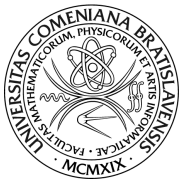
Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky



Mgr. Matej Melo

Autoreferát dizertačnej práce

Charge asymmetry in top-quark pair production at 13 TeV in proton-proton collisions with the ATLAS experiment

na získanie akademického titulu philosophiae doctor

v odbore doktorandského štúdia:

4.1.5. Jadrová a subjadrová fyzika

Bratislava, 2019

Dizertačná práca bola vypracovaná v dennej forme doktorandského štúdia na

Katedre jadrovej fyziky a biofyziky
Fakulty matematiky, fyziky a informatiky
Univerzity Komenskho v Bratislave

Predkladateľ: Mgr. Matej Melo
KJFB FMFI UK
Mlynská dolina F1
842 48 Bratislava

Školiteľ: prof. RNDr. Stanislav Tokár DrSc.
KJFB FMFI UK
Mlynská dolina F1
842 48 Bratislava

Študijný odbor: 4.1.5 Jadrová a subjadrová fyzika

Predseda odborovej komisie:

prof. RNDr. Jozef Masarik DrSc.
KJFB FMFI UK
Mlynská dolina F1
842 48 Bratislava

Abstract

We report a measurement of the charge asymmetry in top-quark pair production using 139 fb^{-1} of proton-proton collision data collected at the centre-of-mass energy of 13 TeV by the ATLAS experiment. Events are reconstructed in the so-called resolved topology and in a topology with highly boosted top quarks. Both topologies are combined and a fully bayesian unfolding method is used to correct for limited detector acceptance and resolution. The charge asymmetry is measured inclusively and differentially as a function of the top-quark pair mass and longitudinal boost. The measured values are in good agreement with the Standard Model NNLO in QCD + NLO in EW predictions and a non-zero inclusive asymmetry is observed at a 4σ confidence level.

Keywords: top quark, charge asymmetry, fully bayesian unfolding, ATLAS

Abstrakt

Prezentujeme meranie nábojovej asymetrie v produkcii top kvarkových párov s použitím 139 fb^{-1} protón-protónových zrážok pri ťažiskovej energii 13 TeV zozbieraných experimentom ATLAS. Udalosti sú rekonštruované v tzv. resolved topológii a v topológii s vysokými top-kvarkovými hybnosťami. Obidve topológie sú skombinované a metóda plne bayesovskej dekonvolúcie je použitá s cieľom zohľadniť limitované detektorové rozlíšenie a akceptanciu. Nábojová asymetria je odmeraná inkluzívne a diferenciálne ako funkcia hmotnosti a pozdĺžnej hybnosti top-kvarkového páru. Namerané hodnoty sú v dobrej zhode s predpoveďami Štandardného modelu na NNLO v QCD + NLO v EW úrovni a nenulová inkluzívna asymetria je pozorovaná na úrovni spoľahlivosti 4σ.

Kľúčové slová: top kvark, nábojová asymetria, plne bayesovská dekonvolúcia, ATLAS

1 Top Quark Charge Asymmetry

The charge asymmetry in heavy quark pair production is a phenomenon predicted by the Standard Model, occurring through higher-order diagrams [1–7]. The charge asymmetry manifests itself through different differential cross sections of the heavy quarks and antiquarks. As a consequence, the probabilities of the final state quarks to be produced in forward/backward directions in the centre-of-mass (CM) frame are not equal. This is illustrated in Figure 1 where the final state particles are heavy quarks Q and $\bar{Q}$ and the initial state particles are light quarks q and $\bar{q}$. The forward direction is given by the direction of the incoming light quark q and the backward direction is defined complementary.

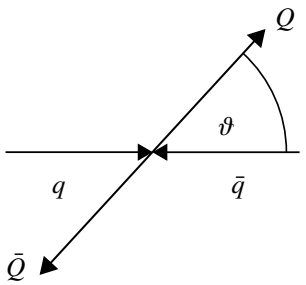


Figure 1: Forward heavy quark production, given by the condition $\cos\theta > 0$.

The definition of the asymmetry in the CM frame is then given by the production cross-sections of the heavy quark Q:

$$A = \frac{\sigma(\cos\theta > 0) - \sigma(\cos\theta < 0)}{\sigma(\cos\theta > 0) + \sigma(\cos\theta < 0)}, \tag{1}$$

where the angle θ is explained in Figure 1.

Beyond the Standard Model (BSM) theories can modify the expected asymmetry due to additional diagrams in which a new particles are exchanged, e.g. axigluons [8, 9], heavy Z bosons [10], or coloured Kaluza-Klein gluon excitations [11]. The charge asymmetry measurements can constrain the parameter space of the BSM theories. A significant deviation of the measured asymmetry from the SM prediction would be considered as an evidence of BSM physics.

Conditions in the hadron colliders do not allow to measure the charge asymmetry using the definition from Equation (1). The interacting partons have different longitudinal momenta and the laboratory frame is not identical with the CM frame as it would be in the case of e^+e^- colliders. To describe the top quark asymmetry at hadron colliders, rapidity y is used instead of the polar angle θ:

$$y = \frac{1}{2}\ln\frac{E + p_z}{E - p_z}, \tag{2}$$

where E is the energy of the final state particle and p_z is the longitudinal momentum. It is clear that in the CM system the condition $\cos\theta > 0$ ($\cos\theta < 0$) is the same as $y > 0$ ($y < 0$).

Additionally, due to the momentum conservation $y_Q = -y_{\bar{Q}}$ and the rapidity difference

$$\Delta y = y_Q - y_{\bar{Q}} \tag{3}$$

is Lorentz invariant under the boosts along the z-axis. As a consequence, a new definition of the charge asymmetry, typically referred to as forward-backward asymmetry A_{FB}, can be defined:

$$A_{\mathrm{FB}}^{Q\bar{Q}} = \frac{N(\Delta y > 0) - N(\Delta y < 0)}{N(\Delta y > 0) + N(\Delta y < 0)}, \tag{4}$$

which is completely equivalent to the CM system definition (1) and can be measured in the laboratory frame.

The definition from Equation (4) was largely used in the Tevatron measurements, where the significant direction was given by the direction of the incident (anti)proton. However, in the symmetric LHC pp collisions it is not possible to use this definition, since the direction of the incident (anti)quark is unknown. On the other hand, in pp collisions the antiquarks are always sea quarks, while the quarks are usually valence quarks. Sea quarks typically carry lower momentum fraction than the valence quarks, therefore in the LHC collisions the $q\bar{q}$ system is mostly boosted in the direction of the incident quark, see Figure 2. This enables to define a complementary forward-central charge asymmetry A_{C} which can be measured at the LHC:

$$A_{\mathrm{C}}^{Q\bar{Q}} = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)}, \tag{5}$$

where $\Delta|y|$ is given by

$$\Delta|y| = |y_Q| - |y_{\bar{Q}}|. \tag{6}$$

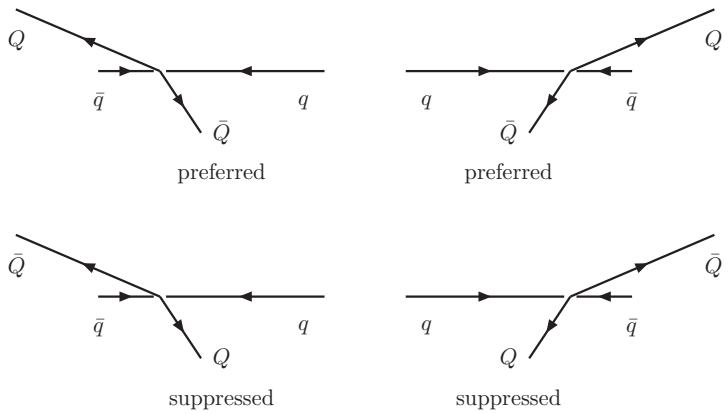


Figure 2: Preferred and suppressed configurations at the LHC assuming positive charge asymmetry.

2 Event Selection and Reconstruction

2.1 Event Selection

A set of selection criteria is applied on full Run 2 ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV, as well as on the simulated samples, in order to reconstruct $t\bar{t}$ ℓ+jets events in the resolved and boosted topology.

Selection Criteria Common to Resolved and Boosted Topologies

- **Primary vertex** with at least two tracks is required.
- **Event quality** – to avoid events affected by detector noise criteria on the calorimeter response and on the jet quality must be fulfilled.
- **Single electron/muon trigger** requirements.
- **Exactly one electron or muon** matched to the trigger with $p_{\mathrm{T}} > 28$ GeV is required. Events containing additional leptons with $p_{\mathrm{T}} > 25$ GeV are rejected.
- **Missing transverse energy and W boson transverse mass (M_{T}^{W})** – in order to suppress fake and non-prompt leptons from multijet background $E_{\mathrm{T}}^{\mathrm{miss}}$ of at least 30 GeV and M_{T}^{W}[1] of at least 30 GeV is required. In the μ+jets channel no cut is applied on $E_{\mathrm{T}}^{\mathrm{miss}}$ but $E_{\mathrm{T}}^{\mathrm{miss}} + M_{\mathrm{T}}^{W} > 60$ GeV is required.
- **b-tagged jet** – at least one of the small-R jets is required to be b-tagged.

Selection Criteria Specific to the Resolved Topology

- **At least 4 small-R jets** with $p_{\mathrm{T}} > 25$ GeV are required.
- **Boosted veto** is implemented in order to remove an overlap between events passing both resolved and boosted criteria. These events are removed from the resolved topology since reconstruction of the top quark four momenta is easier in the boosted topology.
- **Event reconstruction requirements** – the $t\bar{t}$ system is reconstructed using a boosted decision tree (BDT) algorithm, described in detail in Section 2.2. Events are required to have BDT discriminant > 0.3 in order to suppress background processes and combinatorial background due to wrong assignment of jets in $t\bar{t}$ events. This criterion accepts around 52% of $t\bar{t}$ events and only 27% of background events.

[1] $M_{\mathrm{T}}^{W} = \sqrt{2p_{\mathrm{T}}^{\ell}E_{\mathrm{T}}^{\mathrm{miss}}(1-\cos\Delta\phi)}$ where $\Delta\phi$ is the angle between the lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ in the transverse plane with respect to the beam axis.

Selection Criteria Specific to the Boosted Topology

- **At least one small-R jet close to the lepton** with $p_{\mathrm{T}} > 25$ GeV and $\Delta R(\mathrm{jet}, \ell) < 1.5$ is required. If multiple jets satisfy this condition, the one with highest p_{T} is considered for the subsequent reconstruction of the leptonically decaying top quark.

- **At least one top-tagged large-R jet** with $p_{\mathrm{T}} > 350$ GeV and $|\eta| < 2$. In the case of multiple large-R jets satisfying these conditions the one with the highest p_{T} is chosen. Since both top quarks are expected to be back-to-back in the $t\bar{t}$ rest frame, additional requirements related to the large-R jet, isolated lepton and the small-R jet close to the lepton are applied: $\Delta\phi(\mathrm{jet}_{R=1.0}, \ell) > 2.3$ and $\Delta R(\mathrm{jet}_{R=1.0}, \mathrm{jet}_{R=0.4}) > 1.5$.

- **Invariant mass of the reconstructed $t\bar{t}$ system** is required to be larger than 500 GeV. This criterion is imposed to remove a negligible fraction ($\approx 0.1\%$) of poorly reconstructed events which pass the boosted selection criteria despite low $t\bar{t}$ mass.

2.2 Event Reconstruction

After the event selection it is necessary to reconstruct the top quark four momenta from the reconstructed objects, i.e. jets, lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$. In this section the reconstruction techniques in resolved and boosted topologies are described.

Reconstruction in the Resolved Topology In the resoled topology the main challenge of the event reconstruction is to correctly assign individual jets to the four quarks from the $t\bar{t}$ decay. Several reconstruction techniques have been studied, including the Kinematic Likelihood Fitter (KLFitter) [12] and the χ^2-pairing technique [13]. However, these techniques were eventually outperformed by an advanced multivariate Boosted Decision Tree (BDT) technique implemented using the `TMVA` package [14]. The BDT reconstruction technique combines information from the KLFitter with various kinematic variables and the b-tagging information into a single discriminant with value from -1 to 1. Each permutation of jet to quark assignment is evaluated and the permutation with the highest score of the BDT discriminant is chosen. Since the number of possible permutations increases with the number of jets as $\sim n!$, only permutations of up to five jets are considered. If there are more than five jets present in the event, the two jets with the highest b-tagging score are taken into account together with three other jets with the highest p_{T}.

The BDT training is performed separately in the 1-exclusive and 2-inclusive b-tag regions and inclusively in lepton flavours – the electron and muon channels are merged. Together 13 different input variables are used in the BDT. In the analysis, the BDT discriminant of the best permutation is required to be larger than 0.3, which corresponds to $\approx 52\%$ $t\bar{t}$ selection efficiency while only $\approx 27\%$ of background events are kept in the combined 1 b-excl. and 2 b-incl. selection.

After choosing the best permutation it is still necessary to reconstruct the neutrino four-momentum using the $E_{\mathrm{T}}^{\mathrm{miss}}$ information and a constraint from the mass of W boson:

$$m_W = (p_\ell + p_\nu)^2, \tag{7}$$

where p_ℓ and p_ν are four-momenta of the lepton and the corresponding neutrino, respectively. The p_ν^x and p_ν^y components of the neutrino momentum can be obtained from the magnitude and azimuthal angle of the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the problem can be reduced to a quadratic equation for the p_ν^z. If there are two real solutions, the one which leads to the top quark mass closest to 172.5 GeV is chosen. In the case of no real solutions, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is varied by a minimal amount necessary to obtain a real solution.

2.3 Reconstruction in the Boosted Topology

In the boosted topology the four-momentum of the large-R jet satisfying the selection criteria from Section 2.1 is taken as the estimate of the four-momentum of the hadronically decaying top quark. The leptonically decaying top quark four-momentum is constructed from the isolated lepton, selected small-R jet close to the lepton and the neutrino four-momentum. The neutrino four-momentum is calculated using the constraint from Equation (7), similarly as in the case of the resolved topology. The only difference with respect to the resolved topology is that if there are two real solutions for p_ν^z, the one with minimum $|p_\nu^z|$ is taken.

3 Unfolding

In this analysis, the Fully Bayesian Unfolding (FBU) [15] is used to estimate the parton-level charge asymmetry from the reconstruction-level $\Delta|y|$ spectra.

3.1 Fully Bayesian Unfolding Formulation

The FBU is an application of the Bayesian inference to the problem of unfolding: Given the data ($\mathrm{D} \in \mathbb{N}^{N_r}$) and the response matrix $\mathcal{M} \in \mathbb{R}^{N_r} \times \mathbb{R}^{N_t}$, the question is what is the actual parton-level spectrum ($\tilde{\mathrm{T}} \in \mathbb{R}^{N_t}$). In the terms of the Bayesian inference this corresponds to:

$$P(\mathrm{T}|\mathrm{D}, \mathcal{M}) \propto \mathcal{L}(\mathrm{D}|\mathrm{T}, \mathcal{M})\pi(\mathrm{T}), \tag{8}$$

where $P(\mathrm{T}|\mathrm{D}, \mathcal{M})$ is the *posterior* probability of the true spectrum T; $\mathcal{L}(\mathrm{D}|\mathrm{T}, \mathcal{M})$ is the likelihood function of D given T and a *response* matrix $\mathcal{M}$ and $\pi(\mathrm{T})$ is the *prior* probability density for the true spectrum T. In other words, by sampling every possible true spectrum it is possible to obtain a probability density function for the true distributions and in our case also for the A_{C} itself.

The prior, likelihood, treatment of systematic uncertainties and the process of obtaining the posterior probability density are further described in the following sections.

3.2 Prior

The prior probability density $\pi(\mathrm{T})$ is to be chosen according to what is known about T before the measurement is performed. Typically, this corresponds to a reasonably large range which covers all realistic T distributions. If all possibilities are considered with the same probability, the uninformative prior is used:

$$\pi(\mathrm{T}) \propto \begin{cases} 1 & \text{if } \mathrm{T}_t \in [\mathrm{T}_t^{\lceil}, \mathrm{T}_t^{\rceil}], \forall t \in [1, N_t] \\ 0 & \text{otherwise} \end{cases} . \tag{9}$$

If some of the possible T distributions are considered to be more probable, this can by added as an additional information to reduce the variance by introducing a small bias towards the preferred truth distributions:

$$\pi(\mathrm{T}) \propto \begin{cases} e^{\alpha S(\mathrm{T})} & \text{if } \mathrm{T}_t \in [\mathrm{T}_t^{\lceil}, \mathrm{T}_t^{\rceil}], \forall t \in [1, N_t] \\ 0 & \text{otherwise} \end{cases} , \tag{10}$$

where α is an arbitrary parameter and $S(\mathrm{T})$ is a regularisation function. In this analysis an uninformative prior with $[0, 2\tilde{\mathrm{T}}]$ is used, where $\tilde{\mathrm{T}}$ is the parton-level spectrum of the nominal MC sample.

3.3 Likelihood

The response matrix $\mathcal{M}$ with elements $m_{ij} = \epsilon_{t_j} P(r_i|t_j)$ can be obtained from the MC signal sample and contains two pieces of information:

- ϵ_{t_j} - efficiency for a parton-level event produced in bin t_j to be reconstructed in any bin r.
- $P(r_i|t_j)$ - probability of a parton-level event produced in bin t_j to be observed at the reconstruction level in bin r_i. This is summarised in *migration* matrices.

Prediction of the reconstruction-level spectrum $\mathrm{R} \in \mathbb{R}^{N_r}$ corresponding to a given parton-level spectrum T is then

$$r_i(\mathrm{T}, \mathcal{M}) = \sum_{j=0}^{N_r} m_{ij} t_j. \tag{11}$$

The likelihood is defined by comparing the observed spectrum D with the expected one R; assuming Poisson statistics and background prediction $\mathrm{B} \in \mathbb{R}^{N_r}$:

$$\mathcal{L}(\mathrm{D}|\mathrm{T}, \mathcal{M}, \mathrm{B}) = \prod_{i=1}^{N_r} \frac{(r_i + b_i)^{d_i}}{d_i!} e^{-(r_i + b_i)}. \tag{12}$$

3.4 Nuisance Parameter Marginalisation

Treatment of systematic uncertainties is naturally included in the FBU by extending the likelihood with nuisance parameter (NP) terms corresponding to the individual systematic uncertainties.

The marginal likelihood is defined as

$$\mathcal{L}(\mathrm{D}|\mathrm{T}) = \int \mathcal{L}(\mathrm{D}|\mathrm{T},\theta)\pi(\theta)\mathrm{d}\theta, \tag{13}$$

where θ are the nuisance parameters and $\pi(\theta)$ their priors, typically Gaussian distributions G with $\mu = 0$ and σ equal to the magnitude of the corresponding systematic uncertainty. Two main categories of systematic uncertainties are considered:

- Background normalisation uncertainties θ_b which affect only the background predictions.
- Uncertainties related to object identification, reconstruction and calibration θ_s affecting both the signal and the background predictions, $\mathrm{R}(\mathrm{T};\theta_s)$ and $\mathrm{B}(\theta_s,\theta_b)$, respectively.

The signal reconstruction-level prediction is then defined as:

$$r_i(\mathrm{T},\mathcal{M};\theta_s) = r_i(\mathrm{T},\mathcal{M};0)\left(1+\sum_k \theta_s^k \Delta r_i^k\right), \tag{14}$$

where Δr_i^k is the systematic variation corresponding to the uncertainty k.
Similarly, the prediction for each background process:

$$b_i(\theta_s,\theta_b) = b_i(0)(1+\theta_b\Delta b)\left(1+\sum_k \theta_s^k \Delta b_i^k\right), \tag{15}$$

where Δb is the uncertainty on the background normalization.
The marginal likelihood can be then rewritten as:

$$\mathcal{L}(\mathrm{D}|\mathrm{T}) = \int \mathcal{L}(\mathrm{D}|\mathrm{R}(\mathrm{T};\theta_s),\mathrm{B}(\theta_s,\theta_b))G(\theta_s)G(\theta_b)\mathrm{d}\theta_s\mathrm{d}\theta_b. \tag{16}$$

3.5 Signal Region Combination

Utilisation of orthogonal channels, e.g. regions with different background contamination, helps to constrain the individual systematic uncertainties and thus to reduce the total uncertainty. Having the nuisance parameters common to all channels, the likelihood is:

$$\mathcal{L}(\{\mathrm{D}_1\cdots\mathrm{D}_{N_{\mathrm{ch}}}\}|\mathrm{T}) = \int \prod_{i=1}^{N_{\mathrm{ch}}} \mathcal{L}(\mathrm{D}_i|\mathrm{T};\theta)G(\theta)\mathrm{d}\theta, \tag{17}$$

where N_{ch} is the number of orthogonal channels.

The final posterior probability including systematic uncertainties and multiple orthogonal signal regions can be then written as

$$P(\mathrm{T}|\{\mathrm{D}_1 \cdots \mathrm{D}_{N_{\mathrm{ch}}}\}) = \int \prod_{i=1}^{N_{\mathrm{ch}}} \mathcal{L}\left(\mathrm{D}_i|\mathrm{R}_i(\mathrm{T};\theta_s), \mathrm{B}_i(\theta_s,\theta_b)\right) G(\theta_s) G(\theta_b) \pi(\mathrm{T}) \mathrm{d}\theta_s \mathrm{d}\theta_b. \tag{18}$$

In this analysis events are split into four signal regions by topology (resolved/boosted) and b-tag multiplicity (1 b-excl./2b-incl.).

3.6 Sampling

The posterior $P(\mathrm{T},\mathrm{D})$ is determined by sampling the $(N_t + N_{\mathrm{NP}})$-dimensional parameter space and by evaluating for each point the product of $\mathcal{L}(\mathrm{D}|\mathrm{T},\mathcal{M})$ and $\pi(\mathrm{T})$, thus performing a numerical integration.

Advanced techniques specialized for multi-dimensional sampling based on the Markov Chain Monte Carlo method (MCMC) [16] are used within FBU.

The posterior distributions contain the full probability density information; the mean of distribution can be afterwards taken as the unfolded estimate and the RMS as the corresponding uncertainty. Similarly, posterior probability density distribution can be obtained for any quantity that is computed from the spectrum, such as the A_{C}:

$$p(A_C|\mathrm{D}) = \int \delta(A_C - A_C(\mathrm{T}))P(\mathrm{T}|\mathrm{D})d\mathrm{T}. \tag{19}$$

Effectively, this corresponds to calculating the observable of interest from the bin contents for every single sample.

3.7 Binning Choice and Linearity Tests

An important step in the unfolding optimisation is related to proper choice of binning in the $\Delta|y|$ distributions, as well as in the differential variable in measurements of the A_{C} as a function of $m(t\bar{t})$ and $\beta_{\mathrm{z}}(t\bar{t})$.

In the case of differential measurements, the choice of binning in the differential variables is motivated by physics considerations. Fine binning in $m(t\bar{t})$ is desirable in order to discriminate different physics models. Since many BSM theories predict enhancement of the asymmetry in the very high $m(t\bar{t})$ region, it is important to set the last bin threshold as high as possible. At high $\beta_z(t\bar{t})$ the fraction of the $q\bar{q}$ annihilation is larger and the asymmetry is enhanced in a model independent way. Taking into account the statistical limitations the following binning is chosen:

- $\beta_z(t\bar{t})$ - 4 bins: $[0, 0.3, 0.6, 0.8, 1]$.
- $m(t\bar{t})$ [GeV] - 5 bins: $[0, 500, 750, 1000, 1500, \infty]$.

Two competing factors determine the choice of the number of bins in the $\Delta|y|$ distribution:

- Smaller number of bins implies smaller relative statistical uncertainties. At least two bins are necessary to compute A_C (positive and negative side of the $\Delta|y|$ distribution).
- Larger number of bins allows to track the migrations more accurately and thus allows to obtain unbiased estimates. However, only migrations that change the $\Delta|y|$ sign affect the computation of the A_C and these are more likely for small $\Delta|y|$ values. Therefore a fine binning is required in the central $\Delta|y|$ region.

In this analysis four bins in $\Delta|y|$ are used and the same binning is used at both the reconstruction and the parton level in all four signal regions mentioned in Section 3.5. However, the x-value in $\Delta|y|$ binning $[-5, -x, 0, x, 5]$ is optimised separately in the inclusive measurement and in each of the differential bins. The outer bin edges (± 5) are effectively the same as $\pm\infty$, since there are no $t\bar{t}$ events with $\Delta|y|$ values beyond $\approx \pm 4$.

In order to find the optimal x-values in $\Delta|y|$ binning, samples with various parton-level asymmetries are prepared and the unfolding procedure is required to work properly across a range of different parton-level A_C values. In this analysis, the PROTOS [17] generator is used to simulate axigluon (mass of 250 GeV) contribution to the charge asymmetry and the nominal signal sample is subsequently correspondingly reweighted. Parton-level asymmetries of approximately $\pm 1\%$, $\pm 2\%$, $\pm 3\%$ and $\pm 4\%$ are considered. The unfolding response is required to be linear with a slope ≈ 1 and offset ≈ 0. The optimal x-values range from 0.3 to 1.0, depending on the differential bin where the measurement is performed.

4 Systematic Uncertainties

4.1 Experimental Uncertainties

In this section, systematic uncertainties originating from the detector and reconstruction imperfections are described.

Luminosity The relative luminosity uncertainty of the combined 2015-2018 dataset is 1.7%.

Pile-up Scale factors are applied in order to correct for the differences between the pile-up conditions in data and simulated samples. The uncertainties on the corresponding scale factors are considered.

Lepton-Related Uncertainties Lepton (e, μ) identification, reconstruction, isolation and trigger performance, as well as resolution and momentum scale, differ between data and simulations and scale factors are applied to correct these differences. The scale factors are obtained using the tag and probe technique in well understood decays of the Z boson, J/ψ and W boson into leptons [18–20]. The uncertainties on these corrections are treated as systematic uncertainties.

Jet Energy Scale The JES and its uncertainty is estimated from the collision data and MC simulations using techniques described in [21]. Events with a vector boson and additional jets are used to calibrate jets in the central region. Dijet events are used to calibrate forward jets against the jets in the central region and multijet events are used to calibrate high-p_{T} jets. The measurements are combined and decorrelated into a set of 29 nuisance parameters which have different jet p_{T} and η dependencies [22].

Jet Energy Resolution The jet energy resolution (JER) is measured separately in data and MC using in situ techniques [22]. The differences between the jet energy resolutions in data and MC are decorrelated into 8 components with different dependencies on the jet p_{T} and η.

Jet Vertex Tagger The JVT scale factors are calculated using simulated $Z \to \mu^+\mu^-$ and $t\bar{t}$ events [23]. The corresponding JVT uncertainty includes the uncertainty on the pile-up jet contamination and a systematic uncertainty due to the choice of different MC generators.

Large-R Jet Moment Scale and Resolution The scales of the detector response for all large-R jet moments (p_{T}, mass, τ_{32} substructure variable [24]) are obtained using a method described in Ref. [25]. In total, a set of 14 nuisance parameters is used to describe uncertainties on the large-R jet moment scales and resolution.

***B*-tagging** A set of scale factors correcting different b-tagging efficiencies in data and MC is used and the corresponding uncertainties are propagated through the analysis as systematic uncertainties [26–28].

Missing Transverse Energy Scale and Resolution The $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated from several terms corresponding to different types of reconstructed objects. Uncertainty on each object is evaluated and then propagated to the uncertainty on the $E_{\mathrm{T}}^{\mathrm{miss}}$.

4.2 Modelling Uncertainties

Signal and background modelling, based on theoretical understanding of the physical processes, is used in the measurement e.g. to create the response matrix and model the background contributions.

Cross Sections and Normalisation For all processes entering the measurement normalisation uncertainties are considered. Typically, the normalisation uncertainty corresponds to theoretical cross section uncertainties. In some cases, such as for the multijet data-driven background, a conservative normalisation uncertainty estimate is used instead.

Matrix Element Modelling To estimate the uncertainty related to the matrix element (ME) modelling in the $t\bar{t}$ signal the nominal POWHEG BOX 2 generator is compared to the alternative MADGRAPH5_AMC@NLO 2.6.0. For the purpose of this comparison both samples are generated with the simplified ATLFAST-II simulation of hadron showers. The difference is taken as 'shape-only' and the normalisation effect is removed.

Parton Shower Modelling The uncertainty on parton shower and hadronisation modelling (PS) is treated similarly as the ME uncertainty. In this case, the PYTHIA 8.230 generator is compared to the alternative HERWIG 7.04, both generated with ATLFAST-II and interfaced to the POWHEG BOX 2.

Initial and Final State Radiation The uncertainties on the effects of initial state radiation (ISR) and final state radiation (FSR) in top quark production are studied using the nominal MC generators, i.e. POWHEG BOX 2 and PYTHIA 8.230. The fast simulation ATLFAST-II is used in the case of ISR and the full GEANT4 for the FSR.

Top Quark Mass The nominal sample is generated with $m_{\text{top}} = 172.5$ GeV and the effect of the top quark mass uncertainty on the signal modelling is studied using alternative samples generated with $m_{\text{top}} = 172$ GeV and $m_{\text{top}} = 173$ GeV.

Parton Distribution Functions The uncertainty on the PDFs is applied only on the signal sample using the PDF4LHC prescriptions [29].

Single Top Quark Wt-Channel Interference Due to the interference of the Wt-channel with the $t\bar{t}$ production it is necessary to remove the overlap using the diagram removal (DR) technique. The alternative diagram subtraction (DS) method [30] is used as an alternative and the differnce is taken as a corresponding uncertainty.

W+jets Modelling A number of scale variations in W+jets background modelling are considered as shape-only uncertainties [31]. These include the μ_r and μ_f scales, the CKKW scale and QSF scale.

Multijet Shape An alternative parametrisation is taken as a shape-only uncertainty.

4.3 Unfolding Uncertainties

In addition to all systematic uncertainties mentioned above, two uncertainties are related to the unfolding procedure itself. These uncertainties are simply added in quadrature to the total unfolded uncertainty.

Response Matrix Statistical Uncertainty To estimate this uncertainty, 300 pseudo-experiments were performed with the response matrices smeared according the Poisson distribution and the raw number of MC events. The width of the obtained A_{C} distribution is taken and added in quadrature to the total unfolded uncertainty obtained using the nominal response matrix.

Unfolding Bias The non-ideal slope and offset of the unfolding response leads to a bias in the unfolded values. In most cases the bias is found to be negligible.

5 Results

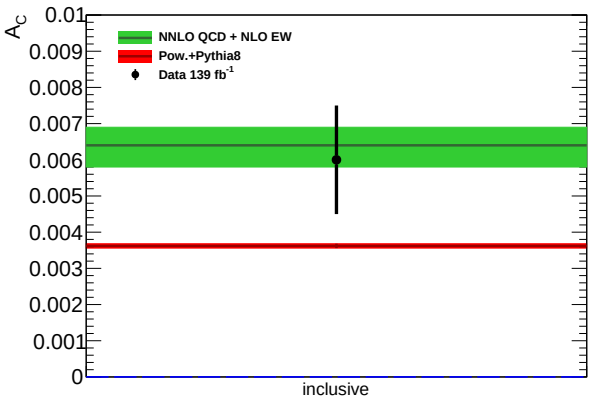
An overview of the A_{C} values with the corresponding uncertainties is provided in Table 1. The uncertainties are dominated by the RMS of the posteriors, followed by the statistical uncertainties in the response matrix and the unfolding bias, which is mostly negligible. The total uncertainty is the sum-in-quadrature of the mentioned uncertainties. Graphically, the results with the total uncertainties are presented in Figure 3. All values are consistent with the NLO Powheg+Pythia 8 prediction. Moreover, a strongest to-date evidence of non-zero inclusive charge asymmetry is observed with a 4 σ confidence level.

		Data 139 fb^{-1}				
		Mean	Post-marg.	Res. mat. stat.	Bias	Total
inclusive		0.0060	0.0014	0.0005	0.0001	0.0015
$m(t\bar{t})$	< 500 GeV	0.0045	0.0044	0.0013	0.0001	0.0045
	500-750 GeV	0.0051	0.0029	0.0009	0.0000	0.0031
	750-1000 GeV	0.0100	0.0067	0.0021	0.0001	0.0070
	1000-1500 GeV	0.0169	0.0077	0.0029	0.0004	0.0083
	> 1500 GeV	0.0121	0.0315	0.0092	0.0005	0.0329
$\beta_{\mathrm{z}}(t\bar{t})$	0-0.3	0.0007	0.0051	0.0020	0.0001	0.0055
	0.3-0.6	0.0085	0.0040	0.0013	0.0003	0.0042
	0.6-0.8	0.0014	0.0044	0.0015	0.0004	0.0047
	0.8-1.0	0.0100	0.0049	0.0013	0.0007	0.0051

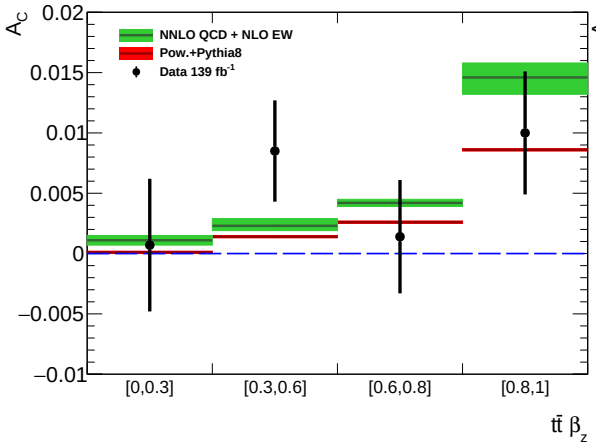
Table 1: Results with uncertainties, including the post-marginalisation uncertainty, uncertainty due to limited number of MC events in the response matrix, uncertainty due to the unfolding bias and the total uncertainty, for the inclusive and differential $A_{\mathrm{C}}^{t\bar{t}}$ measurements.

6 Conclusion

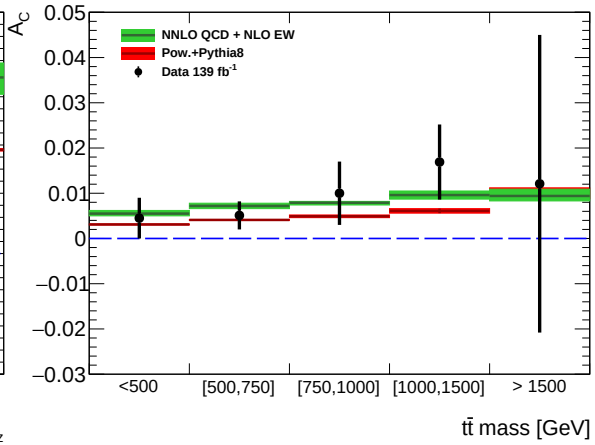
This thesis presents the first measurement of the charge asymmetry in top-quark pair production by the ATLAS experiment using full Run 2 (139 fb^{-1}) proton-proton collision data at $\sqrt{s} = 13$ TeV. The lepton+jets decay channel is investigated. Events are selected in the resolved and boosted topologies which are subsequently combined within the fully bayesian



(a) Data 139 fb^{-1}, inclusive.



(b) Data 139 fb^{-1}, differential $\beta_z(t\bar{t})$.



(c) Data 139 fb^{-1}, differential $m(t\bar{t})$.

Figure 3: The unfolded inclusive and differential ($\beta_z(t\bar{t})$ and $m(t\bar{t})$) A_C values compared to the NLO POWHEG+PYTHIA 8 parton-level prediction. The total uncertainty is shown.

unfolding method. The charge asymmetry is measured inclusively and also differentially as a function of the top-quark pair longitudinal boost and mass. Systematic uncertainties are marginalised within the unfolding procedure. A bootstrapping method is used to remove statistically insignificant systematic uncertainties and a pruning procedure is used to remove negligible uncertainties in order to simplify the unfolding problem. The measurement is limited by statistical uncertainties in all regions and the measured values are in good agreement with the NNLO in QCD + NLO in EW Standard Model predictions. A significant improvement in the precision with respect to the 8 TeV A_C measurements is observed. In the inclusive case, an evidence of a non-zero asymmetry is observed at a 4 σ confidence level.

References

[1] J. H. Kuhn and G. Rodrigo, *Charge asymmetry of heavy quarks at hadron colliders*, Phys. Rev. D **59** (1999) 054017.

[2] M. T. Bowen, S. D. Ellis, and D. Rainwater, *Standard model top quark asymmetry at the Fermilab Tevatron*, Phys. Rev. D **73** (2006) 014008.

[3] W. Bernreuther and Z.-G. Si, *Top quark and leptonic charge asymmetries for the Tevatron and LHC*, Phys. Rev. D **86** (2012) 034026.

[4] V. Ahrens, et al., *Top-pair forward-backward asymmetry beyond next-to-leading order*, Phys. Rev. D **84** (2011) 074004.

[5] N. Kidonakis, *The top quark forward-backward asymmetry at approximate N^3LO*, Phys. Rev. D **91** (2015) 071502.

[6] M. Czakon, et al., *NNLO QCD predictions for fully-differential top-quark pair production at the Tevatron*, JHEP **05** (2016) 034.

[7] M. Czakon, et al., *Top-pair production at the LHC through NNLO QCD and NLO EW*, JHEP **10** (2017) 186.

[8] O. Antunano, J. H. Kuhn, and G. Rodrigo, *Top quarks, axigluons and charge asymmetries at hadron colliders*, Phys. Rev. D **77** (2008) 014003.

[9] G. Rodrigo, *Axigluon signatures at hadron colliders*, PoS **RADCOR2007** (2007) 010.

[10] J. L. Rosner, *Prominent decay modes of a leptophobic Z'*, Phys. Lett. B **387** (1996) 113–117.

[11] P. Ferrario and G. Rodrigo, *Massive color-octet bosons and the charge asymmetries of top quarks at hadron colliders*, Phys. Rev. D **78** (2008) 094018.

[12] E. et al., *A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework*, Nucl. Instrum. Meth. A **748** (2014) 18–25.

[13] ATLAS Collaboration, *Search for $t\bar{t}$ resonances in the lepton plus jets final state with ATLAS using 4.7 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. D **88** (2013) 012004.

[14] A. Hocker et al., *TMVA - Toolkit for Multivariate Data Analysis*, `arXiv:physics/0703039`.

[15] G.Choudalakis, *Fully Bayesian Unfolding*, `arXiv:1201.4612`.

[16] P. Diaconis, *The Markov Chain Monte Carlo Revolution*, Bulletin of the American Mathematical Society **46** (2009) 179–205.

[17] J. A. Aguilar-Saavedra, *Single top quark production at LHC with anomalous Wtb couplings*, Nucl. Phys. B **804** (2008) 160–192.

[18] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, `https://cds.cern.ch/record/2157687`.

[19] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **76** (2016) 292.

[20] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s}$ = 13 TeV*, https://cds.cern.ch/record/2203514.

[21] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002.

[22] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s}$ = 13 TeV*, https://cds.cern.ch/record/2037613.

[23] ATLAS Collaboration, *Tagging and suppression of pileup jets with the ATLAS detector*, https://cds.cern.ch/record/1700870.

[24] ATLAS Collaboration, *Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC*, arXiv:1808.07858.

[25] ATLAS Collaboration, *Identification of boosted, hadronically decaying W bosons and comparisons with ATLAS data taken at $\sqrt{s}$ = 8 TeV*, Eur. Phys. J. C **76** (2016) 154.

[26] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11** (2016) P04008.

[27] ATLAS Collaboration, *Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector*, https://cds.cern.ch/record/2306649.

[28] ATLAS Collaboration, *Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at $\sqrt{s}$ = 13 TeV*, http://cds.cern.ch/record/2314418.

[29] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001.

[30] S. Frixione, et al., *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029.

[31] ATLAS Collaboration, *ATLAS simulation of boson plus jets processes in Run 2*, https://cds.cern.ch/record/2261937.

Selection of Author's Publications

1. Matej Melo (on behalf of the ATLAS Collaboration), *Top-antitop charge asymmetry measurements in the lepton+jets channel with the ATLAS detector*, `arXiv:1901.05034`.

2. 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. B **761** (2016) 136-157.

3. ATLAS Collaboration, *Measurement of the top quark mass in the $t\bar{t} \to$ dilepton channel from $\sqrt{s} = 8$ TeV ATLAS data*, Phys. Lett. B **761** (2016) 350-371.

4. ATLAS Collaboration, *Fiducial, total and differential cross-section measurements of t-channel single top-quark production in pp collisions at 8 TeV using data collected by the ATLAS detector*, Eur. Phys. J. C **77** (2017) no.8, 531.

5. ATLAS Collaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ production cross sections in multilepton final states using 3.2 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **77** (2017) no.1, 40.

6. ATLAS Collaboration, *Search for four-top-quark production in the single-lepton and opposite-sign dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **99** (2019) no.5, 052009.

7. ATLAS Collaboration, *Measurement of the top quark mass in the $t\bar{t} \to$ lepton+jets channel from $\sqrt{s} = 8$ TeV ATLAS data and combination with previous results*, Eur. Phys. J. C **79** (2019) no.4, 290.

8. ATLAS Collaboration, *Search for heavy particles decaying into top-quark pairs using lepton-plus-jets events in protonproton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **78** (2018) no.7, 565.

9. ATLAS Collaboration, *Measurements of $t\bar{t}$ differential cross-sections of highly boosted top quarks decaying to all-hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, Phys. Rev. D **98** (2018) no.1, 012003.

10. ATLAS and CMS Collaborations, *Combination of inclusive and differential $t\bar{t}$ charge asymmetry measurements using ATLAS and CMS data at $\sqrt{s} = 7$ TeV and 8 TeV*, JHEP 1804 (2018) 033.

11. ATLAS Collaboration, *Direct top-quark decay width measurement in the $t\bar{t}$ lepton+jets channel at $\sqrt{s} = 8$ TeV with the ATLAS experiment*, Eur. Phys. J. C **78** (2018) no.2, 129.

12. ATLAS Collaboration, *Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, JHEP 1711 (2017) 191.

13. ATLAS Collaboration, *Top-quark mass measurement in the all-hadronic $t\bar{t}$ decay channel at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP 1709 (2017) 118.

14. ATLAS Collaboration, *Measurements of top-quark pair differential cross-sections in the $e\mu$ channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, Eur. Phys. J. C **77** (2017) no.5, 292.

15. ATLAS Collaboration, *Measurements of charge and CP asymmetries in b-hadron decays using top-quark events collected by the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV*, JHEP 1702 (2017) 071.