# Calibration of electron shower shapes, hadronic recoil reconstruction using deep learning algorithms and the measurement of W boson transverse momentum distribution with the ATLAS detector

Mykola Khandoga

# université
## PARIS-SACLAY

# Calibration of electron shower shapes, hadronic recoil reconstruction using deep learning algorithms and the measurement of W boson transverse momentum distribution with the ATLAS detector

**Thèse de doctorat de l'Université Paris-Saclay**

École doctorale n° 576, Particules Hadrons Énergie et Noyau : Instrumentation, Image, Cosmos et Simulation (PHENIICS)
Spécialité de doctorat: Physique des particules
Unité de recherche: Université Paris-Saclay, CEA, Département de Physique des Particules, 91191, Gif-sur-Yvette
Référent: : Faculté des sciences d'Orsay

**Thèse présentée et soutenue à Gif-sur-Yvette, le 22 septembre 2020, par**

# Mykola KHANDOGA

**Composition du jury:**

| | |
|---|---|
| **David Rousseau** | Président |
| Directeur de recherche, Laboratoire de Physique des 2 Infinis Irène Joliot Curie | |
| **Alessandro Vicini** | Rapporteur & Examinateur |
| Professeur, University of Milan | |
| **Andrew Pilkington** | Rapporteur & Examinateur |
| Professeur, University of Manchester | |
| **Aram Apyan** | Examinateur |
| Chercheur, Fermi National Accelerator Lab | |
| **Maarten Boonekamp** | Directeur de thèse |
| Ingénieur-chercheur, CEA Saclay | |
| **Fabrice Balli** | Co-encadrant |
| Ingénieur-chercheur, CEA Saclay | |

# Synthèse en français

## Introduction

La compréhension de la physique des hautes énergies à notre époque repose sur les succès de ses approches théoriques et expérimentales, réalisés dans la seconde moitié du XXe et au début du XXIe siècles. Cette thèse porte sur l'étude expérimentale des propriétés d'une particule fondamentale appelée boson W et sur l'amélioration des techniques expérimentales utilisées dans l'expérience ATLAS.

### Le Modèle Standard de la physique des particules

Le Modèle Standard (MS) de la physique des particules est une théorie quantique des champs développée dans les années 1960-1970 qui décrit la structure et les interactions de la matière au niveau fondamental. Cette théorie postule l'existence de 12 fermions élémentaires et fournit une description de leurs interactions à travers 3 des 4 forces fondamentales connues: électromagnétique, faible et forte. Les médiateurs de ces forces fondamentales sont également décrits comme des particules élémentaires (voir Fig. 01). Les prédictions du Modèle Standard dépendent de 18 paramètres qui doivent être mesurés expérimentalement.

Les bosons W et Z sont les porteurs de l'interactions faible. Le boson W a été découvert expérimentalement en 1983 au CERN, et les mesures de précision de sa masse et de ses propriétés se poursuivent depuis lors. Le boson W étant l'une des pierres angulaires du Modèle Standard, les mesures expérimentales de précision de ses propriétés permettent de tester et d'améliorer les prédictions théoriques. Ceci est la motivation principale de cette thèse.

### Le Grand Collisionneur de Hadrons (LHC) et lexpérience ATLAS

Le Large Hadron Collider (LHC) est un accélérateur de particules construit près de la ville de Genève en Suisse. Avec une circonférence de 27 km, c'est le plus grand accélérateur de particules au monde. Il est capable d'accélérer des protons jusqu'à une énergie de 6,5 TeV par particule et des ions de plomb jusqu'à 2,76 TeV par nucléon. Au cours de la deuxième phase, le LHC a délivré des faisceaux collimatés en 4 points d'interaction, à un taux de 40 millions de croisements de paquets par seconde. Les quatre points d'interaction correspondent à quatre expériences basées au LHC: ATLAS, CMS, LHCb et ALICE.

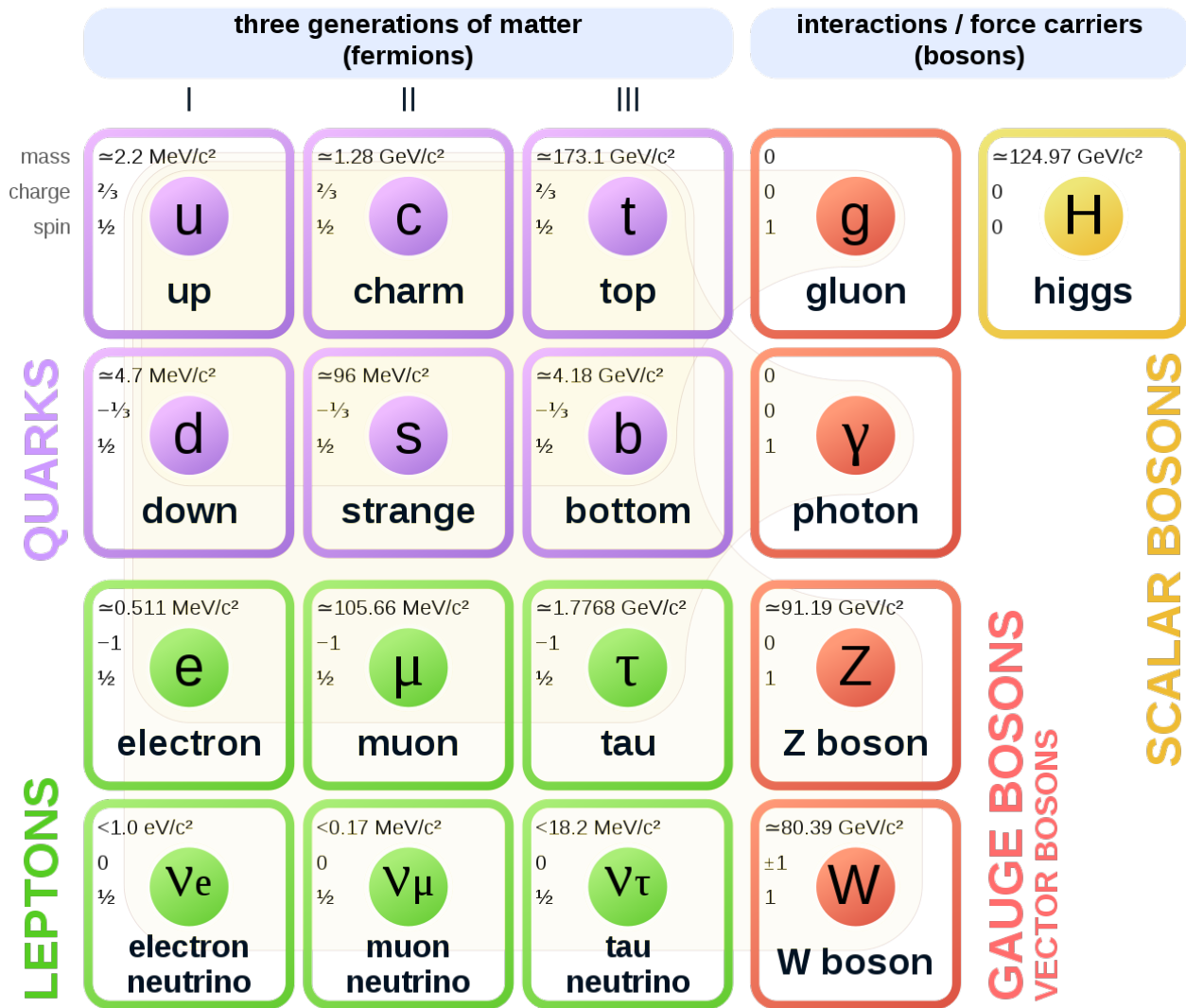# Standard Model of Elementary Particles



**Figure 01:** Liste des particules du Standard Model (SM).

L'expérience ATLAS est un détecteur de particules généraliste de forme cylindrique, de 44 m de long et de 25 m de haut. Il a une structure en forme d'oignon combinant différents types de détecteurs pour permettre une reconstruction des particules la plus efficace possible (voir Fig. 02).

Après la collision des protons au point d'interaction, les particules générées pénètrent dans le détecteur interne, qui fournit des informations sur la direction, l'impulsion et la charge des particules chargées. Ensuite, les particules atteignent le calorimètre, qui est utilisé pour reconstruire l'énergie et l'impulsion des particules neutres et chargées. Enfin, les particules capables de pénétrer dans le calorimètre atteignent les chambres à muons qui constituent le système de détection le plus à l'extérieur du calorimètre ATLAS. Les chambres à muons sont utilisées pour reconstruire la direction, l'impulsion et la charge des muons.
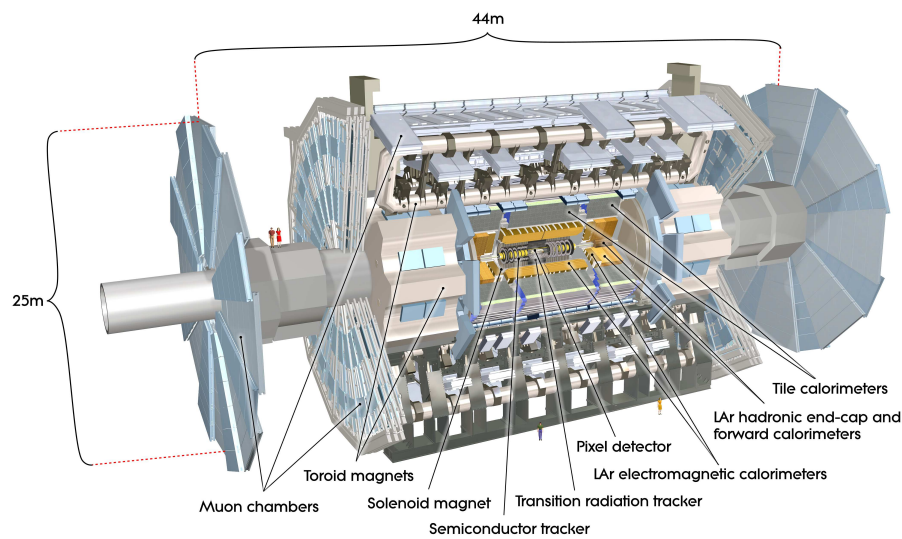


**Figure 02:** Schéma du détecteur ATLAS.

## Correction des formes des cascades électromagnétiques

Le calorimètre électromagnétique, electromagnetic calorimeter (EMC), est conçu pour reconstruire l'énergie des électrons et des photons qui atteignent le calorimètre. Les informations de l'EMC sont également utilisées pour l'identification des particules et le rejet du bruit de fond.

L'EMC se compose de trois couches et d'un pré-échantillonneur (voir Fig. 03). La deuxième couche est la plus épaisse, elle absorbe l'essentiel de l'énergie des électrons et des photons. La deuxième couche a une granularité fine dans les deux dimensions $\eta$ et $\phi$, et fournit des informations sur le développement transverse de la cascade électromagnétique. Ces informations sont utilisées pour calculer un certain nombre d'observables appelées formes de la cascade, puis utilisées comme paramètres d'entrée pour un algorithme MVA qui prend une décision sur l'identification des particules. Il existe trois formes de cascades qui reflètent le développement de la cascade dans la troisième couche du calorimètre:
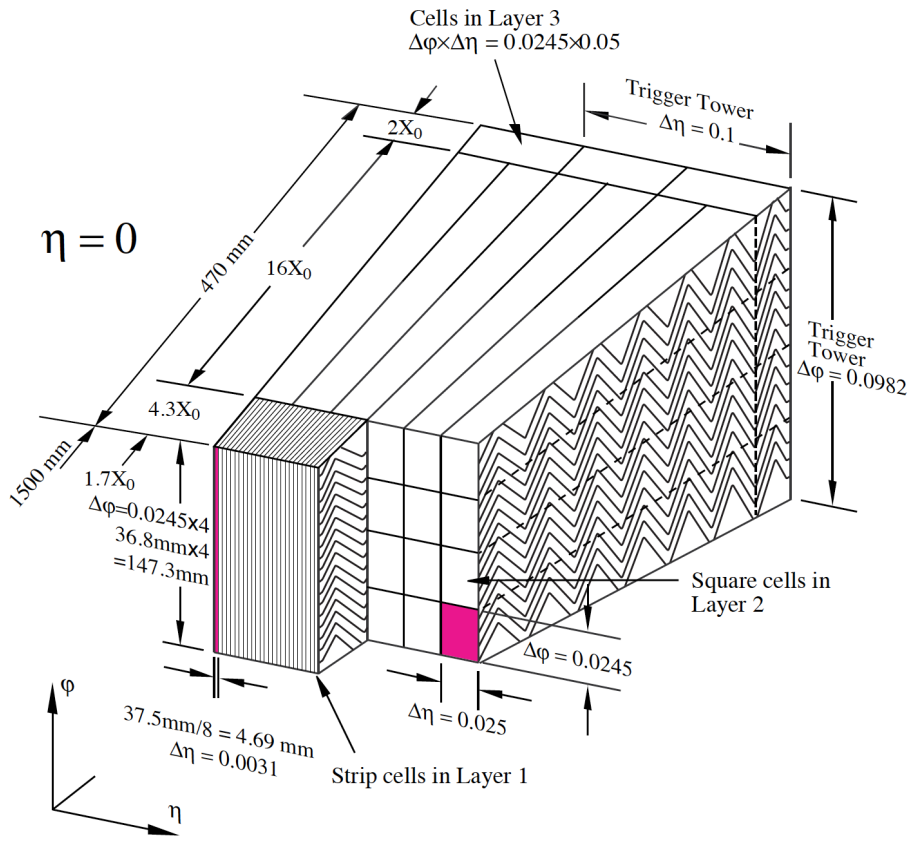
**Figure 03:** Couches du calorimètre électromagnétique d'ATLAS.

- La largeur latérale de la cascade $W_{eta2} = \sqrt{\sum(E_i\eta_i^2) - (\sum(E_i\eta_i)/\sum(E_i))^2}$ calculée dans une fenêtre de 3x5 cellules.

- $R_\phi$ - rapport de l'énergie dans les cellules 3x3 sur l'énergie dans les cellules 3x7 centrées autour de la cellule la plus énergétique.

- $R_\eta$ - rapport de l'énergie dans les cellules 3x7 sur l'énergie dans les cellules 7x7 centrées autour de la cellule la plus énergétique.

Pour des raisons inexpliquées, la modélisation des formes de cascade dans la simulation Monte-Carlo est imparfaite et il y a un écart substantiel avec les données (voir Fig. 04). Ces écarts doivent être corrigés avec des facteurs d'étalonnage, à qui sont associées des incertitudes dépendant du $p_T$. Cette thèse présente une méthode basée sur les données pour corriger ces écarts en redistribuant l'énergie entre les cellules des amas calorimétriques.

Les formes de cascades corrigées dans la simulation sont en bon accord avec les données et se traduisent par un accord significativement meilleur dans les efficacités d'identification (voir Fig. 05). L'effet de la correction est le plus important dans la région des bouchons où il atteint 3%. La méthode
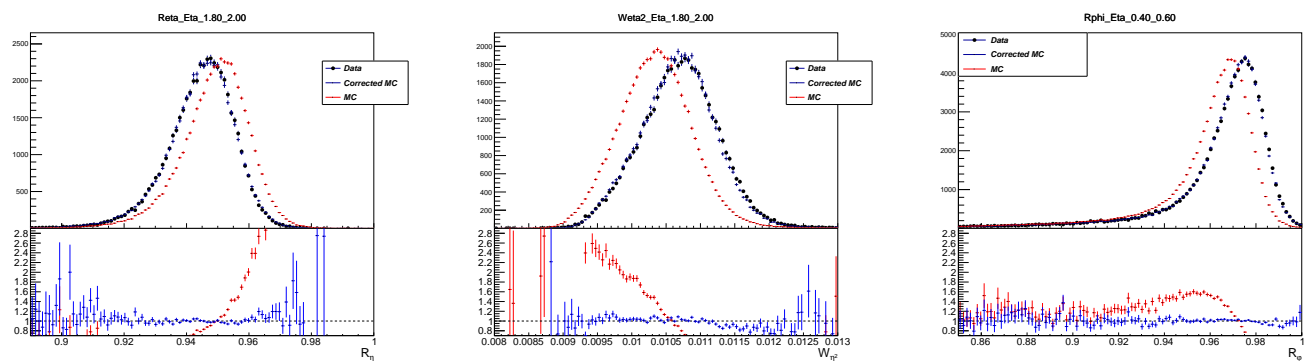
**Figure 04:** **Left**: $R_\eta$ dans $|\eta| = (1.8, 2.0)$. **Central**: $W_{\eta^2}$ dans $|\eta| = (1.8, 2.0)$. **Right**: $R_\phi$ dans $|\eta| = (0.4, 0.6)$.

proposée a été intégrée dans le logiciel de reconstruction officiel de l'expérience ATLAS et sera utilisée par défaut pour les analyses du Run3.
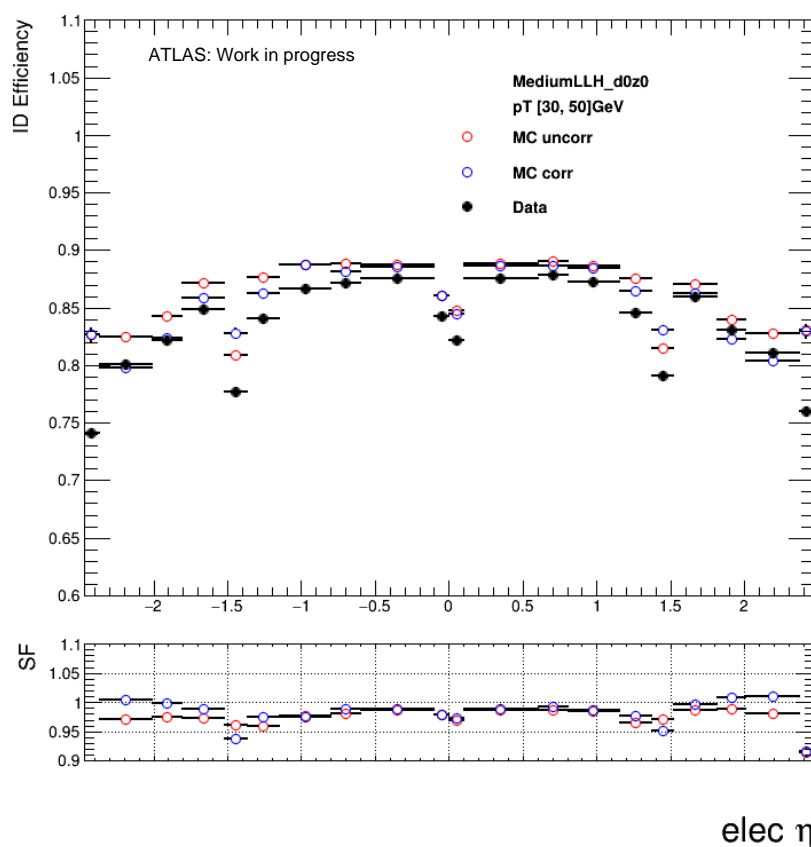


**Figure 05:** Efficacité d'identification des électrons en fonction de leur pseudo-rapidité.

## Spectre en impulsion transverse du boson W

La mesure du spectre en impulsion transverse ($p_T$) du boson W est un objectif difficile mais important. À l'ordre dominant (Fig. 06a), le boson $p_T$ du W est principalement dû aux mouvements intrinsèques du parton et ne dépasse pas 1 GeV. La distribution en impulsion transverse observée, qui atteint des centaines de GeV, est due aux radiations dans l'état initial, qui est un processus survenant à l'ordre suivant, NLO (voir Fig. 06b). Cela permet de tester les prédictions du Modèle Standard, en comparant la simulation aux données.
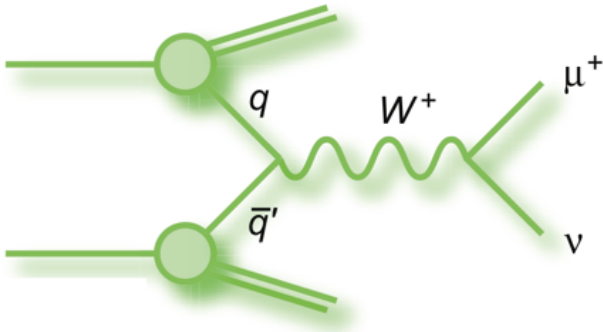


**(a)** Diagramme de Feynman de la production de boson W à l'ordre dominant.
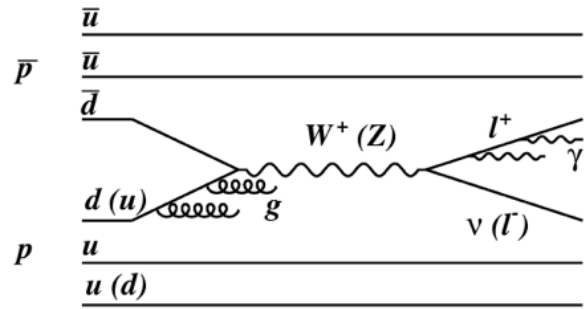


**(b)** Production de boson W à l'ordre NLO.

La deuxième motivation vient du fait que la modélisation précise de la distribution en impulsion transverse du W est nécessaire à la mesure précise de la masse du boson W. Étant l'un des paramètres d'entrée du Modèle Standard, la précision de cette mesure a donc des implications sur toutes les prédictions de la théorie. La modélisation théorique du $p_T$ du W est la deuxième incertitude dominante dans la mesure de la masse du boson W réalisée par la collaboration ATLAS.

La mesure de l'impulsion transverse du boson W est difficile. En raison de la présence d'un neutrino dans l'état final, il est impossible de reconstruire le $p_T$ du boson W directement à partir de ses produits de désintégration. C'est pourquoi il est reconstruit en utilisant une observable appelée recul hadronique. Le recul hadronique est supposé refléter la somme des impulsions transverses des radiations dans l'état initial et être égal en magnitude et antiparallèle au $p_T$ du boson W (voir Fig. 07a). Il est défini pour chaque événement comme la somme vectorielle de tous les *particle flow objects* (PFOs), à l'exclusion des produits de désintégration du boson vecteur. La résolution du recul hadronique dépend directement de l'empilement (Fig. 07b), ce qui a motivé l'utilisation de données à faible taux d'empilement, collectées par l'expérience ATLAS en 2017 et en 2018.

Bien que le recul hadronique soit très corrélé avec le $p_T$ du boson W, la forme de sa distribution peut être très différente en raison des effets de détecteur. Afin de restaurer la distribution sous-jacente, une procédure appelée déconvolution est utilisée. Après avoir déconvolué le recul hadronique, il devient possible de comparer la distribution obtenue avec les prédictions de la simulation.
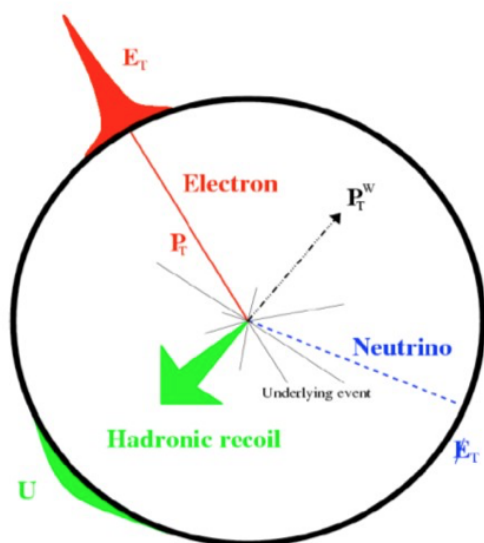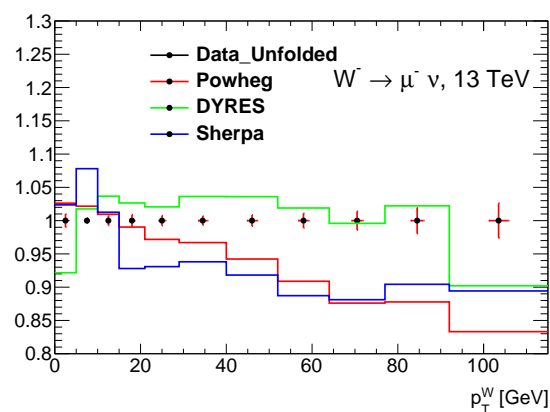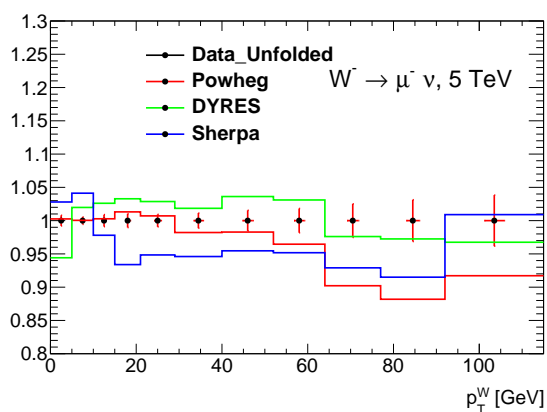
(a) Schéma du recul hadronique.
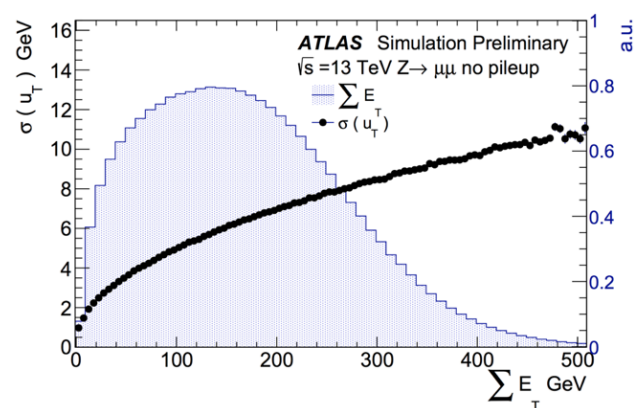


(b) Dépendance en empilement du recul hadronique.
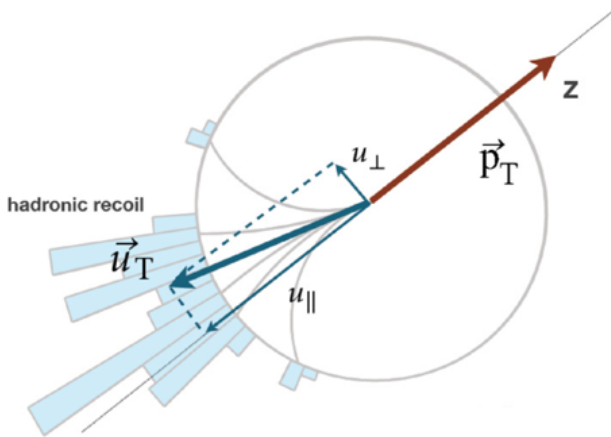
**Figure 07:** Le recul hadronique.





**Figure 08:** Spectre déconvolué comparé aux modèles théoriques pour le canal $W^- \to \mu^- \nu$ à 5 TeV (gauche) et à 13 TeV (droite).

Il apparait que le spectre mesuré est en relativement bon accord avec les modèles à 5 TeV, tandis qu'à 13 TeV aucun des générateurs MC ne fournit une prédiction compatible.

## Reconstruction du recul hadronique à l'aide de réseaux neuronaux profonds

La résolution de la mesure du recul hadronique est cruciale pour la détermination de la distribution du spectre en $p_T$ du W. Les méthodes modernes d'analyse de données permettent d'améliorer la résolution événement par événement. Afin d'améliorer la résolution, un réseau neuronal profond a été utilisé avec les paramètres suivants:

- échantillon d'entrainement : 12 734 109 événements, échantillon de validation: 3 034 130 événements.

- Système : Keras / Tensorflow.

- Fonction objectif: erreur quadratique moyenne.

- Optimiseur: Adam, étape: 0,001.

- Taille du lot: 3900 événements.

- Trois couches denses cachées de 256 nuds chacune.

- couche de normalisation par lots après chaque couche cachée.



(a) Décomposition du recul hadronique.
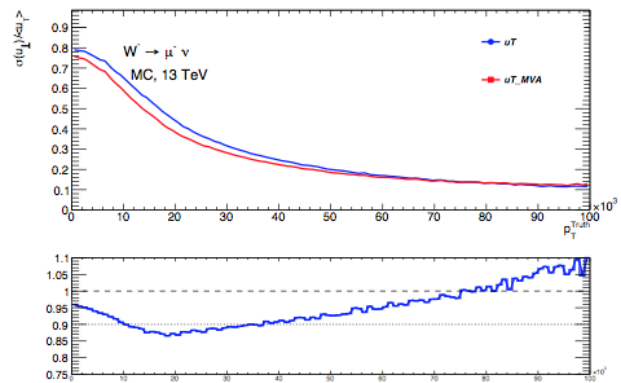
(b) Amélioration de la résolution du recul hadronique par le DNN.

Cette configuration a été testée pour être la plus efficace possible dans des conditions données. La normalisation par lots a amélioré la stabilité de l'ajustement et réduit le temps d'entraînement d'un facteur cinq environ. La liste suivante des paramètres d'entrée a été utilisée:

- Le vecteur du recul hadronique $\vec{u_T}$.

- $\vec{u_T^{charg\omega}}$ - la somme vectorielle des $\vec{p_T}$ des PFOs chargés .

- $\vec{u_T^{neutre}}$ - la somme vectorielle des $\vec{p_T}$ des PFOs neutres .

- $\Sigma E_T$, $\Sigma E_T^{charg\omega}$, $\Sigma E_T^{neutre}$ - les sommes scalaires des énergies transverses.

- Les $p_T$s des deux jets de plus hauts $p_T$s.

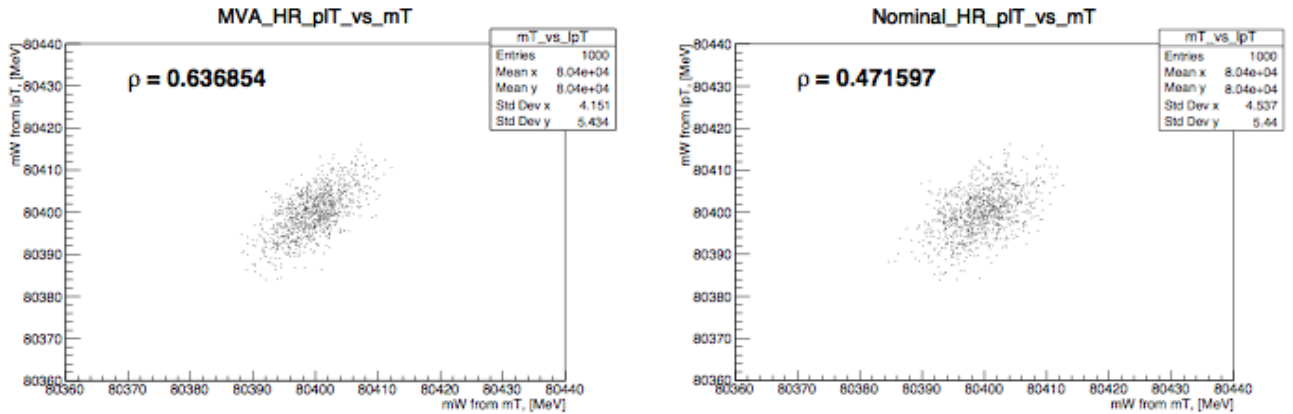- Le nombre de vertex primaires

- Le nombre de PFOs chargés et neutres dans l'événement.

- Les impulsions transverses des cinq PFO neutres et chargés de plus hauts $p_T$s.

L'algorithme somme jusqu'à un total de 38 paramètres d'entrée pour la régression des deux composantes du $\vec{p_T}$ du boson W.



**(a)** Pseudo-expériences pour le recul hadronique nominal.

**(b)** Pseudo-expériences pour le recul hadronique reconstruit avec le DNN.

En raison des effets de détecteur, le vecteur du recul hadronique est différent de celui du $p_T$ du boson W en magnitude, et de plus ne lui est pas exactement antiparallèle (voir Fig. 09a). La résolution du recul hadronique peut être représentée comme l'étalement de la composante perpendiculaire du recul hadronique $\sigma(u_\perp)$. La Fig. 09b montre environ 10% d'amélioration provenant de l'utilisation du réseau neuronal profond pour la reconstruction du recul hadronique.

Une résolution de recul hadronique améliorée se traduit également par une meilleure sensibilité des observables à la masse du boson W. Ceci est testé en utilisant 1000 pseudo-expériences et résulte en une plus petite largeur de la masse transverse lorsque le réseau neuronal est utilisé (voir les figures 010a et 010b).

## Conclusions

Dans cette thèse, trois contributions principales liées à la mesure des propriétés du boson W et aux performances globales du détecteur ATLAS sont présentées.

La correction des formes des cascades dans le calorimètre électromagnétique a permis de corriger l'efficacité d'identification dans la simulation, gagnant 1 à 3% dans la région des bouchons. L'algorithme développé a été adopté comme défaut pour les prochaines analyses de l'expérience ATLAS durant le Run3.

La mesure de la distribution en impulsion transverse du boson W a permis d'effectuer une comparaison avec les modèles MC existants. La comparaison a révélé un bon accord à 5 TeV, mais un écart significatif à 13 TeV. La précision obtenue permettrait de réduire considérablement l'incertitude de modélisation théorique pour la mesure de la masse du boson W.

L'utilisation d'algorithmes d'apprentissage profond a permis d'améliorer la reconstruction du recul hadronique, permettant une amélioration d'environ 10% de la résolution dans la région la plus importante de faible impulsion transverse. Elle a également entraîné une sensibilité accrue des observables à la masse du boson W, qui a été testée en utilisant des pseudo-expériences.

# Contents

# Introduction

The short introduction opens with a historical retrospective of the evolution of theoretical concepts and experimental discoveries, that have shaped the contemporary particle physics. The following section touches a number of opened questions and challenges faced by particle physics. The final section of the introduction outlines the composition of the thesis, providing a brief description of the contents for each of the ten chapters.

## Historical retrospective

The reductionistic idea that the countless varieties of matter types that surround us could be in fact simplified to a combination of much fewer substances has existed at least since the time of Ancient Greece. The thought that you can construct everything you see around out of one or few (e.g. fire, earth, water and air) indivisible elements ($\alpha\tau o\mu o\sigma$ in Greek) is simple, logical and therefore conceptually attractive. Knowing all about these elements could potentially grant us profound understanding of nature. But it wasn't before the XIX$^{th}$ century that this idea has become something more than a philosophical concept and obtained solid scientific evidence.

The composition of the periodic table of elements that has begun in the early XIX$^{th}$ century and concluded in the 1860s [1] was a tremendous step forward, reducing the number of elements to O(100). A brilliant (yet not completely true) hypothesis that all the chemical elements are composed out of hydrogen atoms was published by William Prout as early as 1815 [2]. The elements of the periodic table resembled the ancient Greek concept so much, that they were christened atoms. But the periodic character of the table and strong correlation of atom position in the table with its chemical properties was insinuating on a certain inner structure of the atoms, a possibility for them to be composed out of even smaller objects. The discovery of isotopes in 1913 [3] left little room for other explanation.

Further evidence in favour of the atomistic view kept coming in the late XIX$^{th}$ and early XX$^{th}$ centuries from theoretical and experimental sides. The molecular kinetic theory has been heavily criticized throughout the XIX$^{th}$ century, but the explanation of the Brownian motion [4] has secured its dominance from there on lying a foundation for what is to become the statistical physics. Of particular importance was the discovery of the first subatomic particle in 1897, which was called the electron [5]. It was shortly followed by the identification of a hydrogen ion, subsequently named a

proton [6][1]. Originating from the Greek word $\pi\rho\omega\tau\sigma\sigma$ (primal), the name was inspired by the term *protyle* introduced by William Prout a century before for the hydrogen atom [2].

Further studies of radioactive materials have allowed to compose a seemingly consistent understanding of what matter is composed of. By the time of neutron discovery in 1932 [8] the list of what was called elementary particles was reasonably short: an electron, a proton, and a neutron. The list could arguably be supplemented with a hypothetical very light chargeless particle called neutrino, postulated in 1930 in order to explain the continuous electron spectrum in $\beta$–decay [9]. It was still left to figure out how these elements interact forming the known atoms, molecules and all the matter around. That required additional efforts on the theoretical side, including resolving the inconsistencies between the two new branches of physics supposed to describe the microworld and the fields, namely the quantum theory and the field theory.

To move forward physicists have made use of another source of elementary particles - the cosmic rays. Cosmic rays contained particles of much higher energies comparing to the radioactive materials. Cosmic ray experiments have led to the discovery of the first known antiparticle - the positron [10], confirming the theoretical predictions by Dirac [11]. Further discoveries of the muon [12], pion [13], kaon [14] and $\Lambda_0$ [15] have shown that the list of elementary particles was still far from being completed. The experimental detection of neutrino in 1956 [16] has confirmed the theoretical prediction made over quarter of a century before.

The second half of the XX$^{th}$ century has pronounced a new era in particle physics with the extensive use of particle accelerators. Accelerators have become the main experimental tool in the discovery of new particles and investigation of their properties. Comparing to the cosmic rays, accelerators could offer a more stable flow of high energy particles and better control over the experimental conditions. Thanks to these new tools by the end of 1960s the number of newly discovered particles has exceeded one hundred and kept growing, apparently taking away the reductionistic dream of having a reasonably small number of elementary particles.

On the other hand, the properties of the newly discovered particles (sometimes called "the particle zoo") had provided enough experimental data for theorists to make further assumptions. The particles, if grouped by their properties, have formed patterns - a situation resembling the old story with the atoms of the periodic table. This observation has allowed to assume the existence of even smaller fundamental particles with a fractional charge that would make up all the visible hadrons. These particles were eventually called quarks [17], [18]. By the late 1960s hypothesizing the existence of only three quarks was enough to explain all the visible particles and successfully predict new ones [19]. Since then three more quarks were discovered and as of now all the experimental evidence suggests that the quarks are truly fundamental particles being indivisible in the Ancient Greek sense.

At the same time serious theoretical efforts were taken in order to describe the interactions between

---

[1]The anode rays were discovered in 1886, though they included diverse ions with different charge-to-mass ratios and could not be identified as a single particle [7].

fundamental particles, taking into account the known fundamental forces. In the mid-1970s a theory called The Standard Model was finalized. It included three out of four known fundamental forces (excluding the gravity) and predicted a number of particles which were not discovered by that time. All the key predictions of the theory were successfully confirmed by further experiments, making it a dominant theory in particle physics. The theory was able to describe all the surrounding matter with only 12 fundamental fermions (and their antiparticles) and 5 bosons. The SM is described in more detail in Chapter 1.

Theoretical efforts aimed to further simplify the list of fundamental particles are ongoing, but up to the time of this thesis writing none of them were confirmed experimentally.

## Current challenges

The establishment of the Standard Model was a colossal step forward in the understanding of microworld physics. Nevertheless despite its great success and very good agreement with the vast majority of the experimental data there is a number of inconsistencies and lacunae in the theory which do not allow to think of the SM as of the final theory. Here are most notable of these problematic questions:

1. A number of neutrino experiments have established that the neutrinos have a tiny though non-zero mass. The minimal Standard Model assumes neutrinos to be massless and does not allow to provide mass to the neutrinos.

2. Astrophysical and cosmological evidences confirm the existence of dark matter which does not correspond to any of the particles of the SM.

3. Cosmological observations show a substantial disproportion between observed matter and anti-matter in favour of the former. The SM does not provide an explanation how such an imbalance could have been formed. This fact is probably connected to the problem of CP-violation, which also lacks fundamental explanation from the SM.

4. The discovery of the gravitational waves in 2016 have confirmed the existence of the graviton - the mediator of the gravitational force. The gravitational force is not represented in any way in the SM.

5. No explanation is provided to the vastly different magnitude of the fundamental forces, i.e. why the gravity is $10^{24}$ times weaker than the weak force.

6. Three generations of fermions are postulated with no explanation for number of generations.

In order to attack these and other problems numerous efforts have been taken to either modify the SM or to replace it with a more fundamental theory, but so far none of these Beyond Standard

Model (BSM) theories were ever confirmed experimentally. The SM is still a source of most accurate predictions for any physical process that involves elementary particle interactions. Description of the BSM theories goes beyond the scope of current thesis.

The SM depends on the list of 18 free parameters if the neutrinos are assumed massless (25 parameters if neutrinos are massive). These parameters can not be calculated intrinsically, but they can be related to experimentally measurable quantities. The more precisely we know the values of these parameters - the better is the accuracy of the SM prediction. Precise knowledge of the SM input parameters can also give hints on where to look for a more fundamental theory.

The LHC experiments have already contributed greatly by discovering the last missing piece of the SM, the Higgs boson. This has ended the era of SM particle discoveries but at the same time started the era of LHC precision measurements. The LHC experiments were capable to measure some parameters of the SM for the first time (like the mass of the Higgs boson), but also could improve the existing measurements, boosting the predictive power of the SM.

The scope of this thesis includes the measurement of the W boson transverse momentum spectrum. This measurement may serve as a test for the SM predictions for differential cross-sections. It is also an important part of an ongoing effort at the ATLAS experiment to improve the precision of the W boson mass measurement, which is also among the SM predictions. The mass of the W boson was first measured at Large Electron-Positron (LEP) after its discovery in 1983. The precision of the measurement was further improved by the experiments at Tevatron collider. The only LHC result performed so far was published by ATLAS collaboration in 2018 [20].

Hadron colliders are a challenging environment for the W boson-related measurements, the precision is highly impacted by a number of factors one of them being pile-up. The current analysis is based on the data collected during two special LHC runs with low pile-up, taken in 2017 and 2018.

## Thesis composition outline

The first chapter contains an introduction to the Standard Model, describes its constituents and parameters. It is followed by a more detailed consideration of the electroweak sector, and in particular of the W boson and its properties. In the last section of Chapter 2 the importance and challenges of the W boson transverse momentum measurement are reviewed.

Information on the LHC accelerator and its operations in Chapter 3 is followed by the ATLAS detector description in Chapter 4. Chapter 5 is dedicated to the description of the electromagnetic shower shapes in the ATLAS calorimeter. After the discussion on role of the shower shapes in particle identification I present the method to correct the observed discrepancy between the data and the simulation.

Chapter 6 provides an overview of the object reconstruction techniques used in the ATLAS physics analyses. The following four chapters present the analysis on the precision measurement of the W

boson transverse momentum spectrum . Chapter 7 comprises the calibrations and corrections, used for the reconstructed objects used in the analysis. Chapter 8 contains information on the Monte-Carlo samples, used to model the signal and the backgrounds, event selection details and the Z vertex reweighting procedure. Measurement results are presented and briefly discussed in Chapter 9.

The concluding Chapter 10 is dedicated to the application of deep learning methods for the reconstruction of the hadronic recoil. The obtained results are discussed in the Conclusions. A collection of uncertainty breakdown plots for the main observables in the transverse momentum measurement for all the available channels is presented in Appendix A. The plots containing the comparison between different methods for the hadronic recoil reconstruction are collected in Appendix B.

# 1 The Standard Model

The structure and constituents of the Standard Model (SM) are discussed in the current chapter. The SM of particle physics is a quantum field theory that postulates the existence of three generations of quarks and leptons interacting through three fundamental forces: electromagnetic, weak and strong. From the mathematical point of view the SM is a gauge quantum field theory that has internal symmetries of the unitary product group $SU(3) \times SU(2)_L \times U(1)$. The fourth fundamental force, namely the gravity, is not included in the SM. Nevertheless, since the magnitude of the gravity interaction is negligible on the microscopic scale, it has little to no effect on the precision of the SM predictions. The model has 18 free input parameters[1] - the physical constants that can not be predicted from within the theory and must be measured experimentally. Evidently, the SM predictions are based on these parameters, so the better we know them - the better we can predict how nature behaves on the micro level. The free parameters of the SM are briefly described in section 1.1

A comprehensive description of the quantum field theory formalism goes beyond the scope of the current dissertation and can be found in the corresponding textbooks [21, 22, 23, 24, 25, 26]. In the following section a brief overview of key SM features and constituent parts is provided.

## 1.1   General composition and key parameters

In this section I will describe the fields that enter the SM. Their existence and interactions result in the three fundamental forces that are taken into account by the theory. The quanta of these fields are also called fundamental particles and possess a number of properties like mass, charge (or charges) and spin (see figure 1.1). The fundamental particles are divided into two groups based on their spin: particles with integer spin are called fermions and those with half-integer spin are bosons.

Let's start from the fermion sector. According to the Pauli exclusion principle[27] two fermions can not occupy the same quantum numbers. This in turn, has a consequence that the fermions must occupy a finite volume in space-time and as a result make up matter. Half of the fundamental fermions have colour charge and therefore take part in strong interaction - they are called quarks. The other six

---

[1]There are SM extensions that take into account the non-zero neutrino mass. Then the model gets 7 additional parameters, so their total number reaches 25. Although current thesis only considers the SM where neutrinos are massless.

fermions do not have colour charge and are called leptons (from Greek "$\lambda\epsilon\pi\tau o\sigma$" meaning "little", as they are lighter than the quarks of the same generation). Different types of quarks and leptons are also called flavours, so there are 6 flavours of quarks and 6 flavours of leptons.

For some reason which is yet unknown the twelve elementary fermions make three generations. Particles in the second and third generations have exactly the same charge and spin as the particles of the first generation, but are heavier and also unstable. Normally the particles of higher generations quickly decay down to their lighter kin of the first generation and can only be observed in cosmic rays and particle accelerators. That means all the matter that surrounds us consists of the four fundamental fermions of the first generation[2] (the first column in Fig. 1.1).

The two quarks of the first generation are called up-quark and down-quark (or u-quark and d-quark for short). All the nuclei of the ordinary matter we see around are built out of these two types of quarks. Quarks are capable of interacting through all three SM forces: electromagnetic, weak and strong. Electrons, muons and tau-leptons are sensitive to electromagnetic and weak interaction, while neutrinos can interact (and therefore be detected) only through the weak force. For this reason in particle physics the term "leptons" is sometimes used in a narrow sense referring to electrically charged leptons only. For all quarks and charged leptons the antiparticles were observed as well as the corresponding annihilation phenomena. It is still not clear if neutrinos and antineutrinos of the same flavour are distinct particles.

From our experience we know that matter interacts with matter. But within the SM fermions do not interact with each other immediately. The interaction is mediated by boson-type particles. The SM includes several types of bosons: vector bosons serving as force carriers for electromagnetic, weak and strong interactions, and a scalar Higgs boson whose role will be described in more detail in the corresponding subsection 1.4.1. The Higgs boson, along with the W and Z bosons are massive, while photons and gluons are massless. The masses of the fundamental particles make 12 out of 18 free parameters of the SM[3].

As it was mentioned, bosons interact with fermions through fundamental interactions. The interaction depends on the charge of the interacting particles and on the type of the interaction itself. Each type of interaction has a coupling constant that defines the scale of the interaction. Hence two more parameters to the SM: the strong and electromagnetic coupling constants (the latter is also called the fine structure constant). The weak coupling constant is redundant since it can be obtained from other parameters. The remaining four parameters are coming from the Cabibbo-Kobayashi-Maskawa (CKM) matrix, that contains information on the strength of the flavour-changing weak interaction [28].

An important feature of the Quantum Field Theory (QFT) is that particles also interact with physical vacuum. For instance, a charged particle polarizes the physical vacuum, so the vacuum

---

[2]Strictly speaking we already know that this is not completely true for the neutrinos, as they oscillate between the flavours due to their tiny mass. But in the SM neutrinos are assumed to be massless.

[3]The masses of the W and Z bosons can be replaced by other parameters, e.g. weak mixing angle $\theta_W$ and Higgs potential vacuum expectation value (v. e. v.).

# Standard Model of Elementary Particles

| | three generations of matter (fermions) | | | interactions / force carriers (bosons) | |
|---|---|---|---|---|---|
| | I | II | III | | |
| mass<br>charge<br>spin | ≃2.2 MeV/c²<br>⅔<br>½<br>**u**<br>**up** | ≃1.28 GeV/c²<br>⅔<br>½<br>**c**<br>**charm** | ≃173.1 GeV/c²<br>⅔<br>½<br>**t**<br>**top** | 0<br>0<br>1<br>**g**<br>**gluon** | ≃124.97 GeV/c²<br>0<br>0<br>**H**<br>**higgs** |
| | ≃4.7 MeV/c²<br>−⅓<br>½<br>**d**<br>**down** | ≃96 MeV/c²<br>−⅓<br>½<br>**s**<br>**strange** | ≃4.18 GeV/c²<br>−⅓<br>½<br>**b**<br>**bottom** | 0<br>0<br>1<br>**γ**<br>**photon** | |
| | ≃0.511 MeV/c²<br>−1<br>½<br>**e**<br>**electron** | ≃105.66 MeV/c²<br>−1<br>½<br>**μ**<br>**muon** | ≃1.7768 GeV/c²<br>−1<br>½<br>**τ**<br>**tau** | ≃91.19 GeV/c²<br>0<br>1<br>**Z**<br>**Z boson** | |
| | <1.0 eV/c²<br>0<br>½<br>**νe**<br>**electron neutrino** | <0.17 MeV/c²<br>0<br>½<br>**νμ**<br>**muon neutrino** | <18.2 MeV/c²<br>0<br>½<br>**ντ**<br>**tau neutrino** | ≃80.39 GeV/c²<br>±1<br>1<br>**W**<br>**W boson** | |

QUARKS · LEPTONS · GAUGE BOSONS (VECTOR BOSONS) · SCALAR BOSONS

**Figure 1.1:** The list of particles that enters the SM[29].

changes the charge of the particle [30].This interaction with virtual particles depends on the energy scale and so do the observed quantities like charge, mass etc. The SM is able to predict parameter evolution, so if the value of a certain input parameter $q_0$ is known at the energy $\Lambda_0$ then it is possible to predict its measurable value $q$ at the energy $\Lambda$. This changing of physical parameters is an integral part of the QFT and is called *renormalisation* [22], [31]. In the Figure 1.2 the dependence of the inverted SM coupling constants on the energy is shown.
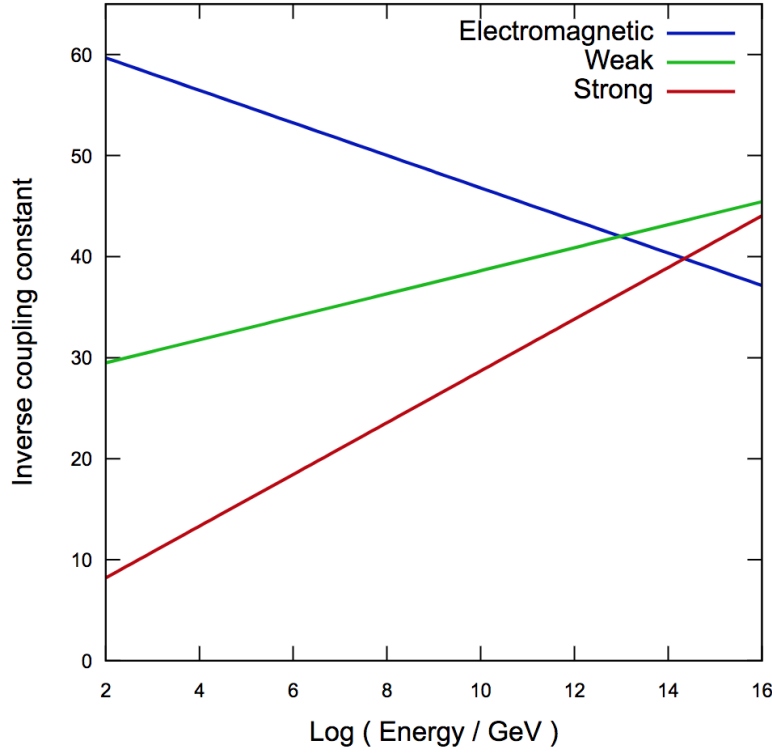


**Figure 1.2:** The running of the inverted SM coupling constants [32].

As we can see from picture 1.2 the strong coupling constant is decreasing with the energy. This phenomenon is called *the asymptotic freedom* [33, 34, 35].

## 1.2 Classical fields and gauge invariance principle

A consistent mathematical description of fields appears to be a more challenging task compared to the description of physical objects that have a definite size and shape even in the classical case. The derivation of Maxwell's equations has been a great success and allowed to obtain the first equations of motion of relativistic fields. It has also subseqently led to the understanding of special relativity [36, 37, 38]. Although for a more general case of fields other than electromagnetic it would be very useful to adopt a more systematic approach like that of Lagrangian or Hamiltonian in classical mechanics.

It has turned out that for the relativistic case the Hamiltonian approach was not quite convenient,

as the dedicated role of time over other degrees of freedom was in discord with relativistic space-time unification. However it was found possible to describe the fields within the Lagrangian approach. In classic mechanics the action of a mechanical system of $i$ mechanical objects is defined as:

$$S = \int L dt = \int \left( \sum_i T_i - U_i \right) dt,$$

where $T_i$ and $U_i$ are the kinetic and potential energies of the $i^{th}$ object. Considering that by definition a field exists in every point of space-time, we need to define the Lagrangian density such that $L = \int \mathcal{L}(\phi, \partial_k \phi, \dot{\phi}) d^3 x$, where $\phi$ is a field and $\partial_k \phi = \nabla \phi$ - the field gradient, $\partial_k = \frac{\partial}{\partial x^k}$, k = 1, 2, 3. Here and further Latin indices run through (1, 2, 3) and are used to denote spatial coordinates, while Greek indices denote space-time coordinates and run though (0, 1, 2, 3). So the action would look like:

$$S = \int L dt = \int \mathcal{L}(\phi, \partial_\mu \phi, \dot{\phi}) d^4 x, \tag{1.1}$$

Now we may use the principle of least action to obtain the equations of motion using the Euler-Lagrange formalism. Let's check it with the example of electromagnetic fields. The Lagrangian density of electromagnetic fields in a vacuum can be written like:

$$S = -\frac{1}{4} \int F^{\mu\nu} F_{\mu\nu} d^4 x. \tag{1.2}$$

The electromagnetic tensor can be defined in terms of electric and magnetic field intensities: $F_{i0} = -F_{0i} = E_i$, $F_{ij} = \epsilon_{ijk} H_k$, where $\epsilon_{ijk}$ is the anti-symmetric Levi-Civita symbol. Alternatively $F_{\mu\nu}$ can be defined in terms of the 4-potential $A_\mu$:

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{1.3}$$

Now we can safely apply the variational principle, and putting $\delta S = 0$ obtain the Maxwell equations in vacuum:

$$\partial_\mu F_{\mu\nu} = 0. \tag{1.4}$$

Noticing the symmetries of the system and using the Noether's theorem[39] we can find the invariants of electromagnetic field. For example, translational symmetry in time and space ensures conservation of energy and momentum. Let's now consider a symmetry of a different kind. The field potential can be shifted by a gradient of an arbitrary function $\alpha = \alpha(x^\mu)$:

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) + \partial_\mu \alpha(x)$$
$$F_{\mu\nu} \rightarrow F'_{\mu\nu} = \partial_\mu (A_\nu(x) + \partial_\nu \alpha(x)) - \partial_\nu (A_\mu(x) + \partial_\mu \alpha(x)) = \partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu}, \tag{1.5}$$

where the commutativity of the derivative operator $\partial_\mu \partial_\nu \alpha(x) = \partial_\nu \partial_\mu \alpha(x)$ was used. Let us now consider the electromagnetic theory in the presence of charges and currents:

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} + j^\mu A_\mu. \tag{1.6}$$

Now we have an interaction of a field potential $A_\mu$ with 4-current $j^\mu = (-\rho, j^i)$. It turns out to be a general property of the field theories: the only form of interaction allowed is between a gauge field and a current. After applying the gradient field transformation and the least action principle we can obtain the corresponding conservation law:

$$\partial_\mu j^\mu = 0. \tag{1.7}$$

So this gradient symmetry [22] or as it is called more often gauge symmetry is connected to the conservation of electric current. If a theory is invariant under gauge transformations then it is called a gauge invariant theory. As we have just seen electrodynamics is the simplest example of such a theory. Taking gauge symmetries into consideration [40] has played a huge role in the development of the SM.

Gauge degree of freedom can be constrained in arbitrary way by applying additional conditions on the gauge function. This is called fixing the gauge and becomes necessary for the quantization. As a result of a non-trivial procedure it can be show that any physical result must be gauge-invariant, i.e. must not depend on the gauge.

## 1.3   Quantum electrodynamics

Quantum Electrodynamics (QED) is a theory of interaction between light and electrically charged particles. Historically it was the first quantum field theory to reach good agreement between quantum mechanics and special relativity. QED vacuum has zero expectation value. Nowadays it is considered to be one of the most precise physical theories ever: theory predictions and experiment results agree up to $O(10^{-11})$. It has also served as a model for the composition of the subsequent parts of the SM, describing other fundamental interactions. Let us consider the free Dirac field based Lagrangian:

$$\mathcal{L} = \bar{\psi}(x)(i\slashed{\partial} - m)\psi(x), \tag{1.8}$$

where $\psi$ and $\bar{\psi}$ are Dirac wave function and its complex conjugate respectively, $\slashed{\partial} \equiv \gamma_\mu \partial^\mu$, $\gamma_\mu$ is one of the four gamma-matrices and $m$ is the mass of the Dirac field. Such a theory, though, would not be physically consistent. This reflects the fact the quantum nature of spin and spinor fields have to be treated as quantum fields. For instance, an attempt to calculate the energy of a Dirac field would lead to a contradiction: the energy would not be positively defined, as some spinors would have negative energies.

This Lagrangian has an internal symmetry to the U(1) transformation: $\psi \to e^{-i\alpha(x)}\psi$, $\bar{\psi} \to e^{i\alpha(x)}\bar{\psi}$. According to Noether's theorem this symmetry implies current conservation: $j^\mu = \bar{\psi}\gamma^\mu\psi$. Now let's get the combined Lagrangian of electromagnetic and Dirac fields, adding the interaction term:

$$\mathcal{L} = \mathcal{L}_{Dirac\,free} + \mathcal{L}_{EM\,free} + \mathcal{L}_{Interaction} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}(x)(i\slashed{\partial} - m)\psi(x) - q\bar{\psi}\gamma^\mu A_\mu\psi, \tag{1.9}$$

where q represents the elementary electric charge. This Lagrangian above is gauge invariant and can be rewritten in a more convenient form:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \bar{\psi}(x)(i\slashed{D} - m)\psi(x), \tag{1.10}$$
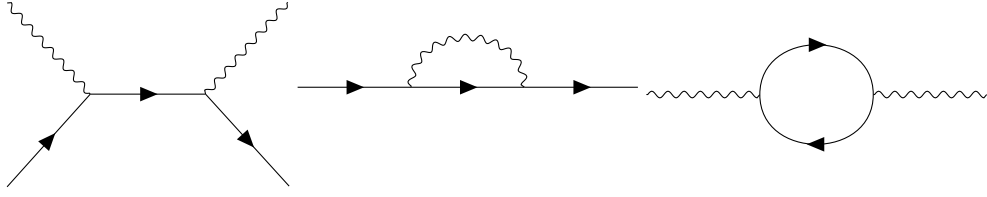
**Figure 1.3:** Examples of QED diagrams: Compton scattering, electron self-energy, photon self-energy.

where $D_\mu = \partial_\mu - iqA_\mu$ is a covariant derivative. If one considers space-time in the presence of a field as curved, then $A_\mu$ would play a role of connectivity. It must be noted that values like $m$ and $q$ meaning electron mass and charge[4] are the SM input parameters mentioned in 1.1.

Further calculations are to be performed by the means of the quantum field theory formalism that treats interaction terms like a perturbation to the free fields, making power series expansion in the coupling constant. In the case of electrodynamics the coupling constant is quite small so good precision is reached soon. Since the photons do not directly interact with other photons, QED allows only one type of vertex - with two electron lines and one photon line.

Although the tree-level processes and diagrams were well understood by 1930th, the loop diagrams were properly explained only by the end of the 1940th making it possible to obtain numerical results of the higher orders of power series expansion and achieve higher precision predictions for QED processes [30, 41, 42, 43, 44, 45, 46, 47]. The examples of QED diagrams are presented in figures below.

It must be noted that although immediate photon-photon interaction is impossible, light-by-light scattering is still possible through loops:
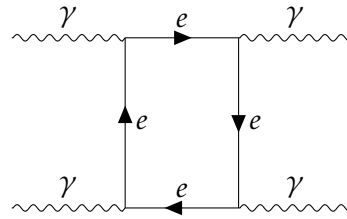


This process was theoretically described in 1936 [48] and experimentally observed 83 years after in heavy ion collisions at the LHC [49].

## 1.4 Electroweak theory and the Higgs mechanism

All the fermions of the standard model are subject to the weak interaction, so its importance for physical processes can not be underestimated. At low energy the weak interaction manifests itself

---

[4]Charge of the electron is related to the electromagnetic coupling constant.

mainly through flavour-changing decays like beta-decay and muon decay. The electroweak theory was created in the end of 1950s [31] [25] [50] thanks to numerous experimental results that allowed to shape its properties. The theory assumed that the electromagnetic and weak fundamental forces are actually manifestation of the same gauge group that has a gauge symmetry $SU(2)_L \times U(1)$ with massive charged and neutral bosons. A few years later the structure of electroweak vacuum was explained along with the mechanism that has allowed the bosons to gain mass [51], [52]. Assuming this the Lagrangian of the electroweak theory must consist of three parts [53]:

- Gauge fields that would mediate the interaction.

- Fermions that interact with gauge fields

- A scalar Higgs field with non-zero vacuum energy that breaks the $SU(2)_L$ symmetry and couples to the fermions.

$$\mathcal{L}_{EW} = \mathcal{L}_{Gauge} + \mathcal{L}_{Higgs} + \mathcal{L}_{Fermions} \tag{1.11}$$

### 1.4.1 Electroweak gauge fields

As it was already pointed out before, knowing the symmetries of a physical system allows one to compose the gauge fields Lagrangian. The part with U(1) symmetry would look like the electromagnetic field from 1.2 having the hypercharge $Y$, a vector potential $B_\mu$ and a gauge coupling $g_1$. The SU(2) field would have 3 vector components $W_\mu^{1,2,3}$, three isospin operators $I_1, I_2, I_3$ and a gauge coupling $g_2$. We can pick the Pauli matrices $\sigma^i$ as the representation of generators of the SU(2) group, then the structure constants are $\epsilon_{abc}$ - Levi-Civita symbol.

$$\begin{aligned} \mathcal{L}_{\mathcal{G}} &= -\tfrac{1}{4} B_{\mu\nu} B^{\mu\nu} - \tfrac{1}{4} W_{\mu\nu}^a W^{\mu\nu,a} B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \\ W_{\mu\nu}^a &= \partial_\mu W_\nu - \partial_\nu W_\mu + g_2 \epsilon_{abc} W_\mu^b W_\nu^c, \end{aligned} \tag{1.12}$$

where the term $g_2 \epsilon_{abc} W_\mu^b W_\nu^c$ appears due to the non-Abelian nature of the SU(2) group (the generators don't commute).

### 1.4.2 Fermion sector

Each fermion generation expressed as left-handed doublet and right-handed singlets is a fundamental representation of the group $SU(2) \times U(1)$:

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, (e_R), \begin{pmatrix} u \\ d \end{pmatrix}_L, (u_R), (d_R), \tag{1.13}$$

$$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, (\mu_R), \begin{pmatrix} s \\ c \end{pmatrix}_L, (s_R), (c_R), \tag{1.14}$$

$$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, (\tau_R), \begin{pmatrix} b \\ t \end{pmatrix}_L, (b_R), (t_R). \tag{1.15}$$

Their quantum states are classified using the following quantum numbers: weak isospin $I_3$, $Q$, weak hypercharge $Y$. Their electric charge can be obtained using the Gell-Mann-Nishijima relation:

$$Q = I_3 + \frac{Y}{2}. \tag{1.16}$$

The fermions are divided by their chirality: only the left-handed particles take part in the charged current of the weak interaction. The left-handed fermion fields of each lepton and quark generation j

$$\psi_j^L = \begin{pmatrix} \psi_{j+}^L \\ \psi_{j-}^L \end{pmatrix} \tag{1.17}$$

make SU(2) doublets, with indices $\sigma = \pm$, while the right-handed fermions can be written as singlets:

$$\psi_j^R = \psi_{j\sigma}^L. \tag{1.18}$$

Like in the the electromagnetic case we can define the covariant derivative that would ensure the gauge invariance of the Lagrangian:

$$D_\mu = \partial_\mu - i g_2 I_a W_\mu^a + i g_1 \frac{Y}{2} B_\mu, \tag{1.19}$$

with $I_a \equiv \frac{\sigma_a}{2}$, then fermion Lagrangian takes the following form:

$$\mathcal{L}_{Fermions} = \sum_f \bar\psi_j^L i \gamma^\mu D_\mu \psi_j^L + \sum_{f,\sigma} \bar\psi_{f,\sigma}^R i \gamma^\mu D_\mu \psi_{f,\sigma}^R. \tag{1.20}$$

### 1.4.3 Higgs field breaking the symmetry

The Higgs field is represented by single complex a scalar doublet field $\Phi(x)$, that has 4 independent components. It spontaneously breaks the $SU(2) \times U(1)$ gauge symmetry, leaving the $U(1)_{EM}$ symmetry intact. The Higgs field doublet has the hypercharge $Y = 1$:

$$\Phi(x) = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix}. \tag{1.21}$$

The Higgs field Lagrangian with non-zero vacuum expectation value is:

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^+ (D_\mu \Phi) - V(\Phi) + \mathcal{L}_{Yukawa}. \tag{1.22}$$

The gauge invariance of the Higgs Lagrangian is ensured in the traditional way by using the covariant derivative:

$$D_\mu = \partial_\mu - i g_2 I_a W_\mu^a + i \frac{g_1}{2} B_\mu. \tag{1.23}$$

The Higgs potential contains the mass term and quartic self-interaction:

$$V(\Phi) = -\mu^2 \Phi^+ \Phi + \frac{\lambda}{4} \partial_\mu (\Phi^+ \Phi)^2, \tag{1.24}$$

where $\lambda$ stands for the quartic Higgs self-coupling constant and $\mu$ is the mass of the $\Phi$ field. The vacuum expectation value $<\Phi>$ does not vanish:

$$<\Phi(x)> = \frac{1}{\sqrt{(2)}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad v = \frac{2\mu}{\sqrt{(\lambda)}}. \tag{1.25}$$

Applying the unitarity gauge [54] we can constrain three out of four degrees of freedom of the Higgs field and rewrite the Higgs doublet in the following way:

$$\Phi(x) = \frac{1}{2} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{1.26}$$

which leaves us with a physical real neutral scalar field $H(x)$ with

$$M_H = \sqrt{2}\mu. \tag{1.27}$$

This real field would couple to itself forming triple and quartic self-coupling vertices, to the gauge fields through the covariant derivatives and to the charged fermions, giving them mass. The Yukawa term in Lagrangian the unitary gauge is:

$$\mathcal{L}_{Yukawa} = -\sum_f m_f \bar{\psi}_f \psi_f - \sum_f \frac{m_f}{v} \bar{\psi}_f \psi_f H, \tag{1.28}$$

where

$$m_f = g_f \frac{v}{\sqrt{2}} = \sqrt{2} \frac{g_f}{g_2} M_W. \tag{1.29}$$

The Higgs coupling constants to the corresponding fermion flavour are denoted as $g_f$.

### 1.4.4 Physical interpretation of gauge fields and parameters

The Higgs coupling to the gauge fields results in the following terms in the Lagrangian:

$$\frac{1}{2} \frac{g_2^2}{2} v^2 (W_1^2 + W_2^2) + \frac{v^2}{4} (W_\mu^3, B_\mu) \begin{pmatrix} g_2^2 & g_1 g_2 \\ g_1 g_2 & g_1^2 \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}. \tag{1.30}$$

In order to get the physical meaning of this expression let us make a transition to the basis of physical fields:

$$\begin{aligned} W_\mu^\pm &= \frac{1}{\sqrt{2}} (W_\mu^\mp \mp i W_\mu^\mp) \\ \begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} &= \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \end{aligned} \tag{1.31}$$

where $\theta_W$ is called the weak mixing angle or the Weinberg angle. In the new basis expression 1.30 has transparent physical sense:

$$M_W^2 W_\mu^+ W^{-\mu} + \frac{1}{2}(A_\mu, Z_\mu)\begin{pmatrix} 0 & 0 \\ 0 & M_Z^2 \end{pmatrix}\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix}, \tag{1.32}$$

with

$$\begin{aligned} M_W &= \tfrac{1}{2}g_2 v \\ M_Z &= \tfrac{1}{2}\sqrt{g_1^2 + g_2^2}\, v. \end{aligned} \tag{1.33}$$

The mixing angle $\theta_W$ also has a very clear physical meaning:

$$\cos\theta_W = \frac{g_2}{g_1^2 + g_2^2} = \frac{M_W}{M_Z}. \tag{1.34}$$

With $A_\mu$ having a sense of electromagnetic potential its coupling to the electron must have a physical meaning of the electric charge $e = \sqrt{4\pi\alpha}$ we can express $e$ in terms of gauge couplings:

$$e = \frac{g_1 g_2}{g_1^2 + g_2^2}, \quad g_2 = \frac{e}{\sin\theta_W}, \; g_1 = \frac{e}{\cos\theta_W}. \tag{1.35}$$

Thus the demonstrated Weinberg rotation (see Fig. 1.4) replaces the original parameters $g_1$, $g_2$, $\lambda$, $\mu^2$, $g_f$ by another set of measurable values $e$, $M_W$, $M_Z$, $M_H$, $m_f$ which are the input parameters of the SM.

## 1.5 Chromodynamics

Quantum Chromodynamics (QCD) is a non-Abelian gauge theory that describes strong interaction. QCD is symmetric under unbroken SU(3) colour symmetry, so the interaction scheme is built in the same way as electromagnetic and electroweak theories. To preserve the gauge invariance the gauge field of gluons is introduced with 8 components, since SU(N) group has $N^2 - 1$ independent elements. The gluons are massless vector bosons like the photons, although because of the non-Abelian nature of the gauge group they couple not only to the fermions but also to the other gluons. The gauge invariant QCD Lagrangian with kinetic term containing covariant derivative would look like:

$$\begin{aligned} \mathcal{L}_{QCD} &= -\tfrac{1}{4}F_{\mu\nu}^a F_a^{\mu\nu} + \bar{\psi}_a(i(\gamma^\mu D_\mu)^{ab} - m\delta^{ab})\psi_b, \\ F_{\mu\nu}^a &= \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g_s f^{abc} A_\mu^b A_\nu^c, \\ D_\mu &= \partial_\mu + i g_s A_\mu^a t_a. \end{aligned} \tag{1.36}$$

with $\psi$ being the quark field, m is the mass of the quark, a,b = 1, 2, ..., 8 are the colour indices, $g_s$ is the strong coupling constant, $f^{abc}$ are the structure constants of the SU(3) group and $t_a$ are the generators of the SU(3) group.

As it was already mentioned in 1.3 quantitative calculations in QFT treat particle interaction as a perturbation to the free field theory. The coupling constant is considered to be a small parameter so every next power of the coupling constant is much smaller than the previous one. Due to asymptotic
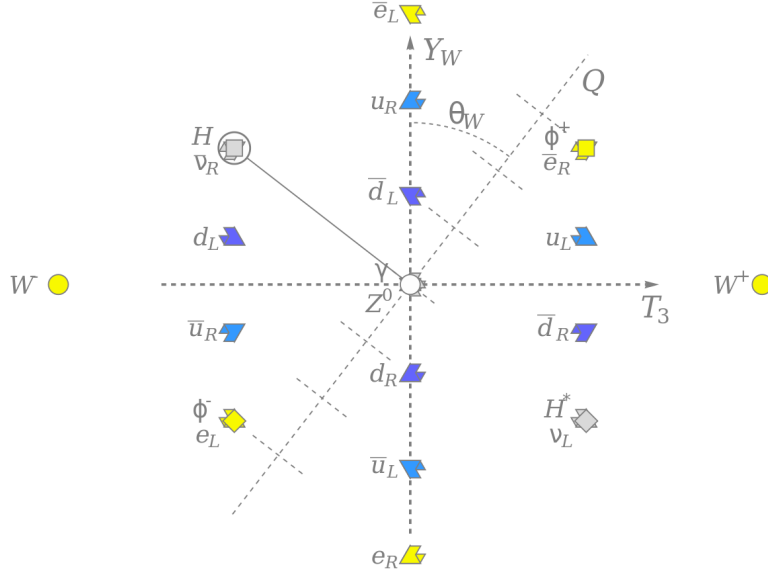
**Figure 1.4:** Electroweak sector and the Weinberg rotation [32].

freedom the constant $\alpha_s$ becomes small at higher energies and allows perturbative calculations. But at a certain energy scale called $\Lambda_{QCD} \approx 200$ MeV, QCD becomes non-perturbative. It means we may no longer assume that interaction is a small perturbation of the free fields. This phenomenon causes the *colour confinement*.

Because of colour confinement we can only observe colourless objects like baryons and mesons, but not quarks and gluons. If a high-energetic parton gets torn out of a hadron then it creates an avalanche-like process creating quark-antiquark pairs until it fully hadronizes (see Fig. 1.5) neutralizing its colour. Such an avalanche is called a hadronic jet.

Currently there is no viable physical theory that would describe QCD vacuum and low-energy behaviour of quarks and gluons. This also means that although nuclear forces are evidently residuals of the QCD interaction of partons within the baryons, there is no continuity between QCD and nuclear physics. Confinement and low-energy QCD remain an unsolved problem of modern physics.
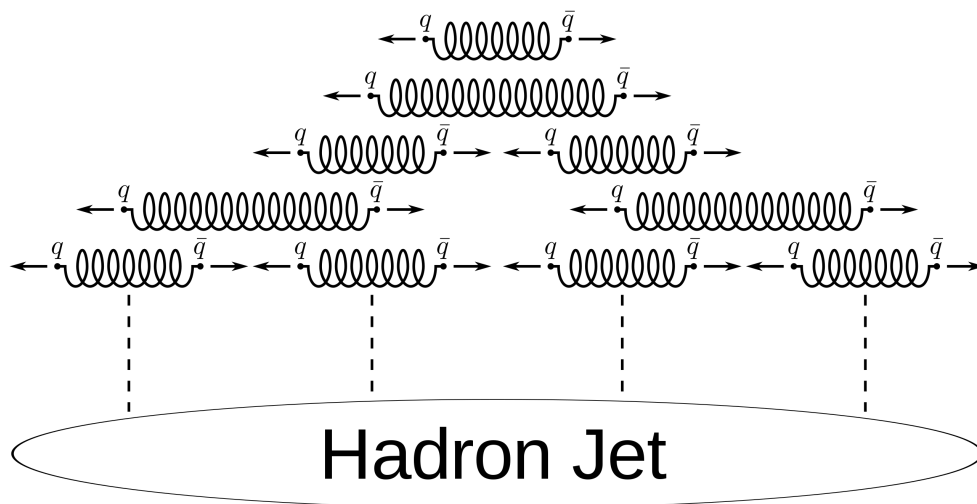
**Figure 1.5:** The formation of a hadron jet [55].

# The W boson

This chapter is dedicated to the electroweak sector of the SM and in particular to the W boson. The aspects of the theoretical modelling of the W boson production and kinematic distributions are discussed. The last section mentions previous measurements of the W boson transverse spectrum and speculates about the target precision for its direct measurement.

## 2.1 The motivation for the W mass measurement

Being one of the cornerstones of the SM, the W boson is tightly connected to the other parameters of the theory. In the leading order of the perturbation theory the W mass depends only on the electroweak parameters [56]:

$$M_W = \sqrt{\frac{\pi \alpha_{EM}}{\sqrt{2} G_F} \frac{1}{\sin \theta_W}}, \tag{2.1}$$

where $G_F$ stands for the Fermi constant, $\alpha_{EM} \approx \frac{1}{137}$ is the electromagnetic coupling constant and $\sin \theta_W$ is the Weinberg angle (see 1.4.3). The factor $\sqrt{\frac{\pi \alpha}{\sqrt{2} G_F}} \approx 40$ GeV sets the lower bound for the possible W mass. Higher order corrections enter the equation in the following way:

$$M_W = \sqrt{\frac{\pi \alpha}{\sqrt{2} G_F} \frac{1}{\sin \theta_W} (1 + \Delta r)}, \tag{2.2}$$

where $\Delta r$ contains the sum of all possible radiative corrections and depends also on other parameters of the SM, first of all on top quark and Higgs boson masses. The correction term is also sensitive to possible BSM effects. As it was mentioned in Chapter 1 the mass of the W boson is one of the input parameters of the SM, so the predictions of the theory directly depend on how precisely we know the
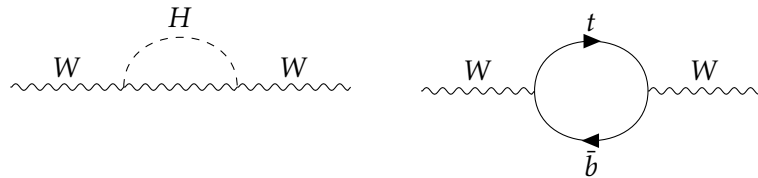


**Figure 2.1:** Next-to-leading order diagrams for W boson propagator containing contributions from heavy quarks and the Higgs boson.

value of the boson mass. On the other hand, we can theoretically constrain the value of the W boson mass assuming the already known values of the other SM parameters. Figure 2.2 demonstrates that the uncertainty on the theoretical estimate for the W boson mass is about two times lower than that of the best available experimental measurement. This motivates the effort for a more precise experimental measurement in order to test the consistency of the SM. Should the improved measurement reveal the inconsistency of the Standard Model - it would also allow to reveal viable BSM theories.



(a) W mass constraint from the global electroweak fit.



(b) Available W mass measurements.
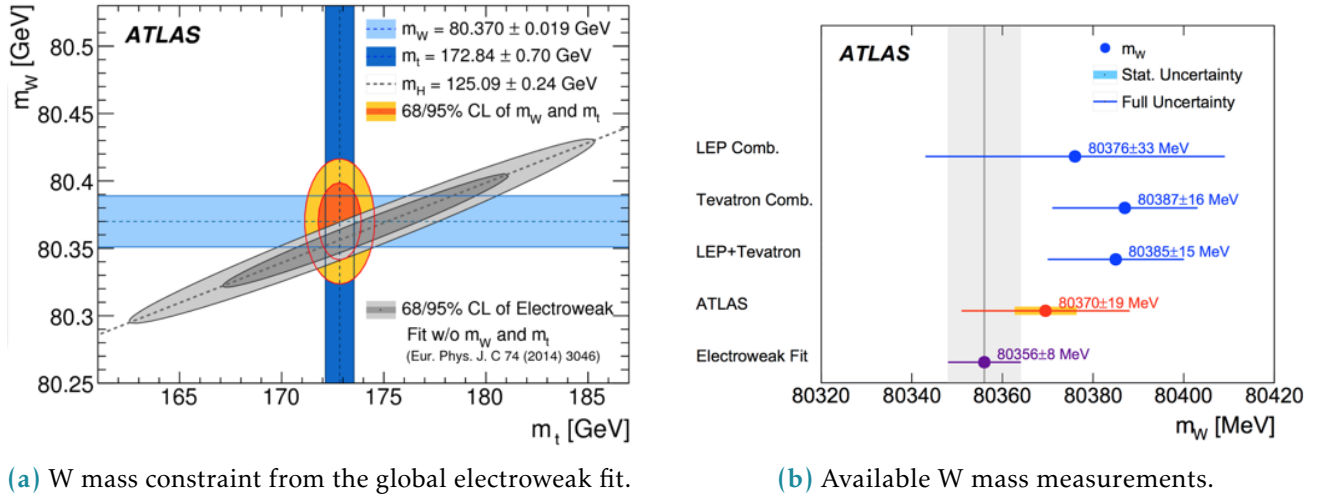
Figure 2.2: W mass measurements and predictions [20].

## 2.2   Massive boson production at hadron colliders

Hadron colliders provide a fruitful environment for the production and study of massive electroweak bosons - all of them were discovered at hadron colliders. Hadron colliders allow to achieve much higher centre-of-mass collision energy and luminosity comparing to their lepton counterparts. At the same time precision measurements at hadron colliders demand much deeper theoretical understanding of different aspects of the SM.
The main theoretical complication of proton-proton colliders lies in the fact that contrary to leptons, protons are complex objects. This raises the following problems:

- A proton-proton collision is in the general case a many-body problem. The accompanying low-energy QCD processes can not be described from the first principles and introduce additional complications for the precision measurements.

- The initial energy of the whole proton is known with good precision, but we don't know how this energy is distributed among the proton constituents. The absence of a consistent theory for the QCD vacuum does not allow to describe the initial states of the proton constituents.

- We know that the proton consists of three valence quarks that have trivially non-vanishing PDFs and interact through gluons. In the course of these interactions all flavours of quarks (called sea quarks) are appearing through radiative mechanisms. The contribution of these sea quarks to the scattering cross-section must also be taken in account.

In order to attack these problems and obtain accurate predictions for the proton-proton collisions it is necessary to take into account the asymptotic freedom that QCD demonstrates at short distances or high energies. At a certain energy scale of the momentum $Q$, transferred during the collision, we can assume that the interacting parts of the proton are asymptotically free and neglect the interaction with the rest of the proton. The factorization occurs only if the transferred momentum $Q \gg \Lambda_{QCD}$ is large, and that is why these processes are called "hard". The physical conditions of the hard processes allow to use the perturbative QCD formalism, since at large energy scale the strong coupling constant $\alpha_s$ becomes small. Processes with lower energy scale of the transferred momentum are called "soft" and do not allow to use the perturbative QCD formalism. As it was mentioned in Section 1, a lot of things in the low-energy non-perturbative sector of the QCD are still unclear.

The production of massive vector bosons occurs during hard processes, however, precise measurements at hadron colliders require understanding of both hard and soft QCD regimes. It is common that the hard scattering of the proton constituents is accompanied by a soft scattering of the remaining proton parts. This forms what is called *underlying event* and also must be taken into account.
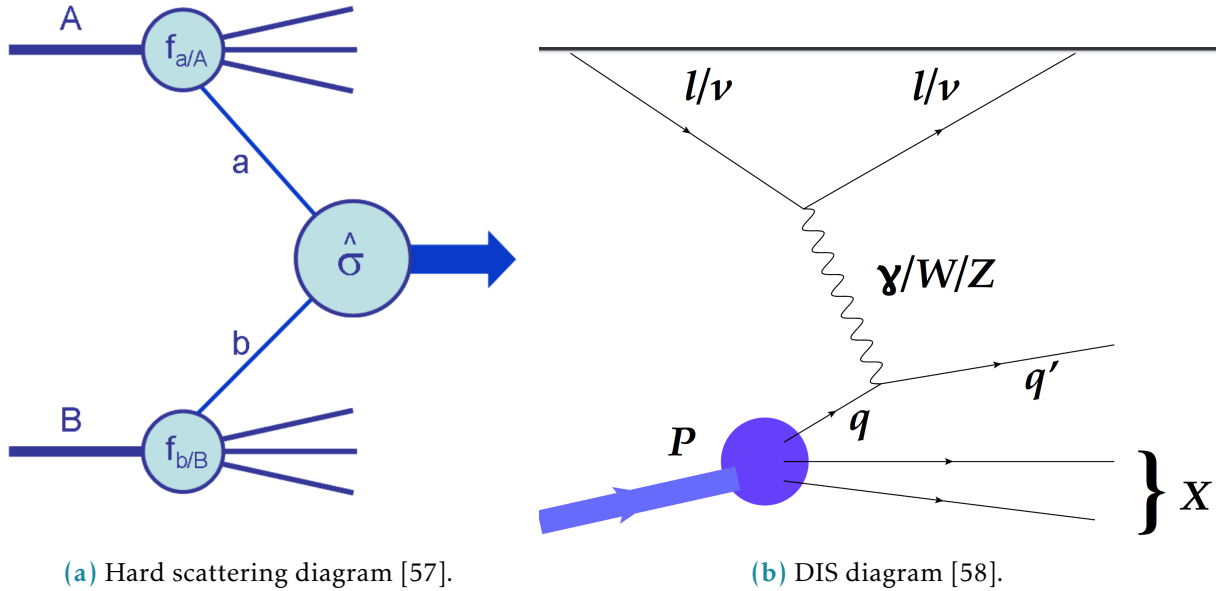


(a) Hard scattering diagram [57].          (b) DIS diagram [58].

**Figure 2.3:** Examples of hard QCD scatterings.

## 2.2.1 Deep Inelastic scattering

In order to better illustrate the factorization approach let us first consider the lepton-hadron process called the Deep Inelastic Scattering (DIS). Historically it was the first experimental evidence for the complex structure of the proton and still serves as an indispensable tool for the proton structure study. Let's try to write a matrix element for a DIS process $e + A \rightarrow e + X$, exchanging a virtual photon with momentum $q^\mu$:

$$|M|^2_{DIS} = 4\pi M_N \frac{\alpha}{q^4} L_{\mu\nu} W^{\mu\nu}_{hadron}, \tag{2.3}$$

where $L_{\mu\nu}$ is the transverse lepton tensor, $q^\mu L_{\mu\nu} = q^\nu L_{\mu\nu} = 0$. The hadronic tensor $W_{\mu\nu}$ along with its normalization factor $4\pi M_N$ is unknown, but we can write it down in general form introducing longitudinal and transverse parts[1] [59]:

$$W_{\mu\nu} = F_1(x, Q^2)\left(-g_{\mu\nu} + \frac{q_\mu q_\nu}{q^2}\right) + F_2(x, Q^2)\frac{\left(p_\mu - q_\mu p \cdot q/q^2\right)\left(p_\nu - q_\nu p \cdot q/q^2\right)}{p \cdot q}, \tag{2.4}$$

with $p_\mu$ being the momentum of the hadron A, $Q^2$ is the exchange momentum, $x = \frac{Q^2}{2p \cdot q}$ and the form-factor functions $F_1(x, Q^2)$, $F_2(x, Q^2)$ are unknown.

The cross-section of the DIS process can be measured experimentally, leaving the possibility to study the form-factor functions. It turned out that these functions do not depend (at least in the first approximation) on $Q^2$ [60]. Further experiments have revealed that the form-factors depend only on the ratio $x$, as it was predicted before [61]. This type of behaviour was called the Bjorken scaling.

These results have led to the idea of partons - point-like constituents of the proton [62]. The factorization theorem states that it is possible to express the hadronic tensor $W_{\mu\nu}$ as a sum of all available partons:

$$W_{\mu\nu}(q_\mu, p_\nu) = \sum_a \int_x^1 \frac{d\xi}{\xi} f_{a/A}(\xi, \mu) H^a_{\mu\nu}(q_\mu, p_\nu, \mu, \alpha_s(\mu)) + NLO. \tag{2.5}$$

The functions $H^a_{\mu\nu}(q_\mu, p_\nu, \mu, \alpha_s(\mu))$ are called the hard scattering structure functions and only depend on parton type $a$, but not on hadron type $A$. These functions describe the high-energy behaviour and can be calculated in the framework of perturbative QCD. At the same time $f_{a/A}(\xi, \mu)$ is called Parton Density Function (PDF) and has a physical meaning of finding a parton of type $a$ (gluon, u-quark, d-quark etc ) in a hadron of type $A$ (proton, neutron, meson) carrying the fraction of $\xi$ of the hadron's momentum. These PDFs contain information on the momentum distribution of quarks and gluons within the hadron. This corresponds to the non-perturbative sector of the QCD which is beyond the reach of theoretical methods available so far. Note that they do not directly depend on the momentum $Q^2$, but only on the energy scale $\mu$.

The DokshitzerGribovLipatovAltarelliParisi (DGLAP) equations show that once the PDFs are known at a certain energy scale $\mu$ they can be perturbatively extrapolated to a different energy scale [63], [64], [65], [66]. The PDFs are universal - they can be measured experimentally at certain conditions in the

---

[1]Given example assumes only electromagnetic interaction. For the more general electroweak case the tensor structure is more complicated and there are more than two scalar structure functions [58].

course of the DIS (or any other) process and then used for numerical calculations of any other process (e.g. Drell-Yan (DY) process) at different conditions. Such a measurement allows a workaround - we may not be able to solve the many-body problem and perform non-perturbative calculations starting from the first principles, yet we still get a theoretical prediction with a good precision. Currently there exists a number of different groups working on the PDF parametrizations and fits, constantly improving the fits using the data coming from hadron and electron-proton colliders. Using different PDF sets may give different results and also helps to estimate the systematic uncertainties implied by the PDFs.

Historically the DIS experiments at HERA electron-proton collider have allowed to perform proton PDFs measurements with a good level of precision in the $x$ region up to $x \backsim 10^{-4}$ at high $Q^2$ of up to 50 000 $GeV^2$ [67]. The HERA experiments operated until 2008, paving the path for precision predictions for the Drell-Yan process. Currently there are prospects for new experiments like Large Hadron Electron Collider (LHeC) that would involve DIS and further improve the PDF precision [68].
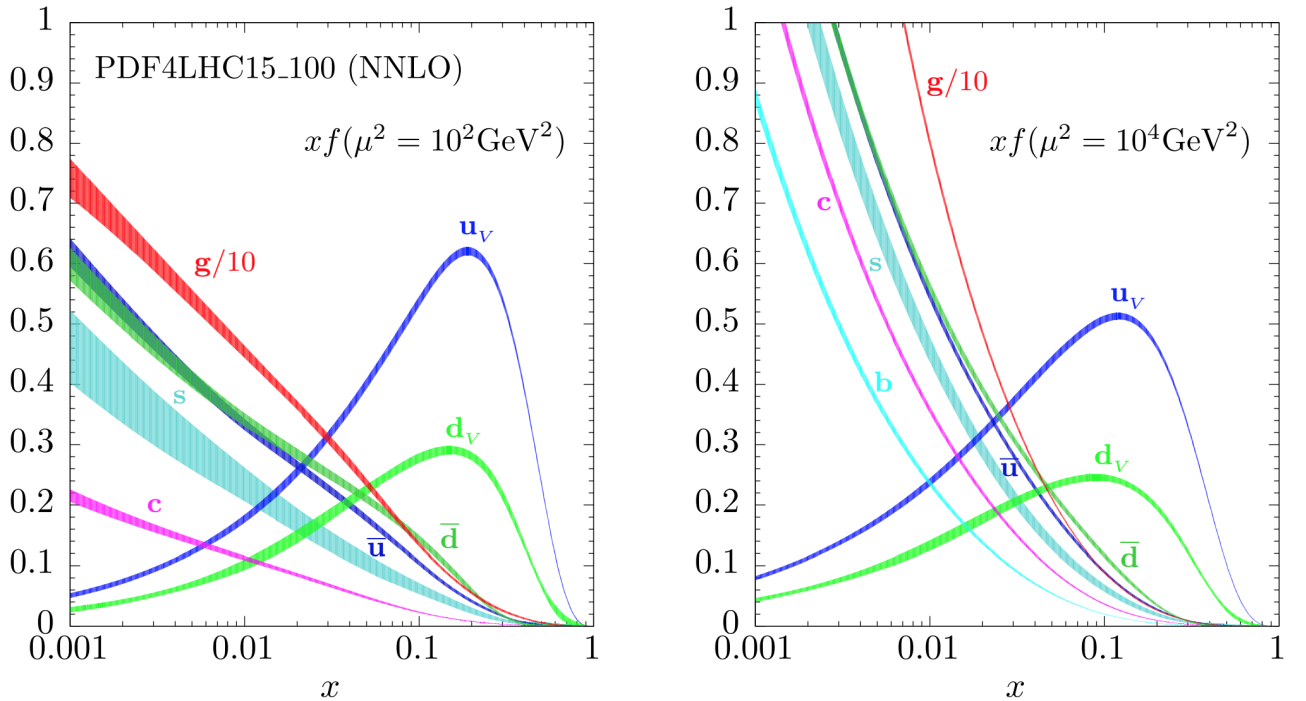


**Figure 2.4:** The evolution of a PDF4LHC15 NNLO Hessian set from $Q^2 = 10^2$ GeV to $Q^2 = 10^4$ GeV using the DGLAP. Notice the increase in the sea quark density. The PDFs include one standard deviation uncertainty band [58].

### 2.2.2 The Drell-Yan process

The DY process happens during the high-energy hadron-hadron scattering when quark and antiquark annihilate to form an electroweak boson [69]. For the neutral DY process $q\bar{q} \rightarrow Z/\gamma^* + X \rightarrow$

$l^+l^- + X$ takes place. In a similar way the charged DY process can happen, generating a W boson: $q\bar{q}' \to W^{\pm} + X \to l^{\pm}\nu + X$. It is postulated that the DY cross-section $\sigma^{DY}$ in a proton-proton scattering can be expressed through the cross-sections of the corresponding parton-parton scattering cross-section convoluted with the PDFs of these partons:

$$\frac{d^2\sigma^{DY}}{dydM^2} = \sum_{a,b=q,\bar{q},g} \int_{\tau_1}^1 dx_1 f_a(x_1,\mu_F^2) \int_{\tau_2}^1 dx_2 f_b(x_2,\mu_F^2) \frac{d^2\hat{\sigma}_{ab}^{DY}}{dydM^2}(x_1,x_2,y,M^2,\mu_R^2,\mu_F^2). \tag{2.6}$$

In this equation $y = \frac{1}{2}\log\frac{E+p_z}{E-p_z}$ represents rapidity, $M^2$ is the invariant mass of the lepton pair, $\mu_F$ and $\mu_R$ are factorization and renormalisation scales correspondingly. Integration limits $\tau_{1,2} = \sqrt{\frac{Q^2}{s}}e^{\pm y}$ with $s$ being the centre-of-mass energy of the two incoming protons. The partonic cross-sections can be in turn computed perturbatively as a series expansion in $\alpha_s$ [58]:

$$\frac{d^2\hat{\sigma}_{ab}^{DY}}{dydM^2}(x_1,x_2,y,M^2,\mu_R^2,\mu_F^2) = \sum_{n=0}^{\infty} \left(\frac{\alpha_s\mu_R^2}{2\pi}\right)^{(n)} \frac{d^2\hat{\sigma}_{ab}^{(n)DY}}{dydM^2}. \tag{2.7}$$

The cross-sections $\hat{\sigma}_{ab}^{(n)DY} \propto \alpha_s^n$ contain only the terms of order $n$ in $\alpha_s$. The exact sum of the expansion does not depend on the $\mu_F$ and $\mu_R$ parameters. However, finite-order calculations demand a specific choice for the two parameters. One of the common choices for the DY process is putting $\mu_F = \mu_R = M$, with $M$ being the mass of the dilepton pair.

From equation 2.6 we can see that the rapidity distribution of the vector boson explicitly depends on the PDFs both in terms of flavour decomposition and in the sense of a particular PDF set. Figure 2.5 demonstrates different rapidity distributions for two centre-of-mass energies and two different PDF sets.

Let us consider partonic cross-sections, which can be constructed using an analogy to QED $e^+e^- \to$
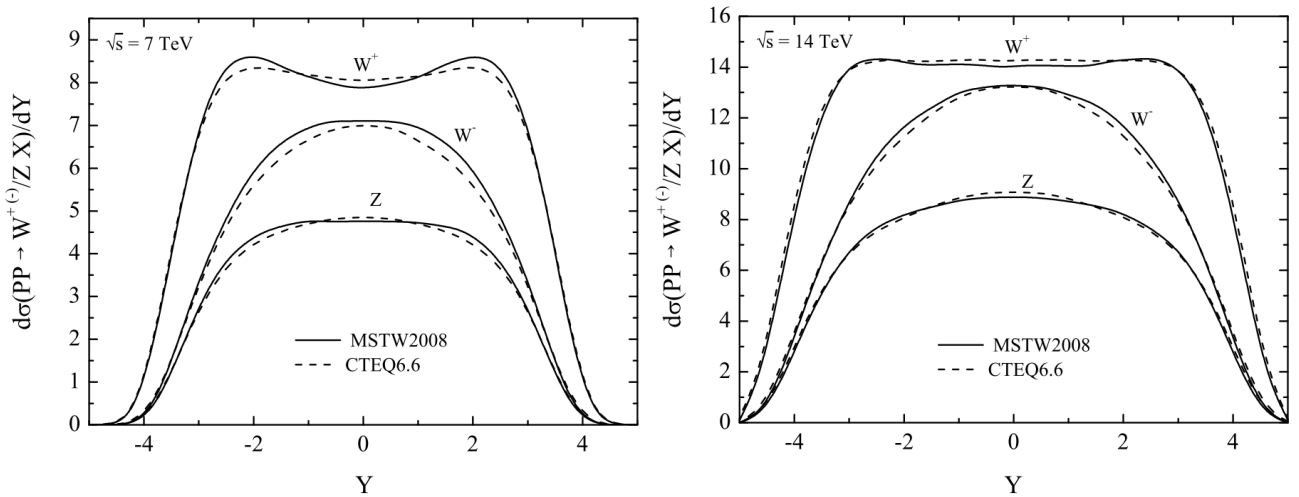


**Figure 2.5:** Rapidity distribution for the vector bosons using MSTW2008 and CTEQ6.6 PDF sets for the centre-of-mass energies of 7 and 14 TeV [70].

$\mu^+\mu^-$:

$$\hat{\sigma}(q\bar{q} \to e^+e^-) = \frac{4\pi\alpha^2}{3s}\frac{1}{N}Q_q^2. \tag{2.8}$$

Here $Q_q$ is the quark charge, $1/N$ stands for the averaging over colour factor and underlines the fact that quark and antiquark must have the matching colour in order to annihilate. In a similar way we can obtain the cross-section of the sub-processes of W and Z bosons production:

$$\begin{aligned} \hat{\sigma}^{q\bar{q}'\to W} &= \tfrac{\pi}{3}\sqrt{2}G_F M_W^2 |V_{qq'}|^2 \delta(s - M_W^2), \\ \hat{\sigma}^{q\bar{q}'\to Z} &= \tfrac{\pi}{3}\sqrt{2}G_F M_W^2 (v_q^2 + a_q^2)\delta(s - M_Z^2), \end{aligned} \tag{2.9}$$

where $V_{qq'}$ is the element of the CKM matrix, $v_q$ ($a_q$) is a vector (axial vector) that couples the Z boson to the quarks. Figure 2.6 shows the contributions of different parton flavours into $W^+$ and $W^-$ cross-sections. An assumption of narrow W resonance was used. The fact that the bosons with opposite charges are formed from different quarks makes a notable difference at the LHC experiments. Figure 2.7 contains the comparison of the results obtained at the LHC experiments with the NNLO theoretical predictions that use different PDF sets.
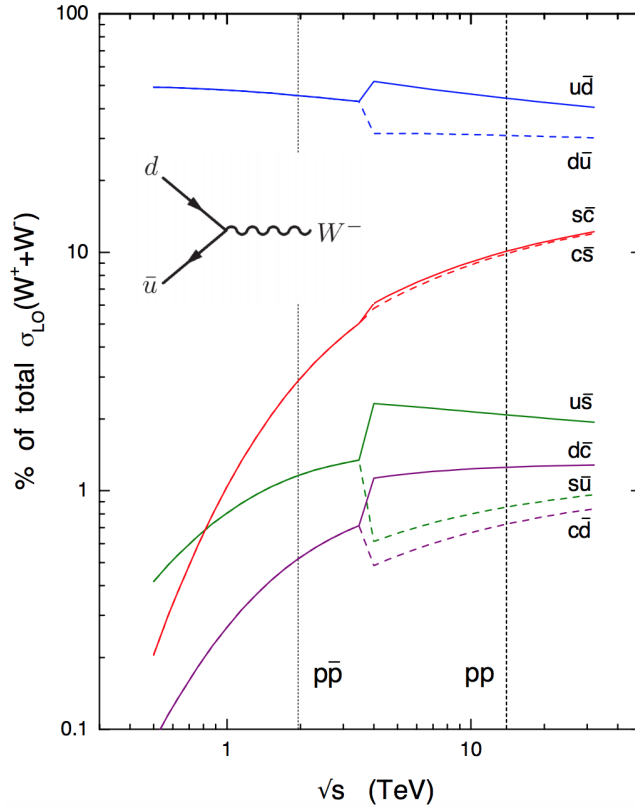


**Figure 2.6:** Parton contributions to the cross-sections of $W^+$ and $W^-$ bosons for LHC and Tevatron cases [71].
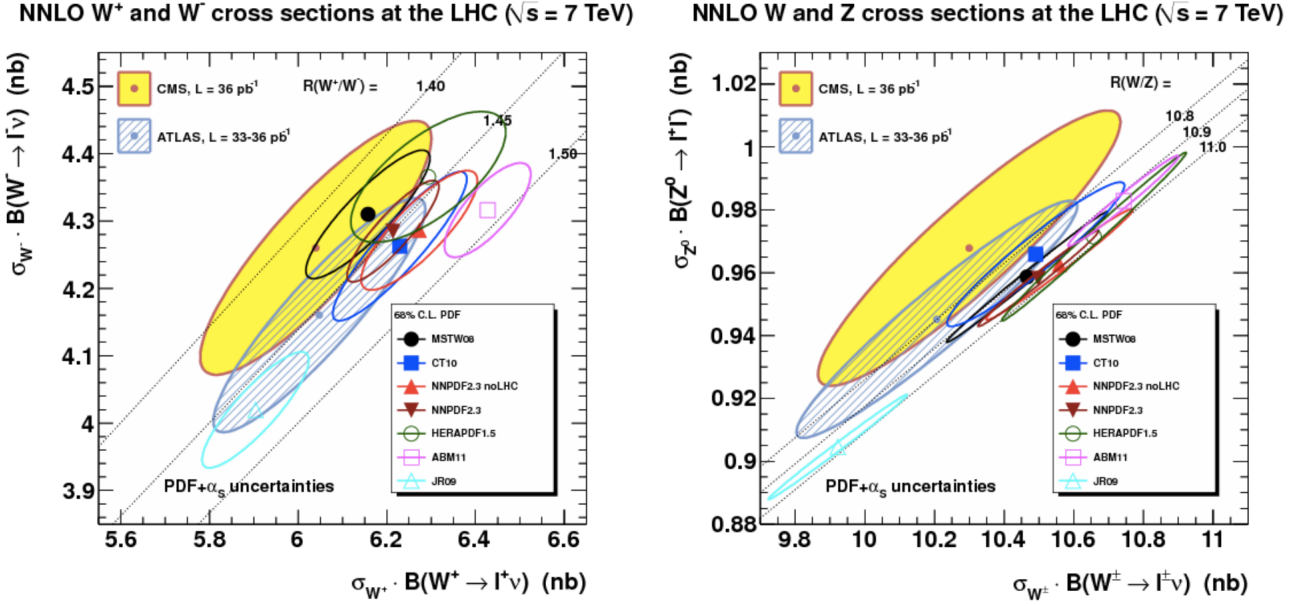
**Figure 2.7:** W and Z boson cross sections LHC at 7 TeV. ATLAS and CMS results, compared to NNLO predictions for various PDF sets [72].

## 2.3 Transverse momentum of massive vector bosons

The leading-order model of the DY process assumes the colliding partons to have their momentum perfectly collinear with the proton as a whole, which would mean that the vector boson $p_T$ should peak at zero. However most of the massive vector bosons produced in the DY process have a small yet non-zero transverse momentum, $p_T \ll M_V$. The main source for the W boson transverse momentum is the initial state radiation by one of the two quarks that create the boson. The spectrum at higher values of $p_T$ is determined by hard perturbative parton emission(s) like $q\bar{q} \rightarrow Vg$, $qg \rightarrow Vq$. The corresponding amplitudes can be conveniently expressed using Mandelstam variables:

$$
\begin{aligned}
\sum |\mathcal{M}^{q\bar{q}' \rightarrow Wg}|^2 &= \alpha_s \sqrt{2}\pi G_F M_W^2 |V_{q\bar{q}'}|^2 \frac{8}{9} \frac{t^2+u^2+2M_W^2 s}{tu}, \\
\sum |\mathcal{M}^{qg \rightarrow Wq'}|^2 &= \alpha_s \sqrt{2}\pi G_F M_W^2 |V_{q\bar{q}'}|^2 \frac{1}{3} \frac{s^2+u^2+2M_W^2 t}{-su},
\end{aligned}
\tag{2.10}
$$

where the summation is performed over colours and spins in the final and initial states. Integrating these partonic matrix elements with the PDFs one can obtain the transverse momentum distribution $d\sigma/dp_T$. Further precision can be obtained by considering corrections from next-to-leading order processes $\backsim O(\alpha_s^2)$ like $q\bar{q} \rightarrow Vgg$ - that would mainly affect the high $p_T$ tail of the distribution.

The matrix elements in 2.10 become singular when the emitted partons become soft or collinear to the initial-state partons - it is related to the poles at $u = 0$ and $t = 0$ in the denominator. Also for the NLO processes like $q\bar{q} \rightarrow Vgg$ a singularity arises if the two final-state gluons are collinear. This creates a problem for the calculation of the low-$p_T$ part of the spectrum. Mathematically it is reflected in the appearance of different powers of logarithms like $\log M_W^2/p_T^2$ in all orders of cross-section expansion in $\alpha_s$, which leads to divergences when $p_T$ is small. This forces us to look for alternative approach that
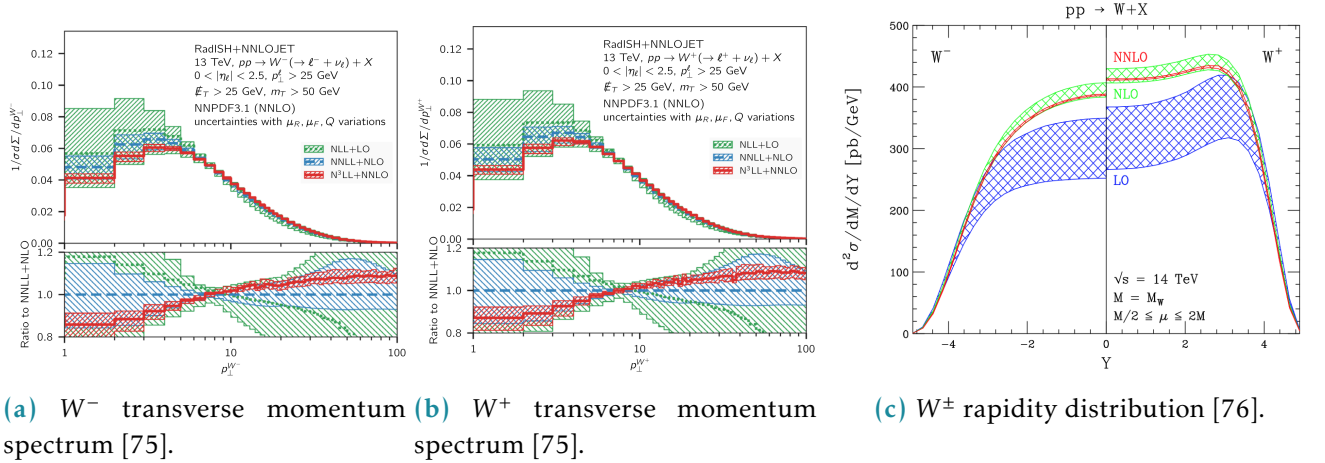
would take into account all the orders of the expansion.

All-order resummation may be performed in a variety of approaches, one of the most popular is provided by parton showers. Its numerical implementation is available in a number of Monte-Carlo generators, PYTHIA, HERWIG and SHERPA are among the most used. It appears that for the case of soft and collinear gluon emission it is possible to factorize and exponentiate the logarithms in a *Sudakov form factor*, such that:

$$\frac{d\sigma}{dp_T^2} = \sigma \frac{d}{dp_T^2} \exp\left\{-\frac{\alpha_s C_F}{2\pi} \log^2 \frac{M_W^2}{p_T^2}\right\},$$
$$\exp\left\{-\frac{\alpha_s C_F}{2\pi} \log^2 \frac{M_W^2}{p_T^2}\right\} = 1 - \frac{\alpha_s}{2\pi} C_F \ln^2 \frac{M_W^2}{p_T^2} + \frac{1}{2!}\left(\frac{\alpha_s}{2\pi}\right)^2 C_F^2 \ln^4 \frac{M_W^2}{p_T^2} - \frac{1}{3!}\left(\frac{\alpha_s}{2\pi}\right)^3 C_F^3 \ln^6 \frac{M_W^2}{p_T^2} + ... \tag{2.11}$$

The exponential $\exp\{G(\alpha_s, L)\}$, where $L = \log M_W^2/p_T^2$ is called the Sudakov form-factor. Its expansion by the powers of $\alpha_s$ defines the resummation accuracy: the term $\backsim O(\alpha_s)$ is called the leading logarithm (LL), term with $\backsim O(\alpha_s^2)$ is the next-to-leading logarithm (NLL) and so on.

The cross-sections obtained with the resummation methods provide a good prediction for soft and collinear emissions at low $p_T \ll M_W$. In order to get a combined cross-section for higher $p_T$ region the resummed cross-section has to be *matched* with the fixed-order cross-sections of the corresponding power in $\alpha_s$. Figure 2.8 contains NNLO resummed predictions for the W $p_T$ spectrum produced with RadISH[73]. In Chapter 9 of this thesis the spectrum generated by another resumming tool - DYRes [74], is used for comparison with POWHEG and SHERPA predictions. Besides the ISR phenomena, there is



(a) $W^-$ transverse momentum spectrum [75].

(b) $W^+$ transverse momentum spectrum [75].

(c) $W^\pm$ rapidity distribution [76].

**Figure 2.8:** Kinematic distributions for $W^\pm$ with corrections.

also an effect of partons moving within the colliding protons, having an intrinsic momentum of their own. This intrinsic momentum $<k_T> \backsim \Lambda_{QCD}$ is well parametrized using a Gaussian distribution with average value of 500 [77] or 700 MeV [78], although there are ongoing efforts for a more sophisticated parametrization that would allow a better modelling of the lower part of vector boson spectrum, at $p_T < 2\text{GeV}$ [79].

## 2.4   The measurement of W boson transverse momentum

As it was shown in Fig. 2.8 the shape of the W transverse momentum distribution is a difficult problem from the theoretical point of view. It heavily depends on the level of theoretical precision and the resummation technique. This is particularly illustrated by the fact that the predictions from different MC generators often do not agree (see, for instance, Figures 9.19 and 9.20). In this situation a precise measurement would help benchmarking the MC predictions and also test our understanding of the Standard Model.
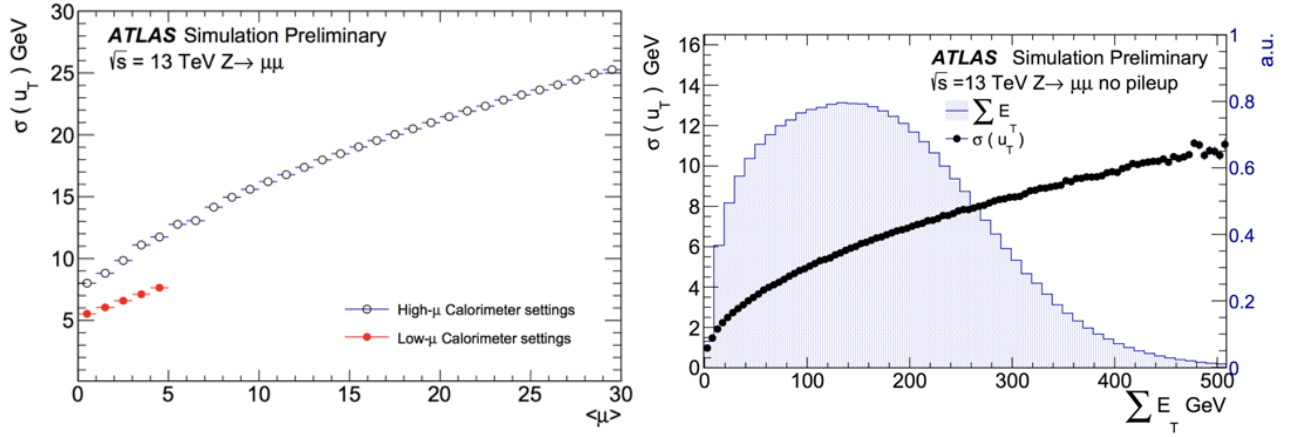
The measurement of the W boson transverse momentum at hadron colliders is a complicated problem. The hadronic decay channels can not be used for the measurement as such decay are very hard to discriminate from the QCD background. The leptonic decays have a clear signature and allow for efficient background rejection, although due to the presence of a neutrino in the final state there is no possibility to measure the transverse momentum spectrum from the final states, like in the $Z \rightarrow l^+l^-$ case.

As it was described in the previous section, the transverse momentum spectrum of vector bosons is caused by the initial state radiation (ISR). The partons produced as a result of the ISR compensate the transverse momentum of the boson. Measuring the combined momentum of these partons allows to reconstruct the vector boson momentum: $\vec{p_T}^V = -\sum \vec{p_T}^{ISR}$. This effectively means the measurement of the transverse momentum of the hadronic final states of the ISR. The corresponding observable is called the hadronic recoil (HR) and is described in more detail in Chapter 7.

The measurement of the HR is strongly dependent on the resolution of the ATLAS hadronic calorimeter. The resolution, in turn, depends on the level of noise in the calorimeter. The main sources of the noise are the pile-up (mean number of primary vertices per bunch crossing) and the underlying event. The right plot on Fig. 2.9 demonstrates a square root dependence between the calorimeter resolution and the sum of the transverse energies on all the objects in the event - $\sum E_T$. This quantity represents the combination of pile-up and underlying event activity. The plot on the right shows the dependence on pile-up and demonstrates two things: first, it is only possible to achieve a precision of 5-6 GeV and having a reasonable resolution of the spectrum peak if the pile-up is around $\langle \mu \rangle \approx 2$. That would also allow to use lower calorimeter threshold improving the resolution (calorimeter threshold is explained in more detail in Chapter 6).

So far there were only two measurements of the W boson $p_T$ at the LHC: in ATLAS [81] and in CMS [82]. Both measurements have used the data collected the special low pile-up runs, and both suffered from low statistics (31 $pb^{-1}$ for ATLAS and 18.4 $pb^{-1}$ for CMS) that resulted in coarse binning (about 8 GeV at low $p_T$) and high relative uncertainties of about 2.5% per bin.

Another strong motivation for a precise measurement of the W boson transverse momentum is its importance for the W boson mass measurement. W boson mass is extracted from the kinematic distributions that directly depend on the $p_T$ distribution of the W boson.

**Figure 2.9:** The dependence of the hadronic calorimeter resolution on pile-up (left) and on the soft event activity (right) [80].

For the Tevatron measurement of the W mass the position of the W boson transverse spectrum peak was taken from the Z $p_T$ spectrum, which was measured with good precision using the leptonic final states of the Z boson decay [83]. However, this approach would lead to much higher uncertainties at the LHC energies due to significantly larger fraction of W bosons induced by the second generation quarks.

The ATLAS measurement of the W boson mass has used the W $p_T$ extrapolated from the Z boson $p_T$ using the differential cross-section ratio:

$$R_{W/Z}(p_T) = \left(\frac{1}{\sigma_W}\frac{d\sigma_W(p_T)}{dp_T}\right)\cdot\left(\frac{1}{\sigma_Z}\frac{d\sigma_Z(p_T)}{dp_T}\right)^{-1}, \tag{2.12}$$

where the W $p_T$ cross-section was taken from the ATLAS 2011 measurement [81]. This has allowed to perform a very precise measurement of the W boson mass, although the $p_T$ spectrum extrapolation uncertainty turned out to be the second largest in the measurement reaching 8-9 GeV. The dominant PDF uncertainty has a comparable magnitude of 9-10 GeV, although it is expected to improve from the Drell-Yan process measurements to come.

The goal for the new measurement of a W transverse momentum spectrum is to have a binning of around 5-6 GeV at low $p_T$ and a relative uncertainty of about 1%. This would allow to reduce the QCD modelling uncertainty by as much as 50%, significantly improving the W boson mass measurement precision [80].

<div align="right"># The Large Hadron Collider **3**</div>

The chapter on the Large Hadron Collider provides a bit of overview on the purpose and operation principle of the collider. It also provides the information on the special low pile-up run which is the main source of experimental data for the W boson transverse momentum measurement analysis.

## 3.1  Introduction

The study of elementary particles naturally demands a stable source of particles. At the dawn of particle physics the two main sources were radioactive materials and cosmic rays. However soon researchers became in need of a more reliable source of particles in terms of particle energy, luminosity and experimental repeatability. This has commenced the era of particle accelerators.

The first examples of particle accelerators were designed in the late 1920s and in the early 1930s. Two different designs emerged: linear and circular. The former accelerates particles via electric field during the single pass through the machine, while the latter uses magnetic field to make accelerated particles go in circles allowing to re-accelerate the same beam many times. On the other hand the circular design comprises energy losses due to Bremsstrahlung radiation.

In the second half of the XX$^{th}$ century the accelerators gradually got bigger and bigger in both size and centre-of-mass energy of the accelerated particles. This has allowed to create an experimental basis for the development of modern particle physics, notably the Standard Model.

Up to this day the biggest particle accelerator with the highest centre-of-mass energy is the Large Hadron Collider (LHC). The LHC is a circular collider that lies in a tunnel of 27 km under the French-Swiss border next to Geneva [84]. In 2012 the two biggest experiments of LHC have claimed the discovery of the Higgs boson, the last elementary particle predicted by the Standard Model which was not yet discovered by that time. [85], [86].

## 3.2  The LHC running sequence

It takes quite a journey for a proton to travel from a hydrogen tank (Fig. 3.1) into one of the LHC's collision points. A resourceful system of pre-accelerators is necessary to make the proton beam ready to get injected into one of the two LHC beam pipes. The LHC accelerator complex was not built from

**Figure 3.1:** A hydrogen tank supplies LHC with protons [87].

scratch - it uses vast CERN infrastructure, that was built for the previous particle physics experiments. After stripping the electrons off the atoms of hydrogen using a magnetic field the yielded protons get accelerated to the energy of 50 MeV by the Linac 2[1] [88]. After that the beam gets into the Proton Synchrotron Booster (PSB) to be accelerated to 1.4 GeV. The next link of the pre-acceleration chain is the Proton Synchrotron (PS) - a true veteran among CERN accelerators that first accelerated protons in 1959 breaking the world record in acceleration energy. Currently thanks to PSB and other modifications it can sustain proton beam intensity 1000 times larger than back in 1959. The PS accelerates the beam up to 25 GeV and conveys it further to the Super Proton Synchrotron (SPS) - the second-largest particle accelerator at CERN. Back in 1983 the massive electroweak bosons were discovered at the SPS but even now it serves as a main accelerator for the NA61/SHINE, NA62 and COMPASS experiments. The SPS raises the beam energy to 450 GeV and finally injects it into the LHC beam pipes (see Fig 3.2).

The LHC has inherited its 27 km tunnel from the predecessor, an electron-positron collider called Large Electron-Positron (LEP). However, all the LEP hardware has been replaced to sustain the conditions of the LHC beam. About 2/3 of the LHC circumference length is occupied by the dipole magnets that bend the trajectory of the proton beam to keep it within the pipe. These magnets use superconducting coils that conduct a current of 11080 amperes to produce a magnetic field of 8.3 Tl. Proton acceleration is maintained by the radio-frequency (RF) cavities (Fig. 3.3a). Besides the acceleration particles the RF cavities are also responsible for beam bunching i.e. separating the beam into a train of separated particle packs, each containing about $10^{11}$ protons. During LHC Run 2 the bunches were separated by 7 meters (25 ns) with a maximum of 2556 circulating bunches. The LHC has four crossing points, where the two beams are crossed in order to collide protons. Naturally, the particle detectors are installed at these four points. Before getting directed at the crossing point the beams get

---

[1]After Run 2 the Linac 2 has been decommissioned to be succeeded by Linac 4.
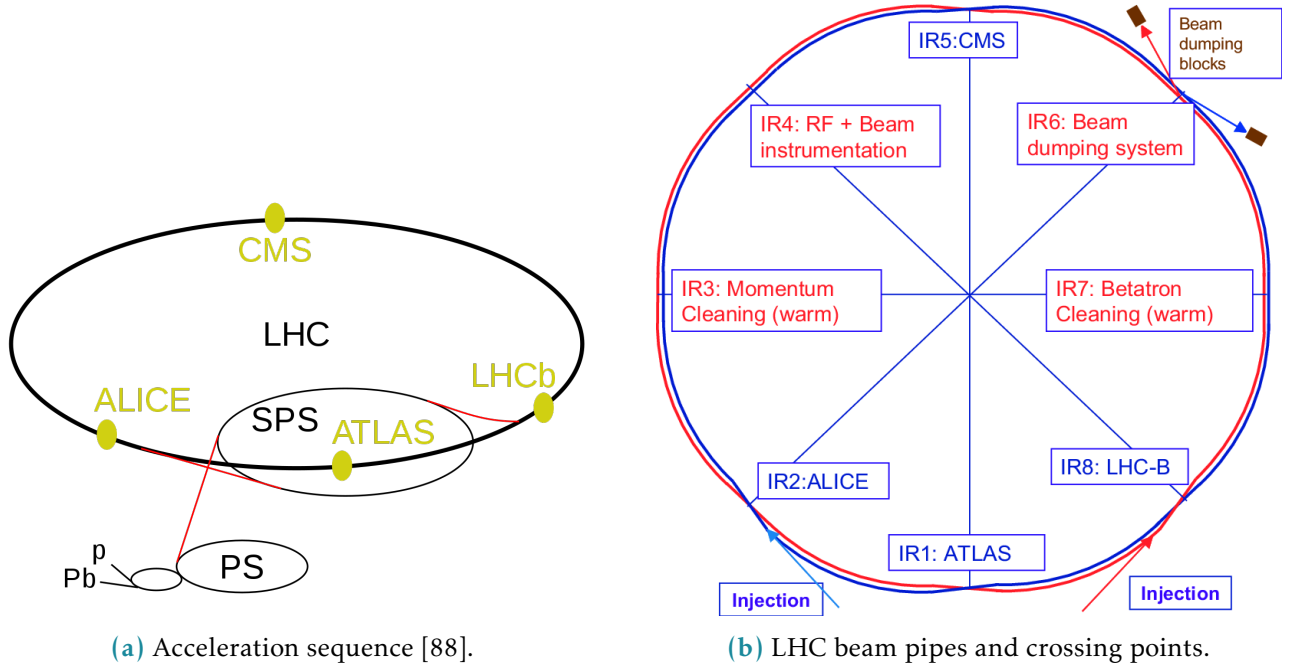
(a) Acceleration sequence [88].



(b) LHC beam pipes and crossing points.

**Figure 3.2:** Schematic depiction of the LHC ring.



(a) The RF cavities [89].



(b) Bunch behaviour at the RF cavities [90].

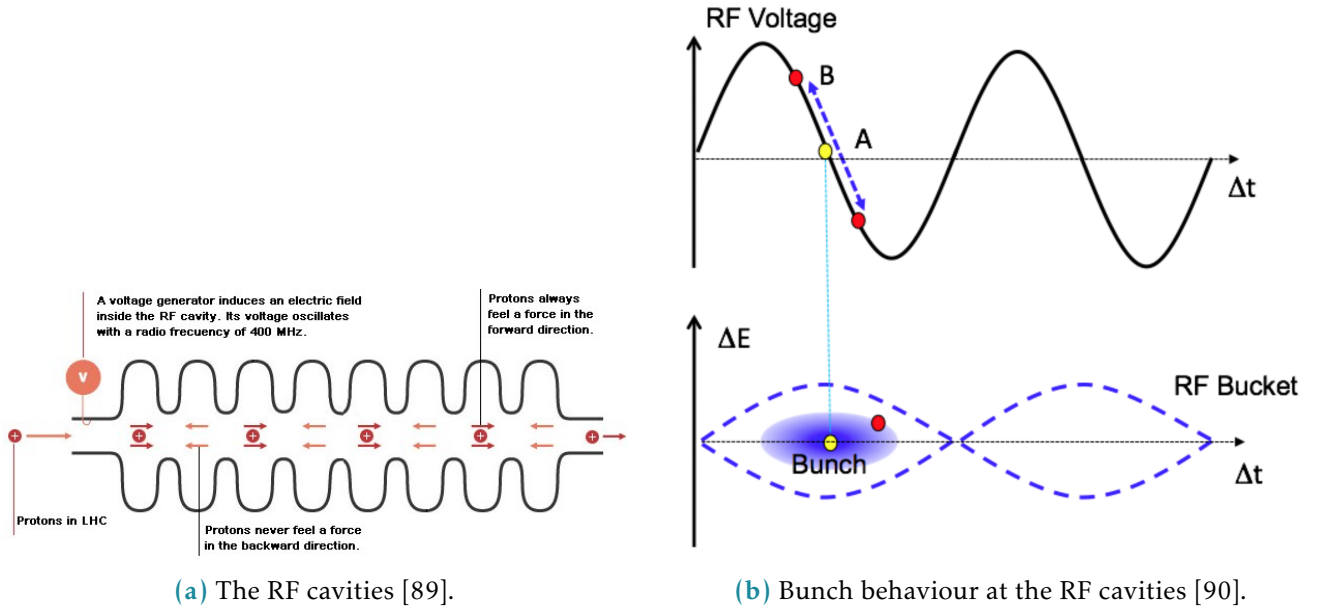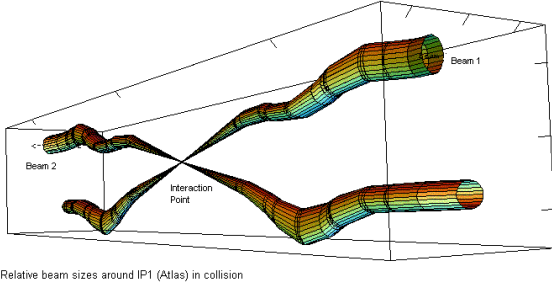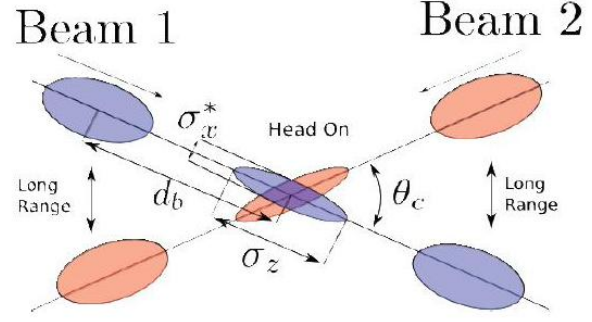**Figure 3.3:** Bunching at RF cavities

squeezed to make their cross-section as small as 16 $\mu m^2$ (Fig 3.4a).



(a) The two beams getting squeezed at the IP [91].



(b) Bunches at the collision point [92].

**Figure 3.4:** Bunch crossing at the LHC.

In order to estimate the number of single proton-proton interactions in the crossing beams a value called instantaneous luminosity (simply called luminosity) is introduced. It is the proportionality factor between the number of events per second $dR/dt$ and the cross-section $\sigma_p$:

$$\frac{dR}{dt} = \mathcal{L} \cdot \sigma_p.$$

For the case of head-on collisions the luminosity would equal to [93]:

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{4 \Pi \sigma_x \sigma_y}, \tag{3.1}$$

with $N_1$ and $N_2$ being the intensities of the two colliding beams, $f$ is the revolution frequency, $N_b$ - the number of bunches per beam, $\sigma_x, \sigma_y$ - the r.m.s. beam widths in the corresponding dimensions, assuming that the bunches in both beams have the same size and Gaussian profiles.

Head-on crossing of the beams would ensure maximal luminosity given the same beams, but on the other hand the measurement would suffer from unwanted beam-to-beam effects. To avoid it the beams at the LHC are crossed at an angle, which is called the crossing angle (see Fig. 3.4b). For the case of head-on collisions the luminosity gets a factor $\mathcal{F}$ [93]:

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{4 \Pi \sigma_x \sigma_y} \cdot \mathcal{F}, \tag{3.2}$$

with geometric factor

$$\mathcal{F} = \frac{1}{\sqrt{1 + \left( \frac{\sigma_s}{\sigma_x} \frac{\theta_c}{2} \right)}},$$

where $\sigma_s$ is the r.m.s. of the bunch length and $\theta_c$ is the crossing angle. Varying the parameters like beam intensity, bunch spacing, beam profile, crossing angle and others becomes a flexible tool for luminosity control. This comes in handy for different physics analysis, as some processes are rare

| Collision energy | Year | Integrated luminosity, $pb^{-1}$ | Total uncertainty, % |
|---|---|---|---|
| 5 TeV | 2017 | 258 | 1.6 |
| 13 TeV | 2017 | 148 | 2.1 |
| 13 TeV | 2018 | 193 | 1.5 |

**Table 3.5:** Energy and luminosity of the special low-mu runs.

and demand as much luminosity as possible (this is true, for example, for most of the Higgs studies), whereas the others suffer from high pile-up conditions. The instantaneous luminosity integrated over a period of time is called the integrated luminosity:

$$\mathcal{L}_{int} = \int_0^T \mathcal{L}(t)dt, \tag{3.3}$$

and is directly related to the number of observed events $\mathcal{L}_{int} \cdot \sigma_p = N_{events}$. A precise measurement of the integrated luminosity is crucial for the LHC results since the uncertainty on it impacts most of the analyses. A comprehensive overview on the luminosity determination at proton colliders can be found here [94]. Absolute luminosity measurements at the LHC are performed predominantly using the van-der-Meer (vdM) scan method [95], [96].

## 3.3   Special low pile-up run during LHC Run 2

During the Run 2 that lasted from 2015 to 2018 the ATLAS experiment has collected 146.9 $fb^{-1}$ of data under different bunch crossing conditions (see Fig. 3.5). However, the precise measurement of the W boson-related processes demands special conditions. High number of proton-proton collisions per bunch crossing leads to contamination of the final state signal with soft collisions products. This effect, known as pile-up, complicates object reconstruction and results in systematic uncertainties growth. For this reason two special runs with low number of interactions per bunch crossing have been performed by the LHC in 2017 and 2018 at the energies of 5 and 13 TeV. Table 3.5 contains information on the data collected at ATLAS experiment during the special low pile-up run with $< \mu > \approx 2$.
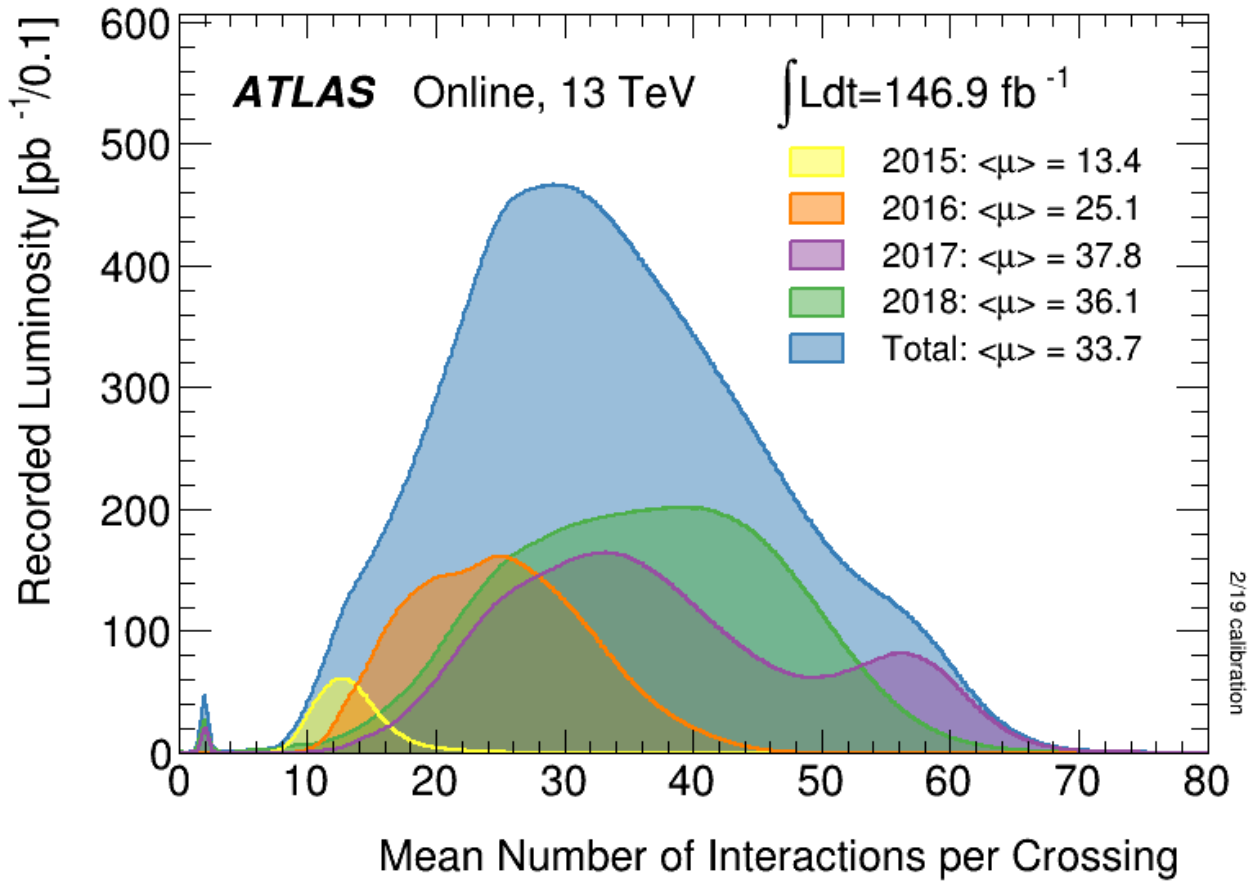
**Figure 3.5:** Number of Interactions per bunch crossing in A Toroidal LHC ApparatuS (ATLAS) Run 2 [97]. The little bump around $\mu \approx 2$ corresponds to special low pile-up runs.

# The ATLAS detector <span style="float:right">4</span>

The ATLAS detector structure is reviewed in the current chapter. After a brief introduction into the general layout of the ATLAS detector all the subsystems are considered it more detail.

## 4.1   General description and layout

The ATLAS experiment is a multipurpose detector at the LHC built, along with its peer Compact Muon Solenoid (CMS), in order to probe the p-p, A-A and p-A collisions using the full LHC luminosity [98]. Being the largest (but not the heaviest) detector ever built for a collider experiment the ATLAS detector is 44m length, 25m tall and weights 7000 tonnes.

The detector has a cylindrical shape and is an onion-like arrangement of several detector systems centred at the Interaction Point (IP) as shown in fig. 4.1. The sub-detectors operate in the magnetic field created by the solenoid and toroid magnets (ATLAS owes its name to the latter). Data acquisition and recording is controlled by the Trigger and Data Acquisition (TDAQ) systems, allowing eventually to lower the event rate to a value acceptable for the data storage. The named systems are described in more detail in this chapter.

## 4.2   Coordinate system

The ATLAS results often reference ATLAS coordinates briefly described in this subsection. The origin of the right-handed coordinate system is placed at the IP with the $z$-axis directed along the beam direction. This, in turn, defines the transverse $x$-$y$ plane with the $x$ axis pointing towards the centre of the LHC ring and the $y$ axis directed upwards. All transverse observables like $p_T$ and $E_T$ are defined in this 2D plane. Besides the mentioned Cartesian coordinates the azimuthal angle $\phi$ is defined in the transverse plane around the beam axis. The polar angle $\theta$ is the elevation angle measured from the beam axis. The following metric quantities are also to be mentioned:

- Pseudorapidity $\eta = -\ln \tan(\theta/2)$,

- Rapidity $y = 1/2 \ln \left[ (E+p_z)/(E-p_z) \right]$
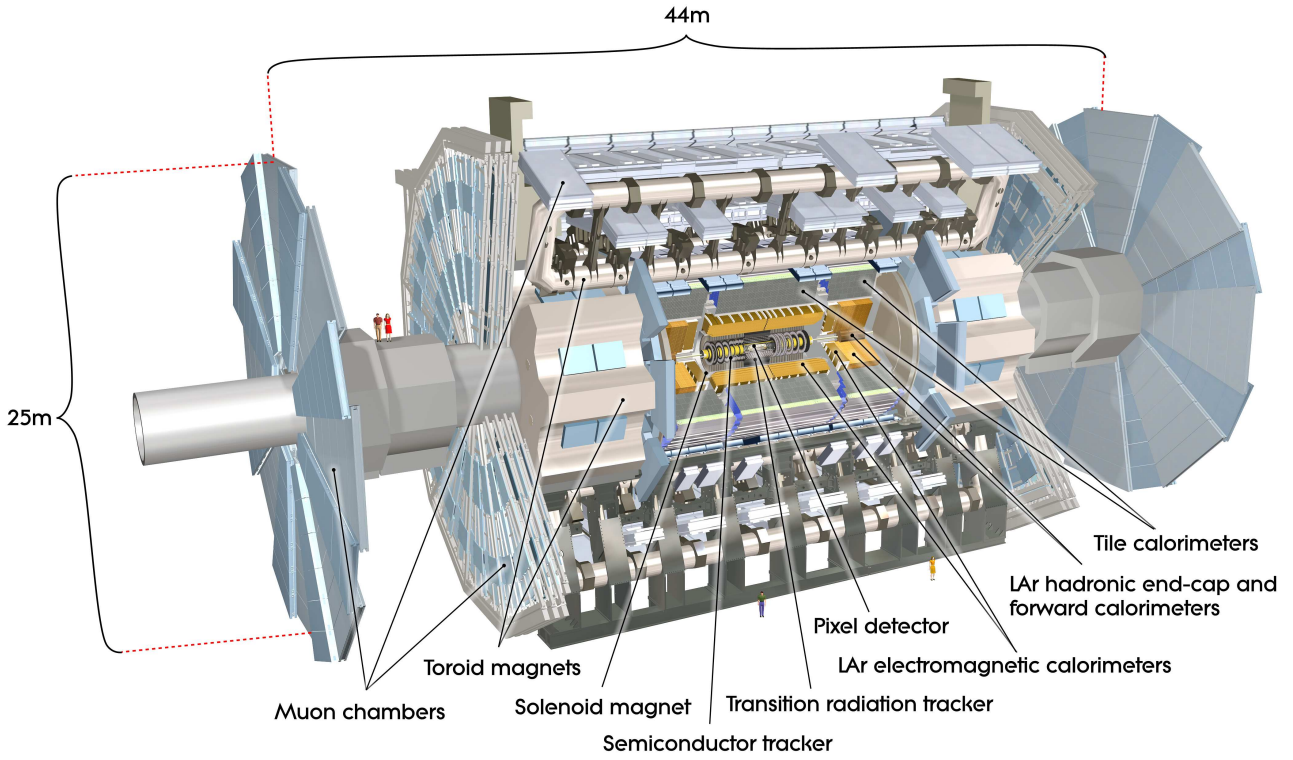
**Figure 4.1:** ATLAS detector general layout

- The distance between particles $\Delta R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$

## 4.3   Magnet system and magnetic field

ATLAS has a hybrid system of four superconducting magnets which has 22 m in diameter, 26 m in length and stores an energy of 1.6 GJ [99]. The windings of the magnets are schematically shown in Fig. 4.2. The four magnets that the magnet system comprises are the following:

- The central solenoid is aligned with the beam axis providing a 2T axial magnetic field for the inner detector.

- A barrel toroid produces a toroidal magnetic field of about 0.5T for the muon detectors in the barrel region.

- Two end-cap toroids produce a toroidal magnetic field of approximately 1T for the muon detectors in the end-cap regions.
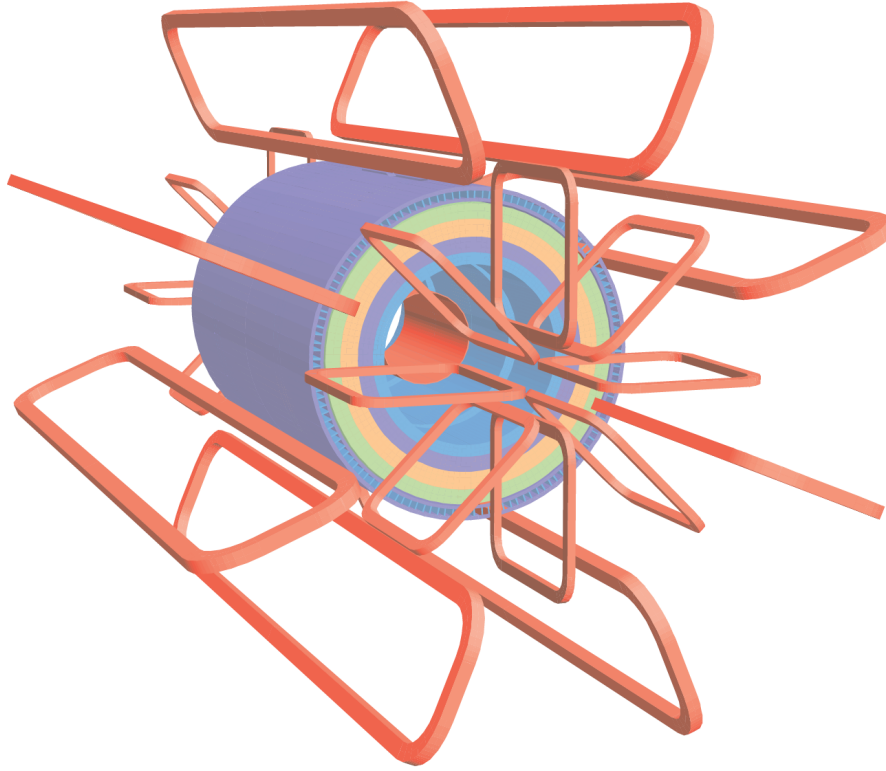
**Figure 4.2:** Geometry of ATLAS magnet windings

## 4.4 Inner detector

The ATLAS Inner Detector (ID) is designed to deliver pattern recognition, high-resolution momentum measurement [100],[101] along with primary and secondary vertex determination for charged particle tracks above a designated $p_T$ threshold of 0.5 GeV (in some cases being capable of going as low as 0.1 GeV) within the pseudorapidity range $|\eta| < 2.5$. The inner detector provides reliable electron identification in the rapidity range of $|\eta| < 2.0$ for energies from 0.5 GeV to 150 GeV.
The ID layout is a result of the technical requirements: it is assembled in a cylindrical envelope of 3512 mm in length and 1150 mm in radius. It is surrounded by the magnetic field of 2T imposed by the superconducting solenoid (see section 4.3).

Three independent sub-detectors complement each other in the inner detector (see Fig. 4.3a):

- Silicon pixel with 3 cylindrical layers for the barrel and 3 discs on each side for the end-cap. It provides the highest granularity around the vertex region. Normally each track hits three pixel layers. The pixel detector has about 80.4 million readout channels. Each of the 1744 identical pixel sensors has 47232 pixels and 46080 readout channels. About 90% of the pixels have a size of 50x400 $\mu m^2$, the remaining pixels are a bit longer: 50x600 $\mu m^2$.

- Silicon microstrip layers (SCT) with 4 cylindrical layers and 9 discs on each side for the end-cap.

A track typically crosses the strip layers in four space points. The SCT has approximately 6.3 millions readout channels from its 15912 sensors. There are 768 active strips of 12 cm length and 80 $\mu$m width per sensor plus two bias potential strips on the sensor edges.

- Transition radiation tracker (TRT) with 73 straw planes in the barrel and 160 straw planes in the end-cap. The TRT has around 351,000 readout channels and detects in average 36 hits per track. The straw tubes that make up the TRT module are 4 mm thick and 1.44 m long (0.37 m in the end-cap) and made out of polyamide films reinforced with carbon fibres. The straws are filled with a gas mixture of 70% Xe, 27% $CO_2$ and 3%$O_2$ and supplied with gilded tungsten anodes which are directly connected to the readout channels. The pixel and SCT sensors are highly radiation-proof and operate in the temperature range from -5°C to -10°C to minimize the radiation damage, while the TRT module operates at room temperature.



(a) Inner detector      (b) A perspective cut-away view of the pixel detector.
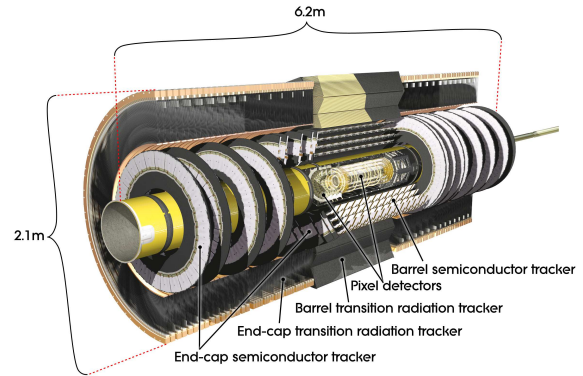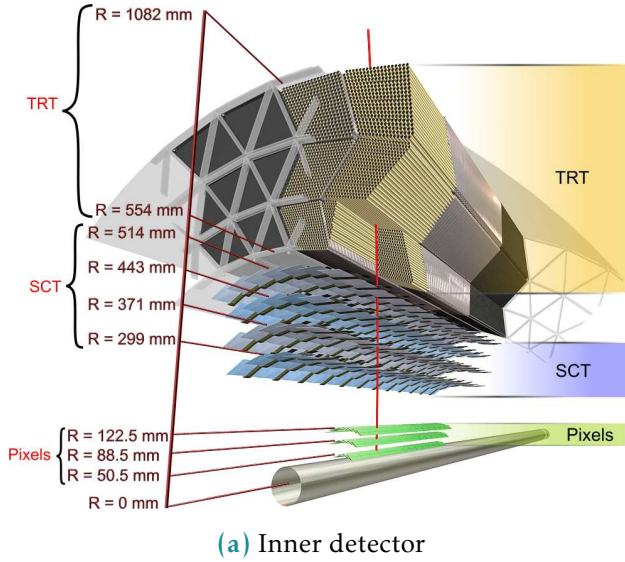
## 4.5 Calorimeter system

The ATLAS calorimeter system covers the rapidity range within $|\eta| < 4.9$ and consists of several different detector systems. The rapidity region matched to the inner detector possesses fine granularity perfectly suited for high-precision measurements of photons and electrons. The remaining part's granularity is coarser but enough to perform hadronic jet reconstruction. The view of the ATLAS calorimeter is presented on fig. 4.4.

Besides measuring the energy of travelling particles calorimeters must also contain electromagnetic and hadronic showers, limiting their ability to penetrate the calorimeter completely and get to the muon chambers. This provides a typical scale for the size of the calorimeter modules: the EMC [102] is >22 radiation lengths ($X_0$) in the barrel and >24$X_0$ in the end-caps. The hadronic calorimeter has a thickness of 9.7 interaction lengths ($\lambda$) in the barrel and 10$\lambda$ of in the end-cap, which is enough to

keep the leakage level below the typical muon background. This size also provides good resolution for the hadron energy measurement. The detailed description of the calorimeter system can be found in table 4.4. The tile calorimeter [103] uses scintillating tiles as active material alternated with steel absorbers. All the other calorimeter systems use liquid argon as an active medium with lead sampling.
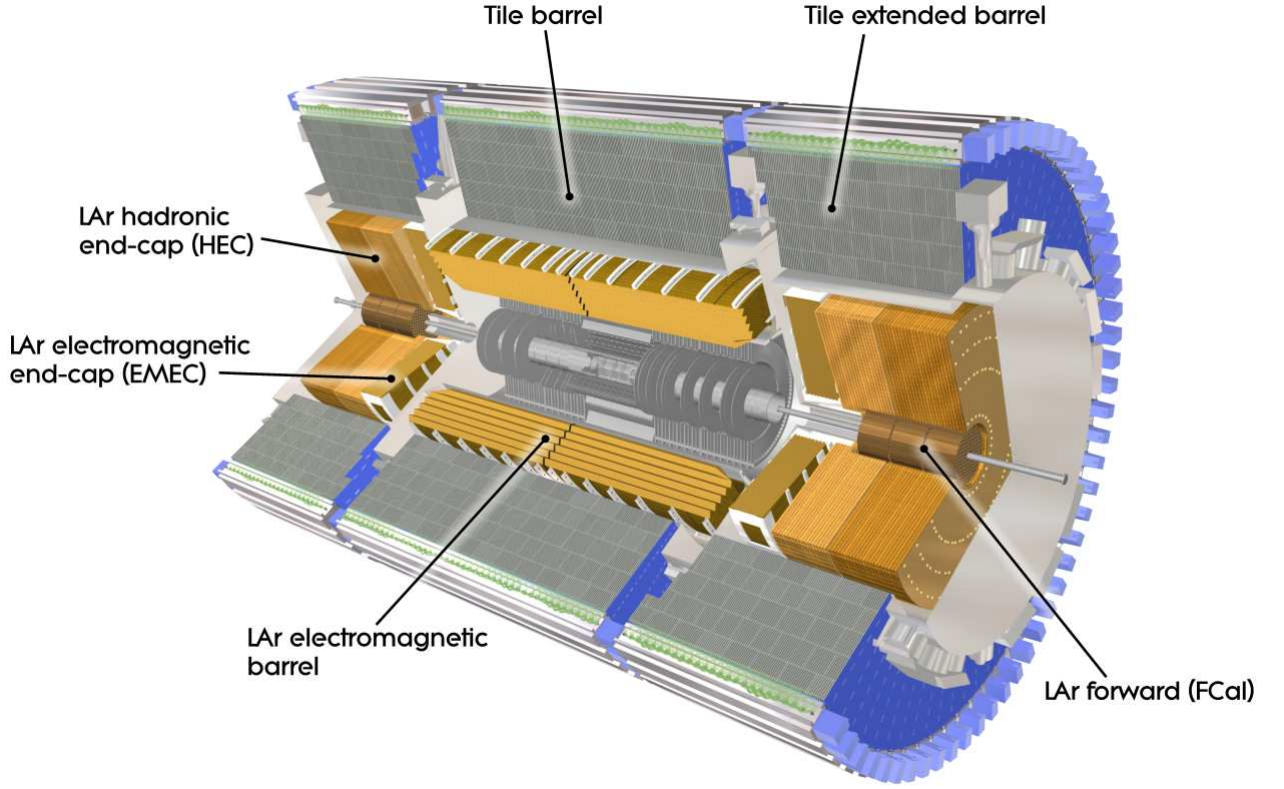


**Figure 4.4:** ATLAS calorimeter general layout

### 4.5.1 Electromagnetic calorimeter

The EMC has two submodules:

- EMC barrel detector.

- Electromagnetic end-cap calorimeter (EMEC) end-cap detector.

The EMC barrel module consists of two identical half-barrels 3.2 meters long with inner and outer radii 2.8 m and 4 m respectively. There is a 4 mm gap at $z = 0$ between the half-barrels. The second crack is situated between the barrel and the end-cap at $1.37 < |\eta| < 1.52$. The EMEC comprises two pairs of coaxial wheels of 63 cm thickness having inner and outer radii of 330 mm and 2098 mm respectively. The crack between the two wheels makes a third crack at $|\eta| = 2.5$. Both barrel and

| | Barrel | | End-cap | |
|---|---|---|---|---|
| **EM Calorimeter** | | | | |
| Number of layers and $|\eta|$ coverage | | | | |
| Presampler | 1 | $|\eta| <1.52$ | 1 | $1.5 < |\eta| <1.8$ |
| Calorimeter | 3 | $|\eta| <1.35$ | 2 | $1.375 < |\eta| <1.5$ |
| Granularity $\Delta\eta x\Delta\phi$ versus $|\eta|$ | | | | |
| Presampler | 0.025×0.1 | $|\eta| <1.52$ | 0.025×0.1 | $1.5 < |\eta| <1.8$ |
| Calorimeter 1st layer | 0.025/8×0.1 | $|\eta| <1.40$ | 0.050×0.1 | $1.375 < |\eta| <1.425$ |
| | 0.025×0.025 | $1.425 < |\eta| <1.5$ | 0.025×0.1 | $1.425 < |\eta| <1.5$ |
| | | | 0.025/8×0.1 | $1.5 < |\eta| <1.8$ |
| | | | 0.025/6×0.1 | $1.8 < |\eta| <2.0$ |
| | | | 0.025/4×0.1 | $2.0 < |\eta| <2.4$ |
| | | | 0.025×0.1 | $2.4 < |\eta| <2.5$ |
| | | | 0.1×0.1 | $2.5 < |\eta| <3.2$ |
| Calorimeter 2nd layer | 0.025×0.025 | $|\eta| <1.40$ | 0.050×0.1 | $1.375 < |\eta| <1.425$ |
| | 0.075×0.025 | $1.4 < |\eta| <1.475$ | 0.025×0.025 | $1.425 < |\eta| <2.5$ |
| | | | 0.1×0.1 | $2.5 < |\eta| <3.2$ |
| Calorimeter 3rd layer | 0.050×0.025 | $|\eta| <1.35$ | 0.050×0.025 | $1.5 < |\eta| <2.5$ |
| Number of readout channels | | | | |
| Presampler | 7808 | | 1536 (both sides) | |
| Calorimeter | 101760 | | 62208 (both sides) | |
| **LAr hadronic end-cap** | | | | |
| $|\eta|$ coverage | | | $1.5 < |\eta| <3.2$ | |
| Number of layers | | | 4 | |
| Granularity $\Delta\eta \times \Delta\phi$ | | 0.1 × 0.1 | $1.5 < |\eta| <2.5$ | |
| | | 0.2 × 0.2 | $2.5 < |\eta| <3.2$ | |
| Readout channels | | | 5632 (both sides) | |
| **LAr forward calorimeter** | | | | |
| $|\eta|$ coverage | | | $3.1 < |\eta| <4.9$ | |
| Number of layers | | | 3 | |
| Granularity $\Delta x \times \Delta y$ | | | FCal 3.0 × 2.6 | $3.15 < |\eta| <4.30$ |
| | | | FCal: ~four times finer | $3.10 < |\eta| <3.15$ |
| | | | | $4.30 < |\eta| < 4.83$ |
| | | | FCal2 3.3 × 4.2 | $3.24 < |\eta| <4.50$ |
| | | | FCal2: ~four times finer | $3.20 < |\eta| <3.24$ |
| | | | | $4.50 < |\eta| < 4.81$ |
| | | | FCal3 5.4 × 4.7 | $3.32 < |\eta| <4.60$ |
| | | | FCal3: ~four times finer | $3.29 < |\eta| <3.32$ |
| | | | | $4.60 < |\eta| < 4.75$ |
| Readout channels | | | 3524 (both sides) | |
| **Scintillator tile calorimeter** | | | | |
| | Barrel | | Extended barel | |
| $|\eta|$ coverage | $|\eta| < 1.0$ | | $0.8 < |\eta| <1.7$ | |
| Number of layers | 3 | | 3 | |
| Granularity $\Delta\eta \times \Delta\phi$ | | 0.1 × 0.1 | 0.1 × 0.1 | |
| | | 0.2 × 0.2 | 0.2 × 0.1 | |
| Readout channels | 5760 | | 4092 (both sides) | |

58

Table 4.4: ATLAS calorimeter in numbers

end-cap electromagnetic calorimeters are designed to have an accordion-shaped absorbers made out of lead plates, coated in stainless steel sheets. The readout electrodes are placed in the gaps between the absorbers. This type of geometry allows full coverage in $\phi$ without cracks together with fast extraction of the signal from both sides of the electrodes. The orientation of the accordion waves is axial in the barrel and radial in the end-caps (see fig. 4.4). These features of the calorimeter lead to virtually uniform performance in the $\phi$ dimension.

Segmentation in $\eta$ is very different in the layers of the calorimeter, but the second layer always has the finest granularity because electrons and photons are supposed to leave most of their energy in the second calorimeter layer. In order to correct for the energy losses upstream the barrel calorimeter is preceded by a thin LAr active layer of 11 mm called presampler. For more details on $\eta$ coverage and granularity see table 4.4.

### 4.5.2 Hadronic calorimeter

The hadronic calorimeter (HC) is composed of three submodules:

- HC scintillating tile detector, a steel sampled detector divided in turn into central barrel having 5.8 m in length and two extended barrels 2.6 m in length each. The extended barrels have inner radii of 2.28 m and outer radii of 4.25 m. The tile calorimeter consists of three layers having about 1.5, 4.1 and 1.8 interaction lengths $\lambda$ in the barrel and 1.5, 2.6 and 3.3 $\lambda$s in the extended barrel.

- Hadronic end-cap calorimeter (HEC) detector is a liquid argon calorimeter sampled with copper. It has two pairs of independent wheels symmetrically located behind the EMEC called the front and the rear wheel. The wheels are cylindrical, their outer radius is 2030 mm.

- Forward calorimeter (FCal) detector modules are located about 4.7 m from the IP and are subjected to very high particle flux and radiation. It consists of three wheels 45 cm deep each. The first one, FCal1 is sampled with copper intended for the measurement of electromagnetic processes. The two other wheels FCal2 and FCal3 are sampled with tungsten and designed for the hadronic showers measurement.

The number of the readout channels as well as the $\eta$ coverage of every module and submodule is described in the Table 4.4.

## 4.6 Muon detectors

Most of the muons produced as a result of the p-p collisions are able to penetrate through the calorimeters and make it to the muon detectors where their tracks are getting measured. The spectrometer provides high-precision measurement of the muon momenta in the rapidity range of
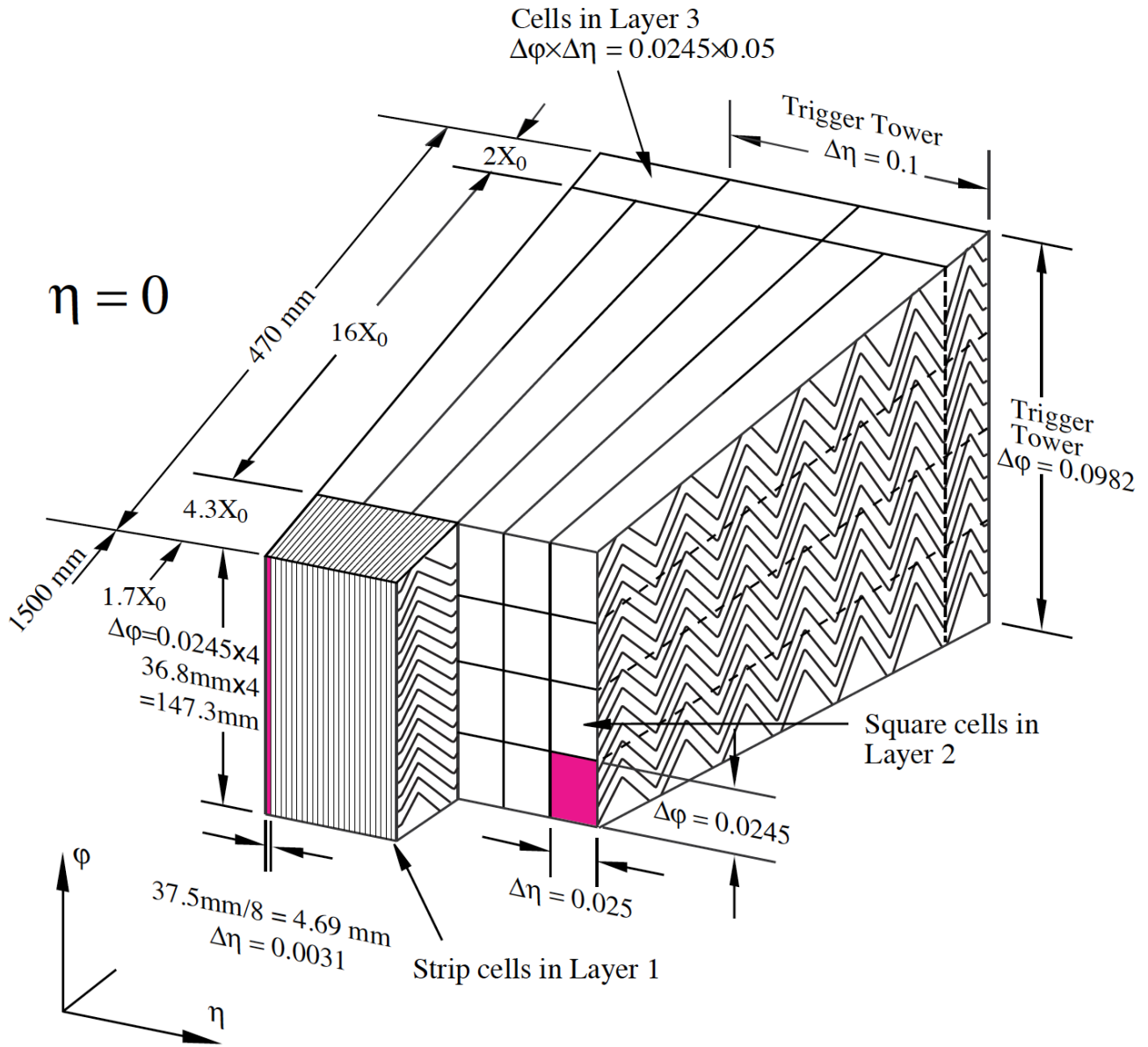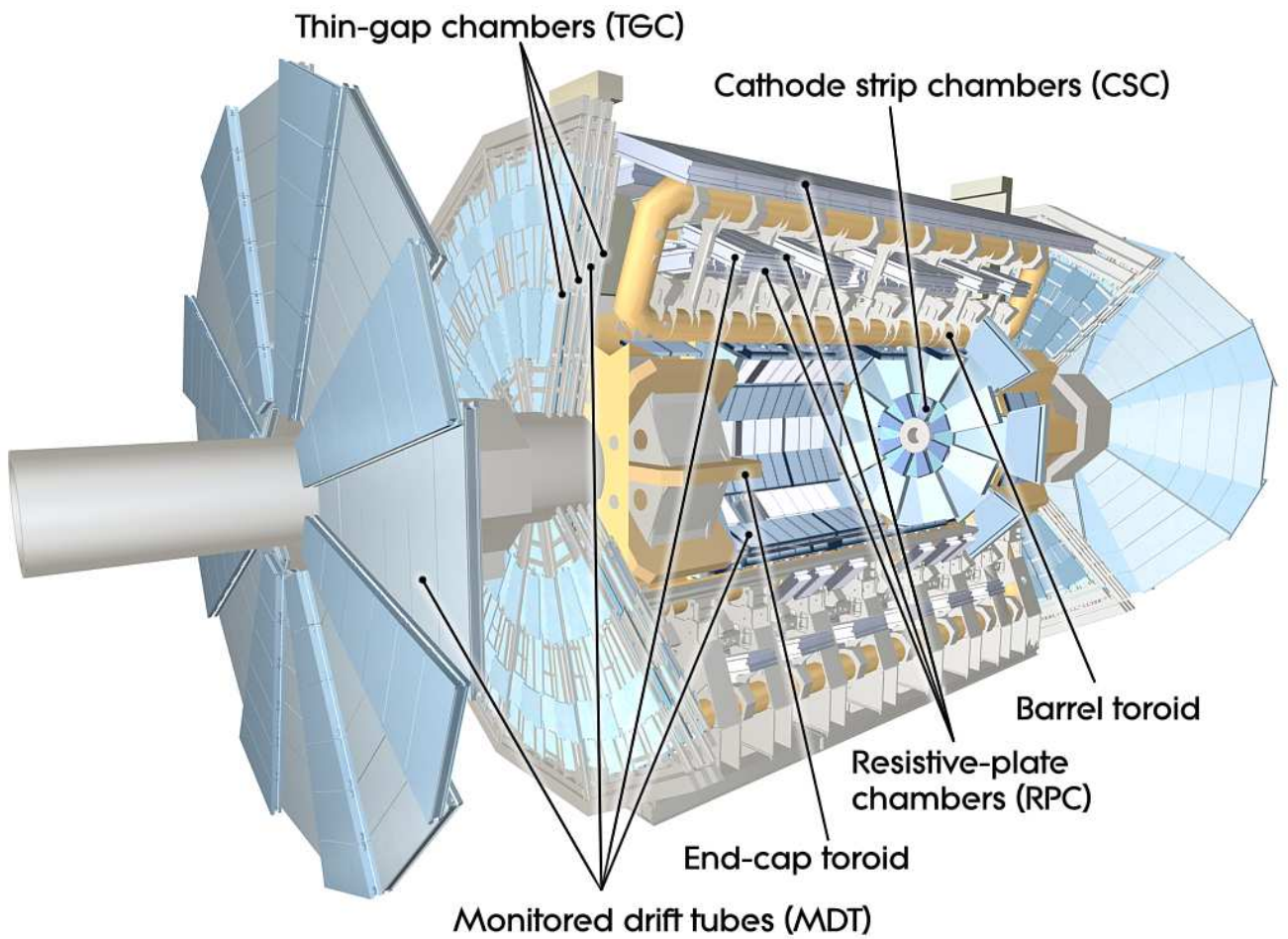
**Figure 4.5:** ATLAS EM calorimeter layers

(a) Barrel



(b) End-cap

**Figure 4.6:** Accordion absorbers of the EMC



**Figure 4.7:** ATLAS muon system

$|\eta| < 2.7$ and approximate transverse momentum range of 3 GeV $< p_T <$ 3 TeV. The lower bound on the momentum is mainly due to energy losses in the calorimeter, while the upper bound is caused by the sagitta bias coming from the tracking chambers alignment. The goal $p_T$ resolution is about 10% for a 1 TeV muon track.

The muon tracks[104],[105] are bent by the toroid magnets allowing to determine muon kinematic properties. The large barrel toroid covers the rapidity range of $|\eta| < 1.4$, while at $1.6 < |\eta| < 2.7$ the tracks are bent by the smaller end-cap magnets. The deflection in the transition region of $1.4 < |\eta| < 1.6$ is provided by the barrel and end-cap fields combined.

The general layout of the muon spectrometer is depicted on fig. 4.7, the parameters of the muon systems can be found in table 4.8. Just like the rest of the detector systems the muon spectrometer is split into the barrel and the end-cap parts.

The muon spectrometer possesses a fast triggering system able to trigger for muons in the rapidity range of $|\eta| < 2.4$. It delivers the track information within a few tens of nanoseconds after the particle passage which also allows to use it for the bunch-crossing identification. The trigger chambers measure both $\eta$ and $\phi$ coordinates of a track of which the former is in the bending plane and the latter is in the non-bending plane.

There are two types of fast triggering detectors used in the muon spectrometer:

- The Resistive Plate Chambers (RPCs) is a gaseous electrode-plate detector filled with a $C_2H_2F_4/Iso-C_4H_{10}/SF_6$ gas mixture (94.7/5/0.3). Two resistive plates of phenolic-melaminic plastic laminate are separated by insulating spacers of 2 mm thickness. The plates contain an electric field of about 4.9 kV/mm such that the ionizing tracks cause avalanches towards the anode. The signal is read out through the capacitive coupling of metallic strips, mounted to the resistive plates. The RPCs have nominal operating voltage of 9.8 kV and provide an excellent time resolution of a few ns with a supported local rate capability of 1000 $Hz/cm^2$

- Thin Gap Chambers (TGCs) are multi-wire proportional chambers with the wire-to-cathode distance of 1.4 mm and wire-to-wire distance of 1.8 mm and wire potential of 2900 V. The 2.8-mm gas gap is filled with highly quenching gas mixture of $CO_2$ and $n-C_5H_{12}$ (55/45). Small distance between the wires allows a very good time resolution of <25 ns in 99% of cases .

The precision-tracking chambers measure the coordinate of a track in the bending plane which is then matched with the second coordinate, measured by the trigger chamber.

There are two types of precision tracking systems used:

- The Monitored Drift Tubes (MDTs) are pressurised drift tubes with a diameter of 29.970 mm filled with $Ar/CO_2$ at 3 bar. Once the muon penetrates the tube it ionises the gas and the ionised electrons are collected at the central tungsten-renium wire of 50 $\mu$m in diameter and at a potential of 3080 V. This type of design carries several advantages: mechanical stiffness hence the alignment precision, reliability coming from the fact that a failure of a single tube would

not cause malfunction of the others. MDTs counting rate is limited to 150 Hz/$cm^2$ which is not sufficient for the innermost layer in the forward region of $2.0 < |\eta| < 2.7$.

- Cathode Strip Chambers (CSCs) are gas detectors filled with $Ar/CO_2$ in 80/20 proportion. The ionised electrons are collected at the wires which are oriented in the radial direction and operate at a potential of 1900 V. They are installed in the so-called Small Wheels and there are 16 CSCs on either side of the ATLAS detector. . The CSCs are able to provide a counting rate of 1000 Hz/$cm^2$ which makes it a reasonable replacement for the MDTs in the region close to the beam.

The precision-tracking chambers in the barrel are positioned between and on the coils of the super-conducting barrel toroid magnet. They form three concentric cylindrical shells around the beam axis at the approximate radii of 5 m, 7.5 m and 10 m. In the barrel region the RPCs were chosen for the fast triggering whereas the MDTs provide the precision tracking. The end-cap muon spectrometer is assembled in the form of large wheels perpendicular to the beam axis and located at distances about 7.4 m, 10.8 m, 14 m and 21.5 m from the interaction point. The triggering in the end-cap is provided by the TGCs. Most of the precision tracking chambers are the MDTs similarly to the barrel, except for the forward region of $2.0 < |\eta| < 2.7$ where the CSCs are installed in the innermost tracking layer. The reason for that is their higher resistance to radiation and increased particle flow which becomes an issue if you get closer to the beam.

Barrel and end-cap alignment is illustrated on Fig. 4.8which contains the side and transverse views of the muon spectrometer.
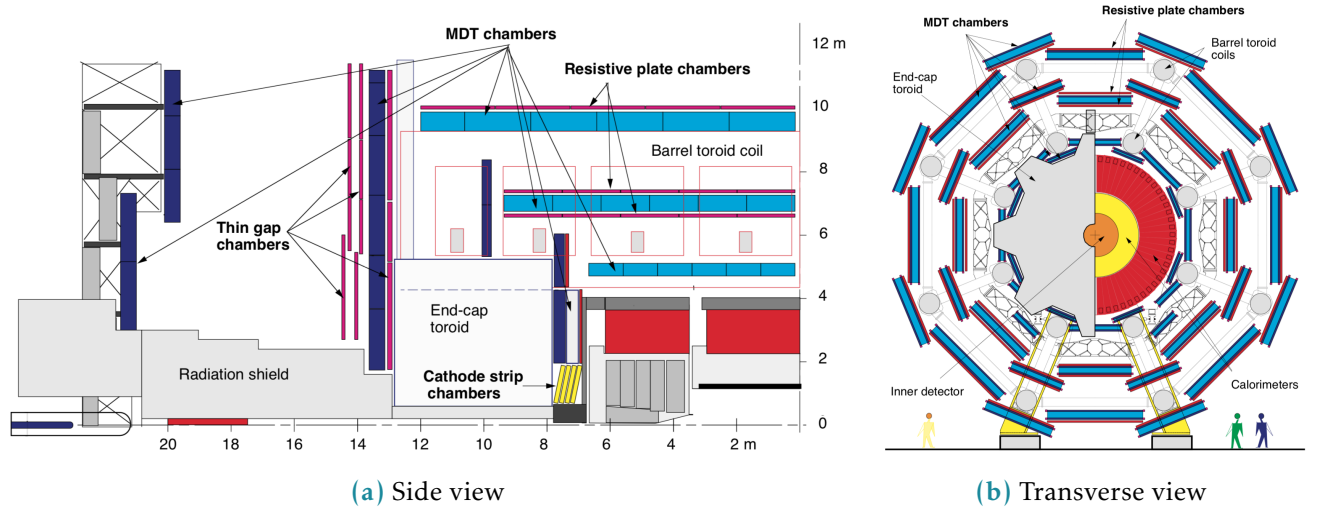


(a) Side view      (b) Transverse view

**Figure 4.8:** Cut views of the muon systems

| **Monitored drift tubes** | **MDT** |
|---|---|
| Coverage | $\lvert\eta\rvert < 2.7$ (innermost layer: $\lvert\eta\rvert < 2.0$) |
| Number of chambers | 1088 (1050) |
| Number of channels | 339 000 (354 000) |
| Function | Precision tracking |
| **Cathode strip chambers** | **CSC** |
| Coverage | $2.0 < \lvert\eta\rvert < 2.7$ |
| Number of chambers | 32 |
| Number of channels | 31 000 |
| Function | Precision tracking |
| **Resistive plate chambers** | **RPC** |
| Coverage | $\lvert\eta\rvert < 1.05$ |
| Number of chambers | 544 (606) |
| Number of channels | 359 000 (373 000) |
| Function | Triggering, second coordinate |
| **Thin gap chambers** | **TGC** |
| Coverage | $1.05 < \lvert\eta\rvert < 2.7$ |
| Number of chambers | 3588 |
| Number of channels | 318 000 |
| Function | Triggering, second coordinate |

**Table 4.8:** ATLAS muon spectrometer subsystems coverage and parameters

## 4.7 Forward detectors

There are three detector systems that cover the ATLAS forward region (see Fig. 4.9): LUminosity measurement using Cherenkov Integrating Detector (LUCID), Absolute Luminosity for ATLAS (ALFA) and Zero-Degree Calorimeter (ZDC). The measurement of luminosity is the main goal of the first two detectors and has fundamental importance: it provides the normalization scale for all the observed processes.

LUCID[106], [107] is the main ATLAS relative luminosity monitor. The main purpose of the LUCID detector is to detect elastic p-p scattering in the forward region measuring the integrated luminosity and performing online monitoring of the instantaneous luminosity and beam conditions with uncertainty of about few percent. It is symmetrically installed at ±17 m from the interaction point and at a radial distance of about 10 cm from the beam line (resulting in $\lvert\eta\rvert \approx 5.8$). On each side four bundles of quartz fibres are used as a medium producing Cherenkov radiation directing the Cherenkov light into the 16 Photomultiplicator Tubes (PMTs) placed outside the radiation shielding.

The ALFA[108] detector is used to measure the absolute luminosity through elastic scattering at small angles. The design of the ALFA detector allows to precisely measure the track of scattered protons at millimetre distance from the beams. This is achieved by placing detectors in Roman Pots i.e. thin walled vessels which allow to operate the detectors inside the LHC beam pipe [109]. In order to perform

such measurement we need to meet the following conditions:

- The beam has to be more parallel than normally. Special collider beam optics allowing high values of the amplitude function at the interaction point $\beta*$ together with reduced beam emittance.

- To be sensitive to small angles the detectors have to be placed as far as possible from the interaction point and close to the beam. This is why the detectors are located inside the Roman Pots at ±240 m from the interaction point. On each side there are two Roman Pots separated by four meters.

The Roman pot windows allow the elastically scattered protons to reach the square scintillating fibres of 0.5 mm width which are in turn connected to multi-anode PMTs through the light-guides. The detector provides a spatial resolution of 30 $\mu$m and allows to measure the absolute luminosity an uncertainty of 1.7% for the Run 2[110].

ZDC are used to detect forward neutrons at $|\eta| > 8.3$ in heavy-ion collisions, which in turn allows to determine the centrality of such collisions. The detector is installed at ±140 m from the interaction point. Every ZDC arm consists of 4 modules: one electromagnetic and three hadronic. These modules are quartz rods shielded by the tungsten plates and connected to the PMTs via the light-guides allowing to measure incending particle energy and position. The EM module has a better position resolution mapping each of 96 quartz rods into a single pixel, while the hadronic modules map a bundle of four rods into a pixel. Only one of the three hadronic modules per arm provide position-sensing rods and only the arm at -140 m has the position-sensing EM module.
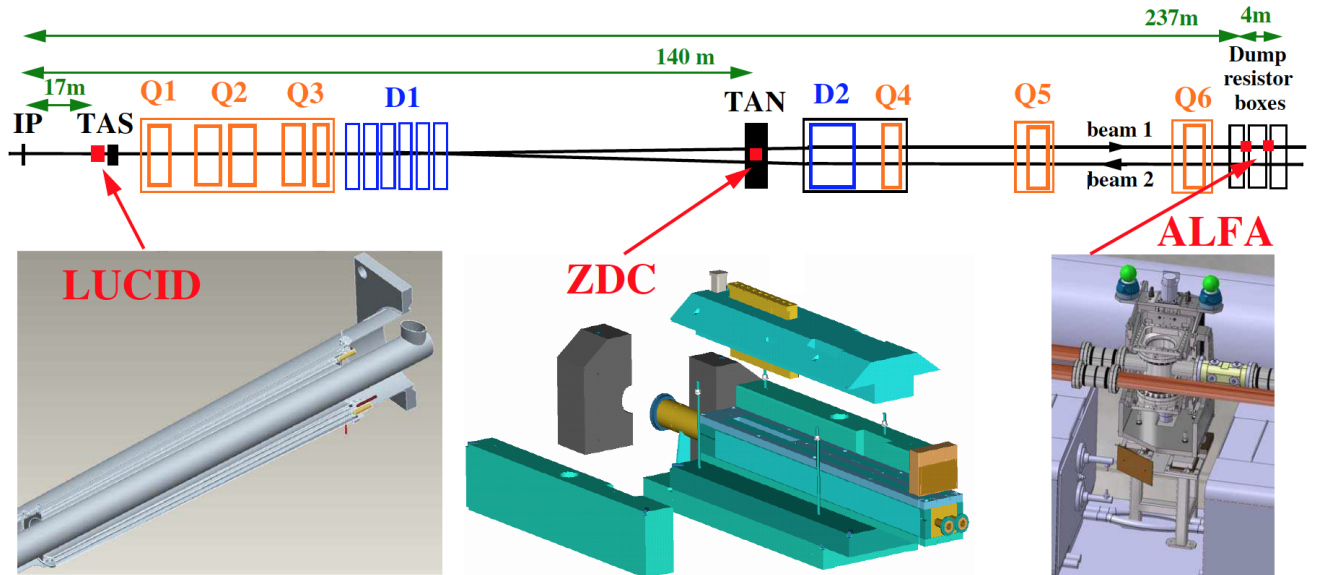


**Figure 4.9:** ATLAS forward detectors

## 4.8   Trigger system

Considering that the bunch crossing rate at LHC is about 40 MHz and that ATLAS detector has over one million read-out channels it would never be possible to store all the raw data without significant preselection that would decrease the data rate. The selection criteria are picked to retain and store only the events which might be interesting for the LHC physics. The preselection and storage is conducted with the help of TDAQ systems.

The trigger system has three distinct levels: L1, L2 and the event filter, the two latter levels are also called High-level Trigger (HLT). Each next level refines the decisions made before and, if necessary, applies additional selection, further lowering the event rate. The data acquisition system receives and buffers the event data from the readout electronics at the L1 trigger accept rate which for Run 2 is about 100 kHz [111]. The HLT then lowers the rate down to 1.5 kHz which is then stored for the offline analysis.

The L1 trigger looks for muons, electrons, photons and hadrons from $\tau$-lepton decays with high transverse momentum, large missing and total transverse energy. The muons of interest are identified using the muon spectrometer trigger system described in section 4.6. The rest of the particles are selected using the information from all the calorimeters with reduced granularity. During the Run 2 an intermediate L1Topo trigger was also added allowing to combine the information from the spectrometer and calorimeter and extend possible trigger selections. Results from these triggers get processed by the central trigger processor which implements the trigger menu made up of different combinations of trigger selections. The decision latency for the L1 trigger must not exceed 2.5 $\mu$s after the corresponding bunch crossing.

For every selected event the L1 defines one or more regions called Region of Interest (RoI) which include the $\eta$ and $\phi$ coordinates of these regions for their subsequent use by the HLT. The L2 selection is seeded by RoI and uses full granularity and precision along with other detector data available. The trigger block diagram is presented in fig. 4.10.
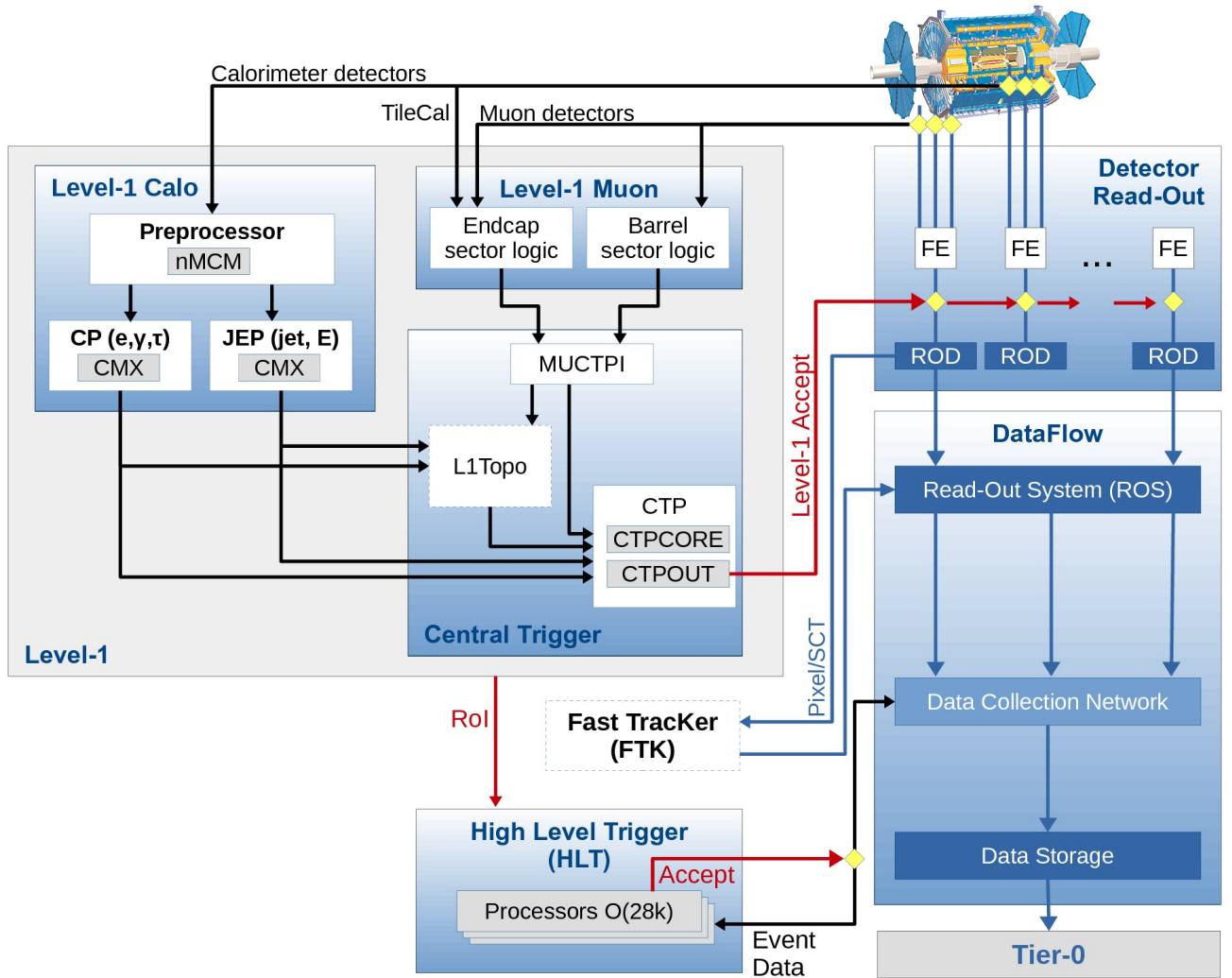
**Figure 4.10:** The scheme of ATLAS trigger systems

# Electromagnetic shower shapes correction 5

The chapter considers the electromagnetic shower development in the ATLAS EMC and its role in particle identification (ID). The existing mismodelling of the shower development in the Monte-Carlo simulation causes discrepancies in electron ID. A correction method that allows to achieve good correspondence between the data and the simulation is proposed, implemented and tested.
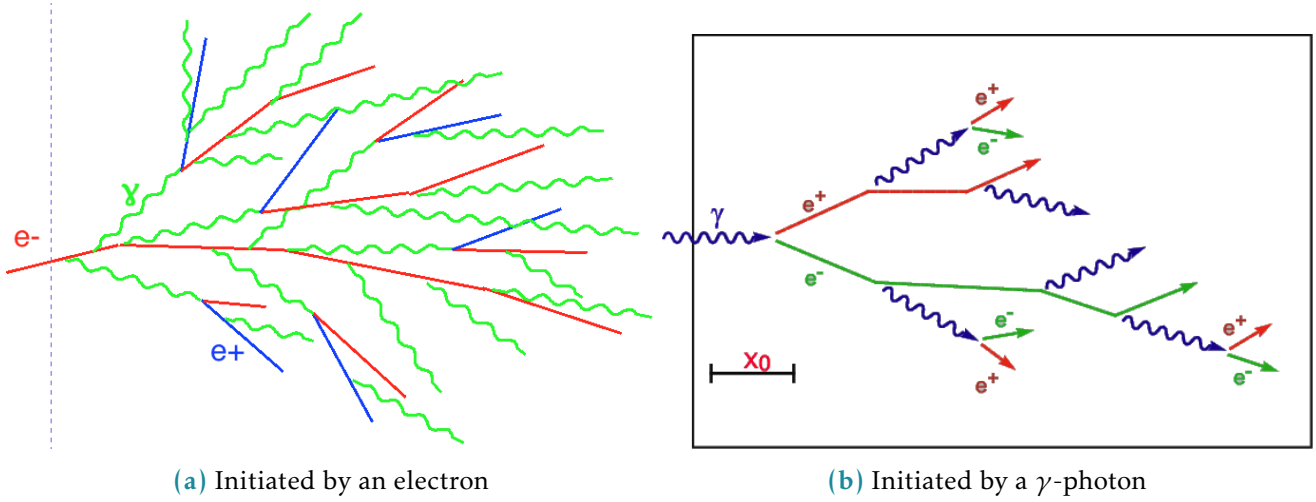
## 5.1  Introduction

The design and functionality of the ATLAS electromagnetic calorimeter was described in 4.5.1. Let's consider a bit more in detail the physical processes happening in the EMC.

It order to measure particle's energy within the calorimeter we must make the particle to loose its entire energy within the calorimeter. For the electrons and photons with energies over few MeV (which is the case for the ATLAS experiment) the primary energy loss mechanism lies in bremsstrahlung radiation and pair creation. The two processes complete each other, so when a high-energy electron or photon gets into the calorimeter, it creates an avalanche-like process called the electromagnetic shower when a bremsstrahlung-radiated photons create more electron-positron pairs which in turn radiate more bremsstrahlung photons and so on and so forth (see Fig. 5.1.)
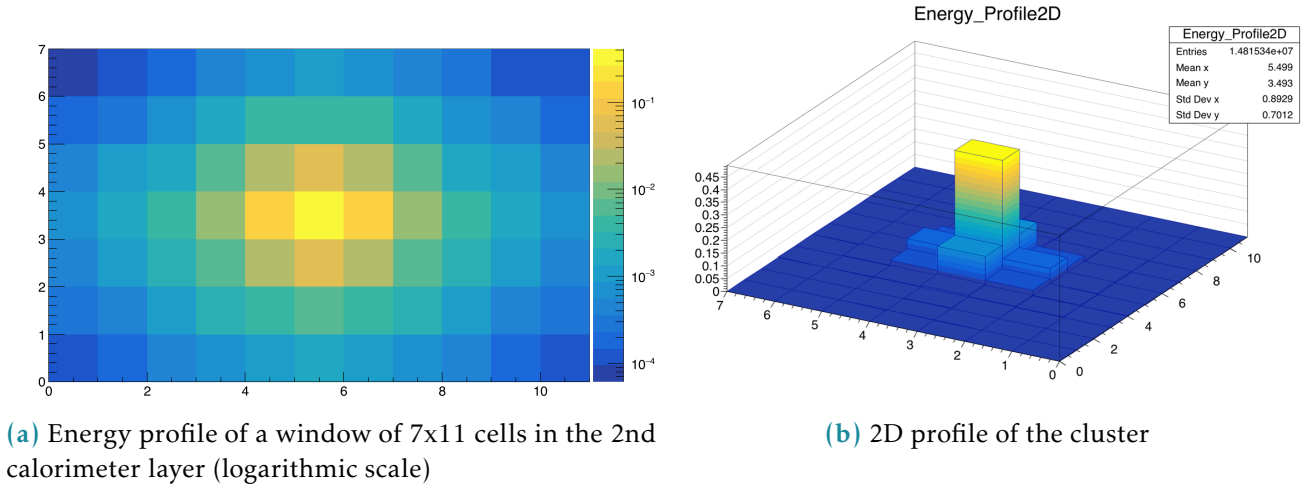 The longitudinal and transverse development of the shower depends on the type of the initial particle and on its energy. The energy is well measured by the calorimeter, but identifying the particle still remains a challenging task. The transverse granularity of the ATLAS calorimeter allows to resolve the energy distribution within the electromagnetic shower in the transverse plane. This information can later be used for particle identification.

When an e/$\gamma$ particle hits the calorimeter its footprint in the second layer of the calorimeter is visible as a cluster of calorimeter cells centered at the central cell having the most energy deposited (sometimes referred to as "the hottest cell"). Roughly 90% of shower energy is contained in the core 3x3 cells. We have considered a cluster of 7x11 ($\eta$ x $\phi$) cells, which is schematically depicted on Fig. 5.2a.

In order to characterise the energy distribution within the shower profile a number of observables called shower shapes are used. They are then used as an input for particle identification MVA algorithm. Current study focuses on the second layer of the calorimeter for which there are three shower shape

(a) Initiated by an electron



(b) Initiated by a $\gamma$-photon

**Figure 5.1:** The schematic portrayal of EM shower development



(a) Energy profile of a window of 7x11 cells in the 2nd calorimeter layer (logarithmic scale)



(b) 2D profile of the cluster

**Figure 5.2:** Visualisations of the 7x11 calorimeter cluster

observables described below [112]:

- Lateral shower width $W_{\eta 2} = \sqrt{\sum(E_i \eta_i^2) - (\sum(E_i \eta_i)/\sum(E_i))^2}$ calculated within a window of 3x5 cells.

- $R_\phi$ - ratio of the energy in 3x3 cells over the energy in 3x7 cells centered around the hottest cell.

- $R_\eta$ - ratio of the energy in 3x7 cells over the energy in 7x7 cells centered around the hottest cell.

The shower shapes distributions for different types of particles is shown in Fig. 5.4 - although the distributions overlap, combining the shower shapes information with the inputs from other detectors allow to identify the particle.

(a) Lateral shower width $W_\eta 2$

(b) $R_\phi$ and $R_\eta$
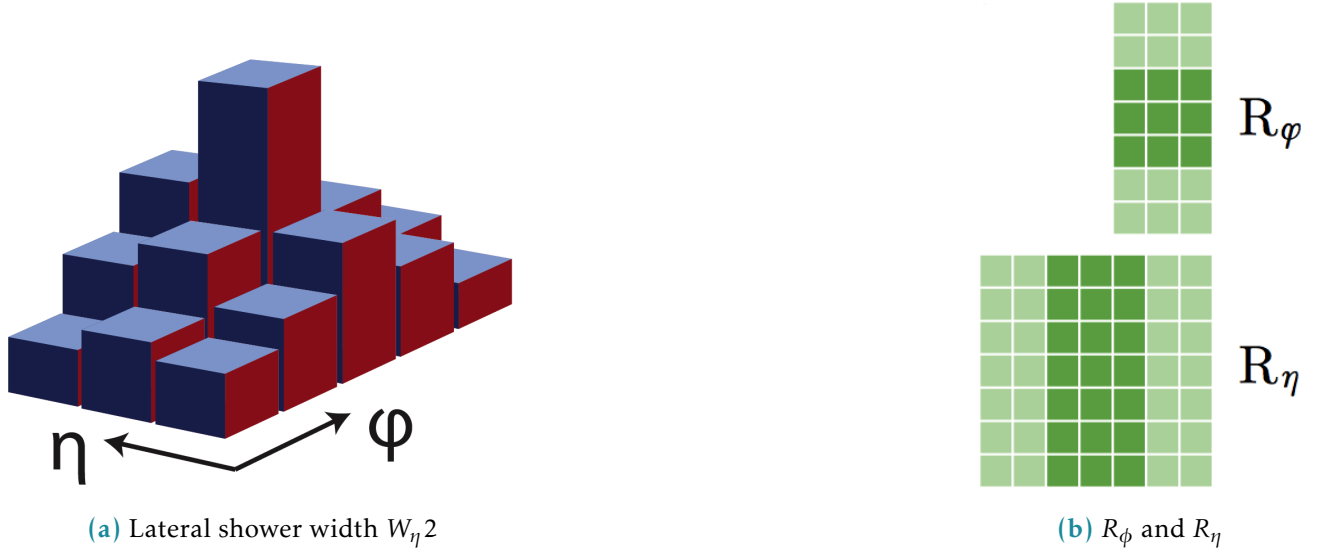
**Figure 5.3:** Shower shapes in the second layer of the electromagnetic calorimeter



(a) $W_{\eta 2}$ distribution simulation

(b) $R_\eta$ distribution simulation
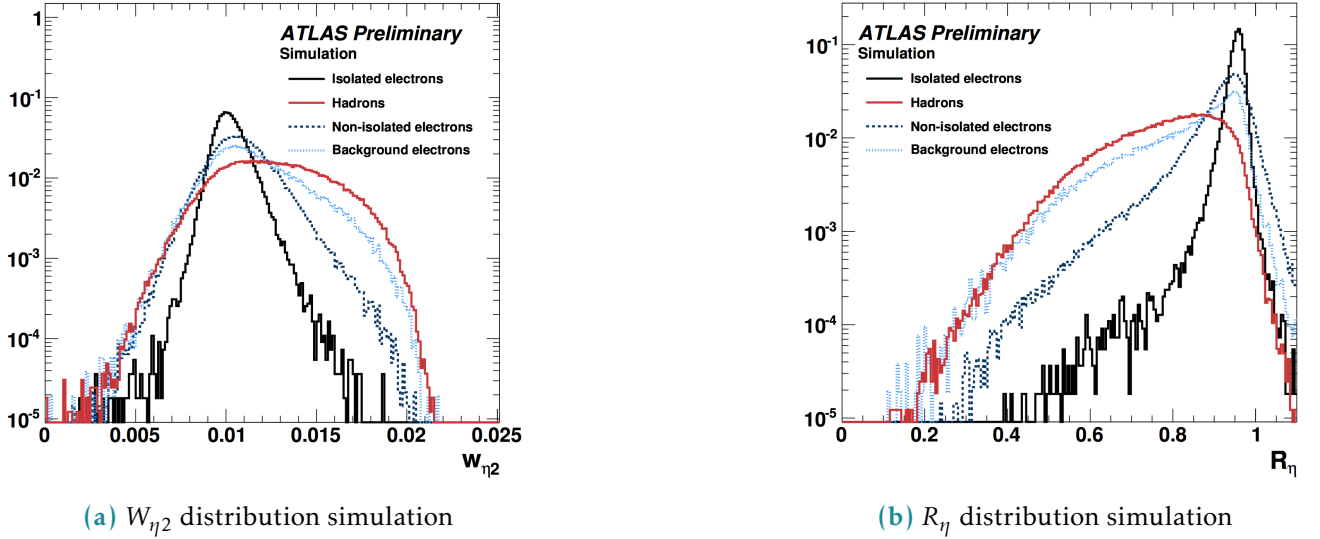
**Figure 5.4:** Distribution of $R_\eta$ in simulation (GEANT4) for electrons and jets [113].

Figure 5.4 shows how $R_\eta$ distribution is different in jets, signal electrons and background electrons. Background electrons denote non-prompt electrons which are not originated from primary vertex.

The shower shapes appear to be extremely sensitive to the detector material modelling. A simplification in the geometry of the EMCal absorber geometry in GEANT4 9.2 (a layered structure of the accordion was represented as a homogeneous material) has lead to visible discrepancies in the shower shapes between the data and MC. This was corrected in GEANT4 9.4 significantly improving the agreement, although not eliminating it completely (see Fig. 5.5). The origin for the remaining discrepancy is not clear.



(a) $W_{\eta 2}$

(b) $R_\eta$
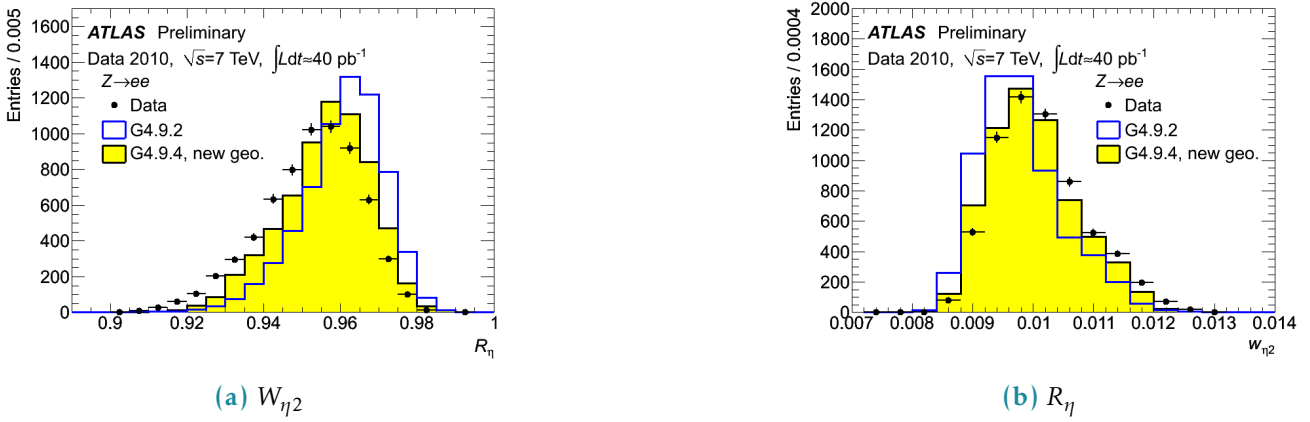
**Figure 5.5:** Data/MC Comparison for Calorimeter Shower Shapes of High Et Electrons [114].

Disagreement in shower shapes between the data and MC leads to discrepancies in particle ID which are later fixed using $\eta-$ and $p_T$-dependent scale factors. Correction of the shower shapes aims to get the scale factors closer to unity, reducing the corresponding systematic uncertainties and improving the precision of the measurements with electrons in the final states.

## 5.2 Shower shapes measurement and correction

### 5.2.1 Event selection

For this study we have considered electrons from the $Z \rightarrow ee$ decay. A set of recommended single electron triggers was used (HLT_e26_lhtight_nod0_ivarloose,
HLT_e60_lhmedium_nod0,
HLT_e140_lhloose_nod0,HLT_e300_etcut). Each event was required to have 2 electrons at least one of which has $p_T > 25$ GeV. In order to suppress the background both electrons had to pass gradient isolation. Z invariant mass cut was applied with a window of $80 - 120$GeV. To avoid identification bias

from triggering the tag and probe approach was used with only probe electrons taken into consideration [115]. The electron cluster in the second calorimeter layer was required to contain information from 77 calorimeter cells. No pile-up reweighting has been applied. Datasets of 264786295 events in data (2017 proton-proton collisions) and 79340000 events in MC (Powheg+Pythia8) were used.

### 5.2.2 Data/MC discrepancies

Our consideration begins with the energy deposit of an electron in the second layer of the calorimeter. A window of 7 cells in $\eta$ and 11 cells in $\phi$ is centred around the cell with the highest energy.

Shower shapes were considered in 14 $\eta$ bins in the range between $|\eta| = (0, 2.4)$ in order to investigate how the discrepancy depends on $\eta$.



(a) $R_\eta$ in $|\eta| = (0.4, 0.6)$      (b) $R_\eta$ in $|\eta| = (1.8, 2.0)$
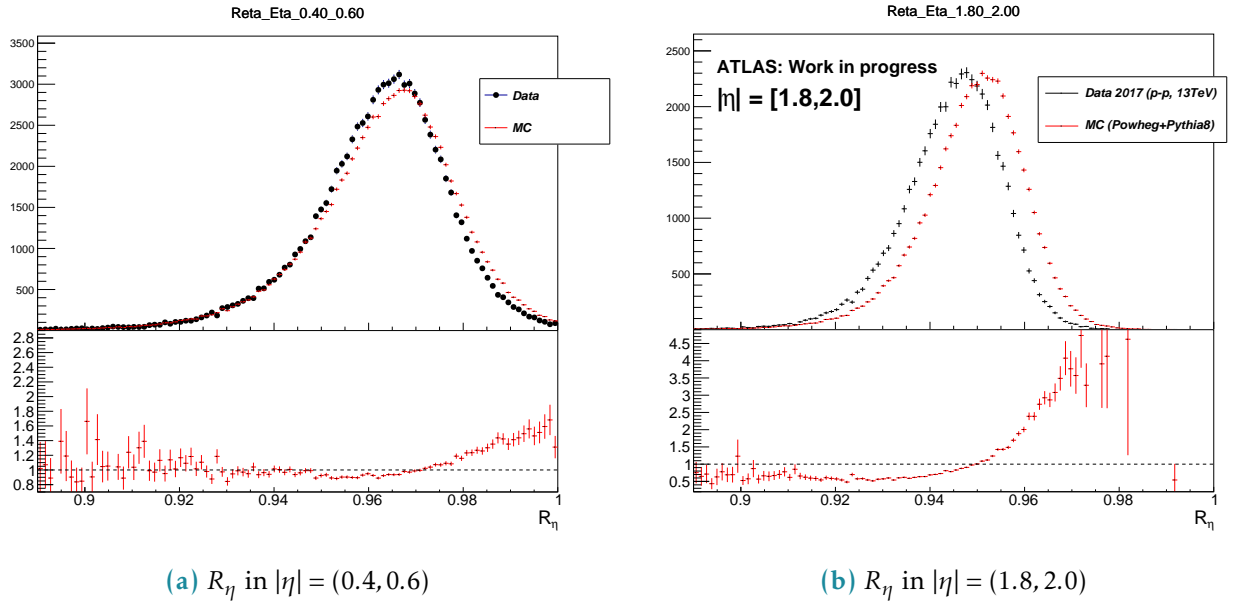
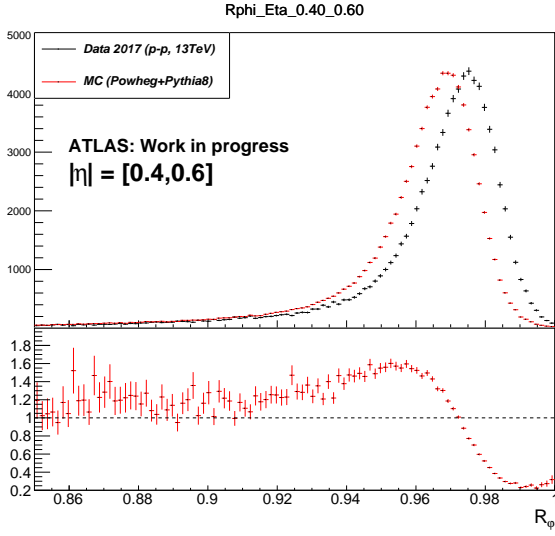**Figure 5.6:** $R_\eta$ in the barel and in the end-cap, Data vs MC

The $\eta$-dependent shower shapes in data are wider than the MC and show a larger discrepancy in the endcap ($|\eta| = (1.52, 2.4)$). For $\phi$ dimension the situation is the opposite: MC is wider than the data and the barrel ($|\eta| = (0, 1.52)$) shows larger discrepancy. Figures 5.6, 5.7, 5.8 contain examples of shower shapes in different eta bins.
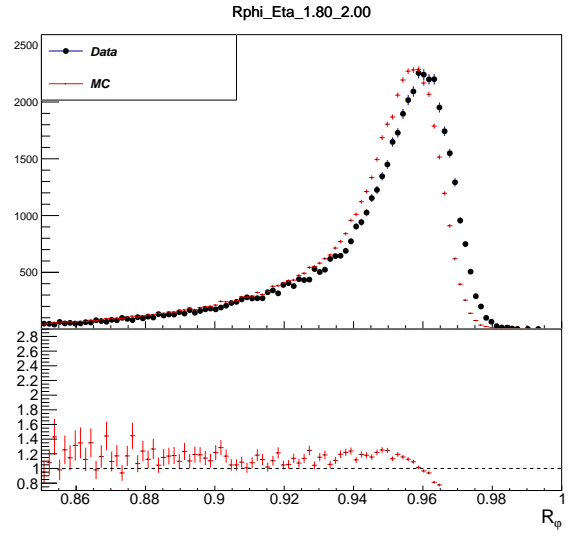
### 5.2.3 The correction procedure

#### The correction matrix

The correction procedure is based on the redistribution of energy between the cluster cells in MC so that the distribution becomes consistent with the data. For every $\eta$ bin a correction matrix is derived
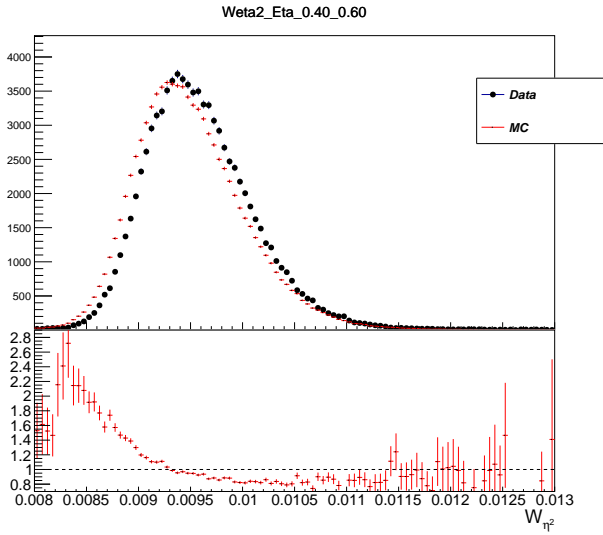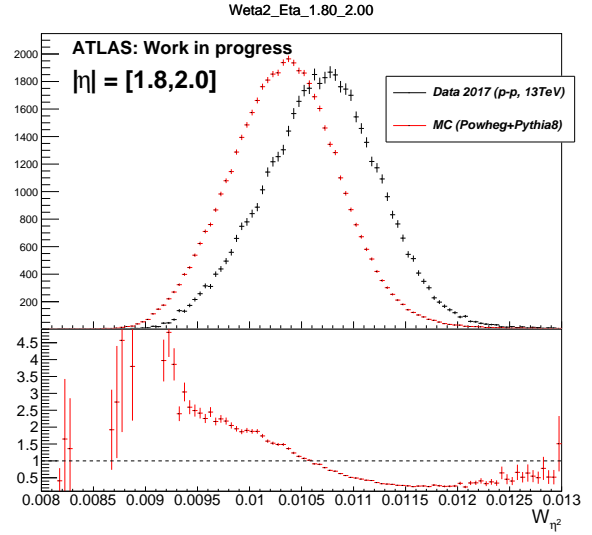
(a) $R_\phi$ in $|\eta| = (0.4, 0.6)$

(b) $R_\phi$ in $|\eta| = (1.8, 2.0)$

**Figure 5.7:** $R_\phi$ in the barel and in the end-cap, Data vs MC



(a) $W_\eta^2$ in $|\eta| = (0.4, 0.6)$

(b) $W_\eta^2$ in $|\eta| = (1.8, 2.0)$

**Figure 5.8:** $W_\eta^2$ in the barel and in the end-cap, Data vs MC

in the following way:

$$M_i^{Correction} = \frac{E_i^{Data}}{\Sigma E^{Data}} - \frac{E_i^{MC}}{\Sigma E^{MC}}$$

$\Sigma_i M_i^{Correction} = 0$, $i = 1..77$.

$E_i^{Data}$, $E_i^{MC}$ - matrix elements of the averaged energy profiles. The correction is then applied to the electron cluster cells on event-by-event basis:

$$E_i^{Reweighted} = E_i^{Non-reweighted}(1 + M_i^{Correction}).$$

This redistributes the energy among the cells keeping the total energy exactly the same.
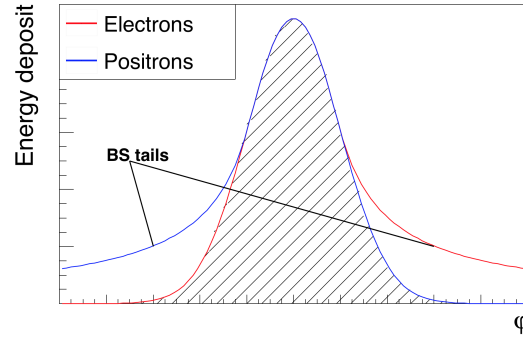
### Bremsstrahlung tails

The magnetic field directed along the $\phi$ dimension leads to a significant asymmetry in the energy deposits for electrons and positrons (Fig. 5.9).



(a) $R_\phi$ in $|\eta| = (0.4, 0.6)$    (b) $R_\phi$ in $|\eta| = (1.8, 2.0)$
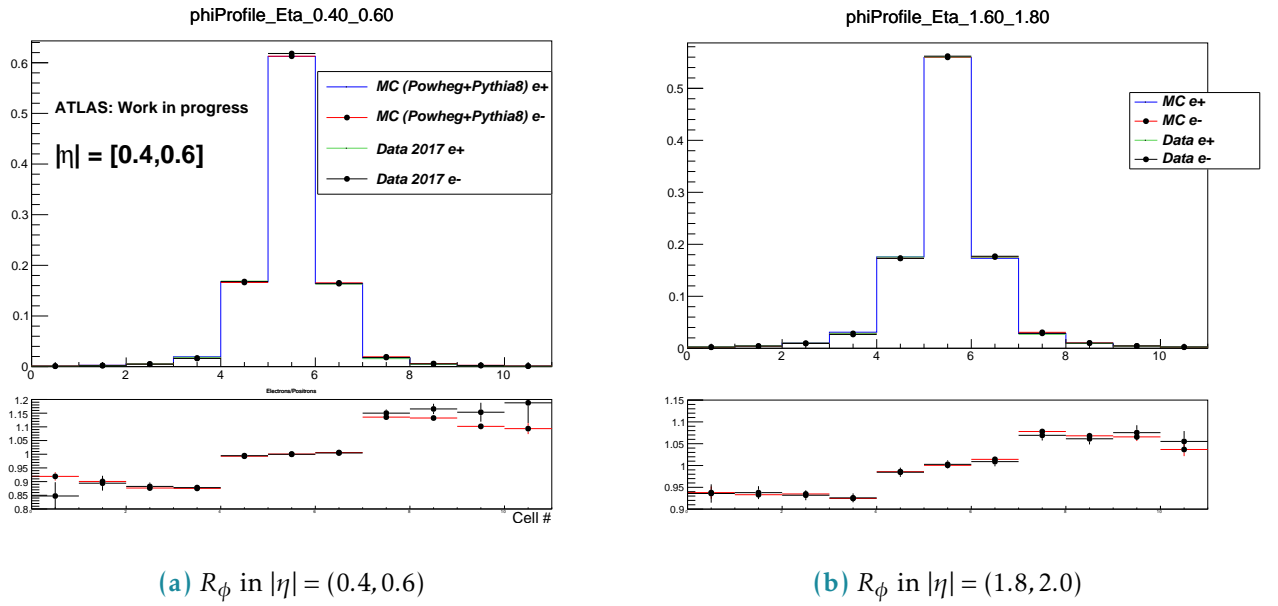
**Figure 5.9:** $R_\phi$ in the barrel and in the end-cap for $e^+$ and $e^-$ in Data and MC. The ratio panel shows $e^+/e^-$ energy deposits in Data (black) and MC (red).

Considering the fact that the reweighting is intended to correct for the data/MC discrepancies themselves and not for the bremsstrahlung effect it makes sense to develop the bremsstrahlung-free correction function based on $e^+$ and $e^-$ correction matrices. The principle is schematically explained on figure 5.10.

Good agreement of data and MC description of $e^+$ and $e^-$ asymmetry gives a hint that the material mismodelling cannot be the main source of the data/MC disagreement.

**Figure 5.10:** Schematic energy profile in $\phi$ dimension. Bremsstrahlung tails subtraction based on $e^+$ and $e^-$ energy profiles.

## 5.3   Results

Figs. 5.11, 5.12, 5.13 show the effect of the correction. The shower shapes in MC become very close to the data, correcting a significant discrepancy.

Fig. 5.14 contain shower shapes vs $p_T$ integrated over $\eta$. They demonstrate that the correction does not depend on the $p_T$ which allows to expect the decreased systematic uncertainties for $p_T$ regions distant from $40 - 50\text{GeV}$.

Finally, Fig. 5.15 shows the effect of the correction on electron ID efficiency. We can see a visible improvement, notably in the endcap region. Nevertheless the barrel region shows little improvement. It can be explained by the fact that electron ID MVA relies on many variables while only a number of them were corrected during current study.

The proposed method is getting integrated into ATLAS Athena framework as an option and is planned to be used as a baseline for Run 3.

## 5.4   Appendix: control plots

(a) Reweighted $R_\eta$ in $|\eta| = (0.4, 0.6)$.

(b) Reweighted $R_\eta$ in $|\eta| = (1.8, 2.0)$.
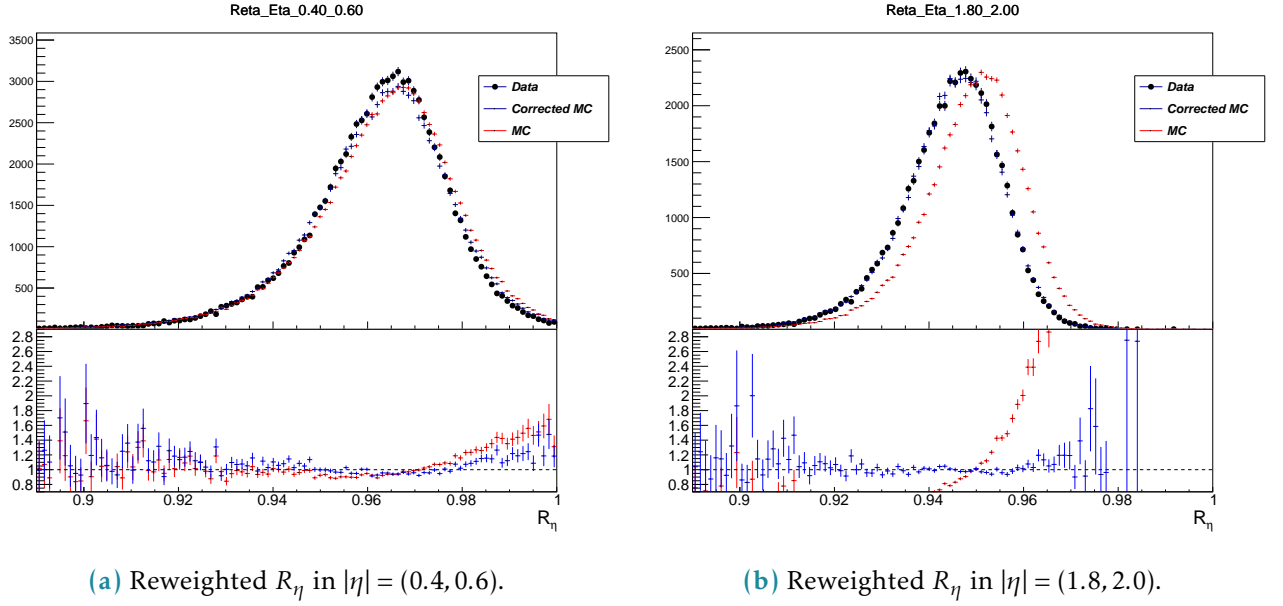
**Figure 5.11:** $R_\eta$ in the barrel and in the end-cap



(a) $R_\phi$ in $|\eta| = (0.4, 0.6)$

(b) $R_\phi$ in $|\eta| = (1.8, 2.0)$
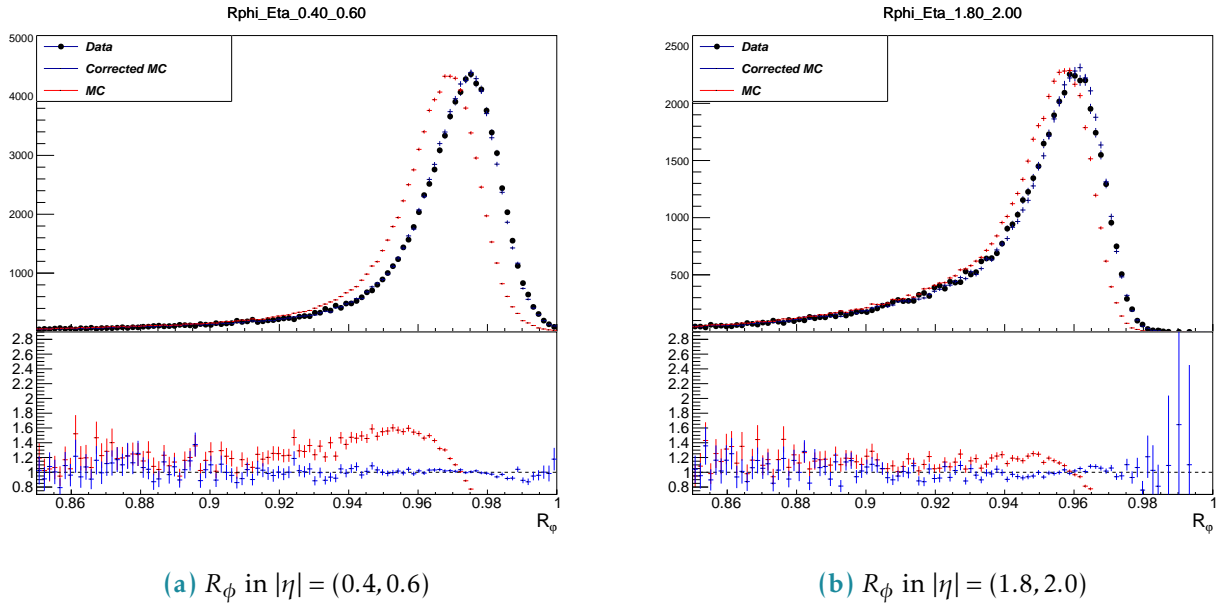
**Figure 5.12:** $R_\phi$ in the barel and in the end-cap, Data, MC, reweighted MC

(a) $W_\eta^2$ in $|\eta| = (0.4, 0.6)$

(b) $W_\eta^2$ in $|\eta| = (1.8, 2.0)$
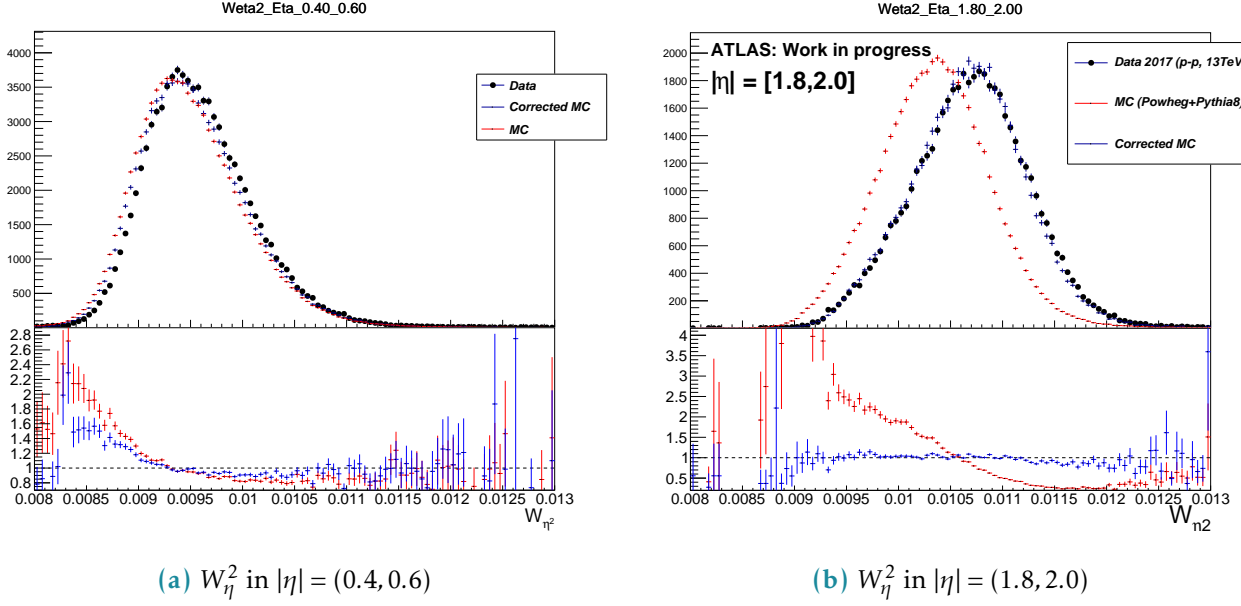
**Figure 5.13:** $W_\eta^2$ in the barel and in the end-cap, Data, MC, reweighted MC
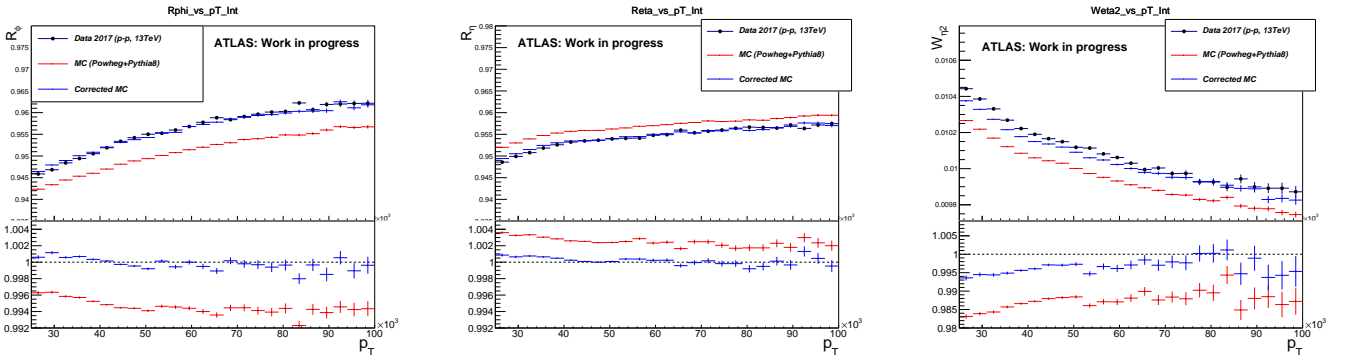


**Figure 5.14:** Distributions integrated over pT (a) $R_\phi$; (b) $R_\eta$; (c)$W_{\eta 2}$.
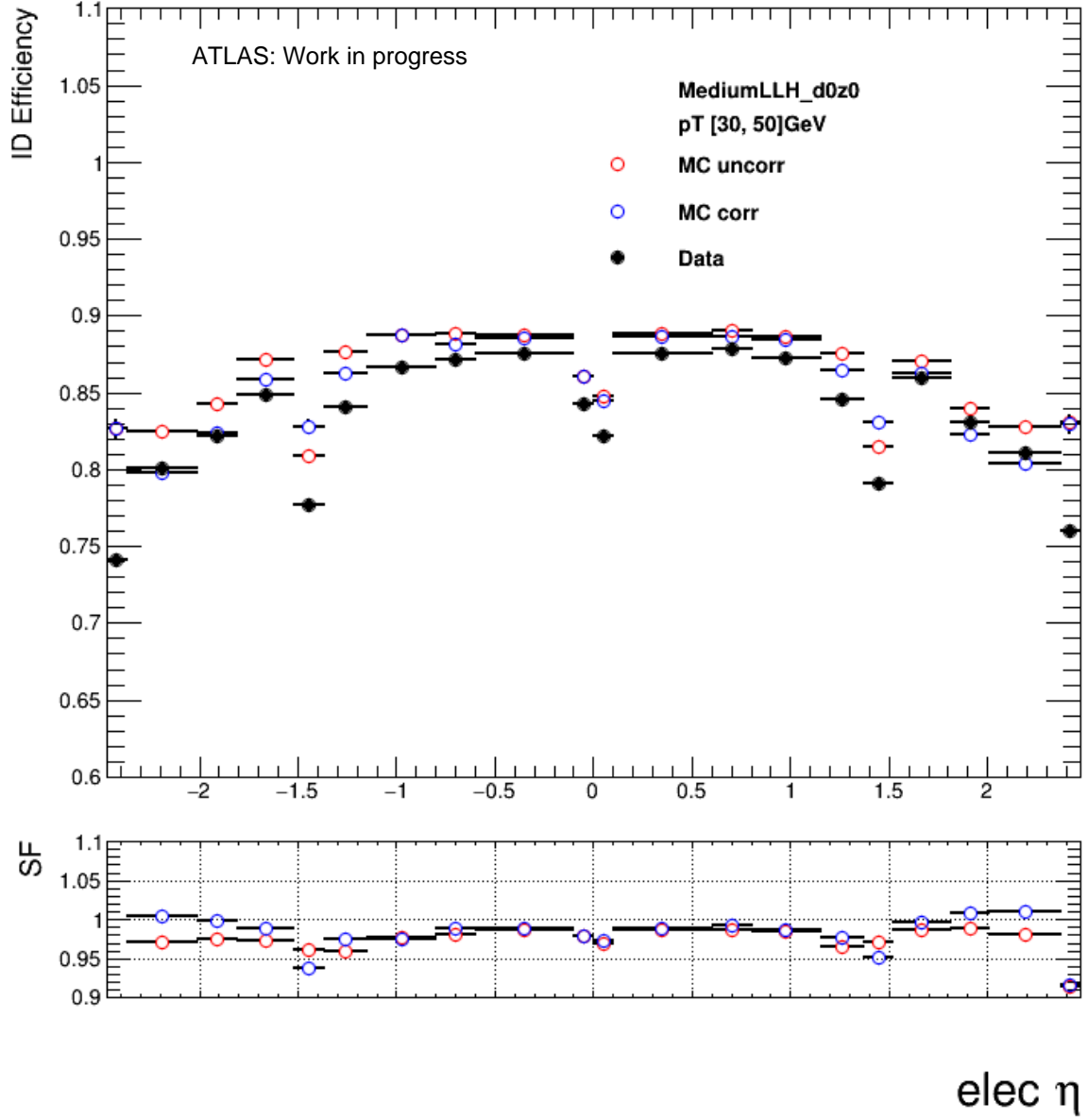
**Figure 5.15:** Electron identification efficiency as a function of the electron pseudo-rapidity
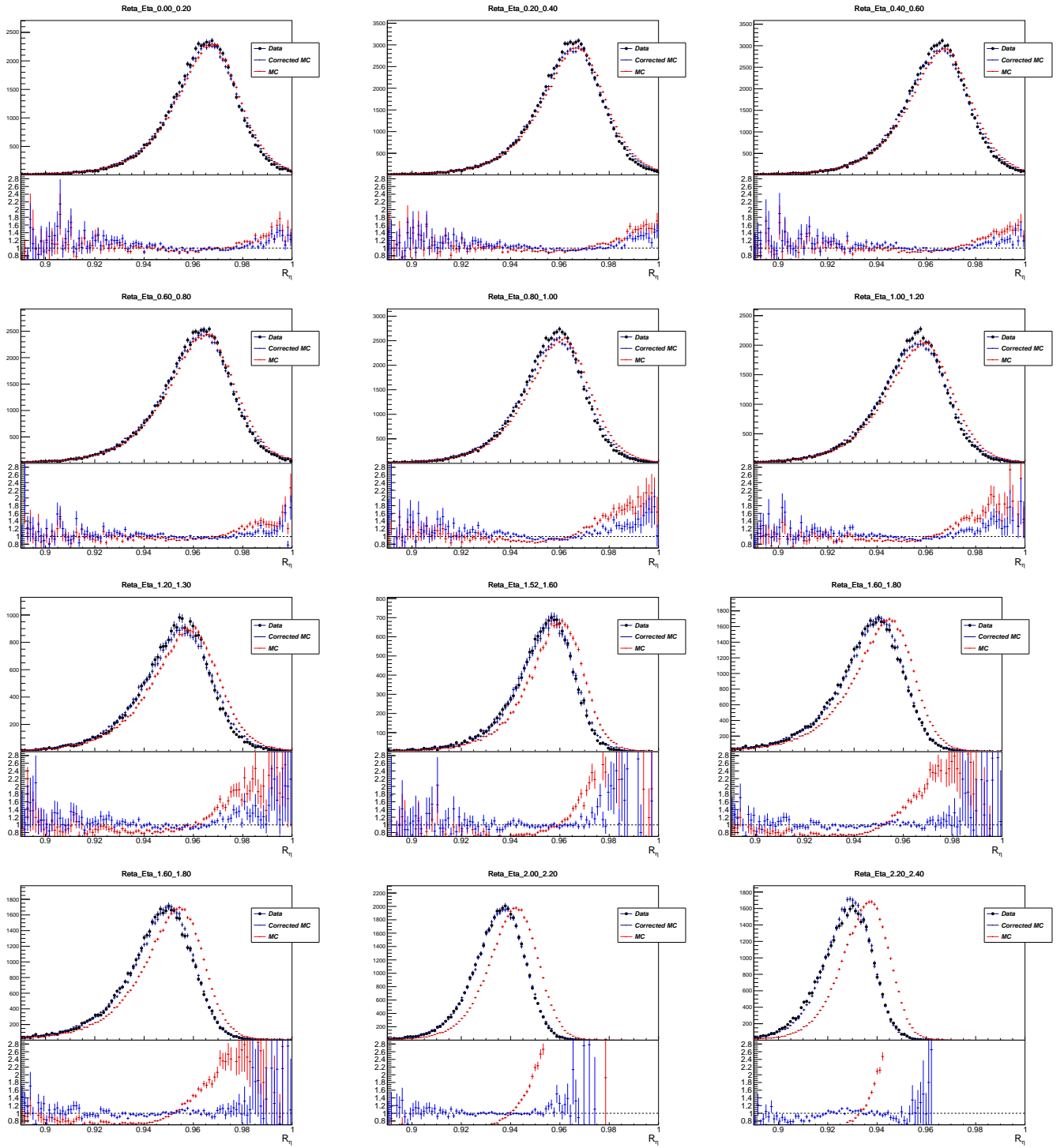
**Figure 5.16:** $R_\eta$ in all eta slices.
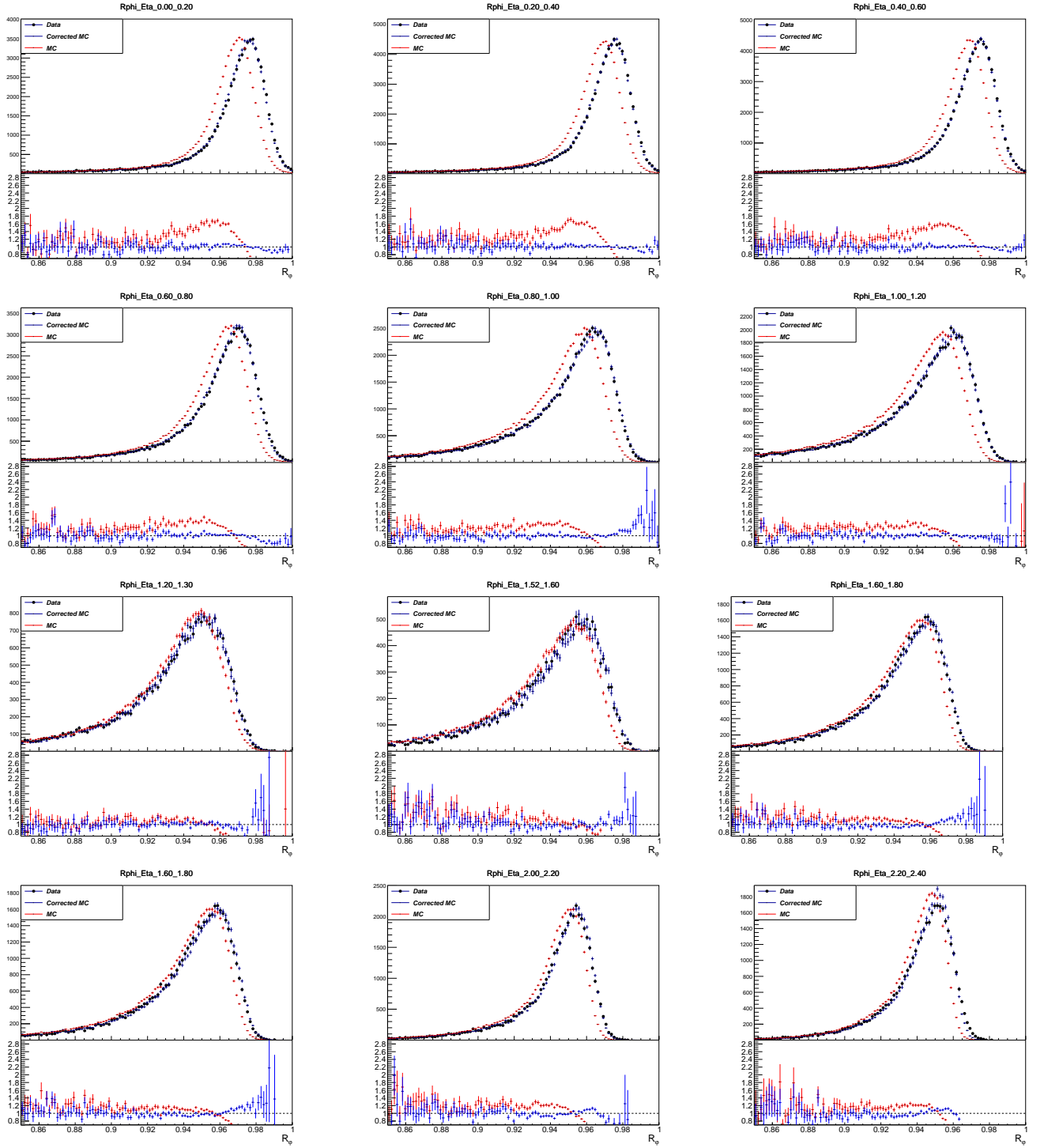
**Figure 5.17:** $R_\phi$ in all eta slices.
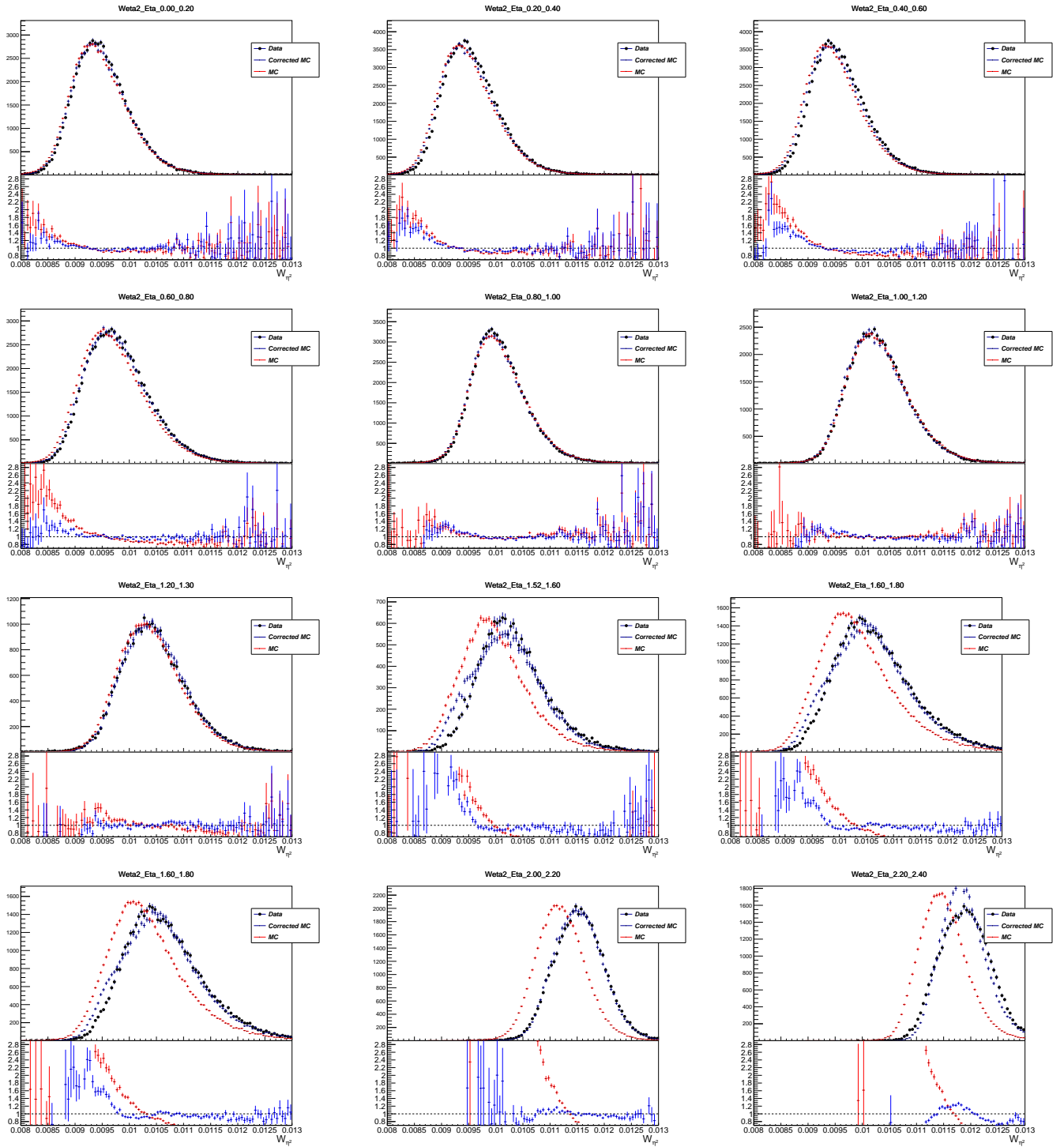
**Figure 5.18:** $W_{\eta 2}$ in all eta slices.
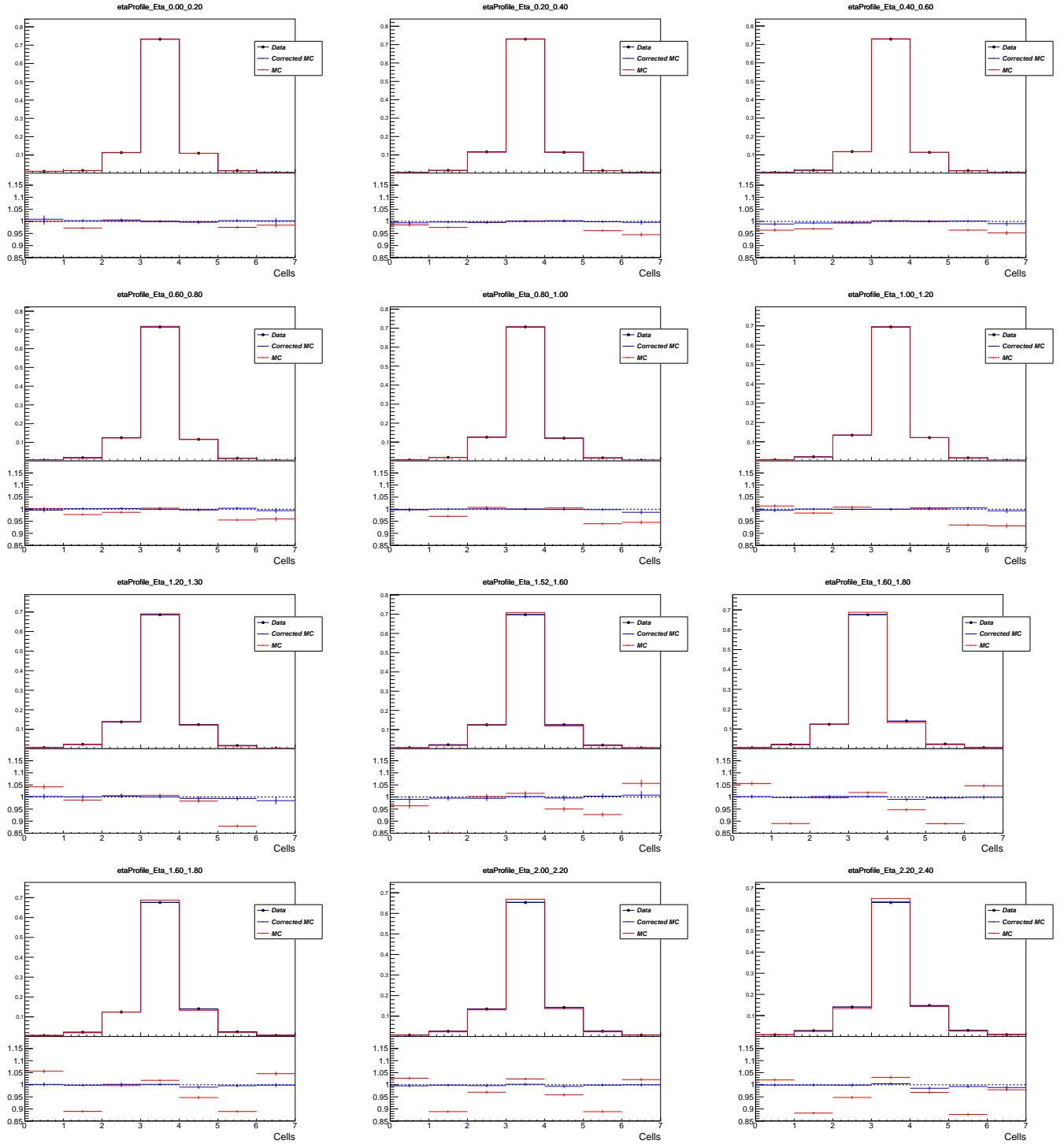
**Figure 5.19:** Energy profile in $\eta$ dimension for all $\eta$ slices.
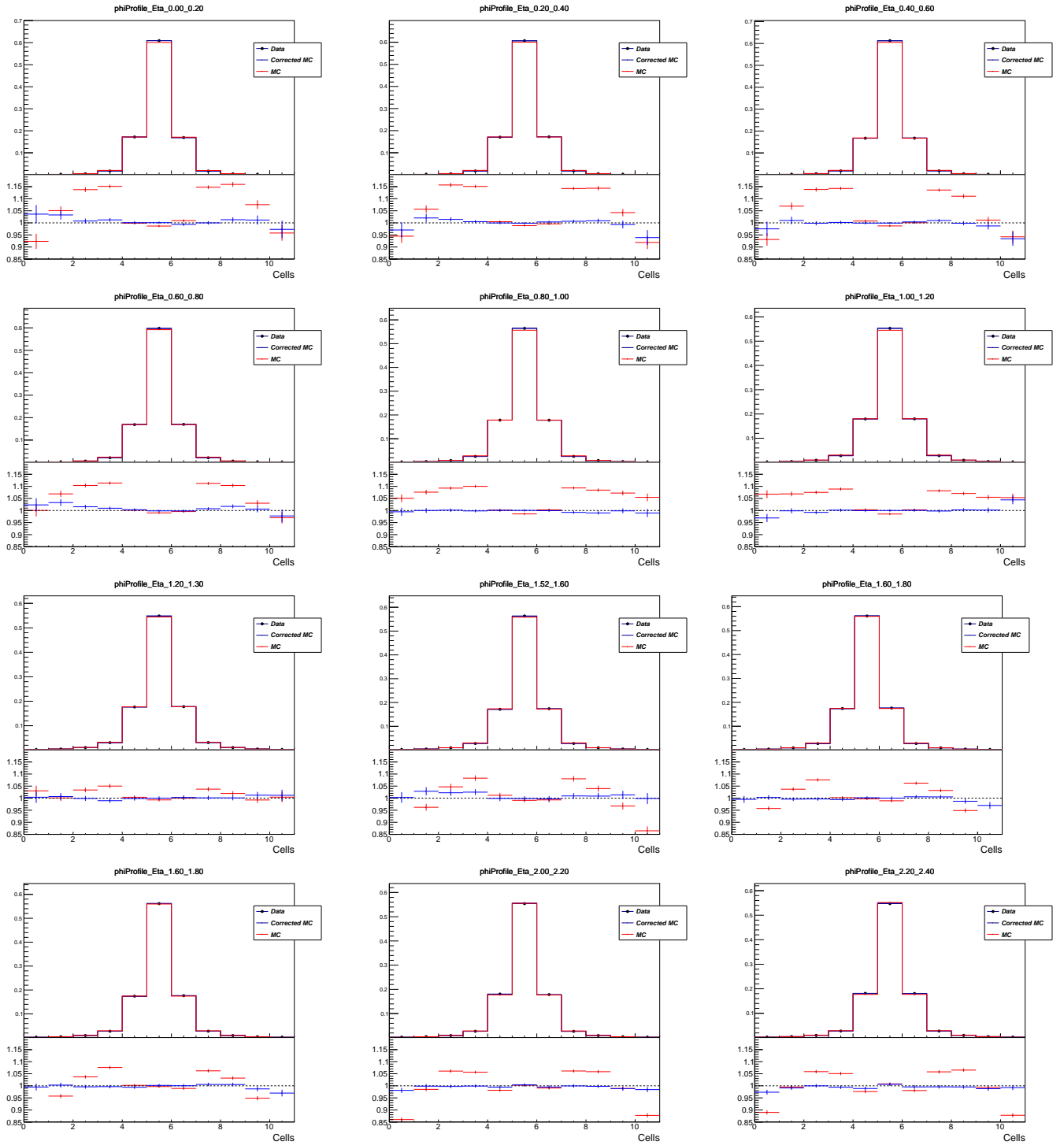
**Figure 5.20:** Energy profile in $\phi$ dimension for all $\eta$ slices.

# 6 Event reconstruction

Interpreting the detector signals and reconstructing the event along with the associated objects relies on a number of sophisticated algorithms. The current chapter provides superficial overview of these algorithms that allow to reconstruct particles (electrons, photons, muons etc) and event-related parameters like the coordinates of the primary vertex.

## 6.1 Charged particles track reconstruction

A track $q$ is formed based on the information from the ID and is represented by five parameters: $q = (d_0, z_0, \phi, \theta, q/p)$, where $d_0$ is the distance from the track to the Z axis (transverse impact parameter), $z_0$ is the Z coordinate of the perpendicular dropped from the track onto the Z axis (longitudinal impact parameter) (see fig. 6.1), $\phi$ and $\theta$ are the azimuthal and polar angles correspondingly and $q/p$ is the charge to momentum ratio of the particle. The process for track reconstruction is the same for lepton and charged hadron candidates.
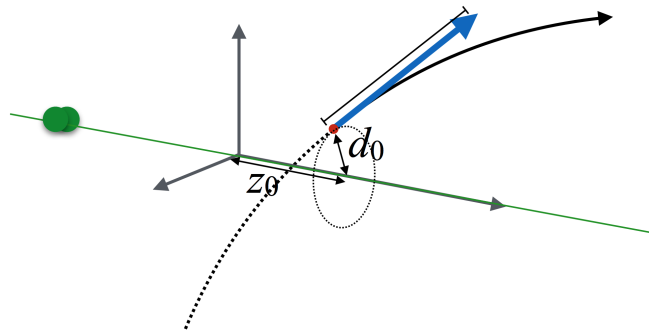


**Figure 6.1:** Impact parameters $z_0$ and $d_0$ [116].

To form tracks using the detector response information the following steps are performed [117]:

- **Clustering** single hits in the pixel and SCT detectors. Neighbouring hits are combined to form a single cluster, clusters are then transformed into *space points* that have having 3D coordinates. A cluster may be identified as a single-particle cluster or as merged cluster, created by two or more particles. Identification of a cluster as a merged one and separation of energy deposits between

the particles (possible only for two particles) is performed by means of a Neural Network (NN) algorithm.

- **Forming seeds** out of the space points. To form a seed three space-points originating from unique layers of the silicon detectors (pixel or SCT) are used. All possible combinations of seeds are formed at this stage. For every seed a crude estimate of the track parameters is performed.

- **Track candidates** are formed out of the seeds by extending them within the silicon sub-detectors following the most likely path. A combinatorial Kalman filter [118] is used to build the track candidates. The purity of the seeds depends significantly on the sub-detector that recorded the corresponding space-points. SCT-only seeds are considered the most reliable, followed by the seeds that origin only from the pixel detector space-points, and the least reliable are the seeds originating from both of these sub-detectors - that determines the order of seed consideration when composing track candidates.
  Some fraction of the seeds that meet the necessary requirements become track candidates, the rest are discarded. A seed may be used for more than one track candidate if more than one space-point extension exists on the same layer.

- **Ambiguity solving** is the next step necessary to eliminate incorrectly assigned space-points or resolve conflicting track candidates that have an overlapping space-point. At this stage the track candidates are assigned a *track score*. The track score depends on the number of clusters associated to the track and which sub-detector these clusters originate from, the existence of holes (the absence of a cluster associated to a detector layer crossed by the track), the quality of the $\chi^2$ fit of the track and track momentum.

  The tracks are ordered by their track score and consequently fed to the ambiguity resolving sequence. A track must pass a number of kinematic cuts, impact parameters cuts, number of holes, number of clusters and shared clusters cuts, otherwise the track candidate is rejected. If a track candidate has no shared clusters with other candidates it is accepted after that. If there are merged clusters then it is up to the NN to either accept the track, reject it or eliminate a space-point and recycle the updated track candidate (see Fig. 6.2a).

- **TRT extension** means matching of the track, composed using the information from silicon sub-detectors to the trace in the TRT tracker. This allows to improve momentum measurement benefiting from extended track length.

- Final high-resolution **track fit** is performed using all available information. Position and uncertainty of each cluster are determined by an additional NN allowing for more precise track parameters. The curvature of the particle track also serves for charge sign identification.
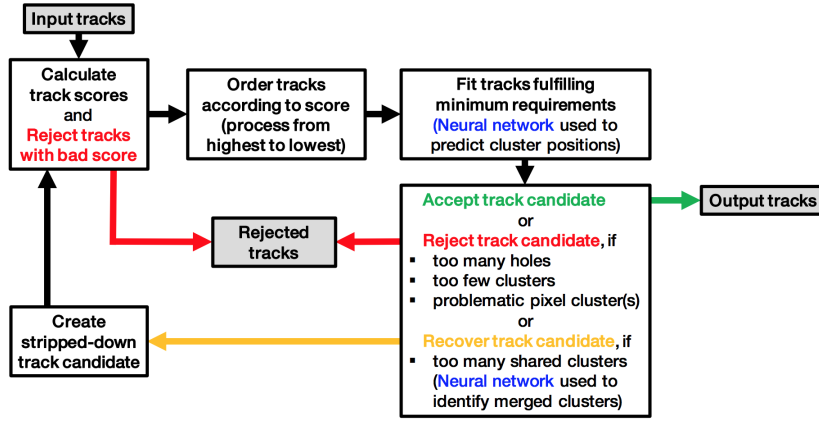
(a) Track ambiguity resolver algorithm.
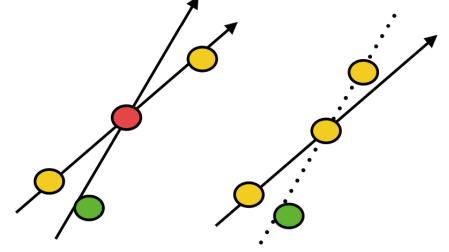
(b) Tracks sharing space-points.

Figure 6.2: Ambiguity solving process [117].

## 6.2 Determining the primary vertex of the event

Primary vertex determination is crucial for physics analyses for many reasons. One of them is the necessity to separate particles originating from hard events from pile-up. Another reason is to keep track of long-lived decay chains and distinguish between prompt and non-prompt particles. Flavour tagging, background suppression and decay reconstruction also rely heavily on the primary vertex determination.

After reconstructing the tracks of individual particles the obtained information is used to reconstruct the Primary Vertex (PV) of the event [119]. The procedure relies on the reconstructed tracks and goes as follows:

- A seed from the first vertex is selected. The transverse position of the seed is taken as the centre of the beam spot. The z-coordinate of the seed is calculated as the mode of $z_0$ coordinates of the tracks.

- Using the seed and the available tracks an iterative fit is performed in order to find the best position for the PV. In each iteration the tracks that are less compatible with the vertex are down-weighted and the vertex position gets recomputed. With every iteration the spread in the weight increases, separating track set into compatible tracks that mostly determine the vertex position and incompatible tracks that have little weight and therefore very little influence on the track position.

- After the fit is done compatible tracks remain assigned to the vertex, while incompatible tracks are removed from it. These incompatible tracks can be used in the determination of a different vertex.

- The procedure is repeated with the remaining tracks of the event.

- The primary vertex is a vertex with the highest sum of the assigned tracks transverse momenta $\sum_{tracks} p_T^2$.

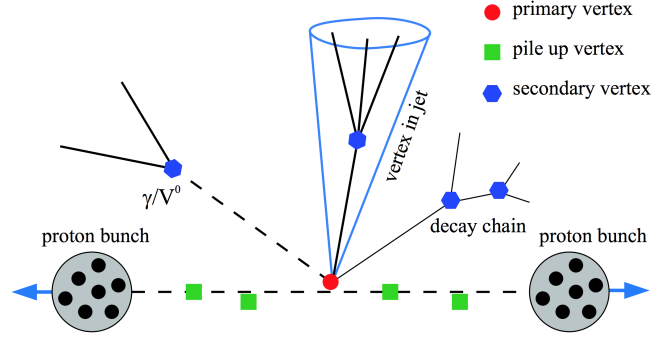For the upcoming Run 3 of the LHC certain improvements and modifications are foreseen [120].



**Figure 6.3:** Primary, secondary and pile-up vertices [121].

## 6.3 Muon reconstruction and identification

Muon reconstruction relies primarily on the information from the ID (the muon track) and the Muon Spectrometer (MS), sometimes also using additional information from the calorimeter. At the first stage a muon is independently reconstructed in the tracker and in the spectrometer, and then the two reconstructed tracks are combined to compose a muon track used in the physics analyses [122]. Track reconstruction is described in section 6.1.

### 6.3.1 Muon reconstruction

Muon reconstruction in the muon spectrometer begins with a search for hit patterns in each muon chamber and forming of the segments. Using the Hough transform [123] the hits in each MDT chamber and nearby trigger chamber are aligned on trajectories in the bending plane. The orthogonal coordinate is measured with RPC and TGC detectors. A separate combinatorial search is conducted in the CSC detectors in $\phi$ and $\eta$ detector planes.

Then the track candidates are built by fitting hits from different layers. This algorithm starts a combinatorial search first using the segments from the middle layers as seeds, as there are more trigger hits in the middle layer. The search is later extended to include the segments from other layers as seeds. Segment selection criteria are based on hit multiplicity and fit quality. The segments are matched using their relative positions and angles. In all the regions, except barrel-endcap transition region, at least two matching segments are needed to build a track (one segment is enough in the transition region).

A single segment can be used by two or more track candidates. An overlap removal algorithm decides to which track should a segment belong or shares a segment between two tracks. A global $\chi^2$
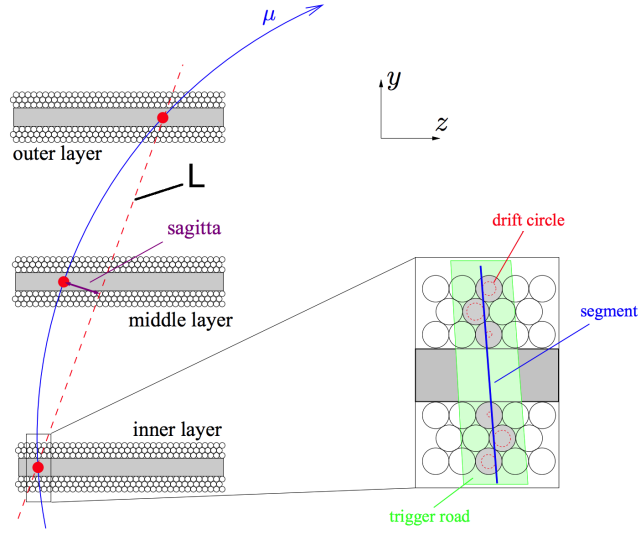
**Figure 6.4:** Sagitta used for the determination of the muon momentum [124].

fit is used to fit all the hits associated to every track. If the $\chi^2$ fit meets the designated criteria then the track is accepted. If a hit impair the $\chi^2$ fit significantly, then this hit may be removed and the fit is repeated. On the other hand, new hits may be recovered if they fit the track candidate trajectory.

Accurate fitting of the track trajectory is extremely important for the measurement of muon momentum. A quantity called *sagitta* is measured by the MS (see Fig. 6.4). Knowing the length $L$ and the sagitta $S$ we can determine the momentum:

$$p = \frac{BL^2}{8S},\tag{6.1}$$

where $B$ is the magnetic field strength.

After the muon gets reconstructed in every detector system separately, the obtained information is combined to form a reconstructed muon object. Depending on the detectors used for the combined reconstruction there are *four types of muons* defined (see Fig. 6.5):

- **Combined (CB) muon** is formed from a global refit of the tracks reconstructed independently in the ID and in the MS. During this global refit the hits from both detectors are used and also new hits may be added. Normally the outside-in pattern is used, when MS track is extrapolated inwards to match ID track. Inverse inside-out procedure is used as a complementary approach.

- **Segment-tagged (ST) muon** is a particle with an ID track that was extrapolated to the MS and associated with at least one local track segment in the MDT or CSC chambers. Normally these are muons with low $p_T$ or their trajectory crosses regions with reduced MS acceptance.

- **Calorimeter-tagged (CT) muon** has a valid ID track that can be associated to an energy deposit in the calorimeter compatible with minimum-ionizing particle. The CT muons have the lowest

purity among the muon types although they provide acceptance where the MS coverage may be absent, like the very central region with $|\eta| \leq 0.1$ for $15 < p_T < 100$ GeV.

- **Extrapolated (ME) muon** (standalone muon) trajectory is reconstructed base only on the MS track and a loose requirement to match the IP. ME muons allow to extend the muon acceptance to the region which is not covered by the ID, namely $2.5 < |\eta| < 2.7$.

In case of overlap between different muon types the preference is given to CB muons, then to ST and then to CT muons. ME muons overlaps are resolved based on the MS track quality.
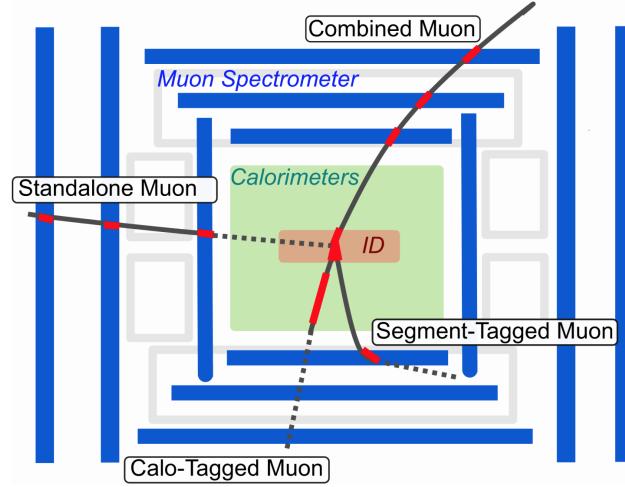


**Figure 6.5:** The four types of reconstructed muons [122].

## 6.3.2 Muon identification

Muon identification is a set of measures to ensure that the registered particle has indeed the characteristics of a muon and to identify the mechanism of its production. Muons created in the course of decay of a short-lived particle (e.g. a massive boson) are called *prompt muons*, while those originating from hadron or tau decays are called *non-prompt*. Muon identification plays an important role in background suppression and guaranteeing a robust momentum measurement.

Muons that are created during the in-flight decay of the charged hadrons in the ID usually have a distinctive "kink" topology in their reconstructed track. This results in a decreased quality of the resulting track fit and the incompatibility between the results of momentum measurement in the ID and MS. Muons originating from W boson decays are called *signal*, while those coming from hadron decays are called *background*. For CB muons the three main identification variables are the following:

- *q/p significance* is defined as $\frac{|(q/p)_{ID} - (q/p)_{MS}|}{\sqrt{\sigma^2(q/p)_{ID} + \sigma^2(q/p)_{MS}}}$ - an absolute difference between $q/p$ measured in the two detectors over the combined uncertainty.

- Relative transverse momentum difference $\rho = \frac{|p_T^{ID} - p_T^{MS}|}{p_T^{combined}}$.

- Normalized $\chi^2$ fit of the combined track.

Robust momentum measurement is ensured by specific requirements to the number of hits in the ID and MS. A number of muon identification selections (working points) is developed to address specific analyses.

### 6.3.3  Muon isolation

Isolated muons are a defining signature of massive boson decays. In the decays of W, Z and Higgs bosons muons are created separated from the rest of the particles. Quantitative measurement of detector activity around a muon candidate is called *muon isolation* and serves as an invaluable tool for background suppression. Muon isolation is assessed through two observables: one is track-based, another is calorimeter-based.

The track-based observable $p_T^{varcone20}$ is defined as a scalar sum of all the particles with $p_T > 1$ GeV in a cone $\Delta R = min(10 GeV/p_T^{\mu}, 0.2)$ around the muon with transverse momentum $p_T^{\mu}$ excluding the proper track of the muon. The $p_T$ dependence helps this definition to perform better for the muons created in the decay of the particles with high transverse momentum.

The calorimeter-based isolation observable $E_T^{topocone20}$ is defined as the sum of the transverse energy of all the topological clusters in a cone of a size $\Delta R = 0.2$ around the muon after subtracting the proper muon energy deposit and correcting for the pile-up effects.

In both cases the size of the cone may be varied, normally in the range between 0.2 and 0.4, depending on the analysis needs. Isolation criteria are typically defined using the relative isolation variables, using the ratio of $p_T^{varcone20}$ and $E_T^{topocone20}$ to the transverse momentum. A number of working points exist, each having a certain requirements for one or both of the isolation variables.

## 6.4  Electron reconstruction and identification

### 6.4.1  Electron reconstruction

Electron reconstruction starts with two separate parts: track reconstruction in the ID and cluster reconstruction in the calorimeter, which are then matched to each other in order to make an electron candidate [125]. During Run 2 two algorithms were used for the cluster reconstruction, both of them are described below.
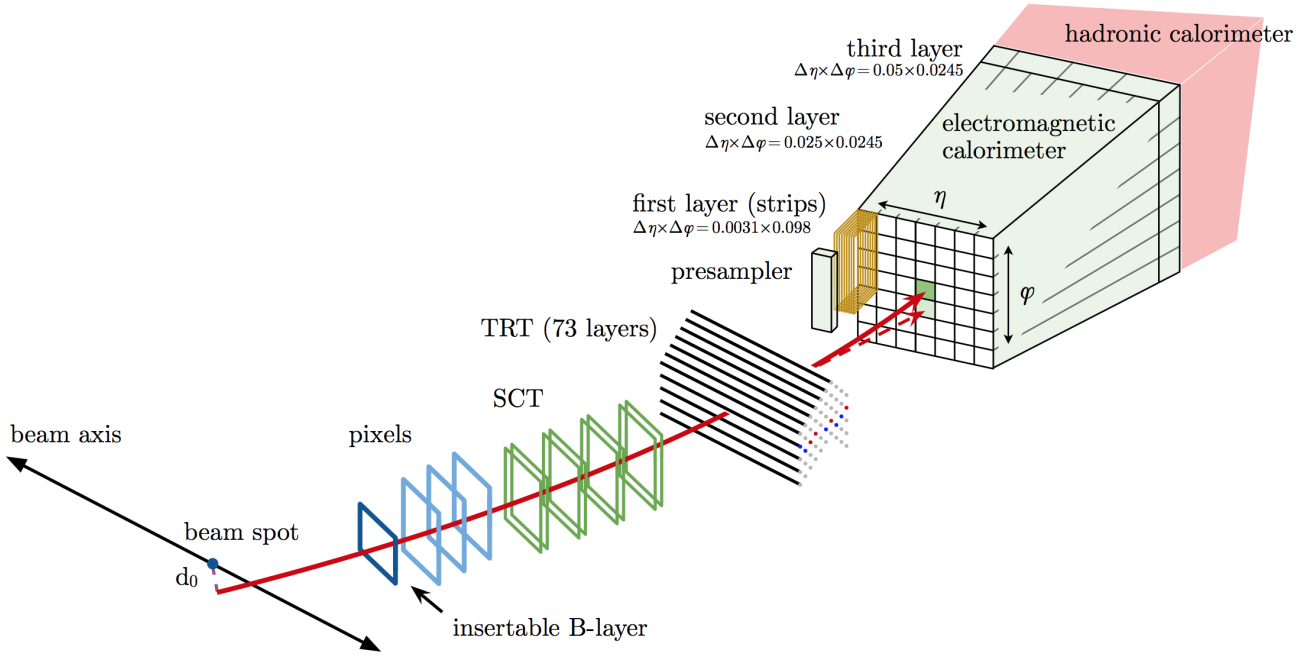
**Figure 6.6:** The path of an electron through the detector is shown by solid red line. The dashed red line denotes the trajectory of a photon, produced as a Bremsstrahlung radiation in the TRT [125].

### Sliding window

It must be mentioned that this method is deprecated and starting from 2017 is replaced by the topocluster method described in the next section. The EMC is divided into a grid of 200x256 towers in $\eta \times \phi$ plane, each tower having a size of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$, reproducing the granularity of the second layer in the EMC. Energy deposits in all available calorimeter layers (first, second and third layers of the EMC in the region $|\eta| < 2.47$ and the presampler in the region $|\eta| < 1.8$) are approximately calibrated at the EM scale and summed up for each tower. If the cumulative energy deposit in a certain tower exceeds 2.5 GeV then this tower is used as a seed. Then for every seed a sliding window algorithm of size $3 \times 5$ is used [126], forming a cluster around every seed.

It happens that two seed-cluster candidates are found in close proximity. When their towers overlap within an area of $\eta \times \phi = 5 \times 9$ in units of $0.025 \times 0.025$ the two clusters are considered overlapping. In this case two options are possible:

- If the transverse energies of the two clusters are more than 10% different then the cluster with higher $E_T$ is retained.

- If the difference in the transverse energies is within 10% then the cluster with higher value of the $E_T$ in the central tower is kept.

After the overlap is resolved the duplicate cluster is removed.

## Topocluster reconstruction
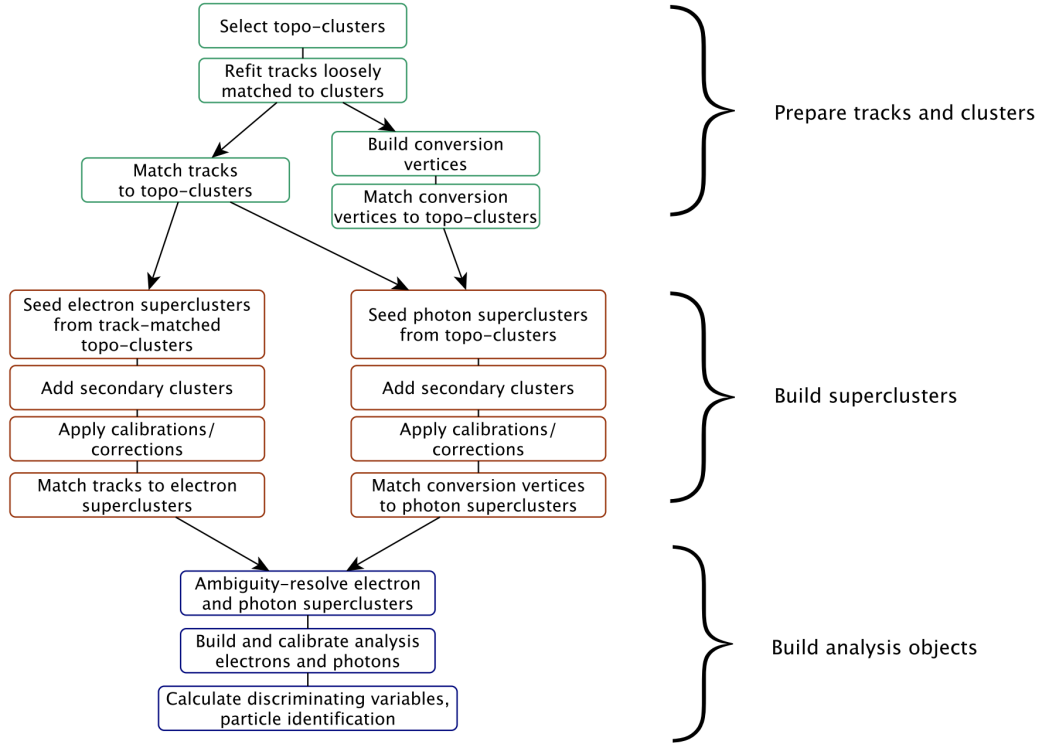


**Figure 6.7:** The algorithm scheme for topocluster reconstruction.

The algorithm for topocluster reconstruction [127], [128] starts with composing proto-clusters in the calorimeter using the noise threshold:

$$\zeta_{cell}^{EM} = \frac{E_{cell}^{EM}}{\sigma_{noise,cell}^{EM}}, \tag{6.2}$$

where $E_{cell}^{EM}$ is the cell energy at the EM scale and $\sigma_{noise,cell}^{EM}$ is the expected cell noise. The latter comprises the electronic and the pile-up noise estimate based on the expected instantaneous luminosity. The proto-cluster is formed around a cell with $|\zeta_{cell}^{EM} \geq 4|$. Then the neighbouring cells are added to the proto-cluster. If an added cell passes the requirement of $|\zeta_{cell}^{EM} \geq 2|$ then it serves as a seed for the next iteration, collecting all of its neighbours to the proto-cluster. If the two proto-clusters share a cell with $|\zeta_{cell}^{EM} \geq 2|$ then these proto-clusters are merged together. Proto-clusters with two local maxima are split into two clusters. For a proto-cluster to be considered as the EM topocluster it must have at least 50% of its energy being contained in the EMC. At the stage of track reconstruction the tracks are first extended and fitted with the global $\chi^2$ fitter using the pion hypothesis [129]. If it fails, then a more complicated pattern reconstruction algorithm based on Kalman filter is used [130]. This algorithm

uses the electron hypothesis and allows up to 30% energy loss at each material surface. Then the tracks are loosely matched to the EM clusters if they meet one of the following criteria:

- The tracks extrapolated to the second layer of the EMC are consistent in $\phi$ and $\eta$ (matching in $\eta$ is not required for TRT-only tracks).

- The extrapolated tracks are consistent in $\phi$ (with a bit tighter requirements) and $\eta$ after rescaling the track momentum to cluster momentum.

Matching in $\phi$ coordinate assumes charge asymmetry to account for different direction of possible Bremsstrahlung radiation for positive and negative particles. Then the loosely matched tracks that have at least four silicon hits are refitted using the optimized Gaussian-sum filter (GSF) [131], that allows to better take into account the energy losses in solid material.

### Track-cluster matching

Once the track is fitted with the GSF algorithm the final matching with the cluster is performed using tighter matching requirements between the track and the cluster barycentre. If matching criteria are met with two or more tracks then the ambiguity resolving algorithm is used. This algorithm takes into account a number of parameters like the distance between the cluster barycentre and the track in $\phi$ and $\eta$, number of hits in the silicon detector and in the innermost silicon layer, association to photon conversion vertex, $E/p$ ratio and $p_T$. This allows to exclude converted photons as electron candidates and also helps to maintain high photon reconstruction efficiency. After track-cluster matching the electron cluster is extended around the seed to $3 \times 7$ in the barrel region or $5 \times 5$ in the end-cap region by adding one row of the cells on each side.

### Supercluster reconstruction

The composition of an electron supercluster is performed in two stages: first, the candidate EM topocluster is tested to be used as a seed for the supercluster. In the second stage the nearby EM topoclusters can be identified as satellite clusters, emerging from Bremsstrahlung radiation or topocluster splitting.

First the EM topoclusters are sorted by their $E_T$ in descending order. For the cluster to be considered a seed it must have $E_T > 1$ GeV, must be matched to a track with at least four hits in the silicon detectors and should not be assigned as a satellite cluster to any other seed. If these requirements are met then the algorithm described in Fig. 6.8 is started. First, all topoclusters within a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ around the seed cluster barycentre are added as satellite cluster, as they most probably represent secondary EM showers coming from the same initial electron. Also, if a cluster within $\Delta\eta \times \Delta\phi = 0.125 \times 0.3$ window around the seed cluster barycentre share the "best-matched" track with the seed cluster - it is also added as a satellite. Finally the energy of the reconstructed cluster
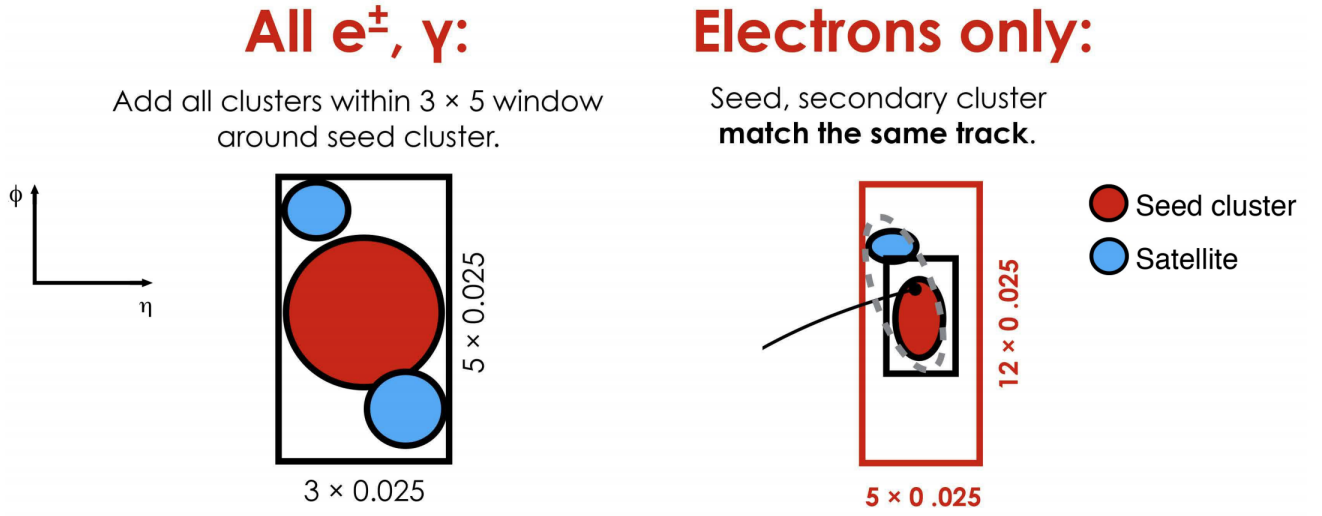
**Figure 6.8:** Supercluster reconstruction for electrons. Seed clusters are shown in red, satellite clusters in blue.

must be calibrated. The calibration is performed using a multivariate technique based on data and MC samples using $Z \to ee$ events [132], [133]. The shower shapes and other discriminating variables are computed at this stage.

### 6.4.2 Electron identification

Prompt electrons in the central region of the ATLAS detector ($|\eta| < 2.47$) are selected using a likelihood-based (LH) identification. The LH uses a number of inputs from ID and calorimeter detectors, as well as combined information from both detectors (see Table 1 in [125]). The probability density functions (pdfs) for the likelihoods of Run 2 were obtained using the simulated events.

The electron LH is based on the products of n pdfs P for signal $L_S$ and background $L_B$:

$$L_{S(B)}(\mathbf{x}) = \prod_{i=1}^{n} P_{S(B)}^{i}(x_i), \tag{6.3}$$

where $\mathbf{x}$ is the vector of the LH input parameters, $P_S^i$ and $P_B^i$ are the pdf values for parameter $i$ at value $x_i$ for signal and background respectively. The LH operates at a number of working points, the higher the likelihood - the lower is the efficiency. For example, the efficiencies for identifying a prompt electron with $E_T = 40$ GeV for Loose, Medium and Tight working points are 93%, 88% and 80% respectively. Prompt electrons are assumed to come from the signal, while background includes the jets that mimic the prompt electrons, electrons from photon conversions and non-prompt electrons from hadron decays. For each electron candidate a discriminant $d_L$ is composed:

$$d_L = \frac{L_S}{L_S + L_B}, \tag{6.4}$$

that defines the electron likelihood identification. This discriminant $d_L$ has a sharp peak at unity for the signal and at zero for the background, which is not very convenient for picking working points. That is why the discriminant distribution is transformed using the inverse sigmoid function:

$$d'_L = -\tau^{-1} \ln\left(d_L^{-1} - 1\right), \tag{6.5}$$

where $\tau = 15$. Each operating point is assigned with a $d'_L$ value - if a discriminant exceeds this value for a given electron then this electron is considered signal.

There are two advantages of using likelihood-based approach comparing to selection-criteria-based ("cut-based") identification:

- The drawback of a cut-based approach is that if an electron fails to pass one of the cuts - it is definitely removed from the selection, while in the LH approach it is still possible for this electron to pass the selection thanks to other parameters. This quality promotes the selection efficiency.

- In case of a significant overlap in signal and background distribution of a certain parameter using it in a cut-based identification would entail large losses in efficiency. In the likelihood-based identification this parameter may be added without penalty.

The likelihood input parameters were obtained from the simulated events, which means that real distributions in data may differ due to various mismodelling effects. These effects must be corrected in order to get an accurate and efficient identification. Mismodelling may depend on coordinates or energy. Chapter 5 of this dissertation is dedicated to the correction of electromagnetic shower shapes in the calorimeter, which are among the likelihood input parameters.

### 6.4.3 Electron isolation

Electron isolation plays a very important role in background suppression in physics analyses. Since electrons are reconstructed using the information from two different detectors - two different isolation definitions are possible, track-based and calorimeter-based. Let's first consider calorimeter-based isolation.

As depicted in Fig. 6.9 the raw isolation energy $E_{isol}^{T,raw}$ includes the energy of all the topoclusters, barycentres of which fall within the isolation radius $\Delta R$. It also includes core energy of the electron candidate $E_{isol}^{T,core}$ which comprises the $5 \times 7$ cells within the area of $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$. The fixed size of the core ensures simplicity and stability, although it may happen that the topocluster is larger than the size of the core resulting in attributing the proper energy of the electron to the outside activity. This leakage effect is corrected for using no pile-up simulated events, parametrizing the leakage with a Crystal Ball function as a function of the transverse energy $E_{T,leakage} = E_{T,leakage}(E_T)$.

Another effect that must be corrected for is the pile-up and underlying event contribution. This contribution is estimated from from the ambient energy density [134]. In every event all positive-energy
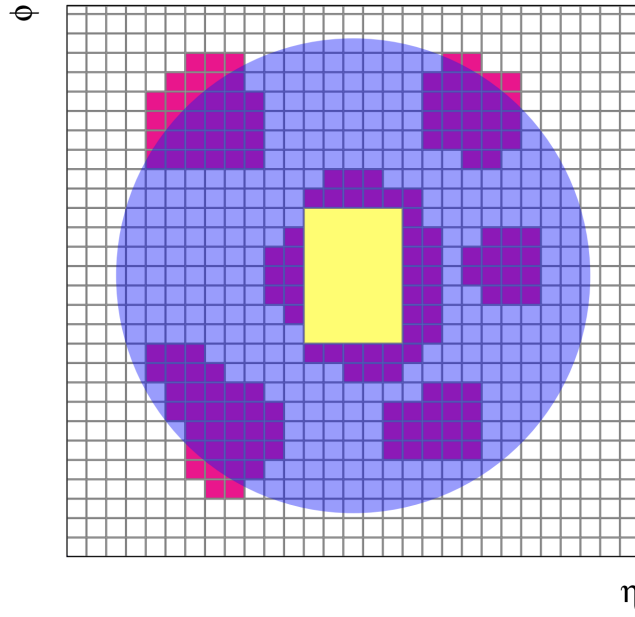
**Figure 6.9:** The isolation cone is centred at the candidate electron. All topological clusters, shown in red, are included in the raw isolation variable. The $5 \times 7$ cells included into core subtraction method are marked in yellow [128].

topological clusters are taken into account in the entire range of calorimeter acceptance $|\eta| < 5$ using the $k_t$ jet clustering algorithm with a radius parameter $R = 0.5$ and no jet $p_T$ threshold. Then for every jet its area $A$ is estimated and the jet energy density $\rho = p_T/A$ is computed. Using the information on the jet energy density together with the location of every jet one can obtain the median energy density $\rho_{median}(\eta)$ - a rapidity-dependent estimate of jet densities for every event. Then the pile-up correction can be evaluated in the following way:

$$E_{T,pile-up}(\eta) = \rho_{median}(\eta) \times (\pi \Delta R^2 - A_{core}), \qquad (6.6)$$

where $\Delta R$ is the radius of the isolation cone, and $A_{core}$ is the area of the subtracted signal core. Finally the calorimeter isolation variable may be defined as follows:

$$E_{T,cone}^{isol} = E_{T,raw}^{isol} - E_{T,core} - E_{T,leakage} - E_{T,pile-up}. \qquad (6.7)$$

The track-based isolation includes all tracks with $p_T > 1$ GeV, within the fiducial region of the ID, that satisfy basic track quality requirements. Pile-up is mitigated by requiring that $|z_0 \sin \theta| < 3$ mm, to ensure that the track originates from the primary vertex. The track-based isolation is composed of all the tracks that fall within the radius $\Delta R$ excluding the candidate electron track.

The own contribution of the candidate track into the isolation must also include possible Bremsstrahlung radiation emitted by the candidate electron. For that reason the tracks are extrapolated to the second layer of the EMC and if they fall within a window of $\Delta \eta \times \Delta \phi = 0.05 \times 0.1$ around the cluster position. The resulting variable is called $p_T^{isol}$.

The track-based isolation allows to use variable-size cone, making the cone smaller for boosted particles. The cone size for the $p_{T,var}^{isol}$ would be:

$$\Delta R = min\left(\frac{10GeV}{p_T[GeV]}, R_{max}\right), \tag{6.8}$$

where $R_{max}$ is the maximum cone size and may vary depending on the analysis needs, typically between 0.2 and 0.4.

## 6.5    Particle flow objects

The measurement of hadronic objects and particle showers remains a complicated task due to the large variety of particle types and properties they posses and because of the large energy/momentum span of the measured objects. For the low-energy charged particles the ID shows better momentum resolution and angular resolution. On the other hand, the calorimeter shows better performance at high energy and is also capable of detecting neutral particles. The idea behind the Particle Flow (PF) algorithm [135] is to combine the information from the two detectors to obtain the best result possible. To properly take into account every particle it has to be ensured that every particle detected in both detectors is counted only once. This means that for a charged particle its deposit in the calorimeter must be found and subtracted. The Particle Flow Object (PFO) reconstruction process is schematically presented in Fig. 6.10. The process starts with getting *tight* tracks from the ID, meaning these tracks
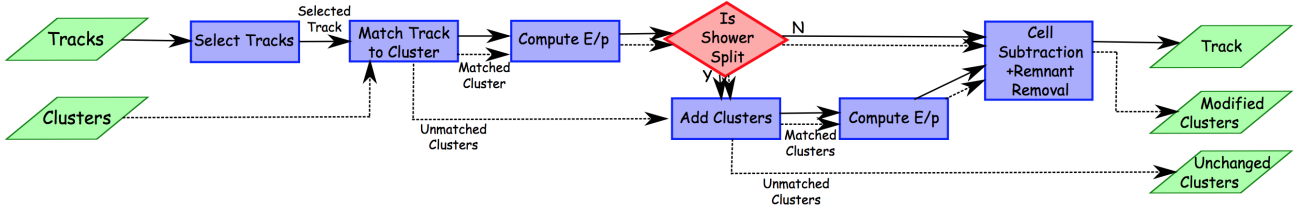


**Figure 6.10:** The algorithm scheme for particle flow object reconstruction [135].

must have at least nine hits in the silicon detectors and no holes n the pixel detector. The tracks must have $|\eta| < 2.5$ and $0.5 < p_T < 40$ GeV, corresponding to the kinematic region where tracks offer better resolution than the calorimeter. The tracks associated to leptons are removed.

The calorimeter topoclusters reconstructed like it was described in section 6.4.1 and calibrated using the EM scale are matched to the tracks based on their spacial position and measured momentum. First the ranked based on a distance metric:

$$\Delta R' = \sqrt{\left(\frac{\Delta\phi}{\sigma_\phi}\right)^2 + \left(\frac{\Delta\eta}{\sigma_\eta}\right)^2}, \tag{6.9}$$

where $\Delta\phi$ and $\Delta\eta$ are the angular distances between the topocluster barycentres and the track, $\sigma_\phi$ and $\sigma_\eta$ are uncertainties in topocluster width. Preliminary matching is reached by requiring that

$E^{clus}/p^{trk} > 0.1$, where $E^{clus}$ is the cluster energy and $p^{trk}$ is the track momentum.

It often happens, that energy deposit of a particle is split between two (most often) or more clusters. Then a split shower recovery procedure is initiated, looking for matching clusters in the radius of $\Delta R = 0.2$ around the track extrapolated to the second layer of the EMC. Then it is estimated if the energy of the track and the energy of the associated topocluster is consistent. If it is the case then the topoclusters matched to the tracks are removed.

Eventually two particle collections are obtained: a collection of charged particle flow objects (cPFOs) each with an associated track and neutral particle flow objects (nPFOs) with a calorimeter deposit. The former must also match the primary vertex, having $|z_0 \times \sin\theta| < 2$ mm. The full procedure is described in detail in [135].

<div style="text-align: right; font-size: 3em;">7</div>

# Calibrations and corrections

The object reconstruction procedure uses the same algorithms for the real data and for the simulation. However, the simulation is never flawless en every aspect. So in order to put the Monte-Carlo simulated events on the same page with the data a number of calibrations and corrections are introduced to the reconstructed objects and to the observables. The current chapter describes the calibrations and corrections, applied in the framework of the W boson transverse momentum measurement analysis.

## 7.1 Electron corrections

This section describes the electron calibrations and corrections used in $p_T^W$ analysis. They were derived from the low-$\mu$ special run data collected in 2017 and 2018 at 5 and 13 TeV and the dedicated Monte-Carlo (MC) samples [136], as well as from the standard ATLAS high pileup data collected during the Run 2.

### 7.1.1 Energy scale and resolution correction
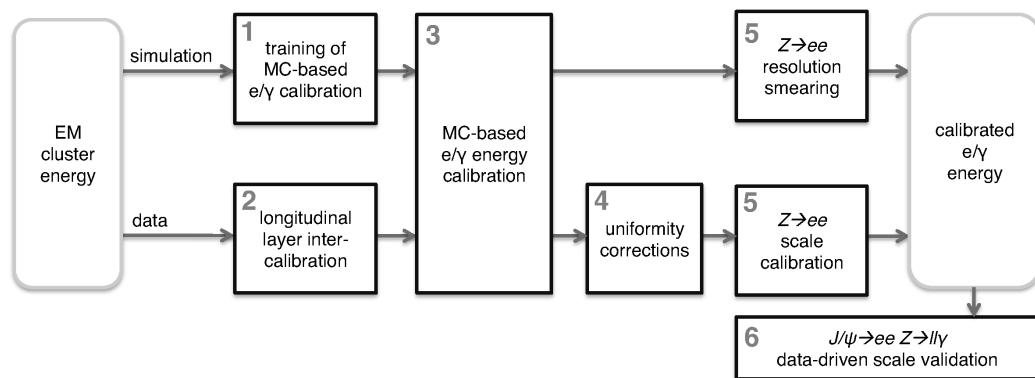


**Figure 7.1:** Schematic overview of energy response calibration procedure for electrons and photons.

In order to obtain the energy scale and resolution corrections, electrons from $Z \rightarrow ee$ process were used. The selection criteria were the same for data and MC simulation. For high-$\mu$, electron candidates must pass the corresponding triggers and are required to have $p_T > 27\,\text{GeV}$ and $|\eta| < 2.47$, satisfying the medium LH ID criteria and loose isolation criteria as described in Ref. [132]. In the low pile-up case electron the requirements are loosened: the candidates must pass the triggers requiring to have $p_T > 15$ and satisfy loose isolation criteria. The full description for the selection criteria for low pile-up electrons is provided in Section 8. The energy scale correction follows the method described in detail in Ref. [112] and schematically described in Fig. 7.1. The energy in both data and MC is corrected using the MVA-based algorithm, then the data is corrected for pile-up and uniformity. The energy response in data is calibrated using the $Z \rightarrow ee$ peak to match exactly the Z resonance in the simulation. Two correction factors are introduced: the energy scale factor $\alpha$ and the constant term $c'$. The correction factors are extracted using the template method described in Ref. [137]:

- The calorimeter is split into $i$ slices in $\eta$ and for each slice the energy response in data is corrected in the following way:
$$E^{data,corr} = E^{MC} = E^{data,uncorr}/(1 + \alpha_i),$$
  where $E^{data,uncorr}$ and $E^{MC}$ are the energy response in data and MC respectively, $\alpha_i$ is the energy correction factor for the $i^{th}$ calorimeter slice in $\eta$.

- The relative energy measurement resolution can be represented as a quadratic sum of three uncorrelated terms:
$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c,$$
  where $b$ term stands for electron and pile-up noise term, $a$ is the stochastic term related to the development of the electromagnetic shower and $c$ is constant. The constant term $c$ includes all corrections that do not depend on the energy (e.g. non-uniformities, mechanical imperfections, temperature gradients) and is normally within 1% in the barrel and about 1-2% in the end-cap. In order to widen the MC mass peak and match it to the data in each rapidity bin an additional constant term $c'$ is added:
$$\left( \frac{\sigma(E)}{E} \right)_i^{data} = \left( \frac{\sigma(E)}{E} \right)_i^{MC} \oplus c_i'.$$

Normally in the standard high-pileup data, the energy scale factors corrections are obtained in 68 $\eta$ bins. For the low pile-up runs wider bins were also considered due to smaller number of $Z \rightarrow ee$ events. Figure 7.2 demonstrates the need for wider bins, as 68 bins result in high uncertainty, especially in the endcap.

Two binnings were considered:

- 48 bins with smaller bins in the barrel and wider bins in the endcap

- 24 bins of equal size, as shown in Table 7.1.

| -2.47 | -2.4 | -2.1 | -1.8 | -1.55 | -1.37 | -1.2 | -1 | -0.8 | -0.6 | -0.4 | -0.2 | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 1 | 1.2 | 1.37 | 1.55 | 1.8 | 2.1 | 2.4 | 2.47 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|

**Table 7.1:** Values of $\eta_{\text{calo}}$ bin frontiers for energy scale factors for low pile-up runs [138].

As can be seen from Fig. 7.2, the statistical instability for the endcap bins is significantly reduced if wider bins are used. Since the $\alpha$ factors are quite similar in 48 and 24 bin cases, the latter is chosen as the baseline.
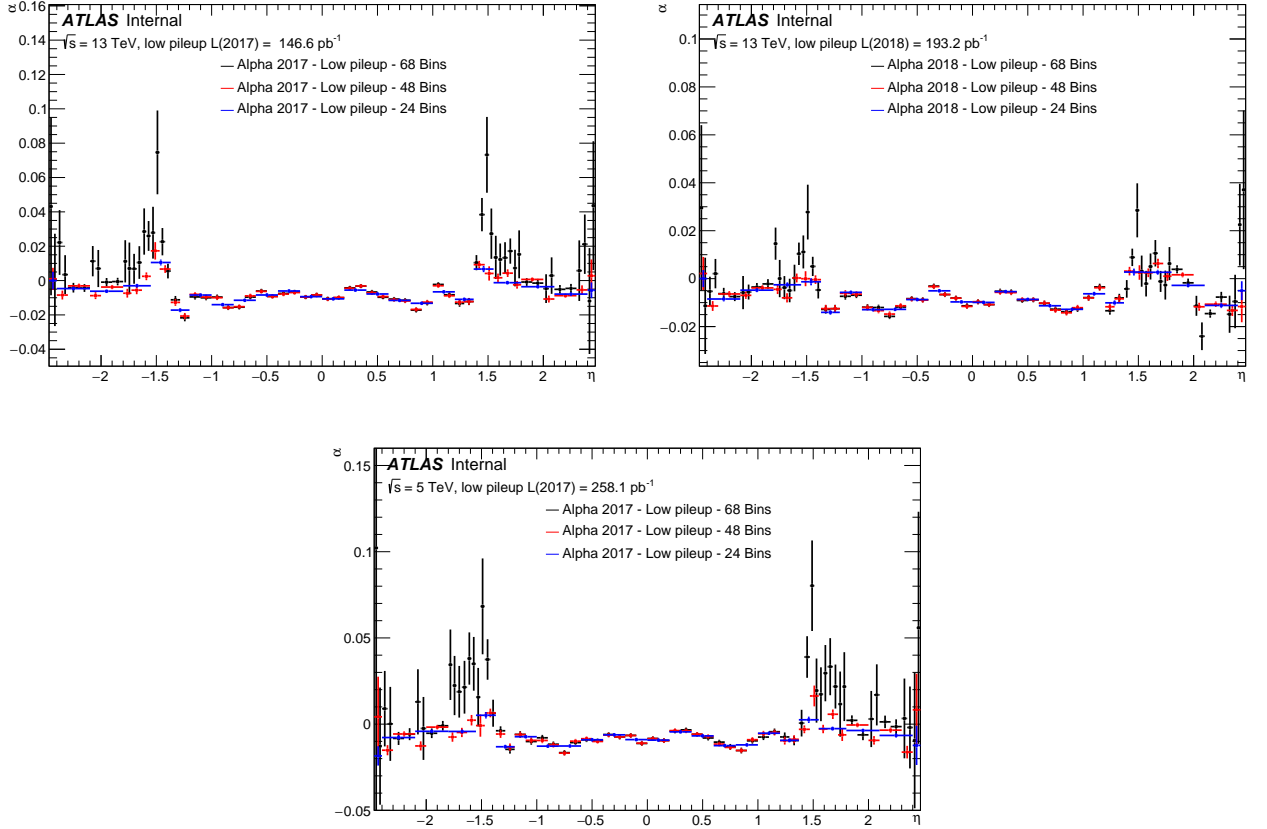


**Figure 7.2:** Energy scale factors $\alpha$ for low pile-up runs of 2017 (left), 2018 (right) and 2017 at 5TeV (bottom) using 68, 48 and 24 $\eta$ bins. It can be seen, that the extraction is unstable in case of 68 bins, resulting in $\alpha$ factors with very large uncertainties [138].

The extracted constant $c_i'$ correction term is presented in Fig. 7.3. The constant term $c'$ depends on the data taking conditions and pile-up, so its extrapolation from a dataset obtained under different conditions appears problematic. This issue is discussed in Ref. [139].

This correction entails experimental uncertainty, caused primarily by the statistical uncertainty of $\alpha_i$ and $c_i'$ factors measurement, other uncertainties are significantly smaller and therefore neglected. The comparison between the data and MC invariant mass distributions around the $Z \rightarrow ee$ peak are
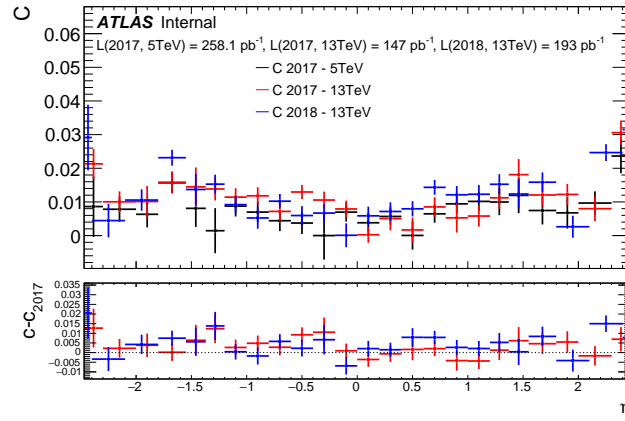
**Figure 7.3:** Additional constant term $c_i'$ for low pile-up runs of 2017 (13 TeV), 2018 (13 TeV) and 2017 at 5 TeV using 24 bins. The lower panel shows the difference of $c_i'$ to the 2017 5 TeV run [138].

presented in Figure 7.4 and Fig. 7.5. The agreement is satisfactory and within the uncertainty.
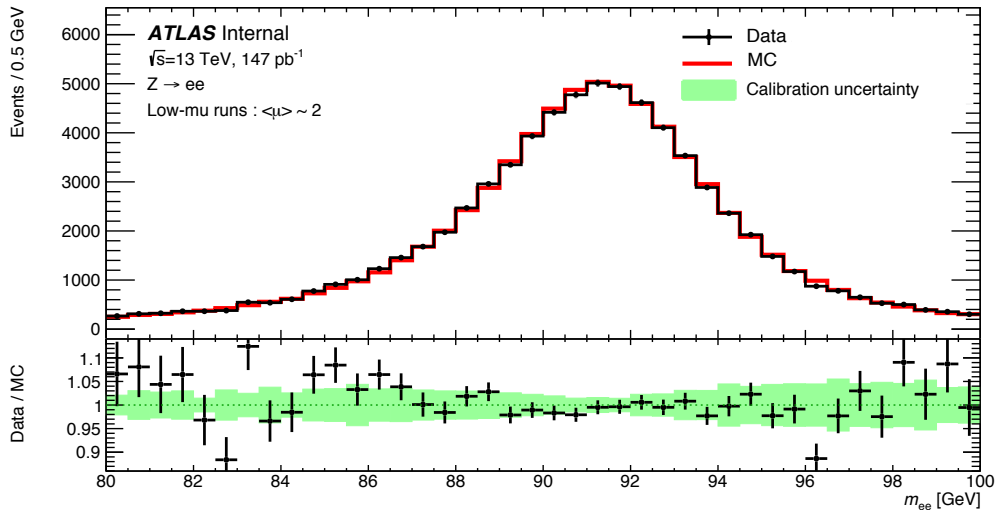


**Figure 7.4:** The invariant mass distribution around the Z-mass for low pile-up Data for 2017 (13 TeV) [138]
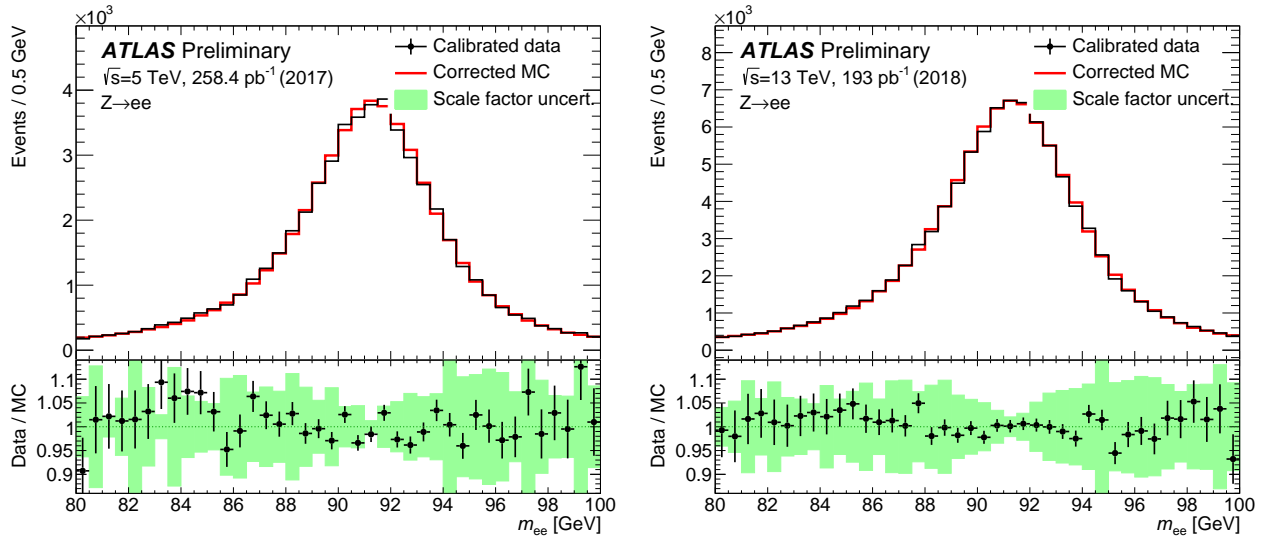
**Figure 7.5:** The invariant mass distribution around the Z-mass for low-pileup data for the $\sqrt{s} = 5\,\text{TeV}$ data taken in 2017 (a) and the $\sqrt{s} = 13\,\text{TeV}$ 2018 data (b) [138].

### 7.1.2 Electron selection efficiency

The electrons used in the analysis are selected based on the corresponding requirements to the quality of their reconstruction, kinematic characteristics, and passing certain identification, isolation and trigger matching criteria. A tag-and-probe method is used to measure efficiencies in data and MC simulation, which may be different due to various aspects of physics and detector modelling (some of them are addressed in Chapter 5 of the current thesis). In order to match the MC simulation and the data in each of the aforementioned aspects the corresponding scale factor (SF)s are introduced. The SF is defined as the ratio of the data efficiency to MC efficiency:

$$SF_{(a)} = \frac{\epsilon^{data}_{(a)}}{\epsilon^{MC}_{(a)}},$$

where $\epsilon$ stand for efficiency and index $a$ stands for reconstruction, ID, isolation or trigger. The SF extraction allows for better agreement between data and simulation, but also brings uncertainties. The total efficiency correction is used as an event weight during the analysis:

$$W^{W \rightarrow ev}_{event} = SF_{reco} \cdot SF_{trig} \cdot SF_{ID} \cdot SF_{iso}.$$

The tag-and-probe method used for the measurement of electron efficiencies includes the following steps:

- A kinematic selection is applied to $Z \rightarrow ee$ events (Cut1).

- A tight selection (Cut2) is applied to one of the two electrons along with matching it to the single-electron trigger. From now on this electron is called the *tag*.

- The other electron is called the *probe* and is used to probe the picked efficiency.

- Selections Cut1 and Cut2 are varied in order to evaluate the stability of the method.

The details are presented in Refs. [132, 125, 128].
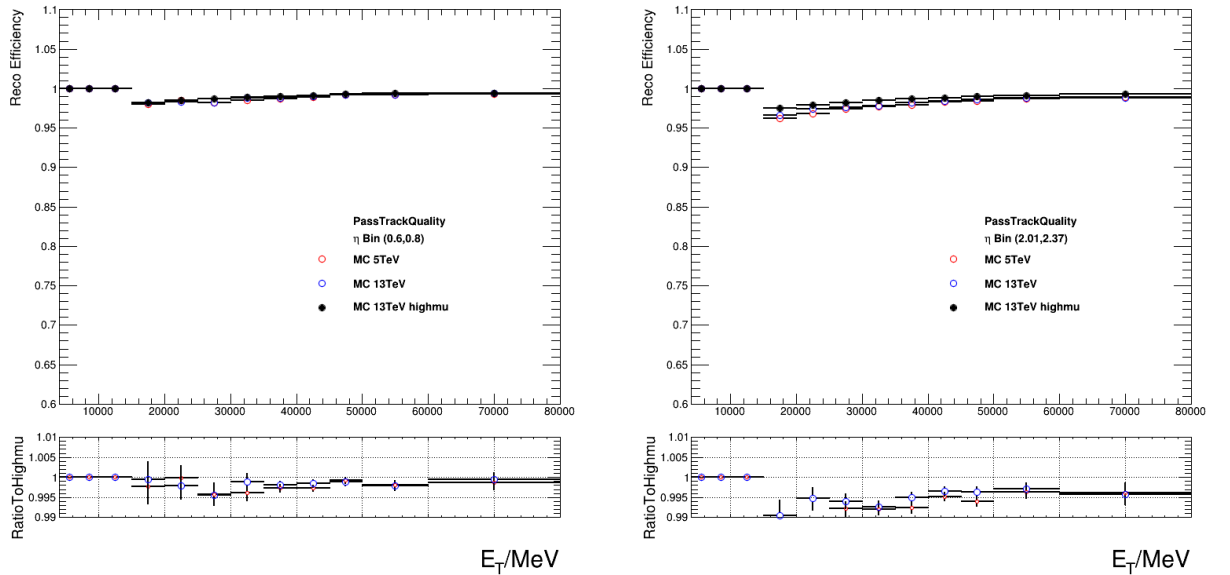
#### Reconstruction efficiency

The fraction of electromagnetic clusters that are matched with the charged particle track from the ID that satisfy the quality requirements define the reconstruction efficiency. An additional "PassTrackQuality" requirement of having at least 1 hit in the pixel detector and and least 7 hits in the silicon detectors is imposed on the track of successfully reconstructed electrons.
So the electron reconstruction efficiency is calculated as:

$$\epsilon^{reco}(p_T, \eta) = \frac{N_{pass} - N^{bkg}_{pass}}{N_{pass} - N^{bkg}_{pass} + N_{fail} - N^{bkg}_{fail} + N_{photon} - N_{fit}}. \tag{7.1}$$

$N_{pass(fail)}$ stands for the number of electrons passing (failing) the "PassTrackQuality" criterion. The background electron candidates $N^{bkg}_{pass(fail)}$ are obtained from the template fits of the background on subsets that pass (fail) the "PassTrackQuality" criterion. The number of electrons that are reconstructed as photons is denoted by $N_{photon}$. The non-electron background that is reconstructed as photons is estimated from analytical fit in the control region away from the $Z \rightarrow ee$ resonance and is called $N_{fit}$. An extrapolation of $SF_{reco}$ from the high-$\mu$ data is used as a baseline for the reconstruction scale factors measurement. The benefits of higher statistics available in high-$\mu$ dataset outweigh the losses imposed by the extrapolation and provide lower uncertainty comparing to the SFs measured with low pileup dataset (see Fig. 7.7).

The extrapolation procedure takes into account the two main differences between the high-mu and low-mu datasets, namely the dependence of the efficiency on the $N_{Vtx}$ per event and a different topocluster threshold. Fig. 7.6 shows that the dependence on the number of vertices per event is well modelled by the MC simulation and can be extrapolated using a fit. The dependence on the topocluster threshold was found negligibly small.



**Figure 7.6:** Electron reconstruction efficiencies compared for low-pileup and high-pileup MC16 in $\eta$ range [0.60-0.80] (left) and [2.01-2.37] (right). The statistical uncertainties are shown [138].
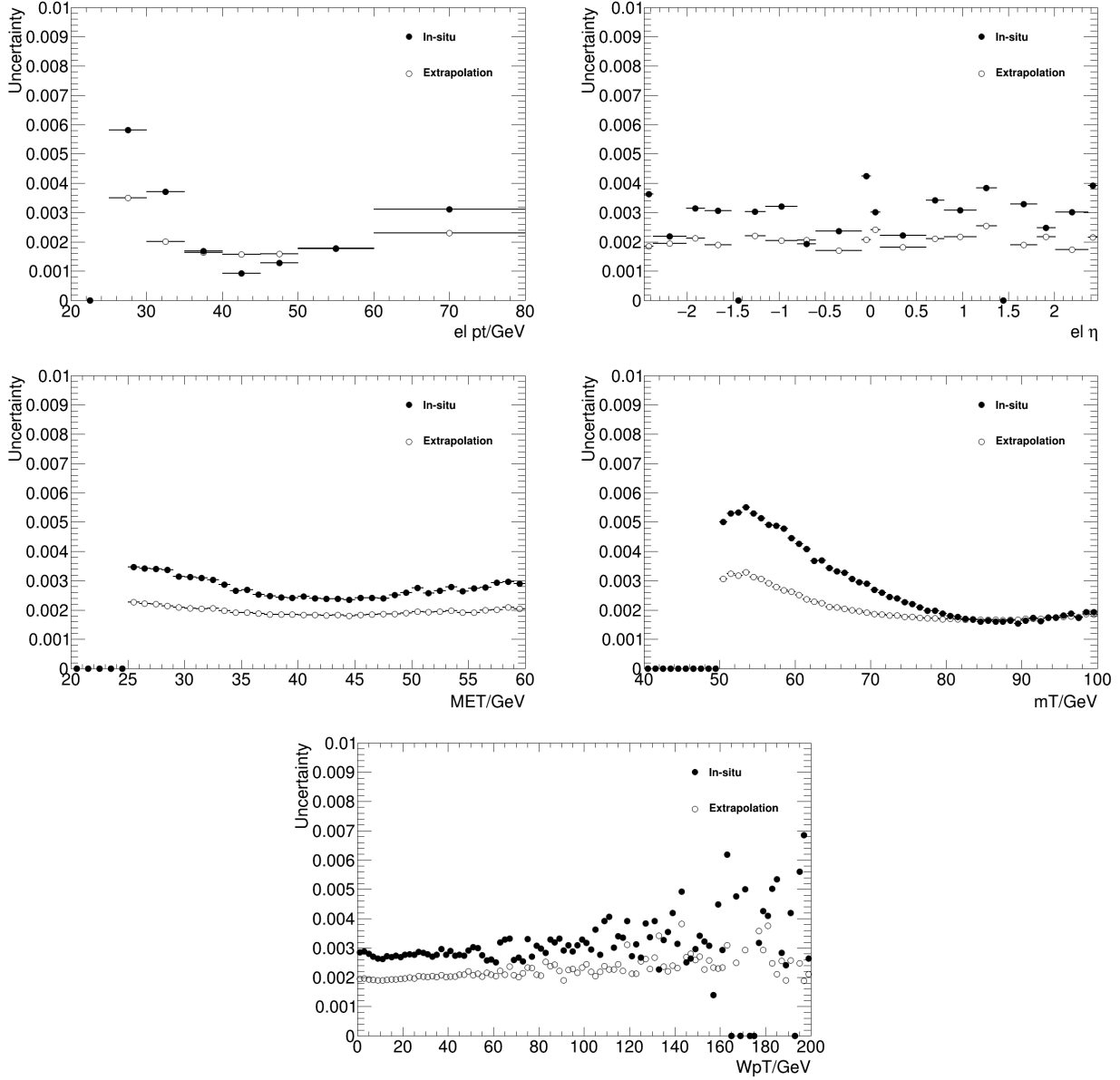
**Figure 7.7:** Comparison of the uncertainties due to electron reconstruction, contrasting the high-$\mu$-extrapolated and in-situ-measured SF uncertainties in a $W^+ \rightarrow e^+\nu$ selection at 13 TeV as function of typical kinematic variables [138].

## Identification efficiency

The fraction of reconstructed electrons that pass a given working point define the electron identification efficiency. The low pile-up $W \rightarrow e\nu$ measurement uses the "Medium LH" working point. The methodology is described in Ref. [125] and includes the combination of two background subtraction methods: *Zmass* and *Ziso*.

In the Zmass method the background is estimated using a template method normalized in $m_{ee}$ side bands. The tag is required to be trigger-matched, pass ID and isolation cuts and have $p_T > 20$ GeV. The probe has to pass the "PassTrackQuality" and the electron/photon ambiguity resolver, have $p_T > 15$ GeV and be separated from jets with $p_T^{jet} > 20$ GeV by $\Delta R > 0.4$.

An alternative method, *Ziso*, uses the calorimeter energy isolation $E_T^{cone}$ of the probe electron to discriminate between background and signal. Signal electrons are expected to have better isolation than the background. On top of the requirements used for the Zmass method the tag and probe pair is required to have opposite sign and to fit into mass window of 15 GeV around the Z boson mass. The background template shape is constructed from the probe electrons that have the same charge as the tag, pass the track quality criteria but fail the shower shape cuts or fail the cut-based loose identification. The fraction of real electrons that pass the described selection is modelled with MC and subtracted from the template. The background template uses the tail region of probe isolation distribution $E_T^{cone0.3}/25\text{GeV} > 0.5$ is scaled to the data luminosity. An example of the probe isolation estimate for the numerator and denominator in eq. 7.1 in the region $25 < E_T < 30$ and $0.8 < \eta < 1.15$ is presented in Fig. 7.8.
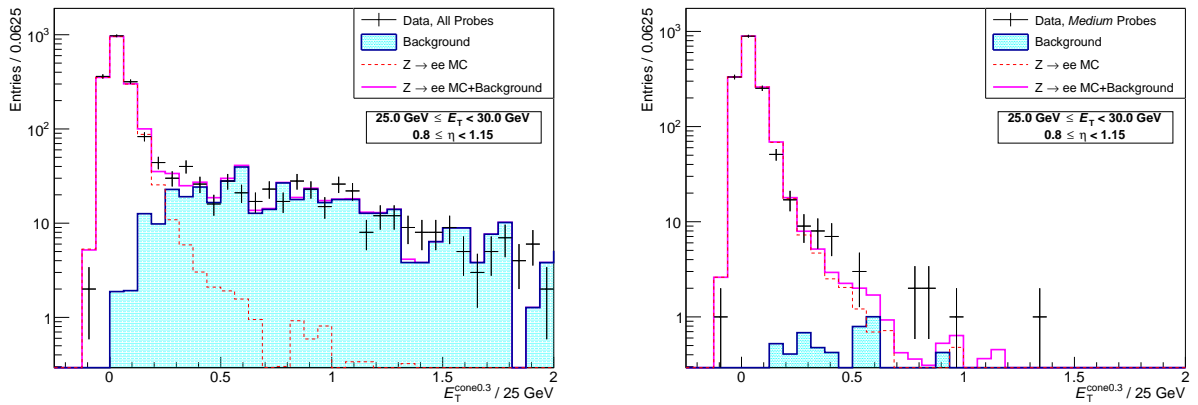


**Figure 7.8:** $E_T^{cone0.3}/25\,\text{GeV}$ isolation distribution of probe electrons in the ZIso-method using 13 TeV $339\,\text{pb}^{-1}$ low-pileup data. Left the denominator and right the numerator distributions are shown, with the data as crosses, the signal $Z \rightarrow ee$ expectation as open line and the background estimate as cyan area (template normalised at high values) [138].

The SF to be used in the analysis is constructed out of both methods. The combination of the results takes into account the high degree of correlation between the two methods and includes the following steps:

- the final SF is defined as an arithmetic mean of the two methods over all systematic variations;

- the statistical uncertainty is calculated as the average of the statistical uncertainties of the variations;

- a covariance matrix is composed from all variations of the two methods and then decomposed into correlated and uncorrelated parts, providing the systematic uncertainty.

The combined results are presented in Fig. 7.9 and show similar results between both methods and the combination. The SFs obtained from 5 and 13 TeV data samples were not combined due to significant differences in measured efficiency.
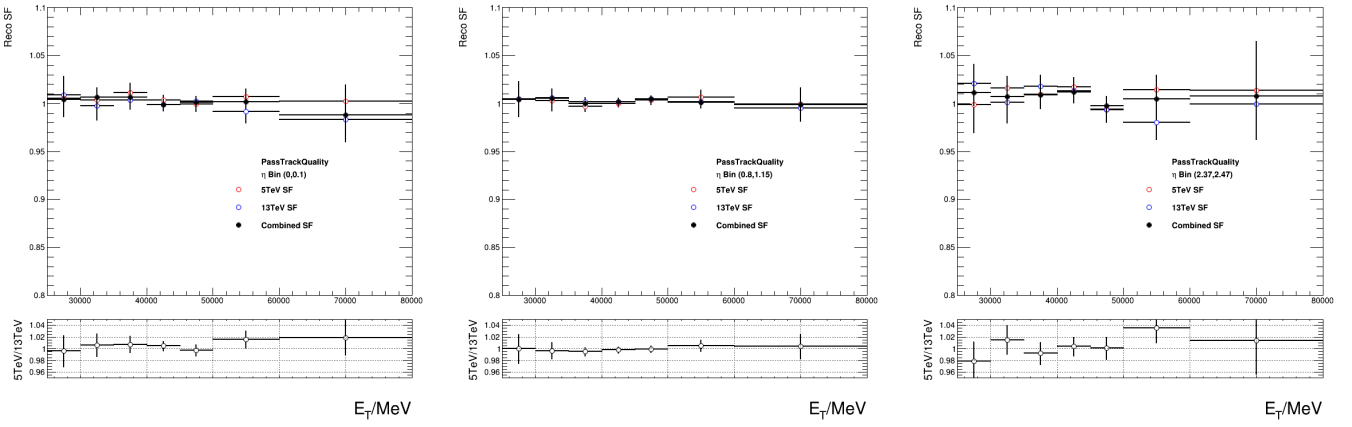


**Figure 7.9:** Comparison of electron reconstruction SFs with 5 TeV and 13 TeV data as well as the 5+13 TeV, selected ranges from most central $\eta = 0 - 0.1$ (left) to most forward $\eta = 2.37 - 2.47$ (right). The bottom panel shows the ratio of 5 TeV and 13 TeV SFs. The total uncertainties are shown [138].

### Isolation efficiency

The electron isolation efficiency is the fraction of reconstructed and MediumLLH-identified electrons that pass a designated isolation requirement. For this analysis the isolation requirement is chosen to be $ptvarcone^{20}/p_T^e < 0.1$. The results are presented in Fig. 7.10 and show that the efficiency is very high. The SFs for 5 and 13 TeV are not combined and used separately.

### Trigger efficiency

During the data-taking at low pile-up the unprescaled trigger HLT_e15_lhloose_nod0_L1EM12 was used. Thanks to the ID and isolation requirements for both tag and probe electrons, the background is negligible in the trigger efficiency measurement. The measurement results are shown in Fig. 7.11 and demonstrate relatively high efficiency in most of the regions. The scale factors are also very close to unity. No combination was performed between 5 and 13 TeV results.
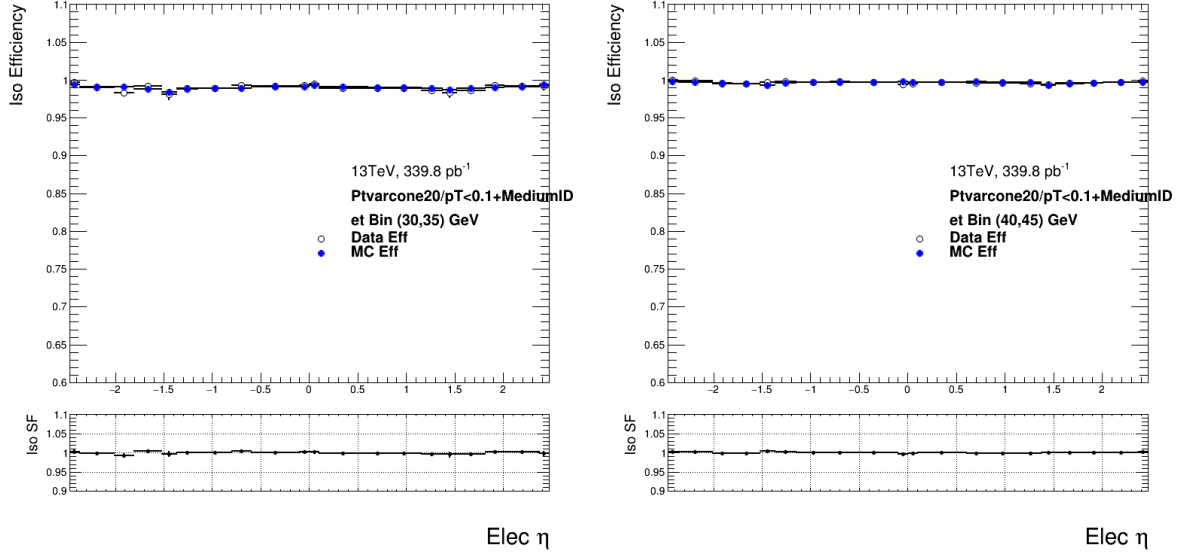
**Figure 7.10:** Electron isolation efficiencies (top panels) and scale factors (lower panels) for the *pt-varcone20/$p_T^e$ < 0.1* working point using 13 TeV 339 pb$^{-1}$ low-pileup data as function of $\eta$ in bins of $p_T$ [138].
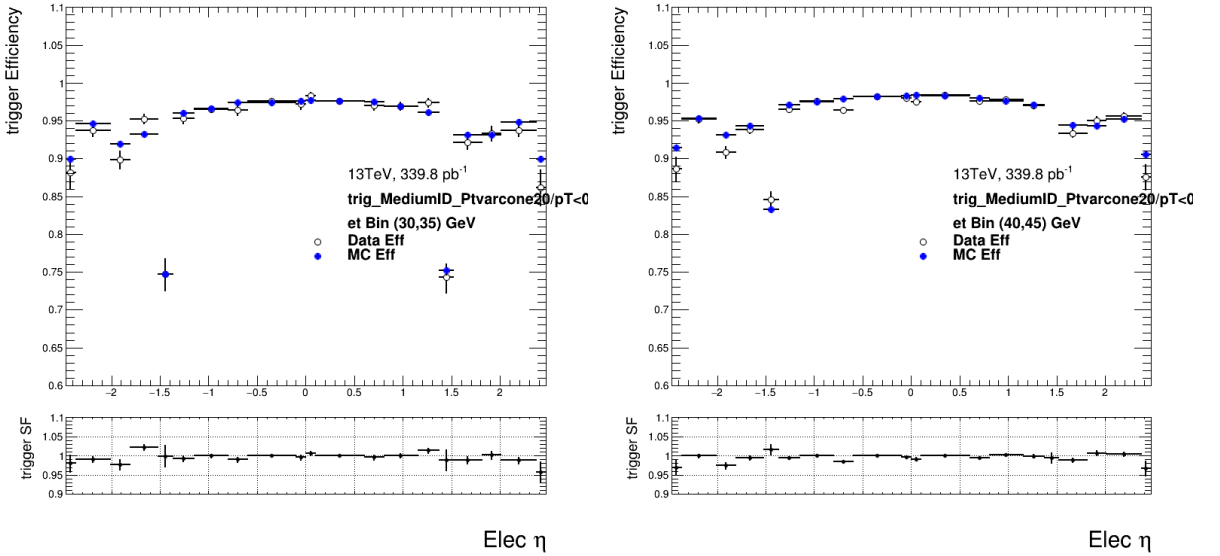


**Figure 7.11:** Electron trigger efficiencies (top panels) and scale factors (lower panels) for HLT_e15_lhloose_nod0_L1EM12 using 13 TeV 339 pb$^{-1}$ low-pileup data as function of $\eta$ in bins of $p_T$ [138].

### 7.1.3 SF uncertainties propagation

The main source of uncertainty for the measurement of the SFs is coming from the background. The uncertainties are estimated by varying the parameters that contribute to background suppression. These parameters include the following variations:

- The Zmass window technique is used in the identification, isolation and trigger efficiencies measurement. The size of the Zmass window was varied in a range of 10, 15 and 20 GeV. This variation dominates at higher values of $p_T$.

- The tag identification and isolation criteria were varied between Medium ID + calorimeter isolation, TightLLH and Tight ID + calorimeter isolation.

- The background template has a major influence on the estimate of signal contamination, especially at $p_T < 30$ GeV. In addition to the nominal range of template extraction in $120 < m_{ee} < 250$ GeV the templates are also normalized using the region of $60 < m_{ee} < 70$ GeV.

- The side band range is varied for the reconstruction efficiency measurement.

- The isolation criteria are varied in the measurement of ID efficiency: $E_T^{cone0.3}/25$ GeV is varied between 0.4, 0.5 and 0.6, also a larger cone isolation around the probe electron was used - $E_T^{cone0.4}/25$ GeV.

Figure 7.12 shows the total relative uncertainties of the electron scale factors at 13 TeV in different $\eta$ bins. Contributions from reconstruction and identification are the dominant ones. The uncertainties are propagated to the observables using the co-called Full correlation model (see Ref. [128]). The idea of the method is to split the sources of SF uncertainty into uncorrelated and correlated sources. Uncorrelated sources are of statistical nature and mostly related to the number of $Z \to ee$ pairs in different bins if $p_T$ and $\eta$ used for SF extraction. Correlated sources of systematic uncertainty arise from the flaws of background subtraction. The Full correlation model includes about 10 sources of systematic uncertainty and around 200 $p_T \times \eta$ bins as sources of statistical uncertainty and allows to propagate these uncertainties to the observables. Figures 7.13 and 7.14 contain the results of error propagation to several observable distributions for 13 and 5 TeV respectively. Again, identification and reconstruction uncertainties have the largest contribution to the total SF uncertainty.
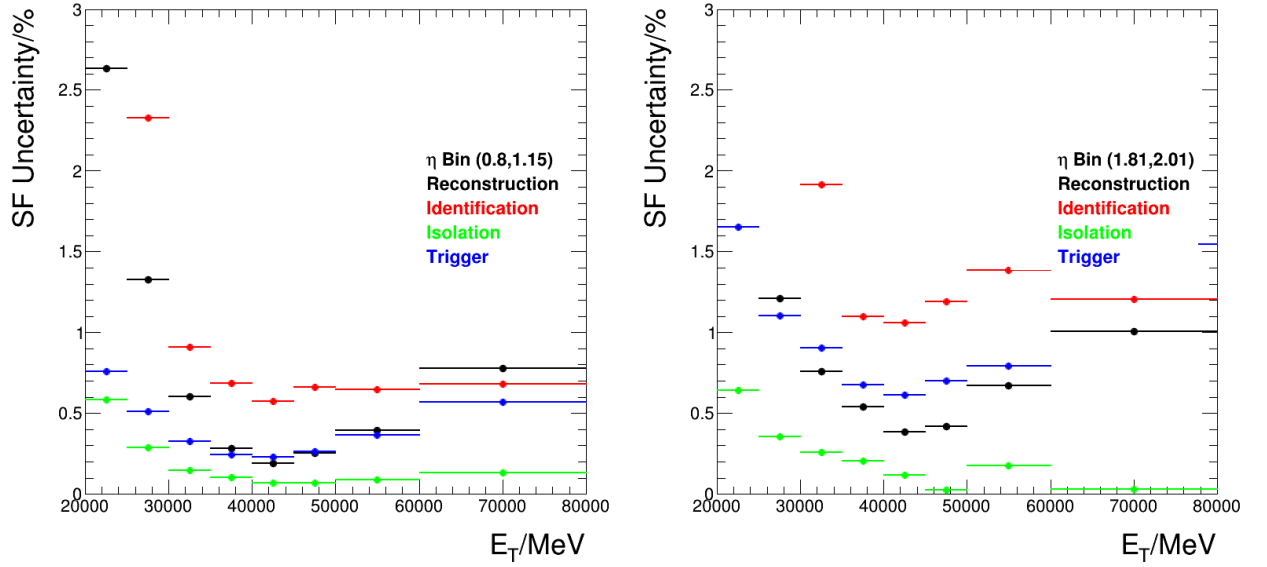
**Figure 7.12:** Examples of total relative uncertainties of electron scale factors at 13 TeV measured with tag-and-probe method [138] in the barrel ($|\eta| = 0.8 - 1.15$) and the end-cap ($|\eta| = 1.81 - 2.01$).
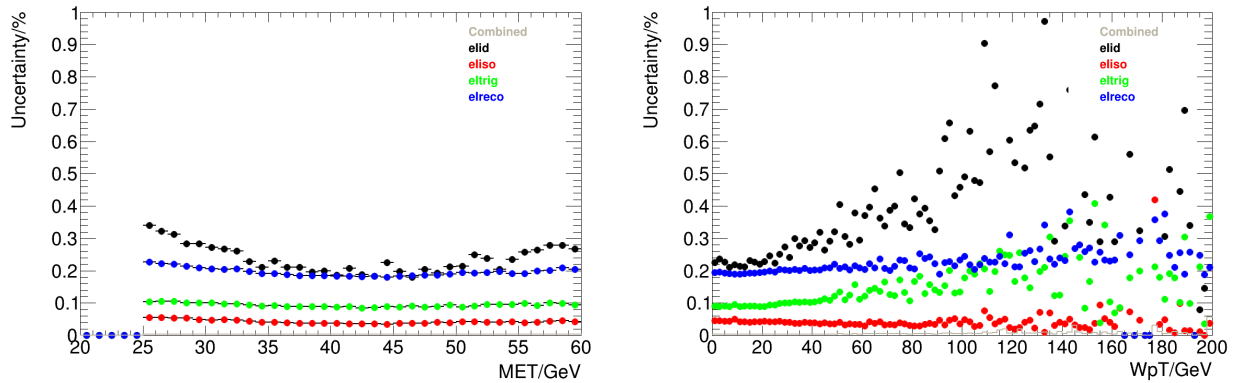


**Figure 7.13:** Examples of contributions to the electron uncertainties related to efficiency SF (reconstruction, identification, isolation and trigger) in a $W^+ \rightarrow e^+ \nu$ selection at 13 TeV as function of MET and W pT [138].
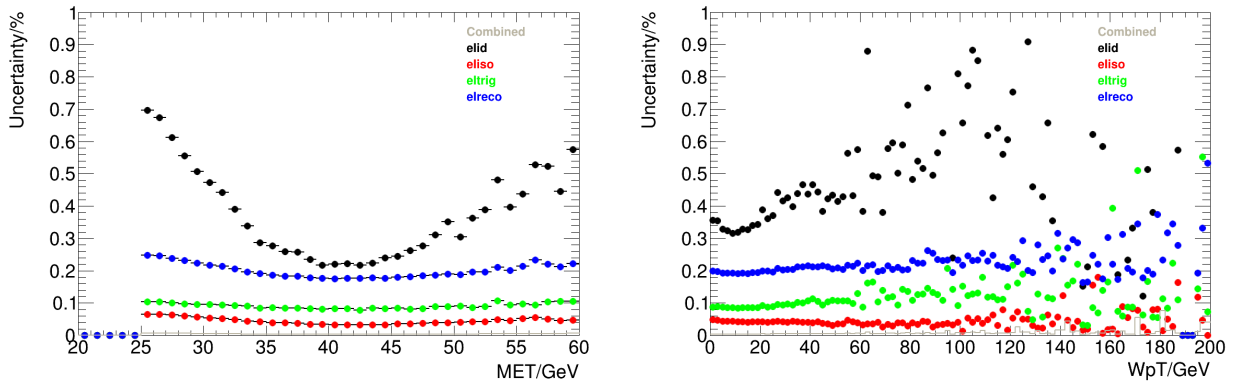
**Figure 7.14:** Examples of contributions to the electron uncertainties related to efficiency SF (reconstruction, identification, isolation and trigger) in a $W^+ \rightarrow e^+ \nu$ selection at 5 TeV as function of of MET and W pT [138].

## 7.2 Muon corrections

Muon corrections are in many aspects similar to the electron corrections described in the previous section. Calibrations are used in order to match the momentum response in the data and in the MC simulation.

### 7.2.1 Muon momentum calibration

The muon momentum calibration comprises corrections to the momentum scale and resolution. At first the ID and MS tracks are reconstructed and corrected separately, and then the two corrections are propagated to correct the CB muon tracks. Low energy muons with $5 < p_T < 30$ GeV are calibrated using the $J/\psi \rightarrow \mu\mu$ resonance, while in higher energy region of $22 < p_T < 300$ GeV the $Z \rightarrow \mu\mu$ resonance is used. The statistical uncertainties are directly linked to the number of $Z$ and $J/\psi$ candidates in the data samples:

- 5.02 TeV data (2017, period M): 660k $J/\psi$ candidates, 75k $Z$ candidates

- 13 TeV data (2017, period N): 1.1M $J/\psi$ candidates, 100k $Z$ candidates

- 13 TeV data (2018, periods G4 and J): 1.5M $J/\psi$ candidates, 130k $Z$ candidates

The corrected transverse momentum $p_T^{corr,Det}$ (where $Det$ may stand for ID or MS) is described by the following equation [122]:

$$p_T^{corr,Det} = \frac{p_T^{MC,Det} + \sum_{n=0}^{1} s_n^{Det}(\eta,\phi)\left(p_T^{MC,Det}\right)^n}{1 + \sum_{m=0}^{2} \Delta r_m^{Det}(\eta,\phi)\left(p_T^{MC,Det}\right)^{m-1} g_m}, \qquad (7.2)$$

where $p_T^{MC,Det}$ is the uncorrected $p_T$ in MC simulation, $g_m$ are normally distributed random variables with zero mean and unit width, $s_n^{Det}$ and $\Delta r_m^{Det}$ are the momentum resolution smearing and scale correction factors applied in the specific $(\eta,\phi)$ detector region respectively. The Z and $J/\psi$ peaks are fitted with a function that is a sum of a Crystal Ball function (that fits the mass peak), a Gaussian (that accounts for effects like multiple scattering) and an exponential that fits the backgrounds. The examples of such fits are presented at Fig. 7.15.

The measured correction parameters for Z and $J/\psi$ peaks for 5 and 13 TeV are presented in Fig. 7.16, 7.17, 7.18 and 7.19. The calibration and resolution corrections were obtained at high pile-up conditions and validated in Z and $J\psi$ for low pile-up.
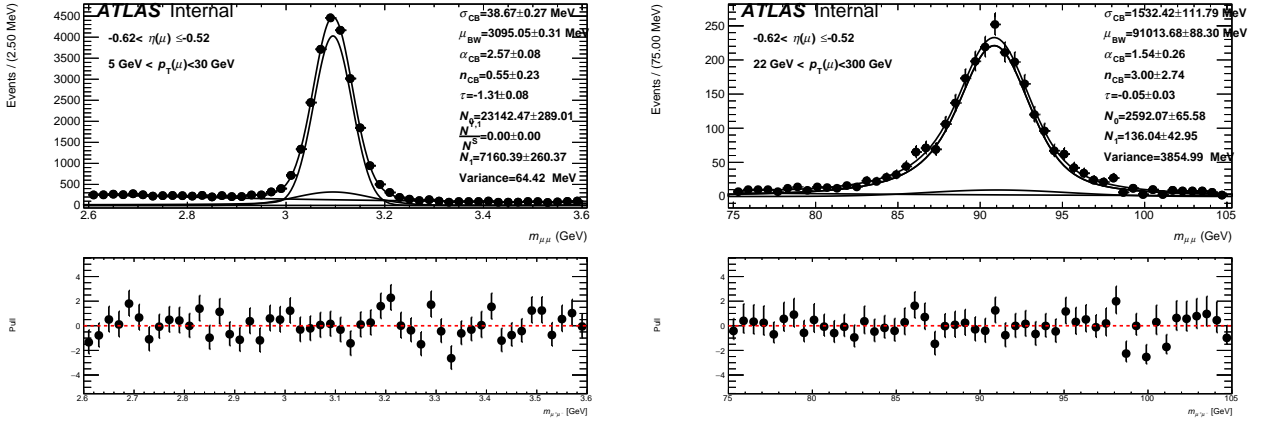
**Figure 7.15:** Example fits to $J/\psi \to \mu\mu$ (left) and $Z \to \mu\mu$ (right) mass peaks for pairs with leading muon pseudorapidity in the range $-0.62 < \eta < -0.52$ in low-pile-up 2017 13 TeV data [140].
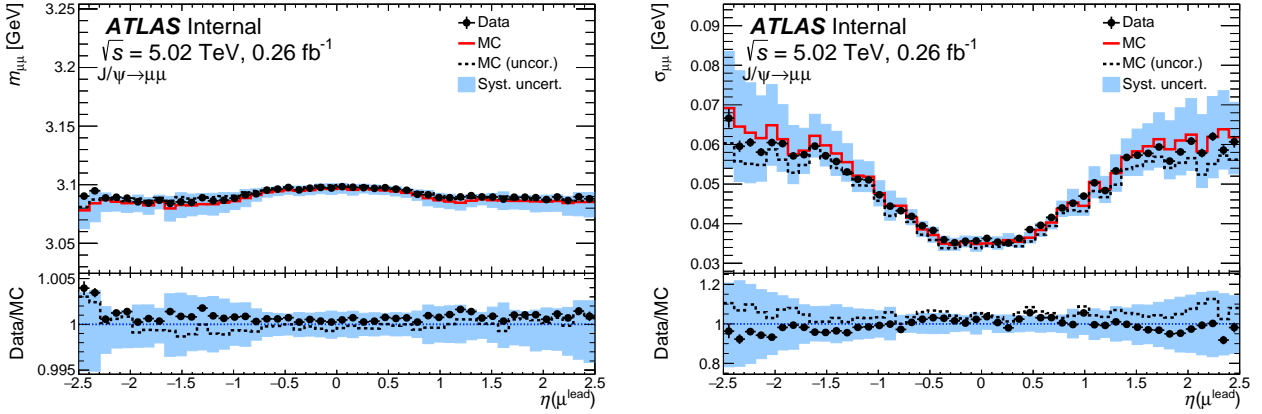


**Figure 7.16:** Mean (left) and width (right) of the $J/\psi \to \mu\mu$ mass peak as a function of the leading muon $\eta$ in 5.02 TeV data and MC. The mean and width are extracted from Crystal Ball components of the fits. In case of the simulation, both the uncorrected (dashed histogram) and corrected parameters (solid histogram) are shown. The fit results are presented for mass peaks constructed using kinematics of the muon CB tracks. The bottom panels in each plot show the data/MC ratio for uncorrected (dashed histogram) and corrected simulation (points) [140].

### 7.2.2 Correction for charge-dependent momentum bias

Misalignment in ID, MS or between the two systems can lead to a charge-dependent bias (also called *sagitta bias*) of muon reconstructed momentum. Its effect can be parametrized as follows:

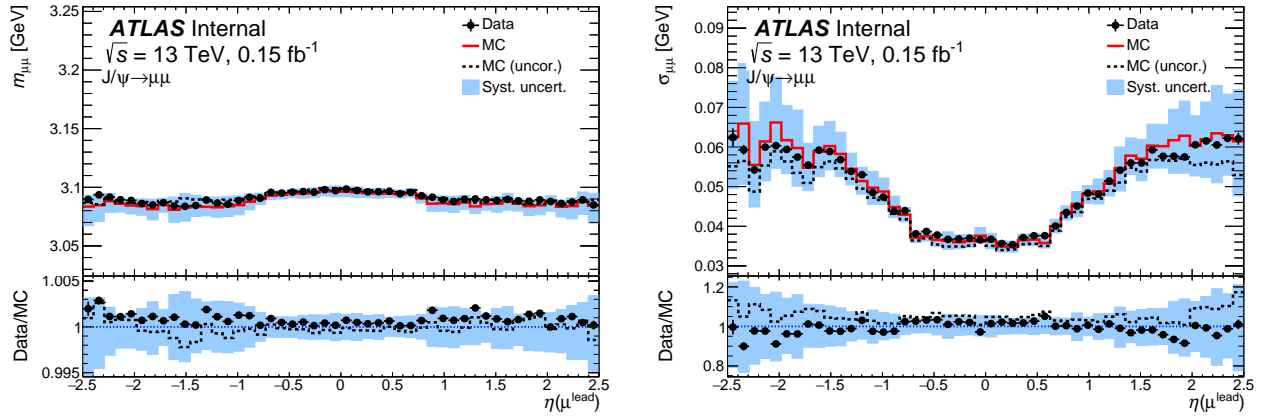$$p_T^{meas} = \frac{p_T^{reco}}{1 + q \cdot \delta_{sagitta} \cdot p_T^{reco}}, \tag{7.3}$$

**Figure 7.17:** Mean (left) and width (right) of the $J/\psi \to \mu\mu$ mass peak as a function of the leading muon $\eta$ in 2017 13 TeV data and MC at low pile-up. The mean and width are extracted from Crystal Ball components of the fits. In case of the simulation, both the uncorrected (dashed histogram) and corrected parameters (solid histogram) are shown. The fit results are presented for mass peaks constructed using kinematics of the muon CB tracks. The bottom panels in each plot show the data/MC ratio for uncorrected (dashed histogram) and corrected simulation (points) [140].
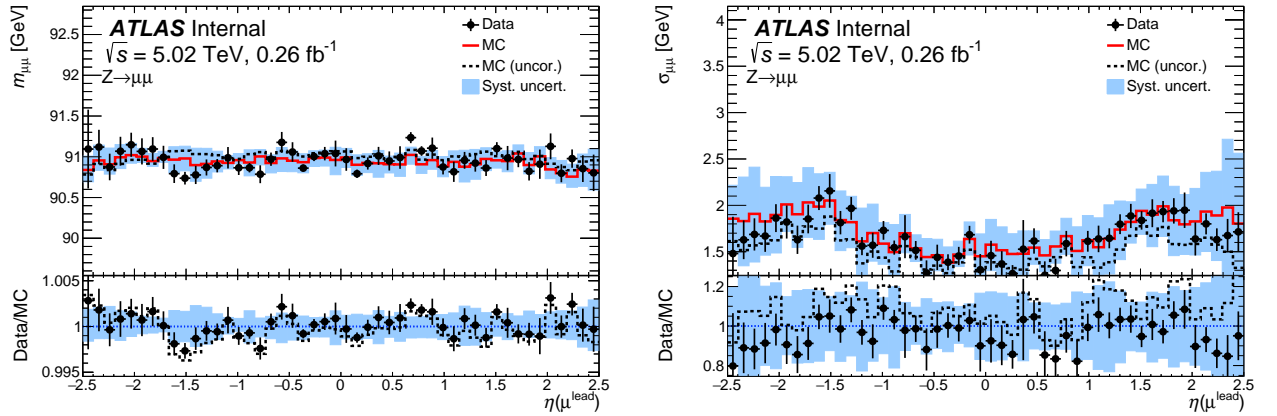


**Figure 7.18:** Mean (left) and width (right) of the $Z \to \mu\mu$ mass peak as a function of the leading muon $\eta$ in 5.02 TeV data and MC. The mean and width are extracted from Crystal Ball components of the fits. In case of the simulation, both the uncorrected (dashed histogram) and corrected parameters (solid histogram) are shown. The fit results are presented for mass peaks constructed using kinematics of the muon CB tracks. The bottom panels in each plot show the data/MC ratio for uncorrected (dashed histogram) and corrected simulation (points) [140].

where $p_T^{meas}$ is the measured momentum that contains a bias, $p_T^{reco}$ is the unbiased reconstructed momentum and the bias to be corrected is denoted as $\delta_{sagitta}$.

The sagitta correction is obtained as a function of $\eta$. There exist three methods of sagitta bias determi-
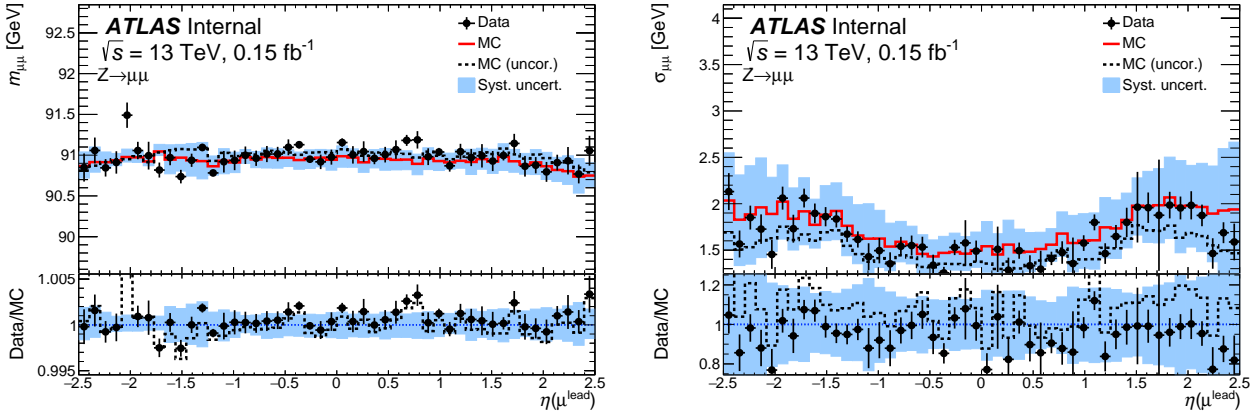
**Figure 7.19:** Mean (left) and width (right) of the $Z \to \mu\mu$ mass peak as a function of the leading muon $\eta$ in 2017 13 TeV data and MC at low pile-up. The mean and width are extracted from Crystal Ball components of the fits. In case of the simulation, both the uncorrected (dashed histogram) and corrected parameters (solid histogram) are shown. The fit results are presented for mass peaks constructed using kinematics of the muon CB tracks. The bottom panels in each plot show the data/MC ratio for uncorrected (dashed histogram) and corrected simulation (points) [140].

nation:

- The ID alignment and momentum measurement can be tested with electrons which have additional information from the charge-independent calorimeter. Using $Z \to ee$ and/or $W \to e\nu$ events it is possible to determine the charge bias of an electron track:

$$\delta_{sagitta} = \frac{< E/p_{track} >^+ - < E/p_{track} >^-}{2 < p_T^{calo} >},\tag{7.4}$$

where $p_{track}$ is the momentum measured in the ID, E is the energy measured in the EMC, from which we can get the $p_T^{calo} = E \cdot \sin\theta$ transverse momentum. The <> brackets denote the averaging in $\eta$ bins.

- The $p_T(\mu)$ method is used by muon combined performance (MCP) and alignment groups in high-$\mu$ data compares the muons and anti-muons spectra in $Z \to \mu\mu$ events. It compares the $p_T^{\mu+}$ and $p_T^{\mu-}$ distributions in data and MC and interprets the discrepancy as a charge-dependent misalignment.

- The Z-mass method uses the Z mass peak. This is the main method used by MCP and alignment groups, the results are denoted as $M_{\mu\mu}\_MCP$ and $M_{\mu\mu}\_Align$ on the plots (see [141]). The sagitta bias is calculated iteratively minimizing the difference between the reconstructed and expected position of the $Z \to \mu\mu$ mass peak position. This method catches the rapidity dependence of the sagitta bias, but is sensitive to overall offset.

The results of these methods are presented in Fig. 7.20. The sagitta bias in low-pile-up data was found to be 10 times higher than in the MC simulation. Further results in the measurement rely on the following method: the differences between the data and MC are averaged over $\eta$ using a fit, and also introduce a global offset. The results are shown in Fig. 7.21.
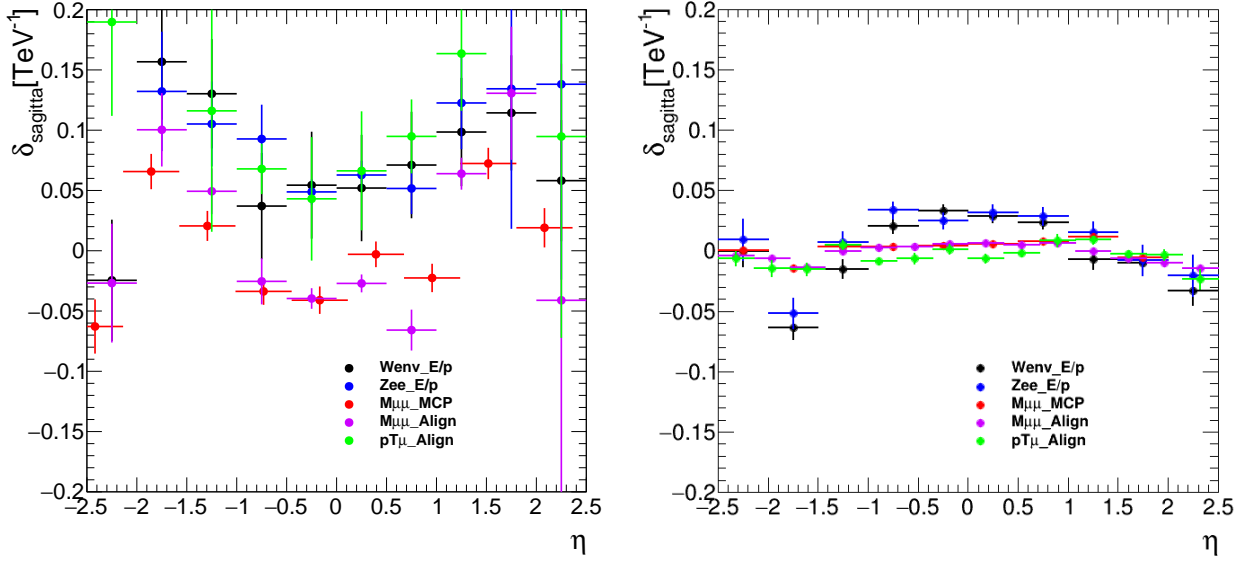
**Figure 7.20:** Sagitta bias corrections derived for 2017 low-pile-up data (left) and simulation (right) at $\sqrt{s} = 13$ TeV. The corrections are evaluated with two $Z$-mass methods ("$M\mu\mu$_MCP" and "$M\mu\mu$_Align"), the $E/p$ method applied to $W \to e\nu$ ("Wenv_E/p") and $Z \to ee$ ("Zee_E/p") events, and the $p_{\mathrm{T}}(\mu)$ method, all of which are discussed in the text [140].
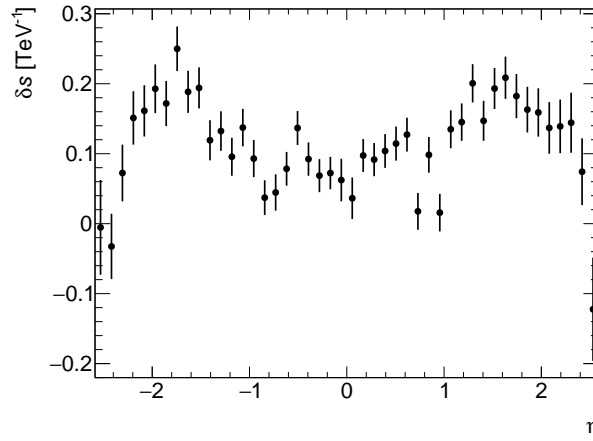


**Figure 7.21:** Sagitta bias correction based on 2017 low-pile-up data at $\sqrt{s} = 13$ TeV. The statistical uncertainty (combined from uncertainties of the $\eta$-dependent correction and the global offset correction) is represented by error bars [140].

### 7.2.3 Muon efficiency measurements

Just like in the case of electrons, muons have to pass a number of quality criteria in order to be used in the analysis:

- Reconstruction and identification: the muon is successfully reconstructed and its ID and MS tracks as well as EMC deposit are matched. The *medium* identification criterion is adopted for the low-$\mu$ analysis. Only CB and ME muons with loose requirements between the tracks from ID and MS are used. The value of $q/p$ significance is required to be $< 7$.

- Isolation: track isolation requirement of $p_T^{varcone20}/p_T < 0.1$ was selected as an isolation criterion.

- Trigger: the muons were required to pass the $HLT\_mu14$ trigger.

- Track-to-vertex association (TTVA) includes requirement for the muon track to match the primary vertex. The muon objects are required to pass $|z_0|\sin\theta < 0.5$ mm and $d_0/\sigma(d_0) < 3$ requirements.

Just like in the case of electrons, possible discrepancies between the data and MC are corrected using the scale factors, which are in turn measured using the tag-and-probe method described in [142]. The product of the scale factors define the event weight:

$$W_{event}^{W \to e\nu} = SF_{reco/ID} \cdot SF_{trig} \cdot SF_{TTVA} \cdot SF_{iso}.$$

All muon efficiencies and scale factors used in current analysis are measured in-situ using the low-$\mu$ datasets at 5 and 13 TeV by the MCP group. The results for their measurements are presented in Fig. 7.22 and 7.23.
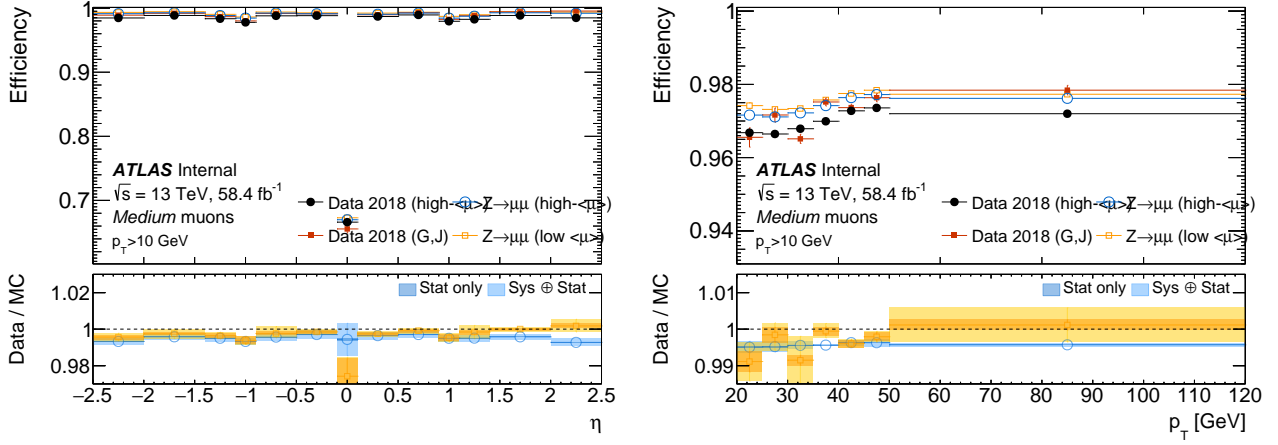
**Figure 7.22:** Comparison of reconstruction efficiencies for Medium muons using the low-$\mu$ run of 2018 at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. Efficiencies are shown as a function of muon $\eta, p_\text{T}$. Red (orange) points correspond to low-$\mu$ data (MC), while the black (blue) points are high-$\mu$ data (MC). The bottom panels show the data/MC ratio for the low-$\mu$ (orange) and high-$\mu$ (blue) sets with statistical and total uncertainties [140].
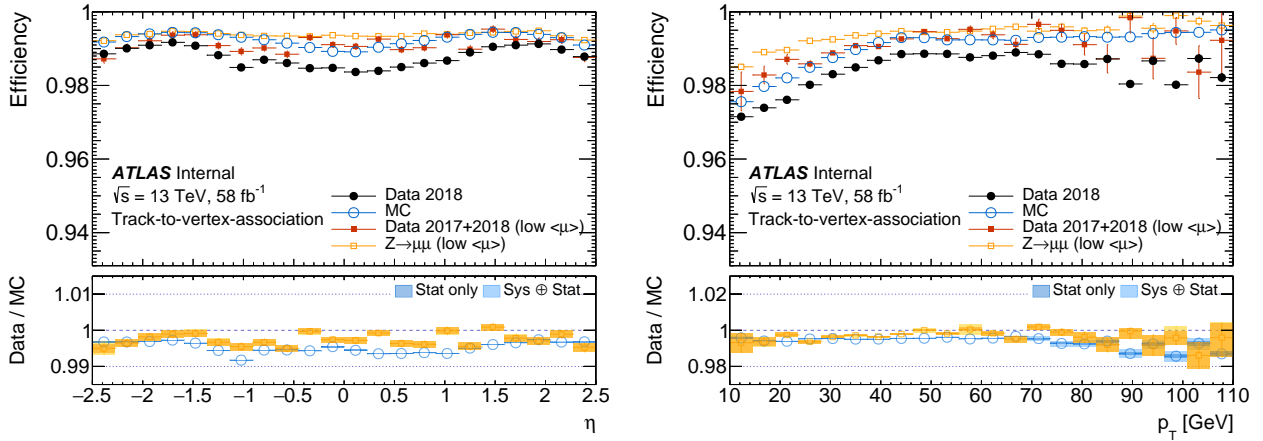


**Figure 7.23:** Comparison of TTVA efficiencies for Medium muons using the low-$\mu$ runs of 2017+18 at $\sqrt{s} = 13\,\text{TeV}$. The low-$\mu$ results compared to a high-$\mu$ data set as specified in the plot legend. Efficiencies are shown as function of muon $\eta$ (left) and $p_\text{T}$ (right). Red (orange) points correspond to low-$\mu$ data (MC), while the black (blue) points are high-$\mu$ data (MC). The bottom panels show the data/MC ratio for the low-$\mu$ (orange) and high-$\mu$ (blue) sets with statistical and total uncertainties [140].

## 7.3 Hadronic recoil calibration

The study of the W boson kinematics by its leptonic decay products $W^{\pm} \to l^{\pm}\nu$ is complicated first of all due to the escaping neutrino that carries away substantial information. However, the W boson transverse momentum can still be measured. As it was shown in Chapter 5, the largest part of the W boson $p_T$ is coming from the initial state radiation. The energy of the created parton shower can be defined as follows:

$$\vec{p_T^V} = \vec{p_T}^{\ell} + \vec{p_T}^{\nu} = -\sum_{i=\mathrm{ISR}\,q,g} \vec{p_T}_i = -\vec{u_T} \,, \tag{7.5}$$

where $\vec{p_T^V}$, $\vec{p_T}^{\ell}$ and $\vec{p_T}^{\nu}$ are the transverse momenta of the vector boson, lepton and neutrino respectively. The vector sum of all the partons from the ISR is called the *hadronic recoil*: $\sum_{i=\mathrm{ISR}\,q,g} \vec{p_T}_i = \vec{u_T}$. Then the missing transverse momentum $\vec{E_T^{\mathrm{miss}}}$ of the escaping neutrino can be defined as:

$$\vec{E_T^{\mathrm{miss}}} = \vec{p_T}^{\nu} = -\left(\vec{u_T} + \vec{p_T}^{\ell}\right). \tag{7.6}$$

The Hadronic Recoil (HR) reconstruction uses the Particle Flow Objects (PFO), which were defined and described in Section 6.5. It is important to exclude lepton(s) from the HR of a W(Z) event to avoid double counting. A cone of $\Delta R < 0.2$ is removed around the lepton(s) and is replaced by a same-size cone taken in the same $\eta$ and $\phi$ region, but $\Delta R > 0.4$ away from any lepton in order to obtain the unbiased pile-up and UE estimate. Only the leptons above $p_T > 10 \mathrm{GeV}$ and passing fiducial cuts in $\eta$ and ID requirements are removed from the HR.

Another important quantity for the HR is the $\sum E_T$ - a scalar sum of the transverse energies of all the PFO. The $\sum E_T$ represents the total event activity, there is a relation between the $\sum E_T$ magnitude and $u_T$ resolution. The underlying event activity, pile-up and soft emissions can be characterized by introducing another quantity: $\Sigma \bar{E}_T = \Sigma E_T - u_T$, which has the meaning of $\Sigma E_T$ with hard activity subtracted.

For the calibration of the HR it is better to introduce quantities that are defined in a natural physical way. The vector boson transverse momentum provides a natural axis which is convenient to use for the 2-component decomposition of the $u_T$ vector. The $u_T$ component parallel to the vector boson $p_T$:

$$u_{\parallel} = \frac{\vec{p_T^V} \cdot \vec{u_T}}{p_T^V} \,, \tag{7.7}$$

and a perpendicular component:

$$u_{\perp} = \frac{|\vec{p_T^V} \times \vec{u_T}|}{p_T^V}. \tag{7.8}$$

Ideally we would like to have $u_{\parallel} = p_T^V$ and $u_{\perp} = 0$, but due to detector effects it is never the case. The perpendicular component $u_{\perp}$ can be thought of as the HR resolution, while $u_{\parallel}$ has a physical meaning of the recoil scale. Another important quantity is called the *bias*:

$$b = u_{\parallel} + p_T^V \,, \tag{7.9}$$
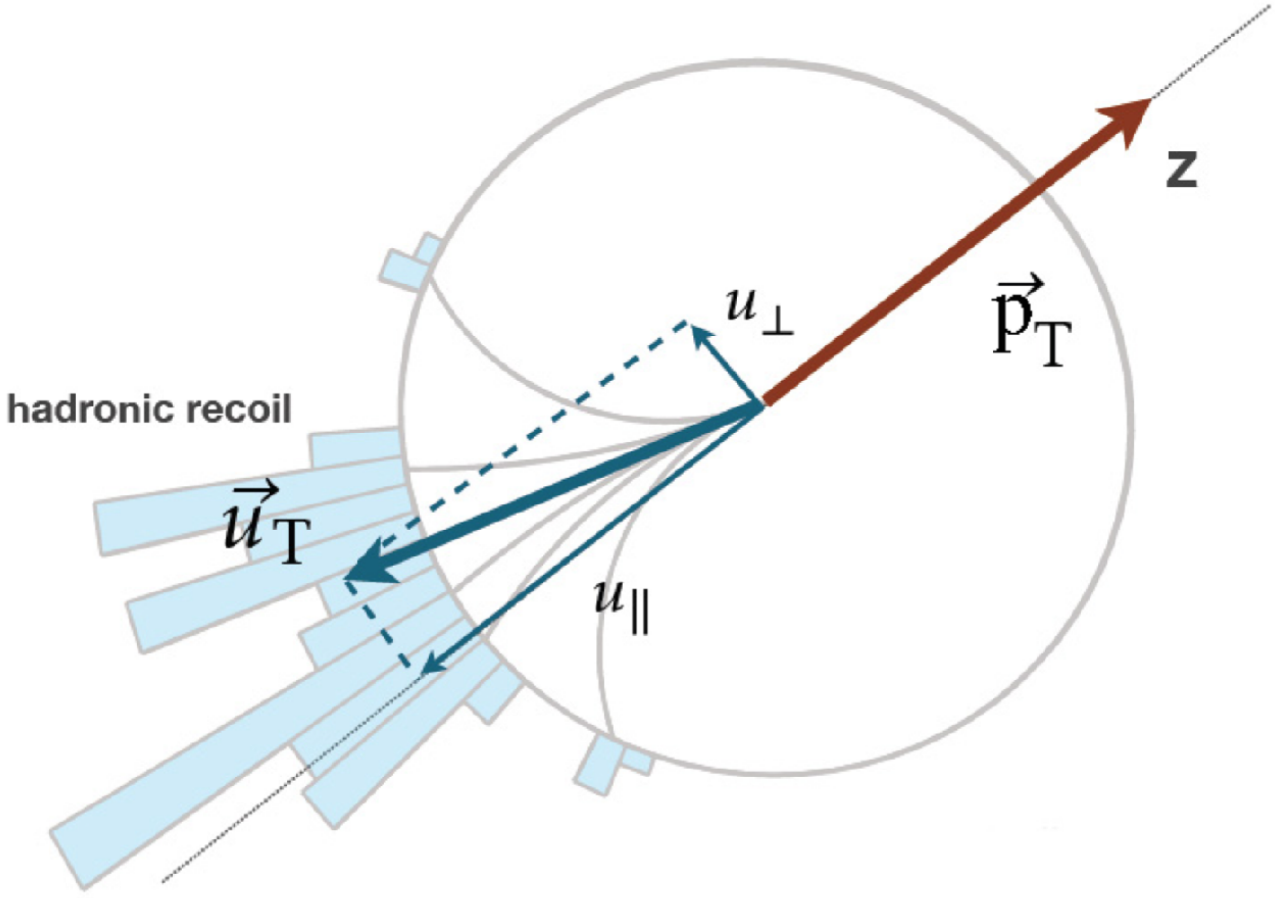
**Figure 7.24:** The hadronic recoil vector in the transverse plane and its components with respect to vector boson $p_T$ [143].

which is expected to be centred around zero. In data, of course, we don't know the $p_T^V$ of the truth boson. However, in Z decay events we can use the dilepton transverse momentum $p_T^{\ell\ell}$ as an axis for $u_\perp$ and $u_\parallel$ decomposition - considering the difference in $u_T$ resolution between $p_T^V$ and $p_T^{\ell\ell}$ is negligible. In W data events it is only possible to use $p_T^l$ for $u_T$ decomposition.

### 7.3.1 SET-$u_T$ reweighting

Despite the fact that the two electroweak bosons, W and Z, share lot of similarities, there are also small but notable differences in valence quark content and PDFs, energy scale, etc. This leads to differences in underlying event and $p_T^V$ spectra, which manifest themselves in the observables like $\Sigma E_T$ and $\Sigma \vec{E}_T$ . For the high-precision measurements it is important to ensure that these quantities as well as their correlations are modelled properly. It is also important to match these correlations in data and MC simulations. Figure 7.25 demonstrates that the baseline Powheg MC simulations lead to a significant mismodelling of the $\Sigma \vec{E}_T$ -$p_T^Z$ correlation, while Sherpa shows much better agreement with
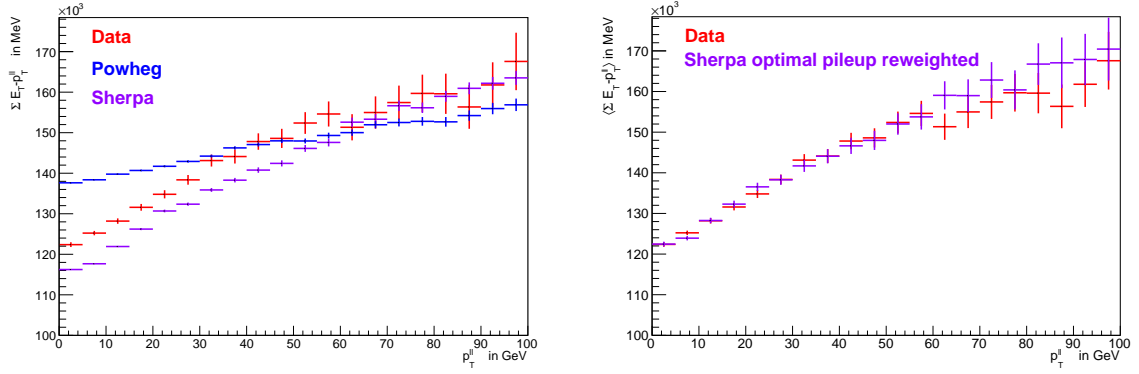
**Figure 7.25:** Comparison of the $\Sigma E_T$ - $p_T^{\ell\ell}$ description in data of the two MC samples Powheg and Sherpa at 13 TeV, showing $\Sigma E_T$ - $p_T^{\ell\ell}$ inclusively (left). Figure on the right shows the comparison of Sherpa to the data after a dedicated pileup reweighting [143].

the data from the very beginning. After applying a special pile-up reweighting of Sherpa samples a very good agreement with the data is achieved.

In order to obtain proper distributions in the MC samples, a three-step reweighting procedure is implemented.

First weight is obtained by from the 2D $\Sigma \vec{E}_T$ -$p_T^V$ distributions ratio in Data and MC:

$$w_{2D}^Z(\Sigma \vec{E}_T, p_T^{\ell\ell}) = \frac{h^{\text{data},Z}(\Sigma \vec{E}_T, p_T^{\ell\ell})}{h^{\text{MC},Z}(\Sigma \vec{E}_T, p_T^{\ell\ell})}, \tag{7.10}$$

where the following binning is used:

- $p_T^{\ell\ell}$ (13 TeV) = $[0, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 16, 20, 25, 30, 40, 50, 55, 65, 80, 100, 200, \infty]$ GeV

- $p_T^{\ell\ell}$ (5 TeV) = $[0, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 16, 20, 25, 27, 30, 40, 45, 50, 60, 70, 100, 200, \infty]$ GeV

- $\Sigma \vec{E}_T$ (13 TeV) = $[0, 10, 20, 30, ..., 380, \infty]$ GeV

- $\Sigma \vec{E}_T$ (5 TeV) = $[0, 10, 20, 30, ..., 280, \infty]$ GeV

This reweighting is obtained from the $Z \to \mu\mu$ and $Z \to ee$ events for 5 and 13 TeV datasets and applied to W Monte-Carlo samples. In W events the $p_T^V$ is used instead of $p_T^{\ell\ell}$ for obvious reasons. This reweighting assures very good agreement for the Z events, but perfect agreement is not guaranteed for the W events. For this reason a second reweighting is derived from the data: $\Sigma \vec{E}_T$ weight is extracted in bins of $u_T$ of 4 GeV width and applied to W MC events on top of the first 2D reweighting:

$$w_{j,\text{sliced}}^{W^\pm}(\Sigma \vec{E}_T) = \frac{h_j^{\text{data},W^\pm}(\Sigma \vec{E}_T)}{h_j^{\text{MC},W^\pm,\text{Z2Dmod}}(\Sigma \vec{E}_T)}, \tag{7.11}$$
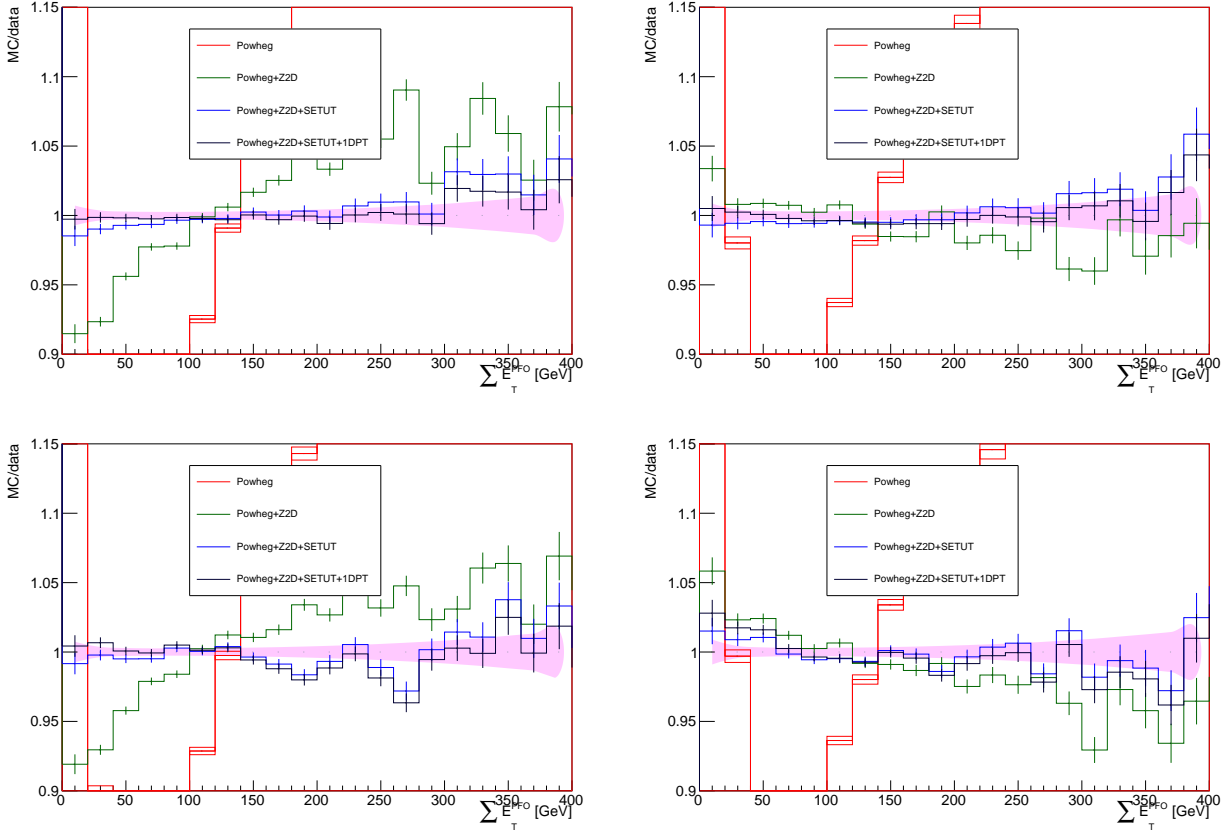
**Figure 7.26:** Ratio of data to predictions in $W \rightarrow \mu\nu$ events at 13 TeV for the $\Sigma \bar{E_T}$ distribution, before and after each $\Sigma \bar{E_T}$ modeling reweighting step. The color band is the data statistical uncertainty. The prediction uncertainty only includes the statistical uncertainty. 'Powheg' uses the baseline MC for the signal. 'Powheg+Z2D' has the 2D ($\Sigma \bar{E_T}$, $p_T^{true,V}$), Z-based reweighting applied. 'Powheg+Z2D+SETUT' adds the $\Sigma \bar{E_T}$ reweighting in bins of $u_T$. 'Powheg+Z2D+SETUT+1DPT' adds the 1D reweighting to recover the initial $p_T^{true,V}$ spectrum [143].

where $h_j$ stands for the normalized $\Sigma \bar{E_T}$ distribution in the $u_T$ bin number $j$ after the standard selection. This reweighting improves the $\Sigma \bar{E_T}$ modelling, but distorts the $p_T^V$ spectrum. This motivates the third correction reweighting with the following weight:

$$w_{1D}^{W^\pm}(p_T^{true,V}) = \frac{h^{\mathrm{MC},W^\pm,\mathrm{mod}}(p_T^{true,V})}{h^{\mathrm{MC},W^\pm,\mathrm{orig}}(p_T^{true,V})}. \tag{7.12}$$

The $h^{\mathrm{MC},W^\pm,\mathrm{orig}}(p_T^{true,V})$ stands for the original $p_T^{true,V}$ spectrum before any reweightings were applied. The total weight applied to an event is the product of the three weights described above: $w_{2D}^Z(\Sigma \bar{E_T}, p_T^{true,V}) \times w_{j,\mathrm{sliced}}^{W^\pm}(\Sigma \bar{E_T}) \times w_{1D}^{W^\pm}(p_T^{true,V})$. The results on the reweighting are shown in Fig. 7.26 and 7.27 for 13 and 5 TeV respectively.

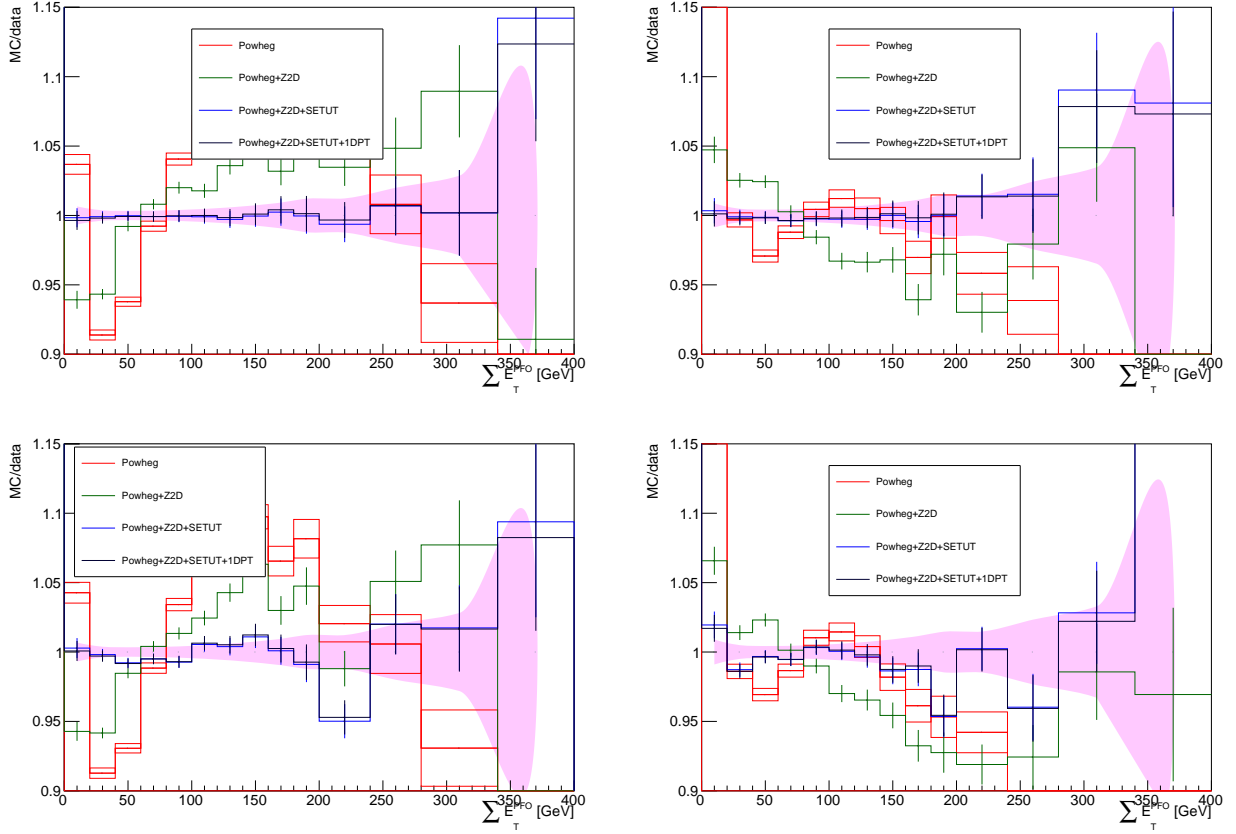The closure of the procedure is checked with Sherpa MC simulation used as pseudo-data, as we

**Figure 7.27:** Ratio of data to predictions in $W \to \mu\nu$ events at 5 TeV for the $\Sigma \vec{E}_T$ distribution, before and after each $\Sigma \vec{E}_T$ modeling reweighting step. The color band is the data statistical uncertainty. The prediction uncertainty only includes the statistical uncertainty. 'Powheg' uses the baseline MC for the signal. 'Powheg+Z2D' has the 2D ($\Sigma \vec{E}_T$, $p_T^{true,V}$), Z-based reweighting applied. 'Powheg+Z2D+SETUT' adds the $\Sigma \vec{E}_T$ reweighting in bins of $u_T$. 'Powheg+Z2D+SETUT+1DPT' adds the 1D reweighting to recover the initial $p_T^{true,V}$ spectrum [143].

don't have the $p_T^V$ distribution from the data. The residual non-closure of less than 1% is treated as a systematic uncertainty.

### 7.3.2 $u_X$ and $u_Y$ correction

The azimuthal angle distribution of the recoil s another source of discrepancy between the MC simulation and the data. While the simulated events have a flat $u^\phi$ distribution, the data events show a non-uniform distribution which is probably caused by detector imperfections or ageing. The correction is performed by introducing additive corrections to the $u_X$ and $u_Y$ components of the HR. The corrections are derived as a mean difference between the data and MC as a function of $\Sigma \vec{E}_T$ in Z

events:

$$u_X^{\text{MC,corr}} = u_X^{\text{MC}} + [(\langle u_X^{\text{data}} \rangle - \langle u_X^{\text{MC}} \rangle)(\Sigma \vec{E}_T)]$$
$$u_Y^{\text{MC,corr}} = u_Y^{\text{MC}} + [(\langle u_Y^{\text{data}} \rangle - \langle u_Y^{\text{MC}} \rangle)(\Sigma \vec{E}_T)]$$
$$(7.13)$$

The dependence of the mean differences $\langle u_X^{\text{data}} \rangle - \langle u_X^{\text{MC}} \rangle$ and $\langle u_Y^{\text{data}} \rangle - \langle u_Y^{\text{MC}} \rangle$ on $\Sigma \vec{E}_T$ is fitted with a linear function. The corrected $\phi$ distributions are shown in Fig. 7.28. It was shown that the correlation between the correction and the magnitude of the recoil is weak and the effect of the correction on the measured W spectrum is of per mille level. For this reason no uncertainty was assigned to this correction.

### 7.3.3 Resolution and response corrections

The correction function for $\sigma(u_\perp)(\Sigma \vec{E}_T, p_T^{\ell\ell})$ is constructed in bins of $p_T^{\ell\ell}$ in the following way:

$$r\left(\Sigma \vec{E}_T, p_T^{\ell\ell}\right) = \frac{\sigma(u_\perp)^{\text{data}}}{\sigma(u_\perp)^{\text{MC}}}, \tag{7.14}$$

where both $\sigma(u_\perp)$ functions are obtained as a linear fit to $\sqrt{\Sigma \vec{E}_T}$:

$$\sigma(u_\perp)(\Sigma \vec{E}_T) = c + d \cdot \sqrt{\Sigma \vec{E}_T}. \tag{7.15}$$

with the following $p_T^{\ell\ell}$ binning:

- 5 TeV: $p_T^{\ell\ell} = [0, 3, 4, 6, 7, 9, 11, 13, 16, 20, 26, 40, \infty]\,\text{GeV}$

- 13 TeV is $p_T^{\ell\ell} = [0, 3, 4, 6, 7, 9, 10, 12, 14, 17, 21, 26, 33, 49, \infty]\,\text{GeV}$

Then the correction for the W boson events is performed using the ratio function as a factor:

$$u_\perp^{\text{MC,corr}} = u_\perp^{\text{MC}} \times r\left(\Sigma \vec{E}_T, p_T^{true,V}\right). \tag{7.16}$$

The correction of the parallel component $u_\parallel$ is done as follows:

$$u_\parallel^{\text{MC,corr}} = \langle u_\parallel^{\text{data}} \rangle + (\langle b^{\text{data}} \rangle - \langle b^{\text{MC}} \rangle) \cdot r_\parallel + (u_\parallel^{\text{MC}} - \langle u_\parallel^{\text{data}} \rangle) \cdot r_\parallel. \tag{7.17}$$

Here the resolution correction factor $r_\parallel$ is in equation 7.14, but reads as $\sigma(u_\parallel)^{\text{data}}/\sigma(u_\perp)^{\text{MC}}$.
The average $\langle u_\parallel^{\text{data}} \rangle$ assumes averaging over all data events in bins of $p_T^{\ell\ell}$ and $\Sigma \vec{E}_T$. The $\Sigma \vec{E}_T$ bins are 10GeV wide for 5TeV and 20GeV wide for 13TeV. Then in each $\Sigma \vec{E}_T$ bin the $p_T^{\ell\ell}$ dependence is fitted with a linear function:

$$\langle u_\parallel^{\text{data}} \rangle(p_T^{\ell\ell}) = e + f \cdot p_T^{\ell\ell}. \tag{7.18}$$

Similarly the difference of the biases $(\langle b^{\text{data}} \rangle - \langle b^{\text{MC}} \rangle)$ is computed in the same bins of $p_T^{\ell\ell}$ and $\Sigma \vec{E}_T$ and fitted in each $\Sigma \vec{E}_T$ bin with a linear function of $p_T^{\ell\ell}$.
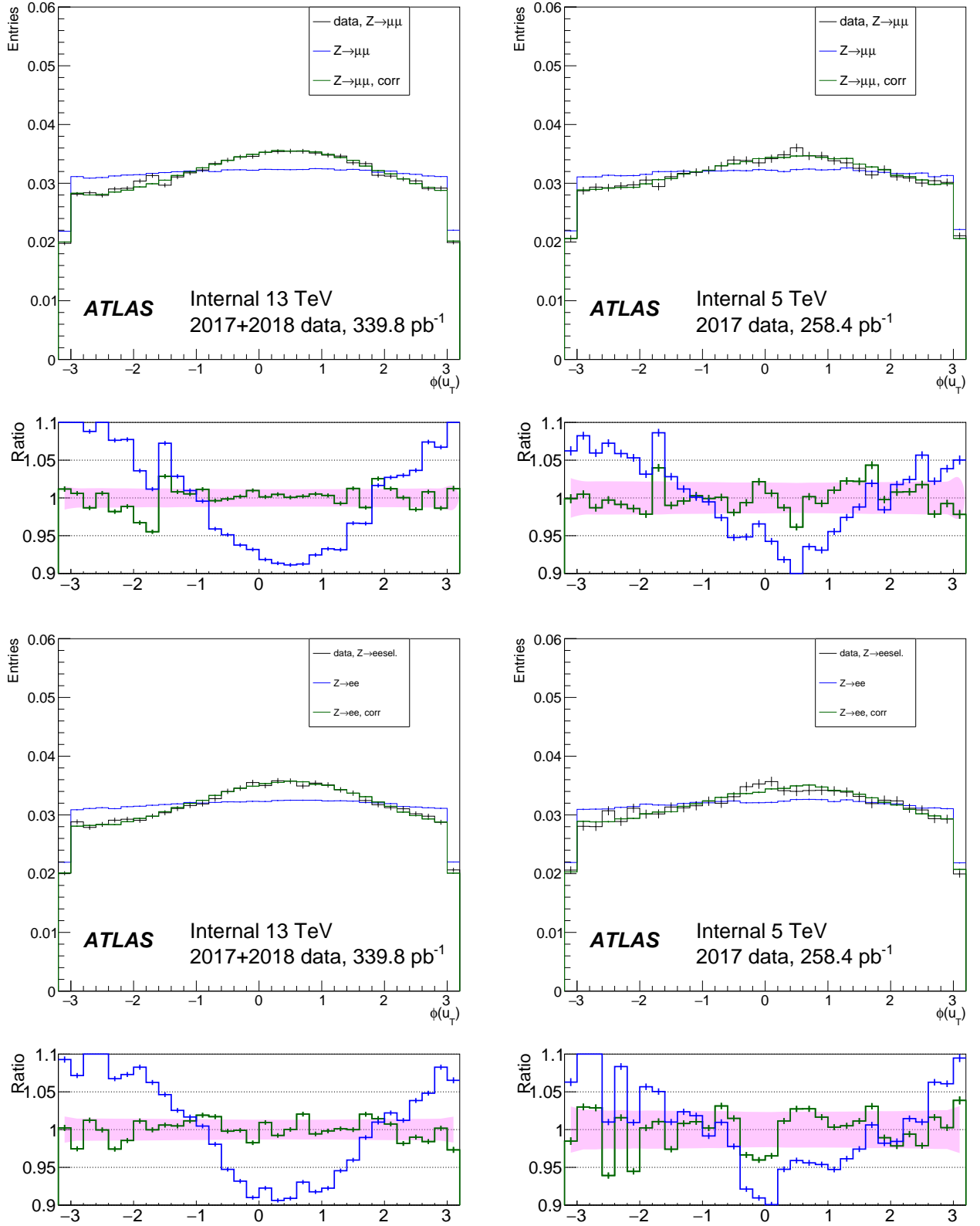
**Figure 7.28:** $\phi(u_T)$ at 5 and 13 TeV, for the data and the simulation before and after $u_X$ and $u_Y$ correction, in $Z$ events. The band in the ratio panel is the data statistical uncertainty [143].

### 7.3.4 Hadronic recoil calibration uncertainties

The uncertainties associated with hadronic recoil calibration are mainly of statistical nature for resolution and response corrections. The uncertainty of the SET-$u_T$ correction comes from the non-closure of Z to W extrapolation mentioned in the dedicated section.

## 7.4 Angular coefficients correction

The fully differential cross-section of the fully leptonic Drell-Yan process can be factorized in the following way [69]:

$$\frac{d\sigma}{dp_1 dp_2} = \left[\frac{d\sigma(m)}{dm}\right]\left[\frac{d\sigma(y)}{dy}\right]\left[\frac{d\sigma(p_T, y)}{dp_T dy}\left(\frac{d\sigma}{dy}\right)^{-1}\right]\left[(1 + \cos^2\theta + \sum_{i=0}^{7} A_i(p_T, y)P_i(\cos\theta, \phi)\right], \qquad (7.19)$$

where $p_1$ and $p_2$ are the 4-momenta of the two leptonic decay products; $m$, $p_T$ and $y$ are the dilepton system invariant mass, transverse momentum and rapidity respectively; $\theta$ is the polar and $\phi$ is azimutal coordinates of the lepton ($e^-$ in $W^- \rightarrow e^- \nu$ or $\nu$ in $W^+ \rightarrow e^+ \nu$). The angular dependence is decomposed in seven spherical harmonics $P_i$ with numerical coefficients $A_i$. The numerical coefficients $A_i$ are in general case are a function of $p_T$, $y$ and $m$, but the mass dependence may be neglected [144]. The $A_i$ coefficients were derived to the NNLO precision using the DYNNLO program for fixed-order cross-section calculations [145].

The default MC samples generated by Powheg+Pythia8 were reweighted at the event level in order to introduce the corrections:

$$w(\cos\theta, \phi, p_T, y) = \frac{1 + \cos^2\theta + \sum_i A_i'(p_T, y)P_i(\cos\theta, \phi)}{1 + \cos^2\theta + \sum_i A_i(p_T, y)P_i(\cos\theta, \phi)}, \qquad (7.20)$$

where $A_i(p_T, y)$ are the angular coefficients, predicted by the Powheg+Pythia8 simulation and $A_i'(p_T, y)$ are the NNLO coefficients evaluated at $O(\alpha^2)$ precision.

# 8 Event selection, datasets and MC samples

This chapter contains the description of the data and MC samples, used in the analysis, as well as the applied cuts. It also describes the event-level correction of the Z axis position of the primary vertex. The chapter concludes with the set of control plots of the main observables.

## 8.1 Data and MC samples

### 8.1.1 Data samples

The data samples for this study were collected with special beam conditions that ensure low pile-up. The MC samples were generated to match these conditions.s The data samples were collected in three runs:

- $\sqrt{s}$ = 5.02 TeV data taken in November 2017, ATLAS data period M, preliminary calibrated luminosity $256.827\,\mathrm{pb}^{-1}$ with an uncertainty of ±1.6%.

- $\sqrt{s}$ = 13 TeV data taken in November 2017, ATLAS data period N, preliminary online luminosity $146.6\,\mathrm{pb}^{-1}$.

- $\sqrt{s}$ = 13 TeV data taken in June 2018, ATLAS data period G4+J, preliminary online luminosity $193.2\,\mathrm{pb}^{-1}$.

The luminosity calibration for the 13 TeV runs is not available yet, the corresponding uncertainty is not known, but is expected to be around 3%. The runs of November 2017 and the run of June 2018 had the same bunch spacing of 25 ns, but a different filling scheme. The two main differences from the high-$\mu$ data collection are the following:

- In order to optimize topo-cluster response for the HR lower topo-cluster thresholds were used.

- Single $e$ and $\mu$ triggers with significantly lower thresholds and looser identification criteria are run without prescale, most notably `HLT_e15_lhloose_nod0_L1EM12` and `HLT_mu14`.

At the beginning of 5 TeV fills the pile-up reached $\mu \sim 5$, slowly descending to $\mu \sim 1$ by the end of the run. In the case of 13 TeV the luminosity was levelled at $\mu = 2$ in the course of the run. The corresponding distributions for $\mu$ and $N_{PV}$ for the 5 TeV and 13 TeV runs are shown in Fig. 8.1 and Fig. 8.2.
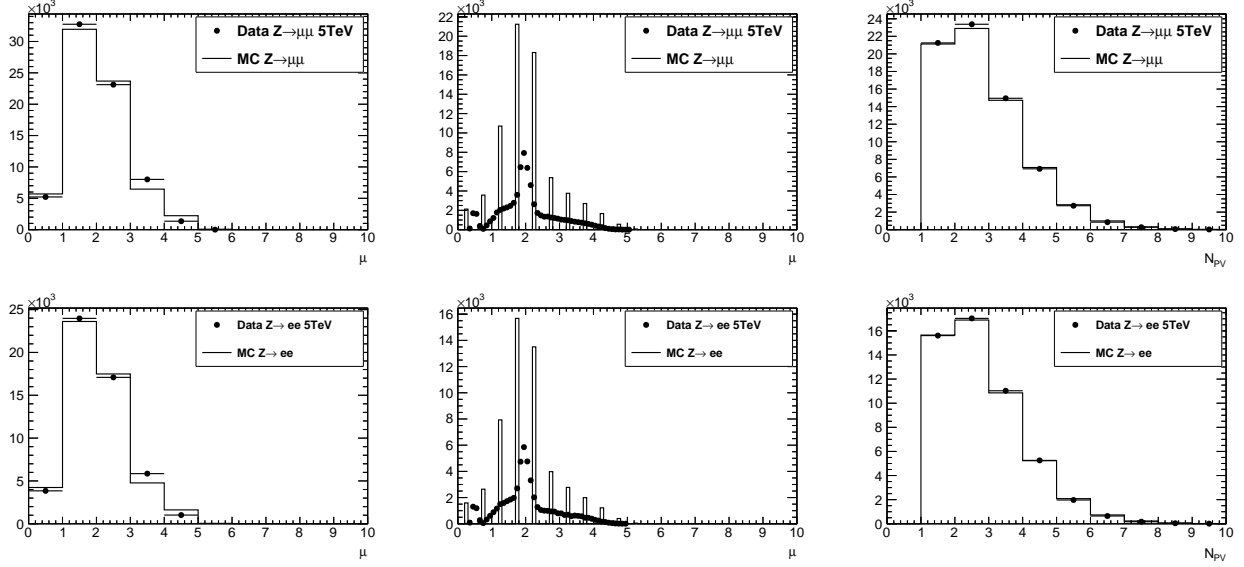


**Figure 8.1:** Distributions for the 5 TeV low-$\mu$ dataset in a $Z/\gamma^* \to \mu\mu$ (top row) and a $Z/\gamma^* \to ee$ (bottom row) selection. The data (points) is compared to $Z/\gamma^* \to \mu\mu$ or $Z/\gamma^* \to ee$ signal MC, respectively. The left and middle plots show the actual $\mu$ in a coarsely-binned and a finely-binned version. The right plot shows the number of reconstructed primary vertices $N_{PV}$ [146].

## 8.1.2   MC samples and cross-sections

Signal and background processes (except for the multijet background) are modelled using fully simulated and reconstructed MC samples, specifically tuned for the special run conditions, namely the low pileup, lower topo-cluster noise thresholds and adapted trigger menu. No pileup reweighting is performed.

The information on the simulated samples and their properties is given in Tables 8.2, 8.2, 8.2, 8.2 [136]. The predicted event counts are normalized to the data luminosity using the cross-sections quoted in the table.

The primary signal event samples for W and Z production are obtained using Powheg [147, 148, 149, 150] event generator with CT10 PDF, linked with Pythia8 [151] with AZNLO tune [152]. Powheg+Pythia8samples are interfaced to Photos++ [153] for final state QED effects simulation.

A set of alternative samples at $\sqrt{s} = 13$TeV was prepared with Sherpa2.2.2 [154] using the NNPDF3.0 PDFs and merging $V + 0, 1, 2$ at NLO accuracy with $V + 3, 4$ at LO accuracy with the
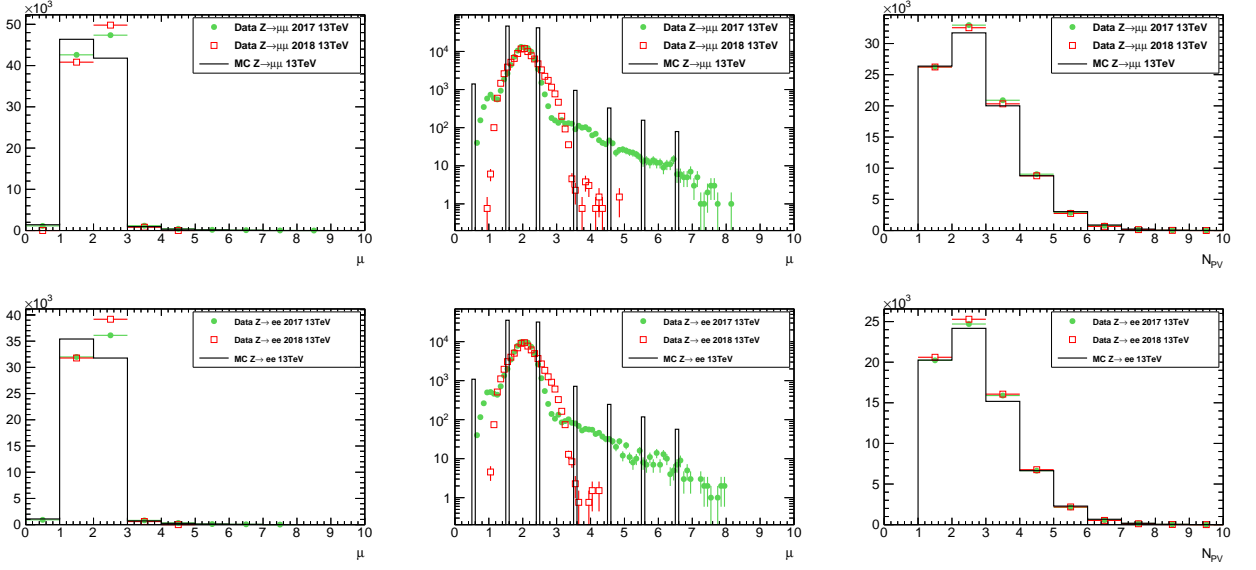
**Figure 8.2:** Distributions for the 13 TeV low-$\mu$ datasets taken in 2017 and 2018 in a $Z/\gamma^* \to \mu\mu$ (top row) and a $Z/\gamma^* \to ee$ (bottom row) selection. The data (points) is compared to $Z/\gamma^* \to \mu\mu$ or $Z/\gamma^* \to ee$ signal MC, respectively. All distributions are (roughly) normalised to the same number of selected events in the 2017 dataset. The left and middle plots show the actual $\mu$ in a coarsely-binned and a finely-binned version. The right plot shows the number of reconstructed primary vertices $N_{PV}$ [146].

MEPS@NLO scheme. A similar set for $\sqrt{s}$ = 5TeV was prepared with SHERPA2.2.5 with a setup similar to 13 TeV samples.

Pileup is modelled by overlaying simulated soft events over the original hard-scattering event. These soft events were modelled using PYTHIA8 with NNPDF2.3LO set of PDFs [155] and the A3 tune [156].

The $W$ and $Z$ processes samples are normalized to NNLO calculations for the cross-sections performed using DYTURBO, an optimised version of DYRES [157, 158] using the MMHT2014nnlo PDF set [159]. Corresponding numerical values were taken from the corresponding ATLAS publications of the 2015 data at 13 TeV [160] and 5.02 TeV [161] and are presented in Table 8.2 for 13 TeV and Table 8.2 for 5 TeV. The uncertainties on those cross-sections arise from the choice of PDF set, from factorization and renormalisation scale dependence, and the strong coupling constant $\alpha_s$ uncertainty resulting in the total uncertainty estimate of about 5%.

Backgrounds from top-quark pair-production $t\bar{t}$ and single-top production ($Wt$, t-channel, s-channel) were generated with POWHEG+PYTHIA8. The 5 TeV $t\bar{t}$ cross section is taken as the TOP++ [162] prediction reported in CMS publication [163]. Di-boson combinations $VV, V = W, Z$ are generated with SHERPA in all decay channels with a requirement of having at least one real lepton in the final state.

| Process | Data set | Generator | $\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}}$ [nb] (th. unc.) | $N_{\text{evt}}^{\text{skim}}$ [$10^6$] | $N_{\text{evt}}^{\text{unskim}}$ [$10^6$] |
|---|---|---|---|---|---|
| $W^+ \to e^+ \nu$ | 361100 | Powheg+Pythia8 | 11.61 (5%) | 40 | 40 |
| $W^+ \to \mu^+ \nu$ | 361101 | Powheg+Pythia8 | 11.61 (5%) | 40 | 40 |
| $W^+ \to \tau^+ \nu$ | 361102 | Powheg+Pythia8 | 11.61 (5%) | 0.28 | 5.0 |
| $W^- \to e^- \bar{\nu}$ | 361103 | Powheg+Pythia8 | 8.630 (5%) | 30 | 30 |
| $W^- \to \mu^- \bar{\nu}$ | 361104 | Powheg+Pythia8 | 8.630 (5%) | 29 | 29 |
| $W^- \to \tau^- \bar{\nu}$ | 361105 | Powheg+Pythia8 | 8.630 (5%) | 0.24 | 4.0 |
| $Z \to ee$ | 361106 | Powheg+Pythia8 | $1.910 \times 1.03$ (5%) | 10 | 10 |
| $Z \to \mu\mu$ | 361107 | Powheg+Pythia8 | $1.910 \times 1.025$ (5%) | 10 | 10 |
| $Z \to \tau\tau$ | 361108 | Powheg+Pythia8 | $1.910 \times 1.025$ (5%) | 0.12 | 1.0 |
| $ZZ(q\bar{q}\ell\ell)$ | 363356 | Sherpa 2.2.1 | $0.01556 \times 0.141$ (10%) | 0.0064 | 0.010 |
| $WZ(q\bar{q}\ell\ell)$ | 363358 | Sherpa 2.2.1 | 0.003433 (10%) | 0.0063 | 0.010 |
| $WW(q\bar{q}\ell\nu)$ | 363359 | Sherpa 2.2.1 | 0.02472 (10%) | 0.0093 | 0.020 |
| $WW(\ell\nu q\bar{q})$ | 363360 | Sherpa 2.2.1 | 0.02472 (10%) | 0.0093 | 0.020 |
| $WZ(\ell\nu q\bar{q})$ | 363489 | Sherpa 2.2.1 | 0.01142 (10%) | 0.0047 | 0.010 |
| $ZZ(4\ell)$ | 364250 | Sherpa 2.2.2 | 0.001252 (10%) | 0.0057 | 0.010 |
| $WZ(3\ell\nu)$ | 364253 | Sherpa 2.2.2 | 0.004583 (10%) | 0.0062 | 0.010 |
| $WW(2\ell2\nu)$ | 364254 | Sherpa 2.2.2 | 0.01250 (10%) | 0.0073 | 0.010 |
| $WZ(\ell3\nu)$ | 364255 | Sherpa 2.2.2 | 0.003235 (10%) | 0.0050 | 0.010 |
| $Wt$ | 410013 | Powheg+Pythia8 | 0.03582 (10%) | 0.0037 | 0.010 |
| $W\bar{t}$ | 410014 | Powheg+Pythia8 | 0.03399 (10%) | 0.0037 | 0.010 |
| $t\bar{t}$ (nominal) | 410470 | Powheg+Pythia8 | $0.8318 \times 0.544$ (7%) | 1.2 | 2.0 |
| $t(\text{t} - \text{chan}.t)$ | 410642 | Powheg+Pythia8 | 0.03699 (10%) | 0.016 | 0.030 |
| $t(\text{t} - \text{chan}.\bar{t})$ | 410643 | Powheg+Pythia8 | 0.02217 (10%) | 0.011 | 0.020 |
| $t(\text{s} - \text{chan}.t)$ | 410644 | Powheg+Pythia8 | 0.002027 (10%) | 0.0050 | 0.010 |
| $t(\text{s} - \text{chan}.\bar{t})$ | 410645 | Powheg+Pythia8 | 0.001268 (10%) | 0.0052 | 0.010 |
| $t\bar{t}$ (syst.) | 410480 | Powheg+Pythia8 | $0.8318 \times 0.438$ (7%) | 0.85 | 1.5 |
| $t\bar{t}$ (syst.) | 410482 | Powheg+Pythia8 | $0.8318 \times 0.105$ (7%) | 0.40 | 0.50 |
| $t\bar{t}$ (syst.) | 410557 | Powheg+Pythia8 | $0.8318 \times 0.438$ (7%) | 0.85 | 1.5 |
| $t\bar{t}$ (syst.) | 410558 | Powheg+Pythia8 | $0.8318 \times 0.105$ (7%) | 0.40 | 0.50 |

**Table 8.2:** Monte Carlo samples at $\sqrt{s} = 13$ TeV. Given is a short description of the process, the ATLAS MC data set number (DSID), the names and version numbers of the MC generator(s), the used value of the higher order cross section times any branching and filter efficiencies ($\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}}$) with the theoretical uncertainty in percent ("th. unc."), and finally the number of events analysed after skimming at derivation production ($N_{\text{evt}}^{\text{skim}}$) as well as the number of events originally processed and simulated ($N_{\text{evt}}^{\text{unskim}}$). In the case of $Z \to \ell\ell$ samples, the given $\epsilon_{\text{filter}} > 1$ is related to the fact, that the cross sections were calculated for $66 < m_{\ell\ell} < 116$ GeV, but the generated mass range is larger. The last section of $t\bar{t}$ samples refers to variation samples for systematics studies. The MC equivalent luminosity $N_{\text{evt}}^{\text{unskim}}/(\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}})$ is generally above 3 fb$^{-1}$ for signal and significant backgrounds, the exception are Powheg $W \to \tau\nu$ and $Z \to \tau\tau$ samples, that have about 0.45 fb$^{-1}$ only [146].

| Process | Data set | Generator | $\sigma \cdot BR \cdot \epsilon_{\text{filter}}$ [nb] (th. unc.) | $N_{\text{evt}}^{\text{skim}}$ [$10^6$] | $N_{\text{evt}}^{\text{unskim}}$ [$10^6$] |
|---|---|---|---|---|---|
| $Z \to \mu\mu$ | 364100 | SHERPA 2.2.1 | $1.932 \times 0.822$ (5%) | 8.0 | 8.0 |
| $Z \to \mu\mu$ | 364101 | SHERPA 2.2.1 | $1.933 \times 0.114$ (5%) | 1.5 | 1.5 |
| $Z \to \mu\mu$ | 364102 | SHERPA 2.2.1 | $1.932 \times 0.0660$ (5%) | 1.1 | 1.1 |
| $Z \to \mu\mu$ | 364103 | SHERPA 2.2.1 | $0.1063 \times 0.690$ (5%) | 1.5 | 1.5 |
| $Z \to \mu\mu$ | 364104 | SHERPA 2.2.1 | $0.1062 \times 0.200$ (5%) | 0.40 | 0.40 |
| $Z \to \mu\mu$ | 364105 | SHERPA 2.2.1 | $0.1063 \times 0.114$ (5%) | 0.25 | 0.25 |
| $Z \to \mu\mu$ | 364106 | SHERPA 2.2.1 | $0.03889 \times 0.593$ (5%) | 0.20 | 0.20 |
| $Z \to \mu\mu$ | 364107 | SHERPA 2.2.1 | $0.03885 \times 0.235$ (5%) | 0.060 | 0.060 |
| $Z \to \mu\mu$ | 364108 | SHERPA 2.2.1 | $0.03889 \times 0.156$ (5%) | 0.035 | 0.035 |
| $Z \to \mu\mu$ | 364109 | SHERPA 2.2.1 | $0.008310 \times 0.561$ (5%) | 0.020 | 0.020 |
| $Z \to \mu\mu$ | 364110 | SHERPA 2.2.1 | $0.008310 \times 0.266$ (5%) | 0.010 | 0.010 |
| $Z \to \mu\mu$ | 364111 | SHERPA 2.2.1 | $0.008320 \times 0.177$ (5%) | 0.0050 | 0.0050 |
| $Z \to \mu\mu$ | 364112 | SHERPA 2.2.1 | $0.001740$ (5%) | 0.0050 | 0.0050 |
| $Z \to \mu\mu$ | 364113 | SHERPA 2.2.1 | $0.0001400$ (5%) | 0.0050 | 0.0050 |
| $Z \to ee$ | 364114 | SHERPA 2.2.1 | $1.933 \times 0.821$ (5%) | 8.0 | 8.0 |
| $Z \to ee$ | 364115 | SHERPA 2.2.1 | $1.932 \times 0.114$ (5%) | 1.5 | 1.5 |
| $Z \to ee$ | 364116 | SHERPA 2.2.1 | $1.932 \times 0.0658$ (5%) | 1.1 | 1.1 |
| $Z \to ee$ | 364117 | SHERPA 2.2.1 | $0.1080 \times 0.694$ (5%) | 1.5 | 1.5 |
| $Z \to ee$ | 364118 | SHERPA 2.2.1 | $0.1077 \times 0.191$ (5%) | 0.40 | 0.40 |
| $Z \to ee$ | 364119 | SHERPA 2.2.1 | $0.1078 \times 0.119$ (5%) | 0.25 | 0.25 |
| $Z \to ee$ | 364120 | SHERPA 2.2.1 | $0.03964 \times 0.616$ (5%) | 0.20 | 0.20 |
| $Z \to ee$ | 364121 | SHERPA 2.2.1 | $0.03967 \times 0.233$ (5%) | 0.060 | 0.060 |
| $Z \to ee$ | 364122 | SHERPA 2.2.1 | $0.04068 \times 0.150$ (5%) | 0.035 | 0.035 |
| $Z \to ee$ | 364123 | SHERPA 2.2.1 | $0.008460 \times 0.569$ (5%) | 0.020 | 0.020 |
| $Z \to ee$ | 364124 | SHERPA 2.2.1 | $0.008450 \times 0.266$ (5%) | 0.010 | 0.010 |
| $Z \to ee$ | 364125 | SHERPA 2.2.1 | $0.008470 \times 0.177$ (5%) | 0.0050 | 0.0050 |
| $Z \to ee$ | 364126 | SHERPA 2.2.1 | $0.001760$ (5%) | 0.0050 | 0.0050 |
| $Z \to ee$ | 364127 | SHERPA 2.2.1 | $0.0001451$ (5%) | 0.0050 | 0.0050 |

**Table 8.2:** Alternative signal $Z \to \ell\ell$ Monte Carlo samples at $\sqrt{s} = 13$ TeV produced with SHERPA. General description of the table see Table 8.2. The samples are split into a long list of orthogonal slices based on "max(pTV,HT)" and filtered further into "b/c/light-jet" subcomponents. For the purpose of this analysis, the number of events in each slice is such that the samples are about $2\text{fb}^{-1}$ each (after application of a penalty factor for negative weight events) and an "inclusive sample" is restored after merging the slices [146].

| Process | Data set | Generator | $\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}}$ [nb] (th. unc.) | $N_{\text{evt}}^{\text{skim}}$ [$10^6$] | $N_{\text{evt}}^{\text{unskim}}$ [$10^6$] |
|---|---|---|---|---|---|
| $W \to \mu\nu$ | 364156 | Sherpa 2.2.1 | $18.58 \times 0.825$ (5%) | 31 | 31 |
| $W \to \mu\nu$ | 364157 | Sherpa 2.2.1 | $18.57 \times 0.131$ (5%) | 8.1 | 8.1 |
| $W \to \mu\nu$ | 364158 | Sherpa 2.2.1 | $18.57 \times 0.0433$ (5%) | 2.6 | 2.6 |
| $W \to \mu\nu$ | 364159 | Sherpa 2.2.1 | $0.9173 \times 0.674$ (5%) | 6.3 | 6.3 |
| $W \to \mu\nu$ | 364160 | Sherpa 2.2.1 | $0.9172 \times 0.244$ (5%) | 2.1 | 2.1 |
| $W \to \mu\nu$ | 364161 | Sherpa 2.2.1 | $0.9163 \times 0.0847$ (5%) | 0.23 | 0.23 |
| $W \to \mu\nu$ | 364162 | Sherpa 2.2.1 | $0.3296 \times 0.600$ (5%) | 0.80 | 0.80 |
| $W \to \mu\nu$ | 364163 | Sherpa 2.2.1 | $0.3297 \times 0.293$ (5%) | 0.27 | 0.27 |
| $W \to \mu\nu$ | 364164 | Sherpa 2.2.1 | $0.3295 \times 0.111$ (5%) | 0.099 | 0.099 |
| $W \to \mu\nu$ | 364165 | Sherpa 2.2.1 | $0.06993 \times 0.548$ (5%) | 0.068 | 0.068 |
| $W \to \mu\nu$ | 364166 | Sherpa 2.2.1 | $0.06995 \times 0.320$ (5%) | 0.034 | 0.034 |
| $W \to \mu\nu$ | 364167 | Sherpa 2.2.1 | $0.06991 \times 0.125$ (5%) | 0.014 | 0.014 |
| $W \to \mu\nu$ | 364168 | Sherpa 2.2.1 | $0.01456$ (5%) | 0.020 | 0.020 |
| $W \to \mu\nu$ | 364169 | Sherpa 2.2.1 | $0.001200$ (5%) | 0.004 | 0.004 |
| $W \to e\nu$ | 364170 | Sherpa 2.2.1 | $18.58 \times 0.825$ (5%) | 31 | 31 |
| $W \to e\nu$ | 364171 | Sherpa 2.2.1 | $18.57 \times 0.131$ (5%) | 8.3 | 8.3 |
| $W \to e\nu$ | 364172 | Sherpa 2.2.1 | $18.57 \times 0.0448$ (5%) | 2.5 | 2.5 |
| $W \to e\nu$ | 364173 | Sherpa 2.2.1 | $0.9168 \times 0.675$ (5%) | 6.4 | 6.4 |
| $W \to e\nu$ | 364174 | Sherpa 2.2.1 | $0.9176 \times 0.244$ (5%) | 2.1 | 2.1 |
| $W \to e\nu$ | 364175 | Sherpa 2.2.1 | $0.9173 \times 0.0851$ (5%) | 0.79 | 0.79 |
| $W \to e\nu$ | 364176 | Sherpa 2.2.1 | $0.3295 \times 0.599$ (5%) | 0.76 | 0.76 |
| $W \to e\nu$ | 364177 | Sherpa 2.2.1 | $0.3297 \times 0.288$ (5%) | 0.28 | 0.28 |
| $W \to e\nu$ | 364178 | Sherpa 2.2.1 | $0.3295 \times 0.111$ (5%) | 0.10 | 0.10 |
| $W \to e\nu$ | 364179 | Sherpa 2.2.1 | $0.06993 \times 0.548$ (5%) | 0.070 | 0.070 |
| $W \to e\nu$ | 364180 | Sherpa 2.2.1 | $0.06996 \times 0.320$ (5%) | 0.034 | 0.034 |
| $W \to e\nu$ | 364181 | Sherpa 2.2.1 | $0.06994 \times 0.137$ (5%) | 0.014 | 0.014 |
| $W \to e\nu$ | 364182 | Sherpa 2.2.1 | $0.01460$ (5%) | 0.020 | 0.020 |
| $W \to e\nu$ | 364183 | Sherpa 2.2.1 | $0.001200$ (5%) | 0.0050 | 0.0050 |

**Table 8.2:** Alternative signal $W \to \ell\nu$ Monte Carlo samples at $\sqrt{s} = 13\,\text{TeV}$ produced with Sherpa. See Table 8.2 for a description of the table. The samples are split into a long list of orthogonal slices based on "max(pTV,HT)" and filtered further into "b/c/light-jet" subcomponents. For the purpose of this analysis, the number of events in each slice is such that the samples are about $1\text{fb}^{-1}$ each (after application of a penalty factor for negative weight events) and an "inclusive sample" is restored after merging the slices [146].

| Process | Data set | Generator | $\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}}$ [nb] (th. unc.) | $N_{\text{evt}}^{\text{skim}}$ [$10^6$] | $N_{\text{evt}}^{\text{unskim}}$ [$10^6$] |
|---|---|---|---|---|---|
| $W^+ \to e^+ \nu$ | 361100 | POWHEG+PYTHIA8 | 4.357 (5%) | 11 | 11 |
| $W^+ \to \mu^+ \nu$ | 361101 | POWHEG+PYTHIA8 | 4.357 (5%) | 11 | 11 |
| $W^+ \to \tau^+ \nu$ | 361102 | POWHEG+PYTHIA8 | 4.357 (5%) | 0.065 | 0.94 |
| $W^- \to e^- \bar{\nu}$ | 361103 | POWHEG+PYTHIA8 | 2.902 (5%) | 7.0 | 7.0 |
| $W^- \to \mu^- \bar{\nu}$ | 361104 | POWHEG+PYTHIA8 | 2.902 (5%) | 7.0 | 7.0 |
| $W^- \to \tau^- \bar{\nu}$ | 361105 | POWHEG+PYTHIA8 | 2.902 (5%) | 0.039 | 0.59 |
| $Z \to ee$ | 361106 | POWHEG+PYTHIA8 | $0.6600 \times 1.025$ (5%) | 6.3 | 6.3 |
| $Z \to \mu\mu$ | 361107 | POWHEG+PYTHIA8 | $0.6600 \times 1.025$ (5%) | 3.4 | 3.4 |
| $Z \to \tau\tau$ | 361108 | POWHEG+PYTHIA8 | $0.6600 \times 1.025$ (5%) | 0.039 | 0.29 |
| $Z \to ee$ | 364381 | SHERPA 2.2.5 | $0.6600 \times 1.12$ (5%) | 5.0 | 5.0 |
| $Z \to \mu\mu$ | 364382 | SHERPA 2.2.5 | $0.6600 \times 1.12$ (5%) | 5.0 | 5.0 |
| $Z \to \tau\tau$ | 364383 | SHERPA 2.2.5 | $0.6600 \times 1.12$ (5%) | 1.5 | 1.5 |
| $W \to e\nu$ | 364384 | SHERPA 2.2.5 | 7.259 (5%) | 25 | 25 |
| $W \to \mu\nu$ | 364385 | SHERPA 2.2.5 | 7.259 (5%) | 25 | 25 |
| $W \to \tau\nu$ | 364386 | SHERPA 2.2.5 | 7.259 (5%) | 6.0 | 6.0 |
| $ZZ(4\ell)$ | 361063 | SHERPA 2.1 | 0.004624 (10%) | 0.017 | 0.049 |
| $WZ(\ell\ell\ell^-\nu\text{SF})$ | 361064 | SHERPA 2.1 | 0.0005324 (10%) | 0.0073 | 0.015 |
| $WZ(\ell\ell\ell^-\nu\text{OF})$ | 361065 | SHERPA 2.1 | 0.001041 (10%) | 0.012 | 0.030 |
| $WZ(\ell\ell\ell^+\nu\text{SF})$ | 361066 | SHERPA 2.1 | 0.0008433 (10%) | 0.010 | 0.020 |
| $WZ(\ell\ell\ell^+\nu\text{OF})$ | 361067 | SHERPA 2.1 | 0.001633 (10%) | 0.016 | 0.039 |
| $WW(2\ell2\nu)$ | 361068 | SHERPA 2.1 | 0.003356 (10%) | 0.068 | 0.090 |
| $WW(q\bar{q}\ell\nu)$ | 361091 | SHERPA 2.1 | 0.006059 (10%) | 0.078 | 0.15 |
| $WW(\ell\nu q\bar{q})$ | 361092 | SHERPA 2.1 | 0.006082 (10%) | 0.14 | 0.26 |
| $WZ(\ell\nu q\bar{q})$ | 361093 | SHERPA 2.1 | 0.002503 (10%) | 0.039 | 0.075 |
| $WZ(q\bar{q}\ell\ell)$ | 361094 | SHERPA 2.1 | 0.0007518 (10%) | 0.017 | 0.025 |
| $ZZ(q\bar{q}\ell\ell)$ | 361096 | SHERPA 2.1 | $0.003789 \times 0.148$ (10%) | 0.0070 | 0.010 |
| $t\bar{t}$ | 410470 | POWHEG+PYTHIA8 | $0.06890 \times 0.544$ (7%) | 1.8 | 2.8 |
| $t(\text{s}-\text{chan.}t)$ | 410644 | POWHEG+PYTHIA8 | 0.0005400 (10%) | 0.028 | 0.050 |
| $t(\text{s}-\text{chan.}\bar{t})$ | 410645 | POWHEG+PYTHIA8 | 0.0002751 (10%) | 0.028 | 0.050 |
| $Wt$ | 410646 | POWHEG+PYTHIA8 | 0.002990 (10%) | 0.018 | 0.050 |
| $W\bar{t}$ | 410647 | POWHEG+PYTHIA8 | 0.002983 (10%) | 0.019 | 0.050 |
| $t(\text{t}-\text{chan.}t)$ | 410658 | POWHEG+PYTHIA8 | 0.005414 (10%) | 0.028 | 0.050 |
| $t(\text{t}-\text{chan.}\bar{t})$ | 410659 | POWHEG+PYTHIA8 | 0.002682 (10%) | 0.028 | 0.050 |

**Table 8.2:** Monte Carlo samples at $\sqrt{s} = 5\text{TeV}$. The table follows the same format as Table 8.2. The MC equivalent luminosity $N_{\text{evt}}^{\text{unskim}}/(\sigma \cdot \text{BR} \cdot \epsilon_{\text{filter}})$ is generally above $2.5\text{fb}^{-1}$ for signal and significant backgrounds, the exception are Powheg $W \to \tau\nu$ and $Z \to \tau\tau$ samples, that have about $0.20\text{fb}^{-1}$ and $0.45\text{fb}^{-1}$ only [146].

## 8.2 Multijet background

The estimate of the multijet background, which contain contributions from fake leptons produced in semi-leptonic decays of heavy quarks, in-flight kaon decays, photon conversions, mis-identified pions, is done using a data-driven technique. The W boson signal region is defined by the following cuts:

- $p_T^\ell > 25$ GeV, $|\eta_\ell| < 2.4$;

- $E_T^{\text{miss}} > 25$ GeV,

- $m_T > 50$ GeV.

- lepton isolation.

The production of multijet events is mainly concentrated at lower values of $p_T^l$, $E_T^{\text{miss}}$ and $m_T$, such that the largest part of the multijet background events is removed by the cuts described above. The background estimate is obtained by fitting the signal and multijet yields in $p_T^l$, $E_T^{\text{miss}}$ and $m_T$ kinematic distributions, but with $E_T^{\text{miss}}$ and $m_T$ cuts relaxed. These kinematic distributions for the signal are modelled using the MC simulation and include the calibrations and corrections presented in the previous chapter. The templates of the multijet distributions are obtained using the data with the same kinematic selection, but with inverted isolation cuts. The multijet yield is obtained in the region with relaxed kinematic cuts and then extrapolated to the signal region, correcting for the efficiency of kinematic cuts.

The first step consists in defining four different regions in phase-isolation space:

- signal region (SR): isolated leptons, signal requirement on $p_T^{lep}$, $E_T^{miss}$ and $m_T$;

- fit region (FR): isolated leptons, relaxed kinematic requirements: $E_T^{\text{miss}} > 0$ GeV, $m_T > 0$ GeV;

- control region 1 (CR1): anti-isolated leptons with FR kinematic requirements;

- control region 2 (CR2): anti-isolated leptons with SR kinematic requirements.

The multijet template is extracted from CR1 and normalized using the fraction fit, obtained from fit region (FR). Then the multijet(MJ) yield is estimated in the SR through the ratio of MJ events in the two control regions: $\epsilon = N_{MJ}^{CR2}/N_{MJ}^{CR1}$. The number of MJ background events is estimated in the following way:

- The number of multijet background events in CR1 ($N_{\text{MJ}}^{CR1}$) and their distributions ($H_{\text{MJ}}^{CR1}$) are derived as follows:

$$N_{\text{MJ}}^{CR1} = N_{\text{data}}^{\text{CR1}} - N_{\text{EW}}^{\text{CR1}}, \tag{8.1}$$

$$H_{\text{MJ}}^{CR1} = H_{\text{data}}^{\text{CR1}} - H_{\text{EW}}^{\text{CR1}} \tag{8.2}$$

where $H^{CR1}$ stands for one of the kinematic distributions used in the fit, namely $p_T^{\ell}$, $E_T^{miss}$ or $m_T$.

- The fraction fit is performed in FR, which has looser kinematics cuts and the same isolation cuts as the signal. The fit has the following form:

$$H_{data}^{FR} = \alpha \cdot H_{EW}^{FR} + T \cdot H_{MJ}^{CR1}. \tag{8.3}$$

The fitting parameter $T$ gives the factor for the MJ contribution in FR: $N_{MJ}^{FR} \approx T \cdot N_{MJ}^{CR1}$. A normalization factor for the EW+top contribution, $\alpha$, is also fitted and should be unity within the uncertainties in the luminosity and the cross-sections of the MC-simulated processes.

- Then the fitted multijet yield is extrapolated to the signal region. The extrapolation factor $\varepsilon$ that was mentioned before can be obtained as follows:

$$\varepsilon \equiv \frac{N_{data}^{CR2} - N_{EW}^{CR2}}{N_{data}^{CR1} - N_{EW}^{CR1}}, \tag{8.4}$$

and assuming that this factor does not depend on the isolation cuts, one obtains

$$N_{MJ}^{SR} = \varepsilon N_{MJ}^{FR}. \tag{8.5}$$

This method relies on the anti-isolation procedure which may introduce a bias into the results. The dependence of the MJ yield on the isolation criteria must be taken into account. In order to do this the control regions CR1 and CR2 are esimated in the slices of anti-isolation with `ptvarcone20/pT` ranging in the following intervals: [0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40].

The change of isolation criterion also biases the hadronic recoil reconstruction procedure, where the cone replacement appears to be isolation-dependent. This bias is overcome by introducing a correction to the hadronic recoil vector:

$$\vec{u}^{corr} = \vec{u}^{baseline} + \vec{u}^{iso}, \text{ where} \tag{8.6}$$

$$\vec{u}^{iso} \equiv \text{ptcone20} \cdot \vec{n}_{\ell}. \tag{8.7}$$

The unit vector $\vec{n}_{\ell}$ is aligned with the lepton direction. This correction vanishes at low isolation in the signal region but introduces a sizable correction in the anti-isolated region (see Fig. 8.3). Some residual dependence of the extrapolated distributions on the isolation criteria is still present and requires shape extrapolation procedure. The shape of the distribution of interest is estimated in three slices of `ptvarcone20/pT` isolation within [0.10, 0.20, 0.30 0.40] in CR2. For every observable $X$ the difference $\Delta[X]$ of the distribution $X$ between consecutive isolation slices is defined as:

$$H_{MJ}^{[0.1,0.2]}[X] = H_{data}^{[0.1,0.2]}[X] - H_{MC}^{[0.1,0.2]}[X]; \tag{8.8}$$

$$\Delta[X] = 1/2 \left[ (H_{MJ}^{[0.1,0.2]}[X] - H_{MJ}^{[0.2,0.3]}[X]) + (H_{MJ}^{[0.2,0.3]}[X] - H_{MJ}^{[0.3,0.4]}[X]) \right], \tag{8.9}$$
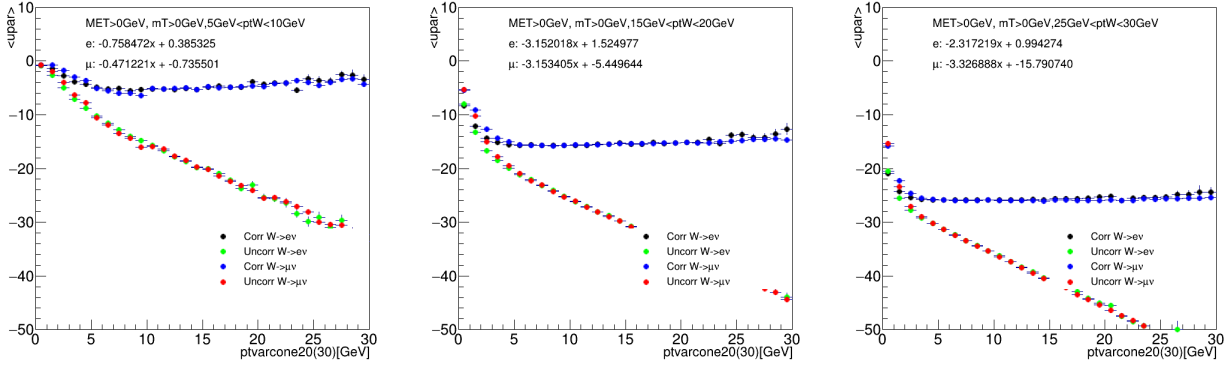
**Figure 8.3:** $< u_\parallel^\ell >$ as a function of `ptcone20`, before and after correction using data samples at $\sqrt{s} = $ 13 TeV.

where $H_X^{[x,y]}$ is the normalized distribution of $X$ in CR2 (anti-isolated signal region) satisfying $x < $ `ptvarcone20`/pT $< y$, estimated from the MC-subtracted data in CR2.

$\Delta[X]$ is supposed to be the difference between MJ sprectrum in the signal region (`ptvarcone20`/pT $< 0.1$) and the isolation slice next to it ($0.10 < $ `ptvarcone20`/pT $< 0.20$). So the extrapolated distribution to the signal region is the following:

$$H_X^{sig} = H_X^{[0.1,0.2]} - \Delta[X]. \tag{8.10}$$

The applied shift $\Delta[X]$ is assigned a 100% relative uncertainty because of the large statistical uncertainty. Figures 8.4 and 8.5 contain the results of the isolation scan in electron and muon channels respectively.
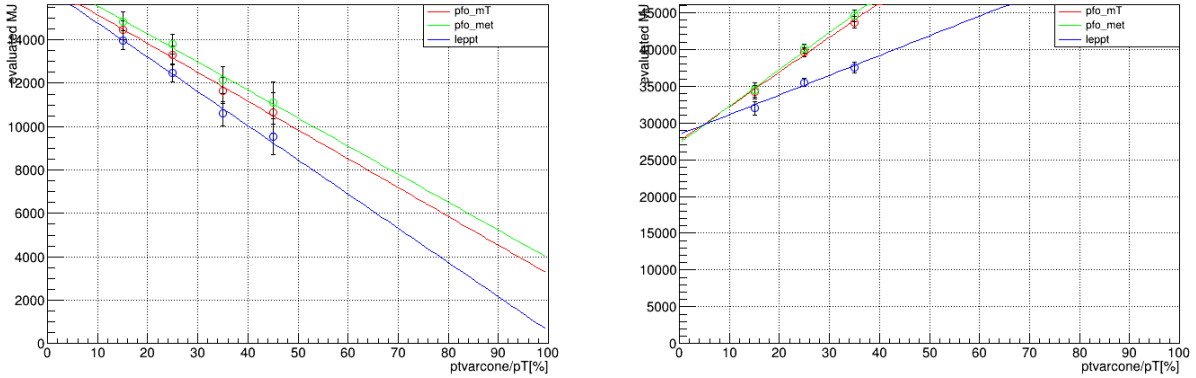


**Figure 8.4:** Isolation scan in the $W^+ \to e^+ \nu$ channel without (left) and with (right) recoil correction [164].

The MJ contributions to the kinematic distributions in slices of isolation along with the extrapolation to the signal region for the 13 TeV for electrons and muons are shown in Figures 8.6 and 8.7. The kinematic distributions with the contributions from the multijets for 5 and 13 TeV are presented in
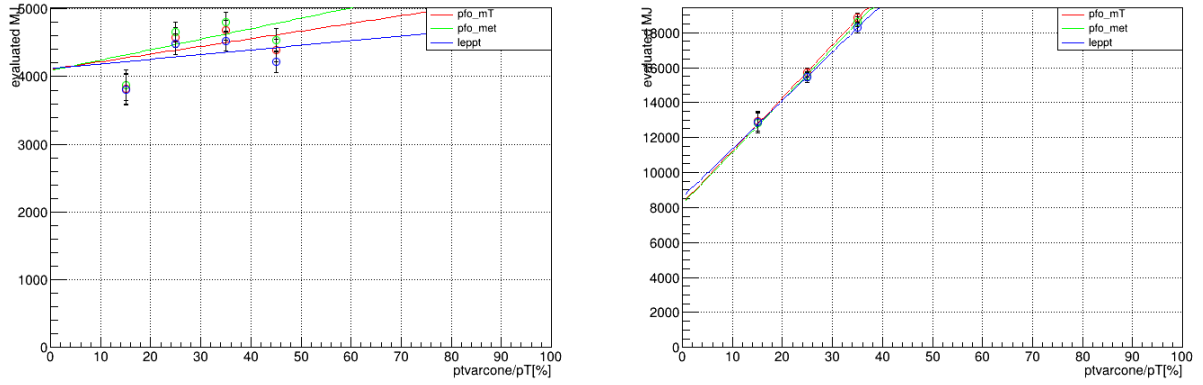
**Figure 8.5:** Isolation scan in the $W^+ \to \mu^+\nu$ channel without (left) and with (right) recoil correction [164].

sections 8.4.3 and 8.4.2 respectively.
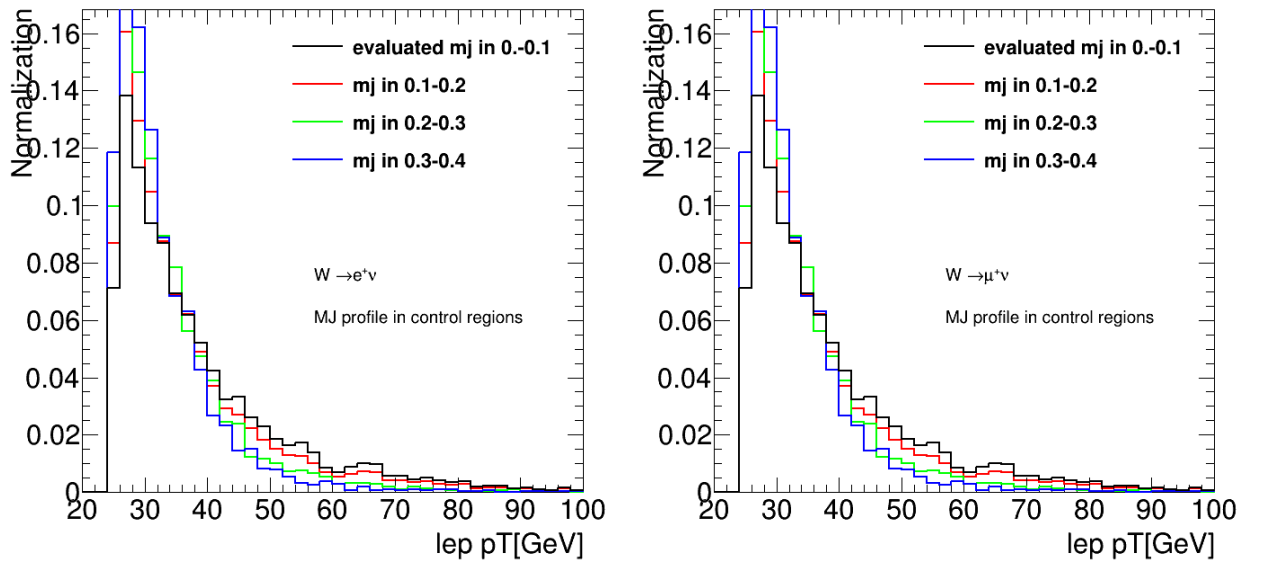


**Figure 8.6:** Extrapolation of the multijet distributions for the lepton transverse momentum in the $W^+ \to e^+\nu$ (left) and $W^+ \to \mu^+\nu$ (right) channels at $\sqrt{s} = 13$ TeV [164].

The multijet yields for all channels and energies are presented in Table 8.7 along with the associated uncertainties. The table shows that the MJ background is significantly higher in the electron channel than in the muon channel.
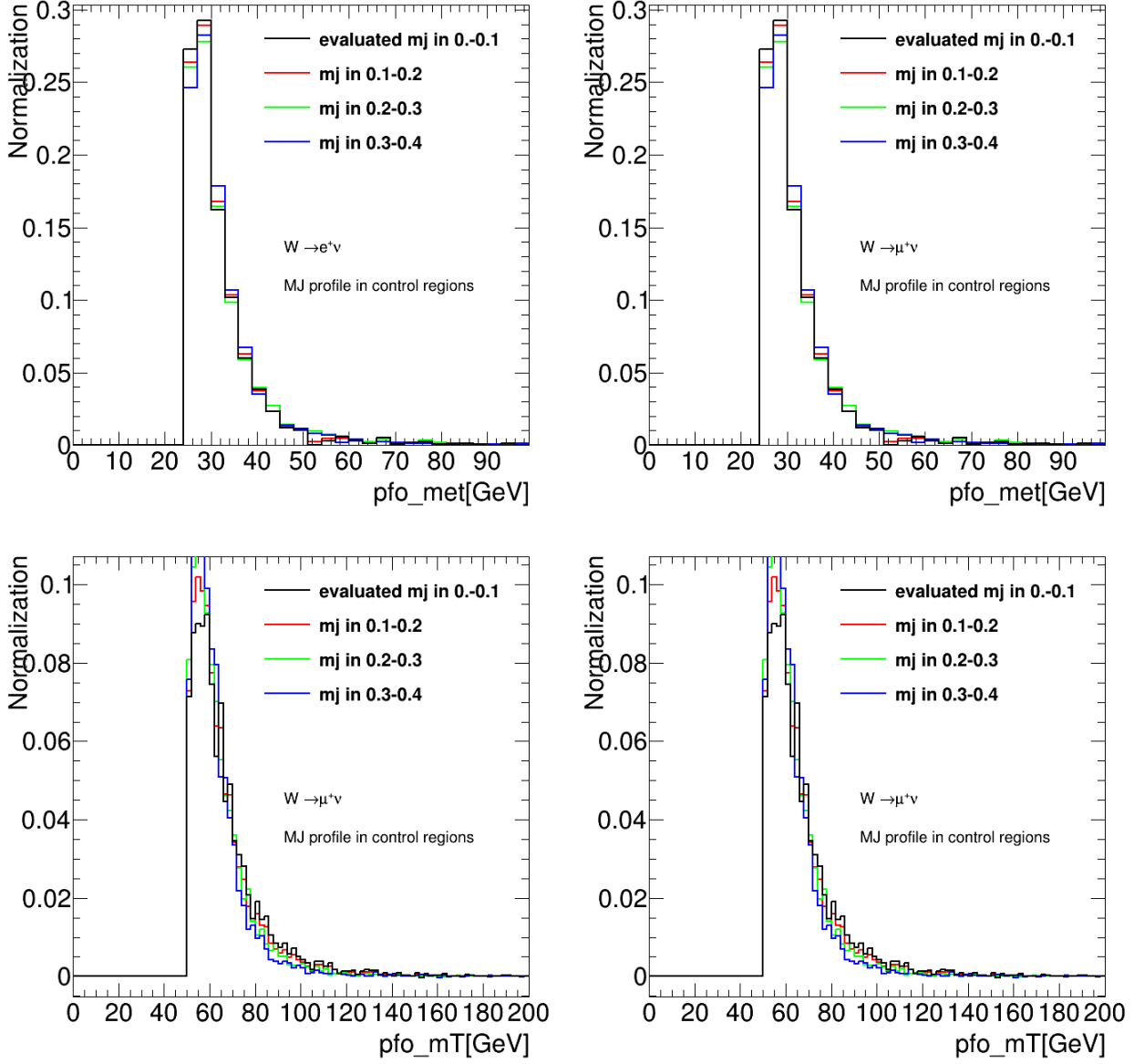
**Figure 8.7:** Extrapolation of the multijet distributions for the missing transverse energy (top) and transverse mass (bottom), in the $W^+ \to e^+\nu$ (left) and $W^+ \to \mu^+\nu$ (right) channels at $\sqrt{s} = 13$ TeV [164].

| Channel | 13 TeV | 5 TeV |
|---|---|---|
| $W^+ \to e^+ \nu$ | $27973 \pm 1756$ | $3027 \pm 554$ |
| $W^+ \to e^- \nu$ | $27388 \pm 1962$ | $2401 \pm 495$ |
| $W^+ \to \mu^+ \nu$ | $9044 \pm 796$ | $724 \pm 192$ |
| $W^+ \to \mu^- \nu$ | $9053 \pm 617$ | $755 \pm 160$ |

**Table 8.7:** Evalution of multijet background yields at 13 TeV and 5 TeV [164].

## 8.3   Z vertex reweighting

The 5 TeV MC samples have been generated to be perfectly matched to the data Although this is not the case for 13 TeV samples, which can be seen at Fig. 8.8. It is also seen from these plots that the 2017 and 2018 data were collected at two different runs under different beam conditions. To avoid possible impact on the acceptance the MC samples were reweighted to the data using $Z \rightarrow ee$ and $Z \rightarrow \mu\mu$ selections. Throughout the analysis the data collected at 13 TeV during the two runs in 2017 and 2018 are used in a combination. For this reason the correction for the MC is obtained from the combination of the two datasets (see Fig. 8.8).
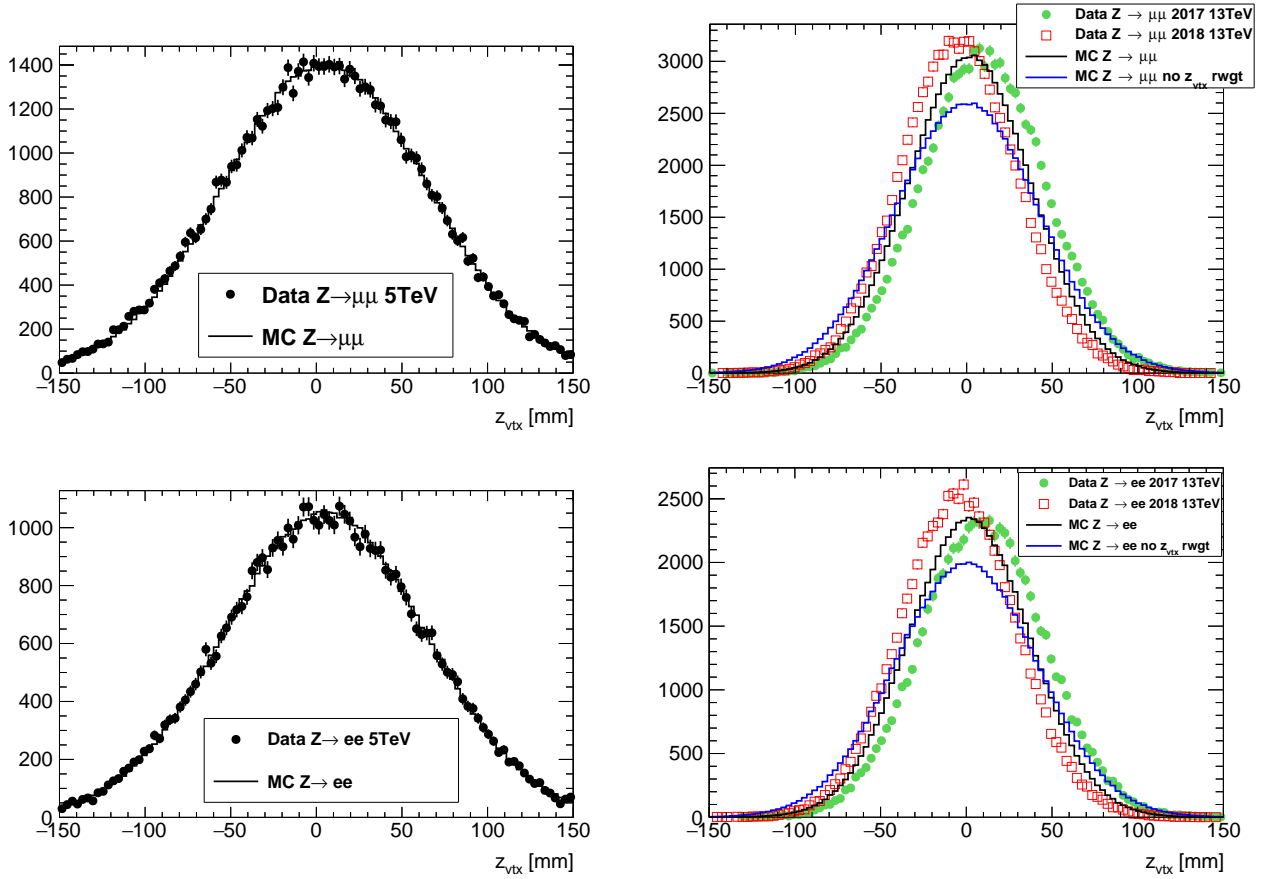


**Figure 8.8:** Distributions for the 5 TeV (left) and 13 TeV (right) low-$\mu$ dataset(s) in a $Z/\gamma^* \rightarrow \mu\mu$ (top row) and a $Z/\gamma^* \rightarrow ee$ (bottom row) selection. The data (points) is compared to $Z/\gamma^* \rightarrow \mu\mu$ or $Z/\gamma^* \rightarrow ee$ signal MC, respectively. The distributions of the $z$-position of the primary vertex selected as the hard interaction are compared for the dataset(s) and the MC simulation before ("no $z_{vtx}$ rwgt", blue, only 13 TeV) and after reweighting (black). For the 13 TeV data the 2017 and 2018 data are shown separately and all distributions are (roughly) normalised to the same number of selected events in the 2017 dataset [146].

The reweighting derived from Z events is also used for the W MC samples events. The plots shown on Fig. 8.9 demonstrate that the correction derived from the Z events works well for all the W channels.
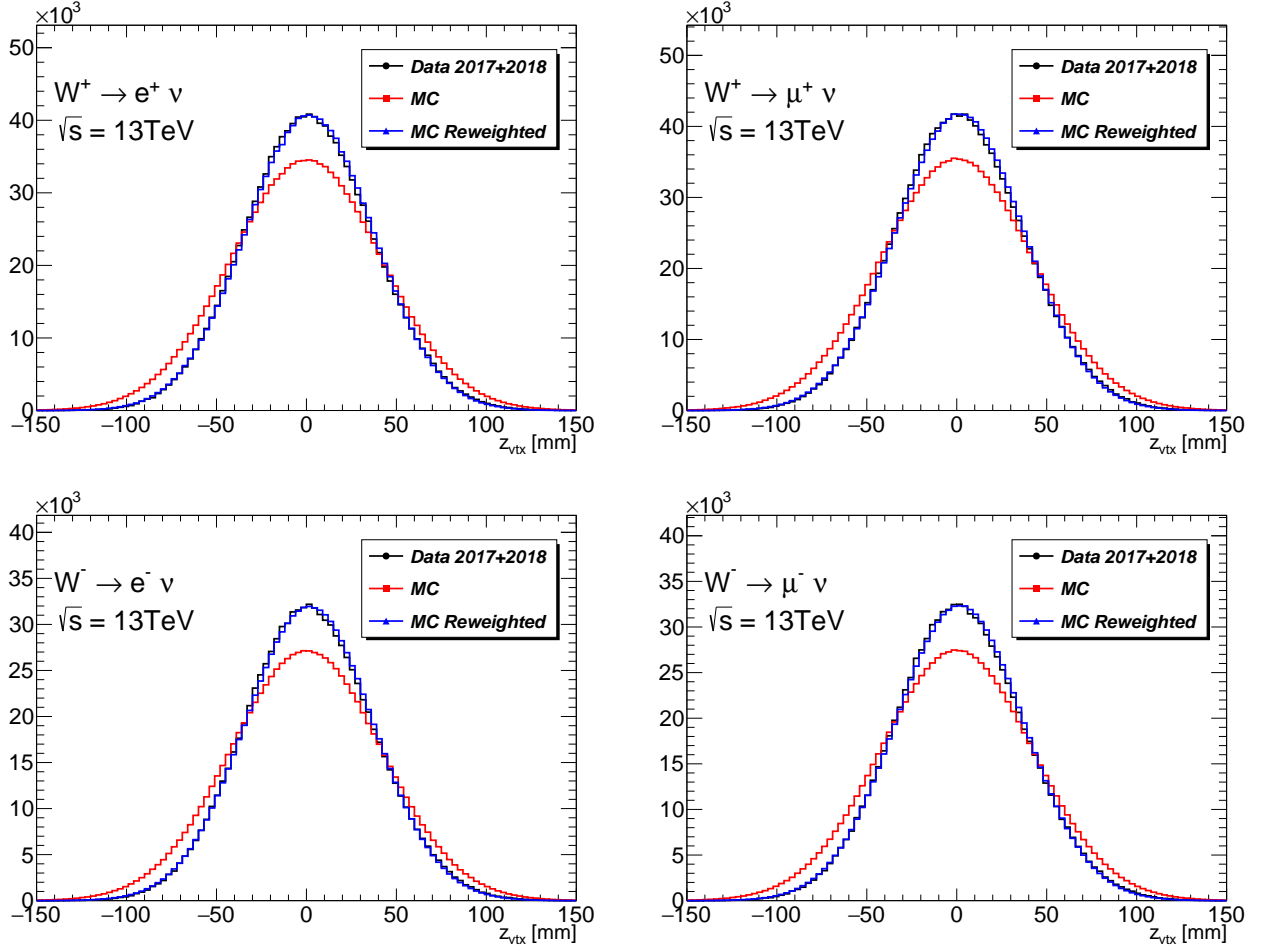
**Figure 8.9:** Distributions for 13 TeV low-$\mu$ dataset(s) in a $W^+$ (top row) and a $W^-$ (bottom row) selection in electron (left) and muon (right) channels. The combined 2017 and 2018 data (black) is compared to W MC signal before (red) and after (blue) the reweighting. All distributions are normalised to the same number of selected events in 2017+2018 dataset [146].

The reweighting causes minor effect on the kinematic distributions. The ratios of these distributions before and after the Z vertex reweighting are demonstrated for the $W^+ \rightarrow \mu\nu$ channel in Fig. 8.10.

The procedure of the vertex reweighting follows the ideology of the W mass measurement performed at 7 TeV [20]. Despite the effect of the reweighting was found to be small for the measurement of the W boson transverse spectrum, the reweighting might still have a notable effect on the W mass measurement, that would use the obtained W pT spectrum as an input. A possible way to estimate the effect of the reweighting on the W boson mass is to use a template fit. The method of the template fits is described in more detail in the internal note for the 7 TeV measurement [165]. Considering the observables that are sensitive to $m_W$ we can estimate the shift of the W mass value due to the
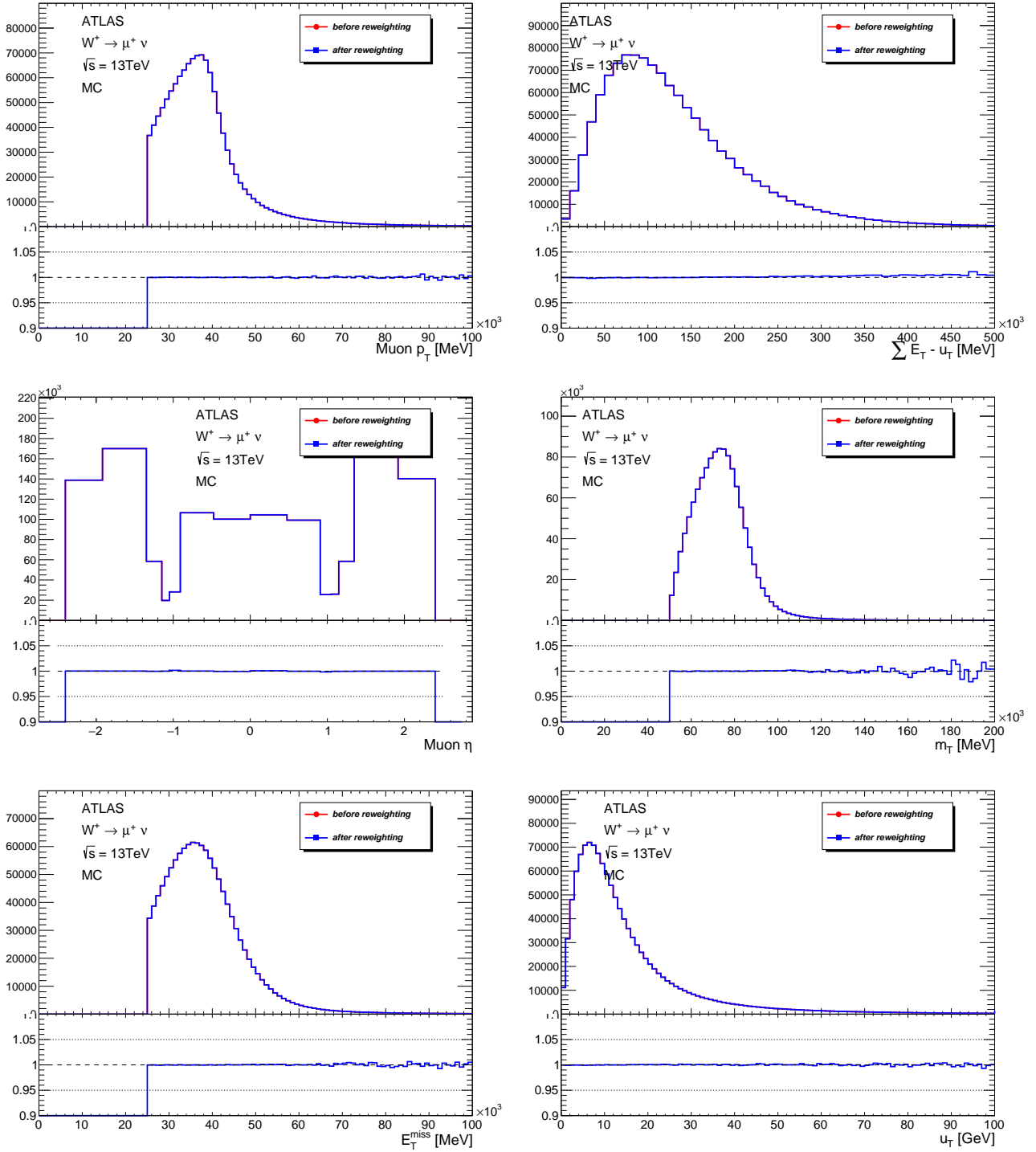
**Figure 8.10:** Distributions for 13 TeV low-$\mu$ dataset(s) in a $W^+$ (top row) and a $W^-$ (bottom row) selection in electron (left) and muon (right) channels. The combined 2017 and 2018 data (black) is compared to W MC signal before (red) and after (blue) the reweighting. All distributions are normalised to the same number of selected events in 2017+2018 dataset [146].

| Observable | $m_T$ | $E_T^{miss}$ | $p_T$ |
|------------|-------|--------------|-------|
| $\Delta m_W$ | 1.4 MeV | 5.1 MeV | 0.2 MeV |

**Table 8.10:** Estimate of the Z vertex reweighting effect on the W boson mass.

reweighting. The results for $W \rightarrow \mu\nu$ channel are provided in the Table 8.10. They demonstrate a non-negligible effect.

## 8.4 W analysis event selection and control plots

### 8.4.1 Event selection

In both cases of 5 and 13 TeV events with $W \rightarrow \ell\nu$ candidate were selected based on a single-lepton trigger requirement. The trigger for $W \rightarrow e\nu$ event candidate `HLT_e15_lh1oose_nod0_L1EM12`, requires at least one reconstructed electron with $E_T$ larger than 15 GeV passing *loose* identification requirements. Candidates for $W \rightarrow \mu\nu$ were triggered by `HLT_mu14` trigger, requiring one medium muon with $E_T$ larger than 14 GeV.

Events are required to contain exactly one lepton (muon or electron) candidate having $p_T > 25$GeV. Electrons are required to have $|\eta| < 2.47$ excluding transition region $1.37 < |\eta| < 1.52$. Muons Events with additional leptons of the same flavour with transverse momentum greater than 20 GeV satisfying some ID criteria are discarded, to better reject the $Z$ background. The ID point is *medium* for the muon channel, and *loose* for the electron channel. There is no requirement on the number of leptons from the different channel than the channel under study.

To suppress background, in particular from multijet processes, events are required to have $E_T^{miss}$ greater than 25 GeV. The $W$ boson transverse mass $m_T$ is demanded to be larger than 50 GeV. This transverse mass is defined as follows:

$$m_T = \sqrt{2p_T^\nu p_T^l (1 - \cos \Delta\phi^\nu)} \tag{8.11}$$

Tables 8.10,8.10,8.10,8.10 contain signal selection event yields for the $W^\pm \rightarrow \ell^\pm\nu$ at $\sqrt{s} = 5$ TeV low-$\mu$ dataset. Similarly the tables 8.10,8.10,8.10,8.10 contain the corresponding numbers for the 13 TeV low-$\mu$ dataset. The listed uncertainties are statistical. Table 8.10 provides a comparison between observed and expected yields. Events denoted as $W \rightarrow \ell\nu$ in the tables and the plots contain the sum of background events coming from $W \rightarrow \tau\nu$ and $W$ leptonic decays other than the signal.

| Cut | Data | Signal | | | $W^{\pm} \to \ell^{\pm}\nu$ BG | | | $Z \to \ell\ell$ | | | Top | | | Diboson | | | Multijet | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| One electron | 1993720 | 643610 | ± | 260 | 32940 | ± | 190 | 44338 | ± | 71 | 1754.4 | ± | 3.9 | 772.2 | ± | 3.7 | - | | |
| Electron trig matched | 1907724 | 612940 | ± | 250 | 30790 | ± | 190 | 42100 | ± | 69 | 1698.5 | ± | 3.8 | 741.1 | ± | 3.6 | - | | |
| Isolation | 1438941 | 610320 | ± | 250 | 30590 | ± | 190 | 41923 | ± | 69 | 1663.6 | ± | 3.8 | 722.5 | ± | 3.6 | - | | |
| $p_T^e > 25$GeV | 720284 | 482240 | ± | 220 | 14790 | ± | 130 | 31955 | ± | 53 | 1464.5 | ± | 3.5 | 592.1 | ± | 3.2 | - | | |
| $E_T^{miss} > 25$GeV | 440605 | 421510 | ± | 210 | 9650 | ± | 100 | 1336 | ± | 20 | 1223 | ± | 3.2 | 420.8 | ± | 2.4 | - | | |
| $m_T > 50$GeV | 430620 | 417430 | ± | 210 | 8800 | ± | 96 | 1047 | ± | 16 | 944.3 | ± | 2.9 | 373.5 | ± | 2.2 | 3030 | ± | 550 |

**Table 8.10:** Analysis cut flow for $W^+ \to e^+\nu$ 5 TeV signal selection. Lepton $p_T$ is required to be over 18 GeV before the final cut.

| Cut | Data | Signal | | | $W^{\pm} \to \ell^{\pm}\nu$ BG | | | $Z \to \ell\ell$ | | | Top | | | Diboson | | | Multijet | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| One electron | 7915023 | 1797340 | ± | 390 | 92520 | ± | 270 | 147490 | ± | 140 | 63207 | ± | 89 | 3069 | ± | 63 | - | | |
| Electron trig matched | 7840239 | 1709140 | ± | 380 | 86370 | ± | 260 | 139760 | ± | 140 | 61110 | ± | 88 | 2967 | ± | 62 | - | | |
| Isolation | 5413483 | 1698430 | ± | 380 | 85560 | ± | 260 | 138890 | ± | 140 | 59834 | ± | 87 | 2939 | ± | 61 | - | | |
| $p_T^e > 25$GeV | 2452868 | 1342200 | ± | 330 | 44450 | ± | 190 | 106270 | ± | 110 | 53811 | ± | 82 | 2565 | ± | 58 | - | | |
| $E_T^{miss} > 25$GeV | 1275513 | 1136520 | ± | 310 | 28580 | ± | 150 | 8313 | ± | 46 | 45707 | ± | 75 | 1990 | ± | 53 | - | | |
| $m_T > 50$GeV | 1207776 | 1117560 | ± | 310 | 24760 | ± | 130 | 6443 | ± | 36 | 34580 | ± | 65 | 1718 | ± | 50 | 28000 | ± | 1800 |

**Table 8.10:** Analysis cut flow for $W^+ \to e^+\nu$ 13 TeV signal selection. Lepton $p_T$ is required to be over 18 GeV before the final cut.

| Cut | Data | Signal | | | $W^{\pm} \to \ell^{\pm}\nu$ BG | | | $Z \to \ell\ell$ | | | Top | | | Diboson | | | Multijet | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| One muon | 2434459 | 760980 | ± | 280 | 35090 | ± | 200 | 37015 | ± | 82 | 2025.3 | ± | 4.1 | 864.7 | ± | 3.7 | - | | |
| Muon trig matched | 2353403 | 664100 | ± | 260 | 30610 | ± | 190 | 32554 | ± | 76 | 1725.6 | ± | 3.8 | 746.6 | ± | 3.4 | - | | |
| Isolation | 1186616 | 659200 | ± | 260 | 30400 | ± | 190 | 32303 | ± | 76 | 1574.6 | ± | 3.7 | 710.1 | ± | 3.3 | - | | |
| $p_T^{\mu} > 25$GeV | 632016 | 508270 | ± | 230 | 13900 | ± | 130 | 22556 | ± | 57 | 1335.3 | ± | 3.4 | 568.2 | ± | 2.9 | - | | |
| $E_T^{miss} > 25$GeV | 470856 | 442600 | ± | 210 | 8700 | ± | 100 | 9959 | ± | 31 | 1111.8 | ± | 3 | 424.5 | ± | 2.5 | - | | |
| $m_T > 50$GeV | 457053 | 438280 | ± | 210 | 7879 | ± | 97 | 9649 | ± | 27 | 879.7 | ± | 2.8 | 381.7 | ± | 2.3 | 720 | ± | 190 |

**Table 8.10:** Analysis cut flow for $W^+ \to \mu^+\nu$ 5 TeV signal selection. Lepton $p_T$ is required to be over 18 GeV before the final cut.

## 8.4.2 $\sqrt{s} = 13$ TeV dataset control plots

Control plots for the 13 TeV low-$\mu$ dataset are provided here after applying all corrections described in Section 7, and after applying the selection described above in this section. In each figure, the right(left)-hand column shows distributions for the $W^+$ ($W^-$) process. The top (bottom) row shows the muon (electron) decay channel. In the ratio panels, the grey band is the total systematic uncertainty, whilst the brown band adds the MC statistical uncertainty in quadrature on top of it. In regions of the distributions insensitive to the modelling of $p_T^W$ there is generally good agreement between data and predictions. The bulk of the $m_T$ distribution is a typical example of distribution that is mostly insensitive to the modelling of $p_T^W$. The $u_T$ distribution depends a lot on the modelling of the W boson transverse momentum and it demonstrates the highest discrepancy between the data and MC. Therefore it can be concluded that the baseline simulation is not modelling $p_T^W$ satisfactorily.

| Cut | Data | Signal | | | $W^\pm \to \ell^\pm \nu$ BG | | | $Z \to \ell\ell$ | | | Top | | | Diboson | | | Multijet | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| One muon | 9570104 | 2100770 | ± | 410 | 83110 | ± | 270 | 2019400 | ± | 2200 | 71602 | ± | 94 | 3442 | ± | 63 | - | | |
| Muon trig matched | 9382783 | 1840550 | ± | 390 | 72820 | ± | 250 | 1750400 | ± | 2000 | 61519 | ± | 87 | 2956 | ± | 59 | - | | |
| Isolation | 3905612 | 1821750 | ± | 380 | 71780 | ± | 250 | 595700 | ± | 1100 | 56849 | ± | 84 | 2916 | ± | 59 | - | | |
| $p_T^\mu > 25$GeV | 1930655 | 1393330 | ± | 340 | 34470 | ± | 170 | 170840 | ± | 490 | 49338 | ± | 78 | 2471 | ± | 54 | - | | |
| $E_T^{miss} > 25$GeV | 1321407 | 1173860 | ± | 310 | 21450 | ± | 140 | 51090 | ± | 180 | 41956 | ± | 72 | 1930 | ± | 49 | - | | |
| $m_T > 50$GeV | 1244892 | 1153800 | ± | 310 | 18270 | ± | 130 | 38304 | ± | 81 | 32375 | ± | 63 | 1705 | ± | 44 | 9040 | ± | 800 |

**Table 8.10:** Analysis cut flow for $W^+ \to \mu^+ \nu$ 13 TeV signal selection. Lepton $p_T$ is required to be over 18 GeV before the final cut.

| Cut | Data | Signal | | | $W^\pm \to \ell^\pm \nu$ BG | | | $Z \to \ell\ell$ | | | Top | | | Diboson | | | Multijet | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| One electron | 1724472 | 374900 | ± | 200 | 24150 | ± | 160 | 41995 | ± | 70 | 1590.5 | ± | 2.9 | 684.8 | ± | 4 | - | | |
| Electron trig matched | 1645694 | 359010 | ± | 200 | 22070 | ± | 160 | 39854 | ± | 68 | 1539.9 | ± | 2.9 | 655.7 | ± | 3.9 | - | | |
| Isolation | 1176976 | 357660 | ± | 200 | 21920 | ± | 160 | 39686 | ± | 68 | 1504.6 | ± | 2.8 | 640.7 | ± | 3.8 | - | | |
| $p_T^e > 25$GeV | 529183 | 302070 | ± | 180 | 11920 | ± | 110 | 30214 | ± | 52 | 1330.8 | ± | 2.6 | 532.9 | ± | 3.5 | - | | |
| $E_T^{miss} > 25$GeV | 281957 | 266750 | ± | 170 | 8084 | ± | 90 | 1293 | ± | 20 | 1112.5 | ± | 2.4 | 380 | ± | 3 | - | | |
| $m_T > 50$GeV | 274329 | 264540 | ± | 170 | 7317 | ± | 84 | 994 | ± | 16 | 855.2 | ± | 2.1 | 338.1 | ± | 2.9 | 2400 | ± | 500 |

**Table 8.10:** Analysis cut flow for $W^- \to e^- \nu$ 5 TeV signal selection. Lepton $p_T$ is required to be over 18 GeV before the final cut.

| Cut | Data | Signal | | | $W^\pm \to \ell^\pm \nu$ BG | | | $Z \to \ell\ell$ | | | Top | | | Diboson | | | Multijet | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| One electron | 7471742 | 1323710 | ± | 330 | 78230 | ± | 230 | 140980 | ± | 140 | 61951 | ± | 86 | 3059 | ± | 58 | - | | |
| Electron trig matched | 7402574 | 1267710 | ± | 330 | 72240 | ± | 230 | 133580 | ± | 140 | 59950 | ± | 85 | 2968 | ± | 57 | - | | |
| Isolation | 4949352 | 1260540 | ± | 330 | 71550 | ± | 230 | 132740 | ± | 140 | 58689 | ± | 84 | 2937 | ± | 57 | - | | |
| $p_T^e > 25$GeV | 2113364 | 1053510 | ± | 300 | 39660 | ± | 160 | 101350 | ± | 110 | 52923 | ± | 79 | 2544 | ± | 53 | - | | |
| $E_T^{miss} > 25$GeV | 1008915 | 900640 | ± | 280 | 25900 | ± | 130 | 7954 | ± | 45 | 45065 | ± | 73 | 1962 | ± | 48 | - | | |
| $m_T > 50$GeV | 949362 | 887810 | ± | 270 | 22400 | ± | 120 | 6052 | ± | 35 | 34177 | ± | 64 | 1695 | ± | 44 | 27400 | ± | 2000 |

**Table 8.10:** Analysis cut flow for $W^- \to e^- \nu$ 13 TeV signal selection. Lepton $p_T$ is required to be over 18 GeV before the final cut.

| Cut | Data | Signal | | | $W^\pm \to \ell^\pm \nu$ BG | | | $Z \to \ell\ell$ | | | Top | | | Diboson | | | Multijet | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| One muon | 2075709 | 440560 | ± | 220 | 22510 | ± | 170 | 34440 | ± | 80 | 1835.6 | ± | 3.1 | 751.5 | ± | 3.3 | - | | |
| Muon trig matched | 2002955 | 383720 | ± | 200 | 19640 | ± | 160 | 30277 | ± | 75 | 1561.6 | ± | 2.9 | 648 | ± | 3.1 | - | | |
| Isolation | 883078 | 381010 | ± | 200 | 19450 | ± | 160 | 30046 | ± | 74 | 1411 | ± | 2.7 | 616.9 | ± | 2.9 | - | | |
| $p_T^\mu > 25$GeV | 426119 | 314370 | ± | 180 | 9370 | ± | 110 | 20749 | ± | 56 | 1202.1 | ± | 2.5 | 505 | ± | 2.5 | - | | |
| $E_T^{miss} > 25$GeV | 298992 | 276060 | ± | 170 | 5893 | ± | 89 | 8716 | ± | 29 | 1004.2 | ± | 2.3 | 372.6 | ± | 2 | - | | |
| $m_T > 50$GeV | 287870 | 273710 | ± | 170 | 5158 | ± | 82 | 8408 | ± | 26 | 788.2 | ± | 2 | 335.6 | ± | 1.9 | 760 | ± | 160 |

**Table 8.10:** Analysis cut flow for $W^- \to \mu^- \nu$ 5 TeV signal selection. Lepton $p_T$ is required to be over 18 GeV before the final cut.

| Cut | Data | Signal | | | $W^\pm \to \ell^\pm \nu$ BG | | | $Z \to \ell\ell$ | | | Top | | | Diboson | | | Multijet | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| One muon | 8773414 | 1518070 | ± | 360 | 64930 | ± | 230 | 2019900 | ± | 2200 | 70580 | ± | 90 | 3230 | ± | 60 | - | | |
| Muon trig matched | 8597493 | 1322980 | ± | 330 | 56520 | ± | 210 | 1750300 | ± | 2000 | 60579 | ± | 84 | 2806 | ± | 56 | - | | |
| Isolation | 3298569 | 1310310 | ± | 330 | 55680 | ± | 210 | 593700 | ± | 1100 | 55949 | ± | 80 | 2751 | ± | 55 | - | | |
| $p_T^\mu > 25$GeV | 1561721 | 1069770 | ± | 300 | 28230 | ± | 150 | 166810 | ± | 490 | 48544 | ± | 75 | 2362 | ± | 52 | - | | |
| $E_T^{miss} > 25$GeV | 1030406 | 910150 | ± | 280 | 17380 | ± | 120 | 47370 | ± | 180 | 41259 | ± | 69 | 1842 | ± | 46 | - | | |
| $m_T > 50$GeV | 963568 | 896850 | ± | 270 | 14710 | ± | 110 | 34572 | ± | 80 | 31772 | ± | 61 | 1598 | ± | 43 | 9050 | ± | 620 |

**Table 8.10:** Analysis cut flow for $W^- \to \mu^- \nu$ 13 TeV signal selection. Lepton $p_T$ is required to be over 18 GeV before the final cut.

| Selection | Observed | Expected | | |
|---|---|---|---|---|
| 5TeV $W^+ \to e^+ \nu$ | 430620 | 431620 | $\pm$ | 600 |
| 5TeV $W^+ \to \mu^+ \nu$ | 457053 | 457790 | $\pm$ | 300 |
| 5TeV $W^- \to e^- \nu$ | 274329 | 276450 | $\pm$ | 530 |
| 5TeV $W^- \to \mu^- \nu$ | 287870 | 289160 | $\pm$ | 250 |
| 13TeV $W^+ \to e^+ \nu$ | 1207776 | 1213000 | $\pm$ | 1800 |
| 13TeV $W^+ \to \mu^+ \nu$ | 1244892 | 1253490 | $\pm$ | 870 |
| 13TeV $W^- \to e^- \nu$ | 949362 | 979500 | $\pm$ | 2000 |
| 13TeV $W^- \to \mu^- \nu$ | 963568 | 988560 | $\pm$ | 690 |

**Table 8.10:** Observed and Expected yield comparison for all signal selections.
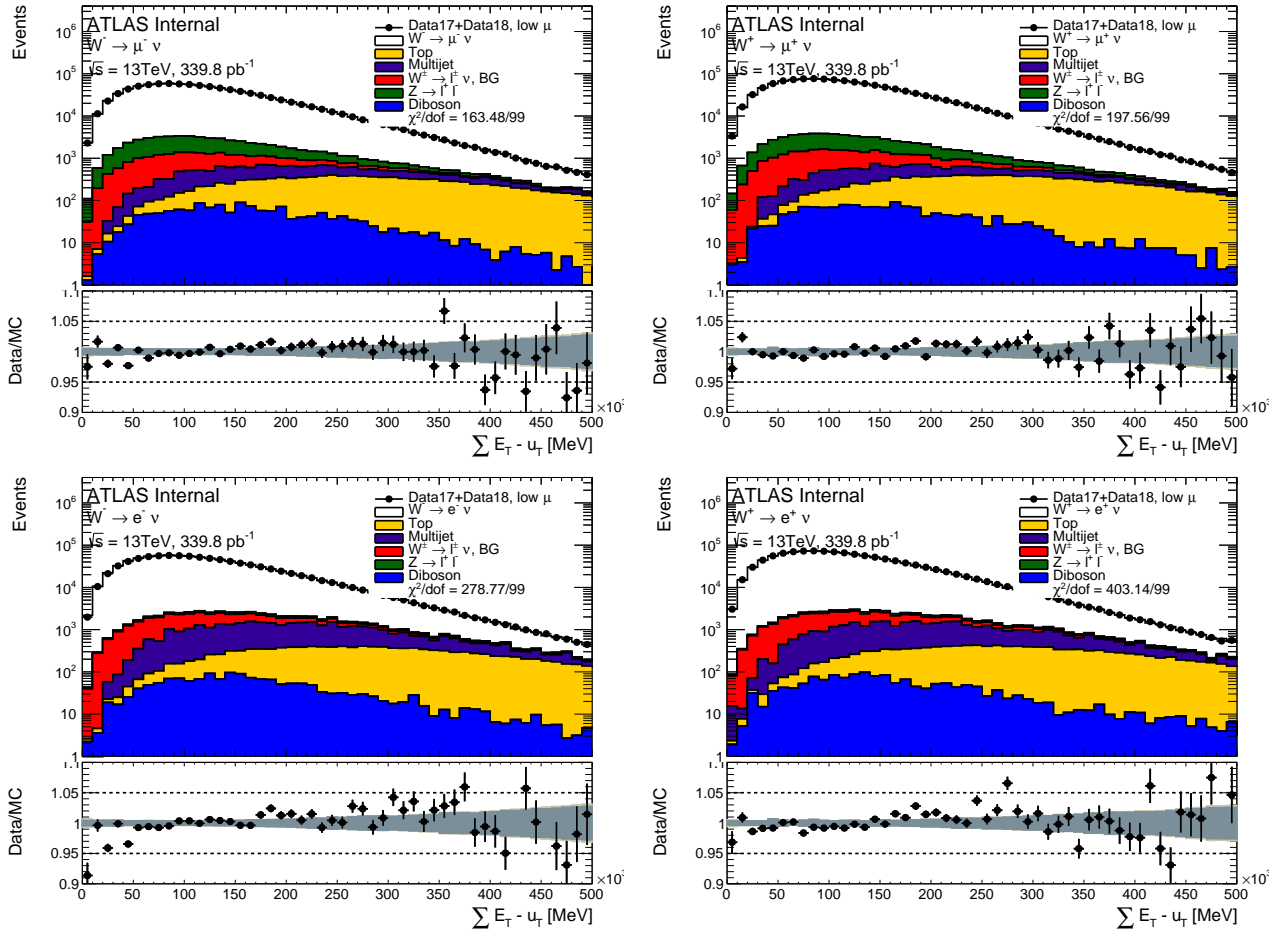


**Figure 8.11:** $\Sigma \bar{E}_T$ distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.
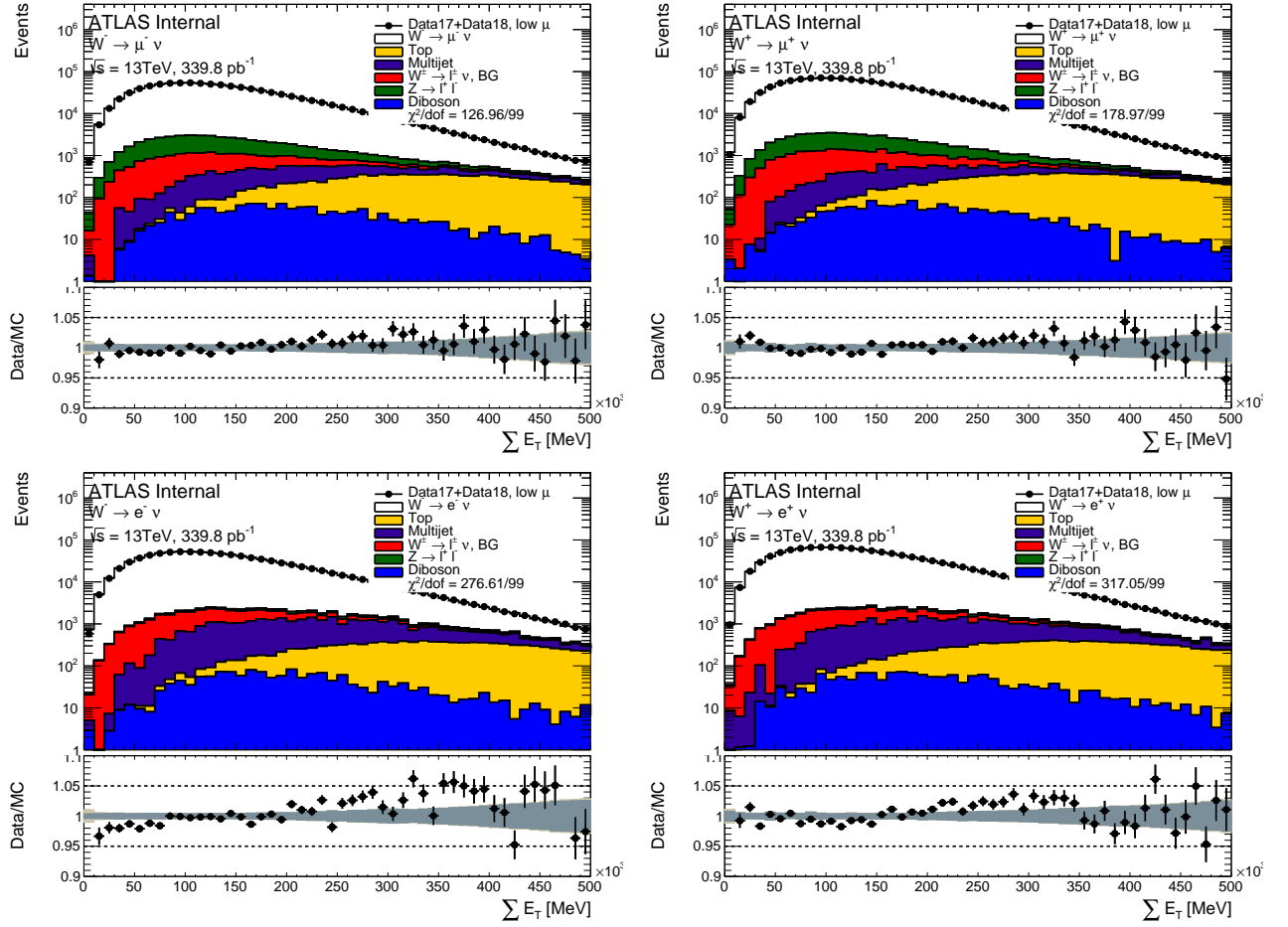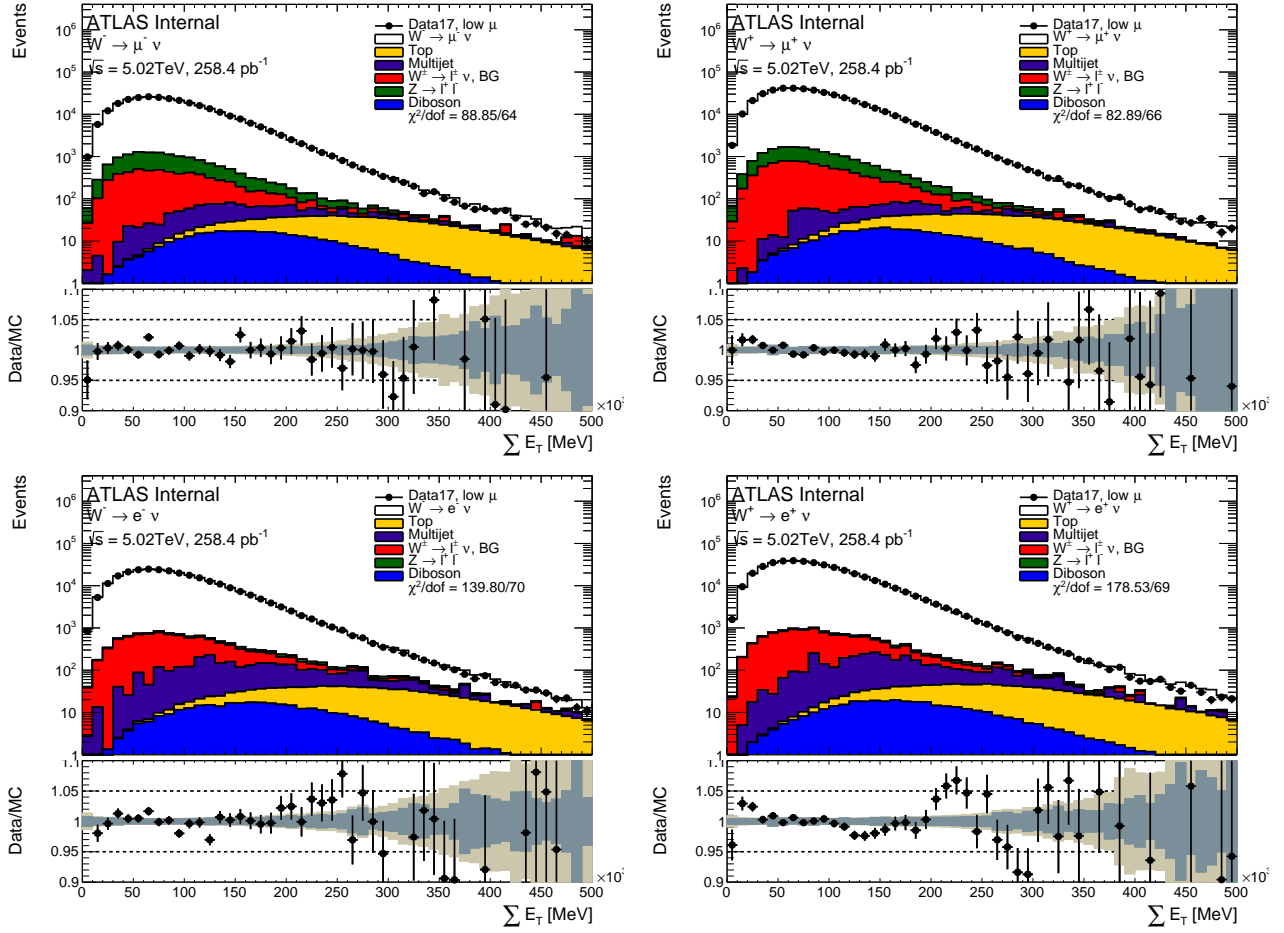
**Figure 8.12:** $\Sigma E_T$ distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.
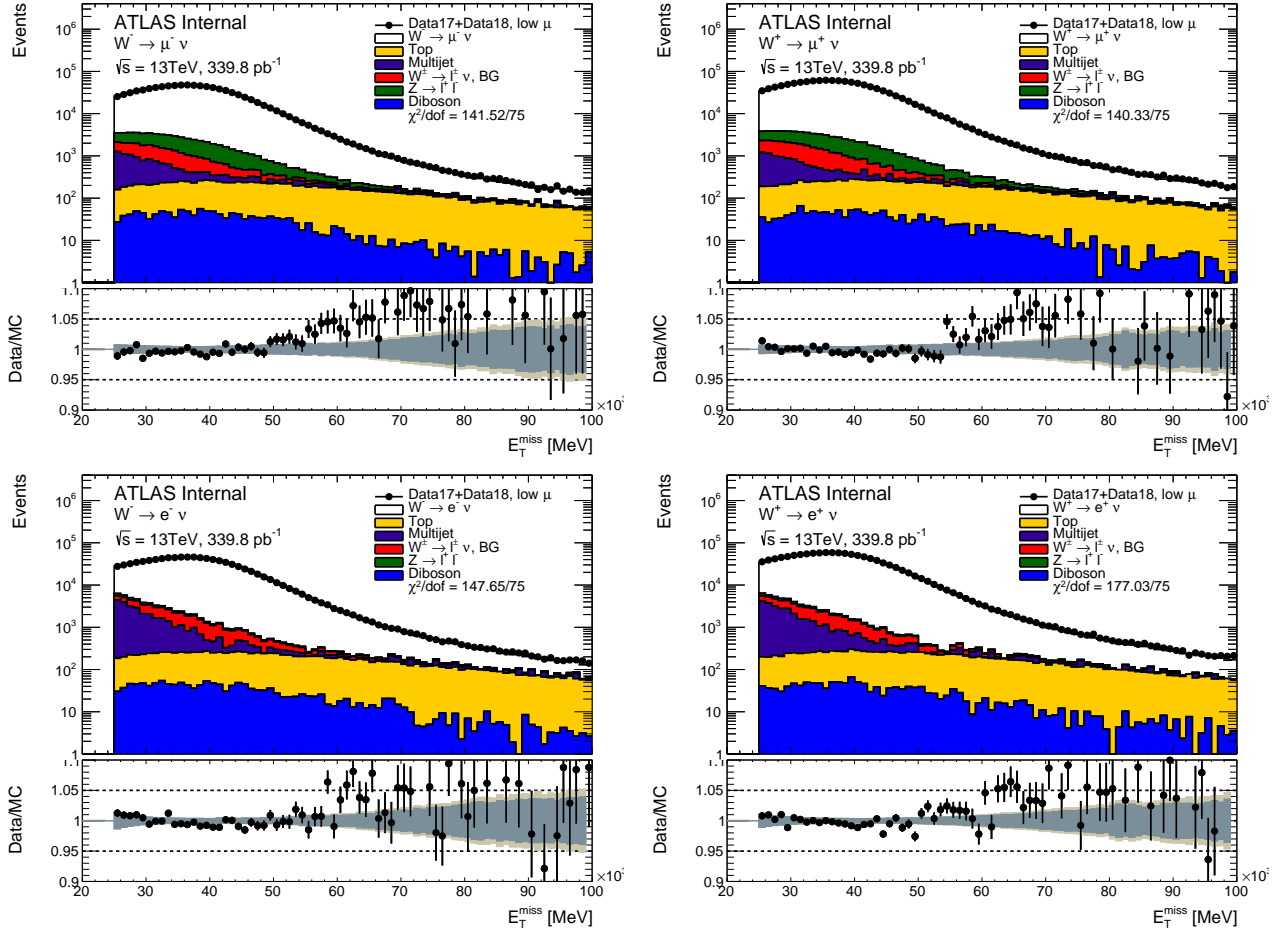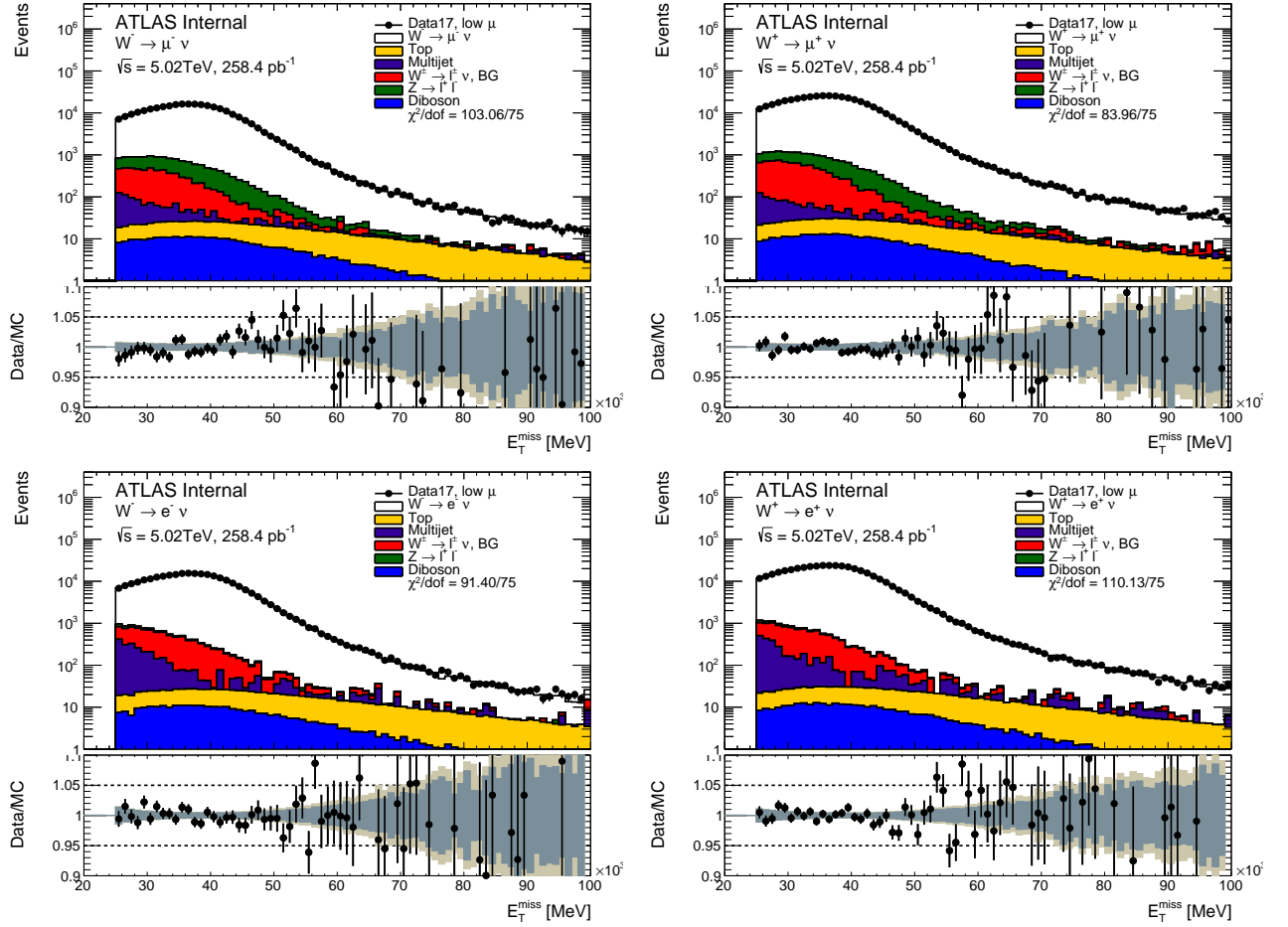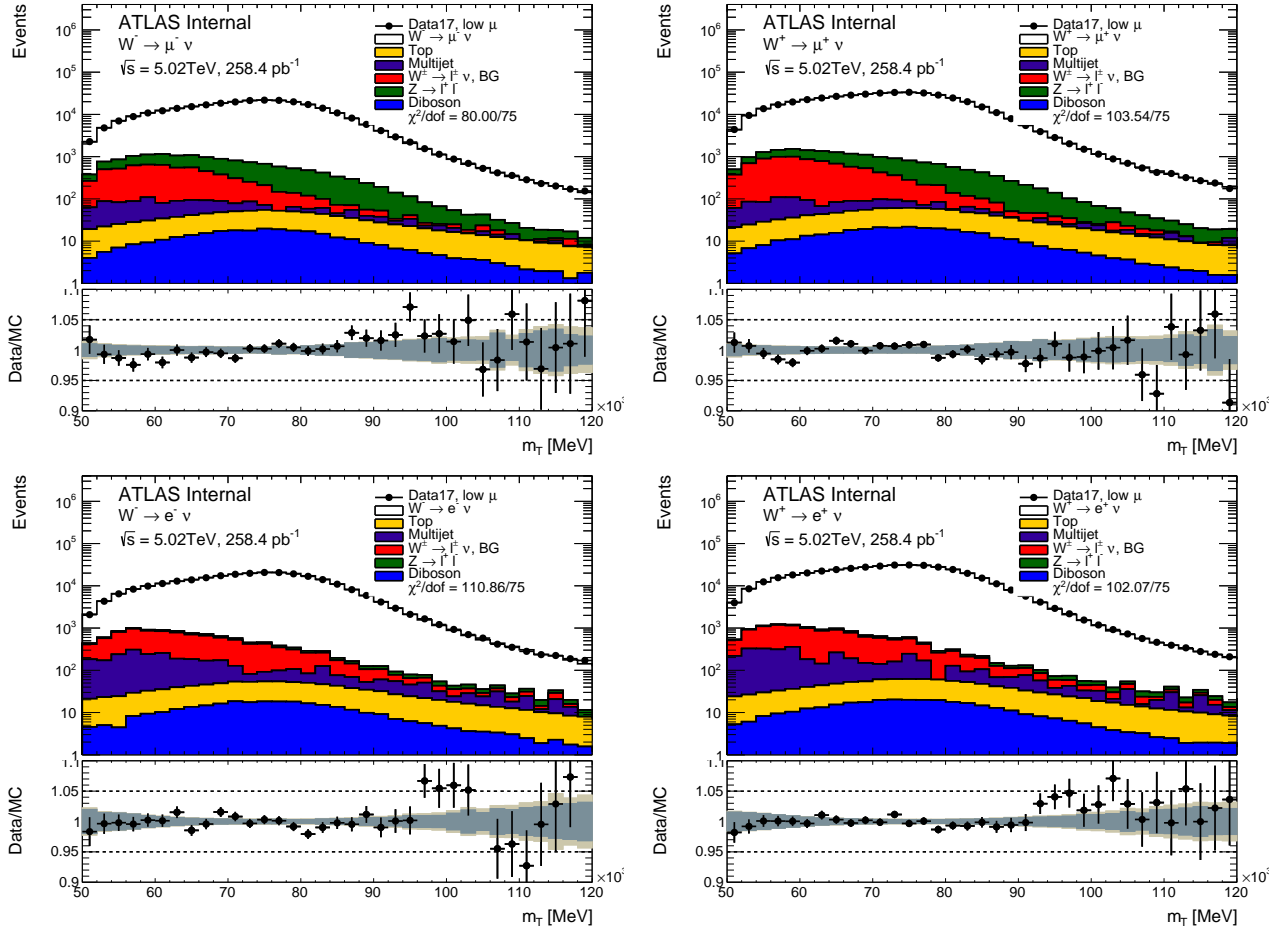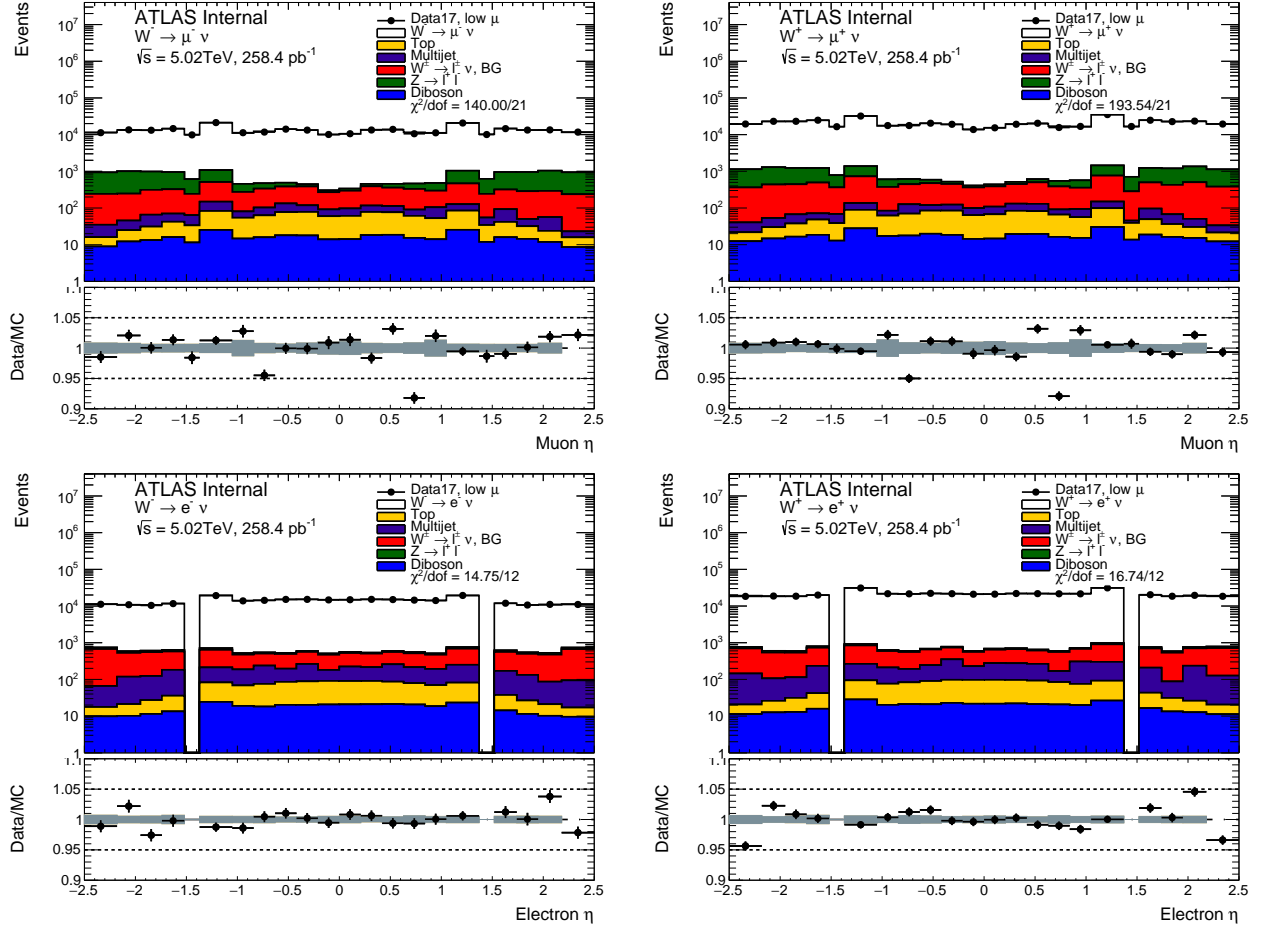
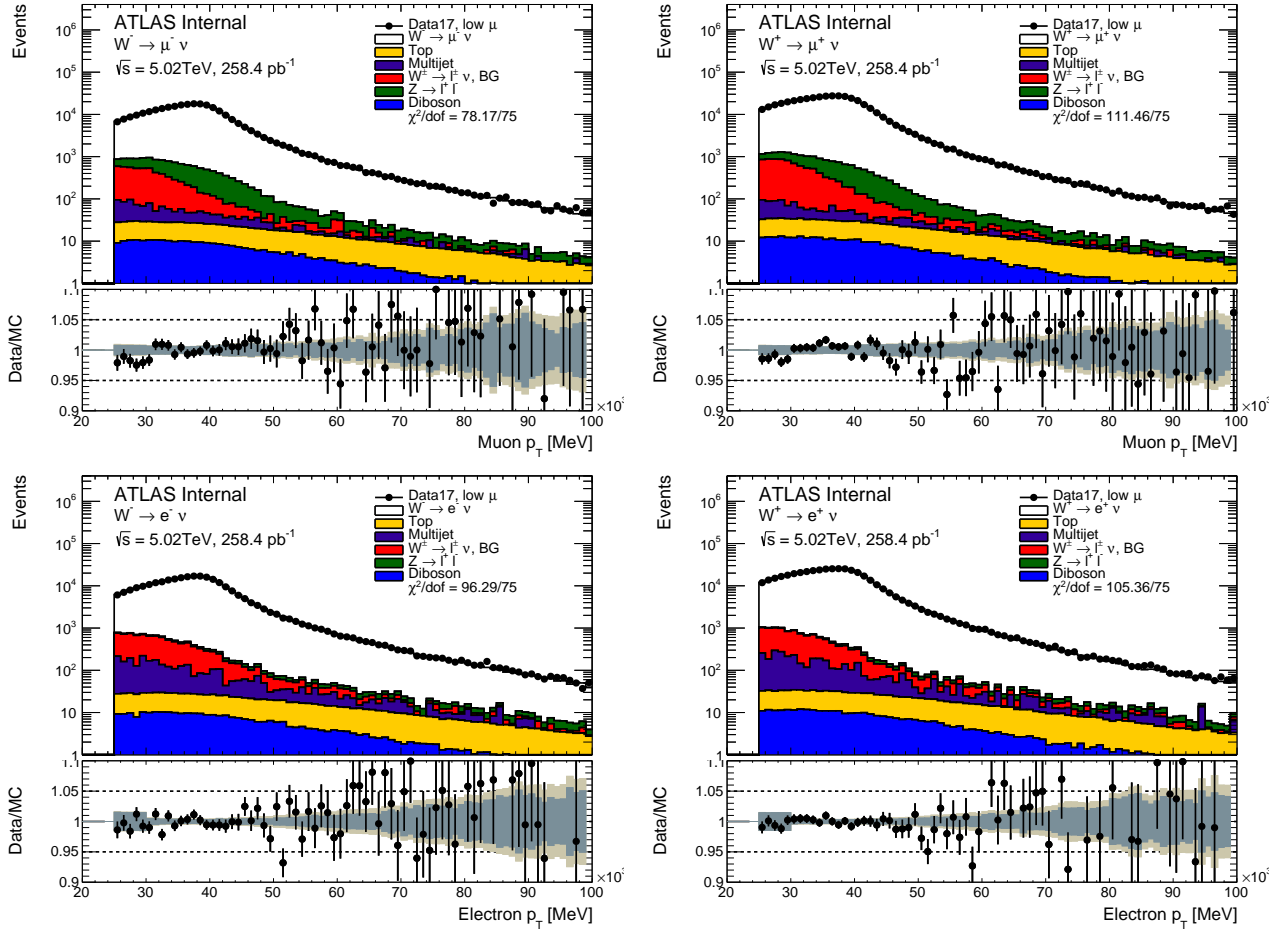**Figure 8.13:** $\vec{E}_T^{miss}$ distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.
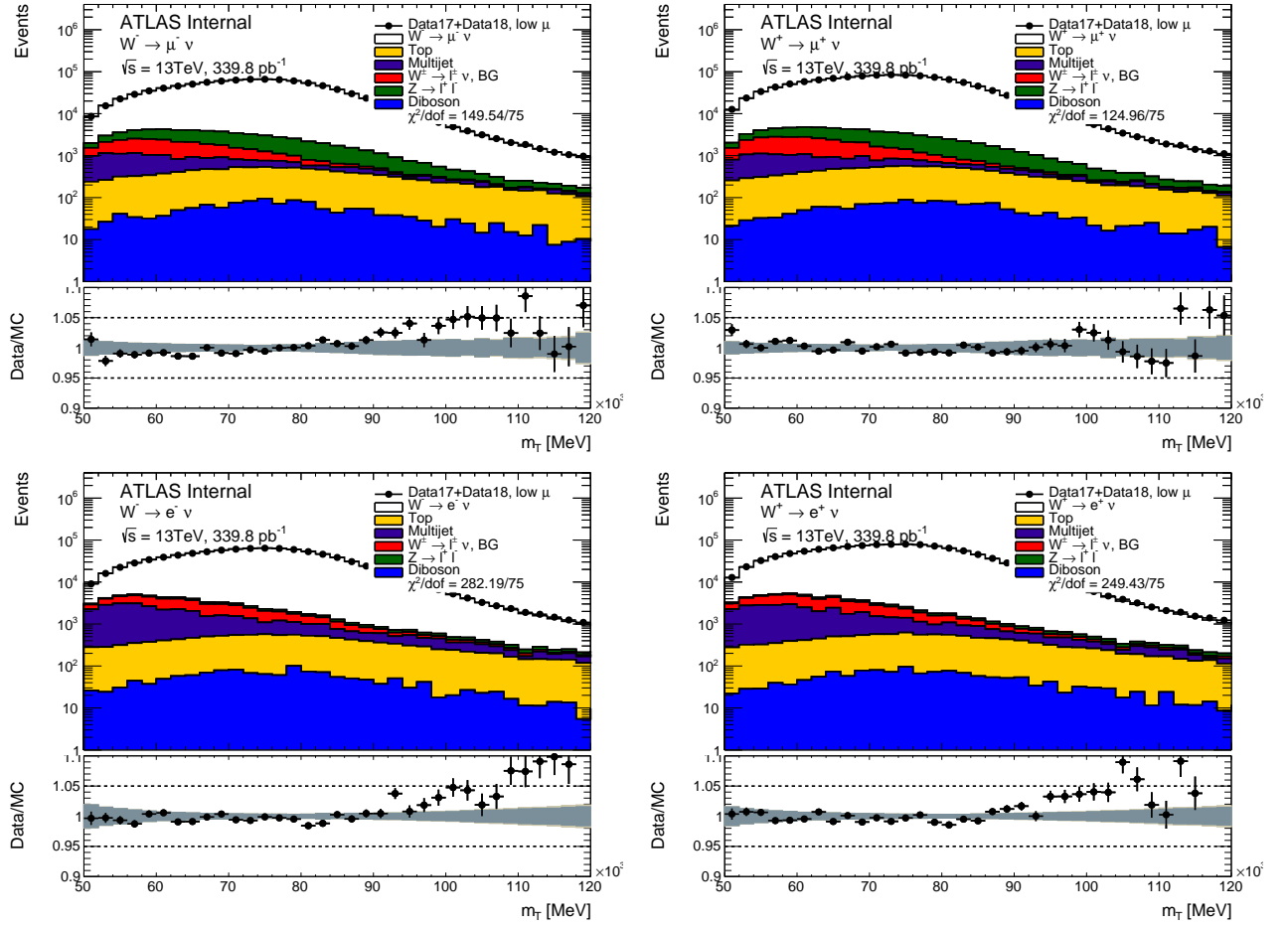
**Figure 8.14:** Transverse mass distribution of the W boson in the muon and electron channel for the $\sqrt{s}$ = 13 TeV dataset.
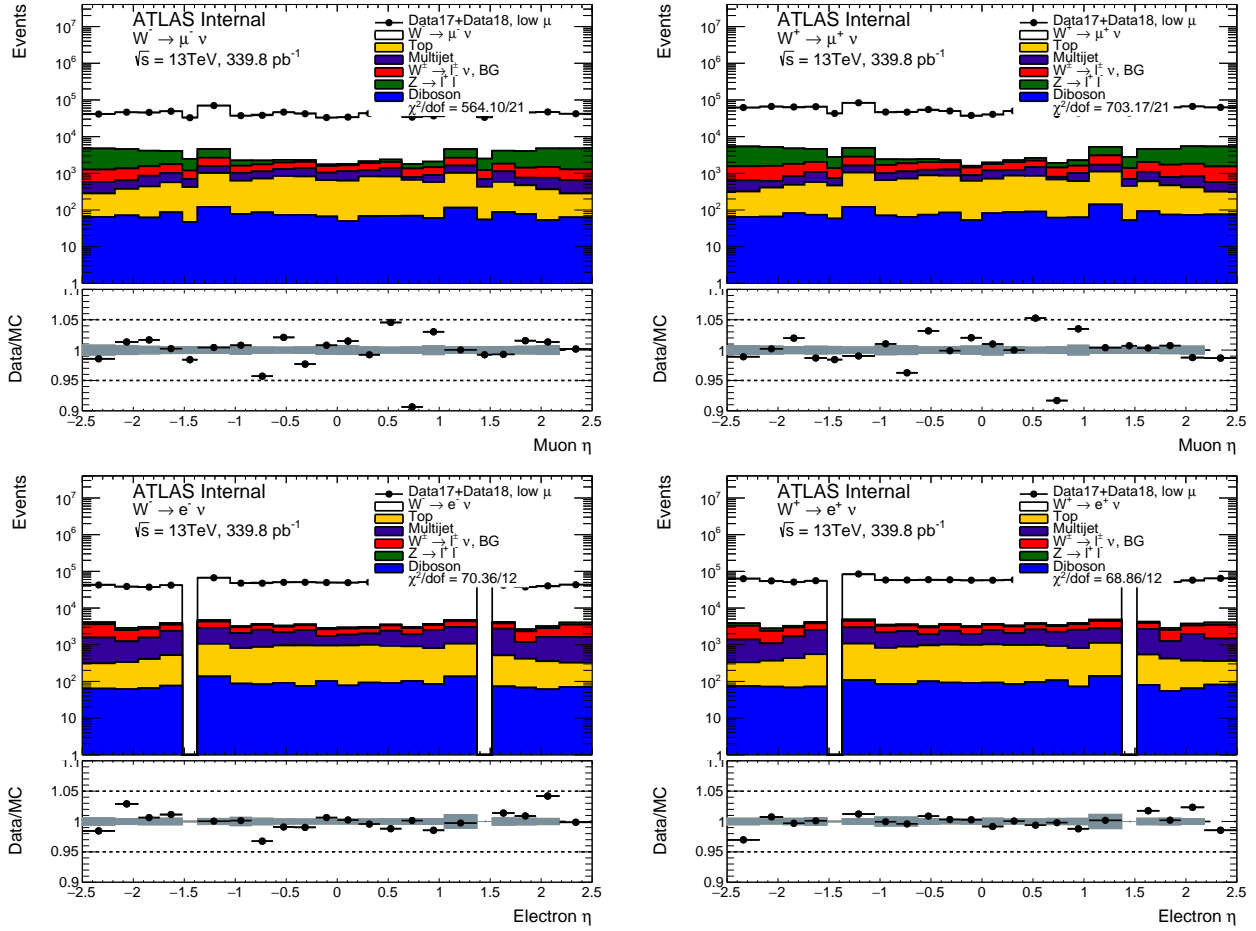
**Figure 8.15:** Lepton pseudorapidity distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.
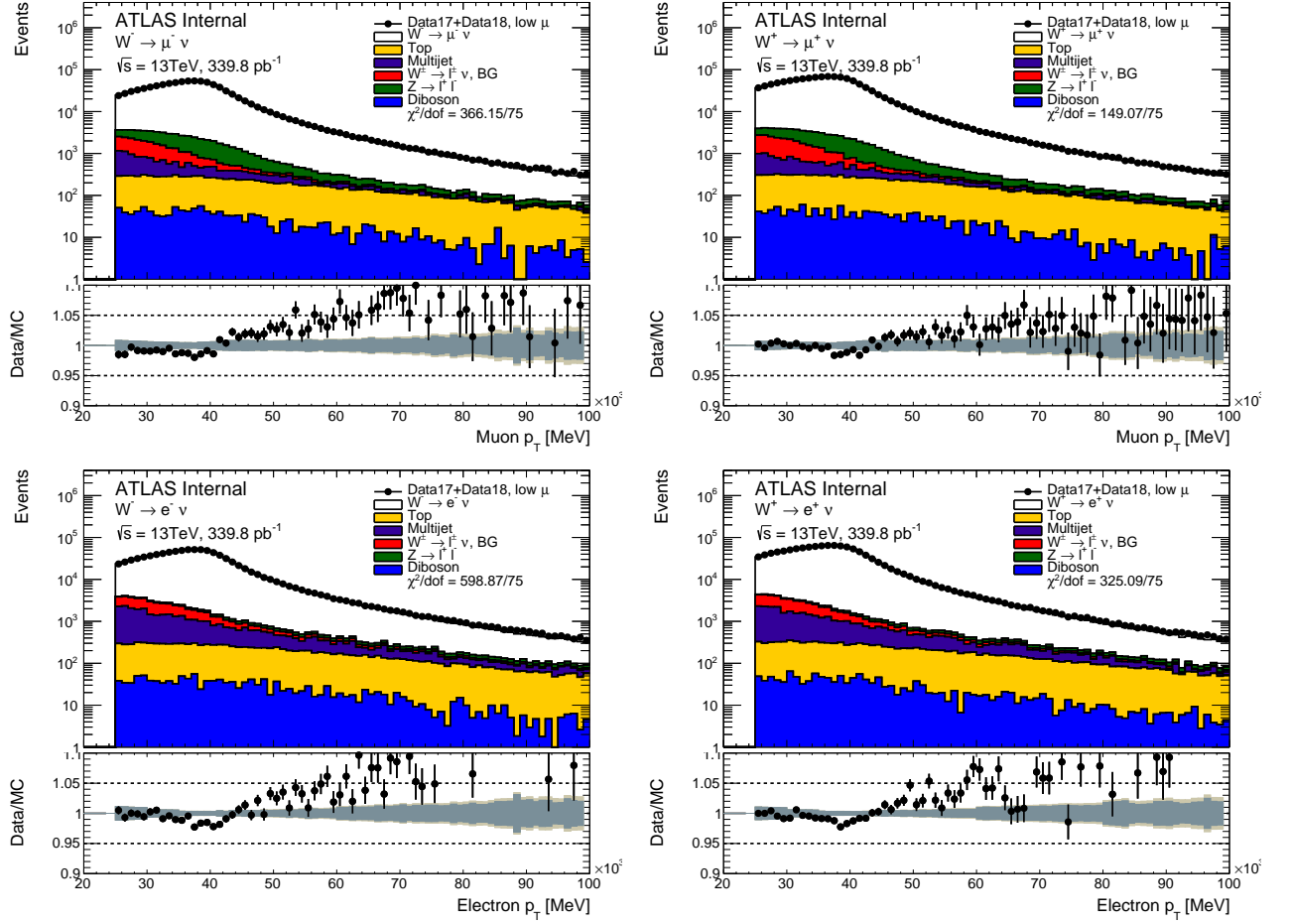
**Figure 8.16:** Lepton transverse momentum distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.
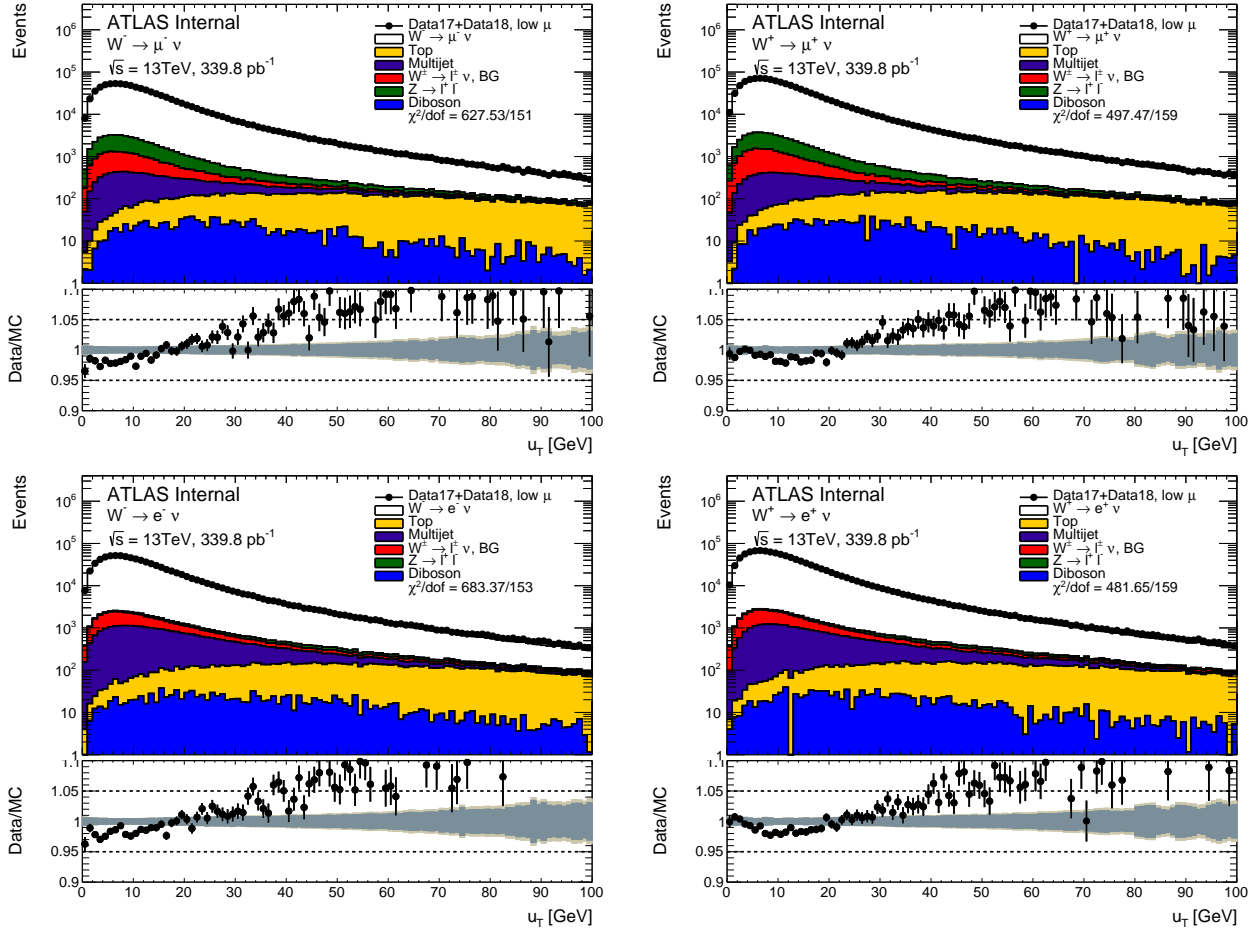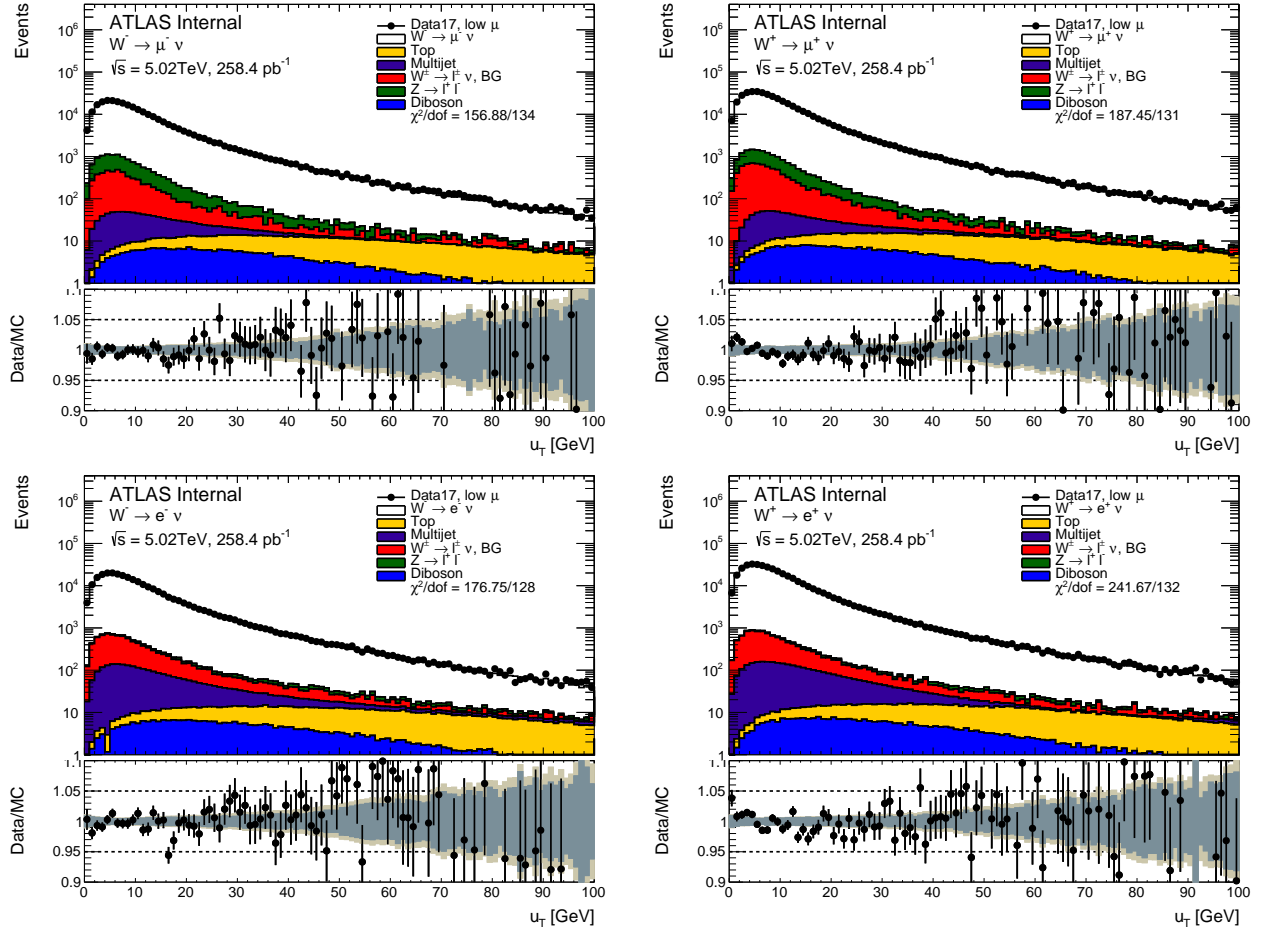
**Figure 8.17:** Hadronic recoil distribution in the muon and electron channel for the $\sqrt{s} = 13$ TeV dataset.

### 8.4.3 $\sqrt{s} = 5$ TeV dataset control plots

Control plots for the 5 TeV low-$\mu$ dataset are provided here after applying all corrections described in Section 7, and after applying the selection described above in this section. In each figure, the right(left)-hand column shows distributions for the $W^+$ ($W^-$) process. The top (bottom) row shows the muon (electron) decay channel. In the ratio panels, the grey band is the total systematic uncertainty, whilst the brown band adds the MC statistical uncertainty in quadrature on top of it. In regions of the distributions insensitive to the modelling of $p_T^W$ there is generally good agreement between data and predictions. The bulk of the $m_T$ distribution is a typical example of distribution that is mostly insensitive to the modeling of $p_T^W$. Compared to the 13 TeV situation, the $u_T$ distribution seems to indicate that the baseline simulation models $p_T^W$ more satisfactorily.
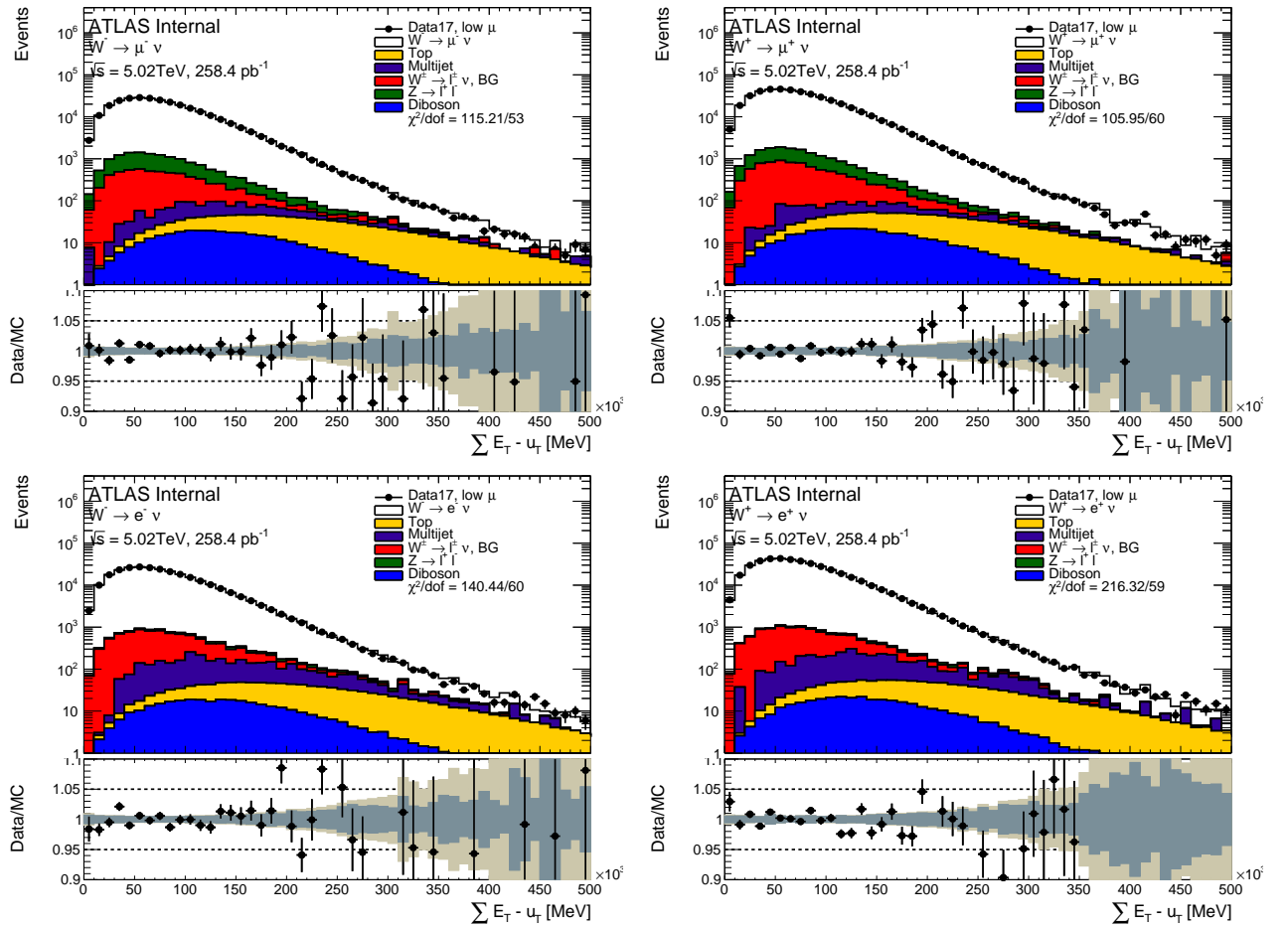


**Figure 8.18:** $\Sigma \bar{E}_T$ distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.

**Figure 8.19:** $\Sigma E_T$ distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.

**Figure 8.20:** $\vec{E}_T^{miss}$ distribution in the muon and electron channel for the $\sqrt{s}$ = 5 TeV dataset.

**Figure 8.21:** Transverse mass distribution of the W boson in the muon and electron channel for the $\sqrt{s}$ = 5 TeV dataset.

**Figure 8.22:** Lepton pseudorapidity distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.

**Figure 8.23:** Lepton transverse momentum distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.

**Figure 8.24:** Hadronic recoil distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV dataset.

# W boson pT spectrum $\boldsymbol{9}$

The final result would require to unfold the hadronic recoil distribution. The unfolding procedure, corresponding uncertainties and the final result are discussed in this chapter. A comparison of the measured underlying spectrum with a number of Monte-Carlo generators is provided.

## 9.1  Unfolding

The measured W $p_T$ spectrum is subject to various detector effects (finite resolution and acceptance, reconstruction efficiency, etc.) that distort the true underlying spectrum. Mathematically, the unfolding problem is an integral equation of the following form:

$$\int K(x,y) \cdot f(x)dx = g(y), \tag{9.1}$$

where we seek the function $f(x)$ assuming that $g(y)$ and the kernel $K(x,y)$ are known [166]. The function $g(y)$ is convoluted (or folded) with the kernel hence the name of the problem. In experimental physics it is more common to use binned distributions instead of continuous functions:

$$\sum_i \mathbf{R}_{ij} \cdot \mathbf{T}_i + \mathbf{B}_j = \mathbf{D}_j, \tag{9.2}$$

with $\mathbf{D}$, $\mathbf{T}$ and $\mathbf{B}$ being vectors that represent detector-level (measured), truth and background distributions respectively; each vector has $n$ components that represent the bins in the corresponding distribution. The response matrix $\mathbf{R}$ represents bin-to-bin migrations caused by the detector effects.

The response matrix is usually obtained through MC simulation, along with the corrections for the fiducial volume acceptance and selection efficiency. Each event is simulated on the truth and reconstructed levels, this means that element $R_{il}$ of the migration matrix contains events that pass both reconstruction and truth cuts (R&T) and would go to bin $i$ of the truth distribution and to bin $j$ of the reconstructed distribution. Detector acceptance for bin $i$ is defined as the ratio $A_i = \sum_j Rij/T_i$. Similarly the reconstruction efficiency for bin $j$ is defined as the ratio $\epsilon_j = \sum_i Rij/D_j$. The underlying distribution estimate is presented in the following way:

$$\sum_i \mathbf{V}_{ij} \cdot (\mathbf{D}_i - \mathbf{B}_i) = \mathbf{U}_j, \tag{9.3}$$

where $U$ vector provides the underlying distribution estimate and $V$ is the unfolding transformation matrix.

There exist a diverse variety of methods to obtain the unfolding transformation. In the current analysis a Bayesian iterative method is used [167], [168]. The method allows to obtain the unfolding transformation provided that the response matrix, acceptance and efficiency corrections are known and the number of iterations is given. The number of iterations as well as the unfolded distribution binning are adjusted in order to minimize the unfolding bias and keep the uncertainty below the designated level.
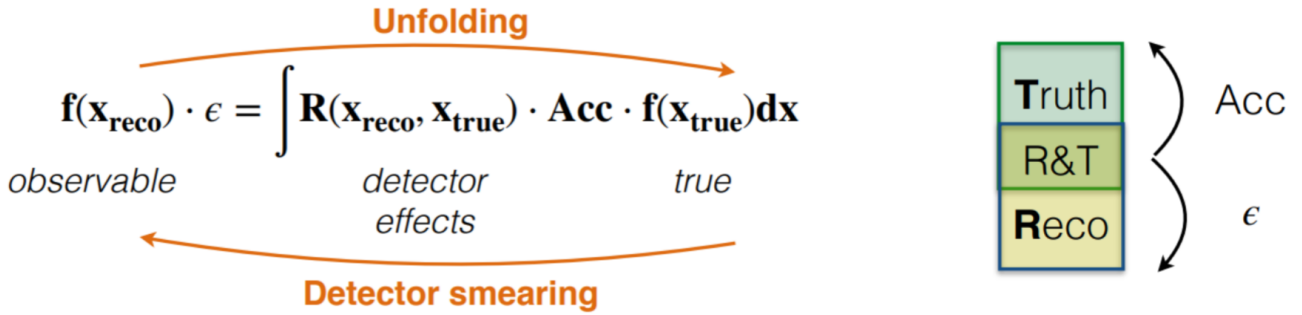


**Figure 9.1:** Schematic description of the unfolding procedure.

The MC simulation of the detector response models the efficiency and fake rate, as well as the correspondence between the truth and the measured W transverse momentum. Here efficiency stands for the fraction of events that pass the truth cuts, but do not pass the reconstructed level cuts. Fake rate stands for the events that pass the reconstructed level cuts, but do not pass the truth cuts. The examples of such distributions are provided on Fig. 9.2.

## 9.2   Propagation of uncertainties

The detector-level uncertainties breakdown for the $p_T$ distribution are presented here. Uncertainties breakdown for the rest of the observables are listed in Appendix A. These uncertainties now have to be propagated to the unfolded level.

### 9.2.1   Statistical uncertainty propagation using Bootstrap method

Bootstrap is a computer-based method of dataset parameters estimate and propagation using the analysis distribution resampling. In particular bootstraping is used for the propagation of statistical uncertainties.

Both data and MC-simulated datasets have limited number of events, hence the statistical uncertainties due to fluctuations. In order to estimate the statistical uncertainty a number of pseudo-data
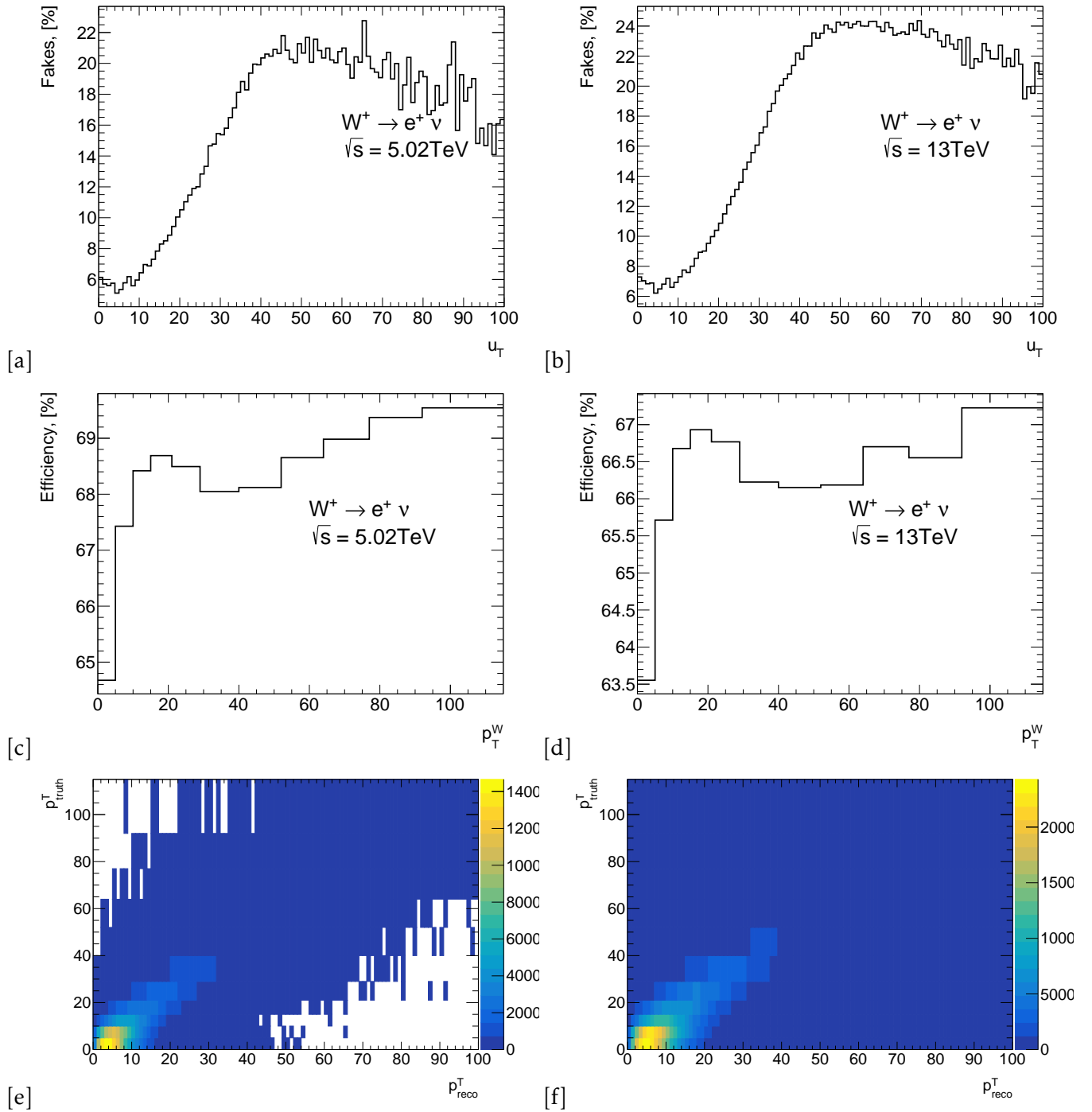
**Figure 9.2:** Distributions for the $W \to e\nu$ channel at 5 TeV (left) and 13 TeV (right). Fake rates are presented in subplots [a] and [b], efficiencies - in [c] and [d]. The two bottom plots contain the response matrices.
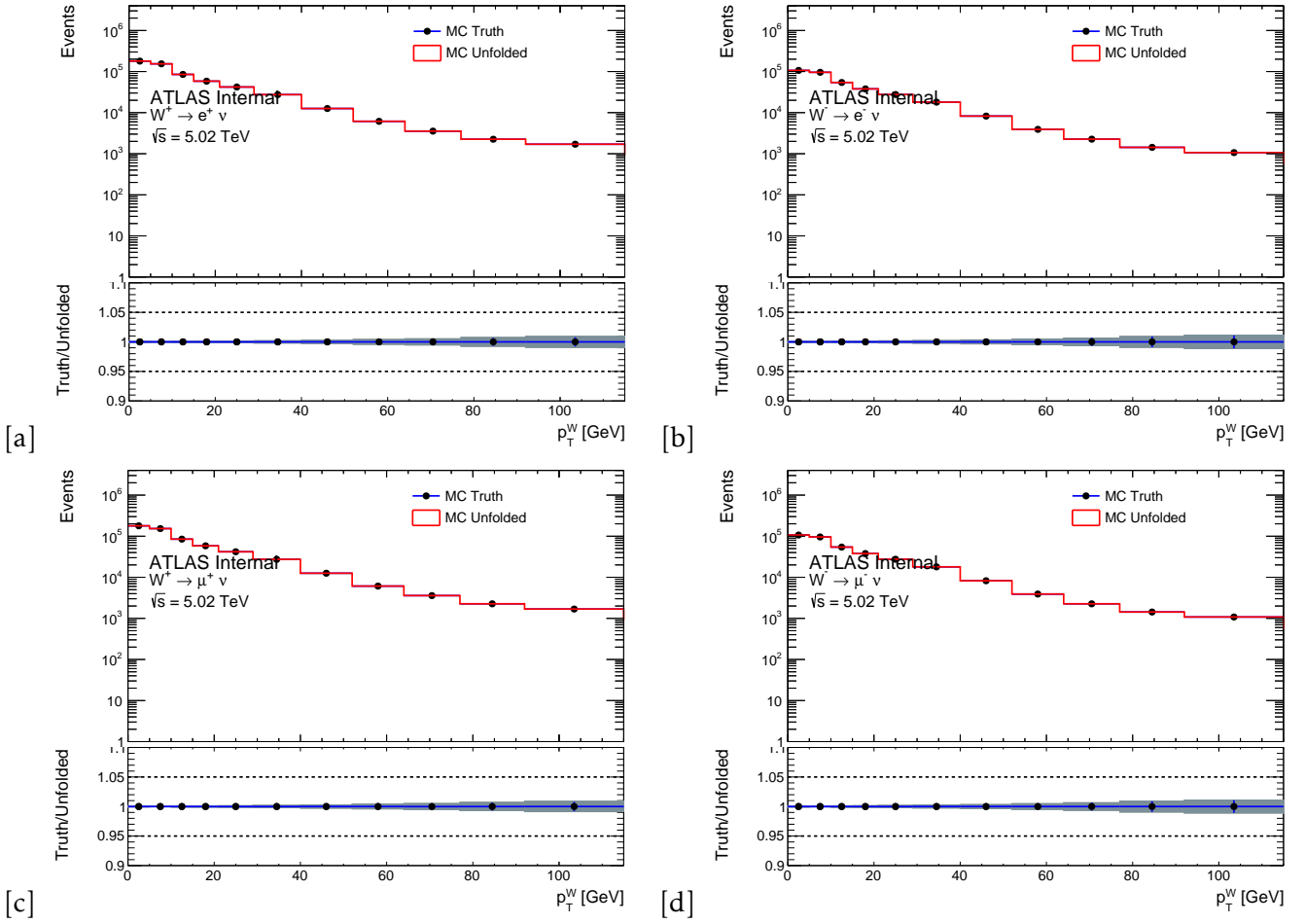
**Figure 9.3:** The technical closure of the unfolding procedure is demonstrated at 5 TeV for all the channels.

sets is generated for both data and MC where each event is assigned a random weight $w$:

$$w = \mathcal{P}(n, 1), \tag{9.4}$$

where $n$ is a random number generated with Poisson distribution with mean $\lambda = 1$, value $\mathcal{P}(n, 1)$ is a Poissonian probability of observing n events while expecting an average of 1 event.

The boostrapping defined in this way allows to take into account the correlated effect of statistical fluctuations across all observables and distributions in the analysis. For the determination of statistical uncertainty of the unfolded spectrum 400 bootstrap samples were generated. In both data and MC cases the statistical uncertainty is estimated by composing the covariant matrix $C_{kl}^{stat}$:

$$C_{kl}^{stat} = \frac{1}{N_{bs}-1} \sum_{\alpha=1}^{N_{bs}} (U_k^\alpha - \langle U_k \rangle)(U_l^\alpha - \langle U_l \rangle), \tag{9.5}$$

where $N_{bs}$ is the number of the Bootstrap toys used, vector $U$ stands for the varied underlying distribution, $\langle U_k \rangle$ is the average underlying distribution. However, the variation is performed in a different
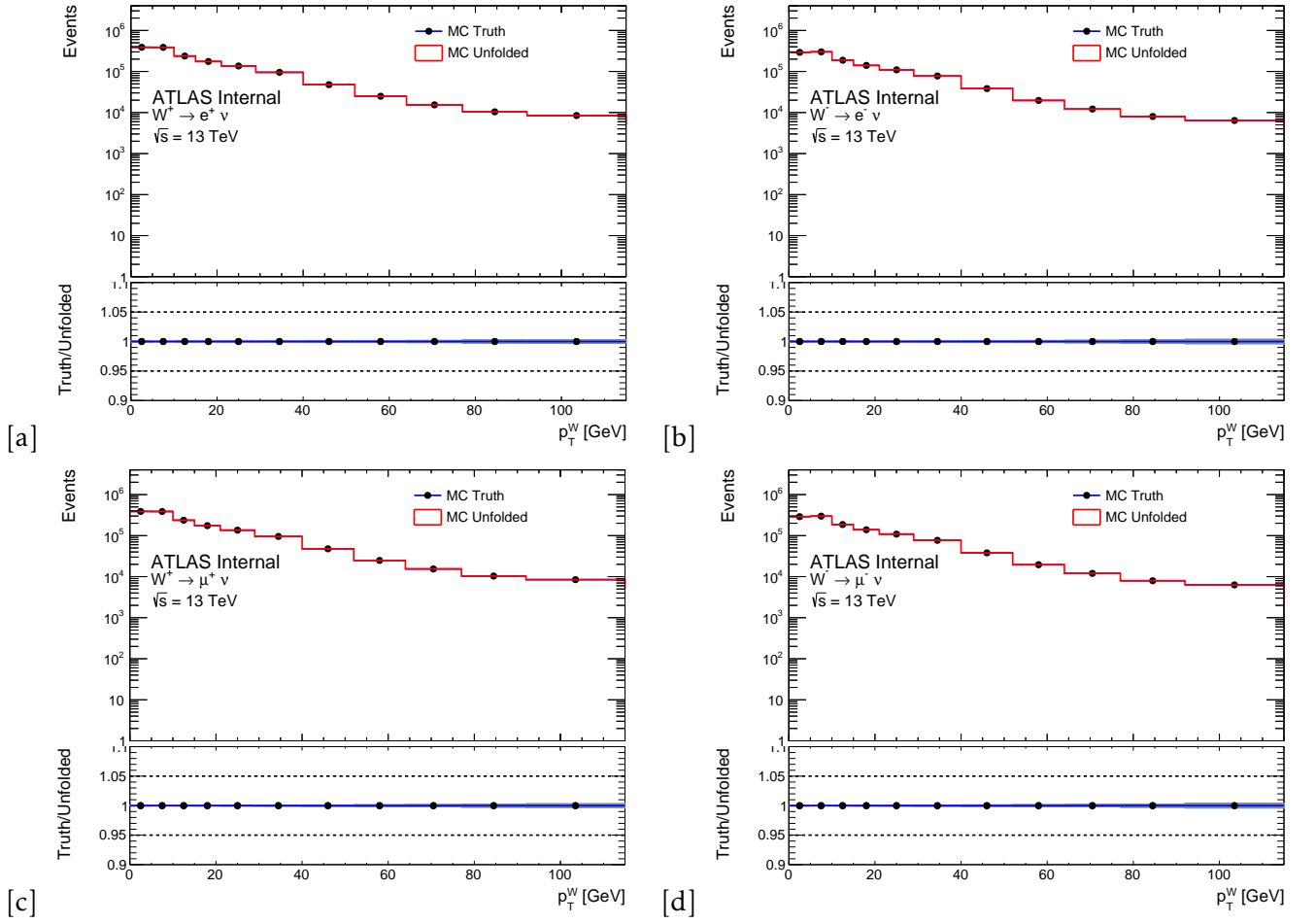
**Figure 9.4:** The technical closure of the unfolding procedure is demonstrated at 13 TeV for all the channels.

way for Data and MC:

$$U_j^{\alpha,(MC)} = V_{ij}^\alpha \sum_i (D_i - B_i),$$

$$U_j^{\alpha,(Data)} = V_{ij} \sum_i (D_i^\alpha - B_i).$$

In the MC case it is the response matrix $V^\alpha$ to be varied ($\alpha$ index corresponds to the variation number), whereas in Data the toys are obtained by varying the measured distribution $D_i^\alpha$. The statistical uncertainty for both cases is defined as:

$$\delta U_k = \sqrt{C_{kk}^{stat}}. \tag{9.6}$$
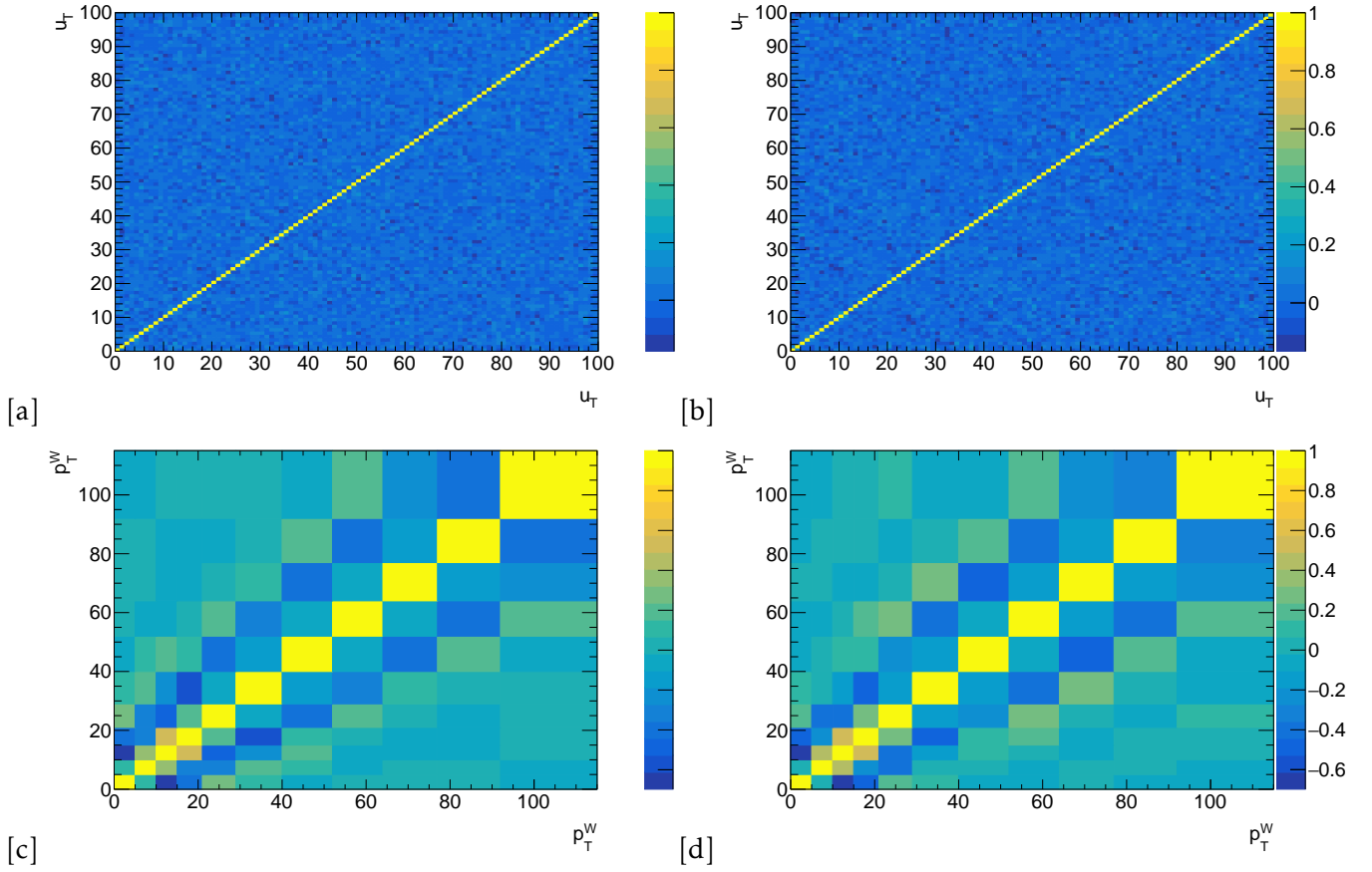
**Figure 9.5:** Correlation matrices for the $W \rightarrow e\nu$ channel at 5 TeV for MC (left) and data (right). The upper plots contain correlation matrices at the reconstructed level, bottom plots show the unfolded level.

## 9.2.2  Systematic uncertainty propagation

Systematic uncertainties are broken down into a number of uncorrelated uncertainty sources, which include signal and background modelling uncertainties, calibration and efficiency uncertainties, physics modelling uncertainties. The systematic variations used for uncertainty estimate on the detector level are propagated to the level of underlying distribution in two different ways. For the background uncertainties:

$$U_j^a = V_{ij} \sum_i (D_i - B_i^a),$$

total background estimate $B_i^a$ is varied in luminosity and cross-section of every back-ground (index $a$ numbers the sources of uncertainty). For other sources of systematic uncertainty:

$$U_j^a = V_{ij}^a \sum_i (D_i - B_i),$$

response matrix variation is created. The corresponding covariance matrix is defined as:

$$C_{kj}^a = \delta U_k^a \delta U_l^a,$$
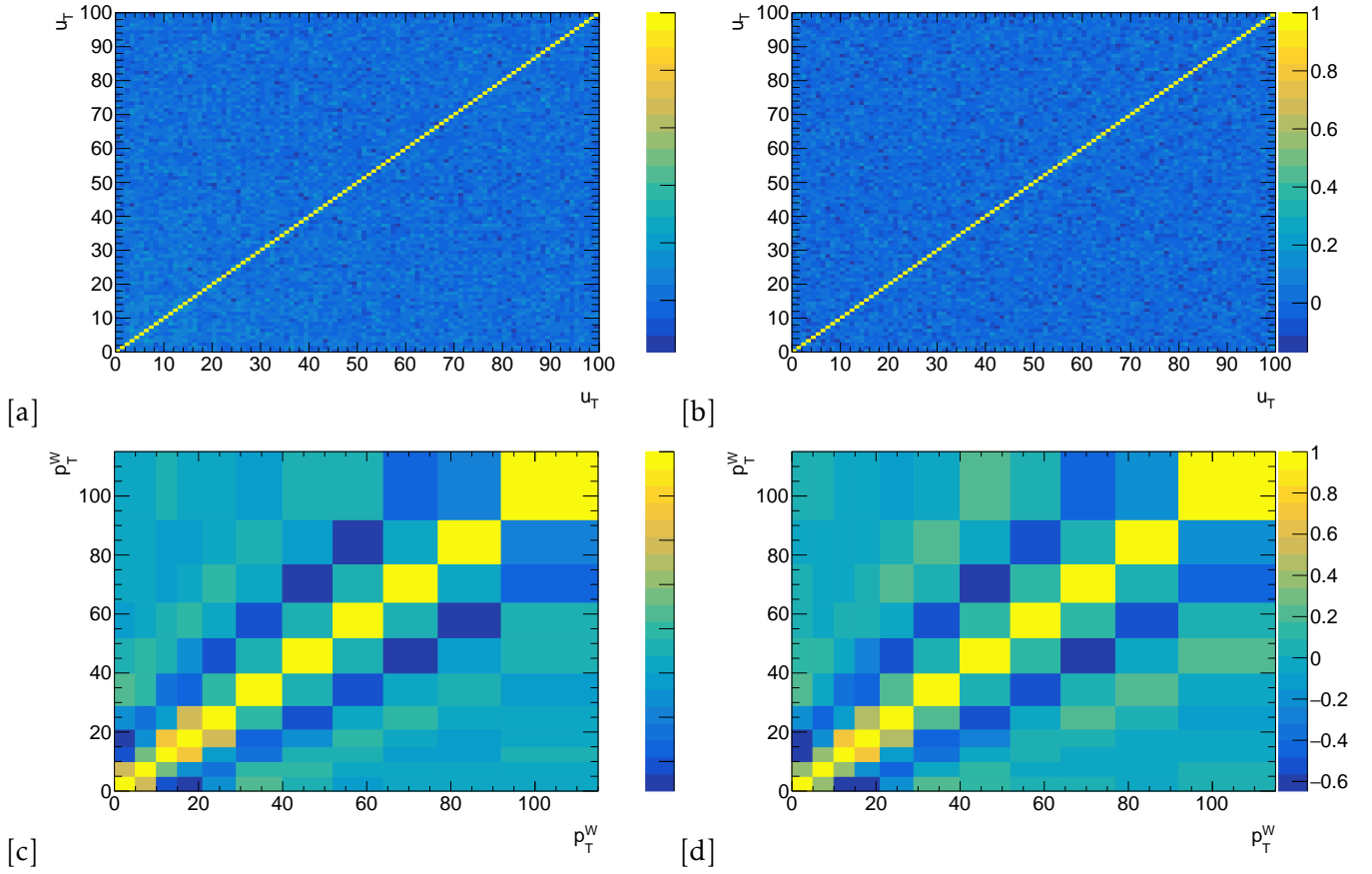
**Figure 9.6:** Correlation matrices for the $W \to e\nu$ channel at 13 TeV for MC (left) and data (right). The upper plots contain correlation matrices at the reconstructed level, bottom plots show the unfolded level.

where the deltas are $\delta U_k^a = U_k^a - U_k^{Nom}$. The total covariance matrix is calculated as a sum:

$$C_{kl}^{tot} = C_{kl}^{stat,Data} + C_{kl}^{stat,MC} + \sum_a C_{kl}^a. \tag{9.7}$$

### 9.2.3  Unfolded uncertainty breakdown

Figures 9.11, 9.13, contain the systematic uncertainties breakdown for electron and muon channels for the reconstructed level distributions for 5 and 13 TeV. Similarly figures 9.12, 9.14 contain unfolded-level uncertainties.

At the detector level the designated level of uncertainty of below 1% is preserved up to 25 GeV for 5 TeV datasets and up to 50 GeV for 13 TeV samples in every channel. An increased role of background uncertainty is observed at 13 TeV due to the significantly higher cross-sections of diboson and top-antitop backgrounds. The scale and hierarchy of uncertainties are preserved at the unfolded level.
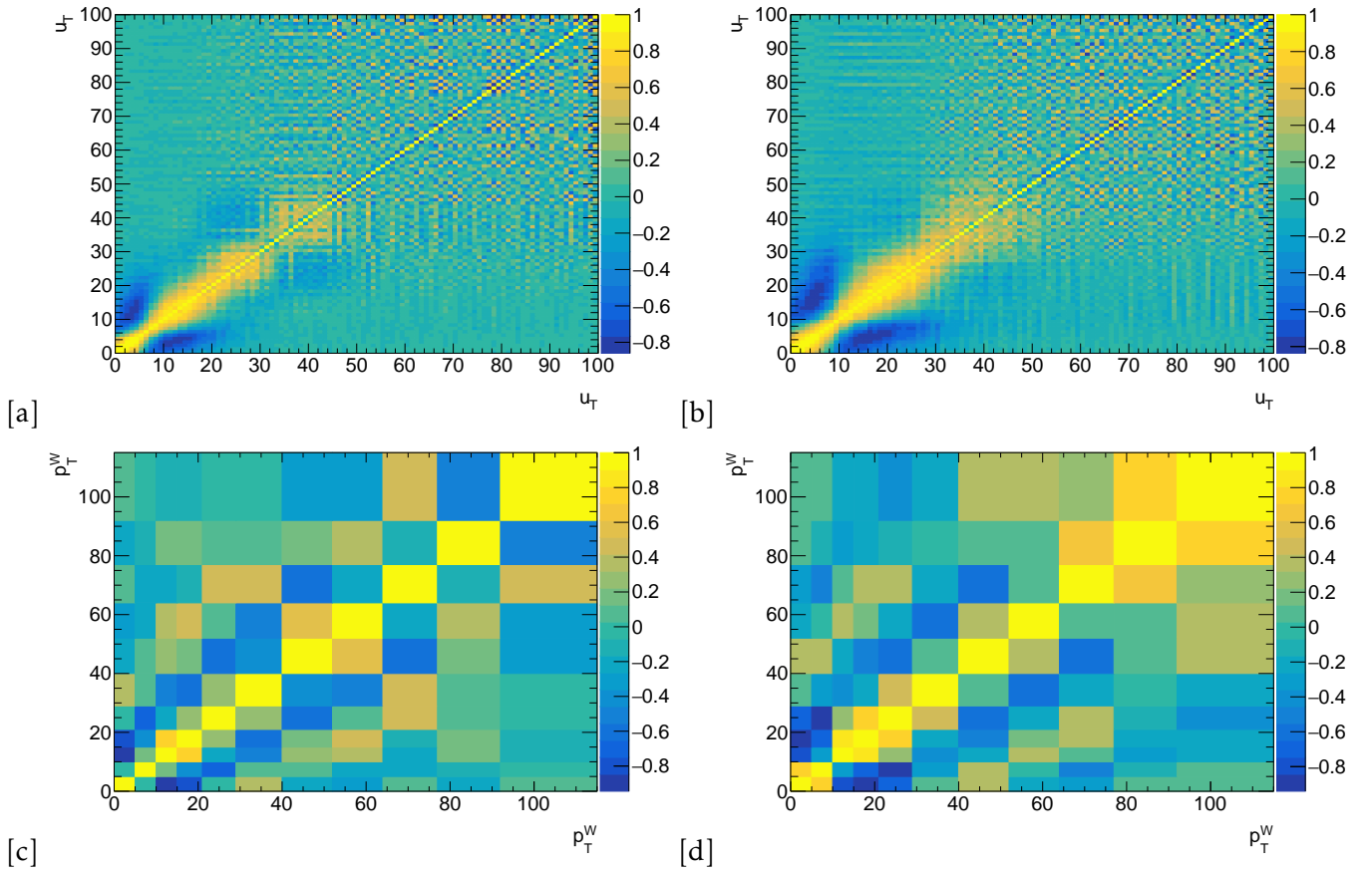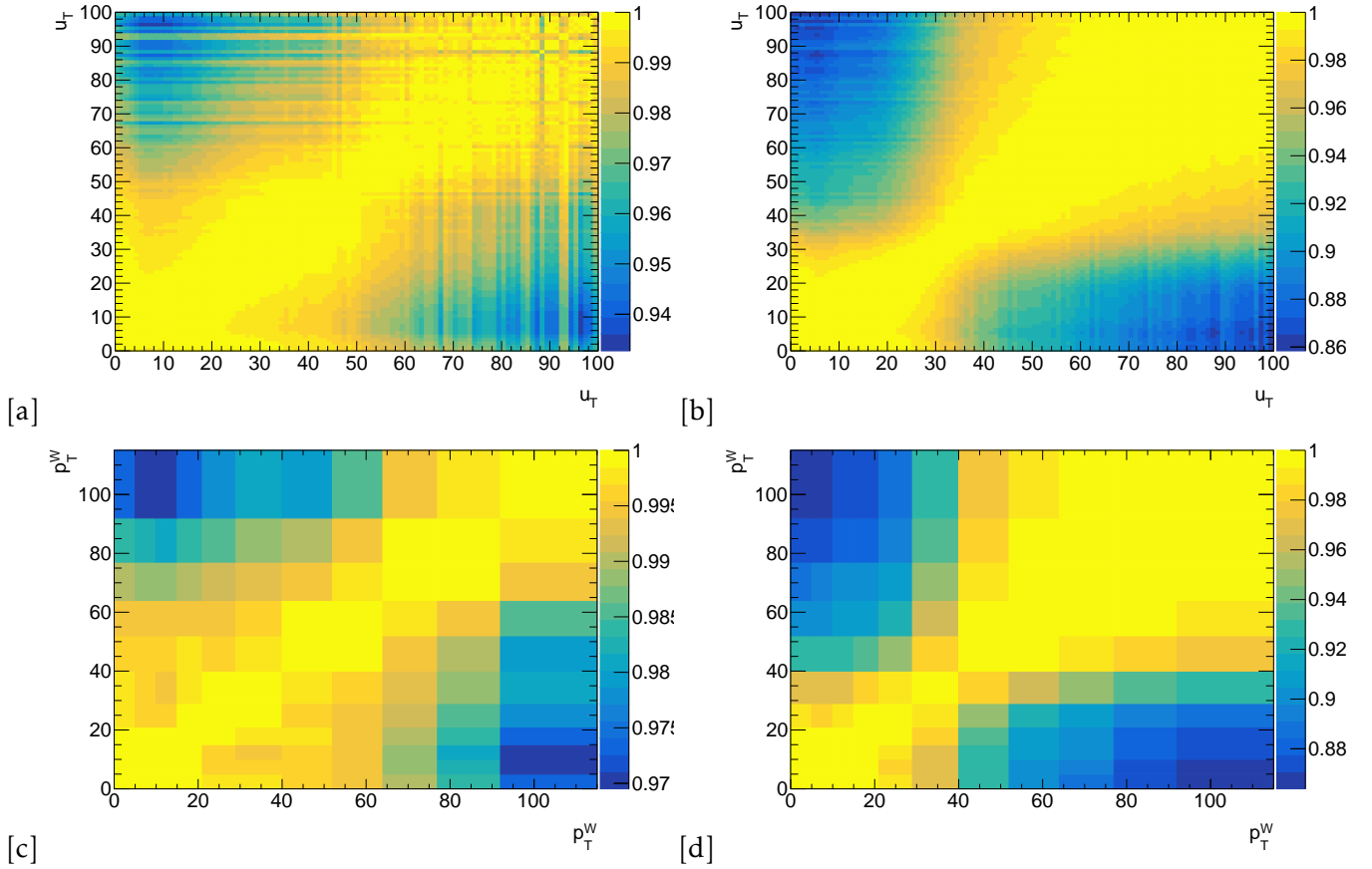
**Figure 9.7:** Correlation matrices for the hadronic recoil systematic variations for $W \rightarrow e\nu$ channel at 5 TeV (left) and 13 TeV (right). The upper plots contain correlation matrices at the reconstructed level, bottom plots show the unfolded level.

The uncertainties in 5 TeV are dominated by the statistical uncertainty and by the contribution from the hadronic recoil calibration, which is also of statistical nature. On the other hand, at 13 TeV background uncertainty also starts to play a significant role due to increased cross-sections of leading backgrounds.
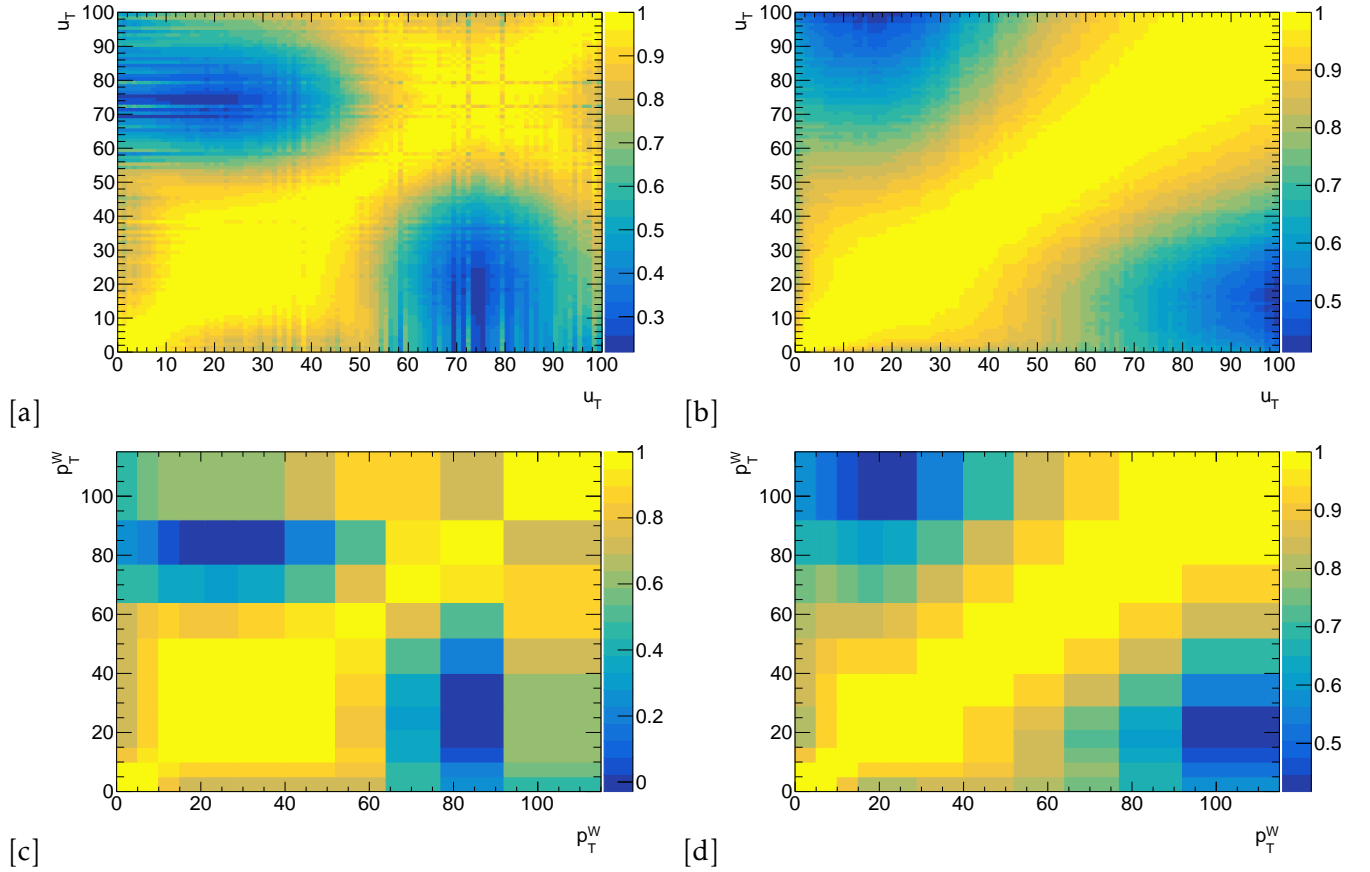
**Figure 9.8:** Correlation matrices for the background systematic variations for $W \to e\nu$ channel at 5 TeV (left) and 13 TeV (right). The upper plots contain correlation matrices at the reconstructed level, bottom plots show the unfolded level.

**Figure 9.9:** Correlation matrices for the scalefactors systematic variations for $W \rightarrow e\nu$ channel at 5 TeV (left) and 13 TeV (right). The upper plots contain correlation matrices at the reconstructed level, bottom plots show the unfolded level.
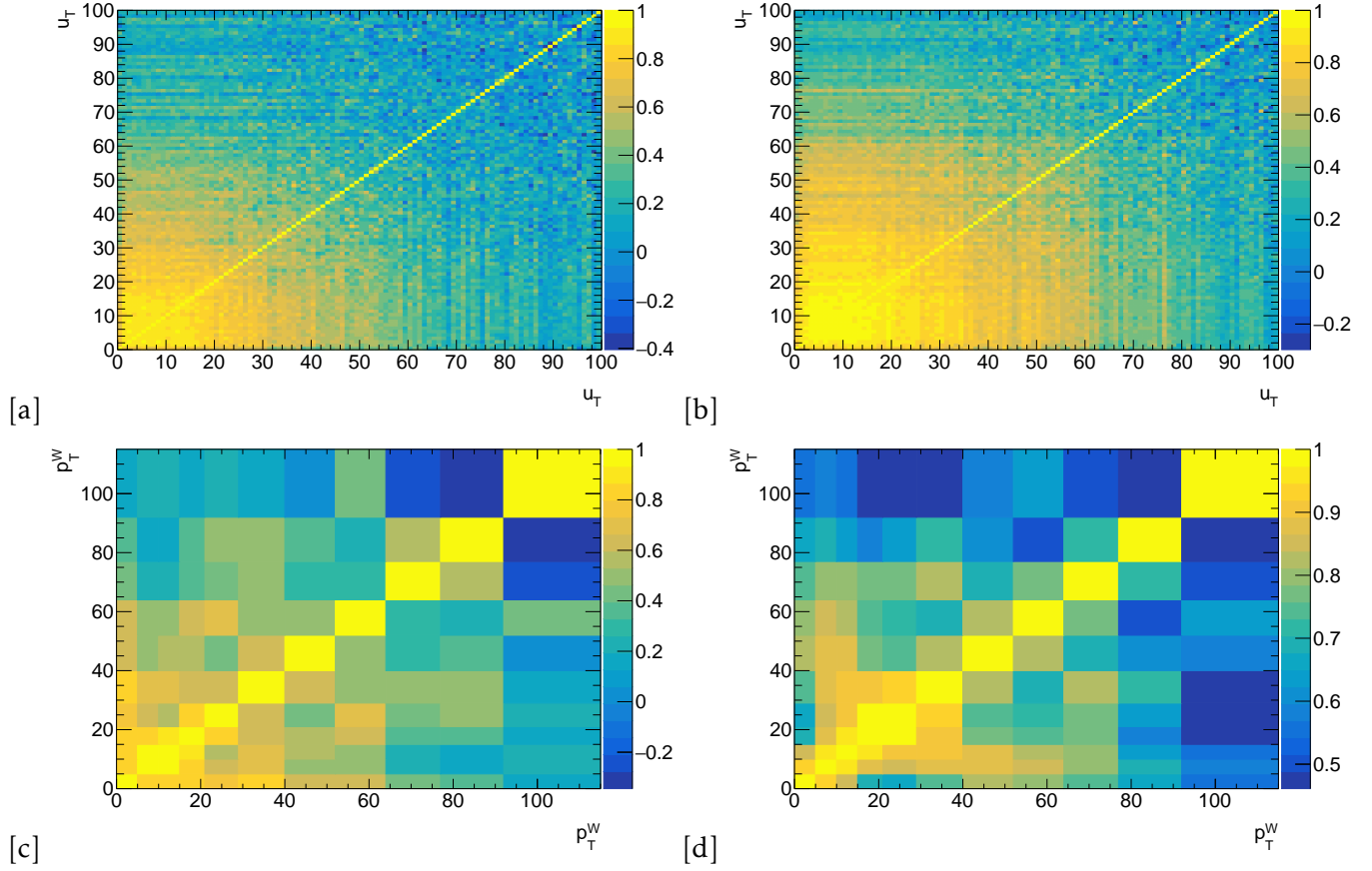
**Figure 9.10:** Correlation matrices for the electron calibration systematic variations for $W \rightarrow e\nu$ channel at 5 TeV (left) and 13 TeV (right). The upper plots contain correlation matrices at the reconstructed level, bottom plots show the unfolded level.
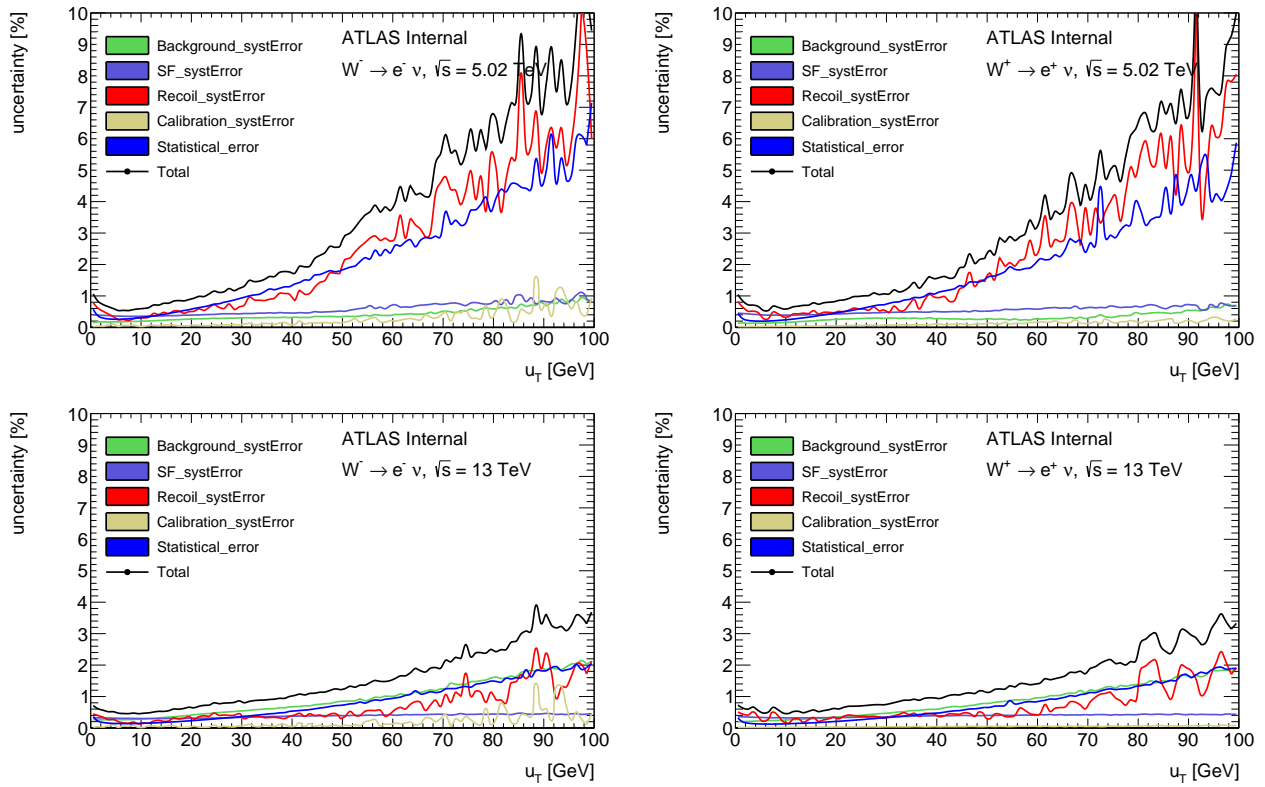
**Figure 9.11:** Breakdown of systematic uncerainties for 5 (a,b) and 13 (c,d) TeV in the electron channel at the reconstructed level.
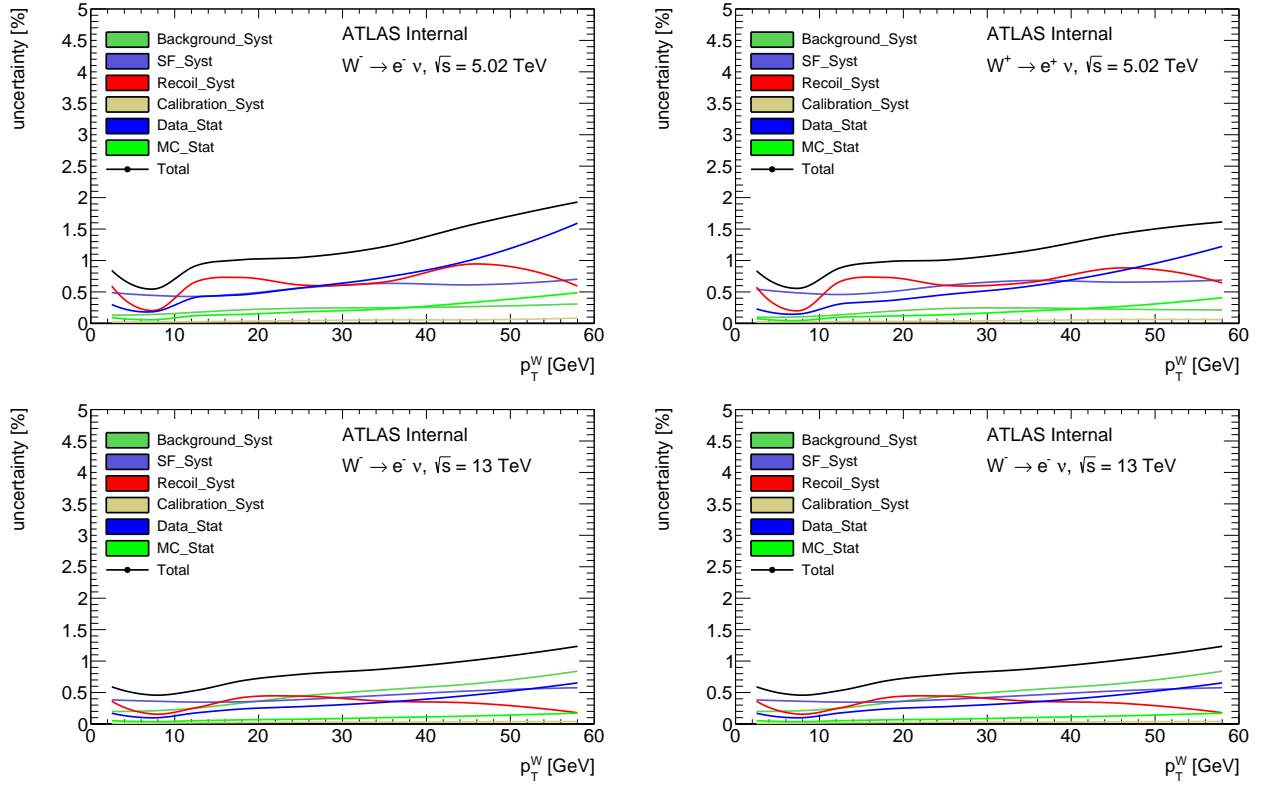
**Figure 9.12:** Breakdown of systematic uncertainties for 5 (a,b) and 13 TeV (c,d) in the electron channel at the unfolded level.
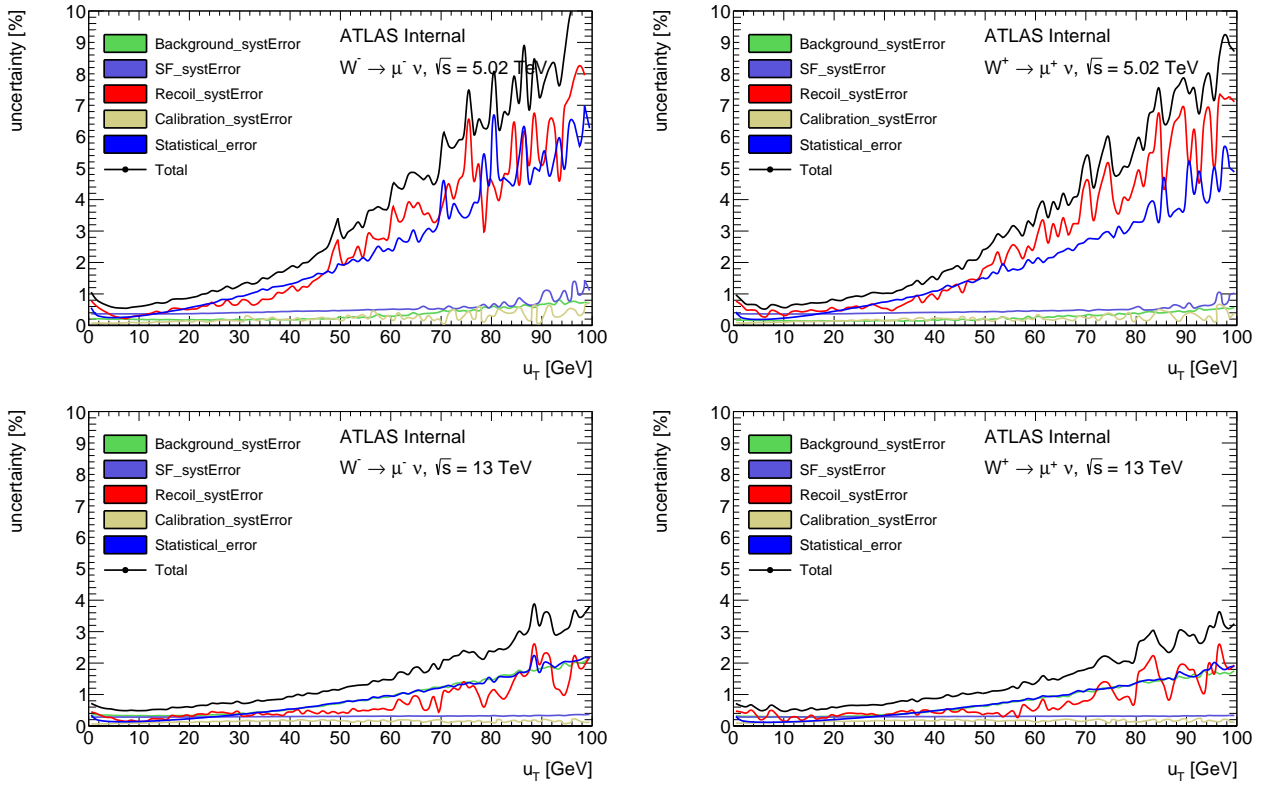
**Figure 9.13:** Breakdown of systematic uncertainties for 5 (a,b) and 13 (c,d) TeV in the muon channel at the reconstructed level.
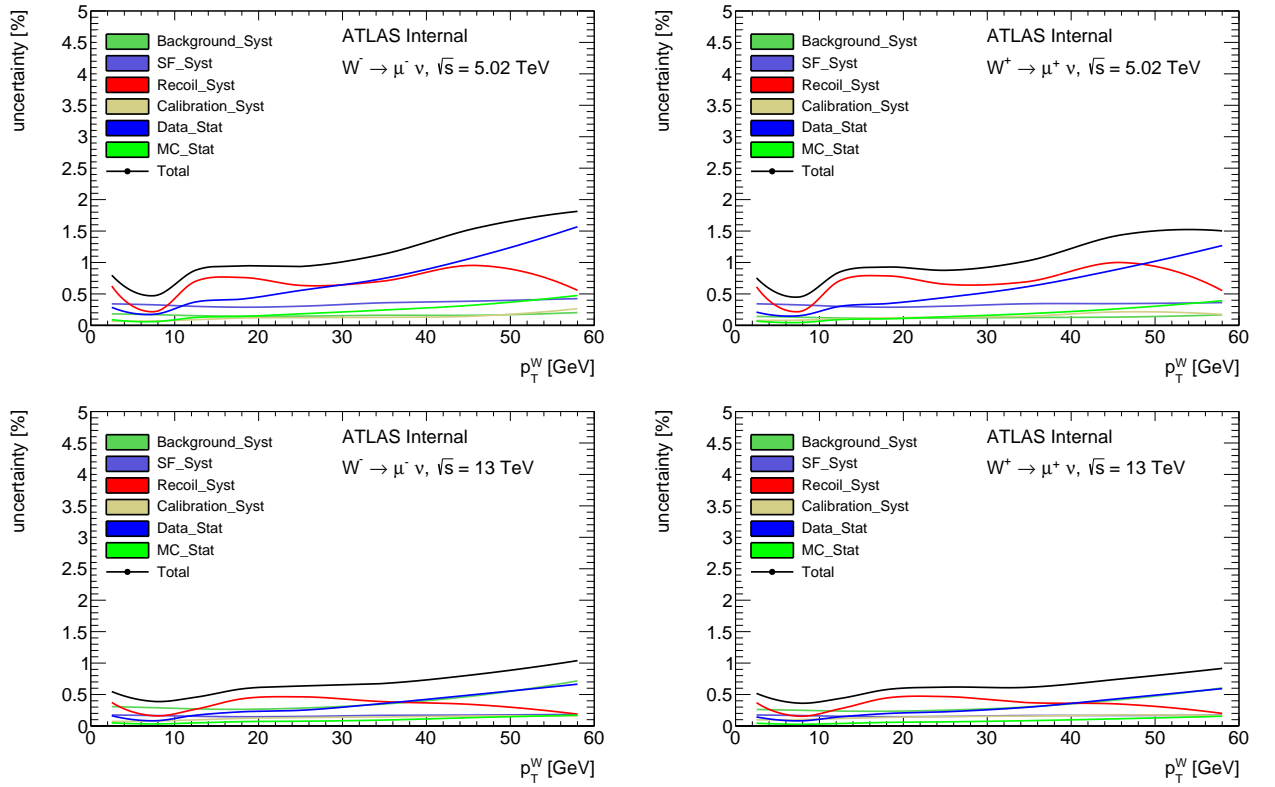
**Figure 9.14:** Breakdown of systematic uncertainties for 5 (a,b) and 13 TeV (c,d) in the muon channel at the unfolded level.
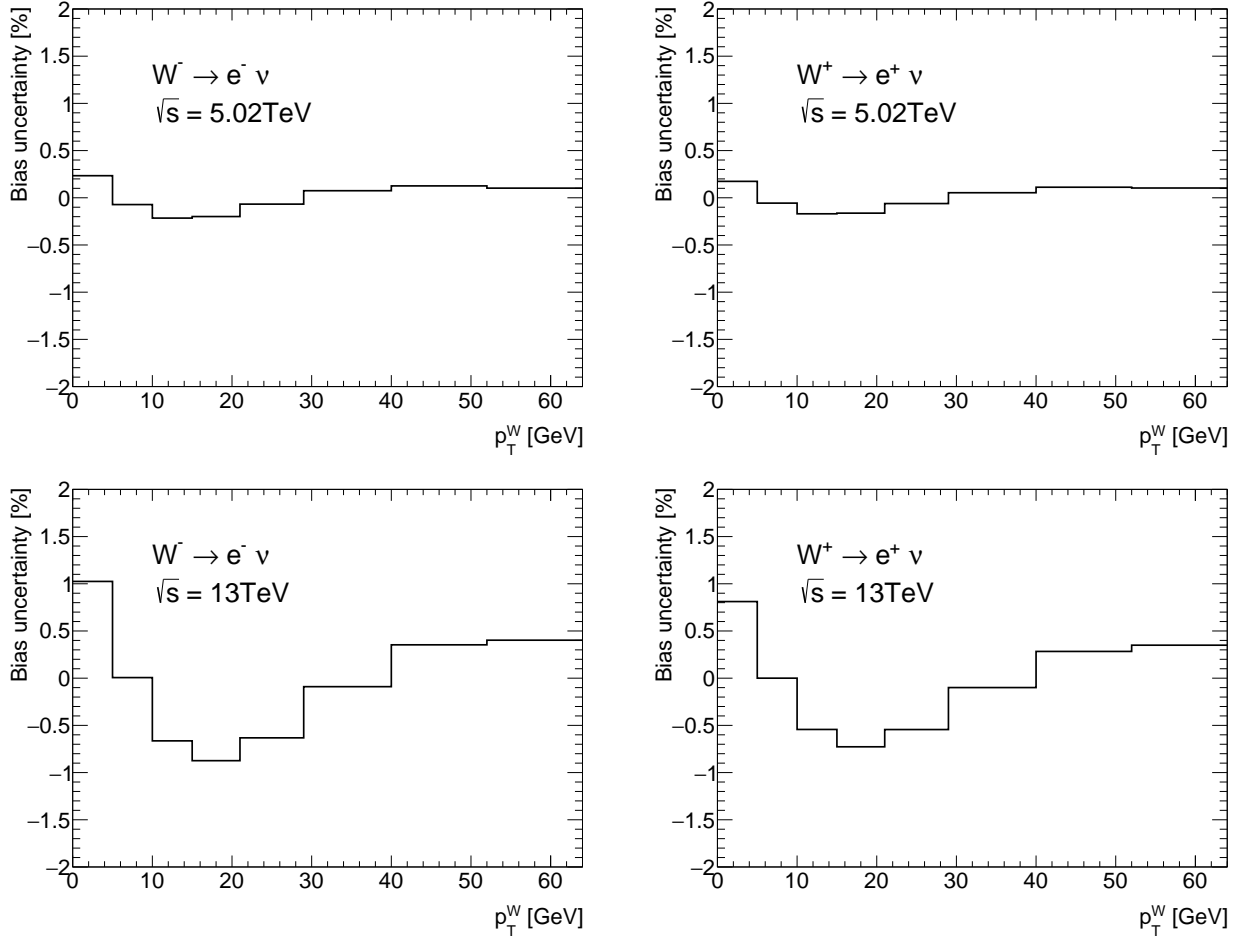
## 9.3 Unfolding bias

One of the uncertainties associated with unfolding usage is called unfolding bias and may arise because the procedure relies on the MC simulation of the distribution, which is used as a prior hypothesis for the Bayesian algorithm. Possible discrepancies between the modelled and true distribution lead to erroneous bin-to-bin migrations and can lead to distortions of the spectrum.

In order to estimate the bias induced by the unfolding procedure it is necessary to quantify how much the unfolded result is impacted by the assumed MC distribution. A set of samples with a different distribution at the truth level though compatible at the detector level is generated.

The truth distribution is reweighted until a good agreement between the data and MC is reached at the reconstruction level. The agreement is estimated in the kinematic region of $u_T < 100\text{GeV}$ using the $\chi^2$ criterion. The truth reweighting procedure is applied to MC samples with a different distribution: Pythia8, Sherpa and DYRES were used. Fig. demonstrates the initial difference in the distributions.

The results are presented on fig 9.15 for 5 GeV bins and 3 unfolding iterations. The obtained bias is close to the precision goal of the measurement ($\sim$ 1%) for the 5 TeVdataset. The 13 TeV dataset shows a larger bias, which can be explained by a larger discrepancy between data and Monte-Carlo. Worse resolution in 13 TeV suggests a neccesity to try a broader binning comparing to 5 GeV.

**Figure 9.15:** Unfolding bias on $p_T^W$ in the electron channel after 3 iterations, for $W^-$ (left) and $W^+$ (right), at 5 TeV (top) and 13 TeV (bottom) [169].
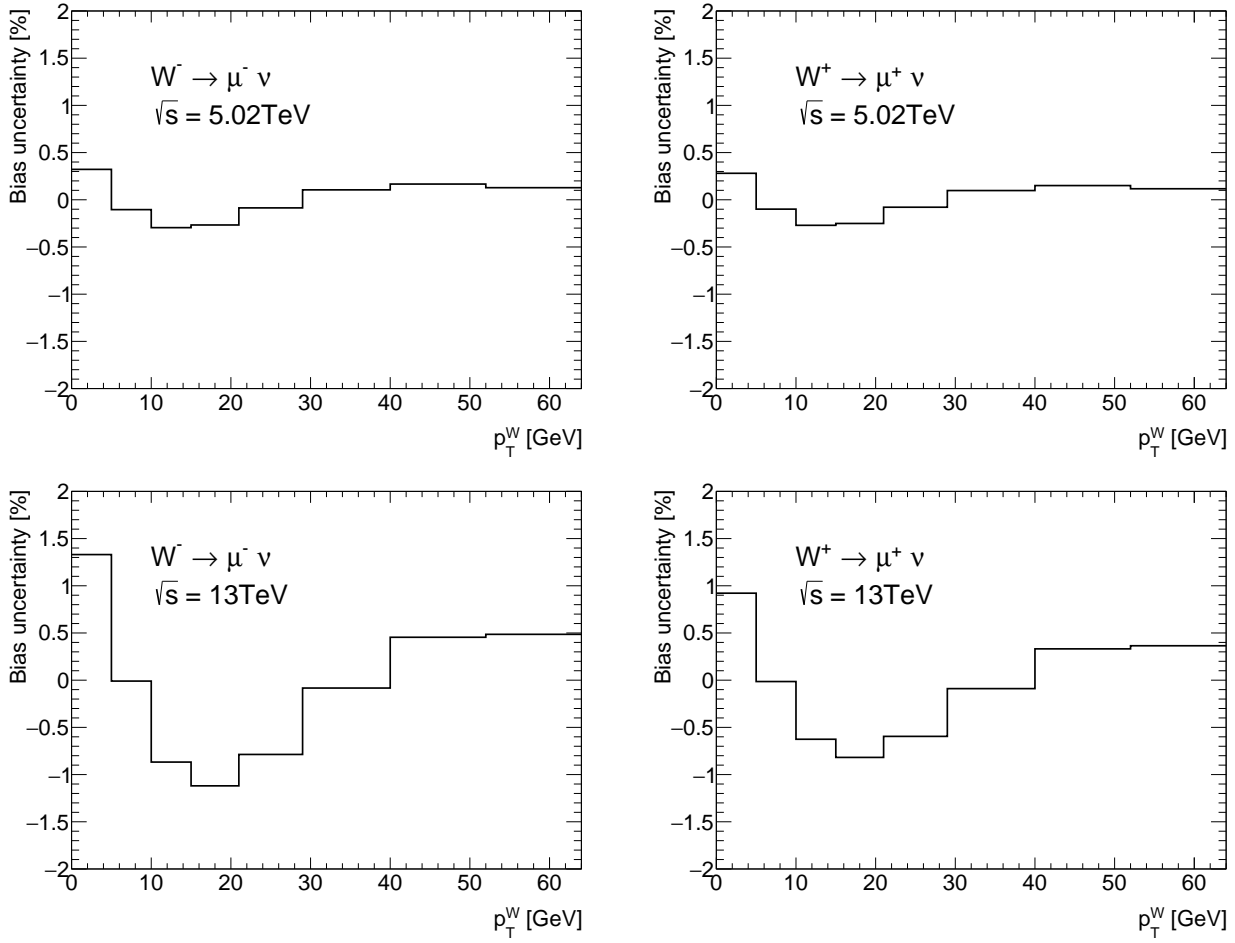
**Figure 9.16:** Unfolding bias on $p_T^W$ in the muon channel after 3 iterations, for $W^-$ (left) and $W^+$ (right), at 5 TeV (top) and 13 TeV (bottom) [169].

## 9.4   Results

The plots containing the unfolded spectrum are presented in Figs. 9.17 and 9.18. The 5 TeV plots demonstrate a decent degree of agreement between the data and Powheg MC simulation. For the 13 TeV the agreement is considerably worse, the data demonstrates higher cross-section in the harder part of the spectrum comparing to the simulation. The assumed reason for this is because the generator was tuned using 7 TeV data which scales fine to 5 TeV but the extrapolation becomes unreliable at 13 TeV.
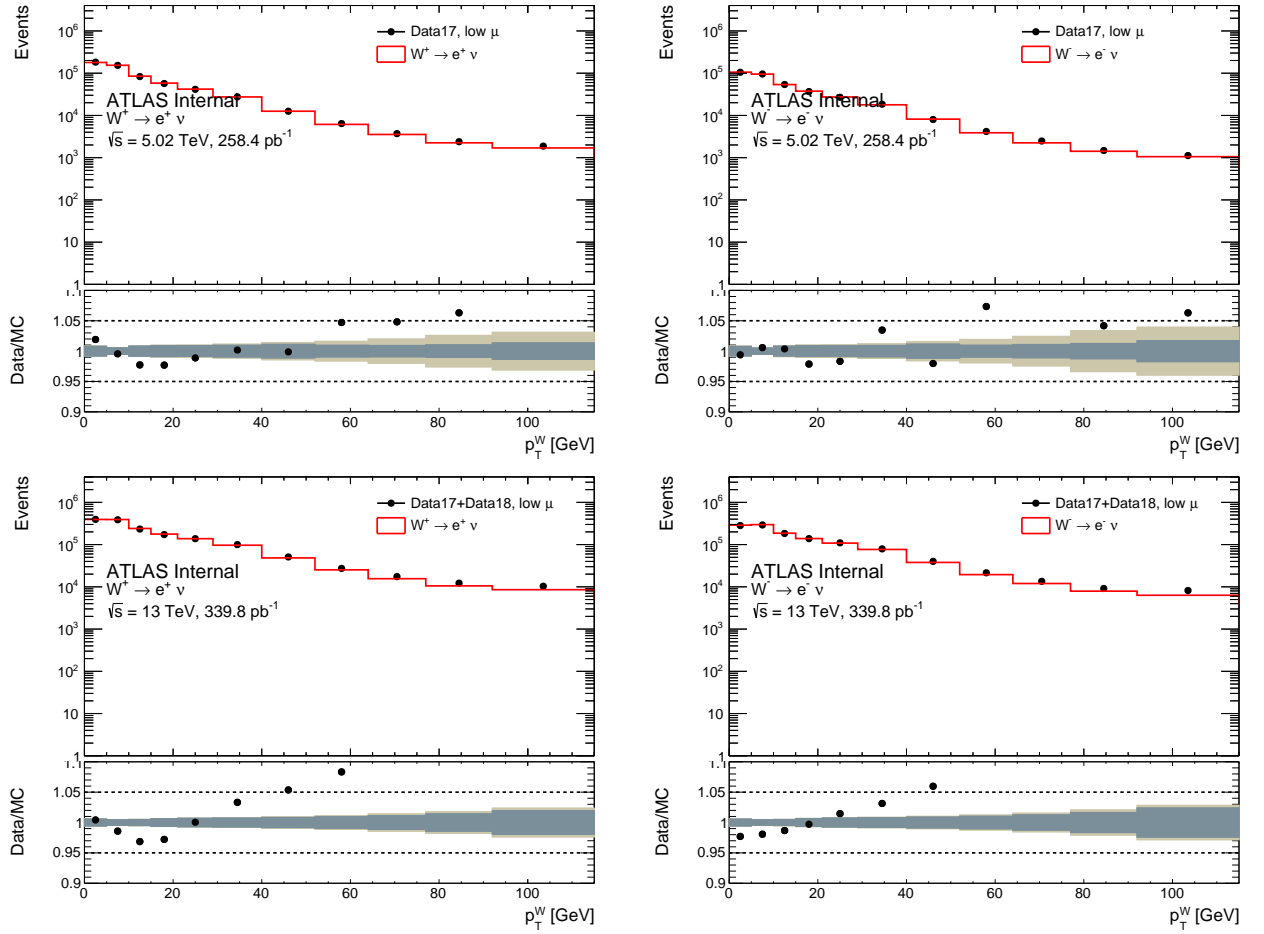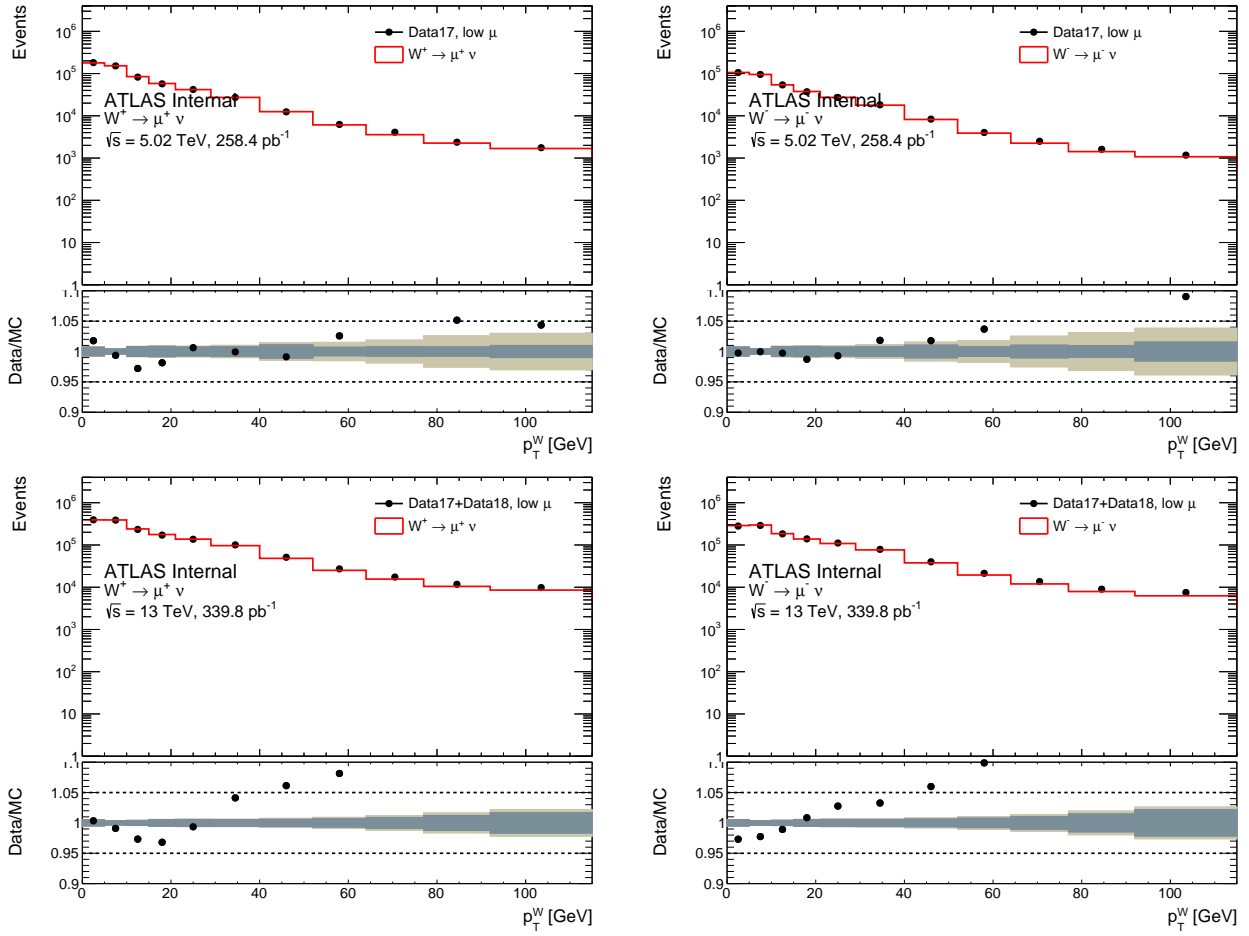


**Figure 9.17:** Unfolded cross-section in the $W^-$ (left) and $W^+$ (right) electron channels, at 5 TeV (top) and 13 TeV (bottom).

The comparison of unfolded spectrum to different theoretical predictions is presented at Figure 9.19 for electron channel and at 9.20 for the muon channel. The estimated experimental uncertainties raise from 1% at low $p_T^W$ to about 5% (2%) at $p_T^W$ =100 GeV, at 5 TeV (13 TeV).

The predictions are generated using Powheg AZNLO, Pythia AZ, Sherpa and DYRES. Powheg and Pythia agree with the data to a similar extent. A softer spectrum is predicted by Sherpa, while DYRES is on the opposite side compared to the data. The observed behaviour holds for both energies, both

**Figure 9.18:** Unfolded cross-section in the $W^-$ (left) and $W^+$ (right) muon channels, at 5 TeV (top) and 13 TeV (bottom).

charges and both decay channels.

## 9.5 Conclusions

There are still some ongoing studies for the optimization of number of unfolding iterations and binning of the truth distributions. But some conclusions could already be made based on the obtained results:

1. The Powheg+Pythia8 AZ tune tuned at the Z $p_T$ spectrum measured at 7 TeV shows a reasonable agreement with the data at 5 TeV at both reconstructed and unfolded levels. On the other hand, at 13 TeV none of the tested MC generators was able to demonstrate good agreement with the data.

2. The observed discrepancies can not be explained by the unfolding bias and lie beyond the
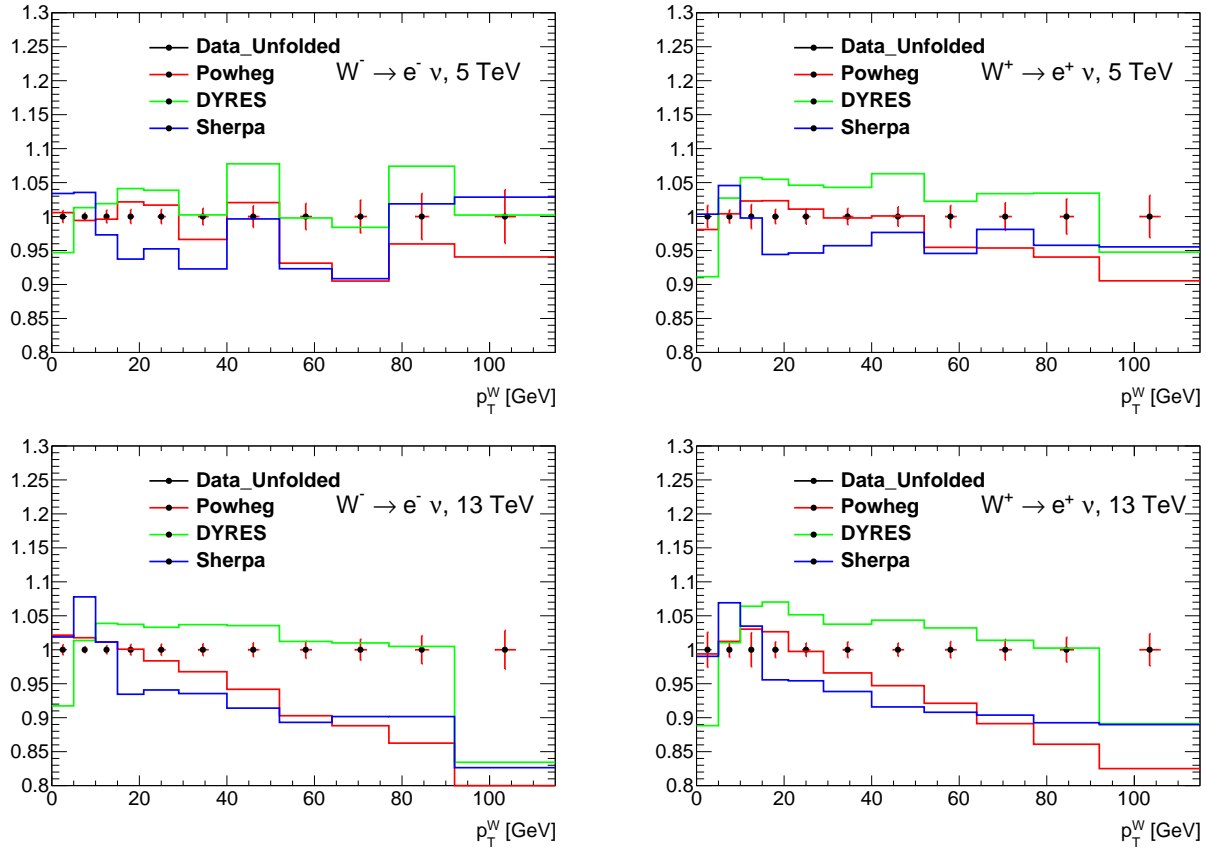
**Figure 9.19:** Unfolded measurement results in the $W^-$ (left) and $W^+$ (right) electron channels, at 5 TeV (top) and 13 TeV (bottom).

measurement uncertainties.

3. The aim of having 1% relative uncertainty in each bin at $p_T < 30\ GeV$ is achievable at both energies.

4. Obtained result can be used to reduce the uncertainty of the W boson mass measurement.
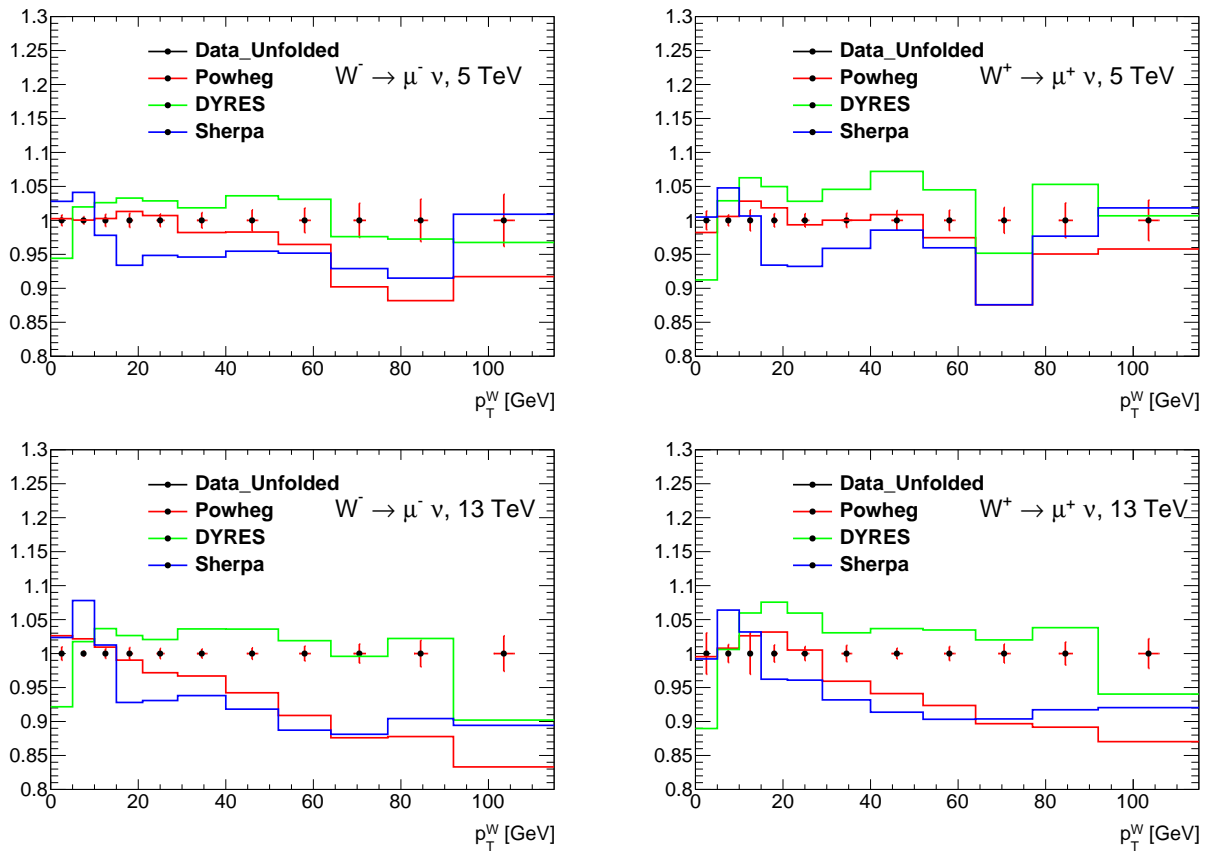
**Figure 9.20:** Unfolded measurement results in the $W^-$ (left) and $W^+$ (right) muon channels, at 5 TeV (top) and 13 TeV (bottom).

# 10

# Hadronic recoil regression with deep neural networks

In the recent years a significant progress was achieved in the field of big datasets analysis. The multivariant analysis techniques, given enough input data, are able to infer very complicated dependencies between the input and target variables. This has motivated me to use deep neural network (DNN) for the regression of the 2-component hadronic recoil vector.

The chapter opens with a theoretical introduction into the underlying principles of the feedforward neural networks used in the analysis. It is followed by the description of the network setup: inputs, hyperparameters and the training process. The results of the regression are benchmarked and discussed in the end of this chapter. The resulting plots for all the W channels are collected in Appendix B.

## 10.1   Deep neural networks

Normally a machine learning problem has a number of ingredients: a dataset $\mathbf{X}$, a set of parameters $\theta$, a model $g(\theta)$ and a loss function $C(\mathbf{X})$ that tells us how well the model $g(\theta)$ describes the dataset. Finding the values of $\theta$ that would minimize the loss function we fit the model.

### 10.1.1   Gradient descent optimization

One of the most powerful and used class of methods in minimizing the loss function is called the *gradient descent*, [170] especially its sub-class, the stochastic gradient descent (SGD) [171], [172]. One of its modifications called ADAM [173] was used as an optimization algorithm in the work presented in this thesis.

Let's assume that a loss function $E(\theta)$ may be estimated as a sum over n data points:

$$E(\theta) = \sum_{i=1}^{n} e_i(x_i, \theta), \tag{10.1}$$

where $x_i$ is a data point and $e_i$ is an estimate of performance. In the simplest case of the gradient descent (GD) algorithm we start looking for the values of parameters $\theta$ such that the sum of functions

$\sum_{i=1}^{n} e_i$ is minimal. We start with a certain value $\theta_0$ and then iteratively perform the following:

$$v_t = \eta_t \, \nabla_\theta \, E(\theta_t),$$
$$\theta_{t+1} = \theta_t - v_t, \tag{10.2}$$

where $\nabla_\theta E(\theta_t)$ is the gradient of $E(\theta)$ with respect to $\theta$; factor $\eta_t$ is called the *learning rate* and defines the length of the step in the direction of $\theta$ performed with every iteration. Balancing learning rate is very important for learning process and convergence. A value too low can make the convergence "stuck" in a local minimum, it also increases the number of iterations. Picking a very high learning rate we risk to miss the minimum so the algorithm would never converge to a minimum. Also, if the number of data points $n$ is high, calculating the gradient is a costly task in terms of CPU time.

Some of the problems accompanying the use of GD are dealt with by using its modification - the SGD. The idea is the following: instead of using all the available data points $n$ at each iteration of the GD, we split the data into $k$ *minibatches*, each having $M$ data points, such that $k = n/M$. Normally the size of the batch is few hundreds of data points, to provide a certain degree of variance and incorporating stochasticity. The transition to SGD algorithm is done in the following way:

$$\nabla_\theta E(\theta) = \sum_{i=1}^{n} \nabla_\theta e_i(x_i, \theta) \rightarrow \sum_{i \in B_l} \nabla_\theta e_i(x_i, \theta), \tag{10.3}$$

where $B_l$ is a set of data points belonging to a minibatch $l \in 1, ..., n/M$. Now every next iteration of $\theta$ parameters update is performed over a different batch, consecutively running over all the batches:

$$\nabla_\theta E^{EM}(\theta) = \sum_{i \in B_l} \nabla_\theta e_i(x_i, \theta),$$
$$v_t = \eta_t \, \nabla_\theta \, E^{EM}(\theta_t),$$
$$\theta_{t+1} = \theta_t - v_t. \tag{10.4}$$

A full iteration over all the $n/M$ batches is called an *epoch*. Now stochasticity prevents the gradient algorithm from getting stuck in a local minimum. Also computing the gradient over fewer data point notably decreases the CPU time spent.

The algorithm may be further improved, adding a "memory", that is to say making every next step $t$ dependent on the direction of the previous step $t-1$:

$$v_t = \gamma v_{t-1} \eta_t \, \nabla_\theta \, E^{EM}(\theta_t),$$
$$\theta_{t+1} = \theta_t - v_t. \tag{10.5}$$

Because of an analogy from physics the parameter $\gamma$ is called a *momentum*, having $0 \leq \gamma \leq 1$ [174], [175]. This parameter provides a certain "inertia" in the change of the direction of the gradient descent. Introduction of the momentum helps for quicker convergence in the case of a slow but steady change of a certain parameter during the gradient descent.

The convergence of the GD may be significantly improved if the learning rate could be different in different directions, depending on the landscape of the parameter space $\theta$: the steeper the gradient

in a certain direction - the smaller the corresponding step. The optimal step could be estimated by obtaining the *Hessian matrix* in the vicinity of a point $\theta_0$, providing a description of the local curvature in a multidimensional space (although calculating Hessian matrix is complicated and slow-converging process [176]). However, a number of methods use the second moment of the gradient to efficiently estimate the optimal learning rate. One of such methods is called ADAM (ADAptive Momentum) [173], its iterative relations are the following:

$$
\begin{aligned}
g_t &= \nabla_\theta E(\theta_t) \\
m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t \\
s_t &= \beta_2 s_{t-1} + (1 - \beta_2) g_t^2 \\
\hat{m}_t &= \frac{m_t}{1 - (\beta_1)^t} \\
\hat{s}_t &= \frac{s_t}{1 - (\beta_2)^t} \\
\theta_{t+1} &= \theta_t - \eta_t \frac{\hat{m}_t}{\sqrt{\hat{s}_t} + \epsilon}.
\end{aligned}
\tag{10.6}
$$

Here the parameters $\beta_1$ and $\beta_2$ set the memory lifetime for the first and second moment; $\eta$ is the learning rate and $\epsilon$ is a small regularization constant keeping the denominators from vanishing. Like in other cases of the SGD here the iterations are performed batch-wise. Parameter $s_t$ is linked to the variance of the gradient size. This basically means that the learning rate is proportional to the first momentum of the gradient and inverse proportional to its standard deviation.

### 10.1.2 DNN structure and training

A neural network is composed of single neurons, also called nodes, arranged in layers. The first layer is called the input layer, the last one is called the output layer; all the layers in between are named hidden layers (see Fig. 10.1).

A single node $i$ takes a vector of k input features $\mathbf{x} = (x_1, x_2, ..., x_k)$ and produces a scalar input $a_i(\mathbf{x})$. Function $a_i$ may have a different form, although it normally can be decomposed into two steps. The first step is a linear transformation of the inputs into a scalar value assigning each input a weight:

$$
z^i = w_k^i \cdot x_k + b^i,
\tag{10.7}
$$

where $\mathbf{w}^i = (w_1^i, w_2^i, ..., w_k^i)$ is a set of $k$ weights assigned to corresponding inputs. The weights $\mathbf{w}^i$ are specific to a neuron $i$, as well as the scalar bias $b^i$. The next step is where the non-linear function $\sigma_i$ comes into play: we can express the output function $a_i(\mathbf{x})$ as follows:

$$
a_i(\mathbf{x}) = \sigma_i(z^i).
\tag{10.8}
$$

There exists a number of options for the non-linear function $\sigma$; in current thesis a tanh is used. When the neurons are arranged in layers in a feed-forward neural network - the outputs from neurons of the previous layer serve as inputs for the succeeding layers neurons. The universal approximation theorem states, that a neural network with a single hidden layer can approximate any continuous

multi-parametric function with arbitrary accuracy [177], [178]. However, in practice it is easier to reach the possible precision having more hidden layers.

So in terms of a Deep Neural Network (DNN) fitting the model means tuning the weights and biases $(\mathbf{w}^i, b^i)$ in such a way that a loss function applied to the new dataset would be minimal. It is reached through iterative process called *training*, that involves the GD with an algorithm called *backpropagation* [179]. The backpropagation algorithm allows to calculate the gradients and adjust the corresponding parameters in a very computation-efficient way.
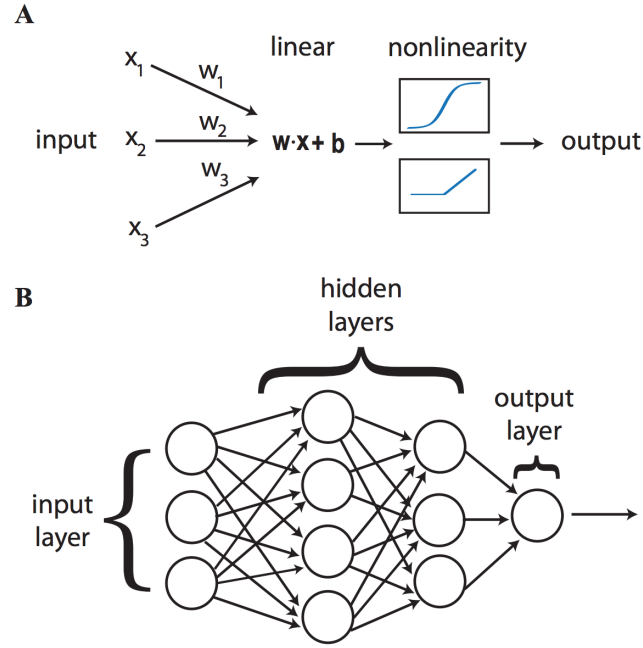


**Figure 10.1:** A: The nodes perform a linear transformation of the inputs, then apply a non-linear activation function. B: The architecture of a deep neural network: neurons are arranged into layers [180].

Let us assume that there are L layers in the network $l = 1, ..., L$, that $w_{jk}^l$ and $b_j^l$ are the weight of an input parameter k and the bias for node $k$ in layer $l$ respectively. The layered structure of the neural network ensures that the inputs for the nodes in layer $l$ depend only on the outputs of the nodes from layer $l - 1$, hence:

$$a_j^l = \sigma\left(\sum_k w_{jk}^l a_k^{l-1} + b_j^l\right) = \sigma(z_j^l), \tag{10.9}$$

where the linear weighted sum is denoted as:

$$\sigma(z_j^l) = \sum_k w_{jk}^l a_k^{l-1} + b_j^l. \tag{10.10}$$

The cost function E is computed from the output of the neural network, so it directly depends only on the values of $a_j^L$. Let us define the error $\Delta_j^L$ of the j-th node in the output (L-th) layer as a change in the

cost function with respect to the weighted output of the last layer:

$$\Delta_j^L = \frac{\partial E}{\partial z_j^L}. \tag{10.11}$$

At the same time the loss depends indirectly on all the preceding layers, so keeping in mind eq. 10.9 we can define the error of an arbitrary node $j$ in arbitrary layer $l$ as the change in the cost function E with respect to the weighted input $z_j^l$:

$$\Delta_j^l = \frac{\partial E}{\partial z_j^l} = \frac{\partial E}{\partial a_j^l}\sigma'(z_j^l), \tag{10.12}$$

where $\sigma'(z_j^l)$ is the derivative of the non-linear activation function $\sigma$ with respect to its input at $z_j^l$. But on the other hand we can also interpret the error function $\Delta_j^L$ in terms of bias partial derivatives:

$$\Delta_j^l = \frac{\partial E}{\partial z_j^l} = \frac{\partial E}{\partial b_j^l}\frac{\partial b_j^l}{\partial z_j^l} = \frac{\partial E}{\partial b_j^l} \cdot \mathbf{1}. \tag{10.13}$$

So starting from the output layer we can compute the error in any layer $l$, provided we know it for the subsequent layer $l + 1$:

$$\Delta_j^l = \frac{\partial E}{\partial z_j^l} = \sum_k \frac{\partial E}{\partial z_j^{l+1}}\frac{\partial z_j^{l+1}}{\partial z_j^l} =$$
$$= \sum_k \Delta_j^l \frac{\partial z_j^{l+1}}{\partial z_j^l}\left(\sum_k \Delta_j^l w_{kj}^{l+1}\right)\sigma'(z_j^l). \tag{10.14}$$

And finally we can get the gradient of the cost function E with respect to a weight of an arbitrary neuron:

$$\frac{\partial E}{\partial w_{jk}^l} = \frac{\partial E}{\partial z_j^l}\frac{\partial z_j^l}{\partial w_{jk}^l} = \Delta_j^l a_k^{l-1}. \tag{10.15}$$

Using these four equations (10.11, 10.13, 10.14, 10.15) it is possible to "backpropagate" the error back from the output layer and once we can compute the gradient - we know how we should tune the weights and biases in order to minimize the loss function.

### 10.1.3 Batch normalization

Batch normalization is a regularization scheme that helps to improve the speed and stability of the DNN training. The main idea behind the method is to prevent an *internal covariant shift* - a change in the distribution of network activations due to the change in network parameters during training by means of normalization of the parameters transferred from layer $l$ to layer $l + 1$ [181]. So let us consider a layer $l$ that has $d$ inputs $\mathbf{x} = (x^1, x^2, ..., x^d)$, then for every $x^k$ we perform the following transformation:

$$\hat{x}^k = \frac{x^k - E[x^k]}{\sqrt{Var[x^k]}}, \tag{10.16}$$

where $E[x^k]$ and $Var[x^k]$ are the expectation and variance of the parameter $x$, calculated over the training dataset, respectively. At the same time, we have to be sure that we preserve the non-linearity of the activation function output. In order to do this the two additional parameters are introduced:

$$y^k = \gamma \hat{x}^k + \beta^k, \tag{10.17}$$

where the parameters $\gamma$ and $\beta$ are trained just like the rest of the network parameters. Practically if the training is performed within the mini-batch scheme with batch size $B = x_1,...,x_m$ the batch normalization layer is inserted between the DNN layers the transformations for the input $x$ are the following:

$$\begin{aligned}
\frac{1}{m} \sum_{i=1}^{m} x_i &\rightarrow \mu_B \\
\frac{1}{m} \sum_{i=1}^{m} (x_i - \mu_B)^2 &\rightarrow \sigma_B^2 \\
\frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} &\rightarrow \hat{x}_i \\
\gamma \hat{x}_i + \beta &\rightarrow y_i \equiv BN_{\gamma,\beta}(x_i),
\end{aligned} \tag{10.18}$$

where $\epsilon$ is a small regularization constant.

## 10.2 The hadronic recoil regression

Considering that hadronic recoil is an observable that uses many inputs from ID, EMC and HC it is reasonable to expect improvement of the result using modern MultiVariate Analysis (MVA) techniques.

### 10.2.1 Input features and model

Training, testing and validation was performed using a simulated MC sample $W^+ \rightarrow \mu\nu$ at 13 TeV, following the selection presented in Section 8. Out of 15 768 239 events that have passed the selection 12 734 109 were used for training and 3 034 130 for testing the performance. Below is the list of 38 input features:

- **Vector sums** of charged PFO, neutral PFO (see Section 6.5) and the vector sum of both. All three vector sums are included into the input features having two Cartesian components each, making 6 input features.

- **Transverse energy sum** $\sum E_T$ is also defined in three similar ways, adding three input features.

- **Cartesian components of the two leading jets momenta in the transverse plane. The jets were demanded to have $p_T > 20$ GeV. If one or both jets don't make the cut or there is less than two jets in the event - the corresponding features were assigned zero value.

- Cartesian components of the five leading Neutral Particle Flow Objects (nPFOs) and five leading Charged Particle Flow Objects (cPFOs) momenta in the transverse plane.

- Number of primary vertices in the event.

- Pile-up value $\mu$.

- Total number of jets in the event.

- Total number of nPFOs and cPFOs in the event.

All input features were pre-processed using the StandardScaler module from Scikit Learn package [182].

The model contains 3 dense layers with 256 neurons each, alternated with batch normalization layers (see Fig. 10.2). Using batch normalization layers has allowed to reduce the training time by a factor 10. The model has used Adam optimizer with learning step 0.001 and batch size of 4000 data points.
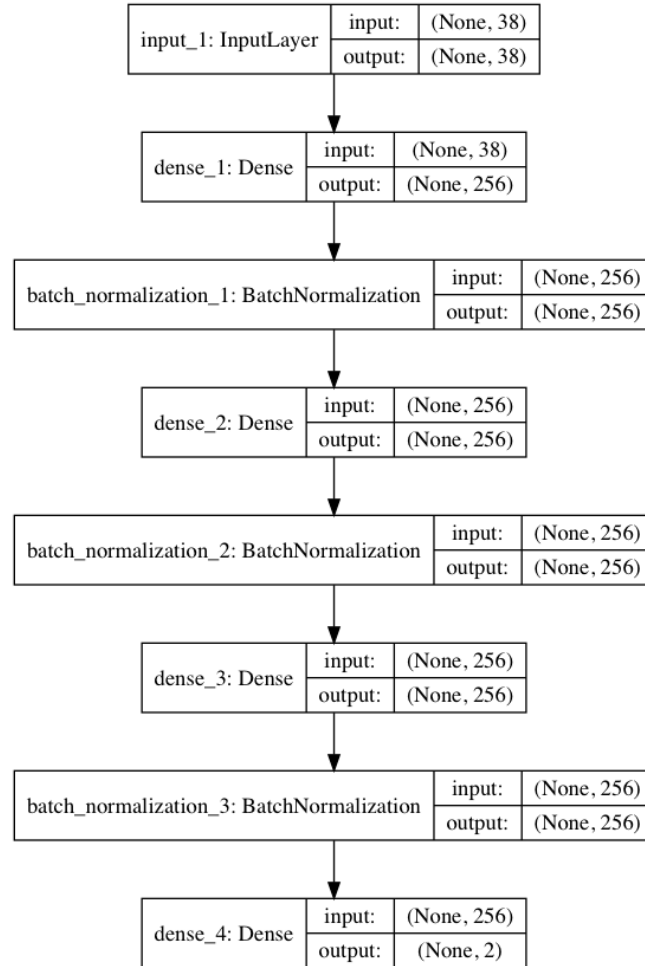


**Figure 10.2:** A model of the DNN used in the analysis.

Twenty percent of events were used for validation. The two target values were Cartesian components of the truth $p_T^W$ vector.

The loss function used is the mean square error, which means that for every batch of size $B$ the loss is defined as:

$$\mathcal{L} = \sum_{i=1}^{B} (x_i^{pred} - x_i^{target})^2, \tag{10.19}$$

where $x_i^{target}$ and $x_i^{pred}$ are the target (truth) and predicted values for the $i^{th}$ event respectively. For our case of two target values, namely $u_X$ and $u_Y$ the loss takes the following form:

$$\mathcal{L} = \sum_{i=1}^{B} \left( (x_i^{pred} - x_i^{target})^2 + (y_i^{pred} - y_i^{target})^2 \right). \tag{10.20}$$

Figure 10.3 shows dependence of loss on training epochs. Eventually the model with weights obtained after 38 epochs of training was used in the analysis.
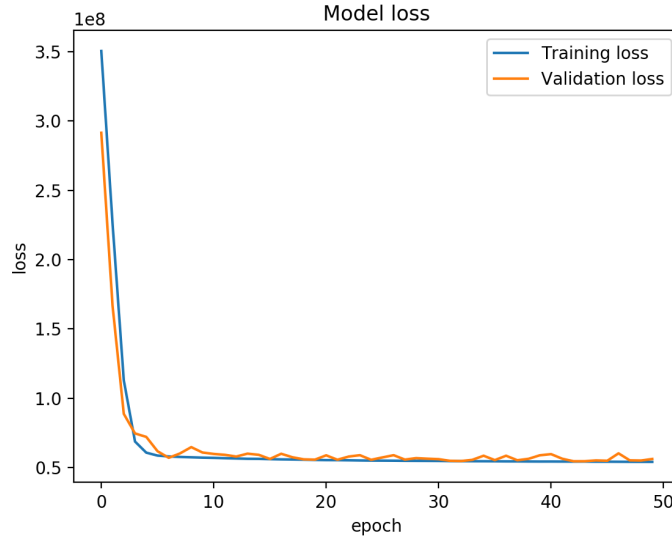


**Figure 10.3:** Learning curve of the model.

## 10.2.2 Kinematic distributions

The results presented here show the regression plots obtained with the trained DNN. The regression was tested for the four W channels in MC and for the two Z channels in both data and MC. Below the plots for $W^- \to \mu\nu$ channel are presented, the rest of the results are presented in Appendix B.
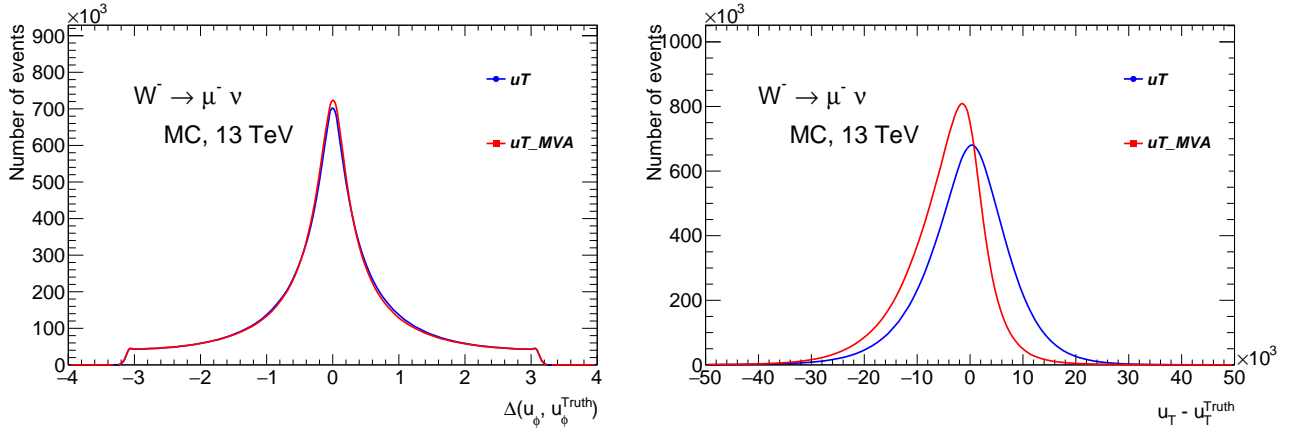
In Figure 10.4 the difference between the reconstructed hadronic recoil and truth distributions (targets) are shown. For the two target components $u_X$ and $u_Y$ a sharper peak centred at zero is observed in the MVA-reconstructed recoil comparing to the standard algorithm. Similar comparison is shown in Fig. 10.5 for HR vector components in polar coordinates. The polar angle $u_\phi$ of the HR vector shows a very small, nearly negligible improvement from the MVA. The $u_T$ vector magnitude demonstrates a

shift from zero for the $u_T^{MVA}$, indicating a bias. Figure 10.6 shows that the MVA-reconstructed recoil has a softer spectrum comparing to both $p_T^{truth}$ and standard recoil spectra.

The bias is shown in Fig. 10.7 together with the $u_{perp}$ component. Indeed the $u_T^{MVA}$ demonstrates a larger bias as compared to the standard algorithm. At the same time the standard deviation of the $u_\perp$ component indicates an improvement in the resolution.
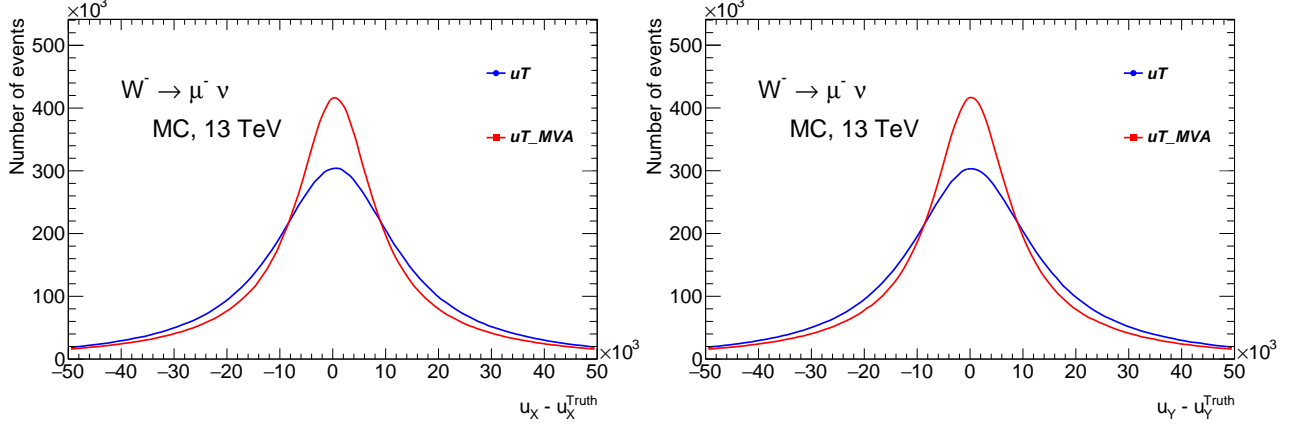


**Figure 10.4:** The difference between the Cartesian components of the HR vector for the standard $u_T$ and $u_T^{MVA}$.



**Figure 10.5:** The difference between the polar components of the HR vector for the standard $u_T$ and $u_T^{MVA}$.

The dependence of the bias and $u_\perp$ on the momentum is studied in Fig. 10.8. The $u_T^{MVA}$ recoil demonstrates improvement in the resolution and a larger bias in the region of $p_T < 80\text{GeV}$. However, for a quantitative resolution comparison we need to make sure that $u_\perp$ and $u_\perp^{MVA}$ are on the same scale. A possible way to achieve this is to normalize them to average recoil $< u_T >$. The resulting normalized curves on Fig. 10.9 show that the MVA provides 5-10% resolution improvement at $p_T < 10\text{GeV}$ and a bit more than 10% in $10 < p_T < 30\text{GeV}$ transverse momentum region.
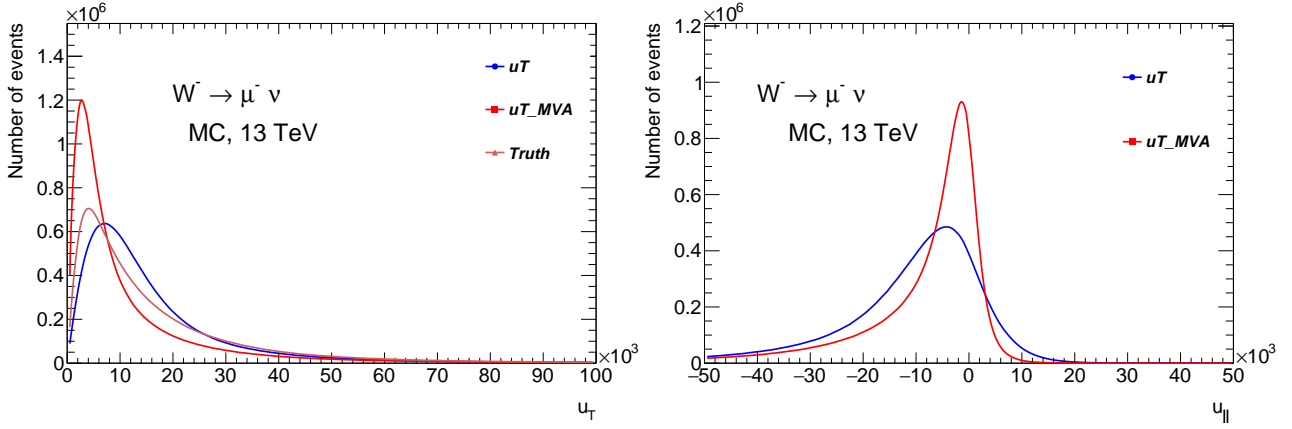
**Figure 10.6:** Comparison of $u_T$ and $u_T^{MVA}$ spectra and response.
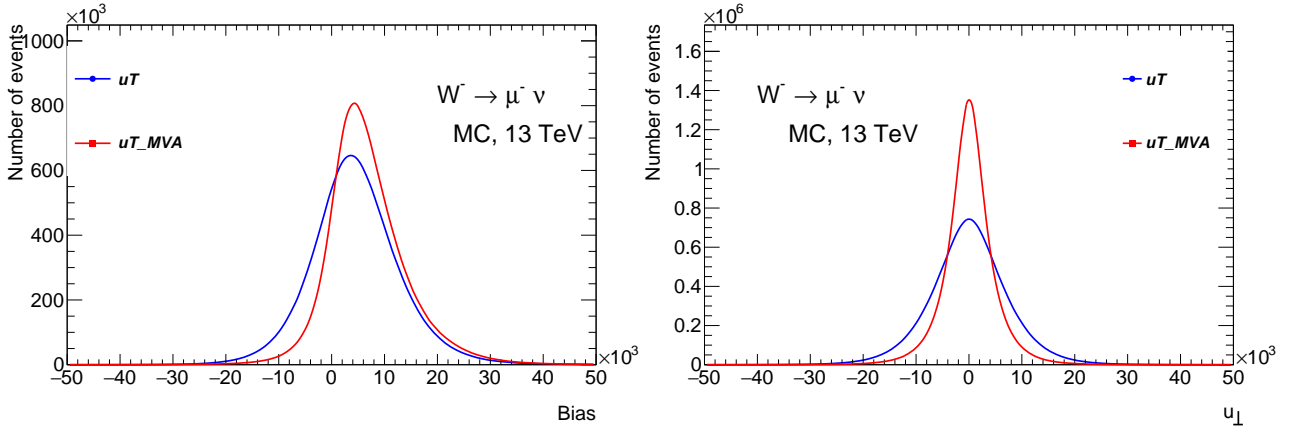


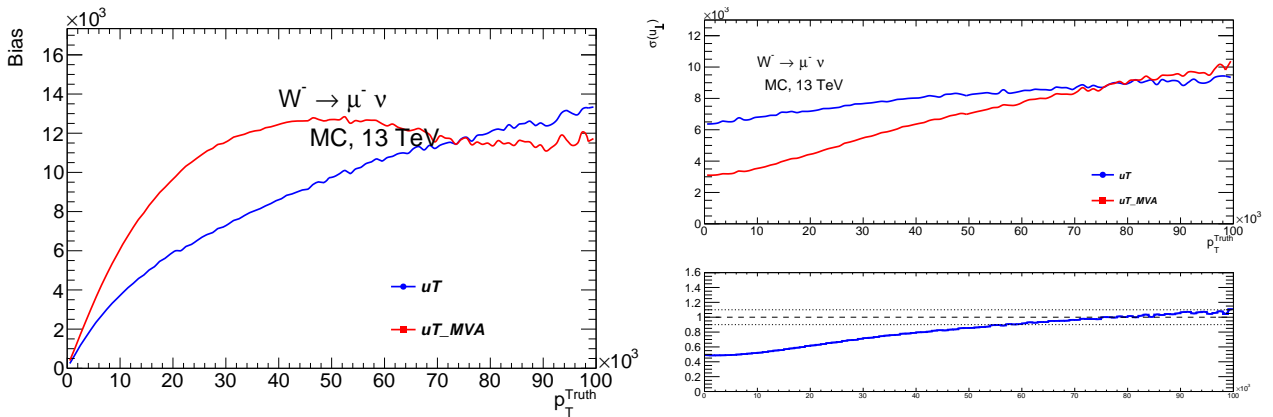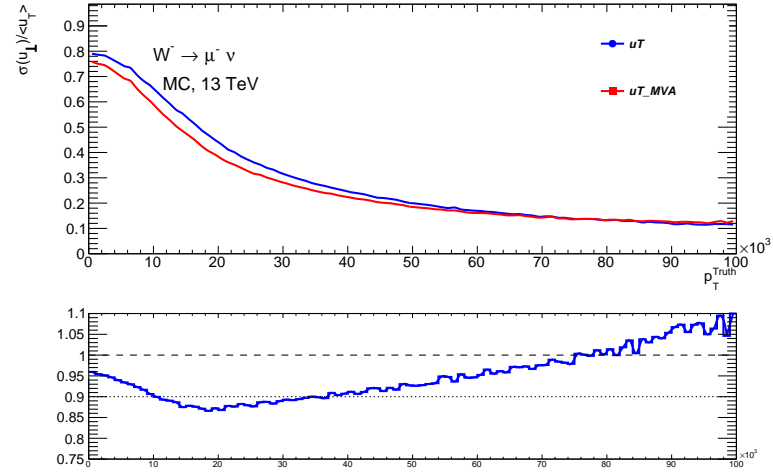**Figure 10.7:** Bias and $u_\perp$ for the standard $u_T$ and $u_T^{MVA}$.



**Figure 10.8:** Bias and $u_\perp$ as a function of $p_T^{truth}$.

**Figure 10.9:** Normalized $\frac{u_\perp}{\langle u_T \rangle}$ as a function of $p_T^{truth}$.

The same neural network was applied for the HR regression for both Z channels, for data and MC events. In case of data the $p_T^{truth}$ vector was replaced with $p_T^{\ell\ell}$. A qualitatively similar picture holds for both Z channels, confirming recoil universality for W and Z events (see Fig. 10.10). The complete set of $u_T$ vs $u_T^{MVA}$ comparison plots for all W and Z channels can be found in Appendix B.
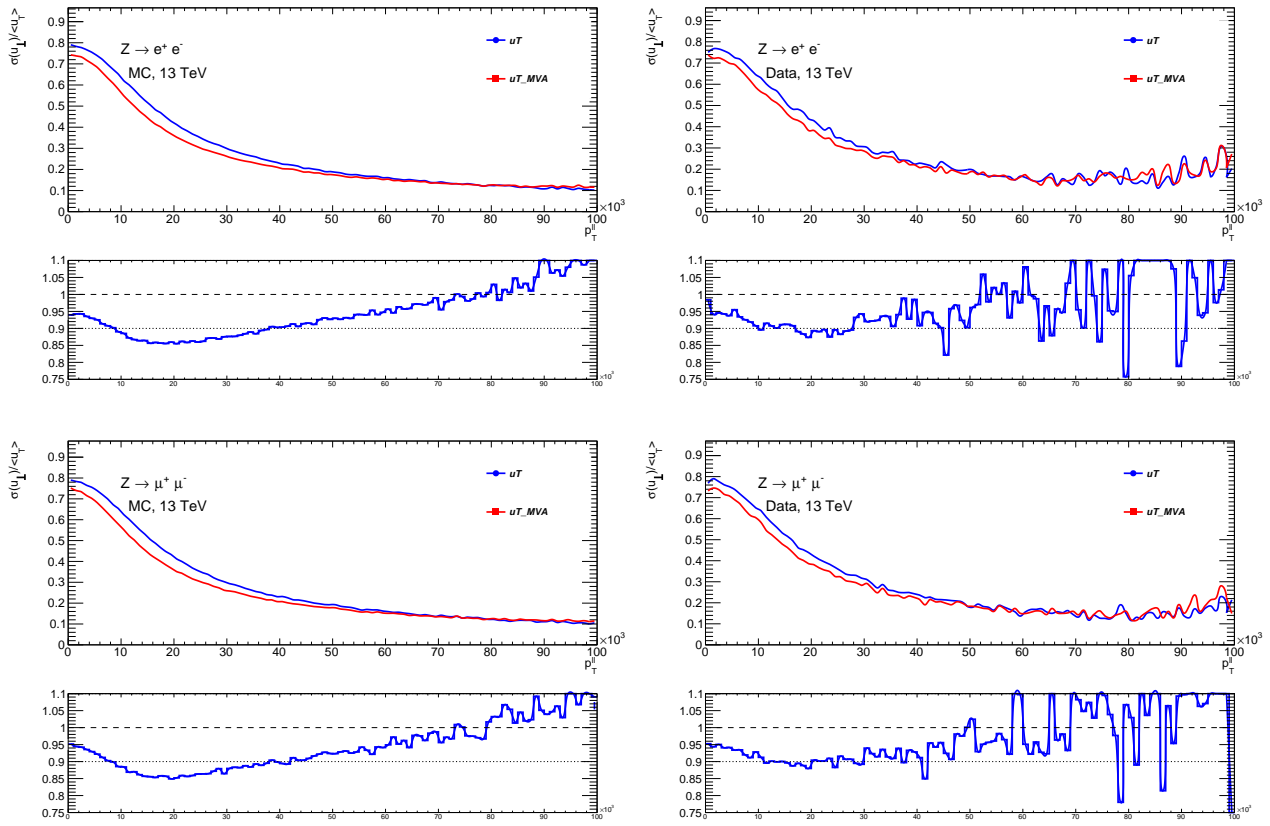
**Figure 10.10:** Relative resolution for Z channels in MC (left) and in data (right).

## 10.3 Rescaling of the MVA hadronic recoil

As we can see from the pT plot the average scale of the MVA-reconstructed hadronic recoil is smaller than that of the nominal hadronic recoil. This leads to an increased bias. Using the $u_\parallel - p_T$ dependence we can rescale the MVA hadronic recoil to match the nominal recoil scale.
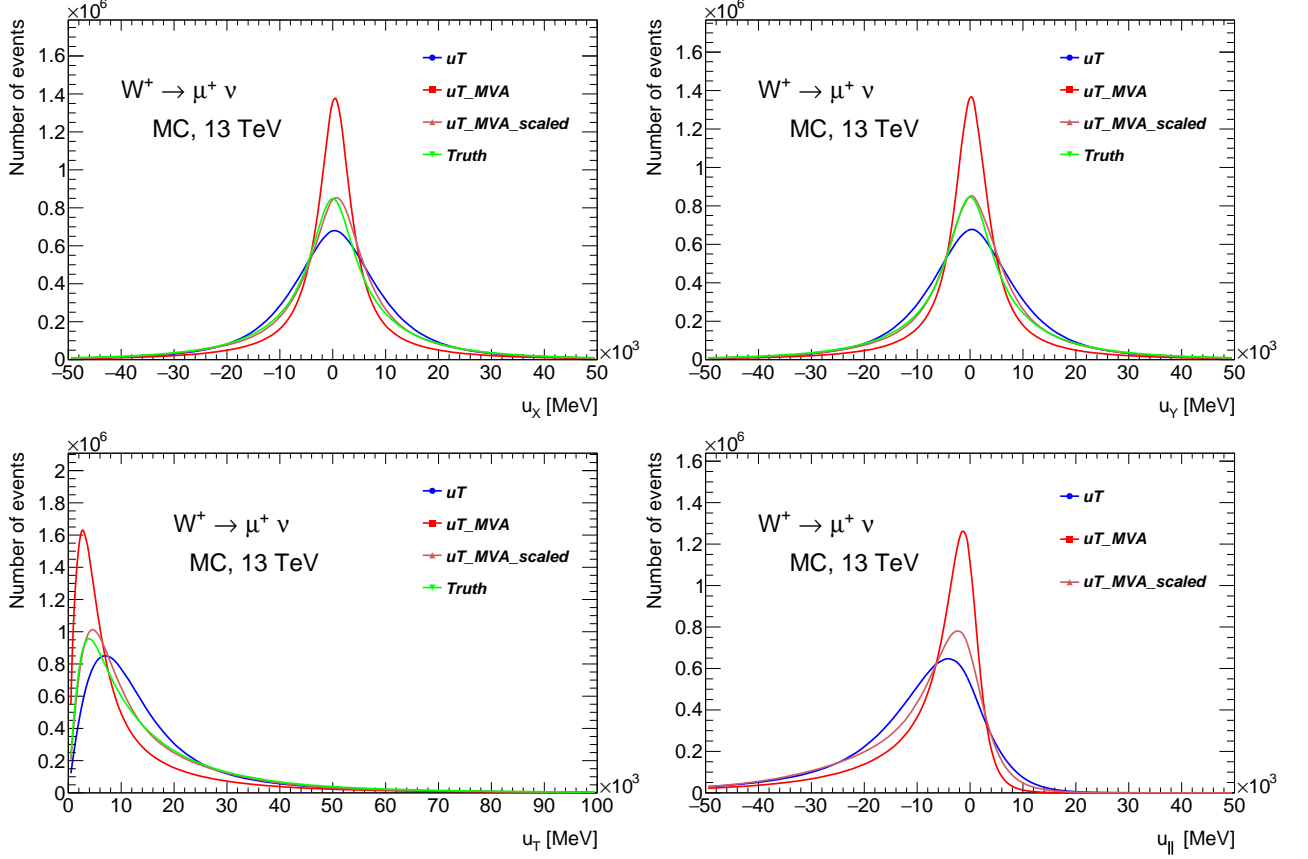


**Figure 10.11:** The target distributions $u_x$ and $u_y$ (top plots). Bottom plots show th uT spectra (left) and $u_\parallel$ distributions (right).

The plots demonstrate that rescaling of the MVA-derived hadronic recoil makes the average bias roughly the same as in the nominal hadronic recoil. At the same time the relative resolution retains the improvement seen in the non-scaled MVA hadronic recoil.
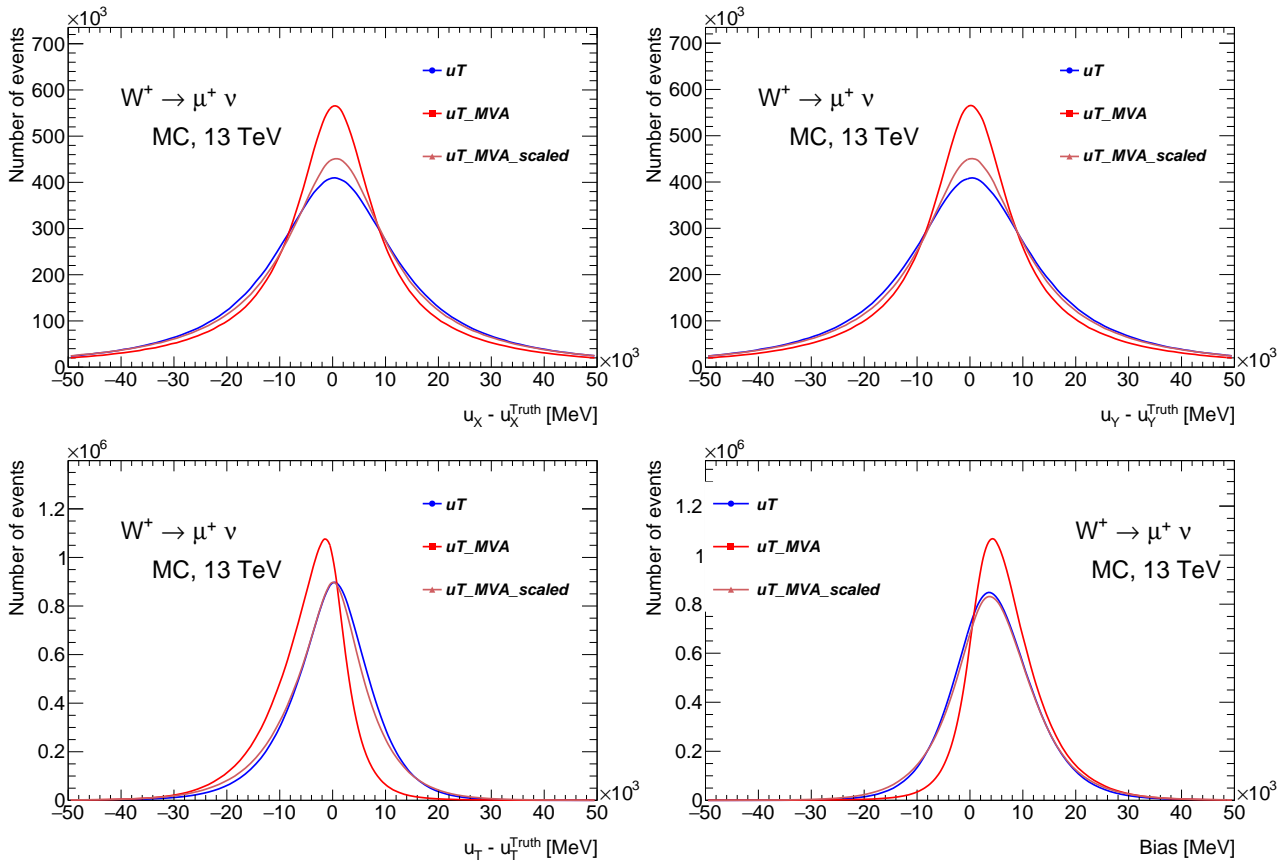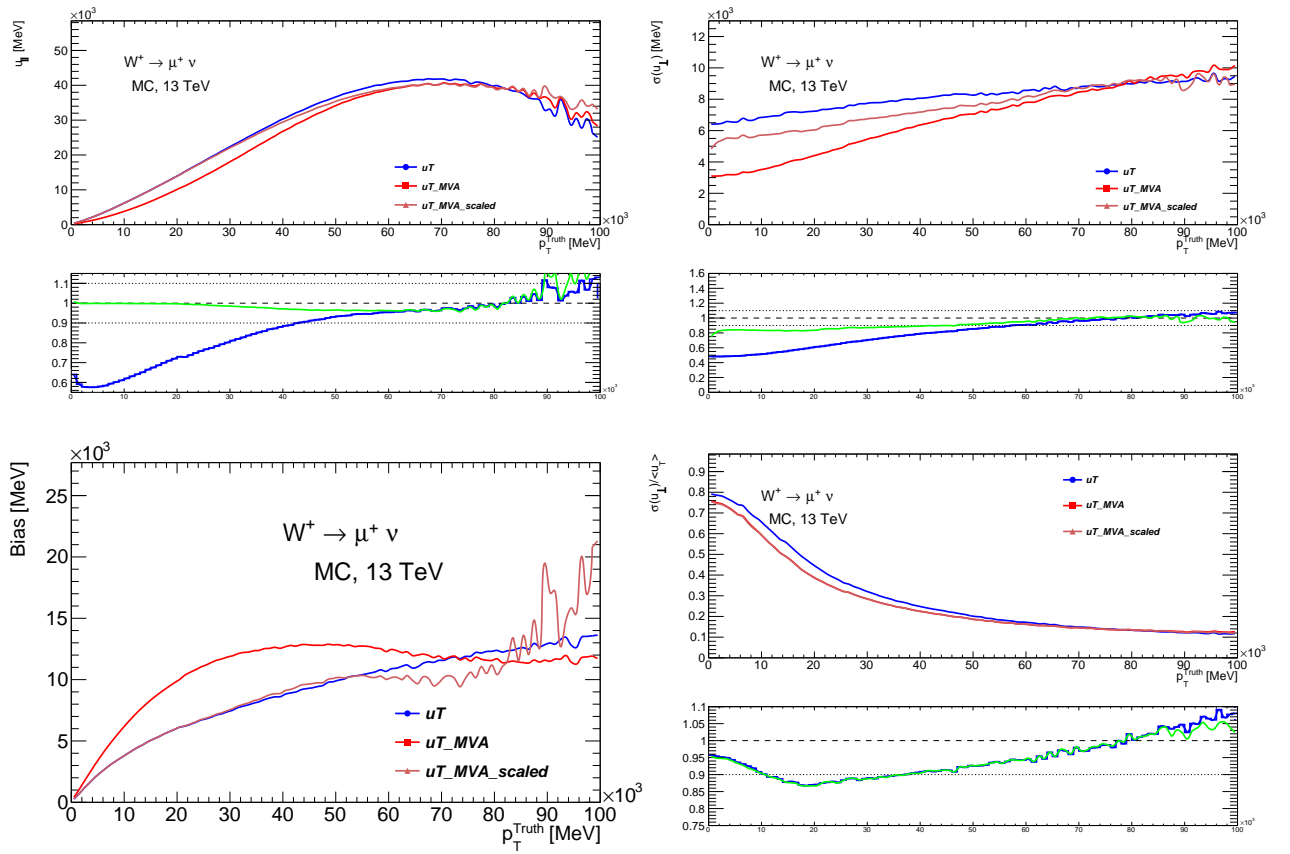
**Figure 10.12:** Difference plots between the target and truth values.

**Figure 10.13:** Bias, $u_{\parallel}$, $\sigma(u_{\perp})$ and $\sigma(u_{\perp})/<u_T>$ dependencies on the $p_T^{truth}$.

## 10.4   Template fits for W boson mass

A way to estimate the improvement is to test if the new hadronic recoil definition results in the improved sensitivity for the W boson mass. In order to do this a set of templates is produced where the mass of the W boson is shifted from its nominal value. The kinematic distributions are reweighted to reflect the shift of the W mass. Every template is compared to the nominal distribution using the $\chi^2$ test. The more sensitive is the observable to the $m^W$ change - the larger is the $\chi^2$ value of the corresponding template. The procedure is described in detail in the W mass paper internal note [165]. In Fig. 10.14 an example of the $\chi^2$ parabola fit is demonstrated. The MVA-derived hadronic recoil shows a considerably higher sensitivity to the W boson mass shift.
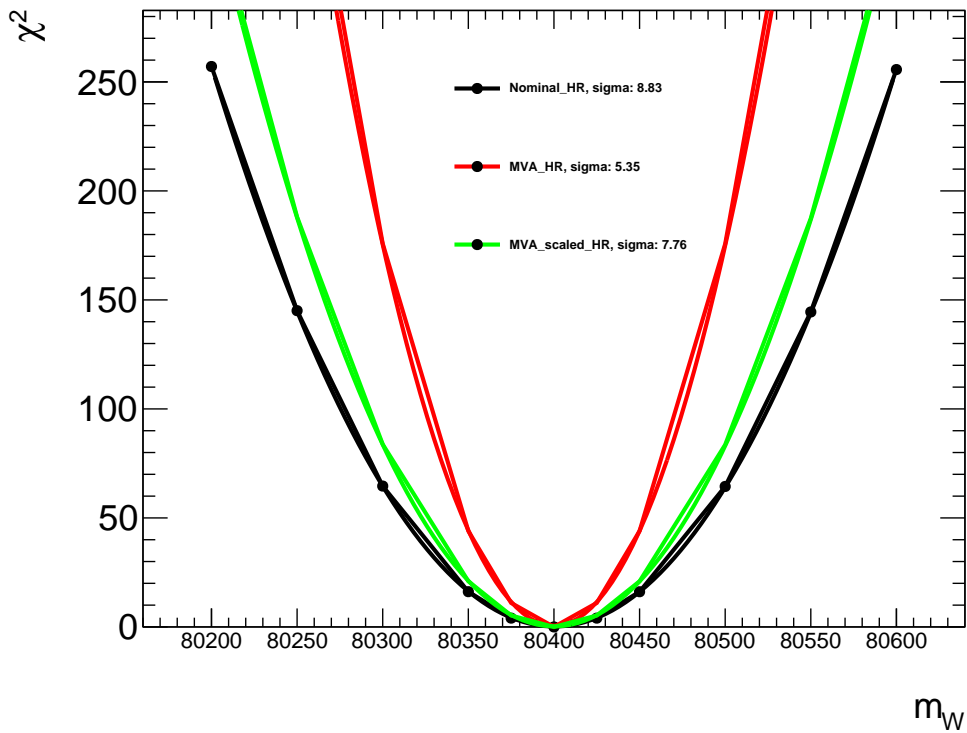


**Figure 10.14:** An example of a $\chi^2$ fit of the $E_T^{miss}$ distribution templates for the three hadronic recoil definitions.

The next step is to produce bootstrap toys in order to estimate the statistical uncertainty of the W mass fits and the correlation between the observables. For every HR definition a set of 1000 toys has been generated. Every toy is fitted using the generated templates, providing a value for the W mass. Putting these values onto a 2D scatter plot allows to estimate the correlation. The standard deviation of the distributions provides the statistical uncertainty.

Here we see that the smaller scale of the MVA hadronic recoil increases the correlation between the observables. On the other hand the transverse mass and missing transverse energy uncertainties
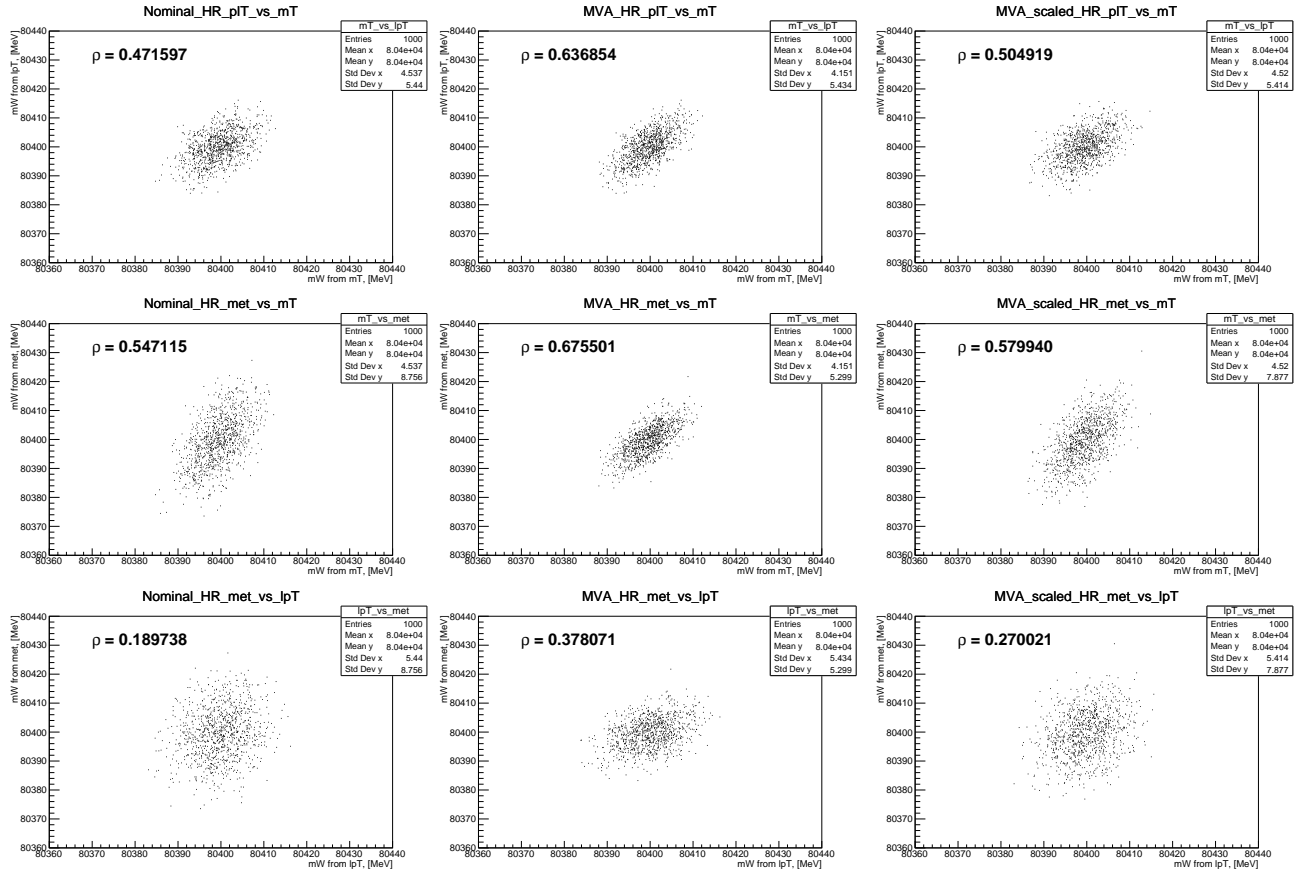
**Figure 10.15:** Scatter plots that illustrate the correlation between the observables for the three hadronic recoil definitions: Nominal (left column), MVA (central column) and scaled MVA (right column).

| Stat. uncertainty, [MeV] | $m_T$ | $E_T^{miss}$ | $p_T^l$ |
|---|---|---|---|
| Nominal | 4.54 | 8.76 | 5.44 |
| MVA | 4.15 | 5.3 | 5.43 |
| Scaled MVA | 4.52 | 7.88 | 5.41 |

**Table 10.15:** Estimate of the Z vertex reweighting effect on the W boson mass.

are smaller in the MVA case.

# Summary

This thesis presents author's contributions to the ATLAS experiment combined performance and to the measurement of the W boson transverse momentum distribution. In this small chapter I have summarized the main conclusions of my work.

The correction of the electromagnetic shower shapes presented in Chapter 5 can be summarized in the following way:

- As it was expected, the correction has a positive effect on the the efficiencies (see Fig. 5.15).

- In the end-cap the correction of efficiency reaches 1-3%, in the barrel the change is smaller. This reflects the fact that the MVA algorithm that makes the ID decision is more sensitive to the shower shapes in the end-cap region, while in the barrel it relies more on the other inputs.

- The proposed algorithm was adopted as baseline for the official data analysis framework of the ATLAS experiment.

- The correction is introduced on the cell level, which makes it also useful for alternative electron identification algorithms that rely directly on the cell energies rather than on the shower shapes.

The measurement of the W boson transverse momentum spectrum has lead to a number of results:

- At $\sqrt{s} = 5$ TeV the obtained results demonstrate fair agreement with the Powheg+Pythia8AZ simulation, tuned at 7 TeV data. At $\sqrt{s} = 13$ TeV a significant discrepancy with the Powheg+Pythia8AZ tune predictions is observed and none of the tested MC generators are able to demonstrate agreement with the data.

- Target precision of 1% in every bin is achieved.

- The obtained direct measurement of the W boson transverse momentum would allow to reduce the theoretical modelling uncertainty for the W boson mass measurement.

Using the DNN algorithm for the hadronic recoil reconstruction also allows to make some conclusions:

- The application of deep learning methods is justified and improves the hadronic recoil resolution.

- The regression turns out to be channel-independent and demonstrates similar performance for different Z and W channels, both in MC simulation and in the data.

- The hadronic recoil reconstructed with the DNNs demonstrates better sensitivity to the W boson mass.

# Appendix A: Experimental uncertainties
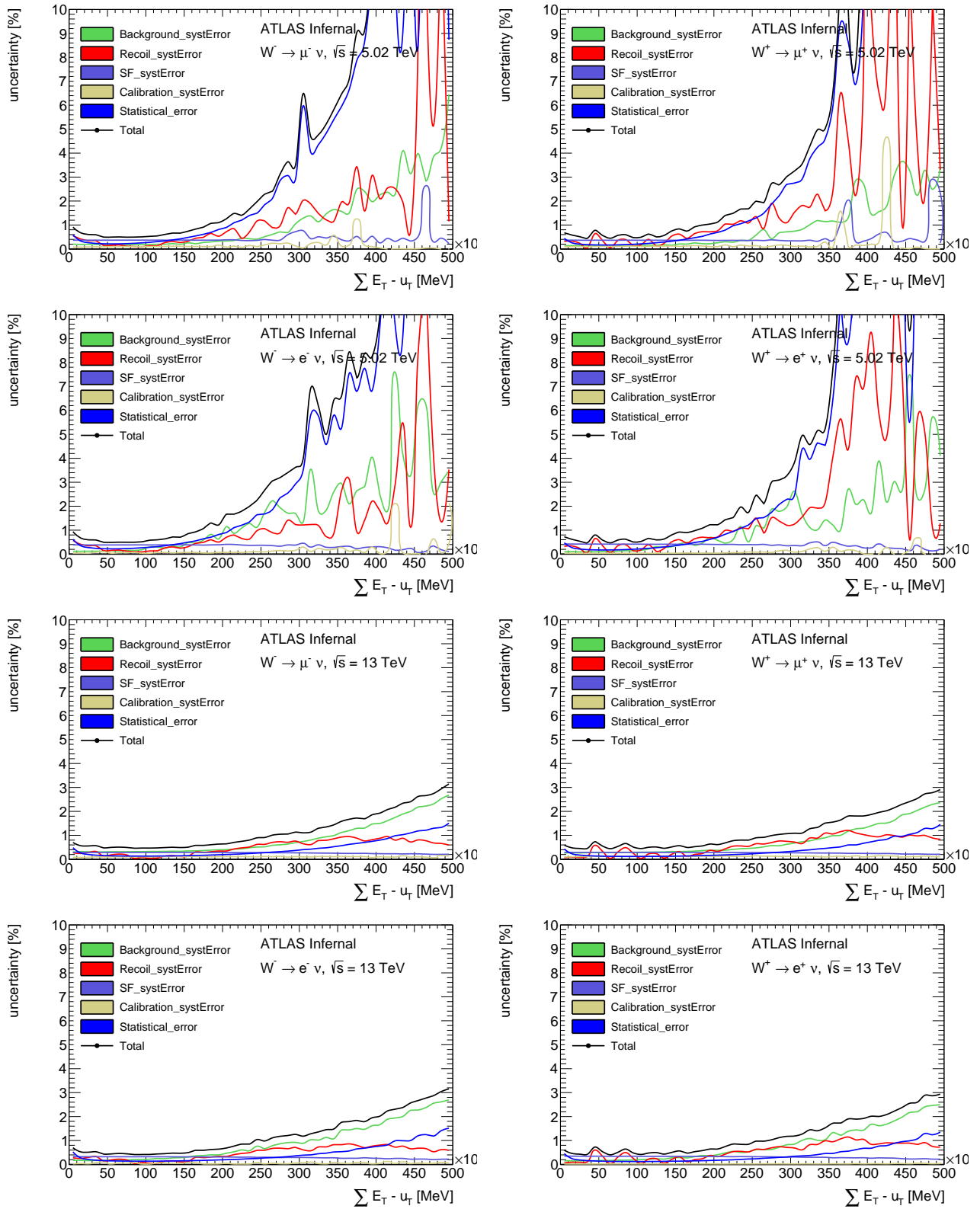
**Figure 10.16:** $\Sigma \bar{E}_T$ systematic error breakdown in the muon and electron channel for the $\sqrt{s} = 5$ TeV and $\sqrt{s} = 13$ TeV datasets.
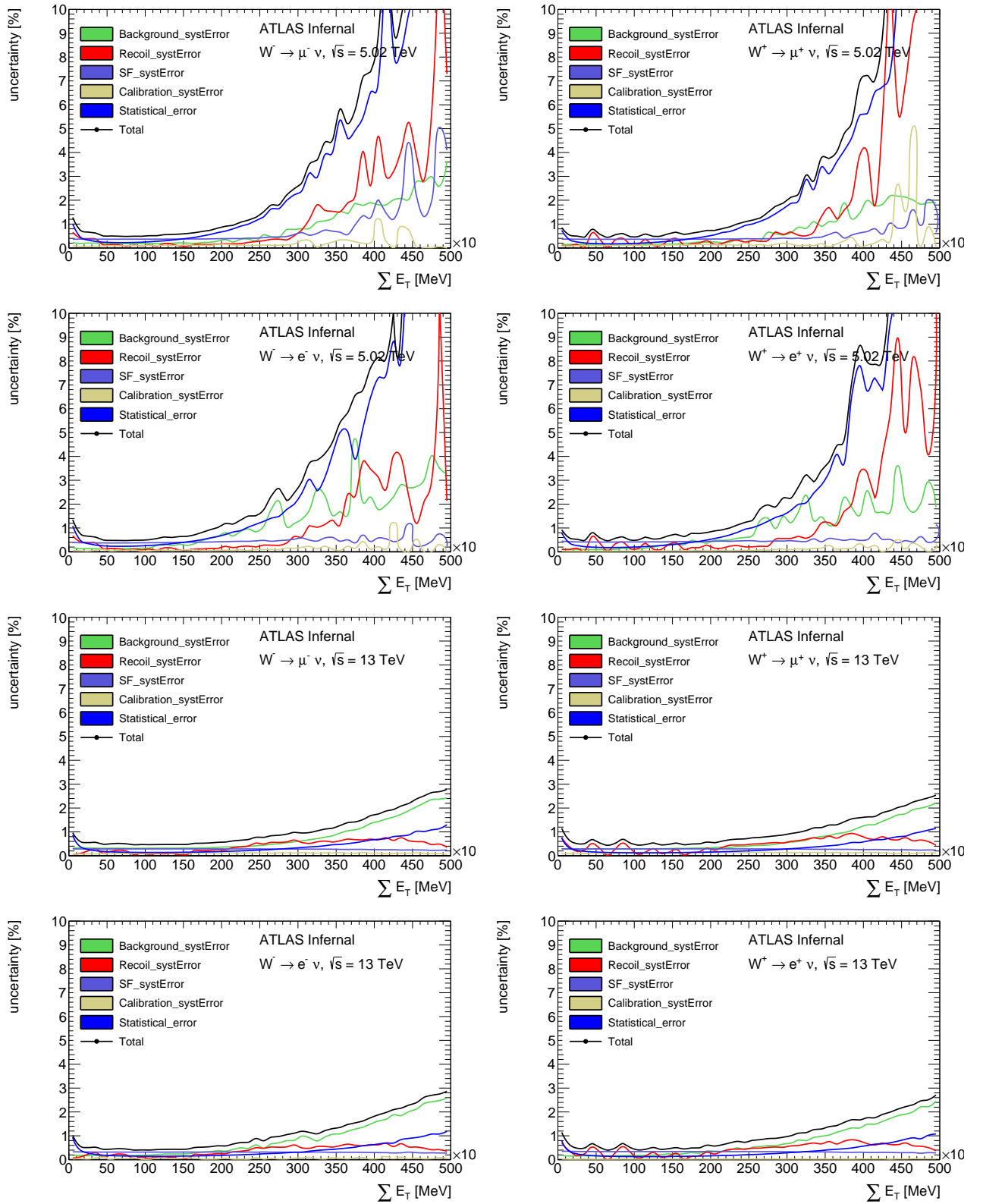
**Figure 10.17:** $\Sigma E_T$ systematic error breakdown in the muon and electron channel for the $\sqrt{s} = 5$ TeV and $\sqrt{s} = 13$ TeV datasets.

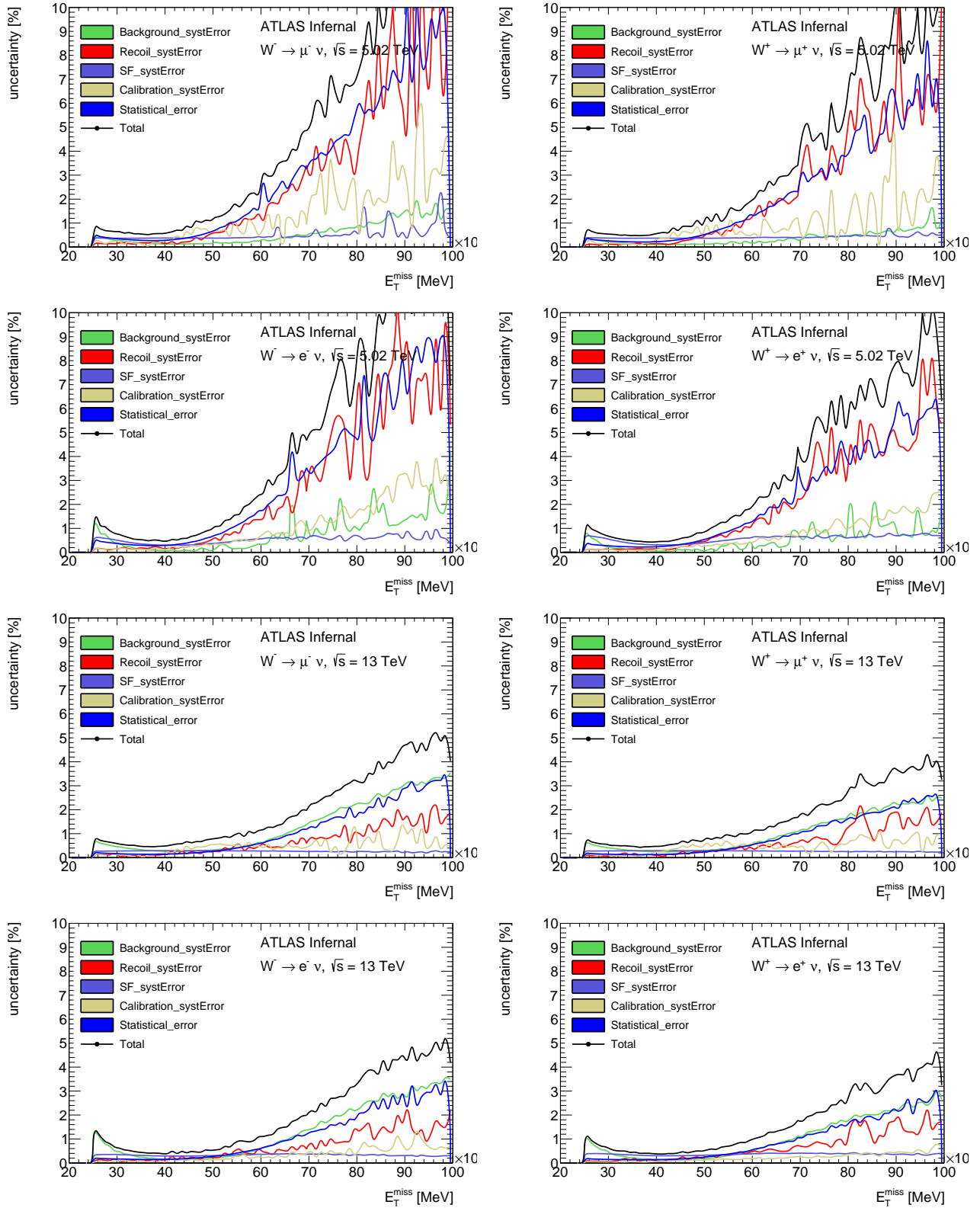**Figure 10.18:** $\vec{E}_T^{miss}$ systematic error breakdown in the muon and electron channel for the $\sqrt{s} = 5$ TeV and $\sqrt{s} = 13$ TeV datasets.

**Figure 10.19:** Transverse mass systematic error breakdown of the W boson in the muon and electron channel for the $\sqrt{s} = 5$ TeV and $\sqrt{s} = 13$ TeV datasets.

**Figure 10.20:** Lepton pseudorapidity systematic error breakdown in the muon and electron channel for the $\sqrt{s} = 5$ TeV and $\sqrt{s} = 13$ TeV datasets.

**Figure 10.21:** Lepton transverse systematic error breakdown distribution in the muon and electron channel for the $\sqrt{s} = 5$ TeV and $\sqrt{s} = 13$ TeV datasets.
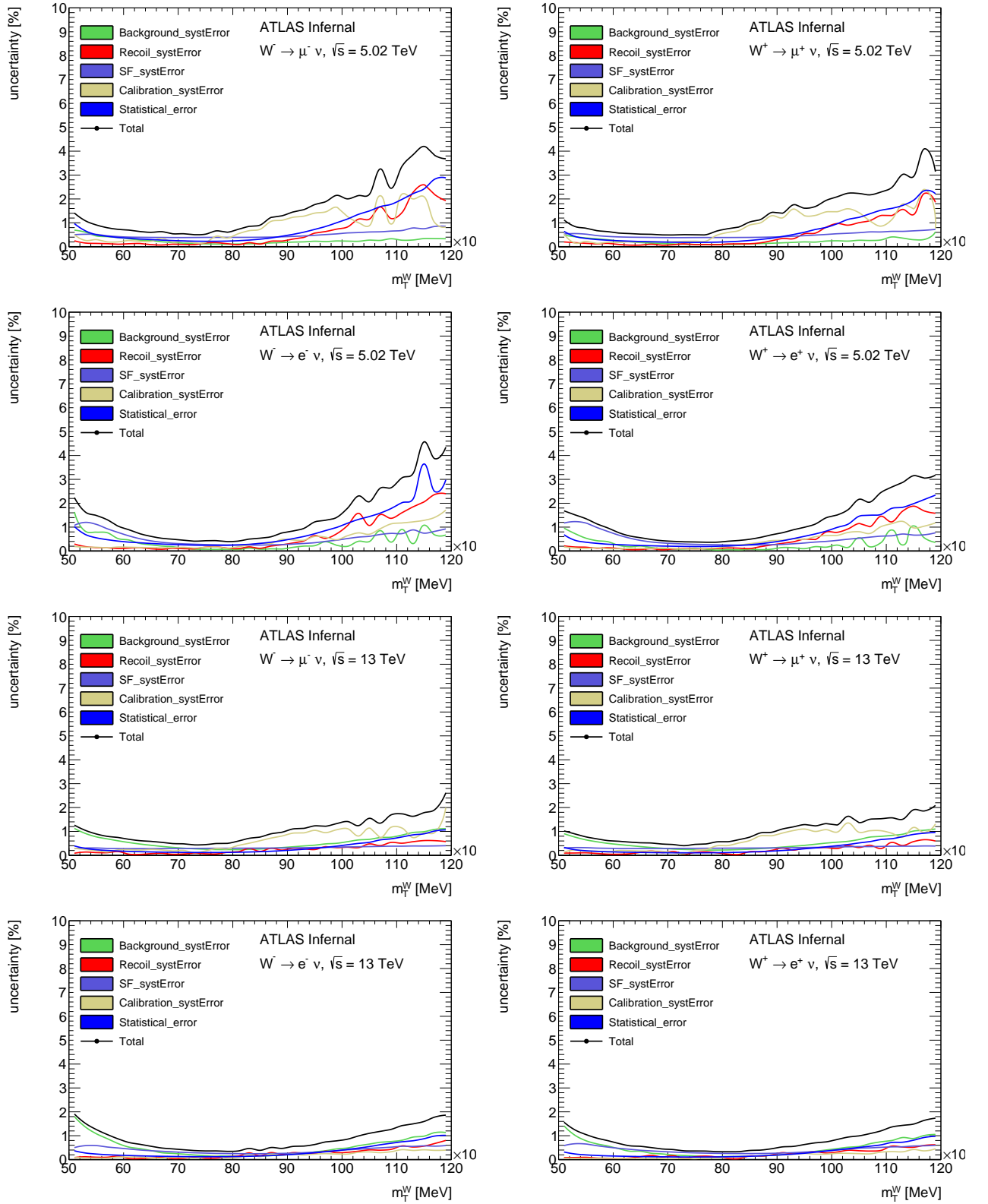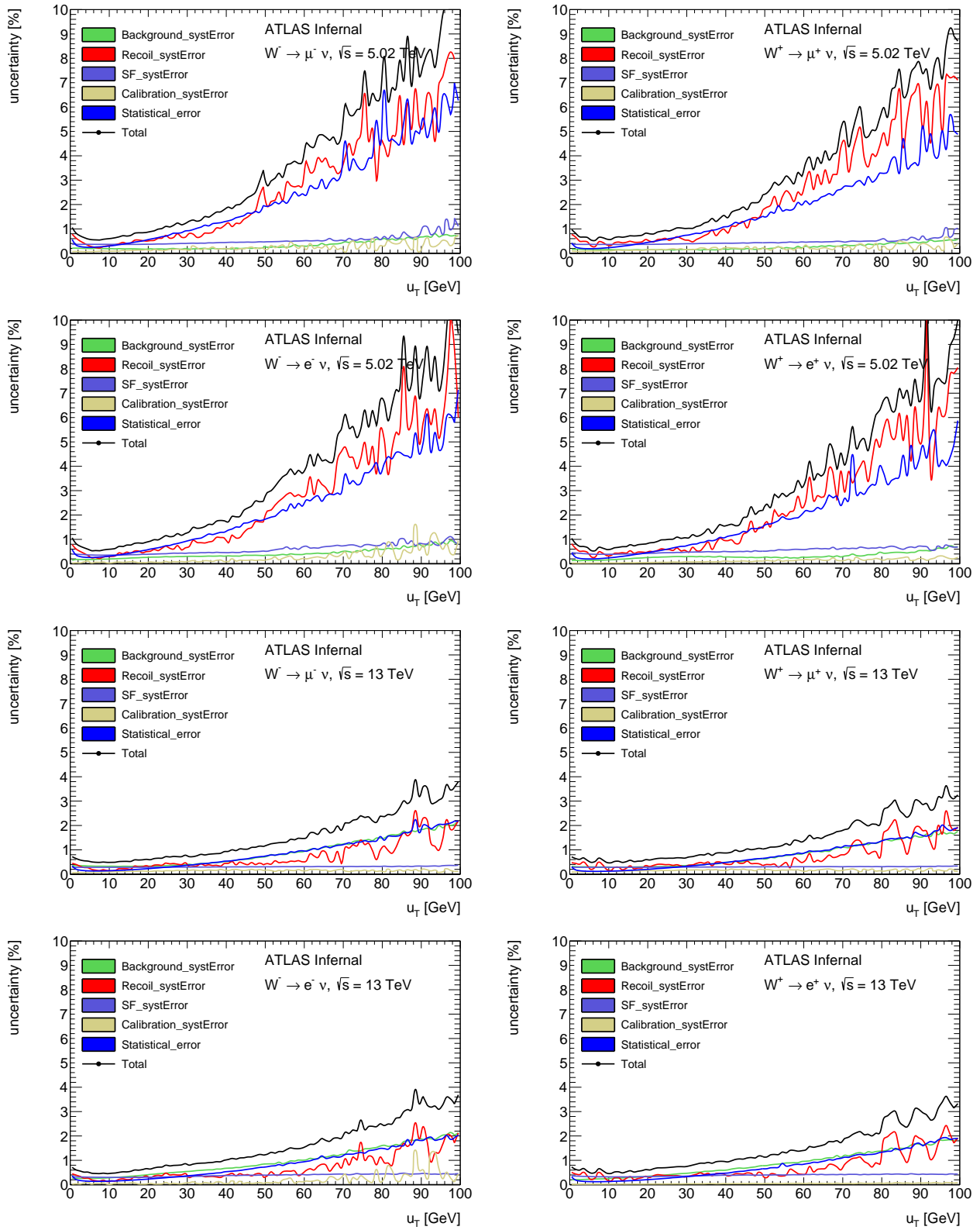
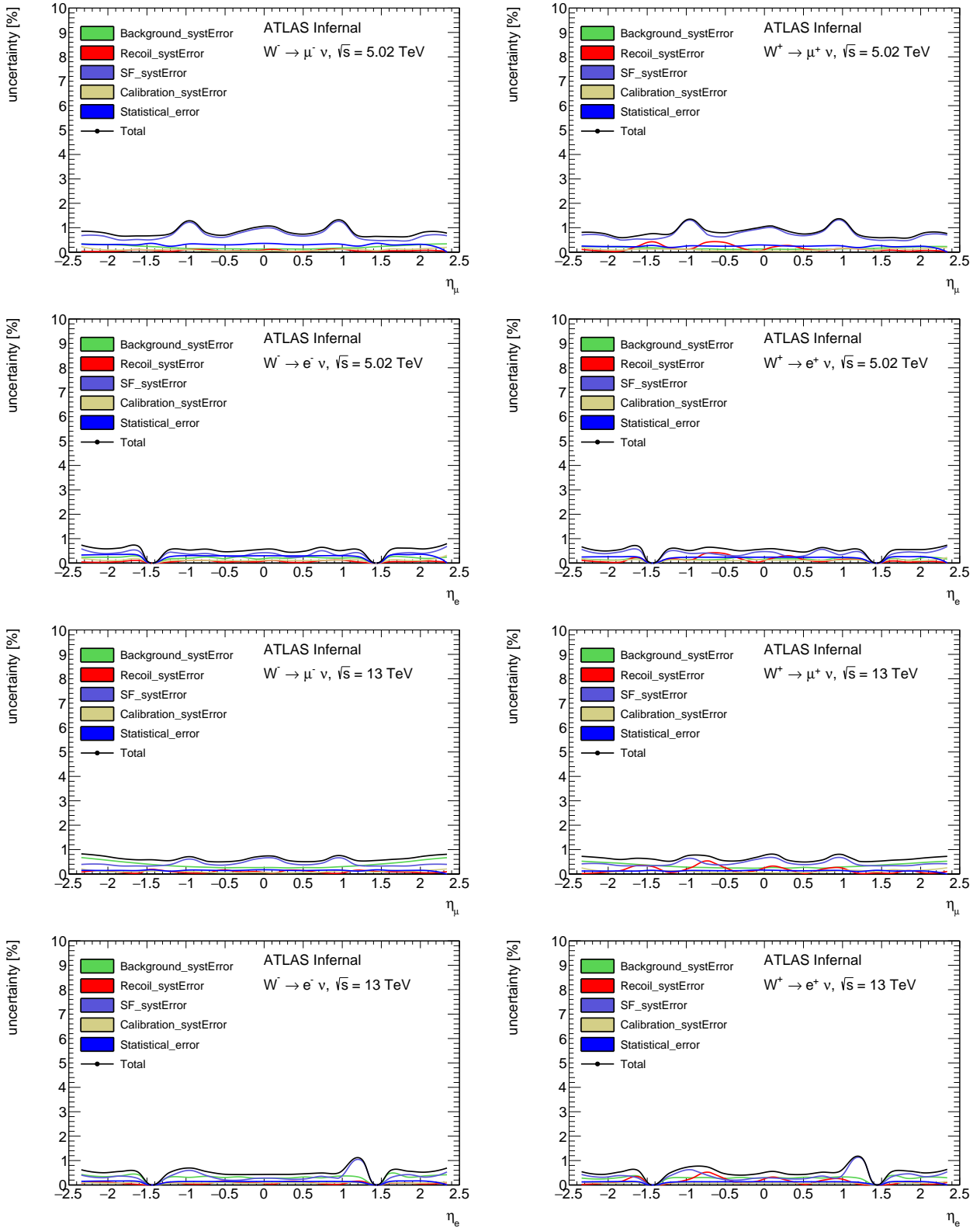**Figure 10.22:** W transverse momentum systematic error breakdown in the muon and electron channel for the $\sqrt{s} = 5$ TeV and $\sqrt{s} = 13$ TeV datasets.

# Appendix B: Hadronic recoil reconstruction using MVA
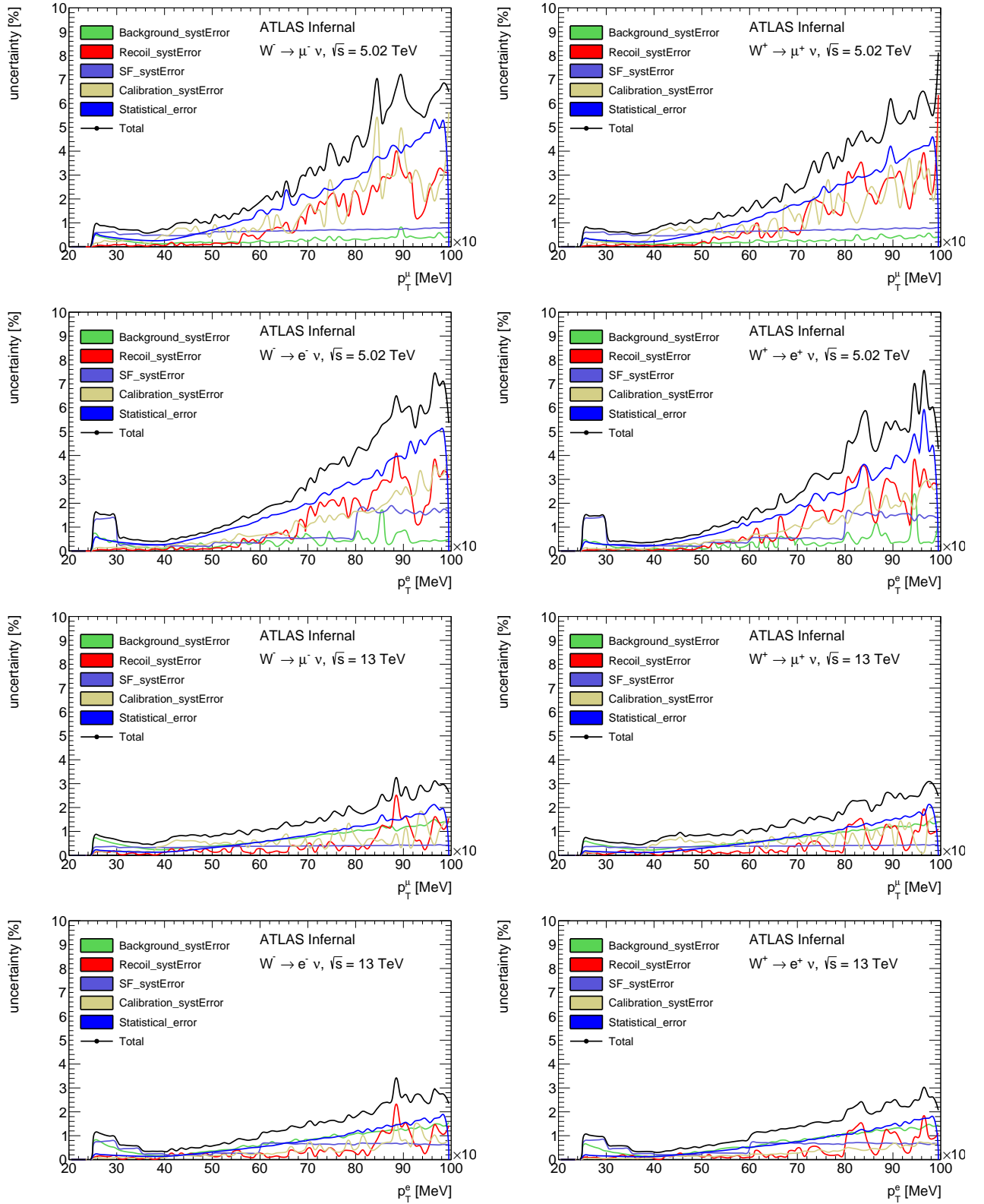
Current appendix contains the comparison of all the kinematic distributions for all W and Z channels.
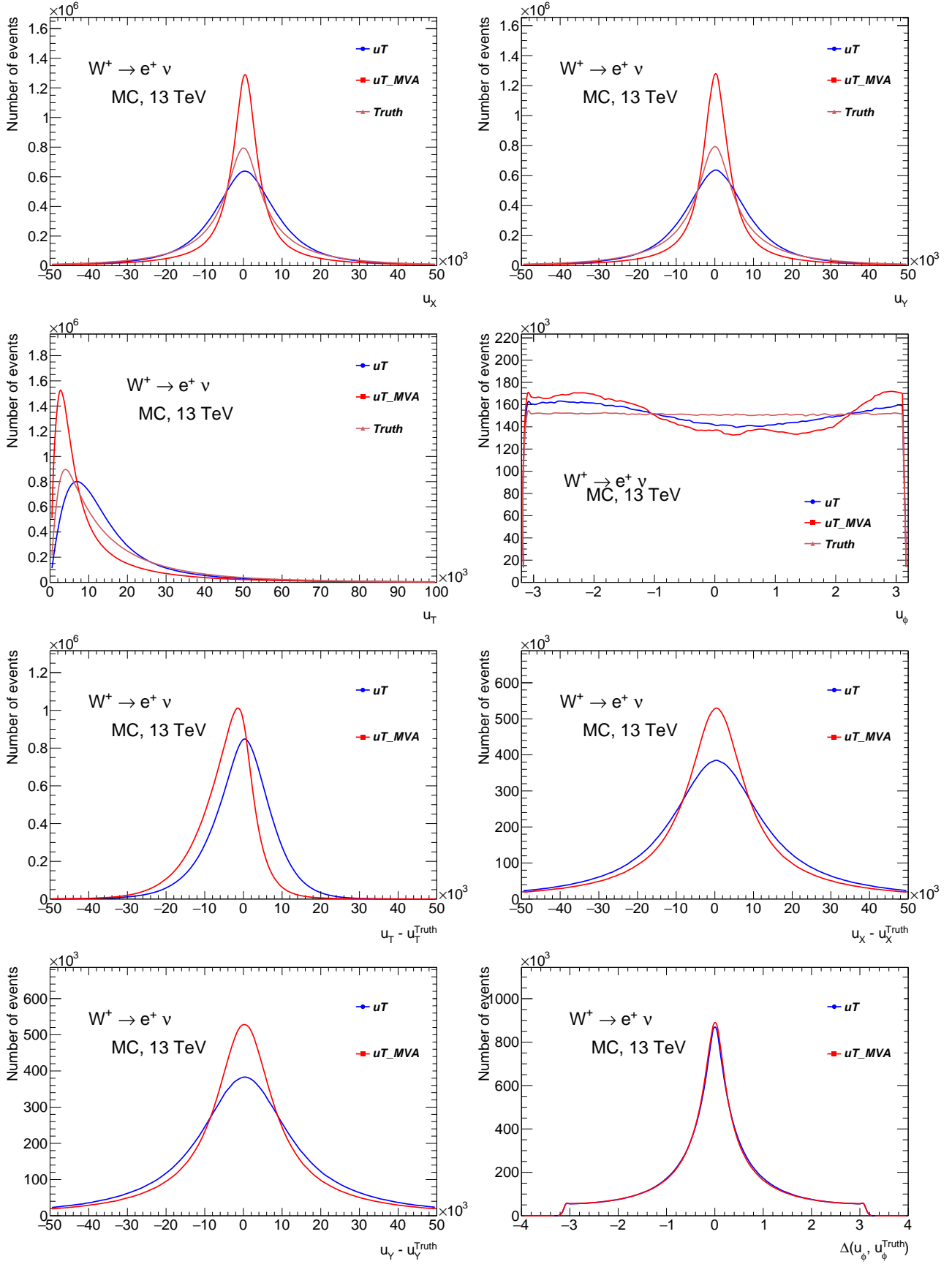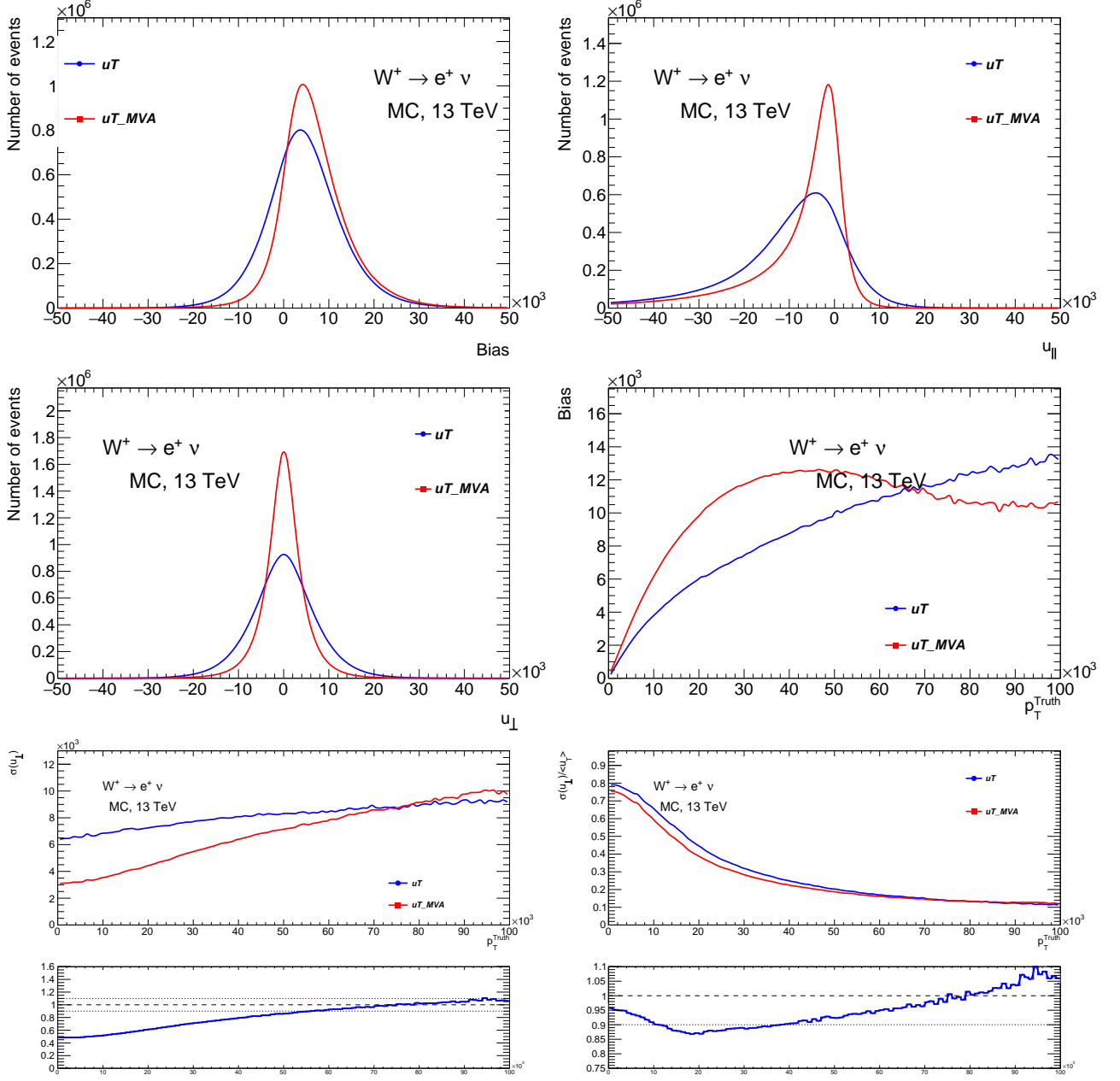
**Figure 10.23:** Comparison of kinematic distributions of $u_T$ vs $u_T^{MVA}$ for $W^+ \rightarrow e^+\nu$ data sample.
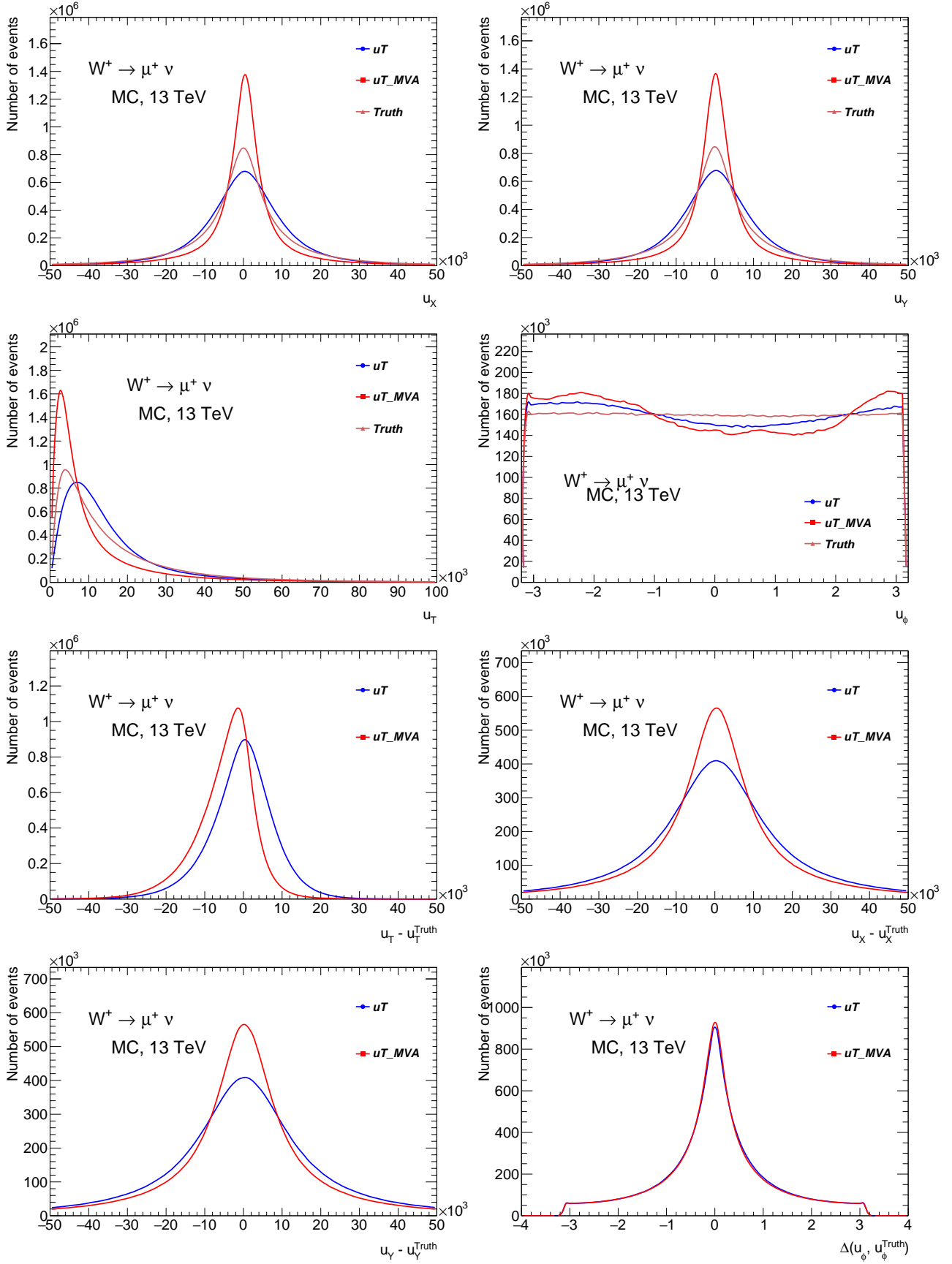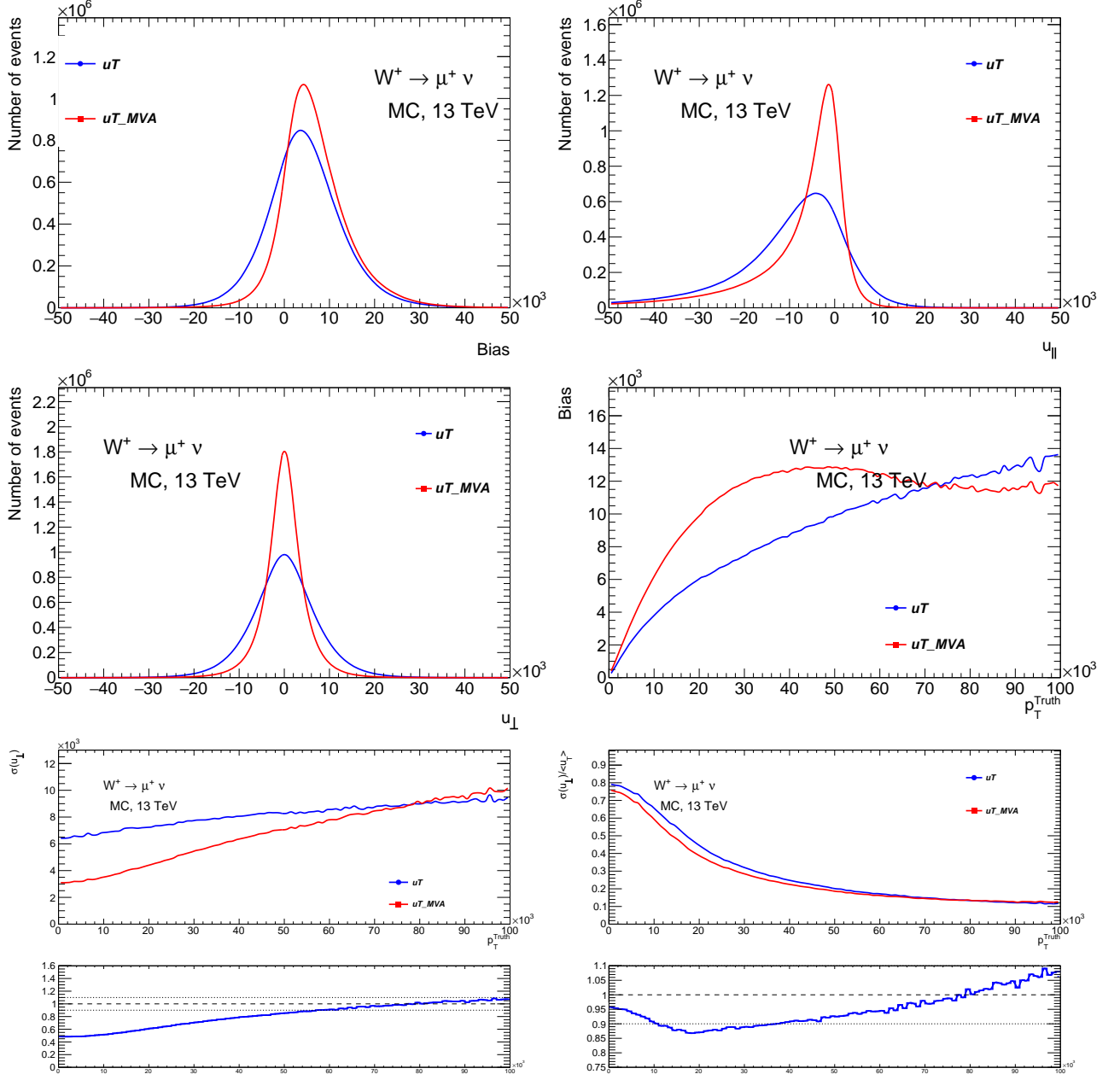
**Figure 10.24:** Comparison of kinematic distributions of $u_T$ vs $u_T^{MVA}$ for $W^- \to e^- \nu$ data sample.
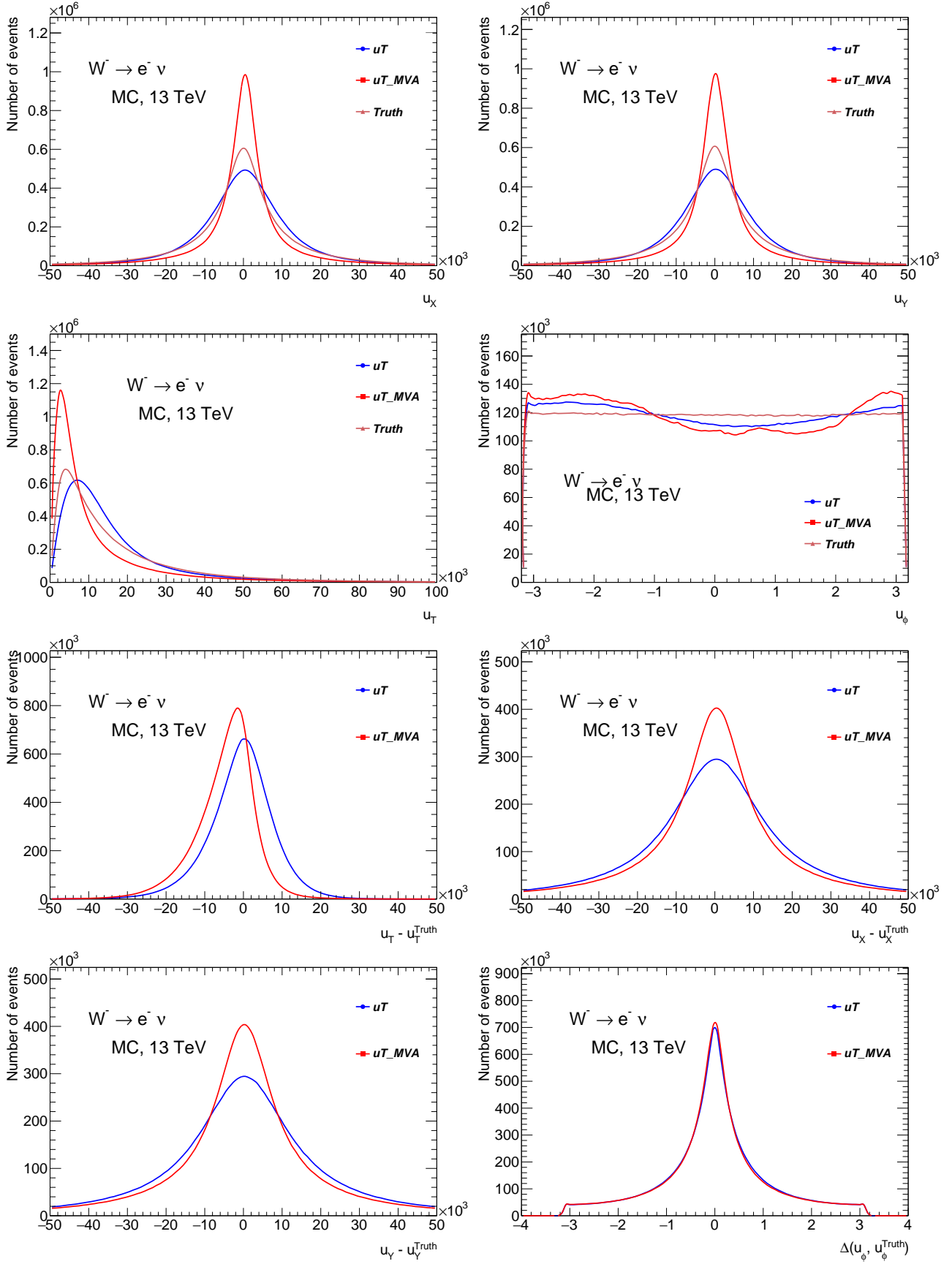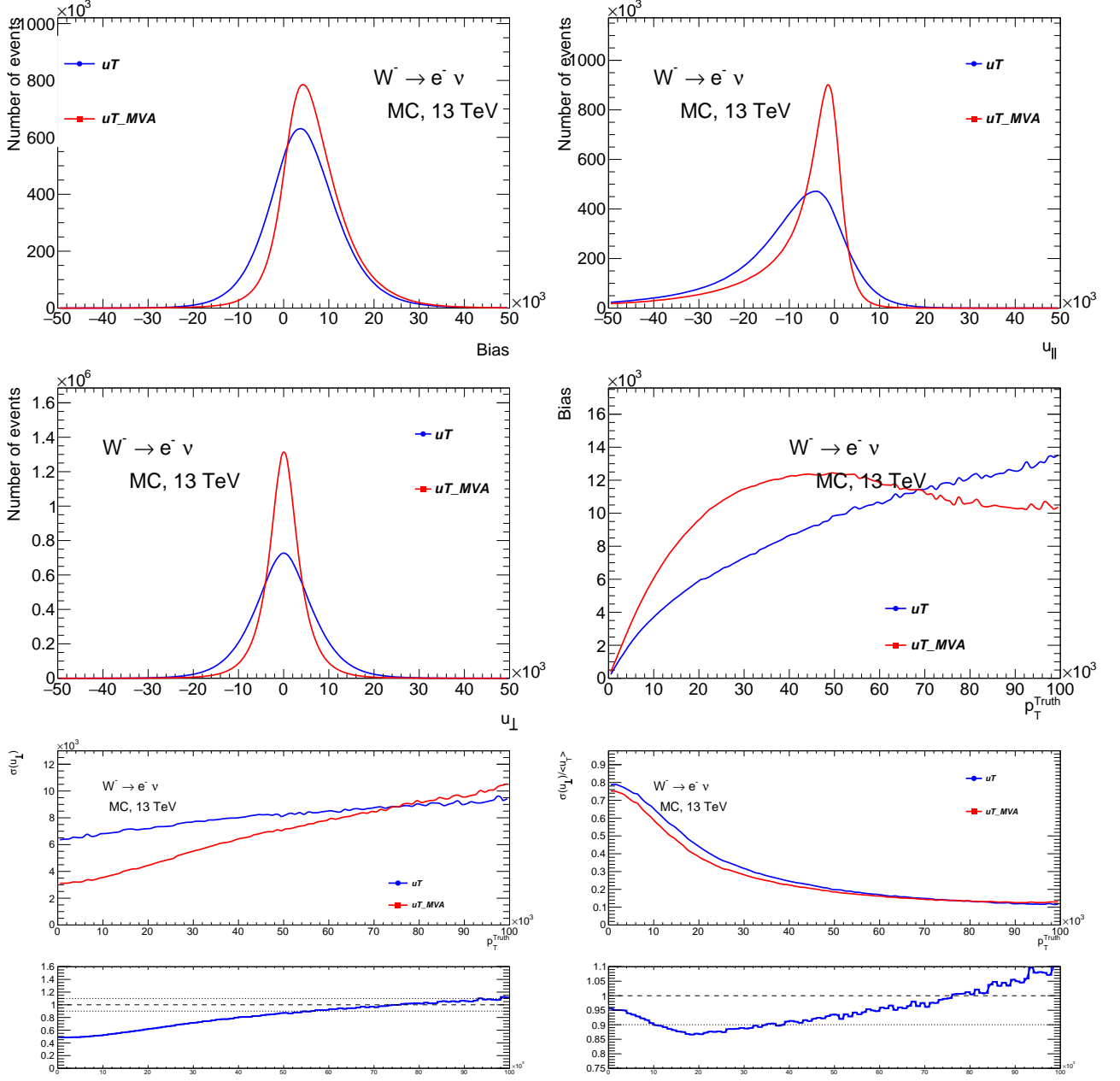
**Figure 10.25:** Comparison of kinematic distributions of $u_T$ vs $u_T^{MVA}$ for $W^+ \rightarrow \mu^+ \nu$ data sample.
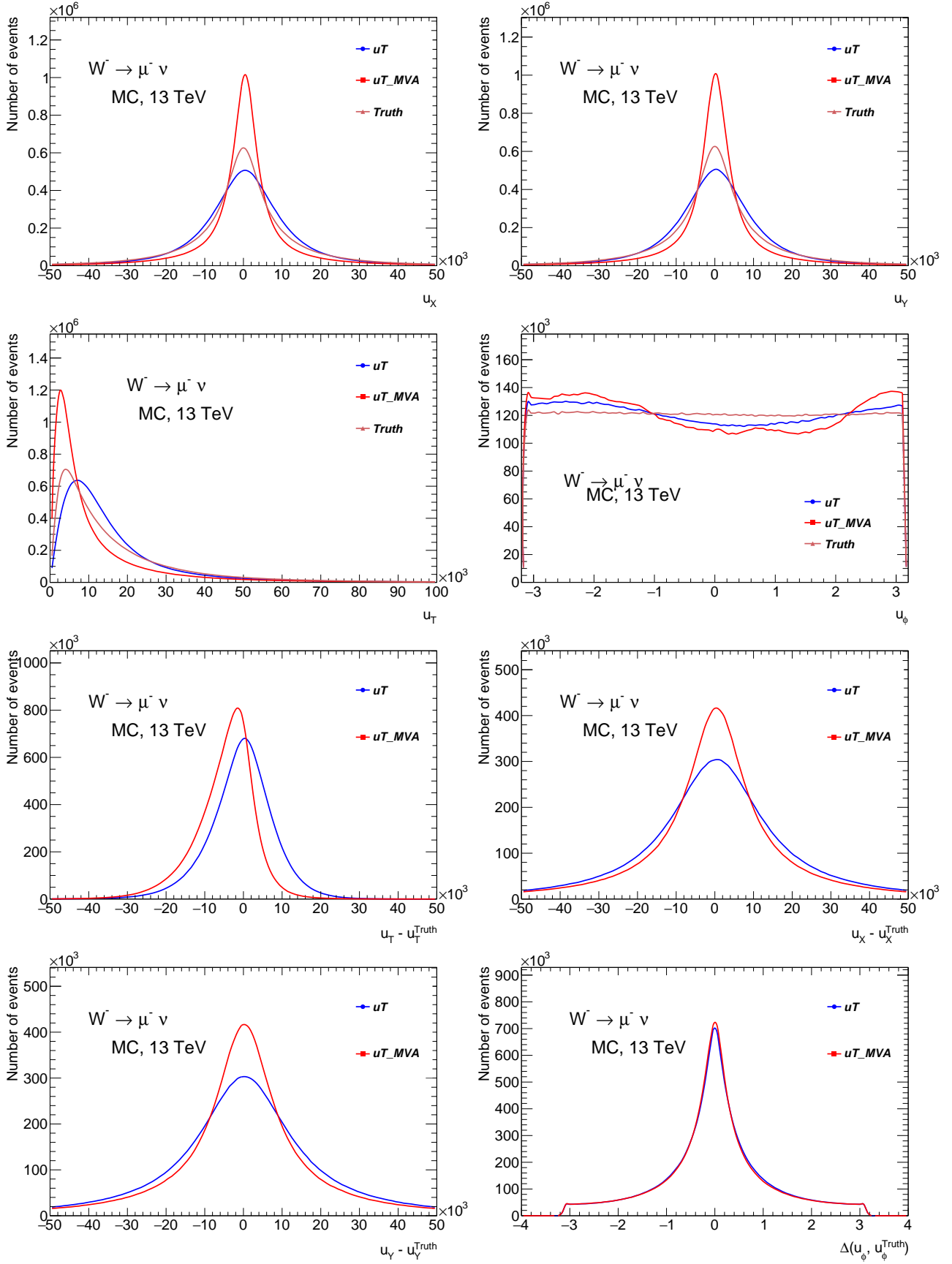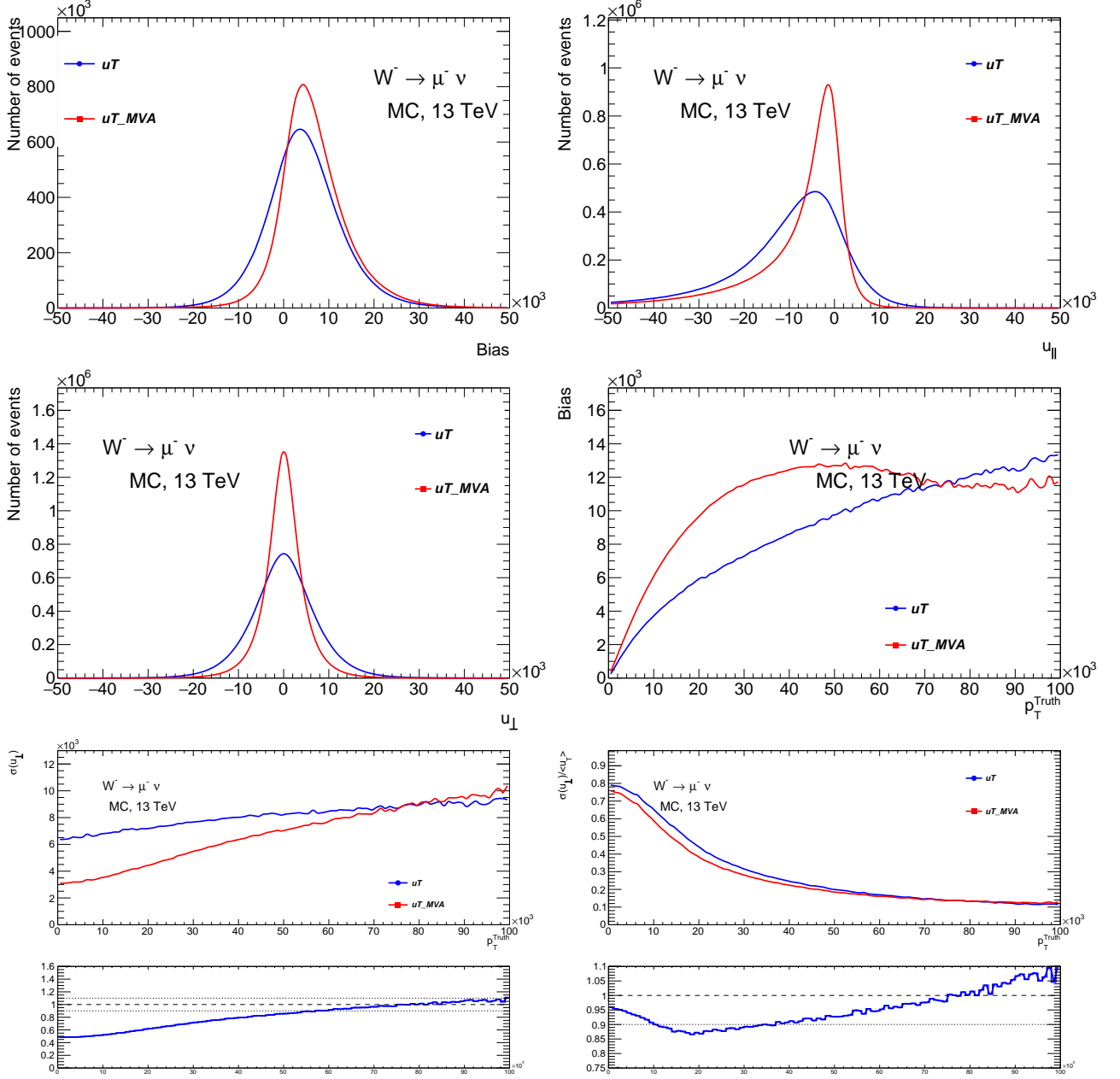
**Figure 10.26:** Comparison of kinematic distributions of $u_T$ vs $u_T^{MVA}$ for $W^- \to \mu^- \nu$ data sample.
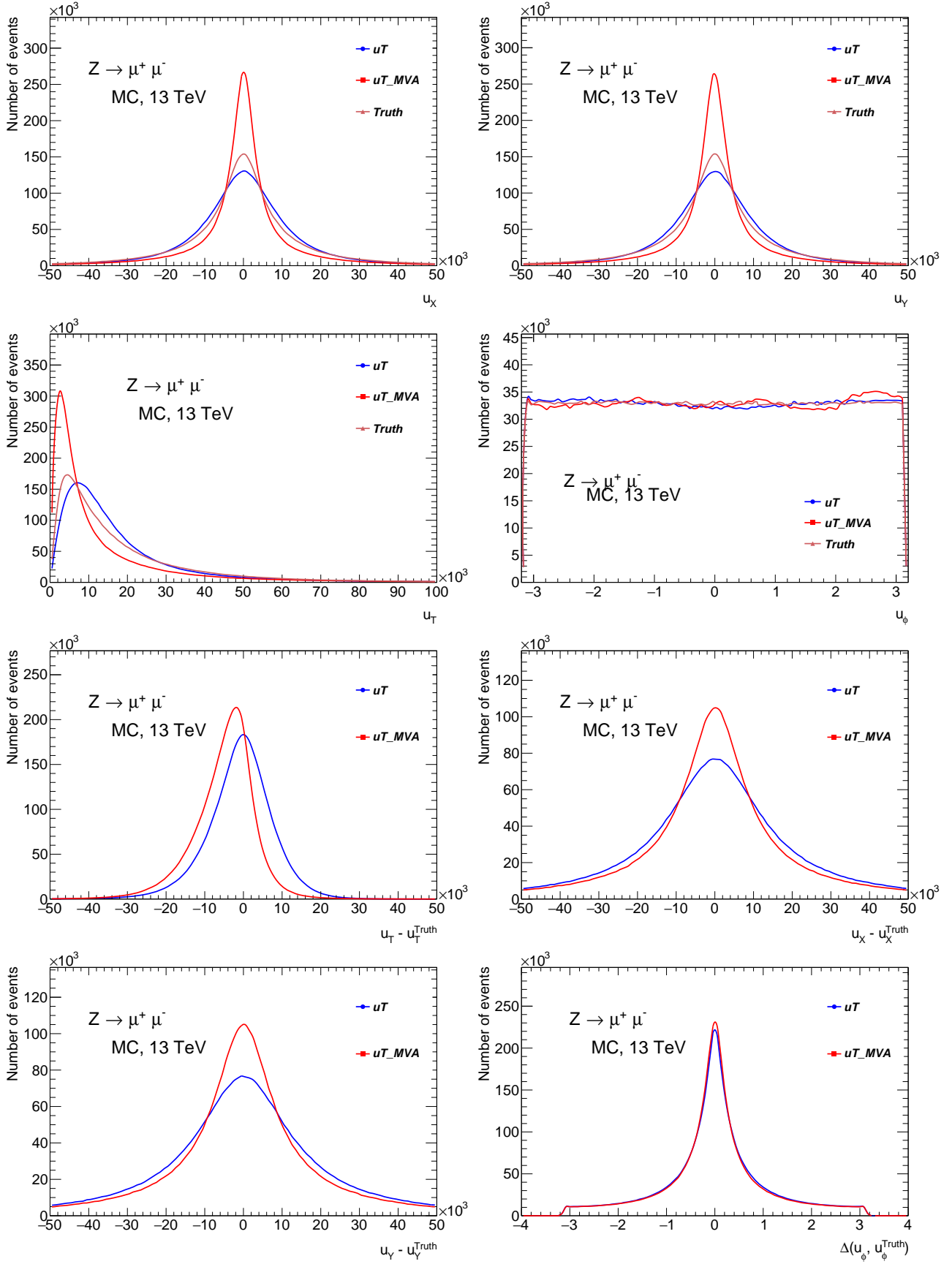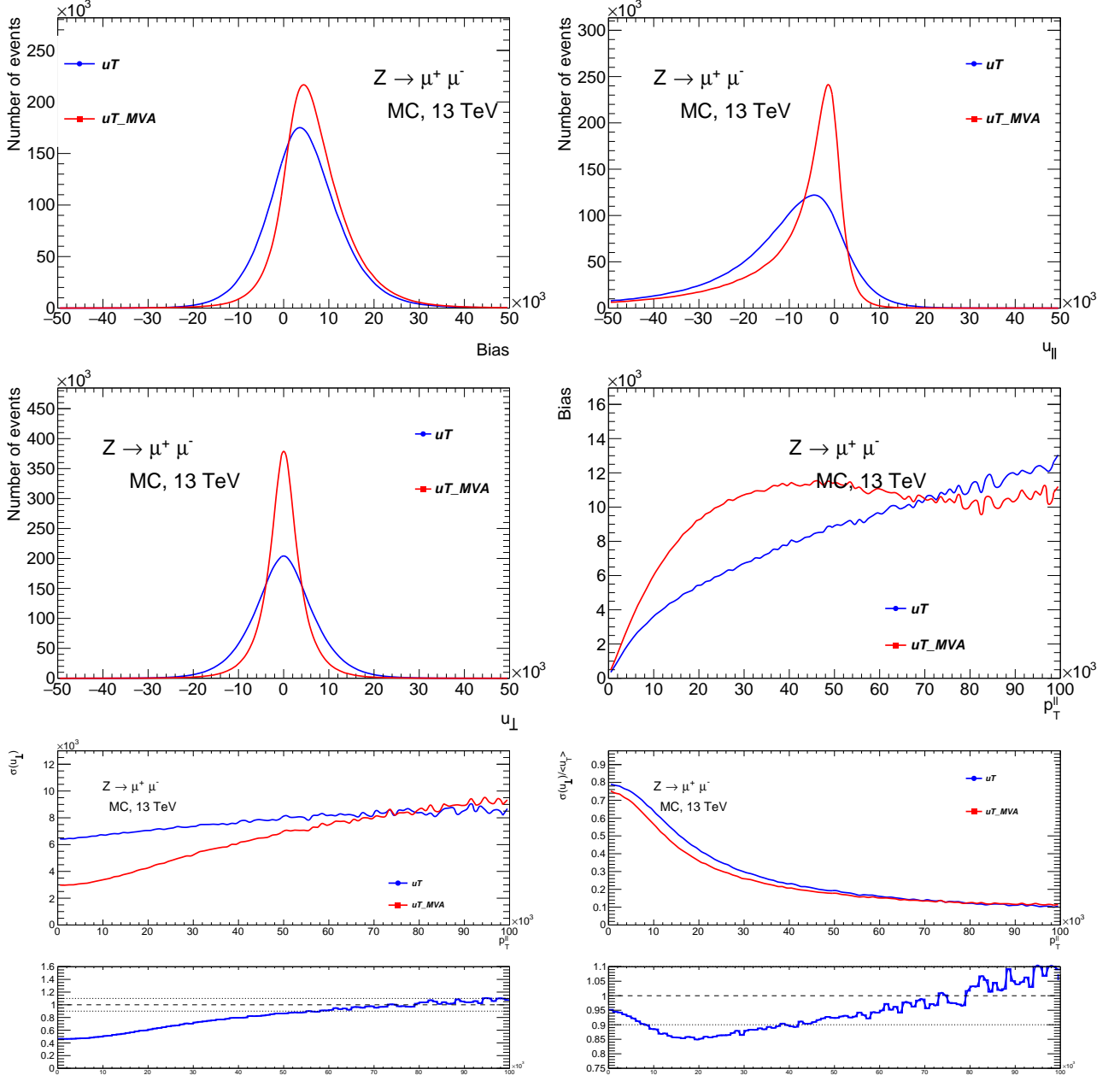
**Figure 10.27:** Comparison of kinematic distributions of $u_T$ vs $u_T^{MVA}$ for $Z \to \mu\mu$ MC sample.
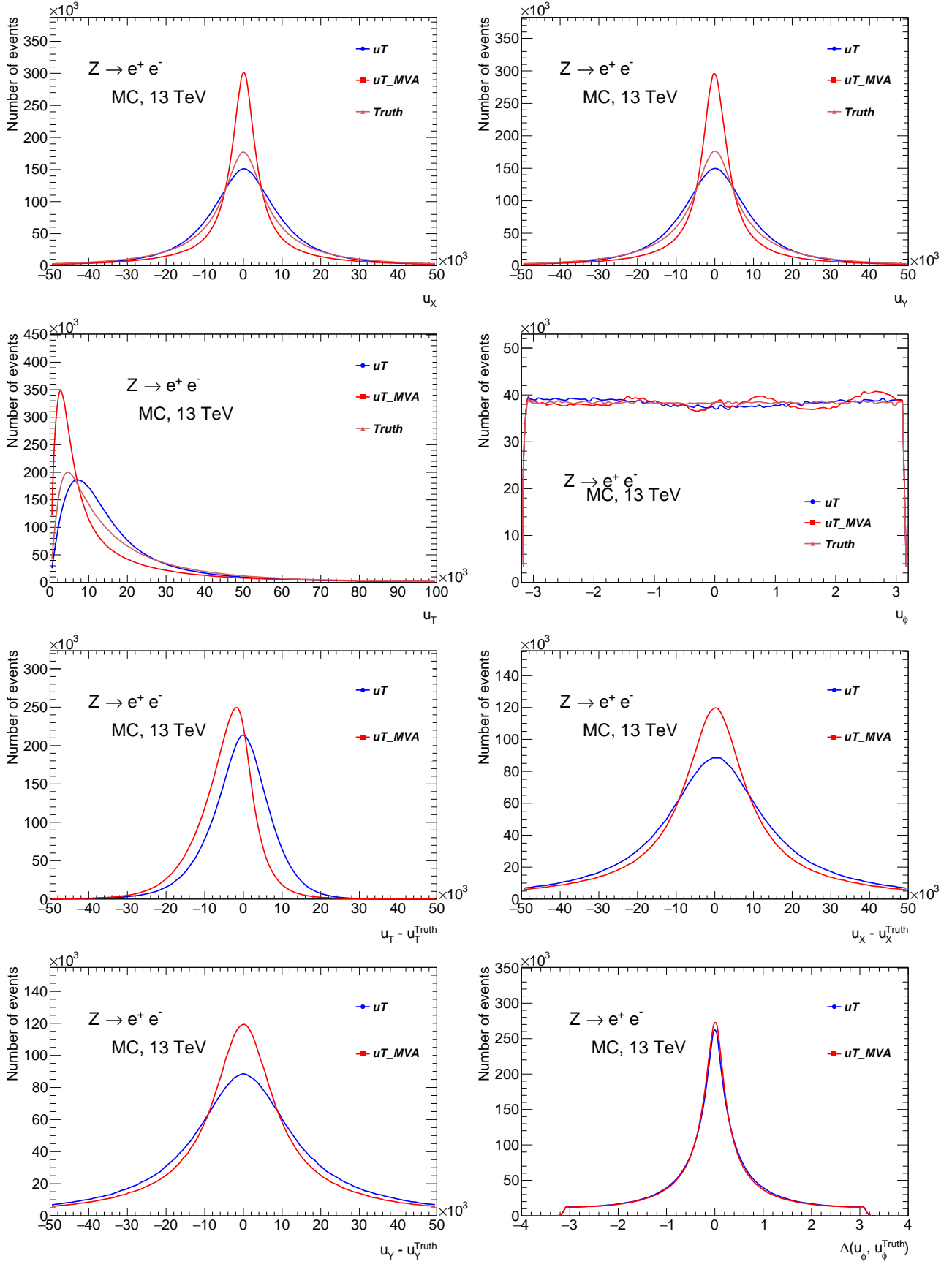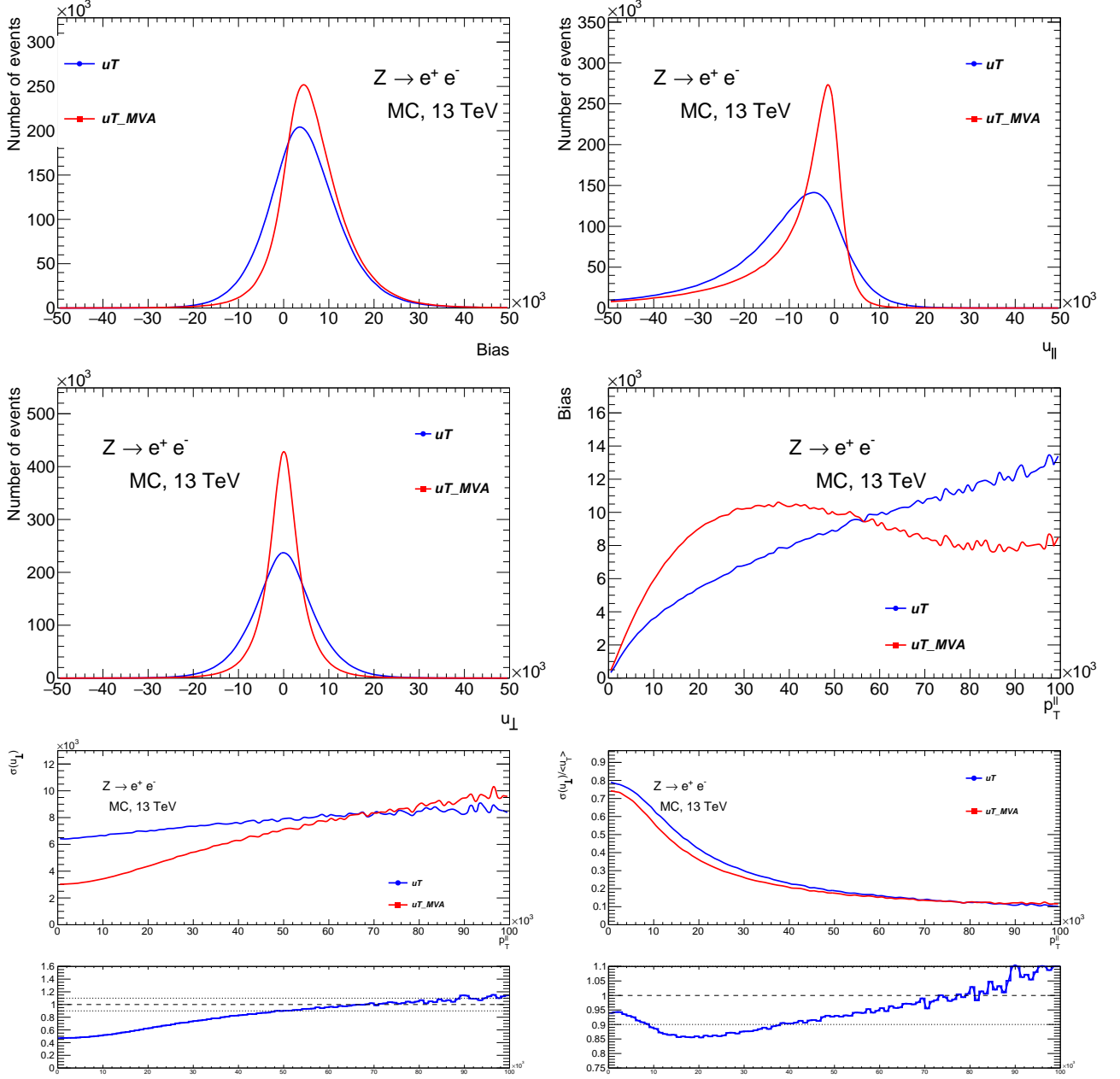
**Figure 10.28:** Comparison of kinematic distributions of $u_T$ vs $u_T^{MVA}$ for $Z \to ee$ MC sample.
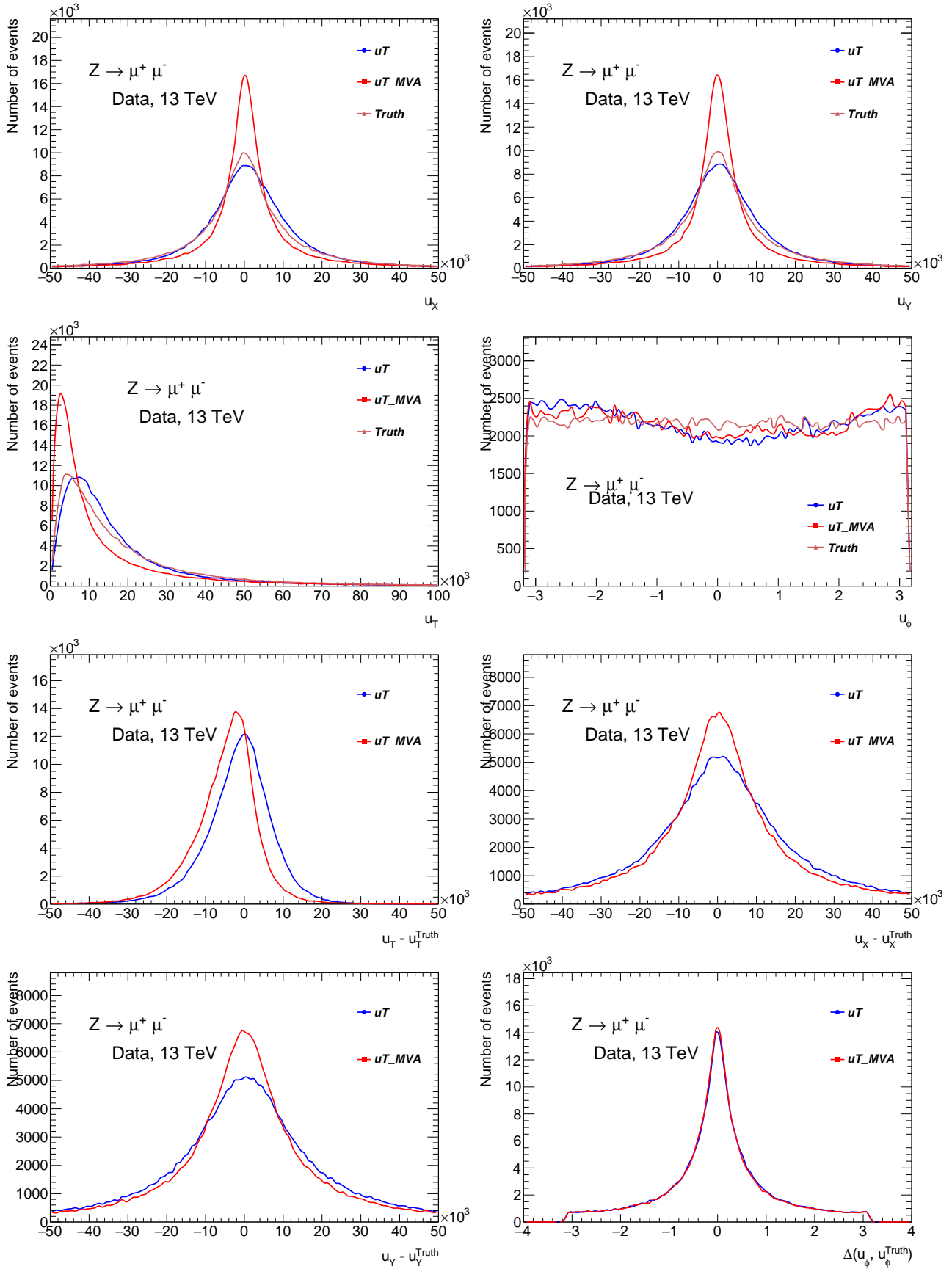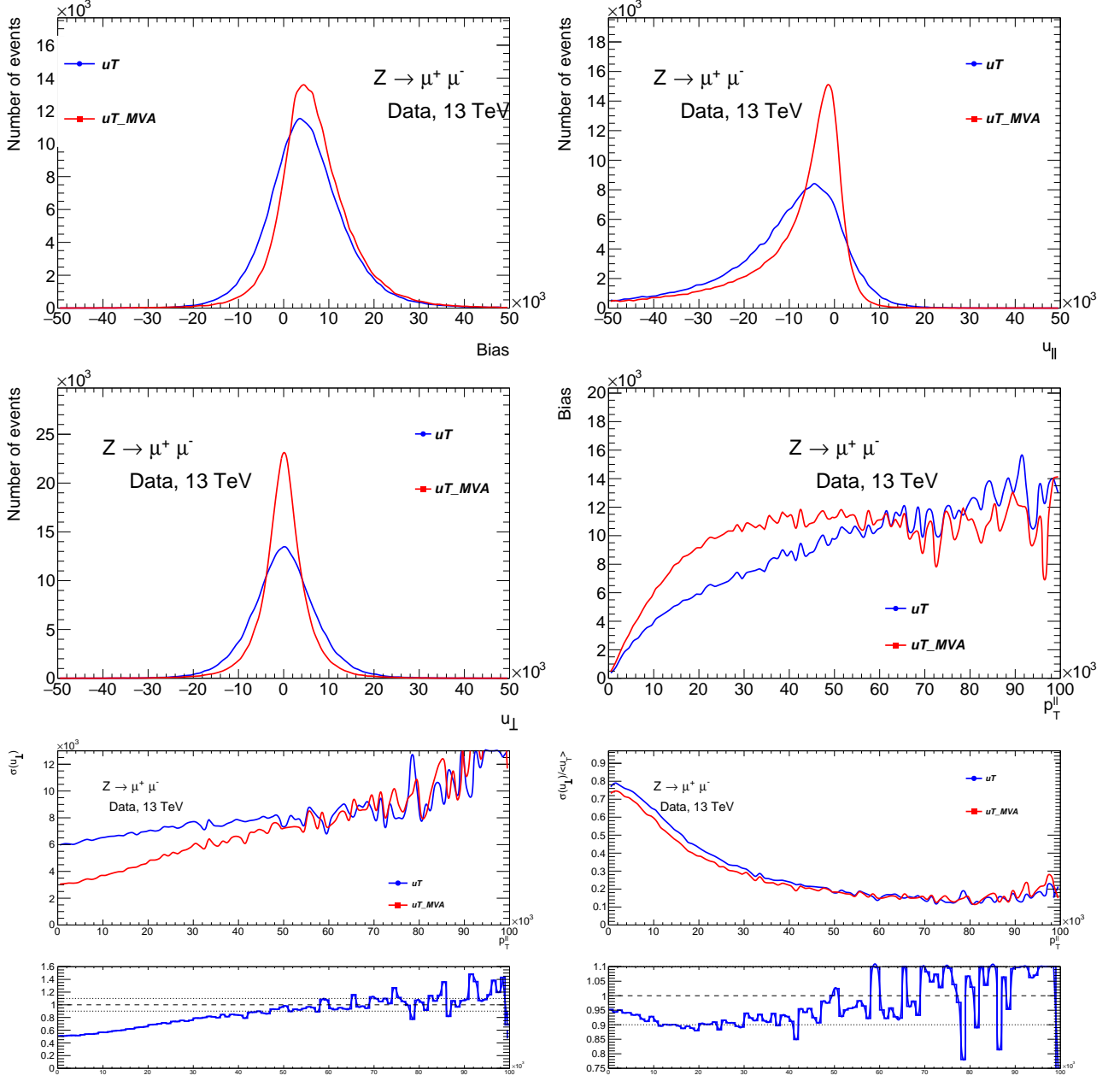
**Figure 10.29:** Comparison of kinematic distributions of $u_T$ vs $u_T^{MVA}$ for $Z \rightarrow \mu\mu$ data sample.
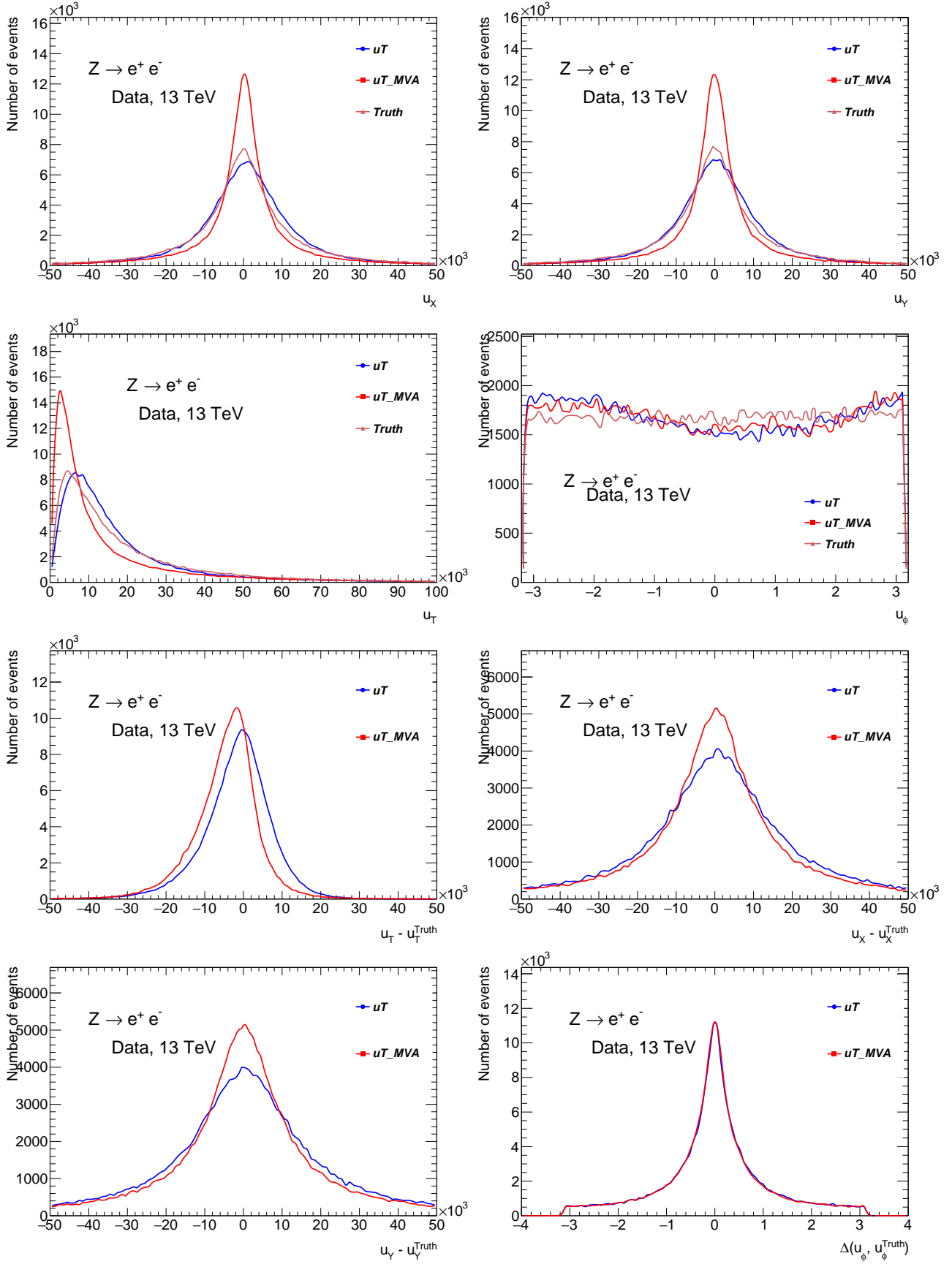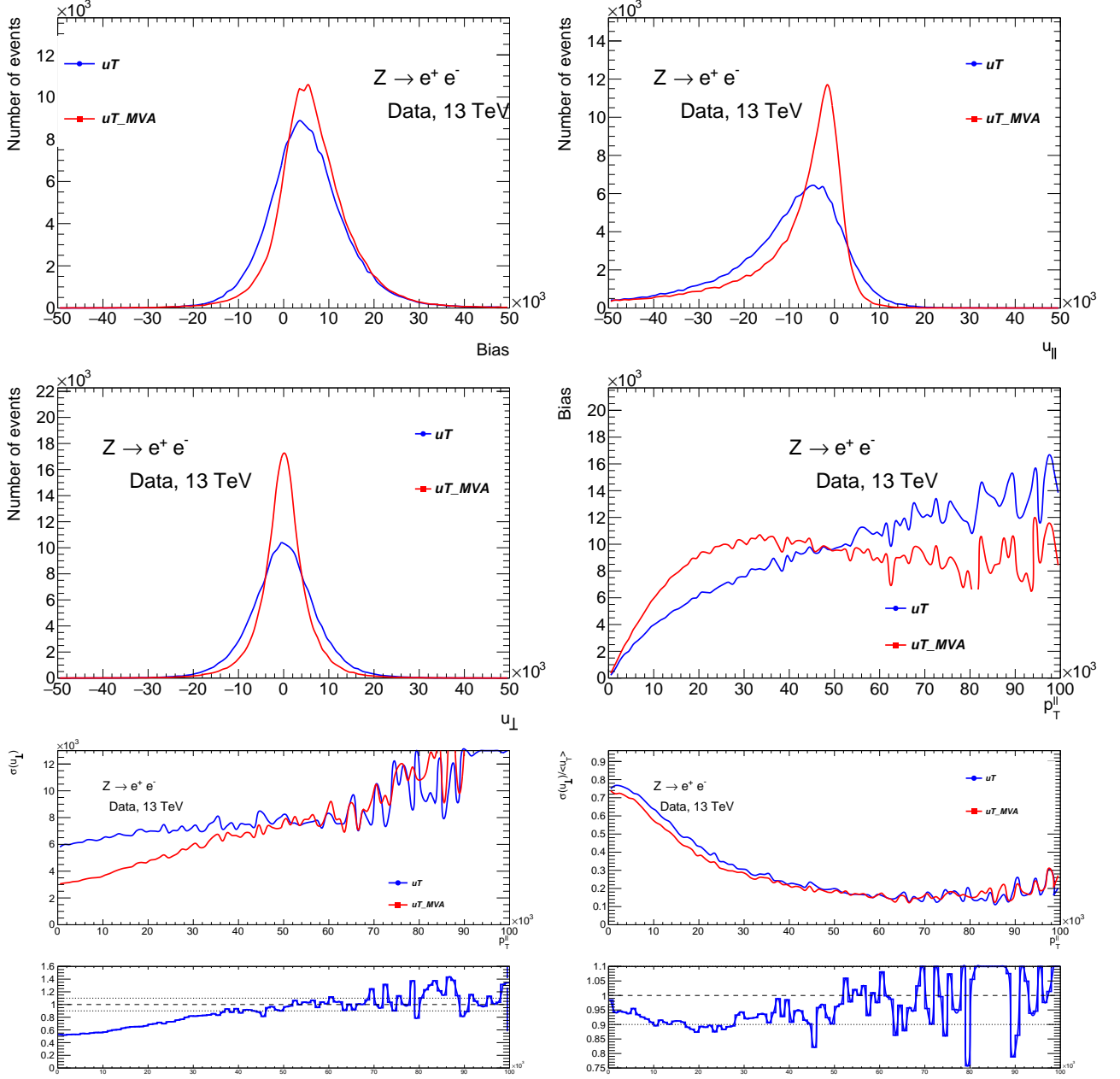
**Figure 10.30:** Comparison of kinematic distributions of $u_T$ vs $u_T^{MVA}$ for $Z \to ee$ data sample.

# Bibliography

[1]   Dmitri Mendeleev. *The principles of chemistry*. Saint Petersburg, 1869.

[2]   William Prout. "On the Relation between the Specific Gravities of Bodies in their Gaseous State and the Weights of their Atoms." In: *Annals of Philosophy* 6 (1815), pp. 321–330. URL: http://web.lemoyne.edu/~giunta/PROUT.HTML.

[3]   F. Soddy. "The Radio-Elements and the Periodic Law". In: *Nature* 91 (20 Mar. 1913), pp. 57–58. DOI: 10.1038/091057a0. URL: https://doi.org/10.1038/091057a0.

[4]   A. Einstein. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen". In: *Annalen der Physik* 322 (5 May 1905), pp. 549–560. DOI: 10.1002/andp.19053220806. URL: https://doi.org/10.1002/andp.19053220806.

[5]   J.J. Thompson. "Cathode Rays". In: *Philosophical Magazine* 44 (Oct. 1897). DOI: 10.1002/andp.19053220806. URL: https://doi.org/10.1002/andp.19053220806.

[6]   W. Wien. "Über positive Elektronen und die Existenz hoher Atomgewichte". In: *Annalen der Physik* 318.4 (1904), pp. 669–677. DOI: 10.1002/andp.18943180404. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.18943180404. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.18943180404.

[7]   Carl E. Moore, Bruno Jaselskis, and Alfred von Smolinski. "The proton". In: *Journal of Chemical Education* 62.10 (Oct. 1985), p. 859. ISSN: 0021-9584. DOI: 10.1021/ed062p859. URL: https://doi.org/10.1021/ed062p859.

[8]   J. Chadwick. "The existence of a neutron". In: *Proceedings of the Royal Society A.* 151 (June 1932), pp. 479–493. DOI: 10.1098/rspa.1932.0112. URL: https://doi.org/10.1098/rspa.1932.0112.

[9]   *A letter by W. Pauli.* URL: https://www.symmetrymagazine.org/article/march-2007/neutrino-invention.

[10]  C. D. Anderson. "The Positive Electron". In: *Physical Review* 46 (Mar. 1933), pp. 491–494. DOI: 10.1103/PhysRev.43.491. URL: https://doi.org/10.1103/PhysRev.43.491.

[11]   P. A. M. Dirac. "The quantum theory of the electron". In: *Proceedings of the Royal Society A*. 117 (Feb. 1928), pp. 610–624. DOI: 10.1098/rspa.1928.0023. URL: https://doi.org/10.1098/rspa.1928.0023.

[12]   Seth H. Neddermeyer and Carl D. Anderson. "Note on the Nature of Cosmic-Ray Particles". In: *Phys. Rev.* 51 (10 May 1937), pp. 884–886. DOI: 10.1103/PhysRev.51.884. URL: https://link.aps.org/doi/10.1103/PhysRev.51.884.

[13]   C. M. G. LATTES et al. "PROCESSES INVOLVING CHARGED MESONS". In: *Nature* 159.4047 (May 1947), pp. 694–697. ISSN: 1476-4687. DOI: 10.1038/159694a0. URL: https://doi.org/10.1038/159694a0.

[14]   G. D. ROCHESTER and C. C. BUTLER. "Evidence for the Existence of New Unstable Elementary Particles". In: *Nature* 160.4077 (Dec. 1947), pp. 855–857. ISSN: 1476-4687. DOI: 10.1038/160855a0. URL: https://doi.org/10.1038/160855a0.

[15]   V. D. Hopper and S. Biswas. "Evidence Concerning the Existence of the New Unstable Elementary Neutral Particle". In: *Phys. Rev.* 80 (6 Dec. 1950), pp. 1099–1100. DOI: 10.1103/PhysRev.80.1099. URL: https://link.aps.org/doi/10.1103/PhysRev.80.1099.

[16]   C. L. Cowan et al. "Detection of the Free Neutrino: a Confirmation". In: *Science* 124.3212 (1956), pp. 103–104. ISSN: 0036-8075. DOI: 10.1126/science.124.3212.103. eprint: https://science.sciencemag.org/content/124/3212/103.full.pdf. URL: https://science.sciencemag.org/content/124/3212/103.

[17]   Murray Gell-Mann. "A Schematic Model of Baryons and Mesons". In: *Phys. Lett.* 8 (1964), pp. 214–215. DOI: 10.1016/S0031-9163(64)92001-3.

[18]   G Zweig. "An $SU_3$ model for strong interaction symmetry and its breaking; Version 2". In: CERN-TH-412 (Feb. 1964). Version 1 is CERN preprint 8182/TH.401, Jan. 17, 1964, 80 p. URL: http://cds.cern.ch/record/570209.

[19]   V. E. Barnes et al. "Observation of a Hyperon with Strangeness Minus Three". In: *Phys. Rev. Lett.* 12 (8 Feb. 1964), pp. 204–206. DOI: 10.1103/PhysRevLett.12.204. URL: https://link.aps.org/doi/10.1103/PhysRevLett.12.204.

[20]   M et al. Aaboud. "Measurement of the $W$-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 78.CERN-EP-2016-305 (Jan. 2017). 86 pages in total, author list starting page 70, 30 figures, 13 tables, submitted to EPJC, All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/STDM-2014-18, 110. 86 p. DOI: 10.1140/epjc/s10052-017-5475-4. URL: https://cds.cern.ch/record/2242923.

[21]   Michael E Peskin and Daniel V Schroeder. *An introduction to quantum field theory*. Includes exercises. Boulder, CO: Westview, 1995. URL: https://cds.cern.ch/record/257493.

[22]   D.V. Shirkov N.N. Bogolyubov. *Quantum fields*. Addison-Wesley, 1982.

[23] Mark Srednicki. *Quantum Field Theory*. Cambridge: Cambridge Univ. Press, 2007. URL: `https://cds.cern.ch/record/1019751`.

[24] V.B. Berestetskii. A.I. Akhiezer. *Quantum Electrodynamics*. Interscience Publishers, 1965.

[25] Steven Weinberg. *The Quantum Theory of Fields, Volume 1: Foundations*. Cambridge University Press, 2005. ISBN: 0521670535.

[26] David J Griffiths. *Introduction to elementary particles; 2nd rev. version*. Physics textbook. New York, NY: Wiley, 2008. URL: `https://cds.cern.ch/record/111880`.

[27] W Pauli. "Über den EinfluSS der Geschwindigkeitsabhängigkeit der Elektronenmasse auf den Zeemaneffekt". In: *Z. Phys.* 31 (1925), pp. 373–385.

[28] J. Beringer et al. "Review of Particle Physics". In: *Phys. Rev. D* 86 (1 July 2012), p. 010001. DOI: `10.1103/PhysRevD.86.010001`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.86.010001`.

[29] *Standard Model particles*. URL: `https://en.wikipedia.org/wiki/Standard_Model`.

[30] Julian Schwinger. "Quantum Electrodynamics. II. Vacuum Polarization and Self-Energy". In: *Phys. Rev.* 75 (4 Feb. 1949), pp. 651–679. DOI: `10.1103/PhysRev.75.651`. URL: `https://link.aps.org/doi/10.1103/PhysRev.75.651`.

[31] Sheldon L. Glashow. "The renormalizability of vector meson interactions". In: *Nucl. Phys.* 10 (1959), pp. 107–117. DOI: `10.1016/0029-5582(59)90196-8`.

[32] *Coupling constants*. URL: `https://commons.wikimedia.org/wiki/File:Running_coupling_constants.svg`.

[33] David J. Gross and Frank Wilczek. "Ultraviolet Behavior of Non-Abelian Gauge Theories". In: *Phys. Rev. Lett.* 30 (26 June 1973), pp. 1343–1346. DOI: `10.1103/PhysRevLett.30.1343`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.30.1343`.

[34] H. David Politzer. "Reliable Perturbative Results for Strong Interactions?" In: *Phys. Rev. Lett.* 30 (26 June 1973), pp. 1346–1349. DOI: `10.1103/PhysRevLett.30.1346`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.30.1346`.

[35] V. S. Vanyashin and M. V. Terent'ev. "The Vacuum Polarization of a Charged Vector Field". In: *Soviet Journal of Experimental and Theoretical Physics* 21 (Aug. 1965), p. 375.

[36] Albert Einstein. "Zur Elektrodynamik bewegter Körper. (German) [On the electrodynamics of moving bodies]". In: *Annalen der Physik* 322.10 (1905), pp. 891–921. DOI: `http://dx.doi.org/10.1002/andp.19053221004`.

[37] M. H. Poincare. "Sur la dynamique de l'electron". In: *Rendiconti del Circolo Matematico di Palermo (1884-1940)* 21.1 (Dec. 1906), pp. 129–175. ISSN: 0009-725X. DOI: `10.1007/BF03013466`. URL: `https://doi.org/10.1007/BF03013466`.

[38] H. A. Lorentz. "Simplified Theory of Electrical and Optical Phenomena in Moving Systems". In: *Koninklijke Nederlandse Akademie van Wetenschappen Proceedings Series B Physical Sciences* 1 (Jan. 1898), pp. 427–442.

[39] E. Noether. "Invariante Variationsprobleme". ger. In: *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1918 (1918), pp. 235–257. URL: http://eudml.org/doc/59024.

[40] C. N. Yang and R. L. Mills. "Conservation of Isotopic Spin and Isotopic Gauge Invariance". In: *Phys. Rev.* 96 (1 Oct. 1954), pp. 191–195. DOI: 10.1103/PhysRev.96.191. URL: https://link.aps.org/doi/10.1103/PhysRev.96.191.

[41] J Schwinger. "Quantum Electrodynamics. I. A Covariant Formulation". In: *Physical Review (U.S.) Superseded in part by Phys. Rev. A, Phys. Rev. B: Solid State, Phys. Rev. C, and Phys. Rev. D* 74 (Nov. 1948). DOI: 10.1103/PhysRev.74.1439.

[42] R. P. Feynman. "Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction". In: *Phys. Rev.* 80 (3 Nov. 1950), pp. 440–457. DOI: 10.1103/PhysRev.80.440. URL: https://link.aps.org/doi/10.1103/PhysRev.80.440.

[43] R. P. Feynman. "The Theory of Positrons". In: *Phys. Rev.* 76 (6 Sept. 1949), pp. 749–759. DOI: 10.1103/PhysRev.76.749. URL: https://link.aps.org/doi/10.1103/PhysRev.76.749.

[44] R. P. Feynman. "Space-Time Approach to Quantum Electrodynamics". In: *Phys. Rev.* 76 (6 Sept. 1949), pp. 769–789. DOI: 10.1103/PhysRev.76.769. URL: https://link.aps.org/doi/10.1103/PhysRev.76.769.

[45] S. Tomonaga. "On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields*". In: *Progress of Theoretical Physics* 1.2 (Aug. 1946), pp. 27–42. ISSN: 0033-068X. DOI: 10.1143/PTP.1.27. eprint: https://academic.oup.com/ptp/article-pdf/1/2/27/24027031/1-2-27.pdf. URL: https://doi.org/10.1143/PTP.1.27.

[46] F. J. Dyson. "The Radiation Theories of Tomonaga, Schwinger, and Feynman". In: *Phys. Rev.* 75 (3 Feb. 1949), pp. 486–502. DOI: 10.1103/PhysRev.75.486. URL: https://link.aps.org/doi/10.1103/PhysRev.75.486.

[47] F. J. Dyson. "The *S* Matrix in Quantum Electrodynamics". In: *Phys. Rev.* 75 (11 June 1949), pp. 1736–1755. DOI: 10.1103/PhysRev.75.1736. URL: https://link.aps.org/doi/10.1103/PhysRev.75.1736.

[48] A. Akhieser, L. Landau, and I. Pomeranchuk. "Scattering of Light by Light". In: *Nature* 138.3483 (Aug. 1936), pp. 206–206. ISSN: 1476-4687. DOI: 10.1038/138206a0. URL: https://doi.org/10.1038/138206a0.

[49] G. et al. Aad. "Observation of Light-by-Light Scattering in Ultraperipheral Pb + Pb Collisions with the ATLAS Detector". In: *Phys. Rev. Lett.* 123 (5 July 2019), p. 052001. DOI: 10.1103/PhysRevLett.123.052001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.123.052001.

[50]   A. Salam and J. C. Ward. "Weak and electromagnetic interactions". In: *Il Nuovo Cimento (1955-1965)* 11.4 (Feb. 1959), pp. 568–577. ISSN: 1827-6121. DOI: 10.1007/BF02726525. URL: https://doi.org/10.1007/BF02726525.

[51]   F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 Aug. 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[52]   Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[53]   Wolfgang Hollik. "Electroweak Theory". In: *Journal of Physics: Conference Series* 53 (Nov. 2006), pp. 7–43. DOI: 10.1088/1742-6596/53/1/002. URL: https://doi.org/10.1088%2F1742-6596%2F53%2F1%2F002.

[54]   Steven Weinberg. "Physical Processes in a Convergent Theory of the Weak and Electromagnetic Interactions". In: *Phys. Rev. Lett.* 27 (24 Dec. 1971), pp. 1688–1691. DOI: 10.1103/PhysRevLett.27.1688. URL: https://link.aps.org/doi/10.1103/PhysRevLett.27.1688.

[55]   *Standard Model particles*. URL: https://en.wikipedia.org/wiki/Color_confinement.

[56]   M. Awramik et al. "Precise prediction for the W boson mass in the standard model". In: *Phys. Rev. D* 69 (2004), p. 053006. DOI: 10.1103/PhysRevD.69.053006. arXiv: hep-ph/0311148.

[57]   John M. Campbell, J.W. Huston, and W.J. Stirling. "Hard Interactions of Quarks and Gluons: A Primer for LHC Physics". In: *Rept. Prog. Phys.* 70 (2007), p. 89. DOI: 10.1088/0034-4885/70/1/R02. arXiv: hep-ph/0611148.

[58]   Jun Gao, Lucian Harland-Lang, and Juan Rojo. "The Structure of the Proton in the LHC Precision Era". In: *Phys. Rept.* 742 (2018), pp. 1–121. DOI: 10.1016/j.physrep.2018.03.002. arXiv: 1709.04922 [hep-ph].

[59]   John C. Collins, Davison E. Soper, and George F. Sterman. "Factorization of Hard Processes in QCD". In: vol. 5. 1989, pp. 1–91. DOI: 10.1142/9789814503266\_0001. arXiv: hep-ph/0409313.

[60]   M. Breidenbach et al. "Observed Behavior of Highly Inelastic Electron-Proton Scattering". In: *Phys. Rev. Lett.* 23 (16 Oct. 1969), pp. 935–939. DOI: 10.1103/PhysRevLett.23.935. URL: https://link.aps.org/doi/10.1103/PhysRevLett.23.935.

[61]   J. D. Bjorken. "Asymptotic Sum Rules at Infinite Momentum". In: *Phys. Rev.* 179 (5 Mar. 1969), pp. 1547–1553. DOI: 10.1103/PhysRev.179.1547. URL: https://link.aps.org/doi/10.1103/PhysRev.179.1547.

[62]   Richard P. Feynman. "Very High-Energy Collisions of Hadrons". In: *Phys. Rev. Lett.* 23 (24 Dec. 1969), pp. 1415–1417. DOI: 10.1103/PhysRevLett.23.1415. URL: https://link.aps.org/doi/10.1103/PhysRevLett.23.1415.

[63] L.N. Lipatov. "The parton model and perturbation theory". In: *Sov. J. Nucl. Phys.* 20 (1975), pp. 94–102.

[64] G. Altarelli and G. Parisi. "Asymptotic freedom in parton language". In: *Nuclear Physics B* 126.2 (1977), pp. 298–318. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(77)90384-4. URL: http://www.sciencedirect.com/science/article/pii/0550321377903844.

[65] V.N. Gribov and L.N. Lipatov. "Deep inelastic e p scattering in perturbation theory". In: *Sov. J. Nucl. Phys.* 15 (1972), pp. 438–450.

[66] Yuri L. Dokshitzer. "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics." In: *Sov. Phys. JETP* 46 (1977), pp. 641–653.

[67] Zhiqing Zhang. "New measurement and QCD analysis of DIS data from HERA". In: *PoS* ICHEP2012 (2013). Ed. by Antonio Limosani, p. 289. DOI: 10.22323/1.174.0289. arXiv: 1302.1898 [hep-ex].

[68] Max Klein. "Future Deep Inelastic Scattering with the LHeC". In: *From My Vast Repertoire ...: Guido Altarelli's Legacy*. Ed. by Aharon Levy, Stefano Forte, and Giovanni Ridolfi. 2019, pp. 303–347. DOI: 10.1142/9789813238053\_0015. arXiv: 1802.04317 [hep-ph].

[69] Sidney D. Drell and Tung-Mow Yan. "Massive Lepton-Pair Production in Hadron-Hadron Collisions at High Energies". In: *Phys. Rev. Lett.* 25 (5 Aug. 1970), pp. 316–320. DOI: 10.1103/PhysRevLett.25.316. URL: https://link.aps.org/doi/10.1103/PhysRevLett.25.316.

[70] Francis Halzen, Yu Seon Jeong, and C.S. Kim. "Charge Asymmetry of Weak Boson Production at the LHC and the Charm Content of the Proton". In: *Phys. Rev. D* 88 (2013), p. 073013. DOI: 10.1103/PhysRevD.88.073013. arXiv: 1304.0322 [hep-ph].

[71] A D Martin et al. "Parton Distributions and the LHC: W and Z Production; rev. version". In: *Eur. Phys. J. C* 14.hep-ph/9907231. DTP-99-64. OUTP-99-31-P. RAL-TR-99-047 (July 1999), 133–145. 30 p. DOI: 10.1007/s100520000324. URL: https://cds.cern.ch/record/392675.

[72] Michelangelo L. Mangano. "Production of electroweak bosons at hadron colliders: theoretical aspects". In: vol. 26. 2016, pp. 231–253. DOI: 10.1142/9789814733519\_0013. arXiv: 1512.00220 [hep-ph].

[73] Stefan Kallweit et al. "Accurate single- and double-differential resummation of colour-singlet processes with MATRIX+RadISH: $W^+W^-$ production at the LHC". In: (Apr. 2020). arXiv: 2004.07720 [hep-ph].

[74] Stefano Camarda et al. "DYTurbo: Fast predictions for Drell-Yan processes". In: *Eur. Phys. J. C* 80.arXiv:1910.07049. 3 (Oct. 2019), 251. 15 p. DOI: 10.1140/epjc/s10052-020-7757-5. URL: https://cds.cern.ch/record/2698978.

[75] Wojciech Bizon et al. "The transverse momentum spectrum of weak gauge bosons at N$^3$LL + NNLO". In: *Eur. Phys. J. C* 79.10 (2019), p. 868. DOI: 10.1140/epjc/s10052-019-7324-0. arXiv: 1905.05171 [hep-ph].

[76] Charalampos Anastasiou et al. "High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO". In: *Phys. Rev. D* 69 (2004), p. 094008. DOI: 10.1103/PhysRevD.69.094008. arXiv: hep-ph/0312266.

[77] A. Bermudez Martinez et al. "Production of $Z$ bosons in the parton branching method". In: *Phys. Rev. D* 100 (7 Oct. 2019), p. 074027. DOI: 10.1103/PhysRevD.100.074027. URL: https://link.aps.org/doi/10.1103/PhysRevD.100.074027.

[78] R.Keith Ellis, W.James Stirling, and B.R. Webber. *QCD and collider physics*. Vol. 8. Cambridge University Press, Feb. 2011. ISBN: 978-0-511-82328-2, 978-0-521-54589-1.

[79] Francesco Hautmann, Ignazio Scimemi, and Alexey Vladimirov. "Non-perturbative contributions to vector-boson transverse momentum spectra in hadronic collisions". In: *Physics Letters B* 806 (2020), p. 135478. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2020.135478. URL: http://www.sciencedirect.com/science/article/pii/S0370269320302823.

[80] *Prospects for the measurement of the W-boson transverse momentum with a low pileup data sample at $\sqrt{s}$ = 13 TeV with the ATLAS detector*. Tech. rep. ATL-PHYS-PUB-2017-021. Geneva: CERN, Dec. 2017. URL: https://cds.cern.ch/record/2298152.

[81] Georges et al. Aad. "Measurement of the Transverse Momentum Distribution of W Bosons in $pp$ Collisions at $\sqrt{s}$ = 7 TeV with the ATLAS Detector. Measurement of the Transverse Momentum Distribution of W Bosons in pp Collisions at sqrt(s) = 7 TeV with the ATLAS Detector". In: *Phys. Rev. D* 85.arXiv:1108.6308. CERN-PH-EP-2011-134 (Sept. 2011). Comments: 36 pages plus author list (64 pages total), 8 figures, 2 tables, submitted to Physical Review D, 012005. 36 p. DOI: 10.1103/PhysRevD.85.012005. URL: http://cds.cern.ch/record/1379863.

[82] Vardan et al. Khachatryan. "Measurement of the transverse momentum spectra of weak vector bosons produced in proton-proton collisions at $\sqrt{s}$ = 8 TeV". In: *JHEP* 02.CMS-SMP-14-012. CMS-SMP-14-012. CERN-EP-2016-152 (June 2016). Replaced with the published version. Added the journal reference and DOI. All the figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/SMP-14-012/, 96. 42 p. DOI: 10.1007/JHEP02(2017)096. URL: https://cds.cern.ch/record/2162036.

[83] Timo Antero Aaltonen et al. "Precise Measurement of the W -Boson Mass with the Collider Detector at Fermilab". In: *Phys. Rev. D* 89.7 (2014), p. 072003. DOI: 10.1103/PhysRevD.89.072003. arXiv: 1311.0894 [hep-ex].

[84] Oliver Bruning and Lucio Rossi. "The High-Luminosity Large Hadron Collider". In: *Nature Rev. Phys.* 1.CERN-ACC-2019-0039 (Mar. 2019), 1. 3 p. DOI: 10.1038/s42254-019-0050-6. URL: https://cds.cern.ch/record/2668521.

[85] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[86] Serguei Chatrchyan et al. "Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[87] *The source of protons*. URL: https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.proton_source.

[88] CERN. *CERN's accelerator complex*. URL: https://home.cern/science/accelerators/accelerator-complex.

[89] *RF cavities*. URL: https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.rf_cavities.

[90] Edmund J N Wilson. *An introduction to particle accelerators*. Oxford: Oxford Univ. Press, 2001. DOI: 10.1093/acprof:oso/9780198508298.001.0001. URL: http://cds.cern.ch/record/513326.

[91] Riccardo De Maria, Oliver Brüning, and Rivkin Leonid. "LHC Interaction region upgrade". In: (Jan. 2008).

[92] LHC outreach. *CERN's accelerator complex*. URL: https://lhc-machine-outreach.web.cern.ch/collisions.htm.

[93] Werner Herr and B Muratori. "Concept of luminosity". In: (2006). DOI: 10.5170/CERN-2006-002.361. URL: https://cds.cern.ch/record/941318.

[94] P Grafström and W Kozanecki. "Luminosity determination at proton colliders". In: *Prog. Part. Nucl. Phys.* 81 (2015), 97–148. 52 p. DOI: 10.1016/j.ppnp.2014.11.002. URL: http://cds.cern.ch/record/2153734.

[95] S van der Meer. *Calibration of the effective beam height in the ISR*. Tech. rep. CERN-ISR-PO-68-31. ISR-PO-68-31. Geneva: CERN, 1968. URL: https://cds.cern.ch/record/296752.

[96] Carlo Rubbia. *Measurement of the luminosity of p−overlinep collider with a (generalized) Van der Meer Method*. Tech. rep. CERN-p$\overline{p}$-Note-38. Geneva: CERN, Nov. 1977. URL: https://cds.cern.ch/record/1025746.

[97] The ATLAS collaboration. *Number of Interactions per Crossing*. Jan. 2018. URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[98] G. Aad et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[99] *ATLAS magnet system: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/338080.

[100] *ATLAS inner detector: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331063.

[101] S Haywood et al. *ATLAS inner detector: Technical Design Report, 2*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: http://cds.cern.ch/record/331064.

[102] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331061.

[103] *ATLAS tile calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331062.

[104] *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331068.

[105] ATLAS Collaboration. *Technical Design Report for the Phase-II Upgrade of the ATLAS Muon Spectrometer*. Tech. rep. CERN-LHCC-2017-017. ATLAS-TDR-026. Geneva: CERN, Sept. 2017. URL: https://cds.cern.ch/record/2285580.

[106] g chiodini g, n orlando n, and s spagnolo s. *ATLAS RPC time-of-flight performance*. Tech. rep. ATL-MUON-PROC-2012-002. Geneva: CERN, Apr. 2012. URL: https://cds.cern.ch/record/1437020.

[107] Marco Bruschi. *The new ATLAS/LUCID detector*. Tech. rep. ATL-FWD-PROC-2015-002. Geneva: CERN, June 2015. URL: https://cds.cern.ch/record/2025000.

[108] S. Abdel Khalek et al. "The ALFA Roman Pot Detectors of ATLAS". In: *JINST* 11.11 (2016), P11013. DOI: 10.1088/1748-0221/11/11/P11013. arXiv: 1609.00249 [physics.ins-det].

[109] U. Amaldi et al. "Measurements of the proton proton total cross-sections by means of Coulomb scattering at the Cern intersecting storage rings". In: *Phys. Lett. B* 43 (1973), pp. 231–236. DOI: 10.1016/0370-2693(73)90277-3.

[110] *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*. Tech. rep. ATLAS-CONF-2019-021. Geneva: CERN, June 2019. URL: https://cds.cern.ch/record/2677054.

[111] William Panduro Vazquez and ATLAS Collaboration. *The ATLAS Data Acquisition system in LHC Run 2*. Tech. rep. ATL-DAQ-PROC-2017-007. 3. Geneva: CERN, Feb. 2017. DOI: 10.1088/1742-6596/898/3/032017. URL: https://cds.cern.ch/record/2244345.

[112] "Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data". In: *JINST* 14.arXiv:1908.00005. 12 (Aug. 2019). 31 figures, 3 tables. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/EGAM-2018-01, P12006. 70 p. DOI: 10.1088/1748-0221/14/12/P12006. URL: https://cds.cern.ch/record/2684552.

[113] *Expected electron performance in the ATLAS experiment*. Tech. rep. ATL-PHYS-PUB-2011-006. Geneva: CERN, Apr. 2011. URL: http://cds.cern.ch/record/1345327.

[114] *Data/MC Comparison for Calorimeter Shower Shapes of High Et Electrons*. Tech. rep. ATL-COM-PHYS-2011-1299. Geneva: CERN. URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/EGAMMA/PublicPlots/20111005/ATL-COM-PHYS-2011-1299/index.html.

[115] "Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data. Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton-proton collision data". In: *Eur. Phys. J. C* CERN-PH-EP-2014-040. CERN-PH-EP-2014-040 (Apr. 2014). Comments: 38 pages plus author list (62 pages total), 20 figures, 4 tables, submitted to JHEP, All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/PERF-2013-03/, 74. 38 p. DOI: 10.1140/epjc/s10052-014-2941-0. URL: https://cds.cern.ch/record/1694142.

[116] "Performance of the ATLAS Silicon Pattern Recognition Algorithm in Data and Simulation at $\sqrt{s}$ = 7 TeV". In: (July 2010).

[117] M. Aaboud et al. "Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2". In: *Eur. Phys. J. C* 77.10 (2017), p. 673. DOI: 10.1140/epjc/s10052-017-5225-7. arXiv: 1704.07983 [hep-ex].

[118] R. Fruhwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nucl. Instrum. Meth. A* 262 (1987), pp. 444–450. DOI: 10.1016/0168-9002(87)90887-4.

[119] S. Boutle et al. "Primary vertex reconstruction at the ATLAS experiment". In: *J. Phys. Conf. Ser.* 898.4 (2017). Ed. by Richard Mount and Craig Tull, p. 042056. DOI: 10.1088/1742-6596/898/4/042056.

[120] *Development of ATLAS Primary Vertex Reconstruction for LHC Run 3*. Tech. rep. ATL-PHYS-PUB-2019-015. Geneva: CERN, Apr. 2019. URL: https://cds.cern.ch/record/2670380.

[121] E. Bouhova-Thacker et al. "Vertex reconstruction in the ATLAS experiment at the LHC". In: *2008 IEEE Nuclear Science Symposium and Medical Imaging Conference and 16th International Workshop on Room-Temperature Semiconductor X-Ray and Gamma-Ray Detectors*. 2008, pp. 1720–1727. DOI: 10.1109/NSSMIC.2008.4774734.

[122] Georges Aad et al. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV". In: *Eur. Phys. J. C* 76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[123] J. Illingworth and J. Kittler. "A survey of the hough transform". In: *Computer Vision, Graphics, and Image Processing* 44.1 (1988), pp. 87–116. ISSN: 0734-189X. DOI: https://doi.org/10.1016/S0734-189X(88)80033-1. URL: http://www.sciencedirect.com/science/article/pii/S0734189X88800331.

[124] Steffen Kaiser. "Search for the Higgs Boson in the Process $pp \rightarrow Hqq$, $H \rightarrow WW$ with the ATLAS Detector". PhD thesis. Munich, Technische Universitaet Muenchen, 2010.

[125] Morad Aaboud et al. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s}$ = 13 TeV". In: *Eur. Phys. J. C* 79.8 (2019), p. 639. DOI: 10.1140/epjc/s10052-019-7140-6. arXiv: 1902.04655 [physics.ins-det].

[126] W Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003. Geneva: CERN, Apr. 2008. URL: https://cds.cern.ch/record/1099735.

[127] Georges et al. Aad. "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1". In: *Eur. Phys. J. C* 77.CERN-PH-EP-2015-304 (Mar. 2016). Comments: 64 pages plus author list + cover page (87 pages in total), 41 figures, 3 tables, submitted to EPJC. All figures including auxiliary figures are available at http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/PERF-2014-07/, 490. 87 p. DOI: 10.1140/epjc/s10052-017-5004-5. URL: https://cds.cern.ch/record/2138166.

[128] Georges Aad et al. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data". In: *JINST* 14.12 (2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].

[129] Thijs G. Cornelissen et al. "The global chi**2 track fitter in ATLAS". In: *J. Phys. Conf. Ser.* 119 (2008). Ed. by Randall Sobie, Reda Tafirout, and Jana Thomson, p. 032013. DOI: 10.1088/1742-6596/119/3/032013.

[130] T Cornelissen et al. *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*. Tech. rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002. Geneva: CERN, Mar. 2007. URL: https://cds.cern.ch/record/1020106.

[131] *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*. Tech. rep. ATLAS-CONF-2012-047. Geneva: CERN, May 2012. URL: https://cds.cern.ch/record/1449796.

[132] Morad Aaboud et al. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data". In: *JINST* 14.03 (2019), P03017. DOI: 10.1088/1748-0221/14/03/P03017. arXiv: 1812.03848 [hep-ex].

[133] Morad Aaboud et al. "Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016". In: *Eur. Phys. J. C* 79.3 (2019), p. 205. DOI: 10.1140/epjc/s10052-019-6650-6. arXiv: 1810.05087 [hep-ex].

[134] Matteo Cacciari and Gavin P. Salam. "Pileup subtraction using jet areas". In: *Phys. Lett. B* 659 (2008), pp. 119–126. DOI: 10.1016/j.physletb.2007.09.077. arXiv: 0707.1378 [hep-ph].

[135] Morad et al. Aaboud. "Jet reconstruction and performance using particle flow with the ATLAS Detector. Jet reconstruction and performance using particle flow with the ATLAS Detector". In: *Eur. Phys. J. C* 77.CERN-EP-2017-024. 7 (Mar. 2017). 67 pages in total, author list starting page 51, 37 figures, 1 table, final version published in Eur. Phys. J C, all figures including auxiliary figures are available at http://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/PERF-2015-09/, 466. 67 p. DOI: 10.1140/epjc/s10052-017-5031-2. URL: https://cds.cern.ch/record/2257597.

[136] Jan Kretzschmar. *Samples and Physics modelling for low pile-up runs taken in 2017 and 2018*. Tech. rep. ATL-COM-PHYS-2019-075. Geneva: CERN, Feb. 2019. URL: https://cds.cern.ch/record/2657141.

[137] ATLAS Collaboration. "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data". In: *Eur. Phys. J. C* 74 (2014), p. 3071. DOI: 10.1140/epjc/s10052–014–3071–4. arXiv: 1407.5063 [hep-ex].

[138] Tairan Xu, Hicham Atmani, and Ludovica Aperio Bella. *Electron corrections for low pile-up runs taken in 2017 and 2018*. Tech. rep. ATL-COM-PHYS-2019-077. Geneva: CERN, Feb. 2019. URL: https://cds.cern.ch/record/2657152.

[139] Nansi Andari et al. *Electron and photon energy calibration with the ATLAS detector using 2015-2017 LHC proton-proton collision data*. Tech. rep. ATL-COM-PHYS-2018-1720. Geneva: CERN, 2018. URL: https://cds.cern.ch/record/2651890.

[140] Alexander Sydorenko, Jakub Andrzej Kremer, and Tairan Xu. *Muon corrections for low pile-up runs taken in 2017 and 2018*. Tech. rep. ATL-COM-PHYS-2019-072. Geneva: CERN, Feb. 2019. URL: https://cds.cern.ch/record/2657116.

[141] *Study of alignment-related systematic effects on the ATLAS Inner Detector tracking*. Tech. rep. ATLAS-CONF-2012-141. Geneva: CERN, Oct. 2012. URL: https://cds.cern.ch/record/1483518.

[142] Nicolas Maximilian Koehler et al. *Muon Efficiency Measurements on the Full Run 2 dataset*. Tech. rep. ATL-COM-PHYS-2019-176. Geneva: CERN, Mar. 2019. URL: https://cds.cern.ch/record/2665704.

[143] Mengran Li et al. *Hadronic recoil reconstruction and calibration for low pile-up runs taken in 2017 and 2018*. Tech. rep. ATL-COM-PHYS-2019-078. Geneva: CERN, Feb. 2019. URL: https://cds.cern.ch/record/2657182.

[144] E. Mirkes. "Angular decay distribution of leptons from W-bosons at NLO in hadronic collisions". In: *Nuclear Physics B* 387.1 (1992), pp. 3–85. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550–3213(92)90046–E. URL: http://www.sciencedirect.com/science/article/pii/055032139290046E.

[145] Stefano Catani et al. "Vector Boson Production at Hadron Colliders: A Fully Exclusive QCD Calculation at Next-to-Next-to-Leading Order". In: *Phys. Rev. Lett.* 103 (8 Aug. 2009), p. 082001. DOI: 10.1103/PhysRevLett.103.082001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.103.082001.

[146] Jan Kretzschmar. *Samples and Physics modelling for low pile-up runs taken in 2017 and 2018*. Tech. rep. ATL-COM-PHYS-2019-075. Geneva: CERN, Feb. 2019. URL: https://cds.cern.ch/record/2657141.

[147] Paolo Nason. "A New method for combining NLO QCD with shower Monte Carlo algorithms". In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv: hep-ph/0409146.

[148] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method". In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].

[149] Simone Alioli et al. "NLO vector-boson production matched with shower in POWHEG". In: *JHEP* 0807 (2008), p. 060. DOI: 10.1088/1126-6708/2008/07/060. arXiv: 0805.4802 [hep-ph].

[150] Simone Alioli et al. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].

[151] T. Sjöstrand, S. Mrenna, and P. Skands. "Brief Introduction to PYTHIA 8.1". In: *Comput. Phys. Comm.* 178 (2008), p. 85. DOI: 10.1016/j.cpc.2008.01.036. arXiv: 0710.3820v1 [hep-ph].

[152] ATLAS Collaboration. "Measurement of the $Z/\gamma^*$ boson transverse momentum distribution in $pp$ collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *JHEP* 09 (2014), p. 145. DOI: 10.1007/JHEP09(2014)145. arXiv: 1406.3660 [hep-ex].

[153] Piotr Golonka and Zbigniew Was. "PHOTOS Monte Carlo: A Precision tool for QED corrections in $Z$ and $W$ decays". In: *Eur. Phys. J.* C45 (2006), pp. 97–107. DOI: 10.1140/epjc/s2005-02396-4. arXiv: hep-ph/0506026.

[154] Stefan Höche et al. "NLO matrix elements and truncated showers". In: *JHEP* 1108 (2011), p. 123. DOI: 10.1007/JHEP08(2011)123. arXiv: 1009.1127 [hep-ph].

[155] Richard D. Ball et al. "Parton distributions with LHC data". In: *Nucl. Phys. B* 867 (2013), p. 244. DOI: 10.1016/j.nuclphysb.2012.10.003. arXiv: 1207.1303 [hep-ph].

[156] ATLAS Collaboration. *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie–Landshoff diffractive model.* ATL-PHYS-PUB-2016-017. 2016. URL: https://cds.cern.ch/record/2206965.

[157] S. Catani and M. Grazzini. "An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC". In: *Phys. Rev. Lett.* 98 (2007), p. 222002. DOI: 10.1103/PhysRevLett.98.222002. arXiv: hep-ph/0703012 [hep-ph].

[158] S. Catani et al. "Vector boson production at hadron colliders: A Fully exclusive QCD calculation at NNLO". In: *Phys. Rev. Lett.* 103 (2009), p. 082001. DOI: 10.1103/PhysRevLett.103.082001. arXiv: 0903.2120 [hep-ph].

[159] L.A. Harland-Lang, A. D. Martin, P. Motylinski, R. S. Thorne. "Parton distributions in the LHC era: MMHT 2014 PDFs". In: *Eur. Phys. J. C* 75.5 (2015), p. 204. DOI: 10.1140/epjc/s10052-015-3397-6. arXiv: 1412.3989 [hep-ph].

[160] ATLAS Collaboration. "Measurement of $W^{\pm}$ and $Z$-boson production cross sections in $pp$ collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Phys. Lett. B* 759 (2016), p. 601. DOI: 10.1016/j.physletb.2016.06.023. arXiv: 1603.09222 [hep-ex].

[161] ATLAS Collaboration. "Measurements of $W$ and $Z$ boson production in $pp$ collisions at $\sqrt{s} = $ 5.02 TeV with the ATLAS detector". In: *Eur. Phys. J. C* 79 (2019), p. 128. DOI: 10.1140/epjc/s10052-019-6622-x. arXiv: 1810.08424 [hep-ex].

[162] Michal Czakon and Alexander Mitov. "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders". In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv: 1112.5675 [hep-ph].

[163] CMS Colaboration. "Measurement of the inclusive $t\bar{t}$ cross section in $pp$ collisions at $\sqrt{s} = 5.02$ TeV using final states with at least one charged lepton". In: *JHEP* 03 (2018), p. 115. DOI: 10.1007/JHEP03(2018)115. arXiv: 1711.03143 [hep-ex].

[164] Tairan Xu and Maarten Boonekamp. *Multi-jet background in low-pile-up runs taken in 2017 and 2018*. Tech. rep. ATL-COM-PHYS-2019-076. Geneva: CERN, July 2019. URL: https://cds.cern.ch/record/2657146.

[165] Nansi Andari et al. *Measurement of $m_W$ with 7 TeV data: W boson mass measurement*. Tech. rep. ATL-COM-PHYS-2014-1569. Geneva: CERN, Dec. 2014. URL: https://cds.cern.ch/record/1976186.

[166] Stefan Schmitt. "Data Unfolding Methods in High Energy Physics". In: *EPJ Web Conf.* 137 (2017). Ed. by Y. Foka, N. Brambilla, and V. Kovalenko, p. 11008. DOI: 10.1051/epjconf/201713711008. arXiv: 1611.01927 [physics.data-an].

[167] G. D'Agostini. "A multidimensional unfolding method based on Bayes' theorem". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 362.2 (1995), pp. 487–498. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(95)00274-X. URL: http://www.sciencedirect.com/science/article/pii/016890029500274X.

[168] G. D'Agostini. "Improved iterative Bayesian unfolding". In: *Alliance Workshop on Unfolding and Data Correction*. Oct. 2010. arXiv: 1010.0632 [physics.data-an].

[169] Hicham Atmani et al. *Measurment of the $p_T$ spectrum of W- and Z-bosons produced in pp collisions at $\sqrt{s} = 5$ TeV and 13 TeV in low-pileup runs*. Tech. rep. ATL-COM-PHYS-2018-1084. Geneva: CERN, July 2018. URL: https://cds.cern.ch/record/2632159.

[170] Yann Lecun et al. "Gradient-based learning applied to document recognition". In: *Proceedings of the IEEE*. 1998, pp. 2278–2324.

[171] Léon Bottou. "Stochastic Gradient Descent Tricks". In: vol. 7700. Jan. 2012, pp. 421–436. DOI: 10.1007/978-3-642-35289-8_25.

[172] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors". In: *Nature* 323 (1986), pp. 533–536.

[173] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. cite arxiv:1412.6980Comment: Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015. 2014. URL: http://arxiv.org/abs/1412.6980.

[174] Yu Nesterov. "A method of solving a convex programming problem with convergence rate $O(1/k^2)$". In: vol. 27. Jan. 1983, pp. 372–376.

[175] Boris Polyak. "Some methods of speeding up the convergence of iteration methods". In: *Ussr Computational Mathematics and Mathematical Physics* 4 (Dec. 1964), pp. 1–17. DOI: 10.1016/0041-5553(64)90137-5.

[176] Yann LeCun et al. "Efficient BackProp". In: *Neural Networks: Tricks of the Trade*. Ed. by Genevieve B. Orr and Klaus-Robert Müller. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 9–50. ISBN: 978-3-540-49430-0. DOI: 10.1007/3-540-49430-8_2. URL: https://doi.org/10.1007/3-540-49430-8_2.

[177] Kurt and Hornik. "Approximation capabilities of multilayer feedforward networks". In: *Neural Networks* 4.2 (1991), pp. 251–257. DOI: 10.1016/0893-6080(91)90009-T. URL: http://www.sciencedirect.com/science/article/pii/089360809190009T.

[178] G. Cybenko. "Approximation by superpositions of a sigmoidal function". In: *Mathematics of Control, Signals and Systems* 2.4 (Dec. 1989), pp. 303–314. ISSN: 1435-568X. DOI: 10.1007/BF02551274. URL: https://doi.org/10.1007/BF02551274.

[179] D. E. Rumelhart and D. Zipser. "Feature Discovery by Competitive Learning". In: *Parallel Distributed Processing*. MIT Press, 1986, pp. 151–193.

[180] Pankaj Mehta et al. "A high-bias, low-variance introduction to Machine Learning for physicists". In: *Phys. Rept.* 810 (2019), pp. 1–124. DOI: 10.1016/j.physrep.2019.03.001. arXiv: 1803.08823 [physics.comp-ph].

[181] Sergey Ioffe and Christian Szegedy. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift". In: *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37*. ICML15. Lille, France: JMLR.org, 2015, pp. 448–456.

[182] F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

# Acknowledgements

I would like to express my sheer gratitude to the people who have not only made the writing of this thesis possible but also turned it into an extremely interesting and pleasant experience.

First of all, I would like to thank my supervisors Fabrice Balli and Maarten Boonekamp for their constant support, friendly attitude, and endless patience. During the three years of the doctorate, Fabrice and Maarten have spent many hours of their priceless time in enlightening discussions with the thesis author. It was a joy and an honour for me to have been supervised by these exceptionally nice and intelligent people.

My deepest thanks to the jury members: president of the jury David Rousseau, jury member Aram Apyan and thesis reviewers Alessandro Vicini and Andrew Pilkington. Your profound questions and keen comments have helped to greatly improve this thesis.

It was great to be a part of the DPhP division of CEA Saclay and ATLAS group in particular. I would like to thank ATLAS group leaders Claude Guyot and Frederic Deliot, DPhP directors Gauthier Hamel de Monchenault and Nathalie Besson, vice director Georges Vasseur and direction assistant Martine Oger.

I was extremely lucky to work with an amazing analysis team: Luis Fayard, Jan Kretschmar, Ludovica Aperio Bella, Hicham Atmani, thank you. Special cheers to my beloved office-mates and friends, Nancy and Artur. Nancy, you

are the kindest person I have ever encountered. Artur, I am so thankful for your amazing sense of humour, even when things were hard.

It was an invaluable experience to make friends during the M2 NPAC study and I am very happy to have you in my life. Luis, my best binome, I believe there is no such physics problem that we would not be able to crack by just discussing it. Jad, Konie, Manuel, Tetiana, our ATLAS journey was full of fun - and more is to come. Tony, you have become my friend at the moment when I needed it most and I will never forget that moment.

I would like to thank Volodymyr Skalozub, my first supervisor and extraordinary physicist who has helped me to find my way to physics. Oleksii Gulov, thank you for becoming my friend after being a great teacher and mentor. My fellow physicists, colleagues in translation, and dear friends Andrii and Ivan, you are the best team I have ever had.

Last but definitely not least sincere thanks to my family for all the love and support, which was vital. Dad, you were the first physicist I have met in my life and my first inspiration. Mom, you have shown me that the world is big and it is worth exploring. My dear brother Ilya, you are my best friend since I was born. My darling fiancée Natalie, thank you for your love, kindness, and patience. We have made it together.

**Titre:** Calibration des cascades électromagnetiques, application de lapprentissage profond a la reconstruction du recul hadronique et mesure de la distribution en impulsion transverse du boson W dans l'expérience ATLAS.

**Mots clés:** Interactions électrofaibles, Modèle standard, Grand collisionneur de hadrons, apprentissage profond, Bosons W

**Résumé:** La première partie de la thèse contient une description de la méthode d'étalonnage du calorimètre électromagnétique, corrigeant les différences entre les données et la simulation pour ce qui concerne le développement des cascades électromagnétiques dans le calorimètre. La méthode améliore l'identification des électrons et réduit l'incertitude systématique associée. La majeure partie de la thèse est consacrée à la mesure précise du spectre en impulsion transverse (pT) du boson W à l'aide des données collectées par l'expérience ATLAS à des énergies dans le centre de masse de 5 et 13 TeV lors de deux prises de données spéciales, à faible taux dempilement, en 2017 et en 2018. La motivation pour la mesure précise du spectre en impulsion transverse du boson W est double. Premièrement, elle sert de test pour les prédictions théoriques obtenues dans le cadre du Modèle Standard et permet de comparer les performances des générateurs Monte-Carlo (MC). La deuxième raison est que ce spectre est un

ingrédient à la mesure de la masse du boson W, qui est un paramètre du Modèle Standard. L'utilisation de données à faible taux d'empilement permet de réduire significativement l'incertitude systématique due au recul hadronique et améliore de ce fait la précision sur la mesure du spectre. La thèse décrit la méthodologie de la mesure du spectre en pT du boson W ainsi que les étalonnages appliqués, les corrections et les incertitudes associées. Le résultat final est obtenu à partir du recul hadronique mesuré à l'aide d'une procédure de déconvolution des effets de détecteur et est comparé aux prédictions théoriques obtenues avec différents générateurs Monte-Carlo. Une méthode alternative pour la reconstruction du recul hadronique, avec l'utilisation de réseaux neuronaux profonds est proposée dans la thèse. Il y est montré que cette méthode améliore la résolution du recul hadronique mesuré d'environ 10% dans la région la plus pertinente, de faible pT. Les observables obtenus par cette approche améliorent la sensibilité à la masse du boson W.

**Title:** Calibration of electron shower shapes, hadronic recoil reconstruction using deep learning algorithms and the measurement of W boson transverse momentum distribution with the ATLAS detector.

**Keywords:** electroweak interactions, Standard Model, Large Hadron Collider, deep learning, W boson

**Abstract:** The initial part of the thesis contains the description of the method for electromagnetic calorimeter calibration, correcting for the Data-MC discrepancy in the development of the electromagnetic showers in the calorimeter. The method improves electron identification and reduces the associated systematic uncertainty. The major part of the thesis is dedicated to the precise measurement of the W boson transverse spectrum using the data, collected by the ATLAS experiment at the energies of 5 and 13 TeV during two special low pile-up runs in 2017 and 2018. The motivation for the precise measurement of the W boson transverse spectrum is twofold. First, it serves as a test for the theoretical predictions obtained within the Standard Model and allows to benchmark the performance of the Monte-Carlo (MC) generators. The second reason is because the W pT spectrum is an input component for the measurement of the W boson mass which is a Standard Model parameter. The use of low pile-up data allows to significantly reduce the hadronic recoil systematic uncertainty improving the precision of the spectrum measurement. The thesis describes the methodology of the W boson pT spectrum measurement as well as the imposed calibrations, corrections and the associated uncertainties. The final result is obtained from the measured hadronic recoil using an unfolding procedure and is compared to the theoretical predictions obtained with different Monte-Carlo generators. An alternative method for the hadronic recoil reconstruction with the use of deep neural networks is proposed in the thesis. The method is shown to improve the resolution of the measured hadronic recoil by about 10% in the most relevant region of low pT. The observables obtained using approach improve the sensitivity to the mass of the W boson.