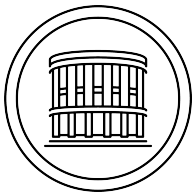
COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

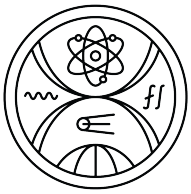
COMBINATION OF TOP-QUARK PAIR CHARGE ASYMMETRY MEASUREMENTS USING ATLAS DATA AT 13 TEV

Dissertation thesis

2022 Mgr. Barbora Eckerová

COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND
INFORMATICS





COMBINATION OF TOP-QUARK PAIR CHARGE ASYMMETRY MEASUREMENTS USING ATLAS DATA AT 13 TEV

Dissertation thesis

Study program:	Nuclear and Subnuclear Physics
Department:	Department of Nuclear Physics and Biophysics
Supervisor:	doc. Mgr. Pavol Bartoš, PhD.
Consultant:	prof. RNDr. Stanislav Tokár, DrSc.

Bratislava, 2022
Mgr. Barbora Eckerová

Mgr. Barbora Eckerová: *Combination of top-quark pair charge asymmetry measurements using ATLAS data at 13 TeV,* Dissertation thesis, © December 2022

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Mgr. Barbora Eckerová
Študijný program: jadrová a subjadrová fyzika (Jednoodborové štúdium, doktorandské III. st., denná forma)
Študijný odbor: fyzika
Typ záverečnej práce: dizertačná
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Combination of top-quark pair charge asymmetry measurements using ATLAS data at 13 TeV
Kominácia meraní nábojovej asymetrie v produkcii top-kvarkových párov použitím 13 TeV dát zaznamenaných experimentom ATLAS

Anotácia: Meranie nábojovej asymetrie je jedným z testov Štandardného modelu. Ak by sa nameraná hodnota nábojovej asymetrie líšila od predpovede Štandardného modelu, mohlo by sa jednať od prejav novej fyziky za Štandardným modelom. Predošlé merania uskutočnené na dátach získaných z protón-protónových zrážok pri energii 7 a 8 TeV zaznamenaných experimentami ATLAS a CMS vykazujú zhodu s teoretickou predpoveďou. Táto práca je zameraná na meranie nábojovej asymetrie v semileptónovom rozpadovom kanáli top kvarkovej párovej produkcie a následnú kombináciu výsledkov s meraním uskutočneným v dileptónovom rozpadovom kanáli. Na meranie sa používajú dáta získané z protón-protónových zrážok pri energii 13 TeV. Získané výsledky je možné interpretovať v efektívnej teórii poľa.

Cieľ: Meranie nábojovej asymetrie v semileptónovom kanáli - štúdium systematických neurčitostí a možnosť ich redukcie. Kombinácia meraní zo semileptónového a dileptónového kanála. Štúdium korelácií neurčitostí pri kombinácii meraní.

Literatúra:
1. M. Beneke et al., Top Quark Physics, Proc. of the Workshop on Standard Model Physics at LHC, CERN 2000-004
2.The ATLAS collaboration, Expected Performance of the ATLAS Experiment, CERN-OPEN-2008-020, Geneva 2009
3. J.F. Donoghue, E.Golowich and B.R. Holstein, Dynamics of the Standard Model, Cambridge University Press, New York 1992

Poznámka: Programátorské zručnosti a znalosť teórie sú vítané

Kľúčové slová: Nábojová asymetria, top kvark, 13 TeV, ATLAS

Školiteľ: Mgr. Pavol Bartoš, PhD.
Konzultant: prof. RNDr. Stanislav Tokár, DrSc.
Katedra: FMFI.KJFB - Katedra jadrovej fyziky a biofyziky
Vedúci katedry: prof. RNDr. Jozef Masarik, DrSc.

Spôsob sprístupnenia elektronickej verzie práce:
bez obmedzenia

53184283

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

Dátum zadania: 29.01.2018

Dátum schválenia: 30.01.2018

prof. RNDr. Jozef Masarik, DrSc.
garant študijného programu

.....
študent

.....
školiteľ

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Mgr. Barbora Eckerová
Study programme:	Nuclear and Subnuclear Physics (Single degree study, Ph.D. III. deg., full time form)
Field of Study:	Physics
Type of Thesis:	Dissertation thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: Combination of top-quark pair charge asymmetry measurements using ATLAS data at 13 TeV

Annotation: The measurement of the charge asymmetry is one of the test of the Standard model. If the measured charge asymmetry is different from the Standard model prediction, it could be a sign for a new physics beyond Standard model. Up to now, the results of the charge asymmetry measurements obtained from the 7 and 8 TeV data recorded by the ATLAS and CMS experiments are consistent with the theory predictions. The focus of the thesis is to measure charge asymmetry in the top quark pair production using the semileptonic decay channel. Moreover a combination with the results from the dileptonic decay channel is expected. The 13 TeV collisions data will be analyzed. The final results can be interpreted within the effective field theory.

Aim: Charge asymmetry measurement in the semileptonic decay channel – study of the systematic uncertainties and possibilities of their reduction. Combination of the results obtained from the semileptonic and dileptonic decay channels. Study of the correlation scheme used for the combination.

Literature:
1. M. Beneke et al., Top Quark Physics, Proc. of the Workshop on Standard Model Physics at LHC, CERN 2000-004
2.The ATLAS collaboration, Expected Performance of the ATLAS Experiment, CERN-OPEN-2008-020, Geneva 2009
3. J.F. Donoghue, E.Golowich and B.R. Holstein, Dynamics of the Standard Model, Cambridge University Press, New York 1992

Comment: Programing skills and knowledge of theory are welcome.

Keywords: Charge asymmetry, top quark, 13 TeV, ATLAS

Tutor:	Mgr. Pavol Bartoš, PhD.
Consultant:	prof. RNDr. Stanislav Tokár, DrSc.
Department:	FMFI.KJFB - Department of Nuclear Physics and Biophysics
Head of department:	prof. RNDr. Jozef Masarik, DrSc.
Assigned:	29.01.2018

Approved: 30.01.2018 prof. RNDr. Jozef Masarik, DrSc.
Guarantor of Study Programme

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

....

Student Tutor

ABSTRAKT

Predložená práca sa zaoberá meraním nábojovej asymetrie z dát nazbieraných detektorom ATLAS počas Run 2. Doposiaľ sa vždy používali dáta zodpovedajúce jednému z rozpadových kanálov top kvarkového páru, buď rozpadový kanál s jedným alebo s dvomi leptónmi. Po prvýkrát sa využila kombinácia dát z oboch kanálov na spresnenie určenia nábojovej asymetrie. Hlavnou analyzačnou metódou je tzv. fully Bayesian unfolding, ktorým sa extrahuje nábojová asymetria z nameraných spektier. Hodnota asymetrie je získaná pre celý súbor dát ale aj pre konkrétne podsúbory, ktoré prislúchajú určitým hodnotám diferenciálnych premenných ako invariantná hmotnosť top kvarkového páru ($m_{t\bar{t}}$), rýchlosť top kvarkového páru v smere osi zväzku v jednotkách rýchlosti svetla, $\beta_z^{t\bar{t}} = v_z^{t\bar{t}}/c$, a priečna hybnosť top kvarkového páru ($p_T^{t\bar{t}}$).

Jednotlivé kroky, z ktorých pozostáva analýza nameraných dát, sú detailne opísané. Použité techniky ako bootstrap metóda, dekorelácie systematických parametrov či preváhovanie spektier buď na základe truth rozdelení alebo NNLO predpovedí, slúžia na korigovanie vplyvu systematických neurčitostí v prípade, že sú tieto neurčitosti nadhodnotené alebo podhodnotené. Rôzne dekorelačné schémy pre súbor systematických neurčitostí zložený zo systematických neurčitostí používaných v jedno- a dvojleptónovom rozpadovom kanáli boli otestované. Pomocou opísaných metód je určená celková hodnota neurčitosti pre namerané hodnoty asymetrie.

Pre dvojleptónový rozpadový kanál bola určená aj hodnota tzv. leptónovej asymetrie. Rovnako ako v prípade nábojovej asymetrie, bola jej hodnota získaná s použitím celého súboru dát a tiež ako funkcia diferenciálnej premennej ako invariantná hmotnosť leptónového páru ($m_{\ell\bar{\ell}}$), rýchlosť leptónového páru v smere osi zväzku v jednotkách rýchlosti svetla, $\beta_{z,\ell\bar{\ell}}$, a priečna hybnosť leptónového páru ($p_{T,\ell\bar{\ell}}$).

Namerané hodnoty nábojovej či leptónovej asymetrie súhlasia v rámci neurčitostí s predpoveďou Štandardného modelu. Výsledky z inkluzívneho a diferenciálneho merania nábojovej asymetrie ako funkcie $m_{t\bar{t}}$ sú použité na odvodenie povolených hodnôt Wilsonových koeficientov prislúchajúcich k operátorom v efektívnej teórii poľa (EFT). Povolené intervaly hodnôt pre Wilsonove koeficienty sú tiež určené za pomoci výsledkov z merania tzv. energetickej asymetrie, ktorej meranie je citlivé na iné oblasti v priestore EFT operátorov a teda prináša dodatočné informácie spresňujúce rozsah intervalov získaných z merania nábojovej asymetrie.

ABSTRACT

The thesis deals with the measurement of charge asymmetry using the full Run 2 data set collected by the ATLAS detector. For the first time, data from single-lepton and dilepton $t\bar{t}$ decay channels are combined. The main technique used in the analysis is fully Bayesian unfolding. Charge asymmetry is unfolded from the full data set, but also using partial data sets corresponding to specific regions of a differential observable. Charge asymmetry is studied as a function of invariant mass of $t\bar{t}$ pair ($m_{t\bar{t}}$), longitudinal boost of $t\bar{t}$ pair along the beam axis ($\beta_z^{t\bar{t}}$) and transverse momentum of $t\bar{t}$ pair ($p_T^{t\bar{t}}$).

In the presented thesis, the whole analysis chain is outlined. Various techniques used to deal with the systematic uncertainties, like bootstrap method, de-correlation technique and truth-based/NNLO-based re-weighting, are described. Numerous correlation scenarios of systematic-uncertainties in combination of single-lepton and dilepton systematic uncertainties are inspected. The total uncertainties on the measured charge asymmetry values are determined.

Additionally, leptonic charge asymmetry is measured in dilepton channel. The inclusive and differential measurements as a function of invariant mass of $\ell\bar{\ell}$ pair ($m_{\ell\bar{\ell}}$), longitudinal boost of $\ell\bar{\ell}$ pair along the beam axis ($\beta_{z,\ell\bar{\ell}}$) and transverse momentum of $\ell\bar{\ell}$ pair ($p_{T,\ell\bar{\ell}}$) are performed.

Charge asymmetry and leptonic charge asymmetry values are found to be consistent with the latest Standard Model prediction. Unfolded results of inclusive and $m_{t\bar{t}}$ differential charge asymmetry measurements are exploited to derive bounds on Wilson coefficients corresponding to the operators in effective field theory. Complementarity of energy asymmetry and charge asymmetry measurements in their EFT interpretation is utilized.

ABOUT AUTHOR'S CONTRIBUTION

The presented measurement is a continuation of previously performed measurement in single-lepton decay channel [1]. However, basic elements of the analysis process has been updated and many alterations has been implemented. The treatment of modeling systematic uncertainties has changed completely due to their re-definition and usage of truth-based re-weighting.

Furthermore, for the first time, charge asymmetry is measured in combination of single-lepton and dilepton channel. Performing such combination requires to adapt the analysis methods for processing data from both channels. Correlation scheme of systematic uncertainties between two decay channels and also within their inner regions has to be decided. Establishing proper treatment of correlations between systematic uncertainties is a necessity in order not to overestimate or underestimate the final uncertainty of unfolded charge asymmetry. Therefore, detailed studies of constraints and pulls for various correlation scenarios have been carried out.

In contrast with the measurement in [1], another differential measurement has been added - measurement of charge asymmetry as a function of transverse momentum of $t\bar{t}$ pair ($p_T^{t\bar{t}}$).

PUBLICATIONS

During the development of the analysis procedure, progresses and updates were frequently reported in ATLAS internal meetings.

Results obtained in this analysis, are published in [2].

Analysis methods and results were presented at the 15th International Workshop on Top-Quark Physics (TOP2022) [3].

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisor, doc. Mgr. Pavol Bartoš, PhD., who was always willing to give a helping hand and offer new ideas and solutions.

I would like to thank my consultant, prof. RNDr. Stanislav Tokár, DrSc., for valuable comments and insights during discussions about the analysis and continuous support.

Many thanks to all my colleagues, who helped me to grasp some aspects of the analysis during friendly debates.

A lot of thanks and love to all my family members, who endured all the hard times with me and showed a great patience and understanding.

Lastly, I would like to send my thanks to André Miede, who developed the typographical look-and-feel `classicthesis` [4], which was used for typesetting of the thesis.

CONTENTS

1 INTRODUCTION

Compared to other quarks, the top-quark has a lot of exclusive features. One of the most significant attributes is the value of the top-quark mass (m_t). Its Yukawa coupling, which depends on m_t as $\sqrt{2}m_t/\eta = 1$, indicates that the top quark can play a role in the electroweak symmetry breaking ($\eta = 246$ GeV, is the vacuum expectation value of Higgs field). The top quark is an excellent perturbative object for testing QCD because it is produced at small distances ($1/m_t$) characterized by low value of coupling constant $\alpha_S \sim 0.1$.

The value of m_t gives a unique possibility to study properties of a bare quark as closely as possible. Being the heaviest known fermion, its lifetime is much smaller than the time of hadronization. Therefore, top-quark properties are directly translated to their decay products, without being diluted by the presence of other quarks bounded together in a hadron.

One of the features of the top-quark pair production is charge asymmetric cross section under the exchange of top quark and top antiquark. The charge asymmetry is a quantity sensitive to a potential new heavy boson which may mediate the top-quark pair production. Hence, it is a suitable probe of new physics, because these new intermediate bosons couple more strongly to top-quarks than to other fermions due to their large mass.

The thesis is focused on study and measurement of the charge asymmetry in top-quark pair production using data collected by the ATLAS detector at a center-of-mass energy of 13 TeV. Previously performed measurements, using collision data at the center-of-mass energy of 7 or 8 TeV either by the CMS or ATLAS detector, measured the charge asymmetry value consistent with the prediction calculated by the Standard Model. But for the first time, the measurement is performed using data from two $t\bar{t}$ decay channels, single-lepton and dilepton together, thus exploiting information from both to obtain more precise results.

This work is organized in four parts. Firstly, the top quark and studied phenomenon, charge asymmetry, are introduced and described. Second part deals with the ATLAS experiment and outlines the experimental settings of data taking. The reconstruction of physical objects from measured signals in the detector is explained. Next, the strategy of data analysis is described step by step: modelling of signal and background samples, transformation of measured distribution to a true distribution by eliminating unwanted effects of the detector (unfolding). Lastly, methods used for treatment of systematic uncertain-

ties are listed. Specifically, two procedures, so called truth-based re-weighting and re-weighting based on NNLO corrections, which help with proper evaluation of signal modelling uncertainties are tested. The last chapter is dedicated to the summary of obtained results. Charge asymmetry value from the inclusive measurement, but also charge asymmetry as a function of different kinematic observables are presented. Interpretation of results in the effective field theory in a form of limits on various Wilson coefficients is given as well.

Part I

THEORETICAL BACKGROUND

Short description of the Standard Model. Top quark. Charge asymmetry explained in the context of the SM. Sensitivity of charge asymmetry to beyond SM physics. History of charge asymmetry measurements.

2

A BIT OF THE STANDARD MODEL

All stable matter consists of the up, down quarks and electrons, particles belonging to the first generation of fermions. We observe another two heavier "copies" of the first generation particles, the c-quark, s-quark and muon in second and the t-quark, b-quark and tau lepton in the third generation. All fundamental fermions together with their electric charge and color charge are listed in Tab. 1. Similar table can be written for anti-fermions, counterparts to fermions with opposite physics charges, like electric or color charge.

Table 1: Summary of fundamental fermions split to three generations. Value of electric charge and presence of color charge is stated, which indicates possible type of interactions for each fermion. Fermions with non-zero electric charge interact electromagnetically, non-zero color charge signifies strongly interacting particle.

Fermions				electric charge	color charge
quarks	u	c	t	$2/3$	yes
	d	s	b	$-1/3$	
leptons	e	μ	τ	-1	no
	ν_e	ν_μ	ν_τ	0	

Physics of elementary particles can be quite well described by the quantum field theory. The most successful theory still pertaining is the Standard Model (SM).

The key principle of the SM is the gauge invariance. Satisfying local gauge invariance of the SM Lagrangian with respect to the specific transformations provides adequate description of all fundamental forces within the model with the exception of gravity. The only caveat is that all gauge fields need to be massless in order to preserve gauge invariance.

The SM Lagrangian is gauge invariant with respect to the direct product of three groups $SU(3)_C \times SU(2)_L \times U(1)_Y$. The $SU(3)_C$ group, where subscript c implies the *color*, is the gauge group of the quantum chromodynamics (QCD), giving rise to 8 gauge bosons, gluons. Each gluon carries combination of color and anticolor charge. According to this group, each quark state comes with three possible color charges, red, blue or green, which are conserved in the strong interactions. Hence, each quark state is arranged in 3-component vector $\psi_q = (\psi_R, \psi_G, \psi_B)^T$.

Electromagnetic and weak interactions are incorporated together into the SM by constructing the SM Lagrangian locally invariant with respect to the transformations given by $SU(2)_L \times U(1)_Y$ group. In weak interactions, the parity is violated which is reflected in different treatment of left-chiral and right-chiral fermions in the SM. Left fermions, similarly right antifermions, are preferred to take part in electroweak interactions. The invariance with respect to the $SU(2)_L$ group, where L stands for *left*, takes care of the mentioned behavior, allowing only left-chiral spinors to enter into charged weak interactions. Left-fermions are organized into SU(2) doublets, like

$$Q_L^i = \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix} \qquad \text{and} \qquad L_L^i = \begin{pmatrix} \nu_L^i \\ e_L^i \end{pmatrix}, \tag{1}$$

where u_L^i denotes any up-type quark and d_L^i any down-type quark in quark doublet state Q_L^i. Index i represents a generation index of quarks. Similarly for leptons L_L^i, the upper state in the doublet is occupied by leptons with higher electric charge, which are neutrinos since they are neutral, while the lower place belongs to charged leptons with $-1e$ electric charge. Right fermions are represented as $SU(2)$ singlet states with weak isospin value T, which equals to 0. On the other hand, each $SU(2)$ doublet has weak isospin value of $1/2$, the upper state has $+1/2$ value of the third component of weak isospin T_3, the lower state $-1/2$. Analogously as for any other symmetries of the Lagrangian, specific charge is conserved in the weak interactions as well. It is the third component of the weak isospin T_3.

The inclusion of electromagnetic interactions into the SM requires the Lagrangian invariance with respect to $U(1)_Y$ group. $U(1)_Y$ symmetry of the Lagrangian implies conservation of the weak hypercharge Y. Hence, in the electroweak Lagrangian, there are 3 gauge boson fields W_μ^i corresponding to the $SU(2)_L$ group and one gauge boson B_μ for $U(1)_Y$ group.

Due to the definition of the Z boson as a linear combination of W_μ^3 and B_μ, Z boson interacts with both, left and right-chiral fermions. However, the coupling constants of the right and left fermions to Z boson are different resulting in different probabilities of their interactions with the Z boson.

All 4 gauge fields, weak isospin gauge fields W_μ^i and weak hypercharge field B_μ, are not physical fields. Linear combination of W_μ^3 and B_μ fields equivalent to rotation by weak mixing angle θ_W in the plane specified by these fields defines the physical fields of the photon and Z boson,

$$A_\mu = \cos(\theta_W) B_\mu + \sin(\theta_W) W_\mu^3, \tag{2}$$

$$Z_\mu = \cos(\theta_W) W_\mu^3 - \sin(\theta_W) B_\mu, \tag{3}$$

while the $W_\mu^\pm$ bosons are determined as $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2)$. All these physical fields, $W^\pm$, Z and photon, still remain massless, although from the experiments non-zero mass of the $W^\pm$ and Z bosons can be deduced [5, 6].

The most basic mass terms like $m^2 \bar{\psi}\psi$, cannot be added to the Lagrangian, because they are not gauge invariant. Thus, weakly inter-

acting gauge bosons, $W^{\pm}$ and Z, acquire mass by the Brout-Englert-Higgs mechanism [7–10]. The introduction of a new scalar field, the Higgs field, into the Lagrangian, provides necessary ground for the mass generation. The Higgs Lagrangian is gauge invariant, the Higgs boson being $SU(2)$ doublet. After the Higgs field obtains non-zero vacuum expectation value, the mass terms for the $W^{\pm}$ and Z emerge from its interaction terms with the Higgs field. As for fermions, their interaction with the Higgs field is included in the Lagrangian, hence after the spontaneous symmetry breaking ($SU(2)_L \times U(1)_Y$ is broken to $U(1)_{EM}$) , they acquire mass as well. The complete Lagrangian of the SM is comprised of 4 terms:

Local gauge invariance of the Lagrangian requires presence of another two terms in the SM Lagrangian: Fadeev-Popov ghost Lagrangian and gauge fixing Lagrangian, which are necessary for correct calculation of cross sections but do not introduce any real particles and thus are omitted here.

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{mass}} + \mathcal{L}_{\text{Higgs}}.$$

NOTE: Electric charge is defined as $Q = T_3 + \frac{1}{2}Y$, which is the only conserved charge after the spontaneous symmetry breaking. Weak isospin and hypercharge are not conserved in the interactions with the Higgs boson.

$\mathcal{L}_{\text{fermion}}$ contains fermion kinetic terms and their interaction with gauge fields:

$$\mathcal{L}_{\text{fermion}} = i\bar{Q}^i_L \not{D} Q^i_L + i\bar{u}^i_R \not{D} u^i_R + i\bar{d}^i_R \not{D} d^i_R + i\bar{L}^i_L \not{D} L^i_L + i\bar{e}^i_R \not{D} e^i_R, \quad (4)$$

where:

- u^i_R is right-handed up-quark-type SU(2) singlet, i is generation index,
- d^i_R is right-handed down-quark-type SU(2) singlet, i is generation index,

$\not{D} = \gamma^\mu D_\mu$

- similarly e^i_R is right-handed SU(2) singlet of leptons, i is generation index,
- D_μ is the covariant derivative, which consists of ∂_μ and terms with gauge fields, building the fermion-boson interaction terms while acting on the fermion doublet/singlet, i.e.

Exact form of D_μ differs according to fermion doublet/singlet on which it acts.

$$D_\mu Q_L = \left(\partial_\mu + i g_S G^a_\mu \frac{\lambda^a}{2} + i g W^a_\mu \frac{\sigma^a}{2} + i \frac{1}{6} g^{'} B_\mu \right) Q_L, \quad (5)$$

  where

- G^a_μ denotes gluon fields,
- λ^a are Gell-Mann matrices and σ^a are Pauli matrices,
- g, g_S and $g^{'}$ are gauge couplings for $SU(2)_L$, $SU(3)_C$ and $U(1)_Y$, respectively.

The gauge Lagrangian $\mathcal{L}_{\text{gauge}}$ represents kinetic terms for gauge bosons and their self-interaction terms:

$$\mathcal{L}_{\text{gauge}} = \frac{1}{2g_S^2}\text{Tr}\left[G^{\mu\nu}G_{\mu\nu}\right] + \frac{1}{2g^2}\text{Tr}\left[W^{\mu\nu}W_{\mu\nu}\right] - \frac{1}{4g'^2}B^{\mu\nu}B_{\mu\nu}, \quad (6)$$

where:

- $G^{\mu\nu}$ stands for gluon field-strength tensor,
- $W^{\mu\nu}$ and $B^{\mu\nu}$ define field-strength tensors for weak isospin gauge fields W_μ^i and weak hypercharge gauge field B_μ.

Part of the Lagrangian corresponding to the Higgs field, $\mathcal{L}_{\text{Higgs}}$, describes propagation of the Higgs boson and interaction of the scalar Higgs field ϕ with gauge bosons, as well as self-interactions:

$$\mathcal{L}_{\text{Higgs}} = (D^\mu\phi)^\dagger D_\mu\phi + \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2, \quad (7)$$

where:

- term $\mu^2\phi^\dagger\phi$ represents formally "mass" term of the Higgs boson. This term cannot be called real mass term due to wrong sign of the term.
- Parameter λ is a coupling constant of self-interaction of the Higgs boson.

After the Higgs field obtains non-zero vacuum expectation value ϕ_{min}, $W^\pm$ and Z bosons get mass terms in the Lagrangian. The supposed mass term for the Higgs boson, $\mu^2\phi^\dagger\phi$, obtains correct sign, hence it becomes the real mass term.

Specifically, $\phi_{min} = \frac{\mu}{\sqrt{2\lambda}} \equiv \frac{v}{\sqrt{2}}$.

After the spontaneous symmetry breaking, the $SU(2)$ Higgs doublet can be rewritten as

$$\phi(x) = \begin{pmatrix}\phi_+\\ \phi_0\end{pmatrix} \qquad \rightarrow \qquad \phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v+h(x)\end{pmatrix}, \quad (8)$$

where $h(x)$ is real scalar field. The part of the Lagrangian corresponding to the Higgs potential therefore becomes

$$\mathcal{L}_{\text{Higgs potential}} = -\mu^2h^2(x) - \lambda v h^3(x) - \frac{\lambda}{4}h^4(x). \quad (9)$$

The masses of the $W^\pm$, Z and Higgs boson are defined as

$$m_W = \frac{vg}{2}, \qquad m_Z = \frac{v}{2}\sqrt{g^2+g'^2}, \quad \text{and} \quad m_H = v\sqrt{2\lambda} = \sqrt{2}\mu. \quad (10)$$

IMPORTANT NOTE: Vacuum expectation value v and gauge couplings g, $g^{'}$ are not specified by theory, hence precise measurement of the W and Z boson mass is extremely important for determination of the fundamental SM parameters such as coupling constants.

Using the same principle, similar terms to those which allow acquisition of the $W^{\pm}$ and Z boson masses, are added in the Lagrangian for fermions:

$$\mathcal{L}_{\text{mass}} = -\Gamma_u^{ij}\bar{Q}_L^i\tilde{\phi}u_R^j - \Gamma_d^{ij}\bar{Q}_L^i\phi d_R^j - \Gamma_e^{ij}\bar{L}_L^i\phi e_R^j + h.c. \tag{11}$$

where:

- Γ_u^{ij}, Γ_d^{ij} and Γ_e^{ij} are the Yukawa couplings of up-type-quarks, down-type-quarks and leptons to the Higgs field, respectively, i, j are generation indexes
- $\tilde{\phi}$ is conjugated Higgs doublet, which equals to $i\sigma_2\phi^*$

In general, $\Gamma_{u/d/e}$ matrices are not diagonal, hence mixing among generations is present.

Fermion mass matrix can be diagonalized using bi-unitary transformation:

$$\bar{\psi}_L M\psi_R = (\bar{\psi}_L S)S^\dagger MT(T^\dagger\psi_R) = \bar{\psi}_L^{'}M_{\text{diag}}\psi_R^{'}, \tag{12}$$

where S and T are unitary matrices transforming left and right-handed Dirac spinor ψ, respectively. Rewriting the Lagrangian in the quark mass-eigenstate basis affects only one part of the SM Lagrangian, the part responsible for charged-current weak interactions. Corresponding quark current is transformed as:

$$\bar{u}_L^i\gamma^\mu d_L^i = \bar{u}_L^{'k}\gamma^\mu(S_u^\dagger)^{ki}(S_d)^{ij}d_L^{'j} = \bar{u}_L^{'k}\gamma^\mu V_{\text{CKM}}^{kj}d_L^{'j}, \tag{13}$$

obtaining the Cabibbo-Kobayashi-Maskawa (CKM) matrix V_{CKM} [11, 12]. It is 3x3-dimensional complex matrix that can be parameterized using 3 angles and one phase. The element of the CKM matrix determines probability of an up-type quark to decay into a down-type quark via interaction with the W boson. Most favored decays are those within one generation ($u \leftrightarrow d$, $c \leftrightarrow s$, $t \leftrightarrow b$), inter-generation transitions (denoted by non-diagonal elements) are suppressed. The structure of the CKM matrix introduces charge-parity (CP) violation in the SM. CP symmetry violation in the SM is not strong enough to explain matter-antimatter asymmetry.

If a theory does not differentiate between an actual particle state and state with opposite electric charge and parity, it is CP-symmetric theory. The outcome of such theory should be matter-antimatter balance, which is not observed.

IMPORTANT NOTE: Mass generation terms are constructed using a left-handed doublet of fermions, the Higgs field and a right-handed fermion singlet. For quarks, there is no difficulty because both doublet and singlet state of quarks were used to describe their electroweak interactions. The situation is different for neutral leptons, i.e. for neutrinos. Right-handed singlet for neutrino is not interacting weakly, hence has not been used to describe neutrino interactions. However,

neutrino mass term can formally be generated by the standard Higgs mechanism. According to [13], this way of neutrino mass generation is unnatural, commonly accepted opinion prefers some new, beyond the SM mechanism.

One of models suggests, that this neutrino mixing matrix should be the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix [14, 15]

Moreover, observation of neutrino oscillations has arisen requirement to distinguish between neutrino weakly interacting states (flavor states ν_ℓ, ν_μ, ν_τ) and mass states (ν_1, ν_2, ν_3 = eigenstates of free Hamiltonian), as the oscillations imply non-zero neutrino mass. Hence analogously, similar mixing matrix as was found for quarks, can be established for neutrinos.

To sum up, the SM gives us quite powerful framework capable of giving the most precise predictions so far. Nevertheless, it does not supply the answer to the fermionic mass hierarchy, i.e. it does not explain why fermions have such a different variety of masses. Also, it does not give an explanation to why the acquisition of mass for the W, Z boson and fermions is provided by the mechanism of symmetry breaking. Similarly, the values of other fundamental SM parameters like coupling constants are specified by measurements, they are not predicted by theory itself. Other observed phenomena like dark matter and dark energy are not included in the SM as well.

2.1 QCD FUNDAMENTALS

According to the observations, there are no particles or bound states with non-zero color charge. Carriers of the color charge, quarks and gluons, are confined inside hadrons. Within a hadron, they behave like quasi-free particles, when in the interaction with scanning photon, a large momentum is transferred. This attribute of the strong force is the most fundamental and is described by the running coupling constant $\alpha_S(\mu^2)$:

$$\alpha_S(\mu^2) \sim \frac{1}{\log\left(\frac{\mu^2}{\Lambda^2_{QCD}}\right)} \tag{14}$$

Coupling constant α_S is not a measurable quantity, hence the dependence on arbitrary parameter is not worrisome.

where μ denotes renormalization scale parameter. At small distances (i.e. high μ^2), the strength of the interaction decreases, but as quarks/gluons start to move away from each other, the coupling constant increases its value, effectively keeping them trapped inside a hadron. The parameter Λ_{QCD} defines boundary at which the coupling constant approaches infinity. Its value was determined to be approximately 250 GeV. The value depends on number of quark types involved in interaction. At this scale, standard perturbation theory cannot be used anymore, as the calculations will be diverging. Hence, some approximate non-perturbative theories have to be utilized for description of bound states.

2.1.1 *A picture of hadron*

However, at scales sufficiently higher than the Λ_{QCD}, a hadron can be pictured as a group of valence partons bound together by strong force, which emit softer quanta, called sea partons. These secondary partons are created in a way that does not break the coherence of a hadron, i.e. together their quantum numbers equal 0. These quantum fluctuations have limited lifetime, hence afterwards are forced to recombine.

If a hadron is probed by a photon with large momentum, it may interact with one of the partons, which results in ejecting of a parton from hadron, thus breaking the hadron coherence. Described interaction is illustrated in Fig. 1. After their interaction, quantum numbers are not preserved, color field has been violated, so remaining partons cannot recombine to establish the initial hadron. As a result, more complicated final state is formed.

Such photon is able to see the internal structure of the incident hadron, if momenta of photon and particular parton are of the same size.

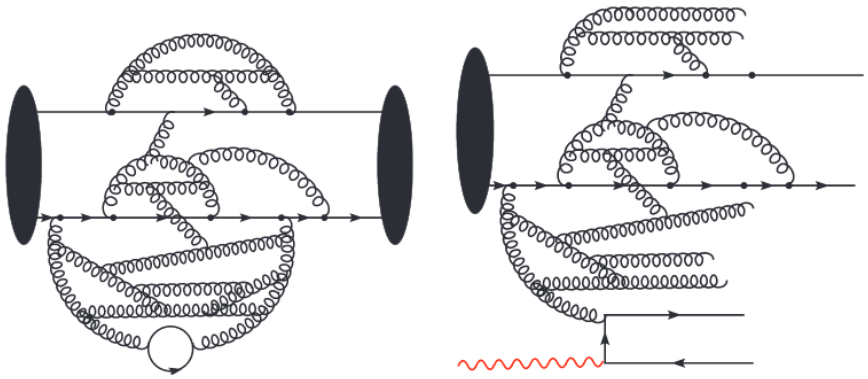


Figure 1: On the left, there is an illustration of a hadron as an ensemble of partons. After hadron-photon interaction, one parton is ejected from hadron therefore the initial hadron cannot be re-established. [16]

2.1.2 *Interaction of hadrons*

An approximation of a hadron being a group of quasi-free partons is valid in case of deep inelastic scattering, when the transferred momentum is high enough for a probe to see individual partons, assuming longitudinal parton momenta being much larger than their transverse momenta (defined with respect to the hadron momentum). Hence a hard collision of two hadrons can be treated as an interaction of two partons; of course, that is just a first step. Effectively, a calculation of the cross section can be broken into parts (assuming partons in final-state): one describing parton interaction and the second part corresponding to the probability of a parton from colliding hadron with specific hadron momentum fraction to participate in the interaction

(parton distribution function or PDF). PDF defines the probability to observe a parton with longitudinal hadron momentum fraction x in a hadron, when probed at a specific scale μ^2. The PDFs cannot be calculated perturbatively, but can be measured for various processes and at various scales assuming mentioned factorization theorem to be valid. Afterwards, the obtained results can be used to derive PDFs for different situations by using evolution equations, DGLAP equations, which take into account processes like emission of a gluon by a quark/gluon and a gluon splitting into a quark-antiquark or gluon pair, hence providing the evolution of the initial-state partons.

DGLAP is an abbreviation for Dokshitzer-Gribov-Lipatov-Altarelli-Parisi equations.

To build the full picture of a hadron-hadron collision, one needs to start with the hard scattering of two partons, accompanied by the radiation of partons from both colliding hadrons (remnants of hadrons). As those partons are losing their energies, they start to approach the scale, at which the hadron formation (hadronization) occurs. This step has to be handled non-perturbatively, hence some approximate model has to be utilized. In case of inclusive single-hadron production, analogously to PDFs, the fragmentation functions are used in the cross section calculation. These functions describe probability of a specific hadron to emerge from a parton at certain scale and carrying particular parton momentum fraction. Visual demonstration of a hadron-hadron interaction is provided in Fig. 2.

Typically, more than just two partons interact. Parton collisions not producing desired final-state are called underlying events.

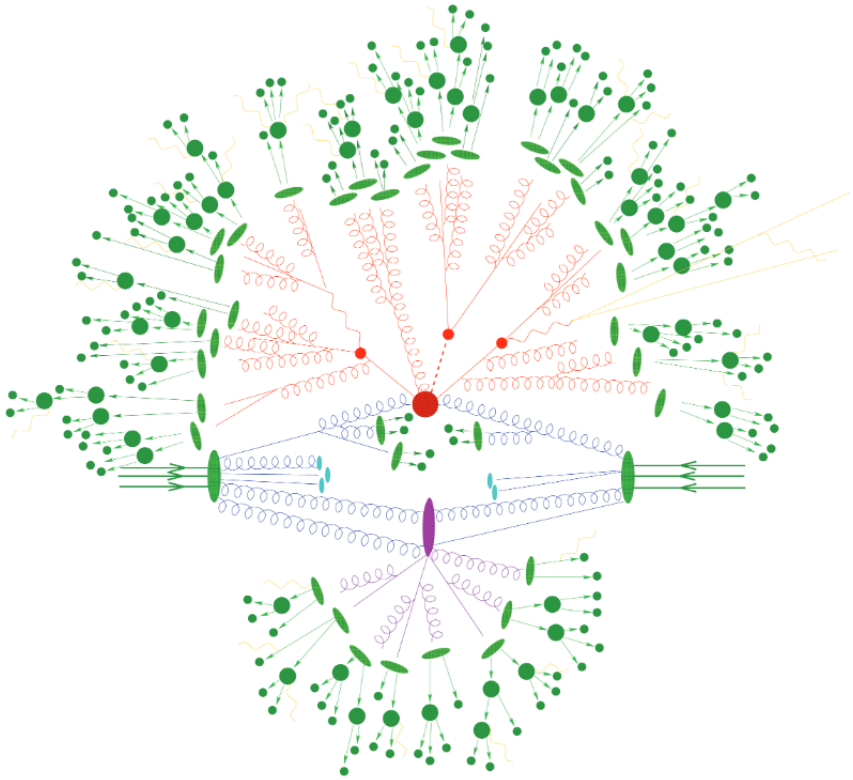


Figure 2: Illustration of a hadron-hadron collision. Red dot signifies the hard scattering of two partons, blue and red smaller markers represent initial-state radiation and final-state radiation, respectively. The underlying event is marked with purple blob. Hadronization process is depicted in green, showing birth of unstable hadrons, which are further decaying. [16]

3

TOP QUARK: POSSIBILITIES & CHALLENGES

3.1 TOP QUARK

Particles from the first generation of fermions, up, down quarks and electrons, are building blocks of all matter in the universe. Fermions from the other two generations create bound states, hadrons, as well, but they are not stable. The only exception is the t-quark. Top-flavored hadron was never detected, because the lifetime of the t-quark is extremely small, smaller than the time needed for hadronization. From the measurement of the top-quark decay width, the top-quark mean lifetime is estimated to be 5×10^{-25} s. The decay width is measured to be $\Gamma_t = 1.76 \pm 0.33(\text{stat})^{+0.79}_{-0.68}(\text{syst})$ GeV [17], while assuming the top-quark mass is equal 172.5 GeV. Assuming the same value of top-quark mass, theoretical calculation to NNLO in QCD gives the value of $\Gamma_t = 1.32$ GeV [18]. On the other hand, time needed for hadronization is approximately 3×10^{-24} s [19], which is about 10 times larger value.

Lifetime of a fermion is defined as the inverse of its decay width.

This behavior is tightly connected with its large mass. A decay width of a particle is basically defined by the number of its possible decays. Therefore, a particle with a larger mass has potentially more decay channels, and consequently, larger decay width. According to the latest measurement, the mass of the top quark is $172.13^{+0.76}_{-0.77}$ GeV [20], which makes the top quark the heaviest elementary particle known so far. Also due to its mass, the top quark is noticeably significant for studies of the Higgs boson, as the coupling constant of any fermion to the Higgs boson is proportional to a fermion mass. The top-quark mass appears at the same scale as the electroweak symmetry breaking, which suggests possible connection between these terms.

The top quark decays almost exclusively to W boson and b-quark. Its decay happens before its spin is depolarized, hence the top-quark spin and other properties can be accessed by measuring attributes of its decay products.

3.1.1 *Top-quark production at the LHC*

In proton-proton collisions, top quark is produced mainly in pairs, together with its antiparticle, top antiquark, as a result of strong interactions. On the other hand, weak force is held responsible for single top-quark production, hence it appears less frequently. Looking at the collision of two protons as an interaction of two partons, we can analyze possible processes producing top-quark pair ($t\bar{t}$). The probability

of each $t\bar{t}$ production mechanism depends on the possibility of a parton with proton momentum fraction x to take part in the interaction. The PDFs are dependent on the renormalization scale μ_r and factorization scale μ_f. The choice of these scales is arbitrary to same extent, hence often the same value is chosen for both scales $Q = \mu_f = \mu_r$. The PDFs are extracted from QCD global fits to experimental data taken in deep inelastic scattering experiments. An illustration of PDFs for proton at scale $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2 is shown in Fig. 3.

For $t\bar{t}$ production, the scale Q is often selected to be m_t.

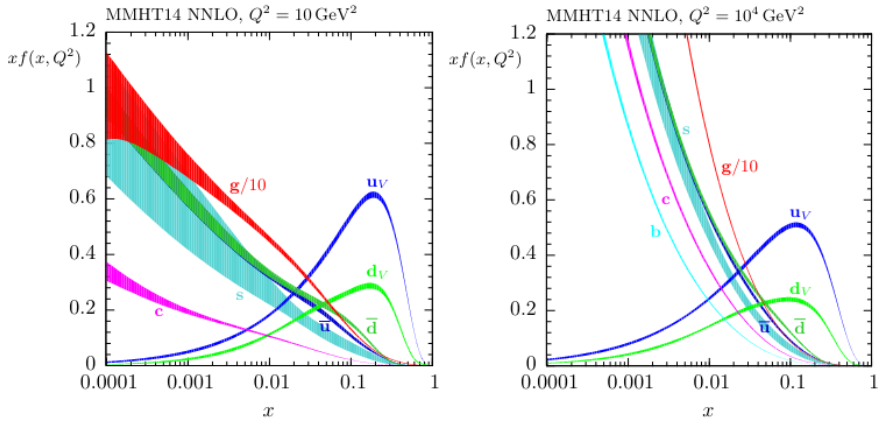


Figure 3: Parton momentum densities in proton as a function of the longitudinal proton momentum fraction x at $Q^2 = 10$ GeV^2 (left) and $Q^2 = 10^4$ GeV^2 (right). [21]

Given minimal energy to produce $t\bar{t}$ pair, $2m_t$, and center-of-mass energy $\sqrt{s}$ of collisions in the LHC, 13 TeV, typical value of parton momentum fraction x for $t\bar{t}$ production is ≈ 0.025 [22]. By looking at the PDF in Fig. 3, one can see higher probabilities of interaction for gluons in low x region than for quarks/antiquarks. That indicates the dominance of gluon-gluon fusion as a production mechanism for $t\bar{t}$ pair at the LHC.

Simplifying situation by assumption that both partons have same x, one can obtain typical values of parton momentum fraction for $t\bar{t}$ production as $x \approx \frac{2m_t}{\sqrt{s}}$. [22]

To sum up, there are three processes responsible for $t\bar{t}$ production:

1. gluon-gluon fusion – dominant production channel, almost 90%,
2. quark-antiquark annihilation – in approximately 10% of cases,
3. quark-gluon scattering – top-quark pair emerges together with additional quark.

NOTE: With rising center-of-mass energy $\sqrt{s}$, partons with even lower values of x will be able to produce $t\bar{t}$ pair, hence gluon-gluon interaction will become more and more prevailing as the $t\bar{t}$ production process.

Top-quark is produced at small distances ($1/m_t$). The value of coupling constant α_S for its production is quite small, approximately 0.1 (see Eq. 14). Therefore, the perturbative calculation of $t\bar{t}$ production

is possible. The latest calculation to NNLO in QCD includes terms with α_S^4 and gluon resummation to next-to-next-to-leading logarithmic (NNLL) accuracy. Therefore, by measurements of $t\bar{t}$-production cross section, the attributes of QCD are tested. In Fig. 4, summary of performed measurements of $t\bar{t}$ cross section are listed together with the theoretical prediction.

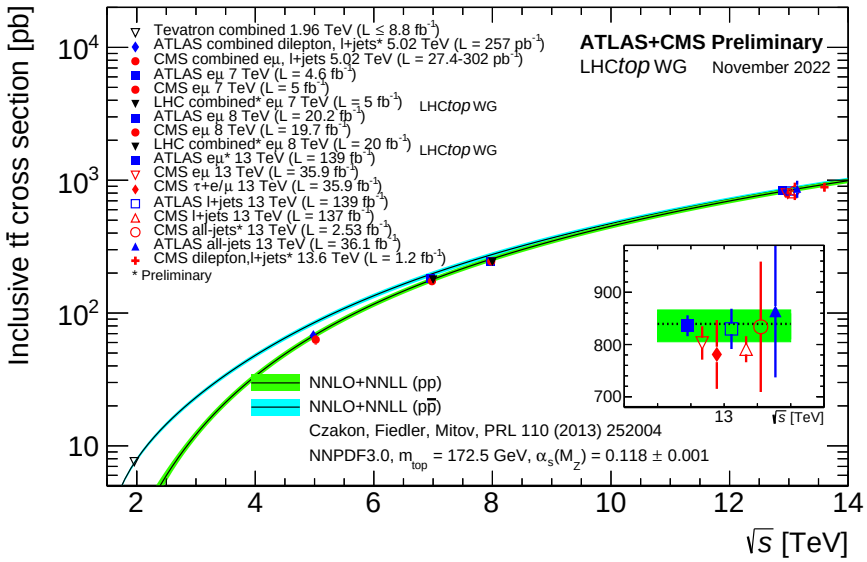


Figure 4: Comparison of $t\bar{t}$ cross section measurements performed by the ATLAS and CMS Collaborations at various center-of-mass energies in pp collisions, and at Tevatron in $p\bar{p}$ collisions, with theoretical calculation in the SM to NNLO in QCD and to NNLL in gluon resummation accuracy [23, 24].

3.1.2 *Top-quark pair decay*

As was previously mentioned, the top quark predominantly decays into W boson and b-quark. When the top quark is produced at the LHC as a $t\bar{t}$ pair, it can be detected only through identification of the decay products, $t\bar{t} \to W^+bW^-\bar{b}$. Depending on the way the W boson decays, three possible decay channels are identified:

SINGLE-LEPTON CHANNEL (ℓ+JETS): In this type of $t\bar{t}$ decay, there is only one lepton in the final state. Each W boson decays differently, one decays into a lepton and its neutrino, the other into pair of quarks (quark and antiquark) with different flavor. The $t\bar{t}$ pair decays through the single-lepton channel in 43.8% of cases [19].

$$t\bar{t} \to W^+W^-b\bar{b} \to \ell^-\,\bar{\nu}_\ell\,q\,\bar{q}'\,b\bar{b} + \ell^+\,\nu_\ell\,q'\,\bar{q}\,b\bar{b}$$

DILEPTON CHANNEL (2ℓ): Significant feature of the dilepton channel is presence of two leptons in the final state. Both W bosons decay leptonically, into a lepton and neutrino. The dilepton decay channel is the least probable of all channels at probability rate only 10.5% [19].

$$t\bar{t} \to W^+W^-b\bar{b} \to \ell^-\,\ell'^+\,\bar{\nu}_\ell\,\nu_{\ell'}\,b\bar{b}$$

ALL-HADRONIC CHANNEL: As the name of the decay channel suggests, there is no lepton in the final state. Both W bosons decay into a pair of quark and antiquark. The all-hadronic channel contributes to the total $t\bar{t}$ decay in 45.7% of cases, which means it is the most probable decay process of $t\bar{t}$ pair [19].

$$t\bar{t} \to W^+W^-b\bar{b} \to\; q\,\bar{q}'\,q''\,\bar{q}'''\,b\bar{b}$$

Given probability rates were estimated assuming lepton universality, i.e. ℓ can be electron, muon or tau lepton [19].

3.2 CHARGE ASYMMETRY

Longitudinal momentum is a momentum projection onto the direction of the beam of colliding particles.

In strong production of heavy-flavor $Q\bar{Q}$ pair, the difference between quark and antiquark distributions of i.e. longitudinal momentum or angle of their flight direction with respect to the beam of colliding particles, is present. This effect is called charge asymmetry.

The term charge asymmetry is used to describe the phenomenon because the cross section of $Q\bar{Q}$ production is not charge symmetric (under the exchange of Q and $\bar{Q}$), which manifests itself in different distributions for Q and $\bar{Q}$.

The production mechanisms of $Q\bar{Q}$ pair in hadron collisions were introduced in Sec. 3.1.1 for $t\bar{t}$ pairs. In the lowest order of calculation (Born approximation), the amplitude of process is charge symmetric. Until the higher order diagrams are considered, there is no charge asymmetry present. As a production mechanism for charge asymmetry, the interaction of $q\bar{q}$ or qg should be examined. The fusion of two gluons produces charge symmetric contribution to the amplitude of $Q\bar{Q}$ pair production. The most important process in terms of charge asymmetry is the quark-antiquark annihilation. The contribution of quark-gluon interaction is less dominant [25, 26]. In Fig. 5, representative diagrams for $q\bar{q}$ annihilation which already produce charge asymmetric contribution to the amplitude are shown. In fact, it is the interference term, which arises when calculating cross section (amplitude squared) from the diagrams shown in Fig. 5, that is odd under the exchange of Q and $\bar{Q}$.

Considering only these diagrams as relevant ones for the first approximation of the charge asymmetric $Q\bar{Q}$ cross section does not give us finite result. To remove the infrared divergence, diagrams with additional emission of gluon have to be added to the evaluation of the

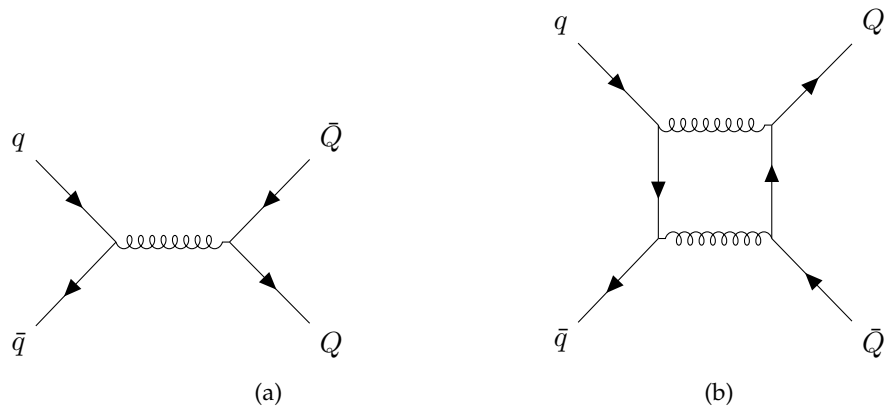


Figure 5: Representative diagrams for quark-antiquark annihilation on the tree level (a) and box diagram (b).

inclusive cross section, see Fig. 6. These diagrams have different final state, $Q\bar{Q}+g$, hence do not contribute directly to the process amplitude, only to the inclusive cross section. Finally, this completes the input to the calculation of $Q\bar{Q}$ cross section from $q\bar{q}$ annihilation at the lowest possible order which exhibits charge asymmetry.

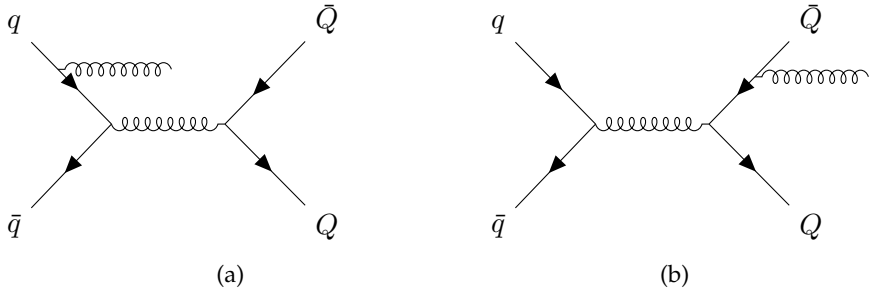


Figure 6: Representative diagrams with real gluon emission in the (a) initial and (b) final state.

Except pure QCD processes, $Q\bar{Q}$ pair can be produced in electroweak interactions, too. By exchanging one virtual gluon in each of diagrams in Fig. 5 for photon or Z boson, we get the electroweak $Q\bar{Q}$ production processes. The mixed QCD-electroweak interference terms, which arise after inclusion of these diagrams into cross section calculation, are charge asymmetric. This provides small corrections up to the discussed order of the calculation [25, 26].

3.2.1 *Charge asymmetry in $t\bar{t}$ production*

Clearly, the size of this phenomenon depends on the type of hadronic collision in which the heavy $q\bar{q}$ pair (in our case $t\bar{t}$) is produced.

Considering two types of collisions, proton-proton (pp) and proton-antiproton ($p\bar{p}$), production mechanisms of $t\bar{t}$ pair in either of these collisions are the same, see Sec. 3.1.1. Contribution of each production channel to the $t\bar{t}$ cross section is different though. Unlike in pp initial state, $t\bar{t}$ pair is dominantly produced by $q\bar{q}$ annihilation in $p\bar{p}$ collisions, making up 85% of total $t\bar{t}$ production. In remaining 15% of cases, gluon fusion is responsible for top-pair production [27]. The $t\bar{t}$ production channel contributing mostly to the charge asymmetric part of the cross section is the $q\bar{q}$ annihilation, hence the value of charge asymmetry is expected to be much higher for $p\bar{p}$ collisions than for pp collisions. It is vital to use generalization that charge asymmetry is exhibited only in $q\bar{q}$ initiated $t\bar{t}$ production regardless of collision type when discussing charge asymmetry definitions.

Collision types for discussion are chosen according to the two large particle colliders, LHC and Tevatron, where pp and $p\bar{p}$ collisions have been performed, respectively.

Charge asymmetry is manifested in different distributions for top quark and top antiquark. Choice of suitable observable is dependent on experimental setup, i.e. type of hadronic collision. In $p\bar{p}$ collisions, it is possible to study charge asymmetry as a function of production angle θ of top quark/antiquark in $p\bar{p}$ rest frame, see Fig. 7. By counting number of events when top quark is produced in the direction of proton (initial-state quark) and when it is top antiquark which arises in proton direction, the forward-backward asymmetry in $p\bar{p}$ rest frame can be defined:

$$A_{FB}^{p\bar{p}} = \frac{N_t(\cos\theta \geq 0) - N_{\bar{t}}(\cos\theta \geq 0)}{N_t(\cos\theta \geq 0) + N_{\bar{t}}(\cos\theta \geq 0)}. \tag{15}$$

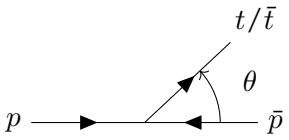


Figure 7: Illustration of $p\bar{p}$ rest frame and production angle θ of top quark/-top antiquark.

Considering $t\bar{t}$ rest frame instead, more convenient observable for charge asymmetry measurement is rapidity difference of top quark and top antiquark. This observable, $\Delta y = y_t - y_{\bar{t}}$, is not affected by Lorentz boost of $t\bar{t}$ along the beam axis, yet still describes differences between top and antitop directions. An illustration of how the rapidity distributions for top and top antiquark look like is presented in Fig. 8. Hence analogously, forward-backward asymmetry for $t\bar{t}$ rest frame is calculated as:

$$A_{FB}^{t\bar{t}} = \frac{N(\Delta y > 0) - N(\Delta y < 0)}{N(\Delta y > 0) + N(\Delta y < 0)}, \tag{16}$$

According to the SM prediction, there is a preference for top quarks to be produced in the direction of initial-state quark. Consequently, top antiquarks tend to be produced in the antiquark direction. In $p\bar{p}$

Rapidity is defined as $y = \frac{1}{2}\ln\left(\frac{E+p_z c}{E-p_z c}\right)$, where E is the energy of a particle and p_z is the projection of particle momentum to the z-axis. Orientation of z-axis is crucial for rapidity sign. For $p\bar{p}$ collisions, z-axis is oriented in the direction of proton, hence rapidity of a particle going in the same direction as proton is positive.

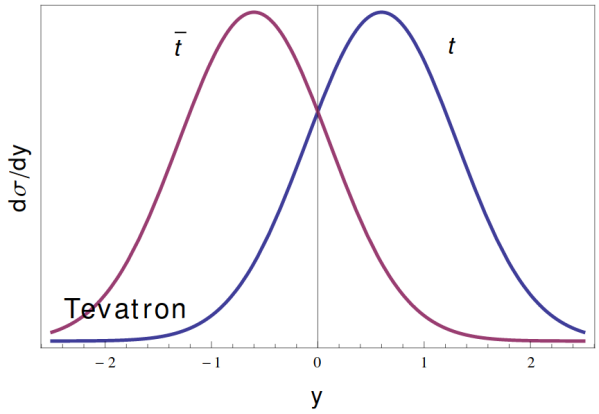


Figure 8: Illustration of rapidity distribution for top and antitop quark in proton-antiproton collision [28].

collisions, quark direction can be unified with proton direction and the same is assumed for antiquark and antiproton.

However, similar assumptions cannot be made for pp collisions. In pp collisions, initial-state is charge symmetric, so the terms "forward" as a direction of quark (from proton) and "backward" as a direction of antiquark (from antiproton) lose their meaning for charge asymmetry definition. Nevertheless, rapidity is still used as an observable, but the distributions for top quark and top antiquark look different than in $p\bar{p}$ collision. Apart from $p\bar{p}$ collision, in pp collision there is a longitudinal momentum imbalance of initial-state quark and antiquark. This difference is transferred to rapidity distributions of top quark and top antiquark. The inclination of top quarks to be produced in initial-state quark direction is expressed as a predominance of top quarks with large absolute rapidity values. On the other hand, top antiquarks are produced more centrally, i.e. with smaller absolute rapidity values, as is depicted in Fig. 9.

Hence more suitable adjective for charge asymmetry in pp collision is "peripheral-central" asymmetry. Using absolute rapidity difference as an observable, $\Delta|y| = |y_t| - |y_{\bar{t}}|$, charge asymmetry is determined as:

$$A_C^{t\bar{t}} = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)}. \tag{17}$$

3.2.2 *Latest charge asymmetry prediction within the SM*

In Sec. 3.2, an idea of how charge asymmetry originates is outlined considering diagrams of the lowest possible order. However, a sizable corrections to charge asymmetry value arise from higher order diagrams. Latest charge asymmetry calculation [29] includes complete set of NLO diagrams describing both strong and electroweak $t\bar{t}$ pro-

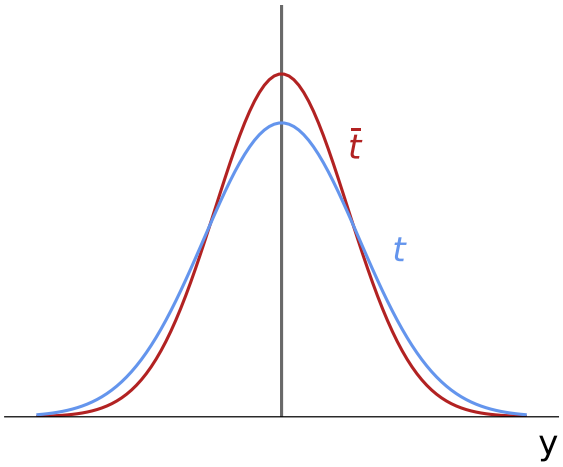


Figure 9: Illustration of rapidity distribution for top quark and top anti-quark in proton-proton collision.

duction and NNLO diagrams in strong interaction. Both NNLO QCD and NLO electroweak (EW) contributions are significant for proper evaluation of charge asymmetry value. Charge asymmetry prediction after inclusion of NNLO QCD ($\mathcal{O}(\alpha_S^4)$) and NLO electroweak corrections is shown in Fig. 10 for pp collisions with a center-of-mass energy of 13 TeV. The effect is tiny in size, below 1% [26]. For $p\bar{p}$ collisions performed at the Tevatron, the value is about 10 times larger [29, 30].

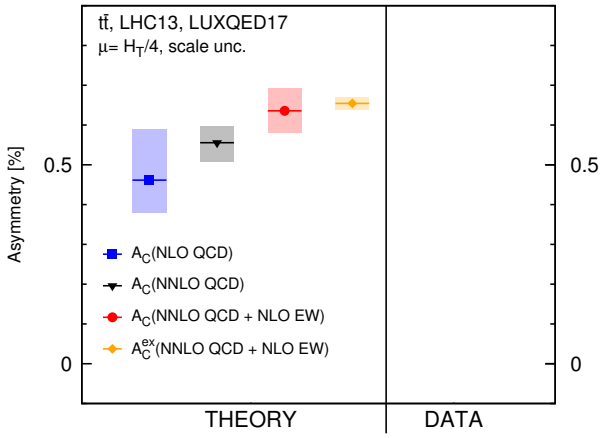


Figure 10: Prediction for inclusive charge asymmetry measurement in $t\bar{t}$ production for pp collisions at 13 TeV, evaluated to either NLO in QCD, NNLO in QCD or to NNLO in QCD + NLO in EW theory [29]. The A_C^{ex} is called expanded charge asymmetry and it denotes different approach to the charge asymmetry calculation. The A_C^{ex} definition is in [29].

Of course there exist specific kinematic regions, where the enhancement of charge asymmetry is predicted. Hence, it is essential to study

charge asymmetry in some deliberately chosen parts of phase space, where contribution of $q\bar{q}$ annihilation is expected to be higher. Setting cuts or slicing phase space according to values of some observable like invariant mass of top pair ($m_{t\bar{t}}$) helps to examine A_C behavior in various phase space regions. Prediction of $m_{t\bar{t}}$-dependent charge asymmetry in pp collisions for 5 regions of $m_{t\bar{t}}$ is shown in Fig. 11. Charge asymmetry value has a tendency to rise with increasing $m_{t\bar{t}}$, because contribution of $q\bar{q}$ initiated $t\bar{t}$ production is rising as well. The reason lies in the PDF of partons. When $t\bar{t}$ pair is boosted along the beam axis with respect to the laboratory frame (which is also the case for $t\bar{t}$ pair with high invariant mass), it is most probably produced by interaction of $q\bar{q}$ or qg due to the fact that quarks carry larger longitudinal momentum fraction of proton than gluons. Selection of these boosted events can be provided by selection requirements on the $m_{t\bar{t}}$, the transverse momentum of $t\bar{t}$ pair ($p_T^{t\bar{t}}$) or directly on longitudinal boost of $t\bar{t}$ pair along the beam axis $\beta_z^{t\bar{t}}$.

In the second panel from the top of Fig. 11, the effects of NLO in EW corrections are presented by comparing of NNLO in QCD + NLO in EW prediction (red band) and NNLO QCD result (grey band), both normalized to NNLO in QCD central value. The impact of EW corrections is quite significant and increases with higher values of $m_{t\bar{t}}$. Higher order QCD corrections change the shape of distribution notably. The size of the effect is demonstrated in the third panel from the top of Fig. 11, where NNLO in QCD and NLO in QCD results are compared.

3.2.3 *Brief history of charge asymmetry measurements in $t\bar{t}$ production*

Forward-backward asymmetry was measured for the first time in $p\bar{p}$ collisions at the Tevatron [31] at a center-of-mass energy of $\sqrt{s} = 1.96$ TeV. Obtained results were within 2σ uncertainty band consistent with the SM prediction [32]. Afterwards, both the CDF [33] and D0 [34] collaborations examined collected data with higher integrated luminosity for more precise A_{FB} determination [35, 36]. These measurements observed discrepancy of more than 3σ from the SM A_{FB} value [36]. At that time, previously discussed higher order corrections were not calculated, hence the measured discrepancies motivated scientists to investigate various beyond the SM theories, which could have explained the excess. Measurement of the charge asymmetry A_C at the LHC in pp collisions was also encouraged by the disagreement between the measured A_{FB} and the SM prediction calculated at that time [37]. Calculation of NNLO in QCD contribution to A_{FB} clarified measured inconsistency. The latest D0 and CDF A_{FB} measurements [38] are in good agreement with the SM prediction [30].

Charge asymmetry was also measured in pp collisions with the CMS and ATLAS detectors at various center-of-mass energies $\sqrt{s} =$

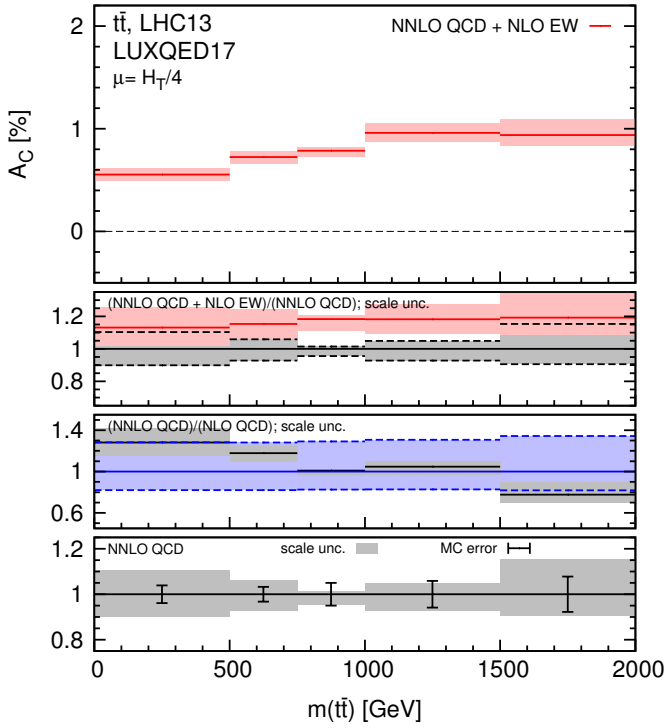


Figure 11: Prediction for charge asymmetry measurement with respect to different values of $m_{t\bar{t}}$ in $t\bar{t}$ production for pp collisions at 13 TeV (first panel from the top). Effect of EW corrections (second panel from the top) as well as higher order QCD corrections (third panel from the top) are inspected [29].

7, 8 or 13 TeV [1, 39–44]. Results from the CMS and ATLAS measurements using data from single-lepton $t\bar{t}$ decay channel at 7 TeV and 8 TeV together with their combinations are summarized in Fig. 12. The comparison of results with theoretical predictions to either NNLO in QCD + NLO in EW theory or just NLO in QCD + NLO in EW theory is also displayed. With larger amount of collected data, from 5 fb^{-1} for 7 TeV to 20 fb^{-1} for 8 TeV, the precision of results is improved, suggesting the statistical uncertainty plays significant role in this measurement.

Charge asymmetry as a function of $m_{t\bar{t}}$ was studied at 8 TeV center-of-mass energy for single-lepton decay channel with the CMS and ATLAS detectors. These measurements were combined to obtain more precise results [45].

Measurement of charge asymmetry was performed for the dilepton $t\bar{t}$ decay channel as well, with the CMS and ATLAS detectors for $\sqrt{s} = 7, 8$ TeV [40, 46–48]. In dilepton channel, there is a possibility to measure leptonic charge asymmetry, which quantifies the same effect, but not from the distributions of t and $\bar{t}$. It is determined from

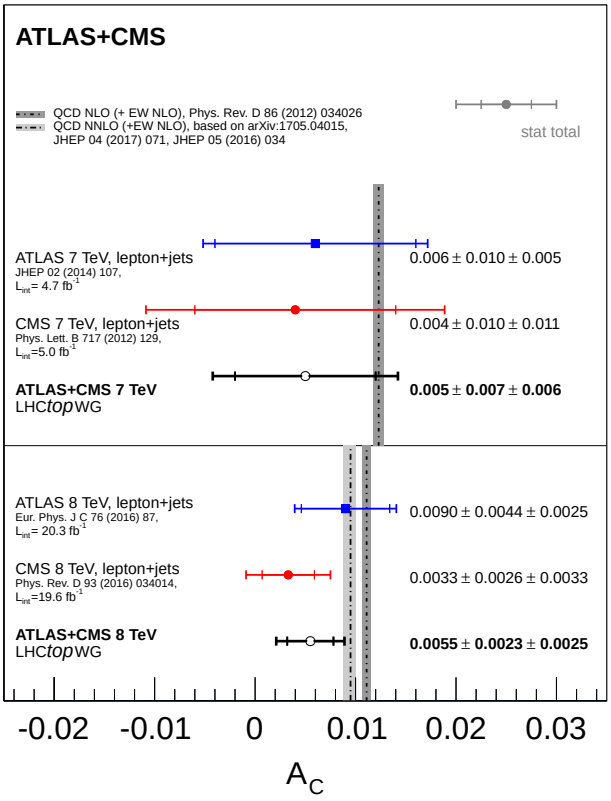


Figure 12: Summary of the ATLAS and CMS charge asymmetry inclusive measurements for single-lepton channel at 7 TeV and 8 TeV and combination of these measurements compared with theory predictions (calculated at (N)NLO in QCD, including NLO electroweak corrections). The inner bars of data points show the statistical uncertainty, the outer bars display the total uncertainty [45].

distributions of leptons arising from their decays. The leptonic $A_C^{\ell\bar{\ell}}$ is therefore defined as

$$A_C^{\ell\bar{\ell}} = \frac{\sigma_{t\bar{t}}(\Delta|\eta(\ell,\bar{\ell})| > 0) - \sigma_{t\bar{t}}(\Delta|\eta(\ell,\bar{\ell})| < 0)}{\sigma_{t\bar{t}}(\Delta|\eta(\ell,\bar{\ell})| > 0) + \sigma_{t\bar{t}}(\Delta|\eta(\ell,\bar{\ell})| < 0)}, \tag{18}$$

where $\sigma_{t\bar{t}}$ denotes the integrated cross section (number of events) with positive/negative difference of absolute values of lepton pseudorapidities.

$A_C^{\ell\bar{\ell}}$ is diluted in comparison with the $t\bar{t}$ charge asymmetry, because leptons do not follow the exact same path as quarks.

The first inclusive $t\bar{t}$ and leptonic charge asymmetry in dilepton channel for the highest achieved center-of-mass energy of $\sqrt{s} = 13$ TeV for data corresponding to integrated luminosity of 35.9fb^{-1} was measured with the CMS detector in 2019 [44]. Later there was a CONF note published by the ATLAS collaboration showing inclusive and differential A_C measurements (with respect to $\beta_z^{t\bar{t}}$ and $m_{t\bar{t}}$) in single-lepton channel at $\sqrt{s} = 13$ TeV as well but with data corresponding to 139fb^{-1} [1]. The latest measurement of $t\bar{t}$ charge asymmetry in single-lepton channel using data collected by the CMS detector at

$\sqrt{s} = 13$ TeV is focused on boosted events with $m_{t\bar{t}} > 750$ GeV [49]. All measurements which were carried out so far show no significant deviations from the SM predicted values.

Limitation of this measurement is of statistical nature. It is not driven by systematic effects, hence the more data is taken, the more precise the result will be. With increasing centre-of-mass energy up to 13 TeV the rate of $q\bar{q}$ annihilation rises as well. However, the ratio of gluon-fusion initial states will be increasing more rapidly. To sum up, the inclusive charge asymmetry is expected to be smaller when comparing 13 TeV and 8 TeV data at the LHC. Nevertheless, the new charge asymmetry measurement at 13 TeV is still worth to perform, mainly due to reduction of the total uncertainty and significant improvement of exclusion limits on operators in effective field theory (EFT) [2]. Moreover, higher centre-of-mass energy provides unique possibility to measure charge asymmetry in high $m_{t\bar{t}}$ region or region of high values of $\beta_z^{t\bar{t}}$. In these areas of specific kinematic variables, charge asymmetry enhancement is predicted [29]. Region of high $\beta_z^{t\bar{t}}$ actually describes region with larger values of $|y_{t\bar{t}}|$ of $t\bar{t}$ pair corresponding to the boosted events [50] as can be seen in equation

$$y_{t\bar{t}} = \frac{1}{2} \ln\left(\frac{1+\beta_z^{t\bar{t}}}{1-\beta_z^{t\bar{t}}}\right). \tag{19}$$

The $t\bar{t}$ events, which are boosted along the z-axis, are typically the final states of $q\bar{q}$ annihilation. That is the reason why is the charge asymmetry enhanced in high $\beta_z^{t\bar{t}}$ region.

Study of charge asymmetry as a function of invariant mass of the $t\bar{t}$ pair, $m_{t\bar{t}}$, has a potential to reveal indications of new physics. Processes with a new heavy intermediate particle can interfere with the SM $t\bar{t}$ production, contributing to the charge asymmetry either positively or negatively [51].

Charge asymmetry measurement with respect to the transverse momentum of $t\bar{t}$ system $p_T^{t\bar{t}}$ mirrors the sensitivity of the differential observable to different asymmetry sources. In the low $p_T^{t\bar{t}}$ region, the $t\bar{t}$ production through the box and Born diagrams (Fig. 5) is dominant, so the interference term of the process amplitudes of these diagrams contributes to the asymmetry by increasing its value. Contrasting with the high $p_T^{t\bar{t}}$ region, where the inital-state-radiation (ISR) and final-state-radiation (FSR) diagrams (Fig. 6) dominate as the $t\bar{t}$ production mechanisms, their interference term produces negative A_C contribution [51].

3.2.4 *Charge asymmetry in effective field theory*

There is no doubt that the SM is not a complete theory, given many issues which are still unsolved. In Chap. 2, some of its imperfections were outlined like arbitrariness of parameters – particle masses, cou-

pling constants. Hence, it is reasonable to think about the SM as an effective theory, theory applicable only to physical processes below some energy scale Λ. Description of a new physics at higher energy level than Λ can be included by formulating new Lagrangian, the effective Lagrangian $\mathcal{L}_{eff}$, whose low-energy limit will be the SM Lagrangian. Basically, the SM Lagrangian can be extended by new terms representing effects of new physics in a form of higher-dimensional operators suppressed by powers of scale Λ:

$$\mathcal{L}_{eff} = \mathcal{L}_{SM}^{(4)} + \frac{1}{\Lambda}\sum_i C_i^{(5)} O_i^{(5)} + \frac{1}{\Lambda^2}\sum_i C_i^{(6)} O_i^{(6))} + \mathcal{O}\left(\Lambda^{-3}\right), \quad (20)$$

where O_i^n are n-dimensional operators (n going from 5 and higher) and C_i^n are their dimensionless coupling constants – Wilson coefficients. This extended Lagrangian describes new physics quite generally, without any dependency on type of new interactions in terms of the SM particles. The only requirement is that new operators should be invariant with respect to $SU(3)_C \times SU(2)_L \times U(1)_Y$ group.

SM Langrangian contains only two- and four-dimensional operators.

Even if the energy scale of new physics is not accessible by our current experiments, it can influence the measurements on lower energy scale. Top quark being the heaviest particle (the closest to the energy scale of new physics) is the best observable for search of new physics. Effective interactions of top-quark can cause deviations in top-quark related measurements. Hence, via measurements of top-quark properties and cross section, the deviations from the SM prediction can be spotted which can be subsequently interpreted in the effective field theory. If no deviations are detected, the measurements can still be helpful with setting constraints on particular Wilson coefficients. Differential measurements selecting $t\bar{t}$ events with high invariant mass $m_{t\bar{t}}$ are of special interest, because of their large sensitivity to physics beyond the SM at high energy scale.

The leading contribution ($\sim \Lambda^{-2}$) from effective operators for any physical observable originates from the interference of the dimension-six operators with the SM Lagrangian. The relevant effective operators for A_C calculation are those contributing to $t\bar{t}q\bar{q}$ interaction – four-fermion operators O_{4q} and corrections to $t\bar{t}g$ coupling O_{tG} [52], see Fig. 13. For charge asymmetry measurement, 15 dimension-six operators are taken into account considering different color structures and chirality of quarks. Operators in the Warsaw basis [53] of structure $(\bar{L}L)(\bar{L}L)$ or $(\bar{R}R)(\bar{R}R)$ are listed here:

Dimension-five operator which is gauge invariant with respect to the SM group is only one – it can produce Majorana mass term for neutrinos after spontaneous symmetry breaking. *Majorana particle = particle and antiparticle are the same*

$$\begin{aligned}
O_{Qq}^{(8,1)} &= (\bar{Q}\gamma_\mu T^A Q)(\bar{q}_i\gamma^\mu T^A q_i) & O_{Qq}^{(1,1)} &= (\bar{Q}\gamma_\mu Q)(\bar{q}_i\gamma^\mu q_i)\\
O_{Qq}^{(8,3)} &= (\bar{Q}\gamma_\mu T^A \tau^I Q)(\bar{q}_i\gamma^\mu T^A \tau^I q_i) & O_{Qq}^{(1,3)} &= (\bar{Q}\gamma_\mu \tau^I Q)(\bar{q}_i\gamma^\mu \tau^I q_i)\\
O_{tu}^{(8)} &= (\bar{t}\gamma_\mu T^A t)(\bar{u}_i\gamma^\mu T^A u_i) & O_{tu}^{(1)} &= (\bar{t}\gamma_\mu t)(\bar{u}_i\gamma^\mu u_i)\\
O_{td}^{(8)} &= (\bar{t}\gamma^\mu T^A t)(\bar{d}_i\gamma_\mu T^A d_i) & O_{td}^{(1)} &= (\bar{t}\gamma^\mu t)(\bar{d}_i\gamma_\mu d_i)\,.
\end{aligned}$$

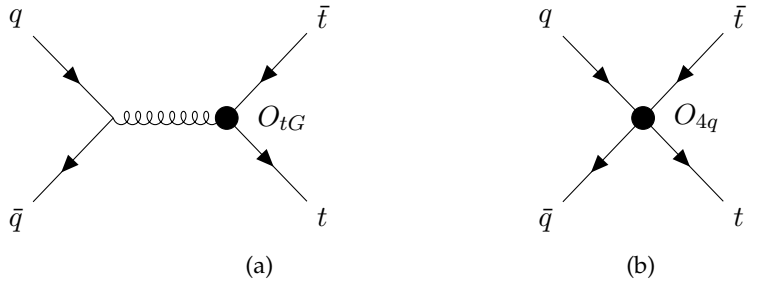


Figure 13: Representative diagrams for quark-antiquark annihilation illustrating impact of EFT operators in $t\bar{t}$ production. Corrections on (a) the $t\bar{t}g$ coupling and (b) the new four-fermion interaction [52].

$T^a = \frac{1}{2}\lambda^a$, *where* λ^a *denotes Gell-Mann matrices,* τ^I *are Pauli matrices.*

Another six operators following $(\bar{L}L)(\bar{R}R)$ structure are:

$$\begin{aligned} O_{Qu}^{(8)} &= (\bar{Q}\gamma_\mu T^A Q)(\bar{u}_i\gamma^\mu T^A u_i) & O_{Qu}^{(1)} &= (\bar{Q}\gamma_\mu Q)(\bar{u}_i\gamma^\mu u_i) \\ O_{Qd}^{(8)} &= (\bar{Q}\gamma_\mu T^A Q)(\bar{d}_i\gamma^\mu T^A d_i) & O_{Qd}^{(1)} &= (\bar{Q}\gamma_\mu Q)(\bar{d}_i\gamma^\mu d_i) \\ O_{tq}^{(8)} &= (\bar{t}\gamma^\mu T^A t)(\bar{q}_i\gamma_\mu T^A q_i) & O_{tq}^{(1)} &= (\bar{t}\gamma^\mu t)(\bar{q}_i\gamma_\mu q_i)\,. \end{aligned}$$

where the doublet of the third generation is given by $Q = (t_L, b_L)$ and q_i denotes quark left doublets for the first two generations (u_L, d_L), (c_L, s_L). The u_i, d_i are right-handed up, down-type quarks for the first two generations, right-handed top quark is denoted as t. The first upper index is a label for color structure: (1) means color singlet operator, (8) means color octet operator. The second index, if present, informs about operator being weak isospin singlet (1) or triplet (3). The last considered operator is tensor operator O_{tG}:

For example, operator $O_{tq}^{(8)}$ *may represent an interaction of light* $q\bar{q}$ *with* t_R *and* $\bar{t}_R$ *as final-state quarks, mediated by heavy gluon-like particle, which can be approximated by a constant (Wilson coefficient* $C_{tq}^{(8)}$*).*

$$O_{tG} = (\bar{t}\sigma^{\mu\nu}T^A t)\tilde{\phi}G^A_{\mu\nu}. \tag{21}$$

Some of previously mentioned operators are redundant, because they can be rewritten as a linear combination of others. Hence there are just 7 independent operators. These can be further combined into 4 linear combinations after assuming initial $q\bar{q}$ state to be most probably $u\bar{u}$ or $d\bar{d}$. As a result, 4 operators are specified with corresponding Wilson coefficients $C_{1,2}^u, C_{1,2}^d$ [52]. Using this combination of effective operators simplifies their impact on charge asymmetry. The charge asymmetry value depends on $C_1^{u,d} - C_2^{u,d}$, whereas $t\bar{t}$ cross section is sensitive to their sum $C_1^{u,d} + C_2^{u,d}$ [52]. In this sense charge asymmetry measurement is complementary to cross section measurement as shown in Fig. 14, where bounds for two Wilson coefficients C_1, C_2 are depicted. Here the assumption of the same coupling of effective operators to u and d quark, $C_1^u = C_1^d = C_1$ and $C_2^u = C_2^d = C_2$, is applied [54].

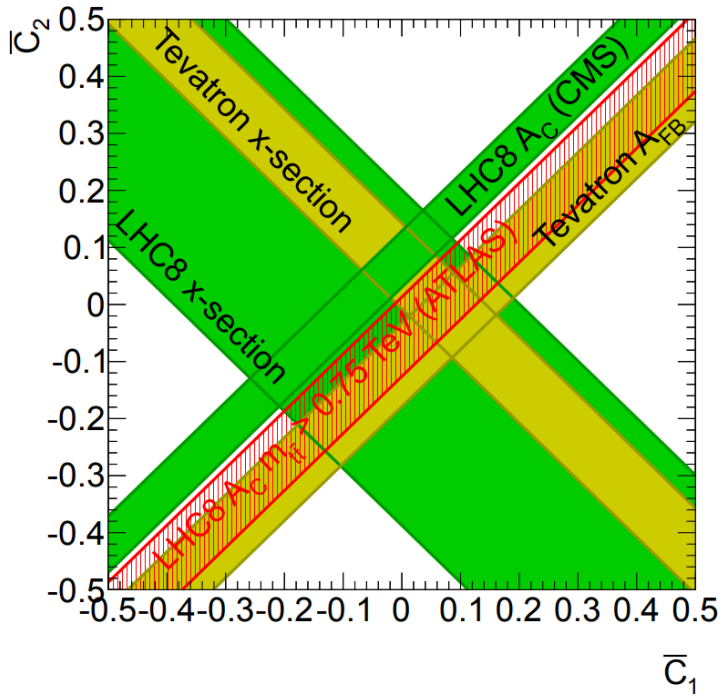


Figure 14: Constraints on C_1 and C_2 Wilson coefficients obtained from LHC measurements at $\sqrt{s} = 8$ TeV: $t\bar{t}$ cross section, inclusive A_C measurement with the CMS detector, differential A_C measurement with the ATLAS detector using events with specific values of $m_{t\bar{t}}$, Tevatron measurements: cross section and A_{FB} measurement [54].

Part II

EXPERIMENTAL SETUP

Description of the Large Hadron Collider. The ATLAS detector and its role. Individual parts of the ATLAS detector. Definition of objects detectable in the ATLAS. Reconstruction of collisions. Selection of events suitable for physics analysis.

4

THE LARGE HADRON COLLIDER

One of the most challenging physics projects humankind has ever undertaken is the construction of the Large Hadron Collider (LHC) [55]. Scientists dedicated many years of research to invent new technologies for building and operating the LHC successfully. The operation of the LHC in Run1 and Run2 proved the excellent technical state of the machine. The evidence of the Higgs boson in 2012 was the biggest satisfaction for all scientists working at CERN for their work spanning over many years. Nevertheless, there are still many challenges for the LHC waiting to be accomplished.

The LHC has been built by European Organization for Nuclear Research (CERN) in the 1990s using the same tunnel which was previously occupied by the Large Electron-Positron Collider. It is located around 100 m underground on the French-Swiss border near Geneva. The purpose of the LHC is to accelerate counter-rotating proton or heavy ion beams which are made to cross at 4 locations around accelerator ring. Hence, there are 4 biggest particle detectors situated at the LHC – ATLAS, ALICE, LHCb and CMS – surrounding each interaction point. The LHC was designed to reach the highest center-of-mass energy of 14 TeV for proton-proton collisions which has not been achieved yet.

This year, 2022, 3-year long shutdown dedicated to upgrades and repairs of the LHC has come to an end. The next period of data taking, Run3 [56], which has already started, is expected to be a transition phase between the LHC and the High Luminosity LHC (HL-LHC) [57]. The plan is to gradually achieve up to two times higher beam brightness comparing to Run2, meeting the beam parameter for the HL-LHC at the end of Run3. Beam energy for pp collisions is 6.8 TeV, hence rising center-of-mass energy $\sqrt{s}$ of collisions to 13.6 TeV in comparison with Run2, when $\sqrt{s}$ was 13 TeV. The expected integrated luminosity to be recorded by the ATLAS and CMS experiments during Run3 could be reaching 200-300 fb^{-1} depending on the beam parameters [56]. In comparison with Run2, it is almost factor 2 higher delivered integrated luminosity.

The installation of the HL-LHC will begin during next 3-year long shutdown, possibly from 2026-2029.

Integrated luminosity is a measure of collected data by the collider for specific process with the cross section σ over specific period of time: $L_{int} = \int \frac{dN}{dt} \frac{1}{\sigma} dt$, where $\frac{dN}{dt}$ is the number of events per second.

4.1 CERN'S ACCELERATOR COMPLEX

The acceleration process of protons up to an energy of 6.8 TeV per beam requires a chain of successive accelerators. The whole process starts with negative hydrogen ions which are at first accelerated by the linear accelerator. Previously, LINAC2 was used where the ions

gain energy of 50 MeV per particle but during the long shutdown it has been switched to LINAC4 due to the higher required luminosity. From 2020, LINAC4 prepares negative hydrogen ions for injection into the next accelerator, the Proton Synchrotron Booster, giving them energy up to 160 MeV. During the process of injection, ions are stripped of electrons leaving just proton beam for subsequent acceleration. The proton beam then proceeds to the chain of circular accelerators in order to boost its energy high enough for successful injection to the largest circular accelerator in the complex with a length of approximately 27 km – the LHC itself.

The succession of accelerators is depicted in Fig. 15. The first circular accelerator in the chain is the Proton Synchrotron Booster which boosts the proton beam up to 2 GeV. Then it is injected into the Proton Synchrotron (PS) which accelerates the beam up to 26 GeV. Afterwards, the beam continues to the Super Proton Synchrotron (SPS) – the last pre-accelerator before the LHC, obtaining energy of 450 GeV.

The usage of the PS is not only for accelerating protons but it serves as one accelerator in the chain of accelerators for heavy ions delivered from the Low Energy Ion Ring (LEIR).

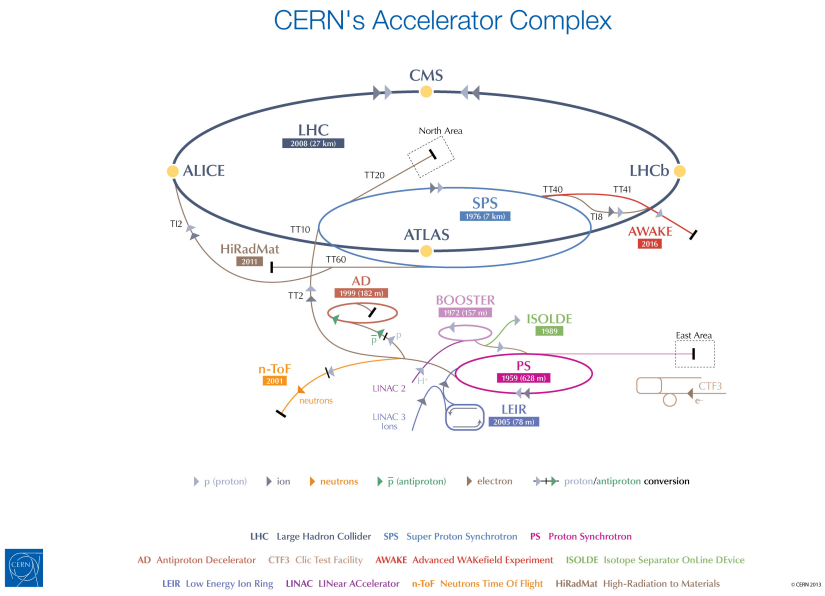


Figure 15: CERN's accelerator complex [58].

The beam circulating in the pipe is not continuous flow of particles. It consists of small (few centimeters long) separated chunks, which are called bunches. They are separated by 25 ns time intervals.

The LHC consists of two completely separate accelerating tubes where protons travel in opposite directions. The beam injection from the SPS to the LHC happens at two different locations in order to fill each proton ring with particle bunches going in the correct direction, see Fig. 16. The LHC was installed in the tunnel previously built for the LEP, therefore many adjustments to the design had to be made. As an example, due to limited space specific common magnet bending system for both rings providing an opposite magnetic flux circulation through two beam channels was adopted. Dipole magnets which are responsible for beam bending, have to be coupled mechani-

cally as well as magnetically because of very small separation of beam pipes. In addition to dipole magnets, there are sextupoles, octupoles or decapoles fixing small defects in dipole magnetic fields.

Except previously mentioned magnets, quadrupole magnets are installed along the beamline as well. Their function is to prevent the beam loss due to collision of particles with the accelerating tube. Four magnetic poles of a quadrupole magnet situated symmetrically around the beam pipe secure the beam squeeze in vertical or horizontal direction. Change of the direction is provided by the rotation of the quadrupole hence the alteration of quadrupole positions along the tube is needed. Another set of quadrupole magnets are used for additional beam size reduction of the order of magnitude before the beam enters the collision area in one of the detectors. Most of magnets used in the LHC are superconducting magnets which operate at a temperature below 2 K provided by cooling system with superfluid helium. Another necessity for successful LHC operation is vacuum system. It supplies insulation of magnets which are cooled down to low temperature below 2 K and also protects helium distribution from overheating. Specifically, ultra-high beam vacuum is maintained in the accelerating tube to reduce undesired beam-gas collisions.

In case of some failure or very low beam intensity caused by repeated proton-proton (pp) or ion-ion collisions, the beam dumping system is used. The beam is deflected from the LHC using extraction kicker magnets and pointed to an external absorber in sufficiently large distance from the accelerator to prevent overheating of the absorber material [55].

Schematic picture of the LHC is shown in Fig. 16. It comprises 8 arcs and 8 straight parts. Each straight section is dedicated to different purpose – 4 serves for experimental usage (beam crossings) and 4 for beam service. Two high luminosity insertions (the ATLAS and CMS experiments) are located on the opposite sides of the accelerating circle (in Octant 1 and Octant 5). Next two beam crossings are utilized by the ALICE and LHCb experiments. In these sections (Octant 2, Octant 8), the beam injection systems are installed. Both beams are entering the LHC reference plane vertically from below. Beam abortion system is housed in one of the remaining 4 straight sections, specifically in Octant 6. Each beam has its own dumping mechanism. In the section located in Octant 4, a 400 MHz superconducting cavity system is used for beam acceleration and last two sections provide beam collimation.

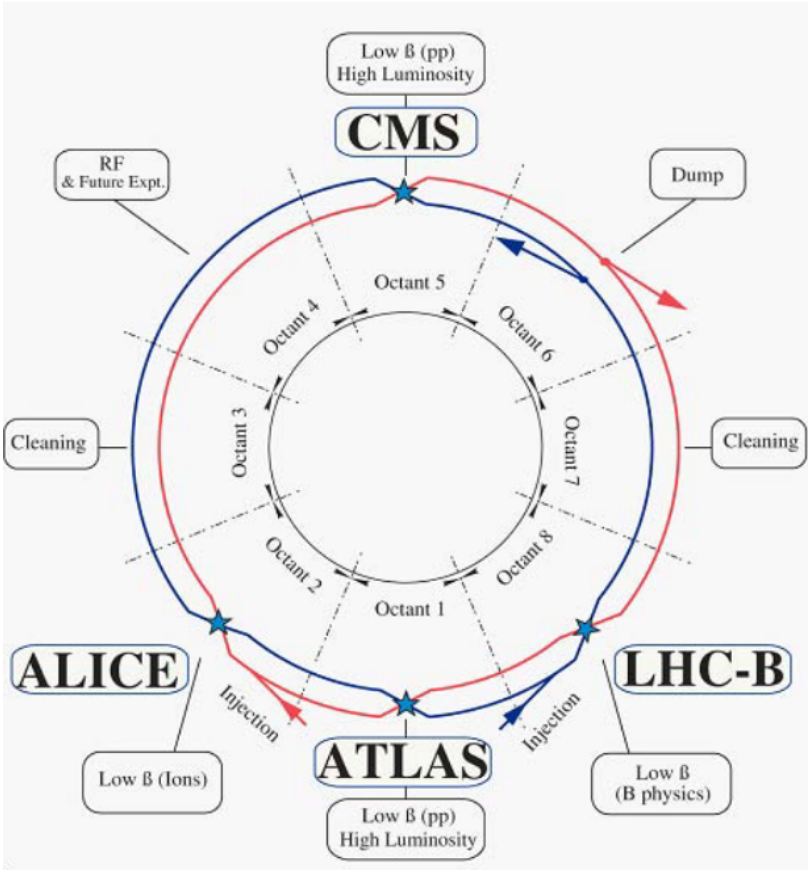


Figure 16: Scheme of the LHC [55].

4.2 THE ATLAS DETECTOR

4.2.1 *Overview*

The ATLAS detector [59] is a multi-purpose detector built for studying particles produced in proton-proton as well as ion-ion collisions. It is the biggest detector from all the detectors located at CERN. The name ATLAS stands for **A** **T**oroidal **L**HC **A**pparatu**S**. In height, it reaches 25 m and in length 44 m. Complete detector weight is 7000 tonnes. Technical realization of the ATLAS detector has to follow requirements given by proposed physics goals. Observation of new phenomena or precise measurement of significant parameters of the Standard Model is challenging due to the nature of proton-proton collisions. Numerous parton-parton collisions appearing during single pp collision can pollute the data and hide the signature of rare processes. In addition, not just two protons interact during the bunch crossing. Each bunch contains around 10^{11} protons to rise the probability of hard scatter interactions, so multiple proton-proton collisions happening during the single bunch crossing produce enormous number of particles effectively polluting the data even more. Hence,

studies of ATLAS data demand sufficient integrated luminosity and particle identification methods.

Overview of the ATLAS detector is shown in Fig. 17. It consists of many different detectors situated around the beam pipe in layers covering almost whole solid angle. Placement of all detection units is forward-backward symmetric with respect to the interaction point. The most important detecting subsystems are highlighted in Fig. 17 and will be described in more detail in following sections.

For description of the ATLAS detector and the events it records, a right-handed coordinate system is used. Its origin is at the nominal interaction point (IP) in the center of the detector and the z-axis is aligned with the beam pipe. The x-axis points from the IP to the center of the LHC ring, and the y-axis points upwards. Alternatively, polar angles can be utilized with the ϕ angle being the azimuthal angle in the plane transverse to the beamline and the θ measuring the angle between the particle trajectory and the beam pipe. Therefore, new coordinate system can be introduced – so called pseudorapidity–azimuthal angle space, where psedorapidity is defined as $\eta = -ln(tan(\frac{\theta}{2}))$. Metric in this space coordinates manifests itself by determining distances as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ [59].

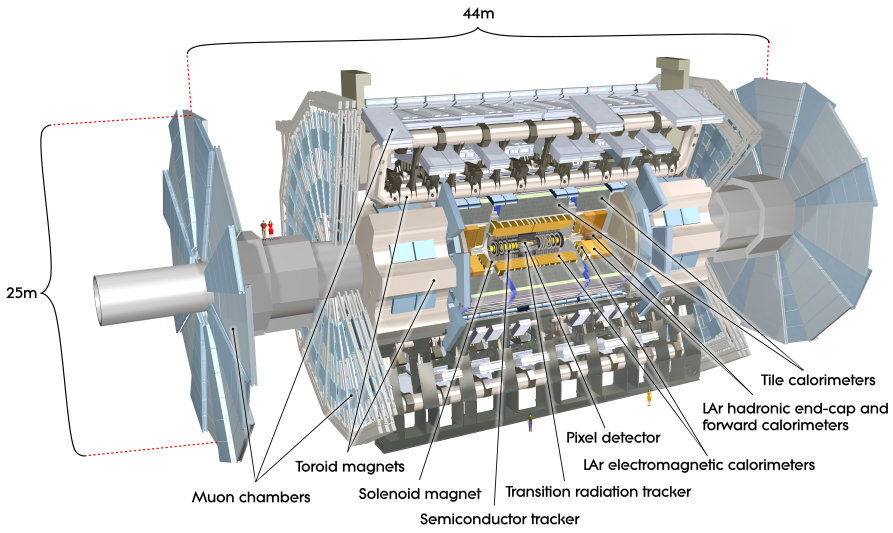


Figure 17: The complete layout of the ATLAS detector and its subdetectors [60].

4.2.2 *The Inner Detector*

The Inner Detector is used for detection of charged particles' tracks.

Particle tracking from the very first moment of a particle emergence from collision is extremely crucial. The detecting systems used for this purpose have to be very close to the interaction point and consequently radiation damage resistant. In the Inner Detector (ID), there

are two types of detectors linked together – the pixel and silicon microstrip trackers (SCT) with the Transition Radiation Tracker (TRT). This combination of detectors creates a sufficient system for momentum and vertex identification.

The most inner layers of the ID consist of the pixel and SCT-based tracking system. In close proximity of interaction point, it covers the acceleration tube in forms of concentric cylindrical planes. Apart from the barrel area, the tracking system in the end-cap regions is made of disks perpendicular to the tube. The most precise detecting system is pixel tracker whose parts are situated closest to the beam. All pixel sensors are arranged in a way that each track hits four pixel layers. Similarly, the SCT includes set of strip layers located around the beam tube in the same manner as pixel layers. Typically, a particle crosses eight strip layers. Both trackers are able to track particles in region of $|\eta| < 2.5$.

The last part of the ID is the Transition Radiation Tracker (TRT) implemented in a form of straw tubes with diameter of 4 mm orientated parallel to the beam pipe in the barrel region and radially in the end-cap regions similarly to other end-cap sub-detectors. Each straw is filled with mixture of xenon-based gases functioning as a detection medium. In the center of the straw, there is a golden wire which serves as an electrode. Number of hits registered by the TRT is approximately 36 per track which allows particle tracking up to $|\eta| = 2$. Additionally, detection of transition-radiation photons helps with electron identification. An illustration of charged particle passing though the ID is shown in Fig. 18.

The conjunction of the precise measurement of particle tracks at small radii near the beam tube with the TRT detecting system in bigger distances and at larger radii gives extremely precise measurements. The outer parts of the TRT are essential for momentum measurement. Although the TRT is not as precise as silicon-based detectors, it offers longer particle track measurement which balances out its imperfections. To provide the momentum measurement (details in Tab. 2), all sub-detectors are surrounded by the thin superconducting central solenoid which provides essential magnetic field of 2 T [59].

Table 2: Resolution and η coverage of individual parts of the ATLAS detector [59].

Detector component	Required resolution	η coverage	
		Measurement	Trigger
Tracking	$\frac{\sigma_{p_T}}{p_T} = 0.05\%\ p_T \oplus 1\%$	$-2.5 < \eta < 2.5$	
EM calorimetry	$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$	$-3.2 < \eta < 3.2$	$-2.5 < \eta < 2.5$
Hadronic calorimetry (jets)			
barrel and end-cap	$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$	$-3.2 < \eta < 3.2$	$-3.2 < \eta < 3.2$
forward	$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$	$3.1 < \|\eta\| < 4.9$	$3.1 < \|\eta\| < 4.9$
Muon spectrometer	$\frac{\sigma_{p_T}}{p_T} = 10\%$ at $p_T = 1$TeV	$-2.7 < \eta < 2.7$	$-2.4 < \eta < 2.4$

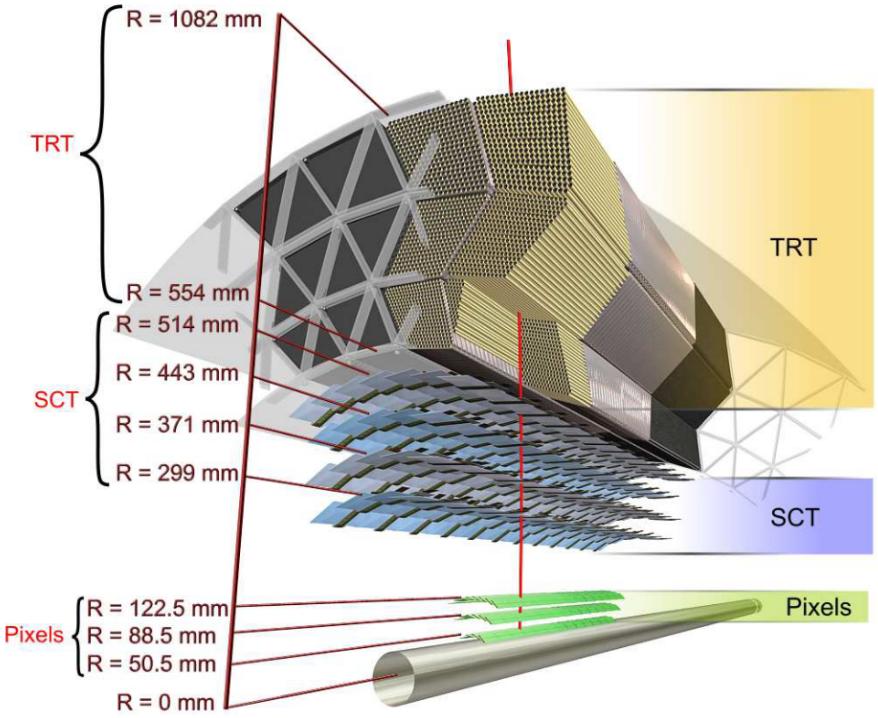


Figure 18: Different sensor layers in the barrel inner detector. Track of a charged particle with $\eta \approx 0$ traversing four pixel layers (the fourth one is missing in the figure, it is the innermost layer called Insertable B-layer (IBL) [61], that has been installed during shutdown in 2016 for more precise particle tracking), four silicon-microstrip sensor layers and approximately 36 straws of the TRT [59].

4.2.3 *Calorimetry*

The whole calorimeter-based detecting system surrounds the ID. There are two basic types of calorimeters distinguished by their purpose. The first one, located closer to the ID, is the electromagnetic calorimeter (EM). Its function is to measure energy of particles like photons and electrons which create electromagnetic showers inside the calorimeter. They are mostly stopped by the substance of the EM calorimeter. The second calorimeter, hadronic calorimeter, provides energy measurement for hadrons, which cannot be absorbed inside the EM calorimeter. It is positioned behind the EM calorimeter both in the barrel region and in the end-cap parts. The finer granularity of the EM calorimeter is necessary for good electron and photon identification and measurement of their properties mostly in η region matching the ID sensitive η region (see Tab. 2). The less granular structure of remaining parts is still adequate for jet reconstruction and E_T^{miss} measurements. The depth of each calorimeter is carefully determined because the calorimeter system must be thick enough to stop and ab-

Sometimes, high-energy electrons or photons penetrate through the whole EM calorimeter and they are stopped by the next detecting layers of hadronic calorimeter.

sorb electromagnetic as well as hadronic showers and minimize the level of punch-through into the outer layers of the ATLAS detector.

Both calorimeters use sampling technology, which means they consist of alternating active and passive layers. A particle passing through these layers loses its energy by creating showers in the passive medium. Created secondary particles are detected in the active medium - their energy is deposited here and converted to signal. Original particle transfers through several layers until it is absorbed and its energy is determined from the secondary-particle's energies.

The placement of the both calorimeter systems in the ATLAS detector is outlined in Fig. 19. Description of each system is given below.

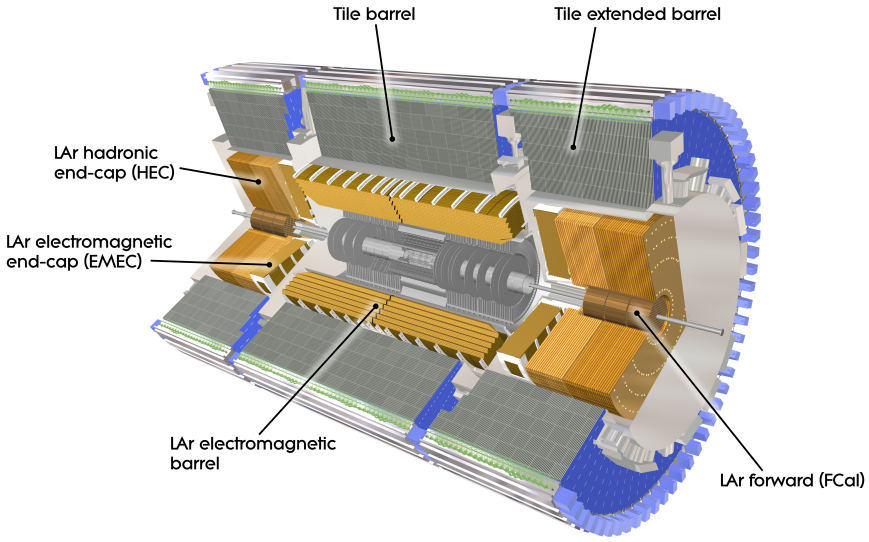


Figure 19: The layout of calorimeters in the ATLAS detector [62].

ELECTROMAGNETIC CALORIMETER The EM calorimeter consists of alternating accordion-shaped liquid argon (LAr) active layers and absorbing layers made of lead. The specific geometry of layers supplies full coverage in ϕ angle. The barrel part of the EM calorimeter is situated directly above the ID covering $|\eta| < 1.375$ range with small 4 mm gap at $z = 0$. Two end-cap LAr calorimeters are divided into two concentric wheels sensitive to $1.375 < |\eta| < 3.2$ region, where inner part covers $1.375 < |\eta| < 2.5$ range and the outer wheel measures particles emerging from IP with $2.5 < |\eta| < 3.2$.

HADRONIC CALORIMETER The layout of hadronic calorimeter is more complicated. Different pasive and active mediums are used in the individual parts of the hadronic calorimeter, depending on their position. The choice of passive and active medium is different for the Tile calorimeter. The absorber medium is created by the steel layers and the function of active material is provided by the scintillating

tiles. In the central barrel region, the Tile calorimeter is set behind the LAr electromagnetic calorimeter and covers $|\eta| < 1.0$ region. In the end cap regions, there are two extended barrels on each side of the central barrel covering $0.8 < |\eta| < 1.7$, using the same materials as in the central barrel. Both the Tile central barrel and extended barrels are divided into 64 modules.

In the end-cap regions, behind the LAr electromagnetic calorimeter in the direction of z-axis, the Hadronic End-cap Calorimeter is located, using the LAr technology and sharing the same cryostats. It consists of wheel-shaped modules using copper layers as passive medium. The copper layers closer to the interaction point are thinner than those in the outer layers of calorimeter. The η coverage of this part of the calorimeter system overlaps the Tile calorimeter η range ($|\eta| > 1.5$) and also the range of the forward calorimeter ($|\eta| < 3.2$).

Scanning the calorimeter system in the z-axis direction, there is another detector in the most forward region – the Forward Calorimeter (FCal), which comprises three parts. The first module is mainly for electromagnetic shower detection using copper as absorbing material and the other two parts made from tungsten are aimed for containment of hadronic showers. Tungsten layers are used as a passive medium in two parts, while liquid argon is used as an active medium in all three parts of the FCal. The FCal shares the same cryostats with the LAr calorimeter. All mentioned calorimeters in the end-cap regions are covered by the Tile calorimeter extended barrels [59]. Resolution of each part of the ATLAS calorimetry is given in Tab. 2 together with their η coverage.

4.2.4 *Muon spectrometer*

Almost all the particles arising from collisions are stopped and absorbed by calorimeter system. Detectable particles which can escape through the calorimeters are mostly muons. Therefore, there has to be another type of detecting system for measuring muon tracks.

The muon spectrometer surrounds all other sub-detectors of ATLAS creating the most outer detection layer. Its primary role is to measure muon properties by bending their tracks in magnetic field using different types of muon detecting chambers and superconducting magnets. In the barrel region, sufficient magnetic field is provided by large air-core toroid magnet consisting of eight coils arranged symmetrically around beam pipe, whereas in the end-cap regions, there are end-cap toroid magnets inserted into both ends of the barrel toroid. In the barrel area, muon-tracking chambers are located inside and on the top of the eight coils of the barrel toroid magnet, whereas in the end-cap regions, they are situated in front of and behind the end-cap magnets, see Fig. 20.

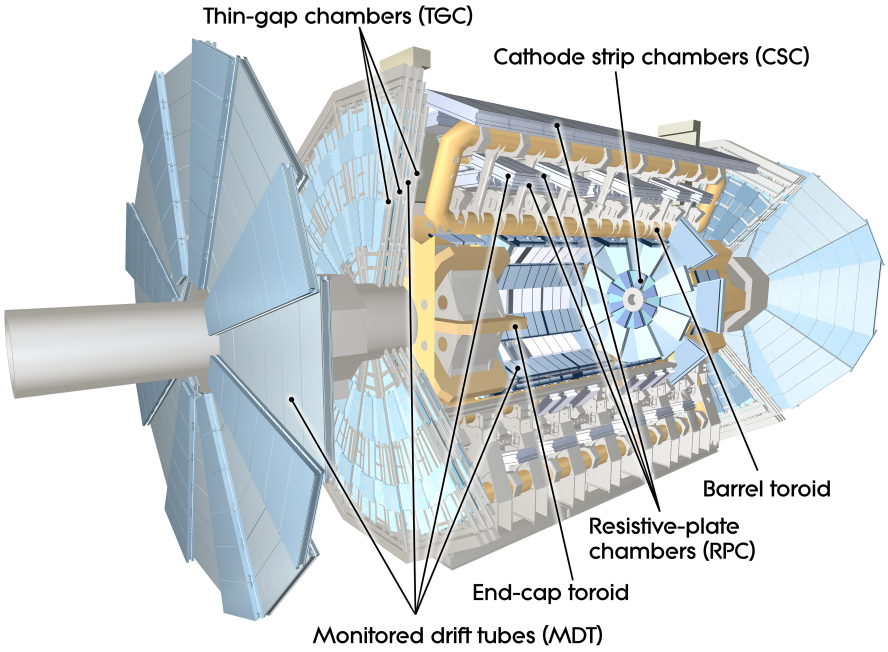


Figure 20: The layout of muon detecting units in the ATLAS detector [63].

Muon chambers are oriented parallel to the beam axis creating three cylindrical layers in the barrel region. In the end-cap regions, the position of muon chambers is perpendicular to the beam pipe creating large disks.

For precise measurement of the muon track coordinates, the Monitored Drift Tubes (MDT) are installed into the muon chambers. MDT chambers are covering range of $|\eta| < 2.7$, except the most inner end-cap layer, where the range is shortened to $|\eta| < 2$. For track measurement, the second coordinate ϕ orthogonal to those measured by MDT (which determines just the η parameter of the track) is determined from the hits recorded in trigger chambers.

The innermost layer in the forward region corresponding to pseudorapidity values $2.0 < |\eta| < 2.7$ is equipped with the Cathode Strip Chambers (CSC). The CSC is a multi-wire proportional chamber with cathodes in form of strips which are placed in orthogonal directions. This arrangement allows measurement in the transverse plane (ϕ coordinate) and in the bending plane (η coordinate) at the same time.

Significant part of muon detecting system is ability to trigger on muon tracks. Hence, aside muon chambers, trigger chambers are installed providing information about a muon track in time interval of $15 - 25$ ns after the particle transition, allowing to match correct bunch-crossing. In the barrel region, Resistive Plate Chambers maintain the triggering for $|\eta| < 1.05$ and in the end-cap regions there are Thin Gap Chambers covering pseudorapidity range $1.05 < |\eta| < 2.4$. Both trigger chamber types are able to measure both coordinates, ϕ

in transverse plane and η in the bending plane [59]. Resolution of the muon system is given in Tab. 2.

4.2.5 *Trigger system*

A necessary part of the data taking in the ATLAS detector is sufficient trigger system. The function of the trigger is to decide which collision data are recorded for further analysis and which are omitted due to limited capacity of the storage system. From the designed bunch-crossing rate of 40 MHz, the trigger system decreases the rate of events to just few hundreds Hz. Two types of trigger systems are used for this purpose: a hardware-based Level-1 [64] trigger and afterwards software-based high level trigger (HLT) [65].

The Level-1 trigger's function is to find important regions (Region-of-Interest (RoIs)) in the detector using information from calorimeter granularity and muon trigger system units. This first step reduces the event rate to approximately 100 kHz within decision time of 25 μs. The information about RoIs is sent to the HLT. Hence, as the second step, sophisticated algorithms implemented in the HLT are executed to again cut off insignificant data using full granularity detector information. The procedure uses either the RoI or the whole event. The HLT reduces the rate from 100 kHz to 1 kHz which takes approximately 200 ms.

The Level-1 trigger collects information from the Level-1 Calorimeter trigger and the Level-1 Muon trigger together with the L1Topo trigger. The information is processed with central trigger processor, which makes the Level-1 trigger decision.

The Level-1 Calorimeter trigger searches for electrons, jets, photons, taus, missing transverse energy (E_T^{miss}) using information from electromagnetic and hadronic calorimeters. At this level, dynamic bunch-by-bunch pedestal corrections are installed for pile-up suppression. The Level-1 Muon trigger provides fast information from muon detectors. The L1Topo trigger takes as an input information like angular separation of objects or sum of p_T of all Level-1 jet objects from both the calorimeters and muon chambers, combining them and using them to make topological selections. Background for many trigger selections is suppressed by usage of the L1Topo trigger by around a factor of 2 [66].

Apart from the HLT and the Level-1 triggers, there is another hardware-based system helping with triggering. The information from silicon tracking detectors after each Level-1 trigger, is gathered and send to the Fast TracKer [67], where the particle tracks from the whole event are reconstructed. This information is then used by the HLT to improve e.g. b-jet triggering or other selections where the full-event tracking information is needed.

5 DATA PROCESSING IN THE ATLAS DETECTOR

Translation between electronic signals from various ATLAS sub-detectors and the actual particles which left the traces, can be quite difficult process. To some extent, this process is arbitrary. It is dependent on how we define each particle/object in terms of tracks, energy depositions in calorimeter clusters or hits in muon chambers. Yet of course, each reconstructed particle still has to have all its attributes given by its nature.

Hard-scatter collision of two protons proceeds through interaction of their partons. First particles emerging from interaction are usually some quarks or gluons. However, these particles cannot exist as free particles, because they carry color charge. Hence, due to the QCD confinement, only colorless objects can be detected. Each quark becomes part of a hadron, while another quarks/gluons arise along the way, creating another hadrons. So instead of one quark/gluon in final-state, there is a whole bunch of hadrons going in approximately same direction forming narrow cone called jet.

EVENT EVOLUTION The stage of an event before hadronization or decay takes place is called *parton-level*. Parton-level particles are mainly quarks or gluons. For $t\bar{t}$ production events, it is of course $t\bar{t}$ pair. An event with already stable particles in the final-state either after hadronization, decay or both, has reached *particle-level* final-state. Both these stages, parton and particle-level can be reached only through simulation of collisions. On the other hand, in the measurement, only *detector-level* objects are accessible - particles and physics objects, which are reconstructed from traces left in the detector. This stage is also called *reconstructed-level*.

Stable particles = leptons and colorless bound states - hadrons and mesons, after decay if possible.

An event is claimed to be a $t\bar{t}$ event, if decay products of $t\bar{t}$ pair are reconstructed. For single-lepton (ℓ + jets) decay channel (see Sec. 3.1.2), one lepton, missing transverse energy E_T^{miss} and at least four jets are required. Reconstructed E_T^{miss} is a sign of neutrino being created in an event. In case of ℓ + jets decay, a neutrino arises from the leptonic decay of the W boson. Required four jets are initiated by b and $\bar{b}$ quark, which are direct decay products of t and $\bar{t}$ quark and another two quarks, which come from the hadronic decay of the W boson. Considered lepton is a muon or an electron, alternatively tau lepton. Specifically, only leptonically decaying tau lepton, because tau is detected solely by its decay products, so in the end a muon or an electron is detected. A tau lepton decays also into hadrons, like kaons or pions. However, the reconstruction of hadronically decaying tau lep-

ton is complicated and therefore not considered here. An illustration of the single-lepton $t\bar{t}$ decay can be found in Fig. 21. Dilepton $t\bar{t}$ decay is identified if an event possesses two leptons and two reconstructed neutrinos from the decay of W bosons, together with jets initiated by b and $\bar{b}$ quark.

The last $t\bar{t}$ decay channel, all-hadronic channel, is not considered for presented study. The identification and reconstruction of this $t\bar{t}$ decay channel is difficult due to complicated final state - there are only jets, no leptons. Therefore, identification of jets corresponding to t and $\bar{t}$ quark cannot be done by measuring charge of the lepton. Hence, reliable methods for assignment of jets to partons have to be developed. The advantage of the all-hadronic channel is that the reconstruction of E_T^{miss} is not necessary.

To sum up, the following objects have to be defined for $t\bar{t}$ reconstruction: an electron, a muon, jets (like b-jet, large-radius (large-R) jet) and missing transverse momentum.

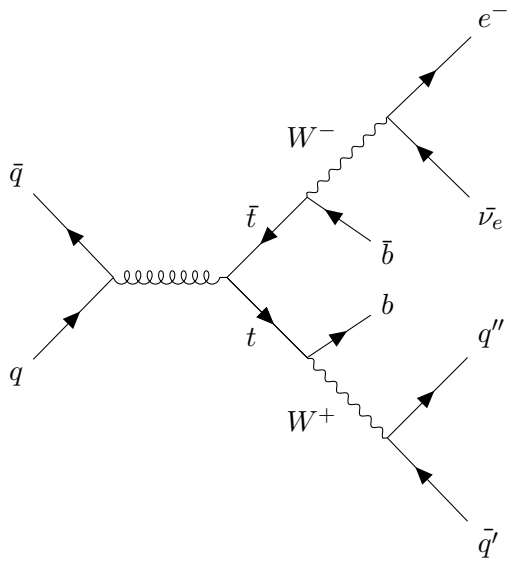


Figure 21: Example of single-lepton decay of $t\bar{t}$ pair produced by quark-antiquark annihilation

5.1 DEFINITIONS OF OBJECTS

Properties of all physics objects are determined with respect to the primary vertex. Events are required to contain at least one vertex satisfying the following criteria:

- At least two tracks per vertex are required to suppress non-collision events.
- Only the tracks with $p_T^{\text{track}} > 0.5$ GeV are considered.

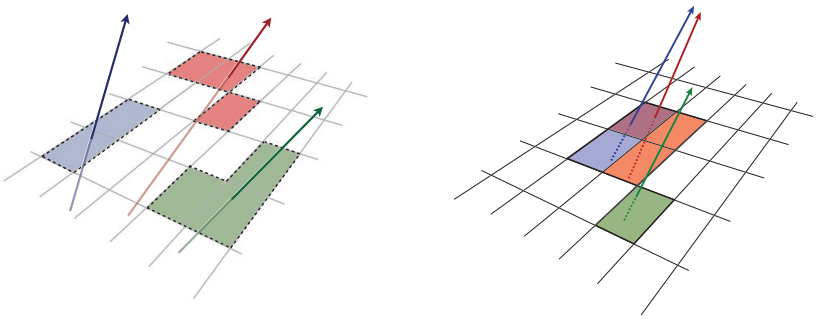
- A vertex with the largest sum of track transverse momenta $\sum p_{T,\text{track}}^2$ is tagged as primary vertex.

5.1.1 *Tracks*

Tracks of charged particles are reconstructed using information from the pixel and SCT sub-detectors. As a particle traverses the layers of the ID, it experiences loss of energy and generates charge in detecting units, signalizing the direction of its flight. Pixels and strips, the basic units of the pixel and SCT sub-detectors, with the charge deposition larger than threshold value sharing common edge or corner, are combined together to create clusters. An intersection point of a particle path with the detecting layers, *space-point*, is represented by a three-dimensional measurement built using these clusters. In the pixel, one cluster is enough to be claimed as a space-point, but in the SCT, clusters from both sides of the strip are needed to create a space-point.

A seed for a track is assembled from 3 space-points. Seeds chosen for a track are used in the Kalman filter [68], which constructs a track candidate by trying different combination of chosen seeds. Additionally, it adds another space-points from other layers of the pixel/SCT, which lie in the preliminary path of a particle. This prevents further investigation of meaningless combinations of space-points, reducing computational time. However, after this step, many track candidates share the same space-points or could be wrongly constructed, which calls for the resolve of the ambiguity.

NOTE: Clusters can be of two types: *single-particle* and *merged clusters*, see Fig. 22. This categorization reflects the way the charge deposit in a cluster has been made: either by one particle or by many of them.



(a) single-particle clusters (b) merged cluster

Figure 22: Schemes illustrating clusters induced by deposits of one particle (a) or more particles (b). Path of each particle and its assigned cluster are drawn in different colors [69].

Expected number of clusters differs depending on the sub-detector as well as the resolution of the sub-detector. Both are considered in the track score. Hole is an intersection of a reconstructed particle trajectory with an active detecting layer but without a corresponding cluster in it.

Moreover, clusters can be shared by at most two track candidates. These *shared clusters* are those used in reconstruction of multiple tracks, which do not meet the criteria to be identified as merged clusters. Hence, the probability of wrong assignment is higher and number of shared clusters per track candidate is limited.

A track score is assigned to each track candidate characterizing its quality. Properties like logarithm of the track p_T (low p_T tracks are more probably incorrectly constructed), number of holes in the track, the goodness of the track fit (χ^2) and sub-detector-based characteristics of a track are considered in evaluation of a track score.

Any track candidate has to fulfill these conditions:

- $p_T > 400$ MeV,
- $|\eta| < 2.5$,
- at least 7 clusters in the pixel and the SCT detector,
- no more than 1 shared pixel cluster or 2 shared SCT clusters on the same layer,
- maximum 2 holes in the combined pixel and SCT detectors, but no more than 1 hole in the pixel detector alone,
- $|d_0| < 2$ mm,
- $|z_0 \sin(\theta)| < 3$ mm,

where d_0 is the transverse impact parameter and z_0 is the longitudinal impact parameter (see Fig. 23) - both measures distances of the point of the closest approach of a track transverse to the primary vertex. The θ is polar angle of a track-candidate momentum [69].

5.1.2 *Electron*

RECONSTRUCTION: For electron reconstruction, the central region of the ATLAS detector corresponding to $|\eta| < 2.47$ is considered. An electron candidate is built from following requirements: a track in the ID and localized energy clusters in the EM calorimeter have to be reconstructed fulfilling track-cluster matching conditions in $\eta \times \phi$ space.

As the first step of electron reconstruction, energy deposits in the EM calorimeter are examined. The EM calorimeter is split into smaller blocks called towers in $\eta \times \phi$ space with size of $\eta^{tower} \times \phi^{\text{tower}} = 0.025 \times 0.025$. Energy corresponding to each tower is evaluated as a sum of energy deposits in all layers of the EM calorimeter for that specific region in η, ϕ coordinates. The whole EM calorimeter is then scanned looking for seed cluster with highest E_{T} value (exceeding

Transverse energy $E_T = E \sin\theta$

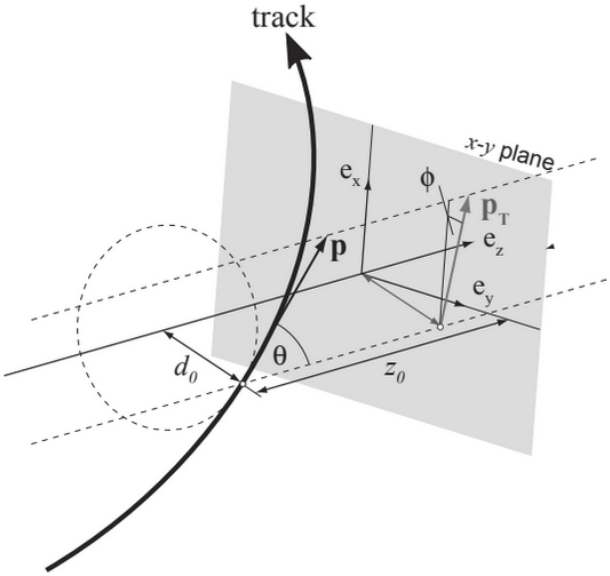


Figure 23: Illustration of transverse and longitudinal impact parameters of a track, d_0 and z_0 [70].

2.5 GeV) constructed from an area corresponding to 3×5 towers. Reconstruction efficiency of the algorithm used for seed cluster reconstruction is dependent on $|\eta|$ and E_{T}.

The second step in finding an electron candidate is the track reconstruction procedure. Track seeds are formed from hits in silicon tracking detectors. Each has to have at least 3 space points constructed from clusters of hits in the pixel and SCT detector.

An electron candidate is reconstructed by matching the track candidate to the calorimeter seed cluster. If more than one track is found within required distance in $\eta \times \phi$ space, the primary electron track is identified taking into consideration also number of hits in silicon tracking detectors, specifically hits in the innermost layer. Furthermore, seed cluster in the EM calorimeter needs to be extended in both coordinates (η, ϕ) to reconstruct the cluster of electron candidate. Energy of the electron candidate is determined from calibrated energy of this extended cluster, while η and ϕ coordinates of the matched track are assigned to this electron candidate defining flight path of this electron. An illustration of an electron traversing various layers of the ID and the EM calorimeter is presented in Fig. 24.

Specifically, difference between ϕ coordinate for cluster and track multiplied by the negative sign of particle charge, $-q \times (\phi_{cluster} - \phi_{track})$ has to belong to the interval $[-0.20, 0.05]$. The separation in η is specified as $|\eta_{cluster} - \eta_{track}|$ and it should be below 0.05.

IDENTIFICATION Prompt electrons need to be distinguished from non-prompt electrons emerging in decays of heavy-flavored hadrons or electrons coming from photon conversion in detector body. Additionally, light-flavored jets could also be misidentified as electrons. Identification process is based on constructing a likelihood function for both signal L_S and background events L_B. Each likelihood func-

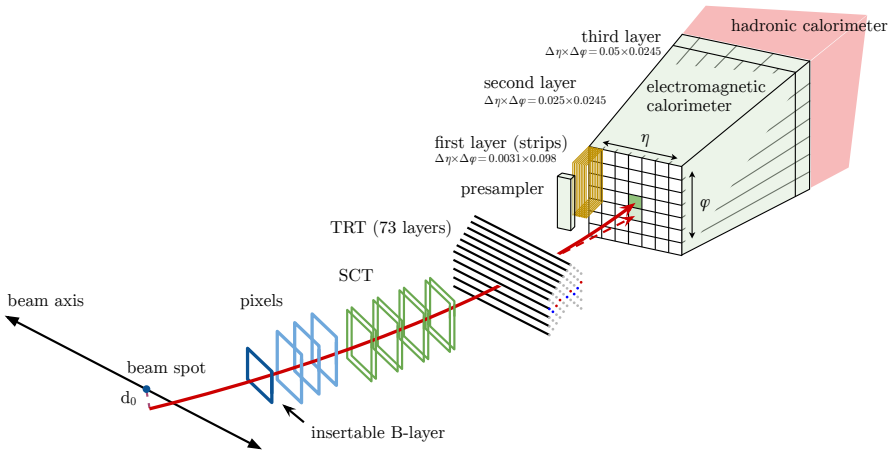


Figure 24: An electron path passing through individual layers of tracking detectors (pixel, SCT, TRT), finally penetrating the EM calorimeter. The EM calorimeter is divided into three layers. The first layer serves for differentiation of photons and pions ($\pi^0 \to \gamma\gamma$). Most of the energy of electrons/photons is deposited in the second layer. The third layer measures the energy of the back part of the shower [71].

tion is a product of probability distribution functions for specific input quantities [71], which collect information from tracking and calorimeter system, like the ones depicted in Fig. 24. For each electron candidate, a discriminant is computed using likelihood functions:

$$d_L \sim -\ln\left(\frac{L_S + L_B}{L_S} - 1\right). \tag{22}$$

Candidates with discriminant values larger than threshold are considered signal. The threshold value of discriminant is chosen according to the operating point required by each physics analysis: *Loose*, *Medium* or *Tight*. Each of them provides different signal efficiencies and background rejection rates, e.g. *Loose* corresponds to the lowest threshold value hence gives highest efficiency of 93%, while *Tight* operating point with the most stringent threshold gives lowest efficiency of 80% for electron transverse energy $E_T = 40$ GeV [71]. Likelihood-based electron identification is applied also during the online selection of events, particularly it is a part of the HLT. The algorithm is as similar as possible to the one used in the offline selection. Quantities which require too much computation time are not used. Also generally, quantities reconstructed in the trigger are not as precise as when they are reconstructed offline.

Signal efficiency is defined as a ratio of signal events selected by the procedure and total number of signal events. A fraction of background events, which are excluded, computed with respect to all background events are called background rejection efficiency.

ISOLATION Another criterion for successful differentiation of prompt electrons and non-prompt electrons or hadrons mimicking electron behavior, is based on mapping the activity in the calorimeter and in the inner detector in close proximity of the candidate electron. A presumption is that there should be little activity in an area around the

candidate electron in $\phi \times \eta$ space. Quantification of this activity is done via calorimeter-based or track-based isolation variables.

Calorimeter-based isolation is established using sum of transverse energies of topological clusters [72] within a specific cone with ΔR radius (typically, $\Delta R = 0.4$) around the electron candidate. A topological cluster is constructed using cells with electromagnetic activity four times higher than the level of expected electronic noise and pile-up as a starting point. Surrounding calorimeter cells are inspected and if a deposited energy is two times higher than the noise threshold for those particular cells, they are considered a part of this topological cluster. Naturally, the energy of an electron candidate itself should be subtracted from the transverse energy of topological cluster by removal of $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ rectangular area around the direction of the electron candidate. This procedure does not take into account potential leakage of the candidate energy outside the predefined area. Hence, particular corrections to the transverse energy of the topological cluster have to be made. Additionally, subtraction of pile-up energy deposited in the discussed cone should be considered as well.

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

Pile-up = multiple pp collisions

Isolation criterion based on the information from tracking system uses reconstructed tracks with $p_T > 1$ GeV and within region of $|\eta| < 2.5$ of the inner detector. The corresponding isolation variable is built in similar manner as in the aforementioned case; the sum of the transverse momenta of tracks located inside the cone assigned to the electron candidate with the exception of electron track itself is taken. The size of the cone can be adapted according to the p_T of the track allowing smaller cone sizes than the ones considered for the calorimeter-based isolation variable:

$$\Delta R = min\left(\frac{10\ \text{GeV}}{p_T[\ \text{GeV}\ ]}, R_{\text{max}}\right). \tag{23}$$

The R_{max} parameter is typically set to 0.2 [73].

Both calorimeter-based and track-based isolation variables can be used for definition of specific operating points. There are three categories of the operating points:

- *Loose isolation* - fixed isolation efficiency required, no dependence on the E_T or η of the candidate electron,
- *Gradient isolation* - fixed isolation efficiency dependent on the E_T required, but independent from η,
- *Fix isolation* - no fixed efficiency, but fixed requirements on isolation variables.

A description of evaluation of electron reconstruction efficiency is given in Sec. 7.1.

5.1.3 *Muon*

RECONSTRUCTION Reconstruction of muons in the ID is similar to a reconstruction of any charged particle [74]. Apart from other charged particles, muons penetrate the whole detector leaving traces in the outermost layer of the ATLAS, the Muon Spectrometer (MS). Muons can be reconstructed using information from the ID and the MS either separately or combined. Muon candidates in the MS are built from reconstructed segments in the MDT chambers constructed from detected hits. A muon track is seeded with a segment found in the middle layers because of higher number of trigger hits. Also, it is fitted with another segments from the inner and outer layers. To build a track, at least two segments have to match when pairing segments from the whole MS, except the barrel-endcap transition region where just one high-quality segment is enough for track reconstruction.

Based on how the information from the ID and the MS are utilized when reconstructing a muon, four muon types are defined:

- *Combined muons:* Potential track is reconstructed separately in the ID and the MS and a combined track is made by fitting the ID and MS track. In most cases, a MS track is constructed first and then it is matched to an ID track.
- *Segment-tagged muons:* An ID track is extrapolated to the MS, where at least one local segment in the MDT or CSC has to be identified as a track-related. This reconstruction is used when a muon has low p_T and crosses just one layer of the MS.
- *Calorimeter-tagged muons:* An ID track is considered as a muon candidate if energy deposit corresponding to the minimum-ionizing particle is found in the calorimeter. This type of muons can be used even if a muon enters the detector region with poor coverage.

Not whole spatial angle is covered by detecting units, there are gaps for cabling and maintenance.

- *Extrapolated muons:* Only the track reconstructed using information from the MS together with loose compatibility of a muon with this track to originate from the interaction point is used. Muons found in the region $2.5 < |\eta| < 2.7$ (not covered by the ID) can be reconstructed by this procedure.

If more than one type of muon can be assigned to an ID track, then the preferred option is to use the *Combined muons*, as a second choice *Segment-tagged muons* and lastly *Calorimeter-tagged muons*. Conflict with *Extrapolated muons* is resolved by further study of track hits in the detector. A track with the best fit and largest number of hits is selected.

IDENTIFICATION Identification procedure is needed for differentiation of prompt muons from background muons, i.e. from decays

of light hadrons like kaons or pions. Specific selection criteria on variables dependent on the transverse momentum and charge measurements in the ID and MS are set to ensure proper identification, together with requirements on the number of hits in the ID and MS for a robust momentum measurement.

In the ID, there should be at least one pixel hit, five hits in the SCT for a candidate combined muon. Also, 10% of TRT hits should be included in the fit of a combined muon track.

According to different requirements applied, there are 4 identification selections for muons:

- *Loose muons* - used specifically for maximization of reconstruction efficiency while still maintaining good-quality tracks; all muon reconstruction types are used,
- *Medium muons* - default selection; only *Combined* and *Extrapolated* muons are used,
- *Tight muons* - suitable when high purity needed, which is accompanied by the reduction of reconstruction efficiency; only *Combined* muons are used,
- *High-p_T muons* - for muons with $p_T > 100$ GeV, should provide good momentum resolution.

First three categories are inclusive, which means that muons which satisfy Tight selection also pass Medium and Loose selection criteria. Applying isolation requirements reduces misidentification by more than an order of magnitude.

An estimation of the efficiency of reconstruction is described in Sec. 7.1.

ISOLATION Similarly to electrons, two isolation variables are used: a track-based and a calorimeter-based variable. Both variables measure activity in the vicinity of a muon candidate with the exclusion of a muon itself. By applying specific requirements, the amount of non-prompt muons should be reduced.

Definition of the track-based variable is the same as for electrons - the sum of the transverse momenta of the tracks with $p_T > 1$ GeV in a cone ΔR, which is p_T dependent and defined according to Eq. 23, with $R_{\text{max}} = 0.3$.

Analogously to electrons, the sum of the transverse energy deposits of topological clusters in a cone with $\Delta R = 0.2$ is called calorimeter-based isolation variable.

Placing different requirements on these isolation variables and/or efficiency of isolation results in 7 isolation working points, designed for various physics analyses: *Loose*, *LooseTrackOnly*, *Tight*, *Gradient*, *GradientLoose*, *FixedCutTightTrackOnly*, *FixedCutLoose* [74].

5.1.4 *Jet*

DEFINITION: A gluon or quark (parton) produced in a proton-proton interaction generates a spray of secondary particles creating a collimated cone around the direction of its mother particle. This so called jet needs to be sufficiently reconstructed in order to recover the original properties of a parton.

RECONSTRUCTION: Generally, jet reconstruction employs the anti-k_t algorithm [75] using specific radius parameter R of the jet cone. The parameter R is defined as a distance in $\eta - \phi$ space, from calibrated topological calorimeter clusters [76]. The algorithm belongs to the group of sequential recombination algorithms which form a jet by grouping the clusters. It starts from the cluster of the hardest particle trying to find clusters for softer ones belonging to the particle according to the specific distance definitions:

A soft particle has lower energy and presumably smaller values of the transverse momentum than particles emerging directly from hard-scattering, "hard" particles.

$$d_{ij} = min(k_{Ti}^{-2}, k_{Tj}^{-2})\frac{\Delta_{ij}^2}{R^2}, \qquad d_{i,B} = k_{T,i}^{-2}, \tag{24}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{Ti} is a transverse momentum of a particle i, y_i and ϕ_i are the rapidity and azimuthal angle of flight direction of the particle i, R specifies radius of the cone in units of angular separation of clusters i and j Δ_{ij}. The d_{ij} evaluates distance between two particles/potential jets i and j, whereas $d_{i,B}$ defines the distance between particle i and the beam (B). The procedure is dependent on comparison of these two distances, d_{ij} and $d_{i,B}$. If $d_{i,j} > d_{i,B}$, clusters i and j are recombined. Otherwise, cluster i is considered a jet and omitted from the algorithm. Both distances are re-calculated for remaining clusters again and the algorithm continues until no particle is left.

If more than one hard particle is found in the close vicinity of the first one ($\Delta_{12} < R$), the resulting shape of a jet can be more complex than just a simple cone (in case of transverse momenta of both particles are similar $k_{T1} \sim k_{T2}$). If $k_{T1} > k_{T2}$, corresponding jet is formed around the first particle. It is always the hard particle which dictates the properties of a jet like direction and shape, while all surrounding soft particles in R radius are creating a conical structure with the hard particle in the center.

When an event is boosted, its longitudinal as well as its transverse coordinates are changed due to this transformation. But the distance ΔR in $\eta - \phi$ space remains the same.

Distances in η, ϕ coordinates are invariant with respect to the longitudinal boosts hence the radius parameter remains unaffected while considering high p_T or low p_T jets. In the analysis presented here, the R parameter is set to 0.4 (small-R). If not stated otherwise, a jet means small-R jet. Detector-level/calorimeter jets, created from the reconstructed EM topological clusters, need to undergo few corrections steps. Firstly, a direction of a jet is corrected to point to the primary vertex. Jet energy and another jet-describing variables need

to be corrected as well. Pile-up contribution is removed from a jet using jet area method [77]. The estimation of pile-up inside a jet cone is done by adding "ghost" particles, (infinitely soft particles) in the jet cone and its vicinity. There are two methods:

1. Adding one ghost particle and studying the dependence of the distance ($\eta - \phi$ or $y - \phi$ space) between ghost particle and hard particle (which defines a jet) on its clustering with the jet. By testing the possible distances the so called *passive area* of a jet is found. This procedure should test the perception of a jet to point-like contributions from underlying events or pile-up.

2. Adding huge density of ghost particles and testing which ones are clustered within a jet. Ghost particles can cluster also among themselves, participating more actively in jet clustering. Number of ghosts clustered inside a jet is a measure of the jet *active* area. Described method should help to evaluate the effect of high density soft radiation background.

Pile-up can be of two types: out-of-time and in-time. In-time pile-up events correspond to additional pp collisions in a single bunch crossing, whereas out-of-time pile-up events comes from other bunch crossing.

The jet energy is corrected using calibration based on MC-data comparison with additional correction due to pile-up effects [78, 79]. Proper jet selection in an area defined by $|\eta| < 2.4$ and $p_T < 60$ GeV is provided by using a Jet Vertex Tagger (JVT) [80]. This procedure helps to differentiate jets originating from primary vertex from those arising in pile-up collisions. The information about the tracks matched to a jet is utilized for the procedure. The JVT method uses two variables R_{pT} and $corrJVF$ to evaluate a probability of a jet being a signal jet. The $corrJVF$ variable estimates a fraction of p_T tracks associated to a jet that can be assigned to the primary vertex. The R_{pT} is a ratio of scalar sum of p_T of all tracks coming from the primary vertex and calibrated p_T of a jet. Both variables have different distribution for hard-scatter jets and pile-up jets, hence providing good separation power. Efficiency of the JVT method was measured by tag-and-probe method for events in $Z \to \mu\mu$+jets sample. A leading jet recoiling against the Z boson is considered a probe. Resulting efficiency is dependent on p_T of the Z boson and JVT cut [80].

Pile-up jets can be of two origins: QCD jets originating in a pile-up vertex and stochastic jets, which are basically random combination of particles arising from multiple pile-up vertices.

B-TAGGING For $t\bar{t}$ reconstruction, it is inevitable to reconstruct and identify a jet originating from a b-quark. Analysis of attributes specific for a jet initiated by a b-quark helps in the process of identification.

Due to quite long lifetime of b hadrons, their decay is identified through the existence of a secondary vertex in the proper distance from the primary vertex corresponding to the b-hadron lifetime. By measuring impact parameters of tracks inside a jet, tracks of remnants from a b hadron decay can be identified. Generally, impact parameters for tracks corresponding to a b-hadron decay have larger values. One of the b-tagging algorithms which profits mainly from the usage of the impact parameters is IP3D [81]. It defines a jet weight by calculat-

For a b hadron with $p_T = 50$ GeV its flight path is ~ 3 mm

ing the sum of the logarithms of each track weight. Weight of a track is a ratio of probabilities for b-jet and light-jet hypothesis obtained by comparing measured impact parameter significance ($d_0/\sigma_{d_0}, z_0/\sigma_{z_0}$) with simulated values.

On the other hand, Secondary Vertex (SV) algorithm specializes in explicit reconstruction of an inclusive three-dimensional secondary vertex. The SV1 algorithm [81] is based on the calculation of the log-likelihood ratio (as IP3D) using different variables characterizing a secondary vertex: energies of tracks which forms the SV, specifically, ratio of their sum and sum of energies of all tracks associated with a jet, invariant mass of these tracks under an assumption that all tracks belong to pions, number of vertices with two tracks, and ΔR distance of a jet direction and direction defined by connecting the primary and secondary vertex.

Another possibility to identify a b-jet is to reconstruct the full decay chain of a b-hadron as it is done by JetFitter algorithm [81]. The secondary vertex from b-hadron decay and tertiary vertex from subsequent decay of c-hadron inside a jet are found using Kalman filter [68]. The algorithm relies on a neural network which is fed with the information about these vertices. Similar variables as are used for the SV1 algorithm are exploited for the JetFitter algorithm. Another used variables are flight length significance (L/σ_L), information about number of tracks assigned to two-track vertices and number of single-track vertices found along b-hadron decay path. The algorithm produces three output values, each defining probability of a b-jet, c-jet and light-flavor jet hypothesis. A discriminant used for b-jet selection is defined as a logarithm of the ratio of b-jet probability factor and light-flavor jet probability factor.

Each of the b-jet identification algorithms mentioned above has its drawbacks and advantages. Algorithm utilizing vertex information has generally lower mistag rate than impact-parameter-based algorithm, but it is limited by the efficiency of secondary vertex finding ($\sim 70\%$) [81]. Therefore a multivariate technique MV2c10 [82], which combines information from the previously developed algorithms, has better performance of b-jet tagging. Variables used as an input are the IP3D and SV1 discriminating variables together with the sum of the IP3D and JetFitter discriminants. By requiring certain b-jet tagging efficiency (60%, 70%, 77% or 85%), four operating points are defined. The efficiency requirement is fulfilled by applying different selection criteria on the output of b-tagging algorithm [82].

LARGE-R JET When a high p_T top quark is produced, its decay products initiate creation of jets which are very close to each other. They can be reconstructed as one large jet. Such a jet is defined by the anti-k_t algorithm [75], which utilizes calibrated topological clusters [72] as basic units for reconstruction. The radius R is set to 1.0 to

capture all radiated particles belonging to the top-quark decay products.

Usage of the large radius of a jet leads to the increase of the amount of background radiation caught inside the jet. Usually, a jet is polluted with particles radiating from the initial states or those arising from multiple parton interactions or pile-up. Ideally, those particles should not be present inside the jet cone because they do not originate from the hard-scattering process. These sources of radiation tends to be more softer, hence, by identification and removal of the calorimeter cells containing soft radiation, the jet can be reconstructed more precisely. The technique developed for this purpose is called *jet trimming* [83].

Apart from initial-state radiation (ISR), it is essential to count in the radiation coming from final states (FSR)

It is based on a re-clustering of tracks inside the large-R jet (seed jet). The clusters of the seed jet are revisited and regrouped into subjets using k_T algorithm [84] with radius R_{sub} smaller than the radius used for a seed jet; typically $R_{sub} = 0.2$. Usage of the k_T algorithm instead of anti-k_T algorithm is well-founded. Due to different handling of jet formation, the k_T algorithm is able to establish subjets better. More precisely, it provides better energy sharing between subjets. The algorithm builds jets by combining softer clusters first, clustering them with harder radiation later, which allows more flat energy distribution among subjets containing FSR. The softness of each subjet is then tested. The p_{Ti} of a subjet is compared to the p_T of the seed jet as follows: $p_{Ti} > f_{cut} * p_T$. Subjets, which pass the criterion, are assembled to form a jet, other subjets are discarded [85]. An illustration of the trimming procedure can be found in Fig. 25.

Anti k_T algorithm tries to assign most of the energy of FSR to one subjet, which may not be possible. Other subjets are therefore less energetic, hence more likely to be discarded.

5.1.5 *Missing transverse momentum*

Particles, which leave no traces in any of the detector layers such as neutrinos, can be spotted by reconstruction of a missing transverse momentum E_T^{miss} [86]. The main idea is to look for a transverse momentum imbalance of the reconstructed detected objects (electrons, muons, τ-leptons, jets, photons), which indicate existence of undetected particles carrying complementary transverse momentum. Total transverse momentum of all particles is expected to be zero under the assumption, that p_T of initial-state partons is negligible.

RECONSTRUCTION E_T^{miss} is reconstructed from all detected particles, utilizing information from all detector subsystems. In addition, much softer detected signals, not connected to any hard objects, are assigned to the primary vertex. Both these signal types need to be accounted for when reconstructing missing transverse momentum. Hence, E_T^{miss} vector is defined as a vectorial sum:

$$E_T^{miss} = -\sum_{i \in \, hard\ objects} p_T^i - \sum_{j \in \, soft\ signals} p_T^j, \tag{25}$$

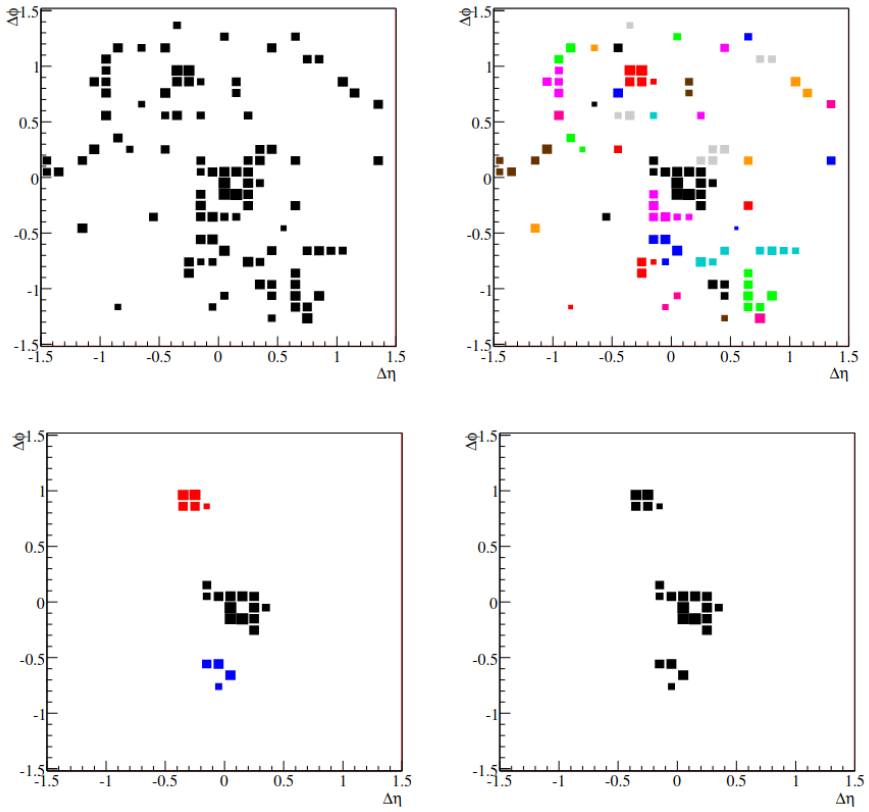


Figure 25: Illustration of the jet trimming procedure. In the upper left plot, a seed jet with $R = 1.5$ reconstructed with anti-k_T algorithm is shown. On the right, the same jet with identified subjets with $R_{sub} = 0.2$ is plotted. Subjets are marked with various colors. The left plot in the second row depicts subjets, which passed the softness criterion ($f_{cut} = 0.03$). They are highlighted with different colors. On the right, the result of the procedure is show - trimmed jet. The area of each cells is dependent on the logarithm of its p_T [83].

where *hard objects* are electrons, muons, photons, hadronically decaying tau leptons τ_{had} or jets. Reconstruction of each hard object is performed independently, so there is a rather big possibility of using the same partial information to reconstruct more than one object. Therefore, the priority order of objects has to be established to avoid double counting. The reconstruction starts with electrons, followed by photons, hadronically decaying tau leptons τ_{had} and jets. Each contribution is rejected if it was reconstructed using the same calorimeter signals/ID tracks as one of the previously added hard objects. Muons have generally very little or no signal overlap with other objects, because they are reconstructed using only ID and MS tracks. A little contribution from calorimeters is expected, because muons are minimum ionizing particles.

Reconstruction of E_T^{miss} is not sensitive to pile-up, because hard objects enter the reconstruction fully calibrated also with pile-up corrections and soft term is built using only ID tracks from the primary vertex.

5.1.6 *Overlap removal*

Finally, after reconstruction of all objects in an event, it is time to look at the whole picture.

Described object reconstruction is performed individually per each object/particle, so there is no consideration of previously reconstructed objects while reconstructing the next one. Thus, there is a need to have a set of rules, which helps to resolve ambiguities like what to prefer, i.e. when there is a reconstructed jet and an electron very close to each other. For this example, there are two scenarios: the reconstructed object is a jet and electron emerges like a secondary particle or the reconstructed object is an electron that initiated a jet. The overlap removal procedure is therefore needed for dealing with the ambiguities:

- Electron vs muon: If an electron and muon share an ID track and if the muon is calorimeter-tagged, the muon is removed. Otherwise the electron is removed.
- Electron vs jet: If there is an electron in $\Delta R < 0.2$ distance from a jet, the jet is discarded. But if there is an electron in $0.2 < \Delta R < 0.4$ distance from a jet, the electron is removed. In case of a large-R jet, if there is an electron within $\Delta R < 1.0$, the large-R jet is removed.
- Muon vs jet: A muon is removed, if there is a jet in $\Delta R < 0.4$ distance with more than two associated tracks. If the jet doesn't have 2 associated tracks, it is removed and the muon remains.

5.2 OBJECT DEFINITION IN CHARGE ASYMMETRY ANALYSIS

A short summary of object definitions used in this analysis is given below.

ELECTRONS Electron candidates should have $E_T > 28$ GeV and $|\eta_{cluster}| < 2.47$. For their identification, Tight operating point is used. The Gradient isolation working point, which provides isolation efficiency of 92% at $p_T = 28$ GeV and reaches 99% at $p_T = 60$ GeV, is applied. An electron candidate is removed if it is reconstructed in transition region between the barrel and the endcap of the LAr calorimeter $(1.37 < |\eta_{cluster}| < 1.52)$. Moreover, impact parameters of electron candidates should fulfill criteria: $|d_0|/\sigma(d_0) < 5$ and $|z_0 \sin\theta| < 0.5$ mm.

MUONS Muons reconstructed as Combined muons are utilized in the analysis. Additionally, they are required to pass identification criteria of Medium muons. Isolation of tracks is chosen similarly to electrons (Gradient isolation working point), which results in an average identification efficiency of 98%. Muon candidates should satisfy following criteria: $|\eta| < 2.5$ and $p_T > 28$ GeV, and for impact parameters $|d_0|/\sigma(d_0) < 3$ and $|z_0 s \in \theta| < 0.5$ mm.

JETS Only jets inside the central region (within $|\eta| < 2.5$) and with $p_T > 25$ GeV are considered. For low p_T jets ($p_T < 60$ GeV) with $|\eta| < 2.4$, JVT is used for differentiation of hard-scatter and pile-up jets. Chosen operating point of JVT provides a rejection factor of 99% for jets originating from pile-up. An efficiency of 92% for identification of hard-scatter jets is provided by selected operating point.

B-TAGGING PROCEDURE The operating point of b-tagging MV2c10 algorithm corresponds to 77% efficiency of tagging a b-quark jet. The rejection factors for jets originating from a c quark, tau lepton, or light quark are 5, 20, and 115 respectively [82].

LARGE-R JETS In this analysis, the trimming parameter R_{sub} is set to 0.2 and f_{cut} to 5% according to previous studies [85].

Part III

ANALYSIS METHODS

MC simulations of signal and background processes. Requirements for selection of $t\bar{t}$ events. Fully Bayesian unfolding and its implementation. Definition of systematic uncertainties: modelling and experimental. The bootstrap method and pruning of the uncertainties. Truth-based and NNLO-based re-weighting of the MC simulations.

6

ANALYSIS STRATEGY: FROM SIMULATION TO REAL DATA

After the reconstruction step is finished (jets, leptons are identified, their attributes calculated), what remains is to select only the events relevant for our measurement. In our case, those events, when $t\bar{t}$ pair is produced. The next step is to use these events to determine cross sections or to look for any effect we would like to study, i.e. charge asymmetry. Naively, one can reconstruct $\Delta|y|$ spectrum from the reconstructed top quark/anti-quark variables and calculate the value of charge asymmetry. But in that case, several important effects are left out. Firstly, the detector-level objects are determined with uncertainties, which have to be included in the measurement. Secondly, the acceptance of the detector is limited. It means that particles crossing the detector in its insensitive parts are not recorded. As every measuring device, the ATLAS detector has some finite resolution which is as well translated to the measured quantities. Therefore, a raw $\Delta|y|$ distribution does not reflect the true nature of the top-quark pair production. A remedy for this situation is proposed – the unfolding procedure. The true spectrum of the measured data distribution is exposed in this way.

Development of the analysis tools (unfolding method, methods for treatment of systematic uncertainties, etc.) is performed using Monte Carlo (MC) simulated samples for signal process and background processes. After all techniques are tested and work sufficiently with the MC samples, the analysis machinery is applied to the real, measured data.

Specifically, the Asimov data set is widely used in most of the studies. This dataset is built from signal and background MC samples to reflect truly the measured distribution without the systematic effects.

6.1 BACKGROUND PROCESSES

At the reconstruction level, processes which mimic to some extent the signature of the signal process, $t\bar{t}$, are included into the group of background processes. The group is slightly different for the single-lepton and dilepton channel.

Background processes are:

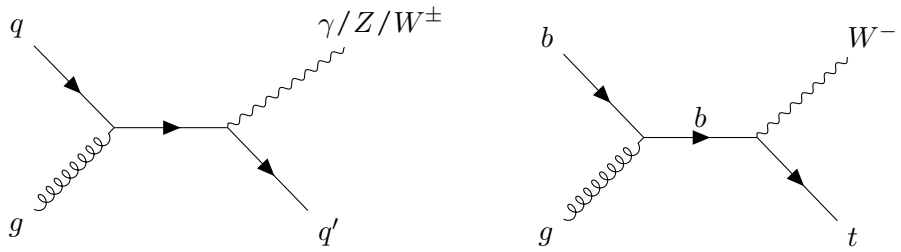
- $W+jets$: W decaying to leptons and additional jets resemble the signal process, although there are no b-jets at first, see Fig. 26a. Additional jets originate from gluon radiation (ISR, FSR), some of them may be tagged as b-jet. It is dominant background pro-

cess in the single-lepton channel. Its contribution in the dilepton channel is negligible.

For single-top process we take into account production of top quarks as well as top antiquarks with associated $b/\bar{b}$ or $d/\bar{d}$ quark.

- *Single top*: There are 3 possible production channels - s, t-channel and Wt channel. The final state particles for the first 2 channels are top quark and b/d quark, see Fig. 27. These channels are only relevant for the single-lepton channel. Last process gives rise to top quark and W, see Fig. 26b, and contributes in both, the single-lepton and dilepton channel.

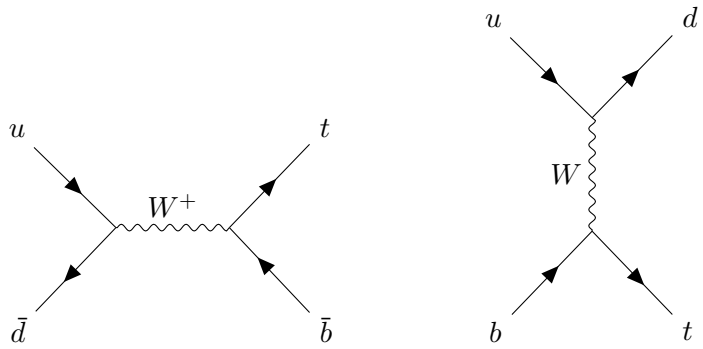
- Z+*jets*: Leptonically decaying Z boson produce 2 leptons, sufficient for mimicking of the dilepton $t\bar{t}$ topology and additional jets can possibly supply b-jets from $t\bar{t}$ decay (Fig. 26a). The process is the most dominant process in the dilepton channel, whereas its contribution in the single-lepton channel is small.



(a) V+jets (b) single top, Wt channel

Figure 26: Representative diagrams of V+jets and Wt channel of single-top production

Analogous diagrams can be drawn for top antiquark production



(a) s-channel (b) t-channel

Figure 27: Representative diagrams of single-top production in the s- and t-channel

- *Diboson* (VV): This background is more important in the dilepton channel than in the single-lepton. Production of 2 bosons

decaying into pair of leptons and some jets imitates the dilepton $t\bar{t}$ decay. (Fig. 28) Contribution in the single-lepton channel is quite small.

- $t\bar{t}X$, tWZ, tZ, *other rare processes*: The processes have small contributions in both, the single-lepton and dilepton decay channels. Examples of $t\bar{t}X$ processes are presented in Fig. 29.

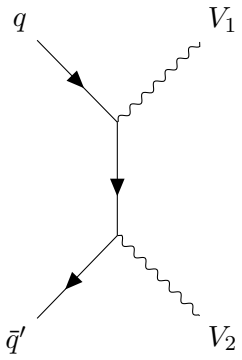
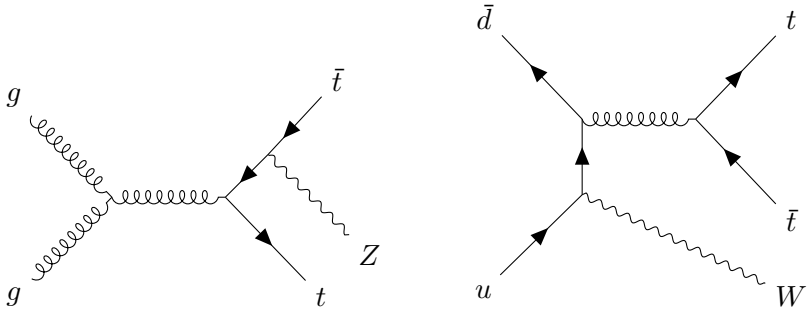


Figure 28: Representative diagram of diboson (WW, WZ, ZZ) production



(a) $t\bar{t}Z$ (b) $t\bar{t}W$

Figure 29: Representative diagrams for $t\bar{t}X$ production: $t\bar{t}Z$ and $t\bar{t}W$

Processes like production of Z+jets, two dibosons (VV, $V = W, Z$, mainly WW, WZ) or $t\bar{t}$ and additional particle (X) have small contribution to background in the single-lepton channel, therefore are merged together in the single-lepton regions.

Analogously as in the single-lepton channel, processes with little contribution, like $t\bar{t} + X$, tWZ, tZ, are merged together in the dilepton channel.

6.2 SIMULATED SAMPLES

Simulation of each process, either signal ($t\bar{t}$) or any background, is not covered by just one MC generator. A structure of a process is

very complicated, see Fig. 2. The generation of a process is divided into several steps: simulation of parton interaction = hard scattering, accompanied by parton showers and followed by hadronization of these partons. Produced hadrons subsequently decay, which needs to be modelled as well. Furthermore, effects of other parton-parton interactions in a proton have to be considered together with influence of pile-up.

Generally, at any stage of event generation, there could be an interaction among created partons/particles and particles/partons within the same or from different interaction, which complicates the simulation even more.

Last part of MC simulations of any process models how each particle is recorded by individual detecting layers of the ATLAS. The full response of the detector [87] is simulated using set of software tools `GEANT4` [88], or using reduced, faster tool – `ATLFAST-II` [89]. Full simulation gives more precise predictions. On the other hand, `ATLFAST-II` utilizes parameterization of particle showers in calorimeters, which helps to produce MC samples faster [87]. For most of the generated samples, full detector simulation is applied. The fast simulation software is largely applied for production of samples used to estimate some of the systematic uncertainties.

For modeling of heavy hadron decays, `EVTGEN 1.6.0` [90] is used for most of the samples. In case of some background samples, the `SHERPA` [91] generator is used instead. Simulation of pileup events is done via `PYTHIA 8` [92] generator with specific PDF set - MSTW20087-LO [93] with A3 [94] tuned parameter set.

6.2.1 *Signal sample*

The nominal signal sample ($t\bar{t}$) is generated using `POWHEG-BOX v2` [95–101], which calculates matrix element (ME) at next-to-leading order (NLO) in strong coupling constant α_S. Further assumptions during the simulation process are being made, like choice of the PDF or the top-quark mass. The `POWHEG` generator uses the NNPDF3.0NLO PDF set [102], the mass of the top-quark is assumed to be 172.5 GeV and additionally a so called h_{damp} parameter is set to 1.5 m_t [103]. The calculation requires choice of renormalization and factorization scale μ_r, μ_f. They are set to nominal value of $\sqrt{m_t^2 + p_{T,t}^2}$.

This generator provides for the hard-scattering part of the simulation and partially for parton shower. There is hardly any strong borderline between ME and parton shower.

h_{damp} parameter is one of the parameters affecting parton shower, it controls additional high p_T emissions coming from NLO processes.

`POWHEG` generator is interfaced with another generator, which simulates subsequent parton shower and hadronization, `PYTHIA 8.230` [92]. This generator uses NNPDF2.3LO PDF set [102] and specific set of tuned parameters called A14 [104].

To check the effect of the choice of parton shower and hadronization model, alternative sample is generated using `POWHEG` interfaced with `HERWIG 7.04` [105, 106]. The nominal $t\bar{t}$ sample (`POWHEG+PYTHIA 8.230`) and `POWHEG+HERWIG 7.04`-generated sample are compared to evaluate the impact. The `HERWIG 7.04` generator uses the MMHT2014LO PDF set [107] and the H7UE set of tuned parameters [106, 108].

For the purpose of testing the analysis procedure, samples with altered charge asymmetry prediction corresponding to heavy axigluon

contribution are simulated. These samples are produced by PROTOS [109] generator with CTEQ6L1 PDF set [110].

6.2.2 *Background samples*

SINGLE TOP Single-top Wt production is simulated using the same generator as the one used for signal modeling, POWHEG BOX v2. The same PDF set, μ_r, μ_f scales are chosen as well. Potential interference with $t\bar{t}$ production is evaluated using diagram removal (DR) scheme [111]. Additionally, samples, which are produced by employing a diagram subtraction (DS) scheme [111] are exploited for estimation of the uncertainty for $t\bar{t}$-single top overlap treatment.

Single-top s-channel and t-channel processes are modelled with the same choice of generator. Following the discussion in [112], the μ_r, μ_f scales are set to $\sqrt{m_b^2 + p_{T,b}^2}$ [100].

Hadronization and parton showers are modelled with PYTHIA 8.230 with the same PDF set and set of tuned parameters as is used for $t\bar{t}$ production. Analogously as for signal sample, the uncertainty covering differences caused by the specific choice of parton shower and hadronization model is estimated by generating alternative single-top sample by POWHEG interfaced with HERWIG 7.04 with the same settings as for $t\bar{t}$ sample.

W+JETS AND Z+JETS W+jets and Z+jets background processes are generated by SHERPA 2.2.1 [91] generator. For matrix element part of simulation, two SHERPA libraries are used, Comix [113] and Open-Loops [114, 115]. ME is calculated to the NLO level for events with up to 2 jets, and only to LO for events with up to 4 jets. The nominal SHERPA parton shower [116] simulation based on Catani-Seymour dipole factorization model together with the cluster hadronization model [117] is utilized. Specific set of tuned parameters are used in this simulation, based on the NNPDF3.0NNLO PDF set [102] and derived by authors of SHERPA.

DIBOSON Diboson samples are simulated also with the SHERPA generator, specifically with version 2.2.1 and 2.2.2. The newer version is used for 2 and 3-lepton samples. Emissions of additional hard-partons [113] are interfaced with a PS based on Catani-Seymour dipoles and using the NNPDF3.0NNLO PDF set. Parton-shower parameters tuned by authors of SHERPA are applied. Matching of matrix element part of simulation with parton shower [118] is performed separately for various jet multiplicities. Afterwards, all samples are merged together using CKKW matching procedure [119, 120]. The extension of this procedure to NLO is provided via MEPS@NLO procedure [121]. The NLO level is reached for up to one additional parton. For up to three additional partons, the LO level is calculated.

$t\bar{t}V$ AND $t\bar{t}H$ The nominal generator, POWHEG BOX v2, and the MADGRAPH5_aMC@NLO 2.3.3 [122] provide simulation of $t\bar{t}H$ and $t\bar{t}V$ processes, respectively. Matrix element simulation is provided to NLO order in α_S using NNPDF3.0NLO PDF set. Specifically, part of the $t\bar{t}H$ events corresponding to 2018-year of data taking, is generated with the MADGRAPH5_aMC@NLO 2.6.0 generator. Subsequent hadronization is modelled with PYTHIA 8.230 ($t\bar{t}V$) or PYTHIA 8.210 [92] ($t\bar{t}H$). The NNPDF2.3lo PDF set is used and the A14 tune is employed.

RARE SM PROCESSES - tWZ, tZ These processes are considered mostly in the dilepton channel. In the single-lepton channel, their contribution is negligible. The simulation is provided by MADGRAPH5_aMC@NLO 2.3.3 generator at NLO with the NNPDF3.0nlo PDF set. Hadronization is simulated with the Pythia 8.212 [92]. Again the A14 tune is used and the NNPDF2.3LO PDF set is chosen.

SCALING FACTORS FOR Z+JETS In the dilepton channel, Z+jets background play an important role. However, its simulation suffers from considerable uncertainties. Decay of the Z boson, $Z \to \ell\bar{\ell}$, contributes mostly to the same-flavor regions, hence the fraction of this background is larger in these channels. For derivation of scale factors, a control region for this background is established by requiring invariant mass of the opposite-sign dilepton pair to be inside an interval of 10 GeV around the mass of the Z boson. Selection criteria on E_T^{miss} are lifted for this control region to increase the number of background events. Finally, a scale factor μ_z is defined as a number of events in data after subtracting other processes divided by the simulated number of Z+jets events.

Contribution of other processes is estimated by the MC predictions.

Due to different contributions of Z+jets background in regions with different b-jet multiplicity ($1b$-tag-exclusive and $2b$-tag-inclusive), independent scale factors are extracted for both regions. Normalization uncertainty of each scale factor is set to 30% in order to cover systematic and statistical uncertainties of the derivation procedure. Derived scale factors are used to correct the MC simulation of Z+jets background in the same flavor regions of the dilepton channel.

6.3 FAKE LEPTON AND NON-PROMPT LEPTON BACKGROUND

Another contribution to the set of background processes creates those events, when a non-prompt lepton or other particle which mimics a lepton is reconstructed and hence treated as a real lepton. Short-term name for these events is fakes or fake events.

Fake events are predominantly those, in which a heavy hadron decays semi-leptonically or a pion/kaon decays into leptons. Additionally, when a neutral pion is produced, it may be mis-reconstructed

as an electron. Similarly, an electron from photon conversion or a prompt photon can be misidentified as prompt electron.

SINGLE-LEPTON CHANNEL: Estimation of fake events is provided by data-driven matrix method [123]. Two subsets of events are created: those which pass loose selection criterion (required only identification) and those which pass tight selection criterion (required identification and isolation). The efficiency scale factors are derived as a ratio of number of events in tight selection subset and number of events in loose selection subset:

$$\epsilon_{real/fake} = \frac{N^{tight}_{real/fake}}{N^{loose}_{real/fake}}. \tag{26}$$

Number of real lepton events are estimated using tag-and-probe method on Z+jets events when Z decays to leptons. On the other hand, fake lepton events are measured in a control region with enhanced number of non-prompt/fake leptons.

In the end, both efficiencies $\epsilon_{real/fake}$ are used to derive weights, which are applied on events from data satisfying loose lepton selection,

$$w_{loose} = \frac{\epsilon_{fake}\epsilon_{real}}{\epsilon_{real} - \epsilon_{fake}}. \tag{27}$$

This weighted distribution is the final estimation of fake events' contribution.

DILEPTON CHANNEL: Contribution of fake events in the dilepton channel is estimated through simulations. The normalization of these processes is obtained by a data-driven method. A scale factor for fake leptons, μ_{fake}, is derived using subset of data events/simulated events passing standard selection criteria (see the beginning of Chap. 5 and Sec. 5.2) together with requirement of two leptons with the same sign of electric charge. The μ_{fake} is then defined as a ratio of simulated events and data events after subtraction of MC-predicted signal fraction in the same-sign region. The scale factor is calculated separately for $ee + e\mu$ events and $\mu\mu$ events.

Additionally in $ee + e\mu$ events, a charge misidentification rate is determined by comparing events with opposite-sign and same-sign charges of 2 leptons, which have the invariant mass inside 10 GeV mass window around the Z mass. The events with the same-sign charges of leptons are mostly signal $t\bar{t}$ events, single top or Z+jets events, when one of the lepton charges has been wrongly identified.

Evaluated effect of charge misidentification is combined with previously derived fake-lepton scale factors to obtain their final values. It is measured separately for events with different b-jet multiplicities ($1b$-tag-exclusive and $2b$-tag-inclusive). A 30% normalization uncertainty in $ee + e\mu$ regions covers statistical uncertainty, charge misidentification and range of estimates of fake leptons in kinematic variables. In

$\mu\mu$ events, a normalization uncertainty of 50% is used, because of low number of same-sign events in this region and high rate of $t\bar{t}V$ background. The rate of fake $\mu\mu$ events is less than 1% in the signal region.

6.4 EVENT SELECTION FOR CHARGE ASYMMETRY ANALYSIS

Charge asymmetry analysis utilizes data from pp collisions at the center-of-mass energy $\sqrt{s} = 13$ TeV collected by the ATLAS detector throughout the years $2015 - 2018$. An event which is considered in the analysis, should contain identified primary vertex and should pass a single-electron or single-muon trigger. The lowest acceptable p_T of a lepton is set in order to improve selection efficiency. In 2015 this threshold was set to 20 GeV for an electron and 24 GeV for a muon candidate. During remaining years of data taking (2016-2018) common threshold of 26 GeV was applied for both lepton types [124, 125]. Additionally, the p_T threshold trigger is accompanied by the isolation requirements which should reduce the trigger rate.

This amount of data corresponds to the integrated luminosity of 139 fb^{-1}.

6.4.1 *Single-lepton channel*

Signature of the single-lepton event comprises of one lepton, 4 jets corresponding to 4 quarks (b, $\bar{b}$ and quarks from W decay) and missing transverse momentum corresponding to a neutrino. Therefore, a single-lepton event should have exactly one lepton which matches the trigger lepton and p_T higher than 28 GeV. Events containing more than one lepton with $p_T > 25$ GeV are discarded. At least one of the small-R jets has to be tagged as a b-jet.

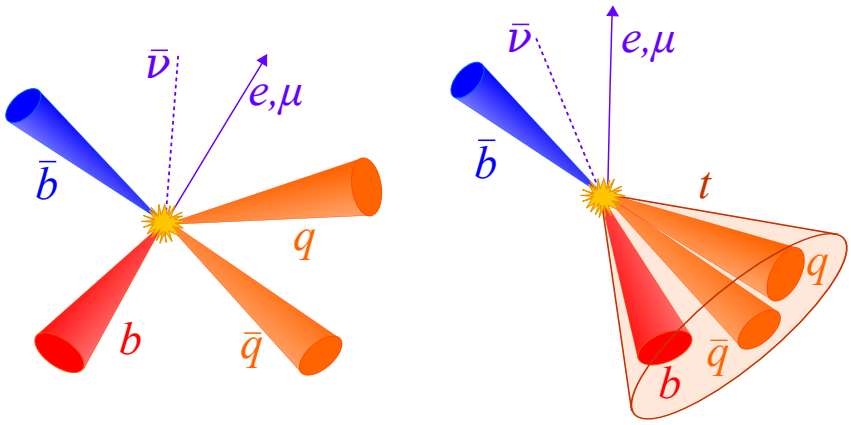
M_T^W is a quantity defined as $M_T^W = \sqrt{2p_T^l E_T^{miss}\Theta}$, where Θ is $(1 - cos(\phi))$ and ϕ is azimuthal angle between a lepton and missing transverse momentum.

Another criteria for E_T^{miss} and M_T^W are applied in order to reduce the fake and non-prompt lepton background. For events with electrons (e+jets), two separate cuts are applied on these variables: each E_T^{miss} and M_T^W has to be higher than 30 GeV. Otherwise, for events with muons (μ+jets), the events should pass combined (triangular) cut: $E_T^{miss} + M_T^W > 60$ GeV. The reason for a separate treatment is a different rate of fake/non-prompt lepton background in e+jets and μ+jets channel, respectively. In e+jets channel, there is a higher probability for a particle/object to be tagged as a proper electron.

Non-prompt leptons do not arise from primary vertex, but rather from another underlying interactions.

Afterwards, the single-lepton events are divided according to b-jet multiplicity to 1b-tag-exclusive and 2b-tag-inclusive region regardless of lepton flavor. Events in e+jets and μ+jets channels are merged. Spatial arrangement of the top-quark pair decay products can differ quite a lot from one single-lepton event to another. We recognize two topologies – resolved and boosted topology. In Fig. 30 there is an illustration of both. For each topology, different, more suitable reconstruction methods are employed.

To sum up, events in the single-lepton channel are divided to four regions: resolved 1-b-tag-exclusive and 2b-tag-inclusive and boosted 1-b-tag-exclusive and 2b-tag-inclusive regions.



(a) Resolved topology (b) Boosted topology

Figure 30: Sketch of topologies of $t\bar{t}$ decay products in the single-lepton channel: resolved (a) and boosted (b) topology. Top-antiquark decay products are in blue color, top-quark decay products in red and orange.

RESOLVED TOPOLOGY Events, which belong to the resolved topology should have good spatial separation of jets. Hence, at least 4 small-R jets with p_T higher than 25 GeV have to be reconstructed in a resolved event. The most challenging task in reconstruction of events in this topology is the correct pairing of reconstructed jets with partons coming from the top-quark pair decay. The technique used for this task is a boosted decision tree (BDT) implemented within TMVA package [126], which combines various kinematic variables describing an event, the information about jet b-tagging together with weight computed by Kinematic Likelihood Fitter (KLFitter) [127] into one single number for each jet-parton pair. The discriminant is evaluated for each permutation of jet-to-parton assignments and only the combination giving the highest value is accepted and used for $t\bar{t}$ reconstruction.

Among the variables used as an input to the BDT, the reconstructed mass of the hadronically/semileptonically decaying top quark or the reconstructed mass of the hadronically decaying W boson are listed. Full list can be found in [2].

The BDT is trained on $t\bar{t}$ simulated sample where a signal permutation is the permutation when all four jets are correctly matched to the partons, requiring that the parton is closer to the jet than $\Delta R = 0.3$. In case of multiple jets satisfying ΔR criterion for one parton, the jet closest to the parton is taken. All other permutations are considered as background. Due to rapid increase in the number of permutations with rising number of jets, only up to 5 jets per event are considered. If there are more than 5 jets in an event, two with the highest b-tagging scores are taken together with the other 3 jets with the highest p_T. The

training is performed separately for $1b$-exclusive and $2b$-inclusive region.

Semileptonic top-quark decay: $t \to W^+b \to \ell^+\nu_\ell b$

One of the discriminating variables used as an input variable to the BDT is the reconstructed mass of semileptonically-decaying top quark. It is calculated using four-momentum vectors of top-quark decay products. However, only the size of neutrino transverse momentum is measured, E_T^{miss}, together with its azimuthal angle ϕ so we know neutrino momentum vector in 2-dimensional transverse plane. If the neutrino mass is neglected, only one unknown coordinate remains - longitudinal momentum of a neutrino, p_z^ν. This coordinate is determined using constraint from the W mass, which provides quadratic equation for p_z^ν. In case of the situation, when this equation has real solutions, the one which gives the top-quark mass closer to the expected value is taken. If an imaginary solution is obtained, the other components of neutrino four-momentum are varied until this equation produces real solutions.

$m_W^2 = (p^\ell + p^\nu)^2$

Events with BDT score higher than 0.3 are considered as signal events. Remaining events with lower score are generally background events. Selection of this threshold on the BDT score helps to reduce combinatorial background from $t\bar{t}$ events and background events when a $t\bar{t}$ is not created. The efficiency of making correct jet-to-parton assignment is 75% for those signal $t\bar{t}$ events where it is possible to assign all jets to partons. The yields of resolved events for signal and background MC samples are listed in Tab.3.

Table 3: Event yields in the resolved topology split by b-tag multiplicity ($1b$-excl., $2b$-incl.) in the single-lepton channel. The presented uncertainty is the total uncertainty, both statistical and systematic uncertainties are included.

Process:	Single-lepton resolved	
	$1b$-excl.	$2b$-incl.
$t\bar{t}$	$1\,540\,000 \pm 140\,000$	$1\,870\,000 \pm 170\,000$
Single top	$90\,000 \pm 11\,000$	$51\,000 \pm 8\,000$
W+jets	$180\,000 \pm 100\,000$	$20\,000 \pm 9\,000$
$Z + VV + t\bar{t}X$	$48\,000 \pm 25\,000$	$14\,000 \pm 7\,000$
Fake	$90\,000 \pm 50\,000$	$47\,000 \pm 24\,000$
Total Prediction	$1\,940\,000 \pm 190\,000$	$2\,010\,000 \pm 180\,000$
Data	$1\,964\,127$	$2\,041\,063$

BOOSTED TOPOLOGY When one of the initial-state partons has much larger value of the longitudinal momentum than the other colliding parton, p_z, one of the final-state particles, $t/\bar{t}$ quark, will also have

large momentum in this direction. Consequently, decay products of that particular quark ($t/\bar{t}$) will be very close to each other forming wider jet - large-R jet, see Fig. 30b. Naturally, different treatment of the reconstruction process for these events is vital for more efficient reconstruction.

It is always the hadronically-decaying quark/antiquark whose decay products collimate into a single large-R jet.

A boosted event has to contain at least one large-R top-tagged jet with $p_T > 350$ GeV and $|\eta| < 2$ and at least one small-R jet close to the identified lepton. Distance of this jet and lepton in $\phi - \eta$ space should be $\Delta R < 1.5$. In $t\bar{t}$ rest frame, large-R top-tagged jet and small-R jet close to lepton together with this lepton should follow the opposite directions. Based on this fact, additional conditions can be applied on boosted events to improve correct $t\bar{t}$ system reconstruction:

- $\Delta\phi(\text{large-R jet}, \ell) > 2.3$,
- $\Delta R(\text{large-R jet}, \text{small-R jet}) > 1.5$.

Top-tagging procedure [128] is performed using information about the attributes of a top-tagged jet candidate, merely jet mass and τ_{32} substructure variable [129]. The large jet should comprise of three smaller jets after trimming procedure (see Sec. 5.1.4). Hence, three calorimeter clusters corresponding to large energy deposits should be found in close proximity ($\Delta R = 1.0$). Light-quark or gluon jets, which are considered as background jets, are associated with just one cluster and soft wide-angle emissions. Analogously, a jet from the W-boson consists of two subjets corresponding to quarks (decay products of the W). These attributes of large jets are utilized by τ_{32} substructure variable, which is tailored to differentiate jets according to the mentioned jet properties.

Obviously, the top-tagged jet should have associated energy deposit around the top-quark mass (172.5 GeV) apart from light background jets, whose energy spectrum should fall exponentially and has its peak around much smaller energies. Additionally, W-boson jets also contribute to the background. Naturally, the peak in energy spectrum for theses jets is at higher value, approximately 80 GeV. An operating point corresponding to 80% efficiency of top-tagging procedure is chosen. Due to bad reconstruction of events corresponding to low $m_{t\bar{t}}$ values the threshold on $m_{t\bar{t}}$ is set to be 500 GeV. Below this threshold only negligible fraction of events passes the boosted selection criteria. It is important to note that those events which satisfy both resolved and boosted selection criteria are considered as boosted events (boosted veto).

The estimate of the hadronically-decaying top-quark momentum is the four momentum of leading-p_T large-R jet. Similarly, the semileptonically-decaying top-quark momentum is build from momenta of isolated lepton, small-R jet and neutrino. Four momentum of neutrino is constructed from constraints on E_T^{miss}, from kinematics of lepton and W boson mass as it is done for events in the resolved topology.

The yields of boosted events for signal and background MC samples are listed in Tab.4.

Table 4: Event yields in the boosted topology split by b-tag multiplicity ($1b$-excl., $2b$-incl.) in the single-lepton channel. The presented uncertainty is the total uncertainty, both statistical and systematic uncertainties are included.

Process:	Single-lepton boosted	
	$1b$-excl	$2b$-incl.
$t\bar{t}$	$50\,000 \pm 12\,000$	$74\,000 \pm 18\,000$
Single top	$3\,600 \pm 1\,100$	$3\,000 \pm 1\,100$
W+jets	$8\,900 \pm 2\,600$	$1\,600 \pm 500$
$Z + VV + t\bar{t}X$	$2\,400 \pm 1\,200$	$1\,400 \pm 700$
Fake	$3\,000 \pm 1\,500$	$2\,300 \pm 1\,200$
Total Prediction	$68\,000 \pm 14\,000$	$83\,000 \pm 18\,000$
Data	54 750	66 571

6.4.2 *Dilepton channel*

Event signature of the dilepton $t\bar{t}$ decay channel consists of two opposite-charge leptons, two b-jets and missing transverse momentum as a sign of the presence of 2 neutrinos. An illustration of the topology is shown in Fig. 31. Hence, basic requirement for a dilepton event is identification of 2 opposite-charge leptons, whose p_T should be larger than 28 and 25 GeV for leading and sub-leading lepton, respectively. Furthermore, at least 2 small-R jets with p_T higher than 25 GeV have to be identified. Additionally, at least one of the small-R jets should be b-tagged.

One of the Z-boson possible decays is a decay to two opposite charge leptons.

One of the background processes which can mimic the signal process very successfully, is production of Z boson and some additional jets, Z+jets process. The fraction of this background process is reduced by requiring the invariant mass of a dilepton pair not to be consistent with Z boson mass, i.e. $|m_{\ell\bar{\ell}} - m_Z| > 10$ GeV. Analogously, a missing transverse energy E_T^{miss} should be higher than $60/30$ GeV in $1b$-tag-exclusive/$2b$-tag-inclusive region in order to reduce Z+jets background. Production of low mass resonances is reduced using similar principle. In $1b$-tag-exclusive region, for events with leptons of the same flavor, $m_{\ell\bar{\ell}}$ is required to be higher than 15 GeV.

In an event, more than just 2 jets can be found. Therefore, possible ambiguities needs to be resolved. In case of only 1-b tagged jet, another jet with the highest p_T is taken for the top quark/anti-quark

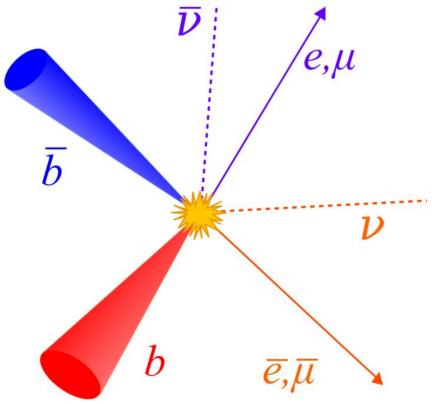


Figure 31: Sketch of topology of $t\bar{t}$ decay products in the dilepton channel. Top-anti-quark decay products are in blue color, top-quark decay products in red and orange.

reconstruction. If 2 or more b-tagged jets are identified, two with the highest b-tagging weight are considered for $t\bar{t}$ reconstruction.

Reconstruction of $t\bar{t}$ system in this decay channel is provided by Neutrino Weighting method [130]. Two unobserved neutrinos in this signature make the reconstruction of top-quark momenta in these events difficult. By imposing a condition of reconstructed invariant mass of the W boson or top quark upon its decay products (lepton and neutrino in case of W boson; lepton, neutrino and b-jet for top quark), the system of equations is established. A solution of this system is obtained for different values of neutrino pseudorapidities η_ν, $\eta_{\bar{\nu}}$, which are drawn from interval $(-5, 5)$ with 0.2 step. A different weight is assigned to each solution according to the level of agreement between the calculated E_T^{miss} and measured E_T^{miss}:

$(p^\ell + p^\nu)^2 = m_W^2$, $(p^\ell + p^\nu + p^{bjet})^2 = m_t^2$

$$w = \exp\left(\frac{-\Delta E_x^2}{2\sigma_x^2}\right) \exp\left(\frac{-\Delta E_y^2}{2\sigma_y^2}\right), \tag{28}$$

where $\Delta E_{x,y}$ express the difference between calculated and measured E_T^{miss} in transverse plane (x and y directions), whereas $\sigma_{x,y}$ signify the resolution of the measured E_T^{miss}. The η_ν, $\eta_{\bar{\nu}}$ values corresponding to the solution with the highest weight are chosen for reconstruction of $t\bar{t}$ system. Unphysical solutions suggesting e.g. negative top-quark energy are discarded. In such cases, the weight approaches 0, therefore the weight corresponding to the chosen solution should be larger than zero.

If, however, the solution cannot be found for any value of neutrino pseudorapidities, p_T of jets is smeared together with the top-quark mass in order to obtain a solution. Those events, for which the solution does not exist even after aforementioned variation of observables, are excluded from the signal distribution. They are accounted for in the method inefficiency. For MC simulated signal sample, the ineffi-

m_t is varied between 171 and 174 GeV, jet p_T is smeared by Gaussian function, which width varies between 8% and 14% of p_T.

ciency value is 15%. The number is higher for background samples, 40%, hence providing reduction of background.

All the events, which are considered as dilepton are further divided into 4 regions according to the b-jet multiplicity (in the same way as in the single-lepton channel) and lepton flavor. There are distinct regions for the same flavor leptons ($ee + \mu\mu$) and different flavor leptons ($e\mu$). The event yields after the selection for signal and background MC samples are listed in Tabs. 5, 6.

Table 5: Event yields in the different lepton-flavour region split by b-tag multiplicity (1-excl., 2-incl.) in the dilepton channel. The presented uncertainty is the total uncertainty, both statistical and systematic uncertainties are included.

Process:	$e\mu$	
	1b-excl.	2b-incl.
$t\bar{t}$	225 000 ± 14 000	262 000 ± 18 000
Single top (Wt)	13 800 ± 1 200	5 600 ± 900
Diboson	650 ± 70	45 ± 6
Z+jets	610 ± 200	77 ± 25
Rare SM ($t\bar{t}X$, tWZ, etc)	690 ± 90	1 040 ± 140
Fake	2 500 ± 700	2 100 ± 600
Total Prediction	243 000 ± 14 000	271 000 ± 18 000
Data	244 258	273 856

Table 6: Event yields in the same lepton-flavour region split by b-tag multiplicity (1-excl., 2-incl.) in the dilepton channel. The presented uncertainty is the total uncertainty, both statistical and systematic uncertainties are included.

Process:	$ee + \mu\mu$	
	1b-excl.	2b-incl.
$t\bar{t}$	105 000 ± 7 000	192 000 ± 13 000
Single top (Wt)	6 100 ± 600	4 100 ± 700
Diboson	296 ± 34	55 ± 6
Z+jets	3 500 ± 1 400	3 800 ± 1 200
Rare SM ($t\bar{t}X$, tWZ, etc)	390 ± 50	850 ± 110
Fake	1 200 ± 400	1 900 ± 700
Total Prediction	117 000 ± 7 000	203 000 ± 13 000
Data	116 096	202 967

6.5 UNFOLDING PROCEDURE

The core of the analysis is an unfolding procedure. It helps to recover a true distribution from the distribution measured by the detector. If one knows how the detector responds on incoming particles, the inversion of this response should provide sufficient tool for obtaining the true distribution. The easiest way of how to encode this response into useful object is to define binned distribution for the measured spectrum and the true spectrum. The response is encoded in the matrix which transforms the true spectrum into the measured one - response matrix.

The response matrix is calculated from simulated reconstructed and true spectrum, while taking into account not only event migrations between bins, but also effects of selection efficiency and acceptance.

However, the matrix inversion in this naive unfolding brings one, quite big problem - the resulting "real" distribution suffers from large fluctuations and high anti-correlations between neighboring bins. Introduction of regularization terms, which modify the bin contents, solves the mentioned issue, but causes additional bias.

Occurrence of these fluctuations is not physical effect, it is caused by the nature of the matrix inversion problem.

Except the simplest unfolding method, there exist many different approaches to this problem of how to obtain the real information from the measured distribution. Specific parts of the unfolding methods are those which exploit Bayesian statistics. The unfolding method used in this analysis is one of them.

6.5.1 *Fully Bayesian Unfolding*

The name "fully Bayesian" is quite fitting for this technique of unfolding. It is basically a practical application of Bayesian inference to the unfolding problem. Many of the unfolding techniques provide as a result one number/parameter, like the strength of the signal or multiple parameters. Apart from these methods, fully Bayesian unfolding (FBU) [131] provides posterior probability *distribution* of a parameter of interest as a final product. It is given by

Both true distribution = what emerges from the collision, and reconstructed distribution = the detector simulation applied to the true distribution, are simulated.

$$p(T|D,\mathcal{M}) \sim \mathcal{L}(D|T,\mathcal{M})\,\pi(T), \tag{29}$$

where the response matrix is denoted as $\mathcal{M}$, true spectrum as T and measured data as D. Following the Bayesian inference, the posterior probability distribution of the true spectrum T given the observed data D is defined by likelihood function $\mathcal{L}$ and by prior probability density $\pi(T)$ of the true spectrum T. The likelihood function describes the probability to measure data D given the true spectrum is T.

LIKELIHOOD FUNCTION $\mathcal{L}$ Assuming Poisson statistics for bin contents of observed data D, the likelihood is defined as a multiplication of Poissonian terms for each bin of spectrum, where expected con-

tent is given by a reconstructed signal spectrum R and background estimation B:

$$\mathcal{L}(D|T,\mathcal{M},B) = \prod_{i=1}^{N_r} \frac{(R_i+B_i)^{D_i}}{D_i!}\, e^{-R_i}, \tag{30}$$

where N_r is number of bins in the reconstructed spectrum R. The reconstructed signal spectrum R is constructed from the true spectrum T and response matrix $\mathcal{M}$:

$$R_j = \sum_{i=1}^{N_t} \mathcal{M}_{ji} T_i, \tag{31}$$

where N_t is number of bins in the true spectrum T. The response matrix $\mathcal{M}$ is $N_t \times N_r$-dimensional matrix, defined as $\mathcal{M}_{ij} = \epsilon_j\, P(R_i|T_j)$, where individual components are

- ϵ_j: probability of a true event from bin T_j to be reconstructed in the detector and passing selection criteria,
- $P(R_i|T_j)$: probability for an event from true bin T_j to be observed in reconstructed bin R_i.

PRIOR PROBABILITY DENSITY OF TRUE SPECTRUM The subjectivity of Bayesian approach enters the procedure in a form of prior probabilities. The prior probability should mirror our knowledge of the true spectrum or our belief/preference for some of them. In this analysis, prior probability density for the true spectrum T is uniform in specific intervals defined by MC prediction:

One can say that prior probabilities fulfill the function of regularization terms, which are used in other unfolding techniques.

$$\pi(T) \propto \begin{cases} 1 \;\text{ if } T_i \in [T_i^{min}, T_i^{max}];\ \forall i \in [1, N_t], \\ 0 \;\text{ otherwise.} \end{cases} \tag{32}$$

By assigning the uniform probability density to the whole collection of true spectra, we state, that none of them is preferred or suppressed.

The T^{min} is the lower bound of the true bin, in our case it is 0. The T^{max} is the upper bound which is set to $2T$, where $2T$ means two times the true-bin content given by the simulated $t\bar{t}$ true-level sample.

The prior probability density defined in this way can be imagined as a constant function in a N_t - dimensional cube, where N_t is number of the true bins in spectrum and length of each edge is $T^{max} - T^{min} = 2T$.

TRUE AND RECONSTRUCTED SPECTRUM IN CHARGE ASYMMETRY ANALYSIS The variable of interest is the absolute rapidity difference $\Delta|y|$. Charge asymmetry can be easily computed from its distribution. The true spectrum T consists of four $\Delta|y|$ bins, two bins above 0 ($\Delta|y| > 0$) and two bins below 0 ($\Delta|y| < 0$), see an illustration in Fig. 32. In the same Figure, there is a sketch of the reconstructed distribution R. The R spectrum comprises sets of 4 $\Delta|y|$ bins,

one set per each region. For example, in the single-lepton channel there are 4 regions (see Sec. 6.4.1), so the total number of bins is $N_{\Delta|y|} \times N_{reg} = 4 \times 4 = 16$ bins. A visualization of response matrix, which transforms the true spectrum into the reconstructed one, is also given in Fig. 32.

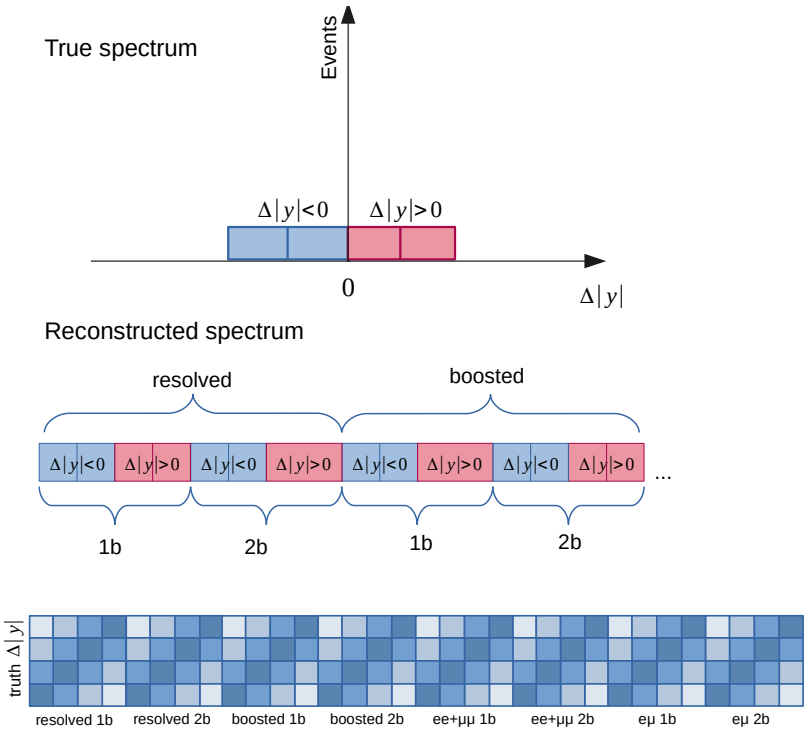


Figure 32: The upper sketch depicts the basic true spectrum T with $4\,\Delta|y|$ bins. The middle diagram illustrates how the number of bins, N_r in R is increased by dividing the sample into multiple regions, N_{reg}. The last picture shows the scheme of the response matrix with the reconstructed spectrum on the horizontal axis, built from the single-lepton and dilepton channels.

6.5.2 *Systematic uncertainties in the FBU*

A significant advantage of the FBU, in contrast to other available unfolding methods, is that the likelihood formalism allows for systematic uncertainties in natural way. The inclusion is performed by extension of the likelihood function by nuisance parameters (NP). For each systematic uncertainty, there is at least one parameter θ

There is no need to create pseudo-data and run multiple pseudo-experiments.

Number of parameters in the FBU grows immensely by inclusion of at least one NP θ per each systematic uncertainty.

$$\mathcal{L}(D|T) \rightarrow \mathcal{L}(D|T,\theta). \tag{33}$$

These parameters are used to describe effects of individual systematic sources on the reconstructed spectrum R and background estimation B (more details in Chap. 7). A prior probability is assigned to each NP. The initial effect of each NP is estimated by Gaussian probability

term with mean value equal to 0, and σ equal to 1, $G(\theta; 0, 1)$. The likelihood function therefore has the form

$$\mathcal{L}(D|T) \rightarrow \mathcal{L}(D|T, \theta)\, G(\theta). \tag{34}$$

The shorter term for NP prior probability term $G(\theta)$ is used from now on.

Effect of each systematic source is parameterized by NPs, hence the reconstructed spectrum has following structure:

$$R_i(T, \theta_s) = R_i(T, 0) \left(1 + \sum_k \theta_s^k \Delta R_i^k\right), \tag{35}$$

where

- $R_i(T, 0)$ is reconstructed-level prediction in i^{th} bin without effects of NPs,
- ΔR_i^k is the relative systematic uncertainty variation on the signal yield in the ith bin corresponding to k^{th} NP,
- pull of the k^{th} nuisance parameter is labeled as θ_s^k.

The lower index of θ_s means 'systematic'.

Similarly, the prediction for each background process is:

$$B_i(\theta_s, \theta_b) = B_i(0)(1 + \theta_b \Delta B) \left(1 + \sum_k \theta_s^k \Delta B_i^k\right), \tag{36}$$

where individual elements are

- $B_i(0)$ is the predicted yield of background in i^{th} bin without effects of NPs,
- ΔB is the relative uncertainty on the background normalization,
- θ_b is the pull of the background normalization NP,
- ΔB_i^k is the relative systematic uncertainty variation on the background yield in the i^{th} bin corresponding to k^{th} NP.

The term pull, used in the previous equations, is used to describe the estimator of a nuisance parameter $\hat{\theta}_s^k$. In the FBU, the estimator $\hat{\theta}_s^k$ is defined as the mean value of the marginal posterior distribution of θ_s^k nuisance parameter, see Sec. 6.5.3 for details.

Finally, because only the posterior distribution of true bins is of interest, all other NP are integrated out:

$$\mathcal{L}_{marginal}(D|T) = \int \mathcal{L}(D|R(T, \theta_s), B(\theta_s, \theta_b))\, G(\theta_s)\, G(\theta_b)\, \mathrm{d}\theta_s\, \mathrm{d}\theta_b. \tag{37}$$

6.5.3 *Combination of multiple regions*

The likelihood-based approach in FBU provides very simple inclusion of different regions into the likelihood function. Because of the form of likelihood function, additional regions are embedded in it by trivial multiplication [131]:

$$\mathcal{L}(D_1, ..., D_{N_{reg}}|T) = \int \prod_{i=1}^{N_{reg}} \mathcal{L}(D_i|T;\theta)\, G(\theta)\, d\theta. \tag{38}$$

In the inclusive charge asymmetry measurement, we leverage mentioned feature by using four different regions in the single-lepton and also in the dilepton channel.

By using multiple regions, the information about correlations of systematic uncertainties is exploited. In addition, the systematic-uncertainty constraints are extracted from the marginal posterior distribution of NPs. In many cases, the prior estimate of the NP, $G(\theta; 0, 1)$, is corrected by the FBU. The posterior distribution of θ, $G(\theta; pull, constraint)$, may give the estimator $\hat{\theta}$ with smaller σ than was assumed before, i.e. $constraint < 1$ (see Fig. 33 for reference). Furthermore, central value of $\hat{\theta}$ may be shifted from nominal value, which is 0, to a different value, which is labeled as $pull$. Combination of both effects, constraints and correlations of NPs, can lead to the reduction of the total uncertainty.

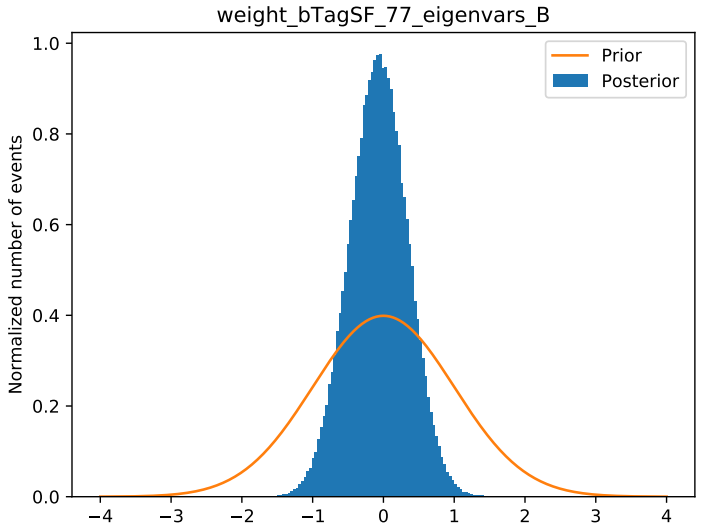


Figure 33: The prior and posterior probability density for one of the NPs, which correspond to scale factors describing efficiency of b-tagging algorithm to identify a b-jet. Width of posterior distribution is smaller in comparison with prior distribution. The effect of NP is reduced after the FBU.

6.5.4 *Posterior probability distribution for A_C*

Posterior distribution of any variable, which can be determined from bins of true spectrum, can be obtained in a similar manner.

From the posterior distribution of the true spectrum T, $p(T|D)$, the charge asymmetry posterior distribution $p(A_C|D)$ is derived. It is computed by sampling the likelihood using Markov-chain Monte Carlo method [132]. In other words, the numerical integration of following formula is performed:

$$p(A_C|\mathrm{D}) = \int \delta(A_C - A_C(T))\, p(T|D)\, \mathrm{d}T. \tag{39}$$

The chosen Markov-chain MC method uses Hamiltonian mechanics to make steps in multidimensional space (in our case $(N_t + N_\theta)$ - dimensional space) - Hamiltonian Monte Carlo (HMC) algorithm. The HMC interprets the likelihood function (more precisely, the negative loglikelihood) as a potential energy and calculates the first derivatives of the Hamiltonian in order to take the steps. Additionally, the sampler does not allow to retrace the steps which contributes to a faster convergence. A simplistic illustration of sampling and retrieving posterior distribution of A_C is pictured in Fig. 34.

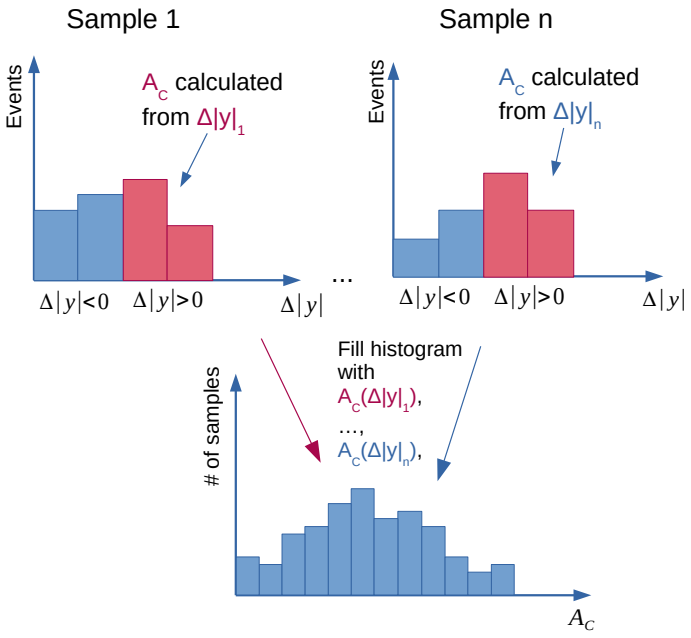


Figure 34: Graphic illustration of how the posterior distribution of charge asymmetry is obtained.

6.6 CHARACTERISTICS OF DIFFERENTIAL MEASUREMENTS

In this analysis, charge asymmetry is studied as a function of three observables: invariant mass of $t\bar{t}$ pair $m_{t\bar{t}}$, longitudinal boost of $t\bar{t}$ pair along the beam axis $\beta_z^{t\bar{t}}$, and transverse momentum of $t\bar{t}$ pair $p_T^{t\bar{t}}$.

Differential measurement of A_C is performed in 4 bins of $\beta_z^{t\bar{t}}$, 3 bins of $p_T^{t\bar{t}}$ and 5 (4) of $m_{t\bar{t}}$ differential observable for events in resolved (boosted) topology. Individual bins of each differential observable are specified as:

- $m_{t\bar{t}}$: $[0, 500, 750, 1000, 1500, \infty]$ GeV,
- $\beta_z^{t\bar{t}}$: $[0, 0.3, 0.6, 0.8, 1]$,
- $p_T^{t\bar{t}}$: $[0, 30, 120, \infty]$ GeV.

The lowest $m_{t\bar{t}}$ bin is removed at the boosted-selection step due to lack of statistics.

Different regions, which were defined in previous sections for the single-lepton and dilepton channel, are defined in the same way for each bin of a differential observable. The reconstructed distribution in any differential measurement consists of considerably larger number of bins. An illustration of the reconstructed spectrum for $p_T^{t\bar{t}}$ differential observable is shown in Fig. 35. The reconstructed distribution for any other differential observable is built accordingly.

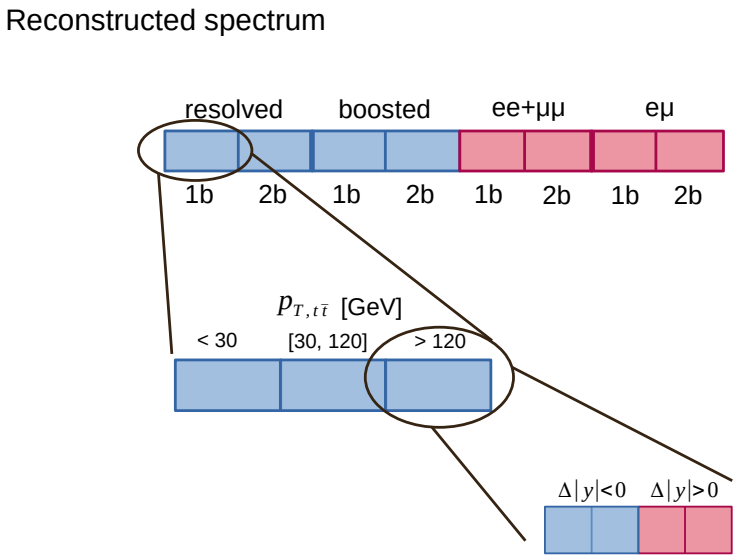


Figure 35: Illustration of the reconstructed distribution used in differential measurement of A_C as a function of $p_T^{t\bar{t}}$.

6.7 OPTIMIZATION OF THE FBU

The structure of the reconstructed distribution was already outlined in Fig.32 for the inclusive measurement and in Fig. 35 for a differential measurement. The $\Delta|y|$ binning in each region looks like $[-5, -x, 0, x, 5]$, where x denotes the inner bin edge. The binning is symmetrical around 0. Outer bin edges are basically identical with $\pm\infty$ because there are no events with $\Delta|y|$ larger than 5 or smaller than -5.

The optimization of the inner bin edges of $\Delta|y|$ distribution is performed in order to get unbiased response of the unfolding procedure. If the binning is not optimized then the method could be unstable and it could not unfold the spectrum correctly.

NOTE: More difficult situation occurs in the case of differential asymmetry measurements. The charge asymmetry value is determined in various bins of the differential observable (see Fig. 35), so binning optimization have to be performed in all differential bins of the differential observable simultaneously, because they influence each other. Change of the bin edge in $\Delta|y|$ distribution in one differential bin can affect migrations between differential bins, hence change the charge asymmetry value.

For each tested binning, the input (true) asymmetry and unfolded asymmetry are compared. The level of bias of the unfolding response is checked by performing the comparison. Besides the distribution corresponding to the SM charge asymmetry prediction, another 8 distributions with different injected charge asymmetry values are unfolded. These alternative samples are obtained from nominal $t\bar{t}$ signal sample, where each event is re-weighted. The re-weighting is based on samples simulated by PROTOS generator. The ratio of alternative and nominal sample is used in the re-weighting. The injected different values of A_C, axigluon-induced asymmetries of $\pm 1\%$, $\pm 2\%$, ± 3 and $\pm 4\%$, are considered.

In samples generated with PROTOS, the $t\bar{t}$ production is mediated via an axigluon.

The unfolding for each injected charge asymmetry value is repeated 300 times, which helps to get more precise parameters of the calibration curve. The Fig. 36 shows the result for the optimal bin edge $x = 0.5$ for the inclusive charge asymmetry measurement in the single-lepton channel.

The binning optimization is performed using data sets without systematic uncertainties and separately for the single-lepton and dilepton channel. After the optimization, $\Delta|y|$ bin edges in the single-lepton channel are listed in Tab. 7.

The same binning is chosen for particular $\Delta|y|$ true spectra.

The impact of systematic uncertainties on FBU response is checked with the most optimal binning. The effect is found to be insignificant. Hence, it is not required to do the optimization of binning again, with systematic uncertainties included. The parameters of the calibration curves obtained in individual differential bins and in inclusive measurement for the single-lepton channel with systematic uncertainties included are summarized in Tab. 8.

The linearity tests described above are run again with the single-lepton and dilepton samples combined with the most optimal binning. In our case, the same binning is used for the single-lepton and dilepton distribution, although in principle, the binning in each of the channels can be completely different. However, the statistics

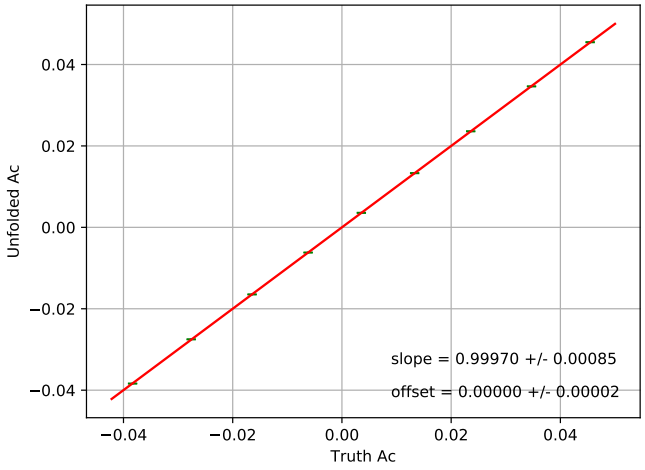


Figure 36: Result of the linearity test for the optimal bin edge $x = 0.5$ for the inclusive A_C measurement in the single-lepton channel. Green points in the plot show the unfolded charge asymmetry values for different true A_C values. The points are fitted with linear function - red line. The linearity test is repeated 300 times so the uncertainties of these points are negligible.

of the single-lepton channel highly outweigh the dilepton channel, hence the binning should be primarily adapted for the single-lepton channel. This "single-lepton-based" choice applied also to the dilepton channel causes negligible effects in combination results.

The same tests for combined channels are executed after the inclusion of systematic uncertainties. The response of the FBU is stable for the combined channels with systematic uncertainties as it is for the single-lepton channel only. The summarized results for the combined channels with the effect of systematic uncertainties are in Tab. 9. From the linearity tests, we assume that the single-lepton channel plays more significant role in combination than the dilepton channel and the inclusion of systematic uncertainties has very little effect on the FBU response.

The ideal parameters of the calibration curve are 1 for the slope and 0 for the offset. Any deviation from these values has to be included as an uncertainty of the unfolded result due to the imperfection of unfolding response. Potential bias, which the unfolding can cause, is evaluated as a difference between unfolded value and true A_C value calculated from the calibration curve:

$A_{C,true}$ *is derived from* $A_{C,unf} = slope * A_{C,true} + offset.$

$$\sigma_{bias} = A_{C,unfolded} - (A_{C,unfolded} - offset)/slope. \tag{40}$$

The estimated bias of the unfolding calculated from the Asimov data set and injected 5% charge asymmetry is listed in the Tab. 10 for the single-lepton channel and in Tab. 11 for combination of the single-

Table 7: Optimized binning in $\Delta|y|$ distribution for inclusive and differential measurements. The same binning is used for dilepton and single-lepton regions.

	$\Delta\|y\|$ binning $[-5, -x, 0, x, 5]$
inclusive	0.5
$m_{t\bar{t}} \in [0, 500]$ GeV	0.4
$m_{t\bar{t}} \in [500, 750]$ GeV	0.6
$m_{t\bar{t}} \in [750, 1000]$ GeV	1.0
$m_{t\bar{t}} \in [1000, 1500]$ GeV	0.9
$m_{t\bar{t}} \in [1500, \infty]$ GeV	0.9
$p_T^{t\bar{t}} \in [0, 30]$ GeV	0.6
$p_T^{t\bar{t}} \in [30, 120]$ GeV	0.5
$p_T^{t\bar{t}} \in [120, \infty]$ GeV	0.4
$\beta_z^{t\bar{t}} \in [0.0, 0.3]$	0.3
$\beta_z^{t\bar{t}} \in [0.3, 0.6]$	0.3
$\beta_z^{t\bar{t}} \in [0.6, 0.8]$	0.5
$\beta_z^{t\bar{t}} \in [0.8, 1.0]$	0.7

lepton and dilepton channel. Overall, the FBU procedure is not significantly biased.

Table 8: Results from linearity tests for the single-lepton channel with all systematic uncertainties included. The tests are repeated 300 times to obtain precise results.

	Protos reweighting	
	slope	offset
inclusive	1.000 ± 0.001	$< 0.0001 \pm < 0.0001$
$m_{t\bar{t}} \in [0, 500]$ GeV	1.002 ± 0.002	0.0002 ± 0.0001
$m_{t\bar{t}} \in [500, 750]$ GeV	0.983 ± 0.001	$-0.0001 \pm < 0.0001$
$m_{t\bar{t}} \in [750, 1000]$ GeV	0.997 ± 0.003	0.0005 ± 0.0001
$m_{t\bar{t}} \in [1000, 1500]$ GeV	0.993 ± 0.004	-0.0005 ± 0.0001
$m_{t\bar{t}} \in [1500, \infty]$ GeV	1.011 ± 0.015	-0.0007 ± 0.0005
$p_T^{t\bar{t}} \in [0, 30]$ GeV	1.006 ± 0.002	-0.0001 ± 0.0001
$p_T^{t\bar{t}} \in [30, 120]$ GeV	1.003 ± 0.002	-0.0001 ± 0.0001
$p_T^{t\bar{t}} \in [120, \infty]$ GeV	0.991 ± 0.004	0.0000 ± 0.0001
$\beta_z^{t\bar{t}} \in [0.0, 0.3]$	1.010 ± 0.006	-0.0003 ± 0.0001
$\beta_z^{t\bar{t}} \in [0.3, 0.6]$	1.005 ± 0.002	$< 0.0001 \pm 0.0001$
$\beta_z^{t\bar{t}} \in [0.6, 0.8]$	1.011 ± 0.002	-0.0001 ± 0.0001
$\beta_z^{t\bar{t}} \in [0.8, 1.0]$	0.993 ± 0.002	0.0002 ± 0.0001

Table 9: Results from linearity tests for the combination of the single-lepton and dilepton channel with all systematic uncertainties included. The tests are repeated 100 instead of 300 times due to very large computing time.

	Protos reweighting	
	slope	offset
inclusive	1.002 ± 0.001	$< 0.0001 \pm < 0.0001$
$m_{t\bar{t}} \in [0, 500]$ GeV	1.005 ± 0.002	0.0001 ± 0.0001
$m_{t\bar{t}} \in [500, 750]$ GeV	0.989 ± 0.001	$-0.0001 \pm < 0.0001$
$m_{t\bar{t}} \in [750, 1000]$ GeV	1.003 ± 0.003	0.0005 ± 0.0001
$m_{t\bar{t}} \in [1000, 1500]$ GeV	0.998 ± 0.004	-0.0005 ± 0.0001
$m_{t\bar{t}} \in [1500, \infty]$ GeV	1.018 ± 0.0014	-0.0003 ± 0.0005
$p_T^{t\bar{t}} \in [0, 30]$ GeV	1.008 ± 0.002	$< 0.0001 \pm 0.0001$
$p_T^{t\bar{t}} \in [30, 120]$ GeV	1.006 ± 0.002	-0.0001 ± 0.0001
$p_T^{t\bar{t}} \in [120, \infty]$ GeV	0.991 ± 0.004	-0.0003 ± 0.0001
$\beta_z^{t\bar{t}} \in [0.0, 0.3]$	1.014 ± 0.006	-0.0004 ± 0.0001
$\beta_z^{t\bar{t}} \in [0.3, 0.6]$	1.007 ± 0.002	0.0002 ± 0.0001
$\beta_z^{t\bar{t}} \in [0.6, 0.8]$	1.014 ± 0.002	-0.0002 ± 0.0001
$\beta_z^{t\bar{t}} \in [0.8, 1.0]$	0.997 ± 0.002	0.0001 ± 0.0001

Table 10: The relative bias of the unfolding of the single-lepton channel is calculated for the Asimov charge asymmetry values and for $A_C^{t\bar{t}} = 5\%$. Uncertainty $\sigma(A_C^{t\bar{t}})$ used for calculation of relative bias is the uncertainty of unfolded Asimov charge asymmetry.

	$\frac{\lvert \text{bias}(A_C^{t\bar{t}} = \text{SM}) \rvert}{\sigma(A_C^{t\bar{t}})}$	$\frac{\lvert \text{bias}(A_C^{t\bar{t}} = 5\%) \rvert}{\sigma(A_C^{t\bar{t}})}$
inclusive	< 0.01	0.01
$m_{t\bar{t}} \in [0, 500]$ GeV	0.05	0.07
$m_{t\bar{t}} \in [500, 750]$ GeV	0.07	0.41
$m_{t\bar{t}} \in [750, 1000]$ GeV	0.08	0.05
$m_{t\bar{t}} \in [1000, 1500]$ GeV	0.06	0.10
$m_{t\bar{t}} \in [1500, \infty]$ GeV	0.02	< 0.01
$p_T^{t\bar{t}} \in [0, 30]$ GeV	< 0.01	0.06
$p_T^{t\bar{t}} \in [30, 120]$ GeV	0.02	0.02
$p_T^{t\bar{t}} \in [120, \infty]$ GeV	< 0.01	0.07
$\beta_z^{t\bar{t}} \in [0.0, 0.3]$	0.05	0.05
$\beta_z^{t\bar{t}} \in [0.3, 0.6]$	< 0.01	0.07
$\beta_z^{t\bar{t}} \in [0.6, 0.8]$	0.03	0.11
$\beta_z^{t\bar{t}} \in [0.8, 1.0]$	0.02	0.04

Table 11: The relative bias of the unfolding of combined channels is calculated for the Asimov charge asymmetry values and for $A_C^{t\bar{t}} = 5\%$. Uncertainty $\sigma(A_C^{t\bar{t}})$ used for calculation of relative bias is the uncertainty of unfolded Asimov charge asymmetry.

	$\frac{\lvert \text{bias}(A_C^{t\bar{t}} = \text{SM}) \rvert}{\sigma(A_C^{t\bar{t}})}$	$\frac{\lvert \text{bias}(A_C^{t\bar{t}} = 5\%) \rvert}{\sigma(A_C^{t\bar{t}})}$
inclusive	0.01	0.07
$m_{t\bar{t}} \in [0, 500]$ GeV	0.04	0.11
$m_{t\bar{t}} \in [500, 750]$ GeV	0.05	0.29
$m_{t\bar{t}} \in [750, 1000]$ GeV	0.10	0.12
$m_{t\bar{t}} \in [1000, 1500]$ GeV	0.07	0.08
$m_{t\bar{t}} \in [1500, \infty]$ GeV	0.01	0.02
$p_T^{t\bar{t}} \in [0, 30]$ GeV	0.02	0.09
$p_T^{t\bar{t}} \in [30, 120]$ GeV	0.02	0.07
$p_T^{t\bar{t}} \in [120, \infty]$ GeV	0.04	0.11
$\beta_z^{t\bar{t}} \in [0.0, 0.3]$	0.07	0.06
$\beta_z^{t\bar{t}} \in [0.3, 0.6]$	0.05	0.14
$\beta_z^{t\bar{t}} \in [0.6, 0.8]$	0.06	0.14
$\beta_z^{t\bar{t}} \in [0.8, 1.0]$	0.02	0.01

7

SYSTEMATIC UNCERTAINTIES: HOW TO CORRECT FOR IMPERFECTIONS

Systematic uncertainties affecting the measurements have different origins. One group are experimental uncertainties, the others come from imperfections of MC simulations and of course, the statistical methods used for data analysis have their uncertainties themselves. Systematic uncertainties are represented in the measurement in a form of nuisance parameters (NPs), which impact the likelihood function and consequently the whole result.

One of method-based uncertainties, which was already mentioned, is the bias of the unfolding.

The effect of each systematic uncertainty is defined by comparing systematically shifted reconstructed distribution with the nominal distribution. The difference between these two distributions is considered as a systematic variation (systematic uncertainty). In case of systematic uncertainties corresponding to the signal modelling, these shifted spectra are derived by using systematically altered response matrix.

The effect of a two-sided systematic variation is symmetrized and the same size of the variation is taken as an upward and downward variation, just with the opposite sign

$$\sigma_{sym} = \pm\frac{1}{2}|\sigma_{up} - \sigma_{down}|. \tag{41}$$

7.1 EXPERIMENTAL UNCERTAINTIES

Experimental uncertainties are connected with reconstruction of leptons, jets and their properties, as well as reconstruction of missing transverse energy.

LEPTONS For every step leading to fully reconstructed electrons, like trigger, reconstruction, identification and isolation, specific correction factors are derived to correct for selection efficiencies. The total efficiency is obtained as a multiplication of efficiencies of individual procedures

$$\epsilon_{total} = \epsilon_{EMcluster} \times \epsilon_{reco} \times \epsilon_{id} \times \epsilon_{iso} \times \epsilon_{trig}. \tag{42}$$

The efficiencies for all mentioned procedures are estimated using the tag-and-probe method on $J/\psi \to \ell\bar{\ell}$ and $Z \to \ell\bar{\ell}$ events. The dilepton pair is selected according to the invariant mass value. One of them should pass strict requirement (tag) and the other one is a probe. Then, by applying the requirement also to the probes, the efficiency for corresponding procedure is evaluated.

Well-known resonances , such as decay of Z boson or J/ψ meson are used to estimate the efficiencies.

Due to differences between data and MC simulations, the scale factors are applied to MC samples in order to imitate data as closely as possible. Events in the simulations are weighted by the ratio of the efficiency obtained in data and in the simulation

$$w_{corr} = \frac{\epsilon_{data}}{\epsilon_{MC}} \sim 1. \tag{43}$$

Finally, the systematic uncertainties of these data-MC correction factors are estimated by varying the respective requirements of tag- and probe-lepton selection [71].

Analogously, differences of the energy scale and resolution between MC simulation and data are present. Therefore, corresponding corrections dependent on η are defined. The α parameter is used to correct energy scale in data [73, 133]

$$E^{data,corr} = E^{data}/(1+\alpha_i), \tag{44}$$

where i denotes individual regions in η. On the other hand, the correction factor c for energy resolution is applied to the MC simulation

$$\left(\frac{\sigma_E}{E}\right)^{MC,corr} = \left(\frac{\sigma_E}{E}\right)^{MC} \bigoplus c_i. \tag{45}$$

where i denotes individual regions in η and $\bigoplus$ means a sum of terms in quadrature. Again, for parameter derivations, events with the Z boson decaying to a pair of leptons are used. Both parameters, α_i and c_i, are extracted from comparison of the invariant mass distribution of lepton pair obtained from data and distribution simulated by generators.

Similarly, specific systematic uncertainties dependent on transverse energy E_T and η connected with the determination of energy scale and resolution are defined [133].

Previously described uncertainties are applicable to electrons. Similar procedures are repeated for muons [74].

SMALL-R JETS, FLAVOR TAGGING [79]

Jets are very complex objects and their calibration is a process consisting of many steps. Naturally, each step of the calibration is potential source of uncertainty. Uncertainties connected with pile-up effects, dependence on the flavor of initial quark, selection criteria and event topology used for the calibration of the jet energy - all these effects are accounted for in the set of nuisance parameters estimating how they affect the measurement.

NOTE: Jets are calibrated using various in situ methods [79]. Some of them use the events with well-calibrated objects like leptons, Z boson, or photons, to correct the jet response by balancing the p_T of a jet against the secondary object in an event. Sometimes, even well-calibrated jet in the central region of the detector is exploited in the

calibration using dijet events. The energy correction obtained from this step is then extended also for jets in the forward region by another step of the calibration - η intercalibration.

The nuisance-parameters set comprises 125 terms covering all mentioned uncertainty sources together with the MC mis-modelling and statistical uncertainty. Many of the parameters come from the in situ calibration, 4 parameters describe pile-up effects, 2 stand for differences between quark- and gluon-initiated jets. Additionally, the set of NPs takes into account also the uncertainty connected with the punch-through of jets (when a jet passes the whole calorimeter and enters the muon detecting system) and single-particle response.

Quark-initiated jet consists of particles with higher p_T, on the other hand, gluon-initiated jets have softer particles and wider transverse spread.

Fortunately, it is not necessary to include all 125 NPs when performing measurement which utilizes jets. For 98 NPs which depend only on the p_T of a jet, an eigenvector decomposition of the covariance matrix is executed and the largest orthogonal terms are taken as new NPs. The covariance matrix of the remaining terms is again inspected to determine how many of them can be combined into one residual parameter. Reduced set of these NPs, which keeps all significant correlations among these parameters is derived, resulting in total number of 29 NPs for the jet energy scale (JES) [78,79].

It is important to evaluate the resolution of the jet energy (JER) not only for jets, but also for measurement of missing transverse momentum. The JER depends on the transverse momentum of jets and can be broken down to 3 terms describing different sources which affect the resolution:

$$\frac{\sigma(p_T)}{p_T} = \frac{N}{p_T} \bigoplus \frac{S}{\sqrt{p_T}} \bigoplus C. \tag{46}$$

The first term accounts for the electronic noise and pile-up contribution, N. The second term influences the JER as a $1/\sqrt{p_T}$ function and characterizes stochastic effects S. Lastly, there is a constant term, C, which is a constant function of the jet p_T.

JER is estimated separately in data and in the MC simulation and to compensate the differences, the smearing procedure needs to be applied. The uncertainties connected with the JER are incorporated into the measurement via smearing of jets by Gaussian function with width σ_{smear}. Hence, the one-standard-deviation of variation is defined as

$$\sigma^2_{smear} = (\sigma_{nom} + |\sigma_{NP}|)^2 - \sigma^2_{nom}, \tag{47}$$

where σ_{nom} is the nominal resolution (JER) of a sample [79]. If the resolution is better in data, no smearing is applied, hence conservative estimate is kept. If the simulation has better resolution than data, then the additional systematic uncertainty is defined as

$$\sigma_{NP} = \sigma^{data}_{nom} - \sigma^{MC}_{nom}. \tag{48}$$

Uncertainty of JER is described by eight nuisance parameters [134].

A nuisance parameter for jet vertex tagger (JVT) efficiency [80] describe two effects. One of them is the uncertainty for remaining pile-up contamination after the removal of pile-up events. The second part consists of the uncertainty coming from usage of different MC generator for decay of Z-boson to $\mu\mu$ and Z-boson to $t\bar{t}$ simulations.

Scale factors (SFs) describing efficiencies of the b-tagging algorithm (MV2c10) to identify b, c, light jets are defined by comparing the efficiencies in data and MC simulations. Uncertainties of the SFs are treated as uncorrelated NPs. There are 9 parameters for b-jets, 4 for c-jets and 5 for light jets [82, 135, 136]. Additional two parameters are needed to cover the uncertainties due to extrapolation of b-jet and c-jet efficiencies for events with high-p_T jets.

LARGE-R JETS In addition to JES and JER uncertainties, the jet mass scale (JMS) and resolution (JMR) together with other uncertainties for jet substructure variables like τ_{32} are considered. The uncertainty on JMR is derived from the widths of the W-jet mass peaks in $t\bar{t}$ events obtained in data and in MC simulation. Systematic variation for JES, JMS and substructure variables is derived from data/MC ratio of that particular variable. For both data and MC simulation, the variable (energy, mass or substructure variable) is obtained for calorimeter jets and for corresponding track-jets, which are well-calibrated and independent of calorimeter-based objects. Used track-jets are matched to calorimeter jets. The following double-ratio is used for derivation of their uncertainties [85, 137]:

$$\frac{data}{MC} = \frac{\left(x_{calo-jet}/x_{track-jet}\right)_{data}}{\left(x_{calo-jet}/x_{track-jet}\right)_{MC}}. \tag{49}$$

In the end, collection of 14 parameters are needed to describe uncertainties connected with the measurement of large-R jets.

MISSING TRANSVERSE ENERGY E_T^{miss} Estimation of E_T^{miss} is dependent on all reconstructed objects (like leptons and jets) in an event, hence uncertainties corresponding to each of them are propagated into the E_T^{miss} uncertainty, while all correlations are preserved. The uncertainties of remaining soft terms, which also contribute to the E_T^{miss} value, are also taken into account. Finally, three nuisance parameters are assigned to missing transverse energy scale and resolution uncertainties [86].

PILE-UP Pile-up related uncertainty is determined through measurement of the luminosity in data [138]. The measurement creates the base for re-weighting algorithm used in MC simulations to match the pile-up measured in data. The uncertainty of pile-up contribution is given by varying the derived correction factors within their uncertainties.

LUMINOSITY The estimation of the luminosity uncertainty is 1.7% [138] for data collected in $2015 - 2018$. The measurement was performed by special detector LUCID-2 [139] designed for luminosity measurement.

Summary table of all experimental uncertainties with number of nuisance parameters assigned to them is given in Tab. 12.

Table 12: Summary table of the number of nuisance parameters used to describe experimental uncertainties.

Experimental uncertainties connected with:	number of NPs
Leptons	19
Small-R JES	29
Small-R JER	8
b-tagging	20
Large-R jets	14
JVT	1
Pile-up	1
E_T^{miss}	3

7.2 MODELLING UNCERTAINTIES

The simulation of collision of particles is a very complex process and it requires to choose the values of its parameters that exhibit a certain freedom of choice. Most of modelling uncertainties are defined though variations of these parameters. There are several different MC generators which differ to a small extent in their predictions. Hence, the source of modelling uncertainty also arises from the choice of the MC generator and its settings. Therefore, most of the modelling uncertainties are evaluated by comparison of two MC generators or by comparing the prediction given by different settings of the same MC generator.

7.2.1 *Signal modelling*

RENORMALIZATION AND FACTORIZATION SCALE The inevitable part of the matrix-element calculation is the choice of the factorization scale μ_f, which controls the scale at which the parton distribution functions are evaluated, together with the renormalization scale μ_r, which defines the scale at which the coupling constants are determined. The uncertainty given by this choice is estimated by compari-

son of the samples simulated with different values of μ_f and μ_r scales. More specifically, two alternative, systematically varied samples are produced for each scale variation. The value of the scale (μ_f or μ_r) is doubled in the first sample and halved in the second.

h_{damp} PARAMETER Another elementary parameter used in the calculation of the process is the h_{damp} parameter, which controls the high p_T emission at the NLO of the calculation. Hence again, the modelling uncertainty corresponding to different values of this parameter is estimated via comparison of the nominal sample and the sample simulated with h_{damp}-parameter value increased to $3m_t$.

α_S IN PARTON SHOWER The modelling uncertainty which accounts for the variations in the strong coupling constant in the simulation of the parton shower, more specifically, in the initial phase of the parton shower, is determined via variations of the Var3c parameter in the A14 tune [104] of the nominal MC generator.

FINAL-STATE RADIATION The uncertainty connected with the estimation of the FSR is determined via variation of the renormalization scale by factor of 2.0 and 0.5 (similarly as for determination of its uncertainty) which influences the QCD emissions, hence also final-state radiation. Technically, the variation of the renormalization scale is implemented through variation of the parton shower weights in the nominal MC generator.

PARTON SHOWER AND HADRONIZATION The uncertainty for the specific choice of the parton shower and hadronization modelling together with the simulation of the underlying events is estimated via comparison of two signal $t\bar{t}$ predictions simulated by different MC generators - the nominal combination of generators is `POWHEG+PYTHIA 8` while the alternative one is `POWHEG+HERWIG 7` [105, 106].

PDF Another modelling uncertainty comes from the choice of the parton distribution functions. The uncertainty is described by 30 nuisance parameters derived using PDF4LHC15 prescription [140].

Signal modelling uncertainties are used in the unfolding without the normalization effect. The overall normalization in $\Delta|y|$ distribution is irrelevant, because the charge asymmetry A_C is defined as a relative ratio (Eq. 17), hence, normalization systematic shifts would get canceled. Therefore, only the shape of the systematic uncertainties matters.

Some modelling uncertainties affect the normalization only in individual regions, like in resolved/boosted $1b/2b$ regions in the single-lepton channel or in similar regions in the dilepton channel. To ac-

count for these effects, separate normalization nuisance parameters are introduced per each region, except one. The reason is that in the unfolding, the normalization of the resulting distribution is a free parameter, hence there has to be at least one degree of freedom of the system to allow the overall normalization to float without constraints. The choice of the region without the normalization NP is arbitrary.

The regions with the normalization NP are: 1b boosted, 2b boosted, 2b resolved, 1b $ee + \mu\mu$, 2b $ee + \mu\mu$, 1b $e\mu$, 2b $e\mu$.

Technically, the signal normalization NPs are implemented as any other NP, but with a flat (uniform) prior probability density. The range of the prior distribution should be wide enough to ensure the freedom of movement for the NPs. Thus, the interval is set to $[-5\sigma, 5\sigma]$. The value of σ for NPs in boosted regions and dilepton regions is assigned to 20%, otherwise to 5%.

The reason behind different values of σ in single-lepton regions lies in the well-known normalization mis-modelling in the boosted region [141]. Larger spread of the prior covers the discrepancy.

7.2.2 *Background modelling*

Implementation of the background modelling uncertainties in the unfolding is the same as for other uncertainties (see Eq. 36), hence for each uncertainty source there is a NP affecting the reconstructed distribution constrained by its Gaussian prior term. However, treatment of the background normalization NPs have to be altered. The prior distribution for these background normalization NPs is truncated in order to prevent pathological behavior - the unfolding producing negative yields of the background processes.

Numbers of all background normalization NPs used in the measurement are listed in Tab. 13.

$W+$JETS The process, which contributes substantially to the single lepton background processes in the $1b$-exclusive region, is the production of W boson with additional jets. The modelling uncertainty of the shape and acceptance for this process is determined by variations of $W+$jets production [142]. The normalization uncertainty of this background process is set to 5% according to the measurement of W boson cross section [143].

Soft-gluon resummation term covers the soft gluon radiation contribution to the calculation of the matrix element. Usually the radiation of soft gluons gives rise to either large corrections or singular terms. Proper handling of the calculation and subsequently simulation is important.

Analogously as for the modelling of signal, the uncertainties corresponding to the different choice of μ_f and μ_r scale are considered. Their values are estimated by using nominal sample with altered dedicated MC weights. The choice of the soft-gluon resummation (QSF) scale is the other source of modelling uncertainty, together with the uncertainty for the matching of the matrix element with the parton shower, governed by CKKW matching procedure [119,144]. Again the corresponding uncertainties are estimated by variations of the MC weights. In case of ME+PS matching uncertainty, the CKKW scale is varied by factor of 0.75 and 1.5. For QSF uncertainty, the alternative samples are produced via variations of the particular scale by 0.5 and 2.

The uncertainty of different jet flavor modelling is determined separately for $W + b/c$/light-jets. The normalization uncertainties for each sample ($W + b/c$/light-jets) is set to $\pm 30\%$.

SINGLE TOP The single-top production has major impact in $2b$-inclusive region in the single-lepton channel and it is the dominant background process of the dilepton channel. According to the calculation [145,146], the normalization uncertainty of the single-top production is set to 5.3%. There are three production channels of the single top: the main contribution is given by Wt channel, the other processes are s- and t-channels. Single-top quark tW production at NLO in perturbative QCD shares final states with LO $t\bar{t}$ production, therefore an interference between these two processes occurs. The double counting is removed by applying the DR scheme. The uncertainty derived by comparing DR and DS scheme is considered [99].

The uncertainty for the choice of PS and hadronization model is estimated similarly as for the $t\bar{t}$ signal. Radiation in the initial and final state is controlled by the renormalization and factorization scale parameters. More exactly, the variation of both scales in the ME simulation causes the variations in the initial-state radiation. Both scales are multiplied by 0.5 or 2 to increase/decrease the rate of the ISR and to give the estimation of the corresponding uncertainty. The final-state radiation is regulated by the value of the renormalization scale in the PS. The uncertainty for the FSR modelling is determined via variations of μ_r by 0.5 or 2 in the PS.

Z+JETS In the $ee + \mu\mu$ channel of the dilepton decay channel, the Z+jets process contributes to the large extent. As for other background/signal samples, the uncertainty for the choice of μ_f and μ_r scale is obtained by the variation of MC weights. The normalization uncertainty of the Z+jets process is assigned to 30% in the dilepton channel. In the single-lepton channel, Z+jet background is merged together with other background processes.

DIBOSON Diboson production is significant only in the dilepton channel. The uncertainty of the diboson cross section (normalization NP) is 6% in the dilepton channel according to [147]. In the single-lepton channel, diboson background is merged together with other background processes.

NON-PROMPT AND FAKE LEPTONS The estimation of the non-prompt and fake leptons is given by data-driven estimate of the real and fake lepton efficiencies. The uncertainty of the procedure is derived by using different parameterization and comparing the results. The shape and normalization uncertainty are treated separately. The uncertainty of the shape is given by previous method (by scaling the

alternative fake estimation to the nominal distribution) and the normalization uncertainty is set to conservative 50% in the single-lepton channel. In the dilepton channel, the same values of normalization uncertainty are assigned to regions with $\mu\mu$ fake-lepton events. For ee and emu regions, the normalization uncertainty is 30%. The effect of the fake and non-prompt lepton background is uncorrelated among the regions. NPs accounting for shape uncertainty are estimated separately for events with electrons and with muons.

OTHER BACKGROUND PROCESSES In the single lepton channel, background processes which contribute to a very small extent, like $Z+$ jets, diboson production, $t\bar{t}V$ and $t\bar{t}H$ are combined together. The normalization uncertainty is fixed to conservative value of 50%. Very rare processes of the dilepton channel, like $t\bar{t}H$ and $t\bar{t}V$ are also merged together. In this case, 13% normalization uncertainty is used [148].

Table 13: Summary table of the number of nuisance parameters used to describe normalization uncertainties of background processes. If the modelling of a background process is different among regions, individual uncorrelated NP is introduced per each region.

channel	background	de-correlation	# of NPs
Single-lepton	$W+$jets	none	1
	Single top	none	1
	Other bckg	none	1
	fake & non-prompt leptons	resolved/boosted $1b/2b$	4
Dilepton	$Z+$jets	$ee + \mu\mu/e\mu$ $1b/2b$	4
	Diboson	none	1
	Rare SM	none	1
	fake & non-prompt leptons	$ee/\mu\mu/e\mu$ $1b/2b$	6

7.2.3 *Gamma factors*

The compensation of the limited size of the MC background samples is provided by the addition of gamma factors [149] in the unfolding. For each bin of the distribution, a gamma parameter is introduced. It is expected to allow the total yield of the background to float around its predicted value. The background yield is modified as $B_i = \gamma_i B_i$. The values of the gamma parameters are constrained by Poisson prior

probability density terms with the mean value defined by the predicted total background yield in the bin $B_i(0)$ and its statistical uncertainty $\sigma_{stat}(B_i)$. The prior probability term is $P(\lambda_i|\gamma_i\lambda_i)$, where $\lambda_i = (B_i(0)/\sigma_{stat}(B_i))^2$. Thus, the most probable value of gamma parameters is unity and it is anticipated to oscillate around unity within the statistical uncertainty of the background bin yields.

7.3 THE BOOTSTRAP METHOD

It is a common problem that the estimation of the systematic uncertainties suffers from low MC statistics. What actually happens is that a systematic variation derived in ways described in the previous sections, has unphysical fluctuations in its distribution. The true nature of these effects is inspected by using the bootstrap method [150].

NOTE: In the charge asymmetry measurement, the shape of the systematic variation is what matters the most - it is what has the largest impact on the charge asymmetry value (see definition in Eq. 17). That is why it is vital to determine the distribution of any systematic variation to its most true self (remove unphysical effects).

The bootstrap method is a type of re-sampling technique. It uses a sampling distribution, from which an estimate of some quantity may be extracted, to generate another distributions in order to estimate the statistical uncertainty of this quantity. The method exploits the assumption, that the sampling distribution sufficiently describes the data set with infinite statistics, from which we would obtain the real uncertainty. Hence, the estimate of real statistical uncertainty of a quantity derived by re-sampling of a nominal sampling distribution is adequate.

A distribution of simulated systematic variation is perceived as a sampling distribution of the real unknown distribution of the systematic variation. Generation of the bootstrap samples is performed via drawing events with altered weights from the 'nominal' systematic distribution, the sampling distribution. The weights are drawn from the Poisson distribution with mean value of unity. This procedure creates the whole ensemble of the potential distributions of the systematic variation.

Taking the weight from the Poisson distribution can be imagined as drawing the same event 0 times or once, twice, ... with the probability given by the Poisson distribution.

NOTE: Systematic uncertainty is determined as a relative difference of the systematically shifted sample and the nominal sample, if the systematic effect is one-sided. In case of two-sided effect, there is an upward variation and downward variation of the nominal distribu-

tion, so the systematic uncertainty is calculated as a relative difference between two systematic shifts

$$\sigma_{up} = \frac{S_{up} - S_{nom}}{S_{nom}}, \qquad \sigma_{down} = \frac{S_{nom} - S_{down}}{S_{nom}}, \tag{50}$$

where $S_{up,down}$ denote an upward/downward systematically shifted spectrum and S_{nom} denotes nominal distribution. The systematic uncertainty is symmetrized afterwards, see Eq. 41.

Therefore, for the estimation of the statistical power of each systematic uncertainty, the bootstrap replicas are generated also for the nominal distribution in a correlated way with the replicas of the systematically shifted sample. The correlations between nominal and systematically shifted sample are conserved via the same initialization seed of the weight generation of the same event. Hence, if the same event appears in both samples, the same weight is assigned to it in both distributions.

In case of two-sided systematic variation, the bootstrap replicas are drawn for both systematically shifted samples (upward, downward variation) and for the nominal sample.

The goal of the bootstrap method is to estimate the precision of systematic uncertainty. In other words, the statistical uncertainty of the systematic variation is obtained. Per each set of replicas (nominal, systematically shifted), the systematic variation is calculated according to the definition in Eq. 50 (for two-sided variations, both σ_{up}, σ_{down}, otherwise σ_{up}), for each bin of the distribution. The values of systematic uncertainty obtained from bootstrap replicas create a distribution itself (again one distribution per each bin). Hence, the statistical uncertainty of the systematic variation is given by the width of the bootstrap-defined distribution of systematic uncertainties. The bootstrap method is illustrated in Fig. 37.

The mean of the bootstrap-defined distribution for a systematic uncertainty should be the same as value of nominal systematic uncertainty (calculated from distributions used as an input to bootstrap procedure), when enough number of replicas are generated.

The number of bootstrap replicas should be large enough for proper determination of statistical significance of the systematic uncertainty. Thus, 500 replicas are drawn for each distribution (nominal, systematically shifted).

NOTE: Systematic uncertainties, which are obtained by re-scaling of nominal $t\bar{t}$ sample, generally do not undergo the bootstrap procedure. For this category of systematic uncertainties, the nominal and systematically shifted sample are fully correlated, hence the statistical uncertainties are expected to be small. However, for some modelling uncertainties obtained by variations of the MC weights, the systematic variations have large magnitude, thus the statistical uncertainty is not small either.

No application of the bootstrap method means these systematic uncertainties remain in the measurement.

7.4 PRUNING OF SYSTEMATIC UNCERTAINTIES

After the statistical uncertainty is estimated, the problem of potential unphysical behavior of some systematic variation can be investigated. The importance of the systematic uncertainty is inspected for

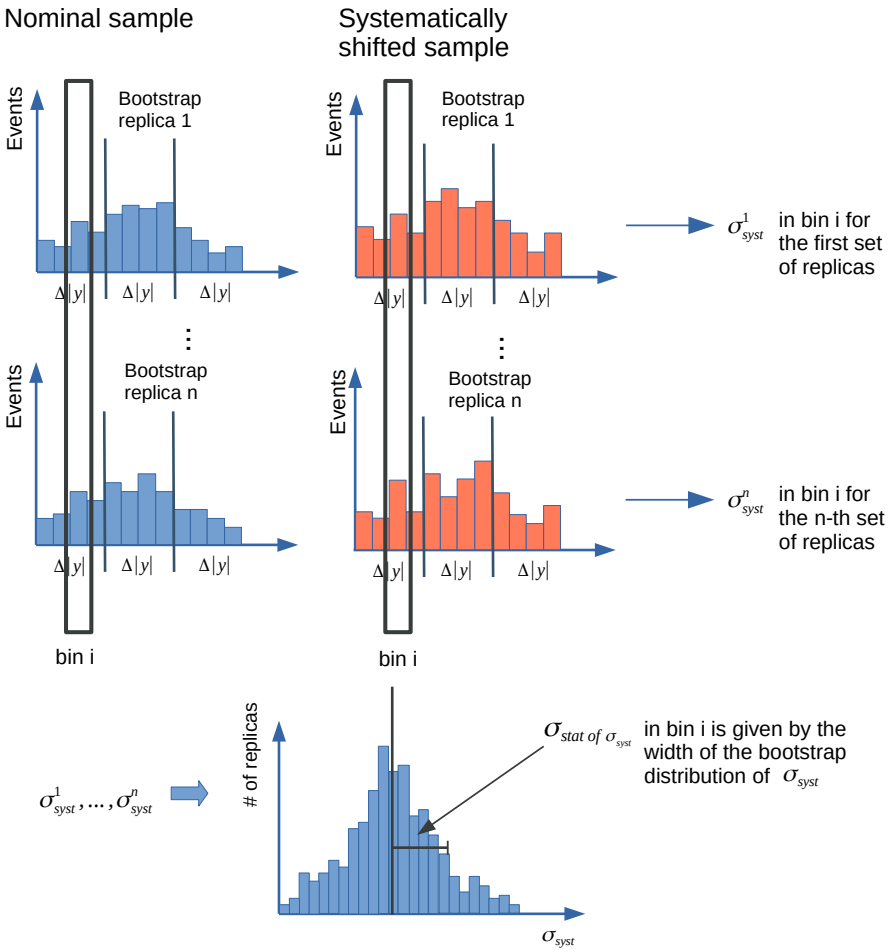


Figure 37: Graphic illustration of the bootstrap method.

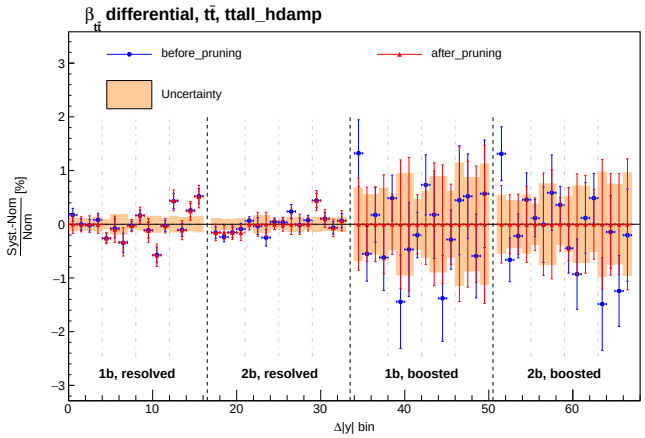
Number of default $\Delta|y|$ spectra (4 bins) in the reconstructed distribution is illustrated in Fig. 32 or in Fig. 35

each 4 $\Delta|y|$ bins in the distribution, because they define the value of charge asymmetry in each region. The systematic uncertainty is considered insignificant if its values in all 4 bins are consistent with zero within 2σ, where σ is the statistical uncertainty of the systematic variation derived by the bootstrap method. When a systematic variation fulfills the criterion, all 4 $\Delta|y|$ bins are merged together and their significance is checked again using the same test. The systematic uncertainty, which is insignificant also after the additional check, is dropped completely from consideration for this particular part of the distribution (inspected 4 $\Delta|y|$ bins). However, if the combined bin is statistically significant, then the value of the combined bin content is set to all 4 bins of $\Delta|y|$ spectrum. Obviously, if all 4 bins are statistically significant, the systematic variation remains unchanged.

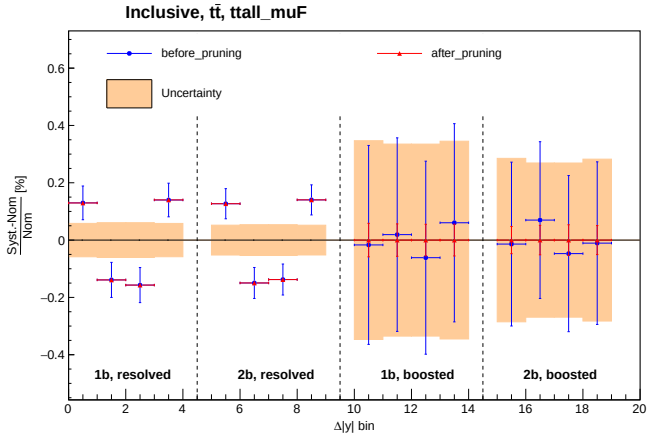
A demonstration of pruning procedure is shown in Fig. 38. For both displayed systematic variations (h_{damp} and μ_f), the systematic uncertainties in boosted regions are pruned. For h_{damp} parameter variation (Fig. 38a, 1.$\beta_z^{t\bar{t}}$ bin in resolved $2b$ region), the effect is found significant

only after all 4 $\Delta|y|$ bins are combined. Therefore, only the normalization effect of the uncertainty is kept.

Using 2σ criterion for four $\Delta|y|$ bins gives $1 - 0.95^4 = 19\%$ chance of insignificant systematic uncertainty to pass the criterion.



(a) h_{damp}, $\beta_z^{t\bar{t}}$ differential measurement



(b) μ_f, inclusive measurement

Figure 38: An example of pruning procedure. Systematic uncertainty for variation of h_{damp} parameter in $\beta_z^{t\bar{t}}$ differential measurement (a) is plotted before pruning (blue points) and after pruning (red points). In boosted regions, the systematic uncertainties are completely discarded by pruning. In first $\beta_z^{t\bar{t}}$ differential bin in $2b$ resolved region, only normalization effect of the variation is kept. Similarly, systematic effects of upward variation of μ_f scale is removed in boosted regions of inclusive measurement (b).

7.5 UNCERTAINTY OF THE RESPONSE MATRIX

The limited MC statistics results in reduced statistical precision of the response matrix. To incorporate this fact as an uncertainty on A_C value, pseudoexperiments with smeared response matrix are performed. In each pseudoexperiment, the Asimov dataset is unfolded with smeared response matrix. The width of the obtained A_C dis-

tribution (e.g. see Fig. 39) is taken as an uncertainty caused by limited statistics of the nominal $t\bar{t}$ sample. The uncertainty is added in quadrature to the uncertainty of A_C obtained by unfolding.

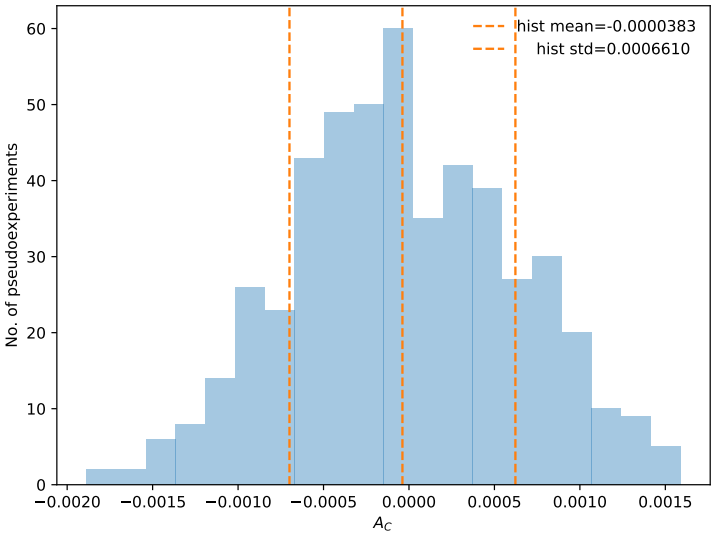


Figure 39: Posterior distribution of A_C from pseudoexperiments with smeared response matrix in the inclusive measurement.

7.6 DEALING WITH CONSTRAINTS OF SYSTEMATIC UNCERTAINTIES

The important information extracted from the FBU unfolding is the information about how the systematic uncertainty is pulled with respect to its nominal value and how much it is constrained after the marginalization performed by the FBU unfolding. The pull and constraint are defined by the marginal posterior distribution (see Sec. 6.5.3).

Typically, if a NP of some modelling systematic uncertainty is constrained too much ($> 30\%$ of its nominal value), it may signalize that the systematic uncertainty lacks statistics and therefore it may have a lot fluctuations, which does not go well with the systematic model and data. Hence, the effects of such systematic uncertainty is suppressed by larger constraints. Sometimes, the less severe reason lies behind the problem of large constraints. Simply, the systematic model may not have enough freedom, which means that a nuisance parameter is constrained by the one specific region but if there is just one NP per whole distribution, then the constraint influences also the whole distribution, which is not desirable. Therefore, in case of large constraints, various systematic de-correlation scenarios are inspected, each of them introduce more NPs for problematic systematic variations. One must not forget on potential correlations among NPs, which can also help to transfer constraints from one NP to another NP. In reality, it is hard to disentangle the effects of correlations and

Ideally, there should be no constraints - that happens when MC modelling gives flawless predictions, which agree with data.

individual constraints for a single systematic uncertainty, if one is interested in the source of a constraint. Usually, it is the combination of all mentioned reasons, each to some extent.

Systematic uncertainties, which suffer from large constraints, are predominantly modelling uncertainties. NPs related to the variations of h_{damp} parameter or μ_r, μ_f scales, variations of radiation in final-state or choice of PS+hadronization model are the most affected ones. For illustration, see constraints for 'default' setup in Fig. 40. Therefore, to deal with the constraints, various de-correlation scenarios of systematic uncertainties have been tested. Choice of the most suitable scheme was decided according to the results of the unfolding of Asimov data set. Also, the increasing complexity of the systematic model due to larger number of NPs complicates the marginalization process, thus unnecessary de-correlations are avoided. If de-correlation of a systematic NP is found to have almost no effect on the size of its constraint, it is not included in the final scenario. In a few cases, de-correlating of NPs across regions leads to significant reduction of constraints, which may cause the increase in total uncertainty by 5-20%. In such situation, the more conservative scheme is used. Generally, de-correlating of NPs has negligible impact on the unfolded value of A_C. Outcome of the tuning of systematic uncertainties' setup is the following de-correlation scheme:

DE-CORRELATION OF SIGNAL MODELLING UNCERTAINTIES

INCLUSIVE MEASUREMENT:

- Due to substantially different kinematics of boosted topology, NPs for $t\bar{t}$ h_{damp}, μ_r, μ_f, FSR in the single-lepton boosted region are de-correlated from other regions (single lepton resolved, dilepton).
- Analogously, NPs for $t\bar{t}$ PS+hadronization are de-correlated between single-lepton and dilepton regions, while only one NP is used for all dilepton regions. NPs in boosted and resolved single-lepton regions are separated.

DIFFERENTIAL MEASUREMENTS:

- Single-lepton boosted region is de-correlated from other regions for $t\bar{t}$ μ_r and μ_f scales variations.
- In case of NPs for $t\bar{t}$ PS+hadronization, h_{damp} and FSR, single-lepton resolved, boosted and dilepton regions are de-correlated. The effect of de-correlation on constraint of NP corresponding to variations of h_{damp} parameter is demonstrated in Fig. 40.
- In addition, $t\bar{t}$ PS+hadronization systematic uncertainty is de-correlated also between $1b$ and $2b$ regions in the single-lepton channel.

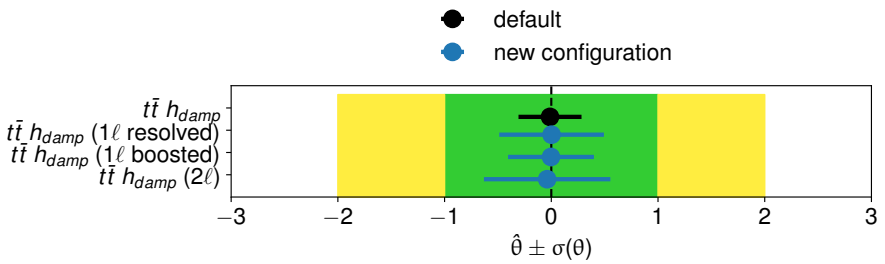


Figure 40: Constraints of NPs for variations of h_{damp} in $m_{t\bar{t}}$ differential measurement for two different configurations: one NP for all regions (default) and separate NPs for single-lepton resolved, boosted and dilepton region (new configuration). The Asimov data set is used.

DE-CORRELATION OF BACKGROUND MODELLING UNCERTAINTIES INCLUSIVE MEASUREMENT:

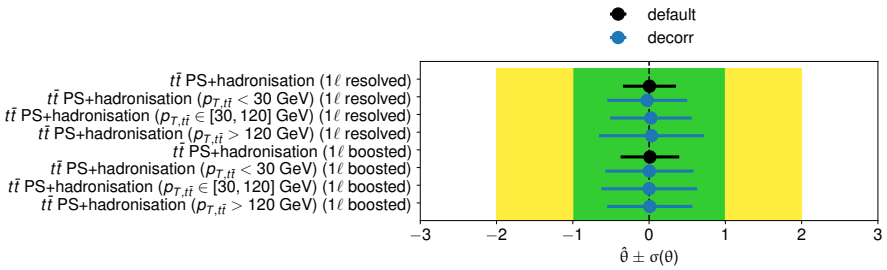
- NPs for FSR variations in single-top Wt-channel are de-correlated between single-lepton boosted region and other regions.

DIFFERENTIAL MEASUREMENTS:

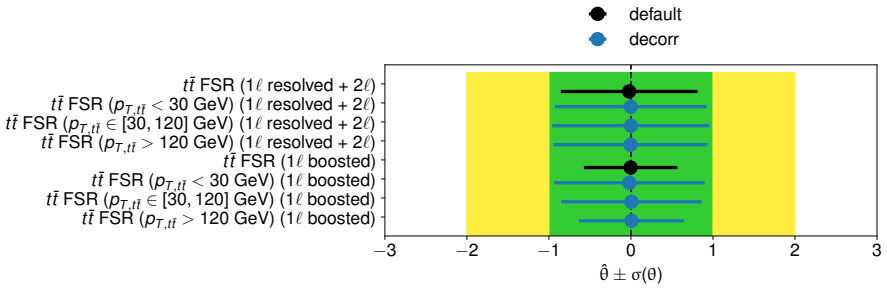
- Single-lepton resolved, boosted and dilepton regions are de-correlated for single-top FSR and single-top PS+hadronization systematic uncertainties.
- In single-lepton channel, W+jets μ_r and μ_f scale variations are de-correlated between resolved and boosted topology.
- Compositions of fake-lepton events differ a lot among the various regions. Thus, normalization uncertainty and shape uncertainty are de-correlated among all regions.
- Separate normalization NPs for the other background processes are introduced for single-lepton and dilepton regions.
- In dilepton channel, the Z+jets normalization uncertainty is de-correlated among all regions.

DE-CORRELATION BY DIFFERENTIAL BINS Additionally, for the differential measurement, there is also a possibility to add one set of NPs per each differential bin, hence multiply the number of parameters per such systematic uncertainty n-times, where n is number of differential bins. The procedure effectively terminates any correlations among differential bins for the investigated systematic uncertainty. This de-correlation is mostly exploited for the signal or background modelling uncertainties, which are considerably constrained despite their de-correlation between resolved/boosted and dilepton

regions, like systematic uncertainty connected with the choice of PS+hadronization model and final-state radiation (for both signal and background samples) or variations of h_{damp} parameter. The effect of de-correlation by differential bins on constraints of PS+hadronization and FSR NPs is demonstrated in Fig. 41.



(a) PS+hadronization



(b) FSR

Figure 41: Effect of de-correlation by differential bins demonstrated for constraints of PS+hadronization and FSR NPs in $p_T^{t\bar{t}}$ differential measurement.

RANKING OF SYSTEMATIC UNCERTAINTIES The significance of each systematic uncertainty is estimated by evaluation of its impact on the unfolded charge asymmetry value. The impact of a systematic uncertainty is determined for two different situations. Firstly, the effect of a systematic uncertainty with the magnitude given by MC simulation is inspected. The tested NP is set to $\theta = pull \pm \sigma$, where $pull$ is the shift of the NP given by previously performed unfolding with default settings. In this way, the impact of a systematic uncertainty is studied for its variation which is already shifted to the value preferred by the unfolded data. Next, the data set with the tested NP-value frozen to either $+\sigma$ or $-\sigma$ variation around the size of $pull$ is unfolded. All NPs are marginalized except the one whose effect on the A_C is tested. Hence, the charge asymmetry value obtained from the unfolding of modified data set declares how the tested NP is able to change the charge asymmetry. In the second scenario, the tested NP is frozen to the constrained size of the uncertainty, $\theta = pull \pm constraint$,

MC-simulated systematic variations = the size of the variation as it enters the unfolding

Unfolded data in this case could be either the Asimov data set or measured data.

so the systematic uncertainty as is determined by the unfolding is studied.

Systematic uncertainties are ordered according to the size of the impact they have on the unfolded charge asymmetry value. The difference of the unfolded asymmetry values, ΔA_C, obtained from the unfolding with default and modified data set is calculated for each tested NP. Both these methods, so called pre-marginalization (a NP frozen to $pull \pm \sigma$) and post-marginalization (a NP frozen to $pull \pm constraint$) ranking are performed for each systematic uncertainty and for both, measured data and the Asimov data set. In Fig. 42, results from ranking of NPs for Asimov data set corresponding to the inclusive A_C measurement are shown. The constraints and pulls of the ranked systematic uncertainties are displayed in the same Figure.

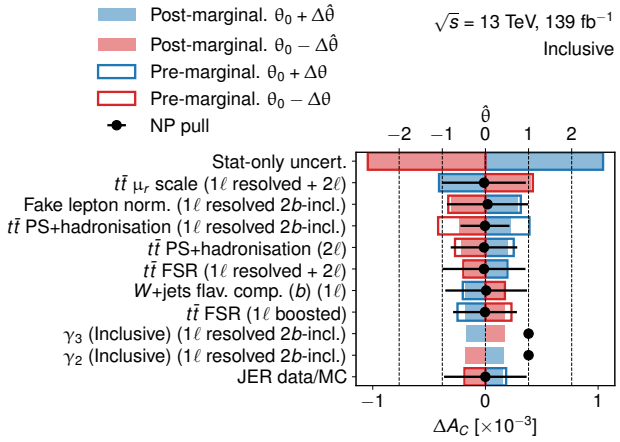


Figure 42: Ranking of the systematic uncertainties in the inclusive measurement for the Asimov data set. Top 10 highest ranked systematic uncertainties are shown in the plot. The order is defined by the size of the effect each of systematic uncertainty has on the A_C value by shifting the variations by their constraints (post-marginalization ranking). Blue and red bars show the effect of upward and downward systematic variations on the A_C value, respectively. The frames and filled bars denote the pre-marginalization and post-marginalization results of the ranking, respectively. Black data points display the constraints and pulls of the ranked systematic uncertainties.

The systematic uncertainties with the most asymmetric distribution are in general the highest in the ranking, although the size of the variation may be small in comparison with other variations.

7.7 TREATMENT OF THE SIGNAL MODELLING UNCERTAINTIES

The signal modelling uncertainties are mostly defined by comparison of the predictions given by different generators or the same generator but with different configuration. The systematic variation affecting

the i-th bin of a distribution is defined as a difference of the reconstructed distributions:

$$\Delta R_i = R_i^{syst} - R_i^{baseline}. \tag{51}$$

The uncertainties defined by the scale variations are simulated with the same generators, but with different values of their parameters.

where R^{syst} is reconstructed distribution for systematically shifted sample and $R^{baseline}$ is nominal reconstructed distribution. However, truth-level distributions, which are simulated by different generators or different configurations of a generator, do not coincide, $T^{baseline} \neq T^{syst}$. The Eq. 51 does take into account the differences between true spectra. In the charge asymmetry measurement, these effects are undesirable, because the reconstructed spectrum is unfolded to the parton level, hence these discrepancies cannot be accounted for. In our case, the whole systematic-uncertainty effects should be encoded in the response matrix M_{ij}^{syst}.

Remainder: reconstructed distribution (for signal) is obtained via folding the true distribution with the response matrix, $R_i = M_{ij}T_j$.

Therefore, the R_i^{syst} has to be corrected for the unwanted effects at the parton level. The event-by-event re-weighting based on the truth-level-prediction differences is applied. The weight is defined as ratio of these predictions, $T_i^{baseline}/T_i^{syst}$ to absorb the differences. The proper definition of the systematic uncertainty is therefore

$$\Delta R_i = R_i^{syst}\frac{T_i^{baseline}}{T_i^{syst}} - R_i^{baseline} = R_i^{syst,re\text{-}weighted} - R_i^{baseline}. \tag{52}$$

It is the same procedure as if the construction of the systematically shifted reconstructed spectrum would be given by the folding of the baseline true prediction with the systematically altered response matrix

$$\Delta R_i = R_i^{syst}\frac{T_i^{baseline}}{T_i^{syst}} - R_i^{baseline} \tag{53}$$

$$= M_{ij}^{syst}T_j^{syst}\,\frac{T_i^{baseline}}{T_i^{syst}} - R_i^{baseline} \tag{54}$$

$$= M_{ij}^{syst}T_j^{baseline} - R_i^{baseline}. \tag{55}$$

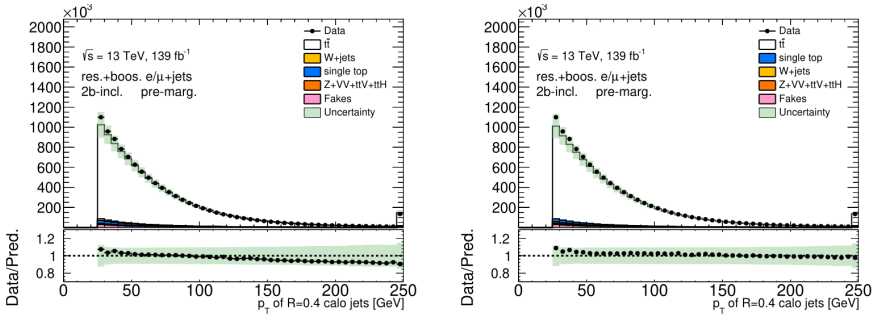
ALTERNATIVE: THE NNLO-BASED RE-WEIGHTING The true-level prediction of the top-quark p_T spectrum suffers from mis-modelling [151]. The MC re-weighting based on the latest, most precise calculation (NNLO in QCD + NLO in EW theory) of the $t\bar{t}$ production [152] can be applied to correct for the discrepancies observed between MC prediction and measured data. Application of such correction has the potential to correct the R_i^{syst} prediction without the use of the truth-level-based re-weighting.

This alternative approach has been tested in the single-lepton channel measurement. The TTbarNNLO Recursive Reweighter package has been implemented into the processing of MC simulated samples. This tool provides the re-weighting based on corrections to three variables: top-quark p_T, mass of $t\bar{t}$ and p_T of $t\bar{t}$. Each event in a spectrum

is re-weighted by the ratio of the prediction calculated at NNLO in QCD + NLO in EW theory and MC simulation. The ratio is derived specifically for the value of the variable (either top-quark p_T, $m_{t\bar{t}}$ or $p_T^{t\bar{t}}$) for that event. The re-weighting is recursive, which means that the resulting distribution is obtained after a finite number of iterations. In each iteration, one of the three variables is used to derive the weight. The order of variables, in which they are used for determination of the weight, is defined by the tool. The number of iterations is pre-defined as well - by imposing a condition that the differences between MC and data are lower than the NNLO-based uncertainty.

The NNLO-based re-weighting follows the same idea as the truth-based re-weighting: it makes the various truth-level distributions more alike.

The re-weighting is developed for the nominal combination of generators (`POWHEG+PYTHIA 8`), but also for another generators like `HERWIG7`. All truth-level predictions are corrected to the same NNLO precision. An application of the recursive NNLO re-weighting positively influences the agreement between the simulation and data, see Fig. 43 for lepton p_T. The improvement is evident mostly in the tail of the distribution. Similar behavior is observed also for other variables like jet p_T or E_T^{miss}.



(a) Without NNLO re-weighting (b) After NNLO re-weighting

Figure 43: Illustration of the data/MC agreement for p_T of a jet in the single-lepton channel with 2 or more b-tagged jets without the NNLO re-weighting (a) and after the NNLO re-weighting (b) is applied.

Of course, the calculation at the precision of NNLO in QCD + NLO in EW theory of $t\bar{t}$ process comes with the choice of various parameters, like PDFs, μ_r or μ_f. Therefore, the usage of this calculation in the re-weighting results in additional uncertainties affecting the measurement. Extra 3 new NPs, which account for variation of scales in the calculation of the NNLO prediction for p_T, $m_{t\bar{t}}$ and $p_T^{t\bar{t}}$, are introduced in the FBU. For the purpose of the test of the NNLO re-weighting, more NPs were not considered. However, proper implementation requires more of them.

The impact of the NNLO re-weighting on the particular signal modelling systematic uncertainties is illustrated by Fig. 44 and Fig.45. The uncertainty for the variations of the radiation in the final state is not affected by the NNLO re-weighting to a large extent. The difference

between the two variations are negligible. However, for example in the inclusive measurement, the PS and hadronization systematic uncertainty in the resolved topology is larger then before, suggesting that previously it was underestimated in this region. The opposite is true for the boosted region. On the other hand, the change in the variation of μ_R scale is more evident, predicting a smaller magnitude of the variations in most of the differential bins and for all measurements. The systematic variation of h_{damp} parameter in $p_T^{t\bar{t}}$ measurement is considerably changed. The variations are much smaller after the NNLO re-weighting.

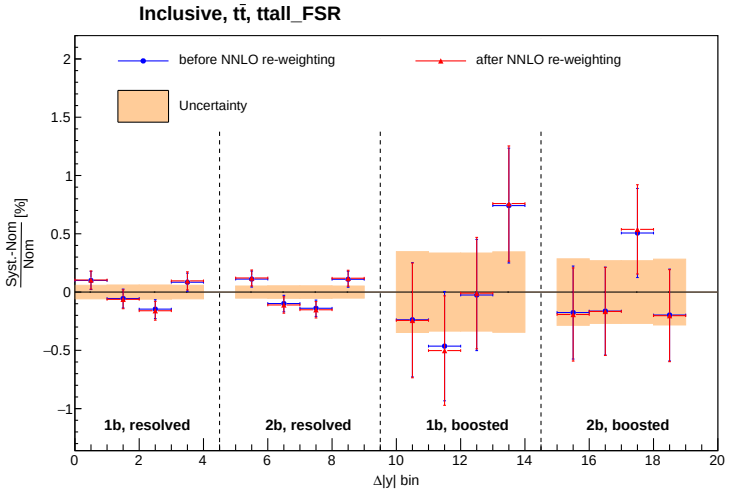
In most of the cases, the effect of the NNLO re-weighting varies among the measurements (inclusive, differential). Therefore, it is not possible to describe the effect of the re-weighting in general.

Different systematic variations subsequently affect the result of the unfolding due to different pulls and constraints. The Asimov data set is constructed using the NNLO re-weighted inputs and this data set is unfolded afterwards. The results are compared with the results obtained from the nominal Asimov data set. The unfolded charge asymmetry values together with the unfolded uncertainties for both setups are summarized in the Tab. 14. The largest difference in the unfolded uncertainty is observed in the differential bin corresponding to the largest values of $p_T^{t\bar{t}}$. The mis-modelling influences the most differential measurement as a function of $p_T^{t\bar{t}}$. The NNLO re-weighting in some cases helps to reduce the systematic uncertainties in this measurement, which may cause that their constraints are smaller. Due to this effect, higher unfolded uncertainties are observed.

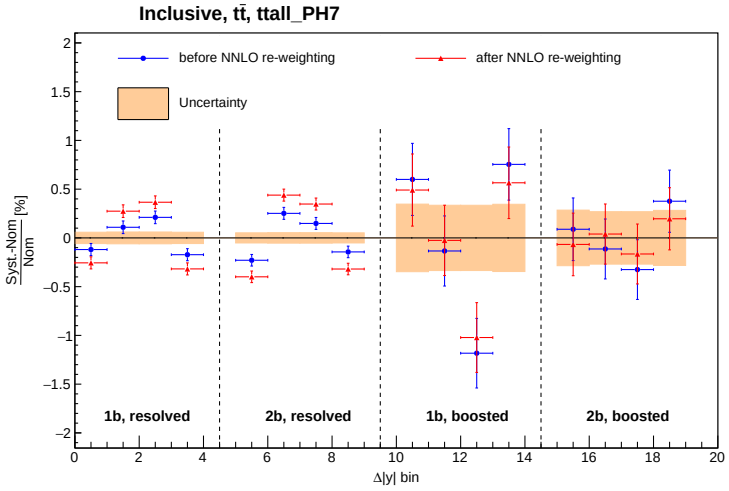
To treat everything consistently, the NNLO re-weighted signal distribution is used for derivation of the experimental uncertainties.

The change in constraints of the modelling systematic uncertainties for the Asimov data set is shown in the Figs. 46a,46b for $m_{t\bar{t}}$ and $p_T^{t\bar{t}}$ differential measurements, respectively. The same plots for inclusive and $\beta_z^{t\bar{t}}$ differential measurements are depicted in Figs. 68a and 68b in App. A.1, respectively.

In general, the NNLO re-weighting influences positively the size of the constraints. For example, in the differential measurement as a function of $p_T^{t\bar{t}}$, the constraint of the NP for h_{damp} variations in boosted region is zero (see Fig. 46b), because the new variation is small enough to fulfill the criteria set for pruning, see Fig. 45b. In resolved region, a little constraint for h_{damp} parameter is observed. On the other hand, constraint for PS+hadronization variation in resolved region is a little worse than before the NNLO re-weighting. For $m_{t\bar{t}}$ measurement, the constraints of h_{damp} parameters are improved the most. However, they are still constrained quite a lot. Analogously, in the $\beta_z^{t\bar{t}}$ differential measurement (see Fig. 67 App. A.1), the PS+hadronization NP in boosted region is also dropped due to the NNLO corrections.

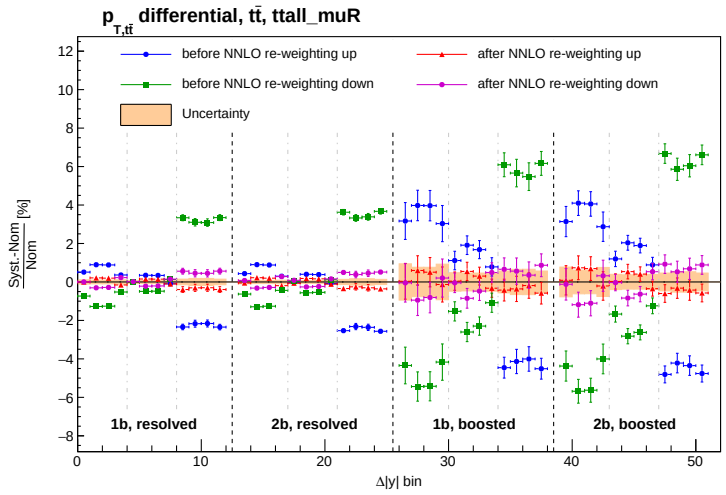


(a) FSR

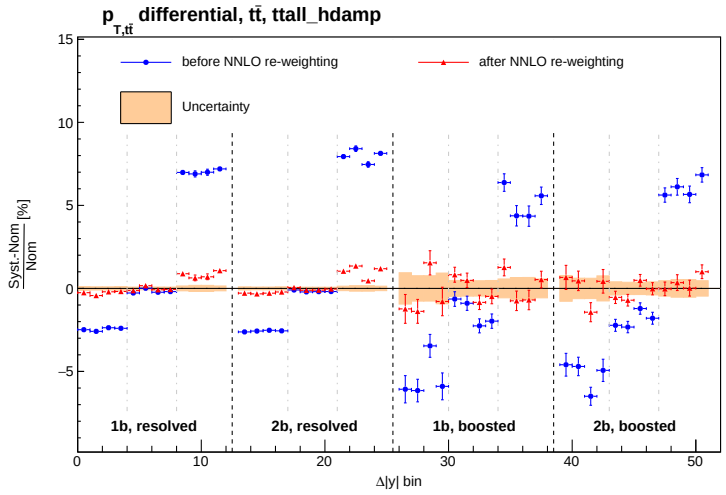


(b) PS+hadronization

Figure 44: The NNLO re-weighting applied to the systematic uncertainty for final-state radiation (a) and for choice of the parton shower and hadronization model (b). The systematic-uncertainty plots for the inclusive measurement are presented. The variation calculated by Eq. 51 is drawn in blue, the variation after the NNLO re-weighting is in red. The orange uncertainty band shows the uncertainty of nominal distribution.



(a) μ_R

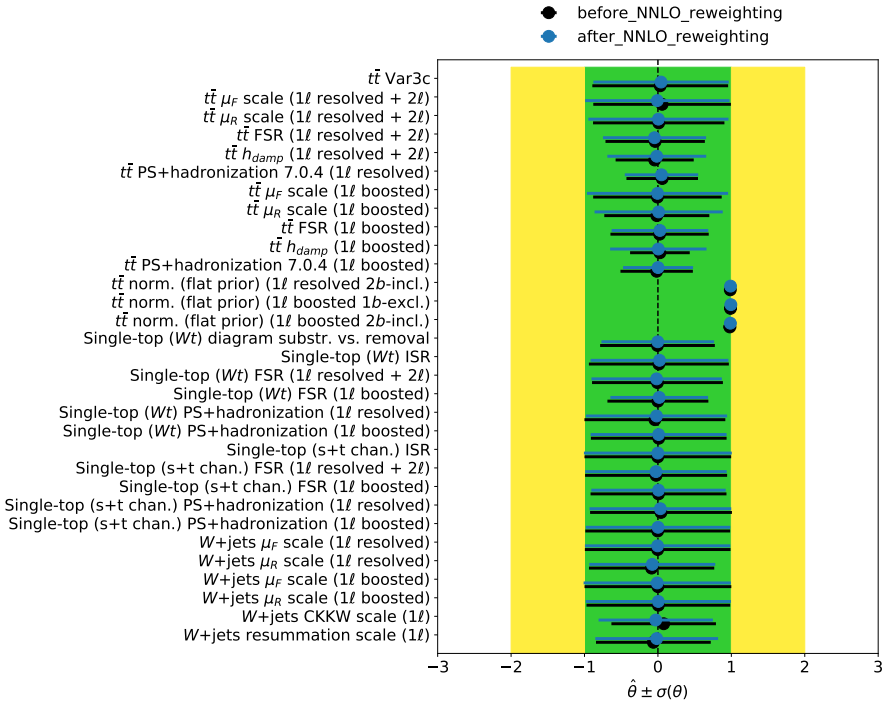


(b) h_{damp}

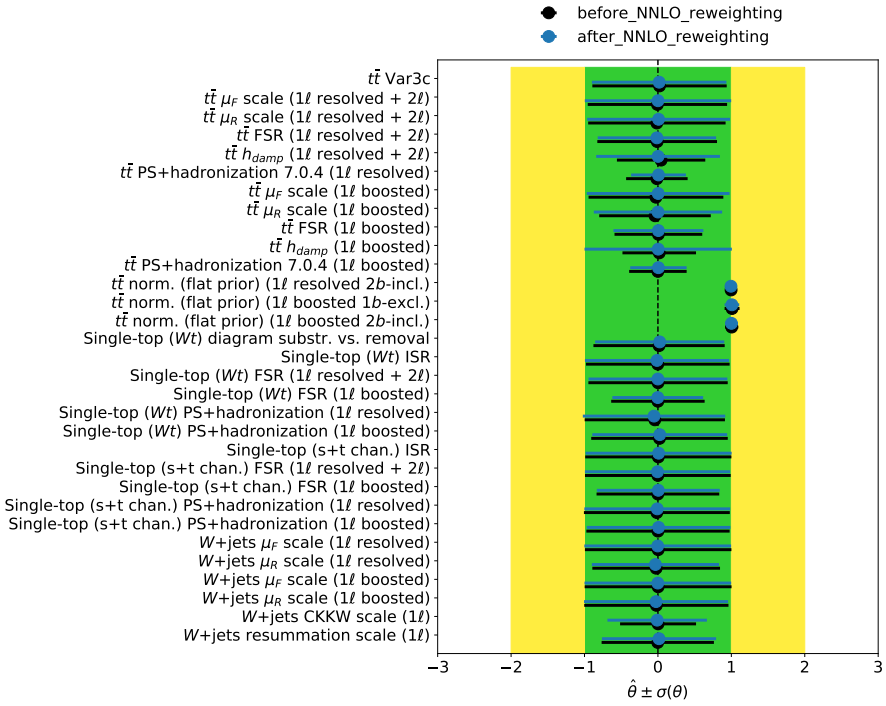
Figure 45: The NNLO re-weighting applied to the systematic uncertainty for variations of renormalization scale (a) and h_{damp} parameter (b). The systematic uncertainty plots for the differential measurement as a function of $p_T^{t\bar{t}}$ are presented. The upward and downward variations of μ_R scale calculated by Eq. 51 is drawn in blue and green, respectively. The variations after the NNLO re-weighting are in red and purple for upward and downward variations of μ_R. Variation of h_{damp} parameter before the NNLO re-weighting is depicted in blue and the NNLO re-weighted variation is in red. The orange uncertainty band shows the uncertainty of nominal distribution.

Table 14: The effect of the NNLO re-weighting on the unfolded charge asymmetry value. The Asimov data set is used. The change in the central value is minimal. The difference of uncertainties with respect to the nominal uncertainty (not applied NNLO re-weighting) in percentage is stated in the last column.

	NNLO re-weighting		
	not applied	applied	
	$A_C^{t\bar{t}} \pm \sigma(A_C^{t\bar{t}})$	$A_C^{t\bar{t}} \pm \sigma(A_C^{t\bar{t}})$	difference [%]
inclusive	0.0036 ± 0.0015	0.0035 ± 0.0016	7.7
$m_{t\bar{t}} \in [0, 500]$ GeV	0.0031 ± 0.0037	0.0031 ± 0.0037	0.2
$m_{t\bar{t}} \in [500, 750]$ GeV	0.0039 ± 0.0025	0.0040 ± 0.0025	2.5
$m_{t\bar{t}} \in [750, 1000]$ GeV	0.0049 ± 0.0057	0.0048 ± 0.0059	4.4
$m_{t\bar{t}} \in [1000, 1500]$ GeV	0.0055 ± 0.0084	0.0053 ± 0.0088	4.4
$m_{t\bar{t}} \in [1500, \infty]$ GeV	0.0072 ± 0.0257	0.0070 ± 0.0266	3.5
$\beta_z^{t\bar{t}} \in [0.0, 0.3]$	0.0001 ± 0.0049	0.0000 ± 0.0052	7.2
$\beta_z^{t\bar{t}} \in [0.3, 0.6]$	0.0014 ± 0.0036	0.0014 ± 0.0039	8.7
$\beta_z^{t\bar{t}} \in [0.6, 0.8]$	0.0028 ± 0.0036	0.0026 ± 0.0037	4.4
$\beta_z^{t\bar{t}} \in 0.8, 1.0]$	0.0082 ± 0.0041	0.0082 ± 0.0042	2.1
$p_T^{t\bar{t}} \in [0, 30]$ GeV	0.0086 ± 0.0045	0.0084 ± 0.0046	3.5
$p_T^{t\bar{t}} \in [30, 120]$ GeV	0.0008 ± 0.0036	0.0007 ± 0.0037	2.2
$p_T^{t\bar{t}} \in [120, \infty]$ GeV	0.0005 ± 0.0063	0.0005 ± 0.0072	13.7



(a) $m_{t\bar{t}}$ measurement



(b) $p_T^{t\bar{t}}$ measurement

Figure 46: Comparison of constraints for the $m_{t\bar{t}}$ (a) and $p_T^{t\bar{t}}$ (b) measurements using inputs with/without the NNLO re-weighting applied. Results from the Asimov data set are shown. The constraints of some systematic uncertainties are smaller for the NNLO re-weighted variations. For both measurements, the constraints of h_{damp} parameters are improved the most.

DE-CORRELATION BY DIFFERENTIAL BINS However, there are still some NPs, which are constrained too much. It is mostly the case for the PS+hadronization NPs (see Fig. 67 in App. A.1) and NPs connected with the variations of the h_{damp} parameter. Therefore, the procedure mentioned in the Sec. 7.6, de-correlation of systematic uncertainty by differential bins, is tested for such modelling systematic uncertainties. The most constrained uncertainties are those connected with the FSR, h_{damp} or PS+hadronization, predominantly in the $m_{t\bar{t}}$ and $p_T^{t\bar{t}}$ differential measurements. Hence, an illustration of the de-correlation procedure is shown for these. In Fig. 47, the constraints for de-correlated PS+hadronization NPs compared with previously obtained constraints are depicted for the $m_{t\bar{t}}$ differential measurement. The obvious reduction of the constraints is present, however, at the cost of more complex systematic model. Another useful information is extracted by using the de-correlation - the regions causing the largest constraints are identified. In case of the PS+hadronization systematic uncertainty, the constraint in resolved topology mostly originates from the second and third $m_{t\bar{t}}$ differential bin.

The de-correlation by differential bins is also applied to the uncertainty of FSR for single top production.

In case of the h_{damp} nuisance parameters in the $m_{t\bar{t}}$ differential measurement (Fig. 48), the situation is even better. Not only the constraints are smaller, but most of NPs are dropped because they are small enough to be considered insignificant by the pruning procedure.

Another situation, which can happen when de-correlation is used, is illustrated by the Fig. 49. In the figure, there are constraints of the FSR NPs shown. All except one NP, corresponding to the boosted topology and the last $p_T^{t\bar{t}}$ differential bin, are negligibly constrained. Hence, there is only one NP bearing the previously observed larger constraint of the boosted-region NP, which size is almost unchanged (slightly smaller).

Overall, the de-correlation by differential bins helps with the reduction of constraints for any differential measurement and NP, or, at least, can be used to point at the region of phase space responsible for the NP constraint.

In Tab. 15, the effect of the de-correlation is evaluated in term of change in the unfolded A_C uncertainty. The differences are quite small, mostly up to 5%. The exception is the last $p_T^{t\bar{t}}$ bin, which shows almost 9% change in the unfolded uncertainty.

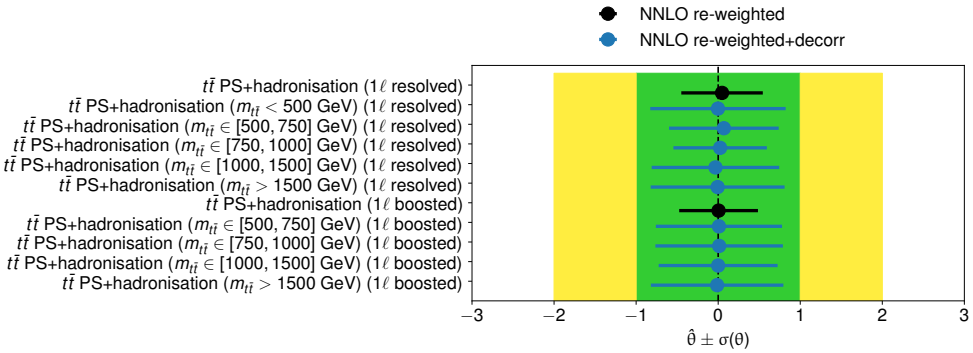


Figure 47: Comparison of constraints for the PS+hadronization systematic uncertainty for the $m_{t\bar{t}}$ measurement using inputs with/without de-correlation of the NPs by differential bins. The inputs are NNLO re-weighted. The constraints are reduced after the de-correlation. Results from the Asimov data set are shown.

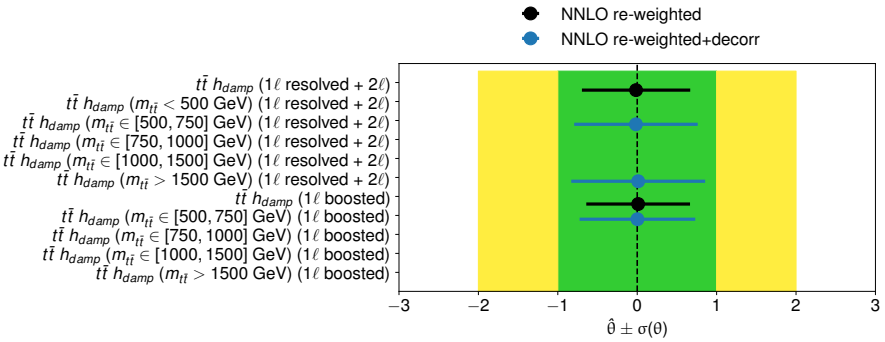


Figure 48: Comparison of constraints for the h_{hdamp} NPs for the $m_{t\bar{t}}$ measurement using inputs with/without de-correlation of the NPs by differential bins. The inputs are NNLO re-weighted. The constraints are reduced after the introduction of additional parameters. Only 3 NPs out of 9 remained in the measurement. Other NPs were removed by pruning. Results from the Asimov data set are shown.

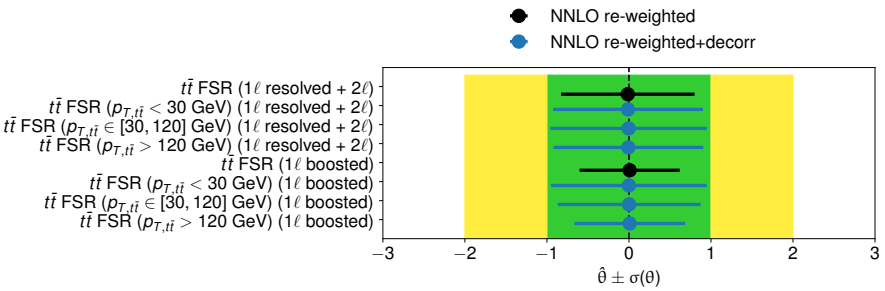


Figure 49: Comparison of constraints for the FSR NPs for the $p_T^{t\bar{t}}$ measurement using inputs with/without de-correlation of the NPs by differential bins. The inputs are NNLO re-weighted. In resolved topology, the new NPs are almost not constrained. The constraint in the boosted topology comes form the last $p_T^{t\bar{t}}$ differential bin. Results from the Asimov data set are shown.

Table 15: The effect of de-correlation by differential bins while using NNLO re-weighting on the unfolded charge asymmetry value. Asimov data set is used. The change in the central value is minimal. The difference of uncertainties with respect to the nominal uncertainty (before de-correlation) in percentage is stated in the last column. The uncertainty after de-correlation is larger according to expectations. The last $p_T^{t\bar{t}}$ bin is affected the most.

	De-correlation		
	before	after	
	$A_C^{t\bar{t}} \pm \sigma(A_C^{t\bar{t}})$	$A_C^{t\bar{t}} \pm \sigma(A_C^{t\bar{t}})$	difference [%]
$m_{t\bar{t}} \in [0, 500]$ GeV	0.0031 ± 0.0038	0.0031 ± 0.0039	4.2
$m_{t\bar{t}} \in [500, 750]$ GeV	0.0040 ± 0.0025	0.0039 ± 0.0025	1.0
$m_{t\bar{t}} \in [750, 1000]$ GeV	0.0048 ± 0.0059	0.0048 ± 0.0060	1.3
$m_{t\bar{t}} \in [1000, 1500]$ GeV	0.0054 ± 0.0088	0.0053 ± 0.0089	1.4
$m_{t\bar{t}} \in [1500, \infty]$ GeV	0.0069 ± 0.0269	0.0068 ± 0.0281	4.3
$\beta_z^{t\bar{t}} \in [0.0, 0.3]$	0.0001 ± 0.0052	-0.0001 ± 0.0055	5.3
$\beta_z^{t\bar{t}} \in [0.3, 0.6]$	0.0014 ± 0.0039	0.0014 ± 0.0040	1.9
$\beta_z^{t\bar{t}} \in [0.6, 0.8]$	0.0026 ± 0.0037	0.0026 ± 0.0038	2.5
$\beta_z^{t\bar{t}} \in [0.8, 1.0]$	0.0081 ± 0.0042	0.0082 ± 0.0043	1.2
$p_T^{t\bar{t}} \in [0, 30]$ GeV	0.0084 ± 0.0046	0.0085 ± 0.0047	2.2
$p_T^{t\bar{t}} \in [30, 120]$ GeV	0.0007 ± 0.0037	0.0006 ± 0.0037	0.3
$p_T^{t\bar{t}} \in [120, \infty]$ GeV	0.0005 ± 0.0071	0.0004 ± 0.0077	8.9

NNLO RE-WEIGHTING VERSUS TRUTH-BASED RE-WEIGHTING

Both re-weighting methods are based on the same principle, but each of them is performing it differently. In Figs. 50,51 and 52, the comparison of constraints obtained for truth re-weighted and NNLO re-weighted modelling systematic uncertainties is summarized, together with the constraints for the uncertainties defined by Eq. 51. The behavior is similar for both new setups, the NPs are constrained to smaller extent. Sometimes, like in the $p_T^{t\bar{t}}$ and $m_{t\bar{t}}$ differential measurement, the re-weighted parameters (h_{damp} NPs in the boosted topology) were pruned. In general, the NNLO re-weighted uncertainties have smaller constraints then the truth re-weighted inputs, which is connected with the larger unfolded uncertainties. In Tab. 16, there is a summary of unfolded results using inputs re-weighted in both ways. The magnitude of uncertainty change varies up to approximately 8%, which is not significant.

CONCLUSION: From now on, the truth-ratio-based (truth-based) re-weighting method is used. Although the NNLO re-weighting gives smaller constraints and subsequently larger uncertainties on the A_C value, it is decided no to use it. The NNLO recursive re-weighting, at the time of its testing, was considered experimental and not explored much, in a sense that not every feature of it was fully understood. Therefore, due to a similar effect of both re-weighting methods on the unfolded results when also the de-correlation by differential bins has been considered, safer and more understood option has been chosen.

Moreover, the impact of the systematic uncertainties on the measurement is evaluated by comparing the unfolded uncertainties obtained by including/not including the systematic uncertainties. From Tab. 17, one can see, the statistics plays a significant role in the uncertainty determination. Therefore, the small changes in the unfolded uncertainties, like those observed while testing different re-weighting methods, are minor effects with respect to the impact of the statistical uncertainty. Nevertheless, it is important to define the systematic model correctly in order to not introduce any bias in the measurement.

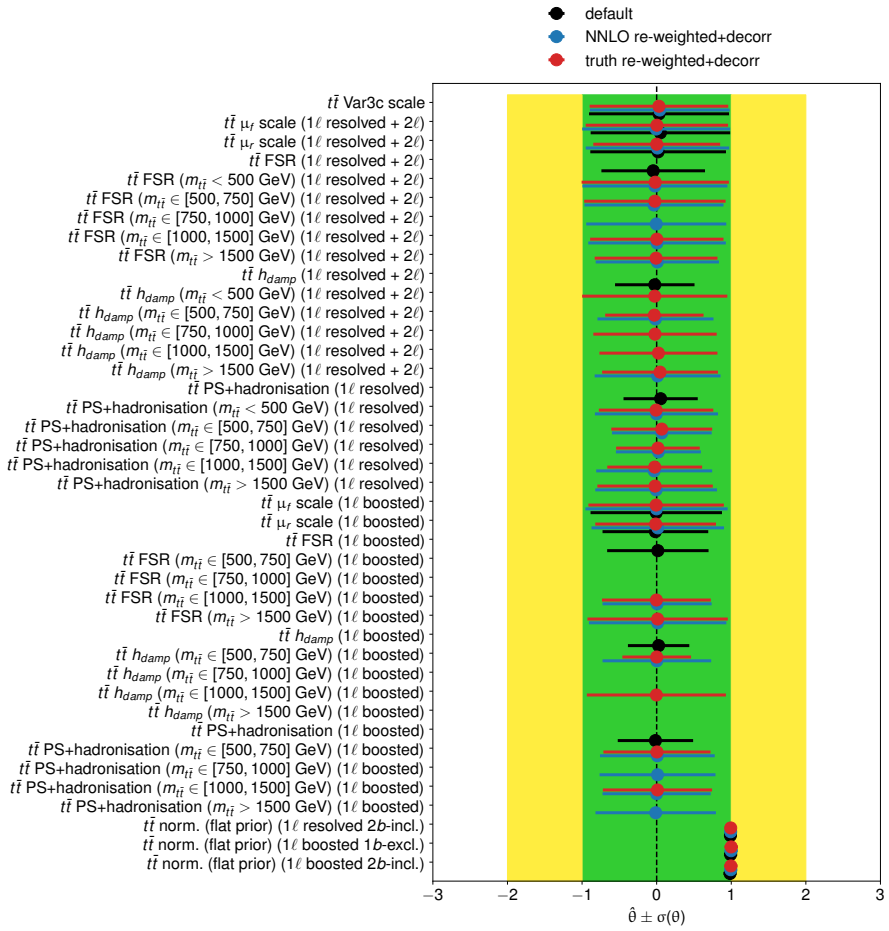


Figure 50: Comparison of constraints for signal modelling uncertainties in the $m_{t\bar{t}}$ measurement using inputs with the NNLO/truth re-weighting applied and without any re-weighting. Constraints of the NNLO re-weighted inputs are in general smaller than constraints obtained using truth re-weighting. If a NP is removed by pruning, it is not shown in the plot, i.e. some of the NPs for h_{damp} parameter variation in resolved topology are pruned after NNLO re-weighting procedures is applied. The situation is opposite in case of some PS+hadronization NPs in the boosted topology, which are not pruned after NNLO re-weighting unlike after truth re-weighting. Default setup denotes constraints of systematic uncertainties defined by Eq. 51. Results from the Asimov data set are shown.

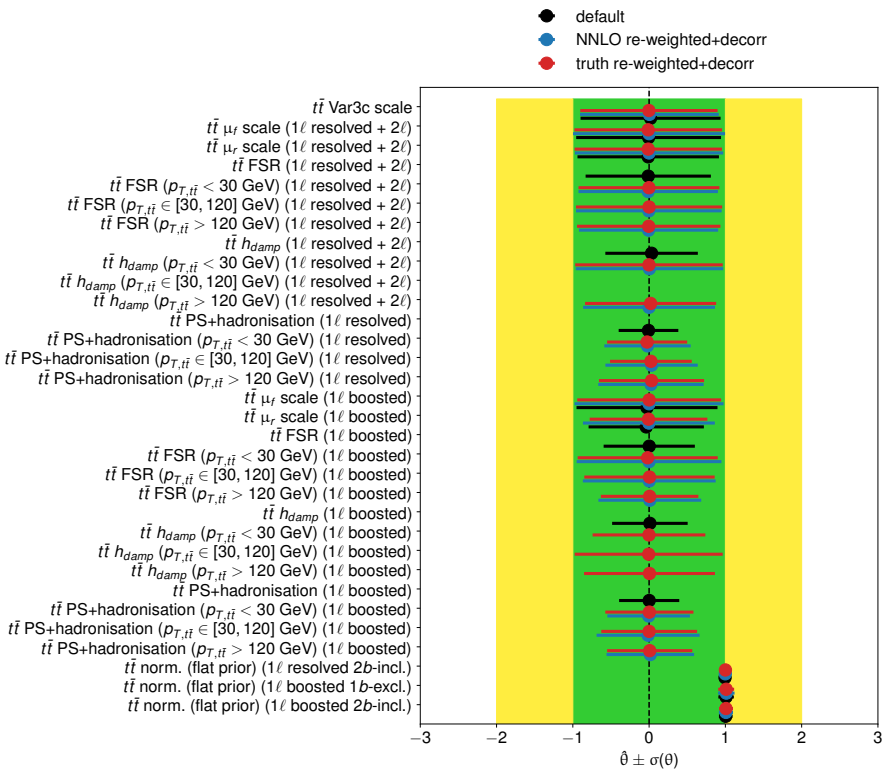


Figure 51: Comparison of constraints for signal modelling uncertainties in the $p_T^{t\bar{t}}$ measurement using inputs with the NNLO/truth re-weighting applied and without any re-weighting. Constraints of the NNLO re-weighted inputs are in general smaller or similar in size as those obtained for the truth re-weighted inputs. NPs for h_{damp} variation in boosted regions are pruned if NNLO re-weighting is used. Therefore they are not drawn. Default setup denotes constraints of systematic uncertainties defined by Eq. 51. Results from the Asimov data set are shown.

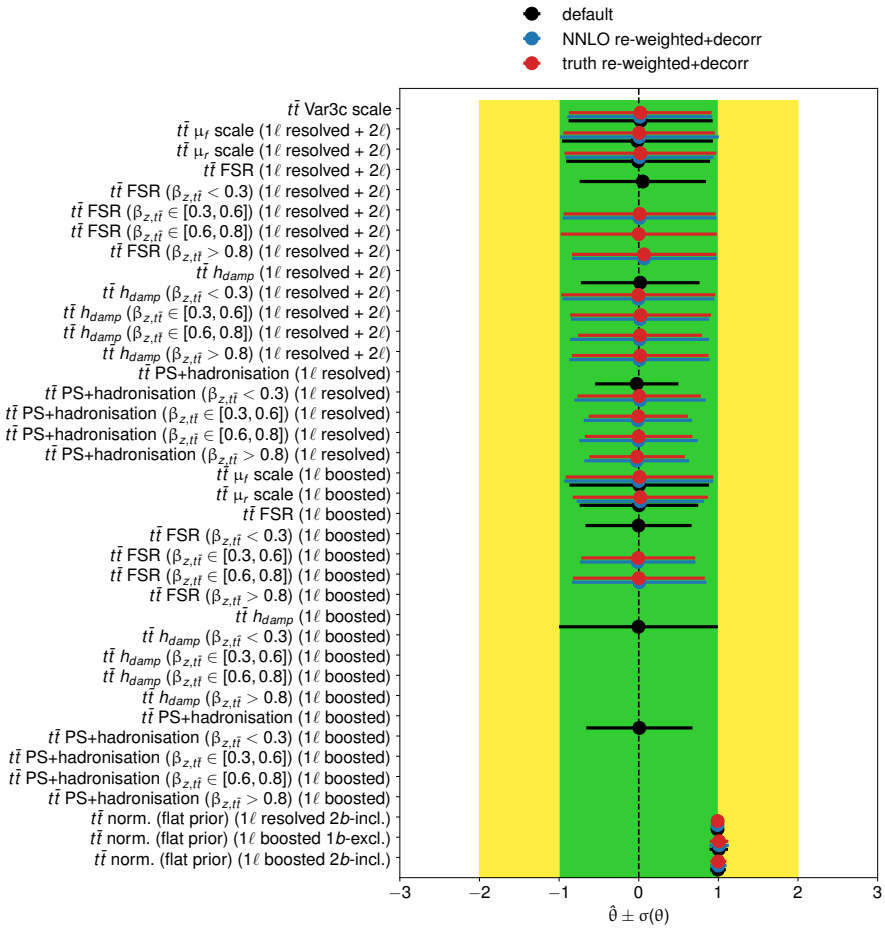


Figure 52: Comparison of constraints for signal modelling uncertainties in the $\beta_z^{t\bar{t}}$ measurement using inputs with the NNLO/truth re-weighting applied and without any re-weighting. Generally, the NNLO re-weighted systematic variations are less constrained. If a NP is removed by pruning, it is not shown in the plot. Default setup denotes constraints of systematic uncertainties defined by Eq. 51. Results from the Asimov data set are shown.

Table 16: The effect of using different re-weighting methods, either NNLO-based or truth-based, together with de-correlation of NPs by differential bins on the unfolded charge asymmetry value is presented. Asimov data set is used. The change in the central value is minimal. The difference of uncertainties with respect to the uncertainty given by unfolding of NNLO re-weighted inputs in percentage is stated in the last column. The A_C uncertainty obtained using the truth re-weighted systematic inputs is smaller than the uncertainty given by the NNLO re-weighted inputs. The first $\beta_z^{t\bar{t}}$ differential bin is affected the most.

	Re-weighting		
	NNLO-based	Truth-based	
	$A_C^{t\bar{t}} \pm \sigma(A_C^{t\bar{t}})$	$A_C^{t\bar{t}} \pm \sigma(A_C^{t\bar{t}})$	difference [%]
inclusive	0.0035 ± 0.0016	0.0036 ± 0.0015	-8.0
$m_{t\bar{t}} \in [0, 500]$ GeV	0.0031 ± 0.0039	0.0030 ± 0.0040	1.8
$m_{t\bar{t}} \in [500, 750]$ GeV	0.0039 ± 0.0025	0.0040 ± 0.0024	-4.1
$m_{t\bar{t}} \in [750, 1000]$ GeV	0.0048 ± 0.0060	0.0048 ± 0.0059	-2.6
$m_{t\bar{t}} \in [1000, 1500]$ GeV	0.0053 ± 0.0089	0.0056 ± 0.0087	-2.8
$m_{t\bar{t}} \in [1500, \infty]$ GeV	0.0068 ± 0.0281	0.0077 ± 0.0263	-6.1
$\beta_z^{t\bar{t}} \in [0.0, 0.3]$	-0.0001 ± 0.0055	$< 0.0001 \pm 0.0050$	-8.4
$\beta_z^{t\bar{t}} \in [0.3, 0.6]$	0.0014 ± 0.0040	0.0015 ± 0.0037	-7.2
$\beta_z^{t\bar{t}} \in [0.6, 0.8]$	0.0026 ± 0.0038	0.0027 ± 0.0037	-2.0
$\beta_z^{t\bar{t}} \in [0.8, 1.0]$	0.0082 ± 0.0043	0.0083 ± 0.0041	-3.8
$p_T^{t\bar{t}} \in [0, 30]$ GeV	0.0085 ± 0.0047	0.0087 ± 0.0046	-3.0
$p_T^{t\bar{t}} \in [30, 120]$ GeV	0.0006 ± 0.0037	0.0007 ± 0.0036	-3.3
$p_T^{t\bar{t}} \in [120, \infty]$ GeV	0.0004 ± 0.0077	0.0004 ± 0.0075	-2.7

Table 17: The results from the Asimov unfolding without the systematic uncertainties (stat-only) and with the inclusion of them (stat+syst). The effect of the systematic uncertainties can be deduced from the unfolded uncertainties. In the last column, there is an estimate of the fraction of the statistical and the 'stat+syst' uncertainty, which evaluates the impact of the systematic uncertainties. The statistical uncertainty creates roughly 45-77% of the 'stat+syst' uncertainty. Thus, the precision of the measurement is statistically driven in the most of differential bins. The truth re-weighting and de-correlation of NPs by differential bins are engaged.

	Stat-only	Stat+syst	
	$A_C^{t\bar{t}} \pm \sigma(A_C^{t\bar{t}})$	$A_C^{t\bar{t}} \pm \sigma(A_C^{t\bar{t}})$	$\left(\frac{\sigma_{stat}}{\sigma_{stat+syst}}\right)^2$ [%]
inclusive	0.0036 ± 0.0011	0.0036 ± 0.0015	57.3
$m_{t\bar{t}} \in [0, 500]$	0.0031 ± 0.0029	0.0030 ± 0.0040	57.1
$m_{t\bar{t}} \in [500, 750]$	0.0041 ± 0.0020	0.0040 ± 0.0024	64.7
$m_{t\bar{t}} \in [750, 1000]$	0.0048 ± 0.0047	0.0048 ± 0.0059	66.6
$m_{t\bar{t}} \in [1000, 1500]$	0.0057 ± 0.0071	0.0056 ± 0.0087	70.8
$m_{t\bar{t}} \in [1500, \infty]$	0.0080 ± 0.0225	0.0077 ± 0.0263	77.2
$\beta_z^{t\bar{t}} \in [0.0, 0.3]$	-0.0001 ± 0.0040	$< 0.0001 \pm 0.0050$	65.8
$\beta_z^{t\bar{t}} \in [0.3, 0.6]$	0.0014 ± 0.0031	0.0015 ± 0.0037	68.5
$\beta_z^{t\bar{t}} \in [0.6, 0.8]$	0.0028 ± 0.0029	0.0027 ± 0.0037	64.9
$\beta_z^{t\bar{t}} \in [0.8, 1.0]$	0.0085 ± 0.0027	0.0083 ± 0.0041	46.5
$p_T^{t\bar{t}} \in [0, 30]$	0.0087 ± 0.0034	0.0087 ± 0.0046	57.7
$p_T^{t\bar{t}} \in [30, 120]$	0.0007 ± 0.0027	0.0007 ± 0.0036	57.9
$p_T^{t\bar{t}} \in [120, \infty]$	0.0006 ± 0.0046	0.0004 ± 0.0075	44.7

Part IV

DATA ANALYSIS

Charge and leptonic asymmetry values extracted from measured data. Results of inclusive and differential measurements. Charge asymmetry measurement in context of EFT interpretation. Bounds on relevant Wilson coefficients determined. Complementarity of energy asymmetry and charge asymmetry measurements in context of EFT interpretation presented.

8

UNFOLDED CHARGE ASYMMETRY

Measurement of the charge asymmetry is performed using the full data set collected at the center-of-mass energy of 13 TeV during years 2015-2018. Data from single-lepton and dilepton $t\bar{t}$ decay channel are unfolded.

8.1 INCLUSIVE MEASUREMENT

The unfolding processes the whole distribution of measured data divided according to the $t\bar{t}$ decay channel, topology/lepton flavor and b-jet multiplicity. Hence, the number of bins in the reconstructed distribution is given by

$1\ell = 4(\Delta|y|\text{bins}) \times 2(b\text{-tag multiplicity}) \times 2(\text{topology}) = 16$ bins

$2\ell = 4(\Delta|y|\text{bins}) \times 2(b\text{-tag multiplicity}) \times 2(\text{lepton flavor}) = 16$ bins, so in total 32 bins in the reconstructed distribution.

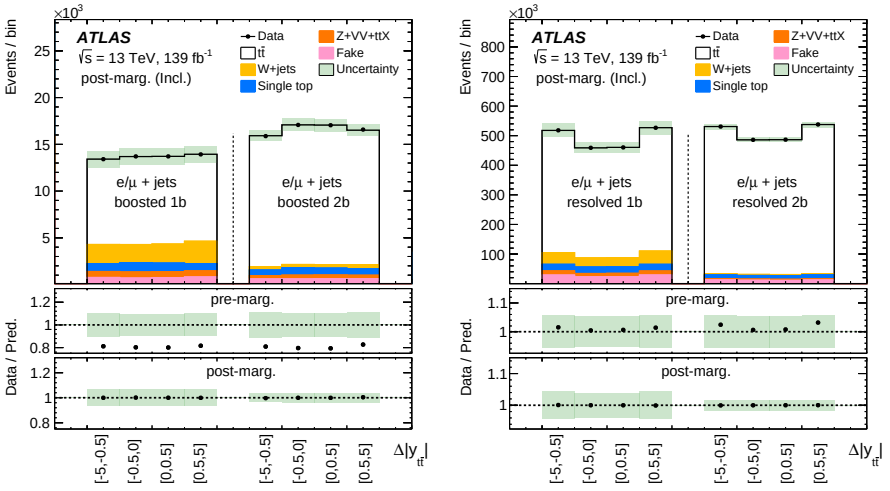
Inevitable part of the unfolding procedure is marginalization of the nuisance parameters. Basically, all nuisance parameters are integrated out from the posterior distribution of the truth-level bins. Pulls and constraints of the NPs extracted from the marginal posterior distribution of each NP are applied to the 'nominal' (MC simulated) prediction of each systematic uncertainty. These corrected variations are used to construct the post-marginalization reconstructed distribution of the $\Delta|y|$ variable. The favorable impact of the marginalization is pictured in the Fig. 53, where $\Delta|y|$ distributions for each region in the single-lepton and dilepton channel are plotted together with the data/MC agreement before and after marginalization. Not only the MC simulation aligns better with the measured data, but the total uncertainty band is also reduced.

The effect of marginalization is illustrated only for the inclusive measurement, but similar behavior is observed for differential measurements too.

The ranking of systematic uncertainties performed for the inclusive measurement shows, that the most important systematic uncertainties are the modelling uncertainties, like variation of the renormalization scale, modelling of parton shower and hadronization, variations in the radiation in the final state for signal $t\bar{t}$ distribution or normalization of fake lepton background, see Fig. 54. However, the effect on the A_C value of any of mentioned systematic uncertainties is of at least one order smaller than the statistical uncertainty. The constraints and pulls of each NP corresponding to the 10 highest ranked systematic uncertainties are plotted also in Fig. 54.

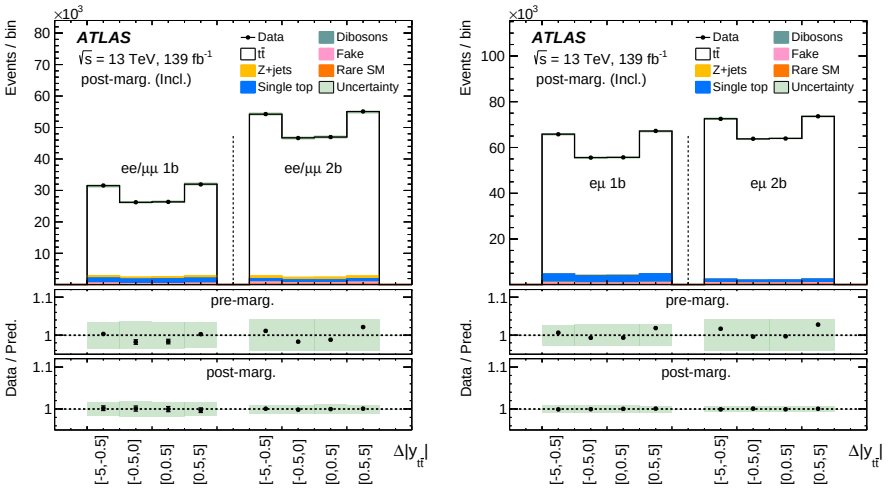
Gamma parameters do not have a constraint, because their role is to correct the background normalization of each bin. By default, the value of their pull is set to 1.

As the main result, the FBU gives the posterior distribution of charge asymmetry, plotted in Fig. 55 for the inclusive measurement.



(a) Boosted topology

(b) Resolved topology



(c) $ee + \mu\mu$ region

(d) $e\mu$ region

Figure 53: $\Delta|y|$ distributions for individual regions in the measurement: (a) boosted, (b) resolved topologies of the single-lepton channel, (c) same-flavor and (d) different-flavor regions of the dilepton channel. Contributions of signal and individual background distributions are marked in different colors. Subplots below the main plots show the agreement of data and simulations together with estimated uncertainty before and after marginalization. Generally, the marginalization improves the data/prediction ratio and reduces the uncertainty.

The mean and sigma of the distribution are extracted and interpreted as the inclusive charge asymmetry value and its unfolded uncertainty. The most up-to-date theoretical prediction given by the SM calculation at NNLO in QCD and NLO in EW theory of A_C is drawn in the same figure as a green band [29]. The unfolded charge asymmetry agrees well with the SM prediction.

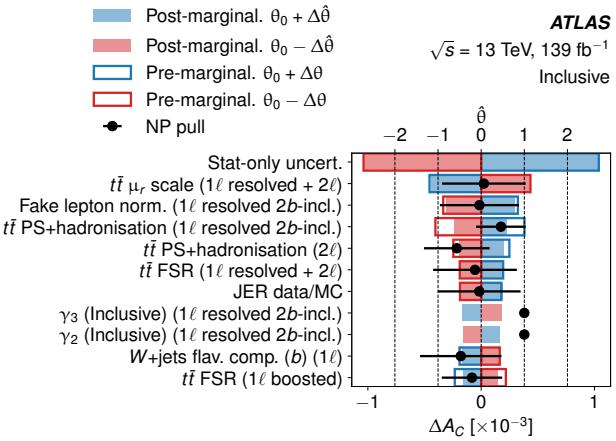


Figure 54: Ranking of the systematic uncertainties in the inclusive measurement for the measured data. Top 10 highest ranked systematic uncertainties are shown in the plot. The order is defined by the size of the effect each of systematic uncertainty has on the A_C value by shifting the variation by their constraints (post-marginalization ranking). Blue and red bars show the effect of upward and downward systematic variation on the A_C value, respectively. The frames and filled bars denote the pre-marginalization and post-marginalization results of the ranking, respectively. Black data points display the constraints and pulls of the ranked systematic uncertainties.

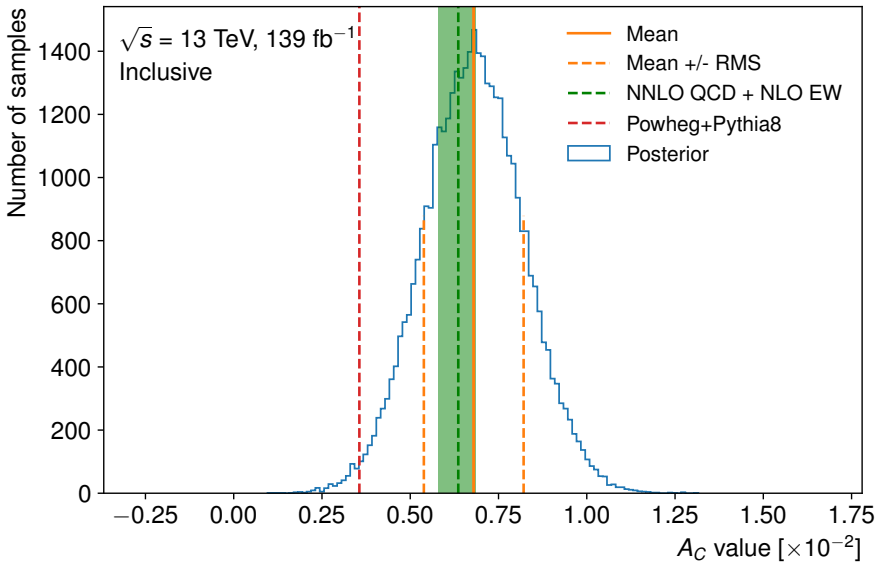


Figure 55: Posterior distribution of the unfolded A_C obtained from measured data in the inclusive measurement. The value of A_C and its unfolded uncertainty are depicted in orange solid and dashed lines. Charge asymmetry prediction given by the POWHEG+PYTHIA8 generators is represented by red dashed line, while the prediction calculated at the NNLO in QCD and NLO in EW is displayed as green band [29].

8.2 DIFFERENTIAL MEASUREMENTS

The dependence of the A_C on different kinematic observables, like $p_T^{t\bar{t}}$, $m_{t\bar{t}}$ or $\beta_z^{t\bar{t}}$ is investigated. The measured data are therefore additionally divided to different intervals of corresponding variable, so the distribution, which is going to be unfolded, has $32 \times N_{diff}$ bins, where N_{diff} is number of differential bins. The truth distribution has also ($4\ \Delta|y|$ bins $\times N_{diff}$) bins. The form of the response matrix responsible for translating the truth spectrum to the measured distribution is illustrated by a sketch in Fig. 56. The unfolding produces A_C posterior

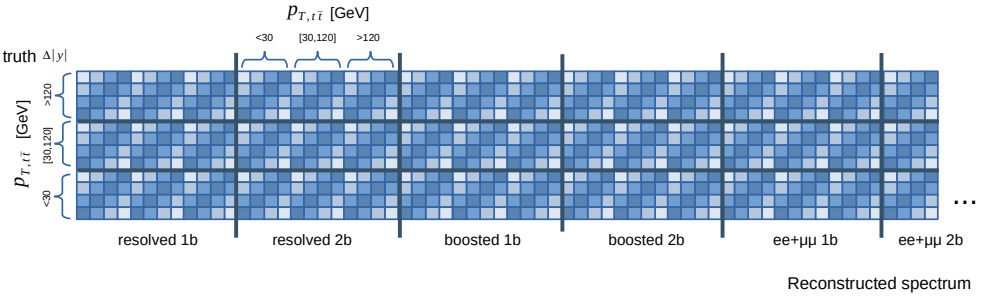
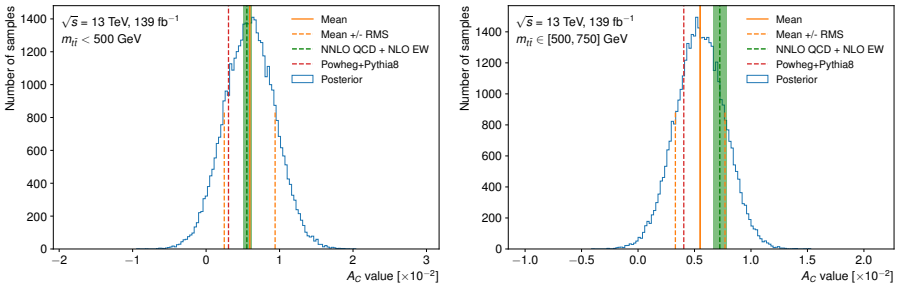


Figure 56: Illustration of the response matrix for differential measurement with 3 intervals for a variable, i.e. 3 differential bins of $p_T^{t\bar{t}}$. The response matrix for any number of differential bins is constructed analogously.

distribution for each differential bin separately. Hence as a result, the A_C value is obtained for each interval of the differential variable ($p_T^{t\bar{t}}$, $\beta_z^{t\bar{t}}$, $m_{t\bar{t}}$) giving an insight into how the charge asymmetry behaves in different parts of phase space or, more importantly, the effects of BSM physics could be displayed by discrepancies of unfolded A_C and A_C calculated by the SM. The unfolded posterior distributions of A_C for the $m_{t\bar{t}}$ differential measurement are given in Fig. 57. The unfolded results do not differ much from the SM calculation in case of $m_{t\bar{t}}$ differential measurement or any other measurements performed ($p_T^{t\bar{t}}$, $\beta_z^{t\bar{t}}$).

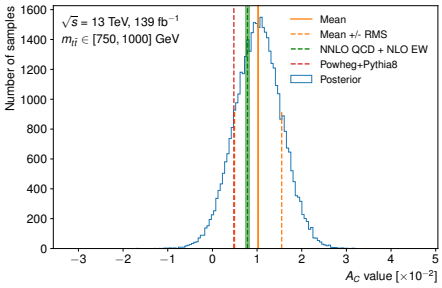
The posterior distributions for differential measurement of A_C as a function of $p_T^{t\bar{t}}$ and $\beta_z^{t\bar{t}}$ are in the App. A.2.

The ranking of systematic uncertainties is performed separately for each differential bin of all differential measurements. Corresponding plots can be found in App. A.3. In general, $t\bar{t}$ and background modelling systematic uncertainties are mostly ranked among the top 10 most influential uncertainties, together with some of JER and JES parameters. Specifically in $m_{t\bar{t}}$ and $p_T^{t\bar{t}}$ measurements, NPs for muon energy scale and resolution bias of muon sagitta or perpendicular resolution of E_T^{miss} are among the top ranked uncertainties. With the exception of last $\beta_z^{t\bar{t}}$ and last $p_T^{t\bar{t}}$ differential bin, the effect of any systematic uncertainty is negligible in comparison with the size of the statistical uncertainty. In case of last $\beta_z^{t\bar{t}}$ differential bin, $0.8 < \beta_z^{t\bar{t}} < 1.0$, or

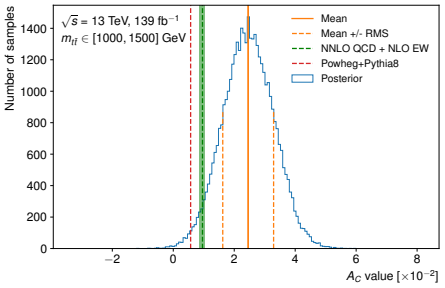


(a) $m_{t\bar{t}} < 500$ GeV

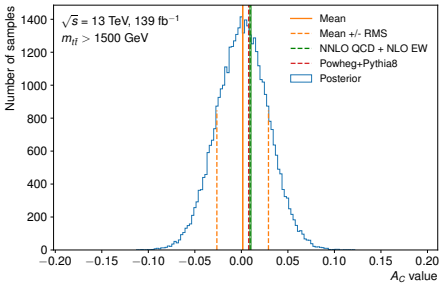
(b) $m_{t\bar{t}} \in (500, 750)$ GeV



(c) $m_{t\bar{t}} \in (750, 1000)$ GeV



(d) $m_{t\bar{t}} \in (1000, 1500)$ GeV



(e) $m_{t\bar{t}} > 1500$ GeV

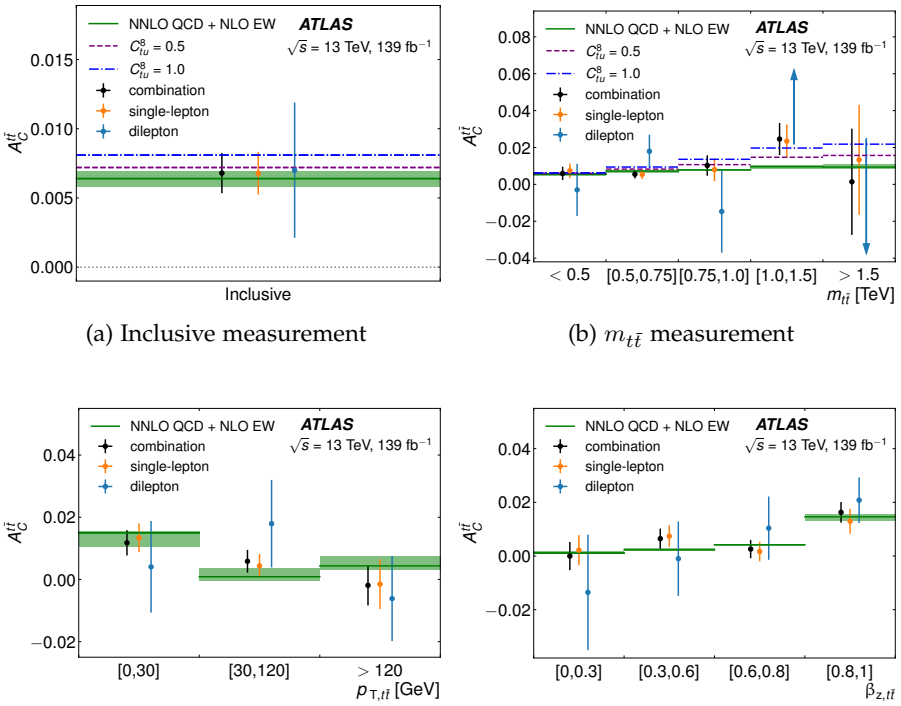
Figure 57: Posterior A_C distributions for $m_{t\bar{t}}$ differential measurement. The value of A_C and its unfolded uncertainty are depicted in orange solid and dashed lines. Charge asymmetry prediction given by the POWHEG+PYTHIA 8 generators is represented by red dashed line, while the prediction calculated at the NNLO in QCD and NLO in EW is displayed as green band [29].

the last $p_T^{t\bar{t}}$ bin, $p_T^{t\bar{t}} > 120$ GeV, the statistical uncertainty is still larger than the effect of the highest ranked NP, but the difference between them is considerably smaller than for other differential bins.

8.3 SUMMARY

The unfolded A_C values obtained using not only combined data from the single-lepton and dilepton channel, but also single-lepton/dilepton subdatasets are depicted in Fig. 58 for inclusive and differential mea-

surements. The vertical lines belonging to each data point, show the total uncertainty for that particular bin.



(a) Inclusive measurement

(b) $m_{t\bar{t}}$ measurement

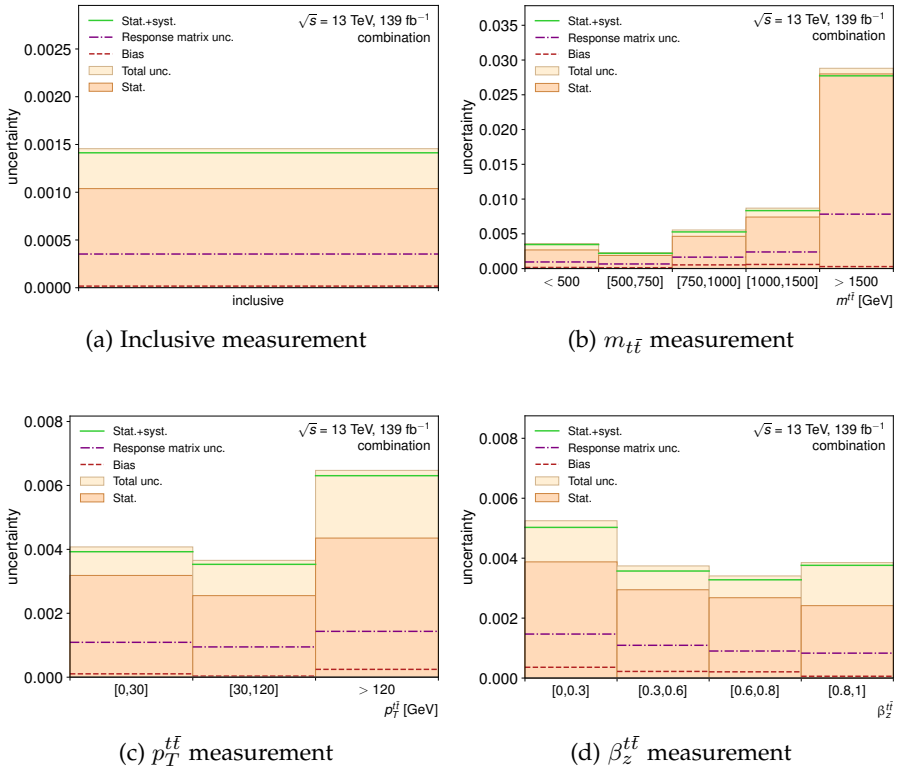
(c) $p_T^{t\bar{t}}$ measurement

(d) $\beta_z^{t\bar{t}}$ measurement

Figure 58: Unfolded A_C using the single-lepton, dilepton channel data separately and in combination for (a) inclusive measurement, (b) $m_{t\bar{t}}$, (c) $p_T^{t\bar{t}}$ and (d) $\beta_z^{t\bar{t}}$ differential measurements. Black points correspond to the combined results, orange to the single-lepton and blue points to the dilepton measurement. The vertical lines show the total uncertainty. The unfolded results are compared with the theoretical prediction at the precision of NNLO in QCD and NLO in EW theory [29]. The prediction is represented by green bands. Moreover, for inclusive and $m_{t\bar{t}}$ measurements, A_C values calculated assuming non-zero values of Wilson coefficient C_{tu}^8 are depicted by dashed and dash-dotted lines in purple and dark blue color, respectively.

NOTE Total uncertainty is given by the unfolded uncertainty, uncertainty coming from the bias of the unfolding (Sec. 6.7) and from limited statistics of the response matrix (Sec. 7.5), which are summed in quadrature. In Fig. 59, the sizes of the effects, by which individual uncertainties contribute to the total uncertainty of the results of combination of $t\bar{t}$ decay channels, are displayed. Size of the statistical uncertainties are presented in the same Figure for comparison. All uncertainties used to calculate the total uncertainty are listed, in App. A.4 in Tab. 20 for inclusive and $m_{t\bar{t}}$ differential measurements

and in Tab. 21 for $p_T^{t\bar{t}}$ and $\beta_z^{t\bar{t}}$ differential measurements, for combination and also for the single-lepton and dilepton channel separately.



(a) Inclusive measurement

(b) $m_{t\bar{t}}$ measurement

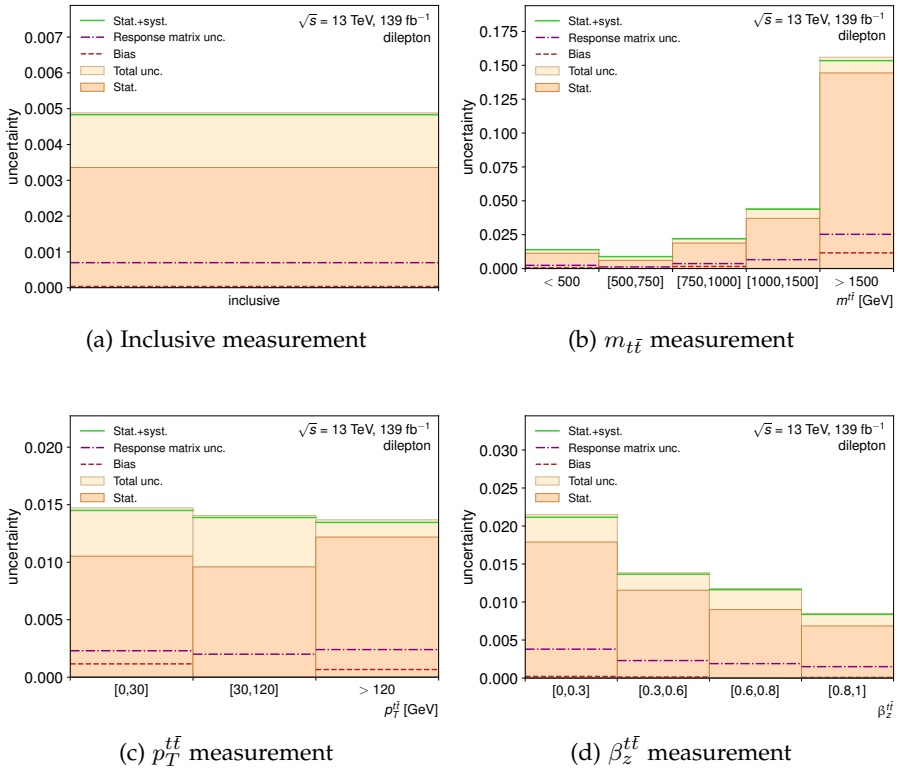
(c) $p_T^{t\bar{t}}$ measurement

(d) $\beta_z^{t\bar{t}}$ measurement

Figure 59: The breakdown of the uncertainties contributing to the total uncertainty for inclusive and differential measurements of analysis of combined single-lepton and dilepton data. The unfolded uncertainties from the unfolding with incorporated systematic uncertainties (stat.+syst.) are depicted by green lines. The bias of the unfolding procedure is estimated using unfolded A_C values from data (red dashed lines). The uncertainties corresponding to the effect of finite-size MC samples on the response matrix determination, are marked with purple dash-dotted lines. The total uncertainties calculated as a sum of squares of previously stated uncertainties are given in light beige color. Additionally, the size of the statistical uncertainty obtained from unfolding performed without systematic uncertainties included is drawn in dark beige color. Only the size of each uncertainty is plotted.

The A_C values extracted using only dilepton data (blue data points in Fig. 58) demonstrate the highest uncertainty values due to the lowest amount of data in this $t\bar{t}$ decay channel. In Fig. 60 (exact values in Tab. 20 and Tab. 21 in App. A.4), it is demonstrated, that the largest contribution to the uncertainty comes from the unfolding, where both, the statistical and systematic uncertainties are accounted for. Moreover, in Tab. 22 and Tab. 23 in App. A.5, the effect of the inclusion of systematic uncertainties is shown for combination and separately for

the single-lepton and dilepton channel. In most cases, the statistical uncertainty is the dominant source of uncertainty.



(a) Inclusive measurement

(b) $m_{t\bar{t}}$ measurement

(c) $p_T^{t\bar{t}}$ measurement

(d) $\beta_z^{t\bar{t}}$ measurement

Figure 60: The breakdown of the uncertainties contributing to the total uncertainty for inclusive and differential measurements of analysis of dilepton data. The unfolded uncertainties from the unfolding with incorporated systematic uncertainties (stat.+syst.) are depicted by green lines. The bias of the unfolding procedure is estimated using unfolded A_C values from data (red dashed lines). The uncertainties corresponding to the effect of finite-size MC samples on the response matrix determination, are marked with purple dash-dotted lines. The total uncertainties calculated as a sum of squares of previously stated uncertainties are given in light beige color. Additionally, the size of the statistical uncertainty obtained from unfolding performed without systematic uncertainties included is drawn in dark beige color. Only the size of each uncertainty is plotted.

In addition to the SM prediction shown for all measurements, the A_C predictions calculated using assumption of non-zero values of the Wilson coefficient C_{tu}^8 are marked with dashed and dash-dotted lines, specifically for inclusive and $m_{t\bar{t}}$ differential measurements in Fig. 58a and Fig. 58b. In the inclusive measurement, we observe the evidence for the charge asymmetry at the LHC – the measured value is not compatible with a hypothesis of zero charge asymmetry at the level of 4.7σ. The exact values obtained from the unfolding together with

the total uncertainty and SM prediction are listed in Tab. 18 for inclusive and $m_{t\bar{t}}$ differential measurements and in Tab. 19 for $p_T^{t\bar{t}}$ and $\beta_z^{t\bar{t}}$ measurements for combined unfolding and single-lepton/dilepton unfolding.

Table 18: Unfolded A_C values for the inclusive and $m_{t\bar{t}}$ differential measurements are given together with the total uncertainties and theoretical prediction from the SM, calculated at NNLO in QCD and NLO in EW theory.

		Data 139 fb^{-1}		SM prediction
	Channel	$A_C^{t\bar{t}}$	Total unc.	
Inclusive	1ℓ	0.0068	0.0015	$0.0064^{+0.0005}_{-0.0006}$
	2ℓ	0.0070	0.0049	
	$1\ell + 2\ell$	0.0068	0.0015	
$m_{t\bar{t}} < 500$ GeV	1ℓ	0.0074	0.0039	$0.0056^{+0.0006}_{-0.0006}$
	2ℓ	−0.0030	0.0141	
	$1\ell + 2\ell$	0.0059	0.0036	
$m_{t\bar{t}} \in (500, 750)$ GeV	1ℓ	0.0054	0.0025	$0.0072^{+0.0006}_{-0.0006}$
	2ℓ	0.0180	0.0089	
	$1\ell + 2\ell$	0.0055	0.0023	
$m_{t\bar{t}} \in (750, 1000)$ GeV	1ℓ	0.0080	0.0062	$0.0079^{+0.0004}_{-0.0006}$
	2ℓ	−0.0147	0.0223	
	$1\ell + 2\ell$	0.0102	0.0056	
$m_{t\bar{t}} \in (1000, 1500)$ GeV	1ℓ	0.0234	0.0090	$0.0096^{+0.0009}_{-0.0009}$
	2ℓ	0.0663	0.0444	
	$1\ell + 2\ell$	0.0246	0.0087	
$m_{t\bar{t}} > 1500$ GeV	1ℓ	0.0133	0.0298	$0.0094^{+0.0015}_{-0.0011}$
	2ℓ	−0.1313	0.1560	
	$1\ell + 2\ell$	0.0014	0.0288	

Table 19: Unfolded A_C values for the $p_T^{t\bar{t}}$ and $\beta_z^{t\bar{t}}$ differential measurements are given together with the total uncertainties and theoretical prediction from the SM, calculated at NNLO in QCD and NLO in EW theory.

		Data 139 fb^{-1}		SM prediction
	Channel	$A_C^{t\bar{t}}$	Total unc.	
$p_T^{t\bar{t}} < 30$ GeV	1ℓ	0.0134	0.0046	$0.0150^{+0.0006}_{-0.0046}$
	2ℓ	0.0041	0.0147	
	$1\ell + 2\ell$	0.0118	0.0041	
$p_T^{t\bar{t}} \in (30, 120)$ GeV	1ℓ	0.0044	0.0038	$0.0009^{+0.0028}_{-0.0012}$
	2ℓ	0.0179	0.0141	
	$1\ell + 2\ell$	0.0058	0.0037	
$p_T^{t\bar{t}} > 120$ GeV	1ℓ	-0.0015	0.0080	$0.0044^{+0.0030}_{-0.0014}$
	2ℓ	-0.0061	0.0137	
	$1\ell + 2\ell$	-0.0019	0.0065	
$\beta_z^{t\bar{t}} \in (0, 0.3)$	1ℓ	0.0022	0.0056	$0.0011^{+0.0005}_{-0.0004}$
	2ℓ	-0.0135	0.0215	
	$1\ell + 2\ell$	< 0.0001	0.0052	
$\beta_z^{t\bar{t}} \in (0.3, 0.6)$	1ℓ	0.0074	0.0040	$0.0023^{+0.0006}_{-0.0004}$
	2ℓ	-0.0010	0.0139	
	$1\ell + 2\ell$	0.0065	0.0037	
$\beta_z^{t\bar{t}} \in (0.6, 0.8)$	1ℓ	0.0017	0.0037	$0.0042^{+0.0003}_{-0.0003}$
	2ℓ	0.0104	0.0118	
	$1\ell + 2\ell$	0.0026	0.0034	
$\beta_z^{t\bar{t}} \in (0.8, 1.0)$	1ℓ	0.0130	0.0046	$0.0146^{+0.0012}_{-0.0014}$
	2ℓ	0.0208	0.0085	
	$1\ell + 2\ell$	0.0163	0.0039	

The precision of the inclusive $A_C^{t\bar{t}}$ measurement is considerably improved if it is compared with the precision of measurements using data collected at lower center-of-mass energies, 7 or 8 TeV [45]. It is demonstrated in Fig. 61, where results from the measurements using data collected by the CMS and ATLAS detector at $\sqrt{8}$ TeV are compared with this measurement.

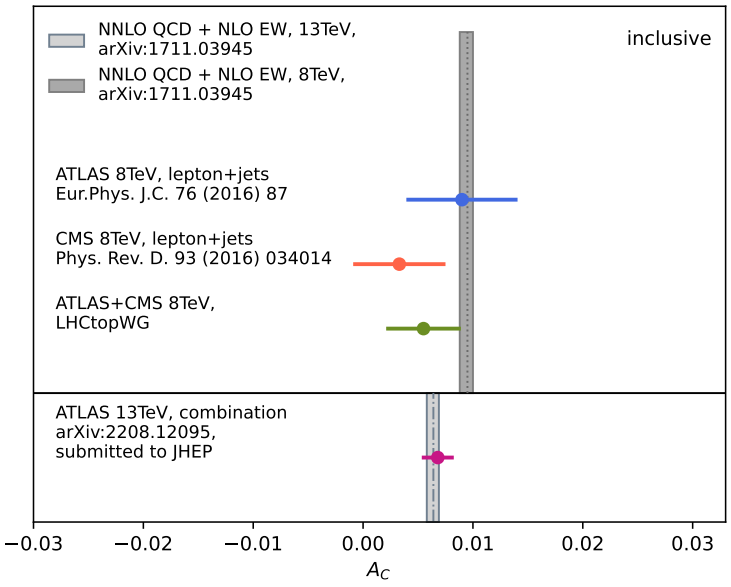


Figure 61: Inclusive charge asymmetry measured in the single-lepton channel at $\sqrt{s} = 8$ TeV by the CMS and ATLAS detector, together with combination of both measurements [45], are depicted in the first panel from the top. Below in the second panel, inclusive charge asymmetry measured at $\sqrt{s} = 13$ TeV with the ATLAS detector using data from the single-lepton and dilepton channel is shown.

In Fig. 62, results from the last three $m_{t\bar{t}}$ differential bins, $m_{t\bar{t}} \in (750, 1000)$ GeV, $m_{t\bar{t}} \in (1000, 1500)$ GeV and $m_{t\bar{t}} > 1500$ GeV, are compared with A_C values obtained from measurement performed by the CMS Collaboration [49], using data collected at $\sqrt{s} = 13$ TeV corresponding to the integrated luminosity of 138 fb^{-1}. The CMS measurement focuses only on boosted $t\bar{t}$ events with $m_{t\bar{t}}$ higher than 750 GeV, which decay into single lepton and jets. Results from both measurements are compatible within uncertainties.

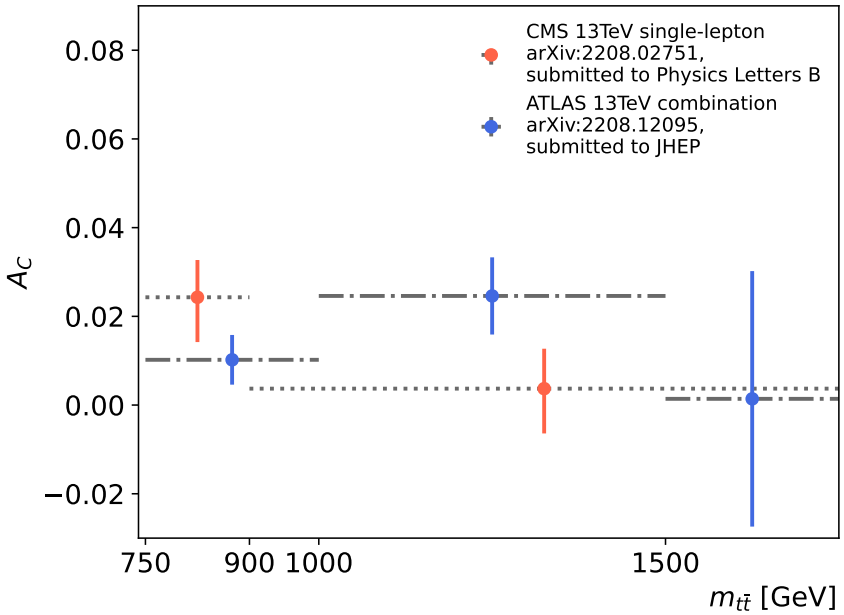
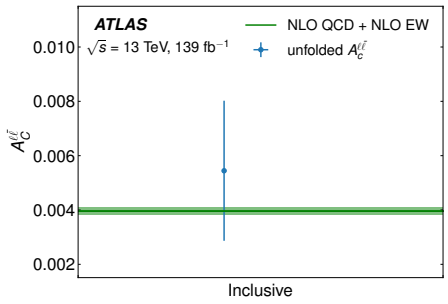


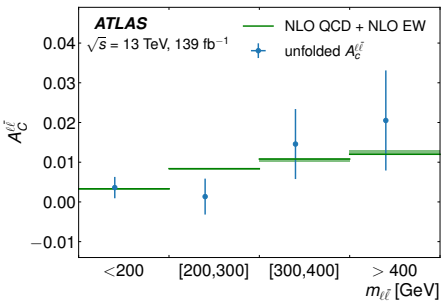
Figure 62: Comparison of $m_{t\bar{t}}$ differential charge asymmetry measurements performed by the ATLAS [2] and CMS Collaboration [49] using data measured at $\sqrt{s} = 13$ TeV. Each measurement uses different binning in $m_{t\bar{t}}$. Measurement using data collected by the CMS detector, marked by orange color, obtains A_C values for events with $m_{t\bar{t}} \in (750, 900)$ GeV and $m_{t\bar{t}} > 900$ GeV. Grey dotted lines show the size of the $m_{t\bar{t}}$ regions. Blue data points represent A_C value unfolded in the measurement presented in this thesis in three $m_{t\bar{t}}$ differential bins, $m_{t\bar{t}} \in (750, 1000)$ GeV, $m_{t\bar{t}} \in (1000, 1500)$ GeV and $m_{t\bar{t}} > 1500$ GeV. Dash-dotted grey horizontal lines label the size of $m_{t\bar{t}}$ bins.

8.4 LEPTONIC CHARGE ASYMMETRY

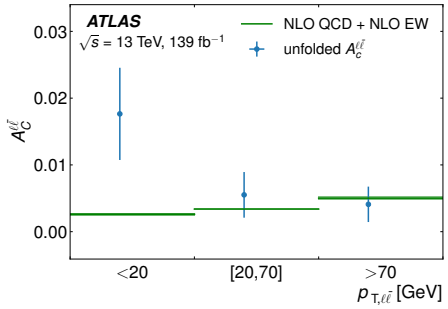
Additionally, in dilepton channel, the leptonic charge asymmetry is unfolded. The same definition, Eq. 18 as is given in Sec. 3.2.3 is utilized. In comparison with the charge asymmetry measured from $t\bar{t}$ pairs, leptonic asymmetry is more diluted because it relies solely on $\ell\bar{\ell}$ pair, which may not follow the same flight direction as their mother particles, top quarks. The results obtained from unfolding of dilepton data are drawn in Fig. 63. Similarly as in case of charge asymmetry, inclusive and differential measurements are performed. The differential variables used in these measurements are invariant mass of dilepton pair $m_{\ell\bar{\ell}}$, longitudinal boost of $\ell\bar{\ell}$ along the beam axis, $\beta_{z,\ell\bar{\ell}}$, and transverse momenta of $\ell\bar{\ell}$ pair, $p_{T,\ell\bar{\ell}}$. The SM prediction of leptonic asymmetry for all measurements is calculated at NLO in QCD and NLO in EW theory. The measured $A_C^{\ell\bar{\ell}}$ values are consistent with this prediction. The exact values of $A_C^{\ell\bar{\ell}}$ and their total uncertainties, together with the theoretical predictions, are given in App. A.6.



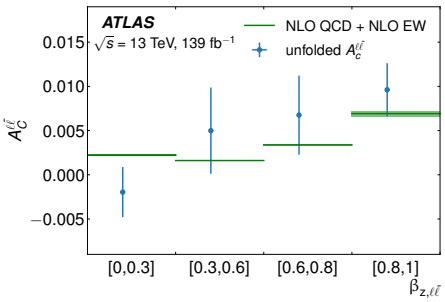
(a) Inclusive measurement



(b) $m_{\ell\bar{\ell}}$ measurement



(c) $p_{T,\ell\bar{\ell}}$ measurement



(d) $\beta_{z,\ell\bar{\ell}}$ measurement

Figure 63: Unfolded leptonic $A_C^{\ell\bar{\ell}}$ asymmetry using data from the dilepton channel for (a) inclusive measurement, (b) $m_{\ell\bar{\ell}}$, (c) $p_{T,\ell\bar{\ell}}$ and (d) $\beta_{z,\ell\bar{\ell}}$ differential measurements. The vertical lines show the total uncertainty. The unfolded results are compared with the theoretical prediction at the precision of NLO in QCD and NLO in EW theory [153]. The predictions is represented by green bands.

8.5 APPLICATION IN EFT

The unfolded A_C results are used to derive constraints on Wilson coefficients, which determine the coupling strength of EFT operators in the Lagrangian. Inclusive and $m_{t\bar{t}}$ differential measurements are exploited for this purpose.

Charge asymmetry prediction calculated in the EFT is provided by `SMEFT@NLO UFO` model [154] in the `MadGraph_aMC@NLO` package [122], which maintains the implementation of EFT operators. Charge asymmetry can be expressed as an analytical function of EFT operators [155] using

$$\sigma^{\pm} = \sigma^{\pm}_{SM} + \sum_i \frac{C_i}{\Lambda^2}\sigma^{\pm}_{1,i} + \sum_i \frac{C_i^2}{\Lambda^4}\sigma^{\pm}_{2,i} + \sum_{i,j} \frac{C_i C_j}{\Lambda^4}\sigma^{\pm}_{3,i,j}, \qquad (56)$$

assuming σ^+ denotes the cross section of events with positive $\Delta|y|$, $\sigma(\Delta|y| > 0)$, and analogously σ^- for events with negative $\Delta|y|$. Two fits of the measured results are performed - linear and quadratic. The linear fit takes into account only the terms proportional to the Λ^{-2} of the $\mathcal{L}_{eff}$, see Eq. 20. The quadratic fit includes also the terms corresponding to the scale Λ^{-4}, taking into consideration the effect of the squared dimension-six operators. The SM cross sections, $\sigma^{\pm}_{SM}$ are calculated at the NNLO in QCD [29], but other terms to the NLO accuracy in QCD.

Constraints on individual relevant Wilson coefficients, C_i, (see Sec. 3.2.4) are derived by minimization of χ^2 function. Only the inspected coefficient (and its term) is present in the function, coefficients corresponding to other EFT operators are set to zero. Covariance matrix used in the minimization takes into account experimental and modelling uncertainties (obtained from the unfolding) and theoretical uncertainties connected with the choice of renormalization and factorization scales in the calculation together with the choice of PDF set.

Limits found for the C^8_{tu} coefficient using each $m_{t\bar{t}}$ differential bin separately are plotted in Fig. 64. The most stringent bound among them is derived for the second highest $m_{t\bar{t}}$ bin, $m_{t\bar{t}} \in (1000, 1500)$ GeV, for the linear fit. However, truly the largest constraint is obtained after utilizing information from all $m_{t\bar{t}}$ differential bins. $A_C^{t\bar{t}}$ measurements performed in parts of phase space with high values of $m_{t\bar{t}}$ are more sensitive on the contribution from the EFT operators, therefore ideal for setting the constraints on Wilson coefficients. Results from previous EFT fits using 8 TeV LHC data [45] and data collected by Tevatron [38] are shown in the plot for reference. The constraint derived from the presented 13 TeV inclusive measurement is noticeably improved, mostly when comparing the linear-fit results with the results from previous measurements.

Minimization of quadratic function can result in two minima, therefore there are 2 data points for some differential bins.

The improvement is evident, although the asymmetry at 13 TeV is more polluted by the $t\bar{t}$ events produced by gluon fusion.

Charge asymmetry differential measurement as a function of $m_{t\bar{t}}$ is used to derive limits on Wilson coefficients corresponding to all four-

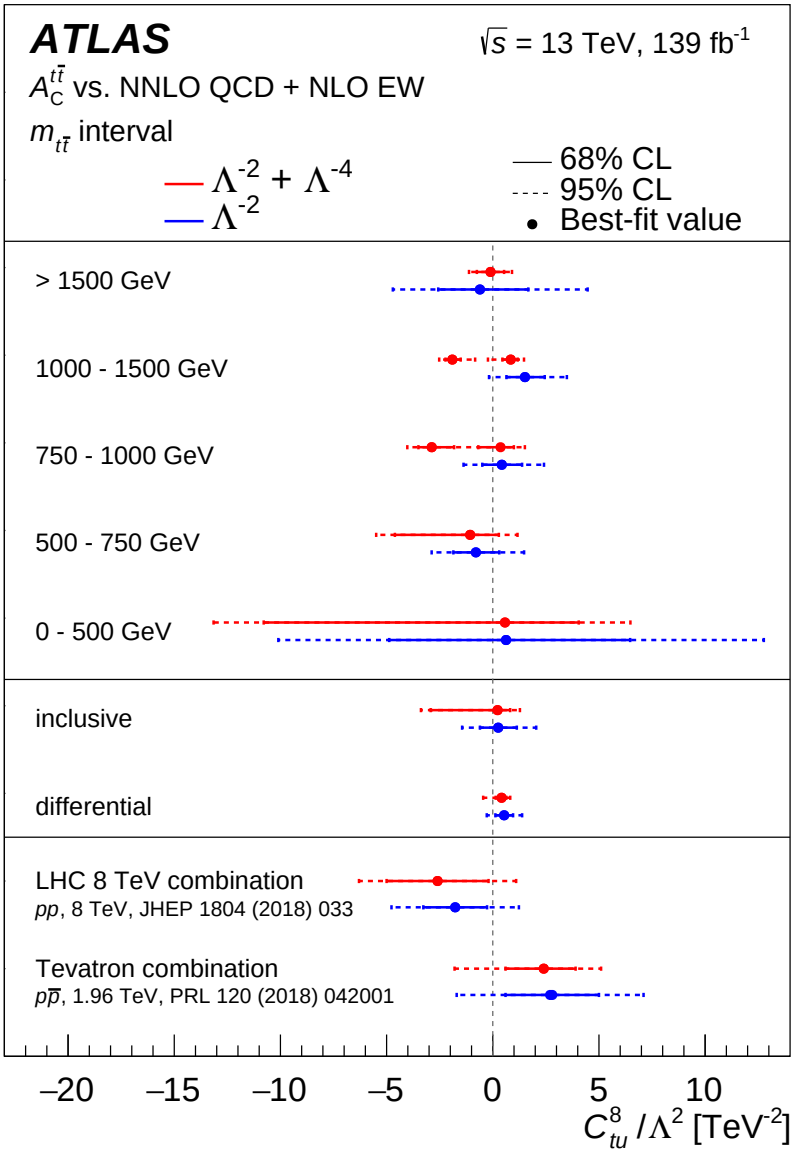


Figure 64: Limits on C_{tu}^{8} Wilson coefficient from inclusive and $m_{t\bar{t}}$ differential measurements. 95% and 68% CL limits from linear/quadratic fit are plotted as horizontal blue/red dashed and solid lines, respectively. The individual $m_{t\bar{t}}$ differential bins are used to derive constraints in the first part of the plot. The second part shows bounds derived from the inclusive measurement and using combined $m_{t\bar{t}}$ differential bins. The tightest limit is achieved by this combined constraint. For comparison, the constraints obtained using data collected by the LHC at 8 TeV and by Tevatron are shown.

fermion EFT operators and operator O_{tG}. Results are summarized in Fig. 65. The color-octet operators are more strongly constrained by the linear fit, because they affect the $t\bar{t}$ production directly at tree level.

Whereas color-singlet operators appear at NLO level, so the quadratic fit provides tighter limits for these [156].

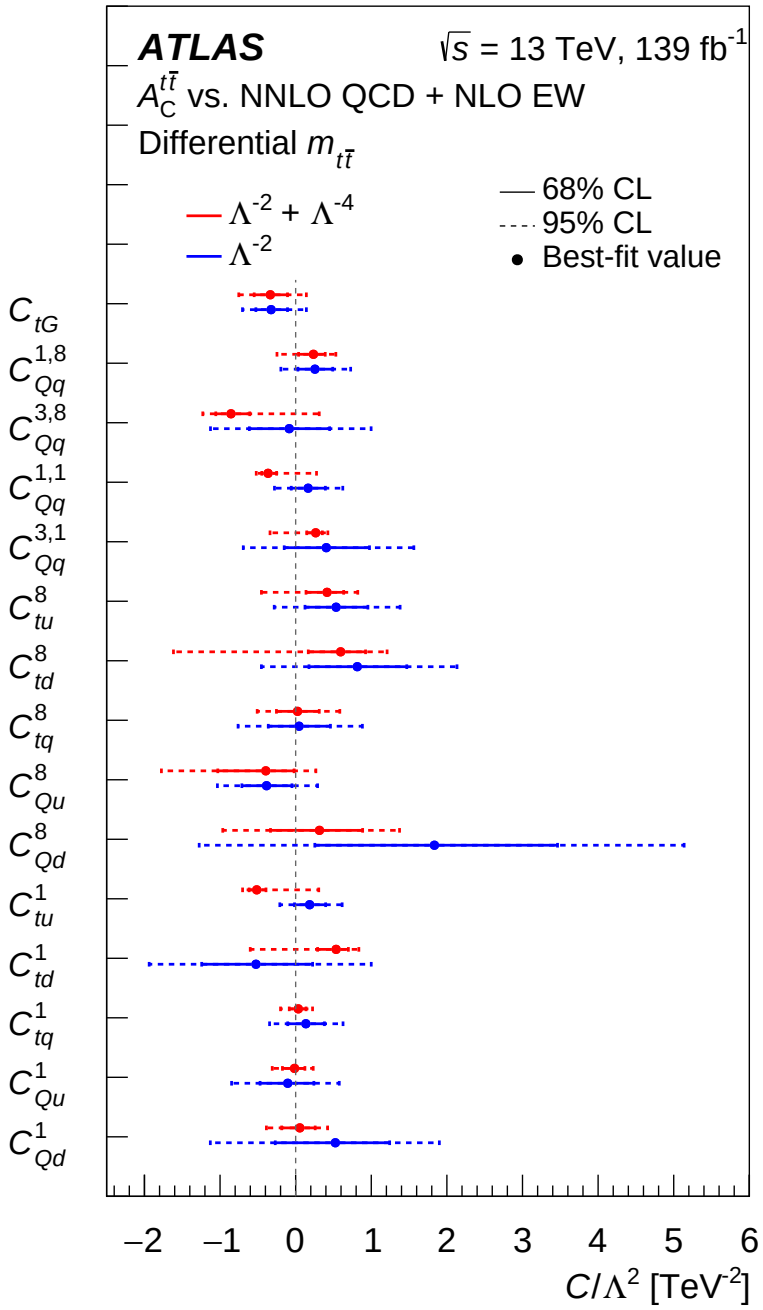


Figure 65: Bounds derived for relevant four-fermion EFT operators and operator O_{tG}. 95% and 68% CL limits from linear fit are marked by blue dashed and solid lines, respectively. Similarly, constraints from quadratic fit are depicted as red dashed lines for 95% CL and solid lines for 68% CL.

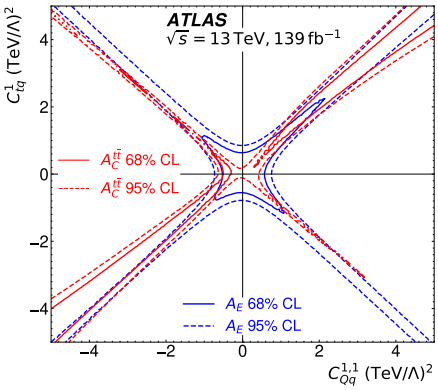
8.5.1 *Charge asymmetry and energy asymmetry*

Measurement of the energy asymmetry in $t\bar{t}+j$ production [157] is considered to be complementary to the charge asymmetry measurement in context of EFT interpretation of their results. The production of $t\bar{t}$ and additional jet includes different terms in QCD amplitude than $t\bar{t}$ production, so the two asymmetries are differently sensitive to the contributions from the EFT operators of different color and chiral type. Interpretation of bounds on Wilson coefficients from both these measurements together benefits from different sensitivity and hence can produce bounds in various directions in the EFT-parameter space.

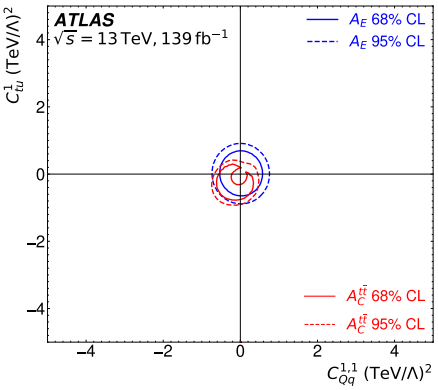
The differences between constraints derived from the charge and energy asymmetry are illustrated by plots in Fig. 66 using quadratic fit considering terms up to Λ^{-4}. Firstly, the sensitivity in the parameter space defined by two color-singlet EFT operators, which differ in the chirality of quark fields, is probed. In Fig. 66a, the bounds for operators of chiral type $(\bar{L}L)(\bar{L}L)$ and $(\bar{R}R)(\bar{L}L)$ are derived from differential $m_{t\bar{t}}$ charge asymmetry measurement and energy asymmetry measurement. Similarly, color-singlet operators of different chirality type, $(\bar{L}L)(\bar{L}L)$ and $(\bar{R}R)(\bar{R}R)$, are inspected and derived limits are drawn in Fig. 66b. In this case, the bounds have similar shapes for both measurements, stating quite the same level of sensitivity for color-singlet operators differing in the chirality.

The situation is considerably different when investigating color octet operators, when again, probing various chirality scenarios. Looking at Fig. 66c and Fig. 66d, one can observe quite distinctive shapes derived from charge and energy asymmetry measurements. Specifically, in Fig 66c, the charge asymmetry measurement does not produce any constraint for the part of EFT-parameter space defined by positive values of C_{tq}^{8} and negative values of $C_{Qq}^{1,8}$. On the other hand, the energy asymmetry measurement is able to give some limit, hence complements the bound obtained from the charge asymmetry measurement quite nicely.

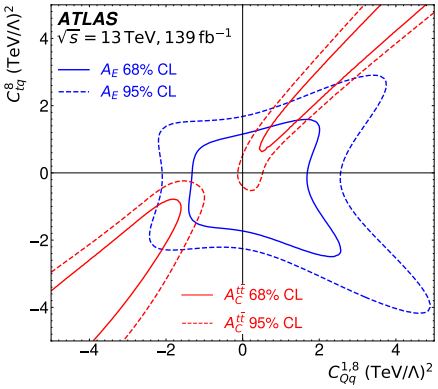
Lastly, the couple of color-singlet and color-octet operators is inspected to complete the picture. The constraints for color-singlet and color-octet coefficients with the same quark chiralities, $(\bar{R}R)(\bar{L}L)$ or $(\bar{L}L)((\bar{L}L)$, are depicted in Fig. 66e and Fig. 66f, respectively. Again, the asymmetry measurements produce different shapes of bounds, which reflects differences in calculation of $t\bar{t}$ and $t\bar{t}+j$ production regarding EFT operators of different color structure.



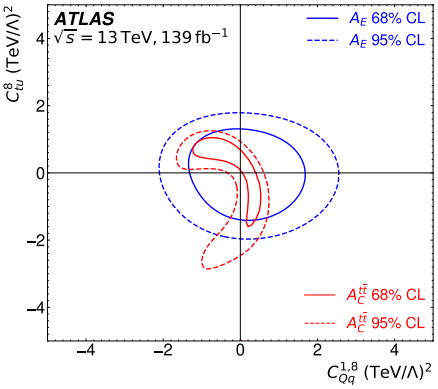
(a) Color-singlet LL and RL operators



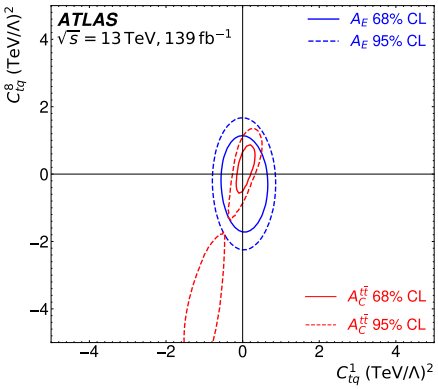
(b) Color-singlet LL and RR operators



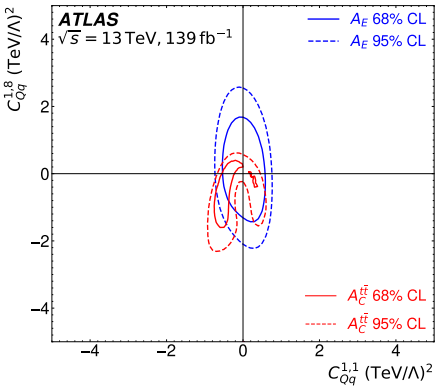
(c) Color-octet LL and RL operators



(d) Color-octet LL and RR operators



(e) Color-singlet and color-octet RL operators



(f) Color-singlet and color-octet LL operators

Figure 66: Bounds on various combinations of Wilson coefficients displaying sensitivity of charge and energy asymmetry measurements. Plots in each row show constraints for representative combination of Wilson coefficients corresponding to various types of EFT operators in terms of color and chirality. LL is an abbreviation for $(\bar{L}L)(\bar{L}L)$, RL for $(\bar{R}R)(\bar{L}L)$ and RR for $(\bar{R}R)(\bar{R}R)$.

9 CONCLUSION

A phenomenon present in $t\bar{t}$ production, charge asymmetry A_C, has been studied using data collected by the ATLAS detector at the center-of-mass energy of 13 TeV, which corresponds to integrated luminosity of 139 fb^{-1}. The charge asymmetry measurement in combination of single-lepton and dilepton $t\bar{t}$ decay channel has been performed. Full analysis chain is described and the newest techniques for handling of systematic uncertainties are introduced and tested.

Improvements in treatment of signal modelling uncertainties have been obtained by implementation of specific de-correlation scheme and application of truth-based re-weighting. Many different systematic scenarios have been tested in order to find the most optimal setting, which reduces constraints of systematic uncertainties and does not significantly influence the A_C uncertainty values. Byproducts of smaller constraints are larger total uncertainties of measurements. Nevertheless, the presented inclusive measurement is the most precise inclusive $A_C^{t\bar{t}}$ measurement ever performed.

Charge asymmetry measured inclusively using the full data set from both $t\bar{t}$ decay channels, single-lepton and dilepton, is in good agreement with the SM prediction calculated at NNLO in QCD and NLO in EW theory. Moreover, for the first time, the measured value, 0.0068 ± 0.0015, is inconsistent with zero charge asymmetry hypothesis at the level of 4.7 σ. Up to now, all measurements using data collected at lower center-of-mass energy, produced A_C values more compatible with zero within their uncertainties.

The charge asymmetry has been studied as a function of different kinematic observables: $m_{t\bar{t}}$, $\beta_z^{t\bar{t}}$, $p_T^{t\bar{t}}$. $A_C^{t\bar{t}}$ values obtained in all differential bins of each variable are compatible with the SM calculation, similarly as for the inclusive measurement.

In the dilepton channel, leptonic charge asymmetry $A_C^{\ell\bar{\ell}}$ has been studied inclusively and also as a function of various observables like $m_{\ell\bar{\ell}}$, $\beta_{z,\ell\bar{\ell}}$ and $p_{T,\ell\bar{\ell}}$. Unfolded $A_C^{\ell\bar{\ell}}$ values are in agreement with the SM prediction calculated at NLO in QCD and NLO in EW theory.

Although the same level of precision is not achieved in differential $A_C^{t\bar{t}}$ measurements, obtained results still give substantial information. Specifically, the results from $m_{t\bar{t}}$ differential measurement are exploited for derivation of bounds on Wilson coefficients corresponding to four-fermion EFT operators and $\mathcal{O}_{tG}$ operator. The complementarity of energy asymmetry and $A_C^{t\bar{t}}$ measurements in aspect of the EFT interpretation of their results, is utilized. Limits derived from both measurements are plotted together for more complete picture

of EFT operators' limits. Furthermore, great potential lies in combination of charge asymmetry, $t\bar{t}$ cross section, and other measurements in order to provide global fits of EFT operator space. The combined constraints are expected to be more stringent than bounds derived utilizing individual measurements.

Part V

APPENDIX

APPENDIX

A.1 NNLO RE-WEIGHTING

An illustration of the influence of the NNLO re-weighting is displayed in Fig. 67. In this Figure, the NNLO re-weighted systematic variation corresponding to the choice of PS+hadronization model and its previous variation are plotted.

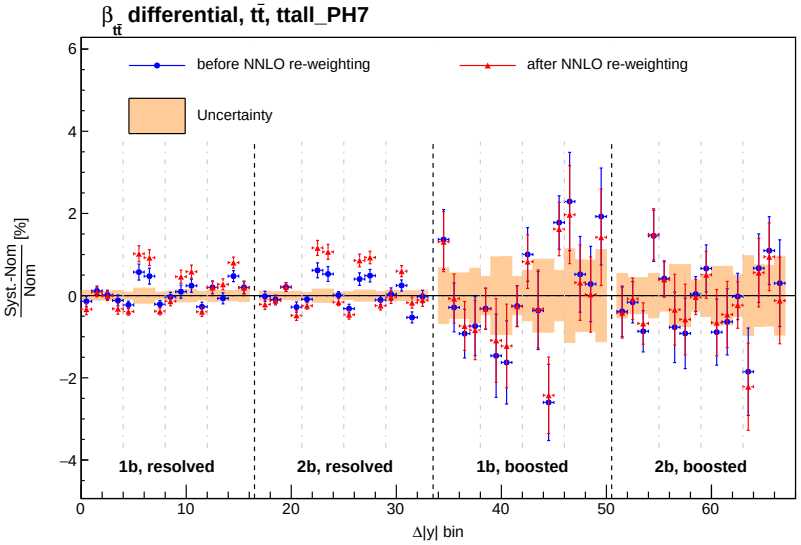
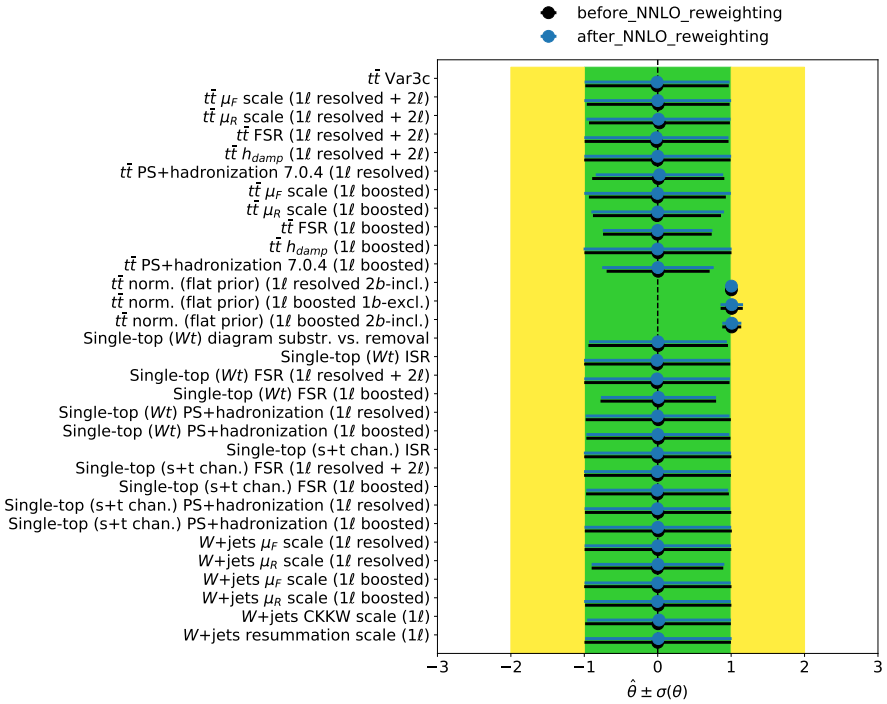
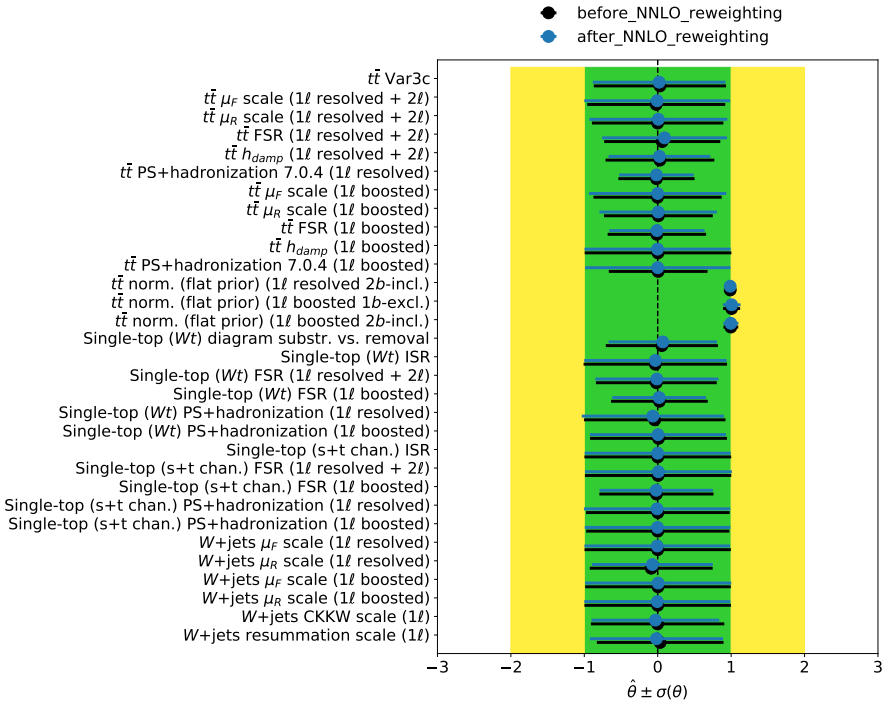


Figure 67: Systematic variation corresponding to the choice of PS+hadronization model before and after application of NNLO re-weighting for $\beta_z^{t\bar{t}}$ differential measurement. In resolved regions, the size of the variations is a little larger, whereas in boosted regions, the actual values get a little smaller. The variation in boosted region after the NNLO re-weighting is small enough to pass the pruning criteria, which results in zero constraint.

In Fig. 68, constraints obtained for systematic variations in inclusive and $\beta_z^{t\bar{t}}$ differential measurements are shown. Two sets of constraints can be compared, for NPs before and after applying of the NNLO re-weighting. Observed differences suggest how the NNLO re-weighting influences individual systematic variations.



(a) inclusive measurement

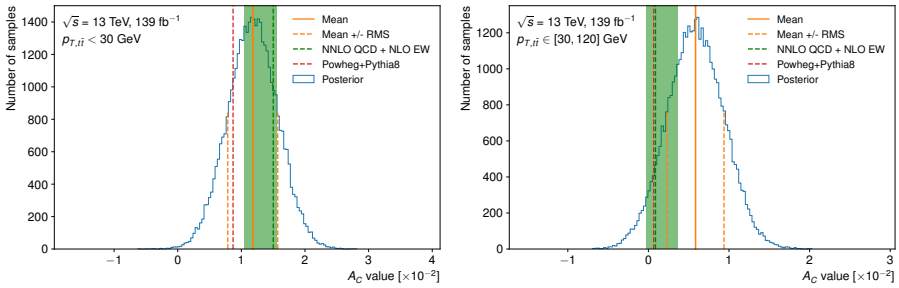


(b) $\beta_z^{t\bar{t}}$ differential measurement

Figure 68: Comparison of constraints for the inclusive (a) and $\beta_z^{t\bar{t}}$ (b) measurement using inputs with/without the NNLO re-weighting applied. Results from the Asimov data set are shown. For inclusive measurement, the constraints are very similar for both setups. However, in case of $\beta_z^{t\bar{t}}$ differential measurement, the constraints of some systematic uncertainties are smaller for the NNLO re-weighted variations. The PS+hadronization uncertainty in boosted region is not constrained, because the NNLO re-weighted systematic variation is removed by pruning.

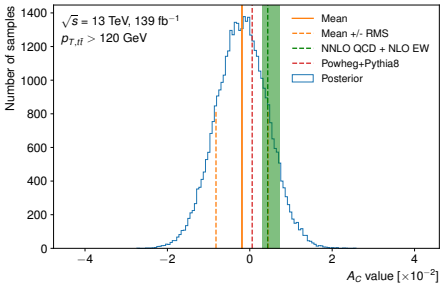
A.2 POSTERIOR A_C DISTRIBUTIONS

Posterior A_C distribution for $p_T^{t\bar{t}}$ and $\beta_z^{t\bar{t}}$ differential measurements are displayed in Fig. 69 and Fig. 70, respectively. Mean value of the posterior distribution is considered as unfolded A_C value for that particular differential bin, whereas width of the distribution is taken as the unfolded uncertainty. The SM predictions calculated at NNLO in QCD and NLO in EW theory are depicted as green bands.



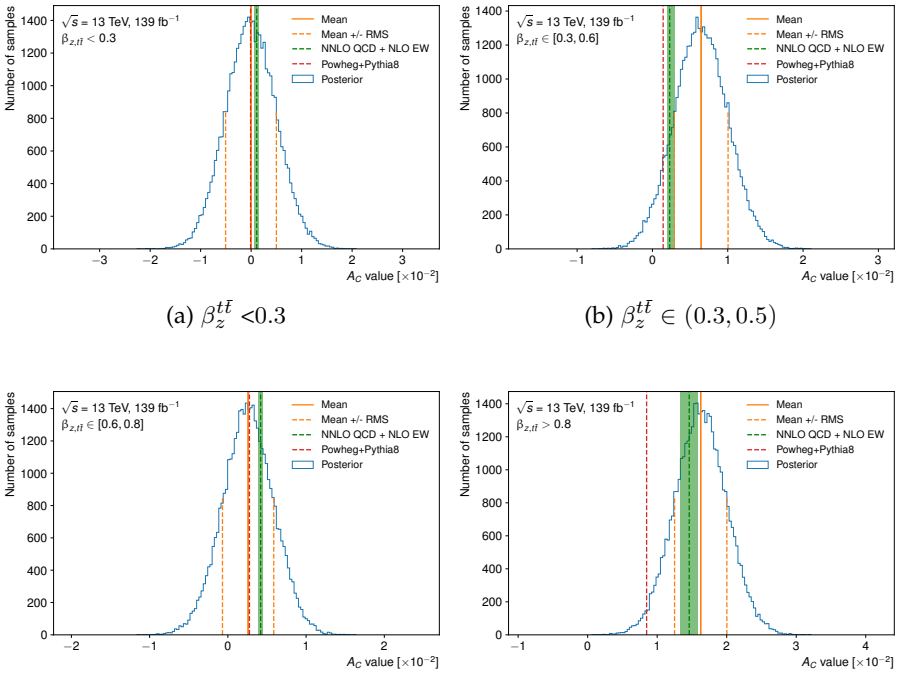
(a) $p_T^{t\bar{t}}$ <30 GeV

(b) $p_T^{t\bar{t}} \in (30, 120)$ GeV



(c) $p_T^{t\bar{t}} > 120$ GeV

Figure 69: Posterior A_C distributions for $p_T^{t\bar{t}}$ differential measurement. The values of A_C and their unfolded uncertainties are depicted in orange solid and dashed lines. Charge asymmetry prediction given by the POWHEG+PYTHIA 8 generators is represented by red dashed line, while the prediction calculated at the NNLO in QCD and NLO in EW is displayed as green band.



(a) $\beta_z^{t\bar{t}}$ <0.3

(b) $\beta_z^{t\bar{t}} \in (0.3, 0.5)$

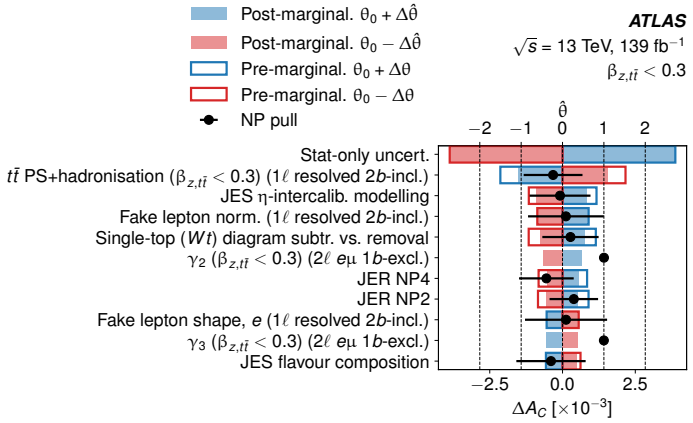
(c) $\beta_z^{t\bar{t}} \in (0.5, 0.8)$

(d) $\beta_z^{t\bar{t}} \in (0.8, 1)$

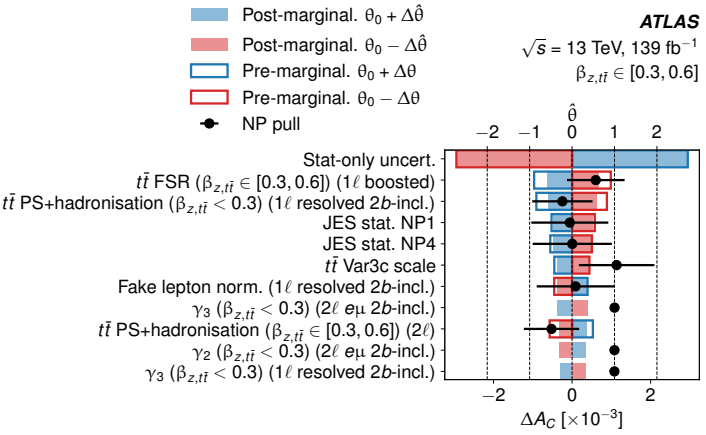
Figure 70: Posterior A_C distributions for $\beta_z^{t\bar{t}}$ differential measurement. The values of A_C and their unfolded uncertainties are depicted in orange solid and dashed lines. Charge asymmetry prediction given by the POWHEG+PYTHIA 8 generators is represented by red dashed line, while the prediction calculated at the NNLO in QCD and NLO in EW is displayed as green band.

A.3 RANKING OF SYSTEMATIC UNCERTAINTIES

Ten highest ranked NPs for each differential bin of charge asymmetry measurements are shown in Fig. 71 and Fig. 72 for $\beta_z^{t\bar{t}}$ differential measurement and in Figs. 73, 74, 75a for $m_{t\bar{t}}$ differential measurement. Lastly, the ranking for differential measurement of A_C as a function of $p_T^{t\bar{t}}$ is depicted in Fig. 75b and Fig. 76. Description of the ranking procedure is given in Sec. 7.6.

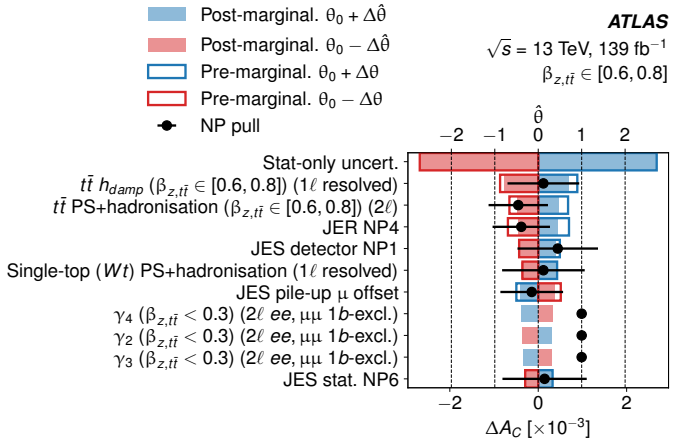


(a) $0 < \beta_z^{t\bar{t}} < 0.3$

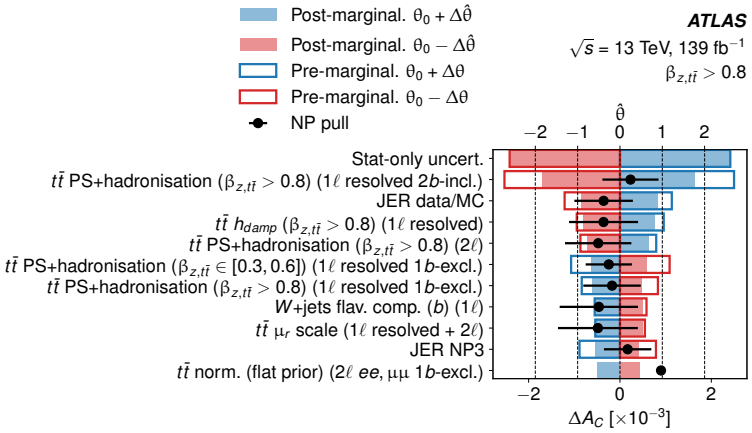


(b) $0.3 < \beta_z^{t\bar{t}} < 0.6$

Figure 71: Ranking of the systematic uncertainties in the $\beta_z^{t\bar{t}}$ differential measurement (a) for $\beta_z^{t\bar{t}} \in (0.0, 0.3)$ and (b) for $\beta_z^{t\bar{t}} \in (0.3, 0.6)$ for the measured data. Top 10 highest ranked systematic uncertainties are shown in the plot. The order is defined by the size of the effect each of systematic uncertainty has on the A_C value by shifting the variation by their constraint (post-marginalization ranking). Blue and red bars show the effect of upward and downward systematic variation on the A_C value, respectively. The frames and filled bars denote the pre-marginalization and post-marginalization results of the ranking, respectively. Black data points display the constraints and pulls of the ranked systematic uncertainties.

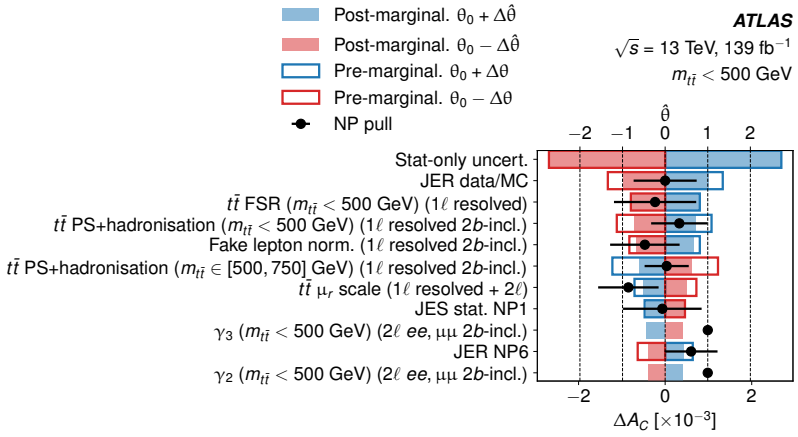


(a) $0.6 < \beta_z^{t\bar{t}} < 0.8$

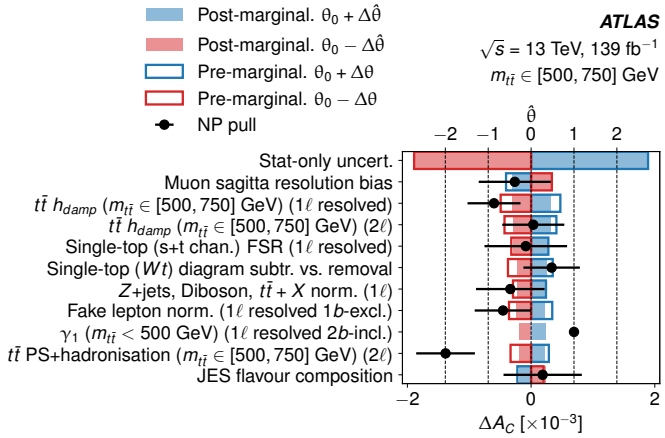


(b) $0.8 < \beta_z^{t\bar{t}} < 1.0$

Figure 72: Ranking of the systematic uncertainties in the $\beta_z^{t\bar{t}}$ differential measurement (a) for $\beta_z^{t\bar{t}} \in (0.6, 0.8)$ and (b) for $\beta_z^{t\bar{t}} \in (0.8, 1.0)$ for the measured data. Top 10 highest ranked systematic uncertainties are shown in the plot. The order is defined by the size of the effect each of systematic uncertainty has on the A_C value by shifting the variation by its constraint (post-marginalization ranking). Blue and red bars show the effect of upward and downward systematic variation on the A_C value, respectively. The frames and filled bars denote the pre-marginalization and post-marginalization results of the ranking, respectively. Black data points display the constraints and pulls of the ranked systematic uncertainties.

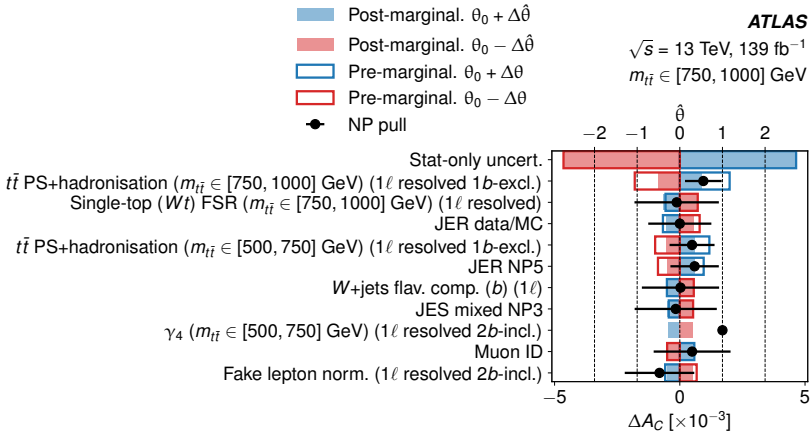


(a) $m_{t\bar{t}} < 500$ GeV

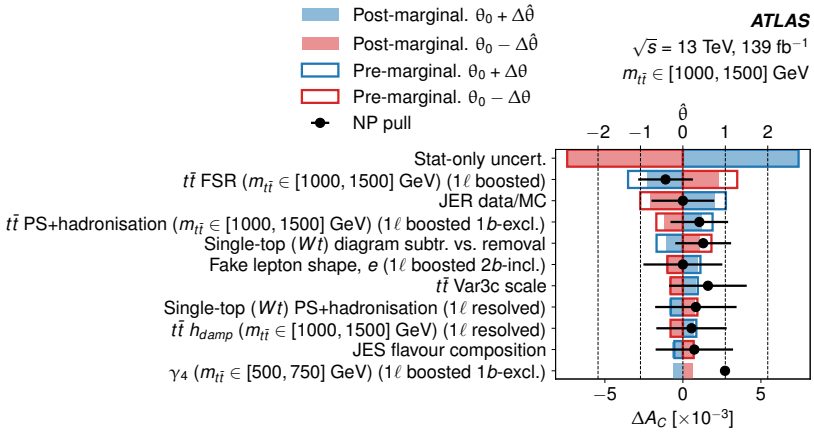


(b) $500 < m_{t\bar{t}} < 750$ GeV

Figure 73: Ranking of the systematic uncertainties in the $m_{t\bar{t}}$ differential measurement (a) for $m_{t\bar{t}} < 500$ GeV and (b) for $m_{t\bar{t}} \in (500, 750)$ GeV for the measured data. Top 10 highest ranked systematic uncertainties are shown in the plot. The order is defined by the size of the effect each of systematic uncertainty has on the A_C value by shifting the variation by its constraint (post-marginalization ranking). Blue and red bars show the effect of upward and downward systematic variation on the A_C value, respectively. The frames and filled bars denote the pre-marginalization and post-marginalization results of the ranking, respectively. Black data points display the constraints and pulls of the ranked systematic uncertainties.

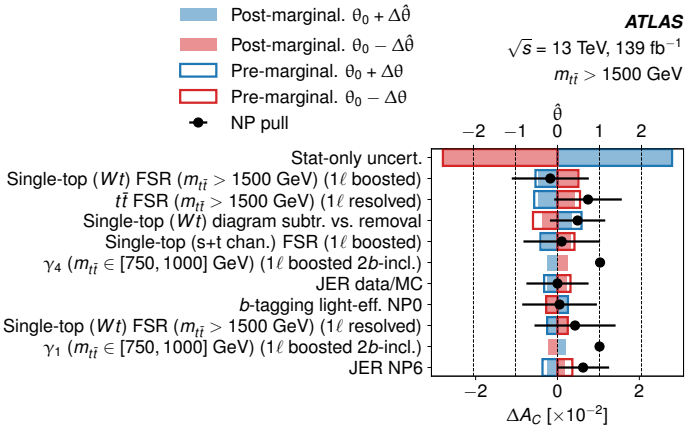


(a) $750 < m_{t\bar{t}} < 1000$ GeV

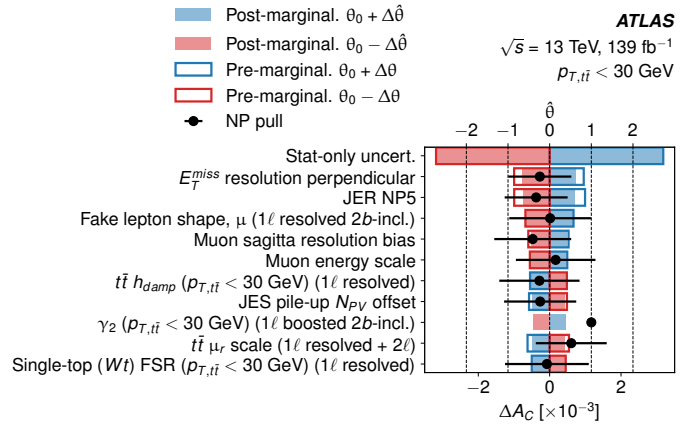


(b) $1000 < m_{t\bar{t}} < 1500$ GeV

Figure 74: Ranking of the systematic uncertainties in the $m_{t\bar{t}}$ differential measurement (a) for $m_{t\bar{t}} \in (750, 1000)$ GeV and (b) for $m_{t\bar{t}} \in (1000, 1500)$ GeV for the measured data. Top 10 highest ranked systematic uncertainties are shown in the plot. The order is defined by the size of the effect each of systematic uncertainty has on the A_C value by shifting the variation by its constraint (post-marginalization ranking). Blue and red bars show the effect of upward and downward systematic variation on the A_C value, respectively. The frames and filled bars denote the pre-marginalization and post-marginalization results of the ranking, respectively. Black data points display the constraints and pulls of the ranked systematic uncertainties.

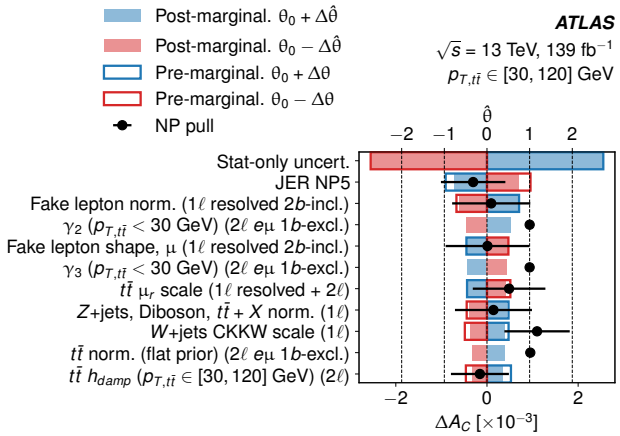


(a) $m_{t\bar{t}} > 1500$ GeV

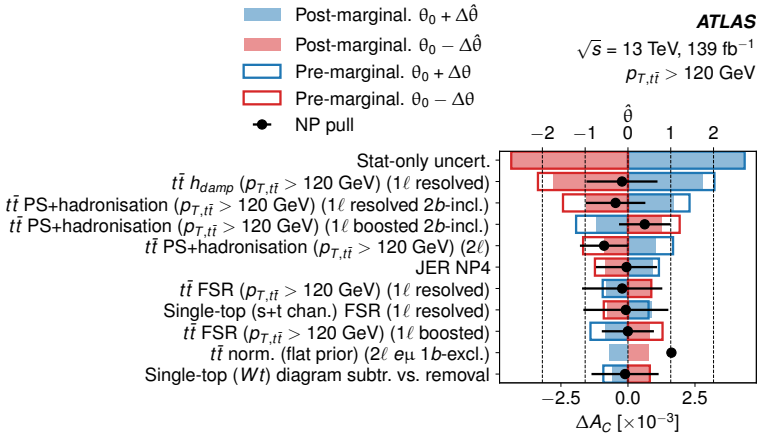


(b) $p_T^{t\bar{t}} < 30$ GeV

Figure 75: Ranking of the systematic uncertainties in the $m_{t\bar{t}}$ and $p_T^{t\bar{t}}$ differential measurement (a) for $m_{t\bar{t}} > 1500$ GeV and (b) for $p_T^{t\bar{t}} < 30$ GeV for the measured data. Top 10 highest ranked systematic uncertainties are shown in the plot. The order is defined by the size of the effect each of systematic uncertainty has on the A_C value by shifting the variation by its constraint (post-marginalization ranking). Blue and red bars show the effect of upward and downward systematic variation on the A_C value, respectively. The frames and filled bars denote the pre-marginalization and post-marginalization results of the ranking, respectively. Black data points display the constraints and pulls of the ranked systematic uncertainties.



(a) $30 < p_T^{t\bar{t}} < 120$ GeV



(b) $p_T^{t\bar{t}} > 120$ GeV

Figure 76: Ranking of the systematic uncertainties in the $p_T^{t\bar{t}}$ differential measurement (a) for $p_T^{t\bar{t}} \in (30, 120)$ GeV and (b) for $p_T^{t\bar{t}} > 120$ GeV for the measured data. Top 10 highest ranked systematic uncertainties are shown in the plot. The order is defined by the size of the effect each of systematic uncertainty has on the A_C value by shifting the variation by its constraint (post-marginalization ranking). Blue and red bars show the effect of upward and downward systematic variation on the A_C value, respectively. The frames and filled bars denote the pre-marginalization and post-marginalization results of the ranking, respectively. Black data points display the constraints and pulls of the ranked systematic uncertainties.

A.4 BREAKDOWN OF UNCERTAINTIES

The total uncertainty of charge asymmetry measurement is a composite of different partial uncertainties. These uncertainties are: the unfolded uncertainty, which is denoted as 'stat+syst', uncertainty accounting for the bias of the unfolding response and lastly, uncertainty of the response matrix, which comes from the limited statistics. Exact values of all contributions to the total uncertainty for inclusive and differential A_C measurements are given in Tab. 20 and Tab. 21.

Table 20: The breakdown of the uncertainties contributing to the total uncertainties for inclusive and $m_{t\bar{t}}$ differential measurements. The unfolded uncertainties from the unfolding with incorporated systematic uncertainties (stat.+syst.) are listed in the first column. In the next column, the bias of the unfolding procedure is estimated using unfolded A_C values from data. The third column contains the uncertainties corresponding to the effect of finite-size MC samples on the response matrix determination. Lastly, the total uncertainties calculated as a sum of squares of previously stated uncertainties are given in the last column.

		Uncertainty			
	Channel	Stat+Syst	Bias	Response matrix	Total unc.
Inclusive	1ℓ	0.0015	< 0.0001	0.0004	0.0015
	2ℓ	0.0048	< 0.0001	0.0007	0.0049
	$1\ell + 2\ell$	0.0014	< 0.0001	0.0004	0.0015
$m_{t\bar{t}} < 500$ GeV	1ℓ	0.0038	0.0002	0.0010	0.0039
	2ℓ	0.0139	0.0006	0.0024	0.0141
	$1\ell + 2\ell$	0.0035	0.0002	0.0010	0.0036
$m_{t\bar{t}} \in (500, 750)$ GeV	1ℓ	0.0024	-0.0002	0.0007	0.0025
	2ℓ	0.0089	-0.0003	0.0012	0.0089
	$1\ell + 2\ell$	0.0022	-0.0001	0.0007	0.0023
$m_{t\bar{t}} \in (750, 1000)$ GeV	1ℓ	0.0060	0.0004	0.0017	0.0062
	2ℓ	0.0220	0.0016	0.0037	0.0223
	$1\ell + 2\ell$	0.0053	0.0005	0.0016	0.0056
$m_{t\bar{t}} \in (1000, 1500)$ GeV	1ℓ	0.0087	-0.0007	0.0025	0.0090
	2ℓ	0.0438	-0.0033	0.0066	0.0444
	$1\ell + 2\ell$	0.0083	-0.0006	0.0024	0.0087
$m_{t\bar{t}} > 1500$ GeV	1ℓ	0.0288	-0.0005	0.0076	0.0298
	2ℓ	0.1535	0.0116	0.0253	0.1560
	$1\ell + 2\ell$	0.0277	-0.0003	0.0078	0.0288

Table 21: The breakdown of the uncertainties contributing to the total uncertainties for $p_T^{t\bar{t}}$ and $\beta_z^{t\bar{t}}$ differential measurements. The unfolded uncertainties from the unfolding with incorporated systematic uncertainties (stat.+syst.) are listed in the first column. In the next column, the bias of the unfolding procedure is estimated using unfolded A_C values from data. The third column contains the uncertainties corresponding to the effect of finite-size MC samples on the response matrix determination. Lastly, the total uncertainties calculated as a sum of squares of previously stated uncertainties are given in the last column.

		Uncertainty			
	Channel	Stat+Syst	Bias	Response matrix	Total unc.
$p_T^{t\bar{t}} < 30$ GeV	1ℓ	0.0044	< 0.0001	0.0013	0.0046
	2ℓ	0.0145	0.0012	0.0023	0.0147
	$1\ell + 2\ell$	0.0039	0.0001	0.0011	0.0041
$p_T^{t\bar{t}} \in (30, 120)$ GeV	1ℓ	0.0037	-0.0001	0.0011	0.0038
	2ℓ	0.0139	-0.0011	0.0020	0.0141
	$1\ell + 2\ell$	0.0035	-0.0000	0.0009	0.0037
$p_T^{t\bar{t}} > 120$ GeV	1ℓ	0.0078	0.0000	0.0018	0.0080
	2ℓ	0.0135	0.0007	0.0024	0.0137
	$1\ell + 2\ell$	0.0063	-0.0002	0.0014	0.0065
$\beta_z^{t\bar{t}} \in (0, 0.3)$	1ℓ	0.0054	-0.0002	0.0015	0.0056
	2ℓ	0.0212	0.0002	0.0038	0.0215
	$1\ell + 2\ell$	0.0050	-0.0004	0.0015	0.0052
$\beta_z^{t\bar{t}} \in (0.3, 0.6)$	1ℓ	0.0038	< 0.0001	0.0011	0.0040
	2ℓ	0.0137	0.0001	0.0023	0.0139
	$1\ell + 2\ell$	0.0036	0.0002	0.0011	0.0037
$\beta_z^{t\bar{t}} \in (0.6, 0.8)$	1ℓ	0.0035	-0.0001	0.0010	0.0037
	2ℓ	0.0116	-0.0005	0.0019	0.0118
	$1\ell + 2\ell$	0.0033	-0.0002	0.0009	0.0034
$\beta_z^{t\bar{t}} \in (0.8, 1.0)$	1ℓ	0.0045	0.0001	0.0010	0.0046
	2ℓ	0.0084	0.0001	0.0015	0.0085
	$1\ell + 2\ell$	0.0038	0.0001	0.0008	0.0039

A.5 STATISTICAL VERSUS SYSTEMATIC UNCERTAINTY

Comparison of uncertainties, which are obtained from unfolding without and with the inclusion of systematic NPs are given in Tab. 22 and Tab. 23 for inclusive and all differential measurements. In the last column, there is an estimation of the statistical power of each differential bin/inclusive measurement. In most bins, the statistics is more dominant source of the uncertainty than systematic uncertainties.

Table 22: Comparison of uncertainties obtained from unfolding without/with systematic NPs included for inclusive and $m_{t\bar{t}}$ differential measurements. In the last column, there is an estimation of the statistical power of each differential bin/inclusive measurement. In most of the bins, the statistics is more dominant source of the uncertainty than systematic uncertainties.

		Uncertainty		
	Channel	Stat.	Stat.+syst.	$\left(\frac{\sigma_{stat}}{\sigma_{stat+syst}}\right)^2$ [%]
Inclusive	1ℓ	0.0011	0.0015	56.1
	2ℓ	0.0034	0.0048	48.3
	$1\ell + 2\ell$	0.0010	0.0014	54.0
$m_{t\bar{t}} < 500$ GeV	1ℓ	0.0028	0.0038	52.0
	2ℓ	0.0114	0.0139	66.5
	$1\ell + 2\ell$	0.0027	0.0035	60.2
$m_{t\bar{t}} \in (500, 750)$ GeV	1ℓ	0.0020	0.0024	68.6
	2ℓ	0.0061	0.0089	47.2
	$1\ell + 2\ell$	0.0019	0.0022	74.1
$m_{t\bar{t}} \in (750, 1000)$ GeV	1ℓ	0.0048	0.0060	64.0
	2ℓ	0.0188	0.0220	73.3
	$1\ell + 2\ell$	0.0046	0.0053	77.4
$m_{t\bar{t}} \in (1000, 1500)$ GeV	1ℓ	0.0075	0.0087	75.8
	2ℓ	0.0371	0.0438	71.8
	$1\ell + 2\ell$	0.0074	0.0083	79.1
$m_{t\bar{t}} > 1500$ GeV	1ℓ	0.0288	0.0288	100.1
	2ℓ	0.1444	0.1535	88.5
	$1\ell + 2\ell$	0.0280	0.0277	102.0

In the last $m_{t\bar{t}}$ differential bin, the statistical uncertainty is even larger than the uncertainty obtained from unfolding with NPs. It signifies that the systematic uncertainties play only negligible role. The excess of the statistical uncertainty is still within the bias uncertainty

of the unfolding (Tab. 20) compatible with the stat.+syst. uncertainty value. Moreover, the uncertainty originating from the limited statistic of the response matrix is even one order larger (Tab. 20), hence these fluctuations are not significant.

Table 23: Comparison of uncertainties obtained from unfolding without/with systematic NPs included for $p_T^{t\bar{t}}$ and $\beta_z^{t\bar{t}}$ differential measurements. In the last column, there is an estimation of the statistical power of each differential bin/inclusive measurement. In most bins, the statistics is more dominant source of the uncertainty than systematic uncertainties.

		Uncertainty		
	Channel	Stat.	Stat.+Syst.	$\left(\frac{\sigma_{stat}}{\sigma_{stat+syst}}\right)^2$ [%]
$p_T^{t\bar{t}} < 30$ GeV	1ℓ	0.0034	0.0044	59.0
	2ℓ	0.0105	0.0145	52.7
	$1\ell + 2\ell$	0.0032	0.0039	65.8
$p_T^{t\bar{t}} \in (30, 120)$ GeV	1ℓ	0.0027	0.0037	53.1
	2ℓ	0.0096	0.0139	47.7
	$1\ell + 2\ell$	0.0026	0.0035	52.2
$p_T^{t\bar{t}} > 120$ GeV	1ℓ	0.0047	0.0078	37.2
	2ℓ	0.0122	0.0135	81.9
	$1\ell + 2\ell$	0.0044	0.0063	47.7
$\beta_z^{t\bar{t}} \in (0, 0.3)$	1ℓ	0.0040	0.0054	56.4
	2ℓ	0.0179	0.0212	71.5
	$1\ell + 2\ell$	0.0039	0.0050	59.5
$\beta_z^{t\bar{t}} \in (0.3, 0.6)$	1ℓ	0.0031	0.0038	65.1
	2ℓ	0.0116	0.0137	71.6
	$1\ell + 2\ell$	0.0029	0.0036	68.0
$\beta_z^{t\bar{t}} \in (0.6, 0.8)$	1ℓ	0.0028	0.0035	64.2
	2ℓ	0.0090	0.0116	60.0
	$1\ell + 2\ell$	0.0027	0.0033	66.8
$\beta_z^{t\bar{t}} \in (0.8, 1.0)$	1ℓ	0.0026	0.0045	33.1
	2ℓ	0.0069	0.0084	67.1
	$1\ell + 2\ell$	0.0024	0.0038	41.3

A.6 LEPTONIC CHARGE ASYMMETRY

In Tab. 24, summary of $A_C^{\ell\bar{\ell}}$ values obtained from unfolding of dilepton data are presented. The table complements the information given in Sec. 8.4. Results of inclusive and differential measurements are included together with the total uncertainties, which are calculated in the same way as is described in Sec. 8.3. Differential measurements are performed as a function of $m_{\ell\bar{\ell}}$, $\beta_{z,\ell\bar{\ell}}$ and $p_{T,\ell\bar{\ell}}$.

Table 24: Leptonic charge asymmetry values in inclusive and differential $A_C^{\ell\bar{\ell}}$ measurements, together with the total uncertainties. The theoretical SM prediction calculated at NLO in QCD and NLO in EW theory is given in the last column.

	Data 139 fb^{-1}		SM prediction
	$A_C^{\ell\bar{\ell}}$	Total unc.	
Inclusive	0.0054	0.0026	$0.0040^{+0.0002}_{-0.0001}$
$m_{\ell\bar{\ell}} < 200$ GeV	0.0036	0.0027	$0.0033^{+0.0001}_{-0.0001}$
$m_{\ell\bar{\ell}} \in [200, 300]$ GeV	0.0013	0.0046	$0.0084^{+0.0002}_{-0.0001}$
$m_{\ell\bar{\ell}} \in [300, 400]$ GeV	0.0146	0.0090	$0.0108^{+0.0003}_{-0.0006}$
$m_{\ell\bar{\ell}} > 400$ GeV	0.0205	0.0128	$0.0120^{+0.0009}_{-0.0002}$
$p_{T,\ell\bar{\ell}} < 20$ GeV	0.0176	0.0070	$0.0026^{+0.0002}_{-0.0002}$
$p_{T,\ell\bar{\ell}} \in [20, 70]$ GeV	0.0055	0.0034	$0.0034^{+0.0001}_{<0.0001}$
$p_{T,\ell\bar{\ell}} > 70$ GeV	0.0041	0.0027	$0.0050^{+0.0002}_{-0.0002}$
$\beta_{z,\ell\bar{\ell}} \in [0, 0.3]$	-0.0020	0.0029	$0.0022^{+0.0001}_{-0.0001}$
$\beta_{z,\ell\bar{\ell}} \in [0.3, 0.6]$	0.0050	0.0049	$0.0016^{+0.0001}_{<0.0001}$
$\beta_{z,\ell\bar{\ell}} \in [0.6, 0.8]$	0.0068	0.0045	$0.0034^{<0.0001}_{-0.0001}$
$\beta_{z,\ell\bar{\ell}} \in [0.8, 1.0]$	0.0096	0.0031	$0.0069^{+0.0003}_{-0.0003}$

BIBLIOGRAPHY

[1] ATLAS Collaboration, *Inclusive and differential measurement of the charge asymmetry in $t\bar{t}$ events at 13 TeV with the ATLAS detector*, ATLAS-CONF-2019-026, CERN, Geneva, 2019. https://cds.cern.ch/record/2682109

[2] ATLAS Collaboration, *Evidence for the charge asymmetry in $pp \to t\bar{t}$ production at $\sqrt{s} = 13$ TeV with the ATLAS detector*, arXiv.2208.12095 [hep-ex]

[3] B. Eckerova, ATLAS Collaboration, *Charge asymmetry in top-pair production with the ATLAS detector*, arXiv:2212.00161 [hep-ex]

[4] A. Miede, *A Classic Thesis Style – An Homage to The Elements of Typographic Style*, https://bitbucket.org/amiede/classicthesis/ Accessed: 2021-12-07

[5] The LHCb Collaboration, *Measurement of the W boson mass*, JHEP 01 (2022) 036

[6] The Working Group on LEP Energy, The LEP Collaborations ALEPH, DELPHI, L3 and OPAL, *Measurement of the mass of the Z boson and the energy calibration of LEP*, Phys. Lett. B 307 (1993) 187-193

[7] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. 13 (1964) 321-323

[8] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. 12 (1964) 132-133

[9] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. 13 (1964) 508

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

[11] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. 10 (1963) 531–533

[12] M. Kobayashi, T. Maskawa and S. Sakata, *CP Violation in the Renormalizable Theory of Weak Interaction*, Prog. Theor. Phys. 49 (1973) 652-657

[13] S. Bilenky, *Introduction to the Physics of Massive and Mixed Neutrinos*, Lecture Notes in Physics, Springer Berlin, Heidelberg, 2010. ISBN 978-3-642-14042-6

[14] B. Pontecorvo, *Inverse beta processes and nonconservation of lepton charge*, Zh. Eksp. Teor. Fiz. 34 (1957) 247

[15] Z. Maki, M. Nakagawa and S. Sakata, *Remarks on the Unified Model of Elementary Particles*, Prog. Theor. Phys. 28 (1962) 870-880

[16] J. Campbell, J. Huston and F. Krauss, *The Black Book of Quantum Chromodynamics: A Primer for the LHC Era*, Oxford University Press, 2019. ISBN 978-0-19-965274-7

[17] 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. 78 (2018) 129

[18] J. Gao, C. S. Li, and H. X. Zhu, *Top-Quark Decay at Next-to-Next-to-Leading Order in QCD*, Phys. Rev. Lett. 110 (2013) 042001

[19] P.A. Zyla et al. (Particle Data Group), *Review of Particle Physics*, Prog. Theor. Exp. Phys. 2020 083C01 (2020)

[20] CMS Collaboration, *Measurement of the top quark mass using events with a single reconstructed top quark in pp collisions at $\sqrt{s}$ = 13 TeV*, JHEP 12 (2021) 161

[21] L.A. Harland-Lang, A.D. Martin, P. Motylinski and R.S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C 75 (2015) 204

[22] A. Quadt, *Top quark physics at hadron colliders*, Eur. Phys. J. C 48 (2006) 835–1000

[23] M. Czakon, P. Fiedler and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_S^4)$*, Phys. Rev. Lett. 110 (2013) 252004

[24] LHC Top Working Group, *LHCTopWG Summary Plots*, https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots, Accessed: 2022-12-01

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

[26] J. H. Kühn and G. Rodrigo, *Charge asymmetries of top quarks at hadron colliders revisited*, JHEP 01 (2012) 063

[27] S. Moch and P. Uwer, *Theoretical status and prospects for top-quark pair production at hadron colliders*, Phys. Rev. D 78 (2008) 034003

[28] G. Rodrigo, *The ttbar asymmetry in the Standard Model and beyond*, arXiv:1207.0331 [hep-ph]

[29] M. Czakon, D. Heymes, A. Mitov, D. Pagani, I. Tsinikos and M. Zaro, *Top-quark charge asymmetry at the LHC and Tevatron through NNLO QCD and NLO EW*, Phys. Rev. D 98 (2018) 014003

[30] 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

[31] Fermi National Accelerator Laboratory, *Accelarator Tevatron*, https://www.fnal.gov/pub/tevatron/tevatron-accelerator.html, Accessed: 2022-11-13

[32] CDF collaboration, *Forward-Backward Asymmetry in Top Quark Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. Lett. 101 (2008) 202001

[33] CDF Collaboration, *The CDF IIb Detector Technical Design Report*, FERMILAB-TM-2198, FERMILAB-PROPOSAL-0924, CDF-DOC-CDF-PUBLIC-6261, FERMILAB-DESIGN-2003-01, Feb. 2003, https://inspirehep.net/literature/613106

[34] Fermi National Accelerator Laboratory, *D0 experiment*, https://www.fnal.gov/pub/tevatron/experiments/dzero.html, Accessed: 2022-11-13

[35] D0 collaboration, *Forward-backward asymmetry in top quark-antiquark production*, Phys. Rev. D 84 (2011) 112005

[36] CDF collaboration, *Evidence for a mass dependent forward-backward asymmetry in top quark pair production*, Phys. Rev. D 83 (2011) 112003

[37] J. A. Aguilar-Saavedra and M. Perez-Victoria, *Asymmetries in $t\bar{t}$ production: LHC versus Tevatron*, Phys. Rev. D 84 (2011) 115013

[38] CDF and D0 Collaboration, *Combined Forward-Backward Asymmetry Measurements in Top-Antitop Quark Production at the Tevatron*, Phys. Rev. Lett. 120 (2018) 042001

[39] ATLAS Collaboration, *Measurement of the top quark pair production charge asymmetry in proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, JHEP 02 (2014) 107

[40] 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

[41] CMS Collaboration, *Inclusive and Differential Measurements of the $t\bar{t}$ Charge Asymmetry in Proton-Proton Collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. B 717 (2012) 129-150

[42] 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

[43] 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

[44] CMS Collaboration, *Measurements of $t\bar{t}$ differential cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons*, JHEP 02 (2019) 149

[45] CMS and ATLAS Collaborations, *Combination of inclusive and differential $t\bar{t}$ charge asymmetry measurements using ATLAS and CMS data at $\sqrt{s} = 7$ and 8 TeV*, JHEP 04 (2018) 033

[46] 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. D 94 (2016) 032006

[47] 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

[48] CMS Collaboration, *Measurements of $t\bar{t}$ charge asymmetry using dilepton final states in pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. B 760 (2016) 365-386

[49] CMS Collaboration, *Measurement of the $t\bar{t}$ charge asymmetry in events with highly Lorentz-boosted top quarks in pp collisions at $\sqrt{s}$ = 13 TeV*, arXiv:2208.02751 [hep-ex]

[50] J. A. Aguilar-Saavedra, A. Juste and F. Rubbo, *Boosting the $t\bar{t}$ charge asymmetry*, Phys. Lett. B 707 (2012) 92-98

[51] J. A. Aguilar-Saavedra, D. Amidei, A. Juste, and M. Pérez-Victoria *Asymmetries in top quark pair production at hadron colliders*, Rev. Mod. Phys. 87 (2015) 421-455

[52] C. Zhang and S. Willenbrock, *Effective-Field-Theory Approach to Top-Quark Production and Decay*, Phys. Rev. D 83 (2011) 034006

[53] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*. JHEP 10 (2010) 085

[54] M. P. Rosello and M. Vos, *Constraints on four-fermion interactions from the $t\bar{t}$ charge asymmetry at hadron colliders*, Eur. Phys. J. C 76 (2016) 200

[55] L. Evans, and P. Bryant, *LHC Machine*, JINST 3 (2008) S08001

[56] S. Fartoukh et al., *LHC Configuration and Operational Scenario for Run 3*, CERN-ACC-2021-0007, CERN, Geneva, November, 2021 https://cds.cern.ch/record/2790409

[57] O. Aberle, I. B. Alonso, O. Brüning, P. Fessia, M. Lamont, L. Rossi, L. Tavian and M. Zerlauth, *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*, CERN Yellow Reports: Monographs, CERN-2020-010, CERN, Geneva, 2020. ISBN 978-92-9083-587-5

[58] *CERN's accelerator complex*, http://cds.cern.ch/record/1621583/files/CERN's-accelerator-complex2013.jpg?version=1, Accessed: 03.02.2022

[59] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 (2008) S08003

[60] J. Pequenao, *Computer generated image of the whole ATLAS detector*, Mar 2008, https://cds.cern.ch/record/1095924

[61] M. Capeans et al., ATLAS Collaboration *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013, ATLAS-TDR-19, CERN, Geneva, 2010. https://cds.cern.ch/record/1291633

[62] J. Pequenao, *Computer Generated image of the ATLAS calorimeter*, Mar 2008, https://cds.cern.ch/record/1095927

[63] J. Pequenao, *Computer generated image of the ATLAS Muons subsystem*, Mar 2008, https://cds.cern.ch/record/1095929

[64] ATLAS Collaboration, *ATLAS level-1 trigger: Technical Design Report*, CERN-LHCC-98-014 ; ATLAS-TDR-12, Technical design report. ATLAS, CERN, Geneva, 1998. https://cds.cern.ch/record/381429

[65] P. Jenni, M. Nessi, M. Nordberg and K. Smith, ATLAS Collaboration, *ATLAS high-level trigger, data-acquisition and controls: Technical Design Report*, Technical design report. ATLAS, CERN, Geneva, 2003. https://cds.cern.ch/record/616089

[66] A. Ruiz-Martínez, ATLAS Collaboration, *The Run-2 ATLAS Trigger System*, ATL-DAQ-PROC-2016-003, CERN, Geneva, Feb, 2016. https://cds.cern.ch/record/2133909

[67] M. Schochet and L. Tompkins, V. Cavaliere, P. Giannetti, A. Annovi and G. Volpi, *Fast TracKer (FTK) Technical Design Report*, CERN-LHCC-2013-007, ATLAS-TDR-021, CERN, Geneva, June, 2013. https://cds.cern.ch/record/1552953

[68] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Methods. Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip. 262 (1987) 444-450

[69] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, Eur. Phys. J. C 77 (2017) 673

[70] ATLAS Collaboration, *ATLAS Track Reconstruction – General Overview*, https://atlassoftwaredocs.web.cern.ch/trackingTutorial/idoverview/, Accessed: 2022-11-24

[71] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at* $\sqrt{s}$ *= 13 TeV*, Eur. Phys. J. C 79 (2019) 639

[72] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C 77 (2017) 490

[73] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST 14 (2019) P12006

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

[75] M. Cacciari, G. Salam and G. Soyez, *The Anti-*k_t *jet clustering algorithm*, JHEP 04 (2008) 063

[76] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, April, 2008. https://cds.cern.ch/record/1099735

[77] M. Cacciari and G. P. Salam and G. Soyez, *The catchment area of jets*, JHEP 04 (2008) 005

[78] 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

[79] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Eur. Phys. J. C 81 (2021) 689

[80] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*, Eur. Phys. J. C 76 (2016) 581

[81] ATLAS Collaboration, *Performance of b-jet identification in the ATLAS experiment*, JINST 11 (2016) P04008

[82] ATLAS Collaboration, *Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV*, JHEP 08 (2018) 089

[83] D. Krohn, J. Thaler and L.-T. Wang, *Jet Trimming*, JHEP 02 (2010) 084

[84] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D 48 (1993) 3160–3166

[85] 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

[86] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C 78 (2018) 903

[87] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C 70 (2010) 823–874

[88] S. Agostinelli et al., *Geant4 - a simulation toolkit*, Nucl. Instrum. Methods. Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip. 506 (2003) 250–303

[89] E. Richter-Was, D. Froidevaux and L. Poggioli, *ATLFAST 2.0 a fast simulation package for ATLAS*, ATL-PHYS-98-131, CERN, Geneva, November, 1998. https://cds.cern.ch/record/683751

[90] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Methods. Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip. 462 (2001) 152–155

[91] E. Bothmann et al., *Event Generation with Sherpa 2.2*, SciPost Phys. 7 (2019) 034

[92] T. Sjöstrand et al., *An Introduction to PYTHIA 8.2*, Com. Phys. Comm. 191 (2015) 159-177

[93] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. c 63 (2009) 189–285

[94] ATLAS Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model*, ATL-PHYS-PUB-2016-017, CERN, Geneva, 2016. https://cds.cern.ch/record/2206965

[95] S. Frixione, G. Ridolfi and P. Nason, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP 09 (2007) 126

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

[97] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP 11 (2007) 070

[98] S. Alioli and P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP 06 (2010) 043

[99] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, Eur. Phys. J. C 71 (2011) 1547

[100] R. Frederix, E. Re and P. Torrielli, *Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*, JHEP 09 (2012) 130

[101] S. Alioli, P. Nason, C. Oleari and E. Re, *NLO single-top production matched with shower in POWHEG:s- and t-channel contributions*, JHEP 09 (2009) 111

[102] NNPDF Collaboration, R. D. Ball et al., *Parton distributions for the LHC run II*, JHEP 04 (2015) 040

[103] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, ATL-PHYS-PUB-2016-020, CERN, Geneva, 2016. https://cds.cern.ch/record/2216168

[104] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, CERN, Geneva, 2014. https://cds.cern.ch/record/1966419

[105] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. C 58 (2008) 639-707

[106] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C 76 (2016) 196

[107] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C 75 (2015) 204

[108] J. Bellm et al., *Herwig 7.1 Release Note*, arXiv:1705.06919 [hep-ph]

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

[110] J. Pumplin, D. R. Stump, J. Huston, H.-L. Lai, P. Nadolsky and W.-K. Tung, *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP 07 (2002) 012

[111] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber, *Single-top hadroproduction in association with a W boson*, JHEP 07 (2008) 029

[112] J. M. Campbell, R. Frederix, F. Maltoni, and F. Tramontano, *Next-to-Leading-Order Predictions for t-Channel Single-Top Production at Hadron Colliders*, Phys. Rev. Lett. 102 (2009) 182003

[113] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP 12 (2008) 039

[114] F. Cascioli and P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. 108 (2012) 111601

[115] A. Denner, S. Dittmaier and L. Hofer, *Collier: A fortran-based complex one-loop library in extended regularizations*, Com. Phys. Comm. 212 (2017) 220–238

[116] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP 03 (2008) 038

[117] J.-C. Winter, F. Krauss and G. Soff, *A modified cluster-hadronisation model*, Eur. Phys. J. C 36 (2004) 381–395

[118] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, JHEP 09 (2012) 049

[119] S. Catani, F. Krauss, B. R. Webber and R. Kuhn *QCD Matrix Elements + Parton Showers*, JHEP 11 (2001) 063

[120] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, JHEP 05 (2009) 053

[121] S. Höche, F. Krauss, M. Schönherr and F. Siegert, *QCD matrix elements + parton showers. The NLO case*, JHEP 04 (2013) 027

[122] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP 07 (2014) 079

[123] ATLAS Collaboration, *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at* $\sqrt{s}$=8 *TeV with the ATLAS detector*, ATLAS-CONF-2014-058, CERN, Geneva, 2014. https://cds.cern.ch/record/1951336

[124] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, Eur. Phys. J. C 80 (2020) 47

[125] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, JINST 15 (2020) P09015

[126] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, arXiv:physics/0703039 [physics.data-an]

[127] J. Erdmann, S. Guindon, K. Kröninger, B. Lemmer, O. Nackenhorst, A. Quadt and P. Stolte, *A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework*, Nucl. Instrum. Methods. Phys. Res. A: Accel. Spectrom. Detect. Assoc. Equip. 748 (2014) 18-25

[128] ATLAS Collaboration, *Performance of top-quark and W-boson tagging with ATLAS in Run 2 of the LHC*, Eur. Phys. J. C 79 (2019) 375

[129] J. Thaler and K. V. Tilburg, *Identifying boosted objects with N-subjettiness*, JHEP 03 (2011) 015

[130] D0 Collaboration, *Precise measurement of the top quark mass in dilepton decays using optimized neutrino weighting*, Phys. Lett. B 752 (2016) 18–26

[131] G. Choudalakis, *Fully Bayesian Unfolding*, arXiv:1201.4612 [hep-ex]

[132] M. D. Hoffman and A. Gelman, *The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo*, JMLR 15 (2014) 1593-1623

[133] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data*, JINST 14 (2019) P03017

[134] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} =$ 13 TeV*, ATL-PHYS-PUB-2015-015, CERN, Geneva, July, 2015. https://cds.cern.ch/record/2037613

[135] ATLAS Collaboration, *Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector*, ATLAS-CONF-2018-001, CERN, Geneva, March, 2018. https://cds.cern.ch/record/2306649

[136] ATLAS Collaboration, *Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at $\sqrt{s} =$ 13 TeV*, ATLAS-CONF-2018-006, CERN, Geneva, April, 2018. https://cds.cern.ch/record/2314418

[137] ATLAS Collaboration, *In-situ measurements of the ATLAS large-radius jet response in 13 TeV pp collisions*, ATLAS-CONF-2017-063, CERN, Geneva, July, 2017. https://cds.cern.ch/record/2275655

[138] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, ATLAS-CONF-2019-021, CERN, Geneva, June, 2019. https://cds.cern.ch/record/2677054

[139] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, JINST 13 (2018) P07017

[140] J. Rojo, *PDF4LHC recommendations for Run II*, PoS DIS2016 (2016) 018 arXiv:1606.08243 [hep-ph]

[141] ATLAS Collaboration, *Measurements of top-quark pair differential and double-differential cross-sections in the ℓ+jets channel with pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, Eur. Phys. J. C 79 (2019) 1028

[142] ATLAS Collaboration, *ATLAS simulation of boson plus jets processes in Run 2*, ATL-PHYS-PUB-2017-006, CERN, Geneva, May, 2017. https://cds.cern.ch/record/2261937

[143] ATLAS Collaboration, *Measurement of $W\pm$ and Z-boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B 759 (2016) 601-621

[144] F. Krauss, *Matrix elements and parton showers in hadronic interactions* JHEP 08 (2002) 015

[145] P. Kant and O.M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz, P. Rieck and P. Uwer, *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, Comp. Phys. Comm. 192 (2015) 74-89

[146] M. Aliev and H. Lacker, U. Langenfeld, S. Moch, P. Uwer and M. Wiedermann, *HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR*, Comp. Phys. Comm. 182 (2011) 1034-1046

[147] ATLAS Collaboration, *Multi-Boson Simulation for 13 TeV ATLAS Analyses*, ATL-PHYS-PUB-2017-005, CERN, Geneva, May, 2017.https://cds.cern.ch/record/2261933

[148] S. Frixione, V. Hirschi, D. Pagani, H.-S. Shao and M. Zaro, *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, JHEP 06 (2015) 184

[149] R. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. 77 (1993) 219-228

[150] G. Bohm and G. Zech, *Introduction to Statistics and Data Analysis for Physicists; 2nd rev. ed*, DESY, Hamburg, 2010. ISBN: 9783935702416

[151] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, Phys. Lett. B 810 (2020) 135797

[152] M. Czakon, D. Heymes, A. Mitov, D. Pagani, I. Tsinikos and M. Zaro, *Top-pair production at the LHC through NNLO QCD and NLO EW*, JHEP10 (2017) 186

[153] W. Bernreuther and Z.-G. Si, *Top quark and leptonic charge asymmetries for the Tevatron and LHC*, Phys. Rev. D86 (2012) 034026

[154] C. Degrande, G. Durieux, F. Maltoni, K. Mimasu, E. Vryonidou and C. Zhang, *Automated one-loop computations in the standard model effective field theory* Phys. Rev. D. 103 (2021) 096024

[155] J. J. Ethier, G. Magni, F. Maltoni, L. Mantani, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, C. Zhang, *Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC*, arXiv:2105.00006

[156] I. Brivio, S. Bruggisser, F. Maltoni, R. Moutafis, T. Plehn, E. Vryonidou, S. Westhoff and C. Zhang, *O new physics, where art thou? A global search in the top sector*, JHEP 02 (2020) 131

[157] ATLAS Collaboration, *Measurement of the energy asymmetry in $t\bar{t}j$ production at 13 TeV with the ATLAS experiment and interpretation in the SMEFT framework*, Eur. Phys. J. C. 82 (2022) 374