Neutral-current Drell-Yan measurements at the ATLAS experiment and their phenomenological interpretation

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Ricardo González López

Department of Physics

Oliver Lodge Laboratory

University of Liverpool

June 16, 2023

Abstract

The neutral-current Drell-Yan production is a process for which precise predictions are currently available, offering a benchmark upon which these state-of-the-art predictions can be tested, both as a way to improve our understanding of Standard Model physics and explore theories that expand on the foundations it establishes. This thesis provides an overview of two analyses focusing on the Drell-Yan production using LHC Run 2 pp collision data collected by the ATLAS experiment during the years 2015 to 2018 at $\sqrt{s} = 13$ TeV. First, a measurement of the single- and double-differential production cross-section in an invariant mass region of 116 to 5000 GeV is presented, using Z/γ^* decays into lepton pairs in the electron and muon channels. The measurement is performed as a function of invariant mass, absolute rapidity of the lepton pair and their angle separation in the Z-boson rest frame. The results obtained are used to perform for the first time tests on lepton flavour universality (LFU) as well as set limits on Effective Field Theory (EFT) coefficients that modify the Drell-Yan production cross-section. Second, a novel search for Lorentz-invariance violating (LIV) signatures using Z-decays into electron and muon pairs is presented. Using the methodology for a luminosity measurement based on the counting of Z-bosons detected, the time-dependence of the Drell-Yan production can be monitored. While the Standard Model predicts no time-dependence on this process, deviations from this behaviour can be interpreted as the effects of LIV-inducing operators in the Standard Model Extension EFT. The results presented in this thesis include an overview of the Z-counting measurement methodology transfer to a flexible analysis framework, as well as the development of new time-dependent simulation tools to estimate expected sensitivities.

Declaration

I hereby confirm this work is my own, except where other works are referenced. This work has not previously been submitted to any institute, including this one. This thesis does not exceed the relevant word count.

Ricardo González López

A mis padres, que me han dado todo el amor que me ha traído hasta aquí. Siempre habéis apoyado todas mis decisiones, que esta tesis sea la prueba de que valió la pena.

To my parents, who've given me all the love that brought me this far. You have always supported every decision I've made, let this thesis prove it was all worth it.

Acknowledgements

I would like to first thank my supervisors, Uta Klein and Jan Kretzschmar, who have provided throughout these years the best guidance I could have ever asked for. You have been an invaluable source of knowledge and encouragement, and I believe you have pushed me to my very best while keeping these years as memories I will always cherish. I would like to extend my gratitude to the whole Liverpool High Energy Physics group. It has been a pleasure to be a part of such an excellent research cluster, both at an academical and personal level. I would like to thank as well all the people I have had the pleasure to work with at the Liverpool ATLAS group, my analysis teams and the UK ITk Pixel testbeam team. It has been thrilling to be a part of such an incredible collaboration and this thesis would have not been possible without the support of these great teams.

Many thanks to my fellow PhD students in the department, in particular the people I was fortunate enough to start this journey with. I came to Liverpool to what should have been a foreign, cold, rainy city lost in Northern England and I found myself feeling at home in no time, surrounded by friends whose support has been invaluable all these years. I have also been lucky enough to meet here Beth, who has turned many months of pressure, tension and uncertainty into the best ones in my life. I could not have been happier since I met you and I cannot wait to see what future holds for us.

I would also like to thank the University of Liverpool basketball club. I have met some amazing people here and I loved playing with every single one of you, it has been great to see everyone grow as a player and as a person. To carry on playing the sport I am so passionate about has given me the headspace I needed for a project as demanding as this PhD, so I will be forever thankful for landing in such a great group of people that shared this passion. Special mention goes to Jorge, happily seating in the middle of the Venn diagram of both groups I just mentioned. It has been an absolute pleasure to call you a colleague, teammate and friend during these years, and thankfully many more.

Last but by no means least, huge thanks to my parents, family and friends back in Spain. The circumstances surrounding this thesis did not allow to see you as much as I would have hoped, but you have made me feel your love and support during these years, being a constant source of inspiration. Despite life continuing to rapidly change, every time I have been back it feels like I was not gone to begin with. You are the reason why home is, and will always be, *home*.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The developments over the last century in particle physics, trying to understand the building blocks of the Universe and the mechanisms ruling their interactions, have led to what is now known as the Standard Model (SM) of particle physics. This framework unites the elementary particles known to date with the electromagnetic, weak and strong forces. Despite being able to predict many results long before their discovery with a high level of accuracy, the SM is known to be an incomplete theory in view of the open questions it fails to answer. These questions have led to the development of a plethora of Beyond Standard Model (BSM) theories, which experiments have been trying to discover for many years with no clear indications of their presence found yet.

The Large Hadron Collider (LHC) has produced an immense amount of proton-proton (pp) collision data since it began operating in 2008, recorded by various experiments situated in different points of the collider. This thesis presents several analyses based on the data recorded by the ATLAS experiment, one of the general purpose detectors located at the LHC. These analyses focus on the measurement of the neutral-current Drell-Yan production at LHC in different regions of the kinematic phase space, both providing a precise measurement of its cross-section and novel interpretations in the context of two different extensions of the SM. The thesis is structured as follows:

Part I: Theory overview Chapter 2 covers the relevant theory foundations for the presented analyses. Chapter 3 provides an overview of the frameworks and tools used to generate predictions of the studied processes.

Part II: Experimental setup Chapter 4 presents an overview of the Large Hadron Collider and the ATLAS experiment in Chapter 5. Chapter 6 covers the reconstruction of the physics objects used in the analyses, as well as the calibration and correction procedures used for the measurement of their properties.

Part III: High-mass Drell-Yan cross-section The measurement of the neutral-current Drell-Yan cross-section at invariant masses above 116 GeV is presented in this part. Chapter 7 presents a general motivation for the measurement, also introducing the existing results for the measurement and how the analysis here presented aims to improve them. The methodology of the measurement is discussed in Chapter 8, covering the simulation, event selection and background estimation used in the analysis, as well as an estimate on the experimental and theory uncertainties. The results of the cross-section measurement are shown in Chapter 9, while Chapter 10 includes an overview of the the Lepton Flavour Universality tests performed with said results and the measurement's Effective Field Theory interpretation. The part concludes with a summary of the results in Chapter 11.

Part IV: Search for Lorentz-invariance violating signatures This part covers the search for Lorentz-invariance violating (LIV) signatures in the dilepton final state, using the Z-boson decays into electron and muon pairs to look for time-dependent features that could arise from the effect of operators in the Standard Model Extension EFT. The part begins in Chapter 12, introducing the motivation for this search. Chapter 13 contains an overview of the methodology used to monitor the production of the Z-boson at the ATLAS experiment in a time-dependent manner and Chapter 14 describes the analysis methodology used to search for Lorentz-invariance violation using these results. Sensitivity studies on LIV signatures are presented in Chapter 15, as well as a methodology for signal blinding in data. The part concludes in Chapter 16 with a summary of the results of the analysis.

Personal Contributions

As part of the research carried out throughout the development of my PhD I was an active member of the high-mass Drell-Yan cross-section measurement team, with significant contributions to its upcoming publication. In Chapter 8 the estimation of the fake electron background (including thorough cross-checks using different methodologies), as well as the development of alternative theory predictions and uncertainties using Sherpa, were performed by myself. The Effective Field theory interpretation presented in Chapter 10 was also developed by myself, including signal generation, fitting and assessment of the impact of experimental systematics. The results presented in this thesis include the EFT interpretation of the double-differential cross-section measurement, which for technical reasons is not part of the measurement publication, improving the limits set on different Wilson coefficients with respect to the single-differential measurement interpretation (which will be included in the upcoming publication).

Furthermore, I was part of the search for Lorentz-invariance violating signatures in dilepton final states. This is a unique search, built by the team from ground up, in which

I took part in most analysis stages. Chapter 13 presents the transfer of the Z-counting methodology to a standard ATLAS analysis framework and the results obtained with it, which was developed by myself. The novel methodologies to simulate time-dependence in Monte Carlo presented in Chapter 15 were also work of my own.

All figures shown in the thesis were created by myself, unless specifically referenced to another source.

Part I

Theory

Chapter 2

Theory foundations

The first part of this chapter covers an introduction to the Standard Model of particle physics, giving an overview of the theory as a whole and its current status. An insight into the limitations of Standard Model and certain theories that try to overcome it are presented next, introducing the Effective Field Theory formalisms on which the measurements' interpretations presented in Parts III and IV will be based upon.

Details on the tools used for theoretical predictions, as well as the basics of proton-proton collision description and simulation, are presented in Chapter 3.

2.1 The Standard Model of Particle Physics

The Standard Model of particle physics is the theory that describes the fundamental interactions between particles and the key properties and characterise said particles [1–4]. It is the culmination of years of theoretical and experimental efforts in order to identify the fundamental particles and test the theories describing the force fields that reign their interactions.

2.1.1 Fundamental particles

The Standard Model includes two types of matter particles: *leptons* and *quarks*. Both have spin-$\frac{1}{2}$ (in units of $\hbar$) and behave point-like, showing no internal structure at distances probed by our current accelerators. For all matter particles described in Standard Model there is an anti-particle, with the same mass as their corresponding particle. The quantum numbers of each antiparticle are the opposite values of those of their respective particle.

Charged leptons are a generalization of the electron, with muons and tau leptons carrying identical values in most of the properties of their lightest partner, with the exception

of lepton flavour charge (L_f, where f represents the three different families). This quantum number differs between lepton generations and is conserved in all Standard Model interactions. For each charged lepton there is a neutral partner of the same generation, the neutrino. Leptons, if charged, interact weakly and electromagnetically, though if neutral, interact weakly exclusively. A summary of the known leptons and their properties can be found in Table 2.1.

Lepton	Symbol	Q	T_3	L_e	L_μ	L_τ	Mass [GeV][†]
Electron	e	-1	$-\frac{1}{2}$	1	0	0	$0.51{\cdot}10^{-3}$
Electron neutrino	ν_e	0	$+\frac{1}{2}$	1	0	0	$< 1.0 \cdot 10^{-9}$
Muon	μ	-1	$-\frac{1}{2}$	0	1	0	0.11
Muon neutrino	ν_μ	0	$+\frac{1}{2}$	0	1	0	$< 0.2 \cdot 10^{-3}$
Tau	τ	-1	$-\frac{1}{2}$	0	0	1	1.78
Tau neutrino	ν_τ	0	$+\frac{1}{2}$	0	0	1	$< 18.2 \cdot 10^{-3}$

Table 2.1: Summary of some properties of the Standard Model leptons. Antiparticles are not shown, their quantum numbers correspond to the opposite of those of their corresponding particle. Q corresponds to the electromagnetic charge in units of e, T_3 is the weak isospin and L_f are the different lepton flavour charges. From Ref. [5].

Quarks are the fundamental constituents of hadrons and they experience all three forces comprised in the Standard Model: electromagnetic, weak and strong. They were proposed in 1964 by Gell-Mann and Zweig [6, 7] to explain the structure that hadrons exhibit, organizing in SU(3) singlets, octets, nonets and decimets. Quarks organize in an analogous way to leptons, finding three families in the Standard Model: *up-down*, *charm-beauty* and *top-bottom*. A summary of the known quarks and some of their properties can be found in Table 2.2.

One of the main properties of quarks is the *colour charge*, the key feature that rules strong interactions. It was first proposed to explain the existence of quarks in seemingly same states [8], which would violate the Pauli exclusion principle. The theory required hadrons to be indistinguishable from each other, imposing the requirement on all hadrons to be *colourless*. The combination of all colour charges gives rise to a colourless state, so does a colour-anticolour combination, as observed in $q\bar{q}$ states. The condition for all hadrons to be colourless determined that no more than three colours could exist, meaning that three was the exact number of colours that quarks could take, namely *red*, *green* and *blue*.

[†]Note that all variables in this thesis will be expressed in natural units, assuming that the speed of light and the Planck constant are equal to unity, $c = \hbar = 1$.

Quark	Symbol	Q	T_3	Mass [GeV]
Up	u	$+\frac{2}{3}$	$+\frac{1}{2}$	$2.16{\cdot}10^{-3}$
Down	d	$-\frac{1}{3}$	$-\frac{1}{2}$	$4.63{\cdot}10^{-3}$
Charm	c	$+\frac{2}{3}$	$+\frac{1}{2}$	1.27
Strange	s	$-\frac{1}{3}$	$-\frac{1}{2}$	0.093
Top	t	$+\frac{2}{3}$	$+\frac{1}{2}$	172.9
Bottom	b	$-\frac{1}{3}$	$-\frac{1}{2}$	4.18

Table 2.2: Summary of some properties of the Standard Model quarks. Antiparticles are not shown, their quantum numbers correspond to the opposite of those of their corresponding particle. Q corresponds to the electromagnetic charge in units of e and T_3 is the weak isospin. From Ref. [5].

Another key characteristic of quarks behaviour is asymptotic freedom and colour confinement. Both can be inferred by studying the running of the strong coupling constant as explained in Section 2.1.2. Asymptotic freedom results into a quasi-free behaviour of quarks present in bound states at small distances, while quark confinement derives into the lack of observation of free quarks and the process called hadronization.

2.1.2 Fundamental interactions

Interactions between elementary particles are represented in the Standard Model as the exchange of virtual quanta, which act as mediators of the force involved in the interaction, an idea first proposed by Yukawa in 1935 [9], with each type of interaction being the consequence of a specific gauge symmetry.

In these theories, the dynamics involved in an interaction are closely related to symmetry principles. These symmetries express the invariance of the studied processes under different transformations. While most invariances appear naturally when considering *global* transformations, applied on all space-time points, *local* invariances are achieved by introducing additional fields into our theory, which ultimately involve the appearance of force mediators and associated charges.

Electromagnetic interaction

The electromagnetic interaction arises from a local U(1) invariance, as described in *quantum electrodynamics* or QED. Considering the following local transformation of a given wavefunction $\psi(x)$,

$$\psi(x) \to \psi'(x) = e^{i\alpha(x)}\psi(x), \tag{2.1}$$

the derivatives of said wave function present in many observables, as well as in the

kinetic terms of the Lagrangian, would transform as:

$$\partial_\mu \psi(x) \to \partial_\mu \psi'(x) = e^{i\alpha(x)}[\partial_\mu \psi(x) + i(\partial_\mu \alpha(x))\psi(x)]. \tag{2.2}$$

This derivative includes a term involving the gradient of phase, $\partial_\mu \alpha(x)$, breaking the invariance. One can overcome this problem by introducing an additional field, A^μ, while promoting the derivative to introduce invariance under U(1) local transformations,

$$\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu. \tag{2.3}$$

A^μ here can be identified as the electromagnetic field, giving raise to an electromagnetic current and an electromagnetic charge, q. This field will transform under U(1) transformations as

$$A_\mu(x) \to A'_\mu(x) \equiv A_\mu(x) - \frac{1}{q}\partial_\mu \alpha(x), \tag{2.4}$$

in such way that the term $D_\mu \psi(x)$ is invariant under U(1) local transformations. This leads to the possibility of introducing a gauge-invariant modification of the free particle Lagrangian,

$$\mathcal{L}_{\text{free}} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi, \tag{2.5}$$

into an expression that introduces the conserved electromagnetic current:

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - qA_\mu \bar{\psi}\gamma^\mu \psi. \tag{2.6}$$

When adding the kinetic term for the gauge field describing the propagation of photons, $\frac{1}{4}F^{\mu\nu}F_{\mu\nu}$ with $F^{\mu\nu} \equiv \partial^\mu A^\mu - \partial^\nu A^\mu$, one finds the usual expression for the QED Lagrangian.

Electroweak interaction and spontaneous symmetry breaking

The weak theory is built in a similar manner to the electromagnetism gauge theory presented above, assuming a SU(2) symmetry instead. However, the Standard Model actually comprises a unification of the SU(2) and U(1) symmetries, presenting a unified *electroweak* theory built upon a $\text{SU}(2)_L \otimes \text{U}(1)_Y$ symmetry. The $\text{SU}(2)_L$ involves a weak-isospin symmetry, while the $\text{U}(1)_Y$ involves a weak-hypercharge symmetry, both corresponding to different quantum numbers characteristic of all matter particles. The gauge-covariant derivative of this group takes the form:

$$D_\mu = \partial_\mu + \frac{ig'}{2}A_\mu Y + \frac{ig}{2}\boldsymbol{\tau} \cdot \mathbf{b}_\mu, \tag{2.7}$$

where g and g' are the coupling constants to each of the emerging fields, A^μ and $\mathbf{b}_\mu$ (note that the latter is a vector, $\mathbf{b}_\mu = (b^1_\mu, b^2_\mu, b^3_\mu)$). Y represents the weak hypercharge,

defined as a function of the electric charge, Q, and the third isospin component, T_3, as $Y = 2(Q - T^3)$. $\boldsymbol{\tau}$ are the SU(2) group generators, the Pauli matrices.

The gauge fields $\mathbf{b}_\mu$ and A_μ do not exactly represent the gauge bosons observed in nature, namely the $W^\pm$ and Z bosons and the photon. Instead, these correspond to a linear combination of the gauge fields:

$$W_\mu^\pm \equiv \frac{{b^1}_\mu \mp i{b^2}_\mu}{\sqrt{2}} \tag{2.8}$$

and

$$\begin{pmatrix} \mathcal{A}_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \cdot \begin{pmatrix} A_\mu \\ b_\mu^3 \end{pmatrix}, \tag{2.9}$$

where the Weinberg or weak mixing angle, θ_W, is introduced to control the mixing of the gauge fields. This parameter can also be expressed in terms of the coupling constants:

$$\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}}, \quad \sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}. \tag{2.10}$$

Fermion fields can be separated into right- and left-handed helicity doublets as:

$$\psi_{R,L} = \frac{1}{2}(1 \pm \gamma_5)\psi, \quad \psi = \psi_R + \psi_L, \tag{2.11}$$

where negative (positive) terms denote right-handed (left-handed) components. The matrix γ_5 is constructed using the Dirac matrices as $\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3$. When the covariant derivative defined in Eq. 2.8 acts on fermion fields, this results in different couplings between gauge fields and fermions:

$$D_\mu\psi_L = \left(\partial_\mu + \frac{ig'}{2}A_\mu Y + \frac{ig}{2}\boldsymbol{\tau}\cdot\mathbf{b}_\mu\right)\psi_L, \quad D_\mu\psi_R = \left(\partial_\mu + \frac{ig'}{2}A_\mu Y\right)\psi_R, \tag{2.12}$$

since the gauge fields $\mathbf{b}_\mu$ only act on the left-handed component and A_μ acts on both components.

However, by simply considering in the Lagrangian the kinetic terms of these gauge fields and the fermions they interact with, one does not retrieve any mass terms for the emerging fields, just as the QED Lagrangian involved the appearance of a massless field. This would give rise to four massless bosons rather than the gauge bosons present in the Standard Model. Moreover, the different coupling of the gauge fields to left- and right-handed components means that, for instance, the electron mass term,

$$e\bar{e} = \frac{1}{2}(1-\gamma_5)e + \frac{1}{2}(1+\gamma_5)e = \bar{e}_R e_L + \bar{e}_L e_R, \tag{2.13}$$

breaks the gauge invariance (and so would all fermion mass terms) and therefore cannot be included. This feature of the $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ gauge symmetry requires an additional mechanism to be added to Standard Model, granting mass to those particles that are known to be massive.

In the SM this mechanism is referred to as the Brout-Englert-Higgs mechanism [10–14], and introduces mass terms through the concept of spontaneous symmetry breaking. The mechanism introduces a doublet Φ of complex scalar fields,

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{2.14}$$

defined with four degrees of freedom, corresponding to different scalar fields ϕ_1, ϕ_2, ϕ_3 and ϕ_4, which are related to the original complex scalar fields as,

$$\phi^+ = \frac{\phi_1 + i\phi_2}{\sqrt{2}}, \quad \phi^0 = \frac{\phi_3 + i\phi_4}{\sqrt{2}}. \tag{2.15}$$

The scalar doublet Φ is used to define the Higgs sector Lagrangian:

$$\mathcal{L}_{\text{Higgs}} = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi); \quad V(\phi) = \mu^2(\phi^\dagger \phi) + |\lambda|(\phi^\dagger \phi)^2. \tag{2.16}$$

The covariant derivative D_μ maintains the electroweak definition presented in Equation 2.7 and the term $V(\phi)$ is the Higgs potential. The potential introduces two free parameters, μ and λ, often referred to as the mass and self-coupling parameters respectively.

If one sets the mass parameter to be defined positively, $\mu^2 > 0$, one finds the minimum of the Higgs potential at

$$|\Phi_{\text{min}}| = \sqrt{\frac{\mu^2}{2\lambda}} = \frac{\nu}{\sqrt{2}}, \tag{2.17}$$

where a new term ν has been introduced, the vacuum expectation value (VEV). Thus, the ground state of the Higgs field Φ can be expressed as

$$\Phi_0 = \begin{pmatrix} 0 \\ \frac{\nu}{\sqrt{2}} \end{pmatrix}. \tag{2.18}$$

If one expands the field around the minimum and applies the unitary gauge [15], fixing the VEVs of three of the scalar fields to zero, the scalar doublet can be rewritten as

$$\Phi' = \begin{pmatrix} 0 \\ \frac{\nu+H}{\sqrt{2}} \end{pmatrix}, \tag{2.19}$$

where H corresponds to a neutral scalar field, namely the Higgs boson. The introduction of this re-defined scalar doublet into the Higgs Lagrangian gives rise to the mass terms that could not be directly introduced due to gauge symmetry breaking in the electroweak theory. Thus, the masses of the fermion and boson fields are expressed via their coupling to the Higgs boson:

$$m_W = \frac{1}{2}\nu g, \quad m_Z = \frac{1}{2}\nu\sqrt{g^2+g'^2}, \quad m_A = 0, \quad m_f = \frac{\lambda_f \nu}{\sqrt{2}}, \tag{2.20}$$

where λ_f represents the Yukawa coupling between a fermion f and the Higgs field, further explained in Section 2.1.3. One can also find the term for the mass of the Higgs field emerging from its self-interacting nature: $m_H = \sqrt{2\nu^2\lambda}$.

Strong interaction

The strong interaction is described by quantum chromodynamics, or QCD, which depicts the interactions between particles carrying *colour charge*, as explained in Section 2.1.1. The theory is described by the gauge group $\mathrm{SU}(3)_C$, introducing a new gauge-covariant derivative,

$$D_\mu = \partial_\mu + i g_s B_\mu, \tag{2.21}$$

where g_s is the coupling constant to the emerging fields and B_μ is a 3×3 matrix composed by the eight colour gauge fields, b_μ^l, and the generators of the SU(3) group, the λ^l matrices [16]:

$$B_\mu = \frac{1}{2}\boldsymbol{\lambda} \cdot \mathbf{b}_\mu = \frac{1}{2}\lambda^l b_\mu^l. \tag{2.22}$$

These give rise to eight massless bosons which will mediate the strong interaction, the *gluons*.

The QCD coupling constant is commonly redefined to a "fine-structure constant"-like coupling:

$$\alpha_s = \frac{g_s^2}{4\pi}. \tag{2.23}$$

An important feature of quantum field theories is the running of the coupling constant, meaning that the coupling strength will change with the energy scale of the interaction. For QCD at lowest order, the effective coupling above can be expressed in terms of the

energy scale, μ, as [17]:

$$\alpha_s(\mu^2) = \frac{\alpha_s(\mu_0^2)}{1 + (\beta_1/12\pi) \cdot \alpha_s(\mu_0^2) \cdot \ln(\mu^2/\mu_0^2)}, \tag{2.24}$$

where $\alpha_s(\mu_0^2)$ is the boundary condition for the coupling at a given energy scale μ_0 and β_1 is the first term in the expansion of the β function. This function measures the growth of the coupling following the renormalization group equation (RGE). It depends on the number of colours, $n_c = 3$, and the number of quark flavours, $n_f = 6$, taking the expression $\beta_1 = 11n_c - 2n_f$.

The energy dependence leads to a decrease of the coupling at high values of μ^2, meaning that at close distances quarks will behave quasi-freely, what is commonly referred to as *asymptotic freedom*. However, when μ^2 decreases, the coupling exponentially increases, meaning that perturbation theory approximations used in its calculation do not hold in this regime. This results into another key characteristic of QCD called *colour confinement*, which explains why there has been no observation of free coloured particles. If two quarks are to be separated from each other, the energy required to overcome the coupling increase when reaching long distances would greatly exceed that needed to form a quark-antiquark pair in the vacuum. This behaviour makes energetically favourable the appearance of hadrons made from $q\bar{q}$ pairs in colourless bound states, in a process called *hadronization*, which is of great importance when studying hadronic final states at particle colliders.

The value of α_s is conventionally considered at the energy scale of the mass of the Z-boson ($M_Z = 91.18$ GeV), at which the world average is [5]: $\alpha_s(M_Z^2) = 0.1179 \pm 0.0010$. A comparison of the running of the strong coupling constant, considering the previously mentioned value as the $\alpha_s(\mu_0^2)$ input for Equation 2.24, with different measurements can be found in Figure 2.1.

2.1.3 The Standard Model Lagrangian

When adding all the different elements previously presented into a single theory that encapsulates our current understanding of elementary particles and their interactions, the Standard Model Lagrangian takes the form:

$$\mathcal{L}_{\text{SM}} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{int}} + \mathcal{L}_{\text{Yukawa}} + \mathcal{L}_{\text{Higgs}} \tag{2.25}$$

- $\mathcal{L}_{\text{gauge}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$
  This term introduces the field strength tensor, $F_{\mu\nu}$, which encodes the dynamics of the interaction mediators of the Standard Model, namely gluons, photons, $W^{\pm}$ and Z^0. The tensor is defined as $F_{\mu\nu} = i\,[D_\mu, D_\nu]\,/g$, where D_μ corresponds to the covariant derivatives introduced in the previous sections for the $\text{SU}(2)_L \otimes \text{U}(1)_Y$ and $\text{SU}(3)_C$

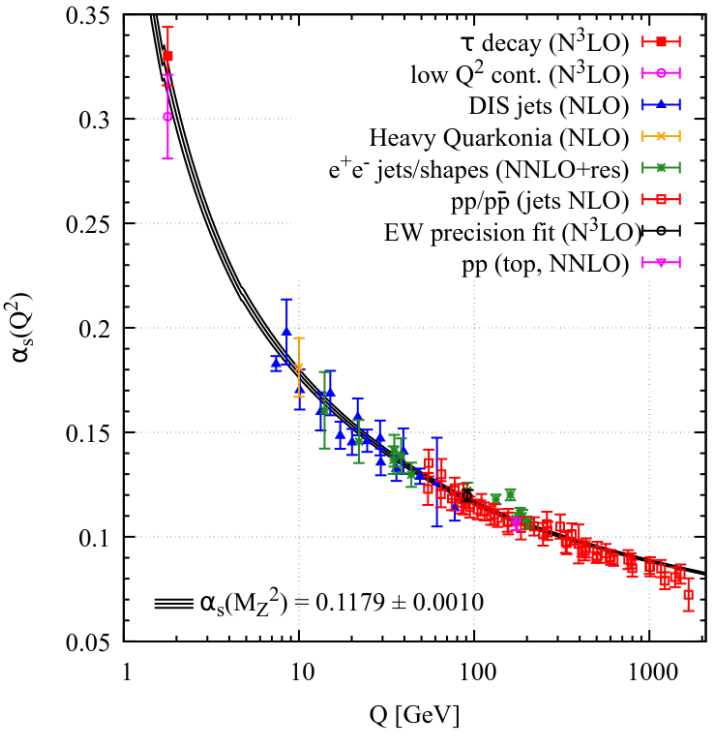


Figure 2.1: Summary of measurements of α_s as a function of the energy scale Q. From Ref. [5].

symmetries introduced in the previous Sections and g are the corresponding couplings to these interactions. The dynamics of the gauge bosons vary across the different mediators: while photons do not interact with themselves, meaning their kinematic term simply describes their free propagation, weak and strong force mediators do interact with each other, which is accounted for in the field strength tensor.

- $\mathcal{L}_{\text{int}} = i\bar{\psi}\not{D}\psi$
  This term was introduced when explaining how one could obtain an electromagnetic interaction invariant under local gauge transformations. The covariant derivative, $\not{D}$, introduces the interaction between the particle fields ψ, representing all fermions in SM, and the interaction bosons. The covariant derivative in Standard Model accounts for the unification of all symmetry groups aforementioned into a single $\text{SU}(3)_C \otimes \text{SU}(2)_L \otimes \text{U}(1)_Y$ symmetry.

- $\mathcal{L}_{\text{Yukawa}} = \psi_i \lambda_{ij} \psi_j \Phi$
  The field Φ corresponds to the Higgs field, as presented in Section 2.1.2. When coupling to fermions via the Yukawa term in the Lagrangian, the Higgs field makes their corresponding mass terms emerge. The strength of the interaction between a fermion, ψ_i and the Higgs field is measured by the Yukawa coupling, λ_{ii}, which is directly related to the mass of the corresponding fermion: the stronger the coupling of that said particle to the Higgs field, the higher its mass will be.

The off-diagonal terms, λ_{ij} with $i \neq j$, represent the coupling between different fermion flavours. The unitary matrix representing the coupling between different quark fermions is of particular relevance, known as the CKM matrix.

- $\mathcal{L}_{\text{Higgs}} = (D^\mu\phi)^\dagger(D_\mu\phi) - V(\phi)$; $V(\phi) = \mu^2(\phi^\dagger\phi) + |\lambda|(\phi^\dagger\phi)^2$
  This is the Higgs sector Lagrangian, as presented in Section 2.1.2, responsible for the spontaneous symmetry breaking mechanism that grants mass to particles in the Standard Model.

2.2 Physics Beyond the Standard Model

Although the Standard Model has been an incredibly successful theory, tested experimentally to very high precision for more than half a century, there are certain phenomena it fails to describe. These observations make apparent the need for a theory that complements and expands the foundations set by the Standard Model, while offering an explanation to those findings it cannot explain. This Section provides an overview of the evidence for the need of Beyond Standard Model (BSM) theories, as well as a brief description of some of their most common proposals.

2.2.1 Unsolved questions

Matter-antimatter asymmetry

The observation of visible matter around us, and the lack of antimatter, indicate that are differences between the behaviour of matter with respect to that of the antimatter, assuming they were produced equally at the Big Bang. However, no mechanism in Standard Model could produce such asymmetry as observed. The baryon asymmetry in the Universe is the name which refers to the excess in the number of observed baryons (N_B) over the observed antibaryons ($N_{\bar{B}}$). Since the end products of matter-antimatter annihilation are mainly photons, and considering there are no antibaryons in the universe today, the asymmetry can be measured using the baryon to photon ratio, $\eta = N_B/N_\gamma$, which is found to be $\sim 6 \cdot 10^{-10}$ [18]. This asymmetry can only be observed if three conditions are met, known as Sakharov conditions [19]: (1) baryon number violation, (2) C and CP (charge conjugation and parity) violation and (3) a deviation from thermal equilibrium. (1) is necessary for a system to evolve from $B = 0$ to $B \neq 0$. With a theory that allows for a process to produce matter asymmetry, C and CP violation denies the possibility of conjugate processes to generate an oppositely signed asymmetry occurring with the same probability, which would restore the original symmetry. Finally, since thermal equilibrium is a state that fulfils time translation invariance, a deviation from it is needed to produce a $B \neq 0$ state starting from $B = 0$.

Neutrino masses

As introduced in Section 2.1.2, neutrinos in the common Standard Model formulation are assumed to be massless, which subsequently means that all neutrinos in SM are considered to be left-handed, while anti-neutrinos are right-handed. Observations from many different experiments of neutrino oscillations, first observed in 1968 in the Homestake experiment [20], require massive neutrinos for a flavour change to be possible. Current experiments are trying to obtain the exact value of the neutrino masses, which would help determine whether they have an associated fermion with opposite chirality (as all massive fermions do in SM), meaning they are Dirac particles and that a right-handed neutrino spinor needs to be included within the Standard Model. Another theory proposes that neutrinos are in fact Majorana particles, meaning that they are their own antiparticle and that what has been described as an anti-neutrino in the Standard Model was actually referring to the same particle. Neutrino physicists keep looking for the answer to this enigma, as well as for an explanation on why would the mass of the neutrino, currently expected to be found in the meV range [5], be orders of magnitude smaller than that of any other fermion.

Grand unification

The development of theories such as electromagnetism or the electroweak theory, which unify different forces into a single theory at a given energy scale, have led to the idea that all forces observed in nature may be different expressions of a single, unified force. The simplest group that could contain all Standard Model forces is the SU(5) group [21] however, as shown in Figure 2.2, the Standard Model predicts no unification of its gauge couplings at any energy scale, requiring for BSM phenomena to alter these couplings and generate a unification scale.

Dark matter and dark energy

Upon contrasting the observed rotation velocity of galaxies with respect to their galactic center with the expected values, scientists found a behaviour that could not be explained with the existence of only the luminous matter in said galaxy [22]. While a lower velocity was expected at greater distances, caused by a decrease in the strength of the gravitational pull from the galactic center, the observations indicated that the velocity was approximately constant, remaining unchanged by the distance to the center. The explanation that emerged to justify this behaviour was the existence of non-luminous (*dark*) matter, which could not be directly observed but increased the gravitational effect observed in the galaxy's halo. The existence of dark matter has further been supported by the observation of gravitational lensing [23] and by cosmic microwave background measurements [24].

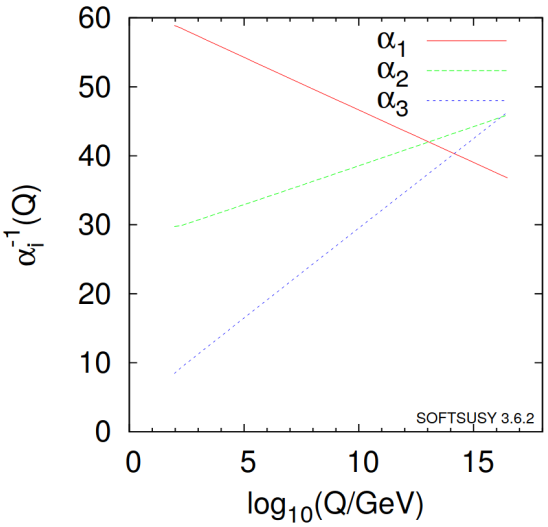


Figure 2.2: Running of all gauge couplings in the Standard Model. From Ref. [5].

The lack of observation of any dark matter candidate that fulfils all necessary conditions strongly constraints its properties, for instance its candidates only being able to interact weakly and gravitationally.

Alternative observations also show that the universe expansion takes place at a rate that our current cosmological theories fail to predict. This acceleration needs for some sort of energy not observed yet, which is analogously referred to as *dark energy*. Little is known about what physics may be behind this phenomena, although observations in the cosmic microwave background have constraint the luminous matter content of the Universe to a 5% [25], with a 27% content being dark matter and the remaining 68% dark energy, further stressing the point of how little is known about the content of the Universe and how important it is to find an explanation to it.

2.2.2 Beyond Standard Model theories

This Section will provide an overview of a selection of Beyond Standard Model theories, with a particular focus on those involving new heavy gauge bosons phenomena.

Sequential Standard Model

The Sequential Standard Model (SSM) [26], assumes the appearance of two charged bosons and a neutral boson which, in analogy to those found in the Standard Model are referred to as $W'^{\pm}$ and Z' bosons. These new bosons share the same quantum numbers than their SM partners, although they are assumed to have a much larger mass, with current limits set in the range beyond the TeV scale [27–29]. Although this model is not believed to be realistic, it provides a common ground to compare constraints coming from different

sources and allow for more complex reinterpretations.

Left-right symmetric models

One of the simplest extensions of the Standard Model involving additional gauge bosons (while providing an explicit description of the gauge theory behind it) is the $SU(2)_{\mathrm{L}} \times SU(2)_{\mathrm{R}} \times U(1)$ model [30]. This theory expands the parity violating left-handed symmetry present in electroweak theory with the addition of right-handed doublets of a $SU(2)_{\mathrm{R}}$ theory. The addition of this group leads to the appearance of three additional bosons, $W_{\mathrm{R}}'^{\pm}$ and Z'. Considering that there is experimental evidence of parity violation in weak interactions, this new symmetry must be broken:

$$SU(2)_{\mathrm{L}} \times SU(2)_{\mathrm{R}} \times U(1) \rightarrow SU(2)_{\mathrm{L}} \times U(1)_{\mathrm{Y}} \tag{2.26}$$

In this process, the additional gauge bosons that emerge from the $SU(2)_{\mathrm{R}}$ group can gain mass via spontaneous symmetry breaking, in a similar manner to the Higgs mechanism with which the Standard Model W and Z bosons obtain mass. As stated before, parity violation has been observed at the energy scales probed so far by colliders, meaning that the emerging $W_{\mathrm{R}}'^{\pm}$ must be very heavy and the interaction mediated by them heavily suppressed at low energies.

Extra dimensional models

Other theories predicting the appearance of additional gauge bosons assume the hypothetical heavy partners to be the Kaluza-Klein (KK) excitations of the Standard Model bosons [31]. In these theories, the usual (3+1) structure of space-time dimensions is expanded to a (3+δ+1) structure. The main (3+1)-dimensional space time is here known as the *brane*, the main structure on which the Standard Model sits upon, to which δ additional spatial dimensions are added, generating excitations associated to the appearance of new gauge bosons [32]. An interesting aspect of these theories is that they also provide an explanation for the hierarchy problem. The graviton (hypothetical mediator of the gravitational force) propagating through extra dimensions, not accessible at the energy range probed by our current experiments, would be the reason for gravity being so weak compared to the other interactions present in the Standard Model, as proposed in the Randall-Sundrum model [33].

Supersymmetry

Within the Standard Model, corrections to the Higgs boson mass emerging from higher order calculations need for fine-tuning to avoid divergent behaviour. The cancellation of those contributions can be naturally achieved by adding a new symmetry relating fermions

and bosons, widely known as *supersymmetry* (SUSY) [34]. This symmetry is achieved introducing a new operator Q, that operates on the Standard Model particles as:

$$Q\,|\text{Boson}\rangle = |\text{Fermion}\rangle\,; \quad Q\,|\text{Fermion}\rangle = |\text{Boson}\rangle \tag{2.27}$$

The resulting new particles when applying said operator are called *super-partners*. If this transformation operated under an exact symmetry, the resulting superpartners would mimic the mass and quantum numbers of their corresponding SM particle, however no integer spin particle has been found at the mass range of the SM particles, an energy range that particle physics experiments have been able to study for decades, from which can be inferred that supersymmetry must be a broken symmetry. This introduces a myriad of free parameters in the theory, with each mass and coupling of any given superpartner being a free variable. This feature incentivised the development of theories like the Minimal Supersymmetric Standard Model (MSSM) [35], where the number of new particles and interactions is minimized, resulting in a reduced numbers of free parameters, while still offering the benefits that all supersymmetric theories bring. These include providing a solution to gauge unification, absorbing divergences in Higgs mass corrections and offering potential candidates for dark matter, since neutrally-charged superpartners fulfil all conditions dark matter candidates need. Unfortunately, no sign of supersymmetry have been found yet [36–39], meaning that the energy scale of said phenomena is likely beyond the TeV frontier.

2.3 Effective Field Theory interpretations

As presented in the previous Section, indications of deviations from SM hint at high energy scales of new physics. One of the approaches to describe phenomena beyond that comprised in the Standard Model is by introducing higher-dimensional operators, whose effects are suppressed by powers of the new physics energy scale (Λ). Effective field theories (EFTs) are the theoretical tools to describe *low*-energy physics, where the definition of "low" depends on the energy scale Λ. In these theories, heavy states (those with $M \gg \Lambda$) are integrated out of the action while only the relevant degrees of freedom ($m \ll \Lambda$) are taken into account. Thus, the information on the heavier states is encoded in the couplings of the resulting low-energy Lagrangian, in which the interactions between light states are organized as an expansion in powers of $1/\Lambda$.

This Section presents the EFT frameworks used for the interpretation of the data presented in Parts III and IV.

2.3.1 Standard Model as an Effective Field Theory

The Standard Model Effective Field Theory (SMEFT) [40] is a generalization of the Standard Model where the low-energy regime corresponds to the Standard Model Lagrangian, as presented in Section 2.1.3, assuming that any extension to it does not involve new particles lighter than the measured vaccuum expectation value. The Lagrangian of this theory is defined as:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \sum_{d>4} \mathcal{L}^{(d)} = \mathcal{L}_{\text{SM}} + \sum_{i} \frac{c_i^{(d)}}{\Lambda^{d-4}} \mathcal{O}_i^{(d)} \tag{2.28}$$

where:

- $\mathcal{L}_{\text{SM}}$ is the EFT description of the Standard Model operators (containing all valid operators with dimension 2[†] and 4).
- $\mathcal{O}_i$ are the different operators of dimension d.
- c_i are the Wilson coefficients, constant terms that measure the impact of the different operators.

An overview of the calculation of the cross-section of any given process (which is the initial basis for Standard Model simulation) can be found in Section 3.1. When introducing a set of additional operators, $\mathcal{O}_i$, in the Standard Model Lagrangian the corresponding amplitude of a given process ($\mathcal{A}$) and hence its cross section, will be modified as follows:

$$|\mathcal{A}_{\text{SM}} + \sum_i c_i \mathcal{A}_i|^2 = \underbrace{|\mathcal{A}_{\text{SM}}|^2}_{\text{SM}} + \underbrace{\sum_i c_i 2\text{Re}(\mathcal{A}_{\text{SM}}^* \mathcal{A}_i)}_{\text{Linear}} + \underbrace{\sum_i c_i^2 |\mathcal{A}_i|^2}_{\text{Quadratic}} + \underbrace{\sum_{i \neq j} c_i c_j 2\text{Re}(\mathcal{A}_{\text{i}}^* \mathcal{A}_j)}_{\text{Cross}} \tag{2.29}$$

In the right hand side of Equation 2.29 different contributions can be distinguished: the first term corresponds to the SM cross section. The second term is the SM-BSM interference, usually referred to as the *linear* term. The third term represents the pure EFT contribution from the operator $\mathcal{O}_i$, namely the *quadratic* term. Lastly, the fourth term is the interference between different EFT operators, known as the *cross* term.

Flavour symmetry assumptions

The SMEFT interpretation carries one main drawback, similar to such present in most SUSY models: in its most accepted formulation, the Warsaw basis [41], SMEFT has nearly 2500 free parameters [40, 42], most of them coming from flavour indices running over the

[†]The only dimension-2 term in the SM Lagrangian is $\phi^\dagger \phi$, the Higgs mass term.

three fermion generations. To reduce this number there are several symmetry assumptions that can be adopted. Flavour symmetry assumptions are well motivated by observations from low energy scale experiments, leading to one of the most commonly used assumptions, the $U(3)^5$ limit [42], which reduces the number of free parameters in the theory down to 93.

This global symmetry is defined by a relation between the weak (unprimed) and mass (primed) basis of the Standard Model SU(2) doublets and singlets:

$$q = \mathcal{U}_q q'; \quad u = \mathcal{U}_u u'; \quad d = \mathcal{U}_d d'; \quad \ell = \mathcal{U}_\ell \ell'; \quad \nu = \mathcal{U}_\nu \nu'. \tag{2.30}$$

Each rotation defines a U(3) flavour group, creating a SM $U(3)^5$ group defined as:

$$U(3)^5 = \mathcal{U}_q \times \mathcal{U}_u \times \mathcal{U}_d \times \mathcal{U}_\ell \times \mathcal{U}_\nu. \tag{2.31}$$

Note how the rotations in the neutrino and quark sectors define the existing PMNS and CKM matrices in Standard Model.

By imposing symmetry under this $U(3)^5$ group the flavour structure of the EFT operators can be factored out, transforming their corresponding Wilson coefficients into scalar quantities and thus significantly reducing the number of free parameters in the theory.

2.3.2 Breaking Lorentz-invariance in SM: Standard Model Extension

One of the fundamental principles of the Standard Model is Lorentz invariance, stating that physical laws should be independent of the relative orientation and velocity of an experiment in spacetime. Lorentz transformations connect observables from experiments oriented or boosted relative to one another and are position independent, meaning that Lorentz invariance applies globally. A consequence of this is the invariance of any process under the combined operation of charge conjugation C, parity P, and time reversal T, as stated in the CPT theorem. The assumption of the invariance of the theory under Lorentz transformations is strongly motivated theoretically and is accompanied by experimental measurements indicating no significant Lorentz invariance violation in a broad range of energies and systems [43]. However, early work in string theory demonstrated that Lorentz invariance could be spontaneously broken by nonzero VEVs of tensor fields [44], meaning that experiments sensitive to the coupling between said fields and SM particles could indirectly probe string interactions through the spontaneous breaking of Lorentz invariance.

In absence of experimental hints, the most conservative approach to test for Lorentz invariance is using a model-independent framework based in Effective Field Theory, commonly known as the Standard Model Extension (SME) [45, 46]. A general SME lagrangian can be expressed as

$$\mathcal{L}_{\mathrm{SME}} = \mathcal{L}_{\mathrm{SM}} + \mathcal{L}_{\mathrm{LV}} \tag{2.32}$$

where the action of the Lorentz invariance-violating term, $\mathcal{L}_{\mathrm{LV}}$, must fulfil

$$|S_{\mathrm{LV}}| = \left| \int \mathcal{L}_{\mathrm{LV}} \, d^4x \right| \ll |S_{\mathrm{SM}}| \tag{2.33}$$

The terms introduced in the SME Lagrangian transform in an unconventional manner, best understood by example. Consider a term in the QED sector,

$$\mathcal{L}_b = -b_\mu \bar{\psi} \gamma_5 \gamma^\mu \psi, \tag{2.34}$$

where $b_\mu = (b_0, b_i)$ is a SME coefficient consisting of an array of four numbers. Since the Lorentz index μ is contracted, $\mathcal{L}_b$ is invariant under transformations of the coordinate system, known as *observer* transformations [45]. However, if the physical system ψ is rotated or boosted, while leaving the coordinates unaffected (performing a *particle* transformation), this SME term would vary, since the coefficients b_μ remain the same under the transformation, but the physical system ψ does not. The term b_μ can be understood as a background vector field sitting at every point in the vacuum. A schematic diagram of this process can be found in Figure 2.3.

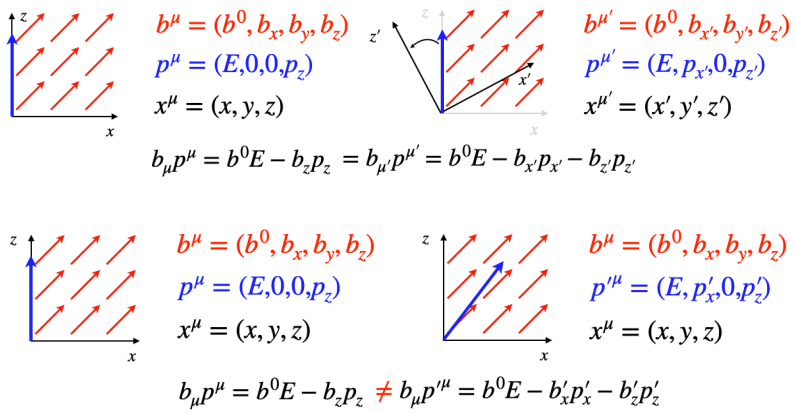


Figure 2.3: Schematics of the (top) observer (bottom) and particle transformations of a given $\mathcal{L}_b$ term in the SME. The presence of a background vector field (red) results into a different interaction with the particle (blue) under each transformation, changing the observation made in the reference plane (black).

Drell-Yan production under SME

The calculation of the Standard Model Drell-Yan production of dilepton pairs from proton-proton collision is presented in-depth in Section 3.1.2. In SME, additional interactions are considered between the quark sector of the Standard Model and the background vector fields, $c_f^{\mu\nu}$, which can be represented as [47]:

$$\mathcal{L}_{\text{LV}} = \frac{1}{2} i c_q^{\mu\nu} \bar{q} \gamma_\mu D_\nu q + \frac{1}{2} i c_u^{\mu\nu} \bar{u} \gamma_\mu D_\nu u \frac{1}{2} i c_d^{\mu\nu} \bar{d} \gamma_\mu D_\nu d, \tag{2.35}$$

where D_ν is the gauge-covariant derivative and the fields correspond to the SU(2) quark singlets and doublets in the SM:

$$q = \begin{pmatrix} u \\ d \end{pmatrix}_L, \quad u = (u)_R, \quad d = (d)_R. \tag{2.36}$$

New coefficients can be defined using those present in Equation 2.35:

$$\begin{aligned} c_U^{\mu\nu} &\equiv (c_q^{\mu\nu} + c_u^{\mu\nu})/2, \quad c_D^{\mu\nu} \equiv (c_q^{\mu\nu} + c_d^{\mu\nu})/2, \\ d_U^{\mu\nu} &\equiv (c_q^{\mu\nu} - c_u^{\mu\nu})/2, \quad d_D^{\mu\nu} \equiv (c_q^{\mu\nu} - c_d^{\mu\nu})/2, \end{aligned} \tag{2.37}$$

in such way that $c_U^{\mu\nu} - c_D^{\mu\nu} = d_U^{\mu\nu} - d_D^{\mu\nu}$. These coefficients are of particular relevance for electroweak-mediated processes, since left- and right-handed fields will couple differently to the SM gauge bosons [47]. The effect of the contributions coming from said operators on the Z-boson production cross-section, calculated at LO and in the full fiducial phase space, can be found in Figure 2.4. All contributions have a decreasing cross-section from low interaction energies until approaching the Z-pole ($Q = 91$ GeV), where parity-violating effects are maximal and d coefficients show the biggest sensitivity, with cross-section increases of up to ~ 3 pb for illustrative values of $d_U^{33} = 10^{-5}$.

When the relative position of the experiment a measurement takes place in changes with respect to the orientation of the SME background fields, the cross-section of the processes involved in the interaction with said fields will change, meaning that the cross-section will become a time-dependent observable. See Section 14.1 for further explanations on the implications of SME on neutral-current Drell-Yan production at the ATLAS experiment.

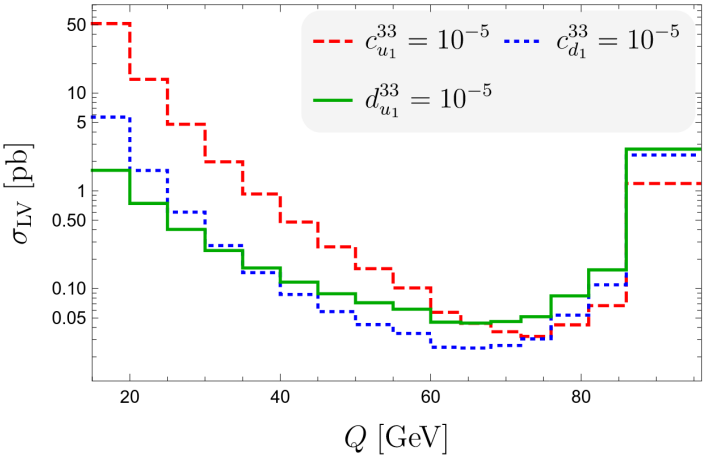


Figure 2.4: Evolution of the Lorentz-invariance violating contributions to the Z-boson production cross section, σ_{LV}, against the energy scale Q. From Ref. [47].

Chapter 3

Modelling of proton-proton interactions

This Chapter provides a description of the phenomenological models used in the description of proton-proton interactions, as well as the steps involved in generating theoretical predictions to compare our data to. Special emphasis is given to the phenomenology of the Drell-Yan process, including an overview of the commonly used kinematic variables to describe the process. Also presented is a summary of the main Monte Carlo generators of relevance for this thesis.

3.1 Phenomenology of proton interactions

Protons are baryons with a complex structure, meaning that hadron-hadron collisions are more complex to describe than lepton-lepton collisions in previous experiments like LEP, as leptons are point-like particles.

The theories used to describe proton-proton interactions are the *parton model* and the *factorisation theorem*. Partons, quarks and gluons found inside the proton, are considered to interact individually, rather than both hadrons taking part in the collision interacting as a whole. Thus, the probability of two protons colliding to produce a certain final state X will be given by the probability of certain partons i and j, each found in each of the interacting protons and carrying a fraction of said proton's momentum, x_i and x_j, to interact times the cross section $\sigma_{ij\to X}$. The cross-section measures the probability of a given process $ij \to X$ to happen, which is calculated from the transition amplitude $\mathcal{A}_{ij\to X}$ following Standard Model calculations [48]. The probability of finding a parton i inside a proton carrying a momentum fraction x is given by the Parton Distribution Functions (PDFs), $f_i(x)$. The total cross section for any given process can be then expressed through the factorisation theorem:

$$\sigma_{pp\to X} = \sum_{i,j} \int_0^1 \int_0^1 \sigma_{ij\to X}(x_i, x_j) f_i(x_i) f_j(x_j) dx_i dx_j, \tag{3.1}$$

where i and j loop through all possible parton combinations that can produce the final state X. PDFs must fulfil the momentum sum rule imposed by momentum conservation,

$$\int_0^1 dx \sum_i x f_i(x) = 1. \tag{3.2}$$

The considerations presented so far do not include processes like those where a parton emits or splits into a different one, which depend on the energy scale considered in the process. When taken into account, the expression above results into:

$$\sigma_{pp\to X} = \sum_{i,j} \int_0^1 \int_0^1 [\underbrace{\hat{\sigma}_0}_{\text{LO}} + \underbrace{\hat{\sigma}_1 \alpha_s(\mu_R^2)}_{\text{NLO}} + \underbrace{\hat{\sigma}_2 \alpha_s^2(\mu_R^2)}_{\text{NNLO}} ...](x_i, x_j, \mu_F^2) \times f_i(x_i, \mu_F^2) f_j(x_j, \mu_F^2) dx_i dx_j, \tag{3.3}$$

where μ_F is the factorization scale, which separates the QCD calculations into short- and long-distance pieces [49] and μ_R is the renormalization scale, which defines the scaling of α_s that the partonic cross section $\hat{\sigma}$ depends on. Note that the latter has been expressed as an expansion in α_s. The first term expansion is known as the *leading order* (LO) contribution, with the following terms being referred to as *next-to-leading order* (NLO), *next-to-next-to-leading* order (NNLO), and so on.

3.1.1 Measurement of the parton distribution function

While Standard Model predictions for the partonic cross-sections in Equation 3.3 can be obtained, there is no way to obtain the non-perturbative PDF shapes in x and their perturbative evolution in Q^2, which must be determined from data. Leading measurements to this day still come from deep inelastic scattering (DIS) experiments, which use a lepton to probe the inner structure of the proton. Most notably, we find the results of the H1 and ZEUS experiments [50] at the HERA ep accelerator. These experiments can be complemented by fixed-target experiments or with the addition of LHC data to achieve higher accuracy.

Recent determinations of PDFs, including up to NNLO QCD corrections, have been led by different groups: MSHT [51], NNPDF [52], CT(EQ) [53] and HERAPDF [50], among others. Most groups use the following input PDFs: $xf = x^a(...)(1-x)^b$, with as many as 28 free parameters. These are fit on the different available datasets, expanding the free parameters around their best fit values and obtaining orthogonal eigenvector sets of PDFs from linear combinations of different parameter variations. An example of the results obtained by the MSHT group is shown in Figure 3.1, where the probability of finding each

parton within the proton is shown against the proton momentum fraction x it carries.

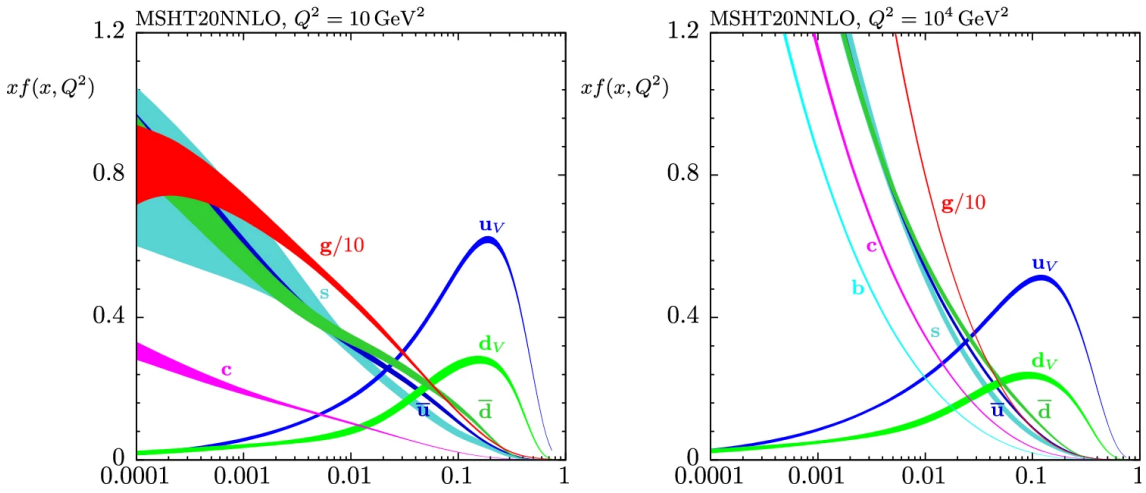


Figure 3.1: Parton Distribution Functions of all proton constituents for the MSHT20 NNLO PDFs [51] at $Q^2 = \mu_F^2 = 10$ GeV^2 (left) and $\mu_F^2 = 10^4$ GeV^2 (right).

3.1.2 Dilepton production in proton-proton collisions

One of the cleanest reactions one can study as a product of proton-proton interactions, which is also of particular importance for the work presented in this thesis, is the production of lepton pairs emerging from the collision. This reaction is called the *Drell-Yan process* (DY) [54], involving a $q\bar{q}$ interaction, which produces a vector boson that decays into two leptons, as displayed in Figure 3.2. Depending on the quarks involved in the interaction the process can be mediated by a Z-boson, resulting into *neutral-current* (NC) Drell-Yan, or a $W^{\pm}$- boson, a process analogously named *charged-current* (CC) DY. The processes can also be mediated by off-mass shell bosons which do not satisfy the usual energy-momentum relationship ($E_{V^*}^2 \neq p_{V^*}^2 + m_{V^*}^2$). This is of particular relevance in the case of neutral-current Drell-Yan since the quantum numbers of virtual photons (γ^*) are indistinguishable from those of a Z-boson. The decay products of each process will differ due to charge conservation, neutral-current DY resulting into a lepton-antilepton pair of leptons (e, μ, τ) and charged-current DY decaying into a (anti)lepton-(anti)neutrino pair.

Neutral-current Drell-Yan kinematics

The measurement of the Drell-Yan process relies on the reconstruction of the dilepton system produced in the vector boson's decay, combining each of the leptons' four-vectors, $p_{\ell,1}^{\mu}$ and $p_{\ell,2}^{\mu}$, into the dilepton four-vector $p_{\ell\ell}^{\mu}$. As shown in Section 3.1, each of the partons that interact to produce the vector boson carries a fraction x_i of the proton's momentum,

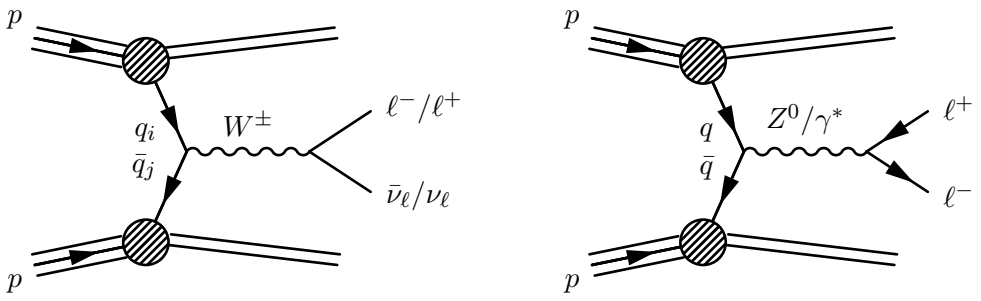


Figure 3.2: Feynman diagram showing the leading order representation of charged- (left) and neutral-current (right) Drell-Yan process.

$p^{\mu}_{p,i}$. The squared transferred momentum in the interaction, Q^2, is expressed as:

$$Q^2 = \left(x_1 p^{\mu}_{p,1} + x_2 p^{\mu}_{p,2}\right)^2 = x_1 x_2 s, \tag{3.4}$$

where s is the squared centre-of-mass energy of the proton-proton collision. In di-lepton final states Q is equal to the dilepton invariant mass, $m_{\ell\ell}$, which expressed in terms of the lepton's four-vectors takes the form:

$$m_{\ell\ell} = \sqrt{\left(p^{\mu}_{\ell,1} + p^{\mu}_{\ell,2}\right)^2}. \tag{3.5}$$

If we define the rapidity y of a particle by:

$$y = \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right), \tag{3.6}$$

where p_z is the component of the momentum along the beam axis, one finds that the rapidity of the dilepton system is related to the momentum fraction carried by the initial partons. At LO this relation can be expressed as:

$$x_{1,2} = \frac{m_{\ell\ell} \cdot e^{\pm y_{\ell\ell}}}{\sqrt{s}}. \tag{3.7}$$

A variable of particular relevance for the description of the Drell-Yan cross-section is the decay angle of the lepton pair. It is useful to define the angles of the outgoing leptons with respect to the incoming partons, but the latter may not necessarily align with the incident proton directions. At leading order, the transverse momentum $\left(p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}\right)$ of the dilepton system is zero, matching that of the incoming protons. However, when introducing higher order corrections (and therefore the presence of initial state radiation) this is no longer true, making the laboratory frame a poor choice for most angular measurements. A common approach to angular measurements in dilepton final states is to adopt the Collins-Soper frame [55]. In this frame, the angles are defined with respect to the negatively charged lepton's momenta and a longitudinal axis that bisects the momentum vectors of the incoming partons (defining the Z-boson rest frame), as shown in Figure 3.3. The z-axis

is defined by the direction of positive longitudinal Z-boson momentum in the laboratory frame, the y-axis as the normal vector to the plane spanned by the parton's momenta, and the x-axis completes a right-handed Cartesian coordinate system with the aforementioned axes. Of particular importance is the polar angle defined in the Collins-Soper frame, θ^*_{CS}, which takes the expression:

$$\cos(\theta^*_{\text{CS}}) = \frac{p_z^{\ell\ell}}{|p_z^{\ell\ell}|} \frac{2\left(p_1^+ p_2^- - p_1^- p_2^+\right)}{m_{\ell\ell}\sqrt{m_{\ell\ell}^2 + \left(p_{\text{T}}^{\ell\ell}\right)^2}}, \tag{3.8}$$

where

$$p_i^{\pm} = \frac{1}{\sqrt{2}\left(E_i \pm p_{z,i}\right)} \tag{3.9}$$

and $i = 1, 2$ refers to the negatively- and positively-charged lepton respectively.

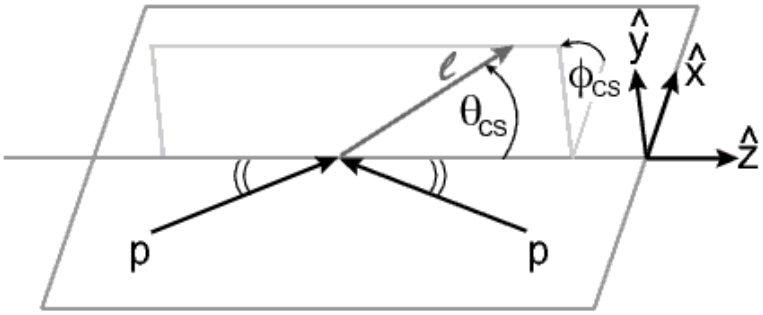


Figure 3.3: Illustration of the Collins-Soper reference frame. Angles defined with respect to the negatively-charged lepton, ℓ. See text for detailed definitions of the axes and angles. From Ref [56].

To increase the information extracted from data, measurements of the Drell-Yan cross-section are usually performed in different kinematic variables, i.e. differentially, as presented in the measurement covered in Part III. At leading order, the Drell-Yan differential cross-section in the previously presented kinematic variables, $m_{\ell\ell}$, $y_{\ell\ell}$ and $\cos(\theta^*_{\text{CS}})$, is given by [57]:

$$\frac{\mathrm{d}^3\sigma}{\mathrm{d}m_{\ell\ell}\mathrm{d}y_{\ell\ell}\mathrm{d}\cos(\theta^*_{\text{CS}})} = \frac{\pi\alpha^2}{3m_{\ell\ell}s}\sum_q P_q\left[f_q(x_1, Q^2)f_{\bar{q}}(x_2, Q^2) + (q \leftrightarrow \bar{q})\right], \tag{3.10}$$

where α is the QED fine structure constant, s is the squared proton-proton centre-of-mass energy, x_i are the incoming parton momentum fractions, and f_q ($f_{\bar{q}}$) are the PDFs for each parton flavour q (anti-flavour $\bar{q}$). The symmetry $q \leftrightarrow \bar{q}$ term accounts for the exchange of the parent protons of the quark and antiquark taking part in the interaction. The rapidity dependence arises from the dependence of x_i on $m_{\ell\ell}$, $y_{\ell\ell}$ and $\sqrt{s}$, as shown

in Equation 3.7. The function P_q is given by:

$$
\begin{aligned}
P_q = & \underbrace{e_\ell^2 e_q^2}_{\gamma^* \text{ exchange}} \quad (1+\cos^2\theta^*_{\text{CS}}) \\
& + \underbrace{e_\ell e_q \frac{2m_{\ell\ell}^2(m_{\ell\ell}^2 - m_Z^2)}{\sin^2\theta_W \cos^2\theta_W \left[(m_{\ell\ell}^2 - m_Z^2)^2 + \Gamma_Z^2 m_Z^2\right]}}_{Z/\gamma^* \text{ interference}} \left[v_\ell v_q (1+\cos^2\theta^*_{\text{CS}}) 2a_\ell a_q \cos\theta^*_{\text{CS}}\right] \\
& + \underbrace{\frac{m_{\ell\ell}^4}{\sin^4\theta_W \cos^4\theta_W \left[(m_{\ell\ell}^2 - m_Z^2)^2 + \Gamma_Z^2 m_Z^2\right]}}_{Z \text{ exchange}} \left[(a_\ell^2 + v_\ell^2)(a_q^2 + v_q^2)(1+\cos^2\theta^*_{\text{CS}}) + 8a_\ell v_\ell a_q v_q \cos\theta^*_{\text{CS}}\right].
\end{aligned}
\tag{3.11}
$$

The first term in the expression corresponds to pure virtual photon (γ^*) exchange, the second arises from the interference of γ^* and Z exchange and the third term corresponds to pure Z boson exchange. The function depends on some of the fundamental parameters of the particles involved in the interaction: m_Z and Γ_Z are the Z boson mass and decay width respectively; e_f are the fermion (leptons, ℓ, and quarks, q) electric charges and a_f and ν_f are the fermion axial and vector couplings:

$$
\begin{aligned}
a_f &= T^3 \\
v_f &= T^3 - 2e_f \sin^2\theta_W,
\end{aligned}
\tag{3.12}
$$

where T^3 is the fermion weak isospin (see Tables 2.1 and 2.2).

The DY invariant mass spectrum is characterized by a $1/m_{\ell\ell}^2$ falling dependence from the γ^* term contribution and a Breit-Wigner peak resonance with center at m_Z and width Γ_Z. The Z/γ^* interference is negative at $m_{\ell\ell} < m_Z$ and changes sign at $m_{\ell\ell} = m_Z$, value above which the contribution is positive. Figure 3.4 shows the relative contribution of each term as a function of $m_{\ell\ell}$.

Photon-induced dilepton production

The calculations presented so far have only taken into consideration the dilepton production via $q\bar{q}$ interactions. However, the quarks in the protons can radiate photons since they are electrically charged, producing a photonic structure within the proton, besides the partonic structure previously discussed. This creates an additional mechanism for dilepton production, namely the *photon-induced* production mechanism, in which two photons scatter to produce a lepton pair. The leading-order Feynman diagrams of the process can be found in Figure 3.5. The contribution from this process becomes larger at high invariant dilepton masses, resulting into an additional background for searches looking for heavy resonances decaying into lepton pairs. Further details can be found in Part III.

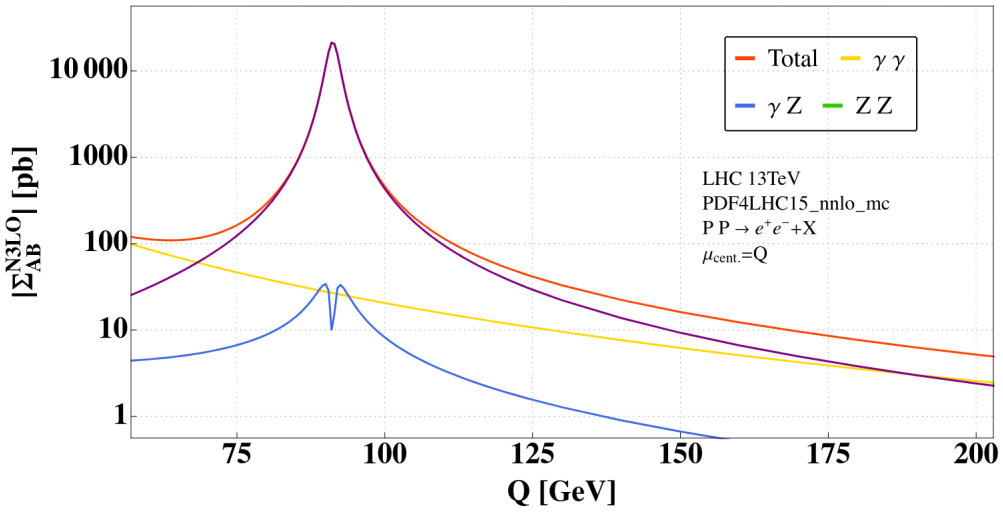


Figure 3.4: Neutral-current Drell-Yan cross-section at N3LO, Σ^{N3LO}, dependence in $m_{\ell\ell}$ at $\sqrt{s} = 13$ TeV. Shown also is the cross-section decomposition into the absolute value of contributions from γ^* and Z boson exchange and their interference. From Ref [58].

Figure 3.5: Leading-order Feynman diagrams showing the photon-induced dilepton production.

The photon contribution in the proton is usually not accounted for in most PDFs used at LHC. However, the increased importance of photon-induced processes at high energies revived the interest in describing the photon contribution in the proton. Few examples of PDFs taking into account this contribution are the MRST2004qed PDF set [59], the NNPDF2.3qed PDF set [60] and the LUXqed17 PDF set [61].

3.2 Monte Carlo simulation

The complexity involved in proton-proton collisions makes the precise description of the final states produced after the collision a very challenging task. Monte Carlo generators split the processes between the initial and final states in different steps, as shown in Figure 3.6. This Section provides an overview of the most relevant parts of these steps. In-depth descriptions of the calculations involved in these processes can be found in Ref. [62].

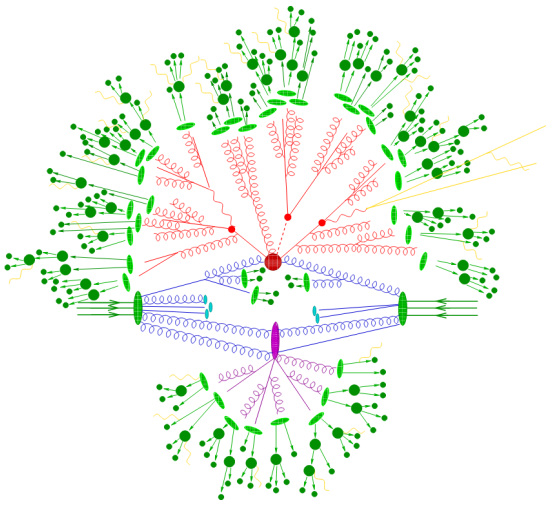


Figure 3.6: Schematic representation of a $t\bar{t}h$ event as simulated by an event generator. After the hard interaction (big red circle), top and Higgs decays are simulated (small red circles). QCD radiation emerges from the different products (red), and a secondary interaction takes place between both protons (purple blob). The final state partons hadronize (light green blobs) and the resulting hadrons decay (dark green circles). QED radiation is produced at all stages of the process (yellow). From Ref [63].

3.2.1 Simulation steps

Matrix element calculation

The partonic cross section of the production of any final state ij can be calculated using the matrix element ($\mathcal{M}$), summing over all possible Feynman diagrams ($\mathcal{F}$) participating in the process. For example, the matrix element of a quark-antiquark interaction decaying into a final state ij will be given by:

$$\mathcal{M}_{q\bar{q}\to ij} = \sum_{a} \mathcal{F}^{a}_{q\bar{q}\to ij} \tag{3.13}$$

This calculation, as presented in Section 3.1, can be performed with different levels of precision. At leading order, Feynman diagrams like those presented in Figure 3.2 are relatively simple and do not allow for many combinations, meaning that they can be quickly calculated by most MC generators. Higher order calculations involving emission of QED and QCD radiation, as well as the appearance of loops in the Feynman diagram (as shown in Figure 3.7) are more challenging. Not only there are more diagrams that need to be accounted for, but higher order effects lead to the appearance of divergences in the calculations, which need to be correct for.

Figure 3.7: Feynman diagrams showing the neutral-current Drell-Yan production at NLO QED (left) and NNLO QCD (right).

Parton showering

The matrix elements only account for interactions involving partons above some hard scale. These can produce QED or QCD radiation (photons and gluons), which can further split or radiate more particles. The modelling of the evolution of the partons, as well as all their subsequent emissions is performed by *parton shower* algorithms. These calculations are done in a step-by-step manner, calculating the probability of a parton to split using the DGLAP equations [64], evolving the momentum transfer from the energy scale of the hard process to the energy scale where the theory breaks down (Λ_{QCD}). In each step, the probability of evolving the process to a lower energy scale without radiating a gluon or splitting is given by the Sudakov form factor, which accounts for virtual effects (coming from quantum loops in the Feynman diagrams) and keep the probability of branching to unity, which removes divergences from the calculation. Once the showering evolves to the energy cut-off, all branching probabilities reach zero and partons become confined to hadrons. Simulation also takes into account the decays of those hadrons produced which are not stable, meaning they have a lifetime $c\tau < 10$ mm.

Additional products from proton-proton collisions

In the previous discussion only the products from the main interaction between the colliding protons have been presented. However, there processes in pp collisions that may produce additional particles that can be detected in the final state together with the hard scatter products. Some of those additional processes are listed below.

- **Final state radiation**: as explained in Section 3.1 the parton model assumes constituents in the colliding protons to behave as free particles. However, gluon emissions generate a large quantity of radiation by further splitting into $q\bar{q}$ pairs. Analogously, particles interacting electromagnetically may emit photons, which can further split into electron-positron pairs which can contribute to the radiation detected together with the hard scatter products. Radiation produced in this manner from particles emerging in the showering and hadronisation processes is known as *final state radiation* (FSR).

- **Initial shower radiation**: while contributions from FSR can be estimated from the evolution of the particles taking part in parton showering, *initial state radiation* (ISR) must preserve the conservation of the momenta in the hard-scatter. Thus, the development of ISR is calculated by evolving backwards from the scale of the hard process to the initial partons [65, 66].

- **Underlying event**: the matrix element calculated from the hard scatter of both colliding protons only considers the interaction between one pair of partons for each pp interaction. However, additional parton interactions typically take place alongside the hard scatter. These additional parton interactions will undergo the same process the main interaction, usually at a lower energy scale, producing what is known as *underlying event*.

- **Pile-up**: all interactions considered so far were consider to originate from the collision of the same proton-proton pair. However, to achieve the highest instantaneous luminosity as possible and produce a more efficient data-taking, LHC collides bunches of protons (see Chapter 4). Due to the density of these proton bunches, it is likely that more than one proton-proton collision occurs upon the bunches crossing. Background particles coming from additional proton collisions other than that producing the hard scatter of interest are referred to as *pile-up*. This behaviour is modelled in Monte-Carlo by overlaying multiple inelastic pp collisions obtained with Pythia 8 [67]. The changing pile-up conditions observed at LHC due to different bunch densities used in different data-taking periods is accounted for by reweighing the events, using pile-up profiles measured in the observed data.

Detector simulation

All products obtained in the steps previously explained are considered to be an ideal representation of the behaviour of particles produced in the proton-proton collisions, referred to as *truth-level*. These are in contrast with what are known as *reconstructed-level* particles. The latter are the result of the outgoing particles interacting with our detector, which smears their properties such as momentum or energy, due to the detector's resolution effects. This is simulated using a toolkit called Geant4 [68, 69], which uses a complete description of the detector and models the particles' trajectories through it and all their subsequent interactions.

3.2.2 Event generators

There is a plethora of Monte-Carlo generators available to describe the processes taking place in LHC collisions. This Section provides a description of those used in the results presented in this thesis.

Herwig [70, 71] is a general-purpose generator which can compute hard processes at LO and NLO with a special emphasis on providing an accurate simulation of QCD radiation. This generator focuses on providing a full description of the showering, hadronization and decay processes that take place after the hard scatter.

Pythia [67] is another generator with a focus on a precise modelling of QCD and QED radiation and showering processes, where approximations used are most accurate when describing small angle radiation. Additionally, Pythia uses a model for hard and soft processes which allows for the simulation of underlying activity. Together with Herwig, it establishes the benchmarks for describing LHC final states, usually interfaced with event generators that allow for higher order matrix element generation such as Madgraph or Powheg.

Madgraph [72] is an event generator used to compute hard scattering processes at LO and NLO in QCD perturbation theory. The introduction of one loop corrections, as well as the need to remove double-counting with the parton shower emissions, causes the appearance of MC events which can be positive or negatively weighted, although final distributions will always be positive. The output from this tool are ASCII files containing the 4-vectors of the particles involved in the hard scatter, which can then be processed by other tools such as Herwig or Pythia.

Powheg [73, 74], rather than an event generator, is a prescription used to interface NLO calculations with parton shower generators. However, the name usually refers to a framework that implements this methodology, PowhegBox [75]. This framework is, similarly to Madgraph, an event generator capable of computing NLO processes which can be later on interfaced with showering and hadronization algorithms. The main difference between both generators resides in the event weights they produce, stemming from the way they treat sub-leading terms in the calculations. Unlike Madgraph, Powheg (POsitive Weight Hardest Emission Generator) tends to veto those events corresponding to emissions that are not allowed, instead of removing the overlap between matrix element generation and showering by assigning negative weights to the events.

Sherpa [63,76] is a generator which unifies all components needed for the description of scattering events taking place at LHC, from matrix element calculation to the subsequent steps. This framework focuses on the best description of final states with large number of isolated jets thanks to the implementation of the MEPS@NLO [77–79] prescription, allowing for matching and merging multiple NLO matrix elements to showering algorithms.

EvtGen [80] is a framework used for an accurate description of bottom and charm

hadrons, allowing for interfacing with other parton showers such as PYTHIA or HERWIG.

3.3 Fixed-order calculations

This Section introduces the tools used for accurate theoretical prediction of the Drell-Yan production cross-section, which is of special interest for this thesis.

FEWZ [81–83] (Fully Exclusive W and Z production) is a framework that allows for the calculation of the fully-differential production cross section of the charged- and neutral-current Drell-Yan processes. It can include corrections at NNLO QCD perturbation theory and, in the case of neutral-current DY, it also allows for NLO electroweak corrections, where calculations can be performed with fiducial kinematic requirements. FEWZ is used to produce the theory predictions used in Part III.

VRAP [84] is another simulation code providing predictions for double-differential cross sections of DY processes at NNLO QCD as a function of dilepton invariant mass and rapidity, with the drawback of not having NLO electroweak corrections available. VRAP was used to calculate the correction *k-factors*, used to correct to NNLO QCD the Monte-Carlo simulations, initially calculated at NLO, used in Part III.

Fixed order predictions exist at N3LO for the neutral-current Drell-Yan production [85], showing a $\sim 2.5\%$ difference with respect to NNLO predictions at $m_{\ell\ell} = 150$ GeV and $< 1\%$ differences at higher dilepton invariant masses. However, a consistent application of such predictions would require N3LO PDF sets. Such sets are in early stages of development [86], limiting the predictions at NNLO accuracy.

Part II

Experimental setup

Chapter 4

The Large Hadron Collider

The Large Hadron Collider (LHC) [87] is a hadron synchrotron located at the European Organisation for Nuclear Research (CERN) in Geneva, Switzerland. It was designed to be the leading particle accelerator in the world, producing proton-proton collisions with a center of mass energy ($\sqrt{s}$) of up to 14 TeV to probe regions of the kinematic phase space previously inaccessible, thus enabling discoveries of new physics phenomena. The accelerator operates mainly in two regimes, depending on what particles are used in the collisions: proton-proton (pp) collisions and lead-lead (Pb-Pb) collisions, as well as a combination of both types of particles. Lead ion collisions are used to study the deposition of large energies in what, in terms of QCD scales, are considered as large spatial regions, providing a unique environment to analyze emerging QCD properties such as the production of *quark-gluon plasma* [88]. The results presented in this thesis will exclusively focus on the analysis of proton-proton collision data.

Section 4.1 provides an overview about particle acceleration at CERN, explaining how the LHC center of mass energy is achieved. The main characteristics of the collisions taking place at the facility are presented in Section 4.2 and a brief summary of the experiments studying LHC collision is discussed in Section 4.3.

4.1 Accelerator complex

The world-leading energies achieved at the LHC are the result of an acceleration process that takes place in a chain of pre-accelerators, as well as the LHC itself. Some of the pre-accelerators previously served for beam delivery to older experiments. A schematic view of the acceleration process is shown in Figure 4.1.

The proton beam preparation begins by accelerating negative hydrogen ions (H^-, hydrogen atoms with an additional electron) in the linear accelerator 4 (Linac4). These are fed to the Proton Synchrotron Booster (PSB), in a process where the atoms are ionized,

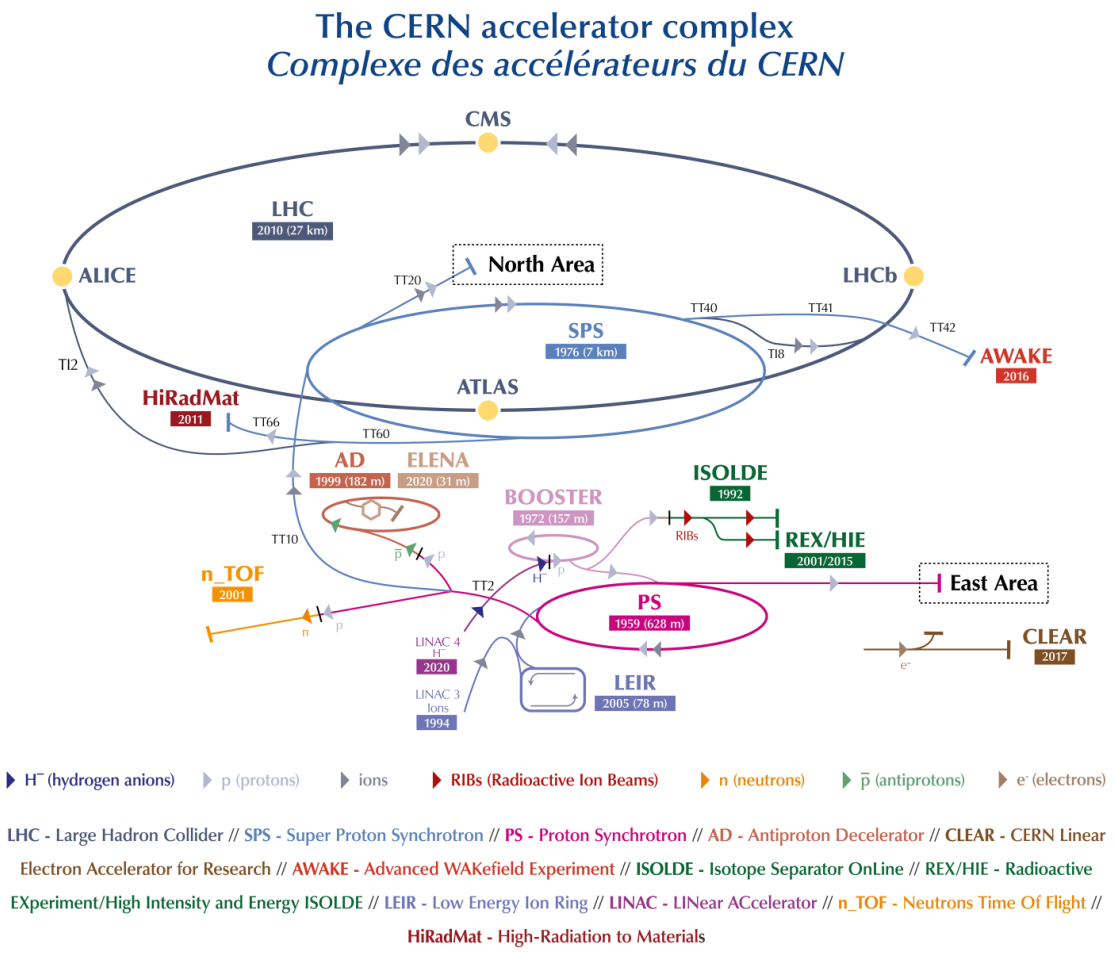


Figure 4.1: Schematic view of the CERN accelerator complex, showing the different accelerators used for proton beam production, as well as the location of the various experiments in the facility. From Ref. [89].

stripping the two electrons and leaving a beam of protons. The Booster accelerates the protons to 2 GeV and injects them into the next accelerator, the Proton Synchrotron (PS), which ramps up the energy and splits the beams into packets known as *bunches*, which are fed into the Super Proton Synchrotron (SPS). This accelerator brings the beam bunches to energies of up to 450 GeV and transfers them into the two main pipes within the LHC ring, one circulating clockwise and the other anti-clockwise. Protons in the bunches are guided through the 27 km circumference by 1232 super-conducting niobium-titanium (NbTi) dipole magnets, and are accelerated up to 6.5 TeV using 8 radio-frequency cavities per ring. Quadrupole magnets are used to focus the bunches and keep the protons within them as confined as possible before injecting the bunches at the collision points inside the main detectors (see Section 4.3) with a total center of mass energy of 13 TeV in Run 2.

4.2 LHC collision conditions

The main two conditions that define the quality of the physics produced by a particle collider are the beam energy it can reach and the amount of produced collisions. Up to 2022, the LHC accelerated proton beams up to 6.5 TeV which set a world-high center of mass energy for proton-proton collisions. The number of collisions is measured by the instantaneous luminosity, $\mathcal{L}$, which is directly related to the event rate (R) of a certain physics process depending on its cross section, $R = \mathcal{L} \cdot \sigma$. The instantaneous luminosity delivered in a collider is parametrized by a series of values which can be accurately measured in the accelerator:

$$\mathcal{L} = \frac{N_p^2 n_b f_r \gamma}{4\pi \epsilon_n \beta^*} \mathcal{F}, \tag{4.1}$$

where N_p is the number of protons per bunch, n_b is the number of bunches per beam, f_r is the revolution frequency of the synchrotron (~ 11.2 kHz [90] in the case of the LHC), γ is the relativistic factor, ϵ_n is the transverse beam emittance and β^* is the beta function, which describes the beam profiles upon colliding. The factor $\mathcal{F}$ is a geometric luminosity reduction which takes into account the angle in which bunches cross when interacting at the crossing point.

The total number of events for a certain process X will be given by the total *integrated* luminosity, $\mathcal{L}_{\text{int}}$, recorded over the total time of data taking:

$$N_{pp\to X} = \sigma_{pp\to X} \cdot \mathcal{L}_{\text{int}} = \sigma_{pp\to X} \cdot \int \mathcal{L} \cdot \mathrm{dt}. \tag{4.2}$$

Considering that the total number of events of a given process is a unit-less quantity, the integrated luminosity is measured in units of inverse cross section, typically fb^{-1} in Run 2 of LHC. The methodology used to measure the luminosity recorded at the ATLAS experiment is further explained in Section 5.6.

Another key factor at high luminosity values at the LHC is the number of simultaneous collisions taking place in each bunch crossing, also known as *pile-up*. It is determined by the center of mass energy of the beam that the total pp inelastic cross-section depends on, together with other beam characteristics such as size, profile or number of protons per bunch and time bunch spacing, which may increase the likelihood of multiple proton collisions to occur.

Table 4.1 summarizes some of the key beam parameters recorded during Run 2 operation of the LHC, compared to the design values of such. A full detailed breakdown of the LHC operation during Run 2, including all alterations to beam conditions across the different years of data-taking, can be found in Ref. [90].

Parameter	Design	2015	2016	2017	2018
Energy [TeV]	7	6.5	6.5	6.5	6.5
Bunches per ring	2808	2244	2556*	1916*	2556*
Bunch spacing [ns]	25	25	25	25	25
Bunch population N_b [10^{11} p/b]	1.15	1.15	1.2	1.25*	1.1
Transverse emmittance ϵ_n [mm·mrad]	3.75	3.5	2.1*	2.3*	2
β-function β^* [m]	0.55	0.8	0.4	0.4*	0.3*
Half crossing angle [μrad]	142.5	145	150*	150*	160*
Peak instantaneous luminosity [10^{34} cm^{-2}s^{-1}]	1	0.55	1.74*	1.9*	2.1
Maximum pile-up μ	∼20	15	45	80	60
Number days of physics operation	n.a.	88	146	140	145
Integrated luminosity delivered [fb^{-1}]	n.a.	4.2	39.7	50.6	66
Integrated luminosity recorded by ATLAS [fb^{-1}]	n.a.	3.2	33.0	44.3	58.4

Table 4.1: Overview of the LHC beam conditions for Run 2, from Ref. [90]. Variables labelled with a star (*) changed across data-taking over the year, see Ref. [90] for more details. Further details on the ATLAS luminosity measurements can be found in Section 5.6.

4.3 LHC experiments

The LHC creates an exciting environment for particle physics analyses, with many experiments investigating different processes taking place at the collider. Out of these, four experiments stand out, each of them located at one of the main underground interaction points (IP) of the LHC where bunch collisions actually take place. ATLAS (full overview provided in Chapter 5) and CMS (Compact Muon Solenoid) [91] are general purpose detectors, aiming to study the largest range of physics possible, with each of them independently working and serving as a confirmation of each other's possible discoveries. The ALICE [92] and LHCb [93] collaborations have detectors specialised in studying specific phenomena of the collisions taking place at the LHC, with the former focusing on the analysis of heavy ion collisions and the latter focusing on physics involving heavy quarks in the forward region.

However, the physics programme at the LHC goes beyond the main experiments aforementioned. The TOTEM experiment [94] consists of a series of detectors spread across half a kilometre around the CMS interaction point, measuring protons staying intact and scattered at small angles in the forward region of the CMS detector. The LHC-forward (LHCf) experiment [95] is divided in two parts, each sitting at either side of the ATLAS

collision point. Similarly to TOTEM, LHCf studies neutral particles produced at nearly zero degrees to the proton beam direction, in the case of LHCf escaping the ATLAS experiment instead of CMS. This offers a way to access the energy region of cosmic ray interactions under a controlled laboratory environment. The MoEDAL detector [96] is an extension to the LHCb detector and focuses on the search for magnetic monopoles as well as highly ionizing Stable Massive Particles (SMPs). The FASER experiment [97] is the most recent one in the LHC physics programme and it focuses on the detection of light and weakly interacting particles produced in the collisions taking place at ATLAS.

Chapter 5

The ATLAS experiment

The ATLAS (A Toroidal LHC ApparatuS) experiment [98] is one of the four main experiments at the facility. The experiment consists of a series of sub-detector systems arranged in symmetrical cylindrical layers around the beam axis. Figure 5.1 shows a schematic view of the ATLAS detector.

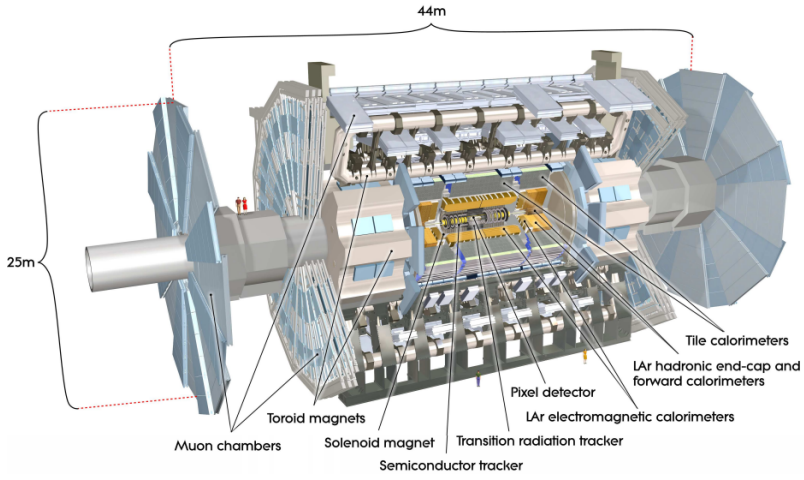


Figure 5.1: Cut-away view of the sub-detectors in the ATLAS experiment. See text for further details. From Ref. [98].

The ATLAS coordinate system as well as a description of some commonly used variables are presented in Section 5.1. Sections 5.2 to 5.4 describe each of the sub-detectors found in the experiment. The trigger system that rules ATLAS data-taking is presented in Section 5.5 and Section 5.6 describes the luminosity monitoring systems the experiment uses to analyze the amount of data recorded by the detector, as well as the methodology followed to measure this quantity.

5.1 ATLAS coordinate system

ATLAS uses a right-handed coordinate system where the z-axis is oriented along the beam pipe, the positive x direction pointing towards the center of the LHC and the positive y-direction pointing to the surface. The origin of all axes is set at the nominal interaction point at the center of the detector.

The spherical coordinates in the experiment are defined with the angles ϕ and θ, where ϕ is the azimuthal angle around the beam pipe and θ is the polar angle between a particle's three-momentum, $\overrightarrow{p}$, and the positive direction of the z-axis, as shown in Figure 5.2. The polar angle, ranging $\theta \in [0, \pi]$, is commonly transformed into the pseudorapidity, η:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right), \tag{5.1}$$

which is zero when θ is perpendicular to the beam-axis. The angular separation between objects in ATLAS, ΔR, is usually expressed in the $\phi - \eta$-plane as:

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}. \tag{5.2}$$

The transverse component of any vector, i.e. $\overrightarrow{p_{\mathrm{T}}}$, is defined as its projection in the x-y-plane, transverse to the beam axis. Given the high boost along the z-axis that most LHC collison products have, the energies and momenta of final states are often more conveniently expressed in terms of their transverse component and their pseudorapidity.

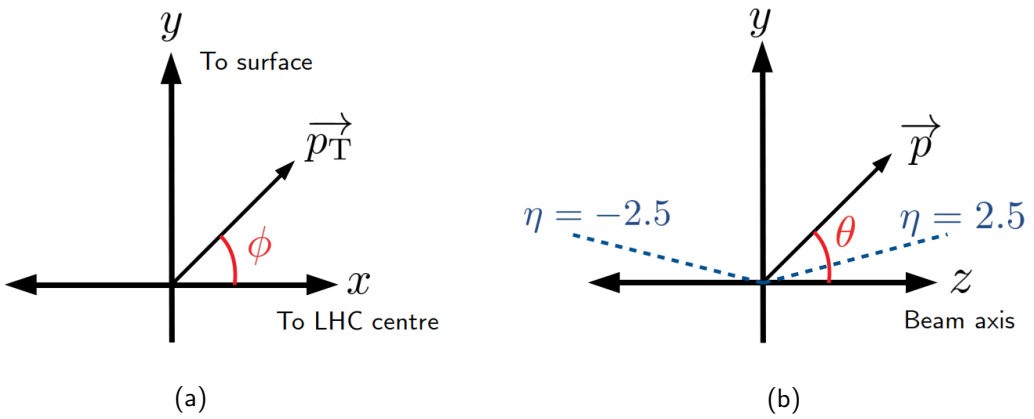


Figure 5.2: Diagram of the ATLAS coordinate system showing (a) the x-y-plane and (b) the y-z-plane. The illustration includes the definitions of the azimuthal and polar angles, ϕ and θ, as well as some sample values of the pseudorapidity η.

5.2 Tracking system

The ATLAS detection of charged particle tracks occurs at the Inner Detector [99], the subsystem closest to the interaction point consisting of three separate sub-detectors: the Pixel detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). The subsystem is immersed in a 2 T magnetic field, generated by a superconducting solenoid magnet that bends the particles trajectories, allowing to determine their momentum by measuring the radius of curvature of the trajectory, ρ. The transverse momentum (p_{T}) of a particle with electric charge q travelling through a magnetic field B can be calculated as:

$$p_{\mathrm{T}} = B \cdot q \cdot \rho. \tag{5.3}$$

A precise measurement of the particle's trajectory and its bending radius results in a better momentum resolution. Figure 5.3 shows the schematics of the ATLAS Inner Detector.

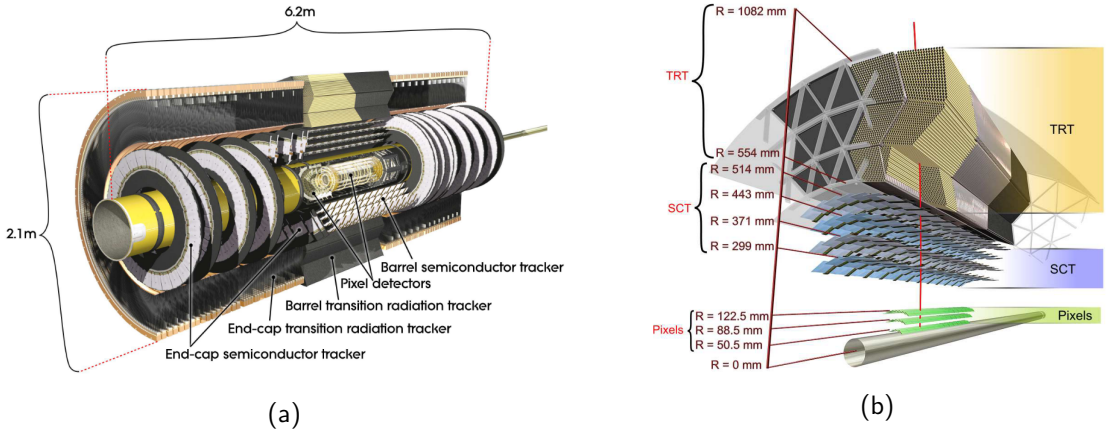


Figure 5.3: Illustration of the ATLAS Inner Detector. (a) Cut-away (longitudinal) view of the subsystem. (b) Close-up (cross-sectional) view of the three sub-detectors used for particle tracking in the barrel region. From Ref. [98].

The Pixel and SCT detectors are silicon-based, meaning that tracking is performed by measuring the charge deposited in the depletion zone created by the junction of a p-doped and n-doped silicon. Such junction creates a zone free of holes and free electrons, expanded by the application of a bias voltage. Upon a charged particle passing through the depletion zone, electron-hole pairs are created, drifting in different directions due to the bias voltage applied on the sensor. This drift then creates a measurable pulse on the electrodes of the detector with which we can locate the passing of a particle.

The Pixel detector is the closest detector to the proton beam. It is made of 1744 mod-

ules, 1456 located in three barrel layers and 288 in three disks on each end, as displayed in Figure 5.3a. Each module contains 46080 pixels, each of them with a typical size of 50×400 μm^2 and thickness of 250 μm. In Run 2 an additional system was integrated into the sub-detector, the Insertable B-Layer (IBL) [100]. The IBL consisted of an additional 224 modules placed between the innermost layer of the barrel and the beampipe, made of pixels of a reduced pitch size of 50×250 μm^2 and a thickness of 200 μm. The increased spatial resolution helps the tracking at increased particle fluences and improves the quality of the track measurements.

The SCT is the sub-detector surrounding the Pixel detector, providing further measurements that improve the precision of the track's momentum and path determination. It comprises 4088 modules arranged in four cylindrical layers in the barrel and two endcaps, each containing nine disks. The modules consist of pairs of strip silicon sensors, measuring along the $R - \phi$ direction in the barrel modules or radially in the end-cap modules. The mean pitch of the strips in all modules is approximately 80 μm. Together with the Pixel detector, they cover the region $|\eta| < 2.5$ with a precision of up to 10 μm ($R - \phi$ plane) and 115 μm (z-axis) in the barrel region, and 10 μm ($R-\phi$) and 115 μm (R) in the end-caps.

In the outermost region of the Inner Detector lies the TRT, a straw tube tracker consisting of 4 mm drift tubes. The inner layer of each tube is made of transition radiation material and gold-plated tungsten wire runs through the core of the tube, acting as an anode. The tubes are filled with Xenon- or Argon-based gas mixtures, which is ionised by the transition radiation emitted when a charged particle crosses the dielectric boundaries of the straw. Since the probability of emission of transition radiation depends on the particle's Lorentz factor $\gamma = E/m$, this sub-detector provides additional information to distinguish electrons from pions.

Tracks detected in the ATLAS Inner Detector are characterized by a set of parameters that are also used as criteria for particle identification, as well as information on the track's quality, helping to distinguish tracks originating from cosmic rays or secondary vertices. These parameters are known as *perigee* parameters and include:

- The charge/momentum ratio of the particle, q/p.
- The z-coordinate of the track at the point of closest approach to the beam-axis, z_0.
- The *impact parameter*, d_0, the distance of closest approach to the beam-axis in the x-y-plane. This measurement is usually used to filter low-quality or non-prompt tracks, using the uncertainty on the impact parameter, $\sigma(d_0)$, to construct the d_0 significance: $d_0^{\text{sig}} = |d_0|/\sigma(d_0)$.

- The polar angle between the particle's momentum and the z-axis in the x-z-plane, θ_0.
- The azimuthal angle between the particle's p_{T} and the x-axis, ϕ_0.

5.3 Calorimeter systems

The ATLAS calorimeter system [101–103] measures the energy of any interacting particles by absorbing them. A schematic view of this subsystem is shown in Figure 5.4. The calorimeters cover a region of $|\eta| < 4.9$ for the full ϕ range $[0, 2\pi]$. The depth of the system is optimised to best contain all showers, covering 20-30 electromagnetic radiation lengths (X_0), defined as the average distance a particle has to travel through the material to reduce its energy by a factor $1/e$, and 11 λ hadronic interaction lengths. The principle of detection is the alternation of layers of high-density material, which will prompt particle showers upon interaction with the material, and layers of *active* material, which measures the deposited energy from said showers. Calorimetry at ATLAS is divided mainly into two subsystems, depending on the particle type each system is optimized to detect: the Electromagnetic (EM) Calorimeter and the Hadronic Calorimeter.

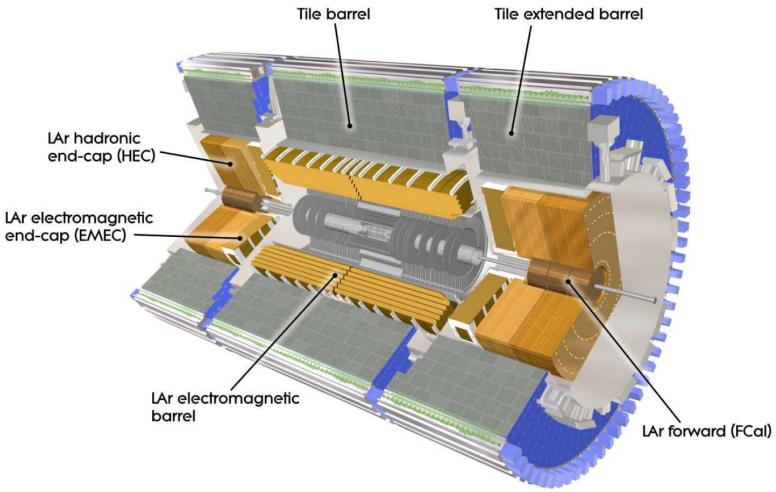


Figure 5.4: Schematic layout of the ATLAS calorimeter system. See text for further details. From Ref. [98].

The EM calorimeter (ECAL) uses Liquid Argon (LAr) for its active layers and is divided into different parts: one covering the range $|\eta| < 1.475$, the *barrel* region and two *end-cap* components which cover the $1.375 < |\eta| < 3.2$ region. Over the η region matching the inner detector, the EM calorimeter presents a finer granularity aimed at precision measurements of electrons and photons. Moreover, this region of the calorimeter is divided into three sections in depth, while the rest of the regions only include two sections in depth.

Upon interacting with the ECAL, incoming electrons and photons will create electromagnetic showers via bremsstrahlung (photon radiation from electrons) and pair production (photon conversion into e^+e^- pairs). These showers then ionise the liquid argon layers such that the energy of the original particles can be reconstructed.

The hadronic calorimeters use a mix of liquid argon and iron and scintillator tile technologies. The system itself is divided into three different parts: the tile calorimeter (HCAL), the LAr hadronic endcap calorimeter (HEC) and the LAr forward calorimeter (FCal). The tile calorimeter is located directly outside the EM calorimeter, in the region $|\eta| < 1.7$, further divided into a central barrel and two extended barrel, as seen in Figure 5.4. Charged shower particles in the scintillating tiles release photons that are guided through optic fibres and are read by photomultiplier tubes. Both the HEC and the FCal use LAr as their detection technology, working in a similar manner to the ECAL. They cover the $1.5 < |\eta| < 3$ and $3.1 < |\eta| < 4.9$ regions respectively.

The energy resolution, $\sigma(E)$, of the calorimeter systems is given by the formula:

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E\ [\text{GeV}]}} \oplus b, \tag{5.4}$$

where a is a stochastic factor that parametrizes the shower evolution and the constant term b quantifies the calorimeter response. The energy resolution of each of the different subsystems can be found in Table 5.1.

Calorimeter	Energy resolution
ECAL	$\frac{10.1}{\sqrt{E}} \oplus 0.2$ %
HCAL Barrel	$\frac{52.0}{\sqrt{E}} \oplus 3.0$ %
HCAL Endcap	$\frac{94.2}{\sqrt{E}} \oplus 7.5$ %

Table 5.1: Energy resolution of the different ATLAS calorimeter systems. HCAL Barrel refers to the combined performance of the tile and LAr calorimeters in the region. From [98].

5.4 Muon system

While the ATLAS calorimeters capture most particles produced in proton-proton collisions, muons are the only interacting particles that pass, leaving only a small amount of energy due to ionisation. The ATLAS Muon Spectrometer (MS) [104] is the sub-detector responsible for the detection of the muons and precisely measure their momenta. An overview of the system can be found in Figure 5.5. It is based on four different detection technologies: Monitored Drift Tubes (MDTs), Cathode Strip Chambers (CSCs), Resistive Plate Cham-

bers (RPCs) and Thin Gap Chambers (TGCs).

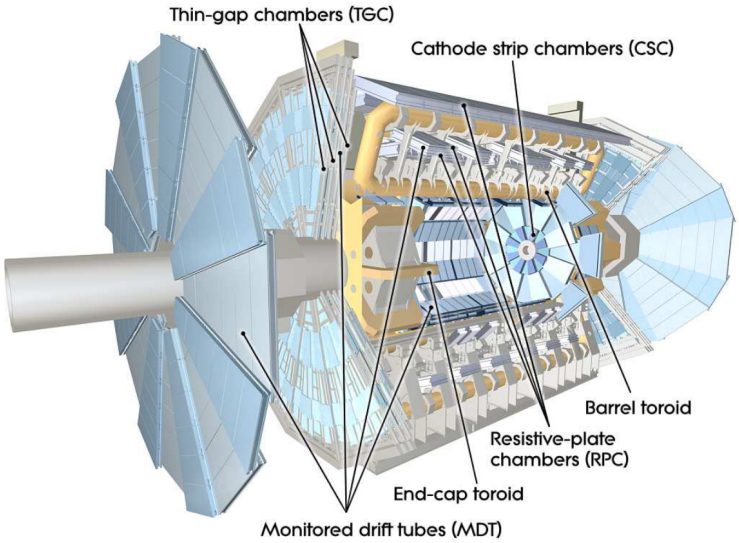


Figure 5.5: Schematic layout of the ATLAS muon detection system. See text for further details. From Ref. [98].

The reconstruction is based on the detection of muon tracks deflected in the toroid magnets allowing to measure their momentum independently of the inner detector track reconstruction. In the barrel region, these tracks are measured in chambers arranged in three cylindrical layers around the beam axis, while the transition and end-cap regions host their chambers in three planes perpendicular to the beam. MDTs and CSCs provide the precision tracking measurements with a combined resolution of 35 μm in the z-axis, 40 μm in R and 5 mm in ϕ, covering a combined range of $|\eta| < 2.7$. RPCs and TGCs are optimised for time resolution over spatial resolution, to provide fast inputs to the trigger system. RPCs measure the ϕ and z components of the tracks with a spatial resolution of 1 cm and a temporal resolution of 1 ns, while TGCs provide a spatial resolution of 2-3 mm in R and 3-7 mm in ϕ with a temporal resolution of 4 ns.

The magnetic field that bends the muon tracks is provided by the toroidal magnet system, composed of three subsystems of eight coils which turns are assembled radially and symmetrically around the beam-axis. The barrel toroid provides a 0.5 T magnetic field in the range $|\eta| < 1.4$, while the end-cap toroids provide 1 T in the $1.6 < |\eta| < 2.7$. The bending power is lower in the transition region between both magnets ($1.4 < |\eta| < 1.6$), where the fringe field of one magnet cancels the bending power of the other.

5.5 Trigger and data acquisition system

As presented in Section 4.2, collisions at LHC took place every 25 ns during Run 2, at a rate of 40 million collisions per second, which makes the storage of every event not possible considering the bandwidth and storage limitations. Moreover, most proton-proton collisions only result into low p_{T} final states with a rate many orders of magnitude higher than the rarer high p_{T} processes of greater interest. The ATLAS Trigger and Data Acquisition system [105, 106] is therefore optimised to identify events containing high p_{T} objects typically selected in physics analyses, while maintaining a sustainable data acquisition rate that abides by bandwidth constraints. The trigger system has different levels, each of them refining the decisions made at the previous level and, if necessary, applying additional selection criteria. An overview of the trigger system used in Run 2 at ATLAS can be found in Figure 5.6.

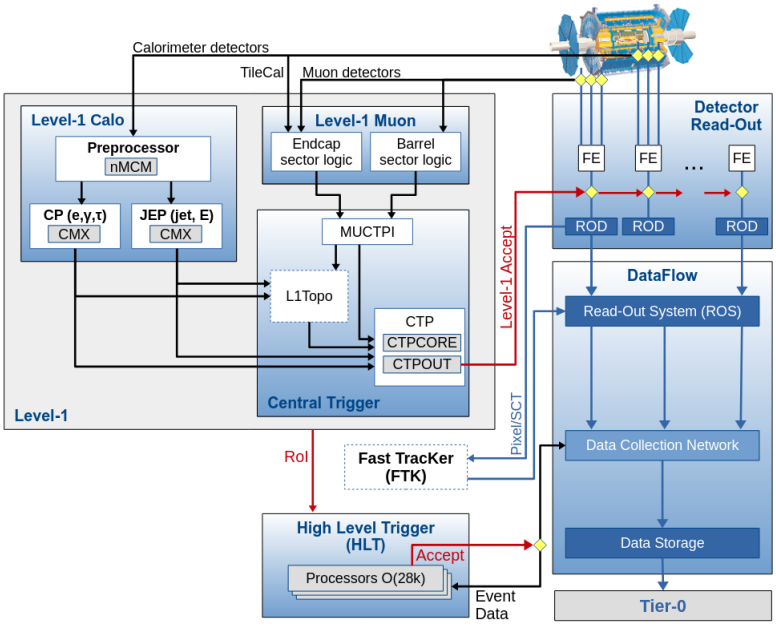


Figure 5.6: Schematic overview of the ATLAS Trigger and Data Acquisition system in Run 2. See text for further details. From Ref. [105].

The first level of the trigger system consists of a hardware-based trigger (Level-1 or L1), using custom electronics to analyse coarse data of the calorimeter (L1-Calo) and muon (L1-Muon) systems. The information provided by both systems is used to construct Regions of Interest (ROIs) in the $\eta - \phi$ plane containing potential physics objects. The calorimeter data is used to identify leptons, photons, jets or transverse missing energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, while the MS provides information on muon candidates. The decisions made by the L1 trigger take place less than 2.5 μs and lowers the event recording rate down to ∼100 kHz.

Once an event is accepted by the L1 system, a full detector readout is performed and the data is temporarily buffered before being passed to the next trigger system, the High-Level Trigger (HLT). This software-based level combines the coarse ROIs provided by the L1 system with the readout containing the information recorded by ATLAS in full granularity to make an improved event selection. This improved selection determines whether the event is stored, at an even further reduced recording rate of 1.2 kHz, after which the data readouts are sent to CERN's storage facilities and processed to allow for offline analysis.

5.6 Luminosity determination

The determination of the number of events of a given process that take place at ATLAS requires for a precise measurement of the integrated luminosity recorded by the experiment. The number of events of a certain process, $N_{pp\to X}$, can be expressed as:

$$N_{pp\to X} = \mathcal{L}_{\text{int}} \cdot \sigma_{pp\to X}, \tag{5.5}$$

where $\sigma_{pp\to X}$ is the process cross-section (see Section 3.1 for more details) and $\mathcal{L}_{\text{int}}$ is the integrated luminosity, the time integral of the instantaneous luminosity registered at the experiment. A precise measurement of this quantity directly translates to the precise measurement of production cross-sections at the experiment and is of great importance for the normalisation of background processes in searches for new physics.

Instantaneous luminosity is usually expressed at the LHC in units of $\text{cm}^{-2}\text{s}^{-1}$, while units for integrated luminosity remove the time dimension. However, the latter is commonly expressed instead in inverse barn, b^{-1}, with 1 b=10^{28} m^2. Thus, it is common for integrated luminosities at the LHC to be discussed in units of pb^{-1} and fb^{-1}.

5.6.1 Luminosity from collider parameters

The instantaneous luminosity delivered to ATLAS upon two bunches colliding, $\mathcal{L}_b$, is related to the collision parameters by

$$\mathcal{L}_b = \frac{f_r n_1 n_2}{2\pi \sum_x \sum_y}, \tag{5.6}$$

where f_r is the LHC revolution frequency ($\sim$ 11.2 kHz [90]) and n_1 and n_2 are the number of protons in each of the bunches. The width and length of the bunches in the $x-y$ plane is given by the parameters $\sum_x$ and $\sum_y$. The value of $\mathcal{L}_b$ can also be expressed as:

$$\mathcal{L}_b = \frac{f_r \cdot \mu_b}{\sigma_{\text{inel.}}}, \tag{5.7}$$

where μ_b is the mean number of inelastic proton collisions per bunch crossing and $\sigma_{\text{inel.}}$ is the pp inelastic cross-section. When summed over all colliding bunches, one retrieves the expression for the total instantaneous luminosity, $\mathcal{L}$:

$$\mathcal{L} = \sum_{b=1}^{n_b} \mathcal{L}_b = \frac{n_b f_r \mu}{\sigma_{\text{inel.}}}, \tag{5.8}$$

where $\mu = \langle \mu_b \rangle$ is the mean number of interactions per bunch crossing, averaged across all bunch pairs. Since the measurement of $\sigma_{\text{inel.}}$ has a relative uncertainty up to 3.7% [107] and μ cannot be directly measured, it is useful to re-define the value of μ to be proportional to an arbitrary quantity μ_{vis} that represents the mean rate of production of some observable quantity. The two will relate as,

$$\mu = \frac{\sigma_{\text{inel.}}}{\sigma_{\text{vis}}} \cdot \mu_{\text{vis}}, \tag{5.9}$$

such that Equation 5.8 can be rewritten as:

$$\mathcal{L} = \frac{n_b f_r \mu_{\text{vis}}}{\sigma_{\text{vis}}}. \tag{5.10}$$

The quantity σ_{vis} represents the cross-section of the production of the observable of choice, which will generally be unknown and will require calibration to obtain its value.

5.6.2 ATLAS baseline luminosity measurement

The baseline luminosity measurement at ATLAS is performed using the Cherenkov Integrating Detector (LUCID) [108]. LUCID consists of two detectors sitting close to the beam at $z = \pm 17$ m from the interaction point, covering $5.6 < |\eta| < 6$. Each of the detectors is comprised by 20 aluminium pipes filled with gas arranged around the beam line. Inelastic pp collisions produce forward particles which interact with the gas inside the pipes to produce Cherenkov radiation. The Cherenkov light is read by PMTs and will be directly proportional to the number of pp collisions taking place in ATLAS, representing the quantity μ_{vis} in Equation 5.10.

The calibration of σ_{vis} for LUCID [109] is obtained in special LHC runs every year using van der Meer (vdM) scans [110]. In these runs the beam conditions are changed with respect to nominal operation, displacing the beams through each other so that the overlap between both beams varies through the scan. This allows to accurately measure the beam's spatial spread ($\sum_x$ and $\sum_y$). Combined with additional measurements of the bunch populations (n_1 and n_2), σ_{vis} is calculated as

$$\sigma_{\text{vis}} = \mu_{\text{vis}}^{\text{max}} \frac{2\pi \sum_x \sum_y}{n_1 n_2}, \tag{5.11}$$

where $\mu_{\text{vis}}^{\text{max}}$ is the interaction rate when the overlap between both bunches is maximal. Since the vdM scans are performed under beam conditions that do not represent the nominal ATLAS data-taking conditions, additional corrections to the luminosity recorded by LUCID are obtained using other luminometers to transfer the calibration to standard beam conditions. These alternative luminometers include: track-counting, which uses the number of inner detector tracks to obtain μ_{vis}; TILE, measuring μ_{vis} from PMT signals in the scintillating layers of the hadronic calorimeter; and EMEC and FCal, which retrieve μ_{vis} from LAr gap currents in the calorimeter systems.

The total ATLAS luminosity recorded during Run 2, provided by correcting the measurements by LUCID for the aforementioned effects, was found to be 140 fb^{-1} [109] with a total uncertainty of 0.83%. A summary of some luminosity related parameters is provided in Table 5.2.

Data period	No. of ATLAS runs	Typical μ	$\mathcal{L}_{\text{int}}$ [fb^{-1}]	Total Unc. [%]
Run 2	582	33.7	140.1	0.83
2015	64	13.4	3.24	1.13
2016	144	25.1	33.4	0.89
2017	183	37.8	44.3	1.13
2018	191	36.1	58.5	1.10

Table 5.2: Summary of the ATLAS luminosity estimate results for pp collision data-taking at $\sqrt{s} = 13$ TeV during Run 2, from [109]. The grouping of the 2015 and 2016 data periods is shown as the corresponding uncertainty is calculated for both periods combined. The number of ATLAS runs containing at least one data-taking segment passing the quality criteria for physics analysis [111] is also presented.

Chapter 6

Particle reconstruction and calibration

Particles produced as a result of the pp collisions taking place at the LHC are observed as signals detected by the different subsystems described in Chapter 5. These signals must be processed to identify the physics objects which analyses seek to study. The reconstructed particles will consist of a four-vector comprising the particle's momentum and energy and certain criteria that allow for the correct identification of the physics object, distinguishing the genuine particles from those that leave similar signatures in our detector. The work described in this thesis focuses on the analysis of charged leptons, electrons and muons. This Chapter provides an overview of the methods employed to reconstruct and identify charged leptons at the ATLAS experiment.

6.1 Particle tracks and Primary Vertices

Charged particles are identified by reconstructing their tracks from the hits produced in the ID. As described in Section 5.2, tracks in ATLAS are characterized using five parameters: q/p, z_0, d_0, θ and ϕ. Tracks are built by joining space-points in the Pixel, SCT and TRT detectors, starting from seeds based on three unique space points with which an initial trajectory is estimated. If the trajectory is compatible with additional space points, it is extrapolated through the pixel and SCT detectors to form a track candidate. These are assigned a score based on quality criteria (such as the presence of clusters in detector layers, shared clusters with other tracks or quality of the extrapolation fit) and fed to a neural network to distinguish between tracks belonging to a same cluster [112]. Once a track has been identified in the silicon detectors it is then extended to matching clusters in the TRT and then refit for an improvement on the resolution of the track parameters. More details on the track reconstruction can be found in Ref. [113].

Once all tracks in an event have been reconstructed, one can identify common intersec-

tion points (vertices) which correspond to proton-proton interactions (primary) or decays of long-lived particles (secondary). The identification of primary vertices is performed using an iterative reconstruction algorithm [114], starting from the identification of an initial seed as the mode of the z_0 value of all tracks in the event. Using this seed as a starting estimate, the best vertex position is found by down-weighting tracks less compatible with the vertex and re-computing its position. Once the position of the vertex converges, the tracks that have not been used in its computation are used to identify additional vertices until no tracks are left in the event.

An event of interest is typically defined as that originating in the primary vertex in the event with the highest sum of squared transverse momentum of its associated tracks, $\sum_i^{\text{N track}} p_{\text{T,i}}^2$, containing a minimum of two tracks with $p_{\text{T}} > 500$ MeV. Any additional vertex passing the requirement of two tracks above the momentum threshold is considered to be an additional primary vertex, originating from simultaneous pp collisions (pileup). Events may also contain secondary vertices, displaced from primary vertices, meaning that they originated from the decays of long-lived particles.

6.2 Electrons

6.2.1 Electron reconstruction

Electrons are identified as a clustered energy deposit matched to a charged-particle inner detector track, as shown in Figure 6.1. Due to the coverage of the inner detector, electrons can only be reconstructed if found in the range $|\eta| < 2.47$[†]. Moreover, as described in Section 5.2, the calorimeter presents a transition region ($1.37 < |\eta| < 1.52$) between the barrel and end-cap modules commonly referred to as the "crack region", in which resolution worsens. Electrons found in this region are typically not used for precision physics analysis.

The matching between ECAL clusters and inner detector tracks presents a major issue, since ATLAS contains a lot of material between both subsystems which, when being traversed by an incoming electron, lead to the production of photons from bremsstrahlung. Since photons and electrons undergo similar processes in the EM calorimeter (photon radiation, pair-production), the showers produced by both present similar features and require careful matching to the ID tracks. Moreover, electron reconstruction algorithms must be designed to account for energy losses coming from the radiation production.

Electron reconstruction at ATLAS [115] starts with the identification of energy deposits (clusters) in the electromagnetic calorimeter. The $\eta \times \phi$ phase space of the calorimeter is

[†]So-called *forward* electrons in the range $|\eta| > 2.5$ can be identified by using information of the FCal exclusively, although with a higher misidentification rate. Such objects are not used in the work described in this thesis.

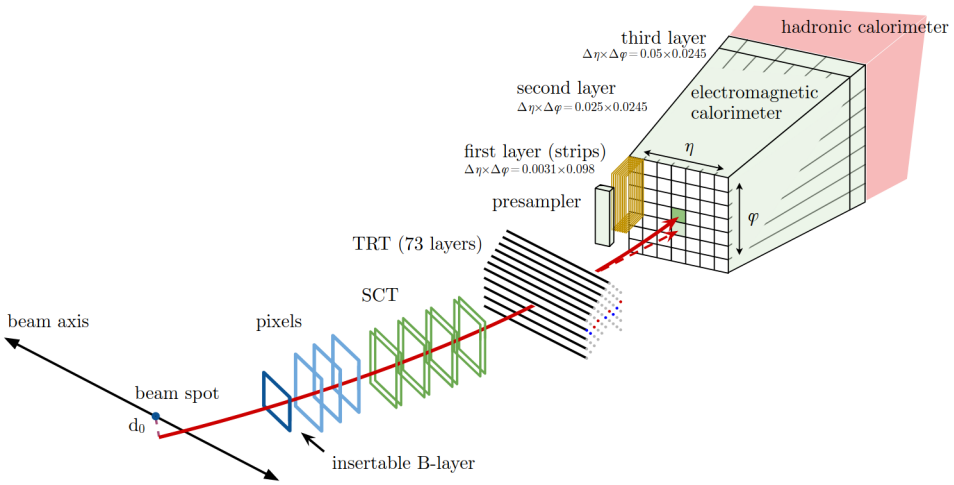


Figure 6.1: Schematic view of an electron's path through the ATLAS detector, shown in red. The dashed line represents the path of a photon produced by the interaction of the electron with the tracking system material. See text for further details. From Ref. [115].

divided into a grid of 200×256 segments (towers), each covering a size of $\Delta\eta\times\Delta\phi = 0.025\times0.025$, given by the granularity of the ECAL. Within each tower, the energy deposits are summed and fed into a sliding-window algorithm [116] that will locate cluster candidates. The algorithm identifies any 3×5 window in $\eta\times\phi$ in the ECAL exceeding a total transverse energy (E_{T}) of 2.5 GeV as a cluster candidate. If two cluster candidates are found in proximity, the candidate with the higher E_{T} is retained.

6.2.2 Electron identification

Once tracks and energy clusters have been matched, electrons are selected using a likelihood-based (LH) identification (ID). The electron likelihood is based on the products for signal (L_S, consisting of prompt electrons) and background (L_B, containing a combination of jets and non-prompt electrons produced in other particle's decays) of n probability density functions, P:

$$L_{S(B)}(\mathbf{x}) = \prod_{i=1}^{n} P_{S(B),i}(x_i), \tag{6.1}$$

where $\mathbf{x}$ is a vector containing different quality measurements derived from the tracks and energy deposits used to construct the electron candidate (see Ref. [115] for a full description). $P_{S,i}(x_i)$ is the value of the signal probability distribution for the quantity i at the value x_i and $P_{B,i}(x_i)$ represents the equivalent value for the background probability distribution. Using both likelihoods, a discriminant d_L is formed for each electron candidate:

$$d_L = \frac{L_S}{L_S + L_B}. \tag{6.2}$$

The distribution of this discriminant peaks at unity (zero) for signal (background) candidates. To simplify the selection of different operating points, a modified discriminant is defined by taking the inverse sigmoid transform of d_L:

$$d'_L = -\tau^{-1} \ln \left(d_L^{-1} - 1 \right), \tag{6.3}$$

where τ is a parameter optimised to 15.

Setting thresholds on the d'_L discriminant, three working points (WPs) are defined: Loose, Medium and Tight. These offer different compromises between identification efficiency and background electron contamination: looser selections have high efficiencies at the cost of higher contaminations, while tighter selections reduce the selection efficiency while offering a higher signal purity. The identification efficiency as a function of E_{T} and η, are shown in Figure 6.2. The efficiencies increase as a function of the electron's E_{T}, peaking at $\sim 93\%$ in the case of the Medium working point at high E_{T}. The agreement between the efficiencies measured in data and MC samples also improves as a function of the electron E_{T}, reaching an agreement better than 2% at high E_{T}. The agreement between data and MC is found to be stable across the η spectrum, with the exception of the crack region measurement, where up to 10% differences are observed.

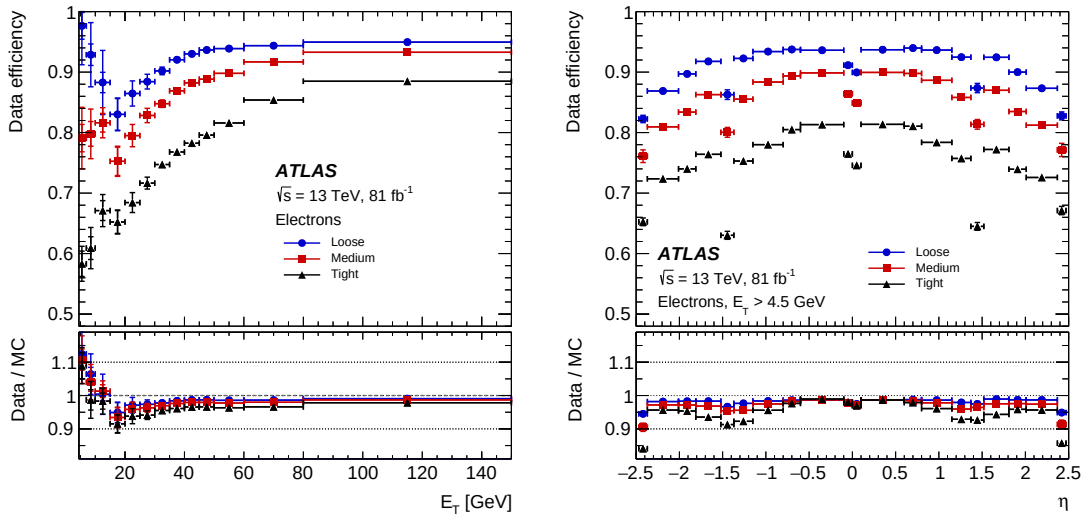


Figure 6.2: Measured LH-based electron ID efficiencies for the Loose (blue dots), Medium (red squares) and Tight (black traingles) working points as a function of transverse energy, E_{T} (left) and pseudorapidity, η (right). From Ref. [115].

6.2.3 Electron isolation

Isolation cuts are used to further improve the purity, distinguishing the production of prompt electrons in signal processes from backgorund processes such as semileptonic de-

cays of heavy quarks or misidentified hadrons. Variables are constructed, quantifying the amount of activity in the vicinity of the object in a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ around the direction of the electron candidate. These quantities can be based on the sum of the transverse energies of clusters in the calorimeter (ie. calorimeter-based isolation) or the sum of the transverse momenta of tracks (track-based isolation). Thus, one can define the quantity $p_{\mathrm{T}}^{\mathrm{coneXX}}$, as the sum of momenta within a cone of radius $\Delta R < XX$ around the electron candidate, excluding the momentum of the candidate itself (and similarly for E_{T} using calorimetry-based isolation).

In the case of track isolation, the high granularity of the inner detector allows for very narrow cone sizes. This is of particular interest in the study of boosted final states, where jets and other objects tend to be produced close to the signal lepton. With these events in mind, a variable cone-size track isolation is commonly used, where the cone radius used to define $p_{\mathrm{T}}^{\mathrm{coneXX}}$ progressively decreases in size as a function of the p_{T} of the electron candidate:

$$\Delta R = \min\left(\frac{10\ \mathrm{GeV}}{p_{\mathrm{T}}\ [\mathrm{GeV}]}, R_{\mathrm{max}}\right), \tag{6.4}$$

where R_{max} is the maximum cone size, typically 0.2 to 0.4.

6.2.4 Electron energy scale and resolution correction

The electron reconstruction performance in data is not perfectly modelled by our simulation. Discrepancies may emerge from degradation of detector components that is not taken into account in simulation or pileup mismodelling, which results in a mismatch of reconstruction efficiencies between data and simulation. A series of calibration techniques are introduced to improve the agreement between simulation and data, including lepton energy and momentum scale calibrations and corrections for identification and isolation efficiencies.

The correction of the energy response to electrons in ATLAS [117,118] combines a series of data- and simulation-based steps. The first step optimises the estimation of the energy of the electron from the "raw" energy deposits in the calorimeter with a multivariate model trained on simulation.

Further corrections improve the non-uniform response of the EM calorimeter. Firstly, adjusted relative energy scales of the different layers in the EM calorimeter are estimated by comparing the deposited energy in each layer using $Z \to \mu\mu$ events. Secondly, corrections for other non-uniform effects (geometric effects at boundaries between modules, corrections for the usage of non-nominal voltage settings...) are evaluated using the ratio

of the measured calorimeter energy to the estimated momentum of electrons and positrons coming from Z decays in the inner detector.

Once these corrections have been applied, the difference in energy scale between data and MC, α_i (where i corresponds to different regions of η), is extracted using the expression:

$$E_i^{\text{data,corr}} = \frac{E_i^{\text{data,raw}}}{1-\alpha_i} \simeq E_i^{\text{MC}}. \tag{6.5}$$

Similarly, the corrections to the energy resolution simulation are parametrised as a constant term c_i that will be added in quadrature to the original modelled resolution:

$$\left(\frac{\sigma_E}{E}\right)_i^{\text{MC,corr}} = \left(\frac{\sigma_E}{E}\right)_i^{\text{MC,raw}} \oplus c_i. \tag{6.6}$$

The values of α_i and c_i are retrieved from samples of $Z \to ee$ decays, optimising the agreement of the invariant mass of the dilepton system between the simulated and observed values in different η regions. The resulting α_i and c_i distributions can be found in Ref. [118], showing sub-percent level corrections in the precision regions of the detector.

Several systematic uncertainties impact the measurement of these corrections, such as pile-up, layer calibration uncertainties, scintillator calibration or uncertainties arising from electromagnetic shower mismodelling. The impact of these uncertainties on the measurement presented in this thesis is further explained in Section 8.5.

6.3 Muons

6.3.1 Muon reconstruction

Muons are generally reconstructed by matching tracks in the inner detector and the muon spectrometer. Tracks observed in the latter are a clear indicator of a muon, since all other particles (with the exception of the neutrino) are entirely absorbed by the calorimeters. Track reconstruction in the inner detector is performed as explained in Section 6.1. An independent track reconstruction is performed in the MS [119], built upon the construction of straight-line track segments within an individual MS station (see Figure 5.5). Different segments across multiple MS stations are combined accounting for the influence of the magnetic field, adding additional hits from trigger detectors to create three-dimensional track candidates. A global χ^2 fit is performed on the candidates taking into account interactions in the detector material and possible misalignment effects between stations. Once a trajectory has been obtained, outlier hits are removed from the track and unused hits along the trajectory are added. The fit is performed using this updated hit collection.

Different reconstruction strategies to combine inner detector and MS tracks are used in ATLAS [119]. All muon objects used in the physics analyses presented in this thesis consist of *combined* muons, which are the results of a combined fit using hits from MS and inner detector tracks. The energy loss of the muon in the calorimeters is accounted in this fitting procedure. To allow for a better acceptance, tracks found in the $|\eta| > 2.5$ MS region are allowed to be combined with short track segments in the pixel and SCT detectors, rather than requiring the reconstruction of a full inner detector track.

6.3.2 Muon identification

Once a muon track has been reconstructed, different selections on the track properties (number of hits, fit quality, variables indicating the compatibility between inner detector and MS tracks...) improve the selection of signal (prompt) muons over background muons, such as those originated from heavy-flavour hadron decays. Thus, different working points are defined, namely Loose, Medium, Tight and High-p_{T}. The muon selection criteria in the physics analysis presented in this thesis use the Medium and High-p_{T} working points. These WPs are optimised to provide an efficiency and purity suitable for most physics analysis, with an emphasis of minimising the systematic uncertainties associated to signal selection. The reconstruction and identification efficiency for the Medium WP is shown in Figure 6.3, reaching efficiencies above 99% at $p_{\mathrm{T}} > 20$ GeV. The measurement of the identification efficiency shows a very good agreement between data and MC, with and agreement of $< 0.5\%$ observed across all p_{T} and η spectra. Efficiencies and their corresponding data/MC agreement for the High-p_{T} WP can be found in Ref. [119].

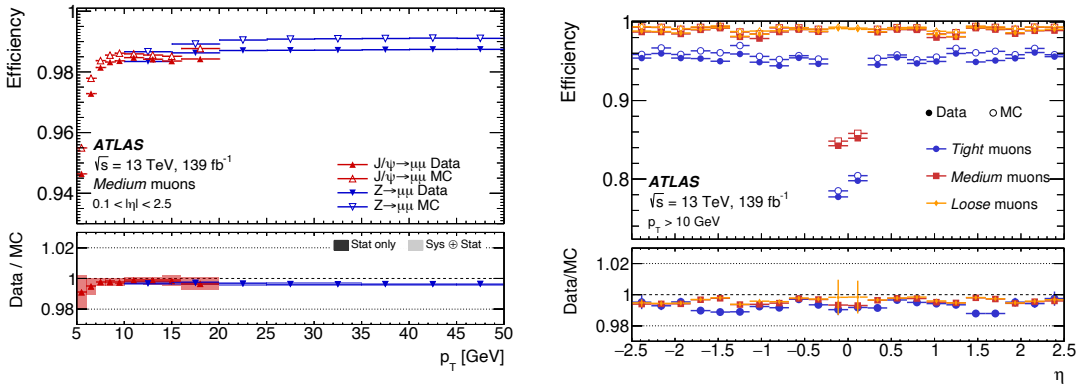


Figure 6.3: Left: muon reconstruction and identification efficiency for the Medium WP, shown as a function of the transverse momentum, p_{T}. Right: muon reconstruction as a function of the different available working points shown as a function of pseudorapidity, η; Loose (yellow), Medium (red), Tight (blue). From Ref. [119].

Muons passing the Medium WP are required to contain at least two so-called *precision*

stations, referring to those stations in the MS in which at least three hits have been detected. To suppress contamination from non-prompt muons, further cuts are placed on the q/p compatibility:

$$q/p \text{ compatibility} = \frac{|(q/p)_{\text{ID}} - (q/p)_{\text{MS}}|}{\sqrt{\sigma^2(q/p)_{\text{ID}} + \sigma^2(q/p)_{\text{MS}}}}, \tag{6.7}$$

where $(q/p)_{\text{ID}}$ and $(q/p)_{\text{MS}}$ are the corresponding values to the tracks in the inner detector and MS respectively and $\sigma^2(q/p)_{\text{ID}}$, $\sigma^2(q/p)_{\text{MS}}$ are their corresponding uncertainties. Muons passing the Medium WP criteria are required to have a q/p compatibility lower than 7.

Muons passing the high-p_{T} WP are required to pass the same cuts as placed on Medium candidates, and additionally contain at least three precision stations to overcome the challenge of reconstructing muons with almost straight trajectories. Moreover, a selection is placed on the relative q/p uncertainty, $\sigma_{\text{rel}}(q/p) = \sigma(q/p)/|q/p|$ to reject low-quality measurements. A p_{T}-dependent threshold is placed on $\sigma_{\text{rel}}(q/p)$, lowering the threshold for muons with lower momenta. Further details on the selection of this WP can be found in Ref. [119].

6.3.3 Muon isolation

The muons purity is further improved by applying isolation cuts in a very similar manner as defined for electrons. A variable $p_{\text{T}}^{\text{coneXX}}$ is defined, adding the momenta of all particles found within a cone of radius ΔR around the muon candidate. Moreover, a varying cut is placed as a function of the muon's p_{T}, as shown in Equation 6.4, exploiting the high granularity of the inner detector in the study of boosted final states. Typically, R_{max} also takes values of 0.2 or 0.4 in the case of muons depending on each analysis needs.

6.3.4 Muon calibration

The momentum scale and resolution of muons at ATLAS is studied using $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ decays [120]. The correction for muon momentum is performed independently in the ID and MS, using the p_{T} of combined muons for this purpose (see Section 6.3.1). A corrected transverse momentum $p_{\text{T}}^{\text{corr,Det}}$, where Det represents either ID or MS, is retrieved as:

$$p_{\text{T}}^{\text{corr,Det}} = \frac{p_{\text{T}}^{\text{MC,Det}} + s_0^{\text{Det}} + s_1^{\text{Det}} \cdot p_{\text{T}}^{\text{MC,Det}}}{1 + \Delta r_1^{\text{Det}} \cdot g_1 + \Delta r_2^{\text{Det}} \cdot p_{\text{T}}^{\text{MC,Det}} \cdot g_2}, \tag{6.8}$$

where $p_{\text{T}}^{\text{MC,Det}}$ is the uncorrected p_{T} as measured in simulation, g_i are normally distributed random numbers with zero mean and unit width and Δr_i^{Det} and s_i^{Det} are the momentum resolution and scale corrections applied. The latter are described as a function

of η and ϕ, allowing for a non-uniform correction throughout the detector.

Each term in the numerator on Eq. 6.8 describes a different correction to the muon momentum scale that needs to be accounted for: s_1^{Det} corrects for inaccuracies in the description of the muons trajectory along the magnetic field in the ID and MS while s_0^{MS} accounts for the energy loss in the calorimeter and material layers. Since this loss is minimal between the interaction point and the ID, s_0^{ID} is set to zero.

The denominator in Eq. 6.8 describes the momentum smearing, assuming that the relative p_{T} resolution can be parametrized as:

$$\frac{\sigma(p_{\text{T}})}{p_{\text{T}}} = \Delta r_1 \oplus \Delta r_2 \cdot p_{\text{T}}. \tag{6.9}$$

The first term accounts for inaccuracies in the muons' trajectory (scattering in the material, magnetic field inhomogeneities, displacement of the track hits) and the second factor describes the intrinsic resolution effects emerging from the spatial resolution of the hit measurements. A summary of the results of the measurements of the s_i and Δr_i parameters can be found in Ref. [120].

Part III

Measurement of the high-mass Drell-Yan cross section

Chapter 7

Motivation

As first presented in Section 3.1.1, a good understanding of the proton content is essential to obtain precise Standard Model predictions. Our knowledge of the parton distribution functions of the proton plays a key role in many predictions and PDF uncertainties are among the leading uncertainties for electroweak measurements such as WW [121], $Z\gamma$ [122] or $Z\gamma$+jets [123] production. Measuring the cross section of the neutral-current Drell-Yan process probing different invariant masses ($m_{\ell\ell}$) and dilepton rapidities ($y_{\ell\ell}$) allows to explore a wide range of x values, as shown in Equation 3.7 and in Figure 7.1. Values of up to $x \simeq 0.91$ can be achieved at high-$m_{\ell\ell}$ and high rapidity regions of the measurement, while low $m_{\ell\ell}$ measurements allow to probe down to $x \sim 10^{-3}$. Thus, measurements of this process in different regions of the kinematic phase space can provide crucial input for PDF fits, helping constrain regions in x that other experiments cannot probe. Moreover, as shown in Section 3.1.2, the cross section of dilepton production in pp collisions is also sensitive to the photon content of the proton. The measurement of the $\gamma\gamma$ initiated dilepton pair production, which has a significant contribution at high invariant masses, can help provide better constraints on the photon content of the proton.

The results of the most recent ATLAS measurement of the Drell-Yan process [124] cross-section at high dilepton invariant masses are shown in Figure 7.2. The measurement was performed single- (measured against $m_{\ell\ell}$) and double-differentially (measured against $m_{\ell\ell}$ and $y_{\ell\ell}$, as well as $m_{\ell\ell}$ and absolute pseudorapidity separation between leptons, $\Delta\eta_{\ell\ell}$) using ATLAS data collected at 8 TeV with an integrated luminosity of 20.1 fb^{-1}, in the range $116 < m_{\ell\ell} < 1500$ TeV. The measurement was later expanded with a triple-differential cross section measurement [57] against $m_{\ell\ell}$, $y_{\ell\ell}$ and $\cos\theta^*_{\text{CS}}$ in a reduced invariant mass range of $46 < m_{\ell\ell} < 200$ GeV.

The following chapters present the analysis of the Drell-Yan cross-section $pp \to Z/\gamma^* \to \ell\ell$, where ℓ represents the light leptons ($\ell = e, \mu$), in the range $m_{\ell\ell} > 116$ GeV. The analysis covers three cross-section measurements, one single-differential measurement as a function

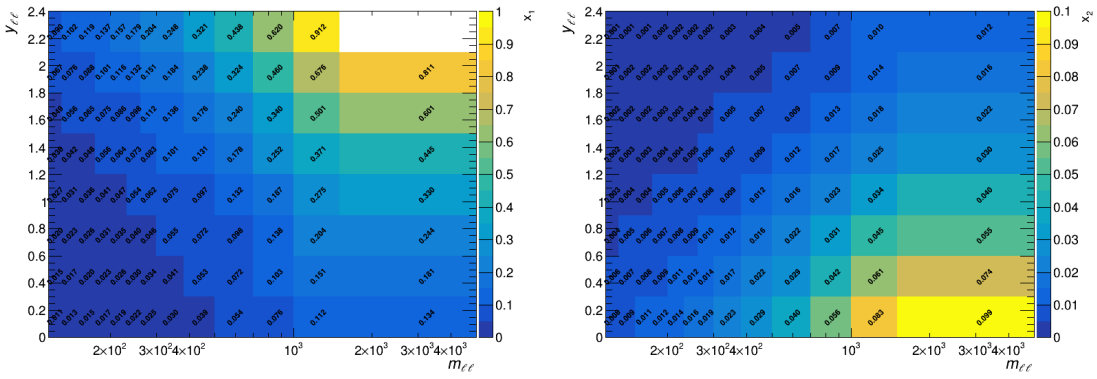


Figure 7.1: Proton momentum fraction x carried by each of the incoming partons, x_1 (left) and x_2 (right), producing a neutral-current Drell-Yan event at dilepton invariant masses beyond the Z-boson resonance at a center of mass energy of $\sqrt{s} = 13$ TeV. Values obtained using Equation 3.7.

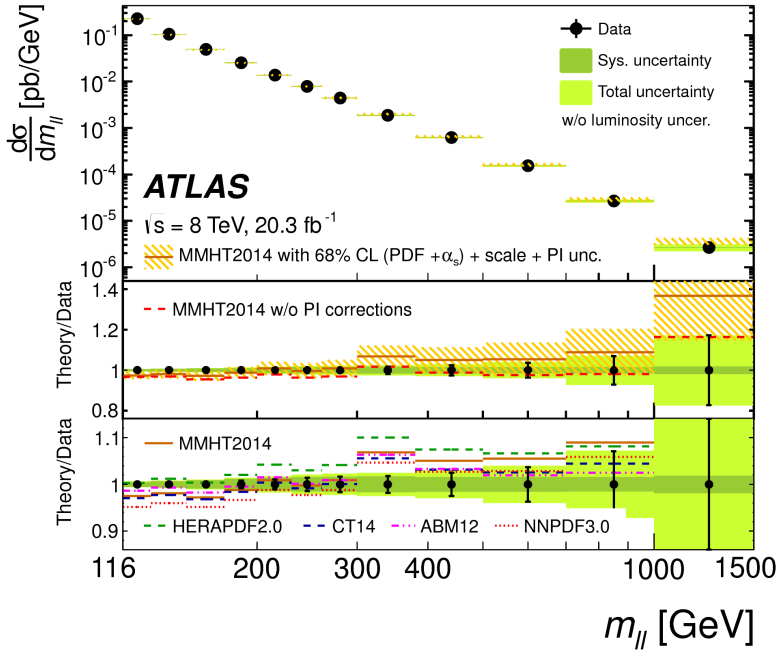


Figure 7.2: Single-differential Drell-Yan cross-section at $\sqrt{s} = 8$ TeV as a function of invariant mass, $m_{\ell\ell}$, using ATLAS Run 1 data. The measured distribution is compared to combined NNLO QCD and NLO EW calculations using different PDFs. From Ref. [124].

of $m_{\ell\ell}$, and two double-differential measurements: one as a function of invariant mass ($m_{\ell\ell}$) and dilepton rapidity ($y_{\ell\ell}$) and another as a function of invariant mass and $\cos\theta^*_{\text{CS}}$ (see Section 3.1.2 for a detailed explanation on this variable). The results presented in this thesis focus on the single-differential measurement against $m_{\ell\ell}$ and the double-differential measurement against $m_{\ell\ell}$ and $\cos\theta^*_{\text{CS}}$, which are used for the Effective Field Theory interpretation. The data used for this analysis was collected by the ATLAS experiment during the LHC Run 2 and amounts to an integrated luminosity of 139 fb^{-1}. The high

dilepton invariant mass phase space of the ATLAS Run 2 dataset was previously probed in the context of searches for BSM resonant [27] and non-resonant phenomena [125] in the dilepton final states, where no deviations from Standard Model where found, placing limits on the masses of potential Z' candidates of $m_{Z'} > 5$ TeV and contact interaction energy scales higher than 35.8 TeV. With respect to the previous ATLAS Drell-Yan cross-section measurements, the one presented in this thesis offers several key differences:

- The increase to $\sqrt{s} = 13$ TeV data allows for the analysis to probe a wider range in x compared to the $\sqrt{s} = 8$ TeV results, which probed a maximum x of $\simeq 0.88$.
- The substantial increase in integrated luminosity used in the analyses results into reduced statistical uncertainties, allowing for a finer binning in the distributions used for the measurement. This increases the amount of information that can be extracted from the measurement, resulting into a better input for PDF fits. The increase in integrated luminosity also allows to expand the range of the measurement against dilepton invariant mass, adding an additional bin covering the range $1500 < m_{\ell\ell} < 5000$ GeV.
- New measurement interpretations have been included, as presented in Chapter 10.
- The addition of the double-differential measurement in $m_{\ell\ell}$ and $\cos\theta^*_{\text{CS}}$ enhances the sensitivity to the photon-induced process and provides a crucial input for Effective Field Theory interpretations.

The following chapters present the different elements involved in the measurement and interpretation of the neutral-current Drell-Yan cross-section at ATLAS. Chapter 8 presents an overview of the analysis strategy, covering the dataset and MC samples used, event selection and background estimation techniques, as well as a summary of the systematic uncertainties of the measurement. In Chapter 9 the cross-section measurement results are shown, including a brief explanation on how electron and muon channel results are combined for a reduction of the total uncertainty of the measurement. These results are used as an input for Effective Field Theory and Lepton Flavour Universality interpretations, as explained in Chapter 10. This part of the thesis concludes with a summary and outlook of the measurement in Chapter 11.

Chapter 8

Measurement overview

This chapter describes the measurement of the cross-section of the Drell-Yan process with the ATLAS experiment. Section 8.1 presents all Monte Carlo samples used for the analysis. The data, as well as the event selection criteria, used in the analysis are explained in Section 8.2. Section 8.3 presents the methodology used to extract the cross-section from the data. The estimation of the different irreducible background processes is discussed in Section 8.4. Finally, Sections 8.5 and 8.6 provide an overview of the different uncertainties on the measurement, explaining the contributions of experimental and theoretical systematic uncertainties respectively. See Chapter 1 for a breakdown of my personal contributions to the analysis.

8.1 Monte Carlo simulation

Several Monte Carlo samples are used in this analysis to model the signal and background processes of interest. These have been centrally provided by the ATLAS Collaboration and were processed through a simulation of the detector geometry and response [126] using GEANT4 [69]. The framework used for reconstruction is the same for both simulation and real data. The effect of multiple interactions in the same and neighbouring bunch crossings (pile-up) was modelled by overlaying the simulated hard-scattering event with inelastic pp events generated with PYTHIA 8.186 [67] using the NNPDF2.3LO set of PDFs [127] and the A3 set of tuned parameters [128]. Correction factors are applied to account for differences between the observed and simulated detector response, as explained in Chapter 6.

8.1.1 Signal processes

Drell-Yan process

The POWHEGBOX v1 generator [73–75] was used for the simulation at NLO accuracy of the hard-scattering processes of Z-boson production and decay in the electron and muon channels. It was interfaced to PYTHIA 8.186 [67] for the modelling of the parton shower,

hadronisation, and underlying event, with parameters set according to the ATLAS AZNLO tune [129]. The CT10 NLO PDF set [130] was used for the hard-scattering processes, whereas the CTEQ6L1 PDF set [131] was used for the parton shower. The effect of QED final-state radiation was simulated with PHOTOS++3.52 [132, 133]. The EVTGEN1.2.0 program [80] was used to decay bottom and charm hadrons. To ensure that enough events are simulated at high invariant masses, such that the MC statistical uncertainty is not a limiting factor for the analysis, the samples are generated in 18 bins of dilepton invariant mass between 120 and 5000 GeV. In addition, one sample is generated simulating all the mass spectrum above $m_{\ell\ell} > 60$ GeV, usually referred to as the *inclusive* sample. To avoid an overlap in event generation with the mass-binned samples, a cut of $m_{\ell\ell} < 120$ GeV is used on the true dilepton invariant mass of the inclusive sample.

While the DY cross-section is calculated at NLO with POWHEGBOX, the prediction is further corrected to NNLO QCD using mass-dependent k-factors [134]. The factors are calculated using the VRAP 0.9 simulation code [84], offering different k-factors that representing various NNLO PDFs. The baseline prediction used in the analysis uses the CT18ANNLO [135] PDF set. Additional corrections, accounting for NLO EW effects (minus the QED FSR corrections) are calculated with FEWZ 3.1.b2 [81–83] and MCSANC [136]. The impact of these corrections can be found in Figure 8.1, against both $m_{\ell\ell}$ and $y_{\ell\ell}$. The combination of the EW and QCD higher-order corrections are performed using the *additive* approach. Further details on the combination of these corrections, as well as an assessment of the impact of the combination approach, can be found in Section 8.6.

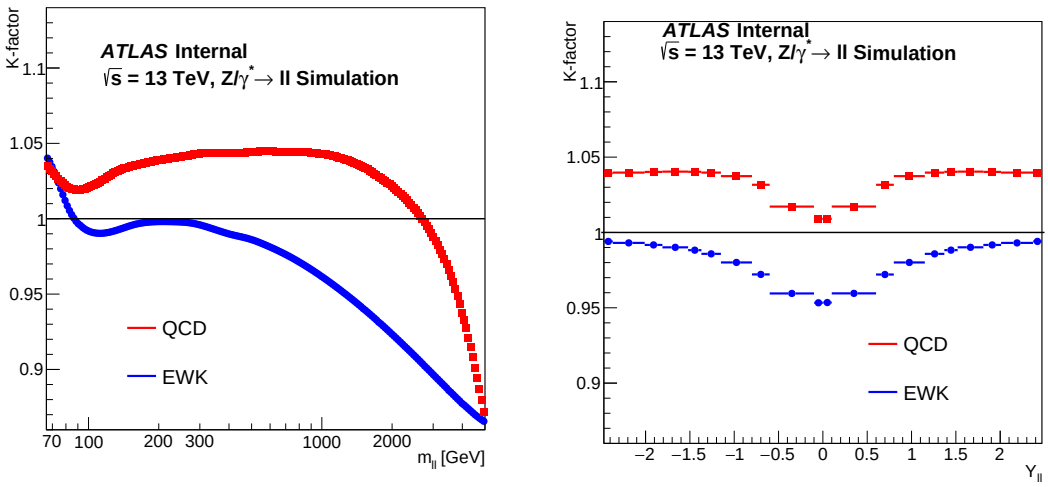


Figure 8.1: Higher-order corrections to the neutral-current Drell-Yan cross-section as a function of dilepton invariant mass (left) and rapidity (right). Corrections account for NNLO QCD differences with respect to NLO QCD (red) and NLO EW (+LO QCD) with respect to LO EW+QCD (blue). From Ref [137].

An alternative sample simulating the Drell-Yan process was obtained using the SHERPA

2.2.11 [138] generator using next-to-leading-order (NLO) matrix elements (ME) for up to two partons, and leading-order (LO) matrix elements for up to four partons calculated with the Comix [139] and OpenLoops [140–142] libraries. They were matched with the Sherpa parton shower [143] using the MEPSatNLO prescription [77–79,144] using the set of tuned parameters developed by the Sherpa authors. The NNPDF3.0nnlo set of PDFs [145] was used as the baseline choice, but the sample offers several re-weighing options accounting for the difference between various PDF sets. This sample is used for systematic checks, as well as the assessment of theory uncertainties presented in Section 8.6.

Photon-induced process

The photon-induced process, which will be considered part of the signal in this analysis, is simulated at leading-order using Pythia 8 and the NNPDF3.1nlo LUX PDF set [146] for the matrix element, and the NNPDF2.3lo PDF set [127] for the showering and hadronisation. The contribution of the sample is below 1% to the total number of events in the lower invariant mass region of the analysis, but can account for $\sim 5\%$ of the events at high mass ($m_{\ell\ell} > 1$ TeV) and high values of $\cos\theta^*_{\text{CS}}$ ($|\cos\theta^*_{\text{CS}}| > 0.8$).

8.1.2 Background processes

Top-quark processes

Top quarks decay into b quarks and a W-boson, which can decay leptonically as $W \to \ell\nu$. Therefore, background to the high-mass Drell-Yan can arise from processes involving top-quark decays. Two main contributions can arise from these processes: the production of top-antitop pairs ($pp \to t\bar{t} + X \to W^+bW^-\bar{b} + X$) and the associated production of a single (anti)top quark with a W-boson ($pp \to tW + X \to WbW + X$).

The production of $t\bar{t}$ and Wt events was modelled using the Powheg Box v2 [73–75] generator at NLO with the NNPDF3.0nlo [145] PDF set. The events were interfaced to Pythia 8.230 [147] to model the parton shower, hadronisation, and underlying event, with parameters set according to the A14 tune [148] and using the NNPDF2.3lo set of PDFs [127]. The decays of bottom and charm hadrons were performed by EvtGen 1.6.0 [149]. The diagram removal scheme [150] was used to remove interference and overlap between $t\bar{t}$ and Wt production. The events were interfaced to Pythia 8.230 [147] using the A14 tune [148] and the NNPDF2.3lo set of PDFs [127].

Higher-order NNLO QCD calculations of the top quark processes were made using Top++ 2.0 [151], deriving a corrected $t\bar{t}$ and single-top cross-section which results in global k-factors of 1.1949 and 1.054 respectively. These samples are used as part of the data-driven background estimation of the top-quark contributions, as explained in Section 8.4.

Diboson processes

Another background contribution stems from the combined production of two vector bosons (WW, WZ and ZZ), namely the diboson production. When the bosons undergo leptonic decays they produce a final state similar to that of Drell-Yan. Though the contribution of this background can be reduced by imposing a cut on the number of leptons present in the events used in the analysis (see Section 8.2), the remaining contributions need to be accounted for.

Samples of diboson final states (VV) were simulated with the Sherpa 2.2.2 [138] generator. Fully leptonic final states and semileptonic final states, where one boson decays leptonically and the other hadronically, were generated using matrix elements at NLO accuracy in QCD for up to one additional parton and at LO accuracy for up to three additional parton emissions. The matrix element calculations were matched and merged with the Sherpa parton shower in the same way as the Drell-Yan signal sample as explained above. The NNPDF3.0nnlo PDF set was used [145], along with the dedicated set of tuned parton-shower parameters developed by the Sherpa authors.

τ-channel Drell-Yan

Additional sources of background include the Drell-Yan τ decay channel, where the Z/γ^* decays into a pair of τ leptons, not included in the signal of this analysis, which may decay leptonically: $\tau \to \nu_\tau \nu_\ell \ell$. If both tau leptons in the pair produced via Drell-Yan decay into the same lepton flavour the signature left in the detector is very similar to that of the Drell-Yan production of light leptons. A key difference is that due to the three-body decay of the τ lepton, the resulting light lepton is produced with a lower momentum than that of the original decay product, resulting in a much lower contribution than the electron and muon decay channels at the same energy range. The $Z/\gamma^* \to \tau^+\tau^-$ background is simulated in the same way as for the Drell-Yan signal and accordingly corrected for higher-order corrections.

8.2 Data and event selection

8.2.1 Data samples

The data used in the analysis was recorded by the ATLAS experiment in the LHC Run 2 at $\sqrt{s} = 13$ TeV, accounting for a total integrated luminosity of 139.0 fb^{-1}. Note that the luminosity measurement used as reference for the results presented in this thesis corresponds to the Run 2 preliminary results [152]. The luminosity used in the final results in the published analysis will correspond to the results reported in Ref. [109]. Further details on the data-taking conditions in this period can be found in Section 4.2. Figure 8.2 shows the sum of the integrated luminosity delivered by the LHC and recorded by ATLAS in the data taking period covered by the analysis. Data shown in blue corresponds to that passing a series of Data Quality checks that ensure that all sub-detectors operated optimally.

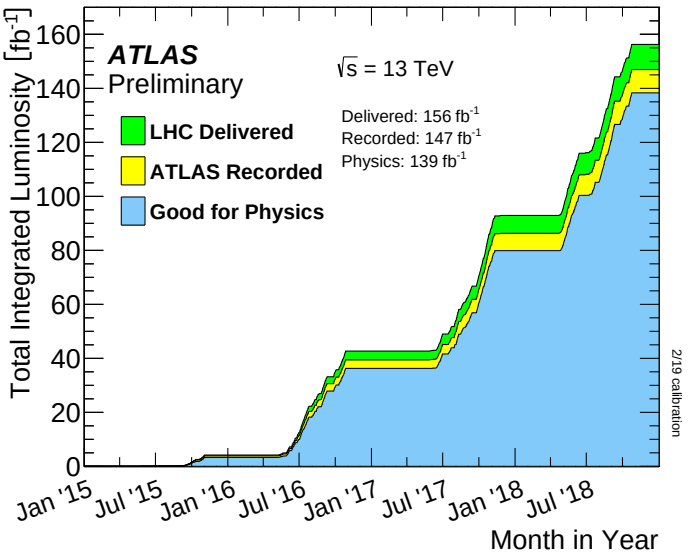


Figure 8.2: Cumulative integrated luminosity delivered to ATLAS (green), recorded by ATLAS (yellow) and determined to be good quality data (blue) during LHC Run 2. From Ref [153].

8.2.2 Event selection

Events used in the analysis are required to belong to a luminosity block passing the good quality selection criteria. They are also required to contain at least one primary vertex with more than two tracks. The events have to fulfil a trigger that requires the presence of at least one or two electrons or muons. A breakdown of the triggers used in the analysis can be found in Table 8.1. Di-electron trigger efficiencies are at least 90% for electrons with $p_{\mathrm{T}} > 30$ GeV, rising to $> 95\%$ for electrons with $p_{\mathrm{T}} > 60$ GeV. Single-muon trigger efficiencies are around 70% and 85% in the barrel and endcap regions respectively for muons with $p_{\mathrm{T}} > 30$ GeV, while di-muon efficiencies increase to 75 to 87% in the same kinematic

phase space. More information about the electron and muon ATLAS Run 2 triggers, as well as their efficiencies, can be found in Refs. [154] and [155].

Year	Trigger	Requirement
2015	HLT_2e12_lhloose_L12EM10VH	2 tracks with p_{T} >12 and Loose ID
	HLT_mu20_iloose_L1MU15	1 CB muon, p_{T} >20
	HLT_mu50	1 CB muon, p_{T} >50
	HLT_2mu10	2 CB muons, p_{T} >10
2016	2e17_lhvloose_nod0	2 tracks with p_{T} >17 and Very Loose ID
	HLT_mu26_ivarmedium	1 CB muon, p_{T} >26 and Medium iso
	HLT_mu50	See above
	HLT_2mu14	2 CB muons, p_{T} >14
2017-18	2e24_lhvloose_nod0	2 tracks, p_{T} >24 and Very Loose ID
	HLT_mu26_ivarmedium	See above
	HLT_mu50	See above
	HLT_2mu14	See above

Table 8.1: Summary of the single- and di-lepton triggers used. CB muons refer to combined muons, see Section 5.2 for more details. Triggers containing the "nod0" chain do not include the transverse impact parameter and its significance in the online trigger particle ID (see Sections 5.2 and 6.1).

The event selection requires the presence of two same-flavour good quality leptons (selection criteria for these are shown in Table 8.2) with opposite charges. If a pair of same-flavour leptons is found in the event, another cut is placed vetoing the presence of a different flavour lepton, meaning that if two good quality electrons are found in an event together with a good quality muon (or viceversa), the event is discarded. Finally, the lepton pair is required to have a minimum invariant mass of 66 GeV.

8.2.3 Lepton objects selection

Lepton pair candidates are further selected to reduce the background from other processes while retaining a high signal selection efficiency. The lepton pair will consist of a so-called *leading* and *subleading* lepton, referring to the candidates with higher and lower p_{T} respectively. Table 8.2 summarises the lepton selection criteria used in the analysis. A detailed description of the methodology for lepton reconstruction and identification is found in Chapter 6.

8.2.4 Binning

The one dimensional binning for the single-differential cross section measurement is:

Selection criteria	**Electron channel**	**Muon channel**
Transverse momentum	$p_{\mathrm{T}}^{\mathrm{lead}} > 40$ GeV $p_{\mathrm{T}}^{\mathrm{sublead}} > 30$ GeV	$p_{\mathrm{T}}^{\mathrm{lead}} > 40$ GeV $p_{\mathrm{T}}^{\mathrm{sublead}} > 30$ GeV
Pseudorapidity	$0 < \lvert\eta^{e}\rvert < 1.37$ or $1.52 < \lvert\eta^{e}\rvert < 2.4$	$0 < \lvert\eta^{\mu}\rvert < 2.4$
Identification	Medium	High-p_{T}, combined muon
Isolation	$E_{\mathrm{T}}^{\mathrm{topo,cone20}}/p_{\mathrm{T}} < 0.06$ and $p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.06$	$p_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.06$
Track-vertex association	$\lvert z_0 \sin\theta\rvert < 0.5$ mm $\lvert d_0\rvert/\sigma(d_0) < 5$	$\lvert z_0 \sin\theta\rvert < 0.5$ mm $\lvert d_0\rvert/\sigma(d_0) < 3$
Invariant mass	$m_{e^+e^-} > 66$ GeV	$m_{\mu^+\mu^-} > 66$ GeV

Table 8.2: Overview of the lepton selection criteria, where each criterion is applied to a single lepton. The fiducial phase space is determined by the invariant-mass, $|\eta^{\ell}|$ and p_{T}^{ℓ} requirements.

$$m_{\ell\ell} = [116, 130, 150, 175, 200, 230, 260, 300, 380, 500, 700, 1000, 1500, 5000] \text{ GeV} \tag{8.1}$$

The binning for the two-dimensional binning of the double-differential measurement in $m_{\ell\ell}$ and $y_{\ell\ell}$ is:

$$\begin{aligned} & m_{\ell\ell} = [116, 150, 200, 300, 500] \text{ GeV} \\ & \quad \times |y_{\ell\ell}| = [0.0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4]; \\ & m_{\ell\ell} = [500, 1500] \text{ GeV} \\ & \quad \times |y_{\ell\ell}| = [0.0, 0.3, 0.6, 0.9, 1.2, 1.5, 1.8, 2.1, 2.4] \end{aligned} \tag{8.2}$$

while the two-dimensional binning for the measurement of the cross-section against $m_{\ell\ell}$ and $\cos\theta^{*}_{\mathrm{CS}}$ is:

$$\begin{aligned} & m_{\ell\ell} = [116, 150, 200, 300, 500] \text{ GeV} \\ & \quad \times \cos\theta^{*}_{\mathrm{CS}} = [-1, -0.8, -0.6, -0.4, -0.2, 0.0, 0.2, 0.4, 0.6, 0.8, 1] \\ & m_{\ell\ell} = [500, 1500] \text{ GeV} \times \cos\theta^{*}_{\mathrm{CS}} = [-1, -0.75, -0.5, -0.25, 0, 0.25, 0.5, 0.75, 1] \end{aligned} \tag{8.3}$$

The binning was chosen to optimise the bin purity, while keeping a fine granularity [137]. The invariant mass binning is extended with respect to the previous $\sqrt{s} = 8$ TeV measurement by adding a high-mass bin covering the range $1500 < m_{\ell\ell} < 5000$ GeV.

8.3 Measurement methodology

The fiducial single-differential cross-section within the kinematic cuts previously presented is given by:

$$\frac{\mathrm{d}\sigma_{\mathrm{fid}}}{\mathrm{d}m_{\ell\ell}} = \frac{N_{\mathrm{sig}}(m_{\ell\ell})}{C_{\mathrm{DY}}(m_{\ell\ell}) \cdot \mathcal{L}_{\mathrm{data}}}, \tag{8.4}$$

where N_{sig} is the number of signal events in data, $\mathcal{L}_{\mathrm{data}}$ is the integrated luminosity used in the analysis and C_{DY} is the correction for efficiency and migration effects calculated from simulation as

$$C_{\mathrm{DY}} = \frac{N^{\mathrm{rec}}(m_{\ell\ell})}{N^{\mathrm{gen}}_{\mathrm{fid}}(m_{\ell\ell})}. \tag{8.5}$$

Here $N^{\mathrm{rec}}(m_{\ell\ell})$ is the sum of weights of Monte Carlo events after simulation of the event reconstruction and selection in data and $N^{\mathrm{gen}}_{\mathrm{fid}}(m_{\ell\ell})$ is the sum of weights at generator level after some common fiducial cuts that will define the phase space of the measurement. In the case of this analysis these cuts are:

$$|\eta^{\ell}| < 2.5,\ p_{\mathrm{T}}^{\mathrm{lead}} > 40\ \mathrm{GeV},\ p_{\mathrm{T}}^{\mathrm{sublead}} > 30\ \mathrm{GeV}. \tag{8.6}$$

The cross-section can be measured at different levels of QED FSR corrections corresponding to various truth level definitions of $N^{\mathrm{gen}}_{\mathrm{fid}}$:

- Born level: leptons used at this level are simulated before QED FSR. Using these leptons fully corrects the distributions for QED FSR effects.
- Bare level: leptons in this level include QED FSR, resulting in a measurement which is not corrected for QED effects.
- Dressed level: hybrid between both levels, where the 4-vectors of the born leptons and all photons produced via QED FSR found within a $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.1$ radius of the lepton are resummed. Using this level performs a partial QED FSR correction, removing the effects of for the collinear emissions but not of the wide angle part.

Figure 8.3 shows a diagram of each of the definitions, while Figure 8.4 shows the impact of each definition on the dilepton invariant mass distribution at the Z-boson mass peak. In the context of this analysis, truth-level leptons will refer to those at Born level. The methodology shown in Equation 8.4 can be expanded to double-differential distributions accounting for the changes in $N^{\mathrm{rec}}(m_{\ell\ell}, ...)$ and $C_{\mathrm{DY}}(m_{\ell\ell}, ...)$.

As opposed to the bin-bin-bin unfolding previously presented, an iterative unfolding method based on Bayes' theorem is used [156], offering more robustness against underlying

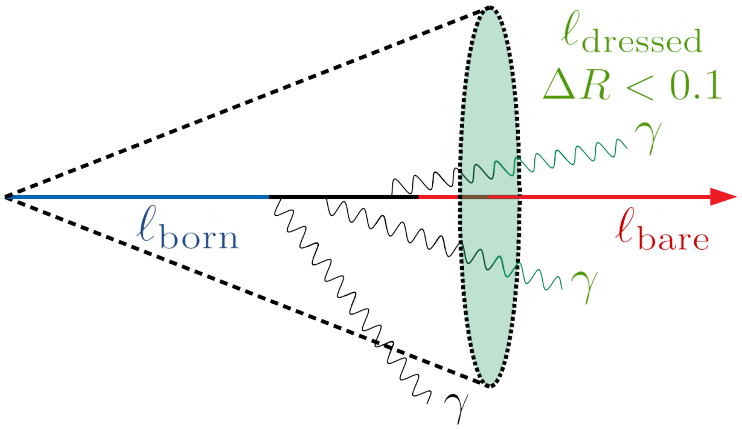


Figure 8.3: Illustration of the different truth-level lepton definitions. Born-level (blue) leptons radiate photons (FSR), resulting in bare-level (red) leptons with lower energy. Dressed-level leptons (green) are defined combining the bare-lepton 4-vector with all radiation found within $\Delta R < 0.1$.

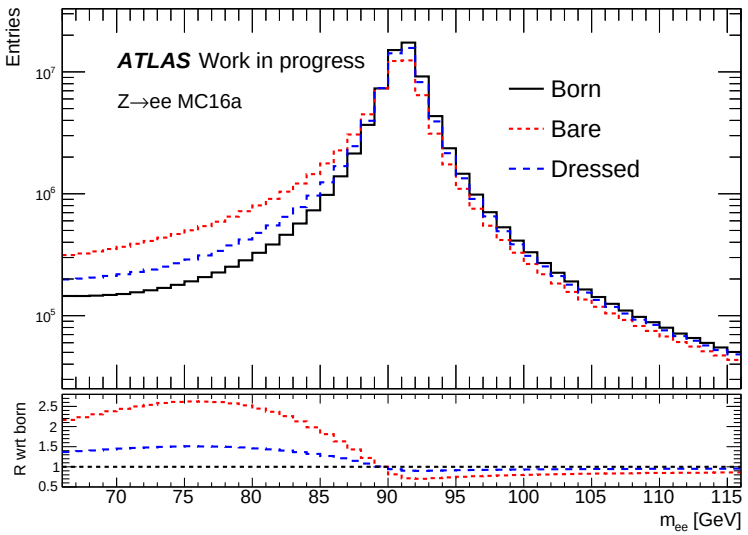


Figure 8.4: Dielectron invariant-mass distribution around m_Z for born-level (black), bare-level (red) and dressed-level (blue) lepton definitions. The bottom panel shows the ratio of the bare- and dressed-level distributions with respect to born-level.

generator distributions. The yield of truth events in each bin of a given distribution are iteratively reproduced starting from a detector response matrix and a simulated truth prior. Thus, the conditional probability for an event in the R_j bin of the distribution at reconstructed level to have originated from an event generated in the T_i bin of the truth distribution is given by:

$$P(T_i|R_j) = \frac{P(R_j|T_i)P_0(T_i)}{\sum_{l=1}^{N_{\text{bins}}} P(R_j|T_l)P_0(T_l)}, \tag{8.7}$$

where $P(R_j|T_i)$ is the conditional probability of an event generated in the bin i of the

truth-level distribution to be detected in the bin j in the reconstructed distribution and $P_0(T_i)$ is the probability for an event to be generated in the i bin at truth-level. The distribution $P(R_j|T_i)$ is obtained from Monte-Carlo samples, simulating the detector response and constructing said probability by normalising each column in the distribution, as shown in Figure 8.5. The distribution is mostly diagonal, which corresponds to the measurement of the same value of $m_{\ell\ell}$ at reconstructed-level as the original truth-level value. Small migrations ($< 10\%$) can be observed across the distribution, but probabilities in the neighbouring bins outside of the diagonal terms are orders of magnitude smaller. Additional response matrices in the muon channel and two-dimensional distributions can be found in Ref. [137].

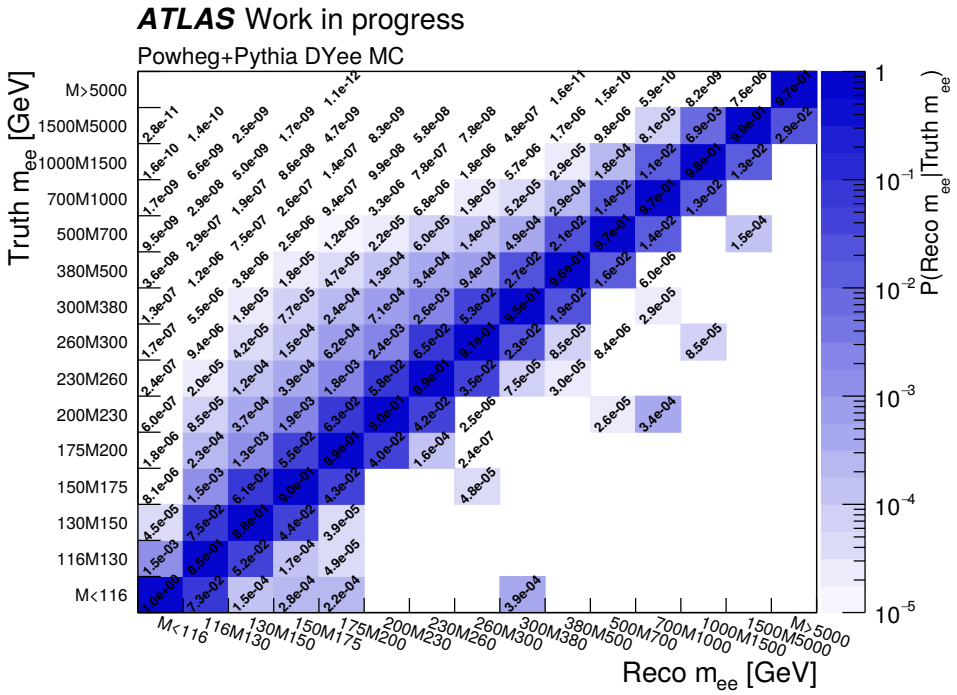


Figure 8.5: Response matrix for the invariant mass for the dielectron channel at born level. The distribution is normalised per column, showing the probability of a reconstructed dilepton invariant mass to originate from each truth-level value.

Starting from a prior truth-level distribution $P_0(T_i)$ one can calculate $P(T_i|R_j)$ in a first iteration using Equation 8.7. This probability can then be used to compute the number of events in the bin i of the truth-level distribution, $\hat{n}(T_i)$, using the number of events in the bin j of the reconstructed distribution, $n(R_j)$:

$$\hat{n}(T_i) = \sum_{j=1}^{N_{\text{bins}}} n(R_j) P(T_i|R_j). \tag{8.8}$$

One can then re-calculate the probability distribution $P_0(C_i)$ by normalising the sum

of $\hat{n}(T_i)$ to unity, which can be used again as input for an updated calculation of the probability in Equation 8.7. The procedure is iterated until a stable spectrum is obtained. In the case of the analysis presented in this thesis the number of iterations required to ensure stability in the unfolding procedure was 5 [137].

8.4 Background estimation

For most background processes, the Monte-Carlo samples introduced in Section 8.1 are used. However, two outstanding contributions benefit from the use of data-driven techniques in their estimation. This Section provides an overview of the procedure used to estimate those backgrounds.

8.4.1 Top quark background

The top-quark processes are the leading background in the Drell-Yan measurement, meaning that theoretical uncertainties on the estimation of this background results in a significant contribution to the total uncertainty in the cross-section measurement. An alternative method to obtain the top quark background is to derive the dilepton (ee and $\mu\mu$) distributions from a mixed-flavour ($e\mu$) control-region. The event selection mimics that used in the nominal analysis (first presented in Section 8.2), changing the requirement of the presence of a same-flavour lepton pair to require events with the presence of exactly one electron and one muon. Thus, the Drell-Yan contribution is not the main source of events, but rather $t\bar{t}$ and single-top processes, where the $e\mu$ events have at truth-level the same properties as ee or $\mu\mu$, as the $W \to \ell\nu$ decays are independent. Once the $e\mu$ distribution (single- and double-differentially) has been obtained, one can derive a transfer factor to convert the $e\mu$ kinematic distribution into that obtained with the nominal same-flavour selection. The transfer factor as a function of invariant mass obtained for both lepton flavours can be found in Ref. [137]. This transfer factor is applied on the $e\mu$ distribution observed in data in the opposite-flavour control region, subtracting all contributions from other SM processes to the final state. The transfer factor corrects for the different detector response to ee, $e\mu$ and $\mu\mu$ final states. By calculating the top-quark background taking the ratio between the $e\mu$ and $ee/\mu\mu$ distributions, theoretical uncertainties largely cancel, resulting in a reduction on the total uncertainty of the cross-section measurement.

8.4.2 Multijet background

Background leptons from misidentified jets passing the lepton candidate selection are estimated using the *matrix method*, as described in Ref. [124]. In this method, the identification criteria for leptons is loosened and the rate with which these looser objects are reconstructed as signal-like electrons is measured separately. Thus, two categories of lepton candidates are defined depending on whether they pass the signal criteria (as described in

Sec. 8.2) or the looser one. The former are referred to as *Tight* leptons, and the latter as *Loose* leptons. Given that we apply a more restrictive set of cuts on this category, N_{Tight} has to be a subset of the loose selection N_{Loose}.

The Tight selection corresponds to the nominal signal selection in the analysis, as described in Sec. 8.2. All objects in the Loose selection have to fulfil the kinematic cuts of the signal selection, but in the case of the electrons they are required to pass the Medium ID criteria and no cut on isolation is applied, while Loose muons are required to fail the isolation cut or the *high-*p_{T} muon quality cut, while still passing the Medium quality cuts.

Since only lepton pairs are considered in the analysis, one can denote them as N_{ij}, with $i, j \in T, L$ and thus having four possible event categories: N_{TT}, N_{TL}, N_{LT}, N_{LL}, the measurable quantities. These can be related to the "true" quantities, that is, the classification of events depending on whether an object was a real (R) or fake (F) lepton, N_{ab}, with $a, b \in R, F$. Both quantities are connected using an efficiency matrix M as follows:

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = M \begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix}. \tag{8.9}$$

The index T refers to signal-like (tight) objects, while more background-like objects are labelled as L for loose. Both left- and right-hand side quantities are defined exclusively, meaning that N_{RR} cannot include other components like N_{RF}, N_{FR} or N_{FF} and similarly N_{TT} does not contain components like N_{TL}, N_{LT} or N_{LL}. This leads to the definition of L to be "pass the loose selection but fail the tight selection", in contrast to the base selection N_{Loose}, which contains all objects included in N_{Tight}.

The matrix M connecting both sides of Equation 8.9 takes the form:

$$M = \begin{pmatrix} r_1 r_2 & r_1 f_2 & f_1 r_2 & f_1 f_2 \\ r_1(1-r_2) & r_1(1-f_2) & f_1(1-r_2) & f_1(1-f_2) \\ (1-r_1) r_2 & (1-r_1) f_2 & (1-f_1) r_2 & (1-f_1) f_2 \\ (1-r_1)(1-r_2) & (1-r_1)(1-f_2) & (1-f_1)(1-r_2) & (1-f_1)(1-f_2) \end{pmatrix}. \tag{8.10}$$

The coefficients f represent the probability that a fake lepton passing the loose selection is reconstructed as a signal-like lepton. The coefficients r represent the probability of a real

lepton passing the loose selection to be reconstructed as a tight one. Note that all objects that pass the signal requirements (*tight* candidates) must fulfil the loose selection too and therefore both must comply with $0 \leq r, f \leq 1$. The first index in Equation 8.9 corresponds to the leading electron, while the second index refers to the subleading one. Similarly, subindex 1 in the efficiencies in Equation 8.10 will refer to the leading electron candidate and 2 stands for the subleading one. The real and fake efficiencies can be expressed as:

$$r = \frac{N_{\text{Tight}}^{\text{real}}}{N_{\text{Loose}}^{\text{real}}}, \; f = \frac{N_{\text{Tight}}^{\text{fake}}}{N_{\text{Loose}}^{\text{fake}}}. \tag{8.11}$$

The real efficiency r is calculated from MC samples, where all objects are true leptons, while the fake efficiency f is measured from data in background-enriched control regions. The number of misidentified signal events corresponds to the number of events in which both leptons pass the signal selection, N_{TT}, while containing at least one fake object. These events can come from events where one of the leptons gets misidentified (RF or FR) or where both leptons are misidentified (FF). The total number of fake leptons is obtained with the addition of these two contributions:

$$\begin{aligned} N_{TT}^{\ell+\text{fake}} &= r_1 f_2 N_{RF} + f_1 r_2 N_{FR} \\ N_{TT}^{\text{multijet}} &= f_1 f_2 N_{FF} \\ N_{TT}^{\text{fakes}} &= N_{TT}^{\ell+\text{fake}} + N_{TT}^{\text{multijet}} = r_1 f_2 N_{RF} + f_1 r_2 N_{FR} + f_1 f_2 N_{FF}. \end{aligned} \tag{8.12}$$

However, the true quantities N_{RF}, N_{FR} and N_{FF} cannot be directly measured so they need to be expressed in terms of events passing the loosened or signal criteria by inverting the matrix shown in Eq. 8.10. Thus, the number of misidentified events in the signal selection containing at least one fake object can be written as:

$$\begin{aligned} N_{TT}^{\text{fake}} = {} & \alpha[r_1 f_2 (f_1 - 1)(1 - r_2) + f_1 r_2 (r_1 - 1)(1 - f_2) + f_1 f_2 (1 - r_1)(1 - r_2)] N_{TT} \\ & + \alpha f_2 r_2 [r_1 (1 - f_1) + f_1 (1 - r_1) + f_1 (r_1 - 1)] N_{TL} \\ & + \alpha f_1 r_1 [f_2 (1 - r_2) + r_2 (1 - f_2) + f_2 (r_2 - 1)] N_{LT} \\ & - \alpha f_1 f_2 r_1 r_2 N_{LL}, \end{aligned} \tag{8.13}$$

where:

$$\alpha = \frac{1}{(r_1 - f_1)(r_2 - f_2)}. \tag{8.14}$$

Using Eq. 8.13 the number of background events contained in our signal can be calculated by applying the corresponding weight depending on whether the event was labelled as TT (both leptons classified as Tight), TL (Tight leading lepton, Loose subleading one), LT (Loose leading lepton, Tight subleading one) or LL (both leptons Loose). The real

and fake efficiencies, r and f, are calculated (in MC and data, respectively) depending on the transverse momentum and pseudo-rapidity of the electron, as explained next. The determination of the efficiencies will only be shown in the electron channel in this thesis. The calculation of the efficiencies in the muon channel can be found in Ref. [137].

Real electron efficiency

The measurement of the real efficiency, r, of the electrons is obtained from high statistics Drell-Yan MC samples. To ensure that only real electrons are chosen for this study a truth-matching is applied, requiring all electrons to be reconstructed within a cone of $\Delta R < 0.2$ around the truth leptons in the MC sample. Measuring the ratio described in Eq. 8.11 the distribution shown in Figure 8.6 (top) is obtained. The efficiency distribution shows a certain independence with respect to η, but there is a substantial difference depending on the p_{T} of the electron, where an increase in the real efficiency can be observed as the electron momentum increases, reaching values compatible with $r = 1$ at $p_{\mathrm{T}}^{e} > 1000$ GeV.

To study the drop in the efficiency in the low-p_{T} range (< 60 GeV), the real efficiency was re-assessed applying only the ID selection to the Tight electrons, separating in this way the effects on the efficiency coming from the ID and the isolation selections. The results can be found in Figure 8.6 (bottom). The rate in the low-momentum range increases in this way up to $\sim 95\%$, showing a good agreement with previous ATLAS results [157]. The change in efficiency when applying the isolation cut matches the expected behaviour, since the chosen operating point for the isolation selection changed between the previous analysis and the one currently applied resulting in lower efficiencies at low-p_{T}.

Fake electron efficiency

The fake efficiency, f, is measured using samples of single fake electrons in data. For this measurement the events are required to pass one of the triggers in a given set of pre-scaled single-electron triggers, as shown in Table 8.3, requiring each trigger to fire in the momentum range specified in the Table. These triggers require the application of a prescale factor as taking the full luminosity would have exceeded the capabilities of the ATLAS data acquisition system.

Real electron contamination is further reduced imposing the requirements listed in Table 8.4. The kinematic cuts used in this selection matches that used in the nominal analysis, as listed in Table 8.2.

Once the events have been selected according to these criteria, the ratio f, can be computed requiring at least one of electron to pass the Loose ID criteria. However, despite the previously explained cuts, the contribution from real electrons cannot be neglected, so it

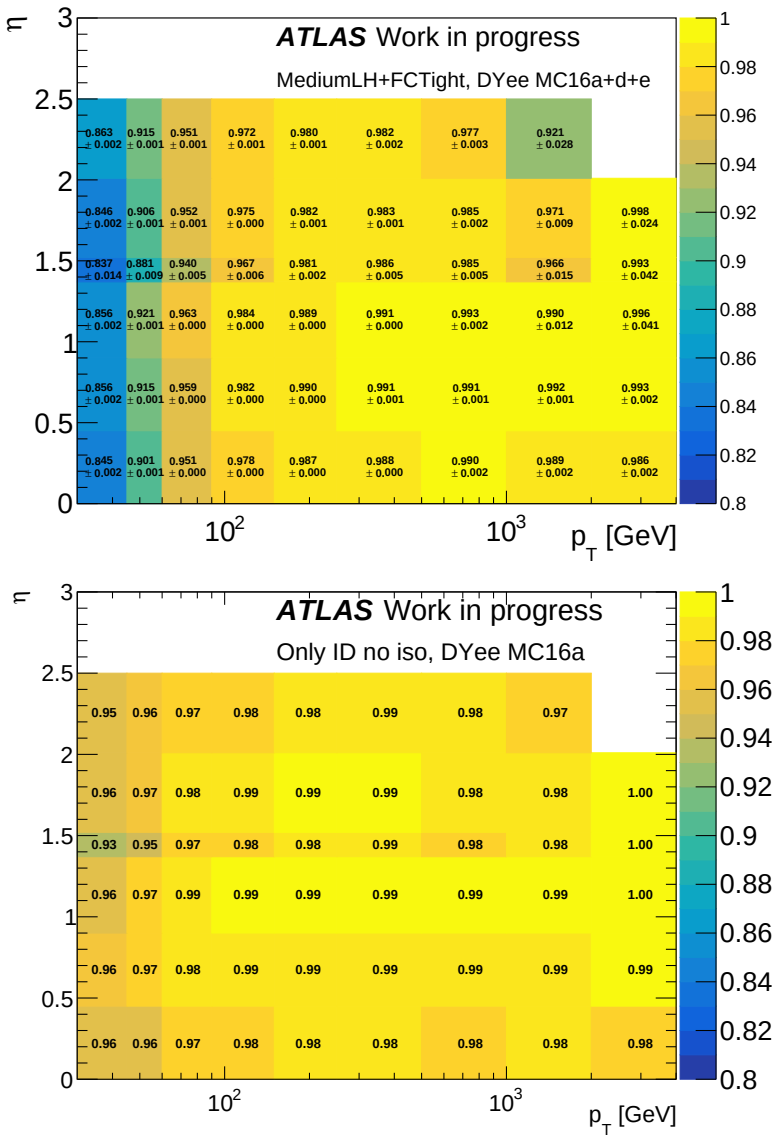


Figure 8.6: Real electron efficiency measured as a function of p_{T} and η. Signal electron candidates selected with cuts applied on ID and isolation (top) and on ID only (bottom).

must be estimated by MC samples of SM processes that can produce electron pairs, namely Drell-Yan, top quark, diboson and triboson processes. The samples used to simulate these processes are presented in Section 8.1. Thus, the fake efficiency f can be re-written as:

$$f = \frac{N_{\mathrm{Tight}}^{\mathrm{fake,data}} - N_{\mathrm{Tight}}^{\mathrm{fake,MC}}}{N_{\mathrm{Loose}}^{\mathrm{fake,data}} - N_{\mathrm{Loose}}^{\mathrm{fake,MC}}} \tag{8.15}$$

where $N^{\mathrm{fake,MC}}$ refers to real electrons produced in the different MC samples that we consider passing the fake-enrichment criteria. The electron p_{T} and η distributions of the Tight and Loose electrons after applying the fake selection can be found in Figures 8.7

Year	Trigger	Average prescale factor	p_{T} range [GeV]
2015	e26_lhvloose_nod0_L1EM20VH	125.0	30-65
	e60_lhvloose_nod0	12.0	65-315
	e200_etcut	1.0	>315
2016	e26_lhvloose_nod0_L1EM20VH	110.02	30-65
	e60_lhvloose_nod0	28.50	65-126
	e120_lhvloose_nod0	15.27	126-147
	e140_lhvloose_nod0	1.90	147-315
	e300_etcut	1.0	>315
2017	e26_lhvloose_nod0_L1EM20VH	472.64	30-65
	e60_lhvloose_nod0	30.82	65-147
	e140_lhvloose_nod0	2.44	147-315
	e300_etcut	1.0	>315
2018	e26_lhvloose_nod0_L1EM20VH	494.32	30-65
	e60_lhvloose_nod0	43.07	65-147
	e140_lhvloose_nod0	2.81	147-315
	e300_etcut	1.0	>315

Table 8.3: Trigger scheme used for the selection of a fake-enriched sample for the full Run 2 dataset. Average prescale factor calculated as the luminosity ratio between the total luminosity registered in a given period and the luminosity recorded with the application of the considered trigger. Trigger naming (and selection) convention follows that presented in Table 8.1.

Cut label	Selection criteria	Targeted process
A	$E_{\mathrm{T}}^{\mathrm{miss}} < 25$ GeV	W decays
B	< 2 Medium electrons	Drell-Yan dilepton production
C	$\lvert m_Z - m_{\mathrm{ee}}\rvert > 20$ GeV	Z-mass resonance

Table 8.4: Overview of the selection criteria used for the derivation of the fake electron efficiency. The transverse missing energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, is calculated as the opposite of the vector sum of the transverse momenta of all reconstructed particles, using $\sum_{\mathrm{reco}} \overrightarrow{p_{\mathrm{T}}} + \overrightarrow{E_{\mathrm{T}}}^{\mathrm{miss}} = 0$.

and 8.8. The data excess observed with respect to the MC distributions corresponds to the fake electrons with which f will be estimated. At $p_{\mathrm{T}} \sim 300$ GeV in the Loose selection a bump can be inferred, corresponding to the application of the "e300_etcut" trigger, see Table 8.3. While the prescale normalisation for this trigger works as expected for good quality (Tight) objects, showing a smoothly falling distribution in p_{T}, the Loose selection criteria is passed by lower quality objects, likely corresponding to hadron activity faking electrons which would not match the objects that originally passed the required trigger.

This increases the number of candidates passing the criteria, causing the bump observed in the distribution. The results of the fake efficiency measurement can be found in Figure 8.9.

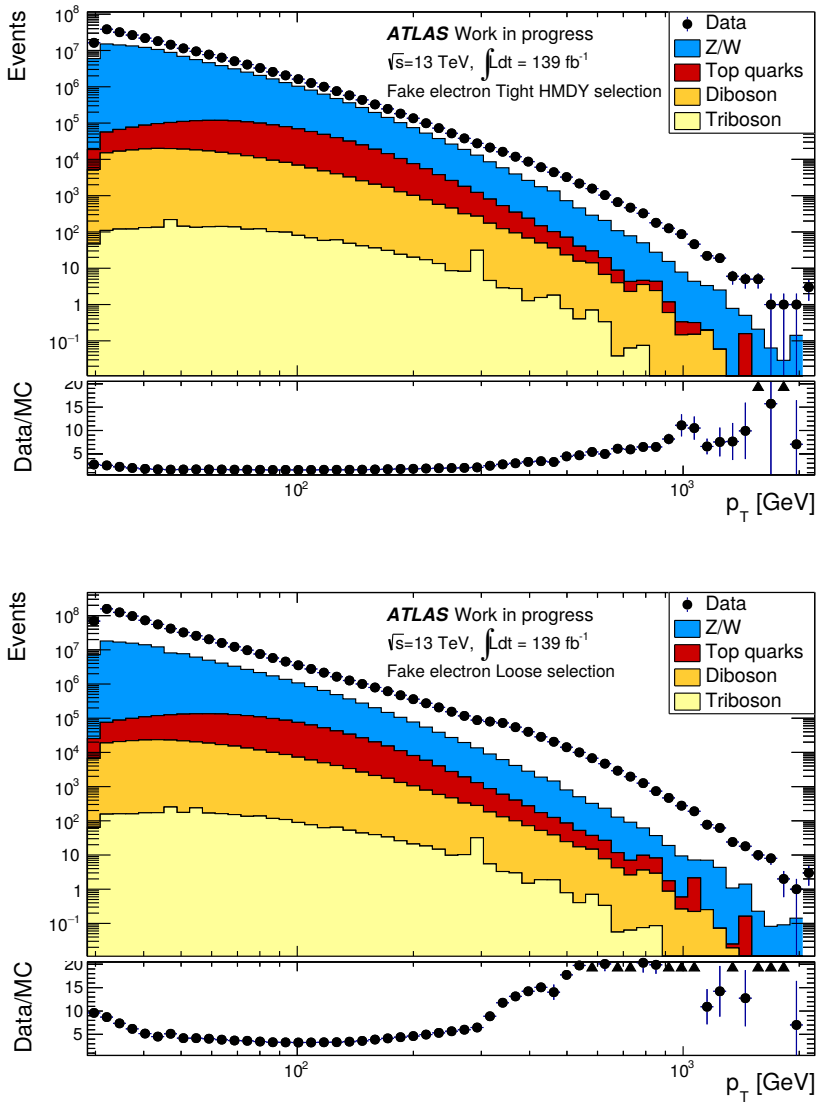


Figure 8.7: Transverse momentum, p_{T}, distribtution of the electrons passing the fake-enrichment criteria and complying with the Tight (top) and Loose (bottom) selection criteria.

Some of the criteria used in the assessment of the real and fake efficiencies can be varied, therefore several systematic uncertainties are considered for the fake electron background:

- Statistics: uncertainties on fake and real efficiencies from the limited number of candidate electrons. Note that this differs from the statistical uncertainty on the background estimation, but rather a systematic variation modifying the efficiencies used in the computation.

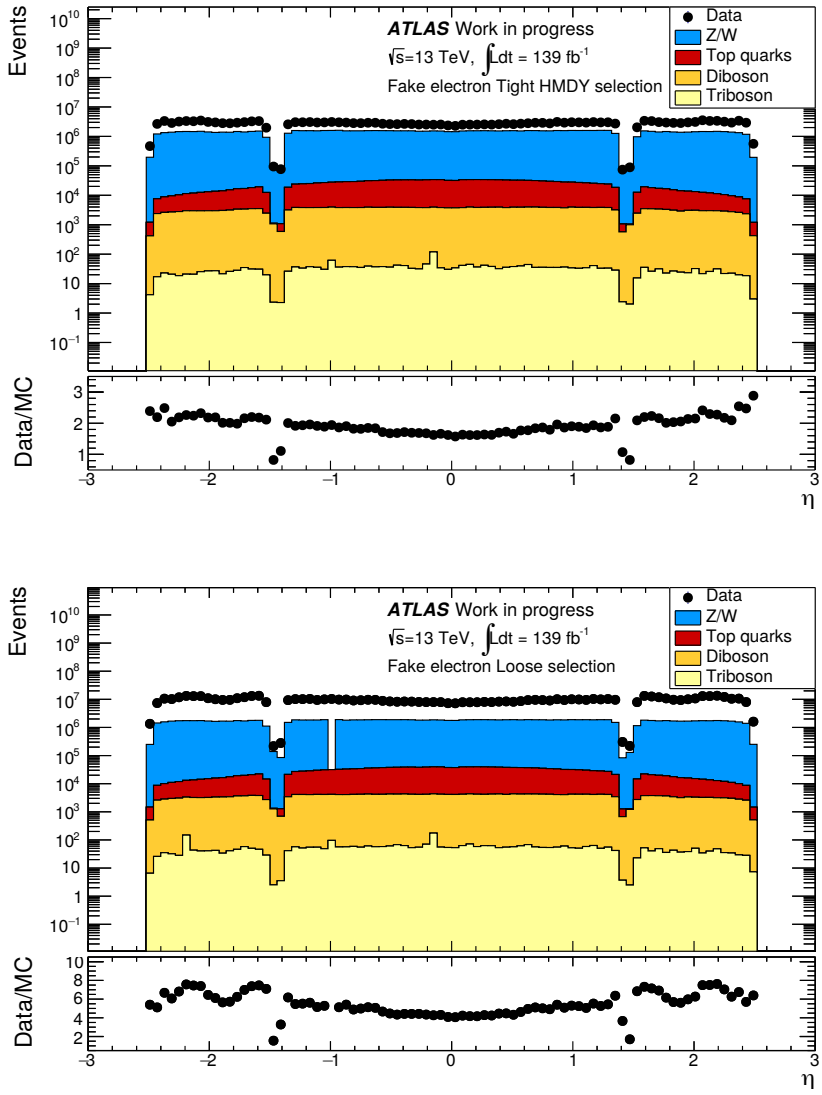


Figure 8.8: η distribution of the electrons passing the fake-enrichment criteria and complying with the Tight (top) and Loose (bottom) selection criteria.

- Prompt lepton subtraction: this source of uncertainty accounts for the prompt lepton subtraction carried out in the fake efficiency estimation when subtracting the MC from data (see Equation 8.15). A flat variation of ±30% was applied in each sample, which is larger than the modelling systematics.

- Composition of control and signal regions: the fake efficiency is calculated in a fake-enriched control region defined by 3 different cuts, as explained in the previous section. This source of uncertainty accounts for the effect of the different cuts, re-calculating the efficiency in the case where each of them is not applied.

The uncertainties due to each of the sources listed above can be found in Figure 8.10,

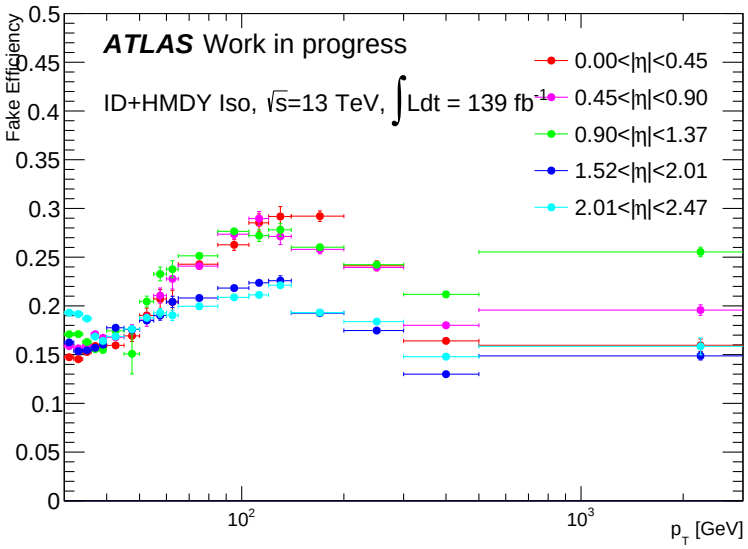


Figure 8.9: Fake electron efficiency measured with the Run 2 ATLAS dataset.

where it can be inferred that the leading contribution to the uncertainty comes from the background subtraction re-scaling, specially in the region $p_{\mathrm{T}} \sim 100$ GeV, where the presence of fake electrons is diminished and MC contributions are more substantial (see Figure 8.7). The total systematic uncertainty on the fake electron efficiency, computed as the quadratic sum of all contributions listed above, is shown in Figure 8.11. Large uncertainties, up to 90% of the fake efficiency, are observed across the distribution. The leading contribution is the modification of the prompt lepton subtraction.

Fake electron background estimation results

Once the real and fake efficiencies have been computed one can retrieve the fake electron background by classifying events as TT, TL, LT or LL according to the criteria passed by the leading and subleading electrons in the event and applying the corresponding weight retrieved using Equation 8.13. The final result of the estimation of the fake electron background can be found in Figure 8.12, expressed as a % of the total number of signal events found in data in each bin. The final background estimation was found to correspond to less than 1% with respect to the signal events in data in most of the mass range covered by the analysis, therefore the large uncertainty in its determination will have a small impact in the total uncertainty of the cross-section measurement. The background was retrieved both with a custom implementation of the matrix method estimation and with a framework provided by the ATLAS Isolation and Fake Forum [158] (IFF), showing good agreement between both estimations. Additional studies were carried out using an alternative methodology for the fake electron background estimation, the *template* method.

The results of these studies are shown in Appendix A. The template method was found to provide a lower background estimation at high dielectron invariant masses than the matrix method, in agreement with findings by other ATLAS analyses studying similar final states. Therefore, the final results of the analysis use the fake electron background estimation obtained with the matrix method, using the implementation of the Isolation and Fake Forum since it allows for an easier interfacing with other analysis tools. Results on the fake muon background estimation can be found in Ref. [137].

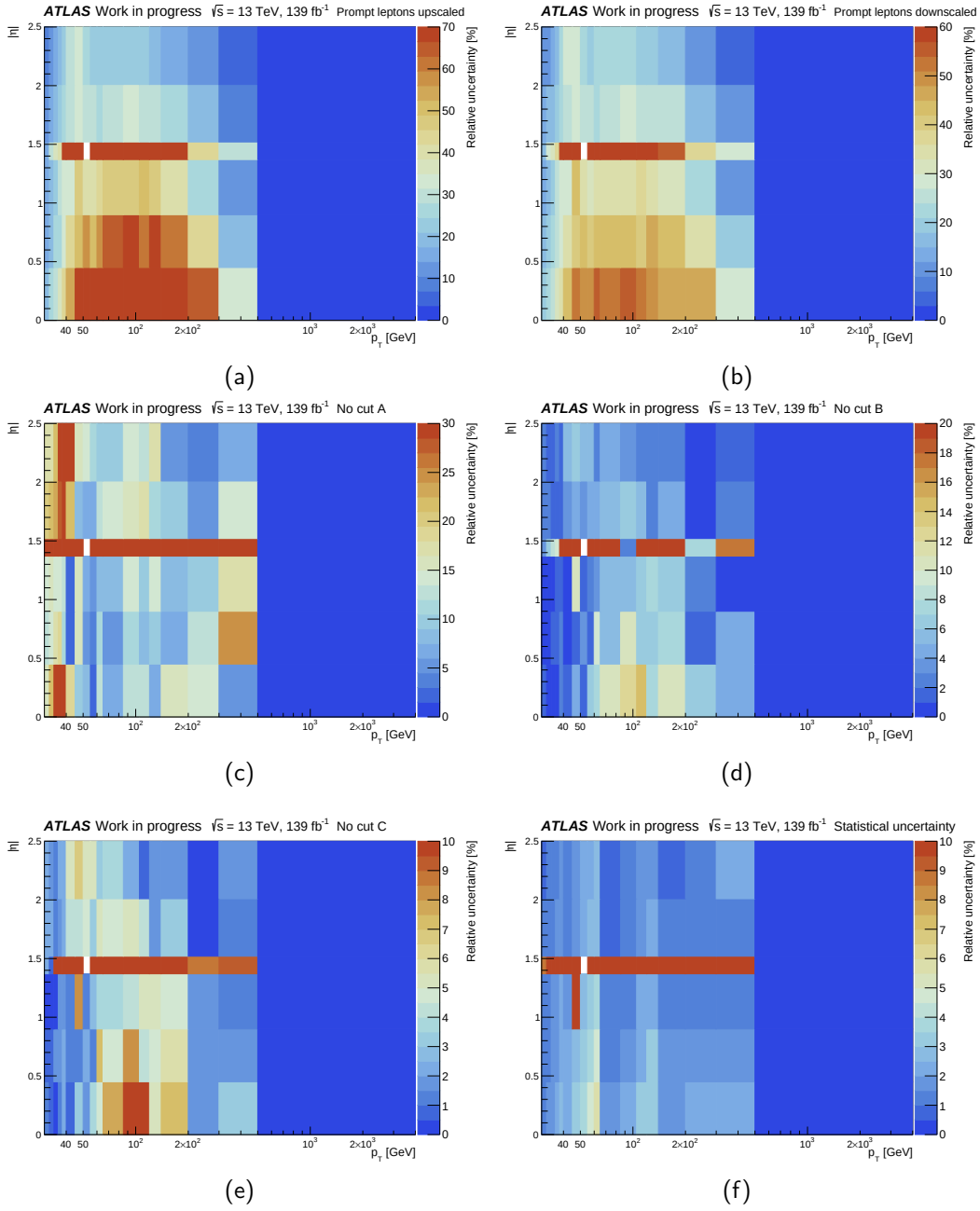


Figure 8.10: Different systematic uncertainty sources considered for the fake electron background: (a) Prompt lepton subtraction scaled up (+30%) (b) Prompt lepton subtraction scaled down (−30%) (c) Effect of removing the W (MET) cut from the fake-enriched CR (d) Effect of removing the DY cut from the fake-enriched CR (e) Effect of removing the Z mass cut from the fake-enriched CR (f) Statistical uncertainty on the fake efficiency. Uncertainties expressed as a percentage of the fake efficiency in every $\eta - p_{\mathrm{T}}$ bin.

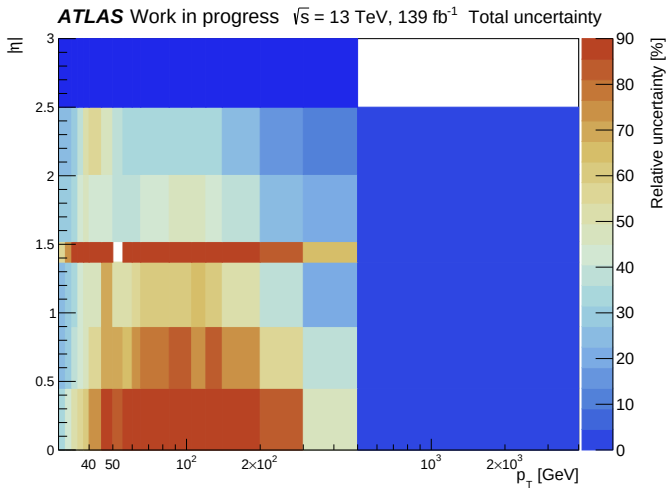


Figure 8.11: Total systematic uncertainty on the fake electron efficiency, calculated as the quadratic sum of values shown in Figure 8.10. Uncertainty expressed as a percentage of the fake efficiency in every $\eta - p_{\mathrm{T}}$ bin.

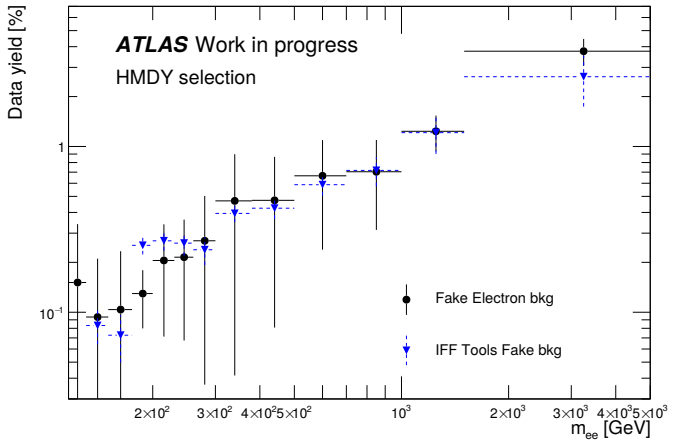


Figure 8.12: Fake electron background, shown as the fraction with respect to events passing the nominal selection in each mass bin. The background is calculated both with a custom analysis framework developed to retrieve the real and fake efficiencies (black) and the framework provided by the Isolation and Fake Forum (IFF, blue), showing good agreement between both estimations. Uncertainties on the IFF background only correspond to the statistical component, systematic uncertainties are equal to those shown for the ones derived with a custom framework.

8.5 Systematic uncertainties

Lepton efficiency and calibration uncertainties

The uncertainty on the efficiencies are generally split into multiple correlated and uncorrelated sources as determined from the performance analyses. The recommendations for the energy scale calibrations are applied to data, which include a total of 60 sources (as described in Ref. [117]). The effects of the uncertainties on the muon scale and resolution corrections (further described in Ref. [120]) are calculated by applying the variations of the corrections.

Background estimation systematic uncertainties

Uncertainties on the top background cover both experimental and theoretical systematic effects. The former include lepton reconstruction, identification and isolation corrections, such as those previously explained, which are correlated to those uncertainties in the signal events. Theoretical uncertainties include the variation of different parameters in the simulation of the process (showering parametrisation, QCD scales, m_{top}) as well as variations on the modelling of the hard scatter estimated using alternative top samples using MADGRAPH.

The uncertainty on the diboson background is given by its cross-section uncertainty, scaling the contributions up and down by 6% [159]. No uncertainty is considered for the $Z \to \tau\tau$ background given its negligible contribution to the total number of events. For an overview on the uncertainty estimation on the fake lepton background see Section 8.4.2.

Systematic uncertainties related to the unfolding procedure

There are several choices in the unfolding procedure explained in Section 8.3 that may vary among different analyses. The effect of varying these choices is estimated and introduced as an uncertainty in the measurement, including:

- Data-reweighting: the prior of the nominal Monte Carlo sample, $P_0(T_i)$ is reweighted to match the distribution found in data $P_0(R_i)$ to account for possible mismodelling effects affecting the unfolding procedure. The difference between the distribution obtained unfolding this pseudo-data distribution with respect to the nominal result is taken as a contribution to the unfolding systematic uncertainty.
- MC generator variation: the prior used in unfolding $P_0(T_i)$ is changed to that obtained at truth-level with an alternative MC generator, SHERPA, resulting into a different unfolded distribution.

- Additional sources: other effects tested in the unfolding procedure included the application of the different PDF variations available in the nominal MC sample to construct alternative response matrices $P(R_j|T_i)$ used to unfold the truth-level prior. Differences with respect to the basic bin-by-bin unfolding first presented in Section 8.3 were also assessed. The impact of these differences was found to be much smaller than the statistical uncertainty of the measurement.

Summary

A summary of these uncertainty contributions is shown in Figures 8.13 to 8.15. Uncertainties on the top background estimation are the dominant source in both channels, highlighting the importance of the data-driven estimation to reduce these uncertainties compared to a full top cross-section modelling uncertainty. Energy scale and identification uncertainties are the dominant experimentally-driven sources in the electron channel, increasing as a function of invariant mass, while muon identification corrections are the dominant source in the muon channel. A similar dependence of the systematic uncertainties can be seen in the two-dimensional distributions, with the top-quark background estimation dominating more prominently the uncertainty at the edge of the $\cos\theta^*_{\text{CS}}$ distribution.

Total systematic uncertainties as low as 1% are achieved in the $m_{\ell\ell} < 200$ GeV range, region in which the systematic uncertainties are greater than the statistical ones. The uncertainties remain at the $\leq 3\%$ level across all distributions with $m_{\ell\ell} < 500$ GeV. The small number of events above this threshold drives the uncertainty, either as direct statistical uncertainty or reflected in the top quark background systematic uncertainty. Total uncertainties below 5% are obtained at $m_{\ell\ell} \leq 1$ TeV. The last bin of the single-differential distribution, $m_{\ell\ell} > 1.5$ TeV, first explored in this analysis, shows a 7% (10%) statistical uncertainty in the electron (muon) channel, while systematic uncertainties are 7.5% (8%).

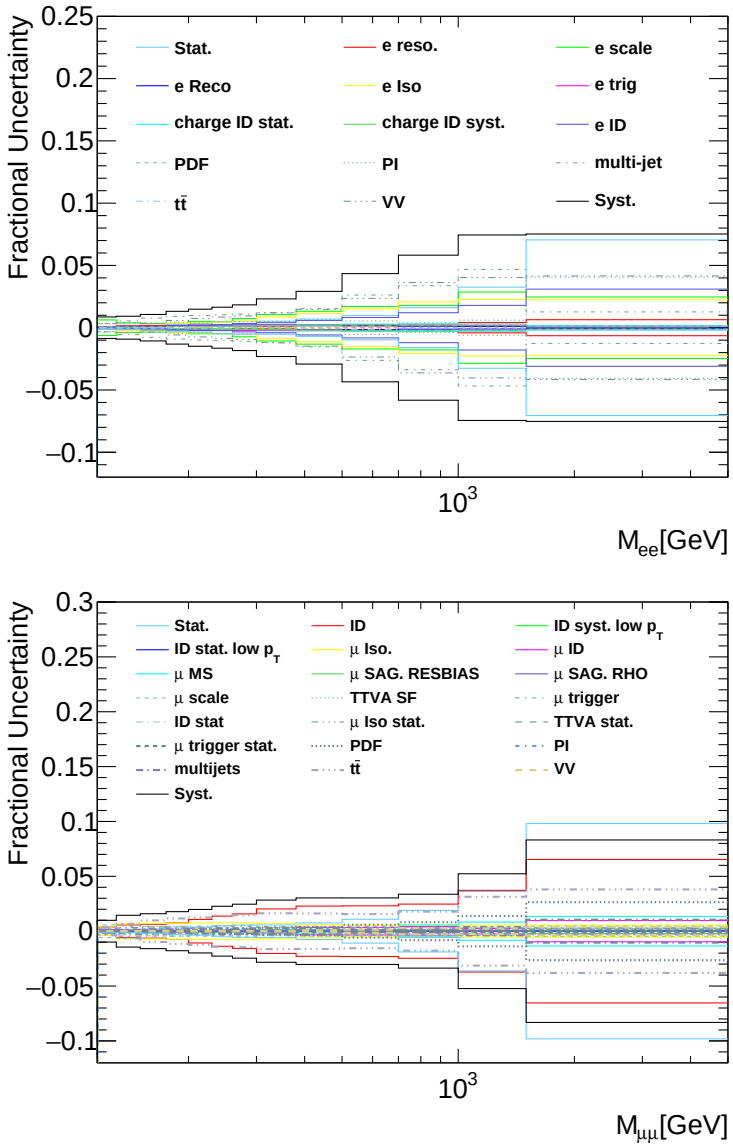
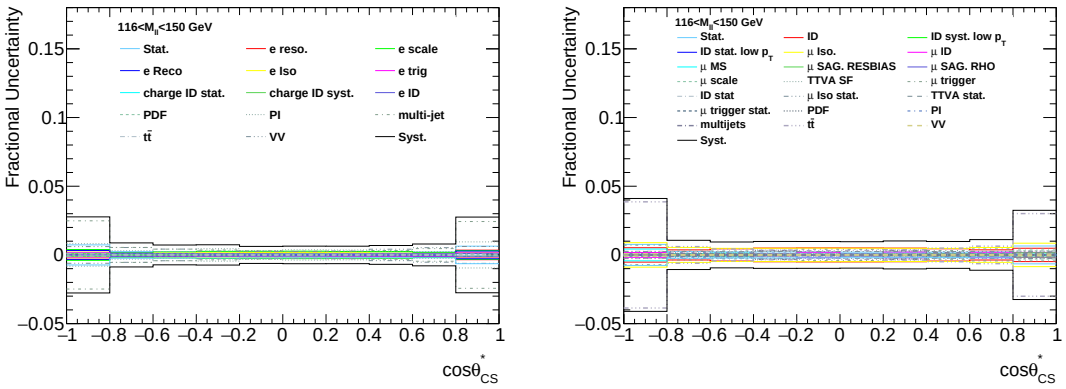
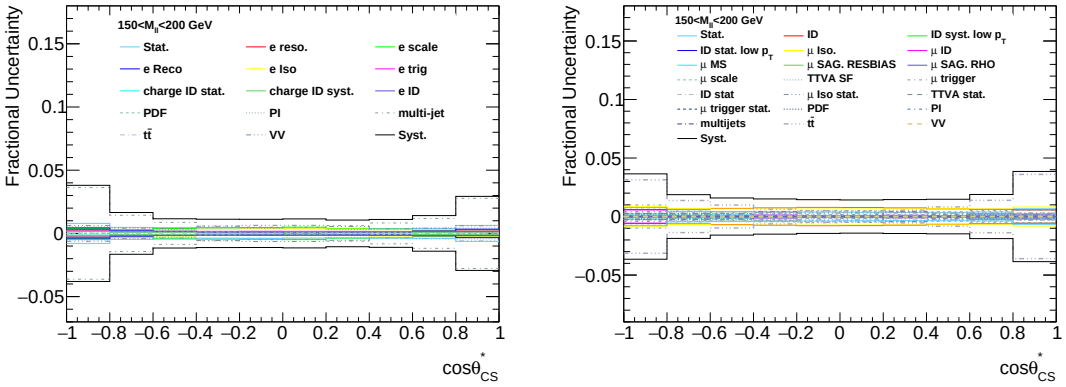


Figure 8.13: Systematic uncertainties against $m_{\ell\ell}$ for the dielectron (top) and dimuon (bottom) final states after the unfolding procedure. From Ref. [137].



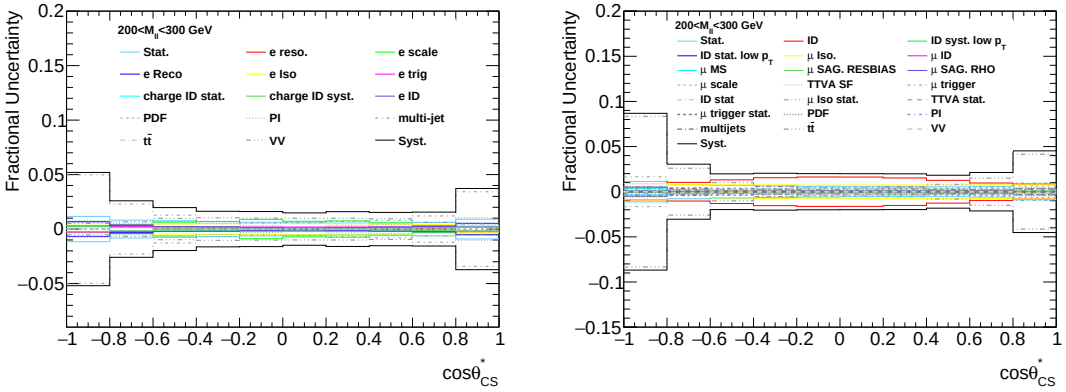
(a) $116 < m_{ee} < 150$ GeV

(b) $116 < m_{\mu\mu} < 150$ GeV



(c) $150 < m_{ee} < 200$ GeV

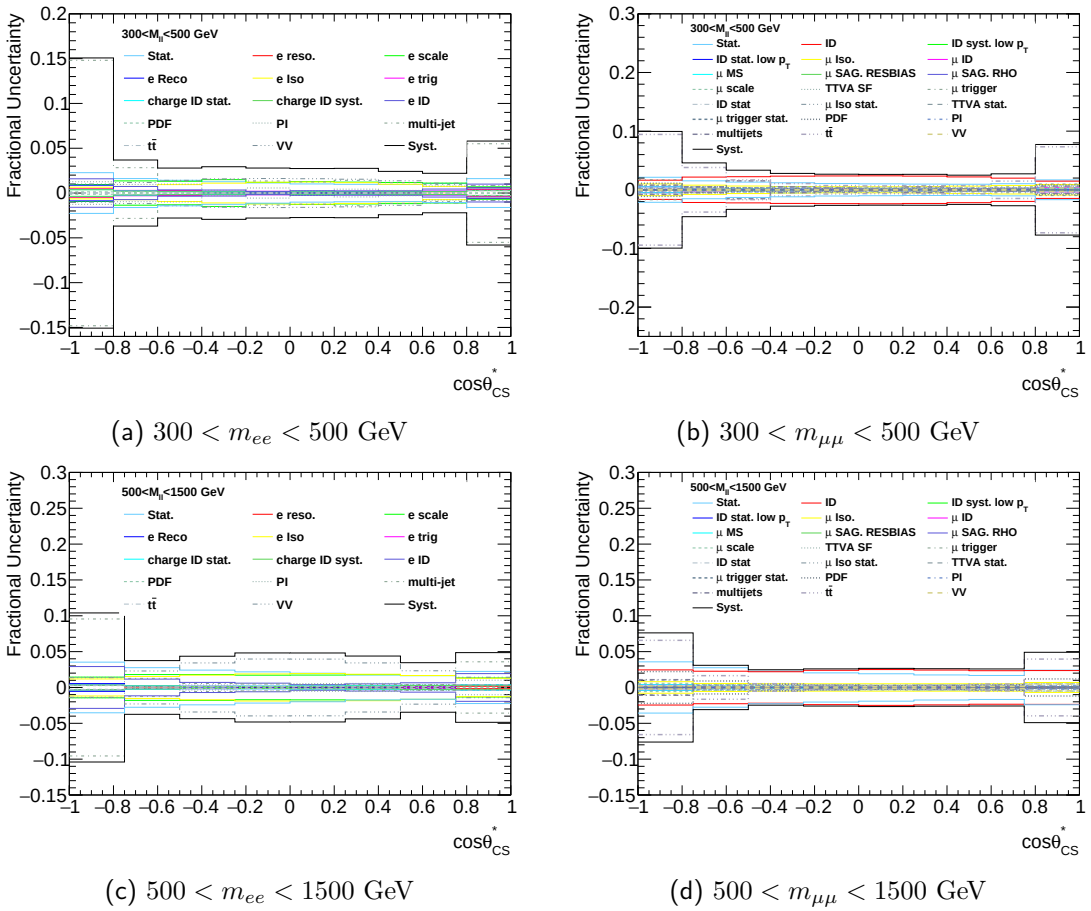
(d) $150 < m_{\mu\mu} < 200$ GeV



(e) $200 < m_{ee} < 300$ GeV

(f) $200 < m_{\mu\mu} < 300$ GeV

Figure 8.14: Systematic uncertainties on the double-differential cross-section measurement as a function of $\cos\theta^*_{CS}$ for the dielectron (left) and dimuon (right) final states after the unfolding procedure for $116 < m_{\ell\ell} < 300$ GeV. From Ref. [137].



(a) $300 < m_{ee} < 500$ GeV

(b) $300 < m_{\mu\mu} < 500$ GeV

(c) $500 < m_{ee} < 1500$ GeV

(d) $500 < m_{\mu\mu} < 1500$ GeV

Figure 8.15: Systematic uncertainties on the double-differential cross-section measurement as a function of $\cos\theta^*_{CS}$ for the dielectron (left) and dimuon (right) final states after the unfolding procedure for $m_{\ell\ell} > 300$ GeV. From Ref. [137].

8.6 Theoretical prediction and uncertainties

The nominal theoretical prediction for the Drell-Yan dilepton production in the analysis is given by the POWHEG $Z \to \ell\ell$ Monte Carlo, details of which can be found in Section 8.1. The samples are generated at NLO QCD and LO EW and reweighed to state-of-art predictions including NNLO QCD and NLO EW corrections. Both corrections are calculated independently and are combined using the *additive* approach [134]:

$$\sigma_{\text{NNLO QCD+NLO EW}} = \sigma_{\text{NNLO QCD}} \cdot \left(1 + \frac{\Delta\sigma_{\text{LO QCD+NLO EW}}}{\sigma_{\text{NNLO QCD}}}\right). \tag{8.16}$$

This approach assumes that the same NLO EW correction term (with the exception of QED FSR) needs to be applied for all orders of QCD, i.e. that the relative higher-order correction changes at each order of QCD. Alternatively, one may follow the so-called multiplicative approach:

$$\sigma_{\text{NNLO QCD,NLO EW}} = \sigma_{\text{LO}} \times \frac{\sigma_{\text{NNLO QCD}}}{\sigma_{\text{LO}}} \times \frac{\sigma_{\text{NLO EW+LO QCD}}}{\sigma_{\text{LO}}}, \tag{8.17}$$

assuming that the higher-order EW corrections are the same for all orders of QCD and can be determined based on LO QCD and then transferred to higher-order QCD. The additive approach will be taken as the nominal choice in the analysis, taking the differences between both approaches as an estimate of the uncertainty on the size of the electroweak correction.

Further theoretical uncertainties in the prediction arise from PDF uncertainties, the renormalisation and factorisation scale choices and the value of the strong coupling constant used to simulate the process. The nominal PDF choice used in the analysis is CT18ANNLO [135], for which the leading seven 90% CL variations were evaluated, as well as differences with respect to other modern PDFs such as NNPDF3.0 [145]. Results on the nominal theory prediction and uncertainties can be found in Ref. [137].

Although the nominal theoretical prediction includes state-of-art predictions and a careful assessment of uncertainties, the calculations were only performed against $m_{\ell\ell}$ and $y_{\ell\ell}$. Moreover, it was observed that the POWHEG Monte Carlo offers a poor modelling of the $\cos\theta^*_{\text{CS}}$ distribution in certain regions of the kinematic phase space, severely limiting the EFT interpretation of the double-differential cross-section measurement presented in Section 10.2. For the purpose of using this distribution for an EFT interpretation, a second theoretical prediction was obtained from the SHERPA sample first presented in Section 8.1. Differences between the nominal prediction and this alternative are shown in Figure 8.16. An invariant mass-dependent difference is observed between both predictions. The first significant difference is in the first invariant mass bin. However the range

$116 < m_{ee} < 120$ GeV is simply missing from the SHERPA sample due to technical reasons, causing the difference shown in the Figure. Additional mass-dependent differences arise from the application of NNLO QCD reweighting to the POWHEG prediction, while SHERPA sample only includes effects up to NLO QCD (and NLO EW, also using the additive approach). For this reason, the mass dependence of the POWHEG is transferred to the SHERPA sample as a bin-by-bin correction based on the ratio of the corresponding predictions in each mass range. The resulting double-differential distribution comparison can be found in Figure 8.17. A difference in the modelling of the $\cos\theta^*_{\text{CS}}$ distribution between both generators is observed in this figure, highlighting the need for an updated prediction with respect to the nominal if EFT fits are going to be performed using the double-differential measurement.

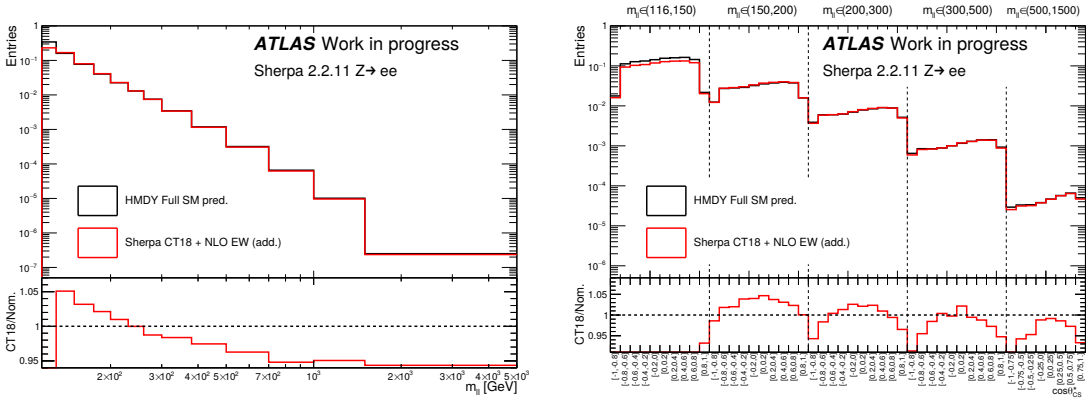


Figure 8.16: Theory predictions available for the single- (left) and double-differential (right) Drell-Yan cross-section against dilepton invariant mass and $\cos\theta^*_{\text{CS}}$.

The nominal distribution obtained in SHERPA uses the NNPDF3.0NNLO PDF set [145] with $\alpha_s = 0.118$, however different assumptions are available in the sample via event reweighing. Some of the available variations include the usage of different modern PDF sets, the 100 eigenvector variations of the NNPDF3.0 PDF set and modification of the α_s, renormalisation and factorization scale parameters, as well as the impact of different approaches to combine NLO EW corrections with higher-order QCD corrections. Figures 8.18 to 8.22 show the distributions obtained varying different theory assumptions, showing the ratio as an estimate of the uncertainty associated to each assumption.

Uncertainties associated to the variation of α_s are found to be 1% across most of the invariant mass spectrum, with little dependence against $\cos\theta^*_{\text{CS}}$. Similarly, the sum in quadrature of all eigenvector variations is found to be rather flat in most of the spectrum, though individual variations show dependence against either of the variables. When used in the interpretation presented in Section 10.2, differences with respect to each eigenvector are introduced as nuisance parameter, allowing to account for all possible distribution shape

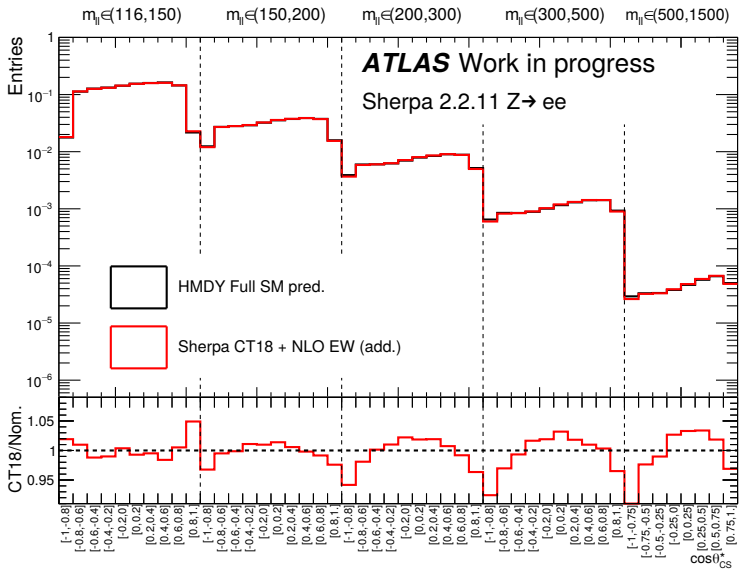


Figure 8.17: Theory predictions available for the double-differential Drell-Yan cross-section against dilepton invariant mass and $\cos\theta^*_{\mathrm{CS}}$. The SHERPA distribution is reweighed to POWHEG using the ratio between both generators in the corresponding mass range for each bin.

differences described in each PDF variations. Different PDF sets are found to show variations within 3% of the nominal set across all the dilepton invariant mass spectrum, with some of the PDF sets such as MSHT20NNLO introducing shape changes in the distribution across $\cos\theta^*_{\mathrm{CS}}$, as shown in Figure 8.19. Uncertainties arising from the variation of the renormalisation and factorization scales are found to be the biggest effects in the sample, reaching up to $\sim 5\%$ at high invariant mass ranges. These show little dependence with respect to $\cos\theta^*_{\mathrm{CS}}$ with the exception of the edge bins in the mass range $116 < m_{\ell\ell} < 150$ GeV, where the phase space is most restricted. It is worth noting that, since the SHERPA Monte Carlo generator includes both matrix element generation and parton shower simulation, modifications of QCD scales tend to have substantial impact in the distributions, as reflected in the results shown in Figure 8.21.

To study the effects of the combination of EW and QCD higher-order corrections an additional approach is included in the sample, namely the exponentiated approach. It consists of an intermediate assumption between the additive and multiplicative approaches, with higher order corrections being parametrised as:

$$\sigma_{\text{NNLO QCD+NLO EW}} = \sigma_{\text{NNLO QCD}} \cdot \exp(\Delta\sigma_{\text{LO QCD+NLO EW}}). \tag{8.18}$$

As shown in Figure 8.22, uncertainties associated to the variation of the higher-order combination procedure are rather small, with $< 0.5\%$ effects in most of the kinematic phase space.

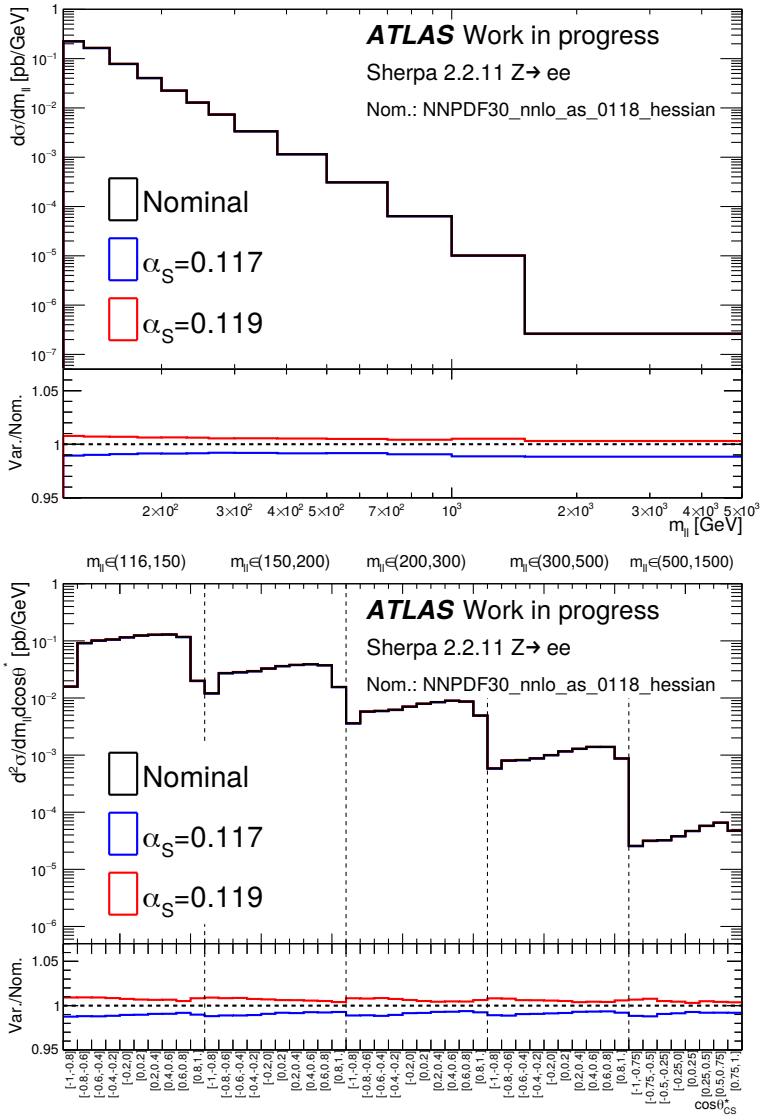


Figure 8.18: Cross-section predictions obtained varying α_s in the mass-rescaled SHERPA sample for the single- (top) and double-differential (bottom) Drell-Yan cross-section against dilepton invariant mass and $\cos\theta^*_{\mathrm{CS}}$. Ratio shown corresponds to an estimate of the uncertainty associated to this variation.

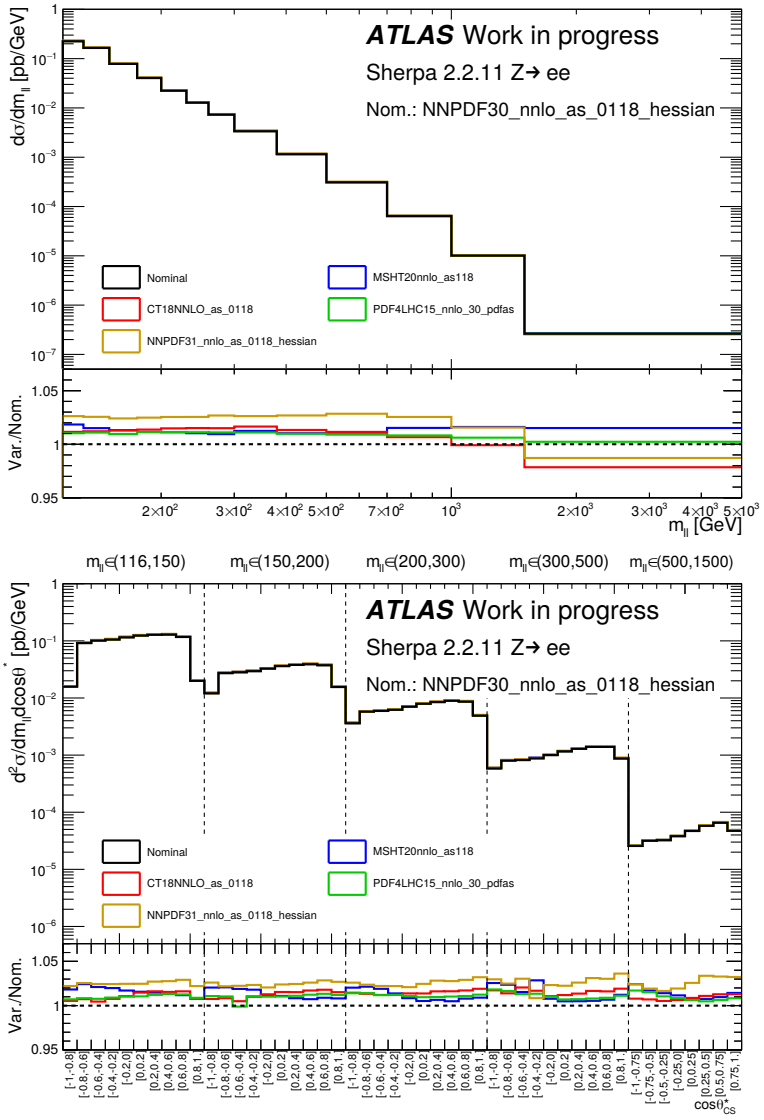


Figure 8.19: Cross-section predictions obtained varying the PDF set used in the mass-rescaled SHERPA sample for the single- (top) and double-differential (bottom) Drell-Yan cross-section against dilepton invariant mass and $\cos\theta^*_{\mathrm{CS}}$. Ratio shown corresponds to an estimate of the uncertainty associated to this variation.

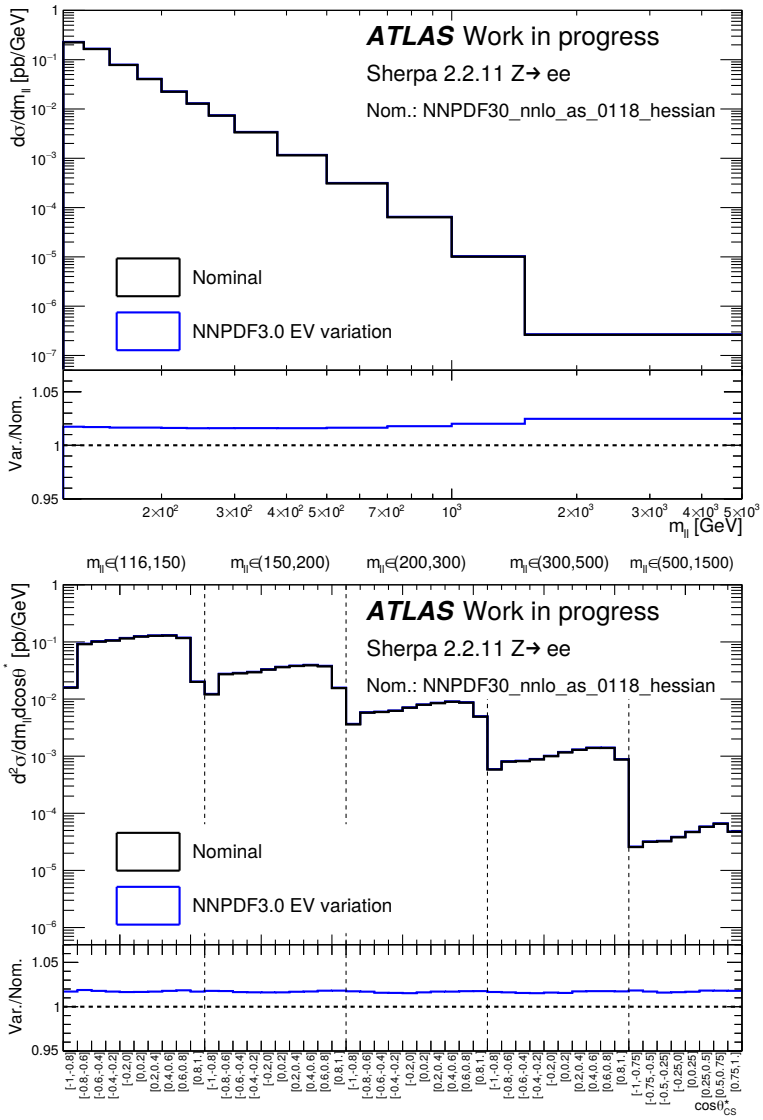


Figure 8.20: Cross-section predictions obtained varying the 100 eigenvectors in the NNPDF3.0 PDF set in the mass-rescaled SHERPA sample for the single- (top) and double-differential (bottom) Drell-Yan cross-section against dilepton invariant mass and $\cos\theta^*_{\mathrm{CS}}$. Distribution shown corresponds to the sum in quadrature of the differences between each eigenvector and the nominal PDF set. Ratio shown corresponds to an estimate of the uncertainty associated to this variation.

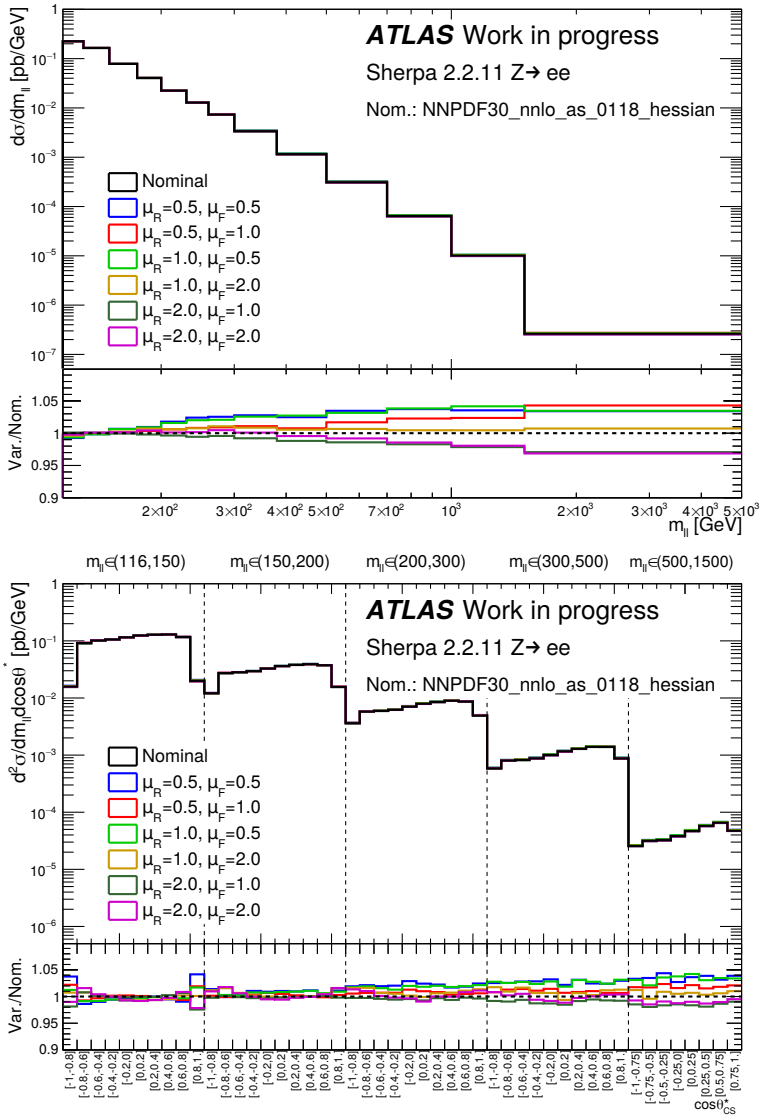


Figure 8.21: Cross-section predictions obtained varying the renormalisation (μ_R) and factorization (μ_F) scales in the mass-rescaled SHERPA sample for the single- (top) and double-differential (bottom) Drell-Yan cross-section against dilepton invariant mass and $\cos\theta^*_{\mathrm{CS}}$. Ratio shown corresponds to an estimate of the uncertainty associated to this variation.

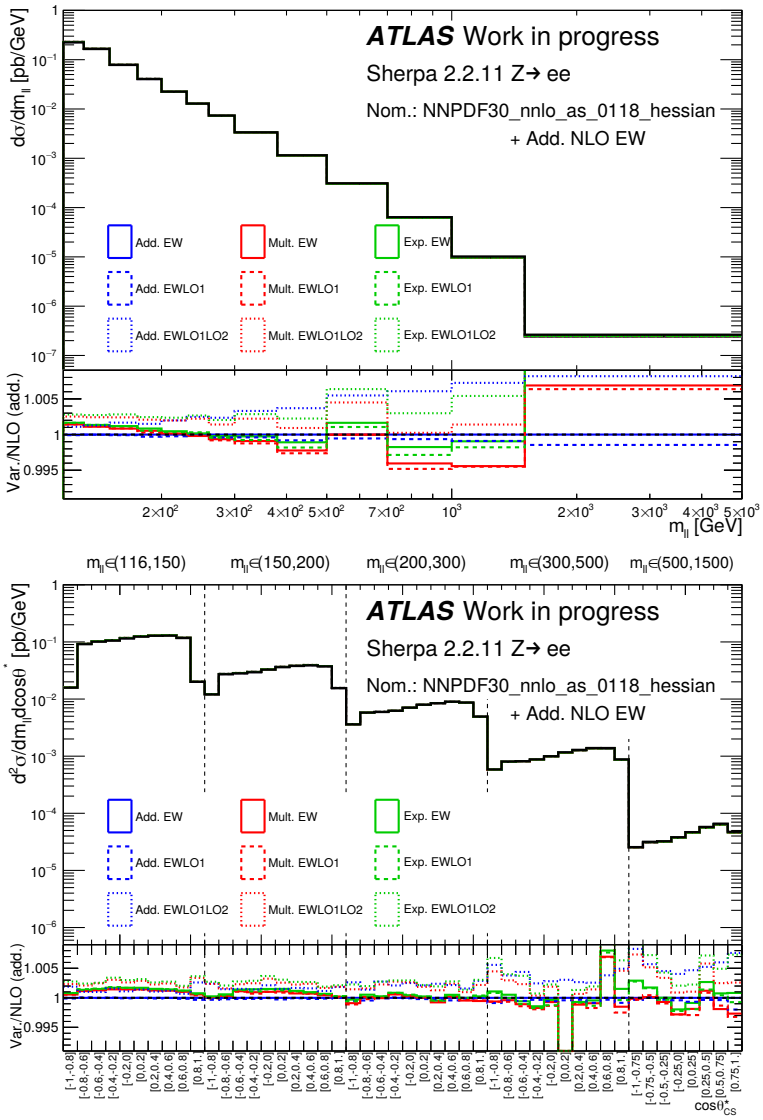


Figure 8.22: Cross-section predictions obtained using different higher-order correction combination schemes in the mass-rescaled SHERPA sample for the single- (top) and double-differential (bottom) Drell-Yan cross-section against dilepton invariant mass and $\cos\theta^*_{\mathrm{CS}}$. Ratio shown corresponds to an estimate of the uncertainty associated to this variation, using the additive approach as the nominal choice. The LO1 and LO1LO2 terms include subleading born corrections.

Chapter 9

Results

This Chapter presents the results on the unfolded Drell-Yan cross section, as well as an overview of the combination procedure of the electron and muon channel results. Additionally, the combined results of the cross-section are compared to different predictions to test the PDF sensitivity of the measurement at high dilepton invariant masses.

9.1 Combination procedure

The result of the cross section measurement in the electron and muon channels can be combined to reduce the statistical and systematic uncertainties on the measurement. This combination is performed using the HERAverager tool [160], first developed for the combination of deep inelastic scattering data [161]. The procedure combines the cross section in each bin i of each channel considered k (here $k = e, \mu$) introducing systematic uncertainties that are correlated between bins and an uncorrelated uncertainty, arising from the statistical uncertainty in data (δ_{stat}) and the uncorrelated components of the systematic uncertainties (δ_{unc}). Thus, a χ^2 function is defined by

$$\chi^2_{\text{tot}}(\bar{\sigma}, \boldsymbol{\theta}) = \sum_{k}^{N_{\text{chan}}} \sum_{i}^{N_{\text{bins}}} \frac{\left(\bar{\sigma}^i - \sum_{j=1}^{N_{\text{sys}}} \Gamma^i_{j,k}\theta_j - \sigma^i_k\right)^2}{\delta^2_{i,\text{stat}} + \delta^2_{i,\text{unc}}} + \sum_{j=1}^{N_{\text{sys}}} \theta_j^2, \tag{9.1}$$

where $\bar{\sigma}$ is the averaged cross section between both channels, σ_k is the cross section in each channel and θ_j are the shifts of the correlated uncertainty source j. A shift of $\theta_j = 1$ corresponds to a shift of the source j by 1σ. The factor $\Gamma^i_{j,k}$ represents the relative uncertainty of the source j in the bin i of the measurement in the channel k. The averaged cross section and the shift on each source of correlated uncertainty can be found by minimising the χ^2, fulfilling the following conditions:

$$\frac{\partial \chi^2_{\text{tot}}(\bar{\sigma}, \boldsymbol{\theta})}{\partial \bar{\sigma}} = 0, \quad \frac{\partial \chi^2_{\text{tot}}(\bar{\sigma}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = 0. \tag{9.2}$$

The minimisation procedure is described in detail in Ref. [160]. The systematic sources

considered are listed in Section 8.5. Due to the data-driven estimation of the top background in the $e\mu$ control region, all lepton reconstruction uncertainties are correlated between both channels.

9.2 Combined cross section results

Figure 9.1 shows the single-differential born-level cross sections in the electron, muon and combined channel, as well as a comparison between the results. The luminosity uncertainty is excluded from this result since it affects both channels in the same way. The pulls of the two individual measurements with respect to their combination is also shown, defined as the difference between the single-channel measurement and the combined result divided by the total uncertainty. As shown in the figure, both channels are in good agreement, with both measurements being found within 2σ of the combined result in all the invariant mass spectrum. The measured cross section presents the expected falling shape against dilepton invariant mass, ranging over five orders of magnitude in the mass range covered by the analysis. At low values of $m_{\ell\ell}$ the result of the combined measurement is dominated by the experimental systematic uncertainties, with a 0.52% contribution in the first mass bin, where a 0.22% statistical uncertainty is found. For $m_{\ell\ell} \geq 700$ GeV the statistical uncertainty is at the level or above of the systematic uncertainty, reaching a 9% contribution in the last bin ($m_{\ell\ell} > 1500$ GeV), where a 5% systematic uncertainty is found. Pulls on the systematic uncertainties, as well as a detailed breakdown of all contributions in each bin of the measurement can be found in Ref. [137]. The χ^2 per degree of freedom is found to be 9.63/13=0.74 ($p = 0.276$) for the single-differential cross-section.

Figure 9.2 shows the measurement of the double-differential cross-section with respect to $m_{\ell\ell}$ and $\cos\theta^*_{\text{CS}}$ in the electron, muon and combined channels. The distributions also show good agreement in this kinematic variable, with pulls in both measurements found within 2σ in all the spectrum except the $-0.8 < \cos\theta^*_{\text{CS}} < -0.6$ range within $116 < m_{\ell\ell} < 150$ GeV, where a pull of 3.4σ is found. The distributions show the effect of the forward-backward asymmetry in the Drell-Yan cross-section, showing an increase in the cross-section at $\cos\theta^*_{\text{CS}} > 0$ with respect to $\cos\theta^*_{\text{CS}} < 0$ as $m_{\ell\ell}$ increases. A total χ^2 per degrees of freedom of 61.0/56=1.09 ($p = 0.699$) is observed for the cross-section measured against $m_{\ell\ell}$ and $\cos\theta^*_{\text{CS}}$ across all invariant mass bins.

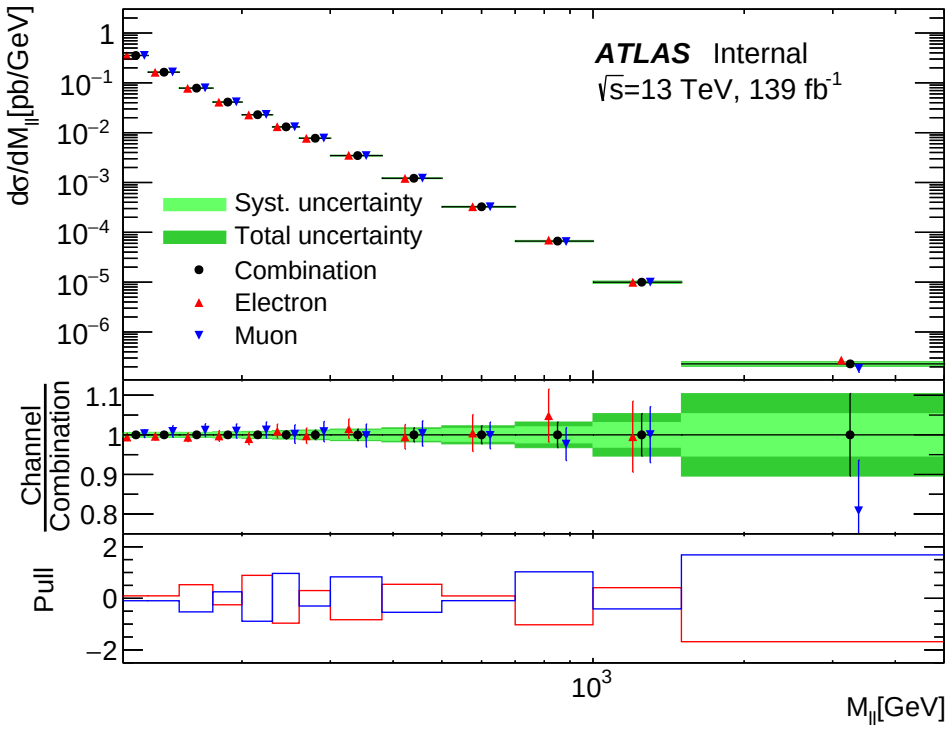
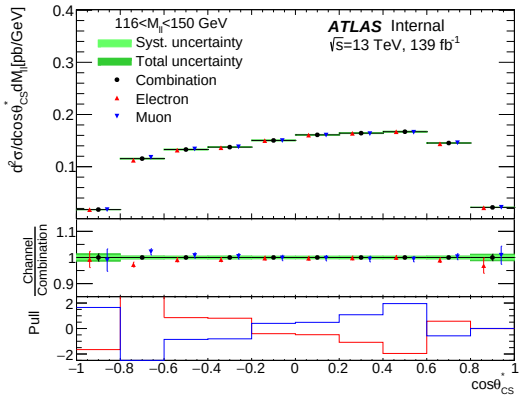
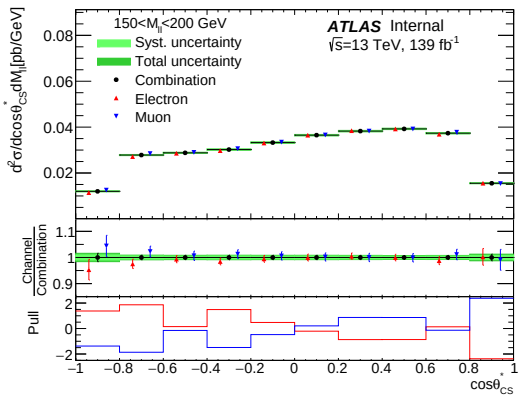


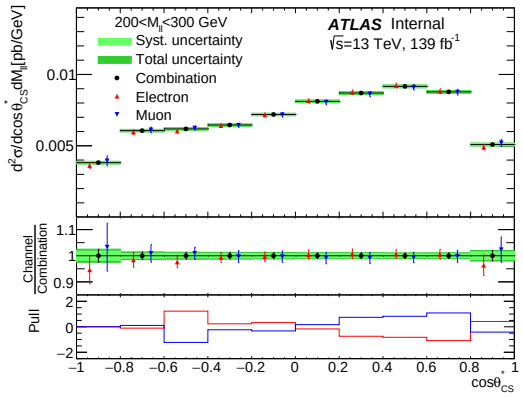
Figure 9.1: Combined born-level single-differential cross-section against $m_{\ell\ell}$ with statistical and systematic uncertainties, excluding the uncertainty on the luminosity. A comparison with the electron (red) and muon (blue) channels is also presented, including the pulls of each measurement with respect to the combined result. From Ref. [137].



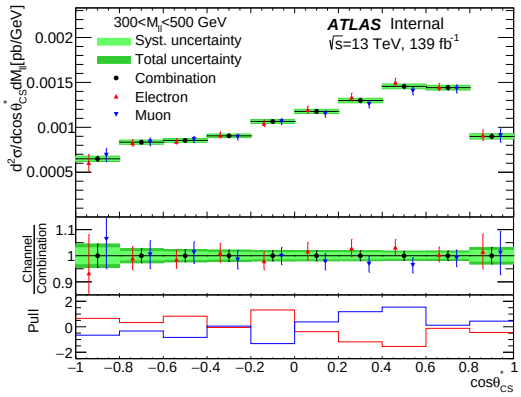
(a) $116 < m_{\ell\ell} < 150$ GeV



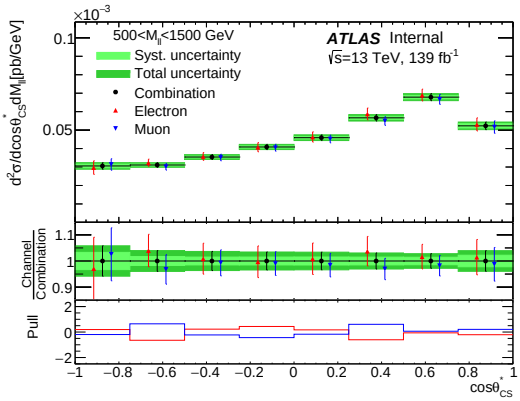
(b) $150 < m_{\ell\ell} < 200$ GeV



(c) $200 < m_{\ell\ell} < 300$ GeV



(d) $300 < m_{\ell\ell} < 500$ GeV



(e) $500 < m_{\ell\ell} < 1500$ GeV

Figure 9.2: Combined born-level double-differential cross-section against $\cos\theta^*_{CS}$ with statistical and systematic uncertainties, excluding the uncertainty on the luminosity. A comparison with the electron (red) and muon (blue) channels is also presented, including the pulls of each measurement with respect to the combined result. From Ref. [137].

9.3 Comparison to theoretical predictions

The combined fiducial cross-section is compared to state-of-the-art predictions, introduced in Section 8.6. The comparison of the single-differential cross-section measurement against $m_{\ell\ell}$ is shown in Figure 9.3. The prediction of the cross-section using the CT18ANNLO PDF set [135] is few % below the measured distributions in most of the invariant mass spectrum, with the exception of the last bin covering masses beyond 1500 GeV, and always compatible with data within 1σ. The uncertainty of the prediction is larger than the uncertainty on the measurement, hinting at the fact that the data should be able to further constrain the theory prediction. When compared to different predictions using various PDF sets, the behaviour depends on the set: when using HERAPDF2.0 [162] the prediction overshoots the measured distribution in the full mass range; when using MSHT20NNLO [51], NNPDF3.1 [52], CT14NNLO [53] or ABMP16 [163] the predictions fall a few % below data at $m_{\ell\ell} < 1000$ GeV, but above this all predictions estimate a larger cross-section than the one measured. Though not shown for visibility, all predictions are found within the 68% CL intervals of each other, agreeing with data as well. The spread between different PDF sets at low invariant mass is larger than the uncertainty of the measurement, highlighting the sensitivity of this measurement to PDFs and the potential to further constrain them.

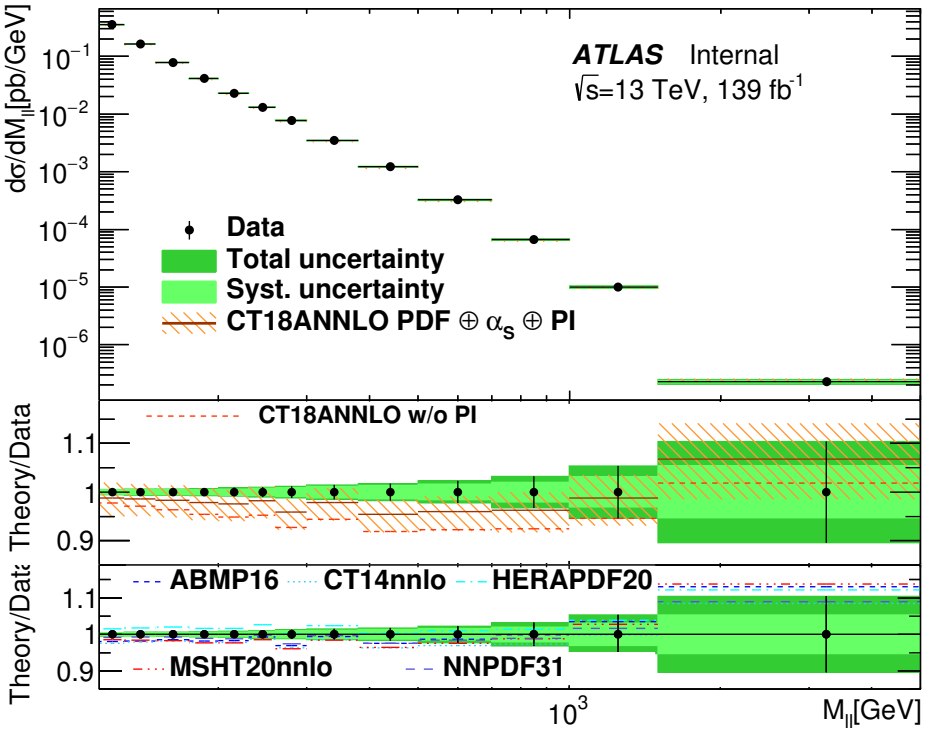


Figure 9.3: Combined born-level single-differential cross-section against $m_{\ell\ell}$ with statistical and systematic uncertainties, excluding the uncertainty on the luminosity. A comparison with NNLO QCD+NLO EW predictions using the CT18ANNLO PDF, where the uncertainty band corresponds to the combined 68% CL PDF, α_s, and photon-induced (PI) uncertainties. The middle ratio panel shows the ratio between said prediction and data, both including (solid line) and not including (dashed line) the PI contribution. The bottom panel shows the ratio with respect to predictions using different modern PDFs. From Ref. [137].

Chapter 10

Interpretation

This Chapter contains an overview of the different interpretations of the results presented in Chapter 9, both in the context of Lepton Flavour Universality (LFU) tests and Effective Field Theory (EFT) interpretations. See Chapter 1 for a breakdown of my personal contributions to the analysis.

10.1 Lepton Flavour Universality tests

Although the Standard Model predicts all lepton generations to couple equally to all SM particles, findings by the LHCb [164] and g-2 [165] collaborations suggest deviations in the lepton sector. As first suggested in Ref. [166] one way to search for SM deviations is to measure the LFU ratio, defined as:

$$R_{\mu\mu/ee} = \frac{\mathrm{d}\sigma_{\mu\mu}/\mathrm{d}X}{\mathrm{d}\sigma_{ee}/\mathrm{d}X}, \tag{10.1}$$

where X is the observable of interest. The Standard Model predicts this ratio to be exactly unity for any value of the dilepton invariant mass, since both muons and electrons couple equally to the Z and γ^* bosons. Measurements at LEP [167] reported excellent agreement between both channels at the Z-pole, with a measured ratio of the leptonic partial widths of 1.0009 ± 0.0028, however the measurement by CMS of the ratio showed a slight excess in the electron channel at $m_{\ell\ell} > 1.5$ TeV [29]. Figure 10.1 shows the measurement of $R_{\mu\mu/ee}$ as a function of $m_{\ell\ell}$, showing good agreement with Standard Model in most of the spectrum with the exception of the range $1500 < m_{\ell\ell} < 5000$ GeV, where a 2.3σ deviation is found. Though not statistically significant in itself, the observation is consistent with the results published by CMS. For the results on the tests using the double-differential cross-section measurements see Ref. [137].

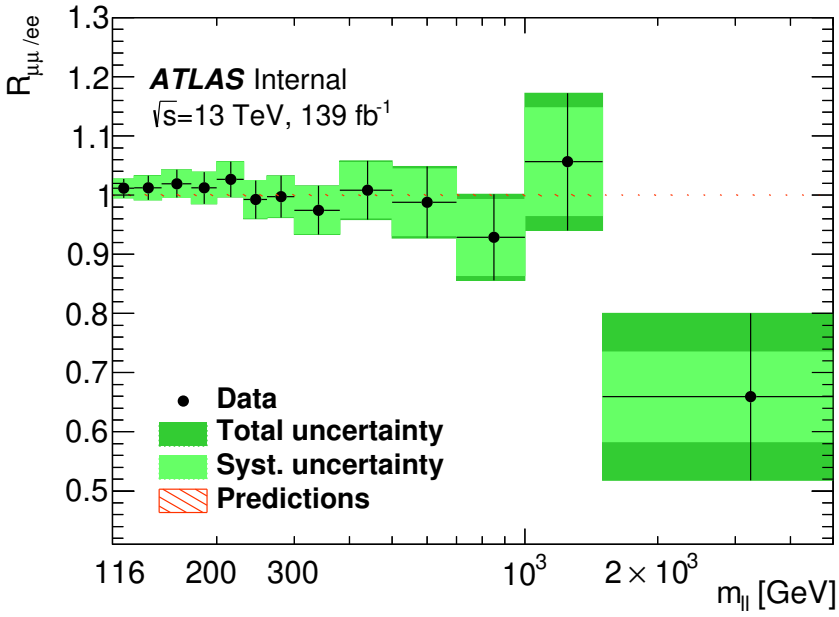


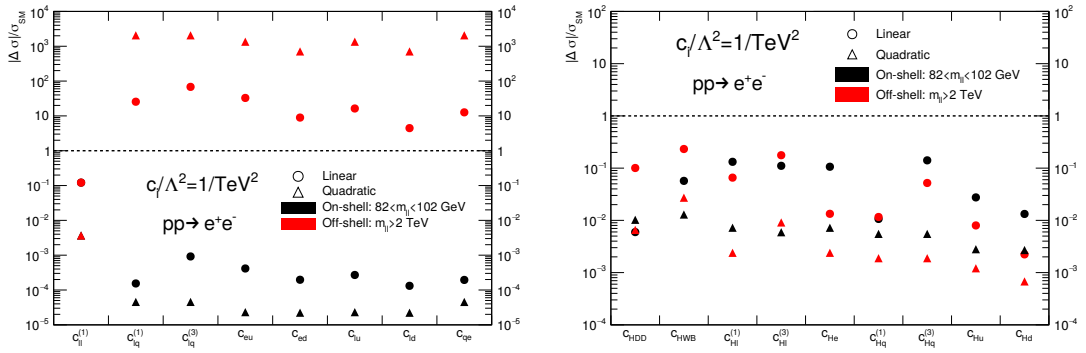
Figure 10.1: Measurement of $R_{\mu\mu/ee}$ as a function of $m_{\ell\ell}$. From Ref. [137].

10.2 Effective Field Theory interpretation

As introduced in Section 2.3.1, the Standard Model Effective Field Theory (SMEFT) framework is a generalization of the Standard Model which adds higher-dimensional operators acting at a high energy scale Λ. When adding these contributions, the Drell-Yan cross-section would change, meaning that the measured cross-section can be used to constrain the Wilson coefficients that parametrise the effects of these operators.

10.2.1 Operator sensitivity

This interpretation uses the $U(3)^5$ flavour symmetry assumption (as explained in Section 2.3.1) reduce the number of free parameters in the theory. The first step in the interpretation was to study which operators change the $pp \to \ell\ell$ cross-section and should be further studied. To do so, the processes $pp \to e^+e^-$ and $pp \to \mu^+\mu^-$ were generated with MadGraph (interfaced with Pythia for parton shower simulation) in the SMEFT framework [42], enabling one of the higher-dimensional operators at a time and rejecting those for which no contribution was found. The size of the EFT effects for the remaining operators was estimated both on-shell ($82< m_{\ell\ell} <102$ GeV) and off-shell ($m_{\ell\ell} >2$ TeV). Figure 10.2 shows the relative effect on the fiducial cross section of the operators we found our final state to be sensitive to in both the dielectron and dimuon decay channels. Some of these effects greatly exceed the SM expectations (e.g. $\times 1000$ cross-section introduced by the quadratic terms of some operators off-shell). This is due to the coupling chosen for this test, c_i/Λ=1 TeV^{-2}, and indicates that limits set on these operators will be several orders smaller than unity.



(a) 4-fermion operators

(b) Vff operators

Figure 10.2: Relative effect on the SM fiducial cross-section with respect to the SM $pp \to e^+e^-$ cross-section of the considered EFT operators. The effect of each operator is separated into the contribution of its corresponding linear and quadratic terms.

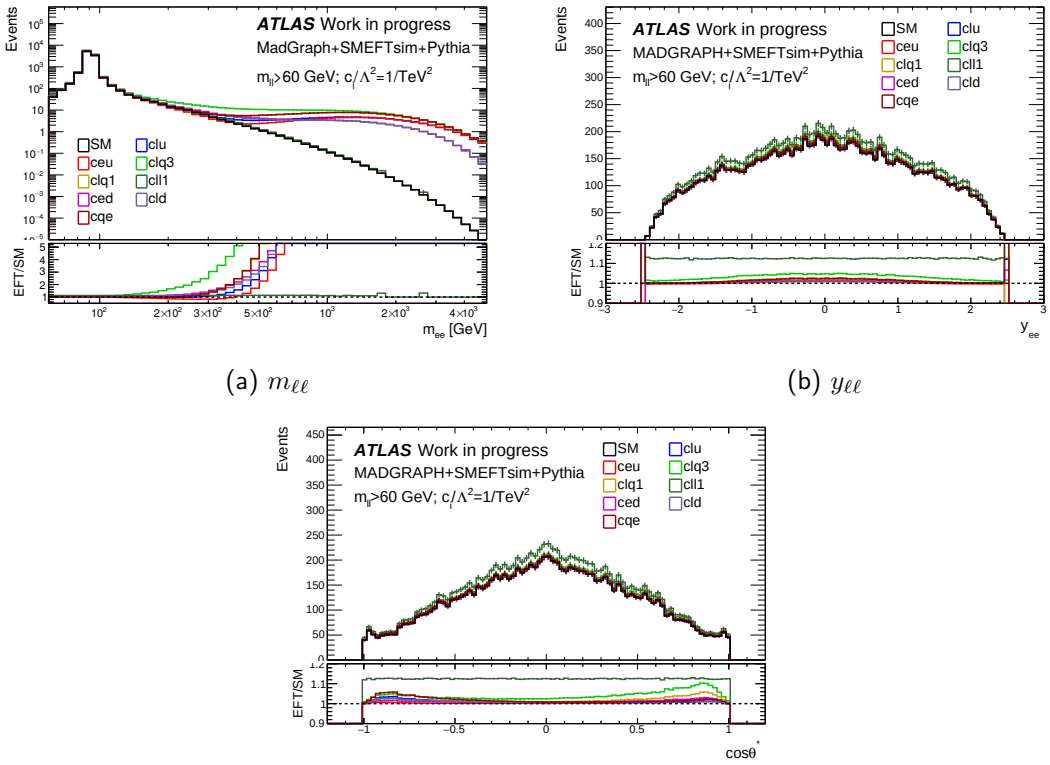
The dilepton final state is found to be sensitive to the EFT operators shown in Tables 10.1 and 10.2. The operators have been broadly divided between those affecting fermion couplings, namely *4-fermion* operators, and those modifying the couplings involving vector bosons, referred to as Vff operators. The effects of the 4-fermion operators over different kinematic distributions can be found in Figures 10.3 and 10.4, evaluated at $m_{\ell\ell} > 60$ GeV and $m_{\ell\ell} > 300$ GeV respectively. Several features can be inferred from these figures: the effect of the SMEFT operators enhances the Drell-Yan cross-section at high invariant masses; the effects of different operators at high (> 300 GeV) invariant masses modify the shape of the $\cos\theta^*_{\text{CS}}$ distribution in different ways; and the effects on $y_{\ell\ell}$ distribution result in an enhancement of the central rapidity cross-section at high dilepton invariant masses. The EFT interpretation presented in following sections will use the single- and double-differential cross-section against $m_{\ell\ell}$ and $\cos\theta^*_{\text{CS}}$. Figures 10.5 and 10.6 show the effects induced by the linear and quadratic terms of all operators with respect to the SM cross-section in each bin of the measurement kinematic phase space. 4-fermion operators introduce a large increase in cross-section at high $m_{\ell\ell}$. The shape differences in the $\cos\theta^*_{\text{CS}}$ spectra introduced by each operator are shown in Figure 10.6. The effect of Vff operators is found to be smaller than the one induced by 4-fermion operators, although shape differences are also observed in the $\cos\theta^*_{\text{CS}}$ distribution.

Wilson coefficient	Operator
cll1	$(\bar{l}\gamma_\mu l)(\bar{l}\gamma^\mu l)$
clq1	$(\bar{l}\gamma_\mu l)(\bar{q}\gamma^\mu q)$
clq3	$(\bar{l}\gamma_\mu \tau^I l)(\bar{q}\gamma^\mu \tau^I q)$
ceu	$(\bar{e}\gamma_\mu e)(\bar{u}\gamma^\mu u)$
ced	$(\bar{e}\gamma_\mu e)(\bar{d}\gamma^\mu d)$
clu	$(\bar{l}\gamma_\mu l)(\bar{u}\gamma^\mu u)$
cld	$(\bar{l}\gamma_\mu l)(\bar{d}\gamma^\mu d)$
cqe	$(\bar{q}\gamma_\mu q)(\bar{e}\gamma^\mu e)$

Table 10.1: Wilson coefficients and corresponding 4-fermion operators considered in this analysis. The fields $\{q, \ell\}$ are the SM quark and lepton left-handed fields respectively, and $\{e, u, d\}$ are the right-handed fields. τ^I are the $SU(2)_{\text{L}}$ SM generators, the Pauli matrices. For further information see Ref. [42].

Wilson coefficient	Operator
cHD	$(H^\dagger D^\mu H)^*(H^\dagger D_\mu H)$
cHWB	$H^\dagger \tau^I H W^I_{\mu\nu} B^{\mu\nu}$
cHl1	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{l}\gamma^\mu l)$
cHl3	$(H^\dagger i\overleftrightarrow{D}^I_\mu H)(\bar{l}\tau^I\gamma^\mu l)$
cHe	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{e}\gamma^\mu e)$
cHq1	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{q}\gamma^\mu q)$
cHq3	$(H^\dagger i\overleftrightarrow{D}^I_\mu H)(\bar{q}\tau^I\gamma^\mu q)$
cHu	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{d}\gamma^\mu d)$
cHd	$(H^\dagger i\overleftrightarrow{D}_\mu H)(\bar{u}\gamma^\mu u)$

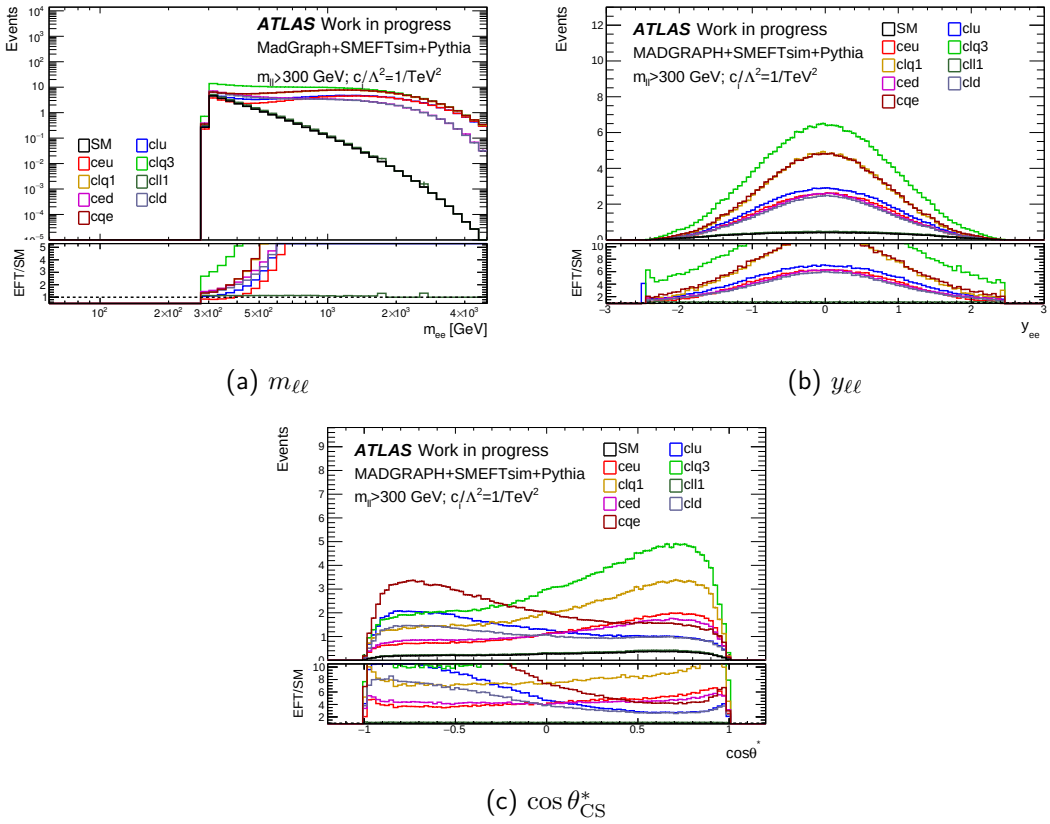
Table 10.2: Wilson coefficients and corresponding Vff operators considered in this analysis. The fields $\{q, \ell\}$ are the SM quark and lepton left-handed fields respectively, and $\{e, u, d\}$ are the right-handed fields. $W_{\mu\nu}$ and $B_{\mu\nu}$ are the SM gauge fields. τ^I are the $SU(2)_\mathrm{L}$ SM generators, the Pauli matrices. H corresponds to the SM Higgs scalar doublet. $H^\dagger i\overleftrightarrow{D}_\mu H = H^\dagger i D^\mu H - (iD^\mu H^\dagger)H$. For further information see Ref. [42].



(a) $m_{\ell\ell}$ (b) $y_{\ell\ell}$

(c) $\cos\theta^*_{\mathrm{CS}}$

Figure 10.3: Different kinematic distributions in the dielectron channel obtained by adding different 4-fermion contributions (both linear and quadratic terms) to the SM prediction. All couplings are set to c_i/Λ=1 TeV^{-2}. The dielectron mass cut is set at 60 GeV.



(a) $m_{\ell\ell}$

(b) $y_{\ell\ell}$

(c) $\cos\theta^*_{\mathrm{CS}}$

Figure 10.4: Different kinematic distributions in the dielectron channel obtained by adding different 4-fermion contributions (both linear and quadratic terms) to the SM prediction. All couplings are set to c_i/Λ=1 TeV^{-2}. The dielectron mass cut is set at 300 GeV.

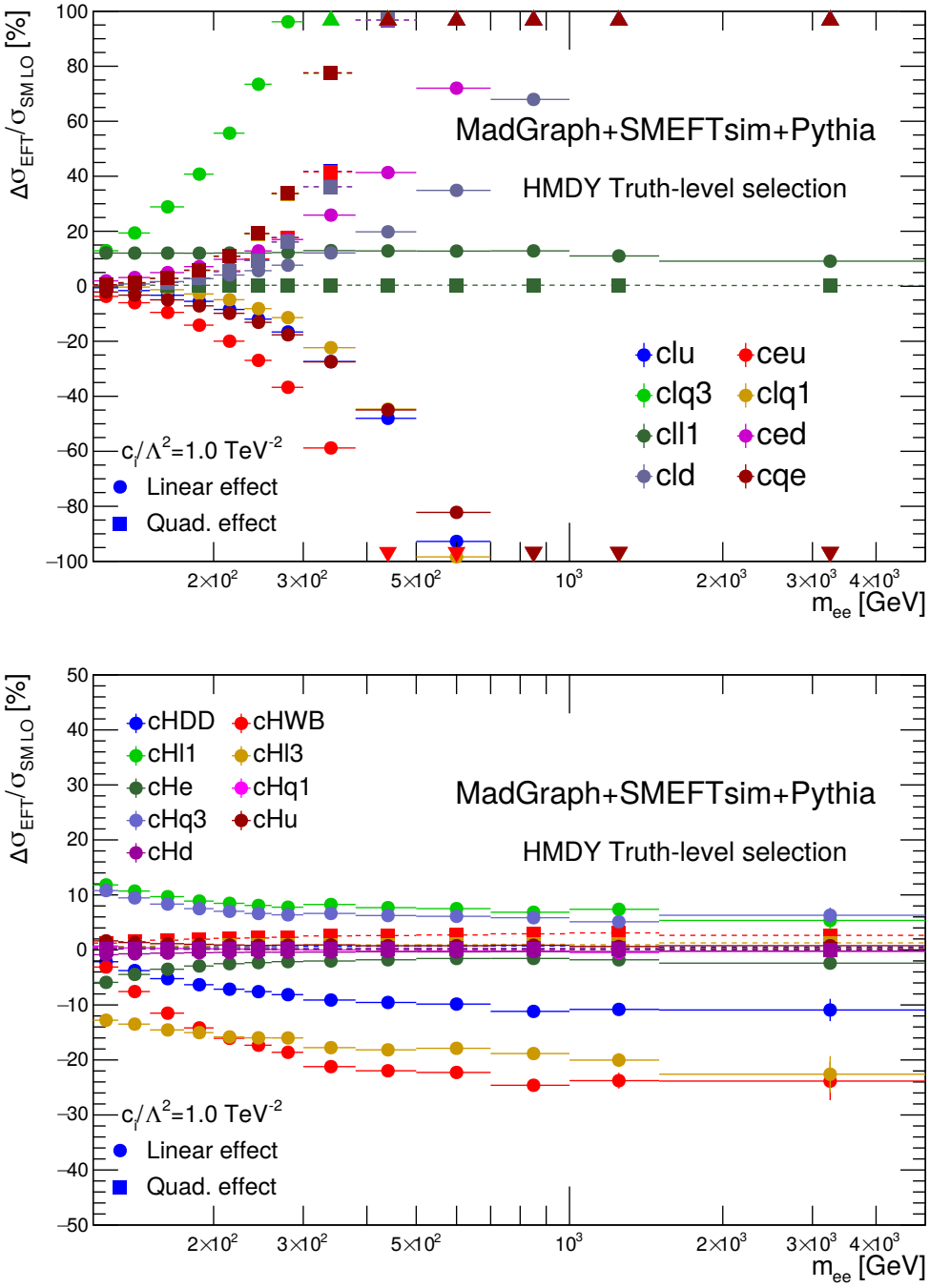


Figure 10.5: Ratio between EFT and SM LO distributions for 4-fermion (top) and Vff (bottom) operators against $m_{\ell\ell}$. All couplings are set to c_i/Λ=1 TeV^{-2}.

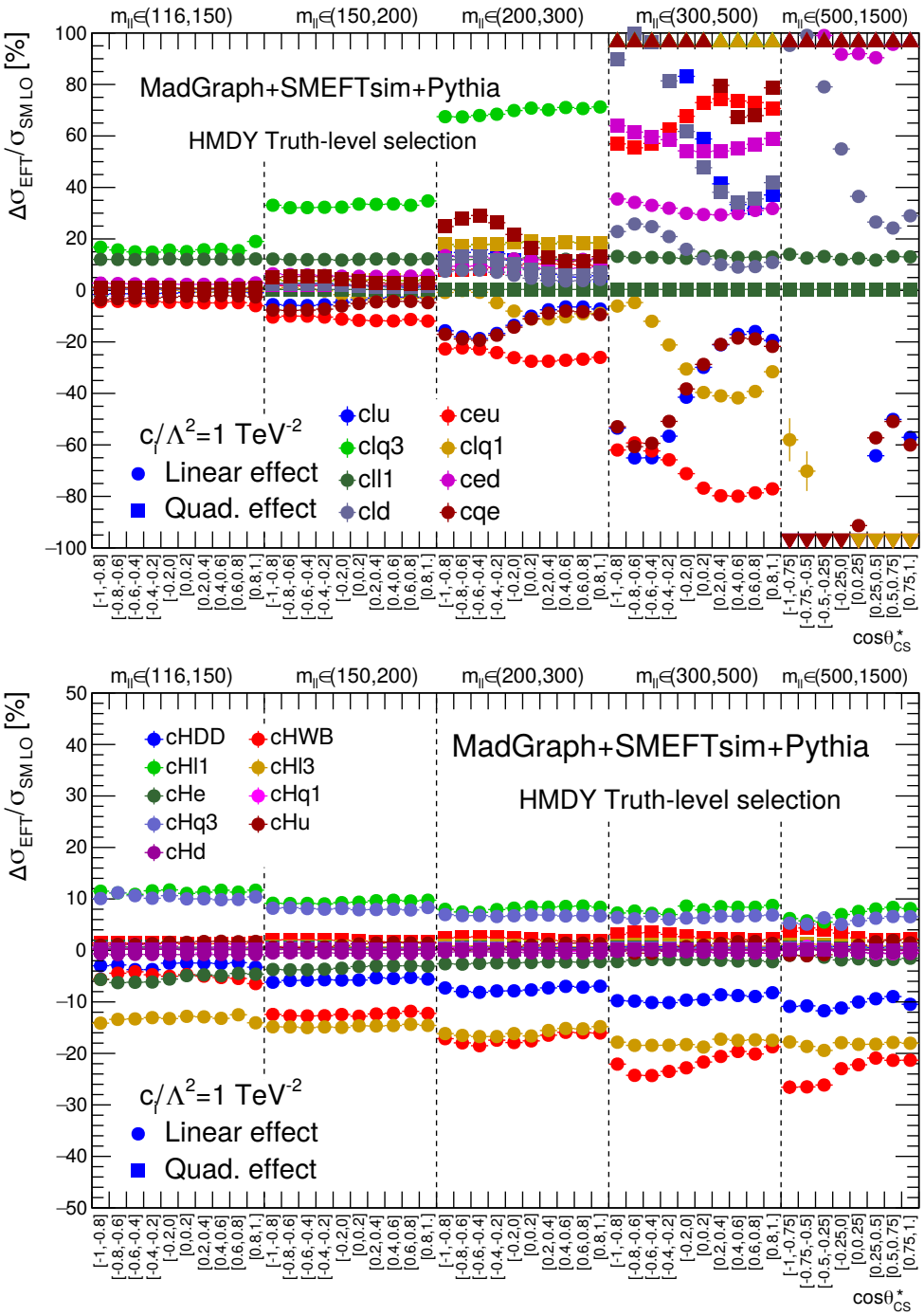


Figure 10.6: Ratio between EFT and SM LO distributions for 4-fermion (top) and Vff (bottom) operators against $m_{\ell\ell}$ and $\cos\theta^*_{\text{CS}}$. All couplings are set to c_i/Λ=1 TeV^{-2}.

10.2.2 Fitting framework

The EFT fitting framework used here was adapted from a standard ATLAS EFT fitting tool [168]. The framework uses ROOFit to compare a combination of SM prediction and EFT signal (linear and quadratic terms) with unfolded measurements to extract limits on the Wilson coefficients of the operators considered. The statistical model that the framework uses is a multivariate Gaussian that represents the measurement's uncertainties. The likelihood used by the model can be written as:

$$L = \frac{1}{\sqrt{(2\pi)^{N_{\text{bins}}}\,|\boldsymbol{\Sigma}|}} \exp\left\{-\frac{1}{2}\left[\vec{x}-\vec{\mu}(\vec{\theta})\right]^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\left[\vec{x}-\vec{\mu}(\vec{\theta})\right]\right\} \times \prod_i \theta_i, \tag{10.2}$$

where $\boldsymbol{\Sigma}$ represents the covariance matrix of the measurement (built from the bin-by-bin correlation of the different uncertainty sources), $\vec{x}$ is the measured distribution and $\vec{\mu}$ is the predicted distribution. The theoretical uncertainties on the prediction are included as nuisance parameters $\vec{\theta}$, where a shift of $\theta_i = 1$ corresponds to a shift of the source j by 1σ, following the procedure used in other ATLAS analyses [169–171]. The prediction in each bin is the combination of both the SM expected contribution (b_i), plus the linear (s_{lin}) and quadratic (s_{quad}) terms of the EFT parameters under study:

$$\mu_i = b_i + c \cdot s_{\text{lin},i} + c^2 \cdot s_{\text{quad},i}, \tag{10.3}$$

where c represents the Wilson coefficient of the operator being considered. However, EFT signals are computed only at LO QCD+EW level, so the prediction is modified to account for higher order corrections in the following manner:

$$\mu_i = b_i^{\text{best}} \cdot \left(1 + c\frac{s_{\text{lin},i}}{b_i^{\text{LO}}} + c^2\frac{s_{\text{quad},i}}{b_i^{\text{LO}}}\right), \tag{10.4}$$

where b^{best} represents the best available SM prediction for the process being studied, including all higher-order corrections available (as presented in Section 8.6), and b^{LO} is a SM prediction computed at leading-order, using the same generator used to simulate the EFT signal. The ratio between the LO SM prediction and the prediction including higher order corrections (namely, the $b^{\text{LO}}/b^{\text{best}}$ ratio), is shown in Figure 10.7, found to be $\sim 30\%$.

The covariance matrix ($\boldsymbol{\Sigma}$) used in Equation 10.2 is built adding the different uncertainty sources in the following way:

$$\Sigma_{ik} = \sum_j^{N_{\text{sys}}} \Gamma_i^j \Gamma_k^j + \delta_{ik}\Delta_i^2, \tag{10.5}$$

where j goes through the different uncertainty sources and i, k represent the different bins in the measurement. Γ_i^j therefore represents the systematic uncertainty in bin i

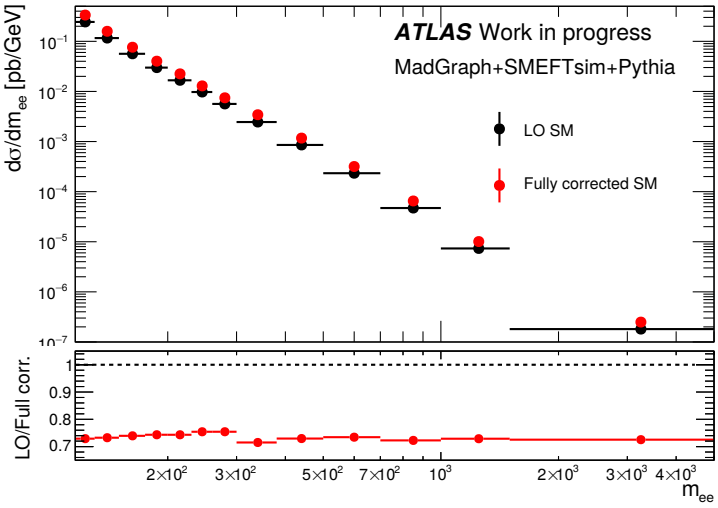


Figure 10.7: Dilepton invariant mass SM predictions calculated at LO and including higher order corrections (top). Ratio plot (bottom) corresponds to the correction factor applied in Equation 10.4.

arising from the source j. Δ_i represents the statistical uncertainty in bin i, assumed to be uncorrelated (hence δ_{ik}, statistical contribution only affects the diagonal terms). In the case of asymmetric systematic uncertainties, the contributions are symmetrised as:

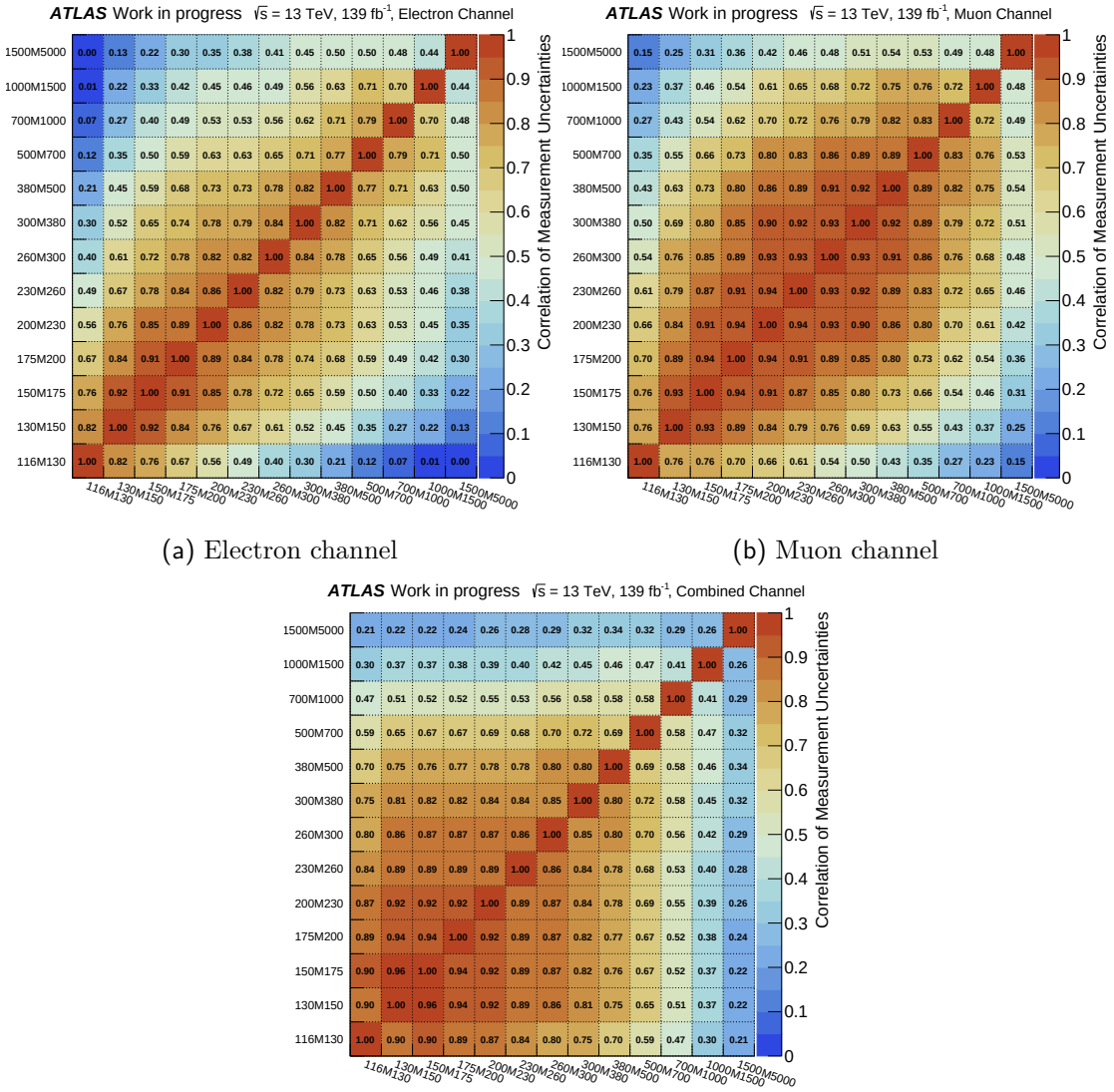
$$\Gamma_i = \frac{u_i^{\text{low}} - u_i^{\text{up}}}{2} \tag{10.6}$$

where u_i^{low} (u_i^{up}) represents the lower (upper) limit of the uncertainty in bin i. Figures 10.8 and 10.9 show the correlation matrices $\mathbf{\Sigma}$ in each of the channels studied in the analysis for the single-differential and double-differential measurements respectively. Large correlations between bins are found in the $m_{\ell\ell} < 300$ GeV region of the analysis, where statistical uncertainties are very small, causing the diagonal term Δ_i^2 in Equation 10.5 to have a much smaller contribution than the correlated sources. High correlations are particularly prominent in the muon channel, where they're observed up to the $m_{\ell\ell} \sim 1$ TeV range.

Experimental systematic uncertainties used in the fits are as described in Section 8.5, including a global 1.7% uncertainty corresponding to the luminosity measurement. Theoretical uncertainties on the fully corrected SM prediction are taken as shown in Section 8.6, including PDF, scale and α_S uncertainties. Moreover, if the relative statistical uncertainty in a bin of a template corresponding to an EFT effect is above 10%, a nuisance parameter is added that parametrizes this uncertainty in the fit.

Limits on the Wilson coefficients can be found by varying their values and computing the expression in Equation 10.2, obtaining the best-fit value, $\hat{c}_i$, at the maximum of L and a set of lower and upper limits on it, given by a difference of $2(3.84) \cdot \log L$ with respect to

the maximum likelihood, at 68% (95%) confidence level.

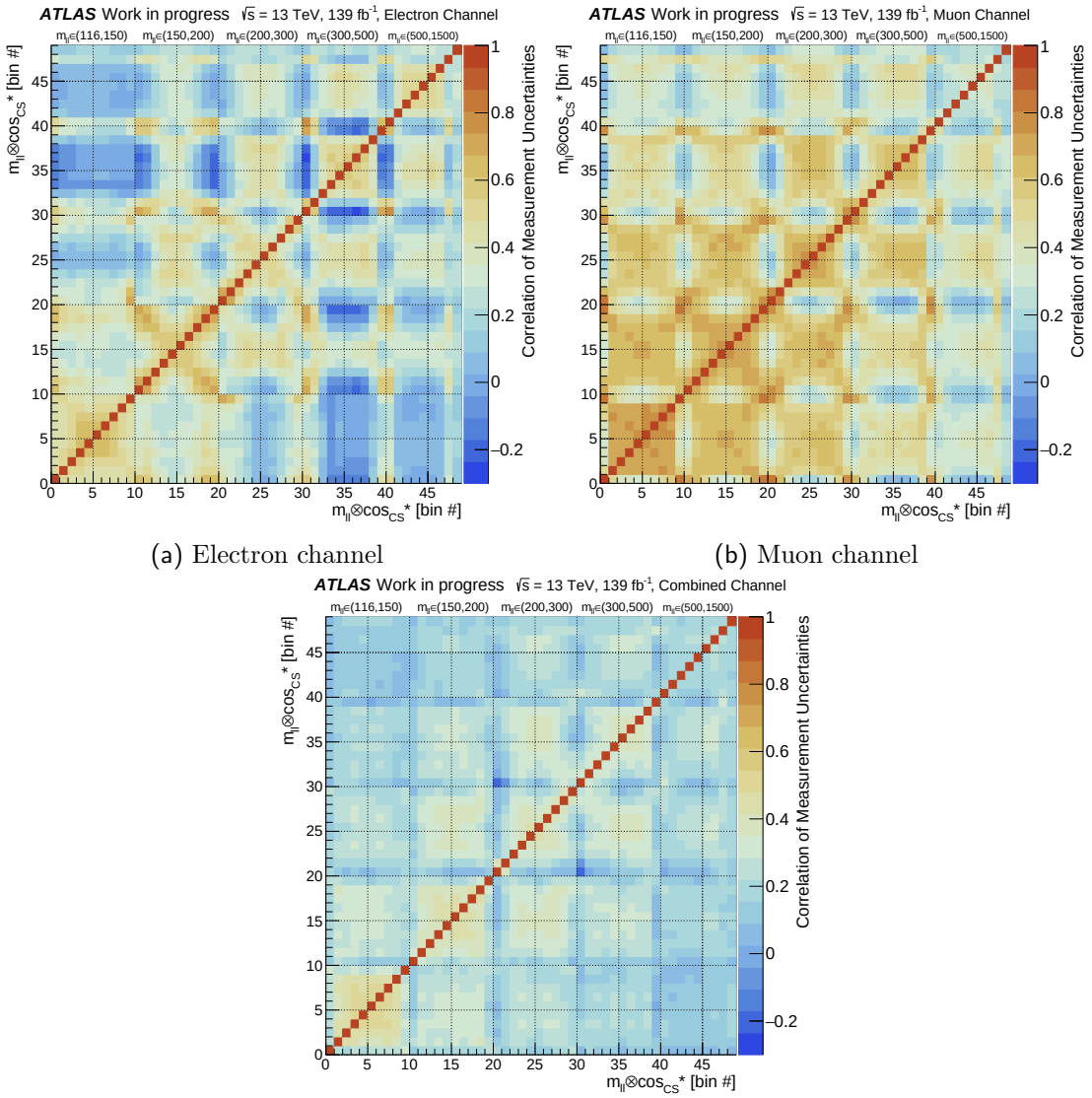


(a) Electron channel

(b) Muon channel

(c) Combined channel

Figure 10.8: Correlation matrix, $\mathbf{\Sigma}$, for the single-differential cross-section measurement against $m_{\ell\ell}$ in the different channels studied in the analysis.



(a) Electron channel

(b) Muon channel

(c) Combined channel

Figure 10.9: Correlation matrix, $\mathbf{\Sigma}$, for the double-differential cross-section measurement against $m_{\ell\ell}$ and $\cos\theta^*_{\mathrm{CS}}$ in the different channels studied in the analysis. The binning corresponds to that presented in Equation 8.3.

10.2.3 Single-differential measurement interpretation

Before performing the EFT fits, a fit to the measured distribution using only the Standard Model prediction fit is performed to ensure the fit methodology is robust. The results of the SM-only fits can be found in Figure 10.10. The fit behaves as expected, with pulls from the theory systematics bringing the prediction to a better agreement with data and reducing the uncertainties after the fit.

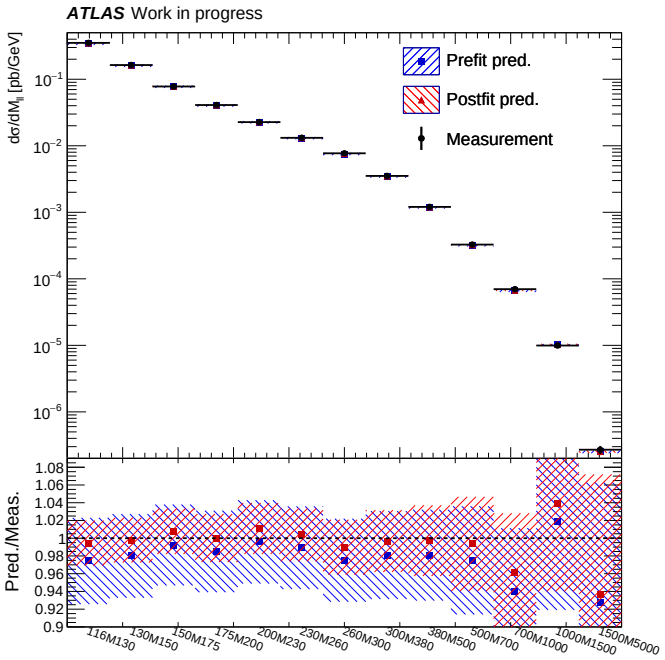


Figure 10.10: Single-differential Drell-Yan cross-section against $m_{\ell\ell}$ as measured in data and predicted by POWHEG (including NNLO QCD+NLO EW corrections). Predictions are shown before (red) and after (blue) the fit to the measured distribution has been performed.

The measurement, the SM predictions and the simulated EFT effects are then fitted jointly to extract best fit values and limits on the Wilson coefficients. The fits are performed in two setups: once using both linear and quadratic EFT contributions and a second time using only linear terms. Dimension-eight operators in the EFT expansion not considered in this analysis are at the same order in an expansion in $1/\Lambda$ as the quadratic terms of the dimension-six operators, so large differences between both fits would indicate that the neglected dimension-eight operators may play a non-negligible role. Expected limits on the coefficients are found by using the Standard Model prediction as pseudo-data while using experimental and theoretical systematic uncertainties observed in the measurement.

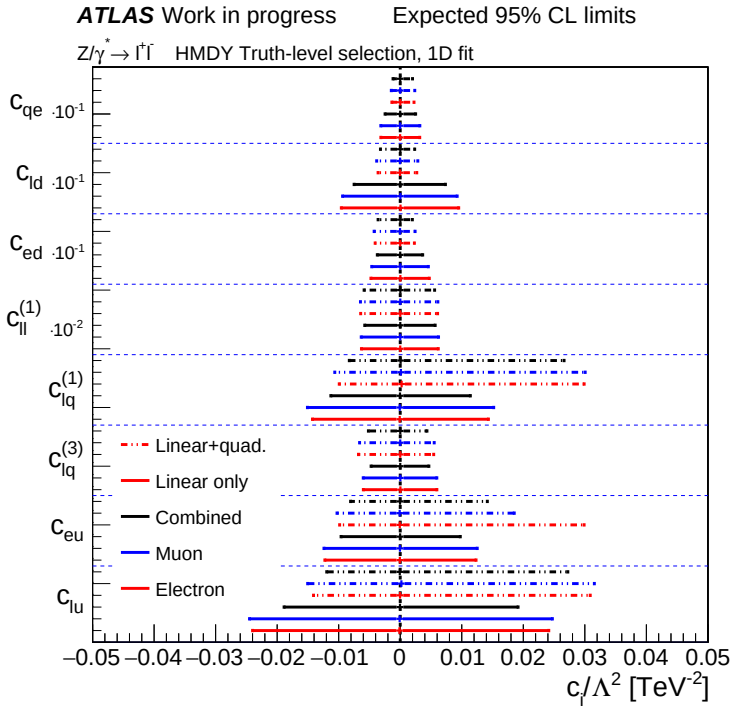
Expected and observed limits on the Wilson coefficients can be found in Figures 10.11 and 10.12 respectively. Detailed results can be found in Appendix B. All operators are compatible with the $c_i = 0$ assumption at 95%CL, but several features stand out from

these results:

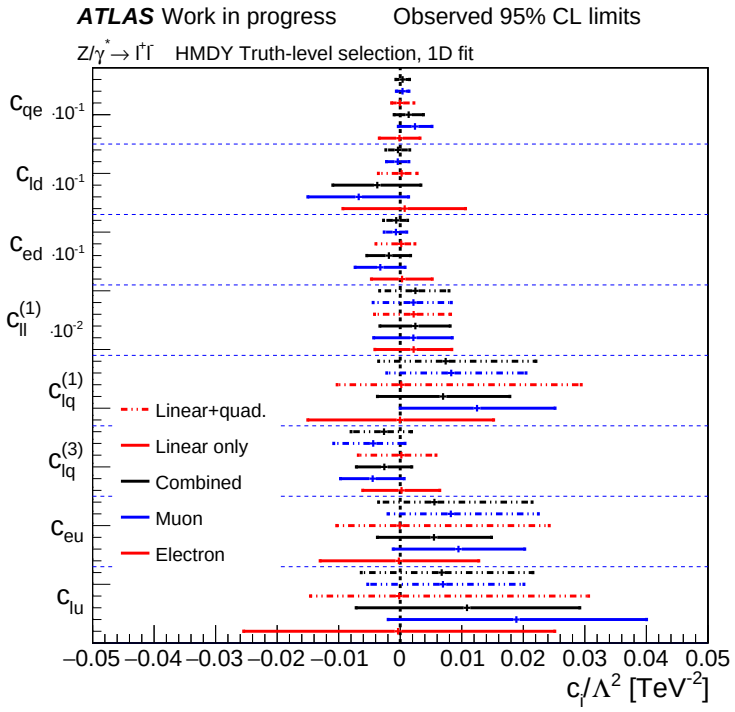
- As expected, stronger constraints can be placed on 4-fermion operators than Vff, since the effects of the former have a bigger impact on the Drell-Yan cross-section. Limits placed on most 4-fermion operators correspond to the strongest limits provided by any ATLAS measurement [169, 172, 173].

- The results using combined electron and muon measurements provide the strongest limits in all operators as expected from the experimental uncertainty reduction.

- When comparing the limits obtained with or without considering quadratic terms tighter limits can be placed on operators such as c_{ed}, c_{ld} or c_{He} when considering quadratic terms. Here quadratic terms enhance the contribution of the linear term, resulting into bigger effects on the DY cross-section and therefore stronger sensitivity. In the case of c_{Hu}, $c_{lq}^{(1)}$ or c_{eu} the linear term provides a negative contribution which, at certain values of the coefficient, can incur in cancellation with the quadratic term, resulting in a second minimum in the likelihood distribution.

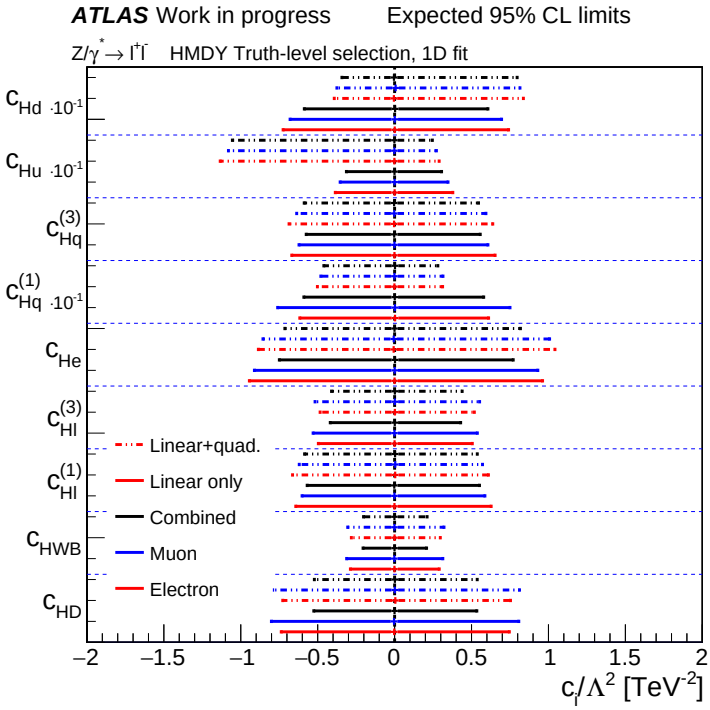


(a) Expected limits

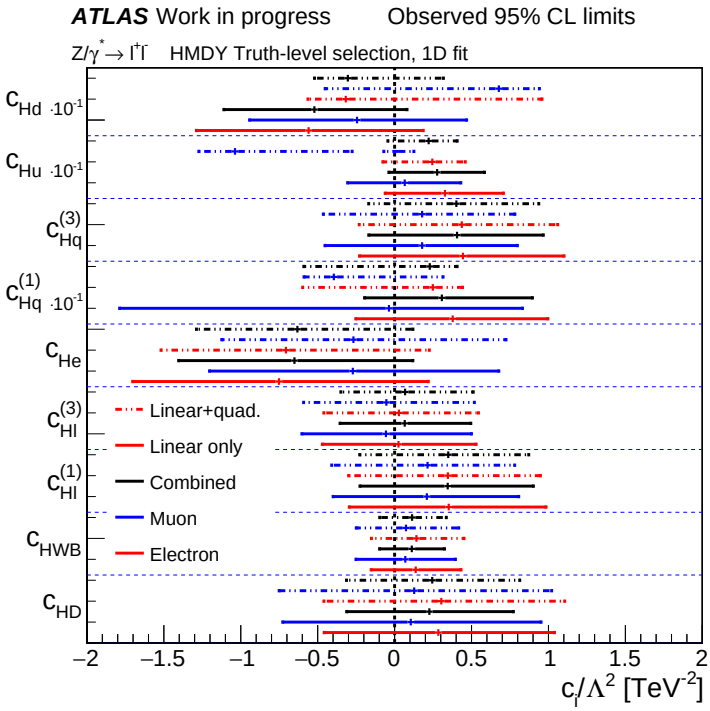


(b) Observed limits

Figure 10.11: 95% CL expected (top) and observed (bottom) limits on the Wilson coefficients corresponding to the operators shown in Table 10.1. Fits are performed using the single-differential measurement of the Drell-Yan cross section against $m_{\ell\ell}$. Black lines correspond to results obtained using the electron channel data, blue to muon channel and red to the combined channel. Solid lines correspond to limits obtained only using linear EFT effects, while dotted lines show results combining linear and quadratic effects.



(a) Expected limits



(b) Observed limits

Figure 10.12: 95% CL expected (top) and observed (bottom) limits on the Wilson coefficients corresponding to the operators shown in Table 10.2. Fits are performed using the single-differential measurement of the Drell-Yan cross section against $m_{\ell\ell}$. Black lines correspond to results obtained using the electron channel data, blue to muon channel and red to the combined channel. Solid lines correspond to limits obtained only using linear EFT effects, while dotted lines show results combining linear and quadratic effects.

10.2.4 Double-differential measurement interpretation

This section presents the results of the EFT interpretation of the double-differential measurement of the Drell-Yan cross-section against $m_{\ell\ell}$ and $\cos\theta^*_{\text{CS}}$. To enhance the sensitivity of the measurement to EFT effects, the last bin in invariant mass, $1500 < m_{\ell\ell} < 5000$ GeV, is added as the effects induced by most operators increase with the dilepton invariant mass. The resulting distribution, including all the kinematic phase space shown in Equation 8.3 and the last invariant mass in the single-differential measurement, will be further referred to as the 2D(+1) distribution. Following the same procedure as previously explained, a consistency check is performed by fitting a SM-only prediction. As discussed in Section 8.6, the prediction used for the double-differential distribution is based on the SHERPA Monte-Carlo generator, applying a mass-dependent correction to POWHEG. The results of the fit consistency check on the 2D(+1) distribution are shown in Figure 10.13. The fit behaves as expected: the fitting procedure brings the prediction to a good agreement with data and reduces the uncertainties after the fit.

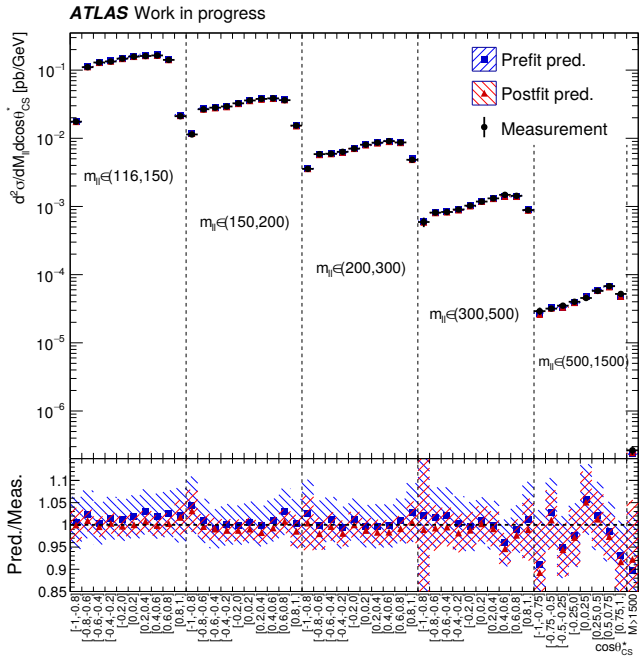
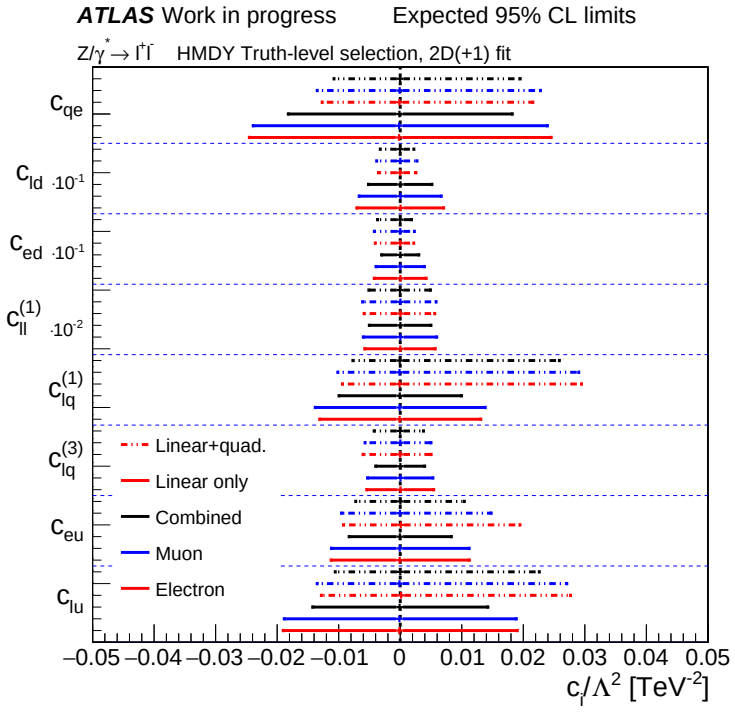


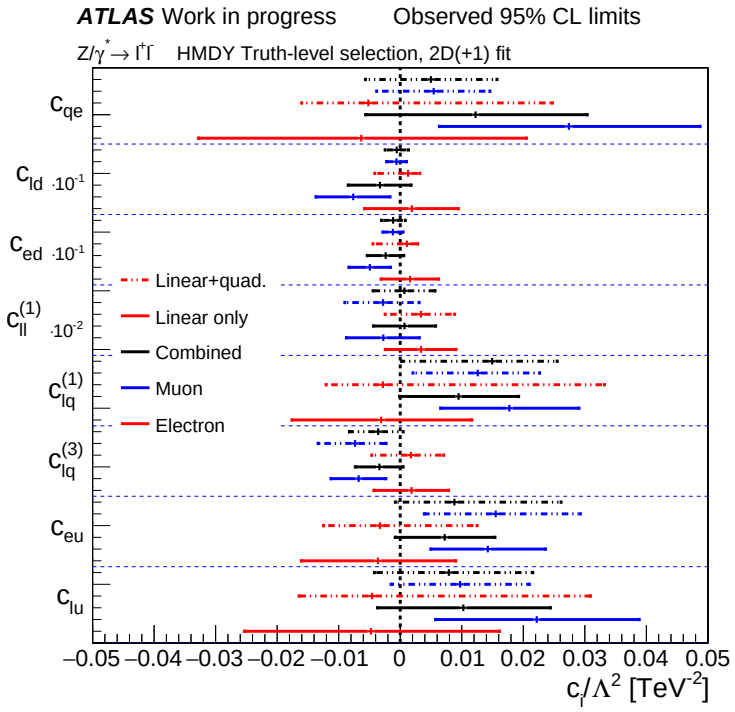
Figure 10.13: Double-differential Drell-Yan cross-section against $m_{\ell\ell}$ and $\cos\theta^*_{\text{CS}}$ as measured in data and predicted by SHERPA (including mass-dependent corrections based on the POWHEG prediction). Predictions are shown before (red) and after (blue) the fit to the measured distribution has been performed.

Fits are performed in the same way as discussed for the single-differential measurement. Expected and observed limits on the Wilson coefficients using the 2D(+1) distribution can be found in Figures 10.14 and 10.15 respectively. Detailed results can be found in Appendix B. Most operators are compatible with the $c_i = 0$ assumption at 95%CL, with the exception of the c_{qe}, $c_{lq}^{(1)}$, c_{eu} and c_{lu} which show up to 3.1σ deviation from $c_i = 0$ in the

linear fits using the results obtained in the muon channel. The sensitivity to Vff operators is enhanced with respect to the single-differential distribution by exploiting the shape dependence of the EFT-induced effects as shown in Figure 10.2. A comparison between fits performed with 1D and 2D(+1) distributions is shown in Table B.5. When comparing the limits obtained with or without considering quadratic terms, similar trends are found as when using the single-differential measurement. In some cases, the tighter limits obtained using the double-differential measurement result in second minima disappearing from the 95% CL interval, as observed in the c_{Hu} operator.

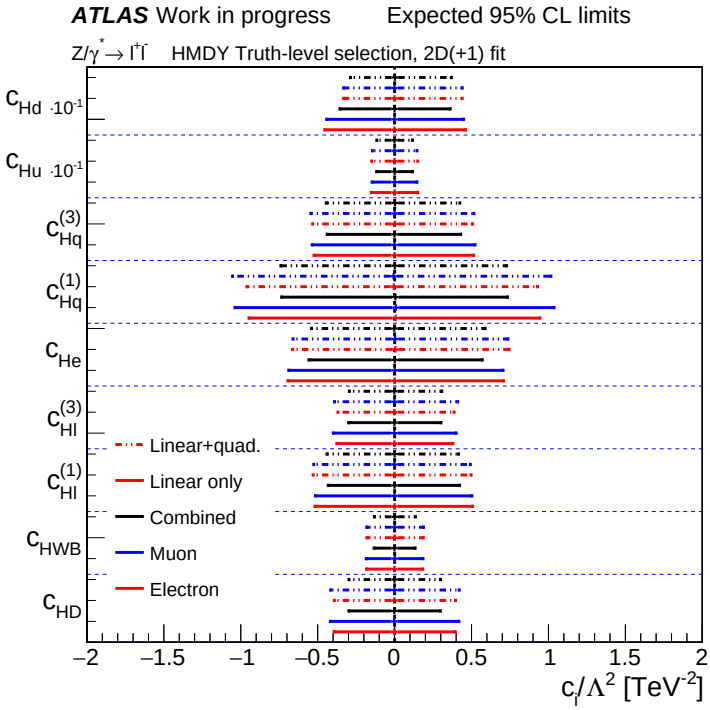


(a) Expected limits

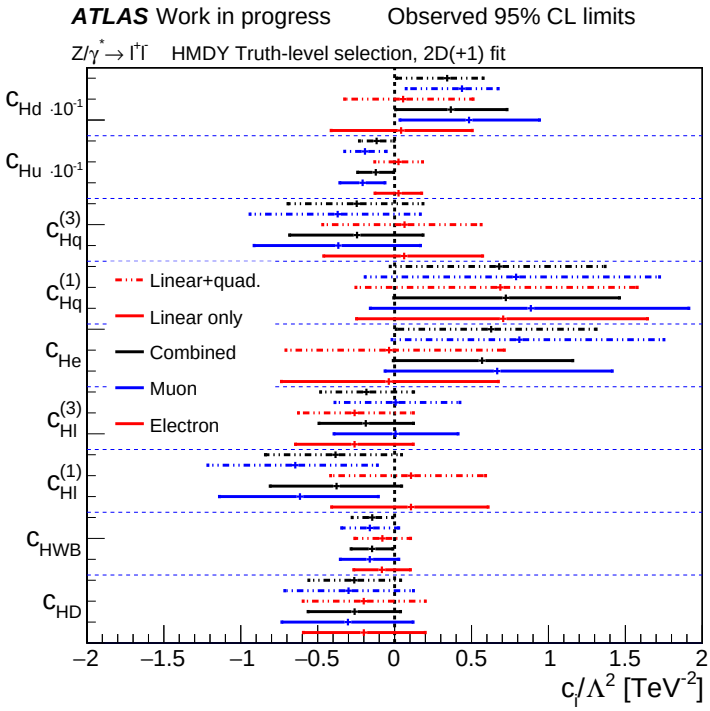


(b) Observed limits

Figure 10.14: 95% CL expected (top) and observed (bottom) limits on the Wilson coefficients corresponding to the operators shown in Table 10.1. Fits performed using the double-differential measurement of the Drell-Yan cross section against $m_{\ell\ell}$ and $\cos\theta^*_{\mathrm{CS}}$. Black lines correspond to results obtained using the electron channel data, blue to the muon channel and red to the combined channel. Solid lines correspond to limits obtained only using linear EFT effects, while dotted lines show results combining linear and quadratic effects.



(a) Expected limits



(b) Observed limits

Figure 10.15: 95% CL expected (top) and observed (bottom) limits on the Wilson coefficients corresponding to the operators shown in Table 10.2 (b). Fits performed using the double-differential measurement of the Drell-Yan cross section against $m_{\ell\ell}$ and $\cos\theta^*_{\mathrm{CS}}$. Black lines correspond to results obtained using the electron channel data, blue to the muon channel and red to the combined channel. Solid lines correspond to limits obtained only using linear EFT effects, while dotted lines show results combining linear and quadratic effects.

10.2.5 Principal component analysis

All results shown in the previous Sections correspond to individual fits of one SMEFT operator at a time, which may not encapsulate all the physics effects this interpretation can provide. However, the results presented in Chapter 9 do not contain sufficient information to perform a simultaneous fit of all operators that the studied final state is sensitive to. Following the methodology presented in Ref. [172], a Principal Component Analysis (PCA) is performed to construct a modified basis consisting of linear combinations of operators in the Warsaw basis. This can be achieved by re-parametrising the Hessian matrix, calculated as the inverse of the measurement's covariance matrix [174], in such way that it is expressed in terms of Wilson coefficients. In the Gaussian limit, the measurement of the eigenvectors of the re-parametrised Hessian matrix are uncorrelated and the expected uncertainty σ of the measurement in the direction of the eigenvectors is related to their eigenvalues, λ:

$$\sigma = \frac{1}{\sqrt{\lambda}}. \tag{10.7}$$

In order to improve the interpretability of the measurement, following the procedure used in Ref. [172], EFT operator groups are defined containing Wilson coefficients with similar physics impact. These groups are those corresponding to 4-fermion and Vff operators respectively. The leading eigenvectors (EVs) of the Hessian matrix in the single-differential measurement are shown in Figure 10.16. Out of all 17 resulting eigenvectors, fits are performed on the most sensitive EVs (those fulfilling $\sigma < 1$) from each category while having very uncorrelated effects, as shown in Figure 10.17. The selected operators are c_0, c_1, c_2, c_3, c_7 and c_8, for which Warsaw basis combinations and sensitivities are shown in Table 10.3.

Individual fits on these combinations of parameters were performed using the single-differential measurement, considering only linear EFT contributions. The consistent addition of quadratic terms would require the computation of all cross-terms of the parameters included in the fit, which due to the variety of sensitive operators leads to a huge number of samples that would need to be processed. Considering that these terms are sub-leading contributions (scaling with Λ^{-4} or higher orders), it was decided to perform the PCA fits using only the linear effects of the operators. The fits are also performed allowing all 6 chosen parameters to float, while profiling the likelihood with respect to a single one of them. The expected limits on the coefficients can be found in Figure 10.18. Very little difference is observed between individual and profiled limits, as expected when considering effects of very uncorrelated operators. The observed limits obtained using the single-differential DY cross-section measurement can be found in Figure 10.19. All coefficient combinations are found to be compatible with the $c_i = 0$ assumption at 95% CL.

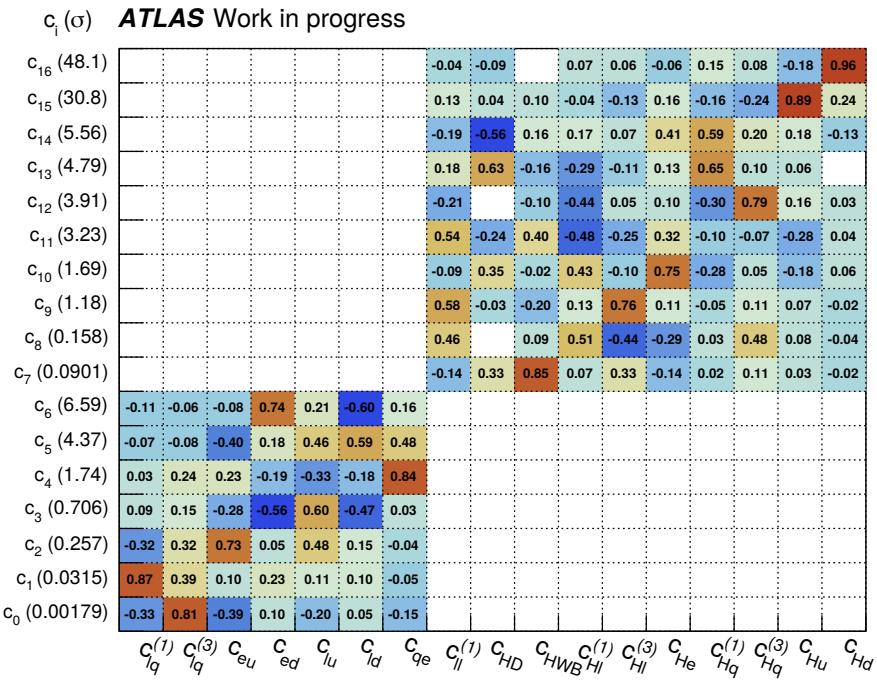


Figure 10.16: Graphical representation of the eigenvectors of the Hessian matrix of the single-differential measurement in the combined channel. Each row corresponds to an eigenvector (c_i), expressed as combinations of the EFT operators. The components of the eigenvectors are rounded to the second decimal place. The expected uncertainty σ (see Equation 10.7) of a measurement in the direction of each eigenvector is also included.

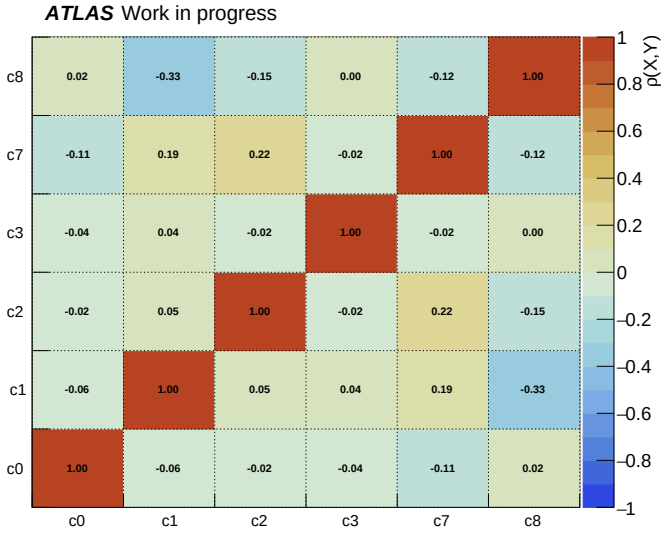


Figure 10.17: Correlation between the effects of the leading PCA eigenvectors included in the likelihood fit.

Operator	Re-redefinition	σ	Warsaw-basis combination
c_0	$c_{4f}^{\prime[1]}$	0.002	$-0.33c_{lq}^{(1)}+0.81c_{lq}^{(3)}-0.39c_{eu}+0.10c_{ed}$ $-0.20c_{lu}+0.05c_{ld}-0.15c_{qe}$
c_1	$c_{4f}^{\prime[2]}$	0.031	$0.87c_{lq}^{(1)}+0.39c_{lq}^{(3)}+0.10c_{eu}+0.23c_{ed}$ $+0.11c_{lu}+0.10c_{ld}-0.05c_{qe}$
c_2	$c_{4f}^{\prime[3]}$	0.257	$-0.32c_{lq}^{(1)}+0.32c_{lq}^{(3)}+0.73c_{eu}+0.05c_{ed}$ $+0.48c_{lu}+0.15c_{ld}-0.04c_{qe}$
c_3	$c_{4f}^{\prime[4]}$	0.706	$0.09c_{lq}^{(1)}+0.15c_{lq}^{(3)}-0.28c_{eu}-0.56c_{ed}$ $+0.60c_{lu}-0.47c_{ld}+0.03c_{qe}$
c_7	$c_{Vff}^{\prime[1]}$	0.090	$-0.14c_{ll}^{(1)}+0.33c_{HD}+0.85c_{HWB}+0.07c_{Hl}^{(1)}+0.33c_{Hl}^{(3)}$ $-0.14c_{He}+0.02c_{Hq}^{(1)}+0.11c_{Hq}^{(3)}+0.03c_{Hu}-0.02c_{Hd}$
c_8	$c_{Vff}^{\prime[2]}$	0.158	$0.46c_{ll}^{(1)}+0.09c_{HWB}+0.51c_{Hl}^{(1)}-0.44c_{Hl}^{(3)}$ $-0.29c_{He}+0.03c_{Hq}^{(1)}+0.48c_{Hq}^{(3)}+0.08c_{Hu}-0.04c_{Hd}$

Table 10.3: Main eigenvectors of the Hessian matrix of the single-differential measurement in the combined channel. The expected uncertainty σ of a measurement in the direction of each eigenvector is also included, ranking the eigenvectors from best constraint (smallest σ) to least constraint. The components of the eigenvectors are rounded to the second decimal place.

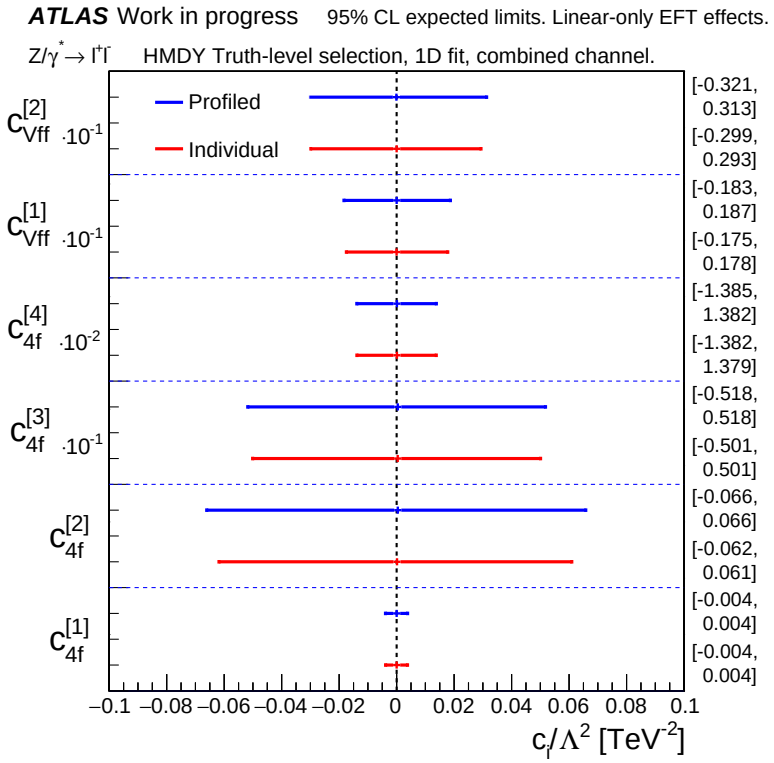


Figure 10.18: 95% CL expected limits on the re-parametrized operators shown in Table 10.3, extracted using the single-differential measurement in the combined channel. Limits are extracted using an Asimov fit, interpreting SM prediction as pseudo-data, while introducing experimental and theoretical systematic uncertainties observed in the measurement (and prediction) on it. All results correspond to limits obtained only using linear EFT effects.

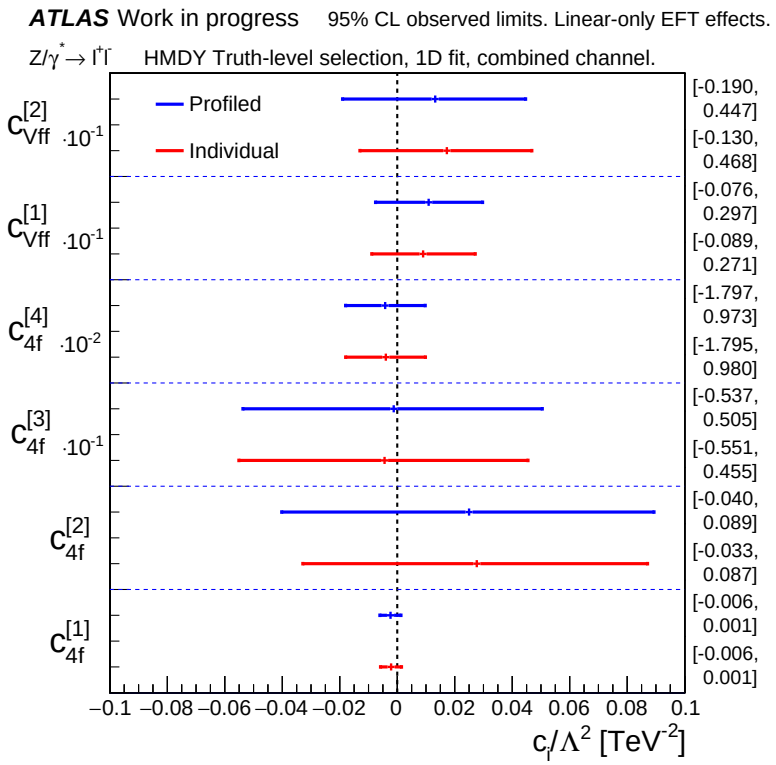


Figure 10.19: 95% CL observed limits on the re-parametrized operators shown in Table 10.3, extracted using the single-differential measurement in the combined channel. All results correspond to limits obtained only using linear EFT effects.

Chapter 11

Conclusion

The measurement of the single- and double-differential neutral-current Drell-Yan cross section was performed, using the data collected by the ATLAS experiment during the years 2015 to 2018. The center of mass energy of the proton-proton collisions recorded was $\sqrt{s} = 13$ TeV. The only available measurement of the Drell-Yan differential cross-section at the energy range explored in this measurement corresponds to early Run 2 results published by CMS [175], for which the smallest uncertainties were about 5%. The results are sensitive to the parton content of the protons, allowing to probe regions of x for which no experimental information was available with the level of precision provided in this measurement. The analysis covered dilepton invariant masses of up to $m_{\ell\ell} = 5000$ GeV, expanding the kinematic range of the measurement with respect to previous ATLAS results. With the inclusion of the photon-induced dilepton pair production as part of the measured signal, the results can also help to constrain the photon content of the proton.

A precision below 1% has been achieved in the phase space of low $m_{\ell\ell}$ (<230 GeV). The measurement is mostly limited by systematic uncertainties. The measured cross-section was compared with theoretical predictions using different PDFs, demonstrating good agreements overall but revealing the potential for better constraints using the data here presented.

The measurement was used for tests on Lepton Flavour Universality and Effective Field Theory interpretations. No significant deviations from the Standard Model were found, but small tensions were found in LFU tests at $m_{\ell\ell} > 1500$ GeV. Limits placed on EFT operators correspond to leading constraints in some of the operators. The measurement will be used as part of the global ATLAS EFT fit that will combine measurements from different processes to enhance the sensitivity to different EFT operators.

Part IV

Search for Lorentz-invariance violating signatures in the dilepton final state

Chapter 12

Motivation

The Standard Model Extension (SME) provides a model-independent approach to the parametrization of Lorentz-invariance violating (LIV) effects that induce a time-dependent variation on the production of the Z-boson at the ATLAS experiment, as first presented in Section 2.3.2. The search for LIV signals has never been done in any LHC experiment, with some of the most recent searches dating back to the ZEUS experiment at HERA [176] and D0 at Fermilab [177, 178]. Using the methodology first developed for luminosity measurements [179], the Z-boson production can be precisely monitored at the ATLAS experiment in a time-dependent manner. Using the dataset collected during the years 2015 and 2018, time modulations in the Z-boson production can be searched for and interpreted in the context of Standard Model Extension (SME) operators. The analysis of the Drell-Yan production at LHC offers sensitivity to the quark sector of SME which has not been explored yet thanks to the difference between vector and axial couplings of fermions to the Z-boson.

The following chapters will present the different elements involved in the time-dependent measurement of the Z-boson production at ATLAS and its interpretation in the SME framework. The structure is as follows: Chapter 13 presents an overview of the methodology for luminosity measurement using Z-counting. Chapter 14 describes the analysis methodology used to search for Lorentz-invariance violation using these results. In Chapter 15 the novel methodology for Monte-Carlo sensitivity studies is introduced, including an overview on how to model time dependence in the simulated samples. These results are used as an input for fits of SME-induced effects, setting the expected limits on the operators this measurement is sensitive to. This part concludes with a summary and outlook of the analysis in Chapter 16.

Chapter 13

Time-dependent Z-boson production measurement

This Chapter contains an overview of the methodology for a time-dependent measurement of the Z-boson production at the ATLAS experiment. Although the methodology follows that presented in the Z-counting luminosity measurement [179], it has been adapted to a newly developed analysis framework which runs on fully calibrated data, as opposed to the original analysis code, which ran on primary data output from the experiment. The transfer of the methodology to a common ATLAS analysis framework allows for systematic checks that the search for Lorentz-invariance violating signals may require, as well as a cross-check for the methodology and a proof of concept for time-dependent analyses in the experiment that may be expanded further to a variety of final states.

13.1 Measurement methodology

The Drell-Yan production of dilepton pairs is a well understood process, with the clean Z-boson signature of two high-transverse momentum, oppositely charged leptons[†] providing robustness against varying data-taking conditions. A luminosity measurement can be obtained by measuring the number of produced Z-bosons, N_Z, using the formula:

$$\mathcal{L}_Z = \frac{N_Z}{\sigma_Z}, \tag{13.1}$$

where σ_Z is the Z-boson production cross section. Experimentally, precise measurements of $\sigma_{Z\to e^+e^-}$ ($\sigma_{Z\to\mu^+\mu^-}$) have been obtained with systematic uncertainties (excluding luminosity) of about 1.0% (1.5%) [180]. The value of the $Z \to \ell^+\ell^-$ cross-section is predicted by calculations at NNLO QCD with a total uncertainty of 3-4% at 90% CL, where the uncertainty is driven by the current knowledge of the proton parton distribution func-

[†] Z-boson decays to τ leptons or quarks are not considered here, as they are much more difficult to reconstruct and prone to very high background levels.

tions (PDFs). As the final Run 2 uncertainties in the baseline ATLAS luminosity are about 0.83% [109], the Z-counting luminosity is not yet competitive as an absolute measurement using the predicted σ_Z, but has many powerful features as a relative luminometer.

The Z-counting procedure is performed as a function of time, in units of "luminosity blocks" (LBs), the smallest time interval for which luminosities are measured in ATLAS, typically corresponding to 60 seconds intervals. For each lepton flavour the raw rate of detected Z-bosons is measured and corrected to account for the trigger and reconstruction efficiencies which are derived in-situ from the same data. The data-driven efficiencies are further corrected using Monte Carlo simulations. A small background subtraction is also applied to remove dilepton pairs originating from diboson and $t\bar{t}$ events.

Thus, the general form for calculating the Z-counting based instantaneous luminosity as shown in Eq 13.1 is modified as follows [179]:

$$\mathcal{L}_{Z\to\ell^+\ell^-}(LB) = \frac{N_{Z\to\ell^+\ell^-}(LB)\cdot(1-f_{\text{bkg}})}{\sigma_{\text{theo}}\times A^{\text{MC}}_{Z\to\ell^+\ell^-}\cdot\varepsilon^{\text{T\&P}}_{Z\to\ell^+\ell^-}(LB)\cdot F^{\text{MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)\cdot t(LB)}, \tag{13.2}$$

where:

- $N_{Z\to\ell^+\ell^-}(LB)$ is the number of selected Z-boson candidates per LB, as defined in the previous section.
- f_{bkg} is the fraction of background events in the signal region. The value is taken as $f_{\text{bkg}} = 0.005$ in both the electron and muon channels, determined from Monte Carlo studies [180].
- σ_{theo} is the inclusive $Z\to\ell^+\ell^-$ cross-section determined for the range $m_{\ell\ell} > 60$ GeV to select the inclusive invariant mass range of the simulated Z events. The value used is $\sigma_{\text{theo}} = 1970$ pb, with a PDF uncertainty of 3.5% at 90% CL, obtained at NNLO QCD theory with the FEWZ 3.1.b2 framework [81–83] using the CT18ANNLO proton PDF [135]. All results presented are either based on $\mathcal{L}_{Z\to e^+e^-}/\mathcal{L}_{Z\to\mu^+\mu^-}$ ratios, or normalised to the integrated baseline ATLAS luminosity values, cancelling the dependence on this value and its uncertainty.
- $A^{\text{MC}}_{Z\to\ell^+\ell^-}$ is an acceptance factor correcting the inclusive cross-section σ_{theo} to the phase space defined in Table 13.1.
- $\varepsilon^{\text{T\&P}}_{Z\to\ell^+\ell^-}(LB)$ is the event trigger and reconstruction efficiency obtained per LB from the data-driven tag-and-probe procedure discussed in Section 13.3.
- $F^{\text{MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)$ is a correction factor for efficiency effects beyond the ones determined

in the tag-and-probe procedure, derived from Monte Carlo as a function of the average pileup, $\langle\mu\rangle$. See Section 13.4 for further details.

- $t(LB)$ is the live time of typically about 60 seconds per LB. The live time is corrected for detector dead time effects.

13.2 Data samples and event selection

The analysis presents the results for the full high-luminosity ATLAS Run 2 dataset, recorded between 2015 and 2018 for pp collisions at $\sqrt{s} = 13$ TeV, amounting to a total integrated luminosity of 139.0 fb^{-1}. Note that the luminosity measurement used as reference for the results presented in this thesis corresponds to the Run 2 preliminary results [152]. The luminosity used in the final results in the published analysis will correspond to the results reported in Ref. [109]. The dataset is divided within ATLAS into 595 runs, where one run typically corresponds to one LHC fill.

Events are recorded by the unprescaled single-lepton triggers with a transverse energy (momentum) threshold of 24 GeV (20 GeV) for electron (muon) candidates for 2015 data and of 26 GeV for 2016-18 data, within the pseudorapidity range of $|\eta| < 2.4$. Events with at least two well-identified and isolated leptons with the same flavour and opposite charge are selected, where either (or both) of the leptons pass the single-lepton trigger requirements. A likelihood based identification criterion is applied to the electron candidates, using the Medium working point. The muon candidates are required to be reconstructed from a combination of inner detector (ID) and muon spectrometer (MS) tracks and pass Medium identification criteria. Both electrons and muons are required to satisfy lepton isolation cuts found in Table 13.1. For more information on lepton identification and isolation at ATLAS see Chapter 6.

The full list of lepton selection requirements are found in Table 13.1, where the phase-space boundaries are determined by the transverse momentum (p_{T}) and pseudorapidity (η) ranges. All leptons passing these requirements will be referred to as "good" leptons, with a same-flavour, oppositely-charged pair of these required to form a Z-boson candidate, with an additional requirement on their invariant mass of $66 < m_{\ell\ell} < 116$ GeV.

13.3 Data-driven efficiency estimation

An important aspect of the Z-counting method is that electron and muon reconstruction and trigger efficiencies can be calculated using the Z events themselves, correcting for time-dependent detector inefficiencies affecting the detection rate. The efficiency to select

Selection criteria	Electron channel	Muon channel
Transverse momentum	$p_{\mathrm{T}}^{e} > 27$ GeV	$p_{\mathrm{T}}^{\mu} > 27$ GeV
Pseudorapidity	$0 < \lvert\eta^{e}\rvert < 1.37$ or $1.52 < \lvert\eta^{e}\rvert < 2.4$	$0. < \lvert\eta^{\mu}\rvert < 2.4$
Identification	LHMedium	Medium, combined muon
Isolation	$E_{\mathrm{T}}^{\mathrm{topo,cone20}}/p_{\mathrm{T}} < 0.2$ and $p_{\mathrm{T}}^{\mathrm{cone30}}/p_{\mathrm{T}} < 0.15$	$E_{\mathrm{T}}^{\mathrm{topo,cone20}}/p_{\mathrm{T}} < 0.3$ and $p_{\mathrm{T}}^{\mathrm{cone30}}/p_{\mathrm{T}} < 0.15$
Track-vertex association	$\lvert z_0 \sin\theta\rvert < 0.5$ mm $\lvert d_0\rvert/\sigma(d_0) < 5$	$\lvert z_0 \sin\theta\rvert < 0.5$ mm $\lvert d_0\rvert/\sigma(d_0) < 3$
Invariant mass	$66 < m_{e^+e^-} < 116$ GeV	$66 < m_{\mu^+\mu^-} < 116$ GeV

Table 13.1: Overview of selection criteria, where each criterion is applied to a single lepton and two oppositely charged leptons of the same flavour are required to form a Z-boson candidate, with invariant mass in the range 66 to 116 GeV.

a Z event is factorised into single-lepton trigger and reconstruction efficiencies, which are measured independently for each lepton flavour. The combination of these factors results in the Z-boson event-level selection efficiency as a function of time (in units of LBs).

13.3.1 Single-lepton trigger efficiency

The single-lepton trigger efficiency is defined with respect to the "good" lepton criteria and determined using events with two "good" leptons passing the selection detailed in Table 13.1. The single-lepton trigger efficiency, $\varepsilon_{\mathrm{trig},1\ell}$ can be derived by counting the number of events where exactly 1 lepton passed the trigger requirement (N_1) and the number of events where both leptons passed the trigger (N_2) as follows:

$$\varepsilon_{\mathrm{trig},1\ell} = \frac{1}{\frac{N_1}{2N_2} + 1}. \tag{13.3}$$

The statistical uncertainty on this efficiency is calculated by propagating the Poisson counting uncertainties on N_1 and N_2. Due to the requirements of two isolated and well identified leptons with a dilepton mass close to the Z-boson mass range, the background contribution is well below 0.1% and neglected. In the approximation that the trigger efficiencies of the two leptons are independent, the event-level trigger efficiency is defined as:

$$\varepsilon_{\mathrm{trig,event}} = 1 - (1 - \varepsilon_{\mathrm{trig},1\ell})^2. \tag{13.4}$$

13.3.2 Single-lepton reconstruction efficiency

For both electrons and muons, the single-lepton reconstruction efficiency is calculated in a tag-and-probe procedure (T&P) similar to that routinely done for ATLAS performance

analyses [118, 120]. A high-quality "tag lepton" is selected along with a loose "probe lepton" candidate with the extra requirement that the invariant mass of the tag-and-probe pair be close to the Z-boson mass. Here, the reconstruction efficiency is defined as the efficiency of the "good" lepton selection with respect to the loose object, and is determined from the fraction of probes that pass the "good" lepton selection. The efficiency $\varepsilon_{\text{reco},1\ell}$ is not the full efficiency to select a "good" lepton. Specifically, it does not include the efficiency to reconstruct the loose "probe" leptons. Any missing components are estimated entirely from simulation and form part of the correction factors discussed in Section 13.4.

The tag lepton is required to pass the single-lepton trigger to avoid a trigger bias, and is also required to pass tight selection criteria to ensure it has a low probability of being a mis-identified lepton. As the probe object has much looser selection criteria, mis-identified leptons and combinatorial backgrounds can be significant, so unlike for the trigger efficiency determination the background needs to be taken into account. The single-lepton reconstruction efficiency (indicated with the subscript "reco,1ℓ") is calculated according to:

$$\varepsilon_{\text{reco},1\ell} = \frac{N^{\text{OS}}_{\text{pass}} - N^{\text{bkg}}_{\text{pass}}}{N^{\text{OS}}_{\text{pass}} + N^{\text{OS}}_{\text{fail}} - N^{\text{bkg}}_{\text{total}}}, \tag{13.5}$$

where $N^{\text{OS}}_{\text{pass}}$ ($N^{\text{OS}}_{\text{fail}}$) is the number of opposite-charge tag-and-probe pairs where the probe satisfies (fails) the "good" lepton selection criteria and $N^{\text{bkg}}_{\text{total}}$ ($N^{\text{bkg}}_{\text{pass}}$) are estimates for the background in the denominator (numerator). The statistical uncertainty on this efficiency is calculated using error propagation on all factors in Equation 13.5.

The selection criteria used to select "tag" and "probe" leptons are listed in Table 13.2 and are discussed further below. The reconstruction efficiency is calculated only for leptons within the fiducial phase space defined in the preselection criteria listed in Table 13.1, while the "probe" selection criteria determines whether a lepton passes or fails.

Electron channel reconstruction efficiency In the electron channel, the tag electron is required to pass tight identification criteria [115] and be matched to a single-electron trigger object, as shown in Table 13.2. The probes are clusters of energy in the electromagnetic calorimeter loosely matched to an ID track with further kinematic requirements, also detailed in Table 13.2. The efficiency for the probe-level selection with respect to true electrons is about 98% and is well described by the simulation [115]. The probability to measure the electron charge correctly is about 99% and is well described by the simulation.

The nominal selection for the tag-and-probe pairs to be used in Equation 13.5 are opposite-charge pairs with an invariant mass of $75 < m_{ee} < 105$ GeV, called the peak

Candidate	Selection criteria	Electron channel	Muon channel
Tag	Lepton ID	LHTight	As in Table 13.1
	Trigger object	Matched	Matched
		Preselection criteria	
Probe	Object	EM cluster matched to ID track	ID track
	p_{T}	$p_{\mathrm{T}}^{e} > 27$ GeV	$p_{\mathrm{T}}^{\mu} > 27$ GeV
	η^{ℓ}	$0 < \lvert\eta^{e}\rvert < 1.37$ or $1.52 < \lvert\eta^{e}\rvert < 2.4$	$0 < \lvert\eta^{\mu}\rvert < 2.4$
	Track-vertex association	-	$\lvert z_0 \sin\theta\rvert < 2$ mm
		Probe selection criteria	
	Requirement	Passes all cuts in Table 13.1	Match to combined muon passing cuts in Table 13.1, with looser $p_{\mathrm{T}}^{\mu} > 21.6$ GeV cut.

Table 13.2: Lepton selection criteria for "tag" and "probe" candidates. Tag candidates are also required to pass all cuts in Table 13.1. Requirements for the probe selection criteria determine whether a lepton passes or fail the reconstruction.

region, where the probe is either passing or failing the "good" electron selection criteria. The background is estimated using a "template" method, where a background-enriched sample with minimal signal contribution is selected from data by inverting the identification and isolation requirements, i.e. requiring probes to fail both isolation and identification criteria, but pass all other selection. The template is normalised to the high-mass sideband within the range of $120 < m_{ee} < 250$ GeV using the statistics of the entire ATLAS run to obtain a smooth distribution. The background yields are obtained integrating the template over the peak region. In the denominator, the normalisation uses probes failing the "good" electron selection ("fail"), while in the numerator same-charge pairs ("SS") are used to ensure a small signal contamination. The resulting expression for the efficiency is:

$$\varepsilon_{\mathrm{reco},1\ell} = \frac{N^{\mathrm{OS}}_{\mathrm{pass,peak}} - N^{\mathrm{OS}}_{\mathrm{template,peak}} \cdot \frac{N^{\mathrm{SS}}_{\mathrm{pass,tail}}}{N^{\mathrm{SS}}_{\mathrm{template,tail}}}}{N^{\mathrm{OS}}_{\mathrm{pass,peak}} + N^{\mathrm{OS}}_{\mathrm{fail,peak}} - N^{\mathrm{OS}}_{\mathrm{template,peak}} \cdot \frac{N^{\mathrm{OS}}_{\mathrm{pass,tail}}}{N^{\mathrm{OS}}_{\mathrm{template,tail}}}}. \tag{13.6}$$

The selection and background estimation are illustrated in Figure 13.1 for an example ATLAS run. The efficiency is calculated using the background-subtracted events between the data points (signal) and background in the peak mass range, taking the ratio between the numerator (left) and denominator (right).

Muon channel reconstruction efficiency All muon candidates that pass the "good" selection criteria and are matched to a single muon trigger object are considered as tags, as shown in Table 13.2. In the muon channel, inner detector tracks passing the preselection cuts listed in Table 13.2 are used as probes. The efficiency for the probe-level selection with respect to true muons is stable and slightly above 99%, well described by simulation [120]. A pass or fail of the probe is determined by matching the ID track to a combined muon passing the "good" criteria with the p_{T} requirement loosened to 21.6 GeV, ensuring that ID tracks at the boundary of $p_{\mathrm{T}} = 27$ GeV can be successfully matched to a combined muon of slightly lower transverse momentum. All tag-and-probe pairs used in Equation 13.5 with the probe passing or failing the "good" muon selection are opposite-charge pairs within a tight invariant mass window of $86 < m_{\mu\mu} < 96$ GeV to reduce background contamination. Following Ref. [120], the background is estimated using a same-charge selection. The signal and background estimation are illustrated in Figure 13.2 for the same example ATLAS run. As explained for the electron channel, the muon reconstruction efficiency is calculated taking the ratio between the background-subtracted signal in the numerator (left) and denominator (right).

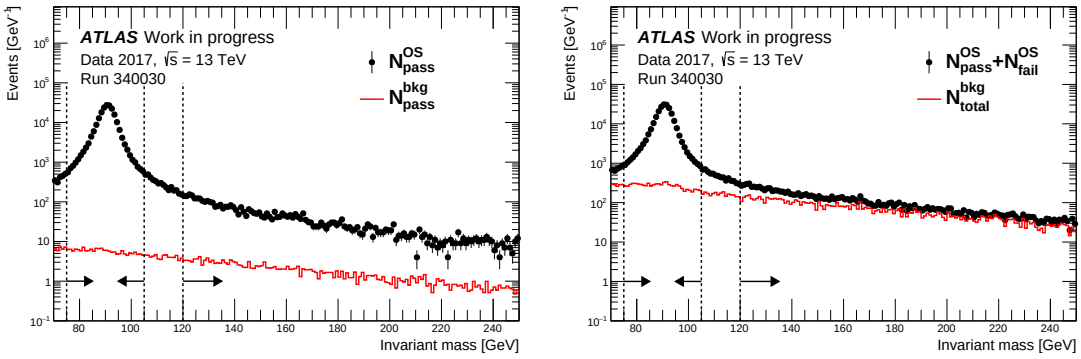


Figure 13.1: Invariant dielectron mass distributions used to calculate the reconstruction efficiency as given in Eq. 13.5, where the left (right) figure shows the contribution to the numerator (denominator). Vertical dashed lines illustrate the "peak" and "tail" mass ranges for the electron template method. The data was recorded from pp collisions at $\sqrt{s} = 13$ TeV in ATLAS Run 340030 on November 4, 2017. The error bars show statistical uncertainties only.

13.3.3 Event-level efficiency

For a $Z \to \ell^+\ell^-$ event there are two leptons and the single-lepton efficiencies must therefore be combined to determine event-level efficiencies. This considers that at least one lepton trigger the Z-candidate event and both leptons are required to pass the "good" lepton criteria:

$$\varepsilon^{\mathrm{T\&P}}_{Z\to\ell\ell} = \left(1 - (1 - \varepsilon_{\mathrm{trig},1\ell})^2\right) \times \varepsilon^2_{\mathrm{reco},1\ell}. \tag{13.7}$$

The event-level efficiency, $\varepsilon^{\mathrm{T\&P}}_{Z\to\ell\ell}$, is applied per decay channel to correct the number

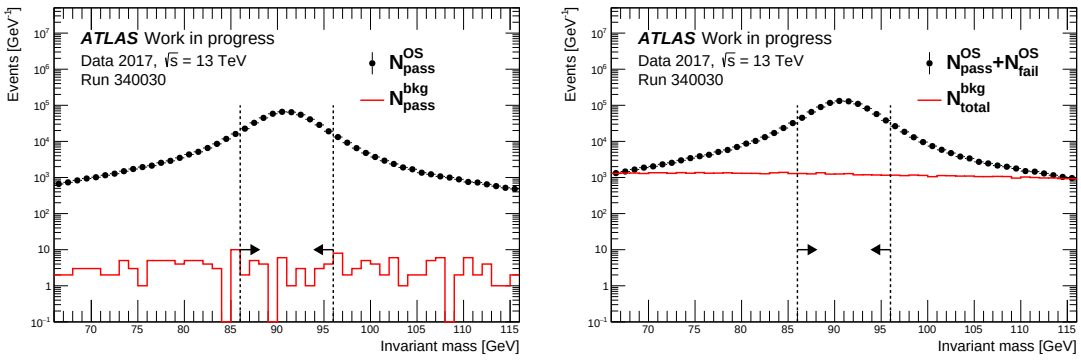


Figure 13.2: Invariant dimuon mass distributions used to calculate the reconstruction efficiency as given in Eq. 13.5, where the left (right) figure shows the contribution to the numerator (denominator). Vertical dashed lines illustrate the "peak" mass range used in the muon channel tag-and-probe procedure. The data was recorded from pp collisions at $\sqrt{s} = 13$ TeV in ATLAS Run 340030 on November 4, 2017. The error bars show statistical uncertainties only.

of raw $Z \to \ell^+\ell^-$ candidates. The corresponding statistical uncertainty is calculated by propagating the uncertainties on each component.

13.4 Correction factors from simulation

The event-level efficiency described in Equation 13.7 accounts for the trigger and reconstruction efficiencies determined from data using a factorised ansatz based on single lepton efficiencies. However, it does not include additional effects due to the ID track efficiency for muons, nor the track-to-cluster matching efficiency or the electron charge mis-identification probability for electrons. Also, non-factorising effects are not included and could lead to a bias in the calculated event-level efficiency, if the efficiencies of the two leptons were correlated. The factorisation is tested and a correction factor is obtained from samples of simulated $Z/\gamma^* \to \ell^+\ell^-$ events.

13.4.1 Samples and procedure

The Monte Carlo $Z/\gamma^* \to \ell^+\ell^-$ signal samples used for this analysis are the same as those used in the high-mass Drell-Yan cross-section measurement, introduced in Section 8.1. Simulated events are required to pass the same kinematic selection criteria as for data, shown in Table 13.1. The closure of the tag-and-probe procedure can then be tested by comparing the full reconstruction efficiency determined in Monte Carlo to the value of $\varepsilon^{\text{T\&P}}_{Z\to\ell\ell}$ determined using the tag-and-probe method on the Monte Carlo events.

The Monte Carlo correction factor, first introduced in Equation 13.2, $F^{\text{MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)$, is defined as:

$$F^{\mathrm{MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle) = \frac{N^{\mathrm{reco,fid,MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)}{N^{\mathrm{true,nocut,MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)} \times \frac{1}{A^{\mathrm{MC}}_{Z\to\ell^+\ell^-}\cdot\varepsilon^{\mathrm{T\&P,MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)}, \tag{13.8}$$

where all quantities are evaluated with Monte Carlo events as a function of the pileup parameter, $\langle\mu\rangle$, and:

- $N^{\mathrm{reco,fid,MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)$ is the number of reconstructed events ("reco") which pass the fiducial phase space ("fid") and event selection requirements presented in Section 13.2.
- $N^{\mathrm{true,nocut,MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)$ is the number of generated Z-boson MC events ("true") without any selection ("nocut").
- $A^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$ is the fiducial acceptance, calculated using leptons originating from a Z-boson as described in Section 13.4.2.
- $\varepsilon^{\mathrm{T\&P,MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)$ is derived by repeating the tag-and-probe (T&P) procedure using reconstructed MC events with small modifications to the background subtraction discussed in Section 13.4.3.

The luminosity calculation in Equation 13.2 depends only on the product $A^{\mathrm{MC}}_{Z\to\ell^+\ell^-}\cdot F^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$, and $A^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$ appears in the denominator of the definition of $F^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$. Hence, the determined luminosity is independent of the value of $A^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$. The separation of these two factors is a matter of convention intended to capture the detector effects in $F^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$, but the Z-counting results do not depend on this separation.

13.4.2 Fiducial acceptance

The acceptance factor $A^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$ is derived from the MC signal samples using true Z-boson events as follows:

$$A^{\mathrm{MC}}_{Z\to\ell^+\ell^-} = \frac{N^{\mathrm{true,fid,MC}}_{Z\to\ell^+\ell^-}}{N^{\mathrm{true,nocut,MC}}_{Z\to\ell^+\ell^-}}, \tag{13.9}$$

where:

- $N^{\mathrm{true,fid,MC}}_{Z\to\ell^+\ell^-}$ is the number of generated events with a pair of opposite-sign leptons with $p_{\mathrm{T}} > 27$ GeV, $|\eta| < 2.4$ (with the additional removal of the crack region $1.37 < |\eta| < 1.52$ for electrons) and $66 < m_{\ell\ell} < 116$ GeV.
- $N^{\mathrm{true,nocut,MC}}_{Z\to\ell^+\ell^-}$ is the number of generated events with a true mass of $m_{\ell\ell} > 60$ GeV.

In order to yield residual factors, $F^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$, that are close to one, an appropriate definition of $A^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$ is needed. This is achieved by using *bare* and *dressed* definition[†] for

[†]Further details on those definitions can be found in Section 8.3.

the generator-level lepton kinematics for muon and electron channels, respectively. The *bare* momentum corresponds to the lepton after emission of QED FSR and is appropriate for muons, where the momentum is measured from the track. The *dressed* momentum is appropriate for electrons, since the energy is measured in the electromagnetic calorimeter, where the photon and electron energy deposits overlap. The obtained acceptance values for $A^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$ are shown in Table 13.3 and are in perfect agreement with those found in the original Z-counting luminosity measurement [179]. As previously mentioned, the calculation details of $A^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$ do not impact the final luminosity determination, therefore the uncertainties on this value are not considered for the Z-counting luminosity uncertainty calculation.

$A^{\mathrm{MC}}_{Z\to e^+e^-}$	$A^{\mathrm{MC}}_{Z\to\mu^+\mu^-}$
0.2996 ± 0.0002	0.3326 ± 0.0002

Table 13.3: Fiducial acceptance values, calculated for the electron and muon channels using the corresponding MC signal samples. The uncertainties reflect the statistical uncertainty of the simulated data.

13.4.3 Tag-and-probe efficiency in Monte Carlo

The pileup-dependent correction factor, $F^{\mathrm{MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)$, and its components are derived using Monte Carlo $Z/\gamma^* \to \ell^+\ell^-$ signal samples. This involves the full application of the efficiency measurement procedure detailed in Section 13.3 for the data. As opposed to the analysis performed with data, each event contains a true $Z \to \ell\ell$ decay and there is no QCD multijet contamination in the signal lepton pairs. However, a small amount of background could be present in the tag-and-probe distributions when a probe candidate corresponds to misidentified hadronic activity, non-prompt lepton production and overlaid pileup events. This is removed by requiring the reconstructed leptons to be matched to a true generated lepton from the $Z \to \ell\ell$ decay. The single-lepton efficiencies obtained with reconstructed MC events according to Equations 13.3 and 13.5 are then combined in the same way as in data using Equation 13.7 to an event-level Z-boson efficiency, $\varepsilon^{\mathrm{T\&P,MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)$. A detailed comparison of the results obtained in the Monte Carlo samples with those obtained in data can be found in Appendix C.

13.4.4 Results for the Monte Carlo correction factor

The pileup-dependent Monte Carlo correction factors, $F^{\mathrm{MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)$ are derived separately for each data-taking periods (each individual year between 2015 and 2018).

The $\langle\mu\rangle$ dependence of $F^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$ is fitted with a second-order polynomial per decay channel and per data-taking period, and the fit results are shown in Figure 13.3. These

parametrisations are used in the determination of the Z-counting based luminosity per data-taking period and as a function of the pileup parameter $\langle\mu\rangle$. The statistical uncertainties of the fits are small and are neglected.

The $Z \to e^+e^-$ correction factor is found to be $> 10\%$ below unity for all $\langle\mu\rangle$, implying that there are additional efficiency effects beyond those captured in the tag-and-probe procedure. The variation between $\langle\mu\rangle = 20$ and $\langle\mu\rangle = 50$ is around 2%. The $Z \to \mu^+\mu^-$ correction factor is few % below unity for all $\langle\mu\rangle$, meaning that a smaller correction is required in this channel. The variation across the pileup range where the majority of data-taking took place is around 1%, with pileup-dependent effects having a smaller impact than in the electron channel.

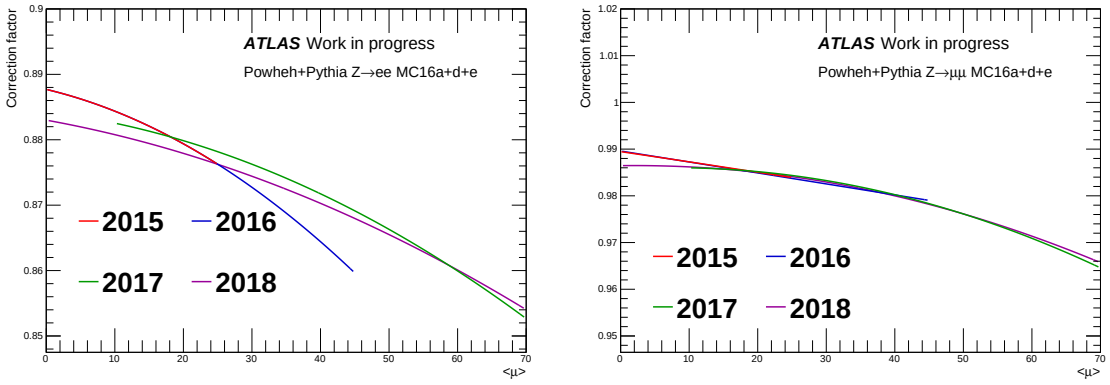


Figure 13.3: $Z \to e^+e^-$ (left) $Z \to \mu^+\mu^-$ (right) correction factors used for each year of data produced using the dedicated Monte Carlo campaigns. The lines show second-order polynomial fits to the correction factor for the corresponding $\langle\mu\rangle$ range per year.

13.5 Results

The following section shows detailed results illustrating the methodology of Z-counting using data recorded for $\sqrt{s} = 13$ TeV pp collisions in ATLAS runs 281411, 300800, 340030 and 360373 in 2015, 2016, 2017 and 2018 respectively. As shown in Table 13.4, the pp luminosity delivered per run, as well as the average number of interactions per bunch crossing, increased significantly from year to year. Throughout the development of the new analysis framework a series of limitations were found to adapt the methodology in the muon channel, as reported in Appendix D. The results here presented will therefore correspond to those obtained in the electron channel, which show an excellent agreement with the original Z-counting results [179] and are an important milestone to the development of time-dependent analyses in ATLAS. However, the final publication of the search for Lorentz-invariance violation will be based on the original results based on the online monitoring tool, which include both electron and muon channel.

Data-taking period	ATLAS Run	Date	Luminosity [pb^{-1}]	Average pileup $\langle\mu\rangle$
2015	281411	11/10/15	163.5	15.1
2016	300800	03/06/16	313.0	21.8
2017	340030	04/11/17	725.3	39.9
2018	360373	09/09/18	416.8	37.5

Table 13.4: Information for the selected ATLAS runs used to illustrate the Z-counting methodology for each of the Run 2 data-taking periods.

The single-lepton reconstruction and trigger efficiencies for each of these runs are shown in Figure 13.4. For display purposes, the efficiencies have been obtained by a weighted average over 20 luminosity blocks to improve the statistical uncertainties. The efficiencies track the variations across the run and reflect the time evolution of the data-taking conditions. After forming the event-level efficiencies and applying the pileup-dependent Monte Carlo correction factors per LB, Z-counting luminosities are combined for $N = 20$ LBs as,

$$\mathcal{L}_{Z\to\ell^+\ell^-} = \frac{\sum_{LB}^{N} \mathcal{L}_{Z\to\ell^+\ell^-}(LB)\cdot t(LB)}{\sum_{LB}^{N} t(LB)}, \tag{13.10}$$

where $\mathcal{L}_{Z\to\ell^+\ell^-}(LB)$ is the individual Z-counting luminosity for each LB and $t(LB)$ is the duration of said luminosity block.

The absolute ratio between the Z-counting and ATLAS baseline luminosities was found to be consistent with unity within the total uncertainty when adding the dominant theoretical and experimental systematic uncertainties in quadrature. The remainder of the results here presented focus on the relative consistency of the Z-counting and baseline ATLAS luminosity measurements as a function of time. Therefore, in all comparisons of the Z-counting and baseline luminosities, the Z-counting luminosity is normalised to the same integrated luminosity as the baseline ATLAS measurement. This can be done over different time periods, for instance a single ATLAS Run, a single data-taking period (year) or the entire Run 2 period. The Z-counting luminosity normalised to the run-integrated baseline ATLAS luminosity is shown in Figure 13.5 for each of the example ATLAS Runs in Table 13.4. The ratio of the normalised Z-counting and ATLAS luminosities has a spread at or below the 2% level, indicating the excellent relative stability between both measurements.

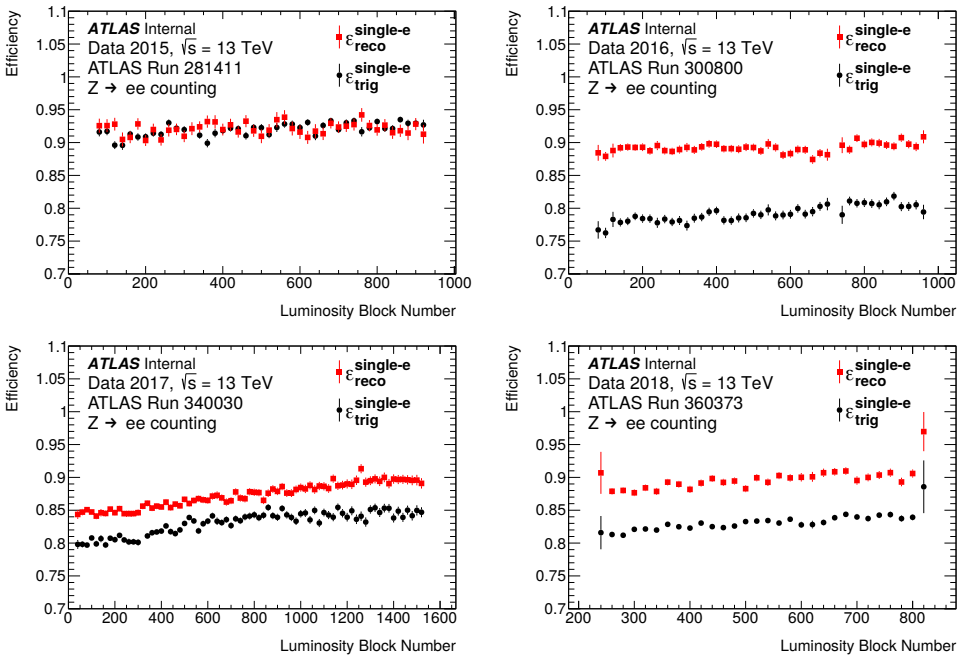


Figure 13.4: Time-dependence of the single electron trigger (black circles) and reconstruction (red squares) efficiencies. Efficiencies were determined using the methodology explained in Section 13.3. Shown is the average of the efficiencies over 20 LBs. Error bars show statistical uncertainties.

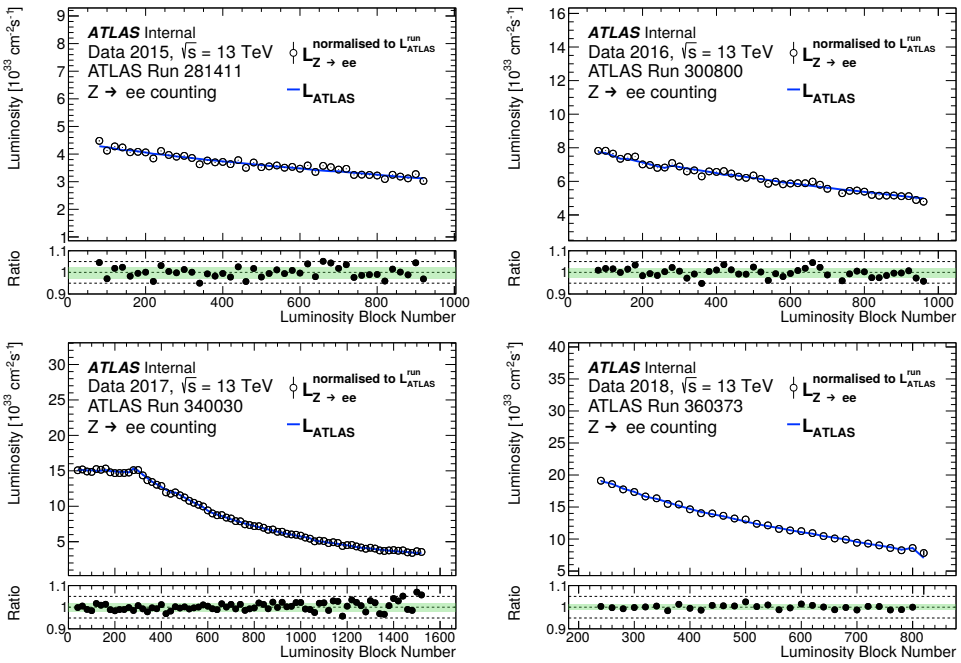


Figure 13.5: Time-dependence in units of LBs of the instantaneous luminosity determined from Z-counting (open circles), the baseline ATLAS luminosity (blue lines) and their ratios (full circles), for the electron channel. The run-integrated Z-counting luminosity is normalised to the baseline ATLAS luminosity in the run. Error bars show the statistical uncertainty of the $\mathcal{L}_{Z\to\ell^+\ell^-}$ determination. The green band in the ratio plot contains 68% of all points centred around the mean.

13.5.1 Time-dependence of $\mathcal{L}_Z/\mathcal{L}_{\mathrm{ATLAS}}$

The ratio of the normalised Z-counting luminosity and the baseline ATLAS measurement [152] is used to study their relative stability over all Run 2 data-taking periods. The spread of the ratio around unity quantifies the relative stability for the two measurements, shown in Figure 13.6 for each data-taking period. The results indicate that the relative stability of the two measurements is very good.

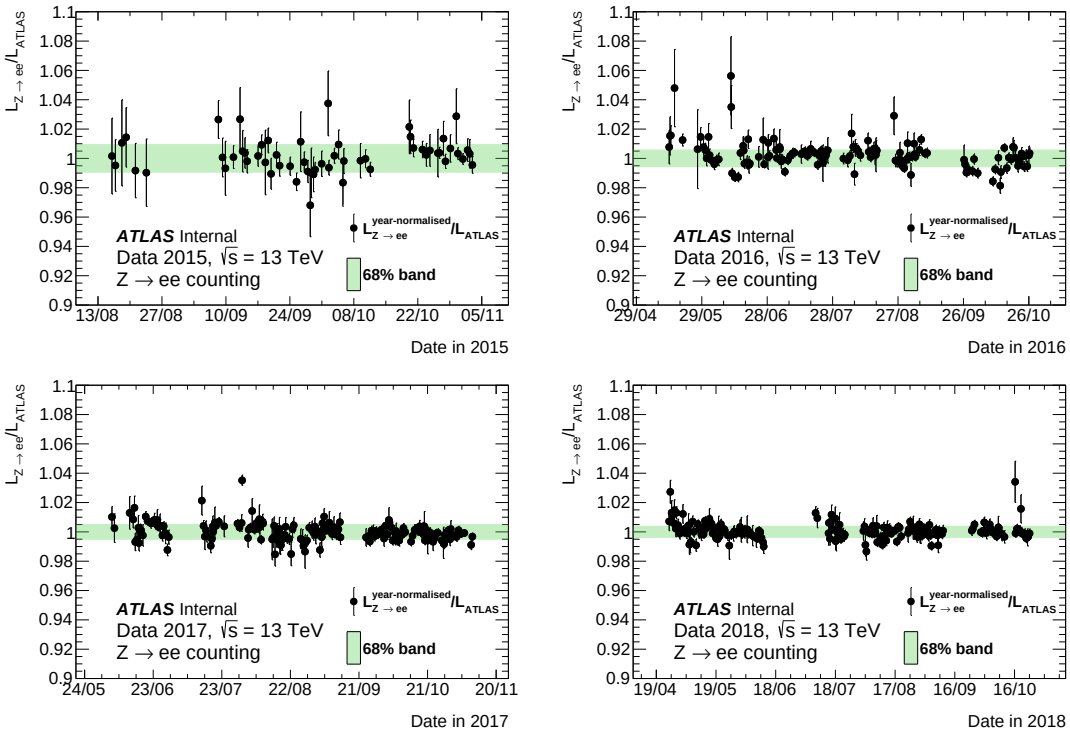


Figure 13.6: Ratio of the integrated Z-counting and baseline ATLAS luminosities per ATLAS run taken from pp collisions at $\sqrt{s} = 13$ TeV for the $Z \to e^+e^-$ channel. The Z-counting luminosity is normalised to the integrated baseline ATLAS luminosity per data-taking period [152]. The x-axis represents the date when the run started. Only ATLAS runs with a minimum length of 40 minutes are included. The error bars show statistical uncertainties and the green bands contain 68% of all points centred around the mean.

The expected improvement in the statistical precision of the Z-counting method with increasing instantaneous luminosity is observed, with the spread around unity decreasing from 2015 to the following years. Furthermore, the year-to-year stability of the ATLAS luminosity scale can be monitored by normalising the Z-counting luminosity such that the total Run 2 integrated luminosity is equal to the corresponding baseline ATLAS value, as shown in Figure 13.7. The spread of these results is approximately 0.66%. The published Z-counting results observed a spread of 0.8% across the Run 2 dataset [179], meaning that the use of calibrated data brings a better stability with respect to the nominal ATLAS luminosity measurement. The variations observed are found within the uncorrelated

year-by-year uncertainties that affect the absolute scale of the baseline ATLAS luminosity, which amount to 1.3% for 2015/16, 1.3% for 2017 and 1.0% for 2018 in the preliminary calibration [152].

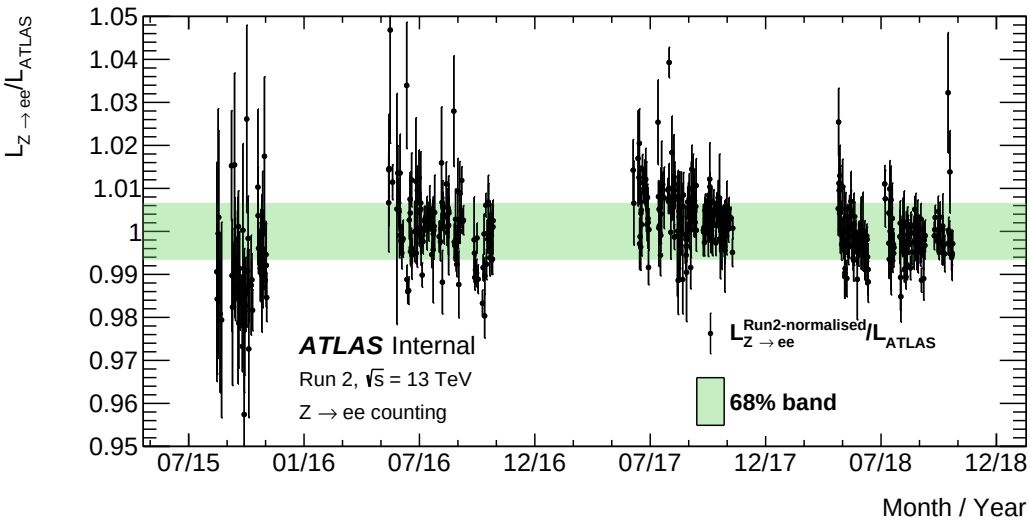


Figure 13.7: Ratio of the integrated Z-counting and baseline ATLAS luminosities per ATLAS run taken from pp collisions at $\sqrt{s} = 13$ TeV for the $Z \to e^+e^-$ channel for the full Run 2 data taking period. The Z-counting luminosity is normalised to the integrated baseline ATLAS luminosity over the Run 2 data-taking period [152]. The x-axis represents the date when the run started. Only ATLAS runs with a minimum length of 40 minutes are included. The error bars show statistical uncertainties and the green bands contain 68% of all points centered around the mean.

13.5.2 Statistical bias studies

The main limitation for the Z-counting methodology as compared to other luminometers resides in the low Z-boson production rates in some of the data-taking periods. Counts of less than 150 bosons per LB may lead to statistical biases in the tag-and-probe efficiency determination. These can account to an overestimation of the luminosity of up to 2% when the number of Z candidates in a luminosity block is below 50, staying under 0.5% in those data periods with over 200 Z events per LB, as observed in previous toy studies [181]. To minimise the effects of this feature of the measurement, the results were re-evaluated aggregating the data through different luminosity blocks. All variables used in Equation 13.2 (and their subsequent elements), are summed together if the luminosity blocks being aggregated meet certain criteria:

- The total number of $Z \to e^+e^-$ reaches 200, at which point the effects of the statistical bias is found to be minimal.
- A total livetime of 15 minutes is recorded among all luminosity blocks being merged.

This condition preserves the time resolution in data taking required to maintain optimal sensitivity to Lorentz-invariance violating effects [182].

- More than 20 luminosity blocks are merged. This condition protects against gaps in data-taking that may cause LBs being merged across large time intervals.

Figure 13.8 shows the ratio between the results obtained using the standard Z-counting methodology and the aforementioned LB-aggregated results. The correction derived by merging quantities across multiple luminosity blocks can be up to 2% in early 2015 runs, while the effects become negligible in 2017 and 2018, where $< 0.2\%$ differences are observed throughout most of the data-taking period. The impact on the aggregation of the results across luminosity blocks is shown in Table 13.5. A 0.06% larger spread is found in the 2015 dataset, while the following years, in which the luminosity block aggregation is expected to have a smaller impact, show an improvement in the overall spread. Across the entire Run 2 dataset, a reduction of 0.02% of the spread of the $\mathcal{L}_{Z\to ee}/\mathcal{L}_{\text{ATLAS}}$ ratio is found, showing that, although LB aggregation may have a big impact on results of individual ATLAS runs where a low number of events were recorded, the stability improvement across the entire dataset is relatively mild.

Data period	Methodology	
	Standard	LB-aggregated
2015	0.98%	1.04%
2016	0.62%	0.61%
2017	0.54%	0.54%
2018	0.41%	0.39%
Full Run 2	0.66%	0.64%

Table 13.5: Summary of the spread (68% of all points centred around unity) of the $\mathcal{L}_{Z\to ee}/\mathcal{L}_{\text{ATLAS}}$ ratio for each of the considered data-taking periods, as well as the full Run 2 dataset.

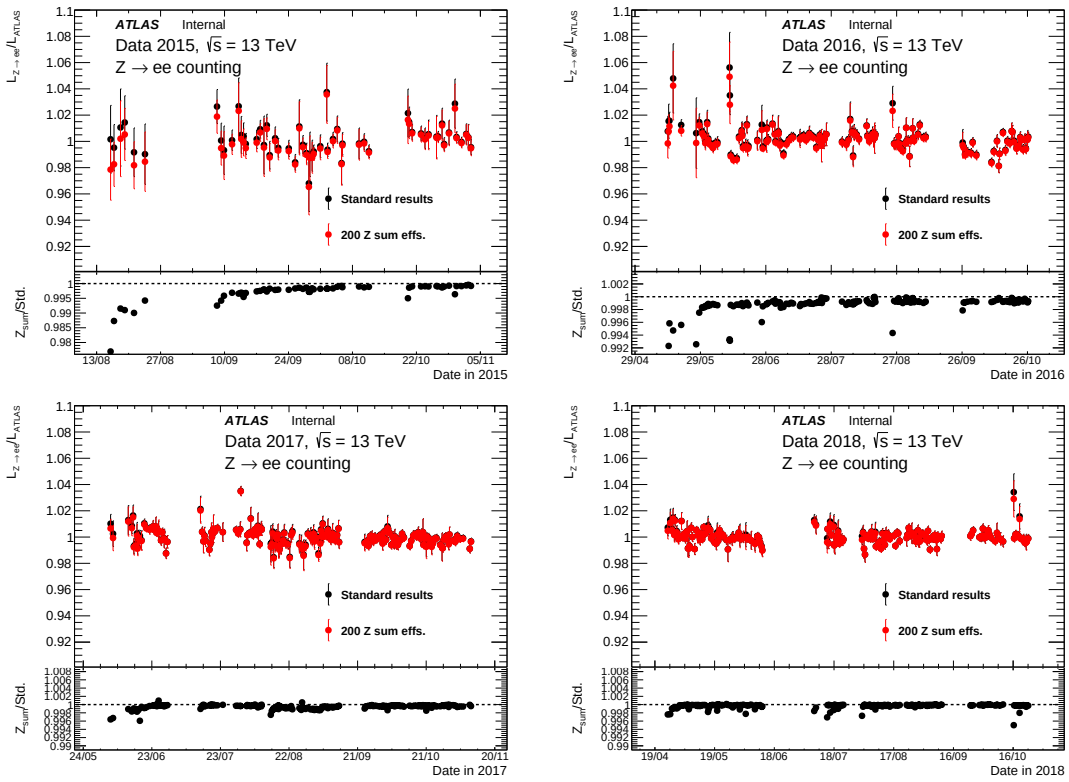


Figure 13.8: Ratio of the integrated Z-counting and baseline ATLAS luminosities per ATLAS run taken from pp collisions at $\sqrt{s} = 13$ TeV for the $Z \to e^+e^-$ channel. The Z-counting luminosity is normalised to the preliminary integrated baseline ATLAS luminosity per data-taking period [152]. Results are shown for the standard Z-counting (black) and the LB-aggregated (red) methodologies. Bottom panels show the ratio between both methodologies.

Chapter 14

Lorentz-invariance violation analysis methodology

This chapter describes the analysis methodology used to search for Lorentz-invariance violation using the Z-counting methodology. Section 14.1 presents an overview of how the effects of the Standard Model Extension, introduced in Section 2.3.2, manifest in a non-inertial frame of reference such as the ATLAS experiment. Section 14.2 presents an overview of how the methodology from the Z-counting is adapted to search for time-dependent LIV signals.

14.1 Lorentz-invariance violation at ATLAS

The measurement here presented is based on data collected by the ATLAS experiment, working in an Earth-based coordinate system which, due to the rotation of the Earth, represents a non-inertial frame. The consequences of this, such as the centrifugal force emerging from this rotation, are many orders of magnitude weaker than the gravitational force which, at the same time, is considerably weaker than the electroweak and strong forces mediating the LHC collisions. Therefore, the non-inertial effects are safe to neglect in most particle physics analyses.

However, under the presence of Lorentz-invariance violating effects, the orientation of the laboratory with respect to the SME coefficients, representing the background fields modifying the Drell-Yan cross-section as introduced in Section 2.3.2, will change over time. Note that the SME coefficients presented in that Section do not present any time dependence, but as Earth changes direction with respect to the directions of said operators over time, a time dependence arises from the Earth's rotation. Laboratory observables are therefore expected to oscillate at harmonics of the Earth's sidereal period, $T_{\oplus} = 23$ hr 56 min 4.091 sec. As each experiment testing these effects possesses different coordinates and orientation with respect to the non-isotropic background fields, it is useful to

introduce a fixed, non-rotating inertial frame. The conventional choice for this frame is the Sun-Centered Frame (SCF) [43,183,184], schematics of which are shown in Figure 14.1.

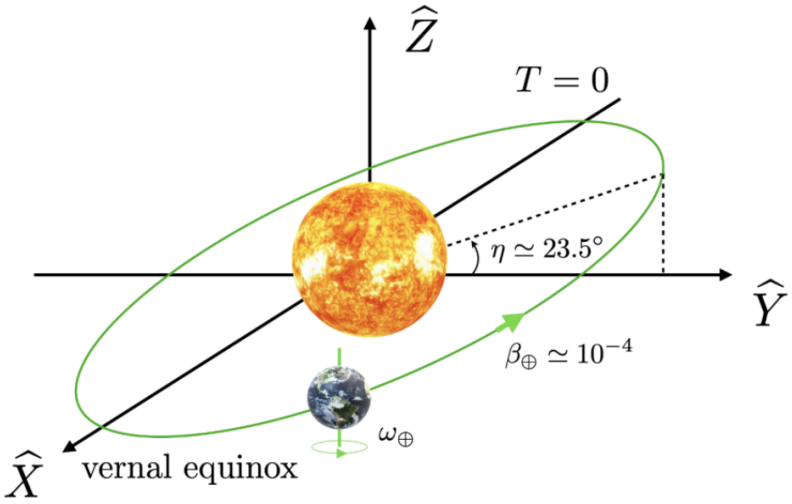


Figure 14.1: Illustration of the orbit of the Earth in the Sun Centered Frame. Adapted from Ref. [184]. η is the angle of the Earth's orbital plane with respect to the $\hat{X}$-$\hat{Y}$-plane and $\beta_\oplus$ is the mean orbital speed.

As the name conveys, the spatial axes of the SCF are centered on the Sun, with the $\hat{Z}$ axis chosen parallel to the Earth's rotation axis. The $\hat{X}$ axis points toward the year 2000 vernal equinox, when the Sun is exactly above the Equator and the $\hat{Y}$ axis completes a right-handed coordinate system with respect to the other two axes. The initial time T=0 of the coordinate system is precisely set by the 2000 vernal equinox, which happened on March 20th, 2000 at 7:35am in the Universal Coordinated Time (UTC).

A translation is needed to express the coefficients, originally given in SCF coordinates $(\hat{X}, \hat{Y}, \hat{Z})$, to the ATLAS coordinate system $(\hat{x}, \hat{y}, \hat{z})$. This can be achieved by a series of rotations, each parametrizing the change required in a given axis. Firstly, one needs to assess the rotations needed to transform the coordinates of *any* laboratory frame on Earth to the SCF. These include the inherent movement caused by the Earth's rotation, parametrized by the sidereal period ($T_\oplus$), and the time of measurement in the SCF frame, t_{SCF}. The time transformation from ATLAS timestamps to the SCF time frame is given by the time shift between the origin of both frames. The time origin in the SCF is set at the 2000 vernal equinox, while ATLAS timestamps are given in *Unix* time, which takes as time reference January 1st, 1970 at 0:00:00 UTC. Therefore, in order to express ATLAS timestamps in the time frame of the SCF, a correction is applied:

$$t(t_{\mathrm{SCF}} = 0) - t(t_{\mathrm{Unix}} = 0) = 953552124 \text{ s}. \tag{14.1}$$

Once a timestamp has been retrieved with respect to the SCF time frame it can be expressed as a sidereal phase $\Phi_\oplus$, expressed as an angle $\Phi \in [0, 2\pi]$ with respect to $T_\oplus$ as:

$$\Phi_\oplus(t_{\text{SCF}}, T_\oplus) = \frac{2\pi \cdot \text{mod}(t_{\text{SCF}}, T_\oplus)}{T_\oplus}, \tag{14.2}$$

where mod represents the modulo operation.

A rotation is needed to make $\hat{z} \cdot \hat{Z} = \cos\chi$, where χ represents the colatitude of the laboratory. For a generic laboratory located at a given colatitude χ this rotation takes the form [185]:

$$R_{\text{LAB}}(\Phi_\oplus) = \begin{pmatrix} \cos\chi\cos\Phi_\oplus & \cos\chi\sin\Phi_\oplus & -\sin\chi \\ -\sin\Phi_\oplus & \cos\Phi_\oplus & 0 \\ \sin\chi\cos\Phi_\oplus & \sin\chi\sin\Phi_\oplus & \cos\chi \end{pmatrix}. \tag{14.3}$$

Additional rotations are needed to connect the generalized expression to the ATLAS coordinate system. While the aforementioned $\hat{y}$ axis is defined as perfectly point east, the ATLAS beam line is oriented at an angle ψ north of east, so a rotation needs to be included to align both axes. Moreover, the ATLAS detector is slightly inclined along the $\hat{z}$ direction by an angle δ, which needs to be corrected for. Thus, the net rotation to transform the ATLAS coordinates to the SCF can be expressed as:

$$R_{\text{ATLAS}}(\Phi_\oplus) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\delta & -\sin\delta \\ 0 & \sin\delta & \cos\delta \end{pmatrix} \begin{pmatrix} +1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \times R_{\text{LAB}}(\Phi_\oplus) \tag{14.4}$$

The second matrix in the previous expression swaps the $\hat{z}$ and $\hat{y}$ directions, as required by the difference in the definitions of the axes between the ATLAS and SCF coordinate systems. The values of the angles[†] that define the rotation R_{ATLAS} are:

$$\chi = 43.764^\circ, \quad \psi = 168.7^\circ, \quad \delta = -0.704^\circ. \tag{14.5}$$

As presented in Equation 2.34, a simple LIV lagrangian can be expressed as a non-isotropic background field fixed under the SCF, $b^\mu_{\text{SCF}} = \left(b^{\hat{X}}, b^{\hat{Y}}, b^{\hat{Z}}, b^{t_{\text{SCF}}}\right)$. This field, as observed from the ATLAS experiment, as:

$$b^\mu = [R_{\text{ATLAS}}(\Phi_\oplus)]^\mu_\nu \, b^\nu_{\text{SCF}} \tag{14.6}$$

[†]Information provided by the Tomography and Geomatics team of the Site and Civil Engineering group of CERN.

The Drell-Yan production cross-section under LIV-inducing effects as parametrized in the SME involves combinations of coefficients as those presented in Equation 2.37. When calculating the production cross-section in the ATLAS laboratory frame, the only time-dependent coefficients are c_f^{33} and d_f^{33}, since the collision between protons happens along the z-axis of the ATLAS laboratory frame. The index f refers to the different quark flavours involved in the process [47]. When expressed in terms of the SCF coefficients using Equation 14.6, the time-dependent components of the coefficients take the form [47]:

$$\begin{aligned} c_f^{33} = &- 2c_f^{\hat{X}\hat{Z}} \sin\chi \sin\psi \left[\cos\chi \sin\psi \cos(\Phi_\oplus) + \cos\psi \sin(\Phi_\oplus)\right] \\ &- 2c_f^{\hat{Y}\hat{Z}} \sin\chi \sin\psi \left[\cos\chi \sin\psi \cos(\Phi_\oplus) + \cos\psi \sin(\Phi_\oplus)\right] \\ &+ c_f^{\hat{X}\hat{Y}} \left[(\cos^2\chi \sin^2\psi - \cos^2\chi) \sin(2\Phi_\oplus) - \cos\chi \sin(2\psi) \cos(2\Phi_\oplus)\right] \\ &+ \frac{1}{2}(c_f^{\hat{X}\hat{X}} - c_f^{\hat{Y}\hat{Y}}) \left[(\cos^2\chi \sin^2\psi - \cos^2\chi) \cos(2\Phi_\oplus) - \cos\chi \sin(2\psi) \sin(2\Phi_\oplus)\right], \end{aligned} \tag{14.7}$$

and similarly for d_f^{33}. The effect of each of these coefficients on the Drell-Yan production cross-section under the SME, as a function of sidereal time, can be expressed as:

$$\frac{\sigma_{\text{SME}}(\Phi_\oplus)}{\sigma_{\text{SM}}} = 1 + c_f^{\mu\nu} \cdot f_{\mu\nu}(\Phi_\oplus), \tag{14.8}$$

where $c_f^{\mu\nu}$ are the different SME coefficients appearing on the right hand side of Equation 14.7 and $f_{\mu\nu}$ are the corresponding sinusoidal functions depending on the ATLAS coordinates and sidereal time. The effect of these coefficients when replacing all angles for their corresponding ATLAS coordinates, as a function of sidereal phase ($\Phi_\oplus$) can be found in Figure 14.2. The plot uses an assumption of $c_f^{\mu\nu} = d_f^{\mu\nu} = 10^{-5}$ for every coefficient, considering the effects of each coefficient at a time. The increased sensitivity of those coefficients involving u-quarks with respect of those involving d-quarks is derived from the dependence of the cross section on the quark charge [47], as well as the dominance of u-quarks in the proton's PDF (see Section 3.1.1).

14.2 LIV search methodology

For the purpose of searching for Lorentz-invariance violating signals, the number of Z-bosons detected by the ATLAS detector, as well as the baseline ATLAS luminosity, are used to construct a simplified cross-section, σ, as introduced in Equation 8.4. This cross section is measured following the Z-counting methodology in each luminosity block passing the data-quality criteria. The quantities measured in each LB are associated to a unique Unix timestamp, which is converted to a sidereal phase using Equation 14.2. Using these variables, the following observable is defined:

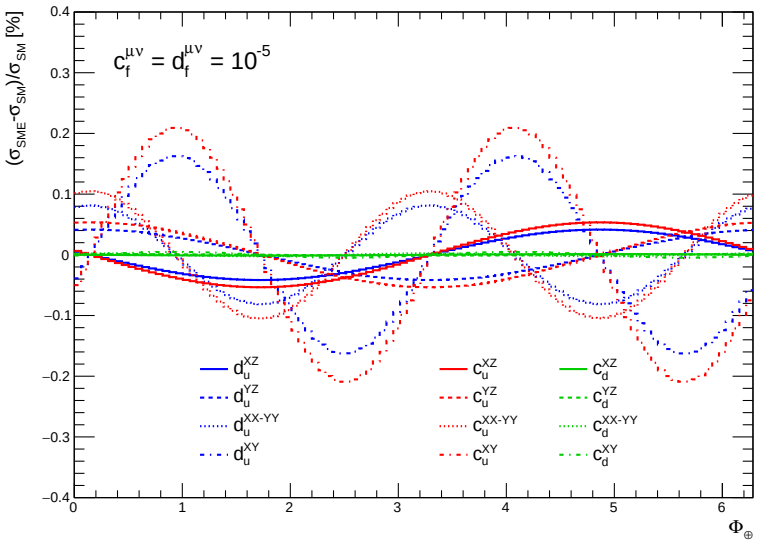


Figure 14.2: Z-boson production cross section change as a function of sidereal phase under the effects of Lorentz-invariance violating effects induced by different SME coefficients. Coefficients are assumed to be $c_f^{\mu\nu} = d_f^{\mu\nu} = 10^{-5}$, considering the effects of each coefficient at a time.

$$R_i(t) = \frac{\sigma_i(t)}{\sigma_T} = \frac{\frac{N_i}{\mathcal{L}_{\mathrm{ATLAS},i}}(t)}{\sum_i^{\mathrm{N\ bins}} \left(\frac{N_i}{\mathcal{L}_{\mathrm{ATLAS},i}}\right)}, \tag{14.9}$$

where N_i corresponds to the number of Z-bosons observed in the i-th bin of the sidereal phase distribution, and $\mathcal{L}_{\mathrm{ATLAS},i}$ is the baseline ATLAS luminosity registered in said bin. By normalising the ratio to the total number of measured Z-bosons and luminosity time-independent systematic uncertainties such as energy calibration or overall luminosity uncertainties will cancel. The analysis covers 100 phase bins, each covering ∼0.063 radians in phase and approximately 15 minutes in time, value chosen to optimise the sensitivity of the analysis [182].

An example of the effect of the sensitive Standard Model Extension operators on the defined observable as a function of sidereal phase is shown in Figure 14.2. All signal shapes introduced by these operators correspond to sinusoidal functions taking the sidereal phase $\Phi_\oplus$ (and $2 \cdot \Phi_\oplus$) as an argument, as shown in Equation 14.7. To generalise the limits extracted from the fits on the observable under study, the following functions are used for signal fitting:

$$f_1 = 1 + p_1^c \cdot \cos(\Phi_\oplus) + p_1^s \cdot \sin(\Phi_\oplus) \tag{14.10}$$

$$f_2 = 1 + p_2^c \cdot \cos(2\Phi_\oplus) + p_2^s \cdot \sin(2\Phi_\oplus) \tag{14.11}$$

$$f_3 = 1 + p_3^c \cdot \cos(3\Phi_\oplus) + p_3^s \cdot \sin(3\Phi_\oplus)\,. \tag{14.12}$$

The limits extracted on the $p^i_{1,2,3}$ $(i = c, s)$ coefficients can thus be re-interpreted to set limits on the SME coefficients $c_f^{\mu\nu}$ on which the Lorentz-invariance violating signals depend on. These functions parametrise the sinusoidal shapes induced by the SME coefficients, probing the frequencies they depend on ($\Phi_\oplus$, $2\Phi_\oplus$) and an additional frequency ($3\Phi_\oplus$) where no SME signal is expected, but can be used for systematic checks, as well as signal tests when using different probing periods. See Section 15.3 for further details.

Chapter 15

Signal sensitivity studies

This Chapter describes novel time-dependent sensitivity studies developed using the $Z \to e^+e^-$ Monte-Carlo samples. Section 15.1 includes an overview of the methodology, including a description of the samples used, the selection, and observable definitions. Section 15.2 describes the studies using data-driven time simulation to test a null signal scenario, presenting two different methodologies to do so. A study on the sensitivity to signal-like signatures, as well as data blinding strategy studies, are presented in Section 15.3.

15.1 Methodology overview

Monte-Carlo simulation of the $Z \to e^+e^-$ production allows to perform SM-only tests, where the sensitivity of the measurement to Lorentz-invariance violating effects can be tested while ensuring that no other BSM effects are in place. The MC samples used include the modelling of particle reconstruction effects in the ATLAS detector, testing the effects that these may have on the measurement. However, Monte-Carlo samples consist of event-by-event simulations, where no time-dependence is introduced. To overcome this issue different approaches have been developed, further explained in Section 15.2.

The Monte-Carlo samples used in these studies are the POWHEG+PYTHIA samples described in detail in Section 13.4. The object and event selection follow that used in the Z-counting luminosity measurement, as presented in Section 13.2. As a proof of concept, only MC samples corresponding to the 2018 dataset are used, equivalent to $\mathcal{L} = 58.5$ fb^{-1}.

The ratio defined in Equation 14.9 can be transferred to the weighted MC sample as follows:

$$R_i = \frac{\tilde{\sigma}_i}{\tilde{\sigma}_{\mathrm{T}}} = \frac{\frac{\sum \omega_i^{\square}}{\sum \mathcal{L}_{\mathrm{MC},i}}}{\sum_i \left(\frac{\sum \omega_i^{\square}}{\sum \mathcal{L}_{\mathrm{MC},i}} \right)}, \tag{15.1}$$

where:

- $\sum \omega_i^{\square}$ corresponds to the sum of weights in the bin i of those events passing the Z-counting selection.

- $\mathcal{L}_{\mathrm{MC},i}$ corresponds to the Monte-Carlo luminosity delivered in each phase bin i. It is defied as the total sum of weights delivered in the bin.

- $\tilde{\sigma}_i$ is a pseudo-cross-section built using the ratio between the sum of weights delivered to each bin i passing the event selection and the MC luminosity, used to construct the double-ratio.

However, one more effect needs to be accounted for in order to perform studies with Monte-Carlo samples at reconstructed level, since particle reconstruction effects may induce a phase dependence that needs to be corrected for. When correcting for these effects, an "unfolded" reconstructed double-ratio can be defined as follows:

$$\tilde{\sigma}_i^{\mathrm{reco,unf}} = \tilde{\sigma}_i^{\mathrm{reco}} \times \frac{\tilde{\sigma}_i^{\mathrm{truth}}}{\tilde{\sigma}_i^{\mathrm{reco}}} = \frac{\sum \omega_i^{\mathrm{reco}}}{\sum \mathcal{L}_{\mathrm{MC},i}} \times \frac{\frac{\sum \omega_i^{\mathrm{truth}}}{\sum \mathcal{L}_{\mathrm{MC},i}}}{\frac{\sum \omega_i^{\mathrm{reco}}}{\sum \mathcal{L}_{\mathrm{MC},i}}}, \tag{15.2}$$

in such way that the ratio $\tilde{\sigma}_i^{\mathrm{truth}}/\tilde{\sigma}_i^{\mathrm{reco}}$ corrects for the effects that particle reconstruction may have on the phase dependence of our observable in each bin of the distribution, performing a so-called *bin-by-bin* unfolding. This way, the central values of the unfolded reconstructed distribution will exactly match those of the truth-level distribution, but will have a bigger statistical uncertainty due to the reduced amount of events passing the reconstruction cuts.

Alternatively, the correction for particle reconstruction effects can be applied following the methodology used in the Z-counting luminosity measurement, where corrections are introduced via event-level lepton efficiencies, $\varepsilon^{\mathrm{T\&P,MC}}_{Z\to\ell^+\ell^-}$, and a Monte-Carlo correction factor, $F^{\mathrm{MC}}_{Z\to\ell^+\ell^-}$. These have previously been calculated as a function of $\langle\mu\rangle$, as shown in Figures C.3 and 13.3 respectively, and are therefore assigned to each MC event at reconstructed level depending on its simulated pileup value.

When accounting for the effects of the lepton efficiencies and MC correction factors, a pseudo-cross-section unfolded following the Z-counting methodology can be defined as follows:

$$\tilde{\sigma}_i^{\mathrm{reco}} = \frac{\sum \omega_i^{\square}}{\sum \mathcal{L}_{\mathrm{MC},i}} \longrightarrow \tilde{\sigma}_i^{\mathrm{reco,unf}} = \frac{\sum \frac{\omega_i^{\square}}{\varepsilon_i^{\mathrm{T\&P}} \cdot F_i^{\mathrm{MC}}}}{\sum \mathcal{L}_{\mathrm{MC},i}}. \tag{15.3}$$

Here, $\varepsilon_i^{\text{T\&P}}$ and F_i^{MC} are respectively the event-level efficiency and MC correction factor, calculated as a function of $\langle\mu\rangle$ as described above, which correct for particle reconstruction effects that each MC event weight assigned.

15.2 Null hypothesis sensitivity studies

This Section presents an overview of the sensitivity studies performed under the assumption of no Lorentz-invariance violating signal, i.e. a *null* hypothesis, where only SM processes are included in the simulation. To perform this analysis, a methodology was developed to introduce time dependence in the Monte-Carlo Drell-Yan samples. Initial studies have been carried out to establish the methodology assigning a uniformly distributed random phase to each MC event. The initial studies were successful, establishing a framework for phase simulation in MC samples and obtaining an initial estimate of the expected limits on the coefficients parametrising the impact of the LIV signals, $p \sim \mathcal{O}(10^{-4})$. Further details of the studies can be found in Appendix E. However, the assumption used in the initial studies does not accurately represent the conditions at which data-taking took place, so a data-driven sidereal phase assignment has been developed to represent the time profile at which data was recorded. To achieve this, the modelling of the pileup profile in the ATLAS Monte-Carlo samples is exploited. Since the number of colliding protons per bunch changed across every year of data-taking, the simulated profile of the pileup parameter can be used to introduce time-dependence in the MC. Two approaches have been developed to introduce time simulation: one based on the mapping of the pileup dependence to the sidereal phase in the ATLAS Run 2 dataset, namely the *phase-based* approach; and one based on the mapping of the timestamps recorded for each pileup value in the Run 2 dataset, namely the *timestamp-based* approach. Both approaches are used to test the sensitivity of the analysis under the null hypothesis assumption.

15.2.1 Phase-based simulation

One way time-dependent effects can be reproduced in our Monte-Carlo simulations is by studying the way pileup and sidereal phase are related in our data. The distribution that these two variables follow can be measured to transform the pileup values, as modelled in MC, into a sidereal phase distribution that corresponds to the mapping observed in data. Figure 15.1 shows the pileup-to-phase mapping observed in the ATLAS Run 2 dataset, obtained by registering each pileup and sidereal phase (retrieved from the ATLAS timestamp) entry recorded in each LB in the dataset. Each entry in the distribution is weighted by the integrated luminosity registered in the LB, meaning that the plot represents how much data was recorded for every pileup value at each sidereal phase. The figure displays non-uniform shapes, from which it can be inferred that data-taking did not take place in uniform time intervals, but instead there is an inherent phase structure that simulation

needs to account for.

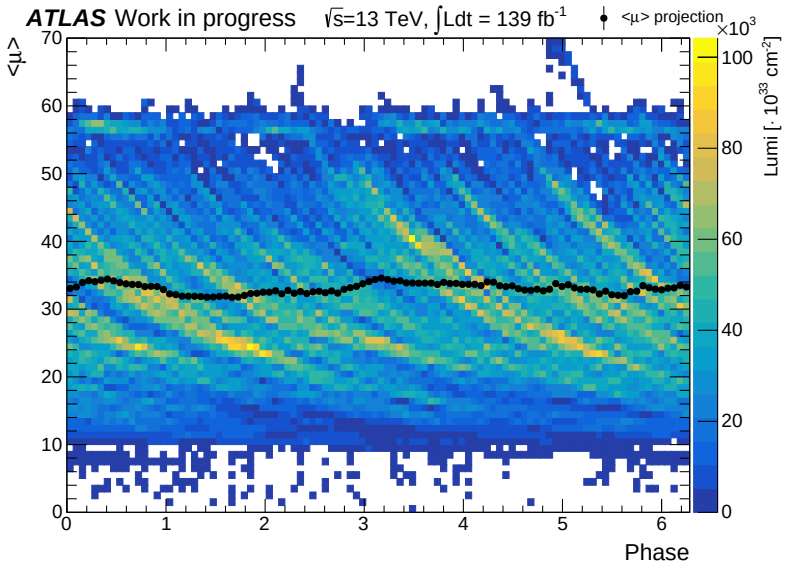


Figure 15.1: Pileup and sidereal phase profile as observed in the ATLAS Run 2 dataset. z-axis shows the total integrated luminosity registered in a given pileup-phase combination. The profile (black dots) shows the average pileup observed in each sidereal phase bin.

The distribution shown in Figure 15.1 can be used to extract the probability of finding each sidereal phase for a given pileup value, $\mathrm{P}(\Phi|\mu)$. Normalising each row in the distribution to unity, as shown in Figure 15.2, we obtain $\mathrm{P}(\Phi|\mu)$ for every pileup value registered in data. An example of a phase probability distribution is also shown in the figure, for a reference value of $\langle\mu\rangle = 30$. Using this distribution, one can assign a sidereal phase to every MC event that follows the pileup-to-phase mapping observed in data in the following manner:

1. Obtain the μ value simulated in the MC event.
2. Read the data phase-pileup mapping, $\mathrm{P}(\Phi|\mu)$, from Figure 15.2 for said μ value.
3. Generate a uniformly distributed random number between 0 and 2π, the probing random phase X.
4. Generate a uniformly distributed random number between 0 and 1, the probing random probability Y.
5. If $\mathrm{P}(X|\mu) > Y$, X is the sidereal phase that the MC event gets assigned.
6. If $\mathrm{P}(X|\mu) < Y$, X is discarded and the process is repeated from step 3 until a valid phase is assigned.

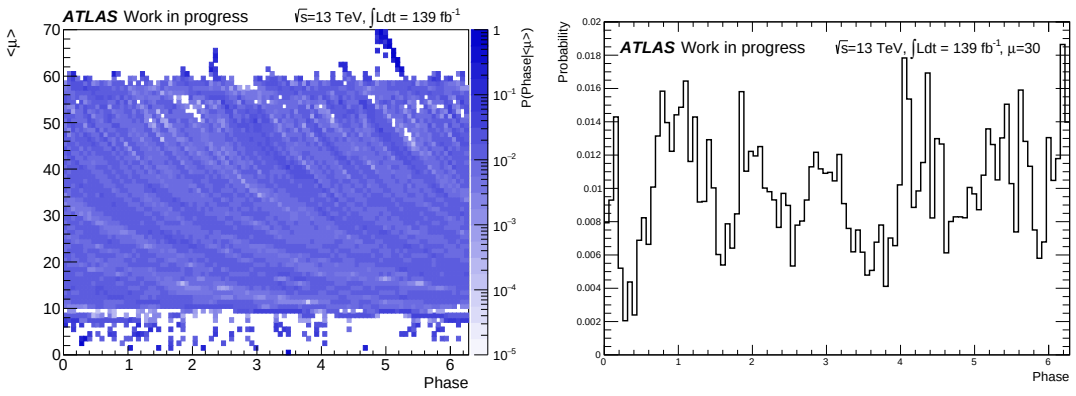


Figure 15.2: (Left) Row-normalised probability distribution of the pileup-phase mapping from the full Run 2 dataset (Figure 15.1). (Right) Phase probability distribution for an example pileup value of 30, $P(\Phi|\mu = 30)$.

Results

When using the assignment of phase-based time simulation to each MC event, one obtains the pileup-phase distribution found in Figure 15.3, showing the same patterns as observed in data (Figure 15.1) which indicates the mapping of both variables is successfully simulated. Figure 15.4 shows the dependence of the distribution of the pseudo-cross-section (defined in Equation 15.2) against sidereal phase. The ratio shown in Figure 15.4 corresponds to the bin-by-bin correction applied to unfold the reconstructed-level distribution to truth-level. These are found to vary across different phase bins at a $\sim 1\%$ level, stressing the importance of this correction in the analysis.

Figure 15.5 shows the double-ratio obtained when using the phase-based time simulation in the MC samples, where both the truth-level and unfolded distributions are displayed. By definition, the central points of the bin-by-bin unfolded distribution exactly match the truth-level ones, the key difference between both distributions being the increased error bars observed in the unfolded distribution. This is due to the lower number of events passing the reconstruction requirements with respect to the total number of generated events passing the kinematic requirements. The distribution obtained when unfolding following the Z-counting methodology displays a larger spread than the other distributions, but all three distributions are compatible with the fit of a straight line, corresponding to the assumption of lack of LIV effects.

The assignment of a sidereal phase to each MC event depending on the simulated pileup parameter is a randomised process. Therefore, processing the same MC samples under different initial random seed choices ruling the randomised phase assignment leads to different distributions. To minimize the impact of this random seed choice the process is repeated multiple times. The final best fit on the coefficients is taken as the mean of the

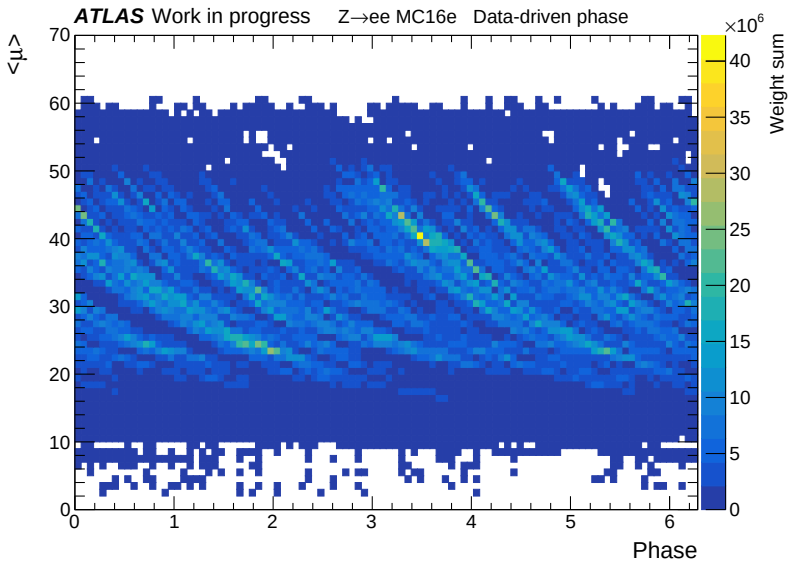


Figure 15.3: Pileup and sidereal phase profile obtained using the phase-based time simulation. The z-axis shows the sum of weights registered in a given pileup-phase combination. Results correspond to MC samples describing the 2018 data profile.

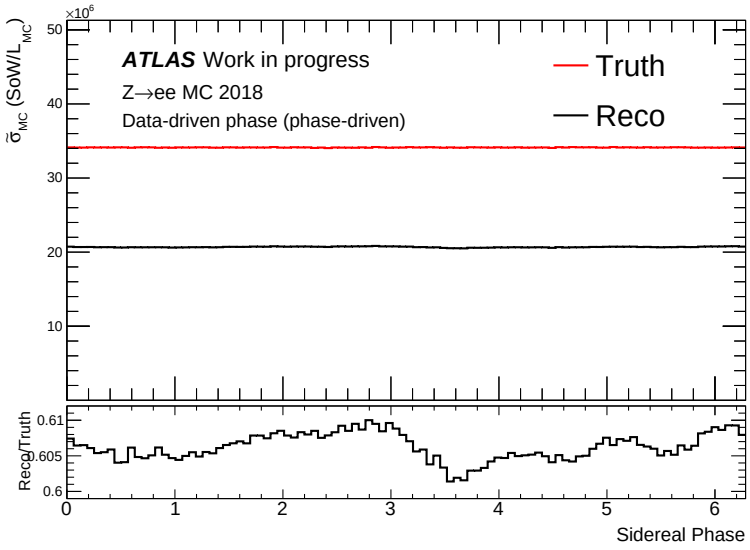


Figure 15.4: Monte-Carlo pseudo-cross-section distribution per sidereal bin obtained using the phase-based time simulation. Distributions are shown at truth (red) and reconstructed (black) level. Results correspond to MC samples describing the 2018 data profile.

results obtained using 100 different random seeds, while the limits set on the parameters correspond to the standard deviation of the distribution of the fit results.

With this simulated observable, and repeating the randomisation process as stated above, bounds on the $p^i_{1,2,3}$ $(i = c, s)$ coefficients can be obtained by performing χ^2 fits (details of which can be found in Ref. [182]) on the signal distributions shown in Equa-

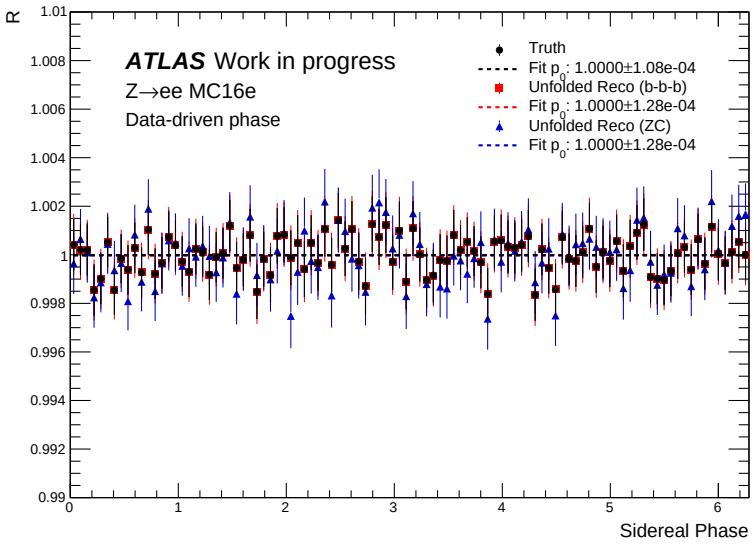
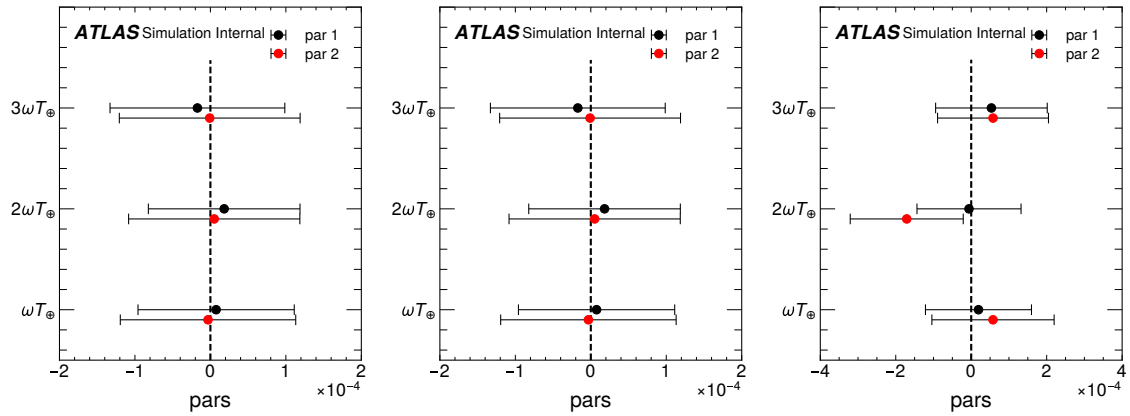


Figure 15.5: Double-ratio obtained using a phase-based time simulation. Figure shows truth-level (black circle) and unfolded distributions. Unfolding performed in a bin-by-bin basis (red square) and following the Z-counting methodology (blue triangle). The fits (dashed lines) correspond to a straight line fit on each of the distributions.

tion 14.10, fitting one function at a time. Figure 15.6 shows the results corresponding to the signal fitting on the truth and reconstructed level distributions from phase-based time simulation. All coefficient limits are compatible with a null hypothesis, with limits at the $\mathcal{O}(10^{-4})$ level. Bounds obtained with the distribution unfolded using the Z-counting methodology, see Equation 15.3, are slightly bigger than those unfolded with the bin-by-bin methodology. A deviation from 0 is observed for the p_2^s coefficient. Though not statistically significant, this deviation may hint at spurious signals induced by systematic effects in the methodology that have to be further studied by the team. One of the possible reasons may be the use of MC corresponding only to the 2018 dataset. The pileup profile varies across each year of data-taking, meaning that a complete sensitivity study will require the application of MC samples describing the entirety of the Run 2 dataset for a correct application of the mapping presented in Figure 15.1. Previous bounds on SME coefficients using ZEUS data [176] translate to limits on p-parameters of the order of $p \sim 2 \cdot 10^{-3}$, meaning that the constraints placed via Drell-Yan production at ATLAS could improve the existing limits by almost an order of magnitude, besides exploring sensitivities to many new coefficients due to the difference in initial states.

15.2.2 Timestamp-based simulation

An alternative to simulate time profile in our Monte-Carlo simulations lies in the mapping of all timestamps associated to each pileup value in the Run 2 ATLAS dataset. To account for the differences in data acquisition across the dataset, the timestamp of each luminosity



(a) Truth-level (b) Bin-by-bin unfolding (c) Z-counting unfolding

Figure 15.6: 68% CL coefficient bounds extracted using the double-ratio distributions at truth-level (left), bin-by-bin unfolded (middle) and Z-counting unfolded (right), with phase-based time simulation. Bounds are obtained using the distributions produced with 100 different random seeds. Limits on the parameters ("pars") are shown for the cosine ("par1", black) and sine ("par2", red) terms of the signal parametrisation. Plots provided by Dr. Yiming Abulaiti[†].

block $t_i(\mu)$ is associated to a weight that corrects for the luminosity that was registered in the luminosity block corresponding to said timestamp, $\mathcal{L}_i$. The timestamp of a luminosity block is defined as the mean between its start and end times. The weight applied takes the expression:

$$\omega_i(\mu) = \frac{\mathcal{L}_i}{\sum_{i=0}^{N} \mathcal{L}_i} \cdot N(\mu), \tag{15.4}$$

where $N(\mu)$ is the number of timestamps corresponding to each pileup value, μ. Thus, the time-profile observed in data can be reproduced by using an approximation of the Inverse Probability Transformation (IPT) method [186]:

1. Obtain the μ value simulated in the MC event.
2. Generate a uniformly distributed random integer number, X, between 0 and $N(\mu)$.
3. The timestamp associated to the MC event will be the X-th entry in the timestamp list of the pileup value μ, $t_X(\mu)$.
4. Weight the MC event by the corresponding probability weight $\omega_X(\mu)$ in addition to all original MC event weights.

Results

When using the assignment of timestamp-based time simulation to each MC event, one obtains the pileup-phase distribution found in Figure 15.7, showing excellent agreement

[†]Contact: yiming.abulaiti@cern.ch

with the phase-based approach as shown in Figure 15.8. The same dependence is also observed for the correction of reconstructed- to truth-level distributions across phase, shown in Figure 15.4.

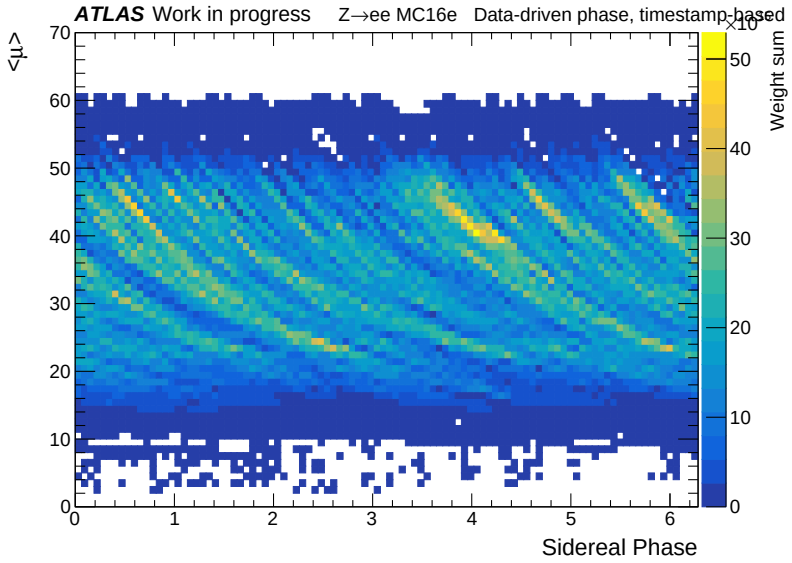


Figure 15.7: Pileup vs sidereal phase profile obtained using the timestamp-based time simulation. The z-axis shows the sum of weights registered in a given pileup-phase combination. Results correspond to MC samples describing the 2018 data profile.

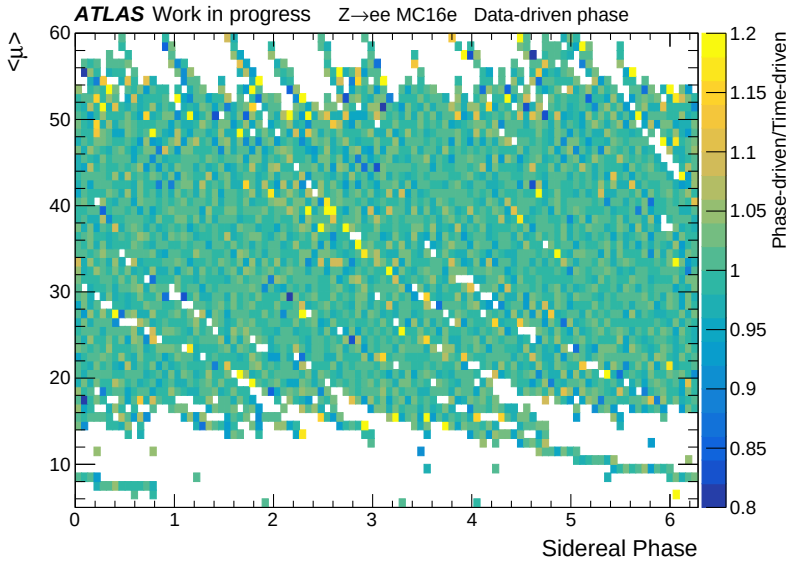


Figure 15.8: Ratio per bin of the phase-pileup profiles obtained using the phase- and timestamp-based time simulation.

Figure 15.9 shows the double-ratio that one obtains when using the timestamp-based time simulation on the MC samples. Both the truth-level and unfolded distributions are

displayed on the Figure. Similar trends are observed as shown for the phase-based results.

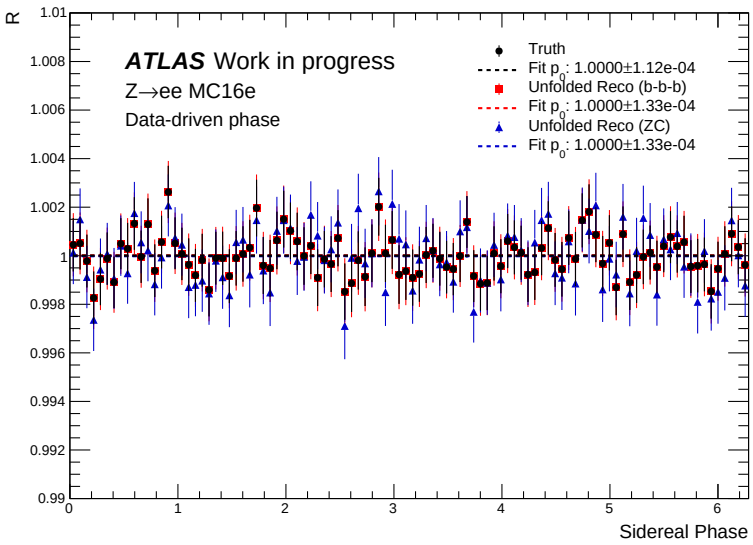
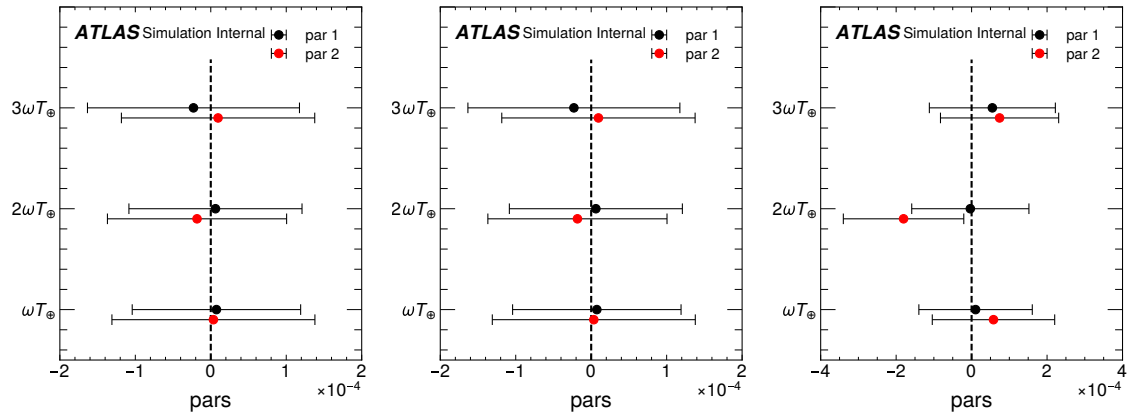


Figure 15.9: Double-ratio obtained using the timestamp-based time simulation. Figure shows truth-level (black circle) and unfolded distributions. Unfolding is performed in a bin-by-bin basis (red square) and following the Z-counting methodology (blue triangle). The fits (dashed lines) correspond to a straight line fit on each of the distributions.

With this simulated observable, bounds on the coefficients are obtained in the same manner as previously described. Figure 15.10 shows the results corresponding to the signal fitting on the truth and reconstructed level distributions using the timestamp-based time simulation. A summary of the fits obtained with both time simulation methodologies can be found in Table 15.1. Although the fits performed on the distributions obtained using the timestamp-based time simulation result into slightly bigger limits ($\sim 10\%$), the results obtained with both simulation techniques are in good agreement.



(a) Truth-level (b) Bin-by-bin unfolding (c) Z-counting unfolding

Figure 15.10: 68% CL coefficient bounds extracted using the double-ratio distributions at truth-level (left), bin-by-bin unfolded (middle) and Z-counting unfolded (right), with timestamp-based time simulation. Bounds are obtained using the distributions produced with 100 different random seeds. Limits on the parameters ("pars") are shown for the cosine ("par1", black) and sine ("par2", red) terms of the signal parametrisation. Plots provided by Dr. Yiming Abulaiti.

Parameters ($\cdot 10^{-4}$)	Simulation					
	Phase-based			Timestamp-based		
	Truth	Binned unfold	Z-count. unfold	Truth	Binned unfold	Z-count. unfold
p_1^c	0.08 ± 1.03	0.08 ± 1.03	0.19 ± 1.4	0.07 ± 1.11	0.07 ± 1.11	0.10 ± 1.50
p_1^s	-0.03 ± 1.16	-0.03 ± 1.16	0.58 ± 1.6	0.03 ± 1.34	0.03 ± 1.34	0.58 ± 1.62
p_2^c	0.18 ± 1.00	0.18 ± 1.00	-0.06 ± 1.4	0.06 ± 1.14	0.06 ± 1.14	-0.03 ± 1.55
p_2^s	0.05 ± 1.13	0.05 ± 1.13	-1.70 ± 1.50	-0.18 ± 1.18	-0.18 ± 1.18	-1.80 ± 1.60
p_3^c	-0.17 ± 1.16	-0.17 ± 1.16	0.53 ± 1.48	-0.22 ± 1.4	-0.22 ± 1.4	0.55 ± 1.67
p_3^s	-0.01 ± 1.20	-0.01 ± 1.20	0.58 ± 1.47	0.10 ± 1.28	0.10 ± 1.28	0.74 ± 1.57

Table 15.1: Summary of the best fit and 68% CL expected limits on the coefficients parametrising the Lorentz-invariance violating signals for both time-simulation methodologies presented in this thesis. Results are shown for fits using distributions at truth- and unfolded reconstructed-level. Note that a factor 10^{-4} precedes all results in the table, not shown for display purposes.

15.3 Signal injection tests

The methodologies established to simulate time-dependence in our MC samples can be used to test the sensitivity to signal-like signatures. They also allow to investigate ways to blind the presence of possible signal in data, in such way that the team can develop the data analysis methodology without biases towards signal-like shapes. All results presented in the following section correspond to timestamp-based time simulation, for reasons presented below. Under the presence of Lorentz-invariance violating (LIV) effects, the total number of Z-bosons detected per sidereal phase, $N_{\mathrm{T}}(\Phi_{\oplus})$, accounts for those generated via SM processes (N_{SM}) and those generated via LIV interactions (N_{LIV}), which can be expressed as:

$$N_{\mathrm{T}}(\Phi_{\oplus}) = N_{\mathrm{SM}} + N_{\mathrm{LIV}} = \mathcal{L}(\Phi_{\oplus}) \cdot \sigma_{SM} \cdot \left[1 + c_f^{\mu\nu} \cdot f_{\mu\nu}(\Phi_{\oplus})\right], \tag{15.5}$$

where $\mathcal{L}(\Phi_{\oplus})$ is the luminosity recorded (or sum of weights generated in the MC sample) in said sidereal phase, σ_{SM} is the Standard Model Z-boson production cross-section (1970 pb, see Section 13.1) and $c_f^{\mu\nu} \cdot f_{\mu\nu}(\Phi_{\oplus})$ are the different SME coefficients and associated sinusoidal functions, introduced in Equation 14.7. In the case of the signal injection studies here presented, the four d-type coefficients are considered as a benchmark signal to test the methodology, since they produce the biggest oscillations:

$$d_u^{XZ} : f_{XZ} = 6.28 \cdot \cos(\Phi_{\oplus}) - 41.05 \cdot \sin(\Phi_{\oplus}) \tag{15.6}$$

$$d_u^{YZ} : f_{YZ} = 41.05 \cdot \cos(\Phi_{\oplus}) + 6.28 \cdot \sin(\Phi_{\oplus}) \tag{15.7}$$

$$\left(d_u^{XX} - d_u^{YY}\right) : f_{XX-YY} = 77.61 \cdot \cos(2\Phi_{\oplus}) + 24.31 \cdot \sin(2\Phi_{\oplus}) \tag{15.8}$$

$$d_u^{XY} : f_{XY} = -48.63 \cdot \cos(2\Phi_{\oplus}) + 155.21 \cdot \sin(2\Phi_{\oplus}). \tag{15.9}$$

The effects of said operators, using $d_u^{\mu\nu} = 10^{-3}$ as a benchmark signal strength, on the observable under study are shown in Figure 15.11. The factors parametrising each sinusoidal function differ between different coefficients due to the way rotations from SCF and ATLAS frames affect each coefficient, as presented in Equation 14.7. Therefore, even if all coefficients are set to the same benchmark value, $d_u^{\mu\nu} = 10^{-3}$, they result into different signal strengths. Note that the parameter values chosen for this test result into big effects, with signal strengths of up to $\sigma_{\mathrm{LIV}}/\sigma_{\mathrm{SM}} = 15.5\%$ in the case of d_u^{XY}. Such effects are bigger than those expected, for they would produce fluctuations that would have been observed by analysis such as the Z-counting luminosity measurement [179], but serve as a hypothesis to test the methodology with.

As shown in Figure 15.11, the effects of unfolding using the Z-counting methodology under the presence of LIV signal shapes are very small, showing excellent agreement with

the distributions obtained at truth-level and unfolding using a bin-by-bin correction factor. These distributions are used to set limits on the $p^i_{1,2,3}$ $(i = c, s)$ coefficients as described in Section 15.2.1. Results on these fits can be found in Figure 15.12. The signal parametrisation correctly reflects the signal being injected in each of the benchmark distributions, with those signal shapes in which the sinusoidal terms depend on $\Phi_\oplus$, i.e. d_u^{XZ} and d_u^{YZ}, obtaining limits compatible with $p^i_1 \neq 0$, while the other terms fulfil $p^i_{2,3} = 0$. In the cases where the signal injected is parametrised as a function of $2\Phi_\oplus$ the limits obtained correspond to $p^i_2 \neq 0$, while the other terms are compatible with $p^i_{1,3} = 0$. The best fit values obtained for each parameter match the prefactors of the sine and cosine functions introduced by the different SME coefficients.

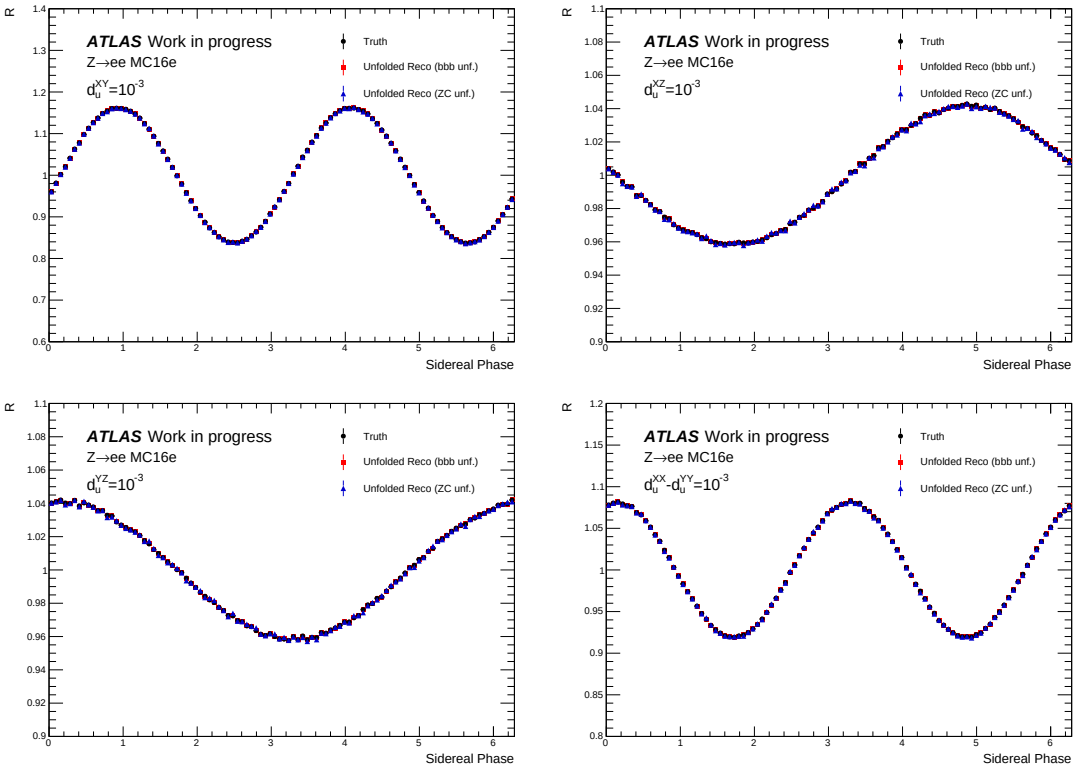
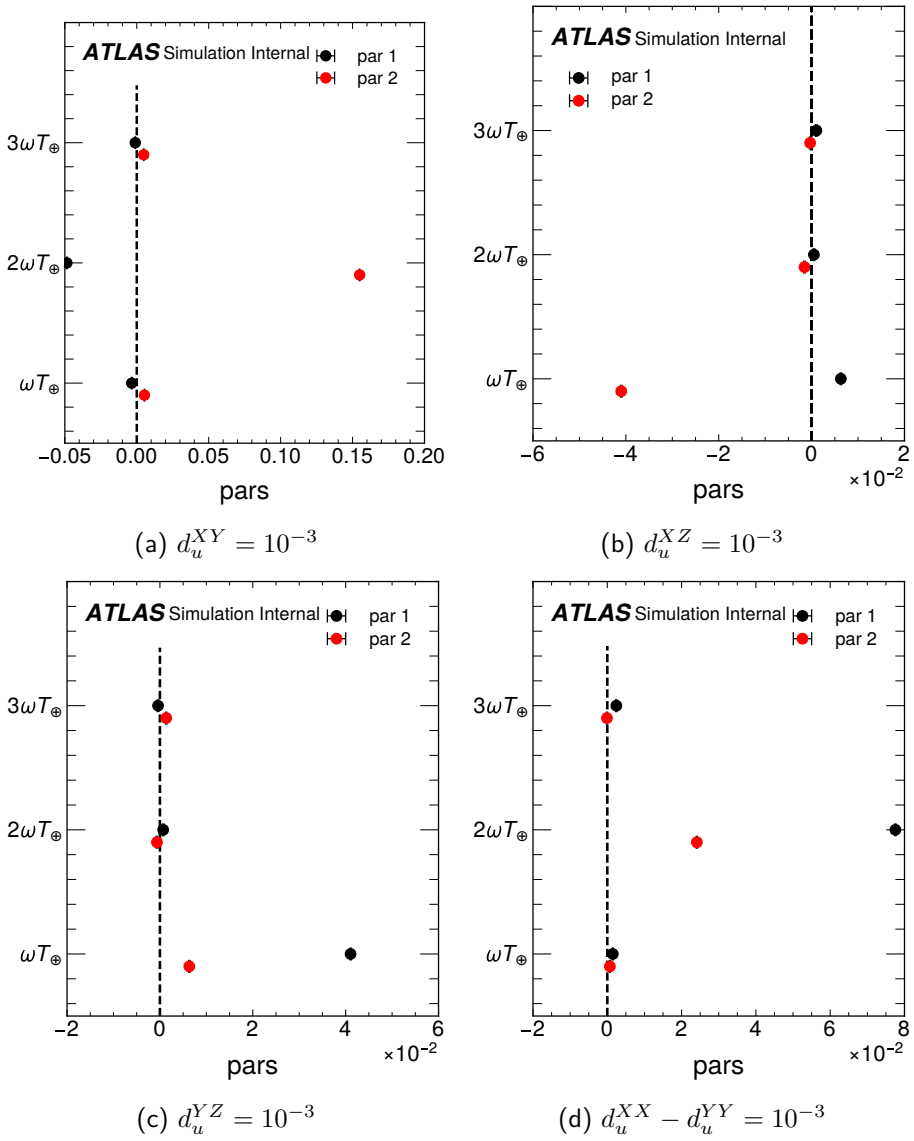


Figure 15.11: Double-ratio obtained using the timestamp-based time simulation under the presence of different SME coefficients. Figure shows truth-level (black circles) and unfolded distributions. Unfolding is performed in a bin-by-bin basis (red squares) and following the Z-counting methodology (blue triangles).

The development of the timestamp-based simulation allows to explore ways to blind the results to the presence of signal-like features in data, which would allow to develop the analysis strategy avoiding biases towards possible signal discoveries. A way to blind the data to the effect of such signals is to transform the timestamps into phase assuming a rotation period different than $T_\oplus$, since SME signals are expected to appear as resonances of the sidereal phase. To implement this, the timestamp assigned to the MC event, t, can



(a) $d_u^{XY} = 10^{-3}$ (b) $d_u^{XZ} = 10^{-3}$

(c) $d_u^{YZ} = 10^{-3}$ (d) $d_u^{XX} - d_u^{YY} = 10^{-3}$

Figure 15.12: 68% CL coefficient bounds extracted using the signal-injected double-ratio distributions using unfolding following the Z-counting methodology. Bounds are obtained using the distributions produced with 100 different random seeds. Limits on the parameters ("pars") are shown for the cosine ("par1", black) and sine ("par2", red) terms of the signal parametrisation. Plots provided by Dr. Yiming Abulaiti.

be transformed into an alternative phase, assuming a different rotation period than the sidereal phase, in a generalisation of Equation 14.2 to accommodate for different rotation period (T_i) assumptions:

$$\Phi_i(t, T_i) = \frac{2\pi \cdot \mathrm{mod}(t, T_i)}{T_i}. \tag{15.10}$$

Thus, the term $\mathcal{L}(\Phi_i)$ in Equation 15.5 depends on this re-defined phase, while SME-induced signals always depend on the sidereal phase, $c_f^{\mu\nu} \cdot f_{\mu\nu}(\Phi_\oplus)$. The considered probing

Earth rotation periods are as follows: T_{solar}=24 hours, corresponding to the duration of a solar day; T_7=7 hours, corresponding to the average duration of an LHC run; and T_1=1 hour, a fast rotation period proposed to ensure blinding to signal-like features. Note that the assumption of different probing periods cannot be achieved in the so-called phase-based time simulation presented in the previous Section, since the choice of period is already fixed by assigning MC events a data-driven sidereal phase choice, $\Phi_{\oplus}$, that cannot be transformed into other periods. Therefore, all results presented in this Section correspond to timestamp-based results.

The effects of the operator $d_u^{XY} = 10^{-3}$ on the observable under study for the different rotation period assumptions are shown in Figure 15.13. It can be inferred that the assumption of a 24-hour rotation period does not completely cancel signal oscillations, possibly due to its similarity with the sidereal period at which signal resonances occur. The assumption of a 7-hour rotation period dampens the induced signal effects, but a clear oscillation can still be distinguished. The only complete blinding to signal achieved with the considered probing periods is observed using $T_1 = 1$ hour. These distributions are used to set limits on the $p^i_{1,2,3}$ $(i = c, s)$ coefficients as described in the previous Section. Results on these fits can be found in Figure 15.14. Although the assumption of different probing periods results into smaller limits of the parameters in the cases of the T_{solar} and T_7 assumptions, as a consequence of signal dampening under wrong rotation periods, the results obtained are still compatible with signal presence. Results obtained in the case of the T_1 are compatible with a null signal hypothesis, with all parameters being compatible with $p = 0$ and similar constraints to those obtained in the no-signal scenario introduced in the previous Section. Therefore, validation studies can be safely performed on data to test the analysis methodology under this assumption, ensuring that signal presence would disappear and, once the analysis is known to be robust, changing the Earth rotation period assumed to ensure signal disappearance to the sidereal period $T_{\oplus}$ at which LIV signals are expected to appear. However, due to time constraints on the development of this thesis the results of the analysis on ATLAS data could not be included as part of this work.

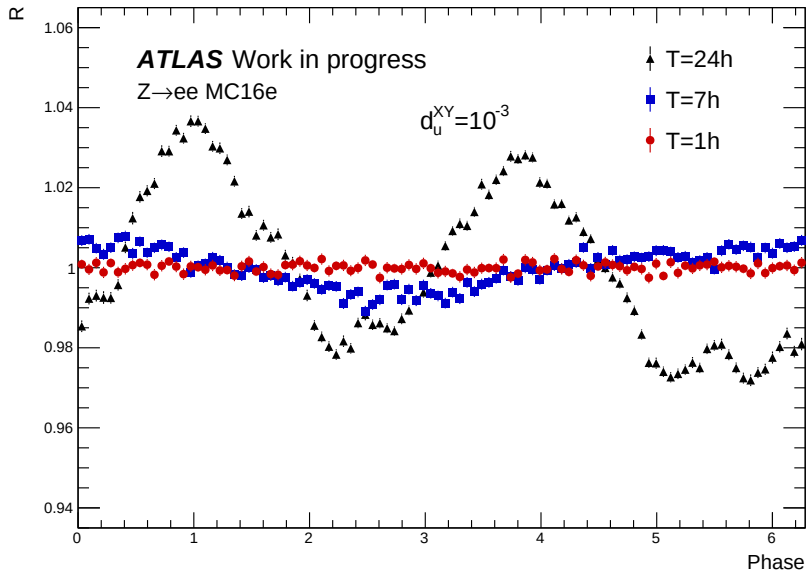
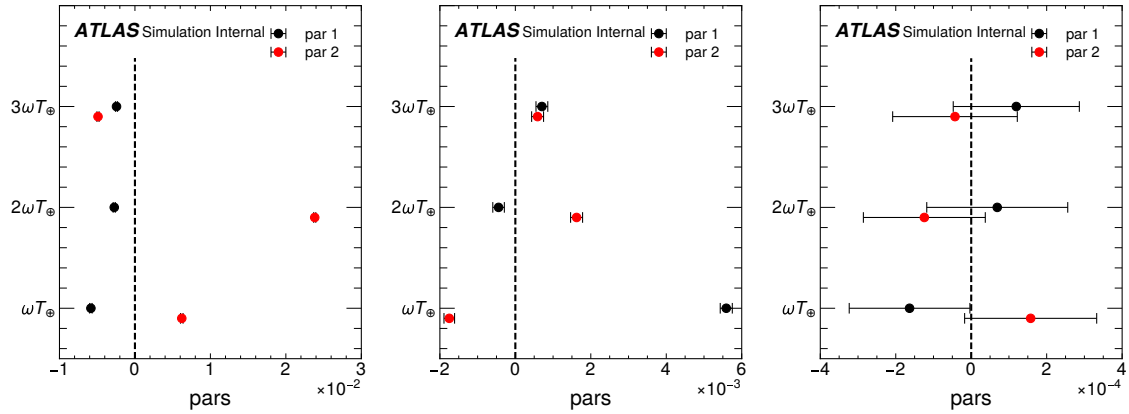


Figure 15.13: Double-ratio obtained using the timestamp-based time simulation under the presence of $d_u^{XY} = 10^{-3}$ signal for different probing period assumptions: 24 hours (black triangles), 7 hours (blue squares) and 1 hour (red circles). Figure shows unfolded distributions. Unfolding is performed in following the Z-counting methodology.



(a) T_{solar} =24 hours. (b) T_7 =7 hours (c) T_1 =1 hours

Figure 15.14: 68% CL coefficient bounds extracted using the $d_u^{XY} = 10^{-3}$ signal-injected double-ratio distributions using T_{solar}=24 hours (left), T_7=7 hours (middle) and T_1=1 hour (right), with timestamp-based time simulation. The distributions used correspond to unfolding following the Z-counting methodology. Bounds are obtained using the distributions produced with 100 different random seeds. Limits on the parameters ("pars") are shown for the cosine ("par1", black) and sine ("par2", red) terms of the signal parametrisation. Plots provided by Dr. Yiming Abulaiti.

Chapter 16

Conclusion

The study of the Z/γ^* production via neutral-current Drell-Yan in pp collisions offers a way to explore the quark sector of the Standard Model Extension that has not been probed before. This can be achieved by measuring the production rate of this process across the LHC Run 2 data-taking period which, under Standard Model assumptions, should not display any time-dependence. Deviations from this behaviour can be therefore interpreted in terms of the effect of operators inducing Lorentz-invariance violation (LIV) due to the presence of non-isotropic background fields. Such search would correspond to the first of its kind at an LHC experiment, demonstrating the potential of novel search techniques for Beyond Standard Model phenomena.

The methodology developed for luminosity measurement at ATLAS by counting the number of detected Z-bosons is suitable for such searches, monitoring the production rate of the neutral-current Drell-Yan per time interval of an average duration of 60 seconds. The methodology was successfully transferred to a newly developed framework, reproducing the original results [179] in the electron channel with an improved stability with respect to the preliminary nominal ATLAS luminosity. The new framework was used to assess the impact of statistical biases on ATLAS runs where the number of Z-boson events per luminosity block was low. These luminosity blocks present a bias in the measurement of the tag-and-probe lepton efficiencies, affecting the luminosity measurement as a whole, which can be minimised by aggregating data from neighbouring luminosity blocks. Corrections for this effect amount up to 2% in specific runs, but the overall correction on the results over the Run 2 dataset were found to be small (0.02%).

To obtain an estimate on the expected limits on the coefficients parametrising the signal shapes, a methodology for time simulation in the ATLAS Monte Carlo samples was developed. The simulation consists of a data-driven assignment of time by studying the relationship between pileup, which is known to be accurately simulated in the samples, and timestamps in the ATLAS dataset. Using this newly developed methodology, expected

limits of $\mathcal{O}(10^{-4})$ were found, improving the best existing limits obtained with ZEUS data by an order of magnitude. Moreover, with respect to this analysis, dilepton production at ATLAS offers sensitivity to new coefficients in the quark sector of Standard Model Extension EFT.

Additionally, tests were performed to explore ways to blind the measured data to LIV signal, as well as test systematic effects on the measurement. By assuming an Earth rotation period (T_i) different than that in which LIV signals are expected to appear ($T_\oplus$), a dampening of the signal strength can be achieved. This offers a way to develop the analysis methodology in data, introducing the assumption of a rotation period in which signals are expected to disappear (eg. $T_1 = 1$ hour) and changing the assumption for the rotation period in later stages of the analysis once the final results are ready to be retrieved.

The methodology here presented will be continue to be used by the analysis team to explore additional ways to blind the measurement to signal presence, as well as to assess the impact of time-dependent systematic effects on the measurement, such as pileup dependence or luminosity-related uncertainties (year-by-year calibrations, use of different luminometers...). The analysis also aims to be a proof of principle for time-dependent analyses at the ATLAS experiment, hoping to expand the methodology for such searches to additional final states.

Summary

This thesis presents two analyses studying the neutral-current Drell-Yan production at the ATLAS experiment using LHC Run 2 pp collision data collected during the years 2015 to 2018 at $\sqrt{s} = 13$ TeV, in the context of a precision Standard Model measurement and novel BSM interpretations.

The measurement of the single- and double-differential neutral-current Drell-Yan cross section was performed, reaching precisions below 1% in the phase space of low $m_{\ell\ell}$ ($<$ 230 GeV), where the leading contributions to the total uncertainty correspond to systematic effects. The measured cross-section was compared with theoretical predictions using different PDFs, observing good agreements overall but revealing the potential for better constraints using the data here presented. The analysis covered dilepton invariant masses of up to $m_{\ell\ell} = 5000$ GeV, expanding the kinematic range of the measurement with respect to previous ATLAS results. The measurement was used for tests on Lepton Flavour Universality (LFU) and Effective Field Theory (EFT) interpretations. No significant deviations from the Standard Model were found, but small tensions were found in LFU tests at $m_{\ell\ell} > 1500$ GeV. Limits placed on EFT operators correspond to leading constraints in some of the operators.

Additionally, the Z/γ^* production was monitored to explore the quark sector of the Standard Model Extension that has not been probed before, searching for the effect of operators inducing Lorentz-invariance violation (LIV) due to the presence of non-isotropic background fields. The methodology developed for luminosity measurement at ATLAS by counting the number of detected Z-bosons was successfully transferred to a newly developed framework, used to check the consistency of the measurement in a standard ATLAS analysis framework and to assess the impact of time-dependent systematic effects. The observed time-dependent ratio of the normalised Z-counting luminosity to the ATLAS nominal luminosity showed a spread of 0.66% across the entire Run 2 dataset, showing the excellent time stability of the luminosity measurement with respect to the nominal value. Moreover, novel techniques were developed for the simulation of time-dependent observables in the Monte-Carlo simulation, with which sensitivity studies were performed, both in the context of a null signal hypothesis and LIV signal presence. Expected limits

of $\mathcal{O}(10^{-4})$ on the coefficients parametrising the LIV signal were found. Tests were also performed to explore ways to blind the measurement to LIV signals, finding ways to develop the analysis methodology in data avoiding biases towards signal-like findings. Introducing the assumption of certain Earth rotation periods (e.g. $T_1 = 1$ hour) LIV signals are found to disappear, allowing to safely develop the analysis and changing the assumption for the rotation period in later stages once the final results are ready to be retrieved.

Appendices

Appendix A

Fake electron background estimation using the template method

This Appendix explains the template method used for an alternative estimation of the fake dielectron background, following the methodology presented in Ref. [187], commonly used for fake lepton background estimation in invariant masses close to the Z-boson mass resonance. In this method, a data-driven approach is used, deriving a multi-jet enhanced selection (also referred to as *background template*), where kinematic cuts are the same as those used in the signal region (presented in Section 8.2), assuming that the kinematics of the Z-boson remains unaffected, but the lepton identification is loosened. Under the assumption that the shape of the background follows that of the signal selection, the goal of the method is to estimate the factor that normalizes the template to the corresponding number of multi-jet events passing the signal selection. Different ID requirements are studied to build the background template, shown in Table A.1. In addition to the selection criteria shown in the table, a same-sign cut can be applied on the template selection, suppressing the remaining signal contamination in our background. Templates containing this cut will be labelled as "nC".

The fit to extract the scale factor needed to derive the number of multi-jet events from our background template has to be performed in a background-dominated distribution, where we can properly scale our background template to the number of observed data events, without suffering from signal contamination. Since most electron candidates coming from mis-identified jets are poorly isolated, the chosen variable to extract the signal contamination was:

$$I_{\text{min}} = \min\left(E_{\text{T}}^{\text{cone20}}/E_{\text{T}}\right), \tag{A.1}$$

Selection	ID	Reverse ID	Isolation
Template	-	Fail Loose	-
	VeryLoose	Fail Medium	
Signal	Medium	-	-

Table A.1: Selection criteria to derive the template and signal samples. The template selection with no cut on ID will also be referred to as "TriggerOnly".

that is, the minimum of the isolation of the two electron candidates in the event. The distribution has a large peak at 0, containing all signal electrons, and shows a tail at positive values, caused by the presence of fake electrons in data, as shown in Figure A.1. Alternative template selections are shown in the Figure, adding a same-sign charge cut and removing the d_0 cut from the selection. The latter was later found to have a small impact on the background template and hence removed from the possible variations.

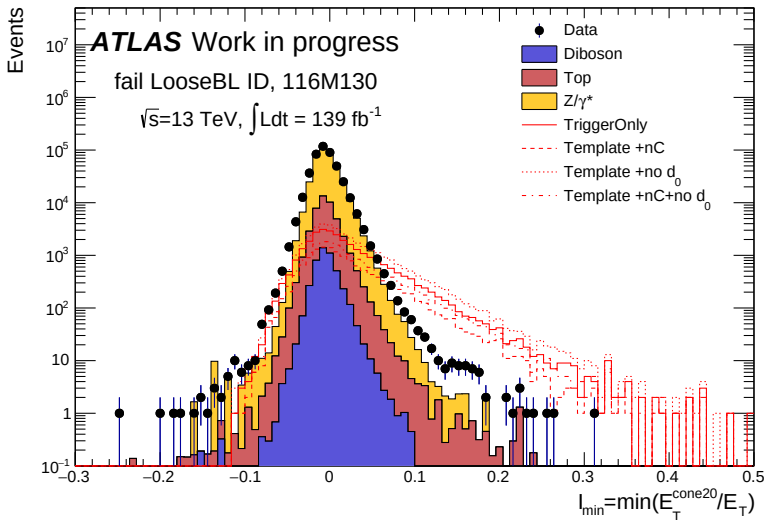


Figure A.1: I_{min} distribution for signal in data and MC samples. Background template selection (no ID cut) also shown, considering different variations by applying same-sign charge cuts ("nC") and dropping the d_0 cut. The results correspond to a mass window of $116 < m_{ee} < 130$ GeV, matching the first bin of the analysis measurement.

Since the variable $E_{\text{T}}^{\text{cone20}}$ is used to derive the scale factor for the background templates, the isolation cut was removed from the signal selection, since the nominal signal selection in the analysis applies a cut on this variable.

To further remove possible signal contamination from our template selection, the remaining events obtained after applying the same cuts on the different MC samples used in

this analysis are subtracted from the background template, as shown in Figure A.2.

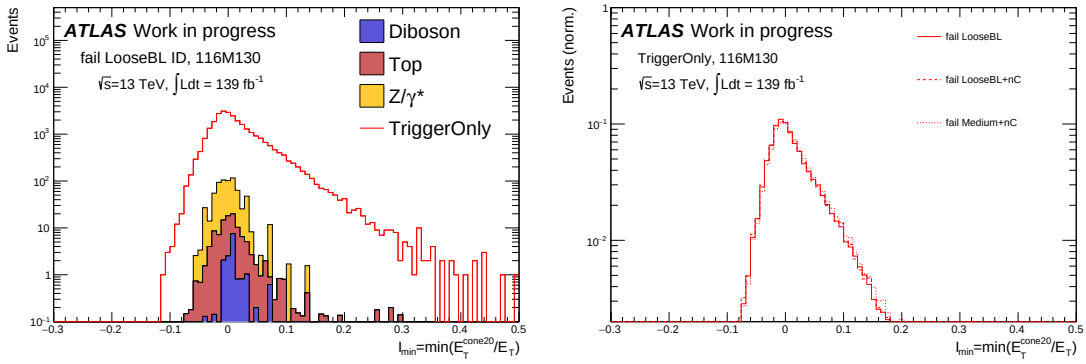


Figure A.2: I_{min} distribution for the template selection. (Left) Signal contamination remaining after template selection is applied. (Right) Signal-subtracted template distributions. Different selections shown, normalized to unit area.

The fit is performed in such way that the number of signal MC events, N_{MC}, in addition to the scaled background template, N_{bkg}, match the total number of observed signal events, N_{data}. This is performed in both the full I_{min} range and in a fit window chosen such that the data is background-dominated (also referred to as *tail*). Thus, the fit can be approached as a two-equation system, solved to extract the background and MC scale factors:

$$\begin{aligned} N_{\text{data}} &= S_{\text{bkg}} \cdot N_{\text{bkg}} + S_{\text{MC}} \cdot N_{\text{MC}} \\ N_{\text{data}}^{\text{tail}} &= S_{\text{bkg}} \cdot N_{\text{bkg}}^{\text{tail}} + S_{\text{MC}} \cdot N_{\text{MC}}^{\text{tail}} \\ \therefore S_{\text{bkg}} &= \frac{N_{\text{data}}^{\text{tail}} \cdot N_{\text{MC}} - N_{\text{data}} \cdot N_{\text{MC}}^{\text{tail}}}{N_{\text{bkg}}^{\text{tail}} \cdot N_{\text{MC}} - N_{\text{bkg}} \cdot N_{\text{MC}}^{\text{tail}}} \\ S_{MC} &= \frac{N_{\text{data}} - S_{\text{bkg}} \cdot N_{\text{bkg}}}{N_{\text{MC}}}. \end{aligned} \tag{A.2}$$

The tail window was chosen such that the fit included as many observed events in the data as possible, while keeping the signal peak out of what we considered the tail. This process was repeated for all mass bins in the analysis. Figure A.3 shows the results corresponding to the first ($116 < m_{\ell\ell} < 130$ GeV) and last ($1500 < m_{\ell\ell} < 5000$ GeV) invariant mass bins of the high-mass Drell-Yan cross-section measurement. The low end of the tail starts at $I_{\text{min}} = 0.1$ for the first bin and is progressively reduced to $I_{\text{min}} = 0.03$ for the last bin. To ensure the correct description of the I_{min} variable, additional studies were carried out to analyse the modelling of $E_{\text{T}}^{\text{cone20}}$ in the Monte-Carlo samples used. The variable was found to be correctly modelled, meaning that the peak observed in the I_{min} MC distribution follows the same as signal electrons observed in data.

The fit procedure is repeated for all template variations and mass bins, obtaining the distributions shown in Figure A.4. All distributions obtained with the different templates

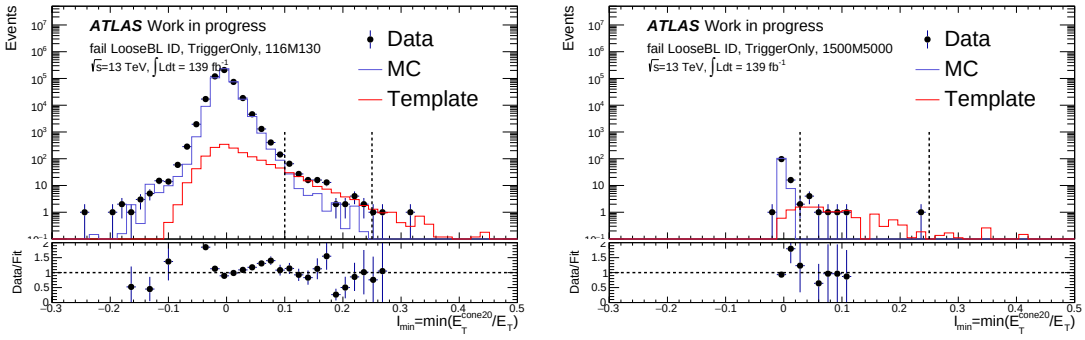


Figure A.3: Post-fit data to background (combination of MC background events and fake electrons derived using the template method) comparison in the (Left) $116 < m_{ee} < 130$ GeV (Right) $1500 < m_{ee} < 5000$ GeV range.

seem consistent with each other, predicting a number of events that are compatible with each other within 1σ variations. The results of the template method were compared with alternative estimations of the fake dielectron background used in the high-mass Drell-Yan analysis, both with a custom framework and the one provided by the ATLAS Isolation and Fake Forum. The comparison can be found in Figure A.5, where predictions are shown as the % of observed signal events found in each mass bin. While matrix method predicts an increasing contribution as the dilepton invariant mass increases, the template method predicts a rather flat contribution against $m_{\ell\ell}$. The matrix method predicts a bigger contribution from $m_{\mathrm{ee}} > 200$ GeV, although both methods are compatible within 2σ in most of the range.

Due to the contrasted validity of the matrix method at high dilepton invariant masses, which results have been cross-checked with other analysis teams studying the same final state, as well as previous findings of the reverse-ID methods failing to describe the fake electron background at high invariant masses (see Ref. [188]), the decision was to keep the results of the matrix method as our nominal fake electron background estimation.

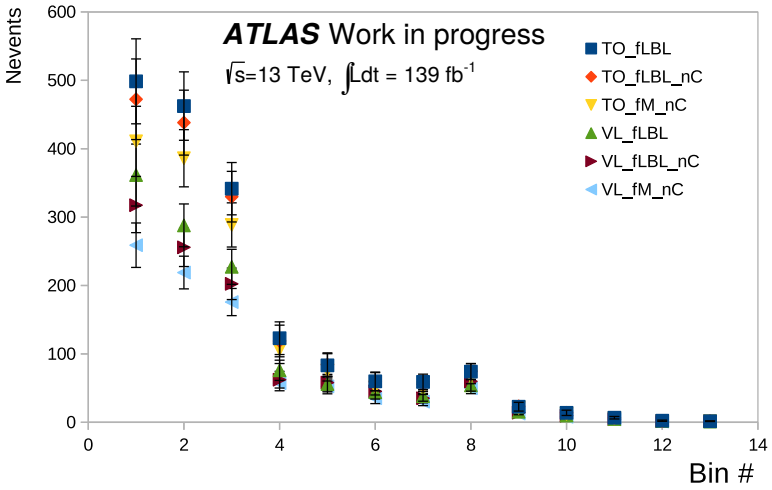


Figure A.4: Number of events predicted by each template for the different mass bins in the analysis. "TO" refers to trigger-only, meaning that the templates do not apply any cuts on ID, while templates labelled "VL" apply VeryLoose ID cuts. "fLBL" and "fM" refer to the reverse isolation cuts, meaning that the templates require electrons to fail the Loose and Medium ID criteria respectively. Templates labelled "nC" apply same-sign charge cuts on the dielectron pair.

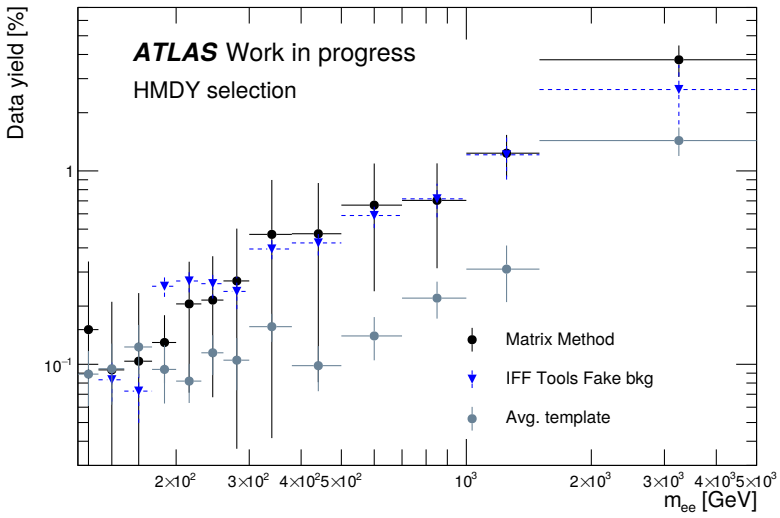


Figure A.5: Fake dielectron background prediction given by each different estimation method, shown as the % of observed signal events found in each mass bin. All 6 templates available in Figure A.4 have been averaged for plotting clarity purposes.

Appendix B

Results on the Drell-Yan EFT fits

This Appendix includes a detailed display of the fit results obtained as described in Section 10.2.

	Electron		Muon		Combined	
Operator	Linear	Linear+quad.	Linear	Linear+quad.	Linear	Linear+quad.
c_{lu}	[-0.024, 0.0124]	[-0.014, 0.031]	[-0.024, 0.025]	[-0.015, 0.032]	[-0.019, 0.019]	[-0.012, 0.027]
c_{eu}	[-0.012, 0.012]	[-0.010, 0.030]	[-0.012, 0.013]	[-0.010, 0.019]	[-0.010, 0.010]	[-0.008, 0.014]
c_{ed}	[-0.048, 0.047]	[-0.041, 0.023]	[-0.046, 0.046]	[-0.042, 0.024]	[-0.037, 0.036]	[-0.036, 0.020]
c_{ld}	[-0.096, 0.095]	[-0.036, 0.027]	[-0.094, 0.093]	[-0.038, 0.029]	[-0.075, 0.073]	[-0.032, 0.023]
c_{qe}	[-0.031, 0.032]	[-0.014, 0.022]	[-0.032, 0.032]	[-0.014, 0.024]	[-0.025, 0.025]	[-0.012, 0.020]
$c_{lq}^{(1)}$	[-0.014, 0.014]	[-0.010, 0.030]	[-0.015, 0.015]	[-0.011, 0.030]	[-0.011, 0.011]	[-0.008, 0.027]
$c_{lq}^{(3)}$	[-0.006, 0.006]	[-0.007, 0.005]	[-0.006, 0.006]	[-0.007, 0.006]	[-0.005, 0.005]	[-0.005, 0.004]
$c_{ll}^{(1)}$	[-0.630, 0.618]	[-0.643, 0.607]	[-0.634, 0.623]	[-0.646, 0.611]	[-0.577, 0.568]	[-0.587, 0.559]
c_{HDD}	[-0.737, 0.746]	[-0.727, 0.753]	[-0.801, 0.807]	[-0.793, 0.812]	[-0.525, 0.534]	[-0.522, 0.537]
c_{HWB}	[-0.287, 0.290]	[-0.281, 0.297]	[-0.312, 0.313]	[-0.305, 0.321]	[-0.205, 0.208]	[-0.202, 0.211]
$c_{Hl}^{(1)}$	[-0.642, 0.628]	[-0.663, 0.610]	[-0.602, 0.587]	[-0.620, 0.571]	[-0.568, 0.552]	[-0.585, 0.537]
$c_{Hl}^{(3)}$	[-0.496, 0.505]	[-0.484, 0.518]	[-0.530, 0.538]	[-0.517, 0.552]	[-0.418, 0.429]	[-0.410, 0.439]
$c_{Hq}^{(1)}$	[-6.164, 6.109]	[-5.019, 3.119]	[-7.602, 7.500]	[-4.796, 3.143]	[-5.899, 5.791]	[-4.600, 2.809]
$c_{Hq}^{(3)}$	[-0.669, 0.653]	[-0.686, 0.638]	[-0.621, 0.607]	[-0.637, 0.594]	[-0.575, 0.557]	[-0.588, 0.545]
c_{He}	[-0.944, 0.962]	[-0.886, 1.042]	[-0.911, 0.931]	[-0.856, 1.007]	[-0.749, 0.771]	[-0.713, 0.818]
c_{Hu}	[-3.856, 3.783]	[-11.337, 2.895]	[-3.540, 3.471]	[-10.802, 2.716]	[-3.134, 3.052]	[-10.545, 2.453]
c_{Hd}	[-7.249, 7.409]	[-3.924, 8.371]	[-6.794, 6.954]	[-3.751, 8.148]	[-5.870, 6.066]	[-3.425, 7.965]

Table B.1: 95% CL expected limits on the Wilson coefficients. Limits correspond to fits performed on the 1-dimensional ($m_{\ell\ell}$) measurement. Limits extracted using Asimov fit, interpreting SM prediction as pseudo-data, while introducing experimental and theoretical systematic uncertainties observed in the measurement (and prediction) on it. The effects of a single operator are considered at a time, setting the rest to zero.

	Electron		Muon		Combined	
Operator	Linear	Linear+quad.	Linear	Linear+quad.	Linear	Linear+quad.
c_{lu}	[-0.025, 0.025]	[-0.015, 0.031]	[-0.002, 0.040]	[-0.005, 0.020]	[-0.007, 0.029]	[-0.006, 0.022]
c_{eu}	[-0.013, 0.013]	[-0.010, 0.024]	[-0.001, 0.020]	[-0.002, 0.022]	[-0.004, 0.015]	[-0.004, 0.021]
c_{ed}	[-0.047, 0.052]	[-0.039, 0.024]	[-0.073, 0.008]	[-0.026, 0.011]	[-0.054, 0.017]	[-0.027, 0.012]
c_{ld}	[-0.093, 0.106]	[-0.036, 0.027]	[-0.150, 0.014]	[-0.022, 0.014]	[-0.110, 0.034]	[-0.023, 0.015]
c_{qe}	[-0.034, 0.032]	[-0.014, 0.022]	[-0.003, 0.052]	[-0.006, 0.014]	[-0.010, 0.038]	[-0.007, 0.015]
$c_{lq}^{(1)}$	[-0.015, 0.015]	[-0.010, 0.029]	[-0.000, 0.025]	[-0.002, 0.020]	[-0.004, 0.018]	[-0.004, 0.022]
$c_{lq}^{(3)}$	[-0.006, 0.006]	[-0.007, 0.006]	[-0.010, 0.001]	[-0.011, 0.001]	[-0.007, 0.002]	[-0.008, 0.002]
$c_{ll}^{(1)}$	[-0.413, 0.838]	[-0.418, 0.818]	[-0.432, 0.846]	[-0.440, 0.830]	[-0.331, 0.812]	[-0.335, 0.793]
c_{HDD}	[-0.458, 1.039]	[-0.460, 1.106]	[-0.727, 0.952]	[-0.751, 1.021]	[-0.312, 0.772]	[-0.313, 0.813]
c_{HWB}	[-0.152, 0.432]	[-0.151, 0.451]	[-0.251, 0.395]	[-0.249, 0.417]	[-0.098, 0.324]	[-0.097, 0.336]
$c_{Hl}^{(1)}$	[-0.295, 0.984]	[-0.299, 0.948]	[-0.400, 0.807]	[-0.409, 0.780]	[-0.225, 0.902]	[-0.227, 0.872]
$c_{Hl}^{(3)}$	[-0.470, 0.530]	[-0.460, 0.545]	[-0.603, 0.500]	[-0.591, 0.518]	[-0.355, 0.495]	[-0.349, 0.511]
$c_{Hq}^{(1)}$	[-2.524, 9.965]	[-5.978, 4.413]	[-17.878, 8.313]	[-5.891, 3.150]	[-1.955, 8.961]	[-5.902, 4.062]
$c_{Hq}^{(3)}$	[-0.228, 1.102]	[-0.230, 1.060]	[-0.453, 0.798]	[-0.463, 0.782]	[-0.168, 0.967]	[-0.169, 0.937]
c_{He}	[-1.703, 0.219]	[-1.519, 0.226]	[-1.203, 0.677]	[-1.124, 0.724]	[-1.404, 0.119]	[-1.288, 0.119]
c_{Hu}	[-0.613, 7.075]	[-0.765, 4.591]	[-3.055, 4.302]	[-12.726, -2.760] ⋃ [-1.679,2.235]	[-0.375, 5.842]	[-0.437, 4.069]
c_{Hd}	[-12.883, 1.862]	[-5.633, 9.582]	[-9.429, 4.683]	[-4.515, 9.411]	[-11.100, 0.806]	[-5.207, 3.172]

Table B.2: 95% CL observed limits on the Wilson coefficients corresponding to the operators. Limits correspond to fits performed on the 1D (m_{ee}) measurement. The effects of a single operator are considered at a time, setting the rest to zero.

	Electron		Muon		Combined	
Operator	Linear	Linear+quad.	Linear	Linear+quad.	Linear	Linear+quad.
c_{lu}	[-0.019, 0.019]	[-0.013, 0.028]	[-0.019, 0.019]	[-0.013, 0.027]	[-0.014, 0.014]	[-0.011, 0.023]
c_{eu}	[-0.011, 0.011]	[-0.009, 0.019]	[-0.011, 0.011]	[-0.010, 0.015]	[-0.008, 0.008]	[-0.007, 0.010]
c_{ed}	[-0.042, 0.042]	[-0.040, 0.022]	[-0.039, 0.039]	[-0.042, 0.023]	[-0.030, 0.030]	[-0.037, 0.018]
c_{ld}	[-0.071, 0.071]	[-0.035, 0.026]	[-0.067, 0.067]	[-0.038, 0.027]	[-0.052, 0.052]	[-0.032, 0.022]
c_{qe}	[-0.025, 0.025]	[-0.013, 0.022]	[-0.024, 0.024]	[-0.013, 0.023]	[-0.018, 0.018]	[-0.011, 0.019]
$c_{lq}^{(1)}$	[-0.013, 0.013]	[-0.009, 0.029]	[-0.014, 0.014]	[-0.010, 0.029]	[-0.010, 0.010]	[-0.008, 0.026]
$c_{lq}^{(3)}$	[-0.005, 0.005]	[-0.006, 0.005]	[-0.005, 0.005]	[-0.006, 0.005]	[-0.004, 0.004]	[-0.004, 0.004]
$c_{ll}^{(1)}$	[-0.577, 0.570]	[-0.588, 0.560]	[-0.600, 0.593]	[-0.611, 0.583]	[-0.502, 0.499]	[-0.510, 0.492]
c_{HDD}	[-0.394, 0.394]	[-0.392, 0.396]	[-0.418, 0.418]	[-0.415, 0.420]	[-0.297, 0.298]	[-0.296, 0.299]
c_{HWB}	[-0.183, 0.183]	[-0.180, 0.186]	[-0.185, 0.185]	[-0.181, 0.188]	[-0.134, 0.134]	[-0.132, 0.135]
$c_{Hl}^{(1)}$	[-0.518, 0.507]	[-0.530, 0.497]	[-0.514, 0.502]	[-0.526, 0.492]	[-0.432, 0.423]	[-0.440, 0.416]
$c_{Hl}^{(3)}$	[-0.376, 0.379]	[-0.368, 0.387]	[-0.399, 0.401]	[-0.390, 0.410]	[-0.299, 0.301]	[-0.294, 0.307]
$c_{Hq}^{(1)}$	[-0.947, 0.945]	[-0.960, 0.929]	[-1.038, 1.037]	[-1.053, 1.016]	[-0.735, 0.734]	[-0.739, 0.728]
$c_{Hq}^{(3)}$	[-0.524, 0.514]	[-0.533, 0.506]	[-0.536, 0.523]	[-0.545, 0.515]	[-0.439, 0.429]	[-0.444, 0.424]
c_{He}	[-0.694, 0.709]	[-0.663, 0.744]	[-0.688, 0.705]	[-0.660, 0.737]	[-0.557, 0.570]	[-0.540, 0.589]
c_{Hu}	[-1.513, 1.500]	[-1.493, 1.505]	[-1.464, 1.449]	[-1.444, 1.456]	[-1.169, 1.159]	[-1.159, 1.162]
c_{Hd}	[-4.544, 4.605]	[-3.318, 4.397]	[-4.418, 4.490]	[-3.318, 4.397]	[-3.557, 3.617]	[-2.868, 3.687]

Table B.3: 95% CL expected limits on the Wilson coefficients. Limits correspond to fits performed on the 2D(+1) ($m_{ll} \otimes \cos\theta^*_{CS} + m_{ee} \in [1500, 5000]$ GeV) measurement. Limits extracted using Asimov fit, interpreting SM prediction as pseudo-data, while introducing experimental and theoretical systematic uncertainties observed in the measurement (and prediction) on it. The effects of a single operator are considered at a time, setting the rest to zero.

	Electron		Muon		Combined	
Operator	Linear	Linear+quad.	Linear	Linear+quad.	Linear	Linear+quad.
c_{lu}	[-0.025, 0.016]	[-0.016, 0.031]	[0.006, 0.039]	[-0.001, 0.021]	[-0.004, 0.024]	[-0.004, 0.022]
c_{eu}	[-0.016, 0.009]	[-0.012, 0.012]	[0.005, 0.024]	[0.004, 0.029]	[-0.001, 0.015]	[-0.001, 0.026]
c_{ed}	[-0.031, 0.062]	[-0.044, 0.028]	[-0.084, -0.016]	[-0.028, 0.005]	[-0.054, 0.006]	[-0.030, 0.008]
c_{ld}	[-0.058, 0.095]	[-0.041, 0.032]	[-0.137, -0.017]	[-0.022, 0.010]	[-0.085, 0.018]	[-0.025, 0.013]
c_{qe}	[-0.033, 0.021]	[-0.016, 0.025]	[0.006, 0.049]	[-0.004, 0.015]	[-0.006, 0.030]	[-0.006, 0.016]
$c_{lq}^{(1)}$	[-0.018, 0.012]	[-0.012, 0.033]	[0.007, 0.029]	[0.002, 0.023]	$[-1.2 \cdot 10^{-4}, 0.019]$	$[2.2 \cdot 10^{-4}, 0.026]$
$c_{lq}^{(3)}$	[-0.004, 0.008]	[-0.005, 0.007]	[-0.011, -0.002]	[-0.013, -0.002]	$[-0.007, -4.8 \cdot 10^{-4}]$	$[-0.008, 4.8 \cdot 10^{-4}]$
$c_{ll}^{(1)}$	[-0.243, 0.908]	[-0.245, 0.884]	[-0.874, 0.310]	[-0.900, 0.308]	[-0.436, 0.576]	[-0.442, 0.566]
c_{HDD}	[-0.593, 0.199]	[-0.595, 0.200]	[-0.733, 0.120]	[-0.714, 0.122]	[-0.563, 0.039]	[-0.558, 0.039]
c_{HWB}	[-0.266, 0.103]	[-0.259, 0.104]	[-0.353, 0.027]	[-0.344, 0.027]	[-0.283, -0.012]	[-0.276, -0.012]
$c_{Hl}^{(1)}$	[-0.410, 0.608]	[-0.416, 0.591]	[-1.139, -0.106]	[-1.214, -0.112]	[-0.809, 0.046]	[-0.840, 0.045]
$c_{Hl}^{(3)}$	[-0.644, 0.122]	[-0.625, 0.122]	[-0.394, 0.413]	[-0.388, 0.424]	[-0.492, 0.123]	[-0.481, 0.123]
$c_{Hq}^{(1)}$	[-0.248, 1.644]	[-0.252, 1.577]	[-0.158, 1.913]	[-0.191, 1.722]	[-0.006, 1.462]	[-0.029, 1.368]
$c_{Hq}^{(3)}$	[-0.461, 0.572]	[-0.467, 0.563]	[-0.913, 0.168]	[-0.940, 0.167]	[-0.683, 0.185]	[-0.693, 0.184]
c_{He}	[-0.736, 0.675]	[-0.706, 0.713]	[-0.062, 1.415]	[-0.017, 1.751]	[-0.009, 1.158]	[0.011, 1.310]
c_{Hu}	[-1.277, 1.760]	[-1.291, 1.815]	[-3.562, -0.624]	[-3.245, -0.546]	[-2.405, -0.062]	[-2.286, -0.039]
c_{Hd}	[-4.126, 5.050]	[-3.249, 5.103]	[0.346, 9.421]	[0.752, 6.733]	[0.064, 7.309]	[0.123, 5.756]

Table B.4: 95% CL observed limits on the Wilson coefficients. Limits correspond to fits performed on the 2D(+1) ($m_{ll} \otimes \cos\theta^*_{CS} + m_{ee} \in [1500, 5000]$ GeV) measurement. The effects of a single operator are considered at a time, setting the rest to zero.

	1D		2D(+1)		Diff. (2D-1D) [%]	
Operator	Linear	Linear+quad.	Linear	Linear+quad.	Linear	Linear+quad.
c_{lu}	[-0.007, 0.029]	[-0.006, 0.022]	[-0.004, 0.024]	[-0.004, 0.022]	-22.2	-7.1
c_{eu}	[-0.004, 0.015]	[-0.004, 0.021]	[-0.001, 0.015]	[-0.001, 0.026]	-15.8	+7.1
c_{ed}	[-0.054, 0.017]	[-0.027, 0.012]	[-0.054, 0.006]	[-0.030, 0.008]	-15.5	-2.6
c_{ld}	[-0.110, 0.034]	[-0.023, 0.015]	[-0.085, 0.018]	[-0.025, 0.013]	-28.5	-2.2
c_{qe}	[-0.010, 0.038]	[-0.007, 0.015]	[-0.006, 0.030]	[-0.006, 0.016]	-25.0	-1.0
$c_{lq}^{(1)}$	[-0.004, 0.018]	[-0.004, 0.022]	$[-1 \cdot 10^{-4}, 0.019]$	$[2 \cdot 10^{-4}, 0.026]$	-13.1	-3.9
$c_{lq}^{(3)}$	[-0.007, 0.002]	[-0.008, 0.002]	$[-0.007, -5 \cdot 10^{-4}]$	$[-0.008, 5 \cdot 10^{-4}]$	-27.6	-15.2
$c_{ll}^{(1)}$	[-0.331, 0.812]	[-0.335, 0.793]	[-0.436, 0.576]	[-0.442, 0.566]	-11.5	-10.6
c_{HD}	[-0.312, 0.772]	[-0.313, 0.813]	[-0.563, 0.039]	[-0.558, 0.039]	-44.5	-47.0
c_{HWB}	[-0.098, 0.324]	[-0.097, 0.336]	[-0.283, -0.012]	[-0.276, -0.012]	-35.8	-39.0
$c_{Hl}^{(1)}$	[-0.225, 0.902]	[-0.227, 0.872]	[-0.809, 0.046]	[-0.840, 0.045]	-24.1	-19.5
$c_{Hl}^{(3)}$	[-0.355, 0.495]	[-0.349, 0.511]	[-0.492, 0.123]	[-0.481, 0.123]	-27.6	-29.8
$c_{Hq}^{(1)}$	[-1.955, 8.961]	[-5.902, 4.062]	[-0.006, 1.462]	[-0.029, 1.368]	-86.6	-86.0
$c_{Hq}^{(3)}$	[-0.168, 0.967]	[-0.169, 0.937]	[-0.683, 0.185]	[-0.693, 0.184]	-23.5	-20.7
c_{He}	[-1.404, 0.119]	[-1.288, 0.119]	[-0.009, 1.158]	[0.011, 1.310]	-23.4	-7.7
c_{Hu}	[-0.375, 5.842]	[-0.437, 4.069]	[-2.405, -0.062]	[-2.286, -0.039]	-62.3	-50.1
c_{Hd}	[-12.883, 1.862]	[-5.633, 9.582]	[0.064, 7.309]	[0.123, 5.756]	-50.9	-63.0

Table B.5: Observed limits difference between 1D and 2D(+1) fits in the combined channel, expressed as the difference in range from the highest to lowest operator limit. The effects of a single operator are considered at a time, setting the rest to zero.

Appendix C

Simulation of the Z-counting lepton efficiencies

This Appendix includes a comparison of the single-lepton trigger and reconstruction efficiencies obtained in Monte Carlo simulation and data samples, as described in Sections 13.3 and 13.4.3.

The simulated single-lepton reconstruction efficiencies agree with those found in data within 2% in most data-taking periods, with the difference between the simulated and observed values exhibiting very little pileup dependence. Trigger efficiencies obtained in MC agree with those measured in data within a 2.5% level, with the exception of 2016. In the muon channel, the agreement worsens to a 7.5% level in most years, while 2016 shows a 5% agreement. The single-lepton efficiencies are combined into an event-level efficiency as shown in Equation 13.7. The agreement between the observed and simulated event-level efficiency is within 5% in most data-taking periods, exhibiting very little pileup dependence. The effect for this mis-modelling of the lepton efficiencies is corrected for by applying the factor $F^{\mathrm{MC}}_{Z\to\ell^+\ell^-}(\langle\mu\rangle)$ in Equation 13.2, as explained in Section 13.4.

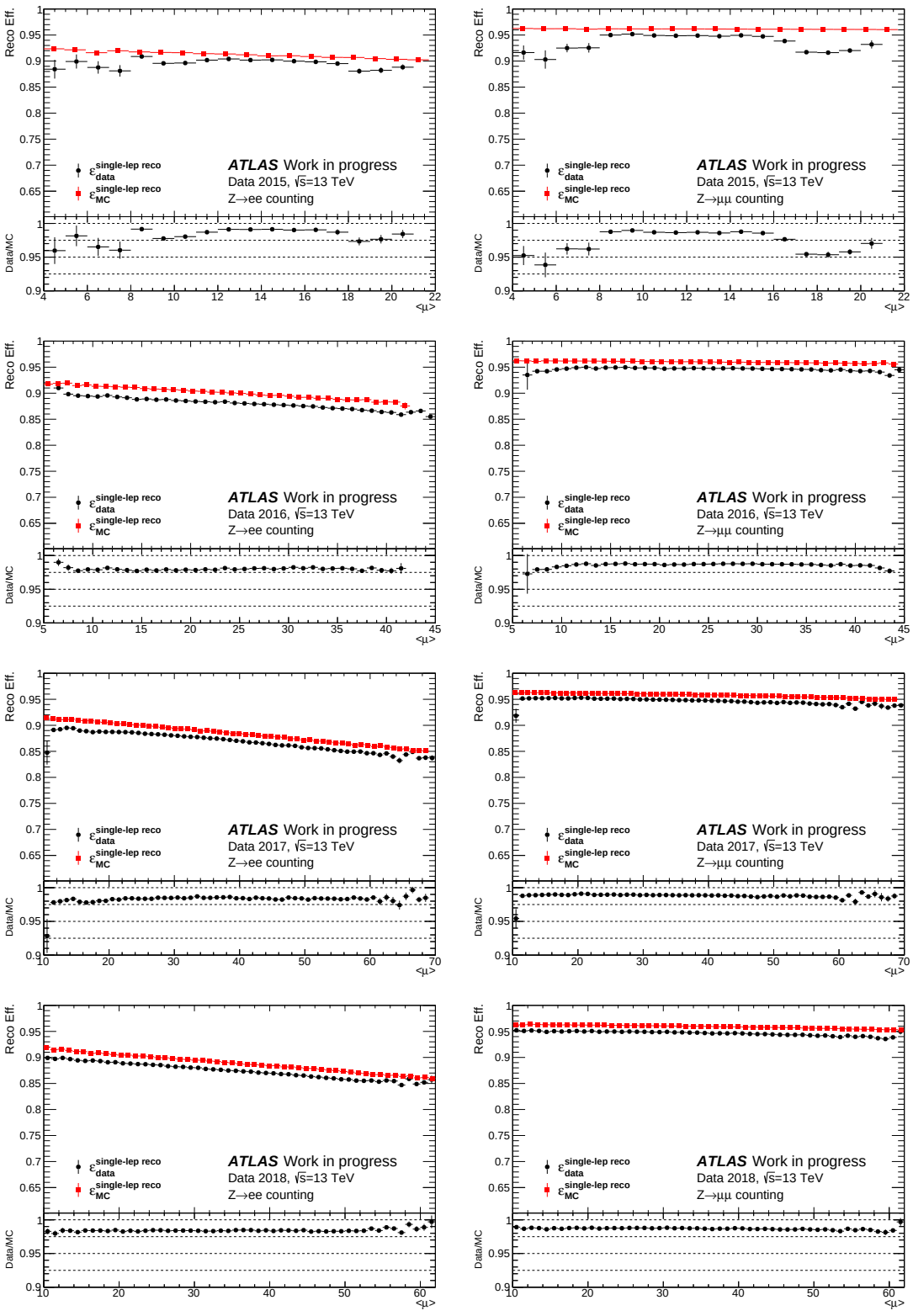


Figure C.1: Single-electron (left) and muon (right) reconstruction efficiencies in data and Monte-Carlo against average pileup.

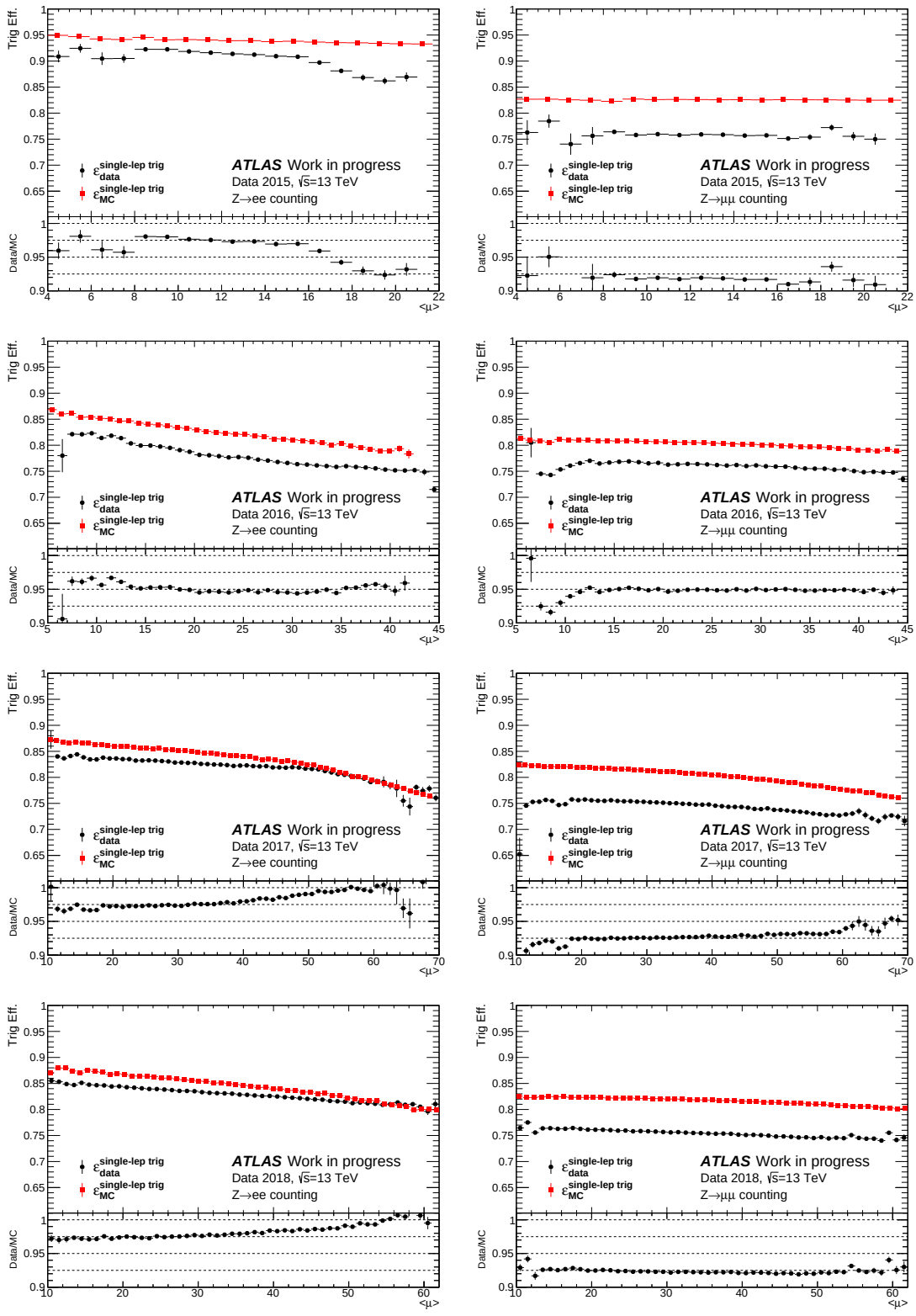


Figure C.2: Single-electron (left) and muon (right) trigger efficiencies in data and Monte-Carlo against average pileup.

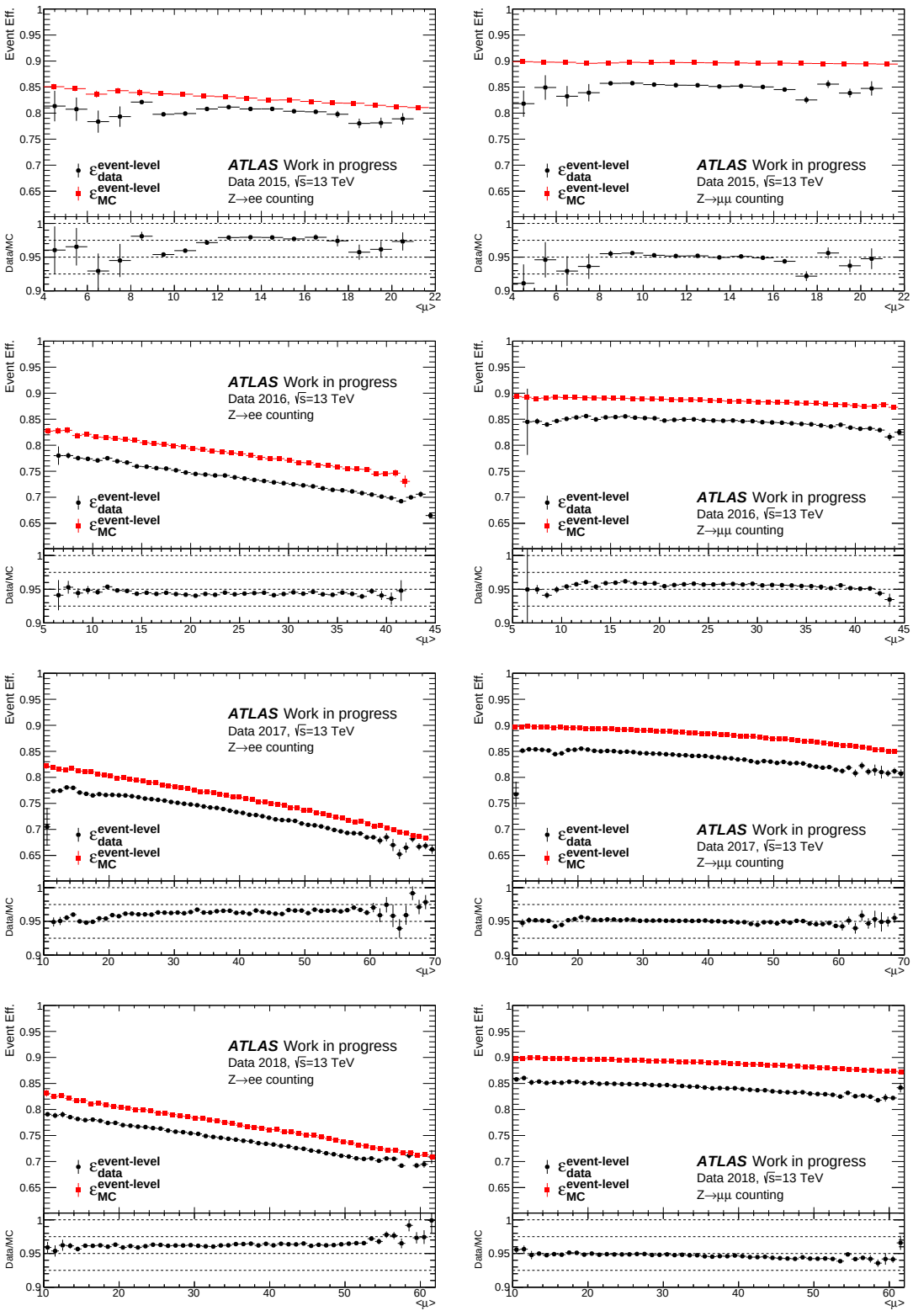


Figure C.3: Event-level electron (left) and muon (right) efficiencies in data and Monte-Carlo against average pileup.

Appendix D

Muon channel limitations in AnalysisBase Z-counting

The results shown in Section 13.5 were obtained adapting the Z-counting methodology to a standard ATLAS analysis framework, AnalysisBase, which runs on data derivations that include some event pre-selection. In the case of the electron channel this pre-selection does not limit the results obtained, but for muon decays it means that reconstruction efficiencies will be inherently biased towards a higher efficiency. The reconstruction efficiency for muons is given by the expression shown in Equation 13.5 and, as shown in Table 13.2, probe candidates consist of inner detector tracks passing some very basic pre-selection cuts. In primary ATLAS analysis object data (AOD), these pre-selection cuts keep $> 99\%$ of the ID tracks, allowing for a good measurement of the reconstruction efficiency. However, in derivated data used by AnalysisBase most of the pre-selected data available includes some track removal, since the chances of ATLAS analyses requiring access to information on all tracks is very unlikely, allowing for the slimming of data files to reduce storage space usage. The tracks that are slimmed away in the derivation process tend to correspond to non-matched objects, since those muon-like candidates pass the pre-selection criteria. This means that the denominator in Equation 13.5 is smaller ($N_{\text{fail}}^{\text{OS}}$ is reduced) when calculated on pre-selected data, resulting in a bigger reconstruction efficiency. Table D.1 shows some example quantities for all elements used in the determination of the reconstruction efficiency for different data derivations available for a given luminosity block in the 2018 data period. The table shows how the different pre-selection requirements applied in different derivations tested drastically reduce the number of tracks fulfilling the $N_{\text{fail}}^{\text{OS}}$ criteria, resulting in a difference in reconstruction efficiency of over 10%.

Moreover, the effects of the track slimming applied in the derivation process on the muon channel Z-counting is dependent on the pileup of the process. The additional proton-proton interactions result in a higher number of low-quality tracks that, although originally

Quantity	Derivation		
	STDM4	MUON1	AOD
$N_{\text{pass}}^{\text{OS}}$	212	212	206
$N_{\text{pass}}^{\text{bkg}}$	0	0	0
$N_{\text{fail}}^{\text{OS}}$	21	42	57
$N_{\text{fail}}^{\text{bkg}}$	6	13	14
$\varepsilon_{\text{reco},1\mu}$ [%]	93.4	87.9	82.7

Table D.1: Summary of the quantities used in the computation of the single-muon reconstruction efficiency, as shown in Equation 13.5 for different derivations. STDM4 is the standard derivation used in single-lepton-based analyses in ATLAS. MUON1 is a muon-specific derivation that lowers the selection criteria on tracks included in the data files. AOD refers to the numbers corresponding to the original results used in the Z-counting luminosity computation [179], obtained using the WZFinder tool in the DataQuality analysis framework.

accounted for in the standard Z-counting methodology using primary ATLAS data, are removed from the stored data in usual ATLAS derivations. This results in a time-dependent effect observed in the muon channel Z-counting using AnalysisBase, since 2017 and 2018 data was taken with a higher pileup than the previous Run 2 years, resulting in a change in efficiency behaviour that biases the luminosity obtained with the Z-counting methodology, as shown in Figure D.1. It can be inferred from the figure that late-Run 2 data analysis results into a Z-counting luminosity dip, stemming from a bigger effect of ID track slimming in the derivation process, increasing the reconstruction efficiency measurement and therefore decreasing the luminosity estimation (see Equation 13.2). The effect can also be inferred from Figure D.2, where the dip in the luminosity obtained in the muon channel results into an increase with respect to that measured in the electron channel in late-Run 2 runs, where higher pileup values where recorded and the effects of track slimming are more severe in the luminosity determination in the muon channel.

Considering the limitations the AnalysisBase framework presents in the muon channel, the final results of the search for Lorentz- invariance violation will be based on the original Z-counting methodology using primary ATLAS data, updated to the final luminosity ATLAS measurement [109].

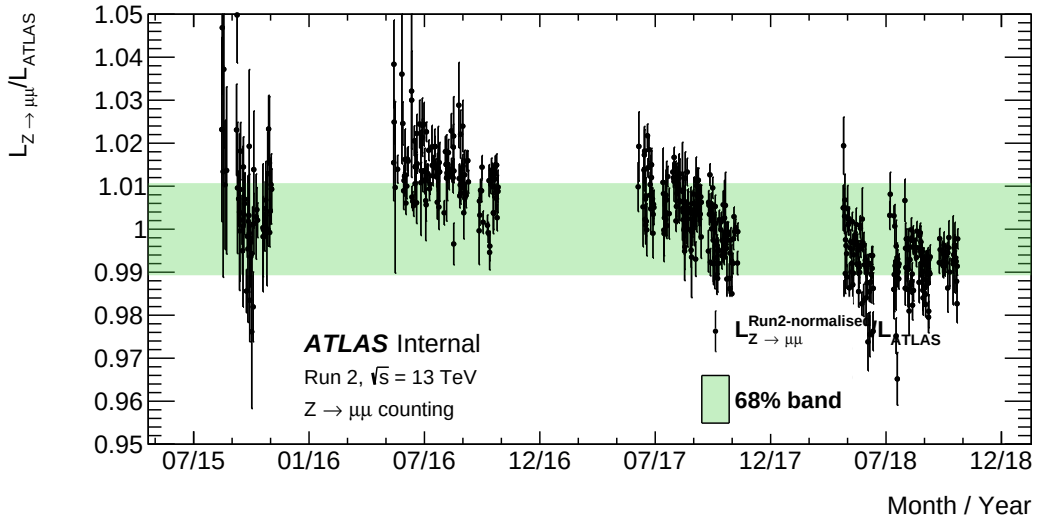


Figure D.1: Biased ratio of the integrated Z-counting and baseline ATLAS luminosities per LHC run taken from pp collisions at $\sqrt{s} = 13$ TeV for the $Z \to \mu^+\mu^-$ channel for the full Run-2 data taking period. The Z-counting luminosity is normalised to the integrated baseline ATLAS luminosity over the Run-2 data-taking period [152]. The x-axis represents the date when the run started. Only ATLAS runs with a minimum length of 40 minutes are included. The error bars show statistical uncertainties only and the green bands contain 68% of all points centred around the mean.

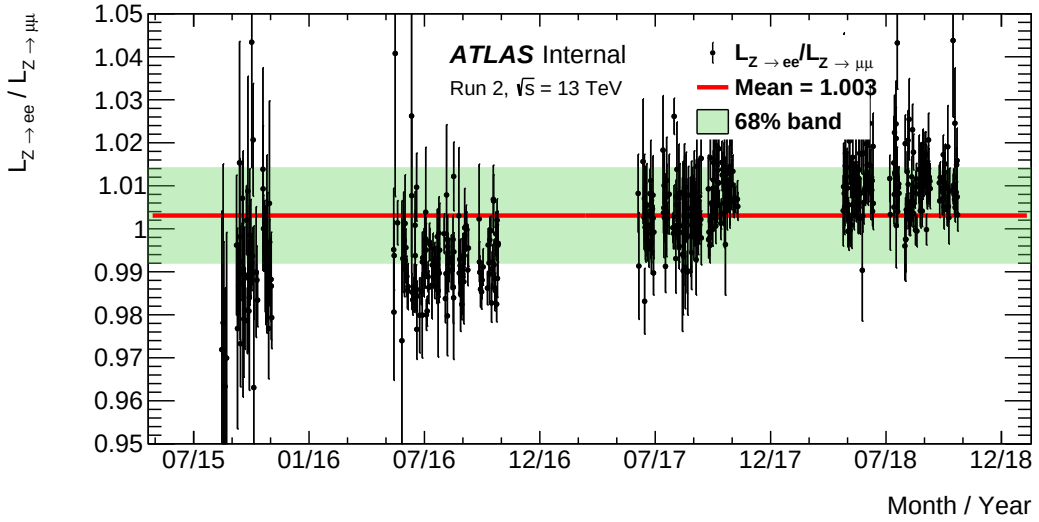


Figure D.2: Biased ratio of the integrated Z-counting $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ luminosities per LHC run taken from pp collisions at $\sqrt{s} = 13$ TeV for the full Run-2 data taking period. The red line indicated the mean of the distribution obtained from a fit to a constant. The error bars show statistical uncertainties only and the green bands contain 68% of all points centered around the mean.

Appendix E

Random phase assignment Monte-Carlo LIV studies

As a first step towards the simulation of time (phase) dependence one can simply assign a random uniformly distributed phase ranging (0,2π). Although this does not accurately represent the time-dependence of data-taking, it establishes a hypothesis to test the analysis framework on.

The pileup to phase mapping obtained when using the assignment of uniformly distributed sidereal-phases to each MC event can be found in Figure E.1. No dependence against phase can be inferred from the distribution, following the expectations when assigning a uniformly distributed phase. The dependence against pileup (y-axis) follows the pileup distribution recorded in 2018 data. The modelling of this variable (and its shape difference across different data-taking periods) can be exploited to introduce an accurate simulation of the phase dependence in MC, as described in Section 15.2. Figure E.2 shows the distribution of MC weights across different phase bins at truth- and reconstructed-level. No dependence against phase is observed for the correction factor shown in the bottom panel, meaning that the correction of reconstructed-level distributions (as described in Equation 15.2), introduces no dependence on the sidereal phase in the distribution.

Figure E.3 shows the double-ratio obtained when using the flat phase assignment on MC simulation. Both the truth- and unfolded reconstructed-level distributions are displayed on the Figure. Due to the nature of the bin-by-bin unfolding correction, the central points of this distribution exactly match the truth-level ones, the key difference between both distributions being the increased error bars observed in the unfolded distribution, due to the lower number of events passing the reconstruction cuts. The distribution obtained when unfolding following the Z-counting methodology displays a bigger spread than the other distributions, but all three distributions are compatible with the fit of a straight line,

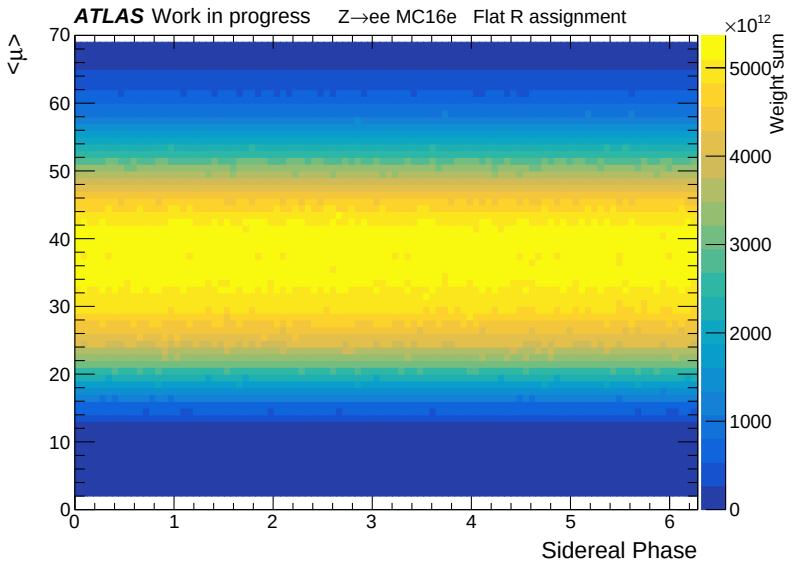


Figure E.1: Pileup vs sidereal phase profile obtained using a flat phase assignment to MC events. z-axis shows the sum of weights registered in a given pileup-phase combination. Results correspond to MC samples describing the 2018 data profile.

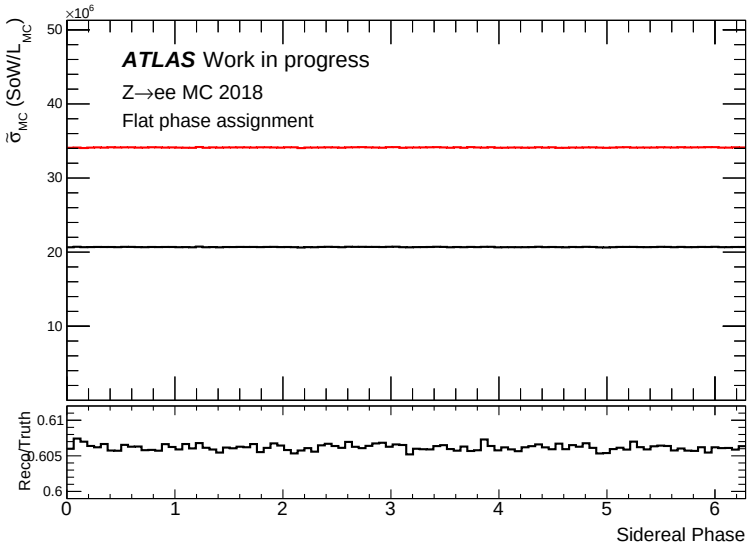


Figure E.2: Sum of weights distribution per sidereal bin obtained using a flat phase assignment to MC events, corrected for MC "luminosity". Distributions shown at truth (red) and reconstructed (black) level. Results correspond to MC samples describing the 2018 data profile.

corresponding to the assumption of lack of LIV effects. The results shown correspond to the processing of the MC samples under one single 1 random seed. The final bounds of the coefficients are retrieved repeating the process for 100 different random seeds, minimising the impact of the initial randomisation.

With this simulated observable, bounds on the coefficients can be obtained by perform-

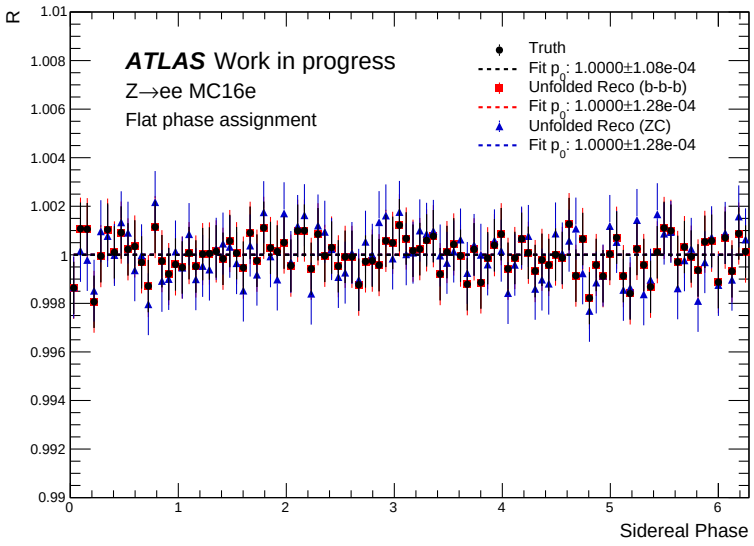


Figure E.3: Double-ratio obtained for the assignment of uniformly distributed sidereal phases. Figure shows truth-level (black circle) and unfolded distributions. Unfolding performed in a bin-by-bin basis (red square) and following the Z-counting methodology (blue triangle). The fits (dashed lines) correspond to a straight line fit on each of the distributions.

ing a χ^2 fit (details of which can be found in Ref. [182]) on the signal distributions shown in Equation 14.7. Figure E.4 shows the results corresponding to the signal fitting on the truth and reconstructed level distributions using a flat phase assignment. All coefficient limits seem compatible with a null hypothesis, obtaining sensitivities of the order $\mathcal{O}(10^{-4})$. The results presented in the analysis were expanded with the addition of data-driven phase simulation, as explained in Section 15.2.

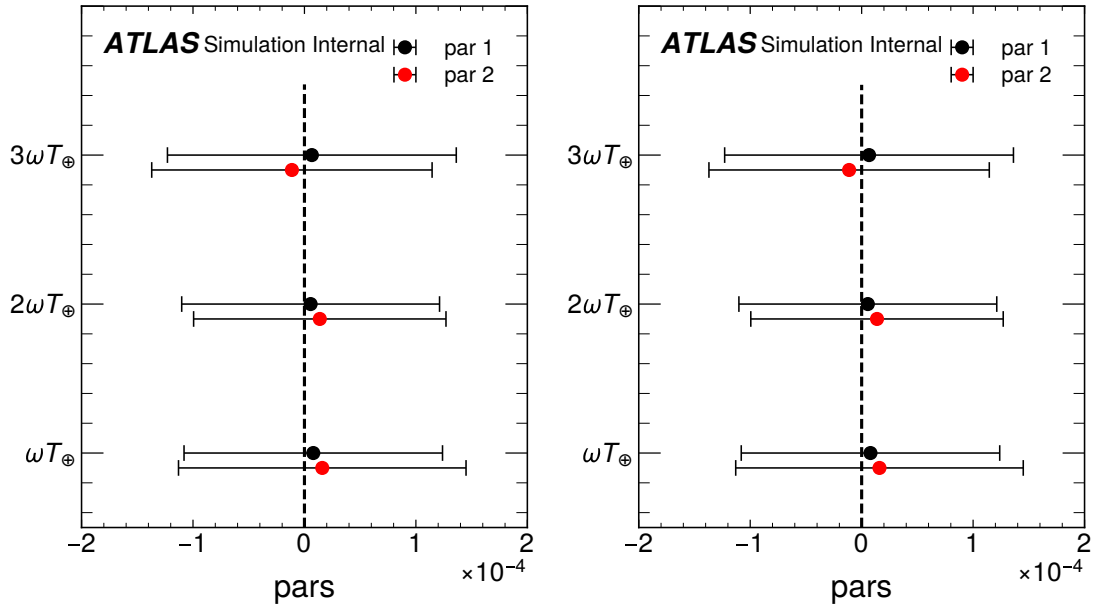


Figure E.4: Coefficient bounds extracted using the double ratio distributions at truth-level (left) and bin-by-bin unfolded (right) with a uniformly distributed sidereal phase assignment. Bounds obtained using the distributions produced with 100 different random seeds. Limits on the parameters ("pars") shown for the cosine ("par1", black) and sine ("par2", red) terms of the signal parametrisation. Plots provided by Dr. Yiming Abulaiti.

Bibliography

[1] I. J. R. Aitchison and A. J. G. Hey, *Gauge theories in particle physics: A practical introduction. Vol. 1: From relativistic quantum mechanics to QED*. CRC Press, 2012.

[2] C. Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions: Second Edition*. Princeton University Press, USA, 9, 2013.

[3] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[4] M. Thomson, *Modern particle physics*. Cambridge University Press, New York, 2013.

[5] Particle Data Group Collaboration, *Review of Particle Physics*, *PTEP* **2020** (2020) 083C01.

[6] M. Gell-Mann, *A schematic model of baryons and mesons*, *Physics Letters* **8** (1964) 214.

[7] G. Zweig, *An SU_3 model for strong interaction symmetry and its breaking; Version 1*, tech. rep., CERN, Geneva, Jan, 1964. http://cds.cern.ch/record/352337.

[8] O. W. Greenberg, *Spin and Unitary-Spin Independence in a Paraquark Model of Baryons and Mesons*, *Phys. Rev. Lett.* **13** (1964) 598.

[9] H. Yukawa, *On the Interaction of Elementary Particles I*, *Proc. Phys. Math. Soc. Jap.* **17** (1935) 48.

[10] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (1964) 321.

[11] P. Higgs, *Broken symmetries, massless particles and gauge fields*, *Physics Letters* **12** (1964) 132.

[12] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508.

[13] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, *Phys. Rev.* **145** (1966) 1156.

[14] F. Englert, R. Brout and M. F. Thiry, *Vector mesons in presence of broken symmetry*, *Il Nuovo Cimento A* **43** (1966) 244.

[15] S. Weinberg, *Physical Processes in a Convergent Theory of the Weak and Electromagnetic Interactions*, *Phys. Rev. Lett.* **27** (1971) 1688.

[16] M. Gell-Mann, *Symmetries of Baryons and Mesons*, *Phys. Rev.* **125** (1962) 1067.

[17] CTEQ Collaboration, *Handbook of perturbative QCD; Version 1.1: September 1994*, .

[18] L. Canetti, M. Drewes and M. Shaposhnikov, *Matter and Antimatter in the Universe*, *New J. Phys.* **14** (2012) 095012 [1204.4186].

[19] A. D. Sakharov, *Violation ofCPin variance,Casymmetry, and baryon asymmetry of the universe*, *Soviet Physics Uspekhi* **34** (1991) 392.

[20] R. Davis, D. S. Harmer and K. C. Hoffman, *Search for Neutrinos from the Sun*, *Phys. Rev. Lett.* **20** (1968) 1205.

[21] H. Georgi and S. L. Glashow, *Unity of All Elementary-Particle Forces*, *Phys. Rev. Lett.* **32** (1974) 438.

[22] V. C. Rubin and J. Ford, W. Kent, *Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions*, *Astrophys. J.* **159** (1970) 379.

[23] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones et al., *A direct empirical proof of the existence of dark matter*, *Astrophys. J. Lett.* **648** (2006) L109 [astro-ph/0608407].

[24] WMAP Collaboration, *Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Data Processing, Sky Maps, and Basic Results*, *Astrophys. J. Suppl.* **180** (2009) 225 [0803.0732].

[25] Planck Collaboration, *Planck 2013 results. I. Overview of products and scientific results*, *Astron. Astrophys.* **571** (2014) A1 [1303.5062].

[26] P. Langacker, *The Physics of Heavy Z' Gauge Bosons*, *Rev. Mod. Phys.* **81** (2009) 1199 [0801.1345].

[27] ATLAS Collaboration, *Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s}$ =13 TeV with the ATLAS detector*, *Phys. Lett. B* **796** (2019) 68 [1903.06248].

[28] ATLAS Collaboration, *Search for a heavy charged boson in events with a charged lepton and missing transverse momentum from pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *Phys. Rev. D* **100** (2019) 052013 [1906.05609].

[29] CMS Collaboration, *Search for resonant and nonresonant new phenomena in high-mass dilepton final states at* $\sqrt{s} = 13$ *TeV*, *JHEP* **07** (2021) 208 [2103.02708].

[30] G. Senjanovic and R. N. Mohapatra, *Exact left-right symmetry and spontaneous violation of parity*, *Phys. Rev. D* **12** (1975) 1502.

[31] E. E. Boos, V. E. Bunichev, M. A. Perfilov, M. N. Smolyakov and I. P. Volobuev, *The specificity of searches for* W', Z' *and* γ' *coming from extra dimensions*, *JHEP* **06** (2014) 160 [1311.5968].

[32] A. Djouadi, G. Moreau and R. K. Singh, *Kaluza-Klein excitations of gauge bosons at the LHC*, *Nucl. Phys. B* **797** (2008) 1 [0706.4191].

[33] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, *Phys. Rev. Lett.* **83** (1999) 3370 [hep-ph/9905221].

[34] S. P. Martin, *A Supersymmetry primer*, *Adv. Ser. Direct. High Energy Phys.* **18** (1998) 1 [hep-ph/9709356].

[35] P. Fayet, *Supersymmetry and weak, electromagnetic and strong interactions*, *Physics Letters B* **64** (1976) 159.

[36] ATLAS Collaboration, *Search for supersymmetry in final states with missing transverse momentum and three or more b-jets in 139* fb^{-1} *of proton−proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, 2211.08028.

[37] ATLAS Collaboration, *Search for direct pair production of sleptons and charginos decaying to two leptons and neutralinos with mass splittings near the W-boson mass in* $\sqrt{s} = 13$ *TeV pp collisions with the ATLAS detector*, 2209.13935.

[38] CMS Collaboration, *Search for chargino-neutralino production in events with Higgs and W bosons using 137* fb^1 *of proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, *JHEP* **10** (2021) 045 [2107.12553].

[39] CMS Collaboration, *Search for supersymmetry in final states with two or three soft leptons and missing transverse momentum in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, *JHEP* **04** (2022) 091 [2111.06296].

[40] I. Brivio and M. Trott, *The Standard Model as an Effective Field Theory*, 1706.08945.

[41] B. Grzadkowski, M. Iskrzynski, M. Misiak and J. Rosiek, *Dimension-Six Terms in the Standard Model Lagrangian*, *JHEP* **10** (2010) 085 [1008.4884].

[42] I. Brivio, Y. Jiang and M. Trott, *The SMEFTsim package, theory and tools*, *JHEP* **12** (2017) 070 [1709.06492].

[43] V. A. Kostelecky and N. Russell, *Data Tables for Lorentz and CPT Violation*, *Rev. Mod. Phys.* **83** (2011) 11 [0801.0287].

[44] V. A. Kostelecký and S. Samuel, *Spontaneous breaking of Lorentz symmetry in string theory*, *Phys. Rev. D* **39** (1989) 683.

[45] D. Colladay and V. A. Kostelecky, *CPT violation and the standard model*, *Phys. Rev. D* **55** (1997) 6760 [hep-ph/9703464].

[46] D. Colladay and V. A. Kostelecky, *Lorentz violating extension of the standard model*, *Phys. Rev. D* **58** (1998) 116002 [hep-ph/9809521].

[47] E. Lunghi, N. Sherrill, A. Szczepaniak and A. Vieira, *Quark-sector Lorentz violation in Z-boson production*, *JHEP* **04** (2021) 228 [2011.02632].

[48] P. Langacker, *The Standard Model and Beyond; 2nd ed.*, High energy physics, cosmology and gravitation. CRC Press, Jul, 2017.

[49] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1 [hep-ph/0409313].

[50] H1, ZEUS Collaboration, *Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA*, *JHEP* **01** (2010) 109 [0911.0884].

[51] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin and R. S. Thorne, *Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs*, *Eur. Phys. J. C* **81** (2021) 341 [2012.04684].

[52] NNPDF Collaboration, *Parton distributions from high-precision collider data*, *Eur. Phys. J. C* **77** (2017) 663 [1706.00428].

[53] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, *Phys. Rev. D* **93** (2016) 033006 [1506.07443].

[54] S. D. Drell and T.-M. Yan, *Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies*, *Phys. Rev. Lett.* **25** (1970) 316.

[55] J. C. Collins and D. E. Soper, *Angular distribution of dileptons in high-energy hadron collisions*, *Phys. Rev. D* **16** (1977) 2219.

[56] ATLAS Collaboration, *Measurement of the angular coefficients in Z-boson events using electron and muon pairs from data taken at $\sqrt{s} = 8$ TeV with the ATLAS detector*, *JHEP* **08** (2016) 159 [1606.00689].

[57] ATLAS Collaboration, *Measurement of the Drell-Yan triple-differential cross section in pp collisions at $\sqrt{s} = 8$ TeV*, *JHEP* **12** (2017) 059 [1710.05167].

[58] C. Duhr and B. Mistlberger, *Lepton-pair production at hadron colliders at N^3LO in QCD*, *JHEP* **03** (2022) 116 [2111.10379].

[59] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, *Parton distributions incorporating QED contributions*, *Eur. Phys. J. C* **39** (2005) 155 [hep-ph/0411040].

[60] NNPDF Collaboration, *Parton distributions with QED corrections*, *Nucl. Phys. B* **877** (2013) 290 [1308.0598].

[61] A. V. Manohar, P. Nason, G. P. Salam and G. Zanderighi, *The Photon Content of the Proton*, *JHEP* **12** (2017) 046 [1708.01256].

[62] T. Sjostrand, *Monte Carlo Generators*, in *2006 European School of High-Energy Physics*, pp. 51–74, 11, 2006, hep-ph/0611247.

[63] T. Gleisberg et al., *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007 [0811.4622].

[64] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, *Nuclear Physics B* **126** (1977) 298.

[65] T. Sjöstrand, *A model for initial state parton showers*, *Physics Letters B* **157** (1985) 321.

[66] T. D. Gottschalk, *Backwards evolved initial state parton showers*, *Nuclear Physics B* **277** (1986) 700.

[67] T. Sjostrand, S. Mrenna and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852 [0710.3820].

[68] R. Brun et al., *GEANT: Detector Description and Simulation Tool*, CERN Program Library. CERN, Geneva, 1993, 10.17181/CERN.MUHF.DMJ1.

[69] S. Agostinelli, J. Allison, K. Amako, J. Apostolakis, H. Araujo, P. Arce et al., *Geant4a simulation toolkit*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **506** (2003) 250.

[70] M. Bahr et al., *Herwig++ Physics and Manual*, *Eur. Phys. J. C* **58** (2008) 639 [0803.0883].

[71] J. Bellm et al., *Herwig 7.1 Release Note*, 1705.06919.

[72] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079 [1405.0301].

[73] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [hep-ph/0409146].

[74] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070 [0709.2092].

[75] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043 [1002.2581].

[76] Sherpa Collaboration, *Event Generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034 [1905.09127].

[77] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD matrix elements + parton showers*, *JHEP* **11** (2001) 063 [hep-ph/0109231].

[78] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *A critical appraisal of NLO+PS matching methods*, *JHEP* **09** (2012) 049 [1111.1220].

[79] S. Hoeche, F. Krauss, M. Schonherr and F. Siegert, *QCD matrix elements + parton showers: The NLO case*, *JHEP* **04** (2013) 027 [1207.5030].

[80] D. J. Lange, *The EvtGen particle decay simulation package*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **462** (2001) 152.

[81] Y. Li and F. Petriello, *Combining QCD and electroweak corrections to dilepton production in FEWZ*, *Phys. Rev. D* **86** (2012) 094034 [1208.5967].

[82] K. Melnikov and F. Petriello, *Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$*, *Phys. Rev. D* **74** (2006) 114017 [hep-ph/0609070].

[83] R. Gavin, Y. Li, F. Petriello and S. Quackenbush, *FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order*, *Comput. Phys. Commun.* **182** (2011) 2388 [1011.3540].

[84] C. Anastasiou, L. J. Dixon, K. Melnikov and F. Petriello, *High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, *Phys. Rev. D* **69** (2004) 094008 [hep-ph/0312266].

[85] C. Duhr, F. Dulat and B. Mistlberger, *Drell-Yan Cross Section to Third Order in the Strong Coupling Constant*, *Phys. Rev. Lett.* **125** (2020) 172001 [2001.07717].

[86] J. McGowan, T. Cridge, L. A. Harland-Lang and R. S. Thorne, *Approximate N^3LO Parton Distribution Functions with Theoretical Uncertainties: MSHT20aN^3LO PDFs*, 2207.04739.

[87] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.

[88] R. Pasechnik and M. Šumbera, *Phenomenological Review on Quark–Gluon Plasma: Concepts vs. Observations*, *Universe* **3** (2017) 7 [1611.01533].

[89] E. Mobs, *The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019*, .

[90] R. Steerenberg et al., *Operation and Performance of the Cern Large Hadron Collider During Proton Run 2*, in *Proc. 10th International Particle Accelerator Conference (IPAC'19), Melbourne, Australia, 19-24 May 2019*, no. 10 in International Particle Accelerator Conference, (Geneva, Switzerland), pp. 504–507, JACoW Publishing, Jun., 2019, DOI.

[91] CMS Collaboration, *The CMS experiment at the CERN LHC*, *Journal of Instrumentation* **3** (2008) S08004.

[92] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, *Journal of Instrumentation* **3** (2008) S08002.

[93] LHCb Collaboration, *The LHCb Detector at the LHC*, *Journal of Instrumentation* **3** (2008) S08005.

[94] TOTEM Collaboration, *The TOTEM Experiment at the CERN Large Hadron Collider*, *Journal of Instrumentation* **3** (2008) S08007.

[95] LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, *Journal of Instrumentation* **3** (2008) S08006.

[96] MoEDAL Collaboration Collaboration, *Technical Design Report of the MoEDAL Experiment*, tech. rep., Jun, 2009. https://cds.cern.ch/record/1181486.

[97] J. L. Feng, I. Galon, F. Kling and S. Trojanowski, *ForwArd Search ExpeRiment at the LHC*, *Phys. Rev. D* **97** (2018) 035001.

[98] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003.

[99] ATLAS Collaboration Collaboration, *ATLAS inner detector: Technical Design Report, 2*, Technical design report. ATLAS. CERN, Geneva, 1997.

[100] ATLAS Collaboration Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, tech. rep., Sep, 2010. https://cds.cern.ch/record/1291633.

[101] ATLAS Collaboration Collaboration, *ATLAS calorimeter performance: Technical Design Report*, Technical design report. ATLAS. CERN, Geneva, 1996.

[102] ATLAS Collaboration Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*, Technical design report. ATLAS. CERN, Geneva, 1996, 10.17181/CERN.FWRW.FOOQ.

[103] ATLAS Collaboration Collaboration, *ATLAS tile calorimeter: Technical Design Report*, Technical design report. ATLAS. CERN, Geneva, 1996, 10.17181/CERN.JRBJ.7O28.

[104] ATLAS Collaboration Collaboration, *ATLAS muon spectrometer: Technical Design Report*, Technical design report. ATLAS. CERN, Geneva, 1997.

[105] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2015*, *Eur. Phys. J. C* **77** (2017) 317 [1611.09661].

[106] W. P. Vazquez and on behalf of the ATLAS Collaboration, *The ATLAS Data Acquisition System in LHC Run 2*, *Journal of Physics: Conference Series* **898** (2017) 032017.

[107] ATLAS Collaboration, *Measurement of the Inelastic Proton-Proton Cross Section at $\sqrt{s} = 13$ TeV with the ATLAS Detector at the LHC*, *Phys. Rev. Lett.* **117** (2016) 182002 [1606.02625].

[108] G. Avoni, M. Bruschi, G. Cabras, D. Caforio, N. Dehghanian, A. Floderus et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, *JINST* **13** (2018) P07017.

[109] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, 2212.09379.

[110] S. van der Meer, *Calibration of the effective beam height in the ISR*, tech. rep., CERN, Geneva, 1968. https://cds.cern.ch/record/296752.

[111] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, *JINST* **15** (2020) P04003 [1911.04632].

[112] ATLAS Collaboration, *A neural network clustering algorithm for the ATLAS silicon pixel detector*, *JINST* **9** (2014) P09009 [1406.7690].

[113] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector in Run-2*, *Eur. Phys. J. C* **80** (2020) 1194 [2007.07624].

[114] S. Boutle, D. Casper, B. Hooberman, K. Grimm, B. Gui, G. Lee et al., *Primary vertex reconstruction at the ATLAS experiment*, *Journal of Physics: Conference Series* **898** (2017) 042056.

[115] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at* $\sqrt{s}$ *= 13 TeV*, *Eur. Phys. J. C* **79** (2019) 639 [1902.04655].

[116] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke et al., *Calorimeter Clustering Algorithms: Description and Performance*, tech. rep., CERN, Geneva, 2008. https://cds.cern.ch/record/1099735.

[117] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data*, *JINST* **14** (2019) P03017 [1812.03848].

[118] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, *JINST* **14** (2019) P12006 [1908.00005].

[119] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at* $\sqrt{s} = 13$ *TeV*, *Eur. Phys. J. C* **81** (2021) 578 [2012.00578].

[120] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s}$ *=13 TeV*, *Eur. Phys. J. C* **76** (2016) 292 [1603.05598].

[121] ATLAS Collaboration, *Measurement of fiducial and differential* W^+W^- *production cross-sections at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *Eur. Phys. J. C* **79** (2019) 884 [1905.04242].

[122] ATLAS Collaboration, *Measurement of the* $Z(\to \ell^+\ell^-)\gamma$ *production cross-section in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *JHEP* **03** (2020) 054 [1911.04813].

[123] ATLAS Collaboration, *Evidence for electroweak production of two jets in association with a* $Z\gamma$ *pair in pp collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, *Phys. Lett. B* **803** (2020) 135341 [1910.09503].

[124] ATLAS Collaboration, *Measurement of the double-differential high-mass Drell-Yan cross section in pp collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector*, *JHEP* **08** (2016) 009 [1606.01736].

[125] ATLAS Collaboration, *Search for new non-resonant phenomena in high-mass dilepton final states with the ATLAS detector*, *JHEP* **11** (2020) 005 [2006.12946].

[126] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, *Eur. Phys. J.* **C70** (2010) 823 [1005.4568].

[127] R. D. Ball et al., *Parton distributions with LHC data*, *Nucl. Phys. B* **867** (2013) 244 [1207.1303].

[128] ATLAS Collaboration Collaboration, *The Pythia 8 A3 tune description of ATLAS minimum bias and inelastic measurements incorporating the Donnachie-Landshoff diffractive model*, tech. rep., CERN, Geneva, Aug, 2016. https://cds.cern.ch/record/2206965.

[129] ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector*, *JHEP* **09** (2014) 145 [1406.3660].

[130] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin et al., *New parton distributions for collider physics*, *Phys. Rev. D* **82** (2010) 074024 [1007.2241].

[131] J. Pumplin et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, *JHEP* **07** (2002) 012 [hep-ph/0201195].

[132] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays*, *Eur. Phys. J. C* **45** (2006) 97 [hep-ph/0506026].

[133] N. Davidson, T. Przedzinski and Z. Was, *PHOTOS Interface in C++: Technical and physics documentation*, *Comput. Phys. Commun.* **199** (2016) 86 [1011.0937].

[134] J. R. Andersen et al., *Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report*, 1405.1067.

[135] T.-J. Hou et al., *New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC*, *Phys. Rev. D* **103** (2021) 014013 [1912.10053].

[136] S. G. Bondarenko and A. A. Sapronov, *NLO EW and QCD proton-proton cross section calculations with mcsanc-v1.01*, *Comput. Phys. Commun.* **184** (2013) 2343 [1301.3687].

[137] M. Cano Bret, X. Li, T. Hryn'ova, M. Lu, E. Rizvi, S. D. Lawlor et al., *Measurement of the high-mass Drell-Yan double-differential cross-section in pp collisions at $\sqrt{s} = 13$ TeV*, tech. rep., CERN, Geneva, 2019. https://cds.cern.ch/record/2703975.

[138] E. Bothmann et al., *Event generation with Sherpa 2.2*, *SciPost Phys.* **7** (2019) 034 [1905.09127].

[139] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, *JHEP* **12** (2008) 039 [0808.3674].

[140] F. Buccioni, J.-N. Lang, J. M. Lindert, P. Maierhöfer, S. Pozzorini, H. Zhang et al., *OpenLoops 2*, *Eur. Phys. J. C* **79** (2019) 866 [1907.13071].

[141] F. Cascioli, P. Maierhöfer and S. Pozzorini, *Scattering Amplitudes with Open Loops*, *Phys. Rev. Lett.* **108** (2012) 111601 [1111.5206].

[142] A. Denner, S. Dittmaier and L. Hofer, COLLIER*: A fortran-based complex one-loop library in extended regularizations*, *Comput. Phys. Commun.* **212** (2017) 220 [1604.06792].

[143] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani–Seymour dipole factorisation*, *JHEP* **03** (2008) 038 [0709.1027].

[144] S. Höche, F. Krauss, S. Schumann and F. Siegert, *QCD matrix elements and truncated showers*, *JHEP* **05** (2009) 053 [0903.1219].

[145] NNPDF Collaboration, *Parton distributions for the LHC run II*, *JHEP* **04** (2015) 040 [1410.8849].

[146] NNPDF Collaboration, *Illuminating the photon content of the proton within a global PDF analysis*, *SciPost Phys.* **5** (2018) 008 [1712.07053].

[147] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten et al., *An introduction to PYTHIA 8.2*, *Comput. Phys. Commun.* **191** (2015) 159 [1410.3012].

[148] ATLAS Collaboration, "ATLAS Pythia 8 tunes to 7 TeV data." ATL-PHYS-PUB-2014-021, 2014.

[149] D. J. Lange, *The EvtGen particle decay simulation package*, *Nucl. Instrum. Meth. A* **462** (2001) 152.

[150] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber, *Single-top hadroproduction in association with a W boson*, *JHEP* **07** (2008) 029 [0805.3067].

[151] M. Czakon and A. Mitov, *Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders*, *Comput. Phys. Commun.* **185** (2014) 2930 [1112.5675].

[152] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, tech. rep., CERN, Geneva, 2019. https://cds.cern.ch/record/2677054.

[153] Public ATLAS Luminosity Results for Run-2 of the LHC. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[154] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, *Eur. Phys. J. C* **80** (2020) 47 [1909.00761].

[155] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, *JINST* **15** (2020) P09015 [2004.13447].

[156] G. D'Agostini, *Improved iterative Bayesian unfolding*, *arXiv e-prints* (2010) arXiv:1010.0632 [1010.0632].

[157] M. Zinser, D. K. Abhayasinghe, G. Artoni, T. Berry, R. Brock, M. K. Bugge et al., *Search for new high-mass resonances in the dilepton final state using proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, tech. rep., CERN, Geneva, 2016. https://cds.cern.ch/record/2217641.

[158] ATLAS IFF fake lepton background estimation tools. https://gitlab.cern.ch/atlas/athena/-/tree/21.2/PhysicsAnalysis/AnalysisCommon/FakeBkgTools.

[159] Standard Model Cross Sections for Diboson/Triboson Productions. https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/SMDC14xsecs.

[160] HERAverager data combination tool. https://wiki-zeuthen.desy.de/HERAverager.

[161] H1 Collaboration, *Measurement of the Inclusive ep Scattering Cross Section at Low Q^2 and x at HERA*, *Eur. Phys. J. C* **63** (2009) 625 [0904.0929].

[162] H1, ZEUS Collaboration, *Impact of jet-production data on the next-to-next-to-leading-order determination of HERAPDF2.0 parton distributions*, *Eur. Phys. J. C* **82** (2022) 243 [2112.01120].

[163] S. Alekhin, J. Bluemlein, S.-O. Moch and R. Placakyte, *The new ABMP16 PDF*, *PoS* **DIS2016** (2016) 016 [1609.03327].

[164] LHCb Collaboration, *Test of lepton universality in beauty-quark decays*, *Nature Phys.* **18** (2022) 277 [2103.11769].

[165] Muon g-2 Collaboration, *Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm*, *Phys. Rev. Lett.* **126** (2021) 141801 [2104.03281].

[166] A. Greljo and D. Marzocca, *High-p_T dilepton tails and flavor physics*, *Eur. Phys. J. C* **77** (2017) 548 [1704.09015].

[167] ALEPH, DELPHI, L3, OPAL, SLD, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group Collaboration, *Precision electroweak measurements on the Z resonance*, *Phys. Rept.* **427** (2006) 257 [hep-ex/0509008].

[168] EFT Fitter for UNfolded measurements. https://gitlab.cern.ch/eft-tools/eft-fun.

[169] ATLAS Collaboration, *Measurements of differential cross-sections in four-lepton events in 13 TeV proton-proton collisions with the ATLAS detector*, *JHEP* **07** (2021) 005 [2103.01918].

[170] ATLAS Collaboration, *Differential cross-section measurements for the electroweak production of dijets in association with a Z boson in proton–proton collisions at ATLAS*, *Eur. Phys. J. C* **81** (2021) 163 [2006.15458].

[171] ATLAS Collaboration, *Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **79** (2019) 535 [1902.05759].

[172] ATLAS Collaboration, *Combined effective field theory interpretation of differential cross-sections measurements of WW, WZ, 4l, and Z-plus-two-jets production using ATLAS data*, tech. rep., CERN, Geneva, 2021. http://cds.cern.ch/record/2776648.

[173] ATLAS Collaboration, *Combined effective field theory interpretation of Higgs boson and weak boson production and decay with ATLAS data and electroweak precision observables*, tech. rep., CERN, Geneva, 2022. https://cds.cern.ch/record/2816369.

[174] *Appendix A: Relationship between the Hessian and Covariance Matrix for Gaussian Random Variables*, in *Bayesian Methods for Structural Dynamics and Civil Engineering*, pp. 257–262, John Wiley & Sons, Ltd, (2010), https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470824566.app1, DOI.

[175] CMS Collaboration, *Measurement of the differential Drell-Yan cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV*, *JHEP* **12** (2019) 059 [1812.10529].

[176] ZEUS Collaboration, *Search for effective Lorentz and CPT violation using ZEUS data*, 2212.12750.

[177] D0 Collaboration, *Search for Violation of Lorentz Invariance in Top Quark Pair Production and Decay*, *Phys. Rev. Lett.* **108** (2012) 261603.

[178] D0 Collaboration, *Search for Violation of CPT and Lorentz Invariance in B_s^0 Meson Oscillations*, *Phys. Rev. Lett.* **115** (2015) 161601.

[179] ATLAS Collaboration, *Luminosity Determination using $Z \to \ell\ell$ events at $\sqrt{s} = 13$ TeV with the ATLAS detector*, tech. rep., CERN, Geneva, Feb, 2021. https://cds.cern.ch/record/2752951.

[180] ATLAS Collaboration, *Measurement of the transverse momentum distribution of Drell–Yan lepton pairs in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **80** (2020) 616 [1912.02844].

[181] H. Lyons, *Measurement of Z-boson production and luminosity monitoring using 5.02 TeV and 13 TeV ATLAS Run-2 pp collision data*, Ph.D. thesis, 2022. 10.17638/03165542, https://livrepository.liverpool.ac.uk/id/eprint/3165542.

[182] U. Klein, E. Lunghi, N. L. Sherrill, R. Gonzalez Lopez, Y. Enari and Y. Abulaiti, *Testing spacetime with dileptons: Search for Lorentz and CPT violation in $Z \to \ell\ell$ events in $\sqrt{s} = 13$ TeV pp collisions using the ATLAS detector*, tech. rep., CERN, Geneva, 2022. https://cds.cern.ch/record/2839176.

[183] V. A. Kostelecký and M. Mewes, *Signals for Lorentz violation in electrodynamics*, *Phys. Rev. D* **66** (2002) 056005.

[184] R. Bluhm, V. A. Kostelecky, C. D. Lane and N. Russell, *Probing Lorentz and CPT violation with space based experiments*, *Phys. Rev. D* **68** (2003) 125008 [hep-ph/0306190].

[185] Y. Ding and V. A. Kostelecký, *Lorentz-violating spinor electrodynamics and Penning traps*, *Phys. Rev. D* **94** (2016) 056008 [1608.07868].

[186] R. Coates, G. Janacek and K. Lever, *Monte Carlo simulation and random number generation*, *IEEE Journal on Selected Areas in Communications* **6** (1988) 58.

[187] Karnevskiy, M and Lin, T-H and Schott, M and Zimmermann, C and Tomlinson, L and Webb, S and Lisovyi, Mykhailo, *Measurement of the transverse momentum and azimuthal decorrelation of leptons in DrellYan events at sqrts = 8 TeV with ATLAS*, Tech. Rep. ATL-COM-PHYS-2014-281, CERN, Geneva, Apr, 2014. https://cds.cern.ch/record/1694674.

[188] D. K. Abhayasinghe, T. Berry, M. K. Bugge, E. Dreyer, P. J. Falke, D. Hayden et al., *Search for new high-mass phenomena in the dilepton final state using the full Run-2 proton-proton collision dataset at $\sqrt{s}$ =13 TeV with the ATLAS detector*, tech. rep., CERN, Geneva, 2018. https://cds.cern.ch/record/2622420.