M}$. The t_i and r_i values represent the expectation values of the number of entries in bin i of the true and reconstructed histograms, respectively. The Eq. 6.1 becomes then:

$$\boldsymbol{R} = \mathcal{M}\boldsymbol{T}. \tag{6.2}$$

The goal of the unfolding is to construct estimators $\hat{\boldsymbol{T}}$ for the N_t parameters $\boldsymbol{T} = \{t_1, ..., t_{N_t}\}$ given a distribution $\boldsymbol{R}$ of the measured values. Figure 6.2 shows an example of a migration matrix, reconstructed, and generated distributions. The reconstructed distribution differs from the generated distribution due to the detector and reconstruction effects. The response matrix, $\mathcal{M}$, is typically derived from simulated samples.

The first approach that can be taken to get the estimators is to invert $\mathcal{M}$, such as $\boldsymbol{T} = \mathcal{M}^{-1}\boldsymbol{R}$. Considering a possible background contribution $\boldsymbol{B}$ to the measured data $\boldsymbol{D}$[1], the estimator $\hat{\boldsymbol{T}}$ can be written as:

$$\hat{\boldsymbol{T}} = \mathcal{M}^{-1}(\boldsymbol{D} - \boldsymbol{B}). \tag{6.3}$$

Although the estimator obtained from matrix inversion is unbiased since the expected value of the estimator $\hat{\boldsymbol{T}}$ is $\boldsymbol{T}$, when the off-diagonal elements of the response matrix are too large, i.e., if the bin size is too small compared to the measurement resolution, the estimators have very large variances and strong negative correlations between neighbouring bins [142]. This effect can be seen in Fig. 6.3, where the the estimator shows very large fluctuations when the matrix inversion

[1] The vector $\boldsymbol{D}$ is considered to be Poisson distributed and independent measurements

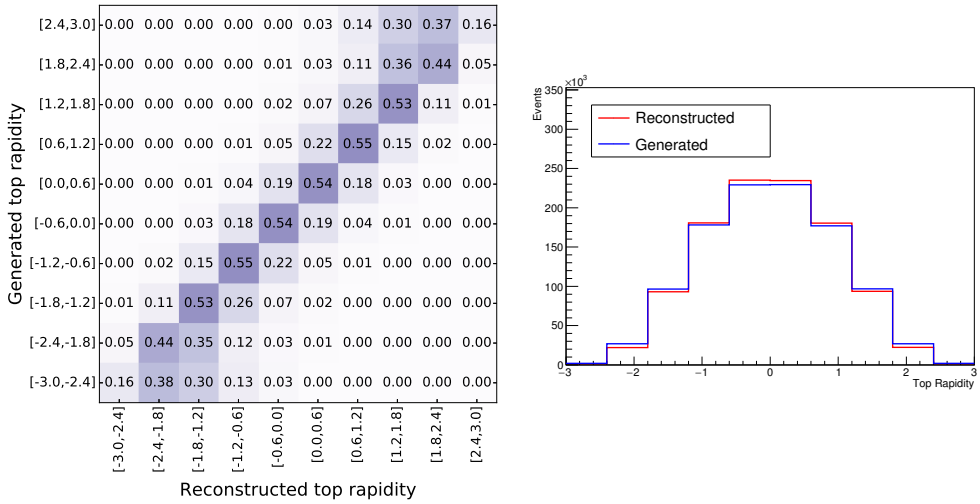


Figure 6.2 – *Migration matrix of the top rapidity (left) normalized by row and comparison between the reconstructed and generated distributions (right). The difference between the reconstructed and generated distributions are coming from the effects introduced by the detector and the reconstruction of the top quark.*

is performed on the observed histogram (signal + background). The reason for this is that fluctuations in data can be interpreted (during the matrix inversion) as a fine structure present in the true distribution. Unfolding methods implement procedures that allow for more precise estimators by reducing statistical fluctuations present in the measurements, at the expense of introducing hypotheses that could lead to a bias. These procedures are known as *regularization*. The regularization commonly uses the information of data and the predicted expectation values to evaluate the smoothness of the distributions.

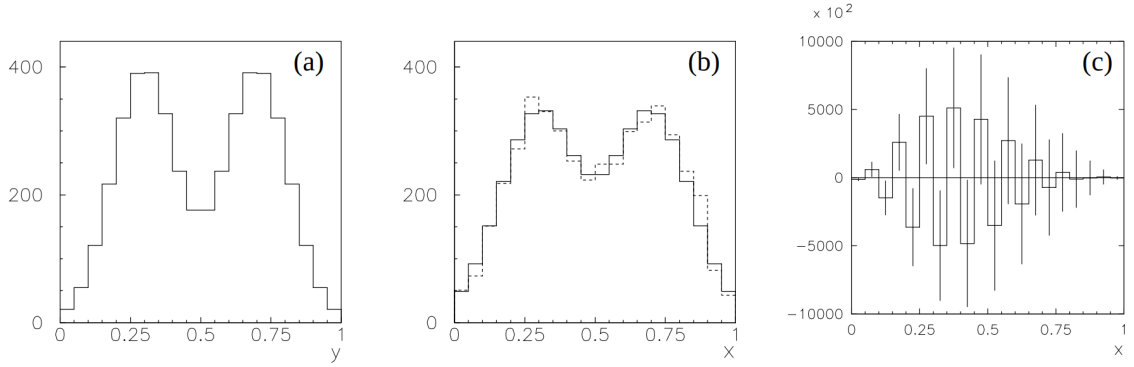


Figure 6.3 – *Example of unfolding using matrix inversion. Figure (a) shows the true histogram. In Fig. (b) the dashed line represents the observed histogram (signal+background) and the solid line represents only the signal histogram. Figure (c) shows the estimators of the true distribution based on Eq. 6.3 [142].*

Several unfolding techniques have been proposed and used in particle physics. The simplest method overcomes the invertibility problem by constructing the estimator for $\boldsymbol{T}$ based on multi-

plicative correction factors obtained from MC simulations. This method is known as the *bin-by-bin* correction. The estimator for bin i is defined as:

$$\hat{t_i} = C_i r_i \quad \text{with} \quad \mathrm{C_i} = \frac{\mathrm{t_i^{MC}}}{\mathrm{r_i^{MC}}}, \tag{6.4}$$

where C_i is the bin-by-bin correction factor derived from simulations as the ratio of the number of true events over the number of measured events. This method neglects the effect of off-diagonal migrations, making it strongly biased towards the simulated model used to derive the correction factors.

6.2 Fully Bayesian Unfolding

The problem of the unfolding can be studied using a Bayesian approach. This is the case of the Fully Bayesian Unfolding (FBU), which consists of the application of Bayes theorem to the problem of unfolding [143]. This application can be stated in the following terms: given an observed spectrum $\boldsymbol{D} \in \mathbb{N}^{N_r}$ and a migration matrix $\mathcal{M} \in \mathbb{R}^{N_r} \times \mathbb{R}^{N_t}$, the posterior probability of the true spectrum $\boldsymbol{T} \in \mathbb{R}^{N_t}$ follows the probability density

$$p(\boldsymbol{T}|\boldsymbol{D}, \mathcal{M}) \propto \mathcal{L}(\boldsymbol{D}|\boldsymbol{T}, \mathcal{M}) \cdot \pi(\boldsymbol{T}), \tag{6.5}$$

where $\mathcal{L}(\boldsymbol{D}|\boldsymbol{T}, \mathcal{M})$ is the conditional likelihood for the data $\boldsymbol{D}$ assuming the true $\boldsymbol{T}$ and the transfer matrix $\mathcal{M}$, and π is the prior probability density for the truth $\boldsymbol{T}$.

Likelihood

Under the assumption that the data are poissonian counts, the likelihood $\mathcal{L}(\boldsymbol{D}|\boldsymbol{T}, \mathcal{M})$ can be computed from the following two pieces of information:

- the transfer matrix $\mathcal{M}$, whose element $\mathcal{M}_{tr}$ is the probability $P(t, r)$ of an event to be produced in the true bin t and to be observed in the reconstructed bin r;
- the efficiency ϵ_t for an event produced in the true bin t and that is reconstructed in any bin r.

The likelihood is then defined by comparing the observed spectrum $\boldsymbol{D}$ with the expected one, considering the background as follows:

$$\mathcal{L}(\boldsymbol{D}|\boldsymbol{T}, \mathcal{M}) = \prod_i^{N_r} Poisson(d_i, r(\boldsymbol{T}, \mathcal{M}) + b_i)), \tag{6.6}$$

where b_i represents the background contribution in bin i and $r(\boldsymbol{T}, \mathcal{M})$ is the predicted reconstructed distribution. Each bin of the predicted reconstructed distribution is given by:

$$r_i(\boldsymbol{T}, \mathcal{M}) = \sum_{j=0}^{N_r} m_{ij} \cdot t_j, \tag{6.7}$$

where $m_{ij} = \epsilon_{t_j} \cdot p(r_i|t_j)$ defines the elements of $\mathcal{M}$. While $\mathcal{M}$ can be estimated from the simulated sample of signal events, the prior probability density $\pi(\boldsymbol{T})$ is to be chosen according to what

is known about $\boldsymbol{T}$ before the measurement is performed. The choice of a prior within the FBU context is similar to the choice of a regularization in other well known unfolding techniques.

Prior Choice

The simplest prior corresponds to the assumption that all points are equally probable:

$$\pi(\boldsymbol{T}) \propto \begin{cases} 1 & \text{if } T_t \in \left[T^{min}, T^{max}\right], \forall\, t \in [1, N_t] \\ 0 & \text{otherwise} \end{cases} \tag{6.8}$$

where $\left[T^{min}, T^{max}\right]$ indicates, for each dimension t, the finite subset of $\boldsymbol{T} \in \mathbb{R}$ over which the numerical integration is being performed. A more general expression can be written as follows

$$\pi(\boldsymbol{T}) \propto \begin{cases} e^{\alpha S(\boldsymbol{T})} & \text{if } T_t \in \left[T^{min}, T^{max}\right], \forall\, t \in [1, N_t] \\ 0 & \text{otherwise} \end{cases} \tag{6.9}$$

where α is an arbitrary regularization parameter, and $S(\boldsymbol{T})$ is an arbitrary function. In some cases, $S(\boldsymbol{T})$ is used as the difference between the true and expected values in quadrature. This particular choice is known as *Tikhonov* regularization functions [144]. The choice of α determines the impact of the prior on $p(\boldsymbol{T}|\boldsymbol{D})$, while $S(\boldsymbol{T})$ determines what additional information is being used to constrain the parameter space, thus reducing the variance of the $\boldsymbol{T}$ parameters.

Sampling and observable extraction

After choosing a prior, and therefore making some well-defined assumptions about the spectrum being unfolded, it is possible to determine the posterior probability density function $p(\boldsymbol{T}|\boldsymbol{D})$. This is done by generating a sufficient number of randomly-distributed points in the N_t-dimensional space, and evaluating for each of them $\mathcal{L}(\boldsymbol{D}|\boldsymbol{T}, \mathcal{M})$ and $\pi(\boldsymbol{T})$. The sampling is performed using the Metropolis-Hastings algorithm [145], which is a Markov Chain Monte Carlo (MCMC) method that samples points in the phase space based on the probability obtained at each sampled point. One million points are sampled in the phase space. Combining the scanned set of points with the probability given by $\mathcal{L}(\boldsymbol{D}|\boldsymbol{T}, \mathcal{M}) \cdot \pi$, it is possible to determine not only the posterior probability density distribution for each bin of the spectrum, but also the posterior probability density distribution for any quantity that is computed from the spectrum, such as the charge asymmetries, the spin correlations, or the top quark polarizations. The mean and RMS of the posterior probability density, which is expected to be a well behaved Gaussian distribution, are taken as central value and statistical uncertainty respectively. Figure 6.4 shows an example of the posterior probability density for the inclusive $A_C^{t\bar{t}}$ in the full phase space using the sum of the signal and background samples from MC (pseudodata). The measured value of the asymmetry is found to be 0.0043 ± 0.0156 which is compatible with the generated value of 0.0041 ± 0.0005.

Marginalization

All measurements are affected by systematic uncertainties. The effects needs to be taken into account in the posterior probability density computation. The effect of the individual systematic

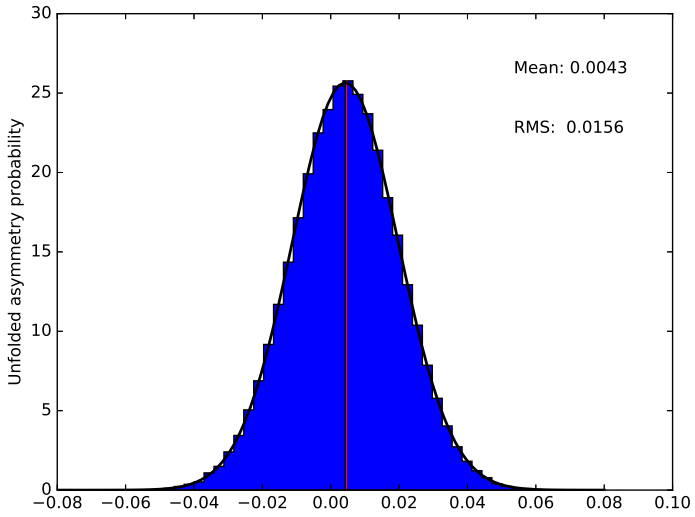


Figure 6.4 – *Posterior probability density for the inclusive $t\bar{t}$ charge asymmetry observable in the full phase space. The mean of the posterior is taken as the asymmetry value, while the RMS is taken as its uncertainty. A Gaussian fit is shown for illustration purposes.*

uncertainty is inferred by exploring the possible values of the systematic effects using a free parameter (called nuisance parameter) that extends the likelihood $\mathcal{L}(\boldsymbol{D}|\boldsymbol{T})$. This procedure is known as the *marginalization* procedure. The marginal likelihood is defined as:

$$\mathcal{L}(\boldsymbol{D}|\boldsymbol{T}) = \int \mathcal{L}(\boldsymbol{D}|\boldsymbol{T},\theta)\cdot\pi(\theta)d\theta, \tag{6.10}$$

where θ are the nuisance parameters, and $\pi(\theta)$ their prior probability densities, which are assumed to be Gaussian distributions $\mathbf{G}$ centered at zero with RMS of one. A nuisance parameter is associated with each of the uncertainty sources affecting the measurement. Two categories of nuisances parameters are considered: the normalizations of the background processes (θ_b), and the uncertainties associated to the object identification, reconstruction and calibration (θ_s). While the first ones only affect the background predictions, the latter, referred to as object systematic uncertainties, affect both the reconstructed distribution for $t\bar{t}$ signal and the total background prediction, referred to as $R(T;\theta_s)$ and $B(\theta_s,\theta_b)$, respectively. The marginal likelihood becomes then:

$$\mathcal{L}(\boldsymbol{D}|\boldsymbol{T}) = \int \mathcal{L}(\boldsymbol{D}|\mathbf{R}(T;\theta_s),\mathbf{B}(\theta_s,\theta_b))\cdot\mathbf{G}(\theta_s)\cdot\mathbf{G}(\theta_b)d\theta_s d\theta_b. \tag{6.11}$$

The differential likelihood $\mathcal{L}(\boldsymbol{D}|\boldsymbol{R}(\boldsymbol{T};\theta_s),\boldsymbol{B}(\theta_s,\theta_b))$ is defined as in Eq. 6.6, with

$$r_i(\boldsymbol{T},\mathcal{M};\theta_s) = r_i(\boldsymbol{T},\mathcal{M};0)\cdot(1+\sum_k \theta_s^k\cdot\Delta r_i^k), \tag{6.12}$$

and, for each background process,

$$b_i(\theta_s,\theta_b) = b_i(0)\cdot(1+\theta_b\cdot\Delta b)\cdot(1+\sum_k \theta_s^k\cdot\Delta b_i^k), \tag{6.13}$$

where $r_i(\boldsymbol{T}, \mathcal{M}; 0)$ and $b_i(0)$ represent the nominal reconstructed-level $t\bar{t}$ signal and background predictions, respectively; Δb is the uncertainty on the background normalization; Δr_i^k and Δb_i^k are the systematic variations for $t\bar{t}$ and background, respectively, in bin i, corresponding to the uncertainty k. The sum runs over all sources of object systematic uncertainty. The prior probability densities $G(\theta_b)$ are truncated at $\theta_{min} = -1/\Delta b$ in order to avoid negative background normalizations.

The marginal posterior probability density for $\boldsymbol{T}$ is computed by sampling the $N_t + N_{np}$ parameter space, where N_{np} is the total number of nuisance parameters, and projecting the sampled points over the $\boldsymbol{T}$ parameter space. The posterior probability density for the asymmetry is computed as described previously with the difference that the RMS of the marginal posterior represents the total uncertainty. Analogously, each nuisance parameter is estimated by the mean and RMS of the corresponding projection of the posterior probability density.

Combination of the channels

The FBU method allows for the mapping from the reconstructed spectrum to the true spectrum without a specific requirement in their dimensions. A response matrix can be chosen to be rectangular. With this choice, a combined measurement of the ee, $\mu\mu$ and $e\mu$ is natural since the generated asymmetry is independent of the chosen channel.

The rectangular response matrix used for the inclusive $t\bar{t}$ asymmetry measurement is shown in Fig. 6.5. The first four bins in the $t\bar{t}$ case and the first sixteen bins in the leptonic case correspond to the ee channel, the next four and sixteen to the $\mu\mu$ and the last ones to the $e\mu$ channel. The matrix is normalized such that the sum of values in each row is equal to one for each channel. With this normalization, it is possible to read what fraction of the reconstructed events migrates out of the diagonal bins. In the migration matrix, around $\sim 50\%$ of the events do not migrate out of the diagonal elements. The migrations for the leptonic asymmetry are around 2%, making the response matrix very diagonal. The binning corresponds to the optimal configuration, which is found by considering several criteria which will be discussed in detail in Chapter 8 and 9 for the charge asymmetry and spin measurements, respectively.
Using a rectangular response matrix, the reconstructed distributions must have the same number of bins as the projection of the response matrix along the reconstructed bins. The reconstructed distribution used for the $t\bar{t}$ asymmetry measurement is shown in Fig. 6.6. The first 4 $\Delta|y|$ bins correspond to the ee channel, followed by $\mu\mu$ and $e\mu$ matching the 4×12 migration matrix shown in Fig.6.5.

6.3 Two-dimensional unfolding

When the behavior of one observable needs to be studied as a function of another observable, it is possible to have nontrivial dependencies due to reconstruction and detector effects. A simple example of such effects can be seen in the migration matrix presented in Fig. 6.7. The first and last four bins of the matrix correspond to the four bin $\Delta|y|$ distribution in two different regions of invariant mass of the $t\bar{t}$ system. It is possible to have migrations between the different invariant

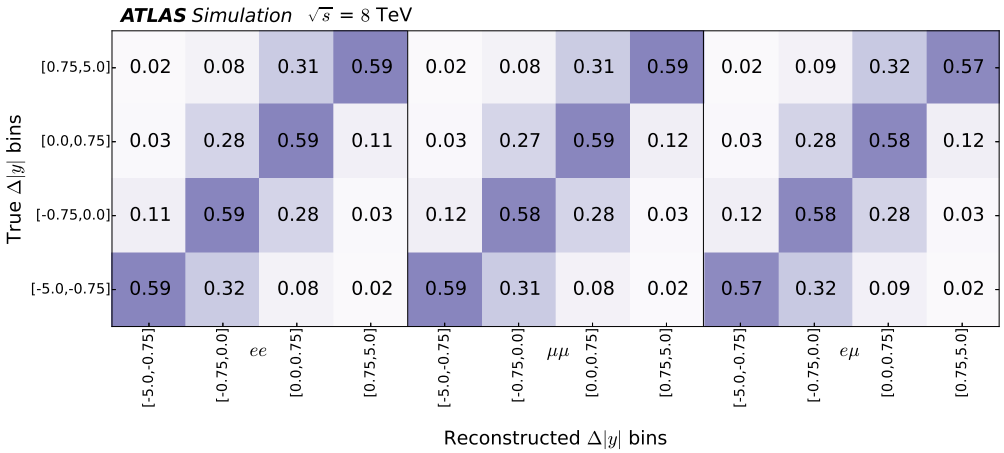
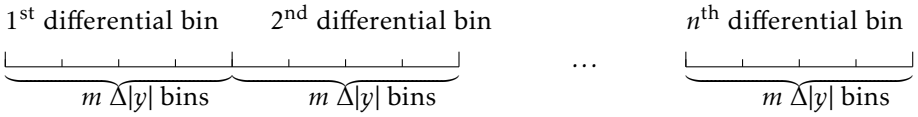


Figure 6.5 – *Rectangular response matrix for the $\Delta|y|$ observable corresponding to the inclusive measurement in the fiducial region. The numbers are normalized such that the sum of values in each row is equal to 1 for each channel.*

mass regions where the bins of the distributions can migrate between each other. For the proper treatment of these migration, an approach known as two-dimensional (2D) unfolding is applied.

A 2D unfolding is needed in order to perform double differential measurements of some observable. For the differential measurements of the charge asymmetry, the same method and setup as for the inclusive measurement (one-dimensional unfolding) is used. The second dimension[2] in this context corresponds to the differential variables to be studied: $m_{t\bar{t}}$, $\beta_{z,t\bar{t}}$, or $p_{\mathrm{T},t\bar{t}}$. The 2D unfolding is performed by transforming two-dimensional histograms into one-dimensional histograms. Assuming that there are m bins for $\Delta|y|$ and n bins for the differential variable, a one-dimensional histogram is created with $m \times n$ bins where the second dimension of the two-dimensional histogram is wrapped as several consecutive sub-ranges of the first dimension. The one-dimensional histogram considered for the unfolding procedure is schematically represented below:



[2]The first dimension corresponds to the $\Delta|y|$ distribution.

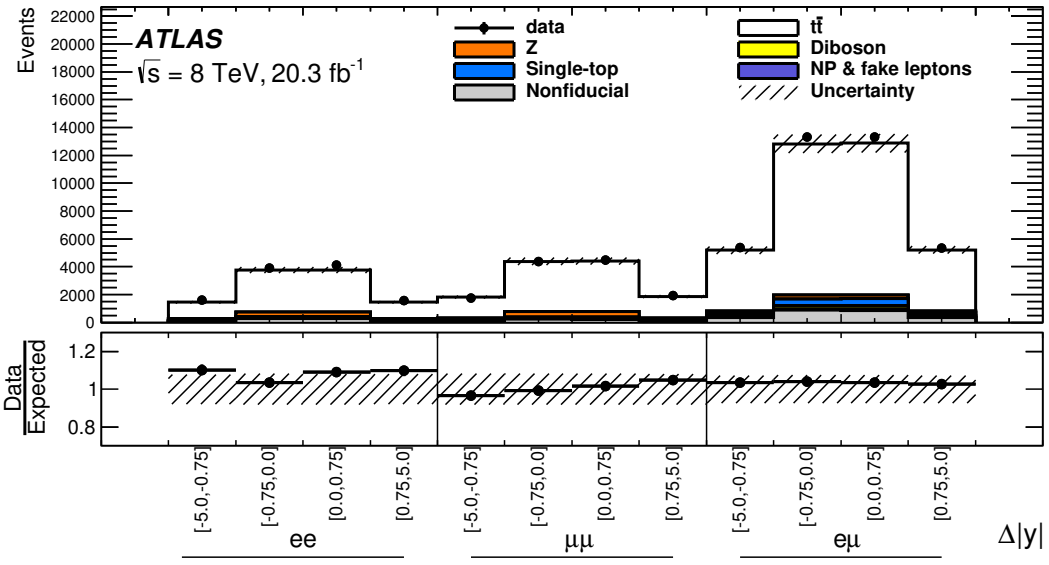


Figure 6.6 – *Input distributions for the inclusive $t\bar{t}$ asymmetry measurement. The first 4 $\Delta|y|$ bins correspond to the ee channel, followed by $\mu\mu$ and $e\mu$. The data/expected ratio is also shown.*

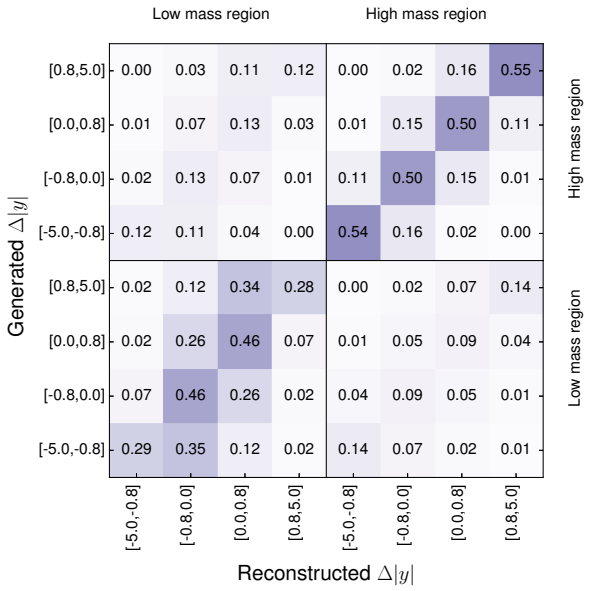


Figure 6.7 – *Example of 2-dimensional migration matrix normalized by row. The first and last four bins of the matrix correspond to the four bin $\Delta|y|$ distribution in two different regions of invariant mass of the $t\bar{t}$ system.*

7

Systematic Uncertainties

Two kind of uncertainties affect the measurements performed in physics: statistical uncertainties and systematics uncertainties. The statistical uncertainties are the result of stochastic fluctuations arising from the fact that a measurement is based on a finite set of observations. Systematic uncertainties arise from uncertainties associated with the nature of the measurement apparatus, assumptions made by the experimenter, or the model used to make inferences based on the observed data. The sources of systematic uncertainties common to the charge asymmetry and top spin measurements are described in this chapter, as well as the technique used for their estimation.

Several classes of systematic uncertainties affect the presented measurements. They can be categorized as follows:

- **The detector modeling uncertainties**, which involve the uncertainties and assumptions made on the reconstructions of the physics objects, such as leptons, jets, and $E_{\mathrm{T}}^{\mathrm{miss}}$.
- **The background modeling uncertainties**, which is related with the assumptions made during the background estimation and uncertainties coming from the cross section estimations.
- **The signal modeling uncertainties**, which come from the assumptions made by the MC generators on modeling of the hard scattering process.
- **Other uncertainties**, which are the uncertainties that are not mentioned in the previous categories, such as those related to the unfolding, kinematic reconstruction, and the limited size of the MC samples.

Most of the systematic sources possess an *up* and *down* variation. In some cases it is only possible to define one variation. This depends on the parameters used to define how the systematic variations affect the measurements. It is possible to visualize this statement in the following example: if a systematic uncertainty is associated with a scale factor, it is possible to reduce or to increase the effect on the central value by the uncertainty on the scale factor. This will lead to an up and down variation. On the other hand, if an uncertainty is associated with a comparison between two different approaches used to estimate a given effect, only one variation is obtained. The uncertainties are expected to correspond to a 68% confidence interval (1σ deviation), such as the

detector modeling, background, and some sources of the signal modeling uncertainties (which will be described below). However, in some cases a clear definition of a 1σ deviation is not possible, due to the lack of knowledge of how the systematic uncertainties impact the measurements. This is the case for instance if the effect comes from the limited size of the MC samples where a different estimation needs to be performed. Some sources of signal modeling uncertainties are affected by the limited size of the MC samples and a conservative approach is taken (as will be described below).

7.1 Detector and background uncertainties

The detector and background systematic uncertainties are estimated within FBU through the marginalization procedure. The total uncertainty obtained from the posterior probability density includes these systematic uncertainties together with the statistical uncertainty. Each systematic source of this category is associated with a nuisance parameter that constrains the likelihood considered in the unfolding procedure. In the following the description of the detector and background uncertainties is presented.

Lepton-related uncertainties

The reconstruction and identification efficiencies of electrons and muons, as well as the efficiency of the triggers used to record the events, differ between data and simulation. Scale factors and their uncertainties are derived using tag-and-probe techniques on $Z \to \ell^+\ell^-(\ell = e, \mu)$ events in data and in simulated samples to correct the simulation for these differences [101, 102, 146, 147]. The accuracy of the lepton momentum scale and resolution in simulation is also checked using reconstructed distributions of the $Z \to \ell^+\ell^-$ and $J/\psi \to \ell^+\ell^-$ masses. In the case of electrons, energy and momentum studies using $W \to e\nu$ events are also used. Small differences are observed between data and simulation. Corrections for the lepton energy scale and resolution, and their related uncertainties are also considered [101, 102, 147].

Jet-related uncertainties

The jet energy scale and its uncertainty are derived combining information from test-beam data, LHC collision data, and simulation [110]. The jet energy scale uncertainty is split into 22 uncorrelated sources that have different jet p_{T} and η dependencies and are treated independently in this analysis. The total jet energy scale uncertainty is one of the dominant uncertainties across the measurements involving the reconstruction of the $t\bar{t}$ system, since small fluctuations on the jet energy could lead to a different reconstruction. The jet reconstruction efficiency is found to be 0.2% lower in simulation than in data for jets below 30 GeV and consistent with data for higher jet p_T. All jet-related kinematic variables (and the missing transverse momentum) are recomputed by removing randomly 0.2% of the jets with p_{T} below 30 GeV and the event selection is repeated. Also, the efficiency for each jet to satisfy the JVF requirement is measured in $Z \to \ell^+\ell^-$ + 1-jet events in data and simulation [121]. The corresponding uncertainty is evaluated in the analysis by changing the nominal JVF cut value and repeating the analysis using the modified cut value.

The uncertainty related to the jet energy resolution is estimated by smearing the energy of the jets in simulation by the difference between the jet energy resolutions for data and simulation [113]. Finally, the efficiencies to tag jets from b- and c-quarks, light quarks, and gluons in simulation are corrected by p_{T}- and η-dependent data/MC scale factors [117,118,148]. The uncertainties on these scale factors are propagated to the measured value by rescaling the distributions by the size of the uncertainties. The uncertainties on the scale factors are split into components depending on the nature of their effect: eight components for b-tagging efficiency and six components for c-tagging efficiency. The mistag rate and the c/τ rates, defined as fraction of jets originating from light flavor, c-quarks, and hadronic tau decays which are tagged by a b-tagging algorithm, are also considered. The c/τ rates are split into six different components depending on the nature of their effect.

Missing transverse momentum

The systematic uncertainties associated with the momenta and energies of reconstructed objects (leptons and jets) are also propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation (called *hard term*). The $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction also receives contributions from the presence of low-p_{T} jets and calorimeter cells not included in reconstructed objects (called *soft term*). The systematic uncertainty on the soft term is evaluated using $Z \to \mu^+\mu^-$ events, comparing data and MC simulation using the energy and momentum response from single tracks, measured by summing the energies of all calorimeter clusters around a single isolated track [119].

Table 7.1 shows the notation used for each systematic uncertainty.

Type	Notation	Name
Lepton	ees	Electron energy scale
	eer	Electron energy resolution
	el_idSF	Electron identification efficiency scale factor
	el_recSF	Electron reconstruction efficiency scale factor
	el_trigSF	Electron trigger efficiency scale factor
	mums	Muon momentum resolution measured with the MS
	muid	Muon momentum resolution measured with the ID
	musc	Muon momentum scale
	mu_idSF	Muon identification efficiency
	mu _recSF	Muon reconstruction efficiency
	mu_trigSF	Muon trigger efficiency
$E_{\mathrm{T}}^{\mathrm{miss}}$	res_soft	E_T^{miss} soft term resolution
	sc_soft	E_T^{miss} soft term scale
Jet	btagN	N-th component of b-tagging scale factors
	mistag	b-tagging mistag rate
	ctautagN	N-th b-tagging c/τ tag rates
	JesEffectiveModel	Jet energy scale effective model
	JesEffectiveDet	Jet energy scale effective detector
	JesEffectiveStat	Jet energy scale effective statistical
	JesEffectiveMix	Jet energy scale effective mix
	EtaIntercalibration	Jet energy scale η inter-calibration
	flavor_comp	Jet energy scale flavor composition
	BJesUnc	Jet energy scale b-JES
	Pileup_Rho	Jet energy scale pileup ρ
	flavor_response	Jet energy scale flavor response
	Pileup_OffsetNPV	Jet energy scale pileup offset (NPV)
	Pileup_OffsetMu	Jet energy scale pileup offset ($<\mu>$)
	Pileup_Pt	Jet energy scale pileup (p_T)
	SinglePart	Jet energy scale single-particle
	jvf	Jet vertex fraction
	jer	Jet energy resolution
	jeff	Jet reconstruction efficiency

Table 7.1 – *Notation for all the detector systematic uncertainties considered in the analyses presented in this dissertation.*

Background-related uncertainties

The background estimation is affected by two different uncertainties: normalization uncertainties and shape uncertainties. Uncertainties on the shape of the background contributions are negligible in all the measurements performed in this work and are therefore not included. The normalization uncertainty on each individual background contribution is described in the following.

The uncertainties on the single-top-quark and diboson backgrounds are about 7% and 5%, respectively [130, 131]. These correspond to the uncertainties on the theoretical cross sections used for the normalization of the MC simulated samples.

The uncertainty on the normalization of the fake-lepton background is evaluated by using different MC generators for each process contributing to this background. The scale factor in the control region is recomputed for each variation and the change is propagated to the expected number of events in the signal region. In the $\mu\mu$ channel, the uncertainty is obtained by comparing the results obtained using the matrix method (described in Sec. 4.2.4), and the estimation used in this analysis. Following a Bayesian procedure assuming constant a priori probability for a non-negative number of events, the resulting total relative uncertainties are $^{+193\%}_{-47\%}$ in the ee, $^{+80\%}_{-53\%}$ in the $\mu\mu$, and $^{+49\%}_{-45\%}$ in the $e\mu$ channel.

In the case of the Drell–Yan events, the detector modeling systematic uncertainties described previously are propagated to the scale factors derived in the control region by recalculating them for all the systematic uncertainty variations. An additional uncertainty of 6% is estimated by varying the Z-mass window of the control region used to obtain the scale factors and is added in quadrature to obtain the final uncertainty on these scale factors.

The background-related uncertainties represent a minor source of uncertainty on the different measurements.

7.1.1 Significance of the variations

The significance of each systematic uncertainty with respect to the MC statistics is evaluated. The purpose of this procedure is to exclude any systematic variation which is not significant with respect to MC statistics. The significance evaluation of each systematic uncertainty is performed by using MC replicas. The replicas are produced by assigning a weight to each event using Poisson fluctuations. The weights are produced such as the sum off all the weights is equal to one. Applying the same weights to the nominal and the systematic sample, makes the replicas statistically correlated. This procedure is known as the *bootstrap* method. Some systematic variations affect the final number of selected events, by allowing the access of more or less events after the selection requirements. The bootstrap method simulates the statistical effect of having more or less events passing the selection after having applied a systematic variation.

After the bootstrapping procedure is applied, the relative difference is calculated between a given systematic variation and the nominal sample for all the replicas on each bin of the distributions. An example of the relative difference in the first bin of the $\Delta|y|$ distribution between a b-tagging component and the nominal sample is shown in Fig. 7.1. The standard deviation of the relative differences is taken as the statistical uncertainty on that component. A systematic uncertainty is considered to be significant on a distributions if there is at least one bin which is not

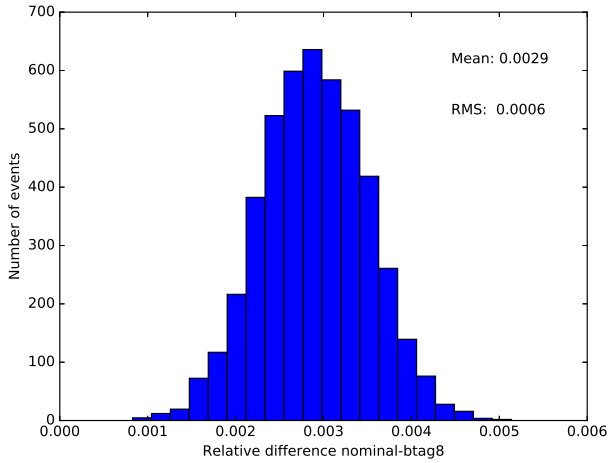


Figure 7.1 – *Relative difference between the nominal sample and btag8 systematic variation after applying bootstrapping, on one example bin of the* $\Delta|y|$ *observable.*

consistent with zero within one standard deviation. This means that the difference between the nominal sample and the systematic variation is larger than the MC statistical uncertainty. All the systematic uncertainties described in Sec. 7.1 are significant with respect to the MC statistics. Figure 7.2 shows examples for the b-tagging, the electron identification scale factors, one component of the jet energy scale, and the jet energy resolution systematic variations for the $\Delta|y|$ distribution. The shaded area represents the data uncertainty and the MC statistical uncertainty correspond to the uncertainty on the points. The x-axis on these distributions correspond to the bins of the distributions, after performing a binning optimization. For the b-tagging systematic uncertainty, a large variation is observed in the ee and $\mu\mu$ channels with respect to the nominal $\Delta|y|$ distribution, while for $e\mu$ this is not the case. This behavior is expected due to the fact that b-tagging is not used in the $e\mu$ channel, making this channel insensitive to that systematic variation. A similar case is observed for the electron reconstruction systematic where the $\mu\mu$ channel is not sensitive to that systematic variation, while in the ee and $e\mu$ channels a big variation is observed. This is due to the electrons present in the ee and $e\mu$ final state. A similar behavior is observed across the different top spin observables, except for the fact that the systematic variations associated with the b-tagging are not larger than the statistical uncertainty on data.

7.1.2 Nuisance parameters

The relative differences between the nominal and the up and down systematic variations are averaged to build the Gaussian prior for the marginalization procedure in FBU. In the cases of systematic uncertainties that do no have an up or a down variation, such as the jet energy resolution, the variation is symmetrized to build the Gaussian prior.

The nuisance parameters obtained after applying the unfolding procedure on pseudodata are shown in Fig. 7.3 for the inclusive $t\bar{t}$ asymmetry and in Fig. 7.4 for the inclusive leptonic

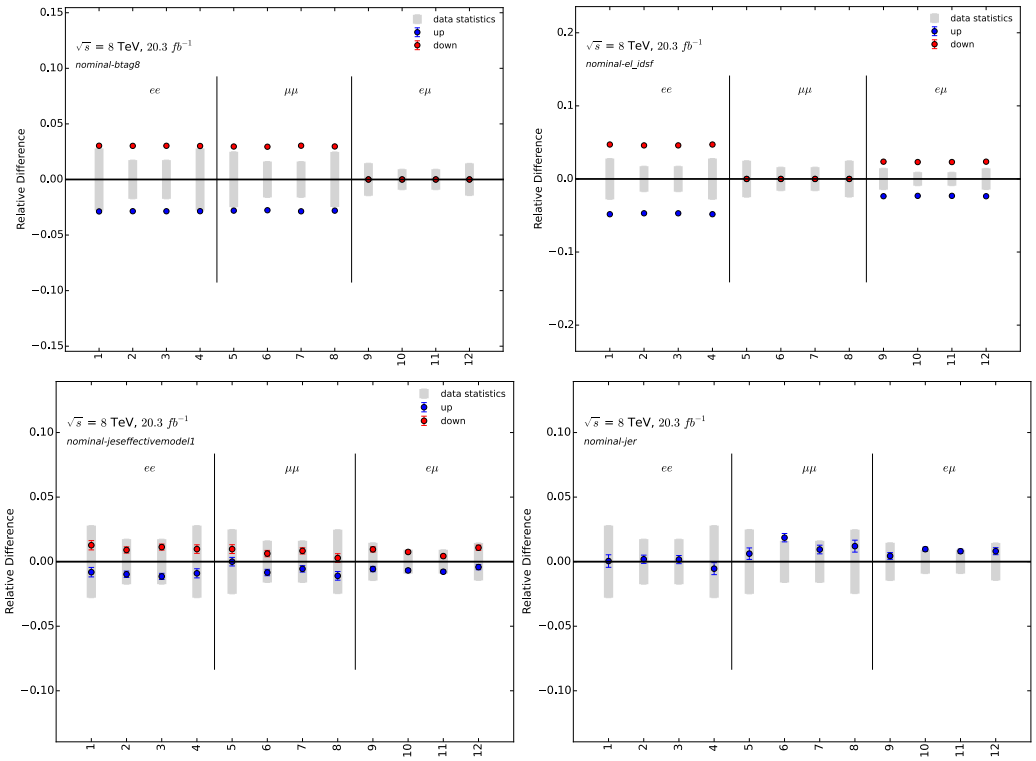


Figure 7.2 – *Relative variations for the 8th component of b-tagging (top left), electron identification (top right), jet energy scale effective model (bottom left), and jet energy resolution (bottom right) with respect to the nominal $\Delta|y|$ distribution for the inclusive $t\bar{t}$ asymmetry. Shaded areas represent the statistical uncertainty on data for each bin. The x-axis corresponds to the bin number of the distribution.*

asymmetry. The pseudodata is made of the sum of signal and background, which implies that the expected value of each nuisance parameter is zero while the standard deviation represents the constraint on the considered systematic variation (a standard deviation of one represents no constraint). There are small fluctuations of the nuisance parameters around zero. This is coming from the stochastic nature of the scanning performed during the unfolding. The strongest constraint is observed in *btag8* and the electron identification scale factor variations. The same effect is observed across the different charge asymmetry measurements. These constraints come from the fact that some of the channels the systematic variations are larger than the data statistical uncertainty. This is the case for the *btag8* component, where the *ee* and $\mu\mu$ channels are sensitive to this variation. The same effect is observed in the electron identification scale factor variation, but in this case the *ee* and $e\mu$ channels are sensitive to the variation (Fig 7.2). For other systematic variations, such as the jet energy resolution and the jet energy scale components, the data is not sensitive to the variation, and therefore no constraint is observed.

For the top spin observables a similar behavior as in the charge asymmetry is observed in all nuisance parameters. The only exception is on the *btag8* component. Across the different top spin observables, the statistical uncertainty on data is larger than the *btag8* systematic variation. This implies that no constraint is expected on that systematic variation. The constraint observed in

the electron identification scale factor for the charge asymmetry measurements is also observed across the top spin observables. The constraints obtained for $C(k,k)$ using pseudodata are shown in Fig. 7.6, while for B_+^k they are shown in Fig. 7.5.

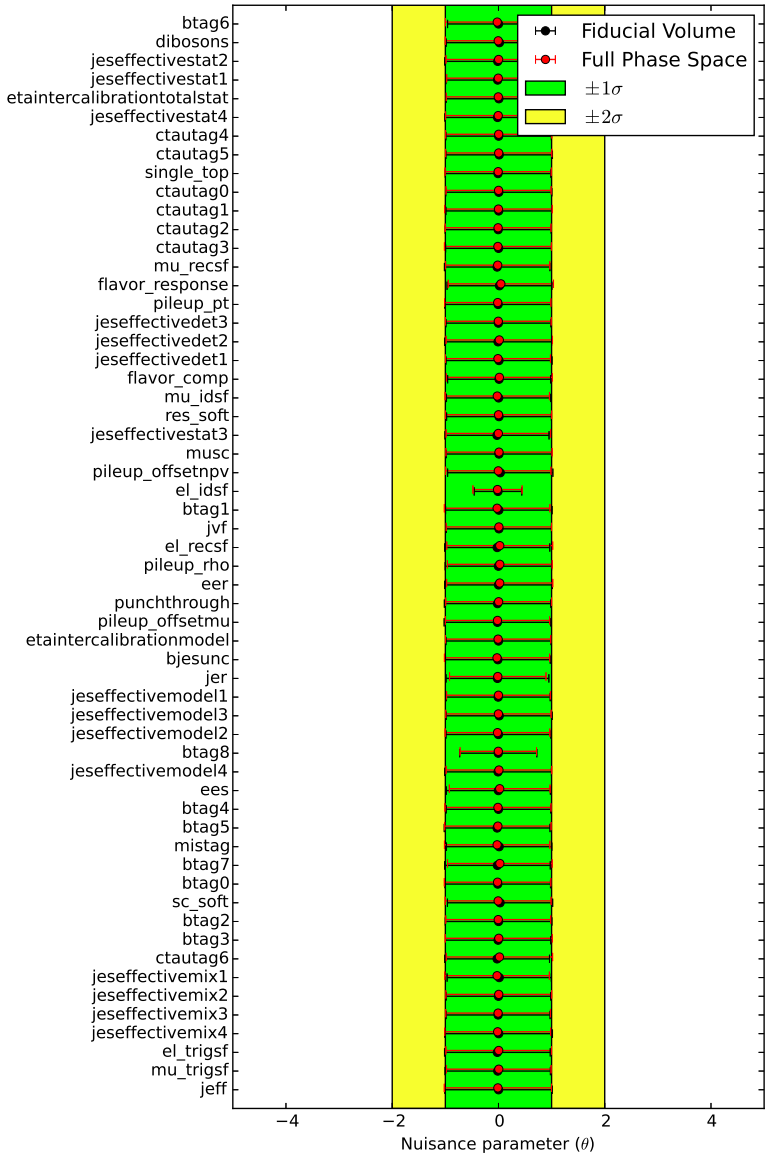


Figure 7.3 – *Nuisance parameters for the inclusive $t\bar{t}$ asymmetry in the fiducial volume and full phase space using pseudodata.*

7.1.3 Individual estimation

The individual estimation of each systematic uncertainty is important to know since it gives information about how much each source of systematic uncertainty is affecting the measurement. This is a useful piece of information that can be used to properly combine measurements be-

tween different decay channels and experiments. Since the systematic estimation for detector modeling is part of the unfolding procedure, it is not possible to obtain the impact of each individual systematic uncertainty on the measurement. A specific procedure is applied to estimate the impact of each source of systematic uncertainty using pseudodata, corresponding to the sum of the nominal signal and background samples. The unfolding procedure with marginalization is applied to the pseudodata and constraints on the systematic uncertainties are obtained as in the examples shown from Fig. 7.3 to 7.6. These constraints are then used to build the $\pm 1\sigma$ variations of the prediction. If a systematic uncertainty is not constrained, the varied sample will correspond to the initial $\pm 1\sigma$ variation. If the systematic uncertainty is constrained, a more precise estimation of the real effect in the final measurement is obtained. The varied pseudodata are then unfolded without marginalization. The impact of each systematic uncertainty is computed by taking half of the difference between the results obtained from the $\pm 1\sigma$ variations built using pseudodata since they are considered to be symmetric (following a Gaussian posterior). This is only an approximate estimate of the individual contribution of each source of systematic uncertainty within the overall marginalization procedure since multidimensional fit is performed is this case.

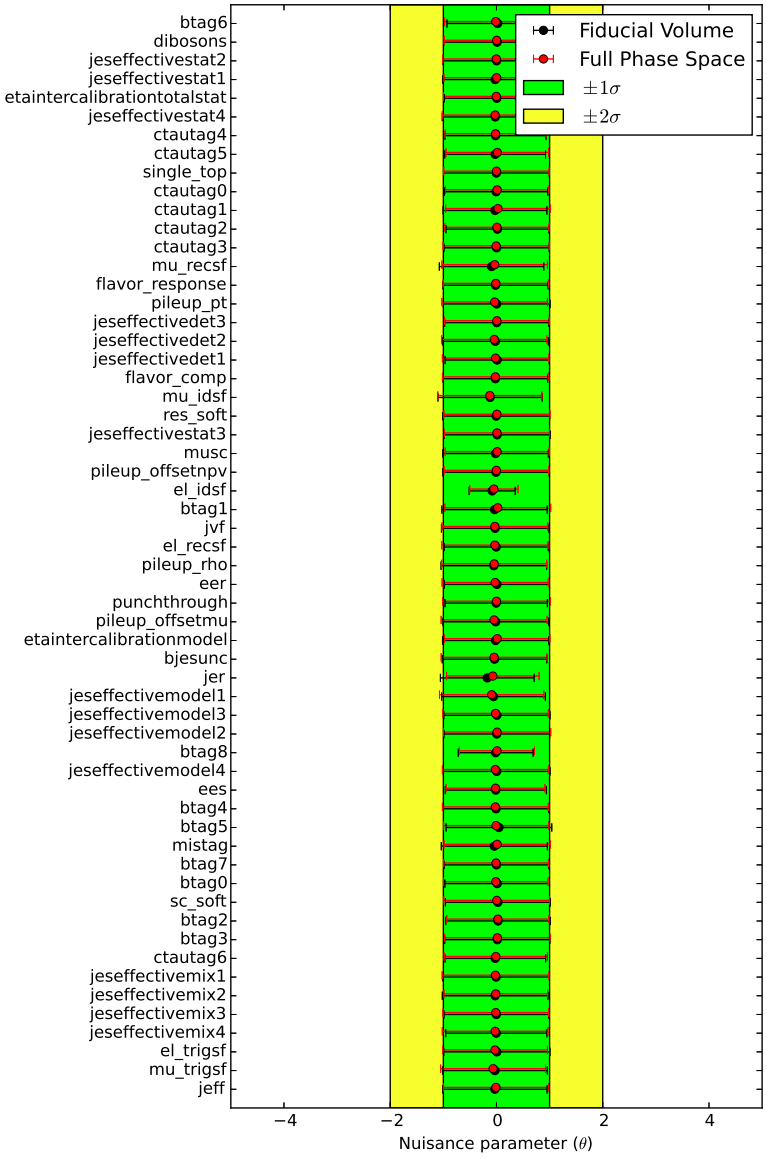


Figure 7.4 – *Nuisance parameters for the inclusive leptonic asymmetry in the fiducial volume and full phase space using pseudodata.*

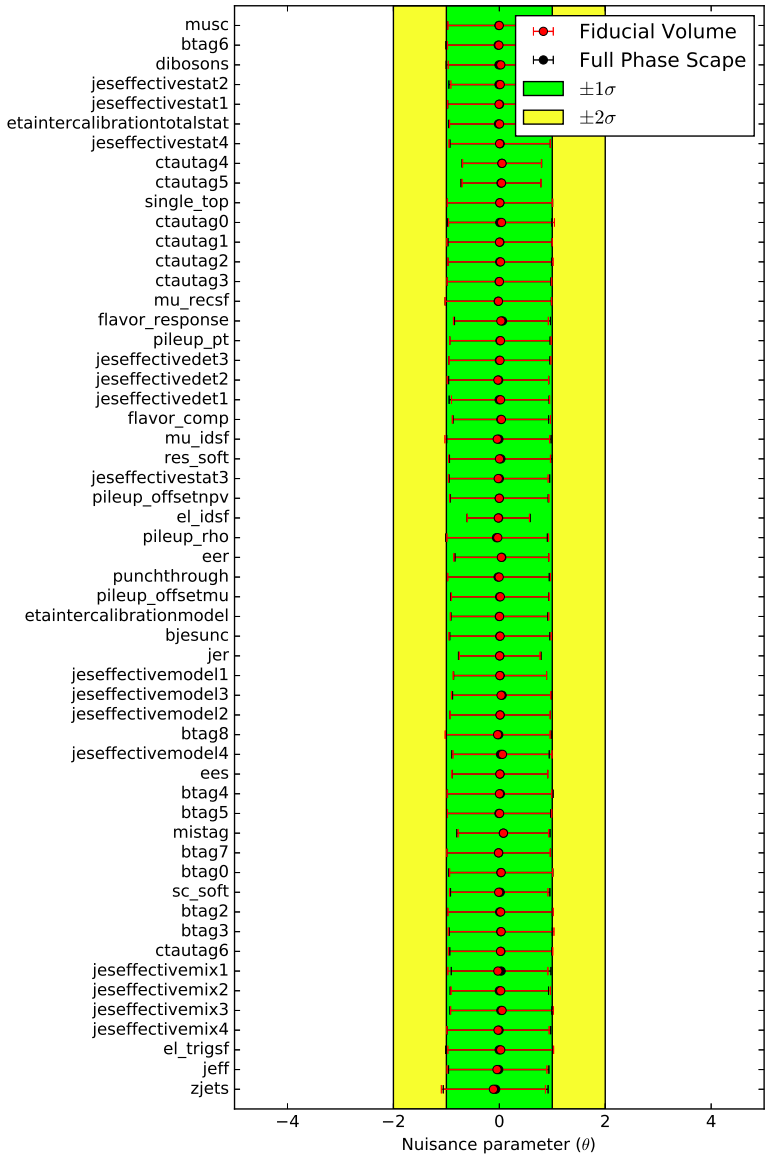


Figure 7.5 – *Nuisance parameters for $C(k,k)$ in the fiducial volume and full phase space using pseudodata.*

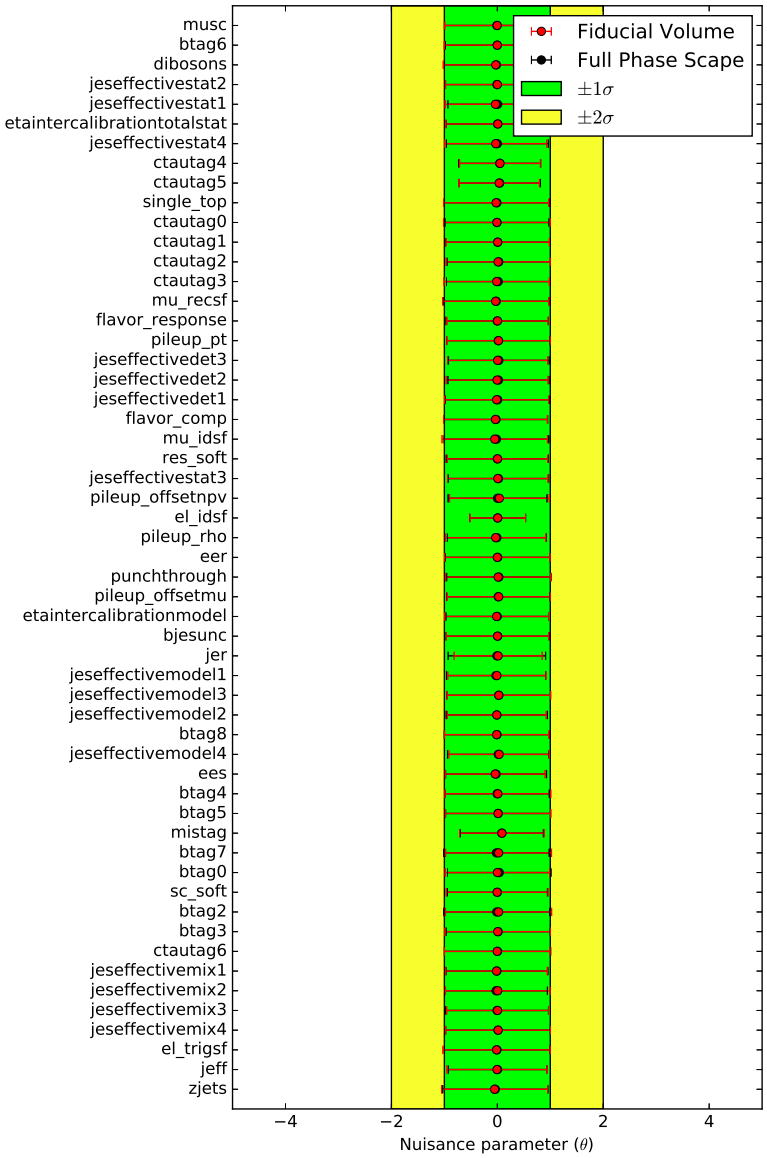


Figure 7.6 – *Nuisance parameters for B_+^k in the fiducial volume and full phase space using pseudodata.*

7.2 Signal modeling uncertainties

The signal modeling uncertainties are not estimated through the marginalization procedure, since they affect the assumptions made on the choice of the models that describe the full chain of the pp collision before any interaction with the detector. For these uncertainties, distributions obtained with different generators (described below) and reweighted to different levels of asymmetries, polarizations, or correlations are unfolded (depending on which observable is studied). The unfolded values are compared with the injected asymmetries, polarizations, or correlations and the calibration curves are obtained by fitting the points using a linear function. This procedure will be explained in more details in Chapter 8. The measured value in data is extrapolated using the slopes and offsets of the calibration curves. This procedure is repeated multiple times by varying the distributions within their MC statistics. A spectrum of estimations that follows a Gaussian distribution is obtained. The mean of the distribution corresponds to the systematic uncertainty estimation and its standard deviation to the MC statistical uncertainty on the systematic uncertainty. The largest value between the mean and the standard deviation is taken as the actual modeling uncertainty. This means that a conservative approach is taken given the limited size of the MC samples. An example of one calibration curve used for the estimation of the parton shower systematic uncertainties (described below) of the $t\bar{t}$ charge asymmetry in the fiducial volume is shown in Fig. 7.7. In the figure, the calibration curves for MC@NLO and Powheg-hvq in the fiducial volume are shown. The response matrix from the nominal Powheg-hvq sample is used. The difference in the slope and in the offset are taken and the measured value in data is used to obtain the estimation. The procedure is repeated varying the input distribution within their MC statistical uncertainty. The corresponding distribution of the systematic estimation is shown in Fig. 7.8. The largest value between the mean of that distribution and the RMS is taken as the actual systematic uncertainty. For each source of uncertainty the following MC generators and comparison are used to obtain the calibration curves and subsequently the systematic uncertainty:

- **MC generator**. The uncertainty is obtained by comparing the Powheg-hvq and MC@NLO samples, both interfaced with Herwig. It is one of the dominant uncertainties on the measurements.

- **Parton shower and hadronization**. This effect is estimated by comparing Powheg-hvq interfaced either with Pythia or Herwig, and is one of the dominant systematic uncertainties.

- **ISR/FSR**. The uncertainty is obtained by taking half of the difference between MC samples producing more or less ISR/FSR. The samples are generated with Powheg-hvq interfaced with Pythia for which the parameters of the generation were varied to span the ranges compatible with the results of measurements of $t\bar{t}$ production in association with jets [149].

- **Color reconnection and underlying event.** The uncertainties associated to color reconnection and underlying event are obtained by comparing dedicated samples with a varied strength of color reconnection and activity of the underlying event to a reference sample.

All samples are generated by Powheg-hvq and interfaced with Pythia. The reference sample uses the Perugia2012 tune, the color reconnection sample the Perugia2012loCR tune and the underlying event sample the Perugia2012mpiHi tune.

- **PDF**. Uncertainties corresponding to the choice of the PDF in the MC generator are obtained by reweighting the nominal signal sample using the error sets of CT10, MWST2008 and NNPDF23, and following the prescriptions recommended by the PDF4LHC working group [150]. The error sets are provided by each of the PDF working groups and they are obtained by varying the parameters of the fit performed to obtain the PDF. The error sets consist of 52 sets of CT10, 40 sets of MWST2008 and 100 sets of NNPDF23. The uncertainty for CT10 is computed using a symmetric Hessian[1] while the uncertainty for MSTW2008 is estimated using an asymmetric Hessian[2]. For NNPDF23 the standard deviation of all the variations is taken [150]. The total uncertainty is obtained by taking half of the extremum of all variations added linearly. This method is called the *envelope* method. An example of the envelope created for the PDF uncertainty estimation is shown in Fig 7.9. Each point represents a reweighted sample using different PDF estimations. The color bands represent the uncertainty estimation for each reweighted sample. The uncertainty is taken as half of the envelope created by adding the color bands linearly.

- **Top quark mass uncertainty.** The top quark mass is assumed to be of 172.5 GeV in MC simulation and in the reconstruction method. A variation of this value could have an impact on the measurement. To estimate this impact, MC samples with different values of the top quark mass are unfolded with the default response matrix. For each observable, the dependence of the unfolded value with respect to the mass is fitted with a linear function. The slope is then multiplied by the uncertainty on the most precise ATLAS top quark mass measurements of 0.70 GeV [151].

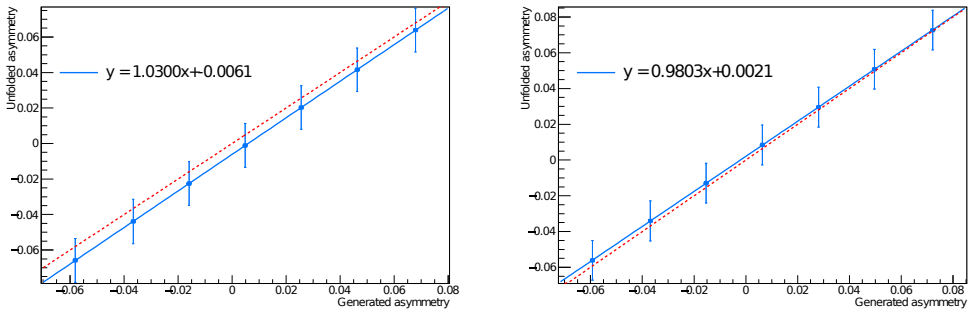


Figure 7.7 – *Calibration curves for the inclusive measurement in the fiducial volume for* MC@NLO *(left) and* Powheg-*hvq (right), both interfaced with* Herwig. *The nominal response matrix is used during the unfolding procedure.*

[1] The symmetric Hessian Δ is defined as $\Delta = \frac{1}{2}\sqrt{\sum_i x_{+i} - x_{-i}}$, where $x_{\pm i}$ correspond the $\pm 1\sigma$ variation on the PDF set.

[2] The asymmetric Hessian Δ is defined as in the symmetric case, but the quadratic difference is performed between the nominal and the $\pm 1\sigma$ variation. This define the upper part and lower part of the asymmetric Hessian.

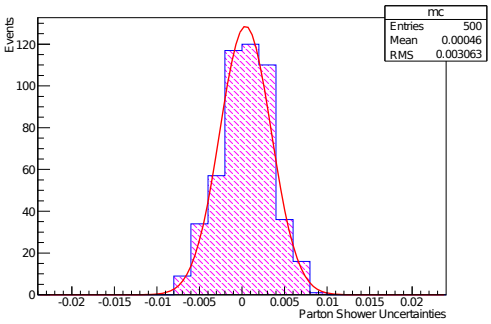


Figure 7.8 – *Distribution used for the estimation of the parton shower uncertainty of the $t\bar{t}$ asymmetry in the fiducial region. The actual estimation corresponds to the highest value between the mean and the RMS of the distribution. In this case the RMS is taken as the systematic uncertainty. The distribution is fit to a Gaussian function for illustration purposes.*

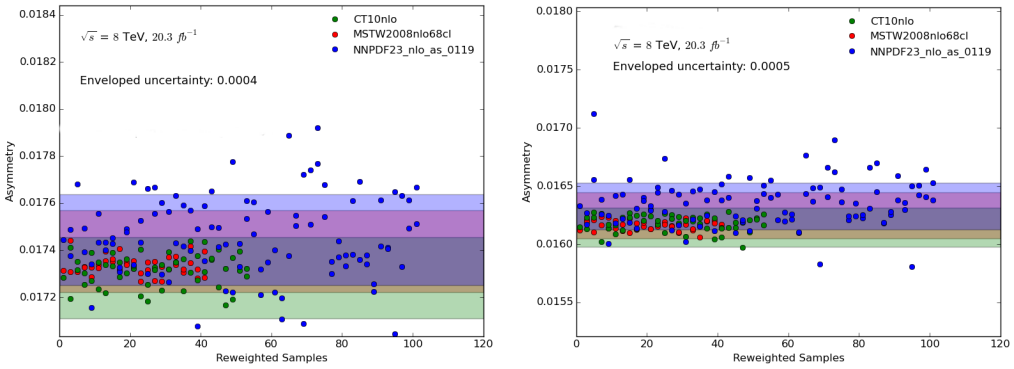


Figure 7.9 – *Asymmetry value for all the reweighted samples used in the PDF uncertainty estimation. The top asymmetry is shown for the inclusive measurement in fiducial volume (left) and full phase space (right). The color bands represents the uncertainty estimation for each PDF set. The total uncertainty is estimated by taking half of the envelope obtained using the different PDF sets.*

7.3 Other uncertainties

This category corresponds to the uncertainties that do not enter in the previous categories and are related to methods specific to the analyzes described here. The common uncertainties affecting the charge asymmetry measurements, spin correlations, and polarizations are the uncertainties related with the bias introduced by FBU (called *nonclosure uncertainty*) and the effect from the MC statistical uncertainty. These sources of systematic uncertainties are discussed below.

7.3.1 Nonclosure uncertainty

When the calibration curve for the nominal signal Powheg-hvq sample is estimated, a residual slope and a nonzero offset are observed. This is due to the stochastic nature of the scanning procedure of the possible true points during the unfolding and the linearity of the method. This bias, introduced by the unfolding procedure, is propagated to the measured values by multiplying the residual slope times the measured value in data and adding the offset of the linear fit.

7.3.2 MC statistical uncertainty

The uncertainty associated with the limited size of the nominal signal Powheg-hvq sample is evaluated by varying the migration matrix within its MC statistical uncertainty and repeating the unfolding procedure for each variation. The standard deviation of the unfolded observables is taken as the uncertainty.

8

Charge asymmetry measurements

In this chapter, the measurements of the charge asymmetry are presented. In the first part, the optimization of the unfolding that leads to the optimal binning configuration for the different observables, is described. This is followed by the estimation of the systematic uncertainties described in Chapter 7, as well as the description and estimation of the systematic uncertainties specific to this analysis. The final part presents and discusses the results obtained together with an outlook for the measurements.

8.1 Unfolding optimization

In this analysis, the unfolding procedure brings the reconstructed distributions back to particle level in the fiducial volume and back to parton level in the full phase space. Both procedures are performed for the $t\bar{t}$ and the leptonic asymmetries.
Several studies are performed in order to validate and optimize the unfolding procedure. The validation and optimization of the unfolding procedure are performed at particle level in the fiducial volume. The configuration obtained after performing the optimization is also checked and used at parton level in the full phase space. The validity of this approach is cross-checked by performing some of the optimization studies at parton level and comparing them with the results obtained at particle level. All tests are performed using the $e\mu$ channel since it is the channel that offers the largest amount of statistics. The optimization is also cross-checked on the combination of the channels.

8.1.1 Inclusive measurement

The performance of FBU depends on the number and size of the bins used to define the histograms. The configurations that are more suitable for the measurements are studied. The different criteria used to discriminate between the bin configurations are presented in the following.

Observable resolution

The resolution of the $\Delta|y|$ and $\Delta|\eta|$ observables, defined in Sec. 5.3, is studied as a function of $\Delta|y|$ and $\Delta|\eta|$, respectively. An optimized configuration of the bin size is selected by studying

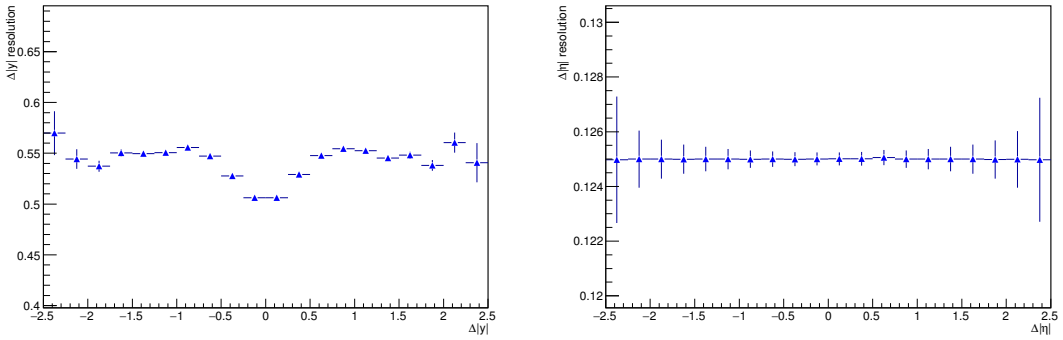


Figure 8.1 – *The $\Delta|y|$ resolution and $\Delta|\eta|$ resolution as a function of $\Delta|y|$ (left) and $\Delta|\eta|$ (right).*

the resolution of the bins, since the migrations between the generated and reconstructed bins are reduced when the width of the bin is larger than the resolution. The resolutions of the $\Delta|y|$ and $\Delta|\eta|$ observables as function of $\Delta|y|_{\text{reco}}$ and $\Delta|\eta|_{\text{reco}}$ are presented in Fig. 8.1, respectively. The events from the three dilepton channels (ee, $e\mu$ and $\mu\mu$) are considered. The resolution as well as its statistical uncertainty are shown. The resolution of the $\Delta|y|$ observable is around of 0.50 in the central region, and increases up to 0.55 in the forward regions. For $\Delta|\eta|$, the resolution is of around 0.12 in the central and forward regions. The main difference between the the resolution of the $\Delta|y|$ and $\Delta|\eta|$ observable is coming from the kinematic reconstruction (KIN Method) applied in the case of $\Delta|y|$, which is not needed for $\Delta|\eta|$.

Using these resolutions, several configurations for the bins are selected and used in different performance tests. These configurations are shown in Tab. 8.1[1] for the $\Delta|y|$ and $\Delta|\eta|$ distributions. For $\Delta|y|$ four, six, and eight bins are considered, while for $\Delta|\eta|$ fourteen, sixteen, eighteen, and twenty bins are considered. The names used to distinguish the configuration are also presented.

[1] More than 10 different configurations were defined and tested. This table presents the configurations of the binning candidates with the best performances.

Name	Configuration
$\Delta\lvert y\rvert$	
fourBin	[0.0, 0.6, 5.0]
fourBin2	[0.0, 0.75, 5.0]
fourBin3	[0.0, 0.8, 5.0]
sixBin	[0.0, 0.75, 1.5, 5.0]
eightBin	[0.0, 0.6, 1.2, 1.8, 5.0]
$\Delta\lvert\eta\rvert$	
bin14_ll	[0.0, 0.3, 0.6, 0.9, 1.1, 1.3, 1.5, 5.0]
bin16_ll	[0.0, 0.3, 0.6, 0.9, 1.2, 1.5, 1.7, 1.9, 5.0]
bin18_ll	[0.0, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.5, 5.0]
bin20_ll	[0.0, 0.1, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.5, 1.8, 5.0]

Table 8.1 – *Binning configurations tested during the optimization are shown. Only the positive part of the binning is shown. Each binning is symmetric with respect to 0.*

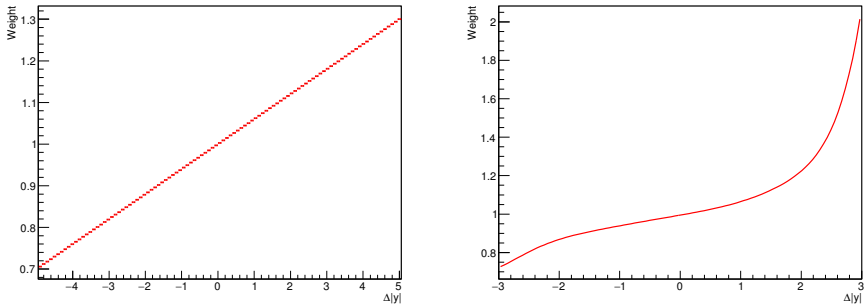


Figure 8.2 – *Reweighting functions used for the calibration procedure that correspond to the linear reweighting function (left) and Protos reweighting function (right). The functions correspond to 6% $t\bar{t}$ asymmetry.*

8.1.1. a Linearity Test

A good test to discriminate binning configurations, as well as to evaluate the performance of the unfolding procedure in general, is performed by using calibration curves that compare unfolded and generated asymmetries at different levels. This procedure is known as *linearity* test. To produce calibration curves, asymmetries are injected at generated level by reweighting the $\Delta|y|$ distribution using specific functions. The weight applied at generated level to reproduce different levels of asymmetry is also applied at reconstructed level. The reweighted distributions at reconstructed level are used as pseudodata and unfolded using the nominal response matrix. The compatibility between the unfolded and injected asymmetries is then checked. Finally, a linear fit is performed and the slope and offset are extracted and used as an estimator for the performance of the test. The quality of the fit is also checked to be linear.

The reweighting functions used to produce the calibration curves are either linear functions or functions based on BSM axigluon models in which the asymmetry value is predicted to be significantly different from the SM value. The axigluon models are generated with the Protos generator [152, 153]. The functions based on axigluon/ models are used for the main tests, while the linear functions are used for the complementary tests. Two reweighting functions can be found in Fig. 8.2. An example of the calibration curves is shown in Fig. 8.3. These calibration curves correspond to the fourBin2 and bin16_ll configuration for the $t\bar{t}$ and leptonic asymmetry. The slopes and offsets are close to one and to zero respectively.

Linearity tests are performed for all the binning configurations. Table 8.2 shows the slopes and offsets obtained, as well as the expected statistical uncertainty for each binning configuration. The expected statistical uncertainty is estimated by applying the unfolding procedure to the sum of the simulated signal and background (called pseudodata) and extracting the uncertainty from the posterior probability density. For the $\Delta|y|$ observable, the binning configuration "fourBin2" gives the best performance. The expected statistical uncertainty increases when the bin sizes are smaller, which is the case for "fourBin", "sixBin", and "eightBin". The slopes follows a similar tendency, increasing when the size of the bins is smaller. The offsets are in all the cases very close to zero. On the other hand, for the $\Delta|\eta|$ observable, the performances for the different configurations are compatible between each other due the good resolution of this observable.

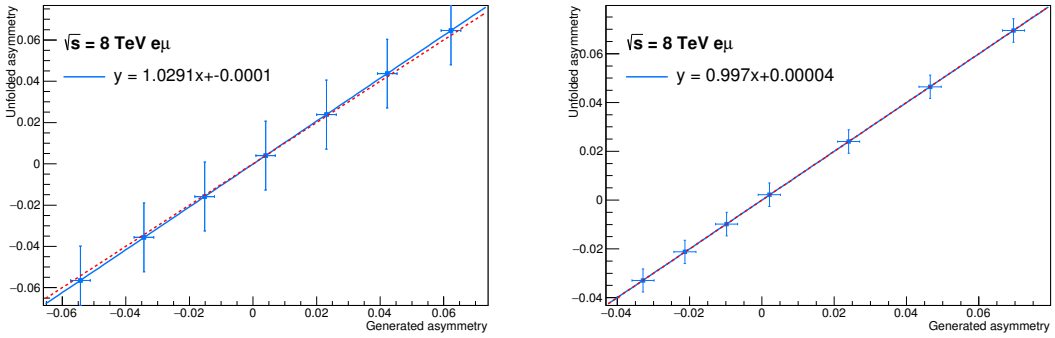


Figure 8.3 – *The calibration curve for the $t\bar{t}$ (left) and lepton (right) asymmetry using the fourBin2 and bin16_ll configuration respectively for the eμ channel in the fiducial region. These calibration curves have been produced using the Protos reweighting function.*

Binning	Expected uncertainty	Slope	Offset
	$\Delta\|y\|$		
fourBin	0.0182	0.981	0.00013
fourBin2	0.0157	1.029	0.00010
fourBin3	0.0151	1.046	-0.00018
sixBin	0.0162	1.065	-0.00028
eightBin	0.0190	1.037	-0.00013
	$\Delta\|\eta\|$		
bin14_ll	0.0053	0.997	0.00000
bin16_ll	0.0056	0.997	0.00004
bin18_ll	0.0056	0.996	-0.00001
bin20_ll	0.0056	0.998	0.00001

Table 8.2 – *Expected statistical uncertainty, slope and offset of the linearity test for the different binning configurations. These results correspond to the eμ channel. The reweighting function corresponding to the Protos generator is used to inject different asymmetry levels for the linearity test. The uncertainty on the slope and offset is approximately* 0.005.

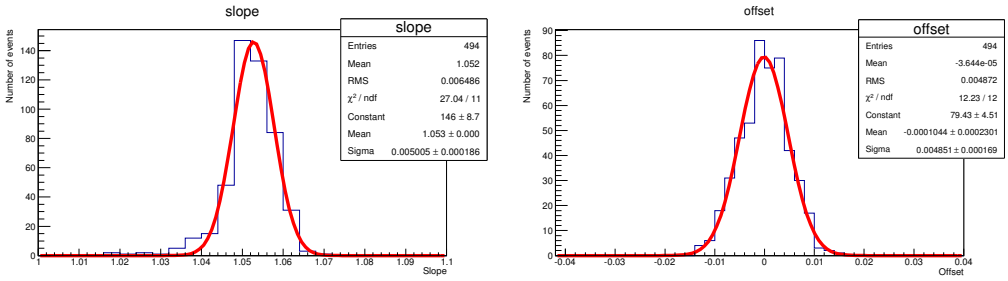


Figure 8.4 – *Slope (top) and offset (bottom) distributions using bootstrap for the uncertainty estimation. The slopes correspond to the calibration curve of $t\bar{t}$ asymmetry in the fiducial region. A Gaussian-fit is applied and its width is taken as the uncertainty.*

The uncertainty on the slope and offset is estimated by performing pseudoexperiments using the bootstrap method in order to keep the correlation between the individual points of the calibration curve. For the calibration procedure, 500 replicas are used. Different levels of asymmetry are injected to each $\Delta|y|$ replica and the unfolding procedure is performed, such that a total of 500 calibration curves are produced. This yields a distribution of slopes and offsets. The uncertainties on the slope and the offset can be estimated from the width of a Gaussian fit applied to these distributions. In Fig. 8.4 the distributions of slope and offset and their respective fits are shown. The width of the fitted Gaussian is found to be approximately 0.005 for both the slope and offset. These uncertainties are independent of the binning configuration, since they only depend on the size of the MC sample. The uncertainty on the slope and offset for $\Delta|\eta|$ is approximately the same as for $\Delta|y|$.

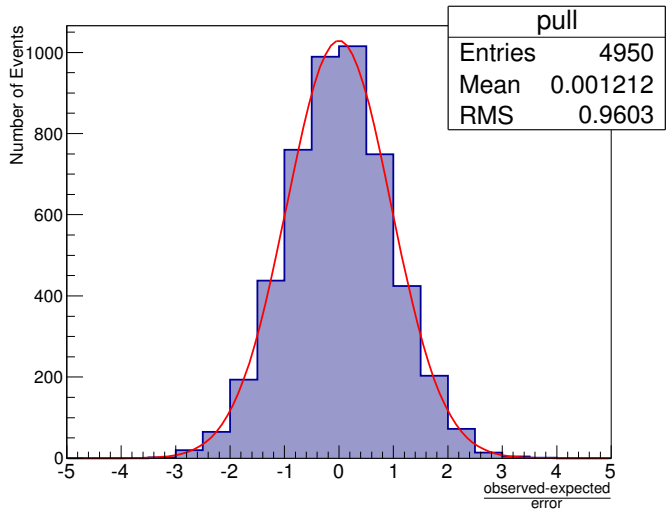


Figure 8.5 – *Pull distribution for the fourBin2 configuration of the $t\bar{t}$ asymmetry in the $e\mu$ channel. The mean is $\sim$0 and the RMS $\sim$1. The red line represents a Gaussian fit, performed for illustration purposes.*

8.1.1. b Pull test

Another important study is to verify that the statistical uncertainty given by the unfolding procedure is compatible with the actual statistical variations. For this test, pull distributions, obtained through ensemble testing, are used. Pulls (p) are defined as follows:

$$p = \frac{A_C^{obs} - A_C^{gen}}{\sigma_{A_C}}, \tag{8.1}$$

where A_C^{obs} is the obtained unfolded asymmetry, A_C^{gen} the generated asymmetry and σ_{A_C} is the statistical error on the unfolded asymmetry. The pull is therefore an estimation of the bias in units of standard deviations on the measured value. Its distribution is expected to be centered at zero (which means that the measurement is not biased), and its width should be one (meaning that the statistical uncertainty is properly estimated). An example of pull distribution is shown in Fig. 8.5. The pull test shows to be compatible with the expected values. The mean and the RMS of the distribution is found to be 0.00 ± 0.03 and 0.96 ± 0.02, for the different four bin configurations of $\Delta|y|$. For $\Delta|\eta|$, the mean and the RMS of the distribution are found to be 0.00 ± 0.01 and 0.97 ± 0.01 for the different binning configurations.

8.1.1. c Complementary tests

In order to complement these studies and select the most optimal configuration for the analysis, the results of the best three binning configurations for the $t\bar{t}$ asymmetry are cross-checked further by using several procedures: a linear reweighting for the linearity test (Tab. 8.3), by unfolding to parton level (Tab. 8.4) and by combining the three decay channels (Tab. 8.5). In the case of the linear reweighting, fourBin2 offers a good balance between the value of the slope and the expected statistical uncertainty. When unfolding at parton level, fourBin3 gives the best per-

formance across the different variables. In combined channels, fourBin2 gives the best slope and the expected statistical uncertainties are quite comparable with the fourBin3 configuration. The offsets across the different tests are very close to zero. The final choice for the analysis is to use fourBin2 for the $t\bar{t}$ asymmetry. For the leptonic asymmetry, the performances of the different tested configurations are very similar. The binning configuration bin16_ll is selected for these measurements, since it was the binning configuration more likely to be used in other experiments[2]. It also gives an improvement in terms of the number of bins with respect to the charge asymmetry analysis performed by ATLAS at 7 TeVwhich was using 14 bins [140]. The final input distribution for the unfolding procedure is shown in Fig 8.6 for the $t\bar{t}$ and leptonic asymmetries. Each of the channels uses the binning configuration "fourBin2" for $\Delta|y|$ and "bin16_ll" for $\Delta|\eta|$.

Binning	Expected uncertainty	Slope	Offset
fourBin	0.0182	1.02	0.0000
fourBin2	0.0165	1.05	-0.0002
fourBin3	0.0157	1.06	-0.0003

Table 8.3 – *Expected statistical uncertainties, slopes and offsets of the linearity test for best three binning configurations in $\Delta|y|$. These results correspond to the eμ channel and unfolding to particle level. A linear reweighting was used to inject different asymmetry levels for the linearity test. The uncertainty on the slope and offset is approximately* 0.005.

Binning	Expected Uncertainty	Slope	Offset
fourBin	0.0179	0.92	0.00038
fourBin2	0.0160	0.95	0.00028
fourBin3	0.0151	0.97	0.00023

Table 8.4 – *Expected statistical uncertainties, slopes, and offsets of the linearity test for the best three binning configurations in $\Delta|y|$. These results correspond to the eμ channel and unfolding to parton level. A protos reweighting was used to inject different asymmetry levels for the linearity test. The uncertainty on the slope and offset is approximately* 0.005.

8.1.1. d Regularization parameters

Different regularization parameters (α) are tested for the fourBin2 configuration. The α parameter is chosen to be in the range: $[-7,-11]$ using the Tikhonov regularization described in Sec. 6.2. An α parameter equal to zero corresponds to a flat prior. The expected statistical uncertainty and the slopes and offsets of the calibration curve are used to evaluate the impact on the measurement. Table 8.6 shows the results of this test. When the regularization parameter gets stronger (i.e towards -7) the statistical uncertainty is reduced as expected, however, the bias which is evaluated

[2] By using the same numbers of bins in the distribution a combination of the measurements is possible

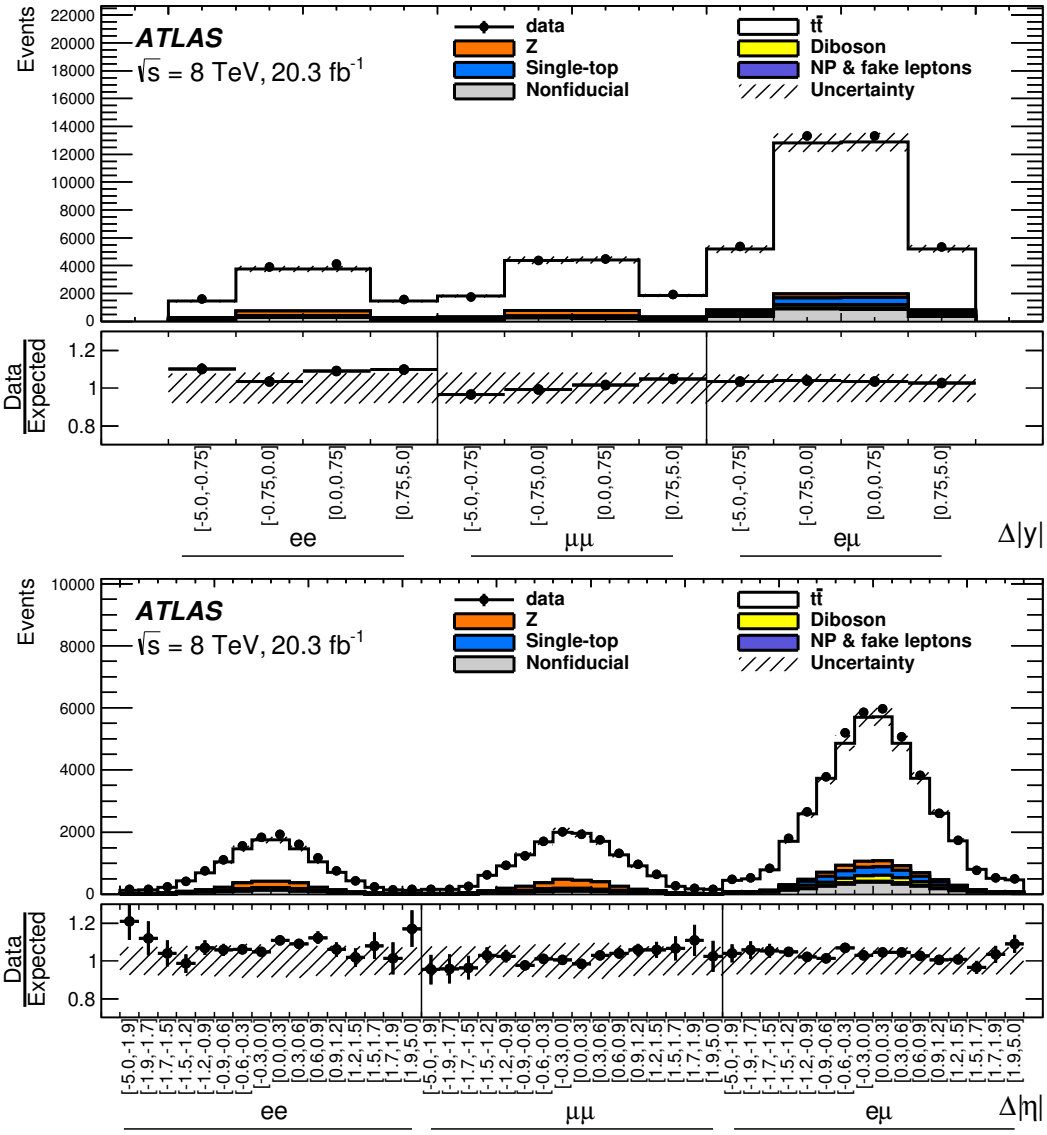


Figure 8.6 – *Input distributions for the inclusive $t\bar{t}$ (top) and leptonic (bottom) asymmetry measurements. The first 4 $\Delta|y|$ and 16 $\Delta|\eta|$ bins correspond to the ee channel, followed by $\mu\mu$ and $e\mu$. The data/expected ratio is also shown.*

Binning	Expected uncertainty	Slope	Offset
fourBin	0.0156	0.98	-0.0004
fourBin2	0.0140	1.01	0.0005
fourBin3	0.0132	1.03	0.0004

Table 8.5 – *Expected statistical uncertainties, slopes, and offsets of the linearity test for the best three binning configurations in*$\Delta|y|$*. These results correspond to the three channels combined. A protos reweighting was used to inject different asymmetry levels for the linearity test. The uncertainty on the slope and offset is approximately* 0.004.

through the slope and offset start to increase. At $\alpha = -8$ the statistical uncertainty starts be to significantly reduced but the bias increases drastically. Values above -7 are not considered since the posterior probability density shows very unstable behaviors, making it impossible to reproduce the generated asymmetry. Since the cost of reducing significantly the statistical uncertainty is too high with respect to the bias introduced by the unfolding procedure, the flat prior is used for all measurements.

α parameter	Expected uncertainty	Slope	Offset
0	0.0157	1.03	-0.0004
-11	0.0149	1.05	-0.0004
-10	0.0133	1.07	-0.0005
-9	0.0128	1.07	-0.0010
-8	0.0097	1.08	-0.0063
-7	0.0072	1.10	-0.0110

Table 8.6 – *Expected statistical uncertainty, slope, and offset from the calibration curve for different values of the regularization parameter.*

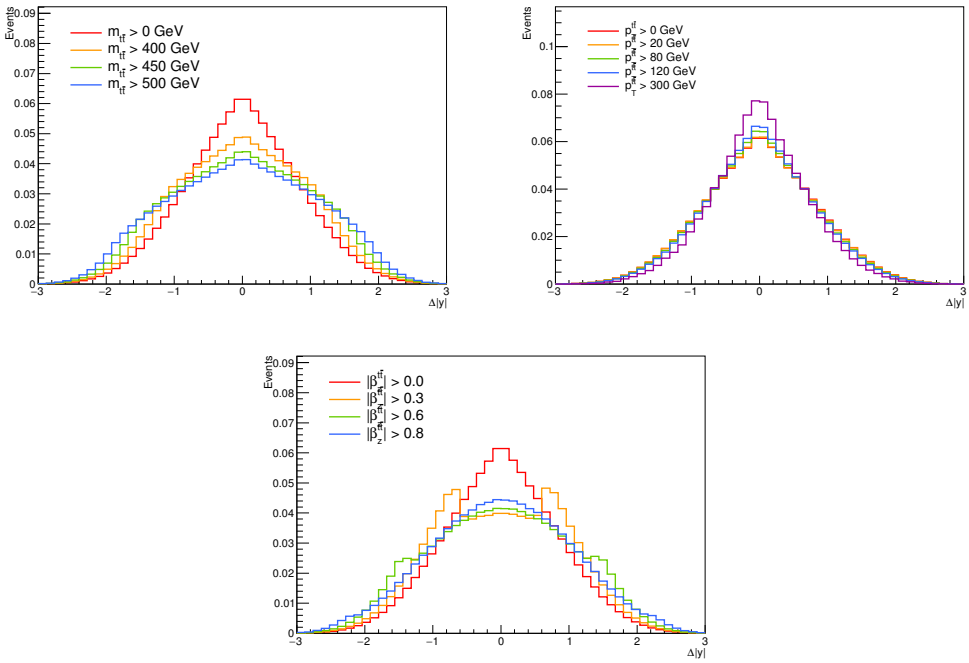


Figure 8.7 – *The distributions of $\Delta|y|$ as a function of $m_{t\bar{t}}$, $\beta_{z,t\bar{t}}$, and $p_{T,t\bar{t}}$. At the top, the $\Delta|y|$ distributions for 6 $m_{t\bar{t}}$ bins in the ranges [0, 400, 450, 500] GeV are shown. In the middle, the $\Delta|y|$ distributions for 5 $\beta_{z,t\bar{t}}$ bins in the ranges [0, 0.1, 0.3, 0.6, 0.8, 1.0] are shown. At the bottom, the $\Delta|y|$ distributions for 5 $p_{T,t\bar{t}}$ bins in the ranges [0, 20, 50, 80, 120, 300] GeV are shown. All the distributions are normalized to the same area.*

8.1.2 Differential measurements

Due to the expected statistical uncertainty, only two bins for each differential variables $m_{t\bar{t}}$, $\beta_{z,t\bar{t}}$ and $p_{\mathrm{T},t\bar{t}}$ are considered. Using more than two bins would lead to very large statistical uncertainties, making the measurements insensitive to the SM and BSM predictions. The major parameter that needs to be optimized is where to set the boundary between the first and second differential bins. However, the configuration of the bins in the $\Delta|y|$ and $\Delta|\eta|$ observables also needs to be taken into account since their shapes depend on the $m_{t\bar{t}}$, $p_{\mathrm{T},t\bar{t}}$, and $\beta_{z,t\bar{t}}$ as shown in Fig. 8.7. The $\Delta|y|$ distribution is broader for high $m_{t\bar{t}}$ and $\beta_{z,t\bar{t}}$ values, while it is slightly narrower for high $p_{\mathrm{T},t\bar{t}}$. The number of bins is chosen to be of four for both observables in order to reduce the expected statistical uncertainty. The main criteria for selecting the best configuration is the expected statistical uncertainty. The linearity test is performed on the best candidate (based on the previous criteria) as a cross check.

8.1.2. a Optimization for $m_{t\bar{t}}$

The boundary between the first and second $m_{t\bar{t}}$ bin is scanned in the range from 400 GeV up to 600 GeV. Four bins are used for the $\Delta|y|$ and $\Delta|\eta|$ distributions, where the size of the inner bins are optimized by scanning them from 0.4 up to 1.4 in steps of 0.1.

The configuration in the range $[0, 500, 2000]$ GeV gives around the same expected statistical

	$\Delta\|y\|$		$\Delta\|\eta\|$	
$m_{t\bar{t}}$ binning	σ (1st $m_{t\bar{t}}$ bin)	σ (2nd $m_{t\bar{t}}$ bin)	σ (1st $m_{t\bar{t}}$ bin)	σ (2nd $m_{t\bar{t}}$ bin)
$[0, 400, 2000]$	0.125	0.021	0.025	0.011
$[0, 450, 2000]$	0.061	0.022	0.015	0.013
$[0, 500, 2000]$	0.035	0.026	0.011	0.016
$[0, 550, 2000]$	0.032	0.030	0.010	0.020
$[0, 600, 2000]$	0.028	0.037	0.009	0.026

Table 8.7 – *The expected statistical uncertainty of $t\bar{t}$ charge asymmetry and the leptonic asymmetry (in $e\mu$ channel) for different $m_{t\bar{t}}$ bins for various binning configurations.*

uncertainty in both differential bins for $\Delta|y|$ and $\Delta|\eta|$ with slightly better performance in the second bin, where the asymmetry value is enhanced in the SM, as well as in BSM models. This configuration is considered optimal and it is used for the measurements. The optimal binning configurations for the $\Delta|y|$ and $\Delta|\eta|$ distribution are found to be

- $m_{t\bar{t}} < 500$ GeV, $\Delta|y| = [-5, -0.6, 0, 0.6, 5]$, $\Delta|\eta| = [-5, -0.8, 0, 0.8, 5]$
- $m_{t\bar{t}} > 500$ GeV, $\Delta|y| = [-5, -0.7, 0, 0.7, 5]$, $\Delta|\eta| = [-5, -1.4, 0, 1.4, 5]$

The larger size of the inner bins in the second differential bin is expected since the $\Delta|y|$ and $\Delta|\eta|$ distributions get broader for larger $m_{t\bar{t}}$. The calibration curves corresponding to the optimal case are shown in Fig. 8.8 for both the first and second $m_{t\bar{t}}$ bins. The slopes are close to 1.0 and the offsets are very small.

8.1.2. b Optimization for $\beta_{z,t\bar{t}}$

Different BSM models predict significantly different charge asymmetries for $\beta_{z,t\bar{t}}$ greater than 0.6, since a more energetic $q\bar{q}$ initial state is inferred. Therefore, the binning $[0.0, 0.6, 1.0]$ is used for the measurement and its optimization is just performed for the binning used in the $\Delta|y|$ and $\Delta|\eta|$ distributions. The optimal configuration is found to be:

- $\beta_{z,t\bar{t}} < 0.6$ GeV, $\Delta|y| = [-5, -0.5, 0, 0.5, 5]$, $\Delta|\eta| = [-5, -0.8, 0, 0.8, 5]$
- $\beta_{z,t\bar{t}} > 0.6$ GeV, $\Delta|y| = [-5, -0.7, 0, 0.7, 5]$, $\Delta|\eta| = [-5, -1.2, 0, 1.2, 5]$

As in the case for $m_{t\bar{t}}$, the larger size of the bins in the second differential bin is expected, since the $\Delta|y|$ and $\Delta|\eta|$ distributions get broader for larger $\beta_{z,t\bar{t}}$. The expected statistical uncertainties of the $t\bar{t}$ charge asymmetry for the optimal configuration are about 2.5% and 2.3% in the first and second $\beta_{z,t\bar{t}}$ bins. For the leptonic asymmetry, the expected statistical uncertainties are about 0.8% and 1.0% in the first and second differential bin. The calibration curves corresponding to the optimal cases are shown in Fig. 8.9 for both the first and second $\beta_{z,t\bar{t}}$ bins. The slopes are close to 1.0 and the offsets are close to zero.

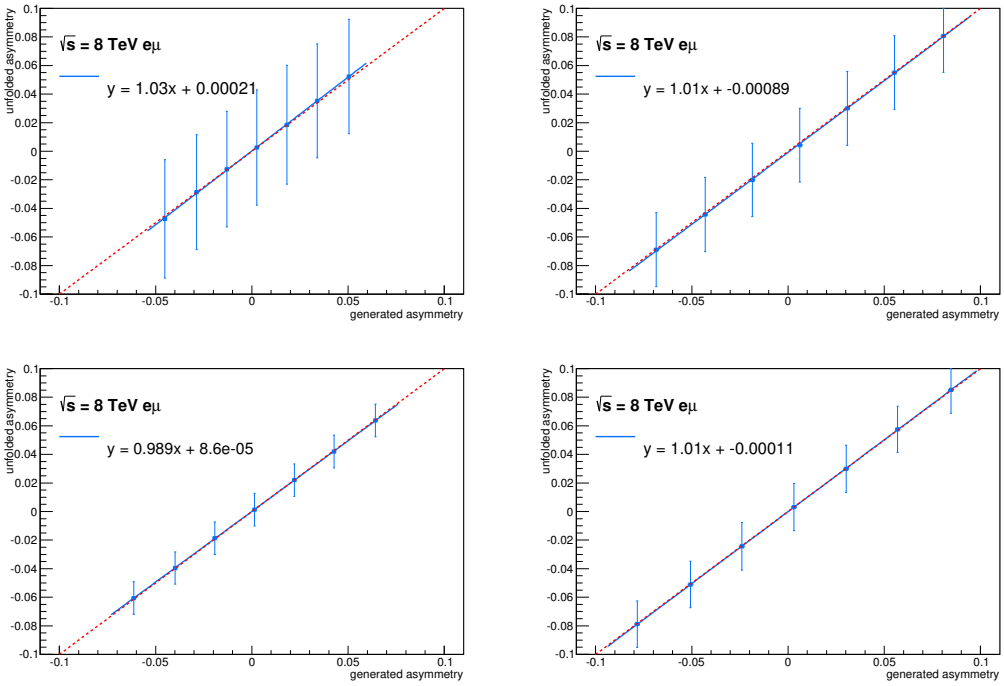


Figure 8.8 – *The calibration curve corresponding to the first $m_{t\bar{t}}$ bin ($m_{t\bar{t}} < 500$ GeV) on the left and to the second $m_{t\bar{t}}$ bin ($m_{t\bar{t}} > 500$ GeV) on right, for the optimized $\Delta|y|$ (top) and $\Delta|\eta|$ (bottom) binning.*

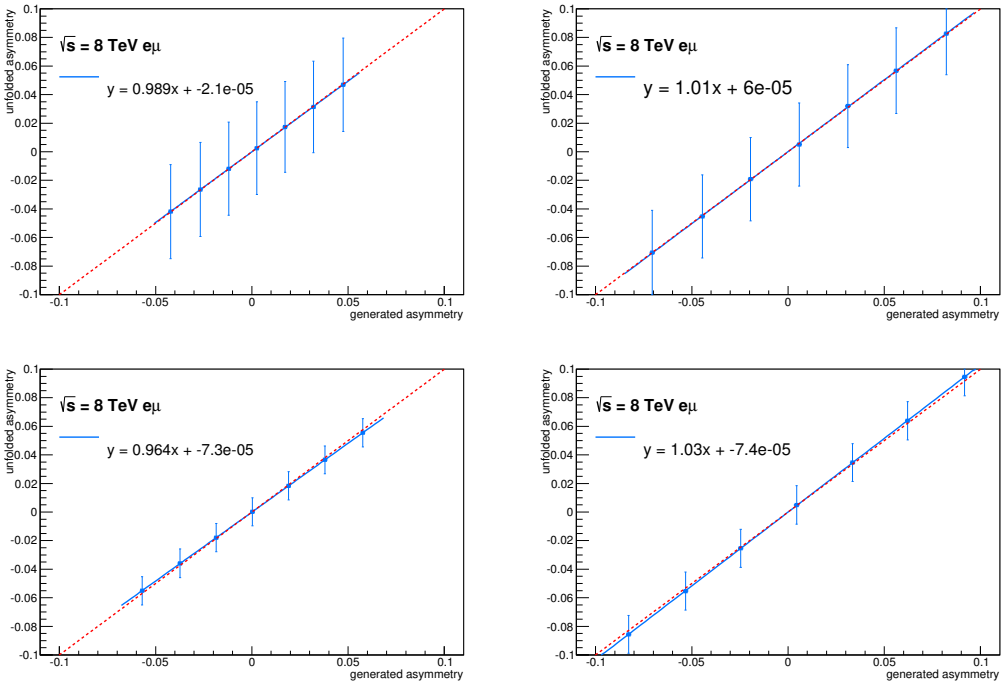


Figure 8.9 – *The calibration curves corresponding to the first (left) and second (right) $\beta_{z,t\bar{t}}$ bin for the optimized $\Delta|y|$ (top) and $\Delta|\eta|$ (bottom) binnings.*

8.1.2. c Optimization for $p_{T,t\bar{t}}$

In the SM, the charge asymmetry is expected to be large for low $p_{T,t\bar{t}}$ and small and roughly constant for higher $p_{T,t\bar{t}}$ due to the different amount of interference provided by the ISR and FSR diagrams and the Born and Box diagrams. At high $p_{T,t\bar{t}}$ the ISR/FSR contribution (which is negative) is dominant. The boundary at $p_{T,t\bar{t}} = 30$ GeV is used since it gives a good balance between the expected effect and the size of the statistical uncertainty. Boundaries at 20 or 25 GeV show a very high statistical uncertainty due to poor statistics in the first bin. Since the boundary is selected to be at $p_{T,t\bar{t}} = 30$ GeV, the optimization for this differential variable is performed only on the $\Delta|y|$ and $\Delta|\eta|$ binnings. The optimal configuration is found to be $[-5, -0.7, 0, 0.7, 5]$ for both distributions and both differential bins. In opposition to the previous two variables, the shape of the $\Delta|y|$ and $\Delta|\eta|$ distributions do not change significantly for different $p_{T,t\bar{t}}$ requirements.

The expected statistical uncertainties on the charge asymmetry for the selected configuration are about 3.8% and 2.9% in the first and second $p_{T,t\bar{t}}$ bins. For the leptonic charge asymmetry the expected statistical uncertainties are about 1.6% and 1.1% in the first and second $p_{T,t\bar{t}}$ bin. The calibration curves corresponding to the optimal case are shown in Fig. 8.10 for both the first and second $m_{t\bar{t}}$ bins. The slopes are close to 1.0 and the offsets are close to zero.

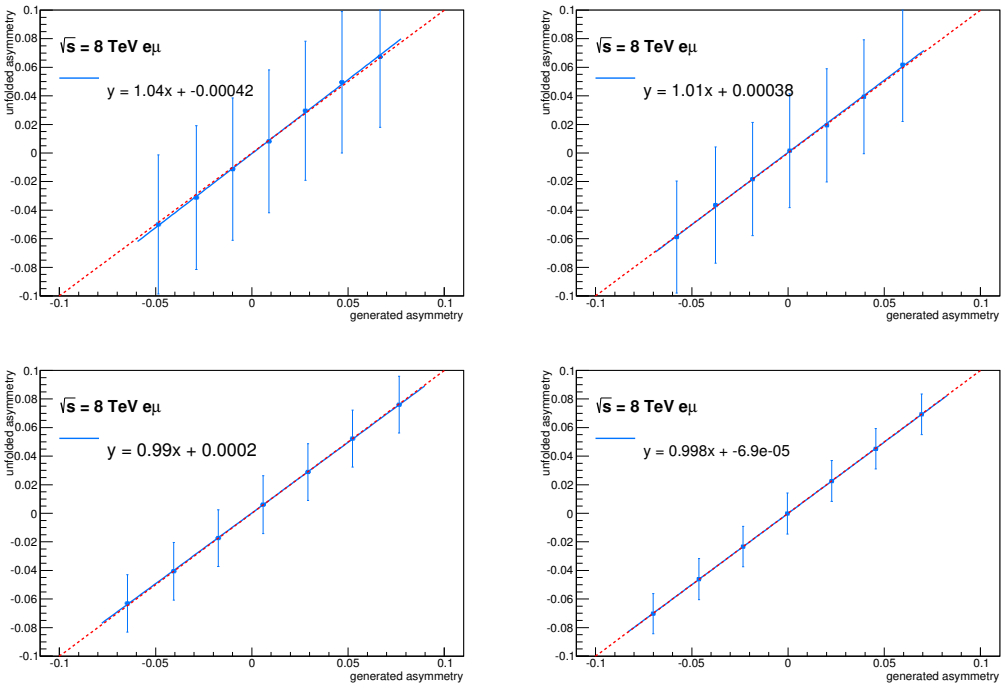


Figure 8.10 – *The calibration curves corresponding to the first (left) and second (right) $p_{T,t\bar{t}}$ bin for the optimized $\Delta|y|$ (top) and $\Delta|\eta|$ (bottom) binnings.*

	$A_C^{t\bar{t}}$		$A_C^{\ell\ell}$	
Source of systematic uncertainty	Fiducial	Full phase space	Fiducial	Full phase space
Missing transverse momentum	0.00187	0.00170	0.00005	0.00011
Lepton energy scale and resolution	0.00146	0.00151	0.00011	0.00011
Lepton reconstruction/identification	0.00009	0.00012	0.00016	0.00014
Jet reconstruction efficiency	0.00225	0.00207	0.00011	0.00002
B-tagging/mis-tag efficiency	0.00050	0.00052	0.00028	0.00027
Jet energy scale and resolution	0.00745	0.00533	0.00044	0.00041
Total Uncertainty	0.00815	0.00618	0.00057	0.00053

Table 8.8 – *Systematic uncertainty estimation for $\Delta|y|$ and $\Delta|\eta|$ for all detector modeling systematic sources for the inclusive measurements in the fiducial and full phase space.*

8.2 Systematic uncertainties

In the following, the estimation of the systematic uncertainties are presented based on the methods explained in Chapter 7.

Detector modeling

Table 8.8 shows the estimation for the different sources of the detector systematic uncertainties in the fiducial region and full phase space for the inclusive $A_C^{t\bar{t}}$ and $A_C^{\ell\ell}$ measurement using the method described in Sec. 7.1.3. The main contribution comes from the jet energy scale. The contribution coming from the systematic uncertainties that are constrained using the marginalization procedure (b-tagging and electron identification) are not the dominant uncertainties. The uncertainties in the fiducial region in some cases are slightly larger than in the full phase space. This effect is coming from statistical fluctuations introduced by the nonfiducial background. Both observations are true for all the differential measurements of the charge asymmetry (see Appendix C).

Signal modeling

In Tab. 8.9 the different contributions of each modeling uncertainty are shown for $t\bar{t}$ and leptonic asymmetry. The MC statistical uncertainty on the estimation is also shown. The actual uncertainty is taken from the maximum value between the estimation and its MC statistical uncertainty. The PDF uncertainties do not have a MC statistical component since their estimation is performed by reweighting the nominal signal sample. The uncertainties corresponding to the color reconnection and underlying event are not shown since their contribution is negligible with respect to the other contributions. As expected, there is a general reduction of the modeling uncertainties in the fiducial volume where the unfolding procedure brings the measurement back to particle level. In the leptonic asymmetry case, the uncertainties are in general smaller since the measurement do not involve a kinematic reconstruction and the charged leptons are defined

at stable particle level, thus their difference between different MC simulation is small. The systematic contributions for all the differential measurements can be found in Appendix D.

	$A_C^{t\bar{t}}$		$A_C^{\ell\ell}$	
Source of systematic uncertainty	Fiducial	Full phase space	Fiducial	Full phase space
MC generator	0.0040 ± 0.0058	0.0065 ± 0.0072	0.0028 ± 0.0018	0.0017 ± 0.0015
PDF	0.0026	0.0022	0.0001	0.0003
Parton shower	0.0029 ± 0.0022	0.0030 ± 0.0020	0.0004 ± 0.0009	0.0017 ± 0.0012
ISR/FSR	0.0020 ± 0.0035	0.0027 ± 0.0035	0.0001 ± 0.0013	0.0012 ± 0.0009
Total	0.0074	0.0085	0.0023	0.0035

Table 8.9 – *Modeling systematic uncertainty contribution of the different sources affecting the inclusive $t\bar{t}$ and leptonic asymmetry.*

The uncertainty on the top quark mass could have an effect on the measurement, since a modification on the top mass could lead to a different invariant mass of the top pair system (affecting the asymmetry value). The dependency of the asymmetry as function of the top mass used in the simulation is studied. A linear fit is performed to the asymmetry values obtained by unfolding MC generators with different top masses, specifically: 167.5 GeV, 170.0 GeV, 175.0 GeV and 177.5 GeV. Figure 8.11 shows the $t\bar{t}$ asymmetry value as a function of different top masses in the fiducial volume. The dependency is evaluated through the slope obtained from the linear fit. The dependency is compatible with zero and thus the uncertainty coming from top mass is not taken into account in the final measurement. A similar behavior is observed across all the measurements.

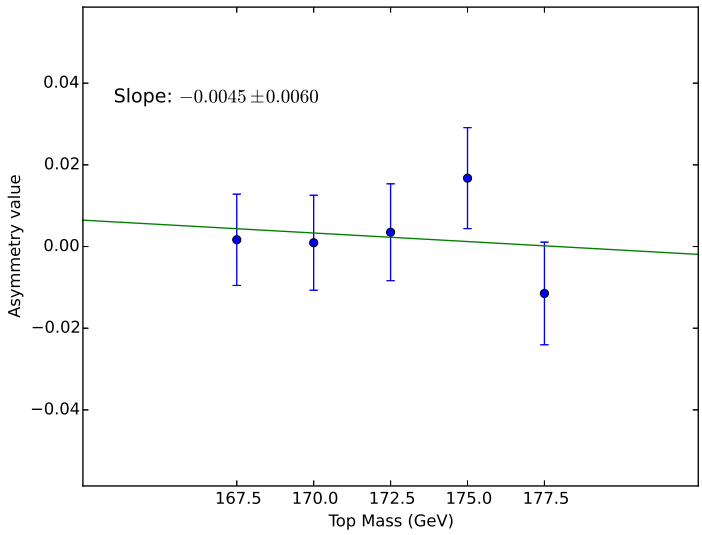


Figure 8.11 – *$t\bar{t}$ asymmetry as a function of different top masses in simulations in the fiducial volume. No dependency is observed*

8.2.1 Other uncertainties

These uncertainties correspond to the bias introduced by the unfolding procedure (nonclosure test) and the impact from the MC statistical uncertainty. The uncertainty related with the KIN method and that only affects the charge asymmetry measurement is also discussed.

8.2.1. a Top-quark kinematic reconstruction

Due to the stopping criteria in the kinematic reconstruction method, it is possible to have different solutions when the smearing of the jets and missing transverse momentum start from a different point of the phase space. To assess this effect, pseudoexperiments are performed by varying the starting point of the smearing procedure. A $\Delta|y|$ distribution is obtained for each variation and the spectrum of each bin is used to compute the uncertainty associated with the reconstruction. Figure 8.12 shows the distribution of the number of events for the second bin of the $\Delta|y|$ distribution in the $e\mu$ channel. The RMS of that distribution is taken as the reconstruction uncertainty for the bin. This uncertainty corresponds to approximately half of the statistical uncertainty.

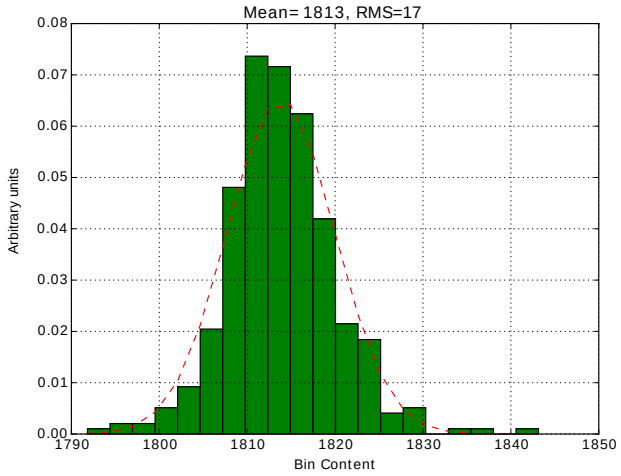


Figure 8.12 – *Distribution of the number of events for the second bin of the $\Delta|y|$ distribution in the $e\mu$ channel after varying the starting point of the smearing procedure. The distribution is normalized to one. A Gaussian fit is performed for illustration purposes.*

Once the uncertainty for each bin is estimated, pseudoexperiments are performed by fluctuating the input distribution of the unfolding procedure within the estimated reconstruction uncertainties. For each fluctuation, the unfolding procedure is repeated and a new asymmetry value is obtained. The correlation between the bins is preserved in the procedure. The correlation matrix is shown in Fig. 8.13. The correlations come from the migrations between the different bins when a different solution is found during the kinematic reconstruction. Figure 8.14 shows the distribution of asymmetry values for the inclusive measurement in the full phase space obtained after 2000 pseudoexperiments. The mean of the distribution corresponds to the expected asymmetry

	$A_C^{t\bar{t}}$	$m_{t\bar{t}}$ [GeV]		$p_T^{t\bar{t}}$ [GeV]		$\beta_z^{t\bar{t}}$	
	Inclusive	0 – 500	500 – 2000	0 – 30	30 – 1000	0.0 – 0.6	0.6 – 1.0
Reconstruction Uncertainty	0.0052	0.0184	0.0068	0.0154	0.0138	0.0165	0.0086
Statistical Uncertainty	0.0110	0.0280	0.0149	0.0305	0.0244	0.0226	0.0179
Reco + Statistical	0.0122	0.0335	0.0164	0.0342	0.0281	0.0279	0.0199

Table 8.10 – *Reconstruction uncertainty in comparison with the statistical uncertainty for the inclusive and differential top asymmetry measurements in the full phase space.*

value from the nominal generator, whereas the RMS corresponds to the uncertainty on the asymmetry related with the reconstruction procedure. Tables 8.10 and 8.11 show the reconstruction uncertainties for the inclusive and differential measurements of the $t\bar{t}$ and leptonic asymmetries, respectively. In most of the cases, the uncertainty is half of the statistical uncertainty.

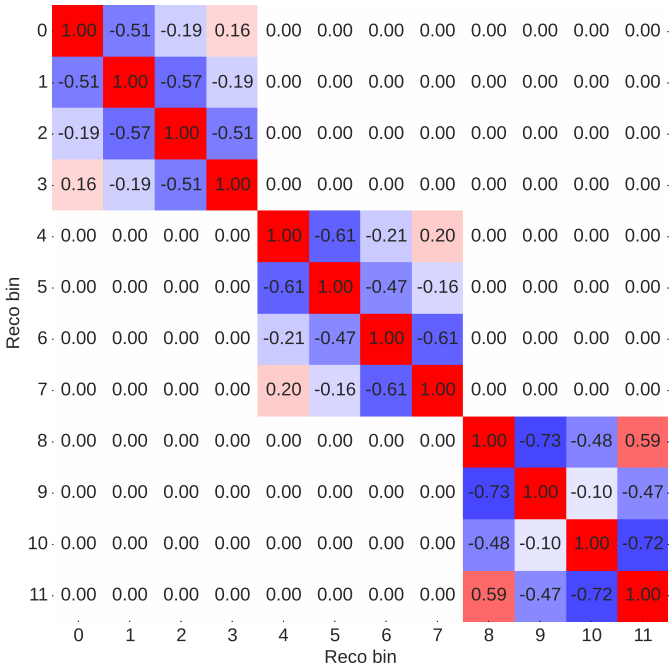


Figure 8.13 – *Correlations between the bins of the $\Delta|y|$ distribution due to the kinematic reconstruction.*

8.2.1. b Nonclosure uncertainties

In Tab. 8.12, the uncertainties for each of the measurements are shown for the $t\bar{t}$ and leptonic asymmetries in the fiducial region and in the full phase space. These uncertainties are added in

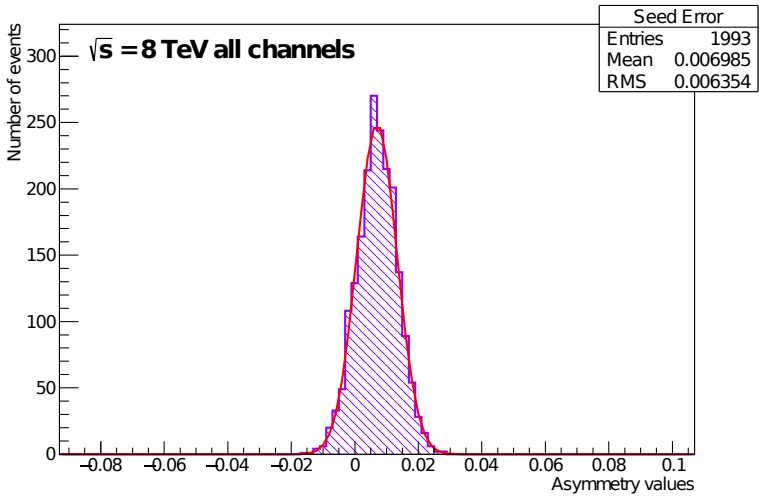


Figure 8.14 – *Distribution of the $t\bar{t}$ asymmetry values for the inclusive measurement in the full phase space obtained after 2000 pseudoexperiments. The RMS of this distribution is taken as the uncertainty related with the reconstruction procedure.*

	$A_C^{\ell\ell}$	$m_{t\bar{t}}$ [GeV]		$p_T^{t\bar{t}}$ [GeV]		$\beta_z^{t\bar{t}}$	
	Inclusive	0 – 500	500 – 2000	0 – 30	30 – 1000	0.0 – 0.6	0.6 – 1.0
Reconstruction Uncertainty	0.0000	0.0049	0.0039	0.0053	0.0044	0.0031	0.0029
Statistical Uncertainty	0.0047	0.0081	0.0114	0.0148	0.0105	0.0072	0.0096
Reco + Statistical	0.0047	0.0095	0.0121	0.0157	0.0114	0.0078	0.0100

Table 8.11 – *Reconstruction uncertainty in comparison with the statistical uncertainty for the inclusive and differential lepton asymmetry measurements in the full phase space.*

quadrature to the total uncertainty and their effect is considered minor.

8.2.1. c MC statistical uncertainty

The uncertainty related with the size of the MC samples in the inclusive and differential measurements of $A_C^{\ell\ell}$ and $A_C^{t\bar{t}}$ is shown in Tab. 8.13. These uncertainties are added in quadrature to the total uncertainties. The impact on the measurements is small.

		$A_C^{t\bar{t}}$		$A_C^{\ell\ell}$	
		Fiducial Region	Full phase space	Fiducial Region	Full phase space
Inclusive		0.0004	0.0006	< 0.0001	0.0001
$M_{t\bar{t}}$	0-500 GeV	0.0012	0.0002	0.0004	0.0026
	500-2000 GeV	0.0004	0.0004	0.0004	0.0014
$\beta_{t\bar{t}}$	0.0-0.6	0.0005	0.0001	0.0006	0.0027
	0.6 - 1.0	0.0001	0.0004	0.0005	0.0013
$p_T^{t\bar{t}}$	0-30 GeV	0.0010	0.0004	< 0.0001	0.0024
	500-1000 GeV	0.0024	< 0.0001	0.0001	0.0027

Table 8.12 – *Non-closure uncertainties for each of the measurements of the $t\bar{t}$ and leptonic asymmetries in the fiducial region and in the full phase space. This contribution is negligible in comparison with the statistical uncertainty.*

		$A_C^{t\bar{t}}$		$A_C^{\ell\ell}$	
		Fiducial Region	Full phase space	Fiducial Region	Full phase space
Inclusive		0.0026	0.0026	0.0007	0.0010
$m_{t\bar{t}}$	0-500 GeV	0.0064	0.0069	0.0016	0.0020
	500-2000 GeV	0.0036	0.0035	0.0024	0.0026
$\beta_{t\bar{t}}$	0.0 - 0.6	0.0049	0.0054	0.0015	0.0018
	0.6 - 1.0	0.0041	0.0040	0.0019	0.0022
$p_T^{t\bar{t}}$	0-30 GeV	0.0076	0.0076	0.0031	0.0035
	500-1000 GeV	0.0058	0.0057	0.0021	0.0024

Table 8.13 – *MC sample size uncertainties for each of the measurements of the $t\bar{t}$ and leptonic asymmetries in the fiducial region and in the full phase space. This contribution is negligible in comparison with the total uncertainty on the measurements.*

8.2.1. d Summary of systematic uncertainties

Tables 8.14 and 8.15 summarize how each category of uncertainty affects the measurements of the leptonic and $t\bar{t}$ asymmetries, respectively. The statistical uncertainty corresponds to the largest contribution, followed by the reconstruction and the signal modeling uncertainties. The signal modeling uncertainties are larger in the differential measurements because of the migrations between the differential bins in the different MC generators used for their estimation. The uncertainty obtained by the sum in quadrature of the individual systematic uncertainties is slightly larger than the total marginalized uncertainty on the measurements. In general there is a reduction of the signal modeling uncertainties in the fiducial volume, however, this is not the case is some measurements where the uncertainties are dominated by the MC statistical effects.

Table 8.14 – *Impact of each source of uncertainty uncertainties on the leptonic asymmetry in the fiducial and full phase space.*

		Absolute uncertainties in $A_C^{\ell\ell}$									
		Fiducial volume					Full phase space				
		Statistics	Detector	Bkg	Signal modeling	Other	Statistics	Detector	Bkg	Signal modeling	Other
	Inclusive	0.005	0.001	0.001	0.002	0.001	0.005	0.001	0.001	0.004	0.001
$m_{t\bar{t}}$	0–500 GeV	0.008	0.002	0.001	0.005	0.005	0.008	0.002	0.001	0.005	0.006
	500–2000 GeV	0.012	0.004	< 0.001	0.013	0.005	0.011	0.004	< 0.001	0.014	0.005
$\beta_{t\bar{t}}$	0–0.6	0.007	0.003	< 0.001	0.004	0.004	0.007	0.002	< 0.001	0.005	0.005
	0.6–1.0	0.010	0.005	0.001	0.005	0.004	0.010	0.003	0.001	0.006	0.004
$p_T^{t\bar{t}}$	0–30 GeV	0.015	0.009	0.001	0.015	0.006	0.015	0.010	0.001	0.017	0.007
	30–1000 GeV	0.011	0.004	0.001	0.012	0.005	0.010	0.004	0.001	0.013	0.006

Table 8.15 – *Impact of each source of uncertainty uncertainties on the $t\bar{t}$ asymmetry in the fiducial and full phase space.*

		Absolute uncertainties in $A_C^{t\bar{t}}$									
		Fiducial volume					Full phase space				
		Statistics	Detector	Bkg	Signal modeling	Other	Statistics	Detector	Bkg	Signal modeling	Other
	Inclusive	0.013	0.008	< 0.001	0.007	0.007	0.011	0.006	< 0.001	0.008	0.006
$m_{t\bar{t}}$	0–500 GeV	0.030	0.024	0.001	0.016	0.021	0.028	0.021	0.002	0.018	0.020
	500–2000 GeV	0.018	0.007	< 0.001	0.015	0.009	0.015	0.006	< 0.001	0.016	0.008
$\beta_{t\bar{t}}$	0–0.6	0.023	0.021	0.002	0.014	0.018	0.023	0.019	0.002	0.015	0.017
	0.6–1.0	0.021	0.009	0.001	0.013	0.011	0.018	0.009	0.001	0.013	0.010
$p_T^{t\bar{t}}$	0–30 GeV	0.035	0.019	0.003	0.018	0.020	0.031	0.015	0.004	0.019	0.017
	30–1000 GeV	0.027	0.015	0.003	0.018	0.017	0.025	0.013	0.003	0.014	0.015

8.3 Results

8.3.1 Unfolded measurements

After the $t\bar{t}$ reconstruction is applied on data, the unfolding procedure is performed. The output of the unfolding procedure is a posterior probability density, where the mean of the distribution corresponds to the measured value and the RMS to the statistical uncertainty, detector modeling, and background systematic uncertainty. The signal modeling uncertainty and the other sources of uncertainties are added in quadrature to the uncertainty. Figures 8.15 and 8.16 show the inclusive and differential results for the leptonic and $t\bar{t}$ charge asymmetry in the fiducial region and in the full phase space. They are compared with the SM expectations [48]. The measured inclusive values in the full phase space are:

$$A_{\mathrm{C}}^{\ell\ell} = 0.008 \pm 0.006, \quad A_{\mathrm{C}}^{t\bar{t}} = 0.021 \pm 0.016. \tag{8.2}$$

They are in agreement with the SM predictions of $A_{\mathrm{C}}^{\ell\ell} = 0.0064 \pm 0.0003$ and $A_{\mathrm{C}}^{t\bar{t}} = 0.0111 \pm 0.0004$ [48]. Figure 8.17 shows the unfolded distributions of the $\Delta|\eta|$ and $\Delta|y|$ observables for the inclusive measurement in the fiducial volume. The distributions are compared with MC predictions at NLO obtained from Powheg-hvq and a good agreement is observed. The correlations between the different bins of the unfolded $\Delta|y|$ and $\Delta|\eta|$ distributions are shown in Fig. 8.18 and 8.20, respectively. The correlation factors between the bins are estimated by taking the normalized covariance of all possible truth values scanned during the unfolding procedure for each bin. Figure 8.19 shows the different scanned points for the second and third bin of the $\Delta|y|$ distribution. These points are used to estimate the correlation between the bins. The statistical and detector systematic uncertainties are taken into account. Neighboring bins are strongly anti-correlated for the unfolded $\Delta|y|$ distribution ($\sim -70\%$). The anticorrelation can be interpreted as follows: when the number of events in one bin increases, the number of events in neighbor bin decreases in order to account for the migrations between the different bins. For the unfolded $\Delta|\eta|$ distribution, the correlation between the bins is negligible since the migration matrix is very diagonal. For the differential measurements, all the unfolded values are compatible with the SM predictions. The largest discrepancy is around 1.2σ and it is present in the high $p_{\mathrm{T},t\bar{t}}$ bin of $A_{\mathrm{C}}^{t\bar{t}}$ in the full phase space.

The statistical uncertainty is the dominant contribution to the total uncertainty in most of the measurements. The dominant systematic uncertainties across all the measurements are the signal modeling and the kinematic reconstruction uncertainties. The signal modeling uncertainties are reduced by performing the measurements in the fiducial region, since the extrapolation from detector acceptance to the full phase space is avoided. The statistical uncertainty is slightly larger in the fiducial region than in the full phase space. This is expected due to the nonfiducial background considered in the fiducial measurements, which is substracted from the data in the unfolding procedure, increasing the size of the statistical uncertainty on the measurements. The behavior of the nuisance parameters are similar between all measurements. Examples for the inclusive measurement of the lepton and $t\bar{t}$ asymmetries in the fiducial region and in the full phase space are shown in Fig. 8.21 and 8.22. The behavior of the nuisance parameters are consistent

between the measurements performed in the fiducial volume and in the full phase space. The nuisance parameters are pulled in the same direction and small fluctuations are present. The constraint on the nuisance parameter corresponding to the electron identification scale factor is the largest and it is around $+1\sigma$. This means that data is more compatible with the $+1\sigma$ electron identification scale factor variation than with the nominal MC sample. A the data/MC comparison shown in Fig. 8.23 in order to understand this behavior. The red points in the ratio plot represent the $+1\sigma$ electron identification scale factor variation, while the black points represent the nominal MC sample. The electron identification scale factor variation describes better the data than the nominal MC sample. The correlation between the nuisance parameters were checked after the unfolding procedure and they are found to be mostly uncorrelated ($< 20\%$).

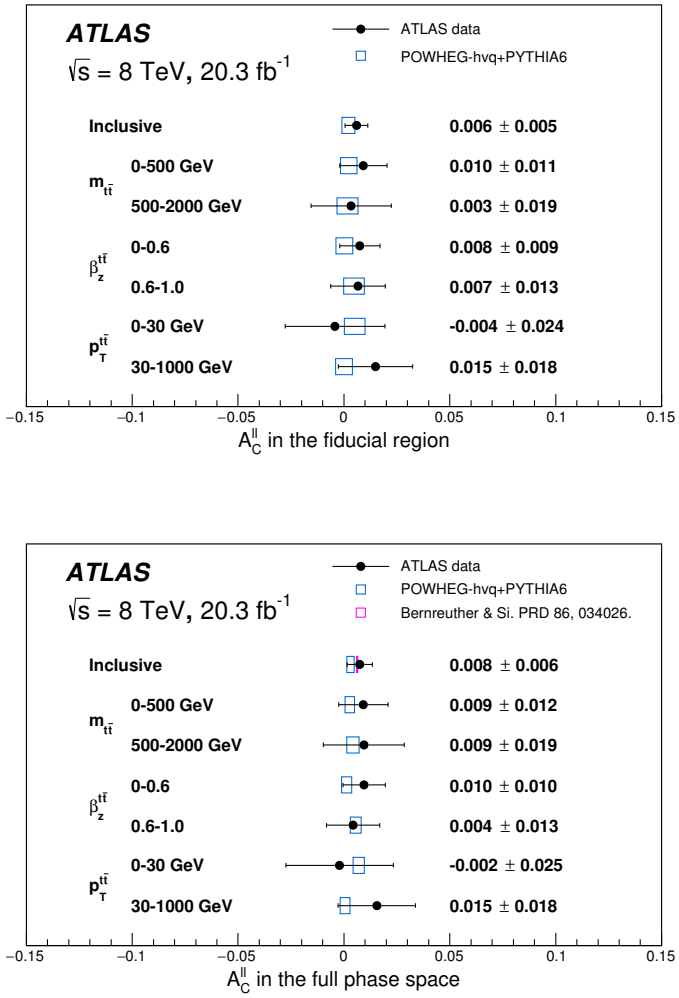


Figure 8.15 – *Summary of all the measurements of the leptonic asymmetry in the fiducial volume (top) and full phase space (bottom). The predictions shown in blue are obtained using* Powheg-*hvq* + Pythia *at NLO and the width of the box correspond to the statistical uncertainty. The inclusive measurement in the full phase space is compared to a NLO + EW prediction [48].*

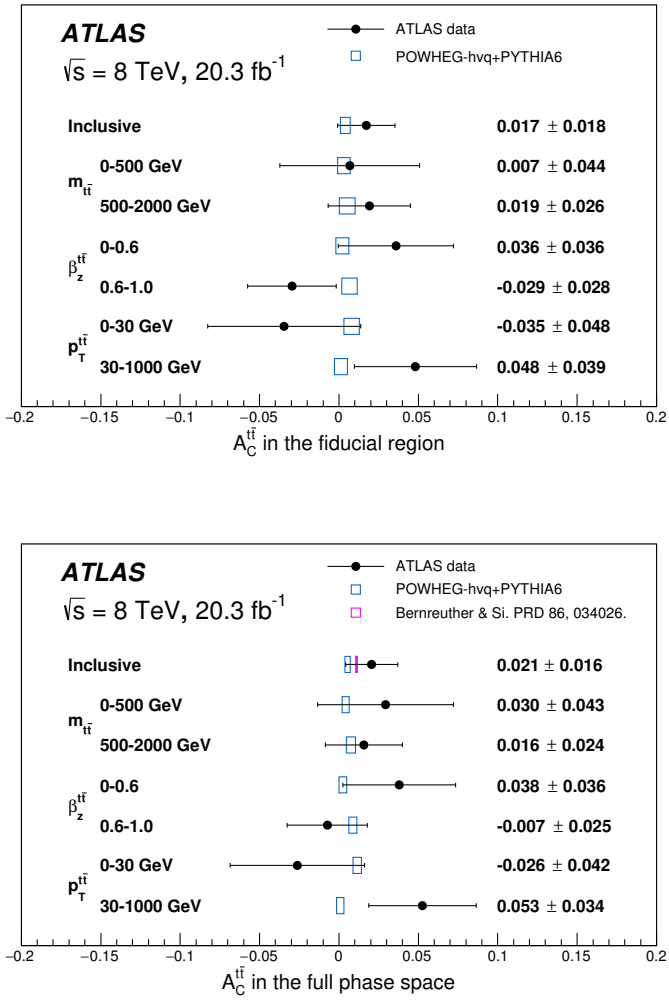


Figure 8.16 – *Summary of all the measurements for the $t\bar{t}$ asymmetry in the fiducial volume (top) and full phase space (bottom). The predictions shown in blue are obtained using* Powheg-*hvq* + Pythia *at NLO and the width of the box correspond to the statistical uncertainty. The inclusive measurement in the full phase space is compared to a NLO + EW prediction [48].*

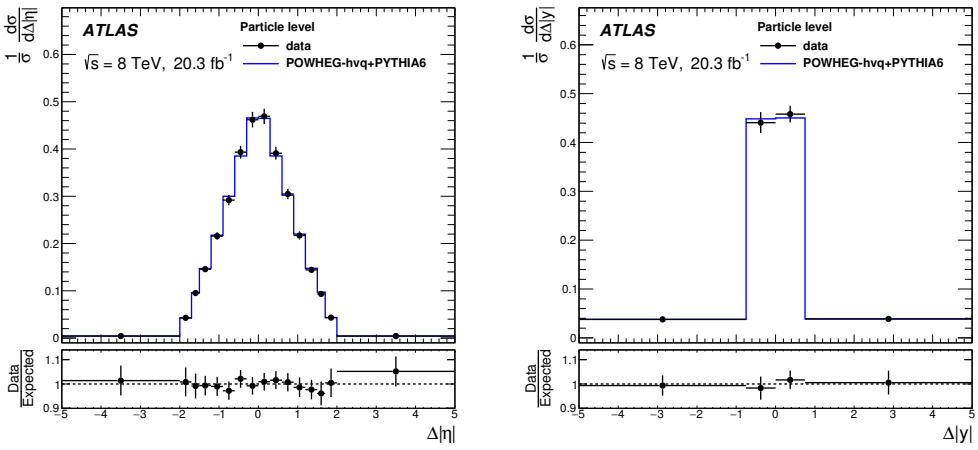


Figure 8.17 – *Distribution of Δ|η| (left) and Δ|y| (right) after the unfolding procedure in the fiducial volume. The predictions from* Powheg-*hvq* + Pythia *prediction at NLO are also shown. The data/expected ratio is shown in the bottom panel.*

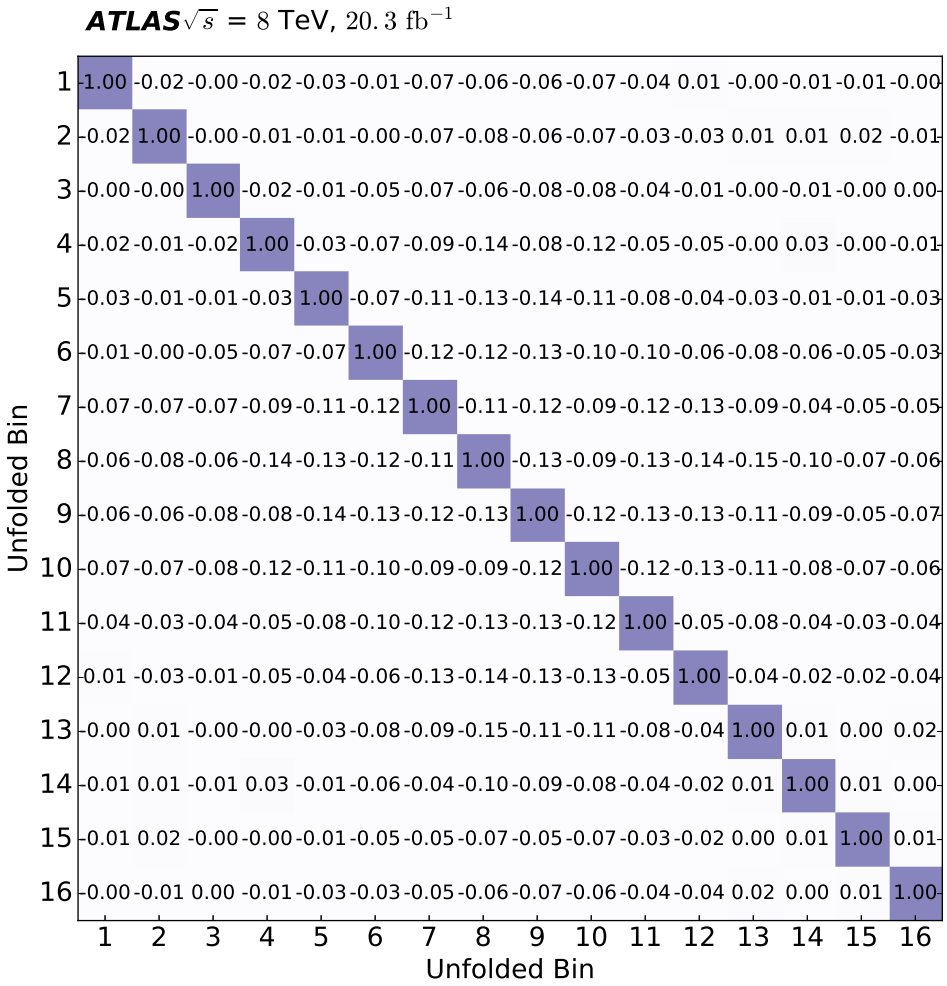


Figure 8.18 – *Correlations between the unfolded bins for the inclusive $\Delta|\eta|$ distribution in the fiducial volume. They are estimated by taking the normalized covariance of all possible truth values that are scanned during the unfolding procedure. The statistical and detector systematic uncertainties are taken into account. The uncertainty on these correlation factor is approximately* 0.1.

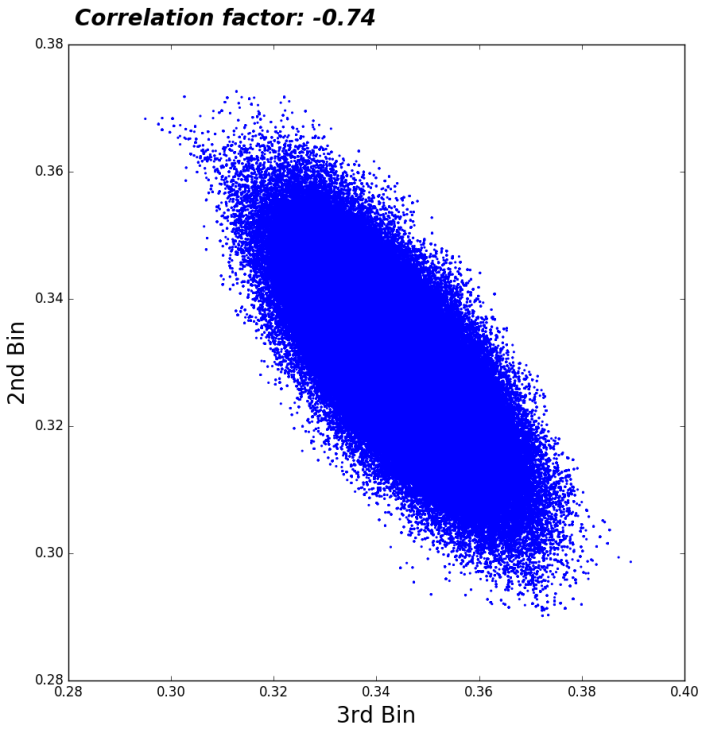


Figure 8.19 – *Scanned points for the second and third bin of the $\Delta|y|$ distribution. The correlation factor between the second and third bin of the $\Delta|y|$ distribution is estimated using these points and found to be -74%.*

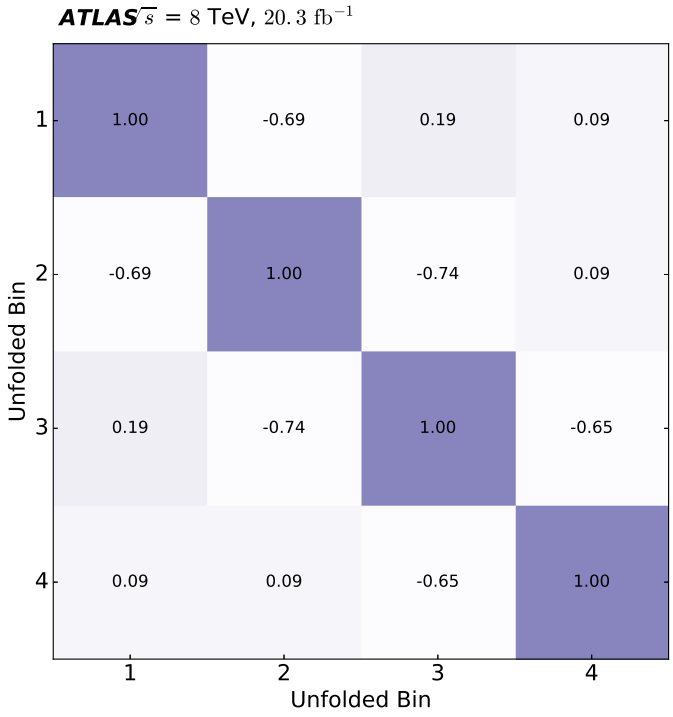


Figure 8.20 – *Correlations between the unfolded bins for the inclusive $\Delta|y|$ distribution in the fiducial volume. They are estimated by taking the normalized covariance of all possible truth values that are scanned during the unfolding procedure. The statistical and detector systematic uncertainties are taken into account. The uncertainty on these correlation factor is approximately* 0.1.

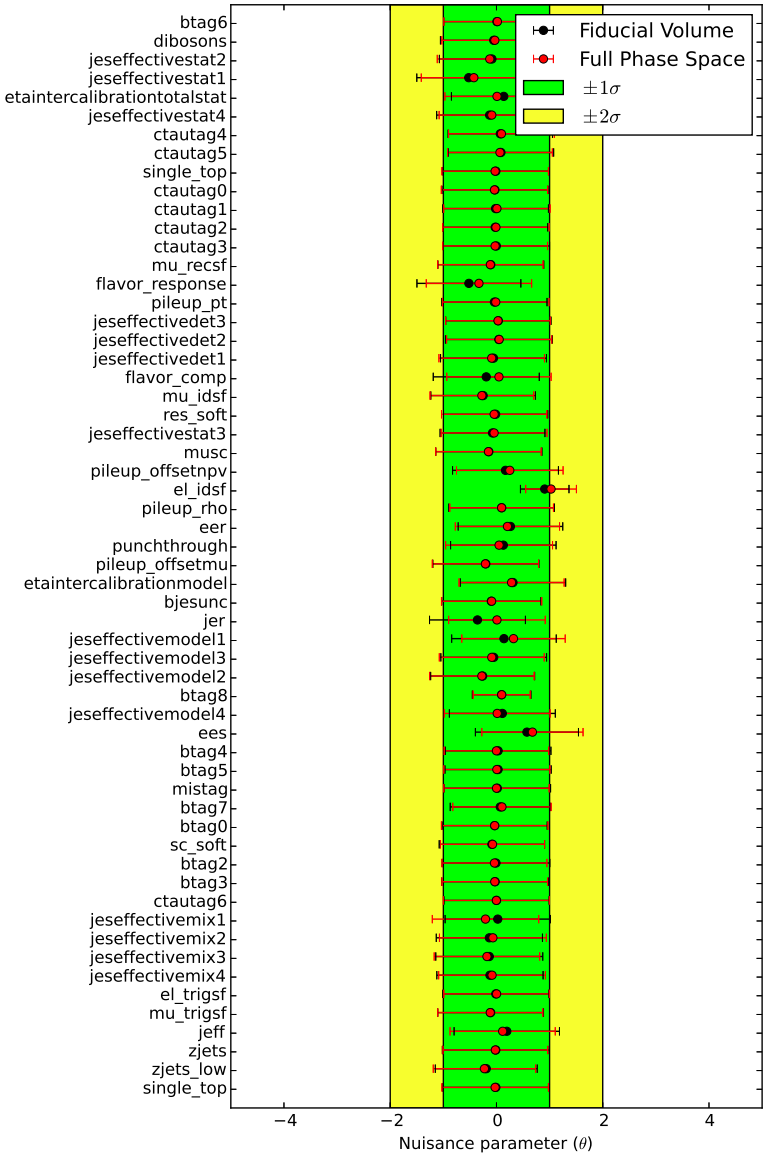


Figure 8.21 – *Nuisance parameters after the marginalization procedure for the inclusive $t\bar{t}$ asymmetry measurements in the fiducial volume (black) and full phase space (red).*

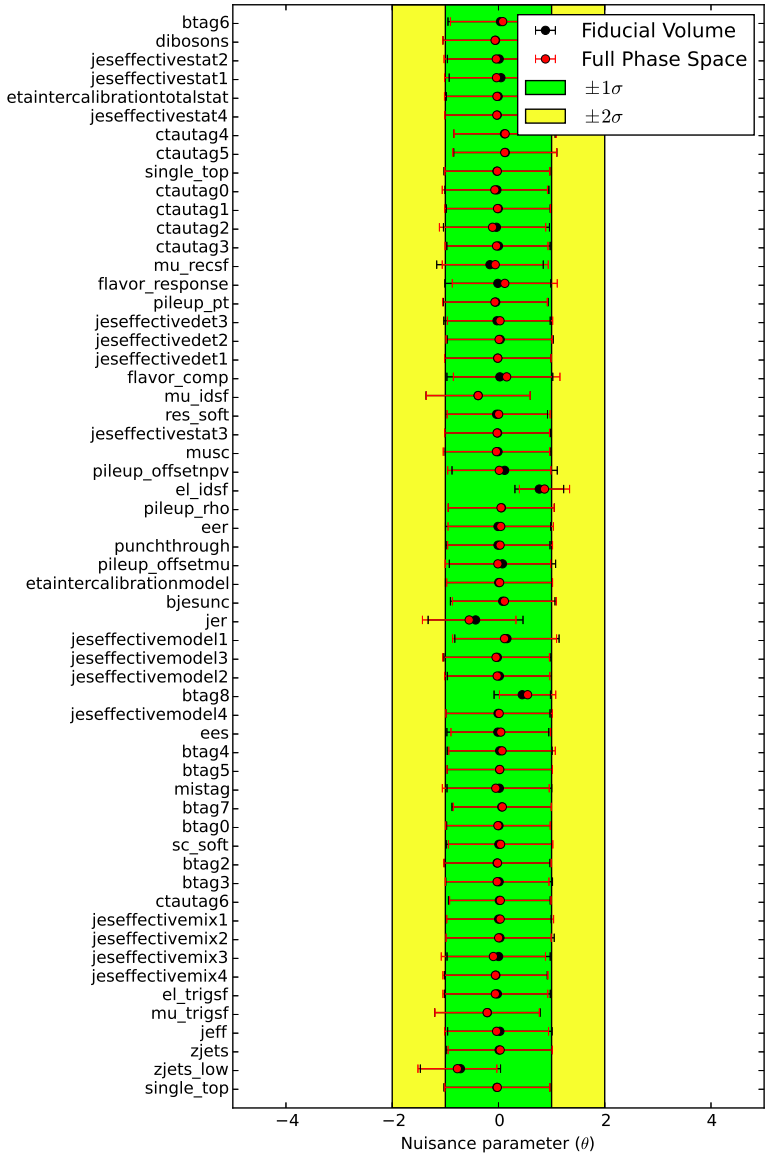


Figure 8.22 – *Nuisance parameters after the marginalization procedure for the inclusive leptonic asymmetry measurements in the fiducial volume (black) and full phase space (red).*

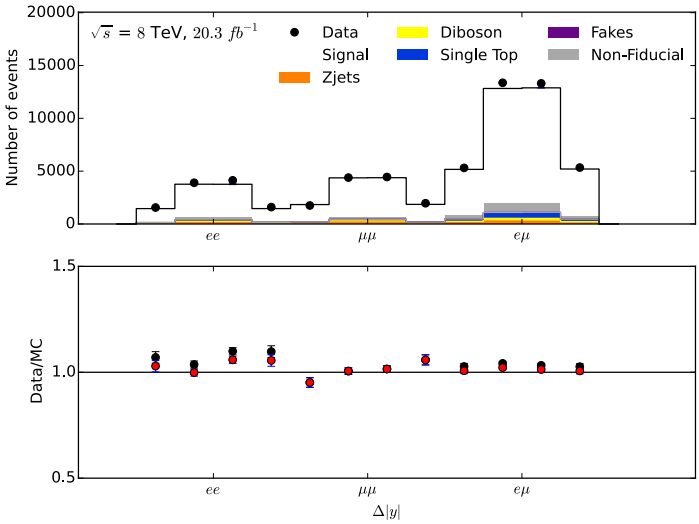


Figure 8.23 – *Data / Monte Carlo comparison of the $\Delta|y|$ distribution before the unfolding. In the ratio plot, the black points represent the ratio to the nominal MC sample while the blue points represent the $+1\sigma$ variation of the electron identification scale factor. A better agreement is observed between data and the $+1\sigma$ variation.*

8.3.2 BSM model interpretation

The charge asymmetry offers a windows to look for new physics, especially because different BSM models predicts different values of the $t\bar{t}$ and leptonic asymmetries. A comparison between the measured values of $A_C^{t\bar{t}}$ and $A_C^{\ell\ell}$ and the values predicted by different BSM models is performed. Figure 8.24 compares the values of $A_C^{\ell\ell}$ and $A_C^{t\bar{t}}$ from the inclusive measurements in the full phase space to the SM predictions and two types BSM models [154]. These models could be invoked to explain an anomalous forward-backward asymmetry at the Tevatron, such as reported by the CDF experiment [34], whilst still being compatible with the measurements performed at the LHC. Two types of BSM models with a new color-octet particle (G) that is exchanged in the s-channel are considered (see Fig. 8.25). In the model with the light color octet, the new particle's mass (m) is assumed to be 250 GeV, which is below the $t\bar{t}$ production threshold and its width is assumed to be $\Gamma = 0.2m$. The model with the heavy color octet uses an octet mass beyond current limits from direct searches at the LHC. The corrections to $t\bar{t}$ production are independent of the mass but instead depend on the ratio of the coupling to the top quarks and the mass, which is assumed to be 1 TeV^{-1}. The new particles in both BSM models would not be visible as resonances in the $m_{t\bar{t}}$ spectrum at the Tevatron or at the LHC. In the Fig. 8.24, model predictions for different left-handed, right-handed, and axial coupling constants to top quarks are shown. The ellipses correspond to the 1σ and 2σ total uncertainty on the measurements. The correlation between the two measurements is taken into account. The correlation taking into account the statistical and detector systematic uncertainty is found to be 30%. The modeling systematic uncertainties are assumed to be 100% correlated. The resulting correlation between the measured $A_C^{\ell\ell}$ and $A_C^{t\bar{t}}$ is

about 48%. The measurements are compatible with the SM and do not exclude the two sets of BSM models considered. More statistics are needed in order to exclude these models.

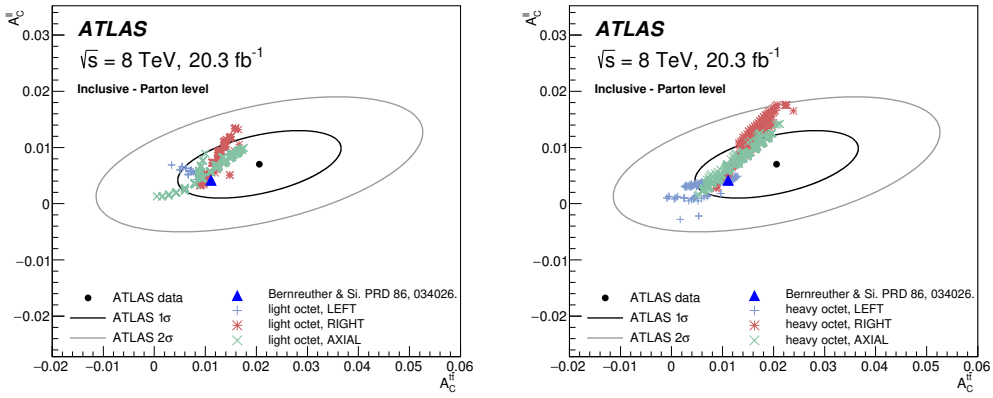


Figure 8.24 – *Comparison of the inclusive $A_C^{\ell\ell}$ and $A_C^{t\bar{t}}$ measurement values in the full phase space with the SM NLO QCD+EW prediction [48] and with two benchmark BSM models [154]. One correspond to a light octet with mass below the $t\bar{t}$ production threshold (left) and the other to a heavy octet with mass beyond the reach of the LHC (right). Ellipses corresponding to 1σ and 2σ are shown taking into account the total uncertainty. The correlation between $A_C^{\ell\ell}$ and $A_C^{t\bar{t}}$ is of 48%.*

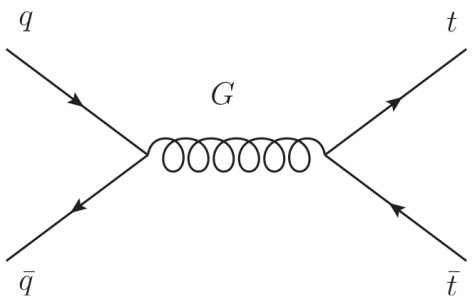


Figure 8.25 – *Tree-level $t\bar{t}$ production for a s-channel color-octet.*

8.3.3 Comparison with other measurements

The measurements are consistent with other LHC asymmetry measurements at 7 and 8 TeV [36–40, 155]. The measurements performed by ATLAS and CMS at 7 and 8 TeV are summarized in Tab. 8.16. The most precise measurement of the inclusive $t\bar{t}$ charge asymmetry is performed by CMS in the semileptonic channel using a template method, which is an approach different than the unfolding procedure. With the template method a more precise measurement can be obtained since it does not unfold but fold the measurement. However, it well known that this method it is highly model dependent and therefore can produce large modeling uncertainties. The unfolding procedure increases slightly the statistical uncertainty with respect to the template

			$A_C^{t\bar{t}}$		$A_C^{\ell\ell}$	
$\sqrt{s}$	Channel	Method	ATLAS	CMS	ATLAS	CMS
7 TeV	lepton+jets	unfolding	0.006 ± 0.010	0.001 ± 0.014	-	-
7 TeV	dilepton	unfolding	0.021 ± 0.030	-0.010 ± 0.019	0.024 ± 0.017	0.009 ± 0.012
8 TeV	lepton+jets	unfolding	0.009 ± 0.005	0.0010 ± 0.0077	-	-
8 TeV	lepton+jets	template method	-	0.003 ± 0.004	-	-
8 TeV	dilepton	unfolding	0.021 ± 0.016	0.011 ± 0.013	0.008 ± 0.006	0.003 ± 0.007

Table 8.16 – *Summary of the inclusive charge asymmetry measurements [36–40, 155].*

method. In terms of precision, this measurement is followed by the measurement performed by ATLAS in the same decay channel but using the unfolding approach. The large difference in the precision between the semileptonic and the dileptonic channel is mainly due to the size of the statistical uncertainty. In the dileptonic channel, the measurement performed by CMS is slightly more precise than the measurement performed in this analysis. The main reason for this difference is coming from the reconstruction uncertainty present in the KIN method. A different reconstruction method together with its optimization could increase the precision of these measurements. For the inclusive leptonic asymmetry, no reconstruction is involved and the result presented in this work corresponds to the most precise measurement to-date.

ATLAS and CMS have also performed differential measurements of the $t\bar{t}$ asymmetry at 7 and 8 TeV as a function of the mass, p_T, boost, and rapidity of the $t\bar{t}$ system. ATLAS has recently explored the high boosted topology, measuring for the first time the asymmetry in $m_{t\bar{t}} > 1.3$ TeV [156]. The results presented in this work are the first differential measurements performed by ATLAS in the dileptonic channel. A direct comparison of the measurements presented and other measurements is not possible since the differential bins are defined differently across the different measurements. However, all the measurements are in agreement with the predictions.

8.4 Outlook on charge asymmetry measurements

The charge asymmetry is an excellent tool for exploring BSM physics. The precise knowledge of this property, especially in boosted regimes, allows for a direct test of the SM. The precision of this measurement strongly depends on the amount of available data, since the statistical uncertainty is one of the dominant uncertainties across the different measurements. In the analysis presented, improvements on the reconstruction method can be achieved since its corresponding uncertainty is one of the dominant ones across the different measurements.

In Run 2, the collision energy was increased to 13 TeV and 150 fb^{-1} are expected to be collected by the end of the data taking period (October 2018). Although, by increasing the collision energy the asymmetry is diluted due to the presence of more $t\bar{t}$ events generated by gg-fusion, the amount of data can make the measurement sensitive to the SM prediction. This is especially true in the semileptonic channel whose branching ratio is significantly larger than the dilepton channel and the precision can increased by a factor of ~ 5.

The boosted $t\bar{t}$ topologies are promising in the near future since in this topology the asymmetry value can be enhanced due to the increased fraction of $q\bar{q}$ events since the initial radiation of the photon cannot be produced in gg-fusion. ATLAS started to explore the boosted topologies at 8 TeV for the charge asymmetry measurement, however, the measurements are strongly dominated by the statistical and modeling uncertainty. The precision of these results is expected to improve at higher energies with more statistics.

The asymmetry in $t\bar{t}+\gamma$ production is another topology where the charge asymmetry can be measured. Compared to inclusive $t\bar{t}$ production, the presence of an additional photon increases the number of $q\bar{q}$ events and therefore enhances the charge asymmetry over the gg background. The measurement of charge asymmetry in this topology is strongly limited by statistics, however, a good level of precision is expected at Run 2 (similar to the precision observed in this work).

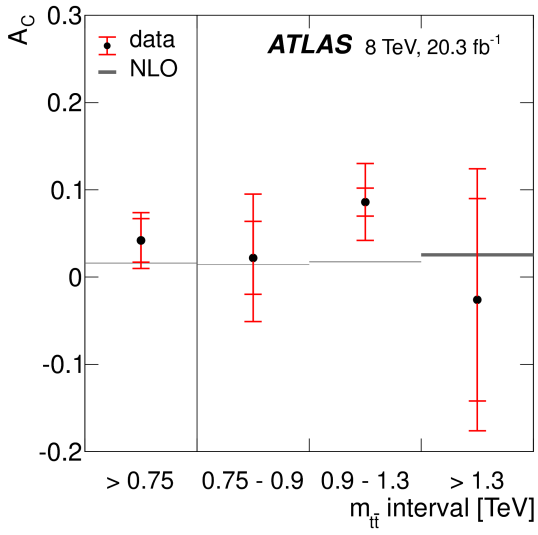


Figure 8.26 – *A summary of the charge asymmetry measurements performed in the boosted topology [157].*

With more data, other differential variables could be used in order to increase the $t\bar{t}$ fraction produced by $q\bar{q}$ annihilation. One interesting variable is the top quark p_{T}. This variable offers an excellent discrimination between gg-fusion and $q\bar{q}$ annihilation, but with a high cost in statistics. With the full dataset at 13 TeV, a first measurement as a function of the p_{T} of the top quark could be possible. Figure 8.27 shows the $q\bar{q}$ fraction for different cuts on the boost of the $t\bar{t}$ system and the top p_{T} at 8 TeV. The selection efficiency after applying the cuts are shown.

The asymmetry could also be defined differently using other observables. These observables are the *incline asymmetry*, based on the inclination between the planes of initial- and final-state momenta, and the *energy asymmetry*, based on the energy difference between top and antitop quarks [158]. The *incline asymmetry* allows for the access of the charge asymmetry in the $q\bar{q}$ channel, while the *energy asymmetry* allows for the access the charge asymmetry in the qg channel. The contribution from the qg channel is approximately 5 times larger than the $q\bar{q}$ channel at 14 TeV. These observables are defined in the $t\bar{t}+1$ jet topology and allows for the enhancement of the asymmetry values up to 12 % with the appropriated kinematic requirements, and offer

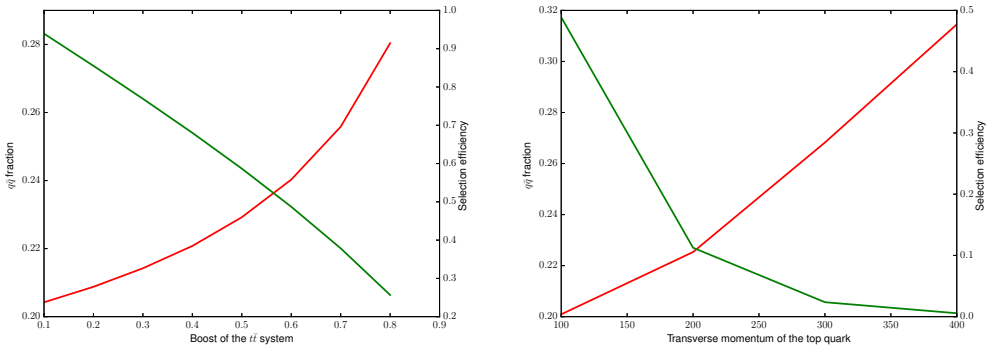


Figure 8.27 – *$q\bar{q}$ fraction (red line) as a function of boost of the $t\bar{t}$ system (left) and p_T of the top quark (right) at 8 TeV. The selection efficiency is also shown (green line).*

an excellent discovery window in proton-proton collisions. With 100fb^{-1} at 14 TeV, it should be possible to measure the observables with a significance of five standard deviationss [158].

9

Top spin measurements

In this chapter, the measurements of the top spin observables described in Chapter 1 are presented. The framework and workflow of this analysis are essentially the same as those used for the charge asymmetry measurements. In the first part, the optimization of the unfolding which leads to the optimal binning configuration for the different observables is described. This is followed by the estimation of the systematic uncertainties described in Chapter 7, as well as the description and estimation of the systematic uncertainties specific to this analysis. The final part presents and discusses the results obtained together with an outlook for the measurements.

9.1 Unfolding optimization

As in the charge asymmetry analysis, two different approaches for the unfolding procedure are used. The observables are unfolded back to parton level and to particle level. The measurement at particle level is performed in the fiducial region, while the measurement at parton level is performed in the full phase space. At parton level, the top quarks are considered by their status code in the MC history after radiation but before decay. Parton level leptons include tau leptons before they decay into an electron or muon and before radiation. With these definitions, the polarization and correlations are extracted (as explained in Chapter 1) from the slope of the $\cos\theta$ distributions and from the mean value of the $\cos\theta_- \cos\theta_+$ distributions, respectively.

The optimization of the unfolding is performed in order to define the best binning configuration for each observable. These studies are performed at parton level and the optimized binning configuration is used both at particle and parton levels. The binning configurations are optimized by evaluating the expected statistical uncertainty, the pulls and the bias from the linearity test. Another test, based on the relative difference between the true and the unfolded distributions in pseudoexperiments, is performed to evaluate the stability of the binning configurations.

For each observable, several binning configurations are tested. The number of bins and the bin widths for each observable are chosen based on its resolution. Due to the large number of observables and tests performed, only some examples are shown in the following. The resolution of the positive helicity polarization as well as the distribution at reconstructed level is shown in Fig. 9.1 as an example. The asymmetric behavior of the resolution is due to the difference between

the shapes of the reconstructed and the true distributions. Fewer events are reconstructed at $\cos\theta^k_+ = -1$, while the true distribution is expected to be flat across the whole range. At $\cos\theta^k_+ = -1$ the leptons are in the opposite direction than the top quarks with very low transverse momentum. By applying a p_T cut at 25 GeV most of these events are removed. In average, between eight and ten different binning candidates are tested for each observable. The binning configurations that are tested for the $\cos\theta^k_+$ observable are presented in Tab. 9.1.

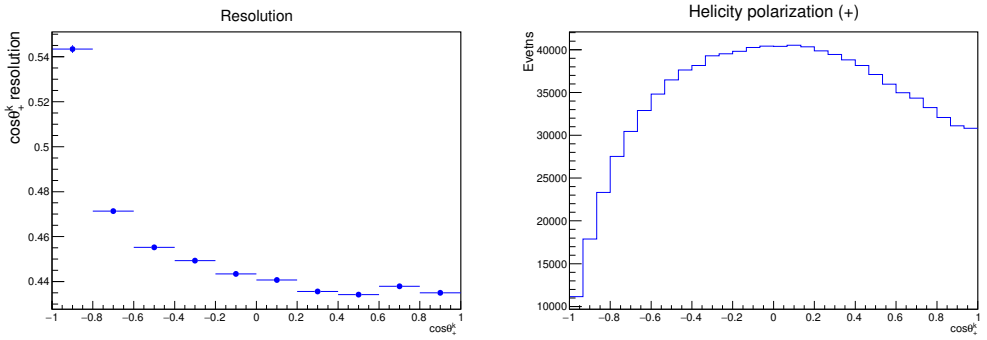


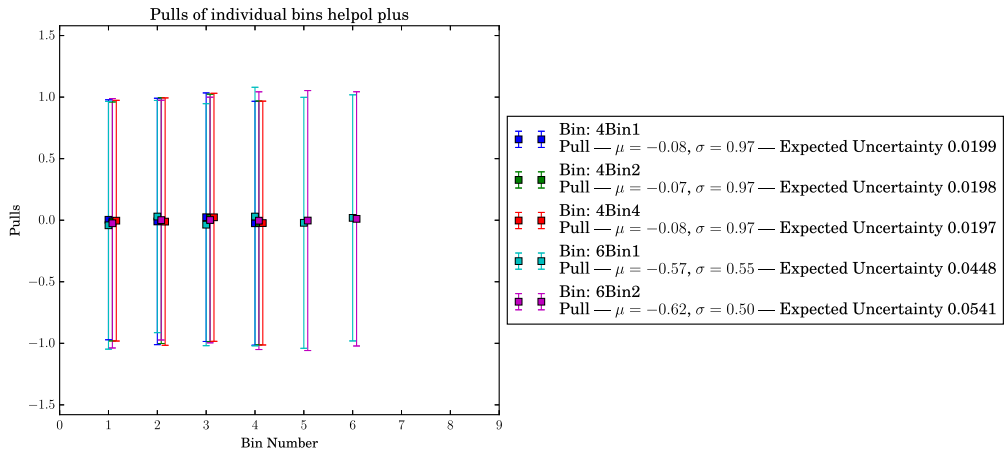
Figure 9.1 – *Resolution of the* $\cos\theta^k_+$ *observable (left) and its corresponding reconstructed distribution using the nominal MC sample (right).*

Name	Configuration
4Bin1	[-1.0, -0.5, 0.0, 0.5, 1.0]
4Bin2	[-1.0, -0.4, 0.0, 0.4, 1.0]
4Bin3	[-1.0, -0.6, 0.0, 0.6, 1.0]
4Bin4	[-1.0, -0.4, 0.0, 0.5, 1.0]
6Bin1	[-1.0, -0.6, -0.25, 0.0, 0.25, 0.6, 1.0]
6Bin2	[-1.0, -0.6, -0.3, 0.0, 0.3, 0.6, 1.0]
6Bin3	[-1.0, -0.66, -0.33, 0.0, 0.33, 0.66, 1.0]

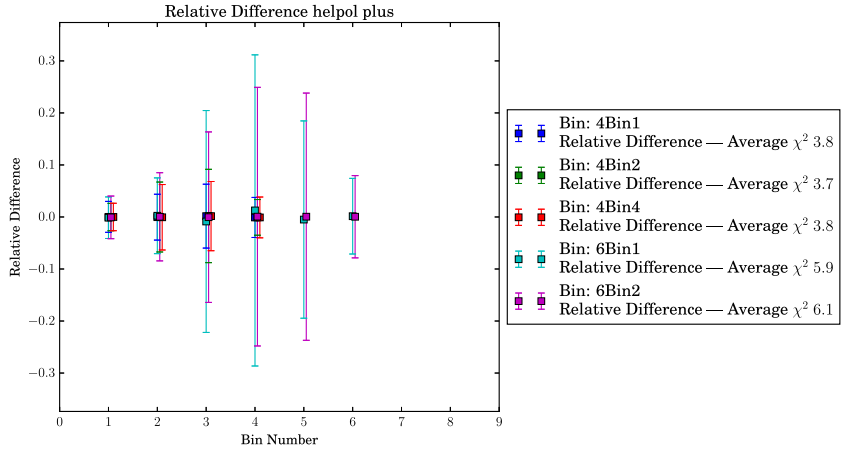
Table 9.1 – *The binning configurations for* $\cos\theta^k_+$ *used for the optimization studies.*

Pseudoexperiments are performed for the pull test on each binning candidate. For this analysis, the pull is estimated on the observables and on each bin of the distribution. Figure 9.2 shows the pulls on each bin and on the observable, as well as the relative difference of the unfolded distribution with respect to the true distribution. The average χ^2 is shown in the legend. The χ^2 is computed between each unfolded and the true distribution. It is possible to interpret the average χ^2 as a test of stability for the bins: if a large average of the χ^2 is found, the distributions is more likely to suffer from statistical fluctuations. For the four-bin configurations, the mean and RMS of the pulls of the distributions are close to zero and one, respectively. This is not the case for the six-bin configurations where the values are not compatible neither with zero nor one on the observable (not observed in the individual bins). The configuration of the bins as a whole may affect the fit performed in order to extract the polarization value, however, the individual behav-

ior of the bins can be consistent with expected pulls. e Small instabilities can be observed in the six-bin configuration since the relative difference fluctuate slightly around zero. This uncertainty can be interpreted as how much the bin content changes if small variations are introduced in the input distributions before the unfolding procedure. Also, it is possible to see that the expected statistical uncertainty increases with the number of bins.



(a) *Pulls*



(b) *Relative difference*

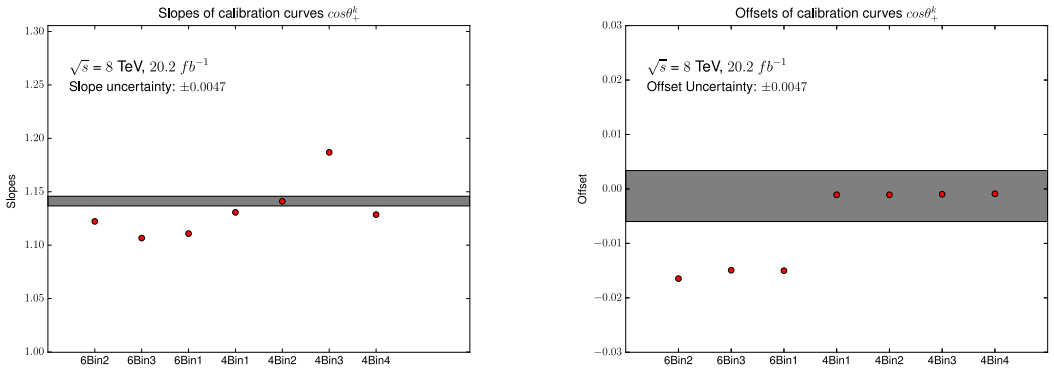
Figure 9.2 – *Pulls (a) and relative difference (b) for the positive helicity polarization in different binning configurations. The pulls on each of the bin (points on the left pad) is shown as well as in the observable (right pad). The relative difference on each bin with respect to the truth information is presented in the bottom figure. The average χ^2 is shown in the right canvas of the bottom figure. The degrees of freedom correspond to the number of bins on each histogram.*

A linearity test is performed in a similar way as in the charge asymmetry analysis. The linearity is determined by reweighting the observables at reconstructed level to different levels of polarizations and correlations at parton level, taking into account their values in the MC simulations. For the polarizations and the spin correlations in the same basis ($C(i,i)$), the double-differential

cross section is used, while a linear reweighting is used for the cross correlations. The double-differential cross section is used in order to remove the polarization and spin correlation in the MC generator and introduce the new values. The following weights w_{pol} and w_{corr} are used:

$$\begin{aligned} w_{\text{pol}} &= \frac{1 \pm P_{\text{rew}} \cos\theta_1 \pm P_{\text{rew}} \cos\theta_2 - C_{\text{th}} \cos\theta_1 \cos\theta_2}{1 \pm P_{\text{MC}} \cos\theta_1 \pm P_{\text{MC}} \cos\theta_2 - C_{\text{MC}} \cos\theta_1 \cos\theta_2} \\ w_{\text{corr}} &= \frac{1 - C_{\text{rew}} \cos\theta_1 \cos\theta_2}{1 \pm P_{\text{MC}} \cos\theta_1 \pm P_{\text{MC}} \cos\theta_2 - C_{\text{MC}} \cos\theta_1 \cos\theta_2}, \end{aligned} \tag{9.1}$$

where $\cos\theta_{1/2}$ refer to the angles between the momentum direction of a lepton in its parent top quark's rest frame and the quantization axis, using the true information of the MC sample. The factors P_{rew} and C_{rew} are the polarization and correlation, to which the distribution are reweighted, respectively. P_{MC} and C_{MC} refer to the polarization and spin correlation present in the MC sample used and C_{th} refers to the SM spin correlation value in the respective quantization axis [55]. The polarization observables are reweighted to values between $-0.1 < P_i < 0.1$, the spin correlations along the helicity and transverse axes between $0.20 < C(i,i) < 0.44$, along the r axis between $-0.1 < C(r,r) < 0.1$. The cross-correlations are reweighted between $-0.15 < C(i,j) \pm C(j,i) < 0.15$. The unfolded value for each reweighted distribution is then compared to the true value and a calibration curve is built by fitting a linear function and extracting the slope and offset. The difference with respect to a slope of one and an offset of zero is propagated to the final results as was done for the charge asymmetry measurements. Figure 9.3 shows the results of slopes and offsets for the calibration curves of the positive helicity polarization at parton level.[1] The offsets are negligible for the four bin configurations, while the slopes get closer to one for the six-bin configurations.



(a) *Slopes* **(b)** *Offsets*

Figure 9.3 – *Slopes and offsets of the calibration curves obtained for different binning configuration for the transverse correlation in the full phase space at parton level. The gray band represents the MC statistical uncertainty on the slopes and offsets.*

A compromise has to be found regarding the expected statistical uncertainties, the instabilities, and the bias introduced by the unfolding procedure. Therefore, the four bin configuration is the best option for the helicity polarization. Specifically, the asymmetry binning 4bin4 is chosen

[1] A good linear behavior is obtained on each case.

Observable	Binning
$\cos\theta^k_\pm$	[-1.0, -0.4, 0.0, 0.5, 1.0]
$\cos\theta^n_\pm$	[-1.0, -0.5, 0.0, 0.5, 1.0]
$\cos\theta^r_\pm$	[-1.0, -0.5, 0.0, 0.5, 1.0]
$\cos\theta^k_+\cos\theta^k_-$	[-1.0, -0.6, -0.25, 0.0, 0.25, 0.6, 1.0]
$\cos\theta^n_+\cos\theta^n_-$	[-1.0, -0.6, -0.25, 0.0, 0.25, 0.6, 1.0]
$\cos\theta^r_+\cos\theta^r_-$	[-1.0, -0.66, -0.33, 0.0, 0.33, 0.66, 1.0]
$\cos\theta^n_+\cos\theta^k_- + \cos\theta^k_+\cos\theta^n_-$	[-1.0, -0.6, -0.3, 0.0, 0.3, 0.6, 1.0]
$\cos\theta^n_+\cos\theta^k_- - \cos\theta^k_+\cos\theta^n_-$	[-1.0, -0.6, -0.3, 0.0, 0.3, 0.6, 1.0]
$\cos\theta^n_+\cos\theta^r_- + \cos\theta^r_+\cos\theta^n_-$	[-1.0, -0.6, -0.3, 0.0, 0.3, 0.6, 1.0]
$\cos\theta^n_+\cos\theta^r_- - \cos\theta^r_+\cos\theta^n_-$	[-1.0, -0.6, -0.3, 0.0, 0.3, 0.6, 1.0]
$\cos\theta^r_+\cos\theta^k_- + \cos\theta^k_+\cos\theta^r_-$	[-1.0, -0.6, -0.3, 0.0, 0.3, 0.6, 1.0]
$\cos\theta^r_+\cos\theta^k_- - \cos\theta^k_+\cos\theta^r_-$	[-1.0, -0.6, -0.3, 0.0, 0.3, 0.6, 1.0]

Table 9.2 – *Chosen binning configuration for each of the observables.*

(because it has a slope closer to one than the others). The same procedure is applied to each of the observable in order to select the best binning configuration.

After the binning optimization is performed for all the observables, four-bin configurations for the polarization observables and six-bin configurations for the different correlation observables are used. The explicit form of the binning can be found in Tab. 9.2. In general, the binnings are symmetric around zero. The only exception is for the polarization along the helicity axis, where an asymmetric binning is chosen due to the shape of its resolution. The input distribution and the response matrix normalized per truth bin are shown for one example of polarization, spin correlation, and cross-correlation in Fig. 9.4 and 9.5. The input distribution for FBU corresponds to the distribution of the observable for each of the channels placed side-by-side. The fraction of events in the diagonal bins of the response matrices are between 0.40 and 0.60 for B^k_+, around 0.40 for $C(n,n)$, and between 0.30 and 0.40 for $C(n,r)+C(r,n)$. The large migrations present in the $C(n,n)$ and $C(n,r)+C(r,n)$ matrices come from the resolution dilution due to how the observables are computed with the different rest frames.

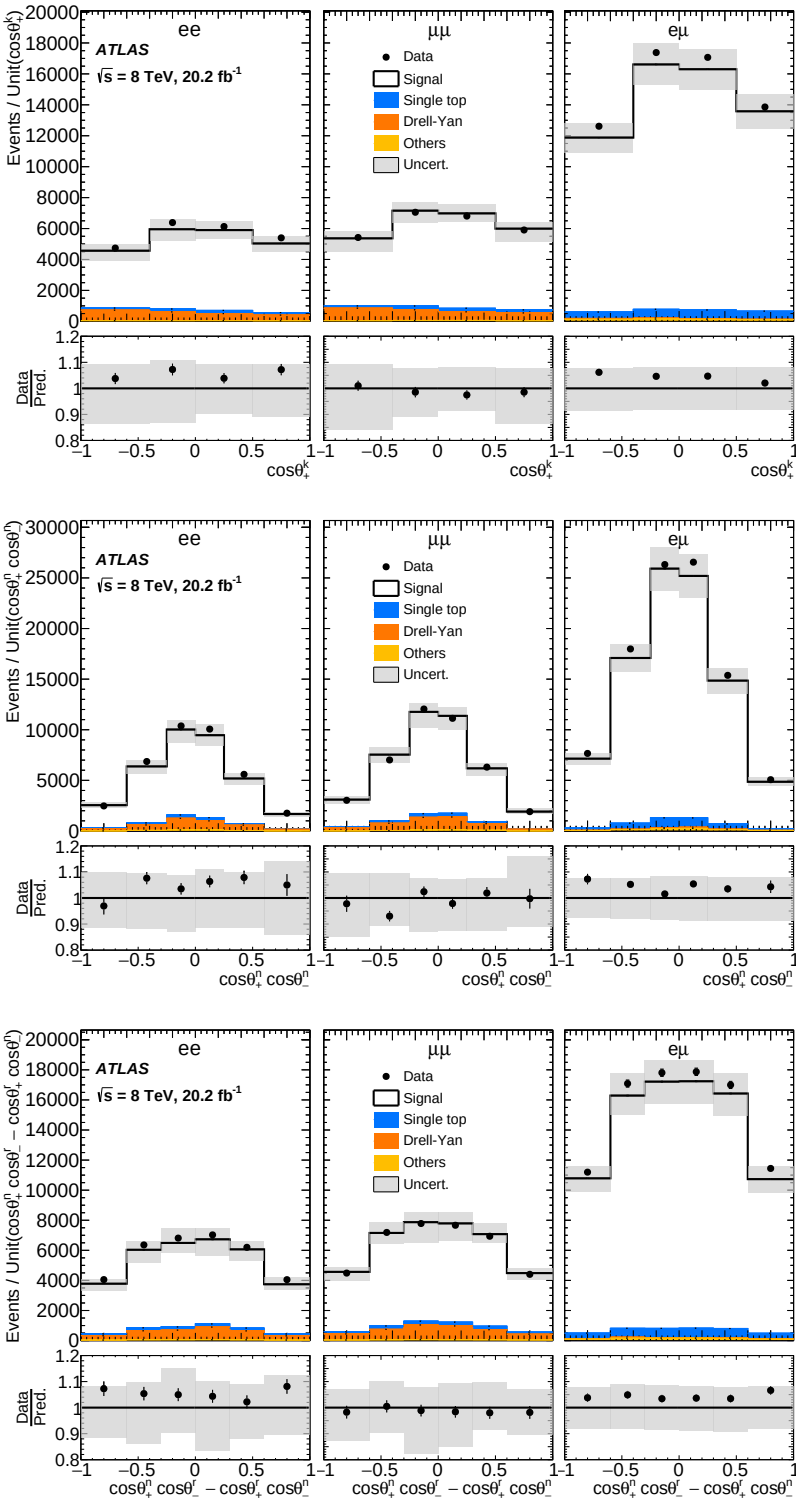


Figure 9.4 – *Input distributions for the unfolding procedure of* $\cos\theta_+^k$, $\cos\theta_+^n \cos\theta_-^n$, *and* $\cos\theta_+^n \cos\theta_-^r - \cos\theta_+^r \cos\theta_-^n$. *The data/expectation ratio is also shown. The gray area shows the total uncertainty on the signal and background. The* $t\bar{t}V$, *diboson and fake lepton backgrounds are shown together in the "Others" category.*

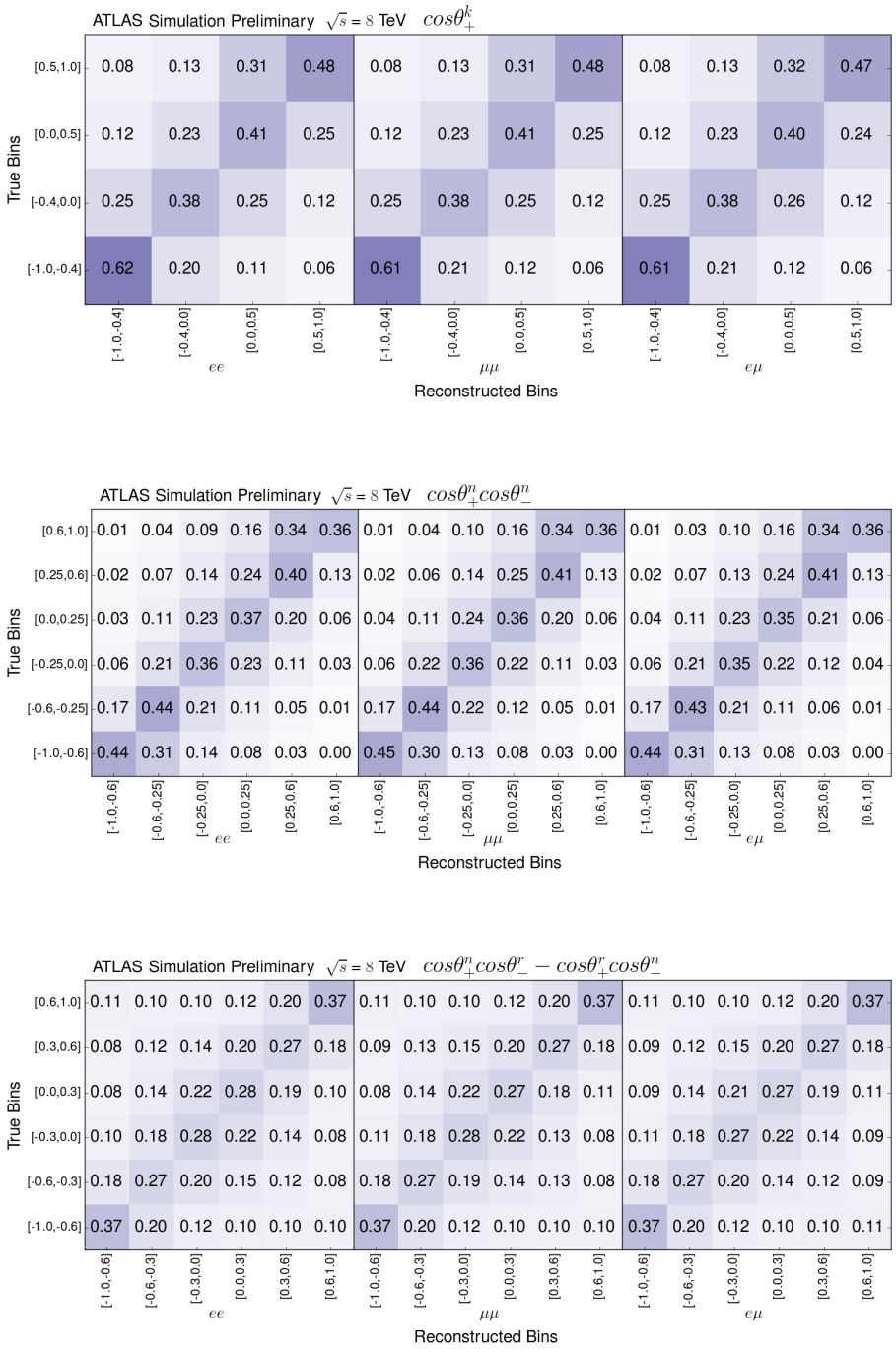


Figure 9.5 – *Response matrices of observables* $\cos\theta^k_+$, $\cos\theta^n_+\cos\theta^n_-$, *and* $\cos\theta^n_+\cos\theta^r_- - \cos\theta^r_+\cos\theta^n_-$ *at parton level. They are divided into the ee, μμ, and eμ channels. The matrices are normalized per truth bin (rows) for each channel separately.*

Source of systematic uncertainty	Fiducial Region	Full Phase Space
Background	0.0007	0.0004
Lepton reconstruction, identification and trigger	0.0013	0.0012
Lepton momentum scale and resolution	0.0009	0.0011
Jet energy scale	0.0034	0.0038
Jet reconstruction and vertex fraction efficiency	0.0003	0.0003
Jet resolution	0.0014	0.0005
B-tagging/mis-tag efficiency	0.0031	0.0025
Missing transverse momentum	0.0016	0.0004
Total	0.0055	0.0050

Table 9.3 – *Detector and background systematic uncertainties estimation for the $cos\theta^r_+ cos\theta^r_-$ observable in the fiducial region and in the full phase space.*

9.2 Systematic uncertainties

In the following, the estimation of the systematic uncertainties are presented based on the methods explained in Chapter 7. Also, the descriptions and estimations of systematic uncertainties specific to this analysis are presented.

9.2.1 Detector modeling and background

Using the method described in Sec. 7.1.3, the individual contributions of the detector modeling and background systematic uncertainties are estimated. Some examples are shown and discussed in this section. The summary of all systematic uncertainty is presented in Sec. 9.3. The estimation for the $\cos\theta^r_+ \cos\theta^r_-$ and the $\cos\theta^n_+ \cos\theta^k_- + \cos\theta^k_+ \cos\theta^n_-$ observable are shown in Tab. 9.3 and 9.4, respectively. The main contribution is observed in the jet energy scale and the b-tagging systematic variation, followed by the jet energy resolution. These variations affect directly the kinematic reconstruction, which can significantly vary the value of the observables. The background uncertainties represent a minor contribution. The total uncertainty is slightly larger in the fiducial region than in the full phase space due to the nonfiducial background. Its subtraction from the signal distribution during the unfolding procedure affects both shape and statistics and the same behavior is observed for the different observables.

Source of systematic uncertainty	Fiducial Region	Full Phase Space
Background	0.0011	0.0006
Lepton reconstruction, identification and trigger	0.0005	0.0013
Lepton momentum scale and resolution	0.0010	0.0005
Jet energy scale	0.0025	0.0033
Jet reconstruction and vertex fraction efficiency	0.0004	0.0007
Jet resolution	0.0001	0.0015
B-tagging/mis-tag efficiency	0.0022	0.0021
Missing transverse momentum	0.0004	0.0014
Total	0.0038	0.0047

Table 9.4 – *Detector and background systematic uncertainties for the* $\cos\theta_{+}^{n}\cos\theta_{-}^{k} + \cos\theta_{+}^{k}\cos\theta_{-}^{n}$ *observable in the fiducial region and in the full phase space.*

9.2.2 Signal modeling

The modeling systematic uncertainties are estimated using the techniques described in Sec. 7.2. As in Sec. 9.2.1, some examples are shown and discussed in this section. The summary of the impact of all systematic uncertainties will be presented in Sec. 9.3. The different contributions are shown for the $\cos\theta_+^k$ and $\cos\theta_+^n\cos\theta_-^k + \cos\theta_+^k\cos\theta_-^n$ observables in Tab. 9.5 and 9.6, respectively. The MC statistical uncertainty associated with the estimation is also shown. The actual uncertainty is taken from the maximum value between the estimation and its MC statistical uncertainty. The PDF uncertainties do not have an uncertainty since the estimation is performed by reweighting the nominal signal sample. In general, the dominant uncertainties across the different observables are the color reconnection and underlying event, followed by the ISR/FSR systematic variation. However, it is not possible to distinguish if these contributions correspond to a real physical effect given the size of the MC statistical uncertainty.

Source of systematic uncertainty	Fiducial Region	Full Phase Space
MC generator	0.0149 ± 0.0162	0.0187 ± 0.0225
Parton Shower	0.0220 ± 0.0087	0.0186 ± 0.0122
ISR/FSR	0.0033 ± 0.0116	0.0043 ± 0.0141
Color Reconnection	0.0087 ± 0.0250	0.0199 ± 0.0317
Underlying Event	0.0301 ± 0.0271	0.0164 ± 0.0358
PDF	0.0064	0.0101
Total	0.0495	0.0537

Table 9.5 – *Modeling systematic uncertainty contributions of the different sources for the $cos\theta_+^k cos\theta_-^k$ observable in the fiducial region and in the full phase space.*

Source of systematic uncertainty	Fiducial Region	Full Phase Space
MC generator	0.0339 ± 0.0307	0.0234 ± 0.0286
Parton Shower/Hadronization	0.0171 ± 0.0169	0.0056 ± 0.0175
ISR/FSR	0.023 ± 0.0219	0.0197 ± 0.0210
Color Reconnection	0.0517 ± 0.0481	0.0602 ± 0.0444
Underlying Event	0.0415 ± 0.0514	0.0379 ± 0.0489
PDF	0.0040	0.0018
Total	0.0854	0.0871

Table 9.6 – *Modeling systematic uncertainty contributions of the different sources for the $cos\theta_+^n cos\theta_-^k + cos\theta_+^k cos\theta_-^n$ observable in the fiducial region and in the full phase space.*

Observable	Fiducial Region	Full Phase Space
$\cos\theta_+^k$	-0.04 ± 0.01	-0.04 ± 0.01
$\cos\theta_-^k$	-0.04 ± 0.01	-0.04 ± 0.01
$\cos\theta_+^n$	0.01 ± 0.01	0.03 ± 0.01
$\cos\theta_-^n$	-0.01 ± 0.01	-0.00 ± 0.01
$\cos\theta_+^r$	0.01 ± 0.01	0.01 ± 0.01
$\cos\theta_-^r$	0.02 ± 0.01	0.03 ± 0.01
$\cos\theta_+^k \cos\theta_-^k$	0.04 ± 0.01	0.06 ± 0.02
$\cos\theta_+^n \cos\theta_-^n$	-0.04 ± 0.03	-0.02 ± 0.03
$\cos\theta_+^r \cos\theta_-^r$	0.05 ± 0.03	0.06 ± 0.03
$\cos\theta_+^r \cos\theta_-^k + \cos\theta_+^k \cos\theta_-^r$	-0.06 ± 0.02	-0.10 ± 0.02
$\cos\theta_+^r \cos\theta_-^k - \cos\theta_+^k \cos\theta_-^r$	0.04 ± 0.03	0.04 ± 0.03
$\cos\theta_+^n \cos\theta_-^r + \cos\theta_+^r \cos\theta_-^n$	0.02 ± 0.01	0.02 ± 0.02
$\cos\theta_+^n \cos\theta_-^r - \cos\theta_+^r \cos\theta_-^n$	-0.08 ± 0.01	-0.07 ± 0.01
$\cos\theta_+^n \cos\theta_-^k + \cos\theta_+^k \cos\theta_-^n$	-0.00 ± 0.04	-0.01 ± 0.04
$\cos\theta_+^n \cos\theta_-^k - \cos\theta_+^k \cos\theta_-^n$	-0.00 ± 0.04	-0.01 ± 0.03

Table 9.7 – *Slopes from the top mass dependency in MC simulation. The uncertainty on the estimation corresponds to the product of the slope and the uncertainty of 0.70 GeV on the ATLAS mass measurement [151].*

The mass uncertainty is evaluated using the method described in Chapter 7 and the dependence of each observable with respect to the MC top quark mass is presented in Tab. 9.7. In several observables no significant dependence is observed. The largest dependency is found in the $C(r,k) + C(k,r)$ and in the $C(n,r) - C(r,n)$ observables. The uncertainties are estimated by multiplying the slopes by the uncertainty on the most precise ATLAS top quark mass measurements of 0.70 GeV [151]. An example of the linear fit performed for the $\cos\theta_+^k \cos\theta_-^k$ and $\cos\theta_+^r \cos\theta_-^r$ observables using different MC masses is shown in Fig. 9.6. The product of the slope and the mass uncertainty is also shown.

Other sources of uncertainties specific to this analysis are studied.

- **Top quark p_{T} modeling.** The top quark p_{T} spectrum is not satisfactorily modeled in MC simulation [30, 159]. The impact of the mismodeling is estimated by reweighting the top quark p_{T} simulation to data. The different distributions are then unfolded using the nominal response matrix. The differences with respect to the nominal values are negligible compared to the other modeling uncertainties. The impact of this mismodeling is thus considered negligible, and no uncertainty is added to the total uncertainty.

- **Polarization/Spin correlation.** The response matrices used in the unfolding are calculated using the SM polarization and spin correlation. An uncertainty related to a different polarization/spin correlation is associated by changing their values and performing a linearity

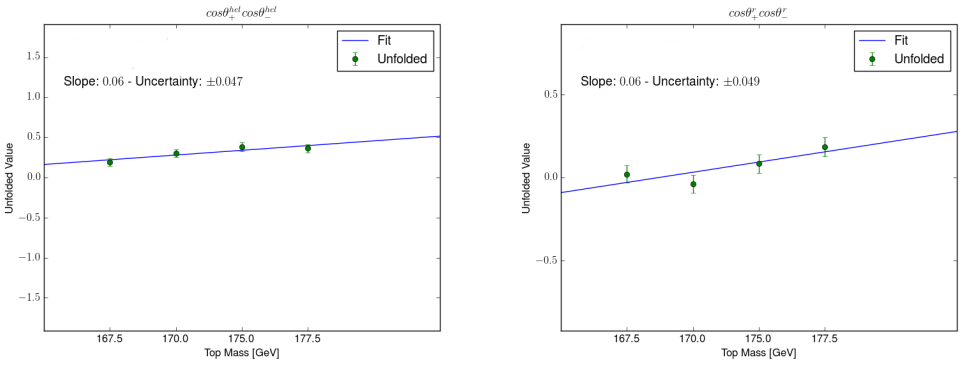


Figure 9.6 – *Mass dependency for the* $\cos\theta^{k}_{+}\cos\theta^{k}_{-}$ *(left) and* $\cos\theta^{r}_{+}\cos\theta^{r}_{-}$ *(right) observables in the full space space. The uncertainty quoted on the figure corresponds to the product of the slope and the uncertainty on the mass measurement of 0.70 GeV [151].*

test using the nominal response matrix. In the reweighting procedure of the spin correlation observables, the polarization is changed by $\pm 0.5\%$, while for the polarization observables, the spin correlation is changed by ± 0.1. This uncertainty cannot be applied to the cross correlation observables as no analytic description of these observables is available. Instead, a linear reweighting is used, not depending on the polarization or spin correlation of any axis. The estimation of this uncertainty is shown in Tab. 9.8 for each observable. This contribution is considered minor and it is added in quadrature to the total signal modeling uncertainty. The largest uncertainty is present on the helicity and r polarizations.

9.2.3 Other uncertainties

Nonclosure uncertainties

The uncertainties related with the remaining bias of the unfolding procedure is evaluated as for the charge asymmetry measurements, described in Chapter 7. In Tab. 9.9 the uncertainties for each of the observables are shown in the fiducial region and in the full phase space. These uncertainties are small and added in quadrature to the total uncertainties.

Observable	Polarization systematic	Correlation systematic
$\cos\theta_+^k$	−−	0.0048
$\cos\theta_-^k$	−−	0.0048
$\cos\theta_+^n$	−−	0.0001
$\cos\theta_-^n$	−−	0.0004
$\cos\theta_+^r$	−−	0.0039
$\cos\theta_+^r$	−−	0.0040
$\cos\theta_+^k \cos\theta_-^k$	0.0011	−−
$\cos\theta_+^n \cos\theta_-^n$	< 0.0001	−−
$\cos\theta_+^r \cos\theta_-^r$	0.0003	−−
$\cos\theta_+^r \cos\theta_-^k + \cos\theta_+^k \cos\theta_-^r$	< 0.0001	−−
$\cos\theta_+^r \cos\theta_-^k - \cos\theta_+^k \cos\theta_-^r$	< 0.0001	−−
$\cos\theta_+^n \cos\theta_-^r + \cos\theta_+^r \cos\theta_-^n$	< 0.0001	−−
$\cos\theta_+^n \cos\theta_-^r - \cos\theta_+^r \cos\theta_-^n$	< 0.0001	−−
$\cos\theta_+^n \cos\theta_-^k + \cos\theta_+^k \cos\theta_-^n$	< 0.0001	−−
$\cos\theta_+^n \cos\theta_-^k - \cos\theta_+^k \cos\theta_-^n$	< 0.0001	−−

Table 9.8 – *Polarization and correlation systematic uncertainties for all the observables in the full phase space. In most of the observables the uncertainties are very small. A similar behavior is observed in the fiducial region.*

Observable	Fiducial Region	Full Phase Space
$\cos\theta_+^k$	0.0134	0.0050
$\cos\theta_-^k$	0.0113	0.0103
$\cos\theta_+^n$	0.0017	0.0014
$\cos\theta_-^n$	0.0037	0.0017
$\cos\theta_+^r$	0.0041	0.0014
$\cos\theta_-^r$	0.0082	0.0011
$\cos\theta_+^k \cos\theta_-^k$	0.0170	0.0109
$\cos\theta_+^n \cos\theta_-^n$	0.0039	0.0007
$\cos\theta_+^r \cos\theta_-^r$	0.0200	0.0160
$\cos\theta_+^r \cos\theta_-^k + \cos\theta_+^k \cos\theta_-^r$	0.0249	0.0289
$\cos\theta_+^r \cos\theta_-^k - \cos\theta_+^k \cos\theta_-^r$	0.0012	0.0019
$\cos\theta_+^n \cos\theta_-^r + \cos\theta_+^r \cos\theta_-^n$	0.0229	0.0098
$\cos\theta_+^n \cos\theta_-^r - \cos\theta_+^r \cos\theta_-^n$	0.0068	0.0037
$\cos\theta_+^n \cos\theta_-^k + \cos\theta_+^k \cos\theta_-^n$	0.0056	0.0108
$\cos\theta_+^n \cos\theta_-^k - \cos\theta_+^k \cos\theta_-^n$	0.0109	0.0048

Table 9.9 – *Nonclosure uncertainties for each of the observables in the fiducial region and in the full phase space. This contribution is generally small in comparison with the total uncertainty on the measurements.*

Observable	Fiducial Region	Full Phase Space
$\cos\theta_+^k$	0.0048	0.0049
$\cos\theta_-^k$	0.0043	0.0044
$\cos\theta_+^n$	0.0047	0.0048
$\cos\theta_-^n$	0.0047	0.0047
$\cos\theta_+^r$	0.0044	0.0048
$\cos\theta_-^r$	0.0057	0.0059
$\cos\theta_+^k \cos\theta_-^k$	0.0012	0.0016
$\cos\theta_+^n \cos\theta_-^n$	0.0008	0.0007
$\cos\theta_+^r \cos\theta_-^r$	0.0011	0.0012
$\cos\theta_+^r \cos\theta_-^k + \cos\theta_+^k \cos\theta_-^r$	0.0018	0.0023
$\cos\theta_+^r \cos\theta_-^k - \cos\theta_+^k \cos\theta_-^r$	0.0016	0.0023
$\cos\theta_+^n \cos\theta_-^r + \cos\theta_+^r \cos\theta_-^n$	0.0022	0.0021
$\cos\theta_+^n \cos\theta_-^r - \cos\theta_+^r \cos\theta_-^n$	0.0017	0.0015
$\cos\theta_+^n \cos\theta_-^k + \cos\theta_+^k \cos\theta_-^n$	0.0021	0.0020
$\cos\theta_+^n \cos\theta_-^k - \cos\theta_+^k \cos\theta_-^n$	0.0015	0.0015

Table 9.10 – *Uncertainty related to the size of the MC samples for each of the observables in the fiducial region and in the full phase space. This contribution is very small in comparison with the total uncertainty on the measurements.*

MC statistical uncertainty

The uncertainty related with the size of the MC samples for each observable in the fiducial and full phase space is shown in Tab. 9.10. The estimation is performed in the same way as for the charge asymmetry measurements, using the method detailed in Chapter 7. This uncertainty is added in quadrature to the total uncertainty. In general, the impact on the measurements is small.

9.3 Results

After the unfolding procedure, the polarizations, correlations, and cross-correlations are extracted. The uncertainty from the unfolding (including marginalization) corresponds to the statistical, detector modeling, and background uncertainties. The signal modeling, the nonclosure, and the MC statistical uncertainties are added in quadrature. Table 9.11 presents the central value and the total uncertainty, as well as a breakdown of the systematic uncertainties for the various categories for all the studied observables. At parton level, the polarizations, correlations, and cross-correlations observables are presented. At particle level, the means of distributions are presented. Because of the limited phase-space used at particle level, the values of the polarization and spin correlations are not proportional to the means of these distributions. In order to compare the size of the uncertainties between the particle and parton level measurements, the means of the polarization observables are multiplied by a factor of 3 and the correlations by a factor of -9. Significant differences are observed in the central value of some observables, such as the polarizations along the helicity and r axis. These differences are expected since the distributions in the fiducial region are significantly different from the distributions in the full phase space. The differences arise due to the kinematic cuts applied to the leptons and jets that define the fiducial region. Figures 9.7 to 9.9 show the observable distributions corrected back to particle level and compared to the true distribution generated with Powheg-hvq+Pythia. In Appendix E, the distributions corrected at parton level are shown. No significant deviation is observed in the shapes of the distributions. The larger uncertainties are observed in the central bins of the cross-correlation distributions, which are caused by the large migrations present in the inner bins. The correlation between the bins is estimated using the same technique as for the $\Delta|y|$ distribution in the charge asymmetry analysis (see Sec. 8.3.1). The bin-to-bin correlations between adjacent bins are typically between -0.9 and -0.7. The correlations between non-adjacent bins range from -0.4 to 0.6.

The signal modeling uncertainties are mostly the dominant source of uncertainties in the different measurements, followed by the statistical uncertainty (see Tab. 9.11). The signal modeling uncertainties are slightly reduced by performing the measurements in the fiducial region. The statistical uncertainties are in some cases slightly smaller than the signal modeling uncertainties. The cross-correlation observables show the larger uncertainties. This is expected due to the large migrations present in the response matrix. The background-related uncertainties are negligible and they are not shown in the table. The "Others" category corresponds to the MC sample size and the nonclosure uncertainties. Its contribution is minor for all the observables. The numbers shown for the "Detector" category correspond to the quadratic sum of the individual estimations obtained for the detector systematic uncertainties. The quadratic sum of the values in the various columns does not necessarily match with the total uncertainty, due to the marginalization procedure. Some nuisance parameters are constrained by the marginalization procedure. The largest constraint is observed on the electron identification scale factor. This nuisance parameter is pulled towards $+1\sigma$ for all the observables, as was observed in the charge asymmetry measurements. An example of the behavior of the nuisance parameters in the fiducial region and in the full phase space is shown in Fig. 9.10, which corresponds to the $C(r,r)$ observable. The mass un-

certainty is shown in a separated column and it is not added to the total uncertainty, in order to be able to improve the precision of the measurement when more precise masses are obtained. In the cases, where the dependency on the mass is compatible with zero, no uncertainty is quoted.

Figure 9.11 shows the predictions at 8 TeV calculated in Ref. [55] and the unfolded results. There is no significant deviation for any of the observables compared to the SM predictions. The largest deviation is approximately 1.5σ. The transverse correlation, which is measured for the first time, differs from the case of no spin correlation by 5.1 standard deviations. The correlations between the different observables is evaluated by performing pseudoexperiments on the nominal MC sample varied within its MC statistical uncertainty. The unfolding procedure with marginalization is applied for each observable and variation. The unfolded values of the observables are used to estimate the correlation between the measurements. The correlations at parton level are shown in Fig. 9.12. The highest correlations are found to be around 10% between the polarization and spin correlation along the helicity and r axes and between some cross correlations. The effect of the detector and background systematic uncertainties are taken into account in the correlation since the marginalization procedure is applied during the estimation.

The measurements of the observables presented in this analysis are compatible with other direct measurements in terms of observable values and uncertainties for the polarizations along the helicity and transverse axis as well as for the spin correlation along the helicity axis. Table 9.12 shows measurements of the polarizations and spin correlation performed by different experiments and using different techniques. The template fits for the polarization observables usually use the information of both the top and antitop quark decay chains. In this case, only one polarization value can be quoted as result and is shown for both columns of polarization along the same axis.

Observables	Central	Total	Statistical	Detector	Modeling	Others	Mass
Full phase space							
B^k_+	−0.044	±0.038	±0.018	±0.001	±0.026	±0.007	0.027
B^k_-	−0.064	±0.040	±0.020	±0.001	±0.023	±0.014	0.027
B^n_+	−0.018	±0.034	±0.020	±0.001	±0.024	±0.005	-
B^n_-	0.023	±0.042	±0.020	±0.001	±0.034	±0.005	-
B^r_+	0.039	±0.042	±0.026	±0.001	±0.029	±0.005	-
B^r_-	0.033	±0.054	±0.023	±0.002	±0.045	±0.006	0.016
$C(k,k)$	0.296	±0.093	±0.052	±0.006	±0.057	±0.011	0.037
$C(n,n)$	0.304	±0.060	±0.028	±0.001	±0.047	±0.001	0.010
$C(r,r)$	0.086	±0.144	±0.055	±0.005	±0.122	±0.016	0.038
$C(n,k)+C(k,n)$	−0.012	±0.128	±0.072	±0.005	±0.087	±0.029	-
$C(n,k)-C(k,n)$	−0.040	±0.087	±0.053	±0.004	±0.058	±0.003	-
$C(n,r)+C(r,n)$	0.117	±0.132	±0.070	±0.003	±0.102	±0.010	0.010
$C(n,r)-C(r,n)$	−0.006	±0.108	±0.069	±0.005	±0.070	±0.004	0.043
$C(r,k)+C(k,r)$	−0.261	±0.176	±0.083	±0.006	±0.135	±0.011	0.065
$C(r,k)-C(k,r)$	0.073	±0.192	±0.087	±0.007	±0.148	±0.005	0.025
Fiducial volume							
$3\langle\cos\theta^k_+\rangle$	0.125	±0.044	±0.018	±0.007	±0.025	±0.020	0.027
$3\langle\cos\theta^k_-\rangle$	0.119	±0.040	±0.022	±0.008	±0.021	±0.014	0.027
$3\langle\cos\theta^n_+\rangle$	−0.025	±0.042	±0.024	±0.001	±0.027	±0.005	-
$3\langle\cos\theta^n_-\rangle$	0.023	±0.046	±0.024	±0.001	±0.036	±0.006	-
$3\langle\cos\theta^r_+\rangle$	−0.104	±0.045	±0.027	±0.008	±0.030	±0.006	-
$3\langle\cos\theta^r_-\rangle$	−0.110	±0.060	±0.024	±0.008	±0.050	±0.010	0.015
$-9\langle\cos\theta^k_+\cos\theta^k_-\rangle$	0.172	±0.078	±0.041	±0.016	±0.050	±0.017	0.027
$-9\langle\cos\theta^n_+\cos\theta^n_-\rangle$	0.427	±0.079	±0.034	±0.011	±0.065	±0.004	0.027
$-9\langle\cos\theta^r_+\cos\theta^r_-\rangle$	0.031	±0.144	±0.055	±0.005	±0.124	±0.020	0.033
$-9\langle\cos\theta^n_+\cos\theta^k_- + \cos\theta^k_+\cos\theta^n_-\rangle$	0.024	±0.132	±0.078	±0.004	±0.085	±0.025	-
$-9\langle\cos\theta^n_+\cos\theta^k_- - \cos\theta^k_+\cos\theta^n_-\rangle$	−0.047	±0.096	±0.059	±0.004	±0.065	±0.002	-
$-9\langle\cos\theta^n_+\cos\theta^r_- + \cos\theta^r_+\cos\theta^n_-\rangle$	0.113	±0.143	±0.076	±0.005	±0.108	±0.023	0.015
$-9\langle\cos\theta^n_+\cos\theta^r_- - \cos\theta^r_+\cos\theta^n_-\rangle$	−0.030	±0.118	±0.076	±0.005	±0.077	±0.007	0.052
$-9\langle\cos\theta^r_+\cos\theta^k_- + \cos\theta^k_+\cos\theta^r_-\rangle$	−0.187	±0.151	±0.069	±0.023	±0.122	±0.006	0.039
$-9\langle\cos\theta^r_+\cos\theta^k_- - \cos\theta^k_+\cos\theta^r_-\rangle$	0.047	±0.128	±0.070	±0.003	±0.082	±0.010	0.023

Table 9.11 – *Results corrected to parton level in the full phase space and to stable-particle level in the fiducial phase-space. The central value with the total uncertainty is shown as well as the contribution from the various systematics categories. The uncertainty related to the top quark mass is presented separately. It is shown as "-" when considered negligible.*

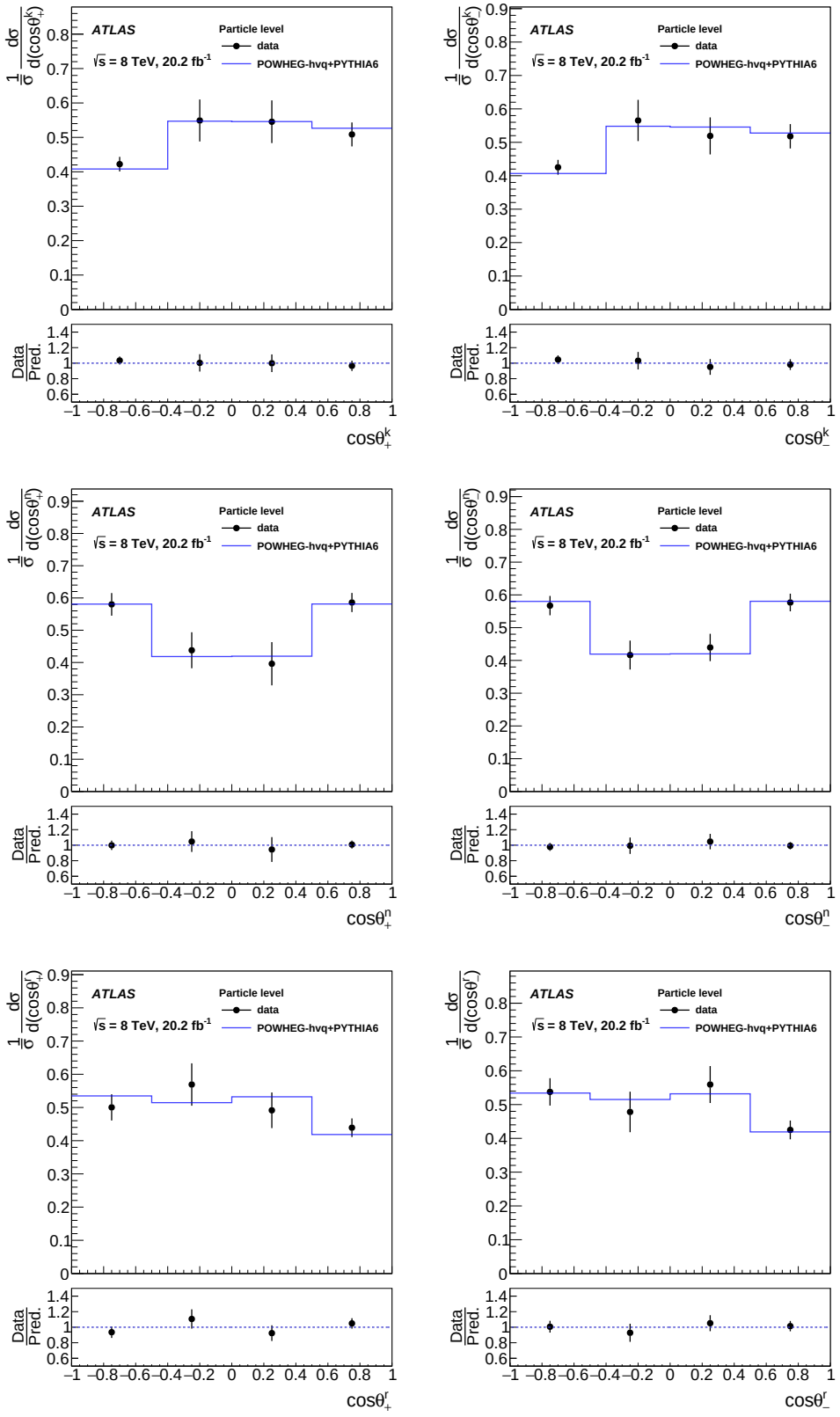


Figure 9.7 – *Comparison of the unfolded polarization distributions and the prediction from the signal MC for the stable-particle measurement. The total uncertainty is shown in each bin as well as the data/expectation ratio in the bottom pads.*

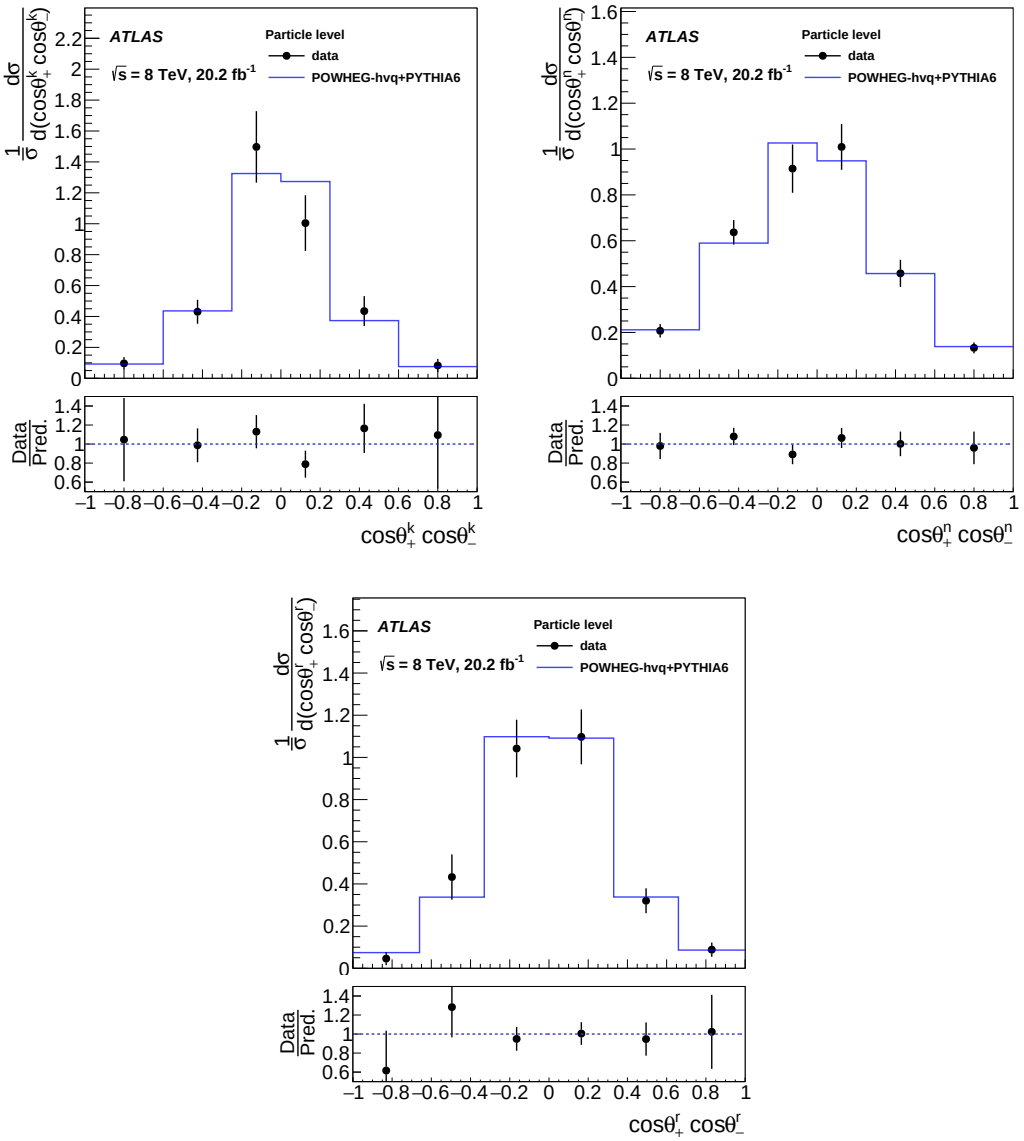


Figure 9.8 – *Comparison of the unfolded spin correlation distributions and the prediction from the signal MC for the stable-particle measurement. The total uncertainty is shown in each bin as well as the data/expectation ratio in the bottom pads.*

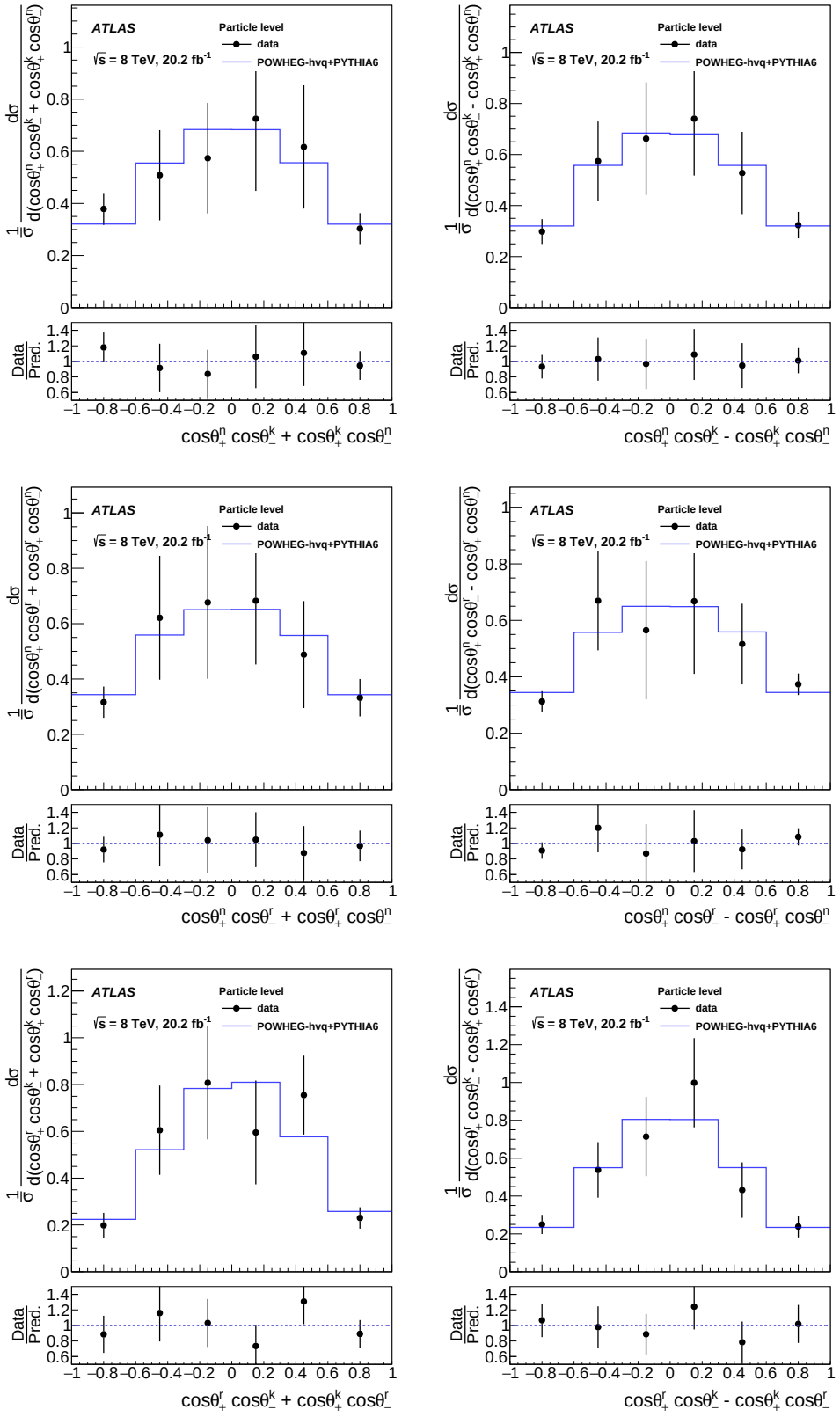


Figure 9.9 – *Comparison of the unfolded cross correlation distributions and the prediction from the signal MC for the stable-particle measurement. The total uncertainty is shown in each bin as well as the data/expectation ratio in the bottom pads.*

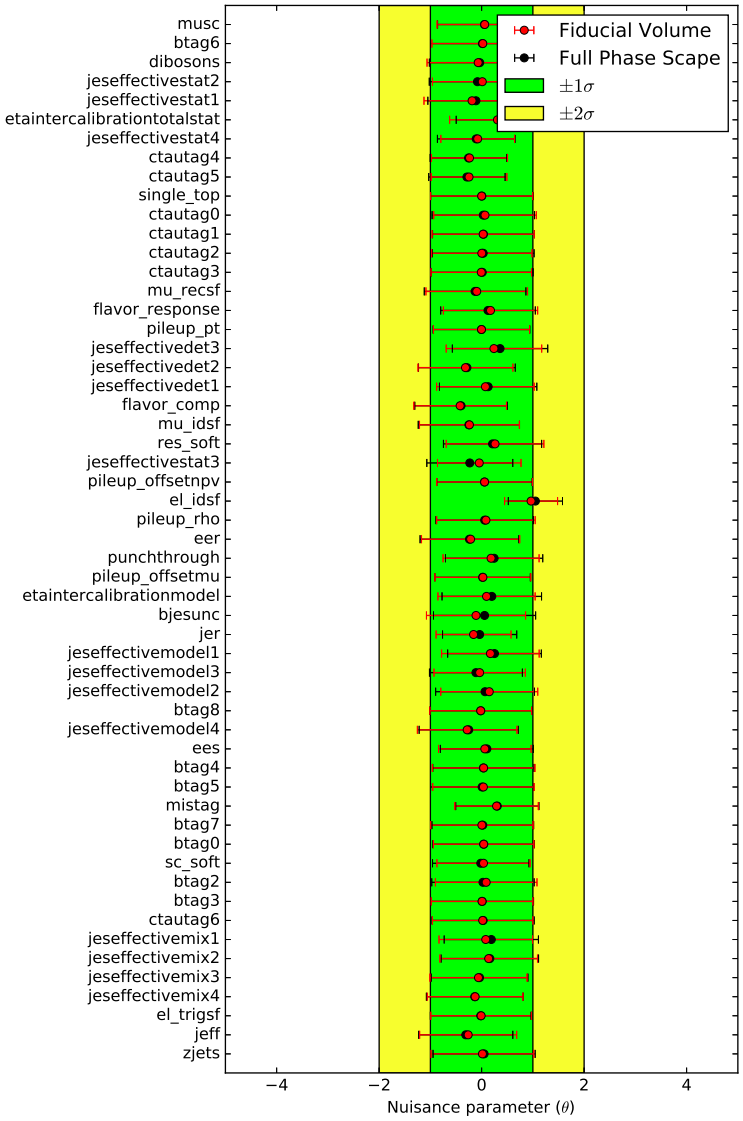


Figure 9.10 – *Nuisance parametes after the marginalization procedure for the $C(r,r)$ observable in the fiducial region (black) and in the full phase space (red).*

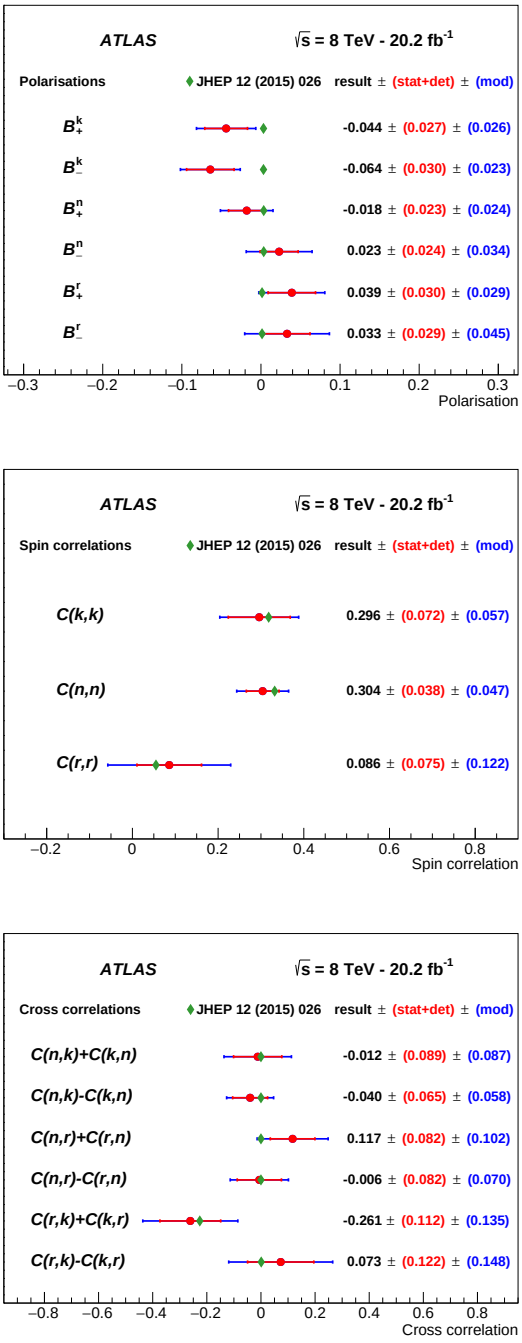


Figure 9.11 – *Comparison of the measured observables (data points) with predictions from the SM (diamonds) for the parton-level measurement. Inner bars indicate uncertainties obtained from the marginalization, outer bars indicate modeling systematic uncertainties, summed in quadrature. The widths of the diamonds are chosen for illustrative purposes only.*

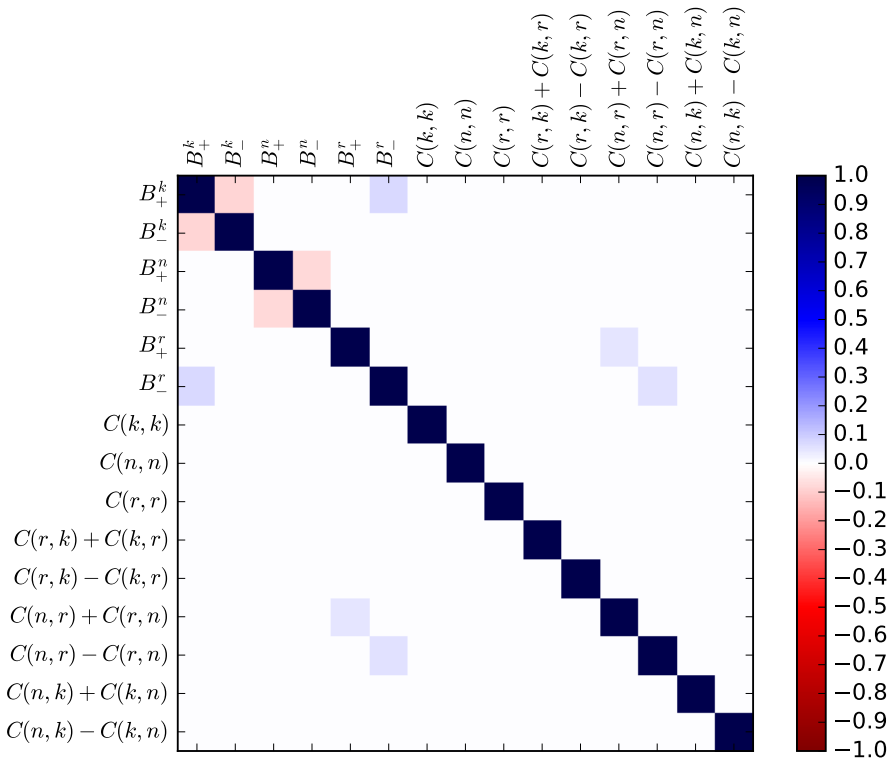


Figure 9.12 – *Correlations between the different observables after the unfolding procedure at parton level. The maximum correlations and anticorrelations are around 10%. They are showed in light blue and light red.*

Experiment	$\sqrt{s}$	Method	B^k_+	B^k_-	$C(k,k)$	B^n_+	B^n_-
ATLAS	8 TeV	Unfolding	-0.044 ± 0.038	-0.064 ± 0.040	0.296 ± 0.093	-0.018 ± 0.034	0.023 ± 0.042
CMS [160]	8 TeV	Unfolding	-0.022 ± 0.058		0.278 ± 0.084	-	-
ATLAS [161]	7 TeV	Template fit	-0.035 ± 0.040		-	-	-
ATLAS [162]	7 TeV	Template fit	-	-	0.23 ± 0.09	-	-
ATLAS [163]	7 TeV	Unfolding	-	-	0.315 ± 0.078	-	-
D0 [42]	1.96 TeV	Template fit	-0.102 ± 0.061		-	0.040 ± 0.034	

Table 9.12 – *Direct measurements of polarizations or spin correlations for different experiments and measurement techniques. The template fits for the polarisation observables usually use the information of both the top and antitop quark decay chains. In this case, only one polarisation value can be quoted as result and is shown for both columns of polarisation along the same axis. The SM predictions of the polarizations at the Tevatron are slightly different [42, 164] due to the different dominant production mechanism, which is $q\bar{q}$ annihilation. Dashes indicate no measurement of the observable for the corresponding analysis.*

9.4 Outlook on top spin measurements

The analysis presented in this work presents a precision test of the SM and opens a new window for exploring new physics using the top spin sector, since many of these observables, such as the polarization and correlations along the r axis, and the cross-correlation are measured for the first time. These measurements will encourage BSM predictions in the top spin sector, since these observables can be used to test the veracity of the proposed models.

Deeper studies and optimizations can be performed in the near future to help improve the precision. The MC signal modeling is shown to be a dominant source of systematic uncertainty. This is in part due to the conservative approach taken in the analysis. With more statistics in the MC simulation, a more precise estimation of the uncertainties can be obtained. In addition, more data would definitely improve the precision of these measurements. With the full dataset of Run 2, a more precise measurement can be performed. The polarizations along the three axes are expected to be measured soon in the semileptonic channel, which will give a more precise measurement and a possible combination with the results presented in this work. Also, template methods could be used in future analyses to improve the statistical uncertainty on the observables. This method could be performed simultaneously to the unfolding approach, since it provides information about the distributions which is not accessible by the template method.

This is the first time that spin density matrix has been fully characterized. The measurement and precision of these observables are of particular interest in effective field theory (EFT),[2] which is a model-independent way to search for new physics. EFT needs to use as many observables as possible to constrain simultaneously any deviation from the SM, which could detect new physics. These observables contribute considerably to the total amount of parameters used in EFT, since they are mostly uncorrelated. Also, new predictions are being produced in double higgs models motivated by this measurement, with will be possible to access with a collision energy of 13 TeV at the LHC [165].

[2] EFT is an extension of the SM where higher dimensional operators sensitive to new physics are considered.

Conclusions

Measurements of $t\bar{t}$ and leptonic charge asymmetry, as well as a full characterization of the spin density matrix have been presented. The measurements have been performed using dileptonic $t\bar{t}$ events from data recorded by the ATLAS detector during 2012 in proton-proton collisions at the LHC with a center-of-mass energy of 8 TeV.

Events were selected by requiring two oppositely charged leptons, two jets, and missing transverse momentum due to the neutrinos present in the final state. Two kinematic reconstruction methods were studied and compared in order to perform a good measurement: the Neutrino Weighting and KIN methods. The KIN method was used for the charge asymmetry measurements, while the Neutrino Weighting method was used for the top spin measurements. These algorithms were improved with respect to previous versions, and optimized for high performance on 8 TeV ATLAS data.

The $t\bar{t}$ system was reconstructed and the measurements were corrected for detector resolution and acceptance by using the Fully Bayesian Unfolding method, which was thoroughly optimized for these analyses. All measurements were corrected to particle level and parton level. At particle level, the measurements were performed in a fiducial region, while at parton level they were performed in the full phase space.

For the charge asymmetry analysis, inclusive and differential measurements were performed using the selected leptons and the reconstructed top and antitop quarks. The inclusive measurements in the full phase space were found to be:

$$A_{\mathrm{C}}^{\ell\ell} = 0.008 \pm 0.006, \quad A_{\mathrm{C}}^{t\bar{t}} = 0.021 \pm 0.016, \tag{9.2}$$

which are in agreement with the SM predictions. In the case of the differential measurements, the transverse momentum, mass, and boost of the $t\bar{t}$ system were used. No significant deviations with respect to the Standard Model predictions were found. The measurements were also found to be consistent with some BSM models. Since the measurements are statistically limited, more data will help to improve the precision and test the veracity of the BSM models. With the current precision, these measurements will gain extra precision in a possible combination with other measurements performed at the LHC.

For the top spin analysis, top-quark polarizations, spin correlations, and cross-correlations were measured. In total, 15 uncorrelated observables, each sensitive to individual elements of the spin density matrix were measured. No significant deviation from the SM expectation was observed. More data and larger MC samples can improve the precision of these measurements and will encourage the exploration of BSM models in the top spin sector. Many of these observables have never been investigated before and this is the first time that a full characterization of the spin density matrix is performed in the history of particle physics.

Appendices

A

Data/MC comparison

In this appendix the distribution of different kinematic variables are shown after event selection and kinematic reconstruction for data and expectations. The selection used in the comparison corresponds to the charge asymmetry analysis.

A.1 Distributions after event selection

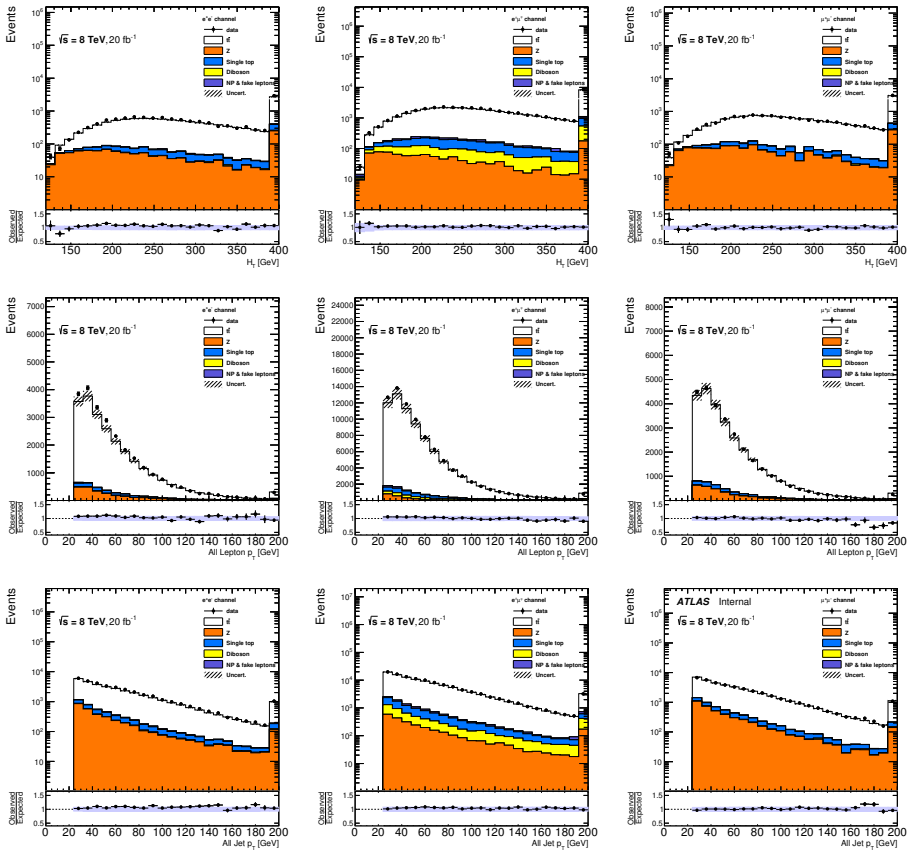


Figure A.1 – *Comparison of H_T distribution, lepton transverse momentum, and jet transverse momentum between data and MC predictions after event selection in the ee (left), eμ (middle) and μμ channels (right). The data/MC ratio is also shown. The shaded area shows the uncertainty on the signal and background normalization.*

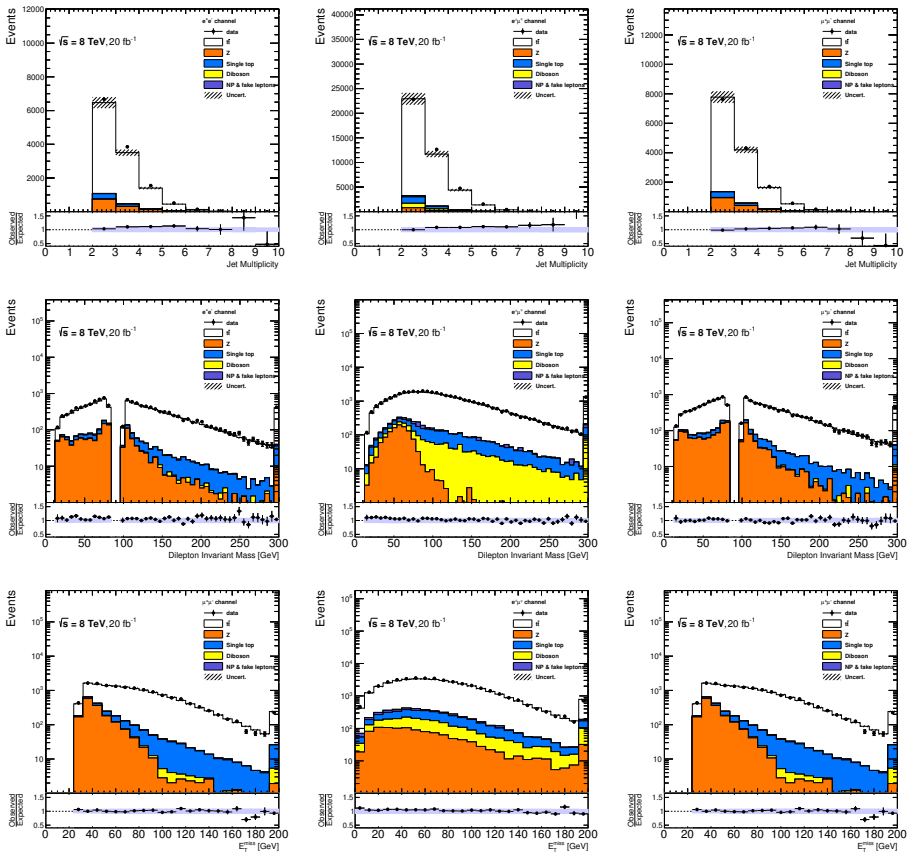


Figure A.2 – *Comparison of number of jets, dilepton invariant mass and E_T^{miss} distributions between data and MC predictions after event selection in the ee (left), eμ (middle) and μμ channels (right). The data/MC ratio is also shown. The shaded area shows the uncertainty on the signal and background normalization.*

A.2 Distributions after event reconstruction

A.2.1 Neutrino Weighting

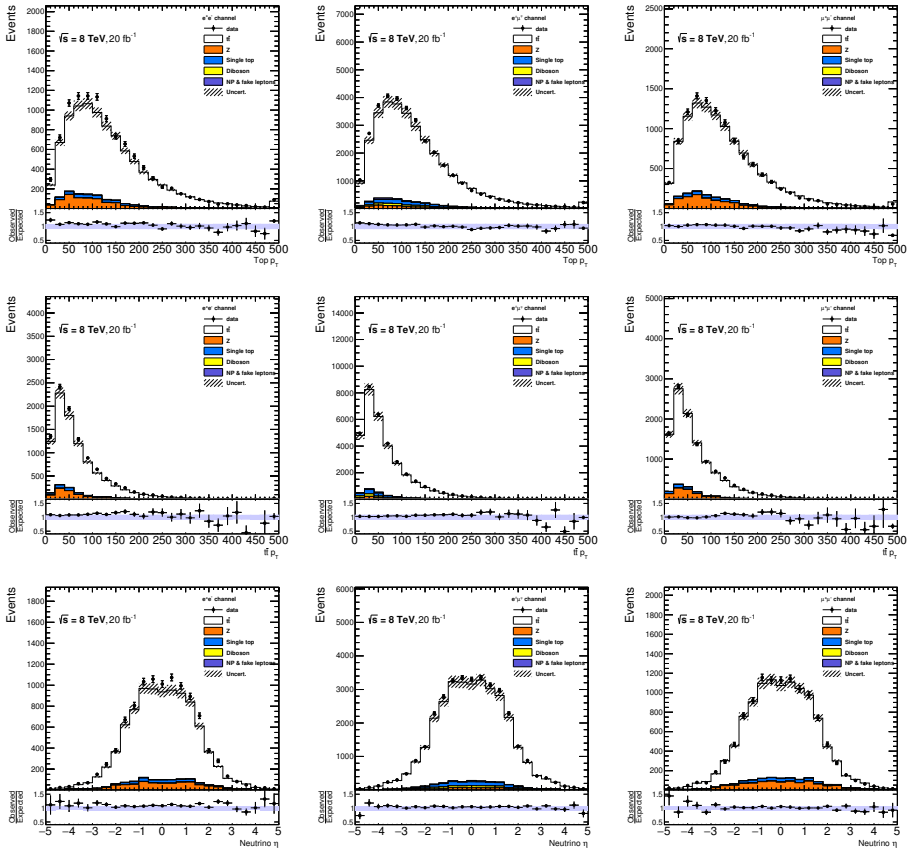


Figure A.3 – *Comparison of top p_T, $t\bar{t}$ p_T and neutrino-η distributions between data and MC predictions after event reconstruction using Neutrino Weighting in the ee (left), eμ (middle) and μμ channels (right). The data/MC ratio is also shown. The shaded area shows the uncertainty on the signal and background normalization.*

A.2.2 KIN Method

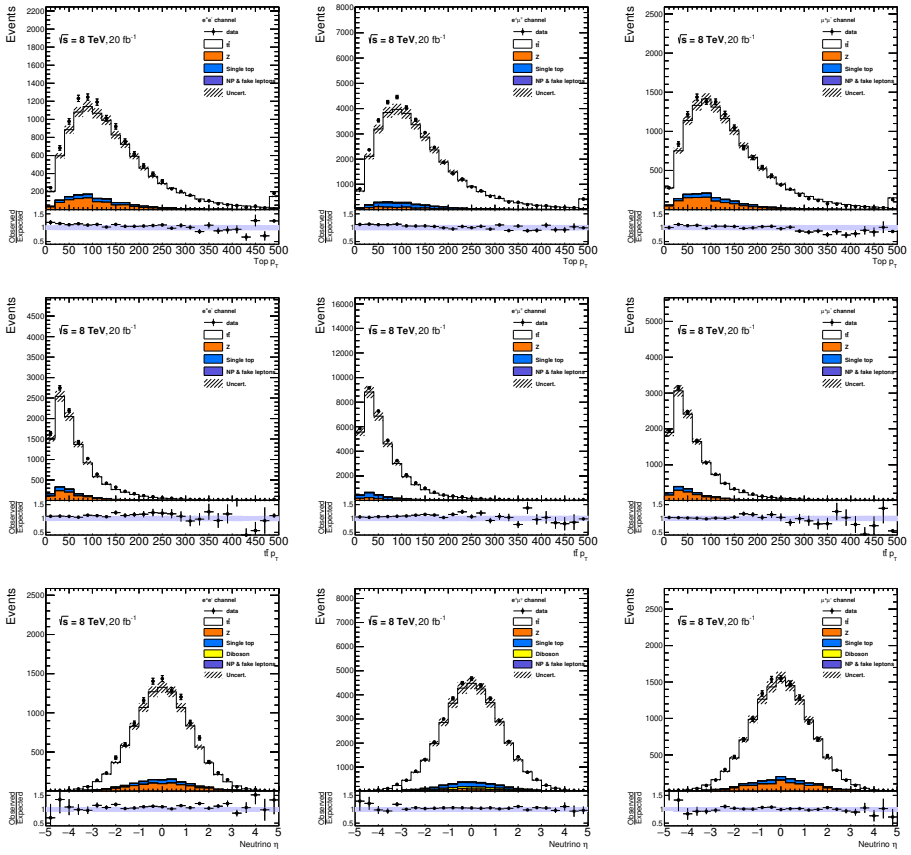


Figure A.4 – *Comparison of top p_T, $t\bar{t}$ p_T and neutrino-η distributions between data and MC predictions after event reconstruction using KIN method in the ee (left), eμ (middle) and μμ channels (right). The data/MC ratio is also shown. The shaded area shows the uncertainty on the signal and background normalization.*

B

Object Resolutions

In this appendix the resolution functions for jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ used during the kinematic reconstruction are presented.

B.0.1 Jet Resolution

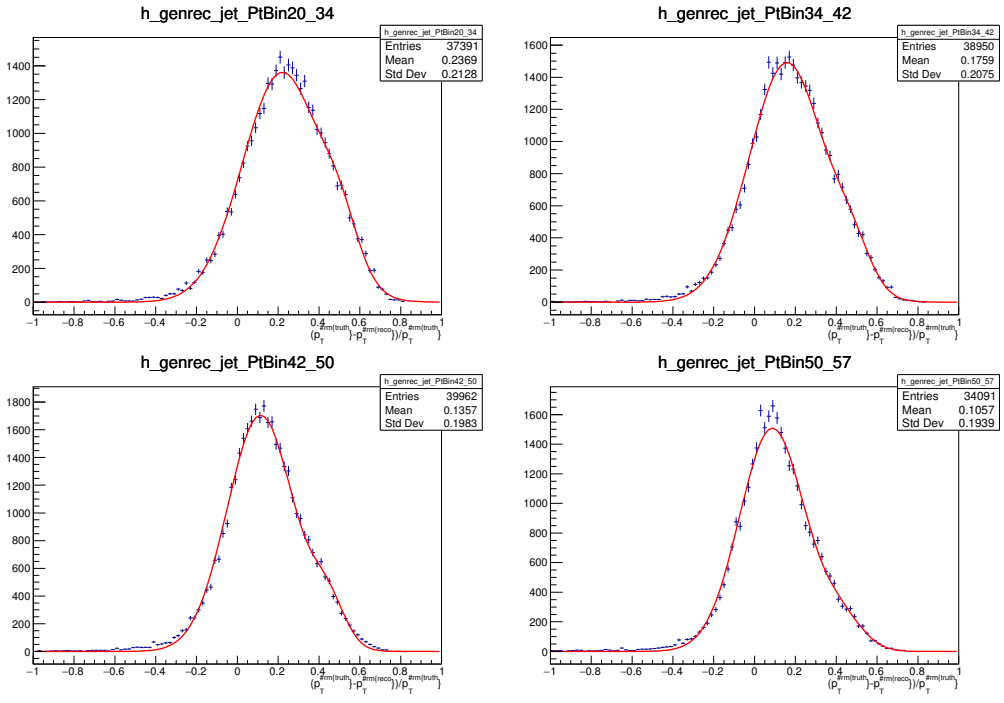


Figure B.1 – *Jet resolution function for p_{T} in the ranges of (top-left)* $[20,34]$, *(top-right)* $[34,42]$, *(bottom-left)* $[42,50]$, *and (bottom-right)* $[50,57]$. *The distributions are fit using a double Gaussian.*

B.0.2 $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution

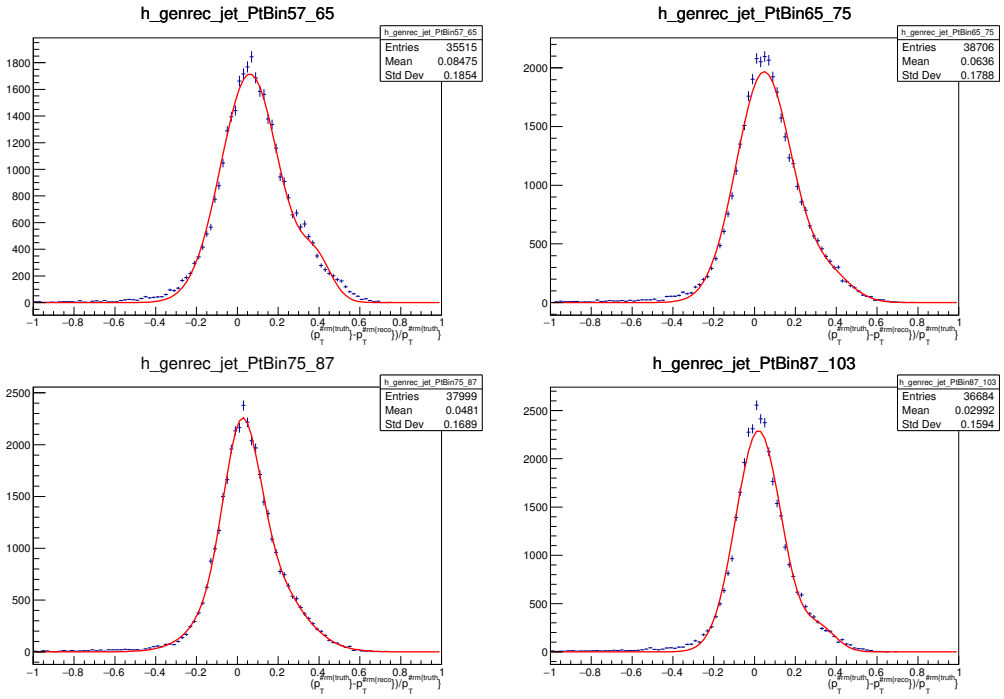


Figure B.2 – *Jet resolution function for p_{T} in the ranges of (top-left)* [57,65], *(top-right)* [65,75], *(bottom-left)* [75,87], *and (bottom-right)* [87,103]. *The distributions are fit using a double Gaussian.*

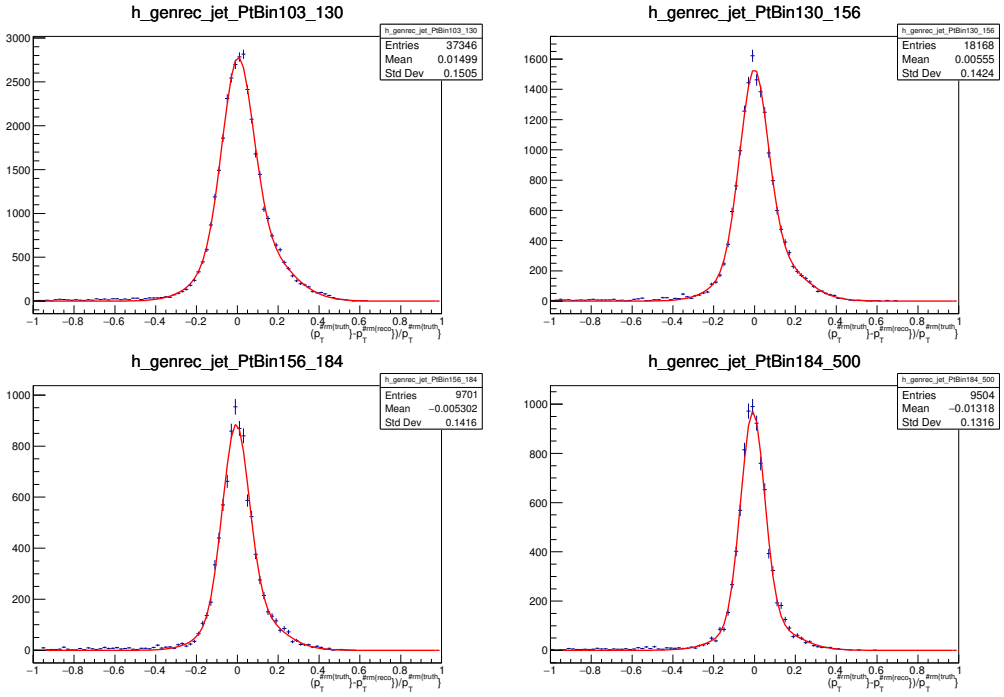


Figure B.3 – *Jet resolution function for p_{T} in the ranges of (top-left)* [103,130], *(top-right)* [130,156], *(bottom-left)* [156,184], *and (bottom-right)* [184,500]. *The distributions are fit using a double Gaussian.*

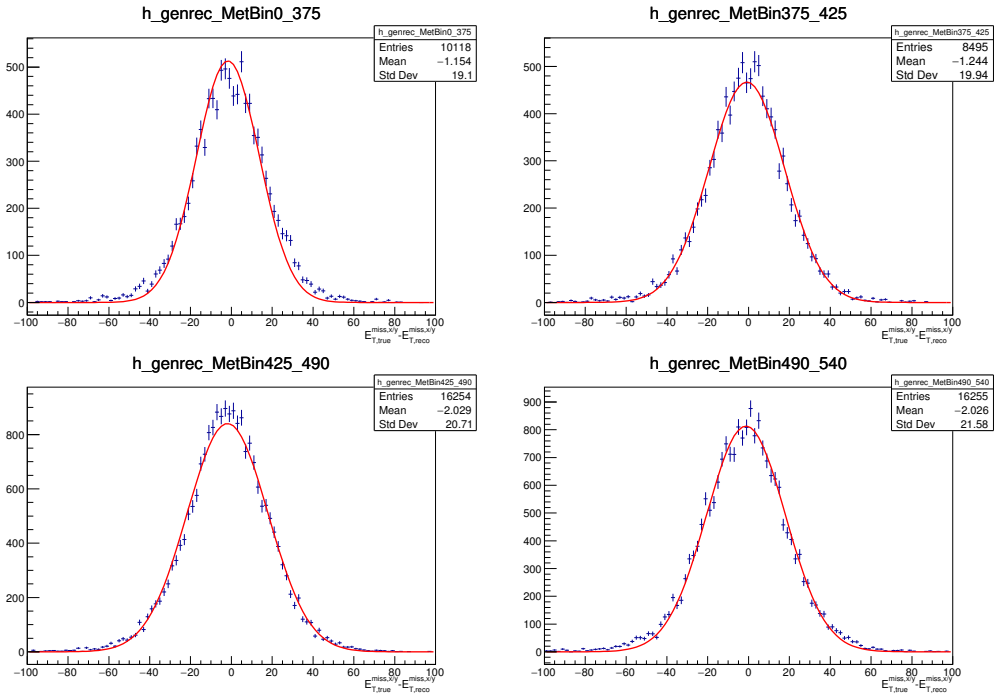


Figure B.4 – E_T^{miss} *resolution function for* E_T^{miss} *in the ranges of (top-left)* $[0, 375]$, *(top-right)* $[375, 425]$, *(bottom-left)* $[425, 490]$, *and (bottom-right)* $[490, 540]$. *The distributions are fit using a Gaussian.*

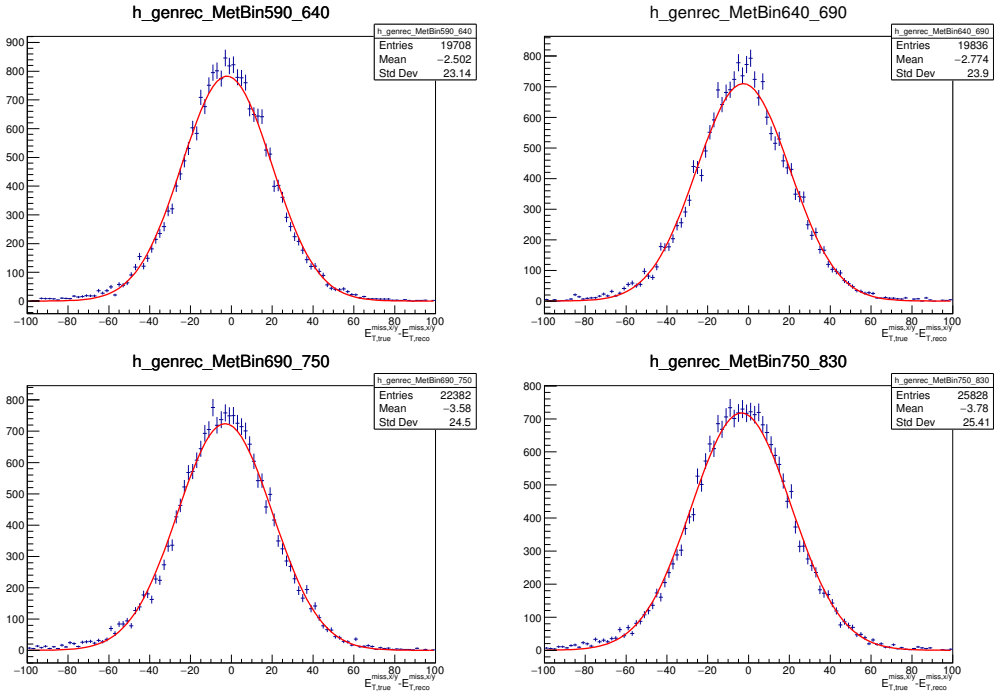


Figure B.5 – E_T^{miss} *resolution function for* E_T^{miss} *in the ranges of (top-left)* $[540, 640]$, *(top-right)* $[640, 690]$, *(bottom-left)* $[690, 750]$, *and (bottom-right)* $[750, 830]$. *The distributions are fit using a Gaussian.*

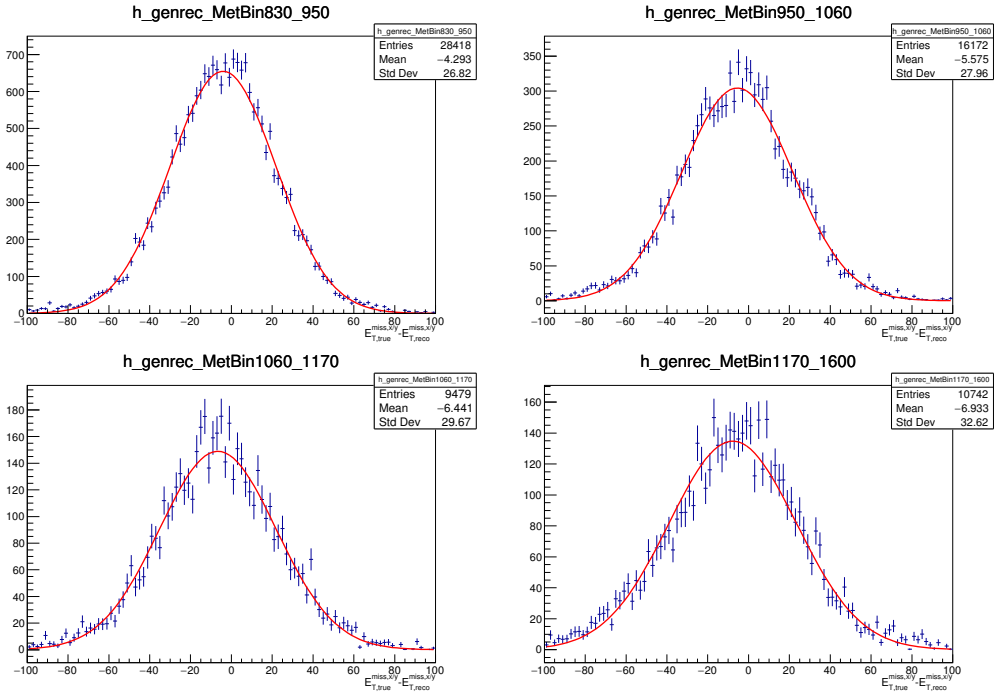


Figure B.6 – E_T^{miss} *resolution function for* E_T^{miss} *in the ranges of (top-left)* $[830, 950]$, *(top-right)* $[950, 1060]$, *(bottom-left)* $[1060, 1170]$, *and (bottom-right)* $[750, 830]$. *The distributions are fit using a Gaussian.*

C

Detector systematic uncertainties

The estimation of the individual detector systematic uncertainties contribution for the differential charge asymmetry measurements is presented in this appendix.

Source of systematic uncertainties	Fiducial Region		Full phase space	
	0 – 500 GeV	500 – 2000 GeV	0 – 500 GeV	500 – 2000 GeV
Missing transverse momentum	0.00286	0.00131	0.00208	0.00093
Lepton energy scale and resolution	0.00550	0.00189	0.00350	0.00260
Lepton reconstruction/identification	0.00022	0.00032	0.00063	0.00014
Jet reconstruction efficiency	0.00191	0.00287	0.00195	0.00312
B-tagging/mis-tag efficiency	0.00157	0.00074	0.00161	0.00079
Jet energy scale and resolution	0.02261	0.00592	0.01995	0.00470
Total Uncertainty	0.02358	0.00701	0.02053	0.00634

Table C.1 – *Estimation of all detector modeling systematic sources for the each of the studied bins in the measurement of the $A_C^{t\bar{t}}$ observable as a function of the mass of the $t\bar{t}$ system.*

Source of systematic uncertainties	Fiducial Region		Full phase space	
	0 – 500 GeV	500 – 2000 GeV	0 – 500 GeV	500 – 2000 GeV
Missing transverse momentum	0.00049	0.00041	0.00034	0.00041
Lepton energy scale and resolution	0.00034	0.00059	0.00035	0.00123
Lepton reconstruction/identification	0.00018	0.00023	0.00016	0.00021
Jet reconstruction efficiency	0.00049	0.00096	0.00055	0.00121
B-tagging/mis-tag efficiency	0.00054	0.00055	0.00037	0.00048
Jet energy scale and resolution	0.00220	0.00328	0.00222	0.00341
Total Uncertainty	0.00240	0.00354	0.00237	0.00388

Table C.2 – *Estimation of all detector modeling systematics sources for the each of the studied bins in the measurement of the $A_C^{\ell\ell}$ observable as a function of the mass of the $t\bar{t}$ system.*

Source of systematic uncertainties	Fiducial Region		Full phase space	
	0 – 30 GeV	30 – 1000 GeV	0 – 30 GeV	30 – 1000 GeV
Missing transverse momentum	0.00697	0.00204	0.00055	0.00263
Lepton energy scale and resolution	0.00685	0.00537	0.00732	0.00652
Lepton reconstruction/identification	0.00122	0.00063	0.00064	0.00029
Jet reconstruction efficiency	0.00208	0.00136	0.00133	0.00086
B-tagging/mis-tag efficiency	0.00166	0.00120	0.00109	0.00051
Jet energy scale and resolution	0.01622	0.01336	0.01315	0.01031
Total Uncertainty	0.01916	0.01467	0.01517	0.01252

Table C.3 – *Estimation of all detector modeling systematics sources for the each of the studied bins in the measurement of the $A_C^{t\bar{t}}$ observable as a function of the p_T of the $t\bar{t}$ system.*

Source of systematic uncertainties	Fiducial Region		Full phase space	
	0 – 30 GeV	30 – 1000 GeV	0 – 30 GeV	30 – 1000 GeV
Missing transverse momentum	0.00089	0.00003	0.00064	0.00015
Lepton energy scale and resolution	0.00346	0.00061	0.00362	0.00085
Lepton reconstruction/identification	0.00030	0.00019	0.00023	0.00017
Jet reconstruction efficiency	0.00340	0.00202	0.00320	0.00178
B-tagging/mis-tag efficiency	0.00127	0.00049	0.00055	0.00047
Jet energy scale and resolution	0.00798	0.00283	0.00854	0.00309
Total Uncertainty	0.00947	0.00357	0.00985	0.00370

Table C.4 – *Estimation of all detector modeling systematics sources for the each of the studied bins in the measurement of the $A_C^{\ell\ell}$ observable as a function of the p_T of the $t\bar{t}$ system.*

	Fiducial Region		Full phase space	
Source of systematic uncertainties	0.0 – 0.6	0.6 – 1.0	0.0 – 0.6	0.6 – 1.0
Missing transverse momentum	0.00352	0.00190	0.00304	0.00223
Lepton energy scale and resolution	0.00291	0.00400	0.00291	0.00393
Lepton reconstruction/identification	0.00036	0.00028	0.00053	0.00028
Jet reconstruction efficiency	0.00244	0.00225	0.00277	0.00214
B-tagging/mis-tag efficiency	0.00120	0.00066	0.00163	0.00073
Jet energy scale and resolution	0.02066	0.00779	0.01823	0.00696
Total Uncertainty	0.02134	0.00927	0.01899	0.00861

Table C.5 – *Estimation of all detector modelling systematics sources for the each of the studied bins in the measurement of the $A_C^{t\bar{t}}$ observable as a function of the boost of the $t\bar{t}$ system.*

	Fiducial Region		Full phase space	
Source of systematic uncertainties	0.0 – 0.6	0.6 – 1.0	0.0 – 0.6	0.6 – 1.0
Missing transverse momentum	0.00025	0.00092	0.00026	0.00076
Lepton energy scale and resolution	0.00066	0.00118	0.00082	0.00089
Lepton reconstruction/identification	0.00015	0.00012	0.00021	0.00027
Jet reconstruction efficiency	0.00040	0.00116	0.00054	0.00102
B-tagging/mis-tag efficiency	0.00064	0.00030	0.00051	0.00044
Jet energy scale and resolution	0.00257	0.00460	0.00146	0.00283
Total Uncertainty	0.00278	0.00499	0.00186	0.00327

Table C.6 – *Estimation of all detector modeling systematics sources for the each of the studied bins in the measurement of the $A_C^{\ell\ell}$ observable as a function of the boost of the $t\bar{t}$ system.*

D

Signal modeling uncertainties

In this appendix all the modeling systematic contributions corresponding to the differential measurements of the charge asymmetry are shown for the fiducial volume and full phase space. The technique used for the estimation is explained in Sec. 7.2.

D.0.1 Uncertainties from the bootstrapping

	$A_C^{\ell\ell}$	$m_{t\bar{t}}$ [GeV]		$p_T^{t\bar{t}}$ [GeV]	
	Inclusive	0 – 500	500 – 2000	0 – 30	30 – 1000
MC generator	0.00276 ± 0.00185	0.00257 ± 0.00376	0.01109 ± 0.00480	0.01606 ± 0.00689	0.01249 ± 0.00436
ISR/FSR	0.00005 ± 0.00129	0.00176 ± 0.00262	0.00015 ± 0.00345	0.00168 ± 0.00464	0.00023 ± 0.00323
Parton Shower	0.00171 ± 0.00121	0.00172 ± 0.00220	0.00762 ± 0.00316	0.00222 ± 0.00417	0.00151 ± 0.00265

	$\beta_z^{t\bar{t}}$	
	0.0 – 0.6	0.6 – 1.0
MC generator	0.00185 ± 0.00336	0.00405 ± 0.00413
ISR/FSR	0.00070 ± 0.00239	0.00141 ± 0.00284
Parton Shower	0.00239 ± 0.00197	0.00136 ± 0.00271

Table D.1 – *Modelling uncertainties for the lepton asymmetry in the full phase space. The PDF uncertainties are not estimated using the bootstrapping procedure since the PDF variations are performed by reweighting the sample.*

	$A_C^{\ell\ell}$	$m_{t\bar{t}}$ [GeV]		$p_T^{t\bar{t}}$ [GeV]	
	Inclusive	0 – 500	500 – 2000	0 – 30	30 – 1000
MC generator	0.00174 ± 0.00147	0.00211 ± 0.00315	0.00997 ± 0.00441	0.01360 ± 0.00593	0.01173 ± 0.00379
ISR/FSR	0.00119 ± 0.00092	0.00346 ± 0.00216	0.00125 ± 0.00323	0.00037 ± 0.00392	0.00227 ± 0.00287
Parton Shower	0.00043 ± 0.00087	0.00206 ± 0.00184	0.00716 ± 0.00294	0.00090 ± 0.00358	0.00172 ± 0.00233

	$\beta_z^{t\bar{t}}$	
	0.0 – 0.6	0.6 – 1.0
MC generator	0.00243 ± 0.00278	0.00190 ± 0.00357
ISR/FSR	0.00178 ± 0.00191	0.00142 ± 0.00250
Parton Shower	0.00240 ± 0.00158	0.00011 ± 0.00228

Table D.2 – *Modelling uncertainties for the lepton asymmetry in the fiducial region. The PDF uncertainties are not estimated using the bootstrapping procedure since the PDF variations are performed by reweighting the sample.*

	$A_C^{t\bar{t}}$	$m_{t\bar{t}}$ [GeV]		$p_T^{t\bar{t}}$ [GeV]	
	Inclusive	0 – 500	500 – 2000	0 – 30	30 – 1000
MC generator	0.00575 ± 0.00480	0.00134 ± 0.01128	0.01206 ± 0.00649	0.00732 ± 0.01328	0.01457 ± 0.00970
ISR/FSR	0.00268 ± 0.00346	0.00268 ± 0.00826	0.00238 ± 0.00483	0.00908 ± 0.00950	0.00030 ± 0.00764
Parton Shower	0.00023 ± 0.00292	0.00689 ± 0.00688	0.00794 ± 0.00404	0.00263 ± 0.00840	0.00082 ± 0.00599

	$\beta_z^{t\bar{t}}$	
	0.0 – 0.6	0.6 – 1.0
MC generator	0.00026 ± 0.00885	0.00987 ± 0.00752
ISR/FSR	0.00856 ± 0.00676	0.00620 ± 0.00540
Parton Shower	0.00087 ± 0.00561	0.00268 ± 0.00461

Table D.3 – *Modelling uncertainties for the top asymmetry in the full phase space. The PDF uncertainties are not estimated using the bootstrapping procedure since the PDF variations are performed by reweighting the sample.*

	$A_C^{t\bar{t}}$	$m_{t\bar{t}}$ [GeV]		$p_T^{t\bar{t}}$ [GeV]	
	Inclusive	0 – 500	500 – 2000	0 – 30	30 – 1000
MC generator	0.00276 ± 0.00185	0.00257 ± 0.00376	0.01109 ± 0.00480	0.01606 ± 0.00689	0.01249 ± 0.00436
ISR/FSR	0.00005 ± 0.00129	0.00176 ± 0.00262	0.00015 ± 0.00345	0.00168 ± 0.00464	0.00023 ± 0.00323
Parton Shower	0.00171 ± 0.00121	0.00172 ± 0.00220	0.00762 ± 0.00316	0.00222 ± 0.00417	0.00151 ± 0.00265

	$\beta_z^{t\bar{t}}$	
	0.0 – 0.6	0.6 – 1.0
MC generator	0.00185 ± 0.00336	0.00405 ± 0.00413
ISR/FSR	0.00070 ± 0.00239	0.00141 ± 0.00284
Parton Shower	0.00239 ± 0.00197	0.00136 ± 0.00271

Table D.4 – *Modelling uncertainties for the top asymmetry in the fiducial region. The PDF uncertainties are not estimated using the bootstrapping procedure since the PDF variations are performed by reweighting the sample.*

D.0.2 Summary Tables

	Fiducial Region		Full phase space	
Source of systematic uncertainties	0 – 500 GeV	500 – 2000 GeV	0 – 500 GeV	500 – 2000 GeV
MC generator	±0.01128	±0.01207	±0.01262	±0.01380
PDF	±0.00056	±0.00048	±0.00151	±0.00073
Fragmentation Model	±0.00688	±0.00794	±0.00956	±0.00727
ISR/FSR	±0.00826	±0.00483	±0.00913	±0.00469
Total	±0.01560	±0.01524	±0.01834	±0.01630

Table D.5 – *Sources of modeling systematic uncertainty for the $t\bar{t}$ asymmetry as a function of the mass of the $t\bar{t}$ system.*

	Fiducial Region		Full phase space	
Source of systematic uncertainties	0 – 500 GeV	500 – 2000 GeV	0 – 500 GeV	500 – 2000 GeV
MC generator	±0.00315	±0.00997	±0.00376	±0.01109
PDF	±0.00019	±0.00042	±0.00052	±0.00044
Fragmentation Model	±0.00206	±0.00717	±0.00220	±0.00762
ISR/FSR	±0.00346	±0.00323	±0.00262	±0.00345
Total	±0.00512	±0.01270	±0.00511	±0.01390

Table D.6 – *Sources of modeling systematic uncertainty for the leptonic asymmetry as a function of the mass of the $t\bar{t}$ system.*

	Fiducial Region		Full phase space	
Source of systematic uncertainties	0.0 – 0.6	0.6 – 1.0	0.0 – 0.6	0.6 – 1.0
MC generator	±0.00885	±0.00984	±0.01028	±0.01057
PDF	±0.00048	±0.00053	±0.00147	±0.00091
Fragmentation Model	±0.00561	±0.00461	±0.00638	±0.00465
ISR/FSR	±0.00854	±0.00619	±0.00838	±0.00547
Total	±0.01353	±0.01252	±0.01479	±0.01281

Table D.7 – *Sources of modeling systematic uncertainty for the $t\bar{t}$ asymmetry as a function of the boost of the $t\bar{t}$ system.*

	Fiducial Region		Full phase space	
Source of systematic uncertainties	0.0 – 0.6	0.6 – 1.0	0.0 – 0.6	0.6 – 1.0
MC generator	±0.00278	±0.00357	±0.00336	±0.00413
PDF	±0.00018	±0.00034	±0.00049	±0.00036
Fragmentation Model	±0.00240	±0.00228	±0.00239	±0.00271
ISR/FSR	±0.00191	±0.00250	±0.00239	±0.00284
Total	±0.00415	±0.00493	±0.00479	±0.00571

Table D.8 – *Sources of modeling systematic uncertainty for the leptonic asymmetry as a function of the boost of the $t\bar{t}$ system.*

	Fiducial Region		Full phase space	
Source of systematic uncertainties	0 – 30 GeV	30 – 1000 GeV	0 – 30 GeV	30 – 1000 GeV
MC generator	±0.01328	±0.01458	±0.01423	±0.01019
PDF	±0.00105	±0.00069	±0.00195	±0.00126
Fragmentation Model	±0.00840	±0.00599	±0.00858	±0.00609
ISR/FSR	±0.00950	±0.00764	±0.00974	±0.00778
Total	±0.01839	±0.01753	±0.01936	±0.01425

Table D.9 – *Sources of modeling systematic uncertainty for the $t\bar{t}$ asymmetry as a function of the p_T of the $t\bar{t}$ system.*

	Fiducial Region		Full phase space	
Source of systematic uncertainties	0 – 30 GeV	30 – 1000 GeV	0 – 30 GeV	30 – 1000 GeV
MC generator	±0.01359	±0.01172	±0.01607	±0.01250
PDF	±0.00034	±0.00025	±0.00104	±0.00048
Fragmentation Model	±0.00358	±0.00233	±0.00417	±0.00265
ISR/FSR	±0.00392	±0.00287	±0.00464	±0.00323
Total	±0.01460	±0.01229	±0.01727	±0.01319

Table D.10 – *Sources of modeling systematic uncertainty for the leptonic asymmetry as a function of the p_T of the $t\bar{t}$ system.*

D.0.3 PDF Uncertainties

The PDF uncertainty is estimated using the *envelop* method described in Sec.7.2, in the following the plots used for the its estimation in all charge asymmetry measurements are presented.

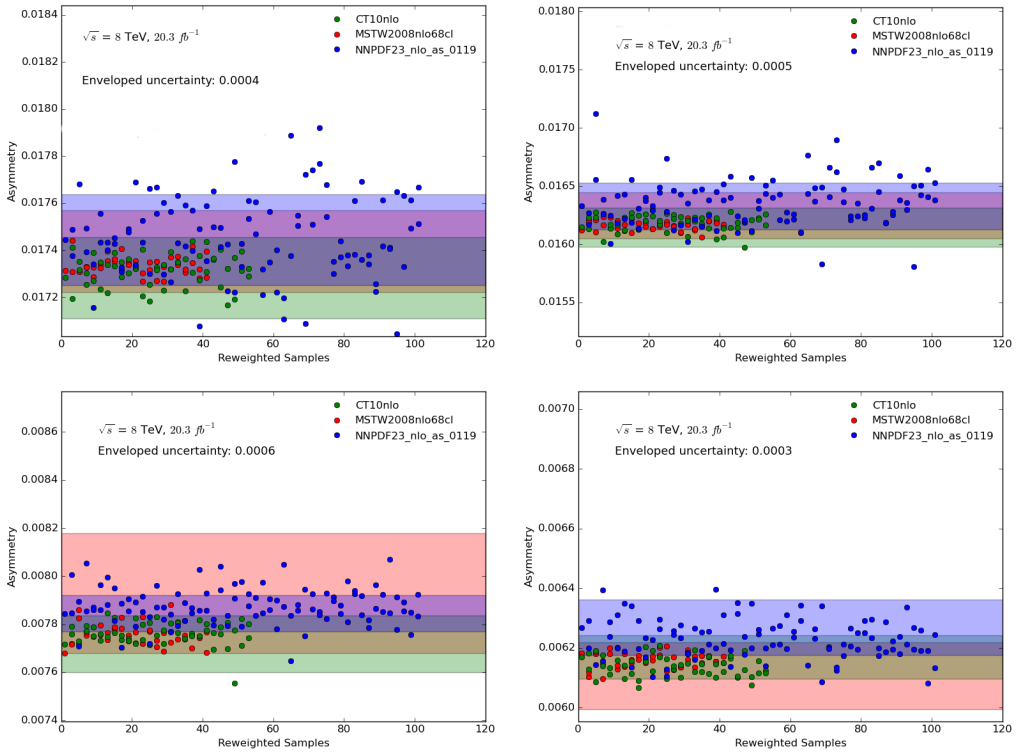


Figure D.1 – *Asymmetry value for all the reweighted samples used in the PDF uncertainty estimation. The top asymmetry is shown for the inclusive measurement in (top left) fiducial volume and (top right) full phase space. The lepton asymmetry is shown for the inclusive measurement in (bottom left) fiducial volume and (bottom right) full phase space.*

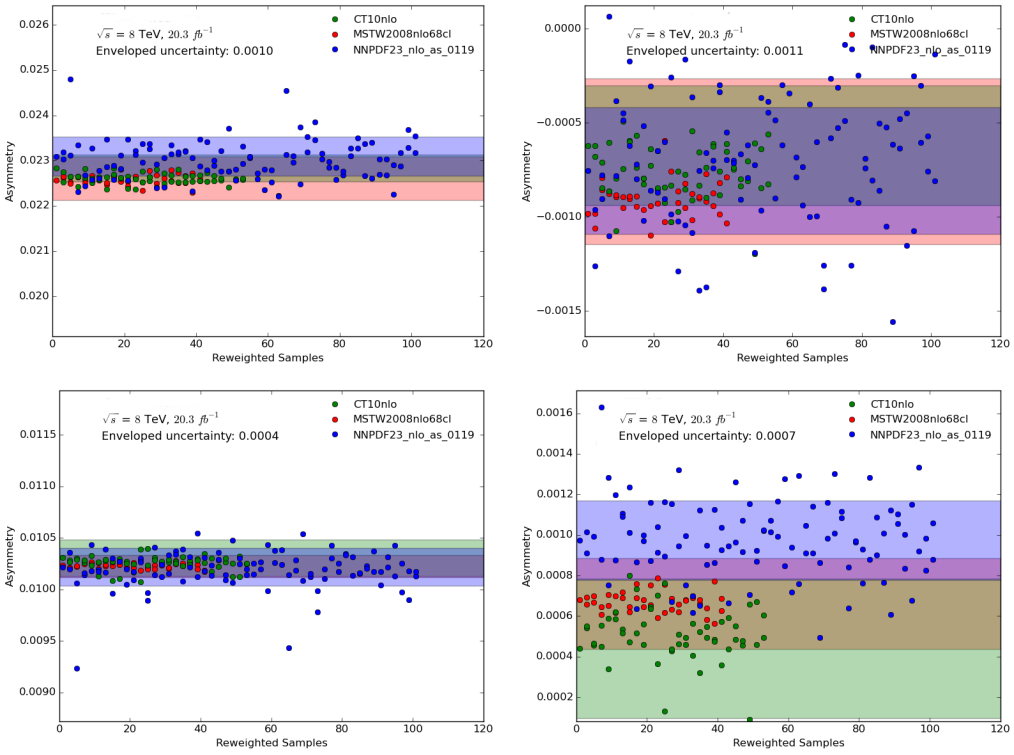


Figure D.2 – *Asymmetry value for all the reweighted samples used in the PDF uncertainty estimation. The top asymmetry as a function of the boost of the $t\bar{t}$ system is shown in the first (top left) and second bin (top right). The lepton asymmetry as a function of the boost of the $t\bar{t}$ system is shown in the first (bottom left) and second bin (bottom right).*

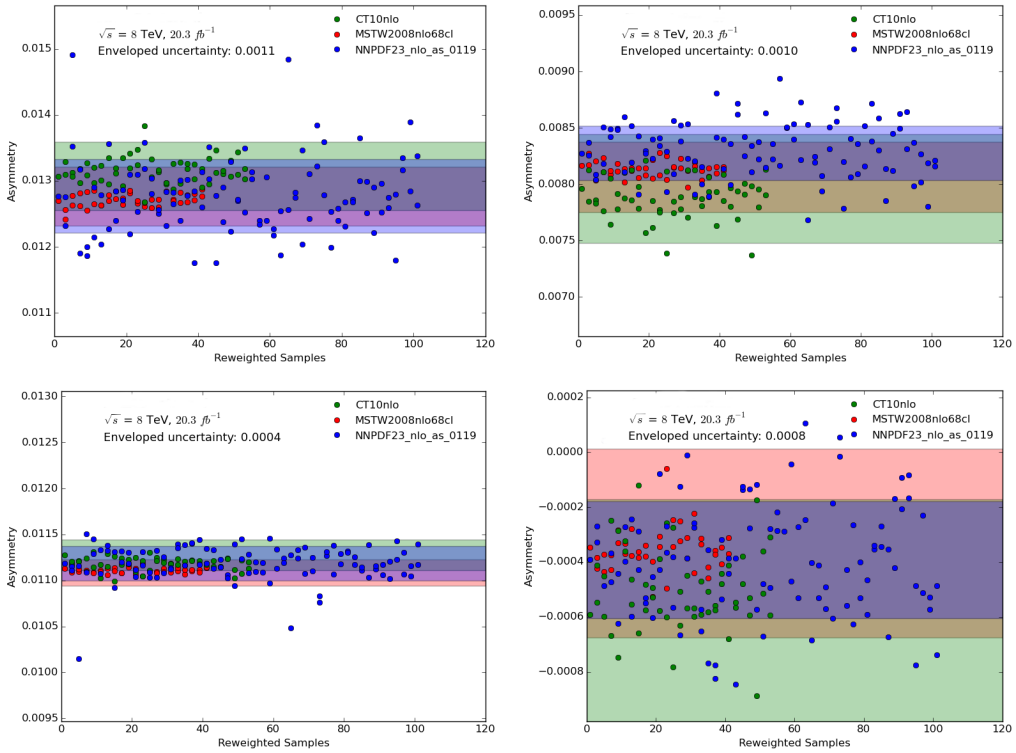


Figure D.3 – *Asymmetry value for all the reweighted samples used in the PDF uncertainty estimation. The top asymmetry as a function of the mass of the $t\bar{t}$ system is shown in the first (top-left) and second bin (top right). The lepton asymmetry as a function of the boost of the $t\bar{t}$ system is shown in the first (bottom-left) and second bin (bottom right).*

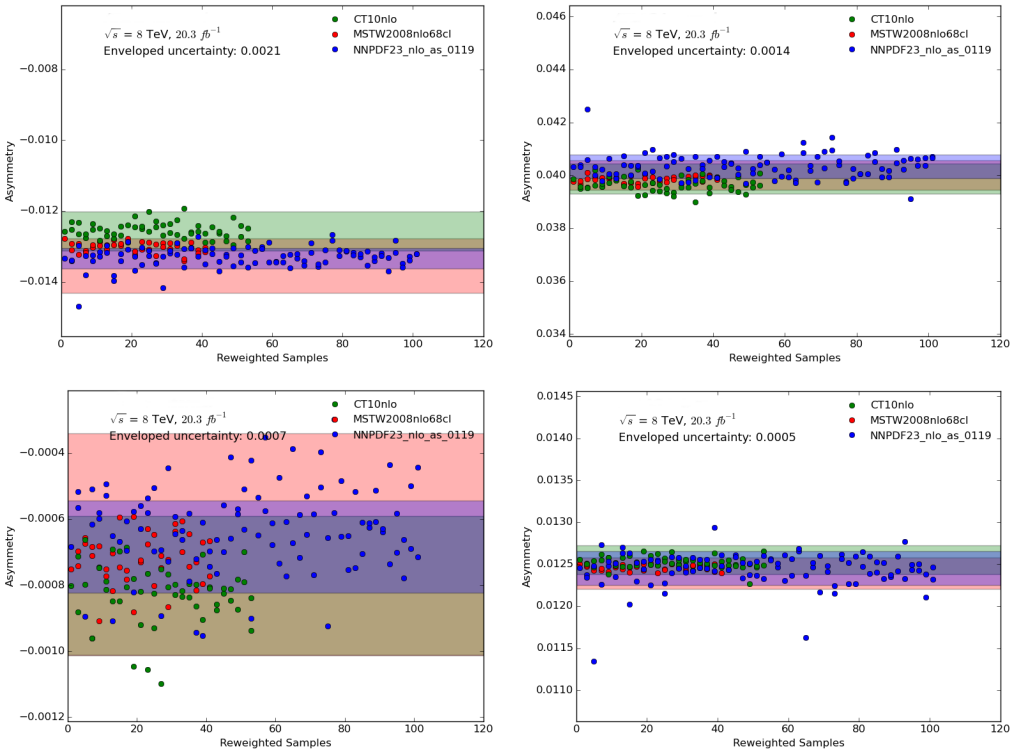


Figure D.4 – *Asymmetry value for all the reweighted samples used in the PDF uncertainty estimation. The top asymmetry as a function of the transverse momentum of the $t\bar{t}$ system is shown in the first (top left) and second bin (top right). The lepton asymmetry as a function of the boost of the $t\bar{t}$ system is shown in the first (bottom left) and second bin (bottom right).*

E

Top spin observable distributions

This appendix presents the unfolded distributions at parton level of the different spin observable. A good agreement is observed between the unfolded results and the SM predictions generated by Powheg-hvq.

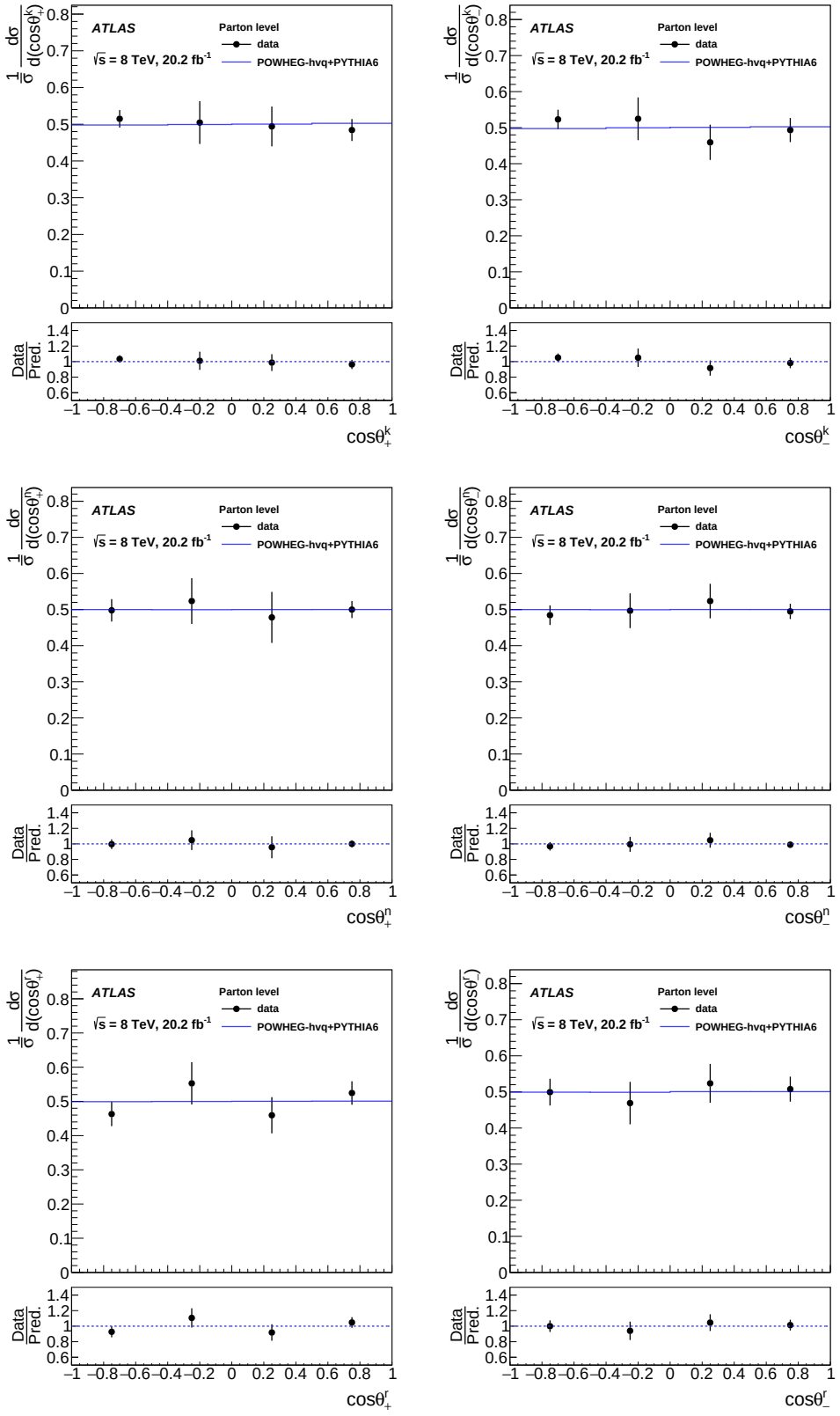


Figure E.1 – *Comparison of the unfolded polarisation distributions and the prediction from the signal MC for the stable-particle measurement. The total uncertainty is shown in each bin. The bin-to-bin correlations between adjacent bins are typically between -0.9 and -0.6. The correlations between non-adjacent bins range from -0.4 to 0.6.*

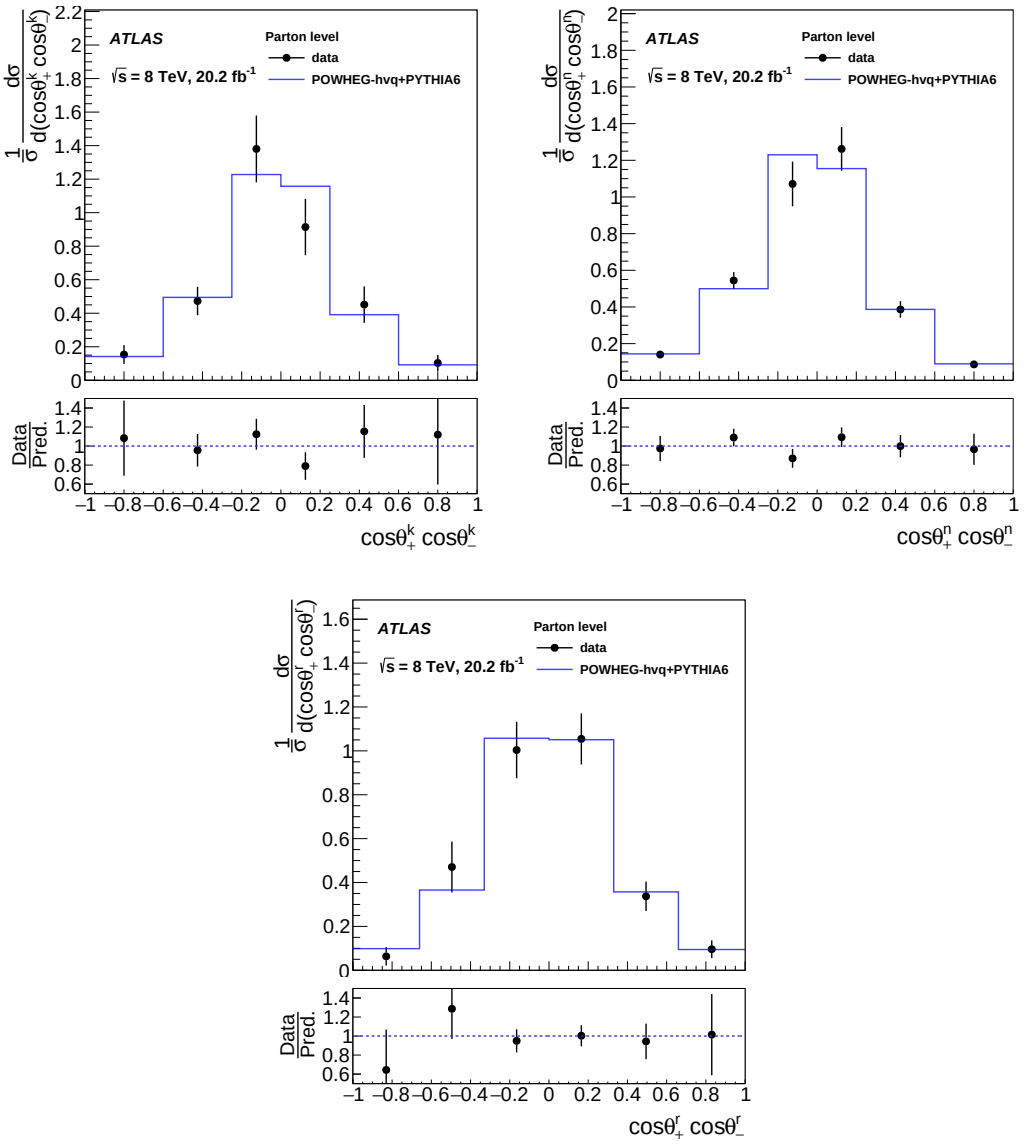


Figure E.2 – *Comparison of the unfolded spin correlation distributions and the prediction from the signal MC for the stable-particle measurement. The total uncertainty is shown in each bin. The bin-to-bin correlations between adjacent bins are typically between -0.9 and -0.4. The correlations between non-adjacent bins range from -0.4 to 0.6.*

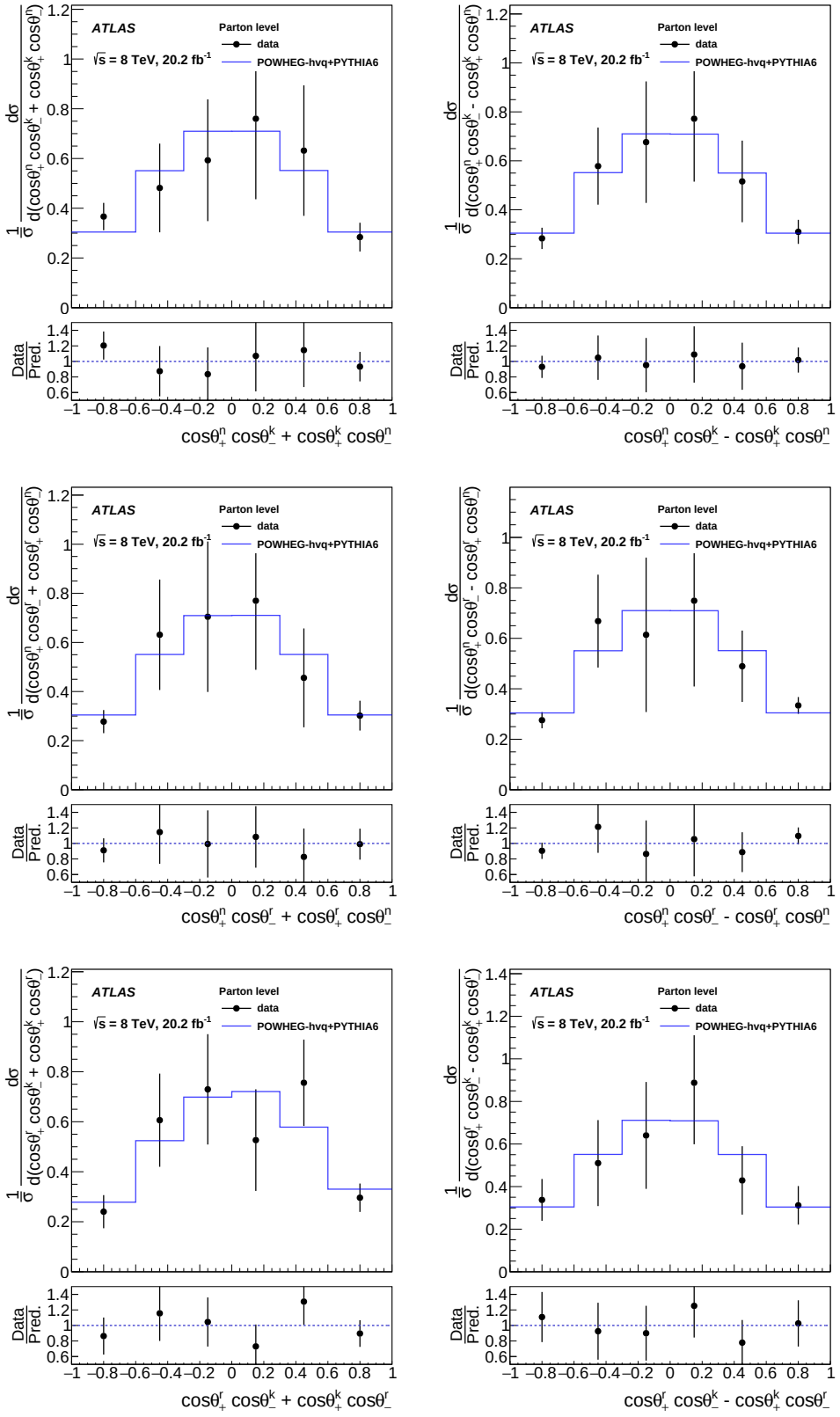


Figure E.3 – *Comparison of the unfolded cross correlation distributions and the prediction from the signal MC for the stable-particle measurement. The total uncertainty is shown in each bin. The bin-to-bin correlations between adjacent bins are typically between -0.9 and -0.7. The correlations between non-adjacent bins range from -0.4 to 0.6.*

Bibliography

[1] W. Marciano and H. Pagels, *Quantum chromodynamics*, Physics Reports **36** (1978) 137 – 276, http://www.sciencedirect.com/science/article/pii/0370157378902089.

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

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

[4] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C680519** (1968) 367–377.

[5] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson, *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105** (1957) 1413–1415, http://link.aps.org/doi/10.1103/PhysRev.105.1413.

[6] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509, http://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[7] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585–587, http://link.aps.org/doi/10.1103/PhysRevLett.13.585.

[8] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323, http://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[9] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1–29, arXiv:1207.7214 [hep-ex].

[10] C. Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30–61, arXiv:1207.7235 [hep-ex].

[11] V. C. Rubin, D. Burstein, W. K. Ford, Jr., and N. Thonnard, *Rotation velocities of 16 SA galaxies and a comparison of Sa, Sb, and SC rotation properties*, Astrophys. J. **289** (1985) 81.

[12] S. P. Martin, *A Supersymmetry primer*, arXiv:hep-ph/9709356 [hep-ph], [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

[13] H. Georgi and S. L. Glashow, *Unity of All Elementary-Particle Forces*, Phys. Rev. Lett. **32** (1974) 438–441, http://link.aps.org/doi/10.1103/PhysRevLett.32.438.

[14] R. N. Mohapatra and G. Senjanović, *Neutrino masses and mixings in gauge models with spontaneous parity violation*, Phys. Rev. D **23** (1981) 165–180, http://link.aps.org/doi/10.1103/PhysRevD.23.165.

[15] S. W. Herb et al., *Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions*, Phys. Rev. Lett. **39** (1977) 252–255, http://link.aps.org/doi/10.1103/PhysRevLett.39.252.

[16] D0 Collaboration, *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995) 2626–2631, arXiv:hep-ex/9503002 [hep-ex].

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

[18] A. D. Martin, W. Stirling, R. Thorne, and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C **63** (2009) 189–285, arXiv:0901.0002 [hep-ph].

[19] A. D. Martin, W. Stirling, R. Thorne, and G. Watt, *Uncertainties on alpha(S) in global PDF analyses and implications for predicted hadronic cross sections*, Eur. Phys. J. C **64** (2009) 653–680, arXiv:0905.3531 [hep-ph].

[20] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, Comput. Phys. Commun. **185** (2014) 2930, arXiv:1112.5675 [hep-ph].

[21] ATLAS and CMS Collaboration Collaboration,, *Combination of ATLAS and CMS top-quark pair cross section measurements using up to 1.1 fb-1 of data at 7 TeV*, Tech. Rep. ATLAS-CONF-2012-134, CERN, Geneva, Sep, 2012. http://cds.cern.ch/record/1478422.

[22] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b -tagged jets in pp collisions at $\sqrt{s}$ = 7 and 8 TeV with the ATLAS detector*, Eur. Phys. J. **C74** (2014) 3109, arXiv:1406.5375 [hep-ex], [Addendum: Eur. Phys. J.C76,no.11,642(2016)].

[23] CMS Collaboration, *Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s}$ = 8 TeV*, JHEP **02** (2014) 024, arXiv:1312.7582 [hep-ex], [Erratum: JHEP02,102(2014)].

[24] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s}$=13 TeV with the ATLAS detector*, Phys. Lett. **B761** (2016) 136–157, arXiv:1606.02699 [hep-ex].

[25] CMS Collaboration, V. Khachatryan et al., *Measurement of the $\mathrm{t\bar{t}}$ production cross section using events in the $e\mu$ final state in pp collisions at $\sqrt{s}$ = 13 TeV*, arXiv:1611.04040 [hep-ex].

[26] *ATLAS Summary Plots*, https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/TOP/. 2016.

[27] Particle Data Group Collaboration, K. A. Olive et al., *Review of Particle Physics*, Chin. Phys. **C38** (2014) 090001.

[28] ATLAS, CDF, CMS, D0 Collaborations, *First combination of Tevatron and LHC measurements of the top-quark mass*, arXiv:1403.4427 [hep-ex].

[29] ATLAS Collaboration, *Measurement of the W boson polarization in top quark decays with the ATLAS detector*, JHEP **06** (2012) 088, arXiv:1205.2484 [hep-ex].

[30] ATLAS Collaboration, *Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector*, Eur. Phys. J. **C76** (2016) 538, arXiv:1511.04716 [hep-ex].

[31] CMS Collaboration, *Measurement of the W boson helicity fractions in the decays of top quark pairs to lepton + jets final states produced in pp collisions at $\sqrt{s}$ = 8TeV*, Phys. Lett. **B762** (2016) 512–534, arXiv:1605.09047 [hep-ex].

[32] ATLAS Collaboration, *Search for the Associated Production of a Higgs Boson and a Top Quark Pair in Multilepton Final States with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2016-058, CERN, Geneva, Aug, 2016. http://cds.cern.ch/record/2206153.

[33] CMS Collaboration, *Search for a Standard Model Higgs Boson Produced in Association with a Top-Quark Pair and Decaying to Bottom Quarks Using a Matrix Element Method*, Eur. Phys. J. **C75** (2015) 251, arXiv:1502.02485 [hep-ex].

[34] CDF Collaboration, *Measurement of the top quark forward-backward production asymmetry and its dependence on event kinematic properties*, Phys. Rev. D **87** (2013) 092002, arXiv:1211.1003 [hep-ex].

[35] D0 Collaboration, *Measurement of the asymmetry in angular distributions of leptons produced in dilepton $t\bar{t}$ final states in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV*, Phys. Rev. D **88** (2013) 112002, http://link.aps.org/doi/10.1103/PhysRevD.88.112002.

[36] CMS Collaboration, *Measurements of the $t\bar{t}$ charge asymmetry using the dilepton decay channel in pp collisions at $\sqrt{s}$ = 7 TeV*, JHEP **04** (2014) 191, arXiv:1402.3803 [hep-ex].

[37] CMS Collaboration, *Inclusive and differential measurements of the $t\bar{t}$ charge asymmetry in proton-proton collisions at 7 TeV*, Phys. Lett. B **717** (2012) 129–150, arXiv:1207.0065 [hep-ex].

[38] CMS Collaboration, *Inclusive and differential measurements of the $t\bar{t}$ charge asymmetry in pp collisions at $\sqrt{s}$ = 8 TeV*, Phys. Lett. B **757** (2016) 154–179, arXiv:1507.03119 [hep-ex].

[39] ATLAS Collaboration, *Measurement of the charge asymmetry in dileptonic decays of top quark pairs in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, JHEP **05** (2015) 061, arXiv:1501.07383 [hep-ex].

[40] ATLAS Collaboration, *Measurement of the charge asymmetry in top-quark pair production in the lepton-plus-jets final state in pp collision data at $\sqrt{s}$ = 8 TeV with the ATLAS detector*, Eur. Phys. J. C **76** (2016) 87, arXiv:1509.02358 [hep-ex].

[41] CDF Collaboration, *Measurement of $t\bar{t}$ Spin Correlation in $p\bar{p}$ Collisions Using the CDF II Detector at the Tevatron*, Phys. Rev. **D83** (2011) 031104, arXiv:1012.3093 [hep-ex].

[42] D0 Collaboration, *Measurement of top quark polarization in $t\bar{t}$ lepton+jets final states*, arXiv:1607.07627 [hep-ex].

[43] ATLAS Collaboration, *Measurement of Top Quark Polarization in Top-Antitop Events from Proton-Proton Collisions at $\sqrt{s} = 7$ TeV Using the ATLAS Detector*, Phys.Rev.Lett. **111** (2013) 232002, arXiv:1307.6511 [hep-ex].

[44] ATLAS Collaboration, *Measurements of spin correlation in top-antitop quark events from proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Phys.Rev. **D90** (2014) 112016, arXiv:1407.4314 [hep-ex].

[45] CMS Collaboration, *Measurements of $t\bar{t}$ spin correlations and top-quark polarization using dilepton final states in pp collisions at $\sqrt{s} = 7$ TeV*, Phys.Rev.Lett. **112** (2014) 182001, arXiv:1311.3924 [hep-ex].

[46] CDF Collaboration, *Evidence for a Mass Dependent Forward-Backward Asymmetry in Top Quark Pair Production*, Phys. Rev. **D83** (2011) 112003, arXiv:1101.0034 [hep-ex].

[47] D0 Collaboration, *Forward-backward asymmetry in top quark-antiquark production*, Phys. Rev. **D84** (2011) 112005, arXiv:1107.4995 [hep-ex].

[48] W. Bernreuther and Z.-G. Si, *Top quark and leptonic charge asymmetries for the Tevatron and LHC*, Phys. Rev. D **86** (2012) 034026, arXiv:1205.6580 [hep-ph].

[49] D0 Collaboration, *Measurement of the forward-backward asymmetry in top quark-antiquark production in ppbar collisions using the lepton+jets channel*, Phys. Rev. **D90** (2014) 072011, arXiv:1405.0421 [hep-ex].

[50] CDF Collaboration, *Measurement of the inclusive leptonic asymmetry in top-quark pairs that decay to two charged leptons at CDF*, Phys. Rev. Lett. **113** (2014) 042001, arXiv:1404.3698 [hep-ex].

[51] M. Czakon, P. Fiedler, and A. Mitov, *Resolving the Tevatron Top Quark Forward-Backward Asymmetry Puzzle: Fully Differential Next-to-Next-to-Leading-Order Calculation*, Phys. Rev. Lett. **115** (2015) 052001, arXiv:1411.3007 [hep-ph].

[52] J. A. Aguilar-Saavedra and M. Perez-Victoria, *Asymmetries in t $\bar{t}$ production: LHC versus Tevatron*, Phys. Rev. **D84** (2011) 115013, arXiv:1105.4606 [hep-ph].

[53] F. Hubaut, E. Monnier, P. Pralavorio, K. Smolek, and V. Simak, *ATLAS sensitivity to top quark and W boson polarization in $t\bar{t}$ events*, Eur. Phys. J. **C44S2** (2005) 13–33, arXiv:hep-ex/0508061 [hep-ex].

[54] ATLAS Collaboration, *Measurement of Spin Correlation in Top-Antitop Quark Events and Search for Top Squark Pair Production in pp Collisions at $\sqrt{s} = 8$ TeV Using the ATLAS Detector*, Phys. Rev. Lett. **114** (2015) 142001, arXiv:1412.4742 [hep-ex].

[55] W. Bernreuther, D. Heisler, and Z.-G. Si, *A set of top quark spin correlation and polarization observables for the LHC: Standard Model predictions and new physics contributions*, JHEP **12** (2015) 026, arXiv:1508.05271 [hep-ph].

[56] CMS Collaboration, *Measurements of $t\bar{t}$ spin correlations and top-quark polarization using dilepton final states in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. Lett. **112** (2014) 182001, arXiv:1311.3924 [hep-ex].

[57] ATLAS Collaboration, *Measurement of Spin Correlation in Top-Antitop Quark Events and Search for Top Squark Pair Production in pp Collisions at $\sqrt{s} = 8$ TeV Using the ATLAS Detector*, Phys. Rev. Lett. **114** (2015) 142001, arXiv:1412.4742 [hep-ex].

[58] C. De Melis, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, https://cds.cern.ch/record/2197559, General Photo.

[59] *ATLAS Luminosity Summary Plots*, https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults. 2016.

[60] T. A. Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3** (2008) S08003, http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[61] ATLAS Collaboration Collaboration,, *ATLAS inner detector: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1997. https://cds.cern.ch/record/331063.

[62] ATLAS Collaboration Collaboration, S. Haywood, L. Rossi, R. Nickerson, and A. Romaniouk, *ATLAS inner detector: Technical Design Report, 2*. Technical Design Report ATLAS. CERN, Geneva, 1997. https://cds.cern.ch/record/331064.

[63] ATLAS Collaboration Collaboration, N. Wermes and G. Hallewel, *ATLAS pixel detector: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1998. https://cds.cern.ch/record/381263.

[64] ATLAS Collaboration, ATLAS Collaboration, *Operation and performance of the ATLAS semiconductor tracker*, JINST **9** (2014) P08009, arXiv:1404.7473 [hep-ex].

[65] T. A. T. collaboration, *The ATLAS TRT Barrel Detector*, Journal of Instrumentation **3** (2008) P02014, http://stacks.iop.org/1748-0221/3/i=02/a=P02014.

[66] ATLAS Collaboration, *The ATLAS TRT end-cap detectors*, Journal of Instrumentation **3** (2008) P10003, http://stacks.iop.org/1748-0221/3/i=10/a=P10003.

[67] ATLAS Collaboration, *Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data*, Eur. Phys. J. **C72** (2012) 1909, arXiv:1110.3174 [hep-ex].

[68] ATLAS Collaboration,, *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996. https://cds.cern.ch/record/331061.

[69] ATLAS Collaboration,, *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996. https://cds.cern.ch/record/331062.

[70] ATLAS Collaboration, *The ATLAS Forward Calorimeter*, Journal of Instrumentation **3** (2008) P02010, http://stacks.iop.org/1748-0221/3/i=02/a=P02010.

[71] ATLAS Collaboration,, *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997. https://cds.cern.ch/record/331068.

[72] Lasagni Manghi, Federico and ATLAS LUCID Collaboration, *The LUCID detector*, Tech. Rep. ATL-FWD-PROC-2015-003, CERN, Geneva, Jun, 2015. https://cds.cern.ch/record/2027724.

[73] P. Jenni, M. Nordberg, M. Nessi, and K. Jon-And, *ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity*. Technical Design Report ATLAS. CERN, Geneva, 2008. https://cds.cern.ch/record/1095847.

[74] W. Buttinger, *The ATLAS Level-1 Trigger System*, Journal of Physics: Conference Series **396** (2012) 012010, http://stacks.iop.org/1742-6596/396/i=1/a=012010.

[75] C. Gabaldon, *Performance of the ATLAS Trigger System*, Journal of Instrumentation **7** (2012) C01092, http://stacks.iop.org/1748-0221/7/i=01/a=C01092.

[76] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, and W. K. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012, arXiv:0201195 [hep-ph].

[77] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, *New parton distributions for collider physics*, Phys. Rev. D **82** (2010) 074024.

[78] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244–289, arXiv:1207.1303 [hep-ph].

[79] S. Höche, *Introduction to parton-shower event generators*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014) Boulder, Colorado, June 2-27, 2014*. 2014. arXiv:1411.4085 [hep-ph]. https://inspirehep.net/record/1328513/files/arXiv:1411.4085.pdf.

[80] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD matrix elements + parton showers*, JHEP **11** (2001) 063, arXiv:hep-ph/0109231 [hep-ph].

[81] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, *Matching matrix elements and shower evolution for top-quark production in hadronic collisions*, JHEP **01** (2007) 013, arXiv:hep-ph/0611129 [hep-ph].

[82] B. Isildak, *Measurement of the differential dijet production cross section in proton-proton collisions at $\sqrt{s} = 7$ tev*. PhD thesis, Bogazici U., 2011. arXiv:1308.6064 [hep-ex]. https://inspirehep.net/record/1251416/files/arXiv:1308.6064.pdf.

[83] T. Sjostrand, *The Lund Monte Carlo for Jet Fragmentation and e+ e- Physics: Jetset Version 6.2*, Comput. Phys. Commun. **39** (1986) 347–407.

[84] R. H. Swendsen and J.-S. Wang, *Nonuniversal critical dynamics in Monte Carlo simulations*, Phys. Rev. Lett. **58** (1987) 86–88.

[85] ATLAS Collaboration, *Measurement of the underlying event in jet events from 7 TeV proton-proton collisions with the ATLAS detector*, Eur. Phys. J. **C74** (2014) 2965, arXiv:1406.0392 [hep-ex].

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

[87] J. M. Butterworth, J. R. Forshaw, and M. H. Seymour, *Multiparton interactions in photoproduction at HERA*, Z. Phys. **C72** (1996) 637–646, arXiv:hep-ph/9601371 [hep-ph].

[88] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, *HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)*, JHEP **01** (2001) 010, arXiv:hep-ph/0011363 [hep-ph].

[89] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003) 001, arXiv:hep-ph/0206293 [hep-ph].

[90] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, *MadGraph 5 : Going Beyond*, JHEP **06** (2011) 128, arXiv:1106.0522 [hep-ph].

[91] S. Frixione, F. Stoeckli, P. Torrielli, B. R. Webber, and C. D. White, *The MCaNLO 4.0 Event Generator*, arXiv:1010.0819 [hep-ph].

[92] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv:0709.2092 [hep-ph].

[93] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. D **82** (2010) 074018, arXiv:1005.3457 [hep-ph].

[94] *First tuning of HERWIG/JIMMY to ATLAS data*, Tech. Rep. ATL-PHYS-PUB-2010-014, CERN, Geneva, Oct, 2010. http://cds.cern.ch/record/1303025.

[95] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250–303.

[96] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, ATL-PHYS-PUB-2010-013, http://cds.cern.ch/record/1300517, 2010. http://cds.cern.ch/record/1300517.

[97] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, and A. Salzburger, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, Tech. Rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar, 2007. https://cds.cern.ch/record/1020106.

[98] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987) 444–450.

[99] *Expected electron performance in the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2011-006, CERN, Geneva, Apr, 2011. https://cds.cern.ch/record/1345327.

[100] ATLAS Collaboration, M. Aaboud et al., *Electron efficiency measurements with the ATLAS detector using 2012 LHC proton-proton collision data*, arXiv:1612.01456 [hep-ex].

[101] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74** (2014) 3071, arXiv:1407.5063 [hep-ex].

[102] ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC protonproton collision data*, Eur. Phys. J. **C74** (2014) 3130, arXiv:1407.3935 [hep-ex].

[103] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter Clustering Algorithms: Description and Performance*, Tech. Rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr, 2008. http://cds.cern.ch/record/1099735.

[104] M. Aharrouche et al., *Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam*, Nucl. Instrum. Meth. **A614** (2010) 400–432.

[105] D. Banfi, M. Delmastro, and M. Fanti, *Cell response equalisation of the ATLAS electromagnetic calorimeter without the direct knowledge of the ionisation signals. Cell response equalization of the ATLAS electromagnetic calorimeter without the direct knowledge of the ionization signals*, http://cds.cern.ch/record/731840, Accepted as Scientific Note SN-ATLAS-2005-054.

[106] J. Colas et al., *Response Uniformity of the ATLAS Liquid Argon Electromagnetic Calorimeter*, Nucl. Instrum. Meth. **A582** (2007) 429–455, arXiv:0709.1094 [physics.ins-det].

[107] ATLAS Liquid Argon HEC Collaboration, B. Dowler et al., *Performance of the ATLAS hadronic end-cap calorimeter in beam tests*, Nucl. Instrum. Meth. **A482** (2002) 94–124.

[108] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

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

[110] ATLAS Collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 17, arXiv:1406.0076 [hep-ex].

[111] *Jet global sequential corrections with the ATLAS detector in proton-proton collisions at sqrt(s) = 8 TeV*, Tech. Rep. ATLAS-CONF-2015-002, CERN, Geneva, Mar, 2015. https://cds.cern.ch/record/2001682.

[112] *Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8\ TeV$*, Tech. Rep. ATLAS-CONF-2015-017, CERN, Geneva, Apr, 2015. https://cds.cern.ch/record/2008678.

[113] ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded in 2010 with the ATLAS detector*, Eur. Phys. J. **C73** (2013) 2306, arXiv:1210.6210 [hep-ex].

[114] ATLAS Collaboration Collaboration,, *Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data*, Tech. Rep. ATLAS-CONF-2010-042, CERN, Geneva, Jul, 2010. https://cds.cern.ch/record/1277682.

[115] ATLAS Collaboration Collaboration,, *Performance of Impact Parameter-Based b-tagging Algorithms with the ATLAS Detector using Proton-Proton Collisions at $\sqrt{s} = 7$ TeV*, Tech. Rep. ATLAS-CONF-2010-091, CERN, Geneva, Oct, 2010. https://cds.cern.ch/record/1299106.

[116] G. Piacquadio and C. Weiser, *A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS*, Journal of Physics: Conference Series **119** (2008) 032032, http://stacks.iop.org/1742-6596/119/i=3/a=032032.

[117] *Calibration of b-tagging using dileptonic top pair events in a combinatorial likelihood approach with the ATLAS experiment*, Tech. Rep. ATLAS-CONF-2014-004, CERN, Geneva, Feb, 2014. https://cds.cern.ch/record/1664335.

[118] ATLAS Collaboration, *Calibration of the performance of b-tagging for c and light-flavour jets in the 2012 ATLAS data*, ATLAS-CONF-2014-046, https://cds.cern.ch/record/1741020.

[119] ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV with ATLAS*, Eur. Phys. J. C **72** (2012) 1844, arXiv:1108.5602 [hep-ex].

[120] ATLAS Collaboration, *Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s}$ = 8 TeV proton-proton collisions in the ATLAS detector*, arXiv:1609.09324 [hep-ex].

[121] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector*, arXiv:1510.03823 [hep-ex].

[122] ATLAS Collaboration, *Performance of the ATLAS muon trigger in pp collisions at $\sqrt{s}$ = 8 TeV*, Eur. Phys. J. **C75** (2015) 120, arXiv:1408.3179 [hep-ex].

[123] G. Pasztor, D. Damazio, and F. Monticelli, *Electron and photon trigger performance plots using 2012 data*, Tech. Rep. ATL-COM-DAQ-2014-058, CERN, Geneva, Jun, 2014. https://cds.cern.ch/record/1706278.

[124] M. Cacciari, M. Czakon, M. Mangano, A. Mitov, and P. Nason, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, Phys. Lett. B **710** (2012) 612–622, arXiv:1111.5869 [hep-ph].

[125] M. Beneke, P. Falgari, S. Klein, and C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, Nucl. Phys. B **855** (2012) 695–741, arXiv:1109.1536 [hep-ph].

[126] P. Baernreuther, M. Czakon, and A. Mitov, *Percent Level Precision Physics at the Tevatron: First Genuine NNLO QCD Corrections to $q\bar{q} \to t\bar{t} + X$*, Phys. Rev. Lett. **109** (2012) 132001, arXiv:1204.5201 [hep-ph].

[127] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, JHEP **12** (2012) 054, arXiv:1207.0236 [hep-ph].

[128] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, JHEP **13** (2013) 080, arXiv:1210.6832 [hep-ph].

[129] 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** (2013) 252004, arXiv:1303.6254 [hep-ph].

[130] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-*, Phys. Rev. D **82** (2010) 054018, arXiv:1005.4451 [hep-ph].

[131] J. M. Campbell and R. K. Ellis, *An Update on vector boson pair production at hadron colliders*, Phys. Rev. D **60** (1999) 113006, arXiv:hep-ph/9905386 [hep-ph].

[132] J. M. Campbell and R. K. Ellis, *$t\bar{t}W^{+-}$ production and decay at NLO*, JHEP **07** (2012) 052, arXiv:1204.5678 [hep-ph].

[133] M. V. Garzelli, A. Kardos, C. G. Papadopoulos, and Z. Trocsanyi, *$t\ \bar{t}\ W^{+-}$ and $t\ \bar{t}\ Z$ Hadroproduction at NLO accuracy in QCD with Parton Shower and Hadronization effects*, JHEP **11** (2012) 056, arXiv:1208.2665 [hep-ph].

[134] ATLAS Collaboration, *Measurement of the top quark-pair production cross section with ATLAS in pp collisions at $\sqrt{s}$ = 7 TeV*, Eur. Phys. J. **C71** (2011) 1577, arXiv:1012.1792 [hep-ex].

[135] M. Cacciari, G. P. Salam, and G. Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) 005, arXiv:0802.1188 [hep-ph].

[136] ATLAS Collaboration, *Differential top-antitop cross-section measurements as a function of observables constructed from final-state particles using pp collisions at $\sqrt{s}$ = 7 TeV in the ATLAS detector*, JHEP **06** (2015) 100, arXiv:1502.05923 [hep-ex].

[137] CDF Collaboration, *Measurement of the top quark mass using template methods on dilepton events in p anti-p collisions at $\sqrt{s}$ = 1.96 TeV*, Phys. Rev. D **73** (2006) 112006, arXiv:hep-ex/0602008 [hep-ex].

[138] CDF Collaboration, *W boson polarization measurement in the $t\bar{t}$ dilepton channel using the CDF II Detector*, Phys. Lett. B **722** (2013) 48–54, arXiv:1205.0354 [hep-ex].

[139] D0 Collaboration, *Measurement of the top quark mass using dilepton events*, Phys. Rev. Lett. **80** (1998) 2063–2068, arXiv:hep-ex/9706014 [hep-ex].

[140] ATLAS Collaboration, *Measurement of the charge asymmetry in dileptonic decays of top quark pairs in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, JHEP **05** (2015) 061, arXiv:1501.07383 [hep-ex].

[141] V. Blobel, *Unfolding Methods in High-energy Physics Experiments*, in *Proceedings, CERN School of Computing: Aiguablava, Spain, September 9-22 1984*. 1984.

[142] G. Cowan, *A survey of unfolding methods for particle physics*, Conf. Proc. **C0203181** (2002) 248–257.

[143] G. Choudalakis, *Fully Bayesian Unfolding*, arXiv:1201.4612 [hep-ex].

[144] A. N. Tikhonov, *Solution of incorrectly formulated problems and the regularization method*, Soviet Math. Dokl. **4** (1963) 1035–1038.

[145] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *Equation of State Calculations by Fast Computing Machines*, The Journal of Chemical Physics **21** (1953) 1087–1092, http://scitation.aip.org/content/aip/journal/jcp/21/6/10.1063/1.1699114.

[146] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. C **72** (2012) 1849, arXiv:1110.1530 [hep-ex].

[147] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data*, Eur. Phys. J. C **74** (2014) 2941, arXiv:1404.2240 [hep-ex].

[148] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11** (2016) P04008, arXiv:1512.01094 [hep-ex].

[149] ATLAS Collaboration, *Measurement of $t\bar{t}$ production with a veto on additional central jet activity in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, Eur. Phys. J. C **72** (2012) 2043, arXiv:1203.5015 [hep-ex].

[150] M. Botje, J. Butterworth, A. Cooper-Sarkar, A. de Roeck, J. Feltesse, et al., *The PDF4LHC Working Group Interim Recommendations*, arXiv:1101.0538 [hep-ph].

[151] ATLAS Collaboration, *Measurement of the top quark mass in the $t\bar{t} \to$ dilepton channel from $\sqrt{s}$ = 8 TeV ATLAS data*, arXiv:1606.02179 [hep-ex].

[152] J. Aguilar-Saavedra, *Single top quark production at LHC with anomalous Wtb couplings*, Nucl.Phys. **B804** (2008) 160–192, arXiv:0803.3810 [hep-ph].

[153] J. Aguilar-Saavedra, *Identifying top partners at LHC*, JHEP **0911** (2009) 030, arXiv:0907.3155 [hep-ph].

[154] J. A. Aguilar-Saavedra, *Portrait of a colour octet*, JHEP **08** (2014) 172, arXiv:1405.5826 [hep-ph].

[155] CMS Collaboration, *Measurement of the charge asymmetry in top quark pair production in pp collisions at $\sqrt{s}$ = 8 TeV using a template method*, Phys. Rev. D **93** (2016) 034014, arXiv:1508.03862 [hep-ex].

[156] ATLAS Collaboration, *Measurement of the charge asymmetry in highly boosted top-quark pair production in $\sqrt{s}$ = 8 TeV pp collision data collected by the ATLAS experiment*, Phys. Lett. **B756** (2016) 52–71, arXiv:1512.06092 [hep-ex].

[157] ATLAS Collaboration, *Measurements of the charge asymmetry in top-quark pair production in the dilepton final state at $\sqrt{s}$ = 8 TeV with the ATLAS detector*, Phys. Rev. **D94** (2016) 032006, arXiv:1604.05538 [hep-ex].

[158] S. Berge and S. Westhoff, *Top-Quark Charge Asymmetry Goes Forward: Two New Observables for Hadron Colliders*, JHEP **07** (2013) 179, arXiv:1305.3272 [hep-ph].

[159] CMS Collaboration, *Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s}$ = 8 TeV*, Eur. Phys. J. **C75** (2015) 542, arXiv:1505.04480 [hep-ex].

[160] CMS Collaboration, *Measurements of $t\bar{t}$ spin correlations and top quark polarization using dilepton final states in pp collisions at sqrt(s) = 8 TeV*, Phys. Rev. **D93** (2016) 052007, arXiv:1601.01107 [hep-ex].

[161] ATLAS Collaboration, *Measurement of Top Quark Polarization in Top-Antitop Events from Proton-Proton Collisions at $\sqrt{s}$ = 7 TeV Using the ATLAS Detector*, Phys. Rev. Lett. **111** (2013) 232002, arXiv:1307.6511 [hep-ex].

[162] ATLAS Collaboration, *Measurements of spin correlation in top-antitop quark events from proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Phys. Rev. **D90** (2014) 112016, arXiv:1407.4314 [hep-ex].

[163] ATLAS Collaboration, *Measurement of the correlations between the polar angles of leptons from top quark decays in the helicity basis at $\sqrt{s} = 7$TeV using the ATLAS detector*, Phys. Rev. **D93** (2016) 012002, arXiv:1510.07478 [hep-ex].

[164] M. Baumgart and B. Tweedie, *Transverse Top Quark Polarization and the $t\bar{t}$ Forward-Backward Asymmetry*, JHEP **08** (2013) 072, arXiv:1303.1200 [hep-ph].

[165] W. Bernreuther, P. Galler, Z.-G. Si, and P. Uwer, *Production of heavy Higgs bosons and decay into top quarks at the LHC. II: Top-quark polarization and spin correlation effects*, arXiv:1702.06063 [hep-ph].