Étalonnage du calorimètre électromagnétique ATLAS et mesure des propriétés du Boson W à $sqrts$ = 5 et 13 TeV avec le détecteur ATLAS au LHC.

Hicham Atmani

Calibration of the ATLAS Electromagnetic Calorimeter and Measurement of W Boson Properties at $\sqrt{s}$ = 5 and 13 TeV with the ATLAS Detector at the LHC

Thèse de doctorat de l'Université Paris-Saclay

École doctorale n° 576, Particules, Hadrons, Énergie, Noyau, Instrumentation, Imagerie, Cosmos, Simulation (PHENIICS)
Spécialité de doctorat: La physique des particules.
Unité de recherche: Université Paris-Saclay, CNRS, IJCLab, 91405, Orsay, France.
Référent: Faculté des sciences d'Orsay.

Thèse présentée et soutenue à Orsay, le 03/12/2020, par

HICHAM ATMANI

Composition du jury:

Maire-Hélène Schune Directrice de recherche, Université Paris-Saclay (IJCLAB)	Présidente de jury
Alexandre Glazov Chercheur, Desy Laboratory	Rapporteur
Aram Apyan Research Associate, Fermi Laboratory	Rapporteur
Anne-Catherine Le Bihan Chargée de recherche, l'Université de Strasbourg (IPHC)	Examinatrice
Maarten Boonekamp Directeur de Recherche, CEA Saclay	Examinateur
Zhiqing ZHANG Directeur de recherche, Université Paris-Saclay (IJCLAB)	Directeur de thèse
Louis FAYARD Directeur de recherche, Université Paris-Saclay (IJCLAB)	Co-directeur de thèse

Thèse de doctorat

NNT: 2020UPASP070

Titre: Étalonnage du calorimètre électromagnétique ATLAS et mesure des propriétés du Boson W à $\sqrt{s} = 5$ et 13 TeV avec le détecteur ATLAS au LHC.

Mots clés: Masse, Mesure de précision, Boson W, ATLAS, LHC

Résumé: Cette thèse se compose de deux parties principales. La première partie correspond à un étalonnage in-situ du calorimètre électromagnétique à l'aide d'électrons et de positrons provenant de désintégrations du boson Z sélectionnées parmi toutes les données nominales collectées par ATLAS lors du Run 2 entre 2015 et 2018, ainsi que pendant des runs spéciaux, caractérisés par un faible nombre d'interactions par croisement afin d'améliorer la résolution expérimentale de la mesure du recul des bosons Z et W et de la reconstruction des différentes variables cinématiques nécessaires à la détermination de la masse du boson W. Les données nominales ont été collectées à une énergie dans le centre de masse de 13 TeV correspondant à une luminosité intégrée d'environ 140 fb^{-1}, alors que les runs spéciaux ont été collectés à 5 et 13 TeV, correspondant à 258 pb^{-1} et 340 pb^{-1}, respectivement. La deuxième partie porte sur la mesure des propriétés du boson W à l'aide des données des runs spéciaux, y compris une mesure de l'impulsion transverse du boson W, des mesures des sections efficaces fiducielles, simple et double différentielles, et une évaluation préliminaire des incertitudes statistiques et systématiques expérimentales sur la détermination de la masse du boson W.

Title: Calibration of the ATLAS Electromagnetic Calorimeter and Measurement of W Boson Properties at $\sqrt{s} = 5$ and 13 TeV with the ATLAS Detector at the LHC.

Keywords: Mass, Precision measurement, W boson, ATLAS, LHC

Abstract: This thesis consists of two main parts. The first part corresponds to an insitu calibration of the electromagnetic calorimeter using electrons and positrons from Z boson decays selected from all nominal data samples taken by ATLAS during Run 2 between 2015 and 2018, as well as during special runs, characterised by a low number of interactions per bunch crossing in order to improve the experimental resolution of the recoil measurement of Z and W bosons and of the reconstruction of various kinematic variables needed for the W boson mass determination. The nominal data samples were taken at a center of mass energy of 13 TeV corresponding to an integrated luminosity of about 140 fb^{-1}, whereas the special runs were taken at 5 and 13 TeV, corresponding to 258 pb^{-1} and 340 pb^{-1}, respectively. The second part is on the measurement of W boson properties using data from the special runs, including a measurement of the transverse momentum of the W boson, a measurement of fiducial, single and double differential cross sections, and a preliminary evaluation of the statistical and experimental uncertainties of the W mass determination.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France

Introduction

After the discovery of W and Z bosons at the Super Proton Synchrotron (SPS) at CERN, particles responsible of weak interactions, the efforts have been geared towards measuring their properties with high precision to test the consistency of the Standard Model. The Standard Model has 25 free parameters to describe particles and their interactions, some of them are measured directly, sometimes with great precision, the other parameters are constrained by the measurement of physical quantities related by the theory. For example, the W boson mass M_W, by including radiative corrections, is related to the masses of the Z boson, Higgs boson and quark top. The Z boson mass is measured with high precision at LEP experiments, and after the discovery of the Higgs boson in 2012 at the LHC, the W boson mass can be predicted theoretically and the comparison between the predicted and measured values is considered as a solid test for the consistency of the Standard Model.

This thesis describes the measurement of the W boson properties using data collected with the ATLAS detector at the LHC at a center-of-mass energy of 5 and 13 TeV, corresponding to an integrated luminosity of 258 pb^{-1} at $\sqrt{s}$ = 5 TeV and 340 pb^{-1} at $\sqrt{s}$ = 13 TeV. The data used are collected using special runs, with low number of interactions per bunch crossing, in order to improve the experimental resolution of the recoil measurement and the reconstruction of the missing transverse momentum and of the transverse mass. In this work, we are interested in the leptonic decays of the W boson, and the charged products of this decay, electron or muon, are accompanied by a neutrino. The neutrino can not be measured directly in the detector but can be measured indirectly using the lepton and the hadronic system which recoils against the W boson.

Chapter 1 describes the Standard Model with a brief review of the spontaneous symmetry breaking mechanism. Also the W boson production in pp collisions at the LHC is described. Finally, a brief history of W boson mass measurements is given, focusing on last results published by the ATLAS collaboration, with a description of all the dominant sources of uncertainties.

Chapter 2 briefly describes the LHC machine, gives a review of the LHC acceleration chain and describes the machine performance. Then, the ATLAS detector in described with its different parts focusing mainly on the electromagnetic calorimeter, an important element of this thesis.

Chapter 3 gives a detailed explanation of the calibration procedure of electromagnetic particles, electrons and photons, in the ATLAS detector. To reach a high precision in our measurement, a precise calibration of electron energy is required. My personal contribution is basically the extraction of the energy scale factors, responsible of the correction of the mis-calibration of the electromagnetic calorimeter, for the nominal runs collected during Run 2, and for special runs collected with low number of interactions per bunch crossing, to be used in a precise measurement of the W mass. For the special runs, and because of the low statistics, we proposed a new approach for the extraction of the energy scale factors.

Chapter 4 presents a theoretical description of the unfolding problem, focusing on the iterative Bayesian unfolding method used in the high energy physics to correct undesired detector effects. The unfolding is used in the measurement of the transverse momentum of the W boson, the measurement of the differential cross sections and the measurement of the W boson mass.

Chapter 5 gives a detailed study on the measurement of the W-boson transverse momentum distribution at 5 and 13 TeV. A precise measurement of p_{T}^{W} will provide a direct comparison with predictions, and a direct measurement may reduce the QCD modeling uncertainties by a factor of two. My personal contribution concerns the unfolding of distributions at the detector level, the propagation of uncertainties through the unfolding, the estimation of the unfolding bias (a bias introduced with the unfolding procedure), and finally an optimisation study of one of the unfolding parameters.

Chapter 6 reports results for the measurement of the W boson fiducial cross sections using the unfolded distributions of p_{T}^{W}.

Chapter 7 presents detailed studies of the measurement of differential and double differential cross section of the W boson at 5 and 13 TeV. The measurements are based on the unfolded distributions of p_{T}^{ℓ} and η_{ℓ}. My personal contribution is the unfolding of p_{T}^{ℓ} and η_{ℓ} distributions at the detector level and the estimation of the corresponding uncertainties.

Chapter 8 gives preliminary results for the uncertainties of the W boson mass determination, using the template method introduced in Run 1, focusing on the statistical and systematic experimental uncertainties.

0.1 Résumé

Motivation. Au 20^{e} siècle, les physiciens ont commencé à construire un modèle qui décrit toutes les particules de la nature et leurs interactions, à l'exception de celles dues à la gravité, que l'on appelle le modèle standard. Ce modèle est la combinaison de deux théories qui décrivent les particules et leurs interactions dans un cadre unique. Les deux composantes du modèle standard sont la théorie électrofaible, qui décrit les interactions via les forces électromagnétiques et faibles, et la chromodynamique quantique (QCD), la théorie de la force nucléaire forte. Ces deux théories sont des théories de champ de jauge, qui décrivent les interactions entre particules en termes d'échange de particules intermédiaires «messagères». Avec le développement technologique au début des années 1970, des expériences ont été construites pour étudier les particules du modèle standard et leurs interactions.

Après la découverte des bosons W et Z au Super Synchrotron à Protons du CERN, particules responsables d'interactions faibles, les efforts ont été orientés vers la mesure de leurs propriétés avec une grande précision pour tester la cohérence du modèle standard. Le modèle standard dispose de 25 paramètres libres pour décrire les particules et leurs interactions, certains d'entre eux sont mesurés directement, parfois avec une grande précision, les autres paramètres sont contraints par la mesure des grandeurs physiques liées par la théorie. Par exemple, la masse M_W du boson W, en incluant des corrections radiatives, est liée aux masses du boson Z, du boson de Higgs et du quark top. La masse du boson Z est mesurée avec une grande précision par les expériences LEP, et après la découverte du boson de Higgs en 2012 au LHC, la masse du boson W peut être prédite théoriquement et la comparaison entre les valeurs prédites et mesurées est considérée comme un test solide pour la cohérence du modèle standard.

Cette thèse décrit la mesure des propriétés du boson W à partir de données collectées avec le détecteur ATLAS au LHC à deux énergies dans le centre de masse de 5 et 13 TeV, correspondant à une luminosité intégrée de 258 pb^{-1} à $\sqrt{s}$ = 5 TeV et 340 pb^{-1} à $\sqrt{s}$ = 13 TeV. Les données utilisées sont collectées à l'aide des runs spéciaux, avec un faible nombre d'interactions par croisement, afin d'améliorer la résolution expérimentale de la mesure du recul et la reconstruction de l'impulsion transverse manquante et de la masse transverse. Dans ce travail, nous nous intéressons aux désintégrations leptoniques du boson W, et les produits chargés de cette désintégration, électron ou muon, sont accompagnés d'un neutrino. Le neutrino ne peut pas être mesuré directement dans le détecteur mais peut être mesuré indirectement à l'aide du lepton et du système hadronique qui recule contre le boson W. Cette thèse est divisée on deux parties, la première partie étudie la calibration du calorimètre électromagnétique du détecteur ATLAS, utilisé pour determiner l'énergie des électrons et des photons avec une grande précision. La deuxième partie est consacrée à la mesure des propriétés du boson W.

Le calorimètre électromagnétique du detecteur ATLAS au LHC. Le LHC est un collisionneur de particules à haute énergie, approuvé en 1996, avec les premiers faisceaux en 2008 à l'organisation européenne pour la recherche nucléaire (CERN) à la frontière entre la France et la Suisse. Avec une circonférence de 27 km et quatre points d'interaction pour quatre grandes expériences (ATLAS, CMS, ALICE et LHCb), le LHC est actuellement le plus grand et le plus puissant accélérateur de la planète. Le LHC est conçu pour accélérer deux faisceaux de protons à plus de 99,99% de la vitesse de la lumière, qui se déplacent dans des directions opposées autour de l'accélérateur et entrent en collision aux emplacements des quatre expériences principales. ATLAS (A Toroidal LHC ApparatuS) est un détecteur polyvalent développé pour étudier différents programmes de physique : interactions électrofaible, production du boson de Higgs, QCD et signatures possibles de la physique au-delà du modèle standard. Le détecteur ATLAS est situé à 100 mètres sous terre au premier point d'interaction du LHC, d'environ 44 mètres de long et 25 mètres de diamètre, pesant environ 7 000 tonnes. Le détecteur ATLAS est composé de différents sous-détecteurs qui presentent une couverture uniforme autour du tube de faisceau et mesurent différentes propriétés des particules dans les collisions proton-proton au LHC. Près du centre, nous commençons par les détecteurs de trace internes, qui mesurent les trajectoires des particules chargées à proximité du point d'interaction. Les calorimètres électromagnétiques (EM) et hadroniques, qui mesurent l'énergie déposée par les électrons, les photons et les jets hadroniques. Les calorimètres sont entourés par le spectromètre à muons, la couche la plus externe, qui est conçu pour mesurer la trajectoire des muons.

Les particules électromagnétiques, électrons et photons, sont utilisées essentiellement dans toutes les analyses notamment dans les études des propriétés du boson de Higgs et dans la mesure de précision des paramètres électrofaible tels que la masse du boson W, permettant un test de cohérence pour le modèle standard. Les particules électromagnétiques sont arrêtées et mesurées dans le calorimètre EM. Pour atteindre une bonne précision dans nos mesures, un étalonnage précis de l'énergie des électrons et des photons est nécessaire. La procédure d'étalonnage est basée sur des échantillons $Z \to ee$, en raison des statistiques élevées et de l'état final propre qui caractérise ce canal. Dans cette thèse, nous discuterons de l'étalonnage de l'énergie des électrons et des photons pour les données nominales en utilisant un étalonnage "in-situ" du calorimètre EM. L'idée principale de cette méthode est de comparer les distributions de la masse invariante m_{ee} des données et de la simulation, et en utilisant cette comparaison nous pouvons déduire deux facteurs de correction que nous appliquons aux données et aux simulations pour la calibration du calorimètre électromagnétique. En plus des runs appelés runs nominaux, on discute aussi de la calibration des runs spéciaux, utilisé pour des mésures de précision, caractérisés par un faible nombre d'interactions par croisement ($\mu \approx 2$). Pour ces runs, nous proposons deux approches différentes pour l'étalonnage des énergies des électrons, la première est similaire à ce que nous faisons pour les runs nominaux, la seconde consiste à faire une extrapolation des résultats des runs nominaux, ce qui permet de réduire les incertitudes statistiques.

La mesure de la distribution d'impulsion transverse du boson W. L'une des sources d'incertitudes théoriques la plus importante dans la mesure de la masse du boson W est l'extrapolation de la distribution en p_{T} du boson Z au boson W ($\approx$ 6 MeV), où les prédictions d'ordre supérieur de la QCD ne sont pas suffisamment précises pour décrire les données. Une mesure précise du p_{T}^{W} fournira une comparaison directe avec les prédictions QCD, cela revient à dire que le remplacement de l'extrapolation théorique de p_{T}^{Z} par une telle mesure directe de la distribution p_{T}^{W} améliorera la précision de la mesure de M_W. La mesure de la distribution du p_{T}^{W} dans la région du p_{T}^{W} faible ($p_{\mathrm{T}}^{W} < 30$ GeV) avec une incertitude $\sim 1\%$ dans un bin de 5 GeV réduira l'incertitude de modélisation QCD dans la mesure de M_W d'un facteur deux. La distribution de p_{T}^{W} est reconstruite à l'aide d'événements $W \to \ell\nu$, où les leptons chargés sont mesurés dans les différents détecteurs de trace ou dans le calorimètre EM, tandis que le neutrino quitte le détecteur sans être directement mesuré. C'est la raison pour laquelle, la distribution p_{T}^{W} est reconstruite par le recul hadronique, u_{T}, défini comme la somme vectorielle de tous les dépôts d'énergie à l'exclusion de l'énergie du lepton. L'impulsion transverse du boson W est définie par:

$$\vec{p}_{\mathrm{T}}^{W} = -\vec{u}_{\mathrm{T}}. \tag{1}$$

Dans la plupart des cas, les distributions des observales physiques sont affectées par des effets de détecteur: efficacité limitée, migration entre les bins etc. Dans cette thèse on discute de la mesure de la distribution d'impulsion transverse du boson W et de la correction des effets indésirables du détecteur avec la méthode d'unfolding. L'utilisation de la technique d'unfolding en physique des hautes énergies permet d'obtenir des résultats indépendants des effets de détection et de reconstruction. Par conséquent, les résultats d'unfolding peuvent être comparés directement à des prédictions théoriques ou à d'autres expériences.

L'idée principale de l'unfolding consiste à construire une matrice à partir de la simulation, appelée matrice de migration, qui contient des informations au niveau de la vérité et de la reconstruction. L'application de l'inverse de la matrice de migration à la distribution des données permet de passer au niveau de la vérité correspondant aux données, qui ne contient aucun effet de détecteur. Aussi cette thèse discute la propagation des différents sources d'incertitude par l'unfolding, en utilisant des techniques de bootstrapping, fit, etc. Au final, les résultats pour la measure du p_{T}^{W} après l'unfolding sont comparés aux differentes prédictions thèoriques.

Les distributions d'impulsion transverse du boson W, p_{T}^{W}, sont utilisées aussi pour la mesure des sections efficaces fiducielles, ce qui permet de comparer nos résultats avec les prédictions disponibles, incluant les corrections de QCD (NNLO) et EW (NLO). Les sections efficaces sont mesurées en utilisant deux méthodes : une avec la correction bin-par-bin qui consiste à appliquer un factor C_i, déduit de la comparaison des niveaux vérité et reconstruit de la simulation. Alors que la deuxième consiste à utiliser les distributions d'unfolding.

Mesure des sections efficaces simple et double différentielles. Les sections efficaces différentielles sont mesurées en fonction de différentes variables (η_ℓ, p_{T}^{ℓ}) en utilisant les distributions après l'unfolding. La mesure des sections efficaces différentielles dans ce processus fournit des tests rigoureux de la théorie QCD, cruciaux pour une compréhension approfondie et la modélisation des interactions QCD. De plus, la dépendance en fonction de la rapidité de la production de boson W dans le processus Drell–Yan fournit des contraintes sur les fonctions de distribution des partons (PDFs), qui sont actuellement la source d'incertitude dominante dans la mesure de la masse W ($\approx$ 9,2 MeV).

En parallèle, un unfolding à 2 dimensions est réalisé pour mesurer les sections efficaces double différentielles dans les bins de ($\eta_\ell - p_{\mathrm{T}}^{\ell}$). Une technique est utilisée pour transférer l'unfolding bidimensionnel à un unfolding unidimensionnel tel qu'utilisé pour les sections efficaces différentielles de η_ℓ et p_{T}^{ℓ} séparément. Dans les deux cas, les différent sources d'incertitudes (statistiques, systématiques et biais) sont propagées par l'unfolding en utilisant la méme approche que celle de l'analyse de p_{T}^{W}.

La mesure de la masse du boson *W*. Le boson W est une particule instable qui se désintègre en un lepton chargé et un neutrino. La masse du boson W est déterminée en utilisant les distributions de la masse transverse du boson W (m_{T}^{W}) et de l'impulsion transverse du lepton (p_{T}^{ℓ}). L'idée de base de la méthode, appelée "templates" (utilisée pour le Run 1), consiste à calculer les distributions simulées par Monte Carlo (MC) de p_{T}^{ℓ} et m_{T}^{W} pour différentes valeurs supposées de M_W ("templates"), et la comparaison entre les "templates" et les données donne la meilleure valeur de la masse du boson W. En plus de la méthode des "templates", il existe une approche différente consistant à utiliser les distributions au niveau unfolded au lieu des distributions au niveau reconstruit. L'idée principale est d'utiliser des distributions déjà corrigées par la procédure d'unfolding et ne conterant pas d'effets de détecteur indésirables. La masse du boson W et les different sources d'uncertitudes (statistiques et systématiques) sont calculées en utilisant les distributions de p_{T}^{ℓ} et m_{T}^{W} séparément puis combinées pour le résultat final. Puisque nos distributions d'intérêt sont générées à partir des mêmes événements, nous devons prendre en compte la corrélation entre ces deux variables. La corrélation est calculée à l'aide des "Toys" de MC, générés en faisant varier les distributions p_{T}^{ℓ} et m_{T}^{W} simultanément avec une variation aléatoire de Poisson. Cette thèse, donne des résultats préliminaires pour la mesure de la masse du boson W avec les incertitudes statistiques et expérimentales correspondantes, en utilisant les données des runs spéciaux collectés avec un faible nombre d'interactions par croisement (μ $\approx$ 2).

Contents

Chapter 1

Theoretical overview

1.1 Introduction

For a long time, understand the nature of the matter that surrounds us was one of the most interesting questions of philosophers. The first idea to explain the nature of matter is due to ancient Greek philosophers in the 6^{th} century B.C., who introduced the term "atom" to describe the small and indivisible object we can find in nature. The next huge step in the understanding of the matter came in the 18^{th} century, where the chemists started to classify the materials on observed properties and also proposed predictions. However, near the end of the 19th century, physicists discovered that atoms are not the fundamental particles of nature, but conglomerates of even smaller particles.

In the 20^{th} century, physicists started to build a model that describes all the particles in nature and their interactions except those due to gravity, which is so called the Standard Model (SM). This model is the combination of two theories that describe particles and their interactions into a single framework. The two components of the SM are the electroweak theory, which describes interactions via the electromagnetic and weak forces, and quantum chromodynamics, the theory of the strong nuclear force. Both these theories are gauge field theories, which describe the interactions between particles in terms of exchange of intermediary "messenger" particles. This chapter gives an overview of the particles and their interactions in the SM.

1.2 The Standard Model

1.2.1 Elementary particles

The particles in the SM, are divided in two groups called fermions and bosons, and are interacting with each other through three known interactions. The classification of particles is based on their physical properties.

1.2.1.1 Fermions

The fermions [118] in the SM are separated into two groups, leptons and quarks. Leptons are assumed to be elementary with no inner structure, while the quarks are constituent of other particles, hadrons, combined by the strong interaction [154].

The SM fermion sector is organised in three generations as shown in Table 1.1. According to the predictions of relativistic quantum mechanics [80], each fermion has a corresponding anti-particle.

Leptons: They are one of the three classes of particles in the SM. There are six known leptons and they occur in pairs called generations. The three charged leptons (e^-, μ^-, τ^-) are: electron, mu-lepton or muon and the tau-lepton or tau. The three charged leptons have the same charge $Q = -\mathrm{e}$. In addition to charged leptons, there are three neutral leptons-neutrinos (ν_{e^-}, ν_{μ^-}, ν_{τ^-}) called the electron neutrino, muon neutrino and tau neutrino respectively, having very small masses.

Quarks: Considered as elementary particles and without inner structure, quarks are combined by the strong interactions to form the hadrons. Quarks can not be isolated because of the "color confinement" (Sec. 1.2.2.2). The strong interaction regroups quarks with different charges and color charges, to form hadrons. The most well known and stable hadrons are protons and neutrons.

TABLE 1.1: Generations of quarks and leptons with their masses and charges [63].

	1st Generation	2nd Generation	3rd Generation	Charge[e]
Quarks	Up (u) m = 2.3 MeV	Charm (c) m = 1.275 GeV	Top (t) m = 173.2 GeV	$+2/3$
	Down (d) m = 4.8 MeV	Strange (s) m = 95 MeV	Bottom (b) m = 4.18 GeV	$-1/3$
Leptons	Electron (e^-) m = 511 keV	Muons (μ^-) m = 105.7 MeV	Tau (τ^-) m = 1.8 GeV	-1
	Electron neutrino (ν_e) m < 2 eV	Muon neutrino (ν_μ) m < 0.19 MeV	Tau neutrino (ν_τ) m < 18.2 MeV	0

1.2.1.2 Gauge bosons

Called also messenger particles or intermediate particles, the gauge bosons (Table 1.2) give rise to the interactions between particles. Photons are the intermediate particles of electromagnetic interactions, which bond the electrons to the nucleus to form the atoms, and which also bond the atoms together to form the molecules. The W and Z bosons, discovered at CERN in 1983 by the UA1 and UA2 collaborations [90], are the weak interaction messengers. Unlike photons, these particles are characterised by a non-zero mass, and their masses were found to be about 80 GeV and 91 GeV, respectively [71]. Exchange of gluons, the intermediate particles for strong interactions are analogous to the exchange of photons in the electromagnetic force between two charged particles, but for strong interactions, they "glue" quarks together, forming hadrons such as protons and neutrons. The Higgs boson has, contrary to the gauge bosons, a spin 0 and has been discovered at CERN in 2012 by the ATLAS and CMS collaborations [4, 126].

TABLE 1.2: The SM bosons with their masses and charges, and corresponding interaction types [130].

Boson	Mass	Charge	Spin	Interaction	Range	Act on
Photon	0	0	1	Electromagnetism	Infinite	Charge
8 gluons	0	0	1	Strong	10^{-15}m	Colour
$W^{\pm}$	80.4 GeV	$\pm$	1	Weak	10^{-18}m	Weak isospin
Z	91.2 GeV	0	1	Weak	10^{-18}m	Weak isospin and hypercharge
Higgs	125 GeV	0	0			

1.2.2 Fields and interactions

In the SM, there are quantum fields associated to the bosons that are responsible of the interactions between particles. The SM is a theory describing the electromagnetic, weak, and strong interactions by a class of quantum field theories constrained by various symmetry principles.

1.2.2.1 Lagrangian formalism and symmetries

The Lagrangian formalism is an efficient method used to describe variety of physical systems including systems with finite (particles) and an infinite number of degrees of freedom (fields). In the SM, we describe all the interactions with the notion of field, and they are built using the Lagrangian formalism.

The easiest way to understand this formalism, is to take an example of an isolated system in classical physics, where the Lagrangian can be written as:

$$\mathcal{L}(x, \dot{x}, t) = T - V, \tag{1.1}$$

where T and V are the kinetic and potential energy, respectively. In the simple case of a system of a particle of mass m moving along a dimension x in a potential $V(x)$, the Lagrangian can be written as:

$$\mathcal{L}(x, \dot{x}, t) = \frac{1}{2}m\dot{x}^2 - V(x). \tag{1.2}$$

On the other hand, the principle of least action [35] tells us that the action:

$$S = \int_{\mathrm{t1}}^{\mathrm{t2}} \mathcal{L}(x, \dot{x}, t)dt, \tag{1.3}$$

must be minimised or maximised [35], which implies that:

$$\delta S = \int_{\mathrm{t1}}^{\mathrm{t2}} \left(\frac{\partial \mathcal{L}}{\partial q}\delta q + \frac{\partial \mathcal{L}}{\partial \dot{q}}\delta \dot{q} \right) dt = 0, \tag{1.4}$$

leading finally to the Lagrangian equation which describes the movement of the system:

$$\frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} = 0. \tag{1.5}$$

For the example introduced in equation (1.2), by injecting equation (1.2) in (1.5) we find:

$$m\ddot{q} = -\frac{\partial V}{\partial q}, \tag{1.6}$$

which is none other than Newton's first law. On the other hand and from the equation (1.5) we find that

$$\frac{d}{dt}\left[\frac{\partial \mathcal{L}}{\partial \dot{q}}\dot{q} - \mathcal{L}\right] = 0, \tag{1.7}$$

which leads us to another important theorem, Noether theorem [88], which means that for any continuous symmetry of a system, there is a constant associated to the movement (in our case here temporal). This notion of symmetry, or invariance, is a key element in the construction of the SM as a gauge theory.

1.2.2.2 Quantum chromodynamics

Quantum chromodynamics (QCD) [84] is the gauge field theory which describes the strong interactions. The QCD is a local gauge symmetry under the $\mathrm{SU}(3)_c$ group [84]. This symmetry generates eight gluons, massless gauge bosons considered as intermediate particle for strong interactions. Gluons are characterised by a conserved quantum number called the color charge (there are eight color states of gluons, composed by the three colors: red, green or blue, and the three anti-colors). The associated Lagrangian of QCD is:

$$\mathcal{L}_{\mathrm{QCD}} = \bar{\psi}_{q,a}(i\gamma^{\mu}\partial_{\mu}\delta_{\mathrm{ab}} - g_s\gamma^{\mu}t^{c}_{ab}A\mu^{c} - m_q\delta_{ab})\psi_{q,a} - \frac{1}{4}F^{A}_{\mu\nu}F^{A\mu\nu}. \tag{1.8}$$

The γ^{μ} are the Dirac γ-matrices, the $\psi_{q,a}$ represent the field of quark of flavor q [84] and a is the color index (quarks come in three colors). The $A\mu^{c}$ correspond to gluon fields, with c running from 1 to $N_c^2 - 1 = 8$ representing the number of existing gluons. The g_s is the strong coupling constant and is universal for all gluons. The constant $g_s(g_s^2 = 4\pi\alpha_s)$ is a fundamental parameter of QCD and can be written as a function of the energy scale Q as:

$$\alpha_s\left(Q^2\right) \approx \frac{1}{\beta_0 \ln\left(Q^2/\Lambda^2\right)}, \tag{1.9}$$

where β_0 is a constant term related to the number of quark flavors and Λ is the scale of the QCD. The dependence of α_s as a function of the energy scale Q, plotted in Figure 1.1, defines the characteristic properties of QCD interactions:

- Asymptotic freedom: it means that at large Q^2 (small distance) the coupling between quarks becomes weak.
- Quark confinement: it means that at small Q^2 (large distance) the coupling between quarks becomes strong and we cannot find a quark as an isolated particle.

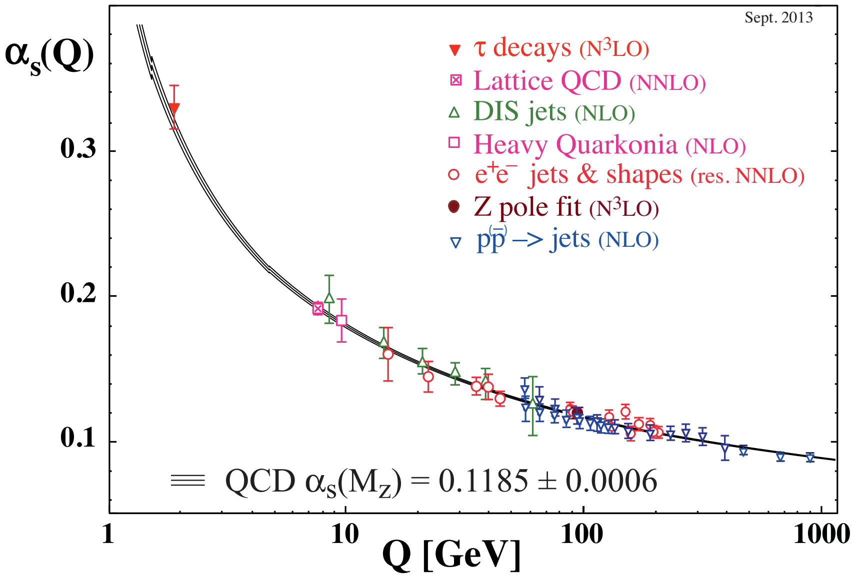


FIGURE 1.1: Evolution of the strong coupling constant as a function of the energy scale measured by various experiments [61].

1.2.2.3 Electroweak interaction

In the SM, the electromagnetic and weak interactions are considered as two different low-energy aspects of a single electroweak (EW) interaction, this theory developed by Glashow, Weinberg, and Salam being known as "GWS theory" [133]. This theory is described by an $\mathrm{SU}(2)_\mathrm{L} \otimes \mathrm{U}(1)_\mathrm{Y}$ gauge group, with the exchange of four mediator bosons: photon, Z, W^+ and W^-. The Lagrangian of the EW theory is described as:

$$\mathcal{L}_\mathrm{EW} = \textstyle\sum_\Psi \bar{\Psi} \left[i\gamma^\mu D_\mu \right] \Psi - \frac{1}{4} W^a_{\mu\nu} W^{\mu\nu}_a - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{1.10}$$

$$D_\mu = \partial_\mu + igT^a W^a_\mu + ig' \tfrac{1}{2} T_Y B_\mu, \tag{1.11}$$

where T^a and T_Y are the generators of $\mathrm{SU}(2)_\mathrm{L}$ and $\mathrm{U}(1)_\mathrm{Y}$, g and g' are the weak and electromagnetic couplings. B_μ and W^a_μ are gauge fields which give rise to the four mediator bosons. The bosons photon, Z, W^+ and W^- can be written as:

$$A_\mu = B_\mu \cos\theta_W + W^3_\mu \sin\theta_W, \tag{1.12}$$

$$Z_\mu = -B_\mu \sin\theta_W + W^3_\mu \cos\theta_W, \tag{1.13}$$

$$W^\pm_\mu = \tfrac{1}{\sqrt{2}} \left(W^1_\mu \mp W^2_\mu \right), \tag{1.14}$$

where θ_W is a mixing parameter called the weak mixing (Weinberg) angle which is precisely measured by experiments: $\sin^2(\theta_\mathrm{W}) = 0.23153 \pm 0.00006$, in the scheme where θ_W is the effective leptonic weak mixing angle [89], and can be expressed in terms of the coupling constants as:

$$\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}}, \tag{1.15}$$

$$\sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}, \tag{1.16}$$

according to the EW Lagrangian (1.11). The fermionic and bosonic fields must be massless to preserve the $\mathrm{SU}(2)_\mathrm{L} \otimes \mathrm{U}(1)_\mathrm{Y}$ gauge symmetry. On the other hand, the experimental observations show the existence of massive bosons and fermions. In 1964, the Brout-Englert-Higgs mechanism [32] proposed a solution to solve this conflict with experiments by adding an additional scalar boson called the (Brout-Englert-) Higgs boson and generating a "Higgs field" which interacts with all the particles. This mechanism is called "spontaneous symmetry breaking". The Lagrangian of the Higgs field can be written as:

$$\mathcal{L}_{\text{Higgs}} = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + \left(D^\mu\Phi\right)^\dagger\left(D_\mu\Phi\right) - V(\Phi), \tag{1.17}$$

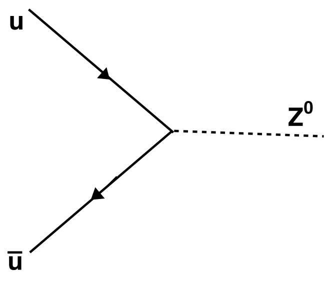
$$V(\Phi) = -\mu^2|\Phi|^2 + \lambda\left(|\phi|^2\right)^2, \tag{1.18}$$

where $\left(D^\mu\Phi\right)^\dagger\left(D_\mu\Phi\right)$ contains the interaction between Higgs and gauge bosons. The Higgs boson mass term is expressed as: $m_H^2 = 2|\mu|^2 = 2\lambda v^2$, while the gauge bosons masses are written as:

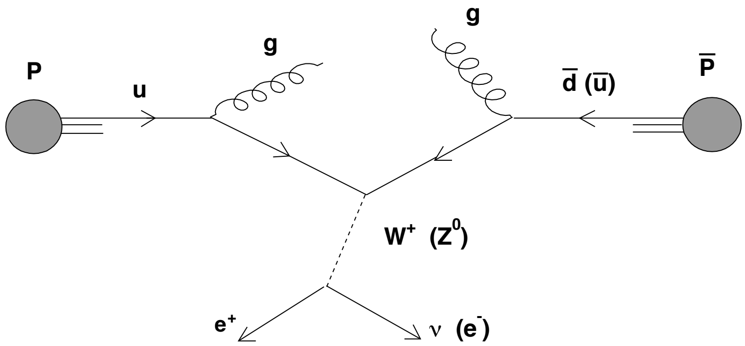
$$m_W^2 = \frac{1}{4}g^2v^2, \quad m_Z^2 = \frac{1}{4}\left(g^2 + g'^2\right)v^2. \tag{1.19}$$

1.3 *W* boson production in *pp* collision

At the LHC [148], the electroweak gauge bosons W and Z are produced from proton−proton collisions (at Tevatron they were produced by proton−antiproton collisions). Each proton is composed of two quarks up (u) and and one quark down (d) which interact through strong interactions by exchange of gluons. Quarks u and d, containing valence quarks, determine the quantum numbers of proton. The production of W and Z bosons at leading order is dominated by quark−antiquark annihilation processes, as seen in Figure 1.2, with $q\bar{q}' \to W$, $q\bar{q} \to Z$ with no momentum in the plane transverse to the beam [132]. However, high order corrections can include radiation of gluons or quarks, where the gluons can self-interact and produce more gluons, and each gluon can also produce a quark and anti-quark pairs, called *sea* quarks, also shown in Figure 1.2.



(A)



(B)

FIGURE 1.2: (A) First-order production diagrams for the W and Z boson. (B) Basic production and leptonic decay for W/Z bosons with radiated gluons.

Eventually, the production of the EW gauge bosons in proton−proton collisions with high order of QCD corrections, is mainly related to the distributions of partons inside each proton. The partonic structure is studied in particular in scattering processes like Deep Inelastic Scattering (DIS) [76], and the resulting Parton Distribution Functions (PDFs) [75] represent the probability density to find partons (quarks and gluons) carrying a momentum fraction x at an energy scale Q. PDF sets cannot be calculated analytically but are obtained by fits to a large number of cross section data points from many experiments [76]. Figure 1.3 shows examples of parton distributions in the proton at two energy scales $Q = 2$ GeV and $Q = 100$ GeV.

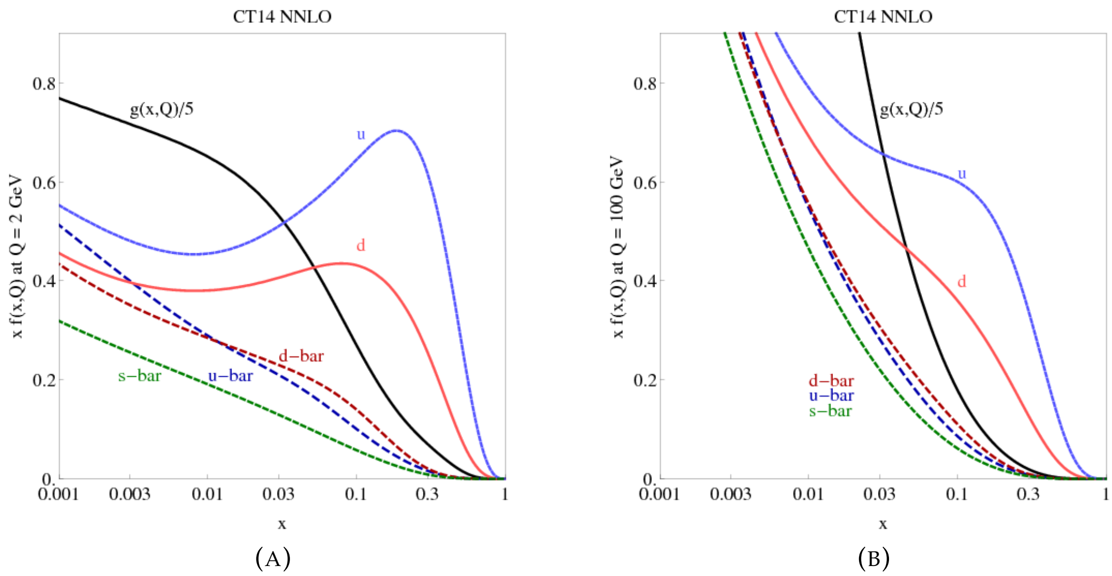


FIGURE 1.3: The CT14 parton distribution functions at $Q = 2$ GeV (A) and $Q = 100$ GeV (B) for u, $\bar{u}$, d, $\bar{d}$, $s = \bar{s}$ and g [65]

In proton−proton collisions, the hadrons interactions can be separated in two types, hard QCD and soft QCD. The hard QCD process for the W boson production corresponds to a production with large momentum transfer Q, and the W boson production cross-section $(p_1 + p_2 \to V + X)$ can be determined using the *Factorization Theorem* [66] and the PDFs $f_i(x, Q)$:

$$\sigma_V\left(h_1\left(p_1\right), h_2\left(p_2\right)\right) = \sum_{a,b} \int_0^1 dx_a dx_b f_{a/h_1}\left(x_a, \mu_F^2\right) f_{b/h_2}\left(x_b, \mu_F^2\right) \times \sigma_{ab\to V}\left(x_a p_1, x_b p_2, \mu_F^2\right). \tag{1.20}$$

where $x_{1,2}$ are the fractions of the momenta of the hadrons and $f_{i,j}$ are the corresponding distributions of quark and anti-quark (a,b), μ_F is the factorization scale that separates hard and soft QCD regimes. The generalisation of equation (1.20) for higher order corrections that can contribute to the W boson production is written as:

$$\sigma_{ab\to V} = \sigma_0 + \alpha_s\left(\mu_R^2\right)\sigma_1 + \alpha_s^2\left(\mu_R^2\right)\sigma_2 + \ldots. \tag{1.21}$$

where μ_R is the renormalisation scale of the QCD running coupling constant, and σ_0 corresponds to the cross section at Leading Order (LO). The terms $\alpha_s(\mu_R^2)\sigma_1$ and $\alpha_s^2(\mu_R^2)\sigma_2$ correspond to the cross-sections at Next-Leading Order (NLO) and Next-to-Next-Leading Order (NNLO). For Drell-Yan processes, the scale parameters μ_F and μ_R are chosen as $\mu_F = \mu_R = M$ [46]. The predictions at NLO order for some important SM cross sections in proton−proton and proton−antiproton collisions are shown in Figure 1.4.

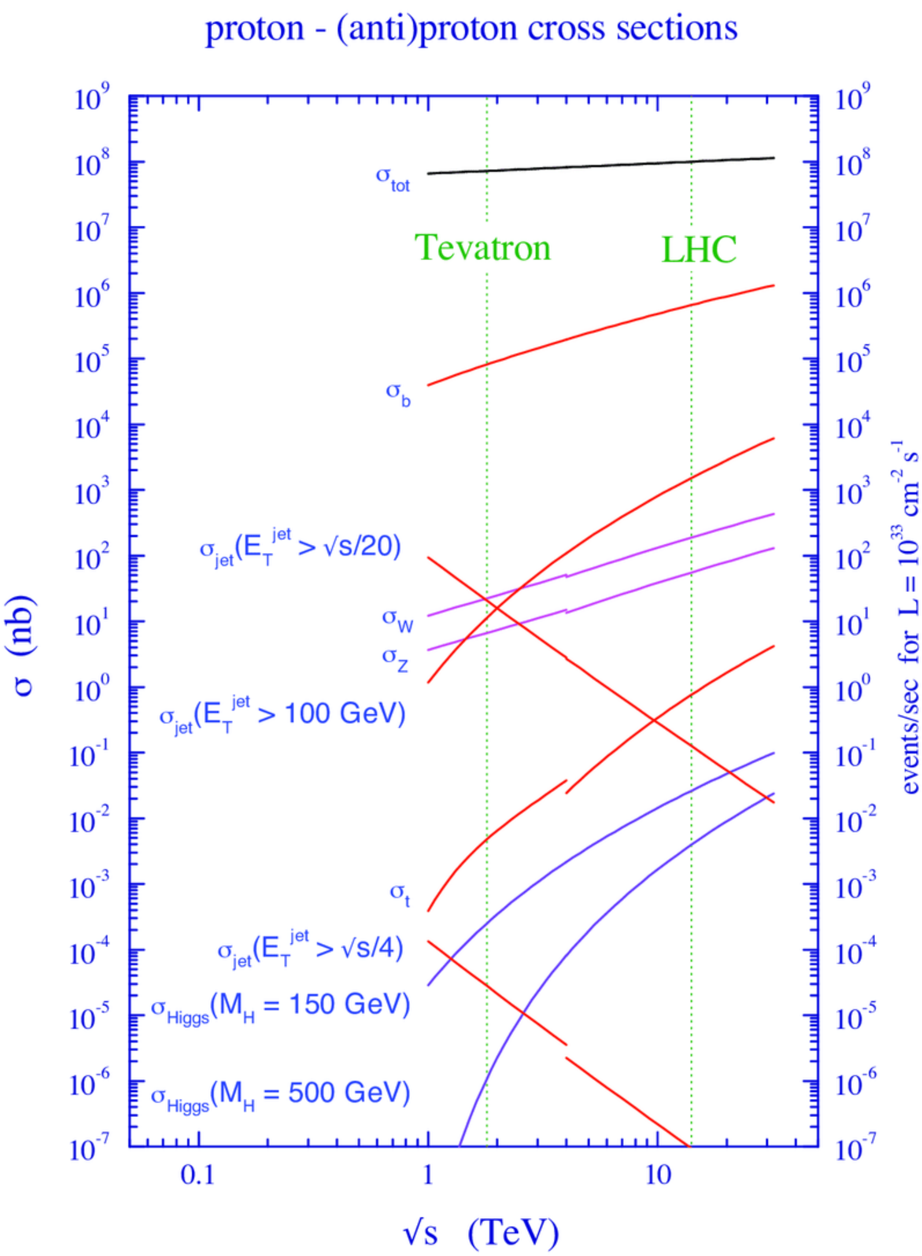


FIGURE 1.4: Standard Model cross sections at the Tevatron and LHC colliders, calculated at next-to-leading order in perturbation theory [141].

1.4 Properties of the W boson

1.4.1 W boson mass

As described in Sec. 1.2.2.3 for the spontaneous symmetry breaking, the fundamental parameters of the electroweak interactions are: the mass of the Higgs boson, the weak mixing angle θ_W, and the coupling constants (g, g'). At lowest order in the EW theory, the W boson mass can be expressed as a function of the fine-structure constant $\alpha(= e^2/4\pi)$, the Fermi constant G_F and the mass of the Z boson. The Fermi constant G_F is a function of the the coupling constant g and calculated using the Fermi model [144]. The Z boson mass is determined with high precision from the Z lineshape scan at LEP1 [138]:

$$\alpha^{-1} = 137.035999074(44), \tag{1.22}$$

$$G_F = 1.1663787(6) \times 10^{-5}\ \mathrm{GeV}^{-2}, \tag{1.23}$$

$$M_Z = 91.1876(21)\ \mathrm{GeV}. \tag{1.24}$$

At this level, the W boson mass can be expressed as:

$$M_W^2 = M_Z^2 \left(\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{\alpha\pi}{\sqrt{2} G_F M_Z^2}} \right). \tag{1.25}$$

It predicts a W mass value of $M_W = 80.939 \pm 0.0026$ GeV. However, higher-order EW corrections introduce an additional dependence on the gauge couplings and the mass of heavy particles of the SM. The W mass boson can be expressed with an additional parameter Δr containing all the high-order corrections:

$$M_W^2 = M_Z^2 \left(\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{\alpha\pi}{\sqrt{2} G_F M_Z^2 (1 - \Delta r)}} \right). \tag{1.26}$$

In summary, the additional parameter Δr depends on the vacuum polarisation contribution of leptons and light quarks, as well as the top-quark and Higgs boson masses and may be sensitive to additional particles and interactions beyond the SM. All those effects make the W mass boson an extremely important parameter of the SM. Producing a W mass measurement with excellent accuracy is accordingly of high importance for testing the overall consistency of the SM, by comparing the experimental measurement of M_W to the theoretical predictions. The determination of the W boson mass at the NLO order (2-loop EW), with some leading NNLO and few N^3LO QCD contributions, is performed with a global fit [28, 89] (see also Refs. [146, 39]) of electroweak parameters. The resulting W mass value is:

$$M_W = 80.359 \pm 0.006\ \mathrm{GeV}. \tag{1.27}$$

1.4.2 Experimental determinations of W boson mass

After the first detection of the W boson by the UA1 [21] and UA2 [29] collaborations [90] in proton−antiproton collisions at the SPS collider in 1983, the obtained M_W value was 81 ± 5 GeV [72] and it was difficult to determine it precisely at this accelerator [60]. However, UA2 produced finally a determination $M_W = 80.35 \pm 0.37$ GeV in 1991 [11]. Later, the W mass was determined [137] in Large Electron-Positron (LEP) collider at CERN [1995-2000] [83]. LEP was accelerating electrons and positrons to reach a center-of-mass energy up to 209 GeV. The direct measurement of the W boson mass at LEP experiments (ALEPH, DELPHI, L3 and OPAL) gives the following result:

$$M_W^{\text{LEP}} = 80.376 \pm 0.025_{\text{stat}} \pm 0.022_{\text{syst}} \text{ GeV}. \tag{1.28}$$

Later, a new determination of M_W was performed in Tevatron experiments (CDF and D0) at Fermilab [2002-2011] [59]. The Tevatron collider was a proton−antiproton collider, where the center of mass energy can reach 1.96 TeV [125]. The M_W was determined from the comparison of kinematical distributions of $W \to l\nu$ with simulated distributions characterised with different M_W values. The direct determination of the W boson mass by the Tevatron experiments (CDF and D0) gives the following result [5]:

$$M_W^{\text{Tevatron}} = 80.387 \pm 0.016 \text{ GeV}, \tag{1.29}$$

The Tevatron and LEP combined results lead to the world average value:

$$M_W = 80.385 \pm 0.015 \text{ GeV}, \tag{1.30}$$

The latest M_W determination is carried out with the ATLAS detector at the LHC [115, 17], using proton−proton collision at a center of mass energy of 7 TeV collected during Run 1 in 2011. The M_W is determined using the lepton transverse momentum (p_{T}^{ℓ}) and transverse mass (m_{T}^{W}) distributions from $W \to \ell\nu$ with the template approach [115]. The W boson transverse mass, m_{T}^{W}, is derived from the missing transverse momentum ($p_{\text{T}}^{\text{miss}}$) and from p_{T}^{ℓ} as follows:

$$m_{\text{T}}^{W} = \sqrt{2 p_{\text{T}}^{\ell} p_{\text{T}}^{\text{miss}} (1 - \cos \Delta\phi)}, \tag{1.31}$$

where $\Delta\phi$ is the azimuthal opening angle between the charged lepton and the missing transverse momentum. The different sources of uncertainties are described in the Table 1.3, and the dominant systematic uncertainties are:

Lepton calibration: The measurement of lepton momentum and energy is derived using information from the Z decay due to its very clean final state. The Run 1 corrections for the leptons with their uncertainties are described in [23, 41, 67] and are studied in detail for this analysis in [158, 152]. Electron energy calibration will be discussed in Chapter 3.

Hadronic recoil uncertainty: It is defined as the uncertainty from the hadronic recoil (HR) calibration [48]. The study for the Run 1 analysis is shown in [140].

The uncertainties in HR calibration are mainly driven by data statistics in the resolution and response corrections [155, 101].

Backgrounds in the W boson sample: The W boson background contributions are estimated using simulation, except for the multijet background using data-driven techniques [155, 157]. The study for the Run 1 analysis is shown in [17].

QCD corrections: The NNLO is used to describe the differential cross-section as a function of boson rapidity and angular coefficient [143, 16]. The QCD uncertainties are coming from the uncertainties in the fixed-order predictions, parton-shower predictions and angular coefficients [115].

Electroweak corrections: Dominated by QED final-state radiation (FSR) [33], the uncertainties are evaluated by comparing the distributions with different computations.

PDF uncertainties: Uncertainties in the PDFs are the dominant source of physics-modelling uncertainty, due to our imperfect knowledge of the PDFs affecting the differential cross section as a function of boson rapidity, the angular coefficients, and the W boson transverse momentum distribution.

The ATLAS experiment gives the following results:

$$M_W^{\text{ATLAS}} = 80.370 \pm 0.019 \text{ GeV}. \tag{1.32}$$

The W boson mass results of ATLAS in comparison with other determinations are shown in Figure 1.5.

TABLE 1.3: The ATLAS M_{W} result with statistical and systematic uncertainties [115].

Combined categories	Value MeV	Stat. Unc.	Muon Unc.	Elec. Unc.	Recoil Unc.	Bckg. Unc.	QCD Unc.	EW Unc.	PDF Unc.	Total Unc.
m_{T}^W, p_{T}^ℓ	80369.5	6.8	6.6	6.4	2.9	4.5	8.3	5.5	9.2	18.5

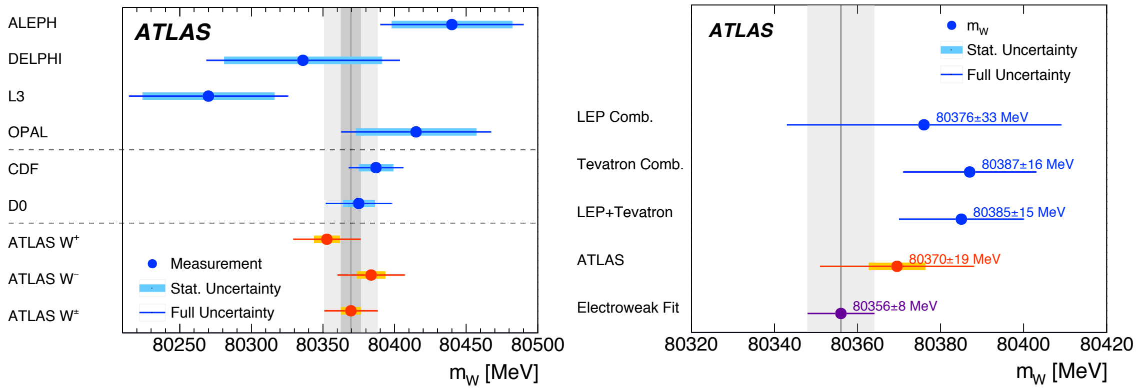


FIGURE 1.5: Left: The M_W results of ATLAS [115] in comparison with other published results from the LEP experiments ALEPH, DELPHI, L3 and OPAL, and from the Tevatron collider experiments CDF and D0. The vertical bands show the statistical and total uncertainties of the ATLAS determination, and the horizontal bands and lines show the statistical and total uncertainties of the other published results. Right: The ATLAS result compared to the SM prediction from the global electroweak fit, and to the combined values determined at LEP and at Tevatron.

Chapter 2

Experimental Setup: The ATLAS experiment at the LHC

The results presented in this thesis are based on data collected during Run 2 by the ATLAS experiment at the Large Hadron Collider (LHC). This chapter provides an overview of the LHC [81] and of the ATLAS experiment [102], with a description focused on the components relevant for the analysis.

2.1 The large hadron collider

The LHC is a high-energy particle collider, approved in 1996, with the first beams in 2008 at the European Organization for Nuclear Research (CERN) [30] at the border of France and Switzerland. With a circumference of 27 km and with four interaction points for four large experiments (ATLAS, CMS, ALICE and LHCb [122]), the LHC is currently the largest and most powerful accelerator on Earth. The LHC is designed to accelerate two beams of protons to more than 99.99% the speed of light, which travel in opposite directions around the accelerator and collide at the locations of the four major experiments. In the LHC, the particles are grouped together in about 2000 bunches in each beam, which can contain 10^{11} particles per bunch [92], and reach an energy up to 6.5 TeV per beam. The beams are therefore at a center of mass energy up to 13 TeV [92] and collide every 25 nanoseconds.

2.1.1 The LHC acceleration chain

Before being injected in the LHC, particles are accelerated through a series of lower energy accelerators that successively increase the energy of the colliding beams [145]. The starting point is a cylinder of hydrogen gas, where the electrons are stripped from hydrogen atoms before injecting the protons in the linear accelerator LINAC2 to begin the first phase of acceleration up to an energy of 50 MeV [81]. Afterwards, the beam of protons is injected into the Proton Synchrotron Booster (PSB) which accelerates them to an energy of 1.4 GeV. The proton bunches are then injected into the Proton Synchrotron (PS) in which they are accelerated to an energy of 26 GeV. After the PS, the 7 km long Super Proton Synchrotron (SPS) accelerates them to reach an energy of 450 GeV. In the last step, the proton beams are injected in the LHC where they are accelerated to their current maximal energy 6.5 TeV [81].

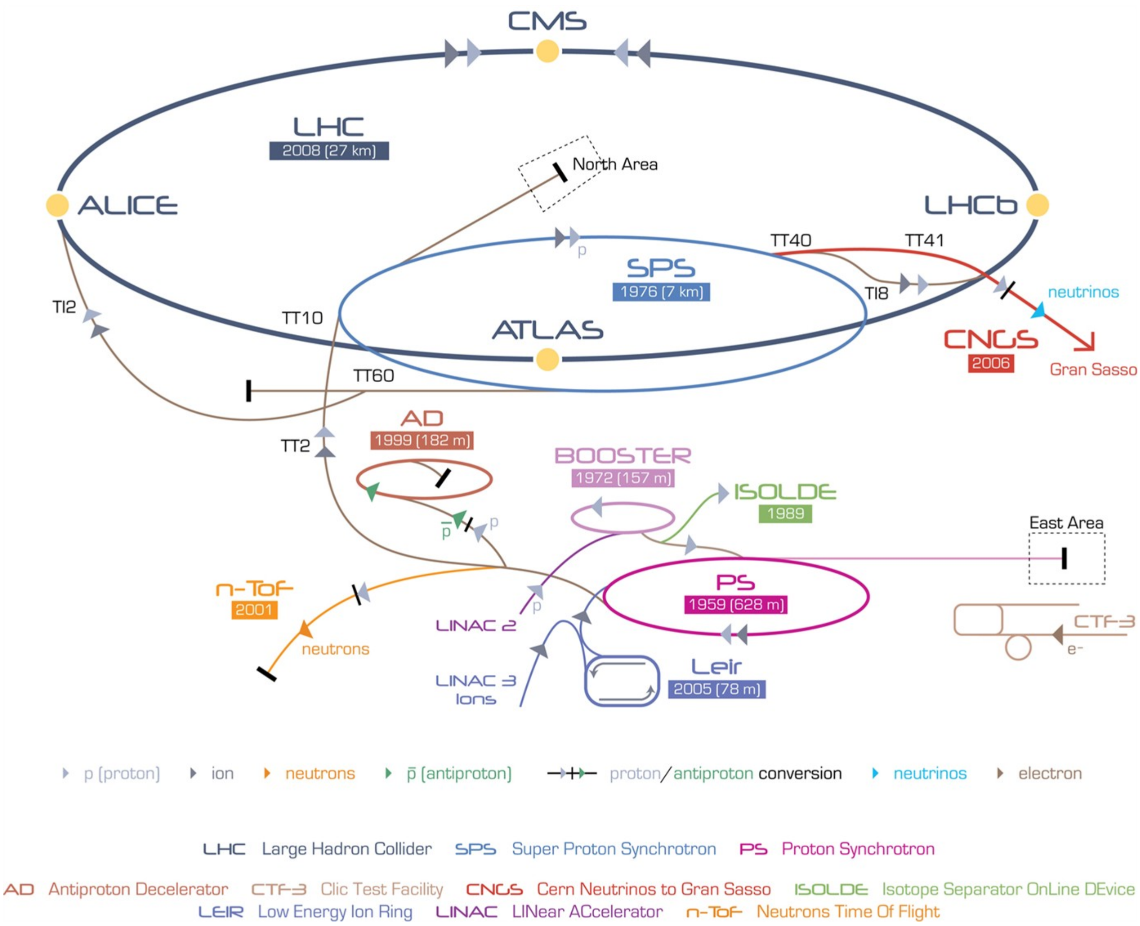


FIGURE 2.1: An overview of the LHC acceleration chain at CERN [119]

2.1.2 LHC performance

The performance of the LHC can be parameterised with two factors, the center of mass-energy which allows to estimate the energy available for the production of new processes, and the instantaneous luminosity [105] (expressed in units $\mathrm{cm}^{-2}\mathrm{s}^{-1}$) which represents the rate of physics process a collider is able to produce. The instantaneous luminosity (in the limit of no crossing angle between the beams) is defined as:

$$L_{\text{inst}} = \frac{N_1 N_2 f_r n_b}{4\pi\sigma_x\sigma_y}, \tag{2.1}$$

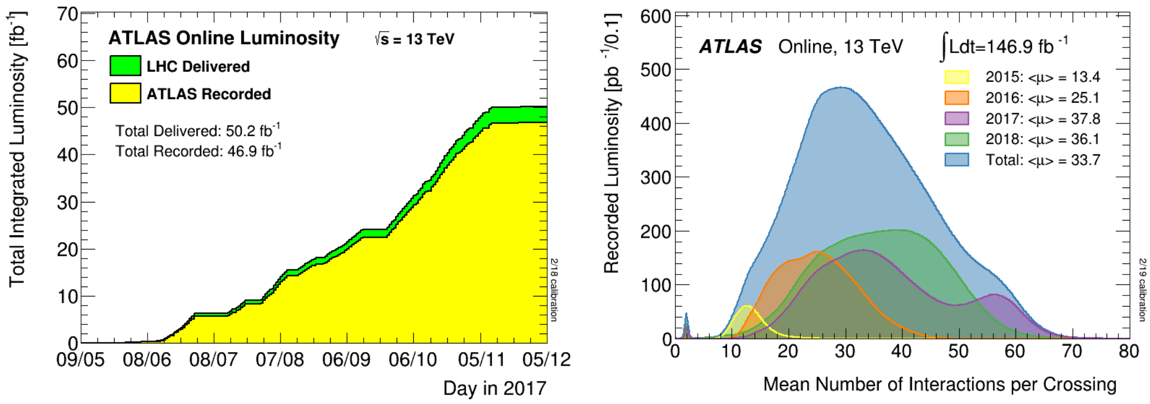
where n_b is the number of bunches in a beam, f_r is the bunch revolution frequency in the LHC, N_1 and N_2 are the number of protons per colliding bunch, σ_x and σ_y are the horizontal and vertical beam size (about 7 μm for the Run 2 in the standard working point). The integrated luminosity is the integral over the data taking time of the instantaneous luminosity:

$$L_{\text{int}} = \int L_{\text{inst}}(t)d(t), \tag{2.2}$$

and it is directly connecting the number of events to the cross-section by:

$$L_{\text{int}} \times \sigma_{\text{process}} = N_{\text{process}}. \tag{2.3}$$

Another significant parameter for our analysis is the pileup, which is the number of inelastic proton−proton interactions per bunch crossing. The average number of proton−proton collisions per bunch crossing is named as $\langle\mu\rangle$ [124]. The dataset used in our analysis is a special dataset characterised with low pileup ($\langle\mu\rangle$ = 2) taken in 2017 and 2018 during Run 2.



(A) (B)

FIGURE 2.2: (A) Integrated luminosity versus time delivered (green) and recorded (yellow) by ATLAS during stable beams for pp collisions at 13 TeV center of mass energy. (B) Mean number of interactions per crossing $\langle\mu\rangle$ per year in Run 2 [106]

2.2 The ATLAS detector

ATLAS (A Toroidal LHC ApparatuS) [52] is a general-purpose detector developed to study different physics programs: SM electroweak interactions, Higgs boson production, hard QCD and possible signatures of BSM physics. An overview of the ATLAS detector components can be seen in Figure 2.3. The ATLAS detector is located 100 meters underground at the LHC first interaction point, approximately 44 meters long and 25 meters in diameter, weighing around 7000 tons [50]. The ATLAS detector is composed of different sub-detectors [52] which give uniform coverage around the beam pipe and measure different properties of particles in proton−proton collisions at the LHC [24]. Near the center, we start by the inner tracker detectors, which measure the trajectories of charged particles close to the interaction point. The electromagnetic and hadronic calorimeters, which measure the energy deposited by electrons, photons and hadronic jets. The calorimeters are surrounded by the muon spectrometer, the outermost layer, which is designed to measure the trajectory of muons.

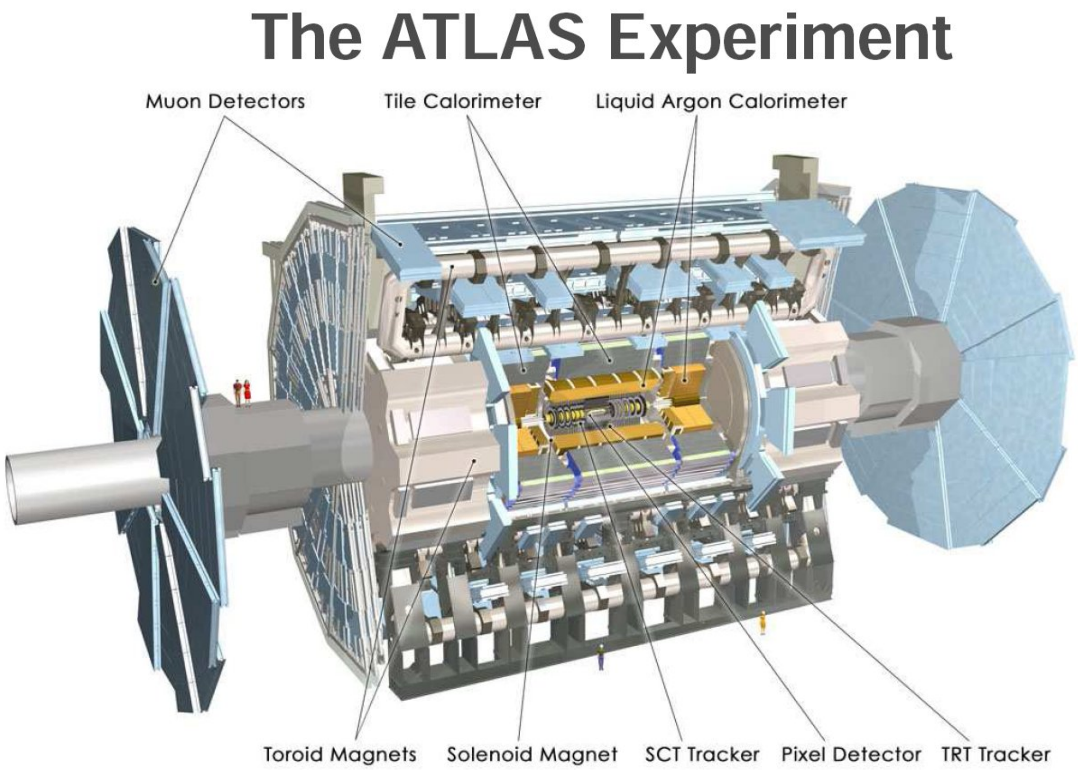


FIGURE 2.3: An overview of the ATLAS detector at CERN [42].

2.2.1 Coordinate system

The nominal interaction point of pp collisions is defined as the origin of the ATLAS coordinate system. The beam direction defines the z-axis where the positive direction is defined as oriented counter clockwise to the LHC ring, while the x is horizontal, orthogonal to the beam pipe and pointing towards the center of the LHC . The y direction is defined as orthogonal to the beam pipe and pointing upwards. The (x, y, z) frame is a right handed frame. Because of the symmetry of the ATLAS detector, a polar coordinate system (ϕ, θ, z) is used, with ϕ being defined with respect to the x-axis and θ with respect to the z-axis, as shown in Figure 2.4. The angle $\theta = 0$ is parallel to the z-axis while $\theta = \pi/2$ is in the xy-plane. In most cases, the pseudo-rapidity η is used to instead of θ and is defined as $\eta = -\ln[\tan(\theta/2)]$, where $\Delta\eta$ is invariant under boosts along the z-axis. Another important variable ΔR is used to calculate the distances between two particles in the $\theta - \eta$ space and is defined as $\Delta R = \sqrt{(\Delta\theta)^2 + (\Delta\eta)^2}$.

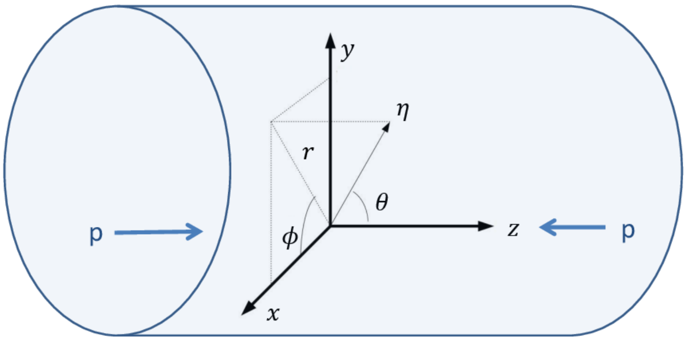


FIGURE 2.4: An overview of the ATLAS coordinate system [128].

2.2.2 Inner tracking detectors

The ATLAS Inner Detector (ID) [149] is the closest detector to the collision point, and it is responsible for the reconstruction of the tracks of charged particles emitted in pp collisions. In the normal (high) pileup mode one has approximately 1000 particles produced in a bunch crossing within the acceptance of the ID (each 25 ns). The inner detector contributes also with the calorimeter and muons spectrometer to the identification of electron, photon and muon. As shown in Figure 2.5, the ID consists of three sub-detectors: the silicon pixel detector, the Semiconductor Tracker (SCT) and Transition Radiation Tracker (TRT):

Silicon pixel detector [150]: It is built with four concentric cylindrical layers around the beampipe (in the barrel). The most-inner layer is called the Insertable B-Layer (IBL) and was installed between Run 1 and Run 2. The pixel detector is reconstructed with a pixel size of $50\mu\text{m} \times 400\mu\text{m}$ ($50\mu\text{m} \times 250\mu\text{m}$ for the IBL). The pixel detector is used for b-tagging and track reconstruction.

Semi-conductor tracker [147]: It is the second part of the inner detector, with four layers of silicon microstrips (in the barrel). The SCT is used for the measurement of momentum and to identify the vertex of charged particles.

Transition Radiation Tracker [129]: This sub-detector surrounds the SCT sub-detector, and consists of multiple layers of straw tubes with a diameter of 4 mm. The TRT is used for momentum measurement and provides discrimination between electrons and hadrons.

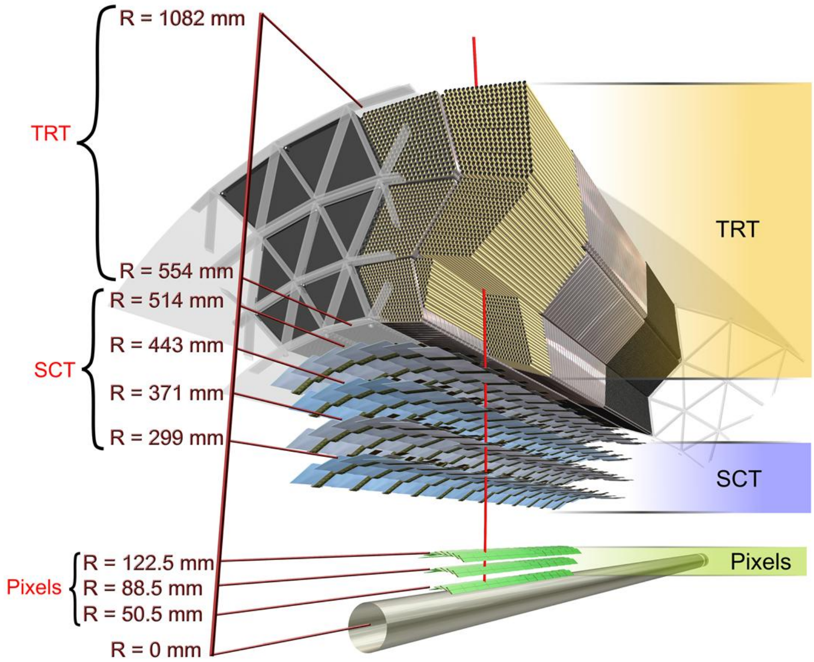


FIGURE 2.5: Schematic showing the ATLAS inner detector [134].

2.2.3 Electromagnetic and hadronic calorimeters

The ATLAS calorimeters system is composed of two main sub-detectors: the electromagnetic (EM) [10] and hadronic [79] calorimeters. The two calorimeters are designed to stop and measure the energy of particles coming from pp collisions (or other processes) and sensitive to electromagnetic or strong interactions: the EM calorimeter, which targets EM showers and measures the energies of electrons and photons, and the hadronic calorimeter, which targets hadronic showers and measures the energy of hadrons. The inner sub-detector is the EM calorimeter, surrounded by the hadronic calorimeter. Both calorimeters are composed of the barrel and two symmetric end-caps, as shown in Figure 2.6, and cover the acceptance up to $|\eta| = 4.9$.

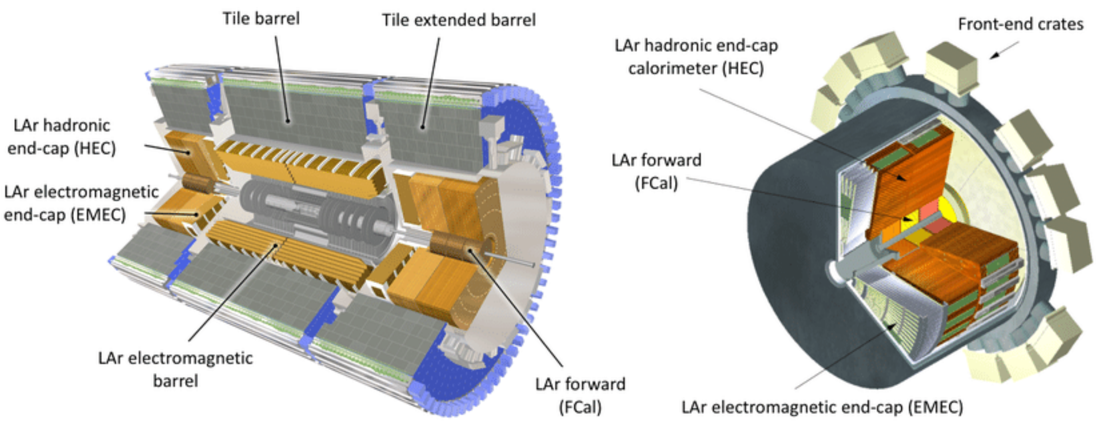


FIGURE 2.6: An overview of ATLAS calorimeter system [142].

2.2.3.1 The electromagnetic calorimeter

The ATLAS EM calorimeter is the most relevant sub-detector for this thesis. In this paragraph, the EM showers and the different components of the ATLAS EM calorimeter are described:

EM shower: An EM shower begins when a high-energy particle (electron, positron or photon) enters a material. Depending on their properties (charge, mass ...), particles interact differently with matter. In our case we are interested in high-energy electrons and photons interactions. Figure 2.7 shows the fraction of energy loss by electrons in lead (used as an absorber in the ATLAS EM calorimeter) and the photon interaction cross-section as a function of their energies. **Electrons** with low energies lose their energy mainly through ionisation and excitation (collisions with the atoms and molecules of the material and the transferred energy is enough to unbind an electron from this atom), while electrons with energies larger than $\simeq$ 10 MeV, lose their energy with bremsstrahlung (interaction of the incoming particle in the electric field of an atom and emission of a high energy photon). **Photons** with low energies, lose their energy through Compton scattering (photons mainly scatter on the electrons of the atoms constituting the medium) and photoelectric effect (emission of electrons). For photons with energies larger than $\simeq$ 10 MeV, interactions result in conversion, produce electron$-$positron pairs. **Electrons and photons** with high energy ($\geq$ 1 GeV) interact with matter to produce secondary photons by bremsstrahlung and electron$-$positron by pair-production with lower energy. These in turn will interact with the matter with the same mechanisms as described before until they lose their energy. This avalanche of produced electrons, positrons, and photons is known as an **EM shower**.

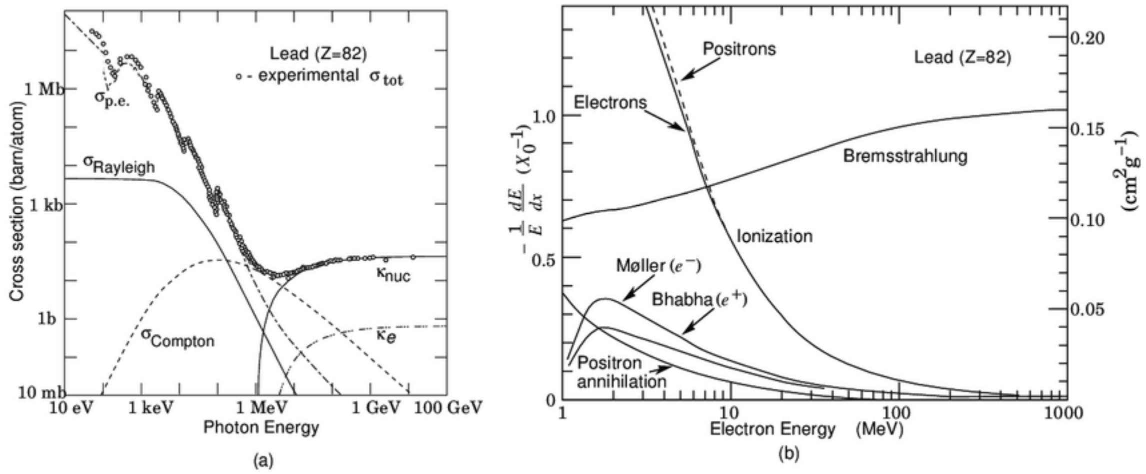


FIGURE 2.7: (a) Photon energy loss in lead as a function of its energy. (b) Electron energy loss in lead as a function of its energy [98].

Energy resolution of an EM calorimeter: The energy resolution of an EM calorimeter can be described by [70]:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (2.4)$$

where the fist term on the right side is the stochastic (S) term, being due to the fluctuations related to the physical development of the shower [73], the second term is a noise (N) term, coming from the electronic noise of the signal readout chain and the pileup noise, the last term is a constant (C) term, coming from instrumental defects that cause variations of the calorimeter response [73], and is independent of the particle energy.

The ATLAS EM calorimeter is a lead liquid Argon (LAr) sampling calorimeter. It is designed with an accordion geometry, an original idea of D. Fournier [78], in order to avoid azimuthal cracks in the detector (ϕ symmetry) [26]. The EM shower is generated when particles interact with the absorber (lead). Secondary particles produced by these interaction ionise the argon and produce ionisation electrons. These ionisation electrons drift towards the anode following the electric field lines produced by the high voltage connected to the electrodes. During their drift, these ionisation electrons induce on the electrodes (see Figure 2.8) an electric current proportional to the number of electrons drifting in the medium, and at the end proportional to the energy deposited.

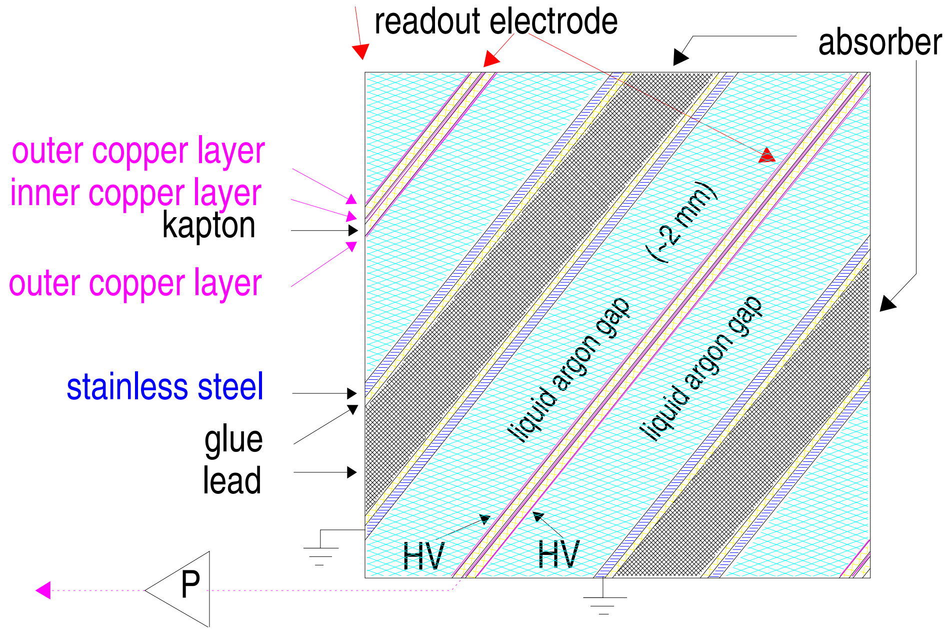


FIGURE 2.8: A sketch of the LAr EM calorimeter [123].

The EM calorimeter [26] has two main parts: the **Barrel** which consists of two half-barrels separated with a gap of 4mm and covers the regions $|\eta| < 1.37$, and two **end-caps** placed at each end of the barrel which cover the regions $1.52 < |\eta| < 3.2$. The part of the end-cap used for precise measurements stops at $|\eta| \approx 2.4$. The region between the barrel and the end-cap is called the transition region and corresponds to $1.37 < \eta < 1.52$, characterised by the presence of a large amount of dead material and is not used in precision measurements like the decay of the Higgs boson into two photons. For the Run 1 W mass analysis [14] a larger part of the detector, corresponding to $1.2 < \eta < 1.82$ was excluded, due to the higher quality wanted and small mismodeling in this region [96]. More details about the EM calorimeter can be found in different theses [6, 13, 141, 31, 82]. The EM calorimeter is divided in three layers: front, middle and back (as shown in Figure 2.9):

- Front layer (L1): It has a very fine segmentation along η: $\Delta\eta \times \Delta\phi = 0.0031 \times 0.1$ in the barrel and $\Delta\eta \times \Delta\phi$ varying between 0.0031×0.1 and 0.0062×0.1 in the end-cap. The fine granularity in η allows to separate a single photon from photons coming from: $\pi^0 \rightarrow \gamma\gamma$.

- Middle layer (L2): it is where the particles deposit most of their energy. The cells in the middle layer are of size η: $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ in the barrel and in the end-cap.

- Back layer (L3): it is where part of the shower leaking after L2 is measured. The cells in the middle layer are of size η: $\Delta\eta \times \Delta\phi = 0.05 \times 0.025$ in the barrel and the end-cap.

In front of the LAr EM calorimeter, there is for $0 < |\eta| < 1.8$ a presampler, which is also based on LAr technology. A detailed description of the LAr EM calorimeter and of the presampler can be found in Table 2.1.

TABLE 2.1: Description of the composition of the LAr calorimeter [25].

	Barrel		End-cap	
EM calorimeter				
Number of layers and $\lvert\eta\rvert$ coverage				
Presampler	1	$\lvert\eta\rvert < 1.52$	1	$1.5 < \lvert\eta\rvert < 1.8$
Calorimeter	3	$\lvert\eta\rvert < 1.35$	2	$1.375 < \lvert\eta\rvert < 1.5$
	2	$1.35 < \lvert\eta\rvert < 1.475$	3	$1.5 < \lvert\eta\rvert < 2.5$
			2	$2.5 < \lvert\eta\rvert < 3.2$
Granularity $\Delta\eta \times \Delta\phi$ versus $\lvert\eta\rvert$				
Presampler	0.025×0.1	$\lvert\eta\rvert < 1.52$	0.025×0.1	$1.5 < \lvert\eta\rvert < 1.8$
Calorimeter 1st layer	$0.025/8 \times 0.1$	$\lvert\eta\rvert < 1.40$	0.050×0.1	$1.375 < \lvert\eta\rvert < 1.425$
	0.025×0.025	$1.40 < \lvert\eta\rvert < 1.475$	0.025×0.1	$1.425 < \lvert\eta\rvert < 1.5$
			$0.025/8 \times 0.1$	$1.5 < \lvert\eta\rvert < 1.8$
			$0.025/6 \times 0.1$	$1.8 < \lvert\eta\rvert < 2.0$
			$0.025/4 \times 0.1$	$2.0 < \lvert\eta\rvert < 2.4$
			0.025×0.1	$2.4 < \lvert\eta\rvert < 2.5$
			0.1×0.1	$2.5 < \lvert\eta\rvert < 3.2$
Calorimeter 2nd layer	0.025×0.025	$\lvert\eta\rvert < 1.40$	0.050×0.025	$1.375 < \lvert\eta\rvert < 1.425$
	0.075×0.025	$1.40 < \lvert\eta\rvert < 1.475$	0.025×0.025	$1.425 < \lvert\eta\rvert < 2.5$
			0.1×0.1	$2.5 < \lvert\eta\rvert < 3.2$
Calorimeter 3rd layer	0.050×0.025	$\lvert\eta\rvert < 1.35$	0.050×0.025	$1.5 < \lvert\eta\rvert < 2.5$
Number of readout channels				
Presampler	7808		1536 (both sides)	
Calorimeter	101760		62208 (both sides)	
LAr hadronic end-cap				
$\lvert\eta\rvert$ coverage			$1.5 < \lvert\eta\rvert < 3.2$	
Number of layers			4	
Granularity $\Delta\eta \times \Delta\phi$			0.1×0.1	$1.5 < \lvert\eta\rvert < 2.5$
			0.2×0.2	$2.5 < \lvert\eta\rvert < 3.2$
Readout channels			5632 (both sides)	
LAr forward calorimeter				
$\lvert\eta\rvert$ coverage			$3.1 < \lvert\eta\rvert < 4.9$	
Number of layers			3	
Granularity $\Delta x \times \Delta y$ (cm)			FCal1: 3.0×2.6	$3.15 < \lvert\eta\rvert < 4.30$
			FCal1: $\sim$ four times finer	$3.10 < \lvert\eta\rvert < 3.15$,
				$4.30 < \lvert\eta\rvert < 4.83$
			FCal2: 3.3×4.2	$3.24 < \lvert\eta\rvert < 4.50$
			FCal2: $\sim$ four times finer	$3.20 < \lvert\eta\rvert < 3.24$,
				$4.50 < \lvert\eta\rvert < 4.81$
			FCal3: 5.4×4.7	$3.32 < \lvert\eta\rvert < 4.60$
			FCal3: $\sim$ four times finer	$3.29 < \lvert\eta\rvert < 3.32$,
				$4.60 < \lvert\eta\rvert < 4.75$
Readout channels			3524 (both sides)	
Scintillator tile calorimeter				
	Barrel		Extended barrel	
$\lvert\eta\rvert$ coverage	$\lvert\eta\rvert < 1.0$		$0.8 < \lvert\eta\rvert < 1.7$	
Number of layers	3		3	
Granularity $\Delta\eta \times \Delta\phi$	0.1×0.1		0.1×0.1	
Last layer	0.2×0.1		0.2×0.1	
Readout channels	5760		4092 (both sides)	

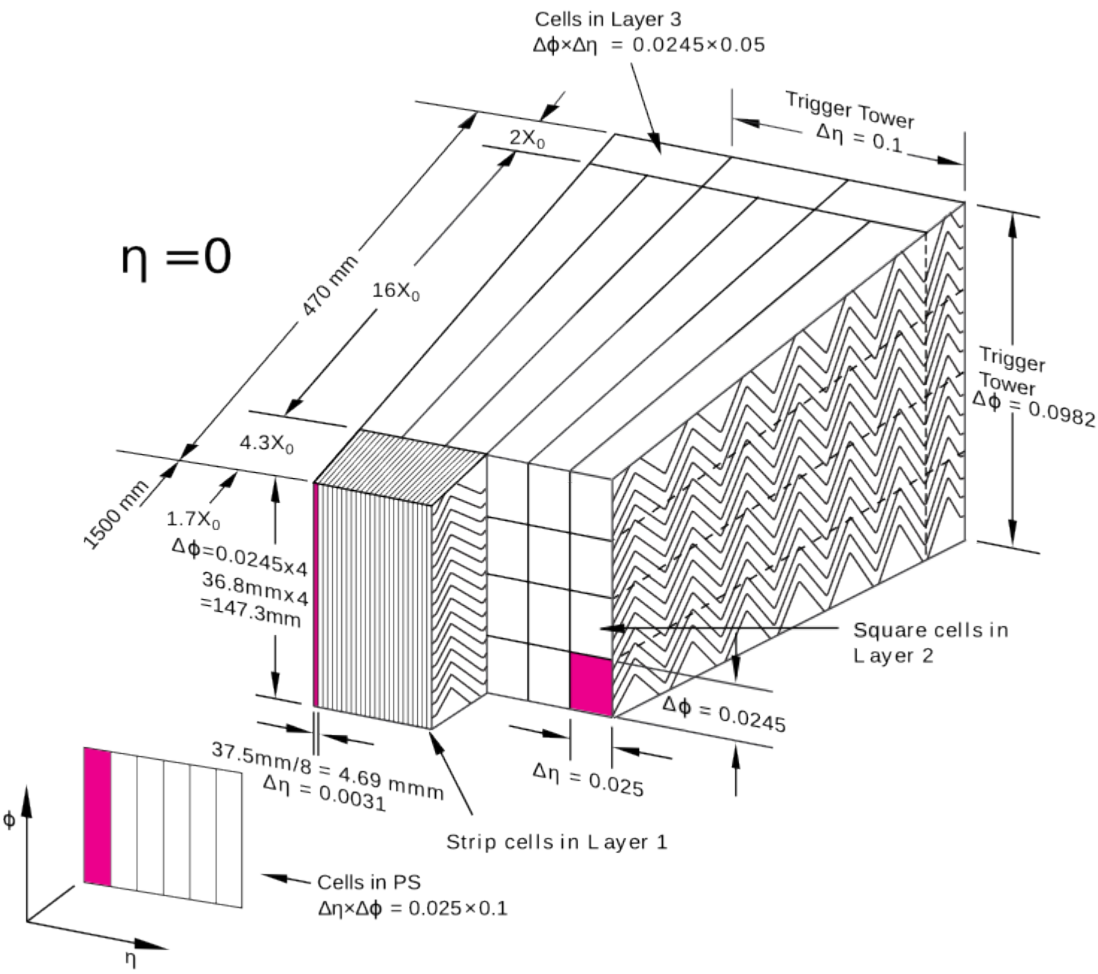


FIGURE 2.9: A sketch of the LAr EM calorimeter layers [135].

2.2.3.2 The ATLAS tile hadronic calorimeter

The tile hadronic calorimeter is located behind the EM calorimeter and operates in a similar way but uses iron as an absorber and scintillating tiles as active material. The tile hadronic calorimeter is composed of three layers covering the range $|\eta| < 1.7$. The first two layers have the same granularity $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ while $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$ is the granularity of the last layer. The tile hadronic calorimeter is used to measure the position and energy of the jets.

2.2.4 Muon spectrometer and toroidal magnets

The ATLAS muon spectrometer, shown in Figure 2.10, is the outermost part of the ATLAS detector and is designed to measure the position and the energy of particles that are able to pass through the inner detectors [86]. Since the muons pass through the calorimeter system with little interaction and therefore conserving most of their initial energy, they are detected with high efficiency in the Muon Spectrometer (MS). It consists of four main types of detectors:

Monitor Drift Tubes (MDTs): They are used for the precision measurement of muon momentum and cover the entire MS detection region $|\eta| < 2.7$. They are built with straw aluminum tubes with 30 mm diameter and each tube is filled with an Ar/CO2 mixture (93% and 7%). The muons ionise the gas and signals of the ionisation electrons are measured.

Cathode Strip Chambers (CSCs): Because of the radiation level in $2.0 < |\eta| < 2.7$ [19], the CSCs replace the MDTs in the most inner layer and provide a precise track measurement.

Resistive Plate Chambers (RPCs): In the RPCs two parallel plates are separated by a thin layer of gas filled with C2H2F4 and SF6. The RPCs provide a track identification and trigger measurements in the barrel region $|\eta| < 1.05$.

Thin Gap Chambers (TGCs): TGCs are multi-wire proportional chambers filled with n-C5H12. The purpose of the TGCs is to replace RPCs in the end-cap regions, $1.05 < |\eta| < 2.4$.

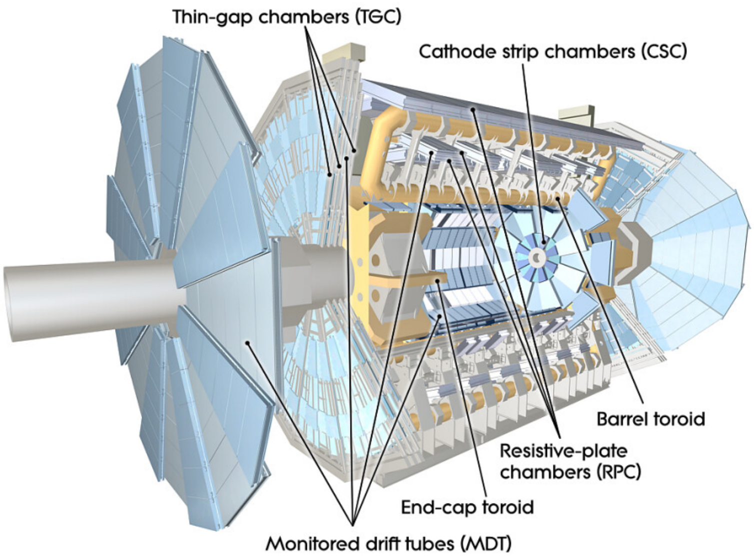


FIGURE 2.10: An overview of ATLAS muon system [3].

2.2.5 ATLAS trigger and data acquisition system

Within the ATLAS detector the proton bunches collide every 25 ns, and can produce nearly 600 terabytes of raw data every second [128]. Because of the limited storage (each event is characterised with a size of the order of 1 Mb) it is impossible to record all these interactions. The aim of the trigger system is to select events having desired signatures. The trigger system selects between 100 and 1000 events per second out of 1000 million in total [128]. During Run 2, the trigger system [136], was divided in two parts (as shown in Figure 2.11):

The L1 trigger: It is a hardware trigger and performs the first stage of the trigger. The L1 trigger uses inputs from the muon spectrometer and the calorimeter systems and searches for signatures from high-p_T muons, electrons/photons, jet and τ-lepton decays in order to choose desired events. The

L1 trigger reduces the event rate from the LHC bunch crossing of 40 MHz to about > 100 kHz.

The High Level Trigger (HLT): The events that have been triggered by the L1 are then filtered by the HLT in order to reduce the rate to 1 kHz. The HLT reconstructs events using a finer granularity of the data with ID tracks to remove most of the pre-selected events.

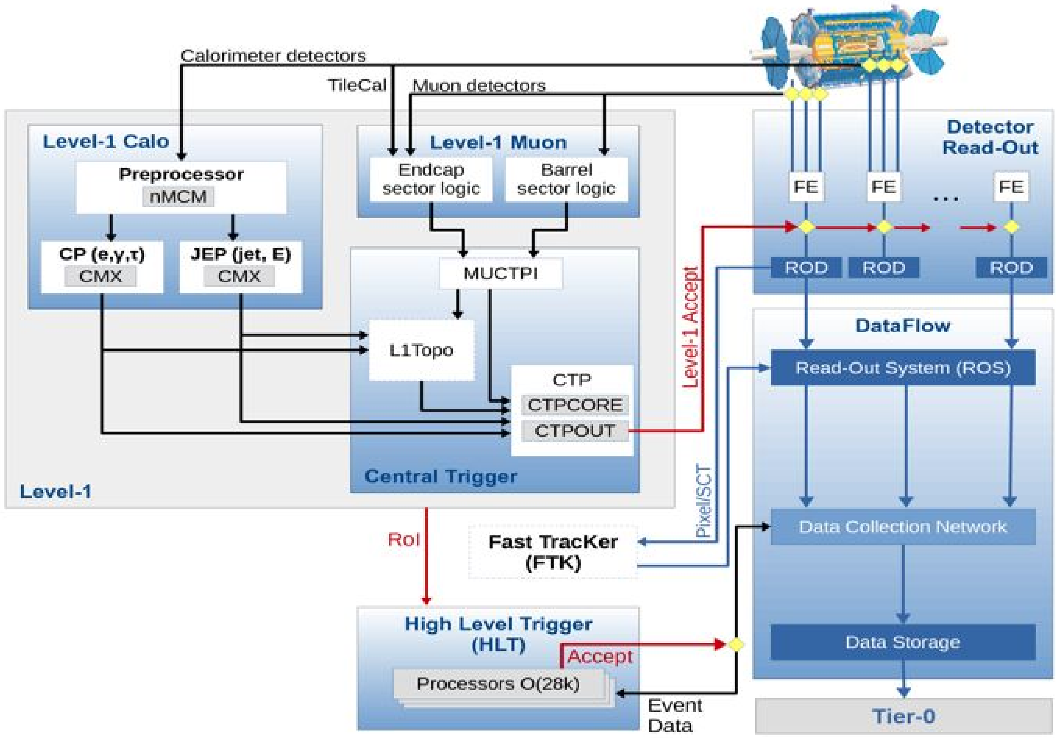


FIGURE 2.11: An overview of ATLAS trigger system [136].

2.3 Reconstruction of physics objects

This section gives an overview of the identification of the electrons and their reconstruction by the ATLAS detector. The electrons will be used for the calibration of the ATLAS EM calorimeter as discussed in Chapter 3.

2.3.1 Electron reconstruction

Electrons and photons are reconstructed in the EM calorimeter (see Chapter 3). When electrons and photons enter to the EM calorimeter, they interact with the lead absorbers and create the EM showers. The EM showers ionise the liquid argon and the ionisation electrons will drift thanks to a high voltage which produces an electric field between the electrodes. During their drift, these ionisation electrons induce on the electrode an electric current. The charge collection time in the electrode is $t_d \approx 450$ ns and the induced signal has a triangular shape as shown in Figure 2.12. Since the charge collection time (450 ns) is longer than the time difference between two bunch crossings at the LHC (25 ns), we will integrate in the charge collection time several bunch crossings and include a lot of pileup events. In order to reduce this effect, the signals are passed through a bipolar filter which shape the signals as shown in Figure 2.12 in order to be more peaked and therefore to have a smaller contribution from pileup.

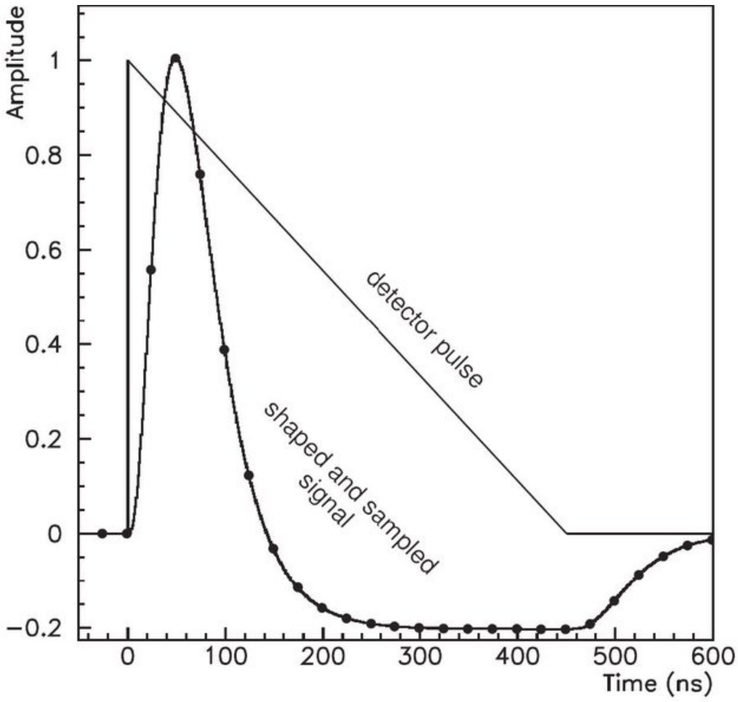


FIGURE 2.12: The pulse shape in the ATLAS LAr calorimeters. The triangular shape is the current pulse generated in the liquid argon by ionisation electrons. The dots shows the positions of the samples separated by 25 ns [127].

The pulses recorded for a cell are used to reconstruct the cell energy in MeV with the formula:

$$E_{\text{cell}} = F_{\mu A \to \text{MeV}} \times F_{\text{DAC} \to \mu A} \times \frac{1}{\frac{M\text{phys}}{M\text{cali}}} \times G \times A \tag{2.5}$$

where:

- $F_{\mu A \to \text{MeV}}$: relates the current from ionisation electrons to the energy deposited in the EM cell [116].
- $F_{\text{DAC} \to \mu A}$: is a conversion factor related to the digital-to-analog converter (DAC).
- $\frac{M\text{phys}}{M\text{cali}}$: is used to correct the gain to take into account the fact that the injected calibration signal is exponential while the physics signal is triangular (see Figure 2.12), and have therefore slightly different maximum amplitudes after the bipolar shaping. It can be obtained from delayed calibration runs, as it is described in [54].

- G: represents the cell gain, measured during the calibration runs (expressed in ADC $\rightarrow$ DAC).
- A: is the signal amplitude extracted using the optimal filtering method [51, 13].

The procedure to reconstruct electrons starts by building a cluster using the measured cell energies in the EM calorimeter, these energies being obtained by equation (2.5). At the beginning of Run 2, a sliding-window clustering algorithm was used, but since 2017, a new algorithm called "dynamical topological cell clustering algorithm" is used. This new algorithm improves the measurement of the electron and photon energy, specially when an electron is emitting a photon by bremsstrahlung [69]. The main difference between the "sliding-window clustering" and the "topological clustering" algorithms, is that the first one is characterised by a fixed-size window, unlike the topological clustering where the selection of cells in a cluster depends on a parameter, ς^{EM}_{cell}, called cell significance, and computed as:

$$\varsigma^{EM}_{cell} = \left| \frac{E^{EM}_{cell}}{\sigma^{EM}_{noise,cell}} \right| \tag{2.6}$$

where E^{EM}_{cell} is the absolute cell energy at the EM scale [69] and $\sigma^{EM}_{noise,cell}$ is the expected cell noise (electronic and pileup noise). This algorithm starts by building clusters of EM cells, called *topo-cluster*. Each *topo-cluster* is built using the same procedure:

- A *topo-cluster* includes cells characterised by $\varsigma > 4$.
- The neighboring cells with $\varsigma > 2$ are added to the *topo-cluster*.
- All neighboring cells with $\varsigma > 0$ are added to the *topo-cluster*.

The procedure of grouping cells in *topo-cluster* is called also $4 - 2 - 0$ which refers to the values of the thresholds on ς. Figure 2.13 shows an overview of the topo-cluster construction.

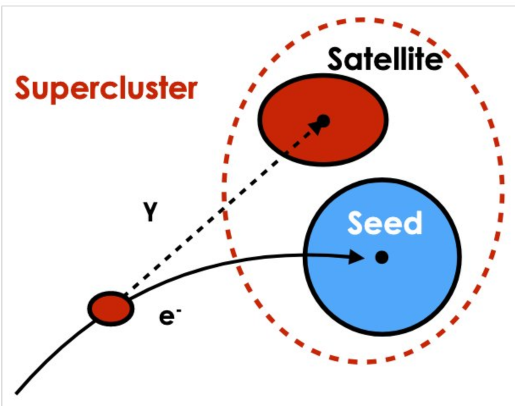
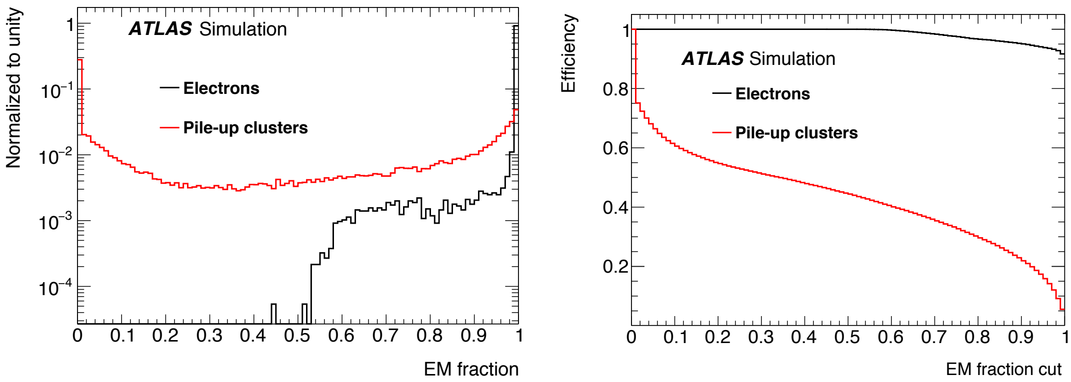


FIGURE 2.13: Illustration of a *topo-cluster* construction [69].

In addition to the procedure described above, there are other selections applied to a topo-cluster to ensure a large rejection of pileup, and to isolate clusters that are primarily from showers in the EM calorimeter. Those selections are based on the factor f_{EM} computed as:

$$f_{\mathrm{EM}} = \frac{E_{\mathrm{L1}} + E_{\mathrm{L2}} + E_{\mathrm{L3}}}{E_{\mathrm{Cluster}}} \tag{2.7}$$

where $E_{\mathrm{L1}}, E_{\mathrm{L2}}$ and E_{L3} are the energies deposited in the first, second and last layers, E_{Cluster} is the energy in the cluster. At the end, only topo-clusters with $f_{\mathrm{EM}} > 0.5$ and $E_{\mathrm{Cluster}} > 400$ MeV are kept. As shown in Figure 2.14, the selection $f_{\mathrm{EM}} > 0.5$ allows to reject over than $\approx 60\%$ of pileup clusters without changing the reconstruction efficiency of true electron topo-clusters [69].



(A) (B)

FIGURE 2.14: (A) Distribution of f_{EM}. (B) Reconstruction efficiency as a function of f_{EM} [69].

2.3.2 Electron identification

In fact, not all of the electrons reconstructed by the "topological clustering" algorithms are prompt electrons. In order to reject background objects, an identification algorithm is used to select prompt electrons and photons from the backgrounds coming from hadronic jets, prompt electrons from photon conversions, and QCD jets. The identification algorithm is based on a likelihood-based (LH) identification, where we use information from the tracking system and the calorimeter system. The discriminant variables are based on the EM shower information, and are shown in Table 2.2.

TABLE 2.2: List of the discrimination variables used in the electron and photon identification [68].

Category	Description	Name	Usage
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the ranges $\|\eta\| < 0.8$ and $\|\eta\| > 1.37$)	R_{had_1}	e/γ
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $0.8 < \|\eta\| < 1.37$)	R_{had}	e/γ
EM third layer	Ratio of the energy in the third layer to the total energy in the EM calorimeter	f_3	e
EM second layer	Ratio of the sum of the energies of the cells contained in a 3×7 $\eta \times \phi$ rectangle (measured in cell units) to the sum of the cell energies in a 7×7 rectangle, both centred around the most energetic cell	R_η	e/γ
	Lateral shower width, $\sqrt{(\Sigma E_i \eta_i^2)/(\Sigma E_i) - ((\Sigma E_i \eta_i)/(\Sigma E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	w_{η_2}	e/γ
	Ratio of the sum of the energies of the cells contained in a 3×3 $\eta \times \phi$ rectangle (measured in cell units) to the sum of the cell energies in a 3×7 rectangle, both centred around the most energetic cell	R_ϕ	e/γ
EM first layer	Total lateral shower width, $\sqrt{(\Sigma E_i (i - i_{max})^2)/(\Sigma E_i)}$, where i runs over all cells in a window of $\Delta\eta \approx 0.0625$ and i_{max} is the index of the highest-energy cell	$w_{s\,tot}$	e/γ
	Lateral shower width, $\sqrt{(\Sigma E_i (i - i_{max})^2)/(\Sigma E_i)}$, where i runs over all cells in a window of 3 cells around the highest-energy cell	$w_{s\,3}$	γ
	Energy fraction outside core of three central cells, within seven cells	f_{side}	γ
	Difference between the energy of the cell associated with the second maximum, and the energy reconstructed in the cell with the smallest value found between the first and second maxima	ΔE_s	γ
	Ratio of the energy difference between the maximum energy deposit and the energy deposit in a secondary maximum in the cluster to the sum of these energies	E_{ratio}	e/γ
	Ratio of the energy measured in the first layer of the electromagnetic calorimeter to the total energy of the EM cluster	f_1	e/γ
Track conditions	Number of hits in the innermost pixel layer	$n_{innermost}$	e
	Number of hits in the pixel detector	n_{Pixel}	e
	Total number of hits in the pixel and SCT detectors	n_{Si}	e
	Transverse impact parameter relative to the beam-line	d_0	e
	Significance of transverse impact parameter defined as the ratio of d_0 to its uncertainty	$\|d_0/\sigma(d_0)\|$	e
	Momentum lost by the track between the perigee and the last measurement point divided by the momentum at perigee	$\Delta p/p$	e
	Likelihood probability based on transition radiation in the TRT	eProbabilityHT	e
Track–cluster matching	$\Delta\eta$ between the cluster position in the first layer of the EM calorimeter and the extrapolated track	$\Delta\eta_1$	e
	$\Delta\phi$ between the cluster position in the second layer of the EM calorimeter and the momentum-rescaled track, extrapolated from the perigee, times the charge q	$\Delta\phi_{res}$	e
	Ratio of the cluster energy to the measured track momentum	E/p	e

Chapter 3

Calibration of the electromagnetic calorimeter

3.1 Introduction

Electromagnetic particles, electrons and photons, are used essentially in all analyses in particular in the studies of the Higgs boson properties and in the precision measurement of electroweak parameters such as the W boson mass, allowing for a consistency test for the Standard Model. As described in Chapter 2, electromagnetic particles are stopped and measured in the EM calorimeter. To reach a good precision in our measurements, a precise electron and photon energy calibration is required. The calibration procedure is based on $Z \to ee$ samples, because of the high statistics and clean final state which characterises this channel. In this chapter, we will discuss the electron and photon energy calibration for the nominal and low pile-up data collected during Run 2 with the ATLAS detector.

3.2 Overview of the calibration procedure

The calibration of the EM calorimeter is a complex procedure and was established during Run 1 [67]. The aim of the calibration procedure, summarised in Figure 3.1, is to measure the energy of electrons and photons with the best precision and resolution. In order to estimate the signal and background contribution, the generated events are passed through a full simulation of the ATLAS detector using GEANT4 [8].

The calibration procedure starts with the energy in EM calorimeter clusters (see Chapter 2), and can be described as follows:

Step 1: based on a MultiVariate Algorithm (MVA) [107], it allows to determine the energy of electrons and photons using the calorimeter cluster properties, measured by the EM calorimeter. The MVA is performed separately for electrons, converted and unconverted photons [67, 100, 151].

Step 2: this step is related to the EM calorimeter design. In fact, the energy of electrons and photons is obtained using the energy deposit in different layers of the EM calorimeter. This step equalises the energy scales of the different longitudinal layers between data and simulation [67].

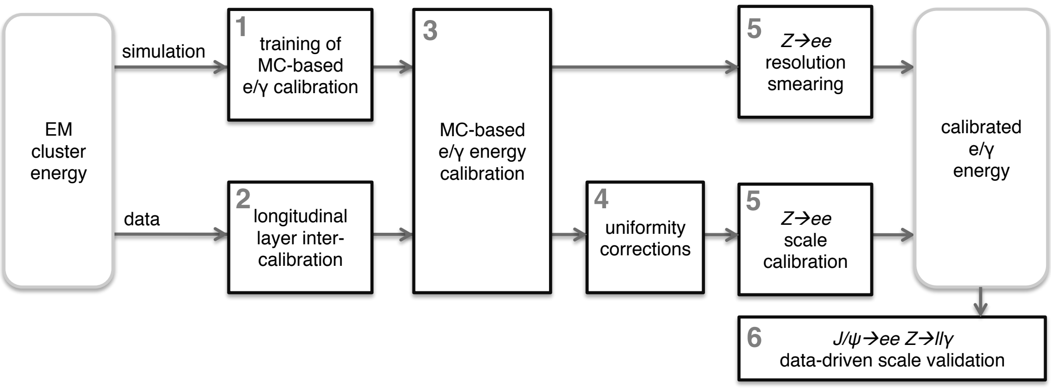


FIGURE 3.1: Schematic overview of the procedure used to calibrate the energy response of electrons and photons in ATLAS [67].

Step 3: the MC-based calibration determined in previous steps is applied to the energy of the clusters in data and simulation.

Step 4: the aim of this step is to include corrections which take into account the uniformity of the calorimeter energy reconstruction as: high-voltage inhomogeneities [18], (where a perfect correction is taken into account in the detector simulation for the zones where there is a "stable" problem) geometric effects such as the inter-module widening (IMW) [113] which are not taken into account in the detector simulation, or biases related to the EM in the detectecelectronic calibration [18].

Step 5: at this step of the calibration procedure, the electron response in data is calibrated to match the expected value in simulation. Also, an additional correction factor aiming to correct the resolution is applied to the simulation, in order to match the data. This step is an important part of this thesis and will be discussed in Sec. 3.3.

Step 6: is the last step, and it does the validation of the scale extracted in step 5 using $J/\psi \to ee$ and $Z \to \ell\ell\gamma$ processes.

In this thesis we will focus on the extraction of the energy scale factors showed in step 5 of Figure 3.1, for the nominal runs and low pile-up runs used for the precise measurement of M_W. As the Z boson mass is precisely measured in LEP experiments [58] and there is a large statistics of Z bosons in ATLAS, the Z boson decay channel ($Z \to ee$) is used for the extraction of the energy scale factors.

3.3 Energy scale and resolution determination with electrons from $Z \to ee$ decays

3.3.1 Overview

After applying the first steps of the calibration procedure (steps 1 to 4 described in Figure 3.1), we still observe an important difference between data and simulation. The sources of the difference are not precisely known. This difference between data and simulation can be seen in the Figure 3.2, which shows the di-electron invariant mass m_{ee} at the step 4 of the calibration procedure, as defined in Figure 3.1, and computed as:

$$m_{ee} = \sqrt{2E_1E_2\left(1 - \cos\theta_{12}\right)}, \tag{3.1}$$

where θ_{12} is the angle between the two electrons measured by the track, and E_1, E_2 are their energies. The discrepancies showed in Figure 3.2 affect the central value

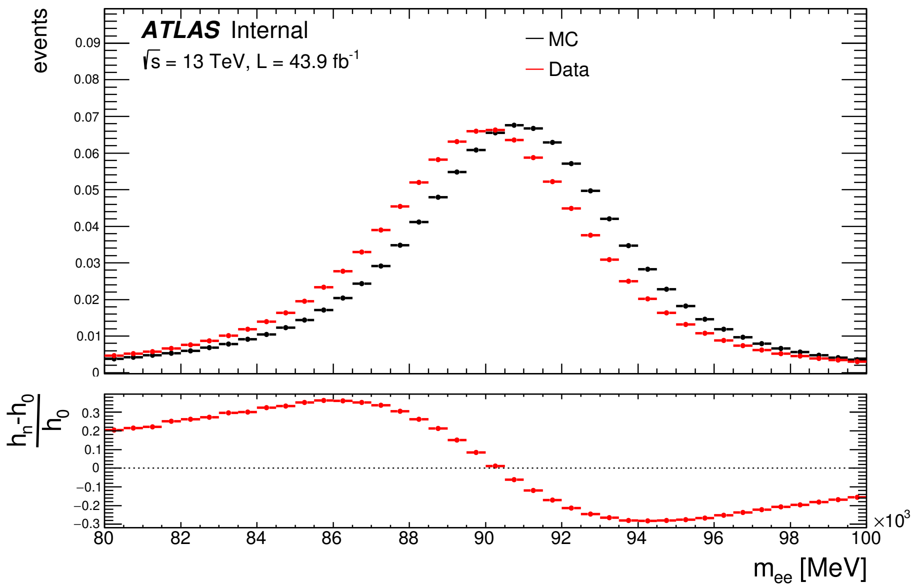


FIGURE 3.2: The di-electron invariant mass m_{ee} after step 4 of the calibration procedure, Figure 3.1, for data and simulation.

of the energy response and the energy resolution. To correct for this difference between data and simulation, two correction factors are extracted. The next paragraph will discuss the methodology used to extract those correction factors.

3.3.2 Definition of the correction factors

As discussed in the previous paragraph, two correction factors are extracted from the $Z \to ee$ channel. The correction factors are called the energy scale factors α and the additional constant term c'. The factors (α, c') will be expressed in η bin i, defined in sec. 3.5.1, as (α_i, c'_i):

- The energy scale factor α: it is applied to the data in order to match the energy response of the simulation:

$$E_i^{\text{corr}} = \frac{E_i^{\text{data}}}{1 + \alpha_i} \tag{3.2}$$

where E^{data} is the measured energy and E^{corr} is the corrected energy.

- The additional constant term c': it is applied to the simulation to be in agreement with the energy resolution of the data:

$$\left(\frac{\sigma(E)}{E}\right)_i^{\text{corr}} = \left(\frac{\sigma(E)}{E}\right)_i^{\text{MC}} \oplus c'_i \tag{3.3}$$

  where $\sigma(E)^{\text{corr}}$ is the resolution of the simulation after applying c', supposed to be equal to $\sigma(E)^{\text{data}}$, which is the resolution of the data, and $\sigma(E)^{\text{MC}}$ is the resolution of the simulation before applying c'.

3.3.3 Effect of the scale factors (α, c') on the di-electrons mass m_{ee}

The scale factors (α, c') are computed using the comparison of the di-electrons invariant mass between data and simulation. Before discussing the method used for the extraction of the scale factors, we will discuss in this part the effect of the scale factors (α, c') on the invariant mass m_{ee}:

$$m_{ee}^{\text{MC}} = \sqrt{2E_1^{\text{MC}} E_2^{\text{MC}} \left(1 - \cos\theta_{12}\right)}, \tag{3.4}$$

by replacing E_1^{MC} and E_2^{MC} with their expressions as shown in equation (3.2). The effect of the scale factor α on m_{ee} is expressed as:

$$m_{ee}^{\text{data}} = m_{ee}^{\text{MC}} \sqrt{\left(1+\alpha_i\right)\left(1+\alpha_j\right)}, \tag{3.5}$$

where i and j are η_{calo} bins where each electron falls in. By neglecting the term of the second order ($\alpha_i \times \alpha_j$) the invariant mass is expressed as:

$$m_{ee}^{\text{data}} = m_{ee}^{\text{MC}}(1 + \alpha_{i,j}), \tag{3.6}$$

and $\alpha_{i,j}$ is written as:

$$\alpha_{i,j} = \frac{\alpha_i + \alpha_j}{2}. \tag{3.7}$$

Contrary to the scale factor α, we can not relate directly the additional constant term c' and the resolution on the invariant mass. Instead, the di-electron invariant mass resolution is expressed in term of the relative energy resolution as:

$$\begin{aligned}\left(\frac{\sigma(m)}{m}\right)^2_{\text{data}} &\simeq \frac{1}{4}\left(\left(\frac{\sigma\left(E_1\right)}{E_1}\right)^2_{\text{data}} + \left(\frac{\sigma\left(E_2\right)}{E_2}\right)^2_{\text{data}}\right) \\ &= \frac{1}{4}\left(\left(\frac{\sigma\left(E_1\right)}{E_1}\right)^2_{\text{MC}} + c'^2_i + \left(\frac{\sigma\left(E_2\right)}{E_2}\right)^2_{\text{MC}} + c'^2_j\right) \\ &= \left(\frac{\sigma(m)}{m}\right)^2_{\text{MC}} + \frac{c'^2_i + c'^2_j}{4}.\end{aligned} \tag{3.8}$$

What is done is to apply to both electrons (with independent random numbers) an effective correction $c'_{i,j}$ and the resolution is therefore expressed as:

$$\left(\frac{\sigma(m)}{m}\right)^2_{\text{data}} = \left(\frac{\sigma(m)}{m}\right)^2_{\text{MC}} + \frac{c_i'^2 + c_j'^2}{4} = \left(\frac{\sigma(m)}{m}\right)^2_{\text{MC}} + \frac{c_{ij}'^2}{2}, \tag{3.9}$$

and $c'_{i,j}$ is written as:

$$c_{ij}'^2 \equiv \frac{c_i'^2 + c_j'^2}{2}. \tag{3.10}$$

Finally, the calibration of the EM calorimeter is simplified to the extraction of the correction factors $\alpha_{i,j}$ and $c'_{i,j}$. To extract these correction factors, the template method in [36, 38] is used for the early Run 2 analysis with a sliding window clustering algorithm, and in [15] for the final Run 2 algorithm with the dynamical topo-cell clustering algorithm (see Chapter 2). The next paragraph will give a detailed explanation of this method.

3.4 Template method for the energy scale factors

3.4.1 Methodology of the template method

The template method described in [82] was established during Run 1 for the extraction of the correction factors $\alpha_{i,j}$ and $c'_{i,j}$. The corrections are determined independently in each (η^i_{calo}, η^j_{calo}) configuration. The idea of the template method is to apply hypothesized values of the scale factors to simulation. For each MC event, the di-electron invariant mass is modified and expressed as:

$$m_{ee}^{\text{template}} = m_{ee}^{\text{MC}}\sqrt{\left(\left(1 + c'_{i,j} \times N_i(0,1)\right)\left(1 + c'_{i,j} \times N_j(0,1)\right)(1+\alpha_{i,j})\,(1+\alpha_{i,j})\right)}. \tag{3.11}$$

For each couple ($\alpha_{i,j}$, $c'_{i,j}$), the new mass distribution is called a template. The comparison between the template and data distributions is done using a χ^2 test, by default for mass values between 80 and 100 GeV:

$$\chi^2 = \sum_{k=1}^{N_{\text{bins}}} \frac{\left(m_{ee,k}^{\text{template}} - m_{ee,k}^{\text{data}}\right)^2}{\left(\sigma_k^{\text{template}}\right)^2 + \left(\sigma_k^{\text{data}}\right)^2}, \tag{3.12}$$

where N_{bins} is the number of mass bins used, $m_{ee,k}^{\text{template}}$ and $m_{ee,k}^{\text{data}}$ are the bin contents of the invariant mass distributions in bin k and $\sigma_k^{\text{template}}$ and σ_k^{data} are the corresponding uncertainties in the bin. By repeating this procedure for all the templates, we can plot a 2D scan of the χ^2 as shown in the Figure 3.3. The minimum of the distribution gives the final correction factors ($\hat{\alpha}_{i,j}$, $\hat{c}'_{i,j}$), which correspond to the best agreement between data and simulation. The determination of the correction factors is related to the determination of the minimum of χ^2.

There is in Figure 3.3 a small correlation between α and c' and the minimum of this 2D distributions is obtained using several 1D fits. The minimisation procedure [82] is shown in Figure 3.4 and can be summarised in the steps below:

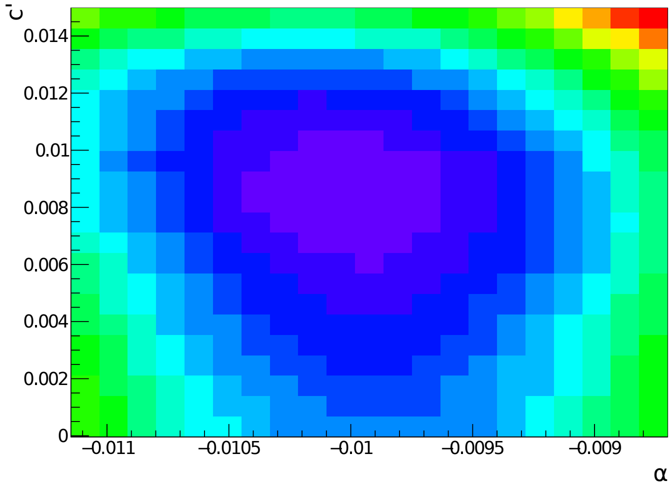


FIGURE 3.3: Distribution of χ^2 test between data and templates, as a function of the energy scale factor and the resolution factor.

- For a fixed value of $c'_{i,j}$, we look at the χ^2 distribution as a function of $\alpha_{i,j}$, as can be seen as a line in Figure 3.3, and the resulting χ^2 is fitted using a parabolic shape function:

$$\chi^2\left(\alpha_{ij}, c'_{ij}\right) = a_0\left(c'_{ij}\right) + \frac{\left(\alpha_{ij} - \alpha_{ij,\text{min}}\left(c'_{ij}\right)\right)^2}{\left(\delta\alpha_{ij}\left(c'_{ij}\right)\right)^2} \tag{3.13}$$

  where $\delta\alpha_{ij}\left(c'_{ij}\right)$ is the uncertainty on $\alpha_{ij,\text{min}}\left(c'_{ij}\right)$ determined by $\Delta\chi^2 = 1$ around the minimum.

- All the c'_{ij} lines of Figure 3.3 are scanned, and the $\chi^2_{\text{min}}\left(c'_{ij}\right)$ is plotted as a function of c'_{ij} and fitted using a 3$^{\text{rd}}$ polynomial function characterised with the parameters (b_0, b_1, b_2):

$$\chi^2_{\text{min}}\left(c'_{ij}\right) = b_0 + \frac{\left(c'_{ij} - \hat{c}'_{ij}\right)^2}{b_2^2} + b_1\frac{\left(c'_{ij} - \hat{c}'_{ij}\right)^3}{b_2^3}. \tag{3.14}$$

  The minimum of this distribution $\hat{c}'_{ij}$ is the most probable value (MPV) for the additional constant term in the configuration (η_1, η_2). The uncertainty on $\hat{c}'_{ij}$ is determined by $\Delta\chi^2 = 1$.

- Finally, $\alpha_{ij,\text{min}}\left(\hat{c}'_{ij}\right)$ is plotted as a function of $\hat{c}'_{ij}$ and a linear fit is performed around $\hat{c}'_{ij}$. The most probable value $\hat{\alpha}'_{ij}$ is defined as the value corresponding to $\hat{c}'_{ij}$.

3.4.2 Inversion procedure

As the values of α_{ij} and $c'_{i,j}$ are computed for each configuration (η_1, η_2) as described above, the correction factors α_i and c'_i must then be computed. For the

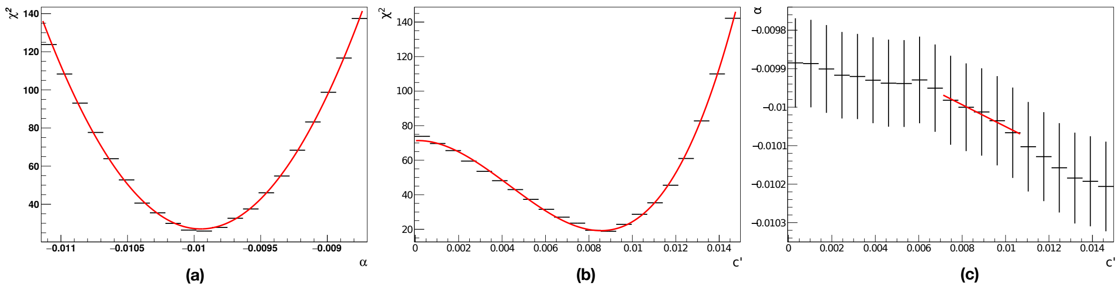


FIGURE 3.4: (a): χ^2 as a function of $\alpha_{i,j}$ for a given value of c'_{ij}. (b): $\chi^2_{\min}\left(c'_{ij}\right)$ as a function of c'_{ij}. (c): $\alpha_{ij,\min}\left(c'_{ij}\right)$ as a function of c'_{ij}.

energy scale factor, and because of the linear equation (3.7), α_i can be computed by the minimisation of a χ^2_α described as:

$$\chi^2_\alpha = \sum_{i,j \leq i} \frac{(\alpha_i + \alpha_j - 2\alpha_{ij})^2}{(\Delta\alpha_{ij})^2}. \tag{3.15}$$

On the other hand, the extraction of the additional constant term c'_i is more complicated because of the non-linearity of equation (3.10) describing the relation between $c'_{i,j}$ and (c'_i, c'_j). The extraction of the constant c'_i is based on the likelihood minimisation [83, 108] using the formula:

$$\chi^2_{c'} = \sum_{i,j \leq i} \frac{\left(\sqrt{\frac{c'^2_i + c'^2_j}{2}} - c'_{ij}\right)^2}{\left(\delta c'_{ij}\right)^2}. \tag{3.16}$$

3.5 Selections and corrections

The results presented in this thesis are based on data collected during Run 2 with the ATLAS detector, corresponding to an integrated luminosity of 139 fb^{-1} collected in 2015, 2016, 2017 and 2018. The data samples are detailed in [156]. To select $Z \to ee$ events, electrons candidates must pass the triggers shown in Table 3.1.

TABLE 3.1: Triggers used for data collected during Run 2.

Year	Trigger
2015	HLT_2e12_lhloose_L12EM10VH
2016	HLT_2e17_lhvloose_nod0
2017	HLT_2e24_lhvloose_nod0
2018	HLT_2e24_lhvloose_nod0 || HLT_2e17_lhvloose_nod0_L12EM15VHI

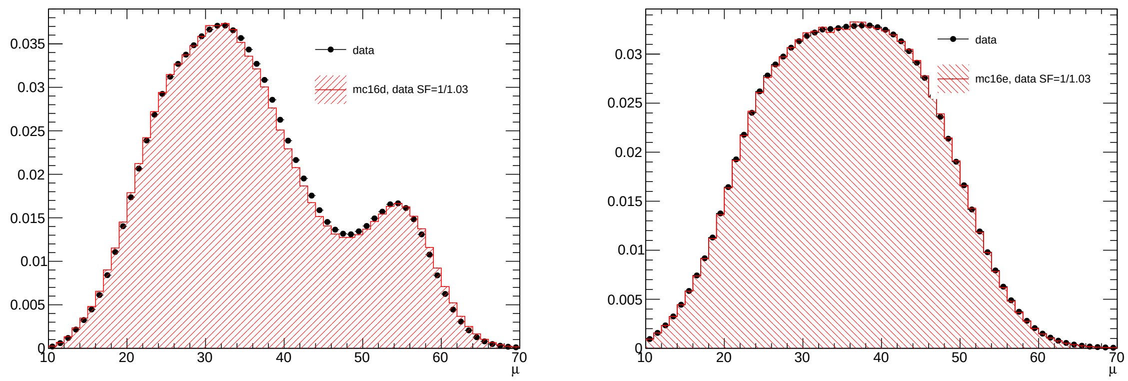
In addition, electron events must pass the MediumLH identification (ID) and loose isolation criteria in order to reduce mis-identified electrons and to suppress the QCD background [82]. Also, electrons are required to have $p_T^\ell > 27$ GeV and

$|\eta_{\rm track}| < 2.47$. Finally, events which pass all the selections mentioned above and with opposite charge are selected in the range $0 < m_{ee} < 180$ GeV. The number of selected events is shown in Table 3.2.

TABLE 3.2: Number of selected events which passes the selections used for the $Z \to ee$ analysis.

	2015	2016	2017	2018
Data	1.62 M	15.6 M	19.2 M	25.4 M
Simulation	6.53 M (MC16a)	18.5 M (MC16a)	20.2 M (MC16d)	28.8 M (MC16e)

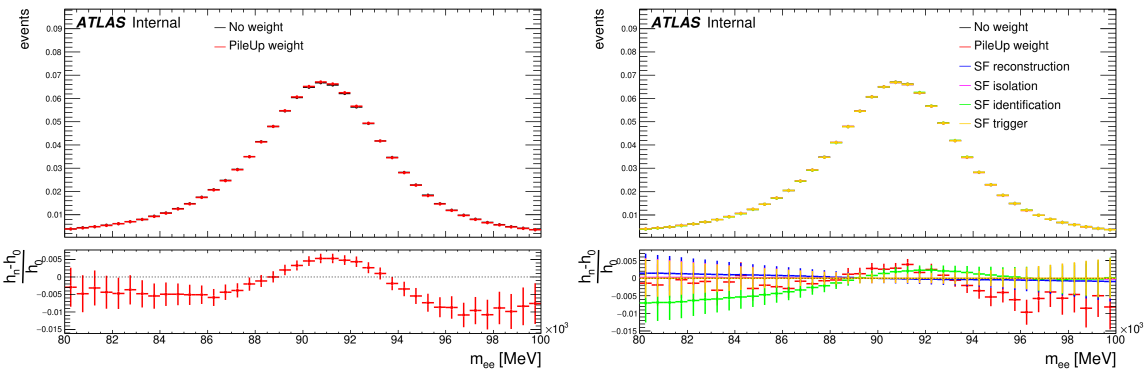
In Table 3.2, MC16a, MC16d and MC16e indicate the tag of simulation samples. Additional corrections in terms of weights to match the data need to be applied to the simulation. One of the corrections is the pile-up reweighting, used to reproduce the distribution of the number of pp collisions per bunch crossing in the data. Figure 3.5 shows an example of the actual number of interactions per bunch crossing in data compared to simulation after the pile-up reweighting.



(A) (B)

FIGURE 3.5: The actual number of interactions per bunch crossing of data which is rescaled by a factor 1/1.03 and simulation for 2017 (A) and 2018 (B), after the pile-up reweighting procedure.

Also, the difference between data and simulation for the reconstruction, identification, isolation and trigger efficiencies is taken into account by applying corresponding scale factors to the simulation. As shown in Figure 3.6, the changes in the invariant mass distribution of the MC before and after applying the pile-up reweighting correction and the different scale factors is typically small. In the current analysis, the electroweak background has been neglected. It is included in the systematic uncertainty (Table 3.4) and its contribution is smaller than 0.05% for invariant mass m_{ee} between 75 and 97 GeV.



(A) (B)

FIGURE 3.6: Effect of the pile-up reweighting (A) and different efficiency scale factors (B) corrections on the normalized $Z \to ee$ mass distribution in simulation. The bottom plot shows the fractional differences of the invariant mass distribution without any scale factors (labelled h_0) and with the application of different efficiency scale factors or reweightings (labelled h_n) separately.

3.5.1 Binning

During Run 1, the energy scale factor α was extracted in 34 bins along η. For Run 2 and because of the high statistics of the data collected, the energy scale factors α are extracted using 68 bins, which correspond to Run 1 binning splitted by two. The small binning allows a better correction of data. For the additional constant term, the Run 1 binning is kept in order to maximise the statistics in each configuration. Table 3.3 shows the new binning used for α and c' in the barrel and end-cap regions.

TABLE 3.3: Absolute values of η_{calo} bin boundaries for energy scale factors α and resolution constant terms c' used in the calibration of electromagnetic calorimeter during Run 2.

	Barrel														
α_i	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	1.1	1.2	1.285	1.37
c'_i	0		0.2		0.4		0.6		0.8		1		1.2		1.37

	End-cap																
α_i	1.55	1.59	1.63	1.6775	1.725	1.7625	1.8	1.9	2	2.05	2.1	2.2	2.3	2.35	2.4	2.435	2.47
c'_i	1.55						1.8		2				2.3				2.47

3.6 Results

3.6.1 Extraction of the correction factors (α, c')

The results of the energy scale factors α for Run 2 data sets are presented in Figure 3.7. The results are extracted separately for each year to take into account the data taking conditions. The observed differences (up to ± 0.005) in the end-cap region between the different years are related to two effects [2]:

Change of luminosity: at high luminosity, a larger current I is induced on HV lines due to a larger amount of energy deposited in the liquid argon gaps. The HV in the detector is reduced by $R \times I$, where R is the resistance between the power supply where the voltage is set to a constant value and the LAr gap. This effect is called high voltage drop and is dominated by luminosity effects.

Change of LAr temperature: this effect is related to the change of liquid argon temperature between the different data taking periods. Studies [97, 31] show that the energy response change by $-2\%/K^0$.

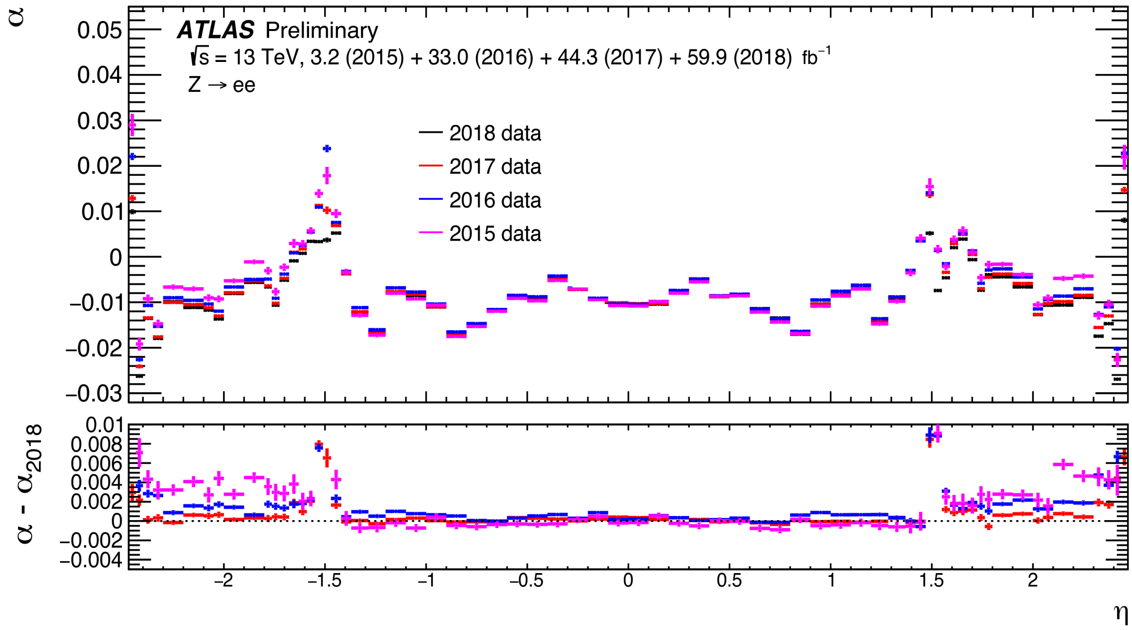


FIGURE 3.7: Energy scale factors extracted for 2015, 2016, 2017 and 2018 data taking during Run 2 as a function of η_{calo}. The bottom panel shows the differences between 2015, 2016 and 2017 to the 2018 data measurements.

For the additional constant term c', the results are shown in the same way for different years of Run 2 in Figure 3.8. Ideally, the additional constant term is independent of luminosity, but, as shown in Figure 3.8, the constant c' decreases as a function of the year. Studies [15] show that this effect is related to mis-modelling of the pile-up noise in simulation: the pile-up noise in the calorimeter increases with $\langle\mu\rangle$ in data and this effect is not well modelled in simulation, and therefore is absorbed by the additional constant c'.

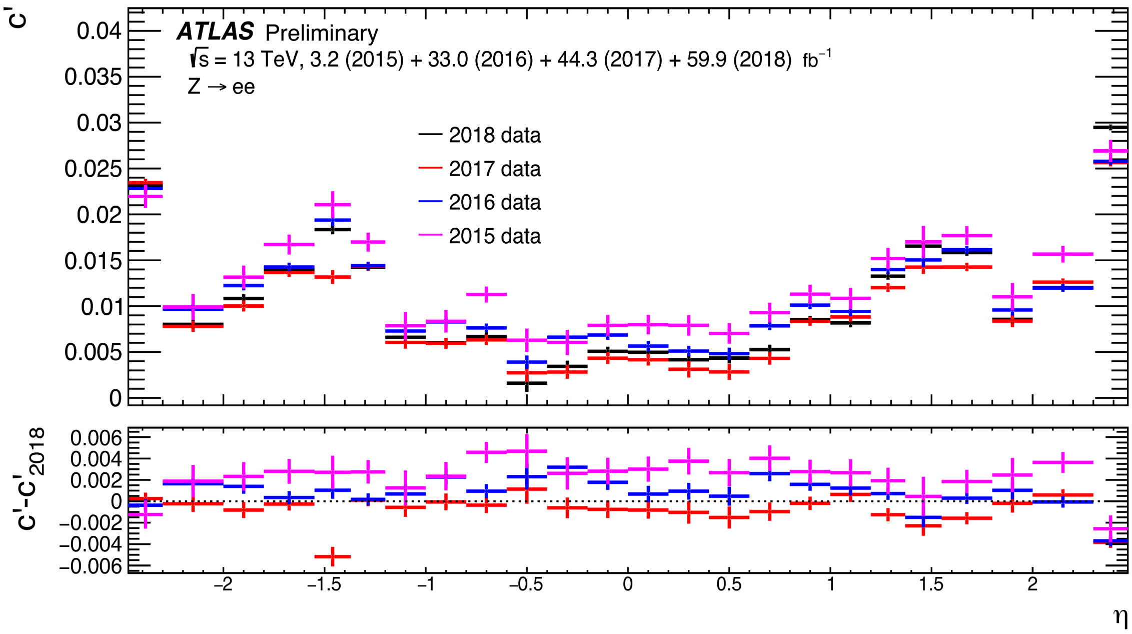


FIGURE 3.8: Additional constant term c' extracted for 2015, 2016, 2017 and 2018 data taking during Run 2 as a function of η_{calo}. The bottom panel shows the differences between 2015, 2016 and 2017 to the 2018 data measurements.

This effect is due to the fact that , in order to simulate the charged distribution and the calorimeter distribution of the data , different tunings of the pile-up reweighting are needed [99]. This is also dependent on the beam crossing configuration at LHC [153] as seen for instance in [121] . The official ATLAS pile-up reweighting correction factor of 1.03 (see figure 3.5) can be changed to 1.2 or 1.3, depending of the beam crossing configuration and the additional constant terms of different years are much more similar [131].

3.6.2 Systematic uncertainties

Different sources of systematic uncertainties of the correction factors (α and c') are evaluated using 64 bins for the energy scale factor α and 24 bins for the additional constant term c' described in Table 3.3, then symmetrised in bins of η_{calo} to reduce the statistical fluctuations. The systematic sources can be summarised in:

Mass window: the energy scale factors depend on the invariant mass window of the fit, due to the fact that the distribution tails are not well modelled in simulation [96]. The impact of the mass window is estimated by changing the window from [80, 100] to [87, 94.5] GeV, and the difference is taken as a systematic uncertainty.

Mass threshold: in the template method, we use only configurations with $m_{ee}^{\text{th}} >$ 70 GeV. This threshold mass is computed [82] based on the fact that the mass of the Z boson, when both electrons have the same E_{T} and are at opposite ϕ, is equal to $M_Z = E_{\text{T}}\sqrt{2 \cdot \cosh(\eta_j - \eta_i)}$. Since the selection requires electrons to have at least E_{T} = 27 GeV, the threshold mass for an (i,j) configuration is defined as $m_{ee}^{\text{th}} = 27 \cdot \sqrt{2 \cdot \cosh(\eta_j - \eta_i)}$. This choice is arbitrary and a systematic uncertainty, defined by comparing the scale factors using $m_{ee}^{\text{th}} >$ 70 and $m_{ee}^{\text{th}} >$ 77 GeV, is added to take into account the impact of the selection.

Background: in the template method, the electroweak background has been neglected. This systematic uncertainty is computed by comparing the scale factors with and without the background.

Electron reconstruction efficiencies: this uncertainty is added to take into account the scale factors (reconstruction, isolation, identification and trigger) applied to MC in order to match data. These efficiency factors are characterised by uncertainties propagated by the template method and considered as systematic uncertainties in α and c'.

Electron reconstruction quality: as shown in Sec. 3.5, electrons must pass medium ID requirement. A systematic uncertainty is evaluated by comparing medium and tight ID electrons. In addition, another systematic uncertainty is added to take into account the uncertainty on the emission of photon by bremsstrahlung in the calorimeter.

Method comparison: in addition to the template method used on this thesis, there is another method called the "lineshape method" [74]. A systematic uncertainty is defined by the difference between the two methods.

Method accuracy: this uncertainty is used to take into account the intrinsic bias of the template method. It is evaluated by injecting known values in a MC sample and try to measure these values using the template method. The difference between the measured and injected values is defined as the systematic uncertainty.

The systematic uncertainties in the scale factors are listed in Table 3.4. The total uncertainty is calculated by the quadratic sum of all the effects described above.

TABLE 3.4: Ranges of systematic uncertainties in α and c' for different η ranges [74].

	Uncertainty in $\alpha_i \times 10^3$			Uncertainty in $c_i \times 10^3$		
$\lvert\eta\rvert$ range	$0-1.2$	$1.2-1.8$	$1.8-2.4$	$0-1.2$	$1.2-1.8$	$1.8-2.4$
Uncertainty source						
Method accuracy	$(0.01-0.04)$	$(0.04-0.10)$	$(0.02-0.08)$	$(0.1-0.7)$	$(0.2-0.4)$	$(0.1-0.2)$
Method comparison	$(0.1-0.3)$	$(0.3-1.2)$	$(0.1-0.4)$	$(0.1-0.5)$	$(0.7-2.0)$	$(0.2-0.5)$
Mass range	$(0.1-0.5)$	$(0.2-4.0)$	$(0.2-1.0)$	$(0.2-0.8)$	$(1.0-3.5)$	1.0
Region selection	$(0.02-0.08)$	$(0.02-0.2)$	$(0.02-0.2)$	$(0-0.1)$	0.1	$(0.2-1.0)$
Bkg. with prompt electrons	$(0-0.05)$	$(0-0.1)$	$(0-0.5)$	$(0.1-0.4)$	0.2	$(0.1-0.2)$
Electron isolation requirement	$(0-0.02)$	$(0.02-5.0)$	$(0.02-0.20)$	$(0.1-0.9)$	$(0.1-1.5)$	$(0.5-1.5)$
Electron identification criteria	$(0-0.30)$	$(0.20-2.0)$	$(0.20-0.70)$	$(0-0.5)$	0.3	0.0
Electron bremsstrahlung removal	$(0-0.30)$	$(0.05-0.7)$	$(0.20-1.0)$	$(0.2-0.3)$	$(0.1-0.8)$	$(0.2-1.0)$
Electron efficiency corrections	0.10	$(0.1-5.0)$	$(0.10-0.20)$	$(0-0.3)$	$(0.1-3.0)$	$(0.1-0.2)$
Total uncertainty	$(0.2-0.7)$	$(0.5-10)$	$(0.6-2.0)$	$(0.3-1.2)$	$(1.0-6.0)$	$(2.0-3.0)$

3.6.3 Data to simulation comparison

After deriving the energy correction factors, they are applied to data and MC events and the final distributions are compared in Figure 3.9. The energy scale factors α are applied to data in order to match the energy response of the simulation and MC events are smeared according to c' factors in order to match the slightly worse resolution in data. The lower panel shows the data to simulation ratio and the total systematic uncertainty on the energy scale and resolution corrections from [36]. The deviations are largest in the tails where they can reach 3% and are mostly covered by uncertainties.

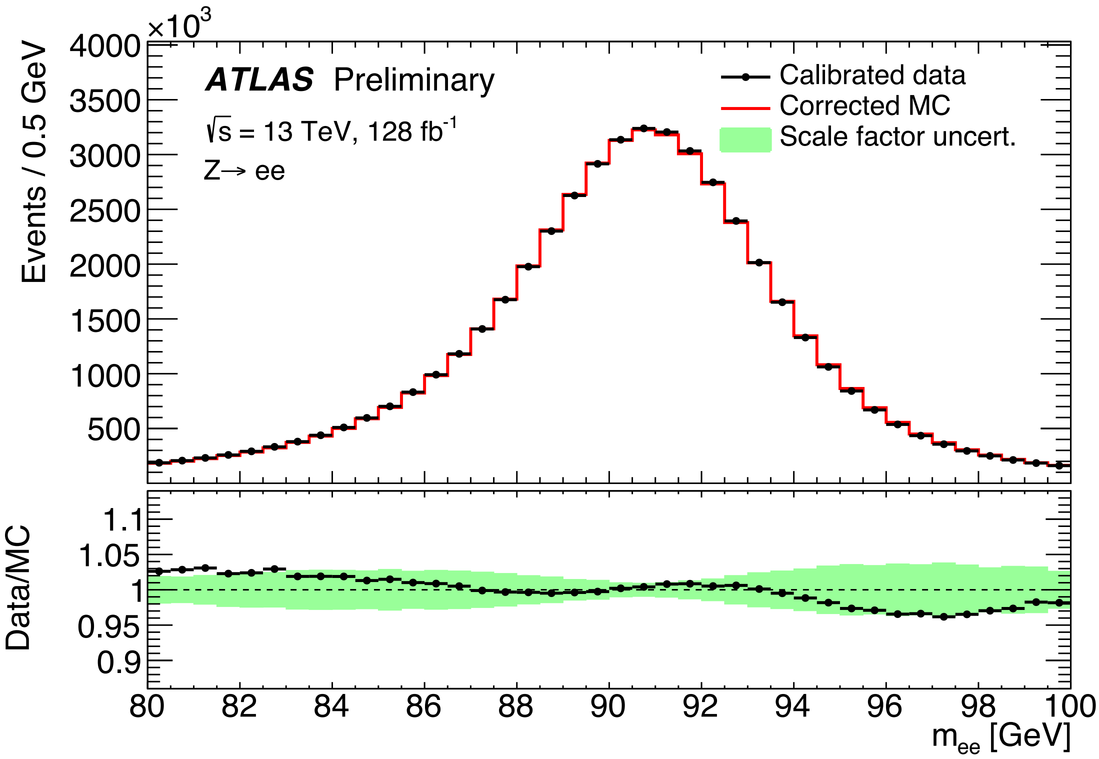


FIGURE 3.9: Inclusive di-electron invariant mass distribution from $Z \to ee$ decays in data compared to MC after applying the full calibration. The simulation is normalized to data. The lower panel shows the data to simulation ratio, together with the uncertainty from the energy scale and resolution corrections.

The systematic uncertainties (see Table 3.4) are dominated by the electron identification, the method comparison, the mass range and the electron Bremsstrahlung removal. However, for the W-mass measurement [115] some improvements [41] were achieved with respect to the Run 1 calibration [67]: in particular restricting the η range by excluding $1.2 < |\eta| < 1.8$, using broader η bins in order to compute the systematic uncertainties and neglecting some uncertainties when we apply the calibration to electrons, like in $W \to e\nu$, for instance the uncertainty related to the electron ID as well as the uncertainty related to Bremsstrahlung emission since the analysis is inclusive in Bremsstrahlung.

3.7 Calibration for low pile-up runs

3.7.1 Introduction

In addition to the data collected for the nominal Run 2 analyses, called nominal runs, there are other runs dedicated to special studies. For the measurement of the W boson mass, we use low pile-up runs characterised with low number of interactions per crossing ($\langle\mu\rangle \approx 2$), as shown in Figure 3.10. These data sets were collected by ATLAS in autumn 2017 (258 pb^{-1} at $\sqrt{s} = 5$ TeV and 148 pb^{-1} at $\sqrt{s} = 13$ TeV) and in summer 2018 (an additional 193 pb^{-1} at $\sqrt{s} = 13$ TeV). The low pile-up samples are detailed in [95]. For low pile-up runs, we use the the same selections as for the nominal runs described in Sec. 3.5, except for the trigger where we use *HLT_e15_lhloose_nod0_L1EM12* for 2017 and 2018 data. The number of selected events for low pile-up runs is shown in Table 3.5.

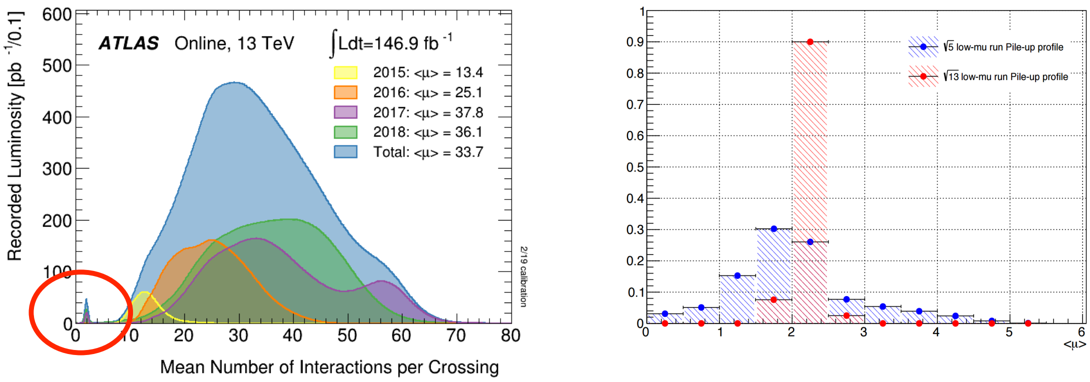


FIGURE 3.10: **left**: low pile-up runs at $\langle\mu\rangle \approx 2$ showed in the red circle. **right**: the pile-up distribution of simulated low pile-up data at $\sqrt{s} = 5$ and 13 TeV.

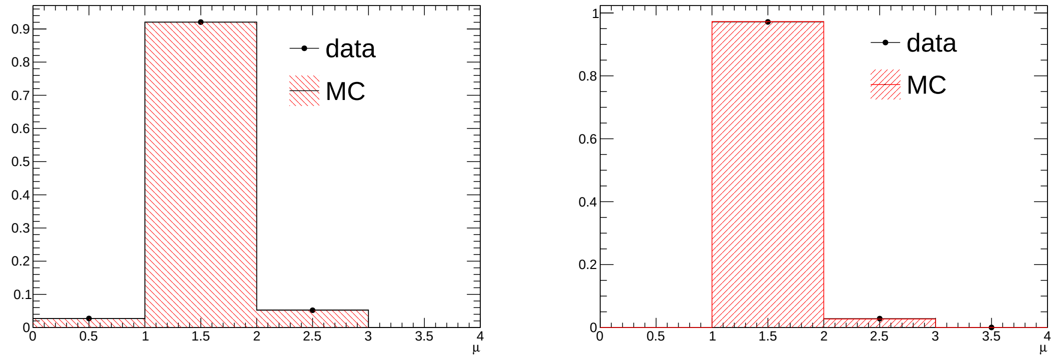
TABLE 3.5: The number of $Z \to ee$ candidate events selected after applying all the selections for low pile-up runs at $\sqrt{s} = 5$ and 13 TeV.

	5 TeV(2017)	13 TeV(2017)	13 TeV(2018)
Data	58.7 k	79.9 k	107.2 k
Simulation	2.14 M	1.38 M	1.41 M

The correction factors related to the reconstruction, identification, isolation and trigger efficiencies are applied also for the low pile-up runs. All these correction factors are obtained from the low pile-up runs except for the reconstruction efficiencies. In the same way, the pile-up reweighting is applied to MC in order to reproduce the distribution of the number of pp collisions per bunch crossing in data. Figure 3.11 shows the distribution of $\langle\mu\rangle$ in data and MC.

3.7.2 Energy scale factors for low pile-up runs

For the low pile-up runs, the same procedure described above is used to derive α and c' correction factors to equalise the response and resolution of data and MC.



(A) (B)

FIGURE 3.11: The average number of interaction per bunch crossing of data and simulation with low pile-up $\sqrt{s}$ = 13 TeV runs for 2017 (A), 2018 (B) after the pile-up reweighting procedure.

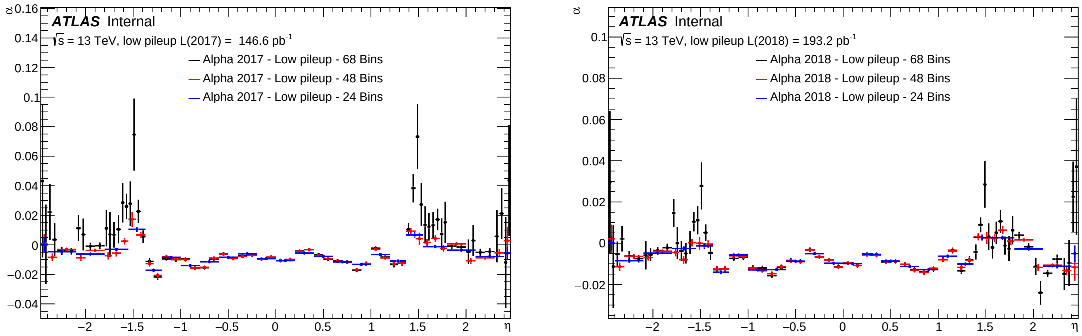
For the nominal high pile-up data, the energy scale factors corrections are derived in 68 η_{calo} bins. Because of the smaller number of $Z \rightarrow ee$ events in the low pile-up runs, the scale factors extracted with the same 68 bins result have large statistical fluctuation and systematic bias, especially in the end-cap region, as shown in the Figure 3.12. To avoid this problem, two binnings were studied combining some bins of the 68 η_{calo} bins:

- either 48 bins in total with bins of larger size only in the end-cap
- or 24 bins in total with bins of larger size in both the barrel and the end-cap regions, as shown in Figure 3.12 and Table 3.6.

TABLE 3.6: Values of η_{calo} bin boundaries for energy scale α for 24 bins.

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

The results obtained from these binnings are shown in Figure 3.12. As can be seen, the instability for the end-cap bins disappears. As the α factors do not vary strongly between the 48 and 24 bins versions, the baseline chosen is 24 bins. The additional constant term c_i' applied to MC to account for the worse resolution in data is shown in Figure 3.13. As the c_i' values were previously observed to be dependent on the pile-up and data taking conditions, it is best to extract and use the constants from the respective data set under study. This is further discussed in Sec. 3 of Ref. [15]. The physics analyses currently use directly the in-situ calibrations as derived in this section. The main uncertainties are given by the statistical uncertainties of the α_i and c_i' factors. As these are significantly larger than other uncertainties, another approach is used to extract the energy scale factor α_i and explained in Sec. 3.7.3.



(A) (B)

FIGURE 3.12: Energy scale factors α for low pile-up runs of 2017 (A) and 2018 (B) using 68, 48 and 24 η bins. It can be seen, that the extraction is unstable in case of 68 bins, resulting in α factors with very large uncertainties.

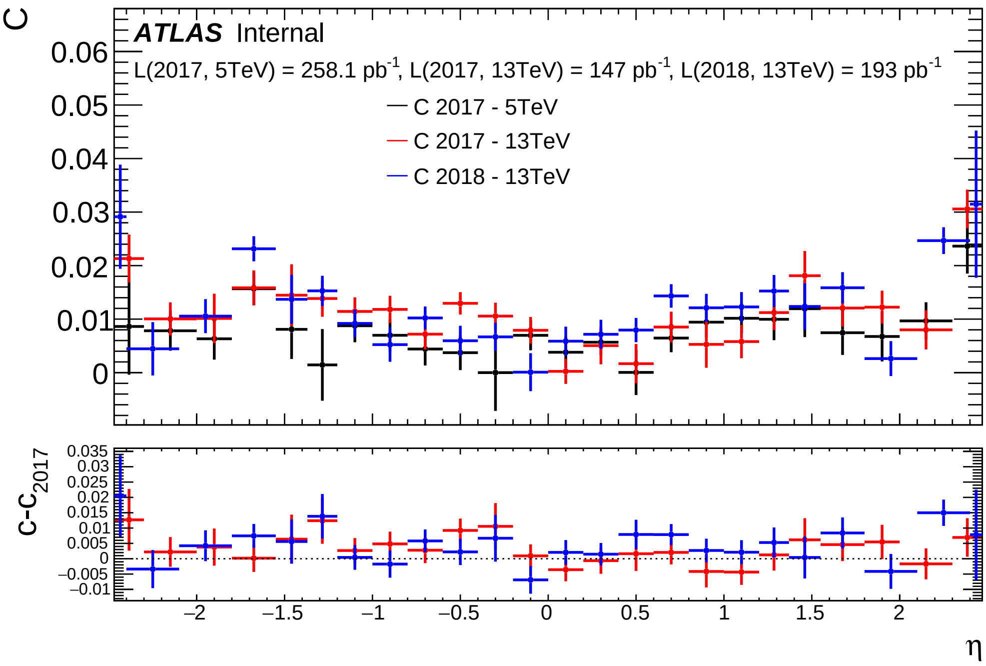


FIGURE 3.13: Additional constant term c_i' for low pile-up runs of 2017 (13 TeV), 2018 (13 TeV) and 2017 (5 TeV) using 24 bins. The lower panel shows the difference of c_i' to 2017 (5 TeV) results.

3.7.3 Extrapolation method

An alternative method to derive the energy correction scale factors for the low pile-up data is to study the dependence of the factor α for high pile-up data sets and to perform an extrapolation to $\langle\mu\rangle \approx 2$. This method exploits the large sample of the high pile-up data, but requires additional work to ensure the extrapolation is under control. The extrapolation proceeds by separating the high pile-up data into intervals of $\langle\mu\rangle$ and applying the template method to extract the energy scale factors as a function of $\langle\mu\rangle$, i.e. $\alpha(\langle\mu\rangle)$. Using a (linear) fit $\alpha(\langle\mu\rangle)$ can be extrapolated to $\langle\mu\rangle \approx 2$. The $\langle\mu\rangle$ intervals are defined in Table 3.7. The extrapolation for two example η bins is shown in Figure 3.14. Figures 3.15 and 3.16 show the extrapolation from high pile-up to low pile-up data for all η bins comparing the negative and positive η bins in a same plot. The asymmetric effect observed between the negative and positive η bins could be due mainly to temperature effects which may not be symmetric. Over the $\langle\mu\rangle$ range samples in the high pile-up data a linear fit is found to be sufficient. In many bins the slope of $\alpha(\langle\mu\rangle)$ is found to be small, but in particularly in the end-cap region the slopes are often significant.

TABLE 3.7: The μ intervals used for extrapolation study.

[0 : 26]	[26 : 33]	[33 : 40]	[40 : 50]	[50 : 80]

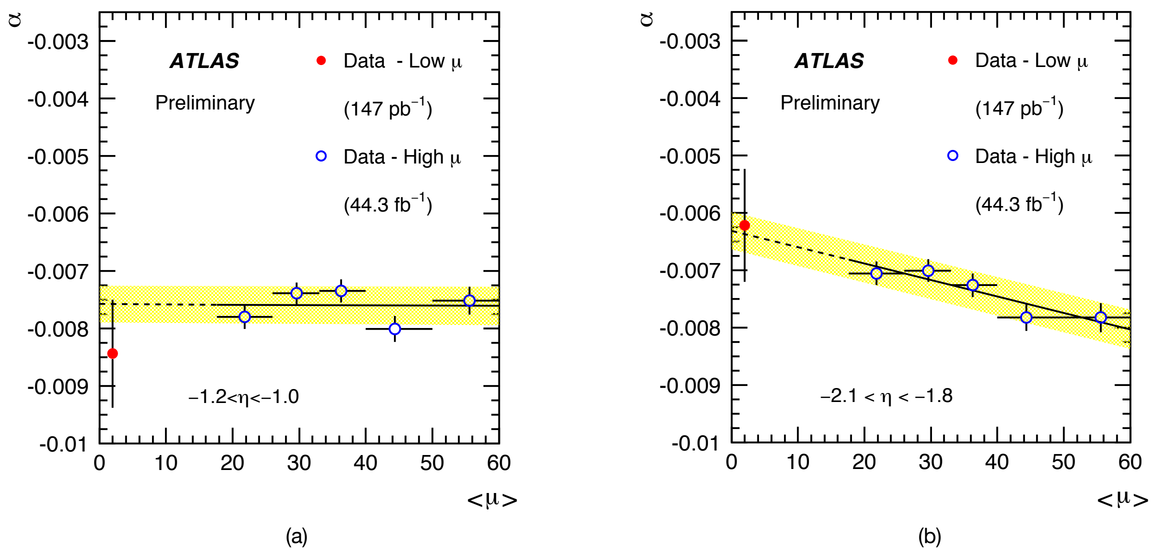


FIGURE 3.14: Examples of the energy scale extrapolation from high pile-up to low pile-up in the barrel (a) and end-cap (b). The blue points show the energy scale factors α for the high pile-up data set as a function of $\langle\mu\rangle$, the black lines show the extrapolation to $\langle\mu\rangle \approx 2$ using a linear function and 5 intervals of $\langle\mu\rangle$, the band represents the uncertainty in the extrapolation. The extrapolation results are compared to the energy scale factors α extracted from the low pile-up data set, represented by the red point.

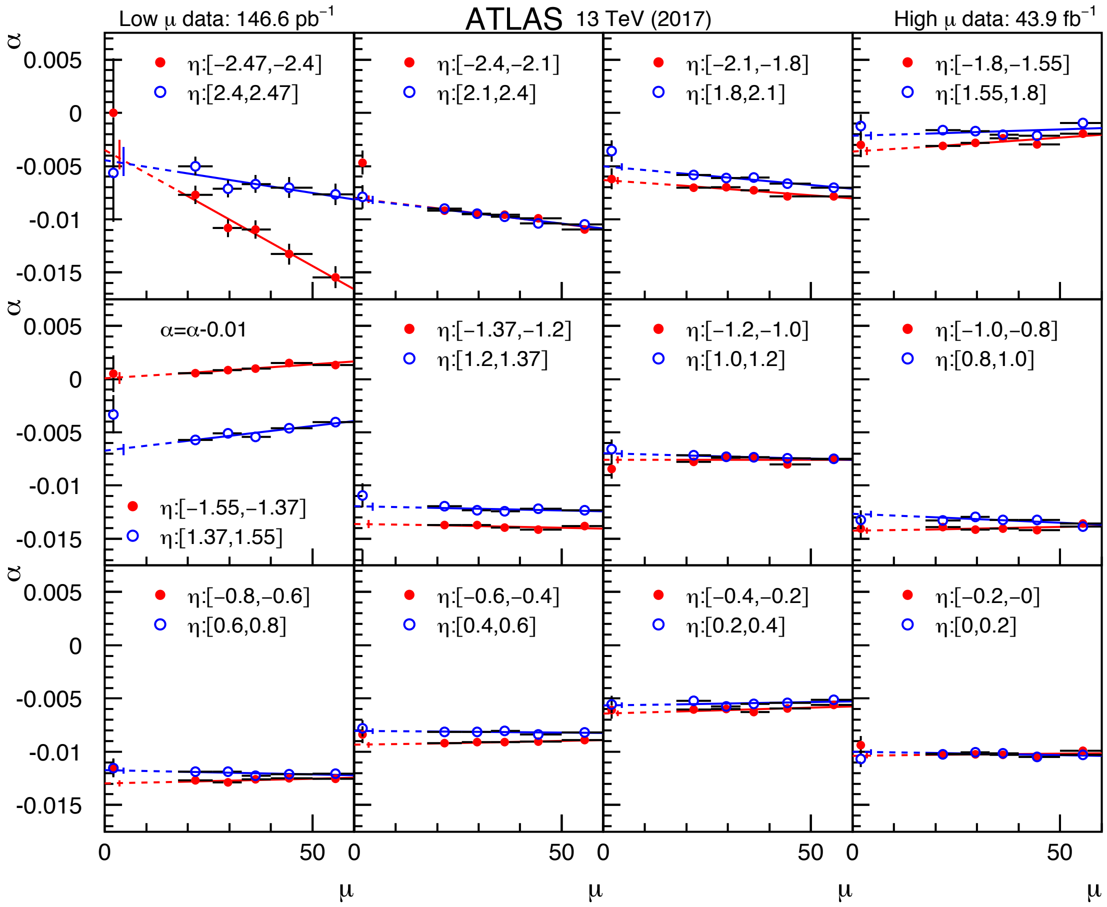


FIGURE 3.15: Energy scale extrapolation from 2017 high pile-up to low pile-up for 2017 (at 13 TeV) low pile-up data. The blue and red points show the energy scale factors α for the high pile-up data set as a function of $\langle\mu\rangle$ for different η regions, the dotted lines show the extrapolation to $\langle\mu\rangle \approx 2$ using a linear function and 5 intervals of $\langle\mu\rangle$. The values of α determined using the low $\langle\mu\rangle$ data sets are also shown at $\langle\mu\rangle \approx 2$. The size of the vertical lines near the low α points represents the uncertainty of the extrapolation.

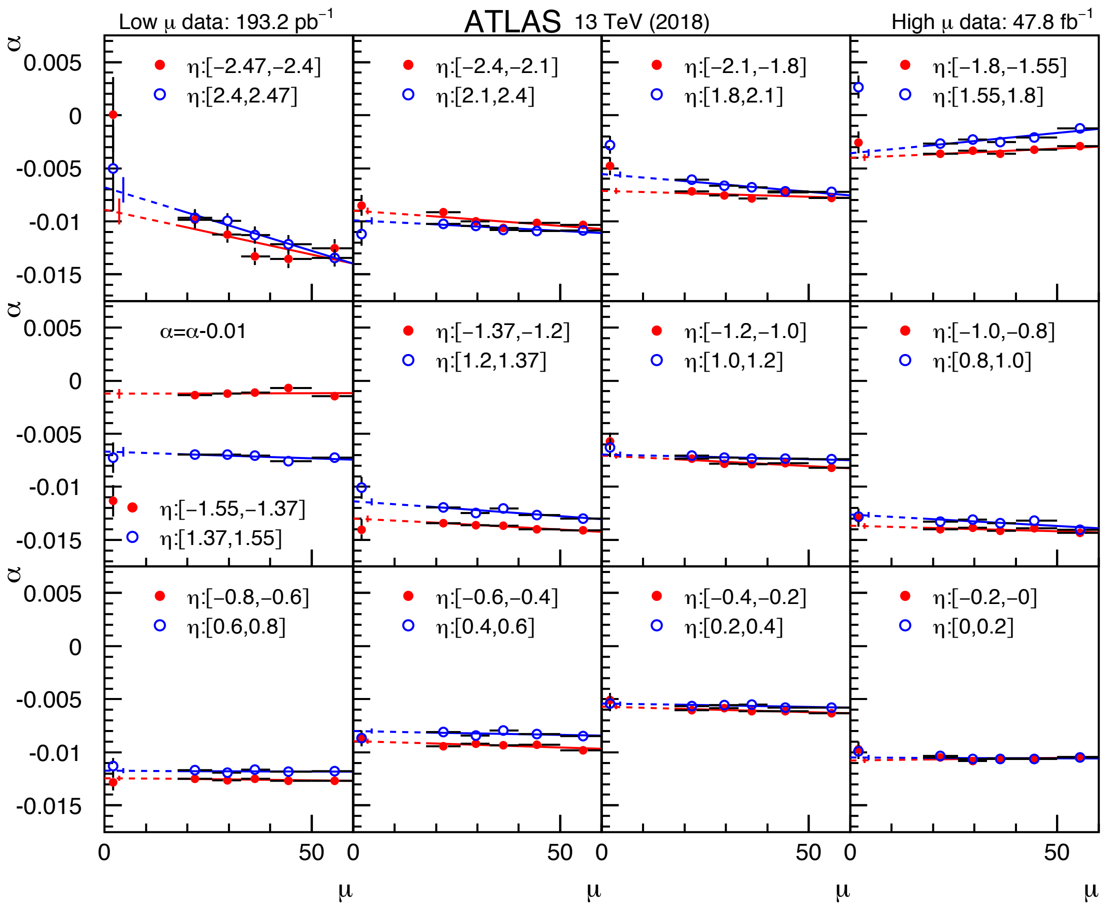
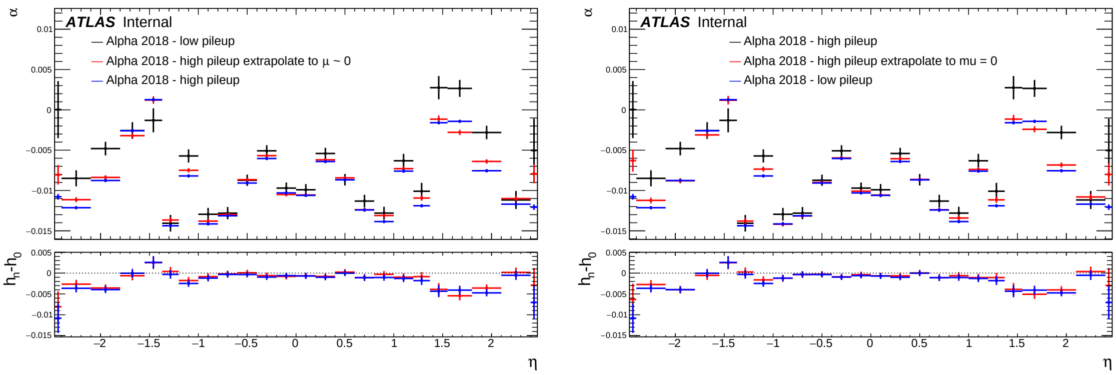


FIGURE 3.16: Energy scale extrapolation from 2018 high pile-up to low pile-up for 2018 (13 TeV) low pile-up data. The blue and red points show the energy scale factors α for the high pile-up data set as a function of $\langle\mu\rangle$ for different η regions, the dotted lines show the extrapolation to $\langle\mu\rangle \approx 2$ using a linear function and 5 intervals of $\langle\mu\rangle$. The values of α determined using the low $\langle\mu\rangle$ data sets are also shown at $\langle\mu\rangle \approx 2$. The size of the vertical lines near the low α points represents the uncertainty of the extrapolation.

3.7.4 Extrapolation results

After extrapolating the results to $\langle \mu \rangle \approx 2$, it could be expected that the energy scale factors coincide with those extracted directly using the low pile-up data within uncertainties. However, this is not always the case as it is shown in Figure 3.17, where it is observed that the extrapolation results are in fact closer to the high pile-up results without extrapolation than to the low pile-up results.



(A) (B)

FIGURE 3.17: The energy scale factors α for 2017 (A) and 2018 (B) data, respectively. The results are shown using directly the low (black) and high (blue) pile-up data and extrapolating the high pile-up results to $\langle \mu \rangle \approx 2$ (red). The bottom panels show the absolute differences between the high-pile-up α factors with and without extrapolation correction (labelled h_n) to the in-situ low pile-up derived α factors (labelled h_0).

This behavior of the extrapolated results was understood to be due to the different settings of the topo-cluster noise thresholds at reconstruction level: for the low pile-up data these were set to correspond to $\langle \mu \rangle = 0$ (to improve the hadronic recoil reconstruction), while the nominal high pile-up data is reconstructed with a threshold corresponding to $\langle \mu \rangle = 40$. The lower noise threshold used for the low pile-up data leads to more cells added to the topo-clusters and thus to a higher energy as shown in Figure 3.18. The effect of different noise thresholds on the energy scale factors α is studied with a dedicated processing of the data and MC (as described in Sec. 3.3) where the noise thresholds are set to the nominal high pile-up values.

Using the template method, the energy scale factors for the low pile-up data are extracted separately for low and high noise thresholds and compared in Figure 3.19 (A). As an alternative method, the difference of the average energy response $(E^{\text{low-threshold}} - E^{\text{high-threshold}})/E^{\text{low-threshold}}$ electron-by-electron reconstructed with low and high noise thresholds can be compared between data and simulation, as shown in Figure 3.19 (B). This second method is chosen because it reduces the statistical fluctuations. After correcting the threshold effect by applying the correction from Figure 3.19 (B), the extrapolation results in 24 bins of η are closer to the low pile-up results extracted directly with the template method, as shown in Figure 3.20. As can be seen from this figure, the difference between the extrapolated and low pile-up results is of the order of 0.1% (absolute difference in α of 0.001) in the barrel region, but increases to 0.5% (absolute difference in α of 0.005)

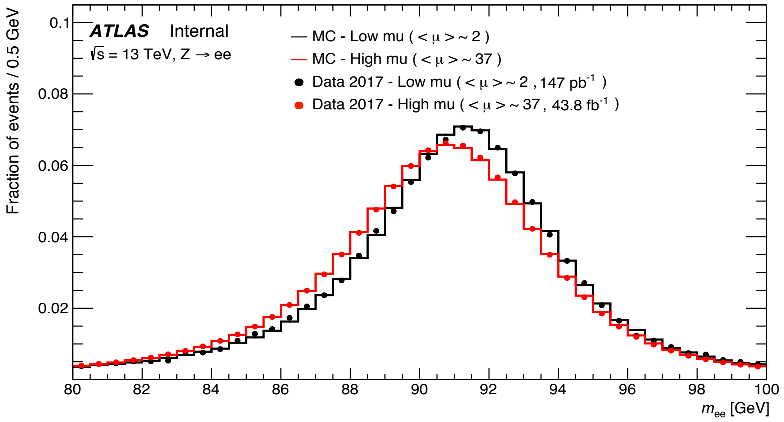
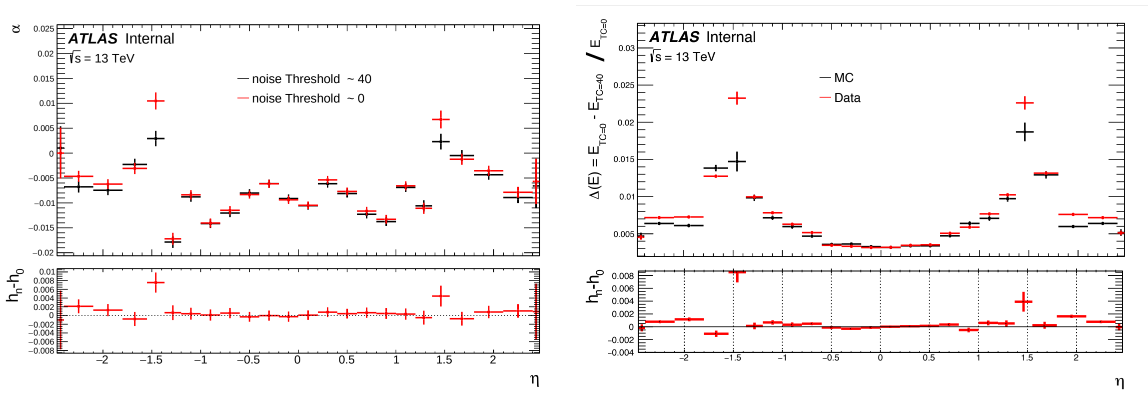


FIGURE 3.18: Comparison of the di-electron invariant mass distribution m_{ee} for data and simulation between high and low pile-up runs.

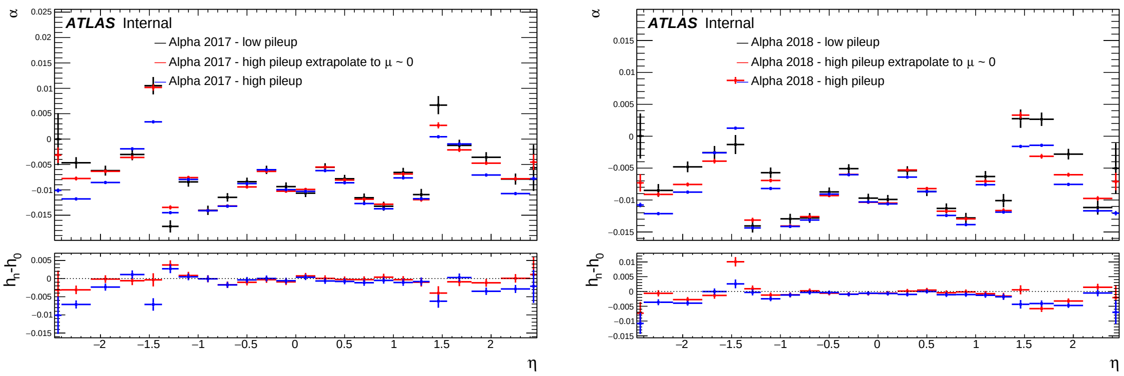


(A) Template α extraction

(B) Electron-by-electron energy response

FIGURE 3.19: (A): The effect of the noise threshold corresponding to $\mu = 0$ (red) or $\mu = 40$ (black) on the energy scale factors α using the template extraction. The bottom panel shows the absolute differences of α-factors obtained with high pile-up (labelled h_n) to those obtained with the low pile-up (labelled h_0) topo-cluster thresholds. (B): The difference in the energy response from the noise threshold settings extracted electron-by-electron on MC (black) and data (red). The bottom panel shows the absolute differences between data (labelled h_n) to MC (labelled h_0).

in the end-cap region (excluding the "crack" region). The additional constant term c' in any case will be taken from the direct results from the template method using low pile-up samples without extrapolation from high pile-up data, as the calibration uncertainties are dominated by the scale factors α_i corrections.



(A) (B)

FIGURE 3.20: The extrapolation results for the energy correction factors α before (blue) and after (red) correcting the effects of the difference of the noise threshold for 2017 (A) and 2018 (B). The results are compared to the in-situ low pile-up results (black). The bottom panels show the absolute differences between the extrapolated results (labelled h_n) to the in-situ results (labelled h_0).

3.7.5 Uncertainties for the extrapolation method

As the high pile-up results are used in the extrapolation procedure, the systematic uncertainties of high pile-up samples evaluated in Sec.3.6.2 are relevant also at low pile-up. In addition to high pile-up systematic uncertainties, there are other uncertainties mainly related to the difference between high and low pile-up runs and to the extrapolation procedure:

Threshold correction: for low pile-up data set a different topo-cluster noise threshold for the energy reconstruction is used, and a systematic uncertainty is evaluated to take into account this difference. This systematic uncertainty is defined as the statistical error on the difference of threshold, shown in the bottom plot panel in Figure 3.19 (B).

Extrapolation systematic uncertainties: The extrapolation uncertainty is considered as the quadratic sum of the following two effects:

1. The choice of the polynomial functions used in the extrapolation: the baseline extrapolation is performed with a polynomial of order 1. The difference between using a first or a second order polynomial function is included as discussed in [15].
2. The number of $\langle\mu\rangle$ intervals used in the extrapolation: for the baseline extrapolation, we used five intervals in $\langle\mu\rangle$. The effect of using three intervals is considered [15].

Temperature uncertainty: for nominal runs, there is a systematic uncertainty which includes the changes of LAr calorimeter response with temperature, but this effect is not linear with μ. Indeed, since it takes some time (few hours) for the liquid argon calorimeter to heat, there is a rough delay between the increase or decrease of luminosity and the corresponding increase or decrease of temperature. This introduces [99] a systematic uncertainty of 0.03% in the barrel and 0.1% in the end-cap region for low pile-up runs.

Figure 3.21 shows an overview of all the sources on the energy scale factor α for the 2017 low pile-up run at $\sqrt{s} = 13$ TeV while using the extrapolation method.

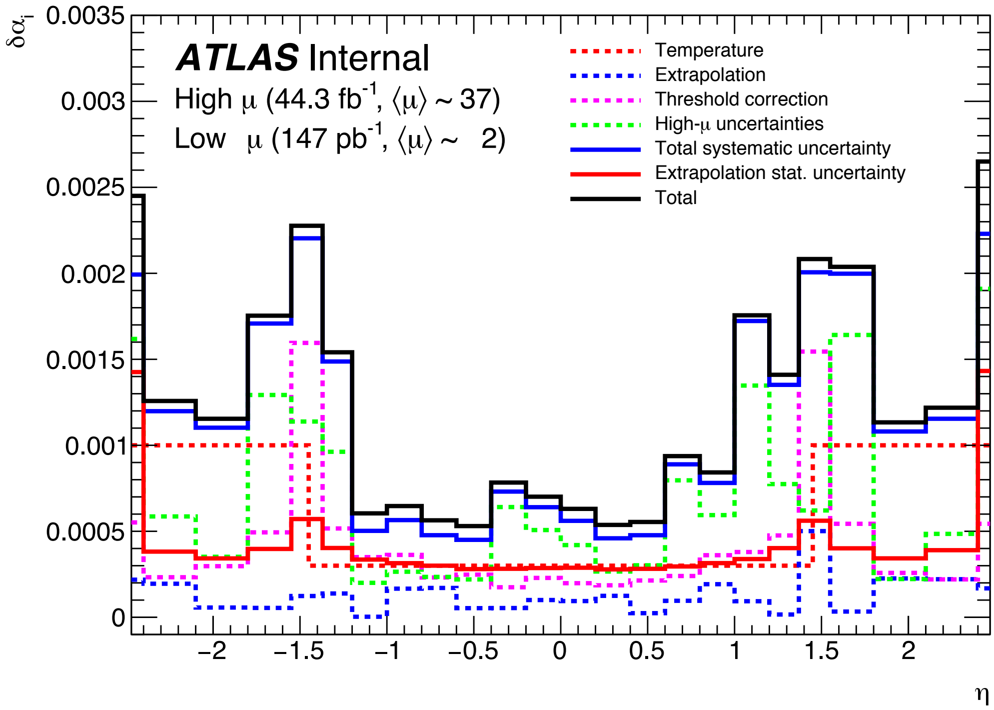


FIGURE 3.21: Uncertainties on the energy scale corrections as a function of η for the 2017 low pile-up runs at $\sqrt{s} = 13$ TeV.

3.7.6 Data to simulation comparison for low pile-up runs

After having calculated the energy correction factor α, we apply them to data and MC events and the final distributions are compared in Figure 3.22. The lower panel shows the data to simulation ratio with the statistical uncertainty in the energy scale.

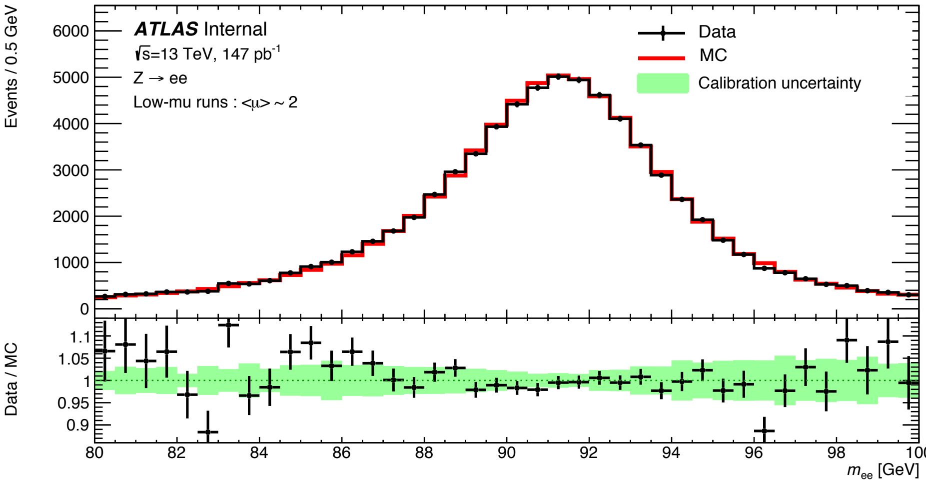


FIGURE 3.22: Inclusive di-electron invariant mass distribution for low pile-up runs from $Z \to ee$ decays in data compared to MC after applying the full calibration. The simulation is normalized to data. The lower panel shows the data to simulation ratio, together with the statistical uncertainty from the energy scale corrections.

3.8 Future of the calibration

A lot of efforts have been made on the e/γ calibration, in Run 1 and Run 2, however there remain several problems. In particular there a small mismodeling of the lineshape by the Monte Carlo as seen in Figure 3.9. Several ideas have been studied (or will be studied) to understand and solve this problem:

- It was noticed [41, 97, 115] that excluding the $1.2 < |\eta| < 1.8$ region gives a better agreement. This was confirmed and scrutinized with more in depth in recent analysis [131, 85].
- Non linearity [57, 85] checks have been performed using a method [38] similar to the template method described in Sec 3.4. The non linearity has been computed but the improvement in the mismodeing is marginal. Additional test of non linearity will be test using E/p as a measure.
- Additional non Gaussian tails could be at the origin of this effect. However simple tests using additional material in front of the calorimeter did not show any significant improvement of the mismodeling [85]. Following work on the forward calorimeter [44], a study as started [94] in order to study these non Gaussian tails in the EM calorimeter.

Chapter 4

Statistical overview: Unfolding

4.1 Introduction

In this chapter we will discuss the theoretical part of the unfolding problem [139], used in chapter 7, to calculate the fiducial and differential cross sections, and in chapter 5 for the measurement of the boson transverse momentum and in the chapter 8 for the measurement of W mass. The need for unfolding stems from the fact that any quantity measured at the LHC detectors is affected by the not completely well known detector effects (like acceptance and resolution). The goal of the unfolding is to correct data distributions and estimate the true physical distributions of the observables of interest without detector effects [40]. In high energy physics, several unfolding methods are used [49], and in our analysis, the iterative Bayesian unfolding [55] is used.

4.2 Unfolding in high energy physics

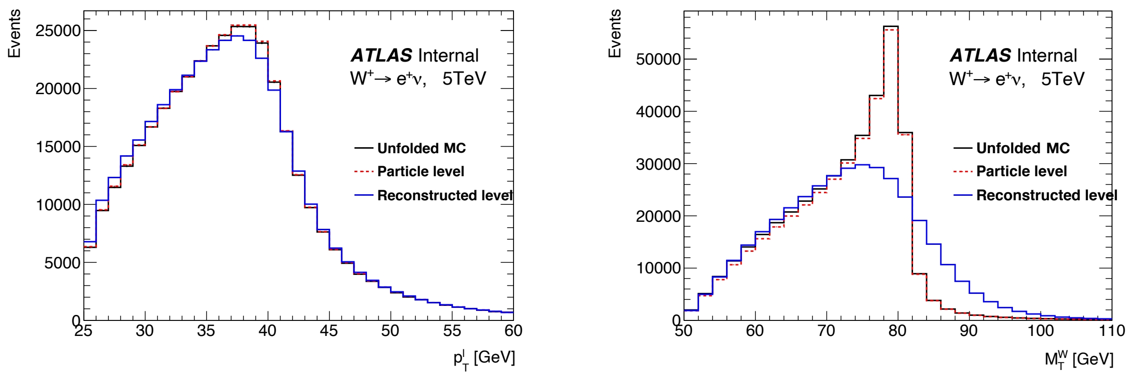
In high energy physics, we are interested in distributions of the observables of interest. In most of the cases, different distributions are affected by detector effects with different sizes. For example, the transverse mass of the W boson is more affected by detector effects than the transverse momentum of the lepton in $W \to \ell + \nu$. Figure 4.1 shows the comparison between simulated distributions without detector effects (particle level), with detector effects (reconstructed level) and after the unfolding for m_{T}^{W} and p_{T}^{ℓ}.

The reconstructed distributions are different from truth distributions because of two effects:

- Limited acceptance: it reflects the fact that not all events are observed by the detector, it is called the detector acceptance and it is smaller than 1 [120].
- Migration: due to limited detector resolution, an event originating from bin i can be measured in another bin j. This effect is taken into account with the migration matrix explained in Sec. 4.3.1.

For a mathematical presentation of the unfolding problem, let's consider that we have just MC simulation vector x (y) of dimension N_x (N_y), where the elements x_i (y_i) represent the number of events in bin i in our distribution at the truth (reconstructed) level. Both vectors x and y are related with a matrix R, called response matrix:

$$R \times x = y. \tag{4.1}$$



(A) (B)

FIGURE 4.1: p_{T}^{ℓ} (A) and m_{T}^{W} (B) distributions before and after detector effects, with the unfolded distributions.

The elements $R_{i,j}$ of the response matrix R represent the probability that an event generated in bin j is measured in bin i. The number of background events must be removed from the vector y. In a real case, the response matrix R is calculated from the migration matrix M, where the $M_{i,j}$ are estimated using information from MC simulation:

$$M_{i,j} = N_{i,j}^{\mathrm{rec}\wedge\mathrm{gen}}, \tag{4.2}$$

where $N_{i,j}^{\mathrm{rec}\wedge\mathrm{gen}}$ represents the number of event generated in truth bin j and reconstructed in bin i. If N_j^{gen} represents the number of event generated in truth bin j, the response matrix is then defined as:

$$R_{i,j} = \frac{M_{i,j}}{N_j^{\mathrm{gen}}}. \tag{4.3}$$

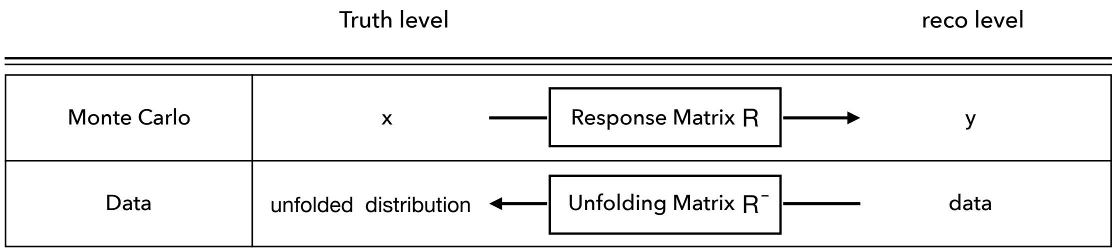
In our case, we are using a slightly modified response matrix $R'i,j$ defined as

$$R'_{i,j} = \frac{M_{i,j}}{N_j^{\mathrm{rec}\wedge\mathrm{gen}}} \tag{4.4}$$

where $N_j^{\mathrm{rec}\wedge\mathrm{gen}}$ is the number of events generated and reconstructed in bin j. The ratio of R by R' is a function of the truth bin j and is equal to the acceptance correction (Sec. 4.3.1)

$$A_j = \frac{N_j^{\mathrm{rec}\wedge\mathrm{gen}}}{N_j^{\mathrm{gen}}}. \tag{4.5}$$

Now let's take the case of real data, where we don't have any information about distributions at the truth level, the idea of unfolding is to apply the inverse of the response matrix calculated using MC simulation to real data to estimate the true physical distributions. At this moment, the unfolding problem is an inversion problem of the response matrix:



The use of the unfolding technique in high energy physics allows to obtain results which are independent from detector and reconstruction effects. Consequently, the unfolding results can be compared directly to theoretical predictions or to other experiments. They also can be used for the precision measurements as the W boson mass M_{W} measurement. On the other hand, there are some cases where the unfolding is not needed. Mainly, the unfolding is used for observables characterised by a large migration between truth and reconstruction distributions. In other words, for the observables with small migration between the truth and reco level, a bin-by-bin correction is sufficient to determine the true physical distributions of the observables of interest. Applying the inverse of the migration matrix to the reconstructed simulation distribution is considered as a closure test for the unfolding.

4.3 Iterative Bayesian unfolding

In this thesis, the iterative Bayesian unfolding [56] is used for the unfolding of our variables of interest with RooUnfold [7]. This paragraph will give an overview of the method, with a detailed description of the propagation of the source uncertainties through the unfolding. The iterative Bayesian unfolding is based on Bayes theorem, which describes the probability of an effect based on prior knowledge of causes related to the effect. Let us consider a list of causes and effects (C, E), where causes (C) correspond to the true values and effects (E) to the values after smearing. Each effect (E) results from several causes. The unfolding problem can be summarised in the estimation of $P(C_i|E_j)$ which corresponds to the probability to observe a cause C_i responsible of observed effects E_j.

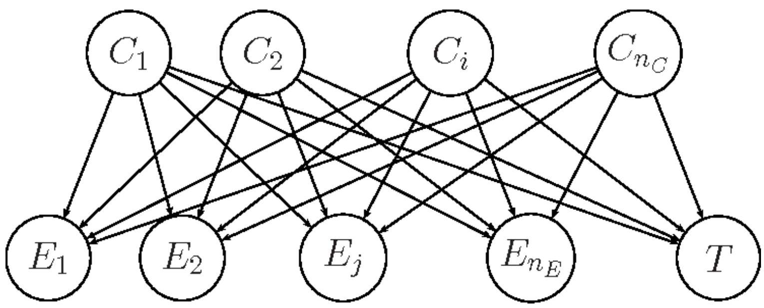


FIGURE 4.2: Probabilistic links from causes to effects. The node T corresponds to undetected events [56].

In the Bayes theorem, the probability $P(C_i|E_j)$ can be calculated as:

$$P\left(C_i|E_j\right) = \frac{P\left(E_j|C_i\right) \cdot P(C_i)}{\sum_{l=1}^{n_C} P\left(E_j|C_l\right) \cdot P\left(C_l\right)}, \tag{4.6}$$

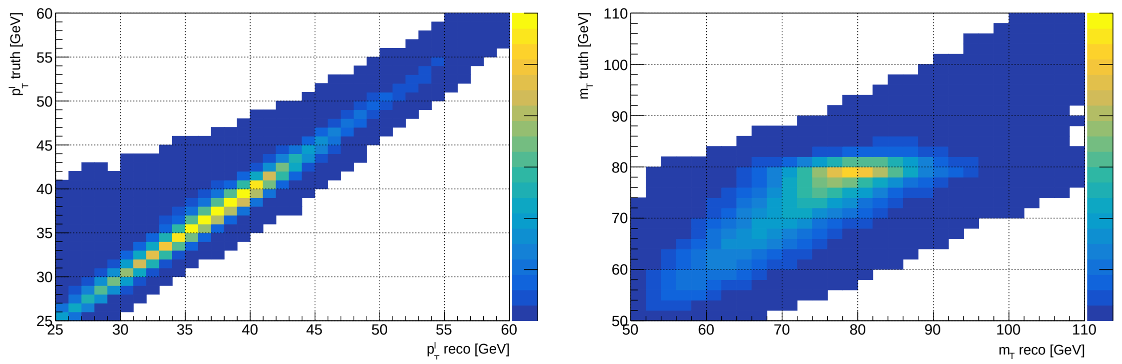
n_c corresponds to the number of possible causes, $P(E_j|C_i)$ represents the probability to observe the effect E_j knowing C_i and $P(C_i)$ is the probability to observe the cause (i). Finally the number of events in the cause bin (i) can be expressed as:

$$\hat{n}\left(C_i\right) = \frac{1}{\epsilon_i} \sum_{j=1}^{n_E} n\left(E_j\right) \cdot P\left(C_i|E_j\right), \quad \epsilon_i \neq 0, \tag{4.7}$$

$n(E_j)$ corresponds to the number of events in the effect bin (j) and $P(C_i|E_j)$ is calculated with formula (4.6) which is using $P(E_j|C_i)$ based on simulation. The iterative Bayesian unfolding is characterised by a bias [139] that we introduce with the unfolding procedure. To reduce the unfolding bias, a regularization parameter is used. The regularization consists in repeating the unfolded procedure several times, as will be discussed later in Sec. 4.4.3. The migration matrix can be determined from simulation by filling a two-dimensional histogram for all selected events with a common matching of truth and reconstructed values (TR) [20].

4.3.1 Migration matrix

The migration matrix is a matrix containing information from the truth and reconstructed level, with e.g. the x-axis corresponding to reconstructed bins and the y-axis to truth bins. The example in Figure 4.3 shows the migration matrix for two variables, p_{T}^{ℓ} and m_{T}^{W}. Comparing m_{T}^{W} to p_{T}^{ℓ} matrix, the transverse mass is characterised with a larger migration between the truth and reconstructed level because of the detector effects which affect more the transverse mass m_{T}^{W}.



(A) (B)

FIGURE 4.3: Example of the migration matrix for p_{T}^{ℓ} (A) and m_{T}^{W} (B).

In addition to the migration and response matrix, there are two important factors, as shown in Figure 4.4, that we apply before and after the unfolding, and will be used later especially for the measurement of the differential cross sections:

The efficiency correction: It is defined as the fraction of events passing reconstructed and truth level selections ($N^{\text{reco,truth}}$) to the number of events that meet the selection criteria at reconstruction level (N^{reco}):

$$\epsilon_i = \frac{N^{\text{reco,truth}}}{N^{\text{reco}}}. \tag{4.8}$$

It is defined as a function of the reconstructed bin number i. The efficiency correction is applied before unfolding to correct data distributions since the data events pass reconstructed selections only.

The acceptance correction: It is defined as the fraction of events that passing reconstructed and truth level selections ($N^{\text{reco,truth}}$) to the number of events that meet the selection criteria at truth level (N^{truth}):

$$A_i = \frac{N^{\text{reco,truth}}}{N^{\text{truth}}}. \tag{4.9}$$

It is defined as a function of the truth bin number i. The inverse of the acceptance is applied to the unfolded distribution in order to extrapolate to the truth fiducial phase space. This has to be done because the unfolding is done with a response matrix R' obtained with events satisfying both truth and reconstructed criteria.

It is worth to note that the events passing $N^{\text{reco,truth}}$ and N^{reco} selections receive both reconstructed and truth weights i.e. SF efficiency, hadronic recoil, calibration, polarisation, generator weights, while the events passing N^{truth} have only truth weights applied.

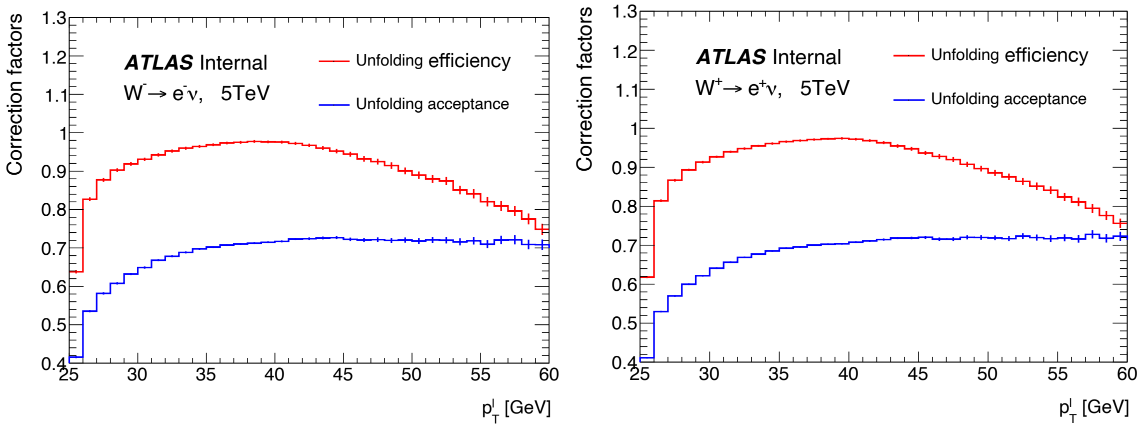


FIGURE 4.4: Example of the acceptance and efficiency factors for p_{T}^{ℓ}.

4.4 Uncertainties with unfolding

The propagation of the statistical and systematic uncertainties through unfolding is a crucial technical aspect when the unfolding is applied to an analysis. In this part, we discuss the propagation of the uncertainties in the iterative Bayesian unfolding.

4.4.1 Propagation of the statistical uncertainty

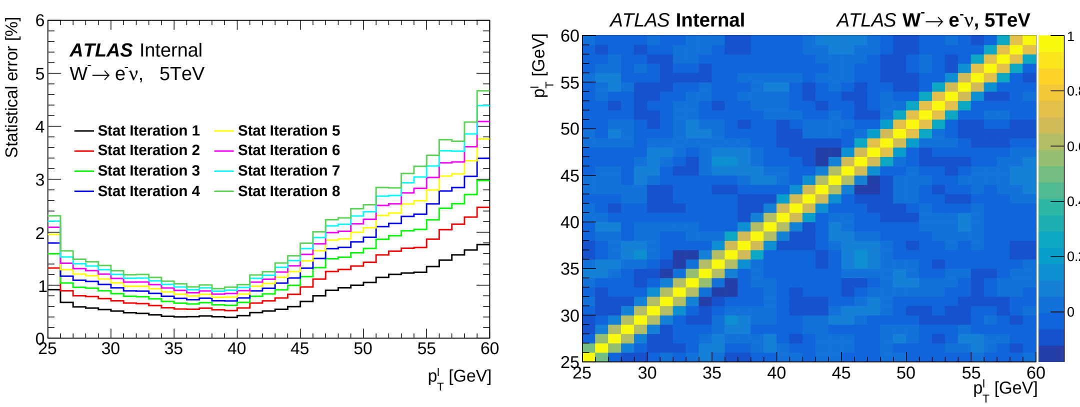
The propagation of the statistical uncertainties through the unfolding is done using pseudo-data (toys). Basically, the idea is to fluctuate the unfolding inputs (data distributions) with Poisson variations [37] to generate toys. Then, for each toy we redo the unfolding procedure using the nominal (not modified) migration matrix. The covariance matrix for the statistical uncertainty is calculated by comparing the unfolded distributions for each toy using:

$$\mathrm{Cov}(i,j) = \frac{1}{n-1}\sum_{k=1}^{n}\left(X_i^{\mathrm{k}} - \bar{X}_i\right)\left(X_j^{\mathrm{k}} - \bar{X}_j\right)^{\mathrm{T}}, \tag{4.10}$$

where X_i^k (X_j^k) corresponds to the content of bin i (j) of the unfolded toy k, $\bar{X}_i$ ($\bar{X}_j$) corresponds to the content of bin i (j) of the average of all toys. The correlation matrix between bins for the statistical uncertainty is calculated using the covariance matrix by the formula:

$$\mathrm{Corr}(i,j) = \frac{\mathrm{Cov}(i,j)}{\sqrt{\mathrm{Cov}(i,i)} \times \sqrt{\mathrm{Cov}(j,j)}}. \tag{4.11}$$

Propagation of the statistical uncertainty for MC simulation is treated differently from data. In fact, the statistical uncertainty for simulation is treated as a systematic uncertainty, and the unfolding for simulation toys is done with a modified migration matrix instead of the nominal migration matrix. Figure 4.5 shows an example of the statistical uncertainty with the correlation matrix for the unfolded distribution.



(A) (B)

FIGURE 4.5: (A) Example of the statistical uncertainty for different iterations. (B) Example of the correlation matrix for the the statistical uncertainty of the unfolding distribution.

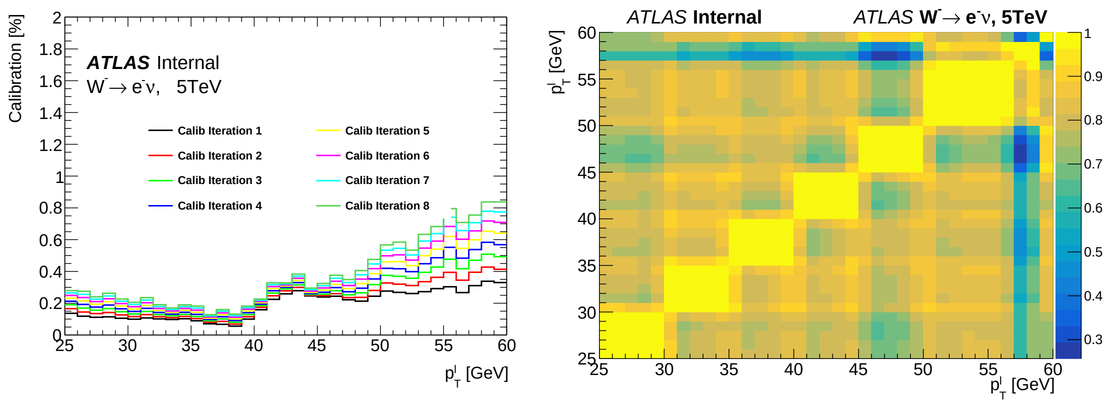
Because of the correlation between truth and reconstruction level for our variables of interest, the statistical uncertainty increases with the number of iterations, as shown in Figure 4.5. Along with the increase of uncertainty with the number of iterations, the anti-correlation between bins increases also to ensure that the statistical uncertainty is independent of the number of iterations when we integrate over all the bins.

4.4.2 Propagation of systematic uncertainties

The estimation of systematic uncertainties at the unfolded level is based on simulated distributions. For a given systematic uncertainty, we varied the inputs distributions (reconstructed distributions and migration matrix) according to this systematic uncertainty. The propagation of the systematic uncertainty through unfolding is estimated as the the difference between the unfolding of the nominal distribution and the unfolding of the modified distribution. For the same reason of migration between bins, the systematic uncertainties increase with the number of iterations as seen in Sec. 4.6. After the unfolding, all the systematic uncertainties are assumed to be fully correlated between the bins, and the covariance matrix (V) is calculated as:

$$V_{i,j} = \sigma_i \times \sigma_j, \tag{4.12}$$

where σ_i (σ_j) is the systematic uncertainty in bin $i(j)$. Figure 4.6 shows as an example the calibration systematic uncertainty as a function of iteration and the corresponding correlation matrix. In fact, the systematic uncertainties must be independent of the number of iterations, and the variation with the number of iterations is related to statistical fluctuations in the systematic variations. For the choice of the number of iterations, the systematic uncertainties are not included in the optimisation study described in Sec. 4.5.



(A) (B)

FIGURE 4.6: (A) Example of the systematic uncertainty for different iterations. (B) Example of the correlation matrix. The calibration uncertainty is defined as the sum of several variations.

4.4.3 Bias uncertainty with unfolding

In addition to the statistical and systematic uncertainties, there is the unfolding bias that we have to take into account. This bias is related mainly to the unfolding method and can be estimated with different approaches. The approach used in this chapter is a simple one used for the unfolding of a variable with small migration between reconstruction and truth level, like for p_{T}^{ℓ} and η_{ℓ}. For the unfolding of a variable with larger migration like p_{T}^{W}, a more involved approach is used and will be described later. The procedure to estimate the bias, through a "data-driven closure test" using the data/MC shape differences for the unfolded observable, can be summarised in two steps: (Figure 4.7) [111]:

- Reweight the MC distribution at truth level with the fitted ratio of data over simulation, in such a way that the reconstructed distribution after the reweighting matches the data in which the background has been subtracted. As shown in Figure 4.8, as we expect, the ratio data/MC is closer to 1 for the reconstruction-weighted distribution.
- The bias is estimated as the difference between the unfolding of the reconstruction-weighted distribution and the truth-weighted distribution.

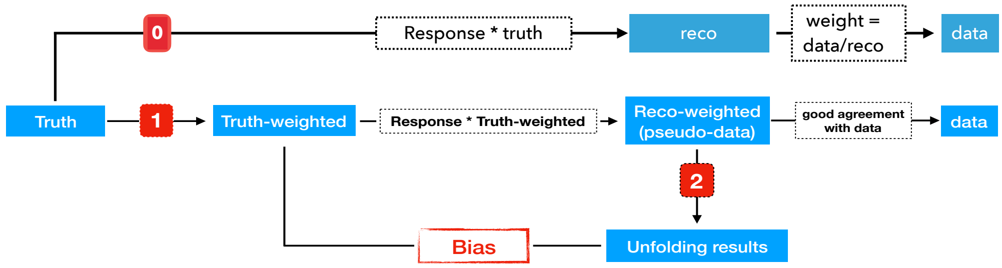
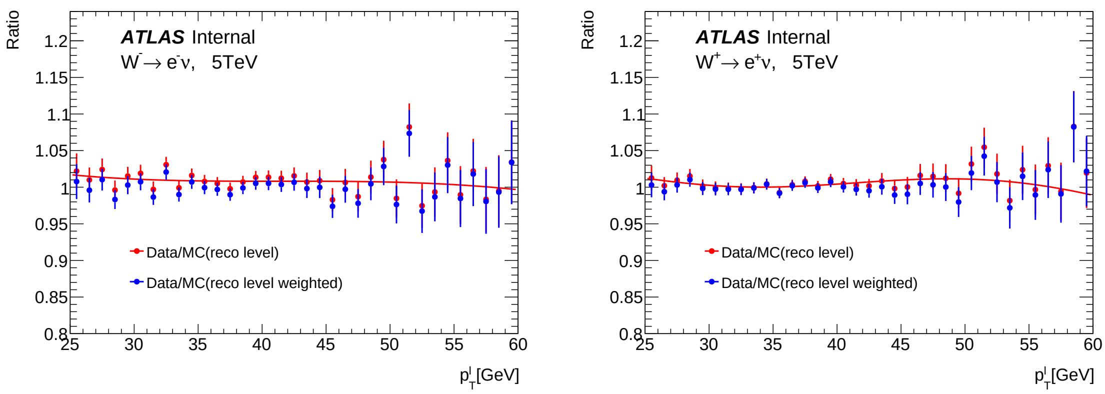


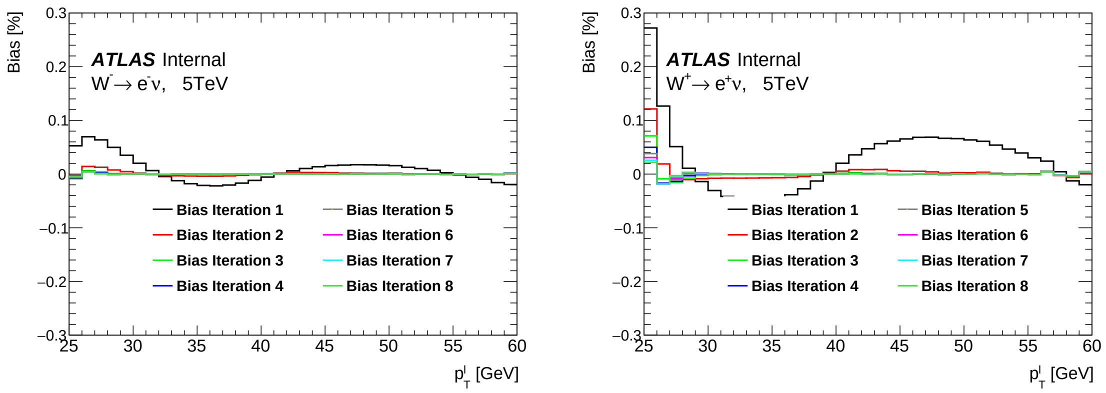
FIGURE 4.7: An overview of the procedure used to estimate the unfolding bias.



(A) (B)

FIGURE 4.8: Comparison of the ratio data/MC using the reconstruction and weighted reconstruction-distributions, for W^- (A) and W^+ (B) at 5 TeV.

In general, the unfolding bias decreases with the number of iterations, as shown in Figure 4.9. Also, as the unfolding does not change the normalisation of the input distributions, the total integrated unfolding bias when we take the correlation (anti-correlation) between bins into account must be equal to 0. Contrary to other source of uncertainties, the bias decreases with the number of iterations and the anti-correlation between bins increases with the number of iterations to ensure that the integrated bias is zero.



(A) (B)

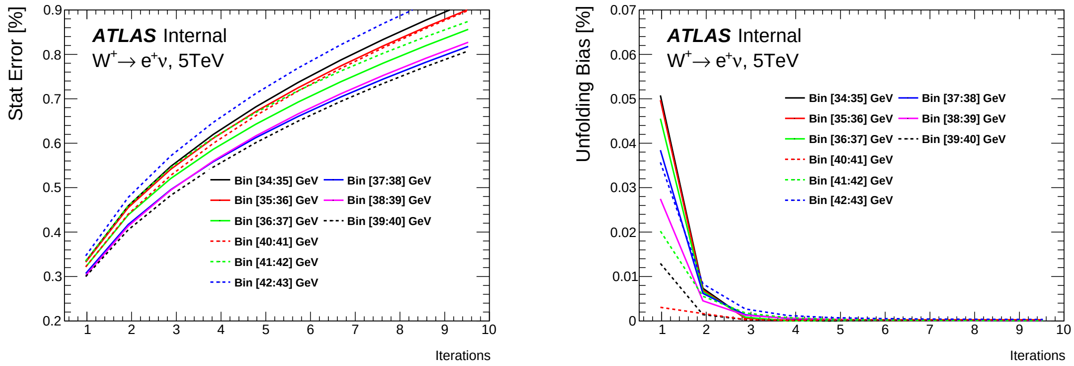
FIGURE 4.9: Comparison of the unfolding bias for different iterations, for W^- (A) and W^+ (B) at 5 TeV.

4.5 Optimisation of the number of iterations

As discussed above, the statistical uncertainty increases with the number of iterations, whereas the unfolding bias, considered as a source of uncertainty, decreases with the number of iterations, as seen in Figure 4.10. Therefore, it is possible to optimise the number of iterations by minimising the combined statistical and bias uncertainties. The other systematic uncertainties are not included in the optimisation as they should be independent of the number of iterations as mentioned earlier. Also, the optimisation should be performed for a selected region of the unfolded distribution since we can not use the whole range of the unfolded distribution (the bias is zero). The example in Figure 4.10 shows the information that can be used for the bin-by-bin optimisation around the peak region:
For our example shown in Figure 4.11, as the bias is very small comparing to other source of uncertainties, the best choice is to use the first iteration. But to avoid the fluctuation/bias in the first iteration, see Figure 4.9, the 2$^{\text{nd}}$ iteration is chosen instead.

4.6 Bin-by-bin unfolding

The bin-by-bin unfolding consists in applying a correction factor that we extract directly from simulation. This unfolding method is used basically in the case where the variable of interest is characterised with a small migration between truth and reconstructed level and when the number of bins is the same between the truth



(A) (B)

FIGURE 4.10: Statistical (A) and unfolding bias (B) uncertainties as a function of the number of iterations for different bins of p_{T}^{ℓ}, around the peak region in our distribution of interest.

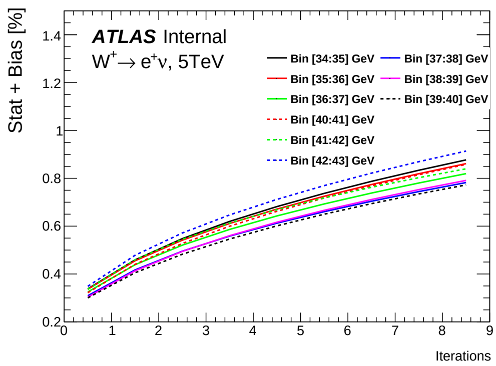


FIGURE 4.11: Sum of the statistical and unfolding bias uncertainties as a function of the number of iterations for different bins of p_{T}^{ℓ}, around the peak region in our distribution of interest.

and reconstructed distributions. Let us consider a MC truth distribution x_i^{gen} and a MC reconstructed distribution y_i^{rec}. The correction factor is calculated as:

$$C_i = \frac{x_i^{\text{gen}}}{y_i^{\text{rec}}}. \tag{4.13}$$

The unfolded data using the bin-by-bin method is calculated as:

$$\text{Unfolded}_i = C_i \times \text{data}_i. \tag{4.14}$$

The bin-by-bin is used only in the case where the detector effects are very small, otherwise this method will introduce a large bias [110]. This method can be used mainly for the unfolding of p_{T}^{ℓ} and η_{ℓ}.

Chapter 5

Measurement of the W-boson transverse momentum distribution

5.1 Introduction

One of the most important theoretical sources of uncertainties in the measurement of the W-boson mass, is the extrapolation of the boson p_{T} distribution from Z-boson to W-boson ($\approx$ 6 MeV [155]), where the QCD high order predictions are not sufficiently precise to describe the data. A precise direct measurement of p_{T}^{W} will provide a direct comparison with QCD predictions, this is equivalent to saying that replacing the theoretical extrapolation from p_{T}^{Z} by such a direct measurement of the p_{T}^{W} distribution will improve the precision of the measurement of M_W. Measuring the p_{T}^{W} distribution in low p_{T}^{W} region (p_{T}^{W} < 30 GeV) with an uncertainty $\approx 1\%$ in bin of 5 GeV will reduce the QCD modelling uncertainty [115] in the measurement of M_W by a factor of two [132]. The p_{T}^{W} distribution is reconstructed using $W \to \ell\nu$ events, where the charged leptons are measured in the different tracking detectors or in the EM calorimeter, as discussed in Chapter 2, while the neutrino leaves the detector unseen. Because of the neutrino, the p_{T}^{W} distribution is reconstructed through the hadronic recoil, u_{T}, defined as the vector sum of all energy deposits excluding the energy of the lepton. The transverse momentum of the W boson is defined by:

$$\vec{p}_{\mathrm{T}}^{W} = -\vec{u}_{\mathrm{T}}, \tag{5.1}$$

and the transverse momentum of the decay neutrino $\vec{p}_{\mathrm{T}}^{\nu}$ is inferred from the vector of the missing transverse momentum $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$ which corresponds to the momentum imbalance in the transverse plan:

$$\vec{p}_{\mathrm{T}}^{\mathrm{miss}} = -(\vec{p}_{\mathrm{T}}^{\ell} + \vec{u}_T). \tag{5.2}$$

For the reconstruction of p_{T}^{W}, a good understanding of $\vec{u}_{\mathrm{T}}$ is needed. The reconstruction of the hadronic recoil is described in [101]. The measurement of p_{T}^{W} is based on low number of interactions per bunch crossing data (low pile-up μ) to ensure a reasonable resolution on the hadronic recoil, as shown in Figure 5.1, which shows the comparison of the resolution on the hadronic recoil between high pile-up runs (black circles) and low pile-up runs (red points). In this chapter, we will describe the measurement of the W-boson transverse momentum through the unfolding of the p_{T}^{W} distributions at the detector level, using the unfolding method described in Chapter 4, with low pile-up data sets collected during Run 2 at $\sqrt{s}$

= 5 and 13 TeV. Also, a different approach is used to estimate the unfolding bias for the p_{T}^{W} analysis, in order to improve our evaluation of the unfolding bias. The new approach, described in [114], consists of using a different reweighing method to get the best data/MC agreement. The main signal events for W and Z boson productions are described in [95]. There generated using the POWHEG event generator using the CT10 PDF interfaced to PYTHIA8 using the AZ NLO tune, and being interfaced to PHOTOS++ to simulate the effect of final state QED radiation.

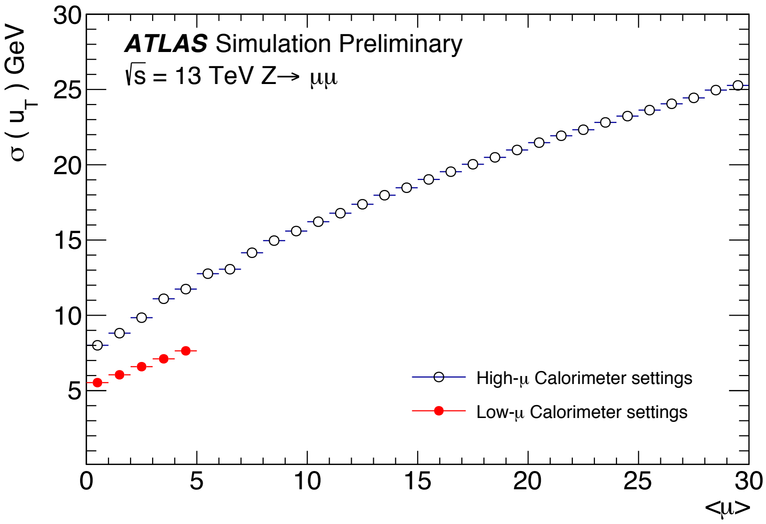


FIGURE 5.1: Hadronic recoil resolution as a function of $\langle\mu\rangle$ for simulated $Z \rightarrow \mu\mu$ events with two different calorimeter settings [132], see Chapter 3 for a discussion of the calorimeter settings.

5.2 Data and simulated distributions

5.2.1 Selections

The selections of $W \rightarrow \ell\nu$ events for the p_{T}^{W} distribution are based on the following two triggers *HLT_e15_lhloose_nod0_L1EM12* and *HLT_mu14*, for electrons and muons, respectively. In addition, events are required to contain one lepton with $p_{\mathrm{T}}^{\ell} > 25$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV to reduce background effects. In addition, the W boson transverse mass defined as $m_{\mathrm{T}}^{W} = \sqrt{2p_{\mathrm{T}}^{\ell}p_{\mathrm{T}}^{\mathrm{miss}}(1-\cos(\Delta\phi))}$, with $\Delta\phi$ being the azimuthal opening angle between the charged lepton and the missing transverse momentum, is chosen to be $m_{\mathrm{T}}^{W} > 50$ GeV. A detailed description of the selections, with the final number of events which pass all the selections, is given for 5 and 13 TeV samples separately in [114].

5.2.2 Control plots for the p_{T}^{W} distribution

Once all the events pass the selections described above, we show the distributions of the W-boson transverse momentum for data compared to MC simulation

TABLE 5.1: Analysis cut flow for $W^+ \to e^+\nu_e$ 5 TeV signal selection.

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

TABLE 5.2: Analysis cut flow for $W^+ \to e^+\nu_e$ 13 TeV signal selection.

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

TABLE 5.3: Analysis cut flow for $W^+ \to \mu^+\nu_\mu$ 5 TeV signal selection.

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

TABLE 5.4: Analysis cut flow for $W^+ \to \mu^+\nu_\mu$ 13 TeV signal selection.

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

TABLE 5.5: Analysis cut flow for $W^- \to e^-\bar{\nu}_e$ 5 TeV signal selection.

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

TABLE 5.6: Analysis cut flow for $W^- \to e^-\bar{\nu}_e$ 13 TeV signal selection.

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

TABLE 5.7: Analysis cut flow for $W^- \to \mu^-\bar{\nu}_\mu$ 5 TeV signal selection.

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

TABLE 5.8: Analysis cut flow for $W^- \to \mu^- \bar{\nu}_\mu$ 13 TeV signal selection.

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

(signal and background) for 5 TeV, in Figure 5.2 and 13 TeV in Figure 5.3 separately. The bottom panels show the ratio data to simulation, with the green band corresponding to the total uncertainty with the statistical and systematic uncertainties added in quadrature. In general, one finds good agreement between the data and the predicted number of events within the uncertainty except for some of the 13 TeV cases.

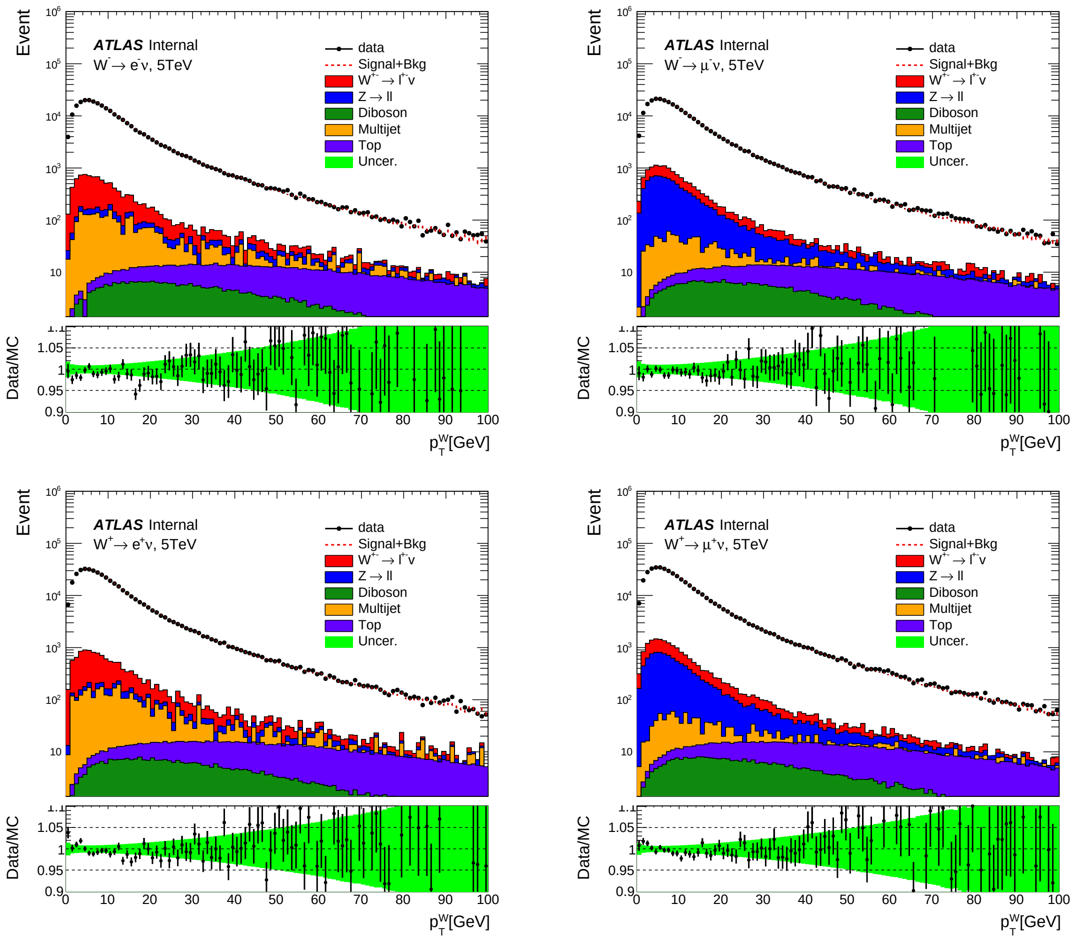


FIGURE 5.2: Reconstructed p_{T}^{W} distributions in data compared to MC (signal and background) in the electron (left) and muon (right) channels for negative (top) and positive (bottom) charges for the $\sqrt{s}$ = 5 TeV data set. The lower panel of each plot shows the data to simulation ratio, together with the total uncertainty at the detector level. The green band is dominated by the uncertainty due to the calibration of the hadronic recoil and the statistical uncertainty. The different sources of uncertainties at the detector level are shown in Appendix B.

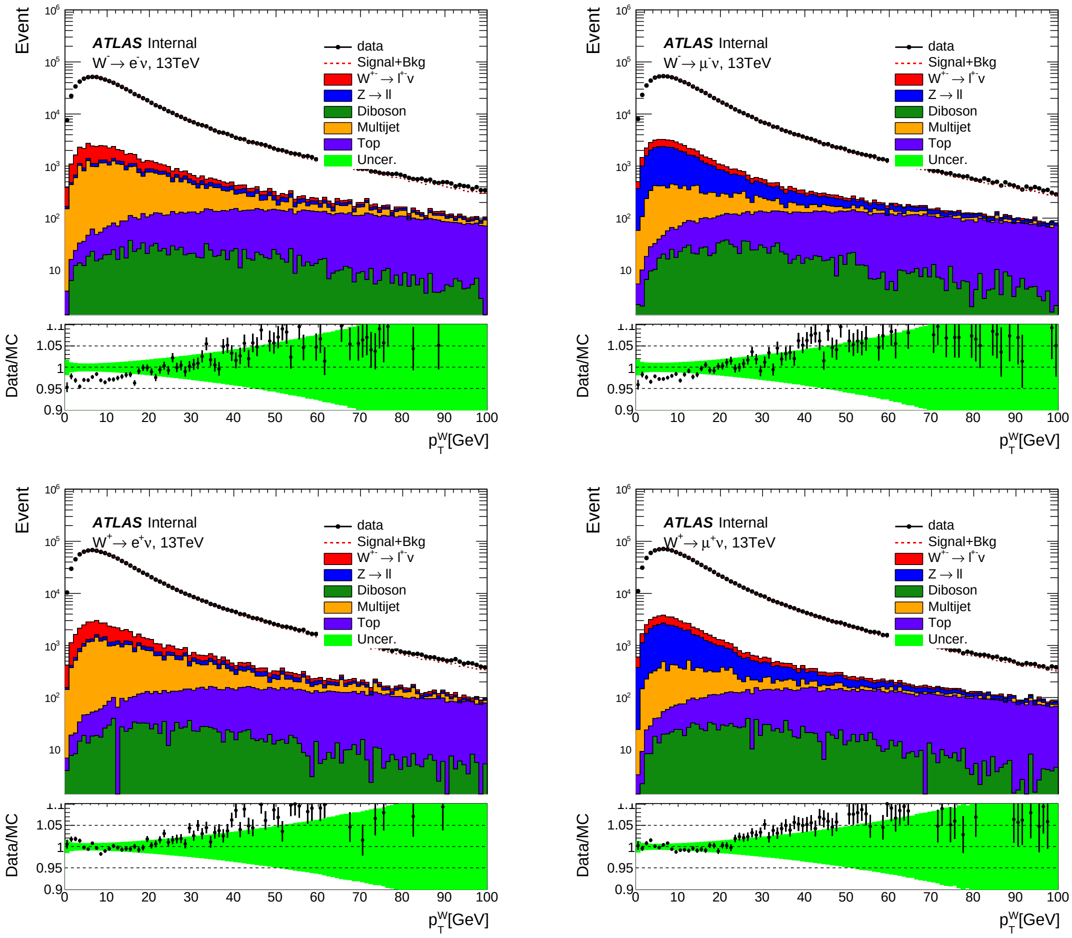


FIGURE 5.3: Same as Figure 5.2 but for the $\sqrt{s} = 13$ TeV data set. The agreement is generally worse at 13 TeV compared to 5 TeV, because for simulation we use the the same tuning, AZ tuned at 7 TeV, which gives a better agreement between data and simulation for 5 than 13 TeV.

5.3 Data unfolding

5.3.1 Unfolding description

As described in Chapter 4, the Bayesian unfolding method is used to unfold data distributions. The unfolding procedure starts by subtracting the background effects from data distributions. The background contribution is based on simulation samples, and their effect on the data is estimated using the formula:

$$\text{data}_i^{\text{corrected}} = \text{data}_i \times \left(1 - \frac{N_i^{\text{Bkgr}}}{N_i^{\text{Sig}} + N_i^{\text{Bkgr}}}\right), \tag{5.3}$$

where N_i^{Bkgr} is the sum of all the background contributions in bin i, showed in Figures 5.2 and 5.3, N_i^{Sig} is the number of events in signal in bin i. Then, the efficiency correction factor, defined in Chapter 4, is applied to data. Figure 5.4 shows an example of such efficiency correction factors. Once the data distributions are corrected, it can be unfolded as described in Chapter 4 using the migration matrix. Figure 5.5 shows an example of the migration matrix used for the p_{T}^{W} unfolding. The migration matrix is characterised by a large migration between truth and detector variables, which makes the unfolding more involved than that of p_{T}^{ℓ} and η_ℓ. The migration matrix is determined using the simulation samples,

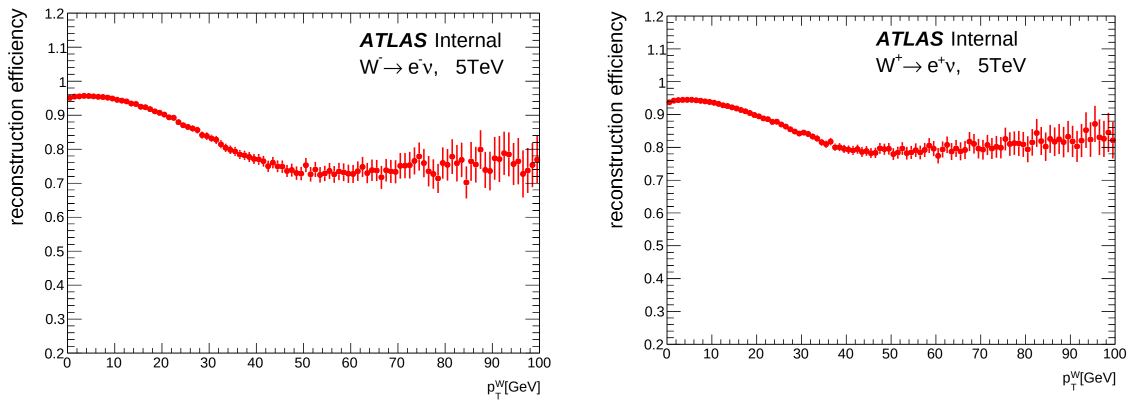


FIGURE 5.4: Example of the unfolding efficiency factor, defined as the ratio of the number of events at the reconstruction level with correspond to a truth level selection divided by the total number of events at the reconstruction level. This efficiency is applied to correct data distributions before unfolding, for electron channels at $\sqrt{s} = 5$ TeV.

Powheg+Pythia8 [43] in our analysis, where the x-axis corresponds to the reconstructed bins and the y-axis to true bins. The migration matrix is constructed in such a way that each event passes both truth and reconstructed selections. The migration between the truth and the reconstructed levels depends on detector effects (such as the finite resolution of the detector and the limited reconstruction efficiency). After the unfolding, the unfolded distribution can not be compared directly to the truth distribution, since this unfolded distribution corresponds to the truth distribution with both truth and reconstructed selections. For a direct

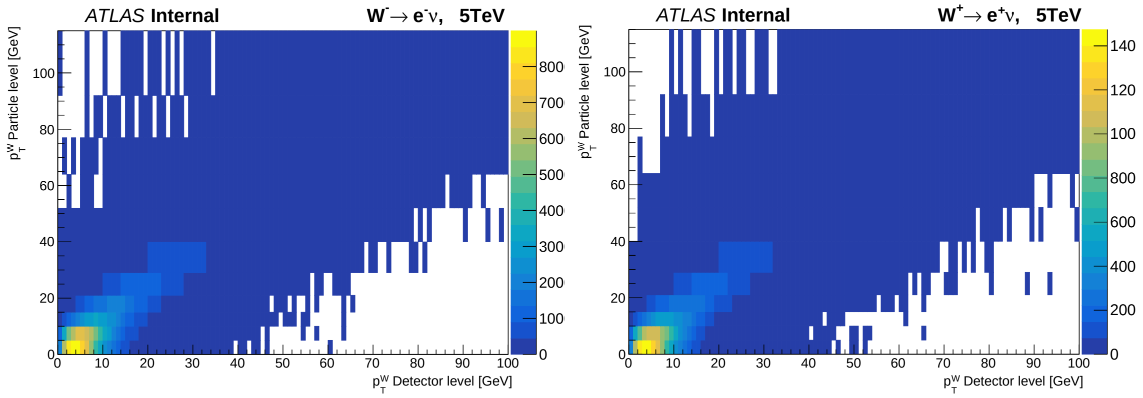


FIGURE 5.5: Example of the migration matrix of p_{T}^{W} for electron channels at $\sqrt{s}$ = 5 TeV. The correlation between bins is more important in the low p_{T}^{W} region (p_{T}^{W} < 30 GeV).

comparison, the unfolded distribution needs the acceptance correction, discussed in Chapter 4. After all the corrections, Figure 5.6 shows an example of the comparison between the truth, reconstructed data and the unfolded data distributions. As described in Chapter 4, the Bayesian unfolding method is characterised by a regularisation parameter, used to reduce the bias that we introduce with the unfolding procedure. This parameter is optimised using statistical and bias uncertainties in Sec. 5.3.6.

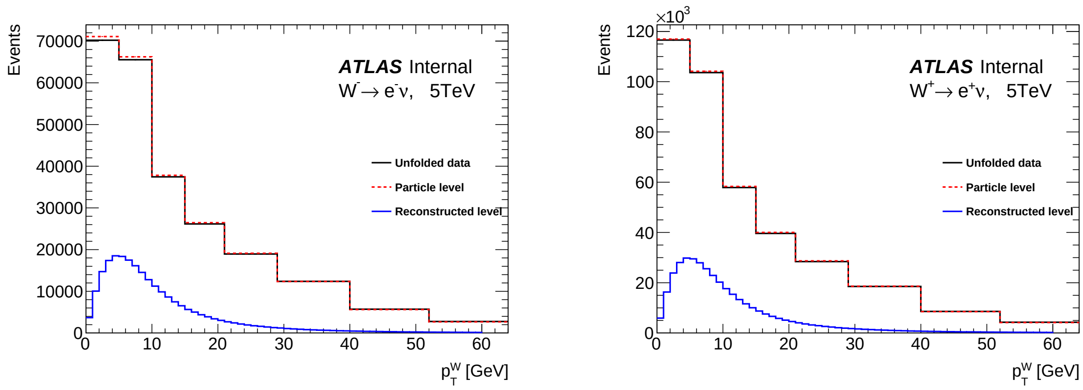


FIGURE 5.6: Example of the unfolded data distribution compared to reconstructed data events and the truth distributions for electron channels at $\sqrt{s}$ = 5 TeV. The acceptance correction is applied to the unfolded distribution to take into account events at the truth level which are not reconstructed.

5.3.2 Experimental systematic uncertainties

In this section, we review different sources of systematic uncertainties affecting the measurement of p_{T}^{W} distributions and the measurement of the differential cross sections in Chapter 7:

Lepton scale factors: As described in Chapter 3, two factors (energy scale and resolution) are applied to data and MC respectively to correct the residual difference observed between data and simulation. The combined effect of all scale and resolution uncertainties on the distributions of p_{T}^{W} is shown in Figures 5.7 and 5.8. The effect on p_{T}^{W} is up to 0.2% in low p_{T}^{W} region.

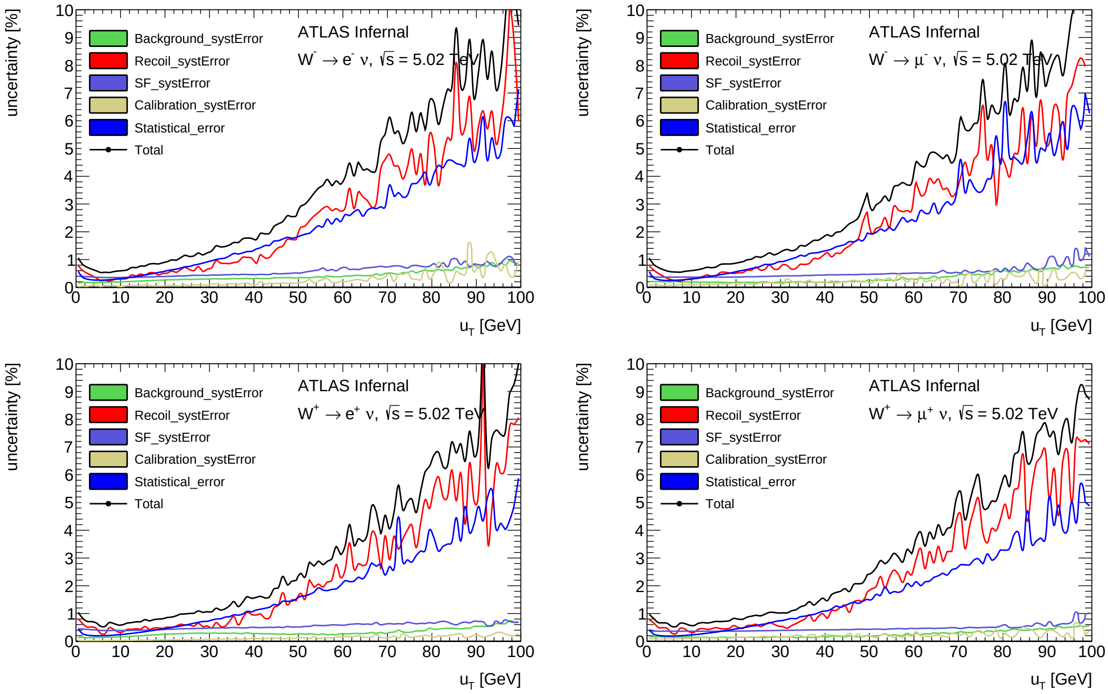


FIGURE 5.7: Different sources of uncertainties on the measurement of p_{T}^{W} distributions at the detector level for the $\sqrt{s}$ = 5 TeV data set. The total uncertainty is less than 1% in the low p_{T}^{W} region (p_{T}^{W} < 30 GeV) and around 5% in the high p_{T}^{W} region ($p_{\mathrm{T}}^{W} \approx 100$ GeV). The total uncertainty is dominated by the hadronic recoil calibration uncertainty and the statistical uncertainty of the data.

Lepton selection efficiency: As detailed in Sec. 7.2, selected leptons are required to pass specific criteria. The efficiency of the selections in the simulation is normalised to that in data and applied to the simulation as product of different scale factors (SFs):

$$W_{\text{event}} = \mathrm{SF}_{\text{reco}} \cdot \mathrm{SF}_{\text{ID}} \cdot \mathrm{SF}_{\text{isolation}} \cdot \mathrm{SF}_{\text{trigger}}\,, \tag{5.4}$$

which correspond to the reconstruction, identification, isolation and trigger scale factors. The SFs are calculated using a "tag-and-probe" method detailed in [22].

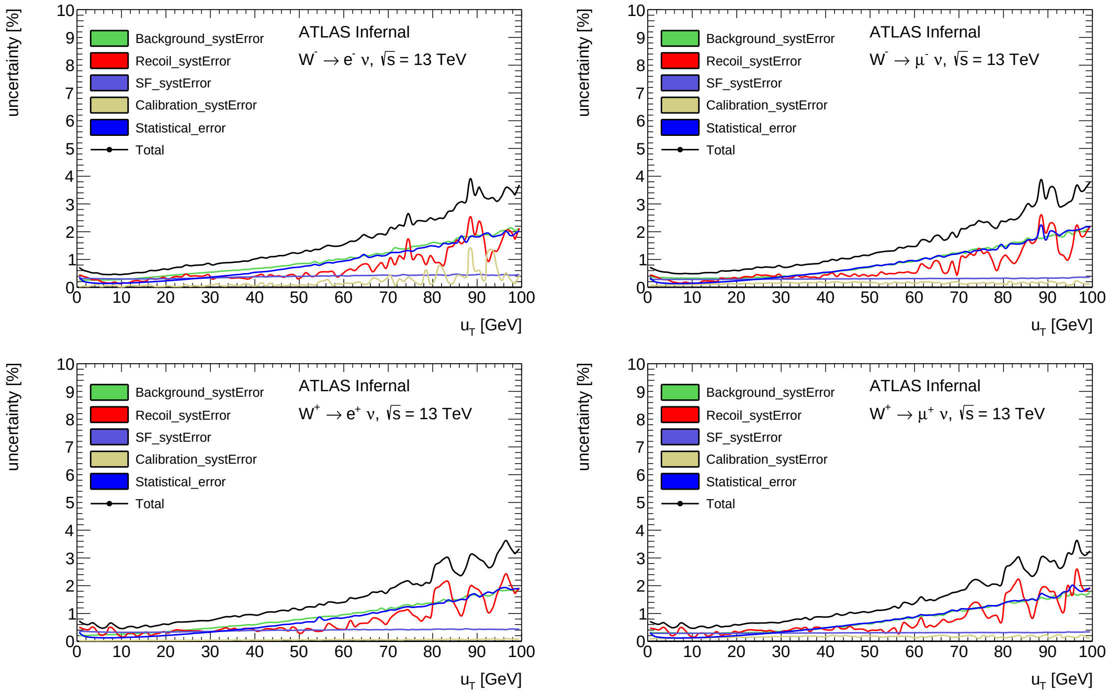


FIGURE 5.8: Different sources of uncertainties on the measurement of p_{T}^{W} distributions at the detector level for the $\sqrt{s} = 13$ TeV data set. The total uncertainty is less than 1% in the low p_{T}^{W} region ($p_{\mathrm{T}}^{W} < 30$ GeV) and around 5% in the high p_{T}^{W} region ($p_{\mathrm{T}}^{W} \approx 100$ GeV). The total uncertainty is dominated by the hadronic recoil calibration uncertainty and the background uncertainty (because of the large background contributions of gauge-boson pair production and top quark production).

Hadronic recoil calibration: Because the neutrino can not be measured in the ATLAS detector, the hadronic recoil, defined as the vector sum of all energy deposits excluding energy of lepton, is used in the W boson analysis to determine p_{T}^{ν} and p_{T}^{W}. The uncertainty coming from the calibration of the hadronic recoil is dominated mainly by data statistics, specially at low p_{T}^{W}. The uncertainty on the hadronic recoil calibration is the dominant systematic uncertainty compared to other source of uncertainties.

Background uncertainty: It is related to the background estimation, in particular to the multi-jet contribution [155], and varies between channels and center-of-mass energies. In general, the background uncertainty is below 0.5% for our regions of interest.

Luminosity: The luminosity uncertainty for 13 TeV low pile-up runs is 1.5 $\%$ for the combination of 2017+2018 data (2.1% for 2017, 1.5% for 2018). The luminosity uncertainty is 1.6% for 5 TeV 2017 low pile-up runs [104].

5.3.3 Propagation of statistical uncertainties

The propagation of the statistical uncertainties of the data through the unfolding is done using pseudo-data, constructed by fluctuating the data distribution with Poisson variations, and the covariance matrix of the statistical uncertainties at the unfolded level is built using the unfolding results for each pseudo-data distribution, as described in Chapter 4. There is also another approach to calculate the covariance matrix at the unfolded level, by using internal toys generated by the *RooUnfoldBayes* class. Figure 5.9 shows an example of the statistical uncertainties at the unfolded level, bin-by-bin, for different iterations. The statistical uncertainties are smaller than 1% in low p_{T}^{W} region ($p_{\mathrm{T}}^{W} < 30$ GeV) and larger than 2 $\%$ at $p_{\mathrm{T}}^{W} = 100$ GeV. Because of the correlation between truth and reconstructed levels (Figure 5.10), the statistical uncertainties increase with the number of iterations as shown in Figure 5.9. In fact, the diagonal elements of the covariance matrix increase with the number of iterations, on the other hand, the correlation between bins (non-diagonal elements) decrease to ensure that the total statistical uncertainties are independent of the number of iterations when we integrate over all bins.

5.3.4 Propagation of systematic uncertainties

The systematic uncertainties are propagated through the unfolding in the same way as described in Chapter 4. In general the propagation of systematic uncertainties is based on simulation samples, where the reconstructed distribution and the migration matrix are modified by their uncertainties. The difference between the unfolding of the modified distribution and the unfolding of the nominal distribution is considered as the systematic uncertainty. The systematic uncertainties increase also with the number of iterations, but contrary to the statistical uncertainties, the increase for the experimental systematic uncertainties is due to fluctuations related to the low statistics. Figure 5.11 shows an example of the recoil systematic uncertainty, the dominant one, as a function of the number of iterations

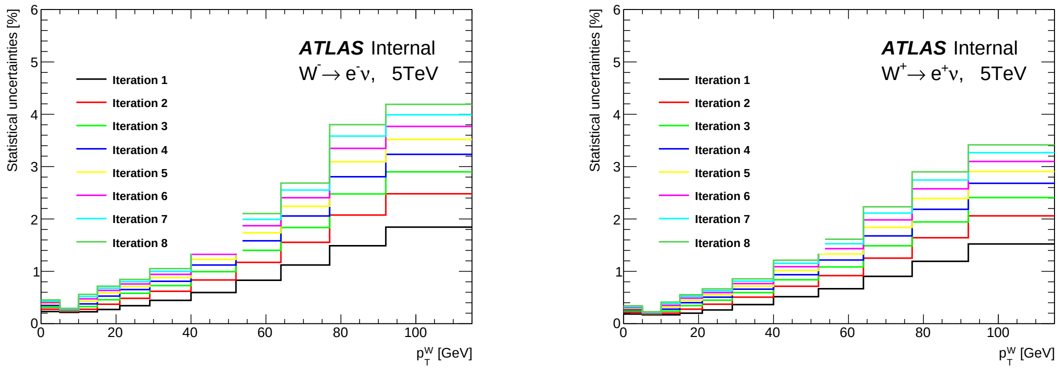


FIGURE 5.9: Statistical uncertainties on the unfolded distribution as a function of the number of iterations, bin-by-bin, for electron channels at $\sqrt{s} = 5$ TeV.

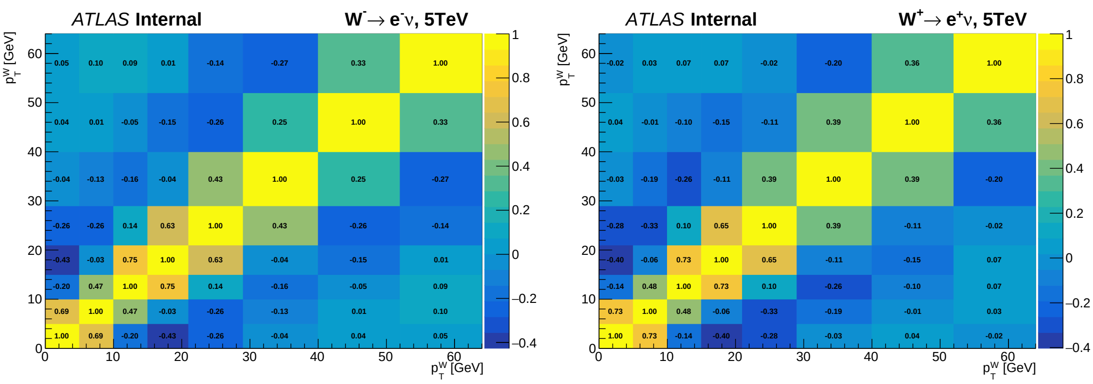


FIGURE 5.10: Correlation matrix of the statistical uncertainties, for electron channels at $\sqrt{s} = 5$ TeV, corresponding to the iteration 4.

for different bins. As for the statistical uncertainties, the correlation between bins (Figure 5.12) decreases with the number of iterations to ensure that the total uncertainty is independent of the number of iterations. All the sources of uncertainties are shown as a function of the number of iterations in Ref. [114]. The different sources of uncertainties at the detector level are shown in Appendix B.

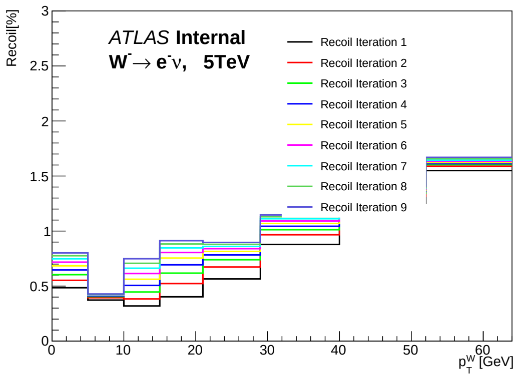


FIGURE 5.11: Example of the recoil systematic uncertainties on the unfolded distribution as a function of the number of iterations, bin-by-bin, for electron channels at $\sqrt{s} = 5$ TeV.

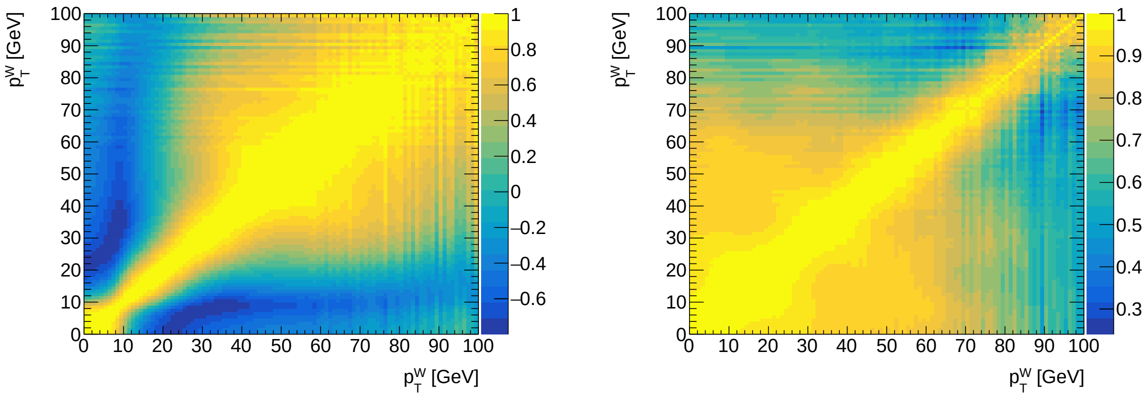


FIGURE 5.12: Correlation matrix for the hadronic recoil calibration uncertainty (left) and lepton calibration uncertainty (right), at 5 TeV, corresponding to iteration 4.

5.3.5 Comparison of the uncertainties

The breakdown of the systematic uncertainties is shown in Figure 5.13 at the unfolded level. The total experimental measurement uncertainty remains below 1% up to $p_{\mathrm{T}}^{W} = 25$ GeV at 5 TeV, and below 2% up to 50 GeV at 13 TeV, for each of the $W^{+} \to e^{+}\nu_{e}$ and $W^{-} \to e^{-}\bar{\nu}_{e}$ channels. The same results are observed also for muon channels [114]. In this range, the statistical uncertainties and recoil calibration uncertainties dominate compared to other sources of uncertainty as shown in Figure 5.13. At 13 TeV the background uncertainty is more important comparing to 5 TeV because of the large contributions of gauge-boson pair production and top-quark production [114]. At 100 GeV, the total uncertainties reach 9% and 3% for 5 and 13 TeV, respectively. The scale and hierarchy of uncertainties are preserved at the unfolded level. The breakdown of the uncertainties for the electron and the muon channels at the detector level are shown in [114]. The uncertainties are calculated using 3 iterations as a parameter of the Bayesian unfolding. The number of iterations is optimised for the measurement of the p_{T}^{W} spectrum in [114].

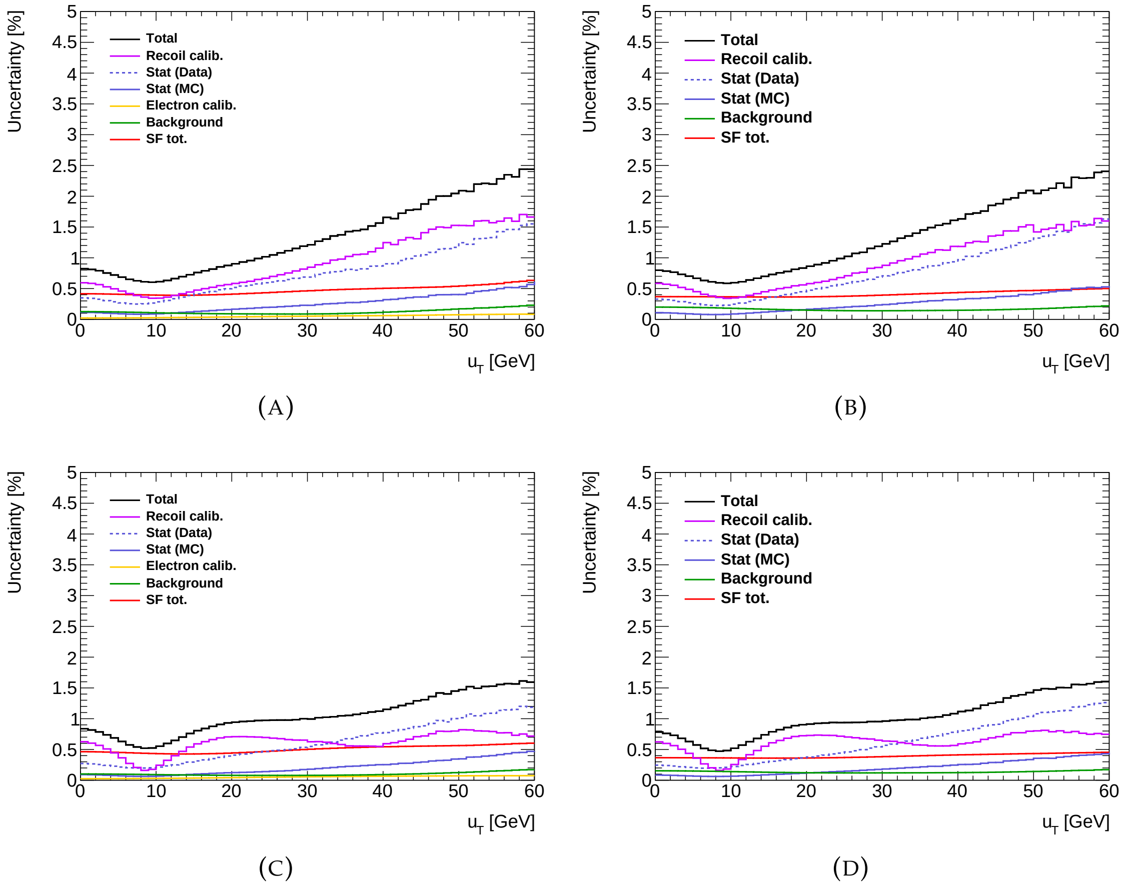


FIGURE 5.13: Different sources of uncertainties on the measurement of p_T^W distributions at the unfolded level for the $\sqrt{s}$ = 5 TeV data set, for electron channels W^- (A), W^+ (B) and muon channels W^- (C), and W^+ (D). The total uncertainty is less than 1% in the low p_T^W region ($p_T^W < 30$ GeV) and around 2% in the high p_T^W region ($p_T^W \approx 60$ GeV). The total uncertainty is dominated by the hadronic recoil calibration uncertainty and the statistical uncertainty of data.

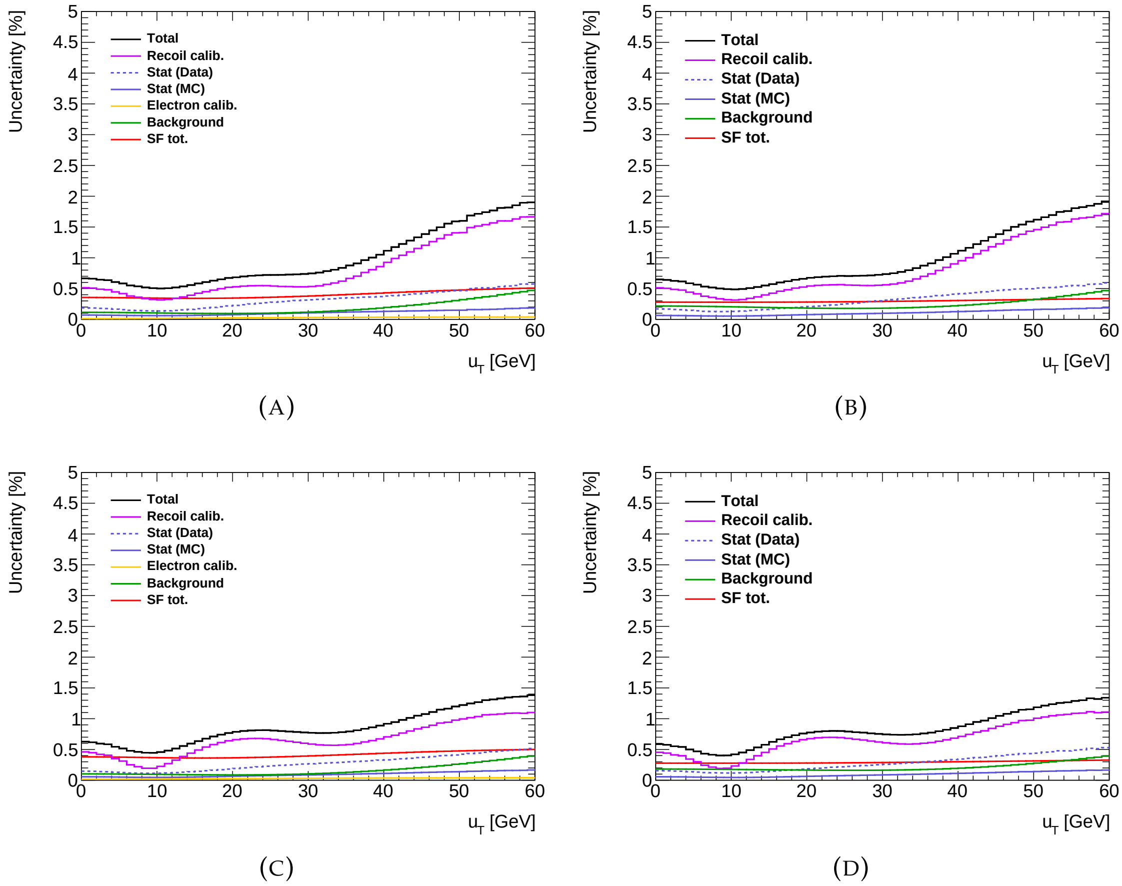


FIGURE 5.14: Different sources of uncertainties on the measurement of p_{T}^{W} distributions at the unfolded level for the $\sqrt{s} = 13$ TeV data set, for the electron channels W^{-} (A), W^{+} (B) and the muon channels W^{-} (C), and W^{+} (D). The total uncertainty is smaller than 1% in the low p_{T}^{W} region ($p_{\mathrm{T}}^{W} < 30$ GeV) and around 1.5% for the high p_{T}^{W} region ($p_{\mathrm{T}}^{W} \approx 60$ GeV). The total uncertainty is dominated by the hadronic recoil calibration uncertainty and the background uncertainty.

5.3.6 Unfolding bias

In the p_{T}^{W} analysis, the unfolding bias estimation is the major concern, because of the large migration between truth and detector levels variables, as shown in Figure 5.5. Contrary to the method used to estimate the bias described in Chapter 4, another more involved approach is used for the p_{T}^{W} analysis. As described in Chapter 4, the unfolding bias can be estimated by:

1. The MC events are reweighted at the truth level to get the best agreement to the data (reconstruction level).

2. The corresponding reconstruction-level distribution is unfolded (as pseudo-data) using the original migration matrix (used for data unfolding).

3. The unfolded result is compared to the reweighted truth distribution, thus providing an estimate of the bias uncertainty.

The new approach is to change the truth level reweighting method. In fact we usually reweight the truth distribution by the data/MC, as discussed in Chapter 4, but for the new approach, we define several reweighting functions at the truth level and we minimise the χ^2 value in order to get the best agreement at the reconstruction-level with data. Figure 5.15 shows an example of the bias uncertainty for 5 TeV. Contrary to the statistical uncertainty, the bias uncertainty decreases with the number of iterations. The bias is important for the first bins, and starts to decrease after 40 GeV because of the large bins size. The bias uncertainty is considered fully correlated, and the correlation between bins increases in order to ensure that the bias is zero when we add all the bins together, as seen in Figure 5.15.

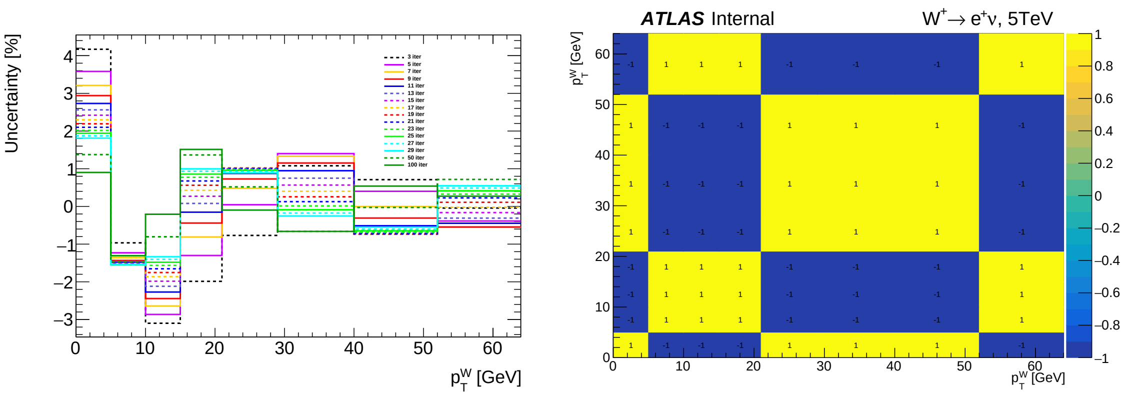


FIGURE 5.15: Relative bias uncertainty for $W^+ \to e^+\nu_e$ at $\sqrt{s} = 5$ TeV for large bins. The truth reweighting is defined based on the new method. Different number of iterations is shown.

5.4 Results of p_{T}^{W} measurement

The results for the unfolded p_{T}^{W} distributions compared to the different predictions are shown in Figures 5.16 and 5.17 for the electron and muon channels, respectively. Excluding luminosity, the experimental uncertainties range from less than 1% at low p_{T}^{W} to about 5% (2%) at p_{T}^{W}=100 GeV, at 5 TeV (13 TeV). These numbers are smaller than those quoted in Appendix B due to the large size of the binning used. The luminosity uncertainty contributes in addition to 1.6% at 5 TeV, and 1.5% at 13 TeV. The predicions include Powheg AZNLO, Pythia AZ, Sherpa and DYRES. An approximately equal level of agreement with data is visible for Powheg and Pythia. Sherpa predicts a softer spectrum, while DYRES is harder than the data. These features are consistently observed in the electron and muon channels, for both W boson charges, and at both energies.

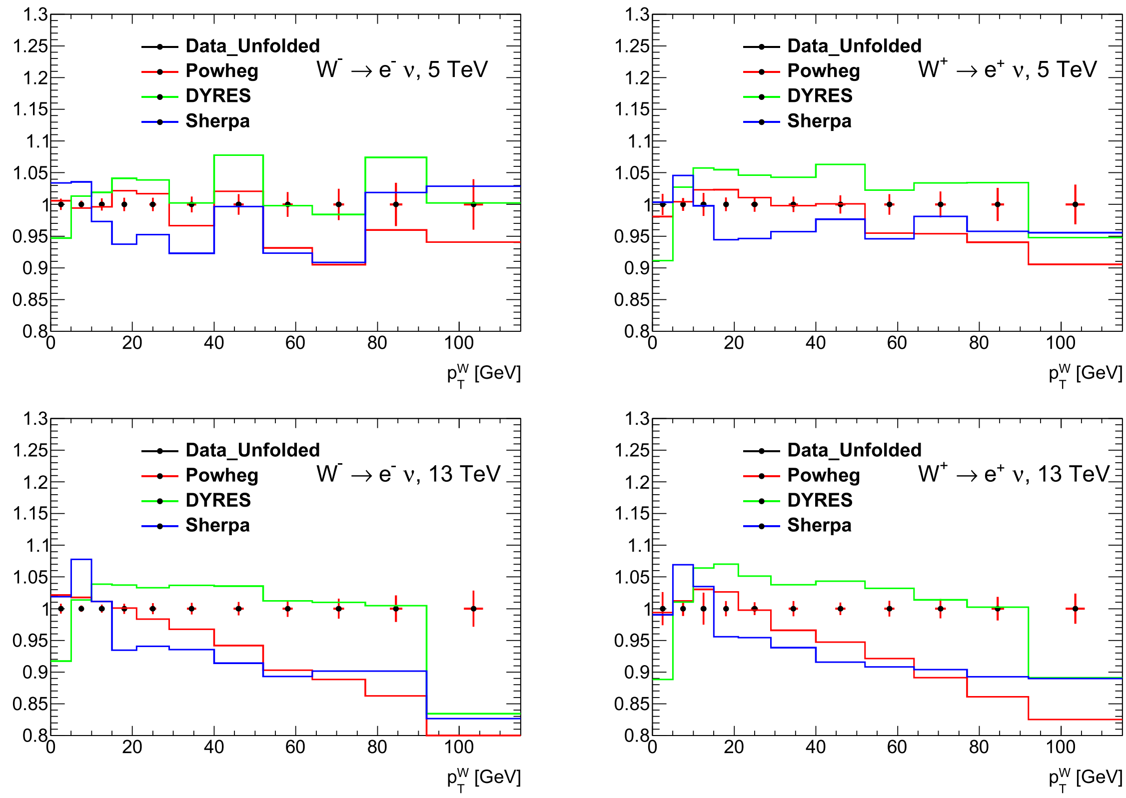


FIGURE 5.16: Unfolded P_{T}^{W} distribution in comparison with various predictions in the W^{-} (left) and W^{+} (right) electron channels, at 5 TeV (top) and 13 TeV (bottom).

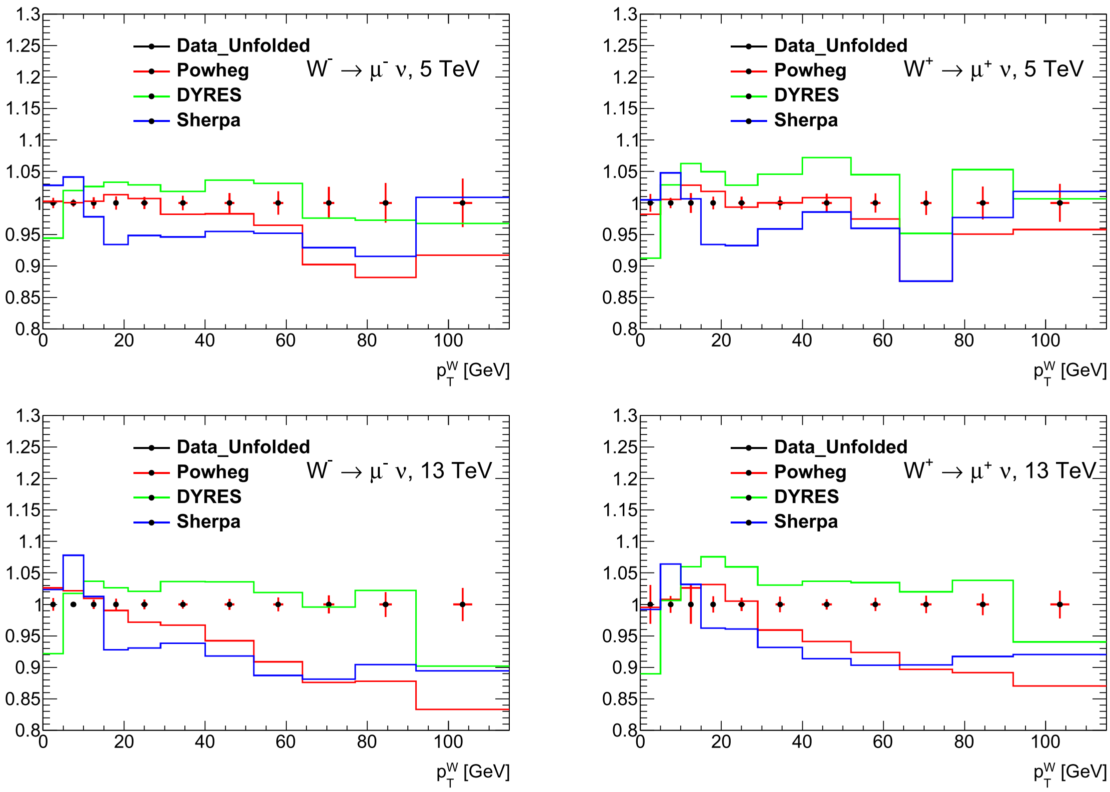


FIGURE 5.17: Unfolded P_T^W distribution in comparison with various predictions in the W^- (left) and W^+ (right) muon channels, at 5 TeV (top) and 13 TeV (bottom).

Chapter 6

W boson production cross sections

6.1 Introduction

The W boson production cross section predictions are available including corrections from QCD (at NNLO in the differential case) and EW (at NLO also in the differential case) [12]. Recently an N3LO computation was performed [64], and a mixed QCD-EW differential computation was also done in [34], see also [62] for an almost complete calculation. Therefore, the measurement of W boson production cross section at the LHC will provide an important test of the SM. Figure 6.1 shows the comparison between the theoretical predictions and measurements from different experiments. The production cross sections are based on p_{T}^{W} distributions described in Chapter 5 with the same selections and corrections, and calculated using two methods: using bin-by-bin correction and using the unfolded distributions. This chapter describes the measurement of the inclusive production cross sections of $W^{\pm} \rightarrow \ell^{\pm}\nu$. The data used correspond to low pile-up runs ($\mu \approx 2$) collected during 2017 and 2018 using proton–proton collisions at $\sqrt{s}$ = 5 TeV and 13 TeV. The bin-by-bin correction is based on a correction factor C extracted from simulation by comparing the truth and reconstruction level, whereas the second method consists of using the unfolded distribution already corrected by the unfolding procedure described in Chapter 4.

6.2 Fiducial cross-section methodology

The fiducial cross section is calculated using the bin-by-bin correction method and compared to the unfolding method, and a brief comparison of the two approaches is shown below.

6.2.1 The bin-by-bin method

The fiducial cross section is calculated from the observed number of events selected in a fiducial phase space after subtracting background contributions and taking into account the detector efficiencies. The resulting fiducial cross section of $W^{\pm}$ for a given channel ($W^{\pm} \rightarrow \ell^{\pm}\nu$) can be expressed with the formula:

$$\sigma_{\mathrm{fid}} = \frac{N^{\mathrm{data}} - N^{\mathrm{bg}}}{\mathcal{L} \cdot C_v} \tag{6.1}$$

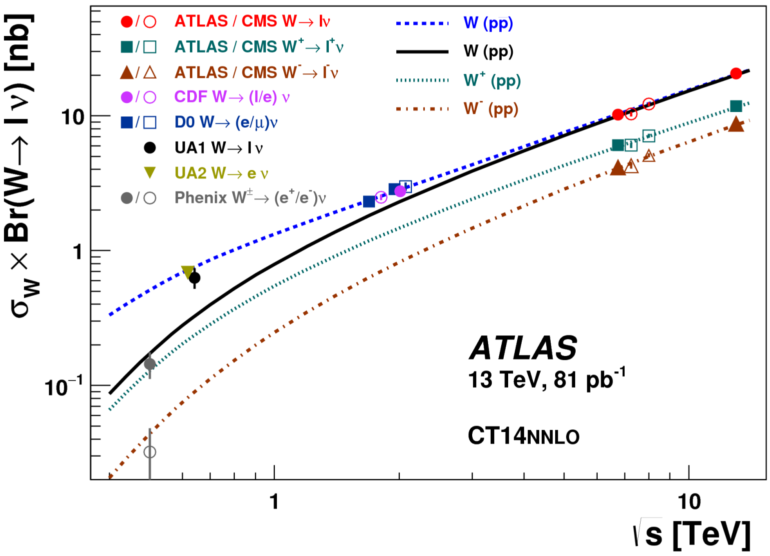


FIGURE 6.1: The measured values of $\sigma_W \times \mathrm{Br}(W \rightarrow \ell)$ for W boson compared to the theoretical predictions based on NNLO QCD calculations [117].

where

- for a given channel, N^{data} and N^{bg} represent the number of events of data in the phase space defined in the section, and the expected number of background events.
- C_v is a correction factor calculated using simulation, corresponding to the ratio of the number of selected events at the detector level and the number of events at the particle level in the fiducial phase space. This correction factor allows to correct the observed difference between data and simulation (due to e.g. reconstruction, identification, isolation, and trigger).
- $\mathcal{L}$ is the integrated luminosity of data.

6.2.2 The Bayesian unfolding method

The second option is to use the unfolding method (the Bayesian unfolding method) defined in the Chapter 4. In general, the idea behind the unfolding is to correct all the detector effects in data distributions, and the total and differential cross sections can be calculated using the unfolded distributions.

For the unfolding approach, the cross section is calculated via the formula:

$$\sigma_{\mathrm{fid}} = \frac{N^{\mathrm{Unfolded}}}{\mathcal{L}} \times \frac{N^{\mathrm{truth\&reco}}}{N^{\mathrm{truth}}} = \frac{N^{\mathrm{Unfolded}}}{\mathcal{L}} \cdot A_{\mathrm{unf}} \tag{6.2}$$

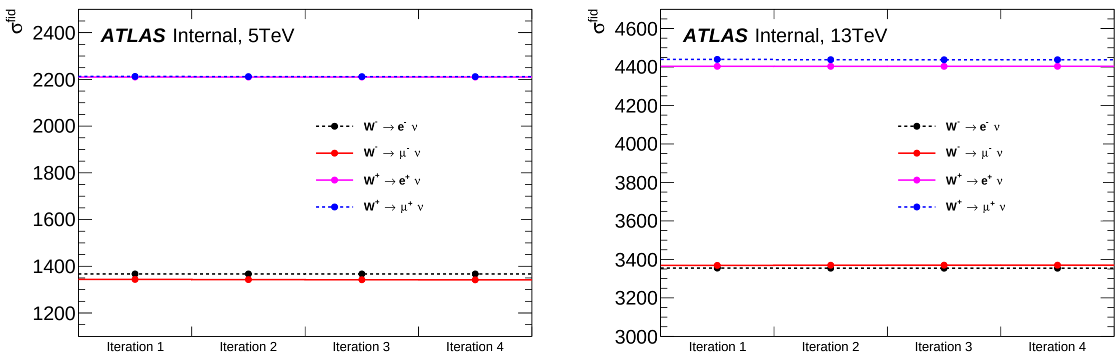
where

- N_{Unfolded} represents the number of events in the unfolded distribution.
- A_{unf} is a correction factor related to the unfolding procedure, defined in the Chapter 4. This factor represents the fraction of events passing reconstructed

and truth selections to the number of events that meet the selection criteria at truth level.

- $\mathcal{L}$ is the integrated luminosity of data.

The unfolding method used in this thesis depends on a regularisation parameter related to the number of iterations (Chapter 4). However as the unfolding does not change the normalisation of the input distributions, the fiducial cross section is independent of the number of iterations. For the different sources of uncertainties (statistical and systematic), the uncertainties are independent of the number of iterations when we take the correlation between bins into account. Also, as the unfolding bias (Chapter 4) depends mainly on the shape of a distribution, when we integer aver all the bins we find no bias. Figure 6.2 shows the fiducial cross sections for different iterations.



(A) (B)

FIGURE 6.2: Fiducial cross section as a function of the number of iterations at 5 TeV (A) and 13 TeV (B).

The propagation of systematic uncertainties using bin-by-bin correction is based on the comparison between the fiducial cross section σ_{fid} and the modified fiducial cross section $\sigma_{\text{fid}}^{\text{var}}$, where:

$$\sigma_{\text{fid}} = \frac{N^{\text{data}} - N^{\text{bg}}}{\mathcal{L}} \times \frac{N^{\text{truth}}}{N^{\text{reco}}}, \tag{6.3}$$

$$\sigma_{\text{fid}}^{\text{var}} = \frac{N^{\text{data}} - (N^{\text{bg}} + \text{var})}{\mathcal{L}} \times \frac{N^{\text{truth}}}{N^{\text{reco}} + \text{var}}. \tag{6.4}$$

The systematic uncertainty can be written as:

$$\text{Systematic} = \frac{\sigma_{\text{fid}}^{\text{var}} - \sigma_{\text{fid}}}{\sigma_{\text{fid}}}. \tag{6.5}$$

For the unfolding procedure, the propagation of systematic uncertainty is done as described in Chapter 5, and the total uncertainty is taken as the sum of all the elements of the unfolding covariance matrix. Good agreement is observed between the two approaches as shown in table 6.1 as example for the $W^+ \to e^+\nu_e$ at 5 TeV.

TABLE 6.1: An example comparison the $W^+ \to e^+\nu_e$ channel at 5 TeV between systematic uncertainties using bin-by-bin correction and the Bayesian unfolding.

	$W^+ \to e^+\nu_e$					
Syst. uncer	Reco SF	Id SF	Iso SF	Trigger SF	e^+ calib	HR calib
Unfolding	0.30%	0.31%	0.33%	0.23%	0.012%	0.08%
Bin-by-bin	0.30%	0.30%	0.33%	0.22%	0.013%	0.08%

6.2.3 Results

The measured fiducial cross sections σ_{fid} for $W^\pm \to \ell^\pm\nu$ are shown in the tables 6.2 and 6.3 with the different sources of uncertainties.

TABLE 6.2: Fiducial cross section with different sources of uncertainties using the bin-by-bin correction and the unfolding approach using 5 TeV samples.

	$\text{W}^- \to e^-\bar{\nu}_e$, 5 TeV, (value $\pm$ stat $\pm$ syst $\pm$ lum) [pb]
σ_{fid} (Unfolding)	$1379 \pm 2.7 \pm 6.4 \pm 22$
σ_{fid} (Bin-by-bin)	$1380 \pm 2.6 \pm 6.3 \pm 22$
	$\text{W}^+ \to e^+\nu_e$, 5 TeV, (value $\pm$ stat $\pm$ syst $\pm$ lum) [pb]
σ_{fid} (Unfolding)	$2227 \pm 3.3 \pm 10 \pm 36$
σ_{fid} (Bin-by-bin)	$2228 \pm 3.4 \pm 10 \pm 36$
	$\text{W}^- \to \mu^-\bar{\nu}_\mu$, 5 TeV, (value $\pm$ stat $\pm$ syst $\pm$ lum) [pb]
σ_{fid} (Unfolding)	$1377 \pm 2.5 \pm 5.6 \pm 22$
σ_{fid} (Bin-by-bin)	$1376 \pm 2.6 \pm 5.5 \pm 22$
	$\text{W}^+ \to \mu^+\nu_\mu$, 5 TeV, (value $\pm$ stat $\pm$ syst $\pm$ lum) [pb]
σ_{fid} (Unfolding)	$2224 \pm 3.3 \pm 8.2 \pm 36$
σ_{fid} (Bin-by-bin)	$2225 \pm 3.3 \pm 8.1 \pm 36$

6.2.4 Comparison with theoretical predictions

Theoretical predictions are calculated for the fiducial cross-sections σ_{fid} using DY-TURBO [45] at NNLO QCD, with different PDF sets: CT18 [93], HERAPDF20 [87], MMHT2014 [91], NNPDF31 [53], ABMP16 [9]. The comparison between measured fiducial cross section and theoretical predictions is shown in Figure 6.3. The uncertainties on the measured σ_{fid} is dominated by the uncertainty on the luminosity, estimated to 1.6% and 1.5% for 5 TeV and 13 TeV, respectively. The CT18 PDF set describes the data best, while the rest of PDFs shows deviation for at least one data set.

TABLE 6.3: Fiducial cross section with different sources of uncertainties using the bin-by-bin correction and the unfolding approach using 13 TeV samples.

	$W^- \to e^-\bar{\nu}_e$, 13 TeV, (value $\pm$ stat $\pm$ syst $\pm$ lum) [pb]
σ_{fid} (Unfolding)	$3445 \pm 3.8 \pm 21 \pm 50$
σ_{fid} (Bin-by-bin)	$3445 \pm 3.8 \pm 20 \pm 50$
	$W^+ \to e^+\nu_e$, 13 TeV, (value $\pm$ stat $\pm$ syst $\pm$ lum) [pb]
σ_{fid} (Unfolding)	$4507 \pm 4.3 \pm 22 \pm 66$
σ_{fid} (Bin-by-bin)	$4505 \pm 4.4 \pm 22 \pm 66$
	$W^- \to \mu^-\bar{\nu}_\mu$, 13 TeV, (value $\pm$ stat $\pm$ syst $\pm$ lum) [pb]
σ_{fid} (Unfolding)	$3444 \pm 3.7 \pm 24 \pm 50$
σ_{fid} (Bin-by-bin)	$3445 \pm 3.8 \pm 25 \pm 50$
	$W^+ \to \mu^+\nu_\mu$, 13 TeV, (value $\pm$ stat $\pm$ syst $\pm$ lum) [pb]
σ_{fid} (Unfolding)	$4504 \pm 4.3 \pm 28 \pm 66$
σ_{fid} (Bin-by-bin)	$4505 \pm 4.3 \pm 28 \pm 66$

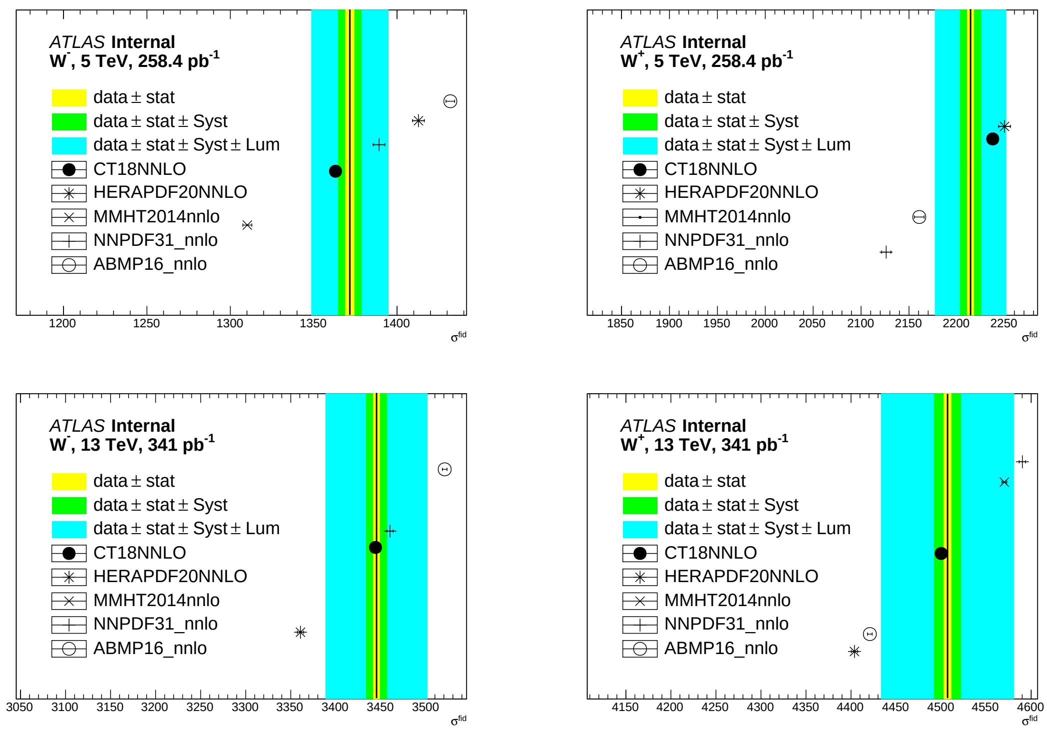


FIGURE 6.3: Measured fiducial cross sections (σ_{fid}) compared to different PDFs set using QCD (NNLO) predictions. The yellow band corresponds to the statistical uncertainty, the middle band to the statistical and experimental systematic uncertainties added in quadrature, while the outer band shows the total uncertainty, including the luminosity uncertainty.

Chapter 7

Measurement of single and double differential cross sections

7.1 Introduction

In this chapter we present detailed studies of the measurement of the differential cross sections of the $W^{\pm}$ boson using the low pile-up runs at $\sqrt{s} = 5$ and 13 TeV, taken in Fall 2017 and July 2018 with the ATLAS detector, corresponding to an integrated luminosity of data of 258 pb^{-1} for $\sqrt{s} = 5$ TeV and 340 pb^{-1} for $\sqrt{s} = 13$ TeV. The data, simulation and all the corrections used in this study are described in Ref. [95]. The differential cross sections are measured in fiducial phase spaces, described in section 4 of Ref. [27], as functions of different variables (η_ℓ, p_{T}^{ℓ}, $\eta_\ell - p_{\mathrm{T}}^{\ell}$) using the unfolded distributions. Different sources of statistical and systematic uncertainties, described in the section 8 of Ref. [27], are propagated via the unfolding procedure. In addition to these sources of uncertainties, there is a bias related to the unfolding procedure, but as the migrations between bins are low for η_ℓ and p_{T}^{ℓ}, the bias in this case is negligible comparing to other sources of uncertainties. The unfolding of data distributions and the propagation of different sources of uncertainties (statistical, systematic and bias) through unfolding, including an optimisation study for the number of iterations needed for the unfolding, are described in Sec. **??**. Section 7.3 shows the results of the differential cross-section measurements and the different sources of uncertainties using the unfolded distributions. In Sec. 7.6, a two dimensional unfolding is used to measure the double differential cross sections in bins of η_ℓ and p_{T}^{ℓ}. A technique is used to transfer the two dimensional unfolding to a one dimensional unfolding as used for differential cross sections of η_ℓ and p_{T}^{ℓ}, separately. All the sources of uncertainties, discussed in Ref. [27], are propagated through unfolding as described in Sec. **??**.

The measurement of differential cross sections in this process provides stringent tests of the QCD theory, and is crucial for a deep understanding and modelling of QCD interactions. Also, the rapidity dependence of the W boson production in the Drell–Yan process provides constraints on the parton distribution functions (PDFs), which are currently the dominant uncertainty source in the W mass measurement (9.2 MeV) [115].

7.2 Data and simulation distributions

7.2.1 Fiducial phase space

The selection of the W candidate events follows the p_{T}^{W} measurement described in Chapter 5. The analysis requires lepton candidates satisfying medium identification criteria based on the EM showers shapes (defined in Rif. [103]). In addition, medium likelihood identification, “ptvarcone20/pT < 0.1” isolation and trigger requirements are applied, trigger requiring the online reconstruction and identification of one lepton passing a p_{T}^{ℓ} threshold of 15 GeV, definitions are shown in Ref. [103]. Candidates within the barrel-end-cap crack $(1.37 < |\eta_\ell| < 1.52)$ are rejected. Also, the selections: $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV and $m_{\mathrm{T}}^{W} > 50$ GeV are applied in order to remove most of the Z-boson and multi-jet backgrounds in the signal phase space.

7.2.2 Experimental systematic uncertainties

In this section, we review the different sources of systematic uncertainty affecting the measurement of the differential cross sections:

Lepton scale factors: As described in Chapter 3, two factors (energy scale and resolution) are applied to data and MC respectively to correct residual difference observed between data and simulation. The energy scale and resolution factors determined from low pile-up runs are applied. The combined effect of all scale and resolution uncertainties on the distributions of η_ℓ, p_{T}^{ℓ} are shown in Figure 7.1. The effect on p_{T}^{ℓ} is up to 2% for large value of p_{T}^{ℓ}, but it’s negligible for η_ℓ.

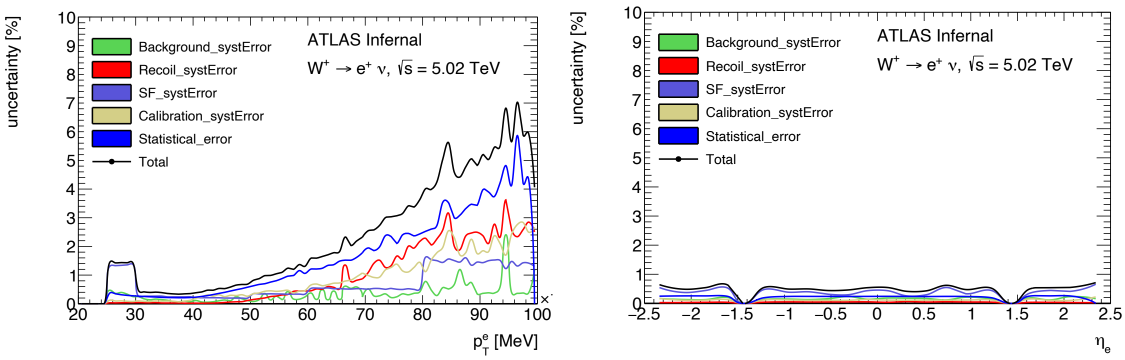


FIGURE 7.1: Uncertainties effect on the distributions of η_ℓ, p_{T}^{ℓ} for $W^{+} \rightarrow e^{+}\nu_e$ at 5 TeV. Uncertainties for 5 and 13 TeV data sets are described in Appendix B.

Lepton selection efficiency: As detailed in Sec. 7.2, selected leptons are required to pass specific criteria. Small differences between data and simulation on the efficiencies of the selections are applied to the simulation as:

$$W_{\text{event}} = \mathrm{SF}_{\text{reco}} \cdot \mathrm{SF}_{\text{ID}} \cdot \mathrm{SF}_{\text{isolation}} \cdot \mathrm{SF}_{\text{trigger}}\,, \tag{7.1}$$

which correspond to the reconstruction, identification, isolation and trigger scale factors (SFs). The SFs are calculated using "tag-and-probe" method detailed in Ref. [22]. The uncertainty on the selection efficiency is found to be the dominant systematic comparing to other source of uncertainties.

Hadronic recoil calibration: Because of the neutrino which can not be measured in the ATLAS detector, the hadronic recoil, defined as the vector sum of all energy deposits excluding energy of lepton, is used in W boson analysis to determine p_{T}^{ν}, p_{T}^{W}, etc. The uncertainty coming from the calibration of the hadronic recoil is related principally to data statistics. This systematic is more important for p_{T}^{ℓ} and is of the order of 2$\%$ for large value of p_{T}^{ℓ}, see Figure 7.1.

Background uncertainty: It is related to the background estimation, in particular to the multi-jet contribution [155], and varies between channels and center-of-mass energies. In general, the background uncertainty is below 0.5$\%$ for our regions of interest.

Luminosity: The luminosity uncertainty for 13 TeV low pile-up runs is 1.5$\%$ for the combination of 2017+2018 data (2.1$\%$ for 2017, 1.5$\%$ for 2018). It is 1.6 $\%$ for 5 TeV low pile-up runs.

7.2.3 Data and MC comparison

The corrections applied during the unfolding are extracted basically from the migration matrix, determined using MC simulation, which connects the particle and detector levels. The idea is that in order to unfold data distribution, the simulation must describe data perfectly. Otherwise, the unfolded data can not be precisely compared to distributions at truth level. More information about objects definitions and all the corrections are described in Section 3 of Ref [27]. Figure 7.2 show the relevant data and MC distributions used for the cross-section measurement.

7.2.4 Unfolding of data distributions

The idea of unfolding is to use a migration matrix built from MC which contains all detector effects and allows us to pass from reconstruction to truth level. As detailed in Section 4, the unfolding is done to correct all detector effects. Contrary to the p_{T}^{W} unfolding described in Ref [27], the η_{ℓ} or p_{T}^{ℓ} unfolding is easier because of the small migration between bins, due to a negligible difference between truth and reconstructed levels (less detector effects), which means that the migration matrix is more diagonal, see Figure 7.3.
The same unfolding method used for p_{T}^{W} is used also for η_{ℓ}, p_{T}^{ℓ} unfolding, the iterative Bayesian unfolding method [55]. Figure 7.4 shows an example of distributions at the unfolding and reconstructed level using 3 iterations. Because of the small migration between bins, the unfolded level distribution is identical to the truth level distribution.

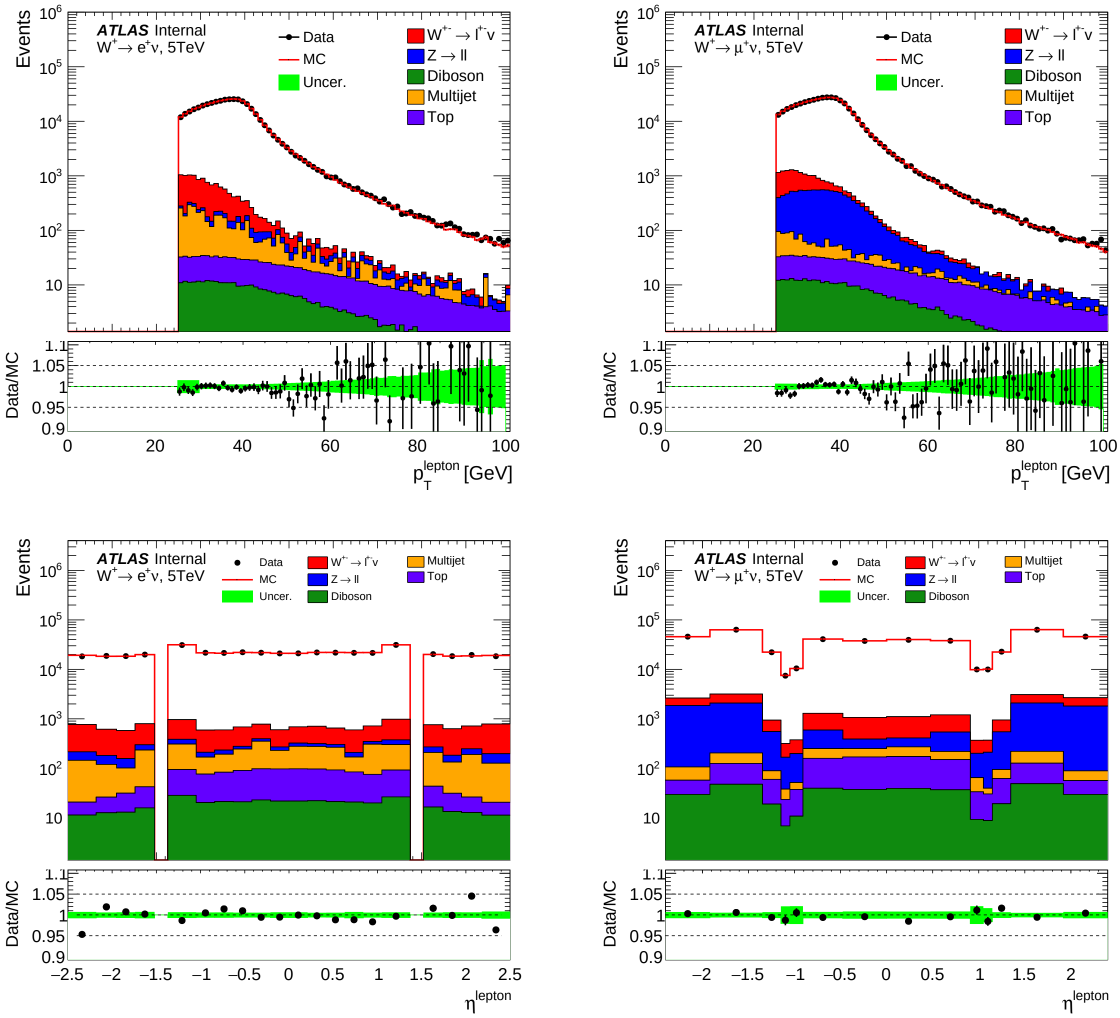


FIGURE 7.2: Example distributions of the observables p_{T}^{ℓ} (top) and η_{ℓ} (bottom) chosen to be unfolded for W^{+} in the electron (left) and muon (right) channels at 5 TeV in the fiducial phase space. The signal and background are normalised to data. The low panel gives the ratio Data/MC in each bin. The green band shows the statistical and systematic uncertainties. Control plots for other channels are shown in Appendix A.

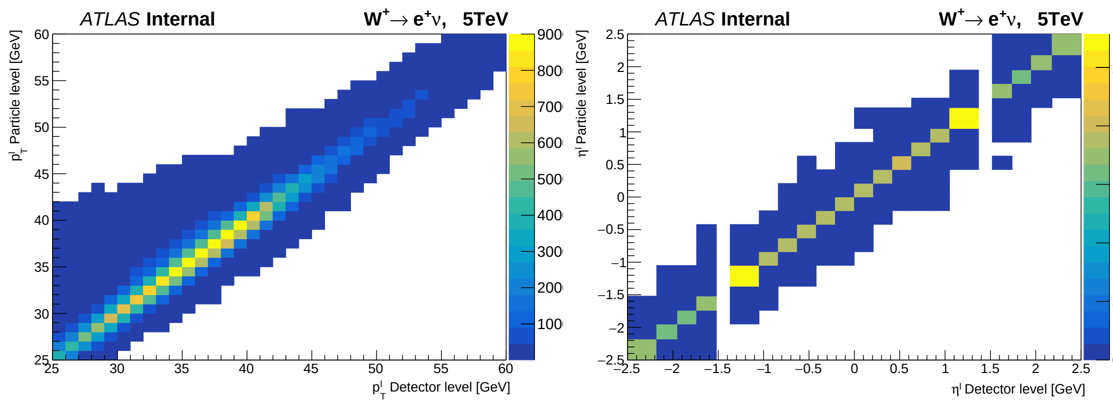


FIGURE 7.3: Example of the migration matrix used in the unfolding of p_{T}^{ℓ} (left) and η_{ℓ} (right) for $W^+ \to e^+\nu_e$ at 5 TeV, the migration matrix is quasi diagonal because of the small difference between particle and detector levels.

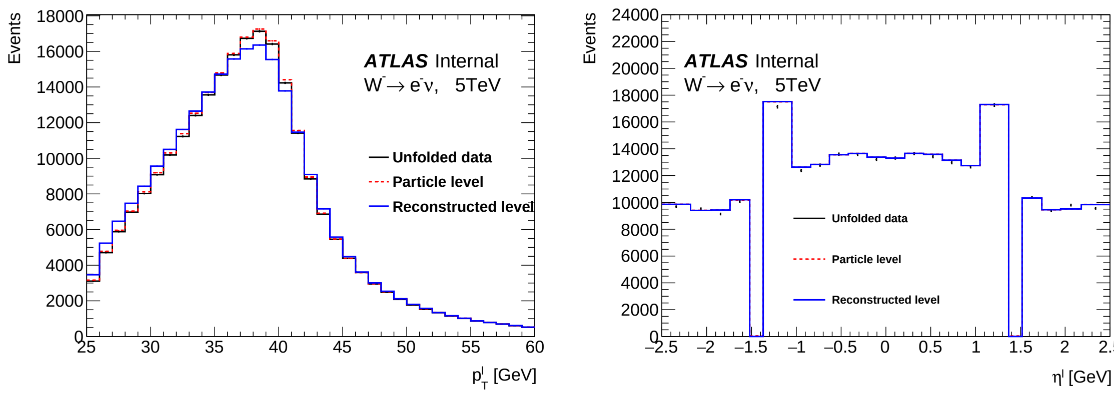


FIGURE 7.4: Example of the unfolded distributions for p_{T}^{ℓ} (left) and η_{ℓ} (right) for $W^- \to e^-\bar{\nu}_e$ at 5 TeV at detector level, particle level and unfolded level.

7.2.5 Propagation of the statistical and systematic uncertainties

The propagation of uncertainties through the unfolding is done in the same way as for p_{T}^{W}, as detailed in Chapter 4. The main difference comes from the degree of migration between bins. Figure 7.5 shows an example of statistical uncertainty at the unfolding level, comparing to η_ℓ, p_{T}^{ℓ} is characterised with slightly larger migration between bins which explains the increase in the statistical error with the number of iterations.

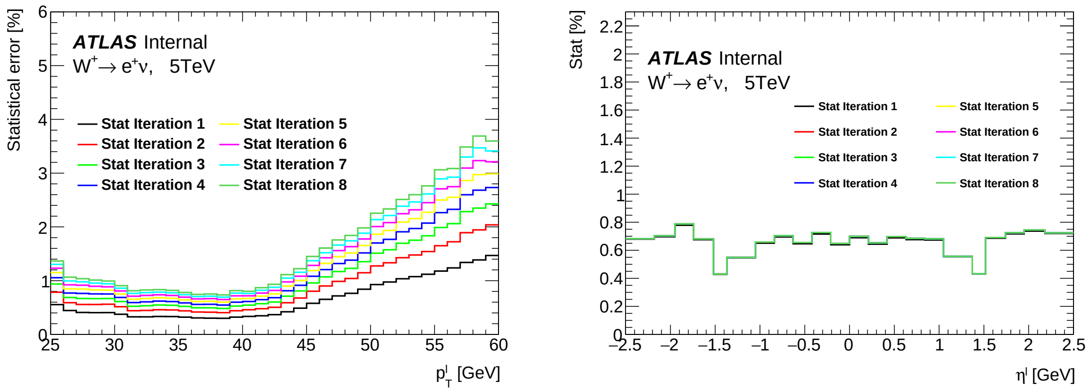


FIGURE 7.5: Example of the statistical uncertainties for p_{T}^{ℓ} (left) and η_ℓ (right), for $W^+ \to e^+\nu_e$ at 5 TeV. Statistical error increases with the number of iterations because of the migration between bins. Statistical uncertainties with their correlation matrices are described in Appendix B.

Contrary to the statistical uncertainty, the systematic uncertainties are more stable with the number of iterations, Figure 7.6 shows an example of the dominant systematic uncertainty at the unfolded level as a function of the number of iterations. For η_ℓ and p_{T}^{ℓ}, the total uncertainty is dominated by the uncertainty from the efficiency scale factors.

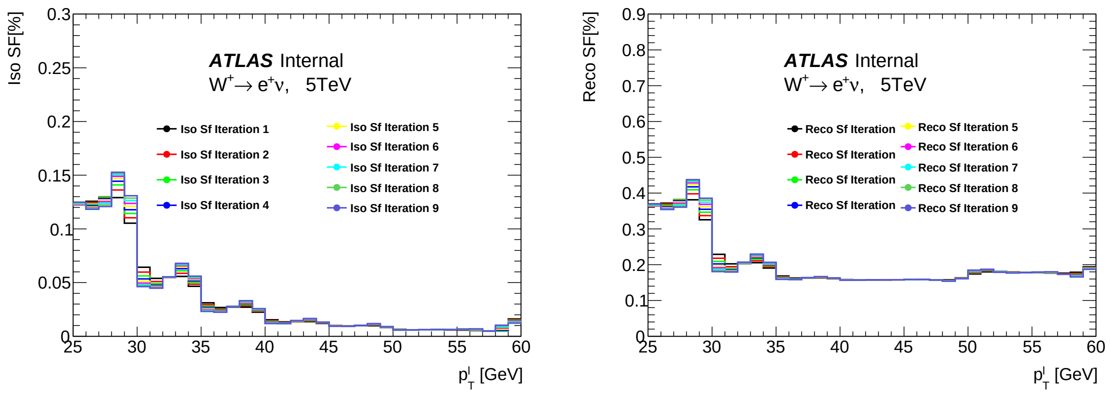


FIGURE 7.6: Example of the systematic uncertainties (isolation (left) and reconstruction (right) SFs) for p_{T}^{ℓ} for $W^+ \to e^+\nu_e$ at 5 TeV.

7.2.6 Unfolding bias

As detailed in Chapter 4, the unfolding method used in this thesis introduces a bias that should not be dominant. The bias is calculated as explained in Chapter 4. The procedure to estimate the bias can be summarised in two steps (Ref [111]):

- Reweight the MC distribution at truth level with the fitted ratio data/MC, in such a way that the corresponding reconstructed distribution, obtained by the truth level reweighted distribution, matches better the data distribution after the background subtraction.

- The bias is estimated as the difference between the unfolded distribution of the reconstruction-weighted distribution and the truth-weighted distribution.

The procedure used to calculate the bias of unfolding is illustrated in Fig. **??**. The reconstruction-weighted distribution must be closer to data compared to the original reconstructed distribution (Fig. 7.7).

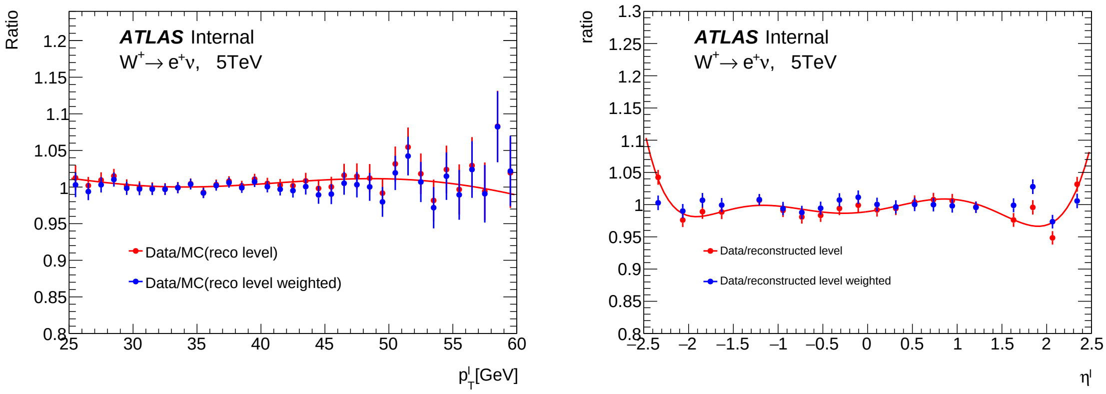


FIGURE 7.7: The ratio of data over the original reconstructed MC distributions for p_{T}^{ℓ} (left) and η_{ℓ} (right) compared to the ratio of data over the weighted one. The latter is in better agreement with data (background subtracted).

As the unfolding does not change the normalisation of input distributions, the total unfolding bias when we take the correlation (anti-correlation) between bins into account must be equal to 0. Contrary to other sources of uncertainties, the bias decreases with the number of iterations and the anti-correlation between bins increases with the number of iterations to ensure that the bias integrated in all bins is zero (Fig. 7.8).

7.2.7 Optimisation of the number of iterations in iterative Bayesian unfolding

As discussed above, the statistical uncertainty increases with the number of iterations, while the unfolding bias decreases with them. Therefore it is possible to minimise the total uncertainty by optimising the number of iterations. As the bias

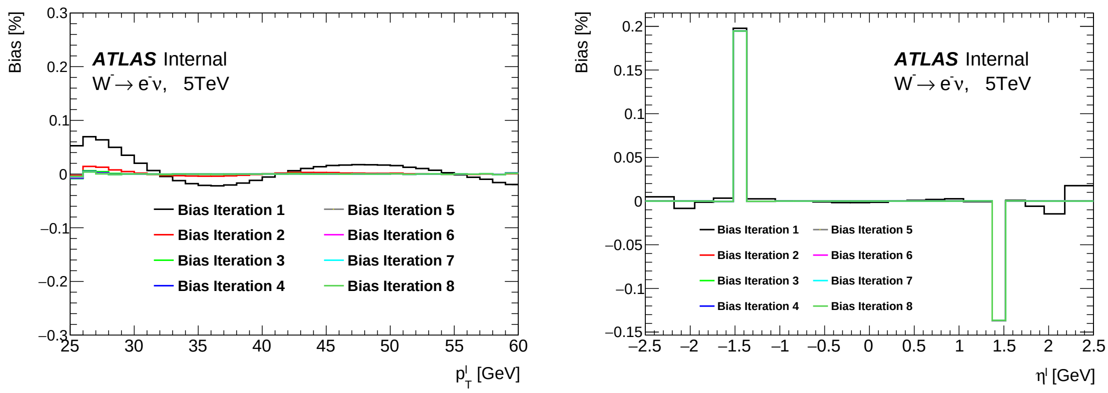


FIGURE 7.8: Example of the unfolding bias of p_{T}^{ℓ} (left) and η_{ℓ} (right) as a function of the number of iterations used in the unfolding for W^- at 5 TeV. After the second iteration, the bias is negligible compared to other sources of uncertainty. The unfolding bias for other channels is shown in Appendix B.

is very small comparing to other sources of uncertainties, the best choice is to use the first iteration. However to avoid the fluctuations in the bias as shown in Figure 7.9, the second iteration is used. As we are interested in the differential cross sections, the optimisation study is done for each bin separately around the peak region.

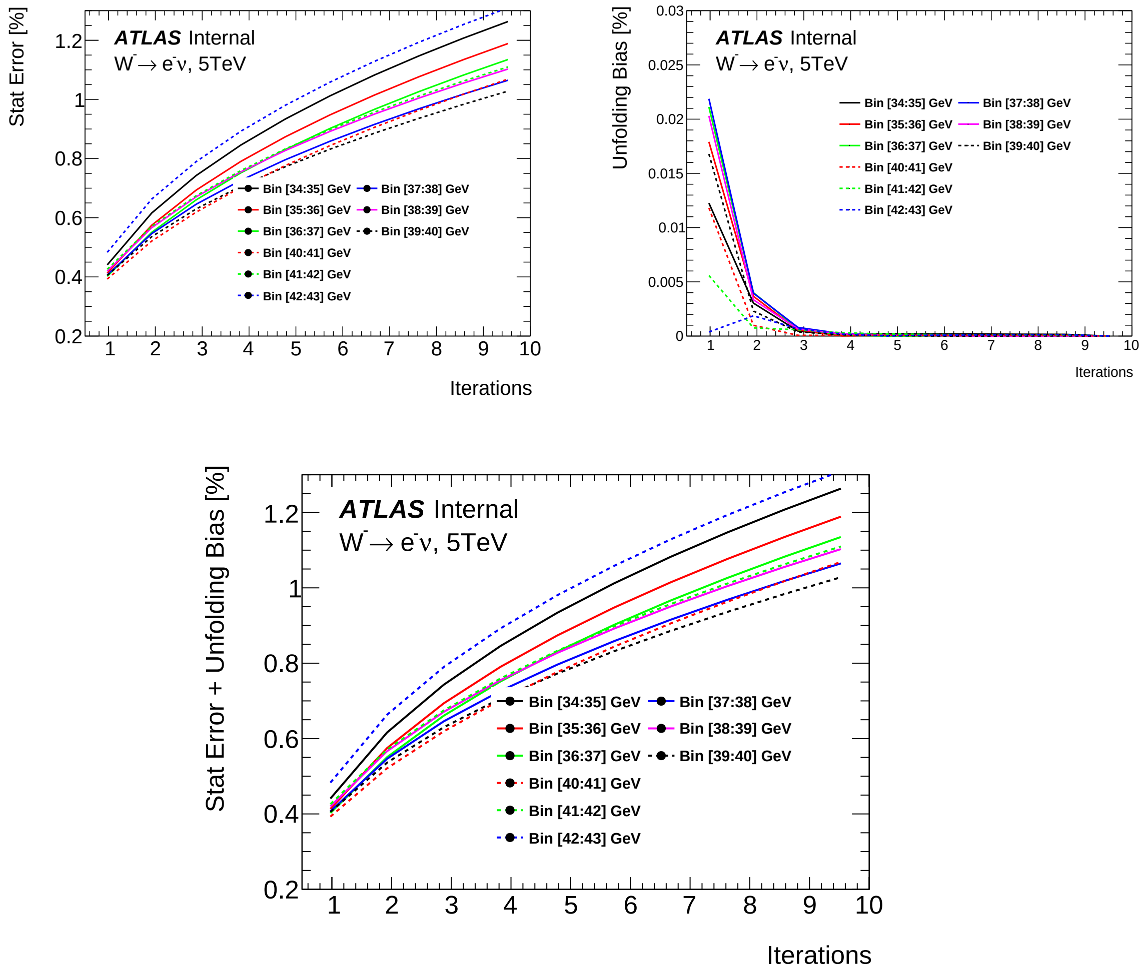


FIGURE 7.9: Example of the statistical uncertainty (top left) and the unfolding bias uncertainty (top right) and their combined uncertainty (bottom) for a few selected bins in p_{T}^{ℓ} as a function of the number of iterations used in the unfolding.

7.3 Differential cross sections

The differential cross sections can be estimated using a correction factor calculated from simulation, the bin-by-bin unfolding, where the differential cross-section formula can be expressed as:

$$\frac{d\sigma_i}{dx^i} = \frac{N^i_{\text{data}}}{\Delta x^i \mathcal{L}} \cdot C_i = \frac{N^i_{\text{data}}}{\Delta x^i \mathcal{L}} \cdot \frac{N^i_{\text{truth}}}{N^i_{\text{reco}}}, \tag{7.2}$$

where Δx^i is the bin width, and N^i is the number of events in bin i. On the other hand, there is another option to calculate the differential cross section, replacing the correction bin-by-bin factor C_i, by the unfolding of the data distribution using the inverse of the migration matrix M_{ij}. The new formula using the unfolded distribution of data is expressed as:

$$\frac{d\sigma_i}{dx^i} = \frac{N^i_{\text{Unf}}}{\Delta x^i \mathcal{L}} \cdot \frac{1}{A_c} = \frac{1}{\Delta x^i \mathcal{L}} \cdot \Sigma_j M^{-1}_{ij} \left(N^j_{\text{reco}} - N^j_{\text{reco,bkg}} \right) \cdot \frac{1}{A_c}, \tag{7.3}$$

where Δx^i is the bin width, N^i_{Unf} is the number of events in the unfolded distribution, A_c is the acceptance correction, used to correct the unfolded distribution and take into account the events that pass the detector-level selection but fail the particle-level selection. Figure 7.10 shows an example of the acceptance correction for p_{T}^{ℓ} and η_ℓ at 5 TeV.

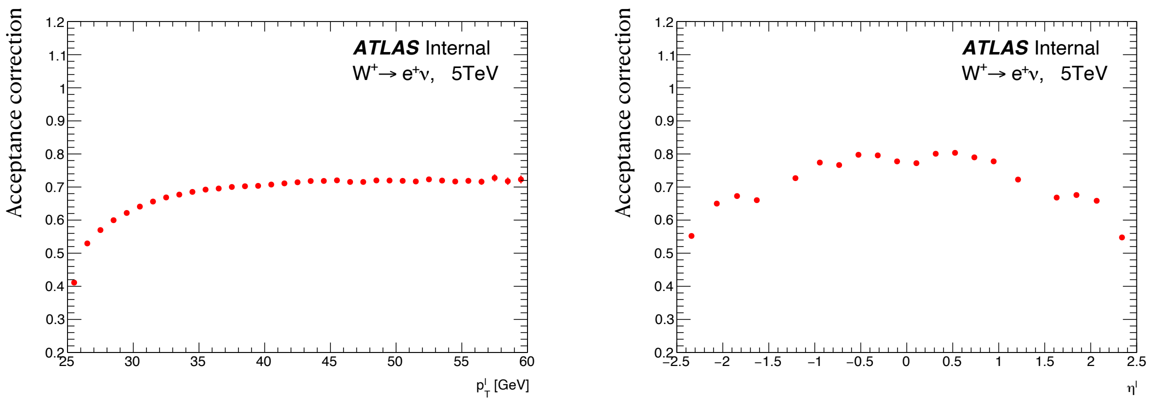


FIGURE 7.10: Fraction of events that pass the detector-level selection but fail the particle-level selection bin-by-bin for p_{T}^{ℓ} (left) and η_ℓ (right).

Differential cross sections in the e channel versus p_{T}^{e} at 5 TeV

TABLE 7.1: Differential cross sections versus p_{T}^{e} for 5 TeV (W^-, e^-). The columns show the bin range, the measured cross section and the corresponding relative uncertainties.

	$W^- \to e^-\bar{\nu}_e$, 5 TeV, uncertainties in (%)			
Range	$d\sigma/dp_{\mathrm{T}}^{e}$ [pb/GeV]	Stat uncertainty	Unfolding bias	Syst uncertainty
[26, 27]	18.300	0.895	0.014	1.583
[27, 28]	22.885	0.802	0.013	1.674
[28, 29]	27.240	0.786	0.008	1.783
[29, 30]	31.219	0.747	0.005	1.344
[30, 31]	35.345	0.706	0.002	0.425
[31, 32]	39.692	0.663	0.001	0.310
[32, 33]	43.592	0.653	0.003	0.315
[33, 34]	48.299	0.616	0.003	0.323
[34, 35]	52.746	0.575	0.004	0.313
[35, 36]	57.124	0.551	0.004	0.286
[36, 37]	61.673	0.546	0.004	0.276
[37, 38]	65.344	0.567	0.003	0.280
[38, 39]	66.728	0.536	0.002	0.275
[39, 40]	63.810	0.521	0.001	0.276
[40, 41]	55.319	0.571	0.001	0.292
[41, 42]	44.373	0.663	0.002	0.344
[42, 43]	34.381	0.705	0.003	0.379
[43, 44]	26.705	0.759	0.003	0.408
[44, 45]	21.292	0.828	0.003	0.385

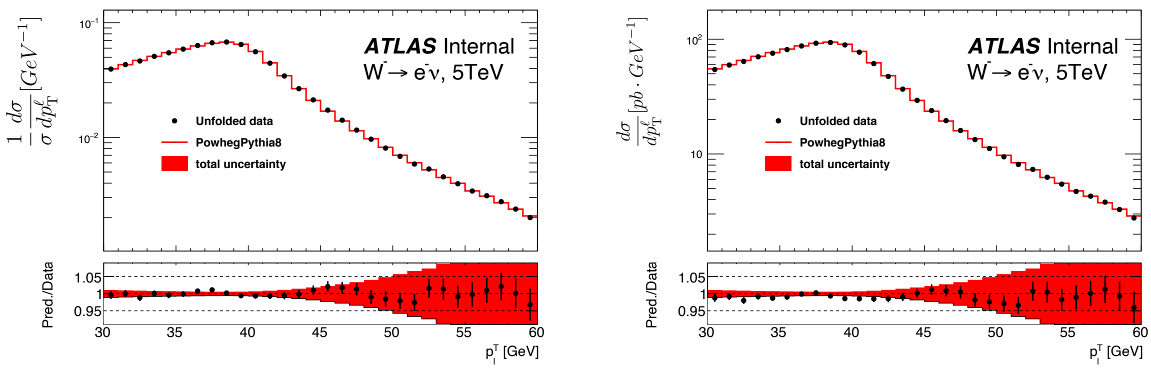


FIGURE 7.11: Differential cross sections (left) and normalised differential cross sections (right) as a function of p_{T}^{e} for 5 TeV (W^-, e^-). The bottom panel shows the ratio data to MC (Powheg+Pethia8) together with the red band showing the total uncertainty.

TABLE 7.2: Differential cross sections versus p_T^e for 5 TeV (W^+, e^+). The columns show the bin range, the measured cross section and the corresponding relative uncertainties.

	$W^+ \to e^+\nu_e$, 5 TeV, uncertainties in (%)			
Range	$d\sigma/dp_T^e$ [pb/GeV]	Stat uncertainty	Unfolding bias	Syst uncertainty
[26, 27]	36.907	0.590	0.019	1.536
[27, 28]	44.556	0.557	0.008	1.644
[28, 29]	51.628	0.557	0.010	1.768
[29, 30]	58.724	0.561	0.009	1.308
[30, 31]	65.036	0.515	0.008	0.389
[31, 32]	71.340	0.441	0.007	0.275
[32, 33]	76.950	0.446	0.008	0.281
[33, 34]	81.804	0.459	0.007	0.297
[34, 35]	87.183	0.455	0.007	0.284
[35, 36]	92.366	0.439	0.007	0.261
[36, 37]	95.698	0.417	0.006	0.258
[37, 38]	98.590	0.413	0.005	0.261
[38, 39]	98.572	0.405	0.001	0.260
[39, 40]	92.989	0.440	0.002	0.253
[40, 41]	79.972	0.459	0.006	0.249
[41, 42]	63.485	0.478	0.008	0.275
[42, 43]	48.760	0.504	0.008	0.294
[43, 44]	37.878	0.589	0.009	0.309
[44, 45]	29.757	0.677	0.006	0.287

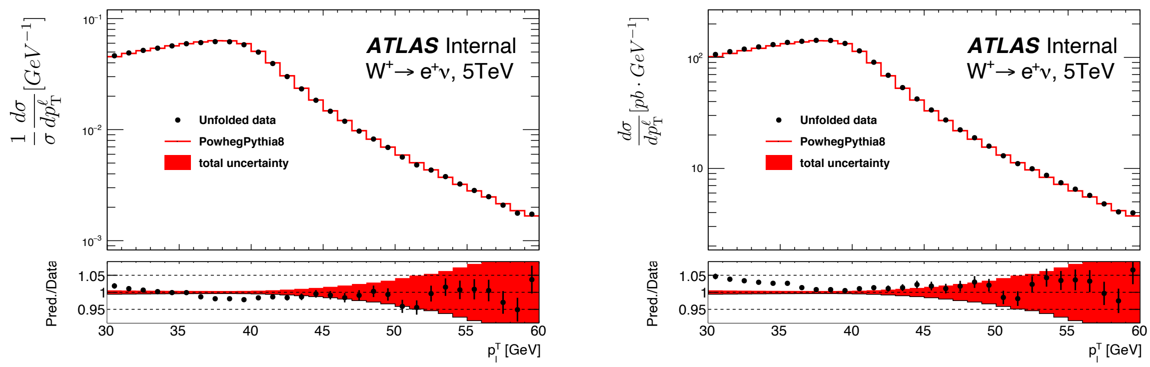


FIGURE 7.12: Differential cross sections (right) and normalised differential cross sections (left) as a function of p_T^e for 5 TeV (W^+, e^+). The bottom panel shows the ratio data to MC (Powheg+Pethia8) together with the red band showing the total uncertainty.

Differential cross sections of the μ channel versus p_{T}^{μ} at 5 TeV

TABLE 7.3: Differential cross sections versus p_{T}^{μ} for 5 TeV (W^{-}, μ^{-}). The columns show the bin range, the measured cross section and the corresponding relative uncertainties.

	$W^{-} \to \mu^{-}\bar{\nu}_{\mu}$, 5 TeV, uncertainties in (%)			
Range	$d\sigma/dp_{\mathrm{T}}^{\mu}$ [pb/GeV]	Stat uncertainty	Unfolding bias	Syst uncertainty
[26, 27]	20.256	0.928	0.010	0.732
[27, 28]	24.586	0.850	0.023	0.764
[28, 29]	28.525	0.770	0.015	0.846
[29, 30]	32.915	0.782	0.011	0.752
[30, 31]	36.989	0.696	0.007	0.542
[31, 32]	41.858	0.638	0.006	0.650
[32, 33]	46.520	0.645	0.004	0.644
[33, 34]	50.704	0.592	0.002	0.676
[34, 35]	55.074	0.557	0.000	0.586
[35, 36]	59.439	0.559	0.001	0.469
[36, 37]	63.851	0.503	0.001	0.564
[37, 38]	67.492	0.499	0.002	0.580
[38, 39]	69.177	0.508	0.002	0.579
[39, 40]	66.828	0.510	0.000	0.498
[40, 41]	58.303	0.531	0.002	0.467
[41, 42]	46.584	0.562	0.005	0.550
[42, 43]	35.949	0.629	0.005	0.563
[43, 44]	27.695	0.736	0.004	0.548
[44, 45]	22.101	0.809	0.004	0.494

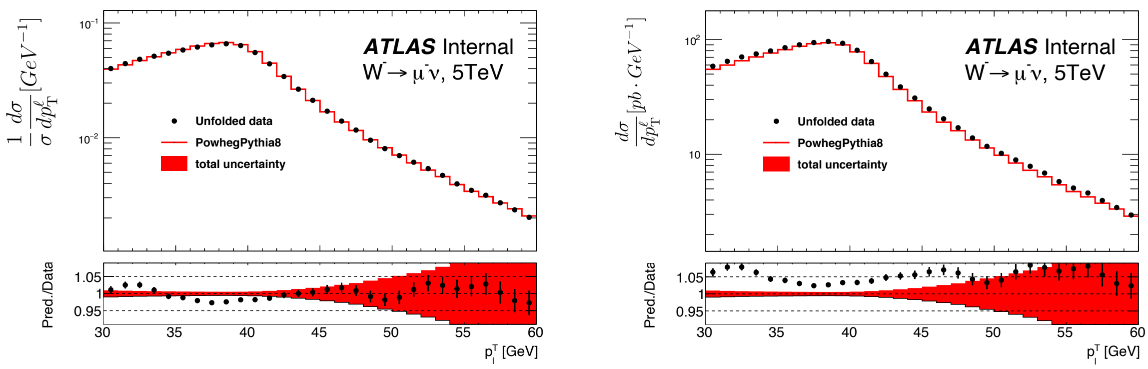


FIGURE 7.13: Differential cross sections (right) and normalised differential cross sections (left) as a function of p_{T}^{μ} for 5 TeV (W^{-}, μ^{-}). The bottom panel shows the ratio data to MC (Powheg+Pethia8) together with the red band showing the total uncertainty.

TABLE 7.4: Differential cross sections versus p_{T}^{μ} for 5 TeV (W^{+}, μ^{+}). The columns show the bin range, the measured cross section and the corresponding relative uncertainties.

	$W^{+} \to \mu^{+}\nu_{\mu}$, 5 TeV, uncertainties in (%)			
Range	$d\sigma/dp_{\mathrm{T}}^{\mu}$ [pb/GeV]	Stat uncertainty	Unfolding bias	Syst uncertainty
[26, 27]	40.943	0.613	0.025	0.707
[27, 28]	49.223	0.592	0.014	0.734
[28, 29]	56.081	0.572	0.013	0.814
[29, 30]	62.909	0.555	0.006	0.713
[30, 31]	69.407	0.518	0.001	0.521
[31, 32]	75.895	0.483	0.001	0.632
[32, 33]	81.579	0.458	0.006	0.634
[33, 34]	87.212	0.439	0.009	0.660
[34, 35]	92.731	0.435	0.013	0.562
[35, 36]	98.205	0.420	0.014	0.457
[36, 37]	102.257	0.414	0.017	0.547
[37, 38]	104.721	0.411	0.017	0.565
[38, 39]	104.533	0.405	0.014	0.567
[39, 40]	97.729	0.432	0.008	0.493
[40, 41]	83.991	0.457	0.002	0.464
[41, 42]	66.658	0.479	0.012	0.539
[42, 43]	51.638	0.534	0.018	0.546
[43, 44]	39.754	0.619	0.020	0.526
[44, 45]	30.943	0.718	0.017	0.476

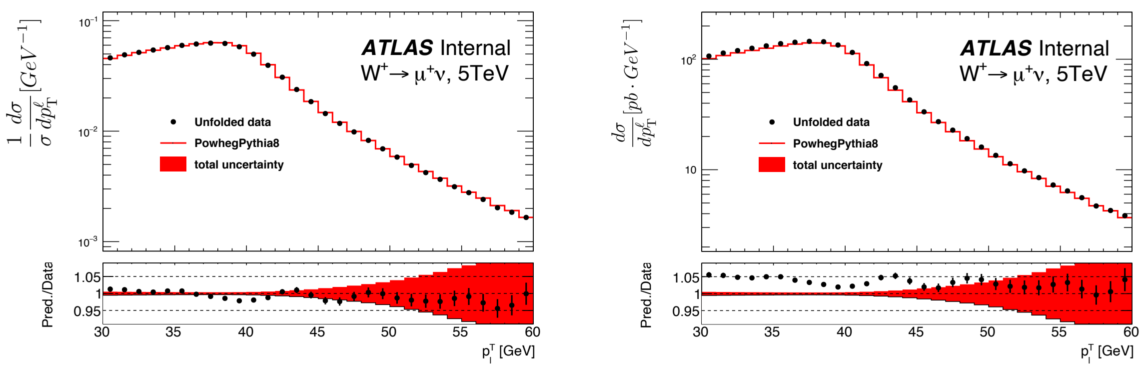


FIGURE 7.14: Differential cross sections (right) and normalised differential cross sections (left) as a function of p_{T}^{μ} for 5 TeV (W^{+}, μ^{+}). The bottom panel shows the ratio data to MC (Powheg+Pethia8) together with the red band showing the total uncertainty.

Differential cross sections of the e channel versus p_T^e at 13 TeV

TABLE 7.5: Differential cross sections versus p_T^e for 13 TeV (W^-, e^-). The columns show the bin range, the measured cross section and the corresponding relative uncertainties.

	$W^- \to e^-\bar{\nu}_e$, 13 TeV, uncertainties in (%)			
Range	$d\sigma/dp_T^e$ [pb/GeV]	Stat uncertainty	Unfolding bias	Syst uncertainty
[26, 27]	48.398	0.448	0.058	0.888
[27, 28]	59.710	0.442	0.026	0.942
[28, 29]	69.166	0.406	0.012	0.997
[29, 30]	78.766	0.376	0.034	0.758
[30, 31]	88.430	0.371	0.041	0.468
[31, 32]	97.272	0.370	0.039	0.501
[32, 33]	106.339	0.360	0.031	0.523
[33, 34]	115.313	0.331	0.019	0.538
[34, 35]	124.252	0.317	0.006	0.418
[35, 36]	133.374	0.312	0.007	0.258
[36, 37]	141.726	0.294	0.020	0.251
[37, 38]	146.909	0.280	0.029	0.256
[38, 39]	148.526	0.273	0.033	0.259
[39, 40]	141.766	0.283	0.030	0.246
[40, 41]	125.030	0.305	0.021	0.230
[41, 42]	104.001	0.351	0.007	0.237
[42, 43]	83.848	0.375	0.006	0.252
[43, 44]	67.814	0.394	0.015	0.263
[44, 45]	55.272	0.458	0.023	0.273

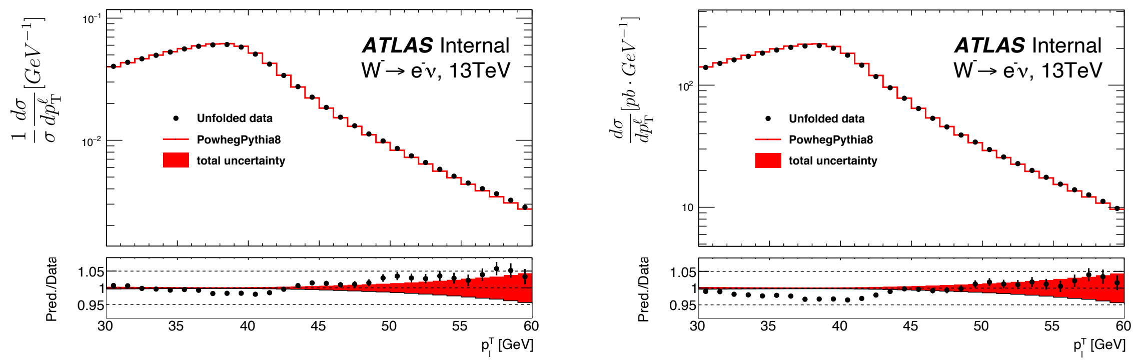


FIGURE 7.15: Differential cross sections (right) and normalised differential cross sections (left) as a function of p_T^e for 13 TeV (W^-, e^-). The bottom panel shows the ratio data to MC (Powheg+Pethia8) together with the red band showing the total uncertainty.

TABLE 7.6: Differential cross sections versus p_{T}^{e} for 13 TeV (W^+, e^+). The columns show the bin ranges, the measured cross section and the corresponding relative uncertainties.

	$W^+ \to e^+\nu_e$, 13 TeV, uncertainties in (%)			
Range	$d\sigma/dp_{\mathrm{T}}^{e}$ [pb/GeV]	Stat uncertainty	Unfolding bias	Syst uncertainty
[26, 27]	73.773	0.372	0.106	0.879
[27, 28]	89.046	0.354	0.123	0.932
[28, 29]	101.290	0.328	0.051	0.988
[29, 30]	112.455	0.316	0.006	0.750
[30, 31]	124.343	0.326	0.035	0.453
[31, 32]	135.555	0.296	0.044	0.481
[32, 33]	145.498	0.293	0.040	0.499
[33, 34]	154.658	0.292	0.029	0.510
[34, 35]	163.823	0.281	0.016	0.398
[35, 36]	172.351	0.276	0.001	0.254
[36, 37]	179.925	0.268	0.016	0.248
[37, 38]	183.698	0.265	0.028	0.252
[38, 39]	182.920	0.270	0.036	0.256
[39, 40]	172.858	0.268	0.034	0.243
[40, 41]	151.985	0.297	0.025	0.226
[41, 42]	125.704	0.317	0.011	0.230
[42, 43]	100.523	0.308	0.002	0.240
[43, 44]	81.060	0.339	0.015	0.251
[44, 45]	65.895	0.395	0.024	0.253

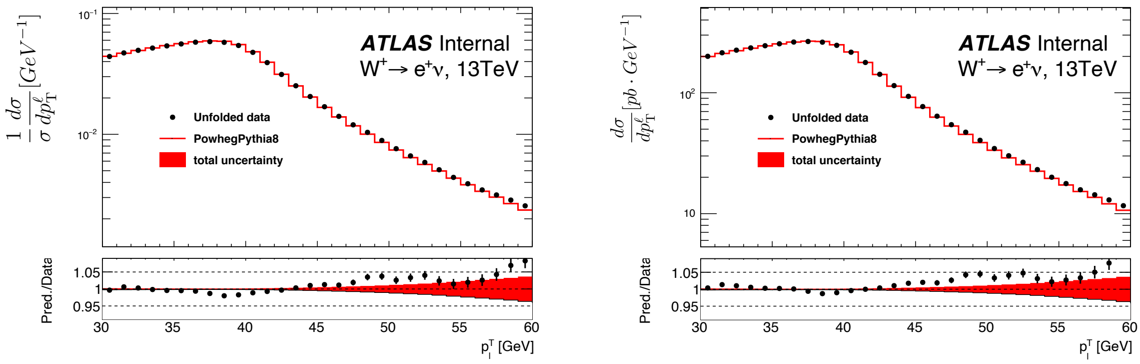


FIGURE 7.16: Differential cross sections (right) and normalised differential cross sections (left) as a function of p_{T}^{e} for 13 TeV (W^+, e^+). The bottom panel shows the ratio data to MC (Powheg+Pethia8) together with the red band showing the total uncertainty.

Differential cross sections of the μ channel versus p_{T}^{μ} at 13 TeV

TABLE 7.7: Differential cross sections versus p_{T}^{μ} for 13 TeV (W^-, μ^-). The columns show the bin range, the measured cross section and the corresponding relative uncertainties.

	$\mathrm{W}^- \to \mu^-\bar{\nu}_\mu$, 13 TeV, uncertainties in (%)			
Range	$d\sigma/dp_{\mathrm{T}}^{\mu}$ [pb/GeV]	Stat uncertainty	Unfolding bias	Syst uncertainty
[26, 27]	52.018	0.476	0.066	0.392
[27, 28]	62.789	0.464	0.085	0.400
[28, 29]	72.624	0.450	0.056	0.431
[29, 30]	81.235	0.390	0.037	0.393
[30, 31]	90.532	0.395	0.023	0.317
[31, 32]	100.091	0.373	0.011	0.360
[32, 33]	109.480	0.328	0.001	0.361
[33, 34]	119.185	0.314	0.009	0.372
[34, 35]	127.754	0.306	0.017	0.340
[35, 36]	136.933	0.291	0.024	0.305
[36, 37]	144.898	0.272	0.029	0.340
[37, 38]	150.402	0.273	0.030	0.347
[38, 39]	152.265	0.289	0.024	0.347
[39, 40]	146.020	0.295	0.011	0.320
[40, 41]	128.869	0.307	0.004	0.311
[41, 42]	107.394	0.334	0.020	0.341
[42, 43]	86.287	0.370	0.026	0.348
[43, 44]	69.418	0.384	0.033	0.346
[44, 45]	56.634	0.403	0.032	0.334

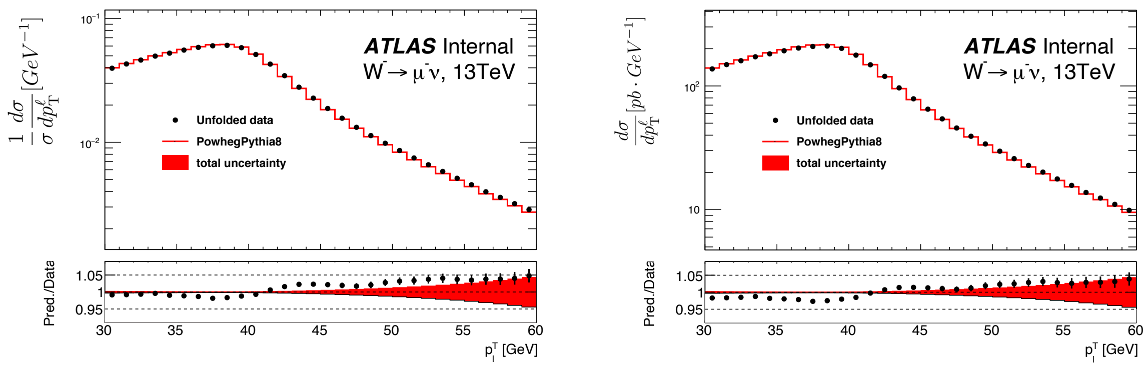


FIGURE 7.17: Differential cross sections (right) and normalised differential cross sections (left) as a function of p_{T}^{μ} for 13 TeV (W^-, μ^-). The bottom panel shows the ratio data to MC (Powheg+Pethia8) together with the red band showing the total uncertainty.

TABLE 7.8: Differential cross sections versus p_{T}^{μ} for 13 TeV (W^{+}, μ^{+}). The columns show the bin range, the measured cross sections and the corresponding relative uncertainties.

	$\mathrm{W}^{+} \rightarrow \mu^{+}\nu_{\mu}$, 13 TeV, uncertainties in (%)			
Range	$d\sigma/dp_{\mathrm{T}}^{\mu}$ [pb/GeV]	Stat uncertainty	Unfolding bias	Syst uncertainty
[26, 27]	80.886	0.359	0.030	0.374
[27, 28]	96.188	0.390	0.007	0.383
[28, 29]	108.856	0.382	0.010	0.411
[29, 30]	120.804	0.352	0.020	0.373
[30, 31]	131.932	0.317	0.024	0.305
[31, 32]	142.772	0.310	0.023	0.345
[32, 33]	152.768	0.294	0.020	0.348
[33, 34]	162.641	0.287	0.014	0.359
[34, 35]	172.353	0.269	0.006	0.325
[35, 36]	180.673	0.265	0.002	0.294
[36, 37]	187.665	0.265	0.011	0.327
[37, 38]	190.965	0.257	0.020	0.336
[38, 39]	190.346	0.262	0.026	0.337
[39, 40]	179.390	0.260	0.025	0.311
[40, 41]	157.157	0.278	0.020	0.305
[41, 42]	128.185	0.303	0.010	0.335
[42, 43]	102.682	0.326	0.000	0.341
[43, 44]	82.328	0.367	0.008	0.335
[44, 45]	66.918	0.413	0.015	0.318

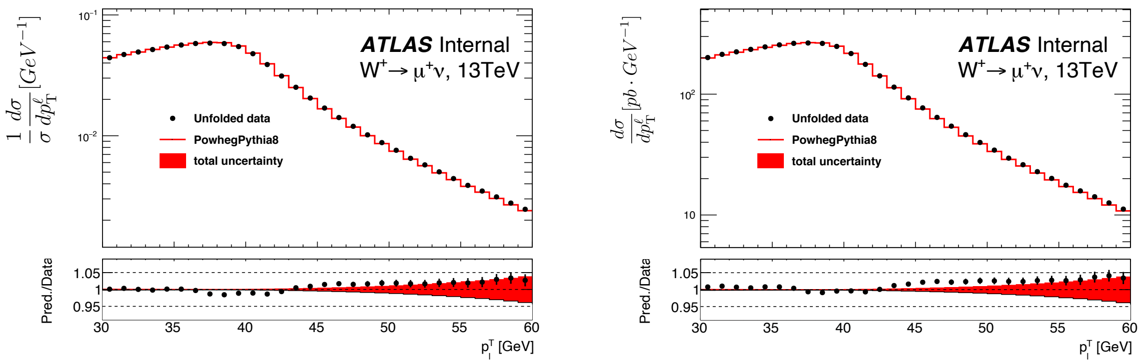


FIGURE 7.18: Differential cross sections (right) and normalised differential cross sections (left) as a function of p_{T}^{μ} for 13 TeV (W^{+}, μ^{+}). The bottom panel shows the ratio data to MC (Powheg+Pethia8) together with the red band showing the total uncertainty.

Differential cross sections of the e channel versus η_e

TABLE 7.9: Differential cross sections versus η_e at 5 TeV (W^+, e^+). The columns show the bin range, measured cross section, relative uncertainties.

	$W^+ \to e^+\nu_e$, 5 TeV, uncertainties in (%)			
Range	$d\sigma/d\eta_e$	Stat uncertainty	Unfolding bias	Syst uncertainty
[-2.50, -2.18]	367.836	0.678	0.000	0.642
[-1.95, -1.74]	461.329	0.775	0.000	0.525
[-1.74, -1.52]	469.095	0.667	0.000	0.641
[-1.52, -1.37]	468.811	0.427	0.146	0.587
[-1.37, -1.05]	464.068	0.543	0.000	0.568
[-1.05, -0.84]	465.592	0.643	0.000	0.492
[-0.84, -0.63]	464.723	0.683	0.000	0.433
[-0.63, -0.42]	460.784	0.637	0.000	0.380
[-0.42, -0.21]	452.088	0.705	0.000	0.375
[-0.21, 0.00]	449.530	0.637	0.000	0.537
[0.00, 0.21]	453.114	0.681	0.000	0.495
[0.21, 0.42]	456.250	0.643	0.000	0.325
[0.42, 0.63]	452.126	0.683	0.000	0.370
[0.63, 0.84]	454.408	0.662	0.000	0.645
[0.84, 1.05]	459.436	0.666	0.000	0.432
[1.05, 1.37]	469.338	0.537	0.000	0.555
[1.37, 1.52]	472.773	0.413	0.030	0.490
[1.52, 1.74]	474.614	0.670	0.000	0.467
[1.74, 1.95]	457.307	0.715	0.000	0.511
[2.18, 2.50]	371.495	0.694	0.000	0.797

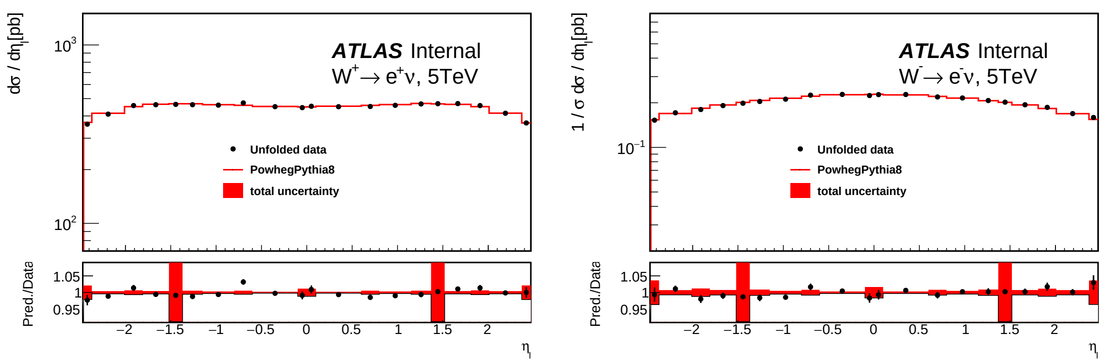


FIGURE 7.19: Differential cross sections (left) and normalised differential cross sections (right) as a function of p_T^e for 13 TeV (W^+, e^+). The bottom panel shows the ratio data to MC (Powheg+Pethia8) together with the red band showing the total uncertainty.

TABLE 7.10: Differential cross sections versus η_e at 5 TeV (W^-, e^-). The columns show the bin range, measured cross section, relative uncertainties.

	$W^- \to e^-\bar{\nu}_e$, 5 TeV, uncertainties in (%)			
Range	$d\sigma/d\eta_e$	Stat uncertainty	Unfolding bias	Syst uncertainty
[-2.50, -2.18]	216.082	0.935	0.944	0.688
[-2.18, -1.95]	249.183	0.929	0.491	0.450
[-1.95, -1.74]	250.450	0.972	0.067	0.531
[-1.74, -1.52]	268.866	0.997	0.578	0.617
[-1.52, -1.37]	275.843	0.602	0.948	0.550
[-1.37, -1.05]	284.339	0.720	0.999	0.519
[-1.05, -0.84]	293.295	0.888	0.591	0.463
[-0.84, -0.63]	305.781	0.838	0.119	0.421
[-0.63, -0.42]	312.623	0.853	0.303	0.353
[-0.42, -0.21]	315.281	0.818	0.569	0.344
[-0.21, 0.00]	311.796	0.802	0.623	0.486
[0.00, 0.21]	317.227	0.796	0.457	0.448
[0.21, 0.42]	314.936	0.842	0.165	0.304
[0.42, 0.63]	308.392	0.770	0.150	0.343
[0.63, 0.84]	302.388	0.800	0.361	0.581
[0.84, 1.05]	298.429	0.797	0.345	0.389
[1.05, 1.37]	289.149	0.712	0.084	0.501
[1.37, 1.52]	280.873	0.552	0.762	0.467
[1.52, 1.74]	272.492	0.908	1.314	0.447
[1.74, 1.95]	256.505	0.972	1.667	0.493
[1.95, 2.18]	250.888	0.912	1.039	0.481
[2.18, 2.50]	215.071	0.924	2.373	0.858

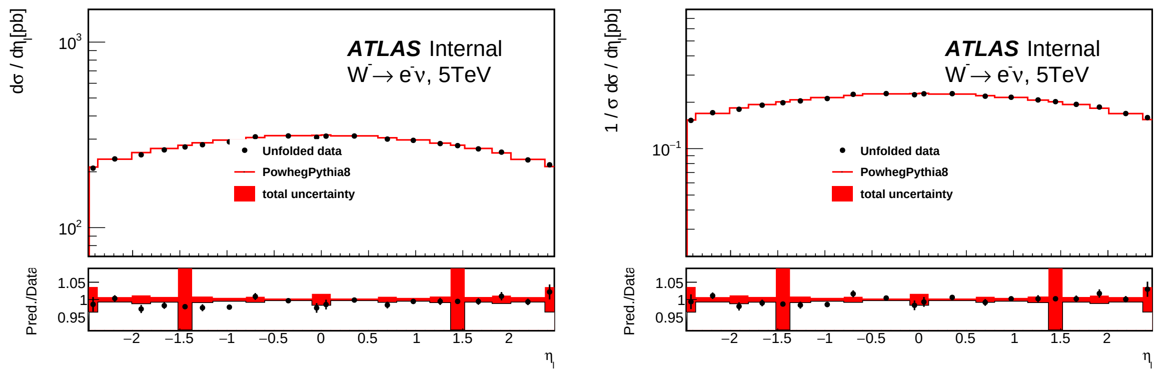


FIGURE 7.20: Differential cross sections (left) and normalised differential cross sections (right) as a function of p_{T}^{e} for 13 TeV (W^-, e^-). The bottom panel shows the ratio data to MC (Powheg+Pethia8) together with the red band showing the total uncertainty.

Differential cross sections of the μ channel versus η_μ

TABLE 7.11: Differential cross-sections versus η^μ at 5 TeV (W^+, μ^+). The columns show the bin range, measured cross section, relative uncertainties.

	$W^+ \to \mu^+\nu_\mu$, 5 TeV, uncertainties in (%)			
Range	$d\sigma/d\eta_\mu$	Stat uncertainty	Unfolding bias	Syst uncertainty
[-2.40, -1.92]	417.143	0.470	0.000	0.761
[-1.92, -1.35]	464.547	0.411	0.000	0.571
[-1.35, -1.15]	462.196	0.670	0.000	0.697
[-1.15, -1.05]	458.964	1.183	0.000	2.246
[-1.05, -0.91]	461.993	0.987	0.000	2.243
[-0.91, -0.48]	453.411	0.504	0.000	0.782
[-0.48, 0.00]	446.514	0.529	0.000	0.896
[0.00, 0.48]	443.203	0.493	0.000	0.886
[0.48, 0.91]	454.335	0.489	0.000	0.893
[0.91, 1.05]	466.137	1.020	0.000	2.233
[1.05, 1.15]	455.781	1.055	0.000	1.775
[1.15, 1.35]	474.952	0.689	0.000	0.664
[1.35, 1.92]	459.009	0.412	0.000	0.543
[1.92, 2.40]	417.343	0.470	0.000	0.769

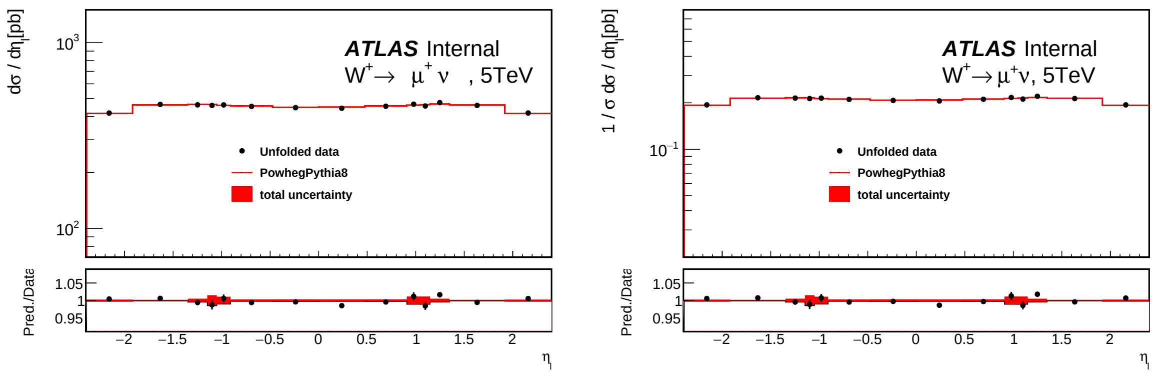


FIGURE 7.21: Differential cross sections (left) and normalised differential cross sections (right) as a function of p_T^μ for 13 TeV (W^+, μ^+). The bottom panel shows the ratio data to MC (Powheg+Pethia8) together with the red band showing the total uncertainty

TABLE 7.12: Differential cross-sections versus η^{μ} at 5 TeV (W^-, μ^-). The columns show the bin range, measured cross section, relative uncertainties.

	$W^- \to \mu^- \bar{\nu}_\mu$, 5 TeV, uncertainties in (%)			
Range	$d\sigma/d\eta_\mu$	Stat uncertainty	Unfolding bias	Syst uncertainty
[-2.40, -1.92]	234.738	0.621	0.000	0.752
[-1.92, -1.35]	268.060	0.508	0.000	0.501
[-1.35, -1.15]	285.380	0.881	0.000	1.307
[-1.15, -1.05]	306.520	1.296	0.000	3.090
[-1.05, -0.91]	297.537	1.314	0.000	0.024
[-0.91, -0.48]	302.886	0.607	0.000	1.061
[-0.48, 0.00]	311.382	0.569	0.000	0.738
[0.00, 0.48]	309.306	0.633	0.000	0.707
[0.48, 0.91]	303.490	0.635	0.000	0.778
[0.91, 1.05]	293.061	1.202	0.000	2.989
[1.05, 1.15]	284.430	1.390	0.000	3.657
[1.15, 1.35]	285.125	0.837	0.000	1.393
[1.35, 1.92]	264.657	0.512	0.000	0.550
[1.92, 2.40]	239.405	0.592	0.000	0.775

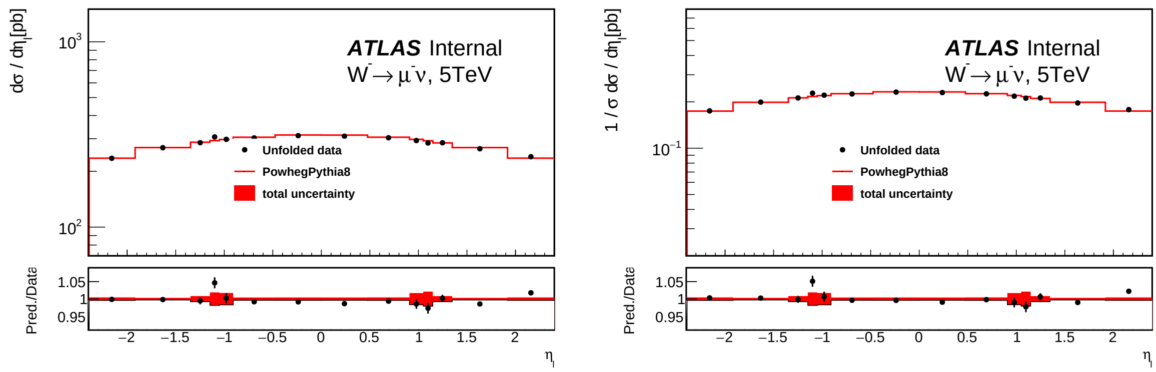


FIGURE 7.22: Differential cross sections (left) and normalised differential cross sections (right) as a function of p_T^{μ} for 13 TeV (W^-, μ^-). The bottom panel shows the ratio data to MC (Powheg+Pethia8) together with the red band showing the total uncertainty

7.4 Comparison of electron and muon channels

The differential cross sections for electron and muon are calculated using different binning in η direction. The choose of η binning is related mainly to the scale factor (SF) binning (reconstruction, trigger, isolation and identification SFs) that we apply to simulation in order to correct the difference between data and simulation. The electron SFs are calculated in the same binning, Table 7.13, while for muon, the SFs are calculated using different binning, Table 7.14.

TABLE 7.13: Values of η bin boundaries for electron SFs.

−2.47	−2.37	−2.01	−1.81	−1.37	−1.15	−0.8	−0.6	−0.1	0	0.1	0.6	0.8	1.15	1.37	1.52	1.81	2.01	2.37	2.47

TABLE 7.14: Values of η bin boundaries for muon trigger SF.

−2.4	−1.918	−1.348	−1.1479	−1.05	−0.908	−0.476	0	0.476	0.908	1.05	1.1479	1.348	1.918	2.4

The idea to compare the differential cross sections in the electron and muon channels, is to keep the binning at the reconstruction level unchanged, to conserve the SFs effects, and change the binning at the unfolded level to a common binning for the two channels. The new binning is chosen in a such a way that we conserve the bin boundaries similar to the SFs binning at the reconstructed level, Table 7.15. Figure 7.23 shows the comparison between the different SFs for electron and muon and the proposed binning at the unfolded level.

TABLE 7.15: Values of η bin boundaries for new binning at the unfolded level.

−2.5	−1.85	−1.36	−1.05	−0.85	−0.5	0	0.5	0.85	1.05	1.36	1.85	2.5

The distributions of η for electron and muon are unfolded to a common unfolded level, Figure 7.24 shows the comparison between distributions at the reconstruction and unfolded level, together with differential cross sections for electron and muon. The comparison of the cross sections shows in a good agreement for electrons and muons, excepting for the around $\eta \approx 1.2$, where the difference is related mainly to the variation of trigger SF for muon, shown in Figure 7.23. The difference observed is around 1.8% and included in the uncertainty. For the comparison with theoretical predictions, the binning defined in the Tables 7.13 and 7.14 are used in order to conserve the effect of scale factors.

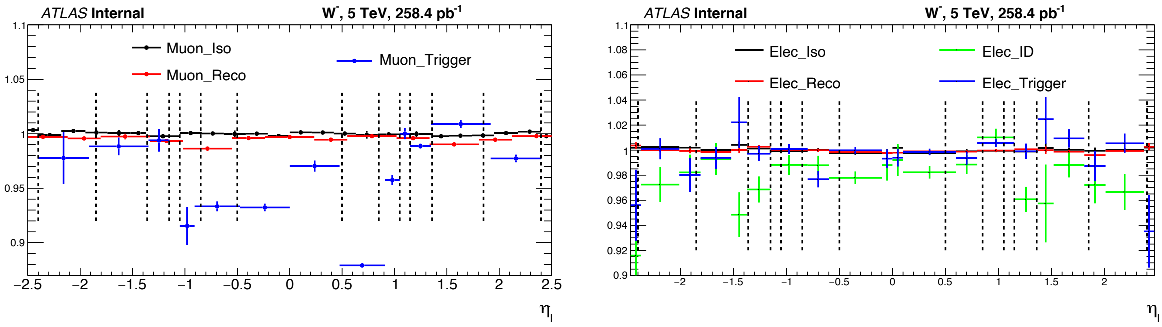


FIGURE 7.23: Muon (left) and electron (right) scale factors (SFs) used to correct simulation. The muons SFs are calculated using different binning, while the electron SFs are calculated in the same binning. The dotted vertical line shows the boundaries for proposed common binning at the unfolded level.

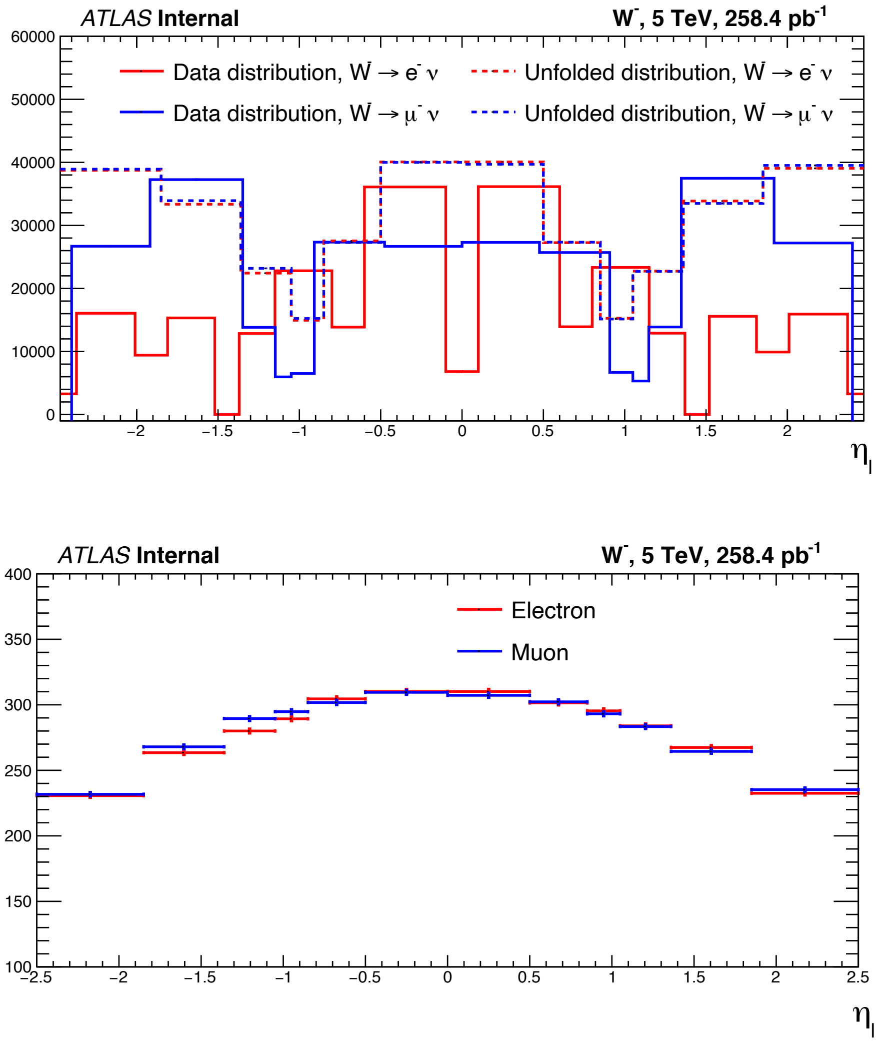


FIGURE 7.24: (Top) Distributions of data as a function fo η_ℓ for electron and muon at the reconstructed and the unfolded levels using the new common binning at the unfolded level. (Bottom) Comparison of the differential cross sections as a function of η_ℓ for electron and muon.

7.5 Comparison with theoretical predictions

The measured differential cross sections for $W^{\pm} \rightarrow \ell^{\pm}\nu$ are compared to theoretical predictions using DYTURBO [45] at NNLO QCD and LO in the EW theory, with different PDF sets: CT18 [93], HERAPDF20 [87], MMHT2014 [91], in the fiducial phase space defined in Section 7.2. The differential cross sections are compared separately for electron and muon without combination. The uncertainties of the theoretical predictions arise from the limited knowledge of proton PDFs. The DYTURBO uses input parameters (G_F, M_W, M_Z) for the theoretical predictions. The PDF sets used were extracted from analyses of various experimental data sets using the corresponding predictions at NNLO in QCD.

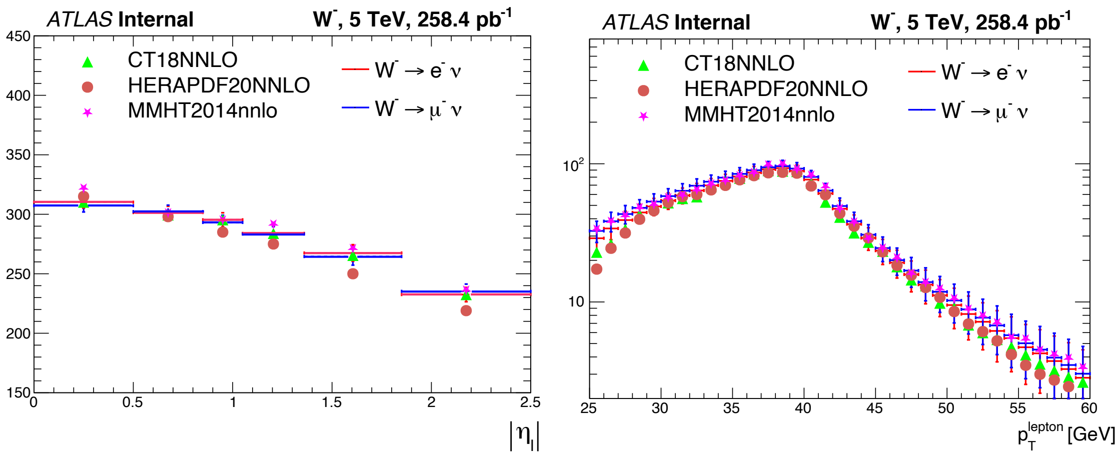


FIGURE 7.25: (left) Differential cross sections as a function of η_ℓ for electron and muon compared to different PDF sets. (right) Differential cross sections as a function of p_{T}^{ℓ} for electron and muon compared to different PDF sets.

The PDF uncertainty on the M_W measurement, the dominant source of physics modelling uncertainty $\approx$ 9.2 MeV, arises from our imperfect knowledge of the PDFs affecting the differential cross section as a function of boson rapidity, the angular coefficients, and the W boson transverse momentum distribution. The measurements of the differential cross sections of the W boson, as a function of η_ℓ, are used to validate and constrain the PDF uncertainty on the measurement of W boson, by comparing the uncertainties on the measured level and the uncertainties on the PDF predictions.

7.6 Double differential cross sections in p_{T}^{ℓ} and η_{ℓ} bins

7.6.1 Introduction

Double differential cross sections in p_{T}^{ℓ} and η_{ℓ} bins are measured using a two dimensional (2d) unfolding of data distributions. The two dimensional unfolding can be transferred to a one dimensional (1d) unfolding by splitting the data distributions of η_{ℓ}, in different ranges of p_{T}^{ℓ} as shown in Fig. 7.26. The statistical and systematic uncertainties are evaluated in the same way as we did for one dimensional unfolding described in Sec. 7.6.2. The bin-by-bin correction method can not be used because of the large migration between bins (Fig. 7.27).

The double differential cross sections can be expressed as:

$$\frac{d\sigma}{dp^{l}_{\mathrm{T},i}d\eta^{l}_{j}} = \frac{N^{\mathrm{Unfolded}}_{i,j}}{\mathcal{L}\Delta p^{l}_{\mathrm{T}}\Delta\eta^{l}} \cdot \frac{1}{A_{\mathrm{unf}}} \tag{7.4}$$

where N_{Unfolded} represents the number of events in the unfolded distribution, and A_{unf} is a correction factor related to the unfolding procedure. This factor represents the fraction of the entries in a truth bin that are in the same bin at reconstruction level, $\mathcal{L}$ is the integrated luminosity of data, and $\Delta p_{\mathrm{T}}^{\ell}$, $\Delta\eta_{\ell}$ are the bin widths.

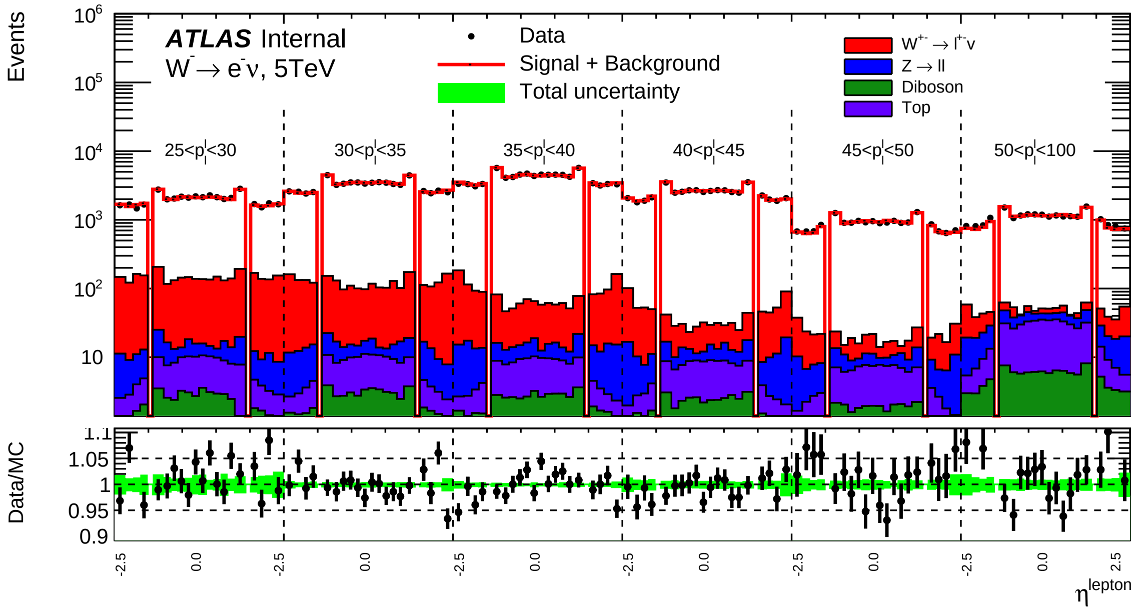


FIGURE 7.26: Distributions of the observables chosen to be unfolded η_{ℓ} in bins of p_{T}^{ℓ} for 5 TeV in the fiducial phase space. The signal and backgrounds are normalised to data. The low panel gives the ratio of the numbers of observed events to the total prediction in each bin. The green band shows the total statistical and systematic uncertainties. All the comparisons data/MC for 5 and 13 TeV are shown in Appendix B.

7.6.2 Migration matrix

In a migration matrix, one axis, e.g. the x-axis corresponds to reconstructed bins, the y-axis to true bins. For the double differential cross sections, the migration matrix is constructed in the same way but we take into account the different ranges of p_{T}^{ℓ}. The x-axis, corresponds to reconstructed η_ℓ in different ranges of reconstructed p_{T}^{ℓ}. The y-axis, corresponds to true η_ℓ in different ranges of true p_{T}^{ℓ}.

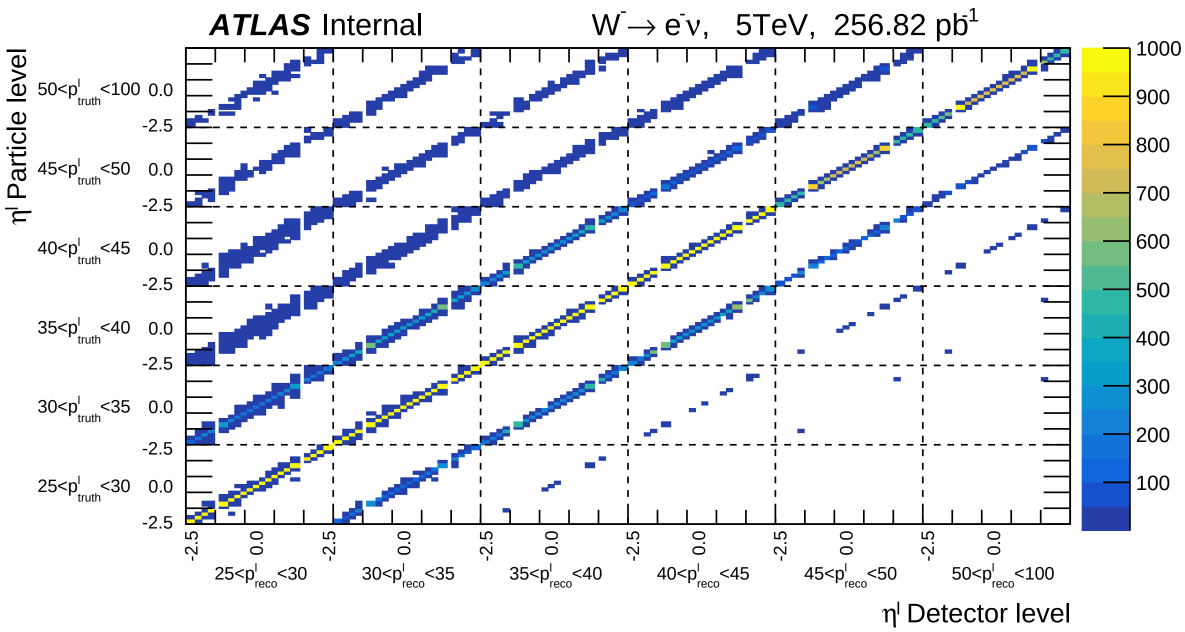


FIGURE 7.27: Example of the migration matrix used to unfold the data distribution for the measurement of the double differential cross sections, for $W^- \to e^-\nu$ at 5 TeV.

7.6.3 Statistical uncertainty

As the 2d unfolding problem is transformed into 1d unfolding, the statistical and systematic uncertainties are calculated as described in Sec. 7.6.2. As shown in Fig. 7.27, the 2d unfolding is characterised with a large migration between bins which explains the variation of statistical uncertainty with the number of iterations (Fig. 7.28).

7.6.4 Unfolding bias

The bias is calculated as described for 1d unfolding (Sec. 7.6.2). The only difference is that we fit the ratio data/MC separately for each range of p_{T}^{ℓ} as shown in Fig. 7.29.

Because of the migration in the 2 dimensional unfolding, the bias is in the order of 1% for the first iteration and decreases with the number of iterations. For the double differential cross-sections results, 4 iterations are used in the final results to ensure that bias contribution is negligible comparing to other source of uncertainties. There are some bins where the bias is in the order 1.5% and does not change

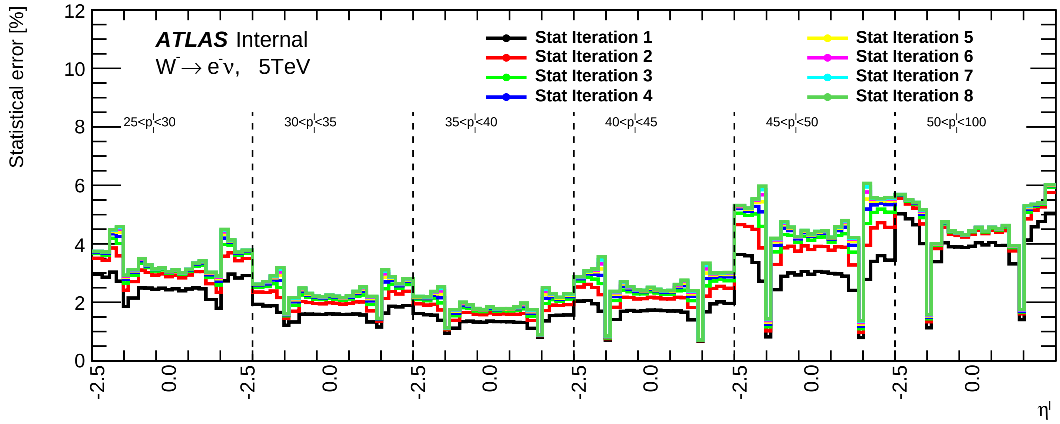


FIGURE 7.28: Example of the statistical uncertainty of unfolded distribution of η_ℓ in different ranges of p_{T}^{ℓ}

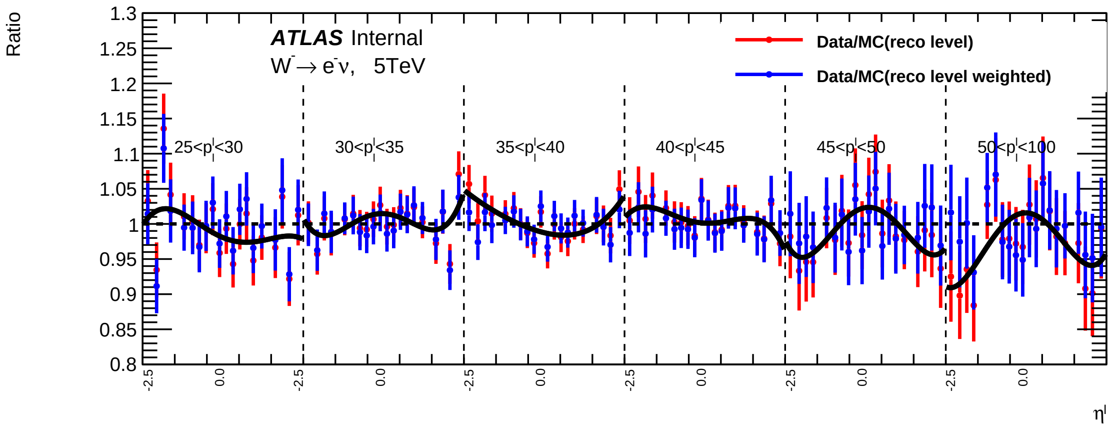


FIGURE 7.29: Example of the fitted ratio data/MC of η_ℓ in different ranges of p_{T}^{ℓ}

2050 with iterations. Basically, the bias values in these bins have no signification as bins correspond to empty bin [1.52, 1.37] (Fig. 7.30).

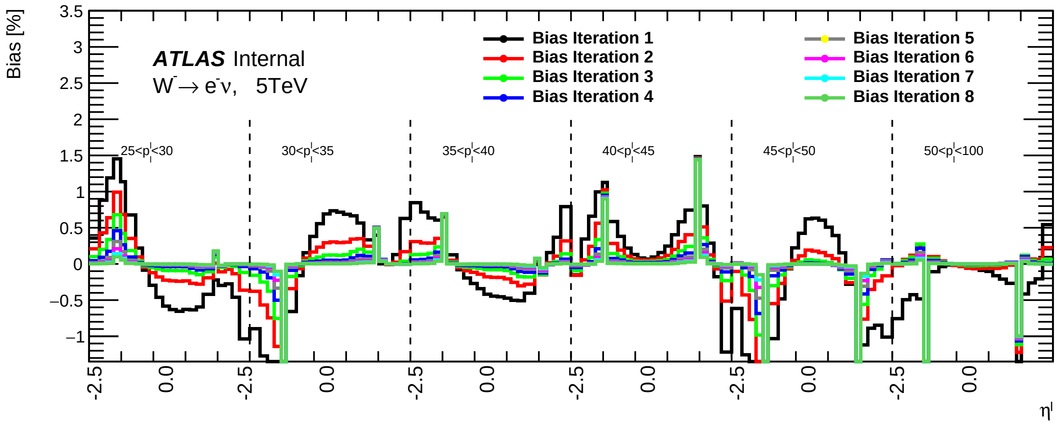


FIGURE 7.30: Example of the bias uncertainty as a function of η_ℓ in different ranges of p_T^ℓ.

2051

7.6.5 Double differential cross sections

The double differential cross section results, together with the statistical, experimental systematic and unfolding bias uncertainties, are shown in Figures 7.31 and 7.32 for 5 TeV and Figures 7.33 and 7.34 for 13 TeV.

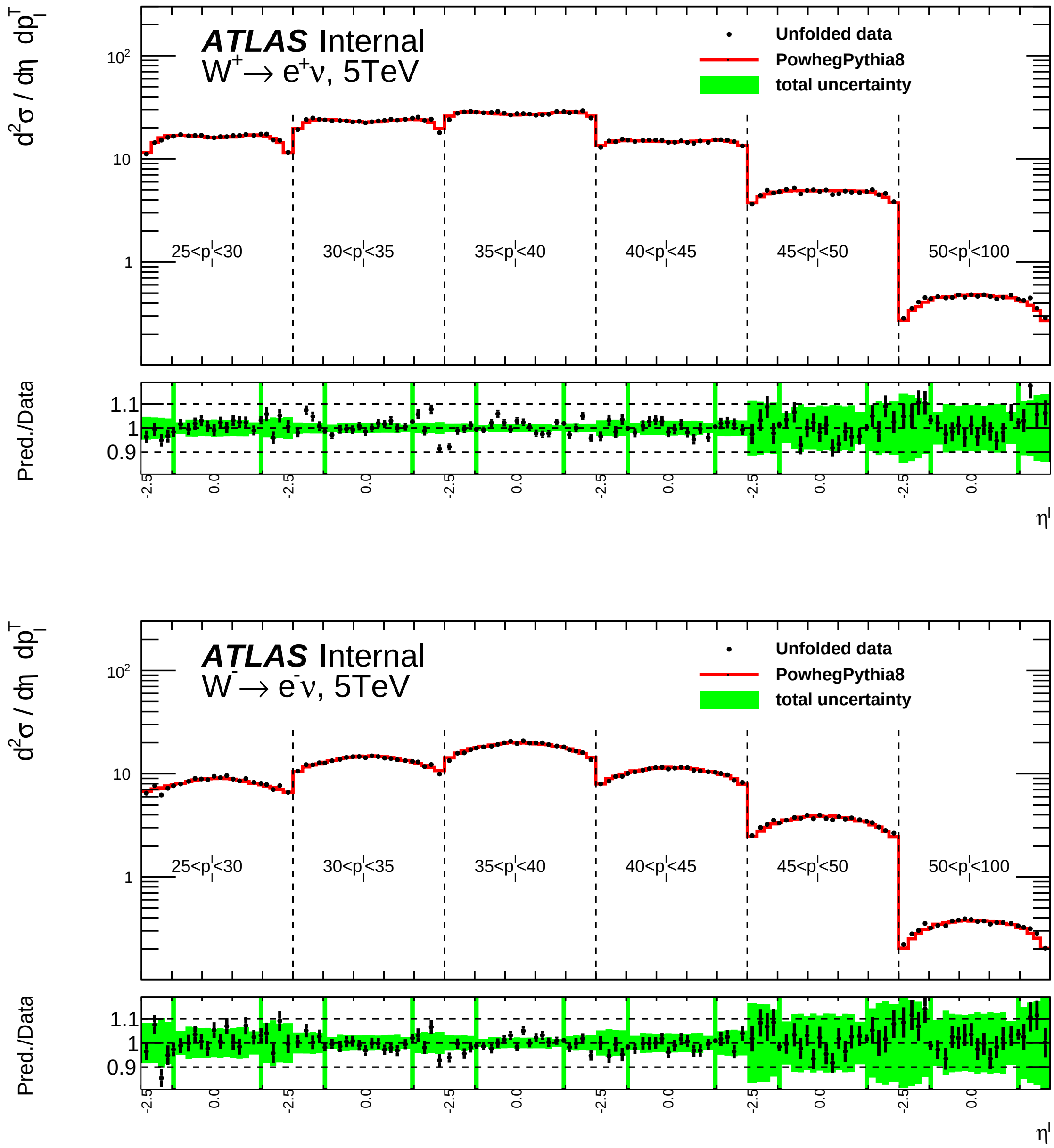


FIGURE 7.31: Double differential cross sections in bins of η_ℓ and p_T^ℓ compared to Powheg+Pythia8 for $W^+ \to e^+\nu$ and $W^- \to e^-\nu$ at 5 TeV. The low panel shows the ratio data/MC and the green band represents the statistical and systematic uncertainties added in quadrature.

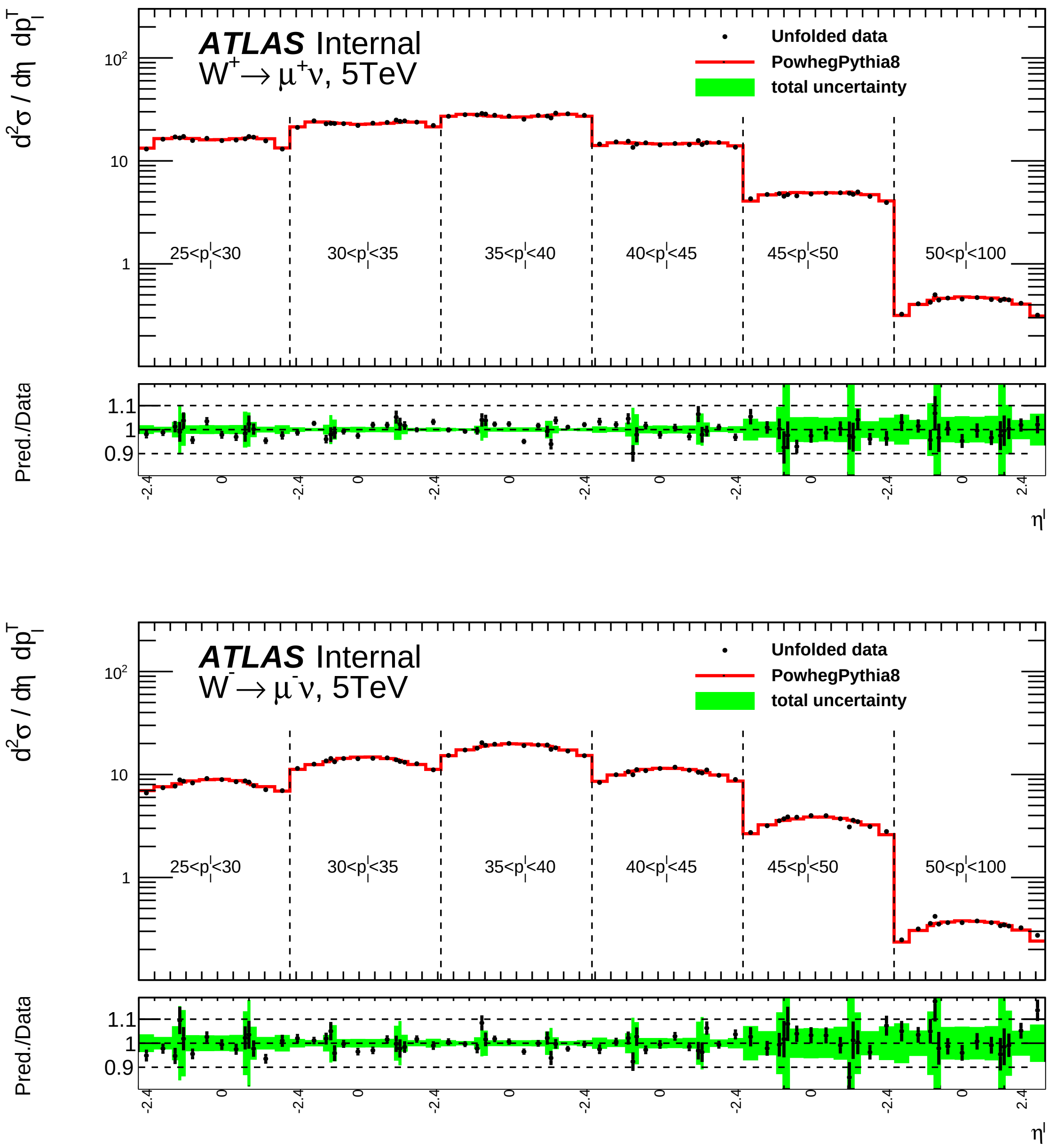
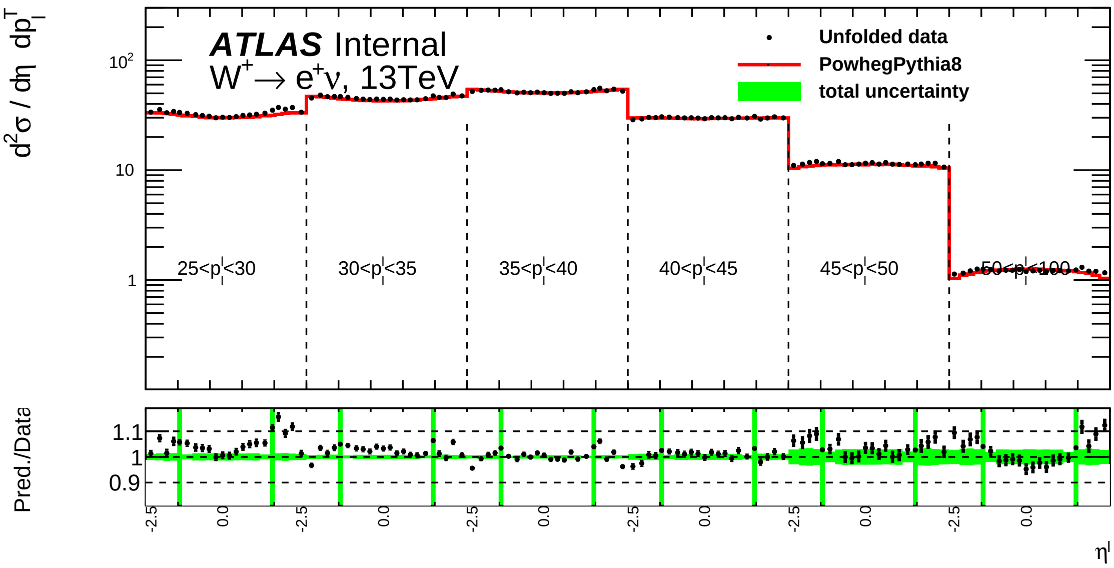


FIGURE 7.32: Double differential cross sections in bins of η_ℓ and p_T^ℓ compared to Powheg+Pythia8 for $W^+ \rightarrow \mu^+\nu$ and $W^- \rightarrow \mu^-\nu$ at 5 TeV. The low panel shows the ratio data/MC and the green band represents the statistical and systematic uncertainties added in quadrature.



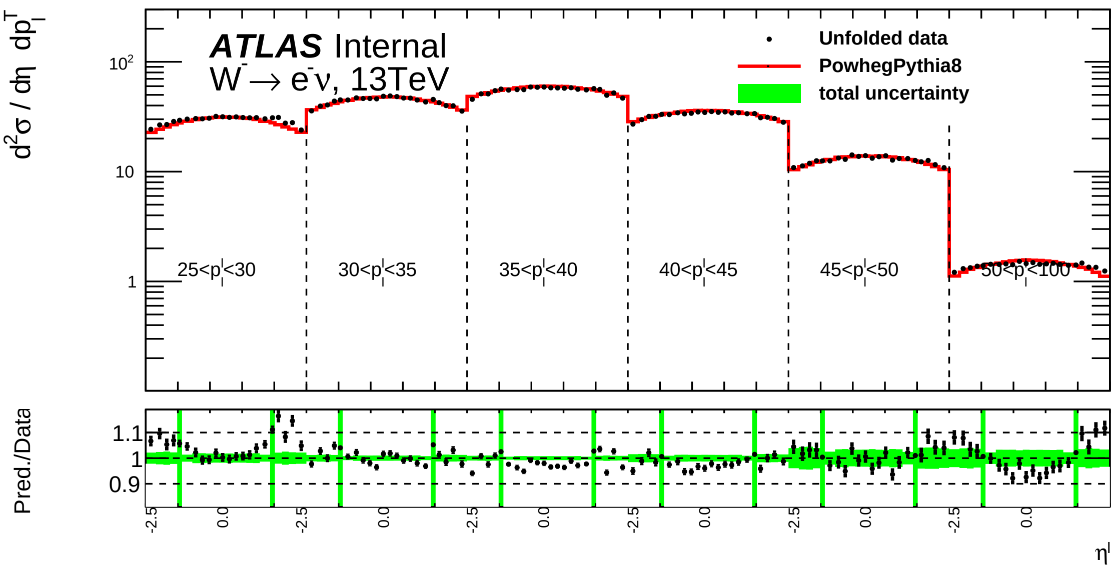


FIGURE 7.33: Double differential cross sections in bins of η_ℓ and p_T^ℓ compared to Powheg+Pythia8 for $W^+ \rightarrow e^+\nu$ and $W^- \rightarrow e^-\nu$ at 13 TeV. The low panel shows the ratio data/MC and the green band represents the statistical and systematic uncertainties added in quadrature.

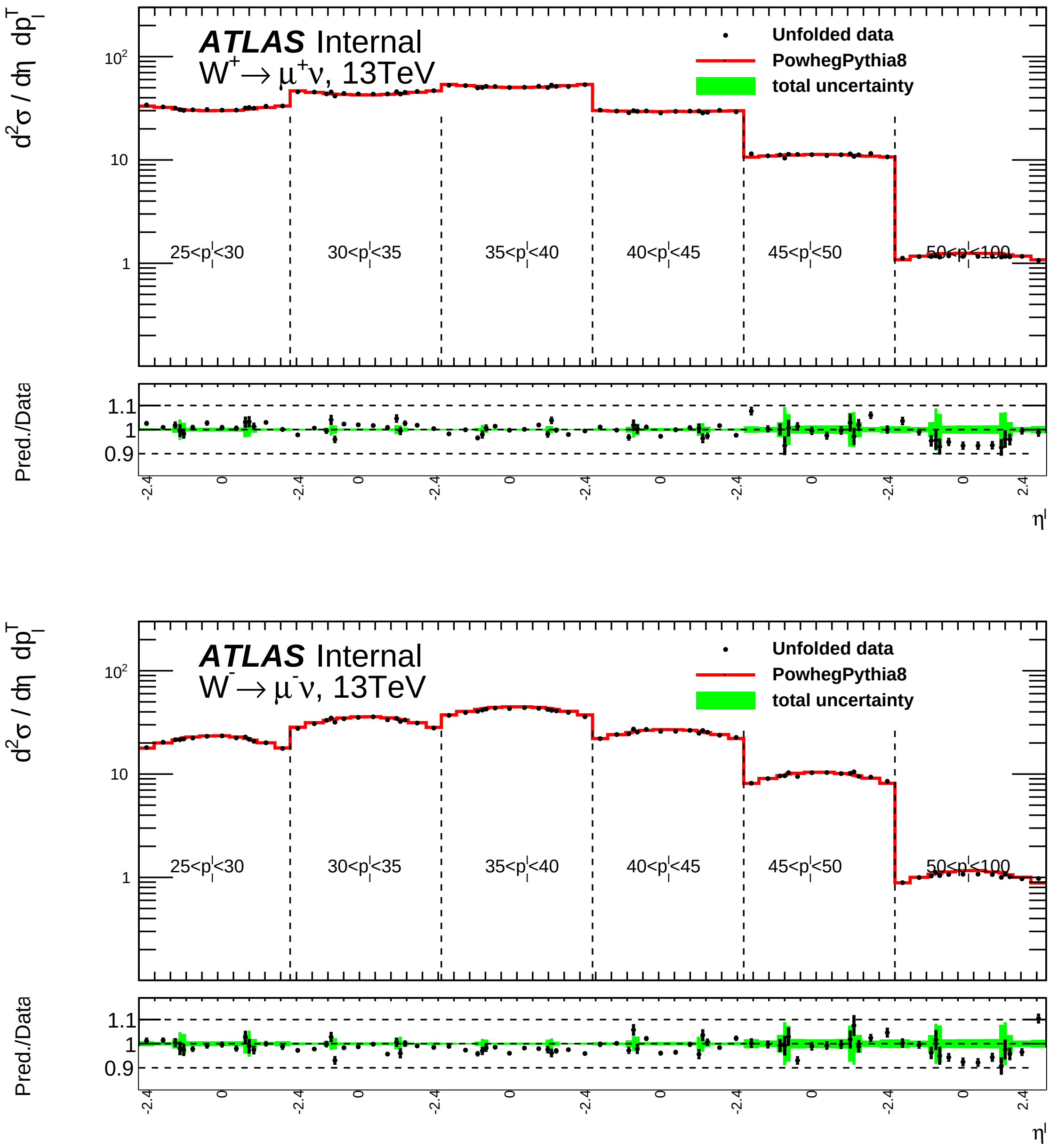


FIGURE 7.34: Double differential cross sections in bins of η_ℓ and p_T^ℓ compared to Powheg+Pythia8 for $W^+ \to \mu^+\nu$ and $W^- \to \mu^-\nu$ at 13 TeV. The low panel shows the ratio data/MC and the green band represents the stat and systematic uncertainties added in quadrature.

Chapter 8

Measurement of the W-boson mass

8.1 Introduction

This chapter will show preliminary results of the measurement of W boson mass using low pile-up data set at $\sqrt{s}$ = 5 TeV and 13 TeV with two approaches: using the templates method [17], developed before for Run 1 analysis, and using the unfolded distributions of our variables of interest. The methodology of using the unfolded distributions for W boson mass is described in Sec. 8.3. In parallel of those methods, there is another approach, using a new fitting algorithm in global W mass, with the profile likelihood approach [77], which treats the correlation between uncertainties differently from the template method. However, in this chapter, we will focus on the evaluation of statistical uncertainty on the W boson mass measurement using the two approaches described above, and the dominated experimental uncertainties: lepton efficiency, lepton calibration and hadronic recoil calibration.

8.2 Template fit method methodology

The W boson is an unstable particle which decays to a charged lepton and a neutrino. The mass of the W boson is determined using the distributions of the transverse mass of W (m_{T}^{W}) and of the transverse momentum of lepton (p_{T}^{ℓ}), where the p_{T}^{ℓ} distribution has a Jacobian peak at $M_W/2$, while the transverse mass peak at M_W, Figure 8.1 shows the distributions of p_{T}^{ℓ} and m_{T}^{W} at the Jacobian peaks. The basic idea of the template method consists in computing the p_{T}^{ℓ} and m_{T}^{W} distributions for different assumed values of M_W, called the templates, and the comparison between templates and data gives the best fit value.

To generate templates with different W masses, the truth level distributions are reweighted using the Breit-Wigner equation:

$$f(m_W) = \frac{d\sigma}{dm} \propto \frac{m^2}{\left(m^2 - m_W^2\right)^2 + m^4\Gamma_W^2/m_W^2}, \tag{8.1}$$

where m_W is the W boson mass, and the weight applied to truth distributions is considered as:

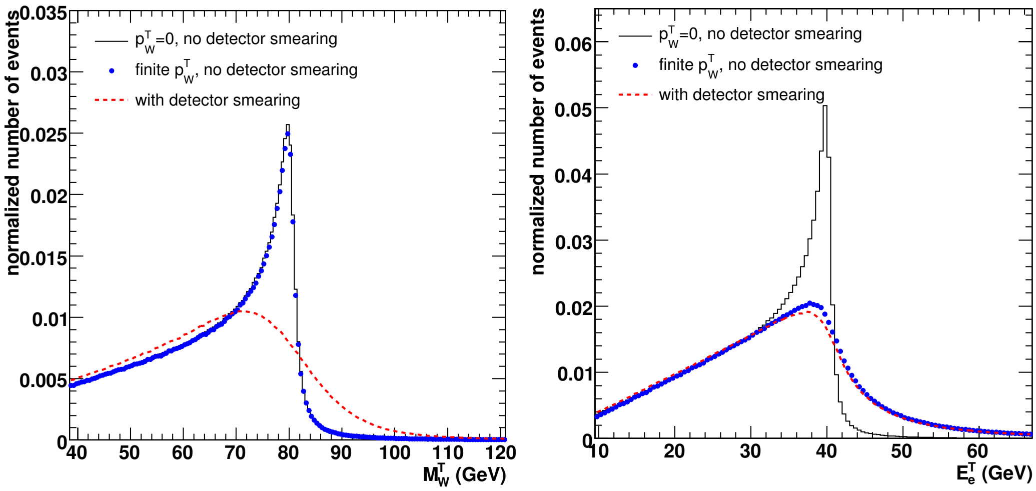
$$\mathrm{weight} = \frac{f(m'_W)}{f(m_W)}, \tag{8.2}$$

where m'_W is the modified mass.

Figure 8.6 shows an example of p_{T}^{ℓ} and m_{T}^{W} distributions compared to the templates generated with different mass values. Then, the comparison between templates and data is based on χ^2 defined as:

$$\chi^2 = \sum_{i=1}^{N} \frac{\left(n_i^{\text{data}} - n_i^{\text{template}}\right)^2}{(\sigma_{n_i}^{\text{data}})^2 + (\sigma_{n_i}^{\text{template}})^2}, \tag{8.3}$$

where n_i^{data} (n_i^{template}) is the number of entries in bin i of data (template), and $\sigma_{n_i}^{\text{data}}$ ($\sigma_{n_i}^{\text{template}}$) is the uncertainty in bin i of data (template). The background is subtracted from the number of entries in data n_i^{data}.



(A) (B)

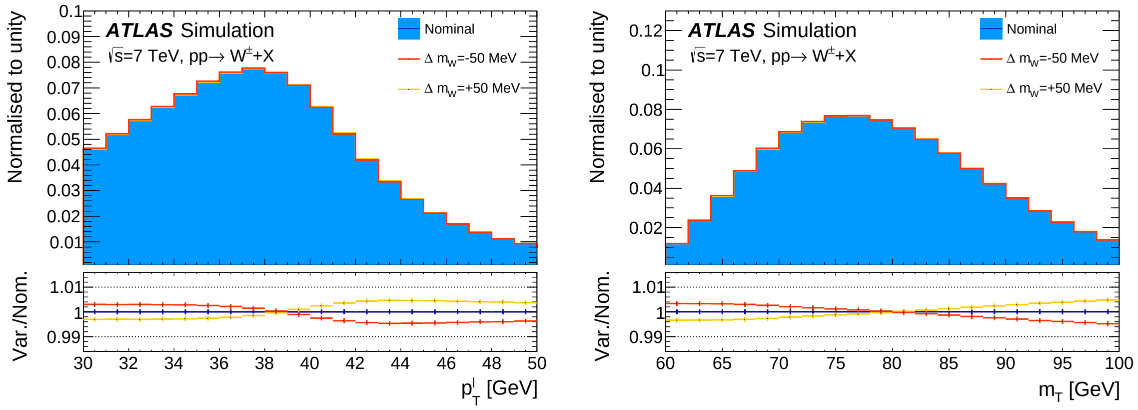
FIGURE 8.1: Transverse mass of W (A) and lepton transverse momentum (B) distributions in W decays. The distributions at the generator level with $p_{\mathrm{T}}^{W} = 0$ (blackline), with finite W boson p_{T}^{W} (blue dots) and including the experimental resolution in the low luminosity phase (red dashed line) are shown [112].

The χ^2 is calculated between data and each template separately, then the computed χ^2 values are fitted using a polynomial function. The minimum of χ^2 distribution gives the best M_W value. Figure 8.3 shows an example of the fitted χ^2 distribution. The templates used in the W-mass fit are signal MC samples reweighted to $M_W \pm$ [0, 25, 50, 100, 150, 200] MeV.

This method has been used in previous experiments (CDF and DO) for the W mass measurement. In parallel to the template method, there is a new method [77] being developed called "profile likelihood" approach, which allows to deal with systematic uncertainties and their correlations in a different way.

8.3 W boson mass using the unfolded distribution

In addition to the method described above, there is a different approach consisting in using the distributions at the unfolded level instead of the distributions at the reconstructed level. The main idea is to use distributions which are already



(A) (B)

FIGURE 8.2: Kinematic distributions of p_{T}^{ℓ} (A) and m_{T}^{W} (B) in simulated events for the W-boson mass nominal value $M_W = 80370$ MeV and the shifted values $M_W = 80320$ MeV and $M_W = 80420$ MeV [115].

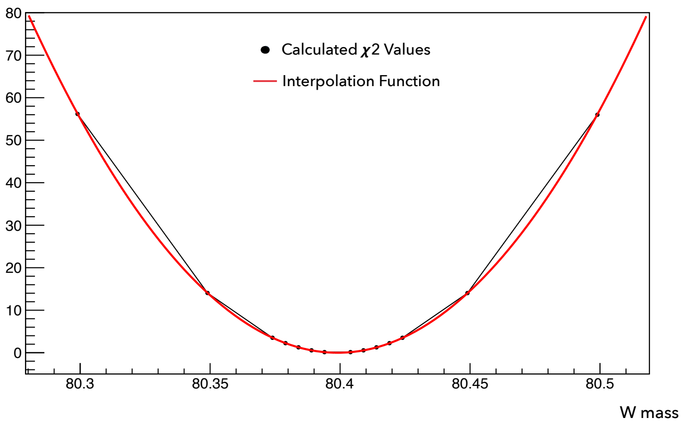


FIGURE 8.3: Fit to χ^2 distribution at different template mass values.

corrected by the unfolding procedure and does not contain undesirable detector effects. The extraction of the M_W boson is the same as described with the template method, except for the χ^2 formula which have to be changed to take into account the correlation between bins at the unfolded level, introduced by the unfolding procedure. The new χ^2 formula is expressed as:

$$\chi^2 = (n^{\text{Unf}}_{\text{data}} - n^{\text{Unf}}_{\text{template}})^T \cdot (V_{\text{data}} + V_{\text{template}})^{-1} \cdot (n^{\text{Unf}}_{\text{data}} - n^{\text{Unf}}_{\text{template}}), \tag{8.4}$$

where $n^{\text{Unf}}_{\text{data}}$ is the unfolded distribution of data, $n^{\text{Unf}}_{\text{template}}$ is the unfolded distribution of template, V_{data} (V_{template}) represents the covariance matrix of the statistical uncertainty for the unfolded distribution of data (template) calculated as described in Chapter 4. Once the χ^2 is calculated for all the unfolded templates, the procedure is the same as described for the template method. Ideally for both methods, we expect to have the same results but with an additional bias for the second method due to the unfolding of the variables of interest.

8.4 Statistical uncertainty

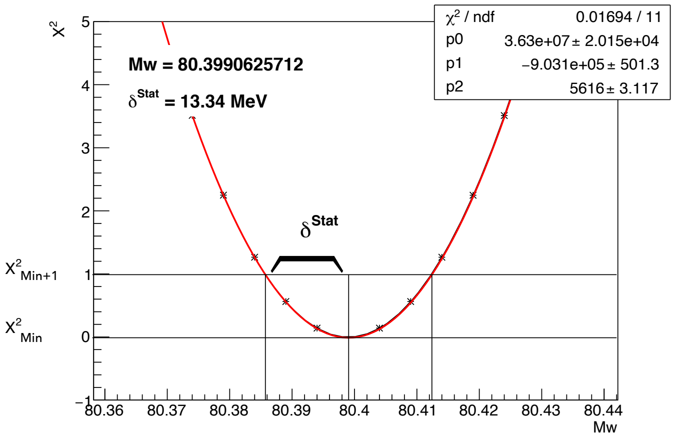
The evaluation of statistical uncertainty is based on data distributions and MC templates, calculated from the χ^2 fit using a parabola function and estimated as the deviation from the measured value of M_W and M'_W correspond to $\chi^2_{\text{min}} + 1$. Figure 8.4 shows an example of the statistical uncertainty estimation. The statistical uncertainties are calculated using distributions of p^{ℓ}_{T} and m^{W}_{T} separately and then combined. Since our distributions of interest are generated using the same events, we have to take into account the correlation between this two variables. The correlation is calculated using toys of MC (400 toys), generated by varying the p^{ℓ}_{T} and m^{W}_{T} distributions simultaneous with a random Poisson variation, and for each toy the W mass is calculated. Then the correlation factor is calculated as:

$$r = \frac{\sum_{i=1}^{N} \left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{N} \left(X_i - \bar{X}\right)^2}\sqrt{\sum_{i=1}^{N} \left(Y_i - \bar{Y}\right)^2}}, \tag{8.5}$$

where N is the number of toys, X_i (Y_i) represent the W mass results for toy i of p^{ℓ}_{T} (m^{W}_{T}), $\bar{X}$ ($\bar{Y}$) is the average of all the measured values X_i (Y_i). The final measured value of the W-boson statistical uncertainty is obtained from the combination of various measurements performed in the electron and muon channels, and in $|\eta|$-dependent categories, as defined in Table 8.1. The boundaries of the $|\eta|$ categories were defined as for Run 1 analysis, driven mainly by experimental and statistical constraints [115]. Figure B.8 shows an example of the correlation between p^{ℓ}_{T} and m^{W}_{T} with the corresponding correlation factor for different ranges of $|\eta|$. The

TABLE 8.1: Summary of categories and kinematic distributions used in the W mass analysis for the electron and muon decay channels [115].

Decay channel	$W \to ev$	$W \to \mu v$
Kinematic distributions	$p_{\mathrm{T}}^{\ell}, m_{\mathrm{T}}$	$p_{\mathrm{T}}^{\ell}, m_{\mathrm{T}}$
Charge categories	W^+, W^-	W^+, W^-
$\|\eta_\ell\|$ categories	[0,0.6], [0.6,1.2], [1.8,2.4]	[0,0.8], [0.8,1.4], [1.4,2.0], [2.0,2.4]



(A)

FIGURE 8.4: Statistical uncertainty calculation from χ^2 distribution.

average is done using BLUE [109].

8.5 Systematic uncertainties

In this section, we will describe the propagation of systematic uncertainties for the W boson mass measurement, focusing on the dominant uncertainties: lepton efficiency corrections, lepton calibration and hadronic recoil calibration. The modeling uncertainties: QCD, Electroweak and PDF's uncertainties are not included in the work described in this thesis. The propagation of uncertainties is based on the templates method introduced in Sec.8.2, where for each uncertainty source, a new set of MC templates is produced. The fitting is then performed separately for the modified and nominal MC templates, and the difference between the fitted values is considered as a systematic uncertainty. The resulting uncertainty for each error source is combined quadratically in order to have for the total uncertainty. The advantage of the template fit method is that it allows a detailed study of the impact of different experimental uncertainties independently, contrary to the profile likelihood approach [77] which gives a total uncertainty.

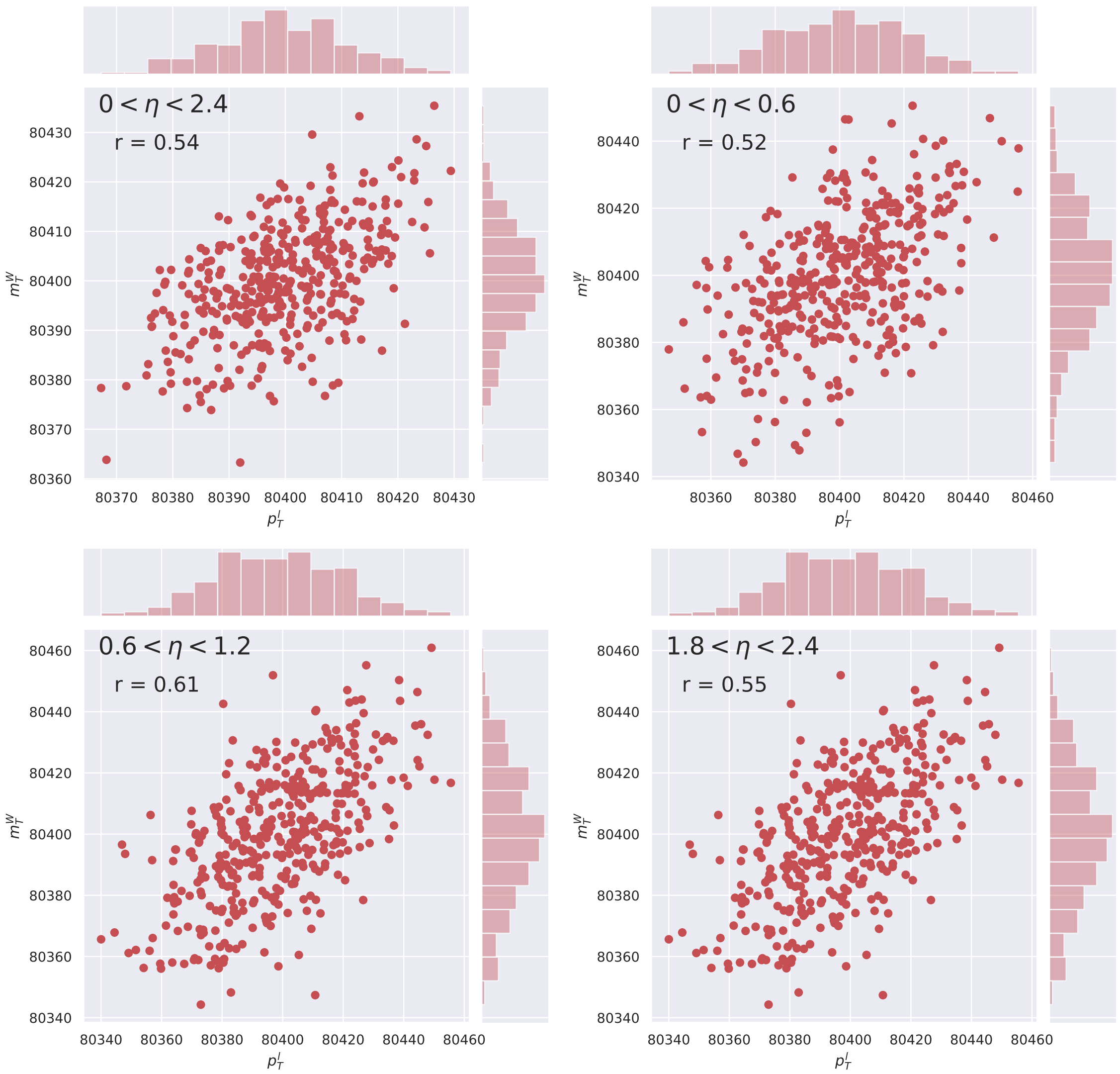


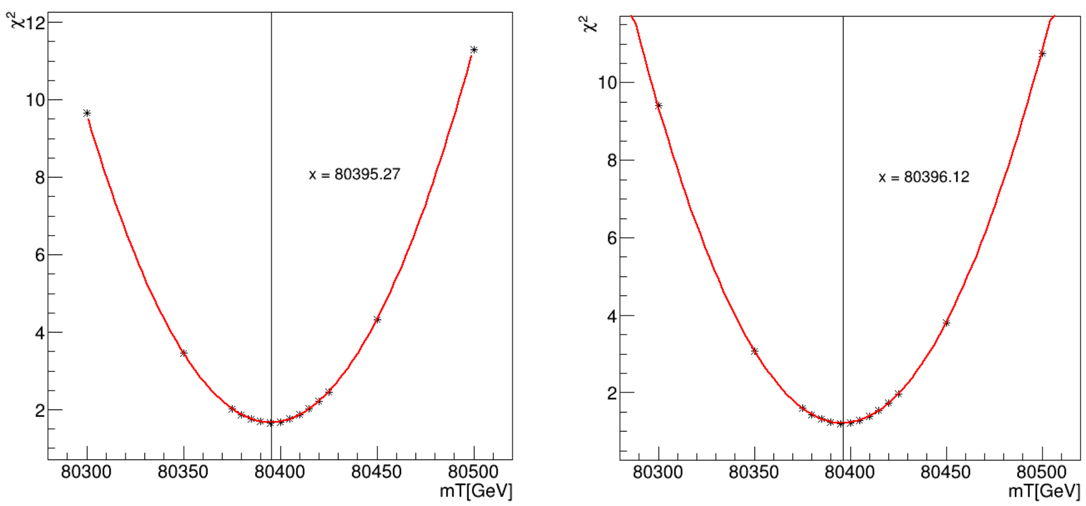
FIGURE 8.5: Correlation between p_{T}^{ℓ} and m_{T}^{W} with the corresponding correlation factor for different ranges of $|\eta_{\ell}|$.

TABLE 8.2: Statistical uncertainties in the M_W measurement for the different kinematic distributions and their combination in $|\eta_\ell|$ regions using data sets of 5 TeV.

η_ℓ range	[0, 0.6]		[0.6, 1.2]		[1.2, 1.8]		[1.8, 2.4]		[0, 2.4]	
Kinematic distribution	p_T^ℓ	m_T^W	p_T^ℓ	m_T^W	p_T^ℓ	m_T^W	p_T^ℓ	m_T^W	p_T^ℓ	m_T^W
Channel	$W^- \to e^-\bar{\nu}_e$, 5 TeV									
Stat[MeV]	55	49	58	53	78	70	61	71	32	29
Correlation	0.52		0.61		0.44		0.55		0.54	
Combined	45		50		63		58		27	
Channel	$W^+ \to e^+\nu_e$, 5 TeV									
Stat[MeV]	54	48	55	49	64	59	53	48	28	25
Correlation	0.57		0.60		0.59		0.57		0.56	
Combined	45		46		55		45		23	
Channel	$W^- \to \mu^-\bar{\nu}_\mu$, 5 TeV									
Stat[MeV]	55	48	59	53	58	55	78	73	31	28
Correlation	0.50		0.52		0.55		0.52		0.53	
Combined	44		49		50		66		26	
Channel	$W^+ \to \mu^+\nu_\mu$, 5 TeV									
Stat[MeV]	51	46	56	50	50	46	54	50	27	25
Correlation	0.51		0.59		0.60		0.50		0.56	
Combined	42		48		43		45		23	

TABLE 8.3: Statistical uncertainties in the M_W measurement for the different kinematic distributions and their combination in $|\eta_\ell|$ regions using data sets of 13 TeV.

η_ℓ range	[0, 0.6]		[0.6, 1.2]		[1.2, 1.8]		[1.8, 2.4]		[0, 2.4]	
Kinematic distribution	p_T^ℓ	m_T^W	p_T^ℓ	m_T^W	p_T^ℓ	m_T^W	p_T^ℓ	m_T^W	p_T^ℓ	m_T^W
Channel	$W^- \to e^-\bar{\nu}_e$, 13 TeV									
Stat[MeV]	37	35	39	36	51	49	44	42	21	20
Correlation	0.56		0.57		0.60		0.61		0.54	
Combined	32		33		45		39		18	
Channel	$W^+ \to e^+\nu_e$, 13 TeV									
Stat[MeV]	36	34	37	36	48	45	40	38	20	19
Correlation	0.59		0.63		0.60		0.67		0.59	
Combined	31		33		41		36		17	
Channel	$W^- \to \mu^-\bar{\nu}_\mu$, 13 TeV									
Stat[MeV]	35	33	39	38	39	37	48	47	20	19
Correlation	0.55		0.60		0.58		0.60		0.63	
Combined	30		31		34		42		18	
Channel	$W^+ \to \mu^+\nu_\mu$, 13 TeV									
Stat[MeV]	35	34	39	37	36	35	46	44	19	18
Correlation	0.57		0.60		0.64		0.64		0.62	
Combined	31		34		32		41		17	



(A) (B)

FIGURE 8.6: Example of M_W fitting results using the nominal MC templates (A) and varied MC templates (B) [155], the difference between the fitted values is considered as the propagated uncertainty on the M_W mass measurement.

- Lepton selection efficiency: lepton efficiency corrections are determined using tag-and-prob [1], and measured separately for electron reconstruction, identification and trigger efficiencies [47], using p_{T}^{ℓ} and m_{T}^{W} separately for different range of $|\eta_\ell|$. For p_{T}^{ℓ} and m_{T}^{W} ranges, and without including the extracted crack region ($1.2 < \eta_\ell < 1.8$), the reconstruction and identification efficiency corrections have an uncertainty of ≈ 4.5 MeV in the barrel region, and around 4 MeV in the end-cap region. The isolation and trigger efficiency corrections are smaller and have an uncertainty of 1 to 2 MeV in the barrel and end-cap.

TABLE 8.4: Lepton selection efficiency uncertainties (in MeV) on the measurement of W-boson mass, for $W^- \to e^-\bar{\nu}_e$ at 5 TeV.

η_ℓ range	[0, 0.6]		[0.6, 1.2]		[1.2, 1.8]		[1.8, 2.4]		[0, 2.4]	
Kinematic distribution	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}
Channel	$W^- \to e^-\bar{\nu}_e$, 5 TeV									
Identification efficiency	4.6	4.7	3.9	3.9	6.8	5.9	4.9	4.2	4.3	4.5
Isolation efficiency	2.1	1.3	2.3	1.3	3.6	2.2	2.3	1.4	2.1	1.2
Reconstruction efficiency	4.7	2.4	5.7	2.9	6.7	3.3	4.5	1.6	5.1	2.2
Trigger efficiency	1.9	1.7	1.3	1.2	2.4	1.9	7.1	4.9	1.4	0.9

- Lepton energy calibration: as shown in Ref. [152] for muons and in Chapter 3 for electrons, lepton energies are calibrated in order to correct the difference between data and simulation. For electrons, the uncertainty related to the lepton energy calibration is in particular due to the limited size of the $Z \to ee$ sample, used in the calibration procedure, while for muons, the uncertainty is related mainly to the limited knowledge of the detector alignment and resolution [155]. The uncertainty for electron channel is in the order of 18 MeV, and larger when we split bins of η_l because of statistical fluctuations.

TABLE 8.5: Lepton energy calibration uncertainties (in MeV) on the measurement of W-boson mass, for $W^- \to e^-\bar{\nu}_e$ at 5 TeV.

η_ℓ range	[0, 0.6]		[0.6, 1.2]		[1.2, 1.8]		[1.8, 2.4]		[0, 2.4]	
Kinematic distribution	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}
Channel	$W^- \to e^-\bar{\nu}_e$, 5 TeV									
Energy scale	27	28	30	33	44	48	44	48	19	20
Energy resolution	1.3	2.3	2.8	4.1	4.1	7.2	3.2	7.7	1.4	2.5

- Hadronic recoil calibration:

TABLE 8.6: Hadronic recoil calibration uncertainties (in MeV) on the measurement of W-boson mass, for $W^- \to e^-\bar{\nu}_e$ at 5 TeV.

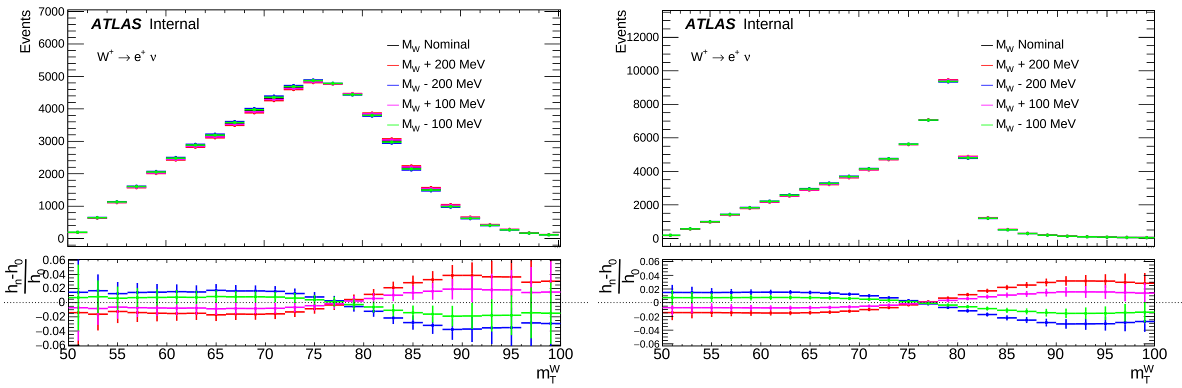
η_l range	[0, 0.6]		[0.6, 1.2]		[1.2, 1.8]		[1.8, 2.4]		[0, 2.4]	
Kinematic distribution	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}
channel	$W^- \to e^-\bar{\nu}_e$, 5 TeV									
$\sum E_{\mathrm{T}}$ reweighting	4.1	6.2	3.9	6.5	6.5	11.2	4.3	7.9	3.8	6.7
Resolution correction	1.9	2.7	2.0	3.4	5.3	8.6	2.3	3.2	2.0	2.9
Response correction	2.9	3.3	3.9	3.2	4.1	7.2	3.5	3.5	3.2	3.3

8.6 Statistical uncertainty with the unfolded distribution

As described in Sec. 8.3, unfolded distributions are already corrected by the unfolding procedure and do not include detector effects. The m_{T}^{W} and p_{T}^{ℓ} distributions with the corresponding templates are unfolded using the iterative Bayesian unfolding. The comparison between the modified templates and the nominal distributions at the reconstructed and unfolded levels is shown in Figure 8.7.

The main particularity of the unfolded distributions is that the unfolding procedure introduces a correlation between bins that we have to take into account in the χ^2 formula, while the statistical uncertainties of the different bins of the reconstructed distributions are fully uncorrelated. The correlation matrix for the statistical uncertainty of the unfolded distribution is calculated with the RooUnfold framework [7]. Figure 8.8 shows an example of the correlation matrix at the reconstructed and unfolded levels for the transverse mass m_{T}^{W}. Then, the template distributions are also unfolded using the corresponding migration matrix. As shown in Sec 8.3, the χ^2 is calculated as:

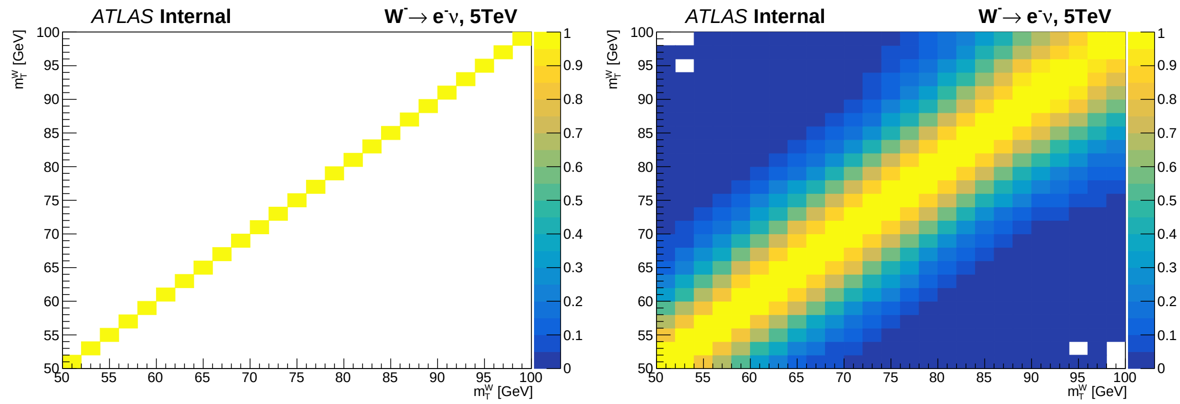
$$\chi^2 = (n_{\mathrm{data}}^{\mathrm{Unf}} - n_{\mathrm{template}}^{\mathrm{Unf}})^T \cdot (V)^{-1} \cdot (n_{\mathrm{data}}^{\mathrm{Unf}} - n_{\mathrm{template}}^{\mathrm{Unf}}), \quad (8.6)$$



(A) (B)

FIGURE 8.7: Distributions of m_{T}^{W} with the corresponding templates at the reconstructed level (A) and at the unfolded level (B).

where the total covariance matrix V is considered as the sum of the covariance matrix of the unfolded data and the unfolded templates, $V = V_{\mathrm{data}} + V_{\mathrm{template}}$. The same procedure is applied also separately for p_{T}^{ℓ} distributions.



(A) (B)

FIGURE 8.8: Correlation matrix for the statistical uncertainty for m_{T}^{W} distribution at the reconstructed level (A) and at the unfolded level (B), the correlation between bins is introduced by the unfolding procedure.

Table 8.7 shows an example of the statistical uncertainties calculated using the unfolded and the reconstructed distributions, for different regions of η. In general, the results are similar for the statistical uncertainty for both methods. Then, the correlation between p_{T}^{ℓ} and m_{T}^{W} is evaluated using the unfolded toys as described for the templates method in Sec. 8.4. In general, using the unfolded distribution does not change the results for the statistical uncertainties, but it is not the case when we treat the systematic uncertainties because of the statistical fluctuations in p_{T}^{ℓ} and m_{T}^{W} and also because of the bias that we introduce with the unfolding procedure.

TABLE 8.7: Statistical uncertainties (in MeV) on the M_W measurement using the unfolded and reconstructed distributions, for p_{T}^{ℓ} and m_{T}^{W} separately, using different regions of η_ℓ at 5 TeV.

η_ℓ range	[0, 0.6]		[0.6, 1.2]		[1.2, 1.8]		[1.8, 2.4]		[0, 2.4]	
Kinematic distribution	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}	p_{T}^{ℓ}	m_{T}^{W}
Channel	$W^- \to e^-\bar{\nu}_e$, 5 TeV									
Stat [Unfolded]	55	49	58	53	78	70	61	71	32	29
Stat [Reconstructed]	54	49	57	53	76	71	62	71	31	29

Chapter 9

Conclusion

This thesis describes mainly my personal work on the in-situ calibration of the electromagnetic calorimeter of the ATLAS detector, and on the measurement of W boson properties using low pile-up data set collected by ATLAS in 2017 and 2018 during Run 2 corresponding to an integrated luminosity of 258 pb^{-1} at $\sqrt{s} = 5$ TeV and 340 pb^{-1} at $\sqrt{s} = 13$ TeV.

The in-situ calibration is the last step in the calibration procedure. It is based on the $Z \to e^+e^-$ event samples and aims for correcting for residual difference in the energy scale and resolution between data and MC simulation. The calibration using the template method developed for Run 1 analysis has been performed for all nominal data samples taken in Run 2 in 2015, 2016, 2017 and 2018 under different running conditions. The number of interactions per bunch crossing μ of these nominal data samples varies typically between 10 and 70, being lower in 2015 and 2016 and higher in 2017 and 2018. Year dependence of the calibration corrections has been studied. The same procedure has also been applied to the low pile-up data showing larger uncertainties due to the limited statistics of the samples. The low pile-up data have a μ value around 2. We have thus developed a new approaching by studying the μ dependence of the energy scale correction of the nominal data samples and extrapolating the correction of the nominal samples to $\mu \sim 2$ to be compared with that obtained directly from the low pile-up data samples. It is found that the two sets of the corrections are consistent and the extrapolated correction has better precision even when the extrapolation uncertainties are taken into account.

The measurement of the W boson properties includes three parts. The first part corresponds to a measurement of the transverse momentum of the W boson, p_{T}^{W}. The modelling uncertainty of p_{T}^{W} by a theoretical extrapolation from Z-boson measurements has been one of the dominant systematic uncertainties of the previous mass determination of the W boson by ATLAS. A direct measurement of p_{T}^{W} would avoid such an extrapolation and the corresponding theoretical modelling uncertainty. The second part is on the measurement of the fiducial, single and double differential cross sections of the W boson production in the electron and muon decay channels at 5 and 13 TeV. The measurement has been compared with a NNLO QCD prediction using different PDF sets, showing its potential in constraining the uncertainty of the PDFs which was the dominant source for the determination of the W boson mass. The third part represents preliminary results for the W boson mass determination using the templates method and a new approach which relies on unfolded distributions. In this thesis, we focused on the measurement of the dominant experimental uncertainties. The final result for the W boson mass must take into account the theoretical and modeling uncertainties that are not studied in this thesis.

Appendix A

Control plots

Comparison data/simulation for p_T^ℓ

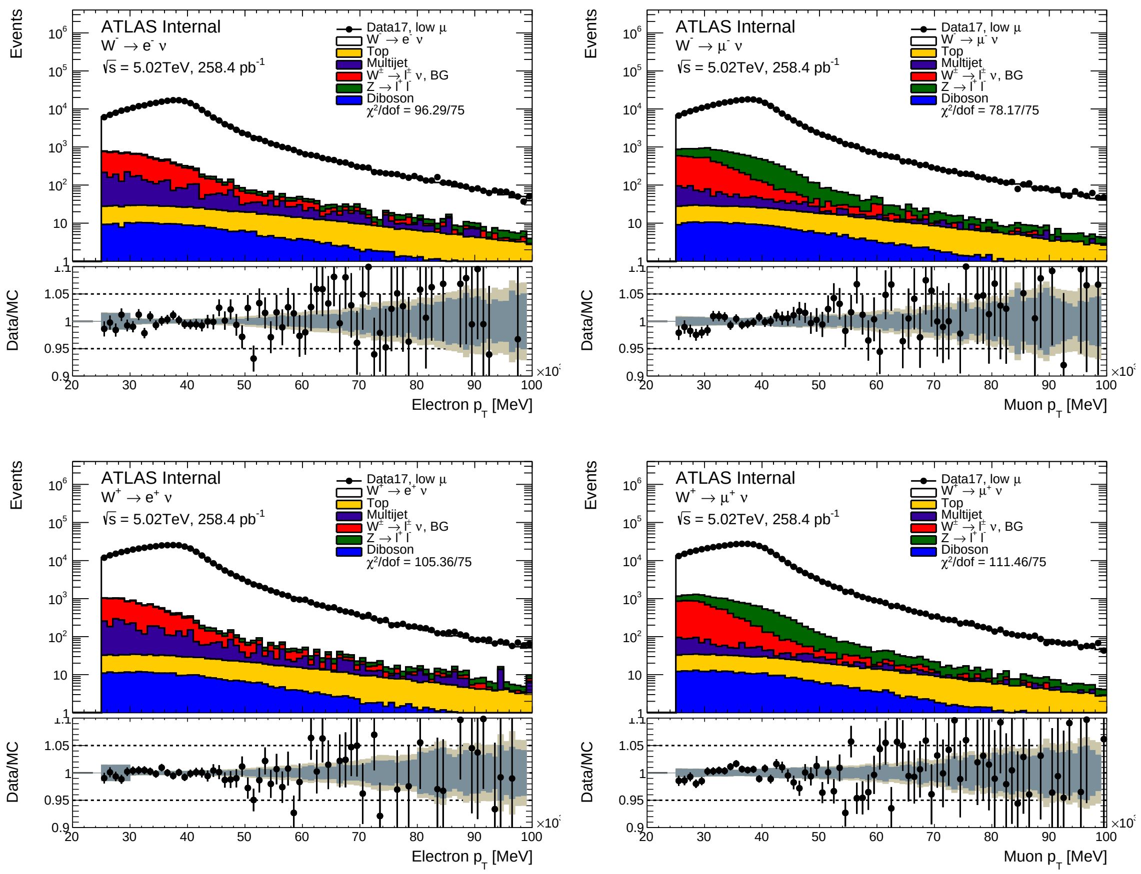


FIGURE A.1: Reconstructed p_T^ℓ distributions at detector level for $\sqrt{s} = 5$ TeV data set in the fiducial phase space. The signal and backgrounds are normalised to data. The low panel gives the ratio data/MC in each bin. The green band shows the statistical and systematic uncertainties added in quadrature. The total uncertainty is dominated by the uncertainty in the efficiency corrections applied to lepton.

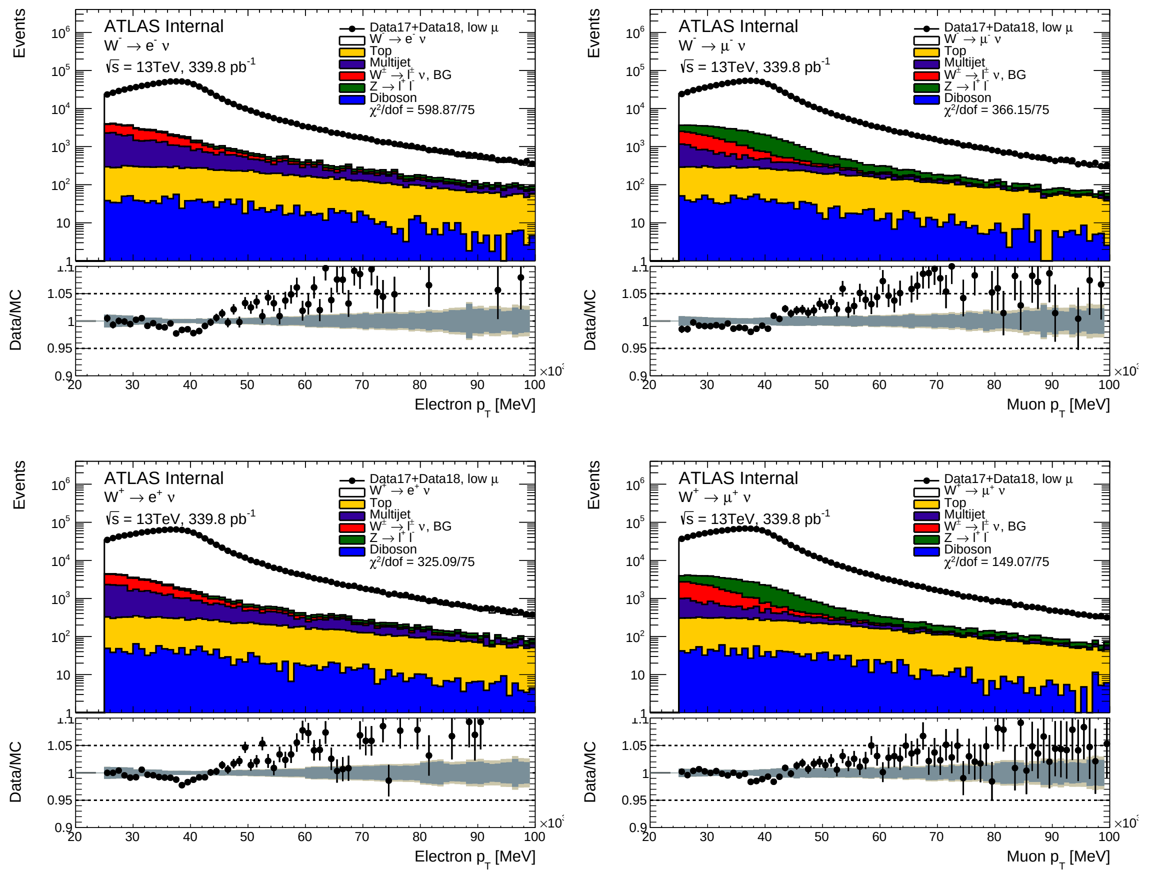


FIGURE A.2: Reconstructed p_{T}^{ℓ} distributions at the detector level for $\sqrt{s} = 13$ TeV data set in the fiducial phase space. The signal and backgrounds are normalised to data. The low panel gives the ratio data/MC in each bin. The green band shows the statistical and systematic uncertainties added in quadrature. The total uncertainty is dominated by the uncertainty on the efficiency corrections applied to lepton.

Comparison data/simulation for m_T^W

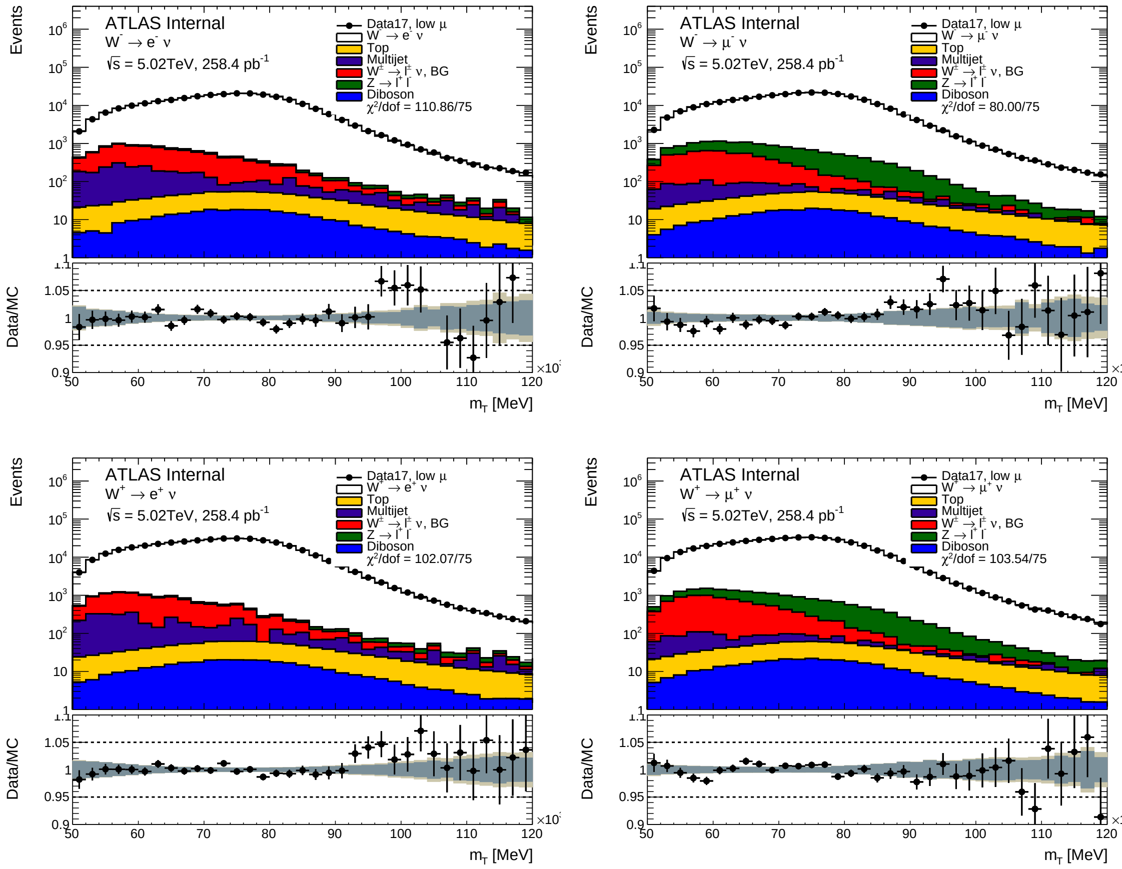


FIGURE A.3: Reconstructed m_T^W distributions at detector level for $\sqrt{s} = 5$ TeV data set in the fiducial phase space. The signal and backgrounds are normalised to data. The low panel gives the ratio data/MC in each bin. The green band shows the statistical and systematic uncertainties added in quadrature. The total uncertainty is dominated by the uncertainty on the efficiency corrections applied to lepton.

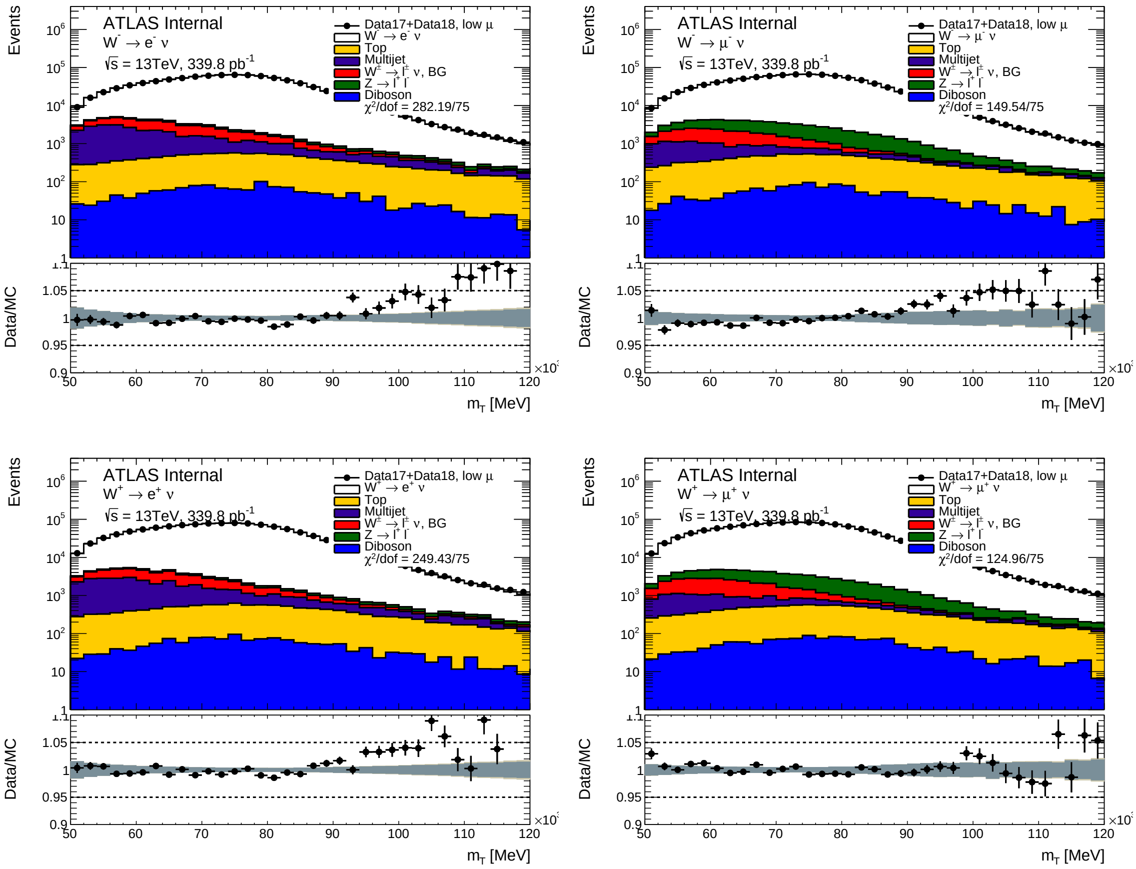


FIGURE A.4: Reconstructed m_{T}^{W} distributions at detector level for $\sqrt{s} = 13$ TeV data set in the fiducial phase space. The signal and backgrounds are normalised to data. The low panel gives the ratio data/MC in each bin. The green band shows the statistical and systematic uncertainties added in quadrature. The total uncertainty is dominated by the uncertainty on the efficiency corrections applied to lepton.

Comparison data/simulation for η_ℓ

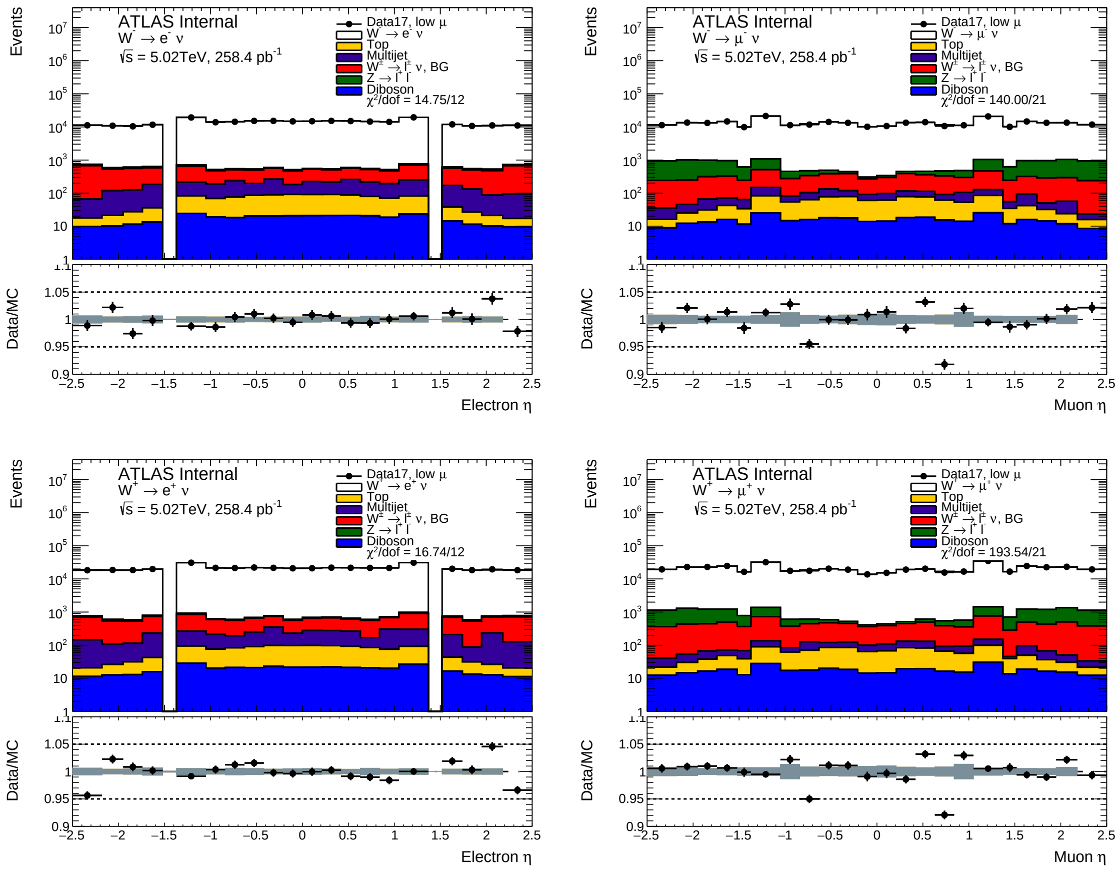


FIGURE A.5: Reconstructed η_ℓ distributions at detector level for $\sqrt{s} = 5$ TeV data set in the fiducial phase space. The signal and backgrounds are normalised to data. The low panel gives the ratio data/MC in each bin. The green band shows the statistical and systematic uncertainties added in quadrature. The total uncertainty is dominated by the uncertainty on the efficiency corrections applied to lepton.

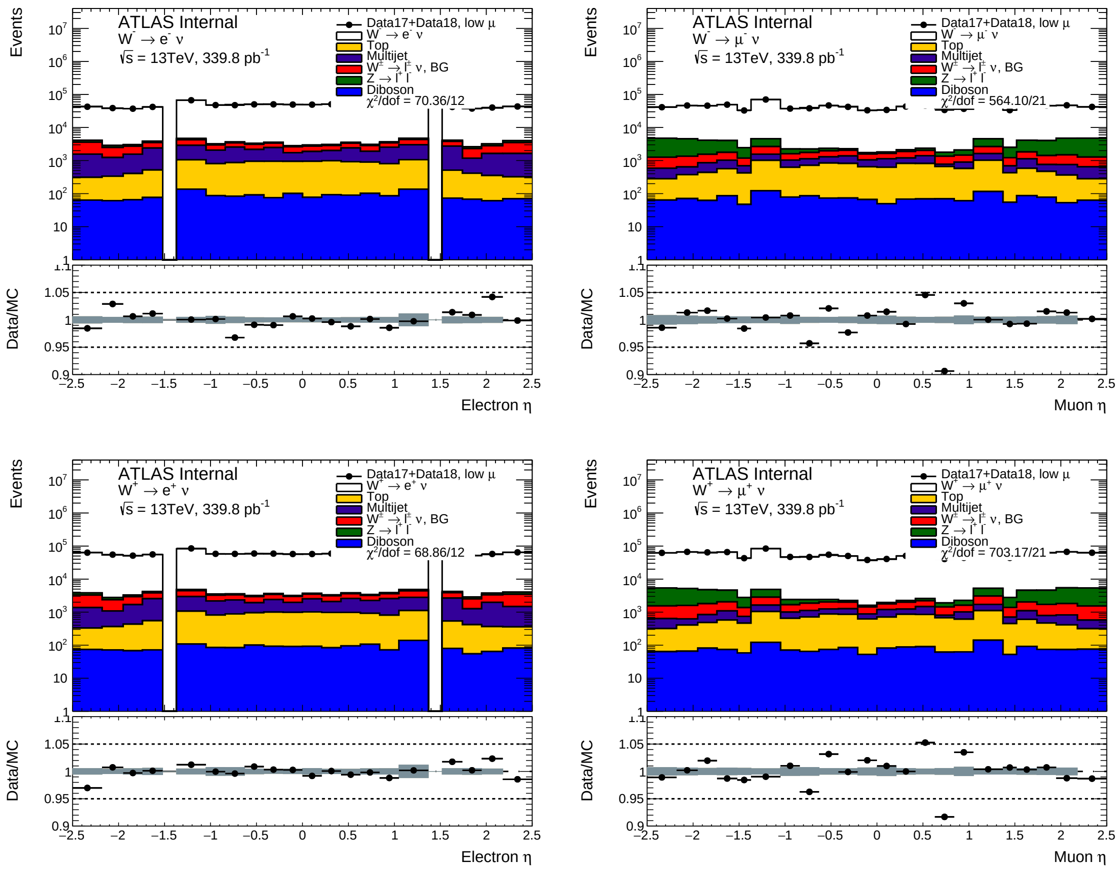


FIGURE A.6: Reconstructed η_ℓ distributions at detector level for $\sqrt{s} = 13$ TeV data set in the fiducial phase space. The signal and backgrounds are normalised to data. The low panel gives the ratio data/MC in each bin. The green band shows the statistical and systematic uncertainties added in quadrature. The total uncertainty is dominated by the uncertainty on the efficiency corrections applied to lepton.

Appendix B

Breakdown of uncertainties

Uncertainties in the measurement of p_T^ℓ at detector level

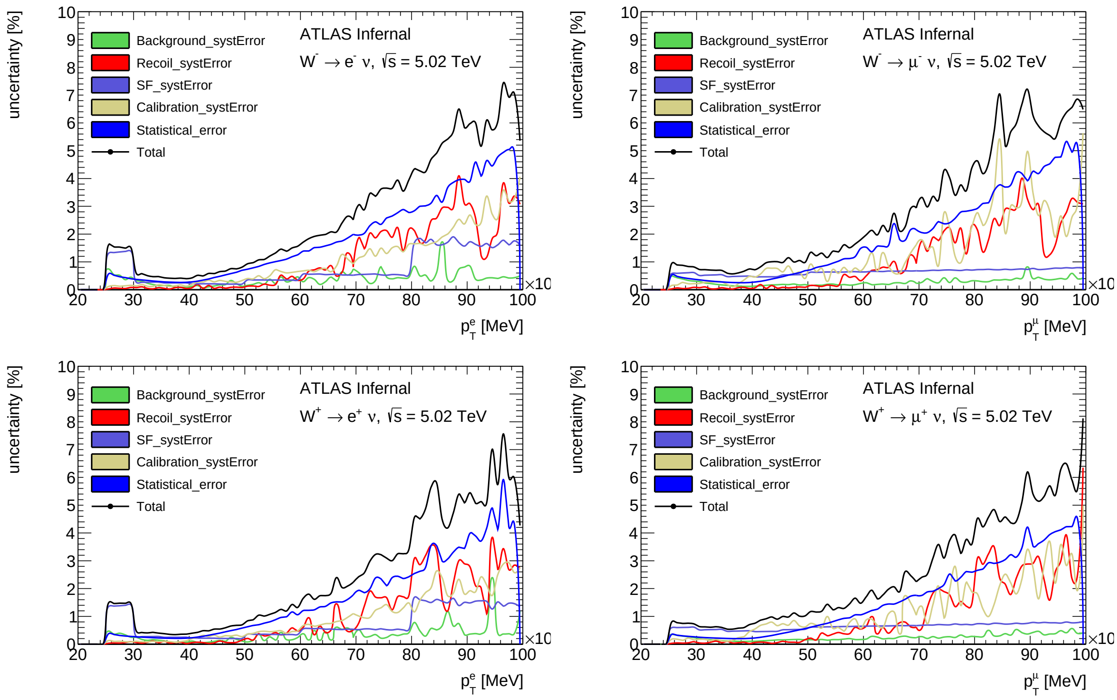


FIGURE B.1: Different sources of uncertainties on the measurement of p_T^ℓ distributions at detector level for the $\sqrt{s} = 5$ TeV data set. The total uncertainty is less than 2% at low p_T^ℓ region ($p_T^\ell < 50$ GeV) and around 6% for high p_T^ℓ region ($p_T^\ell \approx 100$ GeV). The total uncertainty is dominated by SF systematic uncertainty and the statistical uncertainty of data.

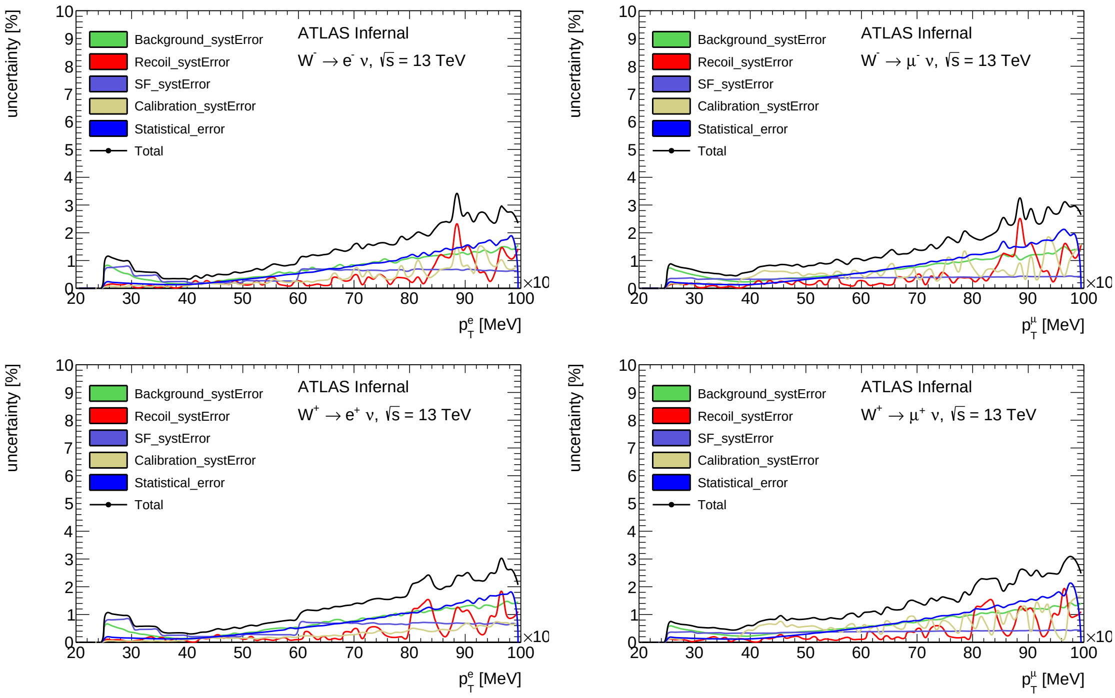


FIGURE B.2: Different sources of uncertainties on the measurement of p_{T}^{ℓ} distributions at detector level for the $\sqrt{s} = 13$ TeV data set. The total uncertainty is less than 1% at low p_{T}^{ℓ} region ($p_{\mathrm{T}}^{\ell} < 50$ GeV) and around 3% for high p_{T}^{ℓ} region ($p_{\mathrm{T}}^{\ell} \approx 100$ GeV). The total uncertainty is dominated by SF systematic uncertainty and the statistical uncertainty of data.

Uncertainties in the measurement of η_ℓ at detector level

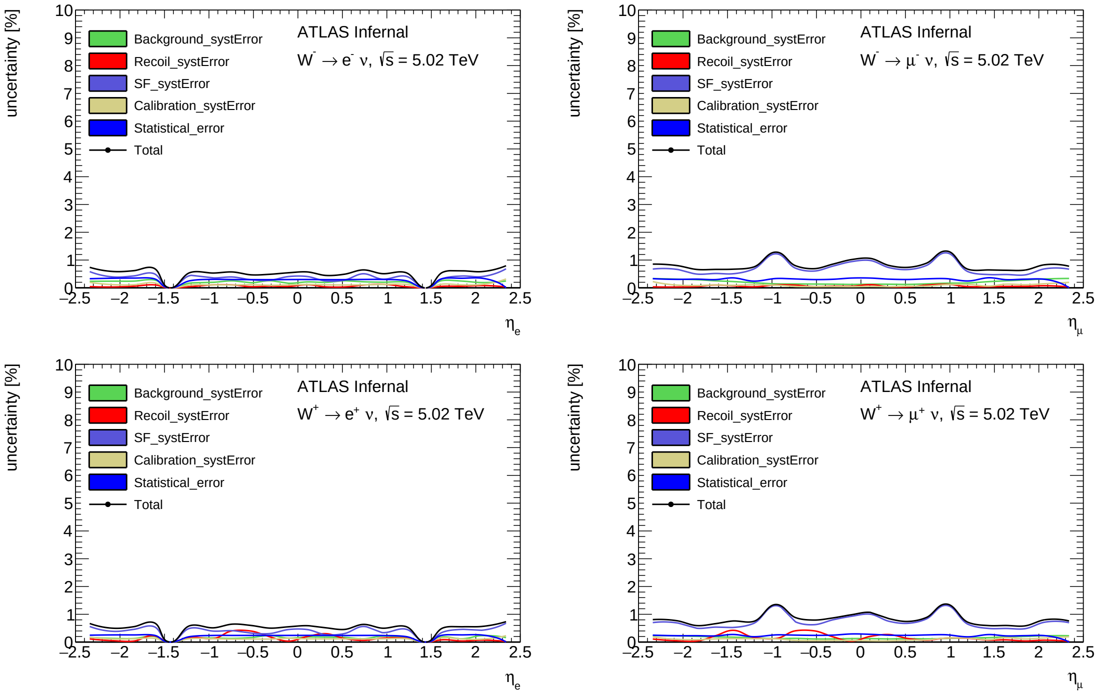


FIGURE B.3: Different sources of uncertainties on the measurement of η_ℓ distributions at detector level for the $\sqrt{s} = 5$ TeV data set. The total uncertainty is less than 1% and dominated by SF systematic uncertainty and the statistical uncertainty of data.

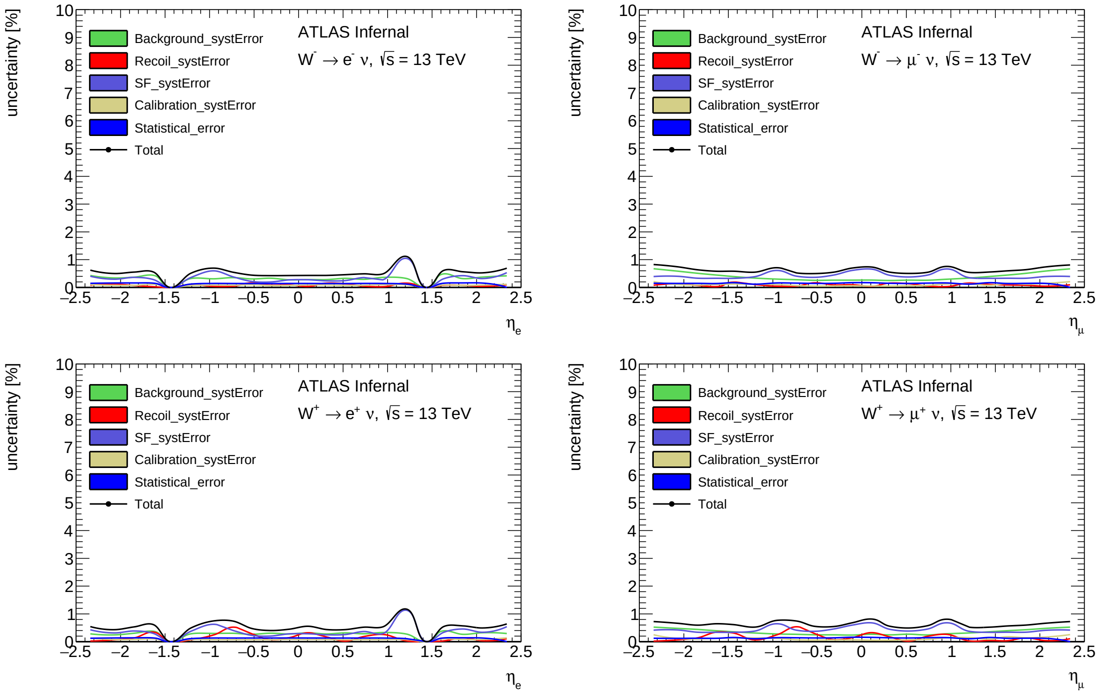


FIGURE B.4: Different sources of uncertainties on the measurement of η_ℓ distributions at the detector level for the $\sqrt{s} = 13$ TeV data set. The total uncertainty is less than 1% and dominated by SF systematic uncertainty and the statistical uncertainty of data.

Uncertainties in the measurement of p_{T}^{W} at detector level

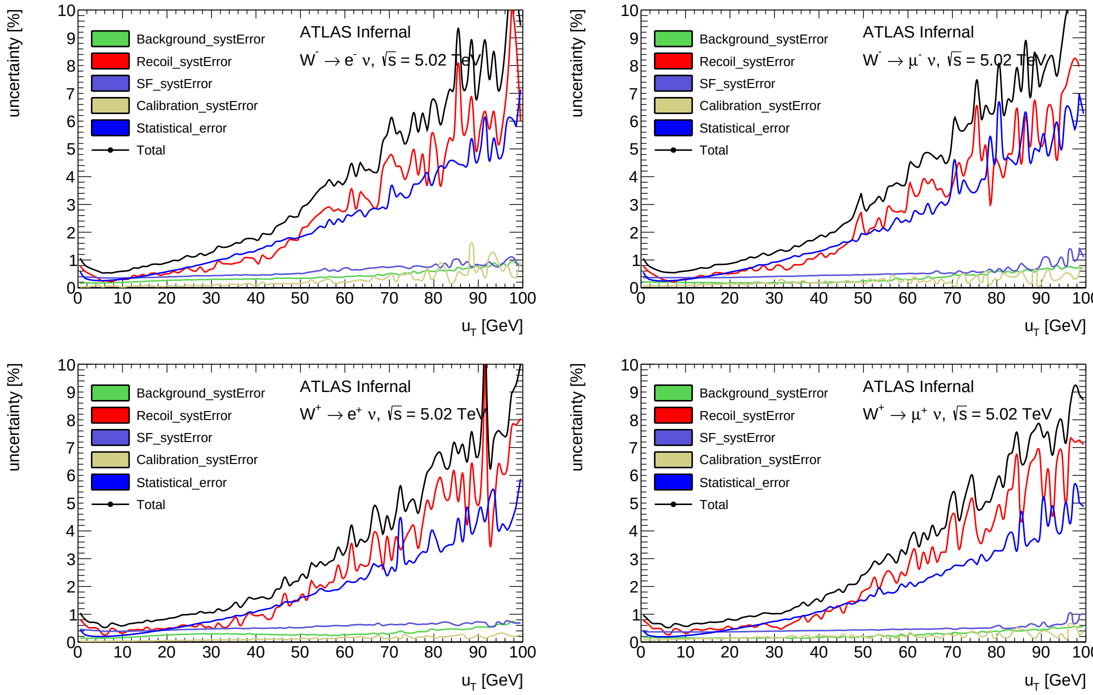


FIGURE B.5: Different sources of uncertainties in the measurement of p_{T}^{W} distributions at detector level for the $\sqrt{s} = 5$ TeV data set. The total uncertainty is less than 1% at low p_{T}^{W} region ($p_{\mathrm{T}}^{W} < 30$ GeV) and around 5% for high p_{T}^{W} region ($p_{\mathrm{T}}^{W} \approx 100$ GeV). The total uncertainty is dominated by hadronic recoil calibration uncertainty and the statistical uncertainty of data.

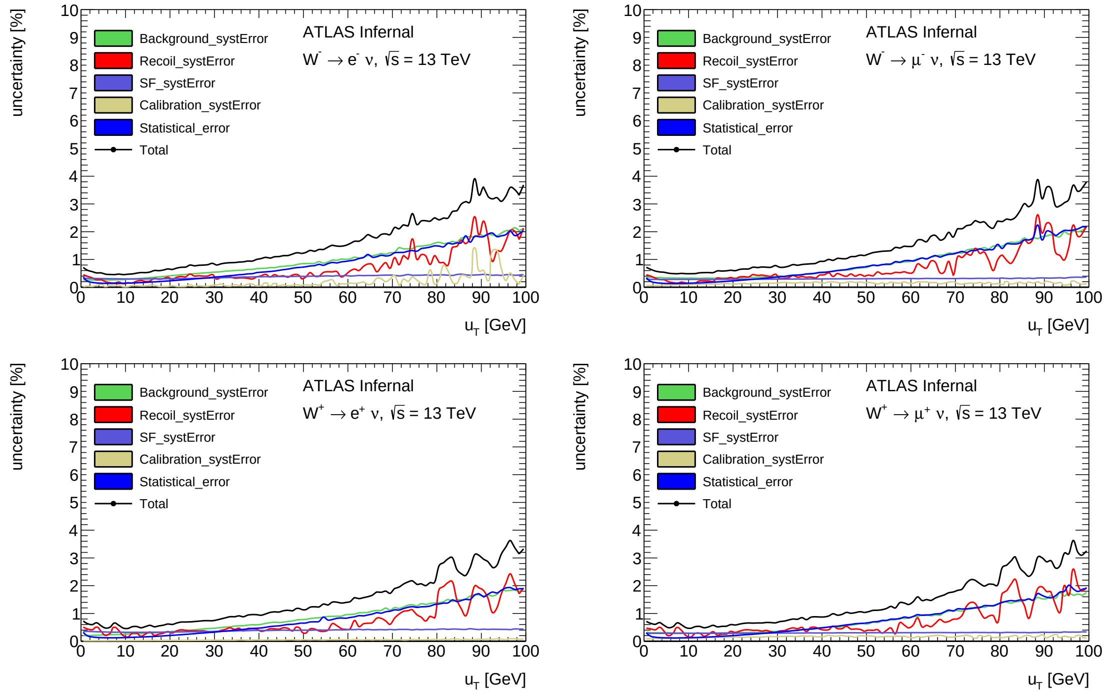


FIGURE B.6: Different sources of uncertainties on the measurement of p_{T}^{W} distributions at detector level for the $\sqrt{s} = 13$ TeV data set. The total uncertainty is less than 1% at low p_{T}^{W} region ($p_{\mathrm{T}}^{W} < 30$ GeV) and around 3% for high p_{T}^{W} region ($p_{\mathrm{T}}^{W} \approx 100$ GeV). The total uncertainty is dominated by hadronic recoil calibration uncertainty and background uncertainty (because of the large contributions of gauge-boson pair production and top-quark production in background).

Uncertainties in the measurement of p_{T}^{W} at unfolded level

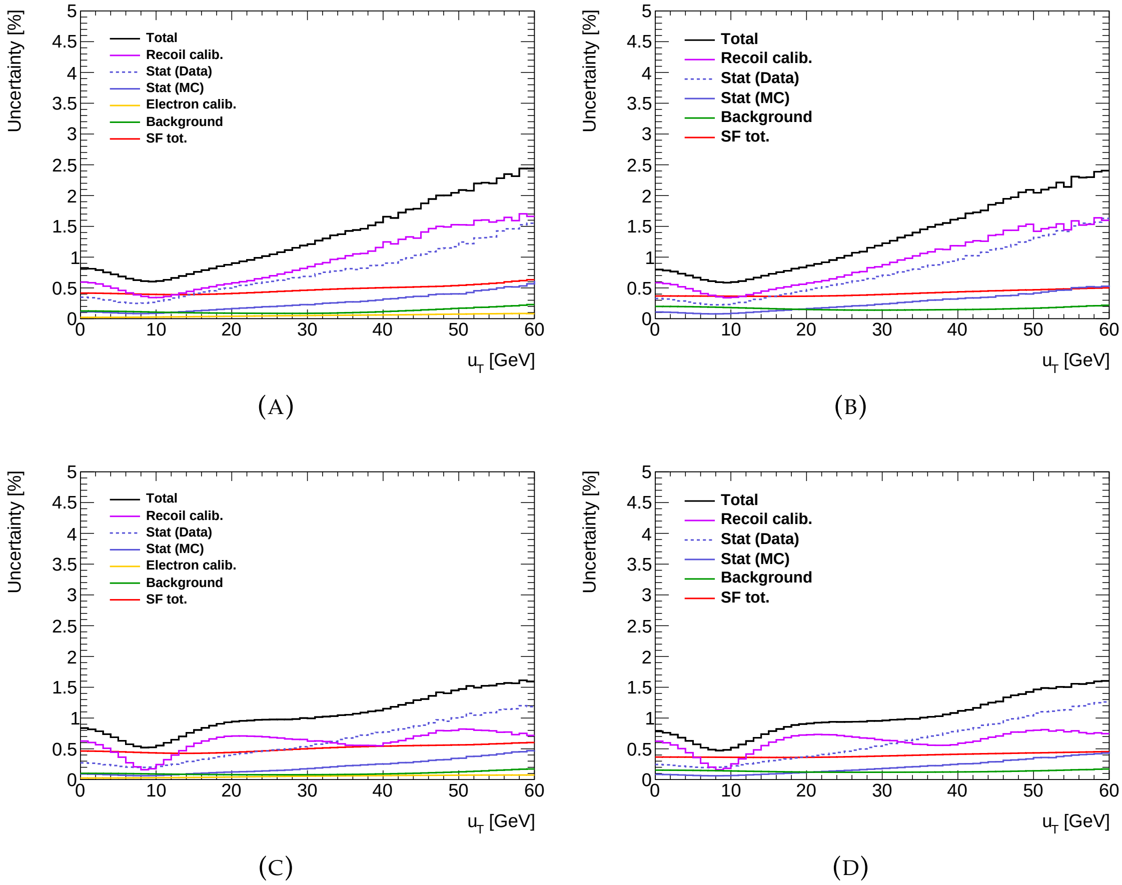
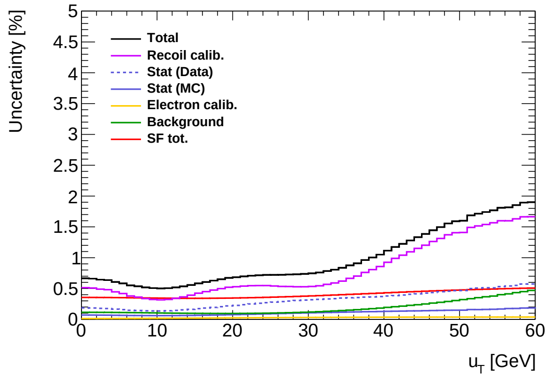
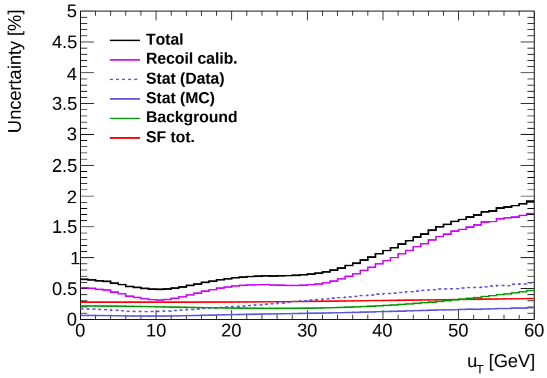


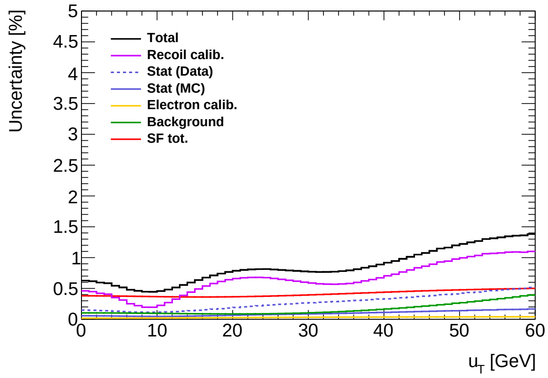
FIGURE B.7: Different sources of uncertainties on the measurement of unfolded p_{T}^{W} distributions for the $\sqrt{s} = 5$ TeV data set, for the electron (A, B) and muon (C, D) channels. The total uncertainty is less than 1% at low p_{T}^{W} region ($p_{\mathrm{T}}^{W} < 30$ GeV) and around 2% for high p_{T}^{W} region ($p_{\mathrm{T}}^{W} \approx 60$ GeV). The total uncertainty is dominated by hadronic recoil calibration uncertainty and the statistical uncertainty of data.



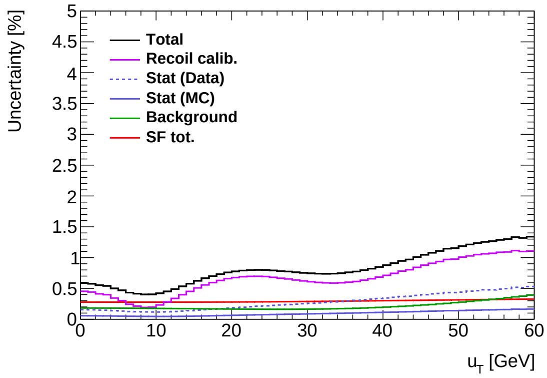
(A)



(B)



(C)



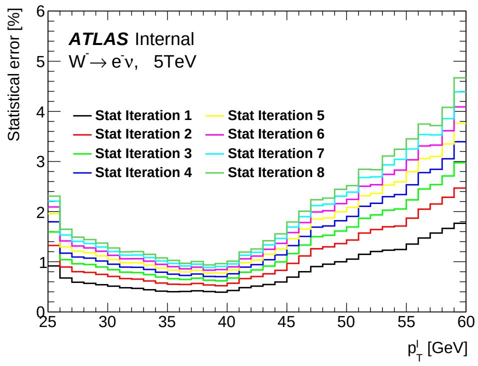
(D)

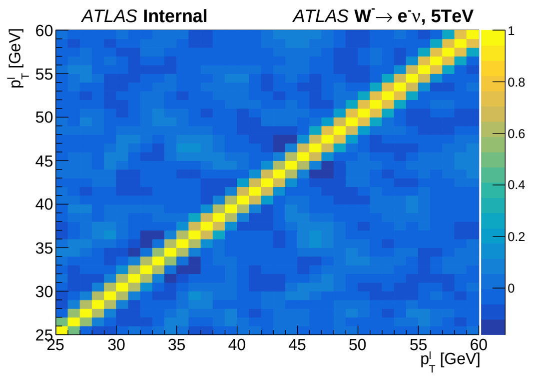
FIGURE B.8: Different sources of uncertainties on the measurement of unfolded p_{T}^{W} distributions at for the $\sqrt{s} = 13$ TeV data set, for the electron (A, B) and muon (C, D) channels. The total uncertainty is less than 1% at low p_{T}^{W} region ($p_{\mathrm{T}}^{W} < 30$ GeV) and around 1.5% for high p_{T}^{W} region ($p_{\mathrm{T}}^{W} \approx 60$ GeV). The total uncertainty is dominated by hadronic recoil calibration uncertainty and the background uncertainty.

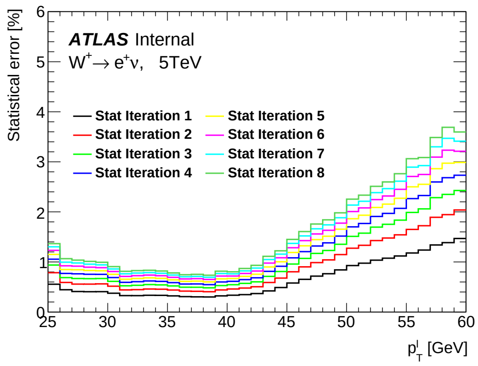
Appendix C

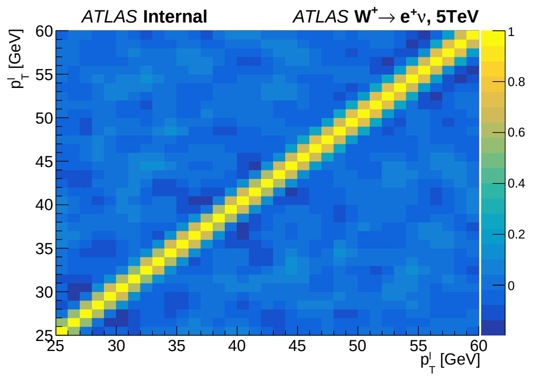
Uncertainties for the differential cross sections

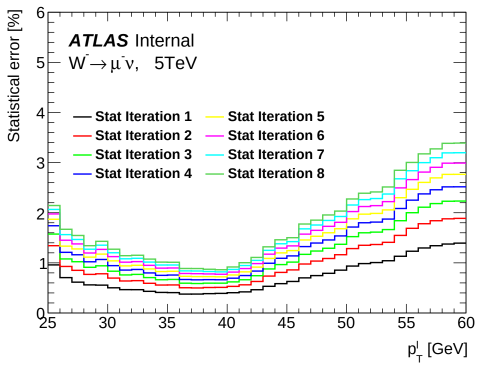
Statistical uncertainties with their correlation matrix of p_{T}^{ℓ} at 5 TeV

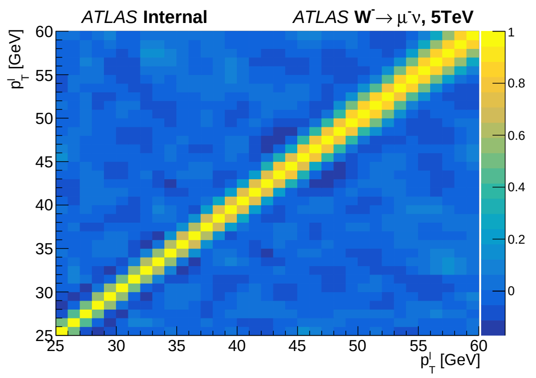


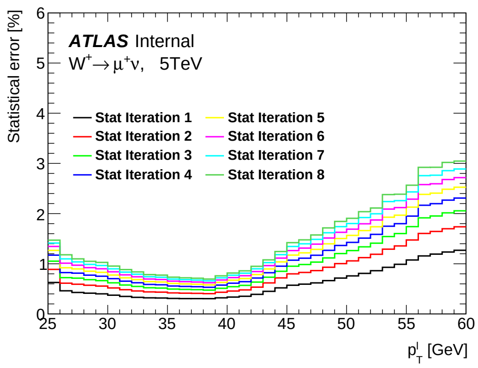


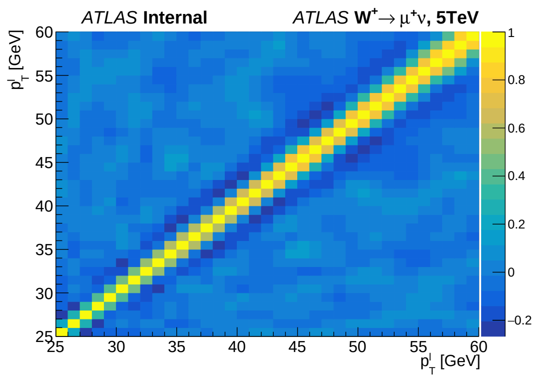




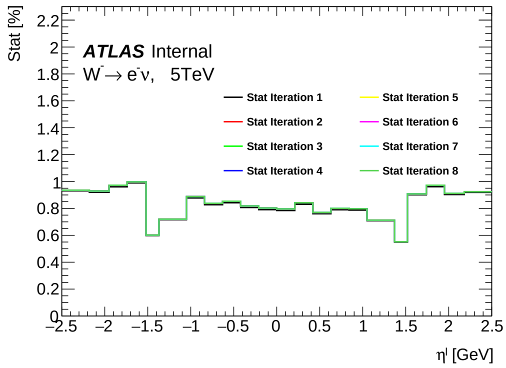


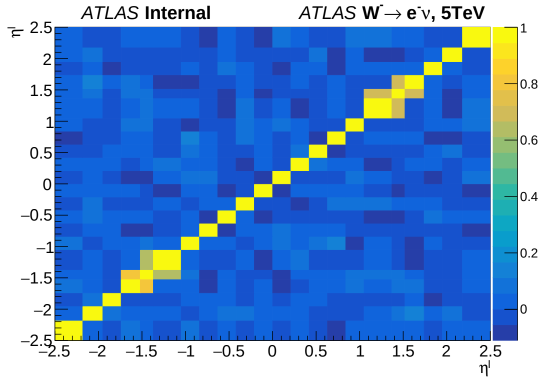


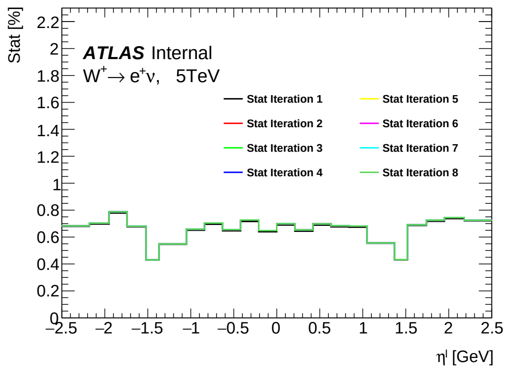


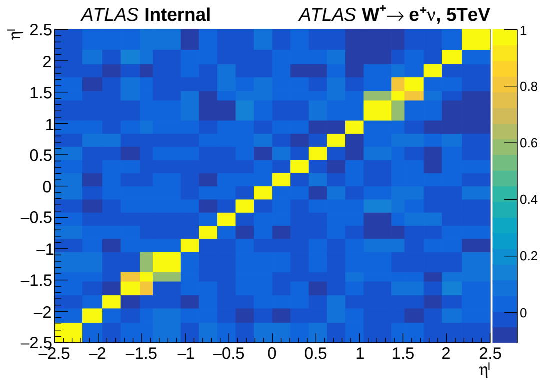


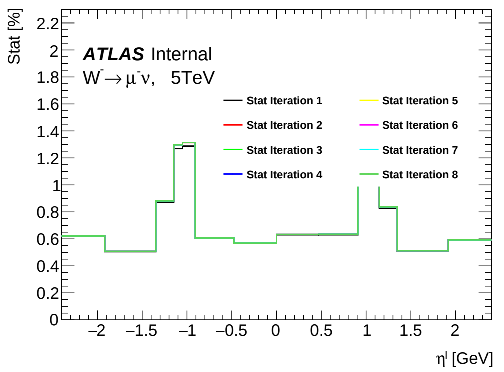
Statistical uncertainties with their correlation matrix of η_ℓ at 5 TeV

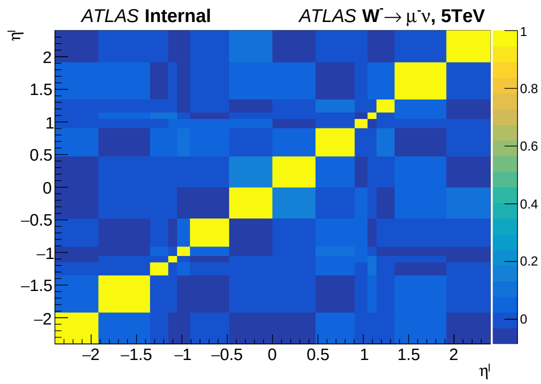


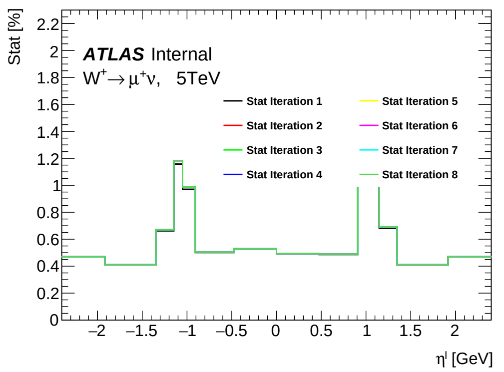


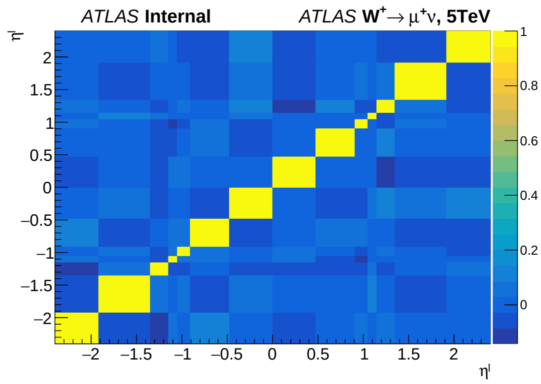




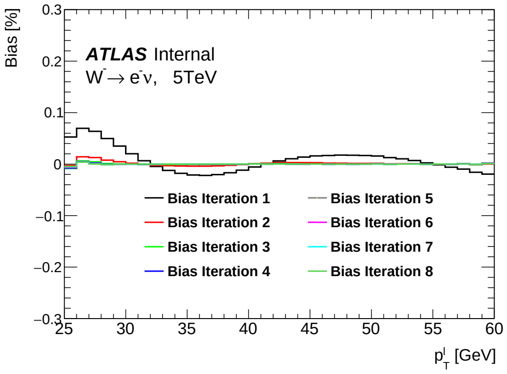


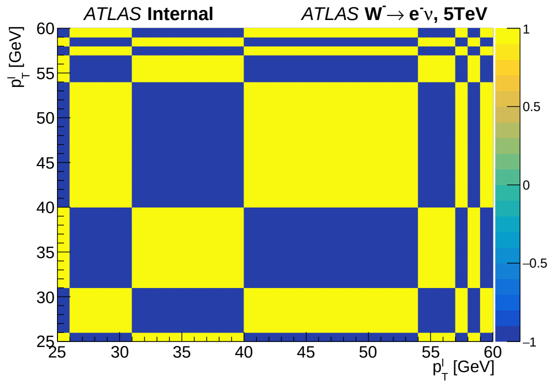


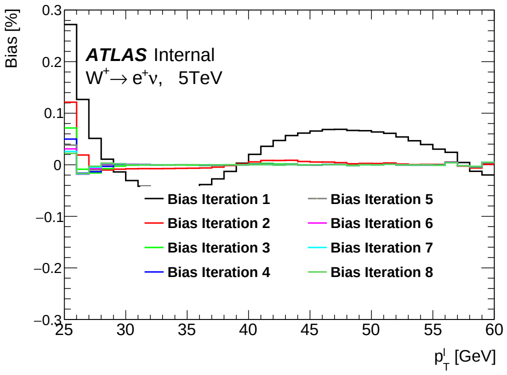


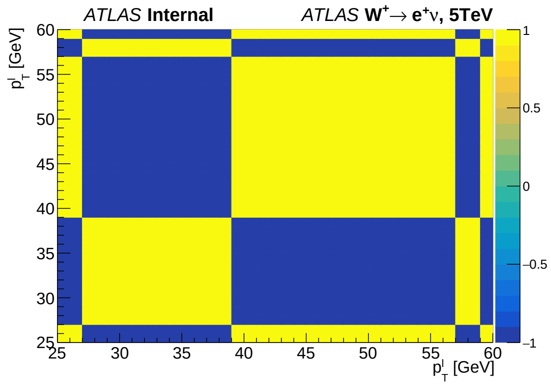


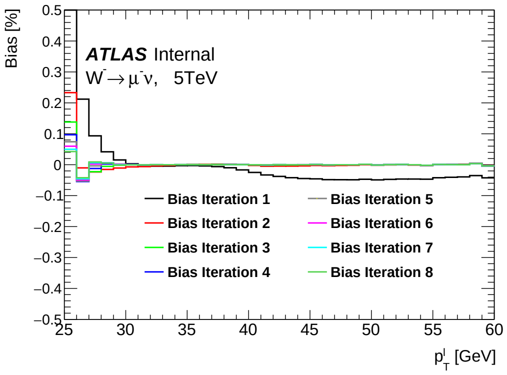
Unfolding bias with their correlation matrix of p_{T}^{ℓ} at 5 TeV

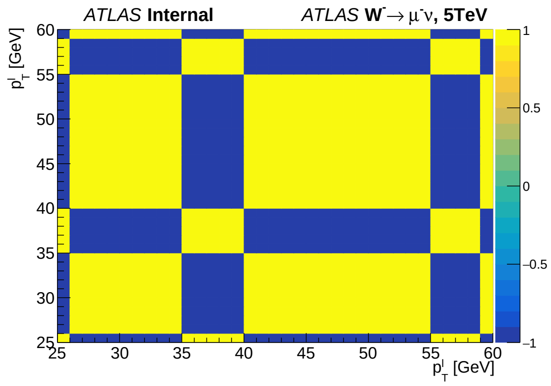


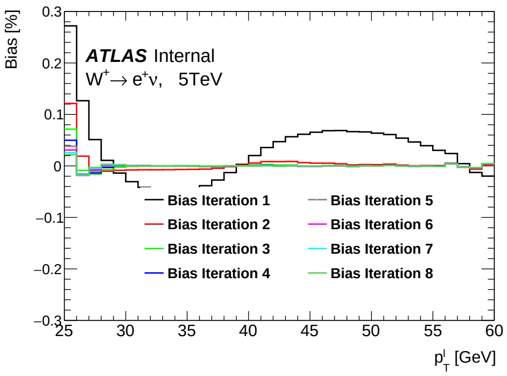


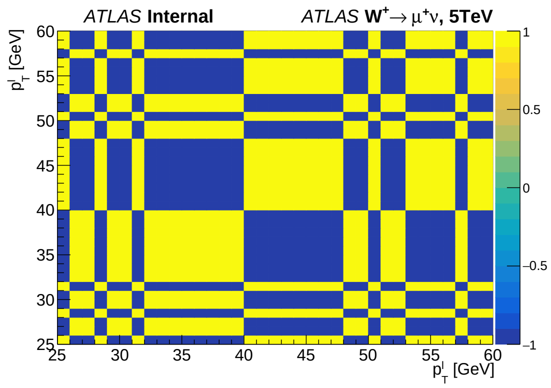




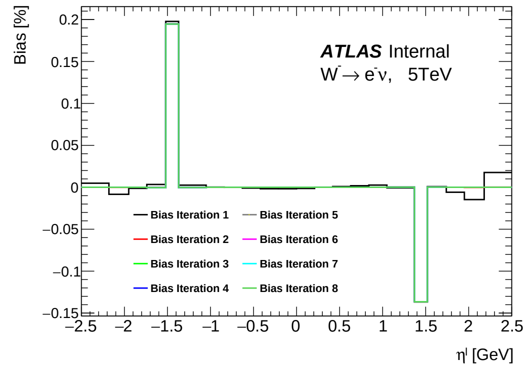


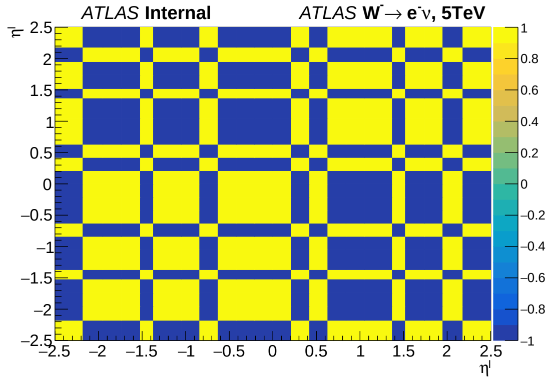


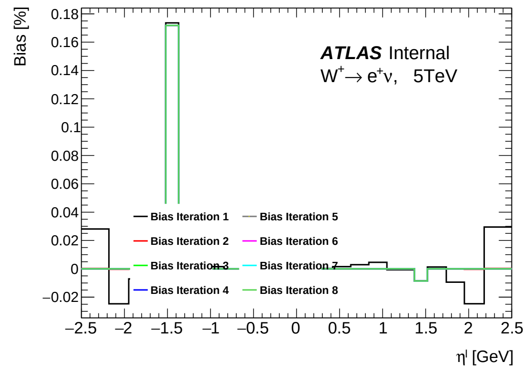


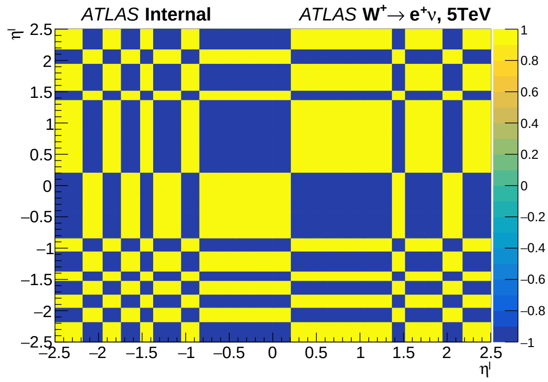


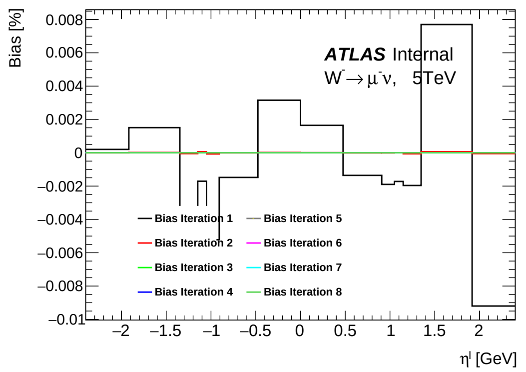
Unfolding bias with their correlation matrix of η_ℓ at 5 TeV

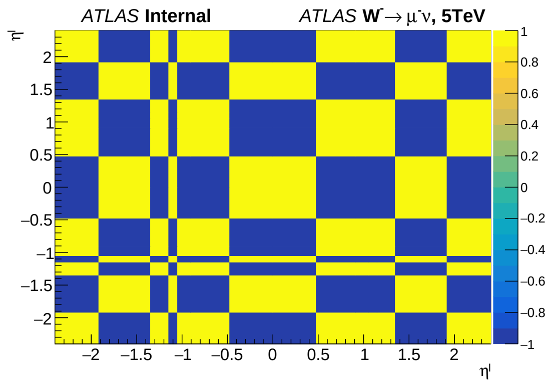


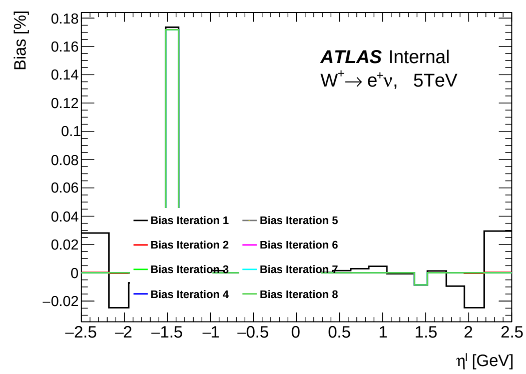


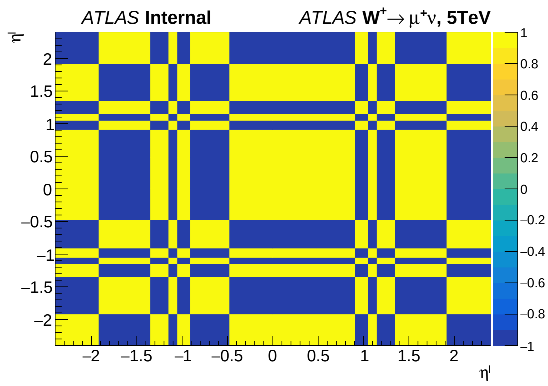












Statistical uncertainties for double differential cross sections at 5 TeV

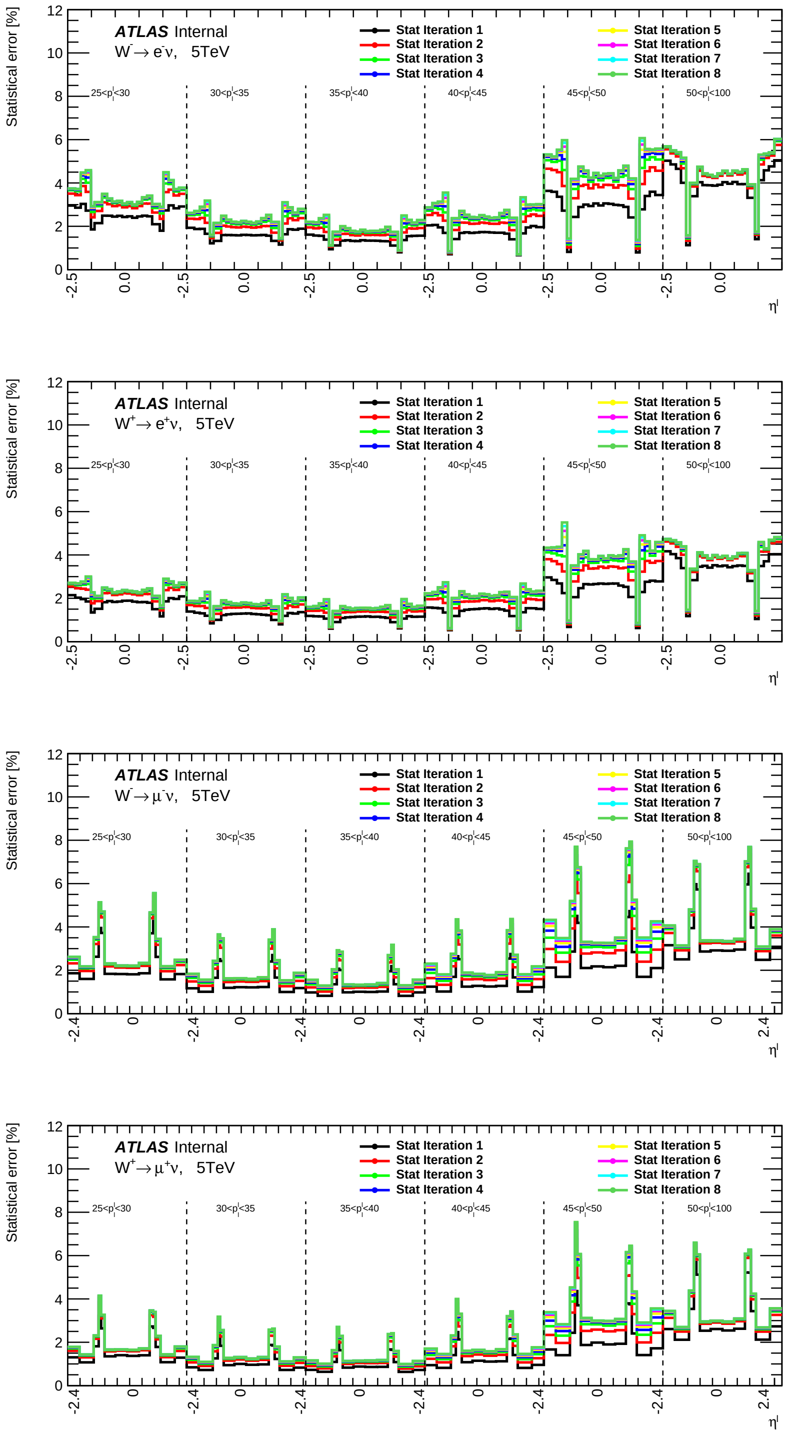


FIGURE C.1: Statistical uncertainties of unfolded distributions used for double differential cross sections at 5 TeV

Statistical uncertainties for double differential cross sections at 13 TeV

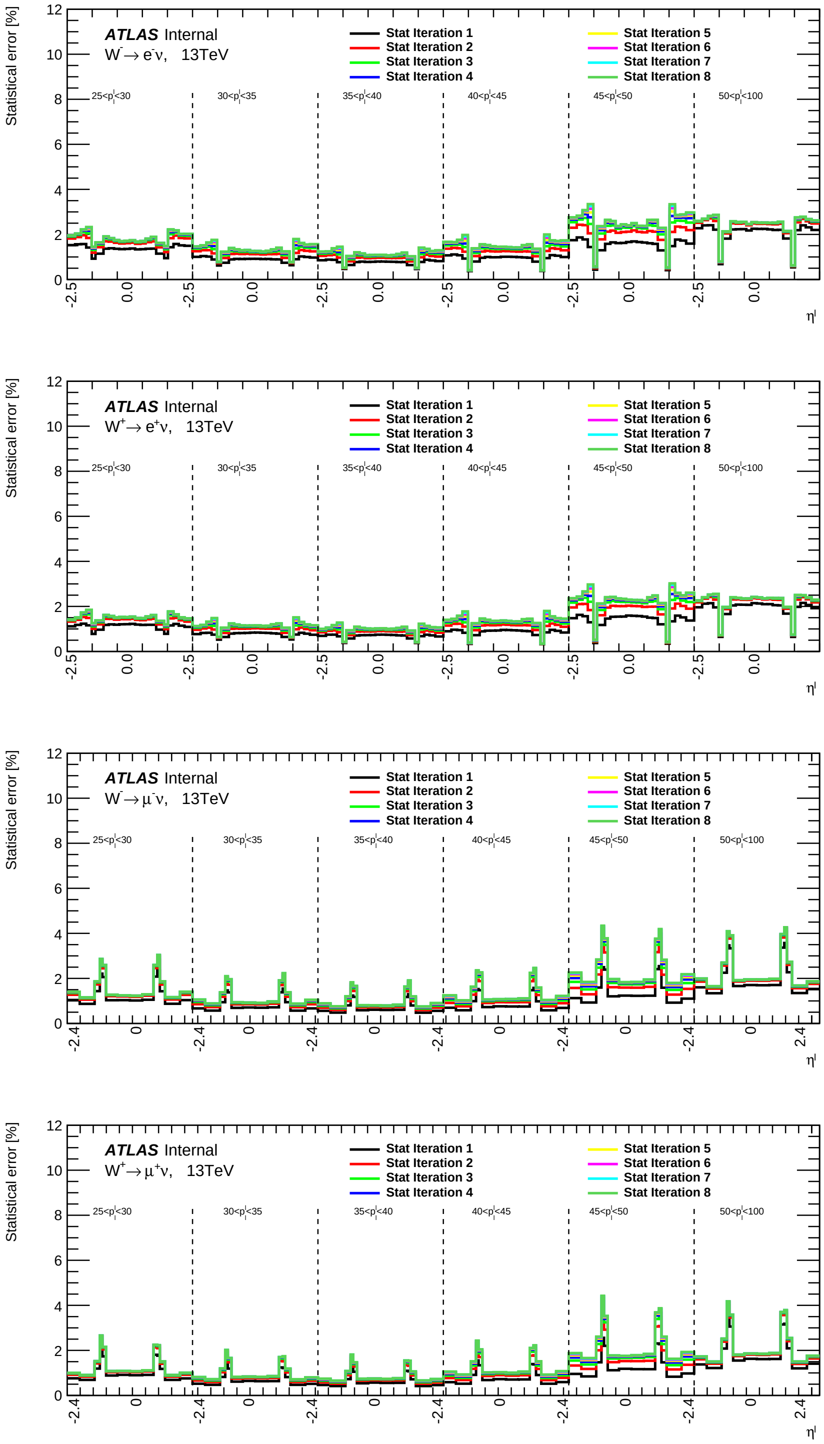


FIGURE C.2: Statistical uncertainties of unfolded distributions used for double differential cross sections at 13 TeV

Unfolding bias for double differential cross sections at 5 TeV

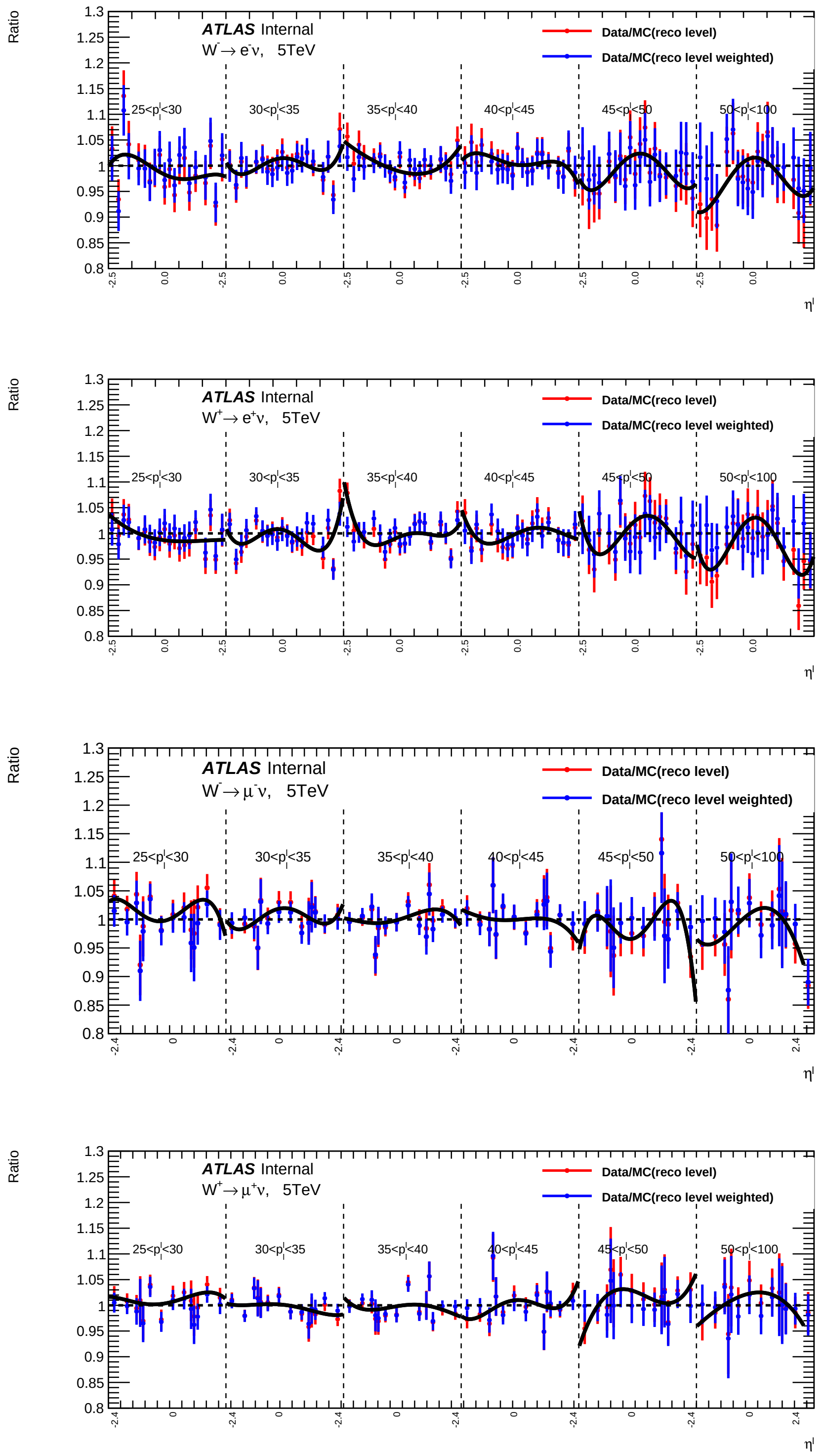


FIGURE C.3: Ratio data/MC used to calculate the unfolding bias for double differential cross sections at 5 TeV

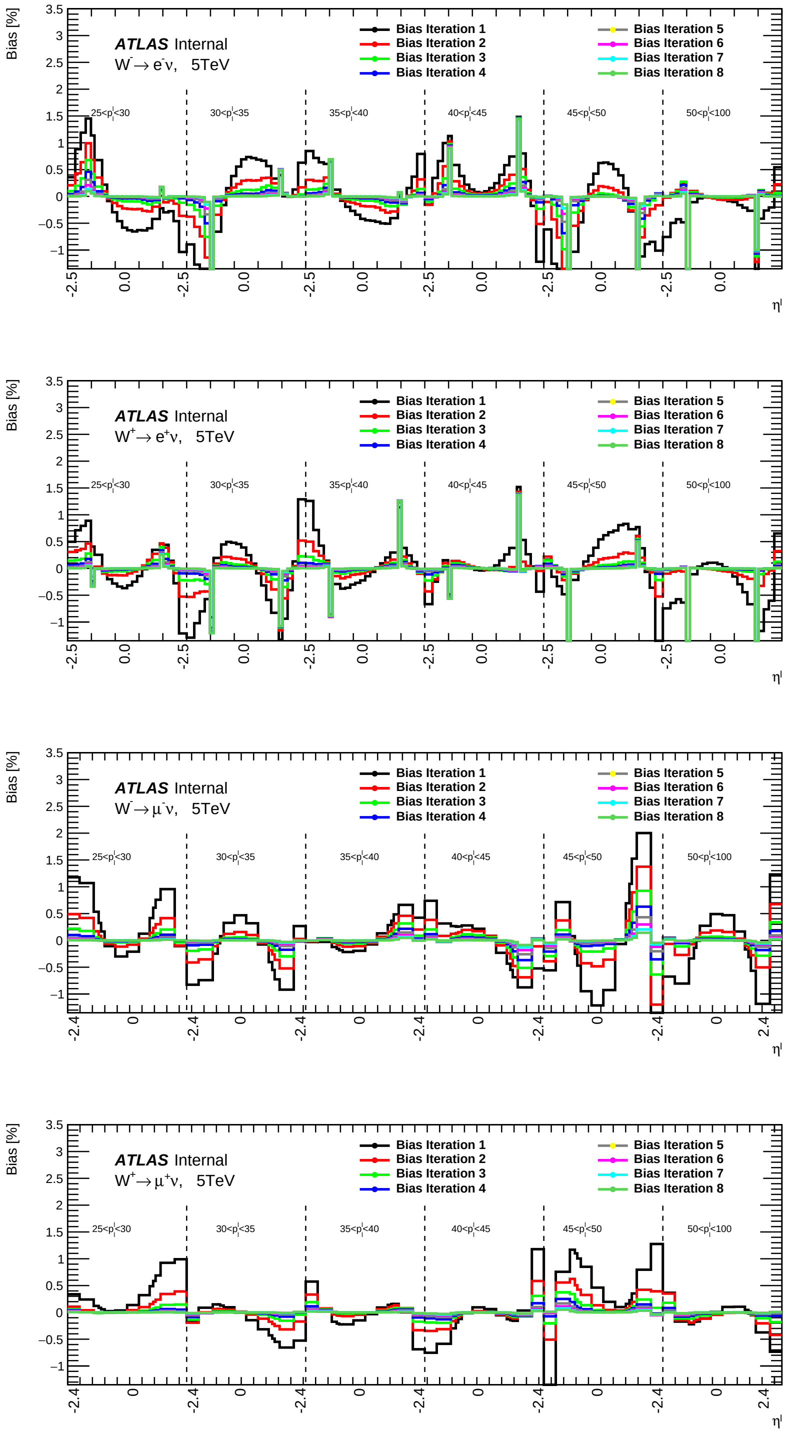


FIGURE C.4: Unfolding bias for double differential cross sections at 5 TeV

Unfolding bias for double differential cross sections at 13 TeV

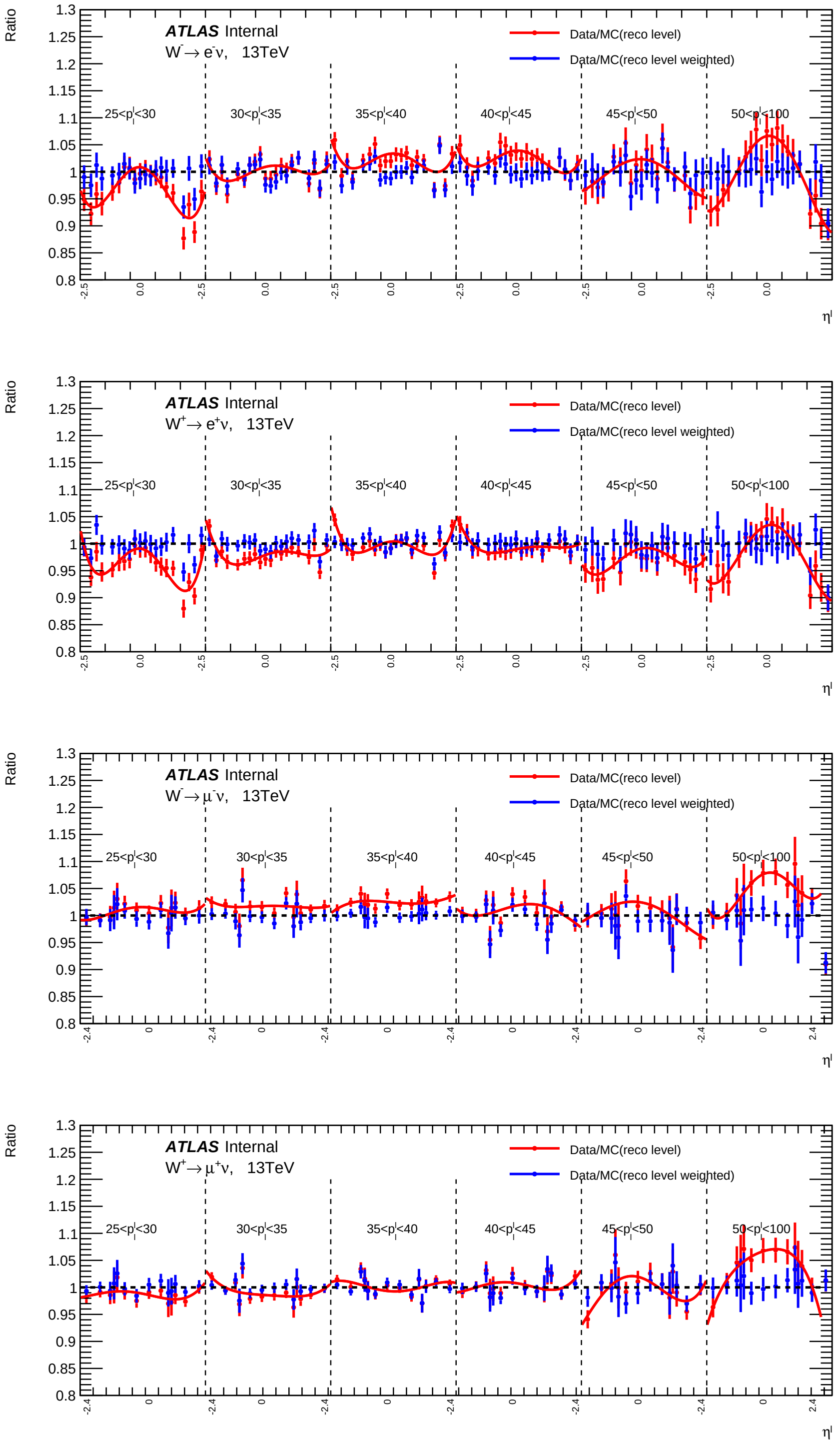


FIGURE C.5: Ratio data/MC used to calculate the unfolding bias for double differential cross sections at 13 TeV

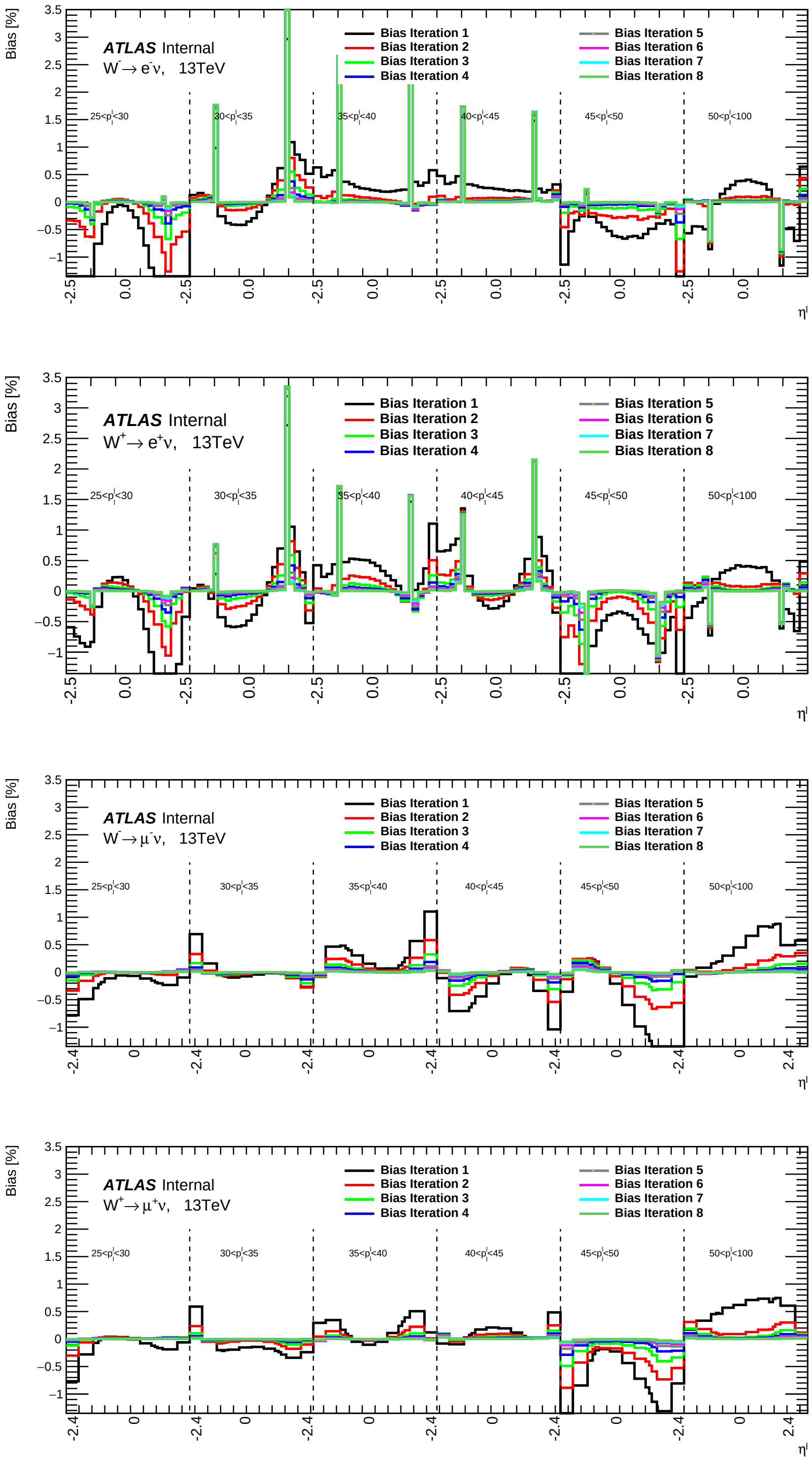


FIGURE C.6: Unfolding bias for double differential cross sections at 13 TeV

Bibliography

[1] M. Aaboud et al. "Electron efficiency measurements with the ATLAS detector using 2012 LHC proton–proton collision data". In: *The European Physical Journal C* 77.3 (Mar. 2017). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-017-4756-2. URL: http://dx.doi.org/10.1140/epjc/s10052-017-4756-2 (cit. on p. 138).

[2] Morad Aaboud et al. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data". In: *JINST* 14.03 (2019), P03017. DOI: 10.1088/1748-0221/14/03/P03017. arXiv: 1812.03848 [hep-ex] (cit. on p. 42).

[3] G. Aad et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003 (cit. on p. 26).

[4] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex] (cit. on p. 2).

[5] Timo Antero Aaltonen et al. "Combination of CDF and D0 W-Boson Mass Measurements". In: *Phys. Rev. D* 88.5 (2013), p. 052018. DOI: 10.1103/PhysRevD.88.052018. arXiv: 1307.7627 [hep-ex] (cit. on p. 11).

[6] Henso Abreu. "Measurement of the inclusive prompt photon cross section and preparation of the search of the Higgs boson decaying into two photons with the ATLAS detector at the LHC". PhD thesis. Orsay, 2011 (cit. on p. 23).

[7] Tim Adye. "Unfolding algorithms and tests using RooUnfold". In: *arXiv e-prints*, arXiv:1105.1160 (May 2011), arXiv:1105.1160. arXiv: 1105.1160 [physics.data-an] (cit. on pp. 61, 139).

[8] S. Agostinelli et al. "GEANT4–a simulation toolkit". In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8 (cit. on p. 33).

[9] S. Alekhin, J. Blümlein, and S. Moch. "NLO PDFs from the ABMP16 fit". In: *The European Physical Journal C* 78.6 (June 2018). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-018-5947-1. URL: http://dx.doi.org/10.1140/epjc/s10052-018-5947-1 (cit. on p. 94).

[10] M Aleksa and M Diemoz. *Discussion on the electromagnetic calorimeters of ATLAS and CMS*. Tech. rep. ATL-LARG-PROC-2013-002. Geneva: CERN, May 2013. URL: https://cds.cern.ch/record/1547314 (cit. on p. 20).

[11] J. Alitti et al. "An Improved determination of the ratio of W and Z masses at the CERN $\bar{p}p$ collider". In: *Phys. Lett. B* 276 (1992), pp. 354–364. DOI: 10.1016/0370-2693(92)90332-X (cit. on p. 11).

[12] Charalampos Anastasiou et al. "High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at next-to-next-to leading order". In: 69.9, 094008 (May 2004), p. 094008. DOI: 10.1103/PhysRevD.69.094008. arXiv: hep-ph/0312266 [hep-ph] (cit. on p. 91).

[13] Nansi Andari. "Observation of a BEH-like boson decaying into two photons with the ATLAS detector at the LHC". PhD thesis. Orsay, LAL, Sept. 2012 (cit. on pp. 23, 29).

[14] Nansi Andari. *W mass measurement*. Tech. rep. ATL-PHYS-PROC-2017-051. Geneva: CERN, May 2017. DOI: oai:cds.cern.ch:2264497. URL: https://cds.cern.ch/record/2264497 (cit. on p. 23).

[15] Nansi Andari et al. *Electron and photon energy calibration with the ATLAS detector using 2015-2017 LHC proton-proton collision data*. Tech. rep. ATL-COM-PHYS-2018-1720. Geneva: CERN, Dec. 2018. URL: https://cds.cern.ch/record/2651890 (cit. on pp. 37, 42, 48, 55).

[16] Nansi Andari et al. *Measurement of m_W at 7 TeV: Physics modeling*. Tech. rep. ATL-COM-PHYS-2014-1436. Geneva: CERN, Nov. 2014. URL: https://cds.cern.ch/record/1966965 (cit. on p. 12).

[17] Nansi Andari et al. *Measurement of m_W with 7 TeV data: W boson mass measurement*. Tech. rep. ATL-COM-PHYS-2014-1569. Geneva: CERN, Dec. 2014. URL: https://cds.cern.ch/record/1976186 (cit. on pp. 11, 12, 131).

[18] Samir. S. Arfaoui. "The ATLAS liquid argon calorimeter high-voltage system: commissioning, optimisation, and LHC relative luminosity measurement". Theses. Université de la Méditerranée - Aix-Marseille II, Oct. 2011. URL: https://tel.archives-ouvertes.fr/tel-00658194 (cit. on p. 34).

[19] Theodoros Argyropoulos et al. "Cathode strip chambers in ATLAS: Installation, commissioning and in situ performance". In: *IEEE Trans. Nucl. Sci.* 56 (2009), pp. 1568–1574. DOI: 10.1109/TNS.2009.2020861 (cit. on p. 26).

[20] A Armbruster et al. *Practical considerations for unfolding*. Tech. rep. ATL-COM-PHYS-2014-277. Geneva: CERN, Apr. 2014. URL: https://cds.cern.ch/record/1694351 (cit. on p. 62).

[21] G. Arnison et al. "Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at s**(1/2) = 540-GeV". In: *Phys. Lett. B* 122 (1983), pp. 103–116. DOI: 10.1016/0370-2693(83)91177-2 (cit. on p. 11).

[22] ATLAS Collaboration. "Electron efficiency measurements with the ATLAS detector using 2012 LHC proton-proton collision data". In: *arXiv e-prints*, arXiv:1612.01456 (Dec. 2016), arXiv:1612.01456. arXiv: 1612.01456 [hep-ex] (cit. on pp. 79, 99).

[23] ATLAS Collaboration. "Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton-proton collision data". In: *arXiv e-prints*, arXiv:1407.3935 (July 2014), arXiv:1407.3935. arXiv: 1407.3935 [hep-ex] (cit. on p. 11).

[24] *ATLAS detector and physics performance: Technical Design Report, 1*. Technical Design Report ATLAS. Geneva: CERN, 1999. URL: https://cds.cern.ch/record/391176 (cit. on p. 17).

[25] *ATLAS Liquid Argon Calorimeter Phase-I Upgrade Technical Design Report*. Tech. rep. CERN-LHCC-2013-017. ATLAS-TDR-022. Final version presented to December 2013 LHCC. Sept. 2013. URL: https://cds.cern.ch/record/1602230 (cit. on p. 24).

[26] *ATLAS liquid-argon calorimeter: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1996. URL: https://cds.cern.ch/record/331061 (cit. on pp. 22, 23).

[27] Hicham Atmani et al. *Measurement of the p_{T} spectrum of W- and Z-bosons produced in pp collisions at $\sqrt{s} = 5$ TeV and 13 TeV in low-pileup runs*. Tech. rep. ATL-COM-PHYS-2018-1084. Geneva: CERN, July 2018. URL: https://cds.cern.ch/record/2632159 (cit. on pp. 97, 99).

[28] M. Baak et al. "The global electroweak fit at NNLO and prospects for the LHC and ILC". In: *European Physical Journal C* 74, 3046 (Sept. 2014), p. 3046. DOI: 10.1140/epjc/s10052-014-3046-5. arXiv: 1407.3792 [hep-ph] (cit. on p. 10).

[29] 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". In: *Phys. Lett. B* 122 (1983), pp. 476–485. DOI: 10.1016/0370-2693(83)91605-2 (cit. on p. 11).

[30] J Beacham et al. "Physics beyond colliders at CERN: beyond the Standard Model working group report". In: *Journal of Physics G: Nuclear and Particle Physics* 47.1 (Dec. 2019), p. 010501. ISSN: 1361-6471. DOI: 10.1088/1361-6471/ab4cd2. URL: http://dx.doi.org/10.1088/1361-6471/ab4cd2 (cit. on p. 15).

[31] Cyril Pascal Becot. "Diphoton lineshape of the BEH boson using the ATLAS detector at the LHC: calibration, mass, width and interferences". PhD thesis. Diderot U., Paris, Sept. 2015 (cit. on pp. 23, 42).

[32] V. A. Bednyakov, N. D. Giokaris, and A. V. Bednyakov. "On the Higgs mass generation mechanism in the Standard Model". In: *Physics of Particles and Nuclei* 39.1 (Jan. 2008), pp. 13–36. DOI: 10.1007/s11496-008-1002-9. arXiv: hep-ph/0703280 [hep-ph] (cit. on p. 6).

[33] W. Beenakker, F. A. Berends, and A. P. Chapovsky. "Final-state radiation and line-shape distortion in resonance pair production". In: *Physics Letters B* 435.1-2 (Sept. 1998), pp. 233–239. DOI: 10.1016/S0370-2693(98)00760-6. arXiv: hep-ph/9805327 [hep-ph] (cit. on p. 12).

[34] Arnd Behring et al. “Mixed QCD-electroweak corrections to W-boson production in hadron collisions”. In: (Sept. 2020). arXiv: 2009.10386 [hep-ph] (cit. on p. 91).

[35] Jacob D. Bekenstein and Bibhas Ranjan Majhi. “Is the principle of least action a must?” In: *Nuclear Physics B* 892 (Mar. 2015), pp. 337–352. DOI: 10.1016/j.nuclphysb.2015.01.015. arXiv: 1411.2424 [hep-th] (cit. on p. 3).

[36] Carsten Bittrich et al. *In-situ scale factors from Zee events*. Tech. rep. ATL-COM-PHYS-2017-757. Geneva: CERN, June 2017. URL: https://cds.cern.ch/record/2268800 (cit. on pp. 37, 46).

[37] Serguei I. Bityukov. “Signal Significance in the Presence of Systematic and Statistical Uncertainties”. In: *Journal of High Energy Physics* 2002.9, 060 (Sept. 2002), p. 060. DOI: 10.1088/1126-6708/2002/09/060. arXiv: hep-ph/0207130 [hep-ph] (cit. on p. 64).

[38] J-B Blanchard, J-B de Vivie, and P Mastrandrea. *In situ scales and smearings from Z and J/Ψ events*. Tech. rep. ATL-COM-PHYS-2013-1653. Geneva: CERN, Dec. 2013. URL: https://cds.cern.ch/record/1637533 (cit. on pp. 37, 57).

[39] Jorge de Blas et al. “The Global Electroweak and Higgs Fits in the LHC era”. In: *PoS* EPS-HEP2017 (2017). Ed. by Paolo Checchia et al., p. 467. DOI: 10.22323/1.314.0467. arXiv: 1710.05402 [hep-ph] (cit. on p. 10).

[40] Volker Blobel. “Unfolding Methods in Particle Physics”. In: (Jan. 2011), 240–251. 12 p. DOI: 10.5170/CERN-2011-006.240. URL: https://cds.cern.ch/record/2203257 (cit. on p. 59).

[41] Maarten Boonekamp and Jean-Baptiste Blanchard. *Measurement of m_W at 7 TeV: Electron performance corrections and uncertainties*. Tech. rep. ATL-COM-PHYS-2014-1434. Geneva: CERN, Nov. 2014. URL: https://cds.cern.ch/record/1966963 (cit. on pp. 11, 46, 57).

[42] Maximilien Brice and Claudia Marcelloni. “View of the Detector ATLAS open.” Feb. 2007. URL: https://cds.cern.ch/record/1206860 (cit. on p. 18).

[43] Andrew Buckley. *ATLAS Pythia 8 tunes to 7 TeV data*. Tech. rep. Dec. 2014, p. 29 (cit. on p. 77).

[44] C.Wang. “presentation at the egamma calibration meeting”. In: (Sept. 2020). URL: https://indico.cern.ch/event/957069/contributions/4031631/attachments/2108689/3546663/ElectronCalibration.pdf (cit. on p. 57).

[45] Stefano Camarda et al. “DYTurbo: fast predictions for Drell–Yan processes”. In: *The European Physical Journal C* 80.3 (Mar. 2020). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-020-7757-5. URL: http://dx.doi.org/10.1140/epjc/s10052-020-7757-5 (cit. on pp. 94, 121).

[46] John M. Campbell, J.W. Huston, and W.J. Stirling. "Hard Interactions of Quarks and Gluons: A Primer for LHC Physics". In: *Rept. Prog. Phys.* 70 (2007), p. 89. DOI: 10.1088/0034-4885/70/1/R02. arXiv: hep-ph/0611148 (cit. on p. 8).

[47] Leonor Cerda Alberich. *Photon and electron identification with the ATLAS detector*. Tech. rep. ATL-PHYS-PROC-2016-230. Geneva: CERN, Nov. 2016. DOI: 10.22323/1.282.1235. URL: https://cds.cern.ch/record/2233097 (cit. on p. 138).

[48] Olmo Cerri. "Hadronic recoil in the W boson production at LHC for a W mass measurement with the CMS experiment". Presented 20 Sep 2017. 2017. URL: https://cds.cern.ch/record/2285935 (cit. on p. 11).

[49] Georgios Choudalakis. "Fully Bayesian Unfolding". In: 2012 (cit. on p. 59).

[50] P.J. Clark. "The ATLAS Detector Simulation". In: *Nuclear Physics B - Proceedings Supplements* 215.1 (2011). Proceedings of the 12th Topical Seminar on Innovative Particle and Radiation Detectors (IPRD10), pp. 85–88. ISSN: 0920-5632. DOI: https://doi.org/10.1016/j.nuclphysbps.2011.03.142. URL: http://www.sciencedirect.com/science/article/pii/S092056321100212X (cit. on p. 17).

[51] W.E. Cleland and E.G. Stern. "Signal processing considerations for liquid ionization calorimeters in a high rate environment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 338.2 (1994), pp. 467–497. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(94)91332-3. URL: http://www.sciencedirect.com/science/article/pii/0168900294913323 (cit. on p. 29).

[52] The ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08003–S08003. DOI: 10.1088/1748-0221/3/08/s08003. URL: https://doi.org/10.1088%2F1748-0221%2F3%2F08%2Fs08003 (cit. on p. 17).

[53] The NNPDF Collaboration et al. *Parton distributions from high-precision collider data*. 2017. arXiv: 1706.00428 [hep-ph] (cit. on p. 94).

[54] *Computation and validation of the electronic calibration constants for the ATLAS Liquid Argon Calorimeters*. Tech. rep. ATL-LARG-INT-2010-007. Geneva: CERN, July 2010. URL: https://cds.cern.ch/record/1278462 (cit. on p. 28).

[55] G. D'Agostini. *Improved iterative Bayesian unfolding*. 2010. arXiv: 1010.0632 [physics.data-an] (cit. on pp. 59, 99).

[56] G. D'Agostini. "Improved iterative Bayesian unfolding". In: *arXiv e-prints*, arXiv:1010.0632 (Oct. 2010), arXiv:1010.0632. arXiv: 1010.0632 [physics.data-an] (cit. on p. 61).

[57] Delgove David. “presentation at the egamma calibration meeting”. In: (Mar. 2019). URL: https://indico.cern.ch/event/804207/contributions/3353868/attachments/1811758/2959326/Linearity.pdf (cit. on p. 57).

[58] W. de Boer. “Precision Experiments at LEP”. In: *60 YEARS OF CERN EXPERIMENTS AND DISCOVERIES. Edited by SCHOPPER HERWIG ET AL. Published by World Scientific Publishing Co. Pte. Ltd* 23 (July 2015), pp. 107–136. DOI: 10.1142/9789814644150_0005. arXiv: 1509.06050 [hep-ex] (cit. on p. 34).

[59] *Design Report Tevatron 1 project*. Tech. rep. FERMILAB-DESIGN-1984-01. 1984. URL: https://cds.cern.ch/record/1478620 (cit. on p. 11).

[60] Luigi Di Lella and Carlo Rubbia. “The Discovery of the W and Z Particles”. In: *Adv. Ser. Dir. High Energy Phys.* 23 (2015), pp. 137–163. DOI: 10.1142/9789814644150_0006. URL: https://cds.cern.ch/record/2103277 (cit. on p. 11).

[61] Günther Dissertori. “The Determination of the Strong Coupling Constant”. In: *The Standard Theory of Particle Physics: Essays to Celebrate CERN’s 60th Anniversary*. Ed. by Luciano Maiani and et al. Oct. 2016, pp. 113–128. DOI: 10.1142/9789814733519_0006. arXiv: 1506.05407 [hep-ex] (cit. on p. 5).

[62] Stefan Dittmaier, Timo Schmidt, and Jan Schwarz. “Mixed NNLO QCD x electroweak corrections of $\mathcal{O}(N_f \alpha_s \alpha)$ to single-W/Z production at the LHC”. In: (Sept. 2020). arXiv: 2009.02229 [hep-ph] (cit. on p. 91).

[63] Daniel Dominguez. “Standard Model. Le modèle standard”. In: (Mar. 2015). General Photo. URL: https://cds.cern.ch/record/2002395 (cit. on p. 2).

[64] Claude Duhr, Falko Dulat, and Bernhard Mistlberger. “Charged Current Drell-Yan Production at N^{3}LO”. In: (July 2020). arXiv: 2007.13313 [hep-ph] (cit. on p. 91).

[65] Sayipjamal Dulat et al. “New parton distribution functions from a global analysis of quantum chromodynamics”. In: *Phys. Rev.* D93.3 (2016), p. 033006. DOI: 10.1103/PhysRevD.93.033006. arXiv: 1506.07443 [hep-ph] (cit. on p. 8).

[66] Miguel G. Echevarria, Ahmad Idilbi, and Ignazio Scimemi. “Factorization Theorem For Drell-Yan At Low q_T And Transverse Momentum Distributions On-The-Light-Cone”. In: *JHEP* 07 (2012), p. 002. DOI: 10.1007/JHEP07(2012)002. arXiv: 1111.4996 [hep-ph] (cit. on p. 8).

[67] “Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data”. In: *European Physical Journal C* 74, 3071 (Oct. 2014), p. 3071. DOI: 10.1140/epjc/s10052-014-3071-4. arXiv: 1407.5063 [hep-ex] (cit. on pp. 11, 33, 34, 46).

[68] "Electron and photon performance measurements with the ATLAS detector using the 20152017 LHC proton-proton collision data". In: *Journal of Instrumentation* 14.12 (Dec. 2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex] (cit. on p. 31).

[69] *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*. Tech. rep. ATL-PHYS-PUB-2017-022. Geneva: CERN, Dec. 2017. URL: https://cds.cern.ch/record/2298955 (cit. on pp. 29, 30).

[70] *Energy Resolution of the Barrel of the CMS Electromagnetic Calorimeter*. Tech. rep. CMS-NOTE-2006-148. Geneva: CERN, Nov. 2006. DOI: 10.1088/1748-0221/2/04/P04004. URL: https://cds.cern.ch/record/1009081 (cit. on p. 22).

[71] Jens Erler and Matthias Schott. "Electroweak precision tests of the Standard Model after the discovery of the Higgs boson". In: *Progress in Particle and Nuclear Physics* 106 (May 2019), pp. 68–119. DOI: 10.1016/j.ppnp.2019.02.007. arXiv: 1902.05142 [hep-ph] (cit. on p. 2).

[72] "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540 GeV$". In: *Phys. Lett. B* 122.CERN-EP-83-13 (Jan. 1983), 103–116. 31 p. DOI: 10.5170/CERN-1983-004.123. URL: https://cds.cern.ch/record/142059 (cit. on p. 11).

[73] Christian Wolfgang Fabjan and F Gianotti. "Calorimetry for Particle Physics". In: *Rev. Mod. Phys.* 75.CERN-EP-2003-075 (Oct. 2003), 1243–1286. 96 p. DOI: 10.1103/RevModPhys.75.1243. URL: https://cds.cern.ch/record/692252 (cit. on p. 22).

[74] Saskia Falke. "Measurement of the Higgs boson properties with Run 2 data collected by the ATLAS experiment. Mesure des propriétés du boson de Higgs avec les données du Run 2 collectées par l'expérience ATLAS". Presented 17 Sep 2019. 2019. URL: https://cds.cern.ch/record/2691944 (cit. on pp. 44, 45).

[75] J. Feltesse. "Introduction to Parton Distribution Functions". In: *Scholarpedia* 5.11 (2010). revision #186761, p. 10160. DOI: 10.4249/scholarpedia.10160 (cit. on p. 7).

[76] Joel Feltesse. "Introduction to deep inelastic scattering: Past and present". In: *20th International Workshop on Deep-Inelastic Scattering and Related Subjects*. 2012, pp. 3–12. DOI: 10.3204/DESY-PROC-2012-02/6 (cit. on p. 7).

[77] "Fits for W mass analysis". In: (). $\sin^2\theta$ and low mu run analyses workshop 2019. URL: https://indico.cern.ch/event/776453/contributions/3284036/attachments/1792310/2920446/PeregoMarta_2019-02-07-workshopOrsay.pdf (cit. on pp. 131, 132, 135).

[78] D. Fournier. "Liquid argon calorimetry". In: *ECFA Large Hadron Collider (LHC) Workshop: Physics and Instrumentation*. Oct. 1990, pp. 356–359 (cit. on p. 22).

[79] Paolo Francavilla. "The ATLAS Tile Hadronic Calorimeter performance at the LHC". In: *Journal of Physics: Conference Series* 404 (Dec. 2012), p. 012007. DOI: 10.1088/1742-6596/404/1/012007. URL: https://doi.org/10.1088%2F1742-6596%2F404%2F1%2F012007 (cit. on p. 20).

[80] Juerg Froehlich. "Relativistic Quantum Theory". In: *arXiv e-prints*, arXiv:1912.00726 (Dec. 2019), arXiv:1912.00726. arXiv: 1912.00726 [quant-ph] (cit. on p. 2).

[81] Cecilia E. Gerber. *LHC Highlights and Prospects*. Sept. 2019. arXiv: 1909.10919 [hep-ex] (cit. on p. 15).

[82] Christophe Goudet. "Etalonnage du calorimètre électromagnétique de l'expérience ATLAS et application à la mesure des couplages du boson de (Brout-Englert-)Higgs dans le canal diphoton dans le cadre du Run 2 du LHC." PhD thesis. Saclay, 2017 (cit. on pp. 23, 37, 39, 44).

[83] "Grand collisionneur électron-positon (LEP)". In: (Jan. 2013). URL: https://cds.cern.ch/record/1997707 (cit. on pp. 11, 39).

[84] Andrey Grozin. "Quantum Chromodynamics". In: *arXiv e-prints*, arXiv:1205.1815 (May 2012), arXiv:1205.1815. arXiv: 1205.1815 [hep-ph] (cit. on p. 4).

[85] Linghua Guo. *Electron energy in-situ calibration and linearity measurements from Z-gt;ee events*. Tech. rep. ATL-COM-PHYS-2020-757. Geneva: CERN, Oct. 2020. URL: https://cds.cern.ch/record/2742791 (cit. on p. 57).

[86] Manuel Guth. "Signal Region Optimisation Studies Based on BDT and Multi-Bin Approaches in the Context of Supersymmetry Searches in Hadronic Final States with the ATLAS Detector". Presented 27 Mar 2018. Mar. 2018. URL: https://cds.cern.ch/record/2308793 (cit. on p. 25).

[87] H1 and 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 [hep-ex] (cit. on pp. 94, 121).

[88] A. K. Halder, Andronikos Paliathanasis, and P. G. L. Leach. "Noether's Theorem and Symmetry". In: *arXiv e-prints*, arXiv:1812.03682 (Dec. 2018), arXiv:1812.03682. arXiv: 1812.03682 [math-ph] (cit. on p. 4).

[89] Johannes Haller et al. "Update of the global electroweak fit and constraints on two-Higgs-doublet models". In: *Eur. Phys. J. C* 78.8 (2018), p. 675. DOI: 10.1140/epjc/s10052-018-6131-3. arXiv: 1803.01853 [hep-ph] (cit. on pp. 5, 10).

[90] P. Hansen. "Results from the UA1 and UA2 Experiments". In: *International School of Physics Enrico Fermi: Elementary Particles*. June 1984, pp. 1–18 (cit. on pp. 2, 11).

[91] L. A. Harland-Lang et al. “Parton distributions in the LHC era: MMHT 2014 PDFs”. In: *The European Physical Journal C* 75.5 (May 2015). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-015-3397-6. URL: http://dx.doi.org/10.1140/epjc/s10052-015-3397-6 (cit. on pp. 94, 121).

[92] B J Holzer. “Introduction to Particle Accelerators and their Limitations”. In: *arXiv e-prints*, arXiv:1705.09601 (May 2017), arXiv:1705.09601. arXiv: 1705.09601 [physics.acc-ph] (cit. on p. 15).

[93] Tie-Jiun Hou et al. “New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC”. In: (Dec. 2019). arXiv: 1912.10053 [hep-ph] (cit. on pp. 94, 121).

[94] Tafoya Juan. “presentation at the egamma calibration meeting”. In: (Nov. 2020). URL: https://indico.cern.ch/event/973758/contributions/4100232/attachments/2141733/3609191/calibration_Egamma.pdf (cit. on p. 57).

[95] Jan Kretzschmar. *Samples and Physics modelling for low pile-up runs taken in 2017 and 2018*. Tech. rep. ATL-COM-PHYS-2019-075. Geneva: CERN, Feb. 2019. URL: https://cds.cern.ch/record/2657141 (cit. on pp. 47, 72, 97).

[96] L.Iconomidou-Fayard. “Status of the E1/E2 investigations with electrons”. In: (cit. on pp. 23, 44).

[97] C de La Taille and L Serin. *Temperature dependance of the ATLAS electromagnetic calorimeter signal. Preliminary drift time measurement*. Tech. rep. ATL-LARG-95-029. ATL-A-PN-29. Geneva: CERN, Nov. 1995. URL: https://cds.cern.ch/record/686091 (cit. on pp. 42, 57).

[98] W. Lampl, Christian Fabjan, and M. Aleksa. “Optimizing the Energy Measurement of the ATLAS Electromagnetic Calorimeter”. In: (Jan. 2006) (cit. on p. 22).

[99] “LAr temperature vs luminosity and impact on energy calibration”. In: (). URL: https://indico.cern.ch/event/748648/contributions/3222929/subcontributions/273194/attachments/1785880/2907510/temperature-2017-ws.pdf (cit. on pp. 43, 56).

[100] B Lenzi and R Turra. *Monte Carlo calibration update for electrons and photons using multivariate techniques*. Tech. rep. ATL-COM-PHYS-2013-1426. Geneva: CERN, Oct. 2013. URL: https://cds.cern.ch/record/1609589 (cit. on p. 33).

[101] Mengran Li et al. *Hadronic recoil reconstruction and calibration for low pile-up runs taken in 2017 and 2018*. Tech. rep. ATL-COM-PHYS-2019-078. Geneva: CERN, Feb. 2019. URL: https://cds.cern.ch/record/2657182 (cit. on pp. 12, 71).

[102] Peilian Liu. “Expected performance of the upgrade ATLAS experiment for HL-LHC”. In: *arXiv e-prints*, arXiv:1809.02181 (Sept. 2018), arXiv:1809.02181. arXiv: 1809.02181 [physics.ins-det] (cit. on p. 15).

[103] "Low Mu W Z Analyses". In: (). URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/LowMuWZAnalyses (cit. on p. 98).

[104] "Luminosity determination for low-pileup datasets at $\sqrt{s} = 5$ and 13 TeV using the ATLAS detector at the LHC". In: (July 2020) (cit. on p. 81).

[105] *Luminosity determination in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector at the LHC*. Tech. rep. ATLAS-CONF-2019-021. Geneva: CERN, June 2019. URL: https://cds.cern.ch/record/2677054 (cit. on p. 16).

[106] "Luminosity determination in pp collisions at sqrt{s} = 7 TeV using the ATLAS detector at the LHC". In: *European Physical Journal C* 71, 1630 (Apr. 2011), p. 1630. DOI: 10.1140/epjc/s10052-011-1630-5. arXiv: 1101.2185 [hep-ex] (cit. on p. 17).

[107] Melina Luthi. "Multivariate analysis techniques for Particle Flow-based neutral pileup suppression at the ATLAS experiment". In: (Jan. 2019). URL: https://cds.cern.ch/record/2655145 (cit. on p. 33).

[108] Simon Lyddon, Chris Holmes, and Stephen Walker. "General Bayesian Updating and the Loss-Likelihood Bootstrap". In: *arXiv e-prints*, arXiv:1709.07616 (Sept. 2017), arXiv:1709.07616. arXiv: 1709.07616 [stat.ME] (cit. on p. 39).

[109] Louis Lyons, Duncan Gibaut, and Peter Clifford. "How to Combine Correlated Estimates of a Single Physical Quantity". In: *Nucl. Instrum. Meth. A* 270 (1988), p. 110. DOI: 10.1016/0168-9002(88)90018-6 (cit. on p. 135).

[110] Bogdan Malaescu. *Discussions on unfolding problems, methods and solutions*. Tech. rep. 2016. URL: http://dpnc.unige.ch/~sfyrla/teaching/Statistics/lectures/Unfolding_Lecture_UGE_Malaescu.pdf (cit. on p. 69).

[111] Bogdan Malaescu. *Standard Model Unfolding twiki-page*. Tech. rep. 2016. URL: https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/StandardModelUnfoldingNew (cit. on pp. 66, 103).

[112] Martina Malberti. "Prospects for the precision determination of the W boson mass with the CMS detector at the LHC". PhD thesis. Milan Bicocca U., 2007 (cit. on p. 132).

[113] Neil Fraser Mcblane. "Energy Dependence of the Intermodule Widening Correction". In: (Aug. 2017). URL: https://cds.cern.ch/record/2281006 (cit. on p. 34).

[114] *Measurement of the* p_{T} *spectrum of* W*- and* Z*-bosons produced in* pp *collisions at* $\sqrt{s} = 5$ *TeV and* 13 *TeV in low-pileup runs*. Tech. rep. ATL-COM-PHYS-2018-1084. Geneva: CERN, July 2018. URL: https://cds.cern.ch/record/2632159 (cit. on pp. 72, 83, 84).

[115] "Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *Eur. Phys. J. C* 78.2, 110 (Feb. 2018), p. 110. DOI: 10.1140/epjc/s10052-017-5475-4. arXiv: 1701.07240 [hep-ex] (cit. on pp. 11, 12, 13, 46, 57, 71, 97, 133, 134, 135).

[116] “Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam”. In: *Nucl. Instrum. Methods Phys. Res., A* 614 (2010), 400–432. 33 p. DOI: 10.1016/j.nima.2009.12.055. URL: https://cds.cern.ch/record/1273537 (cit. on p. 28).

[117] “Measurement of $W^{\pm}$ and Z-boson production cross sections in pp collisions at $\sqrt{\{ s\}}$ = 13 TeV with the ATLAS detector”. In: *Physics Letters B* 759 (Aug. 2016), pp. 601–621. DOI: 10.1016/j.physletb.2016.06.023. arXiv: 1603.09222 [hep-ex] (cit. on p. 92).

[118] Krzysztof A. Meissner and Hermann Nicolai. “Standard Model Fermions and Infinite-Dimensional R Symmetries”. In: 121.9, 091601 (Aug. 2018), p. 091601. DOI: 10.1103/PhysRevLett.121.091601. arXiv: 1804.09606 [hep-th] (cit. on p. 1).

[119] Esma Mobs. “The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019”. In: (July 2019). General Photo. URL: https://cds.cern.ch/record/2684277 (cit. on p. 16).

[120] J.W. Monk and C. Oropeza-Barrera. “The HBOM method for unfolding detector effects”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 701 (Feb. 2013), pp. 17–24. ISSN: 0168-9002. DOI: 10.1016/j.nima.2012.09.045. URL: http://dx.doi.org/10.1016/j.nima.2012.09.045 (cit. on p. 59).

[121] N.Andari. “presentation at the ATLAS W precision analyses informal meeting 27Th October 2020”. In: (Sept. 2020). URL: https://indico.cern.ch/event/969042/contributions/4078527/attachments/2130404/3587822/presentation_271020_andari.pdf (cit. on p. 43).

[122] Jason Nielsen. “Fundamentals of LHC Experiments”. In: *String Theory and its Applications - TASI 2010, From meV to the Planck Scale*. Ed. by Michael Dine, Thomas Banks, and Subir Sachdev. Nov. 2012, pp. 127–152. DOI: 10.1142/9789814350525_0003. arXiv: 1106.2516 [hep-ex] (cit. on p. 15).

[123] Nikiforos Nikiforou. *Performance of the ATLAS Liquid Argon Calorimeter after three years of LHC operation and plans for a future upgrade*. Tech. rep. arXiv:1306.6756. June 2013. DOI: 10.1109/ANIMMA.2013.6728060. URL: http://cds.cern.ch/record/1558820 (cit. on p. 23).

[124] Tadej Novak. “ATLAS Pile-up and Overlay Simulation”. In: (June 2017). URL: https://cds.cern.ch/record/2270396 (cit. on p. 17).

[125] J Nowell. “A Measurement of the W Boson Mass with the ALEPH Detector”. Presented on Nov 2000. 2000. URL: https://cds.cern.ch/record/537297 (cit. on p. 11).

[126] "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 30–61. ISSN: 0370-2693. DOI: https://doi.org/10.1016/j.physletb.2012.08.021. URL: http://www.sciencedirect.com/science/article/pii/S0370269312008581 (cit. on p. 2).

[127] D Oliveira Damazio. *Signal Processing for the ATLAS Liquid Argon Calorimeter : studies and implementation*. Tech. rep. ATL-LARG-PROC-2013-015. Geneva: CERN, Nov. 2013. URL: https://cds.cern.ch/record/1630826 (cit. on p. 28).

[128] ATLAS Outreach. "ATLAS Fact Sheet : To raise awareness of the ATLAS detector and collaboration on the LHC". 2010. URL: https://cds.cern.ch/record/1457044 (cit. on pp. 19, 26).

[129] "Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties". In: *Journal of Instrumentation* 12.5 (May 2017), P05002. DOI: 10.1088/1748-0221/12/05/P05002. arXiv: 1702.06473 [hep-ex] (cit. on p. 19).

[130] Antonio Pich. "The Standard Model of Electroweak Interactions; rev. version". In: hep-ph/0502010. FTUV-2005-0201. IFIC-2005-13 (Feb. 2005), 48 p. DOI: 10.5170/CERN-2006-003.1. URL: https://cds.cern.ch/record/819632 (cit. on p. 3).

[131] "presnetation at the egamma calibration meeting". In: (July 2019). URL: https://indico.cern.ch/event/827055/contributions/3506176/attachments/1885454/3110615/Slide_ES.pdf (cit. on pp. 43, 57).

[132] *Prospects for the measurement of the W-boson transverse momentum with a low pileup data sample at $\sqrt{s} = 13$ TeV with the ATLAS detector*. Tech. rep. ATL-PHYS-PUB-2017-021. Geneva: CERN, Dec. 2017. URL: https://cds.cern.ch/record/2298152 (cit. on pp. 6, 71, 72).

[133] Chris Quigg. "The Electroweak Theory". In: *arXiv e-prints*, hep-ph/0204104 (Apr. 2002), hep–ph/0204104. arXiv: hep-ph/0204104 [hep-ph] (cit. on p. 5).

[134] Giulia Ripellino. *The alignment of the ATLAS Inner Detector in Run 2*. Tech. rep. ATL-INDET-PROC-2016-003. Geneva: CERN, Sept. 2016. DOI: 10.22323/1.276.0196. URL: https://cds.cern.ch/record/2213441 (cit. on p. 20).

[135] Valerio Rossetti. "Performance of the ATLAS Calorimeters and Commissioning for LHC Run-2". In: (July 2015). URL: https://cds.cern.ch/record/2037117 (cit. on p. 25).

[136] Aranzazu Ruiz-Martinez and ATLAS Collaboration. *The Run-2 ATLAS Trigger System*. Tech. rep. ATL-DAQ-PROC-2016-003. Geneva: CERN, Feb. 2016. DOI: 10.1088/1742-6596/762/1/012003. URL: https://cds.cern.ch/record/2133909 (cit. on pp. 26, 27).

[137] S. Schael et al. “Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP”. In: *Phys. Rept.* 532 (2013), pp. 119–244. DOI: 10.1016/j.physrep.2013.07.004. arXiv: 1302.3415 [hep-ex] (cit. on p. 11).

[138] Christoph Schäfer. “The Z lineshape at LEP”. In: *Nuclear Physics B - Proceedings Supplements* 65.1 (1998). The Irresistable Rise of the Standard Model, pp. 93–97. ISSN: 0920-5632. DOI: https://doi.org/10.1016/S0920-5632(97)00982-1. URL: http://www.sciencedirect.com/science/article/pii/S0920563297009821 (cit. on p. 10).

[139] Stefan Schmitt. “Data Unfolding Methods in High Energy Physics”. In: *EPJ Web of Conferences* 137 (2017). Ed. by Y. Foka, N. Brambilla, and V.Editors Kovalenko, p. 11008. ISSN: 2100-014X. DOI: 10.1051/epjconf/201713711008. URL: http://dx.doi.org/10.1051/epjconf/201713711008 (cit. on pp. 59, 62).

[140] Matthias Schott et al. *Measurement of m_W at 7 TeV: Reconstruction of the hadronic recoil*. Tech. rep. ATL-COM-PHYS-2014-1435. Geneva: CERN, Nov. 2014. URL: https://cds.cern.ch/record/1966964 (cit. on p. 11).

[141] Estelle Scifo. “Measurement of the Brout-Englert-Higgs boson couplings in its diphoton decay channel with the ATLAS detector at the LHC”. PhD thesis. Orsay, LAL, July 2014 (cit. on pp. 9, 23).

[142] Steffen Starz. “ATLAS Calorimeter system: Run-2 performance, Phase-1 and Phase-2 upgrades”. In: (July 2018). URL: https://cds.cern.ch/record/2628123 (cit. on p. 21).

[143] John Strologas and Steven Errede. “Study of the angular coefficients and corresponding helicity cross sections of the W boson in hadron collisions”. In: 73.5, 052001 (Mar. 2006), p. 052001. DOI: 10.1103/PhysRevD.73.052001. arXiv: hep-ph/0503291 [hep-ph] (cit. on p. 12).

[144] Robin G. Stuart. “An Improved Determination of the Fermi Coupling Constant, G_F”. In: *arXiv e-prints*, hep-ph/9902257 (Feb. 1999), hep–ph/9902257. arXiv: hep-ph/9902257 [hep-ph] (cit. on p. 10).

[145] “Summary Report of Physics Beyond Colliders at CERN”. In: *arXiv e-prints*, arXiv:1902.00260 (Feb. 2019), arXiv:1902.00260. arXiv: 1902.00260 [hep-ex] (cit. on p. 15).

[146] M. Tanabashi et al. “Review of Particle Physics”. In: *Phys. Rev. D* 98.3 (2018), p. 030001. DOI: 10.1103/PhysRevD.98.030001 (cit. on p. 10).

[147] The ATLAS Collaboration. “Operation and performance of the ATLAS semiconductor tracker”. In: *Journal of Instrumentation* 9.8, P08009 (Aug. 2014), P08009. DOI: 10.1088/1748-0221/9/08/P08009. arXiv: 1404.7473 [hep-ex] (cit. on p. 19).

[148] “The ATLAS Experiment at the CERN Large Hadron Collider”. In: *JINST* 3 (2008). Also published by CERN Geneva in 2010, S08003. 437 p. DOI: 10.1088/1748-0221/3/08/S08003. URL: https://cds.cern.ch/record/1129811 (cit. on p. 6).

[149] "The ATLAS Inner Detector commissioning and calibration". In: *European Physical Journal C* 70.3 (Dec. 2010), pp. 787–821. DOI: 10.1140/epjc/s10052-010-1366-7. arXiv: 1004.5293 [physics.ins-det] (cit. on p. 19).

[150] "The ATLAS silicon pixel sensors". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 456.3 (2001), pp. 217–232. ISSN: 0168-9002. DOI: https://doi.org/10.1016/S0168-9002(00)00574-X. URL: http://www.sciencedirect.com/science/article/pii/S016890020000574X (cit. on p. 19).

[151] Ruggero Turra, Stefano Manzoni, and Archil Durglishvili. *Monte Carlo energy calibration of electrons and photons for release 20.7*. Tech. rep. ATL-COM-PHYS-2017-761. Geneva: CERN, June 2017. URL: https://cds.cern.ch/record/2268813 (cit. on p. 33).

[152] Nenad Vranjes, Troels Petersen, and Mikhail Karnevskiy. *Measurement of m_W at 7 TeV: Muon momentum corrections and uncertainties*. Tech. rep. ATL-COM-PHYS-2014-1433. Geneva: CERN, Nov. 2014. URL: https://cds.cern.ch/record/1966962 (cit. on pp. 11, 138).

[153] Jorg Wenninger. "Operation and Configuration of the LHC in Run 2". In: (Mar. 2019). URL: https://cds.cern.ch/record/2668326 (cit. on p. 43).

[154] Scott Willenbrock. "Hadron Colliders, the Standard Model, and Beyond". In: *arXiv e-prints*, hep-ph/0212032 (Dec. 2002), hep–ph/0212032. arXiv: hep-ph/0212032 [hep-ph] (cit. on p. 1).

[155] Tairan Xu. "Measurements of W boson properties at $\sqrt{s} = 5$ and 13 TeV with the ATLAS detector at the LHC". Presented 25 Sep 2019. 2019. URL: https://cds.cern.ch/record/2689846 (cit. on pp. 12, 71, 81, 99, 138, 139).

[156] Tairan Xu, Hicham Atmani, and Ludovica Aperio Bella. *Electron corrections for low pile-up runs taken in 2017 and 2018*. Tech. rep. ATL-COM-PHYS-2019-077. Geneva: CERN, Feb. 2019. URL: https://cds.cern.ch/record/2657152 (cit. on p. 39).

[157] Tairan Xu and Maarten Boonekamp. *Multi-jet background in low-pile-up runs taken in 2017 and 2018*. Tech. rep. ATL-COM-PHYS-2019-076. Geneva: CERN, July 2019. URL: https://cds.cern.ch/record/2657146 (cit. on p. 12).

[158] Christoph Zimmermann, Matthias Schott, and Mikhail Karnevskiy. *Measurement of mW at 7 TeV : Muon efficiency corrections and uncertainties*. Tech. rep. ATL-COM-PHYS-2015-073. Geneva: CERN, Feb. 2015. URL: https://cds.cern.ch/record/1987497 (cit. on p. 11).