



**HAL**  
open science

# Measurements of inclusive and differential cross sections for vector boson productions with the ATLAS detector at the LHC

Xiaowen Su

► **To cite this version:**

Xiaowen Su. Measurements of inclusive and differential cross sections for vector boson productions with the ATLAS detector at the LHC. High Energy Physics - Experiment [hep-ex]. Université Paris-Saclay; University of science and technology of China, 2024. English. NNT : 2024UPASP080 . tel-04757343

**HAL Id: tel-04757343**

**<https://theses.hal.science/tel-04757343v1>**

Submitted on 28 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Measurements of inclusive and differential cross sections for vector boson productions with the ATLAS detector at the LHC

*Mesures des sections efficaces inclusives et différentielles  
pour les productions de bosons vecteurs avec le détecteur  
ATLAS au LHC*

**Thèse de doctorat de l'université Paris-Saclay et de  
University of Science and Technology of China**

École doctorale n°576 Particules hadrons énergie et noyau :  
instrumentation, imagerie, cosmos et simulation (PHENIICS)  
Spécialité de doctorat : Physique des particules  
Graduate School : Physique. Référent : Faculté des sciences d'Orsay

Thèse préparée dans l'unité de recherche **IJCLab** (Université Paris-Saclay,  
CNRS), sous la direction de **Zhiqing ZHANG**, directeur de recherche, et la  
co-direction de **Haiping PENG**, Professeur.

**Thèse soutenue à Paris-Saclay, le 26 septembre 2024, par**

**Xiaowen SU**

## Composition du jury

Membres du jury avec voix délibérative

<b>Samira HASSANI</b> Directrice de recherche, Institut de Recherche sur les lois Fondamentales de l'Univers, CEA - Saclay	Présidente
<b>Klaus MÖNIG</b> Directeur de recherche, Deutsches Elektronen- Synchrotron	Rapporteur & Examineur
<b>Shu LI</b> Professeur Associé (équivalent HDR), Shanghai Jiao Tong University	Rapporteur & Examineur
<b>Yusheng WU</b> Professeur, University of Science and Technology of China	Examineur
<b>Guangshun HUANG</b> Professeur, University of Science and Technology of China	Examineur

**Titre :** Mesures des sections efficaces inclusives et différentielles pour les productions de bosons vecteurs avec le détecteur ATLAS au LHC

**Mots clés :** ATLAS, Section efficace, Mesure de précision, Modèle standard, Processus Drell-Yan

**Résumé :** Les mesures précises des propriétés des bosons vecteurs (bosons  $W$  et  $Z$ ) sont cruciales pour comprendre le secteur électrofaible du Modèle Standard. En tant que médiateurs de la force faible, ces mesures fournissent une excellente exploration de la chromodynamique quantique perturbative (pQCD) et de la structure du proton, et peuvent également révéler des déviations par rapport aux prédictions théoriques, indiquant potentiellement une nouvelle physique au-delà du Modèle Standard. Par exemple, la mesure de la production des bosons  $W$  et  $Z$  à travers divers régimes d'énergie pourrait dévoiler des contributions potentielles de nouvelles particules ou interactions observables à des énergies plus élevées mais indétectables à des énergies plus basses.

Dans cette thèse, deux mesures des sections efficaces de production  $W \rightarrow l\nu$  et  $Z \rightarrow ll$  (où  $l$  désigne un électron ou un muon) sont présentées en utilisant deux ensembles de données collectées avec le détecteur ATLAS au Grand Collisionneur de Hadrons (LHC).

La mesure des sections efficaces fiduciales et totales inclusives de production de bosons  $W$  et  $Z$  et de leurs rapports ont été effectuées sur les données collectées en 2022 lors de collisions proton-proton à une énergie dans le centre de masse de  $\sqrt{s} = 13,6$  TeV, correspondant à une luminosité intégrée de  $29 \text{ fb}^{-1}$ . Il s'agit de la première mesure du taux de production des bosons  $W$  et  $Z$  utilisant les données de la Run-3 du LHC prises en 2022 à une énergie sans précédent. De plus, la première mesure du rapport des sections efficaces de pro-

duction de paires quark-antiquark top par rapport aux bosons  $W$  utilisant le même échantillon de données est également effectuée dans cette étude, explorant différentes densités de partons dans le proton (PDFs). Les mesures sont comparées aux prédictions du Modèle Standard calculées à l'ordre supérieur NNLO en  $\alpha_s$ , avec une précision logarithmique NNLL en QCD plus une précision à l'ordre NLO en électrofaible. Un bon accord est observé entre les sections efficaces mesurées et les prédictions pour les sections efficaces des bosons  $W$  et  $Z$ , tandis que le rapport des sections efficaces  $t\bar{t}/W$  est légèrement inférieur à certaines prédictions.

La deuxième analyse incluse dans cette thèse est la mesure de la section efficace différentielle de la production  $W \rightarrow l\nu$  en utilisant les données collectées en 2017 et 2018 lors de collisions proton-proton à des énergies dans le centre de masse de  $\sqrt{s} = 5.02$  et 13 TeV. Ces ensembles de données proviennent de collections spéciales durant la Run-2 au LHC, également appelées collisions à faible empiement caractérisées par un faible nombre d'interactions par croisement de paquets,  $\langle \mu \rangle \sim 2$ . La mesure précise des sections efficaces différentielles du boson  $W$  avec de faibles incertitudes systématiques améliore la connaissance de la structure du proton et est utilisée pour imposer des contraintes sur les PDFs en effectuant la repondération hessienne sur les ensembles de PDFs existants, entraînant des contraintes significatives sur les fonctions de distribution des quarks.

**Title :** Measurements of inclusive and differential cross sections for vector boson productions with the ATLAS detector at the LHC

**Keywords :** ATLAS, Cross section, Precision measurement, Standard Model, Drell-Yan process

**Abstract :** Precision measurements of the vector boson ( $W^\pm$  and  $Z$ -bosons) properties are crucial for understanding the electroweak sector of the Standard Model. As the mediators of the weak force, such measurements provide an excellent probe of perturbative quantum chromodynamics (pQCD) and of the proton structure and may also reveal deviations from theoretical predictions, potentially indicating new physics beyond the Standard Model, for example, measuring  $W^\pm$  and  $Z$ -boson production across various energy regimes could uncover potential contributions from new particles or interactions that are observable at higher energies but undetectable at lower energies.

In this thesis, two measurements of the  $W^\pm \rightarrow \ell\nu$  and  $Z \rightarrow \ell\ell$  (where  $\ell$  refers to electron or muon) production cross sections are presented using two datasets collected with the ATLAS detector at the Large Hadron Collider (LHC).

The measurement of the inclusive fiducial and total  $W^\pm$ - and  $Z$ -boson production cross sections and their ratios are performed on the data collected in 2022 in proton–proton collisions at a centre-mass-of-energy of  $\sqrt{s} = 13.6$  TeV, corresponding to an integrated luminosity of  $29 \text{ fb}^{-1}$ . It is the first measurement of the production rate of  $W^\pm$  and  $Z$ -boson using LHC Run-3 data taken in 2022 at the unprecedented high energy. Moreover, the first measurement of the ratio of top–antitop quark

( $t\bar{t}$ ) pair to  $W^\pm$ -boson production cross section using the same data sample is also performed in this study, which probe different parton densities within the proton. The measurements are compared with Standard-Model predictions calculated at next-to-next-to-leading-order (NNLO) in  $\alpha_s$ , next-to-next-to-leading logarithmic accuracy in QCD plus next-to-leading-order (NLO) electroweak accuracy. Good agreement is observed between measured cross sections and the predictions in  $W^\pm$  and  $Z$  cross sections whereas the  $t\bar{t}$  over  $W^\pm$  cross section ratio is slightly lower than some predictions.

The second analysis involved in this thesis is the measurement of the differential  $W^\pm \rightarrow \ell\nu$  production cross section using the data collected in 2017 and 2018 in proton–proton collisions at centre-mass-of-energies of  $\sqrt{s} = 5.02$  and 13 TeV. This datasets are taken from special runs during Run-2 at the LHC, which are also called low pile-up runs characterized by a low number of interactions per bunch crossing,  $\langle\mu\rangle \sim 2$ . The precise measurement on the  $W^\pm$  differential cross sections with low systematic uncertainties improves the knowledge on the proton structure and is used to place constraints on the PDFs by performing the Hessian reweighting on the existing PDF sets, resulting in significant constraints on the quark distribution functions.

# Synthèse

## Introduction

Le Modèle Standard (MS) est un cadre théorique permettant d'expliquer les éléments constitutifs fondamentaux de la matière, les particules élémentaires, et leurs interactions. Le développement de la théorie du MS au cours du 20<sup>ème</sup> siècle a impliqué des avancées théoriques majeures et des découvertes expérimentales. Il a permis d'expliquer un large éventail d'observations expérimentales et a été testé par diverses expériences au fil des décennies. Le MS s'est avéré être une théorie extrêmement accomplie.

En 1983, les bosons  $W^\pm$  et  $Z$  ont été découverts au collisionneur proton-antiproton dans le SPS du CERN. En tant que médiateurs de l'interaction faible, des mesures précises des propriétés des bosons  $W$  et  $Z$  sont essentielles pour comprendre le secteur électrofaible du Modèle Standard. La production de ces bosons vecteurs constitue un test solide pour la chromodynamique quantique perturbative (pQCD). En mesurant leurs sections efficaces à des énergies élevées, des signes potentiels de nouvelle physique peuvent émerger, en particulier si des écarts sont observés qui ne sont pas significatifs à des échelles d'énergie plus basses. De plus, les processus impliquant les bosons  $W$  et  $Z$  servent souvent de bruits de fond importants dans de nombreuses analyses physiques. Des mesures précises des sections efficaces peuvent aider à réduire les incertitudes dans la modélisation de ces processus de bruit de fond. Par ailleurs, la production des bosons  $W$  et  $Z$  dans les collisions proton-proton offre des informations précieuses sur la structure du proton. Cela peut contribuer à une meilleure compréhension du modèle des partons, en éclairant la manière dont les quarks et les gluons, les constituants fondamentaux des protons, interagissent et distribuent leur impulsion au sein du proton lors de collisions à haute énergie. Le processus de production Drell-Yan (DY), y compris  $W^\pm \rightarrow \ell^\pm \nu$ , est un outil important pour la détermination des fonctions de distribution des partons (PDFs). Le processus DY est bien compris théoriquement et peut fournir une sonde directe des distributions de quarks et d'anti-quarks dans le proton, en particulier la distribution des quarks de mer.

Cette thèse présente deux mesures des sections efficaces de la production de bosons vecteurs. La première analyse concerne la mesure de la section efficace inclusive de la production de bosons vecteurs. Cette étude utilise des états finaux leptoniques, mesurant les sections efficaces fiduciaires et totales inclusives de  $W^\pm \rightarrow \ell^\pm \nu_\ell (\bar{\nu}_\ell)$  et  $Z \rightarrow \ell^+ \ell^-$ , ainsi que les rapports de leurs sections efficaces fiduciaires,  $W^+/W^-$  et  $W^\pm/Z$ . En plus des rapports des sections efficaces de  $W^\pm$  et  $Z$ , les rapports  $R_{t\bar{t}/W^\pm}$ ,  $R_{t\bar{t}/W^-}$  et  $R_{t\bar{t}/W^+}$  sont également mesurés dans cette analyse, en raison de leur sensibilité à diverses densités de partons. Cette analyse utilise les données collectées par le détecteur ATLAS lors des collisions  $pp$  à  $\sqrt{s} = 13,6$  TeV, correspondant à une luminosité intégrée de  $29 \text{ fb}^{-1}$ . Ce jeu de données a été collecté en 2022, qui est la

première année de la prise de données de Run-3. L'analyse des processus bien connus  $W^\pm$  et  $Z$  aide à valider la performance du détecteur et le cadre de reconstruction lors des premières périodes de Run 3.

## Le LHC et l'expérience ATLAS

Le Grand Collisionneur de Hadrons (LHC) est le plus grand et le plus puissant accélérateur de particules au monde. Il est situé dans un tunnel circulaire de 26,7 km de long, au CERN (l'Organisation Européenne pour la Recherche Nucléaire) près de Genève, en Suisse. Le LHC utilise deux anneaux d'aimants supraconducteurs pour accélérer des protons ou des ions lourds à une vitesse proche de celle de la lumière, simulant ainsi des conditions similaires à celles qui existaient juste après le Big Bang.

Deux faisceaux de protons voyageant dans des directions opposées sont amenés à entrer en collision en quatre points spécifiques d'interaction autour de l'anneau du LHC, où quatre détecteurs majeurs — ATLAS, CMS, ALICE et LHCb — sont positionnés pour étudier les collisions résultantes.

Le détecteur ATLAS (A Toroidal LHC ApparatuS) est l'un des quatre principaux détecteurs situés au Point 1 du LHC, près du CERN. Avec le CMS, ATLAS est un détecteur à usage général, conçu pour explorer un large éventail de phénomènes physiques.

ATLAS est une structure cylindrique immense, mesurant environ 25 mètres de long et 44 mètres de hauteur, et pesant environ 7 000 tonnes, ce qui en fait le plus grand détecteur de particules jamais construit en volume.

Le détecteur ATLAS est composé de plusieurs sous-détecteurs, chacun ayant une fonction unique. Au total, il comprend six sous-systèmes de détection distincts, disposés concentriquement autour du point d'interaction. Ces sous-systèmes sont conçus pour enregistrer la trajectoire, l'impulsion et l'énergie des particules, permettant ainsi l'identification et la mesure précises des particules individuelles. En commençant par le centre, le détecteur interne est entouré d'un fin système magnétique générant un champ magnétique de 2 T. Autour de celui-ci se trouvent les calorimètres électromagnétiques et hadroniques, suivis par le spectromètre à muons. La couche la plus externe est un grand système d'aimants qui courbe les trajectoires des particules chargées, permettant des mesures précises de leur impulsion.

## Mesure des sections efficaces de production des bosons $W^\pm$ et $Z$ et de leurs rapports à $\sqrt{s} = 13,6$ TeV

La mesure des sections efficaces inclusives de production des bosons  $W^+$ ,  $W^-$  et  $Z$ , y compris leurs canaux de désintégration en électrons ou muons, leurs rapports, et les rapports des sections efficaces fiduciaires du boson  $W$  avec celle de  $t\bar{t}$ , est effectuée en utilisant les données de collision  $pp$  à l'énergie dans le centre de masse sans précédent de 13,6 TeV au LHC, correspondant à une luminosité intégrée de  $29 \text{ fb}^{-1}$ . Les sections efficaces sont mesurées pour la production de bosons  $Z$  dans un volume fiduciaire correspondant à l'impulsion transverse des leptons  $p_T^\ell > 27 \text{ GeV}$ ,

leurs pseudorapidités  $|\eta_\ell| < 2,5$ , et la masse invariante des deux leptons  $66 < m_{\ell\ell} < 116$  GeV, et pour la production de bosons  $W$  dans une région fiduciaire correspondant à  $p_T^\ell > 27$  GeV,  $|\eta_\ell| < 2,5$ , la masse transverse du boson  $W$   $m_T^W > 50$  GeV, et l'énergie transverse manquante  $E_T^{\text{miss}} > 25$  GeV.

Les contributions des bruits de fond sont minimisées efficacement grâce à des sélections d'objets et d'événements, ce qui aboutit à des contributions relativement faibles. Divers processus de bruit de fond sont pris en compte dans cette analyse, catégorisés en deux classes : les bruits de fond électrofaibles et de quark top, présents dans les canaux  $W$  et  $Z$ , et les bruits de fond multijet (MJ) de QCD. La contribution des bruits de fond multijet de QCD est négligeable dans les canaux du boson  $Z$ , ils sont donc uniquement pris en compte lors de l'analyse des canaux du boson  $W$ . Une méthode basée sur les données est employée pour estimer les bruits de fond multijet de QCD dans les canaux du boson  $W$ , en supposant que les formes des distributions cinématiques du MJ sont relativement stables, indépendamment de l'isolation du lepton. Les événements sont classés en 4 régions selon les critères d'isolation des leptons et les coupures cinématiques sur  $m_T^W$  et  $E_T^{\text{miss}}$ .

La région de signal (SR) est définie en utilisant les mêmes sélections que celles des canaux  $W$  dans le Tableau 5.2, avec des critères stricts d'isolation des leptons, et des coupures spécifiques sur les observables  $m_T^W$  et  $E_T^{\text{miss}}$ . En revanche, la région de fit (FR) utilise le même critère d'isolation mais inverse les coupures cinématiques sur  $m_T^W$  et  $E_T^{\text{miss}}$ , ce qui entraîne une proportion accrue de bruits de fond multijet. Deux régions de contrôle anti-isolées sont définies en inversant l'isolation des leptons. La région de contrôle anti-isolée 1 (CR1) est établie en inversant les coupures sur  $m_T^W$  et  $E_T^{\text{miss}}$ , tandis que la région de contrôle 2 (CR2) utilise les mêmes coupures cinématiques que la SR. Par conséquent, les événements avec des leptons anti-isolés peuplent à la fois CR1 et CR2, les enrichissant de bruits de fond multijet. Le modèle MJ initial est obtenu dans CR1 en supprimant les bruits de fond électrofaibles et de top des données. Ce modèle est ensuite utilisé pour effectuer un ajustement des fractions dans la FR afin de déterminer la fraction MJ. Le résultat de l'ajustement est extrapolé à la SR à l'aide de facteurs de transfert, définis comme le rapport des événements MJ dans CR2 à ceux dans CR1. Une étude dédiée a été menée pour traiter le biais résultant de la dépendance du bruit de fond multijet à l'isolation des leptons.

La technique du profil de vraisemblance est employée pour extraire les sections efficaces fiduciaires inclusives de production des bosons  $W$  et  $Z$  ainsi que leurs rapports. La fonction de profil de vraisemblance est méticuleusement construite en utilisant les taux d'événements obtenus après l'application des sélections d'événements. Les incertitudes sont systématiquement intégrées dans le modèle en tant que paramètres de nuisance. Ces incertitudes tiennent compte de diverses sources de variations systématiques et statistiques, garantissant que les sections efficaces et les rapports extraits sont fiables.

Le modèle statistique est construit sous une forme concrète, comme illustré par

la formule suivante :

$$L(\vec{n}; \mu_s, \vec{\theta}) = \prod_{c \in \text{channels}} \text{Pois}(n_{\text{data}} | \mu_{s,c} S_c(\vec{\theta}) + B_c(\vec{\theta})) \prod_{i \in \text{NPs}} G(\theta_i),$$

dans laquelle  $\mu_s$  est la force du signal, représentant le rapport de la section efficace du signal mesurée sur la valeur prédite. Le rapport des sections efficaces fiduciaires est obtenu de manière similaire, en utilisant une formule de vraisemblance modifiée. Pour le rapport  $R_{\alpha/\beta}$ , qui est le rapport de la section efficace fiduciaire du processus  $\alpha$  sur celui du processus  $\beta$ , la fonction de vraisemblance peut être formulée comme suit :

$$L(\vec{n}; \mu_s, \vec{\theta}) = \prod_{c \in \alpha \text{channels}} \text{Pois}(n_{\text{data}} | R_{\alpha/\beta} \mu_B S_c + B_c(\vec{\theta})) \prod_{c \in \beta \text{channels}} \text{Pois}(n_{\text{data}} | \mu_B S_c + B_c(\vec{\theta})) \prod_{i \in \text{NPs}} G(\theta_i).$$

Les prédictions sont calculées à l'aide de DYTURBO avec une précision NNLO+NNLL en QCD et de RENESANCE pour les corrections électrofaibles (EW) au NLO. Divers PDFs récents sont utilisés pour calculer les prédictions théoriques.

Les résultats des sections efficaces fiduciaires pour la production des bosons  $W^+$ ,  $W^-$  et  $Z$  dans les canaux de désintégration leptoniques sont respectivement  $4250 \pm 150$  pb,  $3310 \pm 120$  pb, et  $744 \pm 20$  pb. Ces incertitudes citées représentent l'incertitude totale, qui inclut l'incertitude de luminosité d'environ 2,2%. Les rapports des sections efficaces fiduciaires mesurés sont :  $R_{W^+/W^-} = 1,286 \pm 0,022$ ,  $R_{W^\pm/Z} = 10,17 \pm 0,25$ , et  $R_{t\bar{t}/W^\pm} = 0,112 \pm 0,003$ , avec les incertitudes représentant l'incertitude totale.

De plus, les sections efficaces totales sont mesurées pour la production de bosons  $Z$  dans la plage de masse  $66 \text{ GeV} < m_{\ell\ell} < 116 \text{ GeV}$  et pour la production du boson  $W$  dans tout l'espace des phases. L'étude comprend les mesures pour  $\sigma(W^+)$ ,  $\sigma(W^-)$ ,  $\sigma(W^\pm)$ ,  $\sigma(Z)$ ,  $R_{W^+/W^-}$ ,  $R_{W^\pm/Z}$ ,  $R_{t\bar{t}/W^\pm}$ ,  $R_{t\bar{t}/W^+}$ , et  $R_{t\bar{t}/W^-}$ .

Les résultats montrent que les sections efficaces mesurées des bosons  $W$  et  $Z$  sont en bon accord avec les prédictions du Modèle Standard. Cependant, les rapports des sections efficaces de  $t\bar{t}$  par rapport au boson  $W$  sont légèrement surestimés par certaines prédictions théoriques.

### Mesure des sections efficaces différentielles des bosons $W^\pm$ avec des données à faible empilement à $\sqrt{s} = 5,02$ et 13 TeV

La mesure des sections efficaces différentielles de production du boson  $W^\pm$  dans les canaux de désintégration leptoniques, incluant à la fois les électrons et les muons, est présentée en utilisant les données de collisions proton-proton à 5,02 et 13 TeV avec le détecteur ATLAS lors de prises des données spéciales au LHC en 2017 et 2018 durant la période de Run 2. Les échantillons de données correspondent à une luminosité intégrée de  $255 \text{ pb}^{-1}$  pour 5,02 TeV et de  $338 \text{ pb}^{-1}$  pour 13 TeV,

respectivement. Les incertitudes correspondantes des luminosités intégrées sont de 1% et 0.9%.

Les sections efficaces différentielles de production du boson  $W$  sont mesurées dans une région fiduciaire correspondant à  $p_T^\ell > 25$  GeV,  $|\eta_\ell| < 2.5$ ,  $m_T^W > 50$  GeV, et  $E_T^{\text{miss}} > 25$  GeV. Les sections efficaces différentielles simples et doubles en  $p_T^\ell$  et  $|\eta_\ell|$  sont présentées, ainsi que l'asymétrie de charge du boson  $W$ .

L'estimation du bruit de fond multijet est effectuée en utilisant une stratégie similaire à celle de la première mesure. Les définitions des régions sont légèrement différentes. Des coupures cinématiques relâchées sont appliquées sur CR1 et FR ( $m_T^W > 0$  GeV et  $E_T^{\text{miss}} > 0$  GeV) afin d'augmenter les statistiques.

Les sections efficaces différentielles sont mesurées en fonction des quantités au niveau vérité. La technique de dépliage (unfolding) est utilisée pour corriger les biais de mesure et les effets de résolution introduits par le détecteur. Dans cette mesure, la méthode de dépliage bayésien est utilisée pour résoudre le problème de dépliage, en extrayant de manière itérative la distribution au niveau vérité à partir de la distribution reconstruite. Le biais (biais de l'hypothèse à priori) de la méthode de dépliage itératif et les incertitudes statistiques sont pris en compte pour déterminer un nombre optimal d'itérations.

Les sections efficaces différentielles simples et doubles sont mesurées en termes de pseudorapidité des leptons  $|\eta_\ell|$  et de l'impulsion transverse des leptons  $p_T^\ell$  dans les canaux électron et muon. La combinaison entre les canaux électron et muon est réalisée en utilisant le meilleur estimateur linéaire non biaisé (BLUE). La fonction  $\chi^2$  est définie comme suit :

$$\chi^2 = (\vec{x} - \vec{\mu})^T \mathbf{C}^{-1} (\vec{x} - \vec{\mu}).$$

Les résultats optimaux de la combinaison sont obtenus en minimisant les différences entre les résultats mesurés et les résultats combinés  $\vec{\mu}$ . Les incertitudes et leurs corrélations sont incorporées dans la fonction  $\chi^2$  en utilisant la matrice de covariance  $\mathbf{C}$ . Le problème de minimisation est résolu de manière itérative afin de réduire le biais dû à l'hypothèse gaussienne dans les incertitudes. Il existe une bonne cohérence entre les différents canaux de saveur de leptons.

Les incertitudes expérimentales typiques dans l'espace des phases global sont inférieures à 1%. Les résultats mesurés sont comparés aux prédictions théoriques en utilisant DYTURBO avec une précision NNLO et NNLL en QCD et des différents PDFs. Un bon accord général entre les deux est observé. Dans la région de haute  $p_T^\ell$ , les prédictions théoriques sont légèrement sous-estimées.

L'impact des mesures sur les PDF est évalué en utilisant la technique de profilage. Cette évaluation repose sur l'hypothèse que les nouvelles données ajoutées sont compatibles avec les prédictions théoriques basées sur les ensembles de PDFs actuels. Les prédictions théoriques sont ajustées aux résultats observés en modifiant les PDFs existants. La fonction  $\chi^2$  est définie pour quantifier l'accord entre les données et les prédictions théoriques en fonction des PDFs et est minimisée en ajustant les PDFs actuels pour obtenir un meilleur accord tout en préservant les contraintes et corrélations

originales. Les informations de grille, qui stockent les sections efficaces partoniques, sont générées en utilisant MADGRAPH5 avec une précision NLO en QCD. Les corrections à NNLO en QCD sont ensuite appliquées. L'impact obtenu varie selon un PDF donné, mais de manière générale, une réduction significative des incertitudes des PDFs pour les distributions de quarks valence et de quarks de mer est observée.

## Résumé

Cette thèse présente deux mesures des sections efficaces de production de bosons vecteurs dans les canaux de désintégration leptoniques. Ces mesures des taux de production de bosons vecteurs testent les prédictions de la QCD perturbative à différentes échelles d'énergie, avec un bon accord observé. La première mesure valide les performances du détecteur, des logiciels et des algorithmes de reconstruction en utilisant le jeu de données du début de Run 3. La mesure de la section efficace différentielle fournit des informations précieuses sur le modèle des partons et soutiendra les analyses futures en contraignant les incertitudes liées aux fonctions de distribution des partons.



# Table des matières

<b>1</b>	<b>Introduction</b>	<b>15</b>
<b>2</b>	<b>Theory</b>	<b>19</b>
2.1	The Standard Model . . . . .	19
2.1.1	Elementary particles . . . . .	19
2.1.2	Electroweak interactions and symmetry breaking . . . . .	21
2.1.3	Quantum chromodynamics . . . . .	25
2.2	Proton–proton collisions . . . . .	25
2.3	Proton structure . . . . .	26
2.4	The Drell-Yan process . . . . .	29
<b>3</b>	<b>The LHC and the ATLAS experiment</b>	<b>35</b>
3.1	The Large Hadron Collider . . . . .	35
3.2	The ATLAS detector . . . . .	36
3.2.1	The ATLAS coordinate system . . . . .	37
3.2.2	Inner detector . . . . .	38
3.2.3	Calorimeter system . . . . .	40
3.2.4	Muon spectrometer . . . . .	44
3.2.5	Magnet system . . . . .	45
3.2.6	Forward detectors . . . . .	47
3.2.7	The ATLAS trigger and data acquisition system . . . . .	47
<b>4</b>	<b>Object reconstruction and identification</b>	<b>51</b>
4.1	Track and vertex . . . . .	51
4.2	Electron and photon . . . . .	52
4.2.1	Electron and photon reconstruction . . . . .	52
4.2.2	Electron and photon energy calibration . . . . .	53
4.2.3	Electron and photon identification . . . . .	54
4.2.4	Electron isolation . . . . .	55
4.3	Muon . . . . .	55
4.3.1	Muon reconstruction and calibration . . . . .	55
4.3.2	Muon identification and isolation . . . . .	56
4.4	Jet . . . . .	57
4.4.1	Jet reconstruction and calibration . . . . .	57
4.4.2	Flavor tagging . . . . .	60
4.5	Missing transverse energy . . . . .	60

<b>5</b>	<b>Measurement of <math>W^\pm</math> and <math>Z</math> production cross sections and their ratios at <math>\sqrt{s} = 13.6</math> TeV</b>	<b>65</b>
5.1	Introduction . . . . .	65
5.2	Data and MC samples . . . . .	66
5.2.1	Data samples . . . . .	66
5.2.2	MC samples . . . . .	66
5.3	Object selection . . . . .	68
5.4	Event Selection . . . . .	69
5.5	Background estimation . . . . .	70
5.5.1	Electroweak and top-quark backgrounds . . . . .	70
5.5.2	Multijet backgrounds for $W^\pm \rightarrow \ell\nu$ events . . . . .	71
5.5.3	Multijet background for $Z \rightarrow \ell\ell$ events . . . . .	81
5.6	Theory predictions . . . . .	81
5.6.1	Electroweak scheme and parameters . . . . .	82
5.6.2	PDFs . . . . .	83
5.6.3	Electroweak corrections . . . . .	84
5.6.4	Uncertainties . . . . .	85
5.6.5	Results . . . . .	86
5.7	Systematic uncertainties . . . . .	87
5.7.1	Experimental uncertainties . . . . .	89
5.7.2	Modelling uncertainties . . . . .	90
5.7.3	Systematic uncertainty correlation in the combination . . . . .	92
5.7.4	Dominant uncertainty sources . . . . .	92
5.8	Kinematic distributions and event yields in the signal region . . . . .	93
5.9	Statistical analysis . . . . .	93
5.9.1	Statistical model . . . . .	94
5.9.2	Asimov fits . . . . .	96
5.9.3	Fit to Asimov data for $Z \rightarrow \ell\ell$ . . . . .	98
5.9.4	Fit to Asimov data for $W^\pm \rightarrow \ell\nu$ , $W^+/W^-$ and $W^\pm/Z$ ratios . . . . .	98
5.9.5	Fit to Asimov data for $t\bar{t}/W^\pm$ ratios . . . . .	99
5.10	Results . . . . .	100
5.10.1	General procedure . . . . .	100
5.10.2	Results of $Z$ -boson cross sections . . . . .	101
5.10.3	Results of $W$ -boson production cross sections . . . . .	102
5.10.4	Results of $\sigma_{W^+}/\sigma_{W^-}$ and $\sigma_{w^\pm}/\sigma_z$ cross-section ratios . . . . .	103
5.10.5	Results of $t\bar{t}$ cross section over $W^\pm$ cross section . . . . .	103
5.10.6	The impact of uncertainty on the results . . . . .	104
5.10.7	Compatibility between electron and muon channels . . . . .	104
5.10.8	Summary of the results and comparison with predictions . . . . .	106
5.11	Conclusion . . . . .	107

<b>6</b>	<b>Measurement of <math>W^\pm</math> boson differential cross sections with low pile-up data at <math>\sqrt{s} = 5.02</math> and 13 TeV</b>	<b>129</b>
6.1	Introduction . . . . .	129
6.2	Data and MC samples . . . . .	130
6.2.1	Data samples . . . . .	130
6.2.2	MC samples . . . . .	130
6.3	Object reconstruction . . . . .	132
6.4	Event selection . . . . .	133
6.5	Background estimation . . . . .	134
6.6	Kinematic distributions . . . . .	139
6.6.1	Kinematic distributions at 5.02 TeV . . . . .	140
6.6.2	Kinematic distributions at 13 TeV . . . . .	141
6.7	Theory predictions . . . . .	143
6.8	Unfolding . . . . .	145
6.8.1	Unfolding procedure . . . . .	147
6.8.2	Binning . . . . .	149
6.8.3	Bias estimation . . . . .	150
6.8.4	Unfolding optimization . . . . .	151
6.9	Uncertainty . . . . .	152
6.9.1	Uncertainty sources . . . . .	152
6.9.2	Uncertainty propagation . . . . .	153
6.9.3	Uncertainty correlations . . . . .	155
6.10	Results . . . . .	156
6.10.1	Combination procedure . . . . .	157
6.10.2	Differential cross sections in $ \eta_\ell $ . . . . .	158
6.10.3	Differential cross sections in $p_T^\ell$ . . . . .	158
6.10.4	Differential cross sections in $p_T^\ell$ and $ \eta_\ell $ . . . . .	159
6.10.5	Charge asymmetry of $W$ bosons . . . . .	159
6.11	PDF profiling . . . . .	159
6.11.1	Methodology . . . . .	159
6.11.2	Results . . . . .	162
6.12	Conclusion . . . . .	163
<b>7</b>	<b>Summary and prospect</b>	<b>189</b>
<b>A</b>	<b>Merged-<math>e^+e^-</math> identification</b>	<b>211</b>
<b>B</b>	<b>Identification efficiency measurement using <math>Z \rightarrow ee</math> tag-and-probe method : <math>Z_{\text{iso}}</math> method</b>	<b>215</b>
<b>C</b>	<b>Auxiliary material for measurement of <math>W^\pm</math> and <math>Z</math> production cross-sections and their ratios at <math>\sqrt{s} = 13.6</math> TeV</b>	<b>219</b>
C.1	The cross section measurement with an event-counting approach . . . . .	219

<b>D</b>	<b>Auxiliary material for measurement of <math>W^\pm</math> boson differential cross sections with low pile-up data at <math>\sqrt{s} = 5.02</math> and 13 TeV</b>	<b>221</b>
D.1	Differential cross sections in the electron and muon channels . . . . .	221
D.1.1	Differential cross sections in $ \eta_\ell $ . . . . .	221
D.1.2	Differential cross sections in $p_T^\ell$ . . . . .	226
D.1.3	Differential cross sections in $p_T^\ell$ and $ \eta_\ell $ . . . . .	230
	<b>Publications</b>	<b>235</b>
	<b>Acknowledgements</b>	<b>237</b>

# 1 - Introduction

The Standard Model (SM) is a theory framework to explain the basic building blocks of matters, the elementary particles, and their interactions. The development of the SM theory throughout the 20 century involved major theoretical progress and discoveries on the experiments. It explained a broad range of experimental observations and has been probed by various experiments over decades. The SM has proven to be a successful theory.

The story of particle physics begins in 1897, when J.J. Thomson discovered the electron [1]. Then other particles, proton was discovered by Ernest Rutherford in 1917 [2] and neutron was discovered by James Chadwick in 1932 [3]. These discoveries offered insights into the atomic structure and opened the door to explore the subatomic particles. Consequently, many elementary particles have been discovered, such as positron, muon, etc. In the meantime, the development of quantum mechanics and quantum electrodynamics provided the theoretical framework to describe the behavior of particles at atomic and subatomic level.

In 1954, Chen-Ning Yang and Robert Mills proposed the Yang-Mills theory [4], which introduced the non-Abelian gauge fields and became the foundation of the modern particle physics theory. In 1960s, the electroweak unification was developed by Sheldon Glashow [5], Steven Weinberg [6] and Abdus Salam [7], which unifies the electromagnetic and weak interactions. Murray Gell-Man and George Zweig developed the quark model [8, 9], providing a theoretical framework to describe the structure and interactions of hadrons. In 1964, Peter Higgs, François Englert, and Robert Brout proposed the Higgs mechanism [10-13] to explain how matter acquires masses. In 1970s, the quantum chromodynamics (QCD) theory is proposed to describing the strong interactions between quarks and gluons governed by SU(3) gauge symmetry. On the experimental side, with the advancement of accelerator and detector techniques, there was an explosion of new particle discoveries during this period. The experiments on the Stanford Linear Accelerator Center (SLAC) [14, 15] showed the evidence of the existence of quarks in 1970s. In 1983, the  $W^\pm$  and  $Z$  bosons were discovered at the CERN SPS collider [16-19], which provided the confirmation of the electroweak theory. In 1995, the top quark was discovered by the CDF and D0 experiments [20, 21], completing the last piece of the quarks within the Standard Model theoretical framework. In 2012, the ATLAS and CMS experiments discovered the Higgs boson at the Large Hadron Collider (LHC) at CERN, confirming the Higgs mechanism. The discovery of the Higgs boson marked the identification of the final missing element of the SM.

The LHC stands as the largest and most powerful particle accelerator ever built by humankind. The ATLAS detector is one of the multiple-purposed particle detectors installed at the LHC. A large variety of physics phenomenon can be studied with the ATLAS detector, providing rigorous tests of the current SM theory and enabling the

search for new physics beyond the SM (BSM).

The main topic of this thesis is related to the carriers of the weak force,  $W$  and  $Z$  bosons. Since their discovery at the CERN SPS collider, their properties have been actively studied over decades. These studies help to achieve a better understanding of the electroweak theory and the potential deviations from the SM predictions may reveal BSM physics. The  $W$  and  $Z$  production processes exhibit large cross sections at the LHC and are often studied in the leptonic channels, as presented in this thesis, due to their clear experimental signatures. The production cross sections, also known as the production rates, are important properties of the  $W$  and  $Z$  bosons. Precision measurement of the cross sections of  $W$  and  $Z$  boson productions is an essential benchmark tool to test the quantum chromodynamics and offers insights into the structure of hadrons. Accurate knowledge of  $W$  and  $Z$  boson cross section measurements can also benefit other physics analyses. These processes often serve as important background components in various measurements, so precise cross section data can help reduce uncertainties arising from  $W$  and  $Z$  boson modeling. The cross sections can be interpreted in QCD analyses to determine the parton distribution functions (PDFs).

During the Run 3 data-taking, the LHC achieves the record-breaking centre-of-mass energy of 13.6 TeV in proton-proton collisions. It provides a unique opportunity to perform the measurement of vector boson production cross sections at the unprecedented energy. Therefore, measurements of the inclusive cross sections of the vector boson productions are performed using the Run 3 datasets collected by the ATLAS experiment. As these processes are well understood, such measurement can help to validate the quality of Run 3 dataset at the early stage. The dependence on the centre-of-mass energies is also checked in this measurement at the new energy regimes. The ratios of cross sections are measured in this analysis as well. Many of the experimental uncertainties cancel out when obtaining the cross section ratios, for example, the luminosity uncertainty, which is dominant uncertainty in the cross section measurement. Therefore, the ratios offer the tests to the predictions with high accuracy.

The other measurement presented in this thesis is about the differential cross section of the  $W$  boson production in leptonic decay channels. The differential cross section provides sensitivity to the structure of hadrons. The rapidity of the lepton final states reflects the ratio of momentum fractions carried by the partons. By measuring the differential cross sections, the PDFs of the quarks involved in the interactions can be determined. During special runs at the LHC, datasets were recorded at centre-of-mass energies of 5.02 and 13 TeV. These runs featured a low number of interactions per bunch crossing (low pile-up), resulting in reduced background noise. These datasets benefit the cross section measurement by allowing for high precision results.

I also contributed to the measurement of electron and photon efficiencies in the LHC Run 2 and the search for new resonances in multiple final states involving a high transverse momentum  $Z$ -boson at  $\sqrt{s} = 13$  TeV.

This thesis is organised as follows. Chapter 2 provides a brief introduction to the theoretical background for the studies presented in this thesis, including the Standard Model, proton structure, and Drell-Yan processes. Chapter 3 presents the Large Hadron Collider (LHC) and the ATLAS detector. Physical objects in the ATLAS experiment are described in Chapter 4, discussing the information about their reconstruction, calibration and corrections. Two cross section measurements are presented in the later chapters. Chapter 5 details the first measurement of the inclusive cross section of  $W$  and  $Z$  boson productions, as well as their ratios, at the unprecedented centre-of-mass energy of 13.6 TeV. Chapter 6 presents the analysis measuring the differential cross section of  $W$  boson productions using the low pile-up datasets at 5.02 TeV and 13 TeV, respectively. Chapter 7 summarizes this thesis. Appendix A describes the merged- $ee$  identification I developed in the resonance search analysis [22]. Appendix B describes the measurement of the electron identification efficiencies.



## 2 - Theory

### 2.1 The Standard Model

The Standard Model (SM) of particle physics is a well-established theoretical framework that describes elementary particles and their interactions. Established in the 20th century, it has been remarkably successful in explaining a broad range of experimental observations and serves as the cornerstone of modern particle physics.

In general, the SM theory provides a framework for describing the fundamental interactions of particles : electromagnetic, weak, and strong interactions. The fundamental particles in the Standard Model are classified into two categories : fermions and bosons. Fermions form the building blocks of matter, while bosons act as carriers that mediate interactions. The SM is formulated as a Quantum Field Theory (QFT) and is based on local gauge symmetry :

$$SU(3)_C \times SU(2)_L \times U(1)_Y \quad (2.1)$$

where  $SU(3)_C$  is the gauge group of Quantum Chromodynamics (QCD) that governs the strong interaction, and the  $C$  subscript denotes color charge in QCD. On the other hand,  $SU(2)_L \times U(1)_Y$  is the gauge group that describes the electroweak (EW) interaction, where  $L$  and  $Y$  denote the left-handed fermion and weak hypercharge, respectively.

In QFT, the Lagrangian is used to describe the dynamics of fields and particles. Therefore, the Standard Model can be formulated as a Lagrangian. The full SM Lagrangian can be presented as :

$$\mathcal{L}_{SM} = \mathcal{L}_{EW} + \mathcal{L}_{QCD}. \quad (2.2)$$

The EW sector associates the unification of the electromagnetic and weak interactions while the QCD sector is responsible for strong interactions.

In this section, the elementary particles in the SM are described and the dynamics of the SM are explained in terms of the electroweak and QCD sectors.

#### 2.1.1 Elementary particles

As mentioned in the previous section, the fundamental particles in the Standard Model theory consists of two categories : fermions and bosons. Figure 2.1 displays the fundamental particles in the SM theory framework.

Fermions are the elementary particles following the Fermi-Dirac statistics. They are characterized by possessing half-integer spins ( $\frac{1}{2}$ ,  $\frac{3}{2}$ , etc.), and thus follow the Pauli exclusion principle. Every fermion is paired with an anti-fermion that has opposite charges and quantum numbers, yet shares the same mass. When a fermion meets its paired anti-fermion, they can annihilate each other, converting their mass into energy in the form of particles. Fermions in the SM comprises quarks and leptons. Quarks

## Standard Model of Elementary Particles

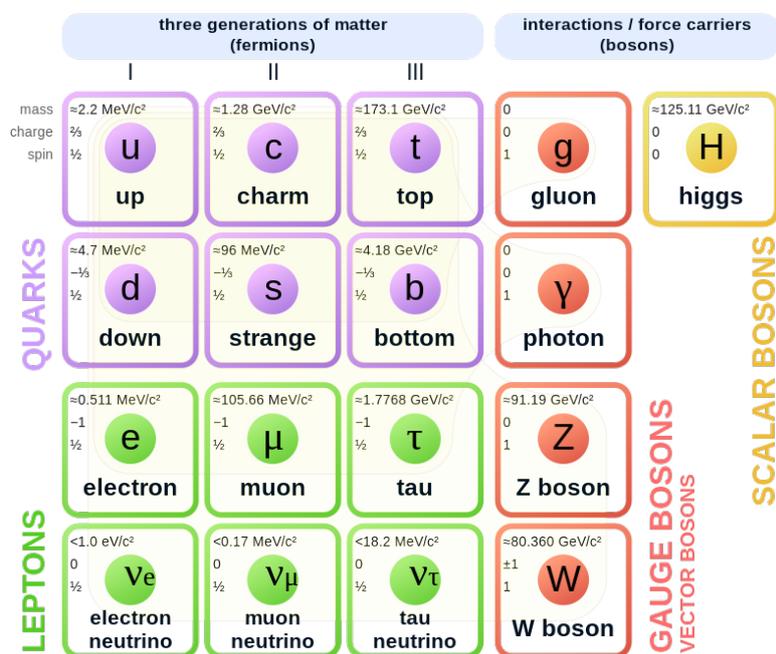


FIGURE 2.1 – The overview of the elementary particles in the SM theory. The particle properties are obtained from Review of Particle Physics in 2019 [23].

form the building blocks of hadrons, for example, protons and pions. As illustrated in the Figure 2.1, the Standard Model includes six types of quarks : up quark ( $u$ ), down quark ( $d$ ), charm quark ( $c$ ), strange quark ( $s$ ), top quark ( $t$ ) and bottom quark ( $b$ ). The six quarks are divided into three generations and each generation consists of two types of quarks. In each generation, one quark ( $u$ ,  $c$  or  $t$  quark) has an electric charge of  $+\frac{2}{3}$ , while the other ( $d$ ,  $s$  or  $b$  quark) has an electric charge of  $-\frac{1}{3}$ . Quarks carry color charges, allowing them to participate in strong interactions, and they can also engage in electroweak interactions. Leptons are classified into two types : electron ( $e$ ), muon ( $\mu$ ), and tau ( $\tau$ ) leptons, which have an electric charge of  $-1$ , and their paired neutrinos ( $\nu_e$ ,  $\nu_\mu$  and  $\nu_\tau$ ), which are electrically neutral. Charged lepton participate in both electromagnetic and weak interactions, while neutrinos, being neutral, only engage in weak interactions. Similar to the quarks, leptons are also divided into three generations and each one contains a charged lepton and its corresponding neutrino. In practice, the term “lepton” often specifically refers to the charged leptons ( $e$ ,  $\mu$  and  $\tau$ ), while the term “neutrino” is used to denote the neutral leptons ( $\nu_e$ ,  $\nu_\mu$  and  $\nu_\tau$ ). A mass hierarchy is evident among fermion generations, with a notable increase in mass from the first generation to the second and third generations, the mass of the first-generation lepton, the electron, is approximately 0.511 MeV, whereas the mass of the second-generation lepton, the muon, is around 105.7 MeV [24].

Bosons follow Bose-Einstein statistics. They have integer spin values (0, 1, 2, etc.) and can occupy the same quantum state. In the SM theory, bosons act as carriers responsible for mediating the fundamental interactions. Bosons are classified into two categories based on their role and properties : Gauge bosons are the mediators of the interactions while Higgs boson plays a crucial role in the mechanism that gives mass to other elementary particles. Gauge bosons consist of photons ( $\gamma$ ), gluons ( $g$ ), and the  $W$  ( $W^+$  and  $W^-$ ) and  $Z$  bosons. Photons are massless, electrically neutral particles that mediate the electromagnetic force. The  $W$  and  $Z$  bosons are massive and mediate the weak interaction;  $W$  bosons are electrically charged, while  $Z$  bosons are neutral. Gluons, which carry the color charge, mediate the strong interaction and can interact with both quarks and other gluons. Higgs bosons are electrically neutral and massive. Unlike other elementary bosons, they do not mediate fundamental interactions. Instead, they generate mass for other particles via the Higgs mechanism.

### 2.1.2 Electroweak interactions and symmetry breaking

The electroweak sector  $\mathcal{L}_{EW}$  unifies the electromagnetic and weak interactions. The Lagrangian of the electroweak sector can be written in terms of :

$$\mathcal{L}_{EW} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{fermion}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}. \quad (2.3)$$

It consists of the gauge fields  $\mathcal{L}_{\text{gauge}}$  for the  $SU(2)_L \times U(1)_Y$  symmetry, fermion field term  $\mathcal{L}_{\text{fermion}}$ , Higgs field term  $\mathcal{L}_{\text{Higgs}}$  and Yukawa term  $\mathcal{L}_{\text{Yukawa}}$ .

The  $SU(2)_L$  gauge group is responsible for the weak isospin symmetry and involves three gauge bosons,  $W_\mu^i$  (where  $i = 1, 2, 3$ ), which interact with left-handed

fermions. The  $U(1)_Y$  symmetry group associates to the weak hypercharge symmetry and involves the gauge boson,  $B_\mu$ .

The gauge Lagrangian,  $\mathcal{L}_{\text{gauge}}$  can be presented as :

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}W_{\mu\nu}^i W_i^{\mu\nu} - \frac{1}{4}B_{\mu\nu}W^{\mu\nu} \quad (2.4)$$

where  $W_{\mu\nu}^i$  is the field strength tensor of the  $SU(2)_L$  gauge bosons, defined as :

$$W_{\mu\nu} = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g\epsilon_{ijk}W_\mu^j W_\nu^k \quad (2.5)$$

and  $B_{\mu\nu}$  is the field strength tensor for the  $U(1)_Y$ , which can be formulated as :

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \quad (2.6)$$

in Eq. (2.5),  $g$  represents the coupling constant for the  $SU(2)_L$  interaction, and  $\epsilon_{ijk}$  denotes the structure constant of the  $SU(2)$  representation, represented as a completely anti-symmetric tensor. The  $SU(2)_L$  gauge bosons mix with the  $U(1)_Y$  to form the observable physical gauge bosons in nature. This results in three massive mediators  $W^+$ ,  $W^-$  and  $Z$  bosons, as well as a massless boson, the photon ( $\gamma$ ). This mixing happens through the Higgs mechanism, which generates the masses of the  $W$  and  $Z$  bosons.

The Higgs field term  $\mathcal{L}_{\text{Higgs}}$  takes the form :

$$\mathcal{L}_{\text{Higgs}} = (D^\mu\Phi)^\dagger D_\mu\Phi - V(\Phi) \quad (2.7)$$

where  $\Phi$  is the Higgs doublet, which is a complex vector in the  $SU(2)_L$  space and can be written as :

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (2.8)$$

where  $\phi^+$  and  $\phi^0$  represent the charge and neutral components, respectively. The initial term in Eq. (2.7),  $(D^\mu\Phi)^\dagger D_\mu\Phi$ , outlines the dynamics of the Higgs field, with  $D_\mu$  representing the gauge covariant derivative that incorporates interactions with the gauge fields :

$$D_\mu = \partial_\mu + i\frac{g}{2}\tau^a W_\mu^a - i\frac{g'}{2}Y B_\mu. \quad (2.9)$$

In this context,  $g$  and  $g'$  are the coupling constants of the  $SU(2)$  and  $U(1)$  gauge fields, respectively. The symbol  $\tau^a$  are the Pauli matrices and  $Y$  is the hypercharge field. The second term in Eq. (2.7),  $V(\phi)$ , is the potential term, which can be written as :

$$V(\Phi) = \mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2 \quad (2.10)$$

where  $\mu^2$  can be negative, leading to spontaneous symmetry breaking, and  $\lambda$  denotes the self-coupling constant, which is positive. The symmetry breaking will be discussed in detail later.

The fermion sector is responsible for the interaction of fermions with the electro-weak gauge field. The fermion Lagrangian,  $\mathcal{L}_{\text{fermion}}$  is written as :

$$\mathcal{L}_{\text{fermion}} = \bar{\phi}_L i\gamma^\mu D_\mu \phi_L + \bar{\phi}_R i\gamma^\mu D_\mu \phi_R \quad (2.11)$$

where  $\phi_L$  and  $\phi_R$  denote the left-handed fermion doublet and right-handed fermion singlet. The covariant derivative for the doublet is presented in Eq. (2.9). The covariant derivative for the doublet is presented in Eq. (2.9), and the corresponding derivatives for  $\phi_R$  can be expressed as follows :

$$D_\mu \phi_R = (\partial_\mu + ig' \frac{Y}{2} B_\mu) \phi_R. \quad (2.12)$$

Fermions gain mass via Yukawa interactions with the Higgs field after spontaneous symmetry breaking, which can be represented as :

$$\mathcal{L}_{\text{Yukawa}} = -y_u \bar{q}_L \tilde{\Phi} u_R - y_d \bar{q}_L \Phi d_R - y_e \bar{l}_L \Phi e_R + h.c. \quad (2.13)$$

where  $y_u$ ,  $y_d$  and  $y_e$  are Yukawa coupling matrices for up-type quarks, down-type quarks and charged leptons ( $e$ ,  $\mu$ ,  $\tau$ ), and  $\Phi$  is the Higgs doublet introduced before and  $\tilde{\Phi}$  is its charge-conjugated counterpart.

Equation (2.7) introduces the spontaneous symmetry breaking within the electroweak sector, facilitated by the Higgs mechanism. Through its potential term, the Higgs field gains a non-zero vacuum expectation value, which breaks the electro-weak  $SU(2)_L \times U(1)_Y$  symmetry and provides mass to the  $W$  and  $Z$  bosons, as well as fermions. The photon remains massless, reflecting the unbroken electromagnetic symmetry.

The potential term of the Higgs field is presented in Eq. (2.10), formulated in terms of Higgs doublet  $\Phi$ . The Higgs doublet, as expressed in Eq. (2.8), can further be represented in terms of real scalar fields. The neutral component is given by  $\phi^0 = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$  and the charged components can be written as  $\phi^\pm = \frac{1}{\sqrt{2}}(\phi_3 \pm i\phi_4)$ . Therefore, the Higgs potential can take the form :

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 = \mu^2 |\phi|^2 + \lambda |\phi|^4$$

$$|\phi|^2 = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2)$$

where the  $|\phi|^4$  term reflects self interaction with an intensity denoted by  $\lambda$ , which is positive to ensure that the total field energy is bounded from below. The vacuum state of the system, representing the lowest energy state, is Lorentz invariant. Therefore,  $\phi$  is considered to be a constant in this vacuum state. If  $\mu^2 > 0$ , the minimum energy is achieved when  $\phi = 0$ . If  $\mu^2 < 0$ , the ground state is achieved when  $|\phi|^2 = -\frac{\mu^2}{2\lambda}$ , which breaks the  $SU(2) \times U(1)$  gauge symmetry. Figure 2.2 illustrates the potential energy  $V(\Phi)$ , which has a ‘Mexican hat’ shape, and the vacuum state corresponds to a circle of minima in the complex plane.

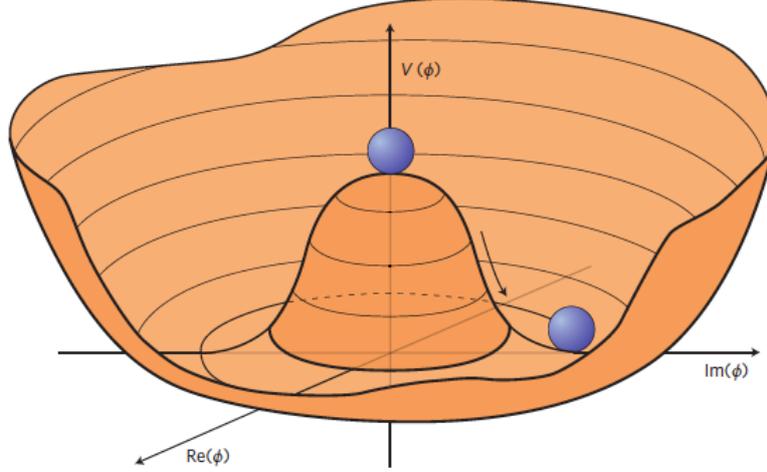


FIGURE 2.2 – The 'Mexican hat' potential induces spontaneous symmetry breaking, with the vacuum state, or lowest-energy state, corresponding to a randomly selected point along the base of the hat's brim [25].

The gauge symmetry  $SU(2) \times U(1)$  permits the selection of a specific vacuum state for the Higgs field. The vacuum expectation value (VEV)  $\langle \Phi \rangle$  can be represented as :

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (2.14)$$

where  $v = \sqrt{-\mu^2/\lambda}$ . The neutral component  $\phi_0$  acquires a vacuum expectation value.

After spontaneous symmetry breaking, the Higgs field  $\Phi$  can be expressed as :

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad (2.15)$$

where  $H(x)$  denotes the physical Higgs boson.

After the Higgs field acquires a vacuum expectation value, the  $W$  and  $Z$  bosons gain mass through the interaction between the Higgs field and gauge fields. By substituting the covariant derivative and the vacuum state into the kinetic term of the Higgs sector in the Lagrangian, the Higgs sector can be expressed as follows :

$$\mathcal{L}_{\text{Higgs}} = \frac{1}{2} \frac{g^2 v^2}{4} (|W_\mu^+|^2 + |W_\mu^-|^2) + \frac{1}{2} \frac{v^2}{4} |g' B_\mu - g W_\mu^3|^2 - V(\Phi). \quad (2.16)$$

The charged gauge bosons  $W^\pm$  are defined as  $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp W_\mu^2)$ . Consequently, the mass of the  $W$  boson is  $M_W = \frac{gv}{2}$ . The neutral bosons, the  $Z$ -boson and photon ( $A$ ) are a mixture of  $W^3$  and  $B$  :

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \quad (2.17)$$

where the  $\theta_W$  is the weak mixing angle, which is defined as  $\tan \theta_W = \frac{g'}{g}$ . Consequently, the  $Z$ -boson is defined as  $\frac{1}{\sqrt{g^2+g'^2}}(gW_\mu^3 + B_\mu)$ , with its mass being  $M_Z = \frac{v}{2}\sqrt{g^2 + g'^2}$ . In contrast, the photon  $A_\mu$  is not included in the Eq. (2.16) and remains massless, corresponding to the unbroken  $U(1)_{\text{EM}}$  symmetry.

For fermions, the Yukawa interactions result in mass terms when the Higgs field acquires the vacuum expectation value. The mass is given by  $m_f = \frac{1}{2}g_f v$ , where  $g_f$  is Yukawa coupling constant.

### 2.1.3 Quantum chromodynamics

The Quantum Chromodynamics (QCD) theory describes the strong interaction between quarks and gluons based on the  $SU(3)_C$  symmetry group, where  $C$  denotes the color charge in QCD. The Lagrangian of QCD can be formulated as :

$$\mathcal{L}_{\text{QCD}} = \sum_f \bar{\psi}_f (i\gamma^\mu D_\mu - m_f) \psi_f - \frac{1}{4} F_{\mu\nu}^a F^{a\mu\nu} \quad (2.18)$$

where the subscript  $f$  denotes the quark flavour,  $\psi_f$  denotes the quark field,  $m_f$  is the mass of the quark, and  $D_\mu$  is the covariant derivative including the interaction with gluons, while  $F_{\mu\nu}^a$  denotes the field strength tensor for the gluon field, which are given by :

$$D_\mu = \partial_\mu - ig_s T^a G_\mu^a$$

$$F_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_\mu^b G_\nu^c$$

where  $G_\mu^a$  are the gluon fields and the subscript  $a$  represents gluons, which carry both color and anti-color, forming eight distinct types through various combinations of color-anti-color pairs,  $g_s$  is the coupling constant, which is related to the QCD coupling constant  $\alpha_s$  through the equation  $\alpha_s = \frac{g_s^2}{4\pi}$ , and  $\alpha_s$  is not a fixed value and depends on the energy scale  $Q$ , therefore,  $\alpha_s$  is often parameterized as  $\alpha_s(Q^2)$ , a phenomenon known as the running of the coupling constant. As the energy scale increases and the distance decreases, the coupling constant reduces, leading to a decrease in the strength of interaction. This property is called asymptotic freedom, which describes how quarks and gluons interact weakly at high energies.

## 2.2 Proton–proton collisions

At the Large Hadron Collider (LHC), a wide range of particle physics phenomena is studied through high-energy proton–proton ( $pp$ ) collisions. As the proton consists of quark and gluons, which are collectively called partons, the  $pp$  collisions can be described in terms of interactions between individual partons from each proton. The  $pp$  collisions involves two primary processes : the hard scattering process and the interaction of proton fragments.

Protons are complex particles composed of valence quarks, sea quarks, and gluons. Specifically, two up ( $u$ ) quarks and one down ( $d$ ) quark define the proton's properties. Sea quarks are virtual quark-anti-quark pairs arising from quantum fluctuations. Gluons are responsible for binding the quarks together and for interacting with each other. The hard scattering refers to interactions between partons that involve large momentum transfer, often resulting in the production of high-energy particles. These processes can be effectively described by perturbative Quantum Chromodynamics (pQCD), as the strong coupling constant decreases at high momentum transfer, making calculations feasible.

Radiation, which involves the emission of additional particles such as gluons or photons, can occur both before and after the hard scattering events. If the emission happens before the hard scattering, the process is called initial state radiation (ISR). If it occurs after the scattering, it is referred to as final state radiation (FSR). The emitted partons lead to a series of successive gluon emissions, known as a parton shower. The partons eventually hadronized to form stable hadrons, resulting in the creation of jets. The interaction of proton fragments also occurs, resulting in processes, which are not associated with the primary vertex, known as underlying events. Figure 2.3 illustrates the  $pp$  collisions using the example of  $t\bar{t}H$  production, which involves several parton-level processes. The hard scattering can be well described with pQCD, while the soft interactions are primarily dominated by non-perturbative effects.

### 2.3 Proton structure

The parton model is a theoretical framework that explains the internal structure of protons through their constituent partons. Protons contain three valence quarks (two up quarks and one down quark) which determine the quantum numbers of a proton. The valence quarks are bound together by the exchange of gluons. In addition to valence quarks, protons also contain a sea of virtual quark-anti-quark ( $q\bar{q}$ ) pairs due to the quantum fluctuations of gluons. As mentioned in Section 2.2, the  $pp$  collision is described in terms of interactions between partons. Each parton carries a fraction of momentum  $x$ , allowing the partonic center-of-mass energy  $\sqrt{\hat{s}}$  to be defined as  $\hat{s} = x_1 x_2 s$ , where  $x_1$  and  $x_2$  are momentum fractions of two incoming partons involved in the interaction, and  $x$  is a key variable introduced by Bjorken in the hadron structure study [27]. Parton distribution functions (PDFs)  $f_p(x, Q)$  is defined to describe the probability of finding a parton with Bjorken  $x$  at a given energy scale  $Q^2$ .

The DGLAP (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi) evolution equations [28-30] govern the evolution of PDFs with the energy scale  $Q^2$  :

$$\frac{\partial}{\partial \ln(Q^2)} \begin{pmatrix} q_i(x, Q^2) \\ g(x, Q^2) \end{pmatrix} = \frac{\alpha_s(Q^2)}{2\pi} \sum_j \int_x^1 \frac{d\xi}{\xi} \begin{pmatrix} P_{q_i q_j}(\frac{x}{\xi}, \alpha_s(Q^2)) P_{q_i g}(\frac{x}{\xi}, \alpha_s(Q^2)) \\ P_{g q_j}(\frac{x}{\xi}, \alpha_s(Q^2)) P_{g g}(\frac{x}{\xi}, \alpha_s(Q^2)) \end{pmatrix} \quad (2.19)$$

where  $P_{ij}(\frac{x}{\xi}, \alpha_s(Q^2))$  represents the probability of a parton  $i$  splitting into a parton  $j$

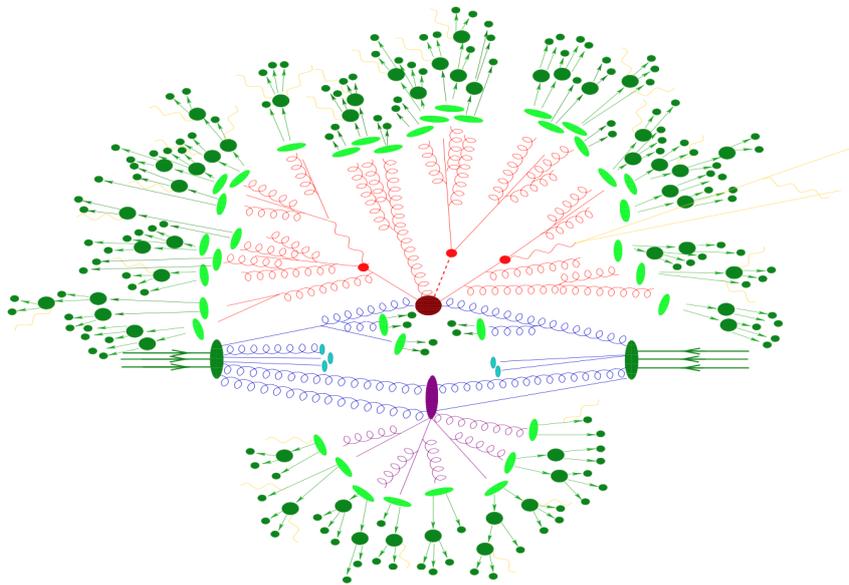


FIGURE 2.3 – A  $t\bar{t}h$  event produced by an event generator. The hard interaction point is marked as big red blob. The produced top quarks and the Higgs bosons are marked using small red blobs. Additional hard QCD radiation is emitted (red), and a secondary interaction occurs (purple blob) before the final state partons hadronize (light green blobs) and the hadrons decay (dark green blobs) [26].

by emitting a parton,  $g$  and  $q$  are the distributions of gluon and quark, and  $\xi$  represents the momentum fraction.

The factorization theorem allows the cross section to be represented as a convolution of the parton distribution functions (PDFs) and the partonic cross sections. The factorization theorem distinguishes between long-distance, non-perturbative effects encapsulated in PDFs and short-distance, perturbative effects represented by hard scattering cross sections. The cross sections of hard scattering process can be factorized as :

$$\sigma = \sum_{i,j} \int dx_1 dx_2 f_i^p(x_1, \mu_F^2) f_j^p(x_2, \mu_F^2) \hat{\sigma}_{ij}(x_1, x_2, \alpha_s(\mu_R), \mu_F^2) \quad (2.20)$$

where  $\hat{\sigma}_{ij}$  represents the parton-level cross section. The subscript  $i$  and  $j$  indicate the functions run over all possible combinations of parton  $i$  and  $j$ .  $\mu_F$  is the factorization scale, where PDFs are evaluated and collinear divergences are absorbed.  $\mu_R$  is the renormalization scale, which is used to renormalize the strong coupling constant  $\alpha_s$ , accounting for its running behavior.  $\mu_F$  and  $\mu_R$  serve as an independent parameters theoretically, allowing for a clearer distinction between the non-perturbative effects in PDFs and the perturbative effects in the hard scattering cross section.  $\mu_F$  reflects the collinear divergences in the PDFs, while  $\mu_R$  handles ultraviolet (UV) divergences in strong coupling constant. In practice,  $\mu_F$  and  $\mu_R$  depend on the process and are often set to be at the same scale of  $Q$ , for example, in the Drell-Yan process, as  $Q^2$  is the invariant mass of the lepton pair final states and the  $\mu_F^2 = \mu_R^2 = Q^2$ .

To achieve better precision, next-to-leading order (NLO) corrections in pQCD are crucial. These corrections incorporate additional parton emissions and loop diagrams, resulting in more accurate predictions of cross sections and kinematic distributions. Equation (2.20) can be expanded as a series in the strong coupling constant  $\alpha_s$  :

$$\sigma = \sum_{i,j} \int dx_1 dx_2 f_i^p(x_1, \mu_F^2) f_j^p(x_2, \mu_F^2) (\hat{\sigma}_{ij}^{\text{LO}} + \alpha_s(\mu_F) \hat{\sigma}_{ij}^{\text{NLO}} \alpha_s^2(\mu_F) \hat{\sigma}_{ij}^{\text{NNLO}} + \mathcal{O}(\alpha_s^3)).$$

At NLO, the cross section involves one-loop virtual corrections and real gluon emission. At next-to-next-to-leading order (NNLO), two-loop virtual corrections, one-loop real corrections and double real emissions are considered.

PDFs are essential tools for making theoretical predictions and derived by performing global fits using measured cross sections from various experiments. The cross sections are predicted using the factorization theorem, as Eq. (2.20) shows. The dependence on the energy scale is governed by the DGLAP equations. Consequently, a global fit is performed using the measured cross sections. The global fit involves a broad range of experimental data from various physics processes and multiple experiments. The input PDFs are parameterized at the initial scale  $\mu_{F0}$ , usually selected to be low to allow for reliable use of pQCD, yet sufficiently high to prevent non-perturbative effects. Take HERAPDF2.0 [31] as an example : the starting scale  $\mu_{F0}$  is set to be 1.9 GeV. PDFs generically take the following analytical form in HERAPDF

determination :

$$xf_i(x) = A_i x^{B_i} (1-x)^{C_i} (1 + D_i x + E_i x^2)$$

where  $A_i, B_i, C_i, D_i$  and  $E_i$  are fitting parameters.  $A_i$  is an overall normalization factor and it is constrained by sum-rules.  $B_i$  and  $C_i$  constrain PDFs at low and high  $x$ , respectively.  $D_i$  and  $E_i$  offer flexibility in shaping the distribution across the entire  $x$ -range. The parameterization of two valence quark PDFs ( $xu_v$  and  $xd_v$ ), the gluon PDF ( $xg$ ) and two anti-quark PDFs ( $x\bar{U}$  and  $x\bar{D}$  for  $u$ - and  $d$ -type anti-quarks, respectively) are considered. At  $\mu_{F0}$ , it is assumed that  $x\bar{U} = x\bar{u}$ , while  $x\bar{D} = x\bar{d} + x\bar{s}$ . An additional term is incorporated in the gluon PDF, which is formulated as :  $xg(x) = A_g x^{B_g} (1-x)^{C_g} - A'_g x^{B'_g} (1-x)^{C'_g}$ . The negative term enhances the parameterization's flexibility at low- $x$ . The DGLAP equations are used to evolve the PDFs to other scales. A  $\chi^2$  value is defined by comparing the theoretical predictions with experimental measurements, and it is minimized considering the uncertainties of both sides.

The experimental uncertainties of the PDF determination are evaluated using approaches such as the Hessian method [32] or a Monte Carlo (MC) method [33]. The Hessian method estimates uncertainties by evaluating the covariance matrix of the fitting parameters, while the Monte Carlo method relies on an ensemble of data set replicas, created by fluctuating the central values within their uncertainties.

PDFs are crucial for understanding the proton structure, making precise predictions at high-energy regimes. Additionally, PDFs ensure global consistency across various experiments and processes and help in comprehending QCD evolution and scaling violations. They also play a vital role in validating and advancing theoretical models in QCD and beyond. The accurate and reliable determination of PDFs is critical for the success of the high-energy physics.

The current parton distribution functions (PDFs) are provided by various research groups, and several widely used sets are introduced in this section. The CT18 PDF set [34], from the CTEQ-TEA collaboration, incorporates a wide range of experimental data and is frequently used for precision QCD tests. The MSHT20 set [35-37], developed by the MSHT collaboration, is similarly based on extensive experimental input and is used in precision calculations for LHC physics. NNPDF sets [38], determined using neural networks to minimize theoretical assumptions, are popular in new physics searches, with the most recent being NNPDF4.0. The ATLASpdf21 set [39] uses data primarily from the ATLAS experiment and gives an LHC-specific perspective that complements global fits.

## 2.4 The Drell-Yan process

The Drell-Yan (DY) process [40] is an essential mechanism describing the production of a lepton pair resulting from the quark and anti-quark annihilation. A typical DY process involves two partons producing a boson, e.g. a virtual photon  $\gamma^*$  or a

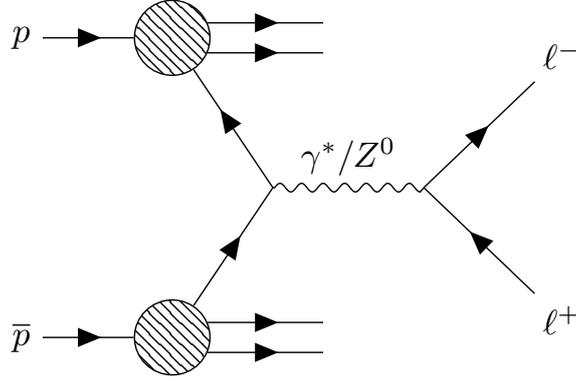


FIGURE 2.4 – The leading-order Feynman diagram for the Drell-Yan process in  $pp$  interactions.

$Z$ -boson, which then decays into a lepton pair, represented as  $q\bar{q} \rightarrow \ell^+\ell^-$ . Figure 2.4 illustrates the leading-order Feynman diagram for Drell-Yan process in  $pp$  collisions.

The quark and anti-quark possess a range of center-of-mass energies,  $\hat{s}$ , which is related to the total collision energy  $s$ , through the equation  $\hat{s} = x_1x_2s$ . The differential cross section for the process can be represented in terms of the invariant mass of the lepton pair  $M$  at the parton level :

$$\frac{d\hat{\sigma}(q\bar{q} \rightarrow \ell^+\ell^-)}{dM^2} = \frac{4\pi\alpha^2}{3M^2} \frac{1}{N} Q_q^2 \delta(\hat{s} - M^2) \quad (2.21)$$

where  $1/N = 1/3$ , accounting for the average of the colors of the initial-state quarks. The cross section of the DY production in the parton model can be calculated following Eq. (2.20) :

$$\begin{aligned} \frac{d\sigma}{dM^2} &= \int_0^1 dx_1 dx_2 \sum_q (f_q(x_1)f_{\bar{q}}(x_2) + f_q(x_2)f_{\bar{q}}(x_1)) \frac{d\hat{\sigma}(q\bar{q} \rightarrow \ell^+\ell^-)}{dM^2} \\ &= \frac{4\pi\alpha^2}{3NM^2} \int_0^1 dx_1 dx_2 \delta(x_1x_2s - M^2) \sum_q Q_q^2 (f_q(x_1)f_{\bar{q}}(x_2) + f_q(x_2)f_{\bar{q}}(x_1)) . \end{aligned}$$

The sum runs over quark flavors, with the additional  $q\bar{q}$  contributions explicitly specified. The overall cross section depends on the PDFs, which are independent of the process scale in the parton model. By defining scaling variable  $\tau = M^2/s$ , the cross section can be written as :

$$\begin{aligned} M^4 \frac{d\sigma}{dM^2} &= \frac{4\pi\alpha^2}{3N} \int_0^1 dx_1 dx_2 \delta(x_1x_2s - \tau) \sum_q Q_q^2 (f_q(x_1)f_{\bar{q}}(x_2) + f_q(x_2)f_{\bar{q}}(x_1)) \\ &= \frac{4\pi\alpha^2}{3N} \tau \mathcal{F}(\tau) . \end{aligned}$$

The DY cross section shows the dependence on the scaling variable  $\tau$  in this equation.

The rapidity of the lepton pair reflects the ratio of the momentum fractions carried by each of the partons involved in the interaction. It can be represented as :

$$y = \frac{1}{2} \ln \left( \frac{x_1}{x_2} \right)$$

therefore, one can calculate the double differential cross sections in lepton pair mass and rapidity :

$$\frac{d^2\sigma}{dM^2 dy} = \frac{4\pi\alpha^2}{3N_s M^2} \sum_q Q_q^2 (f_q(x_1)f_{\bar{q}}(x_2) + f_q(x_2)f_{\bar{q}}(x_1)) . \quad (2.22)$$

The double differential cross sections with respect to the rapidity and mass of the lepton pair reveals internal structure of hadrons. PDFs of quark  $q$  and anti-quark  $\bar{q}$ ,  $f_q(x)$  and  $f_{\bar{q}}(x)$ , can be determined by measuring this cross section.

Considering the leading-order DY process  $q\bar{q} \rightarrow \ell^+\ell^-$ , the parton-level cross section can be expressed as :

$$M^4 \frac{d\hat{\sigma}}{dM^2} = \frac{4\pi\alpha^2}{3N} \tau \hat{\mathcal{F}}(\tau) \quad (2.23)$$

where  $\hat{\mathcal{F}} = Q^2\delta(1 - \tau)$ .  $\hat{\mathcal{F}}$  can be expanded perturbatively in terms of powers of the strong coupling constant  $\alpha_s$  :

$$\hat{\mathcal{F}}(\tau) = \hat{\mathcal{F}}_0(\tau) + \frac{\alpha_s}{2\pi} \hat{\mathcal{F}}_1(\tau) + \dots . \quad (2.24)$$

At NLO, three contributions are considered :

- virtual gluon corrections : include one-loop diagrams where a gluon is exchanged between the quark and anti-quark,
- real gluon emission : refers to the process in which an additional gluon is emitted in the final state,
- quark-gluon scattering : is a process in which a quark from one proton interacts with a gluon from another proton.

The collision energy at the LHC is sufficient to produce on-shell  $W$  and  $Z$  bosons. As the mediators of the weak force,  $W$  and  $Z$  bosons have been extensively studied in research. The leading-order production cross section of the  $q\bar{q} \rightarrow Z$  production takes the form :

$$\hat{\sigma}(q\bar{q} \rightarrow Z) = \frac{\pi}{3} \sqrt{2} G_F M_Z^2 (v_q^2 + a_q^2) \delta(\hat{s} - M_Z^2) \quad (2.25)$$

with  $G_F = \frac{\sqrt{2}}{8} \frac{g^2}{M_W^2}$ , which is the Fermi constant,  $v_q$  and  $a_q$  are the vector and axial-vector couplings of the quark  $q$  to the  $Z$ -boson, respectively.

The leading-order production cross section of the  $q\bar{q}' \rightarrow W$  production takes the form :

$$\hat{\sigma}(q\bar{q}' \rightarrow W) = \frac{\pi}{3} \sqrt{2} G_F M_W^2 |V_{qq'}|^2 \delta(\hat{s} - M_W^2) \quad (2.26)$$

with  $V_{qq'}$  is the CKM matrix element.

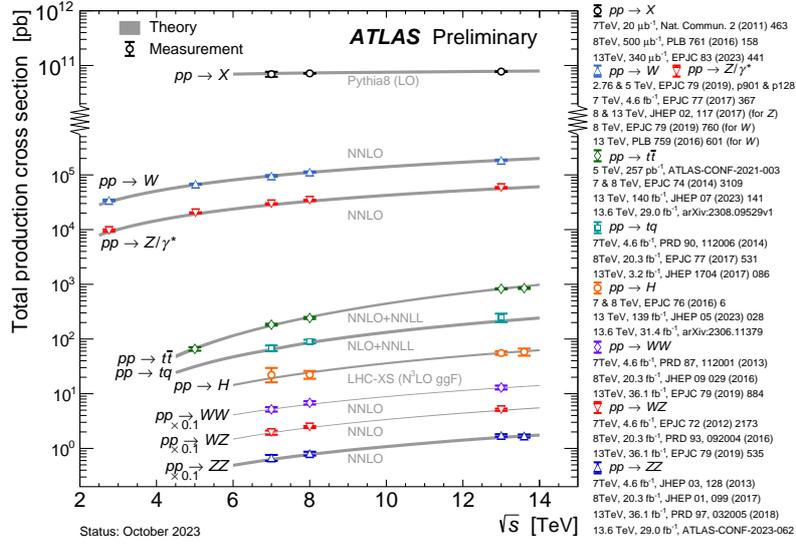


FIGURE 2.5 – Total production cross section measured by ATLAS at different centre-of-mass energies [41].

Figure 2.5 displays the total production cross section by ATLAS at different centre-of-mass energies. The  $W$  and  $Z$  boson productions exhibit large cross sections.

Leptonic decays of  $W$  and  $Z$  bosons are essential processes in the Standard Model, providing clear experimental signatures that are vital for testing electroweak theory and exploring new physics. The production rate of the leptonic final states can be calculated by incorporating the branching ratio (BR), which represents the proportion of a particle's total decays that lead to a particular final state. Therefore, the cross section can be represented as :

$$\hat{\sigma}(q\bar{q} \rightarrow Z \rightarrow \ell\ell) = \hat{\sigma}(q\bar{q} \rightarrow Z) \times \text{BR}(Z \rightarrow \ell\ell),$$

$$\hat{\sigma}(q\bar{q}' \rightarrow W \rightarrow \ell\nu) = \hat{\sigma}(q\bar{q}' \rightarrow W) \times \text{BR}(W \rightarrow \ell\nu).$$

The branching ratios of the leptonic final states are formulated as :

$$\text{BR}(Z \rightarrow \ell\ell) = \frac{G_F M_Z^3}{6\sqrt{2}\pi\Gamma_Z} (|v_\ell|^2 + |a_\ell|^2),$$

$$\text{BR}(W \rightarrow \ell\nu) = \frac{G_F M_W^3}{6\sqrt{2}\pi\Gamma_W}$$

where  $\Gamma_Z$  and  $\Gamma_W$  are the total decay width of  $Z$  and  $W$  bosons, respectively.

The  $W$  and  $Z$  boson production cross-section measurements have been performed across many experiments using the leptonic decay channels. Figure 2.6 presents a comparison between theoretical predictions and the measured results from various experiments as a function of the center-of-mass energy. The comparison offers insights

into the alignment between theoretical models and experimental data across different energy scales.

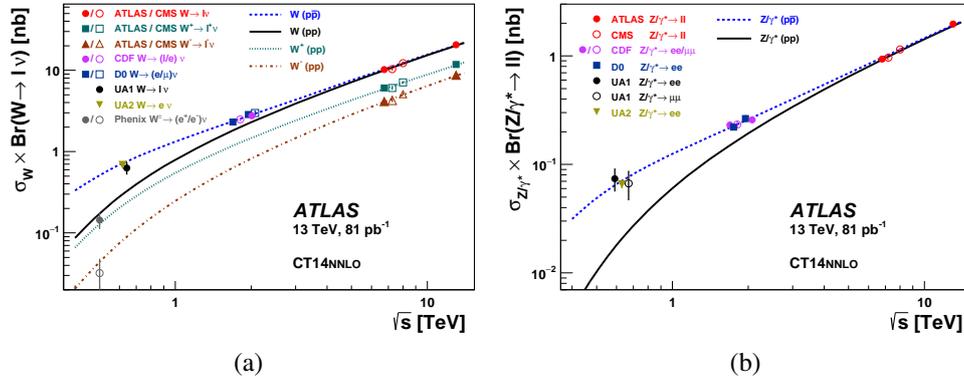


FIGURE 2.6 – The measured cross sections of (a)  $\sigma_W \times Br(W \rightarrow l\nu)$  and (b)  $\sigma_Z \times Br(Z \rightarrow ll)$  as a function of the center-of-mass energy. The measured results are compared with the theoretical predictions. [42]

The Drell-Yan process is also sensitive to different quark flavors. Therefore, the cross sections of such process are vital tools to probe the structure of hadrons, particularly to determine PDFs.



## 3 - The LHC and the ATLAS experiment

### 3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [43] stands as the largest and most powerful particle accelerator worldwide. It is installed in a 26.7 km long circular tunnel situated at CERN (European Organization for Nuclear Research) near Geneva, Switzerland. The LHC comprises two rings of superconducting magnets to increase the energy of protons or heavy ions as they traverse the ring. It accelerates particles to approach the speed of light, thereby simulating the conditions believed to have existed shortly after the Big Bang.

The LHC accelerates particles through a series of steps :

- Initially, protons are generated from hydrogen gas by stripping electrons.
- These protons are accelerated in a linear accelerator (Linac2) to an energy level of 50 MeV.
- Protons are then injected into the Proton Synchrotron Booster (PSB) and accelerated to 1.4 GeV.
- Protons enter the Proton Synchrotron (PS) to be accelerated to 26 GeV.
- Protons are transferred to the Super Proton Synchrotron (SPS), where they are boosted to 450 GeV.
- In the final step, protons are injected into the main LHC ring. In the LHC, protons are accelerated to their final energy, which reaches up to 6.8 TeV per beam during the Run-3 data-taking, corresponding to a centre-of-mass-energy of 13.6 TeV. The designated energy of the LHC is 7 TeV per proton. This acceleration is done with the help of 1232 superconducting dipole magnets bending the particle trajectories.

Two beams of protons traveling in opposite directions are made to collide at four designated interaction points around the LHC ring, where four detectors (ATLAS, CMS, ALICE, and LHCb) are installed. Figure 3.1 displays the layout of the CERN accelerator complex in 2022.

The event rate, the number of events generated by the LHC per second, can be represented as the product of the instantaneous luminosity ( $\mathcal{L}$ ) with the cross section ( $\sigma$ ) for a specific collision process of interest as :  $R = \mathcal{L} \times \sigma$ . In particle physics, luminosity serves as a vital benchmark for assessing the performance of an accelerator. It quantifies the particle collision rate, reflecting the intensity and efficiency of the accelerator operation.

The luminosity in the LHC is affected by many factors and can be expressed in the form :

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{4\pi\sigma_x\sigma_y} \quad (3.1)$$

where  $N_1$  and  $N_2$  are the number of protons in each beam,  $f$  represents the revolution

## The CERN accelerator complex *Complexe des accélérateurs du CERN*

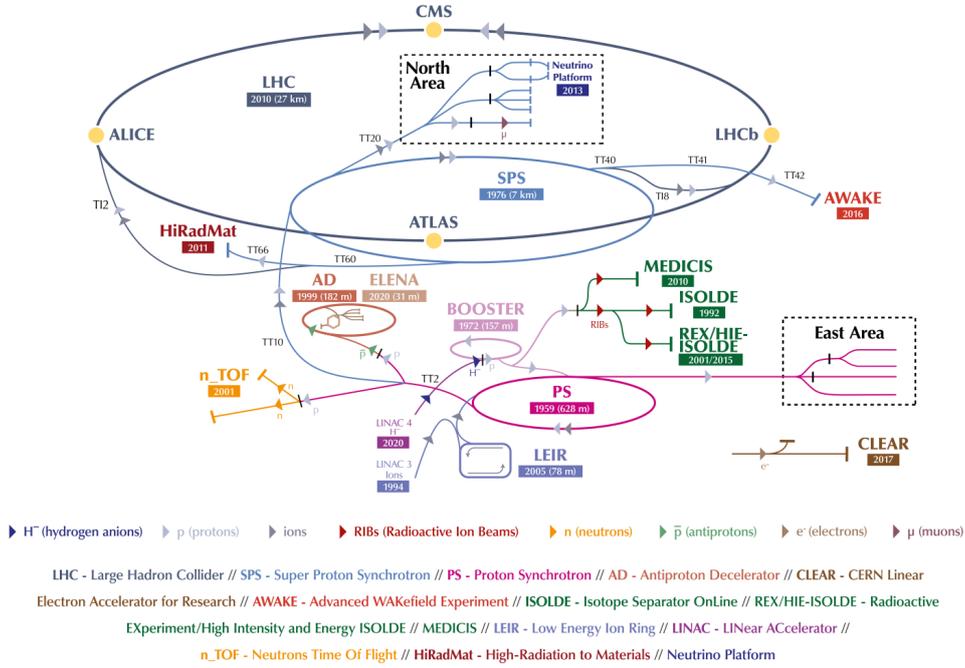


FIGURE 3.1 – The layout of the CERN accelerator complex in 2022. The largest blue ring is the LHC. Taken from Ref. [44].

frequency indicating how frequently the particle completes a full revolution around the collider,  $N_b$  represents the number of bunches in each beam, while  $\sigma_x$  and  $\sigma_y$  denote the size of the beam along the horizontal and vertical directions at the collision point, respectively. The LHC was designed to achieve a peak luminosity of  $1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , however, with continuous upgrades and improvements in the beam dynamics and focusing techniques, the LHC achieved a peak luminosity about  $2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ , which is a factor of two higher than the designed value. When counting the total amount of data collected over a certain period, the integrated luminosity  $\mathcal{L}_{\text{int}}$  is used, typically in unit of inverse cross sections, such as inverse femtobarns ( $\text{fb}^{-1}$ ) or inverse picobarns ( $\text{pb}^{-1}$ ) used in the analyses described in this thesis.

### 3.2 The ATLAS detector

The ATLAS detector (A Toroidal LHC ApparatuS) [45] is one of the four main detectors located at Point 1 at the LHC near CERN. Together with CMS [46], they are two general-purpose detectors in particle physics, which means they are constructed to explore a wide range of physics.

The ATLAS is a massive, cylindrical apparatus, measuring about 25 meters in length and 44 meters in height, with a total weight of around 7,000 tons. It is the largest volume particle detector ever built.

The ATLAS detector consists of a variety of sub-detectors, and In a nutshell, it is composed of six distinct detecting subsystems arranged around the interaction point, which are designed to record the trajectory, momentum, and energy of particles, making it capable of identifying and measuring individual particle precisely. Figure 3.2 displays the overall structure. From the inside out, the inner detector is surrounded by a thin magnetic system generating 2 T magnetic field. Adjacent to it are electromagnetic and hadron calorimeters, along with a muon spectrometer. The outermost side is a massive magnet system bending the charged particle trajectories, enabling the accurate measurement of their momenta.

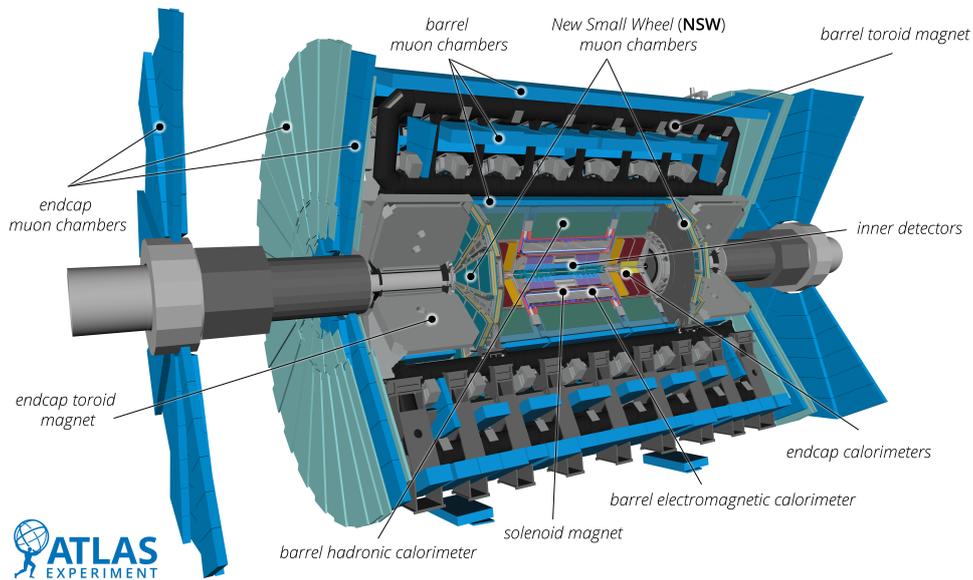


FIGURE 3.2 – The overall structure of the ATLAS detector. Taken from Ref. [47].

### 3.2.1 The ATLAS coordinate system

The coordinate system used by the ATLAS experiment is a right-handed one with its origin positioned at the collision point. In this system :

- The  $z$ -axis is along with the beam direction,
- the  $x$  axis extends toward the center of the LHC ring, and
- the  $y$ -axis points upwards.

The azimuthal angle ( $\phi$ ) provides a measure of the rotation around the  $z$ -axis within  $x$ - $y$  plane, starting from the positive  $x$ -axis. The polar angle ( $\theta$ ) is measured from the positive  $z$ -axis, indicating the direction of a particle in relation to the beamline.

Complementing the angular parameters, the pseudorapidity  $\eta$ , is defined as :

$$\eta = -\ln \tan \left( \frac{\theta}{2} \right)$$

is utilized to describe the direction of a particle. It offers a convenient way to describe the trajectory and energy distribution of particles.

### 3.2.2 Inner detector

The inner detector (ID) plays a crucial role in precisely measuring the direction, momentum, and charge of charged particles produced in the interaction point [48]-[49]. Figure 3.3 displays the layout of the inner detector. From the inside out, the Pixel detector is the innermost component within the Inner detector, surrounded by the Semiconductor Tracker (SCT). The Transition Radiation Tracker (TRT) forms the outermost layer of this detection system.

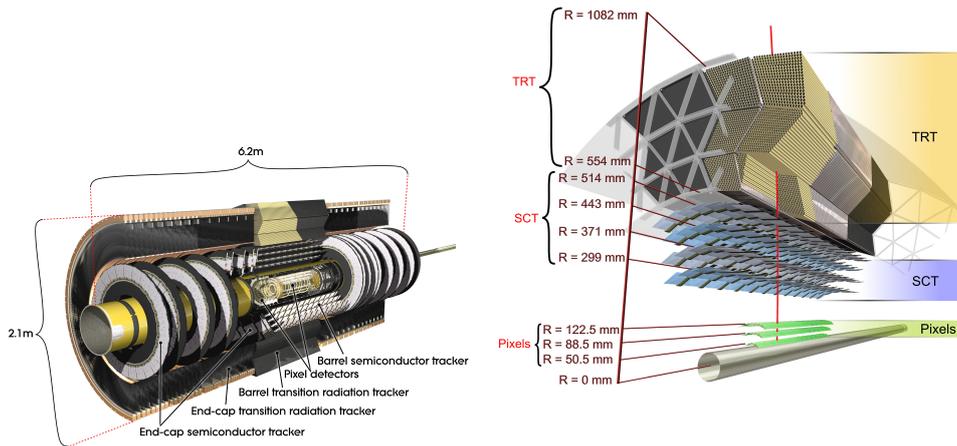


FIGURE 3.3 – The layout of the Inner detector (ID) taken from Ref [50].

### Pixel Detector

The pixel detector is the innermost layer of the inner detector, situated close to the interaction point. It consists of pixel sensors with pixels arranged in a grid pattern. During the shutdown of the LHC in 2013-2014 before Run-2, the pixel detector underwent an upgrade, an insertable Barrel-layer (IBL) was installed between the original pixel detector and a new small-radius beam pipe [51], which aims to maintain or enhance the robustness and performance addressing the higher instantaneous luminosity (the peak luminosity larger than  $2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ ) after the shutdown.

The pixel detector comprises four pixel sensor layers, including the IBL. The innermost part, IBL, is installed only 33.5 mm away from the beam line. For the IBL, pixel size is  $50 \times 250 \mu\text{m}^2$  while it is  $50 \times 400 \mu\text{m}^2$  for other three layers. As shown

in Figure 3.4, the pixel detector has coverage in the region where  $|\eta| < 2.5$  and full coverage in  $\phi$ . Except for the IBL, each of the other three pixel layers consists of a cylindrical layer spanning around 1.4 meters along with the beamline direction in the barrel region and a layer in the end-cap region to extend the coverage.

The pixel detector detects the energy deposited by the charged particles from the collision point via interaction with the silicon atoms. Thanks to the small size of its pixel, it provides high spatial resolution, ensuring the precise measurement of the charged particle trajectories.

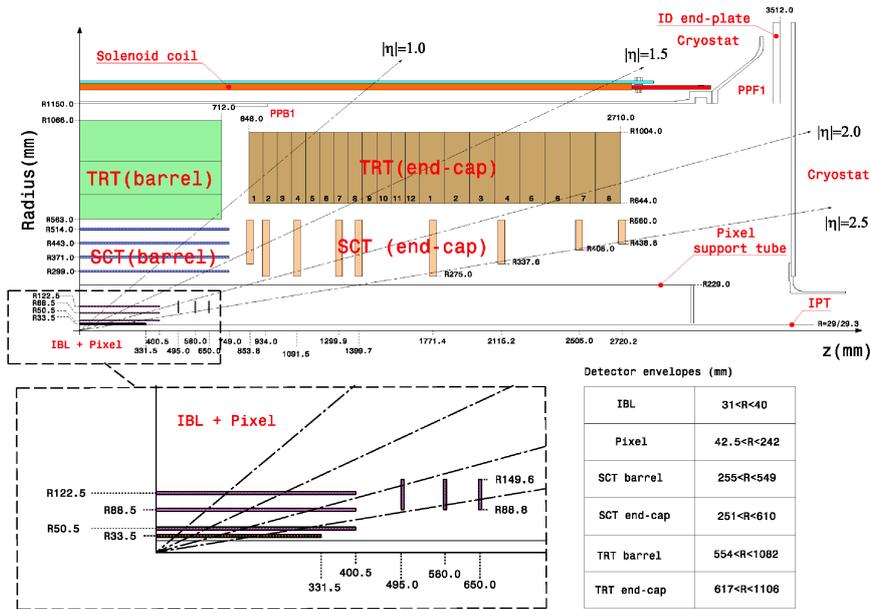


FIGURE 3.4 – The layout of the ATLAS Inner detector (ID) from Ref [51]. The magnified section in the bottom left plot illustrates the layout of the pixel detector.

## Semiconductor Tracker

The semiconductor tracker (SCT) surrounding the pixel detector is a tracking detector made of silicon strips. It comprises four concentric cylindrical layers around the beam axis in the barrel region and two end-cap layers perpendicular to the beam axis. It has coverage in the region with  $|\eta| < 2.5$ . The SCT comprises around 6000 silicon microstrip modules and more than 6 million readout strips. Each particle burst out from the interaction point travels across at least 4 layers of silicon sensors with the optimization of the layout, ensuring the high spatial resolution, with a precision of approximately  $17 \mu\text{m}$  in the transverse plane and  $580 \mu\text{m}$  along the beam line. It allows the SCT capable to provide detailed and precise tracking information.

## Transition radiation tracker

The transition radiation tracker (TRT) is the outermost layer of the ATLAS tracking system. The TRT makes use of the transition radiation to perform electron identification and measure the momentum of charged particles. The transition radiation happens when charged particles travel through materials with different dielectric constants, e.g. the boundary between gas and radiator materials. The TRT is made up of 50 000 straws in the barrel region, which are 144 cm in length, and 250 000 straws in the end-cap regions with 39 cm in length. It covers the region with  $|\eta| < 2.0$ . The TRT consists of 300 000 straws (thin-walled drift tubes), with 4 mm in diameter and a  $30 \mu\text{m}$  gold-plated tungsten wire at its center. The straws are filled with a gas mixture thus when the charged particle travel through the TRT, it ionizes the gas and the ionization signals are detected.

### 3.2.3 Calorimeter system

The calorimeter system is used to absorb the majority of the particles (most of the known particles except muons and neutrinos) originating from the collision and measure their energy deposited in the calorimeter. The ATLAS calorimeter system consists of the electromagnetic calorimeter (ECAL) and hadronic calorimeter (HCAL). The electromagnetic calorimeter absorbs and measures the energy of electrons and photons, while the hadronic calorimeter measures the energy of hadrons. The schematic of the ATLAS calorimeter is displayed in Figure 3.5, the calorimeter surrounds the tracking system and covers the region with pseudorapidity  $|\eta| < 4.9$ .

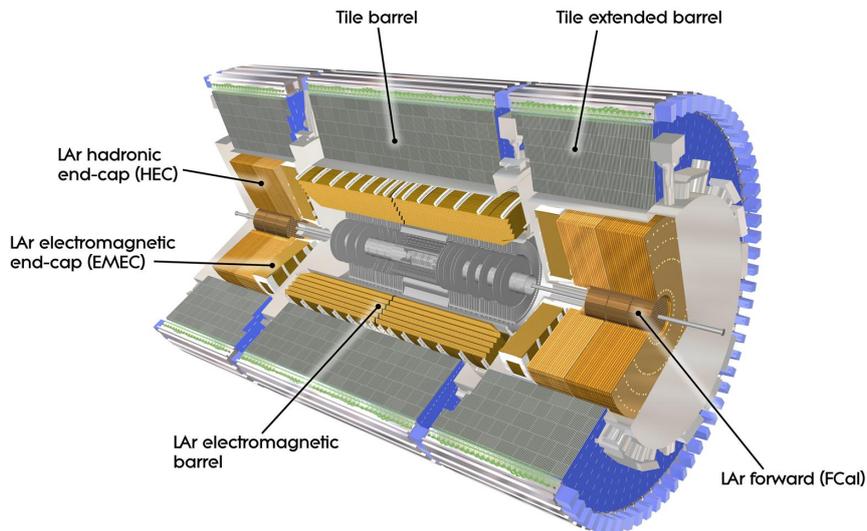


FIGURE 3.5 – The schematic of the ATLAS calorimeter system [52].

## Liquid Argon electromagnetic calorimeter

The liquid argon electromagnetic calorimeter [53] samples the energy of electrons, photons and hadrons coming from the interaction point. It consists of one barrel and two end-cap calorimeters. The barrel calorimeter is a cylindrical apparatus covering the pseudorapidity range  $|\eta| < 1.475$  and full  $\phi$  range, while the electromagnetic end-cap calorimeters (EMEC) has coverage in the region with  $1.375 < |\eta| < 3.2$  in the forward and backward regions of the detector. Each end-cap section of the calorimeter comprises two wheels. The larger wheel extends across the  $|\eta|$  range of 1.375 to 2.5, while the smaller wheel covers the  $|\eta|$  range from 2.5 to 3.2.

When the particles from the collision point enter the electromagnetic calorimeter, they interact with the absorber material and generate a shower of secondary particles. These secondary particles ionize the liquid argon and generate the electron-ion pairs along their path. With strong electric field between the electrodes, the ionized electrons are drifted to the electrodes and this electric current is proportional to the energy deposited by the particle. The detected signal is amplified and digitalized to measure the energy and position of the particle. Figure 3.6 shows the accordion structure of the LAr electromagnetic calorimeter. The utilization of accordion-shaped electrodes in calorimeters ensures uniform coverage and significantly reduces the presence of dead zones.

The thickness of the calorimeter is a crucial parameter in ensuring that particles deposit their entire energy within the calorimeter. The thickness of the LAr electromagnetic calorimeter is more than 22 radiation length ( $X_0$ ) in the barrel region and  $24 X_0$  in the end-cap section, where the radiation length is a parameter characterizing the interaction of the high-energy electrons and photons with materials.

The section with pseudorapidity range  $|\eta| < 2.5$  consists of one pre-sampler layer and three sampling layers, and measures precisely the energy and position of the particles. The pre-sampler layer is the innermost part which is used to correct the energy loss of the particles before they enter the calorimeter. The first layer is with fine granularity ( $\Delta\eta \times \Delta\phi = 0.003 \times 0.1$ ) to provide the precise position information. The second layer of the calorimeter is designed to be significantly thicker to effectively absorb electromagnetic showers and accurately measure the primary fraction of the energy and the granularity is  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ . The third layer of the calorimeter is designed to measure the end tail of the electromagnetic shower, thereby ensuring complete energy capture. It has more coarse granularity of  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ .

The resolution of the calorimeter is described in the form :

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad (3.2)$$

where  $\frac{a}{\sqrt{E}}$  is a stochastic term accounting for fluctuations in the electromagnetic shower,  $\frac{b}{E}$  is the noise term for electronic or background noise, and  $c$  is the constant reflecting imperfection of the calibration and other detector effects that are inde-

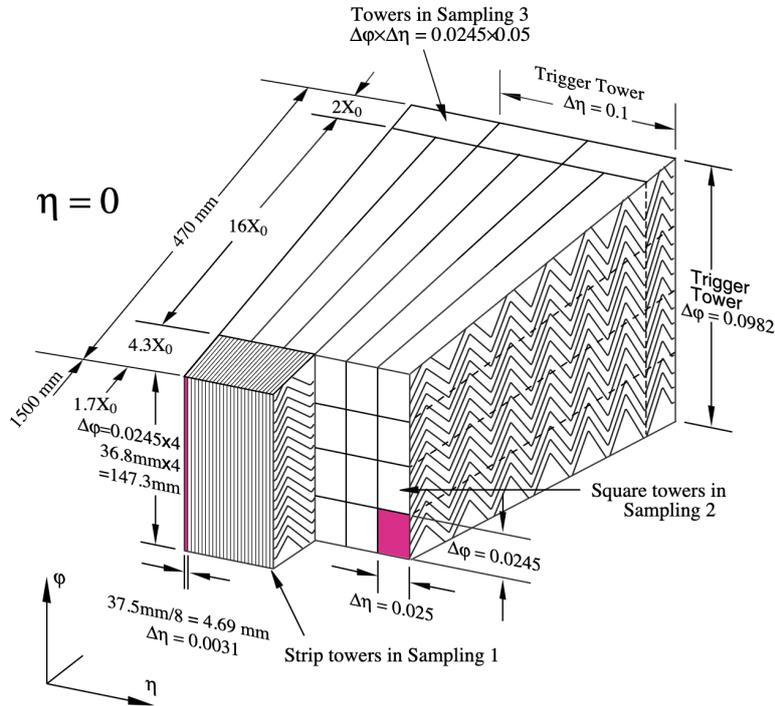


FIGURE 3.6 – The accordion structure of the Liquid Argon calorimeter from Figure 1-2 in Ref [53].

pendent of the energy. The typical energy resolution for electrons in the ATLAS electromagnetic calorimeter is :  $a \approx 10\%$ ,  $b \approx 0.3 \text{ GeV}$ , and  $c$  is less than  $0.7\%$ .

### Hadronic calorimeter

The ATLAS hadronic calorimeter is located between the electromagnetic calorimeter and the muon spectrometer covering the pseudorapidity region with  $|\eta| < 5$ . It is made up of three sub-detectors : the tile calorimeter, hadronic end-cap calorimeters and forward calorimeters. The end-cap and forward calorimeters are equipped with liquid argon (LAr), designed to perform both electromagnetic and hadronic energy measurements.

The tile calorimeter (TileCal) [54] covers the central region, with  $|\eta| < 1.7$ . It consists of three sections : one central section covering the pseudorapidity range  $|\eta| < 1.0$  and two extended barrel sections covering the range  $0.8 < |\eta| < 1.7$ . As shown in Figure 3.7, the tile calorimeter comprises steel and plastic scintillator plates. When a hadron enters the calorimeter, it interacts with the steel absorbers, inducing a shower of secondary particles. The plastic scintillators, positioned between the steel plates, produce photons in response to these secondary particles. These photons are

subsequently converted into an electric current, with a magnitude proportional to the energy of the incoming incident particles. The tile calorimeter achieves an energy resolution for pions of  $\sigma_E/E = 50\%/E \oplus 3\%$ .

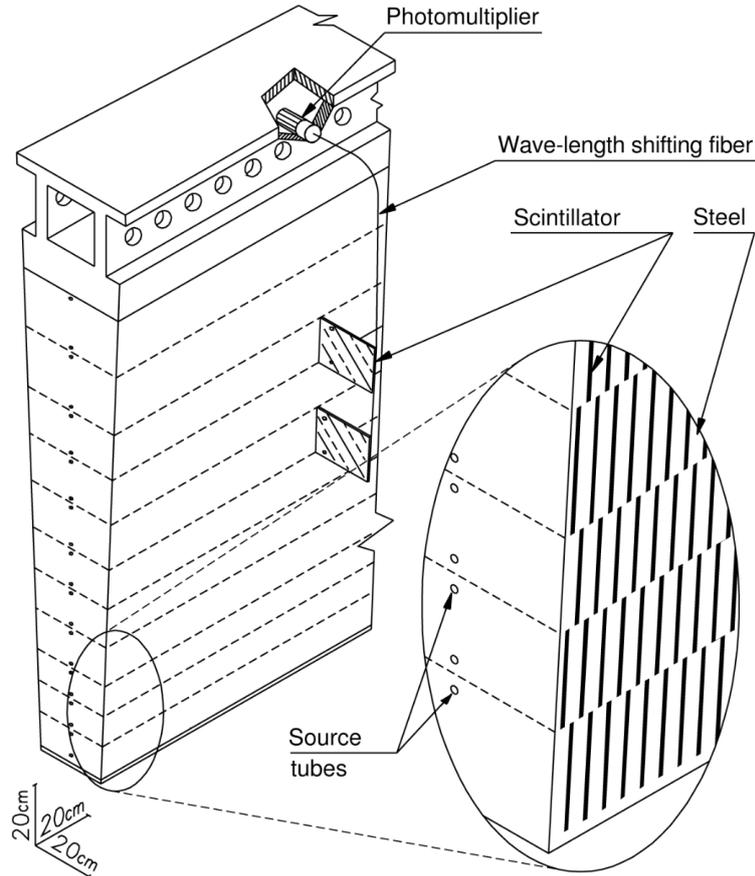


FIGURE 3.7 – The schematics of the tile calorimeter module in Ref. [55] illustrating the assembly of plastic scintillator and steel plates.

The liquid argon hadronic end-cap calorimeter (HEC) [53] is located behind the LAr electromagnetic end-cap calorimeter (EMEC), covering the pseudorapidity range  $1.5 < |\eta| < 3.2$ . Each end-cap consists of two wheels, both constructed with copper plates serving as the absorber material. These copper plates are alternated with liquid argon gaps. The first wheel has a thickness of 0.8 m built from 25 mm copper plates, while the second wheel has a thickness of 1.0 m constructed from 50 mm copper plates.

The liquid argon forward calorimeter (FCal) [53] is situated outside the HEC and covers the forward region with  $3.1 < |\eta| < 4.9$ . It comprises two end-caps and each section consists of three modules. The absorber material of the first module is copper to measure the electromagnetic energy precisely. The other two modules employ

tungsten as the absorber material, offering a high interaction length for hadronic particles.

### 3.2.4 Muon spectrometer

The ATLAS muon spectrometer (MS) [56] is the outermost sub-detector of the ATLAS detector. It is designed for the muon identification and precise measurement of the muon momentum. Muons are highly penetrating particles capable of traversing significant amounts of material with minimal interaction, enabling them to pass through the thick layers of the ATLAS calorimeters. High-momentum muons in final states serve as clear and robust experimental signatures for many physics phenomena at the LHC, therefore, it is crucial to precisely measure muons. The magnetic field provided by the magnet described in Section 3.2.5 bends the trajectory of the muons allowing the tracking and momentum measurement. The muon spectrometer covers  $|\eta| < 2.7$ , and for different regions, sub-detectors with different chamber techniques are used to meet different requirements. The layout of the ATLAS muon spectrometer is displayed in Figure 3.8.

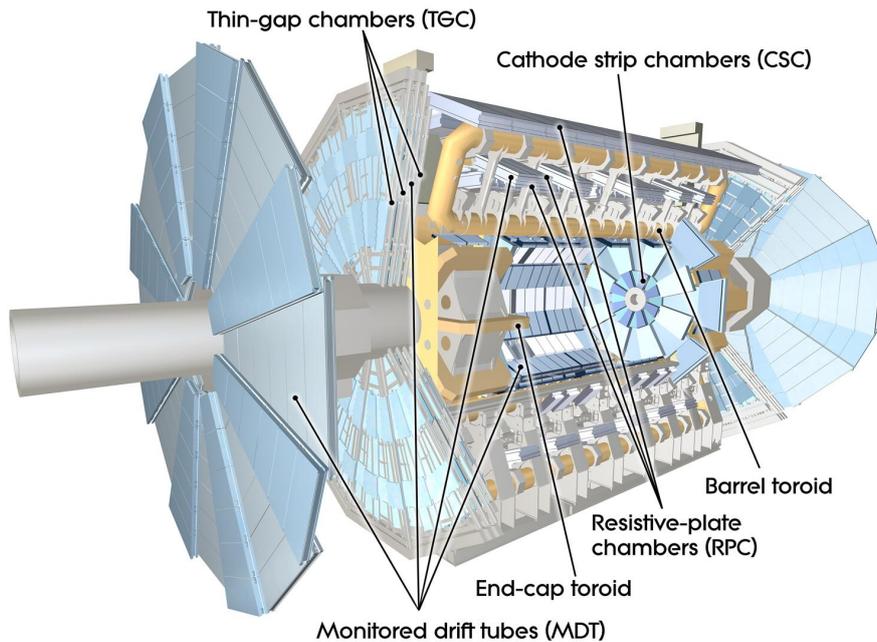


FIGURE 3.8 – The layout of the ATLAS muon spectrometer. [57].

Monitored drift tubes (MDTs) and cathode strip chambers (CSCs) are the sub-detectors capable of measuring the position and momentum of muons with high precision.

The MDTs are composed of aluminium tubes with a diameter of 30 mm and a wall thickness of 400  $\mu\text{m}$ . Each tube contains a central W-Re wire of 400  $\mu\text{m}$  diameter and filled with a mixture of gasses of Ar (93%) and CO<sub>2</sub> (7%). When a muon travels

through a tube, it ionizes the gas and the generated electrons drift to the central due to the electric field. The drift time is recorded to reconstruct the trajectory of the muon. MDTs cover the pseudorapidity range  $|\eta| < 2.7$ , with the exception of the innermost layer, which is limited to  $|\eta| < 2.0$ . The spatial resolution of MDTs is excellent, with  $35 \mu\text{m}$  per chamber.

The CSCs are positioned in the innermost layer of the muon spectrometer, spanning the pseudorapidity range  $2.0 < |\eta| < 2.7$ . CSCs are multiwire proportional chambers featuring a symmetric cell configuration. The function of CSCs is to provide high-resolution tracking in the region with high particle flux. The chambers of CSCs achieve a spatial resolution of  $40 \mu\text{m}$  in the plane parallel to the magnetic field (bending plane) and approximately 5 mm in the plane perpendicular to the magnetic field (transverse plane).

The resistive plate chambers (RPCs) and thin gap chambers (TGCs) are fast-response detectors with nanosecond level time resolution to provide fast muon triggering.

The RPCs comprise of three layers of chambers covering the region with  $|\eta| < 1.05$ . Each RPC consists of two parallel resistive plates made of bakelite and they are separated by a gas gap, which is filled with a mixture of gasses of  $\text{C}_2\text{H}_2\text{F}_4$  (94.7%), iso- $\text{C}_4\text{H}_{10}$  (0.5%) and  $\text{SF}_6$  (0.3%) [58]. The RPCs have the time resolution of 1.5 ns and a spatial resolution of 10 micrometers in both the  $z$ -axis and  $\phi$  directions.

The TGCs are positioned in the end-cap region with pseudorapidity  $1.05 < |\eta| < 2.4$ . They are similar to multiwire proportional chambers, however, the distance between cathode and anode wire is smaller than the anode wire pitch, allowing fast response. The time resolution of the TGC is 4 ns and the spatial resolution is 10 mm in both the  $z$ -axis and  $\phi$  directions.

The new small wheel (NSW) [59] is the major upgrade on the ATLAS muon spectrometer in the Phase-I upgrade and installed during the LHC shutdown in 2018 and 2019. It covers the pseudorapidity  $1.3 < |\eta| < 2.7$  as displayed in Figure 3.9. The NSWs use small-strip thin gap chambers (sTGCs) and Micromegas (micro-mesh gaseous structures) techniques, and sTGCs and Micromegas are arranged in layers to form a robust system for triggering and tracking, which is quick and precise.

### 3.2.5 Magnet system

All other ATLAS detector subsystems are immersed in the magnets field generated by the magnet system [60]. Figure 3.10 displays the schematic view of the magnet system of the ATLAS detector. As the layout illustrates, the ATLAS magnet system comprises a superconducting solenoid magnet and several toroidal magnets.

The central solenoid magnet [61], surrounding the inner detector, generating a 2 T magnetic field aligned with the beam axis at the centre of the tracking volume. This magnetic field is crucial to bend the path of the charged particles, enabling precise tracking and momentum measurement by the inner detector. Since the electromagnetic calorimeter is positioned outside the solenoid, the coil needs to be highly transparent to traversing particles. Thus the magnet incorporates technological advancements

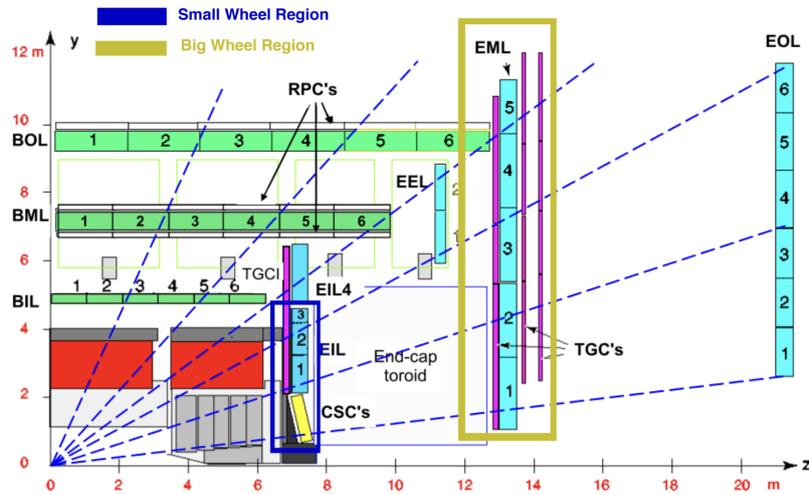


FIGURE 3.9 – The  $z - y$  view of 1/4 of the ATLAS detector. The detector region of the small wheel is also outlined [59].

from previous solenoids, including a new reinforced superconductor that ensures transparency. The solenoid has a diameter of 2.3 meters and a length of 5.3 meters.

The toroidal magnets consist of two sets of toroidal magnets : the barrel toroid and the end-cap toroid. The toroids employ eight coils to generate a magnetic field for the muon momentum measurement. The barrel toroid is 25.3 m long, which is the largest toroidal magnet ever built. It has an inner diameter of 9.4 meters and an outer diameter of 20.1 meters, providing  $\sim 0.5$  T magnetic field on superconductor. Two end-cap toroids are located at the end of the barrel toroids, with 1.65 m inner diameters, 10.6 m outer diameters and 5.0 m length. They generate  $\sim 1.0$  T magnetic fields ensuring the precise measurement of the muon produced at a small angle compared to the beam direction.

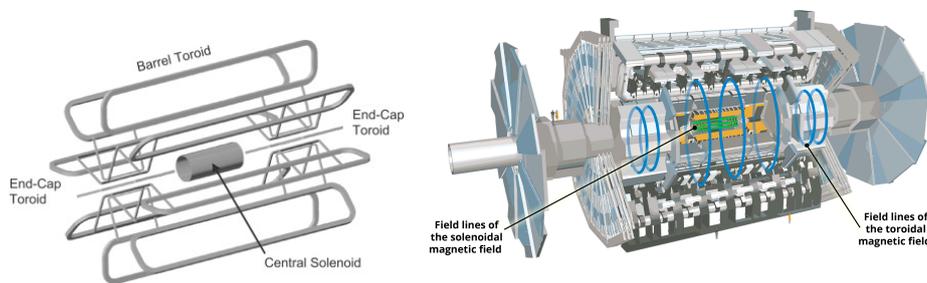


FIGURE 3.10 – The schematic of magnet system of the ATLAS detector. [61, 62]

### 3.2.6 Forward detectors

In addition, the ATLAS experiment includes several forward detectors. These detectors are specifically designed to measure particles emitted at small angles relative to the beam axis. The LUCID-2 is one of the forward detectors crucial for the measurements in this thesis.

The LUCID-2 [63] (luminosity Cerenkov integrating detector), the new ATLAS luminosity detector, comprises multiple small Cherenkov detectors to the precise measurement and continuous monitoring of the luminosity. The LUCID-2 is positioned around the beam pipe, approximately 17 m from the collision point on either side of the ATLAS detector. The detector comprises arrays of Cherenkov tubes designed to detect Cherenkov radiation, which is emitted by charged particles as they pass through the tubes. The integrated luminosity values and their associated uncertainties, utilized in the measurements presented in this thesis, are based on the primary luminosity measurements by the LUCID-2. Figure 3.11 displays the integrated luminosity based on the online luminosity estimate at Run-2 and Run-3, respectively.

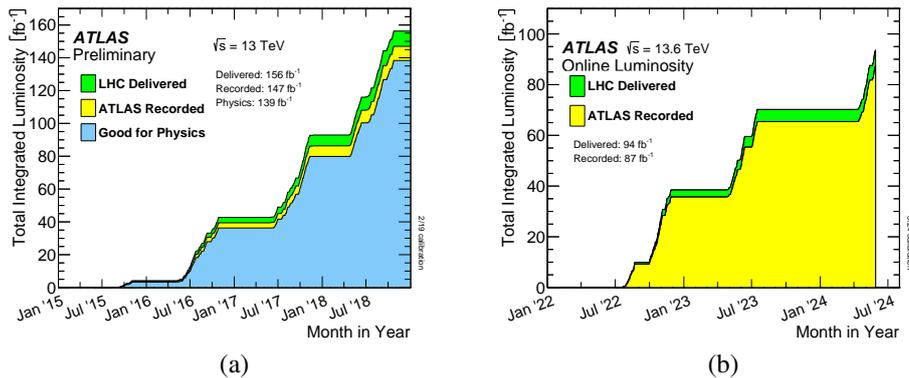


FIGURE 3.11 – Total integrated luminosity as a function of time during  $pp$  collisions (a) at  $\sqrt{s} = 13$  TeV at Run-2 (2015-2018) and (b) at  $\sqrt{s} = 13.6$  TeV at Run-3 (2022-2024). The green area shows the luminosity delivered to ATLAS and the yellow one represents the luminosity recorded by ATLAS in both of figures. In (a), the blue area represents the integrated luminosity certified as good quality data. The figures are taken from Ref. [64, 65].

### 3.2.7 The ATLAS trigger and data acquisition system

The ATLAS trigger and data acquisition (TDAQ) system [66] is designed to process and manage the vast amount of data generated by the detector. The system is capable of online identifying and recording the events with distinguishing characteristics to provide interesting physics information for the further analyses. The LHC operates at a collision rate of up to 40 million interactions per second (40 MHz collision rate), with each event generating approximately 1.5 megabytes of raw data. The

resulting data volume is around 60 million megabytes per second. The TDAQ system is required to reduce the event rate to approximately 1 kHz to meet the requirement of data storage and offline analyses. Figure 3.12 displays the ATLAS TDAQ system in Run-2, the events collected by the detector are filtered by the triggers at different levels and recorded.

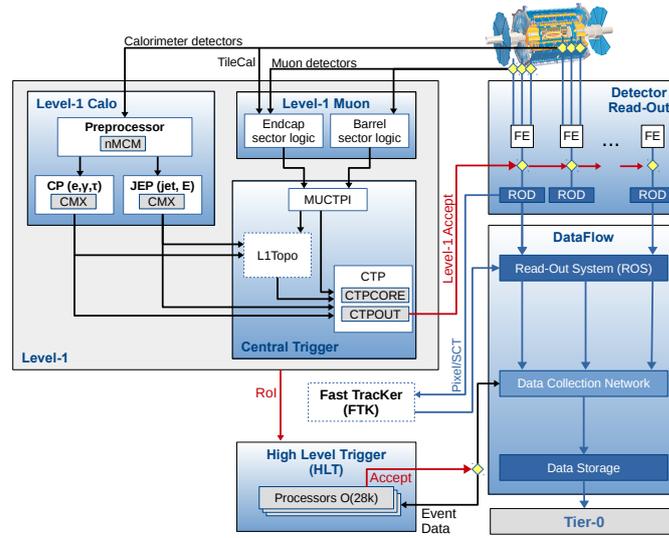


FIGURE 3.12 – The ATLAS TDAQ system in Run 2, with a focus on the components relevant to triggering, from Figure 1 in Ref. [67].

The ATLAS trigger system is composed of level-1 (L1) trigger and high-level trigger (HLT).

The L1 trigger based on the hardware can quickly decide whether to retain an event, and the decision is made within 2.5 ns. The L1 trigger relies on the information collected by several sub-detectors to make judgements. The L1 calorimeter (L1Calo) trigger makes use of the electromagnetic and hadronic calorimeters to identify the physics objects based on calorimeters like electrons, photons, jets,  $\tau$ -leptons with hadronic decay and quantify global event characteristics like missing transverse energy. The L1 muon trigger (L1Muon) depends on the information provided by the RPCs (barrel) and TGCs (end-caps) of the muon spectrometer to identify the muon originating from the interaction point. The L1 Topological (L1Topo) composed of two FPGA (field-programmable gate arrays) processor modules. It combines information from L1Calo and muon central trigger processor interface (MUCTPI) into topological variables to perform selection. The L1 central trigger processor (CTP) receives the

information from L1Calo, L1Muon and L1Topo to make the final decision. The event rate is reduced from 40 MHz to approximately 100 kHz at this level.

Then accepted events are buffered in the read-out system (ROS) and filtered by the HLT in the further step. The L1 trigger provides regional information, known as the region-of-interest (RoI), to the HLT system. This information includes the geometrical region, the type of physics objects, and the thresholds they have passed. The HLT is based on the software and employs more complex algorithm to make final decision. The HLT can reduce the event rate from 100 kHz to around 1 kHz for offline analyses.

The events accepted by the HLT are transferred to the data storage for the offline analyses.



## 4 - Object reconstruction and identification

The information collected by the ATLAS detector is reconstructed and identified as different physics objects based on their characteristics. This process is the initial step in the physics analysis. This chapter discusses the reconstruction and identification of the physics object in the ATLAS experiment.

### 4.1 Track and vertex

Track, the trajectory of a charged particle, plays a crucial role in event reconstruction, as it is a fundamental ingredient utilized by other physics object reconstruction. Tracks can be parameterized using several parameters ( $d_0, z_0, \phi, \theta, q/p$ ) relative to the default reference, an axis aligned with the global  $z$  direction, centered on the collision point. Here,  $d_0$  represents the transverse impact parameter, which is the distance of the closest approach of the particle's trajectory to the collision point in the transverse plane;  $z_0$  is the longitudinal impact parameter, indicating the displacement of the particle's trajectory along the beam axis at the point of closest approach;  $\phi$  and  $\theta$  represent the azimuthal and polar angles, respectively;  $q/p$  is the ratio between the particle electric charge and momentum.

The track reconstruction in the inner detector follows the algorithm introduced in Refs. [68, 69]. The approach consists of several steps [70]: the initial step is clusterization. Clusters measured from detectors are assembled and *space points* are reconstructed based on the position recorded by the inner detector. Track seeds are then defined from the combinations of three space points and selected according to its quality, with criteria depending on the sub-detector. The Kalman-filter [71] is employed to generate track candidates from selected track seeds. The ambiguity of track candidates is estimated in the ambiguity-solver [70], which assigns the track score reflecting the track quality and neural network to determine if a cluster should be split among track candidates. Finally, the candidates are evaluated using a global fit, taking all relevant information into account.

Primary vertex (PV) is defined as the location of the hard-scattering interaction. Vertex reconstruction based on track objects involves both vertex finding and vertex fitting. The vertex reconstruction employs the ATLAS iterative vertex finder (IVF) strategy [72]: determine vertex seed, the most likely position of a new primary vertex, using track information. After finding a seed, nearby tracks are associated with it. A vertex fit is performed to evaluate the vertex candidate, then process the remaining tracks in the further iterations. A new strategy of vertex reconstruction is applied for Run-3 [73] to improve the efficiency and pile-up robustness. The improvement is achieved by using the adaptive multi-vertex finder and fitter (AMVF) [73] with a better seed-finding algorithm and fitting strategy. In practice, the vertex with the

largest sum of the transverse momentum of the associated tracks,  $\sum_{\text{track} \in \text{vertex}} p_{\text{T,track}}^2$ , is chosen as the primary vertex.

## 4.2 Electron and photon

### 4.2.1 Electron and photon reconstruction

Both electron and photon reconstructions start from the construction of clusters, which are combinations of adjacent cells depositing energy in the calorimeter.

An electron is defined as an object constructed from a cluster along with an associated track. Photon reconstruction comprises two cases : converted photon is defined with a cluster and a associated conversion vertex, while unconverted photon is a cluster with no tracks or vertices associated to it. Figure 4.1 displays the trajectory of an electron from the collision point traveling across the ATLAS detector.

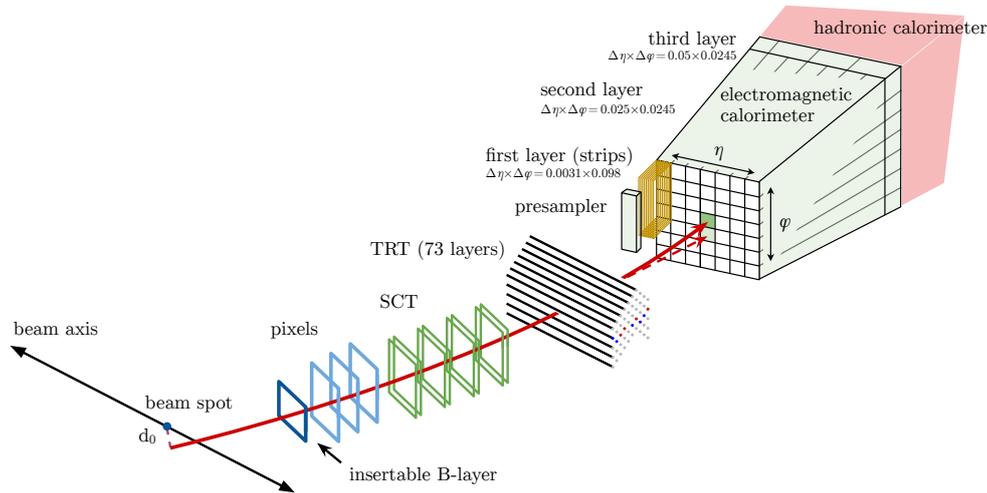


FIGURE 4.1 – The trajectory of an electron from the interaction point as it travels through the detector [74].

The electron and photon reconstruction [75, 76] employs the so-called superclusters [75] technique, which makes use of variable-sized clusters for reconstruction.

The beginning step of the reconstruction is to build the variable-sized topo-cluster [77], which is constructed with clusters that have topological connections. Several criteria are applied to select qualified clusters for further reconstruction. The electromagnetic (EM) energy of the topo-cluster, which represents energy deposited in the electromagnetic calorimeter, must exceed 400 MeV. Additionally, topo-clusters are required to have more than 50 percent of their energy deposited in the electromagnetic calorimeter. Clusters meeting these criteria are referred to as EM topo-clusters.

Tracks in the ID is corrected to take bremsstrahlung energy losses into account. The corrected tracks are paired with fixed-sized clusters in the EM calorimeter and

then refitted with a Gaussian sum filter (GSF) [78] to enhance the precision, which are so-called GSF tracks. Tracks are also used to construct the conversion vertices for photon reconstruction [75].

Electron and photon candidates are then reconstructed separately. The EM topo-clusters with matched tracks serve as electron candidates. The primary track is selected among the matched GSF tracks based on the ranking given by the matching algorithm [74]. For converted photons, at least one conversion vertex matched with the cluster is required. Unconverted photon is based the EM topo-clusters without matched tracks or conversion vertices.

These candidates serve as the input to the superclustering algorithm, the clusters are combined to related split due to bremsstrahlung and then refined through adjustments and re-calibrations to form offline electron and photon objects for further physics analyses. Figure 4.2 shows the correction procedure.

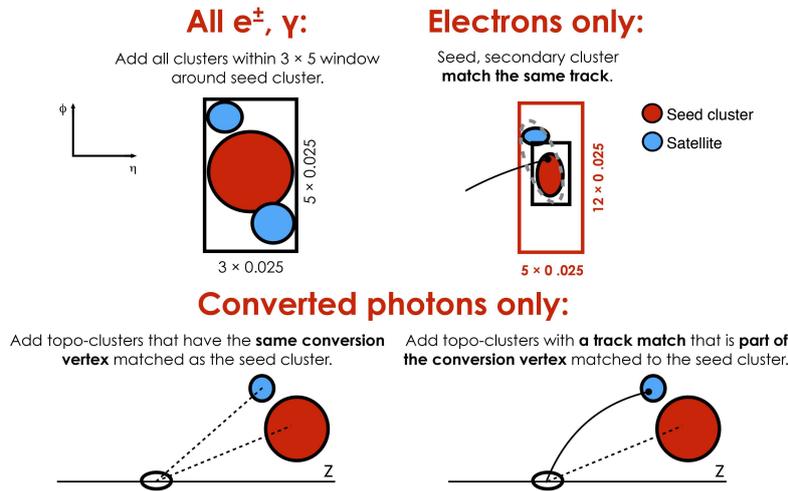


FIGURE 4.2 – Diagram illustrating the superclustering algorithm for electrons and photons, with seed clusters shown in red and satellite clusters in blue. Taken from Figure 3 in Ref. [75].

#### 4.2.2 Electron and photon energy calibration

The energy calibration ensures the energy measurements from the EM calorimeter are reliable and accurate. The calibration strategy is summarized in Figure 4.3.

The energy calculated from clusters is refined using a simulation-based decision tree regression algorithm. Subsequently, the energy from each layer is corrected independently. Corrections derived from simulations are then applied to both data and simulated events. Final calibration tuning relies on the  $Z \rightarrow ee$  process, with adjustments ensuring the  $Z$  resonance peak aligns between real data and simulation. These corrections are applied to both electrons and photons.  $J/\psi \rightarrow ee$  events are used to validate the calibration in the low- $E_T$  case.

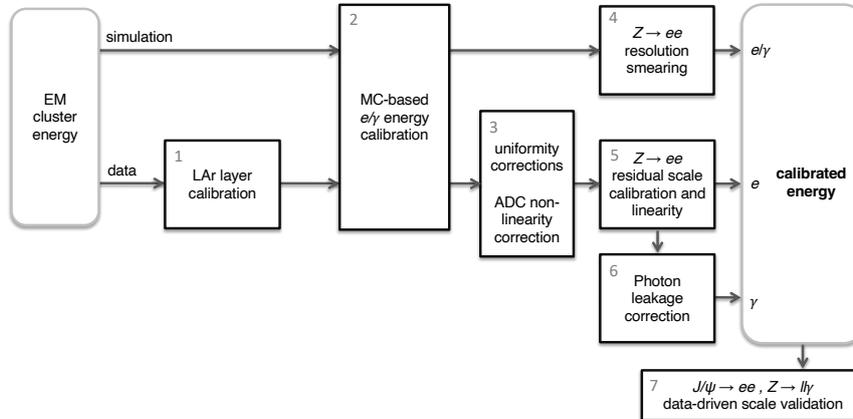


FIGURE 4.3 – The energy calibration strategy for electrons and photons in the ATLAS experiment. Taken from Figure 3 in Ref. [79].

### 4.2.3 Electron and photon identification

Electrons and photons are crucial experimental signatures for many physics processes. To ensure clean final states, it is essential to distinguish prompt electrons and photons – those originating from the primary vertex – from background particles. The discriminant observables are related to track characteristics, shower behavior in calorimeter layers, and track-cluster alignment. Selection criteria for identifying electrons and photons are applied based on  $|\eta|$  and transverse energy ( $E_T$ ).

The likelihood classifier is used for electron identification in ATLAS [75]. The observables characterizing the electron object are used as inputs for the likelihood classifier. Based on identification efficiency, three working points (WP) – Tight, Medium, and Loose – are defined, where the term “WP” refers to a specific set of criteria or thresholds chosen to select particles. Figure 4.4 displays the electron identification efficiency of various WPs. To eliminate the discrepancies between data and simulated samples in the identification efficiencies, corrections on the efficiencies are measured as a function of electron  $E_T$  and  $\eta$ . It is the so-called identification “scale factors” in ATLAS, which are defined as the ratio of efficiency between data and simulated events under the same identification criteria. The identification is available for electrons within the pseudorapidity range  $|\eta| < 2.47$ .

The photon identification depends on the shower shape in the calorimeters. Similar to the electron identification, three levels (Loose, Medium and Tight) of identification criteria are defined for photons within the pseudorapidity range  $|\eta| < 2.37$ .

### 4.2.4 Electron isolation

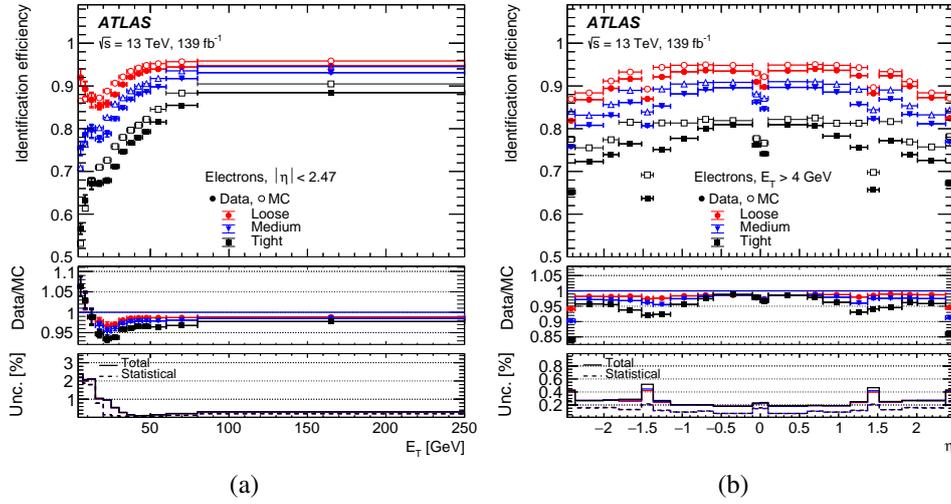


FIGURE 4.4 – Electron identification efficiency as functions of electron (a)  $E_T$  and (b)  $\eta$  from the full Run 2 measurement [79].

Isolation is a measure of the amount of particle activity or energy deposition around the object within a given radius. For electrons, the isolation is a crucial quantity to separate prompt electrons from backgrounds like jet or non-prompt electrons.

Two major isolation variables are used to quantify the isolation in ATLAS : calorimeter isolation,  $E_T^{\text{cone } X}$  and track isolation  $p_T^{\text{cone } X}$ , measuring the energy deposition in clusters and energy of tracks within a radius of  $X$ , respectively, where  $X$  denotes the cone size. In practice, the value 20 and 30 are often used to define the electron isolation criteria, corresponding to a cone radius 0.2 and 0.3. Several WPs are defined based on these variables.

## 4.3 Muon

### 4.3.1 Muon reconstruction and calibration

As muons ionize the matter minimally, they pass through a significant amount of material with minimal energy loss. Therefore, the muon reconstruction mostly relies on the tracking detectors, namely the ID and MS.

Muons could be reconstructed using information from either the ID or the MS. For the ID case, the charged particle reconstruction method described in Section 4.1 also applies to muons. In the MS, the muon is reconstructed in the following steps [80].

Initially, Hough transform [81] is utilized to select short charged track segments within detector regions in MDT chambers. A combination is then performed on these track segments to form a track candidate from the collision point. Track candidates are built with precision measurements in the bending plane and second-coordinate data from the trigger detectors. To evaluate the quality of a track, a global  $\chi^2$  fit

is performed considering the impact of detector material and misalignments. Tracks sharing too many hits with good tracks are eliminated. The selected track candidates are then evaluated by another fit, accounting for energy loss in calorimeters and the capability to be extrapolated back to the beam line.

The global muon reconstruction combines the information from tracking detectors and calorimeters. Five types of muon are defined according to the reconstruction method :

- Combined (CB) muons : detected by aligning MS tracks with those from the ID and executing a combining tracking fit incorporating both ID and MS hits, which accounts for energy loss in the calorimeters.
- Muon-spectrometer extrapolated (MS) muons : MS tracks with no matched ID tracks are extrapolated to the beam line and are defined as MS muons. It recovers the reconstruction efficiency in regions with limited ID coverage, maximizing the MS coverage up to  $|\eta| < 2.7$ .
- Inside-out combined (IO) : IO muons are reconstructed by extending ID tracks to the MS and finding at least three aligned MS hits. A combined track fit is then performed using the ID track, energy loss in the calorimeters, and the MS hits. As the MS tracks are not utilized in the reconstruction, it improves the reconstruction efficiency in areas with limited MS coverage and very low  $p_T$  muons that could not reach MS.
- Segment-tagged (ST) muons : identified by ensuring an ID track, projected to the MS, meets strict angular matching criteria with at least one reconstructed MS segment. Muon parameters are derived from ID tracks.
- Calorimeter-tagged (CT) muons : reconstructed by projecting ID tracks through the calorimeters to detect energy deposits typical of a minimum-ionizing particle. Muon parameters are obtained from ID tracks.

Muon calibrations [82] depends on the physics processes  $Z \rightarrow \mu\mu$  and  $J/\psi \rightarrow \mu\mu$ . Corrections are made to adjust the momentum scale and improve resolution to improve the agreement between actual data and simulated events when measuring the same observables. Figure 4.5 shows the comparison between data and calibrated simulated events in di-muon invariant mass distributions.

### 4.3.2 Muon identification and isolation

Muon identification is utilized to distinguish the prompt muons and the backgrounds, which mainly originate from light hadrons. The ATLAS muon identification is based on hit counts in ID and MS sub-detectors, track fit characteristics and consistency between the ID and the MS. For the CB muons, the compatibility between the ID and the MS can be quantified as a  $q/p$  compatibility [80] :

$$q/p \text{ compatibility} = \frac{|q/p_{\text{ID}} - q/p_{\text{MS}}|}{\sqrt{\sigma^2(q/p_{\text{ID}}) + \sigma^2(q/p_{\text{MS}})}} \quad (4.1)$$

where  $q$  denotes muon's electric charge,  $p$  is muon's momentum, ID and MS represent tracks reconstructed in corresponding detectors,  $\sigma$  in the denominator denotes the

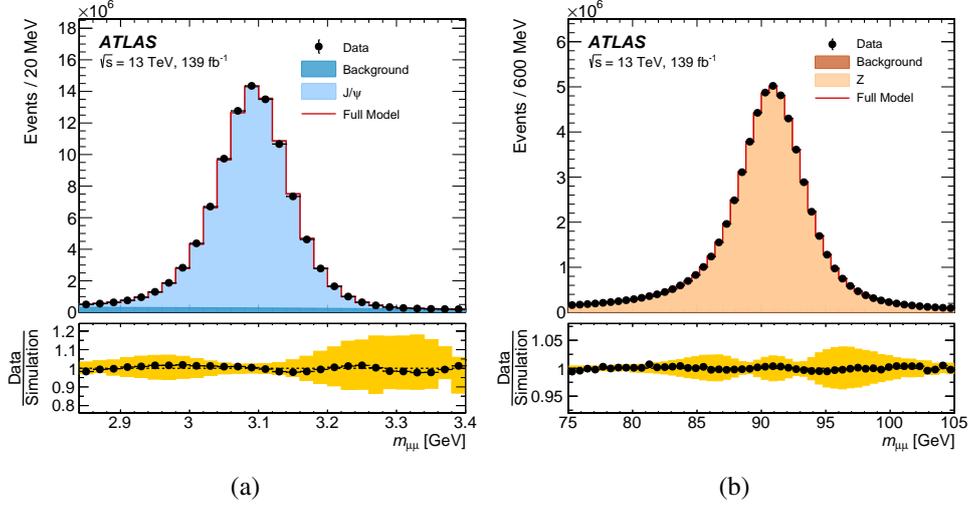


FIGURE 4.5 – Comparison between data and simulated events in the di-muon invariant mass distributions,  $m_{\mu\mu}$ , of (a)  $J\psi \rightarrow \mu\mu$  and (b)  $Z \rightarrow \mu\mu$  events [82].

corresponding uncertainty, and  $\rho'$  is calculated as the absolute deviation between transverse momentum  $p_T$  from ID and MS, normalized by the  $p_T$  of the combined (CB) track.

$$\rho' = \frac{|p_{T,ID} - p_{T,MS}|}{p_{T,CB}}. \quad (4.2)$$

Subsequently, several working points (WPs) are defined to accommodate various use cases. Like electrons, three standard WPs (Loose, Medium, Tight) are defined, each tailored to different levels of identification efficiency and background rejection strength. Muon efficiencies are often measured as functions of muon transverse momentum  $p_T$  and pseudorapidity  $\eta$ , as illustrated in Figure 4.6. The Medium WP is often chosen as the baseline in ATLAS physics analyses. Additionally, high- $p_T$  WPs are specifically adapted for muons with large momentum. The identification is corrected with scaled factors, the ratio of efficiencies between data and simulated events, which is obtained accordingly.

Similar to electrons, the muon isolation also makes use of the calorimeter isolation and track isolation variables described in Section 4.2.4.

## 4.4 Jet

### 4.4.1 Jet reconstruction and calibration

Jet is a collimated spray of particles resulting from the hadronization of quarks and gluons. It plays a crucial role in studying a wide range of physics processes in proton–proton collisions at the LHC.

Jets are reconstructed with the anti- $k_t$  algorithm [83, 84], using a typical radius

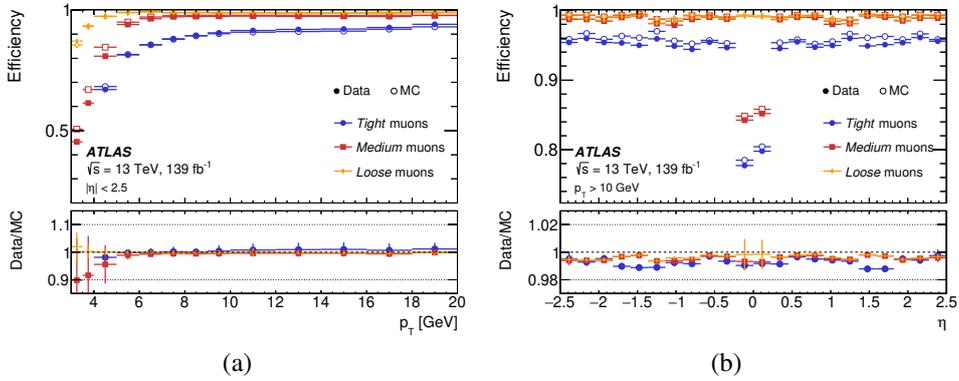


FIGURE 4.6 – Muon reconstruction and identification efficiencies measured as functions of muon transverse momentum (a)  $p_T$  and pseudorapidity (b)  $\eta$  for various ID WPs [80].

parameter of  $R = 0.4$ . The ATLAS experiment employs a jet reconstruction technique called the particle flow algorithm [85], which makes use in an optimal way of energy deposits in the calorimeters and track momenta from the inner tracking detectors. The particle flow algorithm enhances thus jet energy resolution and mitigates pile-up effects.

Similar to electron reconstructions, the topo-clusters are reconstructed as the initial step of the jet reconstruction, which are clusters of topologically connected calorimeter cells with significant energy deposits [77]. To suppress the impact of noise, a threshold is applied to the ratio between energy deposits in the cell and expected noise. This typically requires the measured energy to be at least four times greater than the noise level. Subsequently, the particle flow algorithm is utilized to match tracks with a topo-cluster and apply energy correction to take the difference response of calorimeters and the inner detector. Figure 4.7 illustrates the workflow of the particle flow algorithm.

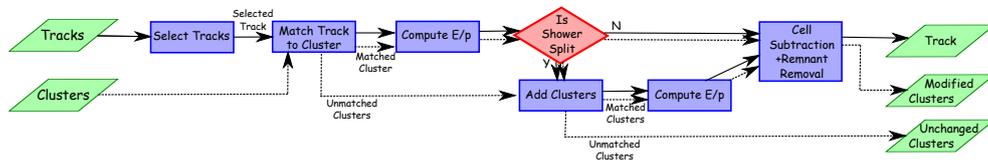


FIGURE 4.7 – The workflow of the particle flow technique [77].

The corrected topo-clusters serve as the input of the anti- $k_t$  algorithm, which reconstructs jets with well-defined shapes and boundaries. The anti- $k_t$  algorithm merges particles surrounding the highest energy constituent identified within the jet

iteratively. First, it defines the distance between particle  $i$  and  $j$  by :

$$d_{ij} = \min \left( \frac{1}{p_{T,i}}, \frac{1}{p_{T,j}} \right) \frac{\Delta R_{ij}^2}{R^2} \quad (4.3)$$

where  $R$  is the distance in the  $y - \phi$  plane. Similarly, the distance between particle  $i$  and the beam  $d_{iB}$  is defined as  $d_{iB} = \frac{1}{p_{T,i}^2}$ . If the distance  $d_{ij}$  satisfies  $d_{ij} < d_{iB}$ , particles  $i$  and  $j$  are merged into a single one, else particle  $i$  is regarded as a jet and removed from the candidates. This iteration continues until all particles are removed, after which jets are reconstructed.

Jet calibration is applied to correct the measured jet energy to accurately reflect the true energy of the original parton that initiated the jet. The accuracy of jet calibration is essential for precise measurement and reducing the related systematic uncertainties in physics analyses.

Figure 4.8 displays the workflow of jet calibration applied to the Run 2 dataset in the ATLAS experiment. Initially, pile-up corrections are applied to eliminate additional energy from extra  $pp$  interactions happening in the same or adjacent bunch crossings, which involve jet area and transverse momentum density adjustments, as well as residual corrections obtained from simulations based the average number of interaction per bunch crossing ( $\langle \mu \rangle$ ) and the number of primary vertices. The absolute jet energy scale (JES) calibration adjusts the four-momentum of jets to ensure that their energy and direction correspond to the truth jets taken from di-jet simulated events. Subsequently, the global sequential calibration is employed to enhance the energy resolution and uncertainties utilizing the track, calorimeter and MS observables. After applying the corrections on both data and simulated events, a data-only residual *in-situ* calibration is employed to eliminate discrepancies between data and simulated jet objects, using the physics processes where the reference objects could be precisely measured, such as  $Z$ +jet and  $\gamma$ +jet.

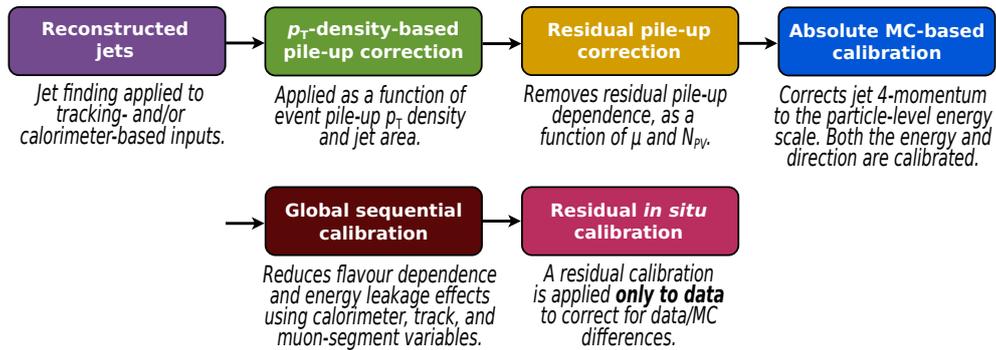


FIGURE 4.8 – The workflow of the jet calibration procedure for ATLAS Run 2 dataset [86].

The jet energy resolution (JER) is then determined, which is expressed as functions of the jet transverse momentum ( $p_T$ ) and pseudorapidity ( $\eta$ ). The resolution can be represented in the form :

$$\frac{\sigma(p_T)}{p_T} = \frac{S}{\sqrt{p_T}} \oplus \frac{N}{p_T} \oplus C \quad (4.4)$$

where  $\frac{S}{\sqrt{p_T}}$  is a stochastic term accounting for statistical fluctuations,  $\frac{N}{p_T}$  is the noise term for electronic or background noise,  $C$  is the constant term for other effects which are proportional to the jet  $p_T$ .

#### 4.4.2 Flavor tagging

Jet flavor tagging is a process to identify the flavor of the quark that initiated a jet. The  $b$ -tagging procedure is applied to jets with  $p_T > 20$  GeV and  $|\eta| < 2.5$  to identify those originating from  $b$ -quarks.

The  $b$ -tagging algorithms used for the Run 2 in the ATLAS experiment are discussed in Refs. [87, 88]. They are classified into two categorizations : low-level and high-level  $b$ -taggers. The low-level  $b$ -taggers, based on the characteristic features used to identify  $b$ -jets, are categorized into two methods. One method utilizes the impact parameters of tracks as inputs for the algorithm to differentiate the vertex of the  $b$ -quark decay from the primary vertex. The other method explicitly reconstructs displaced vertices using a vertex-finding algorithm.

The high-level  $b$ -taggers, DL1, use the output of the low-level  $b$ -taggers as the input to train a deep-learning neural network (DNN) as the classifier. In general, features such as impact parameter data, secondary vertex data, track information, and jet kinematics are utilized to identify  $b$ -jets. The training dataset is obtained with Monte Carlo simulations of the top-antitop pair ( $t\bar{t}$ ) production and high mass  $Z' \rightarrow q\bar{q}$  events. The DL1r neural network produces a set of probabilities for a jet being a  $b$ -jet,  $c$ -jet, or light-flavour jet. The probabilities used to define the  $b$ -tagging discriminant :

$$D_{DL1} = \ln \left( \frac{p_b}{f_c p_c + (1 - f_c) p_{\text{light}}} \right) \quad (4.5)$$

where  $p_c$  and  $p_{\text{light}}$  denote the probabilities of  $c$ -jet and light-jet, respectively, as determined by the DNN, and  $f_c$  represents the proportion of  $c$ -jets in the assumed background model. This method allows the proportion of  $c$ -jets in the background to be determined retrospectively to improve the algorithm's performance at the analysis stage. Figure 4.9 shows the distribution of the  $b$ -jet discriminant, which effectively distinguishes  $b$ -jets from backgrounds.

### 4.5 Missing transverse energy

Not all particles can be measured directly with the ATLAS detector. The most common "invisible" particle is the neutrino, which only has weak interaction with

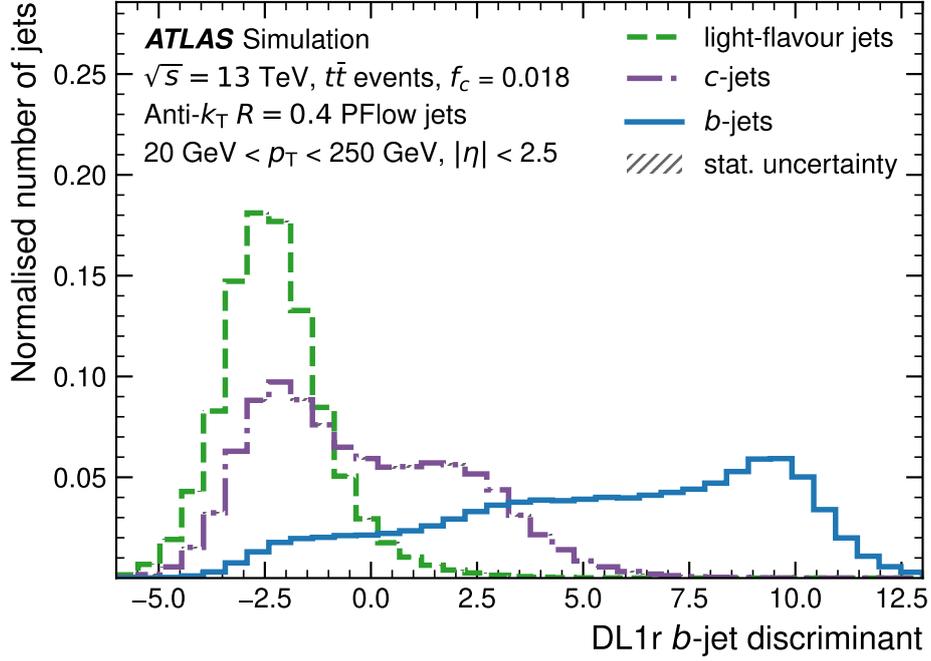


FIGURE 4.9 – The distribution of the  $b$ -tagging discriminator for light-jet,  $c$ -jet and  $b$ -jet in  $t\bar{t}$  simulated events [87].

the matter and does not leave any signals which can be captured by detectors. Therefore, an indirect method is used to measure these “invisible” particles based on the conservation of momenta.

For  $pp$  collisions at the LHC, the transverse momentum (the momentum perpendicular to the beam direction) should ideally equal zero, provided all produced particles are taken into account. Therefore, the transverse momentum of invisible particles can be studied in the transverse plane (perpendicular to the beam line). Momenta of all particles measured by the detector are projected onto this plane and summed vectorially. The deviation from zero indicates the presence of invisible particles. The deviation is so-called  $E_T^{\text{miss}}$ , reflecting a global characteristic of an event.  $E_T^{\text{miss}}$  can be written as a formula :

$$0 = \vec{E}_T^{\text{miss}} + \sum_i \vec{E}_T^i \quad (4.6)$$

where  $\sum_i \vec{E}_T^i$  expresses the vectorial sum of transverse energies of all measured particles. Figure 4.10 illustrates an example of the  $E_T^{\text{miss}}$  reconstruction, a  $W$ -boson decays into a muon and a neutrino in the transverse plane.  $E_T^{\text{miss}}$  is utilized in this event reconstruction to represent the property of the neutrino.

The standard approach to reconstruct the  $E_T^{\text{miss}}$  involves several steps [90]. The measured objects in Eq. (4.6) are classified into two distinct categories. One category

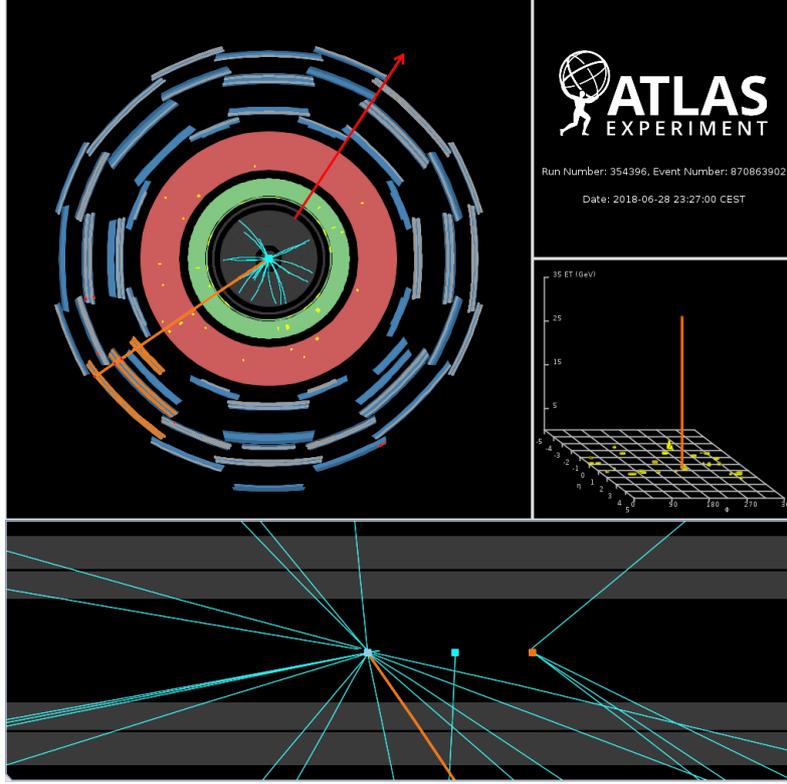


FIGURE 4.10 – The ATLAS event display showing the  $W^- \rightarrow \mu\nu$  process, the red line indicates the direction of the missing transverse momentum [89].

is the ‘hard object’, which includes electrons, muons,  $\tau$ -leptons, photons and jets. The other category, ‘soft tracks’, comprises tracks that align with the hard-scattering vertex but these tracks are not associated with any ‘hard objects’. Therefore, Eq. (4.6) can be re-written in a form that includes the ‘hard term’ and the ‘soft term,’ corresponding to the hard and soft objects, respectively :

$$\vec{E}_T^{\text{miss}} = - \left( \sum_{\text{electrons}} \vec{E}_T^{\text{electron}} + \sum_{\text{muons}} \vec{E}_T^{\text{muon}} + \sum_{\tau} \vec{E}_T^{\tau} + \sum_{\text{photons}} \vec{E}_T^{\text{photon}} + \sum_{\text{jets}} \vec{E}_T^{\text{jet}} \right) - \sum_{\text{soft tracks}} \vec{E}_T^{\text{soft tracks}}. \quad (4.7)$$

The terms in the first parentheses represent the hard term, while the soft term is expressed as the vector sum of the transverse energy of the soft tracks. The magnitude and direction in the transverse plane, denoted as  $\phi_{\text{miss}}$ , are thus obtained :

$$E_T^{\text{miss}} = \left| \vec{E}_T^{\text{miss}} \right| = \sqrt{\left( E_{T,x}^{\text{miss}} \right)^2 + \left( E_{T,y}^{\text{miss}} \right)^2}$$

$$\phi_{\text{miss}} = \arctan \left( E_{T,y}^{\text{miss}} / E_{T,x}^{\text{miss}} \right)$$

where the subscripts  $x$  and  $y$  denote the vector components along the  $x$ -axis and  $y$ -axis, respectively.

The reconstruction of  $E_T^{\text{miss}}$  is significantly influenced by the selection criteria for jets and the pile-up conditions. Therefore, the quality of  $E_T^{\text{miss}}$  is evaluated using

criteria based on jet vertex tagger (JVT) [91] and forward jet vertex tagger (fJVT) [92] discriminants, which are used to distinguish between jets originating from the primary hard-scattering vertex and those arising from pile-up interactions.



## 5 - Measurement of $W^\pm$ and $Z$ production cross sections and their ratios at $\sqrt{s} = 13.6$ TeV

This chapter details the first measurement of inclusive  $W^\pm$  and  $Z$  production cross sections and their ratios at an unprecedented center-of-mass energy at  $\sqrt{s} = 13.6$  TeV. The measurement covers the fiducial and total  $W^\pm$  and  $Z$  production cross sections, as well as the ratios of  $W^+$  to  $W^-$ ,  $W^\pm$  to  $Z$ , and top-antitop-quark pair productions to  $W^\pm$  fiducial production cross sections.

### 5.1 Introduction

Since the discovery of the  $W^\pm$  and  $Z$  bosons at the CERN SPS collider [16-19], the properties of these vector bosons have been actively studied in the past decades. The measurement of  $W$  and  $Z$  boson's properties remains an important topic of research at the LHC. As mediators of the weak force, they offer essential insights to the SM and aid in exploring potential new physics beyond the SM.

Precision measurements of the  $W^\pm$  and  $Z$  production cross sections also provide crucial benchmark tools to probe QCD and reveal the structure of hadrons, in particular, parton distribution functions (PDFs). As illustrated in Figure 2.5, the LHC exhibits large  $W^\pm$  and  $Z$  production rates. Various related measurements of  $W^\pm$  and  $Z$  cross sections are continuously performed at the LHC at different centre-of-mass energies, for instance, the ATLAS experiment [93] measured the cross sections at  $\sqrt{s} = 2.76$  TeV [94], 5 TeV [95], 7 TeV [96], 8 TeV [97, 98] and 13 TeV [42, 99], and the CMS experiment [46] performed the measurements at centre-of-mass energies of 7 TeV [100, 101], 8 TeV [102, 103] and 13 TeV [104]. The LHC achieves a record-breaking center-of-mass energy of 13.6 TeV in  $pp$  collisions in Run 3, which provides an opportunity to carry out such measurement at unprecedented energy.

The leptonic decay provides clean experimental signatures, simplifying the reconstruction of events and the subsequent measurement of boson properties. Therefore, this study makes use of leptonic final states, measuring the inclusive fiducial and total cross sections of  $W^\pm \rightarrow \ell^\pm \nu_\ell (\bar{\nu}_\ell)$  and  $Z \rightarrow \ell^+ \ell^-$ , as well as the ratios of their fiducial cross sections,  $W^+/W^-$  and  $W^\pm/Z$ . In addition to  $W^\pm$  and  $Z$  cross-section ratios, the ratios  $R_{t\bar{t}/W^\pm}$ ,  $R_{t\bar{t}/W^-}$  and  $R_{t\bar{t}/W^+}$  are also measured in this analysis, given its sensitivity to various parton densities. This analysis uses the data collected by the ATLAS detector in  $pp$  collisions at  $\sqrt{s} = 13.6$  TeV, corresponding to an integrated luminosity of  $29 \text{ fb}^{-1}$ . This dataset is collected during 2022, which is the first year of Run-3 data taking. The analysis of the well-known  $W^\pm$  and  $Z$  processes helps in validating the detector performance and reconstruction framework at the early Run 3 periods.

The analysis strategy is as follows : events are initially selected and categorized

into several regions based on the final states :  $e^+e^-$ ,  $\mu^+\mu^-$ ,  $e^-\bar{\nu}$ ,  $e^+\nu$ ,  $\mu^-\bar{\nu}$  and  $\mu^+\nu$ . In addition to these channels, two  $e\mu$  channels from Ref. [105] are involved when extracting the  $t\bar{t}/W^\pm$  ratio. Cross sections and their ratios are subsequently extracted by performing combined profile-likelihood fits in these channels.

The measurements are evaluated against theoretical predictions calculated at NNLO. These predictions incorporate the resummation of logarithmically enhanced contributions in the low transverse-momentum region of the lepton pairs, reaching next-to-next-to-leading-logarithmic (NNLL) accuracy in QCD and NLO precision in EW calculations. Various advanced PDFs are employed in these predictions. Furthermore, the relationship between the cross sections and the center-of-mass energy is examined.

## 5.2 Data and MC samples

### 5.2.1 Data samples

This analysis uses the ATLAS Run-3  $pp$  collisions data sample in 2022 at  $\sqrt{s} = 13.6$  TeV. The dataset has an integrated luminosity of  $29.0 \pm 0.6 \text{ fb}^{-1}$ . Only events that meet data quality standards are selected, verified by their inclusion in the good run list (GoodRunList), which tracks luminosity blocks of data collected under various conditions.

### 5.2.2 MC samples

The production of  $Z$  and  $W$  bosons was simulated using the SHERPA,2.2.12 generator [106]. This simulation employed NLO matrix elements (ME) for configurations involving up to two partons, and leading order (LO) MEs for configurations involving up to five partons, which were calculated with the COMIX [106] and OPENLOOPS [107-109] libraries. The results were matched to the SHERPA parton shower [110] using the MEPS@NLO prescription [111] and a set of tuned parameters developed by the SHERPA authors. The NNPDF3.0NNLO PDF set [112] was utilized, and the samples were normalized to an NNLO prediction [113].

An alternative sample of  $Z$  bosons was simulated using the POWHEG Box v2 generator [114-116] at NLO accuracy for the hard-scattering processes of boson production and decay into the electron, muon, and  $\tau$ -lepton channels. The ME simulation was interfaced with PYTHIA,8.307 [117] to model the parton shower, hadronization, and underlying event, with parameters set according to the AZNLO tune [118]. The CT10NLO PDF set [119] was employed for the hard-scattering processes, while the CTEQ6L1 PDF set [120] was used for the parton shower. The effect of QED final-state radiation was simulated using PHOTOS++,3.64 [121, 122]. Similarly, an alternative sample of  $W$  bosons was simulated using the same POWHEG Box v2 configuration described above for the  $Z$  boson sample. Both the nominal and alternative samples have been used for systematic studies.

Samples of diboson final states were simulated using SHERPA,2.2.12, incorporating

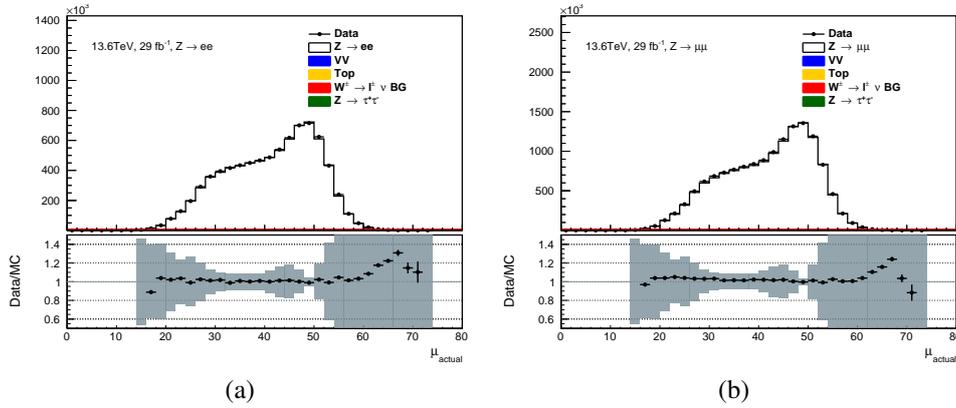


FIGURE 5.1 – Comparison of the actual pile-up distributions  $\mu$  between data and simulations after pile-up reweighting in the (a)  $Z \rightarrow ee$  and (b)  $Z \rightarrow \mu\mu$  events. The lower panel displays the ratio of data to the prediction. The shaded band represents the total systematic uncertainty

off-shell effects and Higgs boson contributions where relevant. Both fully leptonic final states and semileptonic final states, where one boson decays leptonically and the other hadronically, were generated. The matrix elements (MEs) were calculated at next-to-leading order (NLO) accuracy in QCD for up to one additional parton and at leading order (LO) accuracy for up to three additional parton emissions. As there is no dedicated calculation for the cross section currently available, the normalization for 13.6 TeV uses the same  $k$ -factors as in the Run 2 analyses, which are evaluated to be 0.91.

The production of a top-quark pair  $t\bar{t}$  and the associated production of a single top-quark and a  $W$  boson ( $Wt$ ) were modeled using the POWHEG BOX v2 generator, interfaced with PYTHIA 8.307 for parton showering, employing the tune A14 [123]. The MEs were calculated at NLO precision in QCD using the NNPDF3.0NLO PDF set. The  $t\bar{t}$  sample was normalized to the cross-section prediction at NNLO in perturbative QCD, including the resummation of NNLL soft-gluon terms computed using TOP++[2.0] [124-130].

The MC samples are produced before data collection, based on an estimated guess of the data pile-up conditions. To enhance the agreement between data and MC regarding the actual pile-up distribution  $\mu$ , which is the number of interactions per bunch crossing, the pile-up conditions of the MC samples need to be reweighted. The pile-up weight for an MC event is determined using the default MC pile-up profile and the data  $\mu$  histogram corresponding to a specific GoodRunList, which records the qualified events for physics analysis. Figure 5.1 shows the comparison between data and MC after the pile-up weight is applied. Good agreement is observed with  $Z \rightarrow \ell^+\ell^-$  events.

### 5.3 Object selection

The object reconstruction was described in Chapter 4. The relevant object selection for this analysis is discussed here.

The electron candidates should satisfy the Tight identification criteria [74, 75], with a transverse momentum  $p_T$  exceeding 27 GeV and an absolute pseudorapidity  $|\eta|$  within the range of 2.47, excluding the transition region between the barrel and end-cap calorimeters ( $1.37 < |\eta| < 1.52$ ).

The muon candidates should meet the Medium identification criteria [131], with  $p_T$  greater than 27 GeV and  $|\eta|$  less than 2.5.

Leptons, which refers to electron or muon in this analysis, should originate from the primary vertex. Therefore, the selections on the impact parameters are applied. The longitudinal impact parameter of lepton charged track, which is the distance between the track and the primary vertex along the beam line multiplied by the sine of the track's  $\theta$  angle, must be less than 0.5 mm, i.e.,  $\Delta z_0 \times \sin \theta < 0.5$  mm. Additionally, the transverse impact parameter significance, defined as the lepton track's transverse impact parameter  $d_0$  relative to the beam line divided by its uncertainty  $\sigma(d_0)$ , must be less than 5 for electrons ( $|d_0|/\sigma(d_0) < 5$ ) and less than 3 for muons ( $|d_0|/\sigma(d_0) < 3$ ).

To reduce the contamination of non-prompt leptons, leptons are required to pass specific isolation criteria. For both leptons, the TIGHT\_VARRAD isolation working point is applied, requiring  $p_T^{\text{varcone30}}/p_T < 0.06$  and  $E_T^{\text{cone20}}/p_T < 0.06$  for electrons, and  $p_T^{\text{varcone30}}/p_T < 0.04$  and  $E_T^{\text{cone20}}/p_T < 0.15$  for muons.

To correct the lepton energy response of the simulation to match data, the calibration procedures are applied. The electron energy is corrected using calibrations derived based on Run 2 data, while muon momentum scale and resolution are determined using Run 3 data.

The efficiencies of lepton selections are different between data and simulations. Therefore, dedicated scale factors SFs are employed to account for differences in lepton identification, reconstruction, isolation, and trigger efficiencies. The scale factors are determined using specific physics processes, e.g.  $Z \rightarrow \ell\ell$ , based on data collected during specified periods. For electrons, the identifications SFs are derived using Run 2 data. All other lepton SFs, including the muon track-to-vertex-association (TTVA) SFs, are determined using Run 3 data.

Jet objects are reconstructed to determine the missing transverse energy  $E_T^{\text{miss}}$ . Topologically connected calorimeter cells are clustered to form jet candidates using the anti- $k_t$  algorithm, with a radius parameter of  $R = 0.4$  [83]. Information from the calorimeter and the inner detector is integrated using the particle flow algorithm [85]. The jet energy calibration is derived from Run 3 simulations and *in-situ* calibration [86], adapted for data collection at 13.6 TeV. Calibrated jet candidates must have  $p_T > 20$  GeV for  $|\eta| < 2.5$  and  $p_T > 30$  GeV for  $2.5 < |\eta| < 4.5$ . To mitigate the effect of pile-up jets, those with  $p_T$  below 60 GeV must meet the criteria of a neural-network-based jet vertex tagger (NNJVT) discriminant. NNJVT is an improvement over the JVT algorithm [91] utilized during the Run 2 data collection from 2015 to

2018.

The overlap removal algorithm is employed to resolve ambiguities when a single physical object is reconstructed as multiple objects. The algorithm is applied in the following specified order :

- If an electron candidate shares a track with another electron candidate, the one with the lower  $p_T$  is removed.
- Any electron candidate that shares a track with a muon is removed.
- The closest jet within a  $\Delta R$  of 0.2 from an electron candidate is removed.
- Any electron candidate within a  $\Delta R$  of 0.4 from a jet is removed.
- Jets with fewer than three associated tracks within a  $\Delta R$  of 0.2 are removed, as well as jets with a muon inner-detector track ghost-associated to them.
- Muon candidates within a  $\Delta R$  of 0.4 from a jet are removed.

The missing transverse momentum  $E_T^{\text{miss}}$  is employed to infer the presence of neutrinos originating from the  $W$  bosons via  $W \rightarrow \ell\nu$  decay. Tight criteria based on NNJVT are applied to determine  $E_T^{\text{miss}}$ , when forward jets have  $p_T > 30$  GeV.

Jet containing  $b$ -hadrons needed for the  $t\bar{t}$  events are identified using the 77% working point of the DL1d  $b$ -tagging algorithm [87, 88].

## 5.4 Event Selection

Events are selected if they include at least one lepton that is matched to an object identified by a single-lepton trigger, where the lepton refers to either an electron or a muon, as listed in Table 5.1. The single-lepton triggers have varying requirements for object identification, isolation, and  $p_T$ . Triggers with lower  $p_T$  thresholds enforce stricter identification and isolation conditions, while those with higher  $p_T$  thresholds have relaxed criteria. Corrections are applied accounting for the differences of efficiencies between data and simulations. Events are categorized into several categories to address various physics processes.

The di-lepton channels,  $e^+e^-$  and  $\mu^+\mu^-$ , are defined with events with exactly two leptons, having the same flavor and opposite electric charges. The invariant mass of the lepton pair should satisfy  $66 < m_{\ell\ell} < 116$  GeV. The region is enriched with  $Z \rightarrow \ell\ell$  events.

The single lepton channels, with final states  $e^+\nu$ ,  $e^-\bar{\nu}$ ,  $\mu^+\nu$  and  $\mu^-\bar{\nu}$ , correspond to  $W^+ \rightarrow e^+\nu$ ,  $W^- \rightarrow e^-\bar{\nu}$ ,  $W^+ \rightarrow \mu^+\nu$  and  $W^- \rightarrow \mu^-\bar{\nu}$  processes, respectively.

TABLE 5.1 – List of single lepton triggers for event selections, with a logical OR applied between them.

Electron triggers	Muon triggers
HLT_e26_lhtight_ivarloose_L1EM22VHI	HLT_mu24_ivarmedium_L1MU14FCH
HLT_e60_lhmedium_L1EM22VHI	HLT_mu50_L1MU14FCH
HLT_e140_lhloose_L1EM22VHI	

These channels are defined with events with exactly one lepton. Additionally, the transverse mass  $m_T^W$ , which is defined as  $m_T^W = \sqrt{2p_T^\nu p_T^\ell (1 - \cos \Delta\phi^\nu)}$ , should satisfy  $m_T^W > 50$  GeV. The observable carries the information of the  $W$ -boson mass. The missing transverse energy  $E_T^{\text{miss}}$  should be greater than 25 GeV to account for the presence of a neutrino in the final states of  $W$ -boson decay.

In addition to same flavor di-lepton channels,  $e\mu$  channels are defined using events with exactly one electron and one muon with opposite electric charges. These  $e\mu$  channels are exclusively used for extracting the ratio of  $t\bar{t}$  to  $W$ -boson fiducial cross sections. The  $b$ -tag counting method [105] is employed to extract the  $t\bar{t}$  cross section, therefore, events are required to contain  $b$ -jets. The  $e\mu$  channels are thus split into two channels according to the number of  $b$ -jets in the final states :  $e\mu 1b$  and  $e\mu 2b$  denoting the  $e\mu$  channels with exactly one or two  $b$ -jets, respectively.

Table 5.2 summarizes the event selections of  $W$  and  $Z$ -boson channels.

TABLE 5.2 – Event selections for  $W$  and  $Z$  channels.

Electron selections	$p_T > 27$ GeV $ \eta  < 2.47$ and veto of $1.37 <  \eta  < 1.52$
Muon selections	$p_T > 27$ GeV $ \eta  < 2.5$
$W$ -boson selections	Exactly one lepton $E_T^{\text{miss}} > 25$ GeV $m_T > 50$ GeV
$Z$ -boson selections	Exactly two same flavour opposite charged leptons $66 < m_{\ell\ell} < 116$ GeV

## 5.5 Background estimation

Leptonic final states of  $W$  and  $Z$  bosons exhibit large cross sections and provide a clean experimental signature. Therefore, the background contributions are minimized effectively with object and event selections, resulting in relative small contributions. Various background processes are considered in this analysis, categorized into two classes : the EW and top-quark backgrounds, and QCD multijet backgrounds. The EW and top-quark backgrounds are involved in both  $W$  and  $Z$  channels. The contribution of QCD multijet backgrounds is negligible in  $Z$ -boson channels due to the strict requirement of the lepton pair, so they are only taken into account when analysing the  $W$  channels.

### 5.5.1 Electroweak and top-quark backgrounds

MC simulations described in Section 5.2.2 are employed to model the background contributions of the EW and top-quark productions.

EW backgrounds are composed of the contributions from single-boson productions and diboson productions. Single-boson productions, such as  $W^\pm \rightarrow \tau^\pm \nu$  and  $Z \rightarrow \tau^+ \tau^-$ , contribute to both channels, with the leptonic decays of  $\tau$  leptons also considered as background. For the  $W$  channels,  $Z \rightarrow e^+ e^-$  and  $Z \rightarrow \mu^+ \mu^-$  are included, while  $W^\pm \rightarrow e^\pm \nu$  and  $W^\pm \rightarrow \mu^\pm \nu$  are relevant for the  $Z$ -boson channels.

Diboson productions ( $WW$ ,  $WZ$ , and  $ZZ$ ) can resemble  $W$  or  $Z$  events if one of the bosons decays hadronically, invisibly, or if the leptonic decay products do not meet the selection criteria. Likewise,  $t\bar{t}$  pair and  $Wt$  production can produce similar signatures to  $W$  or  $Z$  events with one or both  $W$  bosons decaying leptonically.

For  $W$ -boson selections, the background contribution from diboson and top-quark events is minimal due to their lower production cross sections. For  $Z$ -boson selections, all electroweak background contributions remain below the sub-percent level.

### 5.5.2 Multijet backgrounds for $W^\pm \rightarrow \ell \nu$ events

The QCD multijet (MJ) backgrounds mainly originate from leptons generated in semi-leptonic decays of heavy quarks, in-flight pion decays, and photon conversions, among other sources. Despite effective rejection by isolation criteria and, for  $W$  events, by the  $E_T^{\text{miss}}$  and  $m_T^W$  requirements, the high production cross section and considerable  $E_T^{\text{miss}}$  due to jet energy mis-measurements mean that MJ processes remain a significant background in  $W$ -boson analyses under high pile-up conditions, contributing up to approximately 4% of data events.

Given the difficulties in precisely simulating these processes, data-driven methods are used to estimate the MJ background, employing an advanced ABCD method similar to the one detailed in [42]. The stringent lepton quality requirements, such as Tight identification for electrons and tight isolation working points for both electrons and muons, are expected to reduce the MJ background fraction compared to similar Run 2 analyses. Additionally, the HLT adds an extra isolation requirement, which limits the MJ fraction in dedicated control regions, thereby increasing the challenge of estimating the MJ background in this analysis.

## General procedure

A typical ABCD method is employed to estimate the QCD multijet backgrounds in the  $W$ -boson channels. In this method, one assumes that the shapes of the MJ kinematic distributions are relatively stable, regardless of lepton isolation [42]. Any residual biases are addressed using a fit within the nominal method.

Events are categorized into four regions, defined with regard to the lepton isolations, as well as  $m_T^W$  and  $E_T^{\text{miss}}$  kinematic selections. The signal region (SR) is defined using the same selections as the  $W$  channels in Table 5.2, with tight isolation criteria for the leptons, and specific cuts on the  $m_T^W$  and  $E_T^{\text{miss}}$  observables. In contrast, the fit region (FR) employs the same the isolation requirement but inverts kinematic cuts on  $m_T^W$  and  $E_T^{\text{miss}}$  leading to an increased proportion of multijet backgrounds.

Two anti-isolated control regions are defined with inverted lepton isolation. The

TABLE 5.3 – The definitions of the ABCD regions for the QCD multijet background extraction.

Fit region (FR)	Signal region (SR)
Lepton $p_T > 25$ GeV $E_T^{\text{miss}} < 25$ GeV $m_T^W < 50$ GeV Pass isolation	Lepton $p_T > 25$ GeV $E_T^{\text{miss}} > 25$ GeV $m_T^W > 50$ GeV Pass isolation
Control region 1 (CR1)	Control region 2 (CR2)
Lepton $p_T > 25$ GeV $E_T^{\text{miss}} < 25$ GeV $m_T^W < 50$ GeV Fail isolation	Lepton $p_T > 25$ GeV $E_T^{\text{miss}} > 25$ GeV $m_T^W > 50$ GeV Fail isolation

anti-isolated control region 1 (CR1) is established by inverting the cuts on  $m_T^W$  and  $E_T^{\text{miss}}$ , while control region 2 (CR2) uses the same kinematic cuts as the SR. As a result, events with anti-isolated leptons populate both CR1 and CR2, enriching them with multijet backgrounds.

The initial MJ template is obtained in the anti-isolated CR1 by removing the electroweak and top backgrounds from the data. This template is then used to perform a fraction fit in the FR to determine the MJ fraction. The fit result is extrapolated to the SR using transfer factors, defined as the ratio of MJ events in CR2 to those in CR1.

The definitions of the ABCD regions are summarized in Table 5.3, which guarantee that the kinematic selections for CR1 and FR do not overlap with those for CR2 and SR.

With the ABCD regions defined above, the QCD MJ contribution in the SR can be determined using the following general procedure :

- The first step is to derive the initial MJ template as a function of given discriminant observable in the anti-isolated CR1, which is enriched with QCD multijet backgrounds. The non-QCD components, electroweak and top event yields, are modelled using the MC predictions and subtracted from the data distribution ( $N_{\text{data}}^{\text{CR1}}$ ), resulting in :

$$N_{\text{MJ}}^{\text{CR1}} = N_{\text{data}}^{\text{CR1}} - N_{\text{EW+Top}}^{\text{CR1}} \quad (5.1)$$

where  $N_{\text{MJ}}^{\text{CR1}}$  and  $N_{\text{EW+Top}}^{\text{CR1}}$  represent the estimated MJ background and electroweak and top background distributions, respectively.

- The multijet template obtained from CR1 serves as the input of the fractional fit in the fit region. The data event yields ( $N_{\text{data}}^{\text{FR}}$ ) is considered to consist of two components : electroweak and top events ( $N_{\text{EW+Top}}^{\text{FR}}$ ), and multijet background ( $N_{\text{MJ}}^{\text{CR1}}$ ). The contributions of each component are adjusted in the fits to give the best description of the data distribution in the FR.

$$N_{\text{data}}^{\text{FR}} = \alpha \cdot N_{\text{EW+Top}}^{\text{FR}} + \mu^{\text{MJ}} \cdot N_{\text{MJ}}^{\text{CR1}} \quad (5.2)$$

TABLE 5.4 – A list of single electron (left) or muon (right) pre-scaled supporting triggers used to select events for the multijet background templates. A logical OR is applied between the triggers.

Electron triggers	Muon triggers
HLT_e20_lhvloose_L1EM15VH	HLT_mu22_L1MU14FCH
HLT_e26_idperf_tight_nogsf_L1EM22VHI	

where  $\alpha$  and  $\mu^{\text{MJ}}$  are determined with the fit. MJ event yields in the FR can thus be represented as  $N_{\text{MJ}}^{\text{FR}} = \mu^{\text{MJ}} \cdot N_{\text{MJ}}^{\text{CR1}}$ .

- $N_{\text{MJ}}^{\text{FR}}$  is extrapolated to the signal region with the transfer factor  $\varepsilon$ , which is defined as the ratio between the MJ event yields in CR2 and CR1 :

$$\varepsilon = \frac{N_{\text{MJ}}^{\text{SR}}}{N_{\text{MJ}}^{\text{FR}}} \approx \frac{N_{\text{MJ}}^{\text{CR2}}}{N_{\text{MJ}}^{\text{CR1}}} = \frac{N_{\text{data}}^{\text{CR2}} - N_{\text{EW+Top}}^{\text{CR2}}}{N_{\text{data}}^{\text{CR1}} - N_{\text{EW+Top}}^{\text{CR1}}} . \quad (5.3)$$

Then the final MJ event yields in the SR can be written as :

$$N_{\text{MJ}}^{\text{SR}} = \varepsilon \cdot N_{\text{MJ}}^{\text{FR}} . \quad (5.4)$$

The multijet event yields can be extracted using various discriminant variables. In this measurement, the  $W$  transverse mass  $m_{\text{T}}^W$  and missing transverse energy  $E_{\text{T}}^{\text{miss}}$  are employed. The differences between MJ event yields extracted from two variables are treated as systematic uncertainties.

## Multijet extraction

This section details the procedure to extract the MJ yields in this analysis. As the nominal triggers (Table 5.1) include an isolation requirement, they cannot be used for this study. Instead, pre-scaled supporting triggers are used to select events in the anti-isolated control regions. These supporting triggers correspond to a lower luminosity compared to the nominal ones but have looser or no isolation requirements to minimize bias from isolation cuts at the trigger level. Table 5.4 lists the pre-scaled triggers used in this analysis.

Regarding the general procedure outlined in the previous section, a potential isolation-related bias can be introduced in the normalization factor when the MJ behavior in the isolated regions (SR and FR) is approximated using the anti-isolated regions (CR2 and CR1).

To address this bias, the anti-isolated control regions (CR1 and CR2) are split into several slices with regard to the lepton isolation variables,  $p_{\text{T}}^{\text{varcone30}}$ , which represents the sum of the transverse momentum of charged tracks within a cone around the lepton tracks.  $p_{\text{T}}^{\text{varcone30}}$  has a cone size dependent on the lepton  $p_{\text{T}}$ , which decreases as the  $p_{\text{T}}^{\ell}$  grows and the maximum cone radius is 0.3. Table 5.5 shows the definitions of the isolation slices. The track variables differ between the electron and muon channels

TABLE 5.5 – Width and boundaries of isolation scan slices of anti-isolated control regions (CR1 and CR2), used for the  $W \rightarrow e\nu$  and  $W \rightarrow \mu\nu$  channels. topoetcone20 refers to the sum of transverse energy of topo-clusters within a cone around the lepton. ptvarcone30 denotes the transverse energy of nearby charged tracks.

	$W \rightarrow e\nu$	$W \rightarrow \mu\nu$
Fixed cut	topoetcone20/pT < 0.06	topoetcone20/pT < 0.15
Scan variable	ptvarcone30_Nonprompt_ MaxWeightTTVALooseCone_pt1000/pT	ptvarcone30_Nonprompt_All_ MaxWeightTTVA_pt1000/pT
Scan starting point	0.06	0.04
Scan width	0.06	0.04
Number of slices	4	4

due to varying requirements on the charged tracks. The scan of the track isolation variables starts at  $p_T^{\text{varcone30}} = 0.06$  for the electron channel, with increments of 0.06 per step. For the muon channel, the scan begins at  $p_T^{\text{varcone30}} = 0.04$ , with increments of 0.04 per step. Both channels use four track isolation slices. Figure 5.2 displays the track isolation slices. Calorimeter isolation remains at its nominal values.

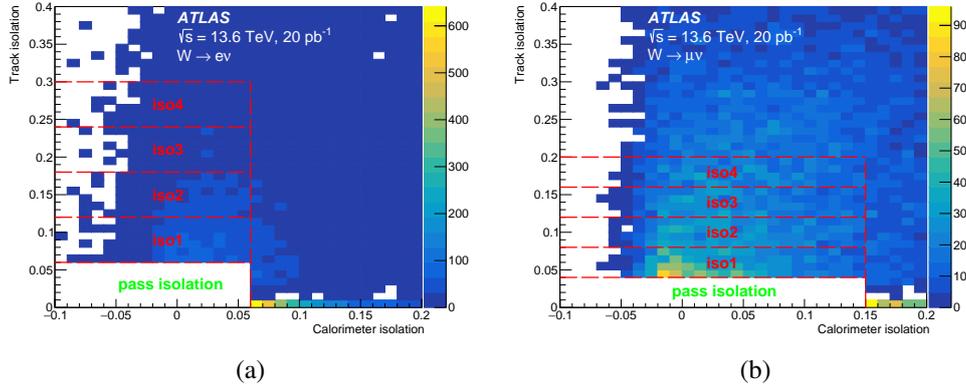
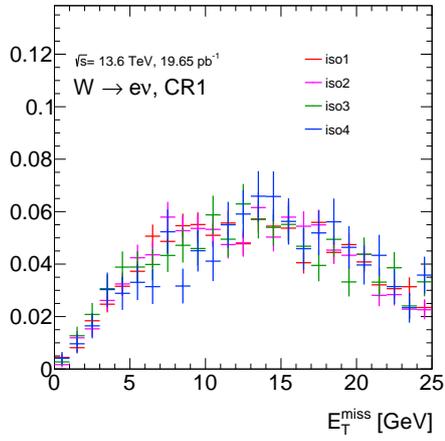


FIGURE 5.2 – Track isolation slices for anti-isolated control regions (a) in the electron channel and (b) in the muon channel. The pre-scaled supporting triggers are used, which corresponds to smaller integrated luminosity.

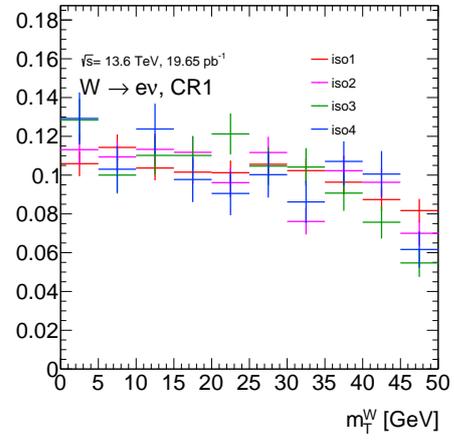
The multijet determination procedure is then repeated for each isolation slices, resulting in a set of multijet yields in the SR with regard to isolation variables. The MJ yields from each isolation slice are used to perform an extrapolation to the track isolation selection applied in the SR to determine the final MJ yields.

The electroweak and top contributions in CR1 are determined with MC simulations and subtracted from the data in CR1 for each track isolation slices. Figures 5.3-5.4 illustrate the comparison between resulting MJ templates after normalizing to unity.

The MJ templates derived from CR1 is used to determine the MJ distributions in

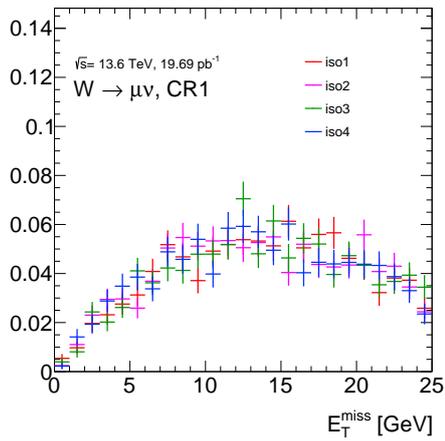


(a)

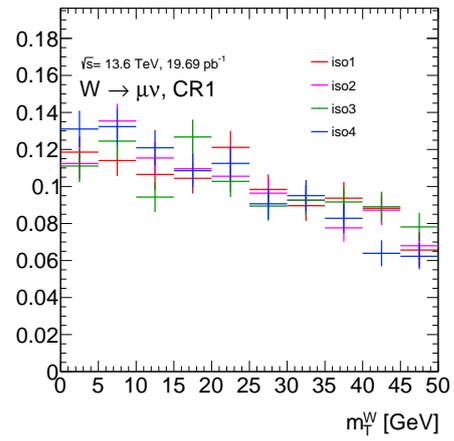


(b)

FIGURE 5.3 – CR1 multijet templates for the electron channel for (a)  $E_T^{\text{miss}}$  and (b)  $m_T^W$ .



(a)



(b)

FIGURE 5.4 – CR1 multijet templates for the muon channel for (a)  $E_T^{\text{miss}}$  and (b)  $m_T^W$ .

the fit region. The profiling likelihood (PLH) technique is used to extract the fraction of the MJ yields. The corresponding likelihood formula can be constructed based on Eq. (5.2) :

$$L(\vec{n}; T, \vec{\theta}) = \text{Pois}(N_{\text{data}}^{\text{FR}} | \mu^{\text{MJ}} \cdot N_{\text{MJ}}^{\text{CR1}}(\vec{\theta}) + N_{\text{EW+Top}}^{\text{FR}}(\vec{\theta})) \prod_{i \in \text{NPs}} G(\theta_i) \quad (5.5)$$

where  $\theta_i$  denotes nuisance parameter, constrained by a Gaussian term  $G(\theta_i)$ .

Systematic uncertainties detailed in Section 5.7 are incorporated into the fit as nuisance parameters. Experimental uncertainties arise from sources such as luminosity, pile-up, lepton efficiency and energy corrections, as well as jet energy corrections. Additionally, modeling uncertainties in the predicted cross sections of electroweak and top productions are also taken into account. As the anti-isolated control regions are created using pre-scaled supporting triggers, the resulting multijet templates experience significant statistical fluctuations. To address this, additional nuisance parameters ( $\gamma$  parameters) are introduced for each bin, incorporating the statistical uncertainty in the multijet background via a Poisson term using Beeston-Barlow method [132, 133].

As an example, Figures 5.5- 5.7 show pre-fit and post-fit distributions of the MJ and other processes obtained from isolation slice 1 (iso1) for the  $W^- \rightarrow e^- \bar{\nu}$  channel based on  $E_{\text{T}}^{\text{miss}}$  and on  $m_{\text{T}}^{\text{W}}$  respectively, as well as the post-fit multijet signal strength, nuisance parameters and  $\gamma$  parameters.

The transfer factor  $\varepsilon$  is calculated separately for each isolation slice based on Eq. (5.3) :

$$\varepsilon_i = \frac{N_{\text{MJ}}^{\text{CR2},i}}{N_{\text{MJ}}^{\text{CR1},i}}$$

where the superscript  $i$  denotes the  $i$ -th isolation slice. Therefore, the multijet yields for the  $i$ -th isolation slice are obtained using  $\varepsilon_i$  and the MJ yields in the FR derived from the  $i$ -th slice of CR1 :

$$N_{\text{MJ}}^{\text{SR}}(i) = \varepsilon_i \cdot N_{\text{MJ}}^{\text{FR},i}.$$

To address the isolation-related bias, the MJ yields in SR obtained from each isolation slice are treated as a function of the isolation of the CR slice. The track isolation variable slices are uniformly spaced, with upper limits set at 0.3 for the electron channel and 0.2 for the muon channel. Fits are then performed to extrapolate the SR MJ yields to those expected from an ideal SR-like CR. Linear and quadratic forms are used for these extrapolations, and the differences between the results of the two forms are considered as systematic uncertainties. The final SR MJ yields for the linear (quadratic) fit are determined by averaging the outcomes of the linear (quadratic) fits derived from  $m_{\text{T}}^{\text{W}}$  and  $E_{\text{T}}^{\text{miss}}$ -based SR yields.

Figure 5.8 illustrates the linear (dashed line) and quadratic (solid line) fits to the relative MJ yields based on the track isolation slice. The central value for each slice is determined by the average track isolation of the events in that slice. The uncertainty

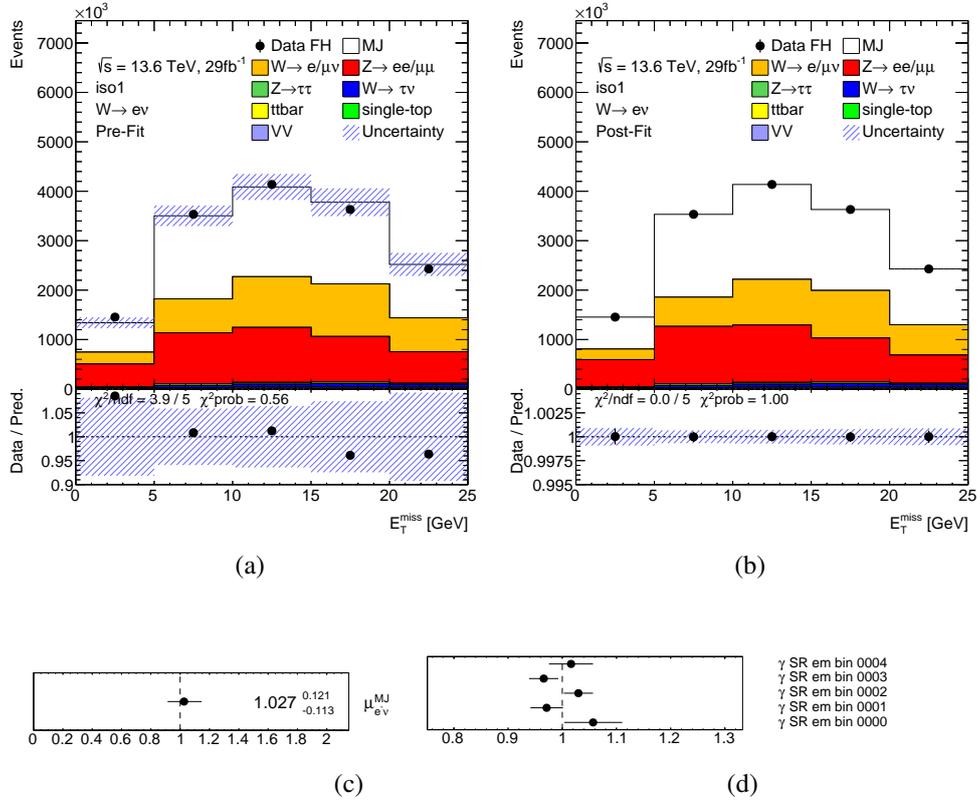


FIGURE 5.5 – Pre-fit (a) and post-fit (b) distributions for the  $W^- \rightarrow e^- \bar{\nu}$  channel on the  $E_T^{\text{miss}}$  variable using the multijet template obtained from isolation slice 1 (iso1). Post-fit results including the multijet signal strength (c) and  $\gamma$  parameters (d) are also illustrated.

from the MJ fit in the FR is propagated to the MJ yield for each point and contributes to the final extrapolation uncertainty.

Quadratic fits are found to give a better description than linear fits in each channel, as confirmed by a  $\chi^2$  criterion. The combined mean of the two quadratic extrapolations, derived from the  $E_T^{\text{miss}}$  or  $m_T^W$  distribution, is used as the central value for the final MJ yield in each channel, labeled  $f_{\text{MJ}}^{\text{SR}}$  in the figure.

Table 5.6 shows the final MJ event yields in the SR determined using the nominal quadratic fits. MJ yields obtained using  $m_T$  and  $E_T^{\text{miss}}$ , as well as the combined average are displayed. The relative MJ event yields with regard to the data event yields in the SR is presented in Table 5.7.

Finally, as the quadratic fits show better consistency between the individual multijet yields in the track isolation scan compared to the linear fits, the combined mean of the  $E_T^{\text{miss}}$  and  $m_T^W$  results from the quadratic fits is adopted as the final multijet yield. The uncertainty for this yield is calculated as the quadratic sum of the uncertainty from the quadratic fit and the difference between the linear and quadratic fit results.

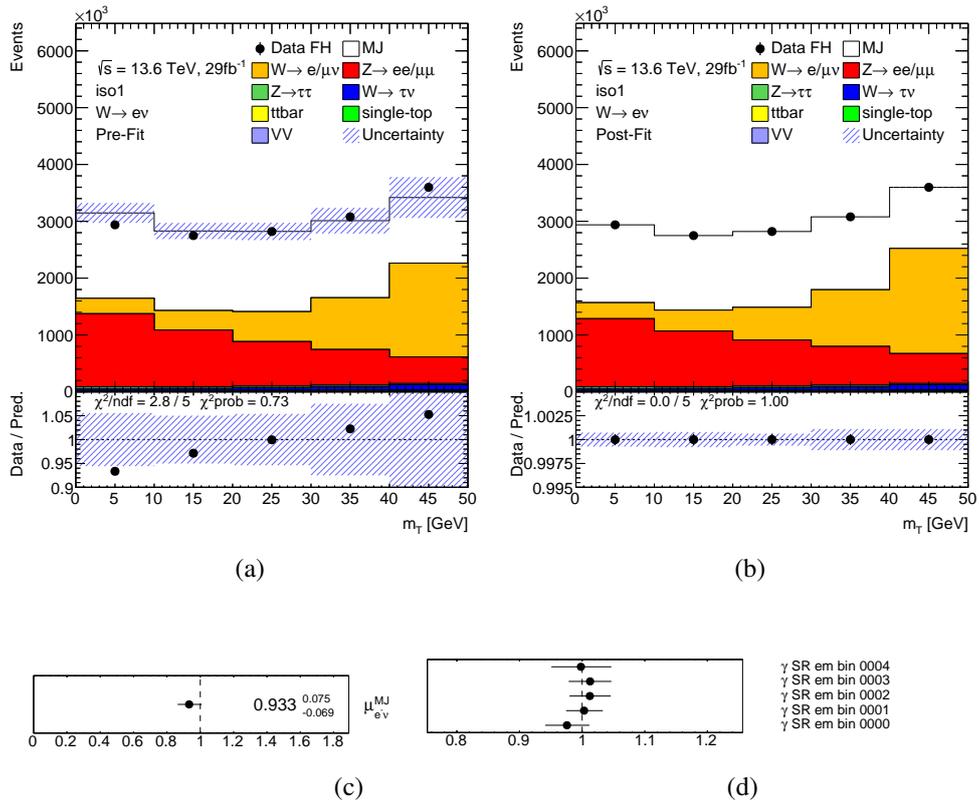


FIGURE 5.6 – Pre-fit (a) and post-fit (b) distributions for the  $W^- \rightarrow e^- \bar{\nu}$  channel on the  $m_T^W$  variable using the multijet template obtained from isolation slice 1 (iso1). Post-fit results including the multijet signal strength (c) and  $\gamma$  parameters (d) are also illustrated.

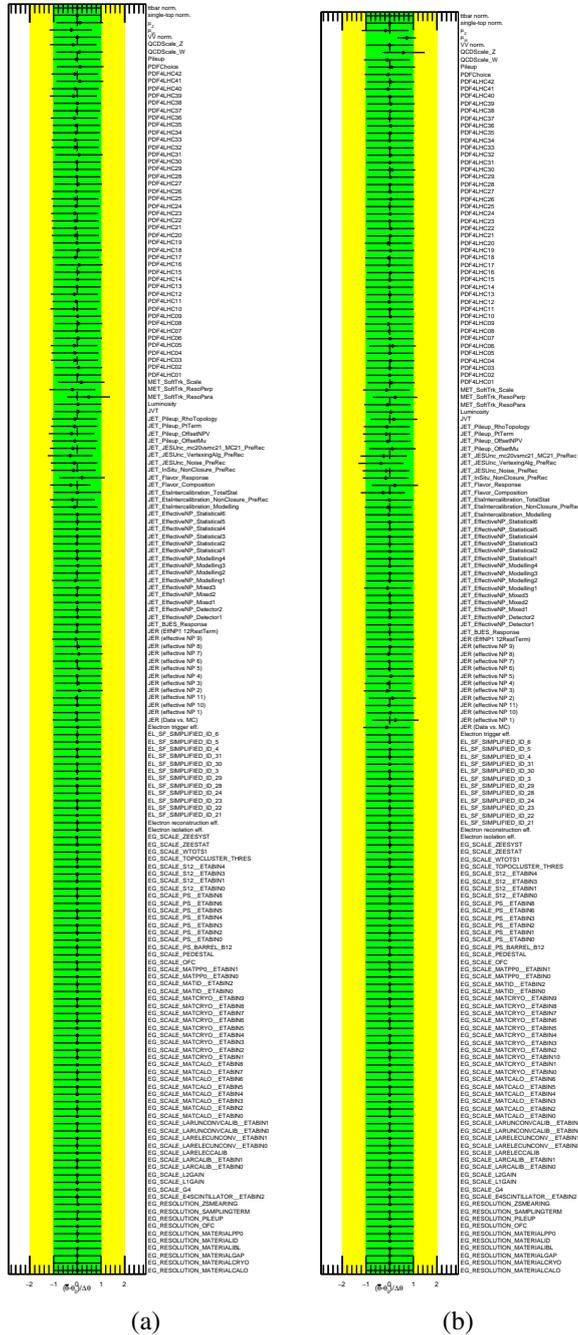


FIGURE 5.7 – Nuisance parameters for the fit in the  $W^- \rightarrow e^- \bar{\nu}$  channel on the (a)  $E_T^{\text{miss}}$  and (b)  $m_T^W$  variables using the multijet template obtained from isolation slice 1 (iso1).

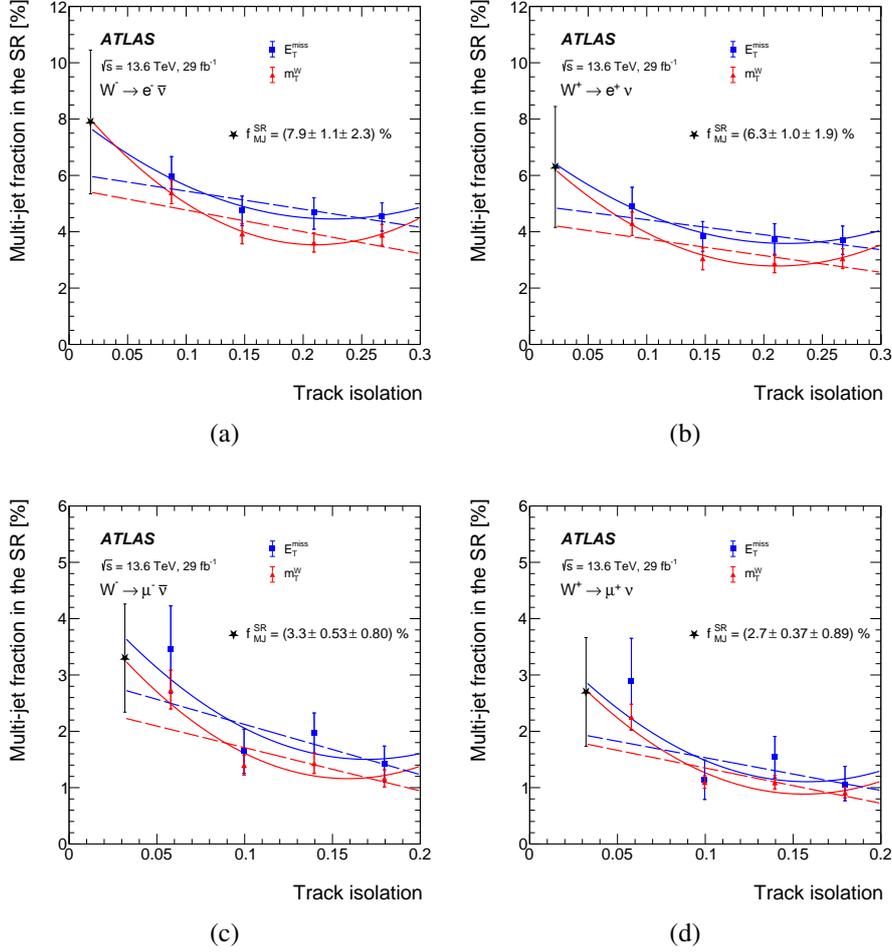


FIGURE 5.8 – The relative multijet yield in the SR is presented as a function of the track isolation variable for the channels : (a)  $W^- \rightarrow e^- \bar{\nu}$ , (b)  $W^+ \rightarrow e^+ \nu$ , (c)  $W^- \rightarrow \mu^- \bar{\nu}$ , and (d)  $W^+ \rightarrow \mu^+ \nu$ . Squares and triangles indicate the  $E_T^{\text{miss}}$  and  $m_T^W$  input measurements, respectively. The solid curves show the quadratic function extrapolation to the SR, while the dashed lines represent the linear function extrapolation. The  $x$ -axis reflects the average position of a track isolation slice. The star denotes the final MJ fraction ( $f_{\text{MJ}}^{\text{SR}}$ ), computed by averaging the quadratic fits for each channel. The first uncertainty represents the combined uncertainty, including the discrepancy between the  $m_T^W$  and  $E_T^{\text{miss}}$  quadratic fits, while the second uncertainty stems from the difference between the linear and quadratic fit results.

TABLE 5.6 – Multijet yield in the SR, derived from  $m_T^W$ ,  $E_T^{\text{miss}}$  and for the combination. The uncertainty is derived from the quadratic fit of the isolation scan, where the error on each of the scan points comes from the post-fit uncertainty on  $\mu_{\text{MJ}}$  from template fits.

Channel	$m_T^W$	$E_T^{\text{miss}}$	Combined mean
$W^- \rightarrow e^- \bar{\nu}$	$4016760 \pm 607784$	$3863240 \pm 986783$	$3974538 \pm 539791$
$W^+ \rightarrow e^+ \nu$	$3853102 \pm 739362$	$3982713 \pm 1172954$	$3889957 \pm 638759$
$W^- \rightarrow \mu^- \bar{\nu}$	$2072496 \pm 371487$	$2326560 \pm 839368$	$2114110 \pm 424202$
$W^+ \rightarrow \mu^+ \nu$	$2202575 \pm 311311$	$2318540 \pm 1146955$	$2210532 \pm 322044$

TABLE 5.7 – Relative Multijet yield in the SR, derived from  $m_T^W$ ,  $E_T^{\text{miss}}$  and from the combination, shown as a fraction [%] of the total yield in the SR. The uncertainty is derived from the quadratic fit of the isolation scan, where the error on each of the scan points comes from the post-fit uncertainty on  $\mu_{\text{MJ}}$  from template fits.

Channel	$m_T^W$	$E_T^{\text{miss}}$	Combined mean
$W^- \rightarrow e^- \bar{\nu}$	$7.98 \pm 1.21$	$7.67 \pm 1.96$	$7.9 \pm 1.03$
$W^+ \rightarrow e^+ \nu$	$6.2 \pm 1.19$	$6.41 \pm 1.89$	$6.26 \pm 1.01$
$W^- \rightarrow \mu^- \bar{\nu}$	$3.26 \pm 0.58$	$3.66 \pm 1.32$	$3.33 \pm 0.53$
$W^+ \rightarrow \mu^+ \nu$	$2.73 \pm 0.39$	$2.87 \pm 1.42$	$2.74 \pm 0.37$

These final multijet yields are also detailed in Table 5.8.

### 5.5.3 Multijet background for $Z \rightarrow \ell\ell$ events

In the  $Z$ -boson channels, a conservative upper limit on the MJ background is determined based on the number of charge misidentified leptons, estimated using the  $m_{\ell\ell}$  sidebands from a same-sign lepton selection. For the electron channel, this contribution is found to be below the sub-percent level, and for the muon channel, it is even smaller. The systematic uncertainties are also negligible. Consequently, the MJ contribution in the  $Z$ -boson channels is not included in this analysis.

## 5.6 Theory predictions

This section details the theoretical predictions of the high order cross sections for the  $W^\pm$  and  $Z$  boson production, their ratios, as well as ratio of  $t\bar{t}$  over  $W$  boson production cross section.

Uncertainties in the cross sections resulting from missing higher-order terms, variations in the strong coupling constant  $\alpha_s$ , and in the input parton distribution functions (PDFs) are also estimated and reported.

The cross sections are calculated with high precision, up to NNLO+NNLL QCD

TABLE 5.8 – Final multijet yield in the SR, with the central values derived with quadratic fits. The uncertainty is given by the quadratic sum of the uncertainty on the quadratic fit and the difference between the linear and quadratic fit results.

Channel	Multijet yield
$W^- \rightarrow e^- \bar{\nu}$	$3974538 \pm 1288162$
$W^+ \rightarrow e^+ \nu$	$3889957 \pm 1322403$
$W^- \rightarrow \mu^- \bar{\nu}$	$2114110 \pm 768724$
$W^+ \rightarrow \mu^+ \nu$	$2210532 \pm 833411$

and NLO EW accuracy. These calculation employs the DYTURBO-1.3.1 [134-137] and ReneSANCe-1.3.3 [138, 139]. The results from these two tools are combined using an additive prescription, which was also applied in the measurement of  $W/Z$  cross sections in the 7 TeV ATLAS data [96].

### 5.6.1 Electroweak scheme and parameters

The predicted cross sections are highly dependent on the chosen electroweak parameter scheme at LO in electroweak coupling. The  $G_\mu$  scheme [140] is employed in the calculation, which uses the particle masses and Fermi constant  $G_F$  as the input parameters. The  $G_\mu$  scheme provides high accurate and reliable theoretical calculations. When considering leptonic final states produced by off-shell electroweak bosons, the complex-mass scheme is utilized. This method employs complex masses for the  $W$  and  $Z$  bosons, and the electroweak mixing angle is defined as  $\cos^2 \theta_W^2 = (m_W^2 - i\Gamma_W m_W) / (m_Z^2 - i\Gamma_Z m_Z)$ . Additionally, the fine-structure constant  $\alpha$  is expressed as  $\sqrt{2} G_\mu m_W^2 \sin^2 \theta_W / \pi$ . The values for  $G_F$ ,  $m_W$ ,  $\Gamma_W$ ,  $m_Z$ , and  $\Gamma_Z$  are sourced from the PDG and are summarized in Table 5.9 for reference.

TABLE 5.9 – Electroweak input parameters used for the calculation of inclusive electroweak vector boson cross sections.

Parameter	Value
$m_W$	80.385 GeV
$\Gamma_W$	2.091 GeV
$\Gamma_Z$	2.4952 GeV
$G_F$	$1.16638 \times 10^{-5} \text{ GeV}^{-2}$
$1/\alpha_{\text{QED}}(m_Z)$	132.233229791

All processes analyzed here use the 5-flavor scheme (FS) where the bottom quark is considered as a light parton and its mass is set to zero ( $m_b = 0$ ) in the

short-distance cross section. It resums large initial state logarithms into the  $b$ -PDFs, resulting in more stable predictions and simplifying the computation of higher-order corrections. The top quark, however, is considered massive and unstable, with a mass of  $m_t = 173.2$  GeV and a decay width of  $\Gamma_t = 1.44262$  GeV. Regarding the production of a single  $W^\pm$  boson, the CKM matrix elements are assigned the following values :

$$\begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} = \begin{bmatrix} 0.97427 & 0.2253 & 0.00351 \\ 0.2252 & 0.97344 & 0.0412 \\ 0 & 0 & 0 \end{bmatrix}. \quad (5.6)$$

In the context of the five-flavor scheme, the elements in the last row are set to zero, as the top mass has considerably larger mass, which is larger than  $W$ -boson mass ( $m_W < m_t$ ). The five-flavor scheme emphasizes the interactions of the lighter quarks ( $u, d, c, s$  and  $b$ ) and ensures that predictions involving these quarks are consistent with both experimental data and theoretical models.

### 5.6.2 PDFs

The calculation of the Drell-Yan cross section in high-energy physics is fundamentally based on PDFs to characterize the initial state of the colliding hadrons. Predictions are calculated with various PDF sets, which are summarized in Table 5.10. In the following list, a brief introduction is given to the PDF sets involved in this calculation, highlighting their technical features and the methods used to compute the corresponding PDF uncertainty for a given observable.

**PDF4LHC21 :** Presented in Ref. [141]. Only NNLO grids are provided. The Hessian set PDF4LHC21\_40 is used, providing the central value (member 0) and the 68% probability intervals (members 1 through 40). Denoting as  $\mathcal{F}^{(k)}$  a physical observable computed using the  $k$ -th eigenvector direction, the PDF uncertainty is given by

$$\delta^{\text{PDF}} \mathcal{F} = \sqrt{\sum_{k=1}^{N_{\text{eig}}} (\mathcal{F}^{(k)} - \mathcal{F}^{(0)})^2}, \quad (5.7)$$

which gives the 68% CL uncertainty according to the symmetric Hessian prescription.

**CT18 :** Presented in Ref. [34]. NNLO grids are provided for the central fit and 58 eigenvector sets, with 29 eigenvectors in both the ‘+’ and ‘-’ directions, representing the 90% confidence level (CL) interval. The 90% CL PDF uncertainty is calculated according to the asymmetric Hessian prescription

$$\delta^{\text{PDF}} \mathcal{F} = \frac{\sqrt{\sum_{k=1}^{N_{\text{eig}}} (\mathcal{F}^{(+)} - \mathcal{F}^{(-)})^2}}{2}. \quad (5.8)$$

In the final results, the 68% CL is quoted, obtained applying a rescale factor to Eq. (5.8).

Additionally, an alternative set (CT18A) which also includes the ATLAS precision 7 TeV  $W/Z$  data [96] is involved as well.

**MSHT20 :** Presented in Refs. [35-37]. NNLO grids are available for the central fit and 64 eigenvector sets (corresponding to 32 eigenvectors in the ‘+’ and ‘-’ directions), yielding the 68% confidence level interval. The PDF uncertainty is determined using Eq. (5.8).

**NNPDF4.0 :** Presented in Ref. [38]. This is the latest release of the NNPDF family sets, providing NNLO grids for the central fit and 100 Monte Carlo replicas. To calculate the best value of the physical observable  $\mathcal{F}$  and its associated PDF uncertainty, the ensemble  $\mathcal{S}$  of  $\mathcal{F}$  values is first computed using the 100 replicas in the set.

$$\mathcal{S} = \{\mathcal{F}^{(1)}, \dots, \mathcal{F}^{(100)}\}. \quad (5.9)$$

Next, the elements from Eq. (5.9) are sorted in ascending order and the lowest and the highest 16% of the ordered replicas are removed, thus retaining the central 68% replicas. The resulting 68% CL interval, denoted as  $[\mathcal{F}_{16}, \mathcal{F}_{84}]$ , is then used to estimate the upper and down PDF errors as follows :

$$\delta^{\text{PDF,up}}\mathcal{F} = \mathcal{F}_{84} - \mathcal{F}_0, \quad \delta^{\text{PDF,down}}\mathcal{F} = \mathcal{F}_0 - \mathcal{F}_{16}. \quad (5.10)$$

The median of the ensemble  $\mathcal{S}$  in Eq. (5.9) is taken as the best value of the observable.

**ABMP16 :** Presented in Ref. [142]. NNLO grids are available for the central fit (member 0) and 29 eigenvectors (members 1 to 30). The 68% confidence level interval can be derived using the symmetric Hessian prescription of Eq. (5.7). The set ABMP16als118\_5\_nnlo is employed for NNLO results.

**ATLASpdf21 :** Presented in Ref. [39]. This set of PDFs is derived from a variety of measurements taken from  $pp$  collisions at  $\sqrt{s} = 7, 8,$  and 13 TeV, alongside deep inelastic scattering data from  $ep$  collisions at the HERA collider. NNLO grids are provided for the central fit (member 0) and 42 eigenvectors (members 1 to 43). The experimental uncertainties for this set have been assessed with an enhanced  $\chi^2$  tolerance,  $\Delta\chi^2 = T^2$ , where  $T = 3$ . The 68% CL interval can be calculated using the symmetric Hessian method described in Eq. (5.7).

### 5.6.3 Electroweak corrections

In the Drell-Yan production, the primary component of higher-order electroweak corrections is the QED radiation emitted from the final-state leptons. This contribution is included in the DY MC samples using PHOTOS [121] and then processed through the comprehensive ATLAS detector simulation. Since the data are unfolded for QED FSR

TABLE 5.10 – Summary of the PDF sets used in the computation of theoretical predictions for the fiducial cross sections.

	Members	Error type	Reference
PDF4LHC21	43	68% CL symmetric Hessian	[141]
CT18, CT18A	59	90% CL asymmetric Hessian (converted to 68%CL)	[34]
MSHT20	65	68% CL asymmetric Hessian	[35-37]
NNPDF4.0	101	Monte Carlo	[38]
ABMP16	30	68% CL symmetric Hessian	[142]
ATLASpdf21	43	68% CL symmetric Hessian	[39]

effects along with other detector effects, this component is not considered separately. The remaining NLO EW corrections consist of contributions from initial-state photon radiation, EW loop corrections, and the interference between initial-state and final-state photon emissions.

To achieve comprehensive predictions, the NLO EW corrections are integrated with the NNLO QCD predictions, which are calculated using DYTURBO, through an additive approach. The additive prescription is based on the assumption that the absolute contribution of the EW correction remains independent of the QCD calculation order. Consequently, the combined cross section can be written as :

$$\sigma_{\text{NNLO QCD}}^{\text{NLO EW}} = \sigma_{\text{NNLO QCD}}^{\text{LO EW}} \left( 1 + \frac{K^{\text{EW}} - 1}{K_{\text{QCD}}} \right) \quad (5.11)$$

where the  $K$ -factors, the electroweak correction  $K^{\text{EW}}$  and QCD correction  $K_{\text{QCD}}$  are defined as :

$$K^{\text{EW}} = \frac{\sigma_{\text{LO QCD}}^{\text{NLO EW}}}{\sigma_{\text{LO QCD}}^{\text{LO EW}}}, \quad K_{\text{QCD}} = \frac{\sigma_{\text{NNLO QCD}}^{\text{LO EW}}}{\sigma_{\text{LO QCD}}^{\text{LO EW}}}.$$

The NLO EW corrections applied to the NNLO+NNLL QCD fiducial cross sections are around  $-0.4\%$  for the  $W^+$  and  $W^-$  channels, and approximately  $-0.3\%$  for the neutral-current channel. Furthermore, the leading contributions of the two-loop EW corrections have been computed, revealing an impact of less than  $0.1\%$ .

#### 5.6.4 Uncertainties

The theory uncertainty in cross-section calculation comprises contributions from statistical fluctuations, PDF uncertainties and QCD scale uncertainties. The estimation of the PDF uncertainty depends on the type of PDF errors, which is detailed in the Section 5.6.2.

QCD scale uncertainty arises from the choice of the renormalization scale  $\mu_R$  and the factorization scale  $\mu_F$ , and  $\mu_R$  is used to renormalize the strong coupling constant  $\alpha_s$ , addressing the running behavior of  $\alpha_s$ . As the factorization theorem in Eq. (2.20) states, the cross section can be expressed in a convolution of the PDFs with the partonic cross section, which are affected by the choice of  $\mu_F$ .

TABLE 5.11 – Fiducial cross sections predictions for  $\alpha_s$  variations obtained using PDF4LHC21.

	PDF4LHC21		
	$\sigma(\alpha_s = 0.117)$ [pb]	$\sigma(\alpha_s = 0.119)$ [pb]	$\delta^{\alpha_s} \sigma$
$pp \rightarrow \ell^+ \ell^-$	736.52	746.92	0.7%
$pp \rightarrow \ell^- \bar{\nu}_\ell$	3248.86	3274.96	0.4%
$pp \rightarrow \ell^+ \nu_\ell$	3922.16	3985.42	0.8%

In practice, the uncertainty arising from the choice of  $\mu_F$  and  $\mu_R$  is evaluated using the standard seven-point variations for the renormalization scale ( $\mu_R$ ) and factorization scale ( $\mu_F$ ). This method involves independently varying  $\mu_R$  and  $\mu_F$  by factors of 1/2 and 2, resulting in the combinations  $(\mu_R/2, \mu_F/2)$ ,  $(2\mu_R, 2\mu_F)$ ,  $(\mu_R, 2\mu_F)$ ,  $(2\mu_R, \mu_F)$ ,  $(\mu_R, \mu_F/2)$ , and  $(\mu_R/2, \mu_F)$ . The corresponding cross sections are calculated for each combination of  $\mu_R$  and  $\mu_F$ . The final uncertainty is obtained by taking the envelope of the resulting predictions with respect to the nominal prediction.

In addition to the uncertainty sources discussed above, the  $\alpha_s$  uncertainty is evaluated by calculating the cross sections using the PDF4LHC21 set. The PDF4LHC21 combination has the assumption:  $\alpha_s(m_Z^2) = 0.1180 \pm 0.0010$  [141], in which the uncertainty corresponds to the 68% CL. Therefore, the  $\alpha_s$  uncertainty can be computed as :

$$\delta^{\alpha_s} \sigma = \frac{\sigma(\alpha_s = 0.119) - \sigma(\alpha_s = 0.117)}{2}. \quad (5.12)$$

Table 5.11 presents the  $\alpha_s$  uncertainties for each processes.

### 5.6.5 Results

The calculated total and fiducial inclusive cross sections for this analysis are reported in this section. Table 5.12 provides the definitions of the fiducial volume in the  $W$  and  $Z$  channels. The BORN level leptons are used, which refers to leptons prior to QED final state radiation. Total cross sections of  $W$  production is calculated without applying any requirement, while for the  $Z$  channel, the invariant mass of the lepton pair is required to be within the mass window  $66 < m_{\ell\ell} < 116$  GeV, which is the same as event selections described in Section 5.4.

Table 5.13 presents the fiducial NNLO+NNLL QCD + NLO EW  $W$  and  $Z$  cross sections. The results are provided for each PDF set described previously, and include errors due to PDF and scale variations. Moreover, the total cross sections have been calculated and are presented in Table 5.14.

Table 5.15 displays the predicted ratios for  $W^+$  to  $W^-$  and  $W^\pm$  to  $Z$  boson production. The ratio of  $W^+$  to  $W^-$  cross sections is mainly sensitive to the differences between the  $u_v$  and  $d_v$  valence-quark distributions at low Bjorken- $x$ . Conversely, the ratio of  $W^\pm$  to  $Z$  cross sections constrains the strange-quark distribution.

TABLE 5.12 – Fiducial cuts for single boson production processes are defined as follows : The lepton kinematics are computed at the Born level, accounting for QED final state radiation. The  $E_T^{\text{miss}}$  is calculated using the neutrinos produced from  $W$ -boson decay.

	$pp \rightarrow \ell^+ \ell^-$	$pp \rightarrow \ell^- \bar{\nu}_\ell / pp \rightarrow \ell^+ \nu_\ell$
Lepton $p_T$ cuts	$p_T > 27 \text{ GeV}$	$p_T > 27 \text{ GeV}$
Lepton $\eta$ cuts	$ \eta  < 2.5$	$ \eta  < 2.5$
Mass cuts	$66 < m_{\ell\ell} < 116 \text{ GeV}$	$m_T^W > 50 \text{ GeV}$
Neutrino cuts	–	$E_T^{\text{miss}} > 25 \text{ GeV}$

TABLE 5.13 – The fiducial cross-section results for single boson production across different PDF sets are provided. The first error listed represents the statistical uncertainty, the second reflects the scale error, and the third denotes the PDF error.

PDF set	$pp \rightarrow \ell^+ \ell^-$ [pb]	$pp \rightarrow \ell^- \bar{\nu}_\ell$ [pb]	$pp \rightarrow \ell^+ \nu_\ell$ [pb]
CT18	733.16 <sup>+0.1%,+0.5%,+5.9%</sup> -0.1%, -0.5%, -6.2%	3372.54 <sup>+0.1%,+1.0%,+1.7%</sup> -0.1%, -0.9%, -3.1%	4385.38 <sup>+0.1%,+0.9%,+1.8%</sup> -0.1%, -0.8%, -3.0%
CT18A	750.52 <sup>+0.1%,+0.5%,+3.7%</sup> -0.1%, -0.5%, -4.3%	3398.85 <sup>+0.1%,+0.9%,+2.4%</sup> -0.1%, -0.9%, -2.5%	4364.99 <sup>+0.1%,+0.9%,+2.3%</sup> -0.1%, -0.8%, -2.4%
MSHT20	747.52 <sup>+0.1%,+0.5%,+2.2%</sup> -0.1%, -0.7%, -2.7%	3360.94 <sup>+0.1%,+0.9%,+1.3%</sup> -0.1%, -1.0%, -1.6%	4317.57 <sup>+0.1%,+0.8%,+1.4%</sup> -0.1%, -0.8%, -1.6%
NNPDF4.0	767.39 <sup>+0.1%,+0.3%,+0.9%</sup> -0.1%, -0.5%, -0.9%	3463.40 <sup>+0.1%,+0.9%,+0.6%</sup> -0.1%, -0.8%, -0.6%	4455.24 <sup>+0.1%,+0.8%,+0.5%</sup> -0.1%, -0.9%, -0.5%
PDF4LHC21	746.14 <sup>+0.1%,+0.4%,+2.8%</sup> -0.1%, -0.6%, -2.8%	3387.04 <sup>+0.1%,+1.0%,+1.5%</sup> -0.1%, -1.1%, -1.5%	4345.91 <sup>+0.1%,+0.9%,+1.5%</sup> -0.1%, -1.0%, -1.5%
ATLASpdf21	787.24 <sup>+0.1%,+0.4%,+3.2%</sup> -0.1%, -0.6%, -4.1%	3545.58 <sup>+0.1%,+1.1%,+2.5%</sup> -0.1%, -1.0%, -2.9%	4579.11 <sup>+0.1%,+1.0%,+2.6%</sup> -0.1%, -1.0%, -3.0%
ABMP16	746.11 <sup>+0.1%,+0.6%,+1.5%</sup> -0.1%, -0.6%, -1.5%	3383.72 <sup>+0.1%,+1.0%,+0.9%</sup> -0.1%, -0.9%, -1.0%	4332.77 <sup>+0.1%,+0.9%,+0.6%</sup> -0.1%, -0.9%, -0.6%

The cross sections are also computed with NNLO QCD + NLO EW accuracy using the MATRIX [143] code. The results show agreement with the MATRIX calculations at the per-mille level, providing additional validation for the predictions.

In addition to the  $W$  and  $Z$  production cross sections, the  $t\bar{t}$  production cross section is also obtained to derive the ratio of  $t\bar{t}$  and  $W$  boson cross sections. The cross section is calculated at NNLO QCD accuracy and NLO in EW coupling using the Top++[2.0] [124-130]. The dependence of the  $t\bar{t}$  cross sections on the top quark mass is evaluated by varying the top mass and calculating the cross sections using the PDF4LHC21 set. The resulting cross sections are presented in Table 5.16.

## 5.7 Systematic uncertainties

This section details the systematic uncertainties considered in this measurement. They are composed of the experimental uncertainties and various modelling uncertainties. The experimental uncertainties are described in Section 5.7.1 and summarized in Table 5.17. The modelling uncertainties arising from the inclusive cross-section calculations are taken into account, which are presented in Section 5.7.2. These un-

TABLE 5.14 – The total cross-section results for single boson production across different PDF sets are provided. The first error listed represents the statistical uncertainty, the second reflects the scale error, and the third denotes the PDF error.

PDF set	$pp \rightarrow \ell^+ \ell^-$ [nb]	$pp \rightarrow \ell^- \bar{\nu}_\ell$ [nb]	$pp \rightarrow \ell^+ \nu_\ell$ [nb]
CT18	1.984 <sup>+0.04%,+0.5%,+5.9%</sup> -0.04%, -0.5%, -6.2%	8.922 <sup>+0.05%,+1.0%,+1.7%</sup> -0.05%, -0.9%, -3.1%	12.055 <sup>+0.05%,+0.9%,+1.8%</sup> -0.05%, -0.8%, -3.0%
CT18A	2.010 <sup>+0.04%,+0.5%,+3.7%</sup> -0.04%, -0.5%, -4.3%	8.962 <sup>+0.05%,+0.9%,+2.4%</sup> -0.05%, -0.9%, -2.5%	12.087 <sup>+0.05%,+0.9%,+2.3%</sup> -0.05%, -0.8%, -2.4%
MSHT20	1.999 <sup>+0.04%,+0.5%,+2.2%</sup> -0.04%, -0.7%, -2.7%	8.866 <sup>+0.05%,+0.9%,+1.3%</sup> -0.05%, -1.0%, -1.6%	11.948 <sup>+0.05%,+0.8%,+1.4%</sup> -0.05%, -0.8%, -1.6%
NNPDF4.0	2.034 <sup>+0.04%,+0.3%,+0.9%</sup> -0.04%, -0.5%, +0.9%	9.069 <sup>+0.05%,+0.9%,+0.6%</sup> -0.05%, -0.8%, -0.6%	12.194 <sup>+0.05%,+0.8%,+0.5%</sup> -0.05%, -0.9%, -0.5%
PDF4LHC21	2.000 <sup>+0.04%,+0.4%,+4.6%</sup> -0.04%, -0.6%, -4.6%	8.911 <sup>+0.05%,+1.0%,+1.5%</sup> -0.05%, -1.1%, -1.5%	12.012 <sup>+0.05%,+0.9%,+1.5%</sup> -0.05%, -1.0%, -1.5%
ATLASpdf21	2.098 <sup>+0.04%,+0.4%,+3.2%</sup> -0.04%, -0.6%, -4.1%	9.334 <sup>+0.05%,+1.1%,+2.5%</sup> -0.05%, -1.0%, -2.9%	12.465 <sup>+0.05%,+1.0%,+2.6%</sup> -0.05%, -1.0%, -3.0%
ABMP16	1.998 <sup>+0.04%,+0.6%,+1.5%</sup> -0.04%, -0.6%, -1.5%	8.847 <sup>+0.05%,+1.0%,+0.9%</sup> -0.05%, -0.9%, -1.0%	11.974 <sup>+0.05%,+0.9%,+0.6%</sup> -0.05%, -0.9%, -0.6%

TABLE 5.15 – Predictions of the ratios of  $W^+$  to  $W^-$  boson and  $W^\pm$  to  $Z$  boson combined production cross sections in the fiducial region based on various PDF sets. The quoted error denote the PDF uncertainty.

PDF set	$W^+/W^-$	$W^\pm/Z$
CT18	1.300 <sup>+0.6%</sup> -0.4%	10.58 <sup>+4.6%</sup> -5.2%
CT18A	1.284 <sup>+0.6%</sup> -0.5%	10.34 <sup>+3.1%</sup> -3.0%
MSHT20	1.285 <sup>+0.7%</sup> -0.7%	10.27 <sup>+2.1%</sup> -1.7%
NNPDF4.0	1.286 <sup>+0.8%</sup> -0.8%	10.32 <sup>+0.7%</sup> -0.7%
PDF4LHC21	1.283 <sup>+0.5%</sup> -0.5%	10.36 <sup>+2.4%</sup> -2.4%
ATLASpdf21	1.291 <sup>+0.5%</sup> -0.4%	10.32 <sup>+2.5%</sup> -2.2%
ABMP16	1.280 <sup>+0.3%</sup> -0.3%	10.34 <sup>+0.9%</sup> -0.9%

TABLE 5.16 – Predictions of the ratios of  $t\bar{t}$  to  $W^+$ ,  $W^-$  and  $W^\pm$  boson production cross sections in the fiducial region based on different PDF sets. The quoted errors are due to statistical fluctuations, QCD scales and PDF+ $\alpha_s$  uncertainties respectively.

PDF set	$t\bar{t}/W^+$	$t\bar{t}/W^-$	$t\bar{t}/W^\pm$
CT18	0.2117 <sup>+0.09%,2.6%,4.6%</sup> -0.09%, 3.7%, 3.4%	0.2753 <sup>+0.11%,2.6%,4.9%</sup> -0.11%, 3.7%, 3.4%	0.1197 <sup>+0.08%,2.6%,4.7%</sup> -0.08%, 3.8%, 3.4%
CT18A	0.2104 <sup>+0.1%,2.6%,4.0%</sup> -0.1%, 3.7%, 4.2%	0.2702 <sup>+0.11%,2.6%,4.1%</sup> -0.11%, 3.7%, 4.3%	0.1183 <sup>+0.08%,2.6%,4.1%</sup> -0.08%, 3.7%, 4.2%
MSHT20	0.2136 <sup>+0.09%,2.6%,3.3%</sup> -0.09%, 3.7%, 2.9%	0.2745 <sup>+0.11%,2.7%,3.4%</sup> -0.11%, 3.7%, 2.6%	0.1201 <sup>+0.08%,2.6%,3.3%</sup> -0.08%, 3.7%, 2.7%
NNPDF4.0	0.2024 <sup>+0.1%,2.6%,1.3%</sup> -0.1%, 3.7%, 2.0%	0.2603 <sup>+0.12%,2.6%,1.3%</sup> -0.12%, 3.7%, 2.2%	0.1139 <sup>+0.09%,2.6%,1.3%</sup> -0.09%, 3.7%, 2.1%
PDF4LHC21 ( $m_t = 171.5$ GeV)	0.2184 <sup>+0.09%,2.7%,2.8%</sup> -0.09%, 3.7%, 2.8%	0.2802 <sup>+0.11%,2.7%,2.9%</sup> -0.11%, 3.7%, 2.9%	0.1227 <sup>+0.08%,2.7%,2.9%</sup> -0.08%, 3.7%, 2.9%
PDF4LHC21 ( $m_t = 172.5$ GeV)	0.2125 <sup>+0.09%,2.6%,2.8%</sup> -0.09%, 3.7%, 2.8%	0.2727 <sup>+0.11%,2.7%,2.9%</sup> -0.11%, 3.7%, 2.9%	0.1194 <sup>+0.08%,2.7%,2.8%</sup> -0.08%, 3.8%, 2.8%
PDF4LHC21 ( $m_t = 173.5$ GeV)	0.2069 <sup>+0.1%,2.7%,2.8%</sup> -0.1%, 3.7%, 2.8%	0.2654 <sup>+0.11%,2.7%,2.9%</sup> -0.11%, 3.8%, 2.9%	0.1163 <sup>+0.09%,2.7%,2.8%</sup> -0.09%, 3.7%, 2.8%
ATLASpdf21	0.2139 <sup>+0.09%,2.7%,5.1%</sup> -0.09%, 3.7%, 4.6%	0.2762 <sup>+0.11%,2.6%,5.1%</sup> -0.11%, 3.8%, 4.5%	0.1205 <sup>+0.08%,2.7%,5.1%</sup> -0.08%, 3.7%, 4.5%
ABMP16	0.1941 <sup>+0.1%,2.6%,3.0%</sup> -0.1%, 3.7%, 3.0%	0.2485 <sup>+0.12%,2.6%,3.0%</sup> -0.12%, 3.7%, 3.0%	0.109 <sup>+0.09%,2.6%,2.9%</sup> -0.09%, 3.8%, 2.9%

certainties are incorporated into the likelihood fits as nuisance parameters detailed in the Section 5.9, and a correlation model is provided to treat the correlations between nuisance parameters in Section 5.7.3. The dominant uncertainties are summarized in Section 5.7.4.

### 5.7.1 Experimental uncertainties

The pre-fit experimental uncertainties are presented in this section and summarized in Table 5.17.

**Luminosity :** The integrated luminosity has an uncertainty of 2.2%, as determined by the LUCID-2 detector [144], which is employed for the primary luminosity measurements.

**Pile-up reweighting :** The pile-up uncertainty is estimated by shifting the pile-up condition and taking the differences between the results obtained using nominal and shifted conditions. In this measurement, the average number of interactions per bunch-crossing is varied by 4% in the simulation to derive the uncertainty.

**Lepton trigger efficiency :** The lepton trigger uncertainties are obtained by comparing the trigger efficiencies of data and simulations with a set of independent variations of systematic sources. They are measured using the Run 3 data and corresponding simulations.

**Electron correction :** The uncertainties related to electron reconstruction, identification, and isolation efficiency corrections are considered, which are detailed in [145]. The electron corrections are adapted from Run 2 dataset and an additional uncertainty is added to account for the differences between Run 2 and Run 3 datasets.

**Electron calibration :** The current electron calibration is derived using MC samples and Run 2 data collected in 2018, based on the methodology described in Ref. [75]. The differences between Run 2 and Run 3 datasets are considered as additional uncertainties.

**Muon correction :** The uncertainties in muon efficiency corrections, including the muon identification, isolation, and TTVA efficiency corrections are considered. They are derived using Run 3 dataset and MC samples, which involve independently varying the components of the inner detector and the muon spectrometer, based on the methodology described in Ref. [80].

**Muon calibration :** The muon calibration uncertainty is measured using Run 3 dataset, using the methodology detailed in Ref. [80].

**Jet energy scale and resolution :** The systematic uncertainties of jet energy scale and resolution affect the calculation of the missing transverse momentum, which change the acceptance of the  $W^\pm$  channels. The uncertainties are measured using Run 2 samples [86] with a dedicated component addressing the difference between Run 2 and Run 3 samples, as described in [105].

**Jet vertex tagger :** The systematic uncertainties arising from NNJVT are addressed by assigning a conservative 10% uncertainty per jet to account for the difference of NNJVT efficiencies between data and MC samples. The details are described in [86].

$E_T^{\text{miss}}$  : The missing transverse momentum uncertainties arising from track soft term are considered, as detailed in Ref. [146].

**Multijet modelling :** The contribution of multijet modeling uncertainty is negligible in the  $Z$  channels, but it affects the  $W$  channels. The uncertainties are detailed in Section 5.5.

The uncertainties related to jet,  $E_T^{\text{miss}}$  and multijet are only considered when measuring the  $W$  production cross sections and related ratios.

TABLE 5.17 – Summary of the relative systematic uncertainties. The uncertainties in the table are the pre-fit uncertainties. The modelling uncertainties of electroweak and top background are described in Section 5.7.2.

$\delta\sigma/\sigma[\%]$	$W^- \rightarrow e^- \bar{\nu}$	$W^+ \rightarrow e^+ \nu$	$Z \rightarrow e^+ e^-$	$W^- \rightarrow \mu^- \bar{\nu}$	$W^+ \rightarrow \mu^+ \nu$	$Z \rightarrow \mu^+ \mu^-$
Pile-up modelling	1.1	1.2	1.2	0.4	0.4	0.3
Electron trigger efficiency	1.2	1.2	0.2	-	-	-
Electron reconstruction efficiency	0.3	0.3	0.6	-	-	-
Electron identification efficiency	0.2	0.2	0.5	-	-	-
Electron isolation efficiency	0.6	0.6	1.2	-	-	-
Electron calibration	0.1	0.1	0.1	-	-	-
Muon trigger efficiency	-	-	-	1.1	1.1	0.4
Muon identification efficiency	-	-	-	0.2	0.2	0.4
Muon isolation efficiency	-	-	-	1.0	1.0	2.0
Muon TTVA efficiency	-	-	-	<0.1	<0.1	0.1
Muon calibration	-	-	-	0.2	0.2	0.1
Jet energy scale	1.7	1.6	-	1.6	1.6	-
Jet energy resolution	0.4	0.4	-	0.4	0.5	-
$E_T^{\text{miss}}$ soft term scale	<0.1	<0.1	-	<0.1	<0.1	-
$E_T^{\text{miss}}$ soft term resolution	0.2	0.2	-	0.2	0.3	-
Electroweak+top background	0.2	0.2	0.2	0.2	0.1	0.2
Multijet background	2.9	2.4	-	1.3	1.1	-
JVT	1.6	1.5	-	1.4	1.3	-
Luminosity	2.2	2.2	2.2	2.2	2.2	2.2

## 5.7.2 Modelling uncertainties

The theoretical uncertainties are considered in terms of the background and signal processes in this section.

### Modelling uncertainties for the background processes

The electroweak and top backgrounds are modeled using the MC simulations. The uncertainties of their cross sections are evaluated conservatively and the contribution of each process is at percent level or smaller. For the  $W$  and  $Z$  productions, the total uncertainty of 5% is assigned, arising from the PDF input,  $\alpha_s$ , and QCD scale variations. The theoretical uncertainty for diboson processes is set at 10%. The modeling uncertainties for  $t\bar{t}$  and single-top processes are 5.1% and 3.5%, respectively, as reported in Ref. [105].

## Modelling uncertainties for the $W$ and $Z$ signals

For the  $W$  and  $Z$  signal samples, the theoretical uncertainties are divided into two main components : one relates to uncertainties associated with the *acceptance*  $A$ , which is the ratio of events within the fiducial volume to those in the total phase space at the particle level. The other relates to the *correction factor*  $C$ , which is the ratio of expected signal events at the detector level to those at the particle level within the fiducial volume. The fiducial volume is defined in Table 5.12.

The overall acceptance  $A$  and correction factor  $C$  are determined using nominal SHERPA signal samples, using the following formula :

$$A = \frac{\sum_{\text{events in fiducial volume}}(\text{weight})}{\sum_{\text{events in total phase space}}(\text{weight})},$$

$$C = \frac{\sum_{\text{events passing event selections}}(\text{weight} \times \text{experimental corrections})}{\sum_{\text{events in fiducial volume}}(\text{weight})},$$

where the weight is defined as the generator weight normalized by the cross section of the sample, while the experimental corrections account for discrepancies between data and simulations. These two components are assessed independently and treated as distinct sources of uncertainty.

The modelling uncertainties arises from various sources :

**QCD scale (ME and parton showers) :** The QCD scale uncertainty is assessed through a seven-point variation of the renormalisation and factorisation scales described in Section 5.6.4. Both variations of the matrix element alone and combined variations of the matrix element and the parton shower scale are evaluated. The final uncertainty is determined by taking the maximum variation from these  $2 \times 7$  sets.

**PDF set :** The PDF uncertainty is estimated by analyzing the internal variations of the PDF4LHC21. The 40 Hessian sets are treated as independent eigenvectors and included in the fit as separate nuisance parameters.

$\alpha_s$  :

The  $\alpha_s$  uncertainty is estimated by considering two different  $\alpha_s$  values in PDF4LHC21, as described in Section 5.6.4.

**PDF choice :** The difference in cross sections obtained with the two PDFs, NNPDF 3.0 NNLO and PDF4LHC21, is regraded as an additional uncertainty.

The theoretical uncertainties in  $A$  and  $C$  are calculated separately. The theoretical uncertainties in the correction factor  $C$  are incorporated into the fit as several independent nuisance parameters, each with a corresponding overall scale, and then propagated to the measured fiducial cross sections. Conversely, the theoretical uncertainties in the acceptance  $A$  are only propagated to the final measured total cross sections. These two components are combined as independent sources of uncertainty in the total cross section calculation.

TABLE 5.18 – The impact of the QCD scale uncertainty in the correction factors of the  $W$  and  $Z$  channels.

$\delta C/C[\%]$	$W^- \rightarrow e^- \bar{\nu}$	$W^+ \rightarrow e^+ \nu$	$Z \rightarrow e^+ e^-$	$W^- \rightarrow \mu^- \bar{\nu}$	$W^+ \rightarrow \mu^+ \nu$	$Z \rightarrow \mu^+ \mu^-$
QCD scale	1.36	1.05	0.59	0.59	0.53	0.13

The theoretical uncertainties are minor sources for the correction factor  $C$ . The uncertainty arising from QCD scale and parton shower variations are summarized in Table 5.18 while the contribution of the other theoretical uncertainties (PDF set,  $\alpha_s$  and PDF choice) is  $\sim 0.2\%$  in  $C_Z$  and  $\sim 0.4\%$  for  $C_W$ . For the acceptance  $A$ , the theoretical uncertainty is  $\sim 2.9\%$  for  $A_Z$  and  $\sim 2.4\%$  for  $A_W$ .

### 5.7.3 Systematic uncertainty correlation in the combination

Simultaneous fits, as described in Section 5.9, for both electron and muon channels yield the combined results and the ratios of fiducial cross sections ( $W^+/W^-$ ,  $W^\pm/Z$ , and  $t\bar{t}/W$ ).

A correlation model for the uncertainties is defined for the combination. Lepton corrections, trigger efficiency, and energy calibration uncertainties are treated as uncorrelated between the electron and muon channels, and the multijet uncertainty is also considered uncorrelated. Other experimental uncertainties (including those related to jets,  $E_T^{\text{miss}}$ , pile-up, and luminosity), as outlined in Table 5.17, are considered fully correlated.

The modeling uncertainties for electroweak and top processes are treated as fully correlated across channels to be fitted. When combining electron and muon channels of the same boson, the uncertainties in signal modeling are fully correlated. However, when combining channels of different bosons, these uncertainties are treated as uncorrelated due to the different production mechanisms at the LHC.

### 5.7.4 Dominant uncertainty sources

The dominant sources of uncertainty varies depending on the physical observables measured in this analysis.

For the  $Z$  cross section, the primary uncertainties are the uncertainties arising from luminosity and lepton corrections. While for the  $W$  cross sections, multijet and jet-related uncertainties are the most significant.

As for the cross-section ratios, many uncertainties, especially experimental ones, mostly cancel out. For the  $W^+/W^-$  ratios, the primary experimental uncertainties such as the luminosity, lepton corrections, and jet corrections largely cancel out. In this case, the multijet uncertainties become the dominant sources since they are treated as independent for each  $W$  decay channel. Similarly, for the ratio of  $W^\pm$  to  $Z$  cross sections, the jet and multijet uncertainties affect only the  $W$  cross section and not the  $Z$  cross section, making them the dominant sources of uncertainty. Regarding the ratio of the  $t\bar{t}$  and  $W$  fiducial cross sections, the uncertainties related to the  $t\bar{t}$  modeling and multijet background are significant because they only contribute to the  $t\bar{t}$  and  $W$

TABLE 5.19 – Event yields of data and predictions after the selections. Statistical uncertainties are shown for the electroweak and top processes. For the multijet process, the normalization uncertainty is displayed instead.

	$W^- \rightarrow e^- \bar{\nu}$	$W^+ \rightarrow e^+ \nu$	$W^- \rightarrow \mu^- \bar{\nu}$	$W^+ \rightarrow \mu^+ \nu$
$W \rightarrow e\nu$	$43,650,000 \pm 70,000$	$55,370,000 \pm 80,000$	–	–
$W \rightarrow \mu\nu$	–	–	$57,760,000 \pm 80,000$	$74,900,000 \pm 90,000$
$W \rightarrow \tau\nu$	$684,000 \pm 8,000$	$819,000 \pm 9,000$	$906,000 \pm 10,000$	$1,120,000 \pm 10,000$
$Z \rightarrow ee/\mu\mu$	$1,416,000 \pm 4,000$	$1,459,000 \pm 4,000$	$4,638,000 \pm 8,000$	$4,903,000 \pm 8,000$
$Z \rightarrow \tau\tau$	$88,000 \pm 1,000$	$91,000 \pm 1,000$	$107,000 \pm 1,000$	$111,000 \pm 1,000$
$t\bar{t}$ +single-top	$863,800 \pm 400$	$905,400 \pm 500$	$802,200 \pm 400$	$843,400 \pm 400$
$VV$	$93,500 \pm 300$	$97,600 \pm 300$	$98,500 \pm 300$	$102,800 \pm 300$
Multi-jet	$4,000,000 \pm 1,000,000$	$4,000,000 \pm 1,000,000$	$2,100,000 \pm 700,000$	$2,200,000 \pm 800,000$
Total predicted	$51,000,000 \pm 1,000,000$	$63,000,000 \pm 1,000,000$	$66,400,000 \pm 700,000$	$84,200,000 \pm 800,000$
Data	50,748,537	62,610,338	65,053,470	82,360,980

	$Z \rightarrow e^+ e^-$	$Z \rightarrow \mu^+ \mu^-$
$Z \rightarrow ee$	$7,710,000 \pm 10,000$	–
$Z \rightarrow \mu\mu$	–	$13,961,000 \pm 10,000$
$Z \rightarrow \tau\tau$	$2,500 \pm 200$	$3,900 \pm 200$
$W \rightarrow \ell\nu$	$139 \pm 57$	$40 \pm 27$
$t\bar{t}$ +single-top	$24,950 \pm 50$	$37,000 \pm 50$
$VV$	$14,960 \pm 80$	$23,800 \pm 100$
Total predicted	$7,750,000 \pm 10,000$	$14,030,000 \pm 10,000$
Data	7,812,978	14,242,875

cross sections respectively and can not cancel out. Additionally, uncertainties from the jet and lepton trigger efficiency, as well as the background modeling (excluding multijet), also do not fully cancel out and therefore become important in the  $t\bar{t}/W$  ratios.

## 5.8 Kinematic distributions and event yields in the signal region

After have implemented the selections and estimated the backgrounds, the kinematic distributions for  $W$  and  $Z$  events that meet the signal event criteria are presented in Figures 5.9- 5.10, with the systematic uncertainties described in Section 5.7. The data and predictions for the primary observables  $m_{\text{T}}$  and  $m_{\ell\ell}$  in both  $W$  and  $Z$  channels are shown. Good agreement is observed, without any significant discrepancies beyond the quoted uncertainties. Table 5.19 provides a detailed summary of event yields for each process, which are used as inputs for the likelihood fits to determine the cross sections and their ratios.

## 5.9 Statistical analysis

This section explains the statistical method employed to derive the cross sections and their ratios. The profile-likelihood function is formulated using the event yields obtained after applying event selections, with uncertainties included in the model as nuisance parameters. These uncertainties account for various sources of systematic and statistical variations. Section 5.9.1 describes the profile-likelihood formula, and the tests using MC predictions are performed to evaluate the performance of the statistical tools in Section 5.9.2.

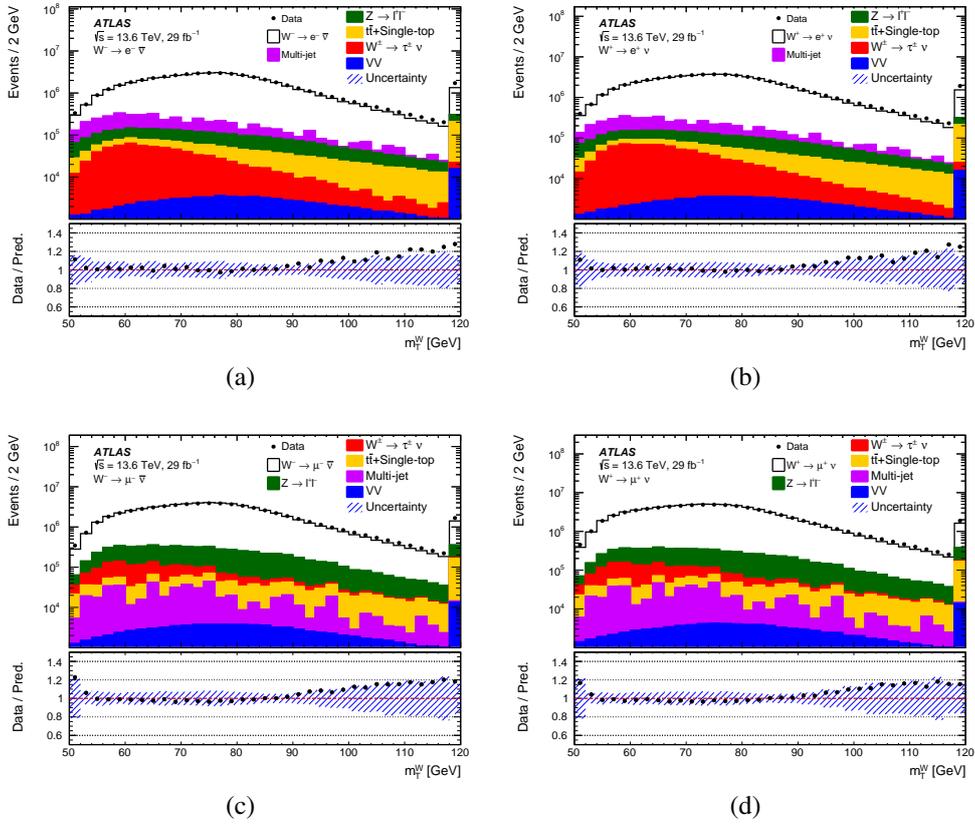


FIGURE 5.9 – A comparison of data (dots) and predictions (histograms) for the  $m_T^W$  distributions is illustrated in the following channels : (a)  $W^- \rightarrow e^- \bar{\nu}$ , (b)  $W^+ \rightarrow e^+ \nu$ , (c)  $W^- \rightarrow \mu^- \bar{\nu}$ , and (d)  $W^+ \rightarrow \mu^+ \nu$ . The hashed band in the ratio plot represents the total systematic uncertainty in the prediction. Additionally, the rightmost bins include the overflow events, ensuring that all data points are accounted for in the analysis.

### 5.9.1 Statistical model

The profile-likelihood technique is employed to extract the inclusive  $W$  and  $Z$  production fiducial cross sections and their ratios. The profile-likelihood function is meticulously constructed using the event yields obtained following the application of event selections. The uncertainties are systematically incorporated into the model as nuisance parameters. The uncertainties account for various sources of systematic and statistical variations, ensuring that the extracted cross sections and ratios are reliable. The methodology involves integrating these nuisance parameters into the likelihood function, which is particularly useful for dealing with these uncertainties and extracting the most information from the data, even in the presence of complex models.

The data event yields can be related the signal cross sections via the following

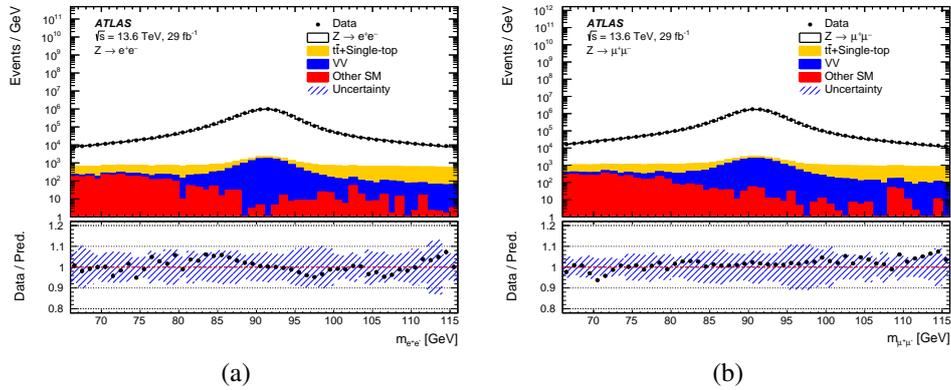


FIGURE 5.10 – A comparison of data (dots) and predictions (histograms) for the  $m_{\ell\ell}$  distributions is illustrated in the following channels : (a)  $Z \rightarrow e^+e^-$  and (b)  $Z \rightarrow \mu^+\mu^-$ . The hashed band in the ratio plot represents the total systematic uncertainty in the prediction.

relationship :

$$N_{\text{data}} = \mathcal{L}\sigma_{\text{sig}}\varepsilon + N_{\text{bkg}} \quad (5.13)$$

where  $\mathcal{L}$  denotes the luminosity,  $\sigma_{\text{sig}}$  presents the signal cross sections, and  $\varepsilon$  is the efficiency for signal events survive after applying the event selections. In the further step, the statistical model is constructed in a concrete form, as shown by the following formula :

$$L(\vec{n}; \mu_s, \vec{\theta}) = \prod_{c \in \text{channels}} \text{Pois}(n_{\text{data}} | \mu_{s,c} S_c(\vec{\theta}) + B_c(\vec{\theta})) \prod_{i \in \text{NPs}} G(\theta_i), \quad (5.14)$$

in which  $\mu_s$  is the signal strength, denoting the ratio of the measured signal cross section to the predicted value.  $S_c$  is the expected number of signal events, and  $B_c$  is the expected background. The normalization of the background contributions is determined by the fit and constrained by the background cross-section uncertainties. The quantity  $\theta$  presents the nuisance parameters, which are constrained by a Gaussian term  $G(\theta_i)$ . This term is determined by an auxiliary measurement. The uncertainties described in Section 5.7 are incorporated in the model as nuisance parameters. The uncertainty arising from the finite number of simulated events of MC samples is considered and included into the likelihood formula via the light weight version of the Beeston-Barlo method [132, 133]. The probability model is constructed as the product of the Poisson distributions (“Pois”) in each channel. This approach ensures that the statistical fluctuations in the number of observed events are accurately modeled, taking into account the expected number of signal and background events in each channel. The theoretical uncertainties of the  $W$  and  $Z$  signal processes are incorporated into the fitting by introducing normalization factors for the signal samples. These normalization factors are calculated by comparing the correction

factors of the nominal and theoretical variations and taking the relative difference. The profile-likelihood is implemented in the `TRExFitter` framework, based on the `Histfactory` [133]. The likelihood function is maximised to determine the best-fit values for all parameters along with their associated uncertainties. The fitting procedure is implemented based on `xRooFit` framework. The uncertainty is derived using MINOS algorithm [147], which allows precise estimation of the uncertainties and the resulting uncertainties are asymmetric. Additionally, the likelihood scans are performed to ensure the robustness and reliability of the uncertainty estimation, as instability in uncertainty estimation using the Hessian algorithm has been observed with large statistics in this analysis.

The ratios of fiducial cross sections are also extracted by maximizing the likelihood functions. However, one needs to apply some modifications to the likelihood formula to extract the cross section ratios. If one wants to extract the ratio  $R_{W^+/W^-}$ , which is the ratio of  $W^+$  to  $W^-$  cross sections, the signal strength of  $W^+$  ( $\mu_{W^+}$ ), can be redefined as  $R_{W^+/W^-}$ . Therefore, the modified likelihood Eq. (5.14) can be formulated as :

$$L(\vec{n}; \mu_s, \vec{\theta}) = \prod_{c \in W^+ \text{ channels}} \text{Pois}(n_{\text{data}} | R_{W^+/W^-} \mu_{W^-} S_c + B_c(\vec{\theta})) \prod_{c \in W^- \text{ channels}} \text{Pois}(n_{\text{data}} | \mu_{W^-} S_c + B_c(\vec{\theta})) \prod_{i \in \text{NPs}} G(\theta_i).$$

Other cross section ratios are extracted in the same manner.

Besides the  $W^\pm$  and  $Z$  results, the study also examines the fiducial cross-section ratio of  $t\bar{t}/W$ . The inputs and uncertainties for  $t\bar{t}$  are sourced from the  $t\bar{t}$  cross-section analysis conducted at 13.6 TeV [105]. The  $t\bar{t}$  production cross section is determined using the well-established  $b$ -tag counting method [148], which takes the advantage of low sensitivity to systematic uncertainties. The event yields of  $e\mu + 1b$ -jets (denoted as  $N_1$ ) and  $e\mu + 2b$ -jets (denoted as  $N_2$ ) channels can be expressed as :

$$\begin{aligned} N_1 &= L \sigma_{t\bar{t}} \epsilon_{e\mu} 2\epsilon_b (1 - C_b \epsilon_b) + N_1^{\text{bkg}}, \\ N_2 &= L \sigma_{t\bar{t}} \epsilon_{e\mu} C_b \epsilon_b^2 + N_2^{\text{bkg}}, \end{aligned}$$

where  $\epsilon_b$  is the  $b$ -tagging efficiency and  $C_b$  is defined as  $\epsilon_{bb}/\epsilon_b$ , with  $\epsilon_{bb}$  standing for the probability to identify both of the  $b$ -jets. The quantity  $C_b$  can be estimated using MC samples via :

$$C_b = \frac{4N_{\geq 0}^{t\bar{t}} N_2^{t\bar{t}}}{(N_1^{t\bar{t}} + 2N_2^{t\bar{t}})^2},$$

where  $N_{\geq 0}^{t\bar{t}}$  denotes the  $t\bar{t}$  event yields passing event selections in  $e\mu$  channels but without any requirements on the  $b$ -jets. The  $t\bar{t}$  cross sections and  $\epsilon_b$  are extracted simultaneously from the likelihood fits.

### 5.9.2 Asimov fits

The performance of the profile-likelihood fits is evaluated by performing fits to the Asimov dataset. An Asimov dataset [149] is a hypothetical dataset in which

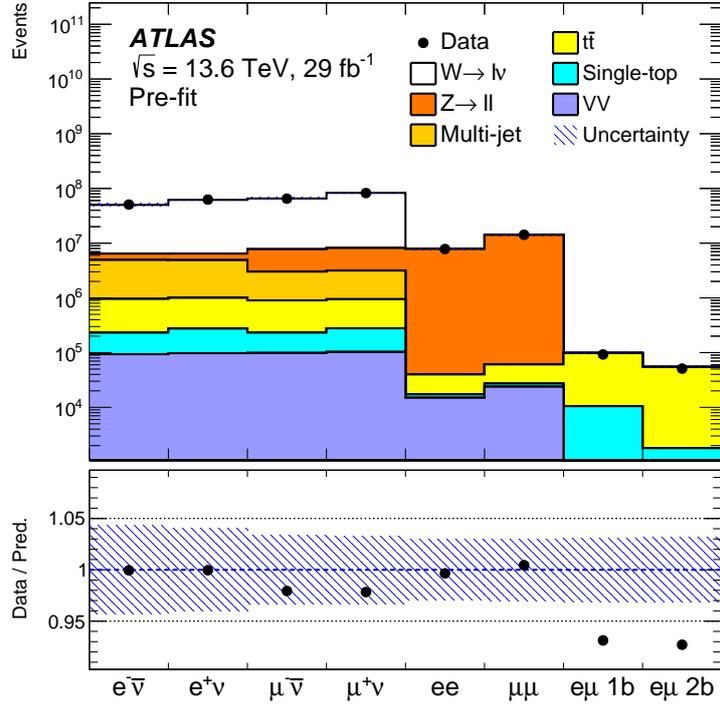


FIGURE 5.11 – Comparison of data event yields in each channel (dots) with the predictions (stacked histograms) is shown before the fits. The dashed error band in the pre-fit figure represents the total systematic uncertainty.

the observed data precisely matches the expected values predicted by the model. This dataset is utilized to evaluate the expected performance of a statistical method. Asimov datasets are constructed as binned datasets, where the event count in each bin is set to the expected event yield for the selected model parameters. In this measurement, a single bin is used per channel.

The inputs are varied depending on the specific cross section to be extracted. To derive the signal strength of the  $Z$  cross section ( $\sigma(Z \rightarrow ee)$ ,  $\sigma(Z \rightarrow \mu\mu)$ , and  $\sigma(Z \rightarrow \ell\ell)$ ), the two same-flavor di-lepton channels ( $ee$  and  $\mu\mu$ ) are used. When fitting the  $W$  cross sections, as well as the  $\sigma(W^+)/\sigma(W^-)$  and  $\sigma(W^\pm)/\sigma(Z)$  ratios, the four single-lepton regions ( $e^-\bar{\nu}$ ,  $e^+\nu$ ,  $\mu^-\bar{\nu}$ , and  $\mu^+\nu$ ) along with the two same-flavor di-lepton regions are utilized as inputs. The combined fits ensure the consistency between cross sections and their ratios. Additionally, to derive the  $t\bar{t}/W$  ratio, two  $e\mu$  regions from selected  $t\bar{t}$  events (split according to the number of  $b$ -jets) are incorporated into the fits.

The dataset used for fitting is categorized by channels and displayed in Figure 5.11, where the two  $t\bar{t} e\mu$  channels are obtained from Ref. [105].

The nuisance parameters represent the sources of uncertainty described in Sec-

tion 5.7. For the electron efficiency and calibration uncertainties, the fine schemes are used which are recommended for precision analyses. Systematic uncertainties below the pruning threshold of 0.01% are excluded. After this systematic pruning, the following numbers of systematic uncertainty sources have been excluded : 104 from the  $ee$  channel, 121 from the  $e^-\bar{\nu}$  channel, 135 from the  $e^+\nu$  channel, 37 from the  $\mu\mu$  channel, 52 from the  $\mu^-\bar{\nu}$  channel, 52 from the  $\mu^+\nu$  channel, 1613 from the  $e\mu$  channel with 1  $b$ -jet, and 1327 from the  $e\mu$  channel with 2  $b$ -jets.

The performance of all likelihood fits are checked using Asimov datasets. The fits to Asimov dataset for  $Z \rightarrow \ell\ell$ ,  $W \rightarrow \ell\nu$ ,  $W^+/W^-$ ,  $W^\pm/Z$  and  $t\bar{t}/W$  are presented in this section.

Nuisance parameter (NP) pulls are checked in fits to Asimov dataset, which represent the deviations of the fit parameters from their nominal values, providing insight into how much each NP had to be adjusted to achieve the best fit. By examining the NP pulls, one can evaluate the goodness of a fit. If a fit is good, the NP pulls should be around zero, indicating that the fit is not systematically biased. If large NP deviations are observed, it may indicate issues like outliers, underestimated uncertainties, or an inadequate fit model. In the fits to Asimov dataset, the expected central values should be at zero.

The impact of nuisance parameters on the signal strength  $\mu$  is evaluated for all fits. For a given nuisance parameter  $\theta$ , if the best-fit value of  $\theta$  is  $\hat{\theta}$  and its uncertainty is  $\Delta\theta$ , then the impact can be determined by fixing the value of  $\theta$  at  $\hat{\theta} \pm \Delta\theta$ , and repeat the fit. The difference between the resulting and nominal signal strength  $\Delta\mu$  reflects the impact of  $\theta$ . The nuisance parameters are ranked in the order of the impact on the signal strengths in both pre-fit and post-fit cases in the following sections. This ranking shows which systematic uncertainties have the most significant effect on the measurement of the signal strength.

### 5.9.3 Fit to Asimov data for $Z \rightarrow \ell\ell$

This section describes the fitting results to Asimov datasets for  $Z \rightarrow \ell\ell$  in the electron and muon channels. The cross sections are obtained using the  $ee$  and  $\mu\mu$  channels as input, respectively. The combined results are obtained by fitting the  $ee$  and  $\mu\mu$  channels simultaneously. Therefore, three fits are performed. The resulting nuisance pulls plots for  $Z$ -boson channels are shown in Figures 5.12. The systematic uncertainties are ranked by the impact on the signal strength  $\mu$ . The impacts are shown, before and after the fits to the Asimov dataset, in Figure 5.13.

### 5.9.4 Fit to Asimov data for $W^\pm \rightarrow \ell\nu$ , $W^+/W^-$ and $W^\pm/Z$ ratios

This section describes the fits to Asimov data for  $W^\pm \rightarrow \ell^\pm\nu(\bar{\nu})$ ,  $W^+/W^-$  and  $W^\pm/Z$  ratios. All of this quantities are extracted by fitting  $e^+e^-$ ,  $\mu^+\mu^-$ ,  $e^+\nu$ ,  $e^-\bar{\nu}$ ,  $\mu^+\nu$ , and  $\mu^-\bar{\nu}$  channels simultaneously. Several quantities can be obtained with one fit by adjusting the parameters of interest, for example, one can obtain the signal strengths of  $W^+/W^-$ ,  $W^- \rightarrow e^-\bar{\nu}$  and  $Z \rightarrow e^+e^-$  at the same time. Fits are repeated several times with the same inputs but changing the parameters

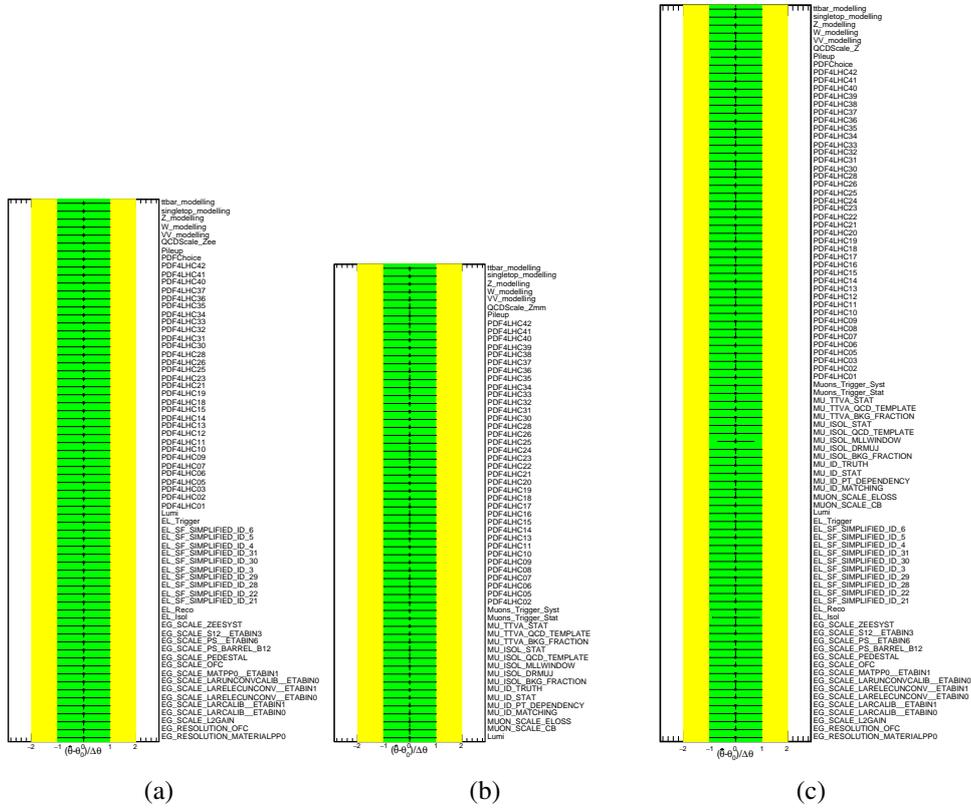


FIGURE 5.12 – Nuisance parameters fitted to the Asimov dataset for (a)  $Z \rightarrow e^+e^-$ , (b)  $Z \rightarrow \mu^+\mu^-$  and (c) the combined  $Z \rightarrow \ell\ell$  cross sections.

of interest each time. Different combinations of parameters of interest are used to obtain the signal strengths of  $W^+ \rightarrow e^+\nu$ ,  $W^+ \rightarrow \mu^+\nu$ ,  $W^- \rightarrow e^-\bar{\nu}$ ,  $W^- \rightarrow \mu^-\bar{\nu}$ ,  $W^- \rightarrow \ell^-\bar{\nu}$ ,  $W^+ \rightarrow \ell^+\nu$ , and  $W^\pm \rightarrow \ell\nu$  cross sections, as well as the  $W^+/W^-$  and  $W^\pm/Z$  cross-section ratios using the same inputs. In addition to these values, the signal strengths of the  $Z$  channels can be determined as well, which are compared to the results obtained from pure  $Z$ -channel fits. Good consistency is found between these fitting results. The nuisance parameter pull plot for the combined fit is shown in Figures 5.14. Figures 5.15-5.17 display the impact of the systematic sources on each signal strength.

### 5.9.5 Fit to Asimov data for $t\bar{t}/W^\pm$ ratios

This subsection describes the fitting results to Asimov datasets for the  $t\bar{t}/W^\pm$  ratios. In addition to  $t\bar{t}/W^\pm$ , the fit to Asimov dataset for  $R_{t\bar{t}/W^+}$  and  $R_{t\bar{t}/W^-}$  ratios are also checked. The pulls plots for the  $W^\pm$ -boson channels are shown in Figure 5.18. The impact of systematic sources ranked by the impact on the signal strength is shown in Figure 5.19.

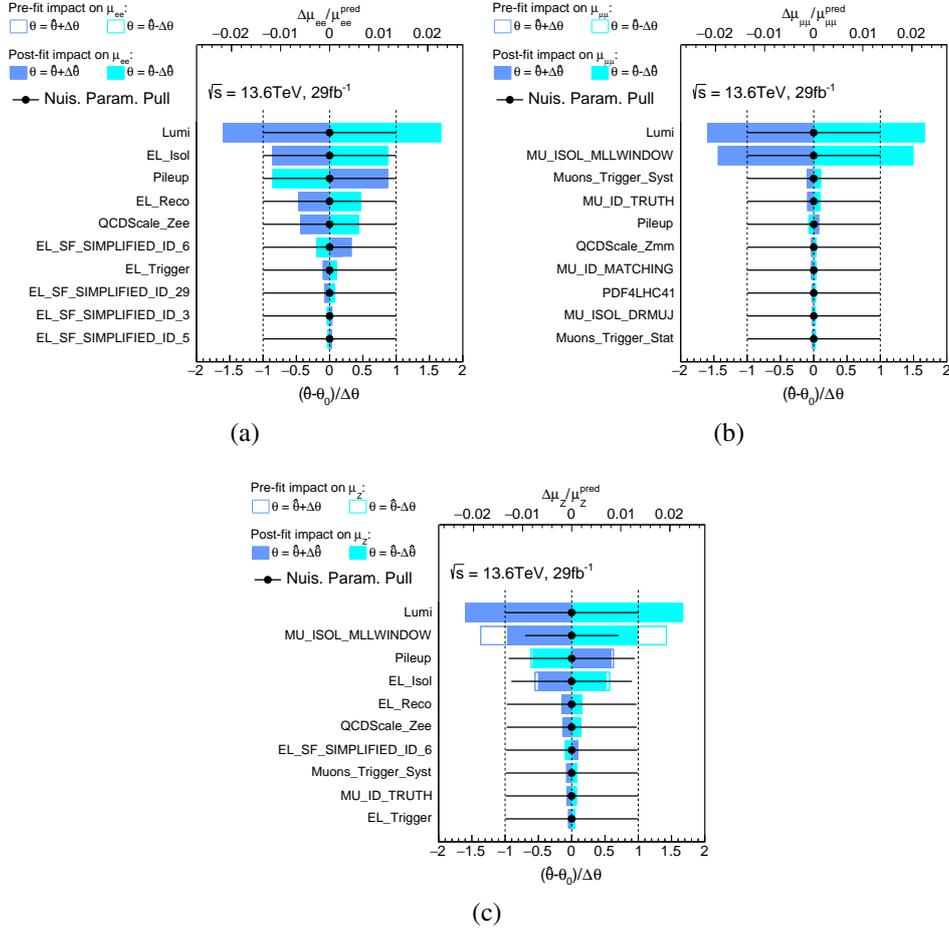


FIGURE 5.13 – Systematic uncertainties ranked by the impact on the signal strength  $\mu$ . The impact is shown before and after the fits to the Asimov dataset for (a)  $Z \rightarrow e^+e^-$ , (b)  $Z \rightarrow \mu^+\mu^-$  and (c) combined  $Z \rightarrow \ell\ell$ . The impact is shown before and after the fits to the Asimov datasets.

## 5.10 Results

### 5.10.1 General procedure

The cross sections  $\sigma_{W^+ \rightarrow \ell^+\nu}$ ,  $\sigma_{W^- \rightarrow \ell^-\bar{\nu}}$ ,  $\sigma_{W^\pm \rightarrow \ell\nu}$ , and  $\sigma_{Z \rightarrow \ell^+\ell^-}$  along with their ratios  $R_{W^\pm/Z}$  and  $R_{W^+/W^-}$  are extracted from profile-likelihood fits to the inclusive data in the four fiducial single-lepton channels  $W^+ \rightarrow e^+\nu$ ,  $W^+ \rightarrow \mu^+\nu$ ,  $W^- \rightarrow e^-\bar{\nu}$ , and  $W^- \rightarrow \mu^-\bar{\nu}$  and the two fiducial di-lepton channels  $Z \rightarrow e^+e^-$  and  $Z \rightarrow \mu^+\mu^-$ . Similarly, combined fits of the four single-lepton channels, the two di-lepton channels, and the two  $t\bar{t}e\mu$  channels from Ref. [105] are used to extract the cross-section ratios  $R_{t\bar{t}/W^\pm}$ ,  $R_{t\bar{t}/W^+}$ , and  $R_{t\bar{t}/W^-}$ .

Fiducial cross sections,  $\sigma^{\text{fid}}$ , are obtained by applying the signal-

strength parameter  $\mu$  to the nominal predicted fiducial cross section, which is derived from SHERPA signal samples. The uncertainties on the signal-strength parameter  $\mu$  are subsequently propagated into the derived fiducial cross sections. Table 5.20 shows the predicted fiducial cross sections used to extract the measured results, which are calculated using the SHERPA signal MC simulations described in Ref. 5.2.2.

TABLE 5.20 – Predicted fiducial cross sections used to extract measured results.

Process	Fiducial cross-section prediction [pb]
$Z \rightarrow e^+e^-$	735.0
$Z \rightarrow \mu^+\mu^-$	735.7
$Z \rightarrow \ell^+\ell^-$	735.3
$W^- \rightarrow e^-\bar{\nu}$	3381.0
$W^- \rightarrow \mu^-\bar{\nu}$	3386.0
$W^- \rightarrow \ell^-\bar{\nu}$	3383.5
$W^+ \rightarrow e^+\nu$	4346.9
$W^+ \rightarrow \mu^+\nu$	4345.6
$W^+ \rightarrow \ell^+\nu$	4346.3
$W^\pm \rightarrow \ell^+\nu$	7729.8

Subsequently, total cross sections,  $\sigma^{\text{tot}}$ , can be derived based on the fiducial cross sections by applying the acceptance correction  $A$  on  $\sigma^{\text{fid}}$ . The acceptance is defined as the fraction of events in the fiducial volume relative to the total number of events in the total phase space at the particle level. The total phase space of  $W$  production is defined without any kinematic cuts while for  $Z$ -boson production, the invariant mass of the lepton pair final states should satisfy  $66 < m_{\ell\ell} < 116$  GeV. The acceptance corrections are calculated using signal SHERPA samples as described in Section 5.7.2.

In the following sections, the results are obtained and compared with the theory predictions computed in Section 5.6. Figure 5.20 displays the comparison between data and predictions both before and after the fits across all regions.

### 5.10.2 Results of $Z$ -boson cross sections

The  $Z$ -boson production cross sections are obtained using pure fits to the di-lepton  $e^+e^-$  and  $\mu^+\mu^-$  channels. Fits are performed separately in the electron and muon channels, as well as a combined fit that includes both channels simultaneously. The signal strength parameters  $\mu_{Z \rightarrow ee}$ ,  $\mu_{Z \rightarrow \mu\mu}$ , and  $\mu_{Z \rightarrow \ell\ell}$ , which represent the ratios between the measured and predicted cross sections, are derived from the fits in the  $e^+e^-$  channel, the  $\mu^+\mu^-$  channel, and the combined fit, respectively. Systematic sources with small uncertainties (less than 0.001) are pruned. The resulting signal

strength parameters are :

$$\begin{aligned}\mu_{Z \rightarrow ee} &= 1.007_{-0.029}^{+0.031}, \\ \mu_{Z \rightarrow \mu\mu} &= 1.016_{-0.030}^{+0.032}, \\ \mu_{Z \rightarrow \ell\ell} &= 1.011_{-0.026}^{+0.028}.\end{aligned}$$

Figure 5.21 shows the nuisance parameters fitted to the real dataset. The systematic uncertainties ranked by the impact on the signal strength are shown in Figure 5.22. The luminosity and lepton uncertainties are dominant sources of the total uncertainties in the  $Z$  cross sections.

### 5.10.3 Results of $W$ -boson production cross sections

The  $W$ -boson production cross sections are obtained through combined fits in the four  $\ell^\pm\nu$  and two  $\ell\ell$  regions. Fits are repeated on the same inputs with different combination of parameters of interest each time. The  $W^+ \rightarrow e^+\nu$ ,  $W^- \rightarrow e^-\bar{\nu}$ ,  $W^+ \rightarrow \mu^+\nu$  and  $W^- \rightarrow \mu^-\bar{\nu}$  cross sections are derived via one fit. The best-fit values and corresponding uncertainties are :

$$\begin{aligned}\mu_{W^- \rightarrow e\nu} &= 1.000_{-0.049}^{+0.052}, \\ \mu_{W^- \rightarrow \mu\nu} &= 0.976_{-0.037}^{+0.039}, \\ \mu_{W^+ \rightarrow e\nu} &= 1.000_{-0.044}^{+0.047}, \\ \mu_{W^+ \rightarrow \mu\nu} &= 0.976_{-0.036}^{+0.037}.\end{aligned}$$

The signal strengths of  $Z \rightarrow ee$  and  $Z \rightarrow \mu\mu$  are obtained at the same time,

$$\begin{aligned}\mu_{Z \rightarrow ee} &= 1.008_{-0.029}^{+0.030}, \\ \mu_{Z \rightarrow \mu\mu} &= 1.016_{-0.030}^{+0.032}.\end{aligned}$$

Similarly,  $W^+ \rightarrow \ell^+\nu$  and  $W^- \rightarrow \ell^-\bar{\nu}$  cross sections are extracted simultaneously, along with the signal strength  $\mu_{Z \rightarrow \ell\ell}$  :

$$\begin{aligned}\mu_{W^- \rightarrow \ell\nu} &= 0.978_{-0.037}^{+0.037}, \\ \mu_{W^+ \rightarrow \ell\nu} &= 0.978_{-0.036}^{+0.035}, \\ \mu_{Z \rightarrow \ell\ell} &= 1.011_{-0.027}^{+0.027}.\end{aligned}$$

$W^\pm \rightarrow \ell\nu$  can be derived with  $\mu_{Z \rightarrow \ell\ell}$  as well :

$$\begin{aligned}\mu_{W^\pm \rightarrow \ell\nu} &= 0.978_{-0.035}^{+0.035}, \\ \mu_{Z \rightarrow \ell\ell} &= 1.011_{-0.027}^{+0.027}.\end{aligned}$$

The  $Z$  production cross section extracted from combined fits are consistent with the results obtained by fitting on the di-lepton channels described in Section 5.10.2.

The pulls of the nuisance parameters of the combined fits are displayed in Figure 5.23. No large deviations from zero are found. The systematic uncertainties ranked by the impact on the signal strengths are shown in Figures 5.24-5.25. The uncertainties related to the jet energy corrections and multijet background estimation are most significant.

#### 5.10.4 Results of $\sigma_{W^+}/\sigma_{W^-}$ and $\sigma_{W^\pm}/\sigma_Z$ cross-section ratios

The ratio between  $W^+$  and  $W^-$  in the fiducial phase space,  $\sigma_{W^+}/\sigma_{W^-}$ , is obtained through a combined fit in the four  $\ell^\pm\nu$  and two  $\ell\ell$  channels. Similarly, the ratio  $\sigma_{W^\pm}/\sigma_Z$  is extracted using the same channels in the fiducial phase space. The fitted values of the corresponding signal strength parameters and their uncertainties are as follows :

$$\begin{aligned}\mu_{W^+}/\mu_{W^-} &= 1.001_{-0.018}^{+0.016}, \\ \mu_Z &= 1.012_{-0.028}^{+0.026}, \\ \mu_{W^-} &= 0.978_{-0.038}^{+0.036},\end{aligned}$$

in the  $\sigma_{W^+}/\sigma_{W^-}$  fit, and

$$\begin{aligned}\mu_{W^\pm}/\mu_Z &= 0.967_{-0.024}^{+0.024}, \\ \mu_Z &= 1.011_{-0.027}^{+0.027},\end{aligned}$$

in the  $\sigma_{W^\pm}/\sigma_Z$  fit.

The pulls of the nuisance parameters are the same as Figure 5.23 as the fits are performed with exactly the same channels. The systematic uncertainties ranked by their impact on the signal strengths are shown in Figure 5.26. As many of the experimental uncertainties such as luminosity cancel out, multijet-related uncertainty is dominant in the measurement of the  $R_{W^+}/W^-$  cross section ratio. As for  $R_{W^\pm}/Z$ , uncertainties related to multijet and jet are most significant.

#### 5.10.5 Results of $t\bar{t}$ cross section over $W^\pm$ cross section

The ratio between the cross sections of  $t\bar{t}$  and  $W^\pm$  in the fiducial phase space, denoted as  $t\bar{t}/\sigma_{W^\pm}$ , is obtained through a comprehensive combined fit. This fit includes data from four  $\ell^\pm\nu$  channels, two same-flavor  $\ell\ell$  channels, and two  $e\mu$  channels. By incorporating data from these various channels, the analysis aims to achieve consistent results among various fitting procedures. The fitted values of the corresponding signal strength parameters, along with their associated uncertainties, provide critical insights into the consistency and reliability of the measurement. These values and uncertainties are as follows :

$$\begin{aligned}\mu_{t\bar{t}}/\mu_{W^\pm} &= 0.939_{-0.026}^{+0.026}, \\ \mu_{W^\pm} &= 0.978_{-0.036}^{+0.034}, \\ \mu_Z &= 1.012_{-0.029}^{+0.026}, \\ \epsilon_b &= 0.544_{-0.004}^{+0.005}.\end{aligned}$$

Here,  $\epsilon_b$  represents the efficiency of  $b$ -jet tagging after the selection, which is extracted simultaneously with the  $t\bar{t}$  signal strength. Detailed information can be found in the  $t\bar{t}$  paper [105].

In addition to the ratio of  $t\bar{t}$  to  $W^\pm$  fiducial cross sections, the  $t\bar{t}/W^-$  and  $t\bar{t}/W^+$  fiducial cross section ratios are measured as well. The best-fit values of the signal strengths, along with their total uncertainties, are :

$$\begin{aligned}\mu_{t\bar{t}}/\mu_{W^-} &= 0.938_{-0.027}^{+0.029}, \\ \mu_{W^-} &= 0.978_{-0.037}^{+0.037}, \\ \mu_{W^+} &= 0.978_{-0.035}^{+0.035}, \\ \mu_Z &= 1.011_{-0.027}^{+0.027}, \\ \epsilon_b &= 0.544_{-0.004}^{+0.005},\end{aligned}$$

for  $t\bar{t}/W^-$ , and

$$\begin{aligned}\mu_{t\bar{t}}/\mu_{W^+} &= 0.938_{-0.026}^{+0.027}, \\ \mu_{W^-} &= 0.978_{-0.038}^{+0.036}, \\ \mu_{W^+} &= 0.978_{-0.036}^{+0.035}, \\ \mu_Z &= 1.012_{-0.028}^{+0.026}, \\ \epsilon_b &= 0.544_{-0.004}^{+0.005},\end{aligned}$$

for  $t\bar{t}/W^+$ , respectively.

There are overlapping parameters of interest in different fits, and the central values and uncertainties are consistent with each other, as well as the results derived in Sections 5.10.2-5.10.3, indicating the reliability of the statistical model.

Figure 5.27 displays the nuisance parameter pulls for the fits. No large deviations from zero are found. The rank of systematic uncertainties by the impact on the signal strength is shown in Figure 5.28. The uncertainties related to the modelling of  $t\bar{t}$  productions and multijet background estimation are dominant.

### 5.10.6 The impact of uncertainty on the results

The impact of uncertainties from various sources is summarized in Table 5.21. These uncertainties are divided into different categories. To evaluate the impact of each category, the nuisance parameters within that category are fixed at their best-fit values and treated as constants. The fitting process is then repeated, leading to a reduced uncertainty compared to the nominal values. The difference between this new uncertainty and the nominal uncertainty represents the impact of that category.

### 5.10.7 Compatibility between electron and muon channels

Simultaneous fits are performed on both electron and muon channels to obtain the combined results.

To check the compatibility between the two lepton-flavor channels, the total uncertainty of the results is decomposed into three components : the  $e - \mu$  correlated component and the electron- and muon-specific components. The combined result, including the total uncertainty (represented by the outer error bar) and the uncorrelated uncertainty (represented by the inner error bar), is compared to the electron channel

TABLE 5.21 – The observed impact (in %) of various sources of uncertainty on the measured fiducial cross sections is detailed. Flavor-tagging uncertainties and top-quark modeling uncertainties are sourced from Ref. [105].

Category	$\sigma(Z \rightarrow ee)$	$\sigma(Z \rightarrow \mu\mu)$	$\sigma(Z \rightarrow \ell\ell)$	$\sigma(W^- \rightarrow e^-\bar{\nu})$	$\sigma(W^+ \rightarrow e^+\nu)$	$\sigma(W^- \rightarrow \mu^-\bar{\nu})$	$\sigma(W^+ \rightarrow \mu^+\nu)$
Luminosity	2.2	2.2	2.2	2.5	2.5	2.5	2.4
Pile-up	1.2	0.3	0.8	1.1	1.1	0.3	0.4
MC statistics	< 0.2	< 0.2	< 0.2	< 0.2	0.4	< 0.2	0.4
Lepton trigger	0.2	0.4	0.2	1.2	1.3	1.0	1.0
Electron reconstruction	1.4	–	0.9	0.7	0.8	–	–
Muon reconstruction	–	2.1	1.4	–	–	1.0	1.0
Multi-jet	–	–	–	2.9	2.4	1.3	1.1
Other background modelling	< 0.2	< 0.2	< 0.2	< 0.2	< 0.2	0.5	0.4
Jet energy scale	–	–	–	1.4	1.4	1.3	1.4
Jet energy resolution	–	–	–	< 0.2	0.3	0.2	0.2
NNJVT	–	–	–	1.6	1.5	1.3	1.3
$E_T^{\text{miss}}$ track soft term	–	–	–	< 0.2	0.4	< 0.2	< 0.2
PDF	0.2	0.2	< 0.2	0.8	0.8	0.6	0.5
QCD scale (ME and PS)	0.6	< 0.2	0.3	1.3	1.2	0.6	0.6
Flavour tagging	–	–	–	–	–	–	–
$t\bar{t}$ modelling	–	–	–	–	–	–	–
Total systematic impact [%]	3.0	3.1	2.7	5.0	4.5	3.8	3.6
Statistical impact [%]	0.04	0.03	0.02	0.02	0.01	0.01	0.01

Category	$\sigma(W^- \rightarrow \ell^-\bar{\nu})$	$\sigma(W^+ \rightarrow \ell^+\nu)$	$\sigma(W^\pm \rightarrow \ell\nu)$	$R_{W^+/W^-}$	$R_{W^\pm/Z}$	$R_{t\bar{t}/W^\pm}$
Luminosity	2.5	2.4	2.4	< 0.2	0.3	< 0.2
Pile-up	0.5	0.7	0.6	< 0.2	< 0.2	< 0.2
MC statistics	< 0.2	0.2	< 0.2	< 0.2	< 0.2	< 0.2
Lepton trigger	1.0	0.9	0.9	< 0.2	0.7	0.8
Electron reconstruction	0.4	0.5	0.4	< 0.2	0.5	0.4
Muon reconstruction	0.6	0.6	0.6	0.2	0.8	0.6
Multi-jet	1.2	1.2	1.2	1.6	1.1	1.0
Other background modelling	0.4	0.4	0.4	< 0.2	0.3	0.9
Jet energy scale	1.3	1.3	1.3	< 0.2	1.3	1.3
Jet energy resolution	< 0.2	0.2	< 0.2	< 0.2	< 0.2	< 0.2
NNJVT	1.4	1.3	1.3	< 0.2	1.3	< 0.2
$E_T^{\text{miss}}$ track soft term	< 0.2	0.3	0.3	< 0.2	0.3	0.3
PDF	0.5	0.5	0.3	0.5	0.2	0.4
QCD scale (ME and PS)	0.8	0.7	0.6	< 0.2	0.7	0.7
Flavour tagging	–	–	–	–	–	< 0.2
$t\bar{t}$ modelling	–	–	–	–	–	1.1
Total systematic impact [%]	3.7	3.5	3.5	1.7	2.4	2.5
Statistical impact [%]	0.01	0.01	0.01	0.01	0.02	0.32

result with its specific uncertainty and the muon channel result with its specific uncertainty. The results are displayed in Figure 5.29, showing good compatibility between the two lepton-flavor channels.

To quantify the compatibility of the measurements, the  $\chi^2$  values are calculated as :

$$\chi^2 = (X_e - X_{\text{comb}}, X_\mu - X_{\text{comb}}) \begin{pmatrix} \sigma_e^2 & \rho\sigma_e\sigma_\mu \\ \rho\sigma_e\sigma_\mu & \sigma_\mu^2 \end{pmatrix}^{-1} (X_e - X_{\text{comb}}, X_\mu - X_{\text{comb}})^T \quad (5.15)$$

where  $X_e$  is the electron cross-section result,  $X_\mu$  is the muon result,  $X_{\text{comb}}$  is the common cross section shared by electron and muon channels obtained from the combined fit,  $\sigma_e$  and  $\sigma_\mu$  are the total uncertainties of electron and muon channel results, and  $\rho$  is the correlation coefficient.

The total uncertainty in each channel is divided into two distinct components : the uncertainty correlated between the electron and muon channels, and the uncertainty uncorrelated between the two channels. The term  $\rho\sigma_e\sigma_\mu$  quantifies this correlation by multiplying the correlated components of uncertainty in the electron and muon channels, i.e.  $\mu_e^{\text{corr}}$  and  $\mu_\mu^{\text{corr}}$ .

The  $\chi^2$  values per number of degree of freedom ( $N_{\text{dof}}$ ) and corresponding  $p$ -

values are summarized in Table 5.22.

TABLE 5.22 – The  $\chi^2/N_{\text{dof}}$  and  $p$ -values of the combination.

Channel	$\chi^2/N_{d.f}$	$p$ -value
$Z$	0.10/1	0.92
$W^+$	1.16/1	0.25
$W^-$	1.03/1	0.30
$Z, W^+$ and $W^-$	2.29/3	0.51

### 5.10.8 Summary of the results and comparison with predictions

The fiducial cross sections and their ratios are extracted by applying the signal strengths from fits on the predicted cross sections shown in Table 5.20. The  $Z$  cross sections ( $\sigma(Z \rightarrow ee)$ ,  $\sigma(Z \rightarrow \mu\mu)$ , and  $\sigma(Z \rightarrow \ell\ell)$ ) are obtained by fitting the same-flavor di-lepton regions as outlined in Section 5.10.2. The  $W$  cross sections and the fiducial ratios ( $W^+/W^-$  and  $W^\pm/Z$ ) are determined using four single-lepton and two same-flavor di-lepton regions. These regions use the same inputs, with the parameters of interest varied for each specific fit. The simultaneous determination of  $W^+ \rightarrow e^+\nu$ ,  $W^+ \rightarrow \mu^+\nu$ ,  $W^- \rightarrow e^-\bar{\nu}$ ,  $W^- \rightarrow \mu^-\bar{\nu}$ ,  $W^+ \rightarrow \ell^+\nu$ ,  $W^- \rightarrow \ell^-\bar{\nu}$ , and  $W^\pm \rightarrow \ell\nu$  cross sections is described in Section 5.10.3. The  $W^+/W^-$  and  $W^\pm/Z$  ratios, are extracted through separate fits, detailed in Sections 5.10.4. For the  $R_{t\bar{t}/W^\pm}$ ,  $R_{t\bar{t}/W^-}$  and  $R_{t\bar{t}/W^+}$ , the signal strengths are described in Section 5.10.5.

In the further step, the total cross sections are obtained by taking the ratio of fiducial cross sections to the acceptance correction. Table 5.23 shows the acceptance correction factors for each channels, along with the theoretical uncertainties arising from the PDF and QCD scale variations, which are detailed in Section 5.7.2. The measured cross-section results, along with their total uncertainties, are presented in Table 5.23.

Figure 5.30 presents a comparison between the measured results for the  $W^-$ ,  $W^+$ , and  $Z$  boson fiducial cross sections and  $R_{W^\pm/Z}$ ,  $R_{W^+/W^-}$ , and  $t\bar{t}/W$  ratios and the NNLO QCD + NLO EW theory predictions, utilizing various PDF sets described in Section 5.6. The theoretical predictions are computed using PDF4LHC21 [141], CT18, CT18A [34], MSHT20 [35-37], NNPDF4.0 [38], ABMP16 [142], and ATLASpdf21 [39].

For the  $W^-$ ,  $W^+$ , and  $Z$  boson fiducial cross sections, as well as the  $R_{W^\pm/Z}$  and  $R_{W^+/W^-}$  ratios, there is an overall good agreement with the Standard Model predictions. However, the  $t\bar{t}/W$  ratio results are slightly lower than the predictions for most of the PDFs considered. This observation is consistent with the findings of the Run 3  $t\bar{t}$  cross-section measurement [105], which also showed the measured  $t\bar{t}$  cross section to be lower than the predicted value. Additionally, Figure 5.30 includes predictions based on the PDF4LHC21 set with  $m_t = 171.5$  GeV and  $m_t = 173.5$  GeV.

TABLE 5.23 – The measured cross sections using the profile-likelihood method. The quoted uncertainty corresponds to the total uncertainty (the statistical uncertainty is negligibly small). Rounding has been applied to all quoted numbers.

Channel	$\sigma^{\text{fid}} \pm \delta\sigma_{\text{stat} \oplus \text{syst}}$ [pb]	Acceptance $A$	$\sigma^{\text{tot}} \pm \delta\sigma_{\text{stat} \oplus \text{syst}}$ [pb]
$Z \rightarrow e^+e^-$	$740 \pm 22$	$0.374 \pm 0.011$	$1981 \pm 82$
$Z \rightarrow \mu^+\mu^-$	$747 \pm 23$	$0.374 \pm 0.011$	$1997 \pm 82$
$Z \rightarrow \ell^+\ell^-$	$744 \pm 20$	$0.374 \pm 0.011$	$1989 \pm 77$
$W^- \rightarrow e^-\bar{\nu}$	$3380 \pm 170$	$0.381 \pm 0.009$	$8880 \pm 490$
$W^- \rightarrow \mu^-\bar{\nu}$	$3310 \pm 130$	$0.381 \pm 0.009$	$8680 \pm 390$
$W^- \rightarrow \ell^-\bar{\nu}$	$3310 \pm 120$	$0.381 \pm 0.009$	$8690 \pm 390$
$W^+ \rightarrow e^+\nu$	$4350 \pm 200$	$0.366 \pm 0.009$	$11880 \pm 620$
$W^+ \rightarrow \mu^+\nu$	$4240 \pm 160$	$0.365 \pm 0.010$	$11620 \pm 530$
$W^+ \rightarrow \ell^+\nu$	$4250 \pm 150$	$0.366 \pm 0.009$	$11620 \pm 520$
$W^\pm \rightarrow \ell^\pm\nu$	$7560 \pm 270$	$0.372 \pm 0.009$	$20310 \pm 890$
<b>Ratio</b>	<b><math>R \pm \delta R_{\text{stat} \oplus \text{syst}}</math></b>		
$W^+/W^-$	$1.286 \pm 0.022$		
$W^\pm/Z$	$10.17 \pm 0.25$		
$t\bar{t}/W^-$	$0.256 \pm 0.008$		
$t\bar{t}/W^+$	$0.199 \pm 0.006$		
$t\bar{t}/W^\pm$	$0.112 \pm 0.003$		

Figure 5.31 displays the total cross sections of  $W$  and  $Z$  productions in the lepton decay channel as a function of the centre-of-mass energy,  $\sqrt{s}$ , along with the theory predictions computed at NNLO QCD accuracy with CT14 [150] PDF set. Good agreement is observed between the SM predictions and the measured cross sections at the new centre-of-mass energy,  $\sqrt{s} = 13.6$  TeV.

## 5.11 Conclusion

A measurement of the inclusive  $W^+$ ,  $W^-$ , and  $Z$  boson production cross sections, including their decay channels into electrons or muons, their ratios, and the ratios of  $t\bar{t}$  to  $W$  boson fiducial cross sections, is performed using  $pp$  collision data at the unprecedented centre-of-mass energy of 13.6 TeV at the LHC, corresponding to an integrated luminosity of  $29 \text{ fb}^{-1}$ . The cross sections are measured for  $Z$  boson production in a fiducial volume corresponding to  $p_T^\ell > 27 \text{ GeV}$ ,  $|\eta_\ell| < 2.5$ , and  $66 < m_{\ell\ell} < 116 \text{ GeV}$  and for  $W$  boson production in a fiducial region corresponding

to  $p_T^\ell > 27 \text{ GeV}$ ,  $|\eta_\ell| < 2.5$ ,  $m_T^W > 50 \text{ GeV}$ , and  $E_T^{\text{miss}} > 25 \text{ GeV}$ .

The fiducial cross-section results for  $W^+$ ,  $W^-$ , and  $Z$  boson productions in leptonic decay channels are  $4250 \pm 150 \text{ pb}$ ,  $3310 \pm 120 \text{ pb}$ , and  $744 \pm 20 \text{ pb}$ , respectively. These quoted uncertainties represent the total uncertainty, which contains the luminosity uncertainty of approximately 2.2%. The ratios of fiducial cross sections are measured as follows :  $R_{W^+/W^-} = 1.286 \pm 0.022$ ,  $R_{W^\pm/Z} = 10.17 \pm 0.25$ , and  $R_{t\bar{t}/W^\pm} = 0.112 \pm 0.003$ , with the uncertainties representing the total uncertainty.

Additionally, total cross sections are measured for  $Z$  boson production within the mass range  $66 \text{ GeV} < m_{\ell\ell} < 116 \text{ GeV}$  and for  $W$  boson production in the full phase space. The study includes measurements for  $\sigma(W^+)$ ,  $\sigma(W^-)$ ,  $\sigma(W^\pm)$ ,  $\sigma(Z)$ ,  $R_{W^+/W^-}$ ,  $R_{W^\pm/Z}$ ,  $R_{t\bar{t}/W^\pm}$ ,  $R_{t\bar{t}/W^+}$ , and  $R_{t\bar{t}/W^-}$ .

The results show that the measured  $W$  and  $Z$  boson cross sections are in good agreement with Standard Model predictions. However, the ratios of  $t\bar{t}$  to  $W$  boson fiducial cross sections are slightly overestimated by some theoretical predictions.

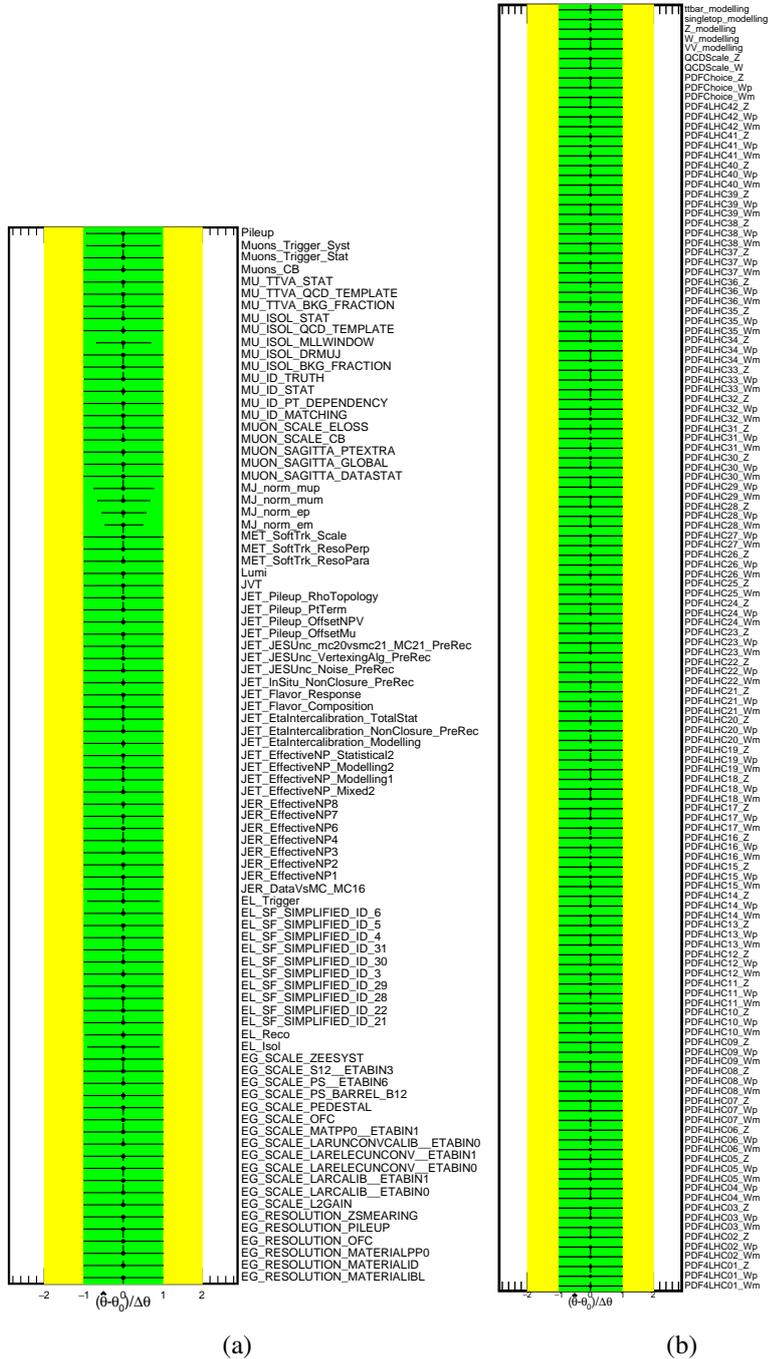


FIGURE 5.14 – Nuisance parameters fitted to the Asimov datasets for the  $W$  and  $Z$  production cross sections and ratios. The nuisance parameters are categorised into 2 classes : (a) experimental uncertainties and (b) theoretical uncertainties.

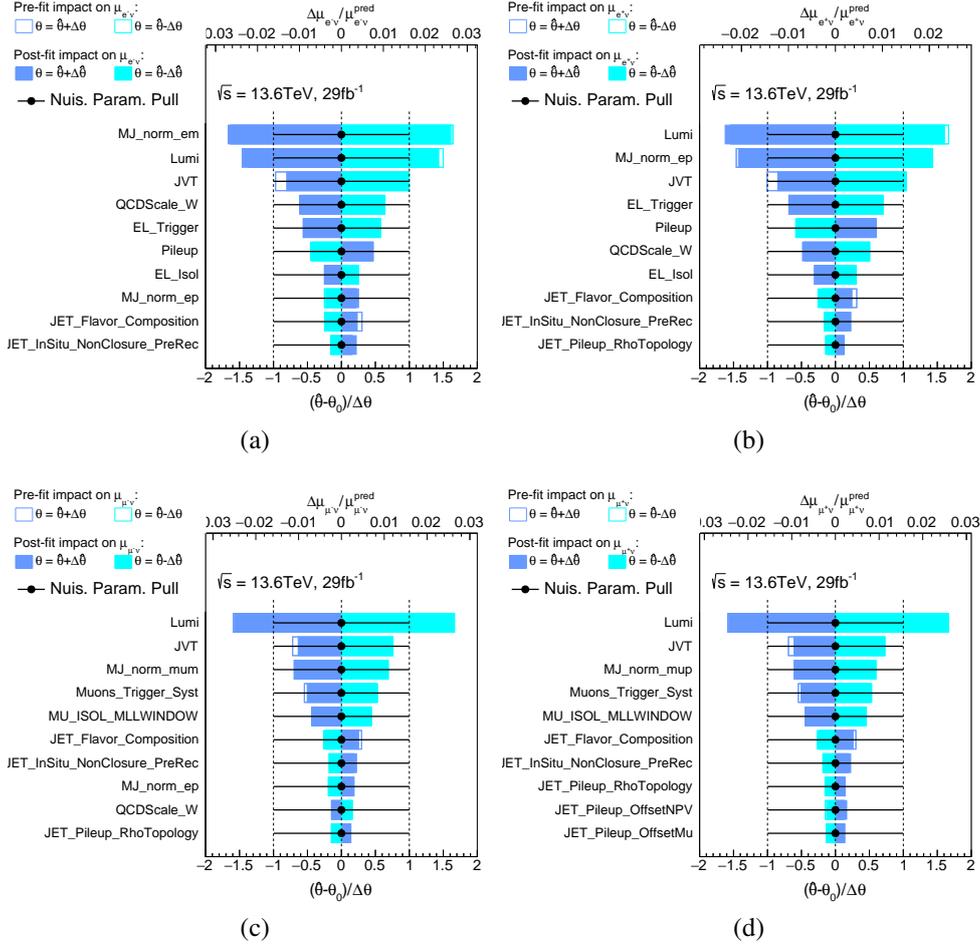


FIGURE 5.15 – Systematic uncertainties ranked by the impact on the signal strength  $\mu$ . The impact is shown before and after the fits to the Asimov datasets for (a)  $W^- \rightarrow e^- \bar{\nu}$ , (b)  $W^+ \rightarrow e^+ \nu$ , (c)  $W^- \rightarrow \mu^- \bar{\nu}$  and (d)  $W^+ \rightarrow \mu^+ \nu$  cross section. The impact is shown before and after the fits to the Asimov datasets.

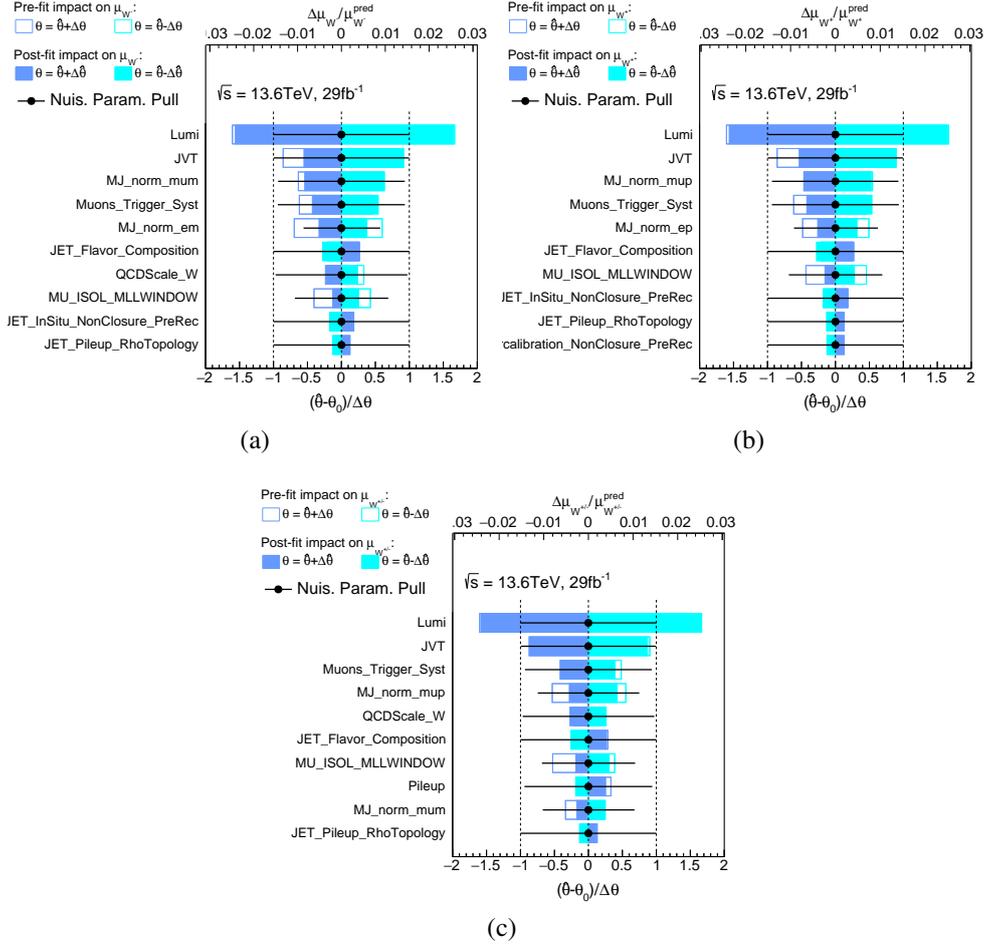


FIGURE 5.16 – Systematic uncertainties ranked by the impact on the signal strength  $\mu$ . The impact is shown before and after the fits to the Asimov datasets for (a)  $W^- \rightarrow \ell^- \bar{\nu}$ , (b)  $W^+ \rightarrow \ell^+ \nu$  and (c)  $W^\pm \rightarrow \ell \nu$  cross section. The impact is shown before and after the fits to the Asimov datasets.

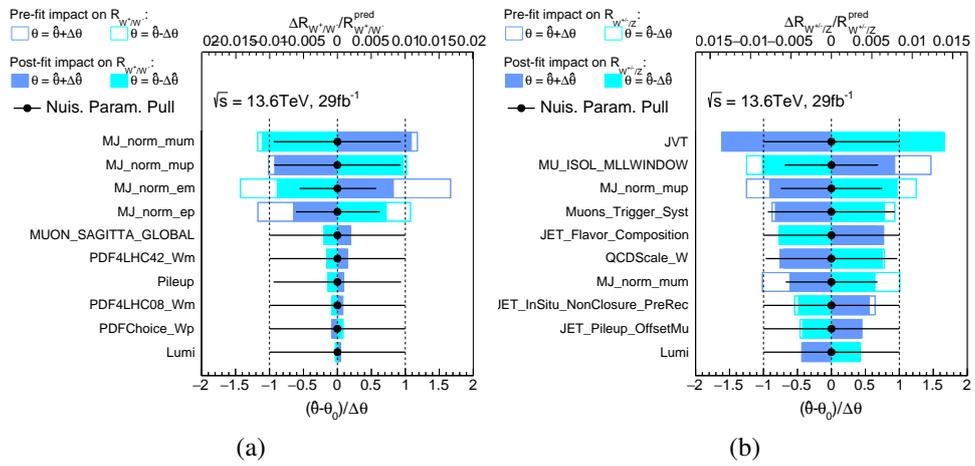


FIGURE 5.17 – Systematic uncertainties ranked by the impact on the ratio of  $W^+/W^-$  and  $W^\pm/Z$ . The impact is shown before and after the fits to the Asimov datasets for (a)  $W^+/W^-$  and (b)  $W^\pm/Z$  cross section ratio. The impact is shown before and after the fits to the Asimov datasets.

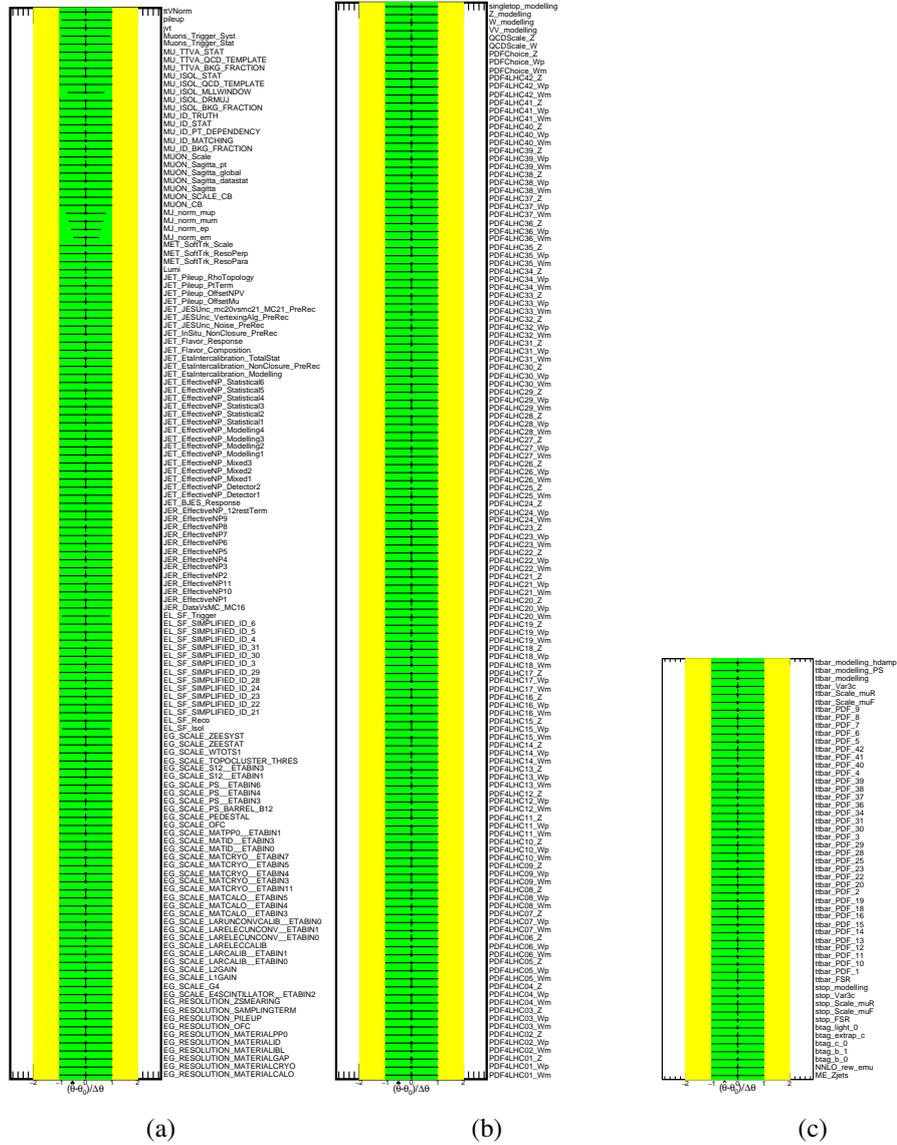


FIGURE 5.18 – Nuisance parameters fitted to Asimov datasets for the  $t\bar{t}$  over  $W^{\pm}$  cross section ratios. The nuisance parameters are categorised into 3 classes : (a) experimental uncertainties on  $W$ -cross section, (b) theoretical uncertainties on  $W$  cross section and (c) additional uncertainties from  $t\bar{t}$  cross section.

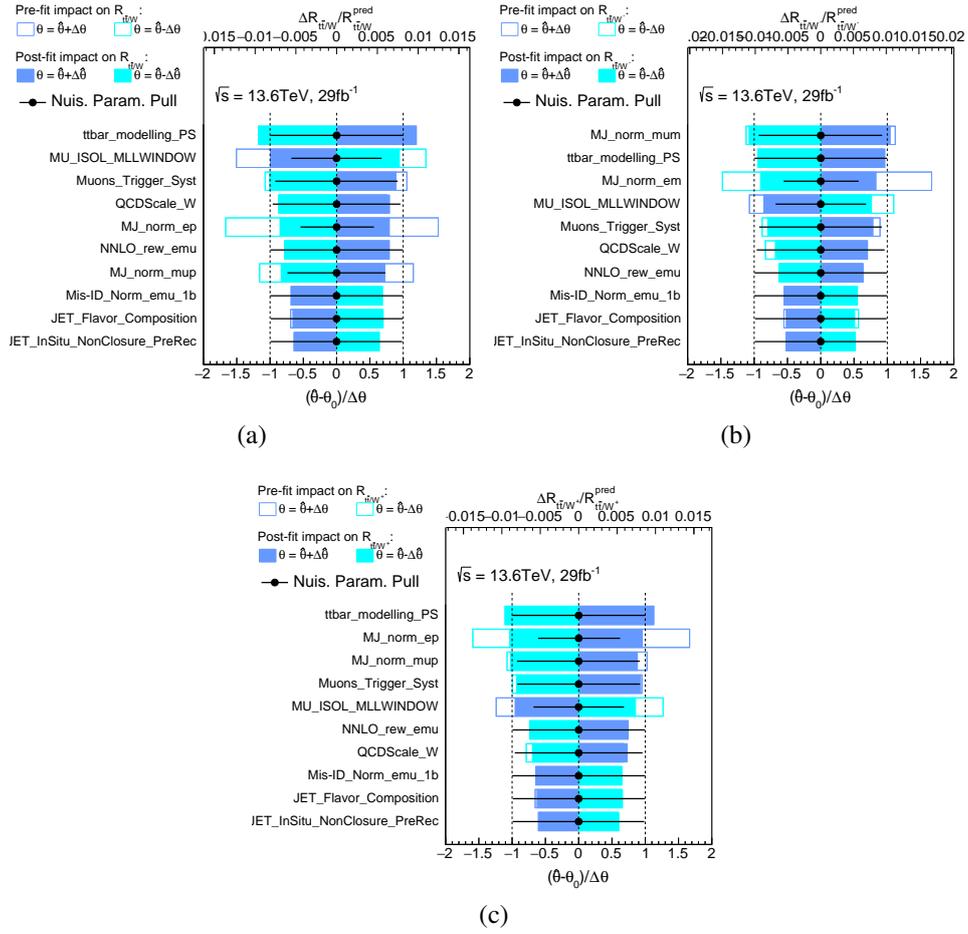


FIGURE 5.19 – Systematic uncertainties ranked by the impact on the ratio of (a)  $t\bar{t}/W^\pm$ , (b)  $t\bar{t}/W^-$  and (c)  $t\bar{t}/W^+$  production cross sections. The impact is shown before and after the fits to the Asimov datasets.

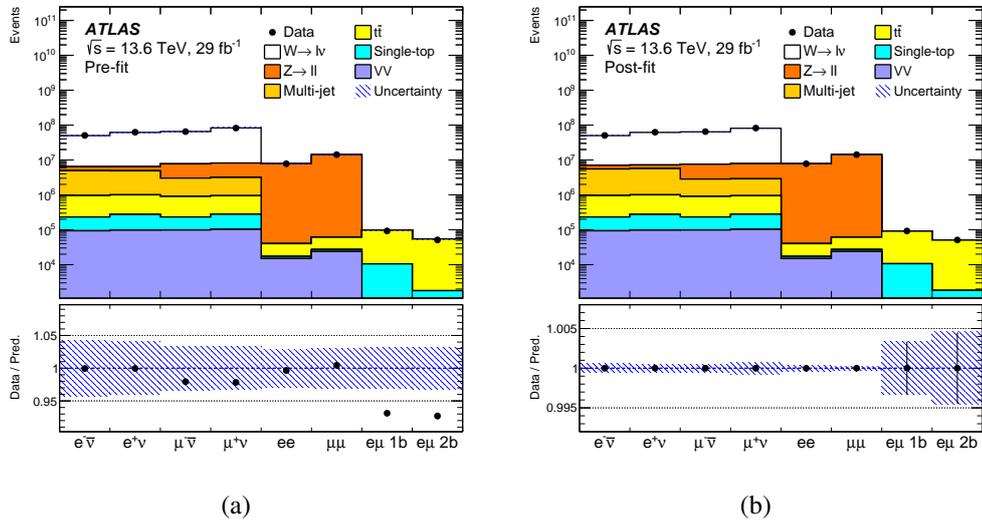


FIGURE 5.20 – The comparison of data and predictions (a) before and (b) after fits in all channels. The dashed error band in the pre-fit plot (a) represents the total systematic uncertainty, while in the post-fit plot (b), it represents the statistical uncertainty derived from the fit.

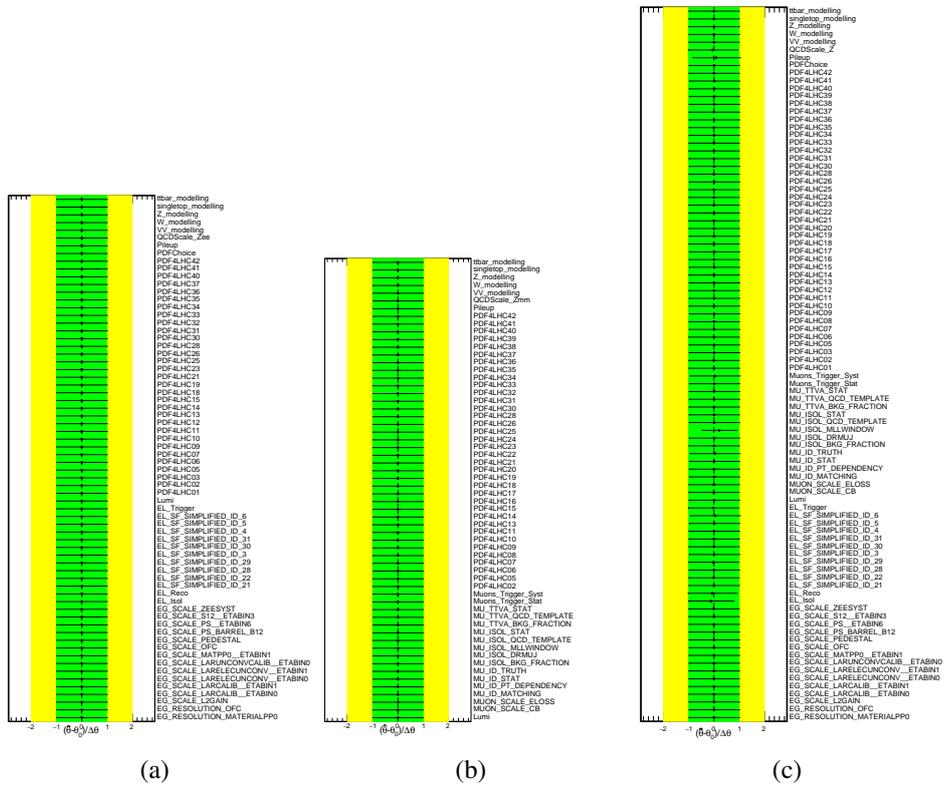


FIGURE 5.21 – Nuisance parameter pulls for the (a)  $Z \rightarrow ee$ , (b)  $Z \rightarrow \mu\mu$  fit and (c) combined fit.

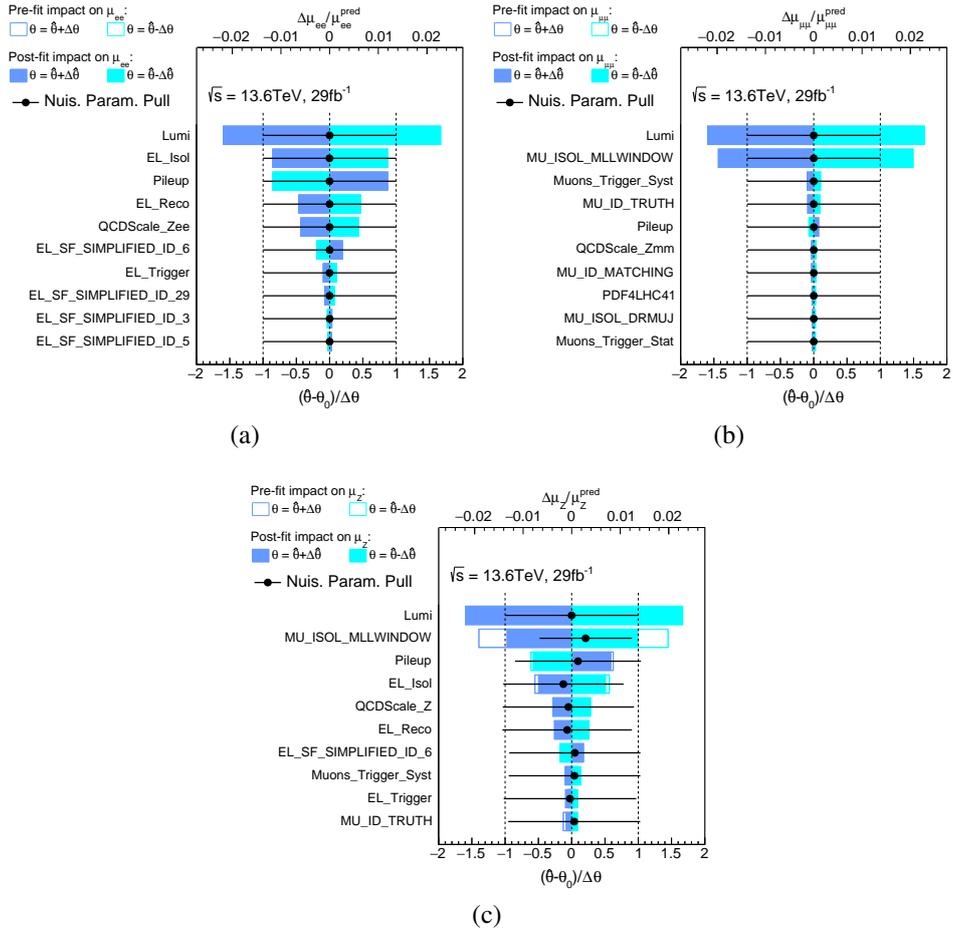


FIGURE 5.22 – Systematic uncertainties ranked by the impact on the signal strength  $\mu$  in (a) the  $Z \rightarrow ee$  fit, (b) the  $Z \rightarrow \mu\mu$  fit, and in (c) the combined fit.



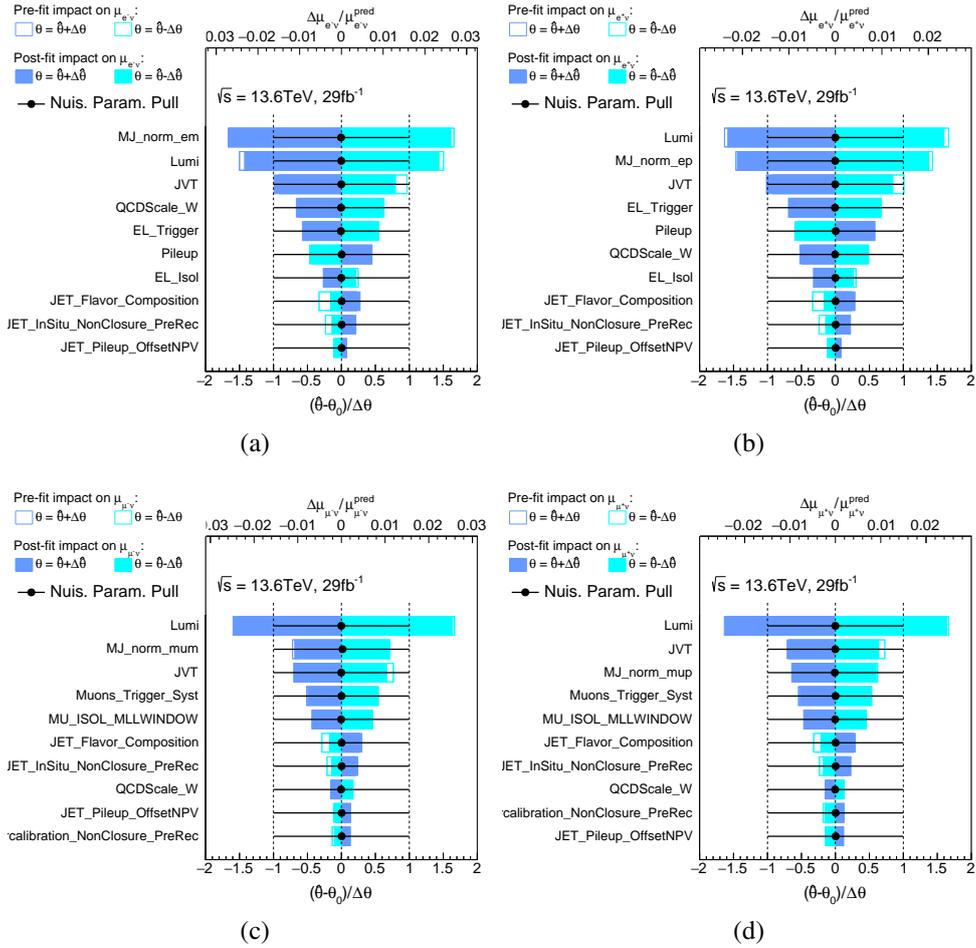


FIGURE 5.24 – Systematic uncertainties ranked by the impact on the signal strength  $\mu$ . The impact is shown before and after the combined fits for (a)  $W^- \rightarrow e^- \bar{\nu}$ , (b)  $W^+ \rightarrow e^+ \nu$ , (c)  $W^- \rightarrow \mu^- \bar{\nu}$  and (d)  $W^+ \rightarrow \mu^+ \nu$  cross section. The impact is shown before and after the fits to the Asimov datasets.

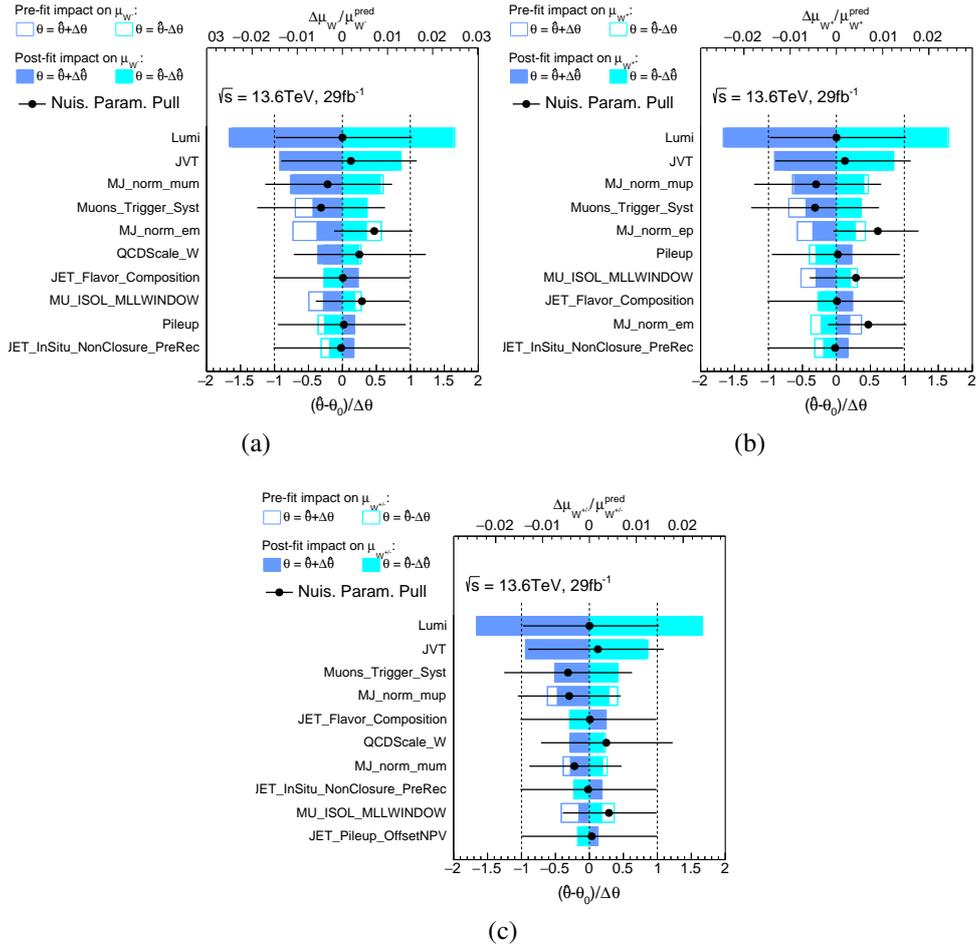


FIGURE 5.25 – Systematic uncertainties ranked by the impact on the signal strength  $\mu$ . The impact is shown before and after the combined fits for (a)  $W^- \rightarrow \ell^- \bar{\nu}$ , (b)  $W^+ \rightarrow \ell^+ \nu$  and (c)  $W^\pm \rightarrow \ell \nu$  cross section. The impact is shown before and after the fits to the Asimov datasets.



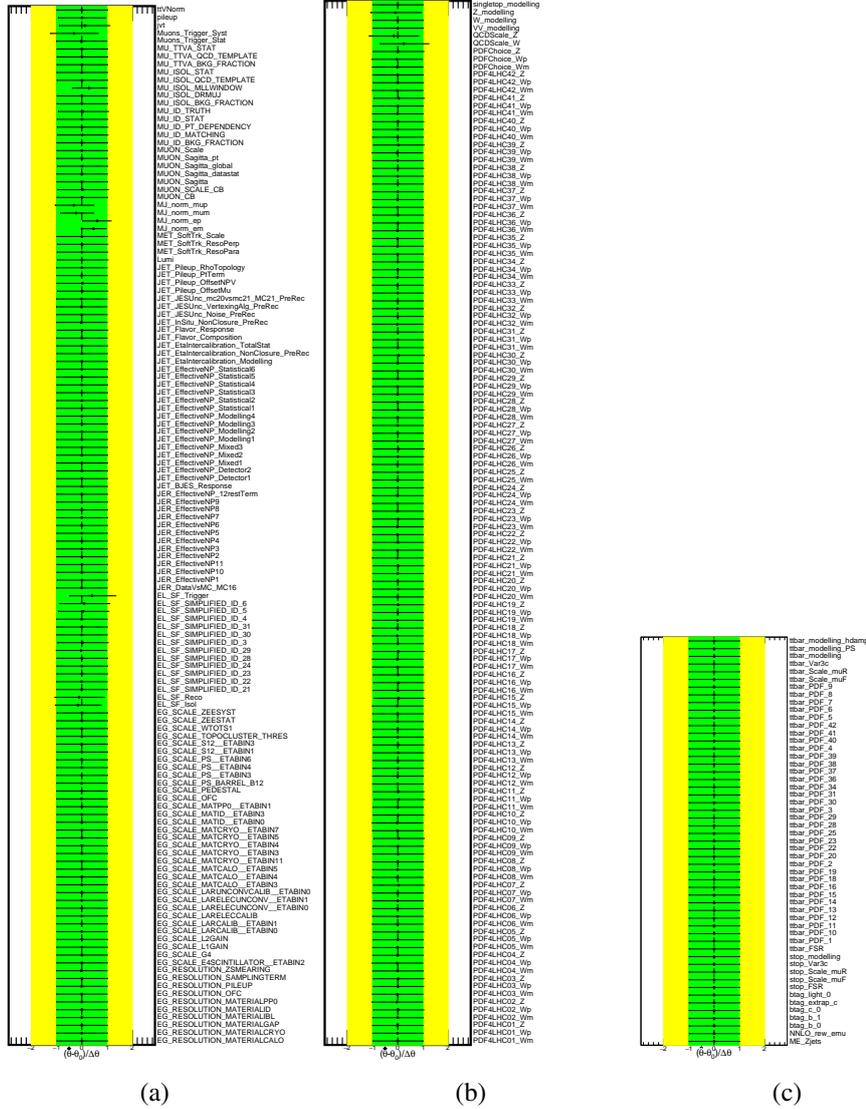


FIGURE 5.27 – Nuisance parameters for the  $t\bar{t}$  over  $W^\pm$  cross section ratios. The nuisance parameters are categorised into 3 classes : (a) experimental uncertainties on  $W$ -cross section, (b) theoretical uncertainties on  $W$  cross section and (c) additional uncertainties from  $t\bar{t}$  cross section.

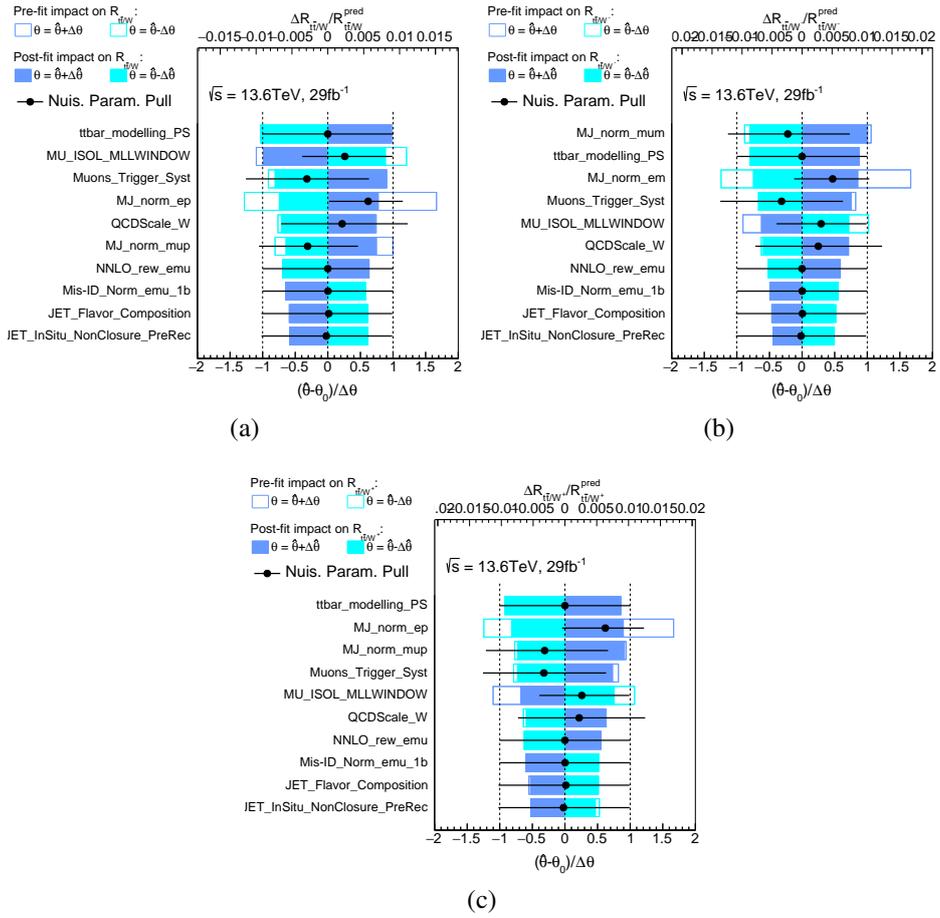


FIGURE 5.28 – Systematic uncertainties ranked by the impact on the ratio of (a)  $t\bar{t}/W^\pm$ , (b)  $t\bar{t}/W^-$  and (c)  $t\bar{t}/W^+$ . The impact is shown before and after the fits.

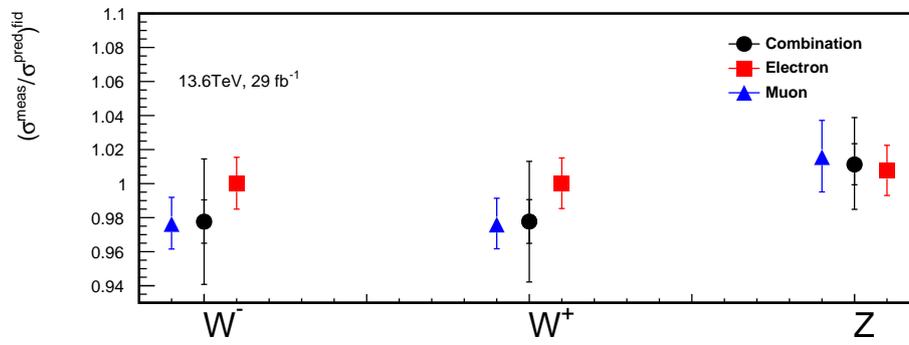
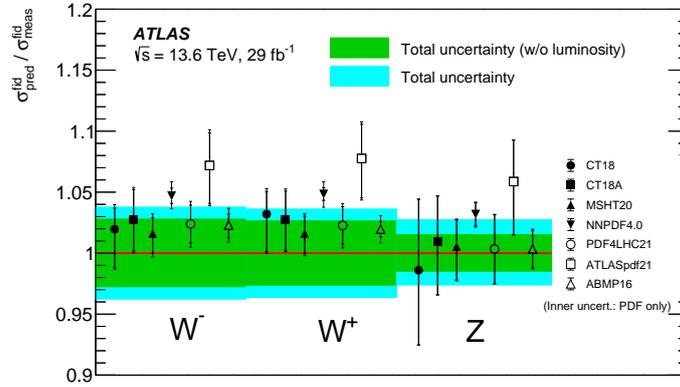
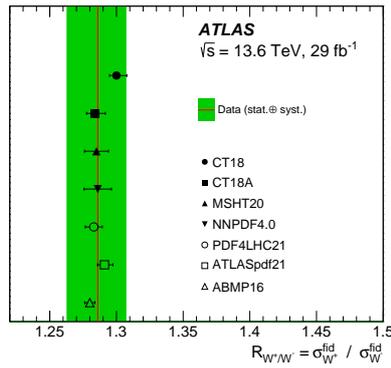


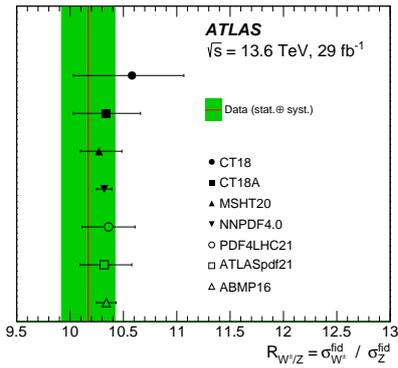
FIGURE 5.29 – Compatibility between electron and muon channels for the  $W^-$ ,  $W^+$ , and  $Z$  processes is illustrated. The central values represent the ratios of the measured fiducial cross sections to the predicted values for each channel. The inner error bar for the combined results represents the uncorrelated systematic uncertainty, which includes both statistical uncertainty and lepton-flavor-specific uncertainties. The outer error bar represents the total uncertainty. The error bars for the electron and muon channels indicate the contributions of uncertainty sources specific to each lepton flavor.



(a)



(b)



(c)

FIGURE 5.30 – The comparison of predictions using different PDF sets to the measured fiducial cross sections is presented for (a)  $W^-$ ,  $W^+$ , and  $Z$  bosons, their ratios (b)  $R_{W^+/W^-}$  and (c)  $R_{W^\pm/Z}$ , and (d) the ratio  $R_{t\bar{t}/W^\pm}$ , (e) the ratio  $R_{t\bar{t}/W^+}$ , and (f) the ratio  $R_{t\bar{t}/W^-}$ . In plot (a), the outer band represents the total uncertainty including the luminosity uncertainty, while the inner band excludes it. In the other plots, the vertical band indicates the total uncertainty, encompassing both systematic and statistical uncertainties. The error bars on the predictions reflect the theoretical uncertainties, with the inner error bars (where available) showing the contributions from the PDF uncertainties.

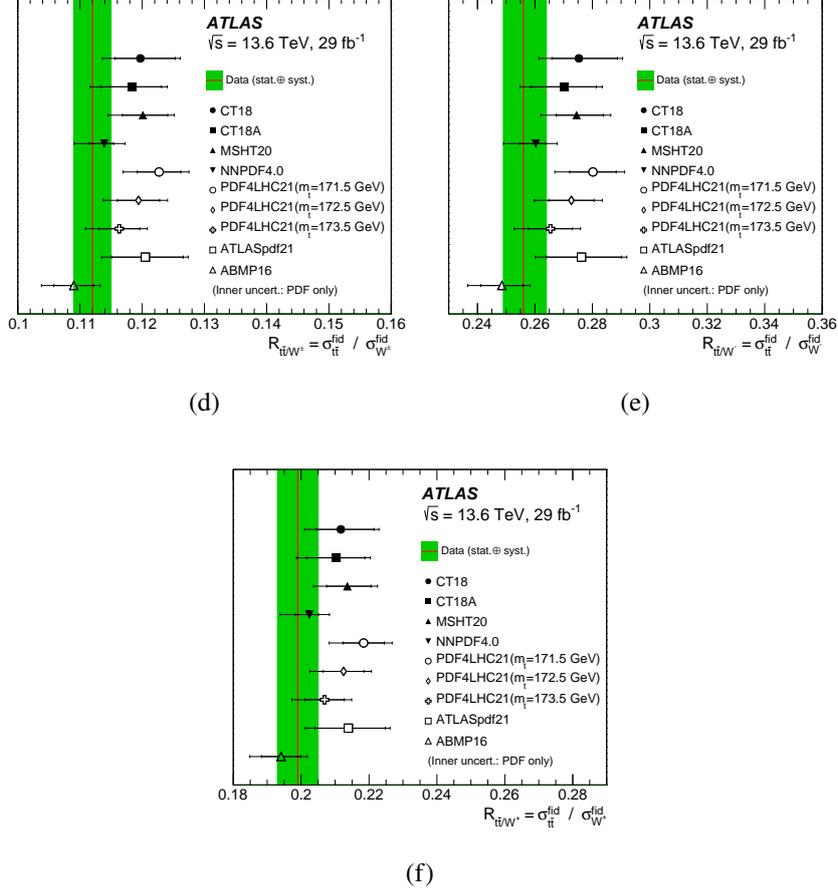


FIGURE 5.30 – The comparison of predictions using different PDF sets to the measured fiducial cross sections is presented for (a)  $W^-$ ,  $W^+$ , and  $Z$  bosons, their ratios (b)  $R_{W^+/W^-}$  and (c)  $R_{W^\pm/Z}$ , and (d) the ratio  $R_{t\bar{t}/W^\pm}$ , (e) the ratio  $R_{t\bar{t}/W^+}$ , and (f) the ratio  $R_{t\bar{t}/W^-}$ . In plot (a), the outer band represents the total uncertainty including the luminosity uncertainty, while the inner band excludes it. In the other plots, the vertical band indicates the total uncertainty, encompassing both systematic and statistical uncertainties. The error bars on the predictions reflect the theoretical uncertainties, with the inner error bars (where available) showing the contributions from the PDF uncertainties.

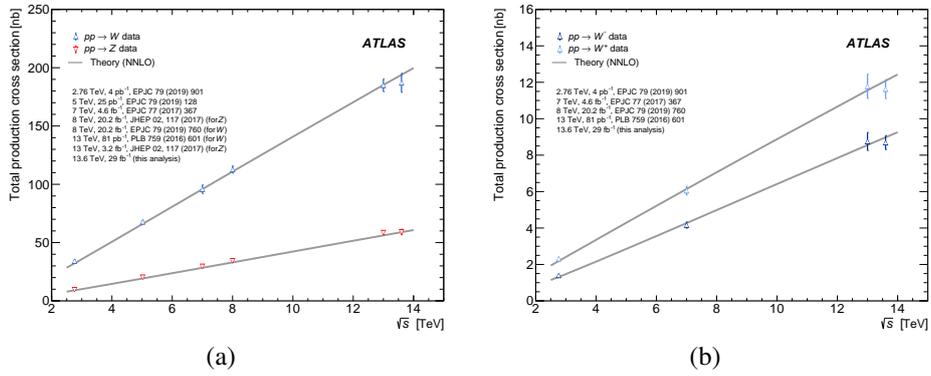


FIGURE 5.31 – The measured (a)  $W^\pm$  and  $Z$  and (b)  $W^-$  and  $W^+$  boson total production cross sections in the leptonic decay channel at different center-of-mass energies are shown. The theory predictions is computed at NNLO QCD accuracy based on the CT14NNLO PDF set.



## 6 - Measurement of $W^\pm$ boson differential cross sections with low pile-up data at $\sqrt{s} = 5.02$ and 13 TeV

This chapter details the measurement of differential cross sections for  $W^\pm \rightarrow \ell^\pm \nu$  processes at center-of-mass energies of  $\sqrt{s} = 5.02$  and 13 TeV. The datasets were collected by ATLAS during the special LHC runs with low pile-up conditions. The measured results are compared with the theoretical predictions calculated at NNLO+NNLL QCD and LO EW accuracy with various sets of PDFs. The PDF sensitivity of the results is evaluated using the PDF profiling technique. The associated charge asymmetry is also measured.

### 6.1 Introduction

The measurement of the differential cross sections of  $W$  boson production are performed using data recorded by the ATLAS detector in 2017 and 2018 from proton-proton collisions at centre-of-mass energies of 5.02 TeV and 13 TeV. The dataset at 5.02 TeV corresponds to an integrated luminosity of  $255 \text{ pb}^{-1}$ , while the dataset at 13 TeV corresponds to an integrated luminosity of  $338 \text{ pb}^{-1}$ . These datasets were recorded during special runs in the Run 2 data-taking period, characterized by low pile-up conditions. The typical mean number of interactions per bunch crossing,  $\langle \mu \rangle$ , is approximately 2. These datasets were specifically collected for precision measurements. The low pileup conditions result in reduced background noise, ensuring highly precise measurements.

The differential cross-section measurement not only probes the pQCD theory but also reveals more information about the structure of protons, as described in Section 2. The results reflect the distribution of momentum fraction carried by each parton, also known as Bjorken  $x$ . The current PDFs can be constrained by incorporating the new cross section results. Improvements in the PDFs can benefit other analyses. For example, in the  $W$  mass measurement, where the PDF uncertainty is one of the dominant sources of error, contributing a 9.2 MeV uncertainty in the ATLAS measurements at 7 TeV [151].

In this analysis, the differential cross sections of  $W$  boson production in the leptonic decay channels are measured in the phase space defined by the lepton transverse momentum  $p_T^\ell > 25 \text{ GeV}$ , lepton pseudorapidity  $|\eta_\ell| < 2.5$ , transverse mass greater 50 GeV and transverse missing energy greater than 25 GeV. The single differential cross sections are presented as a function of  $p_T^\ell$  and  $|\eta_\ell|$ , respectively. The double differential cross sections are also provided as functions of  $p_T^\ell$  and  $|\eta_\ell|$ . In addition to the cross sections, the lepton charge asymmetry is also investigated.

The strategy of the analysis is as follows : events are selected based on specific

criteria to isolate  $W \rightarrow \ell\nu$  process. The backgrounds are estimated and subtracted from the data event yields, ensuring the measured distributions reflect the signal distributions. Detector effects are then corrected using unfolding techniques to retrieve the true differential cross section distributions at the particle level. The differential cross sections are first measured separately in the electron and muon channels, respectively, and combination procedures are then performed to derive the final cross-section results. The measured results are compared to the theoretical predictions computed up to NNLO plus NNLL accuracy in QCD and leading-order precision in electroweak calculation. The sensitivity of the measured cross sections in pseudorapidity to the PDFs is evaluated using PDF profiling techniques.

## 6.2 Data and MC samples

This section provides an overview of the data and MC simulations involved in this measurement.

### 6.2.1 Data samples

The measurement is performed using the data collected during 2017 and 2018 by the ATLAS detector in  $pp$  collisions at the centre-of-mass energies  $\sqrt{s} = 5.02$  and 13 TeV. The dataset at 5.02 TeV corresponds to an integrated luminosity of  $254.9 \pm 2.6 \text{ pb}^{-1}$ , while for 13 TeV, the integrated luminosity is  $338 \pm 3.1 \text{ pb}^{-1}$  [152]. The quoted uncertainty is based on LUCID-2 detector [144] estimate.

These datasets were recorded during special LHC runs, characterized by a low pile-up conditions. The mean number of interactions per bunch crossing,  $\langle\mu\rangle$ , ranges from approximately 1 to 5 for the datasets. Specifically, the 13 TeV data was collected with a stable, leveled luminosity at  $\langle\mu\rangle = 2$ . Figure 6.1 depicts the  $\langle\mu\rangle$  conditions for the Run-2 dataset at 13 TeV, with the leftmost dataset indicating the low pile-up runs. These low pile-up conditions, characterized by fewer interactions per bunch crossing, are crucial for minimizing background noise and enhancing the precision of the measurements. The low pile-up environment allows for more accurate analysis and better quality data, particularly important for determining the  $W$  boson mass and measuring the  $W$  and  $Z$  boson production cross sections.

### 6.2.2 MC samples

The Monte Carlo simulations used in this measurement are specifically generated to address the low pile-up conditions. As a result, unlike previous measurements in Section 5, there is no need to reweight the MC samples based on the  $\mu$  distribution to enhance the agreement with the observed data.

The production of  $W$  and  $Z$  bosons is simulated using the POWHEG BOX v2[114-116, 153] generator. The ME simulation is coupled with PYTHIA[117] to model the parton shower, hadronization, and underlying event. This is done using the CT10 PDF set [119], with parameters configured according to the AZNLO tune [118]. Additionally, the final QED radiation is simulated with PHOTOS [121].

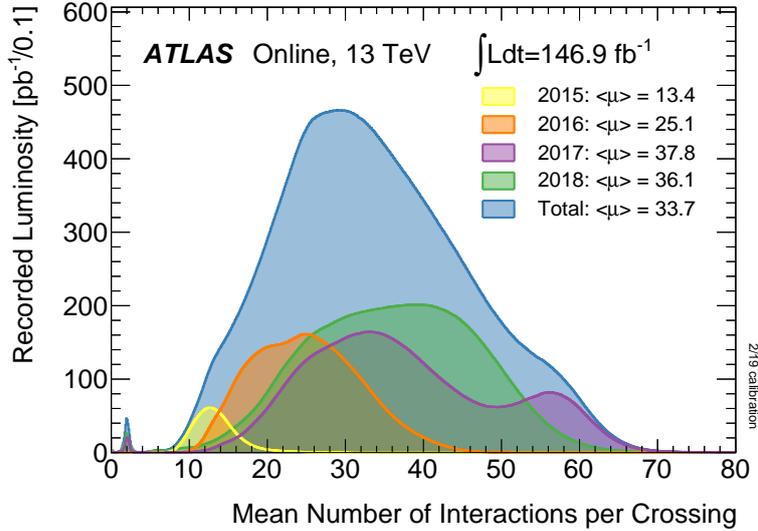


FIGURE 6.1 – The mean number of interactions per bunch crossing for the 2015-2018  $pp$  collision data at  $\sqrt{s} = 13$  TeV. The leftmost peak illustrates the low pile-up conditions of the datasets involved in this measurement.

For  $\sqrt{s} = 13$  TeV, alternative  $W$  and  $Z$  samples are generated using SHERPA 2.2.2 [154]. For the 5.02 TeV dataset, SHERPA 2.2.5 is employed. These simulations use the NNPDF3.0 PDF set [112], with NLO accuracy for configurations involving up to two partons and LO accuracy for configurations with up to four partons, implemented using the MEPS@NLO scheme [110, 111, 155-157]. SHERPA samples are used to estimate the uncertainty due to the choice of generators.

The  $W$  and  $Z$  samples are normalized to NNLO predictions calculated using the DYTURBO [158, 159] programme with the MMHT2014nnlo PDF set [160].

Diboson samples are simulated using SHERPA 2.2.1 and 2.2.2, accounting for off-shell effects and including Higgs boson contributions where applicable. The generation covers fully leptonic final states as well as semileptonic final states, where one boson decays leptonically and the other hadronically. The MEs are calculated with NLO accuracy in QCD for configurations involving up to one additional parton and with LO accuracy for configurations involving up to three additional parton emissions.

The production of top-quark pairs  $t\bar{t}$  and the associated production of a single top-quark with a  $W$  boson  $Wt$  were simulated using the POWHEG BOX v2 generator, which was interfaced with PYTHIA for parton showering and hadronization. The 5.02 TeV  $t\bar{t}$  sample was normalized according to the cross-section prediction provided in Ref. [161], using calculations from TOP++ [124-130].

### 6.3 Object reconstruction

The object reconstructions are detailed in Chapter 4. This section gives a brief description to the selections and corrections on the physics objects involved in this measurement.

Electron candidates are reconstructed following the standard way in the ATLAS experiment, detailed in Section 4.2. Electrons with transverse momentum  $p_T^\ell$  greater than 25 GeV and pseudorapidity  $|\eta| < 2.47$  are selected, excluding electrons in the transition region between barrel and endcap, corresponding to the pseudorapidity range  $1.37 < |\eta| < 1.52$ . The Medium criteria of electron identification [74, 75] are applied to the selected electron candidates to ensure their quality and reduce background contamination.

Muon candidates are reconstructed following the procedure presented in Section 4.3. Muons with transverse momentum greater than 25 GeV and within the pseudorapidity range  $|\eta| < 2.4$  are selected. Muon candidates must pass the medium criteria of muon identification.

The lepton isolation selection is applied to suppress non-prompt lepton and hadron backgrounds. Isolated leptons are selected using the track isolation discriminant  $p_T^{\text{cone20}} / \min(p_T^\ell, 50 \text{ GeV}) < 0.1$  for both electron and muon candidates, where  $p_T^{\text{cone20}}$  refers to the energy of tracks within a radius of 0.2. This criterion ensures the selected leptons are less likely to have charged tracks nearby.

Prompt leptons originating from the primary vertex must pass selections on the impact parameters as defined in Section 4.1. The longitudinal impact parameter  $z_0$  should satisfy  $z_0 \sin \theta < 0.5 \text{ mm}$  for both electrons and muons. The transverse impact parameter  $d_0$  must meet the criteria  $|d_0|/\sigma(d_0) < 5$  for electrons and  $|d_0|/\sigma(d_0) < 3$  for muons, where  $\sigma(d_0)$  refers to the uncertainty of transverse impact parameters.

The calibration procedure is implemented to adjust the lepton energy, accounting for discrepancies in energy response between simulation and observed data. For electron, the calibration correction is determined using  $Z \rightarrow e^+e^-$  events in the low pile-up dataset, following the methodology outlined in Ref.[75]. For muons, the calibration corrections are derived from high  $\mu$  datasets and subsequently validated with  $Z \rightarrow \mu^+\mu^-$  and  $J/\psi \rightarrow \mu^+\mu^-$  events from low pile-up datasets using the approach described in Ref. [80].

The correction factors for trigger and lepton selection efficiencies are derived from dedicated studies to address discrepancies between simulations and the low pile-up datasets. The lepton trigger efficiencies are specifically measured using the low pile-up datasets. The corrections for electron identification and isolation efficiencies are determined using low pile-up data. Similarly, muon isolation and track-to-vertex-association (TTVA) efficiency corrections are obtained from the low pile-up datasets. While the reconstruction efficiency corrections for both electrons and muons are evaluated using a high  $\mu$  dataset and then extrapolated to conditions of low  $\mu$ .

The neutrino originating from  $W$  boson decay cannot be detected directly. Therefore, the hadronic recoil object is reconstructed in this analysis [162], and the missing transverse energy (MET) is calculated based on this hadronic recoil object to account

for the presence of the neutrino in the event.

The hadronic recoil, denoted as  $\vec{u}_T$ , is defined as the transverse momentum of the hadronic jets or other particles recoiling against the vector boson. By definition, the hadronic recoil can be written as  $\vec{u}_T = -\vec{p}_T^V$  in the transverse plane, where  $\vec{p}_T^V$  represents the transverse momentum of the vector boson. In this measurement, the hadronic recoil is reconstructed using the particle flow objects to take advantage of integrating the information from both the calorimeters and the ID tracks. The calibration of the hadronic recoil is performed primarily using  $Z \rightarrow \ell\ell$  events in the low pile-up datasets and is then applied to  $W^\pm \rightarrow \ell^\pm\nu$  events according to the true transverse momentum  $p_T^{\text{true},V}$ . The calibration procedure begins by correcting the event activity variable  $\Sigma\vec{E}_T$ , which is defined as the difference between  $u_T$  and the scalar sum of the transverse energies of the PFO objects,  $\Sigma\vec{E}_T = \sum E_T$ . The procedure is performed in  $W$  and  $Z$  events, respectively. The reweighting factors as a function of the  $(\Sigma\vec{E}_T, p_T^{\ell\ell})$  are obtained by comparing the distributions of  $Z$  events in MC and data samples. Additional reweighting factors are derived as a function of  $\Sigma\vec{E}_T$  in  $u_T$  bins using  $W$  MC samples. Then the factors are applied on  $W$  events. Consequently, corrections are applied to the simulations to match the data. First, the azimuthal distribution of the hadronic recoil is corrected. Then, the recoil scale and core resolution are adjusted using  $Z$  events. The correction involves the variables  $u_\perp$  and  $u_\parallel$ , which denote the projections of the recoil perpendicular and parallel to the  $Z$  boson direction in the transverse plane, respectively. Finally, a tail correction is applied to address the non-Gaussian effects in the distribution tails.

After deriving the hadronic recoil, one can obtain  $E_T^{\text{miss}}$  as :

$$\vec{E}_T^{\text{miss}} = \vec{p}_T^\nu = -\left(\vec{u}_T + \vec{p}_T^\ell\right).$$

## 6.4 Event selection

Events should pass at least one single-lepton triggers, where lepton refers to electron or muon. Both the 5.02 and 13 TeV datasets use unrescaled single-lepton triggers, HLT\_e15\_lhloose\_nod0\_L1EM12 and HLT\_mu14, with an logical OR applied. The electron triggers select events with at least one reconstructed electron with transverse energy above 15 GeV threshold and satisfying the loose identification criteria. The muon trigger HLT\_mu14 requires the presence of a muon with transverse momentum greater than 14 GeV in events.

After the preselection, event candidates with precisely one lepton satisfying the aforementioned criteria in Section 6.3 are selected. If the event has an additional same-flavor lepton with transverse momentum greater than 20 GeV, it is dropped to reduce contamination from the  $Z \rightarrow \ell\ell$  process. Events from the  $W \rightarrow \ell\nu$  process are characterized by the presence of a neutrino from the  $W$  boson decay, so the missing energy is required to be greater than 25 GeV to select such events and suppress backgrounds, in particular multijet events.

Additionally, to enhance the purity of the  $W$  boson sample, events are subjected to a selection on the transverse mass ( $m_T^W$ ) of the  $W$  boson. This transverse mass, defined as  $\sqrt{2p_T^\nu p_T^\ell (1 - \cos \Delta\phi^\nu)}$ , where  $\Delta\phi^\nu$  is the azimuthal angle between the missing transverse energy vector and the lepton, is required to exceed 50 GeV.

The event selections are summarized in Table 6.1.

TABLE 6.1 – Event selections for  $W$  and  $Z$  boson channels.

Trigger	pass HLT_e15_lhloose_nod0_L1EM12 or HLT_mu14
Electron selections	$p_T > 25$ GeV $ \eta  < 2.47$ and veto of $1.37 <  \eta  < 1.52$
Muon selections	$p_T > 25$ GeV $ \eta  < 2.4$
$W$ -boson selections	Exactly one lepton $E_T^{\text{miss}} > 25$ GeV $m_T > 50$ GeV

## 6.5 Background estimation

The backgrounds arising from the electroweak and top-quark productions are evaluated using the simulated samples described in Section 6.2. The electroweak background consists of single boson productions, including  $W^\pm \rightarrow \tau^\pm \nu$  and  $Z \rightarrow \ell\ell$  ( $\ell$  refers to electrons or muons), as well as diboson productions ( $WW$ ,  $WZ$  and  $ZZ$ ). In the case of diboson events, one of the bosons may decay hadronically or invisibly, or the leptonic final states may fail to meet the selection criteria. The top-quark background includes the  $t\bar{t}$  pair and associated  $Wt$  production with similar final states as  $W$ -boson events.

The QCD multijet (MJ) backgrounds are estimated using the ABCD method detailed in Section 5.5.2. Events are categorized into four regions with regard to the lepton isolation discriminants,  $m_T^W$ ,  $E_T^{\text{miss}}$  and  $p_T^\ell$  observables. The signal region (SR) uses the event selection criteria as described in Section 6.4. The fit region (FR) is defined with the same lepton isolation criteria  $p_T^{\text{cone20}} / \min(p_T^\ell, 50 \text{ GeV}) < 0.1$ , but with relaxed cuts on  $m_T^W$  and  $E_T^{\text{miss}}$ , specifically  $E_T^{\text{miss}} > 0$  GeV and  $m_T^W > 0$  GeV. This definition is different from the strategy used in previous measurements, which employed inverted kinematic cuts (Section 5.5.2). The relaxed cuts ensure a higher proportion of multijet backgrounds in the FR and increase statistics considering the limited statistics of the low pile-up datasets.

The two control regions are defined with the anti-isolated leptons, i.e., the lepton has  $p_T^{\text{cone20}} / \min(p_T^\ell, 50 \text{ GeV}) > 0.1$  and fails the isolation criteria. The control region 1 (CR1) shares the same relaxed cuts on  $E_T^{\text{miss}}$  and  $m_T^W$ , while the control region 2 (CR) is defined with  $E_T^{\text{miss}}$  and  $m_T^W$  cuts as in the SR.

TABLE 6.2 – The definitions of the ABCD regions for the QCD multijet background extraction.

Fit region (FR)	Signal region (SR)
Lepton $p_T > 25$ GeV	Lepton $p_T > 25$ GeV
$E_T^{\text{miss}} > 0$ GeV	$E_T^{\text{miss}} > 25$ GeV
$m_T^W > 0$ GeV	$m_T^W > 50$ GeV
Pass isolation	Pass isolation
Control region 1 (CR1)	Control region 2 (CR2)
Lepton $p_T > 25$ GeV	Lepton $p_T > 25$ GeV
$E_T^{\text{miss}} > 0$ GeV	$E_T^{\text{miss}} > 25$ GeV
$m_T^W > 0$ GeV	$m_T^W > 50$ GeV
Fail isolation	Fail isolation

The definitions of the ABCD regions are summarized in Table 6.2. The determination of the multijet background in the SR is detailed in Section 5.5.2. Here a brief description is given. For a certain discriminant variable, the initial multijet template is extracted from the anti-isolated control region CR1 by subtracting the electroweak and top backgrounds from the observed data distribution. Consequently, a fraction fit is conducted within the FR to derive the multijet template in the isolated region. The transfer factor, which accounts for the impact of  $E_T^{\text{miss}}$  and  $m_T^W$  selections, is defined by the ratio of MJ yields in CR2 to those in the CR1. Finally, the contribution of the MJ background is estimated by applying the transfer factor on the MJ template adjusted in the FR.

The procedure mentioned above are repeated for three different discriminant variables,  $m_T^W$ ,  $E_T^{\text{miss}}$  and  $p_T^\ell$ . The average of the results with these variables are regarded as the final MJ background. Figure 6.2 displays an example of the comparison between data and simulations in the FR.

As the transfer factors from the isolated FR to the SR are estimated using the anti-isolated control regions, there may be potential bias due to the dependence on the lepton isolation criteria. Therefore, in this measurement, the same strategy as described in Section 5.5.2 is employed to mitigate this potential bias : the control regions are divided into four slices with the lepton isolation variable  $\text{ptcone20}/\min(p_T, 50 \text{ GeV})$  starting from 0.1 and ending at 0.5, increment by 0.1 each step. The dependence of transfer factors on the lepton isolation is displayed in Figure 6.3. A linear fit is performed and the final transfer factors are thus determined by extrapolation.

The following uncertainties of the multijet estimation are considered :

- Statistical fluctuations due to the finite number of events in data and MC in the fit region.
- Linear extrapolation procedure to mitigate bias related to lepton isolation.
- Mis-modeling of the jet activity in the anti-isolated control regions.

Table 6.3 summarizes the final number of MJ yields in each channel.

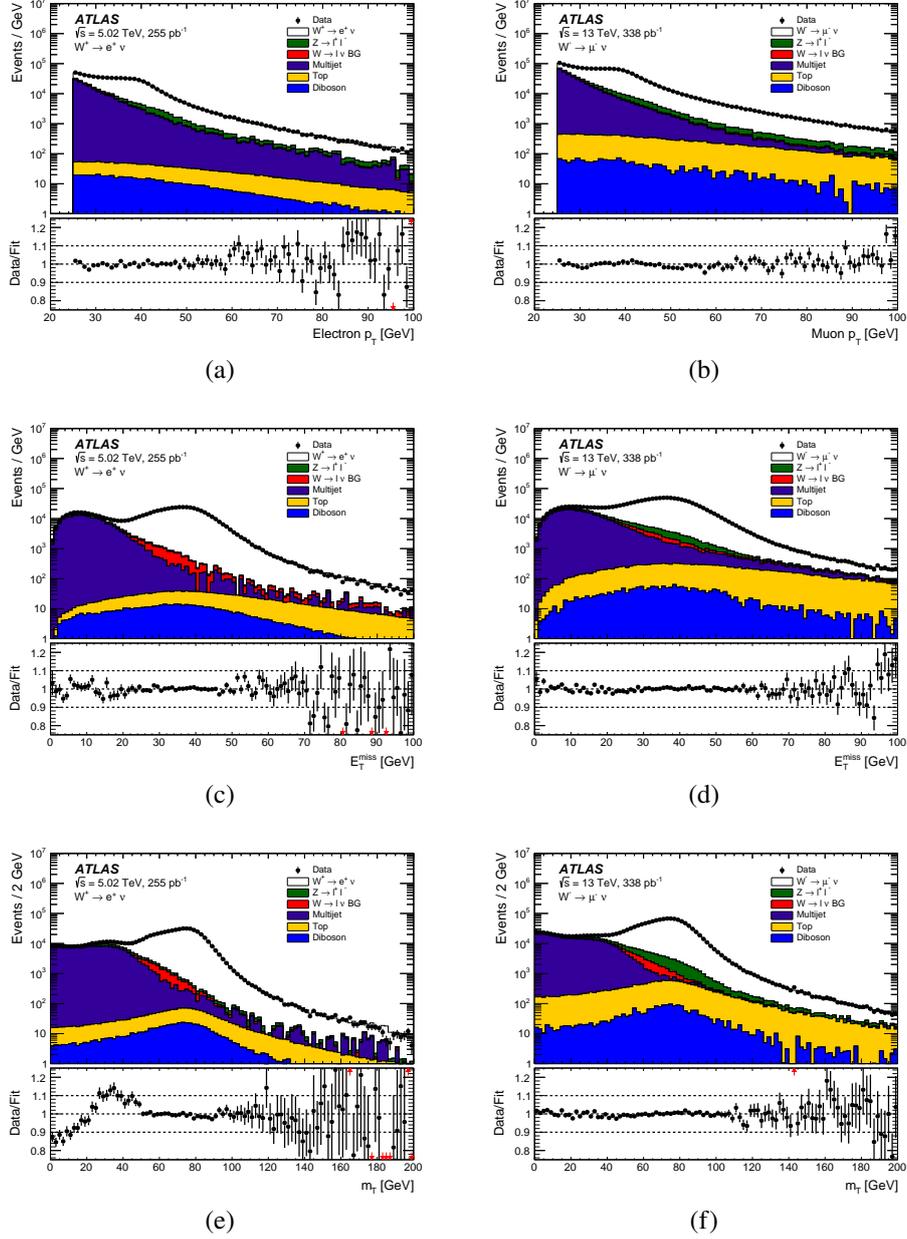


FIGURE 6.2 – A comparison between observed data (dots) and predictions (histograms) for (a,b)  $p_T^l$ , (c,d)  $E_T^{\text{miss}}$  and (e,f)  $m_T^W$  distributions in the fit region. Multijet yields are extracted by the template fits in the fit region for the (a,c,e)  $W^+ \rightarrow e^+ \nu$  events at  $\sqrt{s} = 5.02$  TeV and (b,d,f)  $W^- \rightarrow \mu^- \bar{\nu}$  events at  $\sqrt{s} = 13$  TeV.

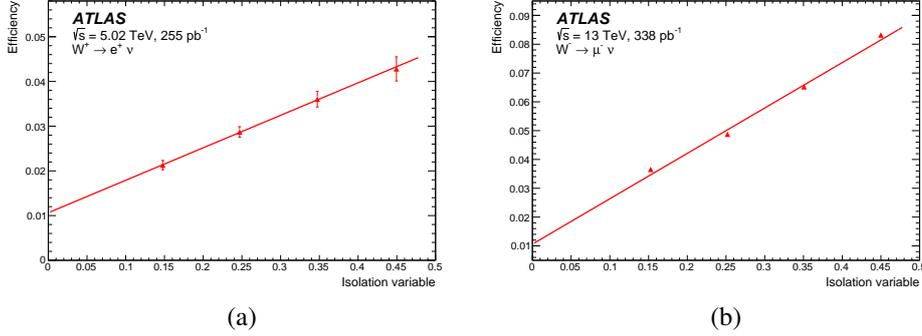


FIGURE 6.3 – The transfer factors used to extend the multijet background from the fit region to the signal region are displayed as a function of the lepton track isolation,  $p_T^{\text{cone}20} / \min(p_T^\ell, 50 \text{ GeV})$ . The red points indicate the lower bounds of each slice of the control regions. The linear fitted extrapolation function is represented by the red line. The errors associated with the points indicate the statistical uncertainty. Examples for (a) the  $W^+ \rightarrow e^+ \nu$  channel at  $\sqrt{s} = 5.02 \text{ TeV}$  and (b) the  $W^- \rightarrow \mu^- \bar{\nu}$  channel at  $\sqrt{s} = 13 \text{ TeV}$  are displayed.

TABLE 6.3 – Event yields of the multijet background in each channel.

Channel	13 TeV Yield	5 TeV Yield
$W^- \rightarrow e^- \nu$	$27,000 \pm 5,000$	$2200 \pm 700$
$W^+ \rightarrow e^+ \nu$	$29,000 \pm 5,000$	$2300 \pm 900$
$W^- \rightarrow \mu^- \nu$	$6,000 \pm 2,000$	$300 \pm 340$
$W^+ \rightarrow \mu^+ \nu$	$6,000 \pm 2,000$	$500 \pm 400$

The multijet templates for the two-dimensional distribution  $(p_T^\ell, |\eta_\ell|)$  are extrapolated from the one-dimensional MJ templates obtained following the procedure described previously. The correlation between  $p_T^\ell$  and  $|\eta_\ell|$  is studied using the MC simulated samples. The JF17 sample for generic  $2 \rightarrow 2$  processes is used for the electron channel, while the  $bb \rightarrow \mu\mu$  and  $cc \rightarrow \mu\mu$  samples are used for the muon channel. To quantify the correlations, 40000 toys are generated to calculate the correlations at 13 TeV. The correlation results are shown in Figure 6.4. Low correlations are observed in both electron and muon channels.

The dependence on  $p_T^\ell$  and  $|\eta_\ell|$  can be visually assessed by comparing the shapes of multijet templates as a function of  $|\eta_\ell|$  across different  $p_T^\ell$  regions. Multijet events are divided into several  $p_T^\ell$  slices, as defined in the double differential cross-section measurement, and the  $|\eta_\ell|$  distributions are renormalized to the same scale for compa-

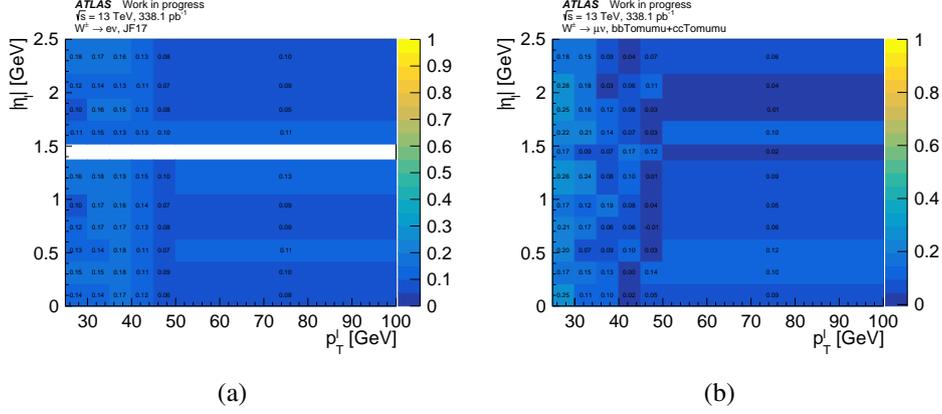


FIGURE 6.4 – Correlation between  $p_T^\ell$  and  $|\eta_\ell|$  of multijet templates estimated using the MC simulated samples at 13 TeV.

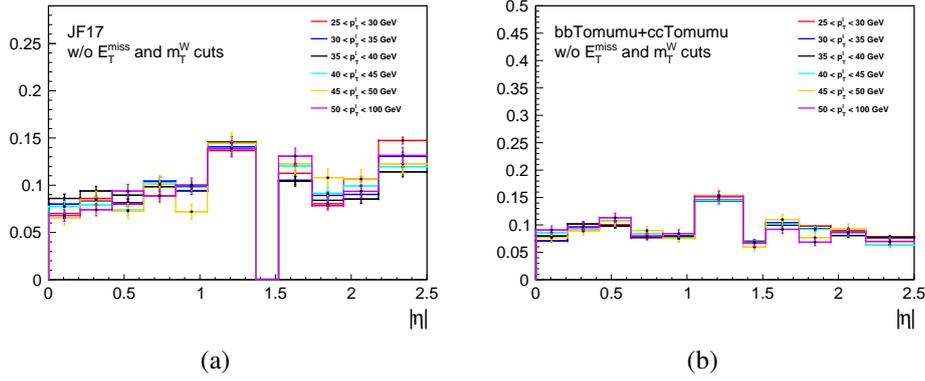


FIGURE 6.5 – The  $|\eta_\ell|$  distribution of multijet templates in various  $p_T^\ell$  regions. The error bars indicate the statistical uncertainties.

risson. However, due to insufficient event yields, one can hardly conclude by comparing shapes after applying the signal selections. Therefore, the kinematic criteria for  $m_T^W$  and  $E_T^{\text{miss}}$  are relaxed. The kinematic selections of the FR defined in Table 6.2 is applied. Figure 6.5 displays the multijet distributions in  $|\eta_\ell|$  in different  $p_T^\ell$  regions, where the error bars indicate the statistical uncertainties. The differences between shapes in various regions fall within the statistical uncertainties.

Given the fact that multijet shape does not have significant dependence on the  $p_T^\ell$  and  $|\eta_\ell|$ , the multijet background template in  $p_T^\ell$  and  $|\eta_\ell|$  are extended from the multijet shape in  $|\eta_\ell|$ . The fiducial phase space are split into several  $p_T^\ell$  slices, the total event yields  $N_{\text{total}}$  of multijet is compared to the event yields in a certain  $p_T^\ell$  slice  $i$ ,  $N_i$ . The resulting multijet template of  $p_T^\ell$  slice  $i$  is  $f(|\eta_\ell|)_i = N_{\text{total}}/N_i f(|\eta_\ell|)$ , where the  $f(|\eta_\ell|)$  is the multijet template in  $|\eta_\ell|$ .

## 6.6 Kinematic distributions

The event yields for each process after the selections are summarized in Tables 6.4 and 6.5. The comparison between data and simulations in various kinematic distributions for  $W$ -boson events is presented in Sections 6.6.1 and 6.6.2, corresponding to 5.02 and 13 TeV, respectively. The observables  $|\eta_\ell|$ ,  $p_T^\ell$ , and  $|\eta_\ell|$  in each  $p_T^\ell$  region are presented, with the differential cross sections measured as functions of these variables.

In Figures 6.6-6.13, the upper panel displays the event yields of data and predictions, which the lower panel shows the data over MC ratio. The ratio panels include two uncertainty bands : the inner (grey) band represents the total systematic uncertainty excluding the luminosity contribution, and the outer (brown) band includes the luminosity uncertainty added in quadrature.

The systematic uncertainties, which are elaborated in Section 6.9, cover a range of factors including lepton trigger corrections, lepton efficiency corrections, hadronic recoil corrections, multijet background uncertainties, and uncertainties in the modeling of signal and background cross sections.

Overall, good agreement between data and predictions is observed. Some differences appear at high  $p_T^\ell$ , but most of these can be accounted for by the uncertainties and there is in general a good consistency between the electron and muon channels.

TABLE 6.4 – The number of data, signal and background (BG) events after selections for different channels at 5.02 TeV. For the multijet backgrounds, the total uncertainty is quoted, whereas for other predictions, only the statistical uncertainty is quoted.

Channel	Data	Signal	$W \rightarrow \ell\nu$ (BG)	$Z \rightarrow \ell\ell$	Diboson	Top	Multijet
$W^- \rightarrow e\nu$ , 5.02 TeV	274375	264500±200	7300±30	1030±6	340±3	863±2	2200 ± 700
$W^- \rightarrow \mu\nu$ , 5.02 TeV	288026	273900±200	5160±20	8520±20	339±2	799±2	300 ± 340
$W^+ \rightarrow e\nu$ , 5.02 TeV	430662	417100±200	9060±30	1102±7	376±2	951±3	2300 ± 900
$W^+ \rightarrow \mu\nu$ , 5.02 TeV	457223	438000±200	7850±30	9770±20	385±2	891±3	500 ± 400

TABLE 6.5 – The number of data, signal and background (BG) events after selections for different channels at 13 TeV. For the multijet backgrounds, the total uncertainty is quoted, whereas for other predictions, only the statistical uncertainty is quoted.

Channel	Data	Signal	$W \rightarrow \ell\nu$ (BG)	$Z \rightarrow \ell\ell$	Diboson	Top	Multijet
$W^- \rightarrow e\nu$ , 13 TeV	949297	876500±300	22360±60	6550±20	1710±40	12780±50	27000 ± 5000
$W^- \rightarrow \mu\nu$ , 13 TeV	964514	897400±300	14650±50	33940±40	1620±40	11950±50	6000 ± 2000
$W^+ \rightarrow e\nu$ , 13 TeV	1207652	1116800±300	24680±60	6880±20	1720±50	13260±50	29000 ± 5000
$W^+ \rightarrow \mu\nu$ , 13 TeV	1245755	1153400±300	18390±50	37690±40	1720±40	12520±50	6000 ± 2000

### 6.6.1 Kinematic distributions at 5.02 TeV

This section presents the  $|\eta_\ell|$ ,  $p_T^\ell$  and  $(p_T^\ell, |\eta_\ell|)$  distributions for data collected at 5.02 TeV.

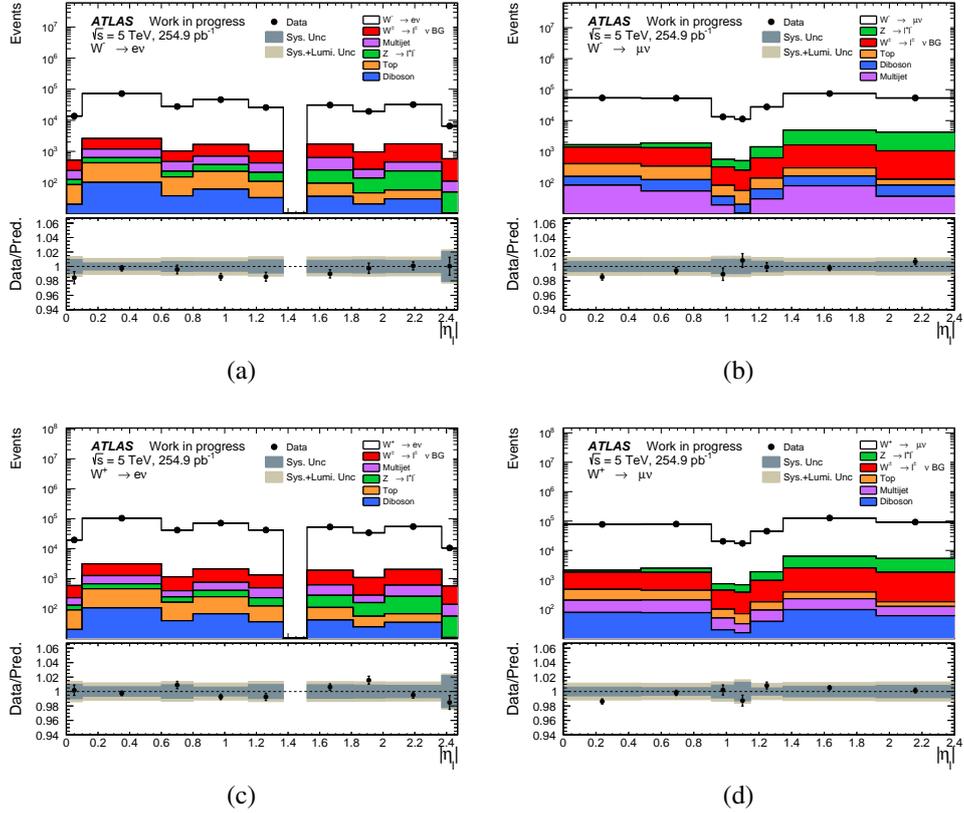


FIGURE 6.6 –  $|\eta_\ell|$  distributions for the (a)  $W^- \rightarrow e^- \bar{\nu}$ , (b)  $W^- \rightarrow \mu^- \bar{\nu}$ , (c)  $W^+ \rightarrow e^+ \nu$ , and (d)  $W^+ \rightarrow \mu^+ \nu$  channels at 5.02 TeV, showing the signal and various background contributions. The bottom panel displays the ratio of data over simulation, and the uncertainty band encompasses all sources of uncertainty (statistical and systematic). The light gray error band represents the total uncertainty, while the dark gray error band indicates the total uncertainty excluding the luminosity uncertainty.

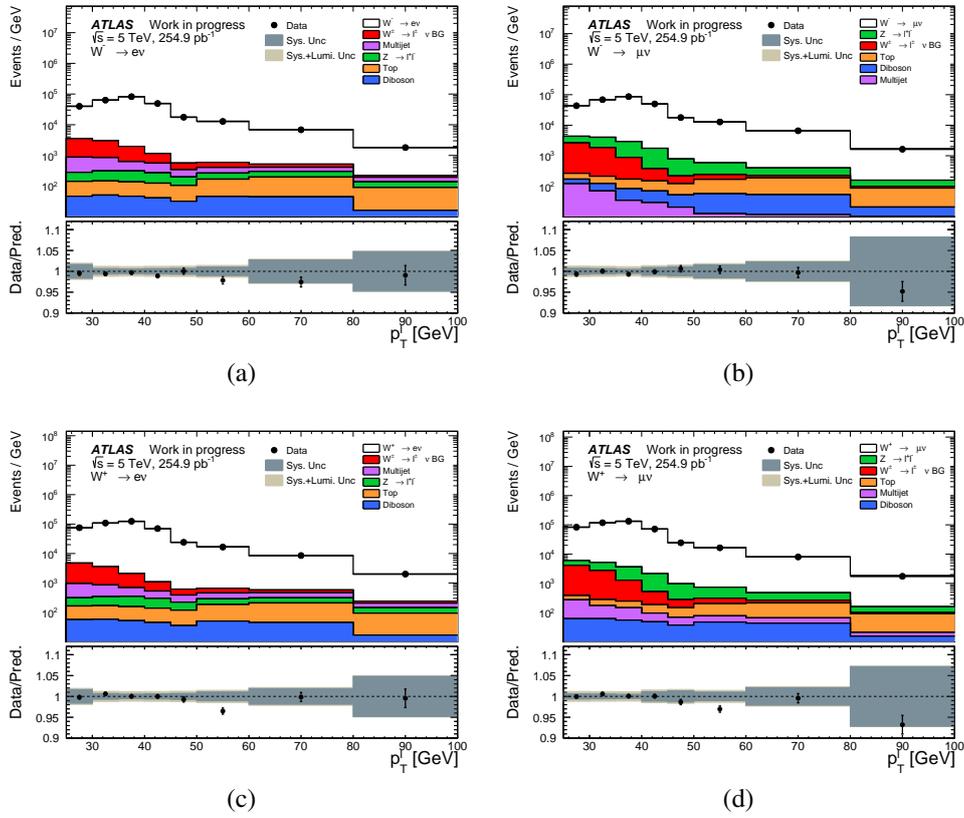
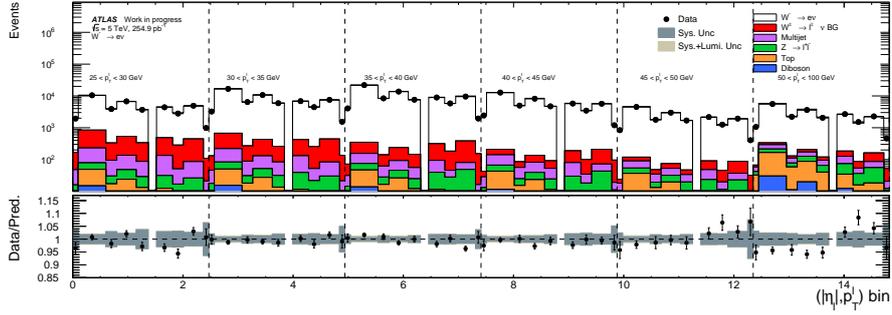


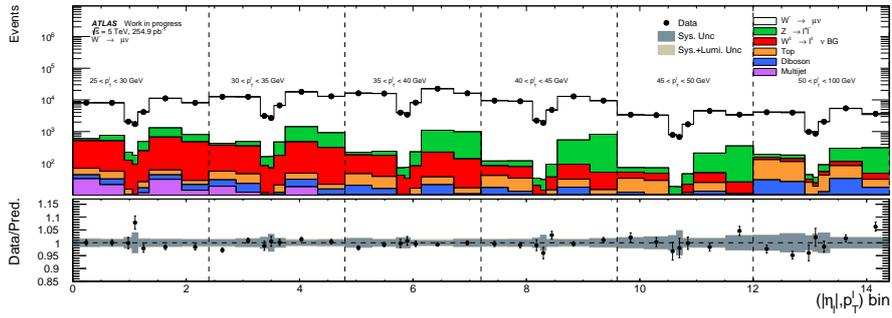
FIGURE 6.7 –  $p_T^\ell$  distributions for the (a)  $W^- \rightarrow e^- \bar{\nu}$ , (b)  $W^- \rightarrow \mu^- \bar{\nu}$ , (c)  $W^+ \rightarrow e^+ \nu$ , and (d)  $W^+ \rightarrow \mu^+ \nu$  channels at 5.02 TeV, showing the signal and various background contributions. The bottom panel displays the ratio of data over simulation, and the uncertainty band encompasses all sources of uncertainty (statistical and systematic). The light gray error band represents the total uncertainty, while the dark gray error band indicates the total uncertainty excluding the luminosity uncertainty.

## 6.6.2 Kinematic distributions at 13 TeV

This section presents the  $|\eta_\ell|$ ,  $p_T^\ell$  and  $(p_T^\ell, |\eta_\ell|)$  distributions for data collected at 5.02 TeV.

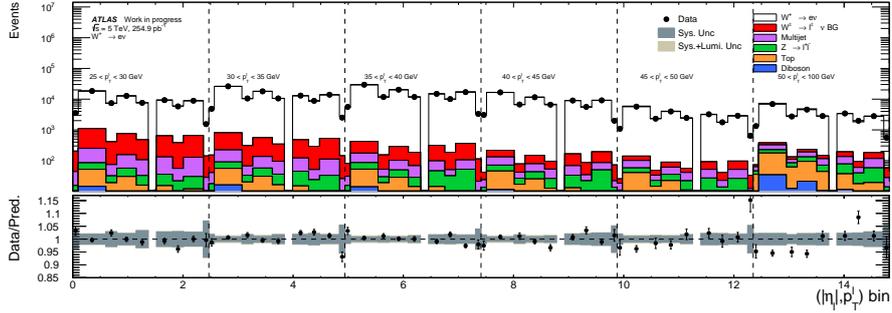


(a)

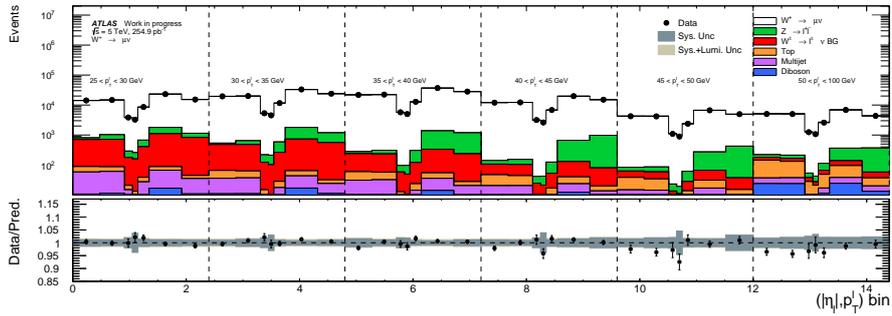


(b)

FIGURE 6.8 –  $\eta^\ell$  distributions for the (a)  $W^- \rightarrow e^- \bar{\nu}$  and (b)  $W^- \rightarrow \mu^- \bar{\nu}$  channels at 5.02 TeV, as a function of various ranges of  $p_T^\ell$ , showing the signal and various background contributions. The bottom panel displays the ratio of data over simulation, and the uncertainty band encompasses all sources of uncertainty (statistical and systematic). The light gray error band represents the total uncertainty, while the dark gray error band indicates the total uncertainty excluding the luminosity uncertainty.



(a)



(b)

FIGURE 6.9 –  $\eta^\ell$  distributions for the (a)  $W^+ \rightarrow e^+\nu$ , and (b)  $W^+ \rightarrow \mu^+\nu$  channels at 5.02 TeV, as a function of various ranges of  $p_T^\ell$ , showing the signal and various background contributions. The bottom panel displays the ratio of data over simulation, and the uncertainty band encompasses all sources of uncertainty (statistical and systematic). The light gray error band represents the total uncertainty, while the dark gray error band indicates the total uncertainty excluding the luminosity uncertainty.

## 6.7 Theory predictions

The section describes the theoretical predictions for the  $W \rightarrow \ell\nu$  processes. Most of the technical details are the same as the theoretical predictions reported in Section 5.6.

The predicted cross sections was computed up to NNLO+NNLL QCD and LO electroweak accuracy. DYTurbo is used to perform the calculations [134–137].

The  $G_\mu$  scheme [140] is chosen to compute the predictions, as it offers highly accurate and reliable theoretical predictions, and five-flavor scheme was used for the QCD calculation. The electroweak input parameters are presented in Table 5.9 and the CKM matrix elements are reported in Eq. (5.6).

The predictions are calculated with various PDFs sets, including : CT18 [34], MSHT20 [35–37], NNPDF3.1 [163], NNPDF4.0 [38] and ATLASpdf21 [39]. The evaluation of the PDF uncertainties follows the description in Section 5.6.2. NNPDF3.1

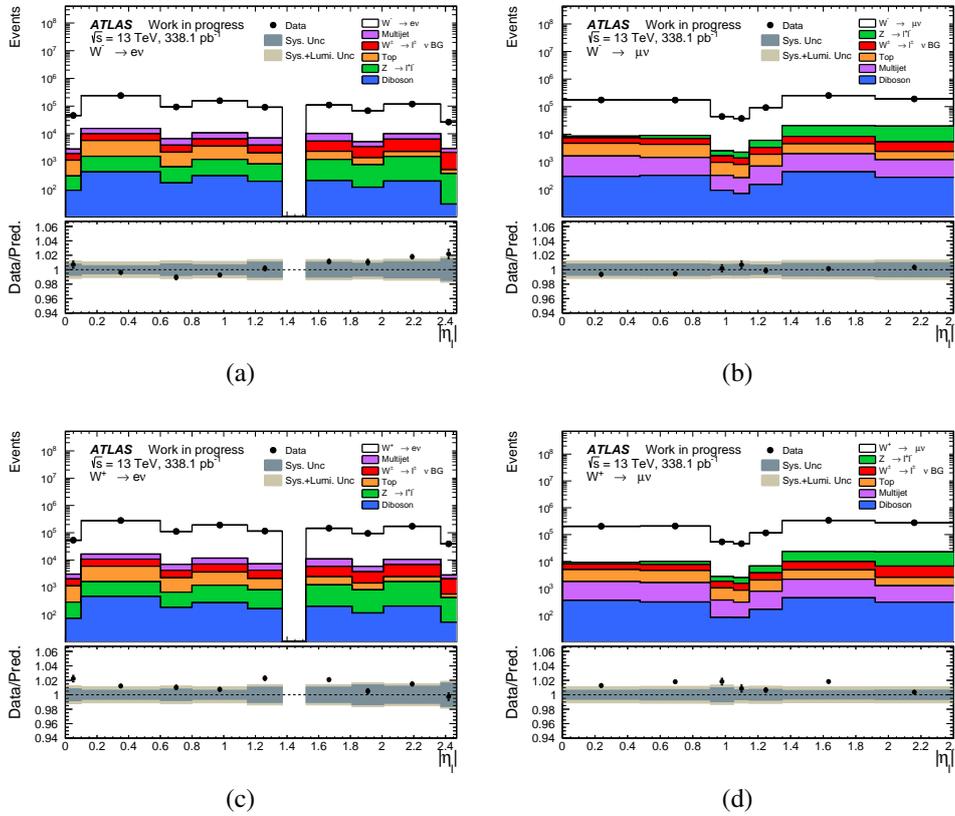


FIGURE 6.10 –  $\eta^\ell$  distributions for the (a)  $W^- \rightarrow e^- \bar{\nu}$ , (b)  $W^- \rightarrow \mu^- \bar{\nu}$ , (c)  $W^+ \rightarrow e^+ \nu$ , and (d)  $W^+ \rightarrow \mu^+ \nu$  channels at 13 TeV, showing the signal and various background contributions. The bottom panel displays the ratio of data over simulation, and the uncertainty band encompasses all sources of uncertainty (statistical and systematic). The light gray error band represents the total uncertainty, while the dark gray error band indicates the total uncertainty excluding the luminosity uncertainty.

is not used in the inclusive vector boson measurement in Chapter 5. As it is also the release of NNPDF family sets, the PDF uncertainty can be assessed from its 100 MC replicas using Eq. (5.9).

The uncertainties arise from statistical fluctuation, QCD scale variations and uncertainties related to PDFs. The QCD scale uncertainties are estimated using the standard seven-point variations, by varying the  $\mu_R$  and  $\mu_F$  by a factor of 2. The envelope of the resulting variations are considered the QCD scale uncertainties.

The differential cross sections are calculated with a fiducial volume defined in Table 6.6. The Born level leptons are used to define the fiducial phase space.

## 6.8 Unfolding

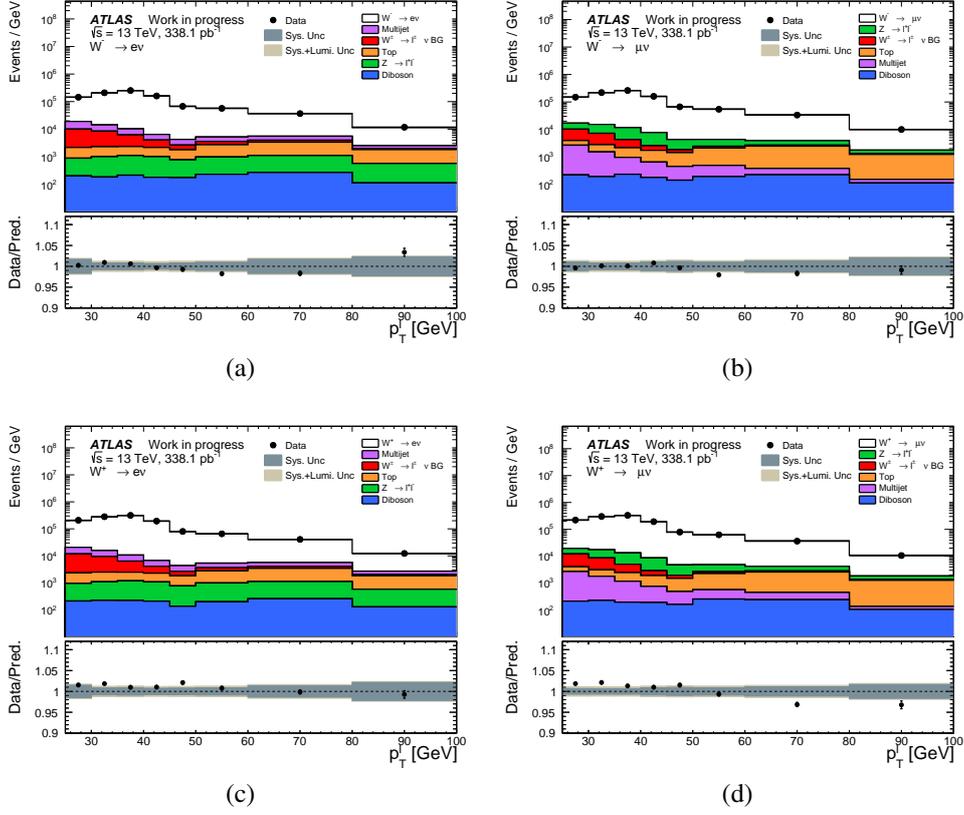
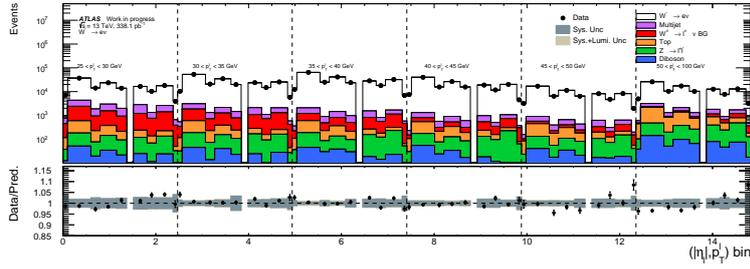


FIGURE 6.11 –  $p_T^{\ell}$  distributions for the (a)  $W^- \rightarrow e^- \bar{\nu}$ , (b)  $W^- \rightarrow \mu^- \bar{\nu}$ , (c)  $W^+ \rightarrow e^+ \nu$ , and (d)  $W^+ \rightarrow \mu^+ \nu$  channels at 13 TeV, showing The bottom panel displays the ratio of data over simulation, and the uncertainty band encompasses all sources of uncertainty (statistical and systematic). The light gray error band represents the total uncertainty, while the dark gray error band indicates the total uncertainty excluding the luminosity uncertainty.

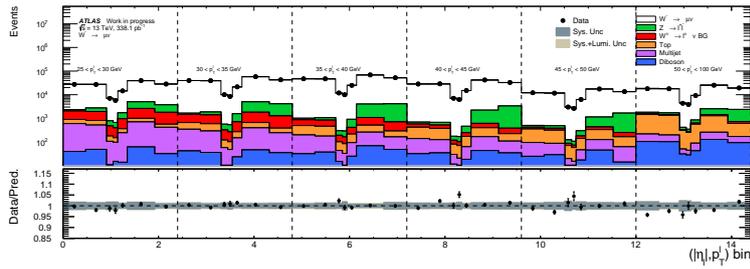
TABLE 6.6 – Fiducial volume for  $W^{\pm} \rightarrow \ell \nu$  processes. The lepton kinematics are computed at the Born level, accounting for QED final state radiation. The  $E_T^{\text{miss}}$  is calculated using the neutrinos produced from  $W$ -boson decays.

	$pp \rightarrow \ell^- \bar{\nu}_{\ell} / pp \rightarrow \ell^+ \nu_{\ell}$
Lepton $p_T$ cuts	$p_T > 25 \text{ GeV}$
Lepton $\eta$ cuts	$ \eta  < 2.5$
Mass cuts	$m_T^W > 50 \text{ GeV}$
Neutrino cuts	$E_T^{\text{miss}} > 25 \text{ GeV}$

This section details the unfolding procedure used in the measurement. The detector has limited acceptance, which means it cannot detect particles outside its coverage. Additionally, its resolution, or ability to distinguish between different particle types,



(a)



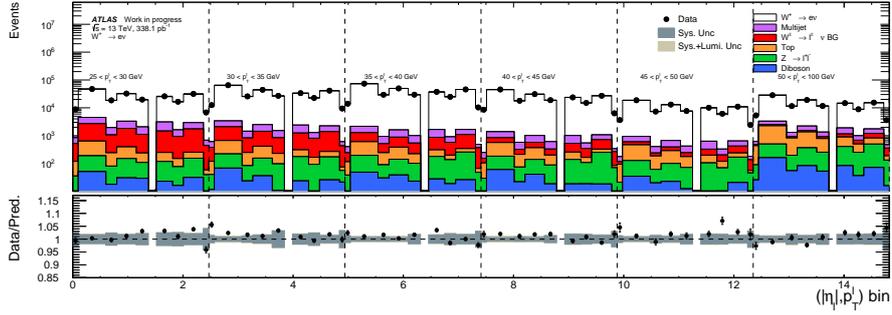
(b)

FIGURE 6.12 –  $\eta^\ell$  distributions for the (a)  $W^- \rightarrow e^- \bar{\nu}$  and (b)  $W^- \rightarrow \mu^- \bar{\nu}$  channels at 13 TeV, as a function of various ranges of  $p_T^\ell$ , showing the signal and various background contributions. The bottom panel displays the ratio of data over simulation, and the uncertainty band encompasses all sources of uncertainty (statistical and systematic). The light gray error band represents the total uncertainty, while the dark gray error band indicates the total uncertainty excluding the luminosity uncertainty.

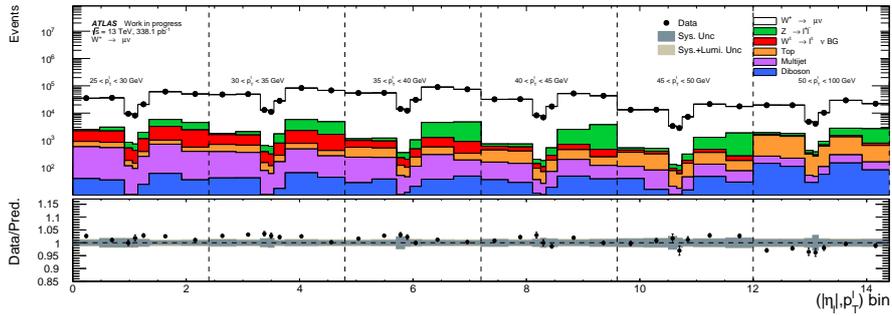
is restricted. This limitation can complicate the accurate reconstruction and identification of detected particles. Consequently, imperfections in coverage, resolution, reconstruction, and identification can cause discrepancies between measured distributions and theoretical predictions, even when both are based on the same physical processes.

To address these issues in this analysis, the unfolding technique is employed. Unfolding is a statistical method designed to correct for measurement bias and resolution effects in the data. This technique involves using a mathematical model to infer the true underlying distribution from the measured distribution, considering the detector's known limitations and uncertainties. By applying this method, the resulting unfolded distribution is expected to align more closely with theoretical predictions, enabling more accurate comparisons and interpretations of the data.

Section 6.8.1 describes the general unfolding procedure applied in this measurement. Section 6.8.2 describes the binning of observables, then displays the migration matrix. Section 6.8.3 details the unfolding bias evaluation. The study of the unfolding optimization is reported in Section 6.8.4.



(a)



(b)

FIGURE 6.13 –  $\eta^\ell$  distributions for the (a)  $W^+ \rightarrow e^+\nu$ , and (b)  $W^+ \rightarrow \mu^+\nu$  channels at 13 TeV, as a function of various ranges of  $p_T^\ell$ , showing the signal and various background contributions. The bottom panel displays the ratio of data over simulation, and the uncertainty band encompasses all sources of uncertainty (statistical and systematic). The light gray error band represents the total uncertainty, while the dark gray error band indicates the total uncertainty excluding the luminosity uncertainty.

### 6.8.1 Unfolding procedure

The physical quantities measured by the detector are distorted by various detector effects, such as resolution, acceptance, and efficiency. It is essential to extract the true distribution to enable accurate comparisons with theoretical predictions and results from different experiments. Thus, it creates an inverse problem : extracting the true distribution from the measured distribution. The unfolding technique [164] is employed to solve this problem, revealing the underlying true distribution by transforming the observed distribution.

Given a true distribution  $f(t)$ , the measured distribution  $g(s)$  can be expressed as :

$$g(s) = \int K(s, t) f(t) dt, \quad (6.1)$$

where  $K(s, t)$  is the kernel function reflecting the transformation of the true distribution into the measured distribution. Limited by the detector resolution and the finite

number of events, the observed distribution is described by histograms. Therefore, the kernel function  $K(s, t)$  is discretized, and the discrete form is represented as the response matrix  $R$ , addressing the bin-to-bin migration.

The continuous function  $f(t)$  and  $g(s)$  are converted into discrete bins, which can be expressed by vectors  $\mathbf{m}$  and  $\mathbf{t}$  for detector and particle level distributions, respectively. The  $i$ -th element in the vector  $\mathbf{m}$  or  $\mathbf{t}$  represents the number of events in bin  $i$ . After the discretization, Eq. (6.1) can be rewritten as :

$$\mathbf{R}\mathbf{t} = \mathbf{m}. \quad (6.2)$$

The response matrix  $R$  can be derived from the MC simulations with full detector effects incorporated. This discrete form enables numerical calculations to solve for the true distribution. If  $\mathbf{d}$  is measured in  $M$  bins while  $\mathbf{t}$  has  $N$  bins, then  $R$  can be presented as an  $M \times N$  matrix. The element  $R_{ij}$  denotes the probability of an event in bin  $j$  of the true distribution appearing in bin  $i$  of the observed distribution,  $R_{ij} = P(m_i|t_j)$ . Equation (6.2) can be written in terms of bins as

$$m_i = \sum_j R_{ij}t_j. \quad (6.3)$$

The problem is to derive the conditional probability  $P(t_j|m_i)$ , representing the probability of an event in bin  $i$  of the measured distribution being observed in bin  $j$  of the true distribution. As the response matrix  $\mathbf{R}$  is often not invertible, this problem cannot be solved by a simple inversion. The iterative Bayesian unfolding (IBU) [164, 165] method is employed in this measurement to provide the solution. Given an initial prior true distribution  $t^0$ , according to Bayes' theorem, one can obtain the conditional probability :

$$P(t_j|m_i) = \frac{P(m_i|t_j)t_j^0}{\sum_k P(m_i|t_k)t_k^0} \quad (6.4)$$

$$= \frac{R_{ij}t_j^0}{\sum_k R_{ik}t_k^0}. \quad (6.5)$$

The estimate of event yields in true bin  $j$ ,  $\hat{t}_j$ , can be obtained as :

$$\hat{t}_j = \sum_i P(t_j|m_i)m_i = t_j^0 \sum_i \frac{R_{ij}m_i}{\sum_k R_{ik}t_k^0}. \quad (6.6)$$

In this way, one can update the true distribution iteratively. For the  $n$ -th iteration, the updated true distribution  $t_j^n$  is written as :

$$t_j^n = t_j^{n-1} \sum_i \frac{R_{ij}m_i}{\sum_k R_{ik}t_k^{n-1}}. \quad (6.7)$$

This procedure is repeated until the true distribution converges.

In practice, several corrections need to be considered. In the measured distribution  $D$ , there are events contributed by backgrounds. These background events  $B$  should be removed from the observed distribution, resulting in the corrected measured distribution  $D - B$ . Consequently, Eq. (6.2) can be represented as :

$$D_i = \sum_j R_{ij} T_j + B_i. \quad (6.8)$$

Another correction after the background subtraction arises from the detection inefficiency. The purity corrections are presented as the ratio of events passing both detector and particle level selections to events passing only the detector level selections in each bin. These bin-by-bin purity corrections are estimated using signal MC simulations and are applied before unfolding, the corrected event yields in bin  $i$  can be written as  $f_i(D_i - B_i)$ , where  $f_i$  denotes the correction factor in bin  $i$ . Finally, after the unfolding procedure, the efficiency corrections addressing the detector imperfect efficiency is considered, which is represented as the ratio of events yields passing both the detector and particle level selections to event yields passing only the particle level selection in each bin. The unfolded distribution should be divided by these efficiency correction factors to match the truth distribution.

The unfolded distribution can be expressed as :

$$\tilde{U}_i = \sum_j U_{ij} f_j (D_j - B_j) \quad (6.9)$$

where  $\tilde{U}$  is the unfolded distribution, giving an estimate of the true distribution,  $f_j$  denotes the purity corrections, and  $U_{ij}$  is the unfolding matrix obtained from the IBU procedure.

The iterative Bayesian unfolding method in this analysis is implemented in the ROOUNFOLD software package [166]. The number of iterations is optimized considering the unfolding bias and statistical uncertainty, which is detailed in Section 6.8.4.

### 6.8.2 Binning

The binning choice for unfolding has been studied for the following considerations.

For the transverse momentum  $p_T^\ell$ , the detector resolution effects as a function of  $p_T^\ell$  are studied using the nominal POWHEG MC samples. Events from  $W$  boson production are selected according to the criteria defined in Section 6.4. The reconstructed lepton candidate is then matched to the generated one using the  $p_T^{\text{true}}$  based on the born definition. The resolution is defined as the RMS value of the difference between the reconstructed and true  $p_T^\ell$ .

Figure 6.14 shows the resolution as a function of  $p_T^{\text{true}}$ , using the 13 TeV sample as an example. It has been verified that the resolution's dependence on  $p_T^\ell$  is consistent regardless of the  $W$  boson's charge and the centre-of-mass energies.

The impact on the sensitivity to the  $W$  mass with different lepton  $p_T$  binning is also investigated, as this variable is crucial in the  $W$  boson mass determination. The effect of lepton efficiency corrections on the  $p_T^\ell$  distribution is examined as well.

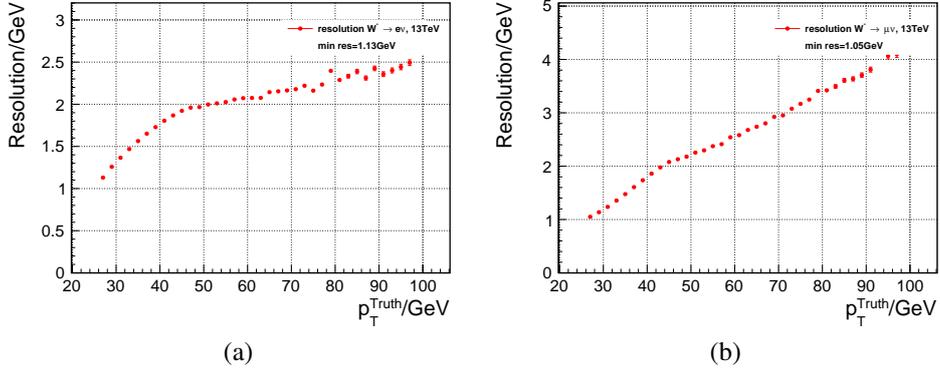


FIGURE 6.14 – Resolution (RMS) as a function of  $p_T^{\text{true}}$  for (a) electrons and (b) muons from  $W$  boson decays selected from POWHEG MC samples at 13 TeV.

For  $\eta^\ell$ , the main challenge arises from the fact that different binnings are used for electrons and muons, due to different scale factor binnings. To allow an easy comparison and combination of the electron and muon channels, a common binning is employed at the unfolded level, which is defined in previous measurement in ATLAS [96].

The final binning used this measurement is given by :

$p_T^\ell$  **bin boundaries** : [25, 30, 35, 40, 45, 50, 60, 80, 100] GeV

$\eta^{\text{electron}}$ -**reconstructed bin boundaries** : [0, 0.1, 0.6, 0.8, 1.15, 1.37, 1.52, 1.81, 2.01, 2.37, 2.47]

$\eta^{\text{muon}}$ -**reconstructed bin boundaries** : [0, 0.476, 0.908, 1.05, 1.1479, 1.348, 1.918, 2.4]

$\eta^\ell$ -**particle level bin boundaries** : [0.0, 0.21, 0.42, 0.63, 0.84, 1.05, 1.37, 1.52, 1.74, 1.95, 2.18, 2.5].

For the double differential cross-section measurement, the width of the  $p_T^\ell$  binning at large  $p_T^\ell$  are more coarse than the binning mentioned above to increase the statistics. Therefore, the bin boundaries of  $p_T^\ell$  for the double differential cross sections are : [25, 30, 35, 40, 45, 50, 100] GeV. While the  $|\eta_\ell|$  binning remains the same as the one dimensional case.

With these defined binnings, the migration effects can be visualized with the corresponding migration matrices, as shown in Figure 6.15. The  $x$ -axis stands for variables at the reconstruction level, while  $y$ -axis at the truth level. The non-diagonal elements are relatively small due to the minor differences between variables at the two levels.

### 6.8.3 Bias estimation

The bias (prior hypothesis bias) of the iterative unfolding method applied to the data is considered and added to the result as a systematic uncertainty.

Since it is impossible to know the true distribution corresponding to the data distribution at the detector level, a data-driven method is used to estimate the unfolding bias. The general idea is [167] :

- As the MC distribution cannot perfectly describe the data, adjustments to the MC are necessary to achieve better agreement between them. A smooth function is used to reweight the true MC distribution, determined by fitting the ratio of background-subtracted data to signal MC distributions at the detector level. After reweighting, the corresponding MC spectra at the reconstructed level should better match the data distribution after removing the background components.
- The reweighted data-like MC distribution serves as the pseudo-data. An unfolding procedure is then performed to unfold the reweighted MC distribution using the original, non-reweighted migration matrix.
- The resulting unfolded distribution is then compared with the reweighted MC distribution at the truth level. The differences between these two spectra are considered as the bias arising from the unfolding procedure.

In this measurement, a cubic polynomial function is utilized to fitting the ratio of background-subtracted data to signal MC distributions at the reconstructed level, which is formulated as :

$$f(x) = A + Bx + Cx^2 + Dx^3 . \quad (6.10)$$

The reweighting factors for each bin  $i$  are obtained by evaluating  $f(x)$  at the center of bin  $i$ , denoted as  $x$ . Figure 6.16 shows examples of the reweighing factors for different observables, as well as the data over MC ratio before and after the reweighing procedure.

#### 6.8.4 Unfolding optimization

The number of iterations for the Bayesian unfolding impacts both the statistical uncertainty and the unfolding bias. As the number of iteration increases, the unfolding bias is anticipated to decrease since it lessens the reliance on the initial prior. However, a greater number of iterations also increases the statistical error and the correlation between adjacent bins, due to reduced regularisation. A trade-off between these error terms can be used to identify the optimal number of iterations.

The standard procedure described in Section 6.8.3 is used to obtain the unfolding bias. The statistical uncertainty is estimated using bootstrap toys generated from the data observables. The Bayesian unfolding procedure is applied on the toys and the variance of the unfolded distribution is taken as the statistical uncertainty. In this measurement, 4000 toys are generated from the data distribution using the resampling method described in Section 6.9.2.

Figure 6.17 shows examples of the unfolding bias with different numbers of iterations. Some spikes are observed in the bias figures, specifically in the bins corresponding to the barrel-endcap transition region, where electron candidates are vetoed

at the reconstruction level. Figure 6.18 displays the impact on the statistical uncertainty. The combination of the unfolding bias and statistical uncertainty are presented in Figure 6.19. It turns out to be that the unfolding bias is smaller than the statistical uncertainty and it is almost unchanged after several iterations. The final iteration numbers are defined accordingly : two iterations are chosen for  $p_T^\ell, |\eta_\ell|$  and the double differential case as the baseline.

## 6.9 Uncertainty

### 6.9.1 Uncertainty sources

The systematic uncertainties in this analysis include those arising from the detector and those associated with modeling. The uncertainty sources related to the detector are categorised into several classes and listed in the following :

**Luminosity :** The uncertainty on the integrated luminosity is 1% at 5.02 TeV and 0.9% at 13 TeV, based on the LUCID-2 detector [144] estimate, which is used for the primary luminosity measurements. Luminosity serves as the dominant source of uncertainty in the cross-section measurement.

**Lepton trigger efficiency :** The trigger efficiency uncertainties are determined by comparing the trigger efficiency results from data and simulations, incorporating various independent systematic variations. These uncertainties are measured using low pile-up datasets from both 5.02 and 13 TeV runs.

**Electron efficiency correction :** The uncertainties arising from electron reconstruction, identification, and isolation efficiency corrections are considered. The correction factors and corresponding uncertainties of identification and isolation efficiencies are evaluated using the tag-and-probe approach with  $Z \rightarrow ee$  events in low pile-up datasets based on the strategy in Ref. [145]. For electron reconstruction, the corrections are determined using a dataset with high pile-up conditions and extrapolated to the low pile-up condition.

**Electron calibration :** The electron energy calibration corrections and uncertainties are estimated through a combination of simulation-based and data-driven procedures using  $Z \rightarrow ee$  events in the low pile-up dataset based on the strategy described in Ref. [75].

**Muon efficiency correction :** The muon reconstruction, isolation, and track-to-vertex-association (TTVA) efficiency correction uncertainties are taken into account. The uncertainties are estimated using tag-and-probe methods based on Ref. [80] with  $Z \rightarrow \mu\mu$  events. The isolation and TTVA uncertainties are derived from low pile-up datasets, whereas the reconstruction uncertainties are extrapolated from high pile-up datasets.

**Muon calibration :** The muon calibration is derived from high- $\mu$  data following the methodology outlined in Ref. [80] and cross-checked using  $Z \rightarrow \mu\mu$  and  $J/\psi \rightarrow \mu\mu$  events in the low pile-up datasets.

**Hadronic recoil :** The hadronic recoil calibration is obtained as functions of  $\Sigma \bar{E}_T$  and  $p_T^V$ , where  $V$  represents vector bosons. The primary uncertainty comes from the corrections applied to  $Z$ -boson events, while the discrepancies between data and MC in the  $\Sigma \bar{E}_T$  distribution contribute a lesser amount of uncertainty.

**Multijet modelling :** The multijet uncertainty consists of two components : statistical uncertainty and methodological variations, as described in Section 6.5.

**Unfolding bias :** The unfolding bias, accounting for the dependence on the initial prior when performing the unfolding procedure to correct the detector effects, is described in Section 6.8.3.

Modeling uncertainties are comprised of two elements : the uncertainty in the signal modeling, which stems from the selection of the signal simulation generator, and the uncertainty associated with the computation of electroweak and top production background cross sections.

The generator uncertainty is handled by comparing the results from the nominal POWHEG signal sample with those from the alternative SHERPA simulation sample, as detailed in Section 6.2. The Sherpa samples are corrected to match the nominal POWHEG sample. Corrections on various observables, including  $p_T$ , rapidity, vertex efficiency, and recoil calibration, are applied. For propagation to the particle level, The corrected SHERPA samples serve as the pseudo data, and the standard unfolding procedures are performed on these datasets using the purity corrections, efficiency corrections, and migration matrices from the POWHEG samples.

The uncertainty arising from the background modeling is evaluated using the same method described in Section 5.7.2. Considering the minor contribution from each process, conservative uncertainties are assigned to each one. For the  $W$  and  $Z$  boson background processes, uncertainties related to the PDF (the internal uncertainty of PDF set and the choice the PDF), the strong coupling constant  $\alpha_s$ , and the QCD scale are incorporated, resulting in an overall uncertainty of 5%. The modelling uncertainty of the diboson process is set to be 10%. The modeling uncertainties for  $t\bar{t}$  and single-top processes are estimated to be 7% and 10%, respectively. These estimates are derived from theoretical predictions by varying the radiation parameters up and down, accounting for the uncertainty due to missing higher-order QCD corrections.

## 6.9.2 Uncertainty propagation

Since the unfolding technique is used to correct detector effects and the observed data are presented as true distributions at the particle level to derive the final results, it is crucial to propagate uncertainties from the detector level to the particle level. This ensures that the uncertainty of the unfolded distributions is accurately accounted for.

In general, the uncertainty due to statistical fluctuations is propagated using bootstrap toys. Systematic uncertainties are evaluated by repeating the unfolding procedure with systematic variations and taking the differences between the resulting unfolded distributions and the nominal distribution.

## Statistical uncertainty

The statistical uncertainty, due to the statistical fluctuation of the finite size of the dataset and simulated samples, are evaluated and propagated to unfolded distributions with the help of the bootstrap method [168].

Pseudo datasets, or bootstrap toys, are generated by resampling the original dataset distribution. The unfolding procedure is applied to this ensemble of bootstrap toys, producing a set of unfolded distributions. The variance of these unfolded distributions is then evaluated to estimate the statistical uncertainty of the unfolded results. The bootstrap method leverages its independence from the data distribution, and the correlation between bins is taken into account after unfolding the toys. In the measurement, 4000 pseudo datasets, referred as bootstrap toys, are generated in different distributions to mimic the variability present in the original dataset for both data and MC samples.

For low-dimensional distributions, the resampling procedure involves adjusting bin contents by randomly varying them around their nominal values within the limits of statistical uncertainty. For higher-dimensional distributions, a weighting method is used, applying a random weight  $w$  drawn from a Poisson distribution,  $w = \text{Pois}(n, 1)$ , which represents the likelihood of observing a specific number of events  $n$  while expecting an average of one event.

To determine the statistical uncertainty of observed data, the variation in the unfolded distribution  $U_i^m$ , representing the content in bin  $i$  of the bootstrap toy  $m$ , is computed by applying the unfolding procedure as follows :

$$U_i^m = \sum_j U_{ij} (D_j^m - B_j)$$

where the  $D_j^m$  is the content in bin  $j$  of the bootstrap toy  $m$ ,  $B_j$  is the content in bin  $j$  due to the background.

When it comes to the statistical uncertainty of the simulated samples, each bootstrap toy results in a different migration matrix. Consequently, for bootstrap toy  $m$ , the variation in the unfolded distribution is expressed as :

$$U_i^m = \sum_j U_{ij}^m (D_j - B_j).$$

Here,  $U_{ij}^m$  is the migration matrix specific to bootstrap toy  $m$ , while  $D_j$  and  $B_j$  are the original data and background contents in bin  $j$ , respectively.

A series of unfolded distributions is generated from the set of bootstrap toys. By calculating the variance across these unfolded distributions, one obtains a robust estimate of the statistical uncertainty in the unfolded distribution. This method ensures that bin-by-bin correlations are properly accounted for. Figure 6.20 displays an example of the correlation matrix of the statistical uncertainties at the particle level for both observed data and MC simulations, after correcting for detector effects.

## Systematic uncertainty

The treatment of the propagation of systematic uncertainties follows a similar approach. Systematic uncertainties are obtained by comparing the systematic variations with the nominal distribution. To propagate these uncertainties, the unfolding procedure is repeated for the systematic variations, and the difference between the resulting unfolded distributions and the nominal unfolded distributions is taken as an estimate of the systematic uncertainties.

Two different cases of systematic uncertainties are considered : one related to correction sources, calibrations, and signal modeling, and the other related to background modeling.

For the first case, where corrections, calibrations, and signal modeling are concerned, the migration matrix varies across systematic variations while the distribution of observed data after background subtraction,  $D - B$ , remains unchanged. Thus, the unfolded distribution for a systematic variation  $k$  can be expressed as :

$$U_i^k = \sum_j U_{ij}^k (D_j - B_j)$$

where  $U_{ij}^k$  represents the migration matrix for the systematic variation.

For the second case, where the background distribution is varied systematically, the unfolded distribution can be expressed as :

$$U_i^k = \sum_j U_{ij} (D_j - B_j^k)$$

where  $B_j^k$  denotes the systematically varied background contribution in bin  $j$ , and the  $U_{ij}$  is the nominal migration matrix, which remains unchanged.

As the unfolding procedure introduces bin-by-bin correlations, the covariance matrix is extracted and used in the combination procedure in the following sections. The corresponding covariance matrix for a given systematic variation is given by :

$$C_{ij} = (U_i - U_i^{\text{nom}}) \cdot (U_j - U_j^{\text{nom}})$$

Considering the statistical covariance described in the previous subsection, the total covariance matrix is :

$$C_{ij}^{\text{total}} = C_{ij}^{\text{stat,Data}} + C_{ij}^{\text{stat,MC}} + \sum_k C_{ij}^k \quad (6.11)$$

where  $k$  runs over all systematic sources. This combined covariance matrix integrates both statistical and systematic uncertainties, offering a comprehensive estimate of the overall uncertainty in the unfolded distributions.

### 6.9.3 Uncertainty correlations

A correlation model is defined to perform the combinations between the electron and muon channels in Section 6.10, as well as the PDF profiling described in Section 6.11. The combination between electron and muon channels involves the electron

and muon channels from the same  $W$  boson at the same collision energy. In contrast, the PDF profiling comprises all the datasets at both 5.02 and 13 TeV.

The uncertainty sources arising from the lepton efficiency corrections and lepton energy calibrations are treated as fully correlated within the same lepton flavor at the same center-of-mass energy.

The recoil calibration uncertainty derived from  $Z$  events is fully correlated across both  $W^+$  and  $W^-$  measurements and between different lepton flavors. However, the calibrations for 5.02 and 13 TeV collision energies are treated as independent. The sub-dominant component, which addresses discrepancies between data and MC in the  $\Sigma \bar{E}_T$  distributions, is considered uncorrelated.

The uncertainties arising from the electroweak and top quark production modeling are treated as fully correlated across different lepton flavors and both  $W^+$  and  $W^-$  channels.

The multijet uncertainty is composed of two main components. The first component is the statistical uncertainty, which arises from the finite number of events and is treated as uncorrelated across different channels. The second component is the methodological uncertainty, which is related to the techniques and procedures used to estimate the multijet background. This component is assumed to be fully correlated between the various channels.

The bias uncertainty is treated as uncorrelated between  $W^+$  and  $W^-$ ,  $e$  and  $\mu$  channels. The uncertainty arising from the choice of the generator for the signal simulation is determined by comparing the nominal POWHEG and alternative SHERPA predictions and regarded as fully correlated across various channels.

The luminosities are determined separately for 5 TeV and 13 TeV. Therefore, they are correlated between electron and muon channels, as well as between  $W^+$  and  $W^-$  channels at the same energy. However, they are uncorrelated between different energies.

Table 6.7 summarises the correlation model.

## 6.10 Results

This section details the differential cross section results. This section presents the combined measured results for the electron and muon channels. The measured results are compared with the predictions calculated in NNLO QCD accuracy. This section is arranged as follows : Section 6.10.1 describes the combination procedure, while Sections 6.10.2–6.10.4 present the measured cross sections, combining electron and muon channels, in  $|\eta_\ell|$  and  $p_T^\ell$ , as well as the double differential case of  $p_T^\ell$  and  $|\eta_\ell|$ . The uncertainty breakdown is presented along with the results. In addition to the cross sections, Section 6.10.5 reports the  $W^\pm$  lepton charge asymmetry.

The differential cross sections in electron and muon channels are provided in Appendix D.1.

### 6.10.1 Combination procedure

TABLE 6.7 – The correlation model of the uncertainties for the combinations between different lepton flavors ( $e$  vs.  $\mu$ ), different  $W$  boson charges ( $W^+$  vs.  $W^-$ ), and different center-of-mass energies (5.02 vs. 13 TeV).

Source of uncertainty	$e$ vs. $\mu$	$W^+$ vs. $W^-$	5.02 vs. 13 TeV
Data & MC statistics	No	No	No
Lepton calibration & efficiencies	No	Full	No
Recoil calibration ( $Z$ -based)	Full	Full	No
Recoil calibration ( $\Sigma \bar{E}_T$ )	Full	No	No
EW & top-quark backgrounds	Full	Full	Full
Multijet background (statistics)	No	No	No
Multijet background (method)	Full	Full	Full
Unfolding bias	No	No	No
Generator uncertainty	Full	Full	Full
Luminosity	Full	Full	No

The best linear unbiased estimator (BLUE) prescription [169, 170] is employed to combine the measured differential cross sections in electron and muon channels. In this method, the combined result is assumed to be the weighted linear combination of the results in the two channels. The weights are determined to result in minimal variance. The BLUE approach incorporates the electron and muon channel cross sections, along with their uncertainties and correlations, into the  $\chi^2$  function, which is defined as :

$$\chi^2 = (\vec{x} - \vec{\mu})^T \mathbf{C}^{-1} (\vec{x} - \vec{\mu}). \quad (6.12)$$

Here,  $\vec{x}$  denotes the measured cross sections in the electron and muon channels, which can be written as :

$$\vec{x} = \{x_1^e, \dots, x_n^e, x_1^\mu, \dots, x_n^\mu\} \quad (6.13)$$

where  $x_i^e$  and  $x_i^\mu$  denote the measured cross sections in  $i$ -th bin in the two channels, respectively, and  $\vec{\mu}$  is the combined results :

$$\vec{\mu} = \{\mu_1, \dots, \mu_n, \mu_1, \dots, \mu_n\} \quad (6.14)$$

determined by minimizing the  $\chi^2$  function. The uncertainties and their correlations are incorporated into the  $\chi^2$  function using the covariance matrix  $\mathbf{C}$ , defined as :

$$\mathbf{C} = \begin{pmatrix} C^e & C^{e\mu} \\ C^{e\mu} & C^\mu \end{pmatrix}. \quad (6.15)$$

where  $C^e$  and  $C^\mu$  denote the covariance matrices for the electron and muon channels, respectively, and  $C^{e\mu}$  is the covariance matrix representing the correlation between the two channels.

When  $\chi^2$  reaches its minimal value, the combined results is given as follows :

$$\vec{\mu} = (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{C}^{-1} \vec{x}, \quad (6.16)$$

where  $\mathbf{H}$  is defined as a  $2N \times N$  matrix, with  $H_{ij} = 1$  when  $i = j$  or  $i = 2j$ , otherwise  $H_{ij} = 0$ . The covariance matrix of the combined result is determined as :

$$\mathbf{C}^{\text{comb}} = (\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-1}. \quad (6.17)$$

The estimate of the combined results from the analytical solution might be biased due to non-normal uncertainty distributions or uncertainties that vary depending on the central values of the cross section. Therefore, in this measurement, Eq. (6.12) is solved using an iterative procedure introduced in Ref. [170], taking advantage of its robustness under these conditions. During the iterations, rescaling procedures are applied to the covariance matrix. Most of the systematic uncertainties undergo linear rescaling, while the statistical uncertainties are rescaled using the square root method. However, uncertainties from background modeling are not rescaled and remain fixed.

### 6.10.2 Differential cross sections in $|\eta_\ell|$

The measured  $|\eta_\ell|$ -dependent differential cross section is obtained by performing the unfolding procedure in the  $|\eta_\ell|$  distribution with the binning described in Section 6.8.2. The combination between the electron and muon channels is performed following the procedure described in Section 6.10.1 and good compatibility is observed between different lepton flavors. Figure 6.21 shows the combination results of the electron and muon channels from  $W^+$  and  $W^-$  boson decays at 5.02 and 13 TeV. The combined result is also compared with the theoretical predictions as shown in Figure 6.22. The theory predictions are calculated with various PDF sets as described in Section 6.7. The size of uncertainties from different sources is illustrated in Figure 6.23. The luminosity uncertainty is dominant, and the contribution from the uncertainty arising from the choice of generator is significant.

### 6.10.3 Differential cross sections in $p_T^\ell$

The measured  $p_T^\ell$ -dependent differential cross section is obtained by performing the unfolding procedure in the  $p_T^\ell$  distribution with the binning described in the Section 6.8.2. The combination between the electron and muon channels is performed following the procedure described in Section 6.10.1 and good compatibility is observed between different lepton flavors. Figure 6.24 shows the combination results of the electron and muon channels from  $W^+$  and  $W^-$  boson decays at 5.02 and 13 TeV. The combined result is also compared with the theoretical predictions as shown in Figure 6.25. The theory predictions are calculated with various PDF sets as described in Section 6.7. The size of uncertainties from different sources is illustrated in Figure 6.26. The luminosity uncertainty is dominant, and the contribution from the uncertainty arising from the choice of generator is significant.

#### 6.10.4 Differential cross sections in $p_T^\ell$ and $|\eta_\ell|$

The double differential cross section in  $p_T^\ell$  and  $|\eta_\ell|$  is obtained by performing the unfolding procedure in the  $(|\eta_\ell|, p_T^\ell)$  distribution with the binning described in the Section 6.8.2. The combination between the electron and muon channels is performed following the procedure described in Section 6.10.1 and good compatibility is observed between different lepton flavors. Figure 6.27 shows the combination results of the electron and muon channels from  $W^+$  and  $W^-$  boson decays at 5.02 and 13 TeV. The combined result is also compared with the theoretical predictions as shown in Figure 6.28. The theory predictions are calculated with various PDF sets as described in Section 6.7. The size of uncertainties from different sources is illustrated in Figure 6.29.

#### 6.10.5 Charge asymmetry of $W$ bosons

In addition to the  $W$  boson differential cross sections, the  $W$  boson charge asymmetry as a function of  $|\eta_\ell|$  is derived by combining the differential cross-section results in  $|\eta_\ell|$ . The  $W$  boson charge asymmetry is defined as :

$$A_\ell = \frac{\frac{d\sigma_{W^+}}{d|\eta_\ell|} - \frac{d\sigma_{W^-}}{d|\eta_\ell|}}{\frac{d\sigma_{W^+}}{d|\eta_\ell|} + \frac{d\sigma_{W^-}}{d|\eta_\ell|}}, \quad (6.18)$$

where  $\frac{d\sigma(W^+)}{d|\eta_\ell|}$  and  $\frac{d\sigma(W^-)}{d|\eta_\ell|}$  denote the differential cross sections for  $W^+$  and  $W^-$  bosons as functions of  $|\eta_\ell|$ , respectively.

Figure 6.30 presents the combined results for the electron and muon channels from  $W^+$  and  $W^-$  boson decays at 5.02 and 13 TeV. These results are compared to theoretical predictions calculated using different PDF sets, as detailed in Section 6.7.

As a large number of the experimental uncertainties of the differential cross sections, such as the dominant luminosity uncertainty, cancels out, the lepton charge asymmetry provides a sensitive probe to the SM. This quantity is particularly sensitive to PDFs, especially the up and down quark distributions. Figure 6.31 displays the contributions of each uncertainty source. The statistical and generator uncertainties become the dominant components.

### 6.11 PDF profiling

This section details the study about the impact of the differential  $W$  boson cross sections in the leptonic decay channels on the PDFs. Section 6.11.1 provides the theoretical framework and preparations to study the impact of the new experimental data via the profiling technique. Section 6.11.2 reports the improvement on PDFs by incorporating the results from this measurement.

#### 6.11.1 Methodology

PDFs describe the momentum distribution of quarks and gluons inside the proton, which play an important role in understanding the proton structure. As described

in Section 2.3, the cross section of the hard scattering process can be derived by convoluting the PDFs with the partonic cross sections according to the factorization theorem in Eq. (2.20). Thus, accurate PDFs are essential to make precise predictions of the hard scattering process in high-energy regimes.

The Drell-Yan production process, including  $W^\pm \rightarrow \ell^\pm \nu$ , is an important tool for determining the PDFs. The DY process is well understood theoretically and can provide a direct probe of the quark and anti-quark distributions in the proton, particularly the sea quark distribution, as described in Section 2.4. In principle, one can determine the PDFs using the ATLAS dataset along with other datasets such as HERA I+II data. However, it requires a full QCD analysis to determine a new PDFs set. In this measurement, the PDF profiling technique [171] is employed to get an initial and quick impact of the new experimental measurements on the existing PDFs.

The PDF profiling technique is an Hessian reweighing method of PDFs. It is under the assumption that the new added data are compatible with the theory predictions based on current PDF sets so the central values of data are set to match the central values of theory predictions. The  $\chi^2$  function is defined to quantify the agreement between data and theoretical predictions depending on the PDFs and minimized by adjusting the current PDFs to achieve better agreement while preserving the original constraints and correlations. The experimental uncertainties and PDF uncertainties are taken into account when constructing the  $\chi^2$  formula along with their internal correlations. The  $\chi^2$  is defined as :

$$\chi^2(\boldsymbol{\beta}^{\text{exp}}, \boldsymbol{\beta}^{\text{theo}}) = \sum_i^{N_{\text{data}}} \left( \frac{\sigma_i^{\text{exp}} + \sum_j \beta_j^{\text{exp}} \Gamma_{ij}^{\text{exp}} - \sigma_i^{\text{theory}} - \sum_j \beta_j^{\text{theo}} \Gamma_{ij}^{\text{theo}}}{\Delta_i} \right)^2 + \sum_i \left( \beta_i^{\text{exp}} \right)^2 + \sum_i \left( \beta_i^{\text{theo}} \right)^2 \quad (6.19)$$

where  $i$  runs over the provided datasets,  $\sigma^{\text{exp}}$  and  $\sigma^{\text{theory}}$  denote the central values of experimental measured results and predictions, and  $\boldsymbol{\beta}^{\text{exp}}$  and  $\boldsymbol{\beta}^{\text{theo}}$  represent experimental and PDF nuisance parameters, while no other theoretical uncertainties are considered during the calculation. The nuisance parameters are associated with the physics quantities  $\sigma^{\text{data}}$  and  $\sigma^{\text{theory}}$  through the matrix elements  $\Gamma_{ij}^{\text{exp}}$  and  $\Gamma_{ij}^{\text{theo}}$  for experimental cross sections and theoretical predictions, respectively. Shifts on the  $\boldsymbol{\beta}^{\text{theo}}$  and the change of their uncertainties impact the theory predictions. The denominator in the first term,  $\Delta_i$ , is defined as :

$$\Delta_i = \left( \delta_{\text{stat}}^i \right)^2 \sigma_i^{\text{exp}} \sigma_i^{\text{theory}} + \left( \delta_{\text{unc}}^i \sigma_i^{\text{theory}} \right)^2 \quad (6.20)$$

where  $\delta_{\text{stat}}^i$  and  $\delta_{\text{unc}}^i$  denote the uncertainty arising from data statistical fluctuations and uncorrelated systematic uncertainties, respectively. This definition allows to minimize the bias related to data statistic uncertainties. The last two terms are penalty terms for the shifts on nuisance parameters  $\beta_j^{\text{exp}}$  and  $\beta_j^{\text{theo}}$ .

The profiled PDFs are obtained by adjusting the  $\beta^{\text{theo}}$  to minimizing the  $\chi^2$  function, updating the PDFs to give a better agreement between data and theory predictions. The updated central values of PDF set,  $\mathcal{F}'_0$  can be expressed as :

$$\mathcal{F}'_0 = \mathcal{F}_0 + \sum_k \left[ \hat{\beta}_k^{\text{theo}} \left( \frac{\mathcal{F}_k^{(+)} - \mathcal{F}_k^{(-)}}{2} \right) + \left( \hat{\beta}_k^{\text{theo}} \right)^2 \left( \frac{\mathcal{F}_k^{(+)} + \mathcal{F}_k^{(-)} - 2\mathcal{F}_0}{2} \right)^2 \right] \quad (6.21)$$

where  $k$  runs over eigenvectors,  $\mathcal{F}_0$  denotes the central value, the symbols ‘+’ and ‘-’ represent the up and down variations, and  $\hat{\beta}^{\text{theo}}$  denote the best-fit nuisance parameters related to the PDF uncertainties.

The PDF profiling is presented using the `xFitter` framework [172-174]. For the theory prediction, the total  $pp \rightarrow X$  cross section  $\sigma_{pp \rightarrow X}$  is factorized as two components according to the factorization theorem described in Section 2.3 :

$$\sigma_{pp \rightarrow X} = \sum_{ij} \int dx_1 dx_2 f_i^p(x_1, Q^2) f_j^p(x_2, Q^2) \times \hat{\sigma}_{ij}(x_1 x_2 s, \alpha_s(Q^2)) \quad (6.22)$$

where  $f_i^p(x, Q^2)$  denotes the PDF term for each parton  $i$ ,  $\hat{\sigma}_{ij}(x_1, x_2, \alpha_s(Q^2))$  represents the contribution of each  $ij$  parton pair combination to the hard scattering cross section. This term is not explicitly dependent on the PDFs and can be pre-computed over a range of parton momentum fractions ( $x_1, x_2$ ), center-of-mass energy  $s$ , and the strong coupling constant  $\alpha_s$ . Cross section grids, which are independent of a specific PDF set, serve as pre-computed tables. The theoretical cross-section predictions are then obtained by convoluting these grids with the PDFs.

The predicted cross-section grids are computed using `MadGraph5_aMC@NLO` [175] at the NLO level, which are interfaced with `APPLgrid` [176] using `aMCfast` [177]. In a further step, bin-by-bin NNLO corrections are obtained by comparing NLO predictions with NNLO predictions using the same PDF. The ratios can be expressed as  $K$ -factors,  $K_i^{\text{NLO} \rightarrow \text{NNLO QCD}}$ , for a generic observable  $X$  :

$$K_i^{\text{NLO} \rightarrow \text{NNLO QCD}} = \left( \frac{d\sigma_{\text{PDF}}^{\text{NNLO QCD}}/dX}{d\sigma_{\text{PDF}}^{\text{NLO QCD}}/dX} \right)_i \quad (6.23)$$

where  $i$  denotes the  $i$ -th bin of  $X$  distribution.

The primary purpose of the PDF profiling study is to evaluate the sensitivity of the new dataset to PDFs. It can not replace the full QCD analysis which is used to determine the PDFs, as the profiling is based on the assumption of compatibility between the dataset and predictions with existing PDFs and the tolerance  $\Delta\chi^2$  is set to 1 in profiling while it can be a different value or even dynamic in PDFs determination.

In this study, the cross sections at NNLO level is computed using `DYTurbo` [134-137] as described in Section 6.7, consequently, the NNLO corrections are derived by comparing the NNLO and NLO predictions from grids.

### 6.11.2 Results

The impact of the newly measured cross sections are tested by applying the profiling on the MSHT20, NNPDF4.0 and ATLASpdf21 PDF sets using the differential cross sections as functions of the  $|\eta_\ell|$  observable at 5.02 TeV and 13 TeV. MSHT20 is determined using enhanced  $\chi^2$  tolerance, which is not the same as the  $\chi^2$  criterion  $\Delta\chi^2 = 1$  applied in the PDF profiling. NNPDF4.0 is the most recent release of the NNPDF family sets and the Hessian conversion [178] is used. Unlike the ATLASpdf21 used for the theoretical predictions described in Section 6.7, which is determined using enhanced tolerance  $\chi^2 = 3$ , the ATLASpdf21 PDFs set for the PDF profiling study is assessed with a  $\chi^2$  tolerance  $\Delta\chi^2 = 1$ .

The electron and muon channels are combined as inputs for profiling, thus one has four sets of data points for  $W^+$  and  $W^-$  at 5.02 and 13 TeV, respectively. For the experimental uncertainties, the uncertainty correlation model used in the combinations, described in Section 6.9, is employed.

The comparison between measured cross sections and the theory predictions using different PDF sets are shown in Figures 6.32-6.34. Red lines denote the predictions with the initial PDFs along with the error band arising from PDF uncertainties. The blue lines with the error band represent the predictions with the updated PDFs after performing the profiling. Good agreement between data and predictions with NNPDF4.0 is observed in all channels. The prediction of ATLASpdf21 is slightly higher than the measured results, resulting the shifts on the theoretical predictions after profiling.

Figures 6.35-6.36 illustrate the MSHT20 results. Figures 6.37-6.38 present the NNPDF4.0 results. Figures 6.39-6.40 displays the ATLASpdf21 results. The valence quark distributions,  $xu_v$  and  $xd_v$ , as well as the strange quark distribution in the light sea, are checked. Additionally, the fraction of the strange quark over the light quarks, defined as  $R_s = x(s + \bar{s})/(2\bar{d})$ , is checked. The impact on the relative uncertainties of various quark distributions is investigated:  $xu_v$ ,  $xd_v$ ,  $x\bar{U}$ ,  $x\bar{D}$ ,  $xs$  and  $R_s$ , where  $x\bar{U}$  and  $x\bar{D}$  denote the anti-quark of  $u$ -type and  $d$ -type in the light sea,  $xs$  represents the sea strange-quark distribution.

The impact of the PDF profiling varies depending on the input PDF. In comparison with HERA data from  $ep$  collision, The ATLAS data is expected to provide more details on the PDFs of sea quarks as well as in the low  $x$  regime based on the previous study [96]. In general, with the measured results, significant reduction in PDF uncertainties on sea quark distributions is observed. For the MSHT20 PDF set, the PDF uncertainties of sea quark PDFs are especially improved in the low- $x$  region. For NNPDF4.0 PDF set, PDF uncertainties are reduced in valence quarks distributions  $xu_v$  and  $xd_v$ , as well as the sea quark distributions  $x\bar{u}$  and  $x\bar{d}$  at the intermediate  $x$ . ATLASpdf21 incorporates the measurements from the ATLAS experiment taken from  $pp$  collisions at  $\sqrt{s} = 7, 8, \text{ and } 13$  TeV, thus the impact of PDF profiling is expected to be small in comparison with other PDF sets.

The  $W^\pm$  boson differential cross section has emerged as a powerful input in the

pursuit of mitigating systematic uncertainties associated with PDF determination and exploitations.

## 6.12 Conclusion

The measurement of the differential cross sections of the  $W^\pm$  boson production in the leptonic decay channels, including both electrons and muons, is presented using the proton-proton collision data at 5.02 and 13 TeV with the ATLAS detector at special LHC runs in 2017 and 2018 during Run 2 data taking. The data samples correspond to an integrated luminosity of  $258 \text{ pb}^{-1}$  for 5.02 TeV and  $340 \text{ pb}^{-1}$  for 13 TeV, respectively. The corresponding uncertainties of the integrated luminosity are 1% and 0.9%.

The single and double differential cross sections are measured in terms of lepton pseudorapidity  $|\eta_\ell|$  and lepton transverse momentum  $p_T^\ell$  in the electron and muon channels. There are good consistencies between different lepton flavour channels. The typical experimental uncertainties in the bulk phase space are below 1%. The measured results are compared with theoretical predictions using DYTURBO at NNLO and NNLL accuracy in QCD with different PDF sets. General good agreement between the two is observed. In the high  $p_T^\ell$  region, the theoretical predictions are slightly underestimated.

The impact of the measurements on PDFs is evaluated using the profiling technique. The effect varies across different PDF sets, but generally speaking, a reduction in PDF uncertainties for both valence and sea quark distributions is observed.

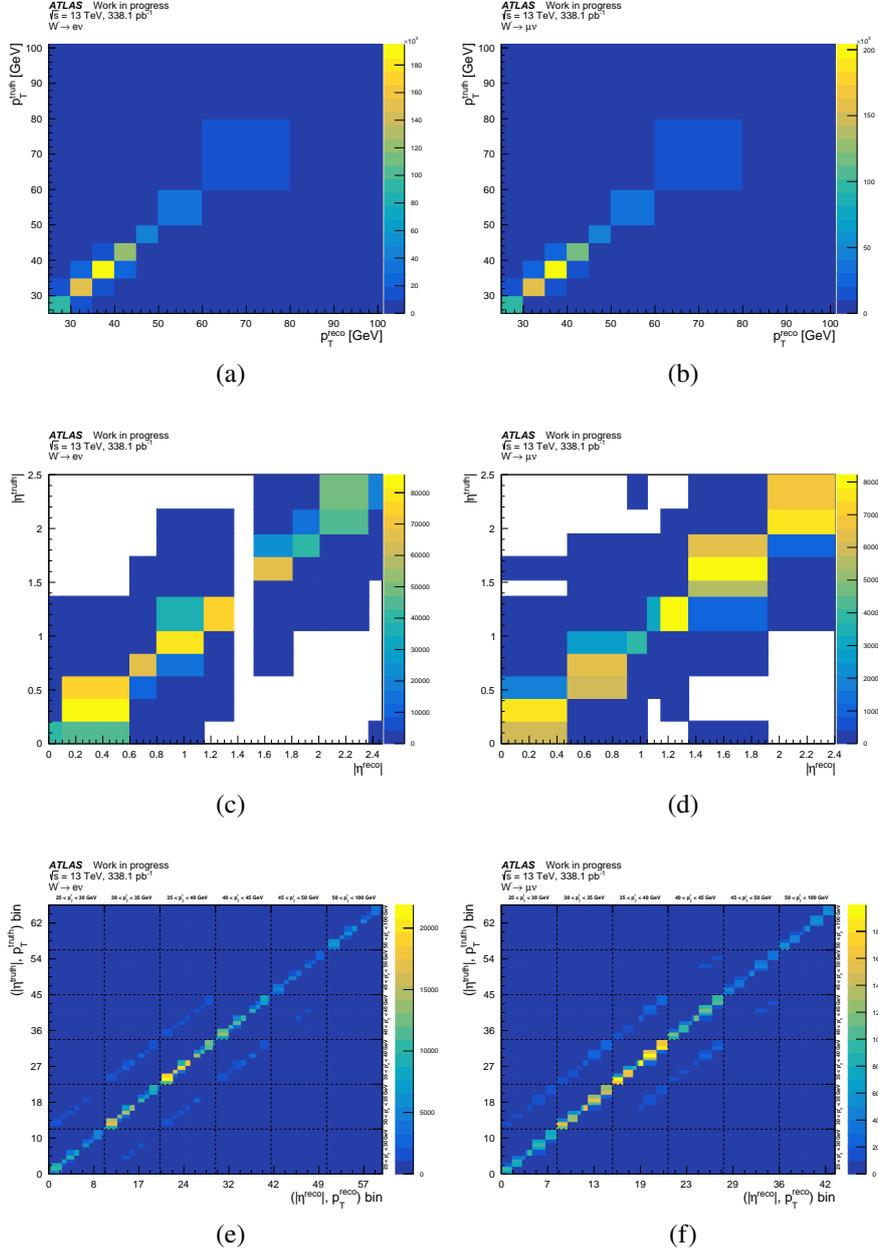


FIGURE 6.15 – Example of the migration matrix used in the unfolding for  $p_T^{\ell}$ ,  $\eta^{\ell}$ , and  $(|\eta_{\ell}|, p_T^{\ell})$  for  $W^- \rightarrow e^- \bar{\nu}$  and  $W^- \rightarrow \mu^- \bar{\nu}$  at 13 TeV

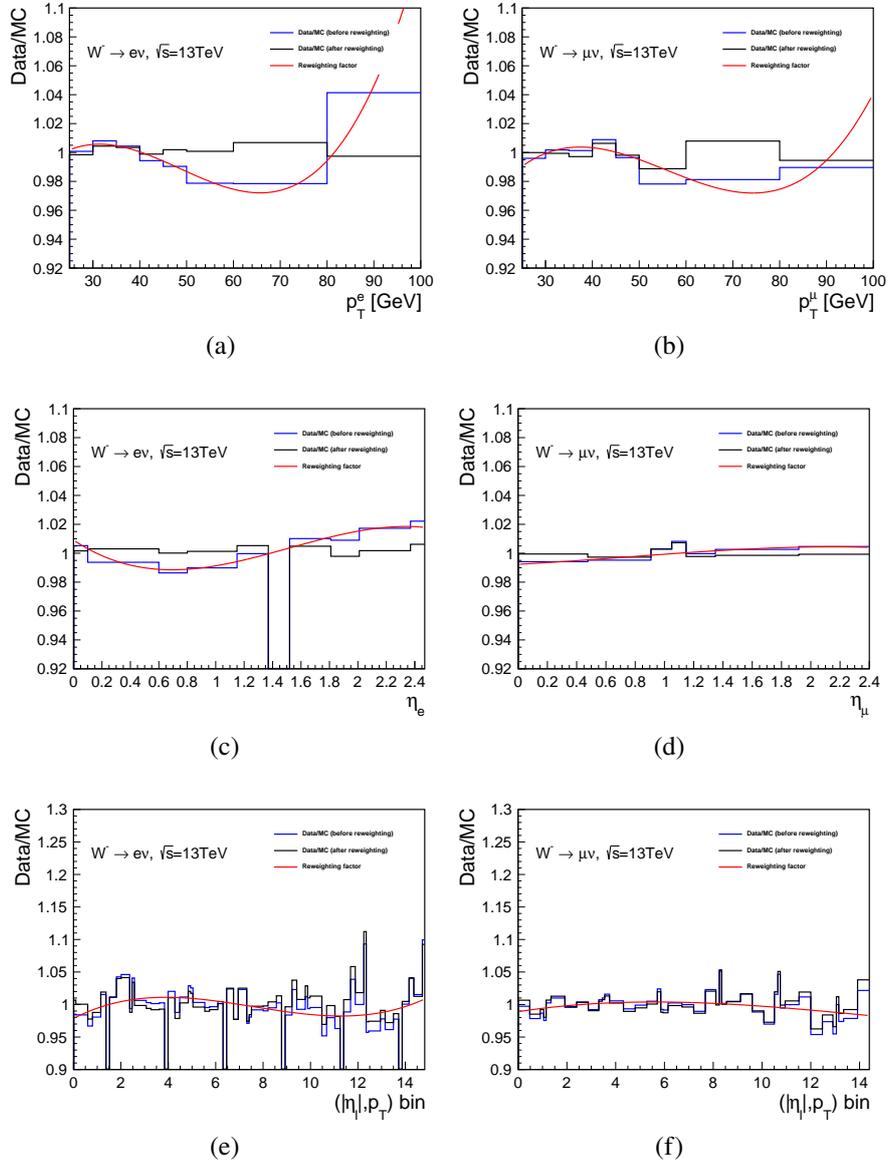


FIGURE 6.16 – The data over reconstructed-level MC distribution ratios before (blue histogram) and after (black histogram) the reweighting as well as the smooth reweighting functions (red curve) of different observables for  $W^- \rightarrow e^- \bar{\nu}$  and  $W^- \rightarrow \mu^- \bar{\nu}$  at 13 TeV.

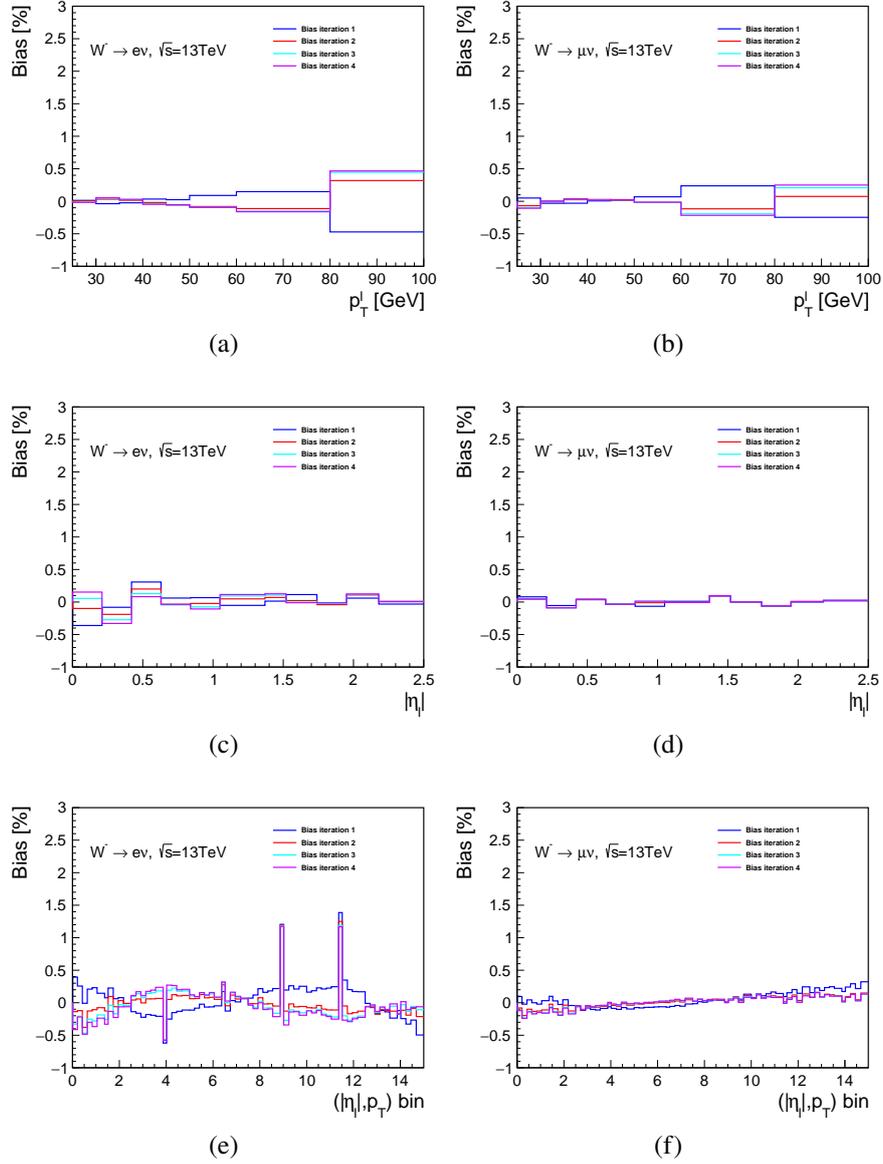


FIGURE 6.17 – The unfolding bias uncertainty for various numbers of iterations using the Bayesian unfolding method. Different observables are presented for  $W^- \rightarrow e^- \bar{\nu}$  and  $W^- \rightarrow \mu^- \bar{\nu}$  at 13 TeV.

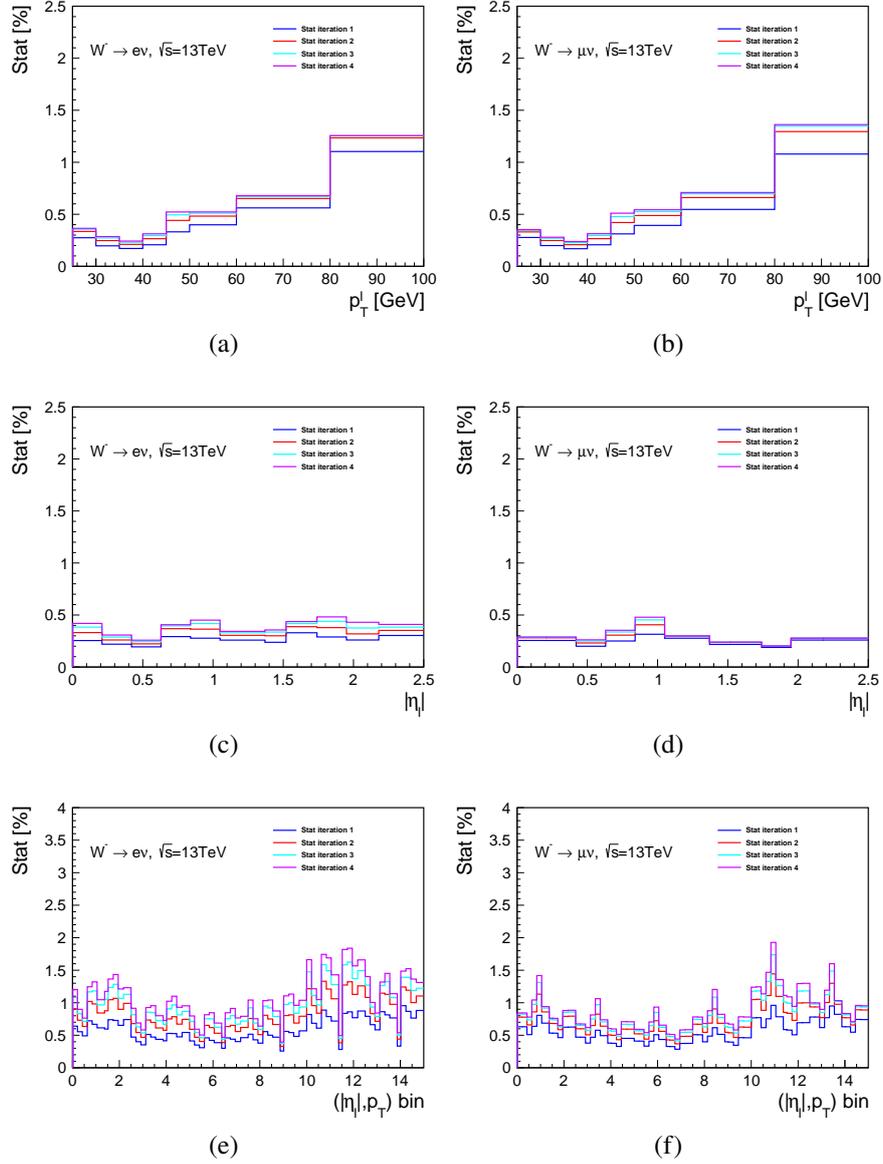


FIGURE 6.18 – The statistical uncertainty for various numbers of iterations using the Bayesian unfolding method. Different observables are presented for  $W^- \rightarrow e^- \bar{\nu}$  and  $W^- \rightarrow \mu^- \bar{\nu}$  at 13 TeV.

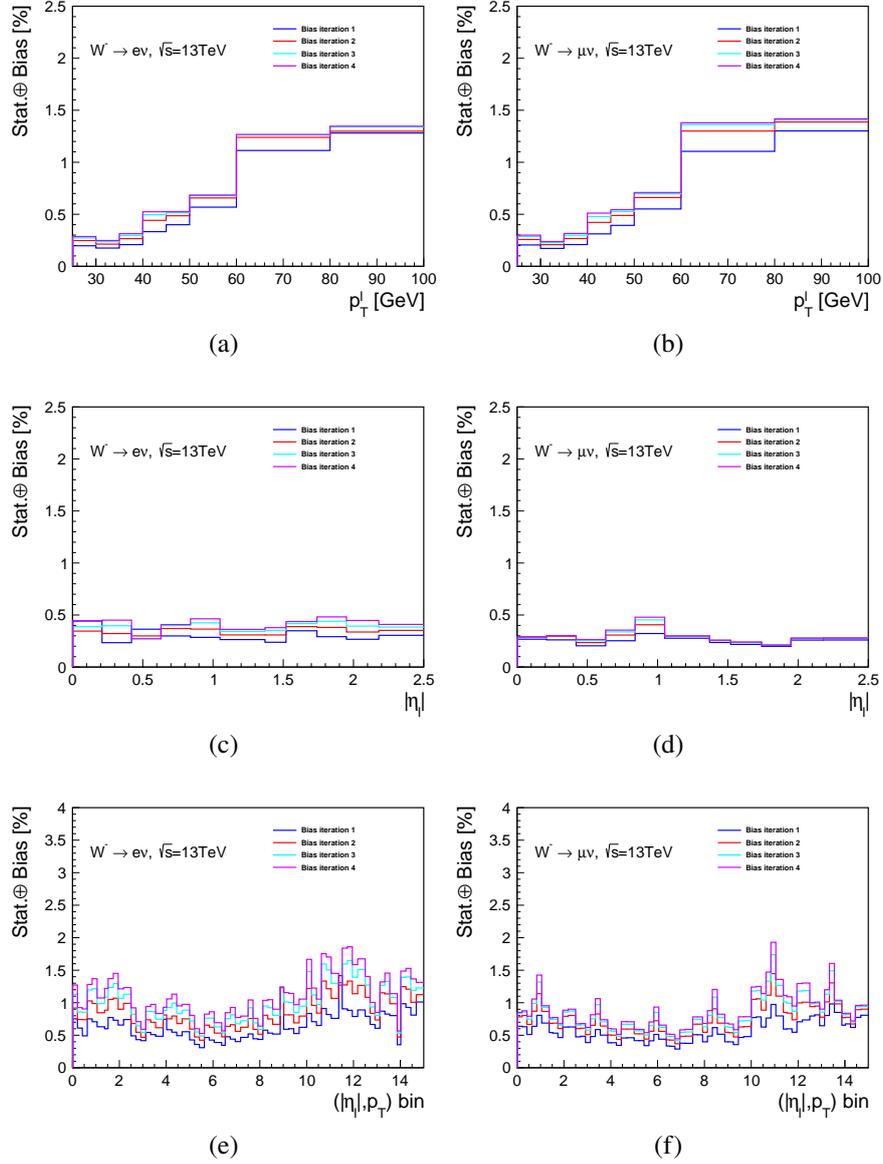


FIGURE 6.19 – The sum of unfolding bias and statistical uncertainty for various numbers of iterations using the Bayesian unfolding method. Different observables are presented for  $W^- \rightarrow e^- \bar{\nu}$  and  $W^- \rightarrow \mu^- \bar{\nu}$  at 13 TeV.

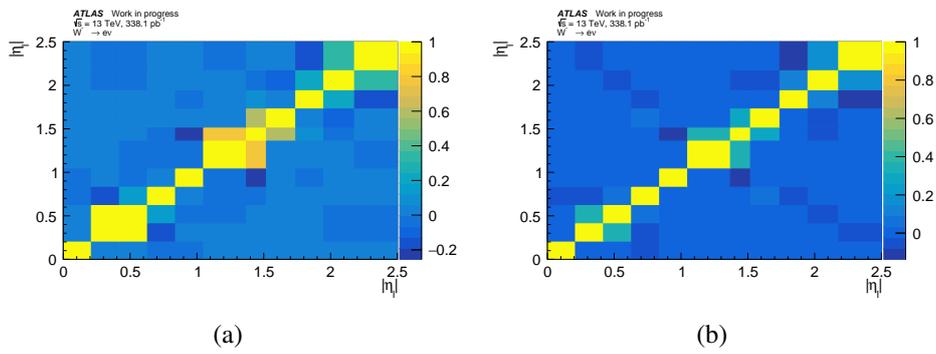


FIGURE 6.20 – The correlation matrix of the statistical uncertainties propagated to the particle level for  $W^- \rightarrow e^- \bar{\nu}$  at  $\sqrt{s} = 13$  TeV is shown for (a) observed data and (b) simulation samples.

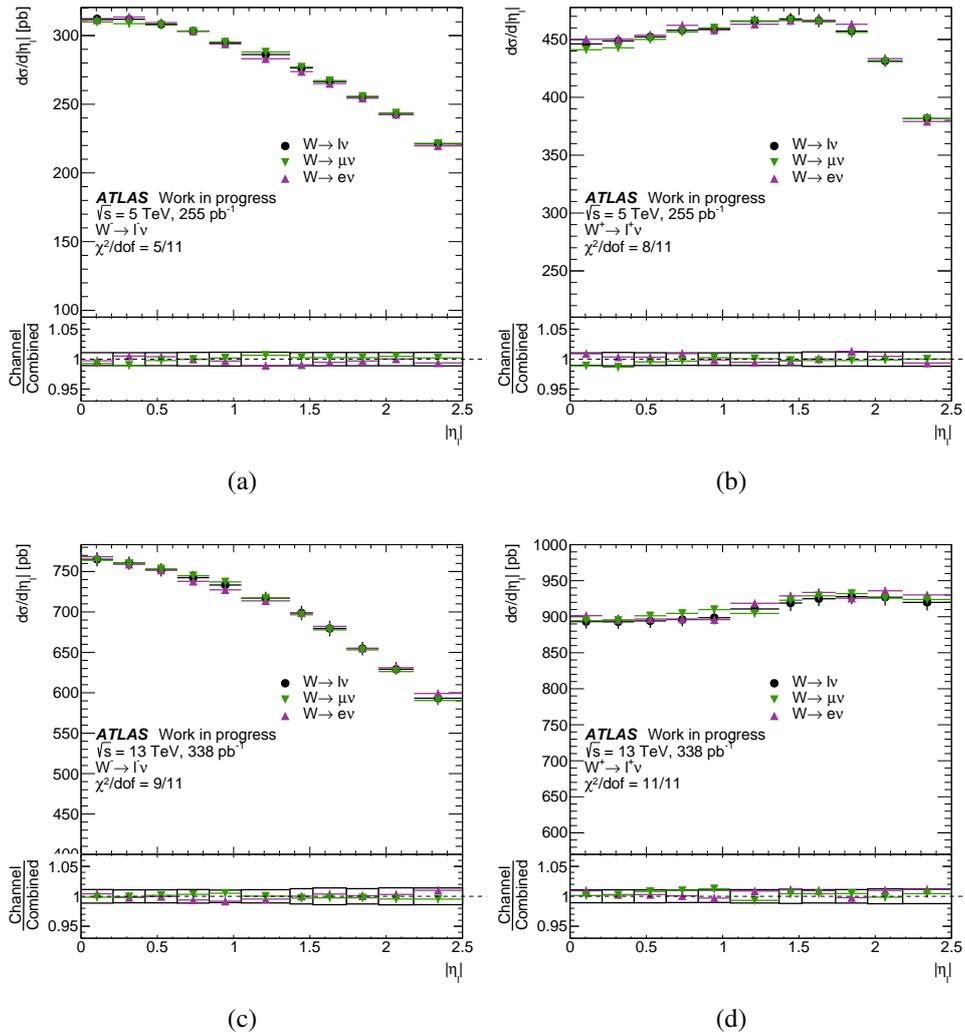


FIGURE 6.21 – Combination of differential cross sections of the electron and muon channels as a function of  $|\eta_\ell|$  at 5.02 and 13 TeV. The electron (purple triangles) and muon (green triangles) results are displayed with uncorrelated uncertainties while the uncertainty associated to the combined results is correlated uncertainty across channels.

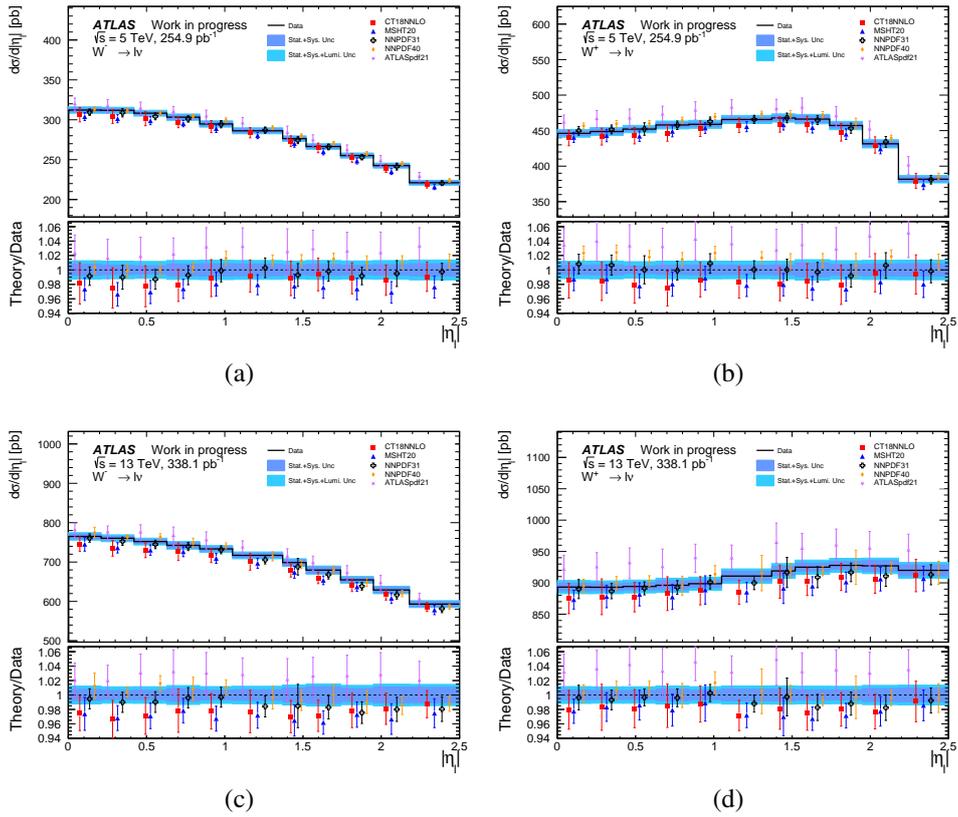


FIGURE 6.22 – Differential cross section as a function of  $|\eta_\ell|$  for the combination between the electron and muon channels at 5.02 and 13 TeV. The data are presented with systematic and total uncertainties and compared with the theory predictions with different PDF sets. The ratios of the data to the corresponding predictions are shown in the lower panels.

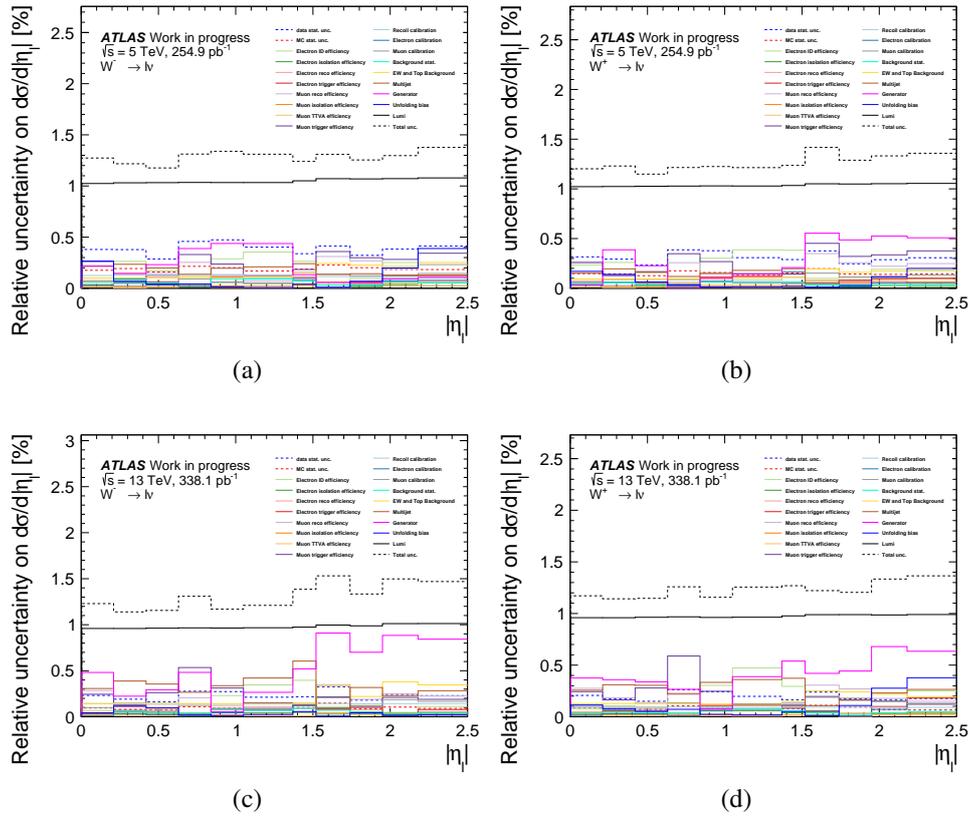


FIGURE 6.23 – The relative systematic uncertainty from each source for differential cross section as a function of  $|\eta^\ell|$  for the combined results between electron and muon channels at 5.02 and 13 TeV.

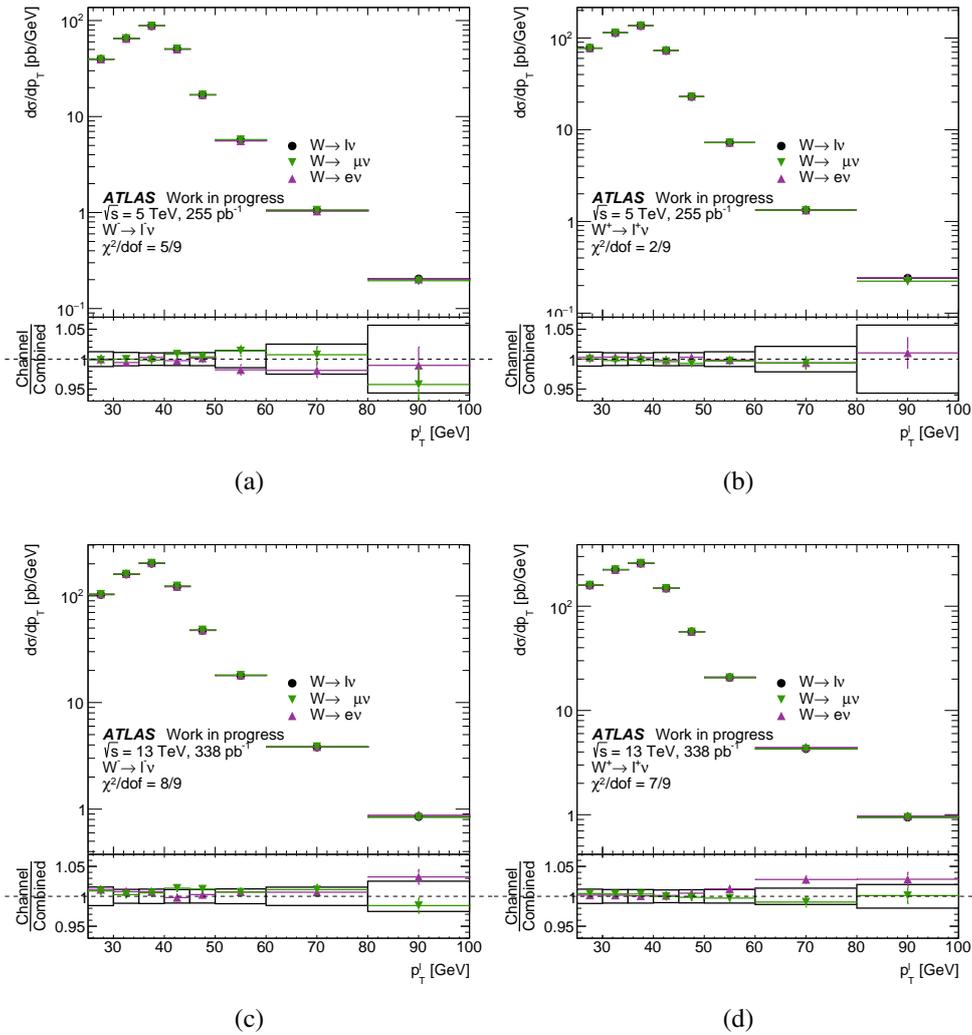


FIGURE 6.24 – Combination of differential cross sections of the electron and muon channels as a function of  $p_T^\ell$  at 5.02 and 13 TeV. The electron (purple triangles) and muon (green triangles) results are displayed with uncorrelated uncertainties while the uncertainty associated to the combined results is correlated uncertainty across channels.

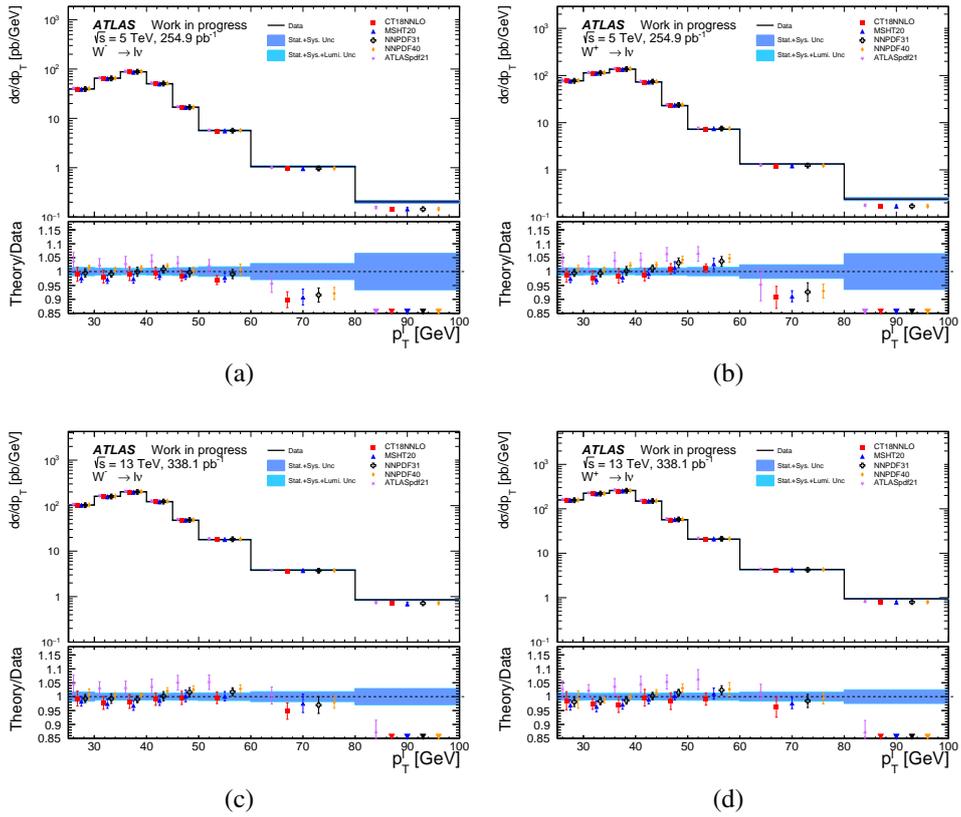


FIGURE 6.25 – Differential cross section as a function of  $p_T^\ell$  for the combination between the electron and muon channels at 5.02 and 13 TeV. The data are presented with systematic and total uncertainties and compared with the theory predictions with different PDF sets. The ratios of the data to the corresponding prediction are shown in the lower panels.

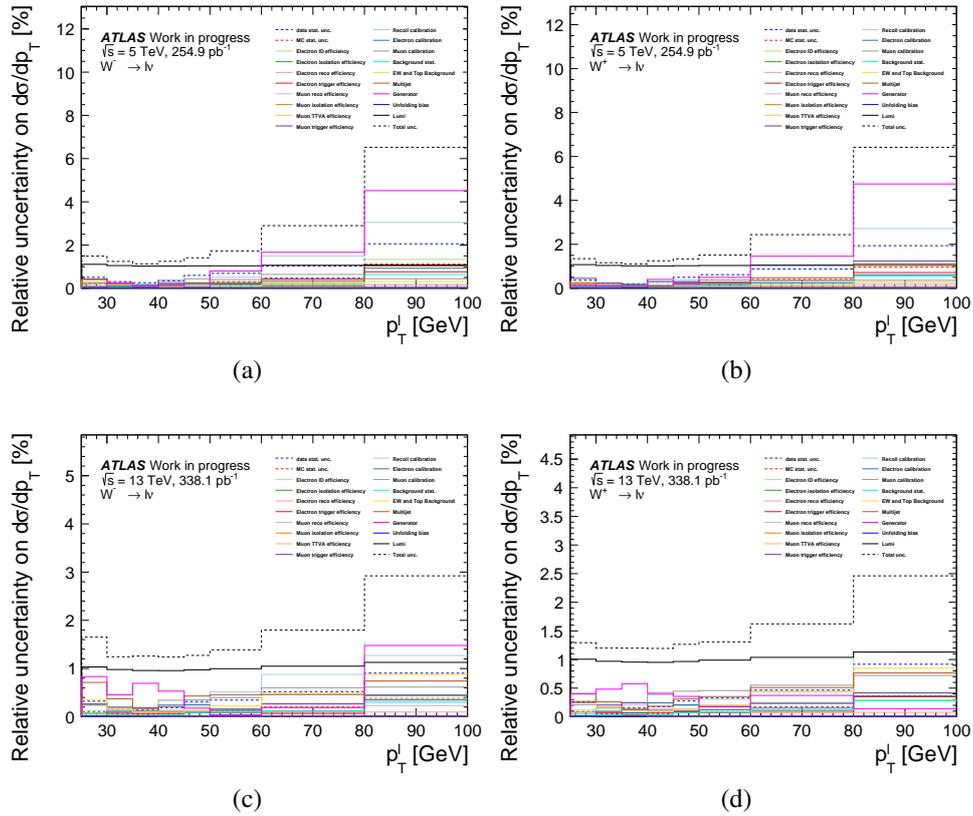


FIGURE 6.26 – The relative systematic uncertainty from each source for differential cross section as a function of  $p_T^l$  for the combined results between electron and muon channels at 5.02 and 13 TeV.

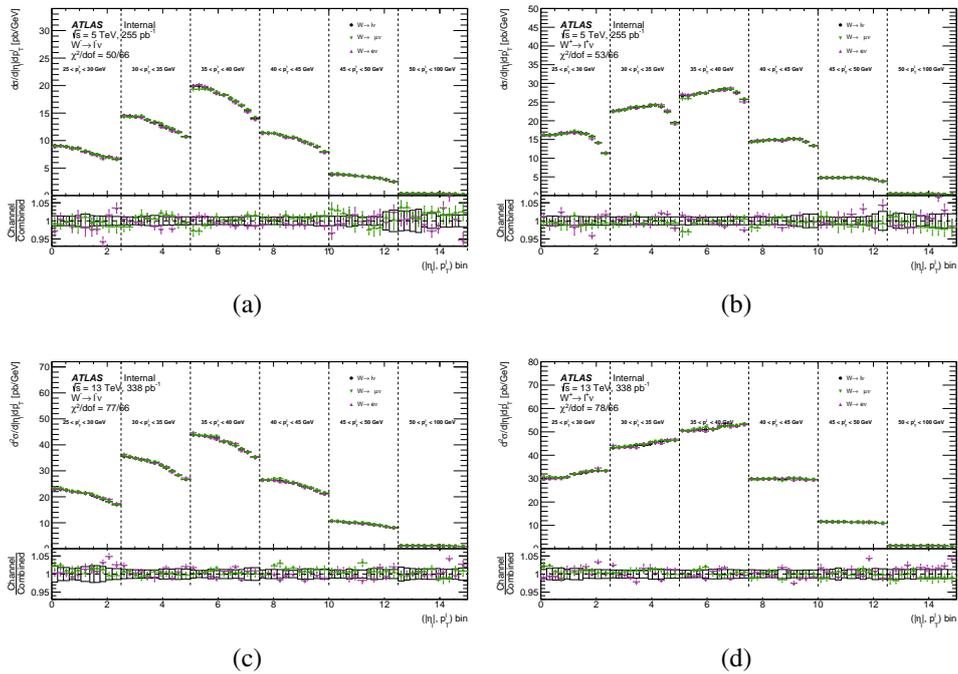
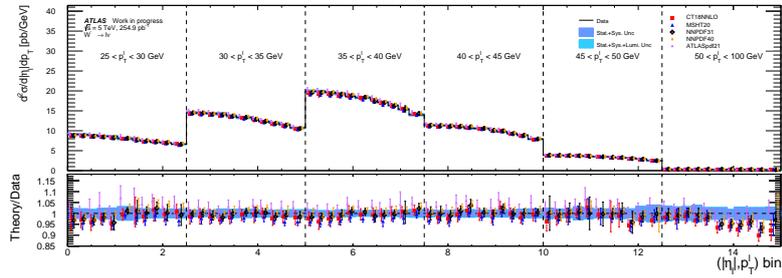
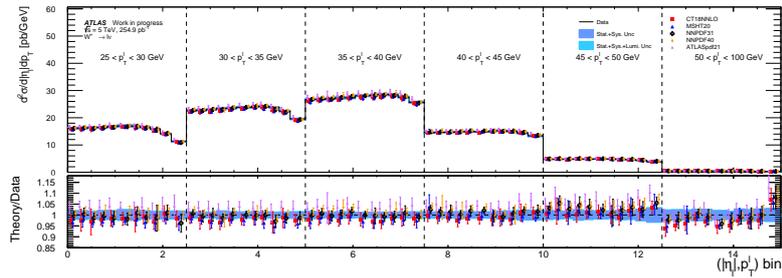


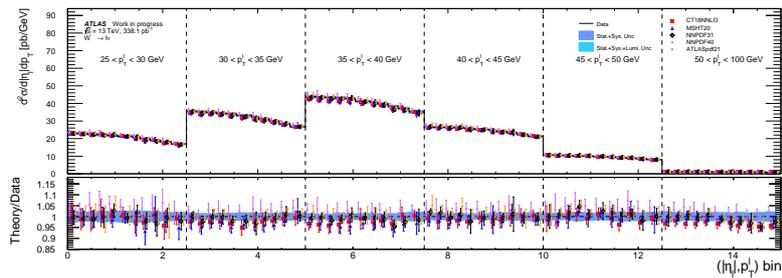
FIGURE 6.27 – Combination of differential cross sections of the electron and muon channels as functions of  $|\eta_\ell|$  and  $p_T^\ell$  at 5.02 and 13 TeV. The electron (purple triangles) and muon (green triangles) results are displayed with uncorrelated uncertainties while the uncertainty associated to the combined results is correlated uncertainty across channels.



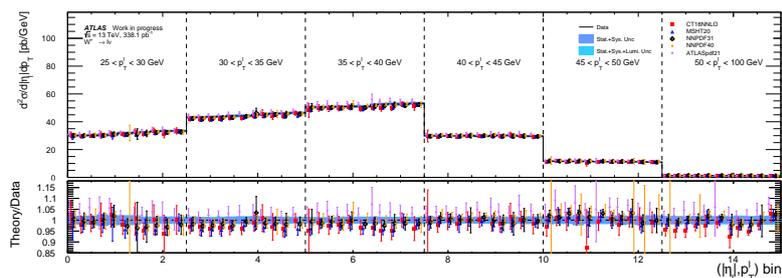
(a)



(b)

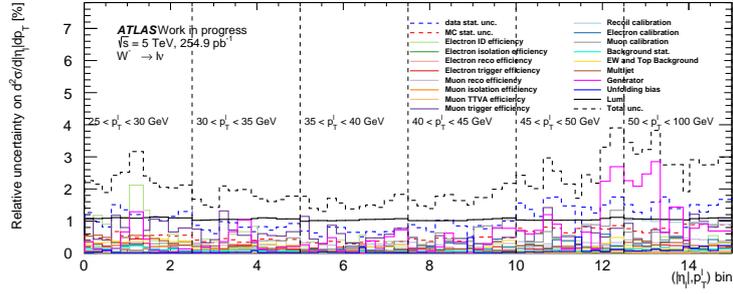


(c)

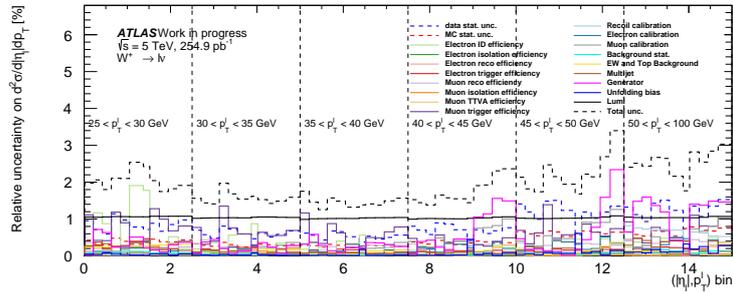


(d)

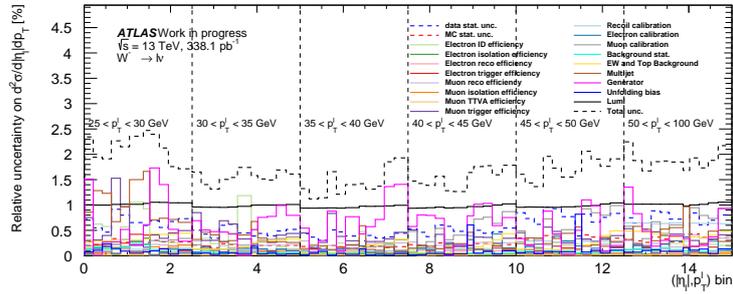
FIGURE 6.28 – Differential cross section as functions of  $|\eta_e|$  and  $p_T^\ell$  for the combination between the electron and muon channels at 5.02 and 13 TeV. The data are presented with systematic and total uncertainties and compared with the theory predictions with different PDF sets. The ratios of the data to the corresponding predictions are shown in the lower panels.



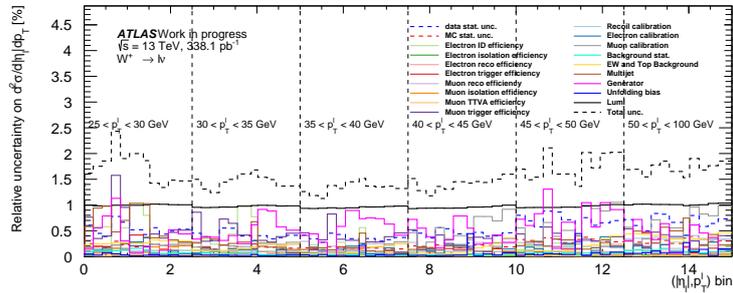
(a)



(b)



(c)



(d)

FIGURE 6.29 – The relative systematic uncertainty from each source for differential cross section as a function of  $p_T^\ell$  and  $|\eta_\ell|$  for the combined results between electron and muon channels at 5.02 and 13 TeV.

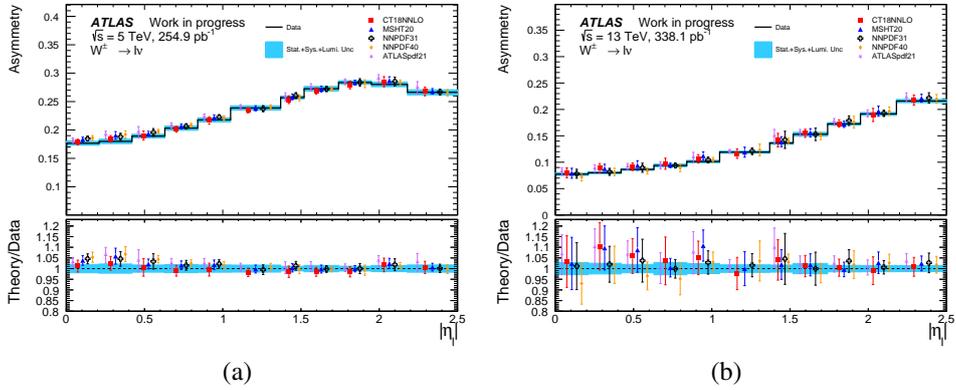


FIGURE 6.30 – The measured  $W$  boson charge asymmetry as a function of  $|\eta_\ell|$  for the combined electron and muon channels at (a) 5.02 and (b) 13 TeV. The data are presented with total uncertainties and compared with theoretical predictions using different PDF sets. The ratios of the data to the corresponding predictions are displayed in the lower panels.

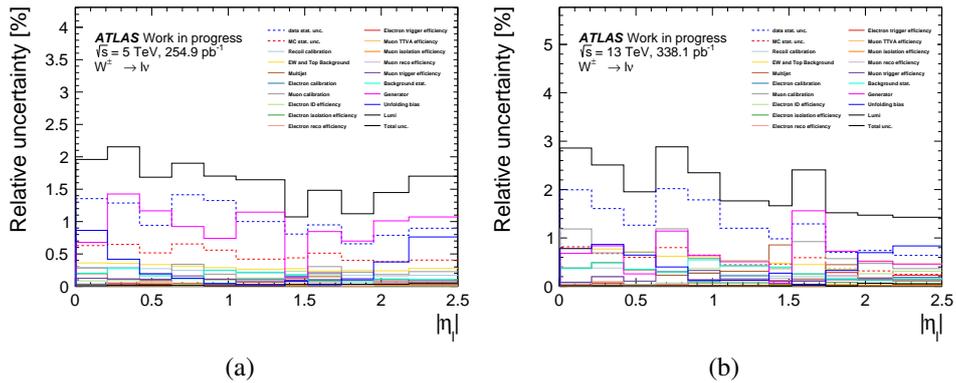


FIGURE 6.31 – The relative systematic uncertainty from charge lepton asymmetry as a function of  $|\eta_\ell|$  for the combined results of the electron and muon channels at (a) 5.02 and (b) 13 TeV.

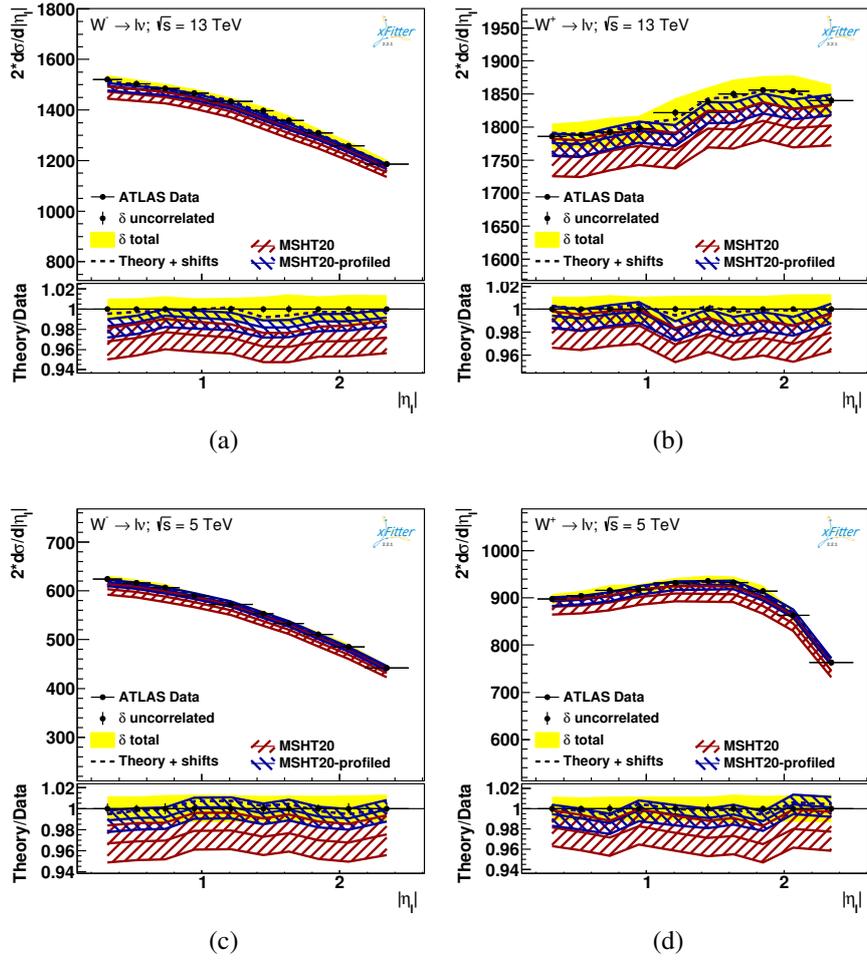


FIGURE 6.32 – Differential cross section as a function of  $|\eta_e|$  for the electron and muon channels at 5 and 13 TeV. The data presented with total uncertainties and compared with the theory predictions with original and profiled MSHT20 PDF sets. The ratios of the data to the corresponding predictions are shown in the lower panels.

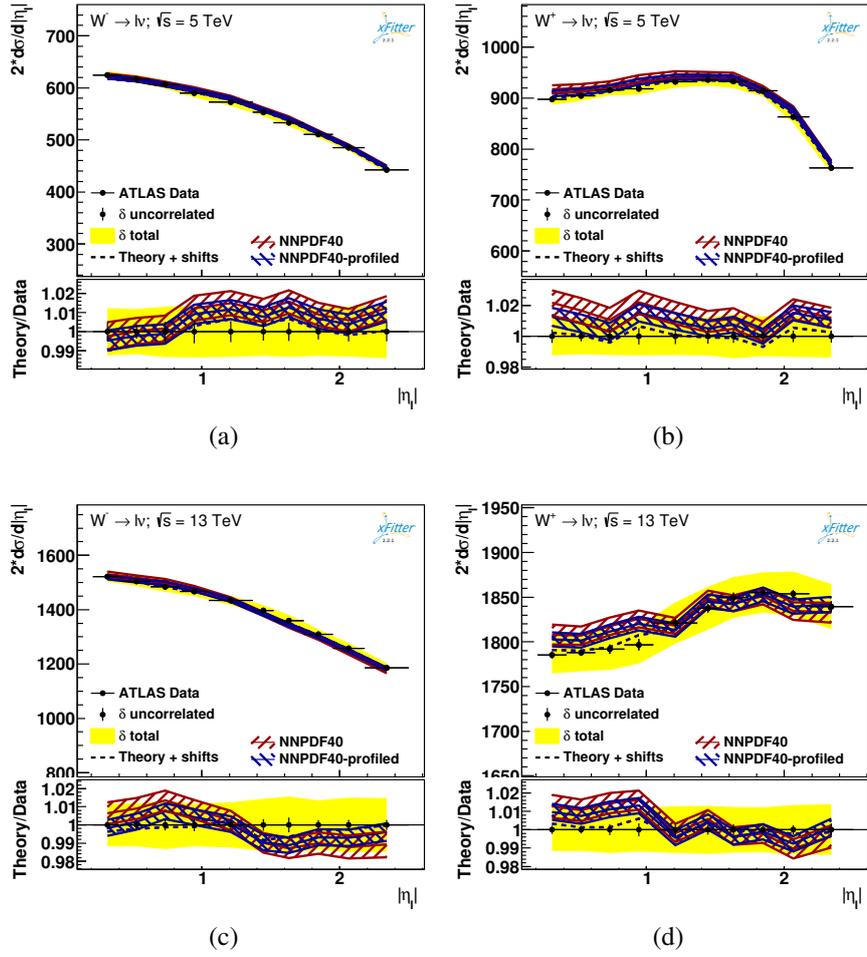


FIGURE 6.33 – Differential cross section as a function of  $|\eta_\ell|$  for the electron and muon channels at 5 and 13 TeV. The data presented with total uncertainties and compared with the theory predictions with original and profiled NNPDF40 PDF sets. The ratios of the data to the corresponding predictions are shown in the lower panels.

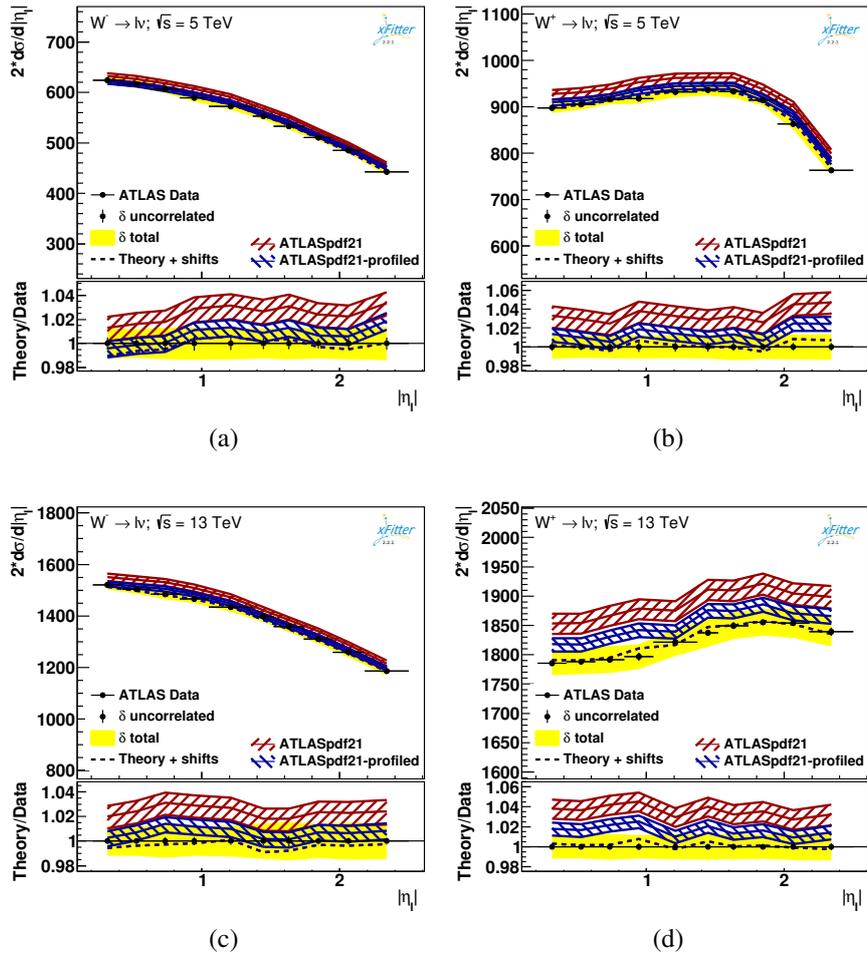


FIGURE 6.34 – Differential cross section as a function of  $|\eta_\ell|$  for the electron and muon channels at 5 and 13 TeV. The data presented with total uncertainties and compared with the theory predictions with original and profiled ATLASpdf21 PDF sets. The ratios of the data to the corresponding predictions are shown in the lower panels.

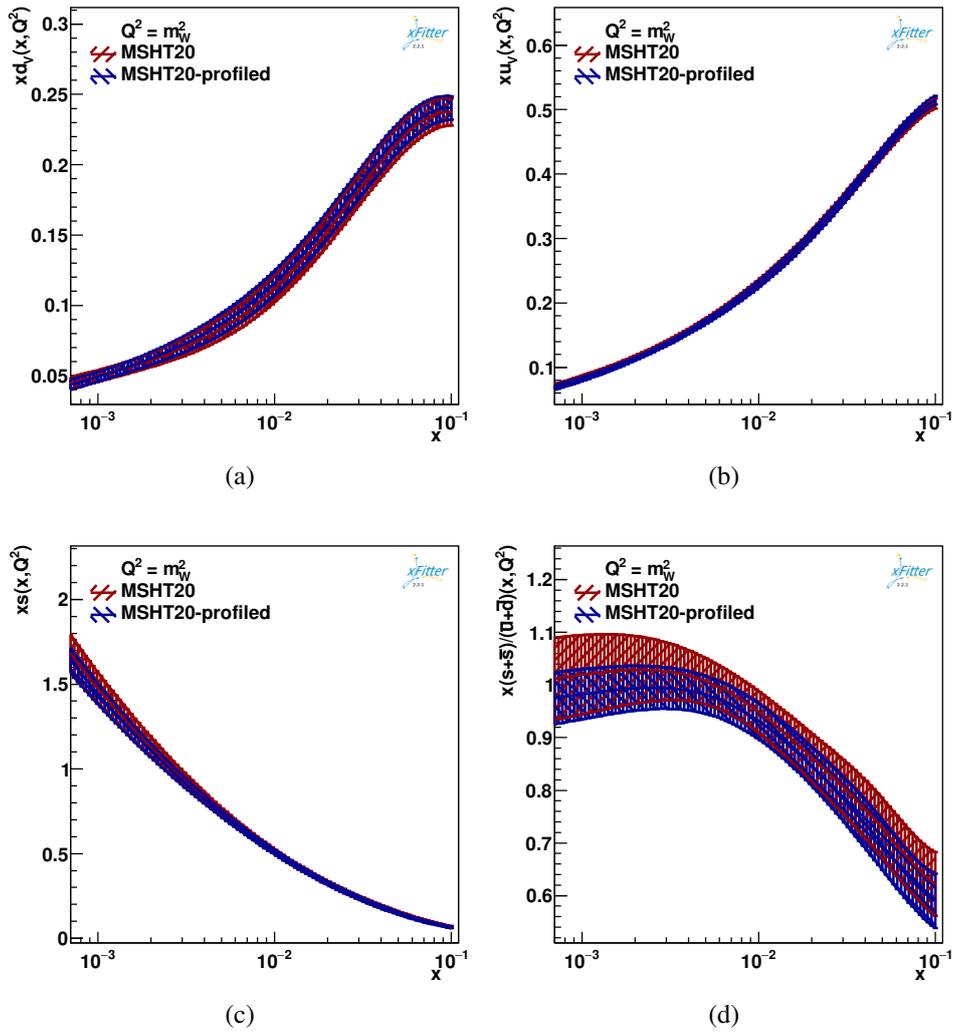


FIGURE 6.35 – Profiled (blue) and original (red) PDF distributions of the MSHT20 PDF set at  $Q^2 = m_W^2$  using the measured differential cross sections in  $|\eta_\ell|$  at 5.02 and 13 TeV. The distributions of (a)  $x d_v$ , (b)  $x u_v$ , (c)  $x s$ , and (d)  $R_s$  PDFs are presented.

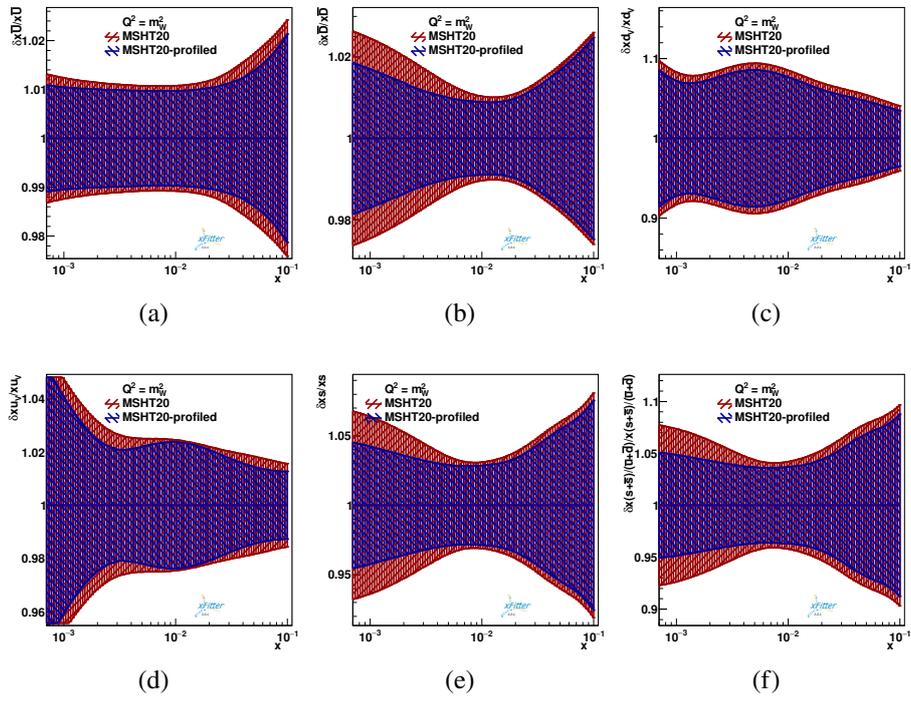


FIGURE 6.36 – The relative uncertainties of profiled (blue) and original (red) PDF distributions of the MSHT20 PDF set at  $Q^2 = m_W^2$  using the measured differential cross sections in  $|\eta_\ell|$  at 5.02 and 13 TeV. The relative uncertainties of (a)  $x\bar{U}$ , (b)  $x\bar{D}$ , (c)  $x d_v$ , (d)  $x u_v$ , (e)  $x_s$  and (f)  $R_s$  PDFs are presented.

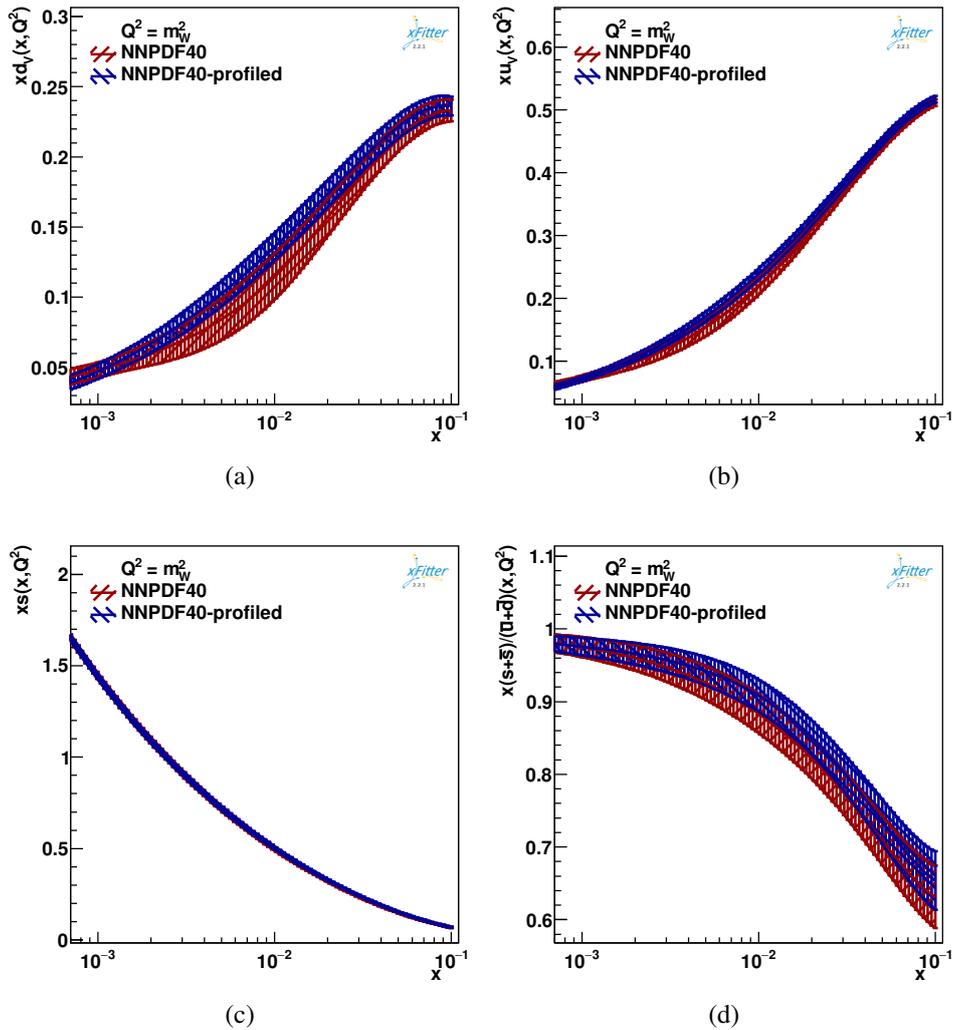


FIGURE 6.37 – Profiled (blue) and original (red) PDF distributions of the NNPDF40 PDF set at  $Q^2 = m_W^2$  using the measured differential cross sections in  $|\eta_\ell|$  at 5.02 and 13 TeV. The distributions of (a)  $x d_v$ , (b)  $x u_v$ , (c)  $x s$ , and (d)  $R_s$  PDFs are presented.

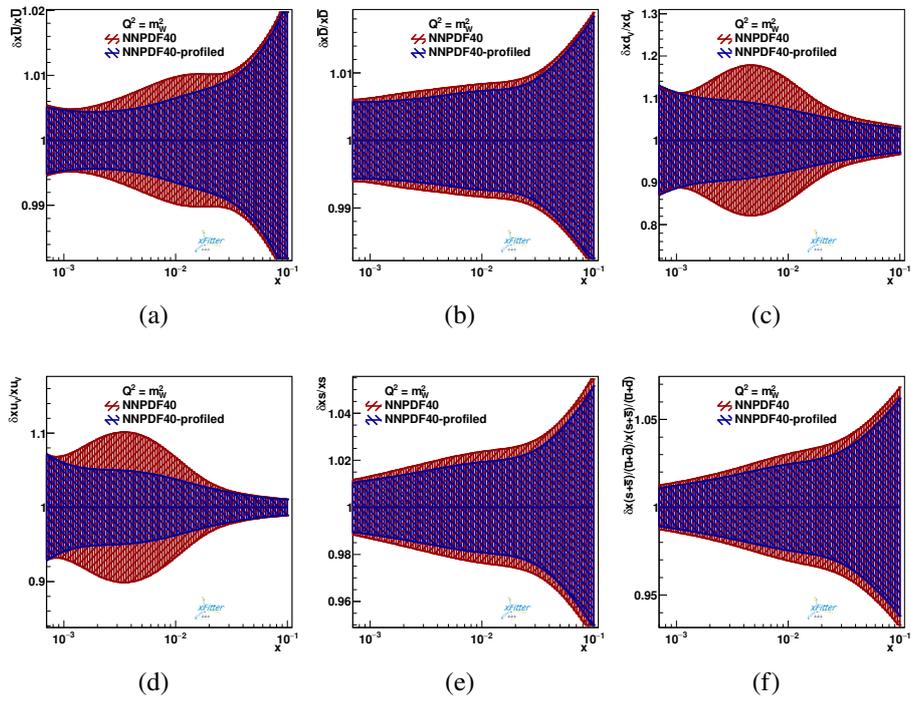


FIGURE 6.38 – The relative uncertainties of profiled (blue) and original (red) PDF distributions of the NNP40 PDF set at  $Q^2 = m_W^2$  using the measured differential cross sections in  $|\eta_\ell|$  at 5.02 and 13 TeV. The relative uncertainties of (a)  $x\bar{U}$ , (b)  $x\bar{D}$ , (c)  $x d_v$ , (d)  $x u_v$ , (e)  $x_s$  and (f)  $R_s$  PDFs are presented.

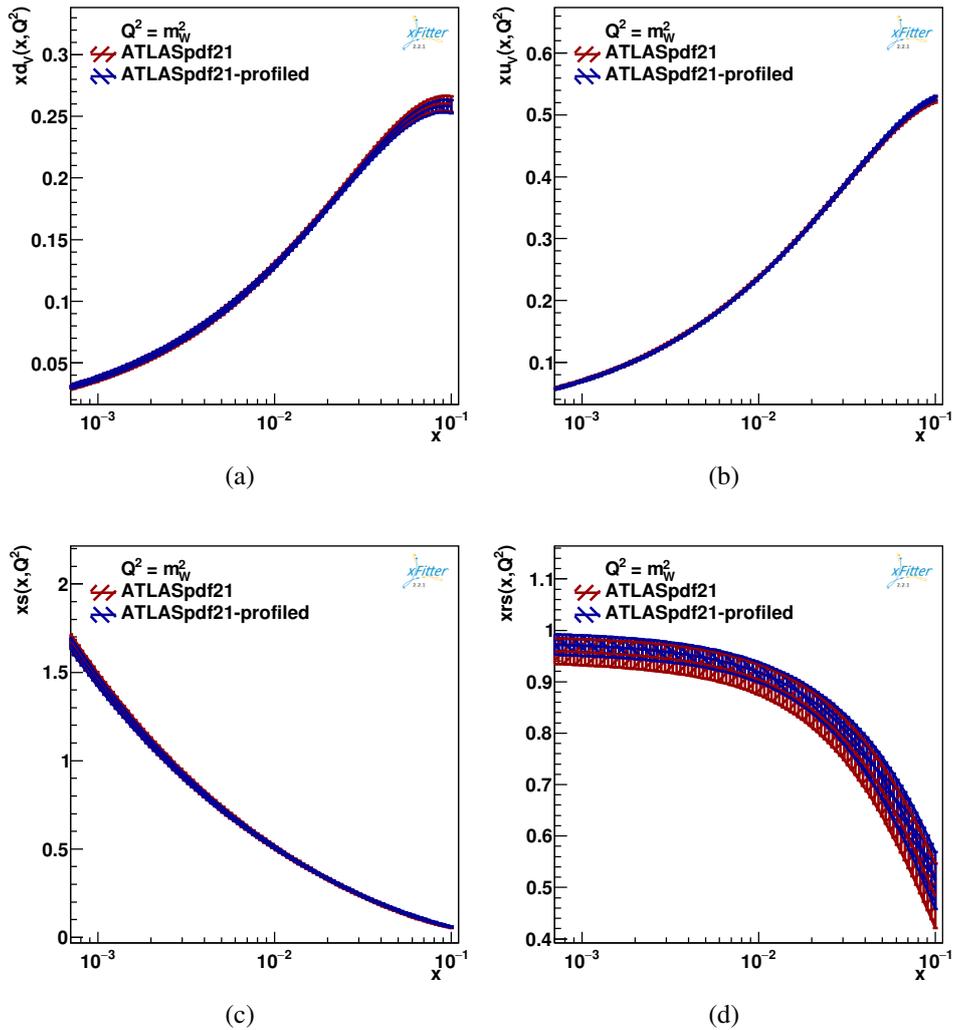


FIGURE 6.39 – Profiled (blue) and original (red) PDF distributions of the ATLASpdf21 PDF set at  $Q^2 = m_W^2$  using the measured differential cross sections in  $|\eta_\ell|$  at 5.02 and 13 TeV. The distributions of (a)  $x d_v$ , (b)  $x u_v$ , (c)  $x s$ , and (d)  $R_s$  PDFs are presented.

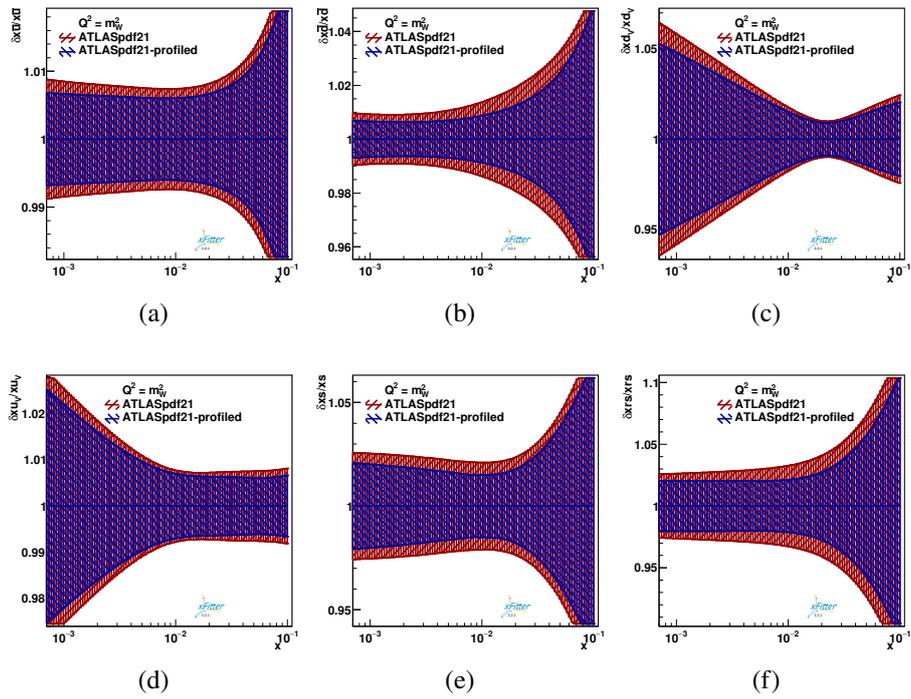


FIGURE 6.40 – The relative uncertainties of profiled (blue) and original (red) PDF distributions of the ATLASpdf21 PDF set at  $Q^2 = m_W^2$  using the measured differential cross sections in  $|\eta_\ell|$  at 5.02 and 13 TeV. The relative uncertainties of (a)  $x\bar{U}$ , (b)  $x\bar{D}$ , (c)  $x d_v$ , (d)  $x u_v$ , (e)  $x_s$  and (f)  $R_s$  PDFs are presented.

## 7 - Summary and prospect

This thesis presents two measurements of vector boson production cross sections in the leptonic decay channels.

The first measurement involves the inclusive  $W$  and  $Z$  boson production cross sections and their ratios using a dataset from  $pp$  collisions at a center-of-mass energy of 13.6 TeV, corresponding to an integrated luminosity of  $29 \text{ fb}^{-1}$ . The fiducial volume for the  $Z$  boson cross-section measurement is defined as lepton transverse momentum  $p_{\text{T}}^{\ell} > 27 \text{ GeV}$ , lepton pseudo-rapidity  $|\eta| < 2.5$ , and invariant mass of di-lepton  $66 < m_{\ell\ell} < 116 \text{ GeV}$ . For  $W$  boson production, the cross section is measured in the fiducial phase space with  $p_{\text{T}}^{\ell} > 27 \text{ GeV}$ ,  $|\eta| < 2.5$ , transverse mass  $m_{\text{T}} > 50 \text{ GeV}$ , and missing transverse energy  $E_{\text{T}}^{\text{miss}} > 25 \text{ GeV}$ . The total cross section for  $Z$  boson production is measured within the mass range  $66 < m < 116 \text{ GeV}$ , while the  $W$  boson production cross section is measured across the entire phase space. In addition to the  $W$  and  $Z$  boson cross sections, the ratios of fiducial cross section  $W^{+}$  to  $W^{-}$ ,  $W$  to  $Z$ , and  $t\bar{t}$  to  $W$  boson fiducial cross sections are also measured in this analysis.

The results are compared with the Standard Model predictions, which are computed at next-to-next-to-leading-order (NNLO) with next-to-next-to-leading-log (NNLL) QCD accuracy and next-to-leading-order (NLO) precision in electroweak calculations with various parton distribution functions (PDFs). Overall, there is a good agreement between the measured results and the theoretical predictions. However, it is noted that the ratio of  $t\bar{t}/W$  is slightly larger than the values predicted by some of the PDF calculations.

This analysis involves the full dataset recorded during 2022, the first year of Run 3 data-taking. Measuring the well-understood  $W$  and  $Z$  boson processes aids in evaluating the performance of the detector and reconstruction frameworks during this early period of Run 3. Given the large production rates of  $W$  and  $Z$  bosons at the LHC, the statistics for this measurement is robust. Consequently, the reported results are primarily limited by systematic uncertainties rather than statistical ones.

The second measurement focuses on the differential cross section of  $W$  boson production, using dedicated datasets recorded by the ATLAS detector during special LHC runs characterized by low pile-up conditions, at center-of-mass energies of 5.02 and 13 TeV, corresponding to integrated luminosities of  $255 \text{ pb}^{-1}$  and  $338 \text{ pb}^{-1}$ , respectively. The measurement is performed in the fiducial phase space defined by  $p_{\text{T}}^{\ell} > 25 \text{ GeV}$ ,  $|\eta| < 2.5$ ,  $m_{\text{T}} > 50 \text{ GeV}$ , and  $E_{\text{T}}^{\text{miss}} > 25 \text{ GeV}$ . The single-differential cross sections of the  $W$  boson production is provided as a function of  $p_{\text{T}}^{\ell}$  and  $|\eta_{\ell}|$ , respectively. The double-differential cross sections are also measured as functions of  $p_{\text{T}}^{\ell}$  and  $|\eta_{\ell}|$ . Additionally, the lepton charge asymmetry is also reported as a function of  $|\eta_{\ell}|$  in this analysis.

The results are obtained in the electron and muon channels, respectively, and show good compatibility between the two lepton-flavor channels. The combination

procedures are applied to the electron and muon channels. The combined results are then compared with theoretical predictions at NNLO+NNLL QCD accuracy using recent PDF sets. The measured results agree well with the Standard Model predictions. However, the predictions are slightly underestimated in the high  $p_T^\ell$  region. Thanks to the low pile-up datasets, the systematic uncertainties are significantly reduced, allowing for the derivation of precise differential cross section results. The impact of the measured cross sections on the PDFs is evaluated using profiling techniques. The uncertainties in some quark distributions are reduced by incorporating the new results.

The first analysis is limited systematically. As the measurement is performed using the early Run 3 datasets, some uncertainties involved in this analysis, such as the uncertainties related to the lepton corrections and calibrations, are estimated using Run 2 datasets with conservative uncertainties. These uncertainties are expected to be reduced with the efforts from the performance groups in the further study using more Run 3 datasets. The second analysis involves low pile-up datasets collected from special LHC runs during Run 2, with lower integrated luminosity in comparison with the high- $\mu$  Run 2 datasets. It is expected or desirable to have more low pile-up data taking during Run 3, thus the cross sections can be measured more precisely as data accumulates. Furthermore, the continuous improvement in performance and analysis techniques will enhance the precision and robustness. The precise measurement of vector boson production cross sections offers opportunities to probe the current Standard Model theory, especially perturbative QCD. Moreover, these measurements can enhance our understanding of proton structure, offering insights into the distribution of quarks and gluons within the proton. This enhanced understanding can benefit other analyses, for example, by providing accurate background modeling and reducing the uncertainties related to PDFs.

## Bibliographie

- [1] J. J. THOMSON, *Cathode rays*, *Phil. Mag. Ser. 5* **44** (1897) p. 293 (cf. p. 15).
- [2] E. RUTHERFORD, *The scattering of alpha and beta particles by matter and the structure of the atom*, *Phil. Mag. Ser. 6* **21** (1911) p. 669 (cf. p. 15).
- [3] J. CHADWICK, *Possible Existence of a Neutron*, *Nature* **129** (1932) p. 312 (cf. p. 15).
- [4] Chen-Ning YANG et Robert L. MILLS, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, *Phys. Rev.* **96** (1954) p. 191, sous la dir. de Jong-Ping HSU et D. FINE (cf. p. 15).
- [5] S. L. GLASHOW, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961) p. 579 (cf. p. 15).
- [6] Steven WEINBERG, *A Model of Leptons*, *Phys. Rev. Lett.* **19** (1967) p. 1264 (cf. p. 15).
- [7] Abdus SALAM, *Weak and Electromagnetic Interactions*, *Conf. Proc. C* **680519** (1968) p. 367 (cf. p. 15).
- [8] Murray GELL-MANN, *A Schematic Model of Baryons and Mesons*, *Phys. Lett.* **8** (1964) p. 214 (cf. p. 15).
- [9] G. ZWEIG, "An SU(3) model for strong interaction symmetry and its breaking. Version 2", *DEVELOPMENTS IN THE QUARK THEORY OF HADRONS. VOL. 1. 1964 - 1978*, sous la dir. de D. B. LICHTENBERG et Simon Peter ROSEN, 1964 p. 22 (cf. p. 15).
- [10] Peter W. HIGGS, *Broken symmetries, massless particles and gauge fields*, *Phys. Lett.* **12** (1964) p. 132 (cf. p. 15).
- [11] Peter W. HIGGS, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) p. 508, sous la dir. de J. C. TAYLOR (cf. p. 15).
- [12] F. ENGLERT et R. BROUT, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) p. 321, sous la dir. de J. C. TAYLOR (cf. p. 15).

- [13] G. S. GURALNIK, C. R. HAGEN et T. W. B. KIBBLE,  
*Global Conservation Laws and Massless Particles*,  
**Phys. Rev. Lett.** **13** (1964) p. 585, sous la dir. de J. C. TAYLOR  
(cf. p. 15).
- [14] Jerome I. FRIEDMAN, Henry W. KENDALL et Richard E. TAYLOR,  
*Deep inelastic scattering of electrons by protons and the structure of  
the nucleon*, *Reviews of Modern Physics* **63** (1991) p. 615 (cf. p. 15).
- [15] J. J. AUBERT et al., *Experimental Observation of a Heavy Particle J*,  
*Physical Review Letters* **33** (1974) p. 1404 (cf. p. 15).
- [16] G. ARNISON et al., *Experimental Observation of Lepton Pairs of  
Invariant Mass Around 95-GeV/c\*\*2 at the CERN SPS Collider*,  
**Phys. Lett. B** **126** (1983) p. 398 (cf. p. 15, 65).
- [17] P. BAGNAIA et al., *Evidence for  $Z^0 \rightarrow e^+e^-$  at the CERN  $\bar{p}p$  Collider*,  
**Phys. Lett. B** **129** (1983) p. 130 (cf. p. 15, 65).
- [18] G. ARNISON et al.,  
*Experimental Observation of Isolated Large Transverse Energy  
Electrons with Associated Missing Energy at  $\sqrt{s} = 540$  GeV*,  
**Phys. Lett. B** **122** (1983) p. 103 (cf. p. 15, 65).
- [19] M. BANNER et al., *Observation of Single Isolated Electrons of High  
Transverse Momentum in Events with Missing Transverse Energy at  
the CERN anti-p p Collider*, **Phys. Lett. B** **122** (1983) p. 476  
(cf. p. 15, 65).
- [20] F. ABE et others (CDF COLLABORATION), *Observation of top quark  
production in  $\bar{p}p$  collisions with the collider detector at Fermilab*,  
*Physical Review Letters* **74** (1995) p. 2626 (cf. p. 15).
- [21] B. ABBOTT et others (DØ COLLABORATION),  
*Observation of the top quark*,  
*Physical Review Letters* **74** (1995) p. 2632 (cf. p. 15).
- [22] Georges AAD et al., *A search for new resonances in multiple final  
states with a high transverse momentum Z boson in  $\sqrt{s} = 13$  TeV pp  
collisions with the ATLAS detector*, **JHEP** **06** (2023) p. 036,  
arXiv : 2209.15345 [hep-ex] (cf. p. 17, 211).
- [23] M TANABASHI et al., *Review of Particle Physics, 2018-2019*,  
**Phys. Rev. D** **98** (2018) p. 030001,  
URL : <https://cds.cern.ch/record/2636832> (cf. p. 20).
- [24] PARTICLE DATA GROUP, P.A. ZYLA et al., *Review of Particle Physics*,  
**PTEP** **2020** (2020) p. 083C01 (cf. p. 21).

- [25] John ELLIS, Mary K. GAILLARD et Dimitri V. NANOPOULOS, *A Historical Profile of the Higgs Boson. An Updated Historical Profile of the Higgs Boson*, (2016) p. 255, arXiv : 1504.07217, URL : <https://cds.cern.ch/record/2012465> (cf. p. 24).
- [26] T. GLEISBERG et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) p. 007, arXiv : 0811.4622 [hep-ph] (cf. p. 27).
- [27] J. D. BJORKEN et E. A. PASCHOS, *Inelastic Electron-Proton and  $\gamma$ -Proton Scattering and the Structure of the Nucleon*, *Phys. Rev.* **185** (5 1969) p. 1975, URL : <https://link.aps.org/doi/10.1103/PhysRev.185.1975> (cf. p. 26).
- [28] Yuri L. DOKSHITZER, *Calculation of the Structure Functions for Deep Inelastic Scattering and  $e^+e^-$  Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) p. 641 (cf. p. 26).
- [29] V. N. GRIBOV et L. N. LIPATOV, *Deep inelastic  $e p$  scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) p. 438 (cf. p. 26).
- [30] Guido ALTARELLI et G. PARISI, *Asymptotic Freedom in Parton Language*, *Nucl. Phys. B* **126** (1977) p. 298 (cf. p. 26).
- [31] H1 et ZEUS COLLABORATIONS, *Combination of Measurements of Inclusive Deep Inelastic  $e^\pm p$  Scattering Cross Sections and QCD Analysis of HERA Data*, 2015, arXiv : 1506.06042 [hep-ex], URL : <https://arxiv.org/abs/1506.06042> (cf. p. 28).
- [32] Joey HUSTON, Jon PUMPLIN, Daniel STUMP et Wu-Ki TUNG, *Stability of NLO global analysis and implications for Hadron collider physics*, *Journal of High Energy Physics* **2005** (2005) p. 080, ISSN : 1029-8479, URL : <http://dx.doi.org/10.1088/1126-6708/2005/06/080> (cf. p. 29).
- [33] Sergey ALEKHIN et al., *The PDF4LHC Working Group Interim Report*, 2011, arXiv : 1101.0536 [hep-ph], URL : <https://arxiv.org/abs/1101.0536> (cf. p. 29).
- [34] Tie-Jiun HOU et al., *New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC*, *Phys. Rev. D* **103** (2021) p. 014013, arXiv : 1912.10053 [hep-ph] (cf. p. 29, 83, 85, 106, 143).

- [35] T. CRIDGE, L. A. HARLAND-LANG, A. D. MARTIN et R. S. THORNE, *An investigation of the  $\alpha_S$  and heavy quark mass dependence in the MSHT20 global PDF analysis*, *Eur. Phys. J. C* **81** (2021) p. 744, arXiv : [2106.10289](https://arxiv.org/abs/2106.10289) [hep-ph] (cf. p. 29, 84, 85, 106, 143).
- [36] S. BAILEY, T. CRIDGE, L. A. HARLAND-LANG, A. D. MARTIN et R. S. THORNE, *Parton distributions from LHC, HERA, Tevatron and fixed target data : MSHT20 PDFs*, *Eur. Phys. J. C* **81** (2021) p. 341, arXiv : [2012.04684](https://arxiv.org/abs/2012.04684) [hep-ph] (cf. p. 29, 84, 85, 106, 143).
- [37] T. CRIDGE, L. A. HARLAND-LANG, A. D. MARTIN et R. S. THORNE, *QED parton distribution functions in the MSHT20 fit*, *Eur. Phys. J. C* **82** (2022) p. 90, arXiv : [2111.05357](https://arxiv.org/abs/2111.05357) [hep-ph] (cf. p. 29, 84, 85, 106, 143).
- [38] NNPDF COLLABORATION, Richard D. BALL et al., *The path to proton structure at 1% accuracy*, *Eur. Phys. J. C* **82** (2022) p. 428, arXiv : [2109.02653](https://arxiv.org/abs/2109.02653) [hep-ph] (cf. p. 29, 84, 85, 106, 143).
- [39] ATLAS COLLABORATION, *Determination of the parton distribution functions of the proton using diverse ATLAS data from pp collisions at  $\sqrt{s} = 7, 8$  and 13 TeV*, *Eur. Phys. J. C* **82** (2022) p. 438, arXiv : [2112.11266](https://arxiv.org/abs/2112.11266) [hep-ex] (cf. p. 29, 84, 85, 106, 143).
- [40] S. D. DRELL et Tung-Mow YAN, *Partons and their Applications at High-Energies*, *Annals Phys.* **66** (1971) p. 578 (cf. p. 29).
- [41] *Standard Model Summary Plots October 2023*, rapp. tech., All figures including auxiliary figures are available at <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2023-039> : CERN, 2023, URL : <http://cds.cern.ch/record/2882448> (cf. p. 32).
- [42] ATLAS COLLABORATION, *Measurement of  $W^\pm$  and Z-boson production cross sections in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Phys. Lett. B* **759** (2016) p. 601, arXiv : [1603.09222](https://arxiv.org/abs/1603.09222) [hep-ex] (cf. p. 33, 65, 71).
- [43] Lyndon R EVANS et Philip BRYANT, *LHC Machine*, *JINST* **3** (2008) S08001, This report is an abridged version of the LHC Design Report (CERN-2004-003), URL : <http://cds.cern.ch/record/1129806> (cf. p. 35).

- [44] Ewa LOPIENSKA, *The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022*, (2022), General Photo,  
URL : <https://cds.cern.ch/record/2800984> (cf. p. 36).
- [45] The ATLAS COLLABORATION,  
*The ATLAS Experiment at the CERN Large Hadron Collider*,  
**Journal of Instrumentation** **3** (2008) S08003,  
URL : <https://dx.doi.org/10.1088/1748-0221/3/08/S08003>  
(cf. p. 36).
- [46] CMS COLLABORATION, *The CMS Experiment at the CERN LHC*,  
**JINST** **3** (2008) S08004 (cf. p. 36, 65).
- [47] Riccardo Maria BIANCHI et ATLAS COLLABORATION,  
“ATLAS experiment schematic or layout illustration”,  
General Photo, 2022,  
URL : <http://cds.cern.ch/record/2837191> (cf. p. 37).
- [48] *ATLAS inner detector : Technical Design Report, 1*,  
Technical design report. ATLAS, Geneva : CERN, 1997,  
URL : <https://cds.cern.ch/record/331063> (cf. p. 38).
- [49] S HAYWOOD, L ROSSI, R NICKERSON et A ROMANIOUK,  
*ATLAS inner detector : Technical Design Report, 2*,  
Technical design report. ATLAS, Geneva : CERN, 1997,  
URL : <https://cds.cern.ch/record/331064> (cf. p. 38).
- [50] Joao PEQUENAO,  
“Computer generated image of the ATLAS inner detector”, 2008,  
URL : <https://cds.cern.ch/record/1095926> (cf. p. 38).
- [51] B. ABBOTT et al.,  
*Production and integration of the ATLAS Insertable B-Layer*,  
**JINST** **13** (2018) T05008,  
arXiv : [1803.00844](https://arxiv.org/abs/1803.00844) [physics.ins-det] (cf. p. 38, 39).
- [52] Joao PEQUENAO,  
“Computer Generated image of the ATLAS calorimeter”, 2008,  
URL : <https://cds.cern.ch/record/1095927> (cf. p. 40).
- [53] ATLAS COLLABORATION,  
*ATLAS Liquid Argon Calorimeter : Technical Design Report*,  
ATLAS-TDR-2; CERN-LHCC-96-041, 1996,  
URL : <https://cds.cern.ch/record/331061> (cf. p. 41-43).

- [54] ATLAS COLLABORATION,  
*ATLAS Tile Calorimeter : Technical Design Report*,  
 ATLAS-TDR-3 ; CERN-LHCC-96-042, 1996,  
 URL : <https://cds.cern.ch/record/331062> (cf. p. 42).
- [55] *Technical Design Report for the Phase-II Upgrade of the ATLAS Tile Calorimeter*, rapp. tech., CERN, 2017,  
 URL : <https://cds.cern.ch/record/2285583> (cf. p. 43).
- [56] *ATLAS muon spectrometer : Technical Design Report*,  
 Technical design report. ATLAS, Geneva : CERN, 1997,  
 URL : <https://cds.cern.ch/record/331068> (cf. p. 44).
- [57] Joao PEQUENAO,  
 “Computer generated image of the ATLAS Muons subsystem”, 2008,  
 URL : <https://cds.cern.ch/record/1095929> (cf. p. 44).
- [58] ATLAS COLLABORATION, *Performance of the ATLAS RPC detector and Level-1 muon barrel trigger at  $\sqrt{s} = 13$  TeV*,  
**JINST 16 (2021) P07029**, arXiv : [2103.01029](https://arxiv.org/abs/2103.01029) [[hep-ex](https://arxiv.org/archive/hep)] (cf. p. 45).
- [59] ATLAS COLLABORATION,  
*ATLAS New Small Wheel : Technical Design Report*,  
 ATLAS-TDR-020 ; CERN-LHCC-2013-006, 2013,  
 URL : <https://cds.cern.ch/record/1552862> (cf. p. 45, 46).
- [60] *ATLAS magnet system : Technical Design Report, 1*,  
 Technical design report. ATLAS, Geneva : CERN, 1997,  
 URL : <https://cds.cern.ch/record/338080> (cf. p. 45).
- [61] A. YAMAMOTO et al., *The ATLAS central solenoid*,  
**Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment**  
**584 (2008) p. 53**, ISSN : 0168-9002,  
 URL : <https://www.sciencedirect.com/science/article/pii/S0168900207020414> (cf. p. 45, 46).
- [62] Ana Maria RODRIGUEZ VERA et Joao ANTUNES PEQUENAO,  
 “ATLAS Detector Magnet System”, General Photo, 2021,  
 URL : <https://cds.cern.ch/record/2770604> (cf. p. 46).
- [63] G. AVONI et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, **JINST 13 (2018) P07017**  
 (cf. p. 47).
- [64] ATLAS COLLABORATION, *Luminosity Public Results : Run 2*, 2024,  
 URL : <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2> (cf. p. 47).

- [65] ATLAS COLLABORATION, *Luminosity Public Results : Run 3*, 2024, URL : <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun3> (cf. p. 47).
- [66] ATLAS COLLABORATION, *ATLAS High-Level Trigger, Data Acquisition and Controls : Technical Design Report*, ATLAS-TDR-16; CERN-LHCC-2003-022, 2003, URL : <https://cds.cern.ch/record/616089> (cf. p. 47).
- [67] ATLAS COLLABORATION, *Performance of the ATLAS trigger system in 2015*, *Eur. Phys. J. C* **77** (2017) p. 317, arXiv : 1611.09661 [hep-ex] (cf. p. 48).
- [68] D WICKE, *A New Algorithm for Solving Tracking Ambiguities*, (1998), URL : <https://cds.cern.ch/record/2625731> (cf. p. 51).
- [69] T. CORNELISSEN et al., *The new ATLAS track reconstruction (NEWT)*, *J. Phys. Conf. Ser.* **119** (2008) p. 032014, sous la dir. de Randall SOBIE, Reda TAFIROUT et Jana THOMSON (cf. p. 51).
- [70] ATLAS COLLABORATION, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC run 2. Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC run 2*, *Eur. Phys. J. C* **77** (2017) p. 673, arXiv : 1704.07983, URL : <https://cds.cern.ch/record/2261156> (cf. p. 51).
- [71] R. FRÜHWIRTH, *Application of Kalman filtering to track and vertex fitting*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **262** (1987) p. 444, ISSN : 0168-9002, URL : <https://www.sciencedirect.com/science/article/pii/0168900287908874> (cf. p. 51).
- [72] Morad AABOUD et al., *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, *Eur. Phys. J. C* **77** (2017) p. 332, arXiv : 1611.10235 [physics.ins-det] (cf. p. 51).
- [73] ATLAS COLLABORATION, *Development of ATLAS Primary Vertex Reconstruction for LHC Run 3*, ATLAS-PHYS-PUB-2019-015, 2019, URL : <https://cds.cern.ch/record/2670380> (cf. p. 51).

- [74] ATLAS COLLABORATION, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J. C* **79** (2019) p. 639, arXiv : 1902.04655 [physics.ins-det] (cf. p. 52, 53, 68, 132).
- [75] ATLAS COLLABORATION, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, *JINST* **14** (2019) P12006, arXiv : 1908.00005 [hep-ex] (cf. p. 52-54, 68, 89, 132, 152, 217, 218).
- [76] ATLAS COLLABORATION, *Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment*, *Journal of High Energy Physics* **2024** (2024), arXiv : 2308.13362 [hep-ex] (cf. p. 52, 215).
- [77] ATLAS COLLABORATION, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, *Eur. Phys. J. C* **77** (2017) p. 490, arXiv : 1603.02934 [hep-ex] (cf. p. 52, 58).
- [78] ATLAS COLLABORATION, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012, URL : <https://cds.cern.ch/record/1449796> (cf. p. 53).
- [79] ATLAS COLLABORATION, *Electron and photon energy calibration with the ATLAS detector using LHC Run 2 data*, *JINST* **19** (2023) P02009, arXiv : 2309.05471 [hep-ex] (cf. p. 54, 55, 217).
- [80] ATLAS COLLABORATION, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J. C* **81** (2021) p. 578, arXiv : 2012.00578 [hep-ex] (cf. p. 55, 56, 58, 89, 132, 152).
- [81] J. ILLINGWORTH et J. KITTLER, *A survey of the hough transform*, *Computer Vision, Graphics, and Image Processing* **44** (1988) p. 87, ISSN : 0734-189X, URL : <https://www.sciencedirect.com/science/article/pii/S0734189X88800331> (cf. p. 55).
- [82] ATLAS COLLABORATION, *Studies of the muon momentum calibration and performance of the ATLAS detector with pp collisions at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J. C* **83** (2023) p. 686, arXiv : 2212.07338 [hep-ex] (cf. p. 56, 57).

- [83] Matteo CACCIARI, Gavin P. SALAM et Gregory SOYEZ, *The anti- $k_t$  jet clustering algorithm*, *JHEP* **04** (2008) p. 063, arXiv : [0802.1189 \[hep-ph\]](#) (cf. p. 57, 68, 211).
- [84] Matteo CACCIARI, Gavin P. SALAM et Gregory SOYEZ, *FastJet user manual*, *Eur. Phys. J. C* **72** (2012) p. 1896, arXiv : [1111.6097 \[hep-ph\]](#) (cf. p. 57).
- [85] ATLAS COLLABORATION, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, *Eur. Phys. J. C* **77** (2017) p. 466, arXiv : [1703.10485 \[hep-ex\]](#) (cf. p. 58, 68, 211).
- [86] ATLAS COLLABORATION, *Jet energy scale and resolution measured in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, *Eur. Phys. J. C* **81** (2021) p. 689, arXiv : [2007.02645 \[hep-ex\]](#) (cf. p. 59, 68, 89, 90).
- [87] ATLAS COLLABORATION, *ATLAS flavour-tagging algorithms for the LHC Run 2  $pp$  collision dataset*, *Eur. Phys. J. C* **83** (2023) p. 681, arXiv : [2211.16345 \[physics.data-an\]](#) (cf. p. 60, 61, 69).
- [88] ATLAS COLLABORATION, *Deep Sets based Neural Networks for Impact Parameter Flavour Tagging in ATLAS*, ATL-PHYS-PUB-2020-014, 2020, URL : <https://cds.cern.ch/record/2718948> (cf. p. 60, 69).
- [89] ATLAS COLLABORATION, “ATLAS Event Display : W boson decays to a muon and a neutrino”, General Photo, 2023, URL : <https://cds.cern.ch/record/2859307> (cf. p. 62).
- [90] ATLAS COLLABORATION, *The performance of missing transverse momentum reconstruction and its significance with the ATLAS detector using  $140\text{ fb}^{-1}$  of  $\sqrt{s} = 13$  TeV  $pp$  collisions*, (2024), arXiv : [2402.05858 \[hep-ex\]](#) (cf. p. 61).
- [91] ATLAS COLLABORATION, *Performance of pile-up mitigation techniques for jets in  $pp$  collisions at  $\sqrt{s} = 8$  TeV using the ATLAS detector*, *Eur. Phys. J. C* **76** (2016) p. 581, arXiv : [1510.03823 \[hep-ex\]](#) (cf. p. 63, 68).
- [92] ATLAS COLLABORATION, *Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector*, *Eur. Phys. J. C* **77** (2017) p. 580, arXiv : [1705.02211 \[hep-ex\]](#) (cf. p. 63).

- [93] ATLAS COLLABORATION,  
*The ATLAS Experiment at the CERN Large Hadron Collider*,  
**JINST** **3** (2008) S08003 (cf. p. 65).
- [94] ATLAS COLLABORATION,  
*Measurement of  $W^\pm$ -boson and  $Z$ -boson production cross-sections in  $pp$  collisions at  $\sqrt{s} = 2.76$  TeV with the ATLAS detector*,  
**Eur. Phys. J. C** **79** (2019) p. 901, arXiv : 1907.03567 [hep-ex]  
(cf. p. 65).
- [95] ATLAS COLLABORATION,  
*Measurements of  $W$  and  $Z$  boson production in  $pp$  collisions at  $\sqrt{s} = 5.02$  TeV with the ATLAS detector*,  
**Eur. Phys. J. C** **79** (2019) p. 128, arXiv : 1810.08424 [hep-ex]  
(cf. p. 65).
- [96] ATLAS COLLABORATION,  
*Precision measurement and interpretation of inclusive  $W^+$ ,  $W^-$  and  $Z/\gamma^*$  production cross sections with the ATLAS detector*,  
**Eur. Phys. J. C** **77** (2017) p. 367, arXiv : 1612.03016 [hep-ex]  
(cf. p. 65, 82, 84, 150, 162).
- [97] ATLAS COLLABORATION, *Measurement of the Drell–Yan triple-differential cross section in  $pp$  collisions at  $\sqrt{s} = 8$  TeV*,  
**JHEP** **12** (2017) p. 059, arXiv : 1710.05167 [hep-ex] (cf. p. 65).
- [98] ATLAS COLLABORATION, *Measurement of the cross-section and charge asymmetry of  $W$  bosons produced in proton–proton collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector*,  
**Eur. Phys. J. C** **79** (2019) p. 760, arXiv : 1904.05631 [hep-ex]  
(cf. p. 65).
- [99] ATLAS COLLABORATION, *Measurement of the transverse momentum distribution of Drell–Yan lepton pairs in proton–proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*,  
**Eur. Phys. J. C** **80** (2020) p. 616, arXiv : 1912.02844 [hep-ex]  
(cf. p. 65).
- [100] CMS COLLABORATION,  
*Measurement of the inclusive  $W$  and  $Z$  production cross sections in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the CMS experiment*,  
**JHEP** **10** (2011) p. 132, arXiv : 1107.4789 [hep-ex] (cf. p. 65).
- [101] CMS COLLABORATION,  
*Measurement of the differential and double-differential Drell–Yan cross sections in proton–proton collisions at  $\sqrt{s} = 7$  TeV*,  
**JHEP** **12** (2013) p. 030, arXiv : 1310.7291 [hep-ex] (cf. p. 65).

- [102] CMS COLLABORATION, *Measurement of inclusive  $W$  and  $Z$  boson production cross sections in  $pp$  collisions at  $\sqrt{s} = 8$  TeV*, *Phys. Rev. Lett.* **112** (2014) p. 191802, arXiv : 1402.0923 [hep-ex] (cf. p. 65).
- [103] CMS COLLABORATION, *Measurements of differential and double-differential Drell-Yan cross sections in proton–proton collisions at  $\sqrt{s} = 8$  TeV*, *Eur. Phys. J. C* **75** (2015) p. 147, arXiv : 1412.1115 [hep-ex] (cf. p. 65).
- [104] CMS COLLABORATION, *Measurements of the  $W$  boson rapidity, helicity, double-differential cross sections, and charge asymmetry in  $pp$  collisions at  $\sqrt{s} = 13$  TeV*, *Phys. Rev. D* **102** (2020) p. 092012, arXiv : 2008.04174 [hep-ex] (cf. p. 65).
- [105] ATLAS COLLABORATION, *Measurement of the  $t\bar{t}$  cross section and its ratio to the  $Z$  production cross section using  $pp$  collisions at  $\sqrt{s} = 13.6$  TeV with the ATLAS detector*, *Phys. Lett. B* **848** (2024) p. 138376, arXiv : 2308.09529 [hep-ex] (cf. p. 66, 70, 89, 90, 96, 97, 100, 103, 105, 106).
- [106] Enrico BOTHMANN et al., *Event generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) p. 034, arXiv : 1905.09127 [hep-ph] (cf. p. 66).
- [107] Federico BUCCIONI et al., *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019) p. 866, arXiv : 1907.13071 [hep-ph] (cf. p. 66).
- [108] Fabio CASCIOLI, Philipp MAIERHÖFER et Stefano POZZORINI, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) p. 111601, arXiv : 1111.5206 [hep-ph] (cf. p. 66).
- [109] Ansgar DENNER, Stefan DITTMAYER et Lars HOFER, *COLLIER : A fortran-based complex one-loop library in extended regularizations*, *Comput. Phys. Commun.* **212** (2017) p. 220, arXiv : 1604.06792 [hep-ph] (cf. p. 66).
- [110] Steffen SCHUMANN et Frank KRAUSS, *A parton shower algorithm based on Catani–Seymour dipole factorisation*, *JHEP* **03** (2008) p. 038, arXiv : 0709.1027 [hep-ph] (cf. p. 66, 131).
- [111] Stefan HÖCHE, Frank KRAUSS, Marek SCHÖNHERR et Frank SIEGERT, *QCD matrix elements + parton showers. The NLO case*, *JHEP* **04** (2013) p. 027, arXiv : 1207.5030 [hep-ph] (cf. p. 66, 131).

- [112] Richard D. BALL et al., *Parton distributions for the LHC run II*, **JHEP 04 (2015) p. 040**, arXiv : [1410.8849 \[hep-ph\]](#) (cf. p. 66, 131).
- [113] Charalampos ANASTASIOU, Lance DIXON, Kirill MELNIKOV et Frank PETRIELLO, *High-precision QCD at hadron colliders : Electroweak gauge boson rapidity distributions at next-to-next-to leading order*, **Phys. Rev. D 69 (2004) p. 094008**, arXiv : [hep-ph/0312266](#) (cf. p. 66).
- [114] Paolo NASON, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, **JHEP 11 (2004) p. 040**, arXiv : [hep-ph/0409146](#) (cf. p. 66, 130).
- [115] Stefano FRIXIONE, Paolo NASON et Carlo OLEARI, *Matching NLO QCD computations with parton shower simulations : the POWHEG method*, **JHEP 11 (2007) p. 070**, arXiv : [0709.2092 \[hep-ph\]](#) (cf. p. 66, 130).
- [116] Simone ALIOLI, Paolo NASON, Carlo OLEARI et Emanuele RE, *A general framework for implementing NLO calculations in shower Monte Carlo programs : the POWHEG BOX*, **JHEP 06 (2010) p. 043**, arXiv : [1002.2581 \[hep-ph\]](#) (cf. p. 66, 130).
- [117] Torbjörn SJÖSTRAND et al., *An introduction to PYTHIA 8.2*, **Comput. Phys. Commun. 191 (2015) p. 159**, arXiv : [1410.3012 \[hep-ph\]](#) (cf. p. 66, 130).
- [118] ATLAS COLLABORATION, *Measurement of the  $Z/\gamma^*$  boson transverse momentum distribution in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, **JHEP 09 (2014) p. 145**, arXiv : [1406.3660 \[hep-ex\]](#) (cf. p. 66, 130).
- [119] H.-L. LAI et al., *New parton distributions for collider physics*, **Phys. Rev. D 82 (2010) p. 074024**, arXiv : [1007.2241 \[hep-ph\]](#) (cf. p. 66, 130).
- [120] J. PUMPLIN et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, **JHEP 07 (2002) p. 012**, arXiv : [hep-ph/0201195](#) (cf. p. 66).
- [121] Piotr GOLONKA et Zbigniew WAS, *PHOTOS Monte Carlo : a precision tool for QED corrections in Z and W decays*, **Eur. Phys. J. C 45 (2006) p. 97**, arXiv : [hep-ph/0506026](#) (cf. p. 66, 84, 130).

- [122] N. DAVIDSON, T. PRZEDZINSKI et Z. WAS, *PHOTOS Interface in C++ : Technical and physics documentation*, *Comput. Phys. Commun.* **199** (2016) p. 86, arXiv : [1011.0937 \[hep-ph\]](#) (cf. p. 66).
- [123] ATLAS COLLABORATION, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, 2014, URL : <https://cds.cern.ch/record/1966419> (cf. p. 67).
- [124] M. BENEKE, P. FALGARI, S. KLEIN et C. SCHWINN, *Hadronic top-quark pair production with NNLL threshold resummation*, *Nucl. Phys. B* **855** (2012) p. 695, arXiv : [1109.1536 \[hep-ph\]](#) (cf. p. 67, 87, 131).
- [125] Matteo CACCIARI, Michal CZAKON, Michelangelo MANGANO, Alexander MITOV et Paolo NASON, *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, *Phys. Lett. B* **710** (2012) p. 612, arXiv : [1111.5869 \[hep-ph\]](#) (cf. p. 67, 87, 131).
- [126] Peter BÄRNREUTHER, Michal CZAKON et Alexander MITOV, *Percent-Level-Precision Physics at the Tevatron : Next-to-Next-to-Leading Order QCD Corrections to  $q\bar{q} \rightarrow t\bar{t} + X$* , *Phys. Rev. Lett.* **109** (2012) p. 132001, arXiv : [1204.5201 \[hep-ph\]](#) (cf. p. 67, 87, 131).
- [127] Michal CZAKON et Alexander MITOV, *NNLO corrections to top-pair production at hadron colliders : the all-fermionic scattering channels*, *JHEP* **12** (2012) p. 054, arXiv : [1207.0236 \[hep-ph\]](#) (cf. p. 67, 87, 131).
- [128] Michal CZAKON et Alexander MITOV, *NNLO corrections to top pair production at hadron colliders : the quark-gluon reaction*, *JHEP* **01** (2013) p. 080, arXiv : [1210.6832 \[hep-ph\]](#) (cf. p. 67, 87, 131).
- [129] Michal CZAKON, Paul FIEDLER et Alexander MITOV, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through  $O(\alpha_S^4)$* , *Phys. Rev. Lett.* **110** (2013) p. 252004, arXiv : [1303.6254 \[hep-ph\]](#) (cf. p. 67, 87, 131).
- [130] Michal CZAKON et Alexander MITOV, *Top++ : A program for the calculation of the top-pair cross-section at hadron colliders*, *Comput. Phys. Commun.* **185** (2014) p. 2930, arXiv : [1112.5675 \[hep-ph\]](#) (cf. p. 67, 87, 131).

- [131] ATLAS COLLABORATION, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at  $\sqrt{s} = 13$  TeV*, *Eur. Phys. J. C* **76** (2016) p. 292, arXiv : 1603.05598 [hep-ex] (cf. p. 68).
- [132] Roger BARLOW et Christine BEESTON, *Fitting using finite Monte Carlo samples*, *Computer Physics Communications* **77** (1993) p. 219, ISSN : 0010-4655, URL : <https://www.sciencedirect.com/science/article/pii/001046559390005W> (cf. p. 76, 95).
- [133] Kyle CRANMER, George LEWIS, LORENZO MONETA, Akira SHIBATA et Wouter VERKERKE, *HistFactory : A tool for creating statistical models for use with RooFit and RooStats*, rapp. tech., New York U., 2012, URL : <https://cds.cern.ch/record/1456844> (cf. p. 76, 95, 96).
- [134] Stefano CAMARDA et al., *DYTurbo : fast predictions for Drell-Yan processes*, *Eur. Phys. J. C* **80** (2020) p. 251, arXiv : 1910.07049 [hep-ph] (cf. p. 82, 143, 161), Erratum : *Eur. Phys. J. C* **80** (2020) p. 440.
- [135] Stefano CAMARDA, Leandro CIERI et Giancarlo FERRERA, *Drell–Yan lepton-pair production :  $q_T$  resummation at  $N^3LL$  accuracy and fiducial cross sections at  $N^3LO$* , *Phys. Rev. D* **104** (2021) p. L111503, arXiv : 2103.04974 [hep-ph] (cf. p. 82, 143, 161).
- [136] Stefano CAMARDA, Leandro CIERI et Giancarlo FERRERA, *Fiducial perturbative power corrections within the  $q_T$  subtraction formalism*, *Eur. Phys. J. C* **82** (2022) p. 575, arXiv : 2111.14509 [hep-ph] (cf. p. 82, 143, 161).
- [137] Stefano CAMARDA, Leandro CIERI et Giancarlo FERRERA, *Drell–Yan lepton-pair production :  $q_T$  resummation at  $N^4LL$  accuracy*, *Phys. Lett. B* **845** (2023) p. 138125, arXiv : 2303.12781 [hep-ph] (cf. p. 82, 143, 161).
- [138] Renat SADYKOV et Vitaly YERMOLCHYK, *Polarized NLO EW  $e^+e^-$  cross section calculations with ReneSANCe-v1.0.0*, *Comput. Phys. Commun.* **256** (2020) p. 107445, arXiv : 2001.10755 [hep-ph] (cf. p. 82).
- [139] Serge BONDARENKO, YAHOR DYDYSHKA, Lidia KALINOVSKAYA, Renat SADYKOV et Vitaly YERMOLCHYK, *Hadron-hadron collision mode in ReneSANCe-v1.3.0*, *Comput. Phys. Commun.* **285** (2023) p. 108646, arXiv : 2207.04332 [hep-ph] (cf. p. 82).

- [140] Stefan DITTMAYER et Max HUBER,  
*Radiative corrections to the neutral-current Drell-Yan process in the Standard Model and its minimal supersymmetric extension*,  
**Journal of High Energy Physics** **2010** (2010), ISSN : 1029-8479,  
 URL : [http://dx.doi.org/10.1007/JHEP01\(2010\)060](http://dx.doi.org/10.1007/JHEP01(2010)060)  
 (cf. p. 82, 143).
- [141] Richard D. BALL et al., *The PDF4LHC21 combination of global PDF fits for the LHC Run III*, **J. Phys. G** **49** (2022) p. 080501,  
 arXiv : [2203.05506 \[hep-ph\]](https://arxiv.org/abs/2203.05506) (cf. p. 83, 85, 86, 106).
- [142] S. ALEKHIN, J. BLÜMLEIN, S. MOCH et R. PLAČÁKYTĚ, *Parton distribution functions,  $\alpha_s$ , and heavy-quark masses for LHC Run II*,  
**Phys. Rev. D** **96** (2017) p. 014011, arXiv : [1701.05838 \[hep-ph\]](https://arxiv.org/abs/1701.05838)  
 (cf. p. 84, 85, 106).
- [143] Massimiliano GRAZZINI, Stefan KALLWEIT et Marius WIESEMANN,  
*Fully differential NNLO computations with MATRIX*,  
**Eur. Phys. J. C** **78** (2018) p. 537, arXiv : [1711.06631 \[hep-ph\]](https://arxiv.org/abs/1711.06631)  
 (cf. p. 87).
- [144] G. AVONI et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, **JINST** **13** (2018) P07017  
 (cf. p. 89, 130, 152).
- [145] Georges AAD et al., *Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment*, **JHEP** **05** (2024) p. 162,  
 arXiv : [2308.13362 \[hep-ex\]](https://arxiv.org/abs/2308.13362) (cf. p. 89, 152).
- [146] ATLAS COLLABORATION,  
*Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at  $\sqrt{s} = 13$  TeV*,  
**Eur. Phys. J. C** **78** (2018) p. 903, arXiv : [1802.08168 \[hep-ex\]](https://arxiv.org/abs/1802.08168)  
 (cf. p. 90).
- [147] F. JAMES et M. ROOS, *Minuit : A System for Function Minimization and Analysis of the Parameter Errors and Correlations*,  
**Comput. Phys. Commun.** **10** (1975) p. 343 (cf. p. 96).
- [148] ATLAS COLLABORATION, *Measurement of the  $t\bar{t}$  production cross-section and lepton differential distributions in  $e\mu$  dilepton events from  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*,  
**Eur. Phys. J. C** **80** (2020) p. 528, arXiv : [1910.08819 \[hep-ex\]](https://arxiv.org/abs/1910.08819)  
 (cf. p. 96).

- [149] Glen COWAN, Kyle CRANMER, Eilam GROSS et Ofer VITELLS, *Asymptotic formulae for likelihood-based tests of new physics*, *Eur. Phys. J. C* **71** (2011) p. 1554, arXiv : [1007.1727 \[physics.data-an\]](#) (cf. p. 96), Erratum : *Eur. Phys. J. C* **73** (2013) p. 2501.
- [150] Sayipjamal DULAT et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev. D* **93** (2016) p. 033006, arXiv : [1506.07443 \[hep-ph\]](#) (cf. p. 107).
- [151] ATLAS COLLABORATION, *Measurement of the  $W$ -boson mass in  $pp$  collisions at  $\sqrt{s} = 7$  TeV with the ATLAS detector*, *Eur. Phys. J. C* **78** (2018) p. 110, arXiv : [1701.07240 \[hep-ex\]](#) (cf. p. 129).
- [152] ATLAS COLLABORATION, *Luminosity determination in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector at the LHC*, *Eur. Phys. J. C* **83** (2023) p. 982, arXiv : [2212.09379 \[hep-ex\]](#) (cf. p. 130).
- [153] Simone ALIOLI, Paolo NASON, Carlo OLEARI et Emanuele RE, *NLO vector-boson production matched with shower in POWHEG*, *JHEP* **07** (2008) p. 060, arXiv : [0805.4802 \[hep-ph\]](#) (cf. p. 130).
- [154] Stefan HÖCHE, Frank KRAUSS, Marek SCHÖNHERR et Frank SIEGERT, *NLO matrix elements and truncated showers*, *JHEP* **1108** (2011) p. 123, arXiv : [1009.1127 \[hep-ph\]](#) (cf. p. 131).
- [155] Stefan HÖCHE, Frank KRAUSS, Marek SCHÖNHERR et Frank SIEGERT, *A critical appraisal of NLO+PS matching methods*, *JHEP* **09** (2012) p. 049, arXiv : [1111.1220 \[hep-ph\]](#) (cf. p. 131).
- [156] S. CATANI, F. KRAUSS, B. R. WEBBER et R. KUHN, *QCD Matrix Elements + Parton Showers*, *JHEP* **11** (2001) p. 063, arXiv : [hep-ph/0109231](#) (cf. p. 131).
- [157] Stefan HÖCHE, Frank KRAUSS, Steffen SCHUMANN et Frank SIEGERT, *QCD matrix elements and truncated showers*, *JHEP* **05** (2009) p. 053, arXiv : [0903.1219 \[hep-ph\]](#) (cf. p. 131).
- [158] Stefano CATANI et Massimiliano GRAZZINI, *Next-to-Next-to-Leading-Order Subtraction Formalism in Hadron Collisions and its Application to Higgs-boson Production at the Large Hadron Collider*, *Phys. Rev. Lett.* **98** (2007) p. 222002, arXiv : [hep-ph/0703012 \[hep-ph\]](#) (cf. p. 131).

- [159] S. CATANI, L. CIERI, G. FERRERA, D. de FLORIAN et M. GRAZZINI, *Vector boson production at hadron colliders : A Fully exclusive QCD calculation at NNLO*, *Phys. Rev. Lett.* **103** (2009) p. 082001, arXiv : [0903.2120 \[hep-ph\]](#) (cf. p. 131).
- [160] L. A. HARLAND-LANG, A. D. MARTIN, P. MOTYLINSKI et R. S. THORNE, *Parton distributions in the LHC era : MMHT 2014 PDFs*, *Eur. Phys. J. C* **75** (2015) p. 204, arXiv : [1412.3989 \[hep-ph\]](#) (cf. p. 131).
- [161] CMS COLLABORATION, *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*, *JHEP* **03** (2018) p. 115, arXiv : [1711.03143 \[hep-ex\]](#) (cf. p. 131).
- [162] Georges AAD et al., *Precise measurements of W- and Z-boson transverse momentum spectra with the ATLAS detector using pp collisions at  $\sqrt{s} = 5.02$  TeV and 13 TeV*, (2024), arXiv : [2404.06204 \[hep-ex\]](#) (cf. p. 132).
- [163] Richard D. BALL et al., *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) p. 663, arXiv : [1706.00428 \[hep-ph\]](#) (cf. p. 143).
- [164] Giulio D'AGOSTINI, *A multidimensional unfolding method based on Bayes' theorem*, *Nucl. Instrum. Meth. A* **362** (1995) p. 487, ISSN : 0168-9002 (cf. p. 147, 148).
- [165] G. D'AGOSTINI, *Improved iterative Bayesian unfolding*, 2010, arXiv : [1010.0632 \[physics.data-an\]](#) (cf. p. 148).
- [166] Tim ADYE, "Unfolding algorithms and tests using RooUnfold", *Proceedings, 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding (PHYSTAT 2011)* (CERN, Geneva, Switzerland, 17-20 jan. 2011) p. 313, arXiv : [1105.1160 \[physics.data-an\]](#) (cf. p. 149).
- [167] Georges AAD et al., *Measurement of inclusive jet and dijet production in pp collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector*, *Phys. Rev. D* **86** (2012) p. 014022, arXiv : [1112.6297 \[hep-ex\]](#) (cf. p. 151).
- [168] B. EFRON, *Bootstrap Methods : Another Look at the Jackknife*, *Ann. Statist.* **7** (1979) p. 1, URL : <https://doi.org/10.1214/aos/1176344552> (cf. p. 154).

- [169] Louis LYONS, Duncan GIBAUT et Peter CLIFFORD,  
*How to Combine Correlated Estimates of a Single Physical Quantity*,  
*Nucl. Instrum. Meth. A* **270** (1988) p. 110 (cf. p. 157).
- [170] A. VALASSI, *Combining correlated measurements of several different physical quantities*, *Nucl. Instrum. Meth. A* **500** (2003) p. 391 (cf. p. 157, 158).
- [171] Hannu PAUKKUNEN et Pia ZURITA,  
*PDF reweighting in the Hessian matrix approach*,  
*Journal of High Energy Physics* **2014** (2014), ISSN : 1029-8479,  
URL : [http://dx.doi.org/10.1007/JHEP12\(2014\)100](http://dx.doi.org/10.1007/JHEP12(2014)100)  
(cf. p. 160).
- [172] F. D. AARON et al., *A precision measurement of the inclusive ep scattering cross section at HERA*,  
*The European Physical Journal C* **64** (2009) p. 561, ISSN : 1434-6052,  
URL : <http://dx.doi.org/10.1140/epjc/s10052-009-1169-x>  
(cf. p. 161).
- [173] H1 et ZEUS COLLABORATIONS,  
*Combination and QCD analysis of charm and beauty production cross-section measurements in deep inelastic ep scattering at HERA*,  
2018, arXiv : [1804.01019](https://arxiv.org/abs/1804.01019) [hep-ex] (cf. p. 161).
- [174] S. ALEKHIN et al., *HERAFitter, Open Source QCD Fit Project*, 2015,  
arXiv : [1410.4412](https://arxiv.org/abs/1410.4412) [hep-ph] (cf. p. 161).
- [175] J. ALWALL et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) p. 079,  
arXiv : [1405.0301](https://arxiv.org/abs/1405.0301) [hep-ph] (cf. p. 161).
- [176] Tancredi CARLI et al.,  
*A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders : the APPLGRID project*,  
*The European Physical Journal C* **66** (2010) p. 503, ISSN : 1434-6052,  
URL : <http://dx.doi.org/10.1140/epjc/s10052-010-1255-0>  
(cf. p. 161).
- [177] Valerio BERTONE, Rikkert FREDERIX, Stefano FRIXIONE, Juan ROJO et Mark SUTTON,  
*aMCfast : automation of fast NLO computations for PDF fits*,  
*Journal of High Energy Physics* **2014** (2014), ISSN : 1029-8479,  
URL : [http://dx.doi.org/10.1007/JHEP08\(2014\)166](http://dx.doi.org/10.1007/JHEP08(2014)166)  
(cf. p. 161).

- [178] Stefano CARRAZZA, Stefano FORTE, Zahari KASSABOV,  
Jose Ignacio LATORRE et Juan ROJO,  
*An Unbiased Hessian Representation for Monte Carlo PDFs*,  
**Eur. Phys. J. C** **75** (2015) p. 369, arXiv : [1505.06736](https://arxiv.org/abs/1505.06736) [hep-ph]  
(cf. p. 162).



## A - Merged- $e^+e^-$ identification

This appendix presents a merged- $e^+e^-$  identification that I developed and used in a new resonance search published in Ref. [22], utilizing data from proton–proton collisions at 13 TeV with the ATLAS detector, corresponding to an integrated luminosity of  $139 \text{ fb}^{-1}$ . This analysis aimed to search for new resonances in events with a highly boosted  $Z$  boson, which decays into a pair of electrons or muons. Therefore, the reconstruction efficiency of the  $Z$  boson decaying leptonically is crucial for increasing the statistical power and enhancing the sensitivity to new physics.

When the  $Z$  boson has high transverse momentum, the two electrons originating from its decay are close to each other. In this case, the two electrons may not pass the isolation selection due to energy contamination from the other prompt electron. Thus, instead of reconstructing the  $Z$  boson from two resolved electrons, one can use a single jet to reconstruct the  $Z$  boson, which is called merged- $e^+e^-$  object. In this analysis, the jet candidates were reconstructed using the anti- $k_t$  algorithm with a radius of 0.4 [83] and particle flow algorithm [85]. The BDT technique is employed to identify the merged- $e^+e^-$  object from the jet candidates.

Events should pass the jet triggers or the electron triggers. When the  $Z$  boson can be reconstructed using the isolated electron pair, the events were not considered to avoid overlapping. Preselections were applied on the jet variables to suppress the hadronic jet backgrounds. As the  $Z$  boson is highly boosted, the transverse momentum of jet candidate should be larger than 450 GeV, ensuring the angular distance of the electron pair is smaller than 0.4. The jet should be within the pseudorapidity range  $|\eta| < 2.47$ . The invariant mass of jet ( $m_{\text{jet}}$ ) should satisfy  $66 < m_{\ell\ell} < 116 \text{ GeV}$ . The number of the charged tracks within the jet cone ( $N_{\text{jet}}^{\text{track}}$ ) each with a transverse momentum greater than 500 MeV and ghost associated to the jet candidate, was required to be at least 1 and no more than 7. A  $p_T$  dependent requirement was applied on the fraction of the jet energy deposited in the electromagnetic calorimeter ( $f_{\text{jet}}^{\text{EM}}$ ). The QCD multijet backgrounds become negligible after the  $f_{\text{jet}}^{\text{EM}}$  selections. At least one GSF tracks ( $n_{\text{GSF tracks}}$ ) with transverse momentum greater than 50 GeV, described in Chapter 4, was required to be matched with the jet. To reduce the background from  $\gamma$ +jets, the jet should satisfy  $N_{\text{PFO tracks}} - N_{\text{GSF tracks}} < 2$ ,  $N_{\text{PFO tracks}}$  denotes the number of particle flow tracks with  $p_T > 1 \text{ GeV}$  within the jet. The candidate should contain at least one  $e/\gamma$  clusters with  $p_T > 25 \text{ GeV}$ ,  $n_{e/\gamma \text{ clusters}} \geq 1$ . The cluster-based mass ( $m_{\text{cls}}$ ) was constructed with the information of the leading  $e\gamma$  cluster and the jet width  $W$ , which is defined as  $\sum(\Delta R(\text{jet}, \text{constituent}) p_T^{\text{constituent}}) / \sum p_T^{\text{constituent}}$ . It is calculated as  $m_{\text{cls}} = 0.5W p_T^2 / \sqrt{z(1-z)}$ , where  $z$  is the fraction of jet energy deposited in the leading cluster. This quantity reflects the two body decay property of the merged object. The selected jet candidate should satisfy the  $m_{\text{cls}} > 60 \text{ GeV}$ .

The following variables serve as the input to train the BDT model :

—  $p_T^{\text{jet}}$ ,  $\eta^{\text{jet}}$ ,  $m_{\text{jet}}$ ,  $f_{\text{jet}}^{\text{EM}}$ ,  $N_{\text{jet}}^{\text{tracks}}$ ,  $N_{\text{GSF tracks}}$  and  $n_{e/\gamma \text{ clusters}}$  used in the preselection.

tions,

- The highest fraction of jet energy deposited in any single calorimeter layer,  $\max(E_{\text{layer}}/E_{\text{jet}})$ ,
- The number of constituents within the jet,  $n_{\text{jet}}^{\text{constituent}}$ ,
- The sum of electric charge of GSF tracks with transverse momentum greater than 50 GeV,  $\sum Q_{\text{GSF tracks}}$ ,
- A variable constructed with jet width,  $W p_{\text{T}}/m_{\text{jet}}$ .

The signal jet candidates were selected from the  $Z \rightarrow e^+e^-$  MC simulations by requiring the jet to be within a radius of 0.1 from the true  $Z$  boson. As the pure hardonic jet backgrounds become negligible after preselections, the backgrounds are mainly contributed by the  $\gamma$ +jets,  $W$ +jets and  $W+\gamma$ . After the preselections, the event yields of signal and backgrounds are 102 593 and 8 572, respectively. These events were used to train the BDT model and the merged- $e^+e^-$  objects were identified by applying selections based on the BDT score. The threshold of the BDT score for selections depends on the transverse momentum of the jet.

Figure A.1 displays the  $Z$ -boson reconstruction efficiency using the merged- $e^+e^-$  identification for the  $Z \rightarrow e^+e^-$  process. Significant improvement is observed for  $Z$  bosons with high transverse momentum when using this identification method, compared to the efficiency achieved with the standard electron reconstruction procedure. The combined reconstruction efficiency reaches 80% for high- $p_{\text{T}}$   $Z$  bosons. The background rejection of the BDT was evaluated and compared with the cut-based method using  $f_{\text{jet}}^{\text{EM}}$  as a function of the signal efficiency. The BDT technique enhances purity when having the same signal efficiency.

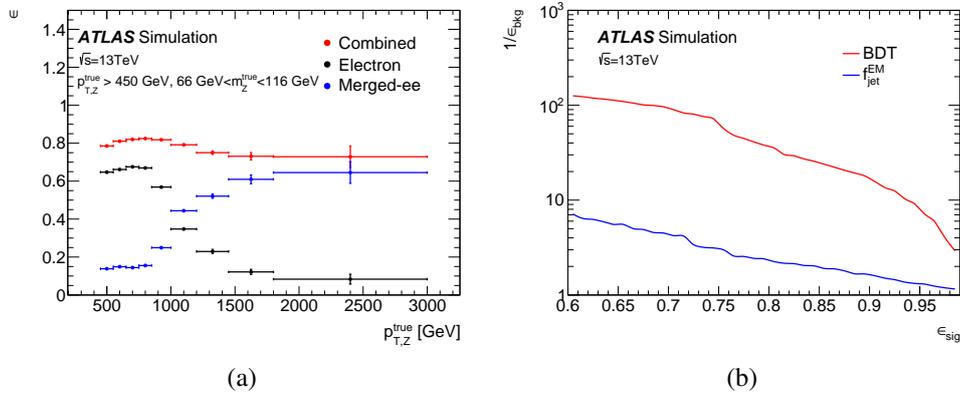


FIGURE A.1 – (a) The  $Z$ -boson reconstruction efficiency as a function of the transverse momentum of  $Z$ -boson at the particle level. The black points indicate the reconstruction efficiencies using the standard electron reconstruction procedure. The blue points denote the efficiencies of merged- $e^+e^-$  identification. The red points are the combined efficiencies of these two methods. (b) The background rejection factor as a function of the signal efficiency. The red curve indicates the BDT results, while the blue curve is for the cut-based method with  $f_{\text{jet}}^{\text{EM}}$ .



## B - Identification efficiency measurement using $Z \rightarrow ee$ tag-and-probe method : $Z_{\text{iso}}$ method

This appendix describes my contribution to the measurement of electron identification efficiencies in LHC Run-2 with the ATLAS detector [76].

As described in Chapter 4, the likelihood classifier is employed to identify electrons. Three working points (Tight, Medium, and Loose) are defined based on their identification efficiency. To address the discrepancies between data and simulated samples in the identification efficiencies, corrections are applied. These efficiency corrections are measured as functions of electron  $E_T$  and  $\eta$ . The tag and probe method is employed to estimate the corrections. This method uses a well-known process. In this analysis, the  $Z \rightarrow ee$  process is used to measure the correction. One electron from the  $Z$  boson decay is identified using the Tight identification criteria, referred to as the “tag” electron, while the other electron, known as the “probe” electron, is used to determine the identification efficiencies. The identification efficiency is determined by taking the ratio of the probe electrons passing a given working point (WP) to all probe electrons.

The  $Z_{\text{iso}}$  method is a background estimation technique used when performing the tag-and-probe approach. In this method, the calorimeter isolation variable  $E_T^{\text{coneXX}}$ , where XX denote the cone size around the electron, is used as the discriminant variable to differentiate between signal and background. Genuine electrons from  $Z$  boson decays are isolated, whereas most backgrounds, such as misidentified jets, are not.  $E_T^{\text{cone30}}$  is used as the default in this analysis, measuring the transverse energy deposits within a radius of 0.3.

The events should pass the electron triggers and contain two electrons. Both the tag and probe electrons are in the central region,  $|\eta| < 2.47$ . The tag electron is selected with tight criteria :  $p_T > 25$  GeV, excluding the region  $1.37 < |\eta| < 1.52$ . It must also satisfy the tight identification and track isolation requirements. The criteria for the probe electron are looser :  $p_T > 25$  GeV and  $|\eta| < 2.47$ . The tag-and-probe electron pair should have opposite electric charges, and the invariant mass of the electron pair should be within the mass window  $75 < m_{\ell\ell} < 105$  GeV.

The  $E_T^{\text{cone30}}$  variable is separated into the signal region ( $E_T^{\text{cone30}} < 12.5$  GeV) and the control region ( $E_T^{\text{cone30}} > 12.5$  GeV). The signal region (SR) is dominated by genuine electrons, while the control region (CR) is enriched with fake electron backgrounds. To estimate the backgrounds in the SR, data events containing tag and probe electron pairs with the same electric charge are selected and the background templates are defined using the probe electrons that fail some of the identification criteria, which are less dependent of the isolation variable to reduce bias. The signal contributions to the control region are determined using the  $Z \rightarrow ee$  simulations and removed from the CR. The background templates are normalized to the events in the

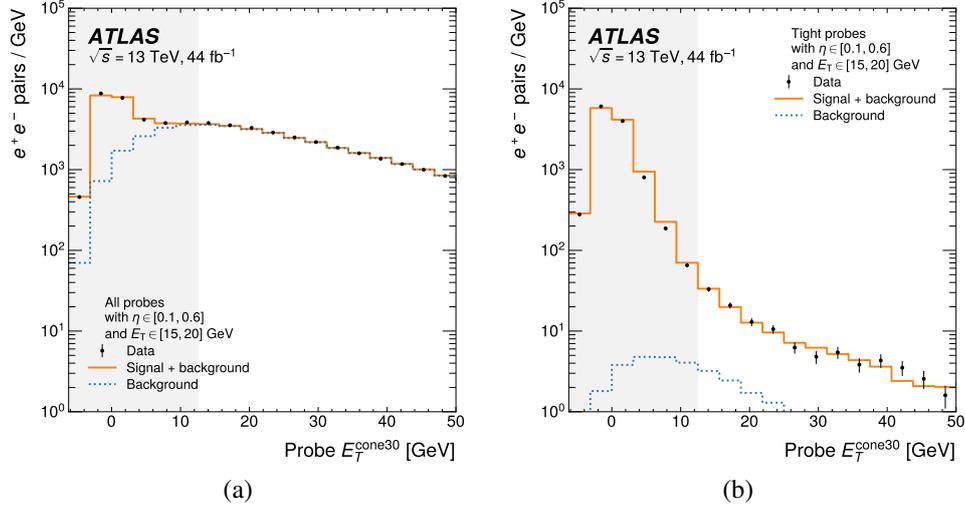


FIGURE B.1 – The isolation variable  $E_T^{\text{cone30}}$  distributions for the probe electrons within  $0.1 < |\eta| < 0.6$  and  $15 < E_T < 20$  GeV are shown. The shaded areas in the plots indicate the signal region. Panel (a) displays the distributions of all probe electrons, while panel (b) shows the distributions of probe electrons satisfying the Tight WP criteria.

CR, after removing the signal leakage, and then extrapolated to the SR. Figure B.1 displays an example of the  $E_T^{\text{cone30}}$  distributions for the probe electrons.

To ensure the background templates provide an accurate description of the backgrounds, the bias is evaluated using fake electrons in the QCD simulations. The bias is defined as follows :

$$\text{bias} = \left| \frac{N^{\text{template}}(E_T^{\text{cone30}} < 7.5 \text{ GeV})}{N^{\text{template}}(E_T^{\text{cone30}} > 7.5 \text{ GeV})} \frac{N^{\text{all}}(E_T^{\text{cone30}} > 7.5 \text{ GeV})}{N^{\text{all}}(E_T^{\text{cone30}} < 7.5 \text{ GeV})} - 1 \right|, \quad (\text{B.1})$$

where the superscripts “template” and “all” indicate the fake electrons passing the background template selection criteria and all fake electrons in the sample, respectively. The bias variable measures the goodness of the background template in replicating the true ratio of the background between the CR and the SR. Several background templates are constructed, and those with minimal bias are selected in each  $E_T$ - $\eta$  bin. The typical signal contamination in the background templates is less than 5%, while it is around 25% for  $|\eta| > 1.81$  and  $E_T > 25$  GeV.

The systematic uncertainties are determined by calculating the differences between the systematic variations and the nominal results. The systematic variations are listed as follows :

**Background templates :** using the alternative background templates.

**Isolation variable :** using  $E_T^{\text{cone40}}$  instead of  $E_T^{\text{cone30}}$ .

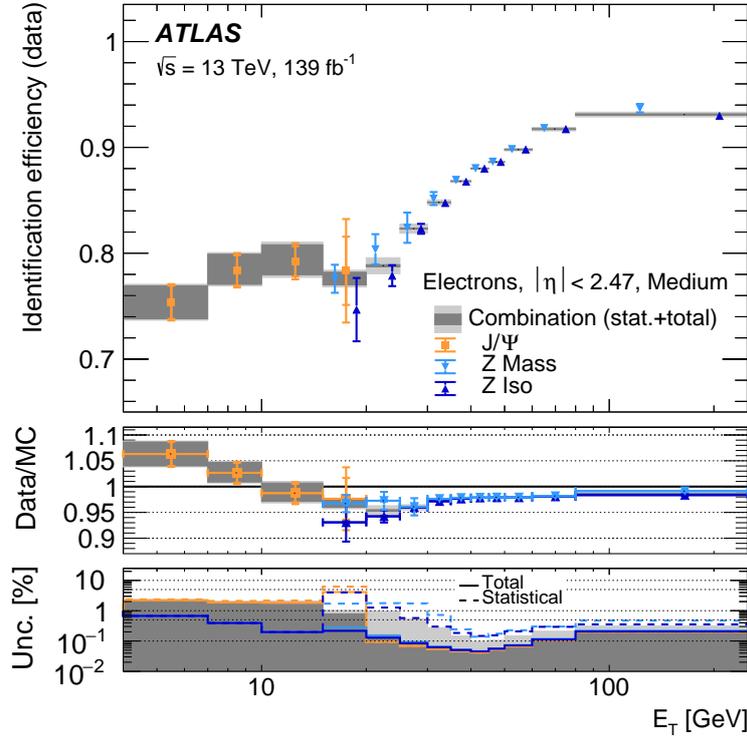


FIGURE B.2 – Comparison of the identification efficiencies for the Medium working point obtained using different approaches.

**Tag electron :** applying an additional isolation requirement of  $E_T^{\text{cone30}} < 7.5$  GeV to the tag electrons.

**The normalization of the signal leakage :** varying the normalisation of the signal contributions in the CR by 30% to address the signal modelling uncertainties.

**The invariant mass of the lepton pair :** using alternative selections of the tag-and-probe electron invariant mass  $80 < m_{\ell\ell} < 100$  GeV and  $70 < m_{\ell\ell} < 110$  GeV.

**The boundary between the SR and CR :** using different definitions of the boundaries between the SR and CR,  $E_T^{\text{cone30}} = 10$  GeV and  $E_T^{\text{cone30}} = 15$  GeV.

The measured results using the  $Z_{\text{iso}}$  method is then combined with ones obtained using  $Z_{\text{mass}}$  method and  $J/\psi$  events [79]. Figure B.2 shows the comparison between the measured results with different approaches. Good consistency are observed. The comparison between the new identification results and the previous results using the 2015-2017 dataset [75] is displayed in Figure B.3. They are consistent, and the uncertainties are reduced.

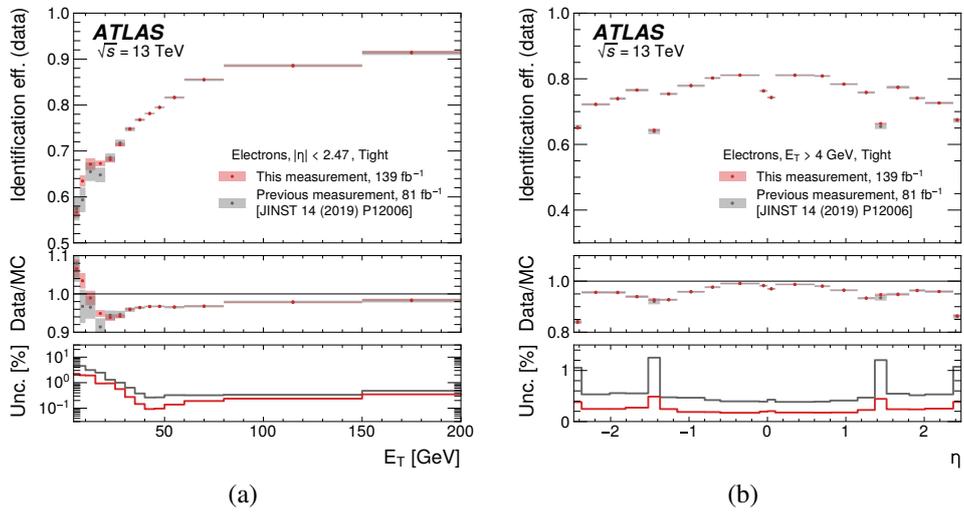


FIGURE B.3—Comparison of the new combined TIGHT identification efficiencies with the previous ATLAS measurements based on data from 2015–2017 [75].

## C - Auxiliary material for measurement of $W^\pm$ and $Z$ production cross-sections and their ratios at $\sqrt{s} = 13.6$ TeV

### C.1 The cross section measurement with an event-counting approach

In addition to the profile likelihood method, the event-counting approach is employed as an alternative method to measure the inclusive cross section, serving as a complementary check. The formula is as Eq. (C.1) shows :

$$\sigma^{\text{tot}} = \frac{N_{\text{data}} - N_{\text{bg}}}{\mathcal{L}CA} \quad (\text{C.1})$$

where  $N_{\text{data}}$  and  $N_{\text{bg}}$  represent the numbers of selected data and background events, respectively. The quantity  $C$  is the correction factor, which is obtained by taking the ratio between the expected signal events at the detector level and particle level, i.e.,  $C = N_{\text{sig}}^{\text{reco}} / N_{\text{sig}}^{\text{true}}$ . The acceptance  $A$  is the ratio between the number of events in the fiducial volume and the total phase space at the particle level. For the  $Z$  channel, the total phase space is defined as the  $Z$ -boson within the range of the invariant mass of the lepton pair,  $66 < m_{\ell\ell} < 116$  GeV.

The results are summarised in Tables C.1.

TABLE C.1 – The inclusive cross sections measured by counting the event yields. The SHERPA  $V$ +jets samples are used to compute the correction and acceptance factors. The first quoted uncertainty is due to statistical fluctuations, the second one is systematic uncertainty excluding the luminosity uncertainty, and the third one is the luminosity uncertainty.

	$\sigma \pm \delta\sigma_{\text{stat}} \pm \delta\sigma_{\text{syst}} \pm \delta\sigma_{\text{lumi}}$		
	$W^- \rightarrow e^- \bar{\nu}$	$W^+ \rightarrow e^+ \nu$	$Z \rightarrow e^+ e^-$
$C$	0.4445	0.4385	0.3613
$\sigma^{\text{fid}}$ [pb]	$3379.4 \pm 0.6 \pm 140.7 \pm 86.5$	$4346.0 \pm 0.6 \pm 156.2 \pm 108.2$	$740.5 \pm 0.3 \pm 14.6 \pm 16.4$
Acceptance $A$	$0.381 \pm 0.009$	$0.366 \pm 0.009$	$0.374 \pm 0.011$
$\sigma^{\text{tot}}$ [pb]	$8875.0 \pm 1.4 \pm 423.4 \pm 227.2$	$11874.7 \pm 1.7 \pm 523.9 \pm 295.7$	$1981.2 \pm 0.7 \pm 69.1 \pm 43.8$
	$W^- \rightarrow \mu^- \bar{\nu}$	$W^+ \rightarrow \mu^+ \nu$	$Z \rightarrow \mu^+ \mu^-$
$C$	0.5872	0.5933	0.6532
$\sigma^{\text{fid}}$ [pb]	$3306.3 \pm 0.5 \pm 92.6 \pm 84.0$	$4240.5 \pm 0.5 \pm 112.6 \pm 105.2$	$747.4 \pm 0.2 \pm 15.4 \pm 16.5$
Acceptance $A$	$0.381 \pm 0.009$	$0.365 \pm 0.010$	$0.374 \pm 0.011$
$\sigma^{\text{tot}}$ [pb]	$8684.7 \pm 1.2 \pm 316.1 \pm 220.5$	$11614.2 \pm 1.4 \pm 434.1 \pm 288.1$	$1998.0 \pm 0.5 \pm 69.8 \pm 44.2$



## **D - Auxiliary material for measurement of $W^\pm$ boson differential cross sections with low pile-up data at $\sqrt{s} = 5.02$ and 13 TeV**

### **D.1 Differential cross sections in the electron and muon channels**

#### **D.1.1 Differential cross sections in $|\eta_\ell|$**

The measured  $|\eta_\ell|$ -dependent differential cross section is obtained by performing the unfolding procedure in the  $|\eta_\ell|$  distribution with the binning described in Section 6.8.2. Figures D.1-D.2 display the results in the electron and muon channels from  $W^+$  and  $W^-$  boson decays at 5.02 and 13 TeV, respectively. The results are compared to the theory predictions calculated with different PDF sets as described in Section 6.7. The size of uncertainties from different sources is illustrated in Figures D.3-D.4.

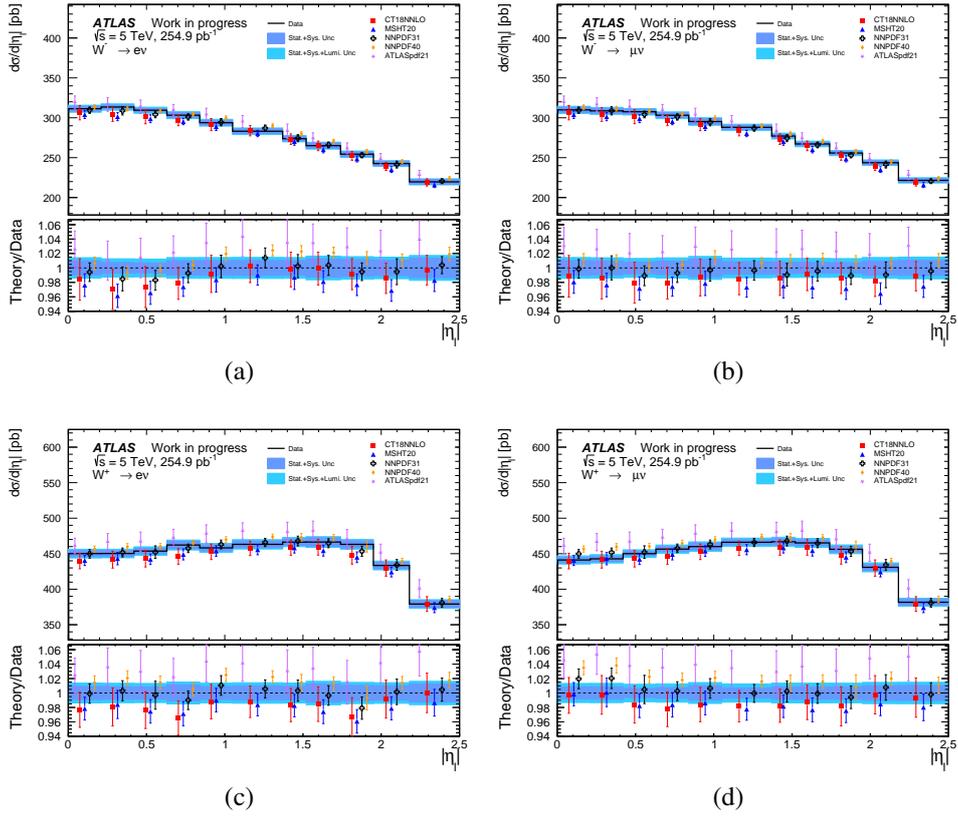


FIGURE D.1 – Differential cross section as a function of  $|\eta^\ell|$  for the electron and muon channels at 5.02 TeV. The data are presented with systematic and total uncertainties and compared with the theory predictions with different PDF sets. The ratios of the data to the corresponding predictions are shown in the lower panels.

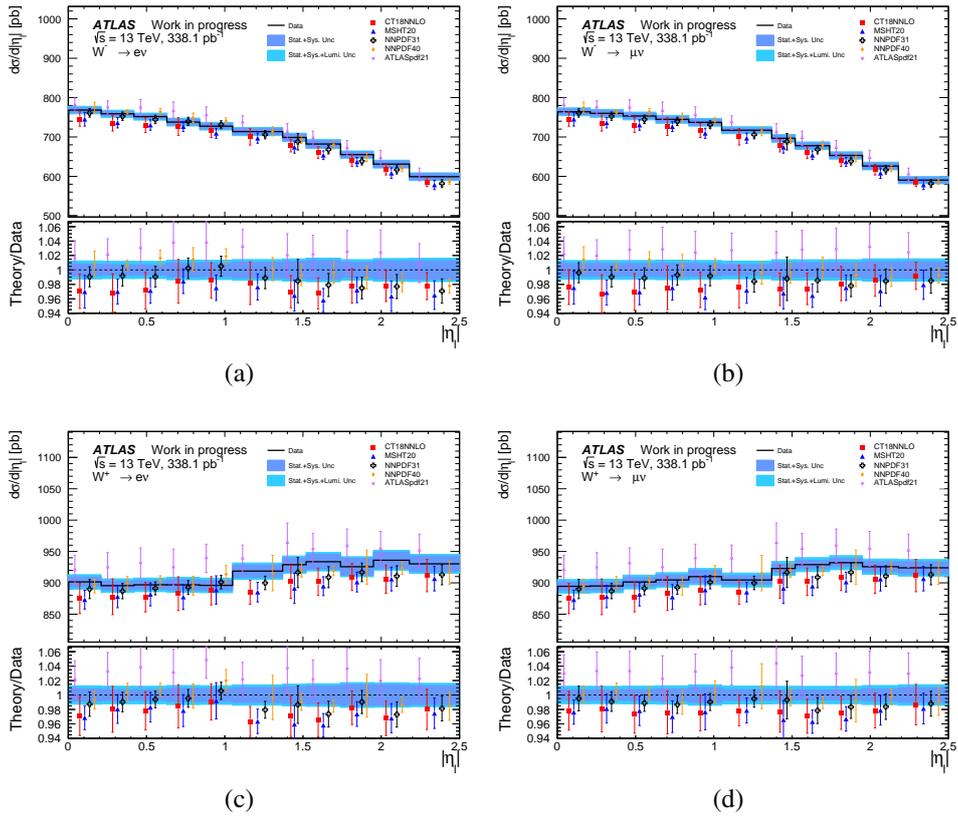


FIGURE D.2 – Differential cross section as a function of  $|\eta^\ell|$  for the electron and muon channels at 13 TeV. The data are presented with systematic and total uncertainties and compared with the theory predictions with different PDF sets. The ratios of the data to the corresponding predictions are shown in the lower panels.

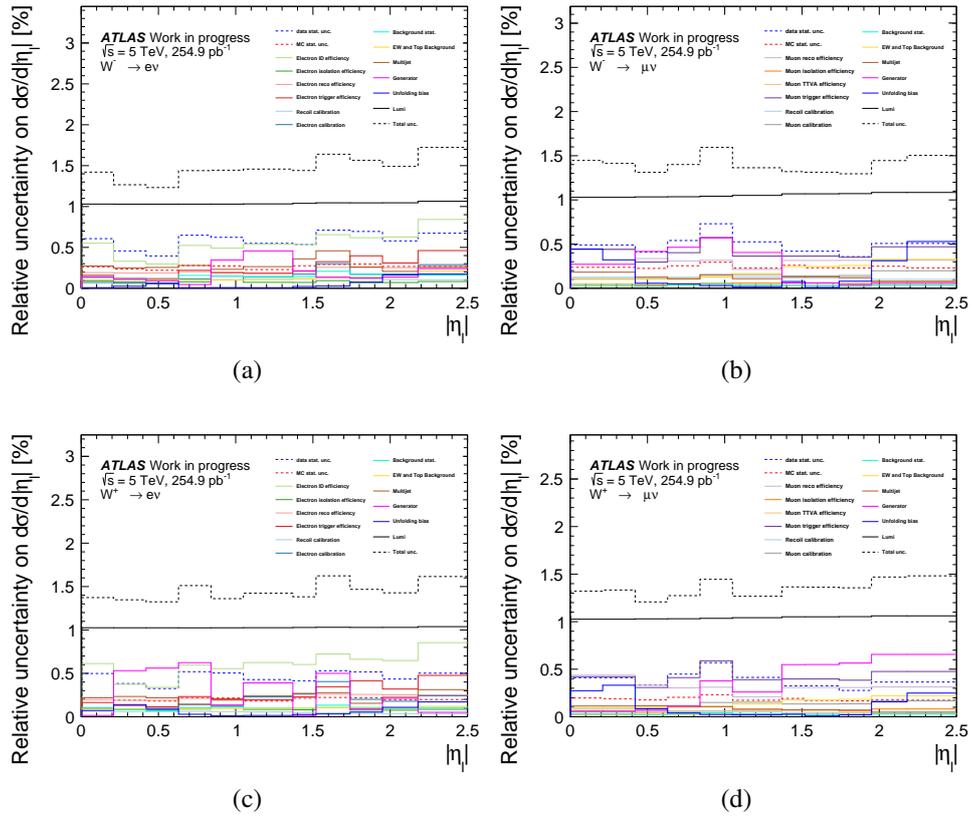


FIGURE D.3 – The relative systematic uncertainty from each source for differential cross section as a function of  $|\eta^\ell|$  for the electron and muon channels at 5.02 TeV.

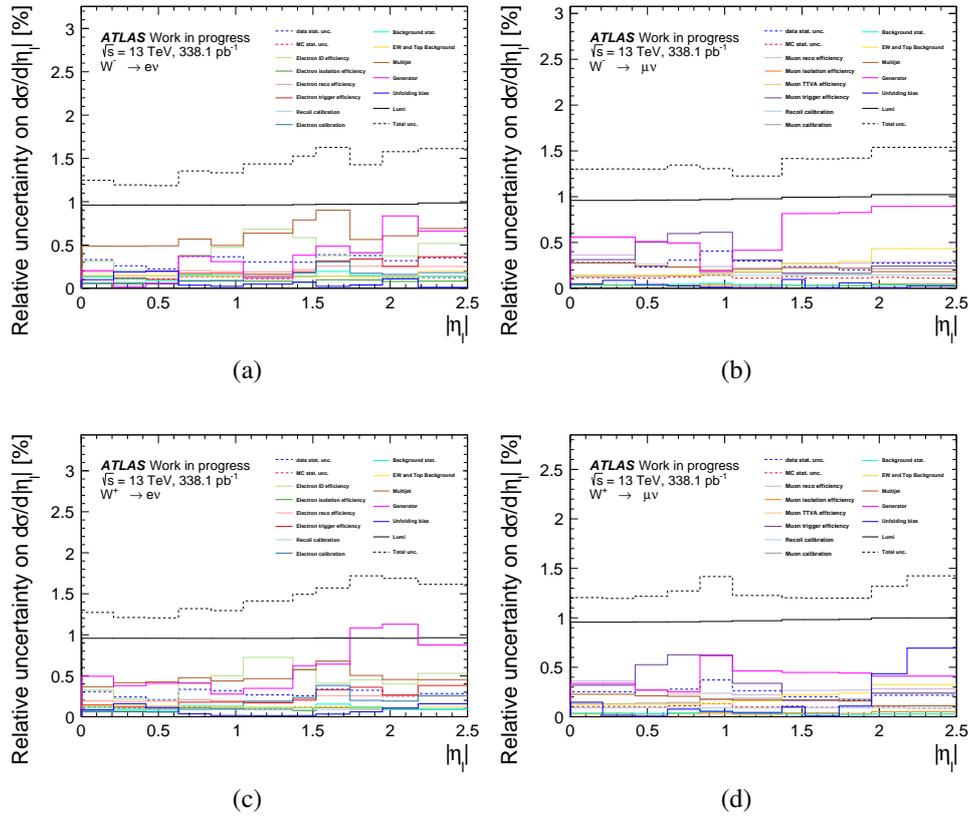


FIGURE D.4 – The relative systematic uncertainty from each source for differential cross section as a function of  $|\eta^\ell|$  for the electron and muon channels at 13 TeV.

### D.1.2 Differential cross sections in $p_T^\ell$

The measured  $p_T^\ell$ -dependent differential cross section is obtained by performing the unfolding procedure in the  $p_T^\ell$  distribution with the binning described in Section 6.8.2. Figures D.5-D.6 display the results in the electron and muon channels from  $W^+$  and  $W^-$  boson decay at 5.02 and 13 TeV, respectively. The results are compared to the theory predictions calculated with different PDF sets as described in Section 6.7. The size of uncertainties from different sources is illustrated in Figures D.7-D.8.

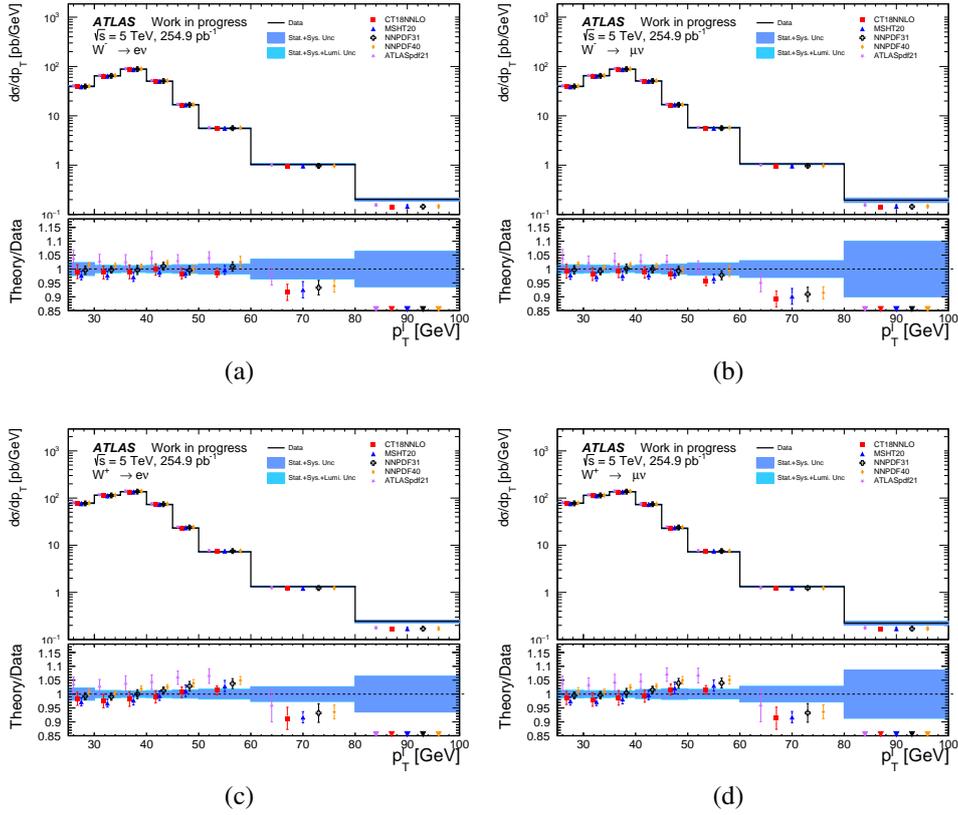


FIGURE D.5 – Differential cross section as a function of  $p_T^\ell$  for the electron and muon channels at 5.02 TeV. The data are presented with systematic and total uncertainties and compared with the theory predictions with different PDF sets. The ratios of the data to the corresponding predictions are shown in the lower panels.

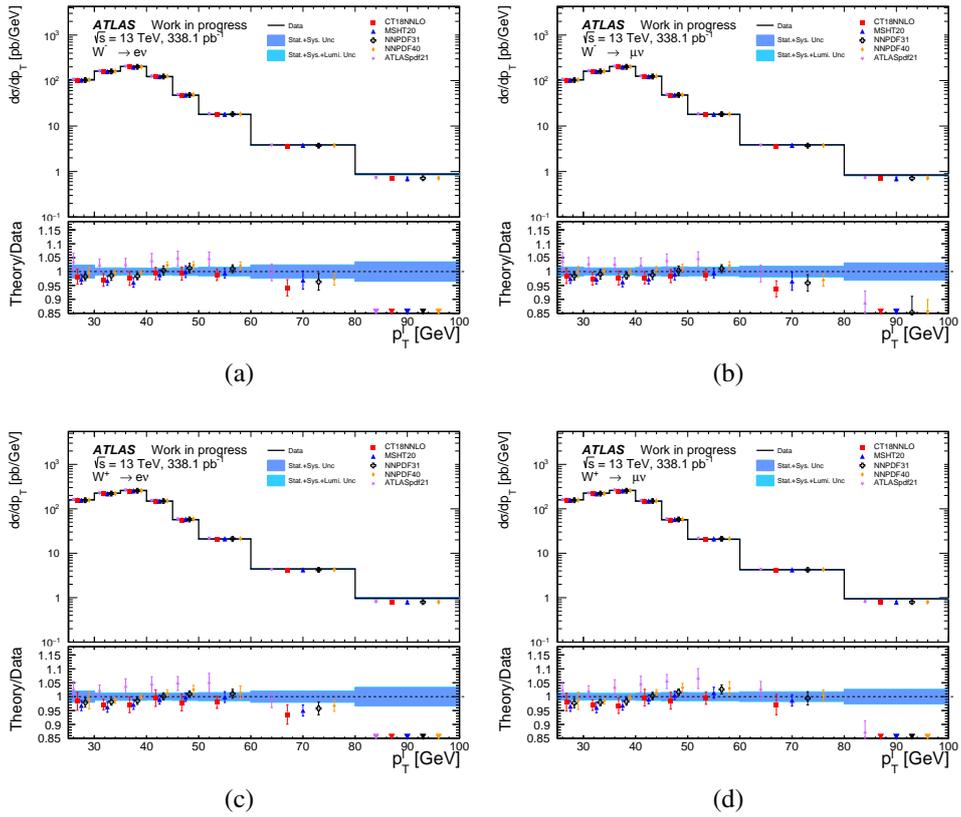


FIGURE D.6 – Differential cross section as a function of  $p_T^\ell$  for the electron and muon channels at 13 TeV. The data are presented with systematic and total uncertainties and compared with the theory predictions with different PDF sets. The ratios of the data to the corresponding predictions are shown in the lower panels.

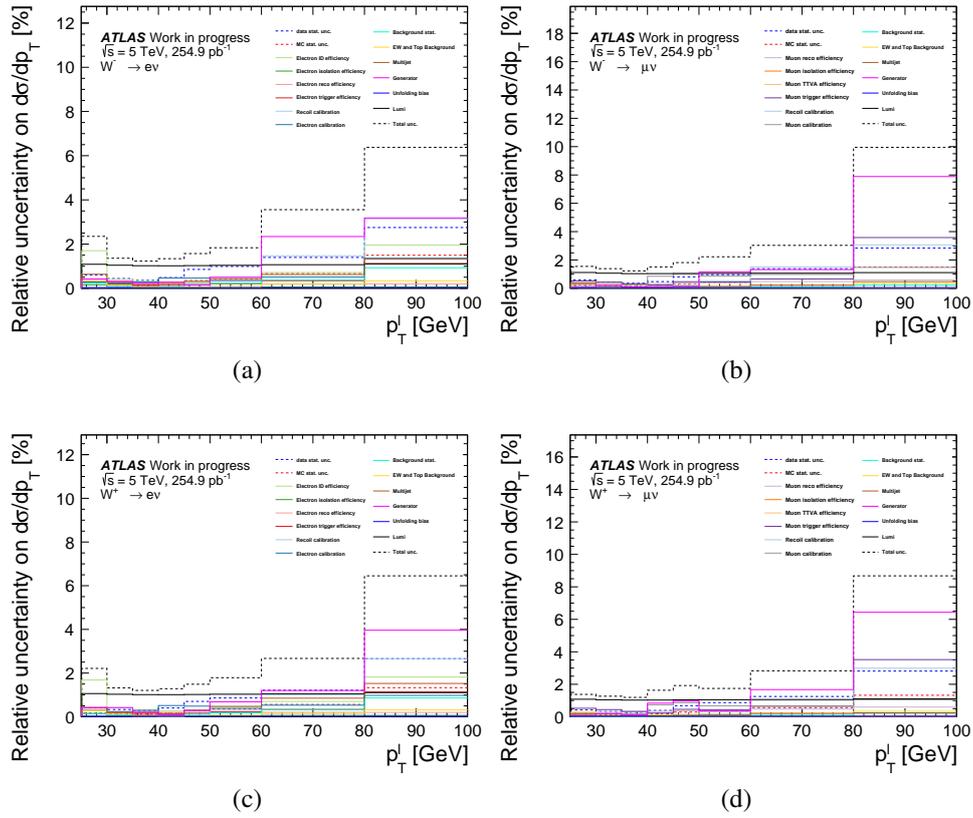


FIGURE D.7 – The relative systematic uncertainty from each source for differential cross section as a function of  $p_T^\ell$  for the electron and muon channels at 5.02 TeV.

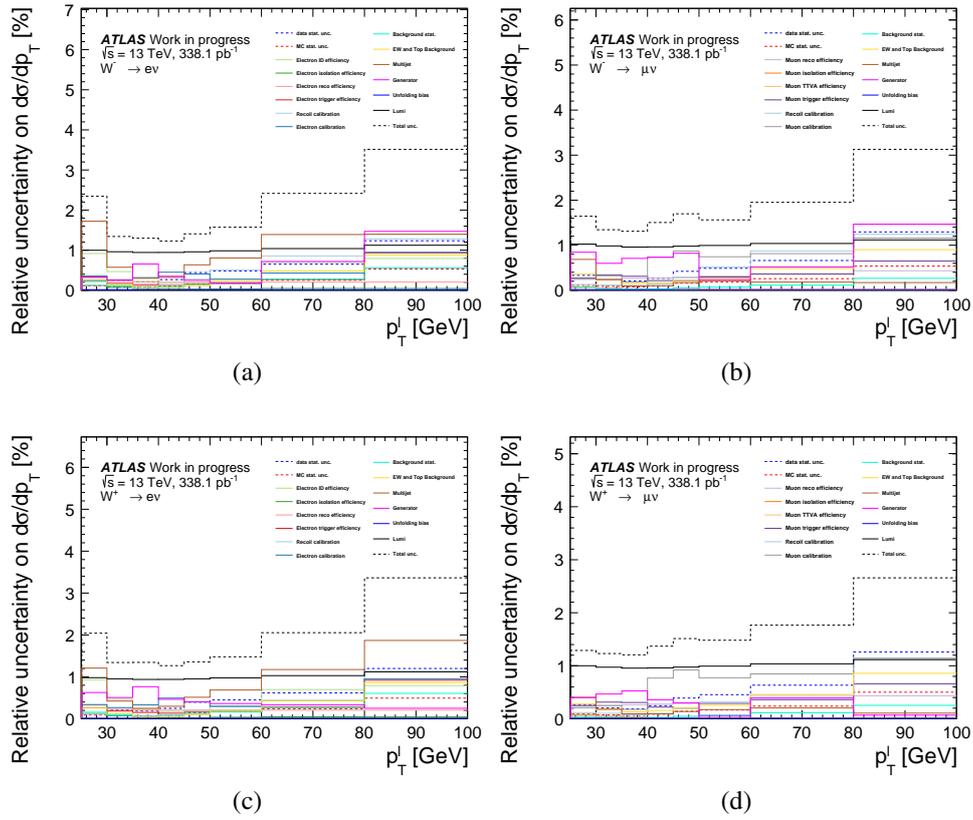
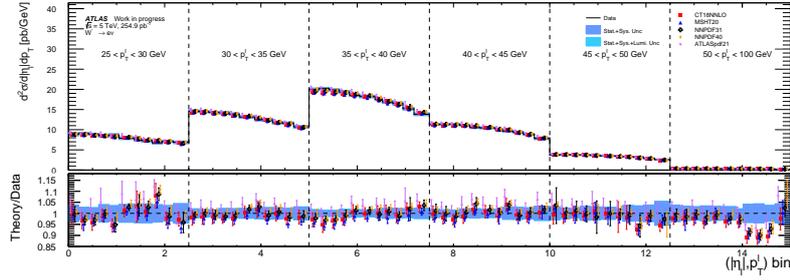


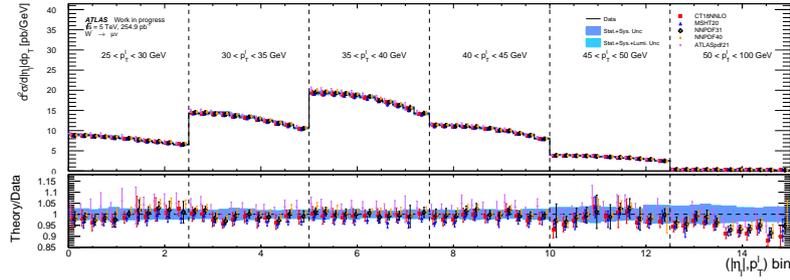
FIGURE D.8 – The relative systematic uncertainty from each source for differential cross section as a function of  $p_T^\ell$  for the electron and muon channels at 13 TeV.

### **D.1.3 Differential cross sections in $p_T^\ell$ and $|\eta_\ell|$**

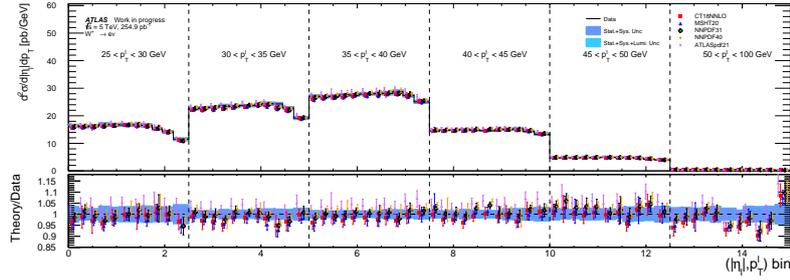
The double differential cross section in  $p_T^\ell$  and  $|\eta_\ell|$  is obtained by performing the unfolding procedure in the  $(|\eta_\ell|, p_T^\ell)$  distribution with the binning described in Section 6.8.2. The results are compared to the theory predictions calculated with different PDF sets as described in Section 6.7. The size of uncertainties from different sources is illustrated in Figures D.11-D.12.



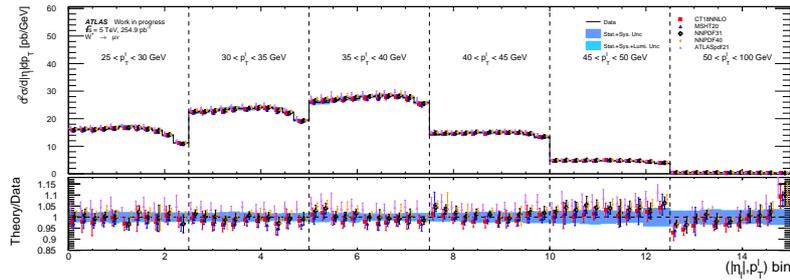
(a)



(b)

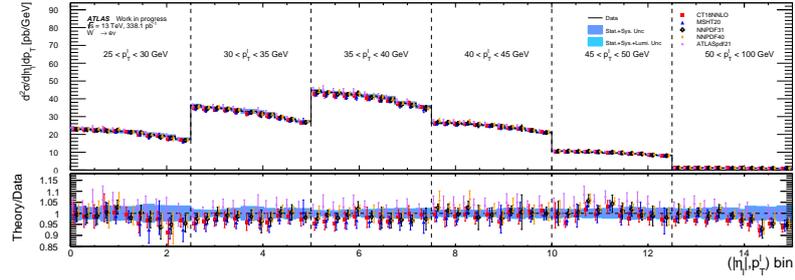


(c)

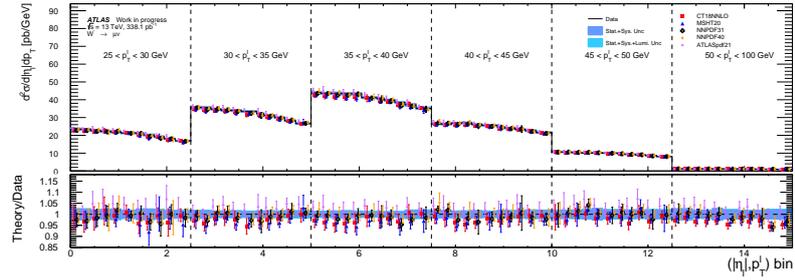


(d)

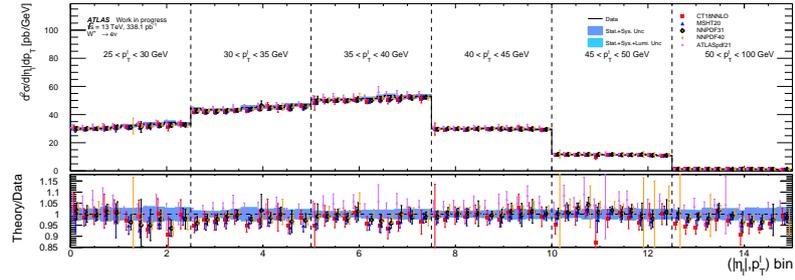
FIGURE D.9 – Differential cross section as functions of  $|\eta_\ell|$  and  $p_T^\ell$  for the electron and muon channels at 5.02 TeV. The data are presented with systematic and total uncertainties and compared with the theory predictions with different PDF sets. The ratios of the data to the corresponding predictions are shown in the lower panels.



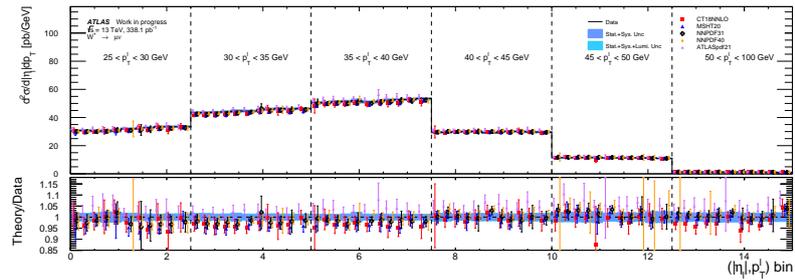
(a)



(b)



(c)



(d)

FIGURE D.10 – Differential cross section as functions of  $|\eta_e|$  and  $p_T^\ell$  for the electron and muon channels at 13 TeV. The data are presented with systematic and total uncertainties and compared with the theory predictions with different PDF sets. The ratios of the data to the corresponding predictions are shown in the lower panels.

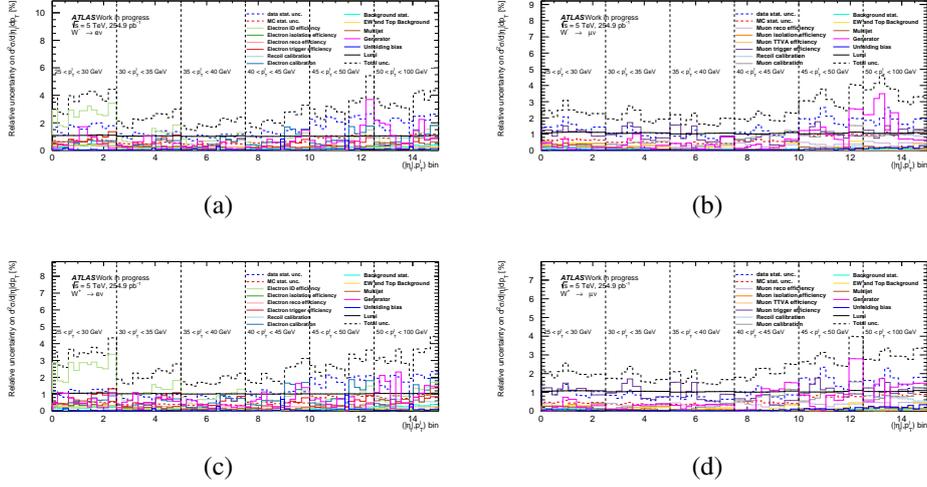


FIGURE D.11 – The relative systematic uncertainty from each source for differential cross section as functions of  $p_T^e$  and  $|\eta_e|$  for the electron and muon channels at 5.02 TeV.

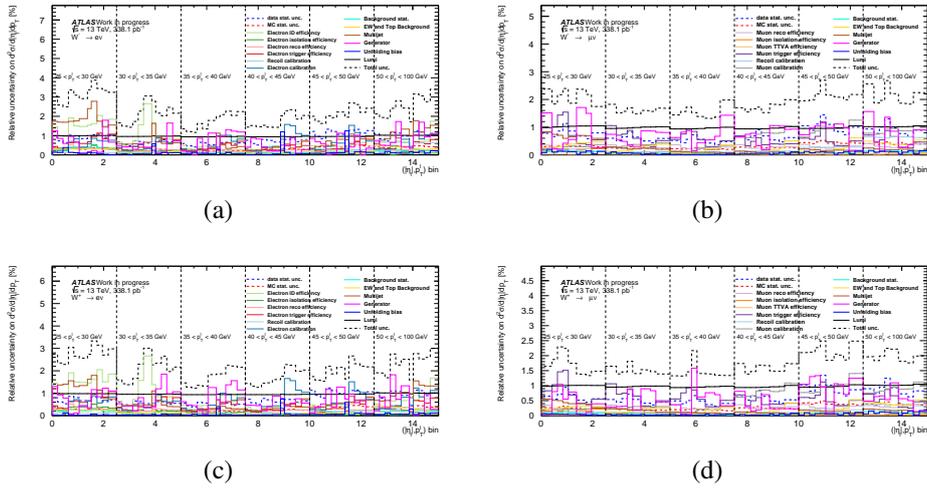


FIGURE D.12 – The relative systematic uncertainty from each source for differential cross section as functions of  $p_T^e$  and  $|\eta_e|$  for the electron and muon channels at 13 TeV.



## Publications

### Published

1. ATLAS Collaboration, *A search for new resonances in multiple final states with a high transverse momentum Z boson in  $\sqrt{s} = 13$  TeV pp collisions with the ATLAS detector*, **JHEP 06 (2023) 036**
2. ATLAS Collaboration, *Precise measurements of W and Z transverse momentum spectra with the ATLAS detector at  $\sqrt{s} = 5.02$  TeV and 13 TeV*, ATLAS-CONF-2023-028, 2023, <https://cds.cern.ch/record/2861057>
3. ATLAS Collaboration, *Top pair production cross section and its ratio with Z production cross section at 13.6 TeV*, **Phys. Lett. B 848 (2024) 138376**
4. ATLAS Collaboration, *Measurement of vector boson production cross sections and their ratios using pp collisions at  $\sqrt{s} = 13.6$  TeV with the ATLAS detector*, **Phys. Lett. B 854 (2024) 138725**
5. ATLAS Collaboration, *Electron and photon efficiencies in LHC Run 2 with the ATLAS experiment*, **JHEP 05 (2024) 162**



## Acknowledgements

First and foremost, I would like to express my deepest gratitude to my supervisors, Prof. Haiping Peng and Prof. Zhiqing Zhang, for their unwavering support, insightful guidance, and constant encouragement throughout the course of my research. Their expertise and constructive feedback have been invaluable in shaping this thesis, and their patience has been greatly appreciated.

A special thanks goes to Prof. Louis Fayard. He offered me a great deal of help while I was in France, assisting with many administrative issues.

I am also deeply grateful to my jury members for carefully reviewing my thesis and providing many thoughtful comments and suggestions.

I would like to extend my heartfelt thanks to my colleagues in the UTSC and IJCLab ATLAS groups, as well as my collaborators within the ATLAS collaboration, particularly those from the Standard Model group. The discussions and collaborations with all of you have been both intellectually stimulating and personally rewarding.

On a personal note, I would like to thank my family for their unconditional love and support.

Thank you to everyone who has contributed to this journey in one way or another.