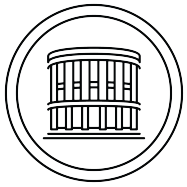
Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics

The associated production of top-quark pair and Z boson in proton-proton collisions at the ATLAS experiment

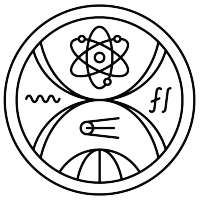
PhD thesis

2022

Mgr. Dominik Babál



Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics



The associated production of top-quark pair and Z boson in proton-proton collisions at the ATLAS experiment

PhD thesis

Study programme:	Nuclear and Subnuclear Physics
Supervisor:	prof. RNDr. Stanislav Tokár, DrSc.
Consultant:	Mgr. Pavol Bartoš, PhD.

Bratislava, 2022

Mgr. Dominik Babál

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Mgr. Dominik Babál
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: The associated production of top-quark pair and Z boson in proton-proton collisions at the ATLAS experiment
Asociovaná produkcia top-kvarkového páru a Z bozónu v protón-protónových zrážkach experimentu ATLAS

Anotácia: Táto práca sa venuje asociovanej produkcii top-quarkových párov a Z-bozónu, pričom top kvark je rekonštruovaný v celo-hadrónovom kanále a Z bosón je rekonštruovaný v dileptónovom rozpadovom kanále. Väzbová konštanta medzi bozónom Z a top kvarkom je citlivá pozorovateľná veličina pre hľadanie novej fyziky. Analýza sa uskutoční na báze použitia protón-protónových dát, ktoré boli získané pri ťažiskovej energie 13 TeV. Tento proces je charakterizovaný vysokým pozadím a nízkym účinným prierezom pre signálny proces. Identifikácia eventov signálu je preto veľmi náročná a je nevyhnutné použiť multivarietné techniky - ako je neurónová sieť alebo BDT (Boosted Decision Tree). Optimalizácia takzvaných "diskriminačných" premenných, ktoré majú rozdielne rozdelenie signálu a pozadia, je dôležitou úlohou, ktorá je potrebná na odlíšenie signálu od pozadia. Cieľom práce je získať procesný prierez s vysokou presnosťou pre účely testovania Štandardného modelu resp. Hľadania novej fyziky.

Cieľ: Základným cieľom práce je získať účinný prierez procesu tt-bar Z s vysokou presnosťou pre účely testovania Štandardného modelu resp. hľadania novej fyziky.

Literatúra: F. Halzen, A. D. Martin: Quarks and Leptons: An Introductory Course in Modern Particle Physics, 1984
O. Behnke, K. Kröninger, G. Schott and T. Schörner-Sadenius: Data Analysis in High Energy Physics, 2013
T. Plehn, Lectures on LHC physics, Springer International Publishing, Switzerland, 2012, 2015.
J.F. Donoghue, E. Golovich, B.R. Holstein, Dynamics of the Standard Model, Cambridge University Press, 1992.

Poznámka: Študent sa bude zaoberať súčasnou špičkovou časticovou fyzikou realizovanou na experimente ATLAS, CERN.

Kľúčové slová: Top kvark, Z bozón

Školiteľ: prof. RNDr. Stanislav Tokár, DrSc.

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

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

Spôsob sprístupnenia elektronickej verzie práce:
bez obmedzenia

Dátum zadania: 26.01.2018

Dátum schválenia: 12.02.2018 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. Dominik Babál
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: The associated production of top-quark pair and Z boson in proton-proton collisions at the ATLAS experiment

Annotation: This work deals with the associated production of top-quark pair and Z-boson, where top quark is reconstructed in all hadronic mode and Z boson is taken in dilepton decay channel. The coupling between Z boson and top quark is a sensitive observable for searches of a new physics. The analysis will be carried out using proton-proton data taken at center-of-mass energy of 13 TeV. This process is characterized by a high background and a low cross-section for the signal process. The identification of signal events is therefore very demanding and multivariate techniques such as the neural network or the boosted decision tree (BDT) must be used. Optimization of so-called “discriminatory” variables, which have different distributions for signal and background, is an important task needed to distinguish the signal from the background. The aim of the work is to obtain the signal process cross section with a high precision for purposes of the Standard model tests and/or searches of a new physics, respectively.

Aim: The aim of the work is to obtain the cross section of the process tt-bar Z with a high precision for purposes of the Standard model tests and/or searches of a new physics, respectively.

Literature:
F. Halzen, A. D. Martin: Quarks and Leptons: An Introductory Course in Modern Particle Physics, 1984
O. Behnke, K. Kröninger, G. Schott and T. Schörner-Sadenius: Data Analysis in High Energy Physics, 2013
T. Plehn, Lectures on LHC physics, Springer International Publishing, Switzerland, 2012, 2015.
J.F. Donoghue, E. Golovich, B.R. Holstein, Dynamics of the Standard Model, Cambridge University Press, 1992.

Comment: The student will have to do with the present front-end particle physics carried out at experiment ATLAS, CERN

Keywords: Top quark, Z boson

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

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

Assigned: 26.01.2018

Approved: 12.02.2018 prof. RNDr. Jozef Masarik, DrSc.
Guarantor of Study Programme

.....
Student

.....
Tutor

Acknowledgments

First of all, I would like to thank my supervisor Stanislav Tokár for his excellent guidance in a field of high energy physics, for giving me opportunity not only to learn new things in physics, statistics and computing, but also to travel to different countries and explore a tiny bit of the world.

The work presented in this thesis would not be possible without great $t\bar{t}Z$ team and especially my closest co-worker Mišo, who helped me with everything from statistics to computing. He greatly broadened my knowledge in neural networks and actually in every step of ATLAS physics analysis. For this and also for being great hiking partner I would like to thank him a lot. I highly appreciate the immense work done by the whole $t\bar{t}Z$ team, especially our patient and hard-working analysis contacts, Fabio and Tom. I thank them for answering all my questions and guiding me through differential analysis. Great gratitude goes also to the main analyzers of the first round of $t\bar{t}Z$ analysis, Tom and Baptiste, for helping me with unfolding and pushing hard when it was needed. Without the effort of these bright physicists I would not be able to write this thesis.

Big thanks goes also to my qualification task supervisor Simone, who helped me to become qualified ATLAS author. The various things I learned about Monte Carlo generators and computing helped me also in my $t\bar{t}Z$ studies.

I thank also to our Fermilab team, George, Igor, Costas and Giorgio, for physical interludes in my CERN work, and for showing me that patience and endurance is a key feature for successful measurement. I believe that our top mass effort will one day conclude CDF legacy.

Of course I cannot forget on the rest of out Bratislava ATLAS group, for creating perfect working environment and being source of wide-ranging information. Therefore I would like to say "thank you" to: Báša for being my longest-standing classmate, for helping me with anything I asked for, and for being good friend; Robo and Paľo for sharing the office, for offering helping hand with any bureaucratic issue I had to solve, and for providing guidance in the restricted zones; Tomáš and Oliver for training my patience in listening, and for broadening my knowledge in many areas, but especially in physics, computing and ancient Rome; Lucia for our running and hiking adventures, and for showing me that there is at least one more person in the world as taciturn as

me; Juro and Sofia for providing BEC fun and bringing a bit of Ukraine to our team; Dano for being the first person from folklore ensemble I had a luck to meet; Adriana for ordering more food than she was able to eat; Maťo for being experienced hiking partner; Mišo R. for our numerous hikes and being my reliable python reference; Tibor for being our server guru and Norika for being best secretary with ability to solve everything.

Last but not least, I greatly thank to my whole family and especially the most important people in my life, my parents, for their never-ending love and support.

Abstrakt

Top kvark predstavuje dôležitú elementárnu časticu Štandardného Modelu kvôli jeho výnimočným vlastnostiam. Vďaka jeho veľkej hmotnosti a zároveň krátkej strednej dobe života sa rozpadá ešte pred vytvorením viazaného stavu, čo z neho robí ideálnu časticu na skúmanie vlastností pseudo-voľného kvarku, a ponúka možnosť na precízne testovanie vnútornej konzistentnosti Štandardného Modelu. Táto práca je zameraná na štúdium asociovanej produkcie top-antitop-kvarkového páru a Z bozónu ($t\bar{t}Z$). Hoci je tento proces považovaný za ojedinelý z hľadiska Štandardného Modelu, veľká ťažisková energia a luminozita dosiahnutá na urýchľovači LHC umožňuje jeho presnú analýzu.

Prvá časť tejto práce je venovaná vôbec prvému meraniu diferenciálneho účinného prierezu $t\bar{t}Z$ procesu uskutčnenému na dátach z detektoru ATLAS. Na meranie je použitý plný dataset Runu 2 urýchľovača LHC, ktorý zodpovedá integrovanej luminozite 139 fb^{-1}. Dáta zodpovedajú pp zrážkam pri ťažiskovej energii 13 TeV. Na meranie diferenciálneho účineho prierezu sú použité rozpadové kanále $t\bar{t}Z$ procesu, ktoré obsahujú vo finálnom stave buď tri alebo štyri nabité leptóny. Meranie je uskutočnené použitím tzv. metódy IBU založenej na Bayesovej formule, pričom účinné preirezy sú zmerané ako funkcie deviatich premenných. Výsledné rozdelenia diferenciálneho účinného prierezu sú v zhode s predpoveďami Štandardného Modelu.

Druhá časť tejto dizertačnej práce je zameraná na meranie inkluzívneho účinného prierezu s použitím rovnakých dát ako pre diferenciálne meranie, avšak v rozpadovom kanále v ktorom sú vo finálnom stave len dva leptóny z rozpadu Z bozónu. Inkluzívny účinný prierez je zmeraný fitovaním profilovej vierohodnosti, pričom výsledok je prezentovaný vo forme sily signálu, ktorá je definovaná ako pomer medzi zmeranou hodnotou účinného preirezu a zodpovedajúcou predpoveďou Štandardného Modelu. Nasledovná sila signálu bola zatiaľ zmeraná len s použitím nasimulovaných Monte Carlo dát a predstavuje očakávanú hodnotu zodpovedajúcu teoretickej predpovedi:

$$\mu_{t\bar{t}Z}^{2\ell} = 1.000^{+0.124}_{-0.116} = 1.000^{+0.100}_{-0.089}(\text{syst.}) \pm 0.074(\text{stat.}). \quad (1)$$

Kľúčové slová: top kvark, Z bozón, $t\bar{t}Z$, inkluzívny účinný prierez, diferenciálny účinný prierez

Abstract

The top quark represents an important elementary particle of the Standard Model because of its extraordinary properties. Due to its large mass, and consequently its short mean lifetime, it decays before the formation of bound states, what makes it an ideal particle for studying properties of the pseudo-bare quark and offers the possibility for precise tests of internal consistency of the Standard Model. This thesis is focused on the study of the associated production of top-antitop-quark pair and Z boson ($t\bar{t}Z$). Although this process is considered very rare in the Standard Model, the large center-of-mass energy and luminosity achieved by the Large Hadron Collider enable its precise analysis.

The first part of this thesis is devoted to the first ever differential cross section measurement of $t\bar{t}Z$ process performed on data detected by the ATLAS detector. The full dataset of Run 2 of the Large Hadron Collider operation, corresponding to integrated luminosity of 139 fb^{-1} obtained in pp collisions at center-of-mass energy of 13 TeV, is employed for measurements in $t\bar{t}Z$ decay channels featuring three or four charged leptons in the final state. The resulting distributions of the differential cross section, measured as a function of nine variables using the method of iterative Bayesian unfolding, are found to be in consistency with the Standard Model predictions.

The second part of this dissertation is focused on the inclusive cross section measurement using the same data as for the differential part, but performed in the final state with only two charged leptons from the Z boson decay. The inclusive cross section is measured by the profile likelihood fitting technique that is used for the extraction of signal strength, representing the ratio between the measured cross section and its corresponding Standard Model prediction. Only the expected signal strength using Monte Carlo simulations is measured, yielding the following result:

$$\mu_{t\bar{t}Z}^{2\ell} = 1.000^{+0.124}_{-0.116} = 1.000^{+0.100}_{-0.089}(\text{syst.}) \pm 0.074(\text{stat.}). \quad (2)$$

Keywords: top quark, Z boson, $t\bar{t}Z$, inclusive cross section, differential cross section

Contents

Introduction

Since its discovery[1] in 1995 [1, 2], the top quark and its physics remains one of the most investigated topics in experimental particle physics. Because of its large mass ($m_t = 173.34 \pm 0.27$ (stat.) $\pm$ 0.71(syst.) GeV [3]) and decay width ($\Gamma_t = 1.76 \pm 0.33$ (stat.) $^{+0.79}_{-0.68}$ (syst.) GeV [4]), and consequently short mean lifetime ($\tau_t \approx 5 \times 10^{-25}$ s)[2], it represents ideal particle for studying properties of pseudo-bare quark. It passes its spin information to its decay products and thus enables to measure observables that depend on top quark spin, providing various possibilities to test phenomena predicted by the Standard Model (SM). Moreover, its large Yukawa coupling constant ($g_t = 0.9902$ [5]) makes it ideal elementary particle for studies involving Higgs boson.

Associated production of the top-quark pair and Z boson ($t\bar{t}Z$) represents rare process predicted by the SM. Its cross section is approximately 1000 times smaller than the top-quark pair production [6]. For this reason, the low amount of data postponed precise measurement of the total cross section at the centre-of-mass energy of 13 TeV until 2018, when the first measurement [7], conducted by the CMS collaboration, was published. This measurement was carried out using limited dataset (the data correspond to the total integrated luminosity of 35.9 fb^{-1}) and yielded total $t\bar{t}Z$ cross section of $\sigma_{t\bar{t}Z} = 0.99^{+0.09}_{-0.08}$ (stat.) $^{+0.12}_{-0.10}$ (syst.) pb. Analogous study [8] was performed also by the ATLAS experiment using 36.1 fb^{-1} of data, resulting in $\sigma_{t\bar{t}Z} = 0.95 \pm 0.08$ (stat.) ± 0.10 (syst.) pb. Both results agree with the SM prediction $\sigma_{t\bar{t}Z} = 0.863^{+8.5\%}_{-9.9\%}\,(\text{scale}) \pm 3.2\%\,(\text{PDF} + \alpha_\text{S})$ pb [6]. First $t\bar{t}Z$ differential cross section measurement[3] was conducted by the CMS experiment in 2019 [9] using more than half of the full dataset (corresponds to about 77.5 fb^{-1}) obtained in the Run 2 of the Large Hadron Collider (LHC) operation.

Differential cross section of $t\bar{t}Z$ has never been measured before using data from the ATLAS detector. This thesis presents the analyses which use the full Run 2 dataset of the LHC operation (corresponds to 139 fb^{-1}), and aims for both inclusive and differ-

[1] Top quark was discovered at the CDF and D0 experiments at the Tevatron accelerator.

[2] Due to the short mean lifetime, the top quark decays before hadronization (formation of bound state with other quarks within framework of quantum chromodynamics).

[3] Total cross section measurement was also included in this analysis, yielding $\sigma_{t\bar{t}Z} = 1.00^{+0.06}_{-0.05}$ (stat.) $^{+0.07}_{-0.06}$ (syst.) pb.

ential measurement in the channels where either 2 charged leptons (*dilepton channel*, $t\bar{t} \to$ jets and $Z \to \ell\ell$), 3 charged leptons (*trilepton channel*, $t\bar{t} \to \ell +$ jets and $Z \to \ell\ell$) or 4 charged leptons (*tetralepton channel*, $t\bar{t} \to \ell\ell +$ jets and $Z \to \ell\ell$) are detected in the final state.

In the first analysis described in this thesis, the differential cross sections as a function of nine variables, are measured in either trilepton or tetralepton decay channels. This analysis is motivated by the assumption that the differential cross section (especially for some $t\bar{t}Z$ observables) is sensitive to the Beyond Standard Model (BSM) physics. Moreover, differential distributions have potential to improve modeling of the Monte Carlo generators. Another reason for studying $t\bar{t}Z$ process is the fact, that it represents important background for associated production of top-quark pair with Higgs boson ($t\bar{t}H$), which is useful for testing the SM prediction of the top-quark Yukawa coupling constant. The differential cross section measurement is a part of ATLAS analysis that was already published in 2021 [10], and includes also measurement of the inclusive $t\bar{t}Z$ cross section, again employing trilepton and tetralepton channels.

Second analysis presented in this dissertation is focused on the ongoing measurement of the inclusive cross section exploiting the dilepton $t\bar{t}Z$ channel. Although the similar analysis in dilepton channel was already performed by the ATLAS collaboration using 36.1 fb^{-1} of LHC data [8], the current effort aims to increase the precision of the previous measurement by using full Run 2 dataset, as well as by introducing sophisticated statistical methods in order to increase sensitivity of the measurement. The physical motivation of this analysis is based on the potential to either verify or exclude current SM prediction for the inclusive $t\bar{t}Z$ cross section.

The thesis is organized as follows: Chapter 1 briefly introduces the SM of elementary particles, its origin and current physics concept. Chapter 2 is devoted to detailed description of the physics related to top quark, Z boson and their associated production. The LHC accelerator complex and the ATLAS detector are described in Chapter 3. Next Chapter 4 focuses on the various reconstruction techniques used for constructing physics objects relevant for the $t\bar{t}Z$ analyses. Observed data and simulated Monte Carlo samples, employed in the discussed analyses, are specified in Chapter 5. Numerous sources of experimental and theoretical systematic uncertainties, considered for the inclusive and differential measurements, are summarized in Chapter 6. Chapter 7 is devoted to the statistical methods employed in either inclusive or differential cross section measurements. The selection criteria defining signal regions, in which the measurements are performed, are summarized in Chapter 8. Dedicated algorithms used for the reconstruction of $t\bar{t}$ pair relevant for constructing either differential variables or discriminating variables used in the dilepton neural network, are described in Chapter 9. The analysis strategy and result of the differential and inclusive measurements are presented in Chapter 10 and Chapter 11, respectively.

Chapter 1

Standard Model of elementary particles

1.1 Introduction

The goal of particle physics from its beginning has been to identify what appear to be elementary (structureless) units of matter surrounding us, and to understand the forces governing their interactions. The expectation is that these elementary particles should interact in the simplest possible way; or that there is substantial connection between the matter particles and forces. This matter/force nature can be illustrated by the discovery of the electron by J.J. Thomson [11] and Maxwell's theory of the electromagnetic field [12], which can be considered a foundation of modern particle physics. The electron was not only recognized as important constituent of matter, but its motion constituted an electromagnetic current, and was thus identified also as a source of the electromagnetic field.

The story of particle physics continued in next one hundred years with discovery and study of two new forces - the strong and weak forces - that triggered the search for new constituents of matter, which would serve also as sources of this new force fields (similarly as electron for electromagnetic force). The studies culminated in the last quarter of twentieth century by identification of new particles which indeed behaved in a similar way as already known electron. More importantly, these new matter units and their observed interactions offered convincing verification of theories of the strong and weak force fields, which elegantly incorporated and generalized the original relationship between electron and electromagnetic field. These theories nowadays constitute the most successful (in terms of experimental verification) concept of particle physics called the Standard Model (SM). Elementary particles of the SM can be divided in three sectors: fundamental fermions, gauge bosons and Higgs boson sector. Particular matter units and basic interactions of the SM (excluding interaction via gravitational force)

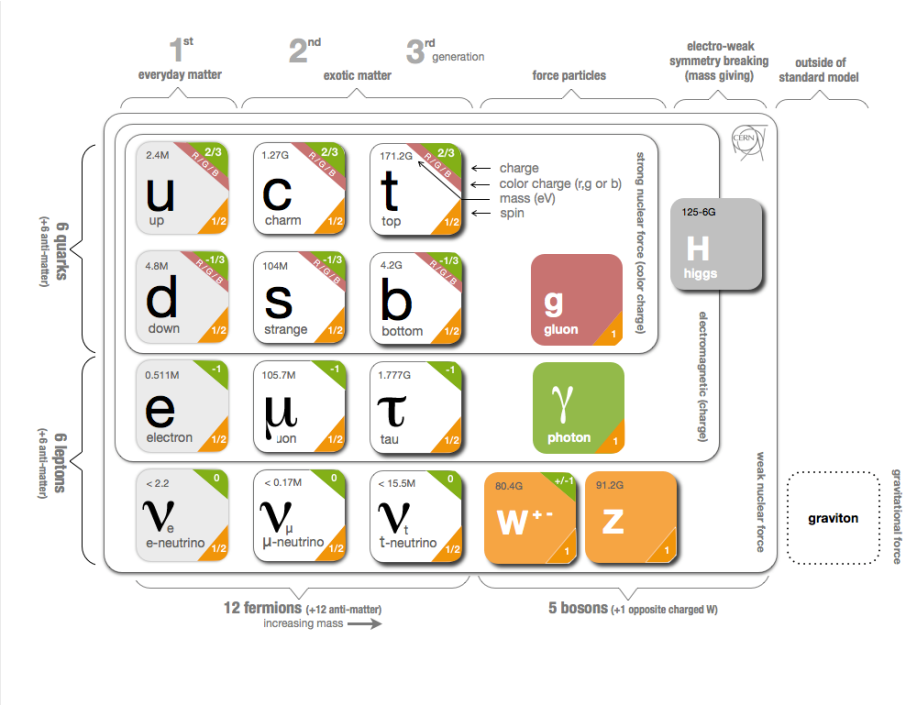


Fig. 1.1: Summary of the elementary particles of the SM [14].

are briefly described in the following sections, which are inspired by Ref. [13].

1.2 Fermions

The matter units of the SM are called fermions, particles with spin-$\frac{1}{2}$ [1]. They are further categorized as either leptons or quarks. While the leptons interact both electromagnetically and weakly (if they are charged, if not, only weakly), the quarks can interact via all three interactions - strong, weak and electromagnetic. Fermions of the SM, together with elementary bosons, mediators of the basic interactions, are summarized in Figure 1.1.

1.2.1 Leptons

Development of the particle physics continued almost forty years after the discovery of the electron, when in 1932 Anderson discovered the positron [15], electron's antiparticle. The properties of the new particle were the same as of electron, but its charge was opposite.

Shorty afterwards, in 1936, experiment led by Anderson and Neddemeyer [16] announced the discovery of the muon (μ^-), first member of the new generation of leptons. The existence of the new particle was confirmed one year later by Street and Stevenson

[1] Natural units with $\hbar = c = 1$ are used throughout whole thesis.

[17]. At first, muon was considered as a particle corresponding to the field quantum of Yukawa's theory of strong nuclear force [18] postulated in 1935. Although muon's mass of 105.7 MeV nicely corresponds with the Yukawa's prediction, experiments led by Conversi in 1947 [19] showed that muon could not be Yukawa's quantum since it did not interact via strong interaction. The μ^- behaved in exactly the same way as electron, interacting only weakly and electromagnetically, with similar interaction strengths as electron, and was therefore identified as lepton.

The third generation of leptons was established after next almost forty years, with the discovery of the tau (τ^-) lepton by Perl [20]. This heaviest lepton (with mass around 1.78 GeV), as well as its antiparticle τ^+ again interacted in the same way as e^- or e^+.

One might naturally think that these new particles are just excitation states of the one particle. However, this hypothesis contradicted with the fact that no other such excitation states were found. Moreover, all these leptons have the same spin ($\frac{1}{2}$) which does not corresponds to typical excitation spectrum. And third argument against this concept was that no γ-transitions were observed between these three hypothetical states [2].

In charged weak interactions [3] each of the three charged leptons are accompanied by its neutral partner, a neutrino. These new types of particles were originally introduced by Pauli in 1930s [22] as a concept to explain violation of conservation laws of four-momentum and angular momentum in β-decay. The electron antineutrino ($\bar{\nu}_e$) was confirmed experimentally by Reines and Cowan in 1956 [23] through studying positrons emitted in nuclear reactor via inverse β-process. The discovery of the muon antineutrino followed soon after in 1962 [24] when Danby studied π^- decay. This important experiment provided evidence that the two neutrinos (ν_μ and ν_e) are two different particles. It was found that the neutrinos form π^- decay always produced muons on interacting with matter and never electrons, suggesting that the created neutrino does not correspond to electron neutrino. Finally, the tau neutrino (ν_τ) was observed in DONUT experiment at Fermilab in year 2000 [25].

One of the most important properties of the neutrinos, their mass, still remains unresolved. As originally suggested by Pauli, the neutrino emitted in β-decay should have very little mass, because the maximum energy carried off by the electron was almost the same as the difference in rest energies of neutron and proton. This was also the reason why the original SM considered neutrinos to be strictly massless. However, there is now clear evidence that neutrinos do have non-zero masses. The phenomenon

[2] For example, current upper limit on $\mu^- \to e^- + \gamma$ branching fraction is 4.2×10^{-13} at 90% confidence level [21]. Similarly, there are less stringent limits for $\tau^- \to e^- + \gamma$ and $\tau^- \to \mu^- + \gamma$.

[3] In case of interactions mediated by neutral Z boson, the charged leptons are not associated with their corresponding neutrinos.

which provides such evidence is called *neutrino oscillations* and was first postulated by Pontecorvo in 1957 [26], and later discovered in 1998 by Super-Kamiokande detector in Japan [27]. It is important to note that neutrino oscillations do not directly measure absolute neutrino masses, but are sensitive only to the differences of squared masses of the neutrinos.

1.2.2 Quarks

Constituents of so-called hadrons, composite particles bound by the strong force, are quarks. Hadrons can be classified as either baryons (if their spin is $\frac{1}{2}$, $\frac{3}{2}$, $\frac{5}{2}$, ...) or mesons (with spin 0, 1, 2, ...). Nucleons (proton p or neutron n) are examples of baryons. SM contains in total 6 quarks organized into three generations in analogy to the lepton categorization. Up-type quarks carry charge of $+\frac{2}{3}$, while down-type quarks carry $-\frac{1}{3}$ of elementary charge [4]. Each quark has its corresponding antiquark sharing the same properties but opposite charge.

The composite nature of hadrons was first revealed in 1963 experiment done by Hofstadter and co-workers [28], which showed that proton was not pointlike but its distribution was approximately exponential with the distance from its center. Further spectroscopy experiments observed sequences of excited states which resembled those present in atomic and nuclear physics. Soon after these findings, Gell-Mann [29] and Zweig [30] proposed that these "spectra" are caused by baryons which contain three spin-$\frac{1}{2}$ particles (called quarks by Gell-Mann) and mesons with quark-antiquark constituents. Important consequence is that quarks carry fractional electric charge, since proton has unit charge. No experiment, however, was able to observe quarks as stable isolated particles, what caused the scepticism that quarks are just mathematical constructs for explaining complicated data rather than objects of physical reality. Indeed even today when hadrons are smashed together into each other, the only products of this collisions are again hadrons [5], and not fractionally charged quarks. This behaviour was later explained by so called confinement of quarks, which is caused by the nature of the strong force acting between quarks [6].

In the time when Gell-Mann and Zweig published their theoretical works, only three quarks (strange s quark in addition to u and d as the constituents of proton and neutron) were necessary to describe known particles such as Λ^0 hyperon (uds) and K^0 ($d\bar{s}$) meson. In 1964 Glashow and Bjorken [31] suggested the existence of the fourth,

[4] Elementary charge corresponds to the electric charge of proton and is typically used as a charge unit in the high energy physics.

[5] In fact also other stable particles (i.e. leptons or photons) can emerge from the high energy collision.

[6] There is one exception, t quark, which can be observed directly through its decay products since it decays before hadronization.

c (charm) quark in analogy with lepton generations. Improved theory by Glashow, Iliopoulos and Maiani [32], which introduced the framework of gauge theories, provided even stronger argument for the existence of the fourth quark. The most precise one-loop calculations [33] in electroweak theory predicted the mass of new quark to lie in vicinity of 1.5 GeV. Experimental evidence followed soon after, in year 1974, with the discovery of the J/ψ meson ($c\bar{c}$) [34, 35] with the mass of around 3 GeV, confirming the theoretical calculations.

Even before the discovery of the c quark, there was already study done by Kobayashi and Maskawa in 1973 [36], in which they predicted existence of the new generation of quark doublet. The necessity of this new generation was based on the problem related to incorporation of the CP symmetry into the quark sector of the electroweak theory. Violation of the CP symmetry in weak interactions was very problematic without third generation of quarks. The search for the new quark doublet has been escalated with already mentioned discovery of the τ lepton in 1975, which established third generation of leptons. Finally, the b (bottom) quark was discovered in 1977 with the observation of the Υ meson ($b\bar{b}$) [37]. Quark sector of the SM was completed almost 20 years later (in 1995) when CDF and D0 experiments at Fermilab observed t quark - the most massive fundamental particle known up to now [1, 2]. One can naturally raise question if there could not exist next generation of quarks without corresponding generation in lepton sector. The SM framework, however, gives definite answer. No. It turns out that these anomalies would cause non-renormalizability of the weak interactions, which require exactly the same numbers of generations for both quarks and leptons. Moreover, the experimental measurements of the Z boson decay width [38] at Large Electron-Positron Collider (LEP) provided evidence for existence of only three light neutrino generations (with mass lower than $m_Z/2$).

It should be noted that there is one important extra property of quarks when compared to leptons. This property (quantum number) is called color and takes a role of generalized charge in strong interactions. Quarks of each flavor come in three varieties, either red, green or blue, thus creating triplets of particles (for example u_r, u_g, u_b). Due to the confinement of quarks mentioned before, only colorless object can be observed in nature.

1.3 Bosons

Another important group of particles in the SM is formed from the mediators of the fundamental interactions, called gauge bosons.

The electromagnetic interaction was extensively studied in the second half of 19th century, when Maxwell [39] proposed for the first time that electromagnetism could

be mediated by light quanta, photons (γ). The wave-particle nature of light was then supported by many renowned physicists of that times, like Einstein, Planck, de Broglie and others. The final confirmation that light does not behave only as a wave, but it needs to be described also through particle concepts, was given by series of Milikan's experiments [40]. Finally, the quantum theory of electromagnetism that was first developed by Dirac in 1927 [41], was completed in 1949 [42] and is known today as quantum electrodynamics (QED). Photons are observed to be massless and carry zero electric charge.

The weak interactions are, on the other hand, mediated by massive $W^{\pm}$ and Z bosons. These intermediate particles were proposed as a part of the unified theory of electromagnetism and weak interactions, which was developed by Glashow, Weinberg and Salam in 1960s (see Section 1.4.3 for more details). $W^{\pm}$ bosons were discovered [43, 44] by UA1 and UA2 experiments at CERN only few months before the observation of Z boson [45, 46]. Both mediators belong among heaviest elementary particles with mass of around 80.4 and 91.2 GeV, respectively [21]. These particles are therefore very unstable and can decay into a pair of less massive particles (except top quark which is even more massive and thus kinematically forbidden). $W^{\pm}$ boson can decay either leptonically to lepton and its corresponding antineutrino, but more probably hadronically to quark-antiquark pair. The Z boson decays into same flavor fermion-antifermion pairs. More details about Z boson decay modes are given in the next chapter devoted to top quark and Z boson.

The mediators of the strong interactions, gluons, were discovered even before $W^{\pm}$ and Z bosons, in spring of 1979 by the PETRA collider at DESY (Deutsches Electronen-Synchrotron) [47]. Their theoretical foundation was laid by Murray Gell-Mann in 1962 [48], who predicted eight independent mediators. Gluons, similar to photons, are massless. However, there is one important property of gluons, which differs it from photons. Gluons themselves carry the color charge, what means that they not only mediate the strong interactions but directly participate in them, what makes quantum theory of strong interactions much more complex than QED. This also implies that gluons cannot be observed as free particles (due to the confinement), and consequently their experimental evidence is only in a form of gluon jets [7].

The last member of the boson family, Higgs boson, plays a special role in the SM. It is a massive scalar boson (with zero spin) with no color and electric charge. Its importance lays in the fact that thanks to the Higgs mechanism (discussed in Section 1.4.4) [49, 50] gauge bosons described above acquire mass. Moreover, fermions acquire mass through their Yukawa coupling to the Higgs field. The Higgs boson has mass of around 125 GeV [21], meaning that it gains mass through the process of self-interaction

[7] Jet refers to the narrow cone of particles produced by the quark or gluon in a process of hadronization.

with its own field. The experimental discovery of the Higgs boson was announced by ATLAS [51] and CMS [52] experiments at CERN in year 2012, hence completing the set of particles predicted by the SM. Since the theoretically predicted decay width of the Higgs boson is around 4.2 MeV [53], it is highly unstable and can be studied only via its decay products. Among many decay channels, the most important are decays into pair of γ quanta and ZZ pair with subsequent decay into two pairs of leptons. Although these decay modes do not have highest branching ratios, they are not contaminated with high background rate and are thus most suitable for experimental measurements.

1.4 Basic particle interactions in the Standard Model

The three fundamental interactions, strong, weak and electromagnetic, are understood as arising due to the exchange of spin-1 bosons between spin-$\frac{1}{2}$ particles that make up matter. Interactions in the SM are based on so called gauge symmetry group

$$SU_c(3) \otimes SU_L(2) \otimes U_Y(1). \tag{1.1}$$

The gauge bosons, mediators of the specific interactions, are associated with the generators of the group algebra. The strong interaction is mediated by eight massless gluons that are associated with eight free parameters of the $SU_c(3)$ group. Subscript "c" is meant to denote color, what suggests that gluons couple only to colored objects, quarks. Since gluons carry both color and anticolor, the $SU_c(3)$ symmetry group leads to the octet of gluons, which mediate the strong interaction by the change of color, and one color singlet state that does not change the color state and thus cannot mediate the strong interaction. There are three generators connected with the weak interactions and one generator corresponding to the electromagnetic force. The subscript "L" in $SU_L(2)$ indicates that only left-handed fermions [8] couple to these gauge bosons. On the other hand, subscript "Y" in $U_Y(1)$ is meant to distinguish the group associated with the so-called weak hypercharge (explained below), from that with electric charge Q. Since weak and electromagnetic interactions were unified in the SM, gauge bosons corresponding to electroweak interaction ($W^{\pm}, Z$ bosons and photon) are constructed as a mixture of the $SU_L(2)$ and $U_Y(1)$ generators.

Particles participating in the interactions arise as a representations of the symmetry groups in Eq. 1.1. Their transformation properties impose the local invariance of the nature's Lagrangian. It is the procedure of promoting global gauge symmetry, θ, of the SM Lagrangian to a local gauge symmetry, $\theta(x)$, that spawns interaction terms by implementing new fields corresponding to the mediating particles. An example of this procedure for a generic theory and global symmetry group will be outlined in

[8] This term is explained in Section 1.4.3

the following section. This part follows naming conventions and is heavily inspired by [54, 55].

1.4.1 From global to local symmetry

Taking a general field ψ in a Lagrangian and requiring invariance under general global symmetry operation θ, the transformation of the field follows the scheme:

$$\psi \rightarrow e^{ig\theta^c\tau^c}\psi, \tag{1.2}$$

where τ^c are the generators of the symmetry group and g is corresponding coupling constant. For the generators τ^c, the following commutation relation holds:

$$[\tau^a, \tau^b] = if^{abc}\tau^c \tag{1.3}$$

where f^{abc} stands for structure constants of the group, which are equal to zero for abelian groups and index c goes through all free parameters of the group. To introduce the local invariance at the top of global invariance, the following additions need to be implemented in the theory in order to preserve the invariance:

1. For each generator of the symmetry group there exists a corresponding gauge field.

2. Interaction terms resulting from 1) imply the transformation of the covariant derivative present in Lagrangian with an additional term.

3. Newly added gauge fields from 1) require additional kinematic and mass terms.

As already mentioned, the SM consists of three symmetry groups, $SU_c(3), SU_L(2)$ and $U_Y(1)$ that give rise to the electroweak and strong gauge sectors, which are orthogonal. Following sections briefly describe basic features of the quantum theories connected with these two interactions.

1.4.2 Strong interaction

The quantum theory underlying the strong interaction is called quantum chromodynamics (QCD). The long series of thoeries and findings which shaped the final formulation of the QCD began in 1950s and 1960s when a large number of strongly interacting particles was discovered (see Section 1.2.2 for more details). The theoretical foundation of the QCD was laid by Gell-Mann [29] and Zweig [30], when they proposed existence of the three quarks as an explanation for the classification of the known mesons and baryons of that time. However, their theories violated the Pauli exclusion principle by requiring the three identical fermions to be in the same quantum state. In reaction

to this findings, Nambu [56] and Greenberg [57] proposed the solution by introducing additional quantum number which originated from the SU(3) gauge symmetry. The theory of strong interactions was finalised in 1973 [58] when one of the key features of the QCD, asymptotic freedom (explained below), was discovered [59, 60].

QCD is a non-abelian gauge theory and is characterized by the SU(3) - special unitary group. Following the recipe described in the previous section, the local gauge invariance of this symmetry group requires introducing quark and gluon fields denoted as $q_{f,\alpha}$ and G^a_μ where f stands for quark flavor (u, d, c, s, t, b), α for color index (red, green, blue) and a for the index of gluon color octet $(a = 1, ..., 8)$. Since leptons and other gauge bosons do not carry color charge, they transform as singlets under $SU_c(3)$ and thus do not participate in strong interactions.

The $SU_c(3)$ group algebra is characterized by the similar commutation relation as in Eq. 1.3. In the case of QCD the generators of the group are T^a (analogous to τ^a from mentioned commutation relation) and can be expressed in terms of the Gell-Mann matrices λ^a, as $T^a = \lambda^a/2$. These traceless Hermitian matrices of the shape 3×3, corresponds to the eight gluons.

The behavior of the free fermion is governed by the so called Dirac Lagrangian that is defined as follows:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \tag{1.4}$$

where the fermion field with mass m is denoted by ψ and γ^μ are the Dirac matrices. To make the Dirac Lagrangian locally invariant, the derivative ∂_μ must be transformed into a covariant derivative, D_μ, in a following way:

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu - ig_s G^a_\mu T^a, \tag{1.5}$$

where g_s represents the strong coupling constant. As a consequence of adding new gluon fields to the Lagrangian, the gluon kinetic term needs to be added in a form:

$$\mathcal{L}_{kin,QCD} = -\frac{1}{4} G^a_{\mu\nu} G^{a\,\mu\nu}, \tag{1.6}$$

where $G^a_{\mu\nu}$ is a gluon field tensor with following definition:

$$G^a_{\mu\nu} \equiv \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc} G^b_\mu G^c_\nu. \tag{1.7}$$

Then the Lagrangian of the QCD takes on the form:

$$\mathcal{L}_{QCD} = -\frac{1}{4} G^a_{\mu\nu} G^{a\,\mu\nu} + \sum_f \bar{q}_f i\gamma^\mu D_\mu q_f \tag{1.8}$$

The resulting strong interactions in the SM consist of couplings between quarks and gluons arising from the second term in Eq. 1.8. However, thanks to the non-abelian nature of the QCD, the first term in the QCD Lagrangian gives rise also to the three- and four-point self-interaction between gluons as shown in Figure 1.2.

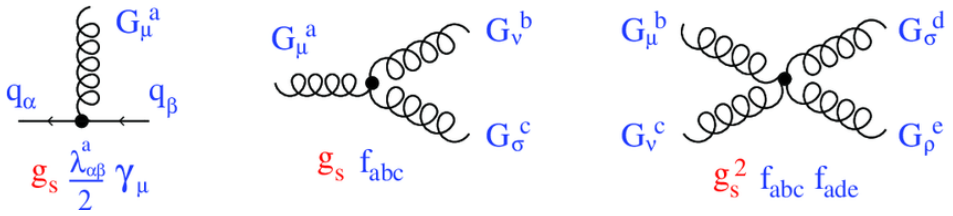


Fig. 1.2: Feynman diagrams of vertices allowed in QCD processes. Quark-gluon coupling on the left, three-point gluon self-interaction in the middle and four-point self-interaction on the right [61].

Renormalizable gauge theories which are essential for the SM have one important property. The dependence of the interaction strength g on the energy scale μ can be expressed by the so-called beta function:

$$\beta(g) \equiv \frac{\partial g}{\partial \log(\mu)}, \tag{1.9}$$

This dependence is usually referred to as the running of the coupling constant (which is in fact not a constant), but it is important to note that it does not predict absolute value of the coupling [9], only its evolution with the energy scale. For the generic non-abelian gauge theory the full derivation of beta function [54] yields the following result:

$$\beta(g) = -\left(\frac{11}{3}T(A) - \frac{4}{3}n_f T(R)\right)\frac{g^3}{16\pi^2} + \mathcal{O}(g^5), \tag{1.10}$$

where $T(A)\delta^{ab} = \mathrm{Tr}(T_A^a T_A^b)$, and $T_A^{(a,b)}$ stands for adjoint representations of the group generators. Factor $T(R)$ denotes the index of the representation and is defined as $T(R)\delta^{ab} = \mathrm{Tr}(T_R^a T_R^b)$, where Hermitian matrices T_R obey the commutation relations valid for particular group generators. n_f in Eq. 1.10 represents number of fermions which participate in the interactions at given energy scale μ.

When considering theory of strong interactions with the group $SU_c(3)$, then $T(A)$ is number of colors (3) and $T(R) = \frac{1}{2}$ resulting in following beta function:

$$\beta(g) = -\left(11 - \frac{2}{3}n_f\right)\frac{g^3}{16\pi^2} + \mathcal{O}(g^5). \tag{1.11}$$

From this equation one can immediately deduce that if $n_f \leq 16$, the beta function yields negative values. This has direct consequence on the behavior of the coupling strength α_S, which is related to the coupling constant g_S as follows:

$$\alpha_S \equiv \frac{g_S^2}{4\pi}. \tag{1.12}$$

For the QCD, with maximum six fermions at energies higher than top-quark mass, there is inverse dependence of the coupling strength on the energy scale Q^2. This can

[9] Absolute value has to be determined experimentally at particular energy.

be obtained by integrating Eq. 1.9 up to the QCD energy scale $\Lambda_{\mathrm{QCD}} \sim 200$ MeV, when the QCD perturbative regime starts to be invalid. The resulting dependency of the coupling strength α_S on the energy scale can be then written as:

$$\alpha_S(Q^2) = \frac{12\pi}{(11N_C - 2n_f)\log\left(\frac{Q^2}{\Lambda^2_{\mathrm{QCD}}}\right)} \tag{1.13}$$

As a consequence of this inverse dependency on the energy scale, at low energies below 1 GeV, the coupling strength increases to around order one and perturbative regime is no more valid for QCD. This is widely known as so-called *confinement* of quarks [62], when they cannot exist as free particles and hence create hadrons. The opposite behavior can be observed at high energies when QCD coupling decreases rapidly and quarks act as free particles. This feature of QCD is called *asymptotic freedom* [63] and was theoretically predicted in 1970s similarly as quark confinement.

1.4.3 Electroweak interaction

The electroweak theory (also known as Glashow Weinberg Salam (GWS) theory) is a unified theory of the weak and electromagnetic interaction. The foundation of this theory was laid by Glashow in 1961 [64], when he proposed the model to combine weak and electromagnetic interaction in the framework of $SU_L(2) \otimes U_Y(1)$ symmetry. The unified theory was completed few years later by Weinberg and Salam [65, 66], who supplemented the Higgs-Englert-Brout mechanism [49, 50] to generate masses of fermions and gauge bosons (more details in the next Section 1.4.4) and thus placed the model in the mathematical framework of gauge theories. Renormalizability of the new theory was proved by 't Hooft [67, 68].

Similar to the $SU_c(3)$ symmetry group, also product $SU_L(2) \otimes U_Y(1)$ is non-abelian. However, the symmetry is spontaneously broken at the electroweak (EW) scale of around 100 GeV, resulting in single $U_Q(1)$ symmetry that governs QED. The process of EW symmetry breaking (EWSB) is further discussed in next section.

In analogy to the generic recipe outlined in Section 1.4.1, one can derive gauge bosons and interaction terms for electroweak theory. As already mentioned, the subscript L in $SU_L(2)$ denotes chirality of the fermions that are allowed in weak interactions. While fermions with left chirality transform as doublets under $SU_L(2)$, the right-handed (from the chirality point of view) fermions transform only as singlets, thus excluding them from charged weak interactions. To transform generic field ψ into either left-handed (ψ_L) or right-handed (ψ_R) component, one can use so-called

projection operators P_L and P_R, which are defined in following way:

$$\begin{aligned} \psi_L &= P_L\psi = \frac{1}{2}(1-\gamma^5)\psi, \\ \psi_R &= P_R\psi = \frac{1}{2}(1+\gamma^5)\psi, \end{aligned} \tag{1.14}$$

where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ and $\gamma^{(0,1,2,3)}$ are the Dirac matrices.

The theory of weak interactions introduced the new quantum number, called weak isospin (T). This number is used to distinguish left-handed doublets with $T = \frac{1}{2}$ (third component of weak isospin $T_3 = \pm\frac{1}{2}$) from right-handed singlets with $T = T_3 = 0$. Since $SU_L(2)$ is the special unitary group of order 2, its generators can be expressed in terms of operators associated to the weak isospin as $\hat{T}_i = \frac{1}{2}\sigma_i$, where σ_i represents 2×2 complex hermitian Pauli matrices.

The electromagnetic interactions are governed, after EWSB, by the $U_Q(1)$ group resulting from the breaking of the $SU_L(2) \otimes U_Y(1)$, where $U_Y(1)$ represents simple unitary group and Y stands for the weak hypercharge, the new quantum number analogous to the weak isospin mentioned above. The third component of the weak isospin T_3 relates to the Y after EWSB as follows:

$$Y = Q - T_3, \tag{1.15}$$

where Q represents electric charge.

To fulfill the requirements of the local invariance, the derivative ∂_μ has to be expanded to covariant derivative D_μ with additional terms representing new gauge fields:

$$\partial_\mu \rightarrow D_\mu \equiv \partial_\mu - igW_\mu^i\frac{\sigma^i}{2} - ig^{'}B_\mu Y, \tag{1.16}$$

where the gauge fields W_μ^i $(i = 1, 2, 3)$ and B_μ are associated with $SU_L(2)$ and $U_Y(1)$ symmetry groups, respectively. Factors g and $g^{'}$ denote coupling strengths of respective fields. As a consequence of this transformation, the new kinetic terms arise in the Lagrangian:

$$\mathcal{L}_{kin,EW} = -\frac{1}{4}W_{\mu\nu}^i W^{i\,\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}. \tag{1.17}$$

The field tensors $W_{\mu\nu}^i$ and $B_{\mu\nu}$ related to $SU_L(2)$ and $U_Y(1)$ are defined as follows:

$$\begin{aligned} W_{\mu\nu}^i &\equiv \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon^{ijk}W_\mu^j W_\nu^k, \\ B_{\mu\nu}^i &\equiv \partial_\mu B_\nu - \partial_\nu B_\mu, \end{aligned} \tag{1.18}$$

where ϵ^{ijk} is fully antisymmetric permutation tensor, called Levi-Civita symbol.

The non-abelian nature of the $SU_L(2)$ enables self-interactions, similarly as gluons self-couple in QCD. On the other hand, the weak coupling strength g does not behave similar to QCD, but it increases rapidly with energy scale at tree level. This

ultimately results in violation of the unitarity at the TeV scale. To preserve the unitarity, the corrections of the higher orders would therefore need to be of the same order of magnitude as contributions from tree level diagrams. This would, however, cause the weak interactions to be strongly-coupled, which apparently does not agree with observed behavior. The solution has been found through implementation of the Higgs-Englert-Brout mechanism described next.

After adding strong and EW interactions to the free fermion field, the Dirac Lagrangian defined in Eq.1.4 becomes:

$$\begin{aligned}\mathcal{L} = i\bar{\psi}\gamma^{\mu}(\partial_{\mu} - ig_s G^a_{\mu} T^a - \frac{i}{2} g W^i_{\mu}\sigma^j - ig^{'} B_{\mu} Y)\psi \\ -\frac{1}{4} G^b_{\mu\nu} G^{b\,\mu\nu} - \frac{1}{4} W^k_{\mu\nu} W^{k\,\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}.\end{aligned} \tag{1.19}$$

1.4.4 Higgs-Englert-Brout mechanism

When considering Lagrangian constructed so far in Eq.1.19, it clearly miss any mass terms of the form $m\bar{\psi}\psi$ for fermions, and $m^2 F^i_{\mu} F^{i\,\mu}$ for gauge bosons. Since various experiments throughout the history of particle physics showed that fermions are not massless, the full Lagrangian of the SM should be able to account for such observations. However, simple addition of the mass terms would break the local invariance of the gauge transformations. Solution was found by Higgs, Englert and Brout in year 1964 [49, 50] when they introduced the new scalar field and the procedure of spontaneous breaking of the EW symmetry. The Higgs-Englert-Brout mechanism not only solves the problem of high masses observed for $W^{\pm}$ and Z bosons, but cures also unitarity violation in weak interactions mentioned above, while still preserving the local gauge invariance.

The process of the EWSB is briefly discussed here for the EW sector, while the similar (but a bit more complex) procedure can be applied to incorporate mass terms for fermions. The process of inference can be started from the Maxwell Lagrangian describing massless gauge boson fields A^{μ}:

$$\mathcal{L}_{Maxwell} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu}, \tag{1.20}$$

where $F^{\mu\nu} \equiv \partial^{\mu} A_{\nu} - \partial^{\nu} A_{\mu}$ is electromagnetic field strength tensor. Standard procedure continues by replacing derivative ∂_{μ} by the covariant derivative D_{μ} in the following way:

$$\partial_{\mu} \rightarrow D_{\mu} \equiv \partial_{\mu} - igA_{\mu}, \tag{1.21}$$

where g again denotes coupling strength. After introducing complex scalar field ϕ the Maxwell Lagrangian becomes:

$$\mathcal{L} = (D^{\mu}\phi)^{\dagger} D_{\mu}\phi - V(\phi) - \frac{1}{4} F^{\mu\nu} F_{\mu\nu}. \tag{1.22}$$

$V(\phi)$ is the scalar potential of field ϕ with following definition:

$$V(\phi) = m^2\phi^\dagger\phi + \frac{1}{4}\lambda(\phi^\dagger\phi)^2, \tag{1.23}$$

where λ and m^2 are free parameters. The Lagrangian in Eq.1.22 is apparently invariant under $U(1)$ transformations of the form:

$$\phi(x) \to e^{-i\alpha}\phi(x), \tag{1.24}$$

The potential in Eq.1.23 is at its minimum when (assuming parameter $m^2 < 0$):

$$\phi(x) = \sqrt{\frac{2|m^2|}{\lambda}}e^{-i\theta} = \frac{1}{\sqrt{2}}ve^{-i\theta}, \tag{1.25}$$

where arbitrary phase θ parametrizes the collection of minima, and at the same time introduces the $U(1)$ transformation in Eq.1.24 by $\theta \to \theta + \alpha$. By choosing $\theta = 0$, the non-zero vacuum expectation value (VEV) yields following value:

$$\langle 0|\phi(x)|0\rangle = \frac{1}{\sqrt{2}}v. \tag{1.26}$$

Since ϕ is a complex field, it can be rewritten through two real scalar fields $\chi(x)$ and $\rho(x)$:

$$\phi(x) = \frac{1}{\sqrt{2}}(v + \rho(x))e^{-i\chi(x)/v}. \tag{1.27}$$

It is clear that dependence on $\chi(x)$ is canceled out in Eq.1.23, and therefore $\chi(x)$ represents massless field, often reffered to as *Goldstone boson*. Thanks to the gauge freedom invoked by the local invariance, the ϕ phase can be arbitrarily shifted in spacetime. Therefore the Goldstone field $\chi(x)$ can be set to zero, what results in the break of the $U(1)$ symmetry. This choice is called *unitary gauge* and such fields, whose VEV break the $U(1)$ symmetry, are called Higgs fields. When now ϕ is substituted into the first term of Eq.1.22, the mass term emerges in the following form [10]:

$$-\frac{1}{2}g^2v^2A^\mu A_\mu = -m_A^2A^\mu A_\mu. \tag{1.28}$$

In this way originally massless field A^μ gained mass $m_A = gv/\sqrt{2}$.

This procedure can be now applied to the previously derived Lagrangian from Eq.1.19 that contains the strong, electroweak and fermion fields. In the first step the analogous complex scalar field Φ has to be introduced in the following form:

$$\Phi = \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_1 + i\phi_2\\ \phi_3 + i\phi_4\end{pmatrix}, \tag{1.29}$$

[10] Also incorporating transformation from Eq.1.21

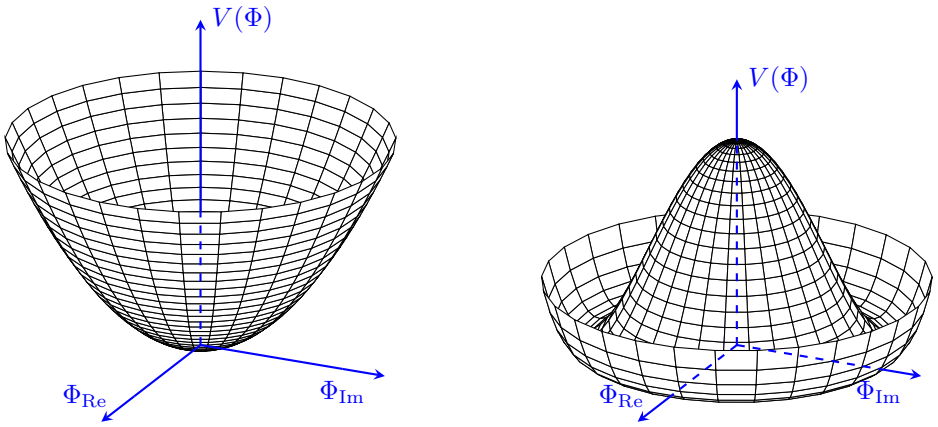


Fig. 1.3: The shape of the Higgs-Englert-Brout potential for $\mu^2 > 0$, $\lambda > 0$ (left) and for $\mu^2 < 0$, $\lambda > 0$ (right).

where the superscripts + and 0 represent electric charges. This field, later identified as Higgs field, can be now incorporated into the Lagrangian in Eq.1.19 using the following additional term:

$$\Delta\mathcal{L}_{Higgs} = (D_\mu\Phi)^\dagger(D^\mu\Phi) - V(\Phi), \tag{1.30}$$

where D_μ follows the same definition as in Eq.1.16. The potential $V(\Phi)$, often referred to as Higgs-Englert-Brout potential is defined in analogy with Eq.1.23:

$$V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2. \tag{1.31}$$

The parameters λ and μ^2 govern the shape of potential as sketched in Figure 1.3. When $\lambda < 0$ the potential does not possess any stable minima and therefore is treated as unphysical. If $\lambda > 0$ and $\mu^2 > 0$, the potential has one stable minimum at $\Phi = 0$. But the most important case (from physics point of view) is with $\lambda > 0$ and $\mu^2 < 0$ when shape of the potential resembles mexican hat and contains infinite set of minima for which:

$$\Phi^\dagger\Phi = \frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}. \tag{1.32}$$

Procedure continues by choosing unitary gauge that fix the VEV. Moreover, one of the components of the complex scalar field is set to zero and resulting VEV can be defined as follows:

$$\langle 0|\Phi(x)|0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} v \\ 0 \end{pmatrix}. \tag{1.33}$$

When the chosen VEV is injected back into Eq.1.30 and the covariant derivative is expanded according to Eq.1.16, the mass terms arise from the following Lagrangian:

$$\mathcal{L}_{mass} = -\frac{1}{8}v^2 \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} gW_\mu^3 - g^{'}B_\mu & g(W_\mu^1 - iW_\mu^2) \\ g(W_\mu^1 + iW_\mu^2) & -gW_\mu^3 - g^{'}B_\mu \end{pmatrix}^2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{1.34}$$

where the gluon fields were dropped since Φ does not carry the color charge, resulting in zero mass for gluons. To obtain the mass eigenstates corresponding to the EW mediators, the matrix in the previous Lagrangian has to be diagonalized. Then the gauge fields corresponsing to the physical $W^{\pm}$, Z bosons and photon arise in the following form:

$$\begin{aligned} W_{\mu}^{\pm} &\equiv \frac{1}{\sqrt{2}}(W^1 \mp iW^2), \\ Z_{\mu} &\equiv c_W W_{\mu}^3 - s_W B_{\mu}, \\ A_{\mu} &\equiv s_W W_{\mu}^3 + c_W B_{\mu}, \end{aligned} \tag{1.35}$$

where $s_W = \sin\theta_W$ and $c_W = \cos\theta_W$ and θ_W is weak mixing angle defined as:

$$\theta_W \equiv \tan^{-1}\frac{g^{'}}{g}. \tag{1.36}$$

When the above mass eigenstates are substituted back into Eq.1.34 the $W^{\pm}$ and Z mass terms finally emerge:

$$\mathcal{L}_{mass} = \frac{(gv)^2}{4} W^{+\mu} W_{\mu}^{-} - \frac{(gv)^2}{8c_W^2} Z^{\mu} Z_{\mu}. \tag{1.37}$$

The masses of $W^{\pm}$ and Z bosons can be then identified as:

$$\begin{aligned} m_{W^{\pm}} &= \frac{gv}{2}, \\ m_Z &= \frac{gv}{2c_W} = \frac{m_{W^{\pm}}}{c_W}. \end{aligned} \tag{1.38}$$

It is important to note that photon gauge field A_{μ} does not acquire mass term and hence the $U_Q(1)$ symmetry is not broken and only $SU_L(2) \otimes U_Y(1)$ breaks.

The only remaining gauge boson that was not yet mentioned in this derivation, and which is observed to be most massive, is Higgs boson. Its mass term emerges immediately when we choose a bit different, but still unitary, gauge:

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{1.39}$$

where H represents a real scalar field. Using this gauge the Higgs-Englert-Brout potential takes on following form:

$$V(\Phi) = \frac{1}{4}\lambda v^2 H^2 + \frac{1}{4}\lambda v H^3 + \frac{1}{16}\lambda H^4, \tag{1.40}$$

where Higgs boson mass:

$$m_H = v\sqrt{\frac{\lambda}{2}}, \tag{1.41}$$

reveals itself in the quadratic term. Note that remaining terms correspond to the three- (cubic term) and four-point (quartic term) interactions of Higgs boson with itself, similar to gluon self-interactions discussed previously.

The derivation of the quark and lepton masses follows the same strategy and is discussed in detail e.g. in Ref. [54], while here only the brief summary is given. The procedure starts by introducing fermion fields corresponding to leptons and quarks, so-called *Weyl fields*. Quarks are represented by fields $q^a_\alpha, \bar{u}^a_\alpha, \bar{d}^a_\alpha$ [11], that correspond to representations of the SM gauge group $SU_c(3) \otimes SU_L(2) \otimes U_Y(1)$ and a denotes index of the generation ($a = 1, 2, 3$), while α stands for color. Analogously, lepton fields are introduced as l^i and $\bar{e}^i$, where i again denotes particular lepton generation.

The generic fields $q^a_{\alpha,L}$ and l^i_L (subscript L denoting left-handed chiral states) are $SU_L(2)$ doublets that can be expressed as follows:

$$q^a_{\alpha,L} = \begin{pmatrix} u^a_\alpha \\ d^a_\alpha \end{pmatrix}_L , \; l^i_L = \begin{pmatrix} \nu^i \\ l^i \end{pmatrix}_L , \tag{1.42}$$

while other fermion fields (with right-handed chirality) are defined as $SU_L(2)$ singlets. To directly obtain a mass term, it is necessary to find product of fermion fields in a form $mf^\dagger f$ that would transform as singlet under the SM symmetry group (Eq. 1.1), thus ensuring gauge invariance and renormalizability. Howver, the mass term can be obtained through so-called *Yukawa* coupling between Higgs field Φ and particular fermion field. This can be expressed for quarks and leptons as follows, respectively:

$$\begin{aligned} \mathcal{L}^q_{\text{Yuk}} &= -\epsilon^{ij}\Phi_i q^a_{\alpha j} y^{ab} \bar{d}^{\alpha b} - \Phi^{\dagger i} q^a_{\alpha i} y'^{ab} \bar{u}^{\alpha b} + \text{h.c.}, \\ \mathcal{L}^l_{\text{Yuk}} &= -\epsilon^{pq}\Phi_p l^i_q y''^{ij} \bar{e}^j + \text{h.c.}, \end{aligned} \tag{1.43}$$

where Yukawa couplings y, y', y'' are represented by 3×3 complex matrices, and ϵ^{ij} denotes two-dimensional antisymmetric Levi-Civita tensor. These Yukawa lagrangians contain field products that yield singlet representation, thus guaranteeing gauge invariance. Moreover, these are only terms of dimension four or less, that feature this property and can be constructed from fermion and Higgs fields.

The procedure continues with the step of EWSB when the Higgs field is chosen to obey unitary gauge from Eq. 1.39. By substituting the components of the Higgs field in unitary gauge into Eq. 1.43 and by writing fermion fields in the spinor notation as Dirac fields:

$$D^a_\alpha = \begin{pmatrix} d^a_\alpha \\ \bar{d}^{\dagger a}_\alpha \end{pmatrix} , \; U^a_\alpha = \begin{pmatrix} u^a_\alpha \\ \bar{u}^{\dagger a}_\alpha \end{pmatrix} , \; L^a_\alpha = \begin{pmatrix} l^i \\ \bar{l}^{\dagger i} \end{pmatrix} , \tag{1.44}$$

then the mass and interaction terms finally emerge in the following form:

$$\begin{aligned} \mathcal{L}^Q_{\text{Yuk}} &= -\frac{1}{\sqrt{2}}(v + H)\bar{D}^{a\alpha} y^{ab} D^b_\alpha - \frac{1}{\sqrt{2}}(v + H)\bar{U}^{a\alpha} y'^{ab} U^b_\alpha + \text{h.c.}, \\ \mathcal{L}^L_{\text{Yuk}} &= -\frac{1}{\sqrt{2}}(v + H)\bar{L}^i y''^{ij} L^j + \text{h.c.}, \end{aligned} \tag{1.45}$$

[11] Note that the bar over fermion field does not mean any kind of conjugation and is used only for naming convention.

The up and down quarks of each generation, and charged leptons thus acquire mass after applying unitary transformations to their respective Dirac fields. These products are then aligned with the diagonal elements of the Yukawa matrices, which are positive and real, what results in following masses:

$$m_{d^a} = \frac{1}{\sqrt{2}} y^{aa} v, \; m_{u^a} = \frac{1}{\sqrt{2}} y'^{aa} v, \; m_{l^i} = \frac{1}{\sqrt{2}} y''^{ii} v. \tag{1.46}$$

As can be seen, the neutrino fields remain massless due to its cancellation with the Higgs field in Eq. 1.43, after applying unitary gauge.

1.5 Limitations of the Standard Model

Although the SM is the theory of nature that helps to understand the universe at the subatomic level, there is number of phenomena which cannot be explained by the framework of the SM. This success of the SM can be illustrated in the Figure 1.4, that shows compatibility between measured cross sections of the various SM processes and their SM predictions. This section, however, briefly summarizes the open and unresolved questions for which SM has no answers.

- **Gravity** - One of the most striking problems of the SM is its inability to include gravitational interactions into its theoretical framework. Although there is extremely successful theory of general relativity that describes macroscopic systems, no such equivalent is currently known for the subatomic particles that could be incorporated into the SM. The so-called *string* theories include also gravity but these cannot be easily experimentally tested nowadays.

- **Dark matter and dark energy** - Many astrophysical observations suggest that the ordinary matter which is governed by the SM makes up only 5% of the observed energy in the universe. The remaining ≈68% is included in the so-called dark energy [70], and 27% in the dark matter. It is expected that dark energy is responsible for the accelerated expansion of the universe. This property of the space was confirmed by the experiments studying cosmic microwave background [71, 70], as well as by supernovae observations [72]. Existence of the dark matter was proved by cosmological observation, such as anomalies in the rotation curves [73] of stars in galaxies or collision of the galaxy clusters [74].

- **Asymmetry between matter and antimatter** - Next phenomenon which is still to be understood is the observed asymmetry between baryonic matter, which dominates, and antimatter. This asymmetry can be explained by taking into account so-called Sakharov conditions [75]. Among these conditions is also

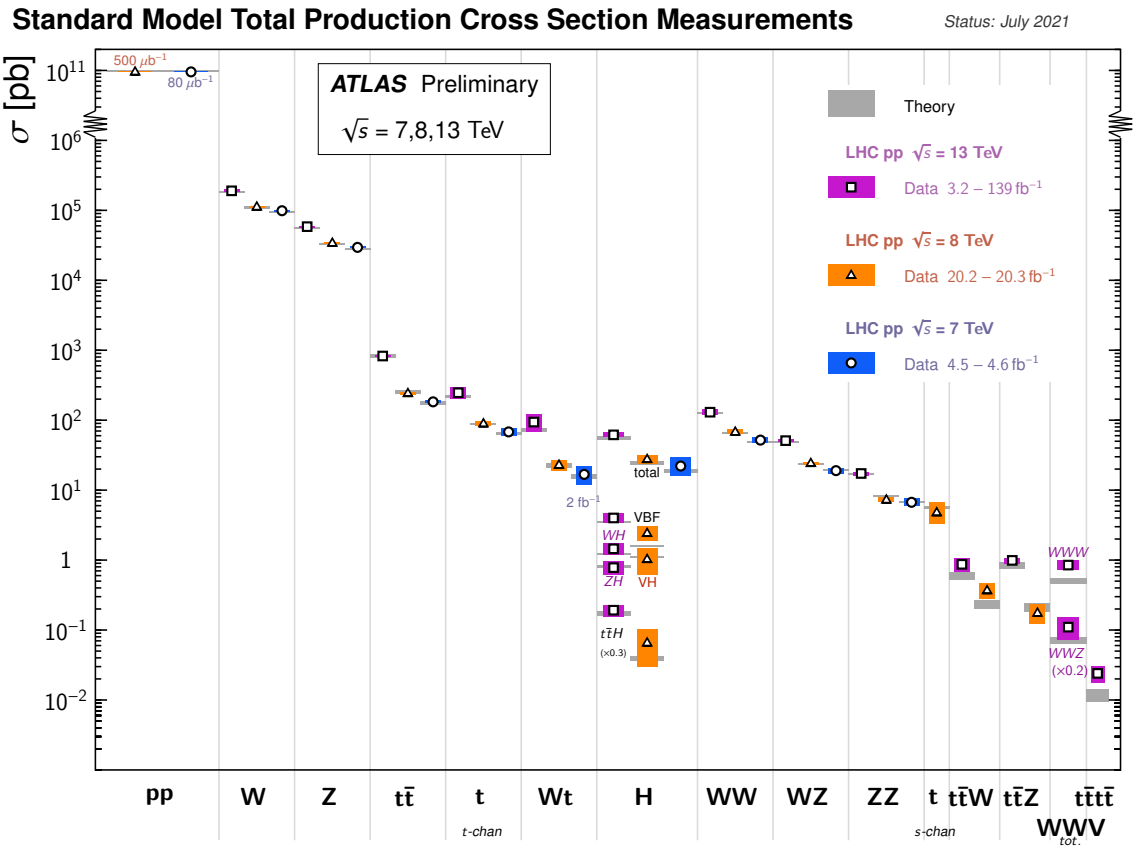


Fig. 1.4: Summary of the theoretical predictions of the cross sections for the various SM processes and their corresponding values as measured by the ATLAS Collaboration. The measured values are corrected for branching fractions. Figure is taken from [69].

violation of the baryon number, what contradicts with the assumptions of the SM.

- **Neutrino masses** - The SM as was formulated in this chapter assumes zero masses of the neutrinos. As already mentioned this was proven to be wrong assumption thanks to the experiments studying neutrino oscillations. Although non-zero mass for the neutrinos can be incorporated relatively easily into the SM through introducing right-handed neutrino fields, no experimental evidence favoring this addition has been found yet.

- **Theoretical issues** - At the top of mentioned phenomenological issues there are also some theoretical problems that arise from the SM framework. One of the most pressing ones is so-called hierarchy problem [76]. The essence of this problem lies in the concept of naturalness that should be present for the complete theory of nature. It requires the ratios between two parameters to be of the order one. This is, however, not the case when value of the parameter depends on the cancellation of two other terms, what then leads to the need of fine tuning. On the other hand, the natural theory should not require a lot of fine tuning. This problem manifests itself in extraction of the Higgs boson mass that requires subtraction of the two large terms, each of the order of 10^{19}. Similar problem is present also for the fermion masses that are dictated by their Yukawa couplings. These couplings span several orders of magnitude, what again violate the naturalness principle.

Since it is clear that the SM is not the definitive theory of nature, alternative frameworks, which try to address the mentioned issues, are actively developed. These so-called Beyond Standard Model (BSM) theories usually suggest existence of new, not yet discovered, particles. For example, the supersymmetry theories (SUSY) [77, 78] predict superpartners for each SM fermion and boson by introducing the new symmetry between them. The problem with most of the SUSY theories lies in the fact that they contain many (> 100) free parameters and no superparticles were discovered so far. Other alternative theories are based on the existence of new, hidden, dimensions (Kaluza-Klein theory [79, 80]) or propose compositeness of the known SM particles (Preon models [81]).

Chapter 2

The physics of top quark and Z boson at the LHC

Since this thesis is aimed at the associated production of top-quark pair and Z boson, their production and decay mechanism will be briefly introduced in this chapter. The motivation for this analysis is also explained in this chapter.

2.1 Top quark and its production at the LHC

Top quark [82], which belongs to the third generation of quarks, plays a special role in particle physics. It has the largest mass among elementary particles, $m_t = 173.34 \pm 0.27(\text{stat}) \pm 0.71(\text{syst})$ GeV [3][1], while it significantly overcomes other quarks[2]. Its mass was measured with precision better than 0.5% what makes it the quark with the most precisely measured mass. Despite its high mass, no experiment observed its internal structure till this time, and thus top quark is considered to be a point-like particle. Moreover, it has also large decay width, its theoretical value in the SM is $\Gamma_t \approx 1.322$ GeV (calulated at next-to-next-to-leading order (NNLO) in QCD using $m_t = 172.5$ GeV) [83]. As a consequence, since decay width is inversely proportional to the mean lifetimes of the particle, the top quark has one of the shortest mean lifetime among elementary particle. Its value is $\tau_t \approx 5 \times 10^{-25}$ s, what is less than the time needed to produce bounded state in QCD, $\tau_{QCD} \approx 3 \times 10^{-24}$ s. This implies that the top quark decays before hadronisation, what enables us to study properties of a 'pseudo-bare' quark.

According to the SM, top quarks are produced at hadron colliders predominantly in form of top-antitop-quark pairs ($t\bar{t}$) through QCD strong interaction [3], either quark-

[1] Result of combined measurement on data from CDF and D0 detectors at the Tevatron accelerator and ATLAS and CMS detectors at the LHC.

[2] For comparison, mass of the second most massive bottom quark is only 4.2 GeV.

[3] Another possibility is the production of a single top quark via electroweak processes, or combina-

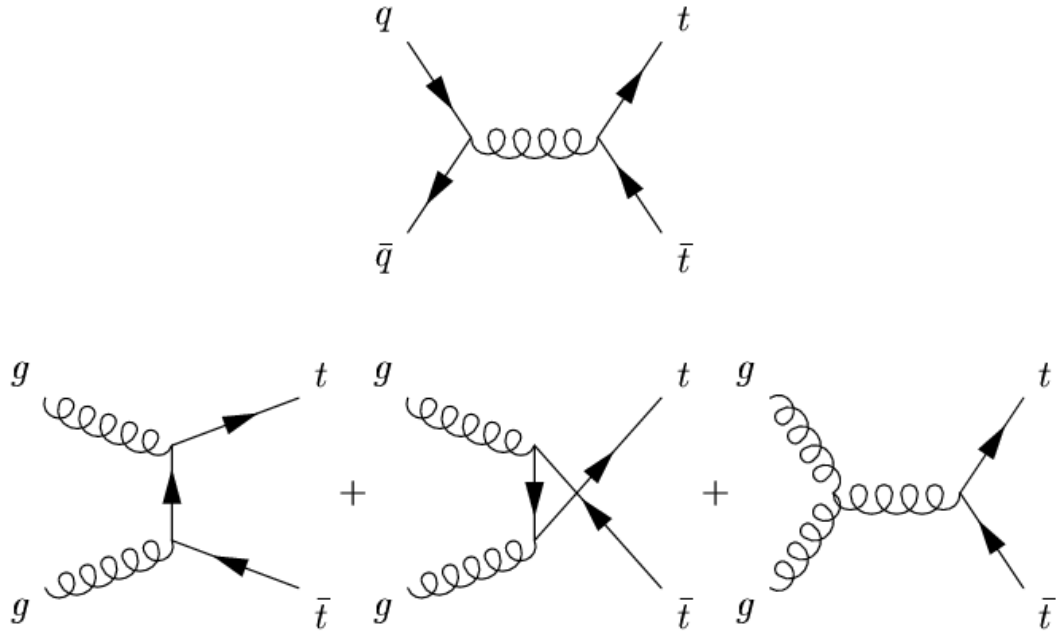


Fig. 2.1: Feynman diagrams representing $t\bar{t}$ pair production via $q\bar{q}$ annihilation (first row) and gluon-gluon fusion (second row).[91]

antiquark annihilation ($q\bar{q} \to t\bar{t}$), or gluon-gluon fusion ($gg \to t\bar{t}$). At the LHC dominant process is the gg fusion ($\approx 90\%$) because for the $q\bar{q}$ annihilation the antiquark from the sea is needed. Moreover, at the energies reached at the LHC, larger part of gluons (when compared to the Tevatron) has sufficient energy to produce $t\bar{t}$ pairs. For example, at the second largest accelerator, Tevatron (already decommissioned), which collided protons with antiprotons, the $q\bar{q}$ annihilation was the main production channel. Feynman diagrams for both processes are depicted in Figure 2.1. At present the $t\bar{t}$ cross section (inclusive and also differential) is known at NNLO (next-to-next-to-leading order) level including the NNLL (next-to-next-to-leading logarithm) soft gluon resummations [84, 85]. In Table 2.1, the total inclusive cross section for $t\bar{t}$ production at the LHC for centre-of-mass energy of 13 TeV, is shown. Values in the table, calculated for various top masses m_t and using PDF (parton distribution function) sets MSTW2008 NNLO 68% CL, CT10 NNLO and NNPDF2.3 NNLO [86, 87, 88, 89], were obtained using TOP++ program [90].

Another way how the top quark can be produced in high-energy collisions is so-called single-top quark production [92]. In this case, top quarks are not produced in the form of $t\bar{t}$ pair, but single particle (either the top or antitop quark) is produced through electroweak interaction. There are three possible electroweak processes which contribute to single-top quark production:

tion of the single and pair production yielding three or four top quarks.

LHC $\sqrt{s} = 13$ TeV				
m_t [GeV]	$\sigma_{t\bar{t}}^{NNLO}$ [pb] central value	δ_{scale}	$\delta_{PDF+\alpha_s}$	δ_{m_t}
172.5	$831.76^{+46.45}_{-50.85}$	$+19.77\ -29.20$	$+35.06\ -35.06$	$+23.18\ -22.45$
173.2	$815.96^{+45.51}_{-49.82}$	$+19.37\ -28.61$	$+34.38\ -34.38$	$+22.67\ -21.95$

Table 2.1: Theory predictions for $t\bar{t}$ production cross section at NNLO in QCD. Values are calculated for different top-quark masses at $\sqrt{s} = 13$ TeV.

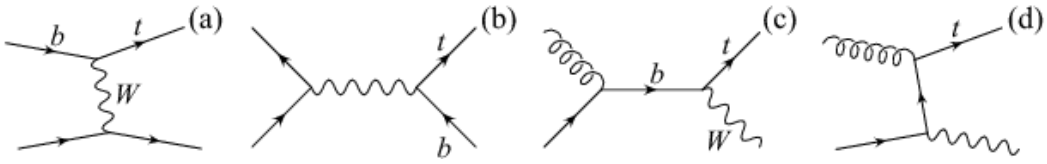


Fig. 2.2: Leading order Feynman diagrams representing single top quark production: t channel (a), s-channel (b) and Wt-channel (c, d) [95].

- t-channel, where top quark is created in the W boson interchange between light (u, d, c, s) and b quark. Shown on the left (a) in Figure 2.2.

- s-channel represents process of quark-antiquark (with different flavor) annihilation mediated by W boson and resulting in t and $\bar{b}$ quarks. Represented by Feynman diagram (b) in Figure 2.2

- Wt-channel, where gluon interacts with b quark from the sea and t together with W boson is produced. Depicted on the rightmost two Feynman diagrams (c, d) in Figure 2.2

Theory predictions for the total production cross section of these processes at $\sqrt{s} = 13$ TeV calculated at NNLO precision using HATHOR package [93, 94] are shown in Table 2.2.

2.2 Top quark decay

Because of its short lifetime, top quark can be detected only through its decay products. According to the SM, the top quark decays via the weak interaction mediated by W boson, resulting in down-type quark and, since W boson is also unstable[4], decay products of W. The flavor of created down-type quark is governed by the elements of the Cabibbo-Kobayashi-Maskawa (CKM) matrix, $|V_{tq}|$. Since $|V_{tb}| = 0.99915 \pm 0.00005$

[4]Decay width of W boson is $\Gamma_W = 2.085 \pm 0.042$ GeV. [96]

LHC $\sqrt{s} = 13$ TeV					
channel	mode	σ^{NNLO} [pb] central value	δ_{scale}	$\delta_{PDF+\alpha_s}$	δ_{total}
t-channel	top	136.02	+4.09 -2.92	+3.52 -3.52	+5.40 -4.57
	antitop	80.95	+2.53 -1.71	+3.18 -3.18	+4.06 -3.61
	top + antitop	216.99	+6.62 -4.64	+6.16 -6.16	+9.04 -7.71
s-channel	top	6.35	+0.18 -0.15	+0.14 -0.14	+0.23 -0.20
	antitop	3.97	+0.11 -0.09	+0.15 -0.15	+0.19 -0.17
	top + antitop	10.32	+0.29 -0.24	+0.27 -0.27	+0.40 -0.36
Wt-channel	top + antitop	71.7	+1.80 -1.80	+3.40 -3.40	-

Table 2.2: Theory predictions for single top quark production cross section at NNLO in QCD. Values are calculated for nominal top-quark mass m_t=172.5 GeV at $\sqrt{s} = 13$ TeV. For Wt-channel top and antitop modes equally contribute to the total cross section.

[96] in the SM, and assuming unitarity of CKM matrix[5], a top quark decays almost exclusively into a b quark and W boson. As mentioned earlier, top quark is produced mainly in form of $t\bar{t}$ pairs, therefore, according to decays of W^+ and W^- bosons, three possible decay channels of the $t\bar{t}$ pair exist:

All hadronic, where both W bosons decays hadronically into a quark and anti-quark (of different flavors). Signature of such events in the detector consists of two b jets and 4 jets originating from light quarks (u, d, c, s). Although this channel allows for full reconstruction of $t\bar{t}$ system (no E_{T}^{miss} - missing transverse energy, signature of neutrinos present in the event), it is contaminated by the large QCD multijet background, which results in low signal-to-background ratio. This mode is responsible for about 45.4% [21] of all $t\bar{t}$ decays. Corresponding Feynman diagram is shown on the left in Figure 2.3.

Lepton+jets (reffered to also as semi-leptonic) channel represents events where one of the W bosons decays hadronically while the second one leptonically into lepton (antilepton) and its corresponding antineutrino (neutrino). This decay thus manifests itself through presence of an energetic isolated charged lepton, alongside with the four energetic jets and missing transverse energy due to escaped neutrino (see the middle diagram in Figure 2.3). Because of only one neutrino in the event, kinematics of the $t\bar{t}$ system can be fully reconstructed[6]. This process accounts for 44.1% [21] of decay rate and represents best balance between a clean event signature and relatively high

[5]Consequently $|V_{td}|$ and $|V_{ts}|$ are of the order of 10^{-3}, and for the purpose of physics analysis can be neglected.

[6]But still there is ambiguity due to longitudinal component of neutrino momentum.

branching ratio.

Dilepton channel offers the best signal-to-background ratio and cleanest signature, because both W boson decay leptonically, producing two energetic leptons, which can be precisely measured. Disadvantage of this channel lies in the presence of two neutrinos in the final state, what makes the full reconstruction of $t\bar{t}$ kinematics very challenging. Moreover, this decay mode suffers from low branching ratio of 10.7% [21]. Corresponding Feyman diagram is shown on the right side in Figure 2.3.

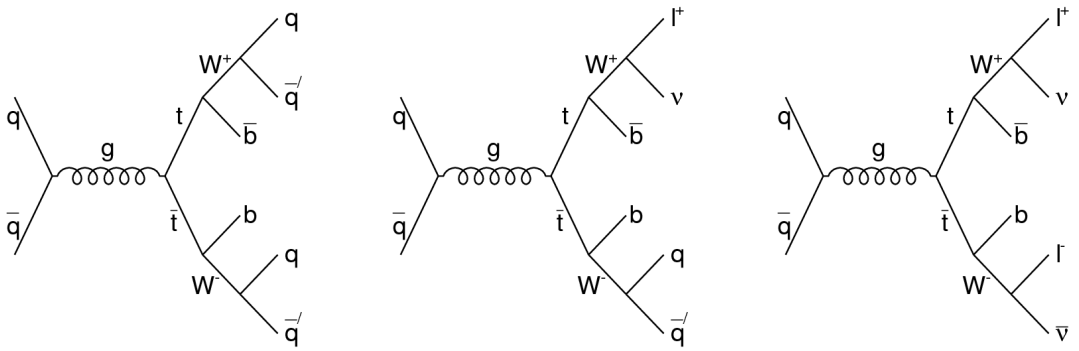


Fig. 2.3: Feynman diagrams representing decay modes of $t\bar{t}$ pair - all-hadronic (left), lepton+jets (middle) and dilepton (right) [97].

2.3 Z boson, its production and decay

The first observation of Z boson, a neutral mediator of weak interaction, occured at UA1 and UA2 experiments at CERN in 1983 [98, 99]. It is the second most massive boson[7], with mass of $m_Z = 91.1876 \pm 0.0021$ GeV [96]. Its large decay width ($\Gamma_Z = 2.4952 \pm 0.0023$ GeV [96]) implies very short mean lifetime ($\tau_Z = 2 \cdot 10^{-25}$ s), what makes this elementary particle highly unstable.

At hadron colliders, a production of the Z boson is dominated by the Drell-Yan mechanism, where a quark and antiquark from colliding protons annihilate producing pair of leptons in the final state (see Figure 2.4, left). This production channel accounts for roughly 65% of all cases, while remaining 35% is covered by quark-gluon interaction, where intermediate fermion radiates Z boson, resulting in 2 leptons and 1 quark in the final state (shown in Figure 2.4, right).

According to the SM, Z boson couples to all elementary particles except of a gluon and photon. For this reason, Z boson can decay via many decay channels. The following decay modes belong among most probable and important from physics point of view[8]:

[7] After Higgs boson with a mass of $m_H = 125.10 \pm 0.14$ GeV [96].

[8] Other decay modes are allowed only at the next-to-leading order (NLO), and thus are greatly

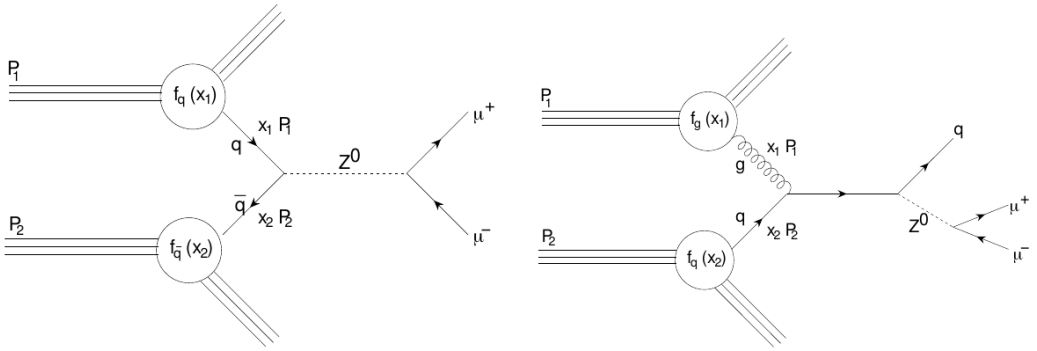


Fig. 2.4: Leading order (left) and next-to-leading order (right) Feynman diagrams representing Drell-Yan process of Z boson production [100].

All-hadronic channel, where Z boson decays into hadrons, accounts for 69.911 ± 0.05 % [96]. The signature of this final state in detector consists of two jets originating from q and $\bar{q}$. However, reconstruction of the two-jet system is very difficult since the calorimeter resolution for dijet pair is low.

Invisible channel with probability of 20.000 ± 0.055 % [96] represents process in which Z boson decays into neutrino and its corresponding antineutrino. Identification of Z from this decay mode is extremely difficult since the only indication is in a form of E_{T}^{miss}, which is present also for other processes.

Leptonic channel is characterized by two leptons with opposite charge in the final state. Moreover, these leptons have same flavor and their invariant mass is close to the nominal mass of Z boson ($m_Z^{nom} = 91.1876$ GeV). Possible final state compositions are then: e^+e^- ($3.3632 \pm 0.0042\%$), $\mu^+\mu^-$ ($3.3662 \pm 0.0066\%$) and $\tau^+\tau^-$ ($3.3696 \pm 0.0083\%$) [96]. As can be seen, contribution for each lepton flavor is the same, due to the principle of *lepton universality*. τ leptons are unstable and can further decay into either hadrons, or less massive leptons, even before reaching the detector. This makes identification of Z boson from decay to $\tau^+\tau^-$ a challenging problem and therefore these decays are not included in this analysis.

2.4 Associated production of the $t\bar{t}$ pair and Z boson

Processes in which $t\bar{t}$ pair is produced together with the Z boson are very rare in the SM. Allowed production mechanisms at leading order include either initial state radiation (ISR), where the Z boson is radiated from one of the initial quarks (in case of $q\bar{q}$ annihilation), or final state radiation (FSR), in which one of the top quarks radiates the Z boson. The Feynman diagrams for these processes are shown in Figure 2.5. Events with $t\bar{t}Z$ signature are predominantly created via gg fusion ($\approx 70\%$), the rest

suppressed [96].

via $q\bar{q}$ annihilation.

According to the SM, in case of Z originating from the FSR there is a possibility to directly probe the coupling of Z boson to top quark. This can be achieved through measurement of $t\bar{t}Z$ production cross section, since this depends quadraticaly on the $t\bar{t}Z$ weak vertex factor containing t to Z coupling constant. Thus, the SM prediction can be either verified or excluded via experimental cross section measurement.

The most general form of the $t\bar{t}Z$ vertex function, which is Lorentz-invariant, can be written in terms of ten form factors. However, under the assumption of coupling between Z and effectively massless fermions, and considering only on-shell top quarks, the number of form factors is reduced to four. The vertex function [101] then takes on the following form:

$$\Gamma^{\mu}_{t\bar{t}Z}(k^2, q, \bar{q}) = -ie\left\{\gamma^{\mu}(F^Z_{1V}(k^2) + \gamma^5 F^Z_{1A}(k^2)) + \frac{\sigma^{\mu\nu}}{2m_t}(q+\bar{q})_{\nu}(iF^Z_{2V}(k^2) + \gamma^5 F^Z_{2A}(k^2))\right\}, \tag{2.1}$$

where e denotes elementary charge, m_t stands for the top quark mass, q and $\bar{q}$ corresponds to outgoing top and antitop quark four-momentum, respectively, $k^2 = (q+\bar{q})^2$, $\sigma^{\mu\nu} = \frac{i}{2}(\gamma^{\mu}\gamma^{\nu} - \gamma^{\nu}\gamma^{\mu})$ and γ^5 is product of Dirac γ matrices. When considering low energy limit, the coefficients $F^Z_{1V}(0)$ and $F^Z_{1A}(0)$ correspond to vector and axial vector $t\bar{t}Z$ form factors, respectively. The other two form factors $F^Z_{2V}(m_Z^2)$ and $F^Z_{2A}(m_Z^2)$, where m_Z represents the Z boson mass, are related to so-called weak magnetic and electric dipole moments and are equal to zero at tree level in the SM. When considering only tree level contributions in the SM, the first two form factors are defined as follows:

$$\begin{aligned} F^{Z,SM}_{1V} &= -\frac{1}{4\sin\theta_W\cos\theta_W}\left(1 - \frac{8}{3}\sin^2\theta_W\right), \\ F^{Z,SM}_{1A} &= \frac{1}{4\sin\theta_W\cos\theta_W}, \end{aligned} \tag{2.2}$$

where θ_W denotes weak mixing angle (see Section 1.4.4).

As a consequence, the hints of new physics can manifest themselves through non-zero values of form factors F^Z_{2V} and F^Z_{2A}, what would indicate deviations from the SM predictions. Moreover, new physics can modify also SM values of F^Z_{1V} and F^Z_{1A}. All these deviations can be effectively studied by measuring distributions of differential cross sections. As can be seen from Figure 2.6, which shows comparison of differential cross section distributions for different values of the mentioned form factors, variables sensitive to the coupling between t quark and Z offer possibility to either verify or exclude predictions of the SM. The measurement of total inclusive and differential cross sections of $t\bar{t}Z$ production is the main motivation and focus of this thesis.

The most accurate theoretical prediction for the total cross section of the $t\bar{t}Z$ production for pp collisions at $\sqrt{s} = 13$ TeV, calculated at NLO [6] is

$$\sigma_{t\bar{t}Z} = 0.863^{+8.5\%}_{-9.9\%}\,(\text{scale}) \pm 3.2\%\,(\text{PDF} + \alpha_{\text{S}})\,\text{pb},$$

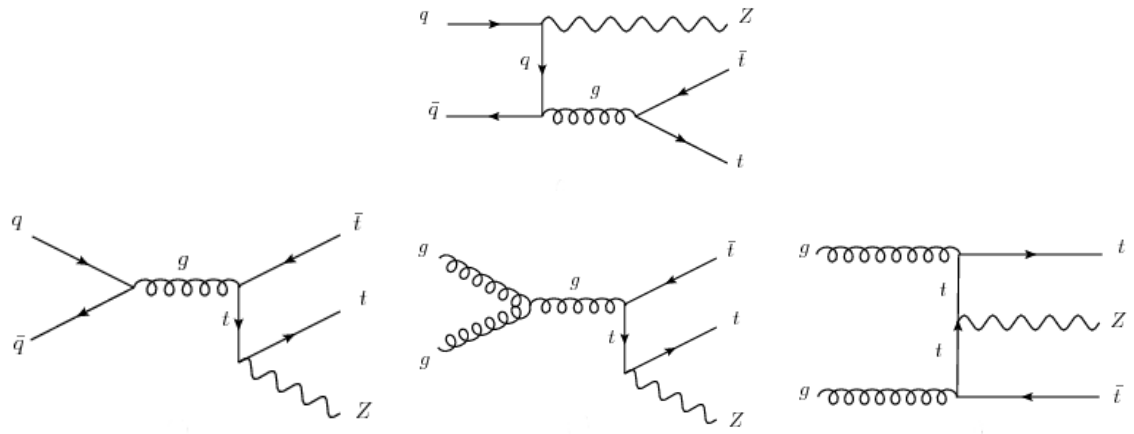


Fig. 2.5: Leading order Feynman diagrams for $t\bar{t}Z$ production via ISR (first row) and FSR (second row) [102].

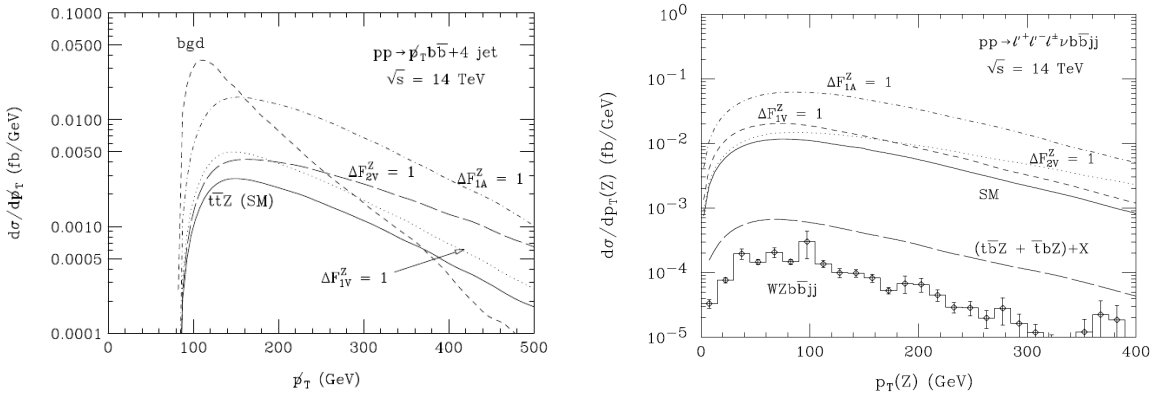


Fig. 2.6: Distributions of the $t\bar{t}Z$ differential cross section as a function of missing transverse momentum $\not{p}_{\mathrm{T}}$ in the decay channel where Z decays into neutrinos and $t\bar{t}$ pair hadronically (left), and as a function of Z transverse momentum p_{T}^{Z} considering Z decay into charged leptons and lepton+jets decay of $t\bar{t}$ pair (right). In the right plot, single top quark background, $(t\bar{b}Z + \bar{t}bZ) + X$, is represented by the dashed line as predicted by the SM, while the non-resonant $WZb\bar{b}jj$ background is obtained from the Monte Carlo simulation and is depicted as histogram, with error bars corresponding to the statistical uncertainty of the simulation. Solid line corresponds to the SM predictions and other lines depict distributions for non-standard $t\bar{t}Z$ couplings, while only one form factor at a time is allowed to deviate from its SM values. Figure on the left is taken from [101] and on the right from [103].

what is ≈ 1000 times less than the cross section of the $t\bar{t}$ pair production.

Both the top quark and the Z boson are unstable particles with a very short mean lifetime. For this reason, the only way how to study them is through the detection of their decay products. There exist various $t\bar{t}Z$ decay channels according to decay modes of individual particles. Since invisible decay of Z boson (to $\nu\bar{\nu}$ pair) is extremely difficult to identify and also hadronic decay has worse momentum and energy resolution compared to leptons, the focus of this analysis is only on channels where Z decays leptonically into two charged leptons (excluding $\tau^+\tau^-$ decays as mentioned in the previous section). Therefore, the following channels are considered for the analyses described in this thesis:

Dilepton channel, where $t\bar{t}$ pair decays hadronically and Z boson decays into lepton pair with opposite charge and same flavor. Although this channel has the highest probability among the studied channels (with the branching ratio of about 3.1%), its disadvantage is high background contamination due to presence of jets. The final state signature then consists of two leptons with an opposite sign, same flavor, and their invariant mass should be close to nominal mass of a Z boson (reffered to as Z-like pair), two b-jets from $t\bar{t}$ decay and four additional jets from decay of W^+ and W^-. Major backgrounds, which resemble dilepton topology, are dilepton decays of $t\bar{t}$ pair and Z associated with jets. Dilepton decay channel is used in measurement of inclusive $t\bar{t}Z$ cross section discussed in this thesis.

Trilepton channel is characterized by presence of three charged leptons in the final state (two oppositely charged leptons from Z boson (e^+e^- or $\mu^+\mu^-$) and one lepton from semileptonic decay of $t\bar{t}$ pair). These are accompanied by two b-jets and two additional light jets from hadronic decay of W boson. One neutrino, registered in form of E_{T}^{miss}, emerges from semileptonic $t\bar{t}$ decay. This channel provides most of the observed data events (however its branching ratio of 2.3% is lower than for dilepton channel), while its background is reasonably small. Dominant backgrounds mimicking signature of this channel are tZ, tWZ and dibosons (WZ, ZZ) [104]. This channel is used in differential cross section measurement presented in this dissertation.

Tetralepton channel includes events where both Z boson and $t\bar{t}$ pair decay leptonically, producing four leptons in the final state. The lepton pair from $t\bar{t}$ is required to be of an opposite sign, while the one from the Z boson must meet the Z-like pair requirements, described above. Two b-jets and two neutrinos arise from $t\bar{t}$ decay. Although this channel has very low background contamination (dominated by fake leptons[9], tWZ and ZZ), it is currently statistically limited due to its low branching ratio of only 0.43%. Tetralepton channel is used for the differential cross section measurement, in the same analysis as the measurement in the trilepton channel.

[9]These are either leptons from decay of b-hadrons, c-hadrons and photon conversions, or jets that were misclassified as electrons.

Chapter 3

The LHC and ATLAS

3.1 Acceleration process at CERN

European Organisation for Nuclear Research (Conseil européen pour la recherche nucléaire - CERN) is an organisation for basic and applied research mainly in the field of particle physics. It is located in Meyrin on the Franco–Swiss border near Geneva. CERN was founded in 1954 and today it affiliates 23 member states, including Slovakia. The accelerator complex at CERN is a succession of machines which serve for increasing energy of the accelerated particles. After the acceleration process, each machine injects the beam[1] into the next one constructed to bring it up to higher energy. The last element in this chain is the Large Hadron Collider (LHC) [105, 106], currently the most powerful accelerator in the world. The majority of pre-accelerators have also its own detectors for experiments at lower energies. Protons, which are subsequently accelerated, are obtained by stripping electrons from hydrogen atoms. Before injection of beam into the LHC, the beam particles are pre-accelerated in 4 stages as is shown in Figure 3.1. Following the light-grey arrow, beams are accelerated first in linear accelerator (LINAC 2 or LINAC 3 in case of ion beams), then passed to either BOOSTER (synchrotron for proton beams) or LEIR (Low Energy Ion Ring for creating short dense bunches from ion beams) and then injected to series of circular synchrotrons (PS - proton synchrotron, and SPS - super proton synchrotron). The acceleration to maximum energy of 6.5 TeV per beam takes about 20 minutes and beams are able to circulate in LHC pipes for many hours.

In addition to protons, the accelerating complex is designed to accelerate also heavy ions (i.e. lead or xenon) [107]. These are produced from highly purified lead sample heated to about 800 °C. Then the lead vapour is ionized by an electric current resulting in many different charge states with majority around Pb^{29+}. Selected ions are pre-accelerated as shown in Figure 3.1 (dark-grey arrow) and after next stripping to Pb^{82+},

[1] Beam refers to focused stream of the accelerated particles concentrated in bunches.

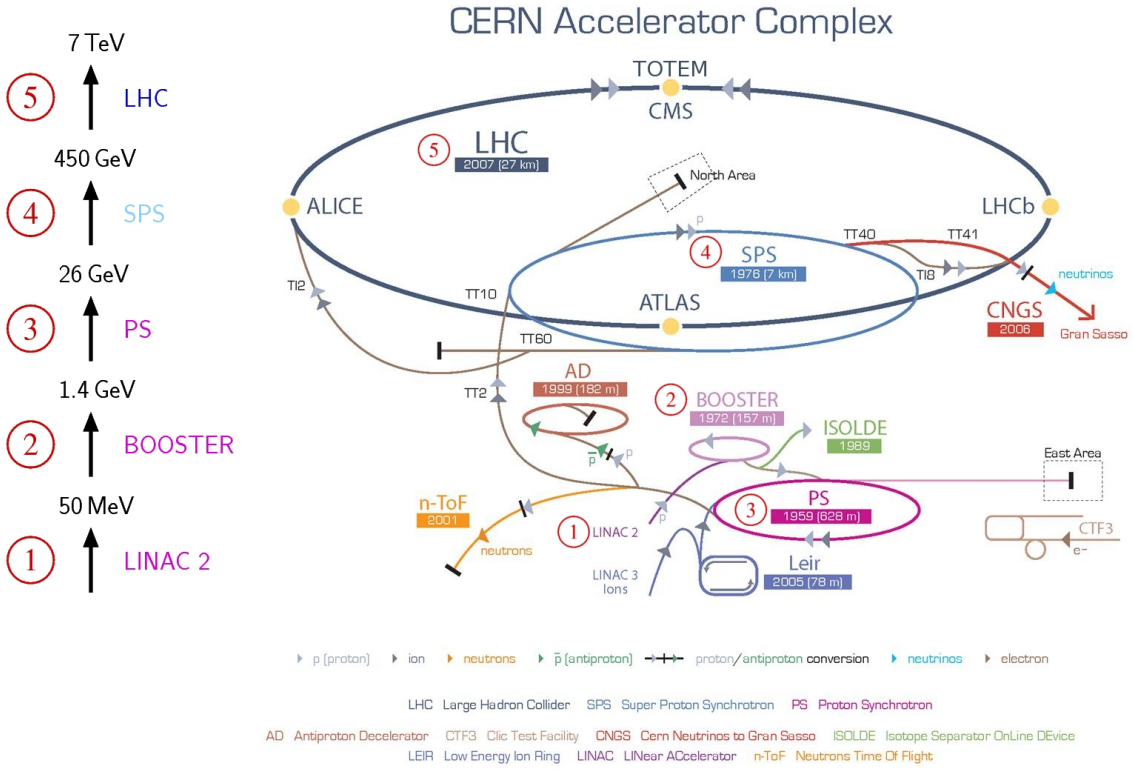


Fig. 3.1: Scheme of accelerating complex at CERN [108].

they are transferred to the LHC. In the final step, the LHC accelerates the ions to maximum energy of 2.56 TeV/u.

3.2 The Large Hadron Collider

The LHC is a two-ring hadron accelerator and collider built for CERN. It was installed in the existing 26.7 km tunnel which originally served as accelerating tunnel for the Large Electron Positron (LEP) collider. The tunnel was constructed between years 1984 and 1989 and consists of eight straight sections and eight arcs which lies between 45 m and 170 m below the earth surface. Four main experiments at the LHC use its detectors for analysing products of particle collisions. Among the largest experiments belong ATLAS (A Toroidal LHC ApparatuS) [109] and CMS (Compact Muon Solenoid) [110], which use multipurpose detectors to study particle physics to greatest possible extent.

The ultimate goal of the LHC is to test the SM and search for evidence of physics beyond the SM. The circular design of the accelerator, where the two proton beams (going in opposite direction) are collided, is much more effective compared to experi-

ments with a stationary target. This makes it possible to study events[2] at maximum construction centre-of-mass energy of 14 TeV. Following topics belong among the most important physics areas which are studied at the LHC:

- In 1964 R. Brout, F. Englert and P. Higgs published a theory [49, 50, 111] that explained origin of mass of elementary particles through their interaction with Higgs field. In mid-1970s, physicists started to search for evidence of Higgs boson, the particle connected with the Higgs field. In July 2012 these searches resulted in a discovery of a particle with properties of Higgs boson. This remains one of the greatest successes of the LHC experiments, but also initiates new research ideas associated with Higgs physics, like detail studies of its properties or searching for its other decays.

- As the SM does not consider theory of gravity as one of the fundamental forces in nature, new theories which offer unification of all forces, like i.e. SUSY, are studied in great extent at the LHC. Existence of least massive supersymmetry partners of known elementary particles could potentially be confirmed when going for maximum energy.

- Cosmological and astronomical observations show that only ≈ 5 % of universe is composed of visible matter. Therefore one of the main focuses of the LHC is to search for a particle or phenomena responsible for a dark matter particles and the dark energy. One of the hypotheses predicts the dark matter to consist of neutral SUSY particles which could be found by the LHC experiments.

- Clarify the cause of matter-antimatter imbalance established in the early universe.

- According to the current theories, after the Big Bang the universe went through a stage of so called quark-gluon plasma (QGP), extremely hot and dense bulk of matter consisting of free quarks. Experiments with ion beams aims for studying QGP and thus understanding the processes in early universe, especially the mechanism of quark confinement.

- Study of compositeness of the SM fundamental particles, especially Higgs boson. The LHC is capable of probing compositeness at a level better than 10^{-21} m.

In the LHC, two proton beams (each beam consisting of 2808 bunches with 1.2×10^{11} protons) collide with frequency of 40 MHz. Crucial variable describing particle accelerator is instantaneous luminosity, which describes ratio of number of events detected in certain time to total interaction cross section. It is thus defined as follows:

$$\mathcal{L} = \frac{\dot{N}}{\sigma A \varepsilon}, [\mathcal{L}] = \frac{1}{\mathrm{cm}^2\mathrm{s}}, \tag{3.1}$$

[2]Event refers to one bunch crossing in detector.

where $\dot{N} = dN/dt$, N is the number of events detected in time t with the corresponding cross section σ, A is the experimental acceptance and ε stands for efficiency of measurement. The expected total number of events is then calculated as product of integrated luminosity $\mathcal{L}_{int} = \int \mathcal{L}dt$ and the total cross section. The design luminosity of proton-proton $(p-p)$ collisions at $\sqrt{s} = 14$ TeV is 10^{34} cm^{-2}s^{-1}. As already mentioned, the LHC was constructed also for accelerating heavy lead ions up to 5.5 TeV for nucleus pair, while the luminosity for ion beams is of the order of 10^{27} cm^{-2}s^{-1}. The main reason for accelerating only protons and ions at the LHC (both are hadrons as indicated in accelerator name) is that the accelerating process, which uses electromagnetic field, requires non-zero electric charge of accelerated particles. Moreover, the particles need to be stable, what constraints the number of suitable particles only to electron, proton and ions (and their anti-particles). Heavy particles, like ions and protons, are more convenient for accelerating to high energies, because they are decelerated less by synchrotron radiation [112] than light particles like electrons. Energy losses in one circulation caused by synchrotron radiation are defined by the following formula:

$$P = \left[\frac{e^4}{6\pi\varepsilon_0 m_0^2 c}\right]\gamma^2 B_\perp^2 \tag{3.2}$$

where e is the elementary charge, ε_0 is the permitivity of vacuum, m_0 is the mass of the accelerated particle, c is the speed of light in vacuum, $B_\perp$ is the component of the magnetic induction perpendicular to the particle's speed vector and $\gamma = E/(m_0c^2)$, where E is the particle's energy. Since energy loss is inversely proportional to squared mass of the particle, electron loses its energy 10^{13} faster than proton.

Two experiments at the LHC, ATLAS and CMS, use full luminosity ($\mathcal{L} \approx 10^{34}$ cm^{-2}s^{-1}), while the other two, LHCb [113], which is focused on study of matter-antimatter imbalance, and TOTEM [114], which detects protons after elastic scattering to small angles, are designed for lower luminosity, $\mathcal{L} \approx 10^{32}$ cm^{-2}s^{-1} and $\mathcal{L} \approx 10^{29}$ cm^{-2}s^{-1}, respectively. Even lower luminosity of $\mathcal{L} \approx 10^{27}$ cm^{-2}s^{-1} is used by the ALICE [107] experiment studying QGP in heavy ion collisions.

3.3 Design of the LHC

The basic building blocks of the LHC are a vacuum tube, where the bunches of particles circulate, dipole magnets, which are used for keeping particles on circular orbits while they are accelerated, quadrupole magnets, which role is to focus the beam to the center of the tube, and electromagnetic resonator that accelerate particles and keep them at a constant energy when acceleration is finished.

The LHC **vacuum system** consists of three parts - isolation vacuum for cryomagnets (largest part of the whole vacuum system ≈ 9000 m^3), beam vacuum to eliminate

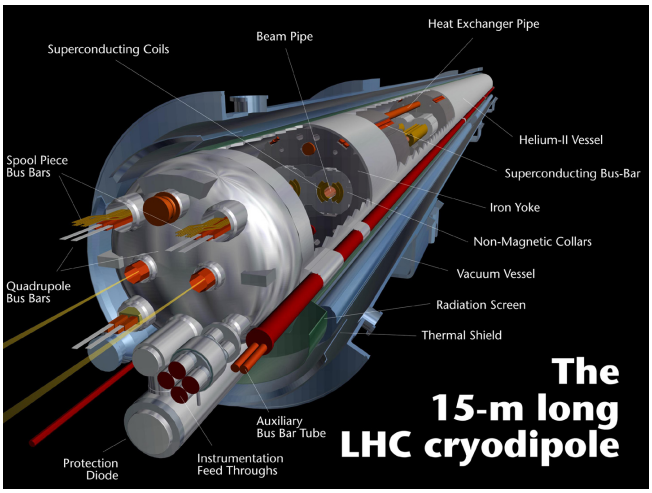


Fig. 3.2: Cryodipole magnet at the LHC [115].

accidental collisions of particles with gas molecules (ultrahigh vacuum corresponding to pressure of 10^{-8} Pa) and isolation vacuum for the helium distribution system.

Magnetic system consists of a large variety of magnets, including already mentioned dipoles, quadrupoles, but also sextupoles, octupoles etc., summing up to total of 9600 pieces. Every magnet in the system is designed to optimise a particle's trajectory. Special insertion quadrupoles are used to squeeze the beam to the smallest possible size in the collision points to increase probability of particles collision. One element of the cryodipole magnetic system and its components is shown in Figure 3.2.

The main role of **electromagnetic cavities** in the LHC is to keep protons tightly bunched to ensure high luminosity, and to deliver radiofrequency power to the beam during the process of acceleration. There are eight superconducting cavities per beam. These are grouped in 2 cryomodules which use superfluid helium to cool them down to operating temperature of 4.5 K (similar to magnets, which operate at 1.9 K).

3.4 The ATLAS detector

ATLAS is a multipurpose detector with forward-backward symmetric (w.r.t. interaction point) and cylindrical geometry. It is one of the largest man-built detectors in the world - its height is 25 m, length 44 m and weights $\approx$ 7000 t. The detector consists of these four cylindrical sub-systems arranged in concentric layers around the interaction point:

- Inner detector (ID), placed in 2 T solenoidal field, is used for precise tracking of electrically-charged particles and measures their charge, direction, momentum and origin [3] (can be from either primary or secondary interaction). Shown in yellow in Figure 3.3.

[3] Origin in this context refers to the underlying interaction from which the particle arises. Hard

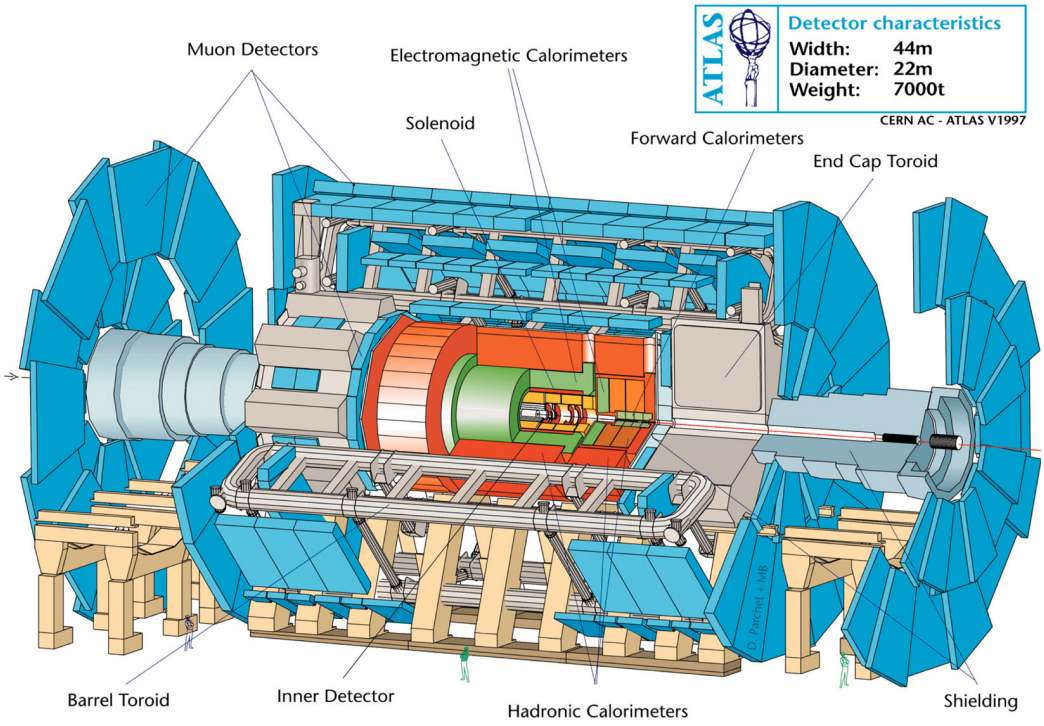


Fig. 3.3: The longitudinal view of ATLAS detector with its sub-systems [116].

- Electromagnetic and hadronic calorimeters absorb and measure energy of electrons, photons and hadrons. This is also crucial for evaluating missing transverse energy, which can be assigned to neutrinos. Depicted in orange and green in Figure 3.3, respectively.

- Muon spectrometer (MS) which identifies muons and measures their momenta. This largest part of ATLAS was built due to very weak interaction of muons with detecting medium of calorimeters. Operation of this spectrometer is independent of the ID system. The MS is contained mostly in the outer magnetic system shown in blue in Figure 3.3.

- Magnetic system, which ensures curving of trajectories of charged particles and thus enables proper measurement of their momenta. Grey color in Figure 3.3.

Particular sub-systems of the detector are described in detail in following sections.

3.4.1 Coordinate system of the ATLAS detector

To be able to consistently describe event topology, the origin of the coordinate system is placed in the interaction point inside ATLAS detector. Beam direction defines z axis, while $x - y$ plane is transverse to the beam direction. The positive x axis points

scatter process represent primary interaction while the subsequent interactions (characterized by the corresponding vertices) are called secondary.

to the center of the LHC ring and positive y axis is defined as pointing upward, making coordinate system right-handed. The azimuthal angle Φ is defined in $x-y$ plane around the beam axis, while polar angle θ defines deflection from the beam axis. Every particle is fully described by its four-momentum $P = (E, p_x, p_y, p_z)$, where E is particle's energy and $\mathbf{p} = (p_x, p_y, p_z)$ is its momentum vector. Transverse momentum of the particle is then defined as:

$$p_T = \sqrt{p_x^2 + p_y^2} = p \sin\theta. \tag{3.3}$$

An important variable used in experimental particle physics is missing transverse energy $E_T^{miss} = E^{miss} \sin\theta$, which defines the total transverse energy corresponding to neutrino(s) in the event and is calculated from the conservation four-momentum. The variable which quantify particle's deflection from beam axis is called rapidity, and is defined as follows:

$$y \equiv \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right), \tag{3.4}$$

Rapidity difference is invariant under Lorentz boost along beam axis and therefore it is widely used in particle physics. In relativistic cases, when $|\mathbf{p}| \approx E$, or in case of massless particles, the above equation becomes:

$$\eta \equiv -\ln\tan\left(\frac{\theta}{2}\right) = \frac{1}{2} \ln\left(\frac{|\mathbf{p}| + p_z}{|\mathbf{p}| - p_z}\right). \tag{3.5}$$

and variable η is called pseudorapidity. The value of η for $\theta = 90°$ is 0, while for $\theta = 0°$ it is ∞ (i.e. $\eta = 2.44$ for $\theta = 10°$). To quantify separation of two objects in the detector, it is common to use their angular difference defined as follows:

$$\Delta R = \sqrt{\Delta\Phi^2 + \Delta\eta^2}, \tag{3.6}$$

where $\Delta\Phi = \Phi_1 - \Phi_2$ and analogously $\Delta\eta = \eta_1 - \eta_2$.

3.4.2 The Inner Detector

Every 25 ns (for Run 2 (2015-2018), for Run 1 (2009-2013) it was every 50 ns) approximately 1000 particles emerge from the interaction point producing high track density in the detector. Fine granularity of the detector, needed for very precise momentum measurements required by studied physics models, is achieved by using pixel and silicon microstrip trackers (SCT) in conjunction with the straw tubes of the Transition Radiation Tracker (TRT). These are concentrated closest to the beam pipe and cover the region with $|\eta| < 2.5$. Pixel and SCT detectors are arranged on concentric cylinders in barrel region[4] and on disks perpendicular to the beam axis in the end-cap region[5]

[4]Barrel region refers to the curved surface of the cylinder placed along beam direction with center in the interaction point.

[5]End-cap region refers to the base of the cylinder placed along beam direction with center in the interaction point.

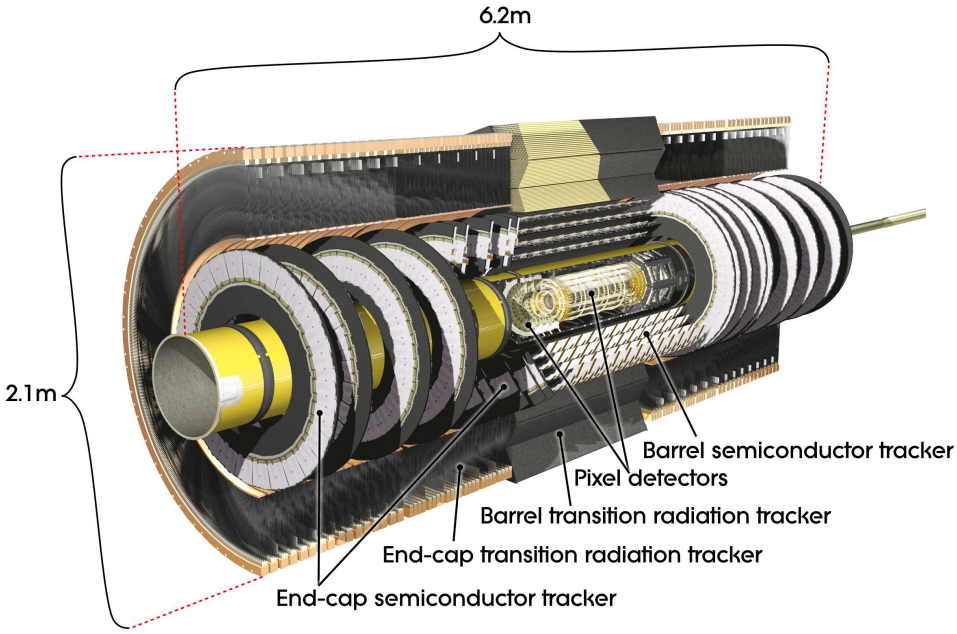


Fig. 3.4: The longitudinal view of the Inner Detector [109].

(see Figure 3.4 for details). The working principle of pixel and microstrip detectors is similar, and is based on freeing electrons from silicon atoms due to interaction with passing particles. These electrons are then collected by readout electronics, producing measurable electric signal.

The Pixel Detector [117] consists of 80.4 million pixels covering total area of 1.7 m^2. Its spatial resolution in $R-\Phi \times z$ plane is 10×115 μm^2. In the upgrade after Run 1, the pixel detector was improved by adding silicon layer to the innermost part of the detector. This additional layer is called Insertable B-Layer (IBL) [118] and it is the first detection device that particles pass into after the collision.

For **the Silicon Microstrip Tracker** [119], eight strip layers are crossed by each track. Tracking in the barrel region ($|\eta| < 1.4$) is provided by small-angle (40 mrad) stereo strips, with additional strips parallel to the beam direction for measuring $R-\Phi$. The same small-angle stereo strips, together with radially aligned strips, are used also in the end-cap region ($1.4 < |\eta| < 2.5$). The total number of channels in SCT is 6.3 million and its spatial resolution is 17×580 μm^2.

The Transition Radiation Tracker [120] operates with 4 mm diameter straw tubes, which provide in total 36 hits per track and cover region up to $|\eta| = 2.0$. Similar to SCT, the straw tubes, filled by xenon-based gas mixture, are arranged in parallel to the beam axis in barrel region (their length is 144 cm and cover $|\eta| < 1.0$), while in the end-cap they are placed radially in wheels (length in end-cap is 37 cm and they cover $1.0 < |\eta| < 2.0$). In total 351,000 readout channels offers intrinsic accuracy of 130 μm per straw.

3.4.3 Calorimetry

In particle physics, energy of a particle is reconstructed usually from its decay products. When these decay products enter particular detecting system (sampling calorimeter in this case), they interact with its medium [6] (active and passive) producing so-called parton shower. Subsequent products of this process are then registered in detecting (active) medium of the calorimeter. Parton shower is usually described by interaction (radiation) length. This quantity evaluates characteristics of the material related to the energy losses of particle passing through it. One radiation length represents the distance traveled by high-energy electron before it decreases its energy $1/e$ times due to brehmstrahlung. Analogously, interaction length is defined for hadronic particle and its inelastic interaction with nuclei of the material. The ATLAS detector posses electromagnetic, as well as hadronic calorimeter, which cover region with $|\eta| < 4.9$ and more than 22 radiation lengths in case of electromagnetic calorimeter (for $\eta = 0$). One of the important characteristics of the calorimeter is its energy resolution [121], which can be defined as follows:

$$\frac{\sigma_E}{E} = \frac{A}{\sqrt{E}} \oplus B \oplus \frac{C}{E} \tag{3.7}$$

where E is energy of the particle in GeV and A, B and C are constants. First term on the right-hand side is so-called stochastic term and corresponds to fluctuations of energy deposited in active medium and its conversion to signal. Second, constant term, represents inhomogenity of the calorimeter and third term, called electronic, describes fluctuations caused by electric noise. Typical values for A are in range 0.5-1.0, for B between 0.03 and 0.05 and for C it is few percent. According to the above equation, the more energetic particles are detected, the better is the resolution.

Electromagnetic calorimeter uses layers of lead plates and chambers with liquid argon (LAr) [122] to absorb and measure electromagnetic showers induced by photons, electrons and positrons. The coverage of the electromagnetic calorimeter in terms of pseudorapidity is $|\eta| < 1.475$ in the barrel region, while the end-cap part covers region with $1.375 < |\eta| < 3.2$. Shower is produced in interaction of particle with lead and thus evolves mainly in lead plates. LAr chambers are used as detection medium, where shower particles ionize molecules of argon. Emitted electrons drift towards charged accordion-shaped electrodes, where they are collected producing electric signal. Intensity of the shower (electric signal) is then proportional to the energy of the incident particle from a hard scatter. To distinguish photons from electrons, the signal from the shower is paired with tracks from ID, where only charged electron leaves signal. Electric neutral photon, beside that it does not produce signal in the tracker, undergo different interaction with calorimeter compared to electron. Photon typically travel some

[6] Medium of the calorimeter is classified as either passive, which absorbs the particle energy and active, which is used for detection and evaluation of particles energy.

distance in the calorimeter before it is converted to electron-positron pair, while electron emits brehmstrahlung immediately after it enters the calorimeter. However, this difference is eliminated because both particles, before they enter calorimeter system, propagate through ID and magnetic system (where photon can already be converted to e^+e^- pair), what makes calorimeter impossible to distinguish between them. For this reason, ATLAS uses so-called pre-shower detector, which identifies particles before they enter calorimeter.

Hadronic calorimeters [123] form more complicated system than electromagnetic one due to far more complex hadron showers. Soft QCD interactions (with lower momentum transfer) between hadrons and nucleons are extremely difficult to model, what makes simulations of hadronic calorimeters less precise. ATLAS hadronic calorimeter system consists of three main parts (also depicted in Figure 3.5): Tile Calorimeter, LAr Hadronic End-cap Calorimeter (HEC) and LAr Forward Calorimeter (FCal). Main components of Tile Calorimeter, which is located right after electromagnetic calorimeter in the barrel region, are iron plates as a passive medium and scintillating tiles, which are used as an active medium. It is divided into three parts. First, barrel part, covers the region with $|\eta| < 1.0$ and two extended barrels on both side cover $0.8 < |\eta| < 1.7$. Hadrons, which are not fully absorbed by electromagnetic calorimeter, produce hadronic showers when interacting with iron plates. These consist mainly from pions and kaons, but also neutrons, low-energy protons and other hadrons. Electric signal is produced when shower particles excite atoms of scintillator, which subsequently emit photons which are then collected by photomultipliers. Again the intensity of the measured electric signal is proportional to energy of particle which iniciated the shower. Hadronic End-cap Calorimeter is placed directly after the end-cap part of electromagnetic calorimeter and thus covers region with $1.5 < |\eta| < 3.2$. It is made of two independent cylinders, each consisting of alternating copper plates (passive medium) and liquid argon (active medium). To extend calorimetry system as much as possible, Forward Calorimeter was designed not only to measure energy of hadrons in region $3.1 < |\eta| < 4.9$, but also to serve for reduction of background radiation in muon spectrometer. The active medium is the same as in HEC, while passive medium consists of three layers - first is from copper (suitable for electromagnetic calorimetry) and next two are from tungsten (for hadronic calorimetry).

3.4.4 Muon Spectrometer

Muons with transverse momentum greater than few hundred MeV produced inside the ATLAS detector pass through whole detecting medium. For their better identification and to precisely measure their momenta, the muon spectrometer [125] was built around calorimetry system. It uses strong magnetic field to curve trajectories

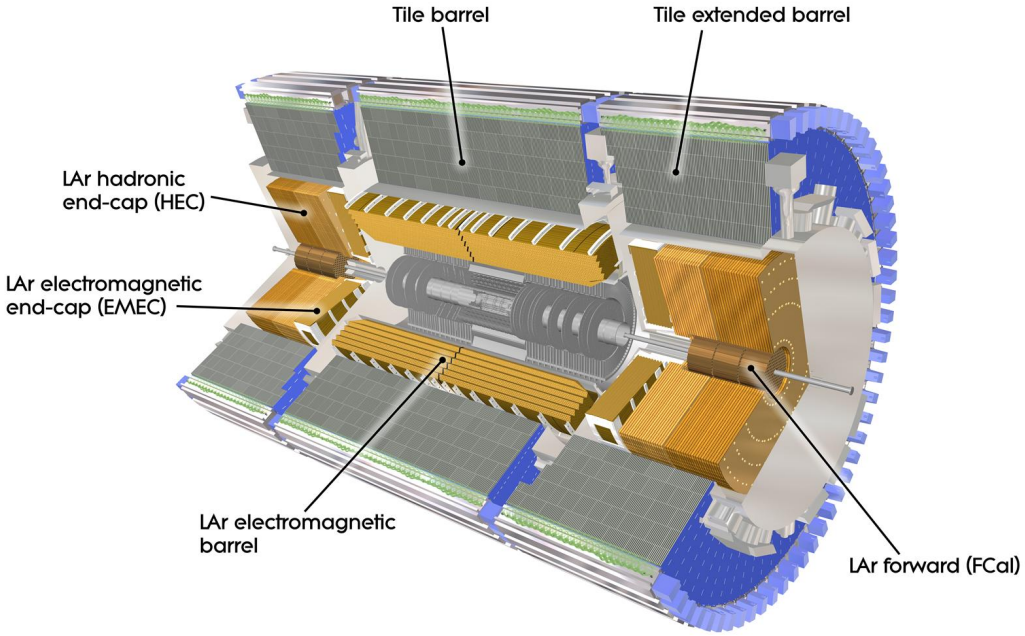


Fig. 3.5: Calorimetry system at the ATLAS detector [124].

of charged muons, what enables calculation of their transverse momenta. The performance of magnetic system can be described by its ability to deflect particle's trajectory and can be quantified by $\int Bdl$, where B is component of magnetic field perpendicular to muon's trajectory and integration goes over muon's path in the chamber. The ATLAS muon spectrometer consists from three magnets: large barrel toroid and two toroids in end-cap regions. All three magnets are superconducting and orientation of their magnetic field is orthogonal to direction of incoming muon as well as solenoid field in the ID. Precise measurements of muon trajectories is provided by several detection sensors (see Figure 3.6):

Monitored Drift Tubes (MDT) are used for precision measurement of track coordinates over range $|\eta| < 2$. They are aligned in groups of 3 to 8 into chambers. Each sense wire in the chamber is isolated from neighboring one, ensuring reliable operation. The resolution of individual tubes is 80 μm.

Cathode Strip Chambers (CSC) replace MDT's in region $2 < |\eta| < 2.7$. These multiwire proportional chambers are designed to withstand demanding rate in this region. Cathodes are segmented into strips, what enables measurement of coordinates in both bending (resolution is 40 μm) and transverse (resolution is 5 mm) plane.

The muon spectrometer has its own triggering system which provides bunch crossing identification, precise measurement of track coordinates and well defined p_{T} thresholds. This trigger system covers pseudorapidity range $|\eta| < 2.4$ and consists of **Resistive Plate Chambers (RPC)**, gas filled detectors triggered by ionisation of passing par-

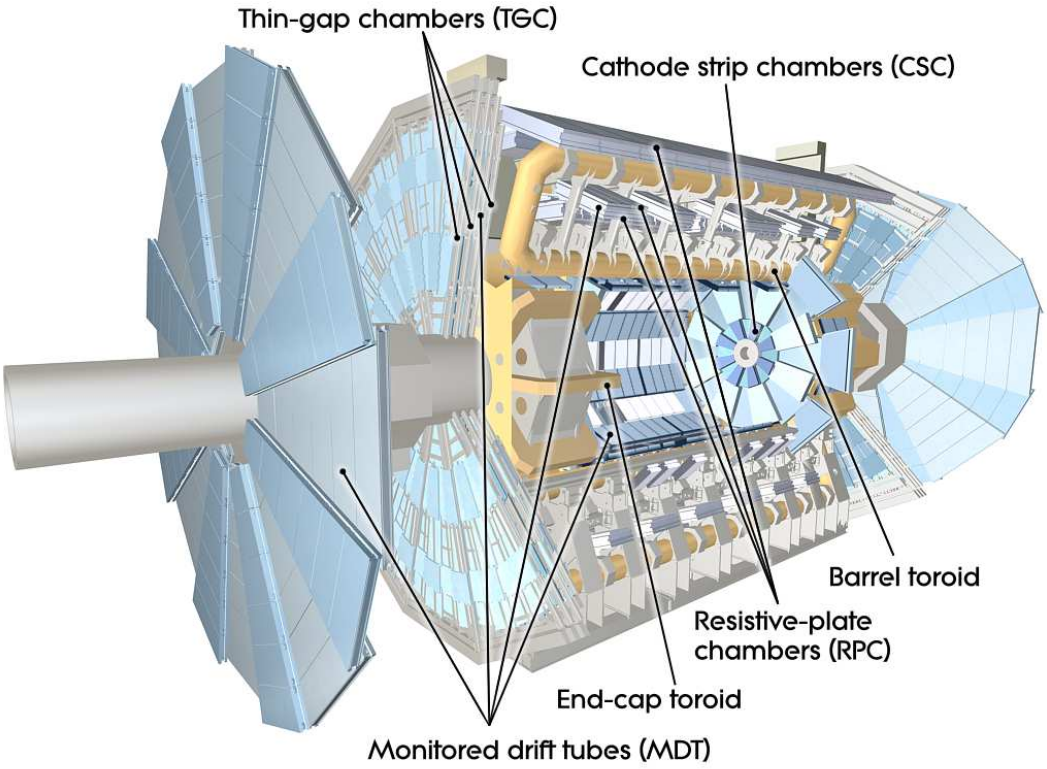


Fig. 3.6: Longitudinal view of the ATLAS muon system [109].

ticle, and **Thin Gap Chambers (TGC)**, multiwire proportional detectors. Both detectors are able to register signal in 15-25 ns, what allows for beam-crossing detection [126].

3.4.5 Triggers

It is not currently possible to store all the data coming from collisions inside the ATLAS detector. This would require to process around 1PB (peta byte) of data per second. For this reason, ATLAS uses trigger system to reduce and select events, which are potentially interesting for offline physics analysis. Each sub-system of the Trigger and Data Acquisition system (TDAQ) is usually connected with particular sub-detector. In Run 1 ATLAS used three-level trigger system - Level 1 (L1), Level 2 (L2) and event filter (EF). Each level was designed to reduce number of events passing the previous level. However, to cope with the higher luminosities of Run 2, L2 and EF have been combined to single High Level Trigger (HLT) [127].

L1 trigger is focused on processing data from calorimeters and muon spectrometer and identification of high p_{T} muons, electrons, photons, jets and τ leptons (which decay into hadrons), as well as large E_{T} and E_{T}^{miss}. These results are then processed by central trigger processor, which apply selection criteria according to trigger 'menu'. Role of the L1 is also to define one or more Region-of-Interest (ROI), which are coordinates in η

and Φ corresponding to regions with physically interesting features. The L1 processing time is around 2.5 μs and at the output the data rate is reduce to around 100 kHz.

This information about ROIs is then passed to **High Level Trigger**. This software trigger is used to further reduce only events in ROIs to the rate of ≈ 1 kHz in about 200 ms. Events passing this filter are written to disks and analysed offline.

The triggers used in the analyses described in this thesis are based on the presence of single lepton (either electron or muon) with p_{T} higher than given threshold in the event. The event is stored only if such leptons are registered, otherwise it is discarded. The efficiency of these single-lepton triggers increases with the lepton p_{T}, reaching almost 100% when the threshold is exceeded.

3.4.6 Magnetic system

As was already briefly mentioned in previous sections, ATLAS uses almost homogenous magnetic system [128], which ensures precise measurement of charged particles momenta. It consists of three large subsystems:

The **Central Solenoid** is located outside of the ID, its radius is 1.22 m and axial length 5.8 m. It is made of NbTi conductor, which creates homogenous 2T magnetic field. It is only 0.66 radiation length thick to absorb as little energy as possible. While for low p_{T}, the particles cannot be measured due to too large curvature of their trajectory, for $p_{\mathrm{T}} > 400$ MeV the detection efficiency reaches almost 100%.

The **Barrel Toroid** is assembled around the calorimeters as well as both end-caps and covers the barrel region ($|\eta| < 1.4$). Its eight coils offer deflecting power between 1.5 Tm and 5.5 Tm. It was designed mainly for purposes of muon spectrometer, where it curves muons trajectories. The barrel toroid is 25.3 m long and its inner and outer diameters are 9.4 m and 20.1 m, respectively.

The **End-Cap Toroids**, placed at both end-caps, extend radially from 1.65 to 10.7 m. They work on the similar principle as barrel toroid and provide even greater deflection 1-7.5 Tm. The system consists also of 8 square coil units assembled in insulation vacuum vessel. They cover the region $1.6 < |\eta| < 2.7$ providing measurements of forward muons.

3.4.7 Forward Detectors

Smaller detector systems, placed on both sides of detector's main body, are designed mostly for luminosity determination. First of the three 'luminosity' detectors is called **LUCID** (LUminosity measurement using a Cherenkov Integrating Detector) [129] and is located at 17 m from interaction point on both sides of the main detector. It is the main online relative-luminosity monitor, and as its name suggests, it detects Cherenkov radiation from inelastic pp scatterings in the forward direction. Second detector, called

ALFA (Absolute Luminosity For ATLAS) [130] is located at ±240 m from interaction point. Scintillating fibre trackers of this detectors, which are placed 1 mm close to the beam, are used for measuring elastic *pp* scatterings at very small angles. For measuring neutral particles (neutrons and photons) from meson decays, ATLAS uses its third system, called **ZDC** (Zero Degree Calorimeter) [131]. These are located ±140 m far from interaction point. The ZDC modules consist of alternating layers of quartz rods and tungsten plates, and are able to measure neutral particles in region $|\eta| \geq 8.2$.

To study soft and hard diffractive events at low luminosities, **AFP** (ATLAS Forward Proton) detector [132] identifies *pp* collisions, in which one or both protons emerge intact. This system is placed at 210 m from the interaction point.

Chapter 4

Object definitions

High energy pp collisions usually result in large number of particles. However, not all of them are stable enough to be directly measured by the ATLAS detector. Actually, there are only 14 particles which lifetime is large enough ($\tau c > 500\mu$m) to reach the detector. The most frequent particles that are registered by ATLAS subdetectors are electrons, muons, photons, $\pi^{\pm}$ and K mesons and protons and neutrons. After detection of particles emerging from the interaction point, these need to be identified and their properties (four-momenta) need to be reconstructed. Since each type of the particle produces specific signal in particular subdetector, various algorithms are employed to extract physics relevant information. This chapter presents basic features of reconstruction algorithms used for the identification of objects included in the $t\bar{t}Z$ signatures.

Figure 4.1 shows how different particles interact with particular detection systems of the ATLAS detector. The signatures observed in only some parts of the detection system can be therefore used for the identification of the type of particle. As can be seen from this figure, muons travel through whole detection system without stopping in any of them and their properties are reconstructed from the tracks they leave in the MS and ID. Photons, on the other hand, escape the ID without leaving the track, because they do not carry electric charge. Photons are therefore identified only if they produce measurable signal in the electromagnetic calorimeter. This differs them from electrons, which also produce electromagnetic shower in the calorimeter, but they leave tracks also in the ID, since they are charged. The similar technique can be used for distinguishing protons from neutrons. Both stop in the hadronic calorimeter but only protons carry electric charge and thus produce signal also in the ID. The only particle which successfully escapes all ATLAS detection systems is neutrino, thanks to its very low interaction cross section with the matter of the detector. Its presence, however, can be deduced from the imbalance in the overall momentum observed in the transverse plane.

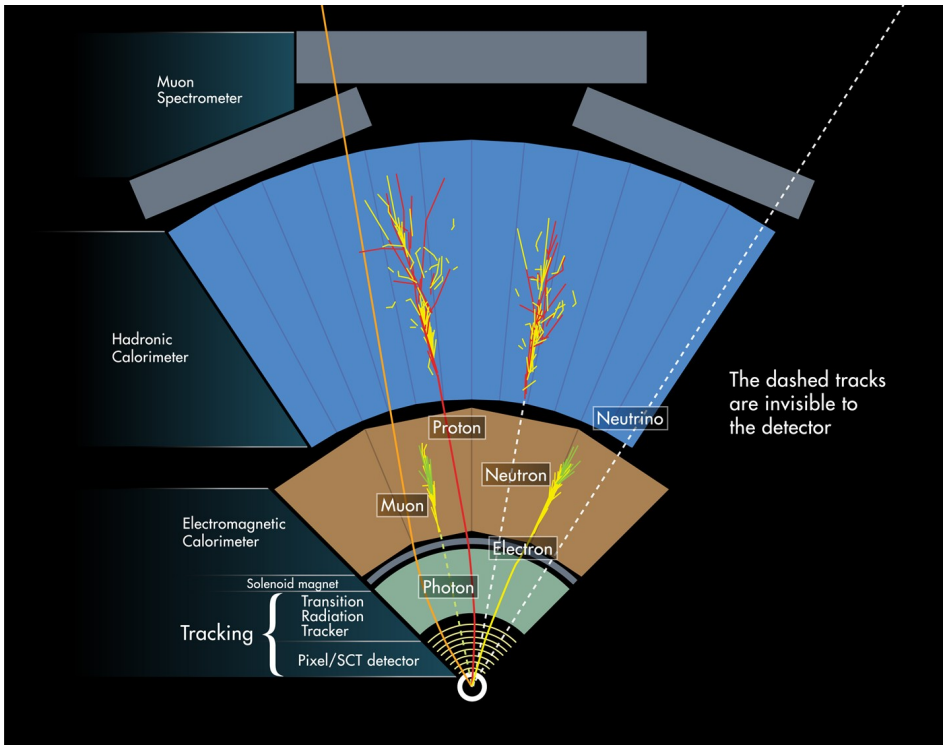


Fig. 4.1: Summary of the various interactions between particles and different subsystems of the ATLAS detector [133].

4.1 Tracks

Although tracks as separate objects are not directly used in the $t\bar{t}Z$ analyses, they play a crucial role in identification of the particles, measurement of their charge and momenta, and provide information about primary and secondary vertices in the event. The reconstruction of the tracks is primarily the role of the ID, in particular pixel detector and SCT. Detailed explanations of the reconstruction algorithms, which are briefly summarized here, can be found in [134, 135].

Track reconstruction begins by clusterization process in which raw measurements are clustered together to produce the final three dimensional curve corresponding to the track. Clusters are formed from pixels and strips in which deposited energy exceed given charge threshold. In the pixel detector, one cluster is enough to constitute one space point. In the SCT, however, clusters on both sides of the strip layer must be present to obtain 3D measurement.

In the next step the sets of three space points are merged into so-called track seeds. The impact parameters of each track seed are estimated from the assumption of a perfect helical trajectory caused by uniform magnetic field. Chosen track seeds are connected with additional space points from remaining layers of pixel detector and SCT which form a preliminary trajectory. Multiple track candidates are then built using combinatorial Kalman filter [136]. As all relevant candidates are identified, there

are usually a large number of candidate tracks which overlap or share common space points. These ambiguities are solved by assigning quality factor to each candidate.

Each track candidate is assigned following set of five parameters:

$$(d_0, z_0, \phi, \theta, q/p), \tag{4.1}$$

where d_0 and z_0 are minimum distances to the vertex in the transverse and longitudinal planes, respectively, ϕ and θ denote azimuthal and polar angle and q/p is the ratio of the track charge and its momentum. So-called *track score* is based largely on weight fractions that reflect intrinsic resolutions and expected cluster multiplicities in each subdetector. Moreover, there are several penalization factors, i.e. if track candidate contains holes (intersection of the track trajectory with sensitive detector element that does not produce cluster), its track score is reduced. Track momentum is also important factor, which promotes more energetic tracks and suppresses numerous tracks with low p_{T} that have usually incorrectly assigned clusters. Ambiguity solver also rejects candidates if they fail in at least one of the following criteria:

- $p_{\mathrm{T}} > 400$ MeV,
- $|\eta| < 2.5$,
- At least 7 pixel and SCT clusters,
- Maximum of one shared cluster in the pixel detector or two shared clusters in the SCT on the same layer,
- Not more than two holes in combined pixel and SCT detectors and not more than one hole in the pixel detector,
- $|d_0^{BL}| < 2.0$ mm,
- $|z_0^{BL} \sin\theta| < 3.0$ mm,

where d_0^{BL} and z_0^{BL} are transverse and longitudinal impact parameters calculated with respect to the beam line.

4.2 Electrons

Electrons, as well as muons, are among the most important physics objects relevant for the analyses described in this thesis. They can be identified and reconstructed with high precision in the ATLAS detector and thus offer a reliable source of physics information.

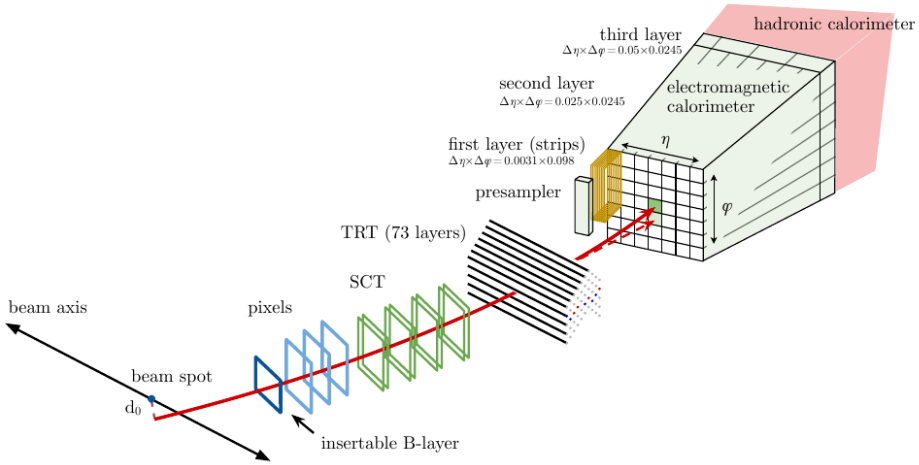


Fig. 4.2: Path of an electron through the ATLAS detector. The red curve represents the electron trajectory, first through the ID (pixel and SCT detectors) and TRT before stopping in the electromagnetic calorimeter. The dashed red line represents photon produced in the process of bremsstrahlung. Figure is taken from [137].

As already mentioned in the beginning of this chapter, electrons leave tracks in the ID, where their momentum and position can be measured, and also in the electromagnetic calorimeter serving for measuring their energies. The following sections present different steps in process of establishing full electron physics object. The more detailed description of mentioned algorithms can be found i.e. in [137, 138].

4.2.1 Electron reconstruction

Electrons can lose their energy due to bremsstrahlung when they interact with the material of the detector. This can occur in calorimeter but also in any other part of detector through which they traverse. In this process electrons emit photons which can be again converted into positron-electron pair that is again able to interact with the detector medium. This process, called electromagnetic shower, is crucial for idetification of electron candidates. The showers are usually collimated and can be reconstructed as a part of the same electromagnetic cluster. Reconstruction of electron candidates is therefore based on three fundamental components: energy clusters found in electromagnetic calorimeter, tracks from the ID, and close matching in $\eta \times \phi$ space between tracks and clusters. Schematic illustration of an electron traversing through the particular subdetectors can be found in Figure 4.2.

Electromagnetic calorimeter is divided into 200×256 elements, called towers, which cover the effective $\eta \times \phi$ space. For each of these elements, the overall tower energy is obtained as a sum of energy deposits in particular layers of the calorimeter. The so-called sliding-window algorithm [139] is used to preselect seed-cluster candidates of

size 3×5 towers, whose summed transverse energy E_{T} exceeds 2.5 GeV. If two seed-cluster candidates overlap, or are in close proximity, only the one with the higher E_{T} (or highest-E_{T} central tower) is kept.

The second step, reconstruction of the tracks from the ID, was already covered in previous Section 4.1. The matching of the seed-cluster candidate with the track candidate is based on their separation in $\eta \times \phi$ space. Algorithm selects only candidate pairs which fulfill the requirement on η separation $|\eta_{\mathrm{cluster}} - \eta_{\mathrm{track}}| < 0.05$ and separation in ϕ plane $-0.10 < -q(\phi_{\mathrm{cluster}} - \phi_{\mathrm{track}}) < 0.05$, where q is the sign of the particle charge [1]. Finally, a seed-cluster candidate with an associated track with four or more hits from the silicon detectors, and no association with a photon-conversion vertex, is considered as an electron candidate.

4.2.2 Electron identification

The role of the identification algorithms is to determine if the electron candidate can be truly considered as prompt electron, or its signature just mimics real electron. A set of criteria is used to filter any non-prompt background processes, dominated by aforementioned photon conversions, but also electrons originating from decays of heavy-flavor quarks and from hadronic jets.

The algorithm relies on the likelihood based approach (LH) in the central detector region (where $|\eta| < 2.47$). A set of total 14 measurable quantities are used as an input to the likelihood, the most of them being extracted from previous measurements in the trackers, electromagnetic calorimeter, and other components depicted in Figure 4.2.

The likelihood function is constructed from the probability density functions (pdfs) $P_{S(B)}$ based on the simulated samples, and is defined with following equation for signal (S) and background (B):

$$L_{S(B)}(\mathbf{x}) = \prod_{i=1}^{n} P_{S(B),i}(x_i), \tag{4.2}$$

where $\mathbf{x}$ is the vector of aforementioned quantities and product goes over n various pdfs. The signal pdfs are constructed from simulated prompt electrons and background contains simulation of all sources of non-prompt backgrounds mentioned previously. To determine if the candidate really corresponds to the electron, discriminant is defined in following form:

$$d_L = \frac{L_S}{L_S + L_B}. \tag{4.3}$$

Based on the chosen discriminant value, various operating points are defined. This definition of the discriminant brings issues with fine binning in the sharp peak and

[1] This asymmetric requirement accounts for the differences in the positively (negatively) charged particles which bend in the negative (positive) ϕ direction in the magnetic field.

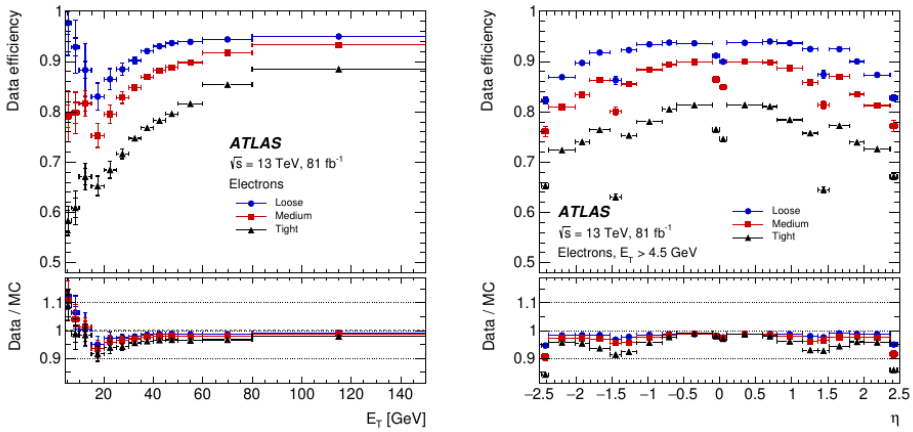


Fig. 4.3: Measured electron identification efficiencies in $Z \to ee$ events as a function of E_{T} (left) and η (right) for the Loose, Medium and Tight operating points. Distributions are obtained by applying data-to-simulation efficiency ratios measured in $J/\psi \to ee$ and $Z \to ee$ events to $Z \to ee$ simulated samples. Uncertainties include both statistical (inner marker) and systematic (outer marker) components. The observed drops in the efficiency for $|\eta|$ around 0 and 1.4 are caused by the gap between two calorimeter half-barrels in the TRT (drop at $|\eta| \approx 0$), and transition region between barrel and end-cap calorimeters (drop at $|\eta| \approx 1.4$). Figure is taken from [138].

therefore the distribution is transformed using inverse sigmoid function:

$$d'_L = -\tau^{-1} \ln(d_L^{-1} - 1), \tag{4.4}$$

where τ denotes parameter, usually fixed to value 15 [140]. If the value of transformed discriminant d'_L is larger than the value determined by the chosen operating point, the electron candidate is accepted as real electron, otherwise it is rejected as coming from non-prompt background. The pdfs for the LH method are derived for two regimes according to transverse energy E_{T}. In low energy regime datasets of $J/\psi \to ee$ events are used for signal pdfs and minimum bias events as non-prompt background, while at high E_{T}, the signal is constructed from $Z \to ee$ events and background from dijet events.

The four operating points usually used in physics analyses are reffered to as Very-Loose, Loose, LooseAndBLayer, Medium and Tight, and correspond to increasing thresholds for the likelihood discriminant. The identification efficiencies for the three oparating points as function of E_{T} and η are shown in Figure 4.3. For all analyses presented in this thesis, Medium operating point is employed.

4.2.3 Electron isolation

To further suppress non-prompt electron contamination in the analysed data, an additional requirement on the electron isolation is imposed. The isolation algorithm is based

on the quantification of activity in the vicinity of the candidate object, since prompt electrons usually produce collimated showers with small radius. Isolation criteria again make use of both calorimeter-based, as well as tracking quantities.

To improve the isolation capabilities of the calorimeter, co-called topological clusters are constructed instead of using single cells. These clusters are seeded by cells whose deposited electromagnetic-scale energy exceeds at least four times the expected noise-level threshold [2] of particular calorimeter cell. Clusters are then expanded with neighboring cells for which deposited energy is more than twice the noise level. In the next step topological clusters within chosen cone radius ΔR are clustered together, and their transverse energies are summed into raw isolation energy variable $E_{\mathrm{T,raw}}^{\mathrm{isol}}$. This variable still includes the energy deposited by the candidate electron that needs to be subtracted, together with correction terms for pileup and leakage (energy contained in topo-clusters originating from cells outside the cone).

Track-based isolation variables are constructed from tracks with $p_{\mathrm{T}} > 7$ GeV that are reconstructed in the fiducial region of the ID ($|\eta| < 2.47$, and excluding LAr crack region with $1.37 < |\eta| < 1.52$) and fulfill basic quality requirements described in Section 4.1. For further pileup suppression, the cut on the longitudinal impact parameter in a form $|z_0 \sin\theta| < 0.5$ mm, is applied. This requirement aims to select only tracks originating from the vertex relevant for the process of interest. Moreover, the cut on the transverse impact parameter $|d_0/\sigma(d_0)| < 5.0$ is employed. The isolation variables are then defined in a similar way as for the calorimeter, but in this case the scalar transverse momenta of the tracks within particular cone of radius ΔR around candidate track, are summed into $p_{\mathrm{T}}^{\mathrm{isol}}$. Thanks to the higher granularity of the ID w.r.t. electromagnetic calorimeter, variable-sized cone radius, depending on the $p_{\mathrm{T}}^{\mathrm{isol}}$, is defined as follows:

$$\Delta R = \min\left(\frac{10\mathrm{GeV}}{p_{\mathrm{T}}[\mathrm{GeV}]}, R_{max}\right), \tag{4.5}$$

where R_{max} is the maximum cone size (between 0.2 and 0.4) and factor 10 GeV is determined from simulated $t\bar{t}$ sample to maximize the rejection of background.

Depending on the different combinations of calorimeter and track-based criteria, following electron isolation operating points are defined: Gradient, HighPtCaloOnly, FixedCutLoose, FixedCutTight, PromptLeptonVetoLoose, PromptLeptonVetoTight [141]. The analyses discussed in this dissertation employ FixedCutTight (FCTight) and PromptLeptonVetoLoose (PLVLoose) operating points. The first one, FCTight, sets fixed cuts on $E_{\mathrm{T}}^{\mathrm{isol}}/p_{\mathrm{T}} < 0.20$ and $p_{\mathrm{T}}^{\mathrm{isol}}/p_{\mathrm{T}} < 0.15$, while PLVLoose is based on the Boosted Decision Tree (BDT) approach, where track jet properties are used for the training. In this approach the training set consists of signal electrons from W and τ decays, while background includes electrons from semi-leptonic B-hadron decays.

[2] Threshold includes both electronic noise and effect of pileup.

4.3 Muons

As a charged particles, muons also leave track in the ID, but since they are far more massive than electrons, they do not predominantly lose their energy through bremsstrahlung and thus do not produce electomagnetic showers in the calorimeter. ATLAS therefore constructed special detector, muon spectrometer, responsible for measuring muon properties. Basic algorithms developed for the reconstruction and identification of muons are briefly described in the following sections. Detailed overview can be found in [142, 143].

4.3.1 Muon reconstruction

In the ID, muon tracks are reconstructed following the same procedure as described for electrons in Section 4.2.1. This section therefore focuses on the second stage that is reconstruction in the MS. It starts with the search for hit patterns inside each muon chamber, what results in formation of sensitive segments. Muon candidates are then built by merging together nearby hits from segments in different layers. Selection of segments is based on the hit multiplicity criteria: there have to be at least two matching segments in central region, while only one high-quality segment is necessary for the barrel-endcap transition region. The hits associated with each track candidate are subsequently fitted with global χ^2 fit, and accepted only if the χ^2 values satisfy given selection criteria.

In addition to the independent reconstruction procedures in the ID and MS the combined reconstruction algorithms, which use both hits from the MS and ID, are usually employed. The following four muon types are defined based on the used subdetectors:

- **Combined muons** (CB) - reconstruction is performed independently in the ID and MS, and a combined track candidate is built from global refit using hits from both the ID and MS subdetectors.

- **Segment-tagged muons** (ST) - a track in the ID is required to be matched with at least one track segment in the MDT or CSC chambers. This category is used when muon traverse only through one layer of the MS due to its low p_{T} or because it falls in low acceptance region.

- **Calorimeter-tagged muons** (CT) - this category uses ID tracks matched with energy deposit in the calorimeter corresponding to the minimum-ionizing particle. Such muons are reconstructed with lowest purity. This type is designed to recover acceptance in the region where the MS is only partially instrumented.

- **Extrapolated muons** (ME) - muons are reconstructed only from the trajectory they leave in the MS, and no track from the ID is required. The trajectory,

however, must be originating from the iteraction point and the muon is required to travel at least two layers (three layers in forward region) of the MS. ME muons are used to extend acceptance for muons into the region $2.5 < |\eta| < 2.7$, which is not covered by the ID.

4.3.2 Muon identification

The goal of the muon identification is to suppress the background signatures mimicking muons, such as pion and kaon decays and to guarantee a robust momentum measurement. Several discrimination variables are used to distinguish prompt muons from background candidates. For CB muons following variables are employed in the identification:

- *q/p significance* - absolute value of the difference between charge/momentum ratio of the candidate measured in the ID and MS divided by their corresponding uncertainties,
- ρ' - absolute value of the difference between p_{T} of the candidate measured in the ID and MS divided by the p_{T} of the combined track,
- *normalised* χ^2 - resulting from the combined track fit.

For a precise momentum measurement, specific criteria are imposed on the number of hits in the ID and MS.

Similar to the electron identification, following selections, which are designed to address specific needs of physics analyses, are defined for muons.

- **Medium muons** - this category, which is considered as the default in ATLAS, is designed to reduce the systematic uncertainties associated with reconstruction and calibration of muons. Medium muons are reconstructed only from CB and ME tracks. To suppress contamination from hadrons misidentified as muons, a loose selection is applied on the compatibility between momentum measured in the ID and MS. To achieve this, *q/p significance* have to be less than seven. This operating point was chosen for the analyses described in this thesis. The reconstruction efficiency of medium muons as a function of transverse momentum and pseudorapidity is shown in Figure 4.4.
- **Loose muons** - this selection provides maximum reconstruction efficiency while preserving good-quality muon tracks and was optimised for the specific needs of analyses involving Higgs boson decay into four leptons.
- **Tight muons** - to provide maximal purity at the cost of efficiency, tight muons are defined. This type uses only CB tracks with hits in at least two layers of

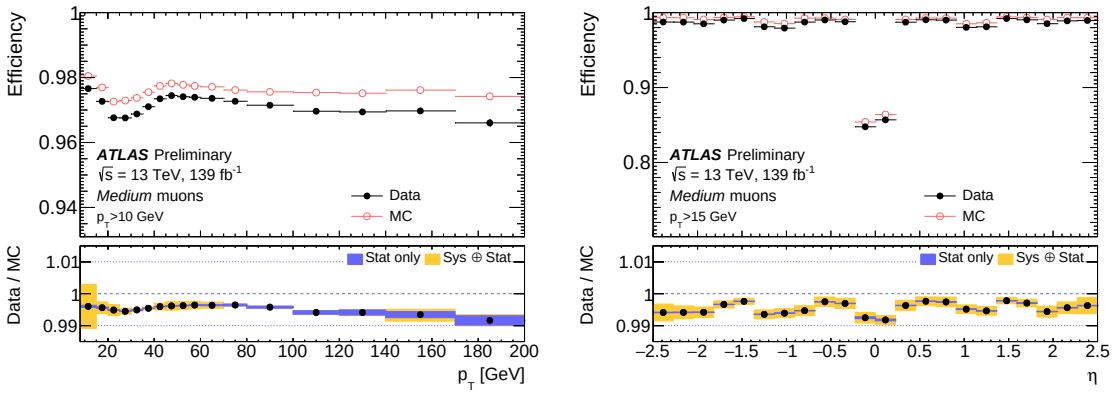


Fig. 4.4: Measured muon reconstruction efficiencies in $Z \to \mu\mu$ events as a function of muon p_{T} (left) and pseudorapidity η (right) using Medium identification algorithm. Only the muons with $p_{\mathrm{T}} > 15$ GeV are used. Data distributions are obtained by applying data-to-simulation efficiency ratios measured in $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ events to $Z \to \mu\mu$ simulated events. Figure is taken from [144].

the MS. Moreover, tight muons has to satisfy all selection criteria required for medium muons and normalised χ^2 of combined track fit less than eight.

- **High-p_{T} muons** - this selection targets searches for high-mass Z' and W' resonances by maximizing momentum resolution for track with p_{T} more than 100 GeV. This selection require passing medium criteria and stricter requirements are used for χ^2 fit and momentum measurements in the ID and MS.

- **Low-p_{T} muons** - This operating point was newly added to offer identification of the lowest-p_{T} muons, which usually do not leave more than one hit in the MS. Selection criteria are focused on reduction of large non-prompt muon background dominated by charged hadron decays. This category was designed for analyses measuring quark-mixing parameters of the SM and supersymmetry searches. There exist two versions of this selection, one cut-based and second employing multivariate techniques.

4.3.3 Muon isolation

Similar to electrons, also muons are subjected to the isolation criteria which aim to further reduce non-prompt candidates emerging from background processes. The same idea as for electrons is used, specifically, that muons originating from decays of massive SM bosons are often isolated from other particles. By measuring detector activity in vicinity of a muon candidate it is possible to effectively eliminate non-prompt muons from semileptonic decays embedded in jets.

Following the procedure already described for electrons, two variables are defined for muon isolation, one associated to track measurement and one calorimeter-based. The track-based variable, denoted as $p_{\mathrm{T}}^{varcone30}$, is defined, in analogy with electrons, as a scalar sum of the track transverse momenta greater than 7 GeV in cone of radius ΔR as defined in Eq. 4.5, where $R_{max} = 0.3$. In analogy with electrons, cuts for longitudinal and transverse impact parameters are used, specifically $|z_0 \sin\theta| < 0.5$mm and $|d_0/\sigma(d_0)| < 3.0$. The calorimeter-based isolation variable is defined analogously, introducing $E_{\mathrm{T}}^{varcone20}$ as a sum of the transverse energies of topological clusters which are present in a cone of size $\Delta R = 0.2$ around the muon candidate. Again the energy deposited by muon itself is subtracted, and the correction for pileup effects is taken into account.

The selection criteria, that define particular isolation operating points, are based on the so-called *relative isolation variables*. These variables denote isolation variables described above, but are divided by the transverse momentum of the muon candidate. In total seven isolation operating points are designed, each optimized for particular physics analyses. The operating points relevant for analyses described in this thesis are *FixedCutTightTrackOnly* and *PLVLoose*. As the name suggests *FixedCutTightTrackOnly* uses fixed cut on transverse momenta in a form $p_{\mathrm{T}}^{varcone30}/p_{\mathrm{T}}^{\mu} < 0.06$ applied for the tracks reconstructed in the ID. On the other hand, *PLVLoose* operating point employs BDT approach for the muon selection and is implemented in the same way as for electrons.

4.4 Jets

Due to the quark confinement (see Section 1.4.2 for details) color-charged particles cannot be directly observed in the ATLAS detector. As a consequence, quarks undergo process of hadronisation in which they couple with other quarks to form hadrons. Similar as electrons, also hadrons produce showers of particles (called jets) which are subsequently registered by the calorimeters. Energy deposited by these showers in the calorimeter are then used to form jets. This section presents algorithms developed for the jets reconstruction, as well as sophisticated techniques for identifying flavor of quark which caused the shower. More detailed description of these algorithms can be found i.e. in [145, 146, 147].

4.4.1 Jet reconstruction

The primary procedure for the reconstruction of jets uses exclusively energy deposits from the calorimeters. In the first step, active calorimeter cells are clustered into three-dimensional topological clusters following the nearest neighbor algorithm [148]. Topo-

clusters are formed from cells according to the cell energy divided by the expected noise factor. Final energy of the topo-cluster is then defined at the electromagnetic scale (EM) which correctly measures energy deposits in the electromagnetic calorimeter. These, so-called *EMtopo* jet candidates, consider only positive energy topo-clusters and are expected to originate from the primary vertex of the hard-scatter process.

Although this algorithm provides robust energy measurement and good resolution in wide kinematic range, the novel technique, combining information from calorimeters and tracker, was developed for ATLAS physics analyses after the end of Run 2. This, so-called *PFlow* algorithm [147] is currently used as default approach for most of the ATLAS analyses. The procedure still uses topo-clusters obtained from calorimeter measurements, but the energy deposited in the calorimeter by charged particles is subtracted and replaced by the momenta of tracks corresponding to these particles as measured in the ID. Since the tracking detectors in the ID have far better resolution, especially for low-p_{T} particles, this replacement improves final energy and angular resolution, as well as reconstruction efficiency and pileup stability, when compared to purely calorimeter-based reconstruction. Subtraction is, however, done only for particles with low energies, because calorimeter resolution for high energetic particles is compatible with tracker resolution. The energy threshold at which PFlow algorithm effectively becomes EMtopo is governed by following relation:

$$\frac{E^{clus} - \langle E_{dep} \rangle}{\sigma(E_{dep})} > 33.2 \log \left(\frac{40\mathrm{GeV}}{p_{\mathrm{T}}^{trk}} \right), \tag{4.6}$$

where E^{clus} is energy of cluster in a cone of size $\Delta R = 0.15$, $\langle E_{dep} \rangle$ is the expected mean energy deposited by the pions and p_{T}^{trk} is transverse momentum of the track. If the above relation is true then the subtraction is not performed and PFlow jet is equivalent to EMtopo jet.

Both PFlow and EMtopo approaches are employed for analyses discussed in this thesis, but in each case only jets with $p_{\mathrm{T}} > 25\mathrm{GeV}$ and reconstructed in pseudorapidity range $|\eta| < 2.5$, are considered.

4.4.2 Formation of jets

The four-vectors resulting from topo-clusters, or topo-clusters combined with tracks in the case of PFlow jets, are combined together by clustering procedure known as k_T algorithm. There exist multiple variations of this algorithm based on the values of free parameters entering its definition. The core of the method lies in determining relative distance $d_{i,j}$ between clusters i and j, which is defined as follows:

$$d_{i,j} = \min(k_{T,i}^{2n}, k_{T,j}^{2n}) \frac{\Delta R_{i,j}^2}{R^2}, \tag{4.7}$$

where k_T denote transverse momenta of clusters (hence the name of the algorithm), ΔR represents their angular separation and R is the radius of the cone in which clusters are merged. Important is the integer parameter n which determines momentum dependence of above equation and is used to differentiate between various forms of the k_T algorithm. Second variable, which is necessary for the jet formation, is relative distance between cluster i and the beam B defined as:

$$d_{i,B} = k_{T,i}^{2n} \tag{4.8}$$

The clustering process begins by calculating both distances $d_{i,j}$ and $d_{i,B}$ for all clusters. If $d_{i,j} < d_{i,B}$ then the clusters i and j are merged together. If the previous relation has to be inverted to be true then the two clusters are not merged and cluster i is defined as jet. This procedure is applied to all clusters iteratively until there are no unmerged input clusters.

As already mentioned, the choice of n determines the characteristic merging pattern. If n is chosen to be zero, then the algorithm is called Cambridge-Aachen [149], and momenta of the particles are effectively excluded from the merging process. The clustering is then performed by simply sorting clusters based on their presence in chosen cone radius R. When n is chosen to be equal one, then the algorithm is simply called k_T [150], and lower-momentum clusters are prioritized and grouped first. The most frequently used procedure is, however, when $n = -1$. This, so-called anti-k_T algorithm [145], used also throughout this dissertation with the cone radius $R = 0.4$, groups first clusters with higher momentum. The advantage of the anti-k_T algorithm lies in the idea that thanks to the inverse dependence on the transverse momentum, relative distance between two low-p_{T} clusters is relatively large, what prevents them from being merged together. On the other side, high-p_{T} cluster combines easily with the low-p_{T} clusters in its vicinity, since relative distance decreases with higher transverse momentum. This results in the circular shapes of all jets reconstructed with this technique as can be seen in Figure 4.5. In this Figure one more algorithm is presented, called SISCone [151]. This jet algorithm was designed to be infrared-safe and uses so-called split/merge procedure for overlapping jets, which assigns the clusters in overlapping regions to the jet with closer center.

4.4.3 Jet Vertex Tagger

Proton-proton collisions in the LHC, which are considered as a source of hard-scatter processes, are usually accompanied by additional collision that need to be filtered. Such additional low-p_{T} collisions are collectively referred to as *pileup* interactions. They consists of two types, so-called *in-time*, which arise due to additional pp collisions present in single bunch crossing, and *out-of-time*, which are caused by energy deposits in calorimeters corresponding to previous, or following, bunch crossing.

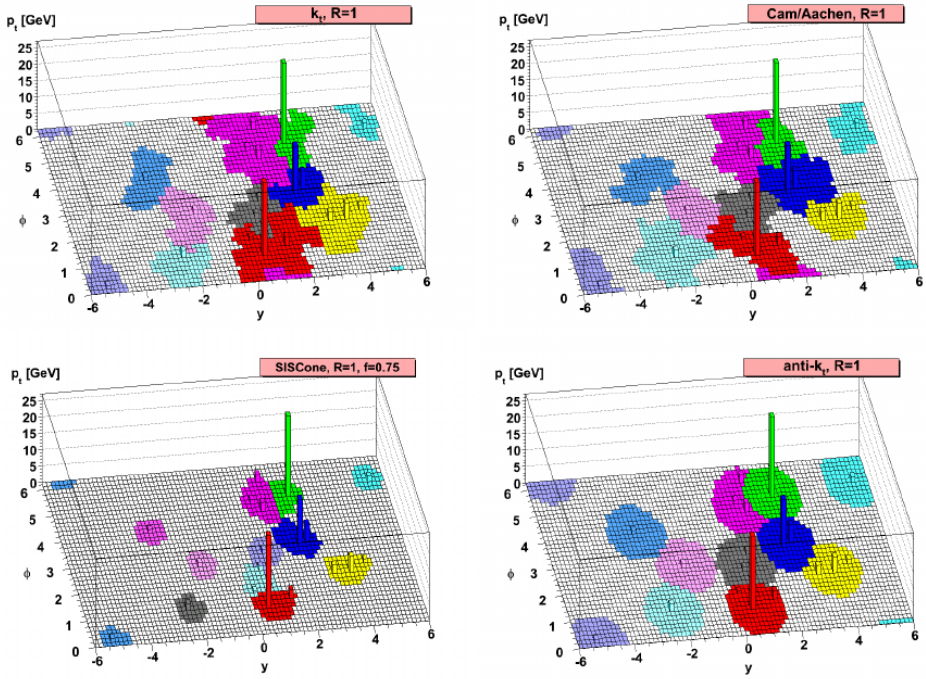


Fig. 4.5: Example of jet clustering algorithms applied to sample event generated with Herwig MC generator. Many random soft "ghost" particles were simulated in the vicinity of jet candidates. These were clustered into jets (colored areas) using four algorithms: k_T (top left), Cambridge-Aachen (top right), SISCone (bottom left) and anti-k_T (bottom right). Typical behavior of anti-k_T technique can be seen in region near $y = 2$ and $\phi = 5$ where green jet with higher p_{T} was prioritized over purple jet with lower p_{T}, while other algorithms place the boundary midway between the two jets. Figure is taken from [145].

The usual procedure is to subtract average pileup-related transverse energy from signal interaction. But, due to the local fluctuations in the pileup activity which may cause spurious jets, alternative approach based on the construction of discriminating variables, is employed.

The so-called *Jet-Vertex-Fraction* (*JVF*) variable is defined as the ratio between scalar sum of the transverse momenta of tracks corresponding to particular jet originating from the primary vertex of the hard-scatter, and the scalar p_{T} sum of all associated tracks:

$$JVF = \frac{\sum_k p_{\mathrm{T}}^{\mathrm{trk}_k}(\mathrm{PV}_0)}{\sum_l p_{\mathrm{T}}^{\mathrm{trk}_l}(\mathrm{PV}_0) + \sum_{n\geq 1}\sum_l p_{\mathrm{T}}^{\mathrm{trk}_l}(\mathrm{PV}_n)}, \tag{4.9}$$

where PV_0 denotes primary vertex and PV_n, $n \geq 1$ represents primary vertices of pileup interactions in the same bunch crossing. The second term in the denominator therefore corresponds to tracks that originate from the pileup interactions. This formula was later improved to take into account also number of pileup tracks n_{trk}^{PU} which are associated with vertices other than primary vertex. Modified discriminant, corrected JVF, is then defined as follows:

$$corrJVF = \frac{\sum_k p_{\mathrm{T}}^{\mathrm{trk}_k}(\mathrm{PV}_0)}{\sum_l p_{\mathrm{T}}^{\mathrm{trk}_l}(\mathrm{PV}_0) + \frac{\sum_{n\geq 1}\sum_l p_{\mathrm{T}}^{\mathrm{trk}_l}(\mathrm{PV}_n)}{k\cdot n_{trk}^{\mathrm{PU}}}}, \tag{4.10}$$

where factor $k = 0.01$. As can be seen in Figure 4.6b, the jet efficiency for *corrJVF* is stabilised at around 90% with increasing number of vertices N_{Vtx}, while for uncorrected *JVF* it degrades from 97% to around 75% for 30 vertices. Figure 4.6a shows distributions of *corrJVF* for pileup (green) and hard-scatter jets (blue). A value of *corrJVF*=-1 in this Figure corresponds to jets without associated tracks.

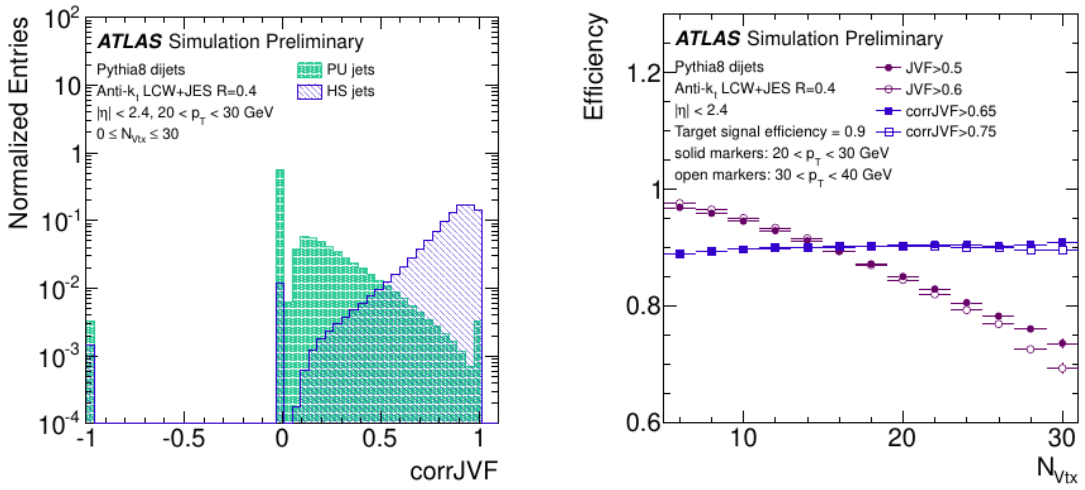
Second discriminating variable, which is then used together with *corrJVF* in construction of final *Jet Vertex Tagger* (JVT) discriminant, is designed to test compatibility between jet and its associated tracks. This variable, denoted as R_{pT}, is defined as the ratio between scalar sum of p_{T} of tracks associated with the jet originating from primary vertex and its calibrated p_{T}:

$$R_{\mathrm{pT}} = \frac{\sum_k p_{\mathrm{T}}^{\mathrm{trk}_k}(\mathrm{PV}_0)}{p_{\mathrm{T}}^{\mathrm{jet}}} \tag{4.11}$$

A value of R_{pT} peaks at 0 and is steeply falling for pileup jets, while for hard-scatter jets its mean value and spread is larger.

JVT discriminant is built from the two variables defined above using multivariate technique, called k-nearest neighbor (kNN) algorithm [140]. Resulting two-dimensional likelihood is used to calculate the probability that particular jet have originated from the hard-scatter process.

Similar algorithm, called *forward* JVT (fJVT) [153], was developed to mitigate effects of pileup in forward region (where $|\eta| > 2.5$). In the first step jets are assigned



(a) (b)

Fig. 4.6: (a) Distribution of the *corrJVF* variable for pileup (green) and hard-scatter (blue) jets in dijet sample generated by Pythia8. Only jets with $20 < p_{\mathrm{T}} < 30$ GeV are used. (b) Dependence of hard-scatter jet efficiency on the number of vertices for *JVF* and *corrJVF* discriminants. Two sets of jets with $20 < p_{\mathrm{T}} < 30$ GeV (solid markers) and $30 < p_{\mathrm{T}} < 40$ GeV (open markers). Figures are taken from [152].

to the vertex PV_i, which corresponds to the highest value of R^i_{pT}. Then the missing transverse momentum for the pileup candidate i is computed as a weighted vector sum of track and jet momenta:

$$\langle \mathbf{p}^{\mathrm{miss}}_{\mathrm{T},i} \rangle = -\frac{1}{2}\left(\sum_{\mathrm{tracks}\in \mathrm{PV}_i} k\mathbf{p}^{\mathrm{track}}_{\mathrm{T}} + \sum_{\mathrm{jets}\in \mathrm{PV}_i} k\mathbf{p}^{\mathrm{jet}}_{\mathrm{T}} \right), \tag{4.12}$$

where factor $k = 2.5$ is used to correct for intrinsic differences between track and jet terms and optimizes the rejection of forward pileup jets. The fJVT discriminant is then constructed for the vertex i as the normalized projection of the missing transverse momentum $\langle \mathbf{p}^{\mathrm{miss}}_{\mathrm{T},i} \rangle$ on the vector of transverse momentum of the forward jet $\mathbf{p}^{\mathrm{fj}}_{\mathrm{T}}$:

$$\mathrm{fJVT}_i = \frac{\langle \mathbf{p}^{\mathrm{miss}}_{\mathrm{T},i} \rangle \cdot \mathbf{p}^{\mathrm{fj}}_{\mathrm{T}}}{|\mathbf{p}^{\mathrm{fj}}_{\mathrm{T}}|^2} \tag{4.13}$$

If the value of $\mathrm{fJVT} = \max_i(\mathrm{fJVT}_i)$ is above certain threshold, then the forward jet is considered as coming from pileup. The distributions of JVT and fJVT can be found in Figure 4.7. In the analyses presented in this thesis, JVT cut of 0.5 for jets with $p_{\mathrm{T}} < 60$ GeV and 0.59 for jets with $p_{\mathrm{T}} < 120$ GeV, is employed.

4.4.4 Flavor tagging

The aim of the flavor tagging procedure is to identify flavor of the quark which initiated formation of jet. This is possible thanks to the differences in hadronisation process and

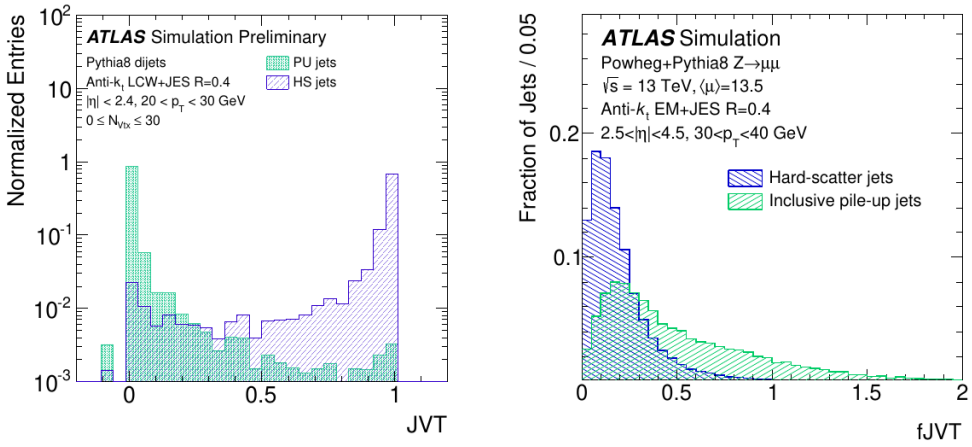


Fig. 4.7: Normalised distributions of JVT (left) and fJVT (right) discriminants for pileup (green) and hard-scatter jets (blue). JVT distribution was obtained from dijet events generated by Pythia8 and only jets with $20 < p_{\mathrm{T}} < 30$ GeV were considered. fJVT discriminant was evaluated from $Z \to \mu\mu$ sample and only for jets with $30 < p_{\mathrm{T}} < 40$ GeV in the region $2.5 < |\eta| < 4.5$. Figures are taken from [152] and [153], respectively for JVT and fJVT.

jet properties for different quarks types. The main focus of this section is identification of jets initiated by b quarks, since these are essential part of $t\bar{t}$ pair decay relevant for analyses presented here. There exist also algorithms designed for identification of jets from c or light (u, d and s) quarks, and procedures for distinguishing quark- and gluon-initiated jets, but these are not further discussed in this section.

The most important characteristic of the B hadrons (from which b jets are formed) that enables their robust identification, is their lifetime of $\tau = 1.5$ ps. This exceptionally long lifetime translates to distance $c\tau$ of about 0.45 mm which b quark travels before its decay. This feature, together with specific topology of the B hadrons, including decays to c quark, serve as a basis for low-level algorithms, which then offer solid inputs for high-level discriminants.

B hadrons are typically identified according to the presence of secondary vertex, which is characterized by mentioned ≈0.45 mm separation from primary vertex. Algorithm for finding such secondary vertices is denoted as SV1 [154], and is based on the construction of likelihood discriminant which uses properties of the tracks to differentiate between B hadrons and other objects such as long-lived particles (i.e. kaons), photon conversions, or hadrons created in interactions with the material of the detector.

Transverse and longitudinal impact parameters of associated tracks, d_0 and z_0 are also successfully employed in identification of B hadrons. These parameters, which measure distance from the primary vertex, create a basis for algorithms IP2D and IP3D [155]. The core of this methods is in construction of conditional likelihood tem-

plates for various initial quark flavors. The output of this method is the reconstructed track trajectory in two (IP2D) or three (IP3D) dimensions. Unfortunately, these algorithms are statistically limited when computing templates, and hence were replaced by algorithm based on recurrent neural networks, RNNIP. This one is trained using variables constructed from track properties and achieves better performance than IP2D and IP3D.

Due to the large branching ratio of b quarks decaying into c quarks, and non-negligible lifetime of outgoing c quarks that translates into relatively long flight path, tertiary vertex is often present in b jets. Although secondary vertex corresponding to b quark is usually too close to tertiary vertex of c quark, so-called JetFitter (JF) algorithm [156] is designed to identify such tertiary vertices. It reconstructs most probable vertex structure of the jet using intersection points of tracks with the jet axis.

The outputs of these low-level algorithms and variables designed in their construction are subsequently used in high-level flavor tagging techniques. ATLAS uses predominantly two categories of tagging algorithms that are briefly described in the following text. Both categories are used for b tagging in the analyses described in this thesis. More details about the algorithms can be found i.e. in [157, 158, 159].

MV2 taggers

This flavor tagging algorithm was typically used as default in ATLAS for analyses conducted during Run 2 of the LHC operation. It was replaced with newly developed DL1 algorithm discussed next. MV2 taggers are based on multivariate techniques, specifically BDT trained with TMVA package [140] from ROOT. The most recent version, MV2c10, which is used also throughout this thesis, uses BDT trained on $t\bar{t}$ and $Z \to t\bar{t}$ samples, in which b jets in final state are treated as signal. In the process of training, 21 variables constructed from combined outputs of SV1, IP2D, IP3D and JF algorithms, are used together with p_{T} and η of jets. To achieve balanced training set, both b and c jets are reweighted according to p_{T} and η of light jets. The sample considered as background in the training consists of 7% of c jets and 93% of light jets. The final MV2c10 discriminant is depicted in Figure 4.8.

DL1 taggers

The most striking difference between MV2 and DL1 taggers lies in the form of output. While MV2 algorithm outputs binary discriminant determining if a jet is compatible with b quark or not, the DL1 procedure results in three-dimensional discriminant representing the probability of a jet corresponding to either b, c or light quark. The DL1 tagger is again based on multivariate analysis, but now it employs deep neural network (DNN) built using state-of-art packages like Keras interfaced with Theano backend [160]. Set of input variables is equivalent to that used in MV2c10, but enhanced with input form newly developed Soft Muon Tagger [161] and c tagging related variables from JF. The algorithm uses multiclassification, meaning that it treats all quark flavors

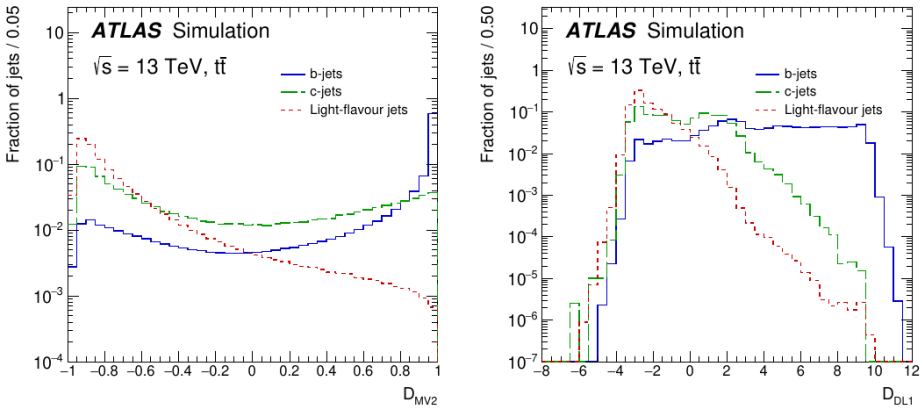


Fig. 4.8: Distribution of the BDT output for MV2c10 (left) and DNN output for DL1 (right) algorithm. In both cases simulated $t\bar{t}$ events were used for training. Figures are taken from [161].

equally by assigning separate output node for each of them. Thanks to this feature, it provides flexibility in choosing output thresholds to achieve desired tagging efficiency. Improved version of this algorithm that uses recurrent neural networks, called DL1r, is employed in this thesis. The DL1r offers significantly better rejection rate for c and light jets when compared to MV2c10, especially in high-p_{T} regions. The distribution of the DL1 discriminant can be found in Figure 4.8.

Similar as for other reconstruction algorithms described in this chapter, also flavor tagging describe its own operating points based on the b jet acceptance efficiency and complementary rejection rate for c and light jets. The aim of various operating points is thus to provide desired b jet acceptance while controlling amount of fake b jets originating from mistakenly classified c and light jets. The ATLAS physics analyses can choose from four operating points according to their specific needs. Operating points, which are common for all b tagging algorithms, are defined based on the following b jet efficiencies: 60% (corresponds to *VeryTight* operating point), 70% (*Tight*), 77% (*Medium*), 85% (*Loose*). The tagging efficiencies for the MV2 and DL1 algorithms corresponding to these operating points are summarized in Table 4.1. In addition to these fixed operating points, so-called pseudo-continuous b tagging (PCBT) constructs exclusive bins in the b tagging discriminant, which correspond to various b jet efficiencies. This approach allows for the use of different calibrated operating points for the jets present in the event or offers possibility to optimize precision of a measurement by constructing orthogonal combinations of selections with different b tagging operating points. In this dissertation, various b tagging operating points (fixed and also PCBT) were used, and will be specified in the next chapters.

ϵ_b	MV2			DL1		
	Selection	Rejection		Selection	Rejection	
		c jet	Light-flavor jet		c jet	Light-flavor jet
60%	> 0.94	4.3%	0.083%	> 2.74	3.7%	0.077%
70%	> 0.83	11.2%	0.33%	> 2.02	10.6%	0.26%
77%	> 0.64	20.4%	0.91%	> 1.45	20.4%	0.77%
85%	> 0.11	37.0%	4.0%	> 0.46	38.5%	3.4%

Table 4.1: Tagging efficiencies for b, c and light-flavor jets corresponding to different operating points for MV2 and DL1 taggers. "Selection" column represents the cut value applied to the respective discriminant. Values are taken from [161].

4.5 Missing transverse momentum

As discussed in Section 2.2, *Lepton+Jets* and *Dilepton* decay mode of $t\bar{t}$ pair results in one or two neutrinos in the final state. Since ATLAS is not designed to register neutrinos, they travel through detector material and their presence can be revealed only indirectly via imbalance in the measured transverse momentum [3]. Thanks to the conservation of total momentum in the event, the sum of four-momenta of all particles in the final state is equal to the net momentum of initial partons. As the consequence, the total longitudinal momentum can yield non-zero value, while the transverse net momentum has to be equal zero.

In ATLAS, the reconstruction of the missing transverse momentum results in a set of observables which are constructed from vectors of transverse momenta corresponding to different objects reconstructed in the detector. Most important observable, from which other variables are derived, is vector of missing transverse momentum defined as follows:

$$\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}} = \underbrace{-\sum_{\substack{\text{selected}\\\text{electrons}}} \mathbf{p}_{\mathrm{T}}^{e} - \sum_{\substack{\text{accepted}\\\text{photons}}} \mathbf{p}_{\mathrm{T}}^{\gamma} - \sum_{\substack{\text{accepted}\\\tau\text{-leptons}}} \mathbf{p}_{\mathrm{T}}^{\tau_{\mathrm{had}}} - \sum_{\substack{\text{selected}\\\text{muons}}} \mathbf{p}_{\mathrm{T}}^{\mu} - \sum_{\substack{\text{accepted}\\\text{jets}}} \mathbf{p}_{\mathrm{T}}^{\mathrm{jet}}}_{\text{hardterm}} \underbrace{- \sum_{\substack{\text{unused}\\\text{tracks}}} \mathbf{p}_{\mathrm{T}}^{\mathrm{track}}}_{\text{softterm}}. \tag{4.14}$$

The *hard term* accounts for fully reconstructed and calibrated particles, including electrons, photons, τ leptons, muons and jets, while *soft term* includes reconstructed tracks from the ID corresponding to charged particles associated with the primary vertex, but not identified with any object mentioned above.

Components of $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$, denoted as $(E_x^{\mathrm{miss}}, E_y^{\mathrm{miss}})$ are used to calculate magnitude of

[3] It should be noted that this imbalance is caused not only by neutrinos but also other particles which deposited part of their energy in various crack regions without sensitive detector elements.

this vector, usually called missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$, and corresponding azimuthal angle ϕ^{miss} representing its direction in transverse plane:

$$E_{\mathrm{T}}^{\mathrm{miss}} = |\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}| = \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2}, \tag{4.15}$$

$$\phi^{\mathrm{miss}} = \tan^{-1}\frac{E_y^{\mathrm{miss}}}{E_x^{\mathrm{miss}}}. \tag{4.16}$$

Since reconstruction of each object entering above formula for $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ is done independently, procedure called *signal ambiguity resolution* prevents from double counting of the same signal, i.e. when the same signal from calorimeter is used to reconstruct both electron and jet. More detailed description of the missing transverse momentum reconstruction in ATLAS can be found i.e. in [162].

4.6 Object overlap removal

After all the objects are identified and reconstructed, overlap removal procedure needs to be applied in order to avoid any possible double-counting of physics objects. Specific sequence of operations is designed to resolve these ambiguities based on track and cluster overlaps, and ΔR (or Δy) separation between reconstructed objects. The overlap removal thus proceeds with the steps in the following order:

- An electron sharing the track or overlapping cluster with other electron is removed.
- A muon reconstructed in calorimeter sharing track with electron is discarded.
- Any subsequent electron that share a track with a muon is dropped.
- A photon found in cone of radius $\Delta R = 0.4$ around an electron or a muon is removed.
- Any jet present in cone of radius $\Delta R = 0.2$ around an electron is excluded.
- An electron found within ΔR of 0.4 of a jet is dropped.
- A jet with fewer than 3 associated tracks found within ΔR of 0.2 of a muon is removed.
- A jet with fewer than 3 associated tracks, which has a muon ID track ghost-associated to it, is excluded.
- Any muon found in cone of radius $\Delta R = 0.4$ around jet is removed.
- A jet, which is found within $\Delta R = 0.4$ of a photon, is discarded.

Chapter 5

Data and Monte Carlo samples

Every ATLAS measurement, which tests compatibility of the physics phenomena hidden in observed data with the SM predictions, relies on an accurate modeling of physical processes, as well as response of the ATLAS detector. The tools, which are usually responsible for these simulations are called Monte Carlo (MC) generators and use pseudo-random numbers to generate the desired distributions corresponding to the underlying physical process.

As mentioned at the end of Section 2.4, two separate $t\bar{t}Z$ analyses are presented in this thesis. First is dealing with differential cross section measurement in trilepton and tetralepton channel, while second one, which is currently ongoing, is focused on inclusive cross section measurement in dilepton decay channel. Both analyses use the same ATLAS data, which are specified in the next section, while the simulated MC samples differ for some processes and will be explicitly specified in the rest of this chapter.

5.1 Data sample

The dataset used in both inclusive and differential cross section measurements was collected during years 2015 - 2018 and hence corresponds to the full Run 2 of the LHC operation. Each year of data taking constitutes one subset of data, which consists of several individual runs when the LHC delivered stable beams. Furthermore, each of these runs contains so-called *Luminosity Blocks* (LB), where each LB represents one minute of data taking.

Depending on the analysis and its specific needs, certain criteria are applied to collected data. Since reconstruction of the physics objects necessary for the analyses discussed in this thesis requires all ATLAS subdetectors to be in the fully operational state, not all LBs fulfill the quality criteria. The total integrated luminosities corresponding to selected LBs, stored at so-called *Good Run Lists* (GRL), are 3.22 fb^{-1} (2015),

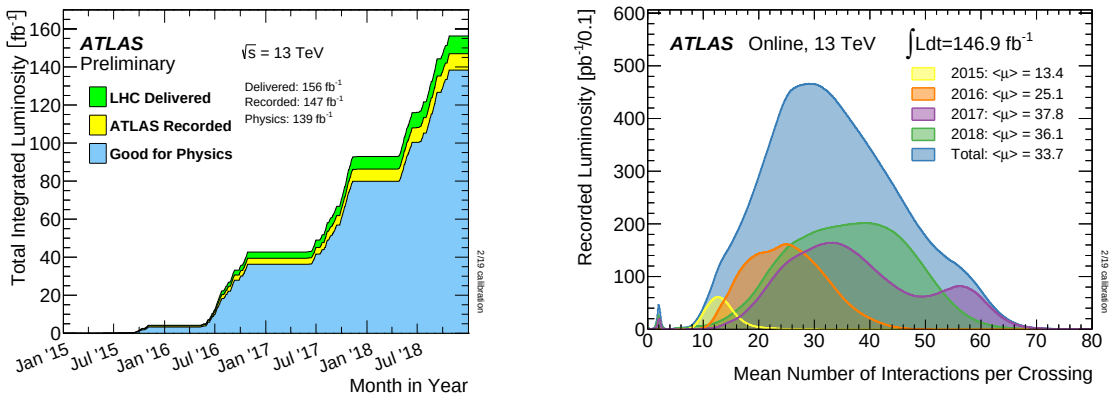


Fig. 5.1: On the left, distribution of cumulative luminosity delivered by the LHC (green), recorded by ATLAS (yellow) and certified to be fulfill "good" quality criteria (blue) over years 2015-2018. On the right, distribution of the mean number of pp interactions per bunch crossing between years 2015 and 2018. Figures are taken from [163].

32.99 fb^{-1} (2016), 44.3 fb^{-1} (2017) and 58.5 fb^{-1} (2018), summing up to 139 fb^{-1} available for the $t\bar{t}Z$ analyses. The increase in total integrated luminostity during full Run 2 of data taking is depicted in Figure 5.1.

Important quantity, which significantly affects event reconstruction, is the average number of pp interactions per bunch crossing $\langle\mu\rangle$, referred to as pileup. This number greatly varies over years as can be seen from Figure 5.1. The mean number of pileup interactions over whole Run 2 was at 34.2, but it peaked at around 13 in year 2015, while in year 2017 $\langle\mu\rangle$ was at around 38 with maximum reaching up to ≈80.

5.2 Monte Carlo samples

MC generators, which are used to simulate signal as well as background physical processes, consists of several successive procedures which are necessary to properly simulate all underlying physics aspects of the event created in the ATLAS detector. Following steps, which are schematically depicted also in Figure 5.2, need to be simulated in order to produce complete event.

5.2.1 Parton Distribution Functions

When the two protons in the ATLAS detector undergo inelastic collision, in fact only one parton in each proton participates in the interaction. It is therefore crucial for the proper MC simulation to determine the probability of a particular parton, which carries certain momentum fraction of proton, to take part in the hard-scattering process. Unfortunately, since non-perturbative nature of QCD does not allow analytical

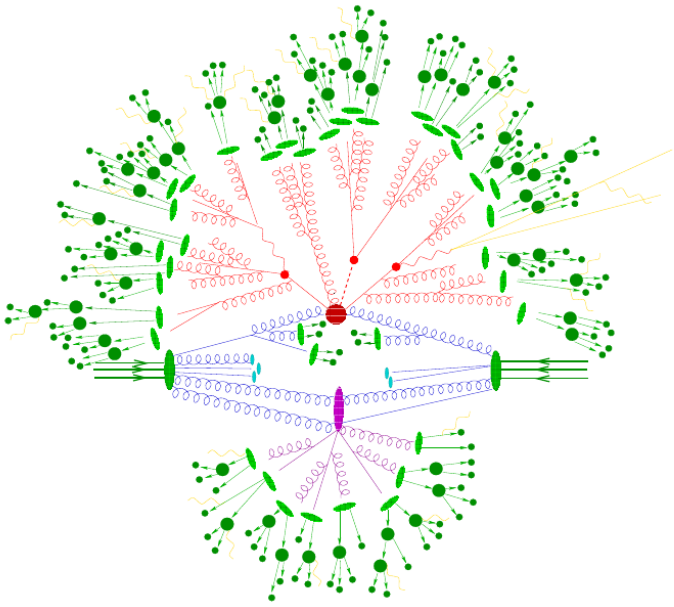


Fig. 5.2: Diagram showing various parts of the event simulation that are modeled by MC generators. Only one *pp* collision is considered in this representation, hence the pileup interactions, which are inevitably present in real collisions are not included here. The red blob in the center, representing hard-scattering process, is surrounded by red tree-like structure corresponding to the parton shower (initial and final state radiations). Purple blob depict secondary hard scattering, called also underlying event. Light green blobs represent hadronization process and dark green blobs hadron decays. Yellow lines correspond to soft photon emissions. Figure is taken from [164].

calculation of these probability distributions at low energies, they need to be measured experimentally.

However, once these probabilities, called *parton distribution functions* (PDF), are measured for a specific energy, their evolution in given energy range can be determined by so-called DGLAP equations. These equations, independently formulated by Dokshitzer [165], Gribov and Lipatov [166], and 5 years later by Altarelli and Parisi [167], separately predict PDFs for quarks (of different flavors) and gluons. Experimentally used PDFs are obtained from combination of measurements conducted by various experiments at different energy scales.

5.2.2 Matrix Elements

Calculation of the matrix elements (ME) corresponding to the given process is essential for determination of the partonic cross section $\hat{\sigma}_{i,j\to X}$. The relation between ME and

$\hat{\sigma}_{i,j\to X}$ is given by the following equation:

$$\hat{\sigma}_{i,j\to X} = \sum_{k=0}^{\infty} \int d\Phi_{X+k} | \sum_{l=0}^{\infty} \mathcal{M}^{l}_{X+k}|^2, \tag{5.1}$$

where k represents number of non-virtual emissions (additional legs in Feynman diagrams), Φ_{X+k} is corresponding phase space for $X + k$ final state, $\mathcal{M}^{l}_{X+k}$ denotes ME amplitude of process $ij \to X + k$ with l loop corrections. Since calculation with k and l going to ∞ are in most cases impossible to do, only limited number of legs and loops, that define order of the calculation, are used. Leading order (LO) calculations then refer to cases where $k = l = 0$, while higher orders, NLO and NNLO corresponds to $l = 1, k = 0$ and $l = 2, k = 0$, respectively.

In general, higher loop corrections use to cause infrared divergences, but these effectively cancel out when taking into account k real emissions. However, configurations when $k > l$ lead to soft and collinear singularities in QCD. These need to be eliminated by restricting full phase space to only its fiducial part where partons are enough separated and have large momentum.

The theoretically calculated cross section of the given process consider production of final-state particles from initial-state partons (denoted by i, j in Eq. 5.1), but in case of proton beams the two collided protons are composite particles consisting of partons. For this reason, the cross section determined by SM theory, $\sigma_{i,j\to X}$, needs to be related to experimentally observed cross section, $\sigma_{p,p\to X}$. This relation is called *factorization theorem* and can be expressed as follows:

$$\sigma_{p,p\to X} = \sum_{i,j} \int \int f_i(x_i, \mu_F) f_j(x_j, \mu_F) \sigma_{i,j\to X}(x_i, x_j, \mu_R, \mu_F) dx_i dx_j, \tag{5.2}$$

where $f_i(x_i, \mu_F)$ represents PDF of the parton of type i at *factorization* scale μ_F carrying fraction x_i of the proton momentum, and the summation goes over all types of partons present in the proton. The free parameters, namely *renormalization* (μ_R) and factorization (μ_F) scales arise as a consequence of the partial calculations described in the previous text (e.g. in $t\bar{t}$ cross section calculated to order NNLO+NNLL), and need to be fixed to some physically motivated values. Since there is still some subjectivity in choosing these scales, they are typically varied by factor 0.5 and 2 of nominal value, and later used as a source of theoretical uncertainty in the analysis.

5.2.3 Parton Shower

As recently mentioned, contributions of the higher orders can result in singularities when using analytical calculations and thus only the regions of the full phase space, which do not cause these singularities, are used in the ME simulation. Therefore these gaps in a phase space need to be filled, what is exactly the aim of the parton

shower (PS) generators. In this process, the event topology is completed by simulating successive emission of gluons and quarks (referred to as quark or gluon splitting) from hard-scatter partons in either initial or final state.

The affected parton undergoing gluon splitting (or final state radiation in case of quark) naturally looses its energy and transfer it to the emitted particle. This intial state radiation (ISR) and final state radiation (FSR) thus continues until the energy scale is lower than so-called *hadronization scale* $Q_0 \approx 1$ GeV, when hadronization simulation (discussed next) takes over.

Since the ME simulations often directly include higher orders (NLO, NNLO), the additional jet emissions need to be "matched" (decided if they are generated by the ME or PS simulation) to the PS simulations in order to avoid double counting in common regions of the phase space. After identification of these phase space regions, the matching algorithm typically prioritize ME emissions with high momentum and large separation from other particles, while PS generator takes care only of configurations where soft and collinear singularities could arise. Many ME-PS matching algorithms use so-called *matching scale* to determine if the emission is accepted or rejected. ME emissions are accepted if they exceed this scale, while PS generator only allow emissions below the scale. The transition region between the two regimes is then smoothed by assigning specific weights to events.

Most frequently used ME-PS matching procedures, which use techniques outlined above, are Catani Krauss Kuhn Webber (CKKW) [168] and Michelangelo L. Mangano (MLM) [169] algorithms.

It should be noted that there exist two types of MC generators. First type represents ME generators that are designed only to perform computations at the ME level and need to be interfaced to other MC generators for simulation of PS and hadronization. The second type, general-purpose MC generators, are capable of full event simulation from calculation of the ME level to hadronization. Examples of ME generators are MadGraph [170] or Powheg [171], while Pythia [172], Herwig [173] or Sherpa [174] represent general-purpose generators.

5.2.4 Hadronization

Below the hadronization scale of 1 GeV, process of emitting quarks or gluons enters non-perturbative QCD regime, and the final part of event generation is thus governed by phenomenological models. Thus in the hadronization step only local formation of hadrons need to be coordinated and no redistribution of momentum, flavor and color of particles, is necessary. Two hadronization models, widely used in modern MC generators, are called *string* and *cluster model*.

The string model [175, 176] assumes virtual gluonic string acting between two par-

ticles that corresponds to the potential energy proportional to the distance between them. The energy stored in the string increases with factor of ≈ 1 GeV/fm as particles propagate through space. Once energy of the gluonic string exceeds value corresponding to particular quark-antiquark pair, the antiquark (or quark in case of original antiquark) is produced, and the string is splitted into two with lower potential energy. This process is repeated until all potential energy available in the string is spend for creation of hadrons.

The cluster model [177, 178] approach, on the other hand, is based on the idea of the color confinement, which allows existence of only colorless states, referred to as clusters. These confined clusters are assigned mass sampled from the distribution peaking at 1 GeV with the mean value of 3 GeV. This procedure, which is independent of hard-scattering process, continues with splitting of the initial cluster into smaller ones until their mass is large enough to produce stable hadrons.

5.2.5 Detector simulation

The hadronization process concludes the simulation of an event from the physics point of view, resulting in the momentum four-vectors assigned to each stable particle [1], and such a topology is referred to as *particle level* [2]. However, the real data events measured in the ATLAS detector are further affected by the interactions between particles and the material of the detector. To ensure fair comparison between real and simulated data, the particle level events need to be propagated through the simulation of the ATLAS detector, and applied the same reconstruction algorithms as used for the real data.

This so-called *reconstruction level*, is usually simulated by GEANT4 package [179, 180, 181] that offers full simulation of all ATLAS subdetectors. Although the *full simulation* (FULLSIM) yields most reliable results, it is computationally very demanding and is thus often replaced by faster and not so CPU intensive approximation, called ATLFAST-II (AF-II). This version still uses GEANT4 simulation for some subdetectors, but employs faster alternatives for specific detector elements, like calorimeter system (ATLAS Fast Calorimeter Simulation - FASTCALOSIM) [182] or tracking devices (ATLAS Fast Tracking System - FATRAS) [183]. The analyses discussed in this dissertation make use of both FULLSIM and AF-II samples.

[1] In the ATLAS detector, particle is considered to be stable if its mean lifetime is higher than 0.3×10^{-10} s.

[2] There exist one more generation stage, called *parton level*, which is frequently used in the ATLAS analyses. This level refers to topology of an event before the decay of final state partons emerging from the hard-scattering but after any QCD emissions in the initial or final state.

5.2.6 Signal $t\bar{t}Z$ samples

The $t\bar{t}Z$ signal process is modeled by the MadGraph5_aMC@NLO [184] generator, which provides a ME calculation at NLO in QCD, with the NLO PDF set `NNPDF3.0NLO` [185]. The $t\bar{t}\gamma^*$ contribution and the Z/γ^* interference is included, such that they feature events with dilepton invariant masses $m_{\ell\ell} > 5$ GeV. When generating these samples, the top-quark mass is set to its nominal value of 172.5 GeV and top-quark decays are simulated at LO using package called MadSpin [186, 187] to preserve spin correlations. The events at ME level are interfaced with Pythia 8 [172] for applying parton showering and hadronization. Specific tune of generator parameters, called `A14` tune [188], is used together with the PDF set `NNPDF2.3LO` at LO. The decays of b and c hadrons are simulated using the EvtGen generator [189].

The theoretical value for the $t\bar{t}Z$ cross section, considering all decay modes of Z boson, including also contribution from $t\bar{t}\gamma^*$ and corresponding Z/γ^* interference, is $\sigma(t\bar{t}Z) = 0.88^{+0.09}_{-0.10}$ pb where the uncertainties are from QCD scale variations, PDF and α_S. This is based on the calculation documented in [190], which predicts $\sigma(t\bar{t}Z) = 0.84^{+0.09}_{-0.10}$ pb and includes QCD and EW corrections at NLO, plus a correction to account for off-shell effects [191].

In the differential cross section analysis in trilepton and fourlepton channel, the results are compared with theoretical predictions obtained from alternative MC generators. Additional $t\bar{t}Z$ samples are thus simulated with Sherpa 2.2.1 [192] using NLO accuracy, while considering both inclusive and multi-leg setups. The same NLO PDF set as for nominal sample is used also for Sherpa sample and, since Sherpa is general-purpose generator, PS is included in the simulation. Another alternative sample is obtained with the same versions of MadGraph5_aMC@NLO and EvtGen generators as for the nominal one, but using a different MC generator for the modeling of the parton-shower and hadronization, namely Herwig 7 [193, 173] instead of Pythia 8. Additionally, to account for the uncertainty related to the chosen set of tuning parameters which are sensitive to ISR, two samples with nominal setup, but using different set of variations in the `A14` tune (called `Var3c` variations [188]), are employed.

5.2.7 Background samples

Following processes, that can resemble $t\bar{t}Z$ topology and thus contaminate the dataset used in the analyses, are considered in the inclusive and/or differential cross section measurements presented in this thesis. The summary of the simulation configurations that were used for particular processes can be found in Table 5.1.

tZq and tWZ

Associated production of the single top quark with Z boson and additional parton (tZq) is simulated with MadGraph5_aMC@NLO at NLO in QCD and employing NNLO PDF set `NNPDF3.0NNLO`. PS is simulated by Pythia 8 with `A14` tune and LO `NNPDF2.3LO` PDF set. Z boson in these samples is allowed to decay into dileptonic pair with invariant mass down to 30 GeV.

The same generators and PDF sets are used for simulation of single top quark production associated with W and Z boson (tWZ). The interference between signal process and tWZ was removed with the so-called DR1 diagram removal procedure [194].

WZ+jets and ZZ+jets

Simulation of the diboson production decaying into three or four charged leptons associated with additional jets, which is important background for trilepton and tetralepton $t\bar{t}Z$ channel, is conducted with Sherpa 2.2.2 generator using NNLO PDFs `NNPDF3.0NNLO`. Similar to alternative Sherpa signal samples, multiple MEs were matched to the Sherpa PS using the so-called MEPS@NLO prescription [195]. OpenLoops package [196] provided QCD corrections at NLO. ZZ/WZ+jets events with no more than one additional parton in the final state were modeled at NLO, while events with two or three additional partons were modeled only at LO precision.

Z+jets

The production of Z boson associated with additional jets, which is relevant mostly for the dilepton channel, is simulated with Sherpa 2.2.1 generator using ME with NLO precision for up to two jets and with LO precision for three and four jets in the final state. Matching to the PS that is provided by Sherpa is done again with MEPS@NLO prescription. PDF set at NNLO precision `NNPDF3.0NNLO` is employed and the samples are normalized to NNLO prediction.

$t\bar{t}$

$t\bar{t}$ pair production, where both top quarks decay leptonically, can mimic topology in the dilepton $t\bar{t}Z$ channel. These processes are simulated with PowhegBox2 generator [197] at NLO precision using `NNPDF3.0NLO` PDF set. The produced events are interfaced with Pythia for the PS using `A14` tune and `NNPDF2.3LO` PDF set. As for all other top-related samples, the mass of the top quark is set to 172.5 GeV. The top quark decays are further modeled at LO using MadSpin, while decays of bottom and charm hadrons are simulated using the EvtGen.

Other processes

This category contains rare processes which form a subdominant contribution to the background composition of $t\bar{t}Z$ process.

Events featuring the production of $t\bar{t}$ pair associated with either W ($t\bar{t}W$) or H boson ($t\bar{t}H$) are simulated using MadGraph5_aMC@NLO NLO MEs with `NNPDF3.0NLO` PDF set. Parton showering is provided by Pythia 8 using `A14` set of tuning parameters.

LO precision is used in generation of samples featuring Higgs boson production with either W or Z boson, which is done by Pythia 8 employing `NNPDF2.3LO` PDF set and `A14` tune.

Samples featuring three- and four-top quark production ($t\bar{t}t$ or $t\bar{t}t\bar{t}$) are produced using MadGraph generator at LO precision that is interfaced to Pythia 8 for PS and hadronization. PDF set that is used for these samples is of LO precision, `NNPDF2.3LO`, and again the generator is set up with `A14` tune. Exactly the same setup is used to generate also $t\bar{t}WW$ process.

Finally, the leptonic decays of triboson processes (WWW, WWZ, WZZ, ZZZ) which can produce up to six leptons, are modeled at NLO precision using Sherpa 2.2.2 with `NNPDF3.0NNLO` PDF set. In case of additional partons (up to three) present in the final state, the precision is reduced to LO.

Process	Generator	Cross section calc.	Parton shower	PDF	MC tune
$t\bar{t}Z$, $Z \to \ell\ell$	MadGraph5_aMC@NLO	NLO	Pythia 8	`NNPDF3.0NLO`	`A14`
	MadGraph5_aMC@NLO 2.3.3	NLO	Herwig 7	`NNPDF3.0NLO`	`H7-UE`
	Sherpa 2.2.1	NLO	Sherpa	`NNPDF3.0NNLO`	Sherpa
$t\bar{t}$	PowhegBox2	NLO	Pythia 8	`NNPDF2.3LO`	`A14`
Z+jets	Sherpa 2.2.1	NLO	Sherpa	`NNPDF3.0NNLO`	Sherpa
$t\bar{t}W$	MadGraph5_aMC@NLO	NLO	Pythia 8	`NNPDF3.0NLO`	`A14`
tZq	MadGraph5_aMC@NLO	NLO	Pythia 8	`NNPDF3.0NNLO`	`A14`
tWZ	MadGraph5_aMC@NLO	NLO	Pythia 8	`NNPDF3.0NNLO`	`A14`
$t\bar{t}H$	MadGraph5_aMC@NLO	NLO	Pythia 8	`NNPDF3.0NLO`	`A14`
$t\bar{t}WW$	MadGraph	LO	Pythia 8	`NNPDF2.3LO`	`A14`
$t\bar{t}t$, $t\bar{t}t\bar{t}$	MadGraph	LO	Pythia 8	`NNPDF2.3LO`	`A14`
$WZ \to \ell\ell\ell\nu$	Sherpa 2.2.2	NLO	Sherpa	`NNPDF3.0NNLO`	Sherpa
$ZZ \to \ell\ell\ell\ell$	Sherpa 2.2.2	NLO	Sherpa	`NNPDF3.0NNLO`	Sherpa
$VVV(V = W/Z)$	Sherpa 2.2.2	NLO	Sherpa	`NNPDF3.0NNLO`	Sherpa
$H \to \ell\ell\ell\ell$	Sherpa 2.2.2	NLO	Sherpa	`NNPDF3.0NNLO`	Sherpa
$H + W/Z$	Pythia 8	LO	Pythia 8	`NNPDF2.3LO`	`A14`

Table 5.1: List of Monte Carlo event generators and their settings used to predict the contributions from SM processes featuring prompt production of two, three or four leptons (electrons or muons).

5.3 Data-driven fake lepton background

Non-prompt leptons, usually called *fake leptons* or *fakes*, represent objects that are unintentionally identified as leptons. They can have physical origin, meaning that they are in fact leptons, but do not come from hard-scatter process or from decays of particles arising from hard-scattering. This category typically contains leptons from meson decays or electrons/positrons from photon conversions. On the other hand, reconstruction and identification algorithms can incorrectly classify signatures of light jets as belonging to leptons. In $t\bar{t}Z$ analysis featuring trilepton and tetralepton final state, majority of fake leptons arise from heavy-flavor decays associated by leptonic decays of W or Z boson.

Estimation of the fake lepton contribution is very challenging since the probability to misidentify jets or photons as leptons is very small, but cross section of background processes featuring multiple jets is very high, meaning that precise estimation of the fake lepton contribution would require large dataset. Therefore if the expected contribution from the fake leptons is high, then the fake yield is usually estimated using a data-driven procedure. Since trilepton and tetralepton channels feature one or two leptons from $t\bar{t}$ decay, these can originate from b hadrons and hence contribute to the fake lepton yield. For this reason, the differential analysis employs data-driven technique to estimate fake contribution and its associated uncertainty. Dilepton channel, thanks to its selection criteria described in Section 8.1 and featuring only leptons from Z boson decay which are easily identifiable, is not affected by large fake lepton contamination and thus the fakes are estimated directly from MC while conservative 50% normalization uncertainty is applied in the inclusive cross section measurement.

The so-called *Matrix Method* (MM) [198, 199], which is used also in the differential cross section measurements, is based on a different response of prompt (real) and fake leptons to identification and isolation techniques, as well as to requirements related to impact parameters. As an input for the method, data events are selected in the same way as for the signal regions, but requirements on the quality of electrons and muons, as defined in Sections 4.2 and 4.3, are relaxed. This selection with looser criteria define *loose* leptons, while leptons passing original selection criteria are referred to as *tight* leptons.

The goal of the MM is to estimate number of fake leptons in the tight sample N_{fake}^{tight}. Number of data events passing loose and tight selection can be defined as follows:

$$\begin{aligned} N^{loose} &= N_{real}^{loose} + N_{fake}^{loose}, \\ N^{tight} &= N_{real}^{tight} + N_{fake}^{tight}, \end{aligned} \tag{5.3}$$

where N_{real}^{loose} stands for number of events passing loose selection and corresponding to prompt leptons, while N_{fake}^{loose} denotes number of events fulfilling loose criteria, but

featuring fake leptons, and analogically for tight selection. If the event passes tight selection it automatically passes also looser criteria, thus the selection efficiency for real (ϵ_{real}) and fake (ϵ_{fake}) contribution can be defined as:

$$\begin{aligned} \epsilon_{real} &= \frac{N_{real}^{tight}}{N_{real}^{loose}}, \\ \epsilon_{fake} &= \frac{N_{fake}^{tight}}{N_{fake}^{loose}}, \end{aligned} \tag{5.4}$$

Combining above equations and extracting N_{fake}^{tight} yields:

$$N_{fake}^{tight} = \frac{\epsilon_{fake}}{\epsilon_{real} - \epsilon_{fake}} \left(\epsilon_{real} N^{loose} - N^{tight} \right) \tag{5.5}$$

Improved version of the MM, referred to as *Likelihood Method* (LM) [200], evaluates the number of fake leptons passing tight selection by introducing likelihood function. Contrary to the MM, this approach ensures non-negative estimate of the fake lepton yields, what is relevant especially for the $t\bar{t}Z$ analyses featuring low number of events, where the MM could result in negative fake yields. The likelihood function is constructed as a product of Poisson probability functions. While the observed number of loose leptons N^{loose}, as well as real and fake efficiencies are fixed, the expectation values of the Poisson functions, N_{fake}^{tight}, are obtained from maximization of the likelihood. Since the total number of data events passing loose selection is known, the only observables that need to be evaluated are real and fake lepton efficiencies. These are measured in dedicated control regions enriched in real and fake leptons. Definitions of the regions for the fake efficiency evaluation are based on the requirements of two equally charged leptons, and can be found in Table 5.2. Both fake and real efficiencies are measured separately for electrons and muons, and are parametrized in p_{T} and $|\eta|$ of the respective lepton, as shown in Table 5.3. For the differential analysis, the fake lepton contributions are estimated separately for each bin of the particular variable used for differential cross section measurement.

Region	N_ℓ ($\ell = e, \mu$)	N_{jets}	$N_{b-\mathrm{jets}}$
e-fakes-CR	$= 2$ ($e^\pm e^\pm$)	≥ 1	≥ 1
μ-fakes-CR	$= 2$ ($\mu^\pm \mu^\pm$)	≥ 1	≥ 1

Table 5.2: Definition of the fake-lepton control regions used for the electron (e-fakes-CR) and muon (μ-fakes-CR) fake efficiency measurements.

Bin ranges	p_{T} [GeV]	$\lvert\eta\rvert$
Muons	[0, 7, 12, 20, 35, 50, ∞]	[0, 0.5, 1, 1.5, 2, 2.5]
Electrons	[0, 7, 12, 20, 35, 50, ∞]	[0, 0.7, 1.37, 1.52, 2, 2.47]

Table 5.3: p_{T} and $|\eta|$ bin ranges chosen for the fake and real efficiency measurements. The symbol "∞" stands for an inclusive last bin. Whereas the p_{T} binning is chosen to be the same for electrons and muons, the $|\eta|$ binning for electrons is not equidistant to have a separate measurement for electrons from the LAr crack region.

Chapter 6

Systematic uncertainties

Final uncertainty on the measured inclusive and differential cross section does not depend only on the amount of the available data (statistical component), but includes also sources related to either detector response and particular reconstruction algorithms, or specific setup used for the generation of the MC samples. These systematic sources, divided into experimental (detector-related) and theoretical (modeling-related) categories, are presented in this chapter.

The treatment of systematic uncertainties in the context of the differential measurements are for the most part analogous to the general prescriptions presented here, but additional details are required and will be presented later in Section 10.3.1.

6.1 Experimental uncertainties

Since the detector response affects all physical objects, algorithms used for their reconstruction, identification or isolation often require calibrations that are accompanied by systematic uncertainties. The specific sources of these uncertainties related to the physics objects relevant for $t\bar{t}Z$ analysis are presented in next sections.

6.1.1 Charged leptons

Systematic uncertainties related to efficiency of the electron and muon reconstruction, identification and isolation algorithms, as described in Sections 4.2 and 4.3, are derived from $Z \to l^+l^-$ and $J/\Psi \to l^+l^-$ events. The efficiencies of the used algorithms and their corresponding uncertainties are determined by so-called *tag and probe* method [201, 142]. This technique is based on the usage of unbiased events corresponding to the leptonic Z boson decays, which are selected according to a tag trigger and subsequently subjected to a probe trigger with predefined selection criteria. The resulting fraction of events passing the probe trigger selection is used to evaluate the efficiency. In case of both electron and muon efficiency measurement, one up and down systematic variation

is associated separately with the reconstruction, identification and isolation. Additional scale factors are derived for muons with p_{T} lower than 15 GeV. These systematics aim to correct for discrepancies between data and MC, and thus are applied for all simulated samples.

Differences observed in reconstruction of lepton energy between measured and simulated data are addressed by scale and resolution corrections. Consistent energy reconstruction in the data and MC is crucial for selection criteria defining analysis regions, which are based on transverse momentum thresholds. Both momentum scale and resolution are checked in reconstructed distributions of dilepton invariant masses in $Z \to l^+l^-$ and $J/\Psi \to l^+l^-$ events. For electrons also events featuring $W \to e\nu$ are considered. While the energy scale is applied to data in order to improve consistency between real and measured energy, the resolution correction factor is applied only to MC samples to ensure agreement with data events [202]. Additionally, the charge-dependent scale corrections, referred to as *sagitta bias*, are applied separately to μ^+ and μ^- in data. Moreover, matching between tracks and vertices for muons also adds one systematic source [142].

6.1.2 Jet energy scale and resolution

The jet reconstruction process, described in Section 4.4, is concluded by several calibrations and corrections [147, 146]. The first correction is applied during the jet reconstruction and is performed by recomputing (event-by-event) η and ϕ coordinates of the topo-clusters, which are subsequently used for constructing four-vectors relative to the primary vertex.

In the next step, correction for an energy scale is applied for the topological clusters. Although the measured energy of clusters is calibrated at the electromagnetic scale, what is correct for energies measured in calorimeters, the PFlow jets require replacing of the topo-cluster energies with momentum of the tracks corresponding to the given jet. This subtraction is often inaccurate because the particles are usually collimated in the jet center, resulting in the additional correction factor.

Next correction is connected with pileup effects, where number of pileup vertices in the event is used to determine average energy carried by the pileup jets which is subsequently subtracted from the overall energy corresponding to the affected jet. Moreover, additional correction factor is applied as a function of number of bunch crossings in order to separately account for in-time and out-of-time pileup components.

Absolute calibrations, which are derived from dijet events modeled with Pythia8 using anti-k_T algorithm with $R = 0.4$, are applied to fully reconstructed jets. The scale factors for transition between generated and reconstructed jet energy and pseudorapidity η are derived from the simulation of the detector response. This is done by

geometrically matching truth-level jets with reconstruction level within a cone of radius $\Delta R = 0.3$. The obtained scale factors, *jet energy scale* (JES), and η intercalibration, are then applied to reconstructed data.

Various fluctuations in the jet response can be caused by jet flavor and composition of particles constituting the jet. Furthermore, the type of particle initiating the jet, if it is either gluon or quark, significantly affects the jet response. The jets initiated by gluons typically have more constituents (more tracks in the ID) and larger radiation spread, than jets initiated by quarks.

The last, data-driven based calibration, is determined by comparing the jet properties with other objects that are typically well modeled. For example η of the non-central jets is calibrated against precisely measured central jets from dijet events. Analogously, calibrations of the transverse momenta are derived from events featuring associated production of Z/γ with jets, or multijet processes for high-p_{T} jets.

All of the above-mentioned calibrations and corrections are associated with their corresponding uncertainties. The JES calibration and its uncertainty have been derived from the combination of test-beam data, MC simulations and LHC collision data [203]. In total there is 80 systematic sources related to JES calibration procedure, but only about 30 of them with the largest impact are used in the measurements presented in this thesis. Remaining sources with lower impact are summed in quadrature and applied as one additional systematic uncertainty.

Another source of systematic uncertainty arises due to the fluctuations in *jet energy resolution* (JER). According to the Eq. 3.7, the energy resolution of the calorimeter can be parametrized with jet p_{T}. The JER has been determined separately for MC simulations and data using techniques described in Ref. [203, 204]. Resulting fractional energy resolution for a particular jet is given as a function of its p_{T} and η. A systematic uncertainty associated to this measurement is obtained as a quadratic difference between JER observed in data and MC. Since JER in MC simulation is determined by smearing the jet energies, hence worsening the energy resolution, it cannot be estimated for better resolution and thus the resulting systematic uncertainty is one sided and has to be symmetrized. The complete set of systematic uncertainties associated with JER consists of 9 source used for the differential, and 13 sources for inclusive cross section measurement.

6.1.3 Jet vertex tagger

Discrepancies in JVT efficiencies for simulated and observed jets are assessed by single up and down uncertainty. The differences are determined from variation of tag-and-probe scale factors which are applied to simulated $Z(\rightarrow \mu^+\mu^-)$+jets events [152]. The overall systematic variation includes contamination caused by pileup jets after applying

pileup suppression procedure, and also uncertainty caused by specific choice of MC generator used for production of the Z+jets sample.

6.1.4 Missing transverse energy

As discussed in Section 4.5, the evaluation of missing transverse momentum consists of two components, hard term representing all reconstructed objects in the detector (electrons, photons, τ leptons, muons and jets), and soft term corresponding to tracks without any assigned object. The systematic uncertainties related to $E_{\mathrm{T}}^{\mathrm{miss}}$ arise due to scale and resolution discrepancies for both the reconstructed physics objects, as well as for unassigned tracks. They are derived using simulated $Z \to \mu^+\mu^-$ events that are compared to data in order to evaluate agreement of the p_{T} balance between hard and soft components [162].

6.1.5 Flavor tagging

Algorithms for b, c and light flavor tagging, described in Section 4.4.4, bring several sources of systematic uncertainties. To account for the differences between tagging performance on data and MC simulations, tagged jets are further calibrated after their reconstruction. These scale factors are applied to MC events and are derived from studying well known processes in data (dileptonic $t\bar{t}$ for b jets, semileptonic $t\bar{t}$ for c jets and multijet processes for light jets). The parametrization is done usually in jet p_{T} and η, and separately for each jet flavor.

Total uncertainty on the tagging scale factors consists of statistical uncertainty in data that are used for their evaluation, statistical uncertainty in MC used for the efficiency calculations, and also systematic uncertainties related to jet reconstruction as mentioned previously. All these uncertainties strongly depend on the p_{T} of the given jet and range from 1 to 5% for high- and low-p_{T} regions, respectively. The MC statistical component, however, reaches much lower values, typically below one percent. In addition to these sources, the modeling systematics affecting used MC samples, are also considered by varying relevant model parameters and determining difference between nominal and adjusted MC samples on a bin-by-bin basis [205]. In total 45 systematic uncertainties are considered for b tagging efficiency, 20 for c mis-tagging and 20 for light-flavor mis-tagging efficiencies in both differential and inclusive cross section analysis. The high number of systematic sources is caused by the usage of PCBT, in which every calibrated operating point is assigned separate set of systematic uncertainties.

6.1.6 Pileup reweighting

In order to account for the differences in the pileup profile between data and MC, dedicated scale factors are applied to simulated samples. These are determined from the simulated minimum-bias events added at the top of events emerging from hard scatter. These additional events are subsequently reweighted according to pileup distribution observed in the LHC data [206]. The related systematic uncertainties, applied to all MC samples, are obtained by rescaling $\langle\mu\rangle$ - the average number of pp interaction per bunch crossing, by the factors corresponding to uncertainties on the reweighting factor.

6.1.7 Luminosity

The luminosity values quoted in Section 5.1, which directly enter the cross section measurements, are measured with the LUCID detector (see Section 3.4.7 for more details). The luminosity measurements are thus associated with the corresponding overall uncertainty yielding 1.7% ($\pm$2.4 fb^{-1}), which is derived following a methodology described in Ref. [207]. This normalization uncertainty is applied as symmetric up and down variation to all MC samples considered for particular analysis.

6.2 Theoretical uncertainties

Systematic sources presented in this section are related to MC modeling of signal and backgorund samples, and include uncertainty on the cross section, choice of PDFs used for the generation, MEs and parameters of parton showering procedure. Moreover, various scales (i.e. μ_R, μ_F), which are chosen with some degree of subjectivity, need to be varied as well.

6.2.1 Signal modeling uncertainties

As already mentioned at the end of Section 5.2.2, the renormalization (μ_R) and factorization (μ_F) scales are varied by factors 0.5 (down variation) and 2.0 (up variation) with respect to their default values. Both individual and simultaneous shifts are considered in the differential measurements, while inclusive analysis employs only simultaneous variation of both scales at the same time [1].

Uncertainties related to the PDF are evaluated following the procedure described by PDF4LHC prescription [208]. This recommendation combines various PDF sets, namely `CT14` [209], `MMHT14` [210] and PDF set used for generating signal samples, `NNPDF3.0` [89]. The resulting set of 100 systematic weights, which are averaged to form

[1] It was shown that simultaneous variation covered individual shifts and thus is considered in order not to overestimate this systematic source.

one up/down variation, includes variations of the $\alpha_S(m_Z^2)$ scale, as well as uncertainties related to the choice of the PDF set.

The next source of uncertainty is related to the PS algorithm and underlying event model. This is addressed by replacing nominal showering procedure, which is done by PYTHIA 8, with alternative interface to HERWIG 7. As mentioned in previous chapter, the MC generator used for producing signal samples uses specific set of tuned parameters, denoted as `A14`. The additional modeling uncertainty is thus derived by comparing nominal samples to equivalent samples with varied `Var3c` parameter governing α_S of ISR in the `A14` tune.

The modeling of the $t\bar{t}Z$ process obtained with nominal setup (MADGRAPH5_aMC@NLO interfaced to PYTHIA 8) is crosschecked with alternative samples generated with SHERPA 2.2.1, but no additional uncertainties are applied for these discrepancies as they are already covered by sources described above.

6.2.2 Background modeling uncertainties

WZ/ZZ+jets background

Modeling uncertainties associated with the diboson processes include, similar as for signal, various scale variations as well as systematics related to the PDF. Since these processes are simulated with SHERPA generator that takes care of all steps - ME calculations, PS and also hadronization - the choice of the matching scale between ME and PS is varied to evaluate its associated uncertainty. Nominal value of CKKW matching scale is set to 20 GeV, while its up and down variations are set to 30 and 15 GeV, respectively. Next uncertainty is related to soft and collinear gluon emission corrections of the higher order. These are governed by the *gluon resummation scale* (QSF), which is varied by factor 2.0 and 0.5 to obtain corresponding up and down uncertainties. Moreover, additional uncertainty is applied also for alternative parton shower recoil scheme (referred to as CSSKIN) [211]. When compared to nominal diboson samples, mentioned variations account for about 1-10% differences, depending on the particular scale.

Systematic uncertainties related to μ_R, μ_F scales and PDF are derived using the same procedures as for signal samples. Renormalization and factorization scales are varied by factors 2.0 and 0.5, both simultaneously and individually (only for differential measurement), and PDF variations follow PDF4LHC prescription.

To account for differences in flavor compositions between dedicated control and signal regions used in the differential measurement, the additional 30% normalization uncertainty is applied on the $WZ/ZZ + c$ component, while 50% is applied on the $WZ/ZZ + b$ events. These uncertainties have been evaluated by comparing data with MC predictions in $Z + b/c$ events, while differences in the heavy-flavor jet fractions

between WZ/ZZ+jets and Z+jets were taken into account as well.

It should be noted that all of the above uncertainties are applied separately to b,c and light-flavor jet components of the WZ/ZZ+jets backgrounds.

Z+jets background

Since Z+jets events are simulated by the SHERPA generator, systematic uncertainties on CKKW matching scale, QSF scale and PDF are evaluated following the same approach as for WZ/ZZ+jets background samples. Renormalization and factorization scales are simultaneously varied by factors 2.0 and 0.5 with respect to nominal values. Additional normalization uncertainty of 10%, motivated by studies in described in Ref. [212], is applied only to light-flavor component ($Z + l$), because normalizations of heavy-flavor components ($Z + b$ and $Z + c$) are treated as free parameters in the inclusive fit (discussed later).

tZq background

Systematic sources for samples featuring tZq topology include variations of PDF, renormalization and factorization scales, which are applied in the same manner as for the WZ/ZZ+jets or signal samples. Moreover, similar as for $t\bar{t}Z$ samples, also tZq background considers up and down variations of `Var3c` parameters in the `A14` tune. The extra normalization uncertainty of 30% is considered as well and is motivated by the dedicated measurements of tZq process presented in Ref. [213, 214].

tWZ background

The approaches to evaluate the μ_R and μ_F scale uncertainties, as well as the uncertainties on the PDF are the same as for the previous samples. Special treatment is employed to account for the interference between signal process and tWZ process. This procedure requires switching from DR1 diagram removal scheme [194] to alternative one, called DR2, what results in corresponding systematic uncertainty.

$t\bar{t}H$ background

For events featuring top-quark pair production associated with Higgs boson, normalization uncertainty of +5.8% and -9.2% is applied, and represents scale variations combined with α_S uncertainties arising from the NLO cross section computation [215]. Moreover, a symmetrized variation of $\pm$3.6% is considered as uncertainty on the PDF, motivated by the previous $t\bar{t}Z$ measurement [8].

Other background processes

Other processes, such as HV (V denotes vector boson - W or Z), VVV, $t\bar{t}W$, $t\bar{t}WW$, $t\bar{t}t$ or $t\bar{t}t\bar{t}$, which are not specified above, have only negligible contribution to the

overall MC yield (typically less than 1%) in the signal regions. Therefore, conservative normalization uncertainty of 50% is applied for each of these processes. The same normalization uncertainty is applied for the fake leptons estimated from the MC that are used in the dilepton inclusive cross section measurement. On the other hand, for the fake leptons obtained from likelihood method that is employed in the differential analysis, uncertainties resulting from the LM are applied for differential variables on the bin-wise level.

Chapter 7

Analysis methods

7.1 Unfolding method used in the differential cross section measurements

The technique of so-called *unfolding* is nowadays widely used in the high energy physics. In any experiment, the distributions of the measured observables (referred to as reconstruction-level spectra) are distorted by detector and physics effects[1], and thus they cannot be directly compared to their theoretical predictions. Therefore, the aim of unfolding is to correct for these distortions, and thus extract truth-level (i.e. particle- or parton-level) physical distributions.

For this purpose several unfolding methods have been developed:

Bin-by-bin method [216] is based on the calculation of generalized efficiency. This evaluates correction factor c_i for events reconstructed in bin i (r_i), which originate from the same bin of truth distribution (t_i) [2]. It is usually calculated using MC simulation. The truth spectrum u_i (estimator of t_i) is then estimated from data d_i as $u_i = c_i \cdot d_i$. Clearly this technique requires the same subdivision in bins for truth-level and reconstruction-level distribution, and hence cannot take into account large migrations between the truth and reconstruction levels. Moreover, it neglects the unavoidable correlations between adjacent bins and therefore this approach is valid only for negligible migrations.

Matrix method [217] tries to address the problem of these migrations by introducing the so-called *migration matrix* M_{ij}. The elements of this matrix represent probability of event originating from bin i (at truth level) to be reconstructed in bin

[1] Such as finite resolution of the detector, its physical limitations, or due to QCD, QED radiative corrections, parton fragmentation or particle decay.

[2] Truth distribution refers to spectrum obtained from the MC generation without including simulation of detector response, thus effectively simulating real physical distribution as predicted by underlying theory model.

j (at reconstruction level). This matrix then needs to be inverted, and estimator of truth spectrum is calculated as $u_i = M_{ij}^{-1} d_j$ (in notation used above for bin-by-bin method). However, this immediately yields problems with M_{ij} inversion, when matrix is singular. Even if migration matrix is constructed in such a way that it is not singular (by modifying binning and evaluating elements from large number of events), the obtained results may be highly unstable and often yields large bin-to-bin fluctuations and correspondingly also large uncertainties. On the other hand, it is only method providing unbiased results, but nevertheless it is not widely used due to the mentioned disadvantages.

Singular Value Decomposition (SVD) [218] was developed to treat the problems with the migration matrix inversion. For this purpose it uses factorisation of migration matrix (of dimensions $m \times n$) in the form $M = ESF^T$, where E represents $m \times n$ orthogonal matrix, F is $n \times n$ orthogonal matrix and S denotes $n \times n$ diagonal matrix with non-negative elements on diagonal (called *singular values* s_i). Using the above notation (from matrix method), the linear system $M\mathbf{u} = \mathbf{d}$, where $\mathbf{u}$ denotes estimator of truth level spectrum with n bins and $\mathbf{d}$ represents data spectrum with m bins, can be then easily diagonalized by introducing two new rotated vectors $\mathbf{z}$ and $\mathbf{y}$, which represent modified truth-level and data spectrum, respectively. Exact solution of unfolding problem is then very simple:

$$\begin{aligned} (ESF^T)\mathbf{u} = \mathbf{d} \quad &\Rightarrow \quad \mathbf{z} \equiv F^T\mathbf{u}, \quad \mathbf{y} \equiv E^T\mathbf{d}, \\ s_i z_i = y_i \quad &\Rightarrow \quad z_i = \tfrac{y_i}{s_i} \quad \Rightarrow \mathbf{u} = F\mathbf{z}, \end{aligned} \tag{7.1}$$

However, under some circumstances, the determination of z_i can be wrong. Firstly, some y_i values can be poorly known due to large uncertainties in data $\mathbf{d}$. Moreover, some singular values s_i may be small or even zero, what would exaggerate the contribution of poorly known y_i.

Different approaches are based on Bayes theorem and therefore provide better interpretation of unfolded results in sense of probability theory. The most widely used "bayesian" techniques are called *Iterative Bayesian Unfolding (IBU)* [219, 220] and *Fully Bayesian Unfolding (FBU)* [221], while the first one is employed for the extraction of differential cross section distributions in trilepton and tetralepton $t\bar{t}Z$ decay channels. Basic ideas and derivation of the IBU method is given in the next section, while its implementation for the differential cross section measurement will be presented later.

7.1.1 Fundamentals of IBU

The Bayes theorem offers the most powerful tool for making statistical inferences. It has several advantages compared to different unfolding methods:

- it can be applied to multidimensional problems,
- it is able to take into account any kind of smearing and migration between reconstruction-level and truth-level distributions,
- it can handle different binning on truth and reconstruction level,
- it does not require inversion of migration matrix.

IBU starts from Bayes formula which is defined as follows:

$$P(T_i|R_j) = \frac{P(R_j|T_i)P_0(T_i)}{\sum_{l=1}^{n_T} P(R_j|T_l)P_0(T_l)}, \tag{7.2}$$

where $P(T_i|R_j)$ is called *posterior* probability and represents conditional probability to observe an event originating from bin i of the truth-level spectrum (T_i) in bin j of reconstruction-level spectrum (R_j). T stands for truth-level spectrum, while R denotes reconstruction-level spectrum. $P_0(T_i)$ is called *prior* probability of truth spectrum (hypothesis). First term on the right hand side, $P(R_j|T_i)$, is called *likelihood* and denotes conditional probability with the meaning, that if the event is reconstructed in bin j, then it originates from bin i of the truth-level spectrum. Term in the denominator serves as normalization factor for posterior probability, while the sum goes over truth bins (n_T is number of bins on the truth level).

On the first sight this formula yields results dependent on prior probability, but iterative approach, which will be described later, ensures its convergence to the correct hypothesis. As usual, $\sum_{i=1}^{n_T} P_0(T_i) = 1$, and if the probability for some bin T_i is zero, it can never change (meaning that no event can originate from this bin). Likelihood term is evaluated from simulated MC reconstruction- and truth-level distributions and plays a role of migration matrix. Note that $0 \leq f_i \equiv \sum_{j=1}^{n_E} P(R_j|T_i) \leq 1$ because there is no need for each event to be reconstructed, due to the detector limitations. f_i is usually called acceptance term [3].

Using the Bayes theorem, the estimator of the truth distribution $\hat{n}(T_i)$ can be evaluated as follows:

$$\hat{n}(T_i) = \frac{1}{f_i} \sum_{j=1}^{n_R} n(R_j)P(T_i|R_j), \tag{7.3}$$

where $n(R_j)$ is measured spectrum (with n_R bins) of variable of interest.

7.1.2 Iterative approach

As was already mentioned, the estimator obtained from Eq. 7.3 depends on chosen prior probabilities $P_0(T_i)$ and therefore this would yield biased results. For this reason, the

[3] But it can be exchanged with efficiency correction following convention used for the ATLAS analyses. The exchange affects only naming pattern without affecting its statistical meaning.

iterative procedure is introduced. The bayesian procedure provides, beside estimator of the truth spectrum, also estimator of the total number of events in the truth spectrum:

$$\hat{N}_{true} = \sum_{i=1}^{n_T} \hat{n}(T_i). \tag{7.4}$$

Estimator of the prior can be then calculated as:

$$\hat{P}(T_i) = \frac{\hat{n}(T_i)}{\hat{N}_{true}}. \tag{7.5}$$

Finally, the estimator of the overall acceptance, which may differ from the acceptance calculated from MC events [4] is defined as follows:

$$\hat{\epsilon} = \frac{N_{obs}}{\hat{N}_{true}}, \tag{7.6}$$

where N_{obs} is the total number of detected events (measured data).

If the initially chosen prior distribution does not correspond to data, its estimator $\hat{P}(T)$ will change. It can be shown that this estimator of prior lies between initial prior $P_0(T)$ and the true prior distribution. This suggests to proceed iteratively with following steps:

1. In the first iteration, choose initial prior $P_0(T)$ either according to the truth-level MC, or if this does not correspond to physical process, or it is in some way unusable (i.e. biased), the process can be started also from the uniform distribution $P_0(T_i) = 1/n_T$.

2. Calculate estimator of the truth spectrum $\hat{n}(T)$ and the prior $\hat{P}(T)$ according to Eq. 7.3 and 7.5.

3. Make a χ^2 test between $\hat{n}(T)$ and $n_0(T)$ ($n_0(T_i) = P_0(T_i)N_{obs}$) in case of first iteration, or between estimators of truth spectra for current and previous iteration.

4. If χ^2 is not reasonably small, replace $P_0(T)$ by $\hat{P}(T)$, and $n_0(T)$ by $\hat{n}(T)$ (or estimators of truth spectra and priors for current and previous iteration), and start the process again with next iteration.

Number of iterations serves as regularization parameter and hence needs to be optimized. This is because the algorithm tends to "learn" for the statistical fluctuations, which then become more pronounced for higher number of iterations. Nevertheless,

[4] This is defined as $\epsilon_0 = N_{reco \wedge gen}/N_{gen}$, where $N_{reco \wedge gen}$ corresponds to the total number of MC events passing both reconstruction-level and truth-level selection criteria and N_{gen} is the total number of MC events passing truth-level selection.

the IBU technique provides relatively quick convergence to the true distribution with using only few iterations (usually 3 or 4).

The original algorithm, proposed by D'Agostini in 1995 [219], has since been improved [220], while the main improvements were made in error calculation and treatment of statistical fluctuation. For the first issue, the standard error propagation used in the original method was replaced by the numerical MC integration considering Dirichlet probability density function for posterior probabilities. The problem of strengthening the statistical fluctuations between iterations was solved by introducing smoothing of newly estimated truth spectrum before passing it to the next iteration[5]. Typically used smoothing function is low order polynomial.

7.2 Methods used in the inclusive cross section measurement

Inclusive cross section measurement in the dilepton channel is based on the usage of profile likelihood fitting technique. The distribution that is used for the extraction of *signal strength* (ratio of the measured cross section and the value predicted by the SM), is obtained as an output of the multivariate technique applied to data in order to separate signal events from background contamination and thus improve sensitivity of the fitting method. Both multivariate technique and profile likelihood method are briefly described in the following sections. Details about implementation of the presented methods in the inclusive cross section measurement will be given later in the dedicated chapter.

7.3 Multivariate techniques

As already mentioned, the dilepton $t\bar{t}Z$ channel suffers from high background contamination, with main backgrounds yielding up to one order of magnitude more events than signal process. Therefore it is crucial to design the selection criteria in such a way as to suppress the background processes as much as possible. Simple cuts on kinematic observables connected to decay products are not always enough to get reasonable separation of signal and background events. Moreover, discriminating between signal and background events improves stability of the fitting procedure since it reduces statistical fluctuations of otherwise sparsely populated bins of fitted distribution.

To achieve optimal signal/background separation (usually far better than selection based on kinematic criteria), the *multivariate techniques* are employed. There exist various types of multivariate analysis (MVA), but all of them are based on combining

[5]Of course smoothing is not applied to the final true distribution.

set of variables into final discriminant during the *training* phase. Trained model can be subsequently applied to perform binary classification (or multiclassification in case of more than two classes in the output) on the unknown dataset (usually real measured data), that is related to the probability of the event to come from signal. The most widely used MVA algorithms are Neural Networks (NN), that are motivated by nervous system of human brain, and Boosted Decision Trees (BDT) that are based on building decision trees using series of simple cuts. Since NN was determined as optimal for the dilepton analysis, its basic concepts are described in the following section.

7.3.1 Artificial neural networks

The basic idea of the artificial NN [222] comes from brain function which uses units called neurons and their connections, called synapses, to send information around the nervous system. The same concept of nodes, playing role of neurons, and synapses is used in the NN architecture. Each node takes an input in form of input variables in case of no previous nodes in the chain, or output from previous nodes. Then the node preforms mathematical operation and passes the result to next node in the chain. Architecture of the NN characterized by number of layers and nodes, its example being shown in Figure 7.1, strongly depends on the analyzed data and should be optimized to get the best performance.

The general idea behind the NN is to improve the separation of events based on some simple cuts (linear boundary) by replacing them with more complicated boundary that would separate events better. Such a boundary can be constructed i.e. by linear combination of the various low-level input variables. Following the notation from Figure 7.1 the output of the node i, denoted by y_i can be then written as follows:

$$y_i(\mathbf{x}) = w_0 + \sum_{m=1}^{M} w_m h_m(\mathbf{x}), \tag{7.7}$$

where $\mathbf{x} = (x_1, ..., x_M)$ denotes node inputs (either input variables or outputs from previous layer), w_m represent weight coefficients of linear combination, h_m is basis function, usually called *activation* function and sum goes over node connections, synapses. The first, constant term w_0 is usually called *bias*, and allows the nodes in the next layer to obtain certain offset.

By allowing enough flexibility for the activation functions, the boundary of any shape can be created. The most widely used activation functions are *sigmoid*, defined as $h(t) = 1/(1+e^t)$, and *hyperbolic tangent* $h(t) = (e^t - e^{-t})/(e^t + e^{-t})$, where t denotes particular input variable (or output from previous layer of NN). The linear combination of such functions then define the shape of decision boundary as shown in Figure 7.2.

The linear combination from Eq. 7.7, which describe only first layer can be generalized for the case of multiple layers (here output from 2 layers corresponding to y_{ANN}

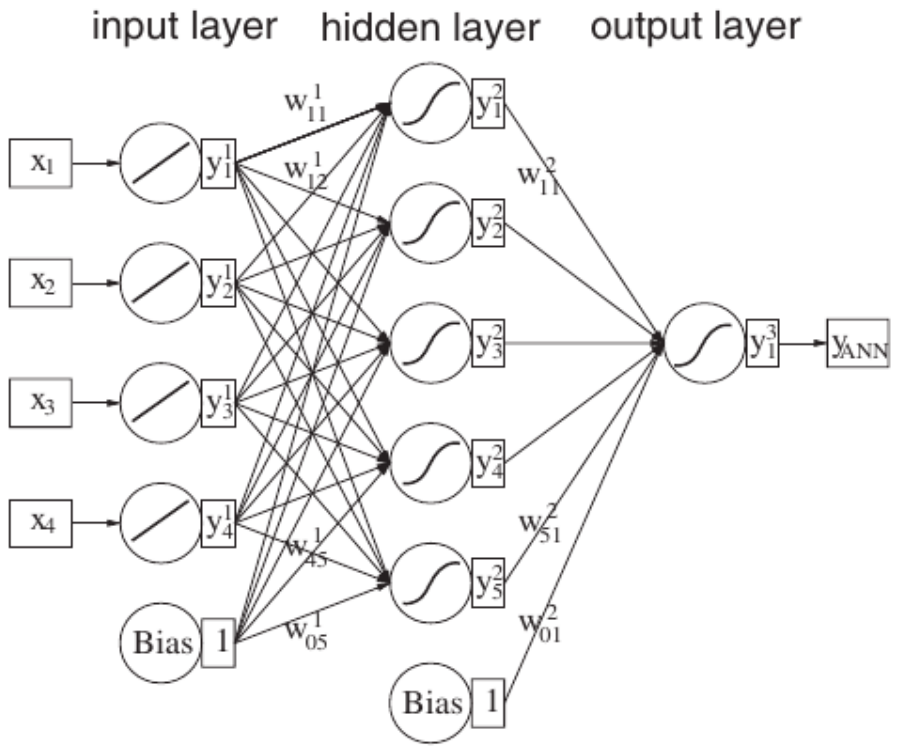


Fig. 7.1: Example scheme of the NN architecture with four input variables $\mathbf{x}$, one hidden layer with five nodes and one output node y_{ANN}. Figure is taken from [222].

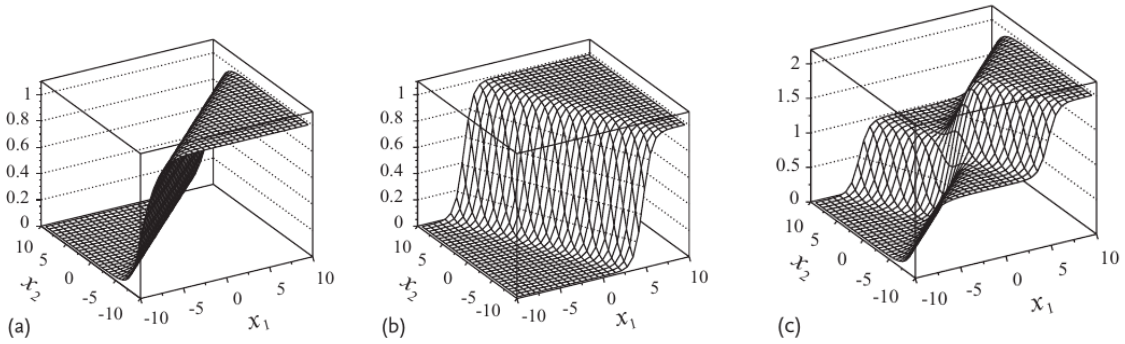


Fig. 7.2: Illustration of the construction of complicated non-linear decision boundary (c). Sigmoid activation function $h(-x_1 + x_2)$ defined using input variables x_1 and x_2 (a) is combined with the similar sigmoid function $h(-2x_1 - x_2)$ defined with different linear combination of inputs (b). By adding more activation functions and changing their input weight, arbitrary decision boundary can be constructed. Figure is taken from [222].

in Figure 7.1):

$$y(\mathbf{x}) = w_0^2 + \sum_{m=1}^{M} \left[w_m^2 h \left(w_{0m}^1 + \sum_{k=1}^{D} w_{km}^1 x_k \right) \right], \tag{7.8}$$

where the same activation function h is considered for all nodes, and term in round brackets (argument of h) represents the output from the first layer, which is employed as input for the activation activation functions h. Architecture from Figure 7.1 could be thus described by Eq.7.8 with $D = 4$ (number of variables) and $M = 5$ (number of basis functions).

It is then straightforward to expand the architecture by adding more layers (middle layers are usually called *hidden*) or more nodes with various activation functions. Number of nodes and layers, as well as used activation functions are treated as free parameters (typically called *hyperparamters*) of the NN and hence need to be optimized. For the activation function of the final node is usually chosen sigmoid in case of binary classification. The reason is that its output value is in a range between 0 and 1, where 1 corresponds to signal-like event while background events peak close to 0.

When looking at the Eq. 7.8 it is natural to ask how to choose the weight coefficient $\mathbf{w}$ to achieve best discrimination between signal and background. To find optimal weights, the metric quantifying the NN performance need to be introduced. This metric is called *loss function*, and it represents quality of separation between signal and background. In the process of NN training, the loss function is minimized, and the weights corresponding to loss minimum are considered optimal for maximal discrimination, thus yielding ideal decision boundary for separating signal from background. Since the weights are typically highly correlated, the resulting loss function has very complicated shape with a lot of local minima. Standard minimization techniques therefore do not yield satisfactory results and more sophisticated method, including technique called *backpropagation*, is used. This technique typically starts from random weights and calculates gradient of the loss function as a function of weights. The procedure continues by moving against gradient direction towards smaller values of loss function. The weights are updated according to the following formula:

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \eta \cdot \nabla_{w_n}, \tag{7.9}$$

where η is called *learning rate* and quantifies how aggressively algorithm proceeds against direction determined by gradient. To increase the convergence speed and prevent it from ending up in some local minimum, the updating of the weights is improved by using so-called *momentum* term [223]:

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \eta \cdot \nabla_{w_n} - \alpha \cdot \nabla_{w_{n-1}}, \tag{7.10}$$

where α is called momentum rate, and $\nabla_{w_{n-1}}$ corresponds to gradient from the previous step. Both learning and momentum rates are free hyperparameters of NN and

thus should be optimized. The backpropagation earned its name because the network output is first evaluated from the current weights using training events, and then the updated weights are fed back into the NN. The backpropagation algorithm can be further improved by using not only first derivatives, but also second derivatives represented by Hessian matrix.

The choice of the loss function depends on the classification problem. In case of simple signal/background discrimination, so-called binary cross-entropy is usually used, which is defined as follows:

$$L(\mathbf{w}) = \sum_{i=1}^{\text{events}} [y^{(i)} \ln(y(\mathbf{x}^{(i)}, \mathbf{w})) + (1 - y^{(i)}) \ln(1 - y(\mathbf{x}^{(i)}, \mathbf{w}))], \tag{7.11}$$

where $y^{(i)}$ denotes class membership (1 for signal and 0 for background event i) and $y(\mathbf{x}^{(i)}, \mathbf{w})$ is network output corresponding to event i and weights $\mathbf{w}$. The definition of the binary cross-entropy emerges from the intention to interpret network output as probability of the event to be signal-like P_s, hence:

$$P_s(\mathbf{x}^{(i)}, \mathbf{w}) = y(\mathbf{x}^{(i)}, \mathbf{w}), \quad P_b(\mathbf{x}^{(i)}, \mathbf{w}) = 1 - y(\mathbf{x}^{(i)}, \mathbf{w}), \tag{7.12}$$

where P_b is the complementary probability of the event to come from background. This last two equations can be written in a more compact form as:

$$P(y^{(i)}, \mathbf{x}^{(i)}, \mathbf{w}) = y(\mathbf{x}^{(i)}, \mathbf{w})^{y^{(i)}} [1 - y(\mathbf{x}^{(i)}, \mathbf{w})]^{(1-y^{(i)})}. \tag{7.13}$$

The likelihood (probability for the entire dataset) then reads:

$$\mathcal{L}(\mathbf{w}) = \prod_{i=1}^{\text{events}} y(\mathbf{x}^{(i)}, \mathbf{w})^{y^{(i)}} [1 - y(\mathbf{x}^{(i)}, \mathbf{w})]^{(1-y^{(i)})}. \tag{7.14}$$

Taking the logarithm of $\mathcal{L}(\mathbf{w})$ directly yields loss function in Eq. 7.11.

The NN architecture used in the dilepton channel analysis is based on KERAS software package [224] with TENSORFLOW [225] backend. Details about the architecture and its hyperparameters will be specified in chapter dedicated to dilepton analysis.

Overtraining and cross-training

Training of the NN in high energy physics analyses is typically performed using MC samples, which contain events of known classification and thus can be employed in NN learning procedure. However, MC samples always contain only finite number of events, often with limited statistics, what results in unavoidable statistical fluctuations. From the above description of the NN it could be tempting to use as many variables, layers and nodes as would be manageable in a reasonable CPU time. Unfortunately, too flexible NN would lead to *overtraining*, when NN would be trained, and consequently

very sensitive, to random statistical fluctuations. The overtrained classifier would then be biased towards fluctuations and classify events incorrectly. There exist several ways how to reduce the overtraining, among the most effective being following:

- Reduce the number of input variables because lower number of variables results in smaller space that can be exploited by the NN. On the other hand, more input variables usually results in higher separation power of the NN and therefore the variables and its number should be chosen carefully.

- Use so-called *early stopping*, when the training is stopped once the value of the loss function does not change by more than predefined factor between the epochs[6]. This prevents the non-linearities in the trained model from creating too detailed decision boundaries.

- Introduce regularization parameters into the loss function, i.e. in a form $L(\mathbf{W}) \rightarrow L(\mathbf{w}) + \sum \mathbf{w}^{\mathrm{T}}\mathbf{w}$ [7], which serve as a penalization term for large weights.

- Add so-called *dropout* layer behind some layers. Its function is to randomly drop (set to zero) predefined fraction r of weights from the previous layer, what again eliminates random fluctuations. The remaining weights are then scaled by factor $1/(1-r)$ such that the original sum over inputs is not changed.

As mentioned, the NN training is performed with MC samples, but once they are used for training, they cannot be used further in the analysis. The reason is that overtraining could bring some bias, which would be translated into incorrect classification if the model would be applied to training dataset. It is therefore necessary to use only subset of MC events for training and the rest for the analysis. However, MC samples are statistically limited and loosing some subset for training (usually one half is used to get at least reasonable NN performance) would lead to increase in statistical uncertainty of the remaining portion of the MC sample. There is, fortunately, approach called *cross-training* that allows to keep whole MC for both training and further analysis. Its idea is to split MC samples into two halves, where one half is used for training the NN model that is subsequently applied to second half, and vice versa second NN is trained on second half and applied to events from the first half. When the trained NNs are then used to classify the measured data, the data sample has to be again splitted into halves so that both independent NNs would be used. Both training sets used in the cross-training has to be independent and contain the same underlying physics effects to ensure that both trained NNs provide the same response. The splitting of

[6] One epoch refers to one update of the weights.

[7] This is an example of so-called *L2* regularization, while *L1* regularization, which adds absolute value of weights magnitude $\sum |w|$ as a penalty term (instead of $\sum \mathbf{w}^{\mathrm{T}}\mathbf{w}$), is also widely used.

the MC sample is performed based on the unique event identifier (usually integer event number). Each event is then assigned to first or second half based on the even or odd value of its event number.

Choice of input variables

The input variables that are chosen for training have the largest impact on the final performance of the NN, and thus their selection is one of the most important steps of the NN optimization. Before variables are used in the training procedure, they often need to be transformed to span only in some predefined range without loosing their separation characteristics. This is necessary because distributions corresponding to some input variables can be defined on the large scales while others on many orders of magnitude smaller scales, what in turn increase the difficulty of classification problem. Large values of input variables can cause that training will start far from the domain of linear response, what would in turn result in the failed convergence of the gradient descent method due to the almost zero gradient. Moreover, the large input values can cause network to learn large weight values resulting in unstable model and consequently its worse performance [226].

There exist many transformation options and its choice depends mostly on the used training algorithm. One of the simplest form is variable *standardization*, when each variable is transformed to span in the vicinity of zero, what can be achieved by subtracting its mean value and dividing by the root mean square of the initial distribution. Similarly, the input variables can be transformed to correspond to certain statistical distribution (i.e. uniform or Gaussian). Other method is based on variable *normalization* when the data are forced to be within range 0 and 1. More sophisticated method of variable transformation is based on *decorrelation* using diagonalized covariance matrix. The NN used in the dilepton analysis employs normalization method for transforming input variables.

As already mentioned the most effective way how to eliminate overtraining is to reduce number of input variables. For this reason variables need to be ranked in order to choose the most important ones that significantly improve NN performance. The ideal way would be to train the network with all combinations of available variables and the best combination according to overtraining and separation power would be taken for final training. However, this is usually not viable, since it would require running training at least $n!$ times, where n is number off input variables, what would be computationally very intensive. For this reason various procedures have been developed for creating variable importance ranking.

The most straightforward and often fastest procedure (because it does not require NN training) is to order the variables according to their separation power S that is

defined as follows:

$$S = \frac{1}{2}\sum_{i=1}^{\text{bins}} \frac{(N_i^{\text{sig}} - N_i^{\text{bkg}})^2}{N_i^{\text{sig}} + N_i^{\text{bkg}}}, \tag{7.15}$$

where N_i^{sig} and N_i^{bkg} corresponds to the number of events in bin i of the signal and background distributions normalized to unity, respectively. The separation power can be used as a metric of the separation between signal and background distribution of the particular variable, since it yields 0 for identical distributions, and 1 for non-overlapping histograms. The high separation power of the variable usually indicates its high importance in the training. Unfortunately, this approach has important drawback, namely that it is calculated considering only one particular variable and hence it neglects its correlation with other variables, what can affect its importance for the training. For this reason, the ranking based on the separation power is often replaced by more sophisticated algorithms involving either correlation statistics (i.e. Pearson correlation coefficient [227]) or training of the NN.

The algorithm chosen for the optimization of the input variable set in the dilepton analysis, is called *permutation importance*. In this approach the variable, which impact is evaluated, is randomly shuffled, meaning that values of given variable are randomly interchanged between events. All other variables from the full set remain unchanged. The NN training is then performed for both original (unshuffled) and updated (shuffled) set and their performances are compared. Comparison is based on the Receiver Operating Characteristic (ROC) curve describing dependency of the signal efficiency on the fraction of rejected background events. Area under the ROC (denoted AUC) corresponds to the discrimination power of the NN. AUC equal to 0.5 represents no separation power, while $AUC = 1$ corresponds to the ideal NN when all background events are rejected for any signal efficiency. The permutation importance ranking is then built according to the relative difference between AUC for original variable set and AUC corresponding to set with shuffled variable, while the larger is the relative difference, the higher is the variable in the ranking. The ROC curve can be used also for checking the overtraining, because AUC of the overtrained NN (usually applied to training dataset) is typically larger than AUC for unknown dataset. The final set of variables is thus selected according to permutation importance ranking while the number of input variables is determined from the overtraining check.

Hyperparameter optimization

Discriminating power of the NN can be further increased by optimization of its free parameters. Set of optimizable hyperparameters depends on the complexity of the NN architecture, while in the dilepton analysis the set of optimized parameters includes number of hidden layers, number of nodes in each layer, activation functions used in

the hidden layers, presence of dropout layers and their rates.

Optimization can be done using various methods. The ideal approach is to perform so-called *grid search*, when all possible combinations of parameter values are used in the training and the set yielding best results is chosen. This is apparently very CPU intensive and thus is replaced in the dilepton analysis by simpler *random search*, in which random combinations of parameters are investigated. The more sophisticated algorithms are based on building Bayesian probability models, or are motivated by the mechanisms of the evolutionary biology.

7.4 Profile likelihood fit

The technique of binned likelihood fit represents powerful and statistically coherent method how to incorporate effects of systematic uncertainties into the extraction of the parameter of interest (POI), in this case the signal strength of the $t\bar{t}Z$ process, denoted as $\mu^{t\bar{t}Z}$. The method is based on the construction of a likelihood and its maximization (typically minimization of the negative logarithm of the likelihood is used) in order to estimate the POI and its uncertainty.

7.4.1 Construction of the likelihood function

In order to construct the final form of likelihood used in the cross section measurement, the simplest case with a single channel, one signal and one background process, neglecting all systematic sources, will be considered first. The following notation and derivation of the likelihood function follows the approach presented in the documentation for the *HistFactory* software package [228]. The number of signal events is denoted by S and its corresponding distribution, playing a role of probability distribution function (pdf), by $f_S(x)$. Similarly, background contamination is represented by B, with $f_B(x)$ being its associated pdf. The case when signal strength $\mu = 0$ represents background only hypothesis, while $\mu = 1$ corresponds to hypothesis with the expected signal associated with the expected background (i.e. as modeled by the MC simulation). The ultimate goal of the fit is to extract the continuous parameter μ from the observed data and its associated uncertainty.

Let us assume that an experiment observed n data events, while for each event e, the value of discriminating variable is denoted as x_e. Then the probability of obtaining n events when the expected yield is $\mu S + B$, while taking into account the pdf of x_e as a relative mixture of $f_S(x)$ and $f_B(x)$, can be defined as follows:

$$\mathcal{P}(x_1 ... x_n|\mu) = \text{Poisson}(n|\mu S + B) \left[\prod_{e=1}^{n} \frac{\mu S f_S(x_e) + B f_B(x_e)}{\mu S + B} \right]. \qquad (7.16)$$

Typically data in the experiment are fixed, what makes the above probability function dependent only on μ. The above Eq. 7.16 is called likelihood $L(\mu)$ and instead of its maximization in order to obtain most probable value of μ, the equivalent procedure of the minimization of its negative logarithm, is usually performed. Substituting Poisson probability mass function of a form $\text{Poisson}(n|\nu) = \nu^n e^{-\nu}/n!$ into above formula gives the following likelihood:

$$-\ln L(\mu) = -n\ln(\mu S+B)+(\mu S+B)+\ln(n!)-\sum_{e=1}^{n}\ln\left[\frac{\mu S f_S(x_e)+Bf_B(x_e)}{\mu S+B}\right] \quad (7.17)$$

Since usually the fitted distribution is in a form of histogram, the above derivation needs to be adjusted for binned data. Contents of bin b of the signal and background histogram can be denoted as ν_b^{sig} and ν_b^{bckg}, respectively. The relation between bin ν_b, and signal and background pdfs can be written as follows:

$$f_S(x_e) = \frac{\nu_{b_e}^{\text{sig}}}{S\Delta_{b_e}},\ f_B(x_e) = \frac{\nu_{b_e}^{\text{bckg}}}{B\Delta_{b_e}}, \quad (7.18)$$

where b_e denotes the index of the bin with width Δ_{b_e} containing value x_e. Since pdfs are normalized to unity the following equations must hold:

$$S = \sum_b \nu_b^{\text{sig}},\ B = \sum_b \nu_b^{\text{bckg}}. \quad (7.19)$$

Expressing Eq. 7.16 in terms of binned histograms using above definitions yields:

$$\begin{aligned}\mathcal{P}(n_{\text{B}}|\mu) &= \text{Poisson}(n|\mu S+B)\left[\prod_{b\in\text{bins}}\frac{1}{\Delta_{b_e}}\frac{\mu\nu_b^{\text{sig}}+\nu_b^{\text{bckg}}}{\mu S+B}\right] \\ &= \mathcal{N}_{\text{comb}}\prod_{b\in\text{bins}}\text{Poisson}(n_B|\mu\nu_b^{\text{sig}}+\nu_b^{\text{bckg}}),\end{aligned} \quad (7.20)$$

where n_{B} now represents binned data histogram, and it can be proved that a constant combinatorial factors can be comprised into the term $\mathcal{N}_{\text{comb}}$, which can be neglected since it does not affect the procedure of finding the likelihood minimum.

This simple example, however, lacks complexity that is required by most of the ATLAS physics analyses. Therefore the probability model needs to be extended in the following way in order to satisfy these requirements:

- Ability to add multiple signal and background samples,
- Ability to free float the normalization of some of the background samples (apply similar scale factors as μ for the signal sample),
- Ability to include systematic uncertainties (i.e. those mentioned in Chapter 6) modifying either the normalization or shape of any sample,

- Ability to account for MC statistical uncertainties for any sample on the bin-wise level,
- Ability to include multiple channels (orthogonal regions of the data phase space defined by specific selection criteria) and properly correlate parameters across them,
- Ability to incorporate samples obtained from data-driven techniques.

After including all above mentioned points the probability density function from Eq. 7.20 becomes:

$$\mathcal{P}(n_{cb}, a_p|\phi_p, \alpha_p, \gamma_b) = \prod_{c\in\text{channels}} \prod_{b\in\text{bins}} \text{Poisson}(n_{cb}|\nu_{cb}) \prod_{p\in A+\Gamma} f_p(a_p|\alpha_p). \tag{7.21}$$

Notation used in this formula, and in the next text, follows these conventions: b iterates over *bins*, c represents *channels*, s denotes *samples* and p is general notation for various *parameters* and factors. Greek letters α, ϕ and γ are the elements of the specific parameter categories: $N = \{\phi_p\}$ represents a set of the free floating normalization factors, $A = \{\alpha_p\}$ is a set of parameters associated with the systematic sources (denoted as nuisance parameters) that are constrained by the external measurement a_p, and $\Gamma = \{\gamma_{csb}\}$ includes parameters connected with the constrained bin-by-bin uncertainties. The term $f_p(a_p|\alpha_p)$ represents the pdf of the nuisance parameter α_p constrained by an auxiliary measurement a_p. Since the uncertainties of these auxiliary measurements are usually assumed to be Gaussian, the corresponding pdf is thus also Gaussian (except γ parameters that follow Poisson distribution).

The expected number of events in the particular bin b corresponding to channel c can be then written as follows:

$$\nu_{cb}(\phi_p, \alpha_p, \gamma_b) = \gamma_{csb}\phi_{cs}(\alpha)\eta_{cs}(\alpha)\sigma_{csb}(\alpha), \tag{7.22}$$

where the meaning of the particular factors is following:

- γ_{csb} - stands for a bin-wise scale factor for given channel and sample, and includes statistical uncertainties, bin-by-bin systematics and data-driven shape extrapolations. In case of no bin-by-bin scale factors, $\gamma_{cb} = 1$,
- ϕ_{cs} - is a product of unconstrained scale factors, typically including POIs,
- η_{cs} - represents parametrized normalization uncertainties associated with a given sample and channel [8],

[8]Note that systematic sources are typically decomposed into component affecting only shape of the distribution (pure shape systematics comprised in σ_{csb} with bin-wise scale factors), and component affecting only normalization of the distribution (pure normalization systematics included in η_{cs} resulting in only one scale factor per sample and channel).

- σ_{csb} - denotes the parametrized histogram (including shape systematic uncertainties), for a given sample and channel.

The auxiliary measurement of the given nuisance parameter usually yields nominal value and $\pm 1\sigma$ variations (referred to as *up* and *down* variations). To get continuously parametrized scale factors as above, i.e. $\eta_{cs}(\alpha)$, the interpolation and extrapolation needs to be introduced between nominal values and up and down variations, as illustrated in Figure 7.3. Before the interpolation, each nuisance parameter is decomposed into a pure normalization part (without shape effect) and shape part (without normalization effect). The normalization and shape effects are then interpolated and extrapolated separately, normalization component with polynomial interpolation and exponential extrapolation, while shape component with linear interpolation.

The *polynomial interpolation* and *exponential extrapolation* for normalization nuisance parameters can be defined as follows:

$$\eta_s(\alpha) = \prod_{p \in \mathrm{Syst}} I_{\mathrm{poly.|exp.}}(\alpha_p; 1, \eta^+_{sp}, \eta^-_{sp}, \alpha_0), \tag{7.23}$$

where

$$I_{\mathrm{poly.|exp.}}(\alpha; I^0, I^+, I^-, \alpha_0) = \begin{cases} (I^+/I^0)^\alpha & \alpha \geq \alpha_0 \\ 1 + \sum_{i=1}^{6} a_i \alpha_i & |\alpha| < \alpha_0 \\ (I^-/I^0)^{-\alpha} & \alpha \leq -\alpha_0 \end{cases} . \tag{7.24}$$

Symbols I^+, I^- and I^0 denote the expected event yields for a given up, down variation and nominal prediction for a particular normalization systematic source, respectively. The value $\pm\alpha_0$ represents the boundaries for the polynomial interpolation and typically corresponds to $\pm 1\sigma$ variation of given normalization nuisance parameter. Coefficients a_i defining the polynomial function are fixed according to the boundary conditions given by $\eta(\alpha = \pm\alpha_0)$, $d\eta/d\alpha|_{\alpha=\pm\alpha_0}$ and $d^2\eta/d\alpha^2|_{\alpha=\pm\alpha_0}$ (see Figure 7.3)

As is apparent from its definition, this strategy cannot yield a negative normalization factor what is the main advantage of this strategy. Some problems, caused by discontinuous first derivative at $\alpha = 0$, can emerge in the likelihood minimization.

The *piece-wise linear interpolation* strategy, employed for the shape systematics, is based on the following definition:

$$\sigma_{sb}(\alpha) = \sigma^0_{sb} + \sum_{p \in \mathrm{Syst}} I_{\mathrm{lin.}}(\alpha_p; \sigma^0_{sb}, \sigma^+_{sbp}, \sigma^-_{sbp}), \tag{7.25}$$

where

$$I_{\mathrm{lin.}}(\alpha; I^0, I^+, I^-) = \begin{cases} \alpha(I^+ - I^0) & \alpha \geq 0 \\ \alpha(I^0 - I^-) & \alpha < 0 \end{cases} . \tag{7.26}$$

This approach allows for incorporation of non-symmetric systematic variations, but in some cases this results in the same problem as for exponential interpolation, with the

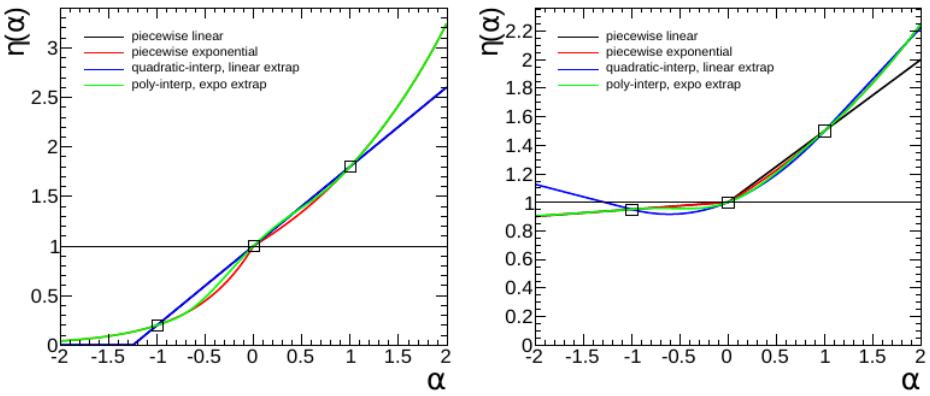


Fig. 7.3: Comparison of the interpolation strategies for different values of normalization variations $\eta^{\pm}$. Left plot corresponds to $\eta^{+} = 1.8$ and $\eta^{-} = 0.2$, while right plot features $\eta^{+} = 1.5$ and $\eta^{-} = 0.95$. Figure is taken from [228].

kink (discontinuity in the first derivative) at $\alpha = 0$. Fortunately, this can be effectively solved by symmetrizing the problematic nuisance parameter.

The approach based on the usage of histograms has its limitations when it comes to the MC simulations. These are often very CPU intensive, thus in some cases the histograms are only sparsely populated. Hence the statistical uncertainty of the MC samples should be taken into account and included in the probability model. A method proposed by Barlow and Beeston [229] suggests to introduce one nuisance parameter for each bin of each sample as an uncertainty on the true rate. This could, however, lead to plenty of nuisance parameter (often several hundreds) for most of the ATLAS analyses. *HistFactory* thus uses a lighter version with one nuisance parameter (γ) per bin associated with the total statistical uncertainty in that bin (including the total MC estimate). Under this approximation the contribution of the statistical uncertainty in bin b can be expressed by following factor:

$$\text{Poisson}(n_b|\nu_b(\alpha) + \gamma_b\nu_b^{\text{MC}}(\alpha))\text{Poisson}(m_b|\gamma_b\tau_b), \tag{7.27}$$

where n_b is the number of observed events in bin b, $\nu_b(\alpha)$ corresponds to the expected number of events for which MC statistical uncertainty does not need to be included (i.e. due to data-driven estimate or statistically large sample), $\nu_b^{\text{MC}}(\alpha)$ on the other hand represents events for which the MC statistical uncertainty cannot be neglected, and has to be associated by the nuisance parameter γ_b. The second term expresses that the MC estimate itself results from an auxiliary measurement and hence is constrained by the mentioned Poisson probability, where m_b corresponds to the total MC yield in bin b that can fluctuate about $\gamma_b\tau_b$.

7.4.2 Minimization of the likelihood function

To find the optimal values of the POIs and desired normalization factors, the likelihood constructed in the previous section need to be maximized. As already mentioned, the finding of POIs is typically done by minimizing negative logarithm of the full likelihood, what yields equivalent results as finding likelihood's maximum. The number of systematic sources and their associated nuisance parameters, that are allowed to float within their constrained limits is, however, very large in majority of the ATLAS analyses. Additionally, the statistics-related nuisance parameters connected with each bin of the fitted distribution need to be considered as well. This makes the number of parameters, that can affect the position of the likelihood minimum, of the order of 100. The minimization problem thus becomes multidimensional and very complex.

The software framework that takes care of the minimization procedure is called *MINUIT* [230] and is based on the underlying method called *MIGRAD*. This method employs a gradient technique and the procedure follows these general steps:

1. In the first step, the inputs are created in the form of starting values of parameters X_i, the first derivatives G_i and covariance matrix V_{ij}. First derivatives are calculated numerically inside MINUIT framework using finite-difference approximation.

2. Then the values of parameters are shifted in the direction of the gradient producing new set of parameters $X' = X - V \cdot G$. If the minimized function F was quadratic and V corresponded to the true covariance matrix, then X' would represent the minimum. However, in general X' is not the minimum, thus the linear search is performed along this direction in order to find α which minimizes $F(X - \alpha \cdot V \cdot G)$.

3. In the next step, matrix V is updated according to a general form $V' = V + f(V, X, X', G, G')$ using either Davidon's formula [231] or Fletcher's dual formula [232].

4. Finally, the original values of X, V and G are replaced by X', V' and G' and steps 1) and 2) are repeated until specific convergence criteria are satisfied. These criteria are based on the calculation of *estimated distance to minimum* (EDM) that is defined as follows:

$$EDM = G^T \cdot V \cdot G, \tag{7.28}$$

   while final value of EDM is required to be less than 0.001.

It is apparent from the above procedure that MIGRAD updates also the covariance matrix V, thus it can provide its final version corresponding to found likelihood minimum. The covariance matrix can be quite trivially transformed to correlation matrix

of the nuisance parameters, from which the post-fit uncertainties corresponding to the nuisance parameters and POIs, can be derived. Since the obtained correlation matrix is symmetric, also the post-fit uncertainties are symmetric. However, the minimized likelihood function does not need to be quadratic and its shape is often asymmetric. To account for this asymmetry the MINUIT package uses so-called *MINOS* algorithm to numerically estimate the asymmetric uncertainties for the POIs. The method is described in detail in Ref. [230], but its basic concept relies on examination of the exact behavior of the likelihood over the specific interval. MINOS determines where the function $F_m(X_i)$ corresponding to POI minimum, but with parameter X_i fixed, attains the value $F_0 + 1/2$ where F_0 stands for overall minimum of the full likelihood. This determination is done for each nuisance parameter considered in the construction of likelihood.

7.4.3 Pruning and smoothing

The large number of nuisance parameters that enter likelihood increases the complexity of the multidimensional fit. Moreover, each nuisance parameter creates a local minimum which makes the fitting procedure unstable and potentially dependent on the starting point of the fit. To solve these issues so-called *pruning* technique is employed to reduce the number of nuisance parameters by choosing only those which have a significant effect on the nominal distribution. As mentioned in Section 7.4.1, each systematic source is decomposed into a pure normalization and pure shape component. Pruning is thus applied separately for both of these components. If the normalization effect is less than 0.01% (of the nominal distribution) the normalization component of the given nuisance parameter for the given sample is dropped. Similarly, if the shape effect, determined as the maximum difference between given systematic variation and the nominal histogram from all bins is less than 0.01%, then the shape component is neglected for the given sample. If both normalization and shape effects are dropped for all considered samples, the affected nuisance parameter is removed completely from the fit. The pruning procedure not only makes the fit more stable and less complex, but consequently it also saves CPU time.

An another issue in the fitting procedure results from the construction of the systematic variations, which are estimated from finite number of events, and are hence associated with non-zero statistical fluctuations with respect to the nominal distribution. These fluctuations can result in artificial constraints of affected nuisance parameters, what can translate into underestimation of the total systematic uncertainty of the measurement. Moreover, if the shape of statistical fluctuations for the given nuisance parameter is similar to fluctuations in the fitted data, the value of the nuisance parameter can be artificially pulled from its nominal value. A procedure called *smoothing*

was developed to remove these unwanted pulls and constraints. Smoothing methods are based on eliminating fluctuation by merging neighboring bins together, or replacing their bin contents by their combined weighted sum. In the final step, special smoothing algorithm (called *353QH*) introduced in Ref. [233], is applied for the distributions in order to ensure smooth transitions between transformed bins.

Chapter 8

Definitions of the signal regions

As mentioned several times the analyses described in this thesis are focused on the $t\bar{t}Z$ inclusive cross section measurement in the dilepton channel, which is currently ongoing, and differential cross section measurement in the trilepton and tetralepton channels, which was already published by the ATLAS Collaboration [10].

The SRs used in the inclusive measurement are designed to offer as high signal sensitivity as possible, while trying to suppress the main background processes. Although the dilepton channel is not the most ideal for the inclusive cross section measurement, due to the largest background contamination, it is included in the ongoing effort to help reduce the total uncertainty when combined with results from the trilepton and tetralepton channels.

The SRs used in the differential analysis are intended to have the highest possible purity of $t\bar{t}Z$ events, as well as sufficient event yields in order to perform a precise characterization of the $t\bar{t}Z$ process and allow for a differential measurement of the $t\bar{t}Z$ cross section which would not be limited by the statistical uncertainties. Therefore only regions with 3ℓ (trilepton) and 4ℓ (tetralepton) signatures described in Section 2.4 are considered in this analysis. The differential cross section measurement is performed in specific regions of the phase space, defined by either parton or particle level requirements [1]. This differential analysis is accompanied also by inclusive cross section measurement (which is not described in this thesis), and since it uses some results from the inclusive analysis, namely the normalization of the WZ/ZZ+light jets backgrounds, the SRs used in the inclusive measurement are described in this chapter as well.

[1] At the particle level, τ leptons are considered to be unstable. Therefore τ decay products (jets and leptons from leptonic decay) are present at particle level, while at parton level τ leptons are excluded.

8.1 Dilepton signal regions

The dilepton channel targets $t\bar{t}Z$ events where the $t\bar{t}$ system decays hadronically while an opposite-sign same-flavor (OSSF) pair of leptons originates from the Z boson. The invariant mass of the lepton pair is required to be within 10 GeV of the Z boson mass [21], with the leptons satisfying the optimised requirements on their transverse momentum (p_{T} greater than 30 and 15 GeV respectively for the leading and subleading lepton in p_{T}).

The dilepton channel generally suffers from a low signal-to-background ratio, with the $t\bar{t}$ and Z+jets processes (both characterized by the presence of two prompt leptons) constituting major backgrounds. Three signal regions are defined, based on jet and b-jet multiplicities, while the b jets are required to pass fixed 77% b-tagging operating point defined by DL1R tagger (see Section 4.4.4 for details). In the SR denoted as 2ℓ-Z-5j2b, exactly 5 jets are required, of which at least two must be b-tagged. Second SR called 2ℓ-Z-6j2b provides its jet-inclusive complement (at least 6 jets). Finally the 2ℓ-Z-6j1b SR, inclusive in jet multiplicity but requiring exactly one b-jet, targets events with the appropriate jet multiplicity for tree-level $t\bar{t}Z$ events but with one non-identified b-jet.

In order to further improve the separation between the $t\bar{t}$ and Z+jets background processes and the $t\bar{t}Z$ signal, one deep NN is trained for each SR on events passing the corresponding selection. Details about the NN architecture, training and validation are given later in Section 11.3.

All SR selection criteria are summarized in Table 8.1. Respective event yields in all three SRs are shown in Table 8.2.

Variable	2ℓ-Z-$5j2b$	2ℓ-Z-$6j2b$	2ℓ-Z-$6j1b$
N_ℓ ($\ell = e, \mu$)	$= 2$		
	$= 1$ OSSF lepton pair with $\lvert m^Z_{\ell\ell} - m^Z_{nom}\rvert < 10$ GeV		
$p_{\mathrm{T}}(\ell_{1,2})$	$> 30, 15$ GeV		
N_{jets}	$= 5$	≥ 6	≥ 6
$N_{b-\mathrm{jets@77\%}}$	≥ 2	≥ 2	$= 1$

Table 8.1: Definitions of the signal regions used in the inclusive cross section measurement in the dilepton channel.

8.2 Trilepton signal regions

The trilepton signal regions target leptonic decay modes of the Z boson accompanied by semileptonic decays of the $t\bar{t}$ pair. Three different regions with exactly three leptons

	2ℓ-Z-$5j2b$	2ℓ-Z-$6j2b$	2ℓ-Z-$6j1b$
$t\bar{t}Z$	279.9 ± 13.7	430.0 ± 38.3	311.8 ± 29.3
$t\bar{t}$ DD	3944.9 ± 70.4	1935.7 ± 47.0	1183.1 ± 36.2
$Z + b$	5093.1 ± 821.7	2629.0 ± 479.4	4802.6 ± 915.5
$Z + c$	557.1 ± 123.7	311.0 ± 74.3	3315.2 ± 695.7
$Z + l$	65.0 ± 35.8	20.0 ± 10.8	1008.0 ± 306.4
tWZ	22.4 ± 2.5	33.3 ± 4.7	39.8 ± 4.5
Diboson (VV)	174.6 ± 87.8	112.3 ± 56.5	403.5 ± 202.8
Fake leptons	28.0 ± 14.1	18.2 ± 9.2	25.0 ± 12.6
Other	46.3 ± 23.2	41.1 ± 20.6	19.1 ± 9.6
Total SM	10211.1 ± 913.5	5530.7 ± 570.1	11108.0 ± 1752.2
Data	-	-	-

Table 8.2: The numbers of expected signal and background events in the dilepton signal regions, obtained for an integrated luminosity of 139 fb^{-1}. The indicated errors include the MC statistical as well as the systematic uncertainties on the different SM components (see Section 6). Category labeled as "Other" includes all other SM processes that feature at least two prompt leptons and are not listed in this table (the complete list can be found in Section 5.2.7). The $t\bar{t}$ background is estimated from data using technique described in Section 11.2 (DD stands for data-driven). Data yields are not quoted in the table since this analysis is in the ongoing state and the results presented in this thesis are based only on the Asimov (MC) dataset.

(electrons or muons) are considered[2].

The OSSF lepton pair, whose invariant mass is reconstructed to be closest to the nominal Z boson mass, is identified as originating from the Z boson. Moreover, the invariant mass of such a pair, $m_{\ell\ell}^{Z}$, is required to be inside 10 GeV window around nominal Z boson mass. Electric charges of the three leptons has to sum up to ± 1. All OSSF lepton pairs are further required to fulfill $m_{OSSF} > 10$ GeV in order to remove contributions due to low-mass resonances, which are not included in the simulation. The SRs are discriminated based on the total number of reconstructed jets and b jets

[2]The requirements on the lepton multiplicities in the trilepton and tetralepton channels always refers to the number of signal and baseline leptons. Signal leptons are required to pass criterium for lepton isolation, while baseline leptons does not. Accordingly, events with three signal leptons and one additional baseline lepton which does not fulfill the signal requirements are not included in the trilepton channel.

present in the event. All jets need to have $p_{\mathrm{T}} > 25$ GeV. The minimum p_{T} requirement for the leading, subleading and third lepton, is set to 27, 20 and 20 GeV, respectively.

A summary of the trilepton SR definitions is provided in Table 8.3. Different approaches are used for the inclusive and the differential measurements: while for the differential selection, the b-jet requirements are based on a fixed 85% operating point (OP) of the MV2c10 discriminant, the inclusive measurement is performed with a combination of two orthogonal SRs which use flexible b-tagging OPs (denoted as PCBT, see Section 4.4.4). The region labeled 3ℓ-Z-$2b3j$ with a fixed b-tagging OP is used for the differential cross-section measurements. For the inclusive measurement, a combination of 3ℓ-Z-$1b4j$-PCBT with a *tight* OP of 60% for the b-jet[3] and 3ℓ-Z-$2b3j$-PCBT (with two b-jets, but a looser OPs of 70%), is used.

The idea behind these different approaches is to optimize the inclusive measurement for precision (by applying rather tight b-tagging OPs) and to suppress the WZ background, whereas for the differential measurement, the priority lies in keeping the SRs populated with many $t\bar{t}Z$ events (in order to have sufficiently high number of events for a differential measurement). Therefore, a looser b-tagging OP is chosen for the differential selection. The event yields for inclusive, as well as differential regions are shown in Table 8.4.

Variable	3ℓ-Z-$1b4j$-PCBT inclusive	3ℓ-Z-$2b3j$-PCBT inclusive	3ℓ-Z-$2b3j$ differential
N_ℓ ($\ell = e, \mu$)	$= 3$		
	≥ 1 OSSF lepton pair with $\lvert m_{\ell\ell}^{Z} - m_{nom}^{Z} \rvert < 10$ GeV		
	for all OSSF combinations: $m_{OSSF} > 10$ GeV		
$p_{\mathrm{T}}(\ell_{1,2,3})$	$> 27, 20, 20$ GeV		
Sum of lepton charges	± 1		
$N_{\mathrm{jets}}(p_T > 25\,GeV)$	≥ 4	≥ 3	≥ 3
$N_{b-\mathrm{jets}}$	$= 1$@60%	≥ 2@70%	≥ 2@85%
	veto add. b-jets@70%		

Table 8.3: The definitions of the SRs defined for the trilepton channel: the inclusive cross section measurement uses a combination of the PCBT regions labeled 3ℓ-Z-$1b4j$-PCBT and 3ℓ-Z-$2b3j$-PCBT, while the differential measurement is performed only in region 3ℓ-Z-$2b3j$, which employs fixed b-tagging OP.

8.2.1 Trilepton parton-level fiducial volume

The parton-level fiducial volume in the 3ℓ channel is defined as the top-quark pair decaying semi-leptonically (e, μ+jets only) and Z boson decaying dileptonically via

[3]With all additional jets required to fail the 70% OP.

$Z \to ee, \mu\mu$. Events featuring τ leptons which originate directly from either the Z boson (via $Z \to \tau\tau$) or the W boson from the $t\bar{t}$ system (via $W \to \tau\nu_\tau$) are removed from the fiducial volume (part of the phase space in which the measurement is conducted) and are not considered for the unfolding, regardless of their subsequent decay[4]. The differential variables are reconstructed from the top quarks after final state radiation, immediately prior to their decays. The invariant mass of the two leptons from the Z boson decay is required to be within ± 15 GeV of the nominal m_Z value[5].

8.2.2 Trilepton particle-level fiducial volume

The particle-level fiducial volume in the 3ℓ channel is constructed to emulate the reco-level 3ℓ-Z-$2b3j$ region. The selection is therefore as follows:

- exactly 3 leptons with $p_T(\ell_1, \ell_2, \ell_3) > 27, 20, 20$ GeV,
- at least one OSSF lepton pair with $|m_{\ell\ell} - m_Z| < 10$ GeV,
- at least three jets (with $p_T > 25$ GeV), with at least two of them to be b jets (at particle level, jets are required to be ghost matched to a b-hadron in order to be considered a b jet).

8.3 Tetralepton signal regions

The tetralepton SRs target the case with Z boson and both W bosons from $t\bar{t}$ pair decaying leptonically. In the final state, two b quarks, four leptons (OSSF lepton pair from Z boson and the other opposite-sign pair from $t\bar{t}$) and two neutrinos from the $t\bar{t}$ pair, are expected. If all objects are reconstructed correctly, the signal should lead to signatures with two b jets, one Z-like lepton pair (OSSF and invariant mass close to the nominal Z boson mass), another opposite-sign (OS) lepton pair and missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. The minimum p_{T} requirement for the leading, subleading, third and forth lepton is 27, 20, 10 and 7 GeV, respectively. In the event reconstruction, the OSSF

[4]The choice of whether or not to include τ decays and their impact on the differential variables was studied extensively; ultimately this choice was seen to have a very small impact on the final measurements but there were small effects which motivated the choice of a harmonized treatment of all τ decays. Since there is no possibility to distinguish between e, μ from τ decay and from other sources (i.e. W, Z bosons), the τ leptons are accepted at detector and particle level, if they pass the respective selection criteria.

[5]The addition of the on-shell requirement for the Z boson in the parton-level fiducial volume was required due to the truth-record information from the alternative signal samples (produced by SHERPA 2.2.1 generator) which are compared with the nominal signal in the final unfolded measurements. For this reason, the width of the Z mass window for acceptance is increased to 15 GeV when compared to detector and particle level, where 10 GeV window is used.

	3ℓ-Z-$2b3j$-PCBT	3ℓ-Z-$1b4j$-PCBT	3ℓ-Z-$2b3j$ (diff.)
$t\bar{t}Z$	210.0 ± 8.9	164.0 ± 14.2	352.3 ± 24.8
$WZ + l$	0.3 ± 0.3	3.1 ± 2.3	29.8 ± 14.5
$WZ + b$	16.0 ± 9.4	29.5 ± 17.9	40.6 ± 24.2
$WZ + c$	2.1 ± 1.0	12.5 ± 5.9	41.2 ± 19.0
$ZZ +$ jets	2.4 ± 1.1	4.6 ± 1.8	11.1 ± 4.4
tWZ	19.3 ± 7.1	23.7 ± 4.2	37.9 ± 10.9
tZq	29.3 ± 9.3	11.7 ± 5.0	48.7 ± 14.9
$t\bar{t}W$	4.3 ± 2.1	1.5 ± 0.7	6.5 ± 3.3
$t\bar{t}H$	5.8 ± 0.5	4.3 ± 0.4	9.5 ± 0.9
Fake leptons	14.6 ± 7.4	29.9 ± 15.1	37.4 ± 14.0
Other	1.5 ± 0.7	0.7 ± 0.4	2.4 ± 1.2
Total SM	305.6 ± 20.9	285.6 ± 31.3	617.5 ± 59.0
Data	343	272	583

Table 8.4: The numbers of observed data and expected background events in the trilepton signal regions, obtained for an integrated luminosity of 139 fb^{-1}. The indicated errors include the MC statistical as well as the systematic uncertainties on the different SM components (see Section 6). Category labeled as "Other" includes all other SM processes that feature at least three prompt leptons and are not listed in this table (the complete list can be found in Section 5.2.7). Reasonable agreement between data and the prediction is observed in all SRs.

lepton pair, whose invariant mass is closest to the nominal Z boson mass, is considered to be the lepton pair from Z-boson decay and it is labelled as $^{Z}_{\ell\ell}$. The other lepton pair is labeled $^{\text{non}-Z}_{\ell\ell}$ and is identified as originating from the $t\bar{t}$ pair. Furthermore, the non-Z lepton pair is required to be opposite-sign, so that the sum of the charges of the four leptons is 0. The signal regions are split according to the flavor of the $^{\text{non}-Z}_{\ell\ell}$ leptons into different flavor (DF) and same flavor (SF) signal regions. This splitting is only relevant for the inclusive fit where the regions are fitted simultaneously, while for the differential cross section all the SRs are merged into one. The dominant ZZ+jets background is further reduced requiring minimal number of jets and b jets[6] and applying cuts on $m_{\ell\ell}^{\text{non}-Z}$ combined with $E_{\text{T}}^{\text{miss}}$. Similarly as for the trilepton selection, all possible OSSF pairs are required to satisfy $m_{OSSF} > 10$ GeV to remove potential contributions from low-mass resonances or photon conversions.

The selection criteria for tetralepton regions are summarized in Table 8.5. The

[6]Note that, unlike the 3ℓ SRs, only b jets with a fixed OP of 85% are used. Again the MV2c10 tagger is employed.

cuts on the $m_{\ell\ell}^{\text{non}-Z}$ and $E_{\text{T}}^{\text{miss}}$, for the ZZ suppression, are applied only in the SF signal regions. If $m_{\ell\ell}^{\text{non}-Z}$ is compatible with Z-boson mass (is inside 10 GeV Z mass window), a higher cut on the $E_{\text{T}}^{\text{miss}}$ is applied, otherwise the $E_{\text{T}}^{\text{miss}}$ cut is less strict (see Table 8.5). The event yields for data and the SM backgrounds in all four tetralepton SRs are shown in Table 8.6.

Variable	4ℓ-DF-1b	4ℓ-DF-2b	4ℓ-SF-1b	4ℓ-SF-2b
$N_\ell(\ell = e, \mu)$	$= 4$			
	$\geq$1 OSSF lepton pair with $\lvert m_{\ell\ell}^{Z} - m_{nom}^{Z}\rvert < 10$ GeV			
	for all OSSF combinations: $m_{OSSF} > 10$ GeV			
p_{T} (ℓ_1, ℓ_2, ℓ_3, ℓ_4)	$> 27, 20, 10, 7$ GeV			
$\ell\ell^{\,\text{non}-Z}$	$e^{\pm}\,\mu^{\mp}$	$e^{\pm}\,\mu^{\mp}$	$e^{\pm}\,e^{\mp}$ or $\mu^{\pm}\,\mu^{\mp}$	$e^{\pm}\,e^{\mp}$ or $\mu^{\pm}\,\mu^{\mp}$
$\lvert m_{\ell\ell}^{\text{non}-Z} - m_{nom}^{Z}\rvert$	–	–	> 10 GeV | < 10 GeV	> 10 GeV | < 10 GeV
$E_{\text{T}}^{\text{miss}}$			> 50 GeV | > 100 GeV	– | > 50 GeV
$N_{\text{jets}}(p_{\text{T}} > 25\,\text{GeV})$	≥ 2	≥ 2	≥ 2	≥ 2
$N_{b-\text{jets}}$@85%	$= 1$	≥ 2	$= 1$	≥ 2

Table 8.5: The definitions of the four tetralepton signal regions. The regions are defined to target different b-jet multiplicities and flavor combinations of the non-Z leptons. Selection criteria related to $E_{\text{T}}^{\text{miss}}$ and $\lvert m_{\ell\ell}^{\text{non}-Z} - m_{nom}^{Z}\rvert$ are applied consistently for both flavor combinations in the SF regions.

8.3.1 Tetralepton parton-level fiducial volume

The parton-level fiducial volume in the 4ℓ channel is defined as top-quark pair decaying dileptonically ($ee, e\mu, \mu\mu$ only) and Z boson decaying also dileptonically via $Z \to ee, \mu\mu$. As in the case of the trilepton fiducial volume, events featuring tau leptons which originate directly from either the Z boson (via $Z \to \tau\tau$) or the W bosons from the $t\bar{t}$ system (via $W \to \tau\nu_\tau$) are removed from the fiducial volume and are not considered for the unfolding, regardless of their subsequent decay. The differential variables are reconstructed from the top quarks after final state radiation, immediately prior to their decays. The invariant mass of the two leptons from the Z boson decay is required to be within ± 15 GeV of the nominal Z mass, similarly as for the trilepton parton level [7].

[7] For the explanation see corresponding footnote in Section 8.2.1.

8.3.2 Tetralepton particle-level fiducial volume

The particle-level fiducial volume in the 4ℓ channel is defined by the following cuts. These are the minimal requirements in order to be able to reconstruct the $t\bar{t}$ system and the Z boson, assuming the dilepton decay of both $t\bar{t}$ pair and Z boson.

- exactly 4 leptons with $p_{\mathrm{T}} > 7$ GeV,
- at least one OSSF lepton pair with invariant mass within ± 10 GeV of the nominal Z boson mass,
- the other lepton pair is required to be opposite-sign,
- p_{T} of the first, second and third lepton are required to be higher than 27, 20 and 10 GeV, respectively,
- at least two jets (with $p_{\mathrm{T}} > 25$ GeV), with at least one of them to be tagged as a b jet (at particle level, a jet ghost matched to a b hadron is considered to be a b jet).

	4ℓ-SF-1b	4ℓ-SF-2b	4ℓ-DF-1b	4ℓ-DF-2b
$t\bar{t}Z$	12.97 ± 1.55	23.32 ± 1.78	16.77 ± 1.40	22.52 ± 1.00
$ZZ + l$	1.80 ± 0.81	0.99 ± 0.52	0.45 ± 0.16	0.02 ± 0.01
$ZZ + b$	1.08 ± 0.69	2.82 ± 1.74	0.22 ± 0.13	0.08 ± 0.05
$ZZ + c$	0.97 ± 0.50	1.20 ± 0.64	0.19 ± 0.09	0.02 ± 0.01
tWZ	2.65 ± 0.48	2.03 ± 0.78	3.57 ± 1.08	2.12 ± 0.88
$t\bar{t}H$	0.46 ± 0.06	0.86 ± 0.10	0.64 ± 0.07	0.75 ± 0.08
Fake leptons	0.68 ± 0.55	0.89 ± 0.75	0.85 ± 0.75	0.26 ± 0.26
Other	0.66 ± 0.34	0.25 ± 0.13	0.71 ± 0.36	0.22 ± 0.11
Total SM	21.27 ± 2.31	32.38 ± 3.08	23.39 ± 2.02	26.00 ± 1.46
Data	19	33	33	32

Table 8.6: Observed and expected event yields in the tetralepton signal regions for an integrated luminosity of 139 fb^{-1}. The indicated errors include the MC statistical as well as the systematic uncertainties on the different SM components (see Section 6). Category labeled as "Other" includes all other SM processes that feature at least four prompt leptons and are not listed in this table (the complete list can be found in Section 5.2.7). Reasonable agreement between data and the prediction is observed in all SRs.

Chapter 9

Event reconstruction and definition of differential variables

Reconstruction of the $t\bar{t}Z$ events from its decay products plays crucial role in both inclusive and differential analysis. In the dilepton inclusive cross section measurement the reconstructed top quarks and Z boson are used to construct discriminating variables that are subsequently used in the training of the NN for signal/background separation.

In order to measure differential cross sections of the $t\bar{t}Z$ process with respect to different variables, reconstruction of both Z boson and $t\bar{t}$ pair is necessary for defining these variables. While the reconstruction of the Z boson is straightforward thanks to the selection criteria defined in the previous chapter, the reconstruction of the $t\bar{t}$ pair is rather challenging, since the decay products need to be correctly assigned to their corresponding top or anti-top quark. The choice of the variables used for the differential measurement is thus limited by the $t\bar{t}$ reconstruction and simultaneously aims to provide as high sensitivity as possible to either verify SM predictions or reveal signatures of new physics.

This chapter presents the algorithms used for the event reconstruction for both inclusive and differential measurement, and also discusses the choice of differential variables.

9.1 Dilepton channel

In the dilepton channel, a dedicated method, referred to as multi-hypothesis hadronic t/W reconstruction, is applied to fully hadronically decaying $t\bar{t}$ system to tests several hypotheses for the number of available and missing final state particles originating from the top quarks. An alternative approach targets the reconstruction of the all-hadronic $t\bar{t}$ system through a neural network.

9.1.1 Multi-hypothesis hadronic t/W reconstruction

Tree-level topology of the dilepton channel consists of two leptons from Z decay and six jets associated to $t\bar{t}$ system. However, thanks to the finite detector resolution for the jet energy and limited acceptance, non-negligible number of $t\bar{t}Z$ events contain at least one jet that either cannot be reconstructed or associated to one of the top quarks. To perform a reasonable kinematic reconstruction of such events, several hypotheses based on the number of missing jets in a given event, are tested.

Following categories of events are defined in order to test their final state topology:

- Both top quarks can be reconstructed (denoted as $2t$ category),
- One b jet is missing, resulting in one fully reconstructable top quark and one hadronically decaying W boson since the second top quark cannot be fully reconstructed ($1t1W$ category) ,
- If both b jets are missing, only the W bosons can be fully recostructed ($2W$ category),
- One top quark can be reconstructed but second W boson is missing some component ($1t$ category),
- Only one W boson can be reconstructed, while second cannot. The reconstruction of either top quark is not possible ($1W$ category).

Reconstructed jets can be assigned to the positions of t and W final states in multiple ways. For each possible jet-quark assignment the invariant masses of involved top quarks and W bosons are interpolated using corresponding pdf distributions built from the reconstruction-level events that are matched to their respective parton-level record. Each event is then investigated independently and tested against each of the categories defined above. This results in output weight representing compatibility of the given event with particular category. The output weight is evaluated for all possible jet-quark assignments and only the combination yielding the highest weight for given category is considered to be correct. Moreover, the dedicated vetoes are applied to the b tagging requirements, specifically optimized for each category, in order to take advantage of the overlap and inherent correlations between the hypotheses.

At the output of the method, each event has assigned five output weights (w_{2t}, w_{1t1W}, w_{2W}, w_{1t}, w_{1W}) corresponding to the defined categories. These weights are then used as the discriminating input variables for the training of the classification NN. As will be later discussed in chapter devoted to the dilepton analysis, these input variables are highly ranked and significantly improve NN performance.

9.1.2 All-hadronic $t\bar{t}$ reconstruction using neural network

To increase sensitivity of the classification NN in the dilepton channel, the reconstruction of the all-hadronic $t\bar{t}$ decay through dedicated Symmetry Preserving Attention Network (SPANet) [234], is employed. SPANet is an attention-based NN originally developed for reconstruction of all-hadronic decay of $t\bar{t}$ events. Since the $t\bar{t}Z$ dilepton channel has very similar topology SPANet can be easily repurposed for the $t\bar{t}Z$ process. Although extra Z boson affects kinematics of the $t\bar{t}$ pair, it does not have any implications on the SPANet functionality.

The aim of the network is to correctly match the jets present in the event to their corresponding partons. Construction of variables based on the jets predicted by SPANet could potentially improve performance of separation NN. Although there are quite a few methods capable of this jet-parton assignment (i.e. χ^2 method or KLFitter [235]), none of them achieve as good performance as SPANet.

SPANet Architecture

Since there is number of symmetries in the $t\bar{t}$ event topology [1], SPANet (as well as other methods) can benefit from them, reducing number of permutations to 90 for six-jet events. The number of permutations increase factorially with higher jet multiplicity, what consequently causes computational complexity (and thus also increases time necessary for the reconstruction) for methods depending on the number of permutations (like χ^2, KLFitter).

The schematic visualization of the SPANet architecture can be found in Figure 9.1. At the input to the network jets are independently embedded to produce latent space representation for each jet. The stack of transformer encoders is used to learn contextual relationships. The main advantage of the SPANet architecture is that it benefits not only from symmetries present in $t\bar{t}$ topology, but also from the tensor attention mechanism. SPANet improves run-time performance over baseline permutation methods by avoiding having to construct all valid assignment permutations. Instead, the jet-parton assignment problem is partitioned into sub-problems for each of the top quarks, and solved using novel technique called symmetric tensor attention.

As shown in architecture scheme in Figure 9.1, tensor attention are final layers in the SPANet chain. The input for these layers is formed from transformer-encoded jets $X_p \in \mathbb{R}^{N\times D}$, where N is total number of jets and D the size of the hidden representation [2].

[1] Namely invariance between quark and anti-quark from hadronic W boson decay and symmetry between top and anti-top quark.

[2] The dimensionality of the hidden representation is treated as a free parameter of the model - hyperparameter. In the implementation used for the dilepton analysis, the dimensionality was chosen to be 128 (specified further in the text).

The encoding of the jets is performed in order to create latent space representation for each jet, thus providing more comprehensible inputs for learning algorithms. A symmetric tensor attention layer contains a single parameter tensor $\theta \in \mathbb{R}^{D\times D\times D}$, where the rank of the tensor corresponds to the number of jets that should make up the top candidate (in this case three). This tensor is built from the inputs represented by embedded jets X_p and combines all relevant contextual relations that are necessary to predict the jet-parton assignments. The tensor attention layer then performs the following computations, using Einstein summation notation:

$$S^{ijk} = \frac{1}{2}(\theta^{ijk} + \theta^{jik}) \tag{9.1}$$

$$O^{ijk} = X^i_n X^j_m X^k_l S^{nml} \tag{9.2}$$

where i, j, k are indices of potential light jets and b jet. Obviously, symmetry tensor S is invariant under changing positions of light jets. Output tensor O is then normalized by performing three dimensional softmax transformation:

$$P(i, j, k) = \frac{e^{O^{ijk}}}{\sum_{i,j,k} e^{O^{ijk}}}, \tag{9.3}$$

yielding probability of the given permutation to be correctly assigned. In case of $t\bar{t}$ process the NN has only two output nodes, P_1 and P_2, representing probability of the jet combination to originate from top or anti-top quark, respectively.

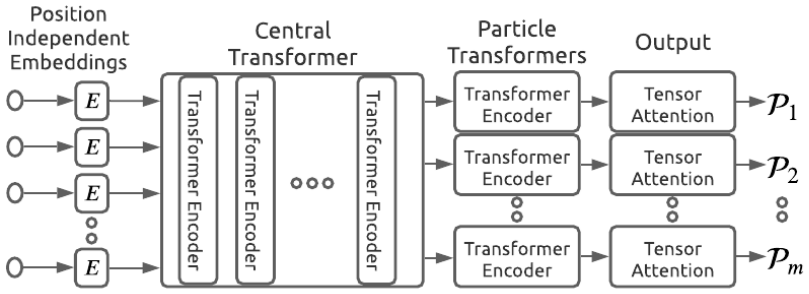


Fig. 9.1: High level structure of the SPANet architecture.

Implementation in the dilepton channel

To find the optimal SPANet model, the training is performed on various dataset configurations. To obtain as many events as possible for the training, all available signal samples are combined. The optimal training set thus consists of nominal MC signal sample, signal sample with alternative parton shower (Herwig 7), signal sample with varied `Var3c` parameter of the `A14` tune, and the LO signal sample designed for the

Effective Field Theory (EFT) interpretation (MADGRAPH5_aMC@NLO + PYTHIA 8). The whole dataset is split to training set which consists of 95% of available events, while remaining MC events are used for the evaluation. The event selection strictly follows the criteria used for the dilepton channel described in Section 8.1, while the training is performed only on events featuring at least 6 jets and at least one b jet. Furthermore, at least one top (antitop) quark has to be correctly matched to corresponding parton-level object [3]. The total number of events available for training sum up to ≈ 4.3 million.

Input for the network is in a form of $p_{\mathrm{T}}, \eta, \phi, m$ (mass) and b-tagging information (if the jet is b-tagged or not) of all jets present in the event. Correct jet assignments are also needed at the input and are constructed by matching detector-level objects to parton level.

Hyperparameters of the NN were optimised for $t\bar{t}$ model using SHERPA hyperparameter optimisation library [236] and are summarized in Table 9.1.

Hyperparameter	**Value**
Optimizer	AdamW
Training Epochs	50
Learning Rate	0.0015
Batch Size	256
Dropout Rate	0.1
Hidden Dimensionality	128

Table 9.1: Optimal values of hyperparameters used for the SPANet training.

Performance of trained model is evaluated based on the *event efficiency* that is defined as the fraction of events in which all jets associated with reconstructable particles are correctly assigned. Additionally, *top-quark efficiency* is defined per-particle, requiring one or two top quarks to be present at the parton level. The obtained efficiencies are shown in Table 9.2.

The output from the SPANet in a form of predictions for each event were used to construct kinematic variables connected with either reconstructed top or antitop quark. Since the predictions are extracted from 3-dimensional loss function, value of the loss function for best permutation could also provide separation power for signal-background discrimination. Full list of designed variables and their definitions can be found in Table 9.3.

[3] The matching between reconstruction- and parton-level objects is performed by requiring reconstructed jet to be within cone of radius $\Delta R = 0.4$ around parton.

		Event	$t\bar{t}Z$ model	
	$\mathbf{N_{jets}}$	**Fraction**	**Event**	**Top Quark**
All Events	==5	0.175	0.562	0.562
	==6	0.203	0.594	0.629
	≥7	0.242	0.530	0.598
	Inclusive	0.619	0.560	0.599
Complete Events	==5	0	-	-
	==6	0.029	0.771	0.823
	≥7	0.070	0.526	0.679
	Inclusive	0.099	0.598	0.722

Table 9.2: Event and top quark efficiencies $t\bar{t}Z$ SPANet model. *All Events* category refers to events with at least one top/antitop quark, while *Complete Events* consists only of events where both top and antitop quarks are present (according to the truth matching information). Column denoted as *Event Fraction* contains fractions of all MC events assigned to given event (either *All Events* or *Complete Events*) and jet multiplicity category.

9.2 Trilepton channel

In the trilepton channel, reconstruction of $t\bar{t}$ system is limited to the leptonically decaying top quark (reffered to as leptonic-side top quark in the following text), hence no full reconstruction is performed. This enables to define set of trilepton variables defined in the Section 9.4.

The assumption for the leptonic-side top quark reconstruction is that the neutrino emerging from the leptonic decay of the W boson represents the dominant source of $\mathbf{E}_{\mathrm{T}}^{\mathrm{miss}}$ in the event. Therefore the neutrino is assigned both magnitude $E_{\mathrm{T}}^{\mathrm{miss}}$ and the

Variable	**Definition**
$p_{\mathrm{T}}^{t1,\mathrm{spa}}, m^{t1,\mathrm{spa}}$	Transverse momentum and invariant mass of the first top quark
$p_{\mathrm{T}}^{t2,\mathrm{spa}}, m^{t2,\mathrm{spa}}$	Transverse momentum and invariant mass of the second top quark
$p_{\mathrm{T}}^{t\bar{t},\mathrm{spa}}, m^{t\bar{t},\mathrm{spa}}$	Transverse momentum and invariant mass of the $t\bar{t}$ pair
$\Delta\phi(Z, t_{\mathrm{spa}})$	Angular separation between Z boson and nearest (in ΔR) top quark
Spanet Loss 1	Value of the loss function for predicted combination of jets for first top quark
Spanet Loss 2	Value of the loss function for predicted combination of jets for second top quark

Table 9.3: Definition of the variables designed from the SPANet outputs.

angle ϕ^{miss} corresponding to the $\mathbf{E}_{\text{T}}^{\text{miss}}$ vector. The next assumption is made for the invariant mass of the neutrino and non-Z lepton, which is required to be compatible with the W boson mass of 80.385 GeV. This requirement allows for the determination of the z component of the neutrino momentum, denoted as $p_{\nu z}$, using the following quadratic form:

$$Ap_{\nu z}^2 + Bp_{\nu z} + C = 0, \tag{9.4}$$

where terms A, B and C are constructed from the transverse momentum and ϕ of the neutrino, fixed W boson mass m_W, and kinematic factors associated with the non-Z lepton.

Depending on the value of the discriminant when solving Eq.9.4, one, two or zero solutions for the neutrino momentum can be obtained. The negative value of the discriminant, and thus no real solution, was found in roughly 35% of all simulated $t\bar{t}Z$ events. In such cases, in order to find real solution for neutrino momentum, the missing transverse energy is sequentially decreased by 100 MeV until at least one real solution for $p_{\nu z}$ (value of the discriminant positive or equal zero) is obtained.

In case of one or two (if value of discriminant is positive) possible neutrino momenta are found, the procedure continues by constructing leptonic-side top quark candidates. The reconstructed W boson (summed four-momenta of neutrino candidate and non-Z lepton) is paired with the b-tagged jets at the reconstruction level, while only the two b jets with the highest b-tagging score are considered. Correct b jet associated to the W boson candidate is selected according to the ΔR separation between b jet and W boson, while the smaller separation is preferred. Note that in case of two neutrino candidates the selected b jets, according to the ΔR criterion, can differ. For the particle-level definition, the same procedure is performed, but the b jets are required to be ghost-matched to the b hadrons. In case of more than two ghost-matched b jets, only the two with the highest p_{T} are considered.

The reconstruction is concluded by assigning output weights to the above defined top-quark candidates based on the invariant mass distribution $m_{b\ell\nu}$, built from the correctly reconstructed top quarks in the simulated events. These "correct" top quarks are reconstructed from the parton-level neutrino, reconstruction-level non-Z lepton and b jet matched to the corresponding parton-level object. The output weight serves as a discriminating metric in case of two neutrino candidates. The two possible top-quark candidates are assigned output weights according to the interpolation of their masses from the $m_{b\ell\nu}$ distribution. The candidate with higher output weight is chosen as more compatible with the leptonically decaying top quark, while the other one is discarded.

9.3 Tetralepton channel

The reconstruction of the $t\bar{t}$ system in the tetralepton channel is limited only to the transverse plane. Again the same assumption as in the trilepton channel is considered for the two neutrinos arising from the $t\bar{t}$ pair, specifically that they carry away whole missing transverse momentum, $p_{\mathrm{T}}^{\mathrm{miss}}$. This quantity directly allows for the reconstruction of the $t\bar{t}$ pair in the transverse plane, enabling to define two differential variables ($p_{\mathrm{T}}^{t\bar{t}}$ and $|\Delta\phi(t\bar{t}, Z)|$), further described in the next section. This partial reconstruction was chosen in order to avoid separate determination of neutrino four-momenta in case of full $t\bar{t}$ reconstruction.

To define the variables mentioned above, it is necessary to distinguish between lepton pair associated with Z boson and $t\bar{t}$ and to identify the two b jets arising from $t\bar{t}$ decay. The identification of the Z-like and non-Z lepton pair is already achieved with the selection criteria described in the Section 8.3.

Selection of b jets follows the same rules as for the trilepton reconstruction. At the reconstruction level the b jets are identified with the two jets featuring the highest b-tagging output weight, while they do not necessarily need to be b tagged. At the particle level the ghost matching of jets to b hadrons is used. In case of only one particle-level b jet, or if there are more than two, the p_{T} ordering is employed in order to choose remaining b jet (if there is only one particle-level b jet) or the two correct b jets (if there are more than two particle-level b jets), while the jets with the highest p_{T} are preferred.

The reconstruction of the $t\bar{t}$ system at both particle and reconstruction level is then straightforward. The momentum vector in the transverse plane, $\mathbf{p}_{\mathrm{T}}^{t\bar{t}}$, is built as a vector sum of missing transverse energy (both magnitude and angle component), with the two non-Z leptons and selected jets.

9.4 Variables for the differential cross-section measurement

Differential cross sections in the trilepton and tetralepton channels are measured as a function of different variables, defined in this section. Two variables, namely the absolute value of the rapidity and transverse momentum of the Z boson are unfolded in the combination of trilepton and tetralepton channels, whereas others are defined separately for the particular channel. Jet multiplicity variable, denoted as N_{jets}, is used separately in both channels although it could be in principle defined for the combined $3\ell+4\ell$ channel. The reason why it is unfolded separately is that the nominal number of jets at the tree level differs between the two channels, and thus the separate treatment

makes this variable more meaningful from the physics point of view. The whole set of variables, their definition and the channel in which they are used for the unfolding, are summarized in the Table 9.4. The selection of the variables was motivated by the aspects discussed in the following text.

In the trilepton channel, the easily reconstructed transverse momentum of the non-Z lepton (denoted as $p_{\mathrm{T}}^{\ell\mathrm{non}-Z}$) is chosen to probe the modeling of the transverse momenta of the top quarks and their decay products in MC simulations, as well as provides a sensitivity to the modeling of ISR and FSR. Leptonic-side top reconsturction allows to define absolute angular and rapidity separation between the Z boson and the leptonically decaying top quark, labeled as $|\Delta\phi(Z, t_{\mathrm{lep}})|$ and $|\Delta y(Z, t_{\mathrm{lep}})|$, respectively. These two variables combine kinematics of Z boson and top quark and thus directly probe the $t\bar{t}Z$ vertex providing sensitivity for testing the $t - Z$ coupling, what could reveal potential non-SM physics effects.

The same potential offers also $|\Delta\phi(t\bar{t}, Z)|$ variable defined in the tetralepton channel, as well as variables defined in the combined $3\ell + 4\ell$ channel, which involve Z boson kinematics. In analogy to the $p_{\mathrm{T}}^{\ell\mathrm{non}-Z}$ in the trilepton channel, the tetralepton channel includes transverse momentum of the $t\bar{t}$ system, $p_{\mathrm{T}}^{t\bar{t}}$, providing sensitivity to MC modeling of the hard-scatter process, as well as ISR, FSR and tuning of the parton showering. The same sensitivity to the MC modeling is offered by the absolute angular separation between the two leptons from the $t\bar{t}$ pair, $|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)|$, which additionally provides sensitivity to potential non-SM effects that could modify the spin correlations between the top quarks.

Finally the jet multiplicity variable unfolded separately in both channels is chosen for its sensitivity to modeling of various MC generators including simulation of hadronization and parton showering.

Variable	Definition	Channel
N_{jets}	Number of reconstructed jets with $p_T > 25$ GeV and η<2.5	3ℓ and 4ℓ
$p_{\text{T}}^{\ell\text{non}-Z}$	Transverse momentum of the lepton not associated with the Z boson	3ℓ
$\lvert\Delta\phi(Z, t_{\text{lep}})\rvert$	Absolute azimuthal separation between the Z boson and the leptonically decaying top quark	3ℓ
$\lvert\Delta y(Z, t_{\text{lep}})\rvert$	Absolute rapidity separation between the Z boson and the leptonically decaying top quark	3ℓ
$\lvert\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)\rvert$	Absolute azimuthal separation between the two leptons associated with the $t\bar{t}$ pair	4ℓ
$p_{\text{T}}^{t\bar{t}}$	Transverse momentum of the reconstructed $t\bar{t}$ system	4ℓ
$\lvert\Delta\phi(t\bar{t}, Z)\rvert$	Absolute azimuthal separation between the Z boson and the $t\bar{t}$ system	4ℓ
$\lvert y^Z\rvert$	Absolute rapidity of the Z boson	$3\ell + 4\ell$
p_{T}^Z	Transverse momentum of the Z boson	$3\ell + 4\ell$

Table 9.4: Variables selected for the differential cross section measurements. As indicated, measurements for some variables are only performed in either the trilepton or tetralepton SR. The jet multiplicity (N_{jets}) is measured separately in both channels, whereas for the Z-related observables ($|y^Z|$ and p_T^Z), both trilepton and tetralepton SRs are combined.

Chapter 10

Differential cross section measurements

This chapter presents the strategy and results of the differential cross section measurements performed in the trilepton and tetralepton SRs, as a part of the $t\bar{t}Z$ analysis published in Ref. [10].

10.1 Analysis strategy

The measured $t\bar{t}Z$ distributions in their respective SRs can be converted into differential cross sections as a function of these variables through the process of unfolding, described in Section 7.1. This allows for the correction of detector effects, as well as signal efficiency and acceptance with respect to a given region of the phase space, and ultimately yields a measurement that can be directly compared to various theoretical predictions, to CMS results or given a sensitive EFT interpretation.

Variables considered for the measurements are described in Section 9.4 and are unfolded to particle and parton level [1] as defined in Sections 8.2 and 8.3.

The differential cross-section is calculated using the following equation:

$$\frac{\mathrm{d}\sigma_{t\bar{t}Z}}{\mathrm{d}X^i} = \frac{1}{\mathcal{L}\cdot\mathcal{B}\cdot\Delta X^i\cdot\epsilon^i_{\mathrm{eff}}}\cdot\sum_j\left[\mathcal{M}^{-1}\right]_{ij}\cdot f^j_{\mathrm{acc}}\cdot\left(N^j_{\mathrm{obs}}-N^j_{\mathrm{bkg}}\right), \tag{10.1}$$

where i is the index of the bin for the observable X with bin width ΔX^i, $\mathcal{L}$ is the integrated luminosity and $\mathcal{B}$ is the branching ratio which is relevant only for the parton-level measurements as described below. The background contribution estimated from MC, N^j_{bkg}, is subtracted from the observed data, N^j_{obs}, in a bin j, to provide the estimated observed signal in that bin. This quantity is then corrected by an acceptance correction term, f^j_{acc}, which accounts for events outside the fiducial phase space satisfying the

[1] Note that jet multiplicity variable N_{jets} is unfolded only to particle level since it is not defined at parton level.

reconstruction-level selection, before being passed through the migration matrix with elements $\mathcal{M}_{ij}$, which provides the bin-to-bin migrations of events between a fiducial bin i and a reconstruction-level bin j. The efficiency correction, ϵ_{eff}, extends the result to events in the fiducial phase-space not being reconstructed in the signal region.

The two correction terms, acceptance and efficiency are defined as follows:

$$f_{\text{acc}} = \frac{N^{\text{reco}\wedge\text{truth}}}{N^{\text{reco}}}, \tag{10.2}$$

$$\epsilon_{\text{eff}} = \frac{N^{\text{reco}\wedge\text{truth}}}{N^{\text{truth}}}, \tag{10.3}$$

where N^{reco} stands for the number of events passing reconstruction-level selection, N^{truth} refers to number of events passing fiducial selection (either at particle or parton level), and the logical symbol $\wedge$ takes their intersection. These corrections, as well as migration matrices, are derived exclusively from the signal MC, and are shown in Figures 10.1-10.3. The same figures depict also the reconstruction-level distributions of the unfolded variables. Note, that main body of this thesis contains only figures for one representative variable in each analysis channel ($|\Delta\phi(Z, t_{\text{lep}})|$ in trilepton, $p_{\text{T}}^{t\bar{t}}$ in tetralepton, and p_{T}^{Z} in their combination), while the figures for the rest of variables can be found in respective Appendix A.1.

Table 10.1 summarizes the branching ratios ($\mathcal{B}$) which are applied in Eq. 10.1 for the parton-level measurements [2] in order to extrapolate the measurements to be inclusive in terms of $t\bar{t}$ and Z boson decays.

Channel	Branching Ratio ($\mathcal{B}$)	Details
3ℓ	0.019279	$t\bar{t} \to bq\bar{q}\bar{b}\ell\nu_\ell, Z \to \ell\ell$ ($\ell = e, \mu$ only)
4ℓ	0.003039	$t\bar{t} \to b\ell\nu_\ell\bar{b}\ell\nu_\ell, Z \to \ell\ell$ ($\ell = e, \mu$ only)
$3\ell + 4\ell$	0.022318	combination of above

Table 10.1: Summary of the branching ratio values used in Eq. 10.1 for the parton-level measurements.

The crucial step in Eq. 10.1 is the inversion of the migration matrix, which can lead to unsatisfactory levels of precision in the case of non-diagonal migration matrices. The unfolding method is therefore chosen to be the iterative ("Bayesian") method [219, 220] as described in Section 7.1.1, which circumvents the problem of inverting the matrix numerically by approximating its inverse iteratively, through repeated applications of Bayes formula. The migration matrix from Eq. 10.1 corresponds to posterior probability entering the estimator of the truth-level spectrum as defined in Eq. 7.3. This unfolding method is implemented as part of the RooUnfold package [237].

[2] For particle level the branching ratios are taken to be 1.

It should be noted that $WZ/ZZ+$ light jets component of the MC background, which contributes to the N_{bkg} term in Eq.10.1, is corrected for potential miss-modeling by applying additional normalization factor derived from the inclusive profile likelihood fit. These normalization factors were obtained from the statistical-only fit in the $3\ell+4\ell$ combined channel, where only statistical uncertainties on data and MC samples were considered. All systematic uncertainties were omitted in order to avoid a potential bias introduced by the pulls of the normalization factors from the nuisance parameters. The following factors with their associated uncertainties (considered as additional source of uncertainty for the affected background processes in the differential measurements) were obtained:

$$\mathcal{N}_{WZ+l}^{\text{stat-only}} = 0.87 \pm 0.03 \tag{10.4}$$

$$\mathcal{N}_{ZZ+l}^{\text{stat-only}} = 1.07 \pm 0.05. \tag{10.5}$$

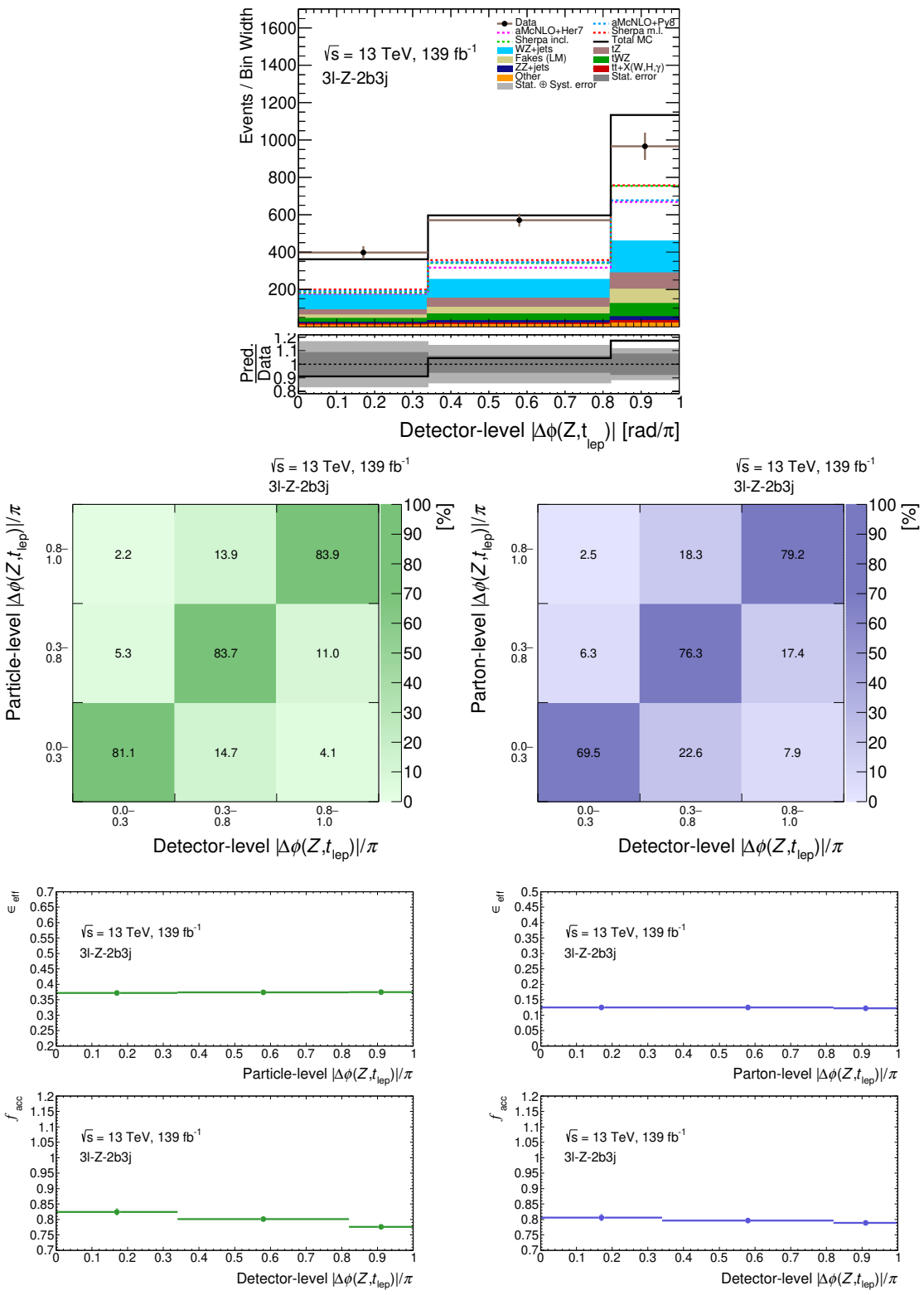


Fig. 10.1: Reconstruction-level distribution (top), together with particle-level (left) and parton-level (right) migration matrices (middle) and efficiency/acceptance corrections (bottom) for $|\Delta\phi(Z, t_{\text{lep}})|$ in the trilepton channel. The dashed lines in the reconstruction-level distribution represent nominal signal (labeled as *aMcNLO+Py8*), that is used also in the total MC prediction, as well as various alternative signal MC predictions (alternative parton shower - *aMcNLO+Her7*, and *Sherpa* samples).

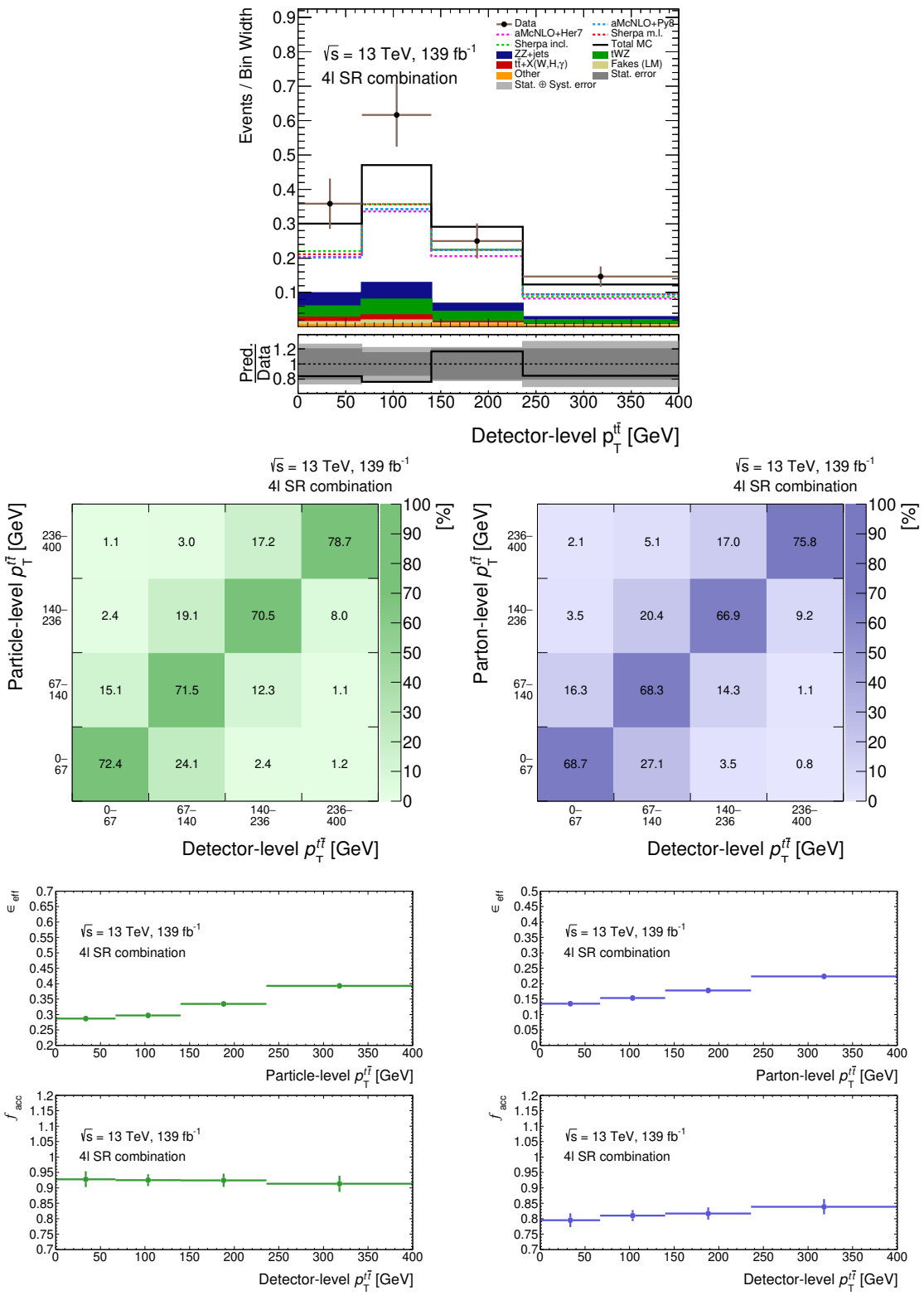


Fig. 10.2: Reconstruction-level distribution (top), together with particle-level (left) and parton-level (right) migration matrices (middle) and efficiency/acceptance corrections (bottom) for $p_{\mathrm{T}}^{t\bar{t}}$ in the tetralepton channel. The dashed lines in the reconstruction-level distribution represent nominal signal (labeled as *aMcNLO+Py8*), that is used also in the total MC prediction, as well as various alternative signal MC predictions (alternative parton shower - *aMcNLO+Her7*, and *Sherpa* samples).

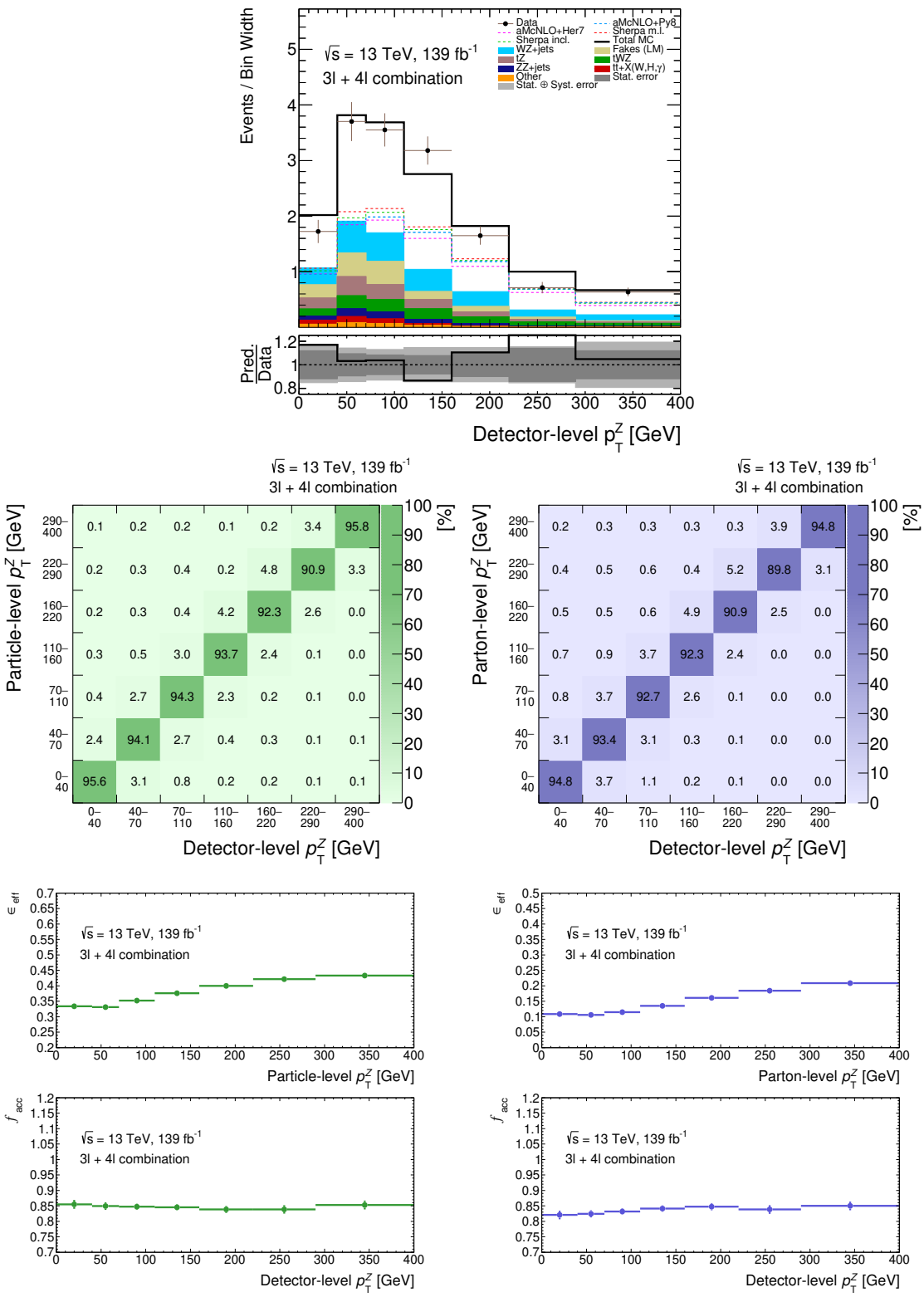


Fig. 10.3: Reconstruction-level distribution (top), together with particle-level (left) and parton-level (right) migration matrices (middle) and efficiency/acceptance corrections (bottom) for p_{T}^{Z} in the combined 3ℓ+4ℓ channel. The dashed lines in the reconstruction-level distribution represent nominal signal (labeled as *aMcNLO+Py8*), that is used also in the total MC prediction, as well as various alternative signal MC predictions (alternative parton shower - *aMcNLO+Her7*, and *Sherpa* samples).

10.2 Optimizations and tests of the unfolding procedure

As outlined in Section 7.1.1, the unfolding technique requires optimization of binning for particular variables in order to achieve optimal sensitivity of the measurements. Moreover, the unfolding procedure relies on the use of iterations, while their number serves as a regularization parameter and needs to be optimized as well. Both optimizations are presented in this section. Furthermore, in order to verify the stability and expected behavior of the unfolding process, several tests, described in the following sections, are performed.

10.2.1 Binning optimization

Algorithm that aims to optimize the number of bins for each particular differential variable has been developed taking into account two main aspects: the lowest possible statistical uncertainty, while keeping the reasonable diagonality of the migration matrix to ensure stable unfolding. Naturally, it is preferred to use as many bins as possible in order to achieve high sensitivity for the underlying physics effects.

At the beginning of the optimization procedure, the maximum tolerable statistical uncertainty in each bin and minimum acceptable value of the diagonal elements of migration matrix, are chosen. The algorithm then performs the optimization giving optimal bin ranges from 3 to 10 bins. The procedure is repeated for the maximum uncertainty in a range from 30% to 50% with 10% step and simultaneously for minimum diagonal elements in a range 50% to 90% using 5% increment. The final binning choice is then based on the physical motivation for particular variable, while taking into account also results of unfolding stability tests.

The optimization algorithm starts from the fine binning (100 equidistant bins) for the given variable and merges the fine bins from left to right (or vice-versa if the variable peaks at higher values) until the requirements defined above are fulfilled for the bin which is being merged. Then such a merged bin is fixed and procedure continues with merging next bunch of fine bins. The last merging process can result in invalid bin (extreme right or left) when there are no more events to fulfill both requirements [3]. If such a bin is produced, it can be either accepted (if it does not cause any further issues in unfolding process and its statistical uncertainty is acceptable), or merged with the neighboring valid bin.

In case of variable with highly diagonal migration matrix (over 90% for each diagonal element), the algorithm above may not be the optimal approach for choosing final

[3] Overflow and underflow bins need to be included in the merging process to ensure that all the events are used in the optimization.

binning. For such variables, the requirement on minimal value of diagonal elements is not relevant, and a modified algorithm, which looks for bins with approximately equal statistical uncertainty, is used. Algorithm merges fine bins similarly to the method described above, until it finds bins with approximately the same statistical uncertainty. This results in possible binnings with 3 to 9 valid bins. The final number of bins is then chosen according to the obtained statistical uncertainty and the physical motivation of particular variable.

The binning choice can in principle differ between particle-level and parton-level unfolding, but for simplicity they are chosen to be identical, while assuming negligible effects on the final results in case of different bin ranges. The final bin ranges for the set of variables to be unfolded were mostly selected according to the description outlined above and are summarized in Table 10.2. In several cases some slight adjustments were made with brief reasoning given below:

- N_{jets} - this variable was not optimized with any of the described algorithms. To offer the highest sensitivity for tuning MC generators and theory comparisons, and since this variable is discrete, the binning was chosen to be as fine as possible while merging higher-multiplicity bins to maintain a reasonable statistical uncertainty. The jet multiplicity for the lowermost bin was chosen according to the minimum number of jets for the given topology (which differs for the trilepton and tetralepton channel).

- $|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)|, p_{\text{T}}^{\ell\text{non}-Z}$ - since the migration matrices for these variables are highly diagonal, the final binning was chosen according to second method (to yield similar relative statistical uncertainties in each bin) and the optimal number of bins (4 in all cases) was chosen in order not to exceed a 25% statistical uncertainty.

- $|\Delta\phi(t\bar{t}, Z)|, |\Delta\phi(Z, t_{\text{lep}})|, |\Delta y(Z, t_{\text{lep}})|$ - since migration matrices for these variables exhibit large off-diagonal elements, the first method (with requirements on the maximum statistical uncertainty and diagonal elements) was used, where in each case two bins (from the original 4 considered) were merged to increase the diagonality of the migration matrix.

- $p_{\text{T}}^{t\bar{t}}$ - for this variable the first method was used. The final two bins were merged to enhance the diagonality of the migration matrix while keeping the statistical uncertainty under 25%.

- $|y^Z|, p_{\text{T}}^Z$ - larger statistics based on the combined selection allowed for a finer binning for these variables. The bin ranges were chosen in such a way that the diagonal elements of the migration matrices lie in the 85-90% range. Under this constraint 8 bins were initially possible. For p_{T}^Z the two uppermost bins were

merged to bring the statistical uncertainty to a level more consistent with the remaining bins.

Channel	Variable	Bin ranges	Units
3ℓ	$p_{\mathrm{T}}^{\ell \mathrm{non}-Z}$	[0, 40, 65, 105, 200]	[GeV]
	$\lvert\Delta\phi(Z, t_{\mathrm{lep}})\rvert$	[0, 0.34, 0.82, 1]	[rad/π]
	$\lvert\Delta y(Z, t_{\mathrm{lep}})\rvert$	[0, 0.57, 1.8, 4.5]	-
	N_{jets}	[3, 4, 5, 11]	-
4ℓ	$\lvert\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)\rvert$	[0, 0.31, 0.59, 0.8, 1]	[rad/π]
	$p_{\mathrm{T}}^{t\bar{t}}$	[0, 67, 140, 236, 400]	[GeV]
	$\lvert\Delta\phi(t\bar{t}, Z)\rvert$	[0, 0.73, 0.93, 1]	[rad/π]
	N_{jets}	[2, 3, 4, 5, 8]	-
$3\ell/4\ell$	$\lvert y^Z\rvert$	[0, 0.10, 0.25, 0.42, 0.64, 0.84, 1.08, 1.33, 2.50]	-
	p_{T}^Z	[0, 40, 70, 110, 160, 220, 290, 400]	[GeV]

Table 10.2: Optimal binnings chosen for the variables used in the differential cross section measurements. The last (first) bin includes always also overflow (underflow) throughout whole analysis.

10.2.2 Optimization of the number of iterations

In order to determine ideal number of iterations, serving as a regularization parameter for the iterative procedure of the IBU, the whole nominal signal MC sample is employed for deriving the migration matrix and unfolding corrections (efficiency and acceptance). The optimization process is based on producing so-called *pseudo-experiments* by smearing events according to the Poisson distribution to simulate the statistical fluctuations. Each of the pseudo-experiments (their total number is 5000) results in smeared reconstruction-level distribution for particular variable, which is subsequently unfolded with various number of iterations (from one to ten).

For every number of iterations N, the unfolded distribution for given pseudo-experiment is compared to the same spectrum, but unfolded with $N-1$ iterations. In case of one iteration the unfolded result is compared to truth-level distribution (either particle- or parton-level). Metric that is used for the comparisons is χ^2/NDF, where NDF denotes number of degrees of freedom. Additionally, the relative statistical uncertainty of the unfolding (relative with respect to the truth-level bin content) is also taken into account. Both the χ^2/NDF values and relative unfolding uncertainty are averaged over bins and pseudo-experiments, and their values are used to built distributions shown in Figure 10.4 (only sample variables are shown here while the rest can be found in Appendix A.2).

The optimal number of iterations is then determined from these distributions as a value at which the χ^2/NDF is stabilized at an approximately constant value, while keeping the statistical uncertainty of the unfolding as low as possible. The optimization is performed separately for the particle and parton level. The obtained ideal numbers of iterations for particular variables are reported in Table 10.3 and are used hereafter for all other unfolding tests described in this chapter, as well as for the final differential results.

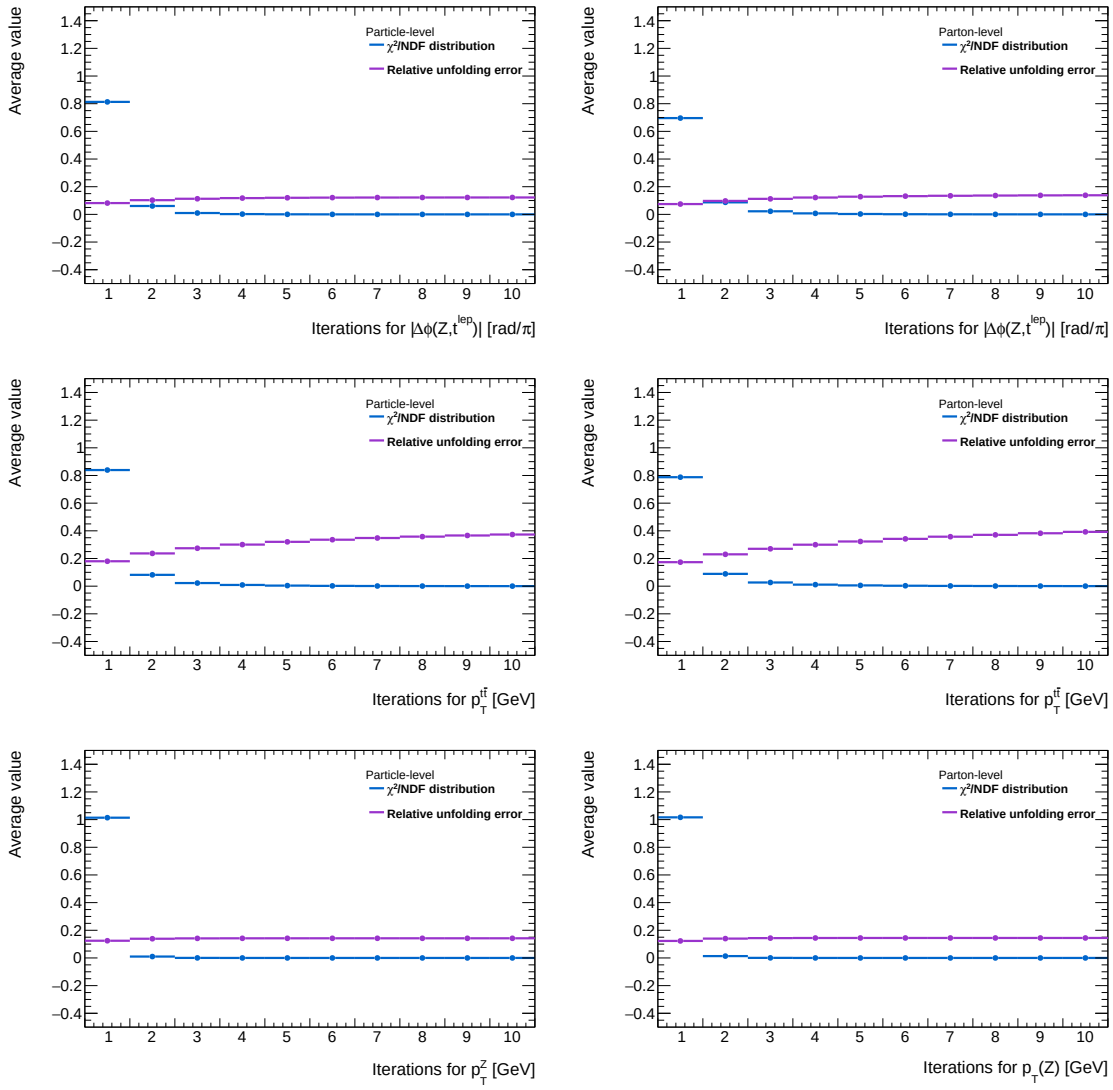


Fig. 10.4: Particle-level (left) and parton-level (right) summary plots for the optimization of the number of iterations for $|\Delta\phi(Z, t_{\text{lep}})|$ in the trilepton (first row), $p_{\text{T}}^{t\bar{t}}$ in the tetralepton (second row) and p_{T}^{Z} in the combined $3\ell + 4\ell$ channel (third row).

Channel	Variable	Particle	Parton
3ℓ	$p_{\mathrm{T}}^{\ell \mathrm{non}-Z}$	3	3
	$\lvert\Delta\phi(Z, t_{\mathrm{lep}})\rvert$	4	5
	$\lvert\Delta y(Z, t_{\mathrm{lep}})\rvert$	4	5
	N_{jets}	4	-
4ℓ	$\lvert\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)\rvert$	2	3
	$p_{\mathrm{T}}^{t\bar{t}}$	3	4
	$\lvert\Delta\phi(t\bar{t}, Z)\rvert$	5	5
	N_{jets}	5	-
$3\ell/4\ell$	$\lvert y^Z\rvert$	3	3
	p_{T}^Z	3	3

Table 10.3: Ideal number of iterations used for the iterative unfolding method, as obtained for each differential variable.

10.2.3 Closure tests

In order to prove stability of the IBU procedure towards statistical fluctuations present in data, so-called *closure* tests are performed. The test is again based on the pseudo-experiments (their total number is 10000), which are constructed by smearing the reconstruction-level distributions of the signal MC samples. Smearing is performed bin-wise by drawing a random number which follows a Poisson distribution with a λ parameter equal to particular bin content. Each of the pseudo-experiments is then unfolded using migration matrix and unfolding corrections derived from the nominal un-smeared sample. The stability of the unfolding is evaluated by calculating *pull* for each pseudo-experiment j in each bin of given variable.

The frequently used definition of the pull assumes highly populated bins of the unfolded spectra, what is not the case in the $t\bar{t}Z$ trilepton and tetralepton analysis. As a consequence, the pull tests employing the original pull definition yield unsatisfactory results for this analysis and thus the definition needs to be modified to account for the limited statistics. The reasoning why the original definition results in the failed tests and the motivation for its modification is described in detail in Appendix A.3.1.

The modified pull that is used for the closure tests presented in this section, is defined as follows:

$$p_i^j = \frac{x_i^j - t_i}{\sigma(x_i^{\mathrm{nominal}})} \tag{10.6}$$

where x_i^j corresponds to unfolded value for pseudo-experiment j in bin i, $\sigma(x_i^{nominal})$ is the uncertainty on the unfolded un-smeared nominal distribution in bin i, and t_i represents the truth-level value. The sample pull distributions for the first two bins of the $|\Delta\phi(t\bar{t}, Z)|$ variable in the tetralepton channel can be found in Figure 10.5.

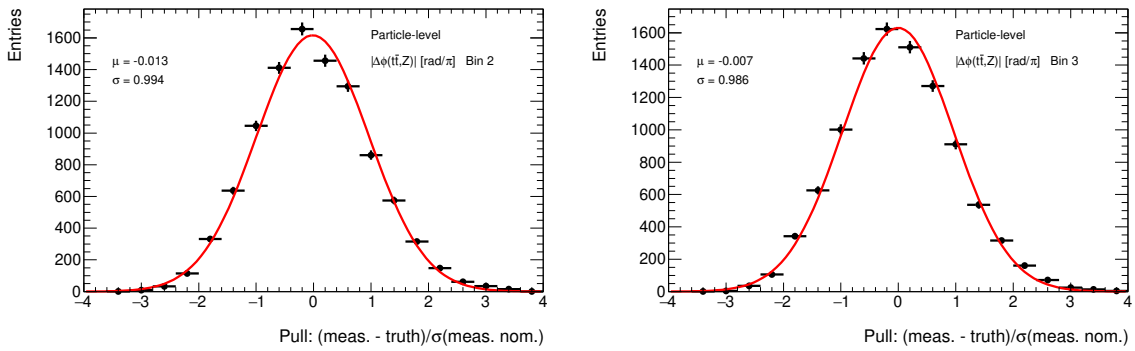


Fig. 10.5: Sample pull distributions for the second (left) and third (right) bin of $|\Delta\phi(t\bar{t}, Z)|$ for particle-level in the tetralepton channel. The solid red line represents the Gaussian fit.

In the original version of the closure tests, presented in Appendix A.3.1, a fit is performed to such distributions based on a Gaussian function which is then superimposed on the same plot. The corresponding mean and standard deviation (width) of the Gaussian fits are quoted for each bin of the tested variable. The unfolding procedure is considered to be stable if the pull means are consistent with zero and standard deviation with one.

However, as discussed in Appendix A.3.1, the Gaussian fit is not suitable for analysis with limited statistics since the pull distribution does not follow Gaussian distribution, but rather discrete Poisson distribution. Moreover, the fitted values corresponding to the pull mean and standard deviation can depend on the chosen binning for the pull distribution. For these reasons, the Gaussian fit is replaced by the calculation of arithmetic mean μ and its corresponding root mean square σ according to the following standard formulas:

$$\begin{aligned} \mu &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N} x_i, \\ \sigma &= \sqrt{\tfrac{1}{N}\textstyle\sum_{i=1}^{N}(x_i-\mu)^2}, \end{aligned} \tag{10.7}$$

where x_i is the pull for pseudo-experiment i and sum goes over all pseudo-experiments. The errors on these estimators are given by:

$$\begin{aligned} \delta\mu &= \tfrac{\sigma}{\sqrt{N}}, \\ \delta\sigma &= \tfrac{\sigma}{\sqrt{2(N-1)}}. \end{aligned} \tag{10.8}$$

Results of the pull tests for sample variables when using the modified definition of the pull (Eq. 10.6) and the extraction of pull mean and width values according to Eq. 10.7 are shown in Figure 10.6. These plots show significant improvements when compared to results employing original strategy described in Appendix A.3.1, and prove that unfolding procedure passes the stability test on the statistical fluctuations in data.

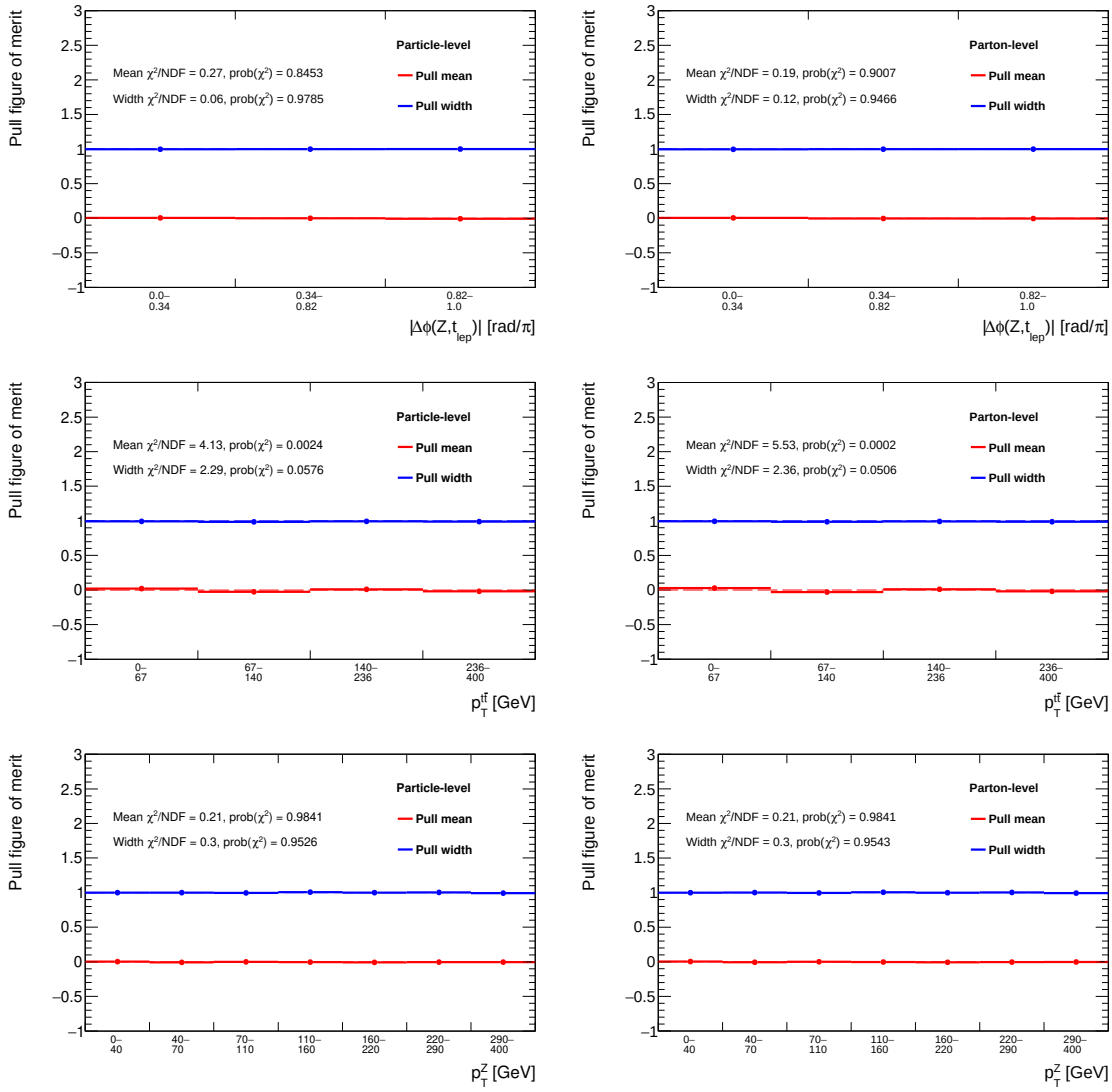


Fig. 10.6: Particle-level (left) and parton-level (right) summary plots for the updated pull tests (using modified pull definition from Eq. 10.6 and formulas from Eq. 10.7 for extraction of mean and width) performed for $|\Delta\phi(Z, t_{\text{lep}})|$ in the trilepton (first row), $p_{\text{T}}^{t\bar{t}}$ in the tetralepton (second row) and p_{T}^{Z} in the combined $3\ell + 4\ell$ channel (third row). Corresponding plots for the rest of differential variables can be found in Appendix A.3.

The χ^2/NDF values quoted in the Figure 10.6 are given by:

$$\chi^2/\text{NDF} = \frac{1}{N_{\text{bins}}} \sum_{i=1}^{N_{\text{bins}}} \frac{\mu_i^2}{(\delta\mu_i)^2} \quad \text{(pull mean)} \tag{10.9}$$

and similarly,

$$\chi^2/\text{NDF} = \frac{1}{N_{\text{bins}}} \sum_{i=1}^{N_{\text{bins}}} \frac{(\sigma_i - 1)^2}{(\delta\sigma_i)^2} \quad \text{(pull width)} \tag{10.10}$$

where μ_i, $\delta\mu_i$, σ_i, and $\delta\sigma_i$ are the estimators of the means and standard deviations (and the estimators of their corresponding uncertainties) for the given bin i, and the NDF corresponds to the number of bins (N_{bins}).

Closure tests using independent datasets

To further ensure that the closure tests presented above are not anyhow biased by using the same MC sample for building migration matrix and unfolding corrections, as well as for actual unfolding, the improved version of these tests using independent datasets is presented in the following. The only difference with respect to original closure tests lies in the usage of half of the signal MC sample for evaluating corrections and constructing migration matrix (referred to as *training* sample), while the second half is used for producing pseudo-experiments, which are subsequently unfolded (referred to as *testing* sample). The splitting of the whole MC sample is performed by generating random number between 0 and 1 for each event and assigning them to either first half (if random number is less than 0.5) or to second half (if random number is more than 0.5).

The pull distributions obtained from the ensemble of pseudo-experiments are constructed using the updated definition from Eq. 10.6 and the pull means and widths, evaluated according to Eq. 10.7, are summarized in plots shown in Figure 10.7. Pull means at these plots (and also plots for other differential variables, which can be found in Appendix A.3) exhibit substantial deviations from zero, what would suggest that the unfolding method is biased and thus not stable when using independent datasets. However, it can be proved that the deviations are not caused by any bias, but again by limited amount of the data available for $t\bar{t}Z$ analysis. When the MC is split into two halves, they result in significantly different reconstruction-level distributions as can be seen in Figure 10.8. By comparing the shape of the ratio between training and testing reconstruction-level distributions (black lines in Figure 10.8) with the shape of the pull means from Figure 10.7, it can be seen that they are similar, suggesting that the effects are strongly correlated.

To quantify the degree to which this bias is truly caused by statistical limitations in the MC sample, the dedicated tests, described in the next Section 10.2.4, were performed. The impact of the limited number of events in the MC sample, which results from these dedicated tests, are applied to the pull means and are presented in a form of gray band shown in Figure 10.7. As can be seen, this effect covers the observed deviations of the pull means, proving that used unfolding method is not biased.

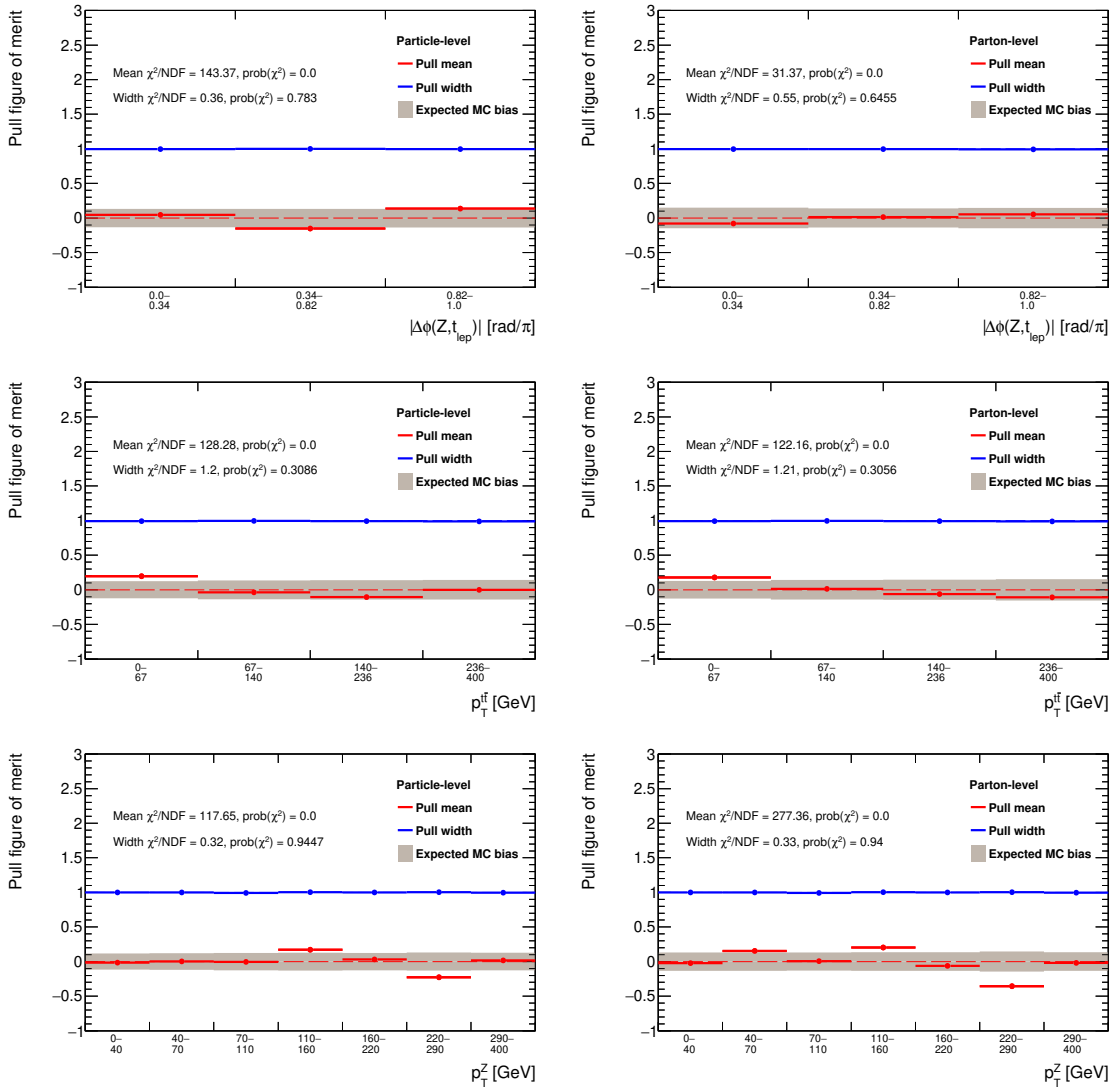


Fig. 10.7: Particle-level (left) and parton-level (right) summary plots for the pull tests using independent sample for building migration matrix and unfolding corrections, performed for $|\Delta\phi(Z, t_{\text{lep}})|$ in the trilepton (first row), $p_{\text{T}}^{t\bar{t}}$ in the tetralepton (second row) and p_{T}^{Z} in the combined 3ℓ+4ℓ channel (third row). Gray band represents expected bias caused by the limited statistics of the MC sample. Corresponding plots for the rest of differential variables can be found in Appendix A.3.

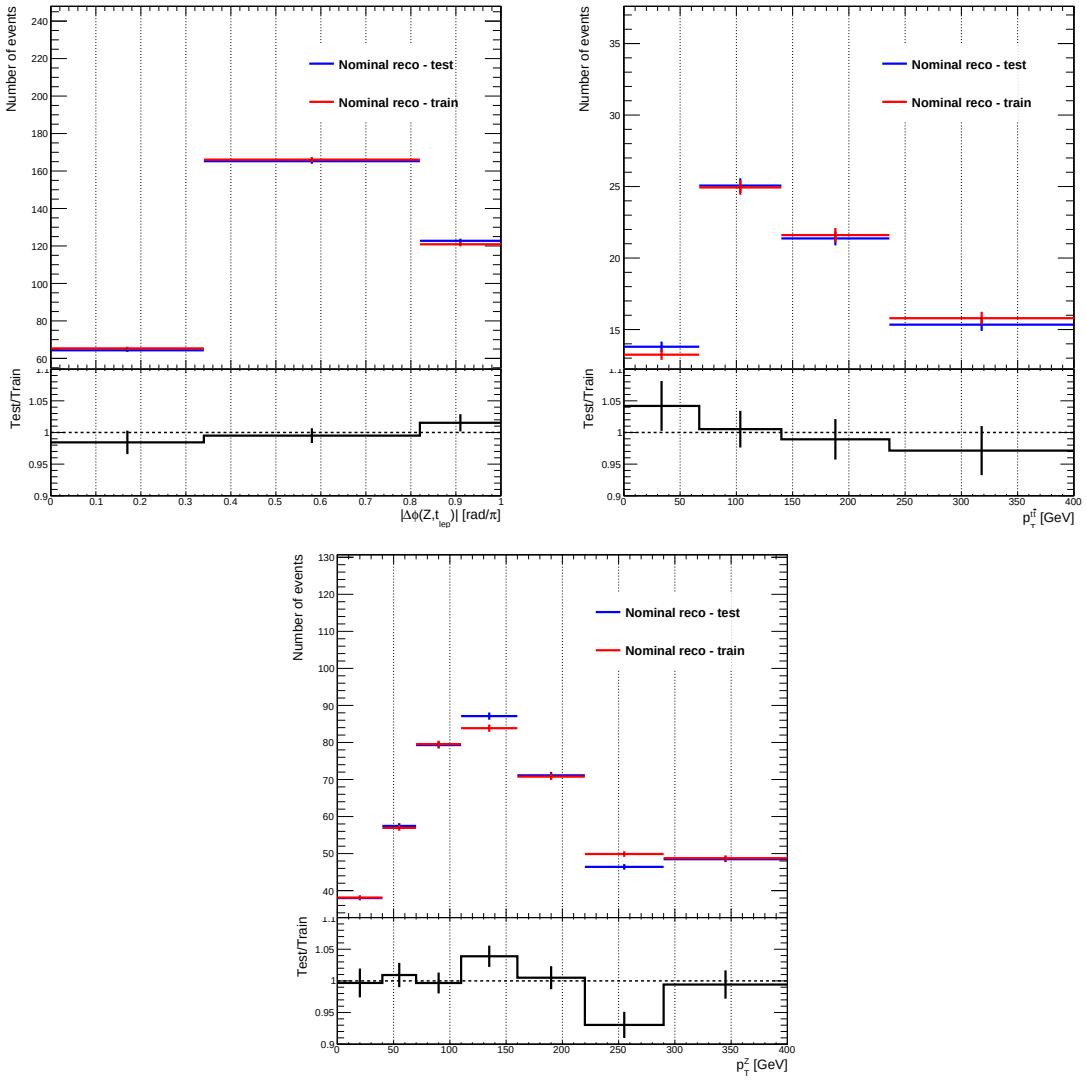


Fig. 10.8: Reconstruction-level distributions for the two independent samples used for the closure tests. Sample denoted as "train" is used for constructing migration matrix and evaluating unfolding corrections, while "test" sample is unfolded. The plots show distributions for $|\Delta\phi(Z, t_{\text{lep}})|$ in the trilepton, $p_{\text{T}}^{t\bar{t}}$ in the tetralepton and p_{T}^{Z} in the combined $3\ell + 4\ell$ channel. Corresponding plots for the rest of differential variables can be found in Appendix A.3.

10.2.4 Impact of statistically limited Monte Carlo samples

As outlined in the previous section, signal MC sample for the differential measurements contains limited number of events, what results in instability of the unfolding procedure when the sample is split. The statistical fluctuations of this MC sample cause non-negligible effects in the migration matrix and unfolding corrections, what can be identified with non-closure observed in the pull tests with independent datasets. Since the actual unfolding on measured data relies on the evaluation of migration matrix and efficiency/acceptance corrections from the MC, this non-closure effect, estimated in this section, needs to be taken into account and applied as additional source of uncertainty in the differential measurements.

The estimation procedure is based on similar approach as the closure test using independent datasets. Again the MC sample is randomly split into training and testing sample, where the first half is used for building migration matrix and unfolding correction, while from the second half, 100000 pseudo-experiments based on bin-wise Poisson smearing, are produced. Unfolded pseudo-experiments are then used to built distributions of relative difference between unfolded and truth-level value, *(unfolded-truth)/truth*, for each bin of the differential variables. Mean of such distribution, example of which is shown in Figure 10.9 for the last bin of p_{T}^{Z}, is taken as an non-closure effect corresponding to the particular splitting.

The procedure described above is then repeated for 2000 different splittings (using different random seed for generating splitting weights), yielding 2000 mean values of mentioned metric. The root mean square (RMS) of the distribution of means represents the expected non-closure effect due to statistical limitation of MC sample for particular bin and variable. Since the obtained values, presented in Table 10.4 for sample variables [4], are evaluated using half of the whole MC sample, the final uncertainty applied to the unfolded data needs to be multiplied by factor $\sqrt{2}$. It should be noted that this uncertainty caused by the statistical fluctuations of the MC sample is substantially smaller (one order of magnitude on average) than the expected statistical uncertainty of the unfolded data. The comparison between these two sources can be found in summary plots in Figure 10.10.

[4] Tables for the rest of differential variables are moved to Appendix A.4

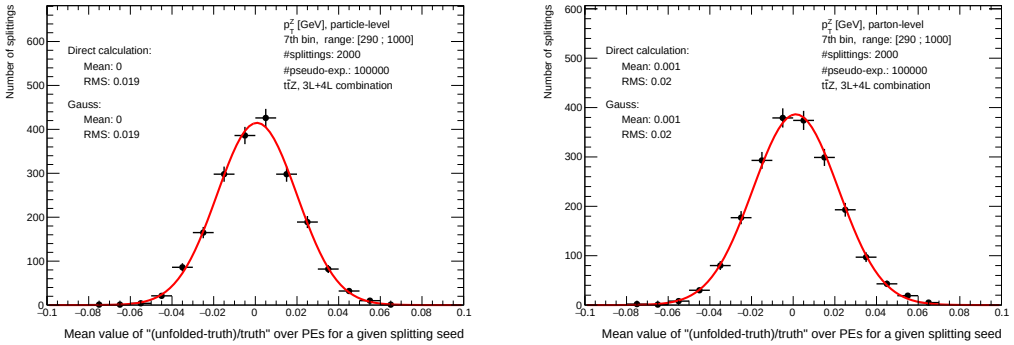


Fig. 10.9: Sample distributions of the average relative difference between unfolded and the truth bin content obtained from 100000 pseudo-experiments for a given splitting seed for the 7th bin of the p_{T}^{Z} variable. Shown are the distributions corresponding to particle level (left) and parton level (right). The red solid line represents Gaussian fit, but the mean that is used further throughout test, is extracted according to the direct calculation defined in Eq. 10.7.

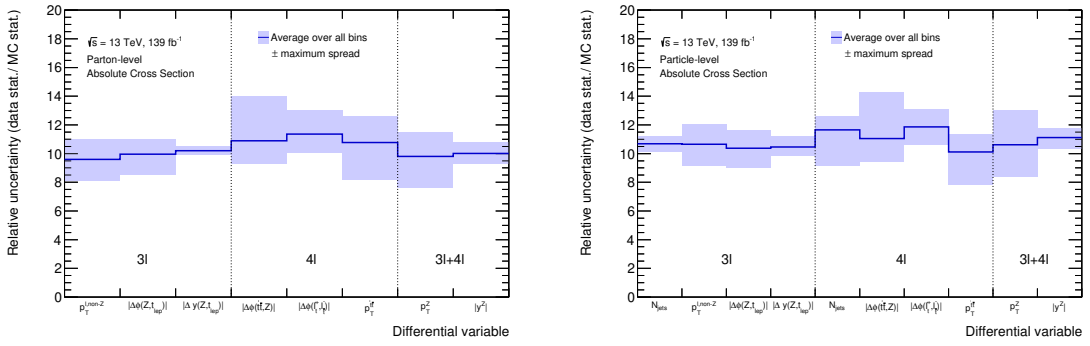


Fig. 10.10: Summary plots showing the ratio of the statistical uncertainties on the unfolded data relative to those due to limited MC statistics. The latter were corrected by a factor $\sqrt{2}$ with respect to those shown in the Table 10.4 (and similar shown in Appendix A.4) to account for the fact that only half the MC events were used for those tests, whereas the full set is used in the analysis. For each variable the average ratio over all bins (solid lines) as well as the full range (shaded region) are shown. The values are shown separately for parton level (left), and particle level (right).

		RMS	
Variable	**Bin**	**Particle level**	**Parton-level**
$\|\Delta\phi(Z, t_{\text{lep}})\|$	1	2.0 %	2.2 %
	2	1.2 %	1.3 %
	3	1.5 %	1.8 %
$p_T^{t\bar{t}}$	1	4.6 %	4.7 %
	2	3.4 %	3.4 %
	3	3.7 %	3.9 %
	4	4.1 %	4.3 %
p_T^Z	1	2.0 %	2.4 %
	2	1.7 %	2.0 %
	3	1.5 %	1.6 %
	4	1.5 %	1.6 %
	5	1.6 %	1.7 %
	6	2.0 %	2.2 %
	7	1.9 %	2.0 %

Table 10.4: Impact of the statistical limitations of the signal MC sample on $|\Delta\phi(Z, t_{\text{lep}})|$ in the trilepton, $p_{\text{T}}^{t\bar{t}}$ in the tetralepton and p_{T}^{Z} in the combined $3\ell + 4\ell$ channel. The quoted values were derived based on half of the simulated events, such that the effective RMS is a factor $\sqrt{2}$ smaller. Values for the rest of differential variables can be found in corresponding tables in Appendix A.4.

10.2.5 Stress tests

The so-called *stress* tests are designed to test the ability of the unfolding procedure to correctly unfold distributions which have significantly different shape than the distributions used for constructing migration matrix and unfolding corrections. This ability is an important feature of differential measurement that needs to provide sensitivity to any physical effect (i.e. the one that would violate the SM) manifesting itself through different shape in measured data distribution, when compared to corresponding MC prediction.

Two different versions of stress tests are performed, both based on shape modifications introduced to reconstruction- and truth-level distributions. The migration matrix and acceptance/efficiency corrections are derived from nominal signal MC sample (without any shape modification). Scale factors, described in the following text, are then applied to the reconstruction-level distributions, which are subsequently unfolded, and compared to correspondingly re-scaled truth-level distributions.

The first set of scale factors S_i, referred to as *data-driven*, is obtained from the difference observed between signal expected in measured data and corresponding MC signal, defined as follows:

$$S_i = \frac{\text{data}^i - \text{MC}^i_{\text{bkg}}}{\text{MC}^i_{\text{signal}}}, \tag{10.11}$$

where MC^i_{bkg} and $\text{MC}^i_{\text{signal}}$ corresponds to the number of background and signal events in bin i, respectively and data_i represents number of observed data events in bin i. The scale factors are derived bin-wise at the truth level and propagated event-by-event to the reconstruction level. In case the event is outside of the fiducial phase space on the truth level, it is not re-weighted.

Second set of scale factors, referred to as *linear*, is obtained from the linear function, which is defined so as to provide $\pm 20\%$, $\pm 40\%$ slope for the extreme events (providing separate scale factor for each event). For example, in case of $+20\%$ slope the event featuring highest value of given variable is re-weighted with factor 1.2, while the event with the lowest value with factor 0.8 (remaining events are assigned scale factor so as to preserve linear function between these two extreme values). In this way, four sets of linear scale factors are derived using the truth-level distributions. Only events in the fiducial volume are re-weighted at reconstruction level.

The plots showing the results of both data-driven and linear stress tests are shown in Figures 10.11-10.13 (and respective figures in Appendix A.5). For data-driven stress tests, the relative discrepancy between unfolded re-weighted distribution and corresponding re-weighted truth-level distribution (red and black lines in bottom ratio plot) is compared to relative statistical uncertainty of unfolded data with respect to re-weighted truth-level distribution (orange band in bottom ratio plot). These compar-

isons suggest, that all observed deviations from the truth-level values are caused by statistical fluctuations in data rather than inconsistency between data and MC. The results of the tests thus show a good ability of the unfolding procedure to recover the re-weighted distribution, even for large values of the slope.

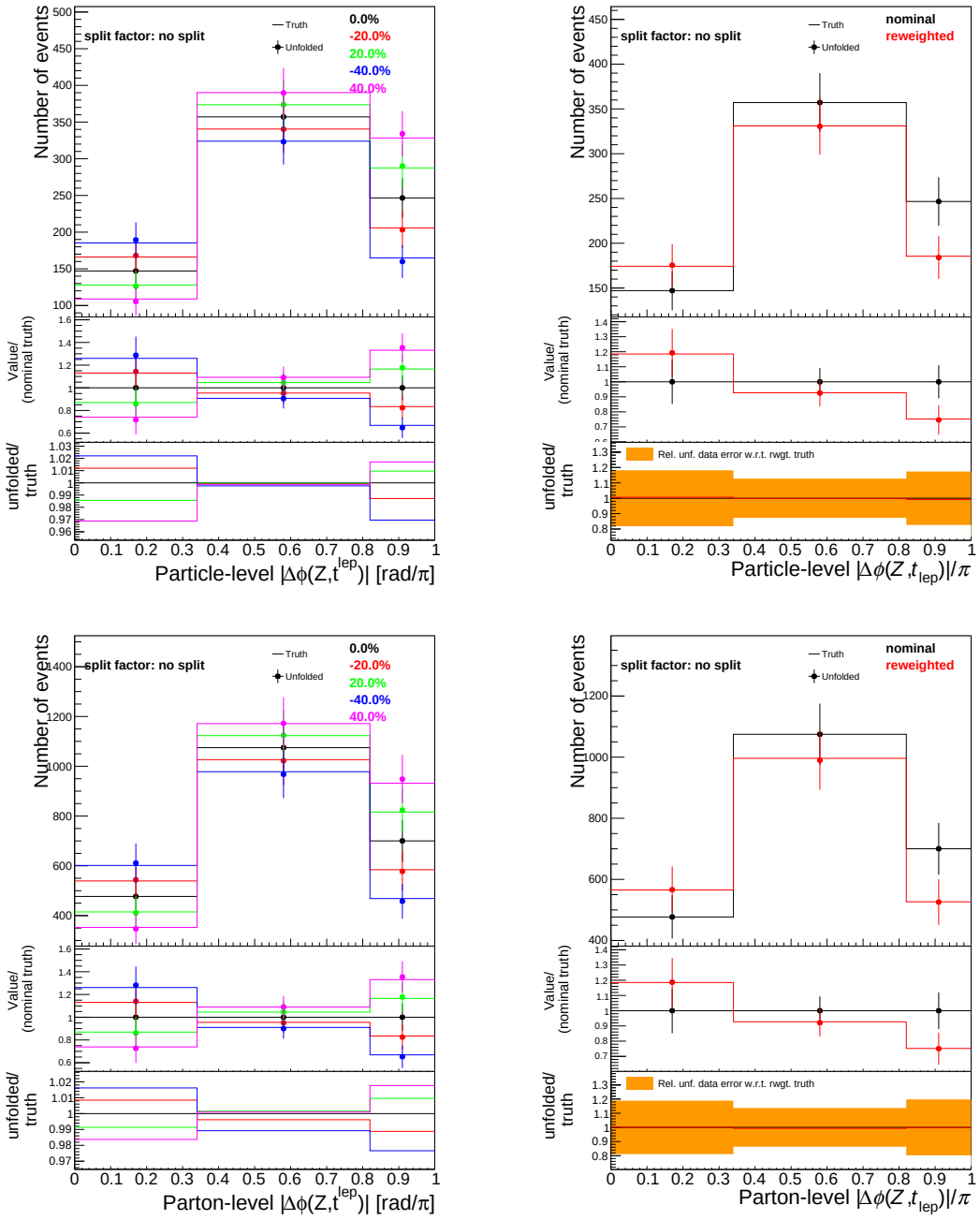


Fig. 10.11: Results of the linear (left) and data-driven (right) stress tests for the $|\Delta\phi(Z, t_{\text{lep}})|$ in the trilepton channel, unfolded to particle level (top) and parton level (bottom).

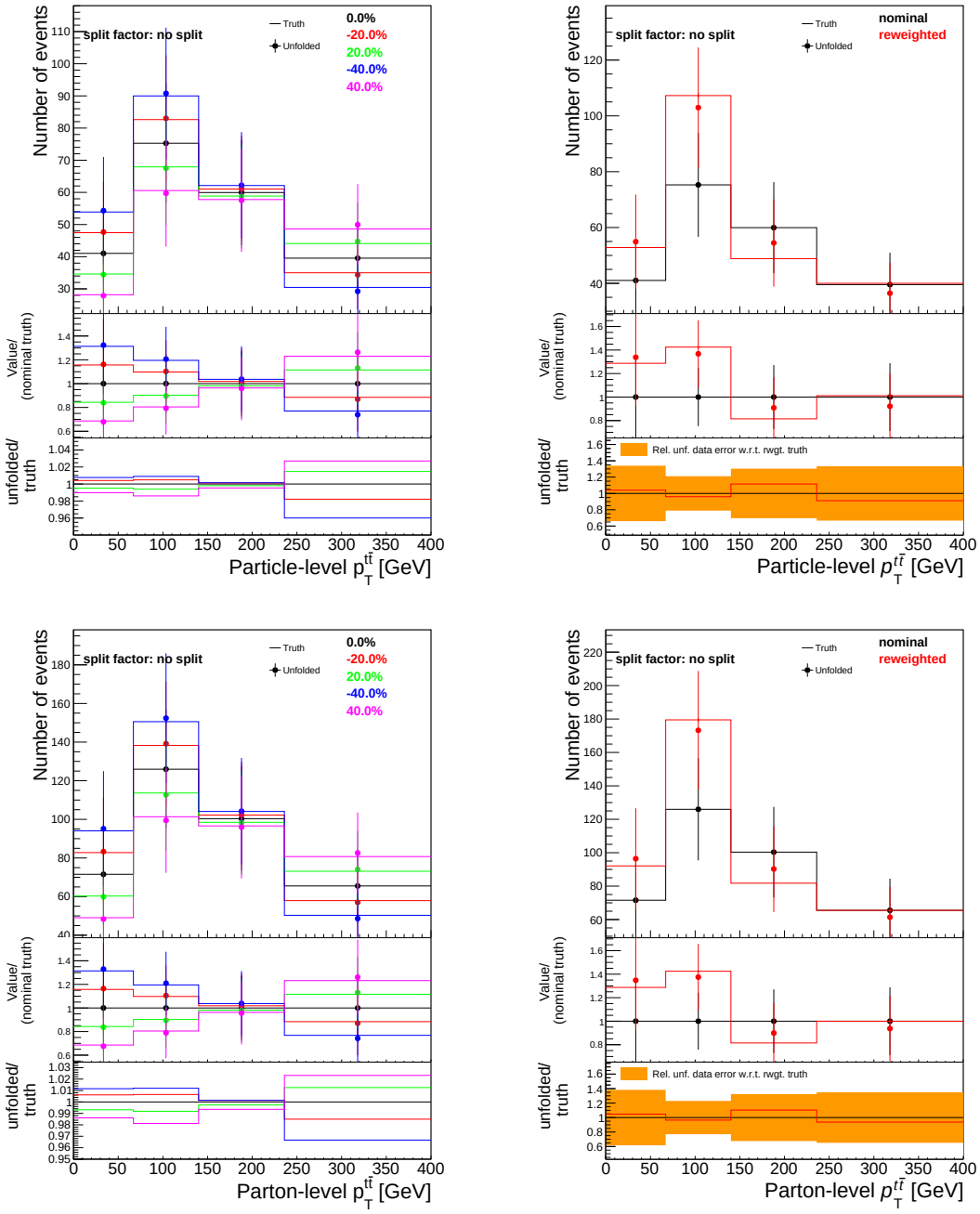


Fig. 10.12: Results of the linear (left) and data-driven (right) stress tests for the $p_{\mathrm{T}}^{t\bar{t}}$ in the tetralepton channel, unfolded to particle level (top) and parton level (bottom).

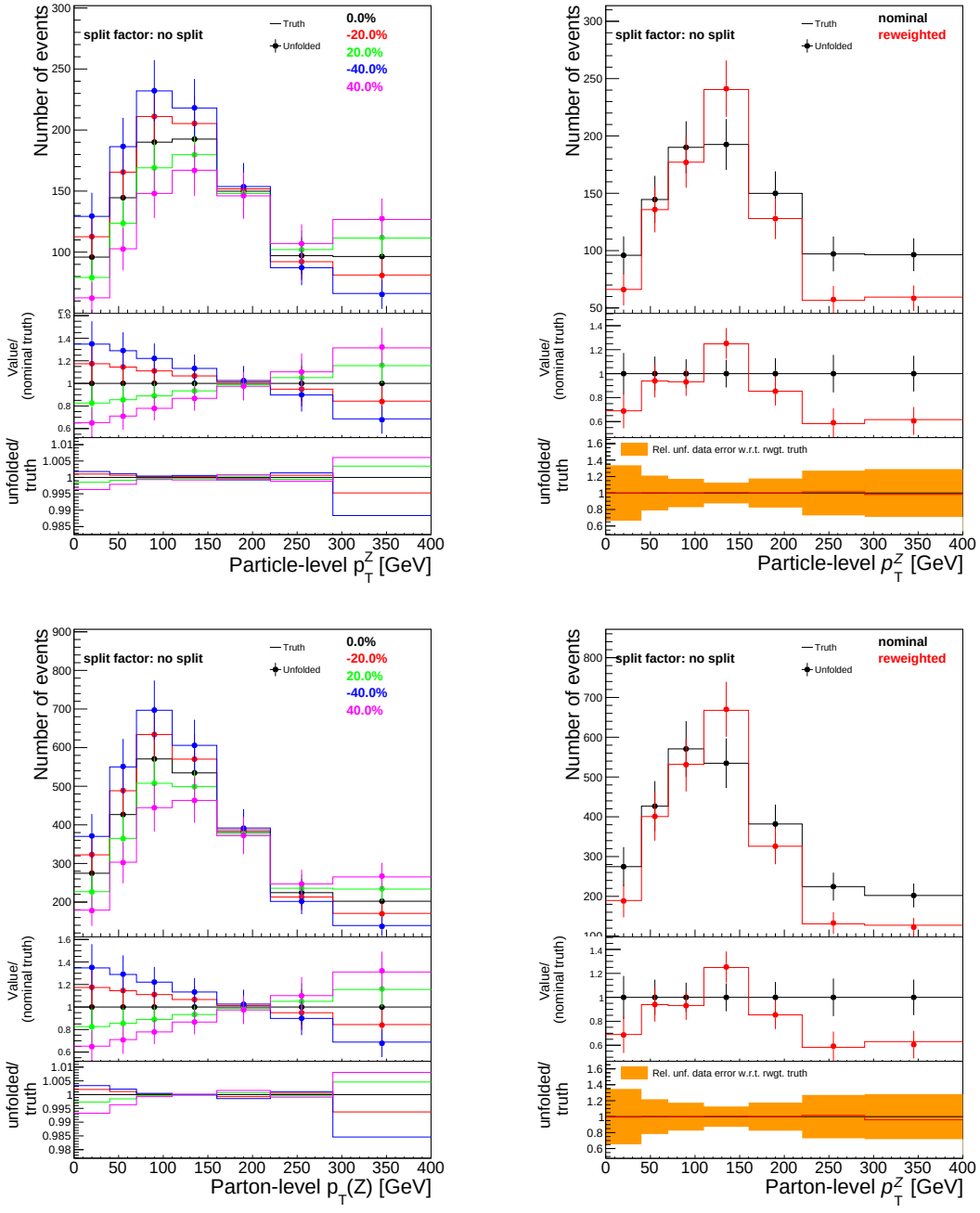


Fig. 10.13: Results of the linear (left) and data-driven (right) stress tests for the p_{T}^{Z} in the combined $3\ell{+}4\ell$ channel, unfolded to particle level (top) and parton level (bottom).

10.3 Results

This section is devoted to the final results of the differential cross section measurements in the trilepton and tetralepton channels. As was already mentioned at the beginning of Chapter 6, the treatment of systematic uncertainties in the context of differential measurement needs more clarification and thus is given in the following section.

10.3.1 Prescription of the systematic uncertainties

The impacts of the various $t\bar{t}Z$ modeling sources are evaluated according to the procedure described in previous Section 10.1, but by performing the unfolding on the modified $t\bar{t}Z$ distribution for each given systematic source separately, and using the migration matrix and unfolding corrections (ϵ_{eff} and f_{acc}) constructed with the nominal signal events. The unfolded modified distribution is then compared to respective truth-level distribution (in case the corresponding truth-level distribution is available, otherwise compared to nominal truth-level distribution) [5]. This observed difference is then divided by the nominal truth-level prediction to get a relative uncertainty for the given systematic source in each particular bin.

For all other sources of systematic uncertainty, such as detector-related ones, which affect both signal and background predictions alike, or those specific to a particular background, the approach is slightly adjusted. The modified signal distribution [6] is first added to that for the modified background. From this the nominal background distribution is subtracted. The resulting distribution is then unfolded, again using the nominal setup (migration matrices and unfolding corrections built from the nominal MC sample) and compared to the nominal truth-level prediction in order to quote the systematic effect for each bin.

In the case of purely one-sided systematic sources (both modeling and detector-related), e.g. parton shower variation, $t\bar{t}Z$ `A14` variations, and some jet energy scale or $E_{\text{T}}^{\text{miss}}$ components, a post-unfolding symmetrization of the quoted uncertainties with respect to the corresponding truth-level distribution [7], is performed.

The individual systematic contributions are grouped into categories based on their origin, and quoted in the uncertainty decomposition plots presented in the next section. It should be noted that the individual systematics effects are further symmetrized

[5] For the systematic uncertainty associated with the $t\bar{t}Z$ `A14` variation of tune parameters, the unfolded up and down variations are compared separately to their respective up and down truth-level predictions. In this sense they are effectively treated as separate systematic sources in order to accommodate the fact that they have different truth-level predictions.

[6] The nominal signal is used in the case of a systematic source affecting only the background prediction.

[7] If the alternative truth-level distribution is not available, the symmetrization is performed using nominal truth-level prediction.

before they are merged into category. If the effect of a particular systematic source is found to be asymmetric in at least one bin (meaning that both up and down contributions have the same sign), then in all bins the largest of the two absolute values (up or down), based on that particular bin, is taken. In such cases the proper sign is always retained: "+" for up and "−" for down variation.

10.3.2 Observed differential cross sections

Differential cross section measurements are performed as a function of variables described in Section 9.4, following the strategy outlined in Section 10.1. The nominal MC signal sample (MadGraph5_aMC@NLO + Pythia 8) is used for determination of migration matrix and unfolding correction, that are subsequently applied in the unfolding performed on the observed ATLAS data. The SM predictions of the differential cross sections, as predicted by various MC generators, are showed in the same plots as measured differential cross sections in Figures 10.14,10.16 and 10.18. Both absolute and normalized cross sections are presented, together with the uncertainty decomposition plots shown in Figures 10.15, 10.17 and 10.19. The plots in the above-mentioned figures present the results for sample variables in each analysis channel, while similar plots for the remaining variables can be found in Appendix A.6. The normalized distributions presented in these figures are obtained post-unfolding by dividing the absolute distributions by measured fiducial cross sections computed by summing up all bin contents of the absolute spectra, thus normalizing the distributions to unit integral.

The MC predictions depicted in the differential cross section plots are specified in Section 5.2.6 and include MadGraph5_aMC@NLO interfaced to either Pythia 8 (red line) or Herwig 7 (magenta line), and Sherpa 2.2.1 inclusive (blue line), as well as multi-leg setup (green line). Moreover, some of the variables feature also dedicated theory predictions calculated at either NLO, NLO+NNLL or nNLO (approximate NLO) precision (including EW corrections). These predictions, shown in a form of gray lines and bands representing the associated errors [8], were calculated according to the method described in Ref. [238], but adjusted in the context of this analysis, using the particular variables and their bin ranges. The dedicated predictions are, however, available only at the parton level, while the two variables involving the leptonic decay products ($p_{\mathrm{T}}^{\ell \mathrm{non}-Z}$ in the trilepton and $|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)|$ in the tetralepton channel) lack these predictions completely.

As can be concluded from the uncertainty decomposition plots, the dominant source of the uncertainty for most of the trilepton variables is limited data statistics when com-

[8] The error bands include uncertainties associated to the scale variations, PDF and MC statistics, which are added linearly to provide conservative estimation. Note that this conservative estimate of the uncertainty is not used further in the compatibility tests discussed next in Section 10.4.

pared to overall systematic uncertainty. Similarly, the statistical uncertainty is even more dominant in the tetralepton channel due to the limited amount of data. In terms of systematic uncertainty, largest contribution in the trilepton channel comes from the b tagging, jet-related sources, as well as MC modeling of WZ+jets, tZq and signal processes. In the tetralepton signal regions the dominant systematic contribution is related to signal modeling, followed by ZZ+jets and tWZ modeling, while jet+$E_{\mathrm{T}}^{\mathrm{miss}}$ also have significant impact on the overall uncertainty. As the combined $3\ell+4\ell$ channel is dominated by trilepton events, the leading systematic sources are similar to those in the trilepton channel, while the statistical component is suppressed due to the largest number of available events. As can be seen by comparing systematic uncertainties between absolute and normalized cross sections, the normalized distributions are less affected by the systematic sources because some of them are effectively canceled out when the normalization is performed. No significant shape dependence of any systematic category can be observed for any differential variable. Furthermore, there is no substantial difference in the impact of particular systematics between parton- and particle-level distributions. Note that the significant difference in the differential cross section scale between absolute particle- and parton-level distributions is caused mainly by applying branching ratio correction only at the parton level (values are quoted in Table 10.1), and by difference in the efficiency and acceptance corrections.

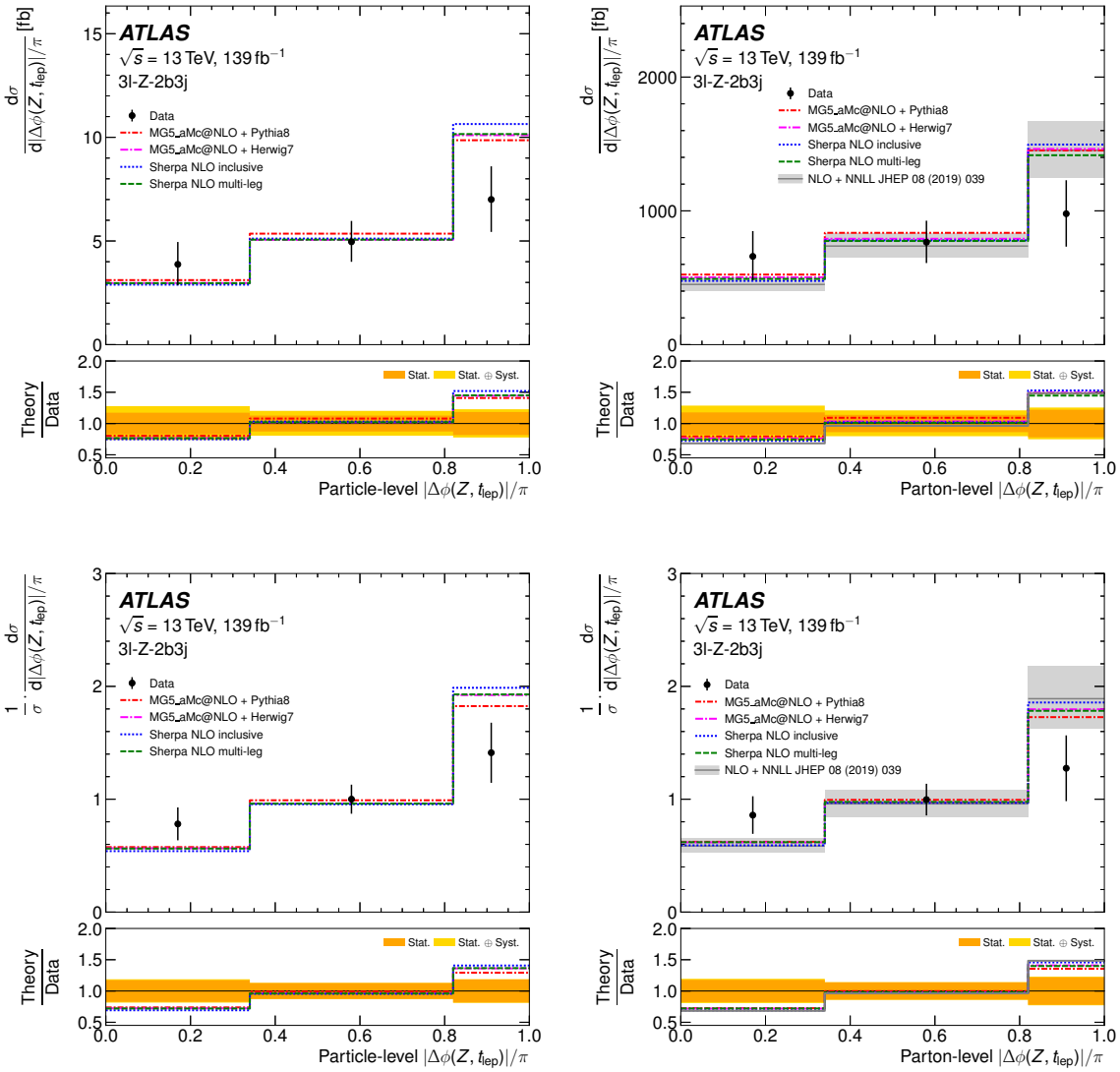


Fig. 10.14: Absolute (top) and normalized (bottom) distributions of the differential cross section as a function of $|\Delta\phi(Z, t_{\text{lep}})|$ in the trilepton channel, unfolded to particle (left) and parton level (right). The bottom ratio plots show statistical and combined (statistical $\oplus$ systematic) uncertainty on the measured differential cross sections.

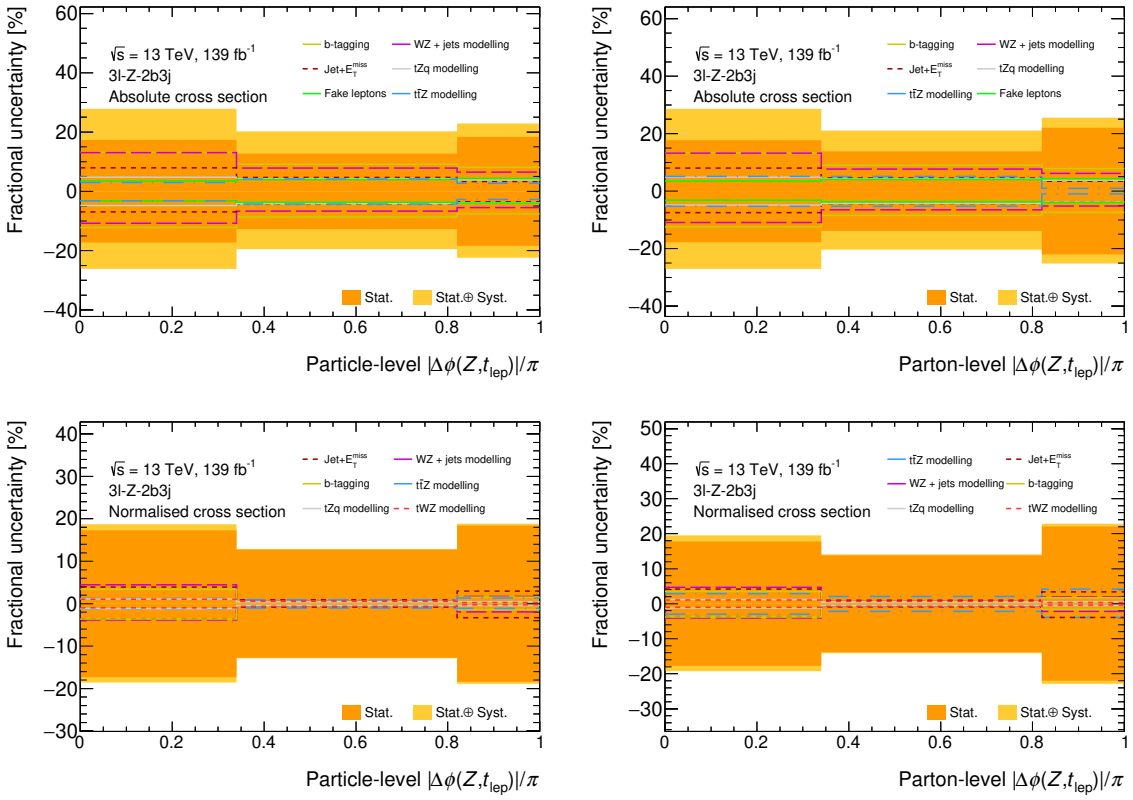


Fig. 10.15: Absolute (top) and normalized (bottom) fractional decomposition of the uncertainties (statistical $\oplus$ systematic) on the differential cross section as a function of $|\Delta\phi(Z, t_{\text{lep}})|$ in the trilepton channel, unfolded to particle (left) and parton level (right).

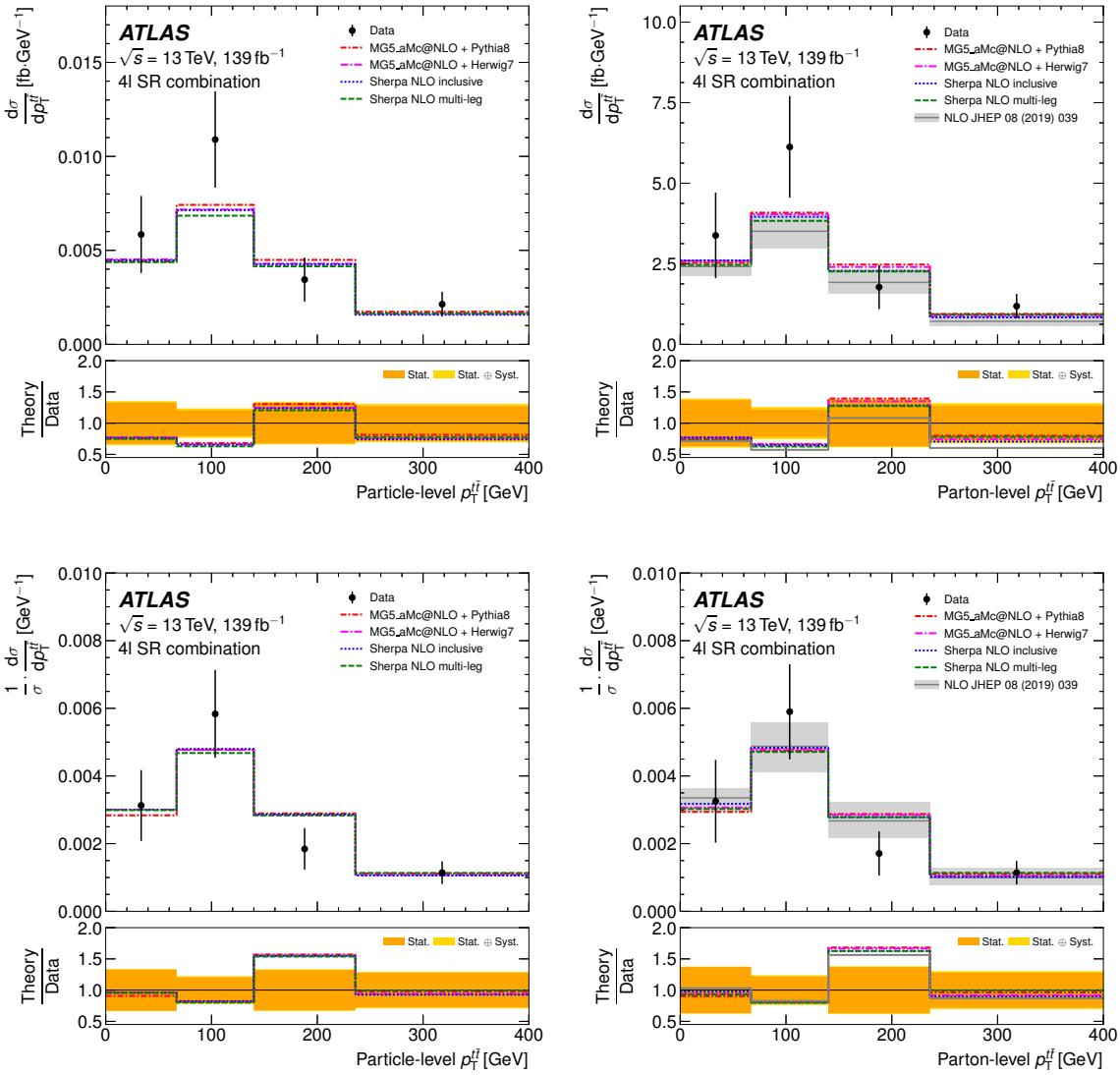


Fig. 10.16: Absolute (top) and normalized (bottom) distributions of the differential cross section as a function of $p_{\mathrm{T}}^{t\bar{t}}$ in the tetralepton channel, unfolded to particle (left) and parton level (right). The bottom ratio plots show statistical and combined (statistical $\oplus$ systematic) uncertainty on the measured differential cross sections.

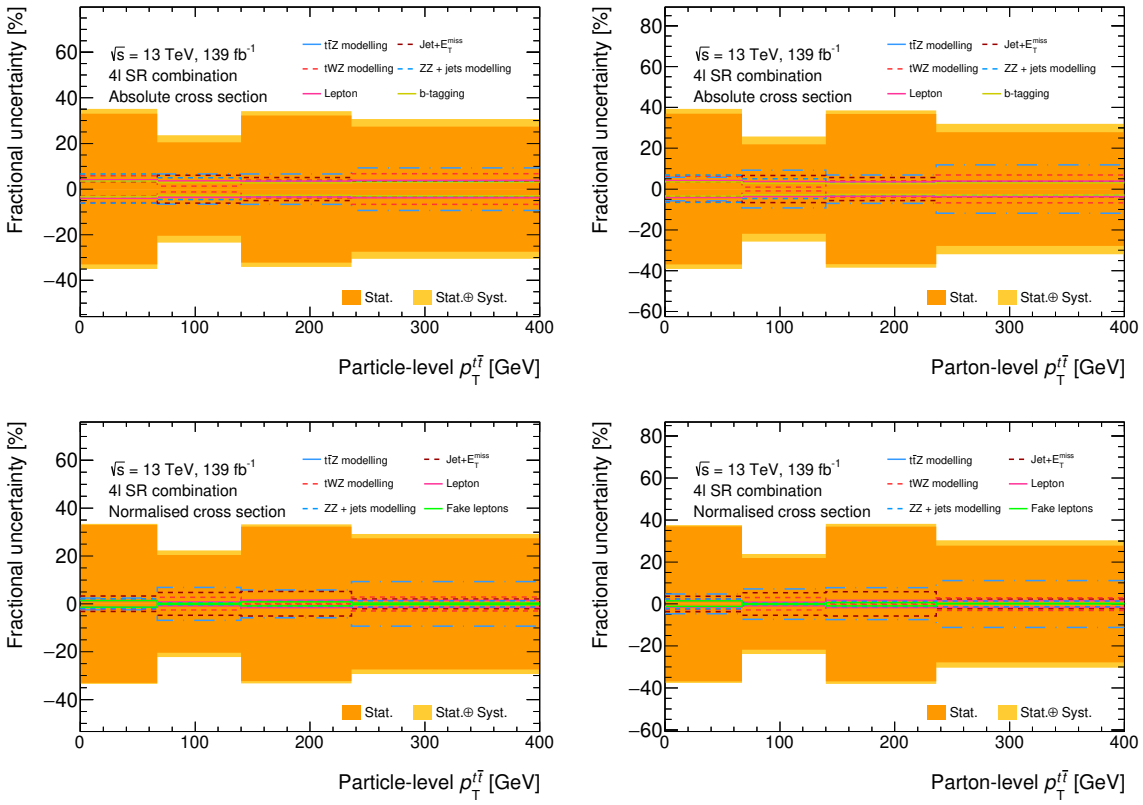


Fig. 10.17: Absolute (top) and normalized (bottom) fractional decomposition of the uncertainties (statistical ⊕ systematic) on the differential cross section as a function of $p_{\mathrm{T}}^{t\bar{t}}$ in the tetralepton channel, unfolded to particle (left) and parton level (right).

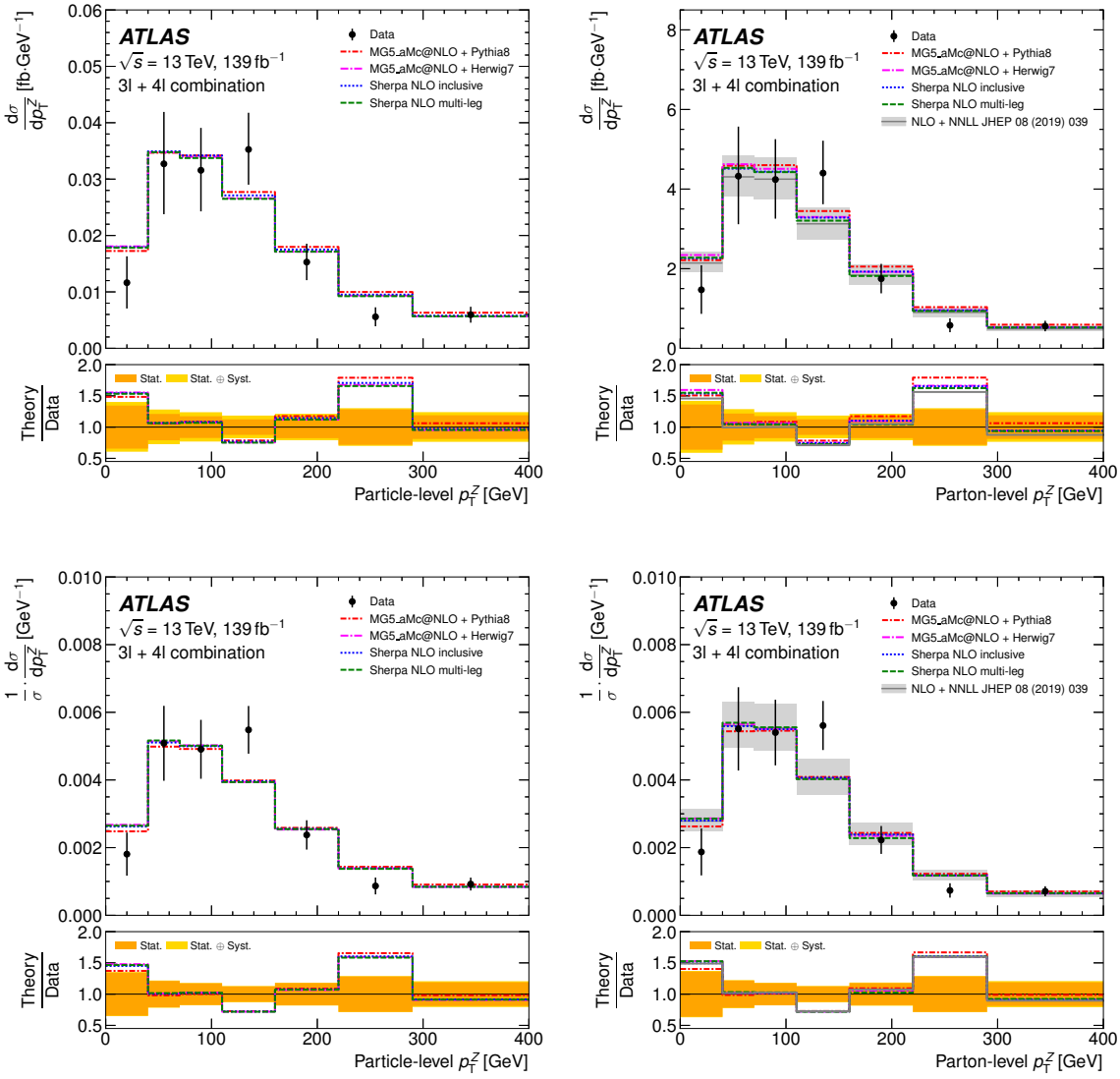


Fig. 10.18: Absolute (top) and normalized (bottom) distributions of the differential cross section as a function of p_{T}^{Z} in the combined $3\ell + 4\ell$ channel, unfolded to particle (left) and parton level (right). The bottom ratio plots show statistical and combined (statistical $\oplus$ systematic) uncertainty on the measured differential cross sections.

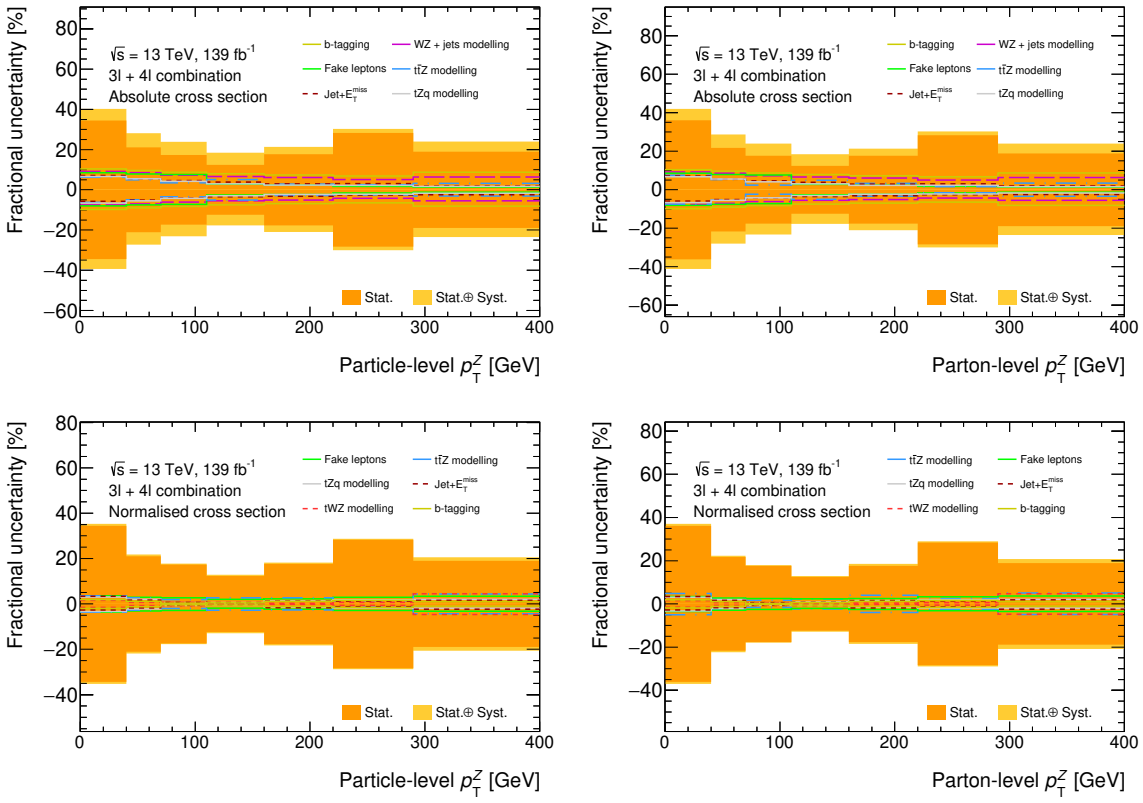


Fig. 10.19: Absolute (top) and normalized (bottom) fractional decomposition of the uncertainties (statistical $\oplus$ systematic) on the differential cross section as a function of p_{T}^{Z} in the combined $3\ell + 4\ell$ channel, unfolded to particle (left) and parton level (right).

10.4 Compatibility between measured and predicted differential cross sections

In order to evaluate the level of compatibility between measured differential cross sections and various MC or theory predictions shown in the differential plots, dedicated calculations of χ^2/ndf as well as corresponding p-value, are performed as described in this section. The calculations are carried out for each differential variable, separately for particle and parton level, and normalized and absolute results.

The standard definition of the χ^2 value is given by the following formula:

$$\chi^2 = \sum_{i=1}^{N_{\text{bins}}} \sum_{j=1}^{N_{\text{bins}}} (n_i - \mu_i)\,(n_j - \mu_j)\,[C^{-1}]_{ij}, \tag{10.12}$$

where n_i denotes content of bin i from the measured differential distribution, while μ_i stands for the corresponding predicted value. Factor $[C^{-1}]_{ij}$ represents ij-th element of the inverse of the covariance matrix for given differential variable. The above-defined χ^2 is reported per number of degrees of freedom (ndf), which is equal to the number of bins (N_{bins}) of particular variable in case of absolute cross section, and to $N_{\text{bins}} - 1$ in the normalized case. The p-values quoted for each variable can be then interpreted as the measure of probability to obtain given value of χ^2/ndf.

The crucial part of χ^2 calculation is determination of the covariance matrix C. The approach employed in this analysis follows the method described in Ref. [239], which ensures that covariance matrix reflects both the statistical as well as all systematic uncertainties. It should be noted that the uncertainties on the theoretical predictions (μ_i from Eq. 10.12) are not incorporated in the covariance matrix and thus omitted in quoted χ^2 and p-values.

The method of estimating the elements of covariance matrix is based on so-called *bootstrap* technique. In the first step, 150000 pseudo-experiments are produced by Poisson-smearing reconstruction-level distribution, in order to simulate effect of statistical fluctuations. Then all the detector-related systematic sources are included by coherently adding Gaussian-distributed shifts to the previously Poisson-smeared bins, where each Gaussian shift corresponds to one uncertainty source governing the amount and direction of particular shift.

The pseudo-experiments constructed in this way are subsequently unfolded using the unfolding correction and migration matrix evaluated from the nominal signal MC sample. In order to ensure correct unfolding result, the bins that acquire negative content due to the sequential shifts, are set to zero prior to the unfolding. After the unfolding, the modeling systematic sources (both for signal and background) are added to the unfolded pseudo-experiments again in a form of Gaussian-distributed shifts [9].

[9] Note the difference between detector-related and modeling systematics. While the detector-related

These are applied as relative variations obtained from the difference between truth-level and unfolded distribution corresponding to the given alternative model.

The changes observed in the unfolded pseudo-experiments, after adding shifts due to modeling systematics, are used to determine the final covariance matrix. Since the covariance matrix is constructed using the MC samples, in order to properly account for described effects in the measured data, the elements of covariance matrix are rescaled according to following prescription:

$$C_{ij} \rightarrow \left(\frac{n_i \cdot n_j}{\mu_i \cdot \mu_j}\right) \cdot C_{ij}, \tag{10.13}$$

where the same notation as in Eq. 10.12, is used.

The obtained χ^2/ndf and p-values for all differential variables unfolded to particle level can be found in Tables 10.5 and 10.6 for absolute and normalized case, respectively. Analogous values for the parton-level measurements are quoted in Tables 10.7 and 10.8. The obtained values indicate overall good compatibility between measured and predicted differential cross sections for most of the variables. The only exceptions, with p-values lower than 0.15, are two trilepton variables ($p_{\mathrm{T}}^{\ell \mathrm{non}-Z}$ and $|\Delta\phi(Z, t_{\mathrm{lep}})|$) and p_{T}^{Z} in the combined $3\ell + 4\ell$ channel. The poorer agreement observed for the p_{T}^{Z} is driven mostly by the 6-th bin (from 220 GeV to 290 GeV), while the effect is caused predominantly by the observed disagreement between unfolded and predicted values and is further enhanced by non-diagonal elements of covariance matrix, as well as the rescaling introduced in Eq. 10.13. The slightly higher compatibility of this variable is observed for the dedicated NLO+NNLL theory prediction (p=0.17). The poor agreement is achieved also between unfolded $|\Delta\phi(t\bar{t}, Z)|$ in the tetralepton channel and corresponding nominal MC prediction (MadGraph5_aMC@NLO + Pythia 8), but the compatibility is increased when the comparison is made with the alternative Sherpa 2.2.1 predictions. Nevertheless, as the statistical uncertainties on the measured values of the differential cross sections for these variables are larger than the observed discrepancies between the predictions in some bins, no definitive statement about the overall compatibility can be made. As can be seen from the values reported in summary tables, the level of agreement (or disagreement for some variables) is in general similar for both the absolute and normalized distributions, and as well for both particle and parton levels.

systematic sources are smeared and added coherently before the unfolding, the modeling systematic sources are smeared and added after the unfolding.

Particle level		MG5_aMC@NLO		MG5_aMC@NLO		Sherpa 2.2.1		Sherpa 2.2.1	
Absolute cross section		+ Pythia 8		+ Herwig 7		NLO multi-leg		NLO inclusive	
	Variable	χ^2/ndf	p-value	χ^2/ndf	p-value	χ^2/ndf	p-value	χ^2/ndf	p-value
3ℓ	N_{jets}	0.8/3	0.85	0.6/3	0.90	0.3/3	0.95	0.5/3	0.92
	$p_{\text{T}}^{\ell,\text{non-}Z}$	7.5/4	0.11	7.2/4	0.13	7.7/4	0.11	7.7/4	0.10
	$\lvert\Delta\phi(Z,t_{\text{lep}})\rvert/\pi$	5.4/3	0.14	6.5/3	0.09	6.7/3	0.08	8.6/3	0.04
	$\lvert\Delta y(Z,t_{\text{lep}})\rvert$	0.9/3	0.83	0.7/3	0.87	0.5/3	0.93	0.9/3	0.81
4ℓ	N_{jets}	1.4/4	0.84	1.7/4	0.79	2.8/4	0.59	2.8/4	0.59
	$\lvert\Delta\phi(\ell_t^+,\ell_{\bar{t}}^-)\rvert/\pi$	2.0/4	0.73	2.3/4	0.69	2.7/4	0.62	2.5/4	0.65
	$\lvert\Delta\phi(t\bar{t},Z)\rvert/\pi$	5.2/3	0.16	4.9/3	0.18	4.1/3	0.25	3.7/3	0.30
	$p_{\text{T}}^{t\bar{t}}$	3.5/4	0.47	3.6/4	0.46	3.8/4	0.44	3.7/4	0.45
$3\ell/4\ell$	p_{T}^{Z}	12.8/7	0.08	12.0/7	0.10	11.6/7	0.11	12.1/7	0.10
	$\lvert y^{Z}\rvert$	2.8/8	0.94	2.9/8	0.94	3.5/8	0.90	2.9/8	0.94

Table 10.5: Summary of χ^2/ndf and p-values representing the compatibility between observed absolute $t\bar{t}Z$ differential cross sections measured at the particle level, and various MC predictions. The quoted values are based on the assumption that statistical fluctuations are distributed according to Poisson, and systematic uncertainties according to Gaussian distribution.

Particle level		MG5_aMC@NLO		MG5_aMC@NLO		Sherpa 2.2.1		Sherpa 2.2.1	
Normalized cross section		+ Pythia 8		+ Herwig 7		NLO multi-leg		NLO inclusive	
	Variable	χ^2/ndf	p-value	χ^2/ndf	p-value	χ^2/ndf	p-value	χ^2/ndf	p-value
3ℓ	N_{jets}	0.3/2	0.88	0.2/2	0.92	0.1/2	0.94	0.2/2	0.89
	$p_{\text{T}}^{\ell,\text{non-}Z}$	6.4/3	0.09	6.4/3	0.09	6.8/3	0.08	6.7/3	0.08
	$\lvert\Delta\phi(Z,t_{\text{lep}})\rvert/\pi$	4.0/2	0.14	5.4/2	0.07	5.5/2	0.06	6.7/2	0.03
	$\lvert\Delta y(Z,t_{\text{lep}})\rvert$	0.4/2	0.81	0.5/2	0.79	0.2/2	0.89	0.5/2	0.77
4ℓ	N_{jets}	0.4/3	0.94	0.3/3	0.96	1.3/3	0.73	1.6/3	0.66
	$\lvert\Delta\phi(\ell_t^+,\ell_{\bar{t}}^-)\rvert/\pi$	1.3/3	0.74	1.1/3	0.78	1.1/3	0.77	1.3/3	0.74
	$\lvert\Delta\phi(t\bar{t},Z)\rvert/\pi$	5.3/2	0.07	4.8/2	0.09	3.3/2	0.19	3.0/2	0.22
	$p_{\text{T}}^{t\bar{t}}$	3.9/3	0.28	3.7/3	0.30	3.6/3	0.30	3.7/3	0.30
$3\ell/4\ell$	p_{T}^{Z}	11.0/6	0.09	10.8/6	0.09	10.6/6	0.10	10.7/6	0.10
	$\lvert y^{Z}\rvert$	2.4/7	0.94	2.6/7	0.92	3.1/7	0.87	2.5/7	0.92

Table 10.6: Summary of χ^2/ndf and p-values representing the compatibility between observed normalized $t\bar{t}Z$ differential cross sections measured at the particle level, and various MC predictions. The quoted values are based on the assumption that statistical fluctuations are distributed according to Poisson, and systematic uncertainties according to Gaussian distribution.

Parton level **Absolute cross section**		MG5_aMC@NLO + Pythia 8		MG5_aMC@NLO + Herwig 7		Sherpa 2.2.1 NLO multi-leg		Sherpa 2.2.1 NLO inclusive		Additional Theory	
	Variable	χ^2/ndf	p-value	χ^2/ndf	p-value	χ^2/ndf	p-value	χ^2/ndf	p-value	χ^2/ndf	p-value
3ℓ	$p_{\mathrm{T}}^{\ell,\text{non-}Z}$	7.6/4	0.11	8.8/4	0.07	8.3/4	0.08	8.6/4	0.07	/	/
	$\lvert\Delta\phi(Z,t_{\mathrm{lep}})\rvert/\pi$	5.5/3	0.14	5.8/3	0.12	5.2/3	0.16	6.9/3	0.07	6.6/3	0.09
	$\lvert\Delta y(Z,t_{\mathrm{lep}})\rvert$	0.9/3	0.82	0.7/3	0.88	0.2/3	0.98	0.5/3	0.92	0.3/3	0.96
4ℓ	$\lvert\Delta\phi(\ell_t^+,\ell_{\bar{t}}^-)\rvert/\pi$	2.1/4	0.72	2.3/4	0.69	2.7/4	0.62	2.6/4	0.63	/	/
	$\lvert\Delta\phi(t\bar{t},Z)\rvert/\pi$	5.2/3	0.16	4.7/3	0.19	3.5/3	0.32	3.4/3	0.33	4.9/3	0.18
	$p_{\mathrm{T}}^{t\bar{t}}$	3.5/4	0.47	3.6/4	0.47	3.5/4	0.48	3.5/4	0.47	4.6/4	0.33
$3\ell/4\ell$	p_{T}^{Z}	12.8/7	0.08	11.7/7	0.11	11.2/7	0.13	11.3/7	0.13	10.4/7	0.17
	$\lvert y^{Z}\rvert$	2.8/8	0.95	2.9/8	0.94	4.0/8	0.85	2.7/8	0.95	2.9/8	0.94

Table 10.7: Summary of χ^2/ndf and p-values representing the compatibility between observed absolute $t\bar{t}Z$ differential cross sections measured at the parton level, and various MC predictions. The quoted values are based on the assumption that statistical fluctuations are distributed according to Poisson, and systematic uncertainties according to Gaussian distribution. Last column corresponds to the dedicated theory predictions based on the calculations presented in Ref. [238], with the variable precisions: NLO ($|\Delta\phi(t\bar{t},Z)|/\pi$, $p_{\mathrm{T}}^{t\bar{t}}$), NLO+NNLL ($|\Delta\phi(Z,t_{\mathrm{lep}})|/\pi$, $|\Delta y(Z,t_{\mathrm{lep}})|$, p_{T}^{Z}), and nNLO ($|y^{Z}|$).

Parton level **Normalized cross section**		MG5_aMC@NLO + Pythia 8		MG5_aMC@NLO + Herwig 7		Sherpa 2.2.1 NLO multi-leg		Sherpa 2.2.1 NLO inclusive		Additional Theory	
	Variable	χ^2/ndf	p-value	χ^2/ndf	p-value	χ^2/ndf	p-value	χ^2/ndf	p-value	χ^2/ndf	p-value
3ℓ	$p_{\mathrm{T}}^{\ell,\text{non-}Z}$	6.6/3	0.09	7.8/3	0.05	7.6/3	0.06	7.7/3	0.05	/	/
	$\lvert\Delta\phi(Z,t_{\mathrm{lep}})\rvert/\pi$	3.9/2	0.14	4.7/2	0.09	4.6/2	0.10	5.9/2	0.05	6.4/2	0.04
	$\lvert\Delta y(Z,t_{\mathrm{lep}})\rvert$	0.4/2	0.80	0.4/2	0.80	0.1/2	0.93	0.4/2	0.83	0.3/2	0.86
4ℓ	$\lvert\Delta\phi(\ell_t^+,\ell_{\bar{t}}^-)\rvert/\pi$	1.2/3	0.75	1.3/3	0.74	1.1/3	0.77	1.2/3	0.75	/	/
	$\lvert\Delta\phi(t\bar{t},Z)\rvert/\pi$	5.4/2	0.07	4.7/2	0.10	2.3/2	0.31	2.6/2	0.28	2.5/2	0.29
	$p_{\mathrm{T}}^{t\bar{t}}$	4.0/3	0.26	3.9/3	0.28	3.5/3	0.32	3.5/3	0.32	3.0/3	0.39
$3\ell/4\ell$	p_{T}^{Z}	11.0/6	0.09	10.8/6	0.10	10.7/6	0.10	10.6/6	0.10	10.5/6	0.11
	$\lvert y^{Z}\rvert$	2.3/7	0.94	2.5/7	0.93	3.5/7	0.84	2.4/7	0.94	2.6/7	0.92

Table 10.8: Summary of χ^2/ndf and p-values representing the compatibility between observed normalized $t\bar{t}Z$ differential cross sections measured at the parton level, and various MC predictions. The quoted values are based on the assumption that statistical fluctuations are distributed according to Poisson, and systematic uncertainties according to Gaussian distribution. Last column corresponds to the dedicated theory predictions based on the calculations presented in Ref. [238], with the variable precisions: NLO ($|\Delta\phi(t\bar{t},Z)|/\pi$, $p_{\mathrm{T}}^{t\bar{t}}$), NLO+NNLL ($|\Delta\phi(Z,t_{\mathrm{lep}})|/\pi$, $|\Delta y(Z,t_{\mathrm{lep}})|$, p_{T}^{Z}), and nNLO ($|y^{Z}|$).

Chapter 11

Inclusive cross section measurement

This chapter is devoted to the measurement of the $t\bar{t}Z$ inclusive cross section in the dilepton channel. Since the dilepton analysis is part of the ongoing $t\bar{t}Z$ measurement, only the preliminary strategy and expected results are presented here.

11.1 Analysis strategy

The inclusive cross section measurement is based on the profile likelihood fitting technique described in Section 7.4 and aims to extract the signal strength $\mu_{t\bar{t}Z}^{2\ell}$ corresponding to the ratio of the observed cross section and its SM prediction. In order to increase sensitivity of the fit, the fitted distribution is chosen to be the output of the neural network designed for the separation of the signal process from irreducible background contamination. The details about the NN architecture and its implementation are given in following Section 11.3.

The extraction of the $t\bar{t}Z$ signal strength is performed via simultaneous fit in the three dilepton signal regions defined in Section 8.1 and include all relevant sources of systematic uncertainties, as described in Chapter 6. In the fit, the leading Z+jets background is decomposed into three sub-samples according to the quark types from which the associated jets originate. If there is at least one jet matched to a b hadron at the particle level, the event is assigned to the $Z + b$ category. Remaining events are categorized as either $Z + c$ (if at least one jet is matched to a c hadron) or $Z + l$ (otherwise, where l denotes jet originating from light-flavor quark). The normalizations of the $Z + b$ and $Z + c$ processes [1] are allowed to free-float in the fit, together with the signal strength parameter. All other background processes are fixed at their expected values and allowed to vary only within their systematic uncertainties. As mentioned in Section 6.2.2, the normalization of the $Z + l$ process is assigned 10% modeling

[1] Note that only normalizations of these background processes are allowed to free-float, while their shapes are fixed according to the corresponding MC samples.

uncertainty. The contribution of the second leading background process, $t\bar{t}$, is estimated using the data-driven technique described in the following section, instead of using its MC estimate.

In this phase of the ongoing analysis, the cross section measurement is not yet performed using real measured data, but only so-called *Asimov* dataset is employed in the fit, what corresponds to using simulated MC events instead of real measured data. The purpose of this fit is to validate the suggested analysis strategy and eliminate potential sources of pulls and constraints of the nuisance parameters used in the fit. Moreover, the Asimov fit provides the estimate of the total expected uncertainty on the POIs, when the fit is performed on the measured ATLAS data.

11.2 Data-driven estimate of the $t\bar{t}$ background

The dileptonic $t\bar{t}$ process with additional jets is particularly important in the dilepton analysis. Since the modeling of these additional jets can suffer from large systematic uncertainties, a data-driven approach is preferred. A $t\bar{t}$-enriched regions (referred to as $t\bar{t}$ validation regions - VR) with low contamination of other backgrounds, are obtained by requiring the lepton pair to be of opposite sign and different flavor ($e^{\pm}\mu^{\mp}$). Additional selection criteria are applied to replicate those of the signal regions defined in Section 8.1, and therefore limit the extrapolation only to the change in lepton flavor ($e^{\pm}\mu^{\mp}$ in the VRs instead of $e^{\pm}e^{\mp}/\mu^{\pm}\mu^{\mp}$ in the SRs). In this way three $t\bar{t}$ VRs are defined and denoted as: VR-$e\mu$-Z-6j1b, VR-$e\mu$-Z-6j2b and VR-$e\mu$-Z-5j2b.

The data-driven estimate of the $t\bar{t}$ background that could be applied in the simultaneous fit in the SRs, is constructed according to following procedure. In the first step, non-$t\bar{t}$ background passing the $e\mu$ selection of the VRs is subtracted from the $e\mu$ data, effectively yielding the $t\bar{t}$ contribution in the selected $e\mu$ data. In order to use these $t\bar{t}$ events in the dilepton SRs, the different acceptances and efficiencies corresponding to different selections need to be taken into account. For this reason, the data-driven $t\bar{t}$ estimate is scaled with the following correction factor:

$$C_{t\bar{t}} = \frac{N_{t\bar{t}}^{\ell\ell}}{N_{t\bar{t}}^{e\mu}}, \tag{11.1}$$

where $N_{t\bar{t}}^{\ell\ell}$ and $N_{t\bar{t}}^{e\mu}$ are the number of expected $t\bar{t}$ events (from MC predictions) after the $\ell\ell$ selection in the SRs and the $e\mu$ selection in the VRs, respectively. A total uncertainty on this correction factor is derived for each VR and reported in Table 11.1. Total uncertainty includes statistical component of the MC, and modeling uncertainty obtained by comparing the value of the ratio for the nominal $t\bar{t}$ sample (PowhegBox2+Pythia 8) and the one obtained from the alternative MC samples,

namely sample with varied h_{damp} parameter[2], different PS generator (HERWIG 7) and different ME generator (MADGRAPH5_aMC@NLO). The overall modeling uncertainty is then obtained as quadratic sum of differences between correction factors obtained from the nominal and alternative MC samples. Since the correction factors are consistent within the quoted uncertainties, the overall value of the correction factor and its associated uncertainty is applied in all dilepton SRs.

	VR-$e\mu$-Z-6j1b	**VR-$e\mu$-Z-5j2b**	**VR-$e\mu$-Z-6j2b**	**Overall**
$C_{t\bar{t}}$	0.990 ± 0.023	0.980 ± 0.011	0.979 ± 0.008	0.982 ± 0.008

Table 11.1: Values of the correction factors extracted from $t\bar{t}$ VRs. Overall value refers to common multiplicative factor applied for data-driven $t\bar{t}$ background. The uncertainty includes both component related to the MC statistics as well as systematic uncertainty corresponding to the modeling of $t\bar{t}$ process.

Table 11.2 shows the yields in the $t\bar{t}$ VRs. Control plots for basic kinematic and jet-multiplicity variables for these VRs can be found in Figure 11.1. Since the modeling of these observables in the MC is in substantial disagreement with the observed data, the decision to use the above-mentioned data-driven approach is well motivated.

	VR-$e\mu$-Z-6j1b	VR-$e\mu$-Z-5j2b	VR-$e\mu$-Z-6j2b
$t\bar{t}Z$	6.80 ± 0.86	17.32 ± 1.27	14.18 ± 1.62
MC $t\bar{t}$	1112.49 ± 296.07	3781.00 ± 735.16	1729.07 ± 389.36
$Z+b$	0.34 ± 0.36	0.19 ± 0.27	0.21 ± 0.14
$Z+c$	0.14 ± 0.06	0.09 ± 0.39	0.01 ± 0.03
$Z+l$	0.01 ± 0.03	0.00 ± 0.00	0.00 ± 0.00
tWZ	0.21 ± 0.03	0.23 ± 0.03	0.21 ± 0.03
Diboson (VV)	6.45 ± 3.25	1.78 ± 0.91	1.29 ± 0.66
Fake leptons	26.13 ± 13.19	31.36 ± 15.85	20.55 ± 10.41
Other	15.16 ± 7.61	35.86 ± 17.97	33.40 ± 16.74
Total SM	1167.72 ± 297.00	3867.82 ± 736.10	1798.92 ± 390.85
Data	1260	4104	2041

Table 11.2: The numbers of observed and expected signal and background events in the $t\bar{t}$ validation regions, obtained for an integrated luminosity of 139 fb^{-1}. The indicated errors include the MC statistical as well as the systematic uncertainties on the different SM components (see Section 6). Category labeled as "Other" includes all other SM processes that feature at least two prompt leptons and are not listed in this table (the complete list can be found in Section 5.2.7).

[2]The h_{damp} factor is parameter of the POWHEG generator which controls the matching between ME and PS, and also effectively regulates radiations with high transverse momenta.

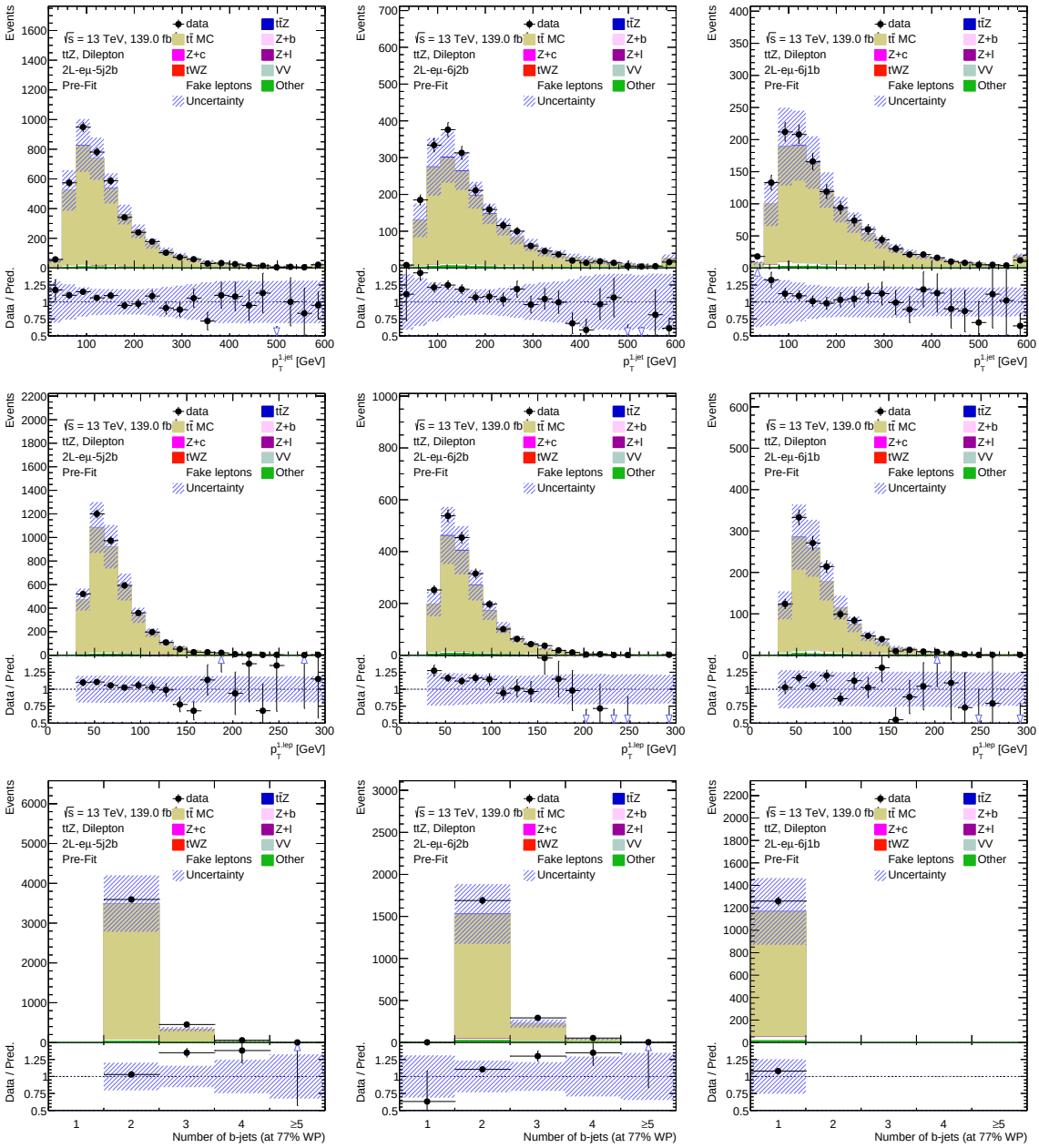


Fig. 11.1: Distributions of the transverse momenta of the leading jet (top row), leading lepton (second row) and b-jet multiplicity (bottom row) in the VR-$e\mu$-Z-5j2b (left), VR-$e\mu$-Z-6j2b (middle) and VR-$e\mu$-Z-6j1b (right).

11.3 Classification neural network

In order to classify if the events belong to either signal or background processes, separate binary classification deep NNs are trained for each SR. Training is based on the Keras [224] framework within Tensorflow [225] package. The optimization of the NN architecture and its hyperparameters is performed separately for each SR after identification of optimal set of input variables.

In the first step, input variables are selected based on the permutation importance

ranking, as described in Section 7.3.1. The complete set of discriminating variables consists of 58 variables, which are related mostly to the kinematics of the top quarks and Z boson, but employ also outputs of various reconstruction methods, i.e. multi-hypothesis hadronic t/W reconstruction (described in Section 9.1.1) or SPANet (described in Section 9.1.2). Results of permutation importance optimizations in a form of variable rankings can be found in Figure 11.2. These ranking plots show already optimal set of variables for each SR, while the definitions of the selected variables are summarized in Table 11.3. Control plots for each variable in a given SR, as well as corresponding separation plots comparing shapes of signal and background, can be found in dedicated Appendix B.1. All control plots indicate reasonably good modeling [3] of the selected variables, approving their usage for NN training.

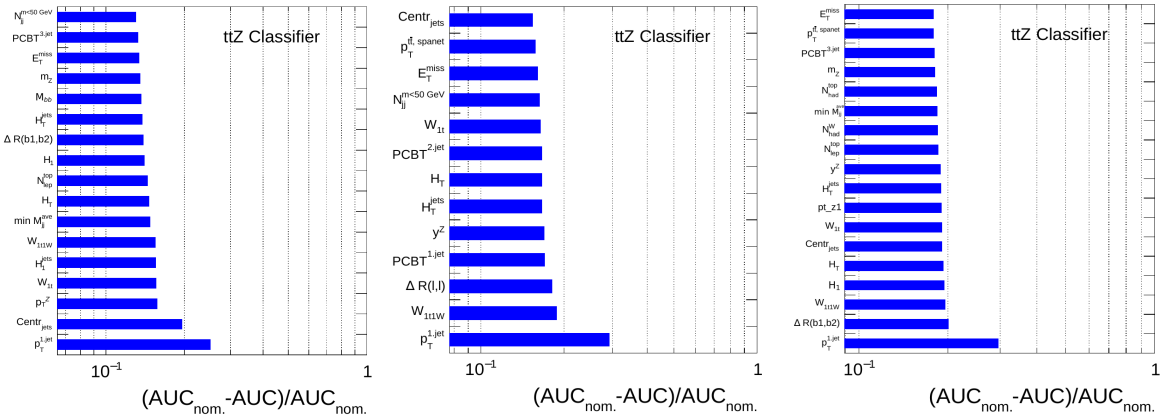


Fig. 11.2: Ranking of the selected variables determined by permutation importance used for the NN training in 5j2b (left), 6j1b (middle) and 6j2b (right) signal region. AUC$_{nom.}$ denotes area under ROC curve of nominal (unshuffled) set of variables, while AUC represents training with set of variables where particular variable was shuffled according to the procedure described in Section 7.3.1.

Ideal set of variables is then used for finding optimal values of hyperparameters that define architecture of the used NNs. The optimization procedure follows the random search described in Section 7.3.1, in which 200 random sets of hyperparameters are used for NN training and best performing NN is used to define optimal hyperparameters. The obtained values of hyperparameters for each SR are summarized in Table 11.4. All NNs are trained with learning rate equal to 0.0001 and batch size of 5000. To suppress overtraining, each NN uses *early stopping* and additional *dropout* layers, with rate of 0.3, after the first layer. The requirements for the early stopping (criterion determining when the training terminates) are based on the value of the loss function. The training is stopped once the loss is not decreased by more than 0.0001 over 30

[3] Slight miss-modeling is found in 2ℓ-Z-$6j1b$ region, but the disagreement between observed data and MC is within uncertainties for most of the variables.

consecutive epochs.

Variable	Definition
H_T	sum of p_T of all objects (jets and leptons) in the event
H_T^{jets}	sum of p_T of all jets in the event
$p_T^{X.jet}$	p_T of the X-th jet
$p_T^{X.lep}$	p_T of the X-th lepton
W_{1t}	weight for $1t$ category from multi-hypothesis hadronic t/W reconstruction (see Section 9.1.1)
$Centr_{jets}$	scalar sum of p_T divided by sum of E for all jets
$\Delta R(b_1, b_2)$	cone between two jets with highest b-tagging weight (output of the DL1r tagger)
W_{1t1W}	weight for $1t1W$ category from multi-hypothesis hadronic t/W reconstruction (see Section 9.1.1)
H_1	first Fox-Wolfram momentum [4]
H_1^{jets}	first Fox-Wolfram moment built only from jets
$N_{jj}^{m<50GeV}$	number of two-jet combinations with mass lower than 50 GeV
p_{T}^Z, m_Z, y^Z	mass and rapidity of Z boson
$minM_{jj}^{ave}$	average (over number of jets in event) minimum invariant mass of jet pairs
M_{bb}	invariant mass of the two jets with the highest b-tagging weight
$\Delta R(l, l)$	cone between two leptons
PCBT_{Xj}	pseudo-continuous bin for X-th jet (where jets are ordered according to the b-tagging weight)
N_{had}^{top}	number of hadronic top candidates (jjb candidates with jj and jjb invariant mass within 15 GeV window around the nominal W and t mass, respectively)
N_{lep}^{top}	number of leptonic top candidates ($b\ell$ pairs with invariant mass below 155 GeV)
N_{had}^{W}	number of hadronic W candidates (jj pairs invariant mass within 15 GeV window around the nominal W mass
E_T^{miss}	missing transverse energy in the event
$p_T^{t\bar{t},spanet}$	transverse momentum of the $t\bar{t}$ system reconstructed from jets predicted by SPANet (see Section 9.1.2)

Table 11.3: Definitions of the discriminating variables used in the dilepton NNs. Jets and leptons used in the definitions are ordered by their p_{T} from the highest one. To suppress effect of the miss-modeling in the events with high jet multiplicity, only first 8 jets ordered by p_{T} are considered when calculating NN input variables.

The input layer consists of chosen variables which are scaled to fill in the range between 0 and 1 in order to reduce the variance of NN weights (referred to as *normalization* method from Section 7.3.1). The output layer consists of two nodes, each representing signal and background class, with sigmoid activation function. Binary cross-entropy, its definition given by Eq. 7.11, is used as a loss function that is minimized during the training.

Each SR employs *cross-training*, described in Section 7.3.1, to obtain models that are subsequently applied to unclassified events. Whole MC sample (signal as well as

[4]The first Fox-Wolfram momentum is defined by following formula:

$$H1 = \sum_{i,j} \frac{\vec{p_i} \cdot \vec{p_j}}{E_{viss}^2}, \tag{11.2}$$

where $\vec{p_i}$ and $\vec{p_j}$ represent 3-momenta of i-th and j-th object (lepton or jet) and E_{viss} corresponds to whole visible energy in the event.

	5j2b	6j2b	6j1b
Number of hidden layers	3	3	3
Number of nodes in layers	90,70,50	80,90,60	80,90,60
Activation functions	`elu`, `softsign`, `softplus`	`softsign`, `selu`, `relu`	`softsign`, `selu`, `relu`
Dropout layer and rate	1. with 0.3	1. with 0.3	1. with 0.3

Table 11.4: Optimized values of hyperparameters used in the dielepton NNs. Definitions of the used activation functions can be found in i.e. Ref [240]. As can be seen, identical architectures are optimal for both SRs featuring six and more jets.

background) is split into two halves (referred to as folds in the following) resulting in training and testing dataset. Two folds are hence used for training, where in the first fold training dataset is used for training and NN performance is evaluated on testing data, while in the second fold the two datasets are interchanged, meaning original testing dataset is used for training and original training dataset for evaluation.

Loss curves shown in Figure 11.3 confirm good convergence towards common loss values for both validation and training set. ROC curves, which can be found in Figure 11.4, prove reasonable separation power of each NN while keeping the overtraining at acceptable level. This can be seen also from Kolmogorov-Smirnov tests shown in Figures 11.5.

The trained NN models are subsequently applied to the MC events on which they were not trained according to the cross-training prescription. The output discriminants of the NNs, used as the inputs for the profile likelihood fit, can be found in Figure 11.6. Binning of these distributions were optimized according to the following procedure. Approach is based on merging initial fine bins using following variable as discriminating metric:

$$Z = z_b \frac{n_b}{N_b} + z_s \frac{n_s}{N_s}, \tag{11.3}$$

where $n_s = \sum_{i=a}^{b} n_s^i$, $N_s = \sum_{i=1}^{N_{fb}} n_s^i$ where n_s^i represents content of fine signal bin i and the sum goes over N_{fb} fine bins (a and b are lower and upper edges of optimized bin), and analogously for background events with subscript b. The free parameters z_b and z_s define the shape of the underlying signal and background distribution, while in dilepton analysis optimal values were found to be $z_s = z_b = 5$. The bin edges are optimized by merging fine bins until $Z > 1$, when the procedure continues with the next bin. As a result, each of the obtained NN output distributions contain in total 10 bins from 0 (background-like events) to 1 (signal-like events).

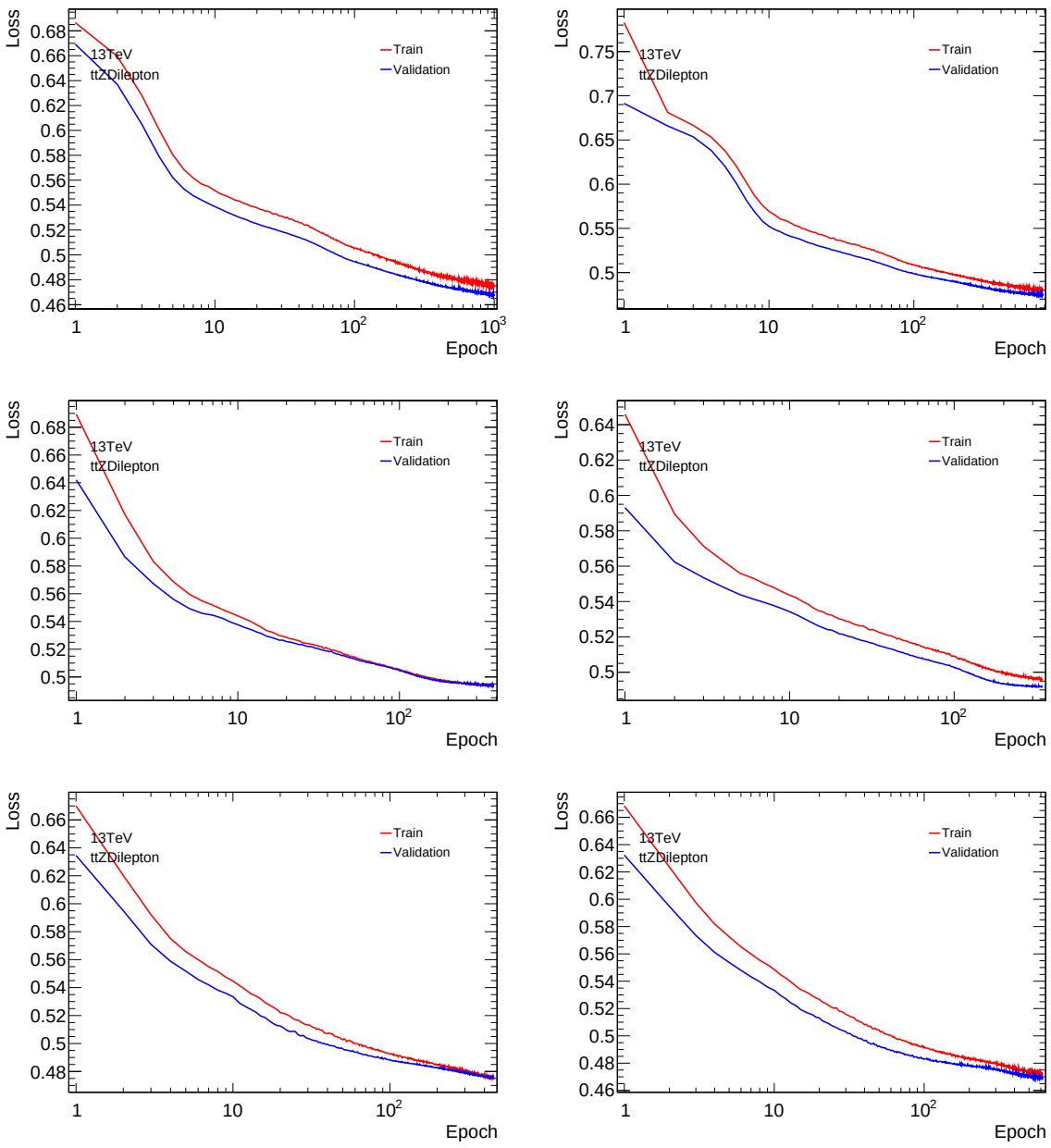


Fig. 11.3: Loss curves for training and validation sets for the first (left) and second (right) fold in 2ℓ-Z-$5j2b$ (top), 2ℓ-Z-$6j1b$ (middle) and 2ℓ-Z-$6j2b$ (bottom) signal region. The low overtraining is achieved by fast convergence of both sets to the same value. Note that the x-axes feature logarithmic scale.

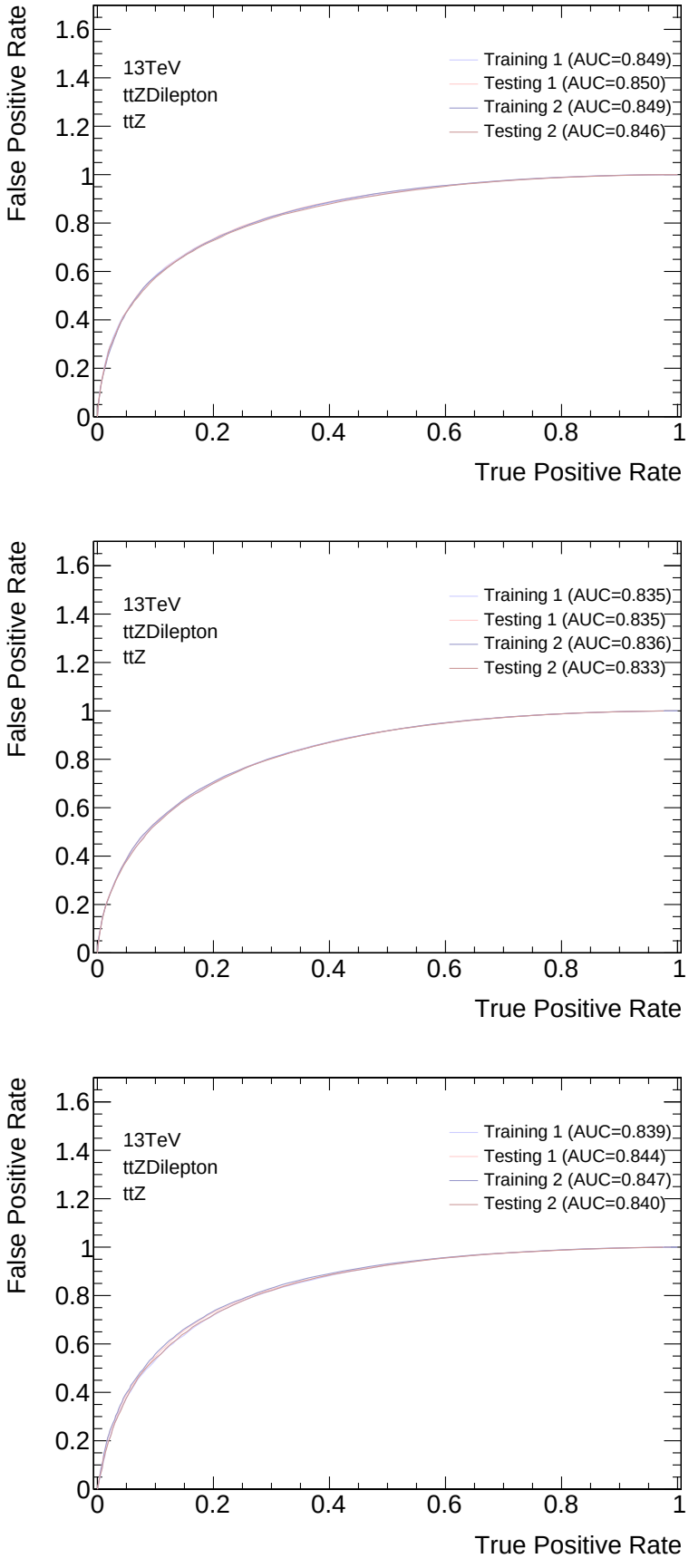


Fig. 11.4: ROC curves for 2ℓ-Z-$5j2b$ (top), 2ℓ-Z-$6j1b$ (middle) and 2ℓ-Z-$6j2b$ (bottom) signal regions. Since training and testing set produce very similar ROC curves, the overtraining is well controlled.

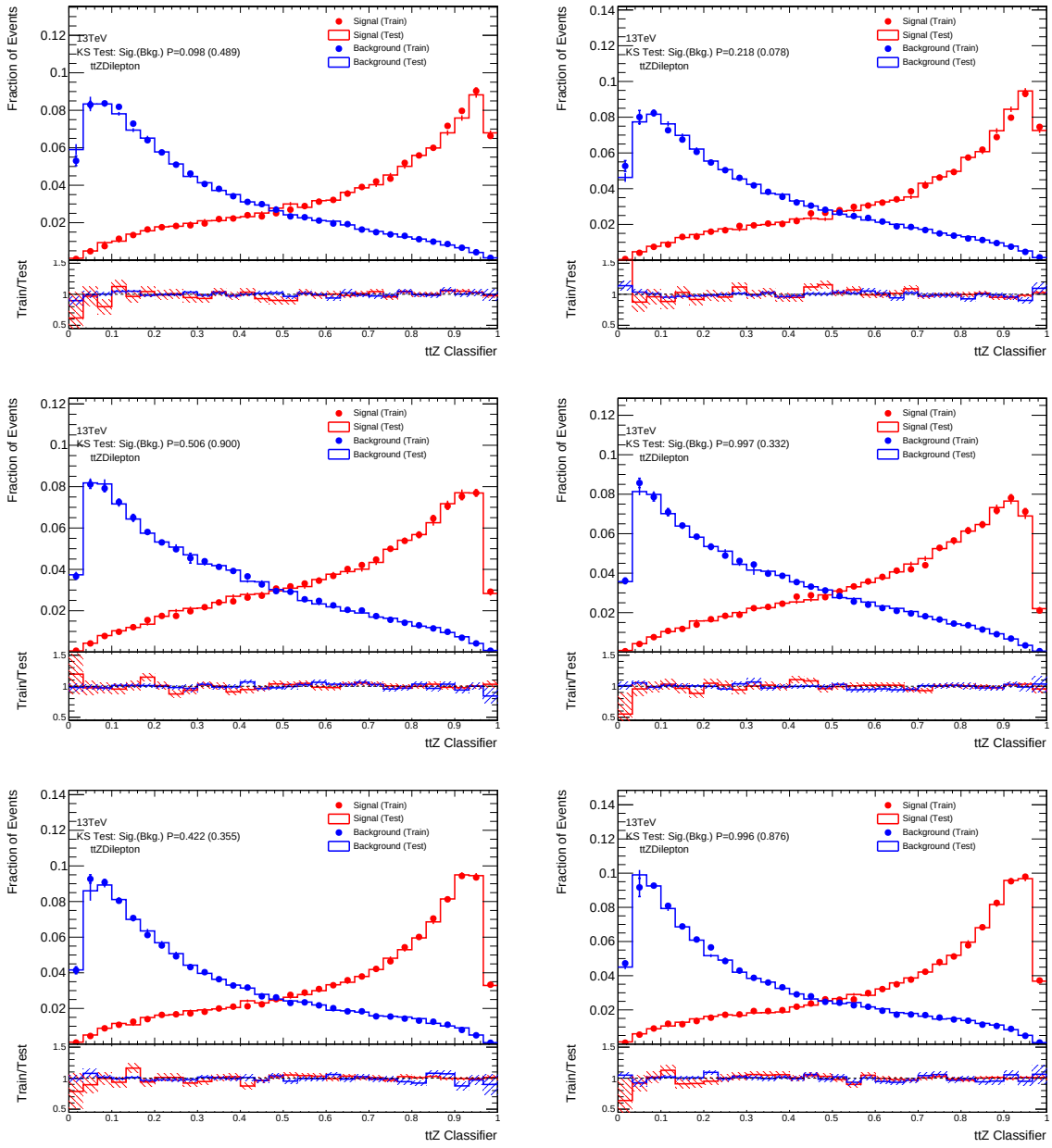


Fig. 11.5: Kolmogorov-Smirnov tests for first (left) and second (right) fold in 2ℓ-Z-$5j2b$ (top), 2ℓ-Z-$6j1b$ (middle) and 2ℓ-Z-$6j2b$ (bottom) signal regions. No significant deviation is observed between testing and training set indicating good generalization performance of the two folds.

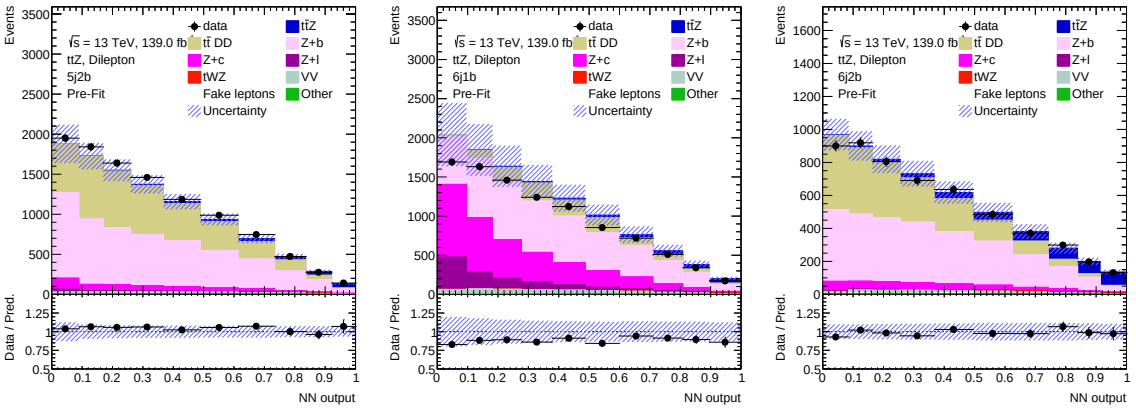


Fig. 11.6: The distributions of the NN output for the three dilepton SRs: 2ℓ-Z-$5j2b$ (left), 2ℓ-Z-$6j1b$ (middle) and 2ℓ-Z-$6j2b$ (right) as used in the fit.

11.4 Expected results

The fitting strategy described in previous Section 11.1 is applied for the Asimov fit in the dilepton SRs. MC distributions of the NN discriminator, shown in Figure 11.6, are simultaneously fitted in all three SRs to get the estimate of the expected signal strength $\mu_{t\bar{t}Z}^{2\ell}$, together with the normalization scale factors for $Z+b$ ($\mathcal{N}_{Z+b}$) and $Z+c$ ($\mathcal{N}_{Z+c}$) processes, which are treated as free parameters of the fit. The following values of the mentioned parameters of interest are obtained:

$$\begin{aligned} \mu_{t\bar{t}Z}^{2\ell} &= 1.000^{+0.124}_{-0.116} = 1.000^{+0.100}_{-0.089}(\text{syst.}) \pm 0.074(\text{stat.}), \\ \mathcal{N}_{Z+b} &= 1.000^{+0.113}_{-0.102}, \\ \mathcal{N}_{Z+c} &= 1.000^{+0.277}_{-0.250}. \end{aligned} \tag{11.4}$$

The fitted values of POIs agree with their MC predictions, verifying the correct convergence of the fit. The nuisance parameter plots, shown in Figure 11.7, summarize the compatibility between pre-fit (θ_0) and post-fit ($\hat{\theta}$) value of the given NP. The values represented by black dots centered at zero indicate no pull of any NP, while the black line corresponding to ± 1 (green bar) indicate no significant constraint (deviation from pre-fit uncertainty $\Delta\theta$) of any NP.

The impact of a NP (θ) on the POI ($\Delta\mu$) is given by the shift in the POI between the nominal fit and modified fit where the given NP is fixed to the value $\hat{\theta} \pm \Delta\theta$, where $\hat{\theta}$ is the post-fit value of the NP and $\Delta\theta$ is the shift introduced by a NP. This is typically evaluated before and after the fit to obtain pre- and post-fit $\Delta\mu$ values in order to reveal potential constraints introduced by the fit. The ranking plot showing leading 20 NPs with the highest impact on fitted POI, resulting from this Asimov fit, can be found in Figure 11.8. The highest-ranked uncertainties arise due to the choice of the μ_R and μ_F scale factors for $Z+b$ and signal processes.

The correlation matrix of the systematic sources (for correlation factors greater than

20%) can be found in Figure 11.9, and highlights expected behaviors: a large correlation (almost 40%) of the fitted $Z + b$ normalization factor with the leading systematic represented by μ_R and μ_F scale variation for $Z + b$. Somewhat large (anti)correlations are found between the fitted $Z + b/Z + c$ normalization factors and leading components of the systematic uncertainties associated with b tagging and Z+jets modeling systematics.

The similar measurements of the inclusive cross sections are conducted also in the trilepton and tetralepton channels, as a part of the ongoing $t\bar{t}Z$ analysis. These measurements are performed in the dedicated SRs and similarly employ neural networks for separating signal and background processes. Since author of this thesis is not directly involved in these measurements, these are not presented in this thesis. However, results of the fit performed with the Asimov datasets in the trilepton and tetralepton SRs are quoted in here in order to provide estimate of the total uncertainty on the cross section obtained from the combined fit in all three $t\bar{t}Z$ decay channels. The results of all three fits, as well as their combination, are quoted in Table 11.5. When the resulting total uncertainty is compared to the uncertainty obtained in the last $t\bar{t}Z$ inclusive cross section measurement (briefly summarized in the next section), the relative improvement of almost 45% can be expected.

Channel	$\mu_{t\bar{t}Z}$
Dilepton	$1.00 \pm 0.074\,(\text{stat.})\,^{+0.100}_{-0.089}\,(\text{syst.})$
Trilepton	$1.00 \pm 0.059\,(\text{stat.}) \pm 0.060\,(\text{syst.})$
Tetralepton	$1.00 \pm 0.116\,(\text{stat.}) \pm 0.048\,(\text{syst.})$
Combination ($2\ell + 3\ell + 4\ell$)	$1.00^{+0.043}_{-0.042}\,(\text{stat.})\,^{+0.054}_{-0.050}\,(\text{syst.})$

Table 11.5: Measured values of the expected $t\bar{t}Z$ signal strength parameters obtained from the Asimov fits in different analysis channels. Value for the combination of the dilepton, trilepton and tetralepton channels is obtained from the combined fit in respective SRs.

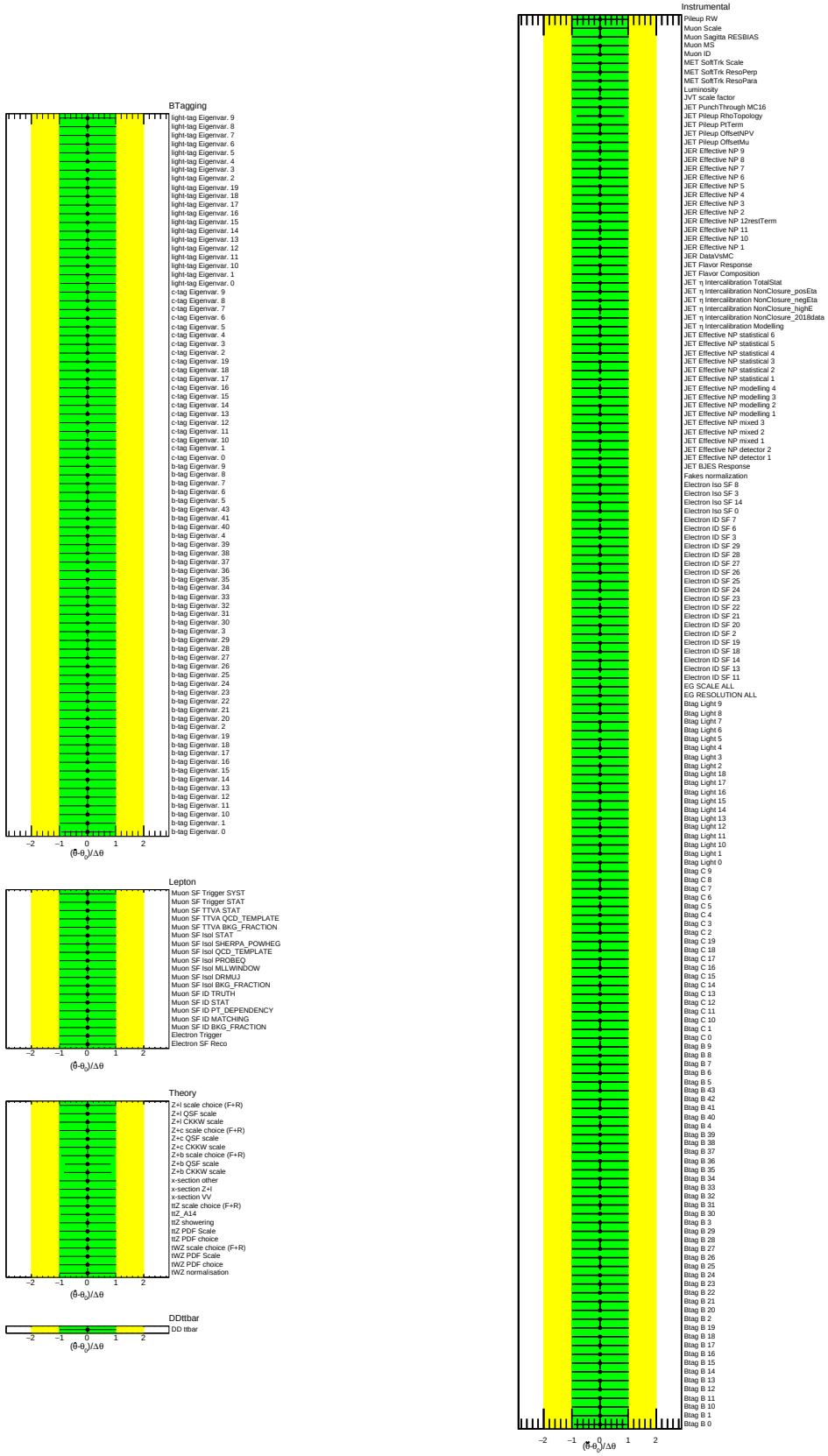


Fig. 11.7: Nuisance parameter plots for systematic uncertainties in the Asimov fit, grouped into several categories based on their origin. Green and yellow bands represent correspondence with $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty, respectively. The black dots and lines indicate agreement between pre- and post-fit value and uncertainty of a given NP.

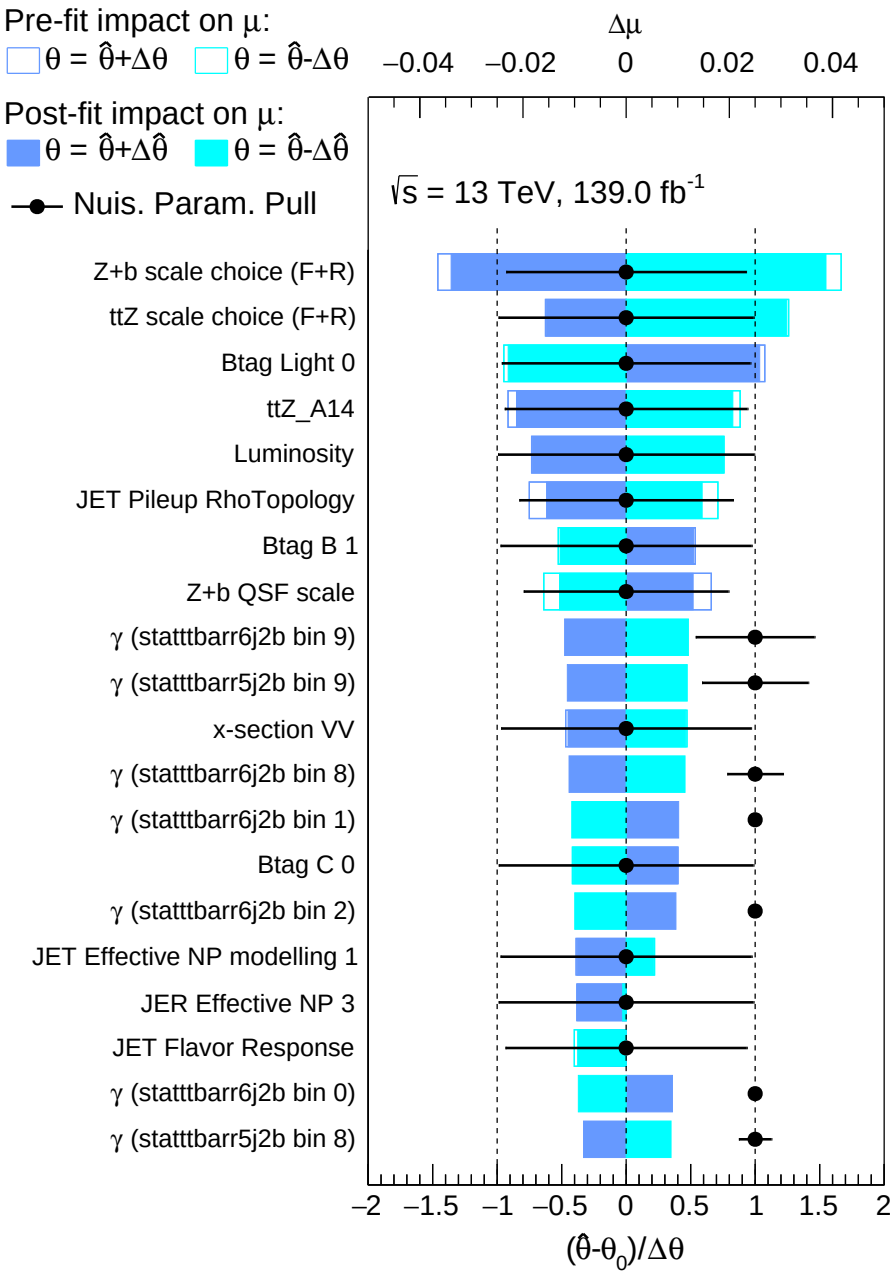


Fig. 11.8: The ranking plot showing impact of leading 20 NPs considered in the Asimov fit. The blue rectangles represent pre-fit (empty) and post-fit (filled) impact of a given systematic source on the measured value of POI. Black dots and error bars indicate pull and constrain of given NP providing actually the same message as NP plots in Figure 11.7. γ scale factors (their nominal value is one since they are applied as multiplicative factors) are related to the data-driven $t\bar{t}$ estimate and correspond to statistical uncertainties due to the limited number of data (see Section 7.4.1 for more details).

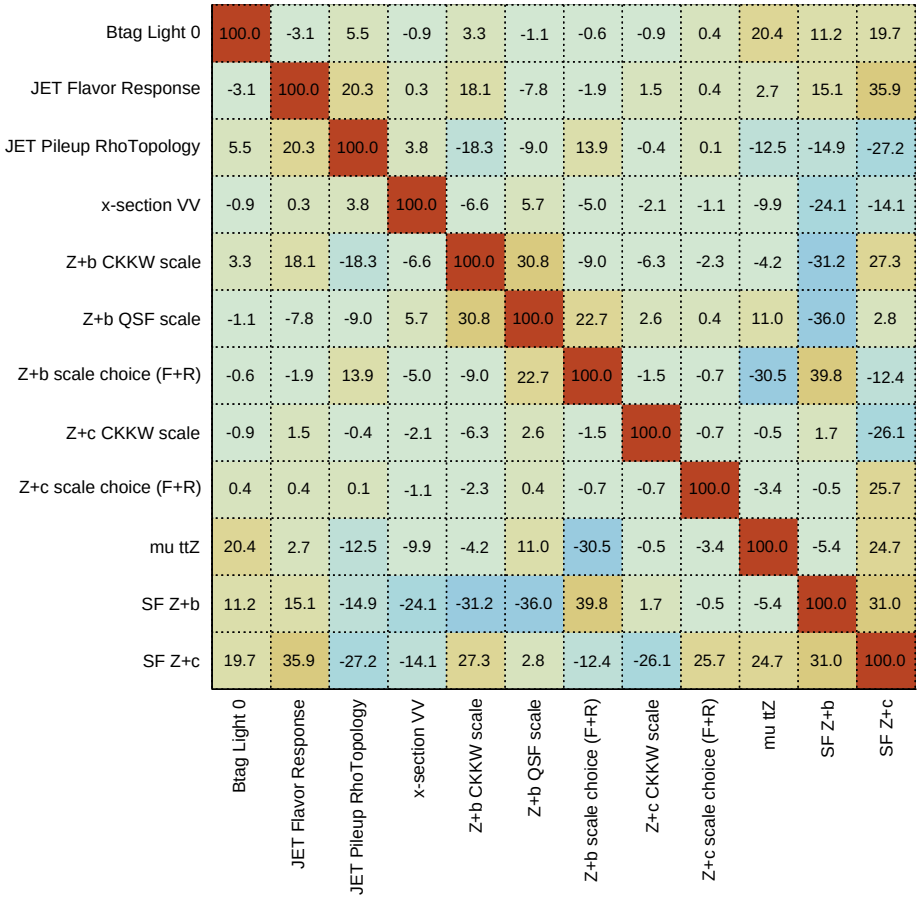


Fig. 11.9: Correlation matrix for systematic uncertainties in the Asimov fit. Only NPs with an absolute value of correlation larger than 20 % with at least one other NP are shown.

11.5 Inclusive cross section measurements in the trilepton and tetralepton channels

This section briefly presents results of the inclusive cross section measurement performed as a part of the previous round of the $t\bar{t}Z$ analysis [10] using full Run 2 data [5]. Although author of this thesis did not directly contribute to these measurements, the analysis is briefly presented here in order to provide complete recent $t\bar{t}Z$ results.

The strategy of the inclusive analysis is based on the same technique of profile likelihood fit as employed for the dilepton channel described in the previous sections. Since the previous $t\bar{t}Z$ analysis was focused only on the trilepton and tetralepton channels, the total inclusive $t\bar{t}Z$ cross section is measured as a combination of these two channels.

The definitions of the inclusive signal regions, which differ from those used in the differential analysis, are described in Sections 8.2 and 8.3, respectively for the trilepton and tetralepton channels. The observed and expected event yields in these inclusive SRs can be found in the same sections. The simultaneous fit is performed in two trilepton SRs (with different jet and b-jet multiplicity) and four tetralepton SRs (distinguished by b-jet multiplicity and flavor of the leptons from the $t\bar{t}$ pair). Contrary to the inclusive measurement in the dilepton channel, the fitted distributions in the trilepton and tetralepton SRs contain only single bin, thus effectively fitting number of events in each SR.

The fit strategy includes usage of two dedicated control regions (CRs), which are designed in order to obtain normalization of the most dominant background processes (WZ+jets in the trilepton and ZZ+jets in the tetralepton channel), from data. The selection criteria which are used for the definition of these CRs, are summarized in Table 11.6. Requirement on the invariant mass for OSSF lepton pairs is applied in both CR in order to select Z bosons. To ensure orthogonality of the WZ+jets CR (3ℓ-WZ-CR) with the trilepton SR, the b-jet veto is applied. Similarly, the orthogonality of the ZZ+jets CR (4ℓ-ZZ-CR) with the tetralepton SRs is ensured by applying cut on the missing transverse energy in the event. The resulting purity in the 3ℓ-WZ-CR achieves 80%, while in the 4ℓ-ZZ-CR goes up to 97%. The event yields obtained in the two CRs can be found in Table 11.7.

In the fit, the signal strength of the $t\bar{t}Z$, $\mu_{t\bar{t}Z}$, and the normalization of the light-flavor components of the WZ/ZZ backgrounds ($WZ/ZZ + l$) are treated as free parameters of the fit and all the systematic sources presented in Chapter 6 are considered as a nuisance parameters. The fitted signal strengths resulting from the separate fits

[5] This inclusive analysis is the part of the same paper as differential cross section measurements presented in this dissertation.

Variable	3ℓ-WZ-CR	4ℓ-ZZ-CR
N_ℓ ($\ell = e, \mu$)	$= 3$	$= 4$
	1 OSSF lepton pair with	2 OSSF lepton pairs with
	$\lvert m_{\ell\ell} - m_Z \rvert < 10GeV$	$\lvert m_{\ell\ell} - m_Z \rvert < 10$ GeV
$p_{\mathrm{T}}(\ell_1, \ell_2, \ell_3, \ell_4)$	$> 27, 20, 20$ GeV	$> 27, 20, 10, 7$ GeV
N_{jets}	≥ 3	–
$N_{b\text{-jets}}$@85%	$= 0$	–
$E_{\mathrm{T}}^{\mathrm{miss}}$	–	20 GeV $< E_{\mathrm{T}}^{\mathrm{miss}} <$ 40 GeV

Table 11.6: Definitions of the two control regions used in the inclusive cross section measurement. The purpose of these control regions is to obtain normalization of the light flavor component of dominant backgrounds (WZ/ZZ+jets) from data.

in the trilepton and tetralepton regions, as well as from the combined $3\ell + 4\ell$ fit, performed on the ATLAS data, are summarized in Table 11.8. None on the nuisance parameter considered in any fit feature significant pull or constrain. The fitted event yields in particular SRs, as well as CRs, resulting from the combined fit, can be found in Figure 11.10.

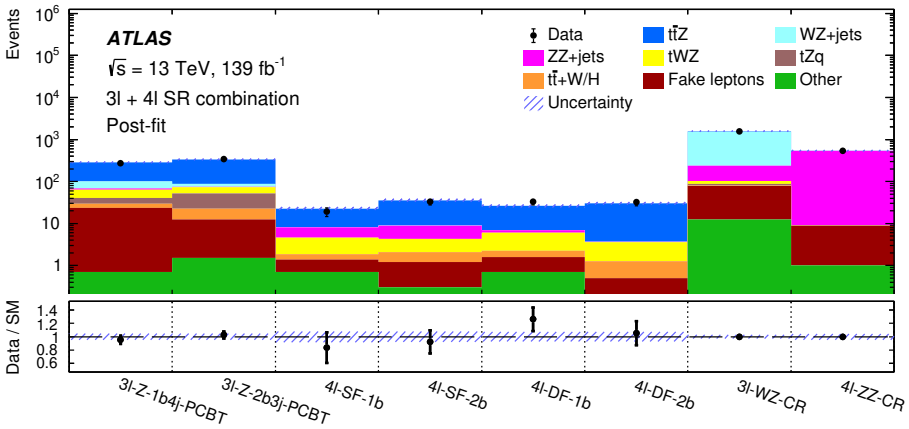


Fig. 11.10: The expected and observed (black dots) event yields as fitted in the trilepton and tetralepton SRs, as well as the WZ/ZZ+jets CRs. The bottom ratio plots shows compatibility between measured data and the total MC prediction. The quoted uncertainties combine both statistical, as well as systematic component.

The obtained signal strengths are subsequently converted to the values of the total cross sections. The combination of the trielepton and tetralepton channel thus yields following value of the cross section, which corresponds to the region of the phase space where invariant mass of the Z boson decay products lies in the range between 70 and

	3ℓ-WZ-CR
$t\bar{t}Z$	43.8 ± 11.3
$WZ + l$	1156.8 ± 368.9
$WZ + b$	16.7 ± 10.0
$WZ + c$	234.5 ± 103.0
ZZ + jets	131.8 ± 21.4
tWZ	13.3 ± 1.2
tZq	9.1 ± 3.5
$t\bar{t}$+$X(W/H)$	1.8 ± 0.38
Fake leptons	86.1 ± 43.3
Other	12.4 ± 6.4
Total SM	1706.3 ± 395.3
Data	1569

	4ℓ-ZZ-CR
$t\bar{t}Z$	0.7 ± 0.1
$ZZ + l$	460.5 ± 36.6
$ZZ + b$	14.3 ± 7.6
$ZZ + c$	20.7 ± 7.5
tWZ	0.2 ± 0.1
ttH	0.01 ± 0.01
Fake leptons	7.7 ± 3.6
Other	0.9 ± 0.6
Total SM	504.9 ± 39.0
Data	539

Table 11.7: Event yields obtained for the 3ℓ-WZ-CR (left) and 4ℓ-WZ-CR (right). The WZ/ZZ+jets backgrounds are decomposed into three categories based on their flavor compositions (featuring b,c, or light-flavor jet). The quoted uncertainties include both statistical component as well as all considered systematic sources described in Chapter 6.

Channel	$\mu_{t\bar{t}Z}$
Trilepton	$1.17 \pm 0.07\,(\text{stat.})\,^{+0.12}_{-0.11}\,(\text{syst.})$
Tetralepton	$1.21 \pm 0.15\,(\text{stat.})\,^{+0.11}_{-0.10}\,(\text{syst.})$
Combination $(3\ell + 4\ell)$	$1.19 \pm 0.06\,(\text{stat.}) \pm 0.10\,(\text{syst.})$

Table 11.8: Measured values of the $t\bar{t}Z$ signal strength parameters obtained from the fits in different analysis channels. Value for the combination of the trilepton and tetralepton channels is obtained from the combined fit in respective SRs.

110 GeV:

$$\sigma(pp \to t\bar{t}Z) = 0.99 \pm 0.05\,(\text{stat.}) \pm 0.08\,(\text{syst.})\,\text{pb}.$$

This value is consistent with the SM prediction of $0.84^{+0.09}_{-0.10}$ pb, calculated at the NLO in QCD and EW [241, 242], as well as with the most recent calculation including NNLL corrections [6, 243]. The total uncertainty is dominated by the systematic component, while the leading systematic sources are related to parton showering (alternative par-

ton shower is modeled by HERWIG 7 instead of PYTHIA 8), modeling of the major background processes, and b tagging. Symmetrized contributions of the particular systematic categories can be found in Table 11.9.

Uncertainty	$\Delta\sigma_{t\bar{t}Z}/\sigma_{t\bar{t}Z}$ [%]
$t\bar{t}Z$ parton shower	3.1
tWZ modelling	2.9
b-tagging	2.9
WZ/ZZ+jets modelling	2.8
tZq modelling	2.6
Lepton	2.3
Luminosity	2.2
Jets + $E_{\mathrm{T}}^{\mathrm{miss}}$	2.1
Fake leptons	2.1
$t\bar{t}Z$ ISR	1.6
$t\bar{t}Z$ μ_F and μ_R scales	0.9
Other backgrounds	0.7
Pile-up	0.7
$t\bar{t}Z$ PDF	0.2
Total systematic	8.4
Data statistics	5.2
Total	10

Table 11.9: Relative contributions of the systematic sources grouped into categories, to the overall systematic uncertainty on the measured $t\bar{t}Z$ inclusive cross section obtained from the combined fit. Note that the quadrature sum of the individual systematic sources is not consistent with the total cross section uncertainty reported in the text due to the correlations resulting from the combined fit.

Conclusion

This thesis presents two analyses related to the associated production of top-antitop-quark pair and Z boson, using the full Run 2 dataset of the LHC operation and corresponding to total integrated luminosity of about 139 fb^{-1} obtained at center-of-mass energy of 13 TeV. Both measurements provide opportunity to precisely test the predictions of the SM, and thus to either verify them or reveal signs of new physics beyond the SM.

The first part of the thesis is devoted to the differential cross section measurements in the trilepton and tetralepton $t\bar{t}Z$ decay channels (and their combination). Since the differential cross section for this SM process has never been measured by the ATLAS detector, the differential analysis requires extensive optimizations and tests of the used analysis techniques. The measurements are performed in two most sensitive $t\bar{t}Z$ decay channels and measure the differential cross sections as a function of nine variables. The observed distributions are corrected for detector and physics-related effects by performing iterative Bayesian unfolding. Both absolute and normalized differential cross sections, presented in Section 10.3.2, are measured at particle and parton level in particular fiducial volumes, and are compared to various MC, as well as dedicated theory predictions (for some variables at parton level) with at least NLO precision. Compatibility between measured data distributions and theory predictions are quantified by calculating χ^2/ndf and corresponding p-values that are summarized in Section 10.4. The obtained values confirm good agreement between data and predictions for most of the differential variables. The slightly poorer agreement is observed for variables $p_{\mathrm{T}}^{\ell\mathrm{non}-Z}$, p_{T}^{Z}, $|\Delta\phi(Z, t_{\mathrm{lep}})|$ and $|\Delta\phi(t\bar{t}, Z)|$, but in all cases p-value exceeds 0.05. Overall uncertainty on differential measurements in both channels is dominated by the statistical component, while in most cases it is larger than the difference between various predictions. For this reason, no definitive conclusion about the compatibility between measured data and prediction can be stated. The presented results were already published as a part of the latest ATLAS $t\bar{t}Z$ measurement [10].

The second analysis discussed in this thesis is focused on the inclusive cross section measurement in the dilepton $t\bar{t}Z$ channel. The profile likelihood fit of the neural network discriminant is performed in the three signal regions in order to extract signal strength $\mu_{t\bar{t}Z}^{2\ell}$. Since this measurement is part of the refined $t\bar{t}Z$ analysis, which has

not been finished yet, the fit is performed using Asimov dataset and yields following value of signal strength:

$$\mu_{t\bar{t}Z}^{2\ell} = 1.000^{+0.100}_{-0.089}(\text{syst.}) \pm 0.074(\text{stat.}). \tag{11.5}$$

The expected uncertainty of the Asimov fit is thus decreased by more than 50% when compared to result of the previous $t\bar{t}Z$ analysis performed in this channel [8], which, however, used only 36.1 fb^{-1} of LHC data.

The current analysis should now proceed with the fit on observed data and the result will be combined with the similar inclusive cross section measurements in the trilepton and tetralepton channels that are also part of the refined $t\bar{t}Z$ analysis. Moreover, current analysis aims to bring improvement also for the differential results presented in this dissertation by introducing novel unfolding technique based on the similar approach as profile likelihood fit.

The original aim of the dissertation, measurement of the $t\bar{t}Z$ inclusive cross section in the dilepton channel, was fulfilled and extended with the differential cross section measurement in the trilepton and tetralepton channels.

Bibliography

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

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

[3] ATLAS, CDF, CMS, and D0 Collaborations. First combination of Tevatron and LHC measurements of the top-quark mass. 2014. arXiv:1403.4427.

[4] 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.*, C78(2):129, 2018. arXiv:1709.04207, doi:10.1140/epjc/s10052-018-5595-5.

[5] CMS Collaboration. Measurement of the top quark Yukawa coupling from $\mathrm{t\bar{t}}$ kinematic distributions in the lepton+jets final state in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *Phys. Rev.*, D100(7):072007, 2019. arXiv:1907.01590, doi:10.1103/PhysRevD.100.072007.

[6] Anna Kulesza et al. Associated production of a top quark pair with a heavy electroweak gauge boson at NLO+NNLL accuracy. *Eur. Phys. J.*, C79(3):249, 2019. arXiv:1812.08622, doi:10.1140/epjc/s10052-019-6746-z.

[7] CMS Collaboration. Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{s}$ = 13 TeV. *Journal of High Energy Physics*, 2018(8), Aug 2018. URL: http://dx.doi.org/10.1007/JHEP08(2018)011, doi:10.1007/jhep08(2018)011.

[8] ATLAS Collaboration. Measurement of the $t\bar{t}z$ and $t\bar{t}w$ cross sections in proton-proton collisions at $\sqrt{s} = 13$ tev with the atlas detector. *Phys. Rev. D*, 99:072009, Apr 2019. URL: https://link.aps.org/doi/10.1103/PhysRevD.99.072009, doi:10.1103/PhysRevD.99.072009.

[9] CMS Collaboration. Measurement of top quark pair production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 03:056, 2020. arXiv:1907.11270, doi:10.1007/JHEP03(2020)056.

[10] ATLAS Collaboration. Measurements of the inclusive and differential production cross sections of a top-quark–antiquark pair in association with a Z boson at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 81(8):737, 2021. arXiv:2103.12603, doi:10.1140/epjc/s10052-021-09439-4.

[11] J.J. Thomson. Cathode Rays. *Philosophical Magazine*, 44:293–316, 1897.

[12] James Clerk Maxwell. A dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 155:459–513, 1865.

[13] Ian J. R. Aitchison and Anthony J. G. Hey. *Gauge theories in particle physics: a practical introduction; 4th ed.* CRC Press, Boca Raton, FL, 2013. URL: https://cds.cern.ch/record/1507184.

[14] Andrew Purcell. Go on a particle quest at the first CERN webfest. Le premier webfest du CERN se lance à la conquête des particules. page 10, Aug 2012. URL: https://cds.cern.ch/record/1473657.

[15] C.D. Anderson. The Positive Electron. *Phys. Rev.*, 43:491–494, 1933.

[16] Carl D. Anderson and Seth H. Neddermeyer. Cloud chamber observations of cosmic rays at 4300 meters elevation and near sea-level. *Phys. Rev.*, 50:263–271, Aug 1936. URL: https://link.aps.org/doi/10.1103/PhysRev.50.263, doi:10.1103/PhysRev.50.263.

[17] J. C. Street and E. C. Stevenson. New evidence for the existence of a particle of mass intermediate between the proton and electron. *Phys. Rev.*, 52:1003–1004, Nov 1937. URL: https://link.aps.org/doi/10.1103/PhysRev.52.1003, doi:10.1103/PhysRev.52.1003.

[18] Hideki Yukawa. On the interaction of elementary particles. *Proc. Phys. Math. Soc. Jap.*, 17:48–57, 1935.

[19] M. Conversi, E. Pancini, and O. Piccioni. On the disintegration of negative mesons. *Phys. Rev.*, 71:209–210, Feb 1947. URL: https://link.aps.org/doi/10.1103/PhysRev.71.209, doi:10.1103/PhysRev.71.209.

[20] M. L. Perl et al. Evidence for anomalous lepton production in $e^+ - e^-$ annihilation. *Phys. Rev. Lett.*, 35:1489–1492, Dec 1975. URL: https://link.aps.org/doi/10.1103/PhysRevLett.35.1489, doi:10.1103/PhysRevLett.35.1489.

[21] Particle Data Group. Review of Particle Physics. *Progress of Theoretical and Experimental Physics*, 2020(8), 08 2020. 083C01. arXiv:https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf, doi:10.1093/ptep/ptaa104.

[22] Wolfgang Pauli. Pauli letter collection: letter to Lise Meitner. Typed copy. URL: https://cds.cern.ch/record/83282.

[23] C. L. Cowan et al. Detection of the free neutrino: A Confirmation. *Science*, 124:103–104, 1956. doi:10.1126/science.124.3212.103.

[24] G. Danby et al. Observation of high-energy neutrino reactions and the existence of two kinds of neutrinos. *Phys. Rev. Lett.*, 9:36–44, Jul 1962. URL: https://link.aps.org/doi/10.1103/PhysRevLett.9.36, doi:10.1103/PhysRevLett.9.36.

[25] DONUT Collaboration. Observation of tau neutrino interactions. *Phys. Lett. B*, 504:218–224, 2001. arXiv:hep-ex/0012035, doi:10.1016/S0370-2693(01)00307-0.

[26] B Pontecorvo. Mesonium and antimesonium. *Soviet Phys. JETP*, 6, 2 1958. URL: https://www.osti.gov/biblio/4344536.

[27] Super-Kamiokande Collaboration. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998. URL: https://link.aps.org/doi/10.1103/PhysRevLett.81.1562, doi:10.1103/PhysRevLett.81.1562.

[28] Robert Hofstadter. *Electron Scattering and Nuclear and Nucleon Structure.* W.A.Benjamin, New York, 1963.

[29] M. Gell-Mann. A schematic model of baryons and mesons. *Physics Letters*, 8(3):214–215, 1964. URL: https://www.sciencedirect.com/science/article/pii/S0031916364920013, doi:https://doi.org/10.1016/S0031-9163(64)92001-3.

[30] G. Zweig. An SU(3) model for strong interaction symmetry and its breaking. Version 1. 1 1964.

[31] J. D. Bjorken and S. L. Glashow. Elementary Particles and SU(4). *Phys. Lett.*, 11:255–257, 1964. doi:10.1016/0031-9163(64)90433-0.

[32] S. L. Glashow, J. Iliopoulos, and L. Maiani. Weak Interactions with Lepton-Hadron Symmetry. *Phys. Rev. D*, 2:1285–1292, 1970. doi:10.1103/PhysRevD.2.1285.

[33] M. K. Gaillard and Benjamin W. Lee. Rare decay modes of the k mesons in gauge theories. *Phys. Rev. D*, 10:897–916, Aug 1974. URL: https://link.aps.org/doi/10.1103/PhysRevD.10.897, doi:10.1103/PhysRevD.10.897.

[34] J. J. Aubert et al. Experimental observation of a heavy particle j. *Phys. Rev. Lett.*, 33:1404–1406, Dec 1974. URL: https://link.aps.org/doi/10.1103/PhysRevLett.33.1404, doi:10.1103/PhysRevLett.33.1404.

[35] J. E. Augustin et al. Discovery of a narrow resonance in e^+e^- annihilation. *Phys. Rev. Lett.*, 33:1406–1408, Dec 1974. URL: https://link.aps.org/doi/10.1103/PhysRevLett.33.1406, doi:10.1103/PhysRevLett.33.1406.

[36] Makoto Kobayashi and Toshihide Maskawa. CP-Violation in the Renormalizable Theory of Weak Interaction. *Progress of Theoretical Physics*, 49(2):652–657, 02 1973. arXiv:https://academic.oup.com/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf, doi:10.1143/PTP.49.652.

[37] S. W. Herb et al. Observation of a dimuon resonance at 9.5 gev in 400 gev proton-nucleus collisions. *Phys. Rev. Lett.*, 39:252–255, Aug 1977. URL: https://link.aps.org/doi/10.1103/PhysRevLett.39.252, doi:10.1103/PhysRevLett.39.252.

[38] Marcela Carena, Andre de Gouvea, Ayres Freitas, and Michael Schmitt. Invisible Z boson decays at e^+ e^- colliders. *Phys. Rev. D*, 68:113007, 2003. arXiv:hep-ph/0308053, doi:10.1103/PhysRevD.68.113007.

[39] J. C. Maxwell. Viii. a dynamical theory of the electromagnetic field. *Philosophical Transactions of the Royal Society of London*, 155:459–512, 1865. doi:10.1098/rstl.1865.0008.

[40] R. A. Millikan. A direct determination of "h.". *Phys. Rev.*, 4:73–75, Jul 1914. URL: https://link.aps.org/doi/10.1103/PhysRev.4.73.2, doi:10.1103/PhysRev.4.73.2.

[41] P. A. M. Dirac and N. H. D. Bohr. The quantum theory of the emission and absorption of radiation. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 114:243–265, 1927. doi:10.1098/rspa.1927.0039.

[42] F. J. Dyson. The radiation theories of tomonaga, schwinger, and feynman. *Phys. Rev.*, 75:486–502, Feb 1949. URL: https://link.aps.org/doi/10.1103/PhysRev.75.486, doi:10.1103/PhysRev.75.486.

[43] G. Arnison et al. Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}$=540 gev. *Physics Letters B*, 122(1):103–116, 1983. URL: https://www.sciencedirect.com/science/article/pii/0370269383911772, doi:https://doi.org/10.1016/0370-2693(83)91177-2.

[44] M. Banner et al. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the cern pp collider. *Physics Letters B*, 122(5):476–485, 1983. URL: https://www.sciencedirect.com/science/article/pii/0370269383916052, doi:https://doi.org/10.1016/0370-2693(83)91605-2.

[45] G. Arnison et al. Experimental observation of lepton pairs of invariant mass around 95 gev/c^2 at the cern sps collider. *Physics Letters B*, 126(5):398–410, 1983. URL: https://www.sciencedirect.com/science/article/pii/0370269383901880, doi:https://doi.org/10.1016/0370-2693(83)90188-0.

[46] P. Bagnaia et al. Evidence for $z^0 \to e^+e^-$ at the cern pp collider. *Physics Letters B*, 129(1):130–140, 1983. URL: https://www.sciencedirect.com/science/article/pii/037026938390744X, doi:https://doi.org/10.1016/0370-2693(83)90744-X.

[47] D. P. Barber et al. Discovery of three-jet events and a test of quantum chromodynamics at petra. *Phys. Rev. Lett.*, 43:830–833, Sep 1979. URL: https://link.aps.org/doi/10.1103/PhysRevLett.43.830, doi:10.1103/PhysRevLett.43.830.

[48] Murray Gell-Mann. Symmetries of baryons and mesons. *Phys. Rev.*, 125:1067–1084, Feb 1962. URL: https://link.aps.org/doi/10.1103/PhysRev.125.1067, doi:10.1103/PhysRev.125.1067.

[49] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321, doi:10.1103/PhysRevLett.13.321.

[50] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508, doi:10.1103/PhysRevLett.13.508.

[51] ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, 2012. URL: https://www.sciencedirect.com/science/article/

pii/S037026931200857X, doi:https://doi.org/10.1016/j.physletb.2012.08.020.

[52] CMS Collaboration. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61, 2012. URL: https://www.sciencedirect.com/science/article/pii/S0370269312008581, doi:https://doi.org/10.1016/j.physletb.2012.08.021.

[53] Abdelhak Djouadi et al. HDECAY: Twenty$_{++}$ years after. *Comput. Phys. Commun.*, 238:214–231, 2019. arXiv:1801.09506, doi:10.1016/j.cpc.2018.12.010.

[54] M. Srednicki. *Quantum field theory.* Cambridge University Press, Geneva, 2007.

[55] Tal Roelof Van Daalen. Searches for heavy top partners with the ATLAS detector and irradiation studies of the Tile hadronic calorimeter, Feb 2021. Presented 23 Mar 2021. URL: https://cds.cern.ch/record/2783771.

[56] M. Y. Han and Y. Nambu. Three-triplet model with double SU(3) symmetry. *Phys. Rev.*, 139:B1006–B1010, Aug 1965. URL: https://link.aps.org/doi/10.1103/PhysRev.139.B1006, doi:10.1103/PhysRev.139.B1006.

[57] O. W. Greenberg. Spin and unitary-spin independence in a paraquark model of baryons and mesons. *Phys. Rev. Lett.*, 13:598–602, Nov 1964. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.598, doi:10.1103/PhysRevLett.13.598.

[58] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. Advantages of the color octet gluon picture. *Physics Letters B*, 47(4):365–368, 1973. URL: https://www.sciencedirect.com/science/article/pii/0370269373906254, doi:https://doi.org/10.1016/0370-2693(73)90625-4.

[59] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, 30:1343–1346, Jun 1973. URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1343, doi:10.1103/PhysRevLett.30.1343.

[60] H. David Politzer. Reliable perturbative results for strong interactions? *Phys. Rev. Lett.*, 30:1346–1349, Jun 1973. URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1346, doi:10.1103/PhysRevLett.30.1346.

[61] Antonio Pich. The standard model of electroweak interactions. 02 2005.

[62] Kenneth G. Wilson. Confinement of quarks. *Phys. Rev. D*, 10:2445–2459, Oct 1974. URL: https://link.aps.org/doi/10.1103/PhysRevD.10.2445, doi:10.1103/PhysRevD.10.2445.

[63] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, 30:1343–1346, Jun 1973. URL: https://link.aps.org/doi/10.1103/PhysRevLett.30.1343, doi:10.1103/PhysRevLett.30.1343.

[64] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588, 1961. URL: https://www.sciencedirect.com/science/article/pii/0029558261904692, doi:https://doi.org/10.1016/0029-5582(61)90469-2.

[65] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264, doi:10.1103/PhysRevLett.19.1264.

[66] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc. C*, 680519:367–377, 1968. doi:10.1142/9789812795915_0034.

[67] Gerard 't Hooft. Renormalization of Massless Yang-Mills Fields. *Nucl. Phys. B*, 33:173–199, 1971. doi:10.1016/0550-3213(71)90395-6.

[68] Gerard 't Hooft. Renormalizable Lagrangians for Massive Yang-Mills Fields. *Nucl. Phys. B*, 35:167–188, 1971. doi:10.1016/0550-3213(71)90139-8.

[69] ATLAS Collaboration. Standard Model Summary Plots June 2021. Technical report, CERN, Geneva, Jul 2021. URL: http://cds.cern.ch/record/2777014.

[70] Planck Collaboration. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020. [Erratum: Astron.Astrophys. 652, C4 (2021)]. arXiv:1807.06209, doi:10.1051/0004-6361/201833910.

[71] Planck Collaboration. Planck 2018 results. V. CMB power spectra and likelihoods. *Astron. Astrophys.*, 641:A5, 2020. arXiv:1907.12875, doi:10.1051/0004-6361/201936386.

[72] Adam G. Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009–1038, sep 1998. doi:10.1086/300499.

[73] K. G. Begeman, A. H. Broeils, and R. H. Sanders. Extended rotation curves of spiral galaxies: dark haloes and modified dynamics. *Monthly Notices of the Royal Astronomical Society*, 249(3):523–537, 04 1991. arXiv:https://academic.oup.com/mnras/article-pdf/249/3/523/18160929/mnras249-0523.pdf, doi:10.1093/mnras/249.3.523.

[74] Maxim Markevitch et al. Direct constraints on the dark matter self-interaction cross-section from the merging galaxy cluster 1E0657-56. *Astrophys. J.*, 606:819–824, 2004. arXiv:astro-ph/0309303, doi:10.1086/383178.

[75] A. D. Sakharov. Violation of CP Invariance, C asymmetry, and baryon asymmetry of the universe. *Pisma Zh. Eksp. Teor. Fiz.*, 5:32–35, 1967. URL: http://dx.doi.org/10.1070/PU1991v034n05ABEH002497.

[76] G.'t Hooft. *Naturalness, Chiral Symmetry, and Spontaneous Chiral Symmetry Breaking*, pages 135–157. Springer US, Boston, MA, 1980. doi:10.1007/978-1-4684-7571-5_9.

[77] S. Dimopoulos and H. Georgi. Softly broken supersymmetry and su(5). *Nuclear Physics B*, 193(1):150–162, 1981. URL: https://www.sciencedirect.com/science/article/pii/0550321381905228, doi:https://doi.org/10.1016/0550-3213(81)90522-8.

[78] H.P. Nilles. Supersymmetry, supergravity and particle physics. *Physics Reports*, 110(1):1–162, 1984. URL: https://www.sciencedirect.com/science/article/pii/0370157384900085, doi:https://doi.org/10.1016/0370-1573(84)90008-5.

[79] Th. Kaluza. Zum Unitätsproblem der Physik. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 1921:966–972, 1921. arXiv:1803.08616, doi:10.1142/S0218271818700017.

[80] O. Klein. The Atomicity of Electricity as a Quantum Theory Law. *Nature*, 118:516, 1926. doi:10.1038/118516a0.

[81] I. A. D'Souza and Calvin S. Kalman. *Preons: Models of leptons, quarks and gauge bosons as composite objects*. 1992.

[82] Eduard Boos et al. The top quark (20 years after its discovery). *Phys. Usp.*, 58(12):1133–1158, 2015. [Usp. Fiz. Nauk185,no.12,1241(2015)]. arXiv:1509.03325, doi:10.3367/UFNe.0185.201512a.1241.

[83] Jun Gao, Chong Sheng Li, and Hua Xing Zhu. Top-quark decay at next-to-next-to-leading order in qcd. *Phys. Rev. Lett.*, 110:042001, Jan 2013. URL: https://link.aps.org/doi/10.1103/PhysRevLett.110.042001, doi:10.1103/PhysRevLett.110.042001.

[84] Michał Czakon, Paul Fiedler, and Alexander Mitov. Total top-quark pair-production cross section at hadron colliders through $\mathcal{O}(\alpha_S^4)$. *Phys. Rev.*

Lett., 110:252004, Jun 2013. URL: https://link.aps.org/doi/10.1103/PhysRevLett.110.252004, doi:10.1103/PhysRevLett.110.252004.

[85] Michal Czakon, David Heymes, and Alexander Mitov. High-precision differential predictions for top-quark pairs at the LHC. *Phys. Rev. Lett.*, 116(8):082003, 2016. arXiv:1511.00549, doi:10.1103/PhysRevLett.116.082003.

[86] Michiel Botje et al. The PDF4LHC Working Group Interim Recommendations. 2011. arXiv:1101.0538.

[87] A. D. Martin et al. Uncertainties on α_S in global PDF analyses and implications for predicted hadronic cross sections. *Eur. Phys. J.*, C64:653–680, 2009. arXiv:0905.3531, doi:10.1140/epjc/s10052-009-1164-2.

[88] Jun Gao et al. CT10 next-to-next-to-leading order global analysis of QCD. *Phys. Rev.*, D89(3):033009, 2014. arXiv:1302.6246, doi:10.1103/PhysRevD.89.033009.

[89] Richard D. Ball et al. Parton distributions with LHC data. *Nucl. Phys.*, B867:244–289, 2013. arXiv:1207.1303, doi:10.1016/j.nuclphysb.2012.10.003.

[90] Michal Czakon and Alexander Mitov. Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders. *Comput. Phys. Commun.*, 185:2930, 2014. arXiv:1112.5675, doi:10.1016/j.cpc.2014.06.021.

[91] Daniel Joseph Sherman. *Measurement of the Top Quark Pair Production Cross Section with 1.12fb^{-1} of $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV.* PhD thesis, Harvard University, USA, 2007.

[92] Andrea Giammanco. Single top quark production at the LHC. *Rev. Phys.*, 1:1–12, 2016. arXiv:1511.06748, doi:10.1016/j.revip.2015.12.001.

[93] M. Aliev et al. HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR. *Comput. Phys. Commun.*, 182:1034–1046, 2011. arXiv:1007.1327, doi:10.1016/j.cpc.2010.12.040.

[94] P. Kant et al. HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions. *Comput. Phys. Commun.*, 191:74–89, 2015. arXiv:1406.4403, doi:10.1016/j.cpc.2015.02.001.

[95] Kevin Kröninger, Andreas B. Meyer, and Peter Uwer. Top-Quark Physics at the LHC. In Thomas Schörner-Sadenius, editor, *The Large Hadron Collider:*

Harvest of Run 1, pages 259–300. 2015. arXiv:1506.02800, doi:10.1007/978-3-319-15001-7_7.

[96] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev.*, D98(3):030001, 2018. doi:10.1103/PhysRevD.98.030001.

[97] Frederic Deliot and Douglas A. Glenzinski. Top Quark Physics at the Tevatron. *Rev. Mod. Phys.*, 84:211, 2012. arXiv:1010.1202, doi:10.1103/RevModPhys.84.211.

[98] UA1 Collaboration. Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c^2 at the CERN SPS Collider. *Phys. Lett.*, 126B:398–410, 1983. [,7.55(1983)]. doi:10.1016/0370-2693(83)90188-0.

[99] UA2 Collaboration. Evidence for $Z^0 \to e^+e^-$ at the CERN anti-pp Collider. *Phys. Lett.*, 129B:130–140, 1983. [,7.69(1983)]. doi:10.1016/0370-2693(83)90744-X.

[100] Matthias Schott. Study of the Z Boson Production at the ATLAS Experiment with First Data, 2007. URL: https://cds.cern.ch/record/2069156.

[101] U. Baur et al. Improved measurement of ttZ couplings at the CERN LHC. *Phys. Rev. D*, 73:034016, 2006. arXiv:hep-ph/0512262, doi:10.1103/PhysRevD.73.034016.

[102] Jinzhong Han, Bingfang Yang, and Xiantu Zhang. Associated production of the Z boson with a pair of top quarks in the left-right twin Higgs model. *EPL*, 105(3):31001, 2014. arXiv:1401.3594, doi:10.1209/0295-5075/105/31001.

[103] U. Baur et al. Probing electroweak top quark couplings at hadron and lepton colliders. *Nucl. Phys. B Proc. Suppl.*, 160:17–21, 2006. arXiv:hep-ph/0606264, doi:10.1016/j.nuclphysbps.2006.09.098.

[104] ATLAS Collaboration. Measurement of ttZ and ttW production at ATLAS in 13 TeV data, using trilepton and same charge dimuon final states. Technical Report ATL-PHYS-PROC-2016-117, CERN, Geneva, Aug 2016. URL: https://cds.cern.ch/record/2211022, doi:10.22323/1.276.0237.

[105] Lyndon Evans and Philip Bryant. LHC Machine. *JINST*, 3:S08001, 2008. doi:10.1088/1748-0221/3/08/S08001.

[106] LHC Guide. Mar 2017. URL: http://cds.cern.ch/record/2255762.

[107] ALICE Collaboration. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008. doi:10.1088/1748-0221/3/08/S08002.

[108] Image - CERN Accelerator Complex. http://inspirehep.net/record/823897/files/figures_chapter_2_fig2_02-lhc_accel_cplx.png. accessed 9.3.2022.

[109] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008. doi:10.1088/1748-0221/3/08/S08003.

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

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

[112] Synchrotron Radiation. http://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.synchrotron_radiation. accessed 9.3.2022.

[113] LHCb Collaboration. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008. doi:10.1088/1748-0221/3/08/S08005.

[114] TOTEM Collaboration. The TOTEM experiment at the CERN Large Hadron Collider. *JINST*, 3:S08007, 2008. doi:10.1088/1748-0221/3/08/S08007.

[115] Image - LHC Images. http://lhc-machine-outreach.web.cern.ch/lhc-machine-outreach/images/cryodipole.jpg. accessed 9.3.2022.

[116] The ATLAS Experiment. http://scipp.ucsc.edu/personnel/atlas.html. accessed 9.3.2022.

[117] A. Grummer. Operational experience with and performance of the atlas pixel detector at the large hadron collider. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 936:684 – 685, 2019. URL: http://www.sciencedirect.com/science/article/pii/S0168900218310830, doi:https://doi.org/10.1016/j.nima.2018.09.002.

[118] ATLAS Collaboration. ATLAS Insertable B-Layer Technical Design Report. 2010.

[119] A Abdesselam et al. The Data Acquisition and Calibration System for the ATLAS Semiconductor Tracker. Technical Report ATL-INDET-PUB-2007-012. ATL-COM-INDET-2007-015. CERN-ATL-COM-INDET-2007-015, CERN, Geneva, Oct 2007. URL: https://cds.cern.ch/record/1064019, doi:10.1088/1748-0221/3/01/P01003.

[120] ATLAS Collaboration. The ATLAS TRT Barrel Detector. *JINST*, 3:P02014, 2008. URL: https://cds.cern.ch/record/1094548, doi:10.1088/1748-0221/3/02/P02014.

[121] C. W. Fabjan and F. Gianotti. Calorimetry for particle physics. *Rev. Mod. Phys.*, 75:1243–1286, 2003. doi:10.1103/RevModPhys.75.1243.

[122] ATLAS Collaboration. *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996. URL: https://cds.cern.ch/record/331061.

[123] ATLAS Collaboration. *ATLAS calorimeter performance: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1996. URL: https://cds.cern.ch/record/331059.

[124] Image - ATLAS Detector Photos. http://hep.phys.sfu.ca/openhouse_2008/kiosk/gallery/calorimeters-combined-barrel.html. accessed 9.3.2022.

[125] Brock Moir. A Correction to the Modelled Jet Energy Resolution of the ATLAS Detector. Master's thesis, University of Alberta, USA, 2009.

[126] ATLAS Collaboration. *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1997. URL: https://cds.cern.ch/record/331068.

[127] Aranzazu Ruiz-Martinez and ATLAS Collaboration. The Run-2 ATLAS Trigger System. Technical Report ATL-DAQ-PROC-2016-003, CERN, Geneva, Feb 2016. URL: https://cds.cern.ch/record/2133909, doi:10.1088/1742-6596/762/1/012003.

[128] ATLAS Collaboration. *ATLAS magnet system: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1997. URL: https://cds.cern.ch/record/338080.

[129] Peter Jenni and Marzio Nessi. ATLAS Forward Detectors for Luminosity Measurement and Monitoring. Technical Report CERN-LHCC-2004-010. LHCC-I-014, CERN, Geneva, Mar 2004. URL: https://cds.cern.ch/record/721908.

[130] Peter Jenni, Markus Nordberg, Marzio Nessi, and Kerstin Jon-And. *ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity*. Technical Design Report ATLAS. CERN, Geneva, 2008. URL: https://cds.cern.ch/record/1095847.

[131] Peter Jenni, Marzio Nessi, and Markus Nordberg. Zero Degree Calorimeters for ATLAS. Technical Report CERN-LHCC-2007-001. LHCC-I-016, CERN, Geneva, Jan 2007. URL: http://cds.cern.ch/record/1009649.

[132] S. Grinstein. The atlas forward proton detector (afp). *Nuclear and Particle Physics Proceedings*, 273-275:1180 – 1184, 2016. URL: http://www.sciencedirect.com/science/article/pii/S2405601415006744, doi:https://doi.org/10.1016/j.nuclphysbps.2015.09.185.

[133] Joao Pequenao and Paul Schaffner. How ATLAS detects particles: diagram of particle paths in the detector. Jan 2013. URL: https://cds.cern.ch/record/1505342.

[134] T Cornelissen et al. The new ATLAS track reconstruction (NEWT). *Journal of Physics: Conference Series*, 119(3):032014, jul 2008. doi:10.1088/1742-6596/119/3/032014.

[135] ATLAS Collaboration. Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2. *Eur. Phys. J. C*, 77(10):673, 2017. arXiv:1704.07983, doi:10.1140/epjc/s10052-017-5225-7.

[136] R. Frühwirth. Application of kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2):444–450, 1987. URL: https://www.sciencedirect.com/science/article/pii/0168900287908874, doi:https://doi.org/10.1016/0168-9002(87)90887-4.

[137] 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(8):639, 2019. arXiv:1902.04655, doi:10.1140/epjc/s10052-019-7140-6.

[138] ATLAS Collaboration. Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data. *Journal of Instrumentation*, 14(12):P12006–P12006, dec 2019. doi:10.1088/1748-0221/14/12/p12006.

[139] W Lampl et al. Calorimeter Clustering Algorithms: Description and Performance. Technical report, CERN, Geneva, Apr 2008. URL: https://cds.cern.ch/record/1099735.

[140] A. Hoecker et al. Tmva - toolkit for multivariate data analysis, 2009. arXiv:physics/0703039.

[141] IsolationSelectionTool - CERN twiki. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/IsolationSelectionTool. accessed 28.02.2022.

[142] ATLAS Collaboration. Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$ =13 TeV. *Eur. Phys. J. C*, 76(5):292, 2016. arXiv:1603.05598, doi:10.1140/epjc/s10052-016-4120-y.

[143] ATLAS Collaboration. Muon reconstruction and identification efficiency in ATLAS using the full Run 2 *pp* collision data set at $\sqrt{s} = 13$ TeV. 8 2020.

[144] ATLAS Collaboration. Muon identification and reconstruction efficiencies in full run-2 dataset. https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/MUON-2019-03/. accessed 01.03.2022.

[145] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008. arXiv:0802.1189, doi:10.1088/1126-6708/2008/04/063.

[146] 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(8):689, 2021. arXiv:2007.02645, doi:10.1140/epjc/s10052-021-09402-3.

[147] ATLAS Collaboration. Jet reconstruction and performance using particle flow with the ATLAS Detector. *Eur. Phys. J. C*, 77(7):466, 2017. arXiv:1703.10485, doi:10.1140/epjc/s10052-017-5031-2.

[148] ATLAS Collaboration. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J. C*, 77:490, 2017. arXiv:1603.02934, doi:10.1140/epjc/s10052-017-5004-5.

[149] Yu.L Dokshitzer et al. Better jet clustering algorithms. *Journal of High Energy Physics*, 1997(08):001–001, aug 1997. doi:10.1088/1126-6708/1997/08/001.

[150] Stephen D. Ellis and Davison E. Soper. Successive combination jet algorithm for hadron collisions. *Phys. Rev. D*, 48:3160–3166, Oct 1993. URL: https://link.aps.org/doi/10.1103/PhysRevD.48.3160, doi:10.1103/PhysRevD.48.3160.

[151] Gavin P. Salam and Gregory Soyez. A Practical Seedless Infrared-Safe Cone jet algorithm. *JHEP*, 05:086, 2007. arXiv:0704.0292, doi:10.1088/1126-6708/2007/05/086.

[152] ATLAS Collaboration. Tagging and suppression of pileup jets with the ATLAS detector. Technical report, CERN, Geneva, May 2014. URL: https://cds.cern.ch/record/1700870.

[153] ATLAS Collaboration. Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector. *Eur. Phys. J. C*, 77(9):580, 2017. arXiv:1705.02211, doi:10.1140/epjc/s10052-017-5081-5.

[154] ATLAS Collaboration. Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data. Technical report, CERN, Geneva, Jul 2011. URL: https://cds.cern.ch/record/1369219.

[155] ATLAS Collaboration. Optimisation and performance studies of the ATLAS *b*-tagging algorithms for the 2017-18 LHC run. Technical report, CERN, Geneva, Jul 2017. URL: https://cds.cern.ch/record/2273281.

[156] G Piacquadio and C Weiser. A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS. *Journal of Physics: Conference Series*, 119(3):032032, jul 2008. doi:10.1088/1742-6596/119/3/032032.

[157] ATLAS Collaboration. Performance of b-jet identification in the ATLAS experiment. *Journal of Instrumentation*, 11(04):P04008–P04008, apr 2016. doi:10.1088/1748-0221/11/04/p04008.

[158] ATLAS Collaboration. Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV. *JHEP*, 08:089, 2018. arXiv:1805.01845, doi:10.1007/JHEP08(2018)089.

[159] ATLAS Collaboration. Optimisation of the ATLAS *b*-tagging performance for the 2016 LHC Run. Technical report, CERN, Geneva, Jun 2016. URL: https://cds.cern.ch/record/2160731.

[160] The Theano Development Team et al. Theano: A python framework for fast computation of mathematical expressions, 2016. arXiv:1605.02688.

[161] ATLAS Collaboration. ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J. C*, 79(11):970, 2019. arXiv:1907.05120, doi:10.1140/epjc/s10052-019-7450-8.

[162] 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*, 78:903, 2018. arXiv:1802.08168, doi:10.1140/epjc/s10052-018-6288-9.

[163] Luminosity Public Results Run 2. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2. accessed 15.03.2022.

[164] Stefan Höche. Introduction to parton-shower event generators. In *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, pages 235–295, 2015. arXiv:1411.4085, doi:10.1142/9789814678766_0005.

[165] Yuri L. Dokshitzer. Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics. *Sov. Phys. JETP*, 46:641–653, 1977.

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

[167] G. Altarelli and G. Parisi. Asymptotic freedom in parton language. *Nuclear Physics B*, 126(2):298–318, 1977. URL: https://www.sciencedirect.com/science/article/pii/0550321377903844, doi:https://doi.org/10.1016/0550-3213(77)90384-4.

[168] Stefano Catani et al. Qcd matrix elements + parton showers. *Journal of High Energy Physics*, 2001(11):063–063, nov 2001. doi:10.1088/1126-6708/2001/11/063.

[169] Michelangelo L Mangano et al. Matching matrix elements and shower evolution for top-pair production in hadronic collisions. *Journal of High Energy Physics*, 2007(01):013–013, jan 2007. doi:10.1088/1126-6708/2007/01/013.

[170] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014. arXiv:1405.0301, doi:10.1007/JHEP07(2014)079.

[171] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007. arXiv:0709.2092, doi:10.1088/1126-6708/2007/11/070.

[172] Torbjörn Sjöstrand et al. An introduction to pythia 8.2. *Computer Physics Communications*, 191:159–177, 2015. URL: https://www.sciencedirect.com/science/article/pii/S0010465515000442, doi:https://doi.org/10.1016/j.cpc.2015.01.024.

[173] Johannes Bellm et al. Herwig 7.0/Herwig++ 3.0 release note. *Eur. Phys. J. C*, 76(4):196, 2016. arXiv:1512.01178, doi:10.1140/epjc/s10052-016-4018-8.

[174] Peter Freeman, Stephen Doe, and Aneta Siemiginowska. Sherpa: a mission-independent data analysis application. In Jean-Luc Starck and Fionn D.

Murtagh, editors, *Astronomical Data Analysis*, volume 4477, pages 76 – 87. International Society for Optics and Photonics, SPIE, 2001. doi:10.1117/12.447161.

[175] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand. Parton fragmentation and string dynamics. *Physics Reports*, 97(2):31–145, 1983. URL: https://www.sciencedirect.com/science/article/pii/0370157383900807, doi:https://doi.org/10.1016/0370-1573(83)90080-7.

[176] Torbjörn Sjöstrand. Jet fragmentation of multiparton configurations in a string framework. *Nuclear Physics B*, 248:469–502, 01 1985. doi:10.1016/0550-3213(84)90607-2.

[177] B.R. Webber. A qcd model for jet fragmentation including soft gluon interference. *Nuclear Physics B*, 238(3):492–528, 1984. URL: https://www.sciencedirect.com/science/article/pii/055032138490333X, doi:https://doi.org/10.1016/0550-3213(84)90333-X.

[178] G. Marchesini and B.R. Webber. Monte carlo simulation of general hard processes with coherent qcd radiation. *Nuclear Physics B*, 310(3):461–526, 1988. URL: https://www.sciencedirect.com/science/article/pii/0550321388900892, doi:https://doi.org/10.1016/0550-3213(88)90089-2.

[179] S. Agostinelli et al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003. URL: https://www.sciencedirect.com/science/article/pii/S0168900203013688, doi:https://doi.org/10.1016/S0168-9002(03)01368-8.

[180] J. Allison et al. Geant4 developments and applications. *IEEE Transactions on Nuclear Science*, 53(1):270–278, 2006. doi:10.1109/TNS.2006.869826.

[181] J. Allison et al. Recent developments in geant4. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 835:186–225, 2016. URL: https://www.sciencedirect.com/science/article/pii/S0168900216306957, doi:https://doi.org/10.1016/j.nima.2016.06.125.

[182] ATLAS Collboration et al. The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim. Technical report, CERN, Geneva, Oct 2010. URL: https://cds.cern.ch/record/1300517.

[183] K Edmonds et al. The Fast ATLAS Track Simulation (FATRAS). Technical report, CERN, Geneva, Mar 2008. URL: https://cds.cern.ch/record/1091969.

[184] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014. arXiv:1405.0301, doi:10.1007/JHEP07(2014)079.

[185] Richard D. Ball et al. Parton distributions for the LHC run II. *JHEP*, 04:040, 2015. arXiv:1410.8849, doi:10.1007/JHEP04(2015)040.

[186] Stefano Frixione et al. Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations. *JHEP*, 04:081, 2007. arXiv:hep-ph/0702198, doi:10.1088/1126-6708/2007/04/081.

[187] Pierre Artoisenet et al. Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations. *JHEP*, 03:015, 2013. arXiv:1212.3460, doi:10.1007/JHEP03(2013)015.

[188] ATLAS Collaboration. ATLAS Run 1 Pythia8 tunes. Technical Report ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov 2014. URL: https://cds.cern.ch/record/1966419.

[189] D. J. Lange. The EvtGen particle decay simulation package. *Nucl. Instrum. Meth. A*, 462:152, 2001. doi:10.1016/S0168-9002(01)00089-4.

[190] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2016. arXiv:1610.07922, doi:10.23731/CYRM-2017-002.

[191] ATLAS Collaboration. Modelling of the $t\bar{t}H$ and $t\bar{t}V(V = W, Z)$ processes for $\sqrt{s} = 13$ TeV ATLAS analyses. ATL-PHYS-PUB-2016-005, 2016. URL: https://cds.cern.ch/record/2120826.

[192] Enrico Bothmann et al. Event generation with Sherpa 2.2. *SciPost Phys.*, 7(3):034, 2019. arXiv:1905.09127, doi:10.21468/SciPostPhys.7.3.034.

[193] M. Bähr et al. Herwig++ physics and manual. *Eur. Phys. J. C*, 58:639, 2008. arXiv:0803.0883, doi:10.1140/epjc/s10052-008-0798-9.

[194] Federico Demartin et al. tWH associated production at the LHC. *Eur. Phys. J. C*, 77(1):34, 2017. arXiv:1607.05862, doi:10.1140/epjc/s10052-017-4601-7.

[195] Stefan Höche, Frank Krauss, Marek Schönherr, and Frank Siegert. QCD matrix elements + parton showers. The NLO case. *JHEP*, 04:027, 2013. arXiv:1207.5030, doi:10.1007/JHEP04(2013)027.

[196] Fabio Cascioli, Philipp Maierhöfer, and Stefano Pozzorini. Scattering Amplitudes with Open Loops. *Phys. Rev. Lett.*, 108:111601, 2012. arXiv:1111.5206, doi:10.1103/PhysRevLett.108.111601.

[197] Simone Alioli et al. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010. arXiv:1002.2581, doi:10.1007/JHEP06(2010)043.

[198] ATLAS Collaboration. Measurement of the top quark-pair production cross section with ATLAS in *pp* collisions at $\sqrt{s} = 7\,\text{TeV}$. *Eur. Phys. J. C*, 71:1577, 2011. arXiv:1012.1792, doi:10.1140/epjc/s10052-011-1577-6.

[199] ATLAS Collaboration. Search for supersymmetry at $\sqrt{s} = 8\,\text{TeV}$ in final states with jets and two same-sign leptons or three leptons with the ATLAS detector. *JHEP*, 06:035, 2014. arXiv:1404.2500, doi:10.1007/JHEP06(2014)035.

[200] Erich W. Varnes. A Poisson likelihood approach to fake lepton estimation with the matrix method. 6 2016. arXiv:1606.06817.

[201] ATLAS Collaboration. Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data. 6 2016.

[202] ATLAS Collaboration. Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s} = 13$ TeV. 2016.

[203] ATLAS Collaboration. Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s} = 13$ TeV. 6 2015.

[204] Craig Sawyer et al. Monte Carlo Calibration and Combination of In-situ Measurements of Jets in ATLAS. Technical report, CERN, Geneva, Jan 2015. URL: https://cds.cern.ch/record/1987296.

[205] ATLAS Collaboration. Optimisation of the smoothing of *b*-jet identification efficiency and mistag rate simulation-to-data scale factors in ATLAS. 2 2020.

[206] Zachary Marshall. Simulation of Pile-up in the ATLAS Experiment. *J. Phys. Conf. Ser.*, 513:022024, 2014. doi:10.1088/1742-6596/513/2/022024.

[207] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J. C*, 76(12):653, 2016. arXiv:1608.03953, doi:10.1140/epjc/s10052-016-4466-1.

[208] Jon Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys. G*, 43:023001, 2016. arXiv:1510.03865, doi:10.1088/0954-3899/43/2/023001.

[209] Sayipjamal Dulat et al. New parton distribution functions from a global analysis of quantum chromodynamics. *Phys. Rev. D*, 93:033006, Feb 2016. URL: https://link.aps.org/doi/10.1103/PhysRevD.93.033006, doi:10.1103/PhysRevD.93.033006.

[210] L. A. Harland-Lang et al. Parton distributions in the LHC era: MMHT 2014 PDFs. *Eur. Phys. J. C*, 75(5):204, 2015. arXiv:1412.3989, doi:10.1140/epjc/s10052-015-3397-6.

[211] Steffen Schumann and Frank Krauss. A Parton shower algorithm based on Catani-Seymour dipole factorisation. *JHEP*, 03:038, 2008. arXiv:0709.1027, doi:10.1088/1126-6708/2008/03/038.

[212] ATLAS Collaboration. Measurements of the production cross section of a Z boson in association with jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 77(6):361, 2017. arXiv:1702.05725, doi:10.1140/epjc/s10052-017-4900-z.

[213] ATLAS Collaboration. Measurement of the production cross-section of a single top quark in association with a Z boson in proton-proton collisions at 13 TeV with the ATLAS detector. *Phys. Lett. B*, 780:557–577, 2018. arXiv:1710.03659, doi:10.1016/j.physletb.2018.03.023.

[214] CMS Collaboration. Measurement of the associated production of a single top quark and a Z boson in pp collisions at $\sqrt{s} = 13$ TeV. *Phys. Lett. B*, 779:358–384, 2018. arXiv:1712.02825, doi:10.1016/j.physletb.2018.02.025.

[215] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2/2017, 10 2016. arXiv:1610.07922, doi:10.23731/CYRM-2017-002.

[216] Georgios Choudalakis. Unfolding in ATLAS. In *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*, pages 297–308, Geneva, 2011. CERN, CERN. arXiv:1104.2962, doi:10.5170/CERN-2011-006.297.

[217] Stefan Schmitt. Data unfolding methods in high energy physics. *EPJ Web of Conferences*, 137, 11 2016. doi:10.1051/epjconf/201713711008.

[218] V Kartvelishvili. Unfolding with Singular Value Decomposition. pages 264–270. 7 p, Jan 2011. URL: http://cds.cern.ch/record/2203260, doi:10.5170/CERN-2011-006.264.

[219] G. D'Agostini. A multidimensional unfolding method based on bayes' theorem. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 362(2):487 – 498, 1995. URL: http://www.sciencedirect.com/science/article/pii/016890029500274X, doi:https://doi.org/10.1016/0168-9002(95)00274-X.

[220] G. D'Agostini. Improved iterative bayesian unfolding, 2010. arXiv:1010.0632.

[221] Georgios Choudalakis. Fully bayesian unfolding, 2012. arXiv:1201.4612.

[222] Olaf Behnke, Kevin Kröninger, Thomas Schörner-Sadenius, and Gregory Schott, editors. *Data analysis in high energy physics: A practical guide to statistical methods*. Wiley-VCH, Weinheim, Germany, 2013.

[223] Raul Rojas. *Neural Networks - A Systematic Introduction*. Springer-Verlag, Berlin, 1996. URL: http://www.inf.fu-berlin.de/inst/ag-ki/rojas_home/pmwiki/pmwiki.php?n=Books.NeuralNetworksBook.

[224] Francois Chollet et al. Keras, 2015. URL: https://github.com/fchollet/keras.

[225] Martín Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org. URL: https://www.tensorflow.org/.

[226] Christopher M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, Inc., USA, 1995.

[227] Karl Pearson. Note on Regression and Inheritance in the Case of Two Parents. *Proceedings of the Royal Society of London Series I*, 58:240–242, January 1895.

[228] Kyle Cranmer et al. HistFactory: A tool for creating statistical models for use with RooFit and RooStats. Technical report, New York U., New York, Jan 2012. URL: https://cds.cern.ch/record/1456844.

[229] Roger Barlow and Christine Beeston. Fitting using finite monte carlo samples. *Computer Physics Communications*, 77(2):219–228, 1993. URL: https://www.sciencedirect.com/science/article/pii/001046559390005W, doi:https://doi.org/10.1016/0010-4655(93)90005-W.

[230] F. James and M. Roos. Minuit - a system for function minimization and analysis of the parameter errors and correlations. *Computer Physics Communications*, 10(6):343–367, 1975. URL: https://www.sciencedirect.com/science/article/pii/0010465575900399, doi:https://doi.org/10.1016/0010-4655(75)90039-9.

[231] R. Fletcher and M. J. D. Powell. A Rapidly Convergent Descent Method for Minimization. *The Computer Journal*, 6(2):163–168, 08 1963. arXiv:https://academic.oup.com/comjnl/article-pdf/6/2/163/1041527/6-2-163.pdf, doi:10.1093/comjnl/6.2.163.

[232] R. Fletcher. A new approach to variable metric algorithms. *The Computer Journal*, 13(3):317–322, 01 1970. arXiv:https://academic.oup.com/comjnl/article-pdf/13/3/317/988678/130317.pdf, doi:10.1093/comjnl/13.3.317.

[233] Jerome H. Friedman. Data Analysis Techniques for High-Energy Particle Physics. In *3rd CERN School of Computing*, page 271, 10 1974.

[234] Alexander Shmakov et al. SPANet: Generalized Permutationless Set Assignment for Particle Physics using Symmetry Preserving Attention. 6 2021. arXiv:2106.03898.

[235] Johannes Erdmann et al. A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework. *Nucl. Instrum. Meth. A*, 748:18–25, 2014. arXiv:1312.5595, doi:10.1016/j.nima.2014.02.029.

[236] Lars Hertel et al. Sherpa: Robust hyperparameter optimization for machine learning, 2020. URL: https://arxiv.org/abs/2005.04048, doi:10.48550/ARXIV.2005.04048.

[237] Tim Adye. Unfolding algorithms and tests using RooUnfold. In *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*, pages 313–318, Geneva, 2011. CERN, CERN. arXiv:1105.1160, doi:10.5170/CERN-2011-006.313.

[238] Alessandro Broggio et al. Top-quark pair hadroproduction in association with a heavy boson at NLO+NNLL including EW corrections. *JHEP*, 08:039, 2019. arXiv:1907.04343, doi:10.1007/JHEP08(2019)039.

[239] ATLAS Collaboration. Measurements of top-quark pair single- and double-differential cross-sections in the all-hadronic channel in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector. *JHEP*, 01:033, 2021. arXiv:2006.09274, doi:10.1007/JHEP01(2021)033.

[240] Keras - layer activation functions. accessed 24.4.2022. URL: https://keras.io/api/layers/activations/.

[241] 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*, 06:184, 2015. arXiv:1504.03446, doi:10.1007/JHEP06(2015)184.

[242] D. de Florian et al. Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector. 2/2017, 10 2016. arXiv:1610.07922, doi:10.23731/CYRM-2017-002.

[243] Alessandro Broggio, Andrea Ferroglia, Giovanni Ossola, Ben D. Pecjak, and Ray D. Sameshima. Associated production of a top pair and a Z boson at the LHC to NNLL accuracy. *JHEP*, 04:105, 2017. arXiv:1702.00800, doi:10.1007/JHEP04(2017)105.

Appendices

A. Differential cross section measurements

This appendix contains additional figures related to the differential cross section measurements. Since the main body of the thesis includes only figures for representative variable in each channel, the corresponding plots for remaining differential variables are shown here.

A.1 Additional pre-unfolding plots

A.1.1 Trilepton channel

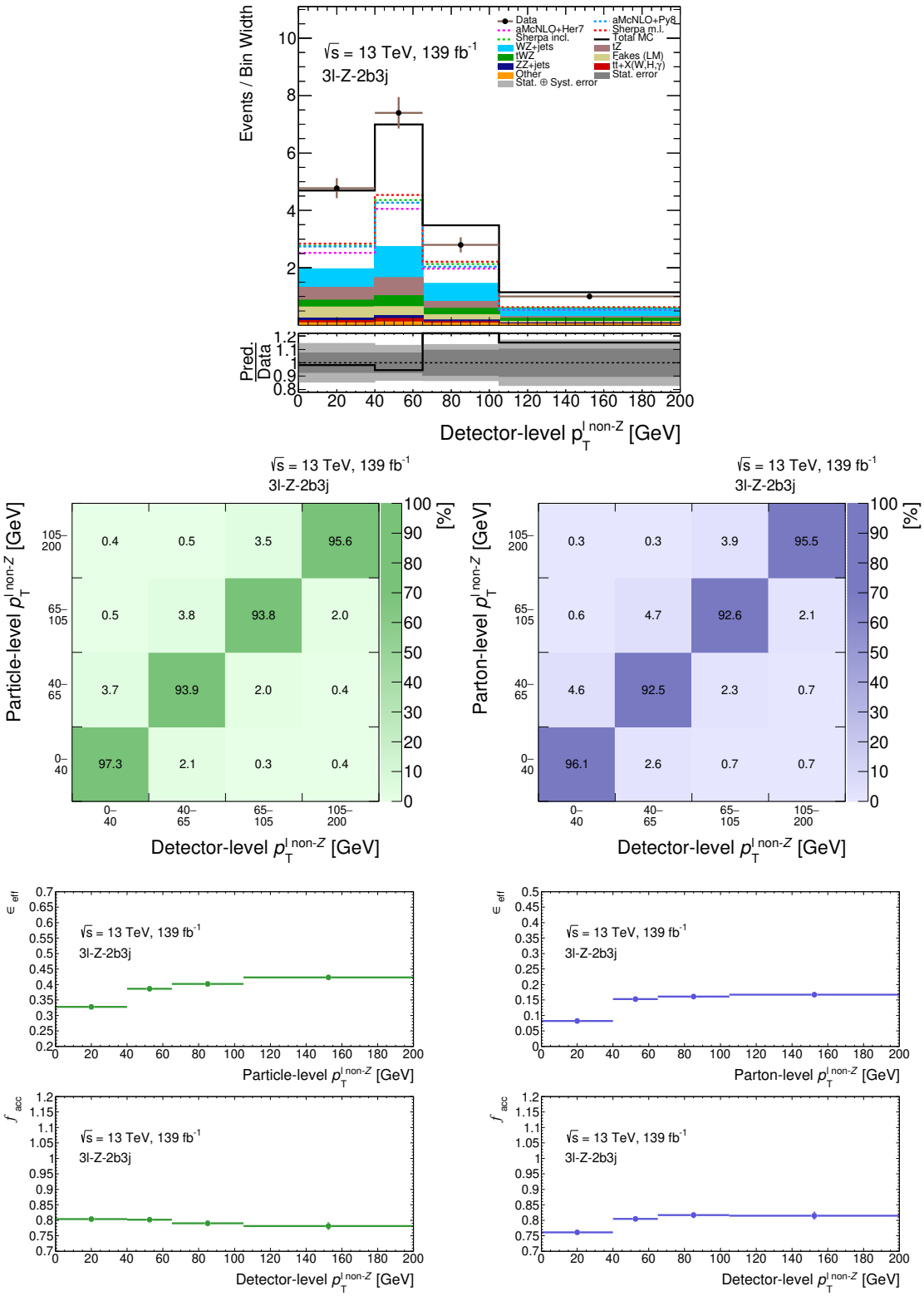


Fig. 11: Reconstruction-level distribution (top), together with particle-level (left) and parton-level (right) migration matrices (middle) and efficiency/acceptance corrections (bottom) for $p_{\mathrm{T}}^{\ell\mathrm{non}-Z}$ in the trilepton channel.

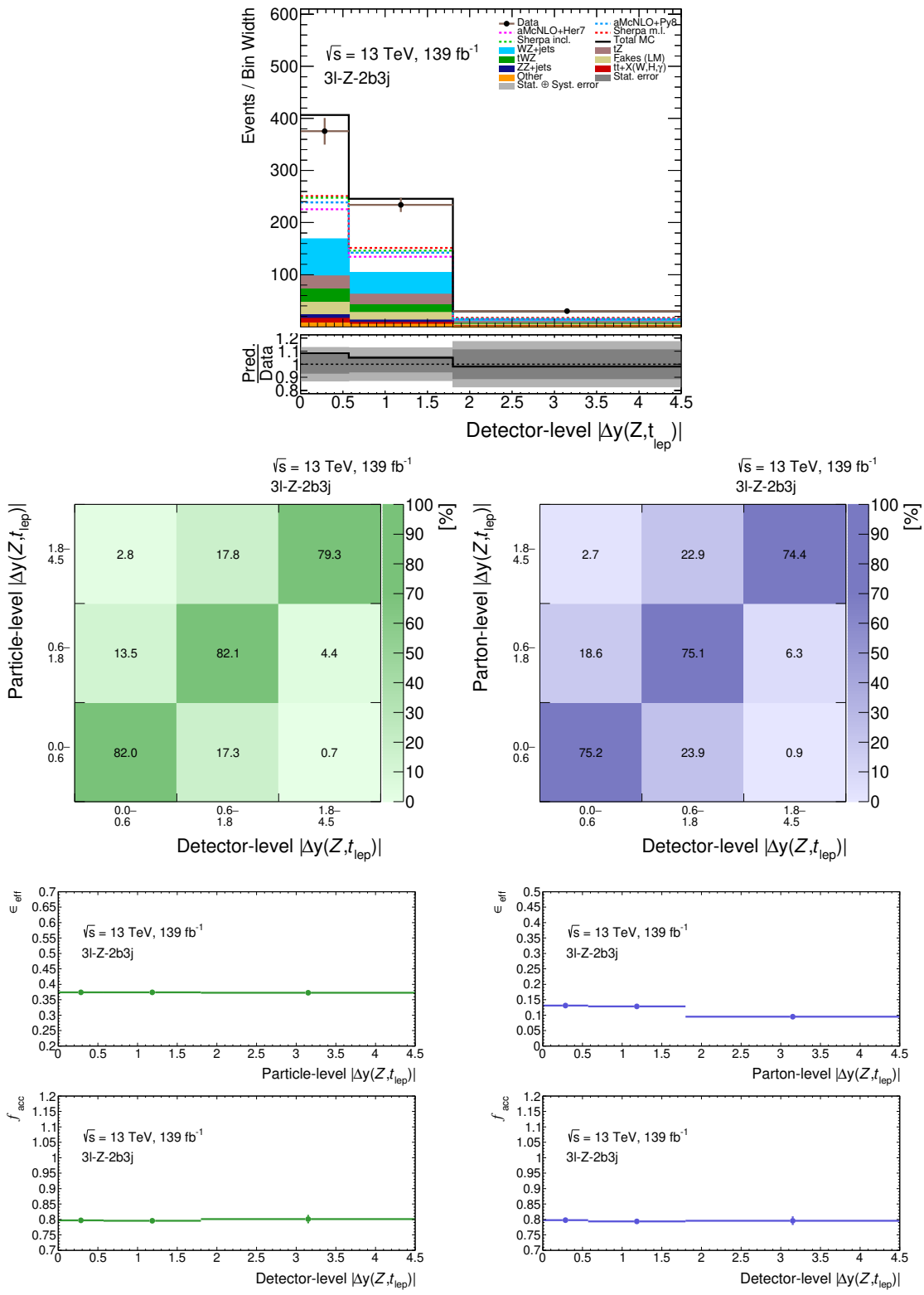


Fig. 12: Reconstruction-level distribution (top), together with particle-level (left) and parton-level (right) migration matrices (middle) and efficiency/acceptance corrections (bottom) for $|\Delta y(Z, t_{\text{lep}})|$ in the trilepton channel.

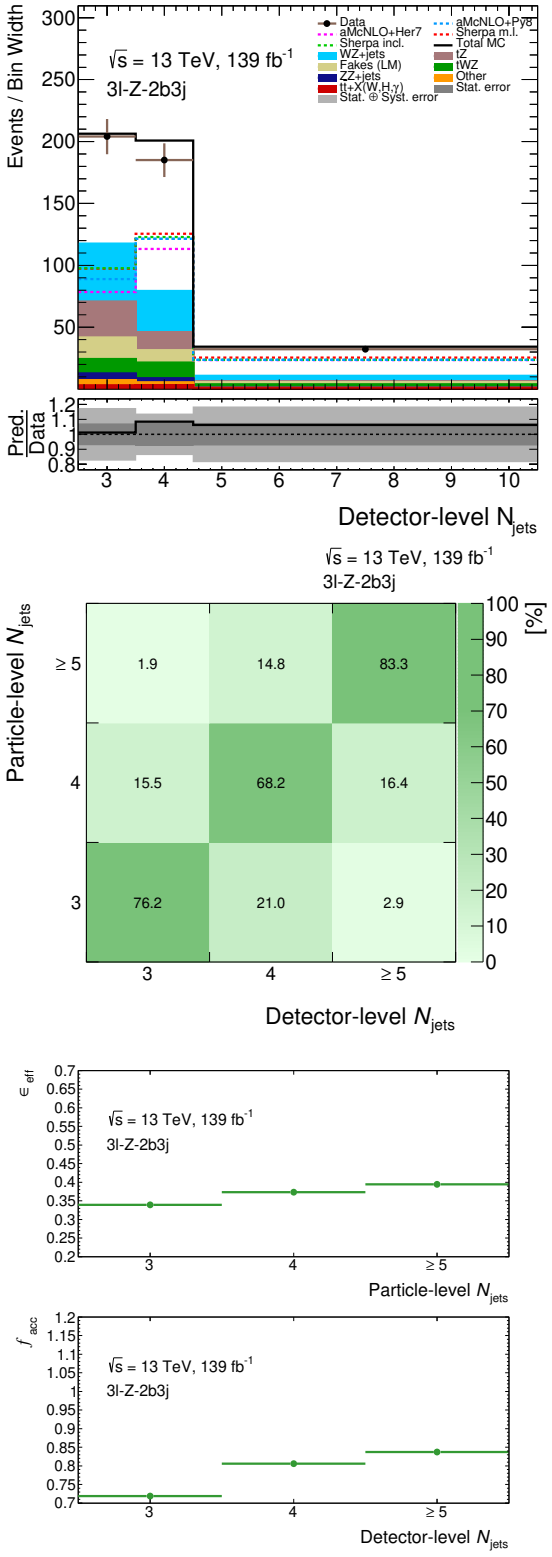


Fig. 13: Reconstruction-level distribution (top), together with particle-level migration matrix (middle) and efficiency/acceptance corrections (bottom) for N_{jets} in the trilepton channel.

A.1.2 Tetralepton channel

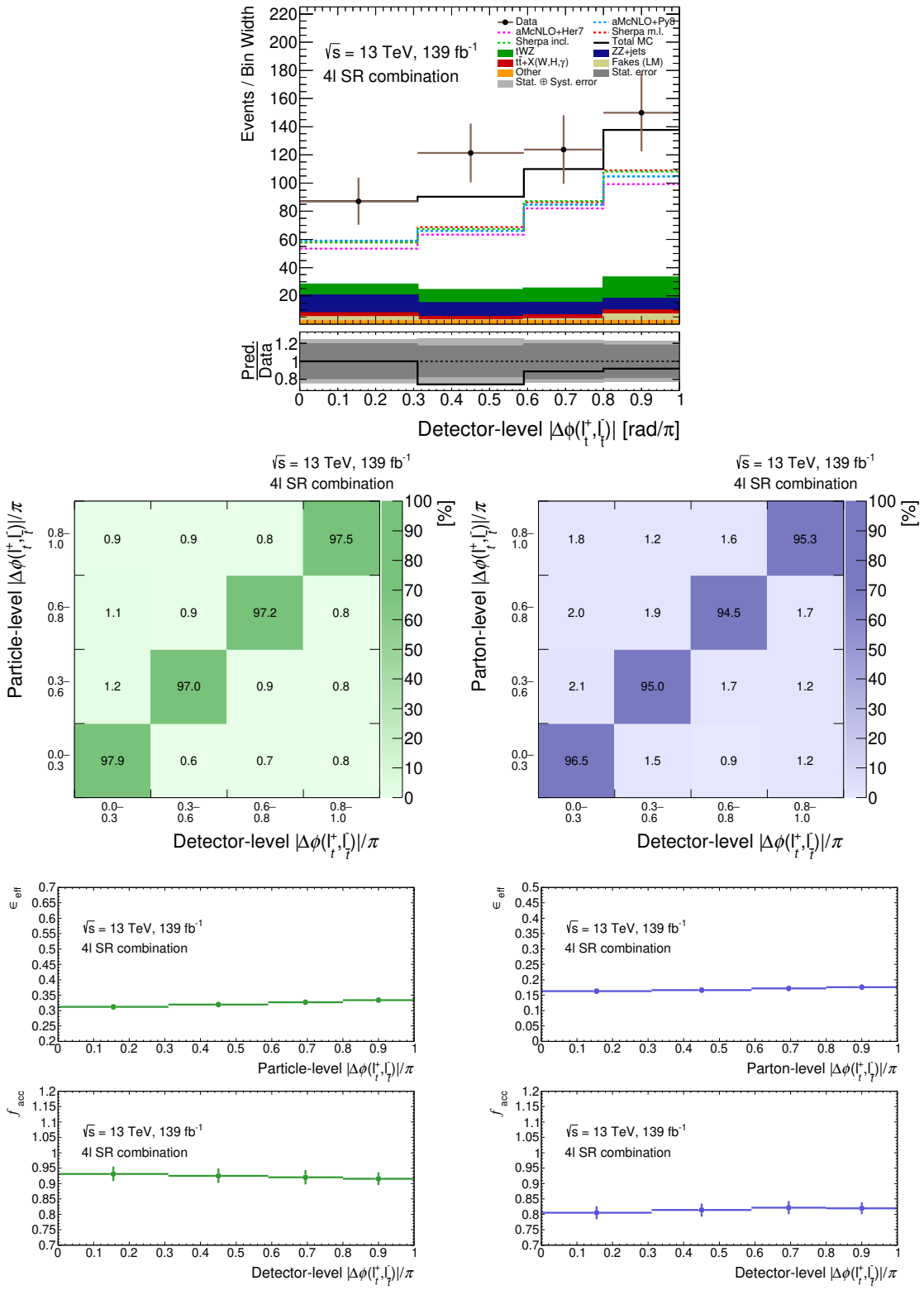


Fig. 14: Reconstruction-level distribution (top), together with particle-level (left) and parton-level (right) migration matrices (middle) and efficiency/acceptance corrections (bottom) for $|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)|$ in the tetralepton channel.

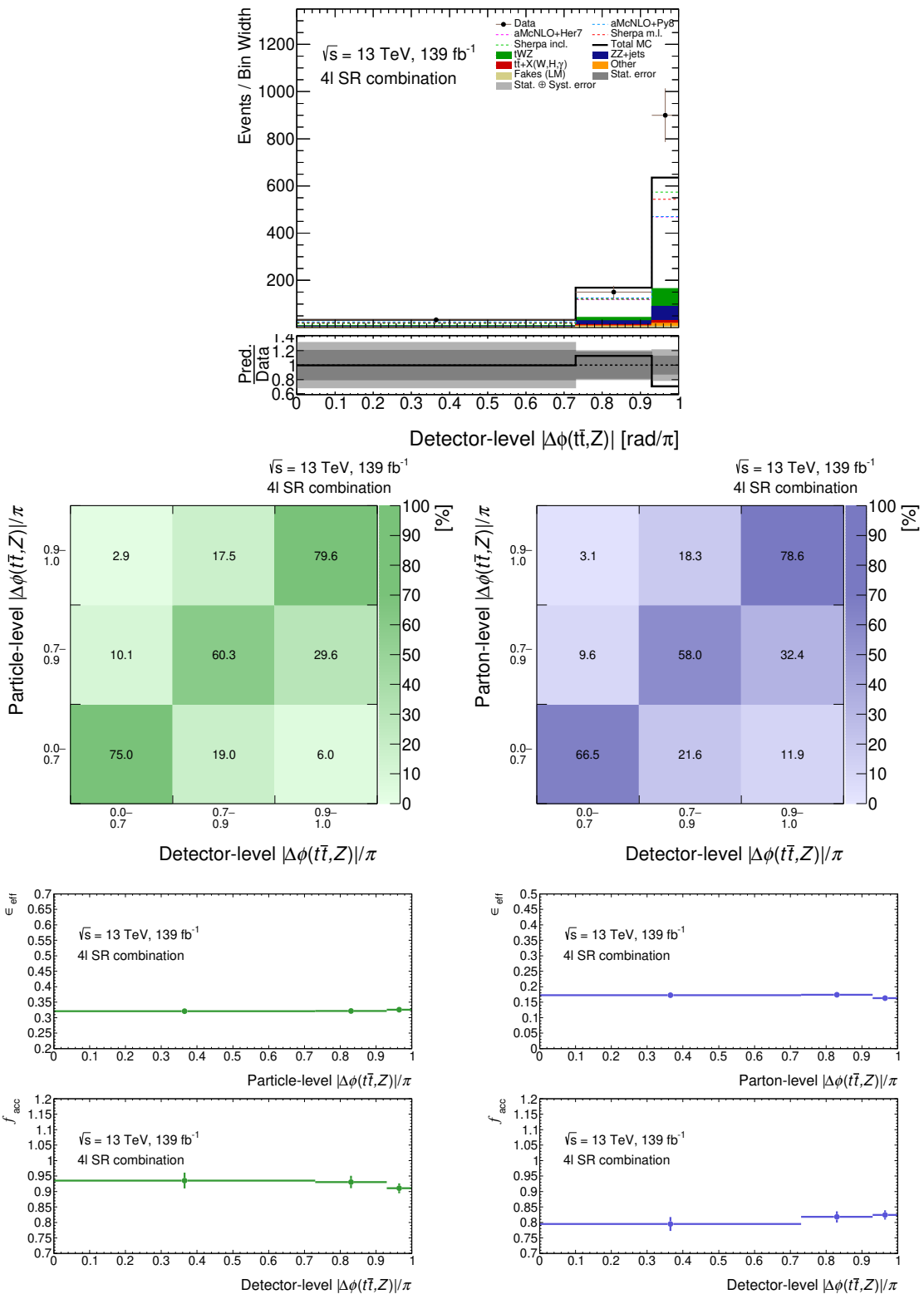


Fig. 15: Reconstruction-level distribution (top), together with particle-level (left) and parton-level (right) migration matrices (middle) and efficiency/acceptance corrections (bottom) for $|\Delta\phi(t\bar{t}, Z)|$ in the tetralepton channel.

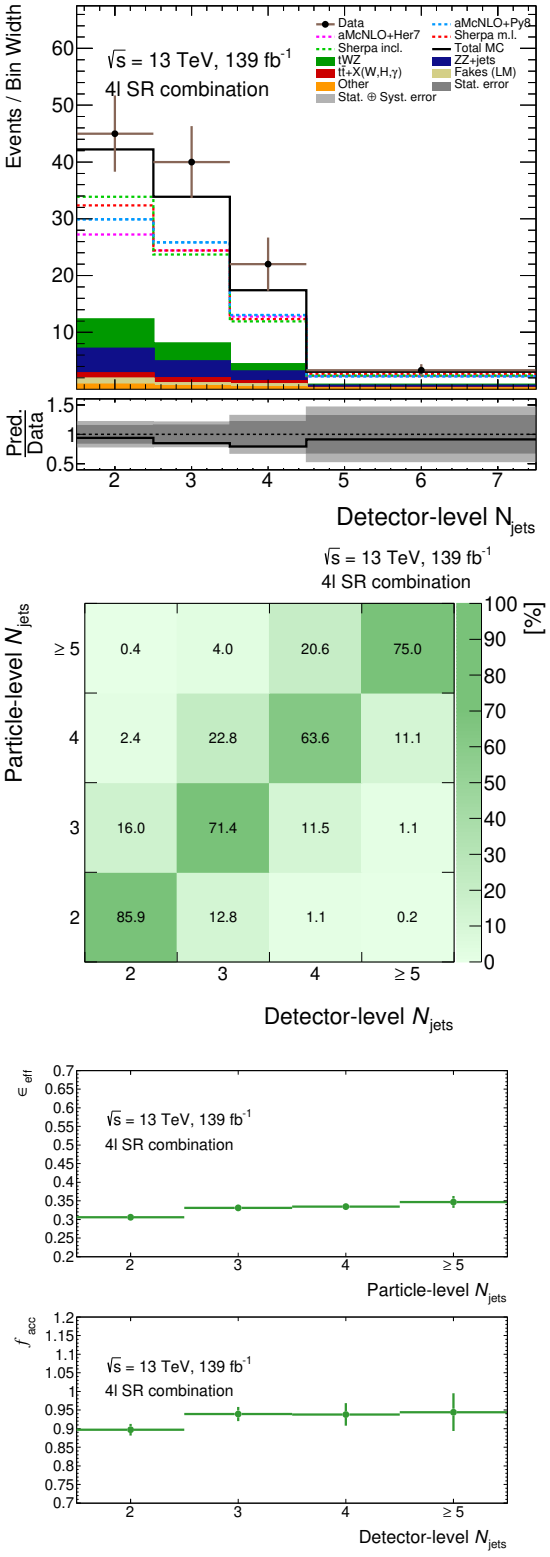


Fig. 16: Reconstruction-level distribution (top), together with particle-level migration matrix (middle) and efficiency/acceptance corrections (bottom) for N_{jets} in the tetralepton channel.

A.1.3 Combined $3\ell + 4\ell$ channel

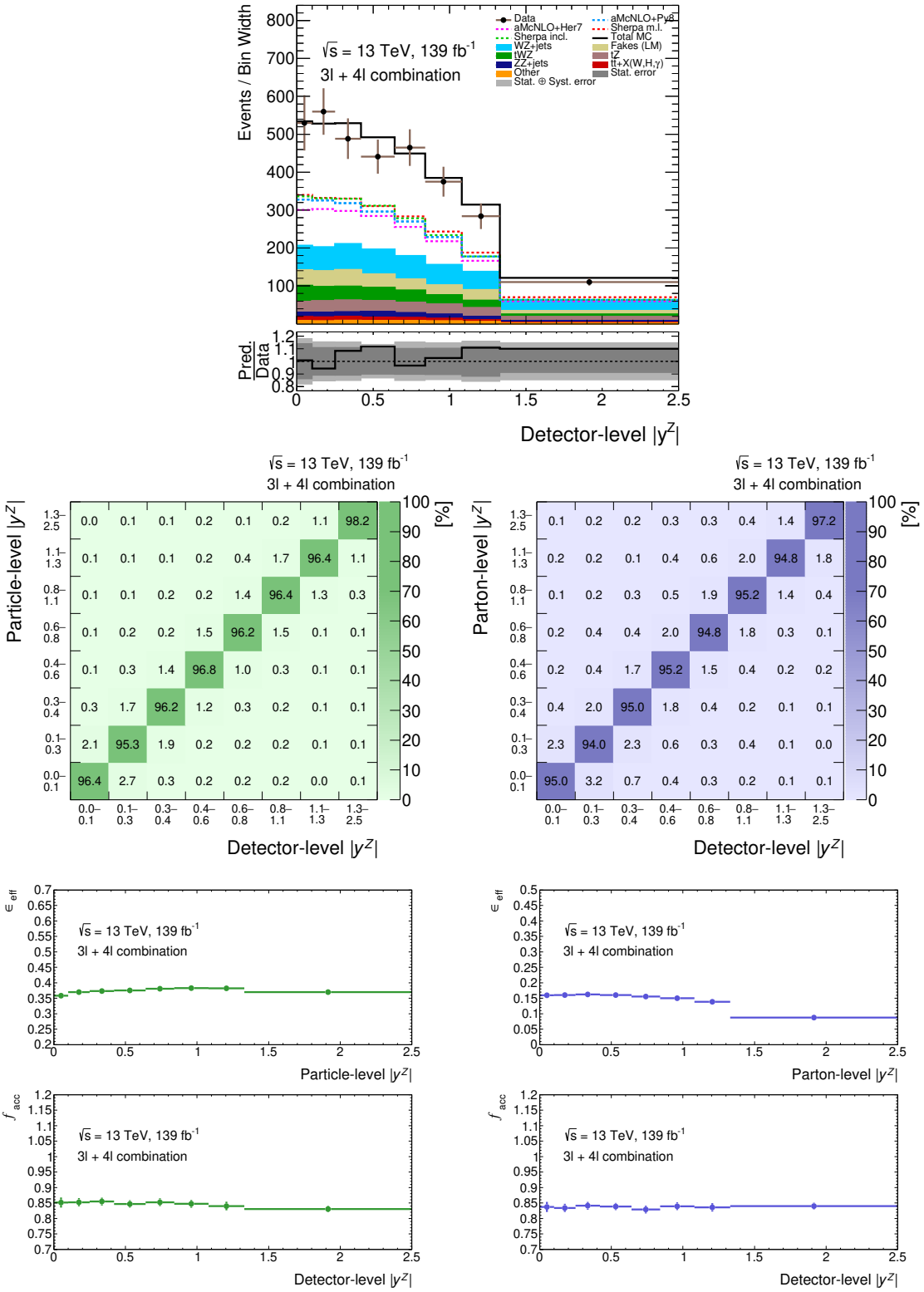


Fig. 17: Reconstruction-level distribution (top), together with particle-level (left) and parton-level (right) migration matrices (middle) and efficiency/acceptance corrections (bottom) for $|y^Z|$ in the combined $3\ell + 4\ell$ channel.

A.2 Optimization of the number of iterations

A.2.1 Trilepton channel

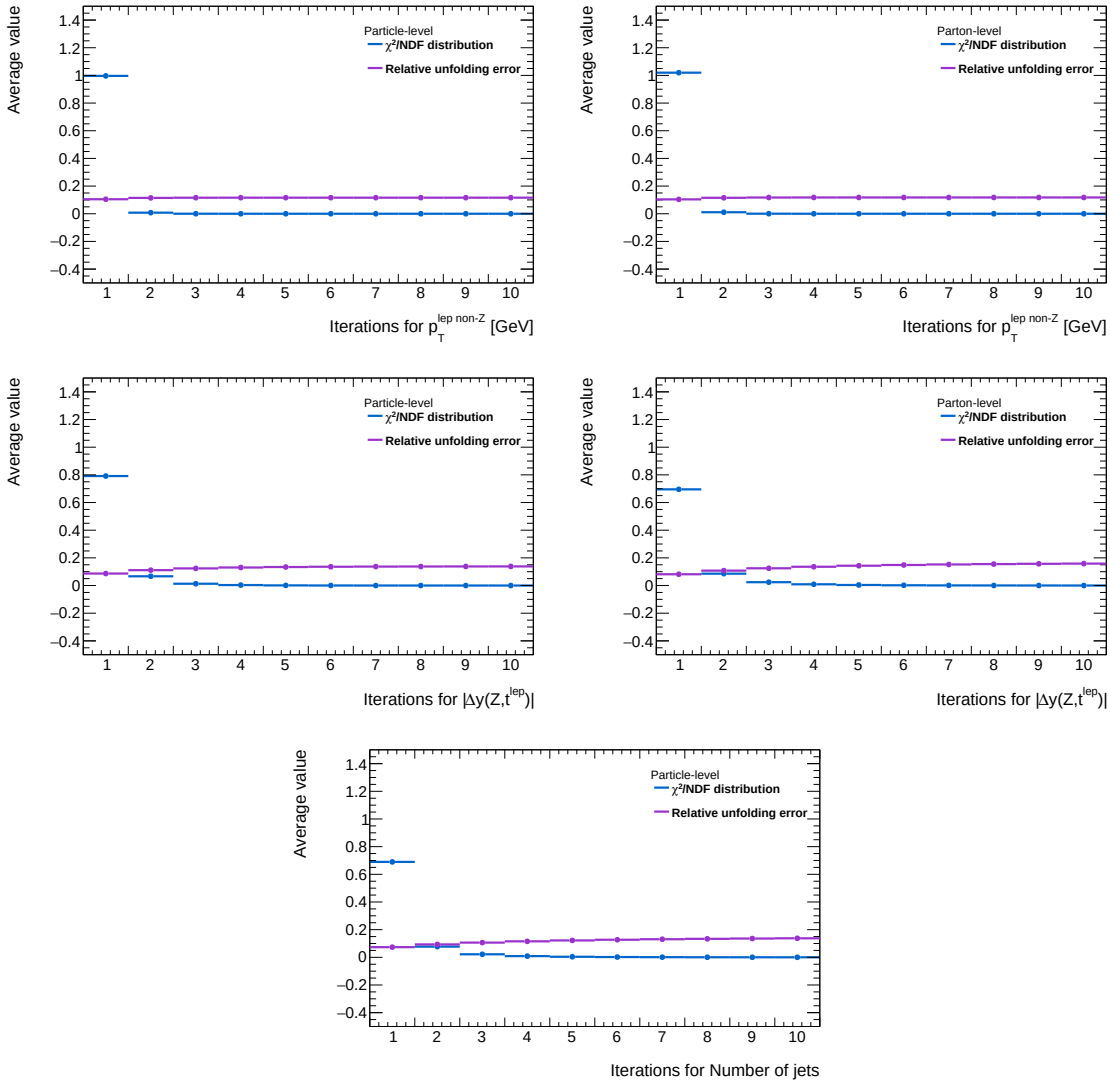


Fig. 18: Particle-level (left) and parton-level (right) summary plots for the optimization of the number of iterations for $p_{\mathrm{T}}^{\ell\mathrm{non}-Z}$ (first row), $|\Delta y(Z, t_{\mathrm{lep}})|$ (second row) and N_{jets} (third row) in the trilepton channel.

A.2.2 Tetralepton channel

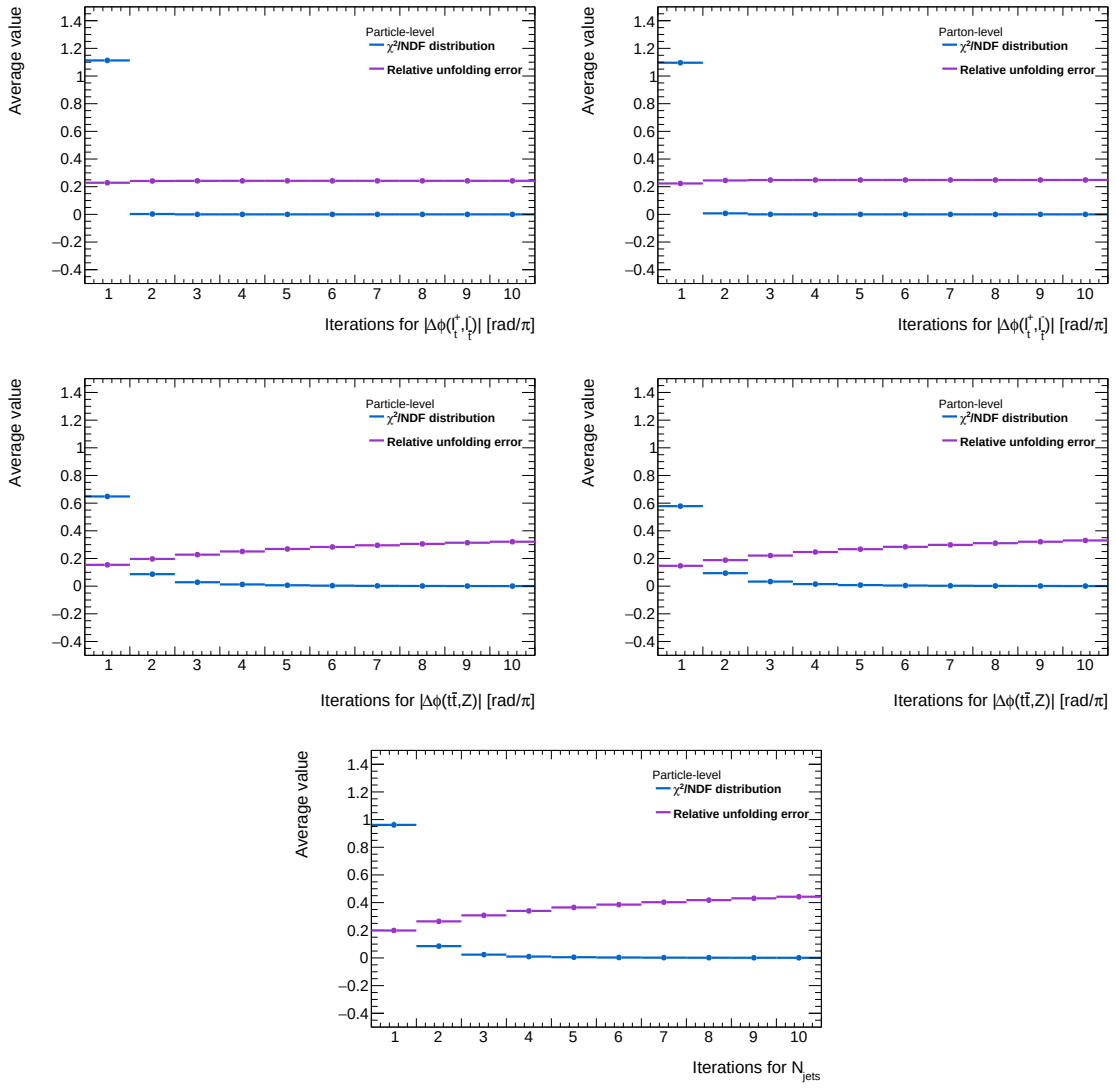


Fig. 19: Particle-level (left) and parton-level (right) summary plots for the optimization of the number of iterations for $|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)|$ (first row), $|\Delta\phi(t\bar{t}, Z)|$ (second row) and N_{jets} (third row) in the tetralepton channel.

A.2.3 Combined $3\ell + 4\ell$ channel

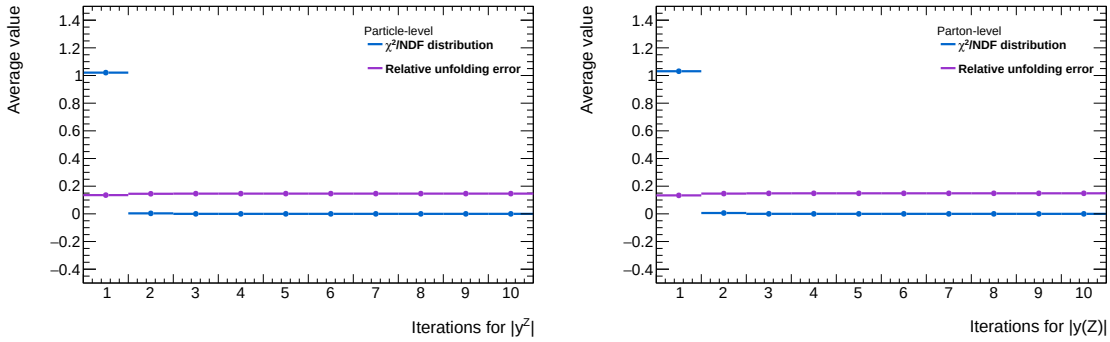


Fig. 20: Particle-level (left) and parton-level (right) summary plots for the optimization of the number of iterations for $|y^Z|$ in the combined $3\ell + 4\ell$ channel.

A.3 Closure tests

This section presents details about the original version of the closure test, as it is usually performed in the analyses featuring high number of events. For the purpose of the differential $t\bar{t}Z$ analysis, the procedure is modified to take into account limited statistics observed in both trilepton and tetralepton signal regions.

Second part of this section presents results of modified closure tests for additional variables considered for the differential cross section measurements in the trilepton, tetralepton and combined channels.

A.3.1 Original closure test and its modifications

The original strategy of performing closure tests follows the procedure presented in the main body of the thesis in Section 10.2.3. The only difference lies in the definition of the pull and extraction of the corresponding mean and standard deviation from the pull distribution.

Typically the pull is defined as follows:

$$p_i^j = \frac{x_i^j - t_i}{\sigma(x_i^j)}, \tag{6}$$

where x_i^j corresponds to unfolded value for pseudo-experiment j in bin i, $\sigma(x_i^j)$ is its associated uncertainty and t_i represents the truth-level value. The sample pull distributions for the first two bins of the $|\Delta\phi(t\bar{t}, Z)|$ variable in the tetralepton channel can be found in Figure 21. Conventionally a fit is performed to such distributions based on a Gaussian function which is then superimposed on the same plot. The corresponding mean and standard deviation of the Gaussian fits are quoted for each bin of the variable, as shown in Figure 22. The unfolding procedure is considered to be stable if the pull means are consistent with zero and standard deviation with one. However, as can be seen from the plots shown in Figure 21, the pull distributions does not agree with the fitted Gaussian, because left tail of the spectrum is populated more than the right tail. This effect, which results in the failed pull tests for some variables with shifted pull means and standard deviations (widths), was found to be caused by the limited statistics in the analysis regions.

To understand the reason for non-Gaussian shape of the pulls in case of limited statistics, the following toy MC example is described. In case of idealized unfolding problem, where the acceptance and efficiency corrections are not needed (they are equal to one) and the migration matrix is perfectly diagonal, the unfolding process is trivial. The resulting unfolded spectrum should by definition agree with the truth-level spectrum. When now Poisson smearing is applied to the reconstruction-level distribution, the smearing will be transferred also to the unfolded distribution. Thus when repeating

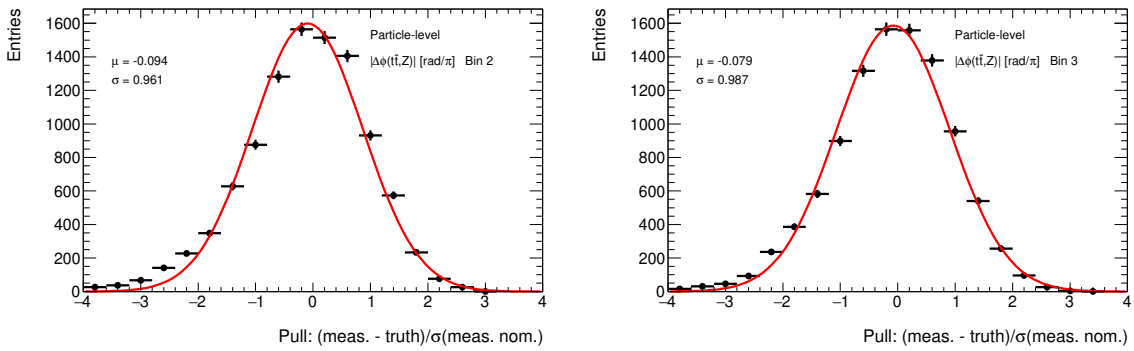


Fig. 21: Sample pull distributions for the second (left) and third (right) bin of $|\Delta\phi(t\bar{t}, Z)|$ for particle-level in the tetralepton channel obtained using original pull definition from Eq.6. The solid red line represents the Gaussian fit.

the smearing for many pseudo-experiments, both the reconstruction-level and unfolded spectrum should resemble the shape of the smearing function - in this case Poisson. Therefore, in case of large number of events in particular bin, a Poisson distribution should be reasonably well approximated by Gaussian distribution, according to the central limit theorem. Unfortunately, this assumption does not hold in case of tetralepton channel, where the average number of events in the differential bin is around 17. The pull distribution, as defined in Eq. 6, then consists of the subtraction of two Poisson distributions what in the end does not change the resulting shape of the pull. However, the denominator of a pull, unfolding error for the particular pseudo-experiment, can change the originally Poisson (Gaussian) shape. In this idealized model, the unfolding uncertainty is simply equal to $\sqrt{x}$ if x is unfolded value. Higher unfolded value then causes higher uncertainty what in consequence leads to the overpopulation on the left tail of the pull. The illustrative plots depicting such pull distributions for the Poisson mean parameter λ=15 and 1500 is shown in Figure 23, where the overpopulation for lower λ is clearly visible. However, this effect is mitigated for highly populated bins as the asymmetry caused by difference in $\sqrt{x}$ for the same positive and negative value of $(x_i^j - t_i)$ is suppressed (see right plot in Figure 23) with increasing t_i (and consequently also x_i^j).

To solve the problem of non-Gaussian shape for the bins with low statistics, while still preserving the purpose of the pull test, the uncertainty on the unfolded pseudo-experiment $\sigma(x_i^j)$ in the denominator of Eq. 6 was replaced by the constant (over pseudo-experiments) uncertainty on the unfolded un-smeared nominal distribution $\sigma(x_i^{\text{nominal}})$. The new definition of the pull thus becomes:

$$p_i^j = \frac{x_i^j - t_i}{\sigma(x_i^{\text{nominal}})} \tag{7}$$

An analogous distributions to that shown in Figure 23, but with the modified pull

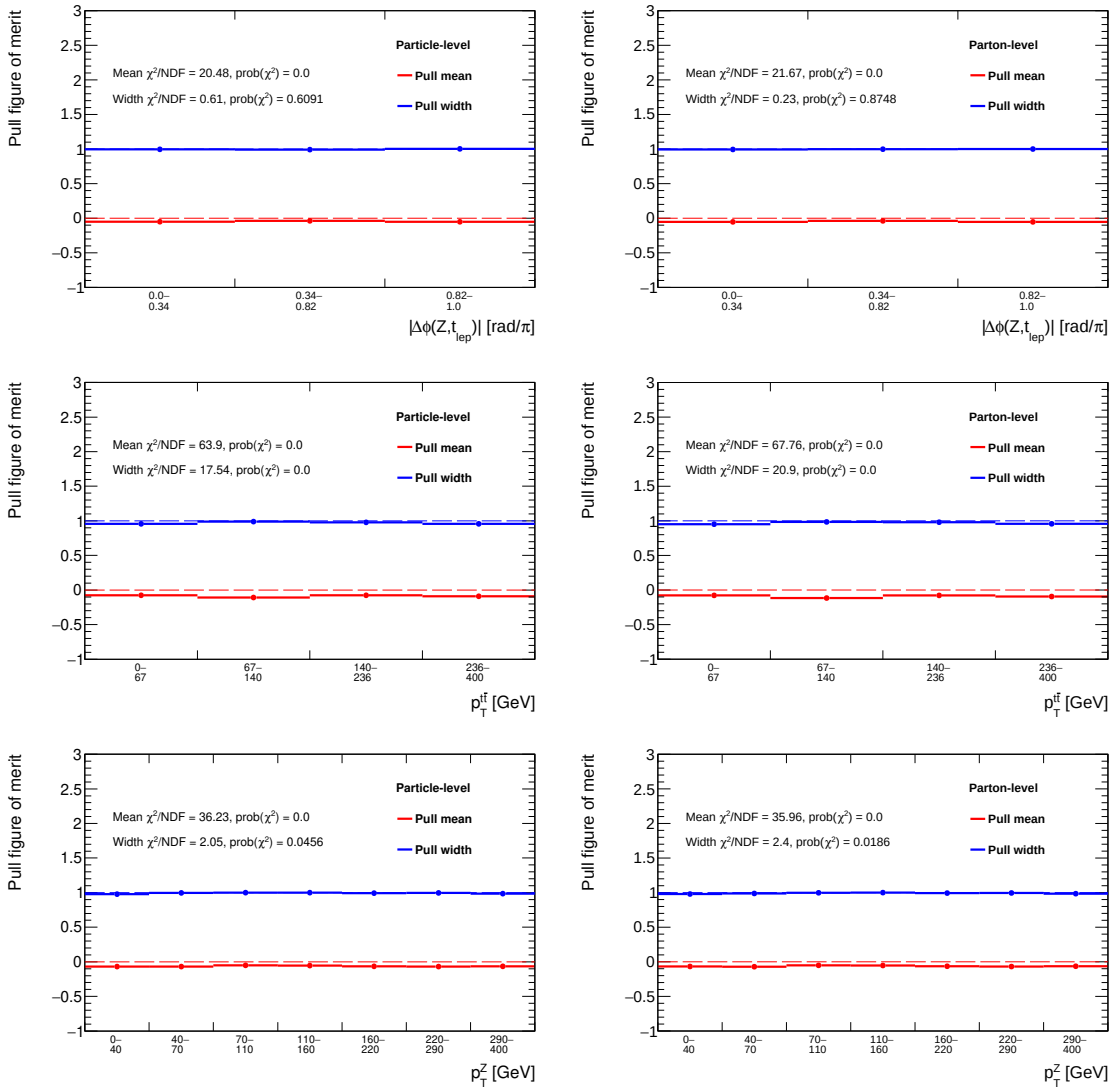


Fig. 22: Particle-level (left) and parton-level (right) summary plots for the pull tests performed for $|\Delta\phi(Z, t_{\text{lep}})|$ in the trilepton (first row), $p_{\text{T}}^{t\bar{t}}$ in the tetralepton (second row) and p_{T}^{Z} in the combined $3\ell + 4\ell$ channel (third row). Corresponding plots for the rest of differential variables can be found in following sections.

definition is shown in Figure 24.

As was already mentioned, a Gaussian fit is not suitable for fitting the pull distribution. One reason is that distribution would, in case of trivial unfolding, be expected to follow a Poisson distribution (cannot be approximated with a Gaussian distribution for low number of events). Moreover, a Poisson distribution is discrete, and consequently unfolded distribution is also discrete when considering trivial unfolding. For this reason, the fit would depend on the choice of binning used for the pull distribution. Additionally, when going from the idealized unfolding in case of the toy model presented above, to real unfolding performed on the MC samples, the unfolding corrections are no more equal to one and the migration matrix contains also off-diagonal

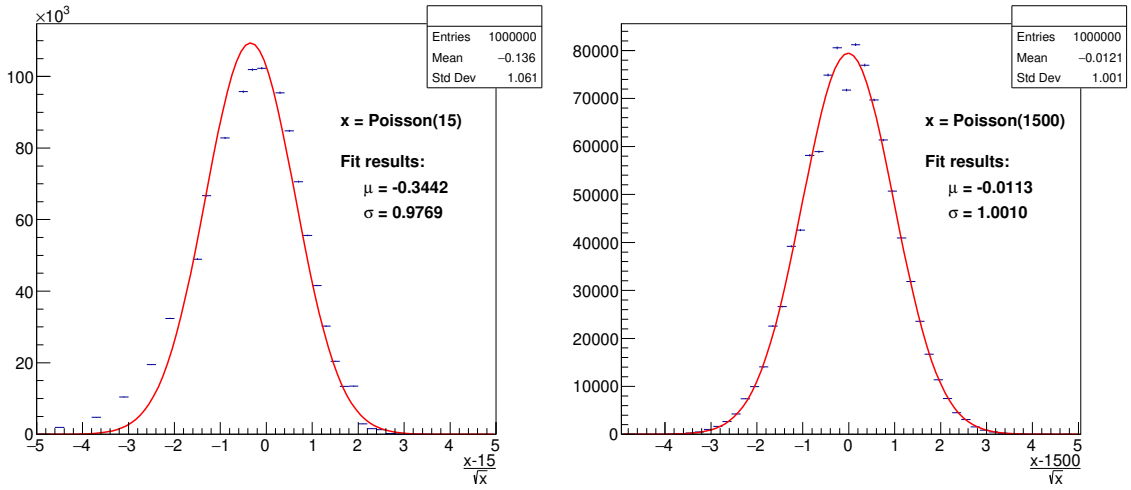


Fig. 23: Simulated toy model pull distributions (based on the definition in Eq. 6) for a Poisson distributed random variable x with $\lambda = 15$ (left) and $\lambda = 1500$ (right). The solid red line corresponds to the fitted Gaussian function.

elements. This brings additional smearing to the unfolded spectrum, what results in smeared discrete values of originally Poisson pull distributions. The larger are the off-diagonal elements of migration matrix the more pronounced is the smearing. This can be demonstrated at Figure 25 by comparing pull distributions using finer binning for the $|y^Z|$ with the highly-diagonal migration matrix, and N_{jets} in the tetralepton channel featuring significantly larger off-diagonal elements.

In order to eliminate above-mentioned effects of pull binning dependence due to the discreetness of Poisson distribution and the smearing effect of the real unfolding, the Gaussian fit is replaced by the calculation of arithmetic mean μ and its corresponding root mean square σ according to the following standard formulas:

$$\begin{aligned} \mu &= \tfrac{1}{N}\textstyle\sum_{i=1}^{N} x_i, \\ \sigma &= \sqrt{\tfrac{1}{N}\textstyle\sum_{i=1}^{N}(x_i-\mu)^2}, \end{aligned} \tag{8}$$

where x_i is the pull for pseudo-experiment i and sum goes over all pseudo-experiments. The errors on these estimators are given by:

$$\begin{aligned} \delta\mu &= \tfrac{\sigma}{\sqrt{N}}, \\ \delta\sigma &= \tfrac{\sigma}{\sqrt{2(N-1)}}. \end{aligned} \tag{9}$$

Results of the pull tests for sample variables when using the modified definition of the pull (Eq. 7) and the extraction of pull mean and width values according to Eq. 8 can be found in the main body of the thesis in Section 10.2.3 for sample differential variables, and in the sections below for the rest of differential variables.

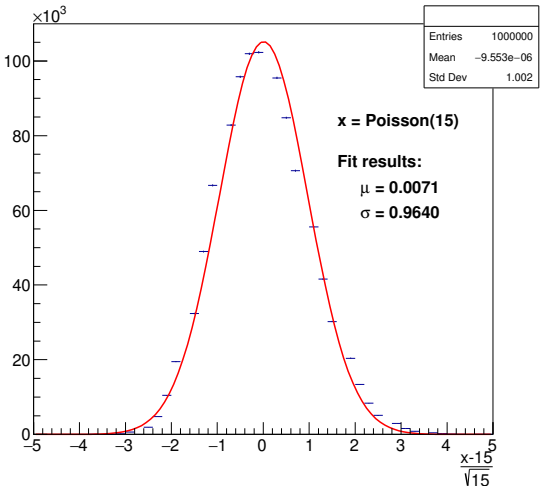


Fig. 24: Simulated toy model pull distributions (based on the modified definition from Eq. 7) for a Poisson distributed random variable x with $\lambda = 15$. The solid red line corresponds to the fitted Gaussian function.

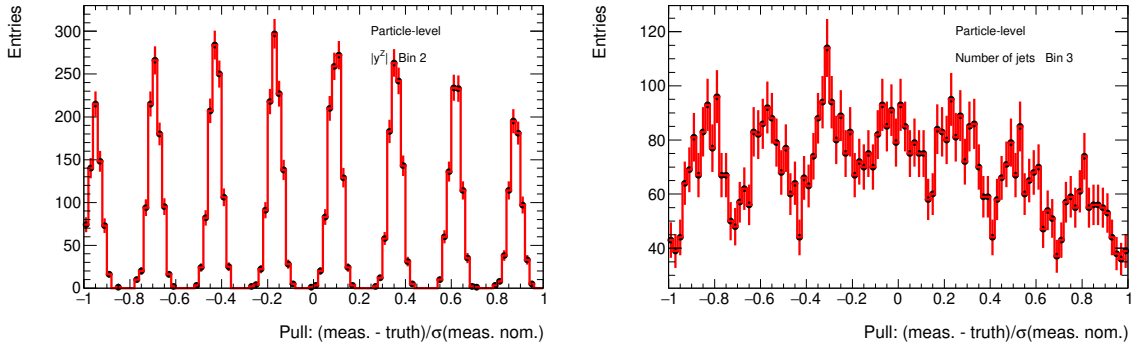


Fig. 25: Pull distributions for second bin of $|y_Z|$ (left) and third bin of N_{jets} (right) in the tetralepton channel, built from 10000 pseudo-experiments unfolded at particle level. Spikes in the distributions corresponds to the Poisson discrete values which are smeared by introducing off-diagonal elements in the migration matrix, and efficiency/acceptance corrections in the unfolding procedure.

A.3.2 Trilepton channel

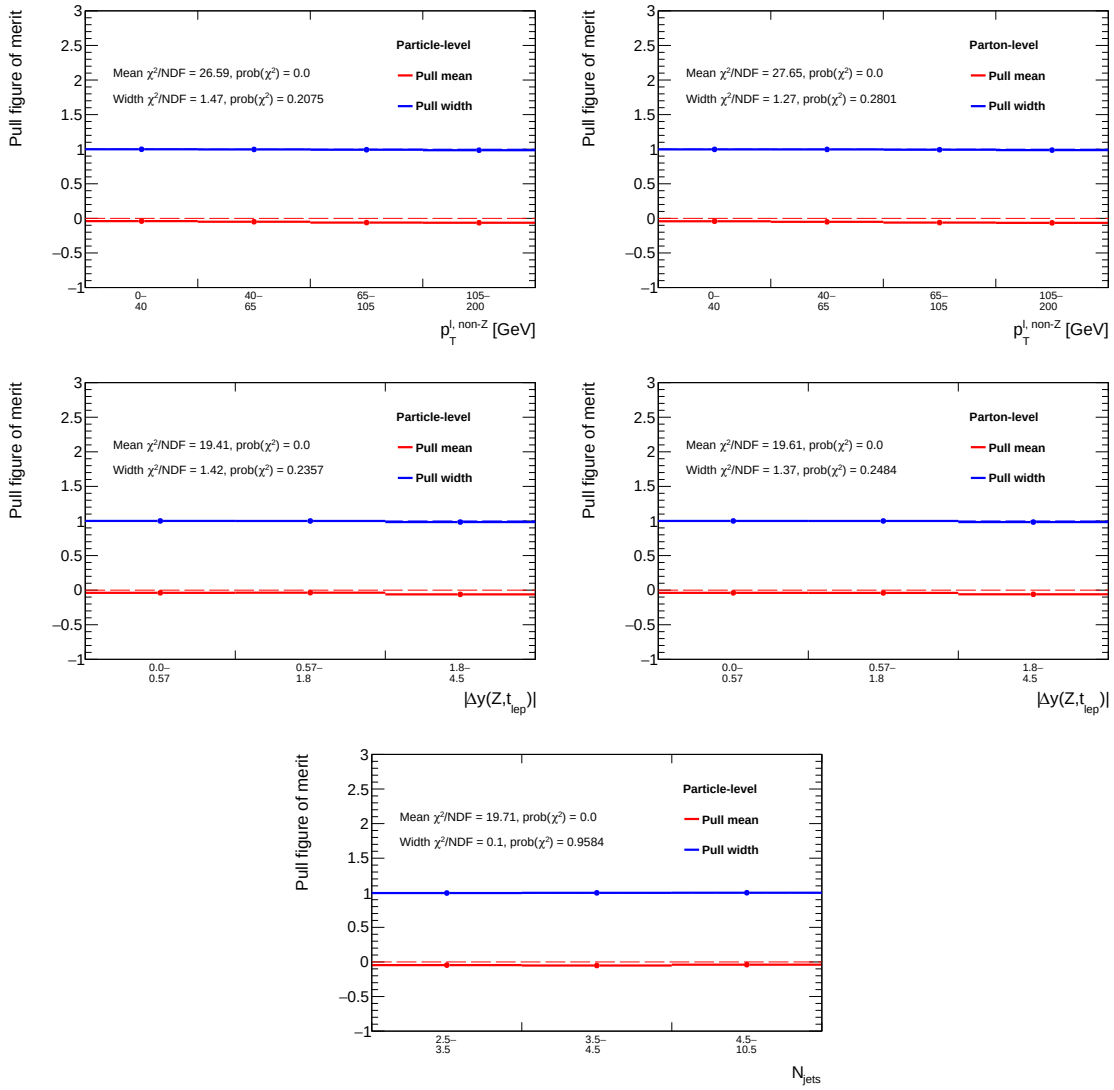


Fig. 26: Particle-level (left) and parton-level (right) summary plots for the pull tests (using original definition of pull from Eq. 6 and Gaussian fitting of the pull distributions) performed for $p_{\mathrm{T}}^{\ell\mathrm{non}-Z}$ (first row), $|\Delta y(Z, t_{\mathrm{lep}})|$ (second row) and N_{jets} (third row) in the trilepton channel.

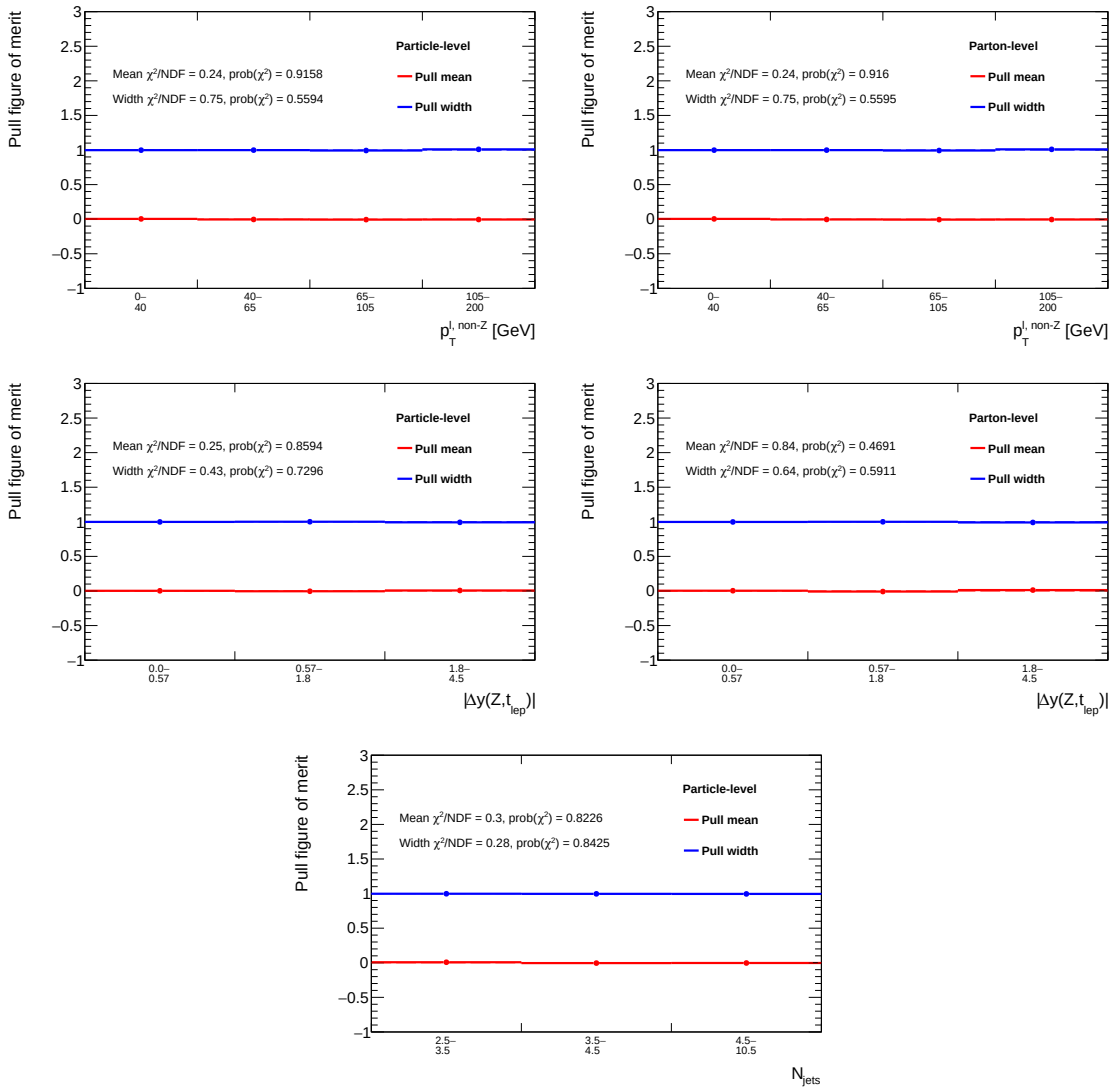


Fig. 27: Particle-level (left) and parton-level (right) summary plots for the updated pull tests (using modified pull definition from Eq. 10.6 and formulas from Eq. 10.7 for extraction of mean and width) performed for $p_{\mathrm{T}}^{\ell\mathrm{non}-Z}$ (first row), $|\Delta y(Z, t_{\mathrm{lep}})|$ (second row) and N_{jets} (third row) in the trilepton channel.

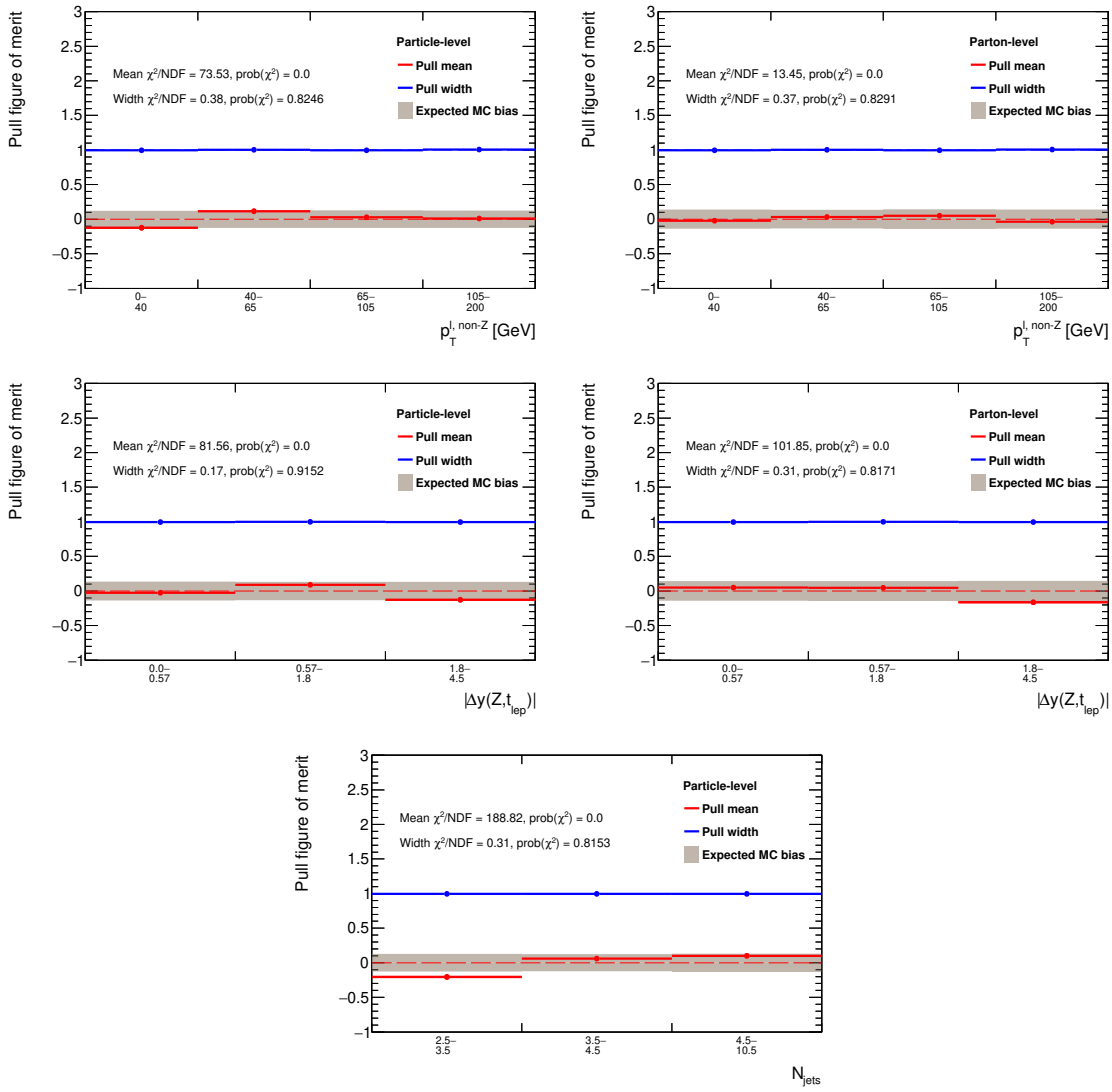


Fig. 28: Particle-level (left) and parton-level (right) summary plots for the pull tests using independent sample for building migration matrix and unfolding corrections, performed for $p_{\mathrm{T}}^{\ell \mathrm{non}-Z}$ (first row), $|\Delta y(Z, t_{\mathrm{lep}})|$ (second row) and N_{jets} (third row) in the trilepton channel.

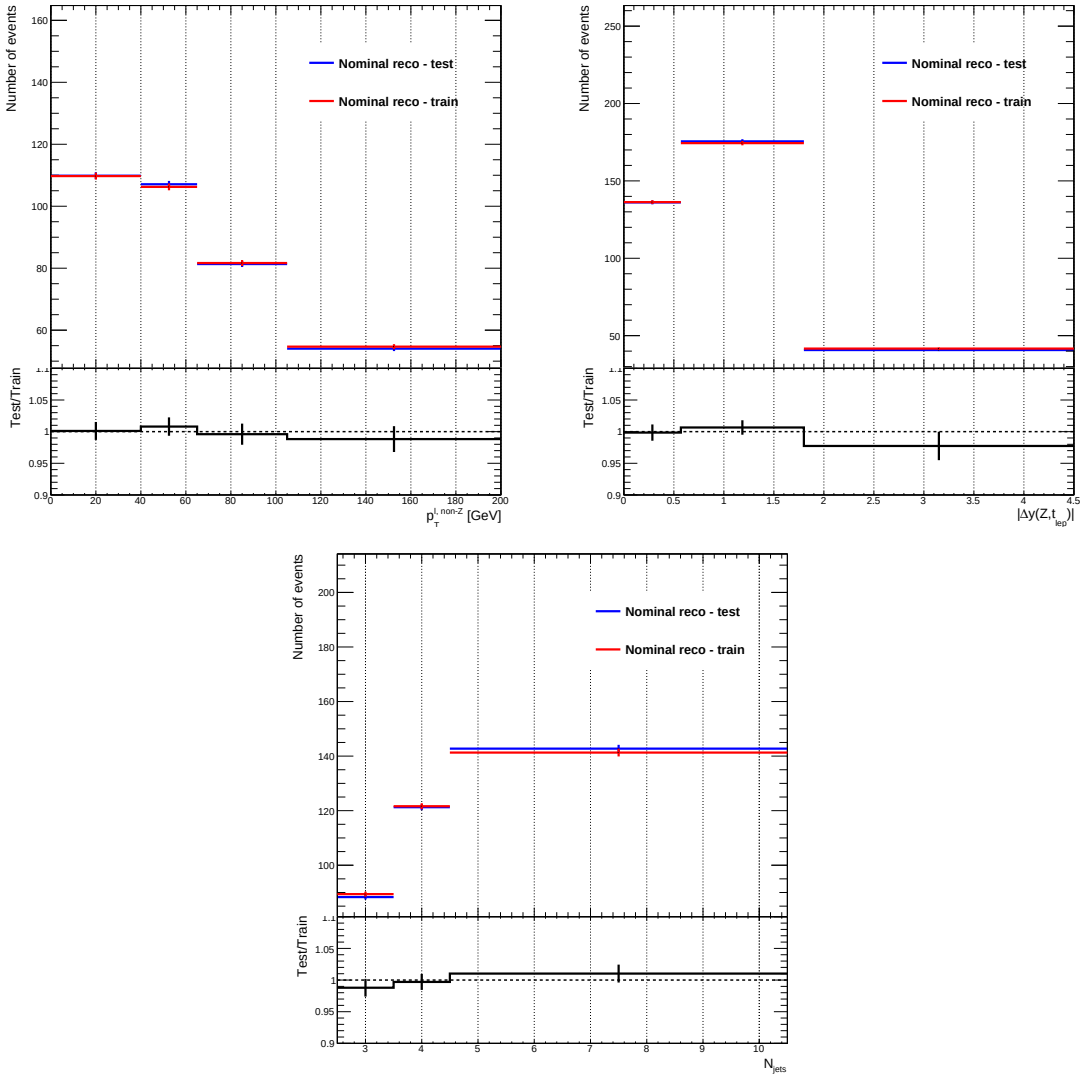


Fig. 29: Reconstruction-level distributions for the two independent samples used for the closure tests. Sample denoted as "train" is used for constructing migration matrix and evaluating unfolding corrections, while "test" sample is unfolded. The plots show distributions for $p_{\mathrm{T}}^{\ell\mathrm{non}-Z}$, $|\Delta y(Z, t_{\mathrm{lep}})|$ and N_{jets} in the trilepton channel.

A.3.3 Tetralepton channel

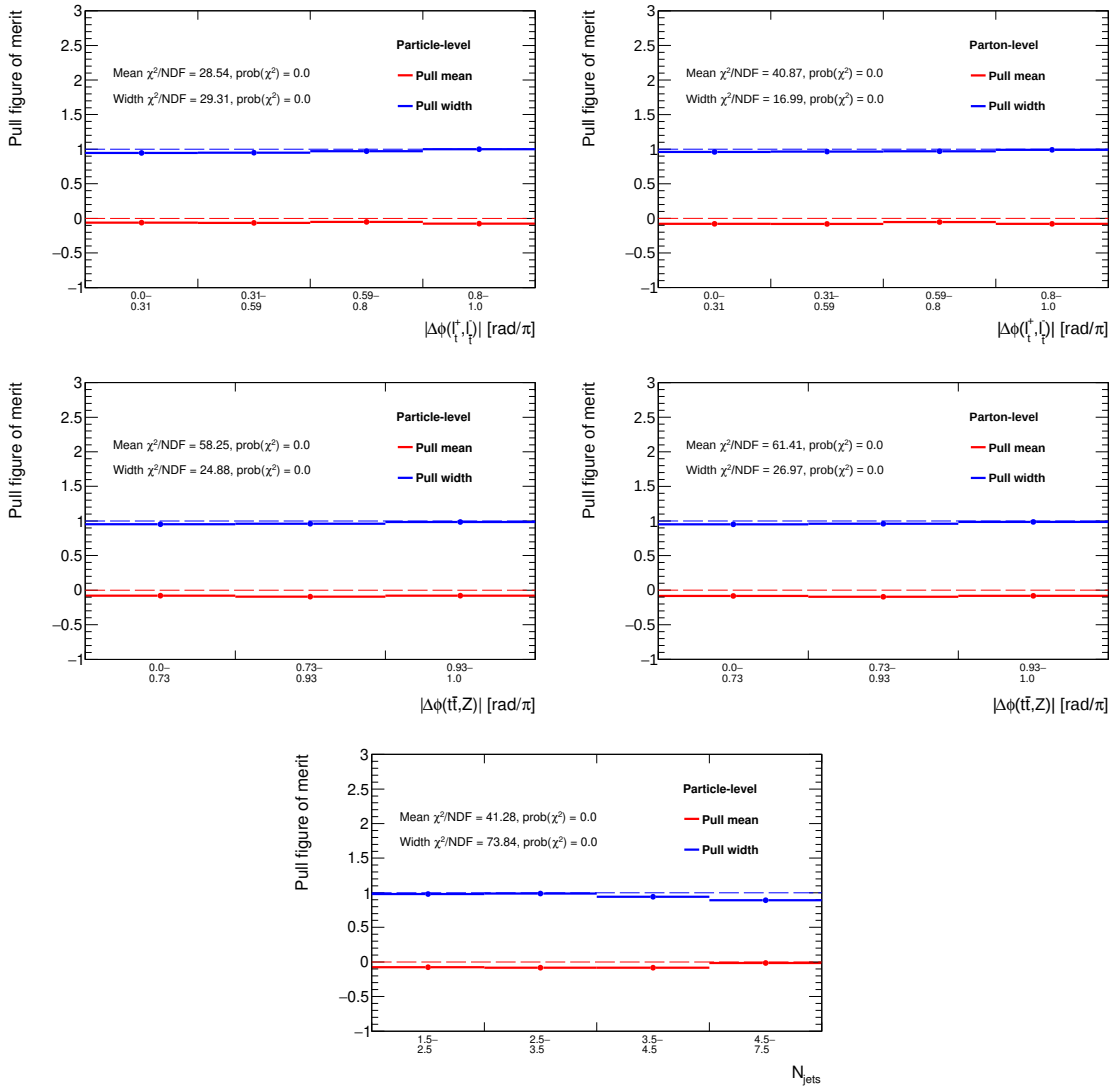


Fig. 30: Particle-level (left) and parton-level (right) summary plots for the pull tests (using original definition of pull from Eq. 6 and Gaussian fitting of the pull distributions) performed for $|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)|$ (first row), $|\Delta\phi(t\bar{t}, Z)|$ (second row) and N_{jets} (third row) in the tetralepton channel.

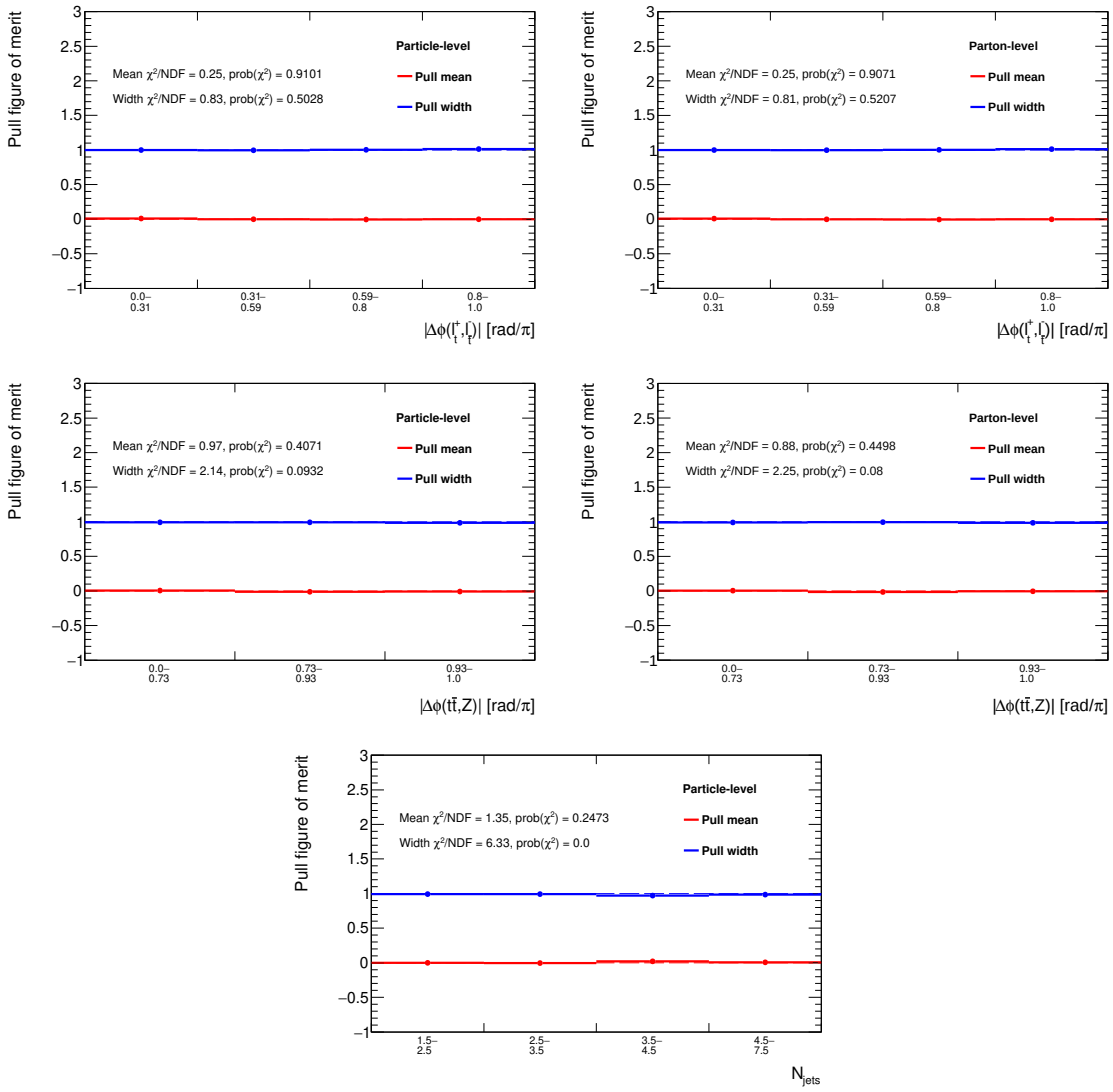


Fig. 31: Particle-level (left) and parton-level (right) summary plots for the updated pull tests (using modified pull definition from Eq. 10.6 and formulas from Eq. 10.7 for extraction of mean and width) performed for $|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)|$ (first row), $|\Delta\phi(t\bar{t}, Z)|$ (second row) and N_{jets} (third row) in the tetralepton channel.

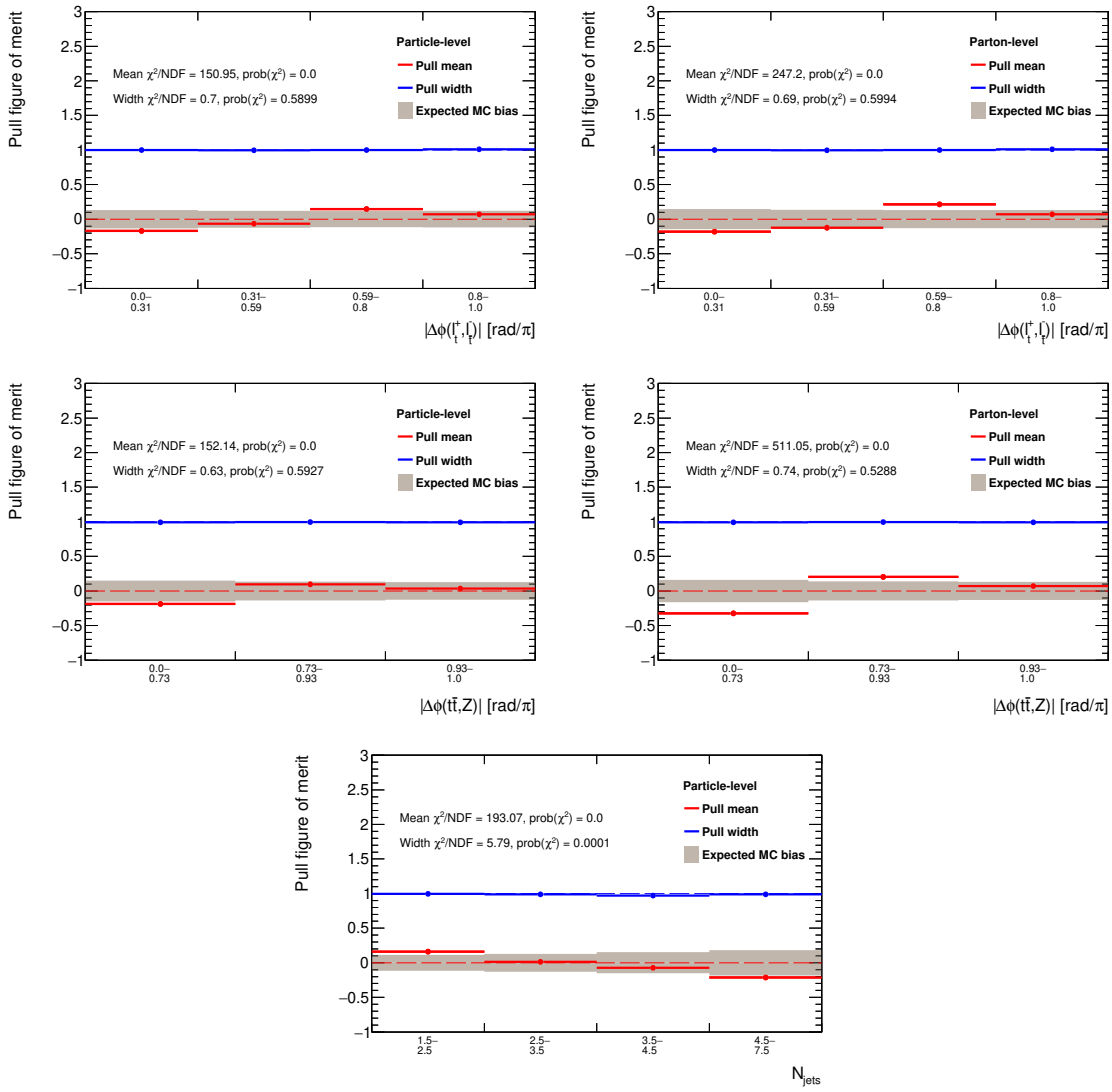


Fig. 32: Particle-level (left) and parton-level (right) summary plots for the pull tests using independent sample for building migration matrix and unfolding corrections, performed for $|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)|$ (first row), $|\Delta\phi(t\bar{t}, Z)|$ (second row) and N_{jets} (third row) in the tetralepton channel.

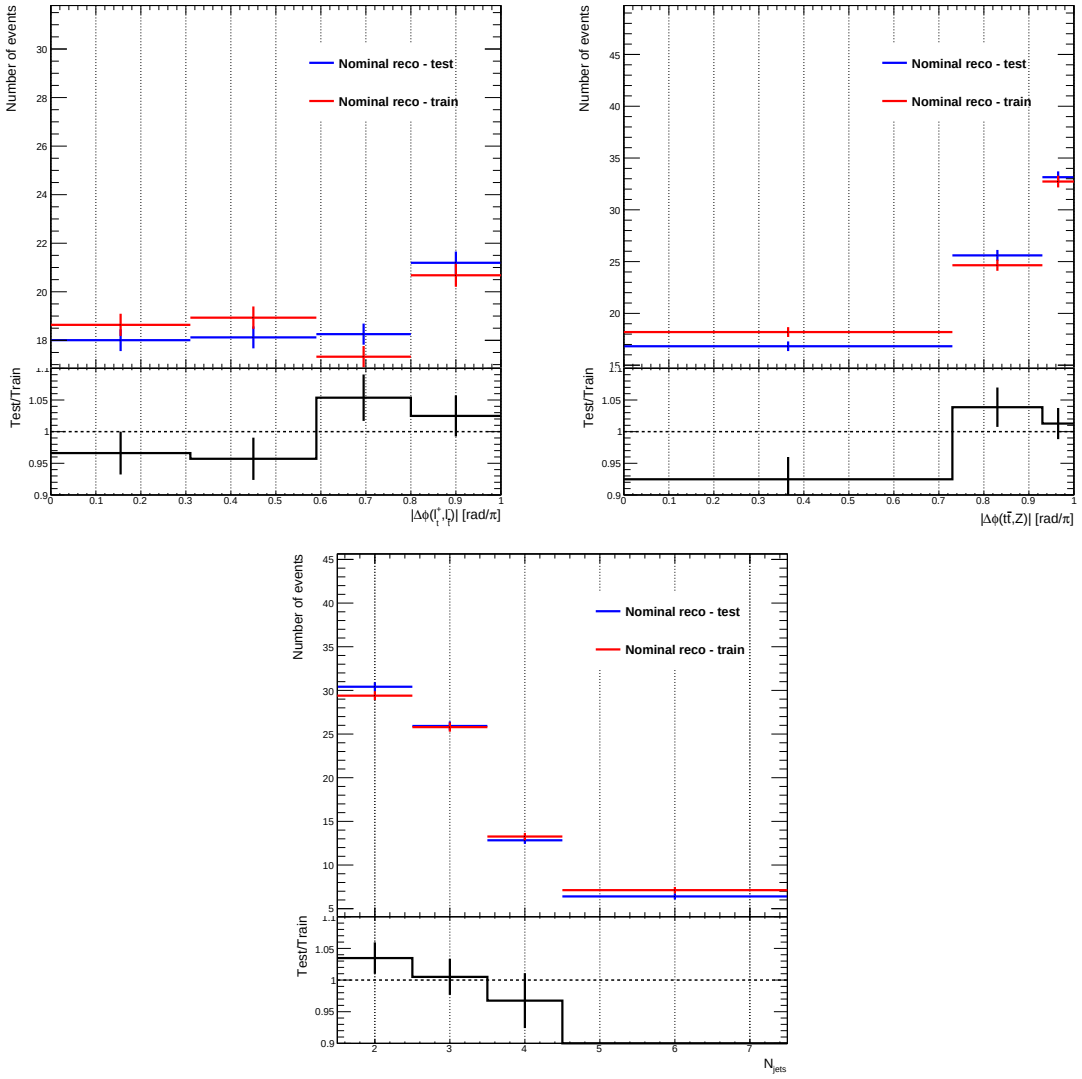


Fig. 33: Reconstruction-level distributions for the two independent samples used for the closure tests. Sample denoted as "train" is used for constructing migration matrix and evaluating unfolding corrections, while "test" sample is unfolded. The plots show distributions for $|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)|$, $|\Delta\phi(t\bar{t}, Z)|$ and N_{jets} in the tetralepton channel.

A.3.4 Combined $3\ell + 4\ell$ channel

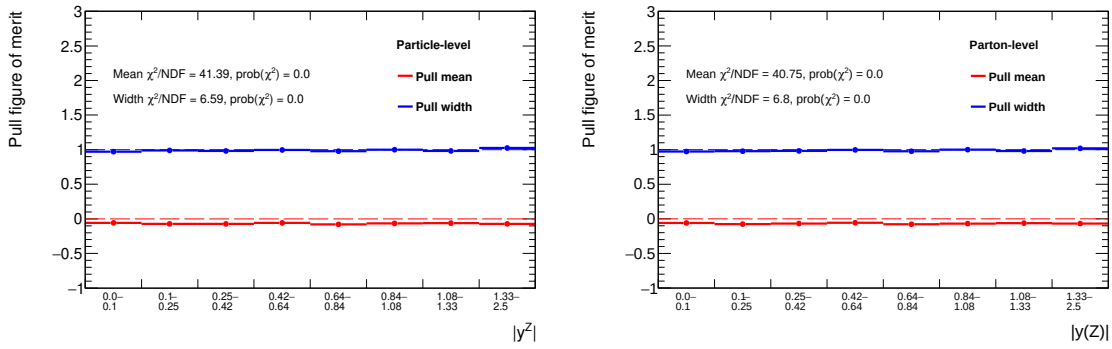


Fig. 34: Particle-level (left) and parton-level (right) summary plots for the pull tests (using original definition of pull from Eq. 6 and Gaussian fitting of the pull distributions) performed for $|y^Z|$ in the combined $3\ell + 4\ell$ channel.

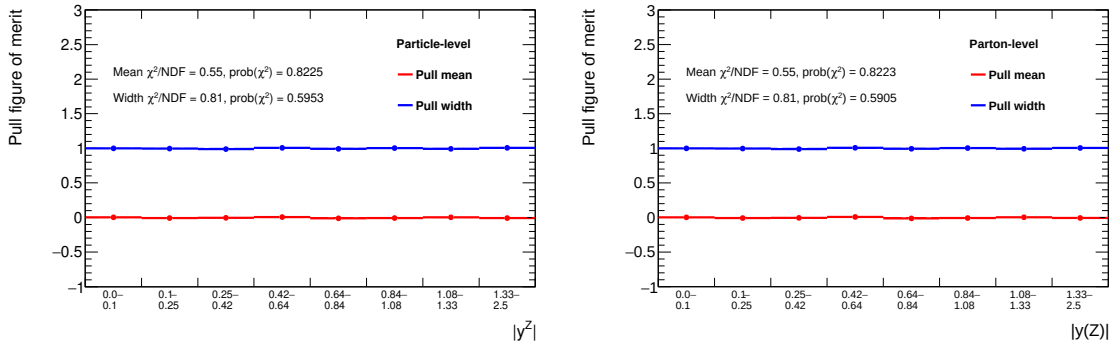


Fig. 35: Particle-level (left) and parton-level (right) summary plots for the updated pull tests (using modified pull definition from Eq. 10.6 and formulas from Eq. 10.7 for extraction of mean and width) performed for $|y^Z|$ in the combined $3\ell + 4\ell$ channel.

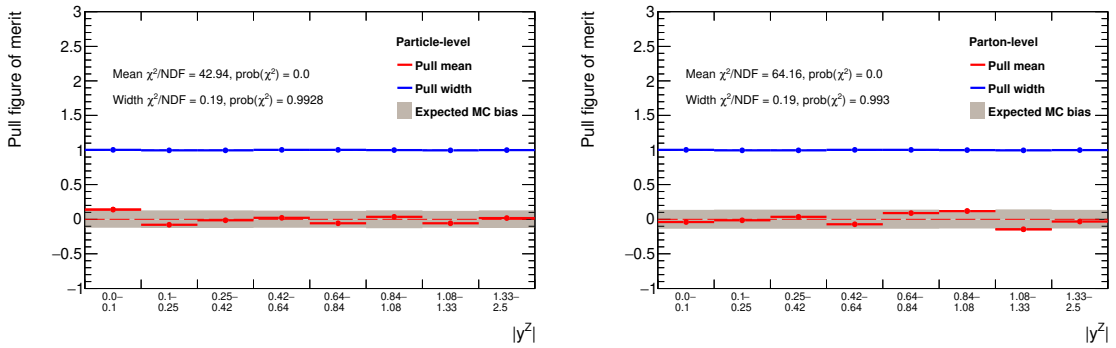


Fig. 36: Particle-level (left) and parton-level (right) summary plots for the pull tests using independent sample for building migration matrix and unfolding corrections, performed for $|y^Z|$ in the combined $3\ell + 4\ell$ channel.

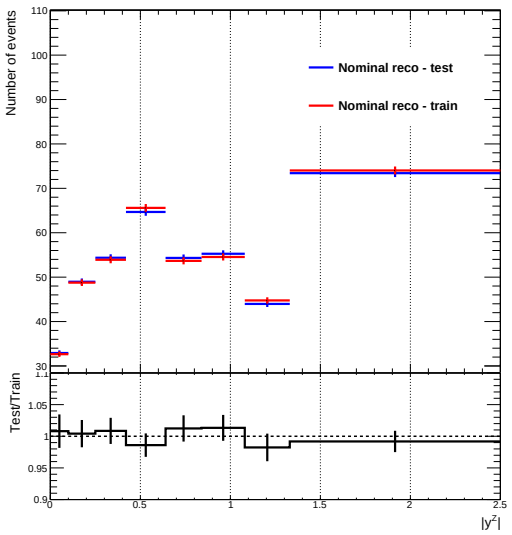


Fig. 37: Reconstruction-level distributions for the two independent samples used for the closure tests. Sample denoted as "train" is used for constructing migration matrix and evaluating unfolding corrections, while "test" sample is unfolded. The plots show distributions for $|y^Z|$ in the combined $3\ell + 4\ell$ channel.

A.4 Impact of statistically limited MC samples

A.4.1 Trilepton channel

		RMS	
Variable	**Bin**	**Particle level**	**Parton-level**
$p_{\mathrm{T}}^{\ell \mathrm{non}-Z}$	1	1.2 %	1.4 %
	2	1.3 %	1.4 %
	3	1.5 %	1.7 %
	4	1.8 %	2.0 %
$\lvert\Delta\phi(Z, t_{\mathrm{lep}})\rvert$	1	2.0 %	2.2 %
	2	1.2 %	1.3 %
	3	1.5 %	1.8 %
N_{jets}	1	1.7 %	-
	2	1.4 %	-
	3	1.3 %	-

Table 10: Impact of statistical limitations of the signal MC sample on the additional trilepton differential variables. The quoted values were derived based on half of the simulated events, such that the effective RMS is a factor $\sqrt{2}$ smaller.

A.4.2 Tetralepton channel

		RMS	
Variable	**Bin**	**Particle level**	**Parton-level**
$\|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)\|$	1	3.1 %	3.5 %
	2	2.9 %	3.2 %
	3	2.9 %	3.4 %
	4	2.7 %	3.0 %
$\|\Delta\phi(t\bar{t}, Z)\|$	1	4.2 %	4.2 %
	2	3.2 %	3.2 %
	3	2.8 %	3.0 %
N_{jets}	1	2.5 %	-
	2	3.1 %	-
	3	5.2 %	-
	4	8.2 %	-

Table 11: Impact of statistical limitations of the signal MC sample on the additional tetralepton differential variables. The quoted values were derived based on half of the simulated events, such that the effective RMS is a factor $\sqrt{2}$ smaller.

A.4.3 Combined $3\ell + 4\ell$ channel

		RMS	
Variable	**Bin**	**Particle level**	**Parton-level**
$\|y^Z\|$	1	2.3 %	2.5 %
	2	1.9 %	2.1 %
	3	1.8 %	2.0 %
	4	1.6 %	1.8 %
	5	1.7 %	2.0 %
	6	1.8 %	1.9 %
	7	1.9 %	2.2 %
	8	1.5 %	1.6 %

Table 12: Impact of statistical limitations of the signal MC sample on $|y^Z|$ in the combined 3ℓ+4ℓ channel. The quoted values were derived based on half of the simulated events, such that the effective RMS is a factor $\sqrt{2}$ smaller.

A.5 Stress tests

A.5.1 Trilepton channel

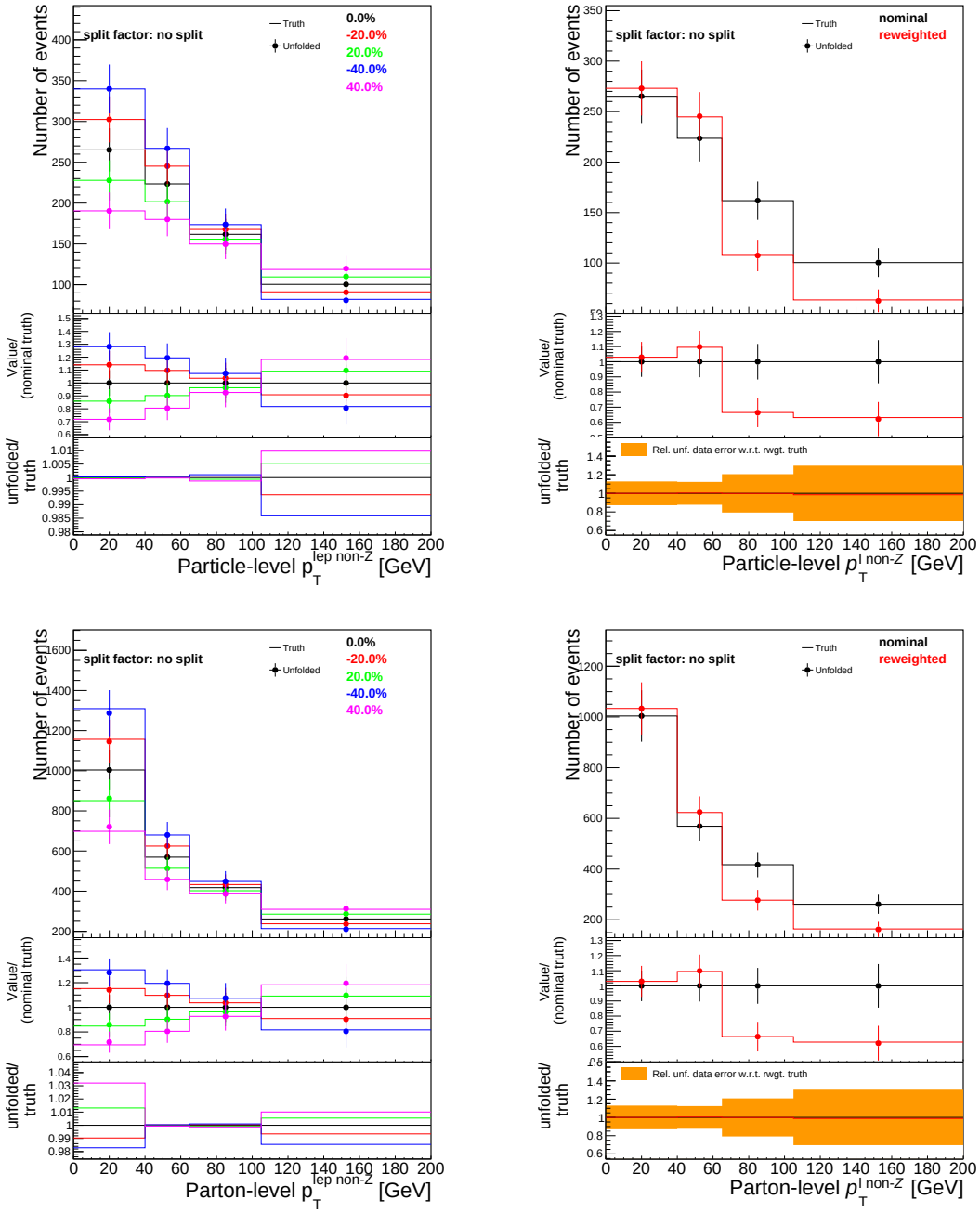


Fig. 38: Results of the linear (left) and data-driven (right) stress tests for the $p_{\mathrm{T}}^{\ell\mathrm{non}-Z}$ in the trilepton channel, unfolded to particle level (top) and parton level (bottom).

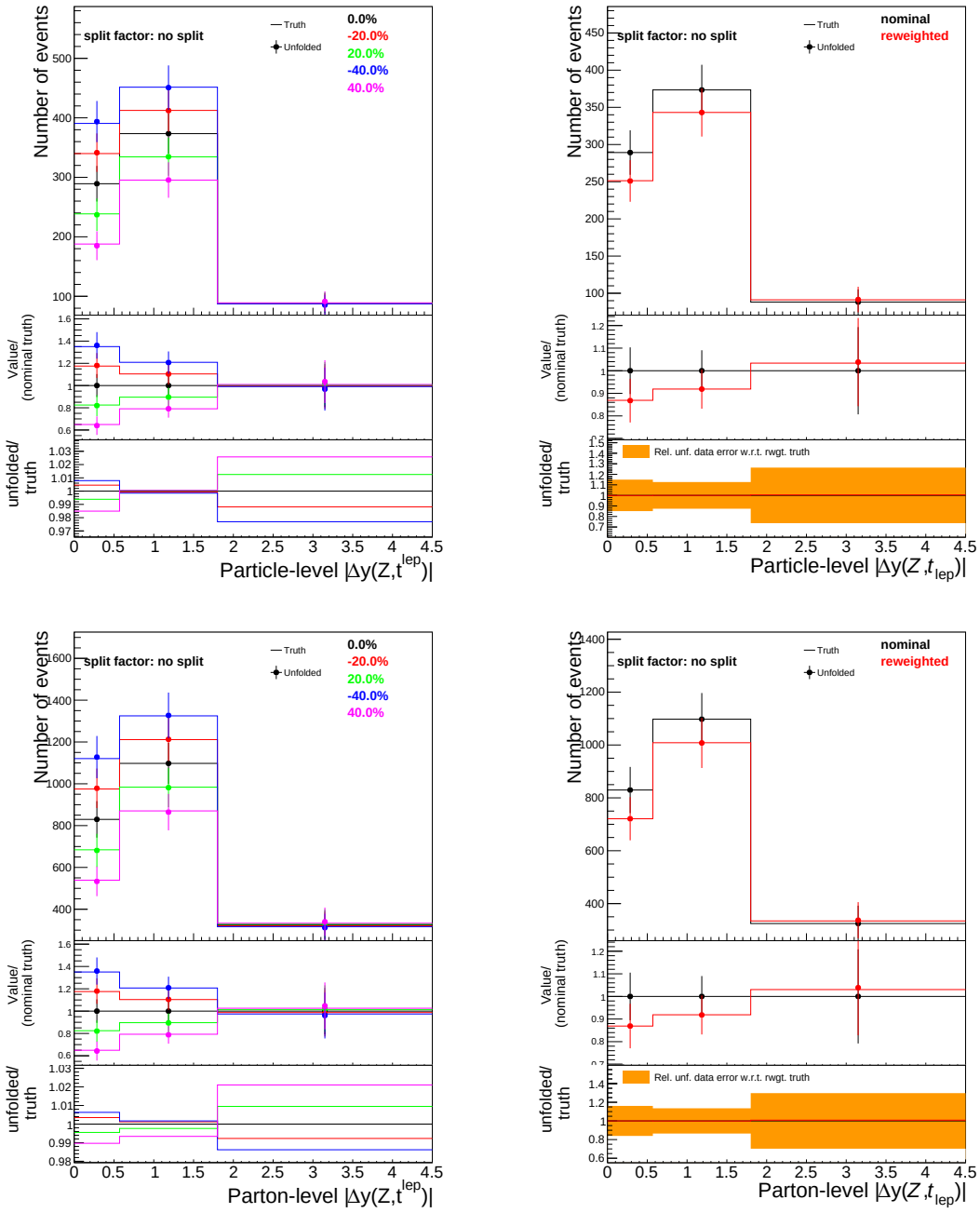


Fig. 39: Results of the linear (left) and data-driven (right) stress tests for the $|\Delta y(Z, t_{\text{lep}})|$ in the trilepton channel, unfolded to particle level (top) and parton level (bottom).

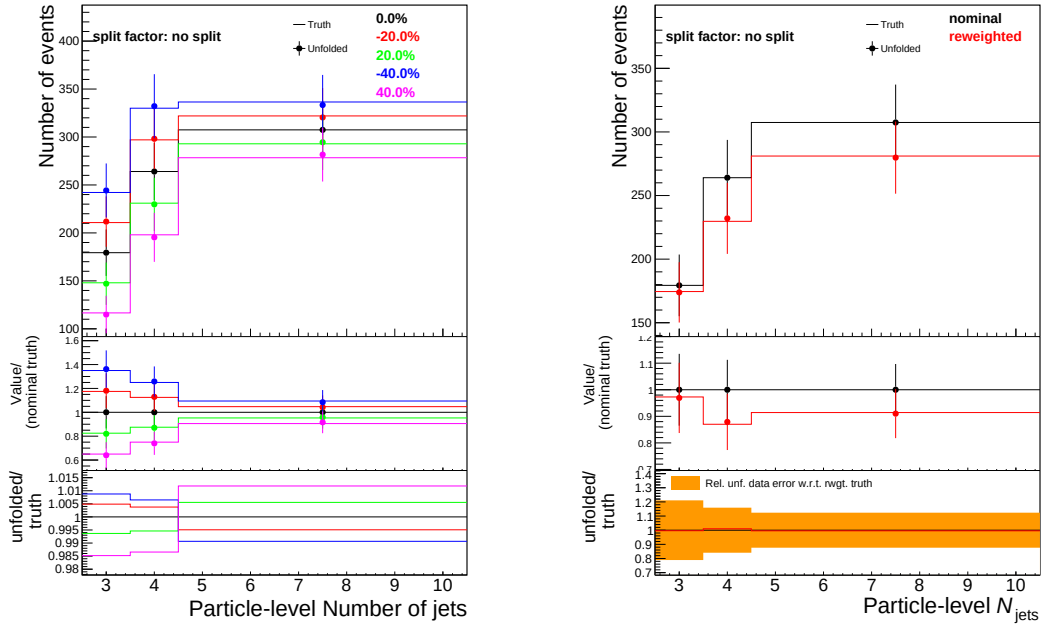


Fig. 40: Results of the linear (left) and data-driven (right) stress tests for jet multiplicity in the trilepton channel.

A.5.2 Tetralepton channel

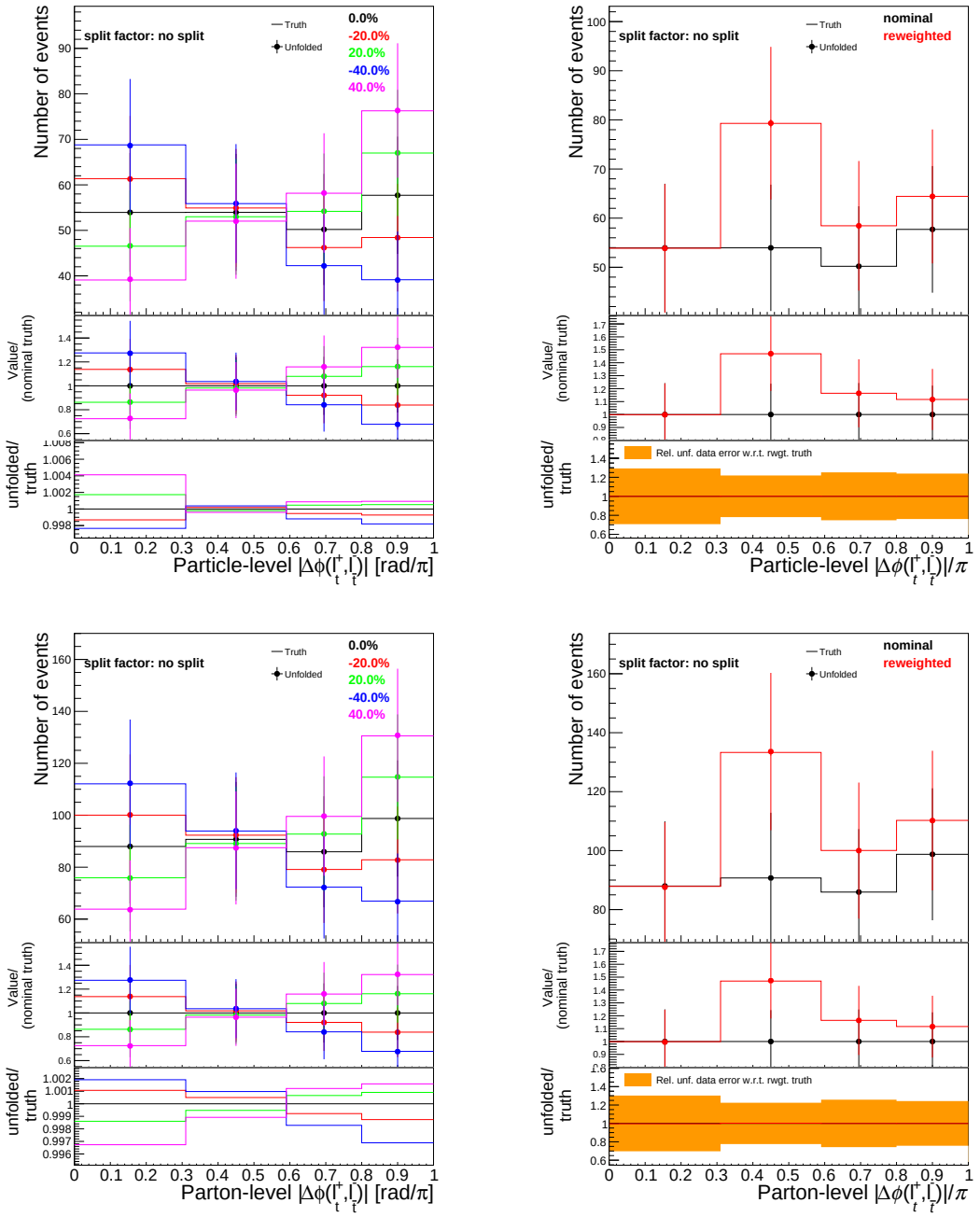


Fig. 41: Results of the linear (left) and data-driven (right) stress tests for the $|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)|$ in the tetralepton channel, unfolded to particle level (top) and parton level (bottom).

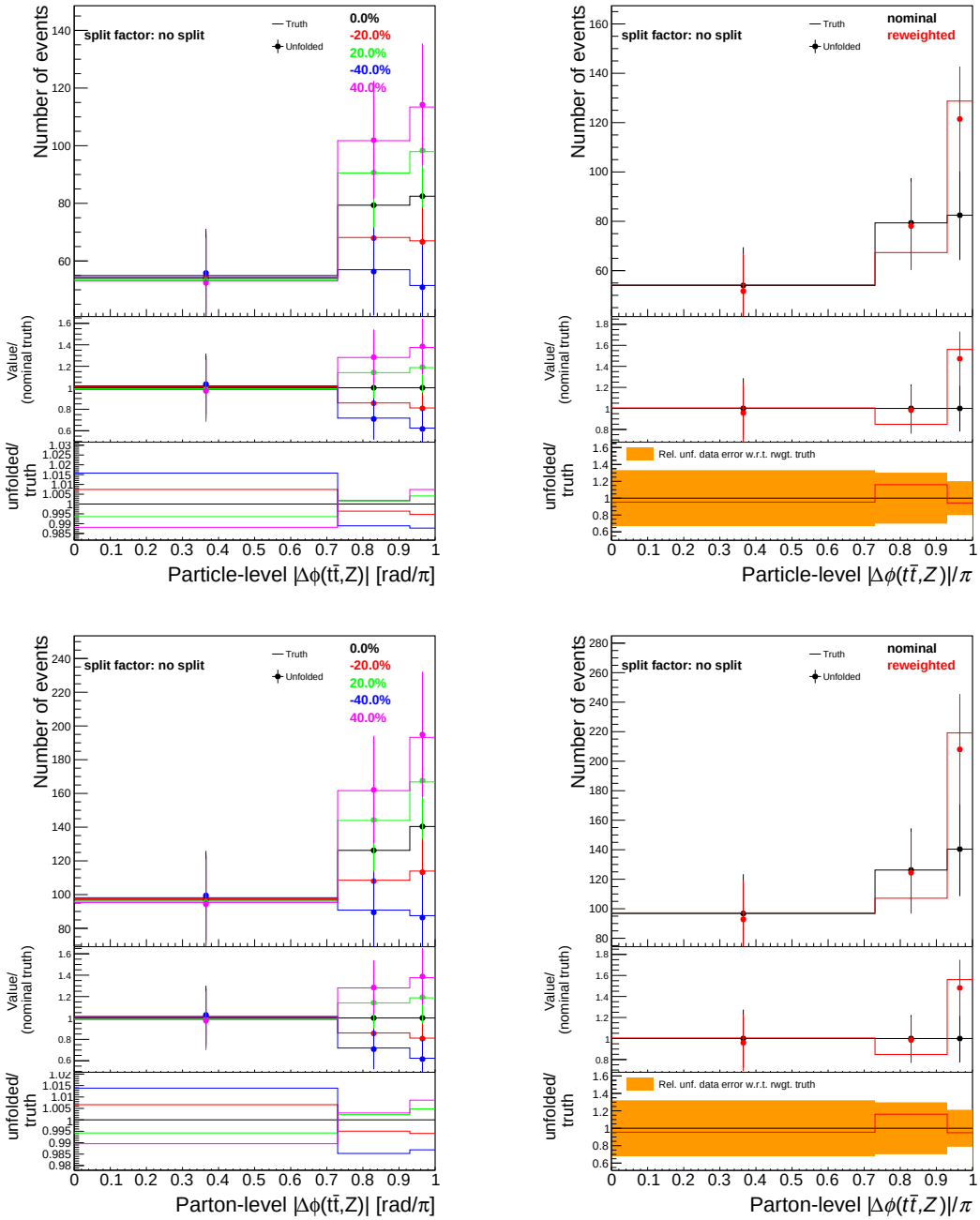


Fig. 42: Results of the linear (left) and data-driven (right) stress tests for the $|\Delta\phi(t\bar{t}, Z)|$ in the tetralepton channel, unfolded to particle level (top) and parton level (bottom).

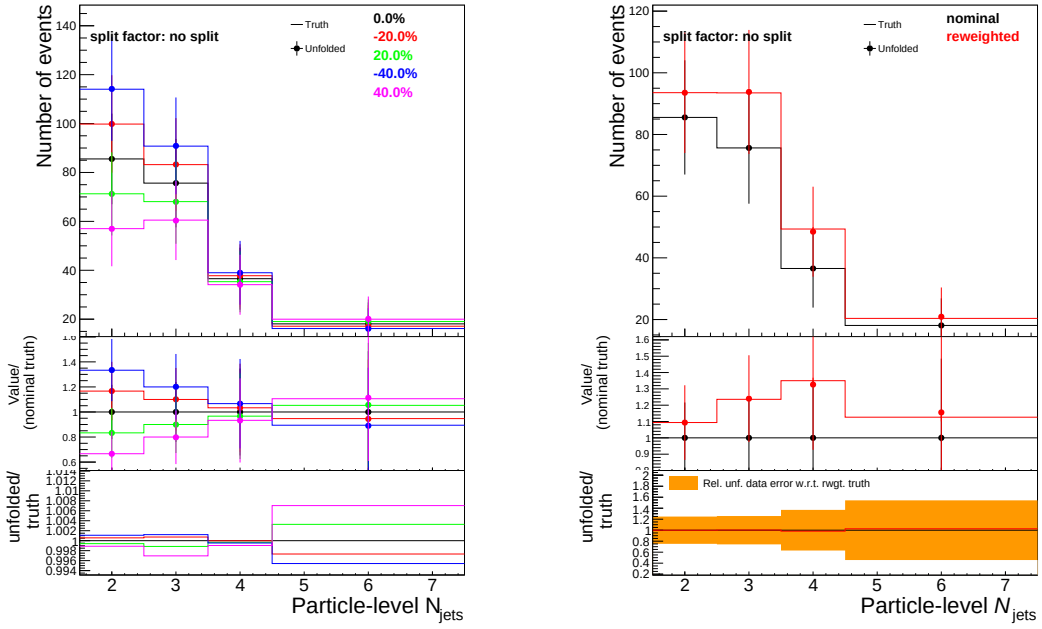


Fig. 43: Results of the linear (left) and data-driven (right) stress tests for jet multiplicity in the tetralepton channel.

A.5.3 Combined $3\ell + 4\ell$ channel

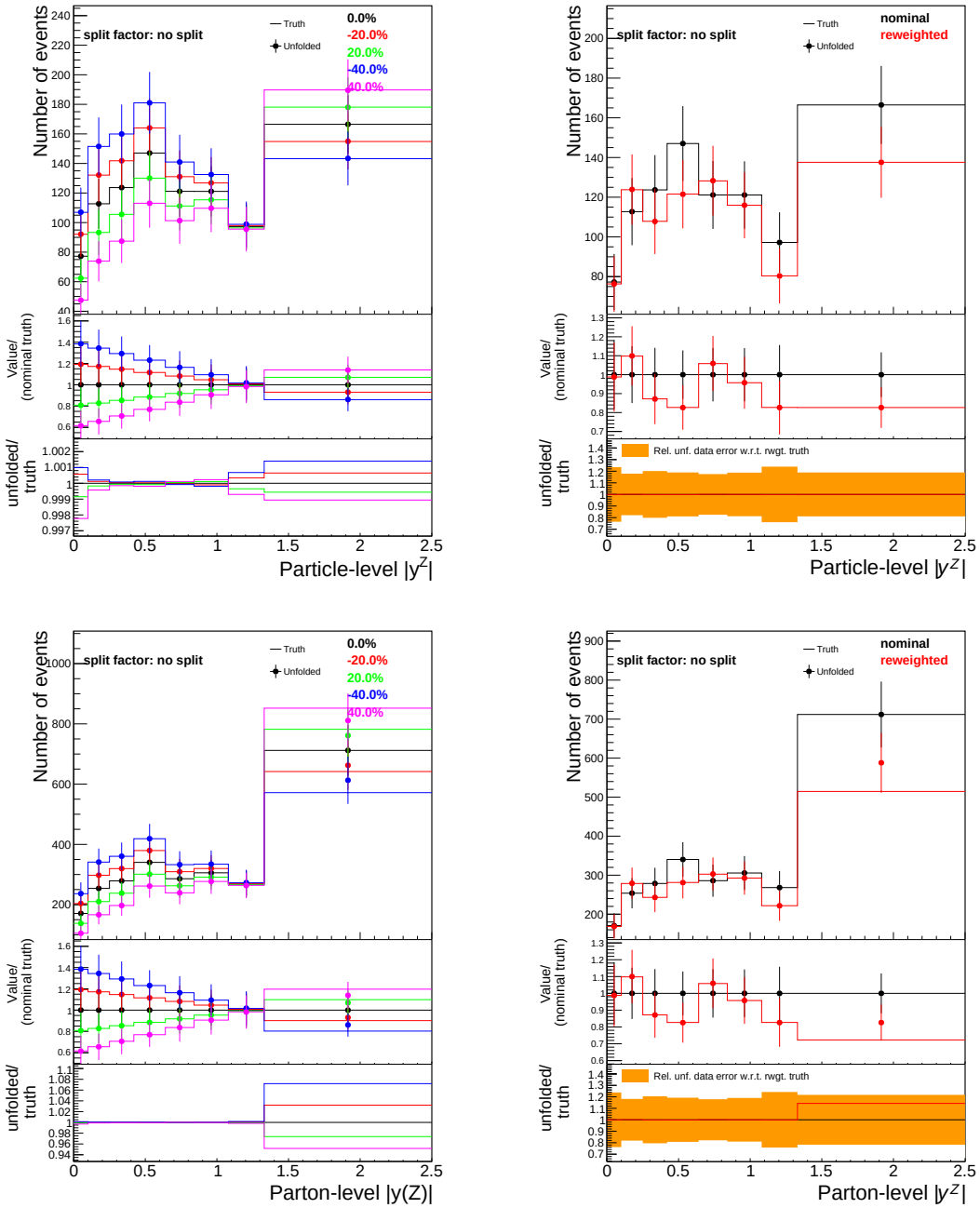


Fig. 44: Results of the linear (left) and data-driven (right) stress tests for the $|y^Z|$ in the combined $3\ell + 4\ell$ channel, unfolded to particle level (top) and parton level (bottom).

A.6 Observed differential cross sections

A.6.1 Trilepton channel

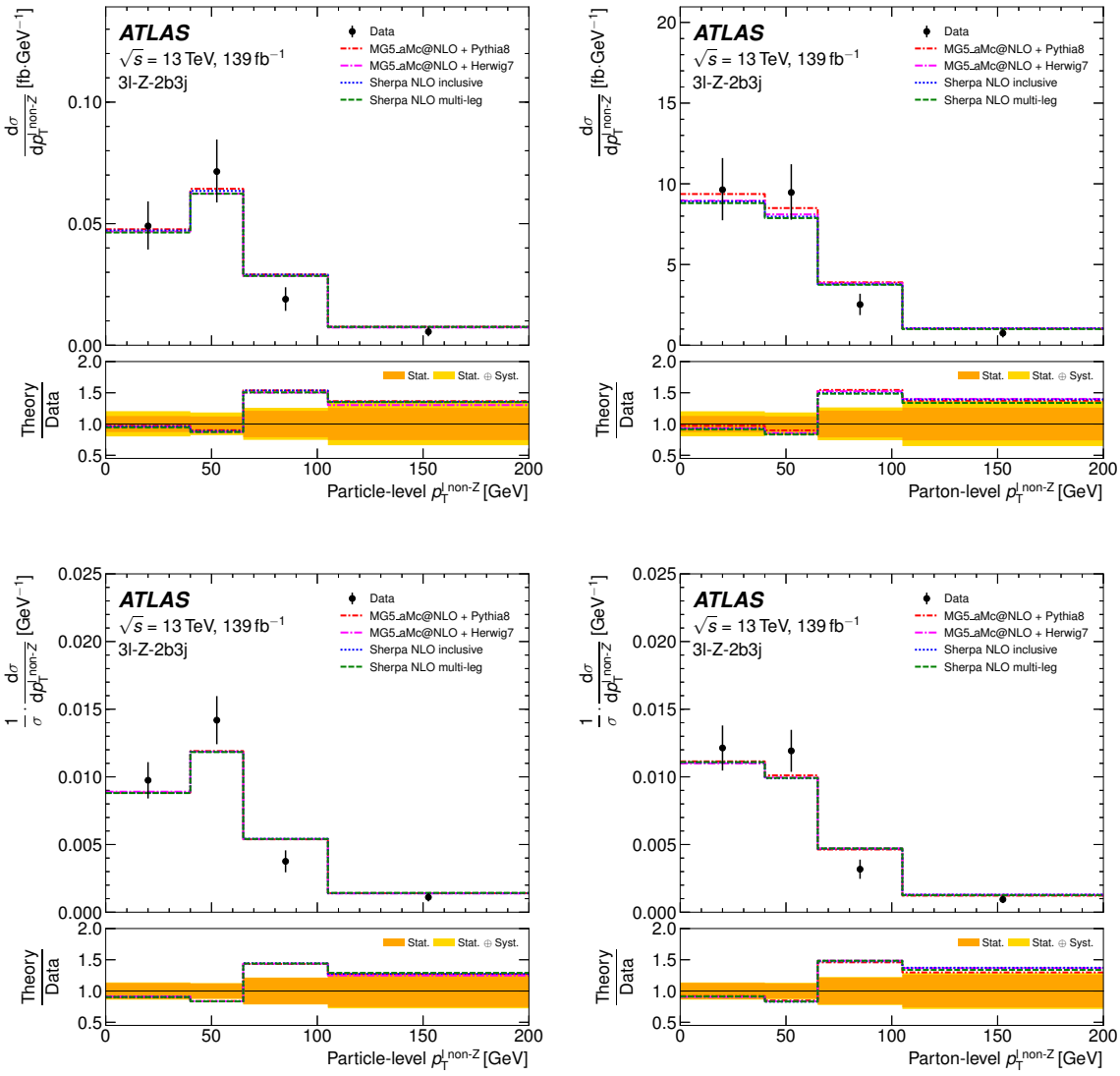


Fig. 45: Absolute (top) and normalized (bottom) distributions of the differential cross section as a function of $p_{\mathrm{T}}^{\ell\mathrm{non}-Z}$ in the trilepton channel, unfolded to particle (left) and parton level (right). The bottom ratio plots show statistical and combined (statistical $\oplus$ systematic) uncertainty on the measured differential cross sections.

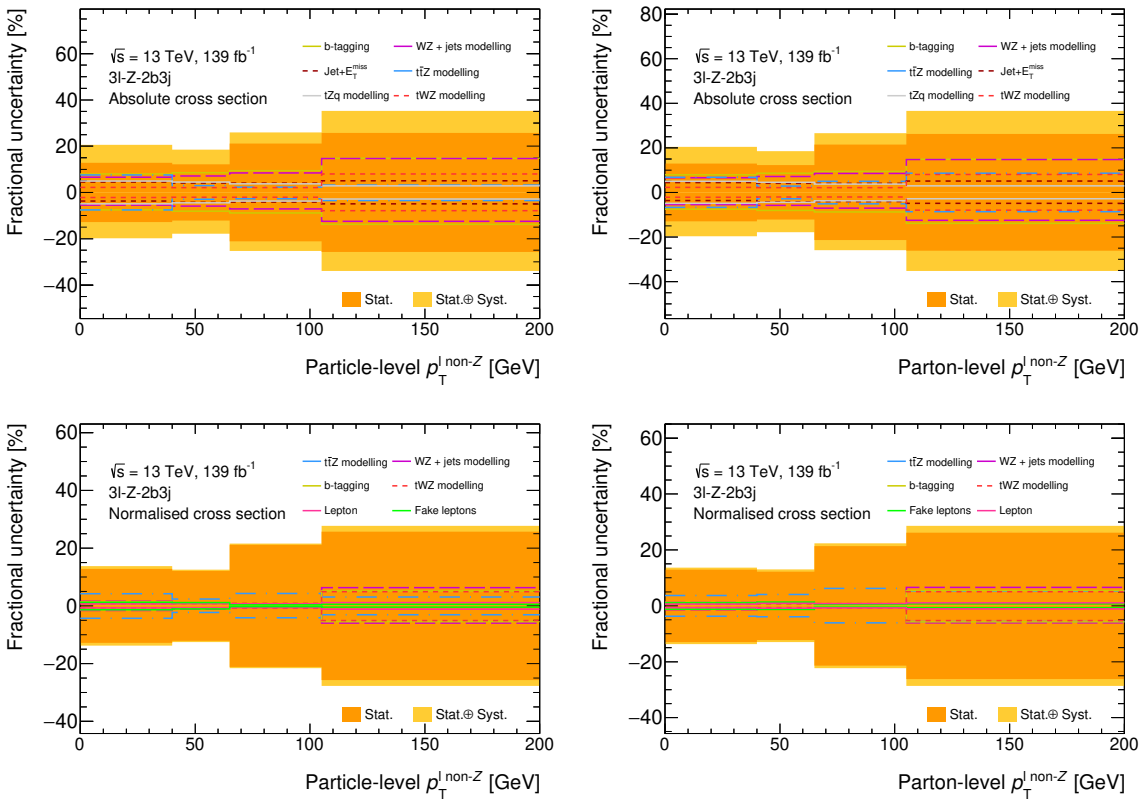


Fig. 46: Absolute (top) and normalized (bottom) fractional decomposition of the uncertainties (statistical ⊕ systematic) on the differential cross section as a function of $p_{\mathrm{T}}^{\ell \mathrm{non}-Z}$ in the trilepton channel, unfolded to particle (left) and parton level (right).

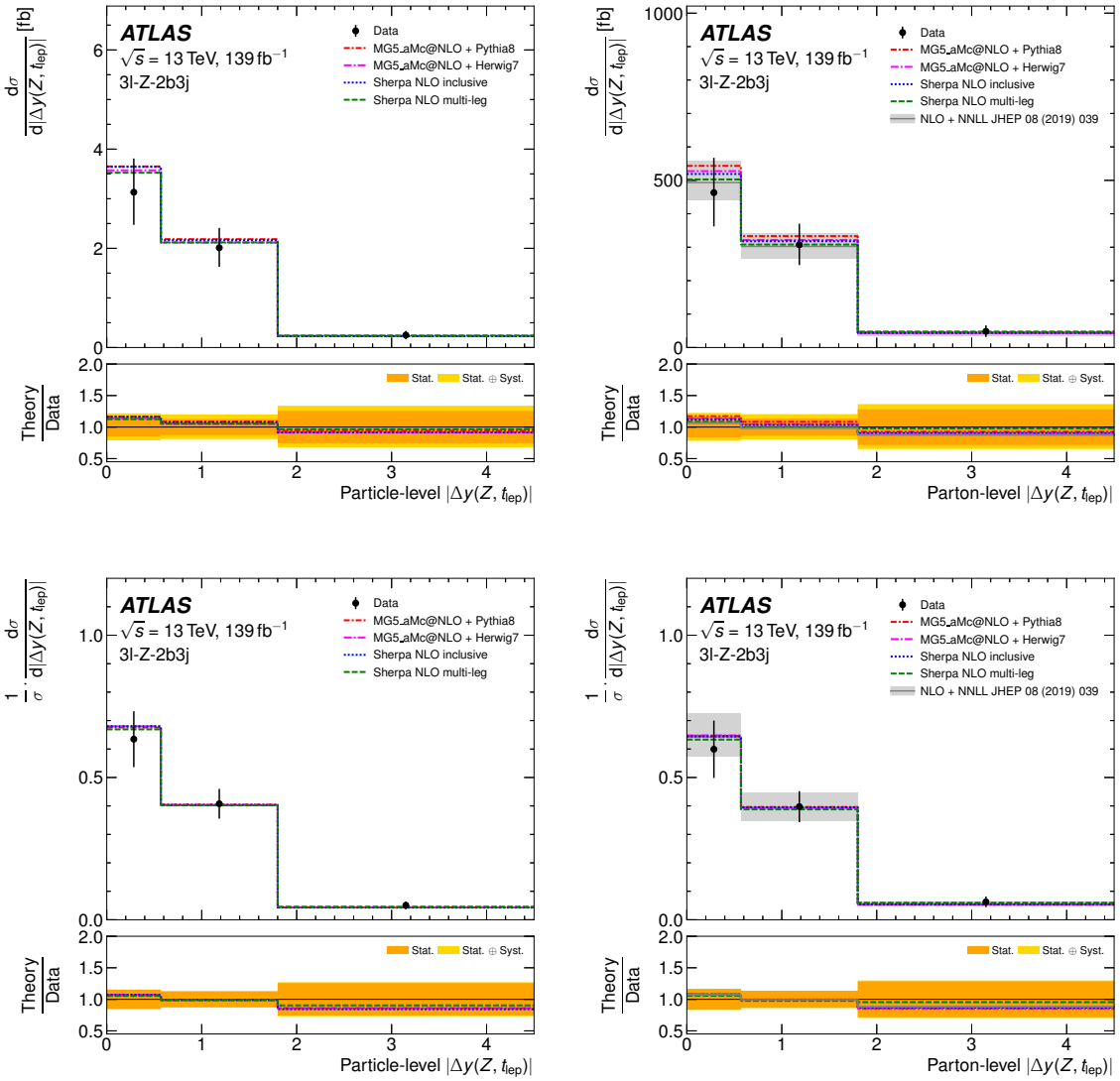


Fig. 47: Absolute (top) and normalized (bottom) distributions of the differential cross section as a function of $|\Delta y(Z, t_{\text{lep}})|$ in the trilepton channel, unfolded to particle (left) and parton level (right). The bottom ratio plots show statistical and combined (statistical $\oplus$ systematic) uncertainty on the measured differential cross sections.

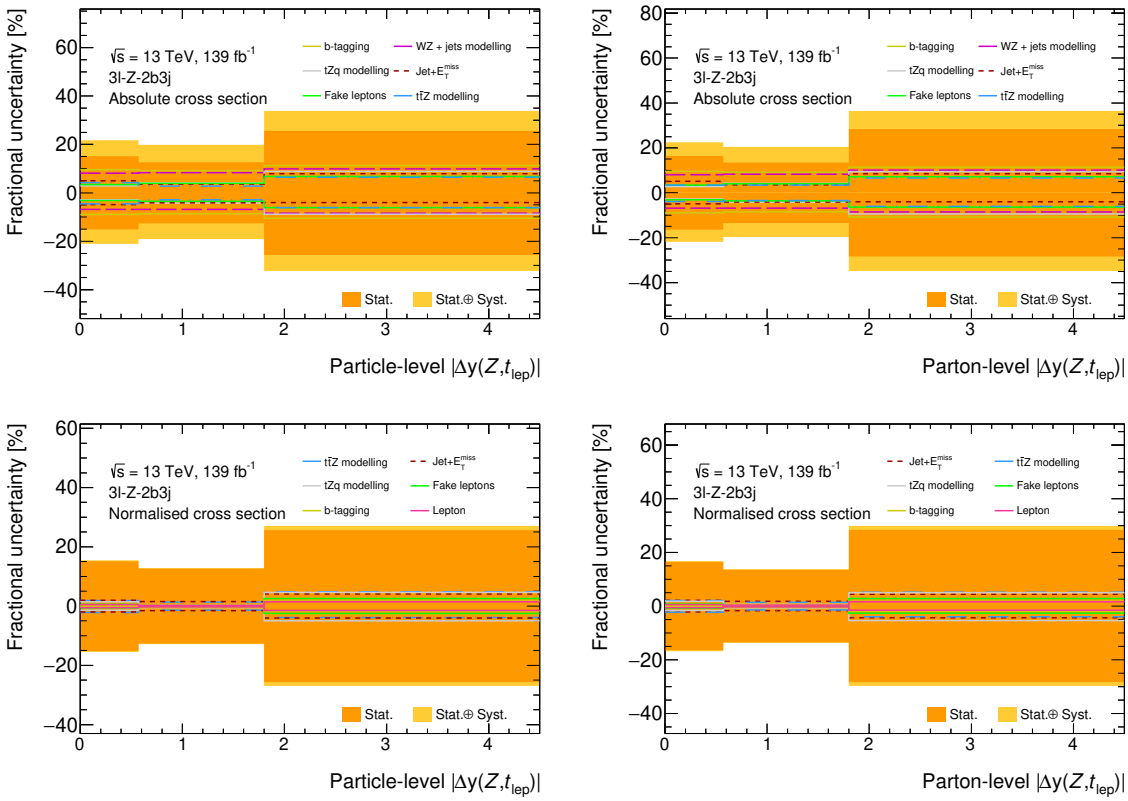


Fig. 48: Absolute (top) and normalized (bottom) fractional decomposition of the uncertainties (statistical $\oplus$ systematic) on the differential cross section as a function of $|\Delta y(Z, t_{\text{lep}})|$ in the trilepton channel, unfolded to particle (left) and parton level (right).

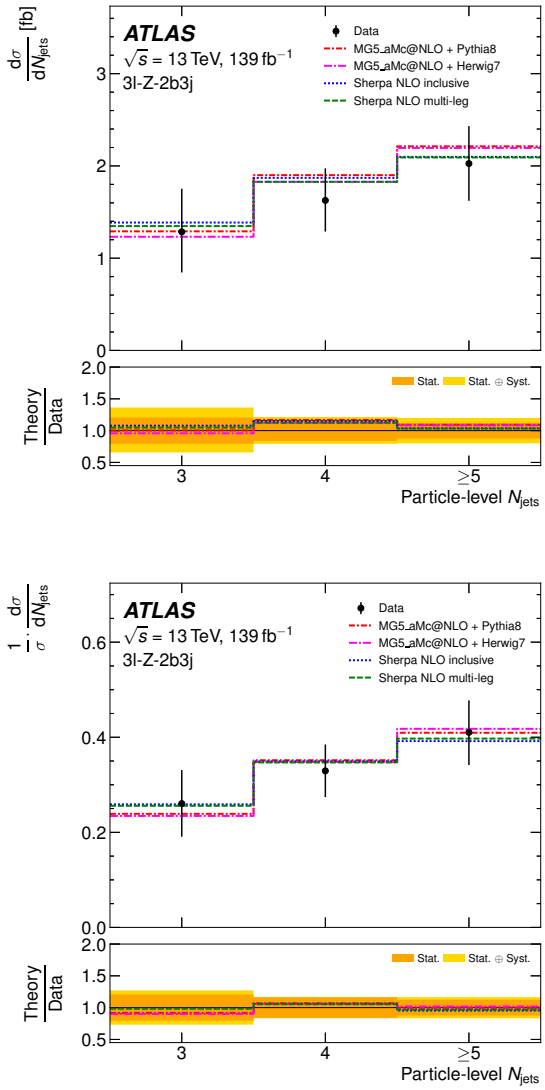


Fig. 49: Absolute (top) and normalized (bottom) distributions of the differential cross section as a function of N_{jets} in the trilepton channel, unfolded to particle level. The bottom ratio plots show statistical and combined (statistical $\oplus$ systematic) uncertainty on the measured differential cross sections.

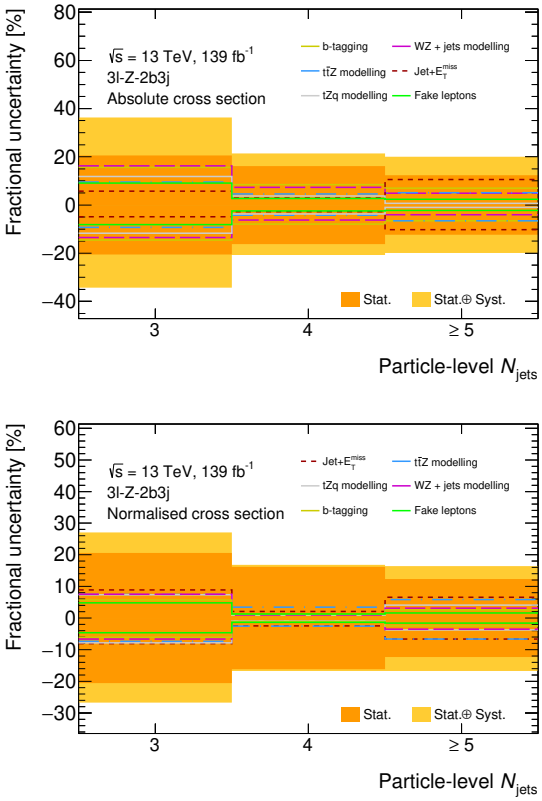


Fig. 50: Absolute (top) and normalized (bottom) fractional decomposition of the uncertainties (statistical $\oplus$ systematic) on the differential cross section as a function of N_{jets} in the trilepton channel, unfolded to particle level.

A.6.2 Tetralepton channel

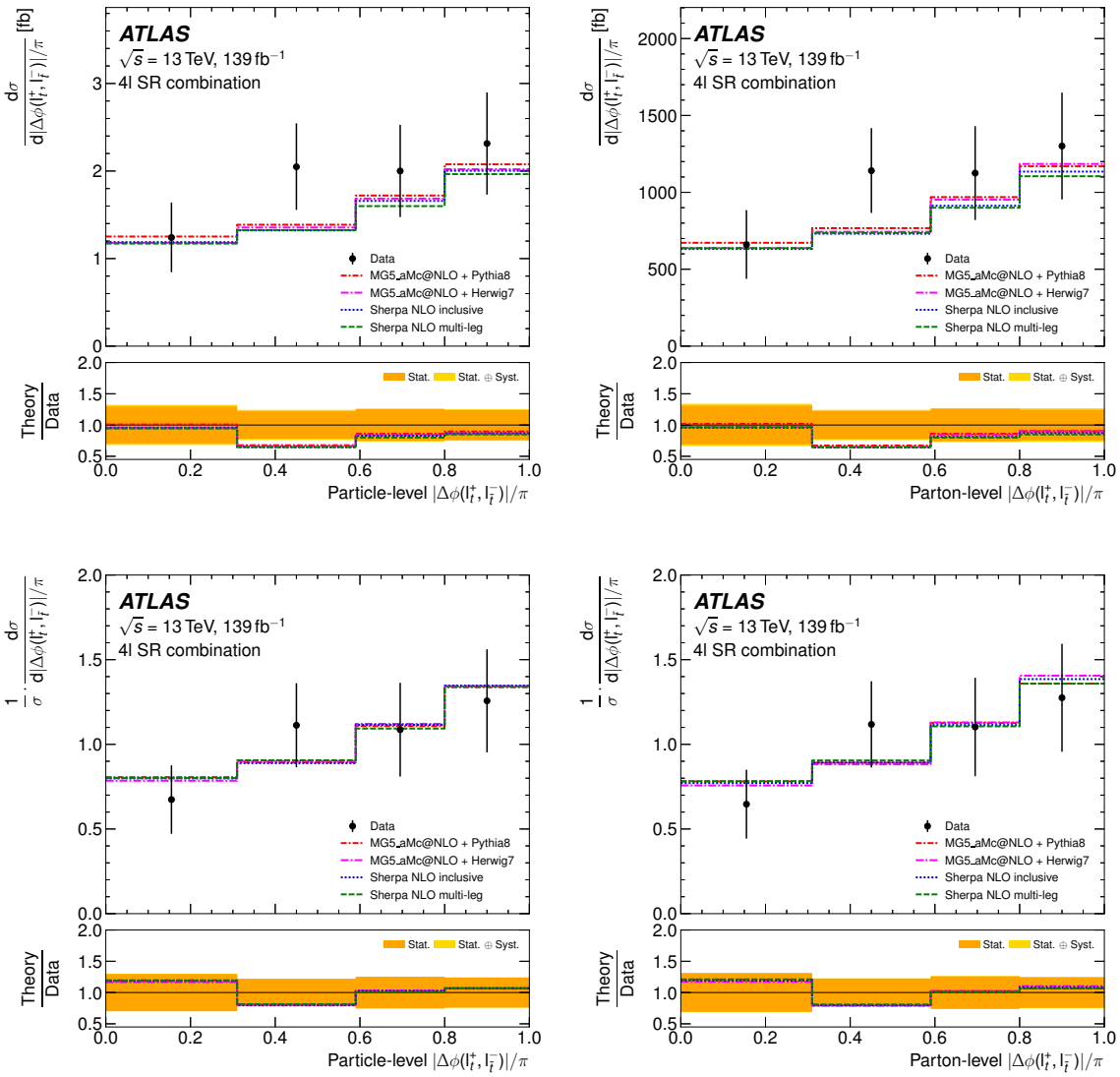


Fig. 51: Absolute (top) and normalized (bottom) distributions of the differential cross section as a function of $|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)|$ in the tetralepton channel, unfolded to particle (left) and parton level (right). The bottom ratio plots show statistical and combined (statistical $\oplus$ systematic) uncertainty on the measured differential cross sections.

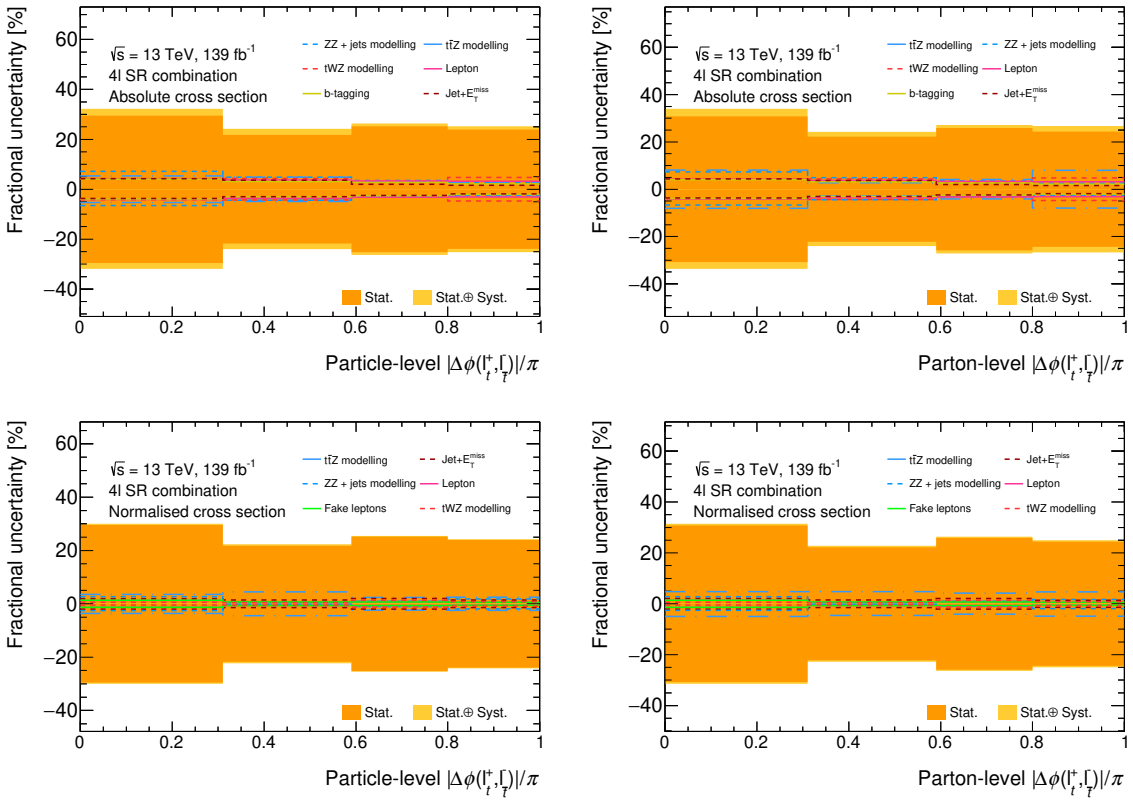


Fig. 52: Absolute (top) and normalized (bottom) fractional decomposition of the uncertainties (statistical $\oplus$ systematic) on the differential cross section as a function of $|\Delta\phi(\ell_t^+, \ell_{\bar{t}}^-)|$ in the tetralepton channel, unfolded to particle (left) and parton level (right).

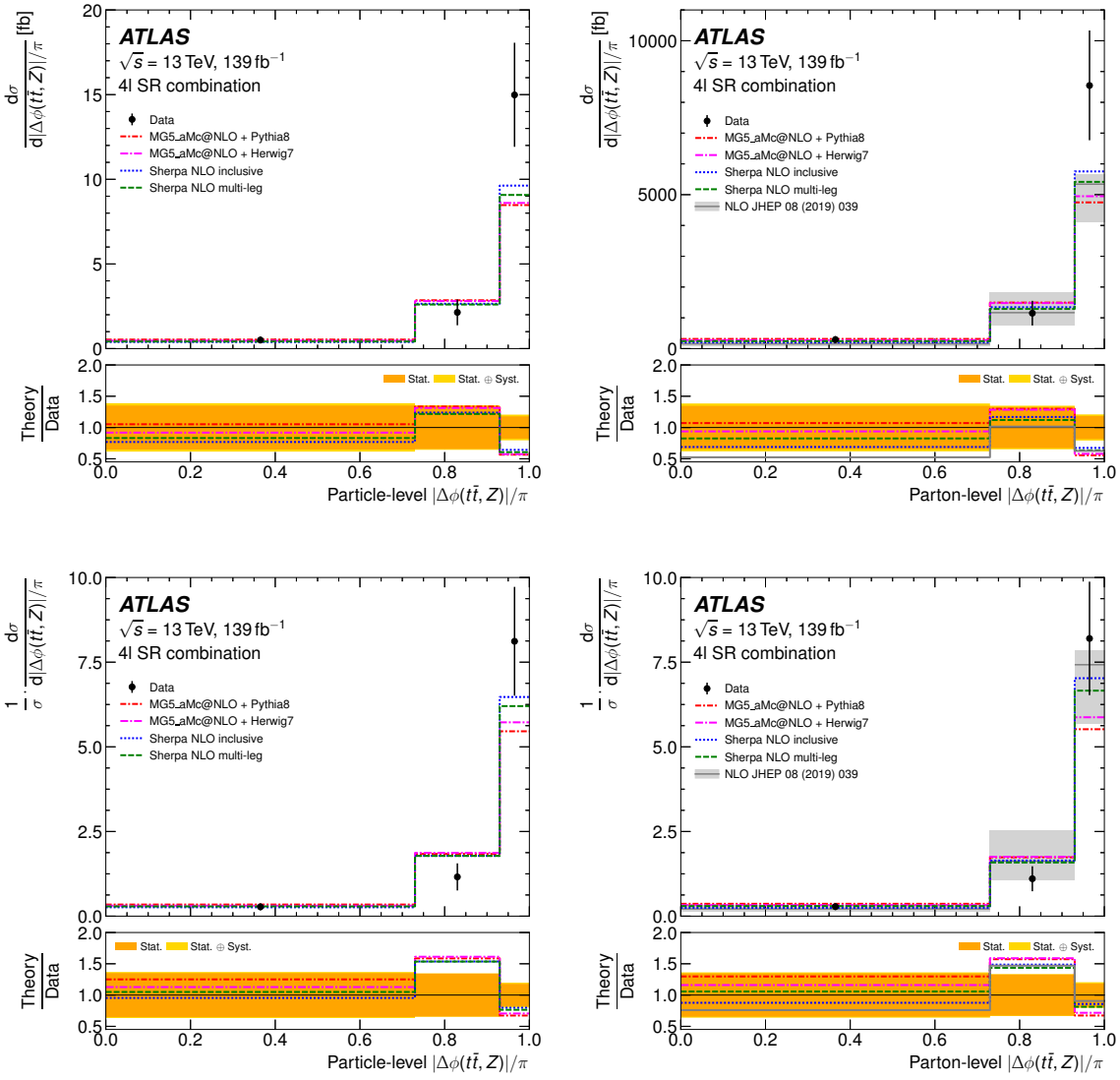


Fig. 53: Absolute (top) and normalized (bottom) distributions of the differential cross section as a function of $|\Delta\phi(t\bar{t}, Z)|$ in the tetralepton channel, unfolded to particle (left) and parton level (right). The bottom ratio plots show statistical and combined (statistical $\oplus$ systematic) uncertainty on the measured differential cross sections.

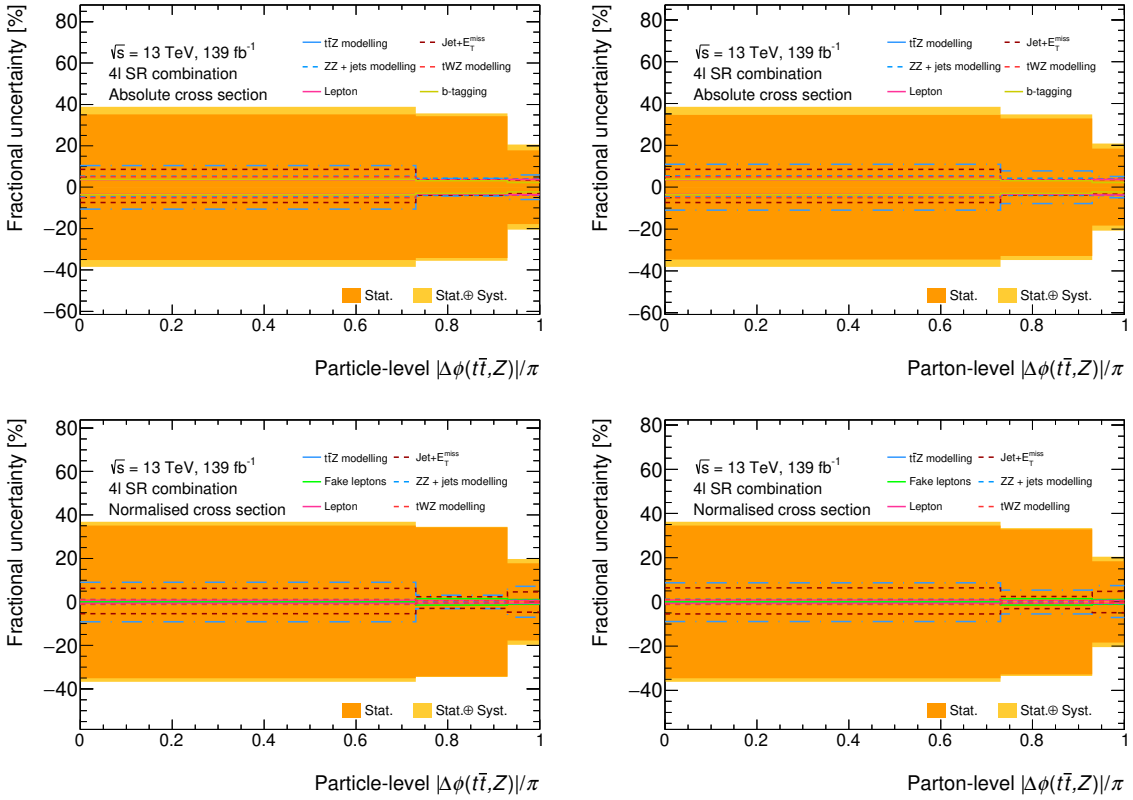


Fig. 54: Absolute (top) and normalized (bottom) fractional decomposition of the uncertainties (statistical $\oplus$ systematic) on the differential cross section as a function of $|\Delta\phi(t\bar{t}, Z)|$ in the tetralepton channel, unfolded to particle (left) and parton level (right).

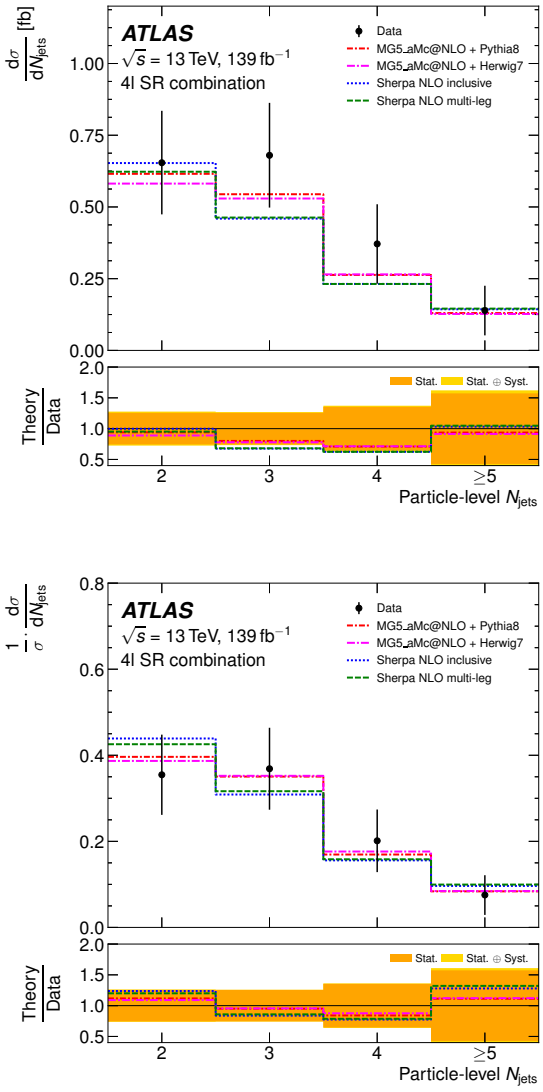


Fig. 55: Absolute (top) and normalized (bottom) distributions of the differential cross section as a function of N_{jets} in the tetralepton channel, unfolded to particle level. The bottom ratio plots show statistical and combined (statistical $\oplus$ systematic) uncertainty on the measured differential cross sections.

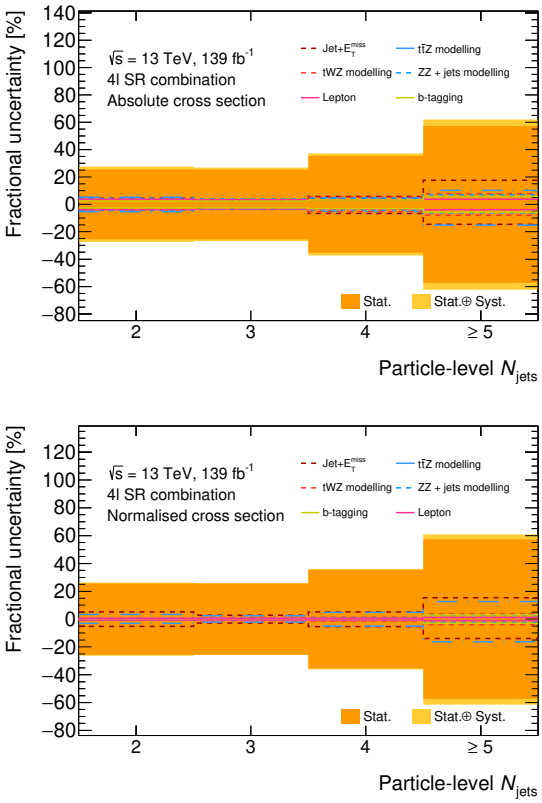


Fig. 56: Absolute (top) and normalized (bottom) fractional decomposition of the uncertainties (statistical $\oplus$ systematic) on the differential cross section as a function of N_{jets} in the tetralepton channel, unfolded to particle level.

A.6.3 Combined $3\ell + 4\ell$ channel

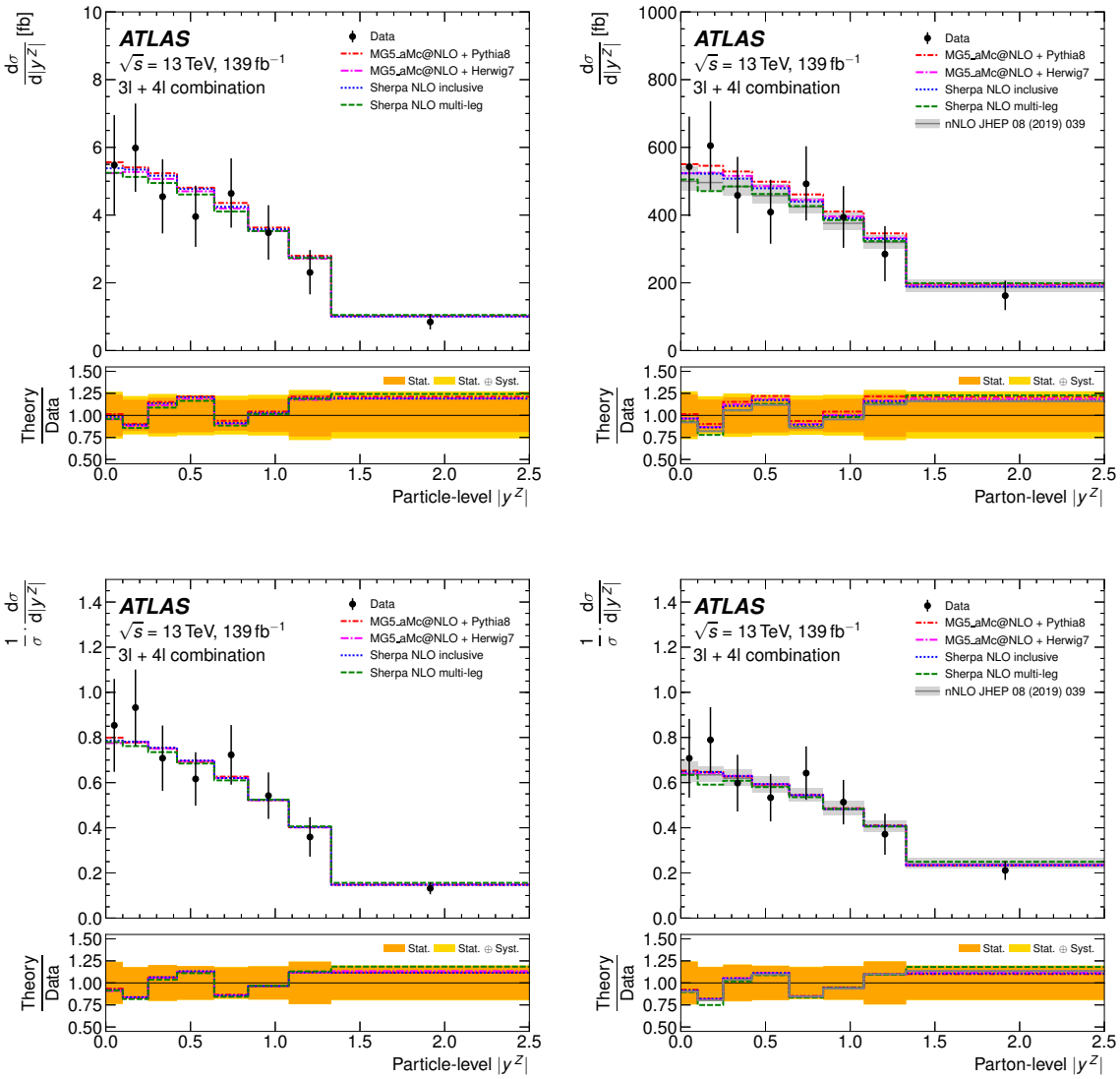


Fig. 57: Absolute (top) and normalized (bottom) distributions of the differential cross section as a function of $|y^Z|$ in the combined $3\ell + 4\ell$ channel, unfolded to particle (left) and parton level (right). The bottom ratio plots show statistical and combined (statistical $\oplus$ systematic) uncertainty on the measured differential cross sections.

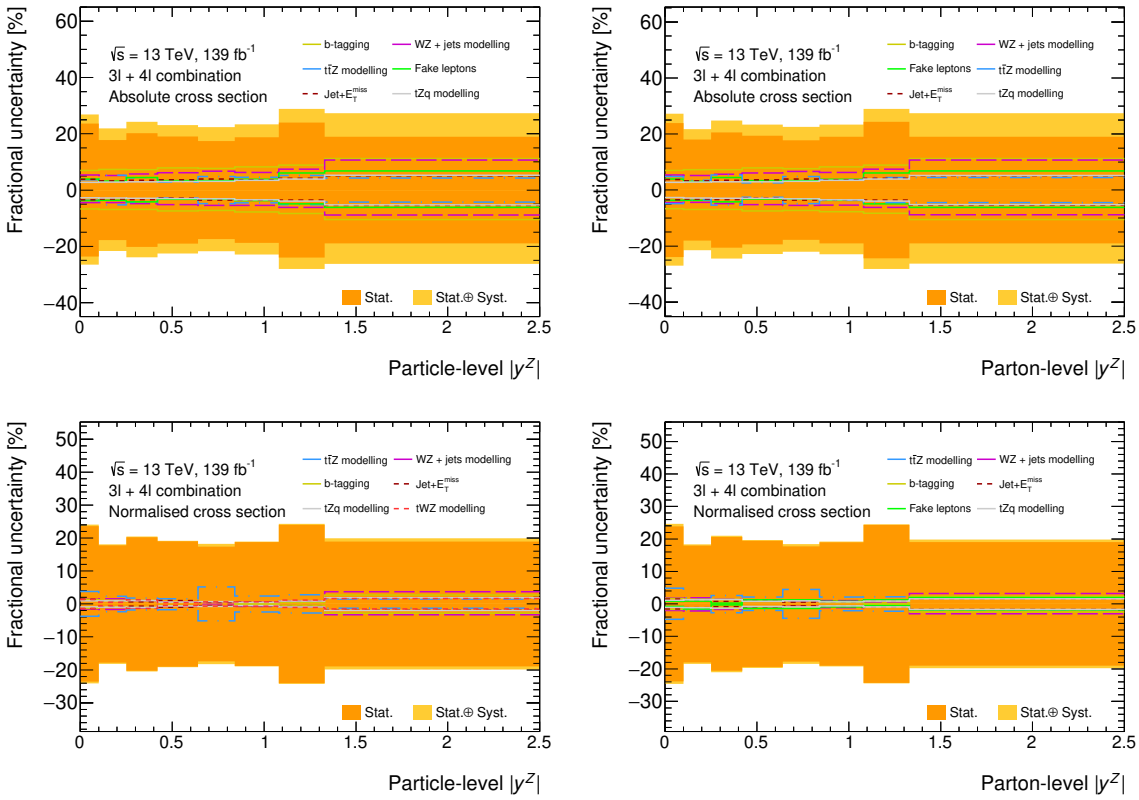


Fig. 58: Absolute (top) and normalized (bottom) fractional decomposition of the uncertainties (statistical $\oplus$ systematic) on the differential cross section as a function of $|y^Z|$ in the combined $3\ell+4\ell$ channel, unfolded to particle (left) and parton level (right).

B. Inclusive cross section measurement

This appendix contains additional figures related to the inclusive cross section measurement in the dilepton channel. Figures include plots relevant for training of the classification NN for each signal region.

B.1 Control and separation plots for selected NN variables

This section present control plots displaying agreement between observed data and simulated MC samples for NN input variables in the first (and third) row, followed by corresponding separation plots in the second (and fourth) row, which compare shape of the signal and background distributions. Uncertainties in the control plots include statistical component as well as contribution from all systematic sources. In the separation plots, both signal and background distributions are normalized to unity. Quoted separation (in %) is calculated according to Eq. 7.15.

B.1.1 2ℓ-Z-$5j2b$ region

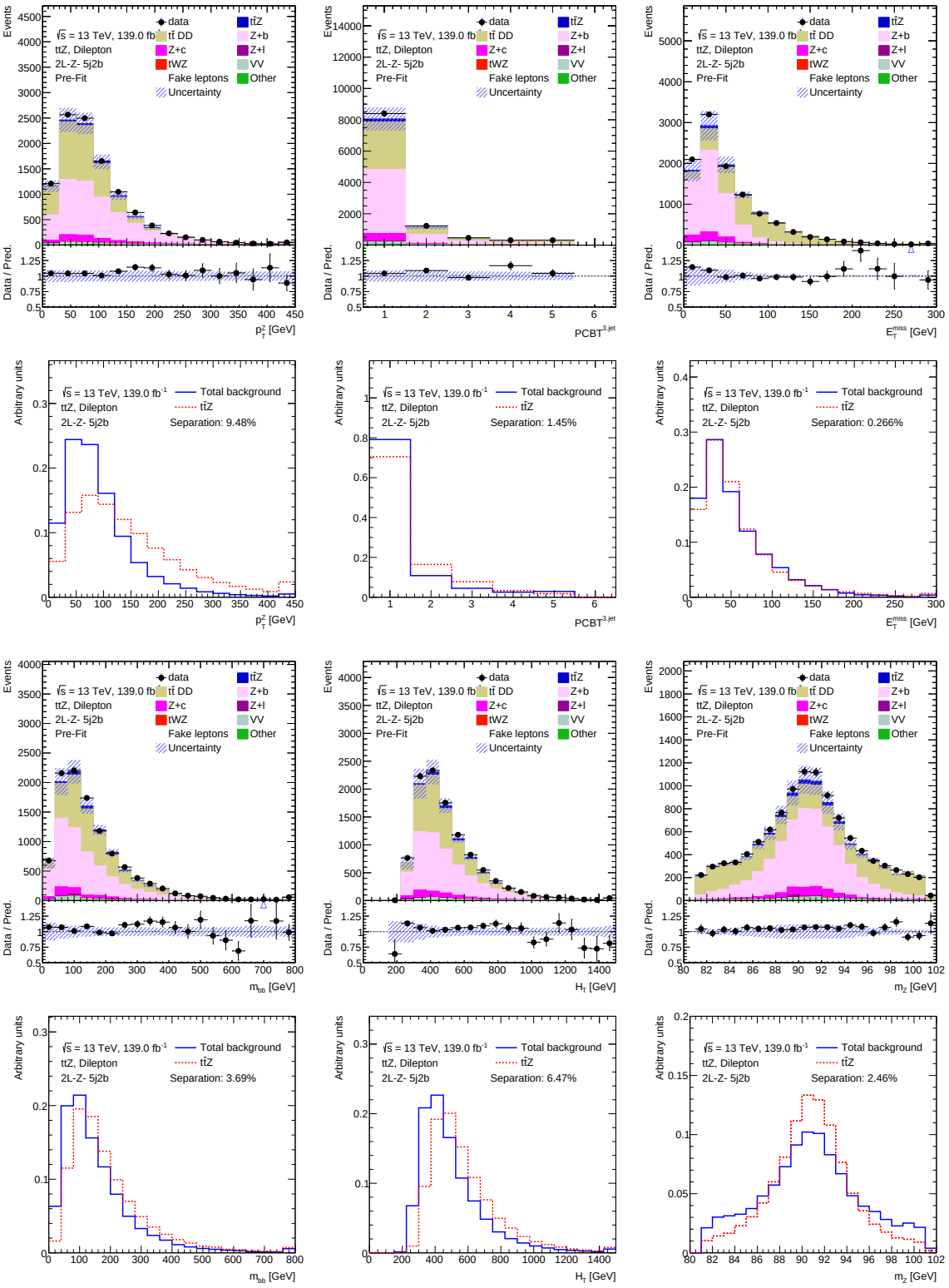


Fig. 59: Neural Network input variables used in 2ℓ-Z-$5j2b$ region.

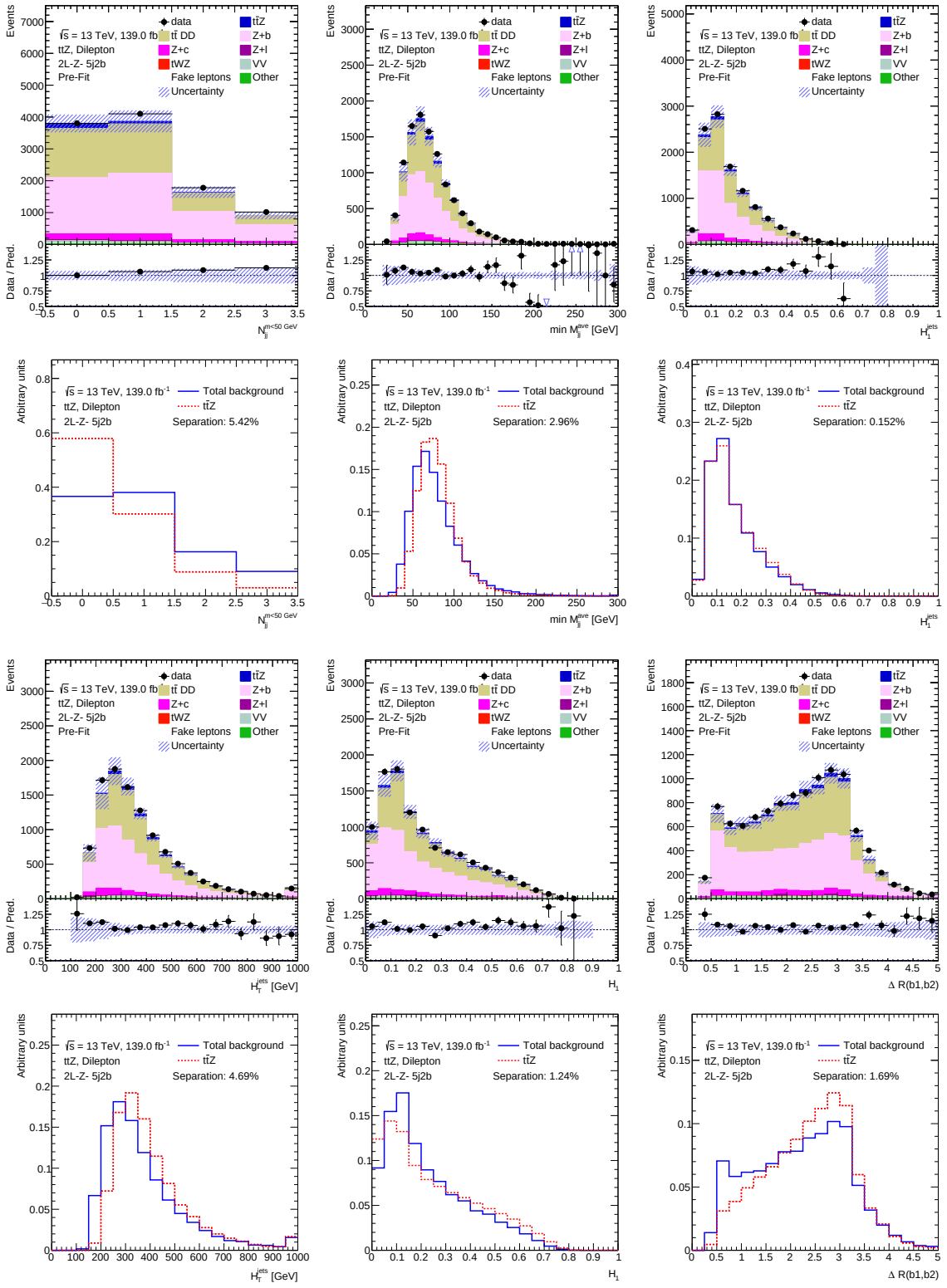


Fig. 60: Neural Network input variables used in 2ℓ-Z-$5j2b$ region.

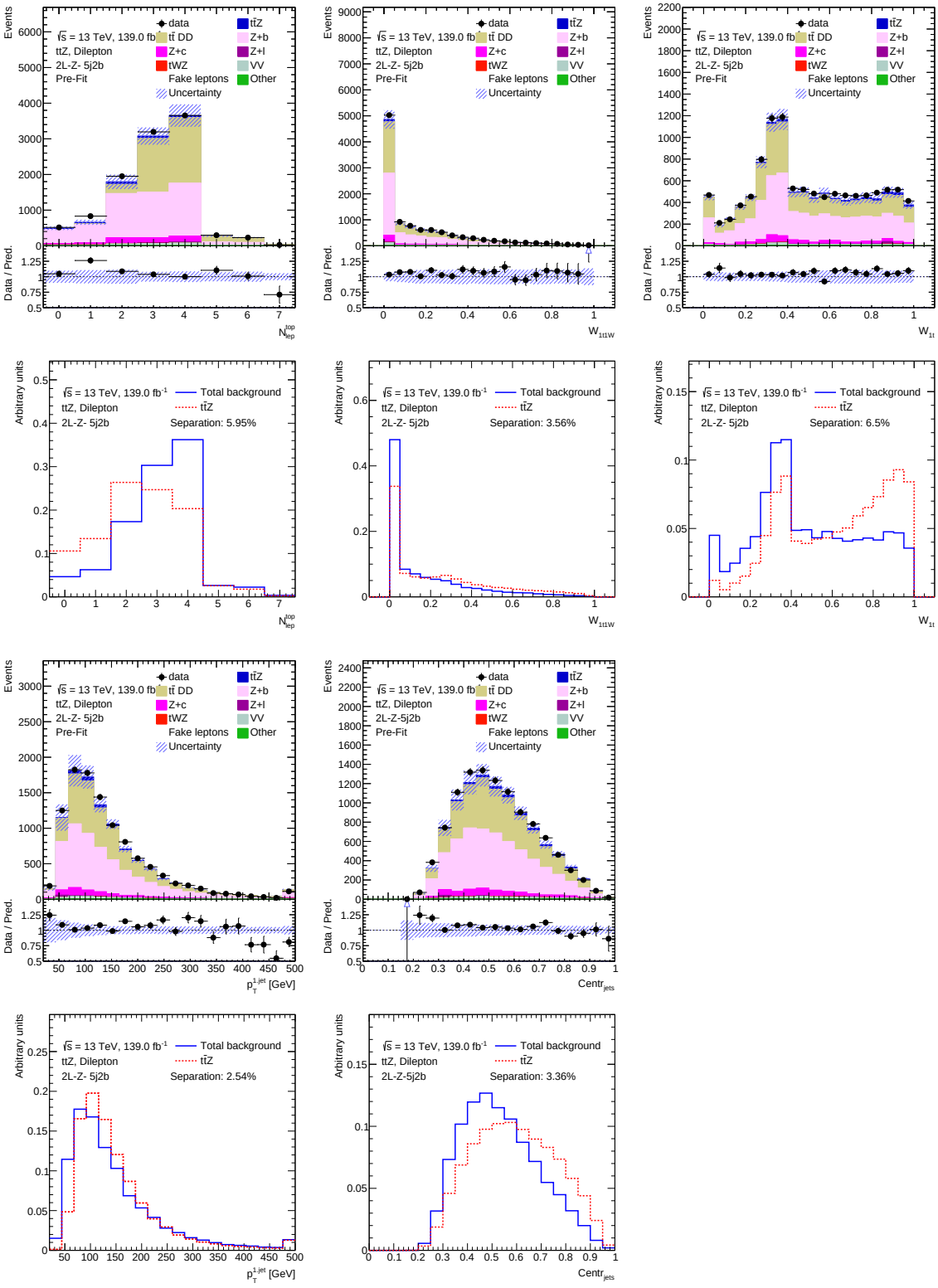


Fig. 61: Neural Network input variables used in 2ℓ-Z-$5j2b$ region.

B.1.2 2ℓ-Z-$6j1b$ region

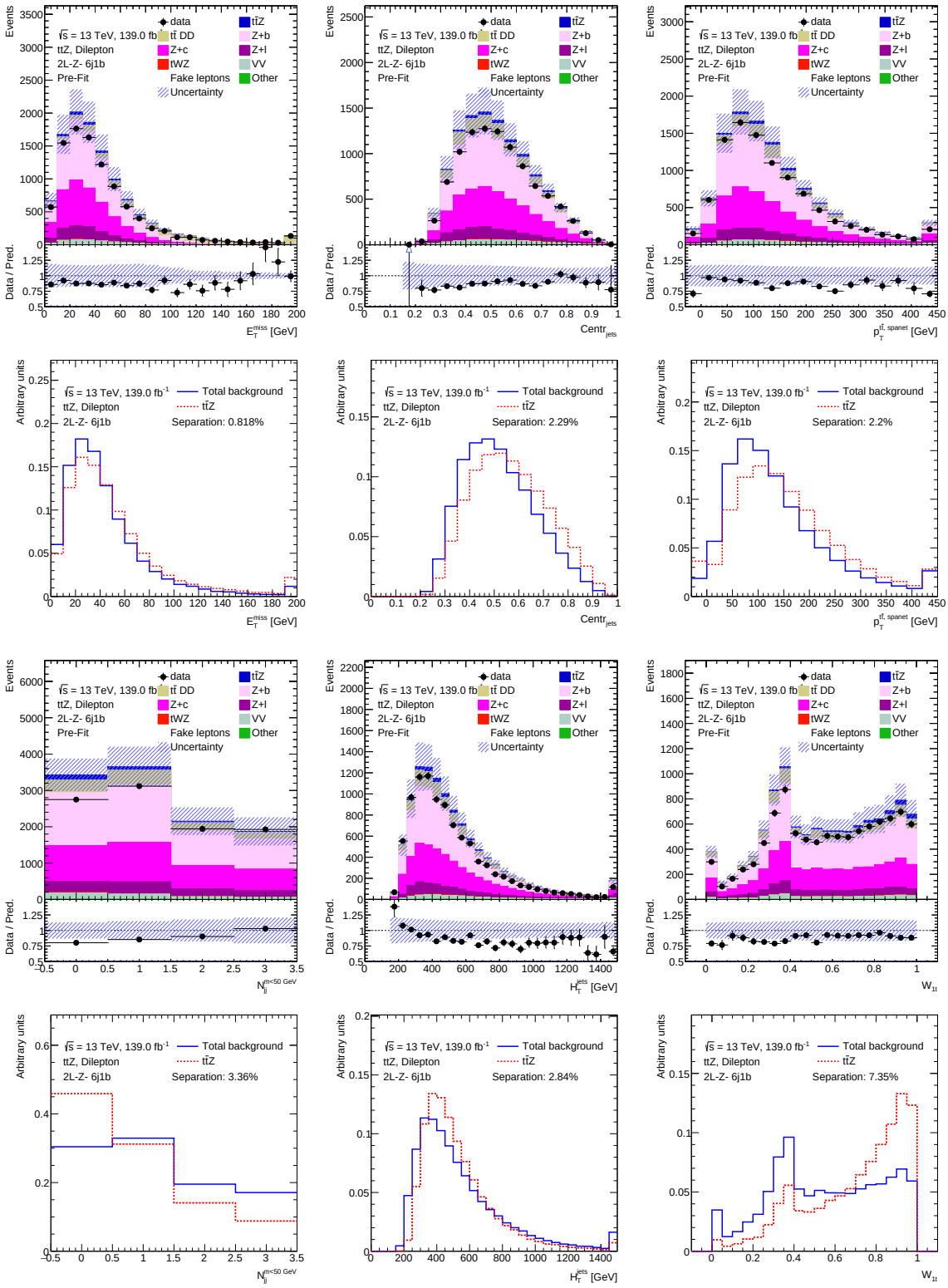


Fig. 62: Neural Network input variables used in 2ℓ-Z-$6j1b$ region.

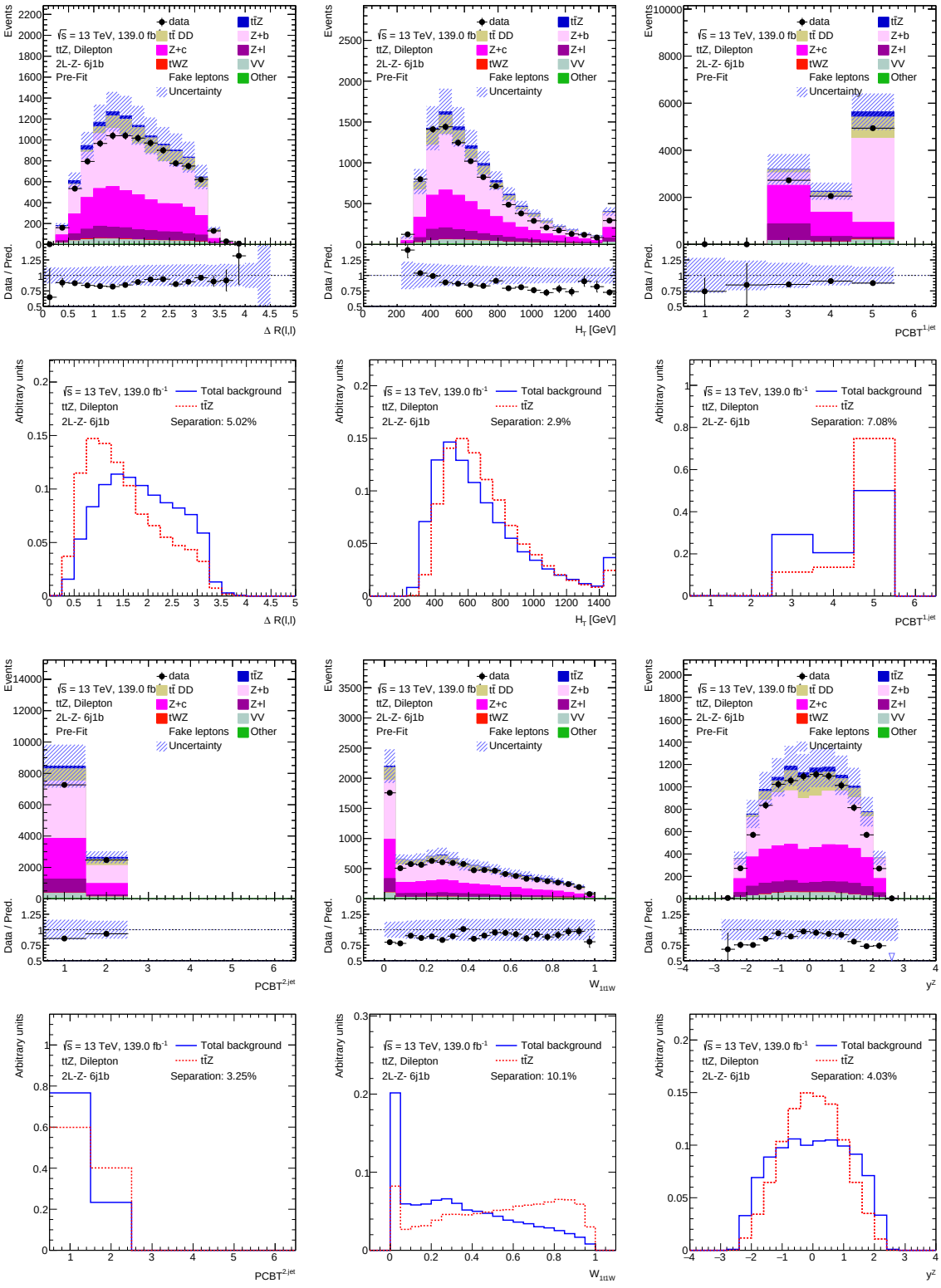


Fig. 63: Neural Network input variables used in 2ℓ-Z-$6j1b$ region.

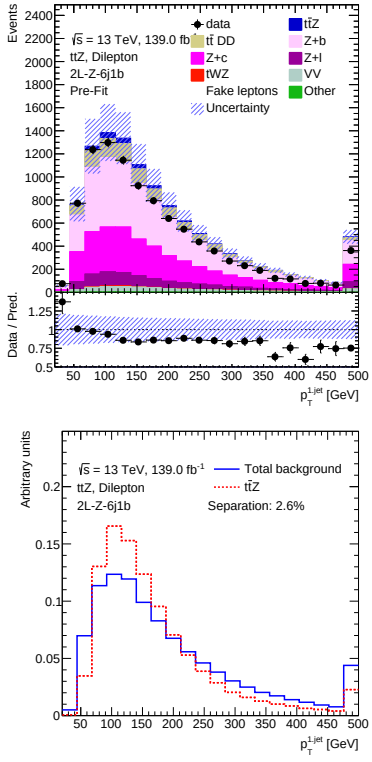


Fig. 64: Neural Network input variables used in 2ℓ-Z-$6j1b$ region.

B.1.3 2ℓ-Z-$6j2b$ region

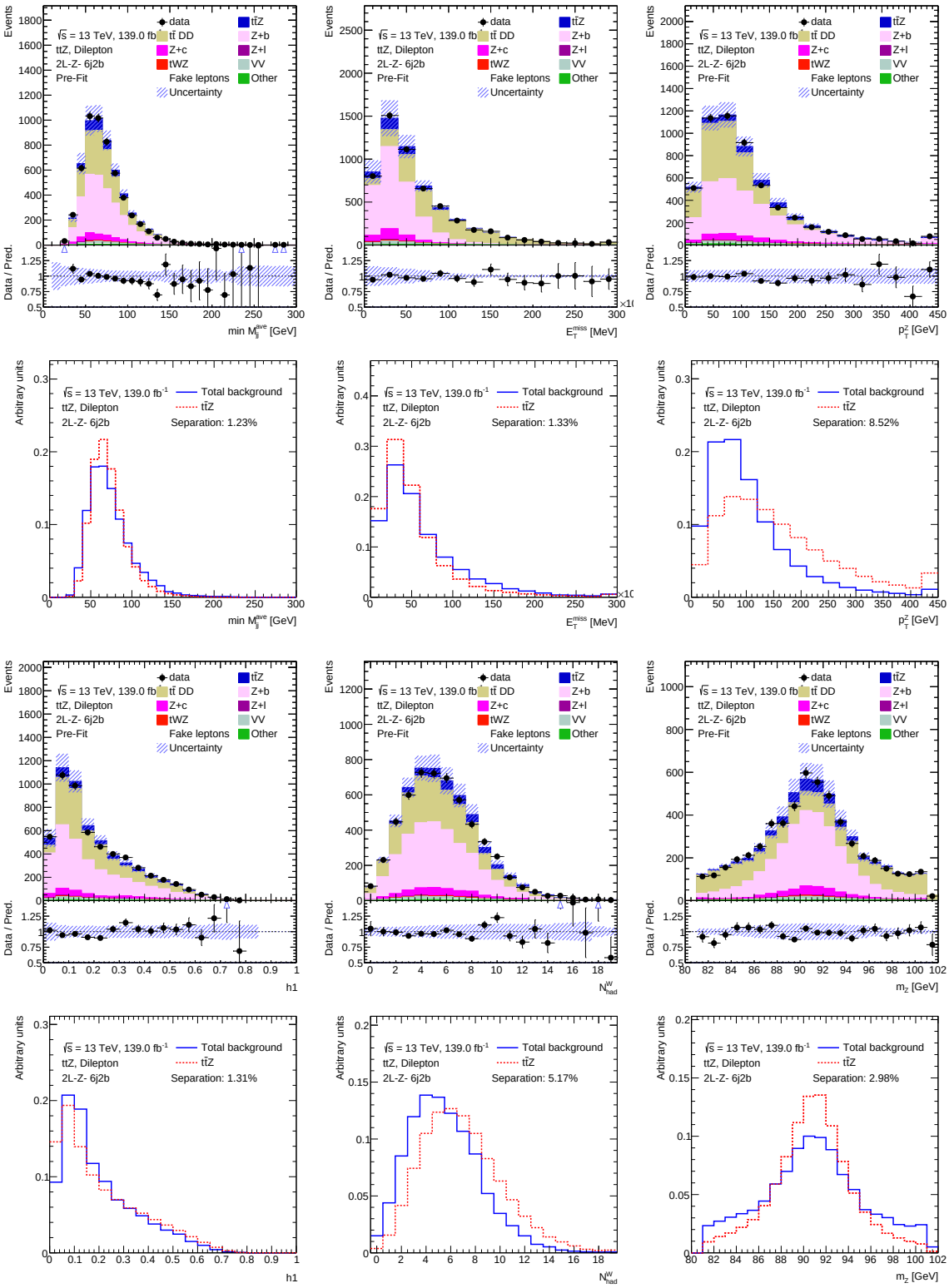


Fig. 65: Neural Network input variables used in 2ℓ-Z-$6j2b$ region.

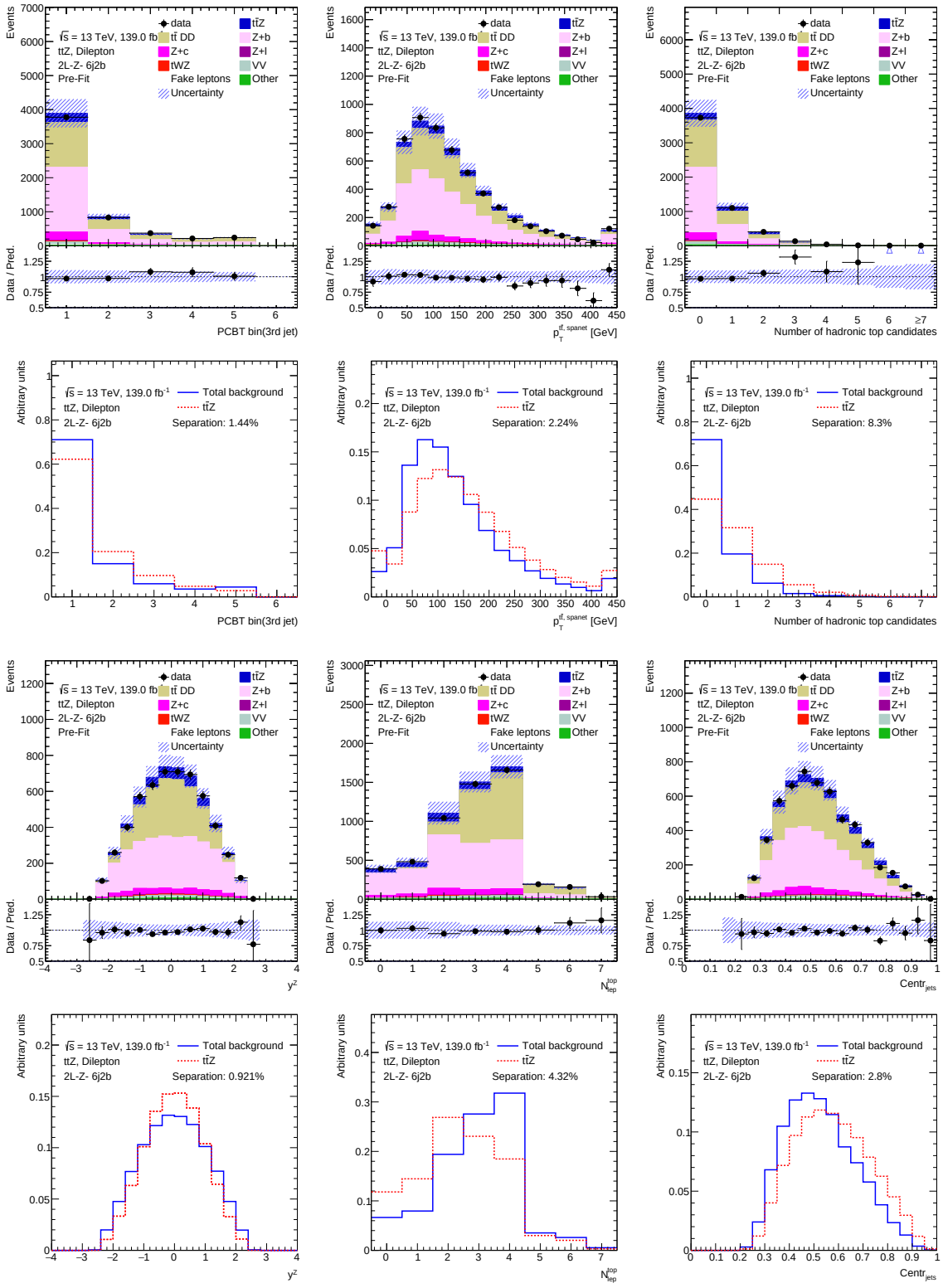


Fig. 66: Neural Network input variables used in 2ℓ-Z-$6j2b$ region.

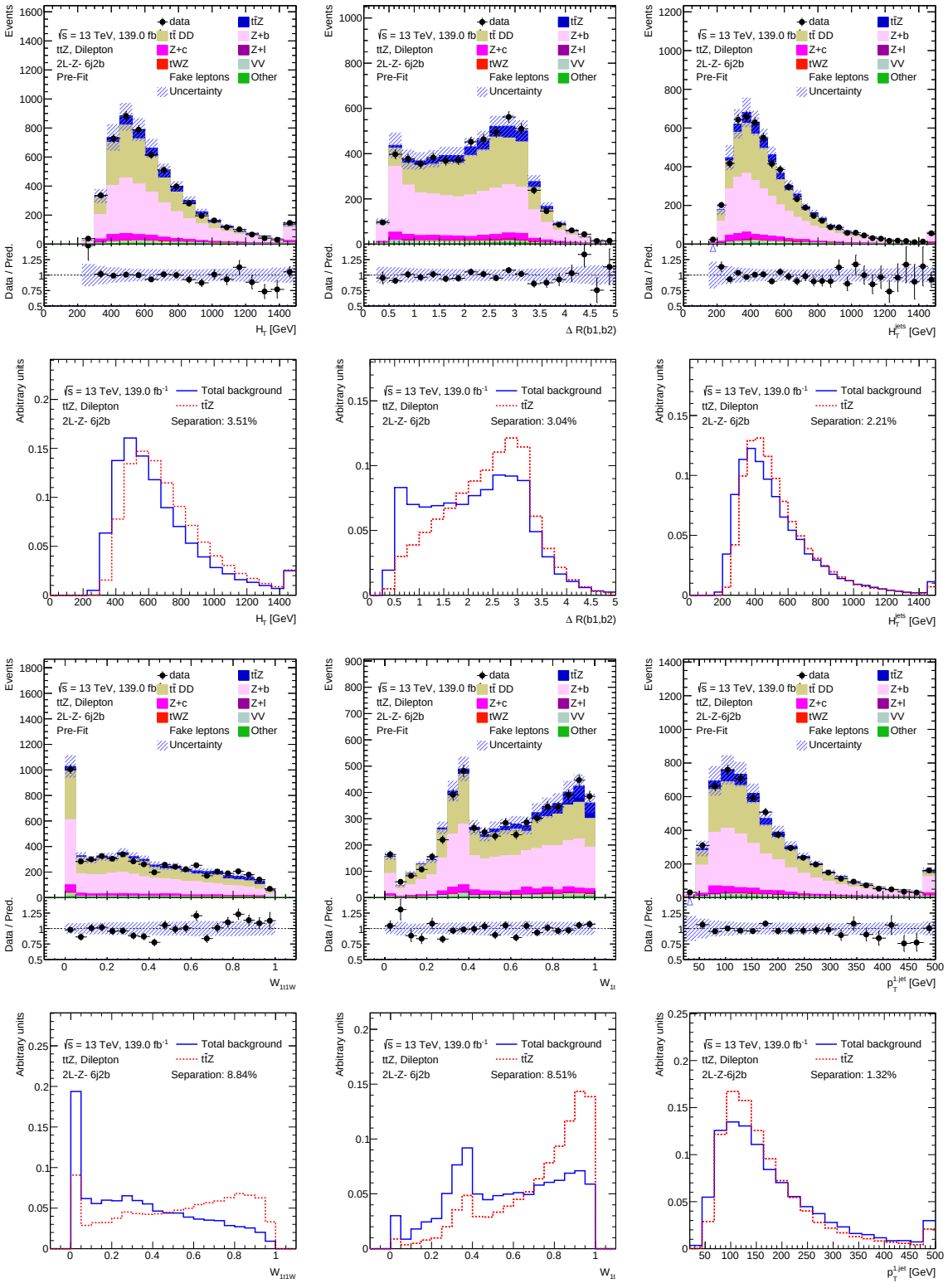


Fig. 67: Neural Network input variables used in 2ℓ-Z-$6j2b$ region.