Measurement of the tZq differential cross-section with the ATLAS detector at the LHC

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

von
Nilima Akolkar
aus
Vadodara, India

Bonn, 06.01.2025

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter/Betreuer:	Prof. Dr. Ian C. Brock
Gutachter:	Prof. Dr. Klaus Desch
Tag der Promotion:	28.02.2025
Erscheinungsjahr:	2025

For my grandparents

Contents

CHAPTER 1

Introduction

Ask the right questions and nature will open the doors to her secrets.

C.V. Raman

The quest to understand the fundamental structure of matter, and ultimately to comprehend the workings of nature, began with the definition of an indivisible entity called an atom. Over the years, advancements in technology, combined with human curiosity, have culminated in the development of the Standard Model (SM) of particle physics. It is a framework that currently stands as our best theory for understanding fundamental particles and forces.

While the Standard Model has successfully explained many phenomena, it fails to address key questions, such as the inclusion of gravity. To probe the SM and delve deeper into the unknown, scientists rely heavily on particle colliders, which recreate high-energy conditions similar to those of the early universe. The Large Hadron Collider (LHC) at CERN, the world's most powerful particle accelerator, has been pivotal in this effort. Situated near Geneva, the LHC accelerates protons to near-light speeds, smashing them together to produce and study rare particles and interactions.

One of the LHC's most significant achievements was the discovery of the Higgs boson in 2012, a fundamental particle that was considered to be the last piece of the SM puzzle. This discovery validated decades of theoretical predictions and opened new avenues for exploring physics beyond the Standard Model. Particles acquire their mass by interacting with the Higgs field. Among them, the top quark, being the heaviest fundamental particle discovered so far, has the strongest interaction with the Higgs field. The top quark's unique properties and behaviour not only serve as an essential probe of the SM but also become a gateway for exploring new physics.

This thesis focuses on one of the rare processes at the LHC, the production of a single top quark and a Z boson. It presents the first differential cross-section measurement of the tZq production using the data collected from the ATLAS detector during 2015 to 2018 at $\sqrt{s} = 13$ TeV. The dataset corresponds to an integrated luminosity of 140.1 fb^{-1}. The tZq production is interesting to study because it allows to probe the coupling of fermion-boson and boson-boson, simultaneously.

The top quark and Z boson, both being heavy particles, decay before being detected. Among the various decay mode combinations, this thesis focuses on the decay mode where both the particles decay into leptons. Due to the presence of three charged leptons in the final state, this channel is called

trilepton decay channel. It is a relatively clean decay channel compared to other modes due to the fact that reconstructing charged leptons is easier and less background-prone compared to reconstructing quarks.

Measuring differential cross-sections can provide deeper insights into the process. Furthermore, any possible new physics might alter the kinematic distributions of particles which can be reflected in the differential cross-section. In this thesis, this measurement is performed using the method of unfolding. In general, any apparatus used for measurement brings with it an extent of reliability on its results. In other words, no apparatus is ideal. The collision data recorded by the detector is distorted due to certain effects and these distortions are undesirable for measurements. Unfolding is a mathematical tool to correct for detector distortions. This thesis presents the implementation of profile likelihood unfolding (PLU) to measure the differential cross-section. Before using PLU for real detector data, the method is validated through a series of statistical tests. The differential cross-sections are presented as a function of a number of observables that probe the kinematics of the tZq system and/or are sensitive to Effective Field Theory (EFT) effects. Moreover, absolute and normalised differential cross-sections are calculated at parton and particle level fiducial volumes, and are compared with theoretical predictions.

This thesis is organised as follows. An overview of the theory of particle physics and some widely used experimental concepts are presented in Chapter 2. It also discusses the top quark and its related processes. Chapter 3 describes particle detection and reconstruction in the context of the ATLAS experiment situated at the LHC. In Chapter 4, the tZq process, which is our signal, is explained along with the procedure of signal extraction. It also outlines the sources of systematic uncertainties. Chapter 5 introduces the concept of unfolding. Moreover, it describes in detail profile likelihood unfolding and its implementation for this analysis. The results of differential cross-section measurement are presented in Chapter 6. The final chapter gives a short summary of this thesis.

CHAPTER 2

Theoretical concepts and experimental basics

This chapter presents an overview of theoretical and experimental concepts widely used in particle physics and are relevant for understanding this thesis. The Standard Model of particle physics, which is by far the best theory explaining fundamental particles and their interactions, is described in Section 2.1. Even though the Standard Model describes and predicts most of the phenomena occurring in nature, it still has some drawbacks which are discussed briefly in Section 2.2. Important concepts required to understand physics manifested at hadron colliders are also discussed. Among the fundamental particles, the top quark is one of main particles studied in this thesis. A detailed description of a top quark and its associated processes are discussed in the end of this chapter.

2.1 Standard Model of particle physics

In the 19th century, John Dalton postulated that matter is made up of small indivisible pieces called atoms. Since then, technological advancements and human curiosity have empowered us to explore various phenomena around us in greater detail. Eventually, our understanding of nature evolved, leading to the development of the Standard Model of particle physics [1], that comprehensively explains the fundamental structure of matter. The Standard Model (SM) of particle physics is a theory that describes most of the phenomena occurring in nature. It is based on fundamental particles and their interactions being governed by Quantum Field Theories (QFTs). The SM has precisely predicted the existence of various particles and their properties. Testing the SM and its predictions plays a crucial role in deciding on the physics program of particle physics experiments.

The particle content of the Standard Model is divided into fermions with half-integer spins and bosons with integer spins. The fermions are further divided into leptons and quarks as shown in Fig. 2.1. Another classification of fermions is into generations. The first generation includes u, d, e^- and ν_e which constitute the matter around us. The second and third generation particles are high energy *siblings* of the first generation particles. These are observed at high energies such as at colliders. The SM also includes anti-particles which are clones of particles with opposite quantum numbers.

The fermions interact with each other by exchanging bosons which are also called *force-carrier* particles. The photon (γ), being massless and electrically neutral, is the messenger of the electromagnetic (EM) force, experienced only by charged particles. The underlying QFT is called Quantum Electrodynamices (QED) based on the U(1) symmetry group. The commonly known electrostatic

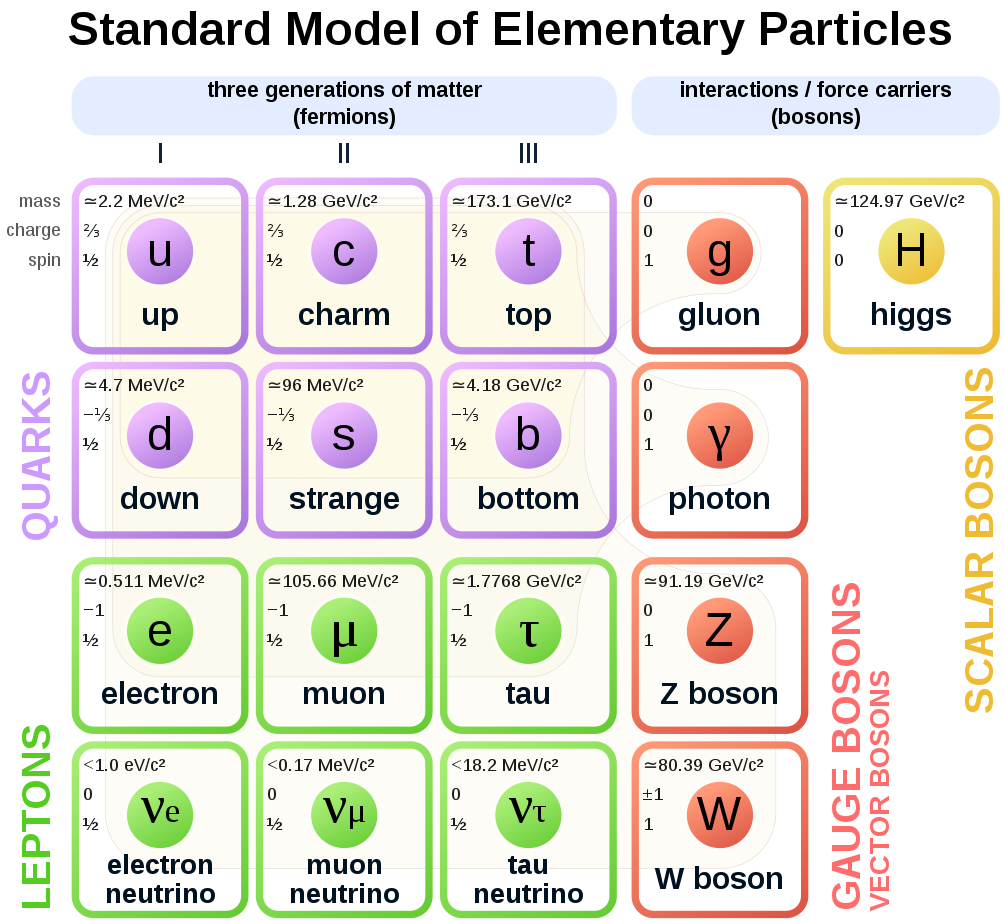


Figure 2.1: Overview of the particles in the Standard Model along with their properties including mass, spin and charge are shown. Particles shown in lavender and green are fermions while the ones shown in red are gauge bosons. The three generations are also highlighted by roman letters. Anti-particles are not shown [2].

attraction between charged particles is the low-energy manifestation of QED. Among the SM, all fermions except neutrinos are sensitive to the EM force. The strength of the EM force is expressed by its dimensionless coupling constant,

$$\alpha \sim \frac{1}{137} \tag{2.1}$$

The strong interaction, mediated by massless gluons, is experienced by particles carrying the so-called *colour* charge. The physics behind the strong interaction is explained by Quantum Chromodynamics (QCD). The quarks, which are the only fermions carrying colour charge, can interact via the strong interaction. In QCD, the gluons themselves also carry colour charge.

The weak force carriers are the vector bosons, $W^{\pm}$ and Z, which unlike γ and gluons, are massive and charged in case of the $W^{\pm}$ boson. The Z boson is electrically neutral. The weak interaction manifests itself in phenomena such as β-decay and fusion processes inside the sun. All the SM particles, including the neutrinos, are sensitive to the weak force. The interaction mediated by the $W^{\pm}$ and Z is called charged-current weak interaction and neutral-current weak interaction, respectively. The famous Wu experiment [3] proved that the charged-current weak interaction violates parity. The parity violating nature of the weak interaction suggests that the interaction configuration must be different from that of QED and QCD. Studies showed that the weak interaction is described using a $V - A$ vertex [4] and this fact dictates that only left-handed chiral particle states and right-handed chiral antiparticle states can participate in charged-current weak interaction.

The theory unifying QED and weak interaction is called the Electroweak theory explained in Section 2.1.3. The last piece of the SM puzzle, which is the last discovered fundamental particle, is the Higgs boson. It is a spin-0 boson, unlike the other bosons. All particles acquire their mass through the Higgs mechanism.

2.1.1 Feynman diagrams

As predicted by the SM, there are different possible ways through which two particles can interact. Having a tool to visualise these interactions would greatly aid in understanding the underlying physics. In particle physics, a tool called Feynman diagrams is used for this purpose.

These diagrams are symbolic representations of particle interactions. They make use of straight lines with arrows to show incoming and outgoing particles and anti-particles. Wavy lines are used to show the boson exchanged between them. It also has a hypothetical time axis which demonstrates the evolution of the process with time. The Feynman diagrams are a pictorial representation showing the evolution of a process.

For a better understanding of Feynman diagrams, consider an example of Bhabha scattering [5]. There are two possible Feynman diagrams for this process as shown in Fig. 2.2. The left diagram represents electron-positron scattering and reads as following: an electron and positron enter, a photon is exchanged between them and the two particles exit. This configuration is called a t-channel diagram. The right diagram, called an s-channel diagram, is obtained by twisting the t-channel diagram. It represents electron-positron annihilation where an electron-positron annihilate into a photon which in turn produces an electron-positron pair. The point of interaction between two particles is called a *vertex*. The time axis is horizontal from left to right. It is vital to note the direction of arrows in particle lines. The arrow directions for particles and antiparticles are opposite to each other. In the shown diagram, the incoming electron points in the forward direction, denoting the evolution of the

interaction with time, whereas the incoming positron points in the backward direction. Following the arrow directions illustrates a continuous current flow.

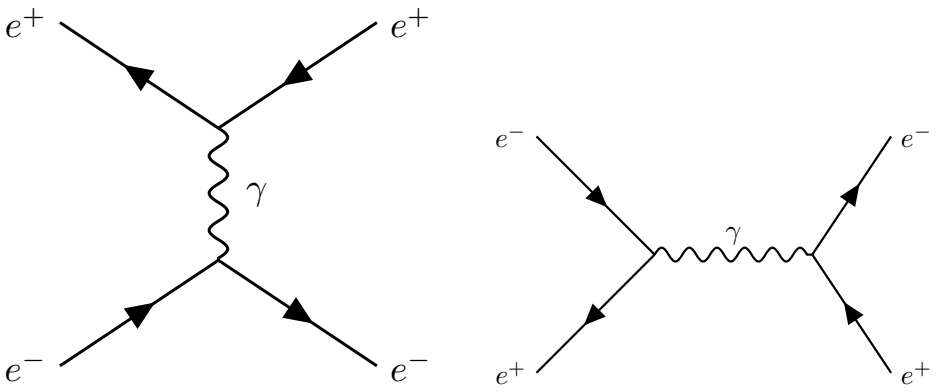


Figure 2.2: The t-channel (left) and s-channel (right) Feynman diagrams showing Bhabha scattering.

The quantitative analysis for a process includes two important steps: accessing the Feynman diagrams to compute the amplitude ($\mathcal{M}$) and together with the phase space, calculating quantities such as decay rate, cross-section and differential cross-section. The Feynman diagrams are analysed through a set of rules called Feynman rules. It is important to note that for each process, there are infinite possible Feynman diagrams which require to be summed to get the accurate final process description. The diagrams shown in Fig. 2.2 are examples in which participating particles are known real particles because they have definite masses. A process can also have intermediate stages where temporary particles are produced which are not observed but simply represent the nature of interaction. Such *virtual* particles can have varying masses. The presence of intermediate states lead to more vertices inside a diagram and therefore give rise to a plethora of Feynman diagrams for a certain process.

In the context of Feynman calculus, each vertex within a diagram contributes a factor equal to the coupling constant of the interaction. Hence, the actual process can be quantitatively described in terms of expansion with respect to the coupling constant. For the expansion to converge, the coupling constant needs to be small. The most basic diagram with the lowest order expansion is called a "tree-level" or a leading-order (LO) diagram. The diagram corresponding to the next order of expansion is called a next-to-leading-order (NLO) diagram. Higher order diagrams with more vertices contribute less owing to the small coupling constant.

2.1.2 Strong Force

Electrons and the nucleus inside an atom are held together by the electromagnetic force. The same force also exists between protons inside the nucleus causing repulsion which can destabilise the nucleus. However, there exists a force which is strong enough to overcome the repulsion and keep the nucleus together. It is called the strong force or the strong nuclear force. The QFT describing the strong force is called Quantum Chromodynamics (QCD) [6] and the underlying symmetry group is SU(3) described by 3×3 matrices. The eight generators of the SU(3) group give rise to eight gluons which are the strong force mediators. The structure of the SU(3) group demands that the wave

function of the strongly interacting particle must be a 3-component vector [7]. This gives rise to a new degree of freedom called "colour", with three states called red, blue and green. Consequently, particles having a non-zero colour charge can feel the strong force. Among the SM particles, only quarks have the colour charge which can be either red, blue or green.

A major differentiating factor between QCD and QED is that the gauge bosons in QCD carry the charge of interaction. In other words, gluons also carry the colour charge which allows them to interact with other gluons as well. As a result of this self-interaction, no coloured object can be found as a free particle in nature. This is hypothesised as colour confinement [8]. To understand colour confinement, consider two quarks being pulled away from each other. Here, a gluon field is created between them which is proportional to the separation. At relatively large distances, the gluon field becomes so strong that at some point, the energy is sufficient to produce new quarks and antiquarks. These newly formed objects combine with the original quarks and antiquarks and continue doing so until they form colourless bound states. This process is called hadronisation. Hence, quarks are always found confined within colourless configurations called hadrons. The possible combinations can be widely categorised into mesons ($q\bar{q}$), baryons (qqq) and antibaryons ($\bar{q}\bar{q}\bar{q}$).

The value of the strong coupling constant, α_S, is relatively larger compared to the coupling constant of QED. As a consequence, contribution of higher order Feynman diagrams increases making calculations difficult. One of the great discoveries in QCD is that the strong coupling constant is in fact not a "constant" but instead the value is dependent on the energy scale of the interaction [9]. The running of α_S means that at low energies, the force between the quarks is stronger (larger α_S) whereas at higher energies, the force becomes weak (smaller α_S). The running of α_S allows us to apply perturbation theory for calculations at high energies. This property of QCD is called asymptotic freedom.

2.1.3 Electroweak theory

In the 1960s, physicists were trying to formulate a gauge theory for weak interactions similar to QED. A theory can be a gauge theory if it has an underlying mathematical symmetry and it is renormalisable. A quantum field theory is renormalisable if the divergences can be absorbed by implementing a finite number of parameters, such as, a coupling constant. Glashow, Salam and Weinberg discovered such a gauge theory by unifying electromagnetic force and the weak force [10].

The electroweak (EW) theory is a unification of QED and the theory of weak interactions. It is described by the symmetry group $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$. The corresponding charges of the electroweak theory are the weak isospin $\mathrm{I}, \mathrm{I_3}$ and the weak hypercharge Y. The weak hypercharge Y determines the interaction under the U(1) transformations. The weak isospin of particles determines their transformation under SU(2) and therefore, it is used to make multiplets of particles. The left-handed leptons (ℓ_L) will form doublets because they transform into each other under the influence of weak force [11]. This is due to the V − A vertex form of the weak interaction. On the other hand, the right-handed particles are singlets (ℓ_R).

$$\ell_R = e^-_R, \mu^-_R, \tau^-_R \tag{2.2}$$

$$\ell_L = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \tag{2.3}$$

The EW model introduces three bosons $W_\mu^{(1,2,3)}$ corresponding to SU(2) and one B_μ corresponding to U(1). The experimentally observed $W^\pm$ are combinations of $W_\mu^{(1)}$ and $W_\mu^{(2)}$ whereas photon (A) and the Z boson are linear combinations of $W_\mu^{(3)}$ and B_μ based on the weak mixing angle (θ_W) as given below:

$$A_\mu = +B_\mu \cos\theta_W + W_\mu^{(3)} \sin\theta_W \tag{2.4}$$

$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^{(3)} \cos\theta_W \tag{2.5}$$

The weak interaction for the quark sector can be explained by creating similar SU(2) doublets(Q).

$$Q = \begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \begin{pmatrix} t \\ b' \end{pmatrix} \tag{2.6}$$

The strength of the weak interactions for quarks is determined experimentally by studying nuclear β-decay. It is observed that the vertices corresponding to different quark flavours have different coupling strengths. The reason for this is given by the Cabibo hypothesis which states that, the flavour eigen states that participate in the weak interactions are a mixture of the mass eigen states [12]. The relation between them is given by the Cabibo-Kobayashi-Maskawa (CKM) matrix [13].

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.7}$$

The values of the CKM matrix elements can be found in [14]. The diagonal of the matrix is close to unity, suggesting that the weak interaction is stronger within the same generation of quarks.

Validation of the EW theory was obtained through experimental observations. The experiments at the Gargamelle bubble chamber [15] in 1973 hinted the evidence of a neutral massive boson responsible for the observed neutrino interactions [16]. In 1983, the Z boson was directly discovered at the Super-Proton Synchrotron at CERN and this discovery verified the electroweak theory [17]. The properties of the Z boson were extensively studied at the Large Electron-Positron (LEP) collider at CERN. The discovery of Z and W bosons are among the crucial tests of the Standard Model.

2.1.4 Higgs mechanism

According to the electroweak theory, weak bosons are required to be massless for the underlying symmetry to be preserved. However, experiments revealed that the weak gauge bosons, W and Z have finite masses [18]. The explanation of this spontaneous symmetry breaking was given by the Brout-Englert-Higgs mechanism [19]. Particles in the SM acquire their mass through the Higgs mechanism. It is a way of spontaneously breaking the SM symmetries by introducing a new field called the Higgs field. The strength of interaction of particles with this field determines how massive the particles will be. The Higgs mechanism implies the existence of a scalar particle, the SM Higgs boson. In a landmark discovery, the Higgs boson was independently discovered by ATLAS and CMS

in 2012 [20, 21]. Since then, the Higgs boson studies are an important part of major collider physics programs.

2.2 Limitations of the SM

The Standard Model is a highly successful theory that has been tested at various collider experiments and almost all the experimental results agree with the predictions, at a high degree of precision. Despite its enormous success, there are still some drawbacks of the SM. They are summarised below:

- Out of the four fundamental forces, only three are explained in the SM. Gravity is not included.
- SM assumes that neutrinos are massless, however, experiments have proved that neutrinos are not massless [22].
- The difference in the mass scale of vector bosons/Higgs boson and the Planck scale is extremely large. This is known as the hierarchy problem [23] and is unexplained by the SM.
- There is no explanation of why is there more matter around us than antimatter.
- The possible existence of dark energy and dark matter is indicated from studies regarding expansion of the universe. There is no explanation in the SM.

2.3 Standard Model Effective Field Theory

Various models have been proposed explaining beyond SM physics. However, no significant experimental evidence in favour of these models, have been found at the LHC. In this scenario, Standard Model Effective Field Theory (SMEFT) provides a model-independent approach to quantify constraints on SM deviations [24]. In SMEFT, additional higher order terms are introduced to the SM, representing new physics beyond the SM at a certain energy scale Λ, higher than the energies possible at the LHC. These terms are built from a set of operators, together with their Wilson coefficients. Observed deviations from the SM can be expressed in terms of these Wilson coefficients.

2.4 Physics at hadron colliders

Colliders are a great place to test the SM and also search for new particles. At hadron colliders, hadrons are accelerated to high energies in order to collide them. The outcomes of these collisions are analysed and compared with theoretical predictions. Some quantities regarding collider physics are discussed in the following section.

Centre-of-mass energy

In a collision between two particles the total centre-of-mass energy is expressed as

$$\sqrt{s} = \sqrt{(\Sigma_{i=1}^{2} E_i)^2 - (\Sigma_{i=1}^{2} \vec{p}_i)^2}, \tag{2.8}$$

where E and $\vec{p}$ are energy and momentum of the two initial state particles. If two beams of the same particle type having the same energy, colliding head-on, then the centre-of-mass energy is $\sqrt{s} = 2E_{\text{beam}}$, neglecting the masses of particles.

Transverse momentum

The axis along which the colliding particles are boosted is referred to as the beam axis or longitudinal axis. The sum of momentum components along the transverse axis (perpendicular to the beam axis), is called transverse momentum,

$$p_{\text{T}} = \sqrt{p_x^2 + p_y^2}, \tag{2.9}$$

where beam is along the z-axis.

Decay rate and branching ratio

A relatively heavy elementary particle often decays into lighter particles through possible decay modes or channels, depending on the conservation laws for quantum numbers and strength of the decay process. The probability per unit time of a particle decaying is called its decay rate (Γ). For N identical particles the change in the number after time dt is given by

$$dN = -\Gamma N dt. \tag{2.10}$$

The lifetime of the particle is the time after which the sample becomes $\frac{1}{e}$ of its original size,

$$\tau = \frac{1}{\Gamma}. \tag{2.11}$$

When multiple decay modes are possible, the total decay rate (Γ_{total}) of the particle is the sum of individual decay rates. In order to learn the dominance of a certain decay mode, we calculate its branching ratio (BR). The branching ratio of a decay mode i is defined as

$$\text{BR} = \frac{\Gamma_i}{\Gamma_{\text{total}}}. \tag{2.12}$$

Parton Distribution Functions (PDFs)

Hadron collisions at high energies give rise to deep inelastic interactions referred to as hard processes. In such cases, the interactions are not between two hadrons but rather between their constituents which are quarks and gluons, collectively known as *partons*. These partons carry a fraction of the total momentum of the hadron. In order to study an interaction, it is important to know the effective energy of the interacting partons and their flavour. This information is encoded in the Parton Distribution Functions (PDFs). It provides quantitative information regarding the distribution of energy carried by the quarks and gluons inside a hadron. A PDF describes the probability of finding a parton of certain flavour i, carrying a momentum fraction x_i at a certain energy scale μ.

With the help of PDFs, we can factorise the cross-section of the process $\sigma_{pp\to X}$ in terms of individual partonic cross-sections $\sigma_{ij\to X}$ as

$$\sigma_{pp\to X} = \Sigma_{i,j=\text{partons}} \int dx_i dx_j \cdot f_1(x_i,\mu^2) f_2(x_j,\mu^2) \sigma_{ij\to X}(s,\mu^2), \tag{2.13}$$

where i,j stand for the partons; f_1 and f_2 are PDFs of the protons defined at energy scale μ; s is the centre of mass energy squared. The PDFs are extracted from combining data collected from various processes [25]. Accurate determination of PDFs is significant in analysis of the Large Hadron Collider (LHC) data because it can affect cross-section calculations. In ATLAS, different sets of PDFs are used depending on the analysis. However, efforts are made towards a global recommendation for all LHC analyses, such as PDF4LHC set [26].

Luminosity and cross-section (σ)

The quantity that measures the ability of a collider to produce particle interactions is called instantaneous luminosity ($\mathcal{L}$). The instantaneous luminosity integrated over the lifetime of collider operation is called integrated luminosity (L). During the operation of the LHC from 2015 to 2018, the delivered integrated luminosity is shown in Fig. 2.3.

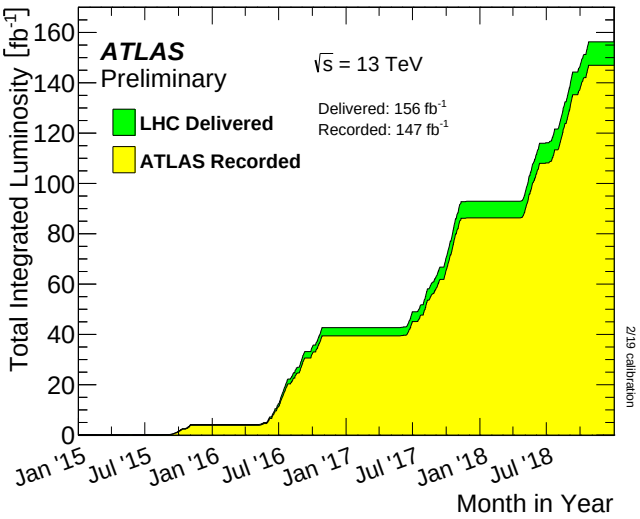


Figure 2.3: Total integrated luminosity delivered by LHC shown in green and that recorded by the ATLAS experiment shown in yellow. The data corresponds to the year 2015 to 2018 [27].

In order to define the event rate for interesting processes, along with luminosity, we require another quantity called the cross-section. At the subatomic scale, the particle interactions are governed by the laws of quantum physics. Therefore, a theory can predict the *probablility* of certain outcomes of collisions. The probability of a certain process to occur is called its cross-section (σ). Finally, the number of events of specific interactions is defined as the product of integrated luminosity and the cross-section (Eq. (2.14)).

$$R = \sigma \cdot \int_{dt} \mathcal{L}(t) \tag{2.14}$$

For a particle collider, beam energies and luminosity are two important figures of merit. High energy allows the production of new heavy particles and high luminosity allows more flux of particles

contributing to large number of collisions.

Differential cross-section

A differential cross-section is a type of cross-section which gives the probability of an interaction with respect to a variable X. In common practice, differential cross-section is defined as $d\sigma/dX$, where dX can be solid angle in scattering experiments or kinematic variables such as p_{T}. The integral over the entire range of X gives the total cross-section.

The differential cross-section offers a deeper insight into the process of interest. For example, in a scattering experiment designed to investigate the internal structure of a target, the scattering profile of the incident particles is analysed. If the scattering rate varies at different solid angles, this variation will be captured in the differential cross-section measurements.

Another advantage of studying differential cross-section is that if there is any new physics, it may manifest itself by altering the kinematic distributions of known SM particles. Measuring the differential cross-sections with respect to kinematic variables and then comparing them to SM predictions is one of the crucial tests of the SM. Any deviations in the comparison can hint towards new physics. A differential analysis is based on the shape of the kinematic distribution and not just the total events in the distribution. Therefore, differential cross-sections can be used for various theory interpretations.

Pileup

The colliding particles are divided into bunches before collisions. Each bunch will give rise to multiple interactions, some of them would be hard interactions resulting from maximum momentum transfer, while others would be soft interactions with minimum momentum transfer. The primary hard scatter collisions, that are usually the main focus of analyses, are contaminated by soft interactions. It is called pileup, defined by the average number of interactions recorded per bunch crossing. Sources of pileup are categorised into in-time and out-of-time pileup. In-time pile up is due to collisions occurring in the same bunch-crossing and out-of-time pile-up is contributed by the collisions from previous or later bunches.

Some of the sub-detectors have sensitivity windows longer than the interval between bunch crossings, this directly affects the recorded number of interactions per bunch. The accurate detection of objects under study becomes difficult due to pile-up events. The higher the luminosity, the more the pileup. The object reconstruction algorithms have dedicated procedures to mitigate pileup in order to accurately reconstruct the interactions of interest. Figure 2.4 shows average number of interactions per bunch crossing with respect to the recorded luminosity during 2015 to 2018 for LHC at CERN.

2.5 Top quark physics

The electroweak theory states that the left-handed particles forming weak isospin doublets can interact weakly. In 1977 the bottom quark was discovered, and it led to the prediction of its weak isospin partner, the top quark t. This prediction was made to keep the electroweak theory internally consistent. After puzzling the scientists for around two decades, the first observation of the top quark was in 1995 at the Tevatron collider at Fermilab by CDF and D0 collaborations [29, 30]. The top quark is the heaviest fundamental particle discovered so far with a mass of (172.57 ± 0.29) GeV [14].

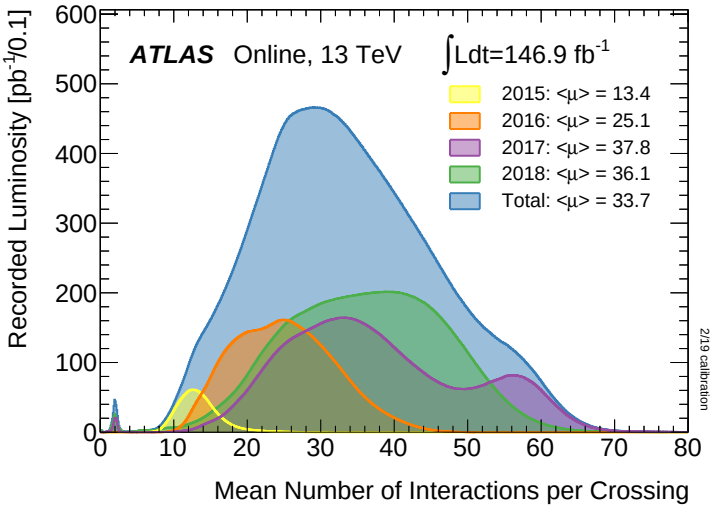


Figure 2.4: Distribution showing average number of interactions versus luminosity per bunch crossing for the 2015 to 2018 collision data at 13 TeV centre-of-mass energy [28].

The top quark being the weak isospin partner of the bottom quark, completes the three generation structure of the SM. Since its discovery, the top quark has been a crucial part of the physics programs at hadron colliders because of its unique properties. The lifetime of the top quark is so short that it decays before it can hadronise. This property gives us a unique opportunity to study a "bare" quark because some of its properties are conserved in the decay process and passed on to its decay products. Regarding the decay products, the top quark decays almost exclusively into a W boson and a b quark.

Study of the top quark is interesting due to various reasons. Its large mass suggests strong coupling to the Higgs field, therefore studying top quarks can give insights into the Higgs sector. In addition, evidence of new physics can be possibly extracted from top quarks because several beyond SM (BSM) physics models suggest that heavy particles might decay into top quarks [31].

After the shutdown of the Tevatron, the LHC remains the only place to produce top quarks in abundance, which is why the LHC is called a top factory. Due to the large centre of mass energy, processes involving top quarks along with heavy bosons are also possible at the LHC. These processes are sensitive to different electroweak and strong couplings. Therefore, precise measurements in this field proves as an important test of the SM.

Production modes

The top quark can be produced either as a top-antitop pair ($t\bar{t}$) in the so-called pair production mode or as a single top quark in association with a light quark or a W boson. The dominant mode is the pair production mode mediated by the strong force. The single top quark production proceeds via the electroweak interaction as discussed below. Studying the different production channels of the t quark offers insights into the QCD sector and also the electroweak sector. Furthermore, measurements of the cross-sections are used to extract important parameters of the SM such as the top quark mass.

Pair production

At LHC energies, gluons inside protons are prone to numerous self interactions and splitting which in turn produces more gluons available for production of heavy particles. The dominant production of top quarks at the LHC is the $t\bar{t}$ pair production [32]. It is initiated either by gluon-gluon fusion, which is the dominant channel, or through quark-antiquark annihilation. The leading-order Feynman diagrams are shown in Fig. 2.5. The inclusive cross-section of the $t\bar{t}$ production measured by ATLAS and CMS at $\sqrt{s}$=13 TeV is shown in Fig. 2.6. The results are found to be in agreement with the SM prediction. The measurements of top pair production are useful for determining PDFs [33].

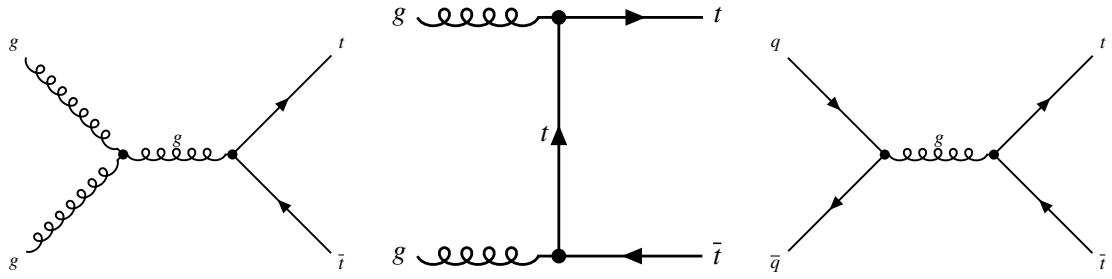


Figure 2.5: Feynman diagrams for $t\bar{t}$ processes at LO in QCD. The gluon-gluon fusion process which is the dominant mode is shown in the leftmost diagram.

Single-top production

There are three separate processes at LO responsible for the production of a single-top quark. Single-top production proceeds only through weak interaction only. Based on the virtuality of the exchanged W boson, the possible modes are t-channel, s-channel and associated Wt-channel as shown in Fig. 2.7. The t-channel process is initiated by a b quark and a spectator quark, producing a single top quark and a light jet. This process has the largest production cross-section at the LHC. The s-channel is characterised by a quark-antiquark pair interacting to produce a single top quark and a b quark. The requirement of an antiquark in the initial state largely suppresses this process [35]. In the associated Wt-channel, a gluon and a bottom quark produce a single top quark and an on-shell W boson. Its final state is same as the $t\bar{t}$ process with one fewer b-jet. The associated Wt process is an important background for other top quark and Higgs related measurements [36].

Due to the electroweak nature of the single top production, it is sensitive to various parameters such as the V_{tb} element of the CKM matrix and the Wtb coupling [37]. The single top production cross-section measurements by ATLAS and CMS are shown in Fig. 2.8.

Rare associated top quark processes

Measurements of rare processes at the LHC are considered to be important tests to validate the SM. Owing to that, the physics program at the LHC is heavily involved in precision measurements related to rare processes, especially those involving the top quarks. The top quark being the heaviest fundamental particle discovered so far, has strong links to the Higgs sector and also to potential BSM physics. Moreover, associated processes involving top quark, probe electroweak couplings of the top

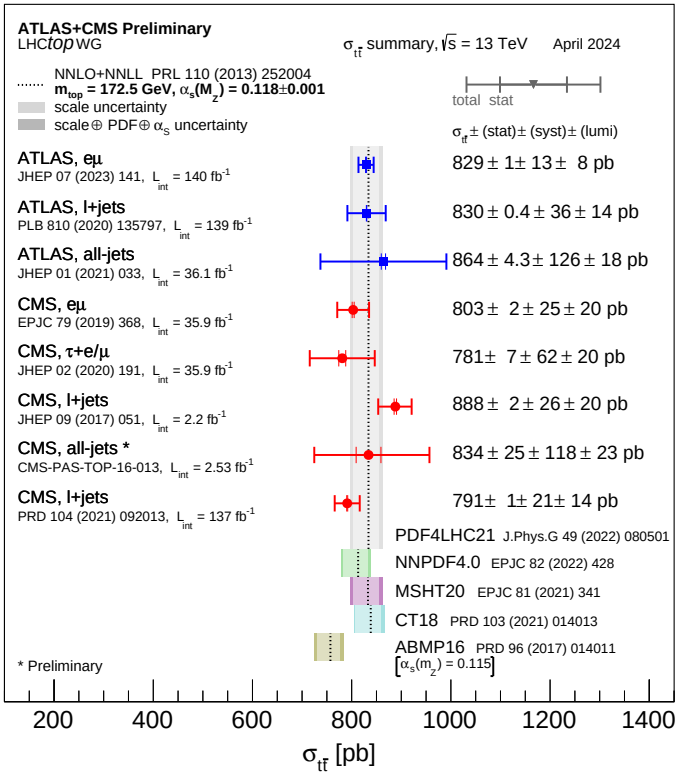


Figure 2.6: Summary of measurements of the $t\bar{t}$ production cross-sections in different analysis channels by ATLAS and CMS at $\sqrt{s}$=13 GeV. The results are compared with NNLO QCD predictions [34].

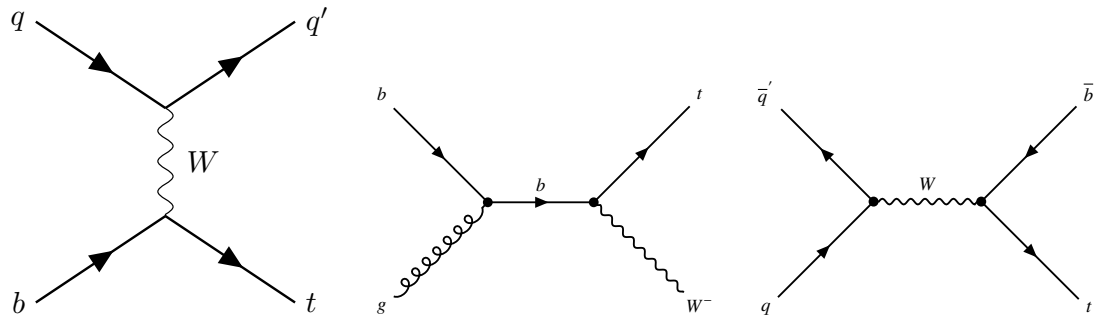


Figure 2.7: Example Feynman diagrams for single top production processes at LO in QCD. The possible production modes namely t-channel (left), associated Wt-channel (middle) and s-channel (right) are shown. The largest cross-section is of the t-channel, followed by the Wt-channel.

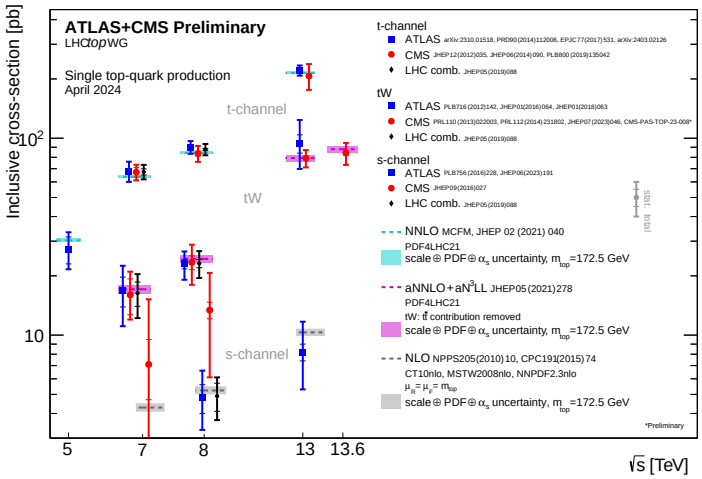


Figure 2.8: Summary plot showing inclusive cross-sections of the single top quark production modes measured at different centre-of-mass energies by ATLAS and CMS [34]. It is seen that the *s*-channel mode is highly suppressed.

quark which are fundamental to the SM. It is therefore relevant to study rare processes involving top quarks. A comparison between the production cross-sections of different top quark related processes measured by ATLAS is shown in Fig. 2.9.

The associated production of a single top quark with a photon, called $tq\gamma$ production, probes the coupling between a top quark and a photon. It is studied using the data collected at 13 TeV centre-of-mass energy, by ATLAS and CMS in the leptonic channel (final states dominated by leptons) due to the higher sensitivity compared to hadronic channels (final states dominated by quarks). The measured cross-section is in agreement with the SM prediction [38].

The evidence for an associated production of a top quark with two heavy bosons (tWZ) is reported by CMS [39]. Studies are performed using the 13 TeV dataset in the leptonic channel and the measured cross-section corresponds to a statistical significance of 3.4 standard deviations.

The associated process of a top quark and a Z boson, called the tZq production, is also one of the rare processes at the LHC. Study of this production is the main focus of this thesis. The tZq production is interesting because it probes two different kinds of couplings through the same interaction. A detailed description is given in Section 4.1. The cross-sections of tZq, tWZ and $tq\gamma$, as measured by ATLAS and CMS, are compared in Fig. 2.10. The tZq process is also useful in studying flavour-changing-neutral-current (FCNC) couplings. The SM forbids interactions of fermions with Z boson, where the flavour of the incoming quark is changed. However, some extension models of SM predict otherwise. A dedicated search for FCNC processes involving a top quark, an up-type quark and Z boson was conducted by ATLAS using the 13 TeV dataset. The goal was to look for events corresponding to either ($gq \rightarrow tZ$) or events where t decayed into Z ($t \rightarrow qZ$). Despite the efforts, no evidence for FCNC processes was found [40].

One of the rarest and heaviest process at the LHC is the production of four tops ($t\bar{t}t\bar{t}$). This process is interesting because it simultaneously probes the coupling of four fermions. In addition, various BSM models predict potential modification in the cross-section, hinting towards new physics. The four tops production was observed at the LHC at 6.1 standard deviation and its cross-section, measured by ATLAS using 140.1 fb^{-1} at 13 TeV centre-of-mass energy, is $(22.5\,^{+6.6}_{+5.5})$ fb [41]. It is found to be

consistent with the SM prediction.

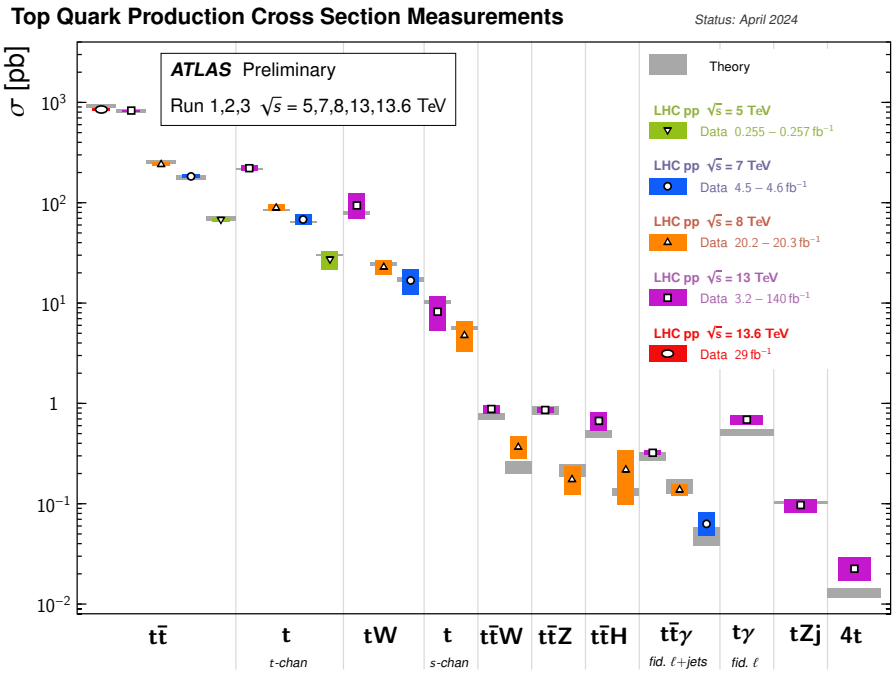


Figure 2.9: Production cross-sections of different processes involving top quarks are shown. It can be seen that the tZq prduction cross-section is smaller than, for instance, the $t\bar{t}$ cross-section [34].

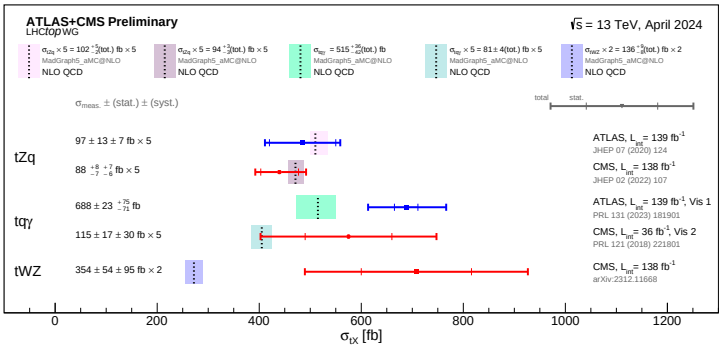


Figure 2.10: Production cross-sections of single top quark associated processes measured by ATLAS and CMS [42].

CHAPTER 3

Particle detection and reconstruction

This chapter describes the particle physics experiment that forms an integral part of this thesis. The data analysed for this study is from the ATLAS experiment described in Section 3.3, situated at the Large Hadron Collider (LHC) which is explained in Section 3.2. Section 3.4 explains in detail particle identification and reconstruction in the ATLAS detector. In addition, the signatures of various objects necessary for this thesis are outlined.

3.1 Particle accelerators and detectors

The matter that we observe around us is made up of the lighter, first generation particles of the Standard Model. To produce heavier particles, a high energy environment is necessary which is created inside a particle accelerator like the LHC. These are devices that drive charged particles to high speeds (close to the speed of light) and then collide them, either with each other or with a target. The results of these collisions enable physicists to study the fundamental particles and the forces governing them.

A typical particle accelerator has two main components: particle source and electromagnetic field. The charged particles traveling in vacuum inside the beam pipe, experience the force from the electric field and are accelerated in the direction of the field. Moreover, the magnetic field bends their path so that the particles travel along the accelerator arcs. In modern machines, metallic chambers called as the RF cavities are installed with the purpose of supplying the required electric field. The field inside the cavity is made to oscillate and the frequency of oscillation is tuned with the arrival timings of the incoming particles. A particle which arrives at the exact time when the field is the highest (resonant frequency), will receive the maximum push. At every revolution, particles arriving sooner or later will be accelerated or decelerated depending on the phase of the electric field.

A direct consequence of applying an oscillating electric field is that the particles will be packed in groups called "bunches", rather than a continuous flow. The acceleration will continue until all the particles reach the desired energy. Once that is achieved, the particles are ready to be collided.

The collisions of these high energy particles give rise to an arena of processes such as production of heavy particles, decay, elastic and inelastic scattering, etc. Various outcomes of collisions are studied with the help of particle detectors. These detectors, like the ATLAS, identify and measure the particles produced after the collisions. It can be said that particle accelerators and detectors are a playground for the Standard Model. For this thesis, data from the ATLAS detector which is at the

Large Hadron Collider (LHC) have been analysed.

3.2 The Large Hadron Collider

The LHC is the world's largest particle accelerator built by the European Organization for Nuclear Research (CERN). It is a proton-proton collider installed in the 26.7 km long tunnel near Geneva across the Swiss-French border [43]. Inside the LHC, two high-energy proton beams travel in opposite directions before they are collided. The particles are guided along the accelerator ring by strong magnetic field created by superconducting magnets. The beams are brought into collision at four locations around the LHC ring. At these locations, the particle detectors are placed aimed for analysing the results of the collisions.

The protons pass through a series of smaller accelerators before entering the LHC ring. These accelerators are collectively known as the CERN accelerator complex as shown in Fig. 3.1. Initially, protons are extracted from hydrogen atoms by applying an electric field and injected into LINAC2, the first accelerator in the series. It accelerates protons to an energy of 50 MeV and feed them to the Proton Synchrotron Booster (PSB), where they achieve an energy of 1.4 GeV. Afterwards the protons enter Proton Synchrotron (PS). It boosts the protons to 25 GeV and injects them into the Super Proton Synchrotron (SPS), which increases the proton energy to 450 GeV. Finally, the protons are inserted into two opposing beams of the LHC, the final machine of the accelerator series. Here, each beam attains an energy of 6.5 TeV.

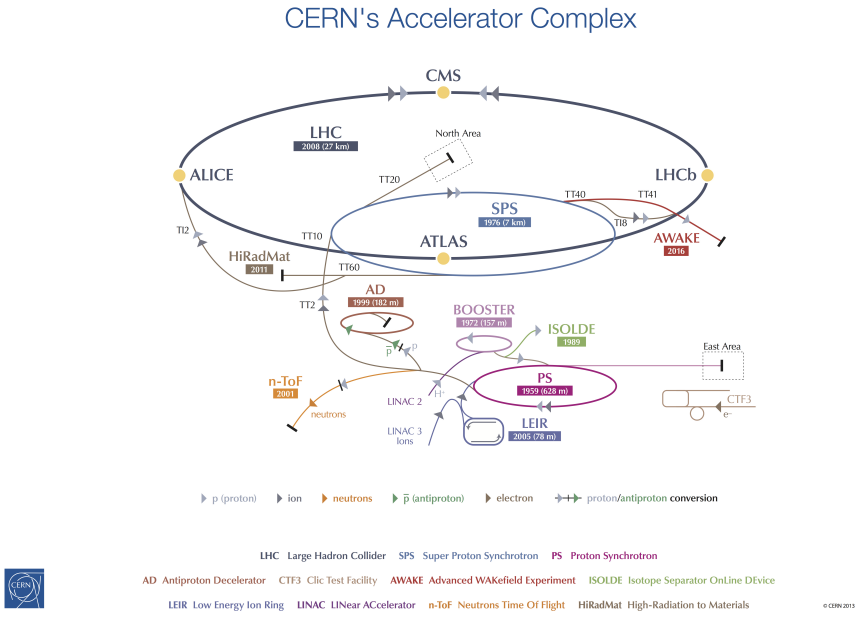


Figure 3.1: Sketch of the LHC ring along with the machines in the CERN accelerator complex [44].

At the LHC, the superconducting dipole magnets have the important responsibility of directing the proton beams along the circular structure. To achieve the desired centre of mass energy, given the circumference of the LHC ring and charge of proton, approximately 8 T magnetic field is required.

This is outside the limit of conventional magnets and that is why superconducting magnets are used at the LHC. Since the available area in the LHC tunnel is limited, a single magnet system is shared by both the beam pipes. The superconducting magnets are immersed in a superfluid helium bath set at a low temperature of 1.9 K in order to achieve superconductivity [45]. When charged particles travel in curved paths, they emit electromagnetic radiation called the synchrotron radiation. The magnet system is shielded from this synchrotron radiation by using beam screens [46] so that low temperature and ultimately superconductivity is preserved.

At the LHC 2808 bunches of protons are injected which are 25 ns apart. This results into 1.2×10^{11} protons per bunch giving rise to different particle interactions. The outcomes hold interesting physics which is studied with the help of particle detectors. Out of the four detectors, two are general purpose detectors namely ATLAS (**A** **T**oroidal **L**HC **A**pparatu**S**) [47] and CMS (**C**ompact **M**uon **S**olenoid) [48]. The main aim of the LHCb experiment [49] is to study B hadrons and their decays whereas ALICE (**A** **L**arge **I**on **C**ollider **E**xperiment) [50] analyses the Pb-Pb collisions at the LHC. This analysis uses data from the ATLAS detector which is described in the next section. A sketch of the LHC and its experiments is shown in Fig. 3.2.

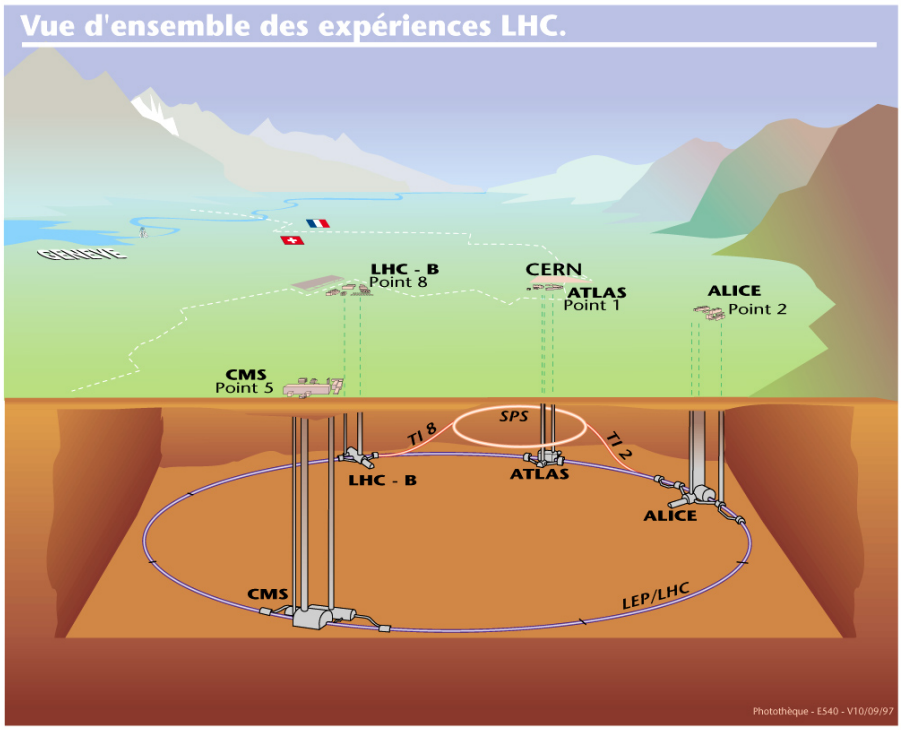


Figure 3.2: Illustration of the LHC ring and positions of the four experiments along with position of CERN [51].

3.3 The ATLAS detector

The ATLAS detector is one of the general purpose detectors at the LHC. It is an assembly of subdetectors constructed around the beam pipe in an onion-like structure as shown in Fig. 3.3. The subdetector

closest to the beam is called the Inner Detector (ID) and it is designed to reconstruct trajectories of charged particles and vertices of interactions. Vertices are locations where the collision/scattering took place. Surrounding the ID, are the calorimeters. The calorimeters measure the energy deposited by particles traversing through them. Beyond the calorimeters, Muon Spectrometers are positioned to mainly detect muons. The ATLAS magnet system generates the strong magnetic field necessary to bend the paths of incoming particles.

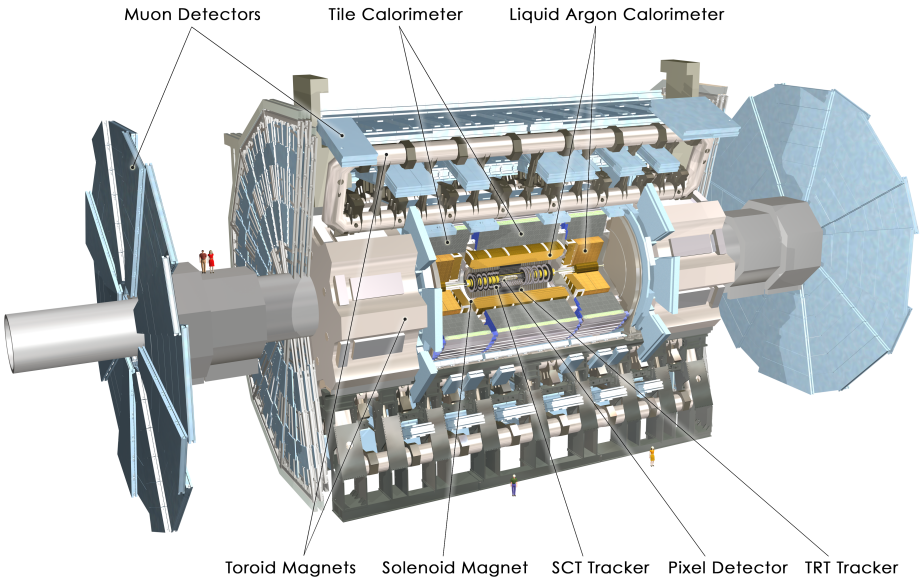


Figure 3.3: A graphic of the ATLAS detector, showing its different sub detectors and its scale. The scale is shown by adding two humans on the left side [52].

Coordinate system

ATLAS uses a spherical coordinate system in which the point where collisions occur, called the Interaction Point (IP), is treated as the origin. The z-axis is along the beam direction, x-axis is directed to the centre of the LHC ring and the y-axis points upwards. The azimuthal angle (ϕ) is measured from the x-axis (around the beam pipe) and the polar angle (θ) is measured from the z-axis (from the beam pipe). A sketch of the coordinate system is shown in Fig. 3.4. It is crucial to note that the momentum of colliding protons is along the beam direction (z-axis) which means that there is zero momentum in the transverse plane before collision. A quantity derived from this fact is known as rapidity defined as follows:

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \tag{3.1}$$

If a collided particle is highly relativistic and launched almost perpendicular to the z-axis, its p_z will be close to zero leading to zero rapidity. Another scenario would be that the particle is moving along the z-axis and in that case the rapidity will be $\pm\infty$. For large collider experiments precise calculation of E and p_z is difficult. Therefore, a quantity called pseudorapidity is used. It is conceptually similar

to rapidity and is defined as,

$$\eta = -\ln \tan\frac{\theta}{2} \tag{3.2}$$

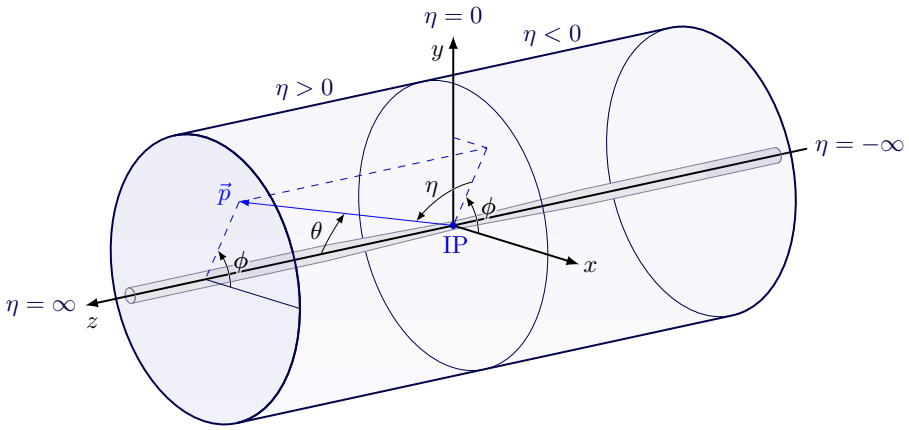


Figure 3.4: Sketch of the ATLAS coordinate system. The origin is the Interaction point (IP) and the coordinates for a particle with momentum $\vec{p}$ are shown. Based on [53].

The Inner Detector

The closest sub detector system to the beam pipe is the Inner Detector (ID) [54, 55]. It has a cylindrical barrel region covering $|\eta| < 1$ and two endcap regions covering $1 \leq \eta \leq 2.5$ [56]. The ID comprises the Pixel Detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT) immersed in a 2 T solenoidal magnetic field. Data recorded by these components are used to measure the charge, direction and momentum of charged particles. The schematic showing the structure of the ID is given in Fig. 3.5.

The Pixel Detector resides along the radii of 33-150 mm [58]. Modules of silicon pixels are arranged in four layers in the barrel region and two layers in the endcap region. It provides a three-dimensional space-point reading that allows for the reconstruction of primary and secondary vertices. The high granularity of these pixels provide high spatial resolution that is beneficial for pattern recognition. The next subcomponent is the Semiconductor Tracker (SCT) which consists of silicon microstrip detectors arranged in four barrel layers and two endcap layers. The radii range of the SCT is 299-560 mm [58]. The outermost subcomponent is the Transition Radiation Tracker (TRT) which is made up of gaseous straw tube detectors. It spans across 563-1066 mm radii [58]. The straws are filled with xenon and carbon dioxide at atmospheric pressure and each straw has a gold-plated tungsten wire along the longitudinal axis [59]. The transition material is a low-density foam surrounding the straws. The data recorded by the TRT is utilised for reconstructing tracks of charged particles. A combination of measurements given by the subcomponents help get a high-resolution vertex and momentum measurement of incoming particles.

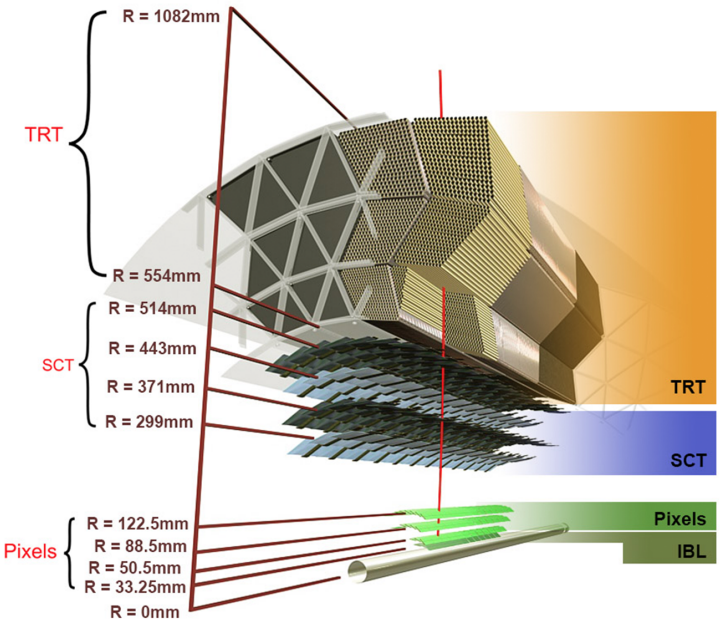


Figure 3.5: Schematic of the structure inside the Inner Detector. It shows the highly granular silicon pixels in the Pixel detector, silicon strips in the SCT and straw tubes in the TRT. Moreover, the distances are shown from the beam pipe in terms of radii [57].

Calorimeters

Calorimeters are detectors that are designed to measure the energy of incident particles. The particles incident particles interact with the active material inside the calorimeters and produce secondary particles. Almost all of their energy is lost inside the calorimeter through shower production. The signals generated by these secondary particles are read-out and recorded for accurate energy measurements. A wide area coverage and sufficient depth are important parameters for calorimeter design. The calorimeters at ATLAS are designed such that they cover a full ϕ-range and $|\eta| < 4.9$. A sliced image of the ATLAS calorimeter is shown in Fig. 3.6.

Electromagnetic Calorimeter (ECAL)

The Electromagnetic Calorimeter (ECAL) is a lead-liquid-argon calorimeter split into a barrel component ($|\eta| < 1.475$) and two endcap components ($1.375 < |\eta| < 3.2$) [47]. Over the precision physics η region, which is essentially the region coinciding with the inner detector, the detectors are highly granular which allows accurate measurement of electron energies. The ECAL is a sampling detector in which liquid argon (active material) and lead absorbers are placed alternatively in an accordion geometry. This geometry allows for a full ϕ-coverage.

The incident particles encounter the absorber, producing charged particles that ionise the argon atoms. The resulting electrons travel towards the electrodes and a signal is recorded. The depth of the ECAL is based on the expected shower length and is estimated by the radiation length. The radiation length X_0 is a characteristic of a material which is defined as the length at which the energy of a

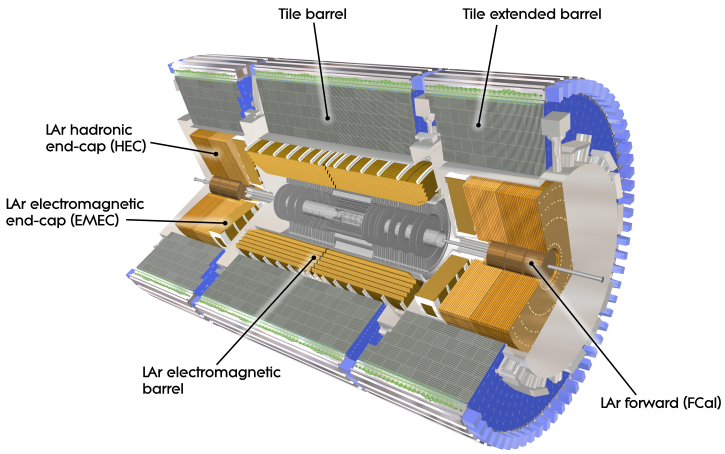


Figure 3.6: A sliced view of the ATLAS calorimeter system [60].

particle traversing the material, is reduced by a factor of $1/e$. The thickness of ECAL is $> 22X_0$ in the barrel and $> 26X_0$ in the endcap regions [61].

Hadronic Calorimeter (HCAL)

The Hadronic Calorimeter (HCAL) consists of the tile calorimeter, the Hadronic Endcap calorimeter (HEC) and the Forward Calorimeter (FCal). The tile calorimeter [62] is a type of sampling detector, covering a range of $|\eta| < 1.7$. The active material consists of scintillating tiles, absorbers are made of steel and wavelength shifting fibers are used for readout. Readout cells are defined by grouping certain number of fibers onto the same photomultiplier tube (PMT).

Overlapping some η region of the tile calorimeter, there is the HEC which occupies $1.5 < |\eta| < 3.2$. It is a copper-liquid-argon sampling calorimeter consisting of two independent wheels per endcap situated right behind the endcap electromagnetic calorimeters. The detector responsible for detecting large $|\eta|$ objects is the FCal which has a coverage of $3.1 < \eta < 4.9$. The active material (liquid-argon) is placed in smaller gaps between the electrodes compared to the gaps in the ECAL. The reason is avoiding ion build-up problems caused by the heavy particle flux encountered by the FCal. Forward calorimetery is essential to increase the detection of high momentum jets that would otherwise escape detection at high η [63]. It allows a hermetic calorimetry that leads to efficient missing E_T measurements [64].

Muon spectrometers

The muon spectrometers play a significant role in accurate measurements of high energy muons. Due to the high mass of muons, they can travel a considerable distance before being detected in the MS. The structure is made up of three superconducting toroid magnets, trigger system and precision tracking chambers [65]. The magnets are situated in the barrel and two endcap regions providing the magnetic field required for muon tracking.

The muon trigger system hunts for muon candidates from the incident particles. It uses the Resistive

Plate Chambers (RPCs) which are placed in the barrel region and the Thin Gap Chambers (TGCs), placed in the endcap region. Different types of detectors are used in barrel and endcap because the muon flux encountered by these regions is different. They select muon candidates with high p_{T} by reconstructing tracks that point directly towards the interaction point. The precision tracking chambers are used to measure track coordinates of muons using two types of chambers: Monitored Drift Tubes (MDT) and Cathode Strip Chambers (CSCs). In the MDTs, each chamber contains two layers, each formed by three or four more layers of drift tubes. The drift tubes are made up of aluminum and are filled with Ar-CO_2 gas mixture.

Magnet system

The ATLAS detector is exposed to large flux of decay products during its operation. A strong magnetic field is required to bend the paths of charged particles for momentum measurements. For this purpose, a magnet system is deployed at ATLAS which includes a central solenoid, barrel toroidal and two end cap toroidal magnets. The central solenoid magnet, operating with a current of 7.6 kA, is wrapped around the beam axis and provides a 2 T magnetic field to the Inner Detector. The barrel and end cap toroidal magnets operate at 20.5 kA and provide a magnetic field ranging from 0.5 to 1 T in the muon chambers. The barrel toroidal magnet is built from eight coils and kept equi-spaced to form a toroid-shaped magnet. The end cap magnets are placed inside the barrel toroidal magnet at both ends of the central solenoid magnet.

Trigger and Data Acquisition System

At the LHC numerous interactions are occurring simultaneously. However, not all events contain interesting characteristics. Therefore, ATLAS uses a dedicated trigger system to select only those events that are interesting from a physics point of view. The selection and storage of events is performed by the Trigger and Data Acquisition system (TDAQ). It has two components, the data acquisition system which reads data from the detector subsystems and the trigger system which decides which data to retain.

ATLAS uses a two-staged trigger system formed by the level-1 (L1) trigger and the high-level trigger (HLT) [66]. The L1 trigger is a hardware based trigger that selects meaningful signals from the detector components and passes on to the HLT which then performs sophisticated algorithms to only retain interesting physics objects which are further stored for offline analysis. The event rate drops from 40 MHz to 100 kHz after the L1 trigger. Moreover, it further drops to 1 kHz after the HLT.

The L1 trigger is composed of L1Calo, L1Muon, L1Topo and Central Trigger Processors (CTPs) [67]. The L1Calo processes signals from the calorimeters whereas L1Muon processes signals from the RPCs and TGCs in the muon spectrometers. In order to avoid particles not originating from the interaction point, L1Muon follows coincidence requirements based on the inner detector and the calorimeter signals. L1Topo which was an upgrade introduced in Run-2 [68], calculates event topological quantities of L1 objects, such as the sum of transverse momenta of jets [69]. The final decision of whether to keep an event or discard it, is made by the CTP in 2.5 µs. It also determines interesting areas to be further investigated by the HLT. A schematic of the Level-1 trigger system is shown in Fig. 3.7.

The HLT is a software based trigger that applies a typical reconstruction algorithm at a preliminary stage followed by more advanced, CPU-intensive algorithms to trigger on events. A farm of CPUs is

running several algorithms in parallel to process signals based on selection criteria similar to offline physics requirements [70].

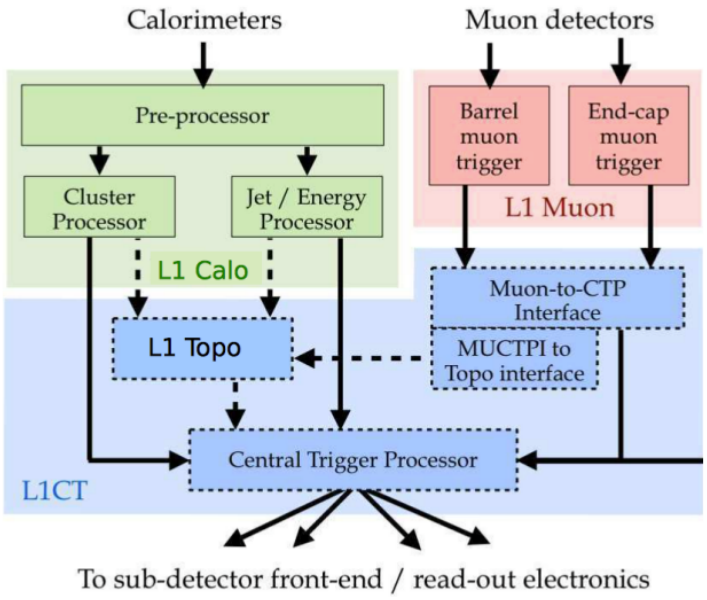


Figure 3.7: A schematic describing the Level-1 ATLAS trigger system [71].

3.4 Physics objects reconstruction

The various decay products from the proton-proton collisions travel through the detector volume leaving behind characteristic trails. Physicists analyse these trails and try to identify objects based on their unique signatures. A schematic showing signatures of different objects is shown in Fig. 3.8.

Tracking

The signals collected from the tracking detectors, the inner detector and muon spectrometers, are used to reconstruct charged-particle trajectories called tracks. The primary-tracking reconstruction from the SCT and pixel detector is performed in three steps: clusterisation, iterative combinatorial track finding and ambiguity solving [58].

In the first step, energy deposits above a predefined threshold are collected from the pixels and strips of the pixel detector and SCT, respectively. These deposits are called “hits”. Hits from pixel detector and SCT are combined to create three-dimensional space points. A linear approximation technique is used to find the intersection point of a particle on the pixel sensor. In a crowded environment, it may happen that the reconstructed cluster has energies from multiple particles. Therefore, it is essential for the algorithms to identify such clusters.

A group of space points are used to create track seeds. These preliminary track seeds pass through a number of selection criteria to ensure purity. For instance, there should be at least one additional space point compatible with the track trajectory estimated from the seed. The task of combining the filtered track seeds to construct track candidates is performed by the Kalman filter [73]. In principle, it

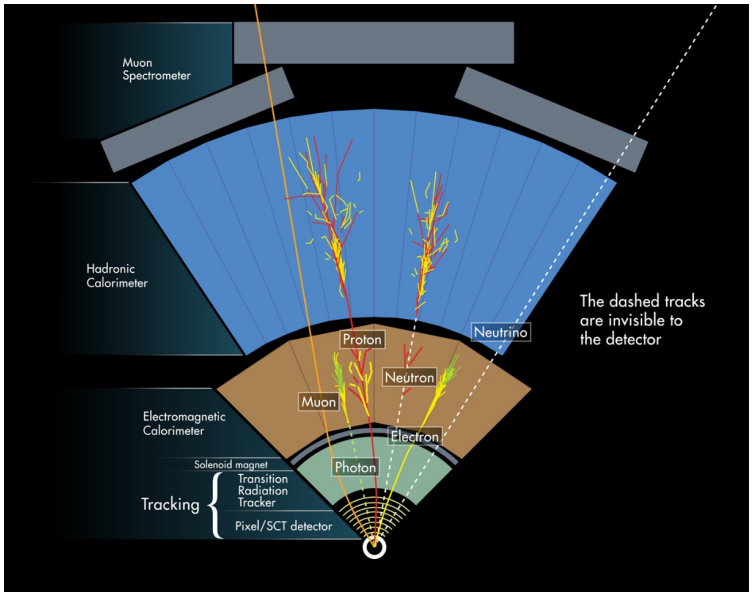


Figure 3.8: The schematic shows the signatures of electron, photon muon and hadrons inside the sub detector systems of the ATLAS detector. The dotted line represents the neutrino which is not detected directly [72].

searches for additional space points from other layers of pixels and strips that are suitable with the track seed. If the filter finds more than one suitable space point on the same layer then it creates multiples track candidates per seed. The Kalman filter is also effective in reducing the amount of track candidates from random space points.

Once all the track candidates are ready, the ambiguity solver is employed. It is required to rectify cases where space points might overlap or have been incorrectly assigned. The ambiguity solver processes track candidates individually based on their scores. Each track gets a score based on how likely it is to be a primary reconstructed track. It also resolves issues where a cluster is assigned to multiple track candidates. Moreover, track candidates are rejected if they fail to meet basic track quality requirements based on p_{T}, $|\eta|$ and number of hits contributing to the cluster formation.

The track candidates that pass the ambiguity solver are used as input to find suitable TRT measurements. The space points in TRT are expressed in r-ϕ and r-z coordinates in the barrel and endcap region respectively. These space points termed as *hits* are employed by the extension algorithm which performs a line fit to find hits suitable with the input track candidates. In this way, the chosen hits form the TRT further extend the track candidate [74].

Vertexing

Vertexing is the process of identifying vertices called primary vertices, which are locations in the detector where proton-proton hard scattering might have occurred. Secondary vertices are positions where particles produced at the primary vertices, further interact.

The vertex reconstruction algorithm takes tracks estimated from the ID as input. These tracks are filtered based on certain quality requirements to ensure they are suitable for vertex reconstruction. The algorithm operates in two steps: vertex finding and vertex fitting. Initially, a seed position is determined from the set of filtered tracks. Then, the seed and the filtered tracks are used to fit the

best vertex position. This process is iterative, starting with equal weights for all tracks. As iterations progress, the weights of more compatible tracks increase, while those of less compatible tracks decrease. Eventually, the less compatible tracks are returned to the pool for potential use in other vertex reconstructions. A vertex candidate must have at least two associated tracks. This procedure is repeated until there are no candidate tracks left, or no more vertices can be found from the remaining tracks. The vertex reconstruction algorithm produces a set of three-dimensional vertex positions and their covariance matrices [75].

The ATLAS reconstruction defines the hard-scatter primary vertex as the primary vertex with the largest $\sum p_{\mathrm{T}}^2$ of the associated tracks. The other primary vertices are associated with pileup contribution.

Topo-clusters

Topo-clusters, as the name suggests, are topological energy clusters produced inside the calorimeter cells caused by particle showers. The technique of collecting calorimeter signals from topologically connected cells and grouping them into clusters is an effective way in highly granular calorimeters such as the one in ATLAS. It is important to suppress the signals from the background noise and pileup contributions. Therefore, a quantity called signal significance is used to decide which signals contribute to cluster formation [76]. It is defined as the ratio of the cell signal to the average expected noise in this cell. The topo-clusters are seeded and expanded on the basis of threshold requirements based on signal significance.

Particle Flow

The momentum and energy measurements from the inner detector and calorimeter, respectively, are prone to overlaps of certain particles. For instance, an event in which an electron leaves a track in the ID and deposits energy in the calorimeter. If information from these sub detectors is used individually, there is a possibility of double counting this particle. In order to remove such overlaps and to utilise the sub detectors to the best of their potential, Particle Flow has been developed. The idea of Particle Flow is to reconstruct an event by combining measurements from the inner detector and calorimeters. The ID is pileup resilient and has a better resolution especially at low p_{T} whereas the calorimeter can detect neutral particles. Particle Flow utilises both of these advantages to produce more reliable measurements.

In the Particle Flow (PFlow) algorithm, one of the inputs is the set of reconstructed tracks that are extrapolated in the direction of calorimeter energy deposits. The algorithm evaluates the probability of the particle energy to be deposited in more than one topocluster. Accordingly, it either stores single topocluster energy or sums up energy of multiple topoclusters to constitute the full energy profile of the particle that created the track. The expected energy deposited by the particle that created the track is subtracted from the matched topoclusters. The remaining energy after subtraction is compared to a threshold value. If the energy is below the threshold, it is removed as it may be due to shower fluctuations. Conversely, if the energy exceeds the threshold, it is retained as it may originate from multiple particles. At the end, a set of selected tracks and remaining topoclusters in the calorimeters represent the reconstructed event with no double counting of signatures. A schematic diagram of PFlow algorithm is shown in Fig. 3.9.

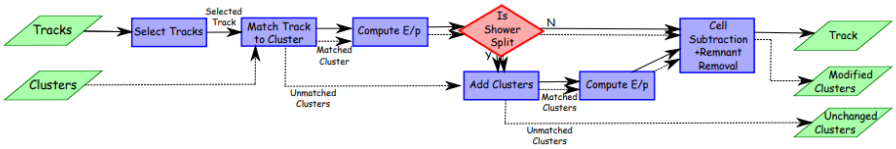


Figure 3.9: A chart representing the Particle Flow algorithm [77]

Electrons

The topoclusters from the electromagnetic calorimeter and the charged particle tracks from the inner detector are used to reconstruct an electron candidate. A sliding-window algorithm was used for the 2015-2016 dataset and the supercluster algorithm was used for the 2015-2018 dataset. In the sliding-window algorithm [78], seed clusters exceeding a transverse energy of 2.5 GeV are identified. The algorithm scans the area in either η or ϕ to locate seed clusters. Furthermore, the reconstructed tracks in the ID that loosely match with the seed clusters are selected on which a fitting procedure is applied. The separation of the cluster-centre and the track must fulfill $|\eta_{\text{cluster}} - \eta_{\text{track}}| < 0.05$. If multiple tracks fulfill the matching criteria, additional requirements are imposed to select the primary electron track. One of these requirements is that the chosen track should not have any association with a vertex from a photon conversion [79].

In the superclusters algorithm, topo cluster construction is improved by better integrating the energy lost due to bremsstrahlung or from electron or photon conversion. Superclusters are constructed in two steps: first step is to identify candidate seed clusters with signal higher than a certain threshold and the next step is to find clusters close to the seed clusters, that may originate from bremsstrahlung. These clusters are added to the seed candidates if required. The superclusters are constructed for photons and electrons separately. Once they are ready, the electron superclusters are matched with ID tracks and photon superclusters are matched with conversion vertices [80].

Muons

There are two types of muon reconstruction algorithms: stand-alone reconstruction from the MS and global reconstruction based on signals from the MS and the ID.

Under the stand-alone method, track segments are identified from different MS stations using a Hough transform. By combining these segments, basic track candidates are constructed. Then a global-χ^2 fit of the muon trajectory through the magnetic field is performed and based on the fit results, unsuitable hits are removed. Moreover, hits that were not considered before are now added, and the fit is re-performed. In the ambiguity resolution step, tracks that share numerous high-quality hits are handled.

The global reconstruction of muons is performed based on the information from the ID and the MS, along with energy loss inside the calorimeters. There are five different strategies leading to different muon types as follows. This analysis makes use of combined muons.

- Combined muon (CB): Here a combined track fit is performed on the hits extracted from both the tracking sub detectors. Moreover, calorimeter energy loss is also considered.

- inside-out muon (IO): The ID tracks are extrapolated to the MS and at least three loosely

alligning MS hits are obtained. After that, a combined fit is performed on the ID tracks, MS hits and calorimeter energy loss.

- Muon-extrapolated (ME): The muons tracks that do not match with any ID tracks are defined for a ME muon. These muons are outside the acceptance range of the ID but still in the acceptance range of the MS.
- Segment-tagged (ST): The ID tracks are required to match at least one MS segment by fulfilling tight angular matching. The ID tracks matched in this way are considered for a muon candidate.
- Calorimeter-tagged (CT): Here the ID tracks are extrapolated to the calorimeters to find clusters consistent with a minimum-ionising particle. The matched ID track is considered for a muon candidate.

Jets

Jets are clusters of particles generated by a quark or a gluon. Jet reconstruction inputs are obtained from the Particle Flow algorithm described in Section 3.4. The matched charged particle tracks and the remaining topoclusters, after the subtraction step, are filtered based on their compatibility with the hard-scatter primary vertex. Charged particle tracks matching this vertex are retained. This is done by requiring $|z_0 \sin\theta| < 2$ mm, where z_0 is the distance of closest approach of the track to the hard-scatter primary vertex along the z-axis [77]. Moreover, the position coordinates of the topoclusters are recalculated relative to the hard-scatter primary vertex. Particle Flow jets are reconstructed using the anti-k_t algorithm with a radius parameter of 0.4. Details about the anti-k_t algorithm can be found in [81].

b-jets

The procedure of identifying a hadronic jet originating from a b-quark is called b-tagging. The b-tagging algorithms are designed to exploit the characteristics of a b-hadron, for instance, its high mass and longer lifetime. The lifetime of a weakly decaying b-hadron is of the order of 1.5 ps and therefore, it can travel a significant distance before decay. Technically, the vertex at which it decays is displaced from the primary vertex where it is produced. Moreover, the decay multiplicity is also higher given the fact that b-hadrons are massive. The algorithms at ATLAS employ a two-staged approach: Firstly, it exploits the large impact parameters[1] of the tracks associated with hadronic jets and their compatibility with the primary vertex. Then, it explicitly reconstructs displaced vertices. Secondly, multivariate classifiers are used to discriminate b-jets compared to c-jets or other light flavour jets [82].

Missing Transverse Energy

Reconstruction of particles crossing the detector volume is crucial to understand what particles are produced as a result of proton-proton collisions. The reconstruction algorithms heavily rely on the signs or traces these particles leave inside the detector. However, some particles leave no traces behind but are still relevant for the analysis. Such particles mostly include SM neutrinos or some

[1] Impact parameter is the distance of closest approach of the track to the collision point.

unknown dark matter particles. The energy corresponding to these "invisible" particles is called missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$). or MET (used in ATLAS software).

The underlying physics is the law of conservation of momentum. The total momentum in the transverse plane is zero before the collision because the initial protons are boosted along the z-axis. Therefore, the total sum of momenta of particles in the transverse plane after the collisions must be zero. The missing transverse momentum algorithm computes two quantities: $p_{\mathrm{T}}^{\mathrm{hard}}$ and $p_{\mathrm{T}}^{\mathrm{soft}}$. The first quantity $p_{\mathrm{T}}^{\mathrm{hard}}$ is the sum of transverse momenta of high-p_{T} objects including reconstructed electrons, photons, muons, jets and τ-leptons whereas $p_{\mathrm{T}}^{\mathrm{soft}}$ includes soft signals from charged particle tracks that are associated with a hard event vertex but not associated with any hard object [83].

The missing energy reconstruction algorithm builds two quantities: missing transverse momentum vector $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}} = (p_x^{\mathrm{miss}}, p_y^{\mathrm{miss}})$ given by Eq. (3.3) and the scalar sum of all transverse momenta that enter the computation of MET given by Eq. (3.4).

$$\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}} = -\left(\sum_{\substack{\text{selected}\\ \text{electrons}}} \boldsymbol{p}_{\mathrm{T}}^{e} + \sum_{\substack{\text{accepted}\\ \text{photons}}} \boldsymbol{p}_{\mathrm{T}}^{\gamma} + \sum_{\substack{\text{accepted}\\ \tau-\text{leptons}}} \boldsymbol{p}_{\mathrm{T}}^{\tau} + \sum_{\substack{\text{selected}\\ \mu}} \boldsymbol{p}_{\mathrm{T}}^{\mu} + \sum_{\substack{\text{accepted}\\ \text{jets}}} \boldsymbol{p}_{\mathrm{T}}^{\mathrm{jet}} + \sum_{\substack{\text{unused}\\ \text{tracks}}} \boldsymbol{p}_{\mathrm{T}}^{\mathrm{track}} \right) \tag{3.3}$$

$$\sum p_{\mathrm{T}} = \sum_{\substack{\text{selected}\\ \text{electrons}}} p_{\mathrm{T}}^{e} + \sum_{\substack{\text{accepted}\\ \text{photons}}} p_{\mathrm{T}}^{\gamma} + \sum_{\substack{\text{accepted}\\ \tau-\text{leptons}}} p_{\mathrm{T}}^{\tau} + \sum_{\substack{\text{selected}\\ \mu}} p_{\mathrm{T}}^{\mu} + \sum_{\substack{\text{accepted}\\ \text{jets}}} p_{\mathrm{T}}^{\mathrm{jet}} + \sum_{\substack{\text{unused}\\ \text{tracks}}} p_{\mathrm{T}}^{\mathrm{track}} \tag{3.4}$$

The magnitude of $\boldsymbol{p}_{\mathrm{T}}^{\mathrm{miss}}$ is the missing transverse momentum ($E_{\mathrm{T}}^{\mathrm{miss}}$) or MET. The quantity $\sum p_{\mathrm{T}}$ represents total activity in an event, and it gives the overall scale for evaluating the hardness of a hard-scatter event.

CHAPTER 4

Signal extraction and background processes

This chapter provides a detailed description of the tZq process and its trilepton decay channel which is the signal in this analysis. The baseline event selection defined to extract maximum signal is given in Section 4.2. The different background processes are described in Section 4.3. In addition, information about simulated samples used for signal and background processes can be found in Section 4.4. The various sources of systematic uncertainties are outlined in Section 4.6. Moreover, a brief explanation of neural networks is also provided at the end.

4.1 tZq production

One of the rare processes at the LHC is the associated production of a top quark and a Z boson through electroweak interactions. It is referred to as tZq production. Its LO t-channel Feynman diagrams in the four-flavour scheme (4FS) are shown in Fig. 4.1. The 4FS is a theoretical description in which active quarks in a proton are u, d, c and s whereas the bottom quark originates from gluon splitting. On the contrary, in five-flavour scheme, a bottom quark is a part of the colliding proton. A Z boson is radiated from any one of the incoming or outgoing quarks (Fig. 4.1(a)) or from the exchanged W boson (Fig. 4.1(b)). In addition to these resonant contributions, there is a small non-resonant contribution in the form of tl^+l^-q (Fig. 4.1(c)) which is also accounted for in this analysis. The associated t quark is produced through interactions such as $u + b \rightarrow d + Z + t$ or $\overline{u} + b \rightarrow \overline{d} + Z + t$ whereas a $\bar{t}$ quark is produced via the charge conjugated processes.

tZq production is interesting to study because it probes the coupling of a top quark and a Z boson as well as the WWZ coupling. In other words, it allows the coupling of two bosons and the coupling of a fermion to a boson to be studied in a single interaction. Moreover, it can provide a solid basis to study similar processes such as the tHq process. The theoretical cross-section based on SM prediction, is calculated at NLO in QCD for a dilepton mass more than 30 GeV to be $(102\,^{+5}_{+2})$ fb. tZq process was observed by ATLAS and CMS during the Run-2 of the LHC. The cross-section, measured by the ATLAS collaboration, is $(97 \pm 13\,(\text{stat.}) \pm 7\,(\text{sys.}))$ fb [84] which is consistent with the SM expectation.

In order to study this process, one has to note that the particles involved are heavy and therefore the only way to spot them is from their reconstructed decay products. Moreover, depending on the branching ratio, there can be several sets of final states. A common practice is to divide the possible

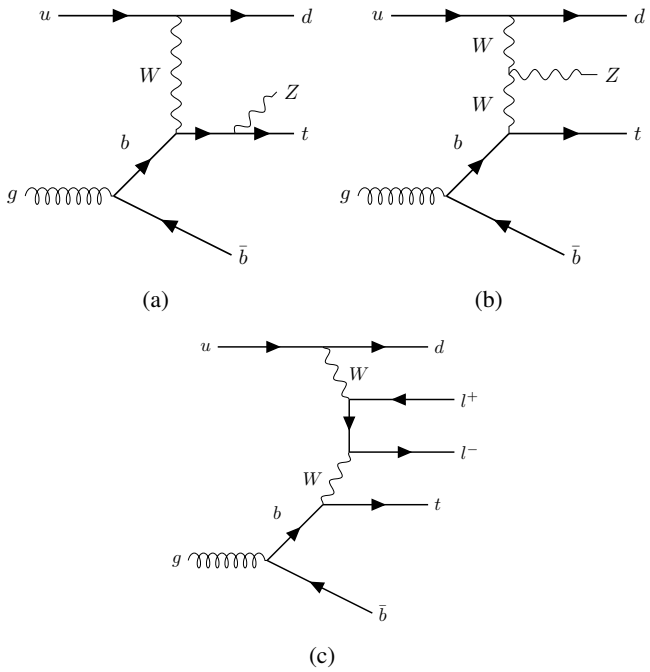


Figure 4.1: Feynman diagrams at LO for tZq production. A Z boson is radiated either from one of the quarks or from the exchanged W boson.

final states into *channels* based on certain combinations of leptons and jets. This analysis focuses on the trilepton decay channel which is described below.

4.2 tZq trilepton channel

As the name suggests, the trilepton decay channel of tZq production contains final states with three charged leptons, as shown in Fig. 4.2. The top quark decays almost exclusively into a bottom quark and a W boson while the corresponding W boson can decay either leptonically or hadronically. Approximately in 25% events, W boson decays into a charged lepton and an associated neutrino. The Z boson can decay either into a pair of charged leptons or into neutrinos (invisible) or into hadrons. In approximately 8% of the produced events Z boson decays into opposite sign same flavour lepton pairs (OSSF). Its probability is equal across the three lepton families (e,μ,τ) owing to lepton flavour universality. It is one of the principles of the SM that the interactions of weak gauge bosons and leptons is the same for all the lepton flavours. This analysis includes Z boson decays resulting into electrons or muons (e^- e^+ or μ^- μ^+). The tau leptons are considered if they decay into lighter leptons (i.e. e or μ).

The branching fractions of the various possible decays of top quark and Z boson are shown in Fig. 4.3. The combination of both of them decaying into leptons occur in only 2% of the produced tZq events. However, selecting the trilepton decay state ensures higher signal purity because a decay state with three leptons is difficult to replicate by background processes. Therefore, the trilepton channel is chosen for studying the tZq process. From this point onward, this will be referred to as the signal.

The accurate reconstruction of top quark and Z boson is essential for the efficient identification

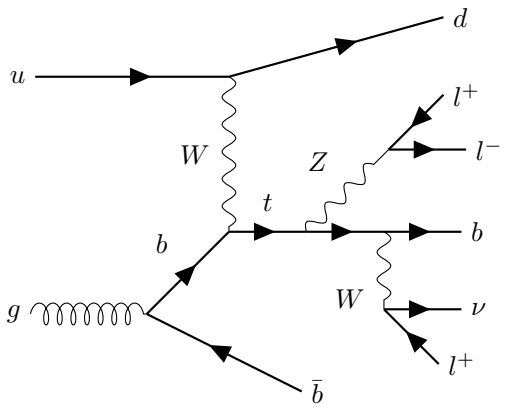


Figure 4.2: *tZq* trilepton final state

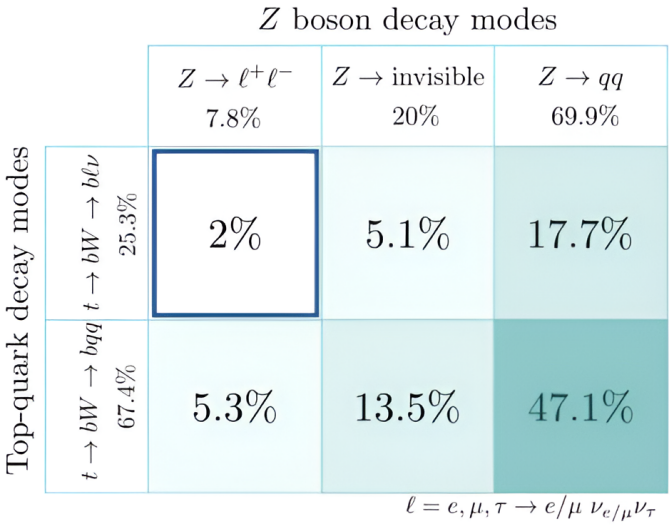


Figure 4.3: Branching ratios of possible decays of top quark and *Z* boson, along with the fractions representing combination of decays [85]

of signal. Z boson is reconstructed from an OSSF lepton pair. If all three leptons have the same flavour, the leptons having invariant mass closest to the mass of Z boson, are chosen for Z boson reconstruction.

The lepton failing the criteria for Z boson reconstruction is used to reconstruct the W boson along with missing transverse energy. Finally, the top quark is reconstructed from the W boson and b-tagged jet having the maximum p_{T}.

Due to the t-channel production, one or more jets are produced close to the beam direction. The forward-jet is defined as the non b-tagged jet yielding the maximum invariant mass with the leading b-jet.

Event selection

In LHC physics, the outcome of a collision between two incoming particles is called as an "event". An important step in an analysis is to reconstruct the final state of interest from the detector data or in other words, find possible occurrences of this final state within the collision events. In order to achieve that, certain requirements are defined in favour of the signal events. The collection of these requirements is called event selection.

For this analysis, events with exactly three leptons (e or μ) are required. τ lepton is considered if it decays into lighter leptons. These leptons are sorted by their p_{T} which must be at least 27,15 and 10 GeV, respectively. Among the selected leptons, there should be at least 1 OSSF lepton pair with its invariant mass (m_{ll}) within 10 GeV of m_Z. This is to identify which out of the selected leptons originate from the Z boson. Moreover, a cut of 20 GeV is applied on invariant mass, in order to suppress backgrounds not containing a Z boson. A cut of 30 GeV is applied on the transverse mass of the W-boson to account for the missing transverse energy.

The number of jets is required to be between 2 and 5, with p_{T} more than 25 GeV and $|\eta|$ more than 4.5. The number of b-jets is required to be 1 or 2, reconstructed at 85% working point with $|\eta|$ more than 2.5. Events with 2 jets, both b-tagged ($2j2b$) are not considered.

Events selected through these baseline cuts are then arranged into phase space regions which is discussed in Section 6.3. It is important to note here that these requirements are chosen to maximise the probability of selecting signal events but in reality there are background processes that mimic the tZq signature and therefore, contaminate the selected signal events.

4.3 Background processes

The background processes for tZq process can be classified according to the number of prompt (or real) leptons in the final state. A lepton is labelled prompt if it originates from either a τ or a massive boson. On the other hand, non-prompt or fakes are objects misidentified as leptons. The source of non-prompt leptons can be bottom and charm hadron decays, meson decays, photon conversions or light jets creating lepton-like signatures. Backgrounds involving only prompt leptons are diboson, $t\bar{t}$+X, $t\bar{t}H$ and tWZ while backgrounds involving non-prompt leptons are $t\bar{t}$, Z+ jets and tW.

Backgrounds involving prompt leptons

In the diboson process, two massive bosons are produced which can be ZZ, WW or WZ, as shown in Fig. 4.4. As per Fig. 4.4(a), the leptonic decay of bosons result in three prompt leptons which can pass

the signal event selection if additional jets are also found. For the ZZ scenario, as shown in Fig. 4.4(b), one of the leptons needs to fail the requirement for a prompt lepton or is not reconstructed. Due to this strong resemblance of the diboson signature with the signal, it is the dominant background in the tZq production.

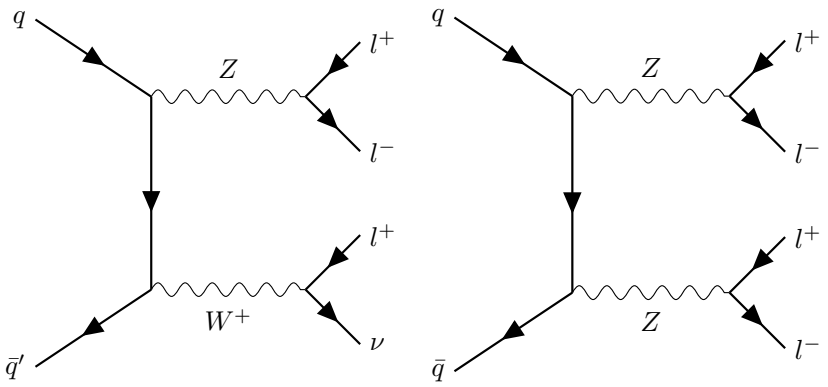


Figure 4.4: Feynman diagrams for the diboson background

The t-quark pair production in association with a heavy boson (Z or W) can be an important source of background. In particular, the $t\bar{t}Z$ process, where the final state already includes a Z boson and a t quark, can produce a very similar signal-like signature. It is shown in Fig. 4.5. $t\bar{t}H$ production contributes less because of its small cross-section.

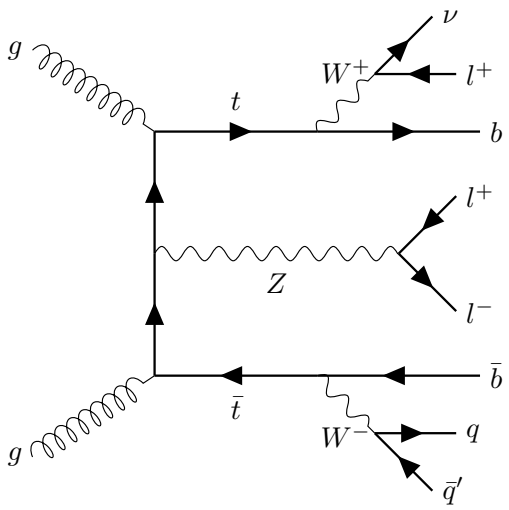


Figure 4.5: Feynman diagrams for the $t\bar{t}Z$ background

Backgrounds involving non-prompt leptons

Backgrounds involving non-prompt or fake leptons are t-quark pair production and the production of Z-boson with jets. As shown in Fig. 4.6(b), there are already two leptons from the Z-boson. If the jets are originating from lighter quarks (quarks from first or second generation), they can be misidentified as leptons leading to a non-prompt lepton contribution. In the $t\bar{t}$ production, as shown in Fig. 4.6(b), if one of the b-jet decays into a lepton, then it can satisfy the signal event selection.

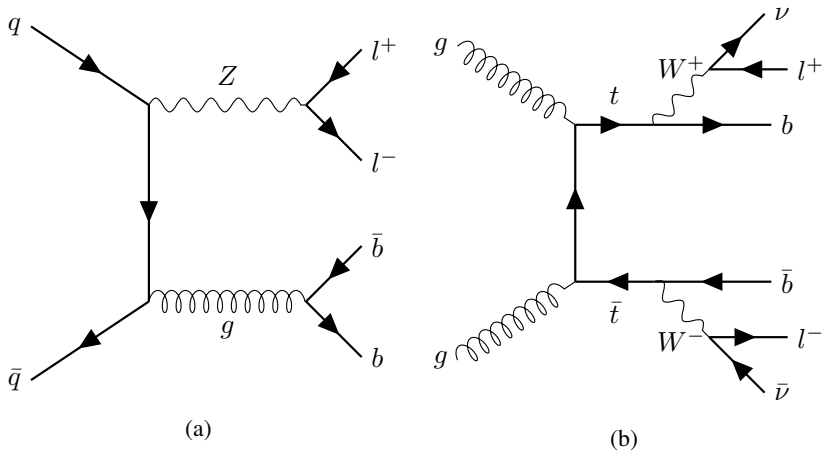


Figure 4.6: Feynman diagrams for non-prompt lepton background

4.4 Monte Carlo simulations and event generation

Monte Carlo simulations are a crucial part of this analysis, as they are used to interpret the detector data. As already discussed before, proton-proton collisions are viewed as collisions between partons. Given the number of incoming protons, fully understanding the collisions is complex. Moreover, the randomness associated with it makes it difficult to use deterministic methods. Here is where Monte Carlo (MC) simulations come to the rescue. MC simulation is a computational technique to model a complex system using random numbers and underlying probability distribution functions. In general, MC simulations use random numbers to sample from the underlying probability distribution functions and then average the result of several iterations of sampling.

In high energy physics, MC simulations are performed by MC generators which are also called event generators. An event consists of a number of outgoing particles produced in accordance with conservation laws. However, quantum processes are inherently random and therefore, the outgoing particles vary from event to event. The task of an event generator is to generate random sequences of events based on given probability distribution functions. This task is not trivial because the structure of an event in hadron collisions is complex. Information about scattering, decays, radiations is required in order to accurately model an event. There is no comprehensive theory that can completely predict the properties of events. Thus, MC generators draw on theoretical models, with the Standard Model as their foundation, while also incorporating aspects derived from experimental data. In analyses, it is advised to test more than one generator because the extent of event modeling might be different for different generators.

The resulting information from an event generator consists of four-momenta of stable particles, known as final state particles. Another use-case of MC generators is detector simulation where the evolution of the final state particles when they travel through the detector is simulated [86].

The simulated data provided by these generators can be thought of as a "digital twin" of the actual observed data. It can be used to predict any experimental observable or a combination of observables. The workflow of MC generators, called as the MC chain, is a step-by-step process that begins with identifying the hard interaction and continues until the final state is achieved. At each stage, the structure of the underlying event evolves. The steps are briefly summarised below:

- **Hard process**: In this step the hard process, defined as the process with the highest momentum transfer, is determined using matrix element calculation combined with the input parton distribution functions. If a resonance is produced in a hard process, such as the t or Z and it shortly decays, then its decay process is also considered within the hard process.

- **Parton shower**: The colliding partons are responsible for emissions that give rise to more partons and subsequently more interactions. The emissions associated with the incoming partons is called initial-state radiation (ISR) while the emissions associated with the outgoing partons is called final-state radiation (FSR). These emissions and their respective interactions are modelled with parton shower (PS) algorithms. The incorporation of PS into the matrix element paints a more accurate picture of the collision process.

- **Multiple-processes**: Until now, only one of the partons from the original hadron is considered but in reality, other partons from the same hadron also interact. Their interactions are termed as multiple-processes. They are calculated at this step of the MC simulation chain.

- **Hadronisation and decay**: The outgoing partons, with sufficient energy, can produce new hadrons due to QCD colour confinement (hadronisation). If these hadrons are unstable, they can also decay into lighter particles. The MC chain also includes these calculations.

The MC chain results in a set of events with well-defined and stable final state particles along with their kinematic distributions. At this stage, the simulated data is referred to as the *truth-level* data. The next step is detector simulation where the interplay between the detector material and final state particles is simulated.

After all the MC events are generated, the interactions of the particles with the detector material is simulated using the GEANT4 software toolkit [87]. The next step is digitisation where detector response is simulated from the interactions and stored in a format compatible with real output from the detector. Moreover, pileup events are also added at this stage [88]. Finally, the simulated data is ready for comparison with the observed detector data. The workflow of the MC chain is illustrated in Fig. 4.7. Among the various available MC generators, Herwig [89], Pythia [90], MadGraph [91] and Sherpa [92] are some of the most commonly used ones.

4.5 Data and simulated samples

This analysis uses collision data collected by the ATLAS detector during 2015 to 2018 (Run-2 of the LHC) at a centre of mass energy of 13 GeV. The total integrated luminosity is of 140.1 fb^{-1}. The

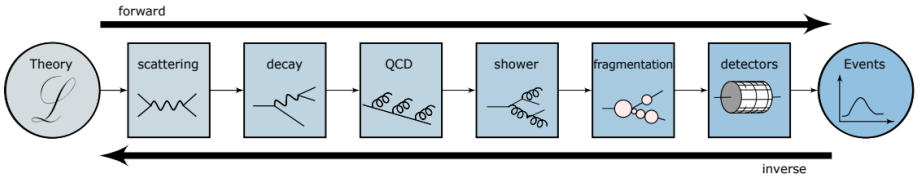


Figure 4.7: An illustration of the steps involved in a Monte Carlo chain [93]

selected data were collected during stable beam operations at the LHC and the ATLAS detector fully functioning.

The ATLAS MC samples for analysis of the Run-2 dataset are divided into three subsets: `mc16a` generated with 2015 an 2016 pileup conditions, `mc16d` generated with 2017 pileup conditions and `mc16e` that includes pileup conditions of 2018 data.

Signal sample

The tZq signal sample is simulated using MadGraph5_aMC@NLO 3.5.1 [94] generated at next-to-leading order (NLO) with NNPDF3.0NLO [95] parton distribution function. In general, the models used for PS and hadronisation contain several free parameters which must be optimised to generate a reasonable description of data. The optimisation is termed as tuning and the resulting sets of parameters are called tune sets. Pythia is used for the parton shower and hadronisation along with the A14 (ATLAS 24 [96]) tune set and the NNPDF2.3LO PDF set. The top quark is decayed at LO using MadSpin [97] to preserve spin correlations. The signal sample is generated using the four-flavour scheme.

Background samples

For processes including $t\bar{t}Z$, $t\bar{t}t\bar{t}$, tWZ and $t\bar{t}t$, MadGraph5_aMC@NLO was used to calculate the matrix element and cross-sections combined with Pythia for parton shower modelling. The generator used for simulating diboson, triboson, Z+jets and $t\bar{t}W$ is Sherpa along with its default parton shower. For $t\bar{t}$, tW and $t\bar{t}H$ processes, Powheg and Pythia were used. The specific versions of generators used for different backgrounds are summarised in Table 4.1. Herwig is used for PS and hadronisation in case of alternate samples required for the theoretical uncertainties as described in Section 4.6.

4.6 Systematic uncertainties

One of the key advantages of using simulated data is their ability to predict how the observed data may appear. However, it is crucial to assess how dependable both the simulated and the measured data are, which is quantified through uncertainties. The proper inclusion of uncertainties is an important part of any analysis.

Uncertainties can be divided into two categories: statistical and systematic. Simply put, statistical uncertainties are related to the statistics of the data whereas systematic uncertainties are complex uncertainties that are not directly from the statistics of the data. Consider an experiment involving a weighted dice. Here, the statistical uncertainty arises from the natural variability in the outcomes of

Table 4.1: Background sample details

Background	Generator	Parton Shower	PDF	Type of Sample
$t\bar{t}Z$	MadGraph5_aMC@NLO 2.8.1	Pythia 8.244	NNPDF3.0NLO NNPDF2.3LO	Nominal
$t\bar{t}Z$	MadGraph5_aMC@NLO 2.8.1	Herwig 7.2.1	NNPDF3.0NLO NNPDF2.3LO	Alternate
tWZ	MadGraph5_aMC@NLO 2.X.X	Pythia 8.235	NNPDF3.0NLO	Nominal
Diboson	Sherpa 2.2.12	Sherpa MEPS@NLO	NNPDF3.0NNLO	Nominal
Triboson	Sherpa 2.2.2	Sherpa MEPS@NLO	NNPDF3.0NNLO	Nominal
$t\bar{t}$	Powheg Box v2	Pythia 8.230	NNPDF3.0NLO NNPDF2.3LO	Nominal
$t\bar{t}$	Herwig 7.2.1	Pythia 8.230	NNPDF3.0NLO	Alternate
tW	Powheg Box v2	Pythia 8.230	NNPDF3.0NLO NNPDF2.3LO	Nominal
Z+ jets	Sherpa 2.2.11	Sherpa MEPS@NLO	NNPDF3.0NNLO	Nominal
$t\bar{t}W$	Sherpa 2.2.10	Sherpa MEPS@NLO		Nominal
$t\bar{t}H$	Powheg Box v2	Pythia 8.230	NNPDF3.0NLO NNPDF2.3LO	Nominal
$t\bar{t}t$	MadGraph5_aMC@NLO 2.2.2	Pythia 8.186	NNPDF2.3LO	Nominal
$t\bar{t}t\bar{t}$	MadGraph5_aMC@NLO 2.3.3	Pythia 8.230	NNPDF3.1LO	Nominal
$t\bar{t}t\bar{t}$	MadGraph5_aMC@NLO 2.3.3	Herwig 7.0.4	NNPDF3.1LO	Alternate

repeated dice rolls. The systematic uncertainty comes from the underlying bias introduced by the weight in the dice which affects the fairness of the dice and systematically shifts the outcomes.

In high energy physics, sources of systematic uncertainties are calibrations of scales, efficiencies of particle identifications and reconstructions, choice of MC generators, etc. These sources of systematic uncertainties are categorised into: instrumental and theoretical uncertainties as described in Section 4.6 and Section 4.6, respectively. In ATLAS, there are analysis groups dedicated in evaluating the size of these systematic uncertainties and providing them for the analysers. In ATLAS, determination of the sizes of systematic uncertainties is performed by dedicated combined-performance (CP) groups. Their results are utilised and validated in physics analysis.

Instrumental or detector uncertainties

This category describes the systematic uncertainties arising due to lack of a perfect detector response. The sources can be luminosity estimation, pileup computation and scales used for object reconstruction. These uncertainties can affect the normalisation and/or shape of detector-level distributions.

- **Luminosity**: The integrated luminosity is 140.1 $\text{fb}^{-1}$and the uncertainty in its calculation is 0.83% [98]. This uncertainty is applied to all simulated signal and background processes as a normalisation uncertainty.

- **Pileup reweighting**: MC generators make use of scale factors to account for differences in pileup distributions between data and simulations. There is an uncertainty associated with these scale factors. It is evaluated by changing the nominal pileup value to a lower and a higher value, then the effect of these changes is calculated to obtain the up and down uncertainty.

- **Jet Energy Scale (JES)**: After the jets are reconstructed, their energies need to be adjusted so that it reflects the energy of the colliding particles. The calibration is done by comparing the reconstructed jets with the true jets which are simulated jets of stable particles without detector effects. Uncertainties originating from the calibration process are categorised as JES uncertainties [99].
- **Jet Energy Resolution (JER)**: JER is the detector's ability to distinguish two jets with similar total energy. The uncertainty associated with the differences in JER for data and simulation is called JER uncertainty [100].
- **Jet-Vertex-Tagger (JVT)**: The JVT algorithm selects jets coming from hard scatter processes and in turn, reduces those coming from pileup. It uses track-based variables that are sensitive to the origin of jets to construct a function called JVT factor. Based on the function values, one can separate jets associated with primary vertex and pileup jets [101]. The uncertainty associated with the JVT factor is one of the systematics.
- **Lepton reconstruction**: Scale factors are applied to correct differences between data and simulation in case of lepton identification, isolation and trigger efficiencies. The uncertainties associated with these scale factors belong to the lepton reconstruction category of systematics.

Theoretical uncertainties

This category involves uncertainties originating from the various models used in the MC simulation chain, and therefore, also called modeling uncertainties or modeling systematics. There are various parameters related to the modelling of a certain process and for each parameter, there is an associated uncertainty which is investigated by varying the values of the parameters. The effect of the variation is estimated and assigned as the systematic uncertainty. In this way, the uncertainties associated with the modelling of the process is extracted.

- **A14**: The uncertainty associated with the A14 tune set is determined by comparing the nominal sample with two alternate samples, both simulated using the same settings as the nominal sample but incorporating the up and down variations of the A14 tune set. This variation is related to the strong coupling constant α_s. This uncertainty is considered for tZq and $t\bar{t}Z$ processes.
- **Scale**: The parameters representing renormalisation and factorisation scales are known as μ_R and μ_F, respectively and their values are μ_R=μ_F=$H_T/6$. The renormalisation scale is for the running of α_S associated with the hard process whereas the factorisation scale for the PDFs. They are introduced to prevent the matrix element from any possible divergences. Moreover, they also separate the perturbative from the non-perturbative parts of the calculation. In order to estimate the uncertainty on the parameters, the values of μ_R and μ_F are varied between μ_R=μ_F=0.5, μ_R=μ_F=1 and μ_R=μ_F=2. This uncertainty is considered for tZq, diboson and $t\bar{t}Z$ processes.
- **Shower and PDF**: The showering uncertainty is calculated by comparing the nominal sample (modeled using Pythia) and the alternate sample (modeled using Herwig). Moreover, uncertainty associated with the PDF is also considered. The shower systematic is included for tZq, $t\bar{t}Z$ and $t\bar{t}$ processes. PDF uncertainty is considered for tZq, diboson and $t\bar{t}Z$ processes.

- **Interference**: This uncertainty accounts for the interference between $t\bar{t}Z$ and tWZ processes. The non-resonant tWZ production can feature a resonant $\bar{t}$ in the intermediate state, leading to overlap with the $t\bar{t}Z$ production. This interference is navigated using various techniques called diagram removal (DR) and diagram subtraction (DS) [102]. Two tWZ samples were generated, one with DR1 and another with DR2 and the difference is taken to be the tWZ modeling uncertainty.

- **Matrix element matching and resummation**: The systematics in this category are taken into account for the diboson process following the recommendations of the ATLAS Physics Modelling Group [103, 104]. The CKKW parameter is related to the calculation of the overlap between jets involved in the matrix element and parton shower computation. The nominal value of the parameter CKKW is 20 GeV and its uncertainty is estimated by varying this value to 30 GeV (up variation) and 15 GeV (down variation) [105]. The QSF parameter determines the scale for the resummation of soft gluon emissions. Its uncertainty is estimated by varying the nominal by 2 and 0.5 [105].

  In the parton shower computation, a recoil scheme refers to how the remaining partons adjust their momenta after emission or splitting, in order to conserve the total momentum. The recoil scheme used for the diboson samples is described in [106]. The uncertainty on the associated parameter CSSKIN is also considered for the diboson process. Finally, approximate NLO EW corrections are included as weights using the electroweak virtual approximation as described in [107].

- **Matching and ISR**: The $t\bar{t}$ background is modelled using Powheg generator in which a parameter called h_{damp} is the damping factor that controls the radiation at which the $t\bar{t}$ system recoils. If this is set to infinity then all the radiation is considered in the computation of ISR. The uncertainty on ISR is estimated by comparing two nominal $t\bar{t}$ samples, one with h_{damp}=1.5m_t and second with h_{damp}=3m_t.

  Depending on the generator and parton shower workflow, there might be overlapping phase spaces which can cause double-counting of events. To prevent this, a parameter called p_{T}hard is defined which refers to the hard scattering transverse momentum scale. This parameter decides the extent of phase space which is vetoed while matching matrix element with the parton shower. The uncertainty on p_{T}hard is estimated by comparing samples with p_{T}hard=1 and p_{T}hard=0.

4.7 Event weights

The collision events generated from the MC simulations must be reweighted in order to reproduce data-taking conditions. The probability of an event, relative to the sum of probabilities for the sample, is given by its MC event weight (w_{MC}). The MC event weights are important because the sum of these weights will determine the correct number of events for that sample.

In addition, a number of correction factors are applied in order to match the data-taking conditions as well as correcting differences in efficiencies of identifying physics objects. Some weights also correspond to a systematic variation. Consider a systematic uncertainty which requires a parameter to be varied. Now, producing the entire event sample with the varied parameter is computationally intensive. Instead, a corresponding weight is included in the event generation [108].

The total event weight can be written as:

$$w_{\text{total}} = w_{\text{MC}} \cdot w_{\text{pileup}} \cdot w_{\text{lepton}} \cdot w_{\text{JVT}} \cdot w_{\text{trigger}} \cdot w_{b\text{-tagging}} \tag{4.1}$$

- w_{MC}: gives the relative probability of producing that event in that sample
- w_{pileup}: to correct the pileup profile of simulated events to match observed data
- w_{lepton}: to correct for the differences in lepton isolation and reconstruction between MC and data
- w_{JVT}: differences in data and MC when applying a cut on the JVT factor are considered by applying this weight
- w_{trigger}: any inconsistency between data and MC related to trigger efficiencies is corrected with this weight
- $w_{b\text{-tagging}}$: since this analysis requires events to contain b-jets, a weight called $w_{b\text{-tagging}}$ is also applied.

4.8 Artificial Neural Networks

A neural network is a computation tool developed to function in a way similar to the human mind. It is widely used in high energy physics for data analysis. The structure of a neural network (NN) is made up of neurons or *nodes*. Their function is to examine unknown systems and identifying interesting features, just like the job of neurons in human mind. Generally, these nodes are arranged in three or more different layers: input layer, hidden layer(s) and the output layer. A list of variables is given as input to the nodes of the input layer. Processing takes place through the subsequent layers and at the end, the output layer returns the conclusions derived by the network. Connections between nodes of different layers are referred to as *synapses*. Each connection between nodes of two consecutive layers, has a weight associated to it. Figure 4.8 shows a diagram of a neural network.

The input received by each node is the sum of weighted output of all nodes of the previous layer. As given in Equation 4.2, y_j is the input to node j, w_{ij} is the weight from the i-th node and x_i is that node's output. The term w_{0j} is called bias.

$$y_j = \Sigma_{i=1}^{n} w_{ij} x_i + w_{0j} \tag{4.2}$$

The output of a node is defined by an *activation* function. A common choice for an activation function is the sigmoid function. It provides output between 0 and 1. A feature that makes a NN special is its capability to *learn* from examples with known inputs and outputs. This is referred to as *training* a neural network. The purpose of training is to find appropriate weights. In supervised training, inputs and outputs are provided to a NN. It processes input and then compares the resultant output with the desired output. Comparison is done by calculating a *loss function*. It is a way to determine how well is the network trained. For better performance of a network the loss function should be minimised. In order to do that, information from the resultant output is propagated back in a model and the initial weights are readjusted so that output is closer to the desired output. This is how a network learns. A dataset flows inside a network several times and each time weights are refined until a minimum value of the loss function is obtained.

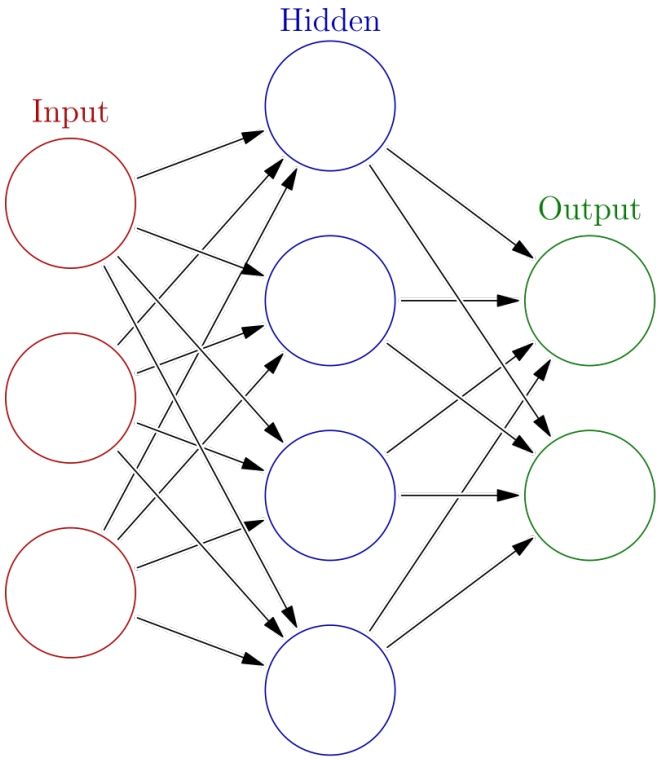


Figure 4.8: Diagram of a neural network including the input, hidden and output layers [109]

CHAPTER 5

Statistical methods

In particle physics, the abundance of available data for exploration has significantly increased with the development of powerful particle colliders. Be it estimating a parameter of the SM or finding evidence of new physics, particle physicists heavily rely on statistical techniques to churn out reliable information from observed data. In this chapter, the statistical methods used in this thesis are discussed. Basic statistical techniques including parameter estimation and method of maximum likelihood are outlined in Section 5.1. The concept of unfolding is introduced along with the techniques used in this thesis are explained in Section 5.2. The main unfolding method used in this analysis, called profile likelihood unfolding is explained in Section 5.3.

5.1 Statistical inference

The main goal of a statistical analysis is to infer key information from the observed data. Commonly used approaches of inference are categorised into frequentist and Bayesian statistics. The main difference between them lies in their interpretation of probability. In statistics, probability is simply the "chance" of an event occurring. In the frequentist approach, probability of an event is interpreted as the occurring frequency of that event, whereas in the Bayesian approach, it is the extent of belief in the occurrence of that event [110].

Statistical problems can be categorised into two types: parameter estimation and hypothesis testing. In parameter estimation, the goal is to determine the "best" possible value of a certain physical quantity or a parameter of a certain mathematical model. It is important to note that the parameter value is always accompanied by an error estimate which quantifies the accuracy of a measurement. A measurement may not be perfect, however the extent of imperfection is concealed in the error values. Therefore, error values play an important role in the interpretation of experimental results.

In hypothesis testing, the main goal is to check if a theory is consistent with the observed data. For this case, the answer is not a numerical value, instead, it is a statement implying how confident we are with the consistency of a theory based on the observed data. In reality, hypothesis testing and parameter estimation are not totally independent of each other. Some problems require estimating a parameter in order to test a hypothesis while in some cases, a parameter is estimated assuming a hypothesis is correct [111]. In this analysis, we focus on parameter estimation. In case of frequentist statistics, method of least squares and method of maximum likelihood are widely used for parameter

estimation. In this thesis, the maximum likelihood method is used which is discussed in the next section.

Method of maximum likelihood

Statistical models are mathematical expressions that relate observations to underlying hypothesis. These models are characterised by a set of parameters. The possibilities of different outcomes, given a mathematical model, are described by probabilities. Another important concept which connects observations to parameters, is the likelihood. Likelihood is a measure of how likely a set of parameters can describe the actual observed data.

Consider a random variable x distributed according to a probability distribution function $f(x;\theta)$ where $f(x;\theta)$ represents our assumed hypothesis. Now suppose the functional form of $f(x;\theta)$ is known, however, some parameters are unknown. Suppose a measurement is performed yielding n independent values, denoted by $x_1, x_2, ..., x_n$. The probability of observing this particular set of values is given by the product of the individual probabilities for each value, as given in Eq. (5.1).

$$P(x_1, x_2, ..., x_n, \theta) = \prod_{i=1}^{n} f(x_i, \theta) dx_i \tag{5.1}$$

Now if the assumed hypothesis is correct, the probability of observing these values should be high. This concept leads to the *likelihood function* $L(\theta)$, which is defined as:

$$L(\theta) = L(\theta | x_1, x_2, ..., x_n) = \prod_{i=1}^{n} f(x_i, \theta) dx_i \tag{5.2}$$

Here, $L(\theta)$ measures how likely it is to observe the set $x_1, x_2, ..., x_n$, given the parameter θ. Unlike a standard probability function, the likelihood function treats the data as fixed and considers θ as the variable.

The objective of maximum likelihood estimation is to estimate the value of θ that makes the observed data most probable. In other words, we search for the parameter values that maximize the likelihood function, a process referred to as the method of maximum likelihood or maximum likelihood fitting. In practice, it is often more convenient to maximize the logarithm of the likelihood function, known as the log-likelihood. This transformation simplifies the product of probabilities in the likelihood function into a sum, which is mathematically easier to handle while preserving the location of the maximum.

5.1.1 Profile likelihood fitting

The profile likelihood approach is built upon the likelihood function described above. In complex problems in high energy physics, the statistical models have many parameters that directly or indirectly affect the measurement. Out of those, at least one parameter is central to the underlying hypothesis, such as the signal strength, which should be estimated. This parameter is called parameter of interest (POI). Furthermore, there are additional parameters that influence the measurement but don't necessarily need to be estimated, for instance, systematic uncertainties. These are called nuisance parameters (NP). Including all these parameters, the likelihood function in Eq. (5.2) can be written as a function of POI (μ) and NP vector (σ) as,

$$L(\theta) = L(x; \mu, \sigma) \tag{5.3}$$

In reality, there are hundreds of NPs affecting the measurement and in such cases it becomes difficult to solve a multidimensional likelihood function. A simpler way is to use a profile likelihood function which is based on the likelihood function but focuses on the POI while considering the effects of the NPs. The profile likelihood function maximises the likelihood over all NPs for each fixed value of the POI. By doing this, the NPs are replaced with their corresponding maximum likelihood estimators $\hat{\sigma}$ calculated by the fit and the likelihood becomes a function of μ. The process of using fixed values of σ for a given μ is called "profiling" out the NPs. This reduces the dimensionality of the likelihood. This procedure of estimating the POIs is called profile likelihood fitting.

It is important to realise that the efficiency of estimating the parameters depends highly on the observed data. In particle physics, this fact imposes a certain level of expectations from particle detectors. Even though our detectors are well-developed and efficient, there are some limitations that may lead to some sort of distortion in observations which may lead to unrealiable results. In order to solve this, physicists often use mathematical techniques to remove the distortions caused by detectors. These techniques fall under the category of *unfolding*.

5.2 Unfolding: basic concept

In any experiment, the accuracy of measurement depends heavily on the performance of the apparatus used. In particle physics, an ideal detector would be able to capture the original and complete information of collisions and accurately preserve the true shape of distributions for any physical observable. However, in reality, the data received from a detector is distorted or smeared due to effects such as limited acceptance and finite resolution of the detector. These distortions may lead to incorrect inferences and therefore, need to be removed.

Unfolding is a mathematical technique used to remove the detector distortions and estimate the original fine structure of observed data. This technique is also called desmearing or deconvolution. The basic concept is to approximate ideal unsmeared data from the observed data with the help of a response matrix. A response matrix is a mathematical construct that characterises the smearing effects introduced by the detector. More details are discussed in Section 5.2.1.

Unfolding is essential in various contexts of particle physics. It enables results to be combined or compared with those from other experiments that may have different levels of smearing. Additionally, it is not always practical to present results alongside their response matrices. Unfolding becomes crucial when comparing experimental results directly to theoretical predictions without accounting for experimental distortions. Furthermore, non-distorted data is needed when fitting specific parameters to data for tuning Monte Carlo simulations [112].

5.2.1 Methodology

There are certain conventions regarding the entities used in unfolding problems within the particle physics community. The distribution of a physical observable obtained from data recorded by the detector is called detector-level distribution. On the other hand, the truth-level distribution represents data that we should have obtained with an ideal detector. This is generated from MC simulations without applying any detector simulations. In unfolding, we also make use of simulated distribution that

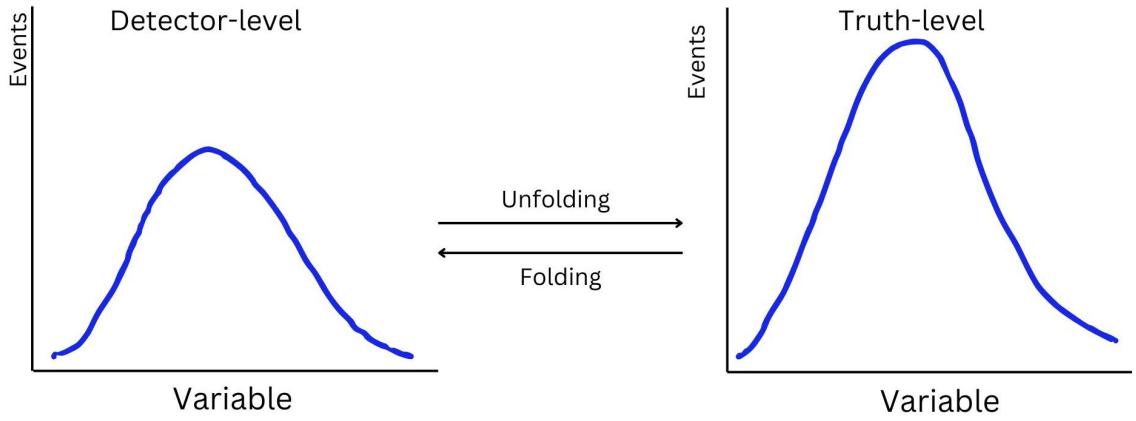


Figure 5.1: An illustration visualising the method of unfolding. The left diagram is an illustration of a histogram obtained from the detector information, called detector-level and the right illustration shows a histogram that is expected form an ideal detector, called truth-level. The difference in the two diagrams depicts the smearing caused by the detector. The procedure to obtain the truth-level information based on the detector-level information is called as unfolding. The reverse procedure is called folding.

includes detector simulation. Since this so-called reconstructed-level distribution is supposed to mimic actual observed data, it is used to validate an unfolding method. Detector-level and reconstructed-level distributions are conceptually the same when discussing unfolding related quantities. Technically, unfolding is a method of estimating the truth-level by correcting detector effects present at detector- or reconstructed-level.

The reconstructed distribution $\vec{x}$ is related to the truth-level distribution $\vec{y}$ by a response matrix R as shown in Eq. (5.4). Here $\vec{b}$ represents backgrounds. An illustration visualising the concept of unfolding is shown in Fig. 5.1. In particle physics problems, data is generally organised into histograms with finite bins. In this case the above given relation is given by Eq. (5.5).

$$\vec{x} = R \cdot \vec{y} + \vec{b} \tag{5.4}$$

$$x_i = \sum_{\mathrm{j}\,=\,1}^{\mathrm{bins}} \mathcal{R}_{ij} y_j + \vec{b} \tag{5.5}$$

A response matrix quantifying detector effects is computed from simulated data. The quantities used to construct the response matrix are migration matrix and two correction factors namely acceptance and efficiency. In order to calculate these quantities, the first step is to define a fiducial truth-level volume. It is basically a subset of the available truth-level phase space, that we are interested to unfold to. For this analysis, fiducial volumes are defined at both parton and particle levels using a set of pre-decided criteria. A schematic showing reconstructed-level and fiducial truth-level volumes is given in Fig. 5.2. Per-bin acceptance adjusts the number of reconstructed events by the fraction of events that are also present at the fiducial truth-level. It is defined as the ratio of events present at the reconstruction- and truth-levels to the total number of events at the reconstruction-level. Acceptance gives an idea of how well the reconstructed data corresponds the true data. The corrected number of

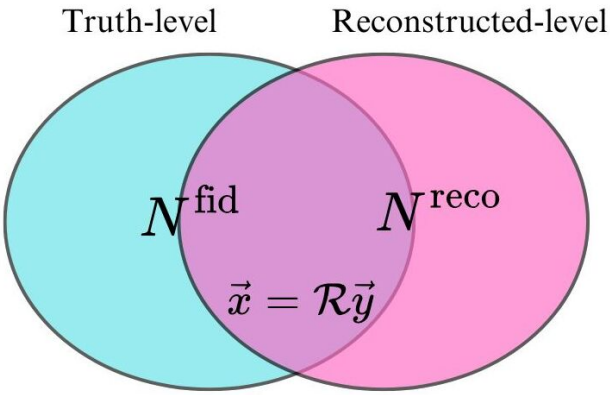


Figure 5.2: A schematic picture showing reconstructed-level volume in blue and truth-level volume in pink.

events at the reconstructed-level is given as,

$$N_i^{\text{reco}} = x_i \cdot a_i \tag{5.6}$$

Per-bin efficiency adjusts the number of truth-level events by the fraction of truth-level events that are also found at the reconstructed-level. It is defined as the ratio of a number of events present at both the reconstruction-level and the truth-level to the total number of events at the fiducial truth-level. This correction factor gives an idea of the detector's efficiency to reconstruct true events. The corrected number of events at the fiducial truth-level is given as,

$$N_j^{\text{fid}} = y_j \cdot \epsilon_j \tag{5.7}$$

The migration matrix describes the bin-to-bin migrations between truth-level and reconstruction-level histograms. For instance, M_{ij} represents the fraction of events found in bin i at the reconstruction-level while being created in bin j at the truth-level. A migration matrix with maximum diagonal components indicates that most events are reconstructed in the same bin in which they were generated. Equation (5.5) can be re-written as,

$$N_i^{\text{reco}} = \frac{a_i}{\epsilon_j} \cdot \sum_{\text{j}=1}^{\text{bins}} M_{ij} \cdot N_j^{\text{fid}} \tag{5.8}$$

Mathematically, the idea of unfolding is to solve Eq. (5.4) for given $\mathcal{R}$, x and b. The resultant values of y can be interpreted as determined true number of events at the truth-level. One would notice a simple way to find the estimators by inverting the response matrix as shown in Eq. (5.9).

$$\vec{y} = \mathcal{R}^{-1}(\vec{x} - \vec{b}) \tag{5.9}$$

Although matrix inversion method is easy to implement, it is a strategy that one should avoid because of its limitations: in some situations, the response matrix is non-invertible then Eq. (5.9) becomes ill-posed. Even though inversion is possible, there are possible statistical fluctuations in the

observed data that may cause negative entries in the inverse matrix. This leads to negative number of events in the unfolded distribution which is unrealistic. When a response matrix acts on a true spectrum, it distorts any fine structure present at the truth-level. Despite that, some residue of this fine structure still remains in the reconstructed spectrum [113]. The inverted matrix, acting on the measured data, assumes its statistical fluctuations are the residual fine structure and restores it. In this way, statistical fluctuations are amplified in the unfolded distribution [114] which is undesirable.

In order to overcome the limitations of matrix inversion, alternate unfolding methods are used in high energy physics. Some of the methods are summarised in [115]. In this thesis, the iterative Bayesian unfolding (IBU) and profile-likelihood unfolding (PLU) are discussed.

5.2.2 Iterative Bayesian unfolding

D'Agostini [116] proposed a method called iterative Bayesian unfolding (IBU) which makes use of Bayes' theorem. To describe this method, consider true events as *causes* ($C_i, i = 1, 2, ..., n_C$) and reconstructed events as *effects* ($E_j, j = 1, 2, ..., n_j$). The conditional probability that a cause C_i gave rise to effect E_j, denoted by $P(C_i|E_j)$, is given by Bayes' theorem:

$$P(C_i|E_j) = \frac{P(E_j|C_i)P(C_i)}{P(E_j)} \tag{5.10}$$

where, $P(E_j|C_i)$ can be interpreted as probability of reconstructed event given true event which is the element M_{ij} of the migration matrix. Consequently, $P(C_i|E_j)$ can be identified as the unfolding matrix. One can determine the number of events ($\hat{n}$) due to cause C_i as

$$\hat{n}(C_i) = \frac{1}{\hat{\epsilon}_i} \sum_{j=1}^{n_E} \hat{n}(E_j)P(C_i|E_j). \tag{5.11}$$

It is important to note that the total number of events due to all causes and all effects are equal because only migration effects are considered so far. By dividing both the sides of Equation 5.11 by total number of events, we obtain

$$P(C_i) = \frac{1}{\hat{\epsilon}_i} \sum_{j=1}^{n_E} P(E_j)P(C_i|E_j). \tag{5.12}$$

Here, $P(C_i)$ is the unfolded distribution. This technique is implemented in the `RooUnfold` [117] software package. The steps performed in iterative Bayesian unfolding to find $P(C_i)$, are shown in Figure 5.3 and explained below:

- An initial guess $P_0(C_i)$ is made and inserted into Equation 5.12. $P(E_j)$ is obtained from the MC reconstructed distribution. The solution provides $P(C_i|E_j)$ which is the unfolding matrix.
- The obtained $P(C_i|E_j)$ is used in Bayes' theorem (Equation 5.10) to get a value of $P(C_i)$ which is different from the initial guess.
- This process is repeated for number of iterations specified by user.

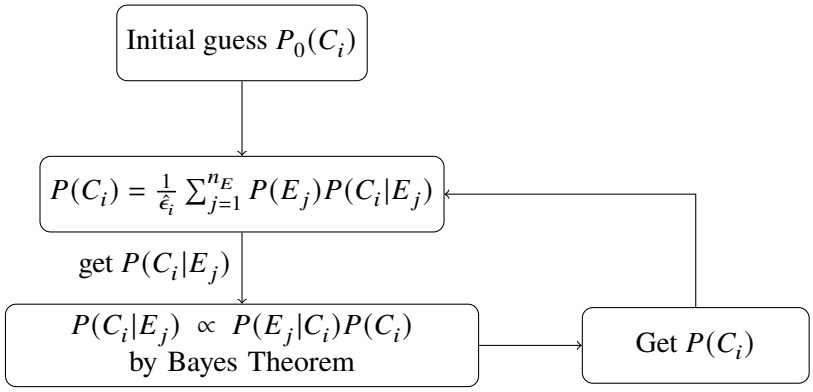


Figure 5.3: Flow chart showing steps followed in iterative Bayesian unfolding

5.3 Profile likelihood unfolding

The method of profile likelihood unfolding (PLU) is based on deconvoluting the observed distribution using profile likelihood fitting (described in Section 5.1.1). In PLU, the underlying mathematical model is defined in such a way that the parameter of interest represents the unfolded distribution. In this analysis, PLU is implemented to calculate the differential cross-section of the tZq production.

5.3.1 Formulation

The first step is to construct a mathematical model that embodies the observed data along with predicted values of model parameters. In high-energy physics analyses, the observed data is typically organised into histograms. As a result, the observable is represented not as individual values but as the content of histogram bins. A likelihood function is constructed from two types of measurements: primary and auxiliary. Primary measurements are based on observed data ($\vec{n}$) and are associated with parameters of interest ($\vec{\mu}$), while auxiliary measurements are existing measurements focusing on systematic variations ($\vec{a}$) of data and are linked to nuisance parameters ($\vec{\sigma}$).

A typical likelihood function for binned distributions is shown in Eq. (5.13). For bin i, the conditional probability of observing n_i events under the prediction of S_i signal events and B_i background events can be given using a Poisson function. Since the bins are independent, the product of their Poisson probabilities forms the first term of the likelihood function. It is a common practice to introduce a signal strength parameter μ that controls the normalisation of signal. The second term is a product of Gaussian probability functions associated with auxiliary measurements.

$$\mathcal{L}(\mu, \vec{\sigma}; \vec{n}, \vec{a}) = \prod_{i=1}^{\text{total bins}} P(n_i|\mu S_i(\sigma) + B_i(\sigma)) \cdot \prod_{j}^{\text{systs}} G(a_j|\sigma_j) \tag{5.13}$$

In PLU, a modified likelihood function is used in which instead of having one parameter of interest representing overall normalisation of signal events, there is a normalisation vector with dimensions equal to number of bins in the truth histogram (truth-level bins). The modified likelihood function is defined in Eq. (5.14).

$$\mathcal{L}(\vec{\mu},\vec{\sigma};\vec{n},\vec{a}) = \prod_{i=1}^{\text{total bins}} P(n_i|\vec{\mu}\vec{S}_i(\sigma)+B_i(\sigma)) \cdot \prod_{j}^{\text{systs}} G(a_j|\sigma_j) \tag{5.14}$$

Once this mathematical model is ready, the next step in PLU focuses on the truth histogram. Each truth-level bin is multiplied by the response matrix resulting in sub-histograms per truth-level bin. In unfolding terminology, this is known as *folding* the truth-level bins. Each sub-histogram or folded truth bin is assigned a normalisation factor or scale factor which is a free parameter in the fit. Finally, a profile likelihood fit of the reconstructed-level distribution is performed on the sum of the sub-histograms yielding best fit values of the normalisation factors. It is important to note that fitting the normalisation of the sub-histograms is equivalent to fitting the normalisation of the truth-level bins, which is the desired unfolded result. In this analysis, the normalisation factors are interpreted as differential cross-section.

Including systematic variations

The Gaussian term in the likelihood is called a constraint term because it restricts the likelihood function based on known information, for instance, systematic uncertainties. Given the fact that there can be many systematic uncertainties, it is difficult to handle nuisance parameters all with different values. To simplify the handling, a standard convention is used where the nuisance parameters are redefined by scaling each Gaussian in such a way that its mean is 0 and standard deviation is 1.

In principle, one can optimise the likelihood based on each value of the nuisance parameter in order to capture the effect of that particular systematic variation. However, in practice, this is computationally intensive. Therefore, for each NP we use template response matrices in the form of nominal, up variation and down variation. A nominal response matrix is built from standard distributions for which the NP value is equal to its mean, i.e. 0. The up (down) variation matrix is built from distributions corresponding to NP value +1 (−1). Based on these inputs, a continuous variation for each NP is calculated using interpolation and extrapolation techniques.

For a kinematic variable X in bin i, a differential cross-section can be calculated by dividing the unfolded yields by bin width ΔX and luminosity $\mathcal{L}$ as presented in Equation 5.15.

$$\frac{\mathrm{d}\sigma_{\text{PLU}}}{\mathrm{d}X^i} = \frac{1}{\mathcal{L}\cdot\Delta X^i} N^i_{\text{unf}} \tag{5.15}$$

where unfolded bin contents N_{unf} are calculated by scaling the truth-level bins with the best fit values of the normalisation factors.

5.3.2 Implementation

A tool called `HistFactory` [118] is used to construct likelihood functions from histograms based on `ROOT`. In this analysis, a software framework called `TRExFitter` [119] is used to perform profile likelihood fitting. `TRExFitter` builds binned template likelihood functions and performs statistical analysis using tools such as `HistFactory`, `RooStats` [120] and `RooFit` [121].

The input to `TRExFitter` consists of template histograms, which are organised into two main categories:

- **Regions**: A region refers to a subset of the data defined by a set of specific event selection. For instance, a region can be defined by selecting events passing a certain neural network cut. The selection criteria for regions should be decided in such a way that the regions are disjoint from one other. This avoids double counting of events. The fitting is performed in each region.

- **Samples**: A sample corresponds to a specific process such as a signal process or a background process. Each sample is represented by a template histogram describing the expected distribution of events for that process in a given region.

Along with the template histograms, response matrices corresponding to nominal and per systematic variations must be also provided. In PLU, the fit calculates all the free and constraint parameters which includes the following:

- Best-fit values of the normalisation factors representing each truth bin.

- Best-fit values of any background normalisation factors for which specific regions might have been defined.

- Statistical and systematic uncertainties on the best-fit values quantifying the precision of the fit. `TRExFitter` uses a tool called `MINOS` [122] to compute the uncertainties.

- A covariance matrix quantifying the correlation between all parameters.

5.3.3 Advantages of PLU

There are some advantages of using profile likelihood unfolding compared to other unfolding methods. Mainly, PLU does not involve any matrix inversion, hence, the chance of amplifying statistical fluctuations is minimum. Since the `TREXfitter` framework is optimised for handling systematic uncertainties, it is quite easy to utilise it for unfolding as well. The formalism also allows adding free-floating normalisation factors for specific backgrounds that can be constrained by adding control regions in the fit. The *tZq* analysis also includes measurement of inclusive cross-section which is also done using the `TREXfitter` framework. Therefore, using PLU for differential measurement allows to have a common statistical framework for the complete analysis.

Asimov dataset

The idea of Monte Carlo simulations (as discussed in Section 4.4) is to repeat random samplings from distribution functions and average the results. In a complex physics problem involving many systematic variations this becomes computationally expensive. Therefore, high energy physics widely uses the concept of an Asimov dataset. It is based on a story by Isaac Asimov where the results of an election were based on a single vote. The idea is to have a representative dataset rather than an ensemble of many simulated datasets [123].

In profile likelihood problems, an Asimov dataset is defined such that the ML best-fit values of model parameters are equal to their generated (true) values. In the context of this analysis, a profile likelihood fit performed on the Asimov dataset should yield $\hat{\mu} = 1$ for the POIs. The uncertainties on μ should be close to the uncertainties obtained for the corresponding fit to observed data. Moreover,

the values of the NPs should be 0 with a variance equal to 1. If the variance is not equal to 1, then the fit is said to have constrained the NP. This procedure is referred to as the Asimov fit in this analysis.

In this analysis, the Asimov dataset is used to validate the profile likelihood unfolding method. The fit values of per-bin normalisation factors and nuisance parameters are expected to align with the true values, ensuring the consistency of the unfolding method.

5.4 Parton and particle level unfolding

For measuring differential cross-section, several kinematic variable distributions are unfolded (discussed in Section 6.1). The MC simulated data can be accessed at three different stages of particles: parton level, particle level and detector level [124]. These stages correspond to the specific steps completed in the MC simulation chain.

The parton level refers to the stage where hard interactions are completed in addition to initial-state and final-state radiation (in case of top quarks) but before the parton shower or hadronisation. Moreover, kinematic distributions of parton level objects, such as the top quark, are directly available from the MC chain.

The particle level refers to the stage after parton shower and hadronisation. The particle level objects are stable particles with longer lifetimes produced in the final-state. In addition, kinematic distributions of objects such as the top quark can be obtained by reconstructing the top quark from stable final state particles. The detector level refers to the simulated data after passing through detector simulation.

Parton and particle level correspond to truth level information whereas detector level data correspond to reconstructed level or observed data. In this analysis, detector level data is unfolded to both parton and particle levels. Furthermore, certain selection is applied to both the levels in order to constitute a fiducial volume. The object definitions tailored for the tZq analysis along with the fiducial phase space are described in the following sections.

5.4.1 Object and phase space definition at parton level

The parton-level fiducial volume is constructed from the Z boson decaying into opposite sign same flavour lepton pairs of e^- or μ^- ($Z \rightarrow e^+ e^-$, $\mu^- \mu^+$) and the on shell t quark decaying into leptons (e^- or μ^-). The kinematic distributions of the t quark are reconstructed after final state radiation and immediately before decay. The leptons are reconstructed before any radiation. Events containing tau leptons coming from either Z boson decay or W boson decay from the t quark, are removed from the definition irrespective of their successive decay. The invariant mass of the lepton pair should be within ± 15 GeV of the mass of the Z boson as per the PDG value of m_Z [125]. As shown in Fig. 5.4, around 66% events lie in this chosen invariant mass window. This criterion follows that used in the $t\bar{t}Z$ analysis [126] and is motivated by wanting to measure the $t\ell\ell q$ cross-section with an on-shell Z boson, tZq. The kinematic distributions are unfolded to this fiducial phase space at the parton-level, referred to as parton-level unfolding.

5.4.2 Object and phase space definition at particle level

This analysis follows the guidelines for particle level objects recommended by the LHC collaborations [127]. Particle level volume is defined by stable objects having a mean lifetime greater than

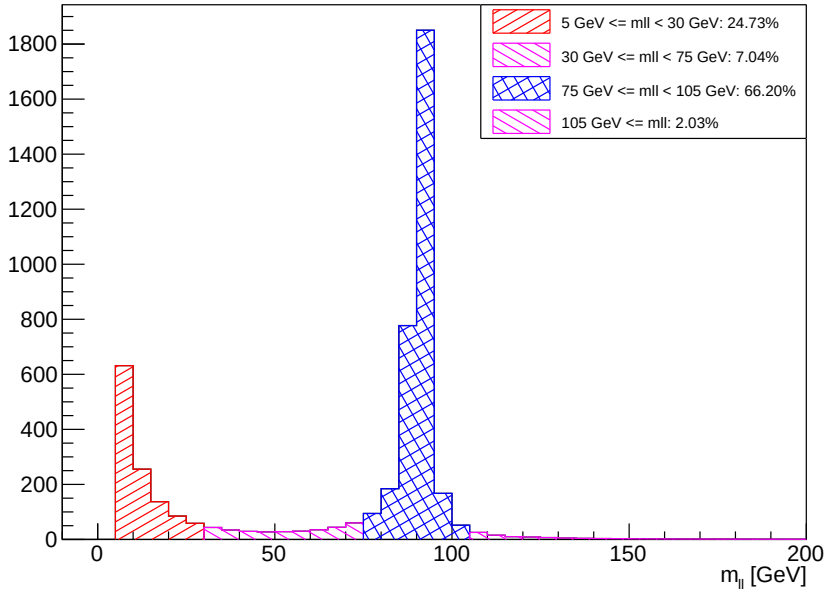


Figure 5.4: Distribution of invariant mass of two leptons not coming from the top-quark, at the parton-level truth. The different invariant mass windows are highlighted, and it is observed that 66% events lie in the ±15 GeV window.

0.03 ns within an observable pseudorapidity range. The particle-level leptons (e^- and μ^-) are accepted if they originate from the t quark or Z boson and not from any other hadrons or quarks. The candidate leptons having the closest invariant mass to m_Z are identified as leptons originating from Z while the remaining lepton is associated with t decay. Within the range $\Delta R < 0.1$, the four-momenta of these leptons are summed with the four-momenta of radiated photons. The missing transverse energy is calculated from the four-vector sum of neutrinos which are selected if they do not originate from hadrons or quarks (same condition as the charged leptons)

Jets are defined with the anti-k_t algorithm, including all stable particles within $R = 0.4$, excluding the leptons used to define selected leptons and neutrinos coming from Z or t decay. If b-hadrons with $p_{\mathrm{T}} > 5$ GeV are found in the MC decay chain, they are clustered in the stable-particle jets with their energies set to zero. Particle-level jets containing one or more of these b-hadrons are considered to originate from a b-quark. A technique called ghost-matching is used to identify b-jets. Low p_{T} tracks called "ghosts" found around the position of b-hadrons are added as inputs to the jet clustering algorithm. If the ghost track is clustered into a jet, it indicates that the jet originated from a b-hadron. More details regarding ghost matching can be found in [128].

The fiducial volume is defined in such a way that it is close to the reconstructed signal region. Data is unfolded to this fiducial volume at the particle level to obtain differential cross-sections, called particle-level unfolding. The selection criteria applied on the particle level objects are detailed in Table 5.1. The $|m_{\ell\ell} - m_Z|$ selection applies to at least one OSSF lepton pair, while the $\min(m_{\ell\ell})$ cut applies to any OSSF lepton pair. It therefore only affects $\mu\mu\mu$ and eee events, where there are two possible combinations of $m_{\ell\ell}$.

Table 5.1: Particle-level selection criteria for unfolding

Variable	Selection
$N_\ell(e,\mu)$	3, $\geq$ 1 OSSF lepton pair
$p_T(\ell_1,\ell_2,\ell_3)$	$\geq$ (27, 15, 10) GeV
$\lvert\eta_e\rvert$	$< 2.47, \notin [1.37, 1.52]$
$\lvert\eta_\mu\rvert$	< 2.5
$\lvert m_{\ell\ell} - m_Z\rvert$	< 10 GeV
$\min(m_{\ell\ell})$	> 10 GeV
$m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$	> 20 GeV
p_{T}(jets)	> 25 GeV
$\lvert\eta$(jets)$\rvert$	< 4.5
N_{jets}	$2-5$
$\lvert\eta$(b-jets)$\rvert$	< 2.5
$N_{\text{b-jets}}$	$\geq$ 1 (ghost-matched)

CHAPTER 6

Differential cross-section estimation

In this chapter, the results of the profile likelihood unfolding in the form of differential cross-sections are presented. The preparatory steps including choosing the variables to unfold and fixing their bin edges, are discussed in Section 6.1 and Section 6.2, respectively. The PLU fit setup is detailed in Section 6.3. Various validation tests performed on the unfolding method are described in Section 6.5. The differential cross-section results for the Asimov dataset are discussed in Section 6.7.

6.1 Variables to unfold

The differential cross-section is measured as a function of kinematic variables of particles associated with tZq production. That includes top quark, Z boson with its associated leptons and the forward jet. The forward jet is defined as the non b-jet yielding the highest invariant mass with the leading b-jet. Variables related to the fundamental kinematics of these particles are included in the list of unfolding variables. Another reason is that the variables related to Z boson and top quark are sensitive to tZ coupling. The list also includes variables sensitive to EFT, for instance $p_{\mathrm{T}}(t)$ and $p_{\mathrm{T}}(Z)$.

The tZq EFT study is based on parametrising and comparing the EFT effects with the Standard Model. For each Wilson coefficient, the events are reweighted at four discrete points. The histograms corresponding to reweighted events are overlayed with the histogram of the Standard Model events. This is done for all the kinematic variables, and it is observed that some variables show quite a significant difference, which means these variables could help study tZq EFT and hence are included in the list of unfolding variables. Overlay histograms[1] for one of the WCs, Re $\left(c_{tW}\right)$, are given in Fig. 6.1. The complete list of unfolding variables and their description is given in Table 6.1. Out of them, those unfolded at the particle level and those unfolded at the parton level are listed in the Table 6.2. At parton level, information was available for quarks but not for jets, hence, variables related to the forward jet are not considered at parton level.

[1] The EFT studies including the overlay plots are performed by Can Suslu, a Master student in the group

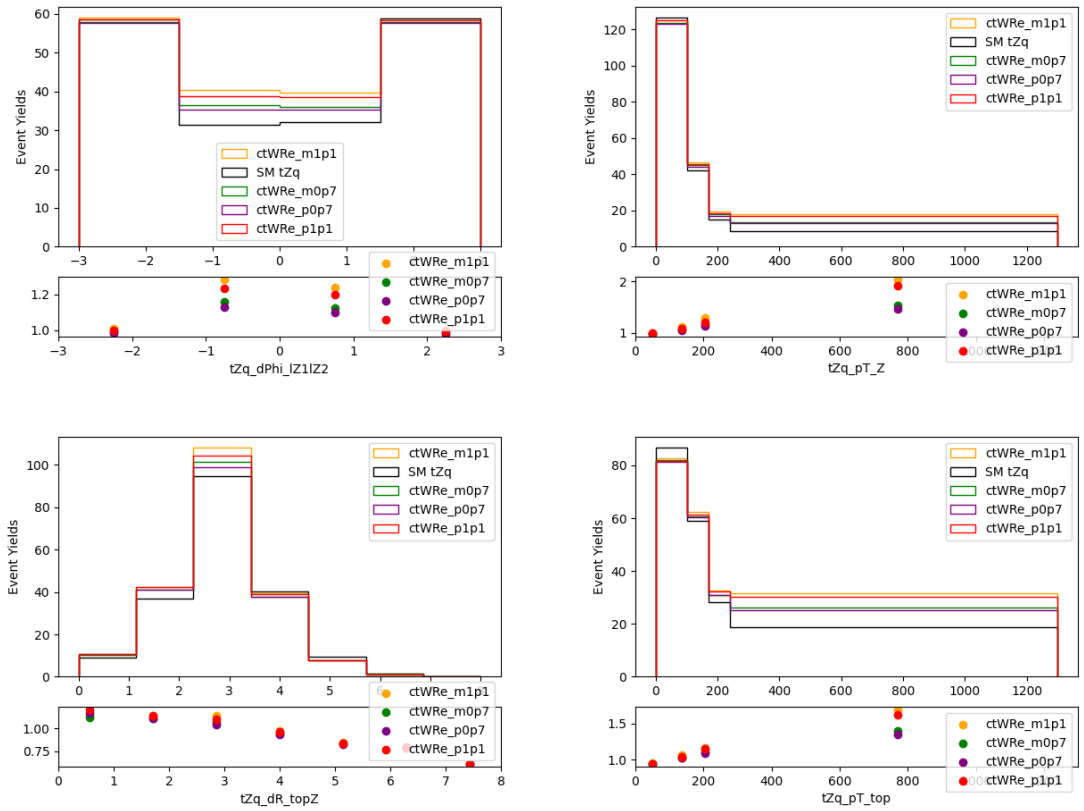


Figure 6.1: Variables showing visible EFT effects. The plots show events with Standard Model weight and events with weight corresponding to different values of Re $\left(c_{tW}\right)$

Table 6.1: Description of variables that are unfolded for measuring differential cross-sections

Symbol	Description
$p_{\mathrm{T}}(t)$	Transverse momentum of the reconstructed t-quark
$\lvert\eta(t)\rvert$	Pseudorapidity of the reconstructed t-quark
$p_{\mathrm{T}}(Z)$	Transverse momentum of the reconstructed Z-boson
$\lvert\eta(Z)\rvert$	Pseudorapidity of the reconstructed Z-boson
p_{T}(f-jet)	Transverse momentum of the forward jet
$\lvert\eta$(f-jet)$\rvert$	Pseudorapidity of the forward jet
$\Delta R(t,Z)$	Angular distance between the t-quark and the Z-boson
$\Delta R(t,$f-jet)	Angular distance between the t-quark and the forward jet
$\lvert\Delta\phi(\ell_1^Z,\ell_2^Z)\rvert$	Azimuthal angular distance between the two leptons originating from the Z-boson

Table 6.2: List of variables unfolded at parton level and particle level

Parton level	Particle level
$p_{\mathrm{T}}(t)$	$p_{\mathrm{T}}(t)$
$\lvert\eta(t)\rvert$	$\lvert\eta(t)\rvert$
$p_{\mathrm{T}}(Z)$	$p_{\mathrm{T}}(Z)$
$\lvert\eta(Z)\rvert$	$\lvert\eta(Z)\rvert$
$\lvert\Delta\phi(\ell_1^Z, \ell_2^Z)\rvert$	$\lvert\Delta\phi(\ell_1^Z, \ell_2^Z)\rvert$
$\Delta R(t, Z)$	$\Delta R(t, Z)$
	p_{T}(f-jet)
	$\lvert\eta$(f-jet)$\rvert$
	$\Delta R(t,$ f-jet)

6.2 Bin optimisation

As already discussed, the binned likelihood function required for PLU is constructed from template histograms. Individual fits are performed for each observable or unfolding variable in this case. For each variable, the fit requires histograms of the following entities:

- Weighted reconstructed level histograms for signal and background processes
- Weighted truth level histogram for the signal
- Migration matrices for the nominal signal distribution and its individual systematic variation
- Acceptances for the nominal signal and its individual systematic variation
- Efficiency for the nominal signal and its individual systematic variation

For each observable, the binning of the histograms is optimised and kept consistent across all entities for that specific observable. To determine an optimised binning, the process begins by identifying the upper limit of the variable being considered. This upper limit is chosen based on the original distribution of the variable, with the aim of ensuring that the majority of events are included within the defined range. Any events that fall outside this specified upper limit are referred to as overflow events. These overflow events are added to the last bin of the respective histogram. One must choose an optimum binning for the observables as it plays an important role in the components of the migration matrix.

The selected phase space is initially divided into fine bins, for instance 40 bins. Then, the fine bin contents are merged into coarser bins so that each bin has more or less similar number of weighted events. The number of bins in the final distribution is decided based on the diagonal of the migration matrix. The goal is to achieve an almost diagonal migration matrix with few migrations, while preserving any structure. Once the binning is decided, the overflow bin content is merged with the last bin for all the unfolding related quantities of all observables. For most of the variables, the best configuration was achieved for four bins. However, for $p_{\mathrm{T}}(t)$, $\lvert\eta(t)\rvert$ and $\Delta R(t, Z)$, four bins resulted in a worse migration matrix which is why a three-bin configuration is decided. The binning is same for observables unfolded at both parton and particle levels. The optimised binnings are given in Table 6.3.

Figure 6.2 shows the optimised bin distribution of the unfolding variables including the composition of signal and background processes.

Table 6.3: Optimised binnings of unfolding variables

Variable	Binning
$p_{\mathrm{T}}(t)$ [GeV]	[0,95,190,450]
$\lvert\eta(t)\rvert$	[0,0.70,1.70,4]
$p_{\mathrm{T}}(Z)$ [GeV]	[0,45,80,135,450]
$\lvert\eta(Z)\rvert$	[0,0.70,1.40,2.10,4.00]
$\lvert\Delta\phi(\ell^Z_1, \ell^Z_2)\rvert$	[0,1.30,2.00,2.70,4.00]
$\Delta R(t, Z)$	[0,2.60,3.50,7.00]
p_{T}(f-jet) [GeV]	[0,65,100,170,400]
$\lvert\eta$(f-jet)$\rvert$	[0,2.00,3.00,3.50,4.50]
$\Delta R(t$, f-jet)	[0,3.40,4.20,5.20,7]

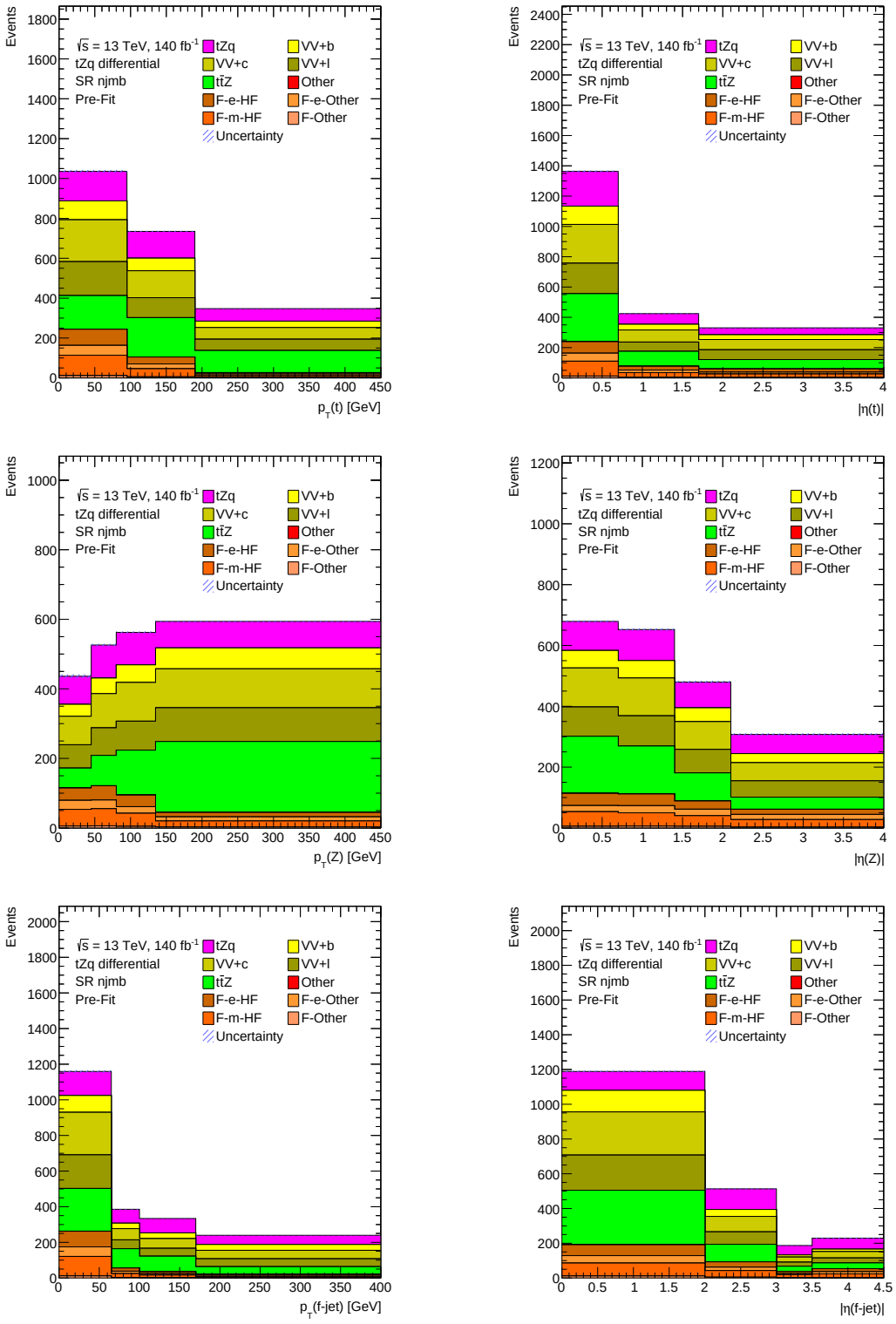


Figure 6.2: Pre-fit distributions of variables that are unfolded. The binning for each variable is optimised for unfolding. These are stack plots that describe the composition of different processes. The pink colour in the plots represents *tZq*.

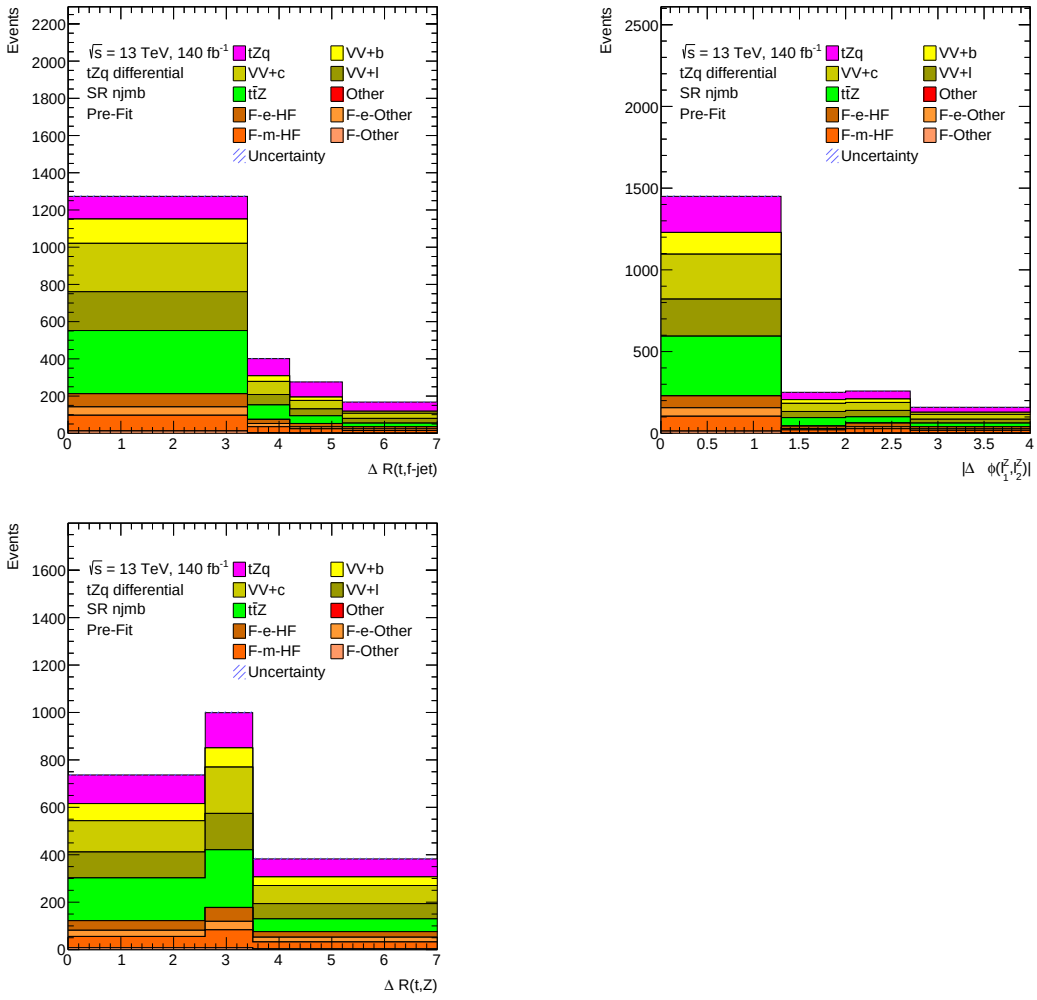


Figure 6.2: Pre-fit distributions of variables that are unfolded. The binning for each variable is optimised for unfolding. These are stack plots that describe the composition of different processes. The pink colour in the plots represents *tZq*.

6.3 PLU fit setup

As discussed in Section 5.3.2, template histograms for a profile likelihood unfolding fit are organised in regions and samples. A profile likelihood unfolding fit differs from a standard profile likelihood fit primarily in its definition of signal regions. In a standard fit, a signal region consists of a distribution of events that meet selection criteria chosen to extract higher signal to background ratio. This region is aimed towards extracting signal normalisation. On the other hand, an unfolding fit extends this concept by including not only the distribution of events passing signal-specific criteria but also additional histograms representing migration matrix, acceptance, and efficiency. A region containing all the aforementioned histograms can be used to extract signal normalisation i.e. unfolding normalisation factors.

6.3.1 Signal regions

In this analysis, a binary neural network (NN) was optimised in which signal events are trained versus all the background events. The NN produces a distribution of scores from 0 to 1 given to each event. These scores represent the possibility of an event to be signal-like or background-like. These studies were performed by colleagues in the analysis team. The NN distribution is shown in Fig. 6.3.

For the differential measurement, the NN distribution is split into two signal regions by applying a cut: $O_{\mathrm{NN}}(\mathrm{SR}) > \mathrm{NNcut}$ and $O_{\mathrm{NN}}(\mathrm{SR}) < \mathrm{NNcut}$. $O_{\mathrm{NN}}(\mathrm{SR}) > \mathrm{NNcut}$ region contains a higher signal to background ratio while $O_{\mathrm{NN}}(\mathrm{SR}) < \mathrm{NNcut}$ does not. However, the signal events in the latter region are still significant and should be utilised for extracting the unfolding normalisation factors. In an unfolding fit, the only way to achieve this is by incorporating this region as a signal region. The complete selection requirements including the baseline criteria as described in Section 4.2, for the signal regions are summarised in Table 6.5.

Deciding the NN cut

The NN cut value is decided on the basis of signal significance. A range of cuts from 0.0 to 1.0 in steps of 0.01 were tested, and for each cut, the number of signal events (S) and background events (B) were recorded to compute the signal significance. The signal significance is calculated as follows:

$$\text{Signal significance} = \frac{S}{\sqrt{S+B}} \tag{6.1}$$

Figure 6.4 shows the cut flow of NN output and corresponding significance for the binary network. The breakdown of the event yields after application of the cut is given in Table 6.4. Events with a NN score higher than the cut value form one of the signal regions in the fit ($O_{\mathrm{NN}}(\mathrm{SR}) > \mathrm{NNcut}$) and the remaining events forms another signal region ($O_{\mathrm{NN}}(\mathrm{SR}) < \mathrm{NNcut}$).

6.3.2 Control regions

The control regions (CR) are defined for mainly two purposes, first is to probe the normalisation of main background processes and second to estimate the non-prompt lepton background.

The CR for the $t\bar{t}Z$ ($\mathrm{CR}(t\bar{t}Z)$) is defined using the same selection criteria as the SRs but with higher number of jets (jet multiplicity) because trileptonic decay of $t\bar{t}Z$ will result into more jets. Therefore,

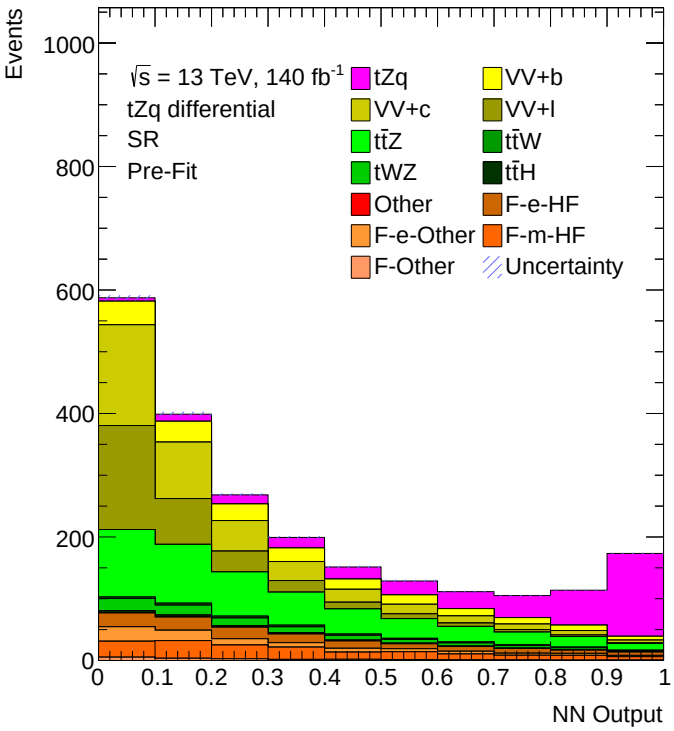


Figure 6.3: Distribution of the binary NN score. The signal tZq, shown in pink, is pushed towards the right side of the distribution due to its high scores as expected.

Table 6.4: Weighted yields of signal and background before and after application of the neural network cut.

	All yields	SR $O_{\text{NN}} > 0.7$
tZq	344.3 ± 1.2	226.3 ± 1.0
$VV + b$	190.0 ± 0.7	25.4 ± 0.2
$VV + c$	403.6 ± 1.6	20.7 ± 0.1
$VV + l$	326.8 ± 1.4	7.6 ± 0.1
$t\bar{t}Z$	476.6 ± 1.7	48.6 ± 0.3
$t\bar{t}W$	19.2 ± 0.1	2.9 ± 0.0
tWZ	86.2 ± 0.3	9.4 ± 0.1
$t\bar{t}H$	14.9 ± 0.1	1.4 ± 0.0
Other	0.9 ± 0.0	0.0 ± 0.0
$t\bar{t}t\bar{t}$	0.3 ± 0.0	0.0 ± 0.0
NPL-e-HF	122.3 ± 0.4	18.4 ± 0.1
NPL-e-Other	81.5 ± 0.3	8.6 ± 0.1
NPL-μ-HF	151.8 ± 0.5	21.1 ± 0.1
NPL-Other	19.1 ± 0.1	1.7 ± 0.1
Total	$2\,237.5 \pm 7.4$	392.2 ± 2.0

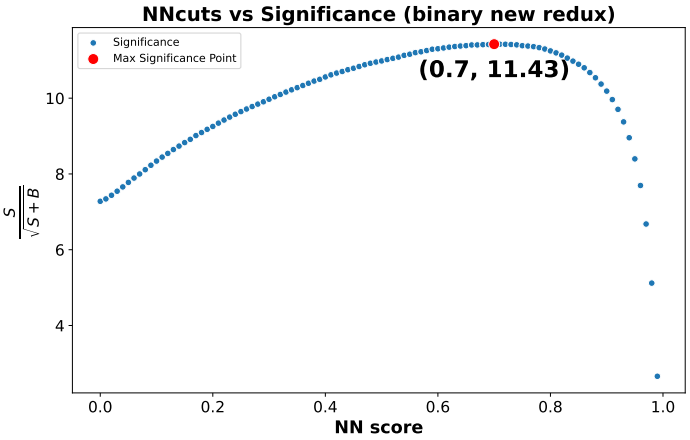


Figure 6.4: Plot of significances obtained after applying a range of cuts on the binary neural network output. The maximum significance of 11.43 is obtained for a cut value of 0.7.

Table 6.5: Definition of the signal regions.

Variable	Preselection	
N_ℓ $(\ell = e, \mu)$	$= 3$	
	≥ 1 OSSF lepton pair	
$p_{\mathrm{T}}(\ell_1, \ell_2, \ell_3)$	$> 27, 15, 10$ GeV	
Sum of lepton charges	± 1	
$\min(m_{\ell\ell})$, applies to any OSSF lepton pair	> 20 GeV	
$\lvert m_{\ell\ell} - m_Z \rvert$, applies to at least one OSSF lepton pair	< 10 GeV	
$m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$	> 30 GeV	
N_{jets} $(p_{\mathrm{T}} > 25$ GeV$)$	2-5	
$N_{b-\mathrm{jets}}$@85%	1-2 (no $2j2b$)	
	$O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$	$O_{\mathrm{NN}}(\mathrm{SR}) < 0.7$
O_{NN}	> 0.7	< 0.7

six or more jets are required, out of which one is a b-tagged jet. The CR for diboson (CR(VV)) has the same jet multiplicity requirement as the SR but with zero b-tagged jets. This region is enriched with light jets. The selection requirements for CR($t\bar{t}Z$) and CR(VV) is summarised in Table 6.6.

Table 6.6: Definition of the signal and control regions.

Variable	Preselection	
N_ℓ $(\ell = e, \mu)$	$= 3$	
	≥ 1 OSSF lepton pair	
p_{T} (ℓ_1, ℓ_2, ℓ_3)	$> 27, 15, 10$ GeV	
Sum of lepton charges	± 1	
$\min(m_{\ell\ell})$, applies to any OSSF lepton pair	> 20 GeV	
$\lvert m_{\ell\ell} - m_Z \rvert$, applies to at least one OSSF lepton pair	< 10 GeV	
$m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$	> 30 GeV	
	CR($t\bar{t}Z$)	CR(VV)
N_{jets} $(p_{\mathrm{T}} > 25$ GeV)	≥ 6	2-5
$N_{b-\mathrm{jets}}$@85%	≥ 1	0

The fit has three CRs related to non-prompt lepton contribution. These regions are defined from the source and type of non-prompt lepton using `MCTruthClassifier` [129] tool. Most of the selection criteria for the CRs are the same as the SRs except the condition that one of the leptons must not satisfy the tight identification and isolation requirement that is applied to the signal region leptons. This lepton is referred to as the "loose lepton". The heavy flavour non-prompt contribution is isolated by defining $t\bar{t}$-enriched regions that choose events with similar number of b-jets as the SR but outside the Z-mass window. This contribution is divided as per the type of non-prompt lepton: CR(NPL-e-HF) for the electron component and CR(NPL-μ-HF) for the muon component. The CR(NPL-e-Other) component is obtained by selecting events with zero b-jets that fall within the Z-mass window. The selection requirements for non-prompt lepton related control regions is summarised in Table 6.7.

6.3.3 Parameters of interest and nuisance parameters

According to the formulation of PLU, the POIs mainly include the normalisation factors for each of the signal sub-histograms (or folded truth bins), which are then translated into differential cross-sections. In addition, the normalisation factors for important backgrounds that are constrained using control regions, are also considered POIs in this analysis.

In total there are 433 NPs associated with theory and instrumental systematic uncertainties as described in Section 4.6. The NPs are divided in subsets in order to study their grouped impact on the POIs. The groups defined for this analysis are : instrumental, b-tagging, leptons and theory. For each group, the impact is calculated by fitting with the nuisance parameters in the group fixed to their best-fit values. The resulting uncertainty on the parameter of interest is then subtracted in quadrature from the uncertainty obtained from the nominal fit. A list of all the NPs in this analysis are given in Appendix C.

Table 6.7: Definition of the non-prompt lepton control regions.

Variable	Preselection		
N_ℓ $(\ell = e, \mu)$	= 3 (of which = 1 loose non-tight)		
	≥ 1 OSSF lepton pair		
p_{T} (ℓ_1, ℓ_2, ℓ_3)	> 27, 15, 10 GeV		
Sum of lepton charges	±1		
$\min(m_{\ell\ell})$	> 20 GeV		
$m_{\mathrm{T}}(\ell, E_{\mathrm{T}}^{\mathrm{miss}})$	> 30 GeV		
N_{jets} $(p_{\mathrm{T}} > 25$ GeV)	≥ 2 & ≤ 5		
	CR(NPL-e-HF)	CR(NPL-μ-HF)	CR(NPL-e-Other)
Loose lepton flavour	Electron	Muon	Electron
$N_{b-\mathrm{jets}}$@85%	1 or 2		= 0
$\lvert m_{\ell\ell} - m_Z \rvert$	> 10 GeV		< 10 GeV

6.4 Uncertainty estimation

The uncertainty estimation on the parameters is performed by `MINOS` [122] tool which is integrated in `TRExFitter`. It uses the profile likelihood ratio, $\lambda(\mu)$, as defined in [130] to compute intervals representing ± standard deviation on the parameters. With the help of the Wilks theorem [131] it can be shown that $\lambda(\mu)$ follows a χ^2-distribution with,

$$-2\ln(\lambda(\mu)) = \frac{(\mu - \hat{\mu})^2}{\sigma_\mu^2} \tag{6.2}$$

The solution of the equation gives the uncertainty on the parameter μ. This uncertainty is referred to as the total uncertainty which can be decomposed into statistical and systematic components.

Apart from the natural statistical uncertainty due to the size of the data, the statistical uncertainty coming from the limited size of the MC samples is also considered in the fit. For each bin and sample, a NP called a gamma factor is created using a Poisson constraint.

The influence of systematic uncertainties on the fit results can be investigated in two ways: direct influence which modifies the yields and indirect influence they have on the results of the maximum likelihood calculation. If a NP affect the overall number of events uniformly across all bins, it is said to have normalisation component. On the other hand, if a NP changes the shape of a distribution while keeping the total number of yields unchanged, it is said to have a shape component. A Nuisance Parameter (NP) can have both shape and normalisation components. A technique called *pruning* is employed to retain the components that significantly affect the fit while removing those that have minimal impact. This is done to protect the fit from small uncertainties likely originating from statistical fluctuations. Furthermore, for each fit region, a smoothing algorithm is applied to remove any effects caused by statistical fluctuations.

After applying smoothing and pruning, the indirect influence of NPs($\vec{\theta}$) on the POIs($\vec{\mu}$) can be examined by calculating the impact. For each NP four additional fits are performed, two for computing pre-fit impact and two for post-fit impact. For the pre-fit impact, the NP value is chosen as the pre-fit

value ± its uncertainty ($\theta \pm \Delta\theta$). Once the fits are done, the resulting values of μ are compared with the nominal best-fit value of μ. The difference quantifies the impact of that NP on μ. Since the impact is calculated after the standard fit has returned best-fit values, the post-fit impact can be calculated. In this case, the NP is set to $\hat{\theta} \pm \Delta\hat{\theta}$.

The fit also gives additional information about the nuisance parameters in the form of pulls and constraints. A constraint is applied on the NP if its uncertainty is decreased from its original value ($\Delta\theta$) during the fitting process. Furthermore, a pull is imposed if the value of the NP differs from its nominal value. Pulls and constraints suggest that the fit is gaining more information about the NP from the preselected regions. In an Asimov fit, a pull or constraint is undesired because the provided information about the NPs should be enough.

In profile likelihood unfolding, the systematic uncertainties affecting backgrounds are handled in the exact same way as described above. For the systematics affecting signal, an alternative response matrix, representing the systematic variation is employed. The associated alternative response matrix for each NP is used to fold the truth distribution resulting into varied distributions for each of the signal sub-histograms. These are then treated as the up/down variations of the signal sub-histograms.

Eventually, the uncertainties on parameters obtained from the fit are propagated to the differential cross-section yields.

6.5 Validation tests of the unfolding method

Before unfolding real data, a number of statistical tests are performed to validate the implementation of the unfolding method. These tests include closure test, stress test and pull test. Depending on the type of test, the goal is either to verify the uncertainty provided by the fit or to check whether the nominal inputs are biasing the unfolding method.

6.5.1 Closure test

The goal of a closure test is to verify the ability of the unfolding technique to recover the underlying distribution in presence of statistical differences between the sample used to create the response matrix and the sample which is going to be unfolded.

In a closure test, the set of reconstructed and truth distributions of an observable is randomly split into two subsets, called training and testing set. The training set is used to compute the acceptance, efficiency and the migration matrix. Hence, the training set is responsible for the building blocks of the response matrix. On the other hand, the testing subset acts as pseudo-data on which the unfolding is performed. In this case, the backgrounds are also divided into training and testing sets and the testing set is used in the fit. Since this test is aimed towards the behaviour of PLU in case of statistical differences, it is performed without including any systematic uncertainties. Otherwise, the fit setup is same as the one used for Asimov fit.

The migration matrix for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$ computed from the training set and the resulting unfolded distribution of $p_{\mathrm{T}}(t)$ at parton-level, are shown in Fig. 6.5. Good agreement is observed between the unfolded result and the truth distribution. The statistical uncertainties in each bin are expected to be higher than a standard Asimov fit (discussed in Section 6.7) given only half of the available signal statistics enter the pseudo-data.

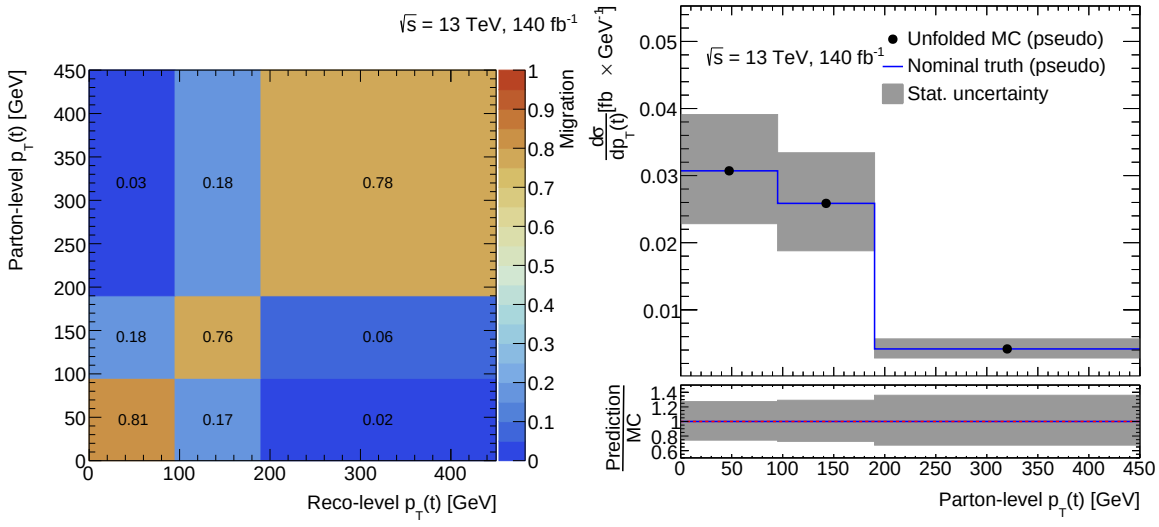


Figure 6.5: Results of the closure test. Migration matrix for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$ (left) and unfolded differential cross-section (right) after parton-level unfolding $p_{\mathrm{T}}(t)$.The uncertainty band includes only statistical uncertainties.

6.5.2 Pull test

The aim of a pull test is to verify the statistical uncertainties on the unfolded results. To perform this test, 10000 pseudo-distributions are created, also called as toy datasets. For each unfolding bin in a toy dataset, the bin content is sampled from a Poisson distribution with a mean equal to the expected bin content given by the standard fit. The production and fitting of the toys are done internally in TRExFitter. Furthermore, all these toys are unfolded using the same response matrix created from the original distributions.

The fit setup for a pull test is intentionally kept minimalist, with no control regions or systematic uncertainties included. This simplified setup is chosen because running a large number of pseudo-experiments with a more complex configuration would be computationally expensive. Despite its simplicity, the minimal setup is sufficient to achieve the objectives of this test. All the nuisance parameters corresponding to systematic uncertainties are fixed, given that the fit only considers statistical uncertainties. The distribution of the resulting 10000 values of the normalisation factors for each bin, is a Gaussian. Now, the standard deviation of the Gaussian is expected to be equal to the uncertainty values given by the fit. The results of toy experiments for parton and particle level $p_{\mathrm{T}}(t)$ are shown in Fig. 6.6 and Fig. 6.7, respectively.

In order to check if the test worked, the standard deviation values are cross-checked with the uncertainty values on the same normalisation factors returned by the fit. These are shown in Fig. 6.8(a) (parton level $p_{\mathrm{T}}(t)$) and Fig. 6.8(b) (particle level $p_{\mathrm{T}}(t)$). It is seen that the uncertainties given by the fit and the standard deviation values are same, concluding that the uncertainties are correctly calculated by the fit. A second interpretation of this test is that if μ follows a Gaussian distribution, then $\frac{\mu-\bar{\mu}}{\sigma}$ must follow a standard normal distribution, which is seen in Fig. 6.9.

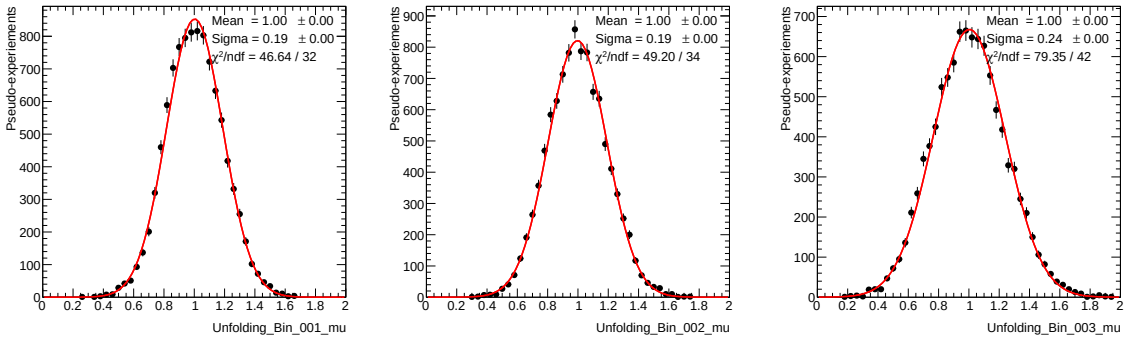


Figure 6.6: Distribution of the normalisation factors for each bin from 10000 toy dataset fits for the observable $p_{\mathrm{T}}(t)$ at parton-level.

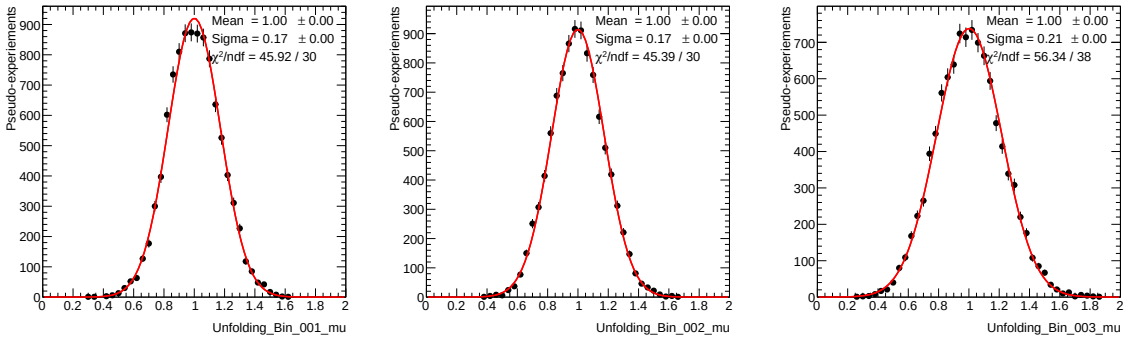
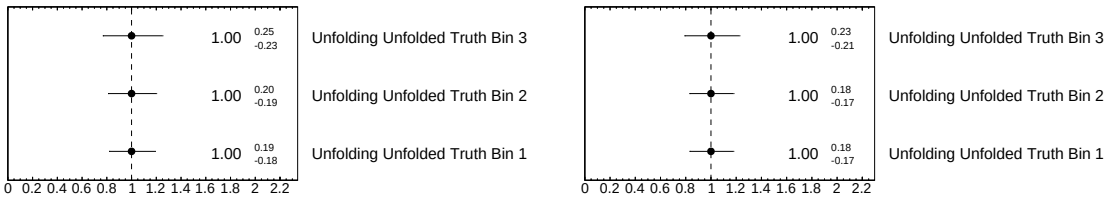


Figure 6.7: Distribution of the normalisation factors for each bin from 10000 toy dataset fits for the observable $p_{\mathrm{T}}(t)$ at particle-level.



(a) NF of $p_{\mathrm{T}}(t)$ at parton level

(b) NF of $p_{\mathrm{T}}(t)$ at particle level

Figure 6.8: Normalisation factors obtained after unfolding $p_{\mathrm{T}}(t)$. Only statistical uncertainties are considered.

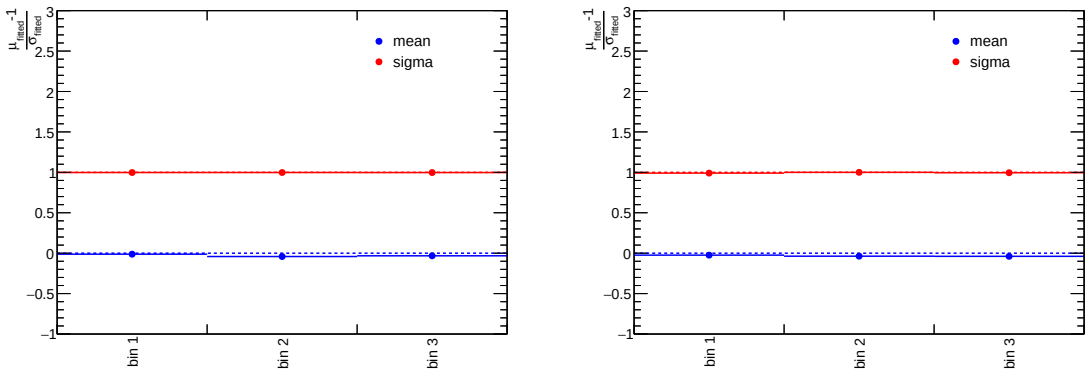


Figure 6.9: Per-bin mean and sigma values of $\frac{\mu-\bar{\mu}}{\sigma}$ from the pull test results in case of $p_{\mathrm{T}}(t)$ unfolding at parton level (left) and particle level (right). As expected, the mean and sigma values imply a standard normal distribution

6.5.3 Stress test

The stress test checks whether the unfolding method is robust against variations in the input distributions. To perform this test, the nominal histograms are reweighted with arbitrary weights to cause shape variations in the inputs. Then the reweighted histograms are unfolded using the nominal response matrix. The desired outcome is that the unfolded reweighted histogram should match the reweighted truth distribution and not the nominal truth distribution. To induce a shape and normalisation variation, the truth-level and reconstructed-level bins were modified by $[+17\%, -12\%, +20\%]$ in case of three-bin histograms and by $[+17\%, -12\%, -14\%, +20\%]$ in case of four-bin histograms. The weights are chosen such that the variations are within the statistical uncertainty of the unfolded yields. The results of this test are in Fig. 6.9. The resulting unfolded distribution matches the reweighted truth and not the nominal truth. Therefore, it can be confidently said that the unfolding method is not biased towards the nominal inputs, and it gives expected results even if the input distribution is fluctuated.

6.6 Cross-check with IBU

The performance of PLU is evaluated against IBU by comparing the statistical uncertainties in the unfolded results produced by both methods. This cross-check is done for $p_{\mathrm{T}}(t)$ and $p_{\mathrm{T}}(Z)$ variables using the Asimov dataset. These two variables are selected because, for $p_{\mathrm{T}}(t)$, the migration matrix is less diagonal, whereas for $p_{\mathrm{T}}(Z)$, it is nearly diagonal. Together, these variables represent two extremes. Unfolding through iterative Bayesian method is performed using three iterations. In this thesis, IBU is used as a method for cross-checking PLU results. Consequently, the development of IBU in terms of optimising number of iterations and including systematic uncertainties falls outside the scope of this work. The differential cross-sections along with the statistical uncertainties given by both methods, are summarised in Table 6.8.

The statistical uncertainties given by IBU and PLU are found to be similar. In some cases, error given by PLU are slightly smaller. The slight difference could be because IBU is not optimised to account for overflow and underflow bins, while PLU is. A more robust comparison can be performed

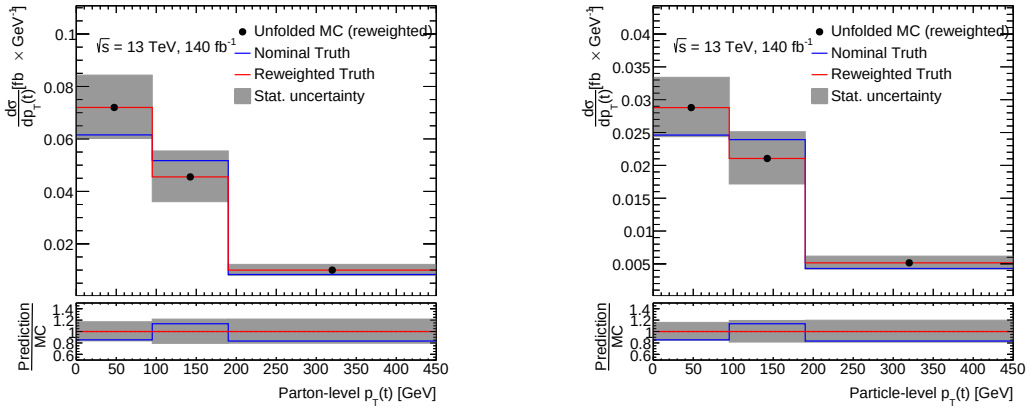


Figure 6.9: Summary of the stress test for $p_{\mathrm{T}}(t)$ unfolded at parton-level (left) and particle-level (right). Here reweighted distributions are unfolded using nominal response matrix.The uncertainty band includes only statistical uncertainties.

after optimising IBU iterations and also systematic uncertainties inclusion. However, based on statistical uncertainty computation, the performance of both these methods is the same.

Table 6.8: Differential cross-section values with statistical uncertainties given by IBU and PLU. The top two tables represent parton and particle level $p_{\mathrm{T}}(t)$ whereas the bottom two tables represent parton and particle level $p_{\mathrm{T}}(Z)$.

Parton-level $\frac{d\sigma}{dp_{\mathrm{T}}(t)}$ (fb GeV^{-1})	IBU error (±)	PLU error (±)
0.0615	0.0125	0.0114
0.0517	0.0100	0.0100
0.0079	0.0023	0.0020

Particle-level $\frac{d\sigma}{dp_{\mathrm{T}}(t)}$ (fb GeV^{-1})	IBU error (±)	PLU error (±)
0.0246	0.0052	0.0043
0.0239	0.0047	0.0042
0.0041	0.0019	0.0009

Parton-level $\frac{d\sigma}{dp_{\mathrm{T}}(Z)}$ (fb GeV^{-1})	IBU error (±)	PLU error (±)
0.0736	0.0170	0.0130
0.1074	0.0235	0.0180
0.0620	0.0137	0.0110
0.0076	0.0022	0.0016

Particle-level $\frac{d\sigma}{dp_{\mathrm{T}}(Z)}$ (fb GeV^{-1})	IBU error (±)	PLU error (±)
0.0304	0.0071	0.0055
0.0459	0.0101	0.0078
0.0282	0.0063	0.0049
0.0037	0.0012	0.0008

6.7 Differential cross-sections for the Asimov dataset

The differential cross-section distributions are obtained by profile likelihood unfolding of the Asimov dataset. In unfolding terminology, it means unfolding the reconstructed-level distributions. The migration matrix, acceptance and efficiencies are computed using the truth-level distribution and the same reconstructed-level distribution. Here, one would expect the unfolded distribution to exactly match the truth-level distribution, given everything is set correctly. This is intuitive because the

reconstruction-level sample is created by smearing the same truth sample. Essentially, this is a way to validate the unfolding method. Hence, unfolding of the Asimov dataset is a consistency test for the unfolding method.

The differential cross-sections are calculated with respect to all the unfolding variables mentioned in Section 6.1 at parton and particle level. The results for parton-level $p_{\mathrm{T}}(t)$ and particle-level $p_{\mathrm{T}}(t)$ are discussed in detail in Section 6.7 and Section 6.7, respectively. The shape of $p_{\mathrm{T}}(t)$ at reconstructed-level and truth-level is overlaid in Fig. 6.10 in order to get a visual idea of the unfolding problem at hand. Compared to parton-level, the reconstructed distribution is closer to the particle-level truth distribution.

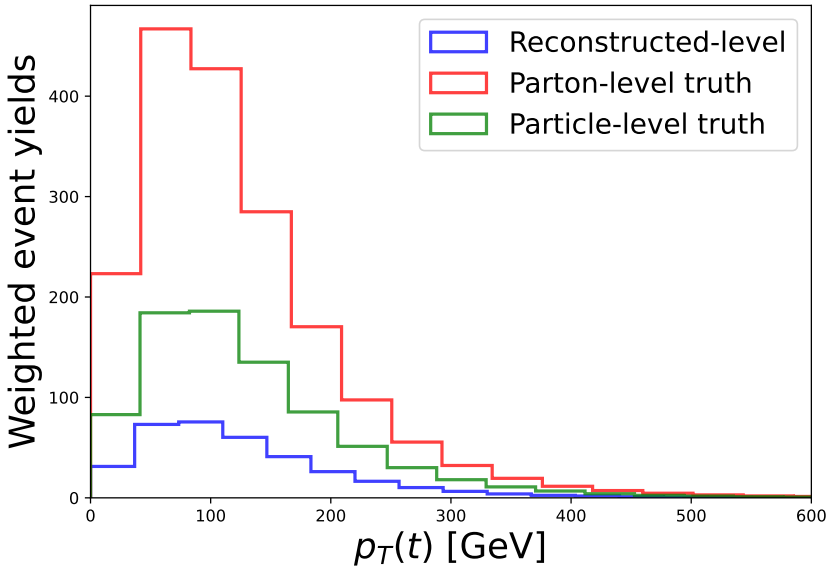


Figure 6.10: Overlay distributions comparing reconstructed-level $p_{\mathrm{T}}(t)$ distribution with parton truth-level (left) and particle-level truth (right).

Parton-level $p_{\mathrm{T}}(t)$

For parton-level $p_{\mathrm{T}}(t)$ unfolding, the unfolding ingredients, which include migration matrix, efficiency and acceptances, are shown in Fig. 6.11. These quantities are calculated for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$ and $O_{\mathrm{NN}}(\mathrm{SR}) < 0.7$ using the same truth-level sample for both. The diagonal of the migration matrices is higher than 65 %, implying that most of the events are reconstructed in the same bin in which they were produced. It also validates that the MC generator correctly performed the reconstruction. It is also crucial to note that, certain particles have inherently better signatures in the detector compared to others. For instance, the Z boson reconstructed from two leptons. For variables related to Z, one can obtain an almost perfectly diagonal migration matrix (as given in Appendix A).

The per-bin acceptance shows that around 95 % of the reconstructed-level events are true events. The remaining may be outside the defined fiducial phase space or reconstructed in another bin which will be accounted for in the migration matrix. The per-bin efficiency is around 14 % for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$

and around 7 % for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$. For each bin, it shows the fraction of events from the truth fiducial volume that are reconstructed in that bin. Remaining events may have potentially migrated to other bins or failed reconstruction. The error bars on acceptance and efficiency plots show 68 % confidence intervals.

The best-fit normalisation factors for each of the folded truth-bin for $p_{\mathrm{T}}(t)$ in addition to the best-fit background normalisation factors obtained from the Asimov fit are shown in Fig. 6.12. Since this is an Asimov fit, the values are expected to be 1, given the unfolding setup is correct. The error on the values includes statistical and systematic uncertainty. The per-bin values of the folded truth bins are used to scale the truth distribution in order to obtain unfolded yields. The main result which is the differential cross-section distribution is obtained by dividing the unfolded yields by bin width of the observable and luminosity equal to 140.1 fb^{-1}. Figure 6.14(a) shows the differential cross-section as a function of $p_{\mathrm{T}}(t)$. The errors (shown with grey bands) include statistical and systematic uncertainties. The average relative uncertainty on the differential cross-section for $p_{\mathrm{T}}(t)$ is 23 %. The breakdown in terms of statistical and groups of systematic uncertainties on the results is shown in Fig. 6.13. The fractional uncertainty is calculated by the ratio of the grouped uncertainty value (obtained from the grouped impact calculation) to the total uncertainty.

In the normalised fit, a new free parameter is introduced μ_{tot}, which represents the total cross-section. The existing normalisation factors for the folded truth bins are redefined in terms of μ_{tot}. The factor corresponding to the last bin is obtained from the difference of total normalisation and the sum of normalisations of the other bins. By doing a normalised fit, the effect of some systematic uncertainties can be reduced, ultimately decreasing the total uncertainty. The result of a normalised fit can be interpreted as the relative cross-sections in each bin of the distribution as shown in Fig. 6.14(b). Instead of the last bin NF, the new free parameter representing total cross-section is shown in the plot. The uncertainty on this parameter is nothing but the uncertainty on the inclusive cross-section.

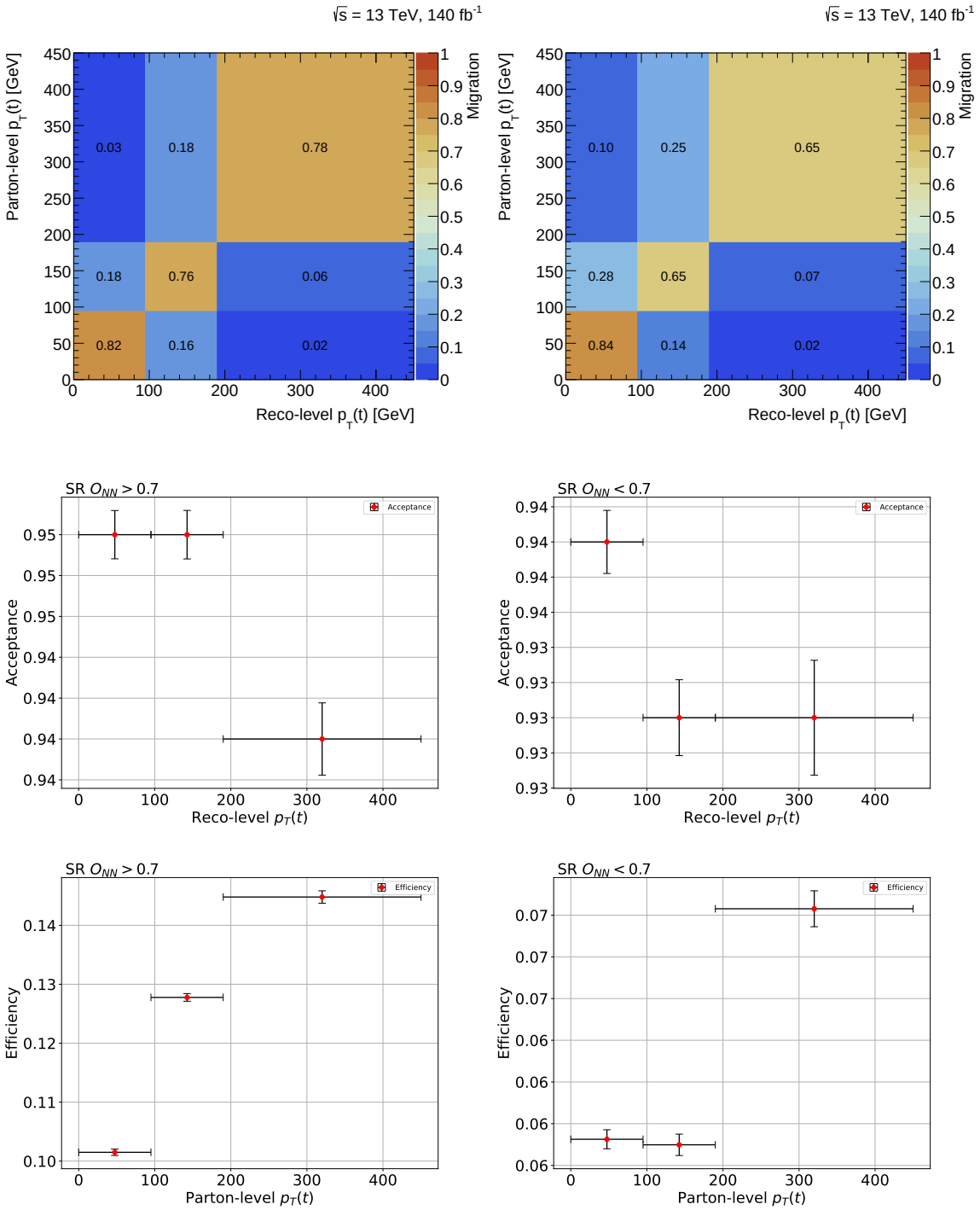


Figure 6.11: Parton-level unfolding ingredients for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$ (left) and $O_{\mathrm{NN}}(\mathrm{SR}) < 0.7$ (right). The migration matrices (top), acceptance plots (middle) and efficiency plots (bottom) that build the response matrices are shown for the observable $p_{\mathrm{T}}(t)$. The last bin includes the overflow events.

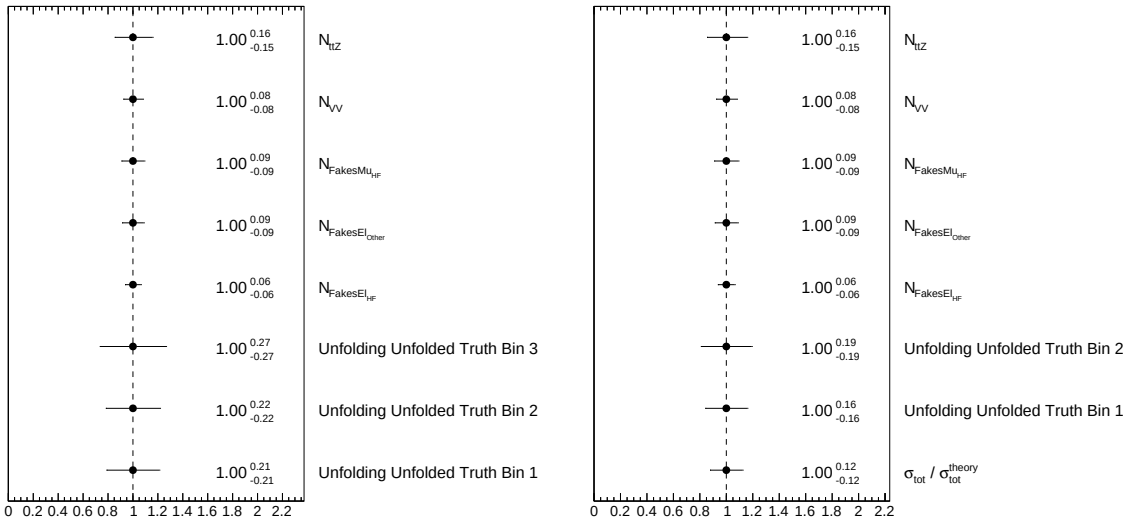


Figure 6.12: Parton-level normalisation factors for the observable $p_{\mathrm{T}}(t)$ from an Asimov fit (left) and a normalised Asimov fit (right). The exact bin edges can be found in Table 6.3. The uncertainties include both statistical and systematic uncertainties.

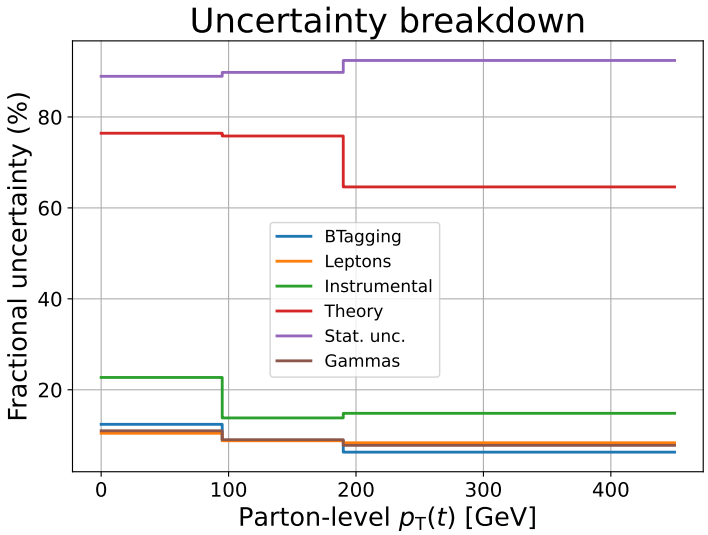
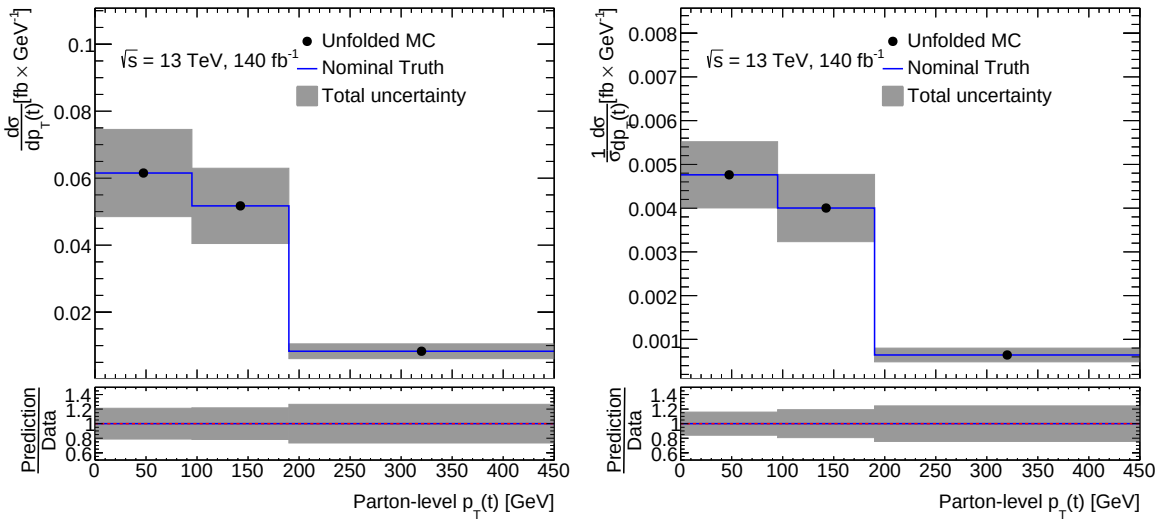


Figure 6.13: Fractional uncertainties on the differential cross-sections for parton-level $p_{\mathrm{T}}(t)$. The statistical uncertainty (shown in purple) is the major contributor in all the bins.



(a) Differential cross-section

(b) Normalised differential cross-section

Figure 6.14: Parton-level unfolded results for the observable $p_{\mathrm{T}}(t)$ from an Asimov fit.The uncertainty band includes both statistical and systematic uncertainties.

Particle-level $p_{\mathrm{T}}(t)$

For particle-level $p_{\mathrm{T}}(t)$ unfolding, the unfolding ingredients, which include migration matrix, efficiency and acceptances, are shown in Fig. 6.15. These quantities are calculated for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$ and $O_{\mathrm{NN}}(\mathrm{SR}) < 0.7$ using the same truth-level sample for both. The diagonal of the migration matrices is higher than 74 %, implying that most of the events are reconstructed in the same bin in which they were produced.

The per-bin acceptance shows that around 93 % of the reconstructed-level events in $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$ are true events. The remaining may be outside the defined fiducial phase space or reconstructed in another bin. The per-bin efficiency is around 28 % for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$ and around 14 % for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$. The efficiency is higher compared to parton-level, this is a consequence of the particle-level fiducial volume being close to the reconstructed-level phase space.

The best-fit normalisation factors for each of the folded truth-bin for $p_{\mathrm{T}}(t)$ in addition to the best-fit background normalisation factors obtained from the Asimov fit are shown in Fig. 6.16. The error on the values includes statistical and systematic uncertainty. The main result which is the differential cross-section distribution is obtained by dividing the unfolded yields by bin width of the observable and luminosity equal to 140.1 fb^{-1}. Figure 6.18(a) shows the differential cross-section as a function of $p_{\mathrm{T}}(t)$. The errors (shown with grey bands) include statistical and systematic uncertainties. The average relative uncertainty on the differential cross-section for $p_{\mathrm{T}}(t)$ is 21 %. The results of a normalised fit are given Fig. 6.18(b). The breakdown in terms of statistical and groups of systematic uncertainties on the results is shown in Fig. 6.17.

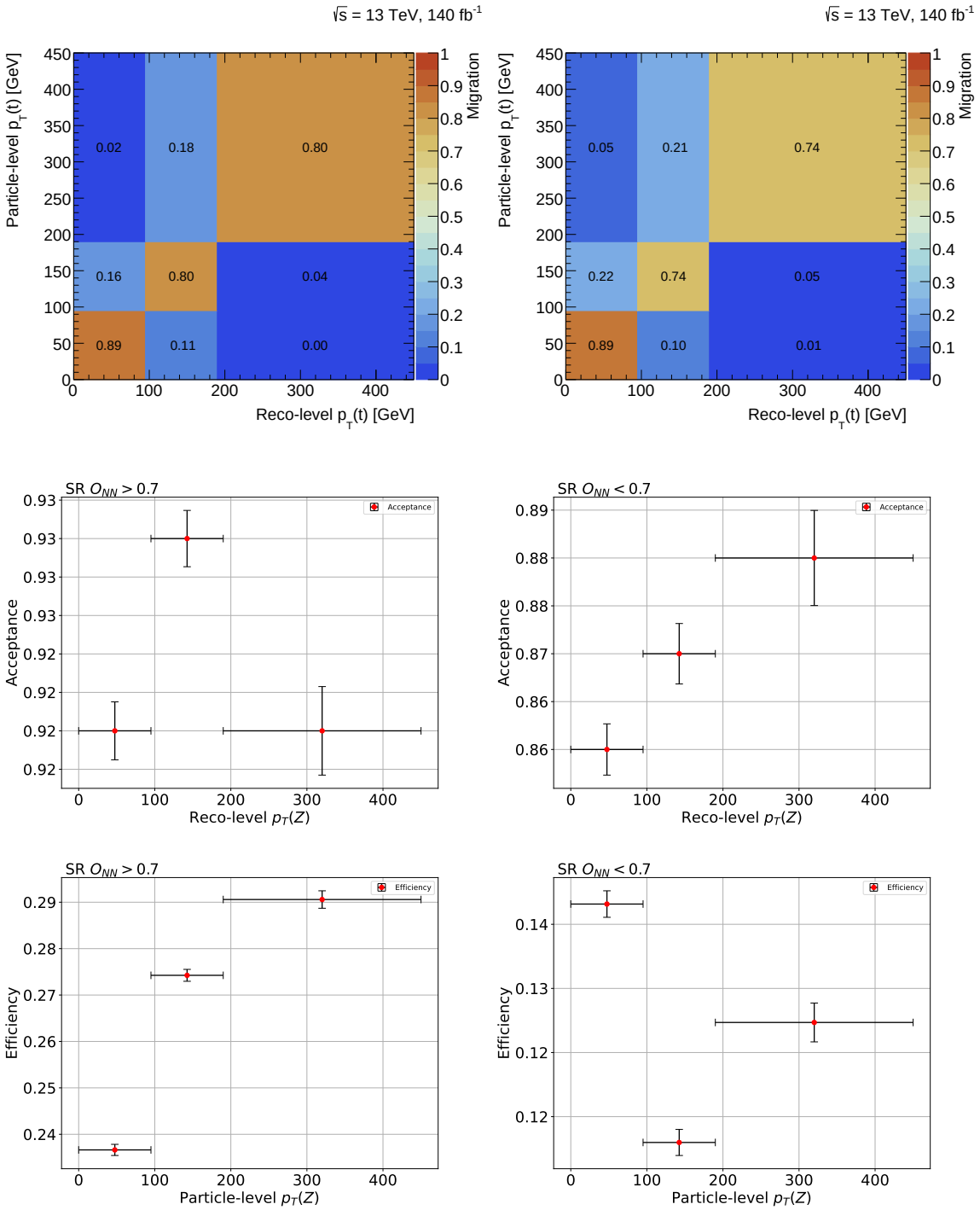


Figure 6.15: Particle-level unfolding ingredients for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$ (left) and $O_{\mathrm{NN}}(\mathrm{SR}) < 0.7$ (right). The migration matrices (top), acceptance plots (middle) and efficiency plots (bottom) that build the response matrices are shown for the observable $p_{\mathrm{T}}(t)$. The last bin includes the overflow events.

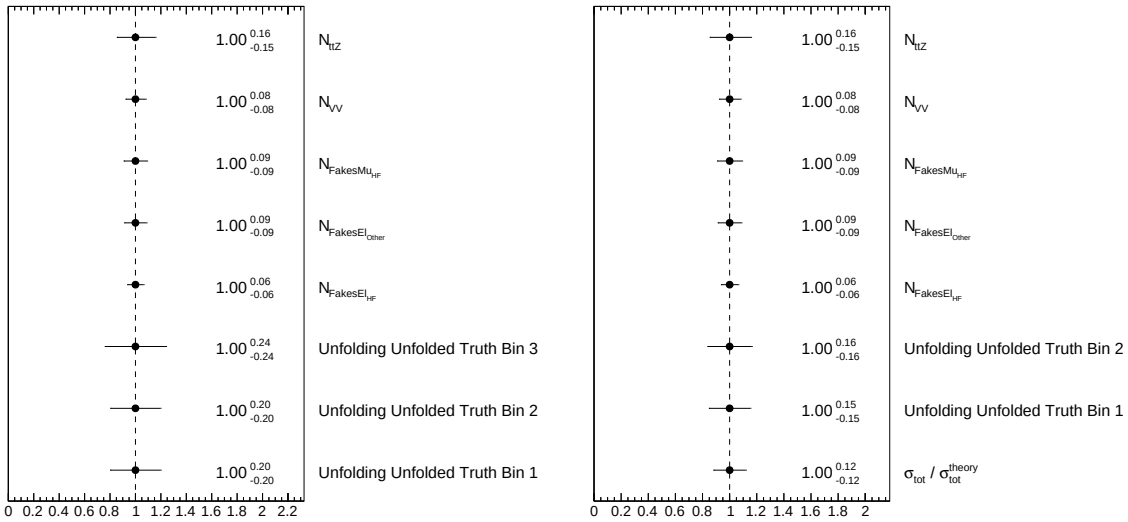


Figure 6.16: Particle-level normalisation factors for the observable $p_{\mathrm{T}}(t)$ from an Asimov fit (left) and a normalised Asimov fit (right). The exact bin edges can be found in Table 6.3. The uncertainties include both statistical and systematic uncertainties.

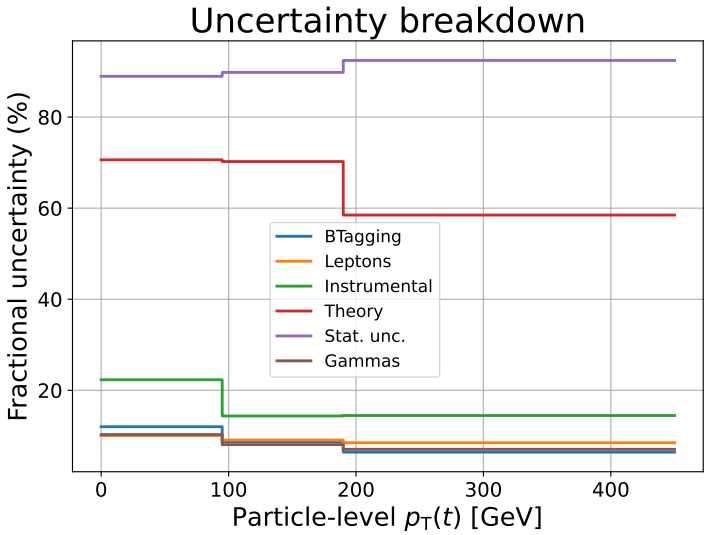
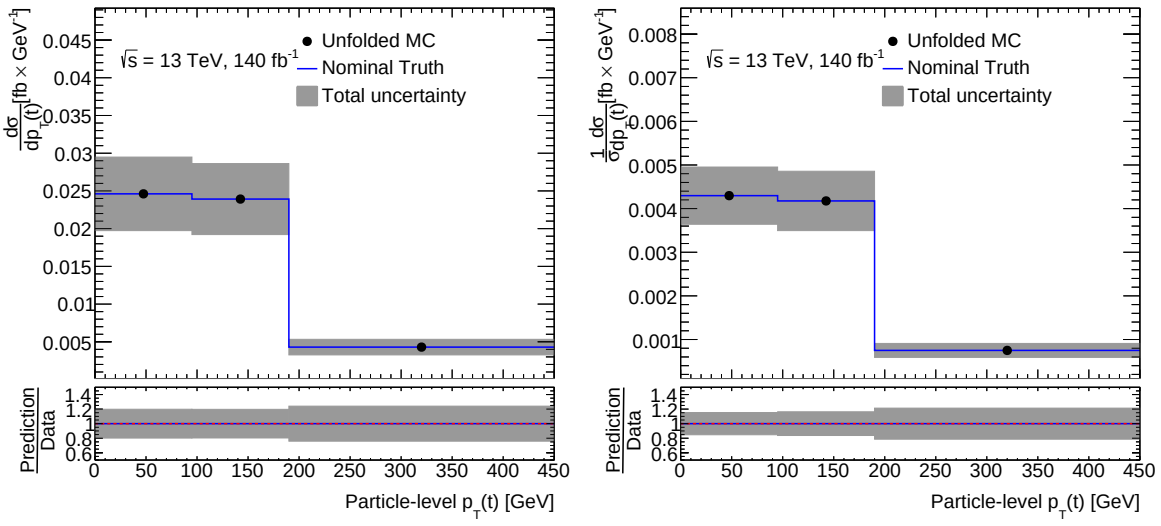


Figure 6.17: Fractional uncertainties on the differential cross-sections for particle-level $p_{\mathrm{T}}(t)$. The statistical uncertainty (shown in purple) is the major contributor in all the bins.



(a) Differential cross-section

(b) Normalised differential cross-section

Figure 6.18: Particle-level unfolded results for the observable $p_{\mathrm{T}}(t)$ from an Asimov fit. The uncertainty band includes both statistical and systematic uncertainties.

Effect of systematic uncertainties

All the nuisance parameters computed from the Asimov fit, except a few, are centered at 0 with a standard deviation of 1. The theory systematics category mainly includes signal and background modelling uncertainties, for instance, tZq showering. The instrumental category includes systematics connected to detector systematics. The remaining two groups cover the b-tagging and leptons related systematics.

In the Asimov fit, no pulls have been observed for any observable. However, notable constraints are observed for $\mu_R + \mu_F(VV + l)$ and in some $t\bar{t}$ modelling uncertainties as shown in Fig. 6.19. The fit might be gaining more information due to the additional CRs associated with non-prompt lepton contributions and diboson. The correlations between the free and nuisance parameters is shown in Fig. 6.20. High correlations are observed between heavy flavour (HF) non-prompt normalisation factors and $t\bar{t}$ modelling. This is expected because the contributions in the associated control regions (CRs) are predominantly from $t\bar{t}$ events with an additional non-prompt lepton. As a result, these factors are similarly sensitive to variations in $t\bar{t}$ modelling uncertainties, leading to the observed correlations.

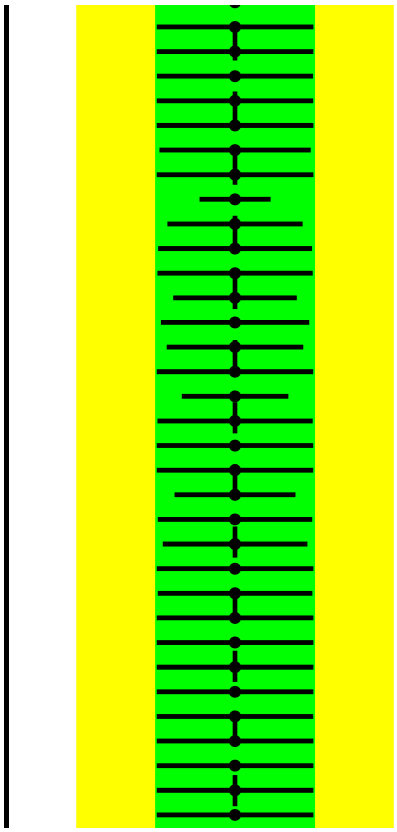


Figure 6.19: A subset of the theory related nuisance parameters. The renormalisation and factorisation scales of diboson light flavour are seen to be non-negligibly constrained.

The ranking plots for each unfolded bin, as shown in Fig. 6.21, list the top 20 nuisance parameters that have the largest impact on the parameters of interest. The pre-fit impact and post-fit impact are shown with empty and filled rectangles around the best-fit values, respectively. The ranking plots also

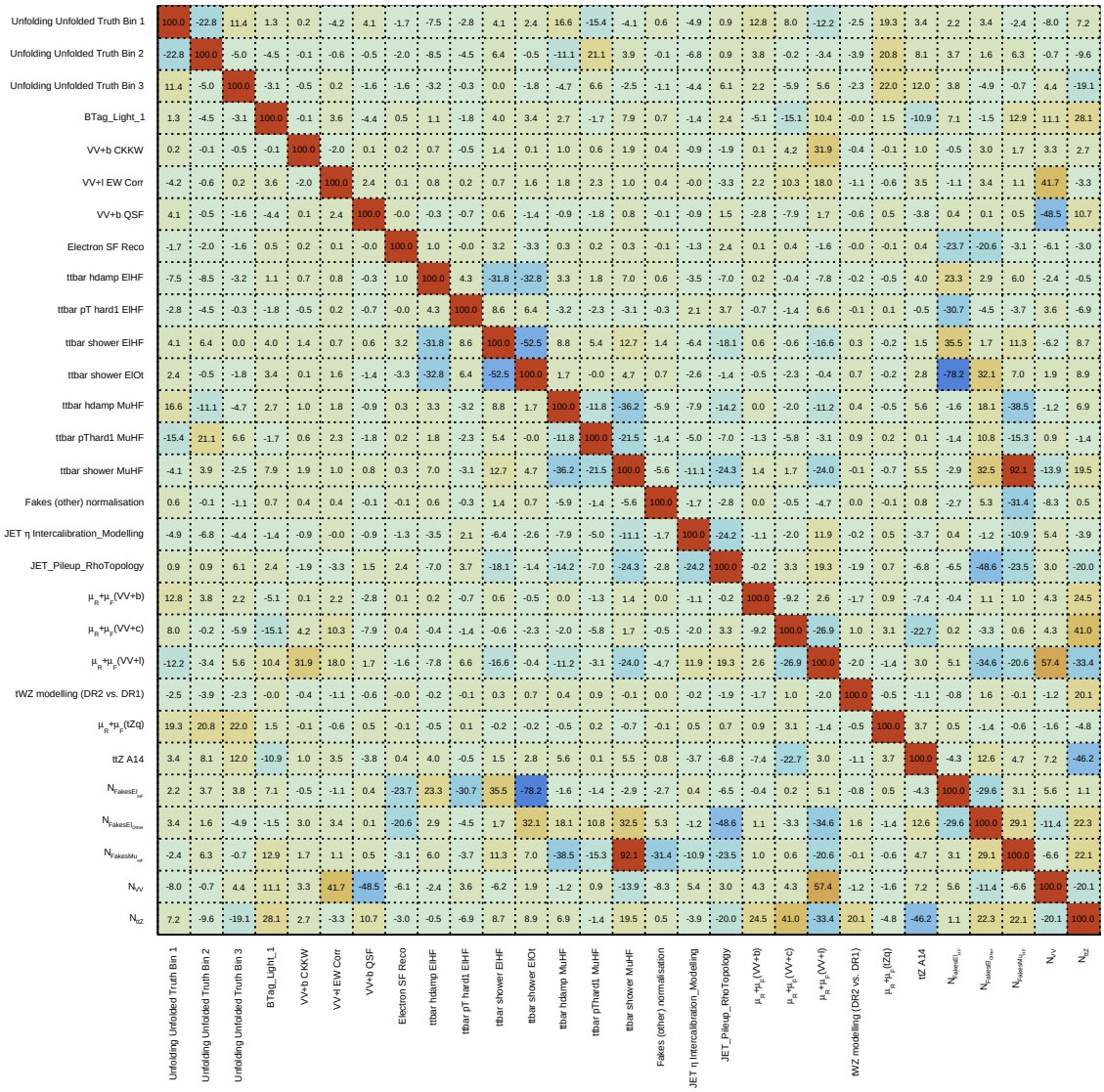


Figure 6.20: Correlation matrix of the systematics for the observable $p_{\mathrm{T}}(t)$ unfolded at parton-level.

show the pulls and constraints on the NPs. The NPs associated with modelling systematics in addition to factorisation and renormalisation scale uncertainties, are ranked high.

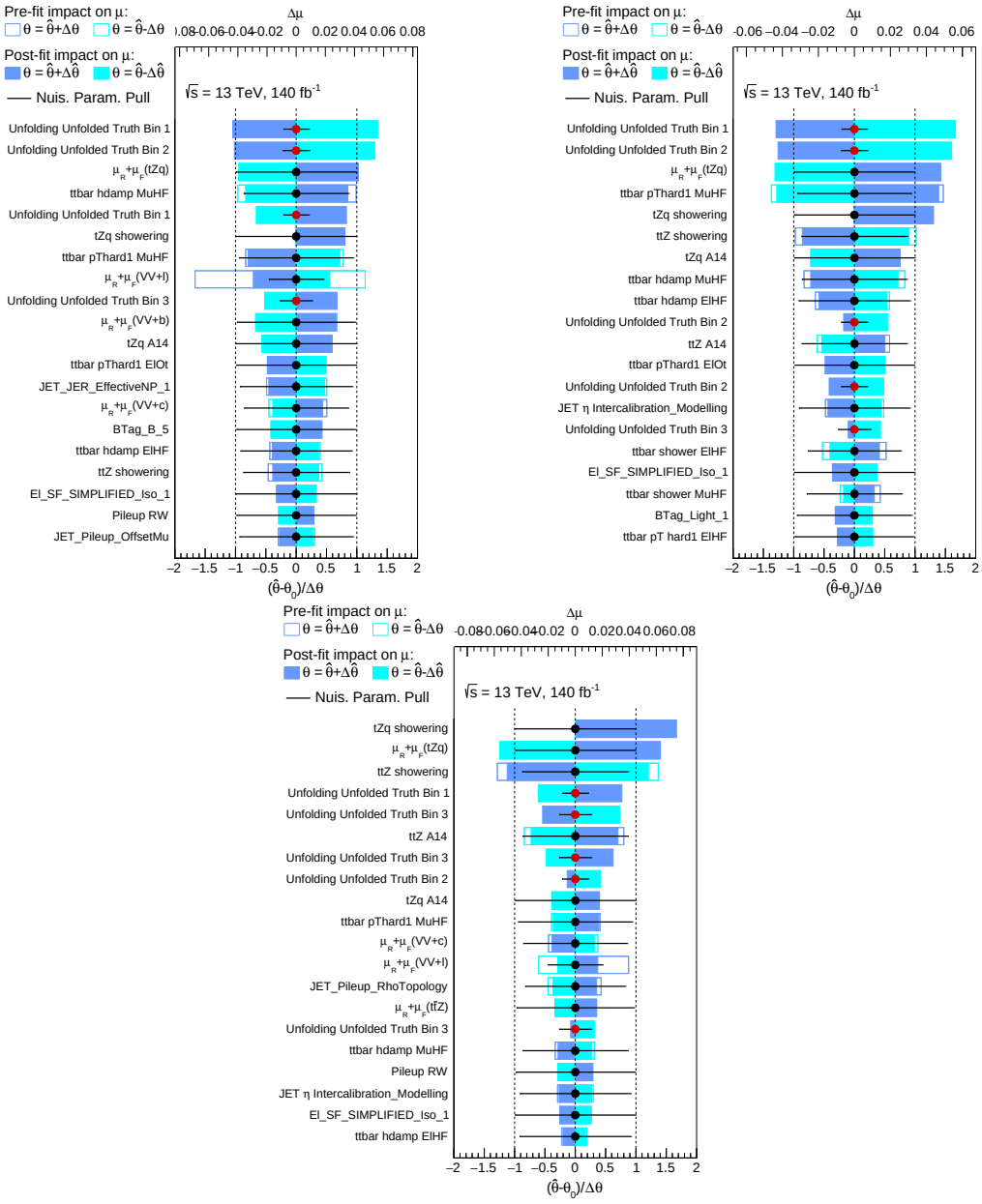


Figure 6.21: Ranking plots for each unfolded bin showing the top 20 systematics having the largest impact arranged for the observable $p_{\mathrm{T}}(t)$ unfolded at parton-level. The red dots correspond to other POIs in the fit.

6.8 Differential cross-sections for the observed data

In this section, the differential cross-sections obtained from PLU implemented on the observed Run-2 dataset, are presented. The normalisation factors calculated in PLU are summarised in Table 6.9 and Table 6.11 for variables unfolded at parton level and particle level, respectively. The values range between 0.78 to 1.22 with a total uncertainty of approximately 20 %, demonstrating that they are close to 1 and indicating a good agreement with the predictions. The resulting differential cross-section yields are presented in Table 6.10 and Table 6.12 for variables unfolded at parton level and particle level, respectively. It can be seen that the observed differential cross-sections are consistent with the theoretical predicted cross-sections.

Table 6.9: Normalisation factors along with total uncertainty calculated by the fit using observed data for all variables unfolded at parton level. The respective binning of variables is also shown.

$p_{\mathrm{T}}(t)$ GeV	Norm factors Value ± Total unc.
[0,95]	1.14 ± 0.22
[95,190]	0.88 ± 0.21
[190,450]	1.11 ± 0.27

$p_{\mathrm{T}}(Z)$ GeV	Norm factors Value ± Total unc.
[0,45]	1.03 ± 0.20
[45,80]	1.05 ± 0.19
[80,135]	0.78 ± 0.19
[135,450]	1.27 ± 0.25

$\lvert\eta(t)\rvert$	Norm factors Value ± Total unc.
[0.00,0.70]	1.22 ± 0.28
[0.70,1.70]	0.57 ± 0.26
[1.70,4.00]	1.42 ± 0.29

$\Delta R(t,Z)$	Norm factors Value ± Total unc.
[0.00,2.60]	1.00 ± 0.23
[2.60,3.50]	0.89 ± 0.22
[3.50,7.00]	1.25 ± 0.25

$\lvert\eta(Z)\rvert$	Norm factors Value ± Total unc.
[0.00,0.70]	1.01 ± 0.20
[0.70,1.40]	1.16 ± 0.20
[1.40,2.10]	0.87 ± 0.19
[2.10,4.00]	1.06 ± 0.22

$\lvert\Delta\phi(\ell_1^Z,\ell_2^Z)\rvert$	Norm factors Value ± Total unc.
[0.00,1.30]	1.13 ± 0.24
[1.30,2.00]	0.80 ± 0.18
[2.00,2.70]	0.78 ± 0.17
[2.70,4.00]	0.73 ± 0.19

Table 6.10: Differential cross-section values calculated for all variables unfolded at parton level. The bin edges and predicted cross-section in each bin is also shown.

$p_{\mathrm{T}}(t)$ GeV	Diff. cross section Obs. ± Total unc.	Pred.
[0,95]	0.070 ± 0.014	0.062
[95,190]	0.046 ± 0.011	0.052
[190,450]	0.009 ± 0.002	0.008

$\Delta R(t,Z)$	Diff. cross section Obs. ± Total unc.	Pred.
[0.00,2.60]	1.544 ± 0.348	1.544
[2.60,3.50]	4.919 ± 1.218	5.513
[3.50,7.00]	1.408 ± 0.283	1.127

$\lvert\eta(Z)\rvert$	Diff. cross section Obs. ± Total unc.	Pred.
[0.00,0.70]	3.866 ± 0.775	3.845
[0.70,1.40]	4.810 ± 0.828	4.158
[1.40,2.10]	3.519 ± 0.755	4.035
[2.10,4.00]	2.513 ± 0.511	2.366

$\lvert\eta(t)\rvert$ GeV	Diff. cross section Obs. ± Total unc.	Pred.
[0,0.70]	5.349 ± 1.234	4.375
[0.70,1.70]	2.644 ± 1.218	4.603
[1.70,4.00]	3.251 ± 0.659	2.285

$p_{\mathrm{T}}(Z)$ GeV	Diff. cross section Obs. ± Total unc.	Pred.
[0,45]	0.076 ± 0.015	0.074
[45,80]	0.113 ± 0.021	0.107
[80,135]	0.048 ± 0.012	0.062
[135,450]	0.010 ± 0.002	0.008

$\lvert\Delta\phi(\ell_1^Z,\ell_2^Z)\rvert$	Diff. cross section Obs. ± Total unc.	Pred.
[0.00,1.30]	2.865 ± 0.595	2.527
[1.30,2.00]	3.639 ± 0.814	4.561
[2.00,2.70]	4.323 ± 0.939	5.514
[2.70,4.00]	1.442 ± 0.384	1.988

Table 6.11: Normalisation factors along with total uncertainty calculated by the fit using observed data for all variables unfolded at particle level. The respective binning of variables is also shown.

$p_{\mathrm{T}}(t)$ GeV	Norm factors Value ± Total unc.
[0,95]	1.13 ± 0.21
[95,190]	0.88 ± 0.19
[190,450]	1.11 ± 0.25

$p_{\mathrm{T}}(Z)$ GeV	Norm factors Value ± Total unc.
[0,45]	1.04 ± 0.20
[45,80]	1.05 ± 0.19
[80,135]	0.77 ± 0.18
[135,450]	1.27 ± 0.24

p_{T}(f-jet) GeV	Norm factors Value ± Total unc.
[0,65]	1.18 ± 0.24
[65,100]	0.82 ± 0.25
[100,170]	0.87 ± 0.21
[170,400]	1.11 ± 0.26

$\Delta R(t,$ f-jet)	Norm factors Value ± Total unc.
[0,3.40]	0.93 ± 0.27
[3.40,4.20]	0.79 ± 0.23
[4.20,5.20]	1.22 ± 0.25
[5.20,7]	1.21 ± 0.30

$\Delta R(t,Z)$	Norm factors Value ± Total unc.
[0.00,2.60]	1.00 ± 0.20
[2.60,3.50]	0.92 ± 0.18
[3.50,7.00]	1.23 ± 0.23

$\lvert\eta(Z)\rvert$	Norm factors Value ± Total unc.
[0.00,0.70]	1.01 ± 0.20
[0.70,1.40]	1.15 ± 0.19
[1.40,2.10]	0.87 ± 0.18
[2.10,4.00]	1.05 ± 0.21

$\lvert\eta$(f-jet)$\rvert$	Norm factors Value ± Total unc.
[0,2]	1.07 ± 0.25
[2,3]	1.00 ± 0.17
[3,3.5]	0.93 ± 0.23
[3.5,4.5]	1.03 ± 0.24

$\lvert\eta(t)\rvert$	Norm factors Value ± Total unc.
[0.00,0.70]	1.09 ± 0.20
[0.70,1.70]	0.75 ± 0.19
[1.70,4.00]	1.40 ± 0.27

$\lvert\Delta\phi(\ell_1^Z,\ell_2^Z)\rvert$	Norm factors Value ± Total unc.
[0.00,1.30]	1.10 ± 0.22
[1.30,2.00]	0.92 ± 0.18
[2.00,2.70]	1.00 ± 0.17
[2.70,4.00]	1.07 ± 0.23

Table 6.12: Differential cross-section values calculated for all variables unfolded at parton level. The bin edges and predicted cross-section in each bin is also shown.

$p_{\mathrm{T}}(t)$ GeV	Diff. cross section Obs. ± Total unc.		Pred.
[0,95]	0.028 ±	0.005	0.025
[95,190]	0.021 ±	0.005	0.024
[190,450]	0.005 ±	0.001	0.004

$\lvert\eta(t)\rvert$ GeV	Diff. cross section Obs. ± Total unc.		Pred.
[0,0.70]	2.998 ±	0.542	2.748
[0.70,1.70]	1.765 ±	0.452	2.340
[1.70,4.00]	0.892 ±	0.173	0.637

$\Delta R(t, Z)$	Diff. cross section Obs. ± Total unc.		Pred.
[0.00,2.60]	0.792 ±	0.156	0.796
[2.60,3.50]	2.473 ±	0.491	2.674
[3.50,7.00]	0.439 ±	0.081	0.358

$p_{\mathrm{T}}(Z)$ GeV	Diff. cross section Obs. ± Total unc.		Pred.
[0,45]	0.032 ±	0.006	0.030
[45,80]	0.048 ±	0.009	0.046
[80,135]	0.022 ±	0.005	0.028
[135,450]	0.005 ±	0.001	0.004

$\lvert\eta(Z)\rvert$	Diff. cross section Obs. ± Total unc.		Pred.
[0.00,0.70]	2.254 ±	0.444	2.233
[0.70,1.40]	2.693 ±	0.454	2.336
[1.40,2.10]	1.750 ±	0.370	2.003
[2.10,4.00]	0.621 ±	0.124	0.594

p_{T}(f-jet) GeV	Diff. cross section Obs. ± Total unc.		Pred.
[0,65]	0.038 ±	0.008	0.032
[65,100]	0.032 ±	0.010	0.039
[100,170]	0.017 ±	0.004	0.020
[170,400]	0.004 ±	0.001	0.004

$\lvert\eta$(f-jet)$\rvert$	Diff. cross section Obs. ± Total unc.		Pred.
[0,2]	1.142 ±	0.266	1.066
[2,3]	1.885 ±	0.319	1.889
[3,3.5]	1.556 ±	0.378	1.671
[3.5,4.5]	0.897 ±	0.212	0.871

$\Delta R(t$, f-jet)	Diff. cross section Obs. ± Total unc.		Pred.
[0,3.40]	0.607 ±	0.173	0.653
[3.40,4.20]	1.476 ±	0.425	1.874
[4.20,5.20]	1.553 ±	0.313	1.271
[5.20,7]	0.493 ±	0.123	0.409

$\lvert\Delta\phi(\ell_1^Z, \ell_2^Z)\rvert$	Diff. cross section Obs. ± Total unc.		Pred.
[0,1.30]	1.286 ±	0.263	1.171
[1.30,2.00]	1.908 ±	0.369	2.082
[2.00,2.70]	2.362 ±	0.410	2.370
[2.70,4.00]	0.896 ±	0.188	0.835

The differential cross-section distributions obtained after unfolding the observed dataset are shown in Fig. 6.22 and Fig. 6.23. The predicted cross-sections in each bin are denoted by blue coloured lines and observed (calculated by PLU) cross-sections are denoted by black dots. The grey blocks represent total uncertainty on the calculated values. Moreover, these plots also represent the ratio of predicted cross-sections to observed cross-section in each bin, which is found to be close to 1. For all the unfolded variables, the observed cross-sections are consistent with the nominal truth i.e. the predicted cross-sections.

The ranking plots showing 20 highly impacting systematic uncertainties are given in Fig. 6.24. Unlike a standard profile likelihood fit, an unfolding fit builds ranking plot for each folded truth bin. In the ranking computation for a specific POI representing a particular bin, other POIs, along with the NPs, are also taken into account. The pulls and constraints on the nuisance parameters are shown in Fig. B.5. All the NPs are found to lie within their standard deviations, with some exhibiting minor pulls and constraints. However, no significant pulls or constraints are observed, thereby validating the auxiliary measurements.

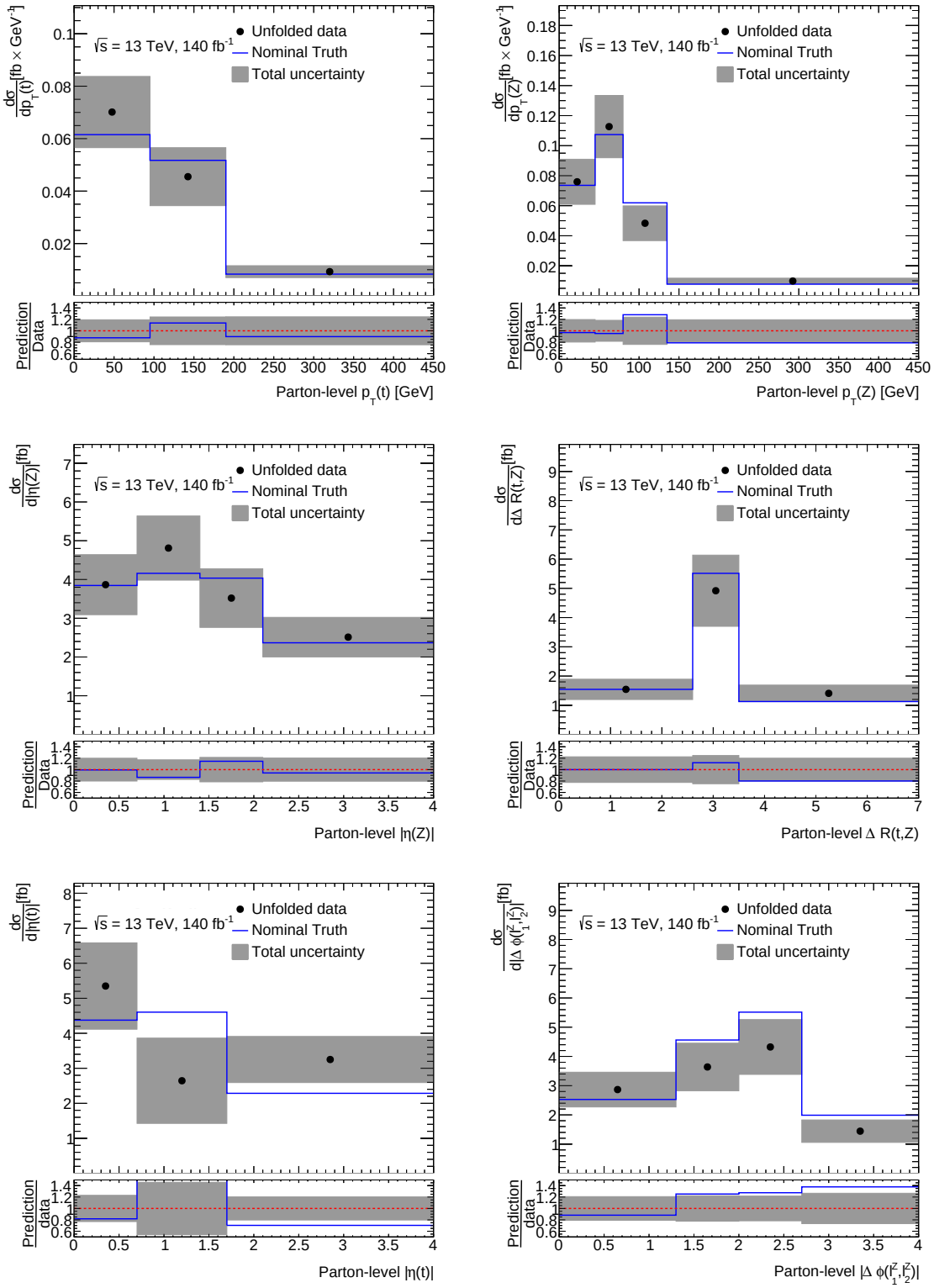


Figure 6.22: Observed differential cross-sections for variables unfolded at parton level. The uncertainty band includes both statistical and systematic uncertainties.

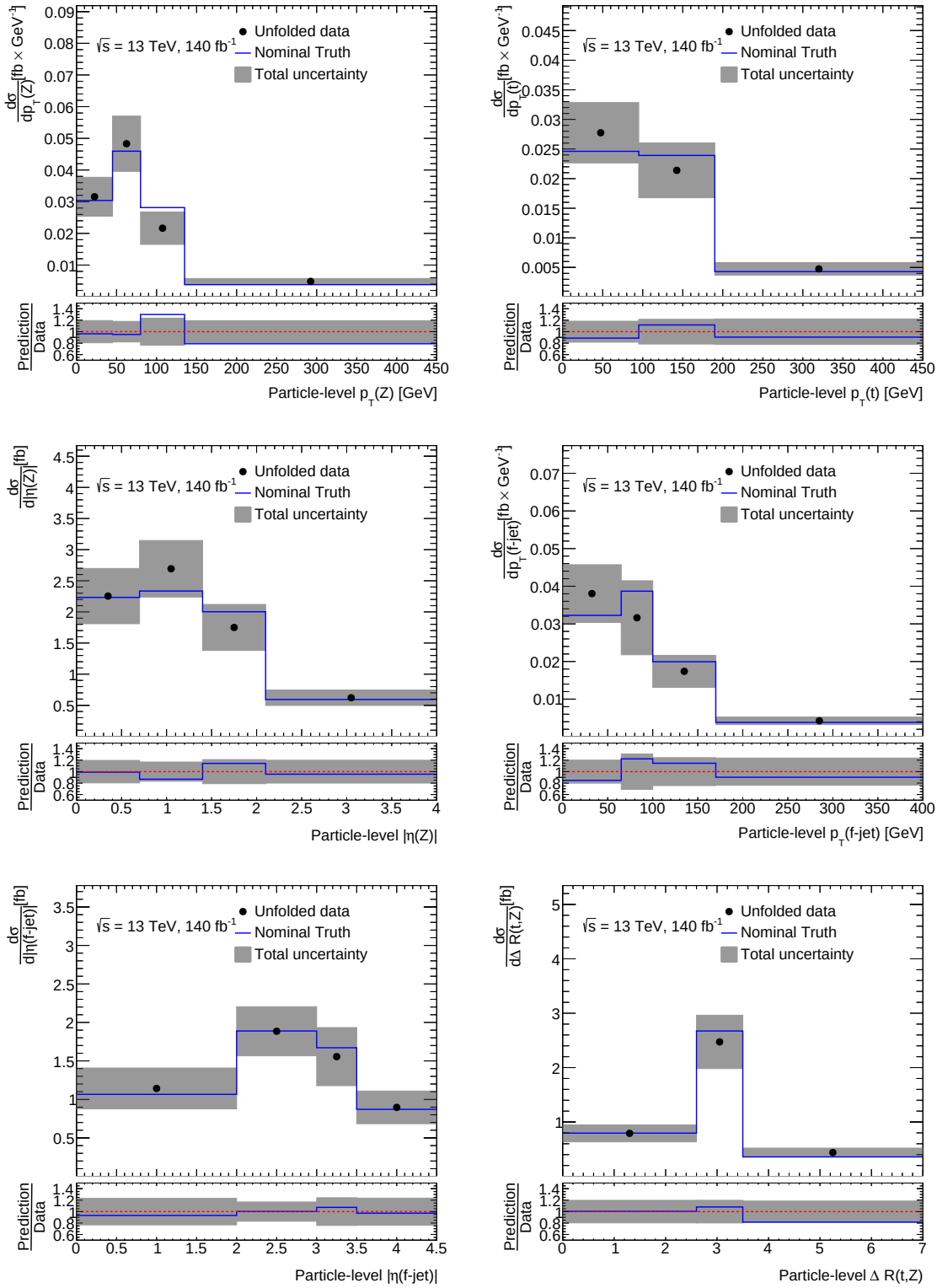


Figure 6.23: Observed differential cross-sections for variables unfolded at particle level. The uncertainty band includes both statistical and systematic uncertainties.

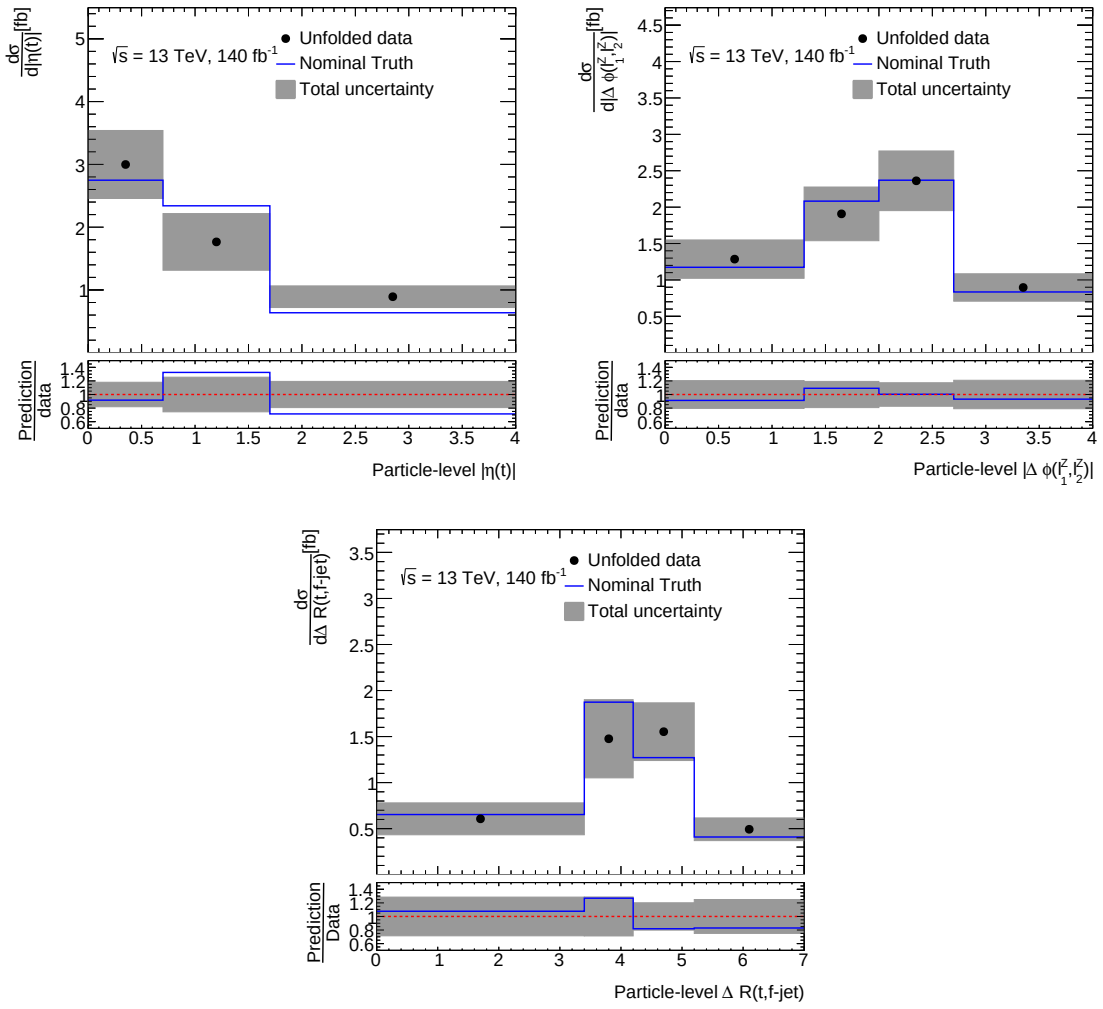


Figure 6.23: Observed differential cross-sections for variables unfolded at particle level. The uncertainty band includes both statistical and systematic uncertainties.

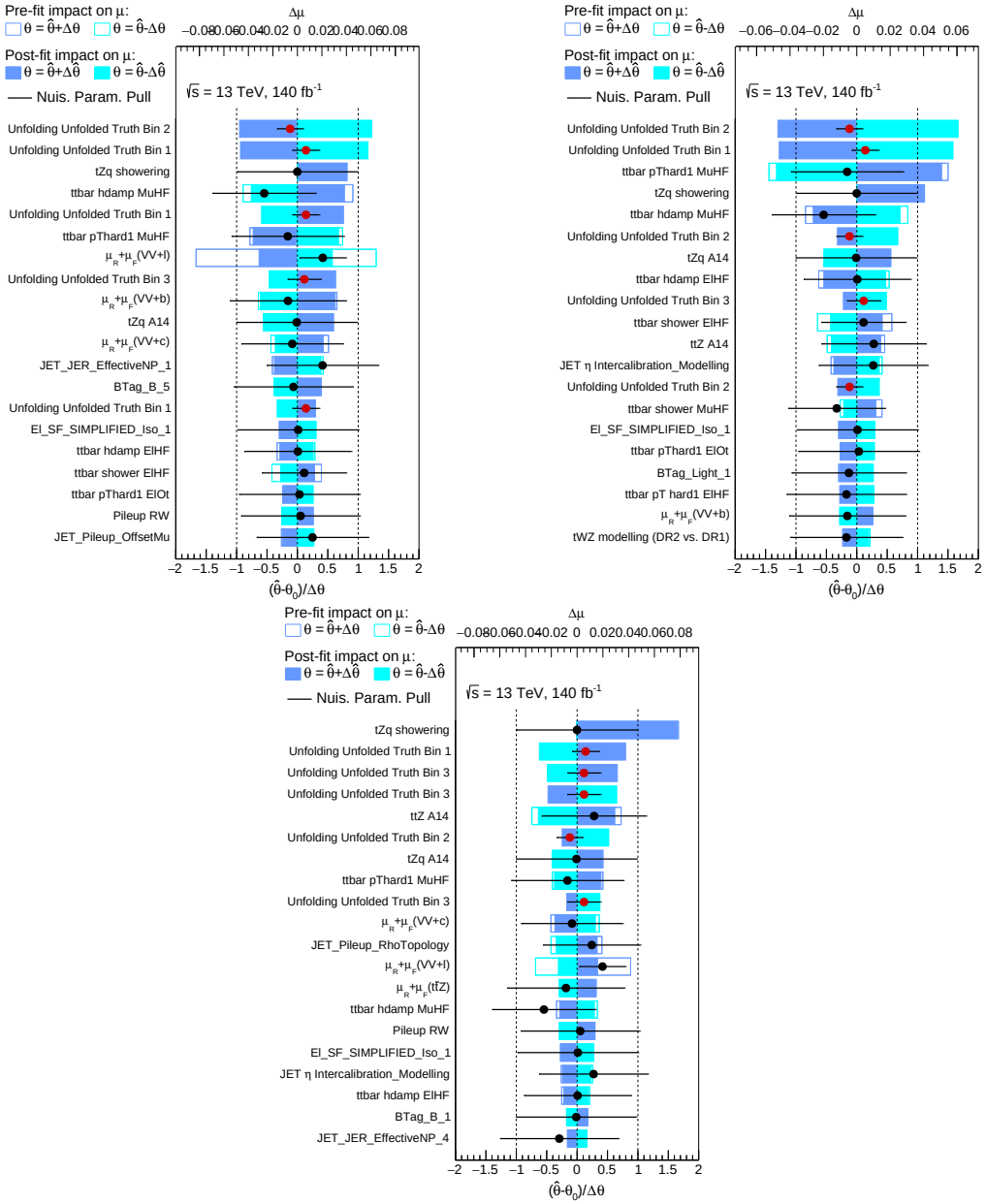


Figure 6.24: Ranking plots for each unfolded bin showing the top 20 systematics having the largest impact arranged for the $p_{\mathrm{T}}(t)$ observed distribution unfolded at parton-level. The red dots correspond to other POIs in the fit.

CHAPTER 7

Conclusion

This thesis presents the differential cross-section measurement of the associated production of a top quark and a Z boson, known as the tZq production, at $\sqrt{s} = 13\,\text{TeV}$. The dataset used has been produced by the ATLAS detector at the LHC during 2015 to 2018, corresponding to an integrated luminosity of $140.1\,\text{fb}^{-1}$. This is the first differential measurement of the tZq production performed within ATLAS. The tZq process has been predicted by the SM and its cross-section measured by ATLAS has a total uncertainty of 14 %. This measurement is performed in the trilepton decay channel.

This analysis uses profile likelihood unfolding to correct for detector effects and measure the differential cross-section. A fiducial phase space is defined at both parton and particle levels and the distributions are unfolded to those. Nine observables are chosen, constructed from η and ϕ of leptons, top quark and Z in addition to transverse momenta of top quark, Z boson, forward jet. These unfolding variables are organised into histograms with binning optimised based on the migration matrix.

Variables associated with the Z boson show an extremely diagonal migration matrix owing to the efficient and precise reconstruction of the Z boson related quantities at ATLAS. Multivariate analysis has been performed within the analysis team, to distinguish signal from background. For the differential analysis, the neural network output is used for defining regions for PLU. This is done by splitting the distribution, and it proved to be effective in reducing the per-bin statistical uncertainties by 40 %.

The robustness and reliability of PLU is checked through a series of validation tests. These tests include stress test, closure test and pull test. Stress test is to check how the method performs in presence of statistical fluctuations while the pull test validates the statistical uncertainty on parameters calculated by the fit. In the closure test, the dataset is divided into two subsets, one is used to prepare the response matrix whereas other is used as a pseudo-dataset for unfolding. The performance of PLU in all these tests has demonstrated strong agreement with the anticipated outcomes. PLU is implemented for the Asimov dataset and unfolding is performed for all the chosen variables. The results are consistent with the expectations, hence, validating the unfolding setup.

Finally, unfolding is performed on the observed data and the measured differential cross-sections are discussed in this thesis. The differential results are found to be in agreement with the theoretical predictions. There is no indication of physics beyond the SM. Moreover, normalised measurements are also presented which cause several of the main systematic uncertainties to cancel out. The results also demonstrate PLU as an effective unfolding method, especially in handling statistical fluctuations in the data. Furthermore, the ease of incorporating systematic uncertainties and additional regions for

controlling backgrounds has led to PLU becoming an increasingly widely used method in ATLAS.

The relative total uncertainties are approximately 20 %, which are reasonable considering the statistical limitations of the signal. The dominant contribution to these uncertainties is statistical errors. Among the systematic uncertainties, tZq and $t\bar{t}$ modelling systematics have higher impact on the fit results. The ability of PLU is compared with iterative Bayesian unfolding. IBU was chosen because it is a widely accepted method of unfolding within ATLAS. The consistency of the results from both these methods serves as an important cross-check of PLU.

This measurement was part of a broader effort aimed at refining the observed cross-section and providing the first differential measurement of the tZq production. A key challenge was fine-tuning the selection criteria to maximise the signal-to-background ratio, particularly given the limited signal statistics available. The selection criteria employed in this measurement represent the most optimised approach achieved through this effort. The tZq analysis team is actively working towards performing an EFT interpretation. The differential cross-sections measured in this thesis, being free from detector effects, play a crucial role in this endeavour of constraining any potential beyond SM physics.

Looking ahead, the commencement of Run 3 data taking at the LHC that began in July 2022 is expected to significantly enhance the precision of measurements, including the differential cross-section of the tZq process. The increased collision energy and improved detector capabilities are anticipated to deliver higher-quality data and larger datasets, effectively reducing statistical uncertainties in these measurements. Higher-quality data, combined with the unfolding method implemented in this thesis, can lead to more precise tests of the Standard Model.

APPENDIX A

Additional results for the Asimov dataset

Differential cross-sections for the Asimov dataset after unfolding variables other than $p_{\mathrm{T}}(t)$, are given in Fig. A.2 and Fig. A.3. Migration matrices for all variables are shown in Fig. A.5 and Fig. A.6.

Sample	SR(tZq, NN > 0)	CR(ttZ)	CR (VV)	CR(NPL-e-HF)	CR(NPL-μ-HF)	CR(NPL-e-Other)
tZq	345.0 ± 11.9	6.6 ± 5.7	77.6 ± 4.4	9.3 ± 1.4	6.8 ± 1.0	11.6 ± 0.6
$VV + b$	190.0 ± 45.3	12.5 ± 5.2	35.7 ± 9.2	7.3 ± 1.6	5.2 ± 1.1	5.2 ± 12.3
$VV + c$	403.6 ± 445.2	16.5 ± 7.1	641.3 ± 138.7	12.5 ± 3.8	8.9 ± 2.2	80.6 ± 22.2
$VV + l$	326.8 ± 90.9	11.6 ± 5.5	5210.7 ± 1077.7	11.6 ± 2.9	8.8 ± 2.3	688.1 ± 152.2
ttZ	476.6 ± 41.4	134.8 ± 33.8	56.7 ± 11.7	26.1 ± 15.0	17.2 ± 0.5	11.5 ± 4.3
ttW	19.2 ± 0.5	1.9 ± 0.2	2.1 ± 0.2	17.1 ± 0.4	8.9 ± 0.2	0.4 ± 0.0
tWZ	86.2 ± 10.5	14.1 ± 4.8	21.2 ± 1.0	2.5 ± 0.2	1.8 ± 0.3	3.5 ± 0.2
ttH	14.9 ± 1.3	4.7 ± 0.6	1.5 ± 0.2	14.1 ± 1.2	8.3 ± 0.7	0.4 ± 0.1
Other	1.2 ± 0.6	1.4 ± 0.7	3.5 ± 1.7	0.7 ± 0.4	0.2 ± 0.1	1.0 ± 0.5
NPL-e-HF	122.3 ± 14.6	3.5 ± 2.0	128.0 ± 13.1	8231.4 ± 394.1	20.0 ± 4.7	2510.3 ± 291.8
NPL-e-Other	81.7 ± 11.8	5.4 ± 1.3	306.8 ± 43.2	1762.8 ± 707.7	19.7 ± 7.4	5376.1 ± 624.4
NPL-μ-HF	151.8 ± 13.4	5.2 ± 1.8	160.8 ± 27.1	38.7 ± 9.4	5651.8 ± 558.0	15.6 ± 4.4
NPL-Other	19.2 ± 9.9	1.2 ± 0.7	59.0 ± 32.5	21.4 ± 10.7	286.0 ± 143.2	10.9 ± 6.9
Total	2238.5 ± 501.7	219.3 ± 42.5	6704.7 ± 1119.9	10155.4 ± 842.2	6043.6 ± 578.3	8715.1 ± 968.3

(a) Pre-fit yields of different processes in the regions used in PLU fit.

Sample	SR(tZq, NN > 0)	CR(ttZ)	CR (VV)	CR(NPL-e-HF)	CR(NPL-μ-HF)	CR(NPL-e-Other)
tZq	345.0 ± 35.7	6.6 ± 5.8	77.6 ± 9.5	9.3 ± 1.6	6.8 ± 1.2	11.6 ± 1.5
$VV + b$	190.0 ± 44.4	12.5 ± 4.7	35.7 ± 8.7	7.3 ± 1.7	5.2 ± 1.1	5.2 ± 9.7
$VV + c$	403.6 ± 112.2	16.5 ± 6.2	641.3 ± 141.8	12.5 ± 3.2	8.9 ± 2.0	80.6 ± 18.3
$VV + l$	326.8 ± 44.9	11.6 ± 2.5	5210.7 ± 163.8	11.6 ± 1.6	8.8 ± 1.2	688.1 ± 31.5
ttZ	476.6 ± 119.8	134.8 ± 18.6	56.7 ± 19.2	26.1 ± 14.1	17.2 ± 3.8	11.5 ± 5.0
ttW	19.2 ± 0.5	1.9 ± 0.2	2.1 ± 0.2	17.1 ± 0.4	8.9 ± 0.2	0.4 ± 0.0
tWZ	86.2 ± 10.2	14.1 ± 4.5	21.2 ± 1.0	2.5 ± 0.2	1.8 ± 0.3	3.5 ± 0.2
ttH	14.9 ± 1.3	4.7 ± 0.5	1.5 ± 0.2	14.1 ± 1.2	8.3 ± 0.7	0.4 ± 0.1
Other	1.2 ± 0.6	1.4 ± 0.7	3.5 ± 1.7	0.7 ± 0.4	0.2 ± 0.1	1.0 ± 0.5
NPL-e-HF	122.3 ± 11.7	3.5 ± 1.4	128.0 ± 11.9	8231.4 ± 560.5	20.0 ± 4.1	2510.3 ± 222.6
NPL-e-Other	81.7 ± 9.4	5.4 ± 0.9	306.8 ± 36.0	1762.8 ± 545.8	19.7 ± 5.5	5376.1 ± 245.6
NPL-μ-HF	151.8 ± 11.5	5.2 ± 1.4	160.8 ± 24.1	38.7 ± 6.1	5651.8 ± 166.8	15.6 ± 3.9
NPL-Other	19.2 ± 9.8	1.2 ± 0.7	59.0 ± 31.4	21.4 ± 10.7	286.0 ± 141.9	10.9 ± 6.5
Total	2238.5 ± 46.4	219.3 ± 14.0	6704.7 ± 81.8	10155.4 ± 100.5	6043.6 ± 77.9	8715.1 ± 92.9
Data	2343	252	6948	9729	6062	7017

(b) Post-fit yields of different processes in the regions used in PLU fit.

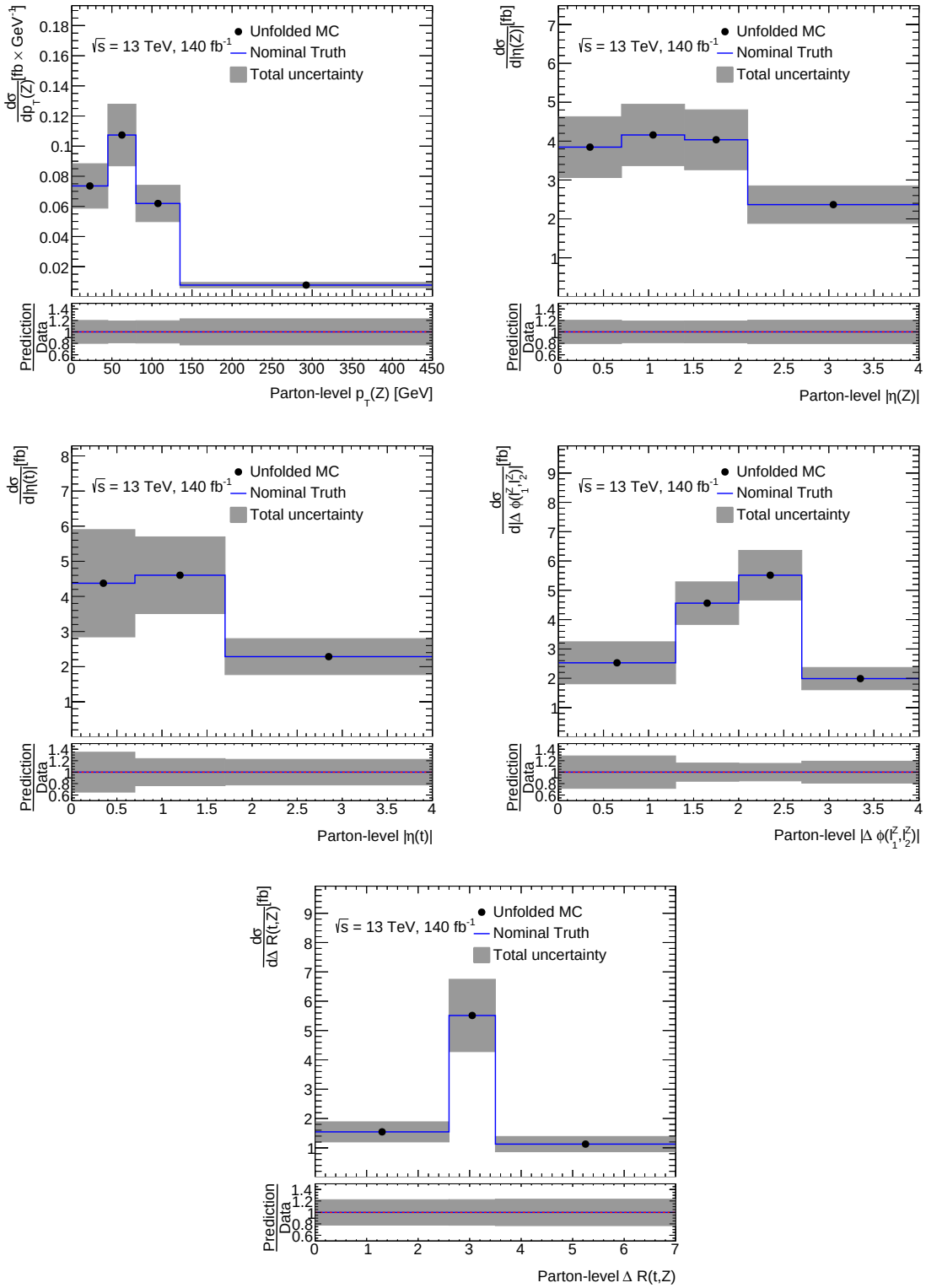


Figure A.2: Parton-level unfolded differential cross-sections for the remaining variables in case of Asimov fit. The uncertainty band includes both statistical and systematic uncertainties.

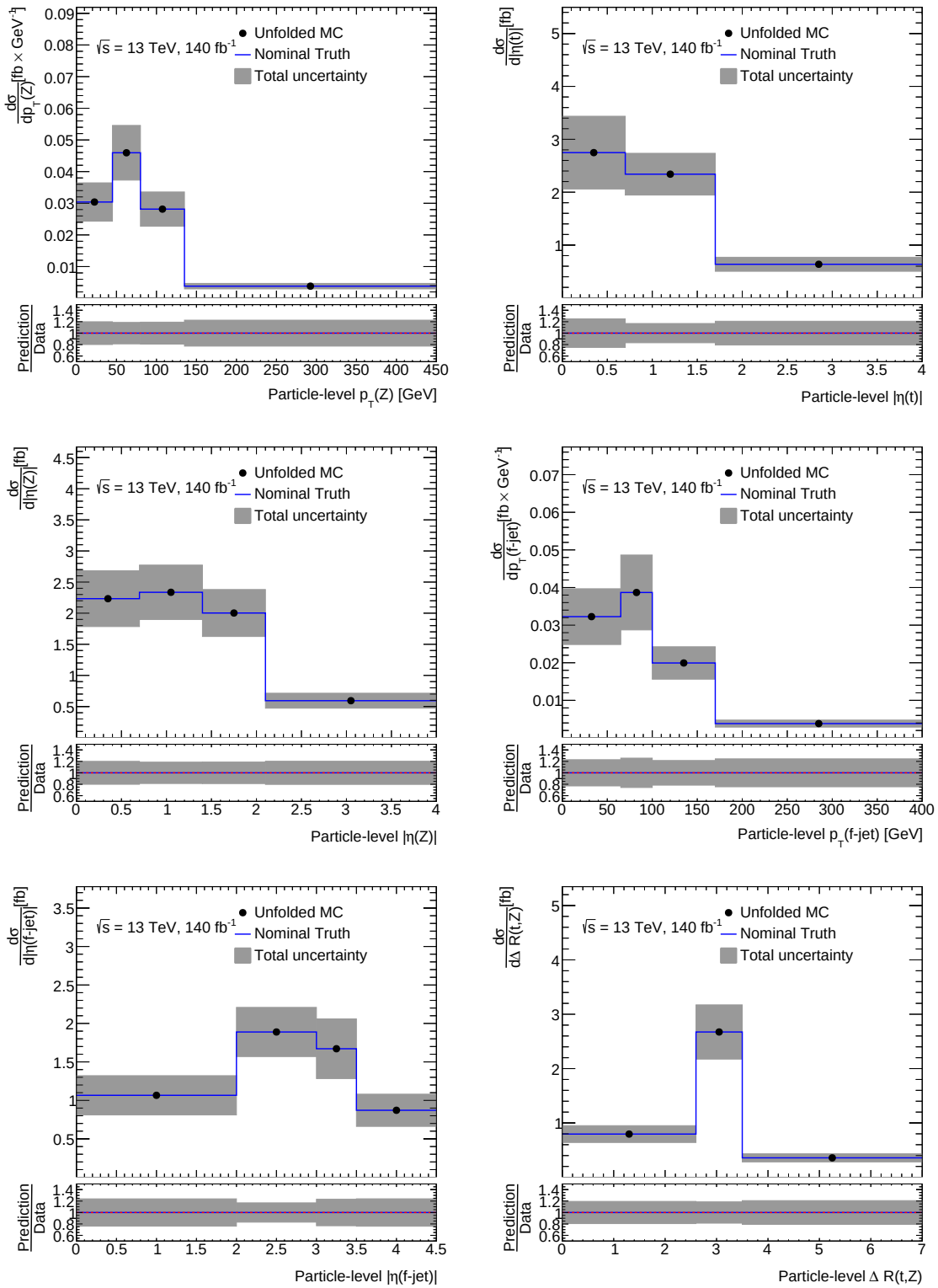


Figure A.3: Particle-level unfolded differential cross-sections for the remaining variables in case of Asimov fit. The uncertainties include both statistical and systematic uncertainties.

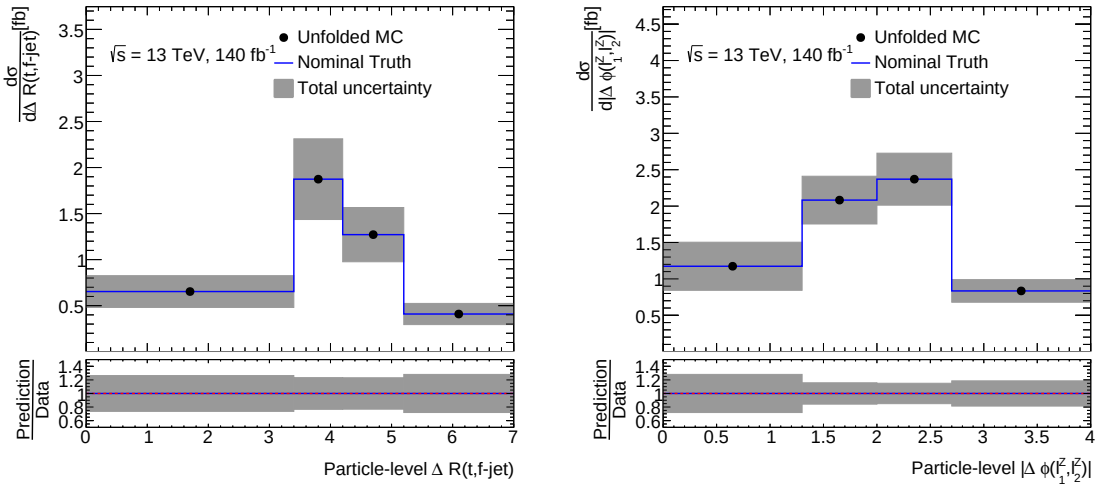


Figure A.3: Particle-level unfolded differential cross-sections for the remaining variables in case of Asimov fit. The uncertainties include both statistical and systematic uncertainties.

Splitting the signal regions for unfolding

For the profile-likelihood fit implemented in the differential cross-section measurement, the signal region is split according to the neural network score into a signal-enriched region and a signal-depleted region. In this way, we gain on the statistical uncertainties on the normalisation factors and eventually the per-bin differential cross-section. Two fits were performed without any systematics: one with a single signal region and the other with two signal regions based on the neural network score. In both cases, the fit setup is simple with no control regions and all backgrounds are constrained. The normalisation factors given by the fits are shown in Fig. A.4. It is observed that the per bin statistical uncertainties in the two-region fit are 40% lower compared to the single-region fit.

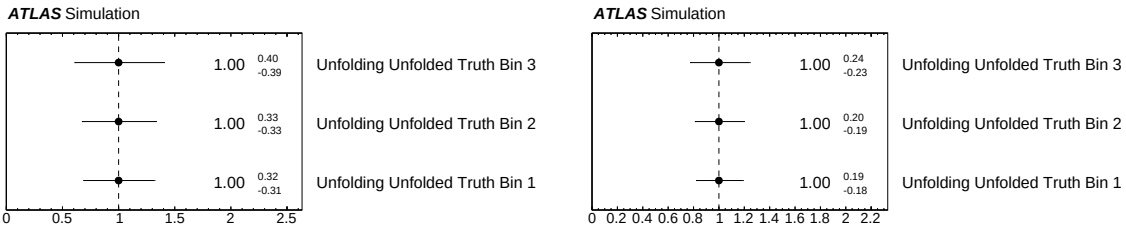


Figure A.4: A comparison between normalisation factor uncertainties in case of one single signal region (left) and in case of two signal regions based on neural network score (right)

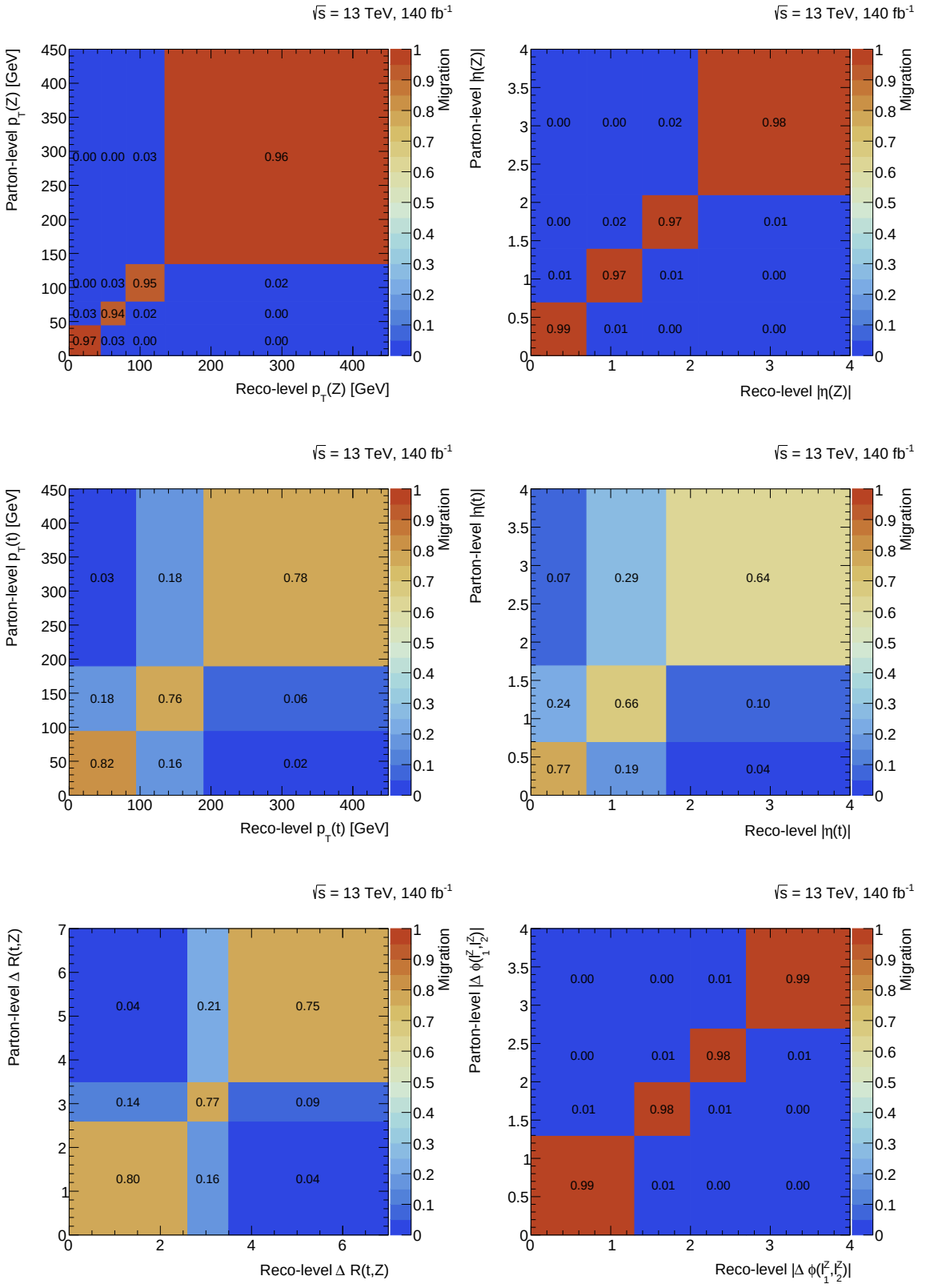


Figure A.5: Migration matrices for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$ for all variables that are unfolded at parton-level.

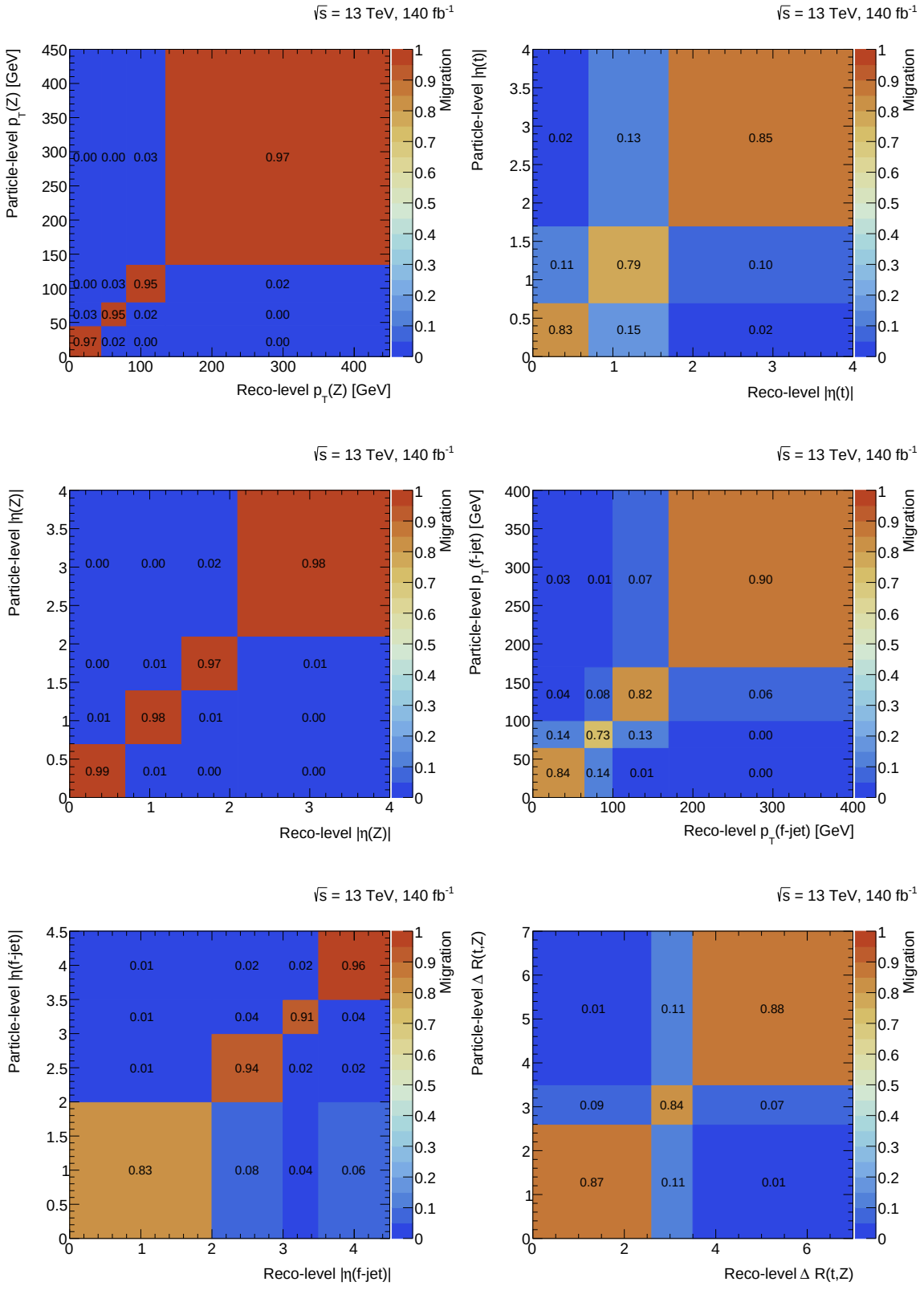


Figure A.6: Migration matrices for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$ for all variables that are unfolded at particle-level.

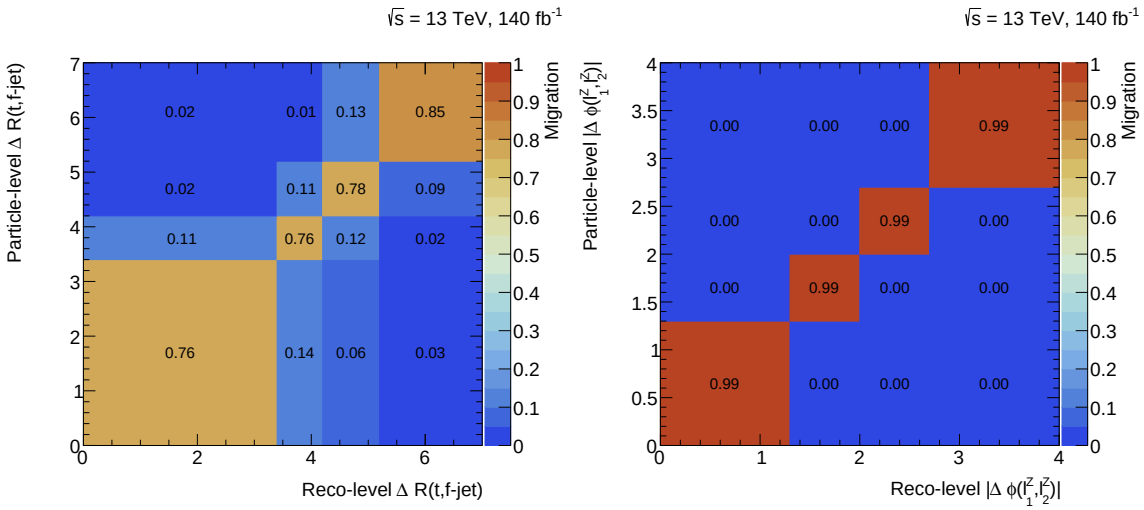


Figure A.6: Migration matrices for $O_{\mathrm{NN}}(\mathrm{SR}) > 0.7$ for all variables that are unfolded at particle-level.

APPENDIX B

Additional unfolding results for observed data

The normalised differential cross-section distributions for the observed data are shown in Fig. B.1 and Fig. B.2. In order to show the data and MC agreement, the prefit and postfit distributions for all the fit regions are shown in Fig. B.3 and Fig. B.4, respectively. The pulls and constraints on different NPs in a data fit are shown in Fig. B.5. The fit results from a normalised differential cross-section fit are summarised in Table B.1 and Table B.2.

Table B.1: Normalisation factors along with total uncertainty calculated by the normalised fit using observed data for all variables unfolded at parton level. In this fit, σ_{tot} is one of the POIs.

$p_{\text{T}}(t)$ GeV	Norm factors Value ± Total unc.
Bin-1	1.10 ± 0.16
Bin-2	0.85 ± 0.19
σ_{tot} / $\sigma_{\text{tot}}^{\text{theory}}$	1.04 ± 0.12

$\Delta R(t, Z)$	Norm factors Value ± Total unc.
Bin-1	0.97 ± 0.18
Bin-2	0.86 ± 0.19
σ_{tot} / $\sigma_{\text{tot}}^{\text{theory}}$	1.03 ± 0.12

$p_{\text{T}}(Z)$ GeV	Norm factors Value ± Total unc.
Bin-1	1.02 ± 0.16
Bin-2	1.03 ± 0.15
Bin-3	0.77 ± 0.15
σ_{tot} / $\sigma_{\text{tot}}^{\text{theory}}$	1.02 ± 0.12

$\lvert\eta(Z)\rvert$	Norm factors Value ± Total unc.
Bin-1	0.98 ± 0.16
Bin-2	1.12 ± 0.16
Bin-3	0.85 ± 0.16
σ_{tot} / $\sigma_{\text{tot}}^{\text{theory}}$	1.03 ± 0.13

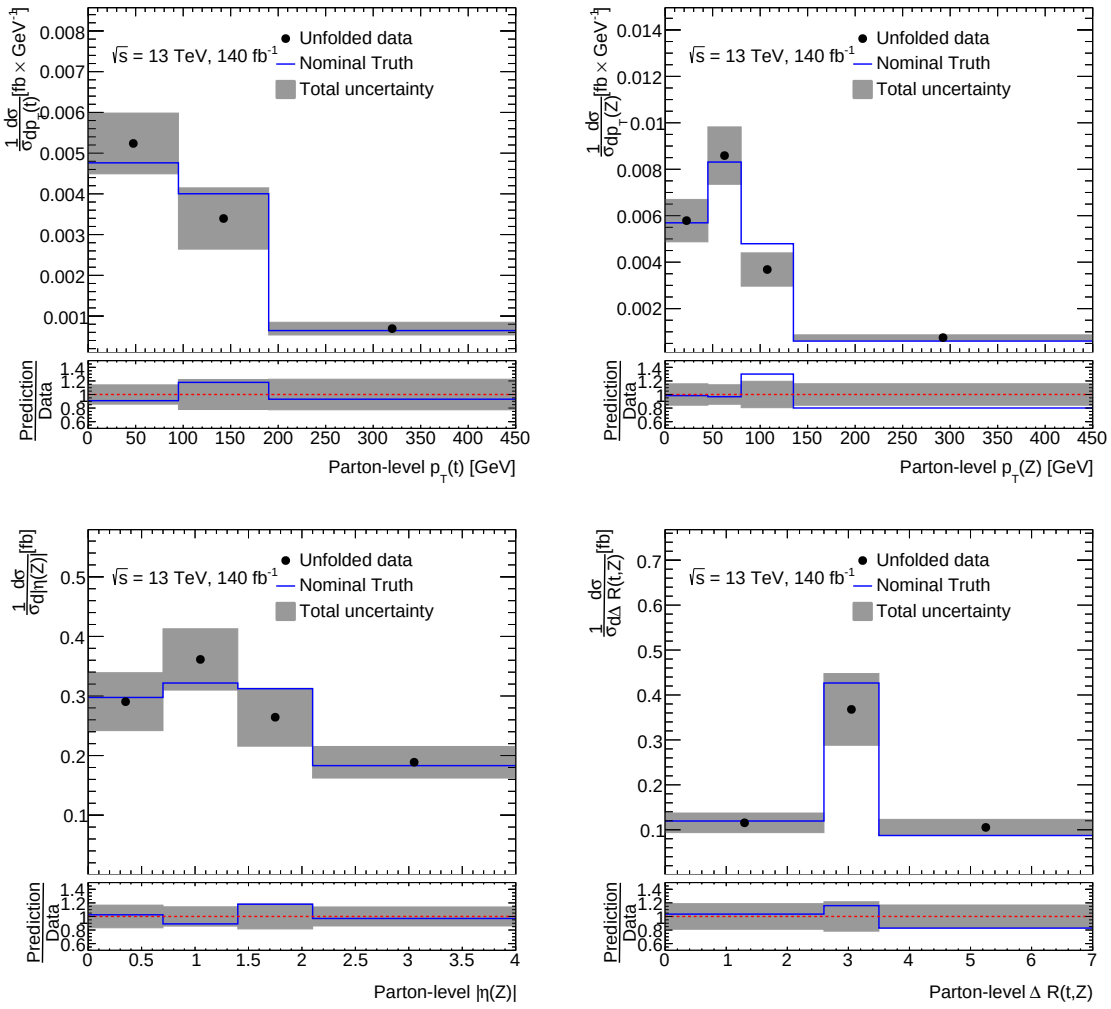


Figure B.1: Observed normalised differential cross-sections for variables unfolded at parton level. The uncertainty band includes both statistical and systematic uncertainties.

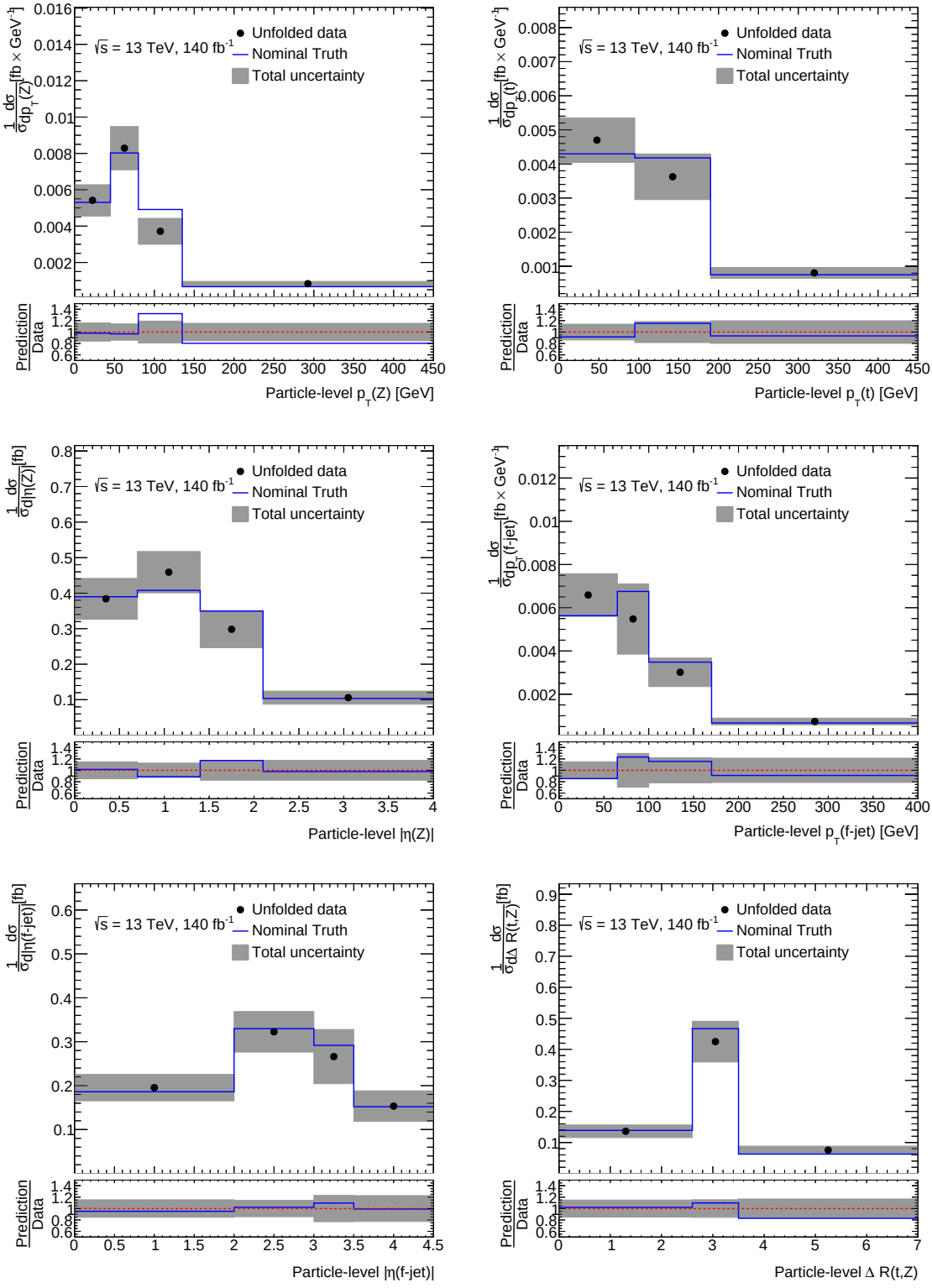


Figure B.2: Observed normalised differential cross-sections for variables unfolded at particle level. The uncertainty band includes both statistical and systematic uncertainties.

Table B.2: Normalisation factors along with total uncertainty calculated by the normalised fit using observed data for all variables unfolded at particle level. In this fit, σ_{tot} is one of the POIs.

$p_{\text{T}}(t)$ GeV	Norm factors Value ± Total unc.
Bin-1	1.09 ± 0.15
Bin-2	0.87 ± 0.16
σ_{tot} / $\sigma_{\text{tot}}^{\text{theory}}$	1.03 ± 0.12

$\Delta R(t, Z)$	Norm factors Value ± Total unc.
Bin-1	0.98 ± 0.15
Bin-2	0.91 ± 0.14
σ_{tot} / $\sigma_{\text{tot}}^{\text{theory}}$	1.02 ± 0.12

$p_{\text{T}}(Z)$ GeV	Norm factors Value ± Total unc.
Bin-1	1.02 ± 0.16
Bin-2	1.03 ± 0.15
Bin-3	0.76 ± 0.15
σ_{tot} / $\sigma_{\text{tot}}^{\text{theory}}$	1.02 ± 0.12

$\lvert\eta(Z)\rvert$	Norm factors Value ± Total unc.
Bin-1	0.98 ± 0.15
Bin-2	1.13 ± 0.14
Bin-3	0.85 ± 0.15
σ_{tot} / $\sigma_{\text{tot}}^{\text{theory}}$	1.02 ± 0.12

p_{T}(f-jet) GeV	Norm factors Value ± Total unc.
Bin-1	1.17 ± 0.18
Bin-2	0.81 ± 0.24
Bin-3	0.86 ± 0.19
σ_{tot} / $\sigma_{\text{tot}}^{\text{theory}}$	1.01 ± 0.12

$\lvert\eta$(f-jet)$\rvert$	Norm factors Value ± Total unc.
Bin-1	1.05 ± 0.16
Bin-2	0.98 ± 0.14
Bin-3	0.91 ± 0.21
σ_{tot} / $\sigma_{\text{tot}}^{\text{theory}}$	1.02 ± 0.13

$\Delta R(t$, f-jet)	Norm factors Value ± Total unc.
Bin-1	0.94 ± 0.19
Bin-2	0.79 ± 0.22
Bin-3	1.23 ± 0.24
σ_{tot} / $\sigma_{\text{tot}}^{\text{theory}}$	0.99 ± 0.13

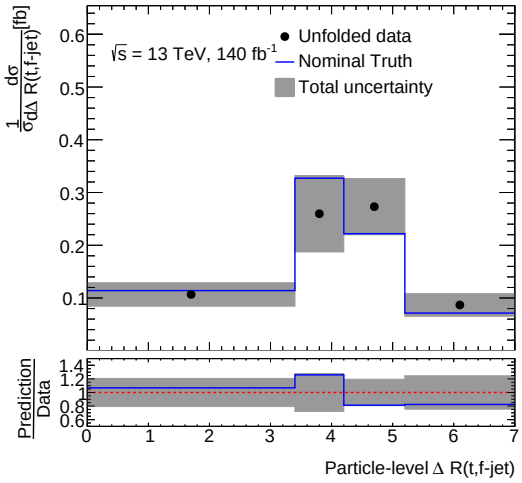
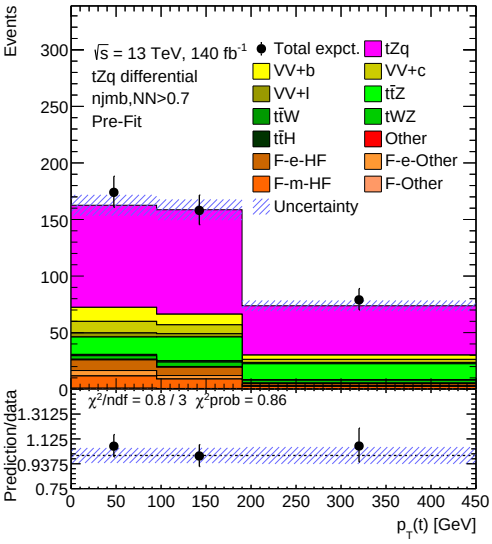
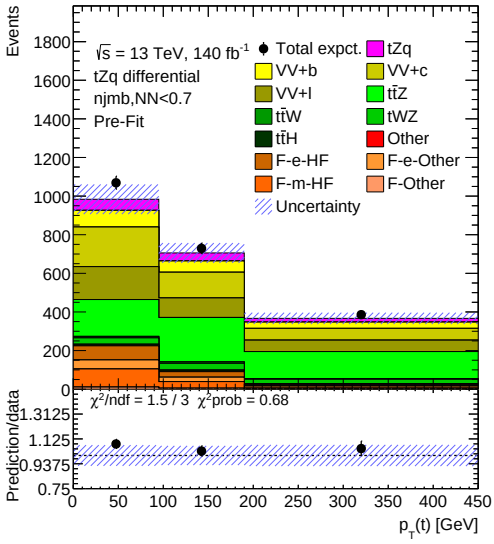
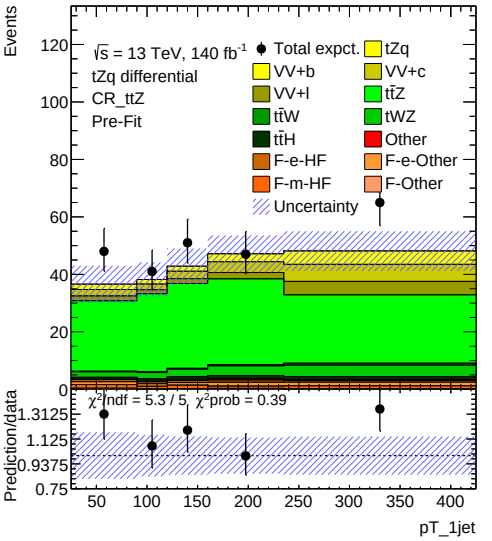
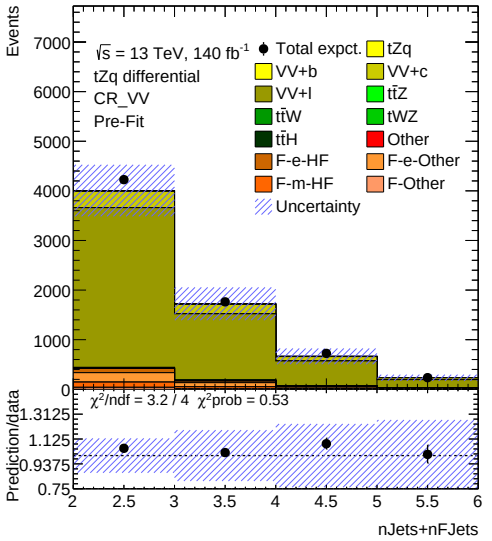


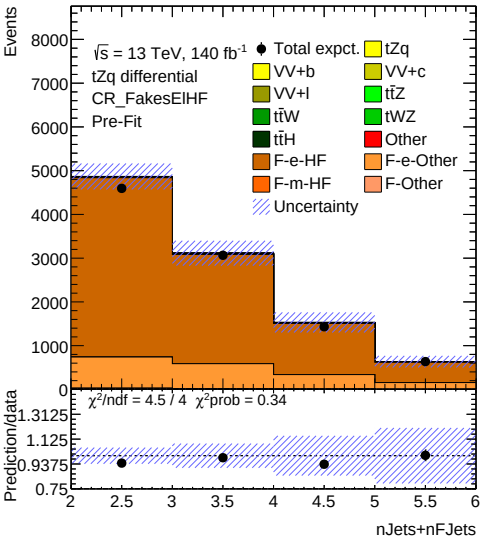
Figure B.2: Observed normalised differential cross-sections for variables unfolded at particle level. The uncertainty band includes both statistical and systematic uncertainties.

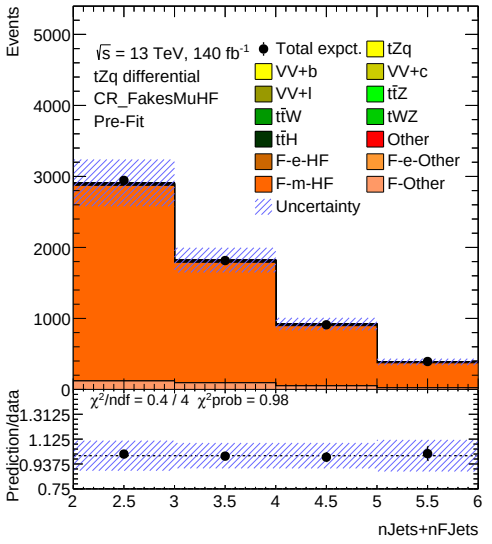

√s = 13 TeV, 140 fb-1
tZq differential
njmb,NN>0.7
Pre-Fit
Total expct.
tZq
VV+b
VV+c
VV+l
tt̄Z
tt̄W
tWZ
tt̄H
Other
F-e-HF
F-e-Other
F-m-HF
F-Other
Uncertainty
Events
χ2/ndf = 0.8 / 3 χ2prob = 0.86
Prediction/data
pT(t) [GeV]


√s = 13 TeV, 140 fb-1
tZq differential
njmb,NN<0.7
Pre-Fit
Total expct.
tZq
VV+b
VV+c
VV+l
tt̄Z
tt̄W
tWZ
tt̄H
Other
F-e-HF
F-e-Other
F-m-HF
F-Other
Uncertainty
Events
χ2/ndf = 1.5 / 3 χ2prob = 0.68
Prediction/data
pT(t) [GeV]


√s = 13 TeV, 140 fb-1
tZq differential
CR_ttZ
Pre-Fit
Total expct.
tZq
VV+b
VV+c
VV+l
tt̄Z
tt̄W
tWZ
tt̄H
Other
F-e-HF
F-e-Other
F-m-HF
F-Other
Uncertainty
Events
χ2/ndf = 5.3 / 5 χ2prob = 0.39
Prediction/data
pT_1jet


√s = 13 TeV, 140 fb-1
tZq differential
CR_VV
Pre-Fit
Total expct.
tZq
VV+b
VV+c
VV+l
tt̄Z
tt̄W
tWZ
tt̄H
Other
F-e-HF
F-e-Other
F-m-HF
F-Other
Uncertainty
Events
χ2/ndf = 3.2 / 4 χ2prob = 0.53
Prediction/data
nJets+nFJets


√s = 13 TeV, 140 fb-1
tZq differential
CR_FakesElHF
Pre-Fit
Total expct.
tZq
VV+b
VV+c
VV+l
tt̄Z
tt̄W
tWZ
tt̄H
Other
F-e-HF
F-e-Other
F-m-HF
F-Other
Uncertainty
Events
χ2/ndf = 4.5 / 4 χ2prob = 0.34
Prediction/data
nJets+nFJets


√s = 13 TeV, 140 fb-1
tZq differential
CR_FakesMuHF
Pre-Fit
Total expct.
tZq
VV+b
VV+c
VV+l
tt̄Z
tt̄W
tWZ
tt̄H
Other
F-e-HF
F-e-Other
F-m-HF
F-Other
Uncertainty
Events
χ2/ndf = 0.4 / 4 χ2prob = 0.98
Prediction/data
nJets+nFJets

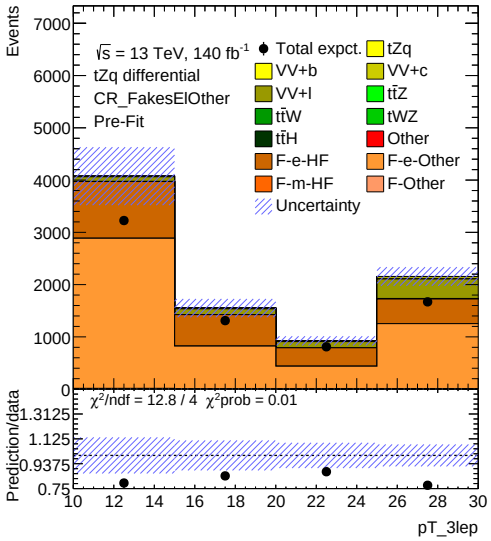
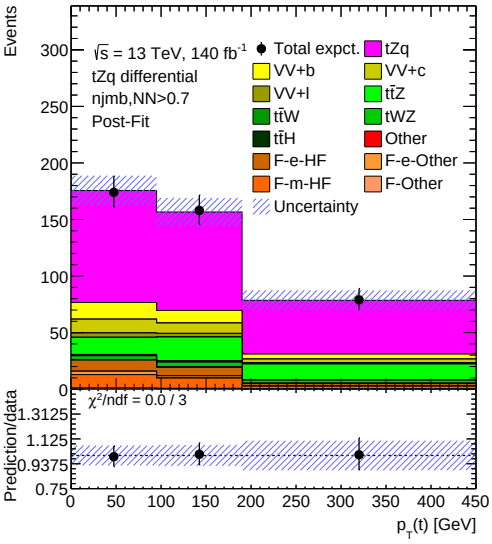
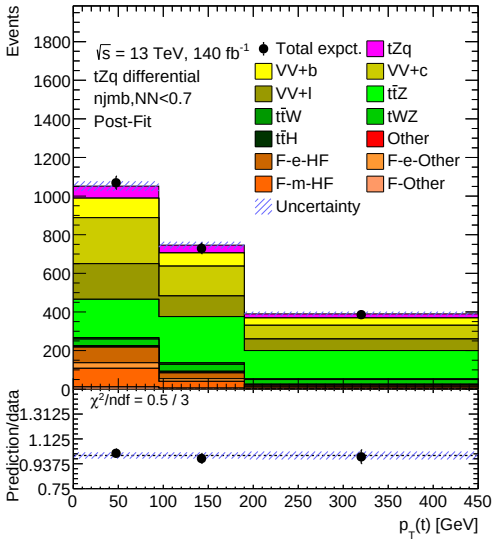
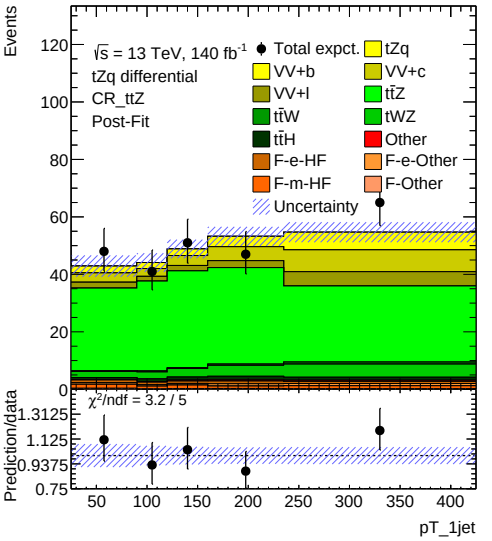
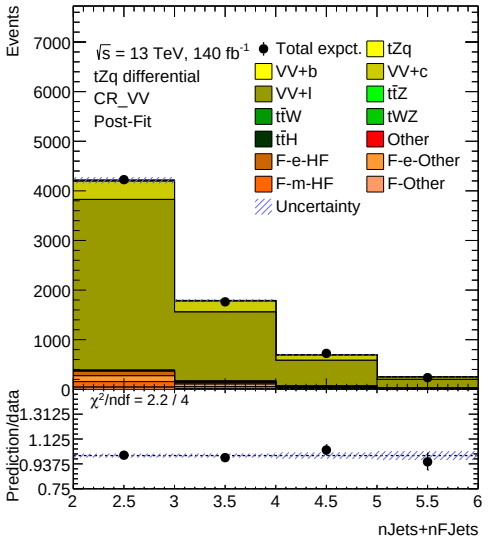


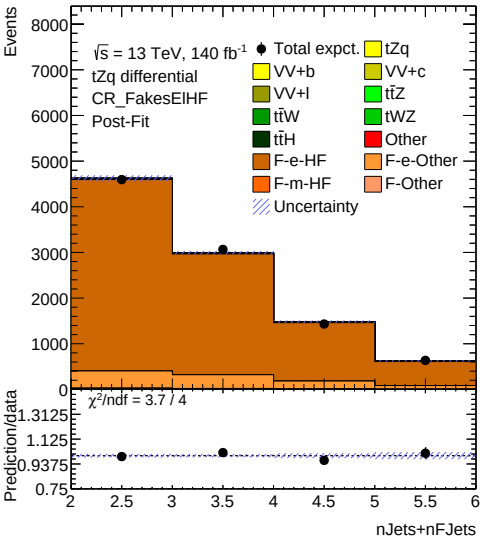
Figure B.3: Pre-fit distributions of the signal and control regions in the fit for $p_{\mathrm{T}}(t)$ unfolding are shown. The top two are the signal regions.

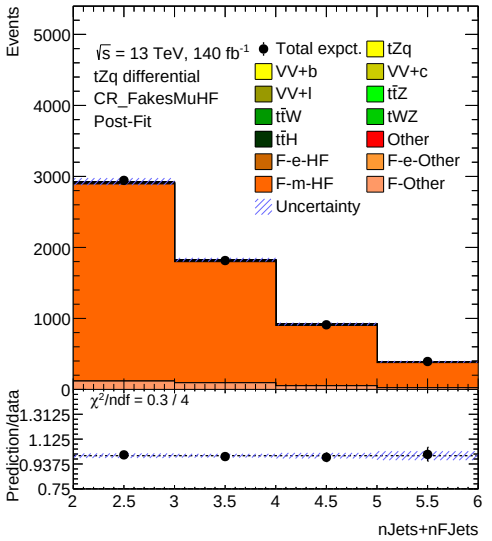

Events
√s = 13 TeV, 140 fb-1
tZq differential
njmb,NN>0.7
Post-Fit
Total expct.
tZq
VV+b
VV+c
VV+l
tt̄Z
tt̄W
tWZ
tt̄H
Other
F-e-HF
F-e-Other
F-m-HF
F-Other
Uncertainty
χ2/ndf = 0.0 / 3
Prediction/data
pT(t) [GeV]


Events
√s = 13 TeV, 140 fb-1
tZq differential
njmb,NN<0.7
Post-Fit
Total expct.
tZq
VV+b
VV+c
VV+l
tt̄Z
tt̄W
tWZ
tt̄H
Other
F-e-HF
F-e-Other
F-m-HF
F-Other
Uncertainty
χ2/ndf = 0.5 / 3
Prediction/data
pT(t) [GeV]


Events
√s = 13 TeV, 140 fb-1
tZq differential
CR_ttZ
Post-Fit
Total expct.
tZq
VV+b
VV+c
VV+l
tt̄Z
tt̄W
tWZ
tt̄H
Other
F-e-HF
F-e-Other
F-m-HF
F-Other
Uncertainty
χ2/ndf = 3.2 / 5
Prediction/data
pT_1jet


Events
√s = 13 TeV, 140 fb-1
tZq differential
CR_VV
Post-Fit
Total expct.
tZq
VV+b
VV+c
VV+l
tt̄Z
tt̄W
tWZ
tt̄H
Other
F-e-HF
F-e-Other
F-m-HF
F-Other
Uncertainty
χ2/ndf = 2.2 / 4
Prediction/data
nJets+nFJets


Events
√s = 13 TeV, 140 fb-1
tZq differential
CR_FakesElHF
Post-Fit
Total expct.
tZq
VV+b
VV+c
VV+l
tt̄Z
tt̄W
tWZ
tt̄H
Other
F-e-HF
F-e-Other
F-m-HF
F-Other
Uncertainty
χ2/ndf = 3.7 / 4
Prediction/data
nJets+nFJets


Events
√s = 13 TeV, 140 fb-1
tZq differential
CR_FakesMuHF
Post-Fit
Total expct.
tZq
VV+b
VV+c
VV+l
tt̄Z
tt̄W
tWZ
tt̄H
Other
F-e-HF
F-e-Other
F-m-HF
F-Other
Uncertainty
χ2/ndf = 0.3 / 4
Prediction/data
nJets+nFJets

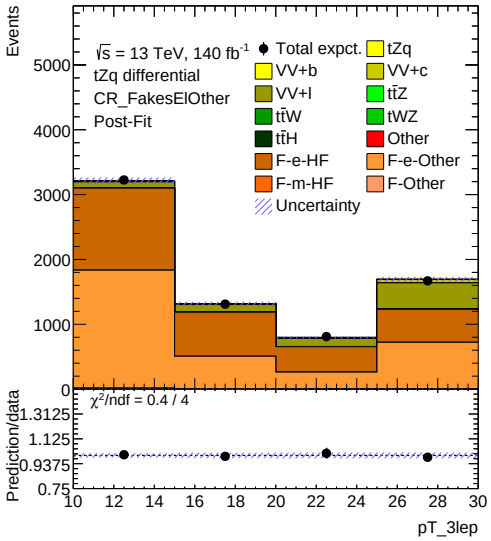


Figure B.4: Post-fit distributions of the signal and control regions in the fit for $p_{\mathrm{T}}(t)$ unfolding are shown. The top two are the signal regions. All regions show a good data to MC agreement.

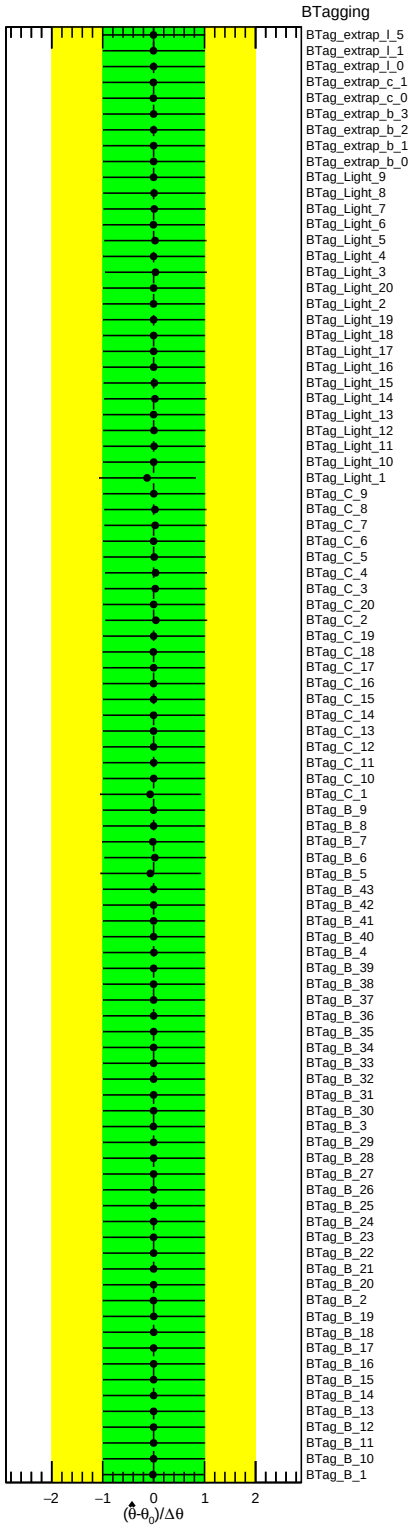


Figure B.5: Nuisance parameters corresponding to b-tagging, leptons, instrumental and theory related systematics in the fit for the $p_{\mathrm{T}}(t)$ observed distribution unfolded at parton-level.

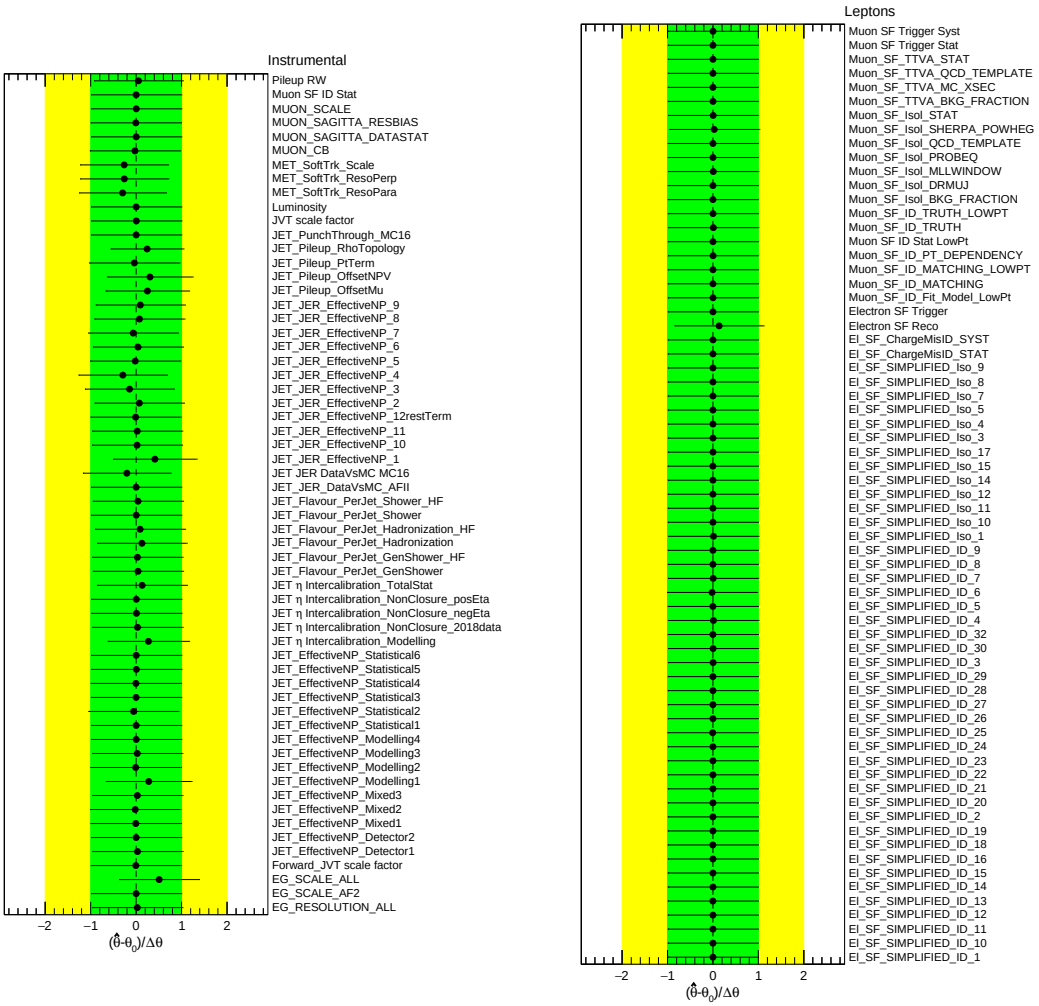


Figure B.5: Nuisance parameters corresponding to b-tagging, leptons, instrumental and theory related systematics in the fit for the $p_{\mathrm{T}}(t)$ observed distribution unfolded at parton-level.

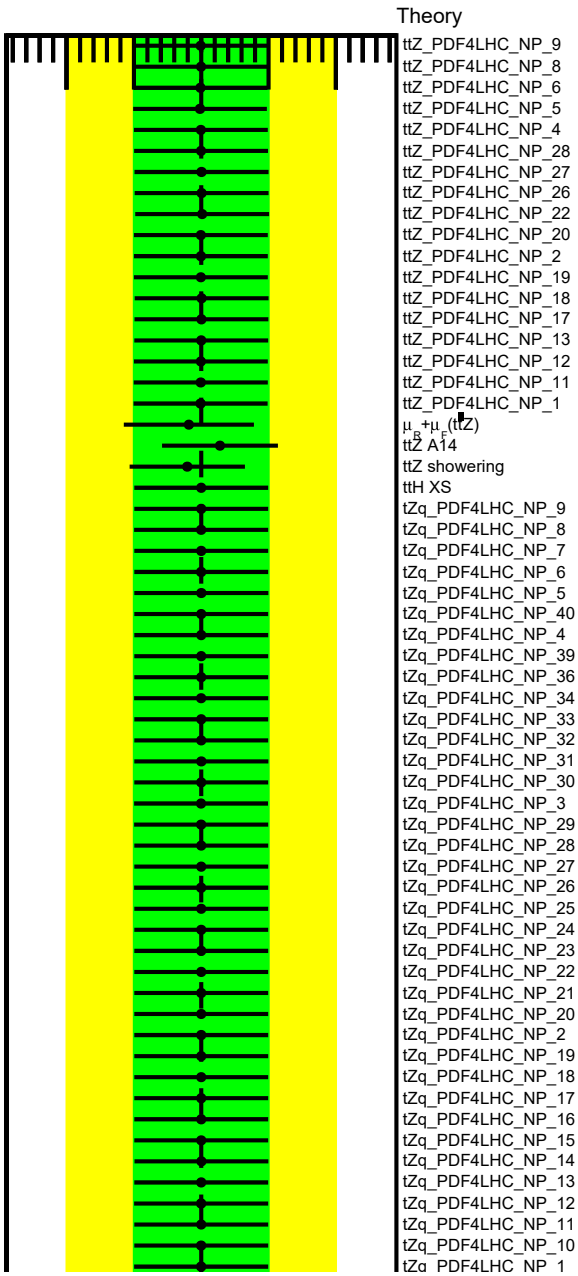


Figure B.5: Nuisance parameters corresponding to b-tagging, leptons, instrumental and theory related systematics in the fit for the $p_{\mathrm{T}}(t)$ observed distribution unfolded at parton-level.

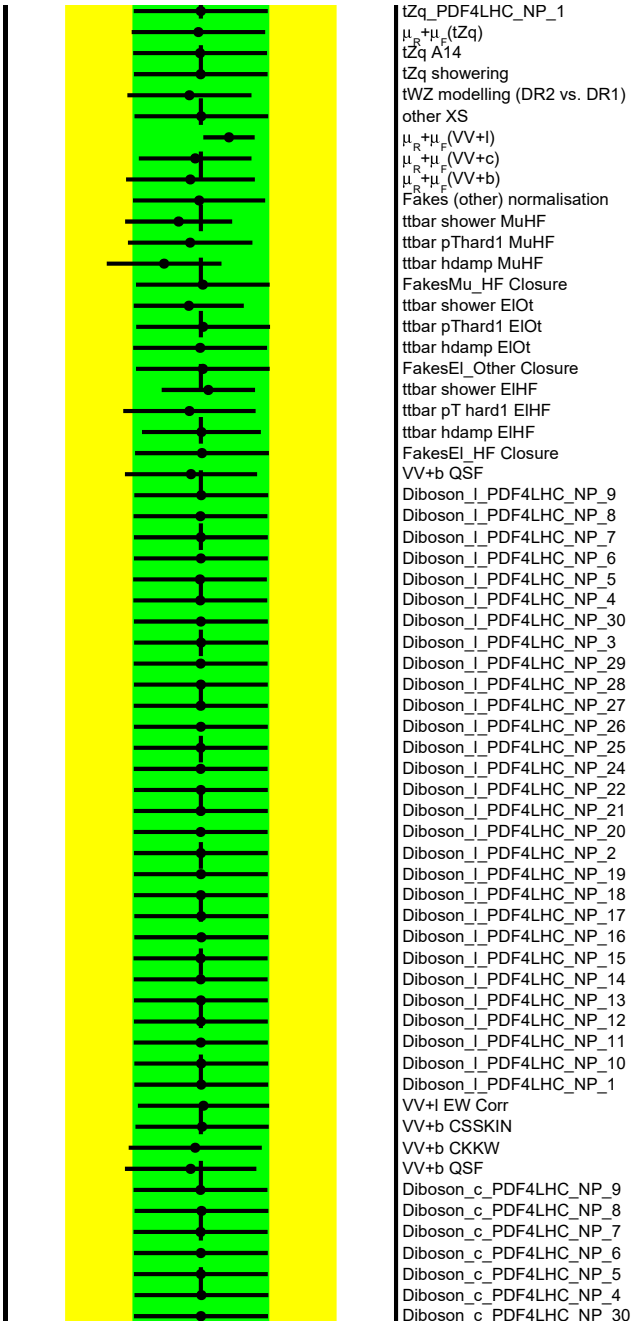


Figure B.5: Nuisance parameters corresponding to b-tagging, leptons, instrumental and theory related systematics in the fit for the $p_{\mathrm{T}}(t)$ observed distribution unfolded at parton-level.

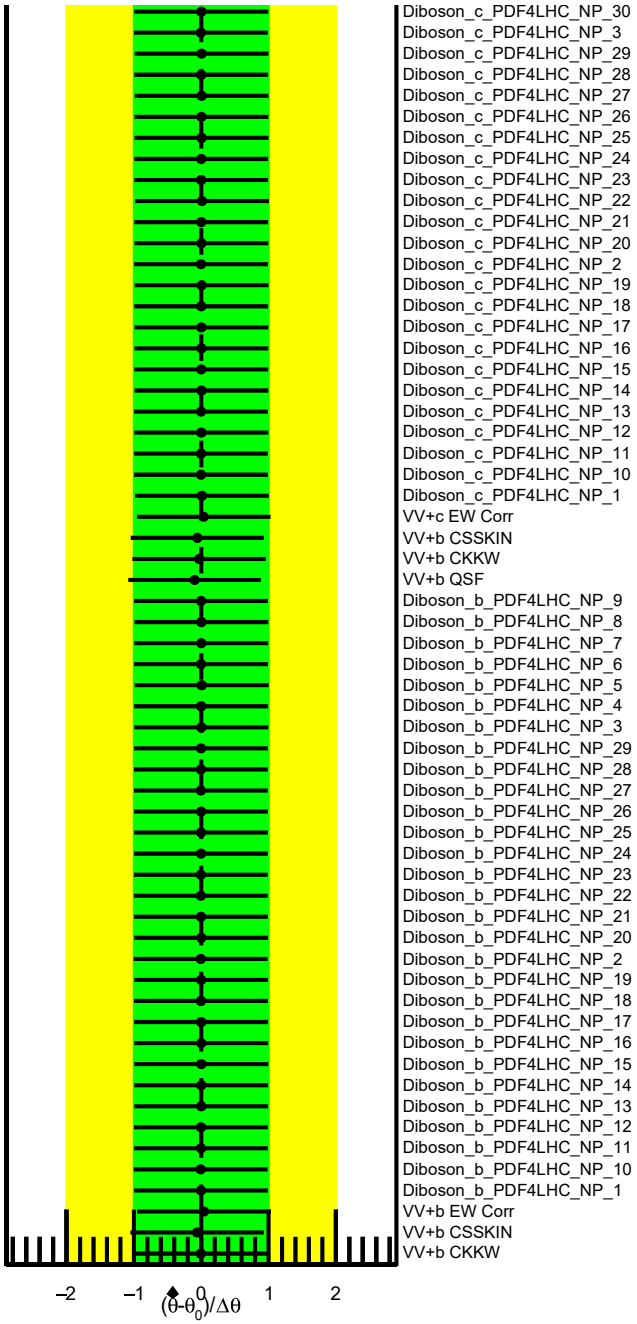


Figure B.5: Nuisance parameters corresponding to b-tagging, leptons, instrumental and theory related systematics in the fit for the $p_{\mathrm{T}}(t)$ observed distribution unfolded at parton-level.

APPENDIX C

List of nuisance parameters

Table C.1: Sources of theoretical uncertainties, with the number of NPs included in the fits.

Process	Source	Number of NPs
tZq	PDF	40
	Showering	1
	A14	1
	Scale	1
$t\bar{t}Z$	PDF	30
	Showering	1
	A14	1
	Scale	1
Diboson	PDF	90 (3 uncorrelated for each flavour)
	Scale	3 (1 uncorrelated for each flavour)
	CKKW	3 (1 uncorrelated for each flavour)
	QSF	3 (1 uncorrelated for each flavour)
	CSSKIN	3 (1 uncorrelated for each flavour)
	EW Corr	3 (1 uncorrelated for each flavour)
tWZ	DR2 vs. DR1	1
$t\bar{t}t\bar{t}$	Showering	1
	Cross-section	1
$t\bar{t}H$	Cross-section	1
$t\bar{t}$	Showering	3 (1 for ElHF, MuHF and ElOther)
	p_{T} hard	3 (1 for ElHF, MuHF and ElOther)
	h_{damp}	3 (1 for ElHF, MuHF and ElOther)
	Normalisation	1

Table C.2: Sources of instrumental uncertainties, with their grouping and the number of NPs included in the fits.

Group	Source	Number of NPs
Luminosity	Luminosity	1
Pile-up	Pile-up reweighting	1
Jets	JVT	1
	Forward JVT	1
	PunchThrough	1
	Effective NP detector	2
	Effective NP mixed	3
	Effective NP statistical	6
	Effective NP modelling	4
	η intercalibration modelling	1
	η intercalibration non-closure	4
	η intercalibration statistical	1
	Pile-up	4
	SingleParticle HighPt	1
	Flavour	6
	JER DataVsMC	2
	JER Effective NP	12
	MET SoftTrk	3
b-tagged jets	Btag B	45
	Btag C	20
	Btag Light	20
	Btag Extrapolation p_{T} B	4
	Btag Extrapolation p_{T} C	2
	Btag Extrapolation p_{T} Light	4
Leptons	Electron SF Simplified ID	32
	Electron SF Simplified Iso	17
	Electron SF Trigger	1
	Electron SF Reco	1
	Electron SF ChargeMisID	2
	Electron Resolution and Scale	3
	Muon SF ID	14
	Muon SF Isol	10
	Muon SF Trigger	2
	Muon CB	1
	Muon SF TTVA	6
	Muon Sagitta	2
	Muon Scale	1

Bibliography

[1] M. K. Gaillard, P. D. Grannis and F. J. Sciulli, *The standard model of particle physics*, Rev. Mod. Phys. **71** (2 1999) S96, URL: https://link.aps.org/doi/10.1103/RevModPhys.71.S96 (cit. on p. 3).

[2] *Standard Model of Particle Physics*, Accessed on: 30.09.2024, URL: https://en.wikipedia.org/wiki/Standard_Model (cit. on p. 4).

[3] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes and R. P. Hudson, *Experimental Test of Parity Conservation in β Decay*, Phys. Rev. **105** (1957) 1413 (cit. on p. 5).

[4] S. Weinberg, *Charge Symmetry of Weak Interactions*, Phys. Rev. **112** (4 1958) 1375, URL: https://link.aps.org/doi/10.1103/PhysRev.112.1375 (cit. on p. 5).

[5] W. Beenakker, F. Berends and S. van der Marck, *Large-angle Bhabha scattering*, Nuclear Physics B **349** (1991) 323, ISSN: 0550-3213, URL: https://www.sciencedirect.com/science/article/pii/055032139190328U (cit. on p. 5).

[6] P. Skands, "Introduction to QCD", *Searching for New Physics at Small and Large Scales* 341, URL: https://www.worldscientific.com/doi/abs/10.1142/9789814525220_0008 (cit. on p. 6).

[7] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, Wiley, 1984, ISBN: 9780471887416 (cit. on p. 7).

[8] F. Gao, C.-Y. Chen and Y.-X. Liu, *Colour Confinement: a Dynamical Phenomenon of QCD*, (2018), arXiv: 1802.08184 [hep-ph] (cit. on p. 7).

[9] A. Deur, S. J. Brodsky and G. F. de Teramond, *The QCD Running Coupling*, Nucl. Phys. **90** (2016) 1, arXiv: 1604.08082 [hep-ph] (cit. on p. 7).

[10] A. Salam, "Weak and Electromagnetic Interactions", *Elementary particle theory. Relativistic groups and analyticity. Proceedings of the Eighth Nobel Symposium*, ed. by N. Svartholm, Stockholm: Almquist & Wiksell, 1968 367 (cit. on p. 7).

[11] S. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961) 579 (cit. on p. 7).

[12] C. M. Ho, *On neutrino flavor states*, Journal of High Energy Physics **2012** (2012) 1029, URL: https://doi.org/10.1007/JHEP12(2012)022 (cit. on p. 8).

[13] *The CKM matrix and the unitarity triangle. Workshop, CERN, Geneva, Switzerland, 13-16 Feb 2002: Proceedings*, CERN Yellow Reports: Conference Proceedings, 2003, arXiv: hep-ph/0304132 (cit. on p. 8).

[14] Particle Data Group, S. Navas et al., *Review of particle physics*, Phys. Rev. D **110** (2024) 030001 (cit. on pp. 8, 12).

[15] L. Alfille, *The Large Heavy Liquid Bubble Chamber" Gargamelle": Technical Aspects and Planning Gargamelle Construction Group*, (1967) (cit. on p. 8).

[16] F. Hasert et al., *Observation of neutrino-like interactions without muon or electron in the gargamelle neutrino experiment*, Physics Letters B **46** (1973) 138, ISSN: 0370-2693, URL: https://www.sciencedirect.com/science/article/pii/0370269373904991 (cit. on p. 8).

[17] L. Di Lella and C. Rubbia, *The Discovery of the W and Z Particles*, Adv. Ser. Dir. High Energy Phys. **23** (2015) 137, URL: https://cds.cern.ch/record/2103277 (cit. on p. 8).

[18] A. Maas, *Brout-Englert-Higgs physics: From foundations to phenomenology*, Prog. Part. Nucl. Phys. **106** (2019) 132, arXiv: 1712.04721 [hep-ph] (cit. on p. 8).

[19] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (16 1964) 508, URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508 (cit. on p. 8).

[20] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1, arXiv: 1207.7214 [hep-ex] (cit. on p. 9).

[21] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30, arXiv: 1207.7235 [hep-ex] (cit. on p. 9).

[22] S. Mertens, *Direct Neutrino Mass Experiments*, Journal of Physics: Conference Series **718** (2016) 022013, URL: https://dx.doi.org/10.1088/1742-6596/718/2/022013 (cit. on p. 9).

[23] C. Burgess, C. de Rham and L. van Nierop, *The hierarchy problem and the self-localized Higgs*, Journal of High Energy Physics **2008** (2008) 061, URL: https://dx.doi.org/10.1088/1126-6708/2008/08/061 (cit. on p. 9).

[24] P. Bechtle et al., *Bottoms up: The Standard Model Effective Field Theory from a model perspective*, Studies in History and Philosophy of Science **92** (2022) 129, ISSN: 0039-3681, URL: https://www.sciencedirect.com/science/article/pii/S0039368122000140 (cit. on p. 9).

[25] R. D. Ball et al., *Parton distributions with LHC data*, Nuclear Physics B **867** (2013) 244, ISSN: 0550-3213 (cit. on p. 11).

[26] J. Rojo, *PDF4LHC recommendations for Run II*, PoS **DIS2016** (2016) 018, arXiv: 1606.08243 [hep-ph] (cit. on p. 11).

[27] ATLAS Experiment, *Public Results*, 2024,
URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2#Online_Luminosity_Summary_Plots (cit. on p. 11).

[28] ATLAS Experiment, *Public Results*, 2024,
URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2#Annual_plots_on_the_Number_of_Pi (cit. on p. 13).

[29] F. Abe et al.,
Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab,
Phys. Rev. Lett. **74** (14 1995) 2626,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.2626 (cit. on p. 12).

[30] S. Abachi et al., *Observation of the Top Quark*, Phys. Rev. Lett. **74** (14 1995) 2632,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.74.2632 (cit. on p. 12).

[31] M. Gallinaro, *Top quark physics: A tool for discoveries*,
Journal of Physics: Conference Series **447** (2013) 012012,
URL: https://dx.doi.org/10.1088/1742-6596/447/1/012012 (cit. on p. 13).

[32] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*,
Phys. Lett. B **761** (2016) 136, arXiv: 1606.02699 [hep-ex] (cit. on p. 14),
Erratum: Phys. Lett. B **772** (2017) 879.

[33] F. Fabbri, *Top pair production measurements at ATLAS*, Nuclear and Particle Physics Proceedings **282-284** (2017) 63, 19th International Conference in Quantum Chromodynamics,
ISSN: 2405-6014, URL:
https://www.sciencedirect.com/science/article/pii/S2405601416302607 (cit. on p. 14).

[34] ATLAS Collaboration, *Top cross section summary plots - April 2024*,
ATL-PHYS-PUB-2024-006, 2024, URL: https://cds.cern.ch/record/2896104 (cit. on pp. 15–17).

[35] M. Cristinziani and M. Mulders, *Top-quark physics at the Large Hadron Collider*,
Journal of Physics G: Nuclear and Particle Physics **44** (2017) 063001,
URL: https://dx.doi.org/10.1088/1361-6471/44/6/063001 (cit. on p. 14).

[36] ATLAS Collaboration, *Measurement of single top-quark production in association with a W boson pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*,
Phys. Rev. D **110** (2024) 072010, arXiv: 2407.15594 [hep-ex] (cit. on p. 14).

[37] J. Andrea and N. Chanon, *Single-Top Quark Physics at the LHC: From Precision Measurements to Rare Processes and Top Quark Properties*, Universe **9** (2023) 439,
arXiv: 2307.14044 [hep-ex] (cit. on p. 14).

[38] G. Aad et al., *Observation of Single-Top-Quark Production in Association with a Photon Using the ATLAS Detector*, Phys. Rev. Lett. **131** (18 2023) 181901,
URL: https://link.aps.org/doi/10.1103/PhysRevLett.131.181901 (cit. on p. 16).

[39] A. Hayrapetyan et al., *Evidence for tWZ production in proton-proton collisions at s=13 TeV in multilepton final states*, Phys. Lett. B **855** (2024) 138815, arXiv: 2312.11668 [hep-ex] (cit. on p. 16).

[40] ATLAS Collaboration, *Search for flavor-changing neutral-current couplings between the top quark and the Z boson with proton–proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **108** (2023) 032019, arXiv: 2301.11605 [hep-ex] (cit. on p. 16).

[41] ATLAS Collaboration, *Observation of four-top-quark production in the multilepton final state with the ATLAS detector*, Eur. Phys. J. C **83** (2023) 496, arXiv: 2303.15061 [hep-ex] (cit. on p. 16).

[42] ATLAS Collaboration, *Top Quarks + X Summary Plots April 2024*, ATL-PHYS-PUB-2024-005, 2024, URL: https://cds.cern.ch/record/2896021 (cit. on p. 17).

[43] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001 (cit. on p. 20).

[44] J. Haffner, *The CERN accelerator complex. Complexe des accélérateurs du CERN*, (2013), General Photo, URL: https://cds.cern.ch/record/1621894 (cit. on p. 20).

[45] L. Rossi, *The LHC Superconducting Magnets*, (2003), URL: https://cds.cern.ch/record/630341 (cit. on p. 21).

[46] V. Baglin, G. Bregliozzi, J. M. Jimenez and G. Lanza, *Synchrotron Radiation in the LHC Vacuum System*, (2011), URL: https://cds.cern.ch/record/1407539 (cit. on p. 21).

[47] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003 (cit. on pp. 21, 24).

[48] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004 (cit. on p. 21).

[49] *The LHCb Detector at the LHC*, Journal of Instrumentation **3** (2008) S08005, URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08005 (cit. on p. 21).

[50] *The ALICE experiment at the CERN LHC*, Journal of Instrumentation **3** (2008) S08002, URL: https://dx.doi.org/10.1088/1748-0221/3/08/S08002 (cit. on p. 21).

[51] J.-L. Caron, "Overall view of LHC experiments.. Vue d'ensemble des experiences du LHC.", AC Collection. Legacy of AC. Pictures from 1992 to 2002., 1998, URL: https://cds.cern.ch/record/841555 (cit. on p. 21).

[52] J. Pequenao, "Computer generated image of the whole ATLAS detector", 2008, URL: https://cds.cern.ch/record/1095924 (cit. on p. 22).

[53] I. Neutelings, "CMS coordinate system", URL: https://tikz.net/axis3d_cms/ (cit. on p. 23).

[54] *ATLAS inner detector: Technical Design Report, 1*, Technical design report. ATLAS, Geneva: CERN, 1997, URL: https://cds.cern.ch/record/331063 (cit. on p. 23).

[55] ATLAS Collaboration, *The ATLAS Inner Detector commissioning and calibration*, Eur. Phys. J. C **70** (2010) 787, arXiv: 1004.5293 [hep-ex] (cit. on p. 23).

[56] D. Barberis, *ATLAS Inner Detector developments*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **446** (2000) 331, ISSN: 0168-9002, URL: https://www.sciencedirect.com/science/article/pii/S0168900200000413 (cit. on p. 23).

[57] ATLAS Collaboration, "Experiment Briefing: Keeping the ATLAS Inner Detector in perfect alignment", General Photo, 2020, URL: https://cds.cern.ch/record/2723878 (cit. on p. 24).

[58] ATLAS Collaboration, *Study of the material of the ATLAS inner detector for Run 2 of the LHC*, JINST **12** (2017) P12009, arXiv: 1707.02826 [hep-ex] (cit. on pp. 23, 27).

[59] A. Vogel, *ATLAS Transition Radiation Tracker (TRT): Straw tube gaseous detectors at high rates*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **732** (2013) 277, Vienna Conference on Instrumentation 2013, ISSN: 0168-9002, URL: https://www.sciencedirect.com/science/article/pii/S0168900213009960 (cit. on p. 23).

[60] J. Pequenao, "Computer Generated image of the ATLAS calorimeter", 2008, URL: https://cds.cern.ch/record/1095927 (cit. on p. 25).

[61] *ATLAS Liquid Argon Calorimeter Phase-II Upgrade: Technical Design Report*, tech. rep., CERN, 2017, URL: https://cds.cern.ch/record/2285582 (cit. on p. 25).

[62] P. Francavilla and O. behalf of the ATLAS Collaboration), *The ATLAS Tile Hadronic Calorimeter performance at the LHC*, Journal of Physics: Conference Series **404** (2012) 012007, URL: https://dx.doi.org/10.1088/1742-6596/404/1/012007 (cit. on p. 25).

[63] J. P. Archambault et al., *Energy calibration of the ATLAS Liquid Argon Forward Calorimeter*, Journal of Instrumentation **3** (2008) P02002, URL: https://dx.doi.org/10.1088/1748-0221/3/02/P02002 (cit. on p. 25).

[64] A. Artamonov et al., *The ATLAS Forward Calorimeter*, Journal of Instrumentation **3** (2008) P02010, URL: https://dx.doi.org/10.1088/1748-0221/3/02/P02010 (cit. on p. 25).

[65] S. Palestini, *The muon spectrometer of the ATLAS experiment (talk presented at the "8th Topical Seminar on Innovative Particle and Radiation Detectors", 21-24 Oct. 2002, Siena, Italy)*, (2003) (cit. on p. 25).

[66] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004, arXiv: 2007.12539 [hep-ex] (cit. on p. 26).

[67] W. Buttinger, *The ATLAS Level-1 Trigger System*, Journal of Physics: Conference Series **396** (2012) 012010, URL: https://dx.doi.org/10.1088/1742-6596/396/1/012010 (cit. on p. 26).

[68] E. Simioni et al., *Upgrade of the ATLAS Level-1 Trigger with event topology information*, Journal of Physics: Conference Series **664** (2015) 082052, URL: https://dx.doi.org/10.1088/1742-6596/664/8/082052 (cit. on p. 26).

[69] Y. Nakahama and on behalf of the ATLAS Collaboration, *The ATLAS Trigger System: Ready for Run-2*, Journal of Physics: Conference Series **664** (2015) 082037, URL: https://dx.doi.org/10.1088/1742-6596/664/8/082037 (cit. on p. 26).

[70] C. Gabaldon, *Performance of the ATLAS Trigger System*, Journal of Instrumentation **7** (2012) C01092, URL: https://dx.doi.org/10.1088/1748-0221/7/01/C01092 (cit. on p. 27).

[71] E. Simioni, *The Hardware Topological Trigger of ATLAS: Commissioning and Operations*, (2018), URL: https://cds.cern.ch/record/2305791 (cit. on p. 27).

[72] S. Mehlhase, *ATLAS detector slice (and particle visualisations)*, (2021), URL: https://cds.cern.ch/record/2770815 (cit. on p. 28).

[73] 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** (1987) 444, ISSN: 0168-9002, URL: https://www.sciencedirect.com/science/article/pii/0168900287908874 (cit. on p. 27).

[74] T. Cornelissen et al., *The new ATLAS track reconstruction (NEWT)*, Journal of Physics: Conference Series **119** (2008) 032014, URL: https://dx.doi.org/10.1088/1742-6596/119/3/032014 (cit. on p. 28).

[75] F. Meloni, *Primary vertex reconstruction with the ATLAS detector*, Prepared for submission to JINST (2016), URL: https://cds.cern.ch/record/2222390/files/ATL-PHYS-PROC-2016-163.pdf (cit. on p. 29).

[76] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex] (cit. on p. 29).

[77] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex] (cit. on pp. 30, 31).

[78] W. Lampl et al., *Calorimeter Clustering Algorithms: Description and Performance*, tech. rep., CERN, 2008, URL: https://cds.cern.ch/record/1099735 (cit. on p. 30).

[79] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv: 1902.04655 [physics.ins-det] (cit. on p. 30).

[80] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 (cit. on p. 30).

[81] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph] (cit. on p. 31).

[82] 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** (2019) 970, arXiv: 1907.05120 [hep-ex] (cit. on p. 31).

[83] ATLAS Collaboration, *The performance of missing transverse momentum reconstruction and its significance with the ATLAS detector using 140 fb^{-1} of $\sqrt{s}$ = 13 TeV pp collisions*, Eur. Phys. J. C (2024), arXiv: 2402.05858 [hep-ex] (cit. on p. 32).

[84] ATLAS Collaboration, *Observation of the associated production of a top quark and a Z boson in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **07** (2020) 124, arXiv: 2002.07546 [hep-ex] (cit. on p. 33).

[85] I. A. Cioara, *Associated Production of a Top Quark and a Z Boson in pp Collisions at $\sqrt{s}$ = 13 TeV Using the ATLAS Detector*, BONN-IR-2018-07, PhD Thesis: University of Bonn, 2018, URL: https://hdl.handle.net/20.500.11811/7636 (cit. on p. 35).

[86] T. Sjöstrand, *Monte Carlo Generators*, 2006 European School of High-Energy Physics, ESHEP 2006 (2006) (cit. on p. 39).

[87] S. Agostinelli et al., *GEANT4 – a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250 (cit. on p. 39).

[88] J. Chapman et al., *Challenges of the ATLAS Monte Carlo production during Run 1 and beyond*, Journal of Physics: Conference Series **513** (2014) 032029, URL: https://dx.doi.org/10.1088/1742-6596/513/3/032029 (cit. on p. 39).

[89] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196, arXiv: 1512.01178 [hep-ph] (cit. on p. 39).

[90] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852, arXiv: 0710.3820 [hep-ph] (cit. on p. 39).

[91] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, *MadGraph 5: going beyond*, Journal of High Energy Physics **2011** (2011) 128, ISSN: 1029-8479, URL: https://doi.org/10.1007/JHEP06(2011)128 (cit. on p. 39).

[92] T. Gleisberg et al., *Event generation with SHERPA 1.1*, Journal of High Energy Physics **2009** (2009) 007, URL: https://dx.doi.org/10.1088/1126-6708/2009/02/007 (cit. on p. 39).

[93] T. Plehn et al., *Modern Machine Learning for LHC Physicists*, (2022), arXiv: 2211.01421 [hep-ph] (cit. on p. 40).

[94] 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** (2014) 079, arXiv: 1405.0301 [hep-ph] (cit. on p. 40).

[95] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244, arXiv: 1207.1303 [hep-ph] (cit. on p. 40).

[96] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419 (cit. on p. 40).

[97] S. Frixione, E. Laenen, P. Motylinski and B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, Journal of High Energy Physics **2007** (2007) 081, URL: https://dx.doi.org/10.1088/1126-6708/2007/04/081 (cit. on p. 40).

[98] A. Collaboration, *ATLAS delivers most precise luminosity measurement at the LHC*, https://home.cern/news/news/physics/atlas-delivers-most-precise-luminosity-measurement-lhc, 2023 (cit. on p. 41).

[99] T. Barillari and O. behalf of the ATLAS Collaboration, *Jet Energy Scale Uncertainties in ATLAS*, Journal of Physics: Conference Series **404** (2012) 012012, URL: https://dx.doi.org/10.1088/1742-6596/404/1/012012 (cit. on p. 42).

[100] ATLAS Collaboration, *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at $\sqrt{s}$ = 13 TeV*, tech. rep., CERN, 2015, URL: https://cds.cern.ch/record/2037613 (cit. on p. 42).

[101] ATLAS Collaboration, *A new tagger for the charge identification of b-jets*, ATL-PHYS-PUB-2015-040, 2015, URL: https://cds.cern.ch/record/2048132 (cit. on p. 42).

[102] S. Frixione, E. Laenen, P. Motylinski, B. R. Webber and C. D. White, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029, arXiv: 0805.3067 [hep-ph] (cit. on p. 43).

[103] ATLAS Internal, *Physics Modelling Group*, 2024, URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PhysicsModellingGroup (cit. on p. 43).

[104] ATLAS Internal, *Weak Boson Processes*, 2024, URL: https://twiki.cern.ch/twiki/bin/view/AtlasProtected/PmgWeakBosonProcesses (cit. on p. 43).

[105] J. K. Anders and M. D'Onofrio, *V+Jets theoretical uncertainties estimation via a parameterisation method*, tech. rep., CERN, 2016, URL: https://cds.cern.ch/record/2125718 (cit. on p. 43).

[106] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani–Seymour dipole factorisation*, JHEP **03** (2008) 038, arXiv: 0709.1027 [hep-ph] (cit. on p. 43).

[107] S. Kallweit, J. M. Lindert, P. Maierhöfer, S. Pozzorini and M. Schönherr, *NLO electroweak automation and precise predictions for W+multijet production at the LHC*, JHEP **04** (2015) 012, arXiv: 1412.5157 [hep-ph] (cit. on p. 43).

[108] C. Bierlich et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3*, SciPost Phys. Codebases (2022) 8, URL: https://scipost.org/10.21468/SciPostPhysCodeb.8 (cit. on p. 43).

[109] *Artificial Neural Network*, Accessed on: 26.09.2024,
URL: https://en.wikipedia.org/wiki/Artificial_neural_network (cit. on p. 45).

[110] H. B. Prosper, *Practical Statistics for Particle Physicists*,
arXiv e-prints, arXiv:1504.00945 (2015) arXiv:1504.00945, arXiv: 1504.00945 [stat.ME]
(cit. on p. 47).

[111] L. Lyons, *Statistics for Nuclear and Particle Physicists*, Cambridge University Press, 1986
(cit. on p. 47).

[112] L. Lyons, "Unfolding: Introduction", *PHYSTAT 2011*, Geneva: CERN, 2011 225
(cit. on p. 49).

[113] G. Cowan, *A survey of unfolding methods for particle physics*, Oxford University Press, 1998
(cit. on p. 52).

[114] Spanò, Francesco, *Unfolding in particle physics: a window on solving inverse problems*,
EPJ Web of Conferences **55** (2013) 03002,
URL: https://doi.org/10.1051/epjconf/20135503002 (cit. on p. 52).

[115] Schmitt, Stefan, *Data Unfolding Methods in High Energy Physics*,
EPJ Web Conf. **137** (2017) 11008,
URL: https://doi.org/10.1051/epjconf/201713711008 (cit. on p. 52).

[116] 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** (1995) 487, ISSN: 0168-9002,
URL: http://www.sciencedirect.com/science/article/pii/016890029500274X
(cit. on p. 52).

[117] *RooUnfold*, Accessed on: 12.10.2024,
URL: https://gitlab.cern.ch/RooUnfold/RooUnfold (cit. on p. 52).

[118] K. Cranmer, G. Lewis, L. Moneta, A. Shibata and W. Verkerke,
HistFactory: A tool for creating statistical models for use with RooFit and RooStats, tech. rep.,
New York U., 2012, URL: https://cds.cern.ch/record/1456844 (cit. on p. 54).

[119] *TRExFitter documentation*, Accessed on: 20.11.2024,
URL: https://trexfitter-docs.web.cern.ch/trexfitter-docs/ (cit. on p. 54).

[120] L. Moneta et al., *The RooStats Project*, PoS (), arXiv: 1009.1003 [physics.data-an]
(cit. on p. 54).

[121] W. Verkerke and D. Kirkby, *The RooFit toolkit for data modeling*, 2003,
arXiv: physics/0306116 [physics.data-an] (cit. on p. 54).

[122] F. James,
MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1, (1994)
(cit. on pp. 55, 69).

[123] G. Cowan, K. Cranmer, E. Gross and O. Vitells,
Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C **71** (2011) 1554,
arXiv: 1007.1727 [physics.data-an] (cit. on p. 55),
Erratum: Eur. Phys. J. C **73** (2013) 2501.

[124] L. Favaro et al., *How to Unfold Top Decays*, SciPost Physics (2024), https://www.thphys.uni-heidelberg.de/~plehn/pics/paper.pdf (cit. on p. 56).

[125] Particle Data Group, P. Zyla et al., *Review of Particle Physics*, Prog. Theor. Exp. Phys. **2020** (2020) 083C01 (cit. on p. 56).

[126] ATLAS Collaboration, *Include and differential cross-section measurements of the $t\bar{t}Z$ production in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector, including EFT and spin-correlation interpretations*, (2023), arXiv: 2312.04450 [hep-ex] (cit. on p. 56).

[127] LHC Top WG Internal, *ParticleLevelTopDefinitions*, 2024, URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/ParticleLevelTopDefinitions (cit. on p. 56).

[128] ATLAS Collaboration, *Flavor Tagging with Track-Jets in Boosted Topologies with the ATLAS Detector*, ATL-PHYS-PUB-2014-013, 2014, URL: https://cds.cern.ch/record/1750681 (cit. on p. 57).

[129] ATLAS Internal, *Monte Carlo Truth Classifier*, 2020, URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/MCTruthClassifier (cit. on p. 68).

[130] G. Cowan, *Statistical data analysis*, Oxford University Press, USA, 1998 (cit. on p. 69).

[131] L. Lyons, "Statistical Issues in Searches for New Physics", *2nd Large Hadron Collider Physics Conference*, 2014, arXiv: 1409.1903 [hep-ex] (cit. on p. 69).

[132] I. C. Brock, *Users Guide to Writing a Thesis in a Physics/Astronomy Institute of the Universität Bonn*, URL: https://www.pi.uni-bonn.de/brock/en/thesis-guide (visited on 25/12/2024) (cit. on p. 139).

List of Figures

List of Tables

Acknowledgements

This journey has been a memorable one. I am deeply grateful to Ian for his unwavering support and motivation, encouraging me to pursue my interests. His consistent supervision, invaluable insights, and steadfast encouragement have helped me grow not only professionally but also personally. My journey would not have been nearly as enjoyable without Ian and Rosi's efforts in fostering a wonderful workplace atmosphere, whether through board game evenings, Kegeln, or even online get-togethers. Thank you, Ian, for developing the thesis guide, which not only provided a flawless template but also spared us the need to worry about the intricate details of thesis formatting and writing. [132]. I would also like to extend my heartfelt thanks to Prof. Klaus Desch, Prof. Bastian Kubis, and Prof. Ala Bunescu for graciously agreeing to serve on my examination committee.

I would like to thank my tZq analysis team, Lidia Dell'Asta, Tom Stevenson, Muhammed Alhroob, Can Suslu, Alberto Plebani. I am grateful to Lidia for always being there for a short meeting, even at 6 pm!. A big thanks to Tom for giving invaluable insights in unfolding and for always taking care of the ntuples, which in my view is a very nasty thing requiring a lot of patience! Thank you Alberto, for our fruitful discussions at CERN and for showing me the best Italian restaurants in Geneva! I am thankful to the TRExFitter and ROOT developers, especially Thomas Dado for their timely and helpful support throughout my thesis. My software framework wouldn't have been possible without the guidance of the IT support in my institute, especially Oliver Freyermuth, who is definitely a computer-whisperer!

I believe for a PhD, getting a good research group makes your life much easier. I consider myself extremely lucky to have worked in the Brock research group who is like a family-away-from-home to me. I am thankful to each and every past and present group member to be a part of my journey. Thank you Oleh Kivernyk for helping me through my qualification task. Thanks to Tanja Holm for being a good office mate. I am thankful to Anjishnu for introducing me to profile likelihood unfolding and for taking the time to proofread my thesis. A heartfelt thank you to Rico and Christian for not only sharing your expertise but also for being there to listen when I needed to vent. I thoroughly enjoyed my time with Richard and Oliver, whether it was during meme competitions, random banter, our CERN visit, conferences, or jamborees. Thanks Oliver for reading my thesis and providing critical feedback that helped me a lot. A big thank you to Florian for going through this crazy journey with me. The banter I had with you coupled with your sharp sense of humour made the workplace really enjoyable for me.

Thanks Varsiha for our annual ATLAS-D lunches talking about our analysis. It sure was like a therapy to me! Thank you Ziad for always being a phone call away. I loved our spontaneous plans and long rides. Thanks for taking care of me since the day I landed in Germany. Thanks Divya for our funny work gossips at jamborees and our potluck meals especially during Covid-19 lockdown. I will always miss you.

Hitesh, thank you for being one of the few people I can talk without any filters. I never knew sitcoms

can help make lifelong friends! You are definitely the Ted Mosby to my Lily Aldrin. Thank you for teaching me Python and making consistent efforts to help me improve not only my code but also my self-esteem.

I would like to thank my best friend, Tilak, who always encouraged me to be the best version of myself. Our discussions on music, coffee, books and so many other things kept me sane amidst my crushing workload. I can't thank you enough for regularly visiting me, checking up on me and being one of the best cheerleaders I could ask for! I loved to share awesome experiences and of course, Spotify, with you.

I am forever grateful to one of the most important people in my life, my partner, Divyanshu. This thesis wouldn't have been in this shape if it weren't for his unconditional support, guidance, motivation and endless patience. I am always inspired by the passion you have for your work. Thank you for teaching me how to deal with lows, keeping my chin up and persevere. Thanks for cooking delicious meals for me, keeping me on track and preventing me from going insane while writing this thesis. I am very lucky to have you in my life.

A big thank you to my sister, Krishna, for her unwavering support. She is a mind-reader who always knows what I need. I couldn't have asked for a better sibling. Thank you Vaidehi for your long video-calls talking about random things that distracted me from my problems. I would like to extend my thanks to my grandparents who love me unconditionally. Weekend conversations with them always brought peace to my mind. I am thankful to my landlords, Mr. and Mrs. Hoyer who treated me as their family member and provided a peaceful environment, which has been a hospitable environment for not only this thesis but for two others as well!

Last but not the least, I am indebted to my parents who guided, supported and encouraged me to fulfill my dreams. I am thankful to them for making it possible for me to study in Germany. A heartfelt thank you to my mother for teaching me how to face challenges and showing me the value of embracing bold risks for great rewards. I am grateful to my father because his work ethic and his work-life balance is an inspiration to me. My parents have always been the *wind beneath my wings*.

I am grateful to all of them, as well as to anyone I may have inadvertently missed, who directly or indirectly crossed paths with me over the past four years. I have learned something from each and every one.