Study of the Drell-Yan process with the ATLAS detector at the LHC

Francesco Giuli
St. John's College, Oxford

A thesis submitted to the University of Oxford
in partial fulfilment of the requirements for the award
of the degree of **Doctor of Philosophy**

Trinity Term, 2018

Cause tramps like us,
baby we were born to run...
(Bruce Springsteen 'The Boss' - 1975)

But I still haven't found
what I'm looking for...
(U2 - 1987)

*Francè, e smettila di dire c@#$%*e!*
(Nonno Lello - 2014)

Study of the Drell-Yan process with the ATLAS detector at the LHC

Francesco Giuli
St. John's College, Oxford

A thesis submitted to the University of Oxford
in partial fulfilment of the requirements for the award
of the degree of **Doctor of Philosophy**

Trinity Term, 2018

Abstract

The Large Hadron Collider (LHC) at CERN, at which protons collide at unprecedented center of mass energies and very high instantaneous luminosity, gives unique possibilities for precise tests of the Standard Model and to search for new physics phenomena. A precise prediction of the processes at the LHC is essential and the knowledge of the parton density functions (PDFs) of the proton plays a key role in this.
In this thesis, three analyses covering different aspects of the Drell-Yan process are presented. In the first analysis a new heavy neutral gauge boson, a so-called Z' boson, is searched for. Those new gauge bosons are predicted by some theories extending the Standard Model gauge group to solve some of its conceptual problems. Decays of the Z' boson in final states with a lepton ($\ell = e^{\pm}, \mu^{\pm}$) and the corresponding anti-lepton are considered. The data in use were collected by the ATLAS experiment in the year 2015+2016 at a center of mass energy of $\sqrt{s}$ = 13 TeV. The collected data corresponds to an integrated luminosity of 36.1 fb^{-1}. The Standard Model prediction for the expected background is estimated with Monte Carlo simulations and methods based on data. The resulting spectrum of the di-lepton invariant mass is tested, using statistical methods, for differences between data and the Standard Model predictions. No significant deviation from the Standard Model predictions is found. Much more stringent limits have been set on the parameters of interest for various models comparing with the Run 1 results. For new gauge bosons, Z'_{SSM} , Z'_{χ} and Z'_{ψ} predicted in extended gauge theory models, the observed lower limits on mass have been set to be 4.5 TeV, 4.1 TeV and 3.8 TeV at 95% confidence level, respectively. While for the Contact

Interactions, the observed lower limits on the corresponding energy scale Λ are set between 24 TeV and 40 TeV depending on the different scenarios.
In the second analysis, the double-differential cross section measurement of the process $pp \to Z/\gamma^* \to \ell^+\ell^-$ at a center of mass energy of $\sqrt{s}$ = 8 TeV is used in a PDF fit to put constraints on the photon PDF of the proton. The analysed data were recorded by the ATLAS experiment in the year 2012 and correspond to an integrated luminosity of 20.3 fb^{-1}. These data are sensitive to the photon structure of the photon and allow a determination of $x\gamma(x, Q^2)$ with uncertainties at the 30% level for $0.02 \leq x \leq 0.1$. The results of this study, is in agreement and exhibit smaller PDF uncertainties that the only other existing photon PDF fit from LHC data, the NNPDF3.0QED analysis, based on previous LHC Drell-Yan measurements. The results are also in agreement within uncertainties with two recent theoretical calculations of the photon PDF, LUXqed and HKR16.
In the third analysis, a measurement of the double-differential Drell-Yan cross section in the di-muon channel is presented differentially in terms of di-muon invariant mass and double differentially in terms of di-muon invariant mass and rapidity. The cross sections are measured using 3.2 fb^{-1} of ATLAS data taken during 2015 at a centre of mass energy of $\sqrt{s}$ = 13 TeV. Dedicated studies to address the main challenges of the analysis, namely the understanding of the trigger SF and of the muon isolation are reported, and the data-driven methods for estimating the QCD/multijet background are described in detail. The extracted fiducial cross sections agree within uncertainties with the theoretical predictions from `Powheg` interfaced with the `Pythia v8.1` parton shower program.
Finally, due to the page limit in this thesis, my work for the ATLAS Collaboration is not presented because it was not strictly related to the physics covered here. This work consists in six public (PUB) notes: two on the MC modelling of V+jets processes [1,2], three on the MC modelling of the $t\bar{t}$+jets process (with particular emphasis on the implementation of new NLO merging schemes in the ATLAS software) [3–5] and one about the inclusion of $t\bar{t}$ data recorded by the ATLAS detector at 8 TeV (both lepton+jets and di-lepton decay channels) in a combined PDF fit, together with the ep deep inelastic scattering data from HERA and the inclusive ATLAS W and Z boson data recorder at 7 TeV [6]. During my PhD course I got involved in other publications which are mentioned in Chapter 4.

Declaration

I, Francesco Giuli, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.
I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.
I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.
I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Francesco Giuli

Acknowledgements

Ho pensato molto se scrivere o meno questa sezione della tesi. Ho tergiversato a lungo ma infine mi sono convinto che ringraziare le persone che hanno reso possibile questo "papello" fosse quantomeno doveroso.

First of all, thanks to my supervisors and mentors, Mandy, Massimo, Eram and Juan for their advices and guidance which have been invaluable. Their dedication and passion inspired me throughout my PhD.
Thanks to the members of the `xFitter` Developers' Team, with a special mention to Sasha, Ringaile and Voica. I cannot thank them enough for the time they spent teaching me and answering my questions and doubts.
Thank you Sue, Kim, Jennifer and Michelle, for your kindness and support with administration. You are really precious to the department and you do a great job.
And now we can switch language to italian, with some brief parts still in english.

Innanzitutto, il primo ringraziamento va alla mia famiglia, a quel sant'uomo di mio padre e quelle due pie donne di mia madre e mia sorella che mi hanno supportato e sopportato sempre e comunque, soprattutto nei momenti meno sereni e più burrascosi della mia vita (vedi mio "secondo genetlico" tanto per dirne una). Perché lo so, non è facile avere a che fare con me in determinate situazioni, o per meglio dire, quasi sempre. Grazie di cuore a Nonno Lello, a Nonna Peppa ed a Nonna Lillina, ai miei zii ed ai miei cugini, per avermi strappato un sorriso anche nei momenti più critici e per essere stati sempre pronti a darmi un consiglio o conforto quando era necessario.

Grazie ai "Preci" (che più preci sono si può), per esserci dal Settembre 2004 e per esserci sempre. Per avere reso indimenticabili gli anni del Socrate, per i nostri aperitivi, le inummerevoli serate e tutti i momenti fantastici trascorsi assieme: se é stato così tosto andare via da Roma, la colpa (probabilmente il merito) é principalmente vostra. Spero possiate far parte della mia vita ancora a lungo, perché molta della forza che ho la traggo dal nostro stare assieme e dal nostro volerci bene.
Grazie agli amici di una vita, a Helenen, Daniele (Barabbasauron) e Stefano (The Mad Speaker), perché mi conoscete così bene che ogni vostra parola va a colpire là, esattamente dove c'è più bisogno.

Grazie a tutti membri della crew italiana del CERN, per avere trasformato Ginevra in una città vivibile ed un luogo in cui si vuole/deve tornare. Per le nostre feste, i nostri venerdì pomeriggio alcolici ad R1, per le succulente linguine all'astice e per le nostre serate a L' Usine. Grazie a tutti voi, veramente: Adele, Secsi, String, Dibbi, Mastro, Giuliano, Peter, Arianna il Capitano, Daniele, Silvia, Beatrice, Camilla, Giulia, Federica, Grazia, Silvia (Pretty Shoulder - Frosinone Culone), Fabio, il Killer, Guido, Lucrezia, Bettooooo, Renado ed il suo gallo calante, Valentina, Francesca, Serena, Chicco, Frankie, Scodella, Simone, Veronica, Alessandra, Simona, Matteo, Elena, Elena, Pippo, Cirotto, Gioacchino, Ghepy, Maurizio ed Umberto. Un grazie

anche ai laureandi e dottorandi di ATLAS della Sapienza, che mi hanno accolto in Saletta, mi hanno reso uno di loro e mi hanno sopportato mentre cercavo di presentare durante vari meetings. E per le partite a tresette con amaro annesso, ovviamente!
Grazie a Peppe ed Emanuele, per le risate e per le nostre aricciate sempre molto ma molto di conforto ogni volta che tornavo in patria.

Thanks to all the past and present members of Emilia, Jamie and Kris, because sharing da hauz with you was simply amazing and the time spent together was marvellous, gorgeous, incredible.

Grazie a tutti membri di "Hola" ad Oxford. Per aver reso e continuare a rendere il Regno Unito ed Oxford un posto accogliente, una seconda casa in cui ho sempre voglia di tornare quasi esclusivamente per rivedere voi. In particolare, grazie a Luca per avermi iniziato ai misteri della QCD, per le ore di sport viste ed ascoltate assieme in ufficio, per il suo "essere svizzero", per le canzoni di Davide Van De Sfroos e per le conferenze sulla neve, sperando di poter assistere ai reciproci talks molto presto. E grazie ad Ale, per essere un amabile ragazzo di Roma Nord e più semplicemente perché è un tenero cucciolo di koala in tutto e per tutto: prepara la brandina a Maryland, che prima o poi vengo a trovarti.
Grazie anche a Marco e Valerio per il loro tempo dedicatomi e per le innumerevoli spiegazioni in questi anni.

Thanks to my Oxford basketball teammates, with a special thanks to the members of the team I joined during my first year: Ludo, Gio, Floris, Sergio, Henry, Phil, Tyler, Joey, Declan and George. Thanks for our incredibly famous "fagiolate" and for playing with me, even when I was really unbearable and I was shouting at everybody. And thanks to my coach Mathias, because getting angry with you was a lot of fun. Thanks for all the hours spent together in the gym and on the court.
Grazie a tutti membri della "Abusive Oxford Italian Society", perché anche se siamo stati poco assieme, siete stati un parte importante dei miei giorni trascorsi ad Oxford.
Thanks to all the PhD students and postdocs of the ATLAS Oxford group I have had the luck to meet. Thanks for all the fruitful discussions, especially the non-physics related ones, for helping me in debugging my code and for your answers to my 'silly' questions. Thanks for all of the lunches, tea breaks and laughs shared.

Grazie ai vari Infected Mushrooms, Paul Kalkbrenner, Knife Party, Mumford & Sons, Bruce Springsteen, U2, Dream Theater, Coro Musicanova, The Who, The Yes, Iron Maiden, The Band, King Crimson, Gentle Giant, Flogging Molly, Dire Straights, Pink Floyd, Gigi D'Ag, Eifell 65, The Clash, Madness, Modena City Ramblers, Lynyrd Skynyrd, Bod Dylan, Eric Clapton, The Bothy Band, Anonima Armonisti, The Dubliners, Francesco Guccini, Fabrizio De Andrè, Invano Fossati, Francesco de Gregori, Kings' Singers, Pentatonix, Florence and the Machine, Daft Punk, The Pogues, El Mago de Oz, The Mahones, Bongo Botrako, La Pegatina, Los Deliquentes, Ska-P, The Chieftains, The Horrorist e tanti tanti altri, per avermi

accompagnato nelle mie interminabili giornate passate davanti al terminale.

Grazie a Luppolo 12, il GB, Renatino, Celestino, il Bar dei Brutti, El Chiringuito Libre, il Macche, il 4:20, L'elephant dans la canette, The Lamb and Flag, The King's Arms, The Turf Tavern and The Gardener's Arms per gli spritz, gli innumerevoli cocktails e le svariate pinte versate, per la ritrovata passione nel bere whisky e per avermi fatto capire cos'è il G&T, per tutte le volte che con solo una serata sono riuscito a ritrovare il buon umore ed anche per aver contribuito al mio alcolismo (non troppo anonimo) in maniera importante ed anche sostanziale direi.

And last but definitely not least... Grazie a Valeria, in arte L.A. Zima (potrei citare innumerevoli diminutivi e vezzeggiativi ma mejo de no!), per avermi preso e semplicemente avermi cambiato la vita. Grazie per il suo amore, i suoi mille sorrisi, la sua totale dedizione e per aver cominciato (E CONTINUATO!) questa avventura assieme a me, nonostante le mille peripezie, i vari travagli affrontati ed i quasi 2000 km tra St. Giles' Street e Via Carlo Conti Rossini. Sono certo che ne vedremo ancora delle belle, siamo solo all'inizio e mo cominciamo a divertisse!
That's all folks!

Contents

List of Figures

List of Tables

Chapter 1

Theory

This chapter introduces the theoretical background needed to understand the measurements presented in this work and to put them into the context of its potential theoretical impact. The first section describes the Standard Model (description based on [7]), while an introduction to the physics of proton-proton collision is discussed in the second section. Then the Drell-Yan (DY) process and the impact of high- and low-mass DY measurements are highlighted in Sec. 1.3. Finally, Sec. 1.4 presents the limitations of the SM and a brief introduction to the beyond the Standard Model (BSM) physics, with a particular emphasis on BSM theories involving a neutral heavy gauge boson.

1.1 The Standard Model of Particle Physics

After years of theoretical studies and experimental research scientists have come to the conclusion that the universe is constructed of a specific number of elementary building blocks, which are governed by four fundamental forces. These perceptions provide a deep insight into the structure of the micro cosmos and help to achieve a better understanding of the construction and coherence of matter. The Standard Model (SM) describes the interactions of the known fermionic matter particles, quarks and leptons, via the strong, the electromagnetic and the weak forces based on the principle of local gauge invariance, e.g. invariance under phase transformations depending on the space-time coordinates. The gravitational force is negligible in atomic and nuclear physics since quantum gravity effects are expected to be comparable with the effects of the other interactions only at very high energies at the Planck scale of $\approx 10^{19}$ GeV.
On the theoretical point of view, the SM is a quantum field theory that is described by the group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, which has 8 + 3 + 1 = 12 generators and gauge fields: the electromagnetic and weak interactions [8, 9] are described by the $SU(2)_L \otimes U(1)_Y$ symmetry group, while $SU(3)_C$ is the symmetry group associated to the strong interactions of Quantum Chromodynamics (QCD) [10]. A vector boson is associated to each generator of the gauge symmetry groups of the SM acting as mediator of the interaction: eight gluons are associated to the $SU(3)_C$ color group and the $W^{\pm}, Z$ and γ are associated to the so-called electroweak

Boson	Mass	Electric Charge	Weak Hypercharge
γ	$< 10^{-18}$ eV	0	0
Z	91.188 GeV	0	0
$W^{\pm}$	80.40 GeV	± 1	0
g	0 eV	0	0

Table 1.1: Gauge bosons in the SM.

Type	First Generation		Second generation		Third generation	
	Flavour	Mass	Flavour	Mass	Flavour	Mass
Lepton	ν_e	< 2 eV	ν_μ	< 2 eV	ν_τ	< 2 eV
	e	512 keV	μ	105.6 MeV	τ	1.777 GeV
Quark	u	$2.3^{+0.7}_{-0.5}$ MeV	c	1.275 ± 0.025 GeV	t	173.07 ± 1.24 GeV
	d	$4.8^{+0.5}_{-0.3}$ MeV	s	95 ± 5 MeV	b	4.18 ± 0.03 GeV

Table 1.2: Fermionic particle content of the SM. The mass values are taken from the Particle Data Group [13].

symmetry $SU(2)_L \otimes U(1)_Y$ (Tab. 1.1).
The particles, which are currently believed to be elementary, can be categorized according to their spin S into *fermions* with half-integer spin, and *bosons* with integer spin values. Fermions are further divided into leptons which do not interact via the strong force and quarks which undergo strong interactions. Furthermore, the fermions come in three different generations; there are always two different flavours of leptons and quarks per generation which yields in total 12 fermions as shown in Table 1.2. Each quark flavour is a colour triplet and carries electroweak charges including electric charges of +2/3 and -1/3 for up-type and down-type quarks respectively. Leptons are colourless and have electroweak charge; the electrons, muons and taus leptons have electric charge equal to -1, while the associated neutrinos (ν_e, ν_μ and ν_τ) are electrically neutral. The weak hypercharge is +1/3 for both the up-type and down-type and it is -1 for both the electrons, muons and taus and the associated neutrinos. Opposite sign electric charges are carried by the respective anti-particles. The masses of the above-mentioned particles are introduced without spoiling the electroweak gauge symmetry via the mechanism of *spontaneous symmetry breaking* [11,12], an additional complex scalar field is required for this purpose and give rise to a new scalar particle, the Higgs boson h, that interacts with other particles with a strength proportional to their masses.

1.1.1 Electroweak Theory

The electroweak interaction is an unified description of the electromagnetic and weak force of the SM of particle physics. Although the two interactions seem to be different at low energies, beyond the unification energy of about 100 GeV,

they are merged together into one force.
From the mathematical point of view, this unification is accomplished under an $SU(2)_L \otimes U(1)_Y$ non-abelian gauge symmetry. Y is the hyper-charge operator and it is defined as $Y = 2(Q - T^3)$, where Q is the conserved charge. The weak isospin operators $T^{1,2,3}$ are the generators of the $SU(2)_L$ symmetry. The corresponding gauge bosons are three W bosons of the weak isospin from $SU(2)_L$ (namely, $W_\mu^{1,2,3}$) and they only couple to left-handed fermions, which are ordered in $SU(2)$ doublets (right-handed fermions are ordered in $SU(2)$ singlets), while B_μ is the gauge field associated with the weak hyper-charge.
The Lagrangian of the electroweak theory is defined as follows:

$$\mathcal{L} = \bar{\psi} i \gamma_\mu D^\mu \psi - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} - \frac{1}{2} Tr(W_{\mu\nu} W^{\mu\nu}) \tag{1.1}$$

where the Dirac field of a spin-1/2 particle is represented by ψ. The kinematic of the fermion field ψ is described by the first term in the equation. Local invariance under $SU(2)_L \otimes U(1)_Y$ symmetry transformation is ensured by introducing the gauge covariant derivative

$$D_\mu = \partial_\mu - i g W_\mu^a T^a + i g' B_\mu Y \tag{1.2}$$

where $W_\mu^a (a = 1, 2, 3)$ and B_μ represents the gauge fields; they are massless vector fields and they are introduced when requiring Lorentz invariance, given the fact that the induced symmetry would be violated by any mass term of the gauge field introduced in the Lagrangian. In Eq. 1.1, g and g' denote the gauge coupling of the $SU(2)_L$ and the coupling constant of the $U(1)_Y$ group respectively.
The kinematics of the vector field B_μ is described by the second term in Eq. 1.1, where $B_{\mu\nu}$ represents the field strength tensor for the weak hypercharge field. This quantity is gauge invariant and it is defined as

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.3}$$

Similarly, we can define the field strength tensor for the weak isospin field, adding an additional term according to the self-interaction of the gauge fields,

$$W_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + g f_{abc} W_\mu^b W_\nu^c, \tag{1.4}$$

where f_{abc} are the *structure constants* defined by the commutation relation

$$\left[T^a, T^b\right] = i f^{abc} T^c. \tag{1.5}$$

The gauge fields represents the $W^\pm$, Z and γ bosons, which are the mediator of the weak and electromagnetic interaction respectively. Gauge bosons and fermions are predicted to be massless in order to preserve local gauge symmetry, but as you can see in Table 1.2, there are experimental evidences that show that the masses of $W^\pm$ and Z are different from 0.
To resolve this issue, the mathematical concept of *spontaneous symmetry breaking* is employed in order to acquire the weak bosons masses. This concept is employed by the so-called *Higgs Mechanism*, which introduces a massive scalar boson field

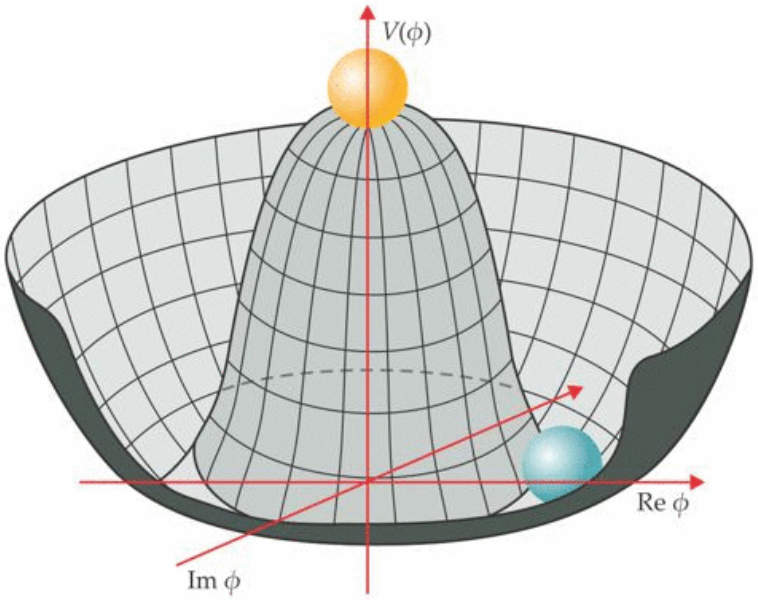


Fig. 1.1: Higgs potential $V(\Phi)$ in the $Re(\phi)$, $Im(\phi)$ plane.

corresponding to the Higgs boson.
The Higgs mechanism [14] postulates an additional doublet Φ of complex scalar fields with the potential

$$V(\Phi) = -\mu^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2, \quad \lambda > 0. \tag{1.6}$$

This additional doublet has four degrees of freedom and its weak hypercharge is chosen to be $Y_\Phi = +1$ such that the lower component of the Φ doublet does not carry any electric charge

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}. \tag{1.7}$$

The Lagrangian for the Higgs sector is given by

$$\mathcal{L}_{Higgs} = (D^\mu\Phi)^\dagger(D_\mu\Phi) - V(\Phi) \tag{1.8}$$

which is invariant under $SU(2)_L \otimes U(1)_Y$ symmetry transformations. Depending on the sign of μ^2, the minimum of the potential is found to be:

$$\mu^2 \leq 0 : \frac{\delta V}{\delta\Phi} = 0 \rightarrow |\Phi_{min}| = 0 \tag{1.9}$$

$$\mu^2 > 0 : \frac{\delta V}{\delta\Phi} = 0 \rightarrow |\Phi_{min}| = \sqrt{\frac{\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{s}} \tag{1.10}$$

The vacuum expectation value $\langle 0|\Phi|0\rangle$ vanishes in the first case and the ground state of the Higgs field is symmetric under $SU(2)_L \otimes U(1)_Y$; however, if $\mu^2 > 0$, the vacuum expectation value differs from 0 and, since no positive charged permanent field is observed, the ground state must take the form

$$\Phi_0 = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix} \tag{1.11}$$

that violates the $SU(2)_L \otimes U(1)_Y$ symmetry. Therefore, three of the original four degrees of freedom of the scalar field are absorbed as longitudinal polarization states of the $W^{\pm}$ and Z bosons, which in this way acquire their masses, while the photon remain massless. The remaining degree of freedom corresponds to a physical massive scalar particle, the Higgs boson. In Fig. 1.1 the form of the Higgs potential in the $Re(\phi)$, $Im(\phi)$ plane is displayed. Considering measurements of the muon decay width Γ_μ, the vacuum expectation value can be constrained to $v \approx 246$ GeV. The fact that the ground state does not exhibit the same symmetry of the Lagrangian shown in Eq. 1.8 is named *spontaneous symmetry breaking.*
The masses of the fermions can be generated by means Yukawa couplings to the Higgs field Φ [15].

1.1.2 Quantum Chromodynamics

The Quantum Chromodynamics (QCD) is the theory of strong interactions and is a non-abelian gauge theory with symmetry group $SU(3)_C$. The interactions between quarks and gluons are described by this theory. The particles affected by the strong coupling require an additional charge to preserve the Pauli principle in QCD. This property, called *colour*, is the QCD analogue of the electric charge in QED. Therefore, quarks and gluons have colour charge, which occurs in red, green and blue. In order to compensate for the local gauge transformations, eight additional fields, which are associated to eight gluons, must be included into the QCD Lagrangian. As already shown in Eq. 1.1, these additional fields are incorporated in the QCD Lagrangian via the covariant derivative D_μ. So, in analogy to the electroweak theory, the gauge invariant QCD Lagrangian can be written as follows:

$$\mathcal{L}_{QCD} = -\frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a + \sum_{i=1}^{N_c} \sum_{f=1}^{N_f} \bar{q}_{f_i} \left(i\gamma^\mu D_\mu - m_f \right) q_{f_i} \tag{1.12}$$

where the sums run over the quark flavours N_f and all colour states N_c. In Eq. 1.12, q_{f_i} represents the quark fermion field. The first term in Eq 1.12 is also called Yang-Mills term and describes the dynamics of the gluon fields A_μ

$$\mathcal{L}_{YM} = -\frac{1}{2} Tr \left(G^a_{\mu\nu} G^{\mu\nu}_a \right). \tag{1.13}$$

The symbol $G^a_{\mu\nu}$, represents the gauge invariant gluon field strength tensor and it is analogous to the electromagnetic field strength tensor, $W_{\mu\nu}$ (Eq. 1.4)

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g f_{abc} G^b_\mu G^c_\nu, \tag{1.14}$$

where f_{abc} are the structure constants of $SU(3)_C$.
QCD enjoys two peculiar properties:

- **Confinement**, which means that the force between quarks does not diminish as they are separated. They are thus forever bound into hadrons such as the proton and the neutron or the pion and kaon. Confinement is seen at large distance e.g. small scale.

- **Asymptotic freedom**, which means that interactions between particles become asymptotically weaker as the energy scale increases and the corresponding length scale decreases. This prediction of QCD was first discovered in the early 1970s by David Politzer, Frank Wilczek and David Gross [16].

The strength of the strong force is represented by α_S and it can be expressed as

$$\alpha_S = \frac{g_S^2}{4\pi}, \tag{1.15}$$

where g_S represents the QCD gauge coupling. This definition is analogous to that of the QED fine structure constant

$$\alpha_{EW} = \frac{e^2}{2\pi} \approx \frac{1}{137}, \tag{1.16}$$

where e is the electron charge. From experiments it has been found that α_S exhibits a different behaviour in the low and high energy regime. Due to its dependence on the energy scale it has acquired the name *running coupling*. The expression of α_S at Leading Order (LO) is [17]

$$\alpha_S(Q^2) = \frac{12\pi}{(33 - 2N_f)\, ln(\frac{Q^2}{\Lambda^2_{QCD}})}, \tag{1.17}$$

Λ_{QCD} indicated the QCD scale and it has been experimentally determined to be $\Lambda_{QCD} \sim 200$ MeV. In high energy collisions, the production process depends on the hard scale Q^2. In the limit of asymptotic freedom, when the coupling constant is very small and tends to be 0 e.g. $\alpha_S(Q^2 \to \infty) \to 0$, the cross section can be expanded perturbatively in powers of α_S. In the limit of confinement, $\alpha_S(Q^2 \to \Lambda^2_{QCD}) \to \infty$ e.g. by separating a quark anti-quark pair, the strong force becomes so large that new $q\bar{q}$ pairs can be produced. In Fig. 1.2 we summarize recent experimental determinations of the QCD running coupling $\alpha_S(Q)$ for different scales, together with the several theoretical prediction. Asymptotic freedom can be seen by the fact that $\alpha_S(Q)$ decreases when Q is increased.

1.2 Proton-Proton collisions

The most powerful tests of perturbative QCD (pQCD) are scattering experiments.
QCD was first established and quantitatively tested studying Deep Inelastic Scattering (DIS) at DIS experiments at SLAC and at collider experiments, such as HERA, LEP and Tevatron. In particular, a three-jet event was first observed by the TASSO experiment at the PETRA accelerator at the DESY laboratory and it was the earliest and most direct evidence for the existence of gluons. A summary of important results by the H1 and ZEUS Collaborations at HERA to test pQCD can be found in [18–24]. Drell-Yan lepton pair production cross sections were also measured by the D0 and CDF Collaborations at the Tevatron and compared to

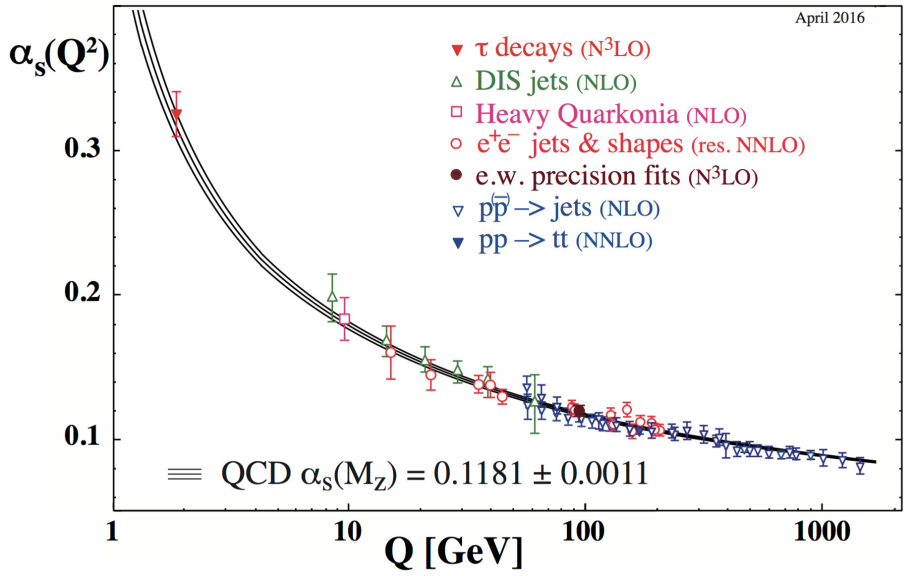


Fig. 1.2: Summary of measurements of α_S as a function of the energy scale Q. The respective degree of QCD perturbation theory used in the extraction of α_S is indicated in brackets (*NLO*: next-to-leading order; *NNLO*: next-to-next-to leading order; *res. NNLO*: NNLO matched with resummed next-to-leading logs; *N3LO*: next-to-NNLO)

pQCD calculations [25, 26]. A summary of measurements by D0 and CDF relevant for QCD is presented in [27]. The experiments at CERN's Large Hadron Collider (LHC) analyse processes at much more higher energy than in previous collider experiments, to allow for a precise test of the theory.
The structure of hadrons (e.g. protons) in high-energy physics and how the partons interact inside the colliding hadrons are described by the so-called *parton model* [28], introduced by R. Feynman in the early 70's.

1.2.1 The structure of the proton

Protons are baryons and therefore composite objects whose complicated dynamics cannot be calculated in the framework of QCD. This compositeness complicates the description of a proton-proton collision with respect to a collision at a lepton collider, where point-like particles collide. The constituents of the proton are three valence quarks (*uud*) embedded in a sea of quark anti quark pairs (called *sea* quarks) and *gluons*.
The main knowledge about the proton structure comes from lepton-proton scattering experiments. A lepton probes a target proton via the exchange of an electroweak boson ($W/Z/\gamma^*$). While elastic lepton-proton scattering allows to investigate the electric and magnetic form factors associated with charge and the magnetic distributions of the proton, the inner proton structure is studied by deep inelastic scattering (DIS) experiments.
Let us recall that neutral current (NC) DIS is the scattering of a highly energetic proton off a proton target

$$e^-(k) + p(p) \rightarrow e^-(k') + X \tag{1.18}$$

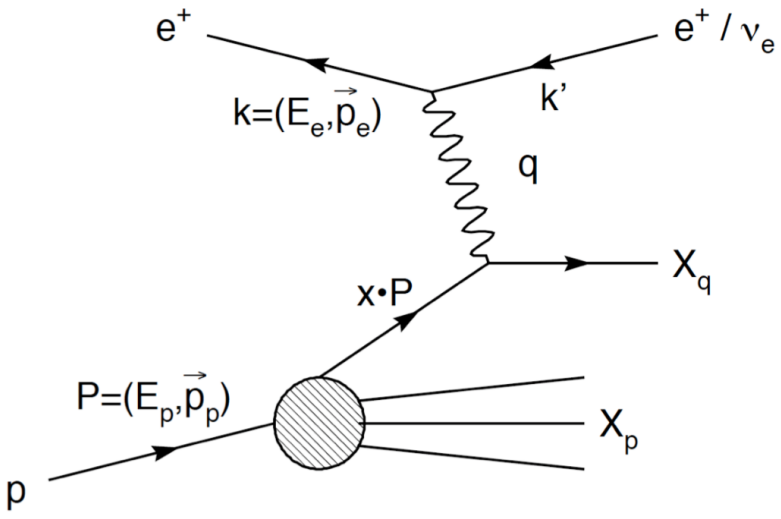


Fig. 1.3: The deep-inelastic scattering process in the QCD parton model.

where in general the proton will be destroyed by the collision (hence the name "inelastic"). The four momentum transferred between the lepton and the proton is

$$q \equiv k' - k. \tag{1.19}$$

The kinematics of the DIS process are completely specified by the following variables

$$x_{B_j} \equiv \frac{Q^2}{2p \cdot q}, \quad Q^2 \equiv -q^2, \quad y \equiv \frac{q \cdot p}{k \cdot p}. \tag{1.20}$$

Here, x_{B_j} is usually known as the *Bjorken* variable. In order to differentiate the *inelastic scattering* from the *elastic* one, the condition must be that $Q^2 \gg M_p^2$, else the proton would not be destroyed. For instance, the center-of-mass energy of the $p - \gamma^*$ collision will be

$$W^2 \equiv (p+q)^2 = M_p^2 + Q^2 \frac{1-x}{x} \simeq Q^2 \frac{1-x}{x}, \tag{1.21}$$

where M_p is the proton mass and it can typically be neglected in the above calculation.
Focusing on the neutral-current (NC) DIS process, the cross section for unpolarized electron-proton scattering can be written as in the following:

$$\frac{d^2\sigma^{NC}_{e^\pm p}}{dxdQ^2} = \frac{2\pi\alpha}{xQ^4}[Y_+\tilde{F}_2 - y^2\tilde{F}_L \mp Y_- x\tilde{F}_3] \tag{1.22}$$

where $Y_\pm = 1 \pm (1-y)^2$ is the so-called *inelasticity* parameter, $y = Q^2/(sx)$ (with x representing the Bjorken scaling variable and s the centre-of-mass energy) corresponds to the relative energy loss of the lepton in the proton rest-frame and $\tilde{F}_2$, $\tilde{F}_3$ and $\tilde{F}_L$ are the generalized structure functions.
In the DIS process, the lepton probes a parton carrying a fraction x of the proton's momentum, with $0 \leq x \leq 1$. The probability to find a given quark in the proton

with a fraction x of the total proton momentum can be expressed in terms of parton distribution functions (PDF), $f(x, Q^2)$. The PDFs are not predicted by the parton model and they must be determined experimentally. The main source of information about the proton structure come from the electromagnetic structure function $F_2(x, Q^2)$:

$$F_2(x, Q^2) = \frac{Q^4 x}{2\pi\alpha_{EW}^2(1 + (1 - y)^2)} \frac{d^2\sigma^{DIS}}{dxdQ^2} \tag{1.23}$$

where σ^{DIS} is the DIS lepton-proton cross-section. In the parton model, the structure function has a relatively simple expression as a convolution between the $\gamma^* q \to X$ partonic cross-section and the PDFs,

$$F_2(x, Q^2) = \sum_{q\bar{q}} \int_x^1 \frac{dz}{z} f_q(z) \hat{\sigma}_{\gamma^* q \to X}(\frac{x}{z}), \tag{1.24}$$

which importantly does not depend on the scale Q^2. At LO, Eq. 1.24 can be written as

$$F_2(x, Q^2) = \sum_{\text{flavours}} e_i^2 \left[q_i(x, Q^2) + \bar{q}_i(x, Q^2) \right], \tag{1.25}$$

with $q_i, \bar{q}_i$ respectively the quark and antiquark distributions in the proton.

1.2.2 Factorisation theorem and total cross section in hadron collisions

In order to calculate the cross section of a QCD process, the collinear factorisation theorem [29] can be used. The main idea is to separate the non-perturbative long-distance interactions, such as PDFs, hadronisation, and multiple soft interactions, and the perturbatively calculable short-distance interactions. This part, calculable with perturbative QCD, is given by the partonic cross section $\hat{\sigma}$ including QCD radiation. The long-distance contributions cannot be computed precisely in pQCD and are described by phenomenological models [30]. In the low energy (or long distances) regime, one encounters infrared (IR) divergences in the limit $Q \to 0$, as well as ultraviolet (UV) divergences in the limit $Q \to \infty$. The IR divergences are absorbed in the definition of the PDFs which thus acquire a dependence on the so-called *factorisation scale*, μ_F. Similarly, the UV divergences are absorbed by the renormalisation procedure in the strong coupling constant definition, which acquire a dependence on the *renormalisation scale*, μ_R. Both these two quantities are not physical so observable quantities should not depend on μ_R and μ_F. The requirement leads to the so-called *renormalisation group equations*

$$\mu_{F,R} \frac{d\sigma}{d\mu_{F,R}} = 0. \tag{1.26}$$

The cross section for quark annihilation into a virtual photon or Z can be written as [30]

$$\sigma_{pp} = \sum_{q,\bar{q}} \int dx_1 dx_2 f_q(x_1, \mu_F^2) f_{\bar{q}}(x_2 \mu_F^2) \hat{\sigma}_{q\bar{q} \to \ell^+\ell^-} \tag{1.27}$$

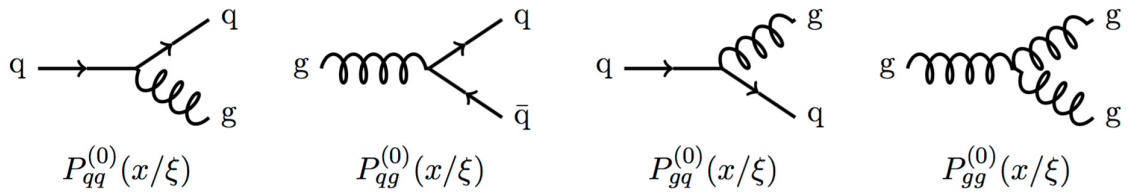


Fig. 1.4: Splitting functions at LO.

where $\hat{\sigma}_{q\bar{q}\to\ell^+\ell^-}$ represents the hadronic cross section for $q\bar{q} \to \ell^+\ell^-$ and the sum runs over all the quarks and anti-quarks combinations. The cross section of the hard process can be written as a perturbative expansion of the strong coupling constant α_S

$$\hat{\sigma}_{q\bar{q}\to\ell^+\ell^-} = \left[\hat{\sigma}_{LO}(\mu_R, \mu_F) + \alpha_S(\mu_R^2\hat{\sigma}_{NLO}(\mu_F, \mu_R)) + ...\right]_{q\bar{q}\to\ell^+\ell^-} \tag{1.28}$$

Typically, these calculations retain only a finite number of terms in the perturbative expansion, so they display a residual dependence on μ_F and μ_R. Therefore, these scales have to be chosen, in order to perform perturbative calculations. To avoid large logarithms in the perturbation series, factorisation scales are usually equalised to the same order of magnitude as the momentum scale of the hard scattering Q^2.

Given their scale dependence, PDFs can be propagated from a defined scale Q_0^2 to an higher scale $Q^2 > Q_0^2$ using the Dokshittzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) parton evolution equation [31–33]

$$Q^2\frac{\partial f_i(x,Q^2)}{\partial Q^2} = \frac{\alpha_S(Q^2)}{2\pi}\int_x^1 \frac{d\xi}{\xi}P_{ij}\left(\frac{x}{\xi}, \alpha_S(Q^2)\right) f_j(\xi, Q^2). \tag{1.29}$$

The LO DGLAP splitting functions $Pij^{(0)}$ represent the probability of a parton j, with momentum fraction ξ, emitting a parton and becoming parton i with a momentum fraction x.

The LO splitting functions shown in Fig. 1.4 are defined as follows:

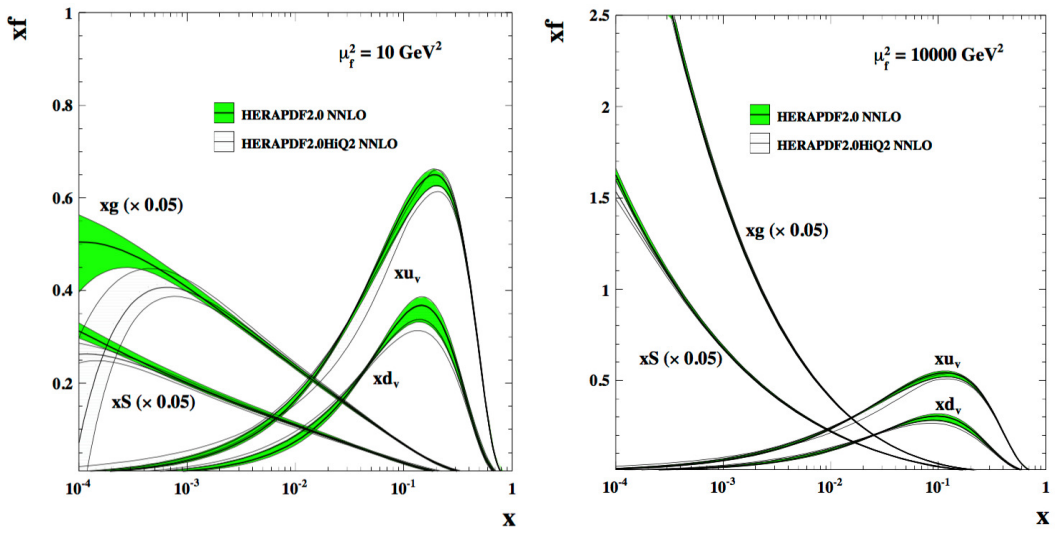
$$P_{qq}^0(z) = \frac{4}{3}\frac{1+z^2}{1-z}, \quad P_{qg}^0(z) = \frac{1}{2}(z^2 + (1-z)^2), \tag{1.30}$$

$$P_{gq}^0(z) = \frac{4}{3}\frac{1+(1-z)^2}{z}, \quad P_{gg}^0(z) = 2N_c\left[\frac{1}{z} + \frac{1}{1-z} - 2 + z(1-z)\right]. \tag{1.31}$$

In the above-described equations, z is defined as x/ξ and $N_c = 3$ represents the number of colour states.

So, while the Q^2 dependence of the PDFs is described by the DGLAP equations, the full x dependency of the PDFs cannot be predicted. Thus this dependency has to be extracted from global QCD fits to several measurements. In general the QCD fits are obtained by parametrising the parton distribution functions at a starting scale Q_0^2 using an analytical form. The generic functional form used for parametrising the quark and gluon distributions is

$$F(x, Q_0^2) = A \cdot x^B(1-x)^C P(x; D, E, ...), \tag{1.32}$$



(a) (b)

Fig. 1.5: PDFs for quarks, anti-quarks and gluons for energy scale Q^2 = 10 GeV^2 (a) and Q^2 = 10^4 GeV^2 (b). The distributions are determined from the H1 and ZEUS Collaborations from global fits to data [34].

where the parameters B and C are physically motivated: the former is associated to the behaviour in the small-x region, while the latter is associated to the large-x behaviour. However, they are not sufficient enough to describe either quark or gluon distributions; that is why the term $P(x; D, E, ...)$ is introduced, in order to add more flexibility in the functional form. Additional constraints are also included, such as the *momentum sum rule*

$$\sum_{i=q,\bar{q},g} \int_0^1 dx x f_i(x) = 1, \tag{1.33}$$

which translates the fact that the total momentum of the proton is distributed among all the quarks and gluons, and the *valence sum rules*

$$\int_0^1 dx (f_u(x) - f_{\bar{u}}(x)) = 2, \tag{1.34}$$

$$\int_0^1 dx (f_d(x) - f_{\bar{d}}(x)) = 1, \tag{1.35}$$

$$\int_0^1 dx (f_s(x) - f_{\bar{s}}(x)) = 0. \tag{1.36}$$

which ensure that the proton has the correct flavour quantum numbers. Once the PDFs are parametrised at Q_0^2, they are evolved in Q^2 and convoluted with the partonic cross section to predict a cross section which can be compared to the actual measurements. For the measured and calculated cross sections a χ^2 function is calculated. The starting parameters are now deduced by minimizing the χ^2. Once these parameters are determined, the PDFs can, starting from the

parametrization scale, be evolved to any Q^2 using the DGLAP equations. Fig. 1.5 shows the PDFs of the partons at two different energy scales Q^2 = 10 GeV^2 (left) and Q^2 = 10^4 GeV^2 (right), calculated from fits to H1 and ZEUS data.
The uncertainties of the extracted PDFs correspond to the experimental uncertainties of the measurements used for the global fit. These uncertainties can be propagated to uncertainties on the deduced PDF parameters. Given the fact that some of these parameters can be highly correlated, the propagation of these uncertainties on PDFs cannot be done straight forward.
There are two main approaches for estimating the uncertainties associated with the fit, namely the *Monte Carlo Replicas method* and the *Hessian method*.

The Monte Carlo Replicas method is a common method used for calculating PDF uncertainties; for example, it is largely used by the NNPDF group. Here, for each of the *n* experimental measurements, *d* Monte Carlo (MC) replicas are created such that any sufficiently large number of them has the same mean, variance and covariance as the experimental measurement [35]. Then, a PDF can be fitted for each of the MC replicas, with the best fit being the average of the obtained fits; the uncertainty associated with this fit is taken as the variance between the fits. This method can be used with any parametrisation of the PDFs and and it is very handy when fitting parametrisation characterized by a large number of free parameters. These above-described replicas comprise systematic uncertainties as well as statistical errors.

The Hessian Method In this method [36–38] the $n \times n$ covariance matrix (with *n* parameters) is build for the up and down variation of the parameters by either 68% or 90% confidence level (CL). This matrix can be rotated into an orthogonal eigenvector basis. The result are *2n* eigenvector sets, one for the up variation and the other for the down variation, which allow the uncorrelated propagation of the fit uncertainties. These eigenvector sets for up (X_i^+) and down (X_i^-) variation can be then combined to an asymmetric uncertainty ΔX_i^+ and ΔX_i^- on the PDF or on an observable using the PDF with the following formula:

$$\Delta X_i^+ = \sqrt{\sum_{i=1}^{2n} [\max(X_i^+ - X_0, X_i^- - X_0, 0)]^2} \tag{1.37}$$

$$\Delta X_i^- = \sqrt{\sum_{i=1}^{2n} [\max(X_0 - X_i^+, X_0 - X_i^-, 0)]^2}, \tag{1.38}$$

where X_0 represents the central value and X_i^+ and X_i^- the up and down variation of source *i*, respectively. ΔX_i^+ adds in quadrature the PDF uncertainty contributions which lead to an increase of a given observable *X*, while ΔX_i^- the PDF uncertainty contributions which lead to a decrease. Large asymmetries in the PDF uncertainties are indicative of a problem with the fit.

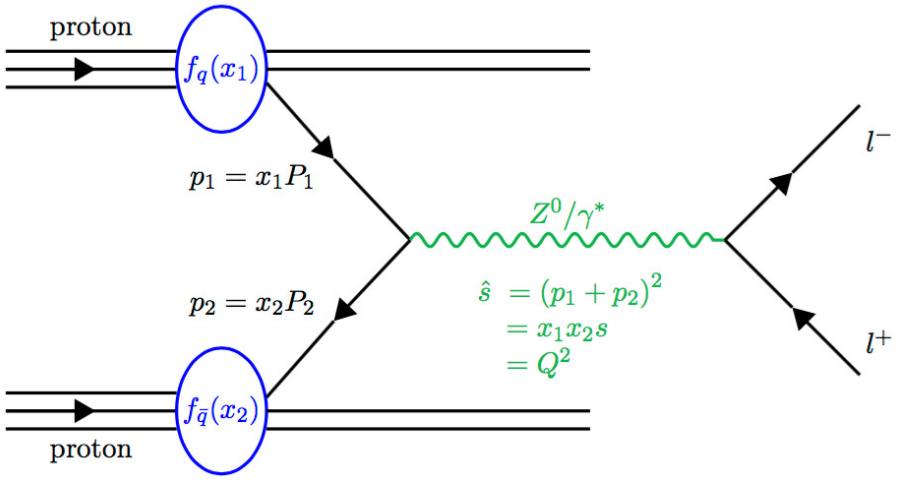


Fig. 1.6: Feynman diagram of the Drell-Yan process. The four-momenta of the protons are defined as $P_{1,2}$, giving the four-momenta of the partons initiating the hard scattering with $p_{1,2}$, depending on the momentum fractions $x_{1,2}$.

1.3 The Drell-Yan process

The neutral-current (NC) Drell-Yan process is the dominant mechanism of Z boson production at the LHC. It was first described by Sidney Drell and Tung-Mow Yan [39] in 1970 and it consists in the production of a lepton pair $\ell^+\ell^-$ at a hadron collider by quark-antiquark annihilation with invariant mass $M^2 = (p_{\ell^+} + p_{\ell^-})^2 \gg 1$ GeV^2. Fig. 1.6 illustrates the leading order process as a Feynman diagram.
Let us take a closer look at the kinematics of this process. The kinematics of vector-boson production is hadronic collisions can be written as follows:

$$p1 = (x_1\sqrt{2}/2, 0, 0, x_1\sqrt{2}/2), \tag{1.39}$$

$$p2 = (x_2\sqrt{2}/2, 0, 0, -x_2\sqrt{2}/2), \tag{1.40}$$

$$q = ((x_1 + x_2)\sqrt{2}/2, 0, 0, (x_1 - x_2)\sqrt{2}/2), \tag{1.41}$$

with $p_{1,2}$ the four-momenta of the incoming partons and $q = p_1 + p_2$ the four-momentum of the gauge boson. In this process, the kinematics is fixed once the gauge boson rapidity is specified. Indeed, the rapidity y of the Drell-Yan lepton pair is defined as

$$y = \frac{1}{2}\ln\left(\frac{E + p_x}{E - p_z}\right) = \frac{1}{2}\ln\left(\frac{x_1}{x_2}\right), \tag{1.42}$$

where p_z is the component of momentum along the beam axis and E represents the total energy of the particle. As it is shown in Eq. 2.2, the rapidity can be written in terms of momentum fractions x_1 and x_2. The kinematic relation of the rapidity and the momentum faction x yields

$$x_1 = \frac{\sqrt{Q^2}}{\sqrt{s}}e^{+y}, \quad x_2 = \frac{\sqrt{Q^2}}{\sqrt{s}}e^{-y}, \tag{1.43}$$

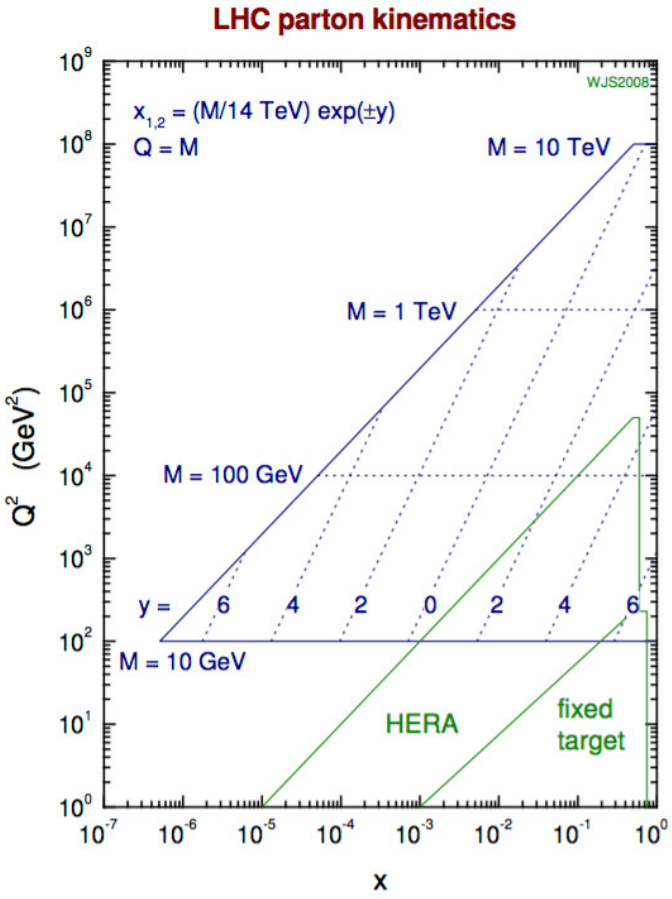


Fig. 1.7: Kinematic relation between partonic quantities: momentum fraction x, hard scale Q^2 and the final-state quantities: invariant mass M and rapidity y, for center-of-mass energy equal to 14 TeV. Figure taken from Ref. [30].

with $M = Q$, and it is illustrated for different invariant masses at $\sqrt{s}$ = 14 TeV in Fig. 1.7. Also shown are the regions of phase space accessible to each experiment. Among the DIS experiments HERA has access to lower x and higher Q^2, whereas fixed target experiments probe lower Q^2 and larger x. A broad range in both variable, x and Q^2, is covered by the LHC. So, the accessible rapidity range for the production of a Z boson is determined by the available center-of-mass energy and the Z mass value:

$$|y_Z^{MAX}| = \ln\left(\frac{\sqrt{s}\ \text{TeV}}{M_Z}\right) \tag{1.44}$$

Thus, for a center-of-mass energy of $\sqrt{s}$ = 8TeV, the Z boson can be produced in the rapidity range $|y_Z| \leq 4.96$, considering M_Z = 91.1876 GeV (value taken Ref. [13]).

1.3.1 Leading Order Cross Section

The cross section for this process at leading order can easily be obtained from the fundamental QED $e^+e \to \ell^+\ell$ cross section, with the addition of appropriate

color and charge factors

$$\hat{\sigma}(q\bar{q} \to \gamma^* \to \ell^+\ell^-) = \frac{4\pi\alpha_{EW}^2}{3\hat{s}}\frac{1}{N_C}Q_q^2 \tag{1.45}$$

where $\hat{s} = M_{\ell\ell}^2$ is the squared center-of-mass energy of the incoming partons, $1/N_C = 1/3$ is a color factor, taking into account that only three colour combinations are possible given the fact that the intermediate state has to be colorless and Q_q is the charge of the quarks.
Eq. 1.45 shows a strongly falling behaviour, $\hat{\sigma} \propto 1/M_{\ell\ell}^2$. If $M_{\ell\ell} \approx 90$ GeV, the process take place via the exchange of a Z boson (namely, $q\bar{q} \to Z \to \ell^+\ell^-$), leading to a Breit-Wigner resonance in the spectrum of the invariant mass. Once $M_{\ell^+\ell^-}$ is comparable to M_Z, the two possible processes, the exchange of a virtual photon and the exchange via a Z boson., interfere. Thus, the differential cross-section distribution is considered. The sub-process cross section yields

$$\frac{d\hat{\sigma}}{dM_{\ell\ell}^2} = \frac{4\pi\alpha_{EW}^2}{3M_{\ell\ell}^2}\frac{1}{N_c}Q_q^2\delta(\hat{\sigma} - M_{\ell\ell}^2). \tag{1.46}$$

The differential cross section of the Drell-Yan process is now considered. In the parton model [17] this is given by a convolution of this sub-process cross-section and the parton densities as follows:

$$\begin{aligned}
\left.\frac{d\sigma}{dM_{\ell\ell}^2}\right|_{LO} &= \int_0^1 dx_1 dx_2 \sum_{q=1}^{N_f}\{f_q(x_1)f_{\bar{q}}(x_2) + f_{\bar{q}}(x_1)f_q(x_2)\}\frac{d\hat{\sigma}}{dM_{\ell\ell}^2}(q\bar{q} \to \ell^+\ell^-) \\
&= \frac{4\pi\alpha_{EW}^2}{3M_{\ell\ell}^2}\frac{1}{N_c}\int_0^1 dx_1 dx_2 \delta(x_1x_2s - M_{\ell\ell}^2)\left[\sum_{q=1}^{N_f} Q_q^2\{f_q(x_1)f_{\bar{q}}(x_2) + f_{\bar{q}}(x_1)f_q(x_2)\}\right] \\
&= \frac{4\pi\alpha_{EW}^2}{3M_{\ell\ell}^2}\frac{1}{N_c}\int_0^1 dx_1 dx_2 \delta(x_1x_2s - M_{\ell\ell}^2)P_{q\bar{q}}(x_1, x_2).
\end{aligned} \tag{1.47}$$

The sum runs over the quark flavours N_f and the dependence on the PDFs is defined in the $q\bar{q}$ probability function $P_{q\bar{q}}(x_1, x_2)$ [40]. Multiplying Eq. 1.47 by $M_{\ell\ell}^4$, the lepton pair cross section can be written as a function of a dimensionless scaling variable $\tau = x_1x_2 = \dfrac{M_{\ell\ell}^2}{s}$

$$\begin{aligned}
M_{\ell\ell}^4\frac{d\sigma}{dM_{\ell\ell}^2} &= \frac{4\pi\alpha_{EW}^2}{3N_c}\tau\int_0^1 dx_1 dx_2 \delta(x_1x_2 - \tau)P_{q\bar{q}}(x_1, x_2) \\
&= \frac{4\pi\alpha_{EW}^2}{3N_c}\tau\mathcal{F}(\tau).
\end{aligned} \tag{1.48}$$

1.3.2 Perturbative QCD corrections to the Drell-Yan process

The source of higher order QCD corrections to the Drell-Yan section are additional Feynman diagrams with additional α_s vertices. These additional terms

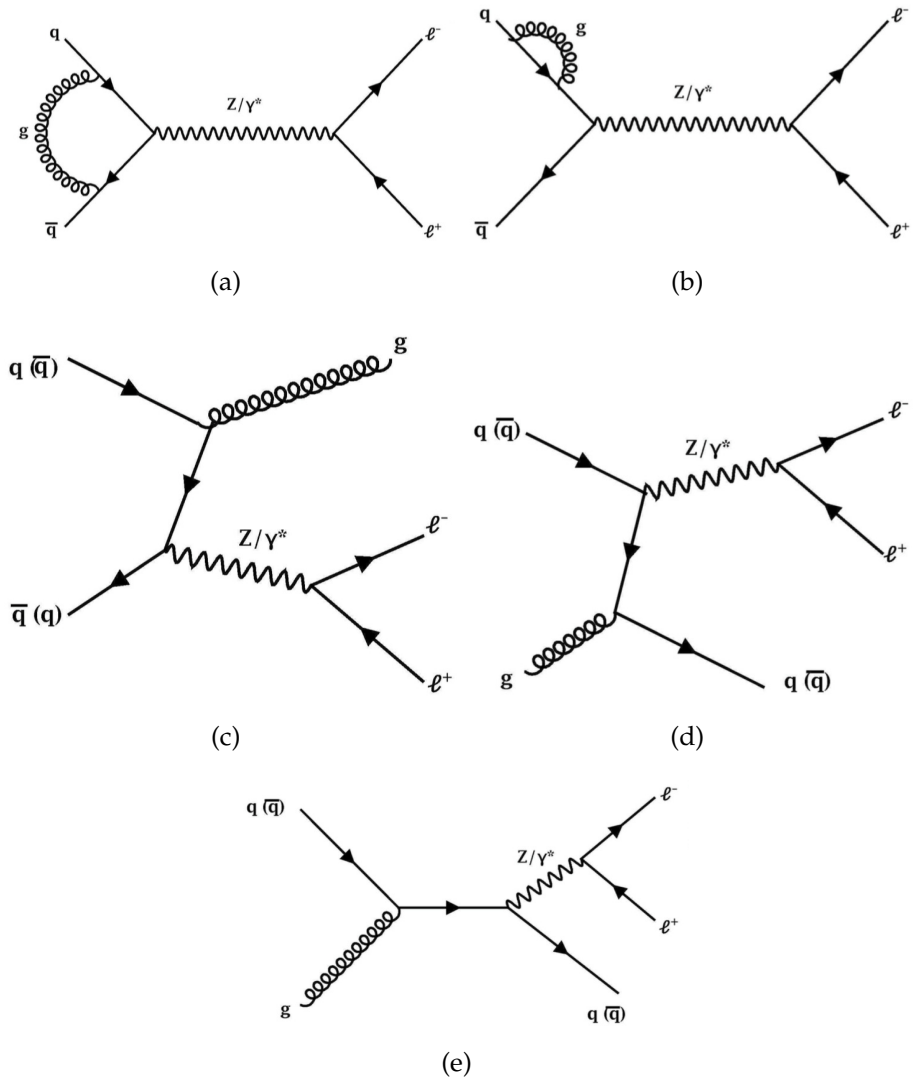


Fig. 1.8: NLO contributions to the Drell-Yan process.

are represented by several Feynman diagrams and constitute real or virtual corrections (see Fig. 1.8). Starting from Eq. 1.48, the LO partonic cross section can be written as

$$M_{\ell\ell}^4 \frac{d\hat{\sigma}}{dM_{\ell\ell}^2}\bigg|_{LO} = \frac{4\pi\alpha_{EW}^2}{3N_c}\tau\hat{\mathcal{F}}(\tau), \tag{1.49}$$

where

$$\hat{\mathcal{F}}(\tau) = Q_q^2\delta(1-\tau). \tag{1.50}$$

In pQCD, the $\hat{\mathcal{F}}(\tau)$ can be expanded in powers of α_s

$$\hat{\mathcal{F}}(\tau) = \hat{\mathcal{F}}_0(\tau) + \frac{\alpha_s}{2\pi}\hat{\mathcal{F}}_1(\tau) + \left(\frac{\alpha_s}{2\pi}\right)^2\hat{\mathcal{F}}_2(\tau) + ... \tag{1.51}$$

The corrections at NLO ($\mathcal{O}(\alpha_s)$) to Drell-Yan process are illustrated in Fig. 1.8. Fig. 1.8 (a) and (b) show the virtual corrections, while the real gluon emission producing a gluon in the final state is shown in Fig. 1.8 (c); finally Fig. 1.8 (d) and

(e) show the $q(\bar{q}) - g$ scattering process which produces an additional parton in the final state along with the Drell-Yan lepton pair.
The real and loop corrections in the cross section calculation introduce divergences of different types, which can be divided as follows:

- *Infrared (IS)* divergences from real and virtual contributions due to soft-gluon emissions;
- *Ultraviolet (UV)* divergences from loop diagram, when the energy of the emitted parton tends to infinity;
- *Collinear* divergences,which are induced by parton splitting of the initial-state partons.

The IR singularities can be regulated by combining the contributions from real and and virtual diagrams in the calculation, while the UV divergences are treated with renormalisation. Several types of regularisation scheme could be used, such as the massive gluon scheme and the dimensional regularisation [40]. In the massive gluon scheme, the real and virtual corrections are regularised by introducing a mass for the gluon. As regards the dimensional regularisation, the integration variables in space-time dimensions are changed from 4 to 4 - ϵ and a renormalisation scale μ_R is introduced. The NLO Drell-Yan cross section measurement as a function of the invariant mass of the di-lepton pair ($Q^2 = M_{\ell\ell}$) at 13 TeV can be found in Fig. 1.9, showing different contributions of the γ^*, Z and interference terms. NNLO ($\mathcal{O}(\alpha_s)$) corrections can also be computed and they include diagrams with loops which produce $q\bar{q}$ pairs [41–43]. Furthermore, theoretical calculations of the Drell-Yan cross-section can be performed at NLO matched to leading-log (LL) and next-to-leading-log (NLL) calculations, to take into account soft and collinear partonic emissions [44–47].

1.3.3 The photon-induced contribution

So far we have only discussed the production of di-lepton pairs $\ell^+\ell^-$ via the Drell-Yan process. The quarks in the proton carry electric charges themselves and hence they can radiate photons, not only gluons. This means that there is also a photonic structure besides the partonic structure of the proton. These photons can also produce lepton pairs, via the so-called *photon-induced* (PI) process. Fig. 1.10 shows the LO Feynman diagrams which contribute to lepton pair production. The left diagram corresponds to the *u*-channel diagram, while the right diagram to the *t*-channel diagram. The PI process does not have a *s*-channel diagram at LO and therefore it has different kinematic properties than the Drell-Yan process: indeed, the photon induced process has a higher contribution at small angles with respect to the direction of the incoming photons.
The photon contribution can be accounted for by introducing a photon PDF and evolving it in Q^2 with modified DGLAP equations, which take into account higher order electro-weak corrections. Many of the standard PDF sets in use at the LHC do not include this photon contribution. The first PDF set which included this photon-induced contribution was MRST2004QED [48]. The modern PDFs which also take

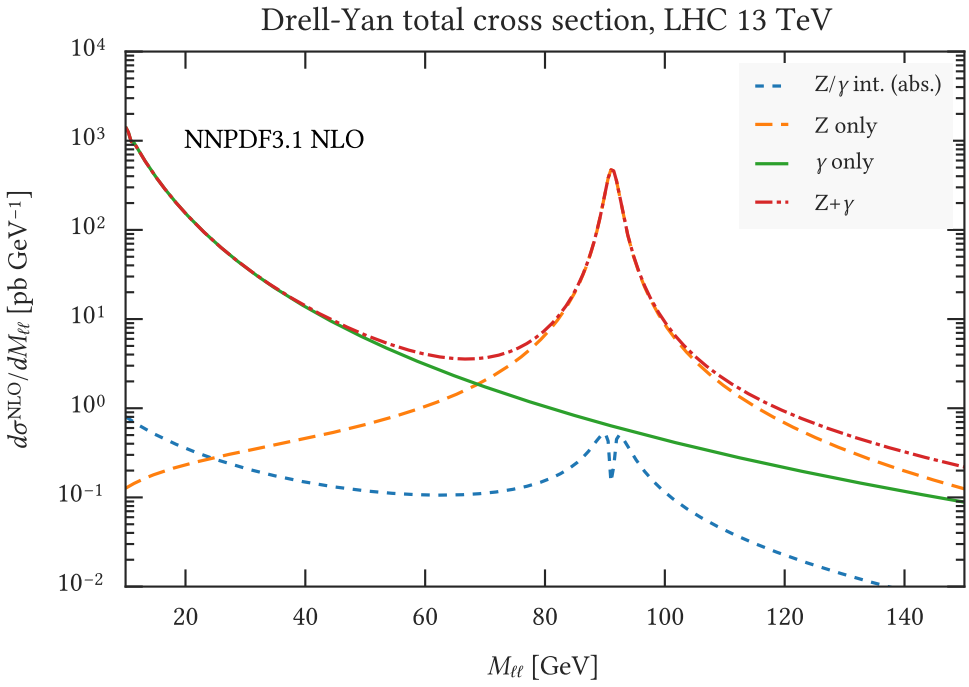


Fig. 1.9: The total NLO DY cross section $d\sigma/dM_{\ell\ell}$ at 13 TeV: the contributions arising from γ^* exchange, Z exchange and the γ^*-Z interference are shown separately.

into accunt this part are: NNPDF3.0 QED [49,50], CT14 QED [51] and two recent theoretical calculations, LUXqed [52,53] and HKR16 [54] (see Chapter 4 for a more detailed explanation). As it will be discussed later, the $\gamma\gamma$-initiated contribution becomes a significant part of the di-lepton production at high invariant masses. The knowledge of this process is therefore an important input for analyses searching for a heavy resonance decaying into lepton pairs.

1.4 Precision tests and limitations of the SM

The SM has been successfully tested in a large variety of experiments over a wide range of energies during the last 40 years. Precision tests of the electroweak interaction performed at LEP or Tevatron accelerators confirmed that the couplings of quark and leptons to the Z and $W^{\pm}$ bosons agree with the the ones predicted by the SM; in addition to that fact, other experimental results, e.g. measurements of weak hadrons decays, provide additional tests of the SM at low energies. Moreover, the QCD sector was first established by the observation of scaling violations in DIS experiments and then many precision DIS measurements followed. The recent discovery, at the Large Hadron Collider (LHC), of a Higgs boson with a mass equal to ≈ 125 GeV [55,56] is the latest success of the SM, for two reasons:

- The measured mass is in agreement with the allowed range from the combined measurement of electroweak observables [57];

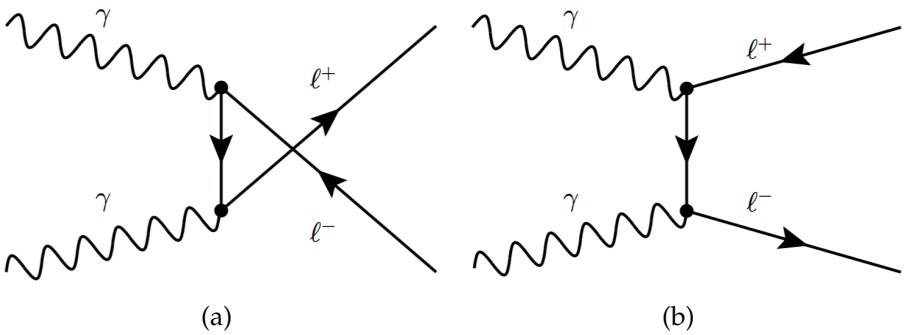


Fig. 1.10: LO photon-induced Feynman diagrams which contributes to di-lepton production. (a) u-channel diagram; (b) t-channel diagram

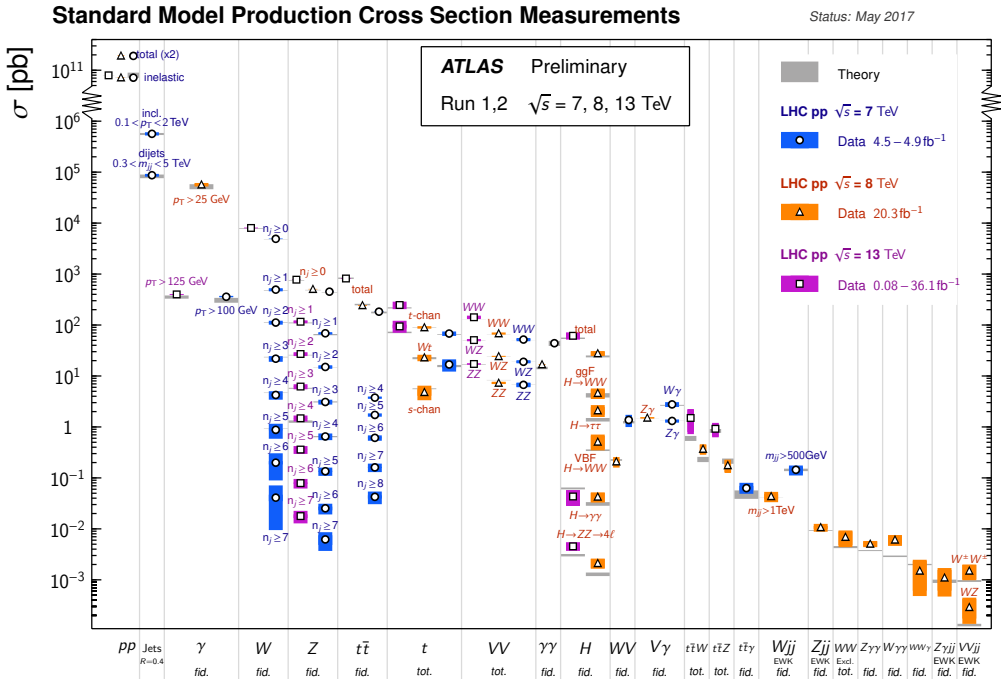


Fig. 1.11: Summary of several Standard Model total and fiducial production cross section measurements, corrected for leptonic branching fractions, compared to the corresponding theoretical expectations. All theoretical expectations were calculated at NLO or higher. The luminosity used for each measurement is indicated close to the data point. Uncertainties for the theoretical predictions are quoted from the original ATLAS papers. Taken from `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/CombinedSummaryPlots/SM/`.

- The spin and coupling strength of the new boson are also in good agreement with the SM predictions for the measured mass.

Disagreement between the SM predictions and experimental data is found for only very few observables; the most significant discrepancies are observed for the anomalous magnetic moment of the muon a_μ and for the forward-backward asymmetry A_{fb} in b-quark production at LHCb. A summary of several SM mea-

surements performed by the ATLAS Collaboration at different center-of-mass energy is shown in Fig. 1.11.
In spite of this success, the SM is widely believed to be an effective theory valid only for energies up to the electroweak scale. In addition to the fact that the SM does not include the gravitational force, this theory does not explain the pattern of fermion masses and does not allow for neutrino masses; thus these and other deficiencies indicate the need for new physics beyond the Standard Model (BSM). Some of the most relevant deficiencies are listed in the following:

1.4.1 Dark Matter

Astrophysical data indicates the existence of non-baryonic, weakly interacting and neutral matter of unknown composition, so-called *Dark Matter*. Supersymmetry, an extension of the Standard Model that postulates the existence of an additional symmetry between fermions and bosons, increases the number of particles and naturally proposes possible dark matter candidates.

1.4.2 Grand Unification Problem

Embedding the SM in a more fundamental framework is theoretically very appealing. The unification of the symmetry structure of the SM into a larger symmetry group leads to the formulation of *Grand Unification Theories* (GUT). Below some energy scale the generic theory could be broken into the symmetry group of the SM which is then interpreted as an effective low energy limit of the larger theory. An indication for this is given by the running couplings which may meet at a scale of $\approx 10^{15}$ GeV. The scale dependence is affected by the possible existence of supersymmetric particles. For masses of supersymmetric particles in the range of 1 - 10 TeV an unification of all three coupling constants at $\Lambda_{GUT} \approx 10^{16}$ GeV could be achieved thanks to some extended BSM theory [58].

1.4.3 Hierarchy and Fine-Tuning Problems

Radiative corrections to the bare Higgs boson mass squared (δm_h^2) diverge quadratically with the cut-off energy scale, Λ, which represents the energy scale up to which the SM is expected to be valid [59]. Putting in the GUT scale $\Lambda_{GUT} \approx 10^{16}$ GeV where new physics is expected to show up, would require a severe *fine-tuning* of order of $(v/10^{15})^2 = \mathcal{O}(10^{-26})$ which seems unnatural. Supersymmetry can provide a solution to the hierarchy problem as each fermion has a bosonic partner yielding the same correction but with opposite sign. Therefore, the cancellation of higher order correction terms would be exact and automatic (this statement is only true for an exact Supersymmetry which implies that all SM particles and their supersymmetric partners have the same mass).

1.5 New physics processes involving charged lepton pairs in the final state

In the following section, we introduce some models which lead to a signature in the final state containing a charged lepton pair ($\ell^+\ell^- = e^+e^-,\ \mu^+\mu^-$).

1.5.1 Neutral Heavy Gauge Boson

The benchmark model for Z' bosons is the Sequential Standard Model (SSM) [60]; in this model, the Z' (Z'_{SSM}) has the same quantum numbers and couplings to fermions as the SM Z boson and no couplings to the W and Z bosons. The Z' boson is hence a copy of the SM Z boson, only with a higher mass and width. A more theoretically-motivated model is a Grand Unification model in which the E_6 gauge group is broken into SU(5) and two additional U(1) groups [61]. The lighest linear combination of the corresponding two new neutral gauge bosons, Z'_ψ and Z'_χ, is considered the Z' candidate: $Z'(\theta_{e_6}) = Z'_\psi \cos(\theta_{E_6}) + Z'_\chi \cos(\theta_{E_6})$, where $-\pi \leq \theta_{e_6} < \pi$ is the mixing angle between the two gauge bosons. The pattern of spontaneous symmetry breaking and the value of θ_{E_6} determine the Z' couplings to fermions. Six different models [60,61] lead to the specific Z' states named Z'_ψ, Z'_N, Z'_η, Z'_I, Z'_S and Z'_χ, respectively. In all the above-described models, the resonances are assumed to have a narrow instrinsic width, meaning that exclusion limits extracted for the widest signal shape can also be used to infer corresponding limits on the other models considered. The expected instrinsic width of the Z'_{SSM} as a fraction of the mass is 3.0%, while for any E_6 model the intrinsic width is predicted to be between 0.5% and 1.2% [62]. The LO cross-section times the branching ratio to a single di-lepton flavour for various Z' models versus polemass at $\sqrt{s}$= 13 TeV is shown in Fig. 1.12.

In the Minimal Z' models [63], the phenomenology is controlled by only two effective coupling constants in addition to the Z' boson mass. This parameterization encompasses many models, including a left-right symmetric model [64,65] and the pure (B - L) model [66], where B (L) is the baryonic (leptonic) number, and B - L is the conserved quantum number.

Previous direct searches for a Z' have been carried out at the Tevatron experiments [67,68] and indirect constraints from LEP [69–72] have resulted in limits on the Z'_{SSM} mass of 1.07 TeV [68] and 1.79 TeV [73] respectively. Searches have also been performed by the ATLAS and CMS Collaborations [74,75]. The strongest ATLAS exclusion limits for the Z'_{SSM} come from the latest search conducted by the ATLAS Collaboration using 3.2 fb^{-1} of *pp* collision data at $\sqrt{s}$ = 13 TeV [76]. That combined analysis of the di-electron and di-muon channels set lower limits at 95% credibility level (C.L.) on the Z'_{SSM} pole mass of 3.36 TeV. The CMS exclusion limits for this search in 2015 are very similar to the one found by ATLAS.

1.5.2 Contact Interactions

The presence of a new interaction can be detected at an energy much lower than that required to produce direct evidence of the existence of a new gauge

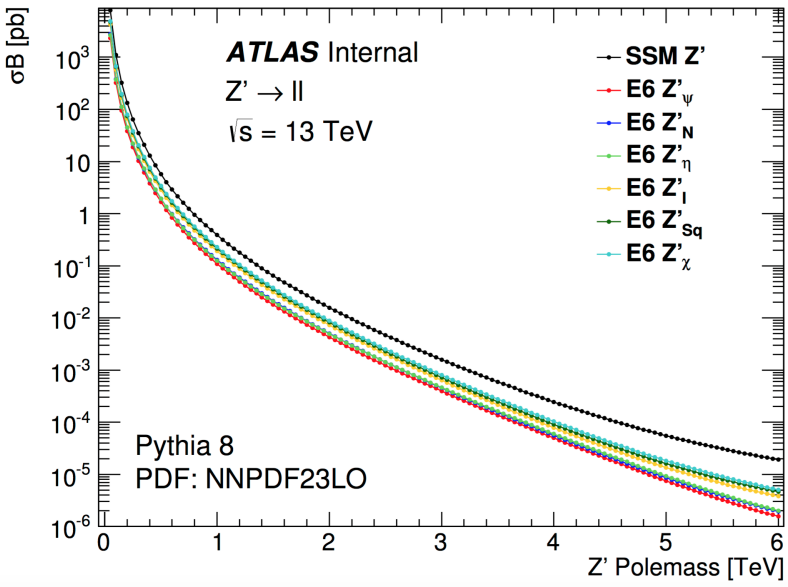


Fig. 1.12: LO cross-section times branching ratio to a single di-lepton flavour for various Z' models versus polemass at $\sqrt{s}$= 13 TeV.

boson. The charged weak interaction responsible for nuclear β decay provides such an example. A non-renormalizable description of this process was succefully formulated by Fermi in the form of a four-fermion contatc interaction [77]. A contact interaction can also accomodate deviations from the SM in proton-proton scattering due to quark and lepton compositeness, where a characteristic energy scale Λ corresponds to the binding energy between fermion constituents. A new interaction or compositeness in the process $q\bar{q} \to \ell^+\ell^-$ can be described by the following four-fermion contact interaction Lagrangian [78,79]

$$\begin{aligned}\mathcal{L} = g^2/\Lambda^2[&\eta_{LL}(\bar{q}_L\gamma_\mu q_L)(\bar{l}_L\gamma^\mu l_L)\\ &+ \eta_{RR}(\bar{q}_R\gamma_\mu q_R)(\bar{l}_R\gamma^\mu l_R)\\ &+ \eta_{LR}(\bar{q}_L\gamma_\mu q_L)(\bar{l}_R\gamma^\mu l_R)\\ &+ \eta_{RL}(\bar{q}_R\gamma_\mu q_R)(\bar{l}_L\gamma^\mu l_L)],\end{aligned} \tag{1.52}$$

where g is a coupling constant chosen by convention in order to satisfy $g^2/4\pi = 1$, Λ is the contact interaction scale, and $q_{L,R}$ and $l_{L,R}$ are left-handed and right-handed quark and lepton fields, respectively. The parameters $\eta_{i,j}$, where i and j and L or R (left or right), define the chiral structure of the new interaction. Different chiral structures could be investigated, with the left-right model obtained by setting $\eta_{LR} = \pm 1$ and $\eta_{RL} = \eta_{RR} = \eta_{LL} = 0$. Likewise, the left-left, the right-left and the right-right models are obtained by setting the corresponding parameters to ± 1, and the others to zero. The sign of η_{ij} determines whether the interference is constructive ($\eta_{ij} = -1$) or destructive ($\eta_{ij} = +1$). The cross-section for the process $q\bar{q} \to \ell^+\ell^-$ is the presence of these contact interaction models can be written as:

$$\sigma_{tot}(m_{\ell\ell}) = \sigma_{DY}(m_{\ell\ell}) - \eta_{ij}\frac{F_I}{\Lambda^2} + \frac{F_C}{\Lambda^4}, \tag{1.53}$$

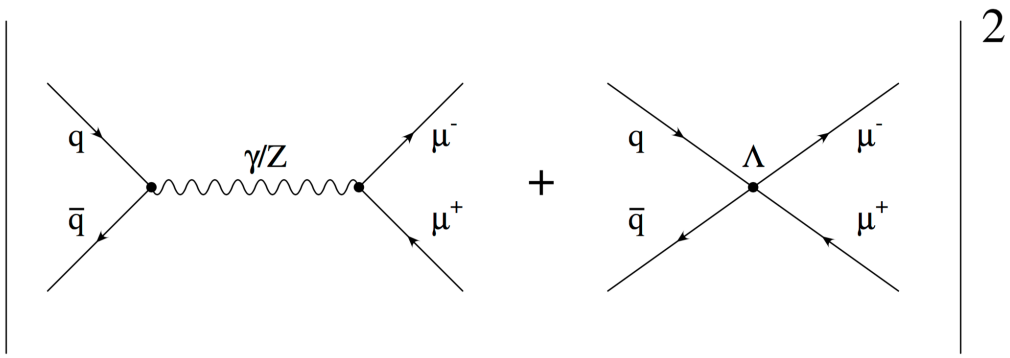


Fig. 1.13: LO production mechanism for DY with additional contact term with scale Λ in the di-muon final state. Similar diagrams apply to the electron channel.

where the first term accounts for the $q\bar{q} \rightarrow Z/\gamma^* \rightarrow \ell^+\ell^-$ SM DY process (see Fig. 1.13), the second term corresponds to the interference between the DY and CI processes, and the third term describes the pure CI process. These two latter terms include F_I and F_C, respectively, which are functions of the differential cross-section with respect to $m_{\ell\ell}$, and do not depend on Λ. The relative impact of the interference and pure CI terms depends on both the di-lepton invariant mass and Λ. Previous searches for CI have been carried out in the neutrino-nucleus and electron-electron scattering [80], as well as electron-positron [72, 81], electron-proton [82], and proton-proton colliders [83, 84]. Searches for CI have also been performed by the ATLAS and CMS Collaborations [75, 85]. The strongest exclusion limits for $\ell\ell qq$ CI in which all quarks flavours contribute come from the latest ATLAS non-resonant dilepton analysis conducted using data recorded during 2015+2016 *pp* collision at $\sqrt{s}$ = 13 TeV [76, 86]. These two searches are part of my thesis work and their results will be described later in Chapter 5.

Chapter 2

The Large Hadron Collider and the ATLAS experiment

The Large Hadron Collider (LHC) is the largest energy particle accelerator ever built. It is a project developed by the European Organization for Nuclear Research (CERN). The main objective of this collider is the investigation of the nature of electroweak symmetry breaking, to measure the properties of the Higgs boson and the search for particles predicted by new physics models beyond Standard Model (BSM) at the TeV scale.

2.1 General remarks

The LHC s a proton-proton (and heavy-ions) collider machine based at CERN near Geneva. This machine uses the 26.7 km tunnel, located underground (between 50 m and 175 m depth), that was built between 1984 and 1989 for the Large Electron-Proton Collider (LEP) machine. In 2000, LEP was decommissioned to give way to the LHC. During the years of ATLAS data taking it has worked at increasing energies, starting from the center of mass energy of 7 TeV in 2011 to 13 TeV from 2015 till 2018.[1]
The LHC, pictured in Fig. 2.1, is designed to deliver proton and ion beam collisions to four main experiments: the two multipurpose detectors ATLAS (A Toroidal Lhc ApparatuS) [87] and CMS (Compact Muon Solenoid) [88], LHCb (LHC beauty experiment) [89] focusing on b-physics and ALICE (A Lhc Ion Collider Experiment) [90] focusing on heavy-ions physics.
Before being injected into the LHC, the particles are accelerated step by step up to the injection energy of 450 GeV, by a series of accelerators shown in Fig. 2.1. For protons, the first system is a linear accelerator (LINAC2), which accelerates the beam at an energy of 50 MeV. The protons then go through the Proton Synchrotron Booster (PSB) and are brought to 1.4 GeV. After that they are injected into the Proton Synchrotron (PS), where they are accelerated to 26 GeV. Finally, the Super

[1] In this thesis, the quoted performances and parameters refer to 2015 and 2016, years when the data analysed in this thesis were collected.

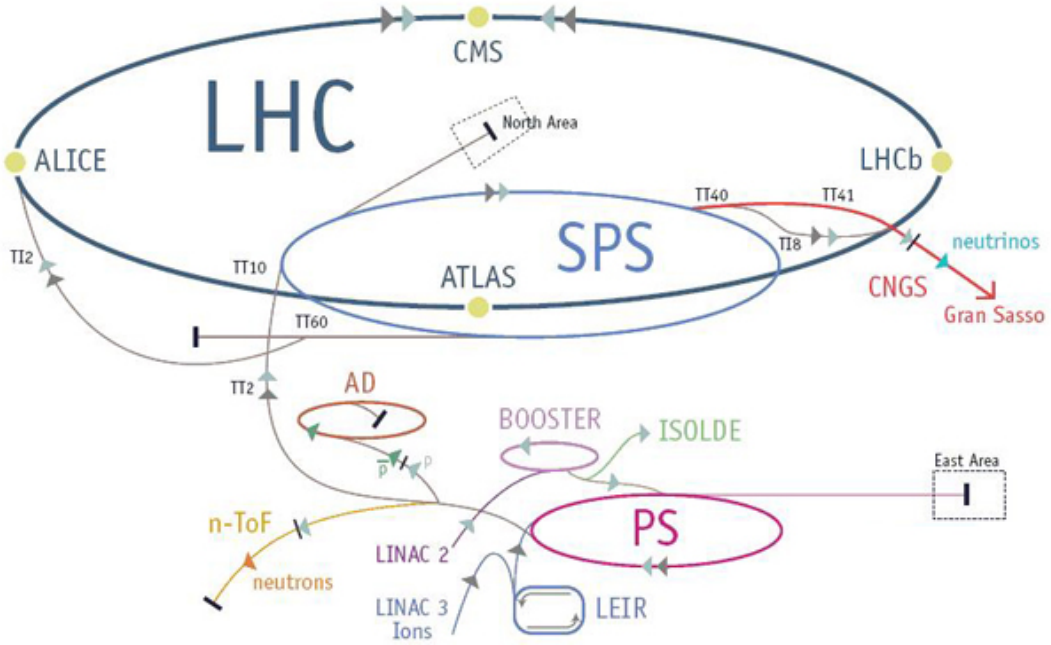


Fig. 2.1: Schematic layout of the LHC.

Proton Synchrotron (SPS) is used to further increase their energy to 450 GeV. From the SPS, two transfer lines inject the proton beams into the LHC.

2.2 Luminosity

A measurement of intensity of the beam is *luminosity*, which depends only on parameters of the beam:

$$\mathscr{L} = \frac{N_b^2 n_b f_r \gamma_r}{4\pi\epsilon_n\beta^*} \cdot \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)\right)^{-1/2} \tag{2.1}$$

where (nominal parameters for the LHC are given in parenthesis):

- N_b is the numbers of particle per bunch ($\approx 1.67 \cdot 10^{11}$),
- n_b is the number of bunch per beam (2808),
- f_r is the revolution frequency (11245 Hz),
- γ_r is the relativistic gamma factor ($\approx$ 7000),
- ϵ_n is the transverse normalized beam emittance (3.75 μm),
- β^* is the beam squeezability at the ATLAS collision point (0.55 m),
- θ_c is the crossing angle between the beams ($\pm$ 142.5 μrad),
- σ_z is the Root Mean Square (RMS) of the bunch length (7.55 cm),

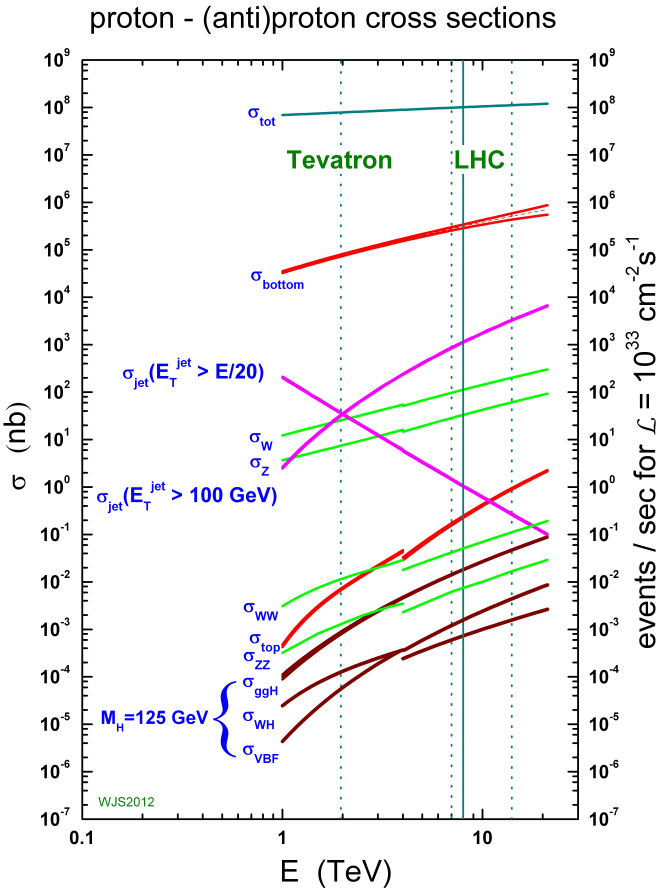


Fig. 2.2: Cross section of SM processes at the Tevatron and LHC colliders. Plot taken from Ref. [91].

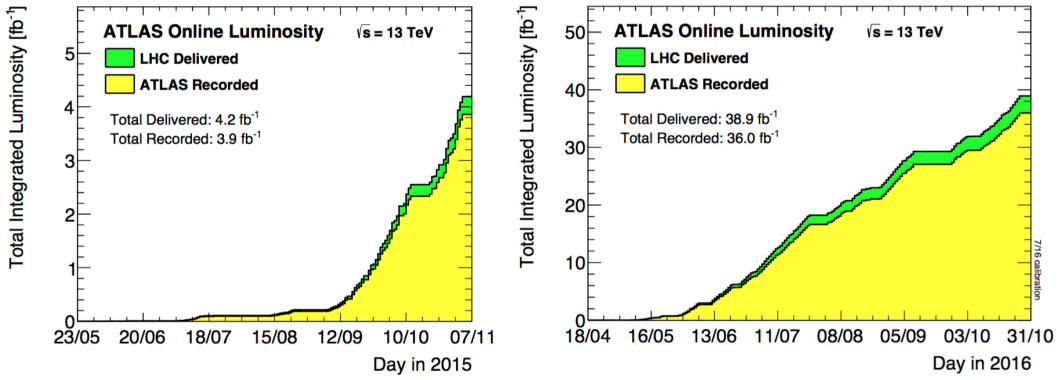
- σ^* is the RMS of the beam size at the interaction point.

All these parameters have been optimized to ensure maximal luminosity. Table 2.1 highlights the typical values of representative beam parameters under operational conditions of the LHC in *pp* collision mode in Run 1 from 2010 to 2012 and in Run 2 from 2015 to 2016. In 2016, the peak luminosity is up to 1.4 $\times$ cm^{-2} s^{-1} which has surpassed the design value.

The numbers of events generated for a specific process can be written as $N_{event} = \sigma_{event} \cdot L \cdot A \cdot \epsilon$, where σ_{event} represents the cross section of the considered process, $L = \int \mathscr{L} dt$ is the luminosity, integrated over time, provided by the machine, A is the geometrical acceptance and ϵ the efficiency of the detector. Since the intent of the LHC is to explore physics which has not been seen so far, the cross section of interesting processes will be very small and thus the luminosity must be maximal. In Fig. 2.2 the cross section for different processes as a function of the center of mass energy is presented.

Parameter	2010	2011	2012	2015	2016
Beam Energy [TeV]	3.5	3.5	4	6.5	6.5
Bunches/beam n_b	348	1331	1380	2244	2220
Bunch crossing time $t_{crossing}$ [ns]	150	150	50	25	25
Protons/bunch N_b [10^11 protons]	0.9	1.2	1.7	1.1	1.1
ϵ_n [μm]	2.6	2.4	2.4	3.5	3.4
β^* [m]	2.0-3.5	1.0-1.5	0.6	0.8	0.4
Peak Luminosity [10^{34} cm^{-2} s^{-1}]	0.02	0.36	0.77	0.51	1.4
<μ>	2	9	21	14	25
Integrated luminosity LHC delivered [fb^{-1}]	0.047	5.5	22.8	4.2	38.9

Table 2.1: Selected proton running conditions in the LHC operation in Run 1 (2010- 2012) and in Run 2 (2013-2016) and the corresponding design parameters.

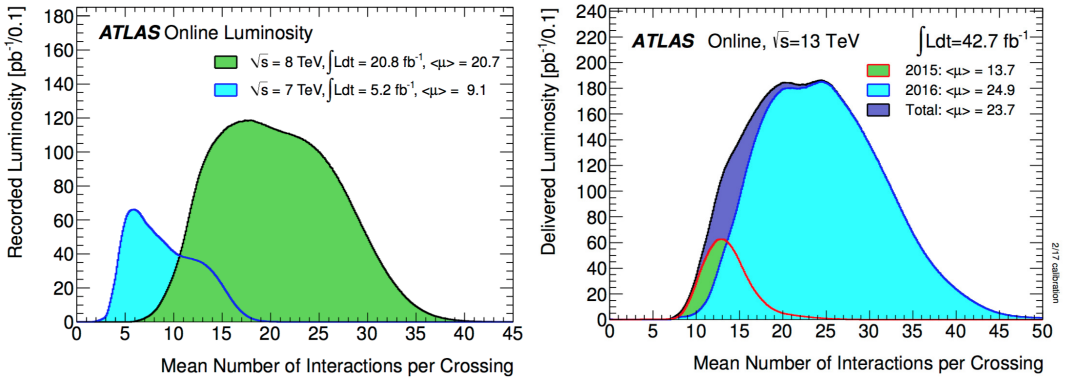


(a) (b)

Fig. 2.3: Luminosities delivered by the LHC (green) and recorded by ATLAS (yellow) at $\sqrt{s} = 13$ TeV in 2015 (left plot) and 2016 (right plot).

2.3 LHC parameters and operation

The LHC design value for instantaneous luminosity in proton-proton collisions at centre-of-mass energy $\sqrt{s} = 14$ TeV is $\mathscr{L} = 10^{34}$ cm^{-2} s^{-1}. With this configuration there are an average of about 23 inelastic scatterings per 25 ns bunch crossing, in which nearly 1000 new particles are produced.
The first beam circulated in the LHC on September 10^{th}, 2008. Nine days later an incident was caused by a faulty electrical connection between two magnets during powering tests of the main dipole circuit. Helium leakage into the tunnel and serious mechanical damage delayed the operations by about a year. In order to prevent this from happening again, the center of mass energy has been reduced to 7 TeV for the data taking periods of 2010 and 2011, while in 2012 the machine was running at $\sqrt{s} = 8$ TeV. During the LHC Run 1 in 2011 and 2012 and Run 2 in 2015 and 2016, the operation of ATLAS detector ensures a high data taking efficiency of about 93%. Fig 2.3 shows the integrated luminosity delivered by the LHC and recorded by the ATLAS detector versus time in 2015 and 2016 with the initial online calibration. On the one hand, the elevated number of particles for



(a) (b)

Fig. 2.4: Luminosity weighted distributions of the mean number of interactions per bunch crossing for (a) 2011-2012 and (b) 2015-2016 data taking, respectively.

bunch increases the number of events of interest but, on the other hand, it increases also the number of *pile-up* events; pile-up is a challenge for the detectors and for the acquisition and analysis of the data. In particle physics, pile-up refers to the situation several collisions are recorded in the same 25 ns bunch crossing. The pile-up events can be generated both by the superimposition of particles interactions coming from previous or following bunch crossing (*out-of-time* pile-up) or by different interactions in the considered bunch crossing (*in-time* pile-up). Because of the pile-up, the mean number of interactions for bunch crossing increases, and thus the detector must be able to separate particles from multiple simultaneous interactions. The distribution of the mean number of interaction for bunch crossing are shown in Fig. 2.4 for both 2011-2012 (left plot) and 2015-2016 (right plot) data taking [92, 93]. In the end, Fig. 2.5 shows the peak instantaneous luminosity delivered to ATLAS during stable beams for pp collisions at 13 TeV centre-of-mass energy for each LHC fill as a function of time in 2015 (left-hand side plot), as well as in 2016 (right-hand side plot).

2.4 The ATLAS detector: a general overview

ATLAS is a general-purpose experiment, designed to cover the widest possible range of physics at the LHC. The formal proposal for ATLAS was introduced in 1994 and 10 years later the detector installation in the cavern began. Around 3000 scientists from 177 institutes in 38 countries work on the ATLAS experiment. In this chapter, the most important features of the ATLAS detector and of its trigger system will be presented; a brief description of ATLAS software will follow. The ATLAS detector is one of the four experiments working at the LHC at CERN; it has been designed to look for the Higgs bosons and, in general terms, to investigate the physics at the TeV energy scale. The ATLAS layout of the whole experiment is similar to other general purpose high energy collider detector, with cylindrical shape (to approach a 4π coverage) and layers of sub-detectors. A cut-away view of the ATLAS detector, which has a length of 44 m, a diameter of 25 m and a

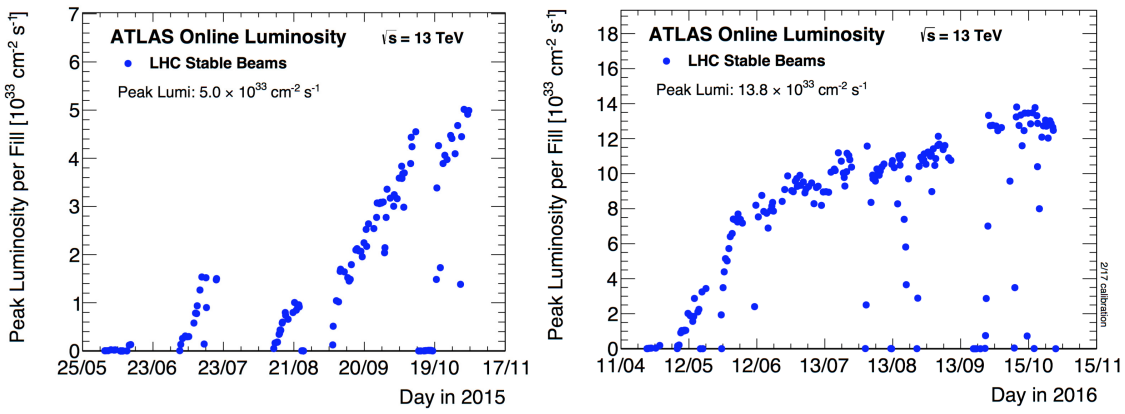


Fig. 2.5: Peak instantaneous luminosity delivered to ATLAS during stable beams for pp collisions at 13 TeV centre-of-mass energy for each LHC fill as a function of time in (a) 2015 and (b) 2016.

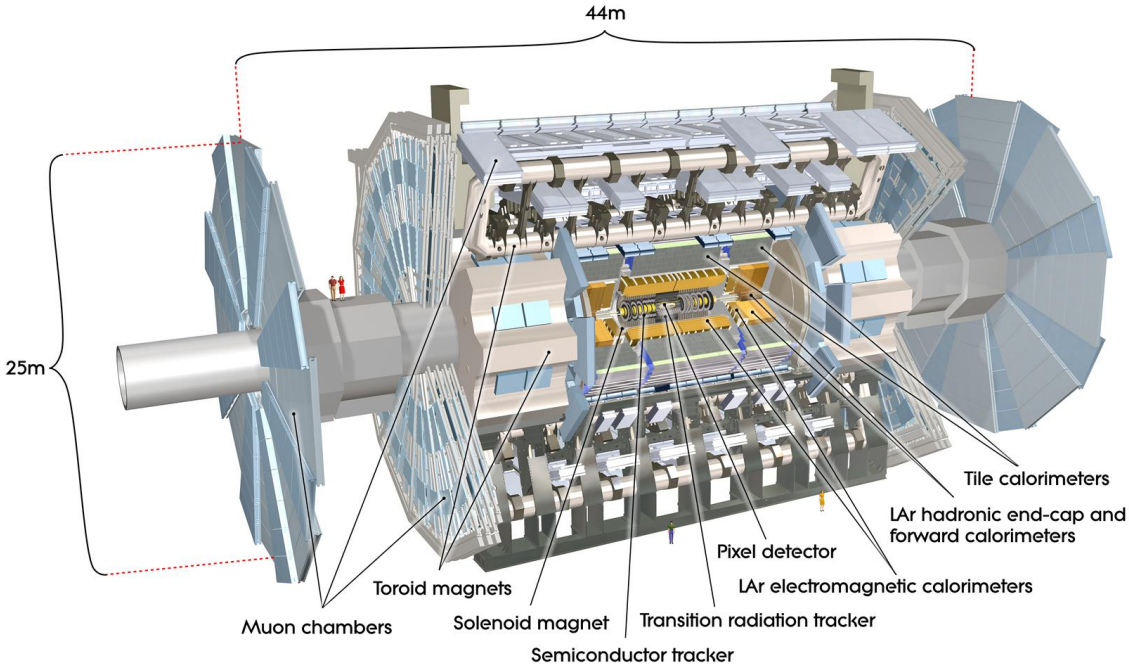


Fig. 2.6: Layout of the ATLAS experiment.

weight of, approximately, 7000 tonnes, is visible in Fig. 2.6. Driven by the physics requirements, the detecting technologies are a precision tracking system (for measuring the momentum of charged particles), calorimeters (for the determination of the energy of the electromagnetic and strongly interacting particles) and muon chambers (for measuring the momentum of muons). Therefore, ATLAS consists of three main subsystems: the tracking system, the electromagnetic and hadronic calorimeters and the muon chambers; the central part around the beam line is called the *barrel* and the wheels perpendicular to the beam axis in the forward

parts are called *end-caps*. In the following, the description of the coordinate system and of the magnets system is presented; later, the different sub-detectors will be described.

2.4.1 Coordinate System

The coordinate system used in the ATLAS detector [94] is a right-handed coordinate system and it is shown in Fig. 2.7. As it is possible to see, the beam

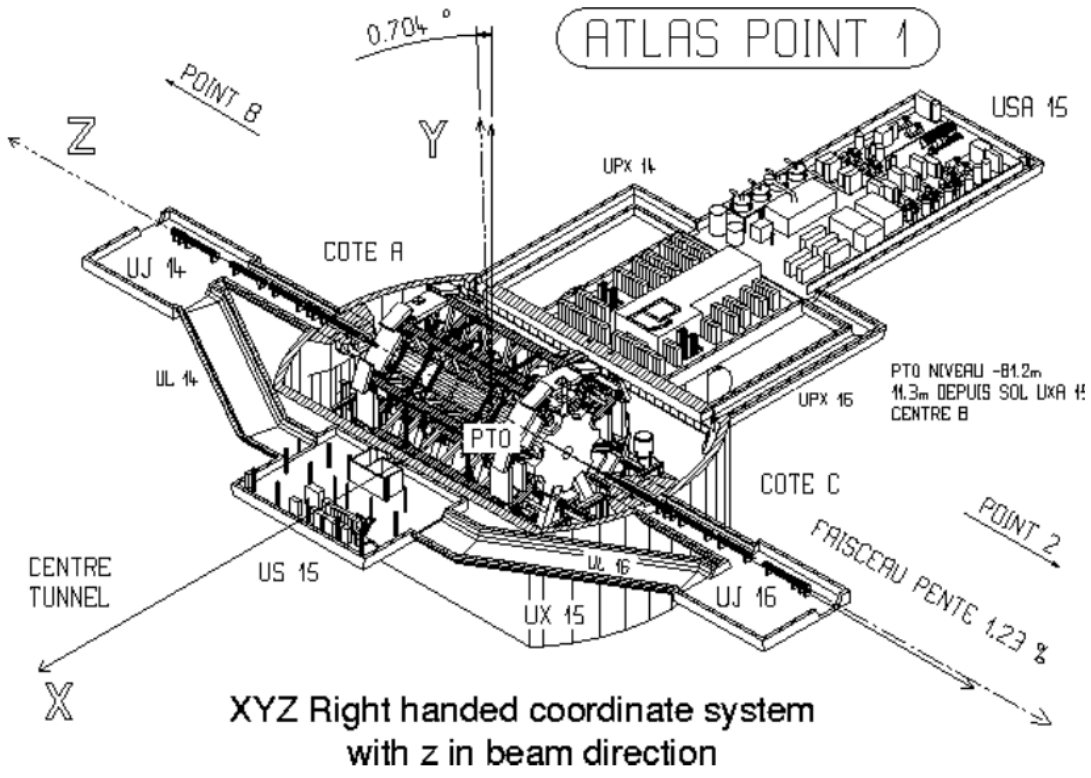


Fig. 2.7: The coordinate system in the ATLAS detector.

direction defines the *z*-axis and the *x*-*y* plane is transverse to the beam direction. The positive *x*-axis is defined pointing to the centre of the LHC ring and the positive *y*-axis is defined pointing upwards. The transverse energy E_T and momentum p_T are defined in the *x*-*y* plane. As in all other experiments at a hadron collider, in ATLAS cylindrical coordinates are used instead of *xyz* coordinates: in cylindrical coordinates the azimuthal angle ϕ is used and, instead of the polar angle θ, the rapidity, defined in Eq. 2.2, for a massive particle, or pseudo-rapidity, defined in Eq. 2.3, in case of the mass of the particle can be neglected.

$$y = \frac{1}{2}\log\left(\frac{E + P_z}{E - P_z}\right) \tag{2.2}$$

$$\eta = -\log\left(\tan\frac{\theta}{2}\right) \tag{2.3}$$

This choice is driven by the fact that, at a hadron collider, the total momentum of the initial system along the z-axis is unknown (since the collision takes place between the partons, not at the proton level); for this reason, a quantity, i.e. the difference of pseudorapidity of two particles, which does not depend on Lorentz boosts along the beam axis is used. In the end, the distance in pseudorapidity-azimuthal angle space is, in general, indicated using the ΔR coordinate, defined as follows:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{2.4}$$

Tracks of charged particles in ATLAS are parameterized with these five parameters:

- The transverse impact parameter d_0, which is the distance to the beam axis in the x-y plane.
- The longitudinal impact parameter z_0, which is the distance to the coordinate system origin in z-direction.
- The azimuthal angle ϕ, measured in the x-y plane.
- The polar angle θ.
- The charge of the particle divided by its momentum, q/p, which characterizes the track curvature.

An illustration of the above-illustrated track parameters is shown in Fig. 2.8

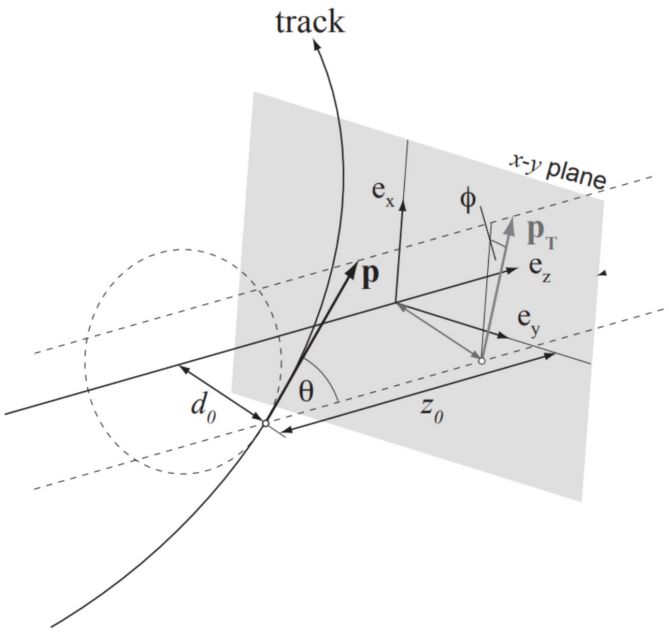


Fig. 2.8: Track parametrization in the ATLAS coordinate system, where the base vectors e_x, e_y and e_z represent the x-, y- and z-axis respectively.

2.4.2 Magnet System

In order to reconstruct the transverse momenta of all the charged particles coming from the interaction point, a magnetic field is fundamental. The ATLAS Magnet System is made up by superconducting magnets of three different kinds [95] :

Central Solenoid

A solenoidal magnet is placed around the Inner Detector (ID) and it has been designed to provide a magnetic field of ≈ 2 Tesla (T) along the beam axis inside the ID [96]. Since it is placed before the electromagnetic (EM) calorimeter, the material budget has to be low in order to not distort measurement in the calorimeter.

Barrel Toroid

The ATLAS Barrel Toroid systems consists of eight coils assembled radially and symmetrically around the beam axis; the peak field provided by the Barrel Toroid coils is 3.9 T, providing 2 to 6 Tm of bending power in the pseudorapidity range from 0 to 1.3 (in other words, the central region of the detector) [97].

Endcap Toroids

A toroidal field, provided by two air-core toroids, characterizes the exterior region of the muon spectrometer [98]. It provides $B \sim 1.0$ T inside the toroid muon spectrometer.

A schematic view of the Magnetic System of the ATLAS Experiment is shown in Fig. 2.9. The magnetic field provided by the ATLAS Magnetic System is not completely

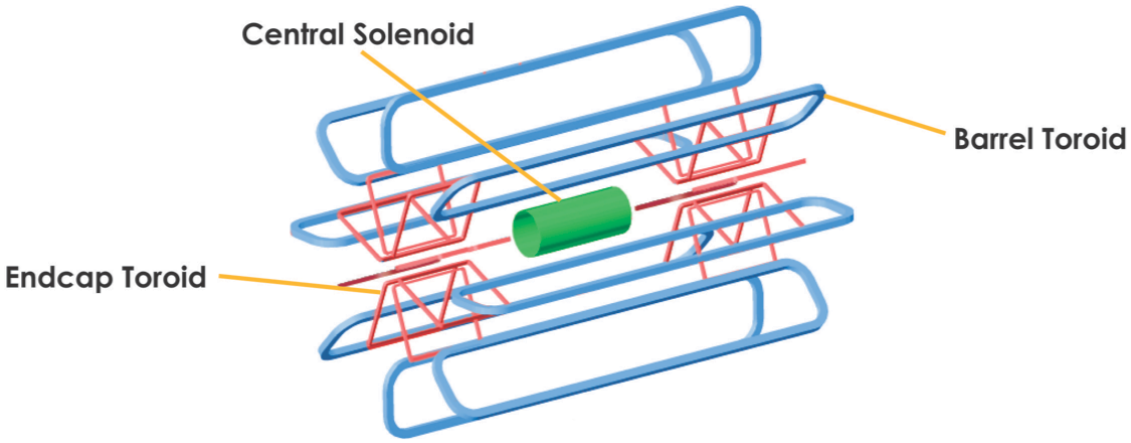


Fig. 2.9: The Magnetic System in the ATLAS detector.

uniform, due to the complexity of the toroids; moreover, the superposition of the magnetic fields of the barrel and endcap toroids originates the so-called "transition

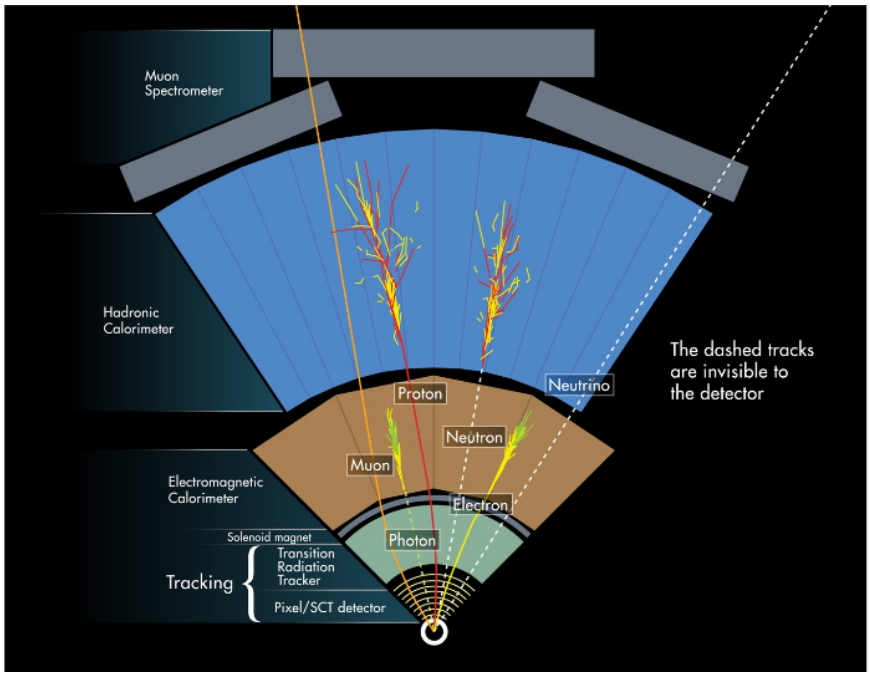


Fig. 2.10: The identification of various particles with the use of the onion structure of the ATLAS detector.

region" (more precisely $1.3 < |\eta| < 1.6$), where the bending power is so much poorer.

2.4.3 Identifying particles with the ATLAS detector

The ATLAS detector has an onion-like structure; the purpose of this design is to be able to discriminate between different type of particles that cross the detector. An illustration of how the various particles are identified in the ATLAS detector is given in Fig. 2.10. Charged particles, curved by the magnetic field, leave a track in the inner detector, indicated by full lines. Photons can convert into electron pairs while crossing the inner detector therefore leaving only signs in some layers of the ID. Electrons and photons are fully stopped in the electromagnetic calorimeter, whereas hadrons, such as protons and neutrons, continue and deposit their energy also in the hadronic calorimeter. Muons, being minimum ionizing particles (MIP), leave a small amount of their energy in the detector, and make it through to the muon spectrometer.In the volume of the muon spectrometer, they are curved by a toroidal magnetic field and are measured before they escape the detector. Dashed lines illustrate particles that do not interact with that part of the detector and are effectively invisible. Some particles escape the detector leaving no trace behind; for example neutrinos can only be inferred by conservation of energy and momentum and show up as missing energy in some part of the detector.

2.4.4 Inner Detector

The Inner Detector (ID) [99,100] is the ATLAS tracker which provides an efficient detection of charged particles with high spatial and momentum resolutions and is capable of identifying primary and secondary vertices; it is the innermost detector of ATLAS and the closest one to the interaction point. It also provides electron identification within η < 2.0 and over a wide range of energies (0.5 GeV - 150 GeV). The ID is contained within a cylindrical envelope of length $\pm$3512 mm and of radius 1150 mm, and located at small radii from the beamline. Its high granularity allows to perform the pattern recognition,vertex and momentum measurements in the pseudo-rapidity range $|\eta| < 2.5$. These capabilities are achieved with a combination of three different sub-detectors: the Silicon Pixel Detectors, the SemiConductor Tracker (SCT) and the Transition Radiation Tracker (TRT) (as it is shown in Fig. 2.12 and in Fig. 2.13). The main feature of the ID is its hardness to the radiation, because it is subject to large doses of highly energetic particles. Passive material in the ID must be minimized in order to avoid inducing particle shower, causing the objects to loose energy before the calorimeter. The resolution of the tracking parameter is given by:

$$\sigma\left(\frac{1}{p_T}\right) = a \oplus \frac{b}{p_T} \tag{2.5}$$

where $\oplus$ indicates sum in quadrature. So one obtains:

$$\frac{\sigma\left(p_T\right)}{p_T} = a \cdot p_T \oplus b \tag{2.6}$$

where the constant *a* is the intrinsic resolution of the detector, and *b* arises from multiple scattering, dominating the resolution at low momentum.
When a particle interacts with one of sub-detector layers it deposits a part of its energy to the sensor (e.g. pixel, straw tube, strips), and this energy is being converted by the sensor readout electronics. If the signal is larger than a predefined threshold a hit is recorded. One wants to have a large number of hits in order to precisely measure the particles track. However, too much of detector material can lead to multiple scattering and conversion of photons. The material budget of the ID is shown in Fig. 2.11. All three sub-detectors provide a very good precision measurement of momenta, such as:

$$\frac{\sigma\left(p_T\right)}{p_T} = 0.005\% \cdot p_T \oplus 1\% \tag{2.7}$$

The parameters of the ID are summarized in Table 2.2.

Silicon Pixel Detector

The Pixel detector is the component closest to the beam pipe and it therefore has to cope with a higher particle flux than any other detector in ATLAS. This requires a high granularity to disentangle tracks from individual charged particles and to identify primary and secondary vertices. It is 1.4 m long and has a diameter

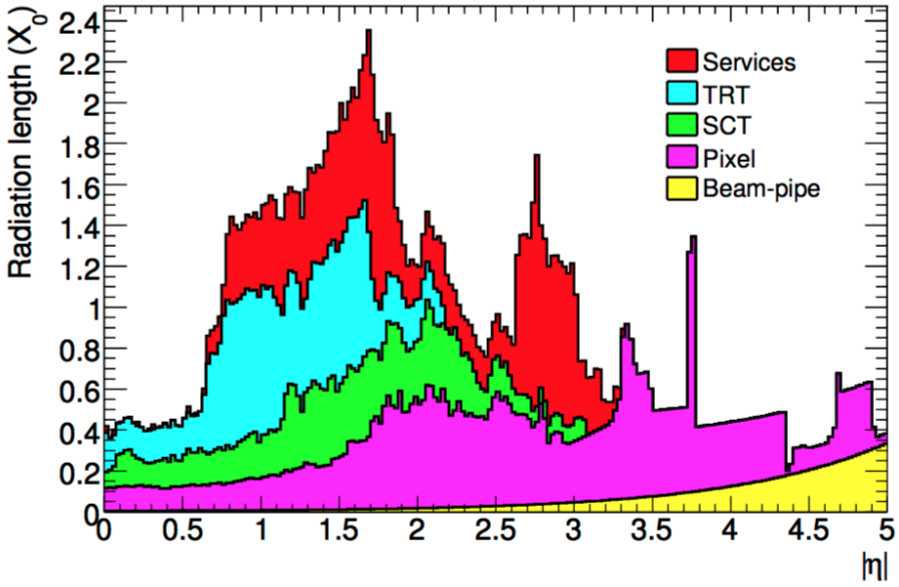


Fig. 2.11: The material budget of the ATLAS Inner Detector as a function of absolute pseudo-rapidity in units of radiation length X_0.

Type	Position	Area [m^2]	Resolution [$\sigma(\mu m)$]	Channels (10^6)	η coverage
Pixels	IBL	0.2	$R\phi$ = 12, z = 66	16	$\pm$ 2.5
	2 barrel layers	1.4	$R\phi$ = 12, z = 66	81	$\pm$ 1.7
	5 end-cap disks	0.7	$R\phi$ = 12, z = 77	43	1.7 - 2.5
SCT	4 barrel layers	34.4	$R\phi$ = 16, z = 580	3.2	$\pm$ 1.4
	9 end-cap wheels	26.7	$R\phi$ = 16, z = 580	3.0	1.4-2.5
TRT	Axial barrel layers		70 (per straw)	0.1	$\pm$ 0.7
	Radial end-cap layers		70 (per straw)	0.32	0.7-2.0

Table 2.2: Parameters of the ID. The resolutions quoted are typical values.

of 0.43 m. The active part of the pixel detector consists of a module composed of silicon sensors, front-end electronics and flex-hybrids with control circuits. There are 1744 modules resulting in a total active area of silicon of approximately 1.7 m^2 and 80.4 million readout channels (this is about 50% of all the readout channels in the entire detector). Each layer consists of several staves containing 13 modules each.The area of each module is $\approx 2 \cdot 6$ cm^2, the thickness is ≈ 250 μm and they contain thousands of pixels with a size of $\approx 50 \cdot 400$ μm^2 [101]. During the first long shutdown of the LHC in 2013-2014, a new innermost pixel detector, the Insertable B-layer (IBL) [102], was added between the first pixel layer and a new smaller radius beam pipe at the radius of 33 mm, to help improve the vertex resolution and efficiency of tagging hadrons composed by a *b* quark (*b-tagging efficiency*), to compensate for inefficiencies in the pixel B-layer which can arise over time due to irreversible radiation damage, as well as to meet the increasing bandwidth requirements resulting from the expected Phase-I LHC peak luminosity. Fig. 2.14 shows the ATLAS IBL detector prior to the insertion and an IBL stave. The spatial resolution of the IBL measured from collision data is 10.0 $\pm$ 0.1 μm in

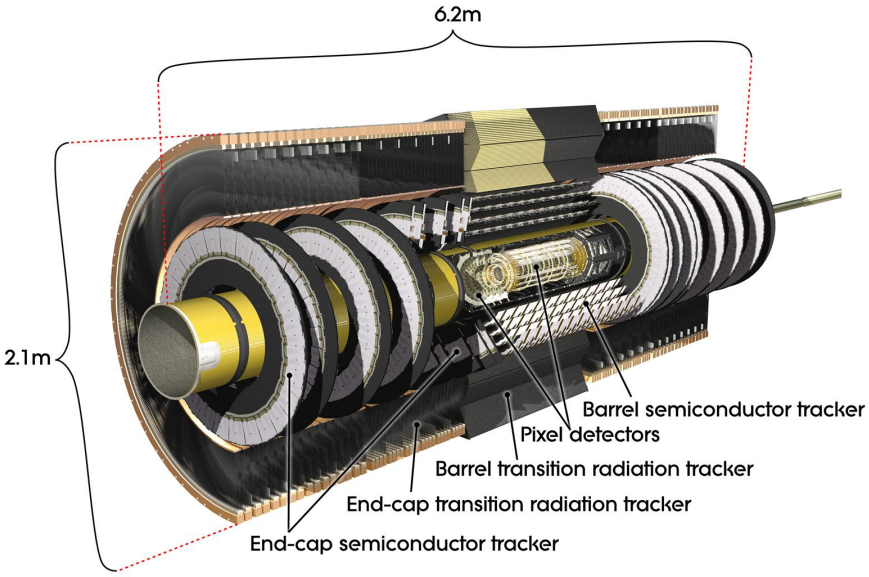


Fig. 2.12: A cut-away view of the ATLAS Inner Detector.

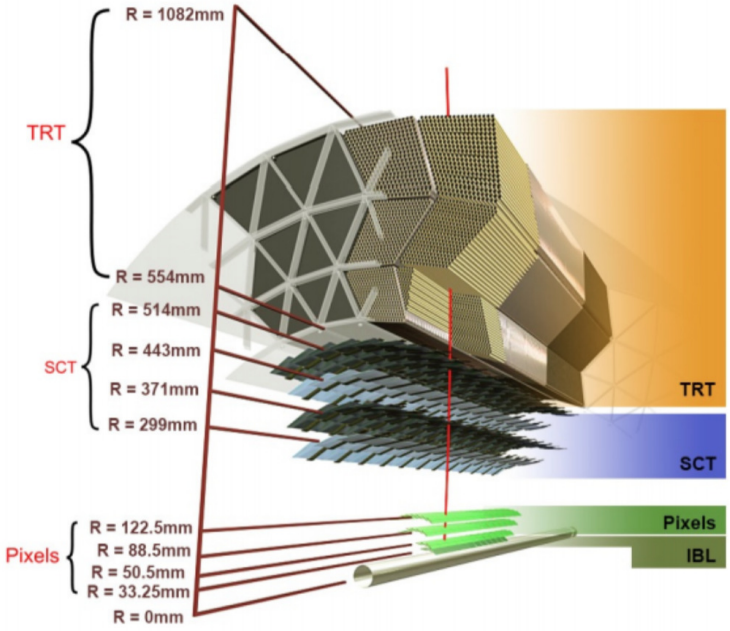
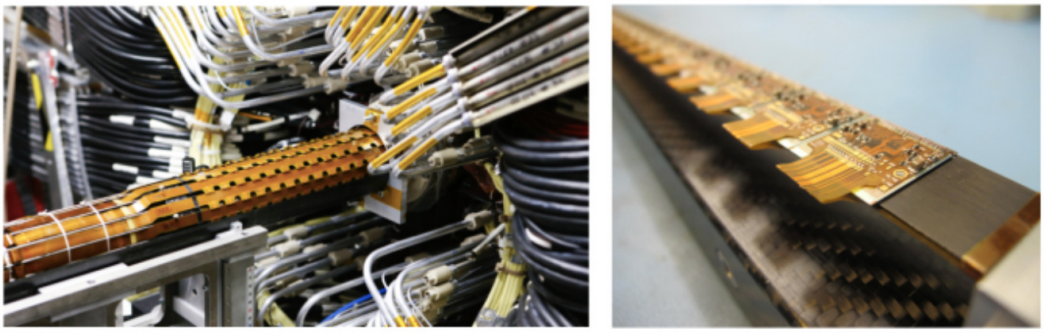


Fig. 2.13: Illustration of the Inner Detector (the barrel part).

the transverse plane ($R-\phi$ plane) and 66.5 ± 0.8 μm in the longitudinal z direction. The resolution of the Cartesian coordinates of reconstructed vertices is less than 0.35 mm for the z-direction and less than 0.16 mm in x and y directions [103]. With increasing number of associated tracks, the resolution improves by up to about one order of magnitude.



(a) (b)

Fig. 2.14: The ATLAS IBL detector before the insertion (left hand-side) and an IBL stave where the single detector modules are mounted on carbon fibre support structures (right hand-side).

SemiConductor Tracker

After the Pixel Detector, the SCT completes the high precision tracking. It is built around the pixel detector and it is designed to provide eight precision measurements per track in the intermediate radial range, contributing to the measurement of momentum, impact parameter and vertex position, as well as providing good pattern recognition by the use of high granularity. The SCT tracker includes four cylindrical layers of modules of silicon microstrips (each made up by two sensor at a ≈40 mrad stereo angle, allowing to measure three-dimensional space points) at average radii of about of 30.0, 37.3, 44.7 and 52.0 cm, and nine disks on each end-cap regions to cover the full tracking pseudo-rapidity range of $|\eta| < 2.5$. It has 4088 modules covering an area of approximately 63 m^2. The 2112 barrel modules are two-sided and use microstrip sensors. Two sensors, each containing 768 active strips, with a nominal thickness of 285 μm and a strip pitch of 80 μm are mounted on both sides the modules. The end-cap disks use 1976 wedge-shaped modules of three different sizes. Including all active readout strips, the SCT layers have more than 6.3 million readout channels in total. The spatial resolution is about 16 μm in the transverse plane and 580 μm in the longitudinal z direction. The system is operated at temperatures between -10 °C and -5 °C in order to attenuate different types of electronic noise.

Transition Radiation Tracker

The TRT [104] is the outermost part of the ID and it consists of straw detectors with diameter of 4 mm each and a 0.03 mm diameter gold-plated tungsten wire in the centre of the tube. In the barrel part, these tubes are arranged parallel to the beam axis placed at a radial region between 554 mm and 1082 mm with respect to the beam axis. It contains 50000 straw tubes with a length of 144 cm, divided into two parts with separate readout systems. The end-caps contain 250000 radial straw tubes of 37 cm length with the readout system connected to the outer radius. The tracking in the TRT is based on transition radiation, which is emitted by charged particles traversing a boundary of two dielectric materials. The tubes are

filled with non-flammable xenon-based gas (70% Xe, 20% CO_2 and 10% CF_4). A potential difference of 1.5 kV is applied to the central wire and the surface of the tube, producing an electric current via gas ionization induced by the transition radiation photons produced in polymer fibres (barrel) and foils (end-caps) surrounding the tubes. This current is detectable as a count in the readout system. Accordingly, this setup allows only to obtain an (R, ϕ) information. The single hit resolution is around 120 μm in the barrel region and 130 μm in the end-caps. Compared to the other parts of the inner detector, this is rather low. However, for each track, the TRT records around 30 hits, which significantly improves the spatial resolution of the track reconstruction if the information of the TRT is combined with the pixel detector and the SCT.
The TRT allows also particle identification by distinguishing between two different types of transition radiation thresholds for the signal readout. A low-threshold, optimized to detect direct ionization from particles penetrating the tube, whereas a high-threshold is sensitive to transition radiation photons. Emission of transition radiation is much more likely for electrons than for charged hadrons, meaning that the high-threshold probability is different over a wide energy range. This allows a discrimination of those particle types which is important, especially for an efficient separation between $\pi^{\pm}$ and $e^{\pm}$ tracks. The total number of readout channels in the TRT is approximately 351000. The barrel part covers pseudorapidities with $|\eta|$ < 1.1. Including the end-caps, its total range is extended to $|\eta|$ < 2.0.

2.4.5 Calorimeters

The ID is surrounded by the calorimeters [105, 106] which cover the range $|\eta| < 4.9$ and have a total diameter of 8.46 m and a length of 13.4 m. Their task is to identify and measure the energy of particles (both charged and neutrals) and jets. It also detects missing transverse energy (MET), produced by particles that escape the detector undetected such as neutrinos) by summing all the measured energy deposits: $E_T^{miss} = \sqrt{(\Sigma E_T \cos\phi)^2 + (\Sigma E_T \sin\phi)^2}$. To be more precise, E_T^{miss} is defined as the momentum imbalance in the plane transverse to the beam axis: the resultant of the negative vectorial sum of the momenta of all the particles, based on the conservation of momentum in the plane transverse to the beam axis z. This subsystem uses different techniques suited to the widely varying requirements of the physics processes of interest. The fine granularity of the electromagnetic calorimeter (ECAL) is ideally suited for precision measurements of electrons and photons. The coarser granularity of the rest of the calorimeter, the hadronic calorimeter (HCAL), is sufficient to satisfy the physics requirements for jet reconstruction and MET measurements. Like many high energy experiments, ATLAS has separate ECAL and HCAL up to $|\eta| < 3.2$. A special combined electromagnetic and hadronic calorimeter covers the range $3.1 < |\eta| < 4.9$. The calorimeters contain dense materials (absorbers), which cause an incoming particle to initiate a shower. Particles created in this shower are detected in the active material, which is interleaved with the absorbers. The total signal in the active material is a measure of the energy of the incoming particle. ATLAS experiment uses two types of active material: Liquid Argon (LAr) and scintillating plastic. Particles

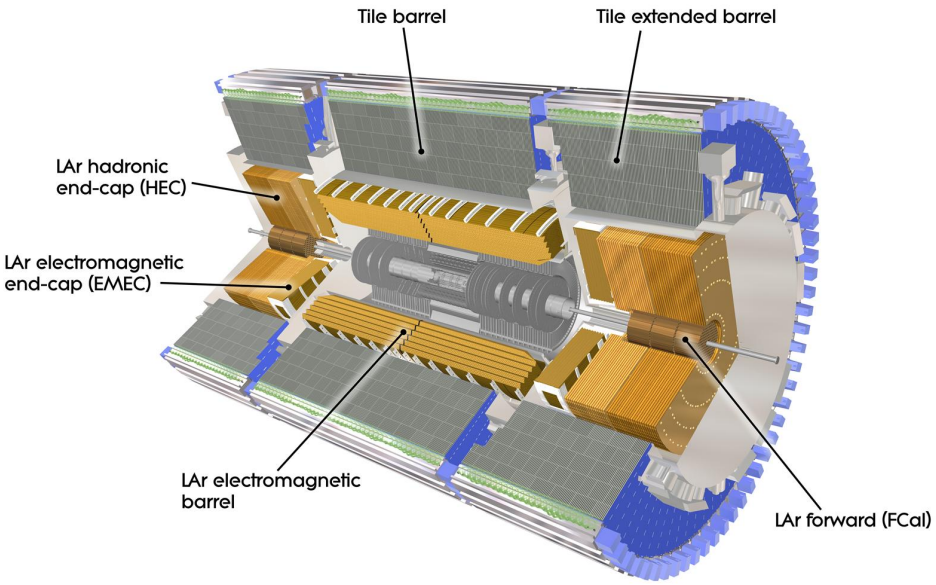


Fig. 2.15: Cut-away view of the ATLAS calorimeter.

that traverse the liquid-argon create charge by ionization, which is collected on readout electrodes. The scintillating plastic is doped with fluorescent molecules, which emit light when the atoms in the plastic are excited by a passing particle. The light is detected and amplified by photomultiplier tubes. For the absorbers several types of material are used: lead, steel, cooper and tungsten. The general layout of the calorimeter is shown in Fig. 2.15.

Electromagnetic Calorimeter

The EM calorimeter is a sampling calorimeter made of LAr as scintillating material and lead as absorbing material, providing complete ϕ symmetry without azimuthal cracks (more details can be found in [107]). As it is visible in Fig. 2.16, the EM calorimeter has an accordion shape, which provides complete coverage and symmetry in the azimuthal angle ϕ. The total depth of the EM calorimeter exceeds 22 radiation lengths(X_0) in the barrel region and 24 in the endcap. The resolution provided by the EM calorimeter is:

$$\frac{\sigma\left(E\right)}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.3\% \tag{2.8}$$

where the first term is the sampling term and the second term is a constant term, and

$$\sigma\left(\eta\right) = \frac{40 \text{ mrad}}{\sqrt{E}} \tag{2.9}$$

where E must be used in GeV. These are design values taken from Ref. [108].
In the region $|\eta| < 2.5$, the EM calorimeter is longitudinally segmented in three layers:

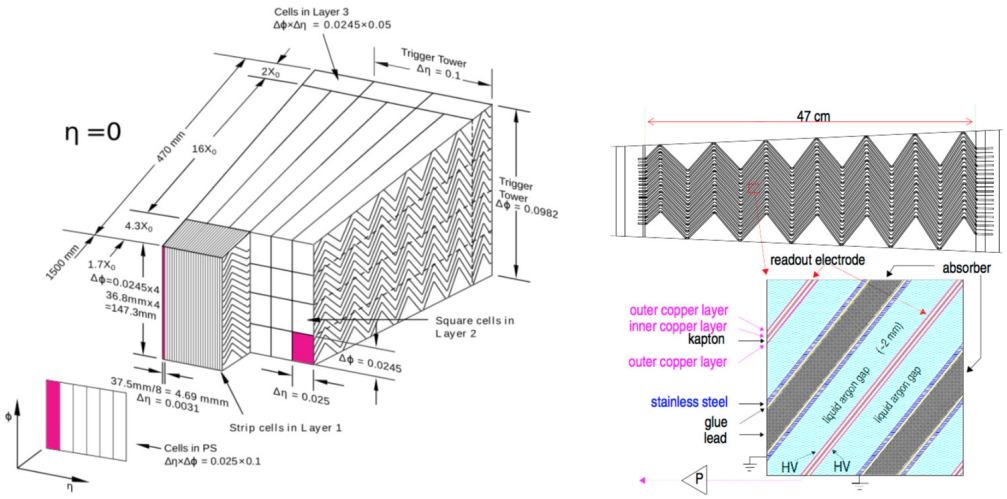


Fig. 2.16: Read-out granularity and accordion shape of the barrel EM calorimeter.

1st Sampling (L1)

The very high segmentation in η of the first layer permits a precise measurement of the shower position and a good rejection of fake photons from π^0 decay, as well as is very useful to separate electrons from $\pi^{\pm}$.

2nd Sampling (L2)

The second layer is the largest and allows to collect the larger fraction of the energy of the incoming particles and provide the most precise measurement in ϕ; it is segmented in square towers ($\Delta\eta \times \Delta\phi = 0.025 \times 0.025$).

3rd Sampling (L3)

The last layer is dedicated to the measurement of the tails of the particle showers.

An additional thin pre-sampler layer covering the $|\eta| < 1.8$ region where the amount of material seen by an incident particle before the calorimeter front face is > 2 radiation lengths (X_0), is used to correct for fluctuations in energy losses of particles before they reach the calorimeter. The total thickness of the EM calorimeter increases from 22 X_0 to 30 X_0 in region from η = 0 to η = 0.8 and 24X_0 to 33 X_0 in region from η = 0.8 to η = 1.3. The amount of material in terms of X_0 is shown in Fig. 2.17.

Hadronic Calorimeters

Surrounding the latter is the HCAL, which is divided in the Tile Calorimeter (TileCal), the Hadronic End-cap Calorimeters (HEC) and the Forward Calorimeter (FCal) with the inner radius of 2.28 m and the outer radius of 4.23 m. Similarly to the EM calorimeter, the Tile calorimeter is longitudinally segmented into three layers, which are needed for triggering and reconstruction of jets. The readout of the tiles is performed by optical fibres. The tiles are grouped into readout cells,

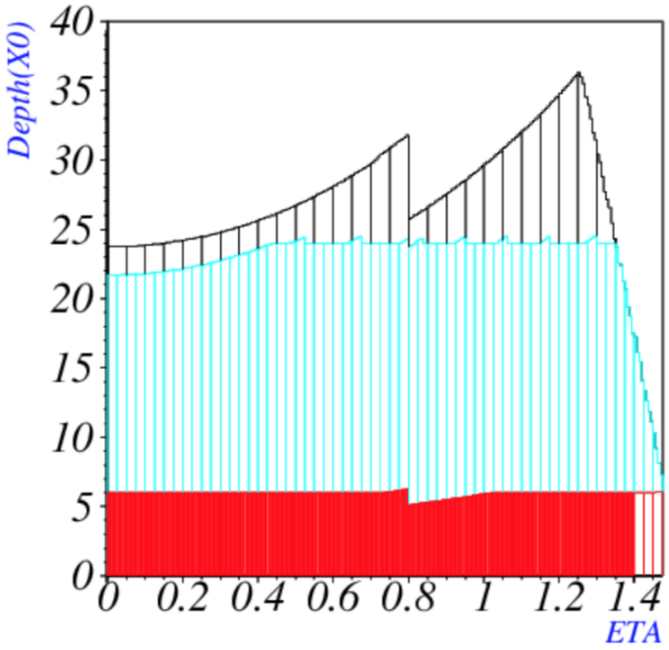


Fig. 2.17: The amount of material traversed by a particle before and in the EM calorimeter, in units of radiation lengths X_0 , as a function of $|\eta|$. Different colors represent three different longitudinal layers of the EM calorimeter.

which are designed to be projective with respect to the interaction point. The TileCal covers the central region ($|\eta| < 1.7$) and uses plastic scintillator plates (called tiles) as active material and steel as absorber. The HEC extends up to $|\eta| = 3.2$ and relies on liquid-argon as active material and copper as absorber material. In this region ($|\eta| < 3.2$), the hadronic calorimeter can provide a resolution which is:

$$\frac{\sigma\left(E\right)}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\% \tag{2.10}$$

In the very forward region up to $|\eta| = 4.9$, the FCal is installed to improve the measurement of the missing transverse energy. Again liquid-argon was chosen as active material while the absorbing material is composed of copper and tungsten. The ultimate energy resolution is expected to be:

$$\frac{\sigma\left(E\right)}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\% \tag{2.11}$$

(in Eq. 2.10 and in Eq. 2.11 the energy must be considered in GeV). Design values taken from Ref. [108].
The cumulative amount of material at the end of the active calorimetry region varies between 10 and 18 nuclear interaction lengths.

2.4.6 Muon Spectrometer

The muon spectrometer (MS) surrounds the calorimeter and it consists of the precision-tracking chambers to accurately and precisely measure the transverse

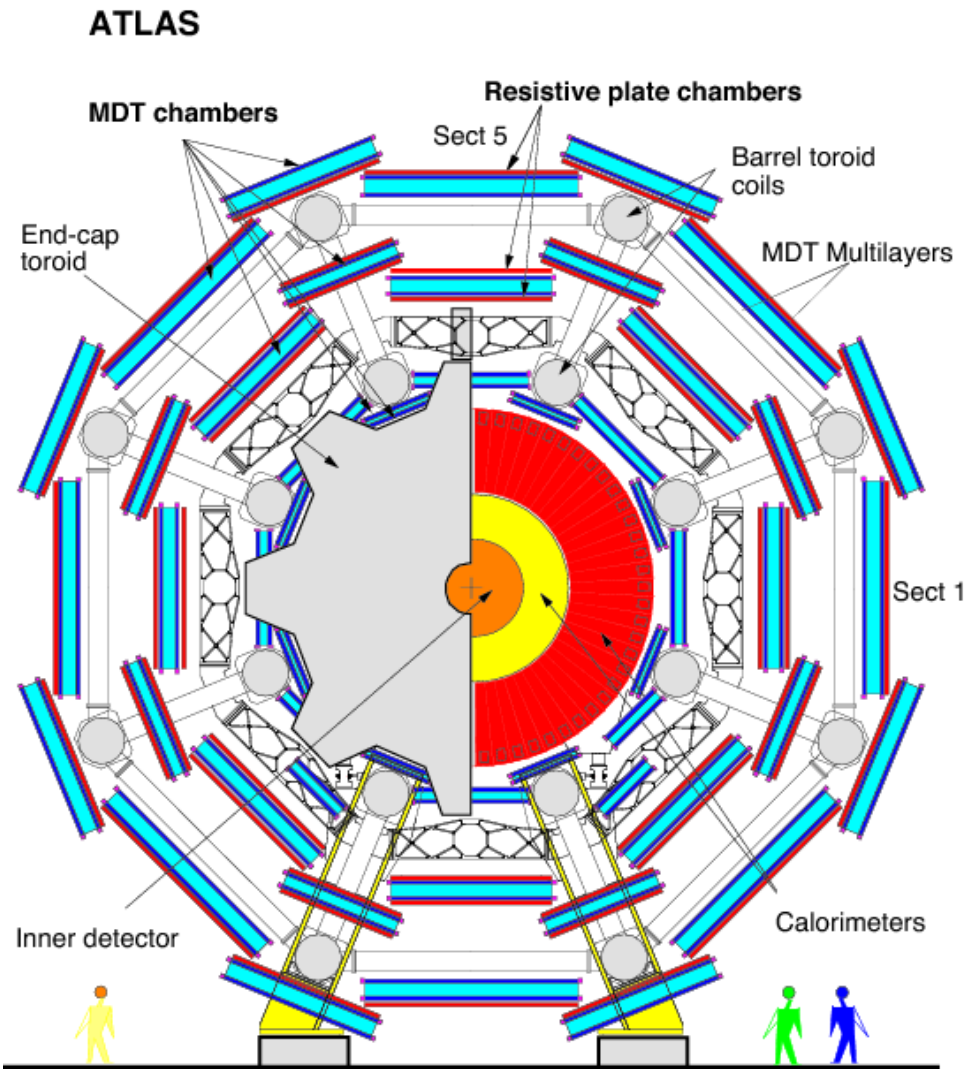


Fig. 2.18: Layout of the muon spectrometer in the *x*-*y* plane (barrel region).

momentum of the muons and the trigger chambers with fast response to trigger on the muons. The MS is immersed in a toroidal magnetic field of approximately 0.5 T and 1 T in the barrel and endcap regions, respectively. The MS is designed to be capable of standalone measurements of the transverse momentum of the muons in a wide range: down to about 3 GeV which is constrained by the muon energy loss in the calorimeters, and up to about 3 TeV. For 1 TeV muon tracks, the performance goal is to achieve approximately 10% stand-alone transverse momentum resolution, which corresponds to a sagitta along the *z*-axis of about 500 μm to be measured with a resolution of 50 μm or less. Fig. 2.18 shows the layout of the muon chambers in the *x*-*y* plane, Fig. 2.19 shows a cut view of one quadrant on the MS. There are two different functions that muon chambers must accomplish: triggering and high precision tracking. The trigger system covers the region up to $|\eta| < 2.4$, and is composed by Resistive Plate Chambers (RPCs) in the barrel and Thin Gap Chamber (TGC) in the end-caps. The triggering system provides bunch-crossing identification (BCID), well-defined p_T thresholds and a measurement of the muon coordinate in the direction orthogonal to the chambers dedicated

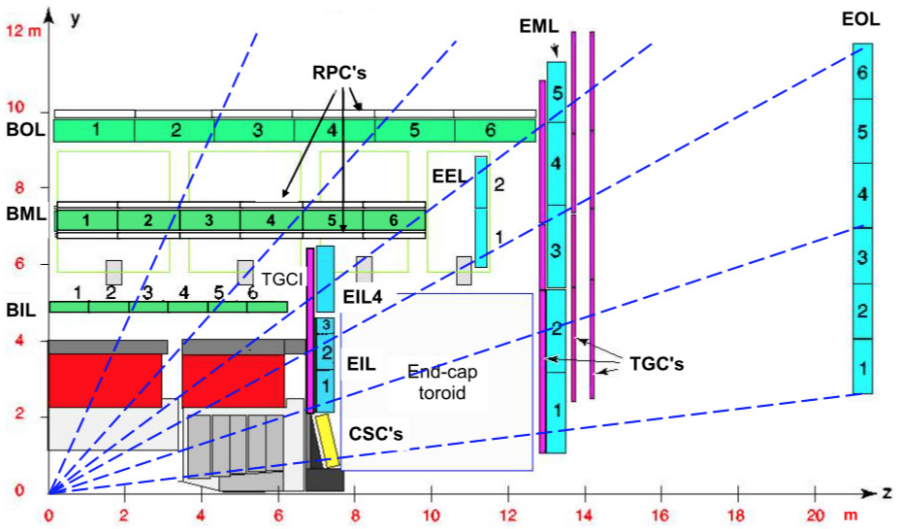


Fig. 2.19: Cut-view of one quadrant of the Muon Spectrometer.

to precision tracking. The tracking is performed by the Monitored Drift Tubes (MDTs) and by Cathode Strip Chambers (CSCs) at large pseudorapidity. High precision mechanical assembly techniques and optical alignment systems provide the essential alignment of the chambers, while the magnetic field reconstruction relies on Hall sensors distributed throughout the spectrometer volume.
Different magnetic fields permit to perform the measure of the momentum:

- In barrel region ($|\eta| < 1.4$) the magnetic field is provided by the barrel toroid.
- For $1.4 < |\eta| < 1.6$, by a combination of barrel toroid and end-cap toroid fields.
- For $1.6 < |\eta| < 2.4$, by two end-cap toroids' fields.

The overall momentum resolution provided by the muon system is:

$$\frac{\sigma\left(p_T\right)}{p_T} = 3\% \text{ at 50 GeV}$$
$$\frac{\sigma\left(p_T\right)}{p_T} = 10\% \text{ at 1 TeV}$$

These are design values and they are taken from Ref. [109]. The MDT and the CSC provide a single hit resolution in the bending plane of about 80 μm and 60 μm, respectively. The muon track in the MS is reconstructed in two steps. In the first step the muons are triggered in the RPC/TGC and local track segments are defined in each layer of chambers. In the next step, the local track segments from different layers are combined through a χ^2-fit forming a full MS track. To reduce the probability of background tracks penetrating the calorimeter, the fitted tracks of the muon candidates are required to point towards the interaction point.
Selected parameters characterize the performance of the muon sub-systems are summarized in Table 2.3. The following sections will briefly describe the technology and performances of the muon sub-systems.

Detector Type	Chamber z/R	Resolution (RMS) ϕ	time	Number of chambers	channels
MDT	35 μm	-	-	1088 (1150)	339k (354k)
CSC	40 μm	5 mm	7 ns	32	30.7k
RPC	10 mm	10 mm	1.5 ns	544 (606)	359k (373K)
TGC	2 - 6 mm	3 - 7 mm	4 ns	3588	318k

Table 2.3: Parameters of the sub-systems of the MS. The quoted spatial resolution (columns 2, 3) does not include chamber-alignment uncertainties. Column 4 shows the intrinsic time resolution of each chamber type. Numbers in brackets refer to the complete detector configuration as planned for 2009.

Monitored Drift-tube Chambers

The MDT chambers are made of drift tubes with a diameter of 30 mm. The wall of the tube is made using aluminium and is 400 μm in thickness. A tungsten-rhenium wire of 50 μm in diameter is located at the centre of the tube. The tubes are operated with a non-flammable mixture of 93% Ar and 7% CO_2 at an absolute pressure of 3 bar. A muon passing the tube is ionized in the drift field. The electrons from the ionization drift towards the wire with a maximum drift time of 700 ns when drifting from the wall to the wire. The position of the muon thus can be determined through the tube's non-linear space and time relation. The spatial resolution of a MDT tube is about 80 μm.

Cathode Strip Chambers

The CSCs are multi-wire proportional chambers with the wires oriented in the radial direction. Both cathodes are segmented, one with the strips perpendicular to the wires (providing the longitudinal coordinate) and the other parallel to the wires providing the transverse coordinate. The position of the track is obtained by interpolation between the charges induced on neighbouring cathode strips while wire signals are not read out. The resolution of a CSC plane is about 60 μm, while in the non-bending direction, the cathode segmentation is coarser leading to a resolution of 5 mm.

Resistive Plate Chambers

In the barrel, RPCs are used to build trigger chambers because of good spatial and time resolution, and adequate rate capability. The RPC is a gaseous parallel electrode-plate detector, operated with a mixture of $C_2H_2F_4$/Iso-C_4H_{10}/SF_6 (94.7/5/0.3), which combines the merits of relatively low operating voltage, non-flammability and low cost, while providing a comfortable plateau for safe avalanche operation. It contains two resistive plates made of phenolic-melaminic plastic laminate. They are kept parallel to each other at a distance of 2 mm by insulating spacers, and the electric field between them is about 4.9 kV/mm, allowing avalanches to form along the ionizing tracks towards the anode. The signal is read out via capacitive coupling to metallic strips, which are mounted on the outer faces of the resistive plates.

Thin Gap Chambers

In the endcap region, TGCs have been selected to build the trigger chambers due to good time resolution and high rate capability. In addition to trigger capability, TGCs provide a second azimuthal coordinate to complement the measurement of the MDTs in the bending (radial) direction. TGCs are multi-wire proportional chambers operated in a highly quenching gas mixture of CO_2 and n-C_5H_{12} (n-pentane), with the characteristic that the wire-to-cathode distance of 1.4 mm is smaller than the wire-to-wire distance of 1.8 mm. This cell geometry allows for operation in a quasi- saturated mode, i.e. with a gas gain of 3×10^5.

2.4.7 Trigger System

The trigger system is an essential component of hadron collider experiments, because the rate of events containing interesting physics phenomena is only a small fraction of the total event rate. Due to the constraint from data storage capacity and rates, it is impossible to store all the events produced in the collisions, therefore, trigger system is employed to rapidly decide with high efficiency potential interesting events produced in a given bunch-crossing interaction to record for later analysis. At the LHC design luminosity, the proton-proton collisions at the LHC take place every 25 ns (40 MHz) with around 23 interactions per crossing.The data-size of one recorded collision is in the order of 1 Mb. These regime can give in total 40 Tb/s which is too high for the modern computer technology. The goal of the ATLAS trigger system is to reduce the rate of candidate collisions from 40 MHz to 100 Hz without a loss of interesting physics events, in order to be handled by the ATLAS computer system, which requires 100 Mb/s [110]. To reduce the amount of data, the online event selection (called trigger system) has to decide whether a collision looks interesting or not and for that employs several strategies. The trigger system selects events by identifying signatures of muons, electrons, photons, tau leptons, jets and B meson candidates, as well as using global event signatures, such as MET. In Run 2, the trigger system consists of two levels of event selections: a Level-1 trigger (L1) reducing event rate to 100 kHz and a high-level trigger (HLT) [111] further reducing event rate to 1 kHz. In LHC Run 1, the HLT includes two levels of event selections: Level-2 and Event Filter farms, which were merged into a single homogeneous farm in Run 2 to improve resource sharing and for simplification; the L1 trigger is implemented as hardware devices in the detector, while the HLT is almost software. With this configuration, the read-out system (ROS) computers save network bandwidth and decrease the read-out request rate. For Run 2, many multivariate analysis techniques have been developed and used at the HLT stage, as for example, in electron and photon trigger systems [112]. The ATLAS Trigger and Data Acquisition (TDAQ) system is illustrated in the block diagram shown in Fig. 2.20 [113].

Level 1 Trigger

There are three L1 subsystems. One is the Level-1 calorimeter trigger (L1Calo), which uses calorimeter information of reduced granularity to trigger electrons,

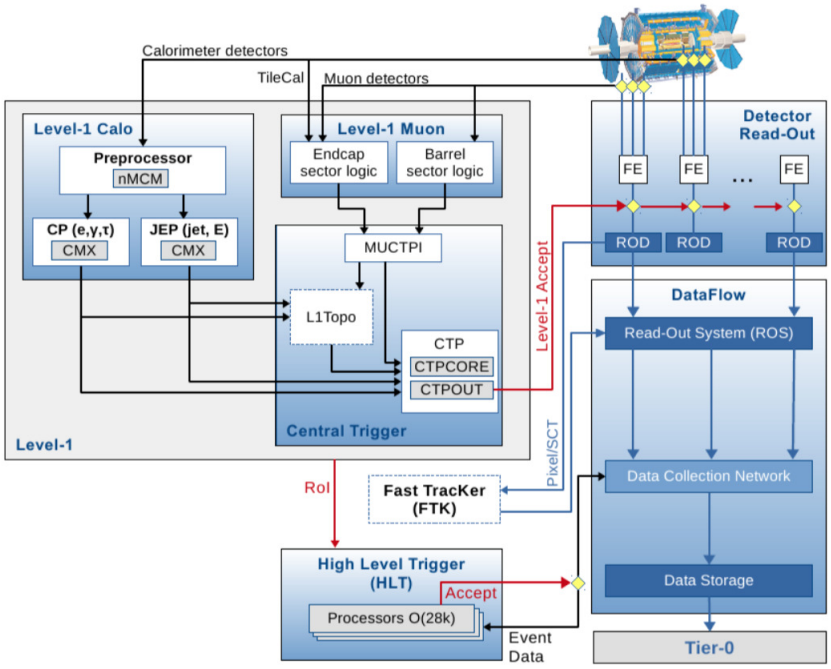


Fig. 2.20: The ATLAS TDAQ system in Run 2 with emphasis on the components relevant for triggering. L1Topo and Fast TracKer (FTK) were being commissioned during 2015.

photons, hadronic tau decays, jets and (missing) transverse energy. The other system with a direct detector interface is the Level-1 muon trigger (L1Muon), which primarily uses TGC and RPC information to make fast decisions on muon trigger items. In the Central Trigger Processors [114], the information of L1Calo and L1Muon is received and the final decision is made. For this it can also use the recently added Level-1 topological trigger (L1Topo) [115] that combines information from multiple L1Calo and L1Muon objects as well as their kinematics (e.g. angular separation, invariant mass); then the CTP makes the final decision.

High Level Trigger

Events accepted by any L1 trigger, are further processed by the HLT, which operates on the full event information (or in some cases on regions of interest pre-selected by the Level 1 trigger). While the Level 1 trigger uses custom-made electronics based on FPGAs, the HLT runs on a conventional computer cluster and has about 0.2 s on average to decide. The target average HLT rate is 1 kHz, a significant improvement compared to 0.6 kHz in LHC Run 1. The set of active triggers usually changes during the data taking, as the intensity of the beam changes. Trigger items may be prescaled, e.g. a specified fraction of positive trigger decisions is randomly ignored, which makes it possible to use triggers which unprescaled rates are too high to record their events. A fast tracker trigger system (FTK) [116] which provides tracking information to the HLT is being prepared for operation later in Run 2. Among others, this will benefit tau and b-jet triggers. Another important pending upgrade will improve the granularity of calorimeter information available to L1Calo [117] and therefore also benefit tau

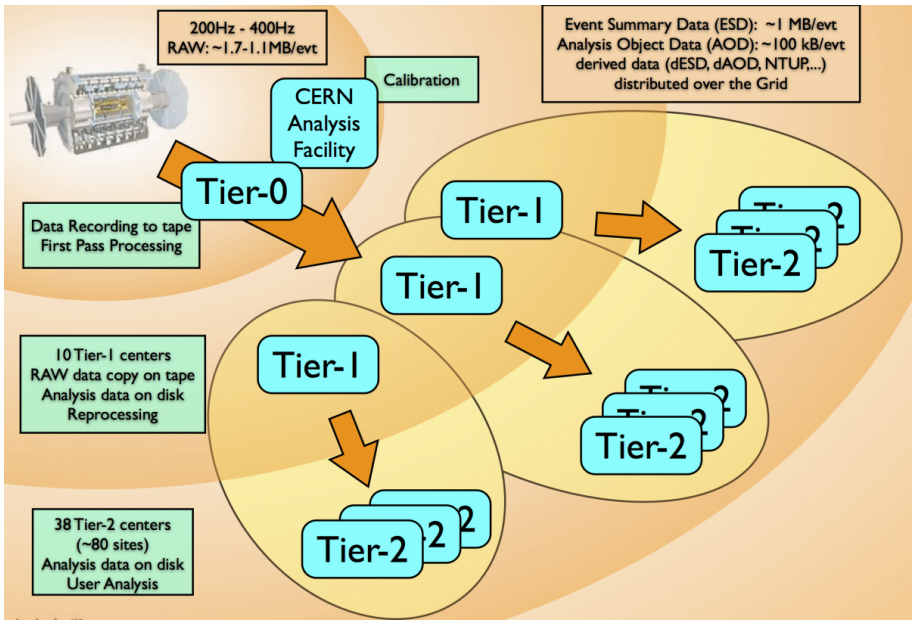


Fig. 2.21: The tier structure of LHC computer grid.

triggers. It will be installed during Long Shut-down 2 (from December 2018 until February 2021).

2.4.8 Computing Models: the GRID

ATLAS works with large amount of data and need sufficient computer facilities to process this data. For the LHC, storage and processing requirements exceed by far the capacities available at a single site and hence a new approach was chosen, which is commonly known under the LHC Computing Grid (LCG) [118, 119]. This computing grid provides an infrastructure for the storage of data and the necessary computing power for the physics analyses and simulations. The data distribution follows a so-called tier-structure (see Fig. 2.21). The LHC data is recorded in a first step on tape at Tier-0 centre at CERN; from there, it is further distributed to worldwide Tier-1 centres, which also stores a large part of data and provides a 24-hours support. On the other hand, the tier-2 centres store only a small part of data since they are designated for user specific physics analysis and simulation. The main advantage of this tier structure is that several copies of data exist. This fact means that a single tier centre is independent from the others and that minimizes the critical points in the infrastructure.

2.4.9 ATLAS Software

The main goal of the ATLAS offline software is to process all data coming from the trigger and DAQ system and to allow users to analyze them with specific tools. In case of the ATLAS experiment, there is an offline software framework called Athena [120], with the underlying architecture starting from the GAUDI one originally developed by the LHCb collaboration. It is written mainly in C++ and

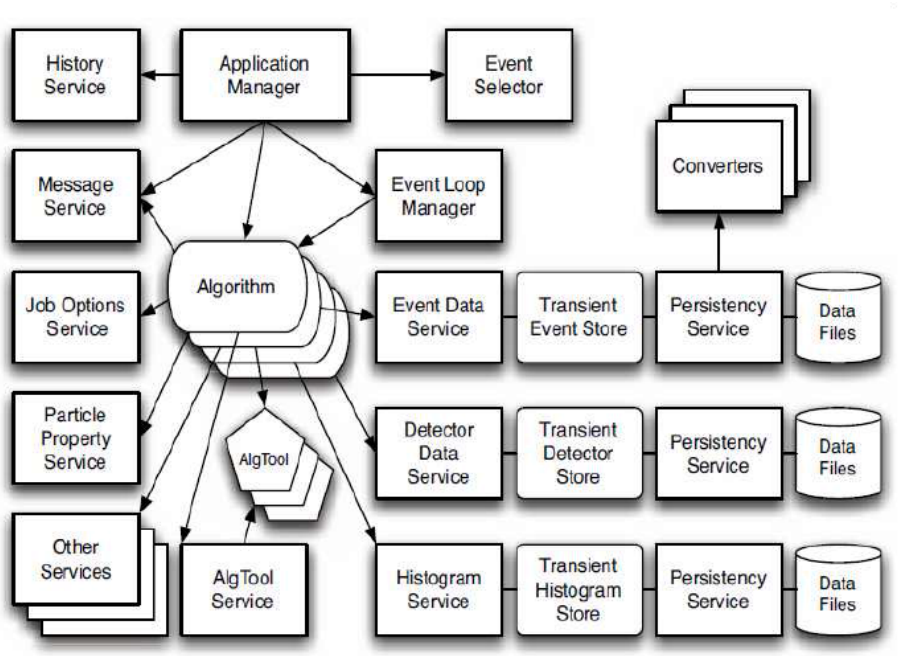


Fig. 2.22: The Athena architecture.

therefore object oriented, with the configuration files written in Python. This framework encompasses not only the reconstruction and analysis algorithms needed for the ATLAS data, but also all the other software needed for the HEP computing; in addition, the Athena framework is used also for generating all Monte Carlo (MC) samples needed by the analysers (the major components of its framework are shown in Fig. 2.22).

Athena provides an integrated communication between various software applications, called algorithms, within the framework. The principle functionality of an algorithm is to take input data, manipulate it and produce new output data. Therefore, Athena processes data event by event. All algorithms cam make their data accessible to access via other algorithms (for example the event information part way-through the reconstruction). Moreover, Athena framework also contains services which are sizeable components that are setup and initialized once at the beginning of the job by the framework and used by many algorithms as often as they are needed, for example the detector description. Additionally, the Athena framework allows dynamic loading of libraries and is organized in form of plug-in modules and allows a flexible configuration of various algorithms to be executed. For data storage, plotting and processing histograms, Athena uses the ROOT framework [121]; all ATLAS data are stored in one of the ROOT formats, which provide a size compressing and fast access of the data.

Chapter 3

Object reconstruction in ATLAS

The unprocessed output of the ATLAS detector or its simulation is not suitable for most physics analyses; indeed, the raw data from the ATLAS detector is a series of readout signals. The task of the offline reconstruction software is to translate this stream of readout signals into the information necessary to perform physics analysis. Reconstruction of physics objects such as electrons, muons, taus and jets begin with reconstruction of basic objects such as tracks (in the ID) and calorimeter clusters. This chapter describes the reconstruction and identification algorithms for the physics objects relevant for the analyses presented in this thesis.

3.1 Low-level objects

To reconstruct the trajectories of charged particles, the Inner Detector (ID) hits are associated to form tracks. In ATLAS, the tracks of primary charged particles are built using *inside-out* track finding [122, 123]. This algorithm is seeded using information of the innermost layers of the pixel detector and then follows the general direction to the outer layers of the Inner Detector to build track candidates. Afterwards, a Kalman filter [124] is applied iteratively to all hits associated to the track candidate and outlier hits are rejected. The numbers of misidentified trajectories and tracks with shared hits is reduced using dedicated ambiguity solving algorithms. In a last step the track is extended to the TRT and its hits here are fitted together with the track from the silicon detectors. The resulting track has five parameters (as already specified in Sec. 2.8), which consist of the location on a plane, as well as two angles and the curvature of the trajectory. Additionally, *outside-in* track finding is performed, which is seeded in the TRT and directed towards the silicon detectors. In particular, this recovers tracks from decays of long-lived particles that decay beyond the silicon detectors. Based on track quality criteria like the number of silicon hits, tracks are usually filtered further for the reconstruction of higher level physics objects. The commonly used "*loose*" track quality selection results in reconstruction efficiencies over 90%, where the exact value depends on the pseudorapidity and track transverse momentum, p_T [125]. Important track properties are the transverse and longitudinal impact parameters d_0 and z_0, the projections of the distance between the track and a reference

point at the closest point of approach. Their resolution is typically well below 0.01 mm. Tracks passing certain quality criteria are considered for the reconstruction of vertices, which are points of particle interaction. In an iterative process, tracks are associated to vertex candidates and fits are performed to determine the vertex position [126] (further details will follow in Section 3.2). The positions of the proton-proton interactions are called primary vertices. The reconstruction efficiency rises from 83% for vertices with two tracks to almost 100% for vertices with at least five tracks. The hard scattering vertex is chosen as the primary vertex with the highest $\sum p_T^2$.

To reduce the impact of electric noise and pile-up, and to facilitate further calorimeter-based object reconstructions, the signals of adjacent calorimeter cells are condensed to form *topological clusters* [127]. Starting from a cell with an energy deposition at least four times higher than its expected noise level, topologically connected cells are added to the cluster using a *growing-volume* algorithm [127]. The cluster energy is calibrated to compensate for inactive detector material and non-compensating calorimeter response. Clusters can be calibrated at the electromagnetic scale, which is tuned to clusters in showers created by electrons (EM topo-clusters). Another common approach, the *local hadronic cell weighting*, corrects for differences between pion and electron showers by estimating the likelihood that a cluster is in an electromagnetic or hadronic shower and weighting the contributions accordingly (LCW topo-clusters).

As regards the tracking in the Muon Spectrometer (MS), the tracks start from segments reconstructed within each muon chamber. In each MDT chamber and nearby trigger chamber, a Hough transform is used to search for hits. Segments are reconstructed by performing a straight-line fit to the hits found in each MDT layer. The RPC or TGC hits measure the coordinate orthogonal to the bending plane. Segments in the CSC detectors are reconstructed by a combinatorial search for hits in the η and ϕ CSC planes, including a *loose* requirement on the compatibility of the segment with the luminous region. Muon track candidates are then built by fitting together hits from segments in different layers. The algorithm used for this task performs a segment-seeded combinatorial search that starts by using as seeds the segments generated in the middle layers of the detector where more trigger hits are available. The search is then extended to use the segments from the outer and inner layers as seeds. The segments are selected using criteria based on hit multiplicity and fit quality and are matched using their relative positions and angles. At least two matching segments are required to build a track, except in the barrel endcap transition region where a single high-quality segment with η and ϕ information can be used to build a track. The same segment can initially be used to build several track candidates. Later, an overlap removal algorithm selects the best assignment to a single track, or allows for the segment to be shared between two tracks. To ensure high efficiency for close-by muons, all tracks with segments in three different layers of the spectrometer are kept when they are identical in two out of three layers but share no hits in the outermost layer. The hits associated with each track candidate are fitted using a χ^2 fit. A track candidate is accepted if the χ^2 of the fit satisfies criteria on e.g. hits position, energy deposit along the tracks etc. The track candidate will be re-fitted in cases where hits making large

contributions to the χ^2 are removed or additional hits consistent with the track candidate are recovered.

3.2 Primary Vertex

The input to the vertex reconstruction is a collection of reconstructed tracks. The procedure of primary vertex reconstruction is divided into two stages: vertex finding and vertex fitting [126]. The former stage generally denotes the pattern recognition process: the association of reconstructed tracks to vertex candidates. The vertex fitting stage calculates the best-fit vertex position and its covariance matrix. The strategy can be briefly outlined in these steps:

- A set of tracks satisfying the track selection criteria is defined.
- A seed position for the first vertex is selected, using a sliding window of 0.7 cm that is moved along the whole interaction region.
- The tracks and the seed are used to estimate the best vertex position with a fit following an iterative procedure. In each iteration, less compatible tracks are down-weighted and the vertex position is recomputed.
- Tracks incompatible with the vertex determined in the previous step are removed and allowed to be used in the determination of another vertex.
- The whole procedure is repeated with the remaining tracks in the event to fit another vertex.

All vertices with at least two associated tracks are retained as valid primary vertex candidates. The output of the vertex reconstruction algorithm is a set of three dimensional vertex positions and their covariance matrices.

3.3 Leptons

In this section the methods used for reconstructing leptons are presented. The attention is focused on electrons and muons because these object characterise the final states of the analyses presented in this thesis.

3.3.1 Muons

Muon Reconstruction

Muons are minimum ionizing particles and they deposit minimum amount of energies when travelling across the calorimeters. The detector components involved in muon reconstruction are the muon spectrometer and the inner detector. However, muons also deposit some energy in the ECAL. The muon system allows the identification of muons with a p_T above 3 GeV. Muons with momenta lower than that are difficult to reconstruct since they do not reach the spectrometer, loose

too much energy in the calorimeter and/or do not leave a significant signal over the noise in the muon spectrometer.
Muon reconstruction is performed based on information from the ID, MS and calorimeters. There are four different types, depending on different reconstruction methods and subdetectors used [128]:

- **Combined muons (CB)**: The combined muon reconstruction algorithm takes the SA muons and matches them to inner detector tracks which satisfy a χ^2 quality test; the muon's momentum is computed as a weighted combination of the momentum measurements in the ID an in the MS. Muons within $|\eta| < 2.5$ can be reconstructed in this way.

- **Segmented-tagged muons (ST)**: This algorithm has been designed to improve the reconstruction of low p_T muons, which do not reach the outer muon station. ST muons are reconstructed by extrapolating a track in the ID to the first station of the MS: the quality of this matching procedure is evaluated using a χ^2 test defined by the difference the MS track and the prediction from the extrapolated ID track (as for CB muons), the $|\eta|$ coverage is $|\eta| < 2.5$.

- **Calorimeter-tagged muons**: In the region of $|\eta| < 0.1$, ID tracks with 15 GeV < p_{T} < 100 GeV are identified as muons if their energy deposits in the calorimeter are consistent with minimum ionizing particles. CT muons recover muon acceptance in the region where the MS is only partially instrumented to allow for cabling and services to the calorimeters and the ID.

- **Extrapolated (ME) muons**: The muon trajectory is reconstructed based only on the MS track and a loose requirement on compatibility with originating from the IP. The parameters of the muon track are defined at the interaction point, taking into account the estimated energy loss of the muon in the calorimeters. In general, the muon is required to traverse at least two layers of MS chambers to provide a track measurement, but three layers are required in the forward region. ME muons are mainly used to extend the acceptance for muon reconstruction into the region 2.5 < $|\eta|$ < 2.7, which is not covered by the ID.

Muon Identification

To distinguish prompt muons from background like light meson decays, a muon identification is performed. For this, several quality requirements are applied, related to the compatibility of the measurement in the Inner Detector and muon spectrometer, the goodness of the combined track fit and the quality of the individual tracks. There are four kinds of muon identification criteria designed to address the specific needs of physics analyses:

- **Loose muons**, are designed to maximize the reconstruction efficiency while provide good-quality muon tracks.

- **Medium muons**, are designed to minimize the systematic uncertainties associated with muon reconstruction and calibration. Only CB and MS muons are considered, with addition requirement on the number of hits in muon chambers and a loose requirement on the compatibility between p_T measurements in the ID and MS.

- **Tight muons**, are selected to maximize the purity of muons at the cost of some efficiency. Only CB muons with at least two hits in MS and satisfying the medium criteria are considered. Additionally, requirements on normalized χ^2 of the track fit and the compatibility between ID and MS p_T measurements is employed to further reject background.

- **High-$\mathbf{p_T}$ muons**, are optimized for searches for high-mass Z' and W' resonances. To maximize the momentum resolution for muons with p_T > 100 GeV, CB muons satisfying medium requirements and having at least three hits in three MS stations are included. Regions of the MS are vetoed, where the alignment is suboptimal.

The '*Medium*' identification criteria is used as default selection for muons in ATLAS and it minimises the reconstruction and calibration uncertainties. Some corrections are applied to MC simulations in form of efficiency scale factors (SF), in order to match isolation and trigger efficiency in data. These corrections are obtained by comparing MC predictions to large samples of $J/\psi \rightarrow \mu^+\mu^-$ and $Z \rightarrow \mu^+\mu^-$ data events using the *tag-and-probe* method. In Fig. 3.1 the efficiency of the reconstruction for Medium muons as a function of their p_T is shown. As can be seen, it is very close to 99%.

Muon momentum scale and resolution

The muon momentum scale and resolution are studied using $J\psi \rightarrow \mu^+\mu^-$ and $Z \rightarrow \mu^+\mu^-$ decays. Although the simulation contains an accurate description of the ATLAS detector, the level of detail is not enough to describe the muon momentum scale to the per mille level and the muon momentum resolution to the percent level. To obtain such of agreement between data and simulation, a set of corrections is applied to the simulated muon momentum. To improve the precision of the procedure, the p_T and η distributions of the Z and J/ψ resonances in simulation are reweighted to the distributions observed in data. The invariant mass distribution for the $Z \rightarrow \mu^+\mu^-$ candidate is shown in Fig. 3.2. After correction, the line-shape of the Z resonance in simulation (red solid line) agrees with the data within the systematic uncertainties, demonstrating the overall effectiveness of the p_T calibration. A better demonstration of the effectiveness of the momentum calibration is obtained by comparing, in data and simulation, the measurement of the position $m_{\mu\mu}$ and resolution $\sigma_{\mu\mu}$ of the di-muon mass peaks, extracted in bins of η and p_T from fits to the $J\psi \rightarrow \mu^+\mu^-$ and $Z \rightarrow \mu^+\mu^-$ invariant mass distributions. Fig. 3.3 shows the position of the peak of the invariant mass distribution, $m_{\mu\mu}$, obtained from the fits to the J/ψ samples as a function of the pseudorapidity of the highest-p_{T} muon for CB pairs. It is visible that the simulation is in very good

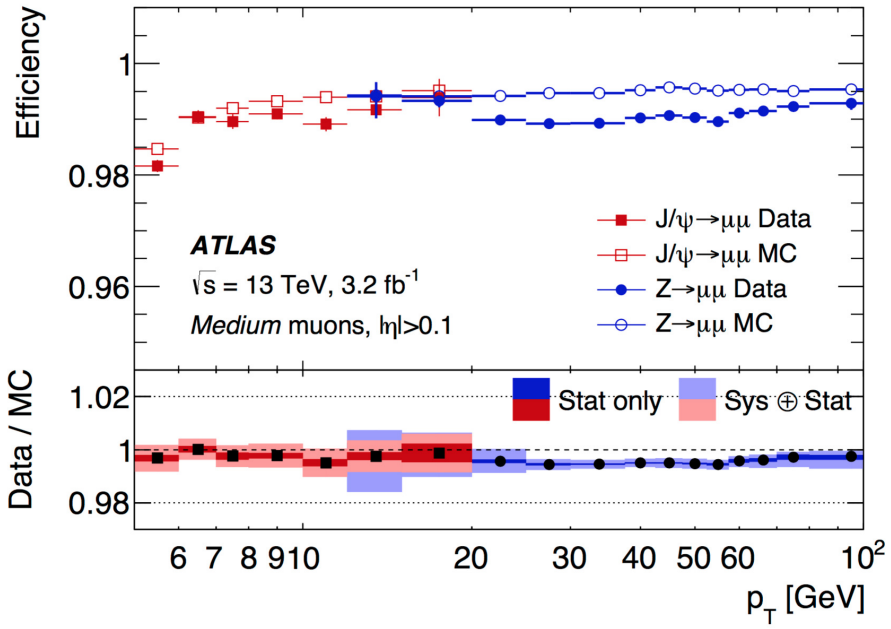


Fig. 3.1: Reconstruction efficiency for the *Medium* muon selection as a function of the p_T of the muon, in the region $0.1 < |\eta| < 2.5$ as obtained with $Z \to \mu^+\mu^-$ and $J\psi \to \mu^+\mu^-$ events. The error bars on the efficiencies indicate the statistical uncertainty. The plot is taken from Ref. [128].

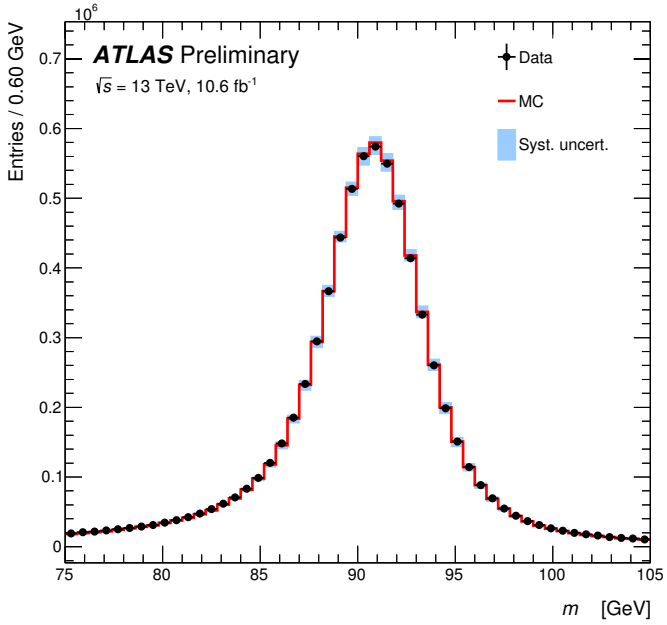


Fig. 3.2: Dimuon invariant mass distribution of $Z \to \mu^+\mu^-$ events reconstructed with combined muons. The points show the data. The continuous line corresponds to the simulation with the momentum corrections applied. The band represents the effect of the systematic uncertainties on the momentum corrections. Simulations are normalised to data. The plot is taken from Ref. [128].

agreement with the data after applying the muon momentum scale and resolution SFs.

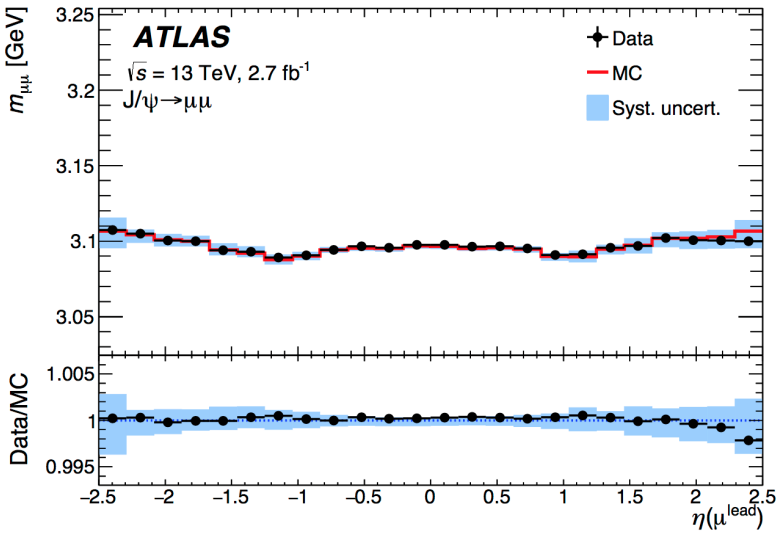


Fig. 3.3: Fitted mean mass of the dimuon system for CB muons for $J\psi \to \mu^+\mu^-$ events for data and corrected simulation as a function of the η of the highest-p_T muon. The upper panels show the fitted mean mass value for data and corrected simulation.The lower panels show the data/MC ratio. The error bars represent the statistical uncertainty; the shaded bands represent the systematic uncertainty. The plot is taken from Ref. [128].

Muon Isolation

Several isolation selection criteria (*isolation working points*) are defined, each optimised for different physics analyses. Table 3.1 lists the most relevant isolation working points for the analysis presented here, specifying the discriminating variables and the criteria used in their definition. The track-based isolation variable,

Isolation WP	Definition
FixedCutLoose	$p_{\mathrm{T}}^{varcone30}/p_{\mathrm{T}}^{\mu} < 0.15$, $E_{\mathrm{T}}^{topocone20}/p_{\mathrm{T}}^{\mu} < 0.30$
Gradient	$\epsilon = (0.1143 \cdot p_{\mathrm{T}}\ [\mathrm{GeV}] + 92.14)\%$, $\epsilon = (0.1143 \cdot p_{\mathrm{T}}\ [\mathrm{GeV}] + 92.14)\%$
FixedCutTightTrackOnly	$p_{\mathrm{T}}^{varcone30}/p_{\mathrm{T}}^{\mu} < 0.06$
FixedCutTight	$p_{\mathrm{T}}^{varcone30}/p_{\mathrm{T}}^{\mu} < 0.06$, $E_{\mathrm{T}}^{topocone20}/p_{\mathrm{T}}^{\mu} < 0.06$

Table 3.1: Definition of the most relevant isolation working points for the analysis presented here. The discriminating variables and the criteria used in the definition are reported in the third column.

$p_{\mathrm{T}}^{varcone30}$, is defined as the scalar sum of the transverse momenta of the tracks with $p_{\mathrm{T}} > 1$ GeV in a cone of size $\Delta_R = \min(10\ \mathrm{GeV}/p_{\mathrm{T}}^{\mu}, 0.3)$ around the muon of transverse momentum p_{T}^{μ}, excluding the muon track itself. The cone size is chosen to be p_{T}-dependent to improve the performance for muons produced in the decay of particles with a large transverse momentum. The calorimeter-based isolation

variable, $E_{\mathrm{T}}^{topocone20}$, is defined as the sum of the transverse energy of topological clusters [129] in a cone of size $\Delta R = 0.2$ around the muon, after subtracting the contribution from the energy deposit of the muon itself and correcting for pile-up effects.
The efficiencies for the isolation working points are measured in data and simulation in $Z \to \mu^+\mu^-$ decays using the *tag-and-probe* method. To avoid probe muons in the vicinity of a jet, the angular separation ΔR between the probe muon and the closest jet, reconstructed using an anti-k_t algorithm [130] with radius parameter 0.4 and with a transverse momentum greater than 20 GeV is required to be greater than 0.4. In addition, the two muons originating from the Z boson decay are required to be separated by $\Delta R_{\mu\mu} > 0.3$. Fig. 3.4 shows the isolation efficiency measured for *Medium* muons in data and simulation as a function of the muon p_T for the *FixedCutLoose* working points, with the respective data/MC ratios included in the bottom panel. The agreement between data and MC is at few percent level, less than 2%.

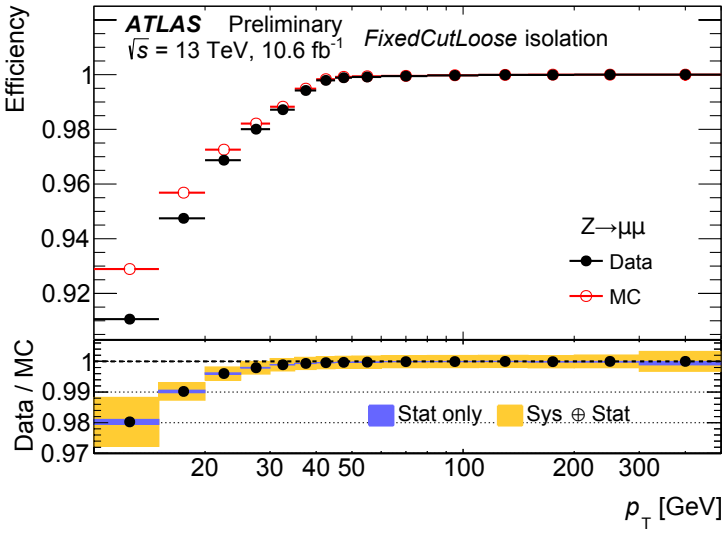


Fig. 3.4: Isolation efficiency for the *FixedCutLoose* muon isolation working points. The efficiency is shown as a function of the muon transverse momentum p_T and is measured in $Z \to \mu^+\mu^-$ events. The full (empty) markers indicate the efficiency measured in data (MC) samples. The errors shown on the efficiency are statistical only. The bottom panel shows the ratio of the efficiency measured in data and simulation, as well as the statistical uncertainties and combination of statistical and systematic uncertainties. The plot is taken from Ref. [128].

3.3.2 Electrons

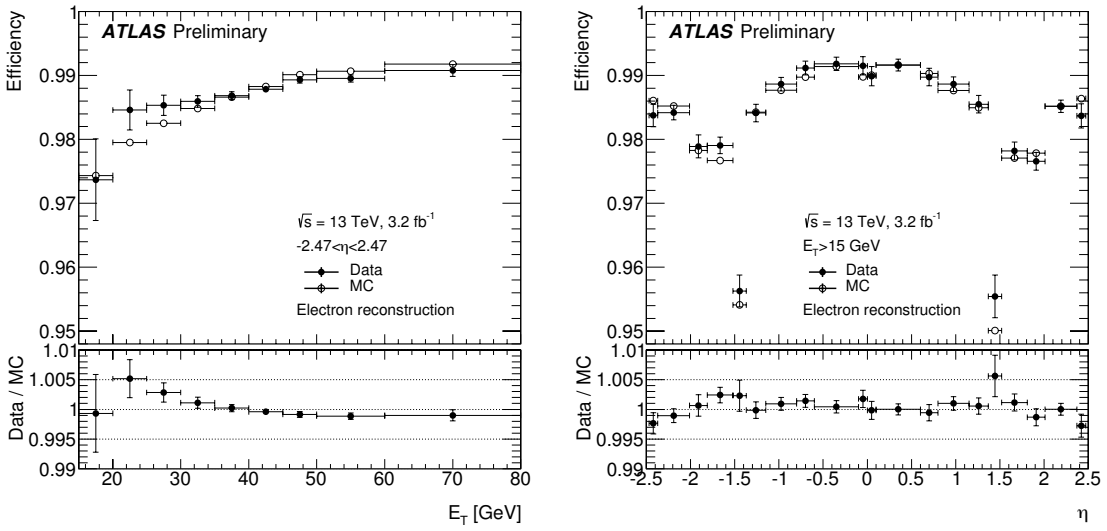
Electron reconstruction method makes use of a combination of the information coming from the ECAL [131] and the ID tracking system. An excellent electron identification is crucial to reject the large backgrounds typically originating from jets leading to fake electrons. The first step, referred to as *cluster reconstruction*, is the identification of energy clusters deposited in the EM calorimeter; these clusters

are reconstructed using a dedicated *sliding-window* algorithm [132]: this algorithm looks at calorimeter cells within a rectangular window of 3×5 in units of 0.025×0.025 in the $\eta \times \phi$ space. This size is fixed and it is usually referred to as "a tower". The energy of each tower is obtained by summing the energy deposit found in all the longitudinal layers of region in the $\eta \times \phi$ space. After the selection of a local maximum in the transverse energy ($E_T > 2.5$ GeV), this energy deposit in treated as a seed cluster, which has to match with one or more tracks coming from the reconstruction in the ID; this procedure of track matching takes account of radiative energy losses of high-energy electrons, which can change the path of an electron and, consequently, modify the track.
The track fitting and extrapolation is done with the so-called *Gaussian Sum Filter* (GSF) algorithm, which takes care of bremsstrahlung energy losses and experimental noise [133]. This algorithm is based on the following idea: the trajectory of an electron can be approximated as a weighted sum of Gaussian functions; this means that the GSF divides the experimental noise and bremsstrahlung energy losses into individual Gaussian components and processes each one of them using a Kalman Filter algorithm [134]. The GSF algorithm finds all the electron candidate tracks (with $p_T > 400$ MeV and $|\eta| < 2.5$), which are fitted again: the tracks obtained by this fit procedure are used to compute electron 4-momentum and match the clusters found in the EM calorimeter (if a energy cluster is matched to a reconstructed track having transverse momentum greater than 0.5 GeV, then an electron candidate is formed).
For Run-2 analyses, the electron measurements are performed by requiring the track associated with the electron to be compatible with the primary interaction vertex of the hard collision, in order to reduce the background from conversions and secondary particles. The track parameters are calculated in a reference frame where the z-axis is taken along the measured beam-line position. The following conditions are also applied: $d_0/\sigma_{d_0} < 5$ and $|\Delta z_0 \sin\theta < 0.5$ mm, where the impact parameter d_0 is the distance of closest approach of the track to the measured beam-line, z_0 is the distance along the beam-line between the point where d_0 is measured and the beam-spot position, θ is the polar angle of the track and σ_{d_0} represents the estimated uncertainty of the d_0 parameter. To assess the compatibility with the primary vertex of the hard collision the Δz_0 between the track and the primary vertex is employed. This vertex is selected from the reconstructed primary vertices (compatible with the beam-line) as the one with the highest sum of transverse momenta of the associated tracks. The efficiency of these requirements in data and MC is estimated together with the efficiency of the various identification operating points.

Electron reconstruction efficiency

The electron reconstruction and trigger efficiencies are measured with the so called *tag-and-probe* method [135] based on $Z \to e^+e^-$, $W \to e\nu$ and $J/\Psi \to e^+e^-$ data and simulated samples: one lepton is selected using the standard cuts (tag), while for the second candidate (probe) a looser set of cuts is applied. The electron identification efficiencies [136] are derived from combined measurements using $Z \to ee$ and $W \to e\nu$ events. Additionally, the energy is corrected for

non-uniformity of the detector response and by using a multivariate regression algorithm trained on MC simulation to correct for loss of energy outside the calorimeter or the clusters. To account for small differences of the calibration between simulation and data, both the energy scale (in data) and resolution (in Monte Carlo simulation) are adjusted based on corrections measured in $Z \to e^+e^-$ events. The resolution is depending on the pseudorapidity and energy of the electron. In the central η < 0.5 region it ranges from 2.5% for 25 GeV to 0.5% for 1 TeV, for larger η it rises up to 10% for low energy electrons [137]. These values have been determined for Run 1 data, but are still approximately valid, as discussed in Ref. [136]. Only small modifications have been made to correct for changes in conditions between Run 1 and Run 2. For most of the pseudorapidity range the reconstruction efficiency is above 97% and increases with the energy of the reconstructed electron. The electron reconstruction efficiency measured in $Z \to e^+e^-$ events of 2015 data, as a function of the transverse energy of the electron cluster (E_T) and the corresponding pseudorapidity η is shown in Fig. 3.5. It is



(a) (b)

Fig. 3.5: Electron reconstruction efficiencies in $Z \to e^+e^-$ events as a function of (a) the transverse energy E_T (integrated over the full pseudorapidity range η) and (b) η (for 15 < E_T 50 GeV). The plots are taken from Ref. [136].

visible that the reconstruction efficiency drops in the so-called *crack region* (1.37 < $|\eta|$ < 1.52), where the barrel and end-cap calorimeters meet.

Electron Identification

After the electron reconstruction, using EM calorimeter and tracking system information, a cut-based approach is used to identify the electrons in the $|\eta| < 2.47$ region (the transition region between the barrel and endcap calorimeters, i.e $1.37 < |\eta| < 1.52$, is excluded). This step is called *electron identification*.
Three different selections, optimized in bins of η and E_T, are defined, according to

the quality of the reconstruction, providing different electron efficiency versus jet rejection working points: loose, medium and tight.

- **Loose**: The loose identification criteria uses only information from the calorimeters. Cuts are applied on the hadronic leakage (the ratio of the E_T in the first compartment of the HCAL and in the whole ECAL) and on shower shape variables (lateral shower shape and width) based on the second layer of the ECAL. This set of cuts provides excellent identification efficiency but low background rejection.

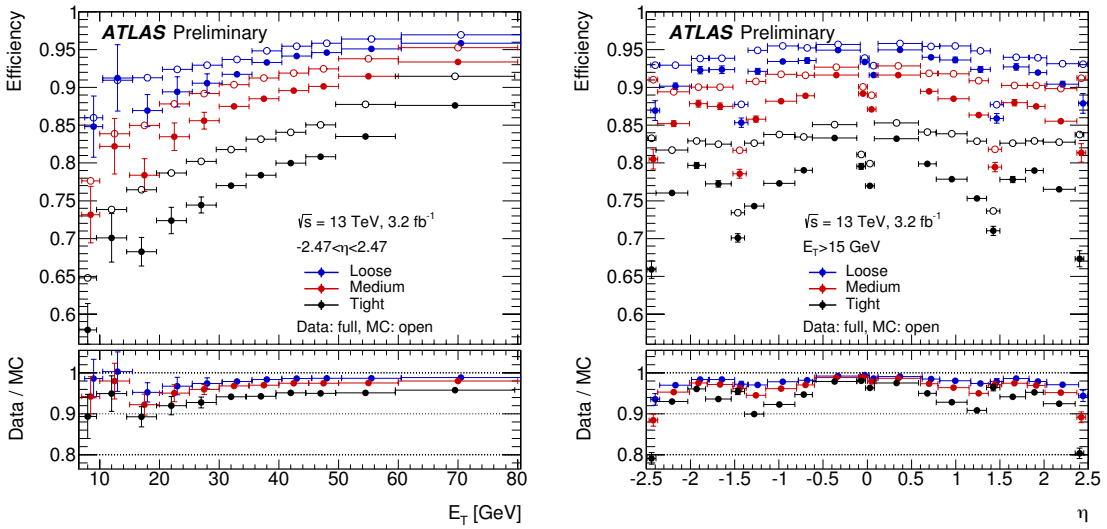
- **Medium**: This selection improves the quality by adding cuts on both the ECAL and the tracking variables. The fine longitudinal segmentation of the first layer of the ECAL is exploited, in particular to reject photons from neutral pion decays, which result into very close deposits of energy; this selection increases the jet rejection by a factor 3-4 with respect to the loose identification and reduces the selection efficiency by $\approx 10\%$.

- **Tight**: This set of cuts makes use of all the particle identification tools; further cuts are applied on the tracking variables and an additional energy (and momentum) isolation cut is applied to the cluster using all cell energies within a cone of $\Delta R < 0.2$ (and 0.3) around the electron candidate. This selection provides the highest isolated electron identification and rejection against jets.

As for the muons, some corrections are applied to MC simulation in form of efficiency scale factors, in order to match isolation and trigger efficiency in data. These corrections are obtained by comparing MC predictions to $J/\psi \to e^+e^-$ and $Z \to e^+e^-$ data samples using the tag-and-probe method. In Fig. 3.6 the combined electron reconstruction and identification efficiencies in $Z \to e^+e^-$ events as a function of the transverse energy E_T, and as a function of pseudorapidity η are shown.

Energy scale and resolution

The electron energy scale and the inter-calibration of the different parts of the EM calorimeter are obtained from $Z \to e^+e^-$, $J/\Psi \to e^+e^-$ and $W^\pm \to e^\pm\nu$ (as shown in Fig. 3.7 for $Z \to e^+e^-$ events). The relative alignment of the calorimeter components with respect to the inner detector has been measured using electron candidates, compatible with the ones coming from *W* and *Z* decays. Fig. 3.7 shows the comparison of the invariant mass distribution for $Z \to ee$ candidates between data and simulation after the energy resolution correction has been applied to the data. Within the uncertainties in the calibration and resolution correction factors, the data and simulation are in fair agreement.



(a) (b)

Fig. 3.6: Electron identification efficiencies in $Z \to e^+e^-$ events as a fucntion of (a) the transverse energy E_T (integrated over the full pseudorapidity range η) and (b) η (for 15 < E_T 50 GeV). The data efficiencies are obtained from the data-to-MC efficiency ratios measured using $J/\psi \to e^+e^-$ and $Z \to e^+e^-$ tag-and-probe method, multiplied by the MC prediction for electrons from $Z \to e^+e^-$ decay. The plots are taken from Ref. [136].

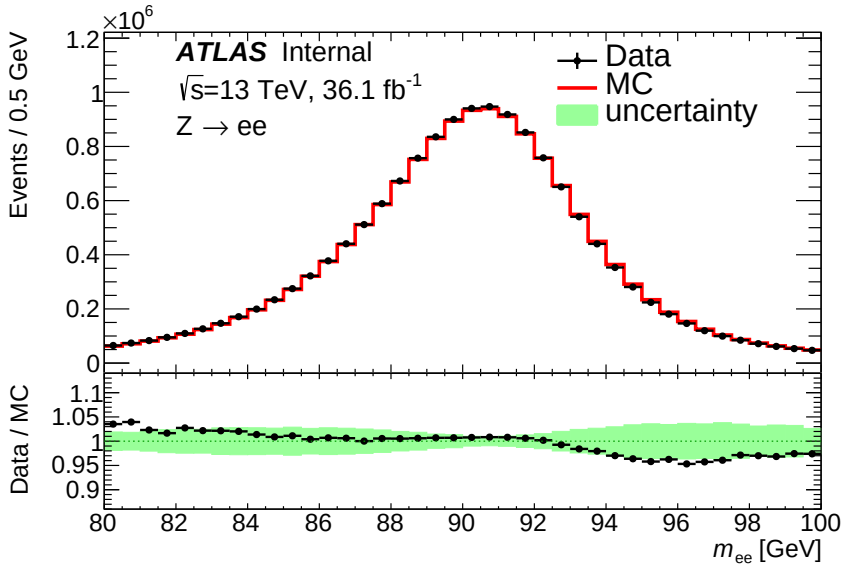


Fig. 3.7: Comparison of the invariant mass distribution of the two electrons in selected $Z \to ee$ candidates, after the calibration and resolution corrections are applied. The total number of events in the simulation is normalized to the data. The ratio is shown in the bottom plot. The uncertainty band of the bottom plot corresponds to the impact of the uncertainties on the calibration and resolution correction factors. This plot is taken from Ref. [138].

Chapter 4

Photon PDF determination

This study has been produced in collaboration between myself and some members of the xFitter Collaboration [139]. In this analysis, I was the main analyser and participated in the weekly/bi-weekly xFitter Developers meetings, where I gave regular updates between the beginning of September 2016 and the end of January 2017. This work resulted in the following publication: "The photon PDF from high-mass Drell-Yan data at the LHC" [140]. During the course of my PhD, I had impact in other publications within this collaboration, namely: "Impact of the heavy quark matching scales in PDF fits" [141] and "Impact of low-x resummation on QCD analysis of HERA data" [142], both published on EPJC.

4.1 Preface

Achieving the highest precision for theoretical predictions at the LHC requires the calculation of hard-scattering cross-sections that include perturbative QCD corrections up to (N)NNLO and electroweak (EW) corrections up to NLO. Parton distribution functions (PDFs) need to be provided with matching accuracy, which in the case of QED effects involves introducing the photon parton distribution function of the proton, $x\gamma(x, Q^2)$. In this chapter, a determination of the photon PDF from fits to recent ATLAS measurements of high-mass Drell-Yan (DY) di-lepton production at $\sqrt{s}$ = 8 TeV is presented [143]. The analysis reported in the following chapter of my thesis is based on the xFitter framework [139] and has required improvements both in the APFEL [144] program, to account for NLO QED effects, and in the aMCfast [145] interface to account for the photon-initiated (PI) contributions in the EW calculations within MadGraph5_aMC@NLO [146].

4.2 Introduction

Precision phenomenology at the LHC requires theoretical calculations which include not only QCD corrections, where NNLO is rapidly becoming the standard, but also EW corrections, which are particularly significant for observables directly sensitive to the TeV region, where EW Sudakov logarithms are enhanced. An

important ingredient of these electroweak corrections is the photon PDF of the proton, $x\gamma(x, Q^2)$, which must be introduced to absorb the collinear divergences arising in initial-state QED emissions.
The first PDF fit to include both QED corrections and a photon PDF was MRST2004 QED [48], where the photon PDF was taken from a model and tested on HERA data for direct photon production [147]. Almost 10 years later, the NNPDF2.3 QED analysis [148, 149], provided a first model-independent determination of the photon PDF based on DY data from the LHC. The resulting photon PDF was however affected by large uncertainties due to the limited sensitivity of the data used as input to that fit. The determination of $x\gamma(x, Q^2)$ from NNPDF23 QED was later combined with the state-of-the-art quark and gluon PDFs from NNPDF3.0, together with an improved QED evolution, to construct the NNPDF3.0 QED set [49, 50]. The CT group has also released a QED fit using a similar strategy as the MRST2004 QED, named the CT14 QED set [51].
A recent breakthrough concerning the determination of the photon content of the proton has been the realization that $x\gamma(x, Q^2)$ can be calculated in terms of inclusive deep-inelastic scattering (DIS) structure functions. The photon PDF from this strategy is called LUXqed [52, 53] and its residual uncertainties are now at the few percent level, not too different from those of the quark and gluon PDFs. A related approach by HKR [54], denoted by HKR16 in the following, also leads to a similar photon PDF as compared to the LUXqed calculation, although no estimate for the associated uncertainties is provided in this case.
In this chapter, I present a direct determination of the photon PDF from recent high-mass DY data collected at the LHC with the ATLAS detector at $\sqrt{s}$ = 8 TeV and I show comparison between this prediction and some existing determinations of the above-described $x\gamma(x, Q^2)$. In the ATLAS publication on Drell-Yan data [143] the Bayesian reweighting method [150, 151] was applied to NNPDF2.3QED, showing that these data can provide significant information on $x\gamma(x, Q^2)$. The goal of this study is therefore to investigate further these constraints from the ATLAS high-mass DY measurements on the photon PDF, this time by means of a direct PDF fit performed within the open-source `xFitter` framework. State-of-the-art theoretical calculations are employed, in particular the inclusion of NNLO QCD and NLO QED corrections to the PDF evolution and the computation of the DIS structure functions as implemented in the `APFEL` program. The implementation of NLO QED effects in `APFEL` is presented here for the first time. The inclusion of NLO QED evolution effects is cross checked using the independent `QEDEVOL` code [152] based on the `QCDNUM` evolution program [153]. The resulting determination of $x\gamma(x, Q^2)$ represents an important validation test of recent developments in theory and data concerning our understanding of the nature and implications of the photon PDF.

4.3 Data and theory

The DY process of lepton pair production in hadronic interactions, $Z/\gamma^* \to \ell^+\ell^-$, is a powerful tool in understanding the nature of partonic interactions and of hadronic structure in detail.

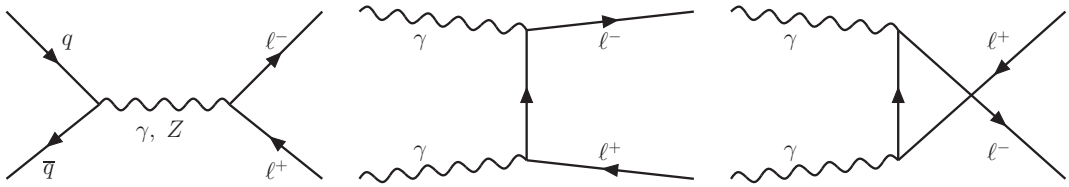


Fig. 4.1: Diagrams that contribute to lepton-pair production at hadron colliders at the Born level.

Here, the photon content of the proton $x\gamma(x, Q^2)$ is extracted from a PDF analysis based on the combined inclusive DIS cross-section data from HERA [34] supplemented by the ATLAS measurements of high-mass DY differential cross sections at $\sqrt{s} = 8$ TeV. The HERA inclusive data are the backbone of modern PDF fits, providing information on the quark and gluon content of the proton, while the high-mass DY data provide a direct sensitivity to the photon PDF.
Two processes contribute to opposite sign, same family, di-lepton production at the LHC: the quark-antiquark s-channel scattering, and from photon-photon t- and u-channel scattering mediated by a lepton. The Feynman diagrams of the above-described processes are illustrated in Fig. 4.1. The ATLAS high-mass Drell-Yan 8 TeV measurements are presented in terms of both the single-differential (1D) invariant-mass distribution, $d\sigma/dm_{\ell\ell}$, as well as double-differential (2D) distributions in $m_{\ell\ell}$ and $y_{\ell\ell}$, namely $d^2\sigma/dm_{\ell\ell}d|y_{\ell\ell}|$, and in $m_{\ell\ell}$ and $\Delta\eta_{\ell\ell}$, $d^2\sigma/dm_{\ell\ell}\Delta\eta_{\ell\ell}$, where $m_{\ell\ell}$ and $y_{\ell\ell}$ represent the invariant mass and the rapidity of the di-lepton final state respectively and $\Delta\eta_{\ell\ell}$ is the difference in pseudorapidity between leading and sub-leading muon. For the invariant-mass 1D distribution, there are 12 bins between $m_{\ell\ell} = 116$ GeV and 1.5 TeV; and for both double-differential distributions, there are five different bins in invariant mass, from the lowest bin with 116 GeV < $m_{\ell\ell}$ < 150 GeV to the highest bin with 500 GeV < $m_{\ell\ell}$ < 1500 GeV. The first three (last two) $m_{\ell\ell}$ bins of the 2D distributions are divided into 12 (6) bins with fixed width, extending up to 2.4 and 3.0 for the $|y_{\ell\ell}|$ and $|\Delta\eta_{\ell\ell}|$ distributions, respectively. The photons which undergo hard scattering in the $\gamma\gamma \to \ell\ell$ process from Fig. 4.1 can be produced by either emission from the proton as a whole (the "elastic" component) or radiated by the constituent quarks (the "inelastic" component). From the theory point of view, the photon PDF extracted from the fit is by construction the sum of the elastic and inelastic contributions, though this analysis is mostly sensitive to the latter. The NLO high-mass DY cross sections have been simulated with `MadGraph5_aMC@NLO` (version 2.4.3), which includes the contribution from photon-initiated diagrams, interfaced to `APPLgrid` (version 01-04-70) through `aMCfast` (version 01-03-00). A tailored release of `APPLgrid` has been used, to account for the contribution of the photon-initiated processes [1]. Both the contributions are generated in the 5-flavours number scheme ($n_f = 5$), where all the quarks, except for the *top* quark, are treated as massless quarks; all the calculations are performed at parton-level fixed-order (FO) without parton shower. Furthermore, mass effects of charm and bottom quarks in the matrix elements, as appropriate

[1] Modified version of `APPLgrid` available at: `https://github.com/scarrazza/APPLgridphoton`

for a high-scale process, have been neglected.
These NLO theoretical predictions are generated using the same selections as in reference [143], namely:

- the invariant mass of the lepton pair is required to be greater than 116 GeV;
- the absolute value of the pseudorapidity of each lepton is required to be less than 2.5;
- the transverse momentum (p_T) of the leading lepton has to be greater than 40 GeV;
- the p_T of the sub-leading lepton has to be greater than 30 GeV.

In modern PDF fitting, many cross-sections which are used cannot be computed analytically in a sufficient short time for input to an iterative PDF fit thus interpolation grid techniques are used. Running `MadGraph5_aMC@NLO` to produce interpolation grids in the `APPLgrid` format using the `aMCfast` interface requires two main steps. The first step consists of setting up and optimizing the interpolation grids that will be subsequently filled up in the second step. As already mentioned above, the production of grids is directly connected to a fixed-order analysis in the `MadGraph5_aMC@NLO` framework. Practically, this is done by doing a preliminary "low statistics" run that allows the code to optimize the interpolation grids based on the particular observables defined in the analysis file. Then, after that this optimisation at low accuracy has been performed, it is possible to proceed to a higher accuracy determination The specific parameters used in the `MadGraph5_aMC@NLO` run card for generating the `APPLgrid` predictions are listed below:

- True = reweight_scale (in order to get scale dependence - scaled varied up and down by a factor of 2);
- True = reweight_PDF (in order to get PDF uncertainty) - reweighting performed using `LHAPDF` (version 6.1.6);
- 29.0 = ptl (minimum lepton transverse momentum p_T);
- 2.6 = etal (maximum lepton |pseudorapidity| $|\eta|$);
- 111.0 = mll (minimum invariant mass of all opposite sign lepton pairs);
- 30.0 = mll_sf (minimum invariant mass of all opposite sign same-flavour lepton pairs);
- 5 = maxjetflavor (5-flavours number scheme).

As discussed below, the NLO calculations are then supplemented by K-factors obtained from `FEWZ` [154] to match higher order (NNLO) calculations. The NLO EW corrections to the DY processes are also estimated using `FEWZ`. The photon-initiated process is taken at LO since this corresponds to the `APPLgrid` implementation and the NLO corrections are very small compared to the data accuracy. The parameters of the input grids are set as follows:

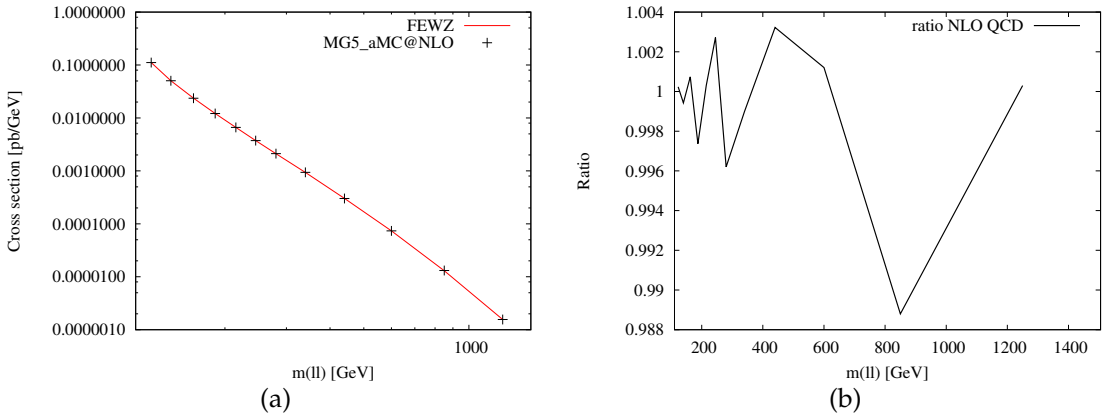


Fig. 4.2: (a) Comparison regarding the cross section as a function of the invariant mass of the lepton pair between NLO QCD + LO EW predictions evaluated with `MG5_aMC@NLO` interfaced with `APPLgrid` and `aMCfast` (black marker) and the ones computed with `FEWZ` (red solid line); (b) ratio between the two different predictions.

- the minimum value of the Q^2 is set to the (lower invariant mass bin edge - 5 GeV)^2;
- the maximum value of the Q^2 is set to the (higher invariant mass bin edge + 5 GeV)^2;
- the minimum value of x in the `APPLgrids` is 10^{-5};
- the maximum value of x in the `APPLgrids` is 1;
- the number of nodes in the Q^2-space is set to 10 (as regards the LO PI contribution, as well as the one-dimensional distribution, this quantity is set to 70);
- the interpolation order in Q^2 is equal to 3;
- the number of nodes in the x-space is set to 30 (as regards the LO PI contribution, as well as the one-dimensional distribution, this quantity is set to 50);
- the interpolation order in x is equal to 3;

The number of nodes in both the Q^2- and x-space is set higher for the PI process mainly because this is a sub-dominant contribution with respect to the $q\bar{q}$-initiated so we want to ensure the same accuracy in the interpolation.
Furthermore, to gain in statistics in the highest invariant mass intervals, the cut on $m_{\ell\ell}$ at the generation level has been optimised and set to the (lower invariant mass bin edge - 5 GeV).

The cross section predictions are evaluated in two different methods: first of all, with a standalone reader of the `APPLgrids` and then within the `xFitter` framework (trunk version[2] - dedicated branch of xFitter that is linked to the adjusted

[2]https://gitlab.cern.ch/fitters/xfitter

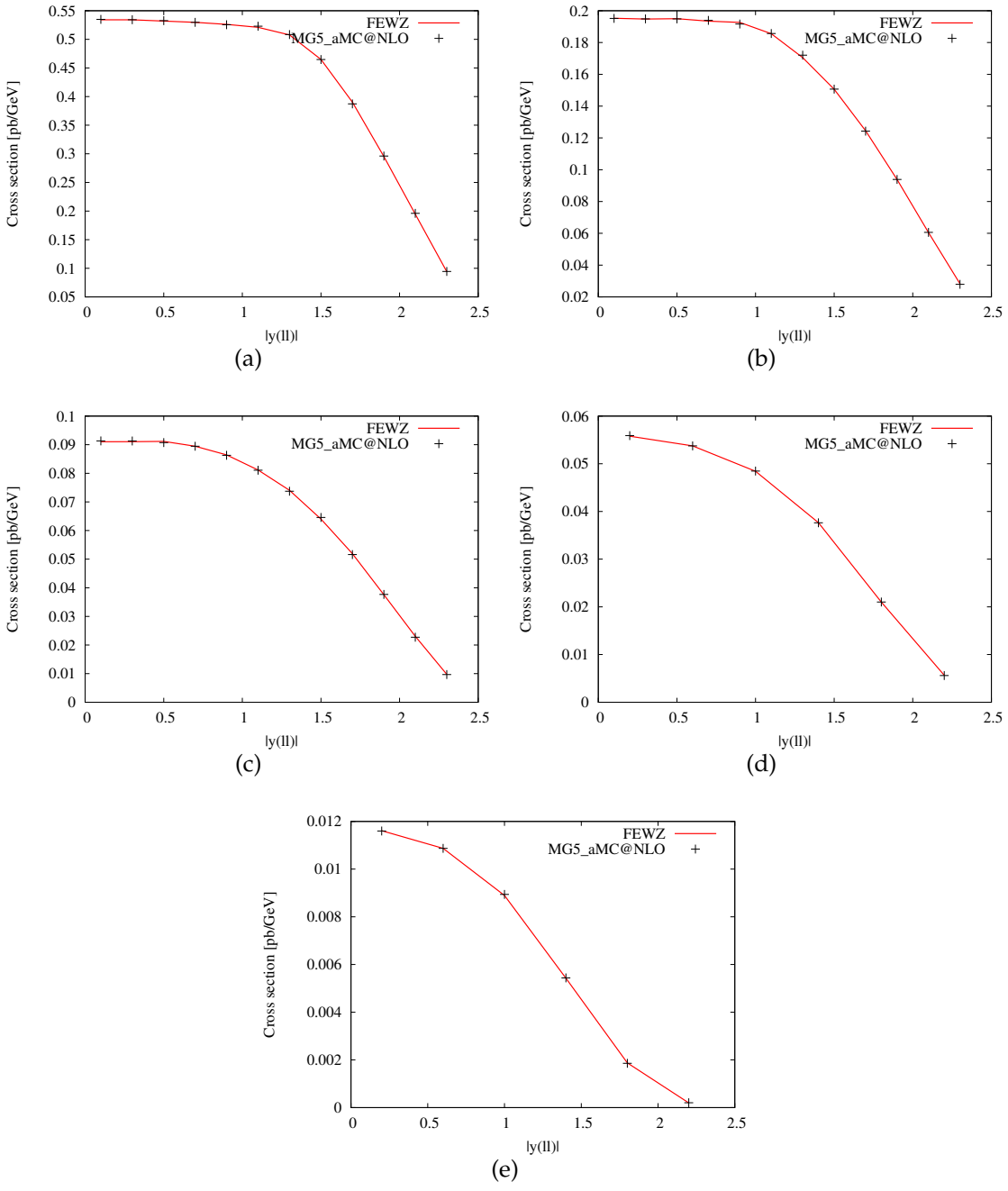


Fig. 4.3: Comparison regarding the cross section as a function of the absolute value of the rapidity of the lepton pair between NLO QCD + LO EW predictions evaluated with `MG5_aMC@NLO` interfaced with `APPLgrid` and `aMCfast` (black marker) and the ones computed with `FEWZ` (red solid line) for different invariant mass intervals: (a) 116 GeV < $m_{\ell\ell}$ < 150 GeV; (b) 150 GeV < $m_{\ell\ell}$ < 200 GeV; (c) 200 GeV < $m_{\ell\ell}$ < 300 GeV; (d) 300 GeV < $m_{\ell\ell}$ < 500 GeV; (e) 500 GeV < $m_{\ell\ell}$ < 1.5 TeV.

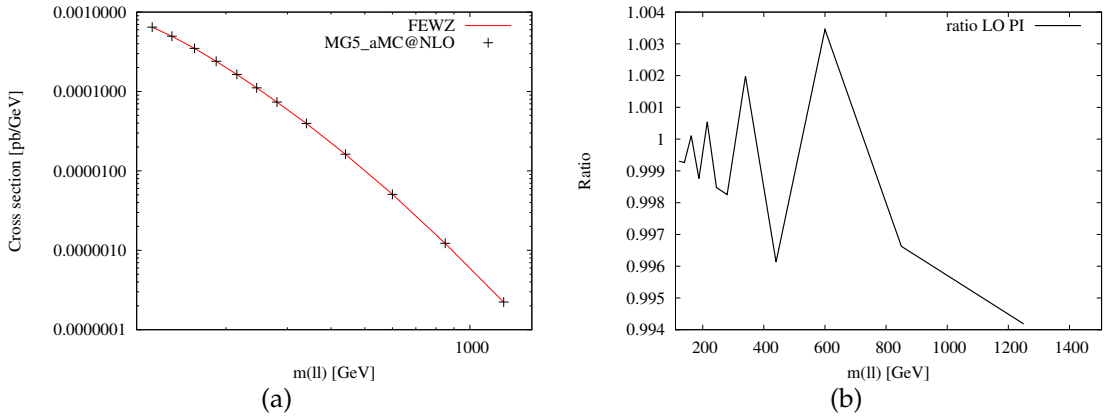


Fig. 4.4: (a) Comparison regarding the cross section as a function of the invariant mass of the lepton pair between LO PI predictions evaluated with `MG5_aMC@NLO` interfaced with `APPLgrid`s and `aMCfast` (black marker) and the ones computed with `FEWZ` (red solid line); (b) ratio between the two different predictions.

`APPLgrid`: PI_apfel_for_lhaGridQED). The difference between the two predictions is at most 1%, for both the 1-dimensional and the 2-dimensional distributions, as shown in Fig. 4.2 (for the $d\sigma/dm_{\ell\ell}$ distribution) and in Fig. 4.3 (for the double differential cross sections - for these only rapidity distributions are shown, since they are the most sensitive one to the photon contribution). The same comparison has been made for the LO PI predictions shown in Fig. 4.4 and Fig. 4.5.
The DIS structure functions and PDF evolution are computed with the `APFEL` program [144], which is currently accurate up to NNLO in QCD and NLO in QED, including the relevant mixed QCD+QED corrections. This means that, on top of the pure QCD contributions, the DGLAP evolution equations [32, 33, 155] are solved including the $\mathcal{O}(\alpha_s\alpha)$ and $\mathcal{O}(\alpha^2)$ corrections to the splitting functions. Corrections of $\mathcal{O}(\alpha)$ are also included leading to a (weak) explicit dependence of the predictions on the photon PDF. Heavy-quark (charm and bottom) mass effects to DIS structure functions are taken into account using the FONLL-B (C) general-mass scheme [156] for the NLO (NNLO) fits. The numerical values of the heavy-quark masses in the mass parameter scheme are taken to be $m_c = 1.47$ GeV and $m_b = 4.5$ GeV as determined in [34], consistent with the latest PDG averages [157]. The reference values of the QCD and QED coupling constants are chosen to be $\alpha_s(m_Z) = 0.118$ and $\alpha(m_\tau = 1.777\text{ GeV}) = 1/133.4$, again consistent with the PDG recommended values. In the calculation of the Drell-Yan cross section, dynamical renormalisation μ_R and factorisation μ_F scales are used, which are set equal to the scale of invariant mass $m_{\ell\ell}$, both for the quark- and gluon-induced and for the photon-induced contributions.
The choice of other values for these scales in the QED diagrams, such as a fixed scale $\mu_R = \mu_F = M_Z$, leads to variations of the photon-initiated cross-sections of a few percent at most. The choice of the scale for the photon PDF is further discussed in [158, 159].
For the kinematics of the ATLAS DY data, the ratio between the photon-initiated contributions and quark- and gluon-induced di-lepton production is largest for

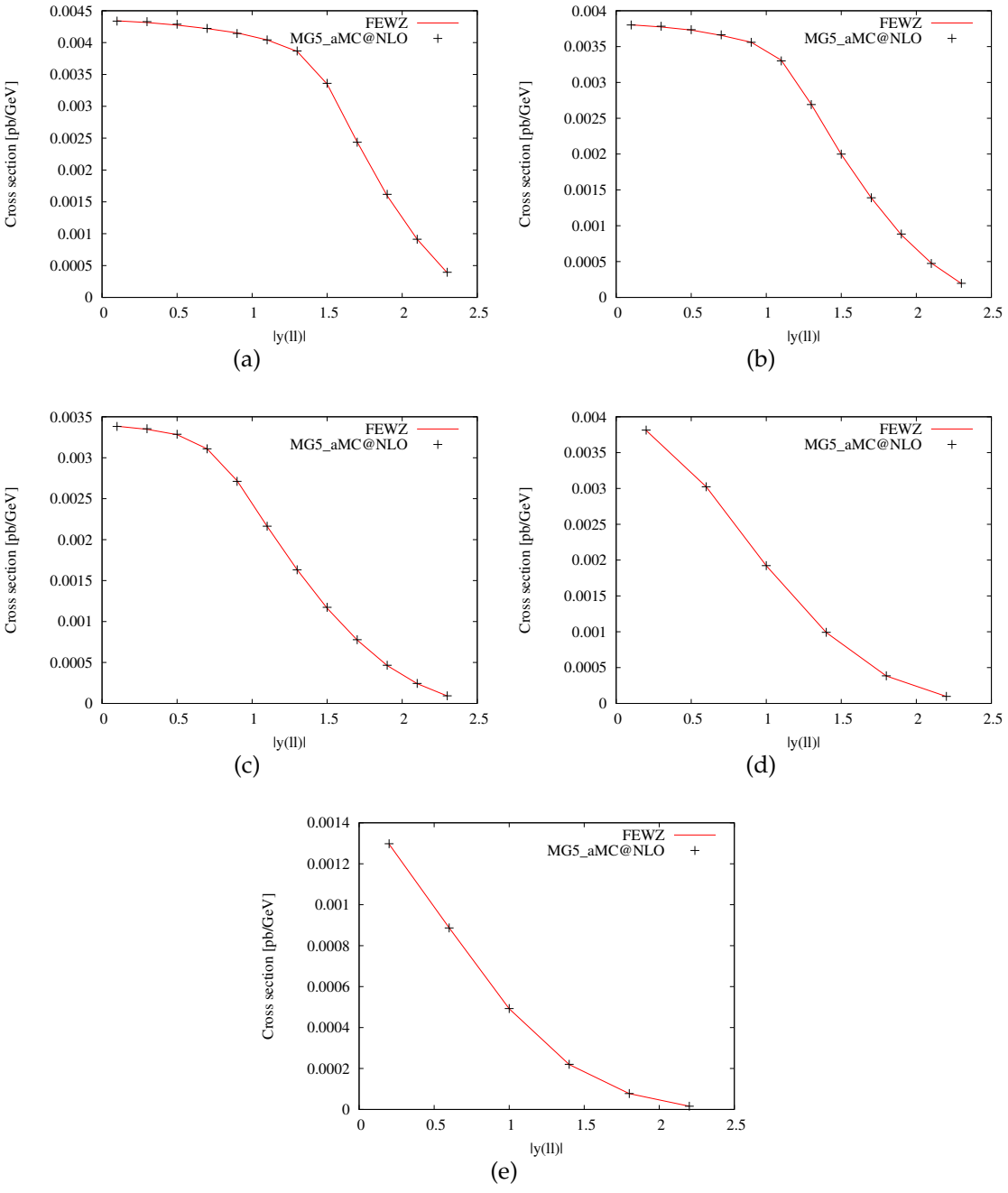


Fig. 4.5: Comparison regarding the cross section as a function of the absolute value of the rapidity of the lepton pair between LO PI predictions evaluated with `MG5_aMC@NLO` interfaced with `APPLgrid` and `aMCfast` (black marker) and the ones computed with `FEWZ` (red solid line) for different invariant mass intervals: (a) 116 GeV < $m_{\ell\ell}$ < 150 GeV; (b) 150 GeV < $m_{\ell\ell}$ < 200 GeV; (c) 200 GeV < $m_{\ell\ell}$ < 300 GeV; (d) 300 GeV < $m_{\ell\ell}$ < 500 GeV; (e) 500 GeV < $m_{\ell\ell}$ < 1.5 TeV.

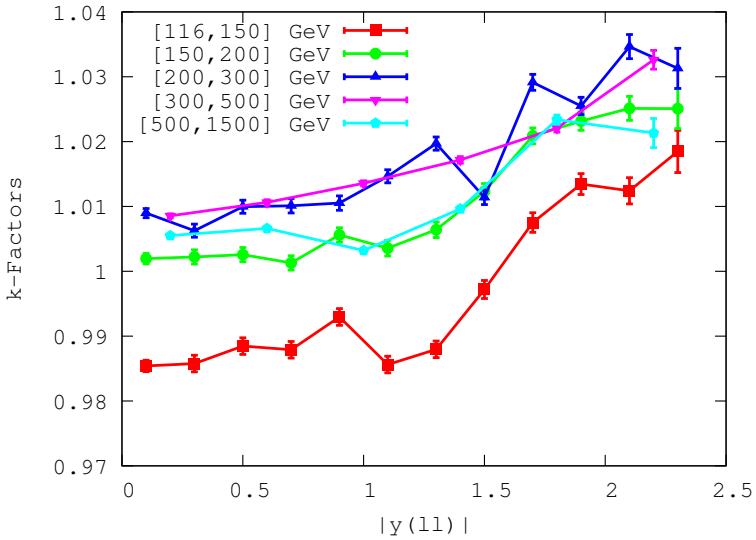


Fig. 4.6: The NNLO/NLO K-factors, defined in Eq. (4.1), that account for higher order QCD and EW effects to the high-mass DY cross sections with the photon induced contribution subtracted, as a function of the di-lepton rapidity $|y_{\ell\ell}|$. Each set of points corresponds to a different bin in the di-lepton invariant mass $m_{\ell\ell}$.

central rapidities and large invariant masses. For the most central (forward) rapidity bin, $0 < |y_{\ell\ell}| < 0.2$ ($2.0 < |y_{\ell\ell}| < 2.4$), the ratio between the QED and QCD contributions varies between 2.5% (2%) at low invariant masses and 12% (2.5%) for the highest $m_{\ell\ell}$ bin. Therefore, data from the central region will exhibit the highest sensitivity to $x\gamma(x, Q^2)$.
The `MadGraph5_aMC@NLO` NLO QCD and LO QED calculations used in this work have been benchmarked against the corresponding predictions obtained with the `FEWZ` code, finding agreement within statistical uncertainties of the predictions for both the 1D and the 2D distributions. In order to achieve NNLO QCD and NLO EW accuracy in our theoretical calculations, the NLO QCD and LO QED cross-sections computed with `MadGraph5_aMC@NLO` have been supplemented by bin-by-bin K-factors defined as:

$$K(m_{\ell\ell}, |y_{\ell\ell}|) \equiv \frac{\text{NNLO QCD} + \text{NLO EW}}{\text{NLO QCD} + \text{LO EW}}\,, \tag{4.1}$$

using the MMHT2014 NNLO [160] PDF set both in the numerator and in the denominator. The K-factors have been computed using `FEWZ` with the same settings and analysis cuts as the corresponding NLO calculations of `MadGraph5_aMC@NLO`. This approximation is justified since NNLO K-factors depend very mildly on the input PDF set, see for example [161]. The photon induced contribution, as provided in [143], has been explicitly subtracted from the `FEWZ` predictions. Fig. 4.6 shows the K-factors of Eq. (4.1) corresponding to the double differential $(m_{\ell\ell}, |y_{\ell\ell}|)$ cross sections as a function of the di-lepton rapidity $|y_{\ell\ell}|$, where each set of points corresponds to a different di-lepton invariant mass $m_{\ell\ell}$ bin. The vertical bars on each point represent the statistical uncertainty on the calculated K-factors.
The K-factors vary between 0.98 and 1.04, highlighting the fact that higher-order

corrections to the Drell-Yan process are moderate, in particular at low values of $m_{\ell\ell}$ and in the central region. Even at forward rapidities, the K-factors modify the NLO result by at most 4%.

4.4 Fit settings

In order to make a full PDF fit, the ATLAS DY data are fitted together with the final combined inclusive cross section data from HERA. The DGLAP equations yield the PDFs at all scales if they are input as functions of x at a starting scale Q_0^2, which should be large enough that perturbative QCD can be assumed to be valid. For the present analysis this value is chosen to be $Q_0^2 = 7.5$ GeV^2. This is also the value chosen for the minimum value of Q^2 for data entering the fit. The value of $\alpha_s(M_Z)$ is chosen to be $\alpha_s(M_Z) = 0.118$.
The charm PDF is then generated perturbatively from quarks and gluons by means of DGLAP evolution, exploiting recent developments in `APFEL` which allow the setting of heavy-quark thresholds μ_h differently from the heavy quark masses m_h, such that $\mu_c = Q_0 > m_c$. Hence a high threshold can be used without having to parametrise the charm PDF [141].
The expression for the χ^2 function used for the fits is that of Ref. [162], which includes corrections for possible biases from statistical fluctuations and treats the systematic uncertainties multiplicatively, and it is also shown in Eq. 4.2:

$$\chi^2 = \sum_i \frac{\left[D_i - T_i\left(1 - \sum_j \gamma_j^i b_j\right)\right]^2}{\delta_{i,\mathrm{unc}}^2 T_i^2 + \delta_{i,\mathrm{stat}}^2 D_i T_i} + \sum_j b_j^2 \sum_i + \ln \frac{\delta_{i,\mathrm{unc}}^2 T_i^2 + \delta_{i,\mathrm{stat}}^2 D_i T_i}{\delta_{i,\mathrm{unc}}^2 D_i^2 + \delta_{i,\mathrm{stat}}^2 D_i^2}, \quad (4.2)$$

where T_i is the theoretical prediction and D_i the measured value of the i-th data point, $\delta_{i,\mathrm{stat}}$, $\delta_{i,\mathrm{unc}}$, and γ_j^i are the relative statistical, uncorrelated systematic, and correlated systematic uncertainties, and b_j are the nuisance parameters associated to the correlated systematics which are determined during the fit.
Alternative forms that do not include these corrections, such as those defined in [24, 34], have also been studied but no significant differences in the results have been observed.
In this analysis, the parametrised PDFs are the valence distributions $xu_v(x, Q_0^2)$ and $xd_v(x, Q_0^2)$, the gluon distribution $xg(x, Q_0^2)$, and the *u*-type and *d*-type sea-quark distributions, $x\bar{U}(x, Q_0^2)$, $x\bar{D}(x, Q_0^2)$, where $x\bar{U}(x, Q_0^2) = x\bar{u}(x, Q_0^2)$ and $x\bar{D}(x, Q_0^2) = x\bar{d}(x, Q_0^2) + x\bar{s}(x, Q_0^2)$. The photon distribution $x\gamma(x, Q_0^2)$ is also parametrised at the starting scale. The following general functional form is adopted:

$$xf(x) = Ax^B(1-x)^C(1 + Dx + Ex^2) \quad (4.3)$$

where some of the normalisation parameters, in particular A_{u_v}, A_{d_v} and A_g, are constrained by the valence and momentum sum rules. Note that the photon PDF also enters the momentum sum rule. The parameters $B_{\bar{U}}$ and $B_{\bar{D}}$ are set equal to each other, so that the two quark sea distributions share a common small-x behaviour. Since the measurements used here are not sensitive to the strangeness content of the proton, strangeness is fixed such that $x\bar{s}(x, Q_0^2) = r_s x\bar{d}(x, Q_0^2)$, where

the strangeness fraction $r_s = 1.0$ is consistent with the ATLAS analysis of inclusive W and Z production [163,164]. The further constraint $A_{\bar{U}} = 0.5A_{\bar{D}}$ is imposed, such that $x\bar{u}(x, Q_0^2) \to x\bar{d}(x, Q_0^2)$ as $x \to 0$. The explicit form of PDF parametrisation Eq. (4.3) at the scale Q_0^2 is determined by the technique of saturation of the χ^2, namely the number of parameters is increased one by one until the χ^2 does not improve further, employing Wilks' theorem [165]. Following this method, the optimal parametrisation for the quark and gluon PDFs found for this analysis is:

$$\begin{aligned} xu_v(x) &= A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} (1 + E_{u_v} x^2)\,, \\ xd_v(x) &= A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}}\,, \\ x\bar{U}(x) &= A_{\bar{U}} x^{B_{\bar{U}}} (1-x)^{C_{\bar{U}}}\,, \\ x\bar{D}(x) &= A_{\bar{D}} x^{B_{\bar{D}}} (1-x)^{C_{\bar{D}}}\,, \\ xg(x) &= A_g x^{B_g} (1-x)^{C_g} (1 + E_g x^2)\,, \end{aligned} \tag{4.4}$$

while for the photon PDF it is used:

$$x\gamma(x) = A_\gamma x^{B_\gamma} (1-x)^{C_\gamma} (1 + D_\gamma x + E_\gamma x^2)\,. \tag{4.5}$$

The parametrisation of the quark and gluon PDFs in Eq. (4.5) differs from the one used in the HERAPDF2.0 analysis in various ways. First of all, a higher value of the input evolution scale Q_0^2 is used, which is helpful to stabilise the fit of the photon PDF. Second, an additional negative term in the parametrisation of the gluon is not required here, because of the increased value of Q_0^2 which assures the positiveness of the gluon distribution. Third, the results of the parametrisation scan are different because of the inclusion of the ATLAS high-mass DY cross-section data. PDF uncertainties are estimated using the Monte Carlo (MC) replica method [166–168], cross-checked with the Hessian method [36] using $\Delta\chi^2 = 1$. The former is expected to be more robust than the latter, due to the potential non-Gaussian nature of the photon PDF uncertainties [149]. In the following, it will be shown that these two methods to estimate the PDF uncertainties on the photon PDF lead to similar results.

In addition, a number of cross-checks have been performed to assess the impact of various model and parametrisation uncertainties. For the model uncertainties, variations of the charm mass between $m_c = 1.41$ GeV to 1.53 GeV, of the bottom mass between $m_c = 4.25$ GeV to 4.75 GeV, of the strong coupling constant $\alpha_s(m_Z)$ between 0.116 to 0.120 are considered, and additionally the strangeness fraction is decreased down to $r_s = 0.75$. For the parametrisation uncertainties, the impact of increasing the input parametrisation scale up to $Q_0^2 = 10$ GeV^2 is considered as well as the impact of including additional parameters in Eq. (4.5). These extra parameters make little difference to the χ^2 of the fit, but they can change the shape of the PDFs in a non-negligible way. Such additional parameters are D_{u_v}, $D_{\bar{u}}$, $E_{\bar{d}}$, as well and the extra negative term in the gluon PDF used in HERAPDF2.0. The impact of these model and parametrisation uncertainties on the baseline results is quantified in Sec. 4.5.3.

In the end, Fig. 4.7 and Fig. 4.8 show the *u* - valence, *d* - valence, $\bar{u}$, $\bar{d}$ and gluon PDF distributions at Q^2 = 7.5 GeV^{2}and Q^2 = 10^4 GeV^2 respectively. Here, the PDFs from the current fit are compared to those of HERAPDF2.0 and NNPDF3.0. The

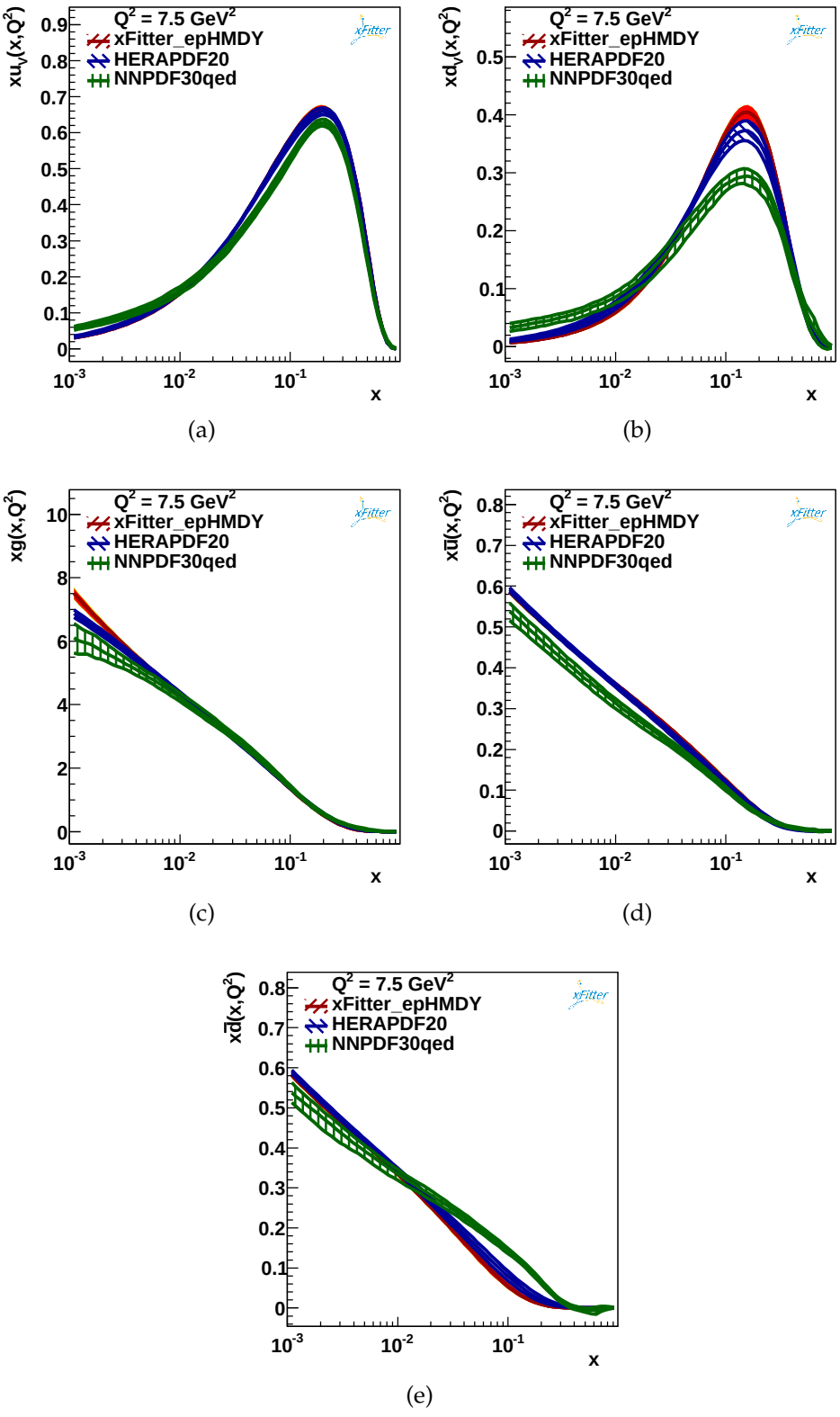


Fig. 4.7: NNLO PDF distributions at $Q^2 = 7.5^2$ GeV^2 for PDFs from the current fit, HERAPDF2,0 and NNPDF3.0: (a) u - valence; (b) d - valence; (c) gluon; (d) $\bar{u}$; (e) $\bar{d}$

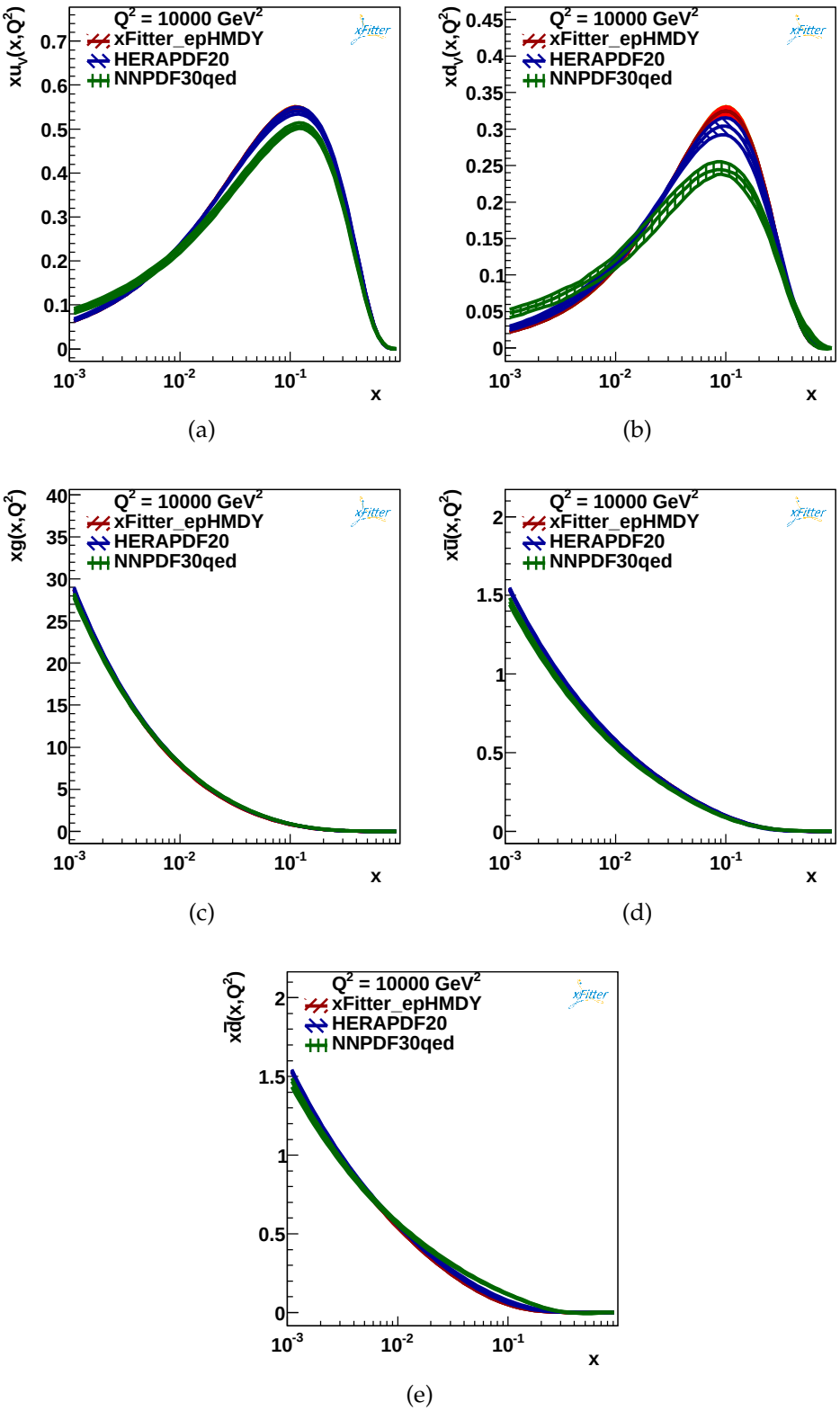


Fig. 4.8: NNLO PDF distributions at $Q^2 = 10000^2$ GeV^2 for PDFs from the current fit, HERAPDF2,0 and NNPDF3.0: (a) u - valence; (b) d - valence; (c) gluon; (d) $\bar{u}$; (e) $\bar{d}$

Dataset	χ^2 /$N_{\rm dat}$
HERA I+II	1236/1056
high-mass DY 116 GeV $\leq m_{\ell\ell} \leq$ 150 GeV	9/12
high-mass DY 150 GeV $\leq m_{\ell\ell} \leq$ 200 GeV	15/12
high-mass DY 200 GeV $\leq m_{\ell\ell} \leq$ 300 GeV	14/12
high-mass DY 300 GeV $\leq m_{\ell\ell} \leq$ 500 GeV	5/6
high-mass DY 500 GeV $\leq m_{\ell\ell} \leq$ 1500 GeV	4/6
Correlated (high-mass DY) χ^2	1.17
Log penalty (high-mass DY) χ^2	-0.12
Total (high-mass DY) $\chi^2/N_{\rm dat}$	48/48
Combined HERA I+II and high-mass DY $\chi^2/N_{\rm dof}$	1284/1083

Table 4.1: The $\chi^2/N_{\rm dat}$ in the NNLO fits for the HERA inclusive structure functions and for the various invariant mass $m_{\ell\ell}$ bins of the ATLAS high-mass DY data. In the latter case, the contribution to the χ^2 arising from the correlated and log-penalty terms are indicated, as well as the overall $\chi^2/N_{\rm dof}$ is provided, where $N_{\rm dof}$ is the number of degree of freedom in the fit.

PDF shapes of the present fit are often closer to the HERAPDF than to NNPDF because of the dominance of HERA data in the fit.

4.5 Results

4.5.1 Fit quality and comparison between data and fit results

In the following, the results that will be shown correspond to those obtained from fitting the double-differential $(m_{\ell\ell}, |y_{\ell\ell}|)$ cross-section distributions. It has been verified that comparable results are obtained if the $(m_{\ell\ell}, \Delta\eta_{\ell\ell})$ cross-section distributions are fitted instead.
For the baseline NNLO fit, the value $\chi^2_{min}/N_{dof} = 1284/1083$ is obtained where N_{dof} is the number of degrees of freedom in the fit which is equal to total number of data points minus number of free parameters. The contribution from the HERA inclusive data is $\chi^2/N_{\rm dat} = 1236/1056$ and from the ATLAS high-mass DY data is $\chi^2/N_{\rm dat} = 48/48$, where $N_{\rm dat}$ the number of the data points for the corresponding data sample. These values for $\chi^2/N_{\rm dat}$, together with the corresponding values for the various invariant mass $m_{\ell\ell}$ bins of the ATLAS data, are summarised in Table 4.1. The quality of the agreement with the HERA cross sections is of comparable quality to that found in the HERAPDF2.0 analysis. Note that in the calculation of the total χ^2 for the ATLAS data, the correlations between the different $m_{\ell\ell}$ bins have been taken into account.
Fig. 4.9 demonstrate a good agreement between ATLAS data and the NNLO theory predictions obtained from the `xFitter_epHMDY` fit. This agreement is also quantitatively expressed by the values of the χ^2 reported in Table 4.1, where for the ATLAS data a $\chi^2/N_{\rm dat} = 1$ is found. This is particularly remarkable given the high precision of the data, with total experimental uncertainties at the few percent level in most of the kinematic range.
The output parameters from the NNLO fit are reported in Table 4.2.

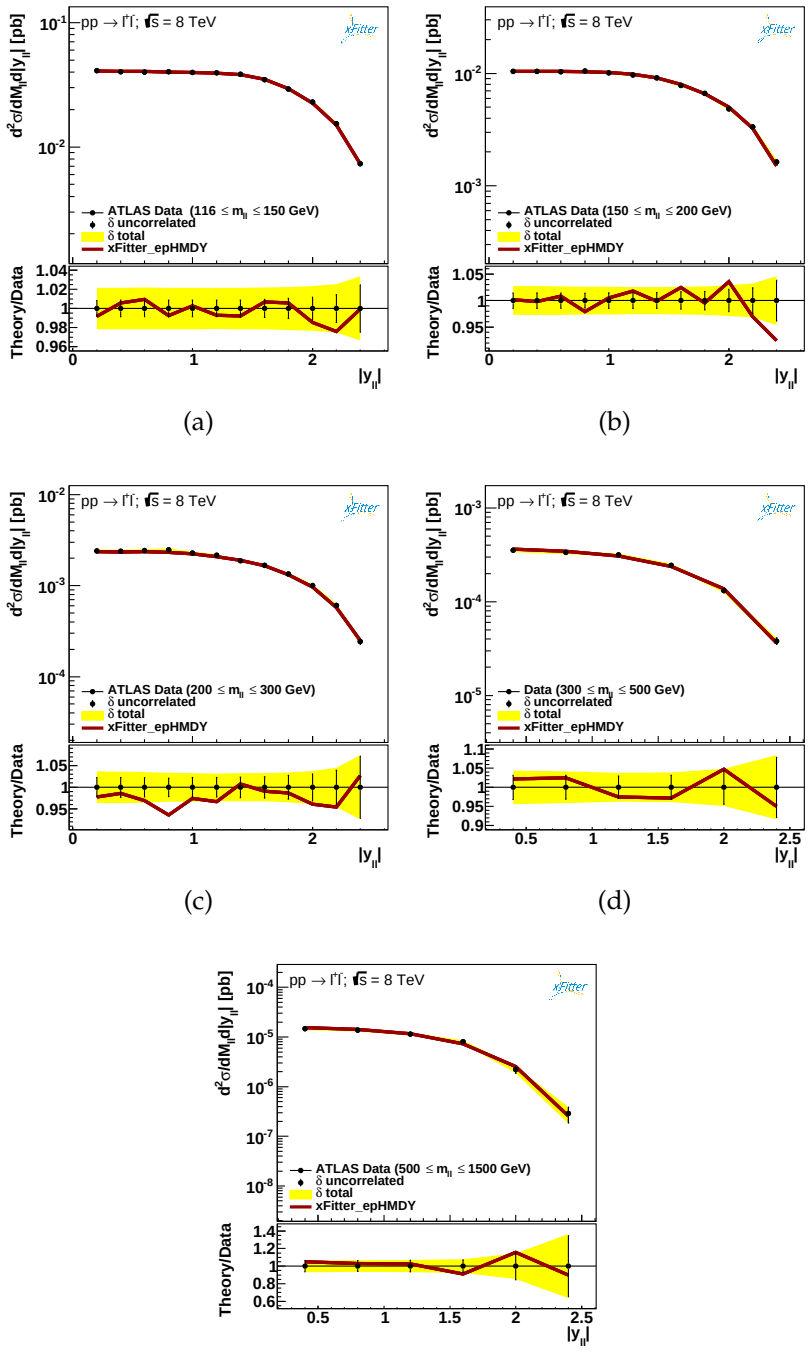


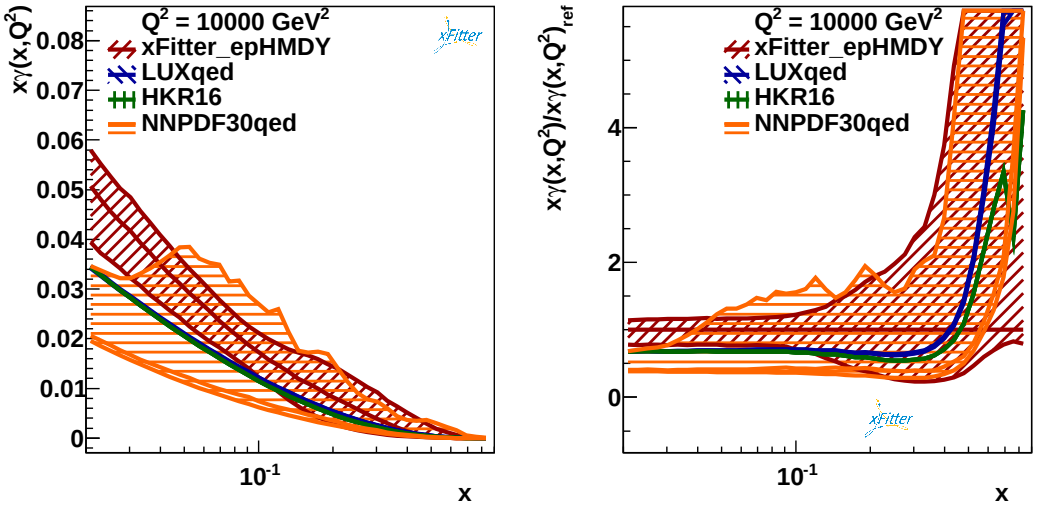
Fig. 4.9: Comparison between the results of the fit and the ATLAS data for the $(m_{\ell\ell}, |y_{\ell\ell}|)$ double-differential DY cross-sections as function of $|y_{\ell\ell}|$. The comparisons are shown both in an absolute scale and as ratios to the central value of the experimental data in each $y_{\ell\ell}$ bin. The error bars on the data points correspond to the bin-to-bin uncorrelated uncertainties, while the yellow bands indicate the size of the correlated uncertainties. The solid lines indicate the results of this fit `xFitter_epHMDY`.

Parameter	xFitter_epHMDY
B_g	$-0.220^{+0.014}_{-0.013}$
C_g	$6.92^{+0.62}_{-0.61}$
B_{u_v}	$0.761^{+0.017}_{-0.015}$
C_{u_v}	$5.060^{+0.064}_{-0.092}$
E_{u_v}	$8.07^{+0.80}_{-0.82}$
C_{d_v}	$5.61^{+0.24}_{-0.30}$
$C_{\bar{u}}$	$6.37^{+0.56}_{-0.38}$
$A_{\bar{d}}$	$0.3226^{+0.0078}_{-0.0083}$
$B_{\bar{d}}$	$-0.1921^{+0.0033}_{-0.0033}$
$C_{\bar{d}}$	$14.0^{+1.0}_{-1.7}$
A_γ	$0.00120^{+0.031}_{-0.00089}$
B_γ	$-0.62^{+0.63}_{-0.36}$
C_γ	$10.0^{+13}_{-5.9}$
D_γ	-4^{+210}_{-15}
E_γ	87^{+257}_{-140}

Table 4.2: PDF parameters for the NNLO fit.

4.5.2 Determination of the photon PDF

In Fig. 4.10, the photon PDF, $x\gamma(x, Q^2)$, is shown at $Q^2 = 10^4$ GeV^2, and it is compared to the corresponding LUXqed, HKR16 and NNPDF3.0QED results. In the left plot the comparison is presented in an absolute scale, while in the right plot the ratio of different results normalized to the central value of the fit is shown. For the present fit, `xFitter_epHMDY`, the experimental PDF uncertainties at the 68% confidence level (CL) are obtained from the MC method, while model and parametrisation uncertainties are discussed below. Likewise, the NNPDF3.0QED PDF set is shown the 68% CL uncertainty band, while for LUXqed the associated PDF uncertainty band is computed according to the prescription of Ref. [52]. For HKR16, only the central value is available. The x-range in Fig. 4.10 has been restricted to the region $0.02 \leq x \leq 0.9$, since beyond that region there is only limited sensitivity to $x\gamma(x, Q^2)$. Fig. 4.10 shows that for $x \geq 0.1$ the four determinations of he photon PDF are consistent within PDF uncertainties at the 1-σ level. For smaller values of x, the photon PDF from LUXqed and HKR16 is somewhat smaller than `xFitter_epHMDY`, but still in agreement at the 2-σ level. This agreement is further improved if the PDF uncertainties in `xFitter_epHMDY` arising from variations of the input parametrisation are added to experimental uncertainties, as discussed in the following. Moreover, the results shown in this chapter and NNPDF3.0QED agree at the 68% CL for $x \geq 0.03$, and the agreement extends to smaller values of x once the parametrisation uncertainties in `xFitter_epHMDY` are accounted for (more details in Sec. 4.5.3). The LUXqed and the HKR16 calculations of $x\gamma(x, Q^2)$ are very close to each other across the entire range of x. Fig. 4.10 shows that for $0.04 \leq x \leq 0.2$ the present analysis exhibits smaller PDF uncertainties as compared to those from NNPDF3.0QED. Indeed, the experimental uncertainty on the `xFitter_epHMDY` turns out to be at the $\sim 30\%$ level for $x \leq 0.1$. At larger x it



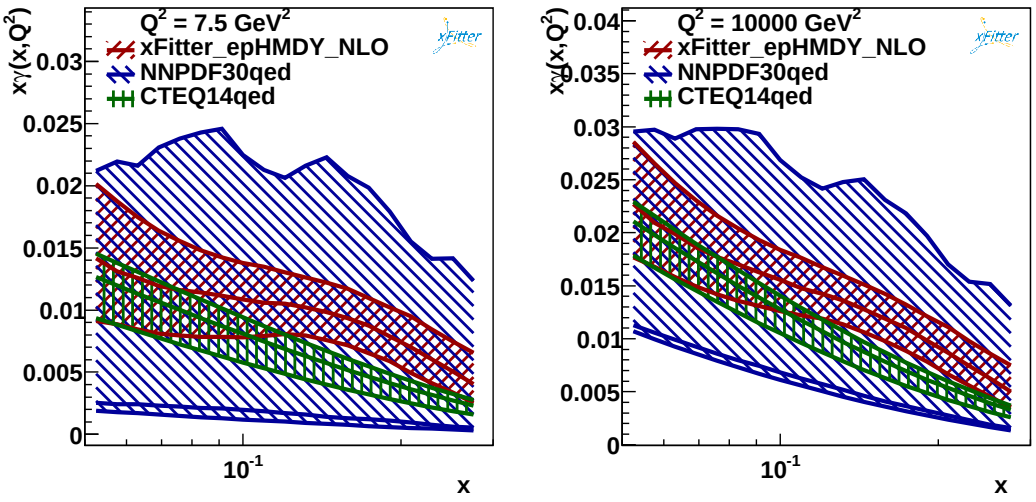
(a) (b)

Fig. 4.10: (a) Comparison between the photon $x\gamma(x, Q^2)$ at $Q^2 = 10^4$ GeV^2 from the present NNLO analysis (`xFitter_epHMDY`) with the corresponding results from NNPDF3.0QED, LUXqed and HKR16; (b) the same comparison, now with the results normalized to the central value of `xFitter_epHMDY`. For the present fit, the PDF uncertainties are shown at the 68% CL obtained from the MC method, while model and parametrisation uncertainties are discussed below. For HKR16 only the central value is shown, while for LUXqed the associated PDF uncertainty band [52] is included.

increases rapidly specially in the positive direction. The reason for this behaviour at large x can be understood by recalling that variations of $x\gamma(x, Q^2)$ in the negative direction are constrained by positivity. The limited sensitivity of the ATLAS data does not allow a determination of $x\gamma(x, Q^2)$ with uncertainties competitive with those of LUXqed, which are at the few percent level. These results became available while this analysis was underway. Nevertheless, it is valuable to see that the present model independent evaluation of the photon PDF, including the high-mass DY data, can improve on the similar model independent analysis of NNPDF3.0, which lacks these data.

A similar comparison between the NLO photon distribution and the NNPDF3.0qed NLO photon PDF in the restricted x-range is shown in Fig. 4.11. The comparison is also made to the CT14qed photon PDF [169]; the bands in the plot represent the experimental PDF uncertainty 68% C.L. . The NLO fit predictions agree with CT14qed photon PDF at the 1-σ level in most of the restricted x-range and the improvement in the uncertainty related to the photon PDF, compared that estimated by NNPDF3.0qed, is quite remarkable.

It is also interesting to assess the impact of the high-mass Drell-Yan 8 TeV measurements on the light quark and gluon PDFs. For this purpose, the fits have been repeated freezing the photon PDF to the `xFitter_epHMDY` shape. This is necessary because HERA inclusive data alone, which are the benchmark for this comparison, have no sensitivity to the photon PDF. In this way, a meaningful comparison between the quark and gluon PDFs from a HERA-only baseline and



(a) (b)

Fig. 4.11: Comparison between the NLO photon PDF distributions for the present analysis, NNPDF3.0QED, CT14qed (68% C.L.): (a) at scale 7.5 GeV^2; (b) at the evolved scale 10^4 GeV^2.

the HERA+HMDY fit can be performed. This comparison is shown in Fig. 4.12 for the up and down anti-quarks $x\bar{u}(x, Q^2)$ and $x\bar{d}(x, Q^2)$ PDFs, for which the effect of the high-mass DY data is expected to be most pronounced, since HERA inclusive cross sections provide little information on quark flavour separation. In Fig. 4.12, the $x\bar{u}(x, Q^2)$ and $x\bar{d}(x, Q^2)$ together with the associated MC uncertainties have been computed at the initial parametrisation scale of $Q^2 = 7.5$ GeV^2 and are shown as ratios to the central value of the `xFitter_epHMDY` fit. The modifications in the medium and large-x antiquark distributions from the high mass DY data are rather moderate. It has been verified that the same conclusions can be derived from fits obtained by switching off the QED effects for both the HERA only fits and the HERA+HMDY fits. Therefore, while the ATLAS high-mass DY measurements have a significant constraint on the photon PDF, their impact on the quark and gluon PDFs is moderate.

4.5.3 Robustness and perturbative stability checks

Following the presentation of the main result shown in this chapter, the robustness of this determination with respect to a number of variations is assessed. Firstly, variations in the values of the input physical parameters, such as α_s or the charm mass are explored. Secondly, variations of the choices made for the PDF input parametrisation are considered. Finally, variations associated to different methodological choices in the fitting procedure are quantified. In each case, one variation at a time is performed and compared with the central value of $x\gamma(x, Q^2)$ and its experimental PDF uncertainties computed using the MC method.

First the impact of uncertainties associated to either the choice of input physical parameters or of specific settings adopted in the fit is considered. Fig. 4.13

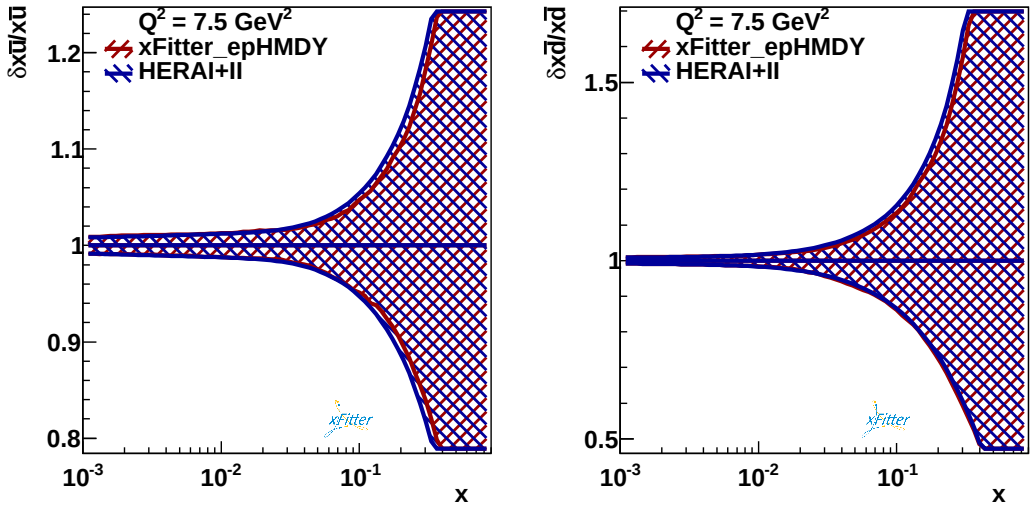


Fig. 4.12: The impact of the ATLAS high-mass 8 TeV Drell-Yan measurements on the $x\bar{u}$ and $x\bar{d}$ sea quark PDFs at the input parametrisation scale $Q^2 = 7.5$ GeV^2. The results are shown normalized to the central value of `xFitter_epHMDY`.

shows the comparison between the `xFitter_epHMDY` determination of $x\gamma(x, Q^2)$ at $Q^2 = 10^4$ GeV^2, including the experimental MC uncertainties, with the central value of those fits for which a number of variations have been performed. Specifically:

- The strong coupling constant is varied by $\delta\alpha_s = \pm 0.002$ around the central value.
- The ratio of strange to non-strange light quark sea PDFs is decreased to $r_s = 0.75$ instead of $r_s = 1$.
- The value of the charm mass is varied between $m_c = 1.41$ GeV and $m_c = 1.53$ GeV, and that of the bottom mass between $m_b = 4.25$ GeV and $m_b = 4.75$ GeV.
- The minimum value $Q^2_{\rm min}$ of the fitted data is decreased down to 5 GeV^2.
- The input parametrisation scale Q_0^2 is raised to 10 GeV^2 as compared to the baseline value of $Q_0^2 = 7.5$ GeV^2.

The results of Fig. 4.13 highlight that in all cases effect of the variations considered here is contained within (and typically much smaller than) the experimental PDF uncertainty bands of the reference fit. The largest variation comes from the strangeness ratio r_s, where the resulting central value turns out to be at the bottom end of the PDF uncertainty band for $x \geq 0.1$. From the χ^2 point of view, all above described model variation has a total $\chi^2/N_{\rm dof}$ which differs just by few units from the one of the nominal fit; so most of various fit are indistinguishable from each other in Fig. 4.13.
Another important check of the robustness of the present determination of $x\gamma(x, Q^2)$ can be obtained by comparing the baseline fit with further fits where a number of

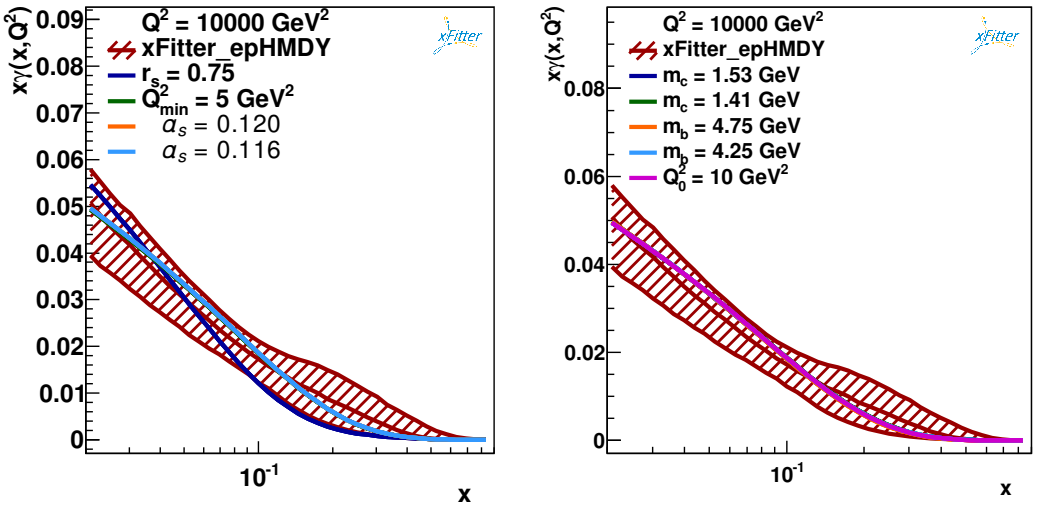


Fig. 4.13: Comparison between the baseline determination of $x\gamma(x, Q^2)$ at $Q^2 = 10^4$ GeV^2 in the present analysis, `xFitter_epHMDY`, with the central value of a number of fits for which one input parameter has been varied. The following variations have been considered: $r_s = 0.75$, $Q^2_{\rm min} = 5$ GeV^2, $\alpha_s = 0.116$ and 0.118 (left plot); and $m_c = 1.41$ and 1.53 GeV, $m_b = 4.25$ and 4.75 GeV, and $Q^2_0 = 10$ GeV^2 (right plot). See text for more details about these variations.

new free parameters are allowed in the PDF parametrisation, in addition to those listed in Eq. (4.5). Fig. 4.14 shows the impact of three representative variations (others have been explored, leading to smaller differences): more flexibility to the gluon distribution, introducing a negative counter term (labelled by "neg"), in addition to D_{u_v}, and then $D_{\bar{u}} + D_{\bar{d}}$. As before, all variations are contained within the experimental PDF uncertainty bands, though the impact of the parametrisation variations is typically larger than that of the model variations: in the case of the $\mathrm{neg} + D_{\bar{u}} + D_{\bar{d}}$ variations, the central value is at the lower edge of the PDF uncertainty band in the entire range of x shown.

A cross-check of the robustness of the estimated experimental uncertainty of the photon PDF in this analysis is provided by the comparison of the MC replicas and Hessian methods. Fig. 4.14 shows this comparison indicating a reasonable agreement between the two methods. In particular, the central values of the photon obtained with the two fitting techniques are quite similar to each other. As expected, the MC uncertainties tend to be larger than the Hessian ones, especially in the region $x \geq 0.2$, indicating deviations with respect to the Gaussian behaviour of the photon PDF.

To complete these studies, an interesting exercise is to quantify the perturbative stability of the `xFitter_epHMDY` determination of the photon PDF $x\gamma(x, Q^2)$ with respect to the inclusion of NNLO QCD corrections in the analysis. To study this, Fig. 4.15 shows a comparison between the baseline fit of $x\gamma(x, Q^2)$, based on NNLO QCD and NLO QED theoretical calculations, with the central value resulting from a corresponding fit based instead on NLO QCD and QED theory. In other words, the QED part of the calculations is identical in both cases. For the NNLO fit, only the experimental PDF uncertainties, estimated using the MC

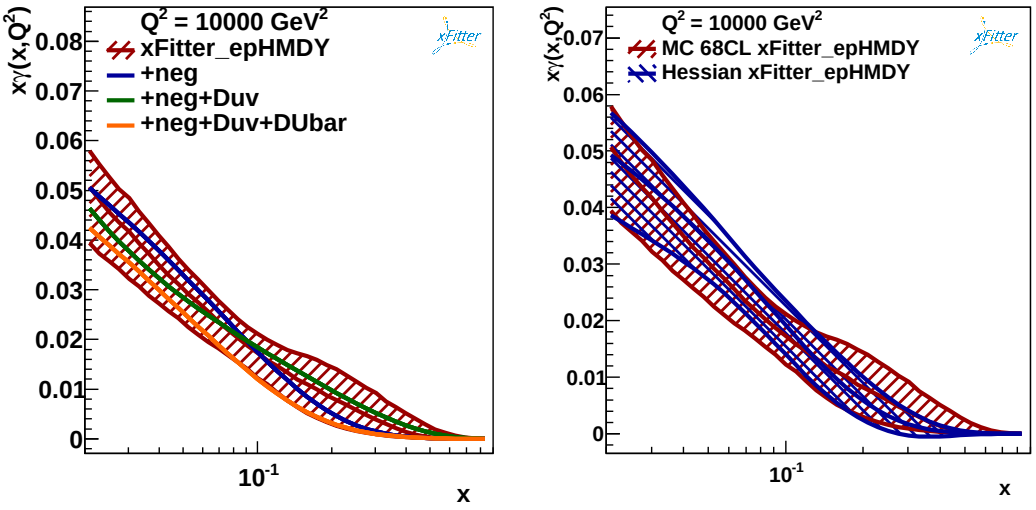


Fig. 4.14: Left: the impact on the photon PDF $x\gamma(x, Q^2)$ from `xFitter_epHMDY` in fits where a number of additional free parameters are allowed in the PDF parametrisation Eq. (4.5). The parametrisation variations that have been explored are: more flexibility to the gluon distribution, introducing a negative counter term (labelled by "neg"), adding on top D_{u_v}, and then adding $D_{\bar{u}} + D_{\bar{d}}$. Right: comparison between the `xFitter_epHMDY` determinations obtained with the MC (baseline) and with the Hessian methods, where in both cases the PDF error band shown corresponds to the 68% CL uncertainties.

method, are shown. From the comparison of Fig. 4.15, it is clear that the fit of $x\gamma(x, Q^2)$ exhibits a reasonable perturbative stability, since the central value of the NLO fit is always contained in the one-sigma PDF uncertainty band of the baseline `xFitter_epHMDY` fit. The agreement between the two fits is particularly good for $x \geq 0.1$, where the two central values are very close to each other. This comparison is shown at low scale, $Q^2 = 7.5$ GeV^2 and high scales $Q^2 = 10^4$ GeV^2, indicating that perturbative stability is not scale dependent.

4.6 Summary of Chapter 4

In this chapter, a new determination of the photon PDF from a fit of HERA inclusive DIS structure functions supplemented by ATLAS data on high-mass Drell-Yan cross sections has been presented, based on the `xFitter` framework. As suggested by a previous reweighting analysis [143], this high-mass DY data provides significant constraints on the photon PDF, allowing a determination of $x\gamma(x, Q^2)$ with uncertainties at the 30% level for $0.02 \leq x \leq 0.1$. The results of the present study, dubbed `xFitter_epHMDY`, are in agreement and exhibit smaller PDF uncertainties that the only other existing photon PDF fit from LHC data, the NNPDF3.0QED analysis, based on previous LHC Drell-Yan measurements. The results are in agreement within uncertainties with two recent calculations of the photon PDF, LUXqed and HKR16. For $x \geq 0.1$, the agreement is at the 1-σ level already including only the experimental MC uncertainties, while for $0.02 \geq x \geq 0.1$ it is important to account for parametrisation uncertainties. The findings indicate

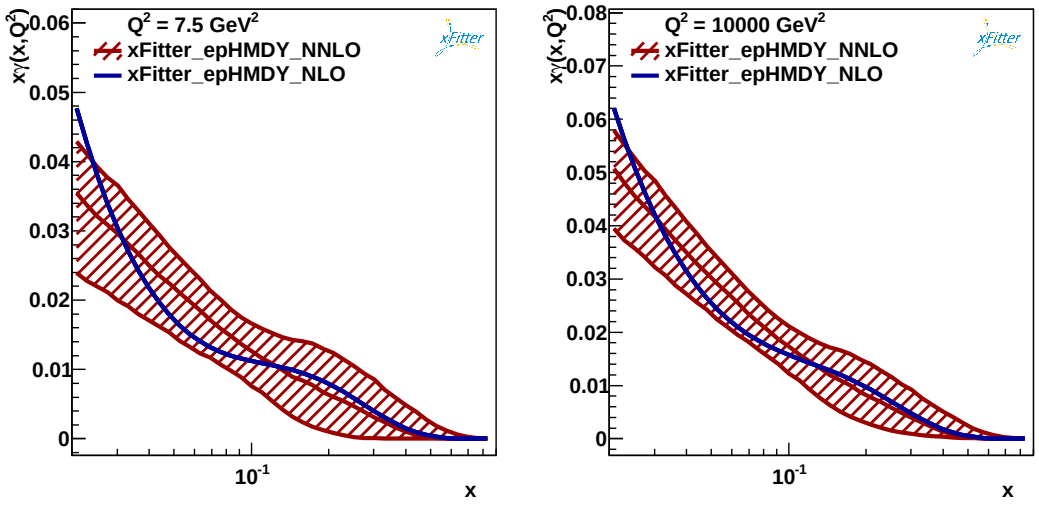


Fig. 4.15: Left plot: comparison between the reference `xFitter_epHMDY` fit of $x\gamma(x, Q^2)$, based on NNLO QCD and NLO QED theoretical calculations, with the central value of the corresponding fit based on NLO QCD and QED theory, at $Q^2 = 7.5$ GeV^2. In the former case, only the experimental MC PDF uncertainties are shown. Right plot: same comparison, now presented at the higher scale of $Q^2 = 10^4$ GeV^2.

that a direct determination of the photon PDF from hadron collider data is still far from being competitive with the LUXqed and HKR calculations, which are based instead on precise measurements of the inclusive DIS structure functions of the proton. The results of this study are available upon request in the `LHAPDF6` format [170].

Chapter 5

Z' BSM search in the di-electron channel at $\sqrt{s}$ = 13 TeV

The di-lepton (ee and $\mu\mu$) final-state signature has excellent sensitivity to a wide variety of new phenomena expected in theories beyond the Standard Model (BSM). It benefits from high signal selection efficiencies and relatively small, well-understood backgrounds. My own contributions were on the optimization of the cuts applied in the event selection in the electron channel, the validation study of the higher order *k*-factors (NNLO QCD + NLO EW)/(NLO QCD + LO EW) and setting exclusion limits on the Z' production cross section times branching ratio to two leptons as a function of Z' pole mass in various different BSM scenarios. Given the fact that my contribution to this search was in the di-electron channel, only results in this specific channel or combined results will be shown in the following, without any particular emphasis on the results in the other decay channel. The focus of this chapter is on my contributions; however, in order to make them understandable and to put them into context, a concise account of all stages involved in obtaining the results is included. During the period I got involved in this BSM search, I contributed to two ATLAS CONF notes [171,172] and two papers [76,86].

5.1 Data and MC samples

The ATLAS software release 20.7 was used for reconstructing both MC events and the data. The detector response is simulated with `GEANT4` [173,174]. Reweighting of events in MC to match the data according to the trigger, isolation, and reconstruction efficiency are also included.

5.1.1 Background processes

The data sample used in this analysis was collected during the 2015 and 2016 LHC runs with *pp* collisions at $\sqrt{s}$ = 13 TeV. After selecting periods with stable beams and requiring that relevant detector systems are functional, the data set used for the analysis corresponds to 36.1 fb^{-1} of integrated luminosity. Event qual-

ity is also checked to remove those events which contain noise bursts or coherent noise in the calorimeters.
Modelling of the various background sources relies primarily on Monte Carlo (MC) simulation. The dominant background contribution arises from the Drell-Yan (DY) [175] process. Other background sources are $t\bar{t}$, single-*t* [3] and diboson (*WW*, *WZ*, *ZZ*) [176] production. In the case of the di-electron channel, multi-jet and *W*+jets processes also contribute due to the misidentification of jets as electrons. A data driven method, described in the following, is used to estimate these background contributions.
DY events are simulated using `Powheg-box_v2` [177] at next-to-leading order (NLO) in Quantum Chromodynamics (QCD), and interfaced to the `Pythia v8.186` [178] parton shower model. The CT10 partonic distribution function (PDF) set [179] is used in the matrix element calculation. The AZNLO [180] set of tuned parameters ("tune") is used, with the CTEQ6L1 PDF set [181], for the modelling of non-perturbative effects. Event yields are corrected with a mass-dependent rescaling to next-to-next-to-leading order (NNLO) in the QCD coupling constant, computed with `VRAP_v0.9` [182] and the CT14NNLO PDF set [183]. The NNLO QCD corrections are a factor of ≈ 0.98 at $m_{ll} = 3$ TeV. Mass-dependent electroweak (EW) corrections are computed at NLO with `MCSANC_v1.20` [184]. The NLO EW corrections are a factor of ≈ 0.86 at m_{ll} = 3 TeV. Those include photon-induced contributions ($\gamma\gamma \to ll$ via t- and u-channel processes) computed with the MRST2004QED PDF set [48]. The quantity $\sigma_{\text{NNLO\_QCD+NLO\_EW}}$ (numerator of the k_F expression) can be obtained in two different ways:

- **factored computation**: the k_F are fully multiplicative and

$$\sigma_{\text{NNLO\_QCD+NLO\_EW}} = k_{\text{QCD}} \times k_{\text{EW}} \times \sigma_{\text{LO\_QCD,LO\_EW}} \tag{5.1}$$

  with $k_{\text{QCD}} = \sigma_{\text{NNLO\_QCD}}/\sigma_{\text{LO\_QCD}}$ (computed with `VRAP_v0.9`) and $k_{\text{EW}} = \sigma_{\text{HO\_EW}}/\sigma_{\text{LO\_QCD,EW}}$ (computed with `MCSANC_v1.20`),

- **additive computation**: the HO cross section is constructed as follows:

$$\sigma_{\text{NNLO\_QCD+NLO\_EW}} = \sigma_{\text{NNLO\_QCD}} \times \left(1 + \frac{\sigma_{\text{LO\_QCD,EW}} - \sigma_{\text{LO\_QCD}}}{\sigma_{\text{NNLO\_QCD}}}\right) \tag{5.2}$$

Di-boson processes with four charged leptons, three charged leptons and one neutrino, or two charged leptons and two neutrinos are simulated using the `Sherpa_v2.1.1` generator [185]. Matrix elements contain all diagrams with four electroweak vertices. They are calculated for up to one ($4l, 2l + 2\nu$) or no additional partons ($3l + 1\nu$) at NLO. Di-boson processes with one of the bosons decaying hadronically and the other leptonically also are simulated using the `Sherpa_v2.1.1` generator. They are calculated for up to one (*ZZ*) or no (*WW*,*WZ*) additional partons at NLO. The CT10 PDF set is used in conjunction with dedicated parton shower tuning developed by the Sherpa authors.
For the generation of $t\bar{t}$ and single top quarks in the Wt-channel and s-channel the `Powheg-box_v2` generator with the CT10 PDF set in the matrix element calculations is used. EW *t*-channel single-top-quark events are generated using the

Powheg-box_v1 generator. This generator uses the four-flavour scheme for the NLO matrix element calculations together with the fixed four-flavour PDF set CT10f4. The parton shower, fragmentation, and the underlying event are simulated using Pythia v6.428 [186] with the CTEQ6L1 PDF set and the Perugia 2012 tune [187]. The top-quark mass (m_{top}) is set to 172.5 GeV. The $t\bar{t}$ and single-top-quark MC samples are normalised to a cross-section as calculated with the Top++2.0 [188] program, which is accurate to NNLO in perturbative QCD, including resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms.

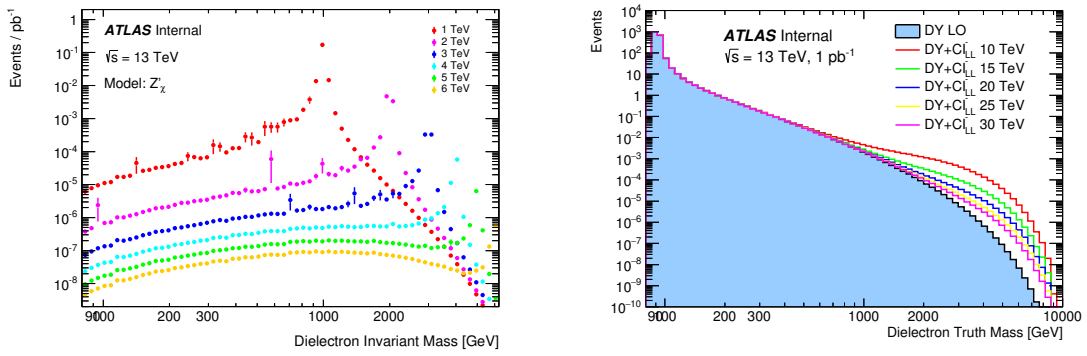
5.1.2 Signal processes

Dedicated Z' signal samples were generated at LO using Pythia8.186 with the NNPDF23LO PDF [148] and A14 tune [189] for event generation, parton showering and hadronisation. Samples were created for four polemasses, respectively 2, 3, 4 and 5 TeV, for benchmarking the Z'_{χ} model and at 3 TeV for the Z'_{SSM}. The Pythia8 generator was also used to produce the CI signal samples at LO using the NNPDF23LO PDF were produced using the Pythia8 generators and five benchmark values of Λ (every 5 TeV starting from 10 TeV and up to 30 TeV) were generated for each of the CI models, in both the destructive and constructive interference scenarios.
The strategy for producing these signal templates is based upon reweighting LO DY to create Z' and CI signal templates for a model at any given value of Z' polemass, or CI value of Λ. The reweighting algorithm to the LO DY samples takes as input on an event-by-event basis: the event born mass, the flavour of the incoming quarks and the desired Z' polemass or Λ value. Then, for that event, the Feynman diagrams at LO for both the SM DY and BSM Z'/CI model and the reweighting function returns a ratio of the two, which is applied to the event as a weight and transforms the SM DY spectrum into the desired model. This reweighting techinque is codified in the LPXSignalReweightingTool, a package avilable in RootCore and designed in such a way to be plugged directly in the analysis workflow. Ref. [190] gives a detailed explanation of this reweighting technique and how to set it up with specific examples.
An example of the Z' and CI templates produced with this technique can be found in Fig. 5.1.

5.2 Event selection

The selected events are required to be in the list of good runs[1,2]. Events are required to have at least one reconstructed primary vertex and at least one pair of same-flavour lepton candidates. Then, the selected events have to pass a two-electron trigger based on the *Loose* identification criteria with an E_T threshold of 17 GeV for each electron and the LAr, Tile and SCT requirements which provide

[1] data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns.xml
[2] data16_13TeV.periodAllYear_DetStatus-v83-pro20-15_DQDefects-00-02-04_PHYS_StandardGRL_All_Good_25ns.xml



(a) (b)

Fig. 5.1: Left plot: Signal templates for the Z'_χ presented at six representative pole masses. Right plot: Signal templates for the CI LL model with constructive interference presented for five representative Λ values.

protection against data corruption and noise bursts in the calorimeters. Finally, an event is selected if it contains at least two electrons.
Electrons are reconstructed in the central region of the ATLAS detector covered by the tracking detectors ($|\eta| < 2.47$), by combining calorimetric and tracking information. The transition region between the central and forward regions of the calorimeters, in the range $1.37 \leq |\eta| \leq 1.52$, exhibits degraded energy resolution and is therefore excluded. Electrons must not flagged as coming from a bad calorimeter cluster, furthermore to search for high-mass di-lepton signatures of new physics, requirements are applied to the data and MC samples to select events with two high-E_T electrons, satisfying the criteria described below:

- They must have $E_T > 30$ GeV;
- They must pass the ID track requirement on the d_0 significance: $|\, d_0^{BL}/\sigma(d_0^{BL})\,| < 5$;
- They must pass the ID track requirement $|\, \Delta z_0 \sin(\theta)\,| < 0.5$ mm with respect to the beam line;
- They must pass the "Likelihood Medium" identification criteria based on track and shower-shape variables;
- They must pass the "Loose" isolation criteria.

Only the electron pair with the highest scalar sum of E_T is retained in each event; the opposite-charge requirement is not applied in the di-electron channel due to the high chance of charge misidentification for high-E_T electrons.
The electron event selection yield for the 2015 and 2016 datasets after requiring the final selection is shown in Fig. 5.2.

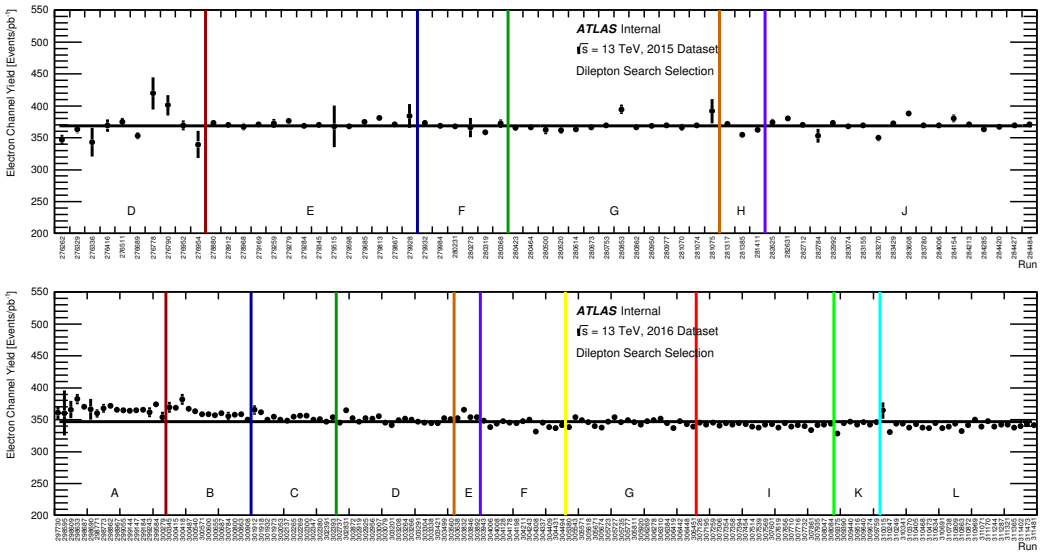


Fig. 5.2: Yield for each run of 2015 (upper plot) and 2016 (lower plot) data taking periods, after requiring the final selection. The capital letter represent the various data taking periods.

5.3 MC background estimation

The SM processes producing two real leptons in the final state are modelled and estimated using Monte Carlo (MC) simulated samples as described in Section 5.1. The processes for which MC simulation is used are: DY, $t\bar{t}$ and single-top-quark, and di-boson (WW, WZ, and ZZ) production. The event selection is applied to each of the MC samples representing these backgrounds, which are subsequently scaled to the same integrated luminosity. Scale factors are applied to the MC distributions to correct them with respect to the collected data, and are mostly derived by the combined performance groups in several bins of E_T and $|\eta|$. The scale factors correct for the trigger, electron ID, isolation, and reconstruction efficiencies.

5.4 Fake electron background estimation

As regards the *W*+jets, multi-jet and other production processes, where at most one real electron is produced, nevertheless contribute to the selected *ee* sample due to their having one or more hadronic jets satisfying the electron selection criteria. The contribution from these processes is estimated simultaneously with a data-driven (DD) technique, called the *matrix method*.

5.4.1 The matrix method

This methodology is based upon loosening certain identification criteria applied in the nominal selection, to estimate the rate with which those looser objects pass the nominal selection. In this way, the signal contamination by misidentified jets can be determined. In this method, two different electron selections are de-

fined: a *tight* (T) one and a *loose* (L) one. Note that the set of objects passing the tight selection, N_{tight}, is a subset of those passing the loose criteria, N_{loose}.
In this analysis, only pairs of electron candidates denoted by N_{xy} (where $x, y \in T, L$) are considered. In this way, the number of electron pairs passing four different selection are measurable: N_{TT}, N_{TL}, N_{LT} and N_{LL}. The first index represents the leading object ($p_{T,lead} > p_{T,sub}$) and the second one the sub-leading. The relation between these reconstructed quantities and the real ones denoted by N_{ab} with $a, b \in R, F$ which refers to whether the object is a real electron or a fake one is shown is Eq. 5.3.

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = \begin{pmatrix} r^2 & rf & fr & f^2 \\ r(1-r) & r(1-f) & f(1-r) & f(1-f) \\ (1-r)r & (1-r)f & (1-f)r & (1-f)f \\ (1-r)^2 & (1-r)(1-f) & (1-f)(1-r) & (1-f)^2 \end{pmatrix} \begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix} \tag{5.3}$$

To be considered, the real objects are required to pass at least the *lhLoose* identification criteria (see Sec. 3.3.2 for the definition of this working point). The quantities in the vector on the right hand side of Eq. 5.3 are true quantities and they are not experimentally measurable. These true quantities are exclusive which means that for example N_{RR} does not contain any other components like N_{RF}. The vector on the left hand side of the equation is exclusive, but measurable, which means for example that N_{TT} does not contain components of N_{TL} as well. In Eq. 5.3, the fake and the real rate are represented by the coefficient f and r respectively and they are determined as follows:

$$f = \frac{N_{tight}^{fake}}{N_{loose}^{fake}} \qquad\qquad r = \frac{N_{tight}^{real}}{N_{loose}^{real}} \tag{5.4}$$

Here, N_{tight}^{real} (N_{loose}^{real}) is the number of real electrons passing the $tight$ ($loose$) selection. N_{tight}^{fake} (N_{loose}^{fake}) give the corresponding numbers for fake electrons. So, the real rate stands for the probability for a real electron to be reconstructed as "tight"-electron and the fake rate denotes the probability for a fake selected as loose to be reconstructed as tight. The objects passing the loose selection need to pass the trigger used in the analysis, while the tight objects have to fulfil all cuts of the loose selection and additionally the $lhMedium$ identification and an isolation cut. N_{TT}, which represents the number of events reconstructed as signal but originated from a pair of objects containing at least one fake, is our quantity of interest. This is described in Eq. 5.5 following the first line of Eq. 5.3.

$$\begin{aligned} N_{TT}^{l+jets} &= rf(N_{RF} + N_{FR}) \\ N_{TT}^{di-jet} &= f^2 N_{FF} \\ N_{TT}^{l+jets\&di-jet} &= rf(N_{RF} + N_{FR}) + f^2 N_{FF} \end{aligned} \tag{5.5}$$

Eq. 5.5 depends on inaccessible truth quantities. These true variables can be expressed as a function of measurable quantities by inverting the matrix in Eq. 5.3:

$$\begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix} = \frac{1}{(r-f)^2} \begin{pmatrix} (f-1)^2 & (f-1)f & f(f-1) & f^2 \\ (f-1)(1-r) & (1-f)r & f(1-r) & fr \\ (r-1)(1-f) & (1-r)f & r(1-f) & fr \\ (1-r)^2 & (r-1)r & r(r-1) & r^2 \end{pmatrix} \begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} \tag{5.6}$$

The combined fake contributions to the signal selection is given by,

$$\begin{aligned} N_{TT}^{l+jets\&di-jets} = \alpha[2rf(f-1)(1-r) + f^2(1-r)^2]N_{TT} \\ +\alpha f r^2(1-f)(N_{TL}+N_{LT}) \\ +\alpha f^2 r^2 N_{LL} \end{aligned} \tag{5.7}$$

where

$$\alpha = \frac{1}{(r-f)^2}. \tag{5.8}$$

5.4.2 Real rate determination

A sample of real electrons is need for the determination of the real rate r. This rate is determined from DY simulation and it is calculated in bins of p_T and $|\eta|$. Fig. 5.3 shows this determination. The real rate is found to be greater than $\approx 90\%$

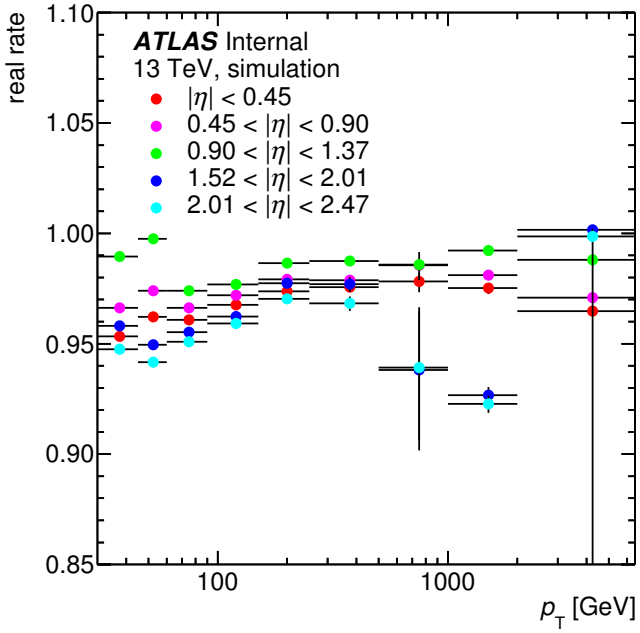


Fig. 5.3: Real efficiencies of electrons (Eq. 5.4) as a function of p_T in different $|\eta|$ regions. The statistical uncertainty is displayed by the vertical bar.

in the overall p_T range and rising at higher p_T.

5.4.3 Fake efficiency determination

The fake rate is calculated using fake electrons in data. A fake enriched sample is obtained by requiring events which satisfy a single electron trigger with a $loose$ identification criteria ($lhLoose$) or a $very\ loose$ one ($lh\nu Loose$). To reduce the real electron contamination in this sample, a cut on $E_T^{miss} > 25$ GeV is imposed to suppress real electrons from W decays. Furthermore, in order to suppress real electrons from DY events, a cut on the reconstructed invariant mass, m_{reco}, is

imposed, vetoing events with $|m_Z^{reco} - m_{reco}| < 20$ GeV. Despite the above-described cuts, the contribution of real electrons cannot be neglected. These contributions are estimated with DY, $W + jets$, Di-boson and Top MC simulation and subtracted from N_{loose}^{fake} and N_{tight}^{fake}. These rates are shown in Fig. 5.4 and they have been determined as a function of p_T and $|\eta|$. The fake rates are $\approx 20\%$ in the low p_T

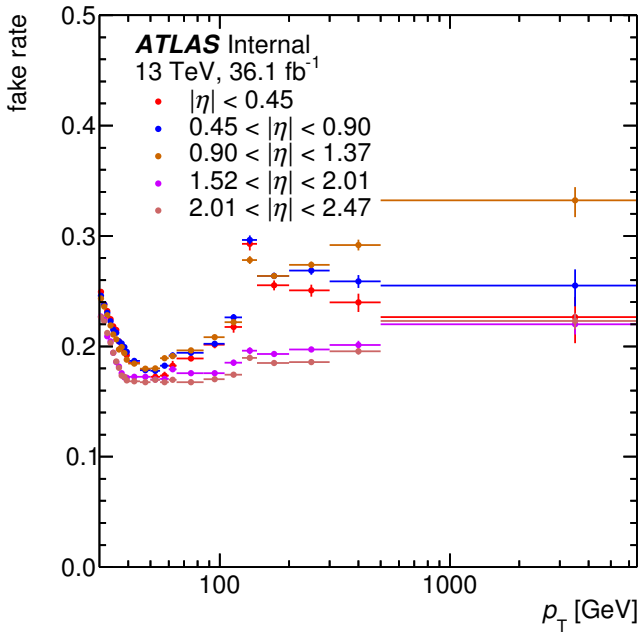


Fig. 5.4: Fake rate of electrons as a function of p_T and $|\eta|$.

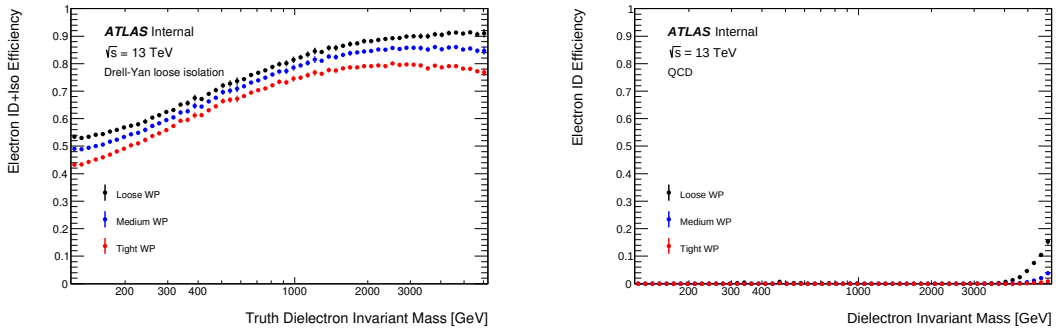
region and they start increasing from around 100 GeV in the low $|\eta|$ region.

5.5 Electron identification

The choice of the electron identification criteria has been optimised comparing the different likelihood-based working points (WPs) provided by the e/γ performance group, namely: Loose, Medium and Tight. The signal efficiency was tested on the SM DY MC, while the background rejection was tested on di-jet MC: Fig. 5.5 shows these efficiencies as a function of the di-lepton invariant mass m_{ll}. This study shows that "*Likelihood Loose*" and "*Likelihood Medium*" provide a similar performance in terms of signal efficiency compared to background rejection for the di-lepton search. Ultimately, the Medium working point is chosen for the nominal event selection, given the fact that it makes the analysis slightly more robust against fake electrons.

5.6 Systematic Uncertainties

A flat uncertainty of 3.2% is assessed due to the uncertainty on the luminosity scale. Mass-dependent systematic uncertainties are considered as nuisance pa-



(a) (b)

Fig. 5.5: Full selection efficiency for various working points as applied to (a) DY MC as a function of di-lepton truth mass. The same study is performed for (b) di-jet MC as a function of di-lepton reconstructed mass.

rameters and include both the theoretical and experimental effects on the total background and experimental effects on the signal. They are assessed as a function of the di-lepton invariant mass and for each source of systematic uncertainty the largest of the up/down variations is taken. This variation is then treated as being a symmetric up/down variation versus di-lepton invariant mass, and is the quantity which is used in the statistical interpretation.
The systematic uncertainties related to electron energy/momentum scale and resolution, as well as the identification, trigger, reconstruction and isolation efficiency, were provided by the e/γ group official recommendations, as described in Sec. 5.3. The data-driven estimate of the number of events containing at least one fake electron which mimics the signal (as described in Sec. 5.4) has a relatively large uncertainty at high di-lepton invariant masses. The uncertainty on this estimate is assessed by variation of the methodology and calculated rate of fake electrons. To cover these effects, the systematic uncertainty on the data-driven estimate is calculated by scaling the default fake background estimate by $\pm 1\sigma$ and taking the ratio between this and the nominal estimate when included in the total background estimate.
Theoretical uncertainties in the background prediction are dominated by the DY background in this search. They arise from the PDF eigenvector variations of the nominal PDF set, as well as variations of PDF scale, α_S, EW corrections, and photon-induced (PI) corrections. The effects of different PDF set choices are also considered. The PDF variation uncertainty is obtained using the 90% confidence level CT14NNLO PDF error set. Rather than using a single nuisance parameter to describe the 28 eigenvectors of this PDF error set, which could lead to an underestimation of its effect, a re-diagonalised set of 7 PDF eigenvectors was used, which are treated as separate nuisance parameters. The sum in quadrature of these eigenvectors matches the original CT14NNLO error envelope well. The uncertainties due to the variation of PDF scale and α_S are derived using `VRAP` with the former obtained by varying the renormalisation (μ_R) and factorisation (μ_F) scales of the nominal CT14NNLO PDF up and down simultaneously by a factor of two.

The value of α_S used (0.118) is varied by $\pm$ 0.003. An additional uncertainty is derived due to the choice of nominal PDF set, by comparing the central values of CT14NNLO with those from other PDF sets as recommended by the PDF4LHC forum [191], namely MMHT14 [192] and NNPDF3.0 [50]. The maximum absolute deviation from the envelope of these comparisons is used as the PDF choice uncertainty, where it is larger than the CT14NNLO PDF eigenvector variation envelope. Theoretical uncertainties are not applied to the signal prediction.
Theoretical uncertainties in the $t\bar{t}$ and di-boson backgrounds were also considered. The $t\bar{t}$ MC sample is normalised to a cross-section of $\sigma_{t\bar{t}} = 832^{+20}_{-29}$ (scale) $\pm$ 35 (PDF + α_S) pb, with the `Top++` program, as described in Section 5.1. The first uncertainty comes from the independent variation (up and down simultaneously by a factor of two) of the factorisation and renormalisation scales, μ_F and μ_R, while the second one is associated to variations in the PDF and α_S, following the PDF4LHC prescription. Normalisation uncertainties in the top quarks and di-boson background were found to be negligible. The uncertainties in the top-quark and di-boson background extrapolations are estimated by varying both the functional form and the fit range, taking the envelope of all variations. These normalisation uncertainties were also found to be negligible with respect to the total background estimate.
The EW correction systematic uncertainty was assessed by comparing the additive versus factored computation of the EW k_F, where the additive computation is used as the nominal in this search (for the definition of these two methods, please see Sec. 5.1.1). The size of the uncertainty is ≈4% at di-lepton invariant masses of 3 TeV.
Systematic uncertainties used in the statistical analysis of the results are summarised in Table 5.1 at di-electron mass values of 2 TeV and 4 TeV.

5.7 Event yields

Expected and observed event yields, in bins of invariant mass, are shown in Table 5.2. Expected event yields are split into the different background sources and the yields for two signal scenarios. The DY process is dominant over the entire mass range. In general, the observed data are in good agreement with the SM prediction, taking uncertainties into account, as described in the previous section. The di-electron invariant mass distributions, m_{ee}, are shown in Fig. 5.6, as well as the E_T, η and ϕ of both selected electrons in the signal region (m_{ee} > 120 GeV) in Fig. 5.7. No significant excess is observed. The highest invariant mass event is found at m_{ee} = 2.90 TeV in the di-electron channel; this event appear to be very clean with little other detector activity, as it is shown in Fig. 5.8.

5.8 Statistical analysis

Two complementary approaches are used to quantify the significance of an excess between the expected and observed data, namely: a log-likelihood ratio (LLR) test, which utilises signal templates for the different models under investi-

Source	di-electron	
	Signal	Background
Luminosity	3.2% (3.2%)	3.2% (3.2%)
MC statistical	< 1.0% (< 1.0%)	< 1.0% (< 1.0%)
Beam energy	2.0% (4.1%)	2.0% (4.1%)
Pile-up effects	< 1.0% (< 1.0%)	< 1.0% (< 1.0%)
PDF Choice	N/A	< 1.0% (8.4%)
PDF Variation	N/A	8.7% (19%)
PDF Scale	N/A	1.6% (2.7%)
α_S	N/A	1.6% (2.7%)
EW corrections	N/A	2.4% (5.5%)
Photon-induced corrections	N/A	3.4% (7.6%)
Top quark theoretical	N/A	< 1.0% (< 1.0%)
Di-boson theoretical	N/A	< 1.0% (< 1.0%)
Reconstruction efficiency	< 1.0% (< 1.0%)	< 1.0% (< 1.0%)
Isolation efficiency	9.1% (9.7%)	9.1% (9.7%)
Trigger efficiency	< 1.0% (< 1.0%)	< 1.0% (< 1.0%)
Identification efficiency	2.6% (2.4%)	2.6% (2.4%)
Electron energy scale	< 1.0% (< 1.0%)	4.1% (6.1%)
Electron energy resolution	< 1.0% (< 1.0%)	< 1.0% (< 1.0%)
Multi-jet & W+jets	N/A	10% (129%)
Total	10.1% (11.0%)	18.1% (132.0%)

Table 5.1: Summary of the relative systematic uncertainties in the expected number of events at a di-electron mass of 2 TeV (4 TeV). The values quoted for the uncertainty represent the relative change in the total expected number of events in the bin containing the reconstructed m_{ee} mass of 2 TeV (4 TeV). For the signal uncertainties the values were computed using a Z'_χ signal model with a pole mass of 2 TeV (4 TeV) by comparing yields in the core of the mass peak (within the full width at half maximum) between the distribution varied by the considered uncertainty and the nominal distribution. The total uncertainty quoted on the last line is obtained from a sum in quadrature of the individual uncertainties. "N/A" represents cases where the uncertainty is not applicable.

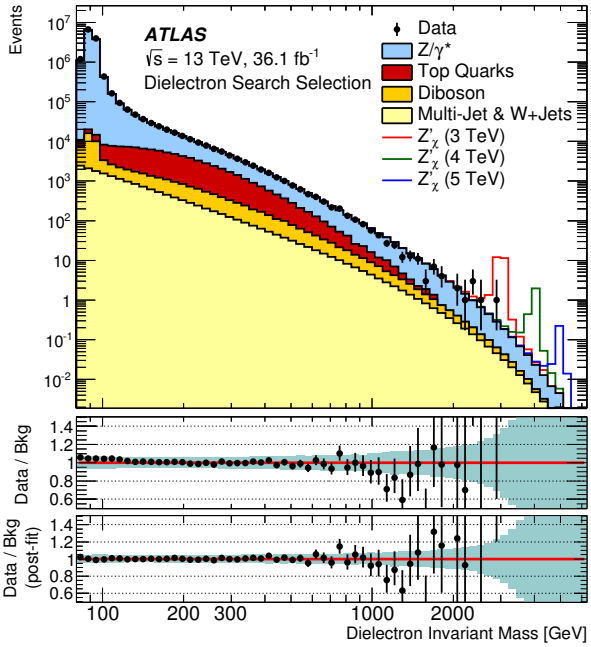
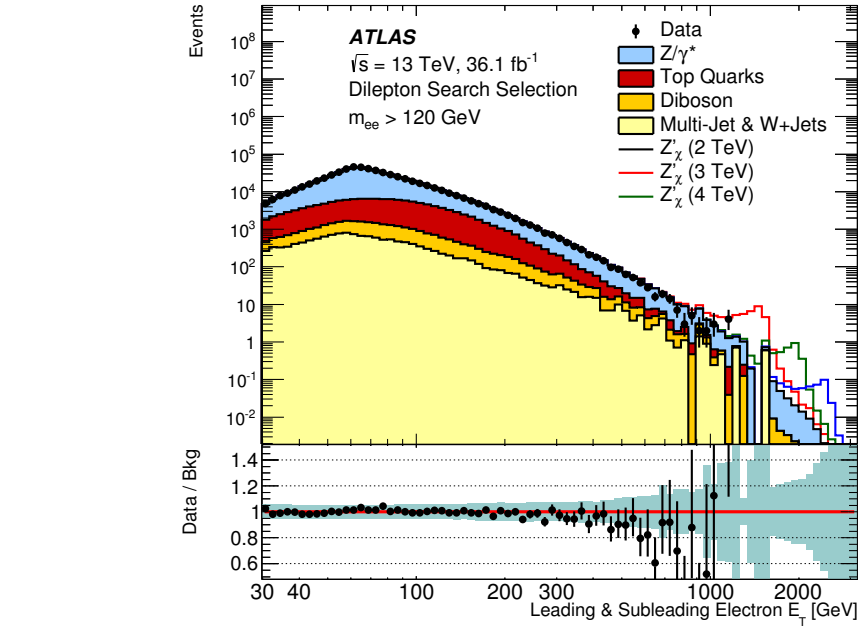
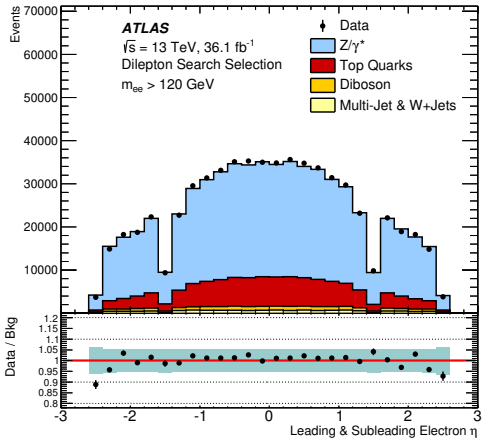


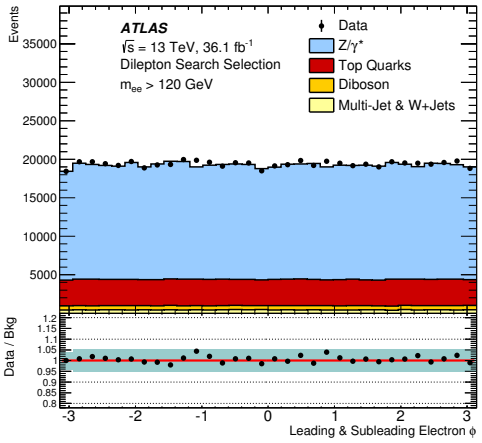
Fig. 5.6: Reconstructed di-electron invariant mass distributions after selection. Data and the SM background estimates as well as their ratio are displayed, along with three selected resonant signals overlaid, Z'_χ with a pole mass of 3,4 and 5 TeV. The bin width of the distributions is constant in $\log(m_{ee})$, and the shaded band in the lower panel illustrates the total systematic uncertainty, as explained in Section 5.6. The data points are shown together with their statistical uncertainty. The MC expectation is normalised to the expected integrated luminosity. The middle panel shows the ratio of the data to the expected background. The lower panel shows the ratio of the data to the adjusted expected background ("post-fit") that results from the statistical analysis.



(a)



(b)



(c)

Fig. 5.7: Reconstructed electron (a) E_T (or p_T), (b) η and (c) ϕ after selection. "Leading" denotes the electron with the highest E_T (P_T) of the selected pair, while "Sub-leading" denotes the second electron. The distributions are shown for di-electron invariant masses greater than 120 GeV. Data and the SM background estimates as well as their ratio are displayed. The shaded band in the lower panel illustrates the total systematic uncertainty, as explained in Section 5.6. The data points are shown together with their statistical uncertainty and the MC expectation is normalised to the expected integrated luminosity.

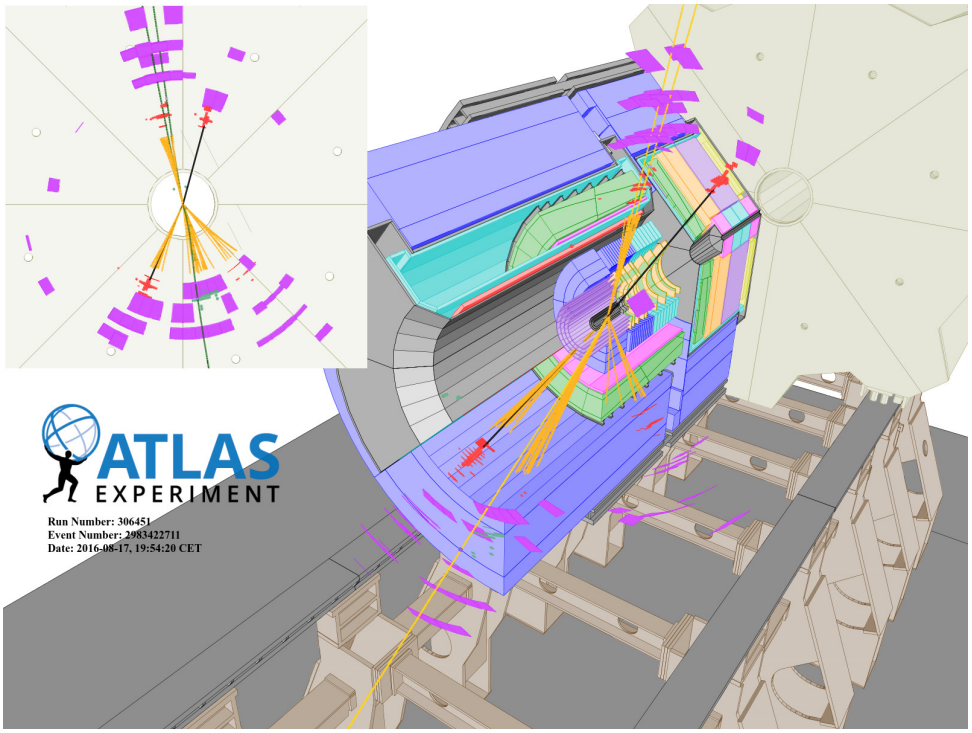


Fig. 5.8: Event display for the selected event with the highest di-electron invariant mass. Red lines represent the tracks and energy deposits in the cells of the electromagnetic calorimeter for the two selected electron candidates. Yellow lines correspond to tracks from secondary interactions. Purple boxes indicate energy deposition in tile cells. Candidate leptons with $E_T > 30$ GeV are displayed. The highest momentum electron has an E_T of 889 GeV and an η of -0.51. The sub-leading electron has an E_T of 868 GeV and an η of 1.14. The invariant mass of the pair is 2.90 TeV.

m_{ee} [GeV]	80–120	120–250	250–400	400–500	500–700
Drell–Yan	11 800 000 ± 700 000	216 000 ± 11 000	17 230 ± 1000	2640 ± 180	1620 ± 120
Top quarks	28 600 ± 1800	44 600 ± 2900	8300 ± 600	1130 ± 80	560 ± 40
Di-bosons	31 400 ± 3300	7000 ± 700	1300 ± 140	228 ± 25	146 ± 16
Multi-jet & W+jets	11 000 ± 9000	5600 ± 2000	780 ± 80	151 ± 21	113 ± 17
Total SM	11 900 000 ± 700 000	273 000 ± 12 000	27 600 ± 1100	4150 ± 200	2440 ± 130
Data	12 415 434	275 711	27 538	4140	2390
Z'_χ (4 TeV)	0.00635 ± 0.00021	0.0390 ± 0.0015	0.0564 ± 0.0025	0.0334 ± 0.0027	0.064 ± 0.004
Z'_χ (5 TeV)	0.00305 ± 0.00012	0.0165 ± 0.0006	0.0225 ± 0.0010	0.0139 ± 0.0007	0.0275 ± 0.0015

m_{ee} [GeV]	700–900	900–1200	1200–1800	1800–3000	3000–6000
Drell–Yan	421 ± 34	176 ± 17	62 ± 7	8.7 ± 1.3	0.34 ± 0.07
Top quarks	94 ± 8	27.9 ± 2.8	5.1 ± 0.7	< 0.001	< 0.001
Di-bosons	39 ± 4	16.9 ± 2.1	5.8 ± 0.8	0.74 ± 0.11	0.028 ± 0.004
Multi-jet & W+jets	39 ± 6	16.1 ± 2.0	7.9 ± 2.3	1.6 ± 1.2	0.08 ± 0.27
Total SM	590 ± 40	237 ± 17	81 ± 7	11.0 ± 1.8	0.45 ± 0.28
Data	589	209	61	10	0
Z'_χ (4 TeV)	0.0585 ± 0.0035	0.074 ± 0.005	0.121 ± 0.011	0.172 ± 0.017	2.57 ± 0.27
Z'_χ (5 TeV)	0.0218 ± 0.0013	0.0295 ± 0.0021	0.040 ± 0.004	0.040 ± 0.004	0.280 ± 0.030

Table 5.2: Expected and observed event yields in the di-electron channel in different di-lepton mass intervals. The quoted errors correspond to the combined statistical, theoretical, and experimental systematic uncertainties. Expected event yields are reported for the Z'_χ model, for two values of the pole mass.

gation, providing local and global significances, and a BumpHunter/TailHunter Tool, which effectively scans the di-lepton invariant mass spectrum for any local resonant/non-resonant signal-like excess, providing the corresponding global significance.

5.8.1 Likelihood-ratio test

The probability that the background fluctuates creating a signal-like excess equal to or larger than what is observed defines the importance of an excess and it is denoted as p_0 (p-value). p_0 is computed using a test statistic defined as follows:

$$q_0 = \begin{cases} 0 & \text{for } \hat{\mu} < 0, \\ -2\ln\left[\frac{\mathcal{L}(\text{data}|0,\hat{\theta}_0)}{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta})}\right] & \text{for } \hat{\mu} \geqslant 0, \end{cases} \tag{5.9}$$

where $\hat{\theta}_0$ represents the nuisance parameters for the null (no signal-like) hypothesis and $\hat{\theta}$ for a given μ value (so called *signal strength*). The parameter μ is defined as a ratio of the signal production cross-section times branching ratio in the di-electron final state ($\sigma \cdot B$) to its theoretically predicted value.
The more important an excess, the larger the difference between the two likelihoods considered in this equation and the larger is the corresponding q_0 value. The probability p_0 corresponding to a given experimental observation q^{obs} is evaluated as follows:

$$p_0 = P(q_0 \geqslant q_0^{obs}|\text{background-only}) = \int_{q_0^{obs}}^{\infty} f(q_0|0,\hat{\theta}_0^{obs})dq_0 \tag{5.10}$$

Usually, the p_0 values are translated to a scale of *significance* (Z), or number of standard deviations. It is done by using the inverse of the cumulative function for the normal Gaussian Φ [193],

$$Z = \Phi^{-1}(1 - p_0) \tag{5.11}$$

For example, in the background-only hypothesis, the probability to have an excess with 5σ significance ($Z = 5$) is $p_0 = 3 \cdot 10^{-7}$; 5σ is conventionally the significance required for claiming the observation of a new particle.
Moreover, to account for the "look elsewhere effect", we report the global *p*-value that is the probability of observing a fluctuation somewhere in the entire background invariant mass spectrum which is at least as significant as the smallest local *p*-value measured in data. The global *p*-value accounts for the increased possibility of observing positive background fluctuations when looking for signals at a range of masses. Extensive cross-checks were performed over the whole mass spectrum, and all the deficits or excesses were quantified by calculating the local Poisson *p*-value, using the sources of systematic uncertainty described in Section 5.6, which gives a significance less than two standard deviations for each mass interval. The distribution of the observed *p*-value for a Z'_χ in the di-electron channel, as well as the one for the combined channel (ee and $\mu\mu$) is presented in Fig. 5.9.
To perform this LLR search, the HistFactory package [194], together with RooStats [195] and RooFit [196] packages are used.

5.8.2 BumpHunter

The BumpHunter [197] method quantifies the probability to observed an excess in the observed data from the background prediction. The strength of this method is in the fact that a precise model is not necessary and it relies just on the assumption that signal will present as an excess over the background prediction in several contiguous bins of the histogram of the variable of interest - in our case the di-lepton mass. The minimal width of search interval is set to two bins and the maximum search interval is set to half the number of bins in the input distribution: these are the defaults search windows explored in this search. The *p*-value is computed in each of these intervals.
The di-electron invariant mass distribution with the observed data overlayed on the combined background prediction as well as the local significance (lower panel) is shown in Fig. 5.10, where a pair of blue lines indicates the interval with the smallest *p*-value to observe as many or more events as actually observed given the background expectation. The common ATLAS BumpHunter implementation is used in this analysis.

5.9 Results

The statistical tests described in the previous section do not reveal a signal. So, in the absence of any significant excess, limits are set on the respective signal model's parameter of interest, Θ. For each signal region the total number of

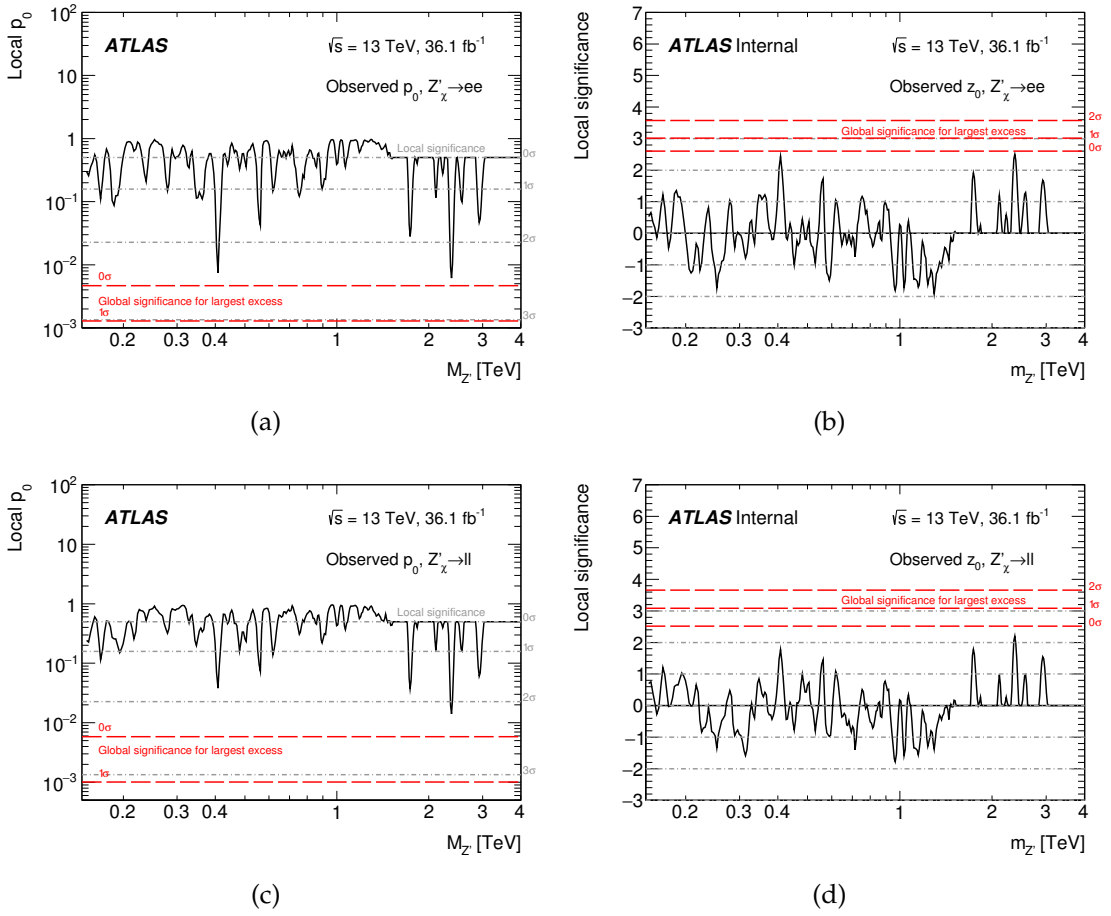


Fig. 5.9: The local (left-hand column) p-value and (right-hand column) significances derived from ideal Z'_χ signals between 0.15 and 3.0 TeV. Shown are results for the (upper) di-electron and (lower) di-lepton channels. Also shown are local (dot-dashed line) and global (broad dashed line) significance levels, where the global levels apply only to the most significant local excess. The binning used in each channel is optimised for the mass resolution.

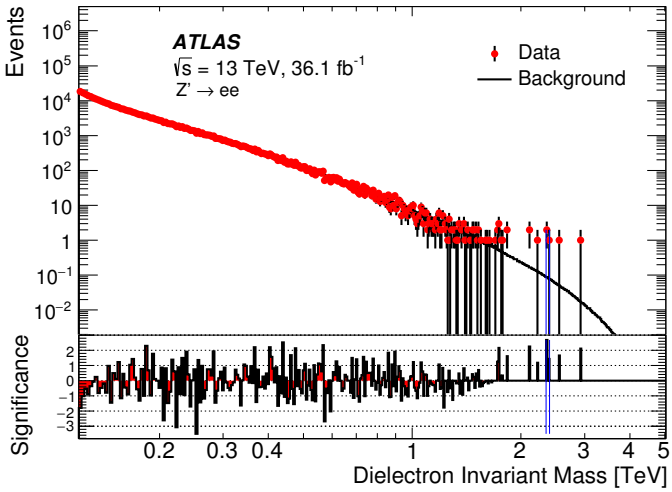


Fig. 5.10: Di-lepton mass distribution in the di-electron channel, showing the observed data, combined background prediction, and corresponding bin-by-bin significance.

expected events μ is given by Eq. 5.12,

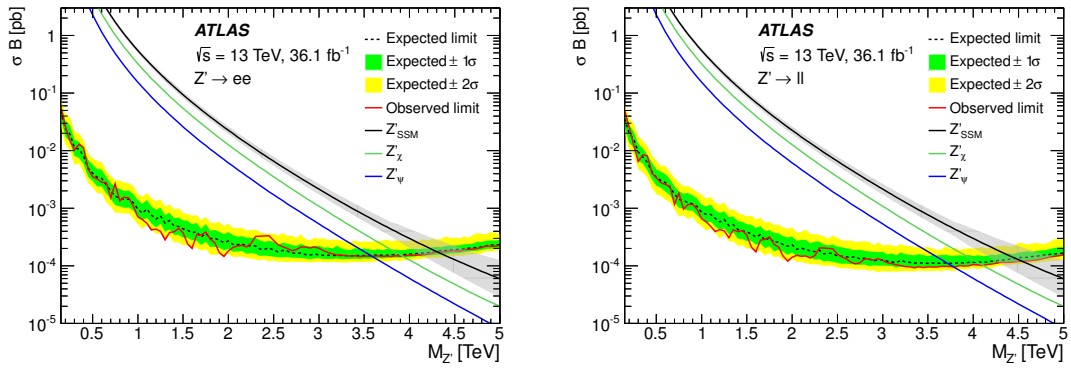
$$\mu = n_s(\Theta, \overline{\Omega}) + n_b(\overline{\Omega}) \tag{5.12}$$

where the predicted number of events by the `Pythia8` Z' or CI MC for a particular choice of model parameter Θ is represented by $n_s(\Theta, \overline{\Omega})$. The total number of events of background events is $n_b(\overline{\Omega})$ and in both cases $\overline{\Omega}$ represents the set of Gaussian nuisance parameters that account for systematic uncertainties on the number of respective signal and background events. The parameter Θ corresponds to a choice of the Z' pole mass in the resonance search, and the energy scale Λ and interference parameter η_{LL} in the case of the CI search.
The likelihood of observing n events given the new physics parameter Θ and nuisance parameters $\overline{\Omega}$ is given by the product of Poisson probabilities for each mass bin k, as shown in Eq. 5.13,

$$\mathcal{L}(n \mid \Theta, \overline{\Omega}) = \prod_{l=1}^{N_{channel}} \prod_{k=1}^{N_{bin}} \frac{\mu_{lk}^{n_{lk}} e^{-\mu_{lk}}}{n_{lk}!}, \tag{5.13}$$

where n_{lk} is the number of events observed in data, and $\mu_l k$ is the total number of expected events (signal plus background), both in mass bin k, and channel l. Then, the limit on n_s is converted into an upper limits on cross-section times branching ratio $\sigma(Z') \cdot B(X \to \ell^+\ell^-)$ by scaling with the observed number of Z boson events and the know value of $\sigma \cdot B(Z \to \ell^+\ell^-)$: $\sigma \cdot B(X) = \dfrac{n_S}{\mathcal{L}\mathcal{A}\epsilon(X)}$, where:

- $\mathcal{L}$ is the integrated luminosity of the data
- $\mathcal{A}\epsilon(X)$ is the total acceptance times efficiency for a given signal template for the parameter of interest X.



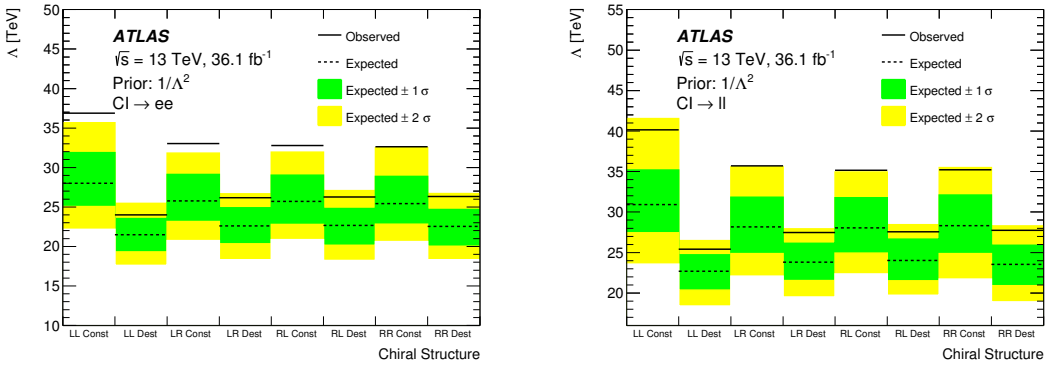
(a) (b)

Fig. 5.11: Upper 95% C.L. limits for $Z^{'}$ production cross section times branching ratio to two leptons as a function of $Z^{'}$ pole mass ($M_{Z'}$). The signal theory lines are calculated with `Pythia8` using the NNPDF23LO PDF set and corrected to NNLO in QCD using VRAP and the CT14NNLO PDF set. The signal theoretical uncertainties are shown as a band on the $Z^{'}_{SSM}$ theory line. Shown are results for: (a) the di-electron channel; (b) the combined channel ($ee, \mu\mu$).

This conversion in a variety of interference and chiral coupling scenarios. The logarithmic m_{ee} histogram binning shown in Fig. 5.6 and made up of 66 bins is chosen for setting limits on resonant signals using $Z^{'}_{\chi}$ signal templates; the prior probability is chosen to be uniform and positive in the cross-section for this limit calculation. The upper 95% percentile of the posterior probability is then quoted as the upper 95% level limit on $\sigma \cdot B$. The above-described calculations are performed with the Bayesian Analysis Toolkit (BAT) [198]. Limit values obtained using the experimental data are quoted as observed limits, while median values of the limits from a large number of pseudo-experiments, where only SM background is present, are quoted as the expected limits. The upper limits on the $\sigma \cdot B$ in a $Z^{'}$ model are set at 95% credibility limit (C.L.) and they are presented in Fig. 5.11. The upper limits on $\sigma \cdot B$ for $Z^{'}$ bosons start to weaken above a pole mass of $\sim$3 TeV. This is mainly due to the combined effect of a rapidly-falling signal cross-section as the kinematic limit is approached, and the natural width of the resonance.
Furthermore, lower limits are set at 95% C.L. on the energy scale scale Λ, for the LL, LR, RL, and RR CI model. Both constructive and destructive interference scenarios are explored, as well as priors of $1/\Lambda^2$. Limits are presented for the di-electron and combined di-lepton channels in Fig. 5.12 using a $1/\Lambda^2$ prior.

5.10 Summary of Chapter 5

The ATLAS detector at the Large Hadron Collider has been used to search for resonant and non-resonant new phenomena in the di-lepton invariant mass spectrum above the Z boson pole. This chapter presented the results of a search conducted with 36.1 fb^{-1} of pp collision data collected at $\sqrt{s}$ = 13 TeV and recorded during 2015 and 2016 data-taking period. The highest invariant mass event is



(a) (b)

Fig. 5.12: Lower limits on the energy scale Λ at 95% C.L., for the CI model with constructive ('Const' in the labels on the x-axis) and destructive ('Dest' in the labels on the x-axis) interference (for all considered helicities). Results are shown for the (a) di-electron and (b) di-lepton channels.

found at 2.90 TeV in the di-electron channel. The observed di-electron invariant mass spectrum is consistent with the Standard Model prediction, within systematic and statistical uncertainties. The data are interpreted in terms of resonant spin-1 Z' gauge boson production and non-resonant contact interactions. Upper limits are set on the cross-section times branching ratio for a spin-1 Z' gauge boson. The resulting lower mass limits are 4.5 TeV for the Z'_{SSM}, 4.1 TeV for the Z'_{χ}, and 3.8 TeV for the Z'_{ψ}. Other Z' models are also constrained in the range between those quoted for the Z'_{χ} and Z'_{ψ}. These are more stringent than the previous ATLAS result obtained at $\sqrt{s}$ = 8 TeV, by up to $\sim$ 1 TeV. In the end, lower exclusion limits are also set on the energy scale Λ for various $qq\ell\ell$ contact interaction models, ranging between 23.5 TeV and 40.1 TeV.

Chapter 6

Low mass Drell-Yan double differential cross section

This study has been produced in collaboration between myself and some members of the ATLAS collaboration. In this analysis, I was the main analyser so I contributed to all the main aspects of the analysis, apart from the studies related to the trigger efficiencies and trigger scale factors determination, performed by a student and a senior researcher from "Sapienza" University of Rome.
The cross section for the process $Z/\gamma^* \rightarrow \mu\mu$ is presented in terms of both the single-differential (1D) invariant-mass distribution, $d\sigma/dm_{\mu\mu}$, as well as double differential (2D) distributions in $m_{\mu\mu}$ and $|y_{\mu\mu}|$, namely $d^2\sigma/dm_{\mu\mu}d|y_{\mu\mu}|$. These differential cross section are measured in proton-proton collision at $\sqrt{s}$ = 13 TeV at the LHC using the ATLAS detector. The measurements are performed in the μ-channel only for invariant masses between 7 GeV and 66 GeV (with the invariant mass interval between 9 and 12 GeV excluded) using an integrated luminosity of 1.6 fb^{-1}, corresponding to the full dataset collected in 2015. The cross sections are determined within fiducial acceptance regions.
The kinematic range of the measurement can be seen in Fig. 6.1, where the region bounded in black represents the kinematic range accessible with ATLAS, for comparison with other experiments. The lower limit in mass, *M*, is fixed by the minimum thresholds for triggering low-p_T leptons, and the limits on the rapidity, *y*, are set by the pseudo-rapidity acceptance for leptons. This accessible region allows a measurement to be made between the kinematic regions accessible to the LHCb and HERA. The region of low *x*, $x \sim 10^{-4}$, is interesting phenomenologically since novel effects in perturbative QCD, such as $\ln(1/x)$ (BFKL) resummation or non-linear evolution equations may become important. Value of $x \sim 10^{-4}$ are only accessible at HERA for low scale $Q = M \simeq 3$ GeV, where non-perturbative effects may mask and new perturbative contributions. However at the LHC, this low-x region is well inside the perturbative regime. This particular region of small *x* and small *M* is interesting phenomenologically where novel QCD effects may become important.
Indeed, some tension was observed at low Q^2 such that the χ^2 for the fits to the final combined inclusive cross-section data from the HERA experiments H1 and ZEUS drops steadily as the minimum four-momentum squared Q^2_{min} of the data

entering the fit is raised up from $Q^2_{min} \approx 3.5$ GeV^2 to $Q^2_{min} \approx 10$ GeV^2 (see Fig. 19 of Ref. [34]). This turns out to be true for all perturbative orders and is not mitigated by going to higher order. In particular, the χ^2 of the NNLO fits is not better than the NLO fit for low values of Q^2_{min}.
A further observation is that the increased χ^2 of the fits to the low-Q^2 data is largely attributable to the kinematic region of low x and high *inelasticity* y (where $y = Q^2/sx$ and $\sqrt{s}$ is the centre-of-mass energy) in the neutral-current reduced cross section data σ_{red}, defined as:

$$\sigma_{\mathrm{red}} = F_2 - \frac{y^2}{Y_+} F_L \,, \tag{6.1}$$

where, F_2 and F_L are the structure functions, which are related to the parton distributions [199], and $Y_+ = 1 + (1-y)^2$. In this kinematic region (low Q^2 and low x) the data take a turn-over (see *e.g.* Fig. 59 of Ref. [34]). This effect can be ascribed to the negative term proportional to F_L in Eq. (6.1). However, fits to data using fixed-order pQCD do not describe this turn-over very well, suggesting that a larger F_L is needed for a better description.
Recently, an alternative approach which can improve the description of low-Q^2 data has been proposed. Since the kinematics of HERA is such that low-Q^2 data is also at low x, it has been suggested that the DGLAP resummation of $\ln Q^2$ terms should be augmented by $\ln(1/x)$ (BFKL) resummation [142, 200]. It is interesting to investigate if such effects are evident in the low-x region probed by ATLAS.

6.1 Data and MC samples

As regards the ATLAS software release, the 20.7 one was used for reconstructing both MC events and the data. The detector response is simulated with `GEANT 4` [173, 174].
The data used for this measurement were recorded between August 16 and November 03, 2015, during data periods D-J. In these periods, the LHC circulated 6.5 TeV proton beams with 25 ns bunch spacing. In Periods D-J, the peak instantaneous luminosity was $\mathcal{L}$ = 5.1 $\cdot$ 10^{33} cm^{-1} s^{-1}, and the average number of pileup events was $< \mu >$ = 13.5 (for the definition of pileup please refer to Sec. 2.3). The runs corresponding at the full 25 ns dataset are found in the range 276262-284484. The data collection is separated into periods in which the trigger conditions remained stable.
A Good Run List (GRL) consists of a list of luminosity blocks, spanning 1-2 minutes of data-taking, in which the LHC is circulating stable colliding beams and all critical detector components are functioning properly. In particular, the inner detector, the electromagnetic and hadronic calorimeters, and the muon spectrometers must deliver data of high quality to ensure that electrons, muons, jets, and missing transverse energy are measured accurately. The data sample is selected using the following GRL: `data15_13TeV.periodAllYear_DetStatus-v79-repro20-02_DQDefects-00-02-02_PHYS_StandardGRL_All_Good_25ns.xml` and the

MC ID	Process	Generator	Mass range [GeV]	Events	Filter	ϵ_{filter}	σ [pb]
361107	$Z/\gamma^* \to \mu^+\mu^-$	Powheg+Pythia8	$60 < m_{\mu^+\mu^-}$	20M	-	1.0	1901
361108	$Z/\gamma^* \to \tau^+\tau^-$	Powheg+Pythia8	$60 < m_{\tau^+\tau^-}$	20M	-	1.0	1901
361666	$Z/\gamma^* \to \mu^+\mu^-$	Powheg+Pythia8	$6 < m_{\mu^+\mu^-} < 10$	7M	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	$5.45{\cdot}10^{-2}$	13960
361667	$Z/\gamma^* \to \mu^+\mu^-$	Powheg+Pythia8	$10 < m_{\mu^+\mu^-} < 60$	10M	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	$2.749{\cdot}10^{-1}$	6621
361668	$Z/\gamma^* \to \tau^+\tau^-$	Powheg+Pythia8	$6 < m_{\tau^+\tau^-} < 10$	983800	$p_T^{\tau 1,2} > 3.5$ GeV, $\lvert\eta^{\tau 1,2}\rvert < 2.7$	$5.93{\cdot}10^{-4}$	13960
361669	$Z/\gamma^* \to \tau^+\tau^-$	Powheg+Pythia8	$10 < m_{\tau^+\tau^-} < 60$	473000	$p_T^{\tau 1,2} > 3.5$ GeV, $\lvert\eta^{\tau 1,2}\rvert < 2.7$	$1.51{\cdot}10^{-2}$	6622
363675	Dd $\gamma\gamma \to \mu^+\mu^-$	Pythia8	$6 < m_{\mu^+\mu^-} < 18$	500000	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	$7.16{\cdot}10^{-2}$	109.31
363676	Dd $\gamma\gamma \to \mu^+\mu^-$	Pythia8	$18 < m_{\mu^+\mu^-} < 60$	750000	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	$2.66{\cdot}10^{-1}$	17.41
363677	Dd $\gamma\gamma \to \mu^+\mu^-$	Pythia8	$60 < m_{\mu^+\mu^-} < 200$	200000	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	$3.43{\cdot}10^{-2}$	1.79
363686	$\gamma\gamma \to \mu^+\mu^-$	Herwig++	$6 < m_{\mu^+\mu^-} < 18$	500000	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	$1.81{\cdot}10^{-1}$	69.67
363687	$\gamma\gamma \to \mu^+\mu^-$	Herwig++	$18 < m_{\mu^+\mu^-} < 60$	750000	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	$3.30{\cdot}10^{-1}$	9.98
363688	$\gamma\gamma \to \mu^+\mu^-$	Herwig++	$60 < m_{\mu^+\mu^-} < 200$	200000	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	$2.92{\cdot}10^{-1}$	0.63
363697	Sd $\gamma\gamma \to \mu^+\mu^-$	LPair	$6 < m_{\mu^+\mu^-} < 18$	500000	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	1.0	17.28
363698	Sd $\gamma\gamma \to \mu^+\mu^-$	LPair	$18 < m_{\mu^+\mu^-} < 60$	750000	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	1.0	7.34
363699	Sd $\gamma\gamma \to \mu^+\mu^-$	LPair	$60 < m_{\mu^+\mu^-} < 200$	200000	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	1.0	0.58
363833	$b\bar{b} \to \mu^+\mu^-$	Pythia8B	$6 < m_{\mu^+\mu^-}$	15M	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.5$	$4.52{\cdot}10^{-4}$	$3.39\text{E}{\cdot}10^{8}$
363834	$c\bar{c} \to \mu^+\mu^-$	Pythia8B	$6 < m_{\mu^+\mu^-}$	2.3M	$p_T^{\mu 1,2} > 3.5$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.5$	$4.22{\cdot}10^{-5}$	$5.19{\cdot}10^{8}$
410503	$t\bar{t}$	Powheg+Pythia8	-	10M	$p_T^{\ell} > 2.0$ GeV	1.0	76.93
424102	$\Upsilon_{1S} \to \mu^+\mu^-$	Pythia8B+Photos++	-	5M	$p_T^{\mu 1,2} > 4.0$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	1.0	474610
424105	$\Upsilon_{2S} \to \mu^+\mu^-$	Pythia8B+Photos++	-	5M	$p_T^{\mu 1,2} > 4.0$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	1.0	172180
424106	$\Upsilon_{3S} \to \mu^+\mu^-$	Pythia8B+Photos++	-	5M	$p_T^{\mu 1,2} > 4.0$ GeV, $\lvert\eta^{\mu 1,2}\rvert < 2.7$	1.0	205890

Table 6.2: Cross sections σ (before applying filters) and numbers of fully simulated MC events for each samples used in the analysis. The MC simulation filter is an event selection at the generator level. The corresponding filter efficiencies are given in the table. MC generators used to produce the MC events are also indicated with the appropriate MC ID run number. 'Sd' and 'Dd' means single- and double-proton dissociative reaction respectively.

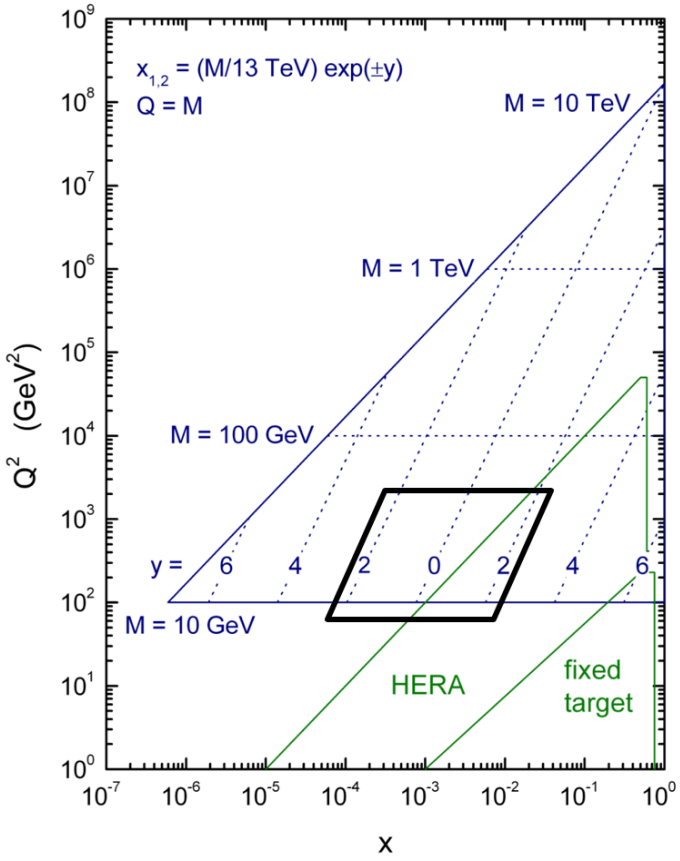


Fig. 6.1: The kinematic coverage of the LHC and HERA experiments. The region of the low mass DY measurement is shown bordered in black.

corresponding lumi tag OflLumi-13TeV-004 is used. The uncertainty on the integrated luminosity is $\pm$ 2.1 %. It is derived, following a methodology similar to that detailed in Ref. [201], from the final calibration of the luminosity scale using *x*-*y* beam-separation scans performed in August 2015.
Simulated event samples from the *MC15c* Monte Carlo production campaign are used. The derivations used are created with release 20.7.5.1.1 for the *MC15c* productions using the setup for 25 ns bunch spacing. These above-mentioned derivations are characterised by the `r7773` tag, which corresponds to the μ-profile of 2015 data only.
Nearly all of the signal and the electroweak background samples are generated with the `Powheg` Monte Carlo (MC) program [177, 202, 203]. Specifically the codes for single-boson production [45], interfaced with the `Pythia v8.1` parton shower program [178]. The programs used the CT10 parton distribution functions [179] (PDF) and the AZNLO CTEQ6L1 tune for `Powheg+Pythia` [204]. This tune was constructed specifically to describe the low-end of the Z-boson p_T spectrum better than achievable with fixed-order or analytically-resummed calculation. A more detailed description is provided in Ref. [175]. The Drell-Yan (DY) event yields are corrected with a rescaling that depends on the di-muon invariant mass from next-to-leading order (NLO) to next-to-next-to-leading order (NNLO) in the strong coupling constant α_S, computed with `VRAP v0.9` [182] and the CT14NNLO PDF

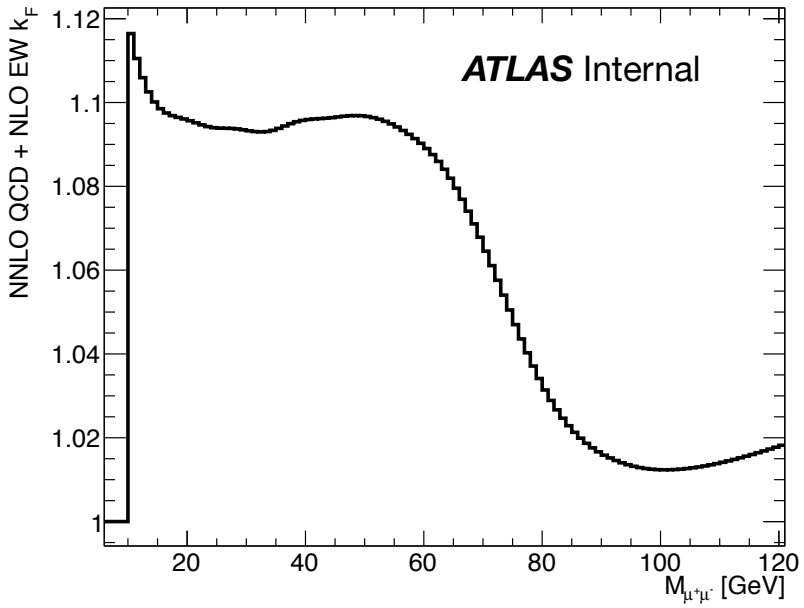


Fig. 6.2: Higher-Order corrections to Neutral Current Drell-Yan, namely: NLO-to-NNLO QCD correction with respect to CT10 generated with `Powheg+Pythia8`, including LO-to-NLO EW correction using CT14. The nominal higher-order PDF which is used to compute the final kFs is CT14NNLO.

set [183]. Two-loops higher order electroweak (EW) corrections, computed with `MCSANC` [184], are also applied. Fig. 6.2 presents the above-mentioned corrections versus di-muon truth mass at Born level[1]. Z bosons decaying to $\tau^+\tau^-$ are also considered as background processes, but in this case no mass-dependent *k*-factors are applied. For invariant masses below 10 GeV, any k_F is applied because the higher-order calculations are not available in that low-mass regime[2].

In order to check the correct normalisation between the various DY signal samples, the invariant mass of the di-muon system at Born level has been studied and it is shown in Fig. 6.3. Here, the following cuts are applied at Born generator level:

- $p_T^{truth} > 4.5$ GeV;
- $|\eta^{truth}| \leq 2.4$,

to match the kinematic cuts applied in the event selection exactly. A smooth transition between the three samples is clearly visible.

The $t\bar{t}$ background was generated at NLO using the `Powheg-Box v2` generator [205] and `Pythia v8.1` [178] (A14 tune [189]), and the production cross section (σ) is normalised to the one calculated at NNLO+NNLL with the `Top++ v2.0` program [188]. For this sample, the `Powheg` model parameter h_{damp}, which controls matrix element (ME) to parton shower (PS) matching in `Powheg` and effectively regulates the high-p_T radiation, was set to 278.75 GeV. A di-lepton filter

[1] Particle-level kinematics before final state radiation (FSR).

[2] `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HigherOrderCorrections2015`

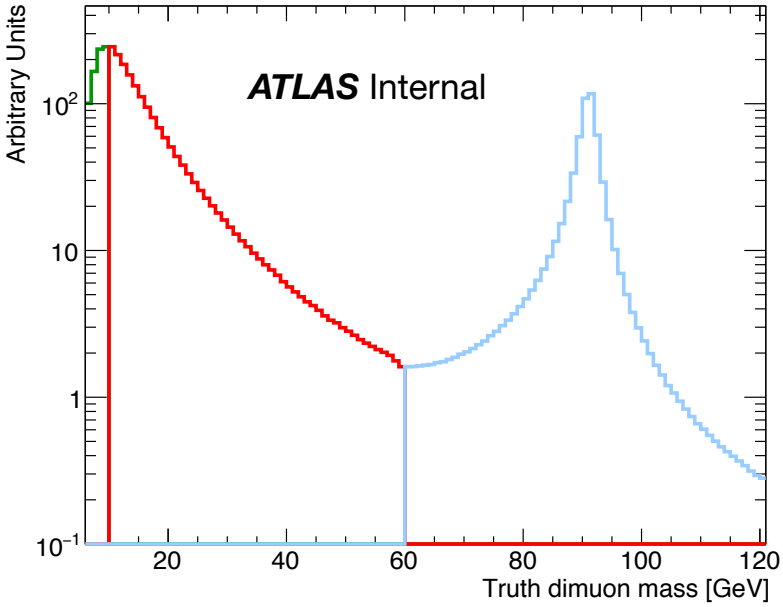


Fig. 6.3: Di-muon invariant mass distribution at Born level. Green solid line: DY sample covering the 6 < $m_{\ell\ell}$ < 10 GeV region; red solid line: DY sample covering 10 < $m_{\ell\ell}$ < 60 GeV; Light blue solid line: Z sample covering the $m_{\ell\ell}$ > 60 GeV region.

is also applied.
The event samples for exclusive $\gamma\gamma \rightarrow \ell^+\ell^-$ production are generated using `Herwig++ v2.5.1` [206] event generator. The dominant background samples for PI single-proton dissociative di-lepton production are generated using `Lpair v4.0` [207]. Depending on the photon virtuality, the proton dissociation in `Lpair` is modelled using Suri-Yenni [208] or Brasse [209] structure functions. The `Lpair` package is interfaced to `Jetset v7.408` [210], where the `Lund` [211] fragmentation model is implemented. The `Herwig++` and `Lpair` generators do not include any corrections to account for proton absorptive effects.
For the double dissociative reactions, `Pythia v8.212` [178] is used with the NNPDF2.3QED [149] photon PDFs as a default option. Depending on the multiplicity of the dissociating system, the default `Pythia8` string or mini-string fragmentation model is used for proton dissociation. The absorptive effects in double-dissociative MC events are taken into account using the default multi-parton interactions model implemented in `Pythia8` machinery [212].
QCD multijet $b\bar{b}$ and $c\bar{c}$ production samples generated using `Pythia` v8.1 [178] (A14 tune [189]) are also considered, as well as the MC samples reproducing Υ (1S)$\rightarrow \mu^+\mu^-$, Υ (2S)$\rightarrow \mu^+\mu^-$ and Υ (3S)$\rightarrow \mu^+\mu^-$ events. The former samples have been filtered at generator level, requiring two muons with $p_T > 3.5$ GeV and $m_{\mu\mu} > 6.0$ GeV. The latter samples are considered just in the lowest invariant mass bin, between 7 and 9 GeV, and they are produced with `Pythia v8.1` [178] interfaced with `Photos++ v3.52` [213] and the A14 CTEQ6L1 tune [189]. Details about the baseline and all the background samples used in the analysis are presented in Table 6.2, where the cross-sections σ (before applying filters) and

numbers of fully simulated MC events, as well as the corresponding filter efficiencies are shown.
In this chapter, the following terminology is in use:

- the Z samples refer to the resonant Z MC samples;
- the DY samples refer to the low-mass Z/γ^* MC samples,

and both together are our MC signal samples is use.

6.2 Event and object reconstruction

This section describes the selection criteria for all objects used in this analysis, namely, charged-particle tracks, vertices and muons. The measurement is based on Analysis Release 2.4.43[3] where the official performance tools are provided. The objects used in the measurement are calibrated and the latest scale factors and corrections are applied to correct the MC selection efficiencies to the ones in data, following the recommendations of the combined performance groups[4].
Muons in ATLAS are reconstructed by combining measurements from the inner detector (ID) and muon spectrometer (MS). The muon momentum calibration is applied before any object selection.
Event level selection:

- Event is required to be in the list of good runs.
- Event passes the OR of the following four triggers[5]
  - `HLT_mu4_iloose_mu4_7invm9_noos`,
  - `HLT_mu4_iloose_mu4_11invm60_noos`,
  - `HLT_mu6_iloose_mu6_11invm24_noos`,
  - `HLT_mu6_iloose_mu6_24invm60_noos`.
- Event is not flagged as being incomplete, and pass LAr and Tile error requirements which provide protection against noise bursts and data corruption in the ECAL and HCAL respectively.
- Event contains at least two combined muons (see Sec. 3.3.1 for the definition of combined muons).

Combined Muon level selection:

- Muon must have $p_T^\mu > 4.5$ GeV,
- Muon must have $|\eta^\mu| \leq 2.4$,

[3] `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/AnalysisBaseReleaseNotes24`
[4] `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCPAnalysisGuidelinesMC15`
[5] A more detailed description of the analysis triggers will follow in Sec. 6.2.1.

- No ID track requirement on $|\sigma(d_0)|$ and on $|\Delta z_0 \sin\theta|$ with respect to the beam line for single muon (see Sec. 2.8 for the definition of these quantities),
- Muon passes the "*FixedCutTight*" isolation criteria (Track isolation requirement: `ptvarcone30`/p_T < 0.06; Calorimeter isolation requirement: `topoetcone20`/p_T < 0.06 - more information about these isolation working points can be found in Sec. 3.3.1),
- Muon passes the "*Medium*" quality selection determined by the `MuonSelectionTool`[6].

Di-muon pair selection:

- Considering only the muon pair with the highest total p_T,
- No opposite-sign (OS) charge requirement applied (in the following, all the presented distributions are made after the (OS - SS) subtraction),
- Di-muon invariant mass between $7 < m_{\mu\mu} < 9$ GeV or $12 < m_{\mu\mu} < 60$ GeV.

Moreover, an additional cut on the transverse momentum of the Z boson, $p_T^Z > 8$ GeV, is applied in the lowest invariant mass bin.
The (OS - SS) subtraction is performed to eliminate the contamination from $W + jets$ events, which is symmetric in final states with OS or SS muons, and part of the multijet background.
A cosmic muon fakes two opposite sign muons by travelling through the detector towards the beam axis and then away from the beam axis. This change of direction relative to the beam axis means the two fake reconstructed muons are reconstructed with opposite charge. Due to the fact the two reconstructed muons are in fact the same cosmic muon, the variable $\eta_\mu^1 + \eta_\mu^2$ describing the sum of the η of the two reconstructed muons will be very close to zero. So, the $|\eta_\mu^1 + \eta_\mu^2|$ has been studied (as in Ref. [143]) but it has been found that the cosmic rays contamination is completely negligible, given the considered low invariant mass range in this analysis.

6.2.1 Trigger

Several trigger chains are used in this analysis, dependent on the instantaneous luminosity and the invariant mass of the di-muon pair. All the triggers used in this analysis were active for instantaneous luminosities below $3 \cdot 10^{33}$ cm^{-2} s^{-1}. These triggers were active at the same time with fixed prescales (PS). The total integrated luminosity recorded by the `HLT_mu6_iloose_mu6_24invm60_noos` and `HLT_mu6_iloose_mu6_11invm24_noos` (`HLT_mu4_iloose_mu4_7invm9_noos` and `HLT_mu4_iloose_mu4_11invm60_noos`) is 1280.28 pb^{-1} (319.68 pb^{-1}). Table 6.3 shows a summary of all the triggers in use.
The characteristics of the above mentioned `2mu6` (`2mu4`) triggers are:

- two muons with p_T > 6 (4) GeV are required at the L1 level.

[6] `https://twiki.cern.ch/twiki/bin/view/Atlas/MuonSelectionTool`

Trigger chain name	Average prescale	Recorded Luminosity [pb^{-1}]
HLT_mu4_iloose_mu4_7invm9_noos	10.0	319.68
HLT_mu4_iloose_mu4_11invm60_noos	10.0	319.68
HLT_mu6_iloose_mu6_11invm24_noos	2.5	1280.28
HLT_mu6_iloose_mu6_24invm60_noos	2.5	1280.28

Table 6.3: Summary of the recorder luminosity by the triggers in use in the analysis. The second column in the table shows average between the part of data taking in which the triggers were active and that in which were off because the instantaneous luminosity were below $3 \cdot 10^{33}$ cm^{-2} s^{-1}

- at HLT, an cut on the invariant mass of the di-muon system is applied, defined by *AinvmB*, which should be read as follows: $A < m_{\mu\mu} < B$,
- at HLT, the compatibility of the two muons tracks with the same vertex is required,
- at HLT, at least one muon passes the *Loose* isolation criteria. The *Loose* identification criteria are designed to maximise the reconstruction efficiency while providing good-quality muon tracks (e.g. calorimeter, track and combined isolation of 99%) [7],
- at HLT, there is no opposite-charge requirement.

6.2.2 Corrections to the Monte Carlo modelling

Several corrections are applied to the MC simulation in order to improve its agreement with data. Reconstruction and isolation efficiencies of muons in MC are corrected using the `MuonEfficiencyCorrections` package with the latest recommendations provided by the MCP group.
Before the reconstruction efficiency can be calculated, a correction must first be applied to the MC simulation to take into account the mis-modelling of the resolution response and momentum scale of the muon reconstruction. Dedicated ATLAS studies have been performed using a "*tag and probe*" method with muons from Z or J/ψ decays in order to measure the difference in momentum scale and momentum resolution between MC and data [128]. The efficiency of the muon reconstruction algorithm is also documented in Ref. [128]. The results from this study are also provided in a software tool (`MuonEfficiencyCorrections`) within the analysis framework, which is used to retrieve a scale factor to correct for the measured discrepancy between data and MC. The tool returns a p_T, η and ϕ dependent scale factor (SF) for each reconstructed muon (calculated after the p_T smearing has been applied).
Trigger SF correcting for the mis-modelling of all the di-muon triggers in the MC are not applied using the `TrigMuonEfficiency` package, but they are extracted using dedicated studies performed by other members of the analysis team. The impact of these SF on the relevant kinematic distribution is addressed in Sec. 6.2.6.

[7] `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/IsolationSelectionTool`

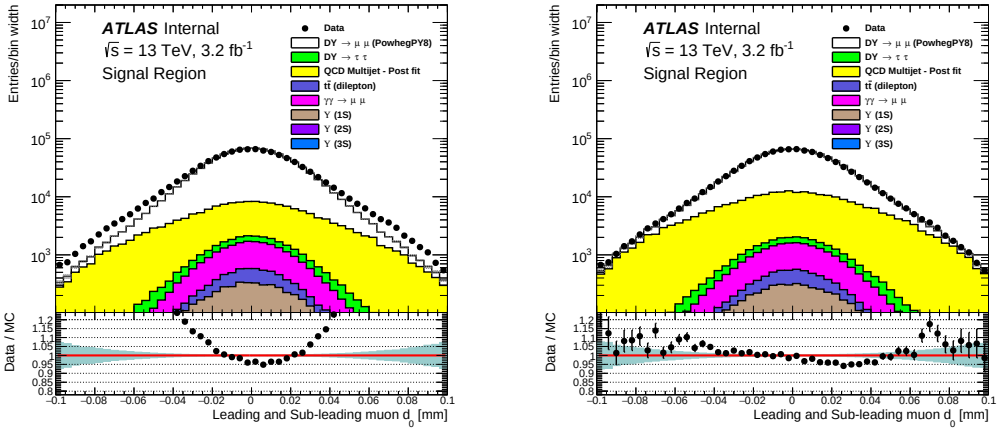


Fig. 6.4: Inclusive d_0 impact parameter distribution before (left) and after (right) the smearing correction applied. The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars).

The ID Tracking Performance group has provided smearing for the ID tracking parameters to account for differences in d_0 and z_0 resolutions between the data and MC. The `InDetTrackSmearingTool`[8] is used to apply this correction to the reconstructed charged-particle tracks. This is done by applying additional smearing for d_0 and z_0 impact parameters to recover shapes observed in data. Fig. 6.4 - 6.5 presents d_0 impact parameters and $\Delta z_0(\mu_1, \mu_2)$ comparison between data and MC before and after applying this supplementary smearing correction.

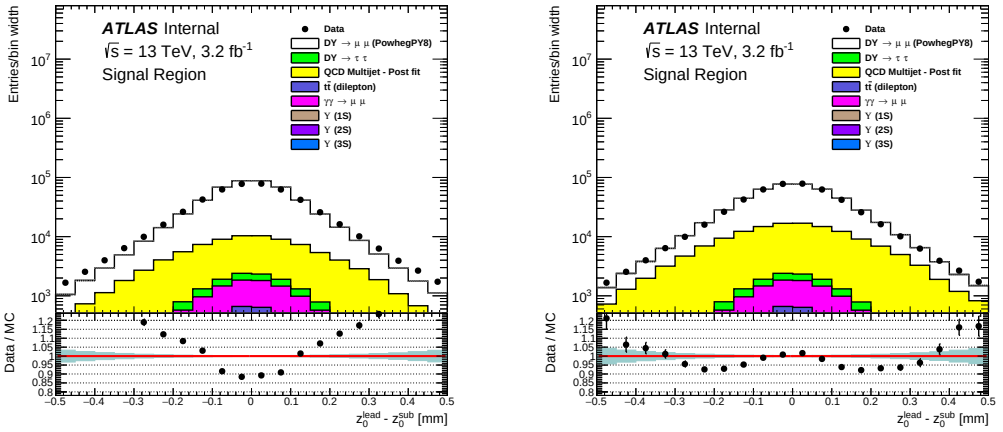


Fig. 6.5: Inclusive Δz_0 impact parameter distribution before (left) and after (right) the smearing correction applied. The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars).

The d_0 distribution shows a reasonable agreement between data and MC predictions, while there are two symmetric dips between 0.2 and 0.3 mm and -0.2 and -0.3 mm in the $\Delta z_0(\mu_1, \mu_2)$ distribution. The method to retrieve the data-driven

[8] `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/InDetTrackSmearingTool`

normalisation of the LO Pythia $b\bar{b}/c\bar{c}$ samples will be described Sec. 6.3.7 and the factor shown in Tab. 6.5 have been applied to produce these plots.

6.2.3 Reconstructed vertices and pile-up corrections

Because of the high LHC peak luminosity and small bunch separation, the number of proton-proton (pp) interactions occurring in the same bunch crossing is large. This is referred to as event *pile-up* and requires the use of dedicated algorithms and corrections to mitigate its impact on the reconstruction of leptons. The variations in pile-up conditions in 2015 are included in the simulation of MC events in order to match the different running conditions and the pile-up distribution observed in data. Residual differences in the pile-up between data and MC simulation have been corrected by re-weighting the MC events to reproduce the average number of interactions per bunch-crossing, $< \mu >$, observed in data. This reweighting of the MC samples is performed using the official PileupReweightingTool [214]. Additionally, the $< \mu >$ value is scaled by 1/1.09 (with respect to data), based on studies of the data/MC agreement for the number of reconstructed vertices vs. $< \mu >$ (as described here[9]). The distribution of the number of reconstructed vertices in data and MC, before and after pile-up reweighting is applied are shown in Fig. 6.6. It can be seen that the sum of the re-weighted MC samples (solid red line) almost accurately describes the data distribution.

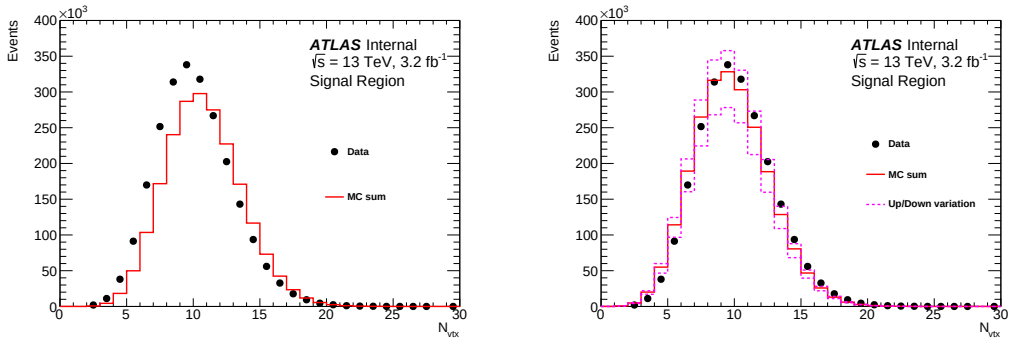


Fig. 6.6: Control distributions of the average number of reconstructed vertices before (left) and after (right) application of the PileupReweightingTool. All MC expectations are scaled to the integrated luminosity of the data using predicted MC cross sections of each sample. The magenta band in the right plot denotes the systematic variation of 1/1.09 scaling factor (varied between 1/1.0 and 1/1.18).

6.2.4 Z boson transverse momentum modelling

The discrepancy in the Z_{p_T} spectrum, due to the well-known mismodelling of this variable in the Powheg+Pythia simulation (see Ref. [215]), has been inspected. To improve the description of the Z p_T distribution, re-weighting is performed using the same data and MC samples as for the comparisons between

[9] https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/ExtendedPileupReweighting

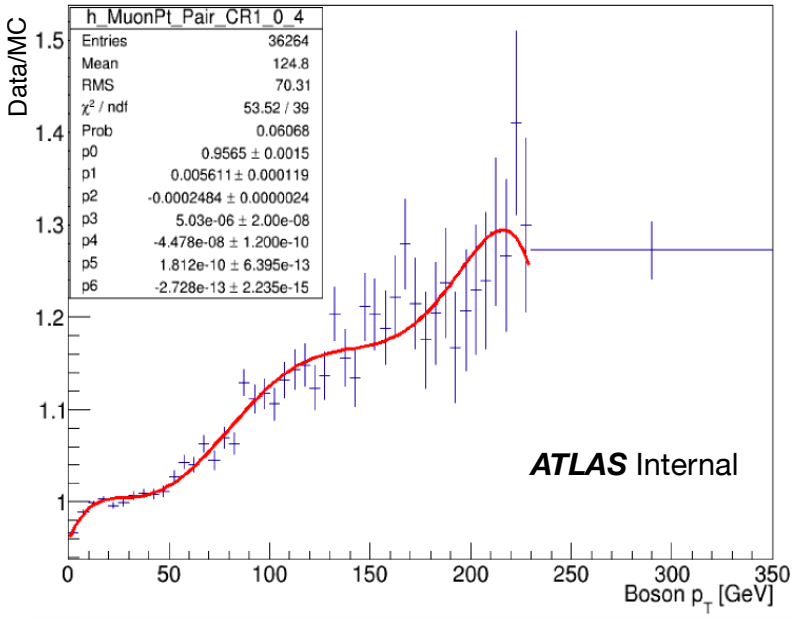


Fig. 6.7: Ratio between data subtracted by all the background MC samples and signal MC of the Z p_T distribution for the muon channel in the Z CR. The red curve corresponds to the polynomial fit performed in the $[0-230]$ GeV range.

data and MC in the control plots. To take into account the discrepancies observed at high p_T in the Z p_T spectrum, the following procedure is used. After having applied all the selection cuts at reco-level (defined as the information after the reconstruction algorithms have been run over the data from the simulated detector), the ratio of the data Z p_T distribution, after all background MC samples have been subtracted, and the signal MC was fitted was fitted with a polynomial function of sixth degree, namely: $f(x) = p0 + p_1 x + p_2 x^2 + p_3 x^3 + p_4 x^4 + p_5 x^5 + p_6 x^6$. The fit has been performed in the p_T range [0 - 250] GeV. Then, an event-by-event weight has been applied using the muon p_T at Born-level. These weights have been applied in such a way that the Z p_T reweighting only shuffles events between bins, but does not overall change the truth level cross section.

This method is first performed in the Z peak control region (CR), defined by applying the same kinematic cuts applied in the ATLAS measurement of $t\bar{t}$ and Z boson cross-section ratio performed at $\sqrt{s}$= 7, 8 and 13 TeV[10], namely:

- Trigger: `HLT_mu20_iloose_L1MU15` or `HLT_mu50`
- Muon p_T > 25 GeV
- Muon $|\eta|$ < 2.4
- ID: Medium muon
- Isolation: `LooseTrackOnly`
- $|d_0$ significance$|$ < 3 and $\Delta z_0 \sin\theta$ < 0.5 mm

[10] `https://cds.cern.ch/record/2117165/files/SupportNoteZ.pdf?version=4`

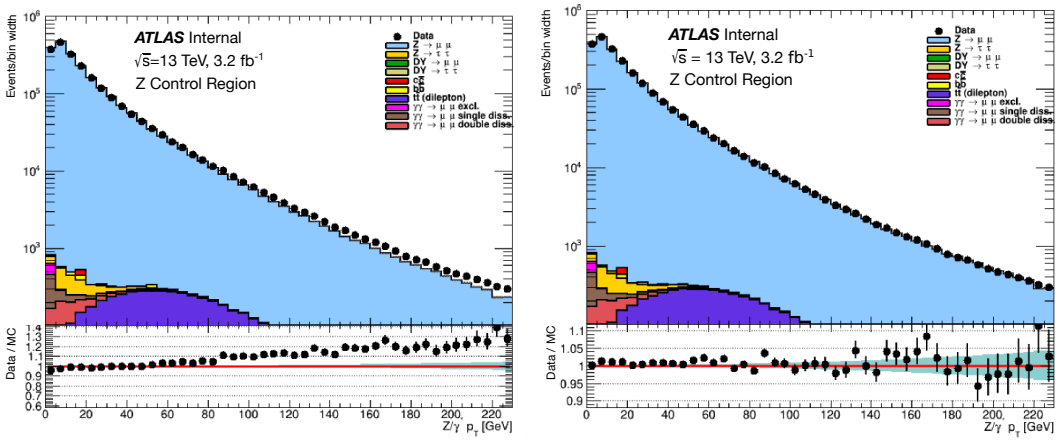


Fig. 6.8: Comparison between data subtracted by all the background MC samples and signal MC of the Z p_T before (left) and after (right) the reweighting procedure in the Z CR. The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars).

- Exactly two muons, same flavour and oppositely charged
- Mass Window: 66 < $m_{\mu\mu}$ < 116 GeV

The result of this validation study is shown in Fig. 6.7, where the the p_T range $[0 - 230]$ GeV has been fitted. A $\chi^2/\text{ndf} = 53.52/39$ is obtained with the following values for fit parameters:

$$
\begin{aligned}
p_0 &= 9.565 \cdot 10^{-1} \pm 0.15 \cdot 10^{-2} \\
p_1 &= 5.611 \cdot 10^{-3} \pm 0.12 \cdot 10^{-3} \\
p_2 &= -2.584 \cdot 10^{-4} \pm 0.24 \cdot 10^{-5} \\
p_3 &= 5.030 \cdot 10^{-6} \pm 0.20 \cdot 10^{-7} \\
p_4 &= -4.478 \cdot 10^{-8} \pm 0.12 \cdot 10^{-9} \\
p_5 &= 1.812 \cdot 10^{-10} \pm 0.06 \cdot 10^{-11} \\
p_6 &= -2.728 \cdot 10^{-13} \pm 0.22 \cdot 10^{-14}
\end{aligned}
$$

The comparison between data and MC before and after the reweighting procedure is shown in Fig. 6.8. The agreement between data and MCs after the Z p_T reweighting exercise is expected by construction.
Having thus checked the methodology, the same study has been carried on in the signal region (SR) using the same polynomial function, and the result of the fit is shown in Fig. 6.9 while the comparison between data and MC before and after the reweighting procedure is shown in Fig. 6.10 for the Z p_T distribution and in Fig. 6.11 for the inclusive lepton p_T distribution.
The following output parameters from the fit have been obtained:

$$
\begin{aligned}
p_0 &= 4.418 \cdot 10^{-1} \pm 1.01 \cdot 10^{-2} \\
p_1 &= 5.795 \cdot 10^{-2} \pm 2.29 \cdot 10^{-3}
\end{aligned}
$$

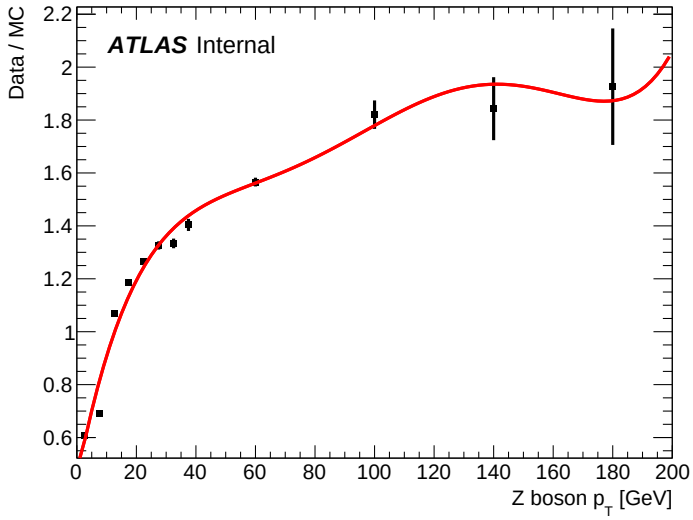


Fig. 6.9: Ratio between data subtracted by all the background MC samples and signal MC of the Z p_T distribution for the muon channel in the SR (defined in Sec. 6.2). The red curve corresponds to the polynomial fit performed in the $[0-200]$ GeV range. The total χ^2/n.d.f. is 178.1/7.

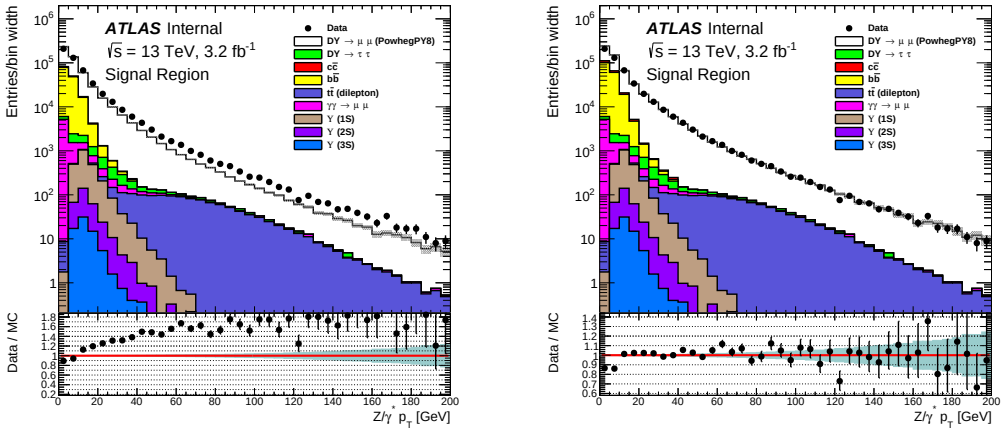


Fig. 6.10: Z p_T before (left) and after (right) the reweighting procedure in the analysis SR (defined in Sec. 6.2). The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars).

$$p_2 = -1.273 \cdot 10^{-3} \pm 1.31 \cdot 10^{-4}$$
$$p_3 = 1.418 \cdot 10^{-5} \pm 2.68 \cdot 10^{-6}$$
$$p_4 = -7.285 \cdot 10^{-8} \pm 2.13 \cdot 10^{-9}$$
$$p_5 = -0.138 \cdot 10^{-9} \pm 0.57 \cdot 10^{-10}$$

It can be seen from Fig. 6.10 and Fig. 6.11 that the data/MC agreement improves up to 100 GeV, where the data/MC ratio is close to unity. Then, for invariant masses greater than 100 GeV, the statistical uncertainties are about 20% or even bigger. For low di-muon masses (e.g. $m_{\mu\mu} < 10$ GeV), where the multijet back-

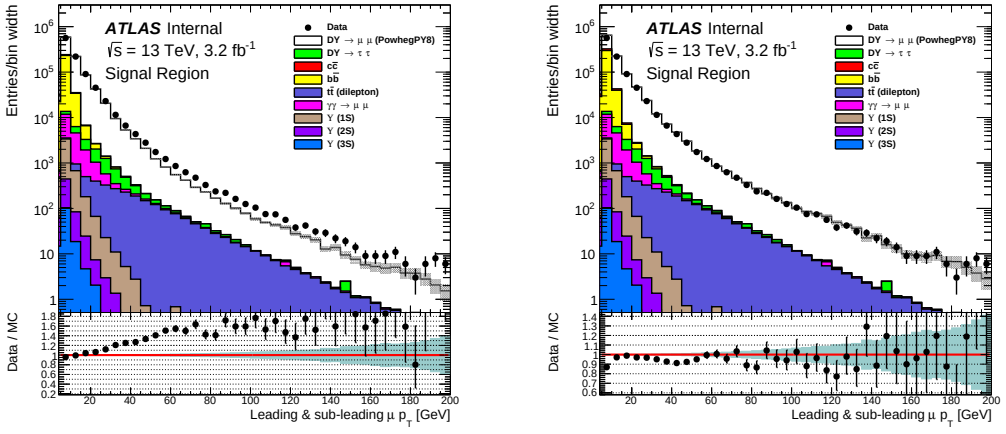


Fig. 6.11: Inclusive muon p_T before (left) and after (right) the reweighting procedure in the analysis SR (defined in Sec. 6.2). The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars).

ground is one of the dominant contributions, the data/MC ratio is not fitted well. It is clear that there was an improvement in the description of the data provided by the the reweighted MC samples but it is also visible that the achieved gain is not sufficient to provide a good modelling of the Z p_T across the whole transverse momentum range, especially in the low-p_T region. For this preliminary study, the additional requirement on $p_T^Z > 8$ GeV in the 7-9 GeV mass bin (as described in Sec. 6.2) has not been applied, because it has been decided to apply this specific cut at a later stage of the analysis.
A further improvement can be obtained by considering a mass-dependent reweighting. The above-described procedure is repeated in each invariant mass bins, namely: 7 to 9 GeV, 12 to 14 GeV, 14 to 17 GeV, 17 to 22 GeV, 22 to 28 GeV, 28 to 36 GeV, 36 to 46 GeV and 46 to 60 GeV. The various fitted functions can be found in Fig. 6.12. The very highest p_T bins are not included in the fit and they are treated as *overflow* bins, so a fixed weight equal to the value of the data/MC ratio in the last bin is applied to every event in that bin. As described in Sec. 6.2, an additional cut requiring $p_T^Z > 8$ GeV is applied for $7 < m_{\mu\mu} < 9$ GeV. Moreover, in this particular given mass bin, also the lowest p_T bin is not included in the fit but it is treated as *underflow* bin (same definition as for the *overflow* bins just above).
The Z p_T distributions after the above-mentioned reweighting are shown in Fig. 6.13, where now it is clearly visible a good description of the data within uncertainties. The method to retrieve the data-driven normalisation of the LO `Pythia` $b\bar{b}/c\bar{c}$ samples will be described Sec. 6.3.7 and the factor shown in Tab. 6.5 have been applied to produce these plots.
In the end, a poor agreement between data and MC can be observed in the 7 to 9 GeV bin, and it is probably due to the wrong normalisation of the MC DY sample (please refer to Sec. 6.3.1 for further discussion). A more detailed analysis of this particular bin will follow in Sec. 6.2.8.

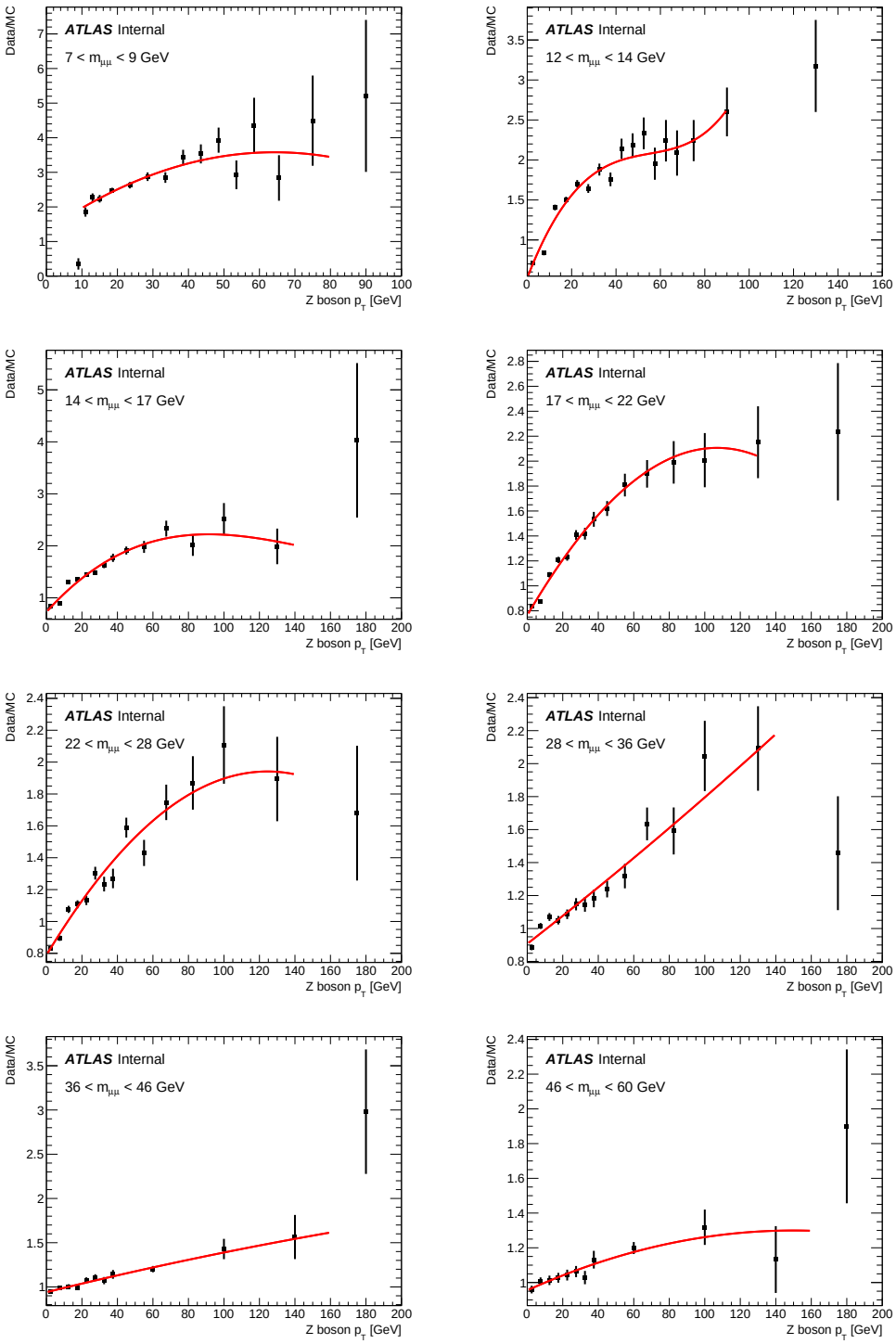


Fig. 6.12: Ratio between data and MC of the Z p_T distribution for the muon channel in the SR (defined in Sec. 6.2). The red curve corresponds to the polynomial fit performed in the different invariant mass bins.

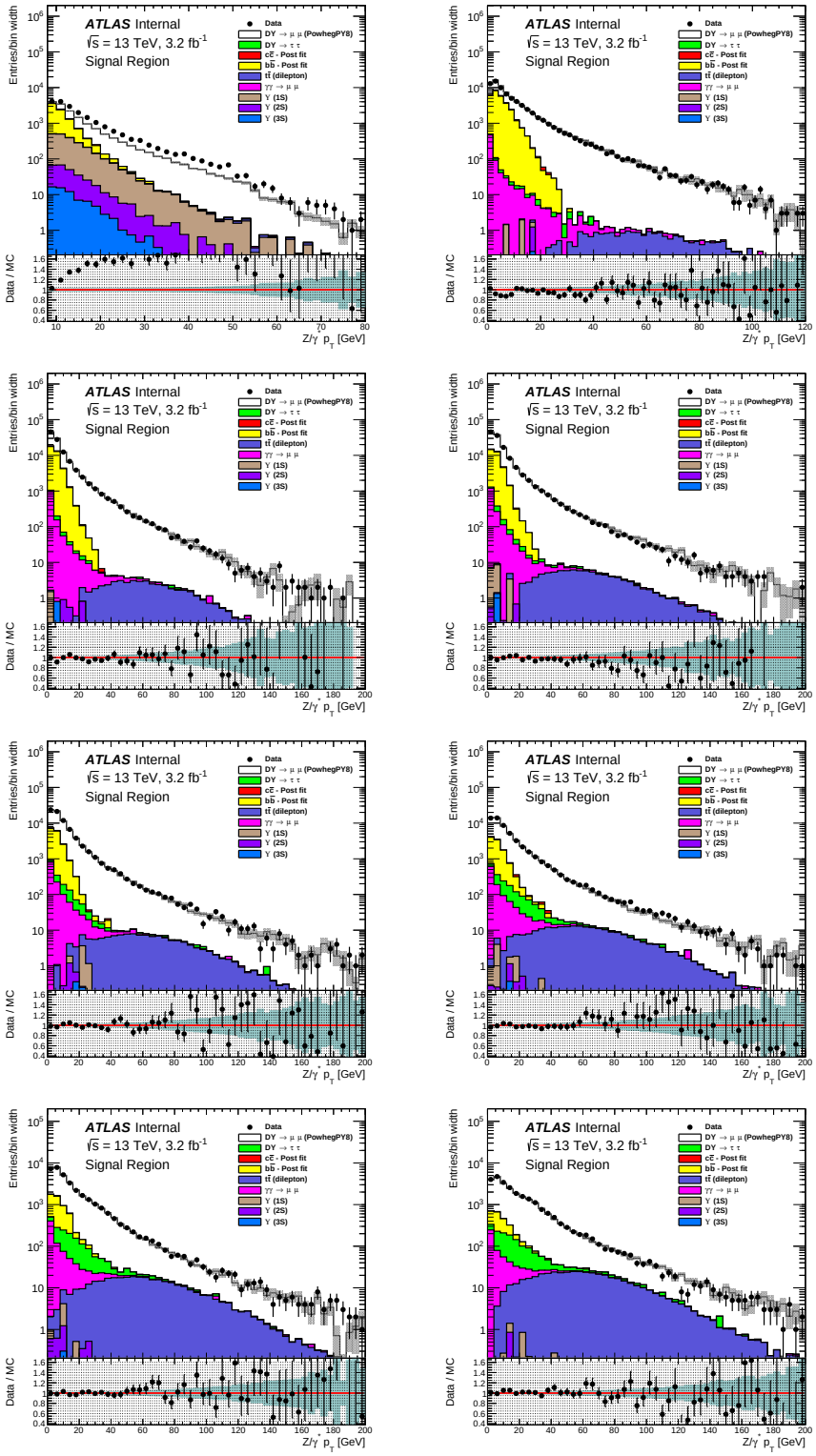


Fig. 6.13: Comparison between data and MC of the Z p_T after the above-described reweighting procedure in the various invariant mass bins of our analysis SR (defined in Sec. 6.2). The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars).

6.2.5 Υ Control Region

In the nominal selection of the analysis, the invariant mass region between 9 and 12 GeV is excluded because there are many resonances decaying to a muon pair; in particular, the dominant contribution to the di-muon final state comes from the following three resonances, namely Υ_{1S}, Υ_{2S} and Υ_{3S}, decaying to $\mu^+\mu^-$. A check if the normalisation of these LO `Pythia` Υ MC samples is required. The di-muon invariant mass spectrum between 7.5 and 12 GeV is shown in Fig. 6.14 (left). In this invariant mass region, all the selection criteria are applied, apart from the trigger selection requirement. The trigger `HLT_2mu4_bUpsimumu` is used as the signal region triggers are disabled in the Υ mass range. The luminosity recorded by this trigger (2.6 fb^{-1}) is higher than the one recorded by the triggers used in the nominal selection, so both data and MC predictions are scaled to the `HLT_mu6_iloose_mu6_11invm24_noos` luminosity, namely 1.3 fb^{-1}. Furthermore, a cut on the invariant mass of the di-muon system is applied, selecting muon pairs having $9 < m_{\mu\mu} < 12$ GeV.
In order to find the correct normalisation of the Υ samples, a `TFractionFitter` method [216] is used, where the data are fitted with the Υ_{1S}, Υ_{2S} and Υ_{3S} MC samples and the sum of all the other MC samples. This method essentially builds a template from an arbitrary number of MC distribution shapes, and then determines the fraction of the data contained in each MC category. The only parameters which can be measured are the fraction for each component, but this method supports a large number of parameters, and requires only a single MC template for each component. It uses Poisson statistics, so it works even with low-statistics or empty bins, in both the data and template. Support for Poisson statistics requires introducing an additional parameter for each bin of each template histogram, as well as the "physics" fraction parameters, but the minimisation with respect to these additional parameters is done analytically rather than introducing them as formal fit parameters. However, the `TFractionFitter` method is limited to linear combinations of fixed component shapes, and cannot handle more general fitting problems, in which some of the fitted parameters determine the shape of the distribution, rather than just the normalisation.
The normalisations found are listed below:

$$\Upsilon_{1S} = 0.2650 \pm 0007$$
$$\Upsilon_{2S} = 0.5408 \pm 0006$$
$$\Upsilon_{3S} = 0.1610 \pm 0005$$

The di-muon invariant mass with the normalisations found in the previous study applied is shown in Fig. 6.14 (right plot). It is evident that the data/MC agreement for invariant masses greater than 8.5 GeV gets better and closer to the unity, while it is still off in the low invariant mass region. This is probably due to the fact that the normalisation of MC predictions for QCD multijet background ($b\bar{b}$, $c\bar{c}$) is not reliable (please refer to Sec. 6.3, where the data-driven normalisation of the multijet background is found). Furthermore, it was studied if the normalisation factors found in the previous study vary with the rapidity. For this purpose, four *pseudo-rapidity* regions are defined:

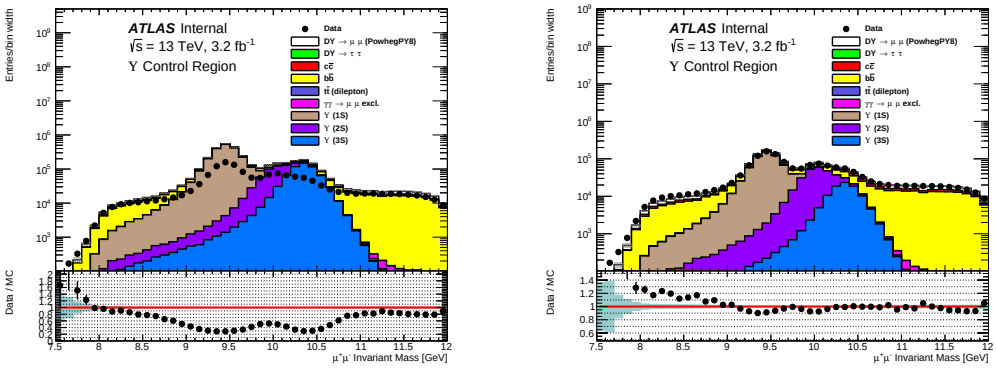


Fig. 6.14: Di-muon invariant mass spectrum from 7.5 GeV and 12 GeV before (left) and after (right) the normalisations found in the previous study have been applied. The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars).

- Υ region A: both the leading and the sub-leading muon in the barrel ($|\eta^{lead}| \leq 1.05$ and $|\eta^{sub}| \leq 1.05$)
- Υ region B: both the leading and the sub-leading muon in the end-cap ($|\eta^{lead}| > 1.05$ and $|\eta^{sub}| > 1.05$)
- Υ region C: leading muon in the barrel and the sub-leading muon in the end-cap ($|\eta^{lead}| \leq 1.05$ and $|\eta^{sub}| > 1.05$)
- Υ region D: leading muon in the end-cap and the sub-leading muon in the barrel ($|\eta^{lead}| > 1.05$ and $|\eta^{sub}| \leq 1.05$)

As before, a `TFractionFitter` method is used to find the correct normalisations and the results are shown in Fig. 6.15. The normalisations in the four pseudorapidity regions have been found to be similar among them; to be more precise, they differ from each other by 2-3%. Moreover, they are also consistent with the normalisation factors found in the study integrated in pseudorapidity, where the difference is ∼2-3% as well.

6.2.6 Trigger SF validation

In this section, the validation of the trigger SFs computed for the analysis triggers will be shown. These SFs have been calculated by other members of the analysis team, so only the studies where I got involved will be reported in the following.
The trigger SFs validation was performed in two separate steps. First of all, the trigger SFs were tested as a function of the transverse momentum of the leading muon of the selected pair. More precisely, we assessed their impact in four different p_T slices, namely:

- $p_T^{\text{lead}} < 6$ GeV

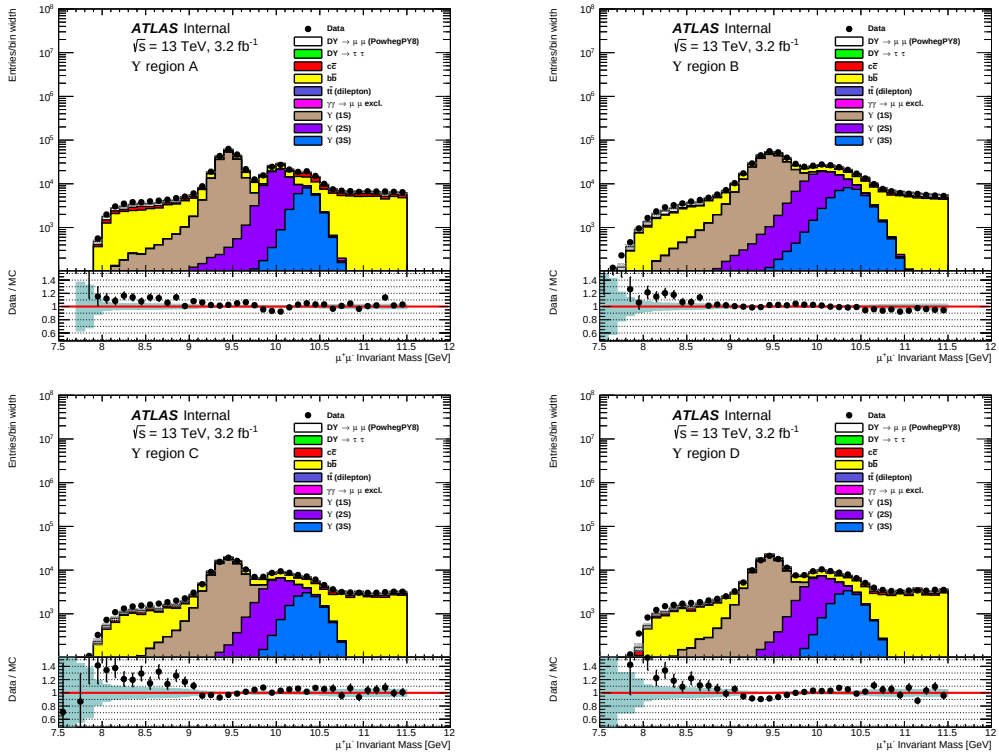


Fig. 6.15: Di-muon invariant mass spectrum from 7.5 GeV and 12 GeV. with the normalisations found in the previous study applied. The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars). Top left: region A; Top right: region B; Bottom left: region C; Bottom right: region D.

- $6 < p_T^{\text{lead}} < 10$ GeV
- $10 < p_T^{\text{lead}} < 15$ GeV
- $p_T^{\text{lead}} > 15$ GeV

The results of this study are shown in Fig. 6.16 for the `OR` between the `HLT_mu6_iloose_mu6_11invm24_noos` and `HLT_mu6_iloose_mu6_24invm60_noos` triggers, while in Fig. 6.17 for the `HLT_mu4_iloose_mu4_7invm9_noos` and `HLT_mu4_iloose_mu4_11invm60_noos` ones. An improvement in the data/MC agreement between 5% - 10% is visible and the $|y|$ dependence is now flat.

Then, the impact of the trigger SFs have been tested as a function of pseudo-rapidity, in four different regions in $|\eta|$ defined as follows:

- $|\eta^{\mu_1}| \leq 1.05$ and $|\eta^{\mu_2}| \leq 1.05$ (both muons in the barrel - region A)
- $|\eta^{\mu_1}| > 1.05$ and $|\eta^{\mu_2}| > 1.05$ (both muons in the end-cap - region B)
- $|\eta^{\mu_1}| \leq 1.05$ and $|\eta^{\mu_2}| > 1.05$ (leading muon in the barrel, sub-leading muon in the end-cap - region C)
- $|\eta^{\mu_1}| > 1.05$ and $|\eta^{\mu_2}| \leq 1.05$ (leading muon in the end-cap, sub-leading muon in the barrel - region D)

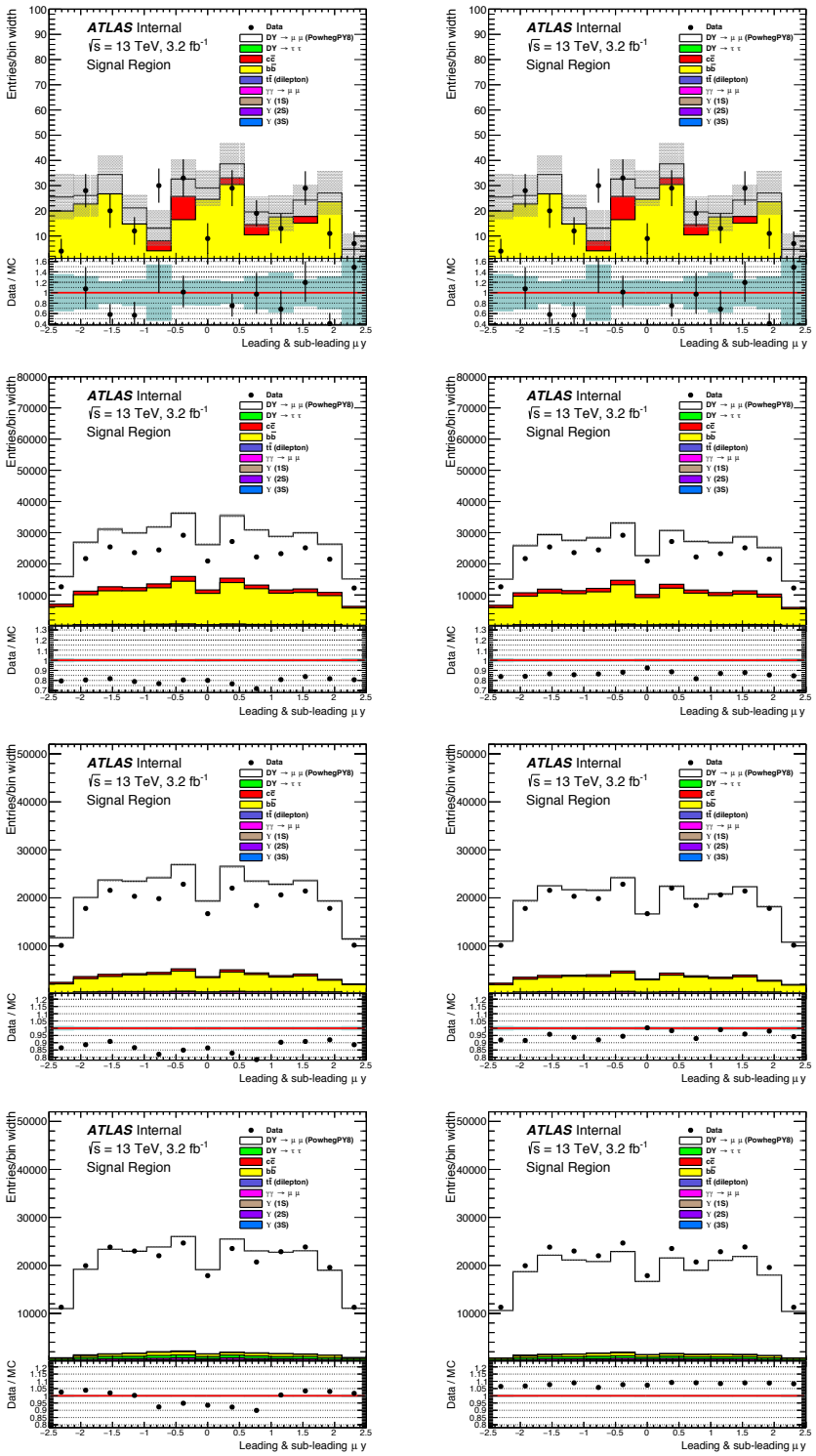


Fig. 6.16: Inclusive rapidity distribution (left) before and (right) after applying the trigger SF for the `2mu6` triggers. First row: $p_T^{\text{lead}} < 6$ GeV; second row: $6 < p_T^{\text{lead}} < 10$ GeV; third row: $10 < p_T^{\text{lead}} < 15$ GeV; fourth row: $p_T^{\text{lead}} > 15$ GeV. The statistical uncertainties for the signal and background distributions are shown in the shaded band in the ratio panel (lower panel), and the statistical uncertainties on the data points are represented by the vertical bars.

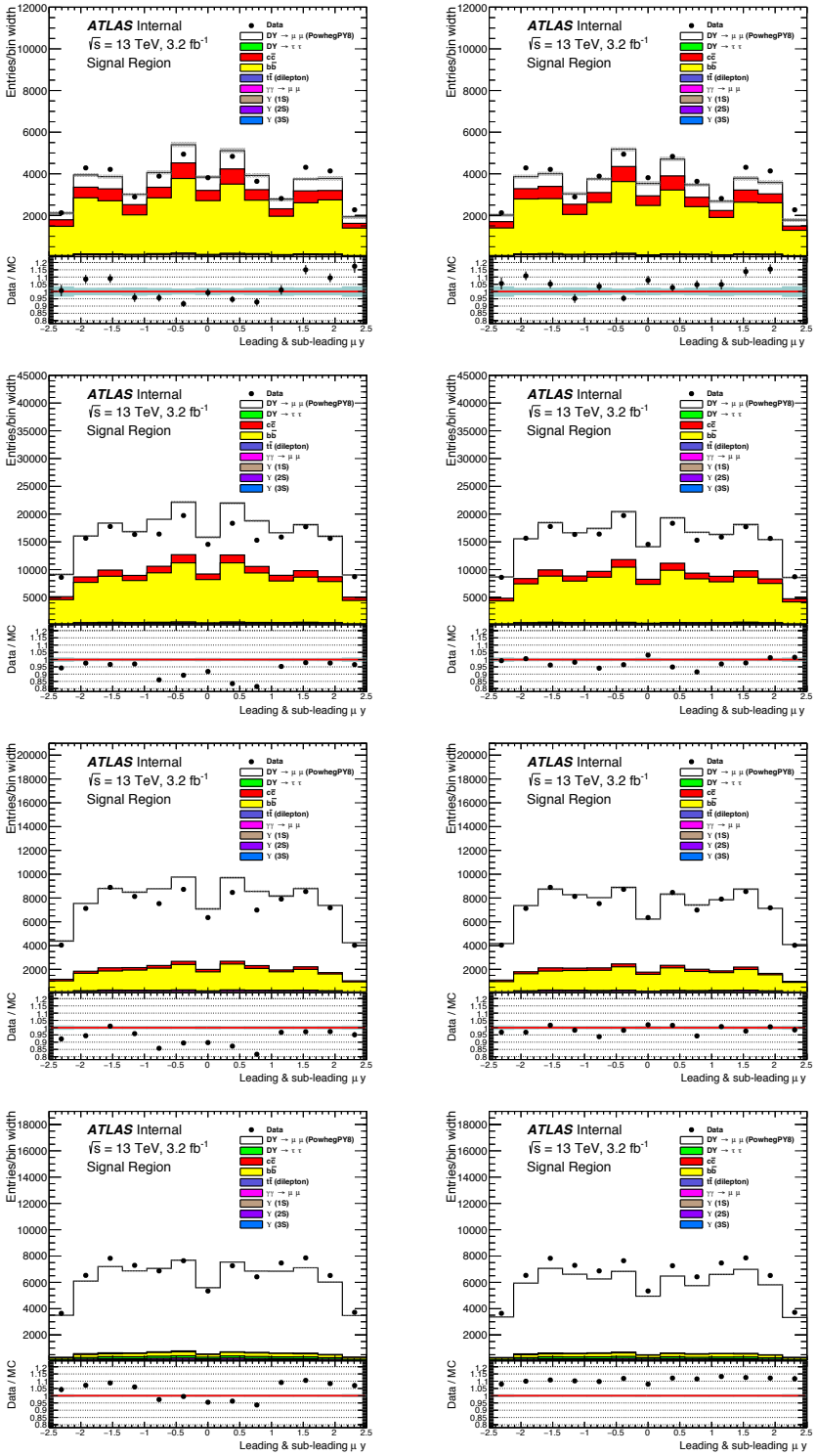


Fig. 6.17: Inclusive rapidity distribution (left) before and (right) after applying the trigger SF for the `2mu4` triggers. First row: $p_T^{\text{lead}} < 6$ GeV; second row: $6 < p_T^{\text{lead}} < 10$ GeV; third row: $10 < p_T^{\text{lead}} < 15$ GeV; fourth row: $p_T^{\text{lead}} > 15$ GeV. The statistical uncertainties for the signal and background distributions are shown in the shaded band in the ratio panel (lower panel), and the statistical uncertainties on the data points are represented by the vertical bars.

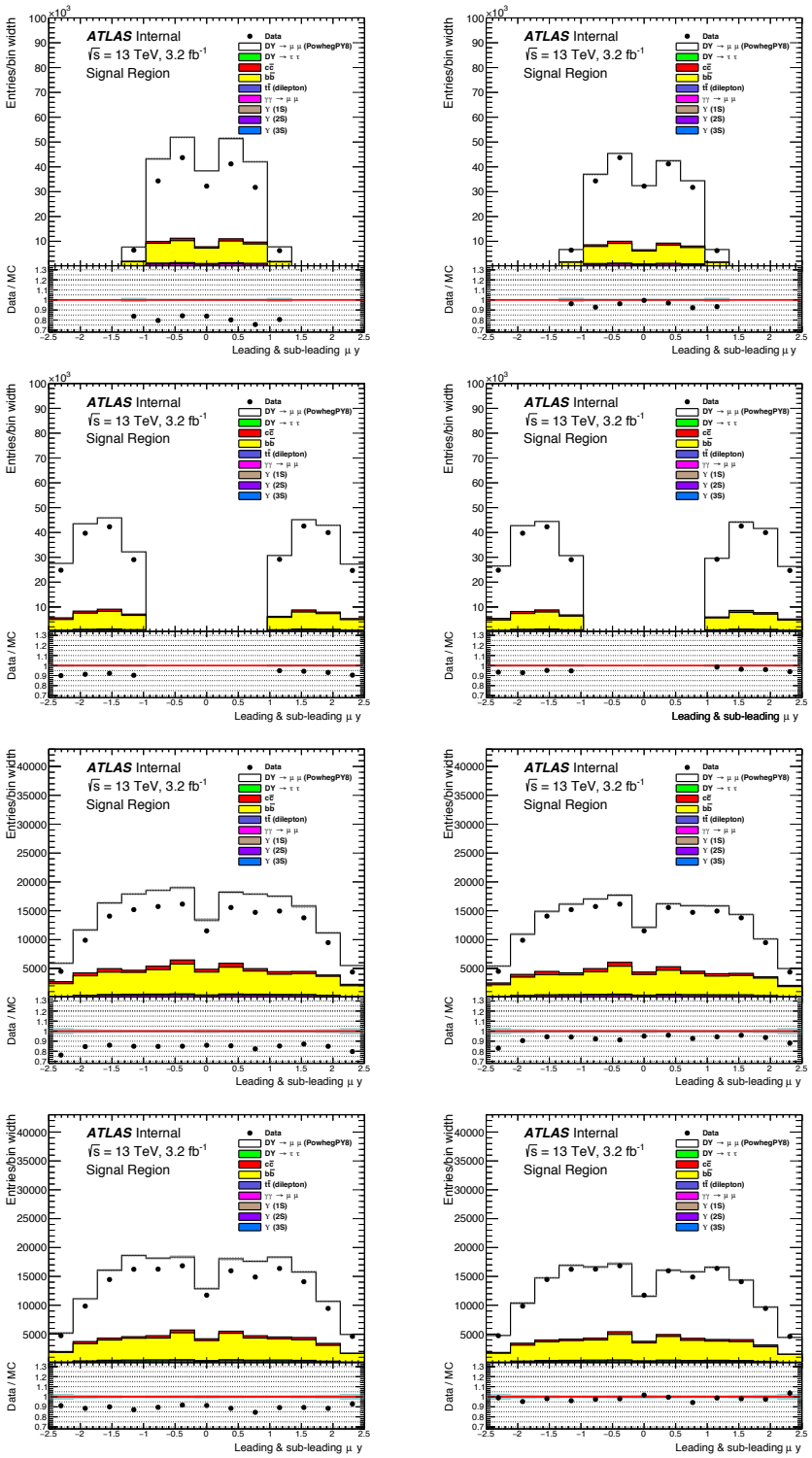


Fig. 6.18: Inclusive rapidity distribution (left) before and (right) after applying the trigger SF for the `2mu6` triggers. First row: region A; second row: region B; third row: region C; fourth row: region D. The statistical uncertainties for the signal and background distributions are shown in the shaded band in the ratio panel (lower panel), and the statistical uncertainties on the data points are represented by the vertical bars.

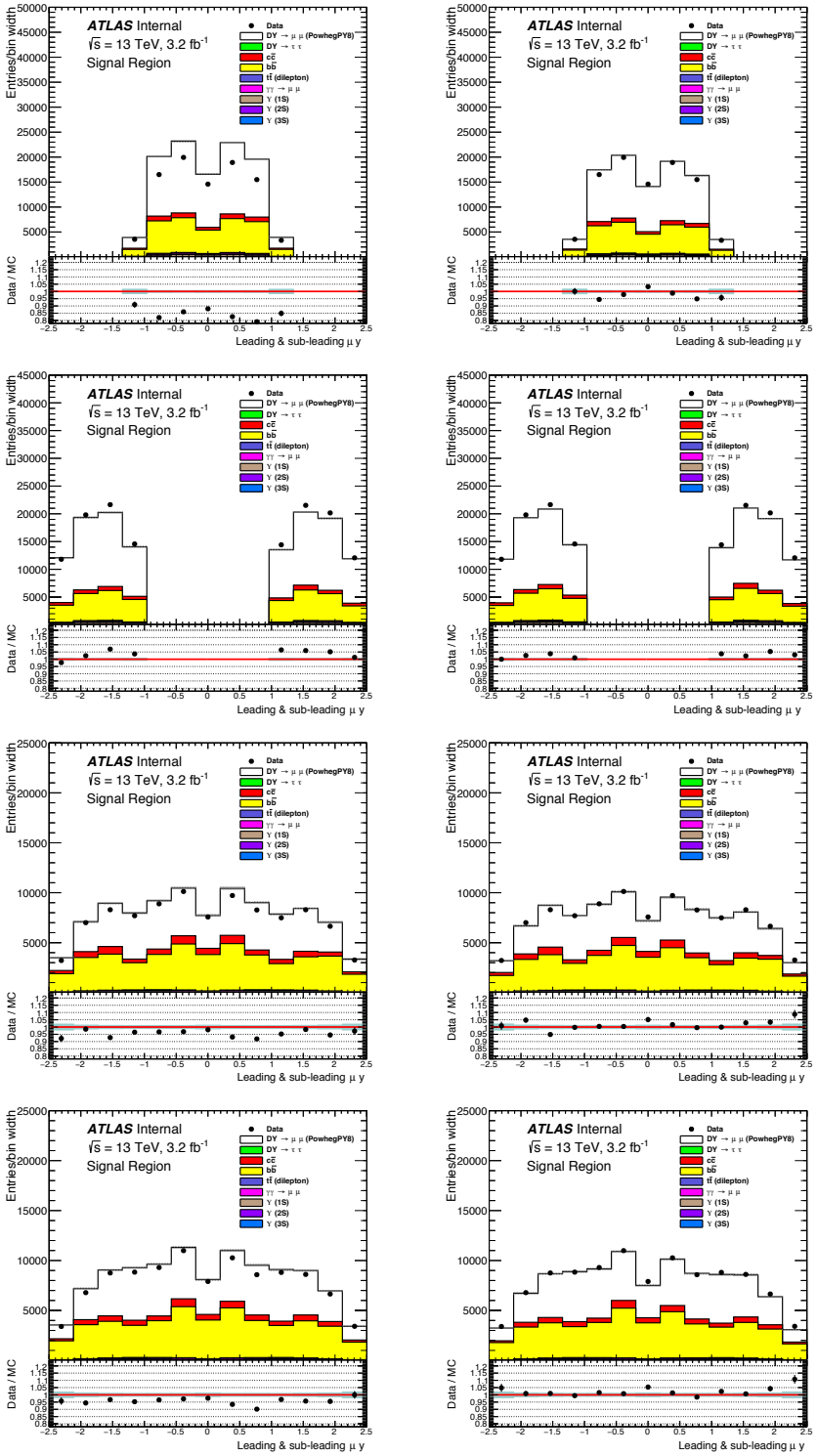


Fig. 6.19: Inclusive rapidity distribution (left) before and (right) after applying the trigger SF for the `2mu4` triggers. First row: region A; second row: region B; third row: region C; fourth row: region D. The statistical uncertainties for the signal and background distributions are shown in the shaded band in the ratio panel (lower panel), and the statistical uncertainties on the data points are represented by the vertical bars.

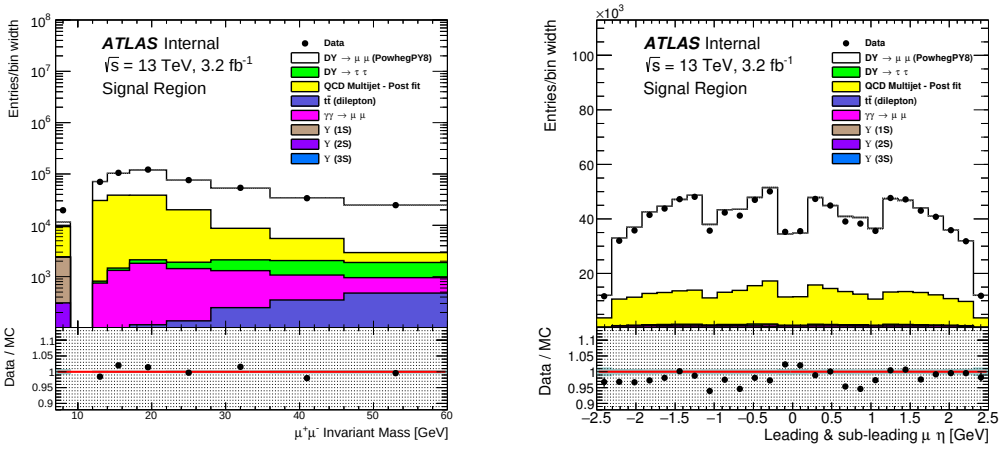


Fig. 6.20: Invariant mass of the di-muon system (left) and inclusive lepton η (right) distributions. All selection criteria are applied. The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars).

The outcome of this validation can be found in Fig. 6.18 for the `2mu6` triggers and in Fig. 6.19 for the `2mu4` ones. A similar improvement in the data description as compared to the one observed in the previous case can be seen. The disagreement between data and MC shown in Fig. 6.16 - 6.19 can be explained by the wrong normalisation of the LO `Pythia` $b\bar{b}/c\bar{c}$ samples. The method to retrieve the data-driven normalisation of the LO `Pythia` $b\bar{b}/c\bar{c}$ samples will be described Sec. 6.3.7.

6.2.7 Data to MC comparisons

All the relevant kinematic distributions for data and Monte Carlo (MC) simulation after all selection criteria are applied are shown in Fig. 6.20 - Fig. 6.22. The invariant mass distribution of the di-muon pair shows a good agreement in the data/MC ratio across the whole spectrum, apart from in the lowest invariant mass bin, where we expect to not have MC providing a good description of the data, given the wrong normalisation of the DY signal MC sample (see Sec. 6.3.1 for further discussion). The inclusive η, ϕ and the Z boson y distributions are in a good agreement between data and MC. This agreement is also visible in the Z p_T and inclusive lepton p_T distribution shown in Fig. 6.22, where the MC predictions agrees with data within uncertainties. The method to retrieve the data-driven normalisation of the LO `Pythia` $b\bar{b}/c\bar{c}$ samples will be described Sec. 6.3.7 and the factor shown in Tab. 6.5 have been applied to produce these plots.

6.2.8 The 7 to 9 GeV bin

The aim of this section is to show the control kinematic distributions in the invariant mass bin between 7 and 9 GeV. In this specific bin, all the selection criteria are applied, apart from the trigger selection requirement. Indeed, the di-muon pairs selected in this bin need to pass and to be matched to the `HLT_mu4_iloose_mu4_7invm9_noos` trigger. In addition to the nominal se-

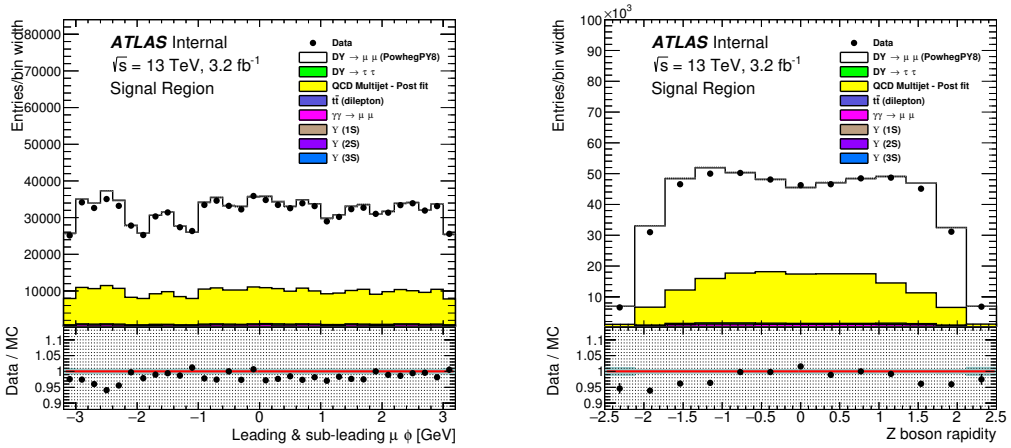


Fig. 6.21: Inclusive lepton ϕ (left) and Z boson rapidity (right) distributions. All selection criteria are applied. The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars).

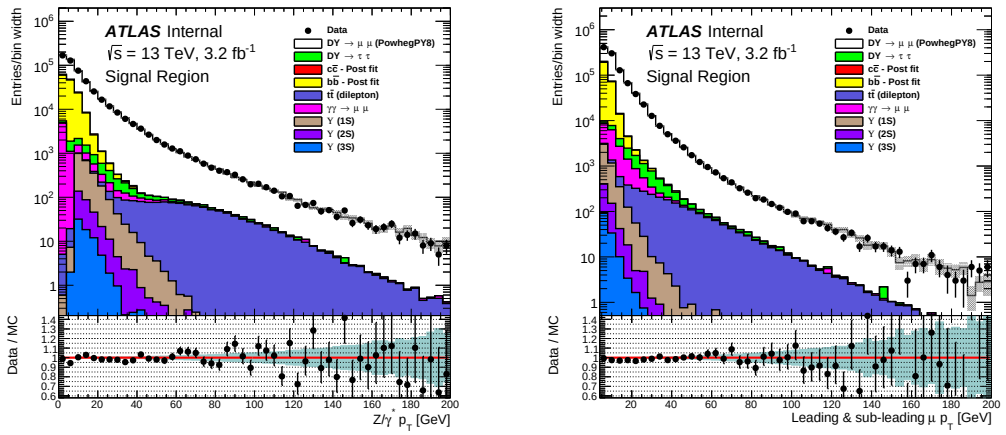


Fig. 6.22: Z boson p_T (left) and inclusive lepton p_T (right) distributions. All selection criteria are applied. The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars).

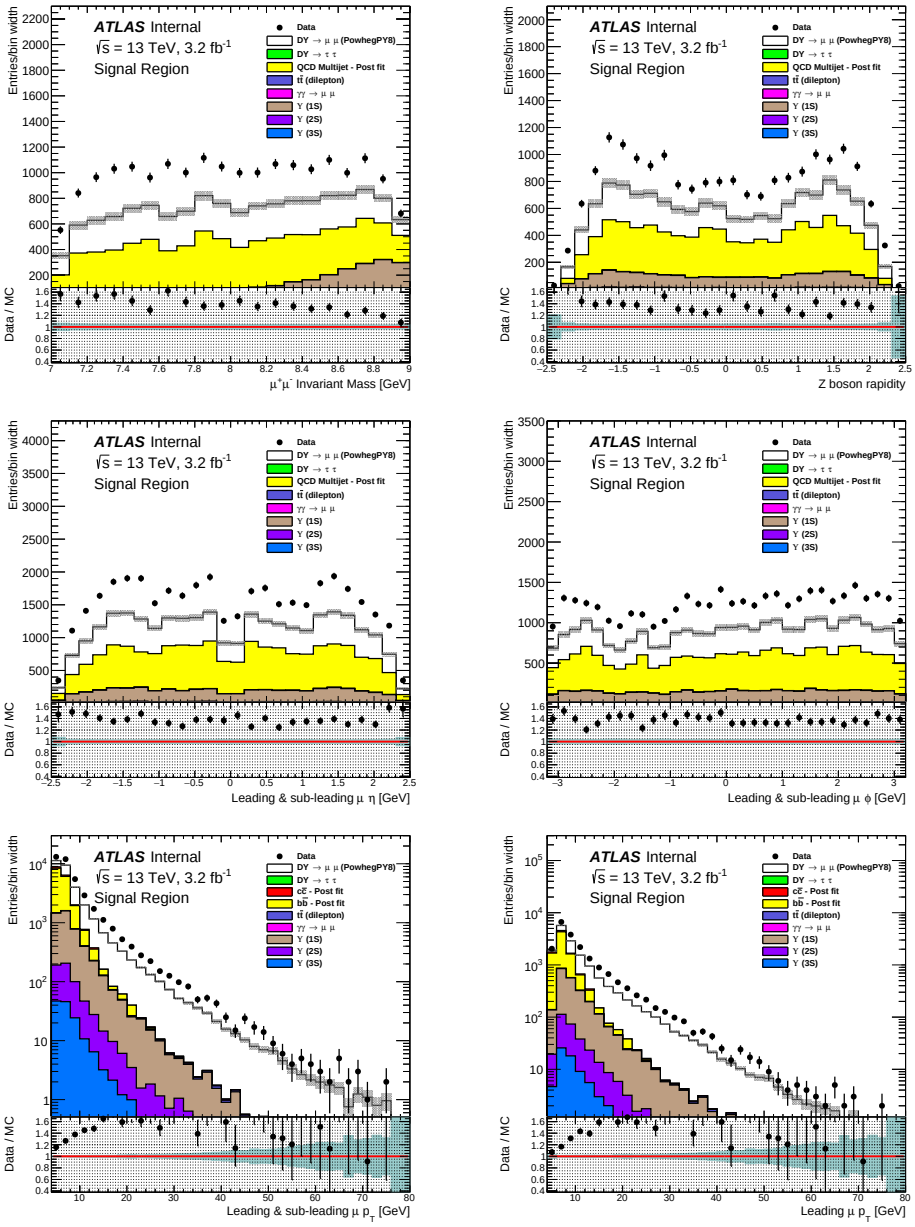


Fig. 6.23: Control kinematic distributions in the invariant mass bin from 7 to 9 GeV. all the selection criteria are applied, apart from the trigger selection requirement. Indeed, the di-muon pairs selected in this bin need to pass and to be matched to the `HLT_mu4_iloose_mu4_7invm9_noos` trigger.The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars). Top left: invariant mass distribution; Top right: Z boson rapidity; Middle left: Inclusive $|\eta|$ distribution; Middle right: Inclusive ϕ distribution; Bottom left: Inclusive p_T distribution; Bottom right: Leading muon p_T distribution.

lection criteria, a cut on $p_T^Z > 8$ GeV is imposed, to suppress the QCD multijet background further. The normalisation factors found in the Υ CR (Sec. 6.2.5) are applied to the Υ samples shown in Fig. 6.23. The Z p_T reweighting previously

described in Sec. 6.2.4 is applied. The method to retrieve the data-driven normalisation of the LO `Pythia` $b\bar{b}/c\bar{c}$ samples will be described Sec. 6.3.7 and the factor shown in Tab. 6.5 have been applied to produce these plots.
It is clearly visible that a poor agreement between data and MC is achieved, probably due to the wrong normalisation of the MC DY sample (please refer to Sec. 6.3.1 for further discussion). The poor description of the data provided by the MC is probably due to the breakdown of the theory predictions, given that the considered phase space is very tiny and close to the cut imposed on the lepton p_T. In addition to that, this invariant mass range probes the very high-p_T tail and it could be that `Powheg+Pythia` does not reproduce kinematic distributions so well, taking into account that the p_T of the di-muon system is much more greater than the invariant mass in this region. Furthermore, there is the possibility that further background sources need to be included. For example, one could be the production of $W^- + c$, with *W* and *c* decaying into μ^- and μ^+ respectively. This specific process could produce a high-p_T muon and a soft one, and give an opposite charge excess at large di-muon p_T. Anyway, it should be significantly suppressed by the isolation requirement and also it should not be appearing predominantly in the low mass bin.

6.3 QCD background estimation

The largest background in the considered invariant mass range is expected to be multijet + heavy flavour jets with leptonic decay of mesons. These mesons can decay semi-leptonically, where one of the quarks decays via a charged current weak decay to produce a charged lepton and a neutrino. This charged lepton can be mis-identified as coming form the DY interaction. Only energetic charged leptons produced from meson decays will be selected by the DY selection. The requirement that the muons are isolated removes a large fraction of potential multijet background events. However, multijet processes have large cross sections and therefore still remain a large background, particularly for lower values of di-muon invariant mass and muon p_T.
Another source of background comes from decays of $\pi^{\pm},\ K$ and hadrons misidentified as muons due to random coincidences between inner detector tracks and hits in the muon spectrometer. This background is reduced greatly by isolation requirements and using high-quality muon reconstruction working points. As this is expected to be symmetric between $+$ and $-$ charges, it can be easily subtracted by performing same-sign muons subtraction. On the contrary, the background from heavy flavours decays is expected to have an excess in the opposite-sign pairs, e.g. $gg \to b\bar{b} \to \mu^+\nu_\mu X, \mu^-\bar{\nu}X$. To extract this heavy flavour component, the fact that muons from heavy flavours decays originate from the decay vertices of the respective *B* hadrons and thus the two muons will not originate from a common vertex, as for the DY pairs, can be exploited.
Monte Carlo (MC) simulations are thought to only perform moderately well at simulating QCD background processes, and indeed are thought not to get the normalisation right (even though the event shapes may be reasonable). For this reason a *template fit* method has been developed to minimise the reliance on MC

predictions. The *template fit* provides a way of improving the QCD background estimation in Monte Carlo, by utilising supplementary information about the shape of the distributions given by the non isolated *Opposite-Sign* (OS) data, subtracted by *Same-Sign* (SS) data component.
In the following this fully data-driven method (we refer to this method as '*SuperFitter*' in the following) for the QCD background estimation is presented.

6.3.1 The SuperFitter method

An almost pure sample of QCD background events is selected in the data by requiring the full event selection except the `FixedCutTight` isolation cut. A *not-isolated* muon pair is made by two muons both failing one or more of the following requirements:

- `ptvarcone30/pt` < 0.06;
- `topoetcone20/pt` < 0.06;

The variable in use for discriminating the QCD background is defined as follows:

$$\chi^2_{QCD} = \left(\frac{d_0(\mu_1)}{\sigma_{d_0}(\mu_1)}\right)^2 + \left(\frac{d_0(\mu_2)}{\sigma_{d_0}(\mu_2)}\right)^2 + \left(\frac{\Delta z_0(\mu_1,\mu_2)}{\sigma(\Delta z_0(\mu_1,\mu_2))}\right)^2 \tag{6.2}$$

where the last term of the equation represents the significance of the Δz_0 of the pair, defined as follows:

$$\text{significance } \Delta z_0 = \frac{|z_0^{\mu_1} - z_0^{\mu_2}|}{\sqrt{(\sigma_{z_0}^{\mu_1})^2 + (\sigma_{z_0}^{\mu_2})^2}}. \tag{6.3}$$

Then, assuming a gaussian resolution on d_0 and z_0, for two muons originating from the same vertex, the distribution of χ^2_{QCD} follows a χ^2 distribution with 3 n.d.f. . Afterwards, the probability to be a 'signal-like' (flat between 0 and 1) or 'background-like' (peaked around 0) event has been evaluated as Prob$(\chi^2_{QCD}, \text{n.d.f.} = 3)$ and it has been used for estimating the QCD background. The distribution of this new variable can be found in Fig. 6.24.
This new variable is preferred to the $|\,d_0$ significance$\,|$ of the muon pair,

$$\frac{|d_0^{\mu_1} - d_0^{\mu_2}|}{\sqrt{(\sigma_{d_0}^{\mu_1})^2 + (\sigma_{d_0}^{\mu_2})^2}}, \tag{6.4}$$

which is dominated by the muon with the worst resolution.
Calling ϵ_{iso} the isolation efficiency and DY the number of signal events before any isolation requirement, the number of DY events after the isolation selection on two or zero muons (namely DY_{ii} and DY_{nn}) can be defined as follows:

$$\begin{aligned} DY_{ii} &\propto \epsilon_{iso}^2 \cdot DY \\ DY_{nn} &\propto (1-\epsilon_{iso})^2 \cdot DY \end{aligned} \tag{6.5}$$

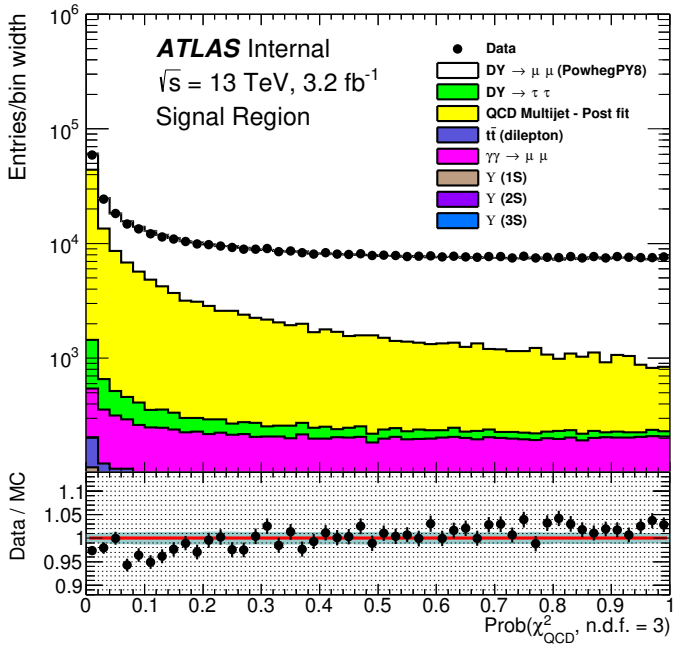


Fig. 6.24: Prob$(\chi^2_{QCD}, \text{n.d.f.} = 3)$ distributions for both data and sum of MC samples. All selection criteria are applied. The statistical uncertainties for the signal and background distributions are combined in the shaded band in the ratio panel (lower panel), and the statistical uncertainties are shown on the data points (vertical bars). The method to retrieve the data-driven normalisation of the LO `Pythia` $b\bar{b}/c\bar{c}$ samples will be described Sec. 6.3.7 and the factor shown in Tab. 6.5 have been applied to produce this plot.

So, under the assumption that the background shape of the variable chosen is independent from the isolation requirement, we can fit the distributions for the following two different cases, namely *isolated-isolated* (ii) and *not isolated-not isolated* (nn), with:

$$\begin{aligned} N_{ii} &= k \cdot DY^j_{ii} + b_{ii} \cdot B^j + EW^j_{ii} \\ N_{nn} &= k \cdot DY^j_{nn} + B^j + EW^j_{nn}, \end{aligned} \tag{6.6}$$

where j indicates a given bin of the Prob$(\chi^2_{QCD}, \text{n.d.f.} = 3)$ distribution, k is the DY signal strength, b_{ii} is the coefficient to normalise the ii background, B is the number of nn multijet background events and EW represents the sum of the following backgrounds: $t\bar{t}$, $Z/\gamma^* \to \tau^+\tau^-$, $\Upsilon \to \mu^+\mu^-$ and $\gamma\gamma \to \mu^+\mu^-$. N_{ii} and N_{nn} represent the number of *isolated-isolated* and *not isolated-not isolated* events respectively, while DY_{ii} and DY_{nn} are the MC predictions for ii and nn DY events. Then, the following χ^2 function can be defined for a given measurement bin in $(m_{\mu\mu}, |y_{\mu\mu}|)$:

$$\begin{aligned} \chi^2 = &\sum_j \left(\text{data}^j_{ii} - N^j_{ii}(k, b_{ii}, B^j)\right)^2 / \sigma^2_{ii}(j) \\ &+ \sum_j \left(\text{data}^j_{nn} - N^j_{nn}(k, B^j)\right)^2 / \sigma^2_{nn}(j), \end{aligned} \tag{6.7}$$

where j runs over all the bins in Prob(χ^2_{QCD}, n.d.f. $= 3$) and

$$\sigma^2_{ii}(j) = \left(\sigma^{data,j}_{ii}\right)^2 + \left(\sigma^{DY,MC,j}_{ii}\right)^2$$
$$\sigma^2_{nn}(j) = \left(\sigma^{data,j}_{nn}\right)^2 + \left(\sigma^{DY,MC,j}_{nn}\right)^2,$$

and

$$\text{n.d.f.} = 2 \cdot (\text{number data bins}) - \text{number background bins} - 2 = \text{N} - 2, \quad (6.8)$$

with σ representing the statistical uncertainty only and the *N* + 2 free parameters in the fits corresponding to B^j, k and b_{ii}.
The result of the fit to the *isolated-isolated* and *not isolated-not isolated* distributions for all the different invariant mass bins is shown in Fig. 6.25. The statistical uncertainties are approximately 2% for $14 < m_{\ell\ell} < 60$ GeV, $\approx$ 4% for $12 < m_{\ell\ell} < 14$ GeV and it is between 11% and 16% in the first invariant mass bins ($7 < m_{\ell\ell} < 9$ GeV). It is also visible that the normalisation of the MC DY samples is always around the unity, but not in the lowest invariant mass bin, where is close to 2. This could explain the disagreement between MC and data in the first bin of the invariant mass distributions shown in Fig. 6.20, in the first plot of Fig. 6.13 or in the kinematic distributions plotted in Fig. 6.23.

6.3.2 Comparison with the TFractionFitter method

In order to cross-check the DY sample normalisations found with the *SuperFitter* method presented above, the following method has been used. The relative fraction of $b\bar{b} + c\bar{c}$ and Drell-Yan (DY) MC required to describe the Prob(χ^2_{QCD}, n.d.f. $= 3$) variable in data are defined by

$$N^{data}_{ii} = f_1 \cdot DY_{ii} + f_2 \cdot (N^{data}_{nn} - N^{other\ MCs}_{nn}), \quad (6.9)$$

where f_1 is the fraction of the isolated DY events and f_2 is the fraction of the not isolated data, after the subtraction of all the other not isolated MC samples, including the DY sample, namely: $Z/\gamma^* \to \mu^+\mu^-$, $t\bar{t}$, $\Upsilon \to \mu^+\mu^-$ and $\gamma\gamma \to \mu^+\mu^-$ (an event is considered *not isolated* when both the muons in the pair do not satisfy the isolation requirement). A likelihood fit using the `TFractionFitter` tool within the `ROOT` framework is used to determine f_1 and f_2. f_1 and f_2 are defined in each measurement bin in ($m_{\mu\mu}$, $|y_{\mu\mu}|$), being f_1 the DY signal strength previously defined as k in Eq. 6.6. To compute the goodness of this QCD background estimate, a χ^2 function is defined comparing the data and MC distributions after this procedure. Eq. 6.9 assumes that the normalisation of the DY MC in $N^{other\ MCs}_{nn}$ is correct. As this is not true, the fit has to be repeated recursively. The convergence of the `TFractionFitter` method has been studied and it has been found that this method needs four iterations to converge e.g the normalisation of the DY MC sample does not change between the third and the fourth iteration within a tolerance of 0.1%.
A comparison between the normalisation of the DY MC sample found with these the two above-mentioned methods is shown in Fig. 6.26. As it is clearly visible

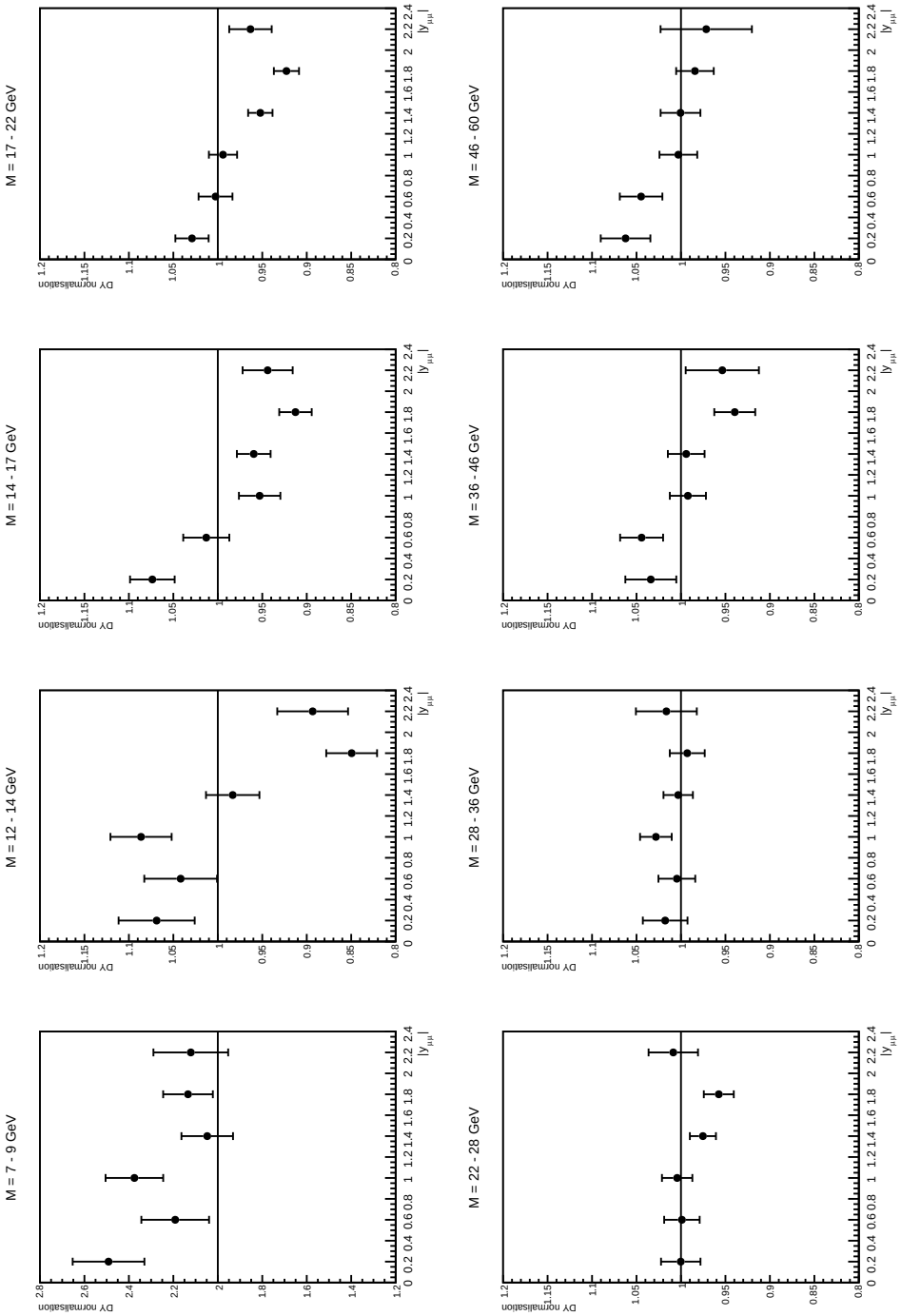


Fig. 6.25: DY sample normalisations found using the *SuperFitter* method (black dots) for all the eight invariant mass bins considered. The vertical bars represent the statistical uncertainty only. The horizontal solid line does not have any physical meaning: it has been drawn just as a reference in the middle of the y-axis to help the reader.

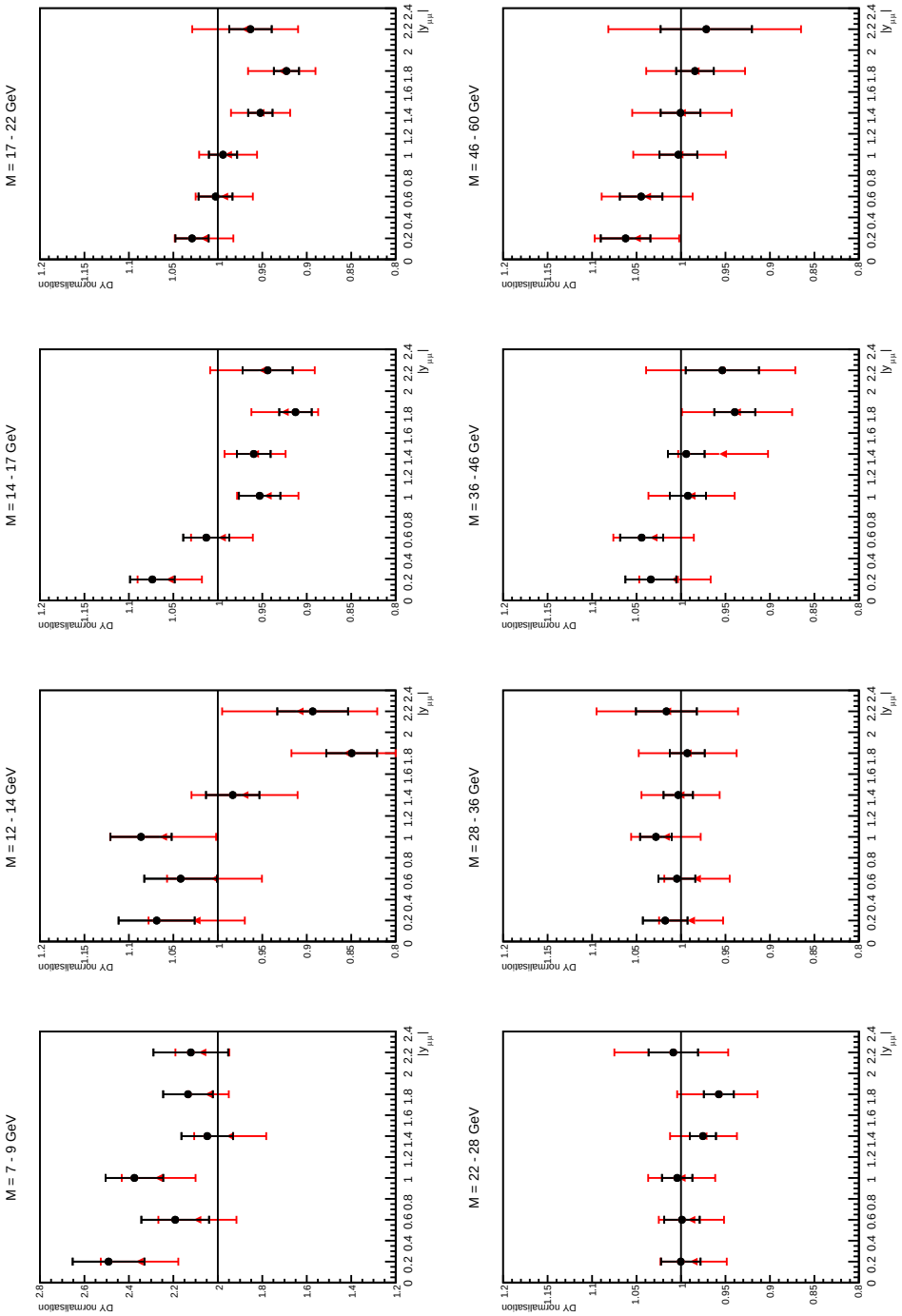


Fig. 6.26: Comparison between the DY sample normalisations found using the *SuperFitter* method (black dots) and the `TFractionFitter` method (red triangles) for all the eight invariant mass bins considered. The vertical bars represent the statistical uncertainty only. The horizontal solid line does not have any physical meaning: it has been drawn just as a reference in the middle of the y-axis to help the reader.

from Fig. 6.26, a nice agreement between the two methods has been observed. It can be also seen that the uncertainty on the DY normalisations computed with the `TFractionFitter` method are always larger than the ones estimated with the *SuperFitter* method.

6.3.3 Closure test

As a closure test, the fit results with pseudo data have been checked or so-called '*Asimov*' sample constructed from MC simulation:

$$\begin{aligned} \mathrm{Asim}_{ii}(x) &= k \cdot \mathrm{DY}_{ii}(x) + b_{ii} \cdot \mathrm{HF}_{nn}(x), \\ \mathrm{Asim}_{nn}(x) &= k \cdot \mathrm{DY}_{nn}(x) + \mathrm{HF}_{nn}(x), \end{aligned}$$

where DY is the MC DY signal, HF represents the $b\bar{b} + c\bar{c}$ MC background and the indices ii and nn stand for isolated-isolated and not isolated-not isolated muon pairs. The HF quantity for the isolated-isolated case is obtained by scaling the HF_{nn} quantity by arbitrary factors k_{ii}, given the assumption that the shape of the multijet background is the same for isolated-isolated and not isolated-not isolated events. Note that the error is not the actual MC error, but it is the error associated to a data sample with the number of events expected by the MC, so: $\sigma(x) = \sqrt{\mathrm{Asim}(x)}$. The results of this closure test are summarised in Fig. 6.27. As expected, the normalisation of the DY MC sample is compatible with unity in every $(m_{\ell\ell}, y_{\ell\ell})$ bin. This closure test proves that our fitting methodology does not introduce any bias in the normalisation of the DY MC sample.

6.3.4 A new tailored isolation working point

The possibility to use 'tailored' isolation working points with different cone sizes has been considered: in particular, it has been investigated if a larger cone size could have helped in discriminating the DY signal from the QCD multijet background better. The following two choices have been explored:

- `ptvarcone40`$/p_T$ < 0.06 (`iso WP1`)
- `ptvarcone40`$/p_T$ < 0.06 and `topoetcone30`$/p_T$ < 0.06 (`iso WP2`)

The former contains a requirement on the track isolation (expected to be easier for evaluating systematic uncertainties), while the latter requires both track and calorimeter isolation. These two above-mentioned isolation working points have been compared to the standard one provided by the `IsolationSelectionTool`[11], namely `FixedCutTight` (FCT) and `FixedCutTightTrackOnly` (FCTTO). The former isolation working point is the one in use in the analysis event selection, described in Sec. 6.2.
In order to understand which isolation working point would be the most suitable

[11] `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/IsolationSelectionTool`

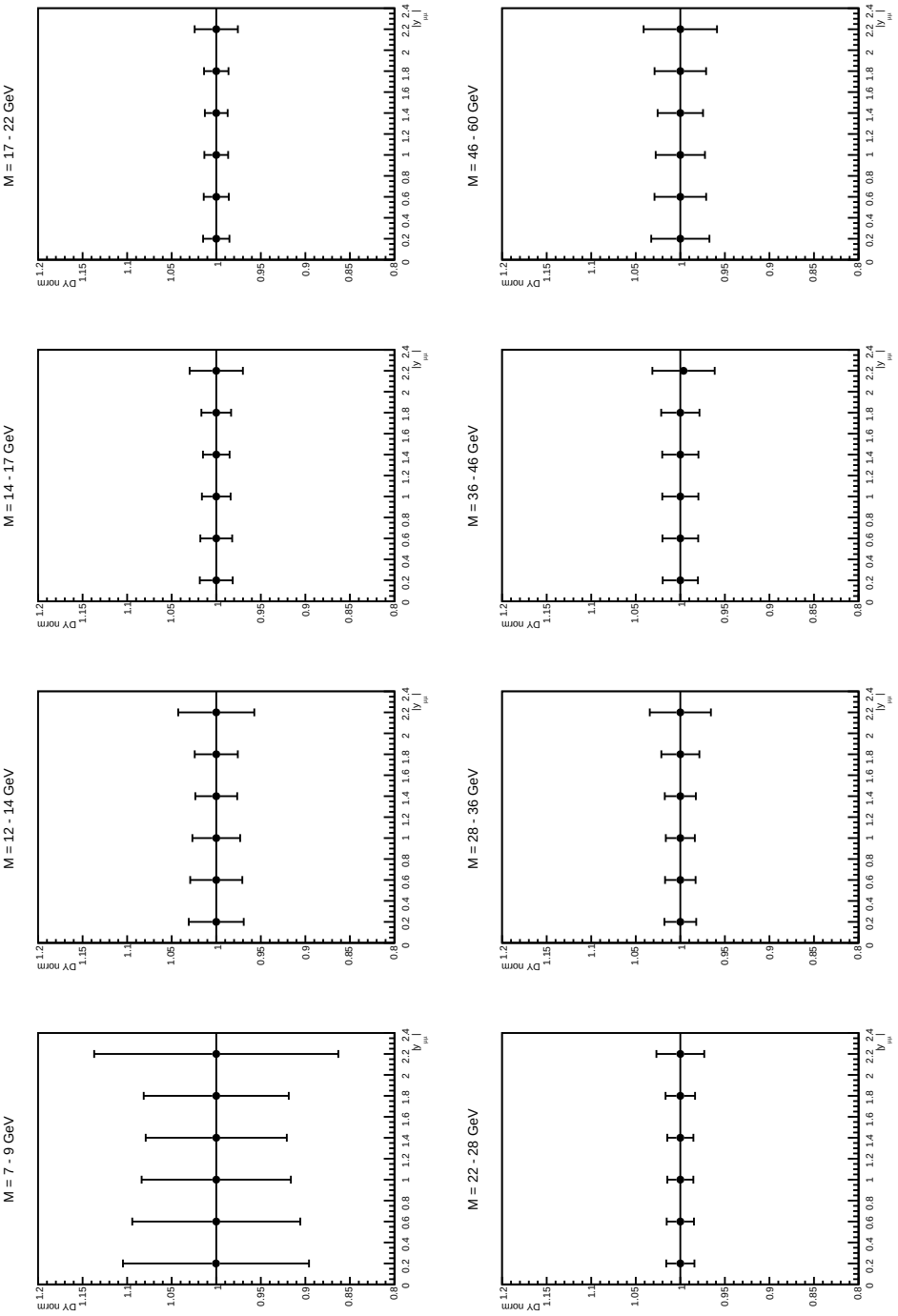


Fig. 6.27: Results of the closure test when replacing data sample with '*Asimov*' sample for all the eight invariant mass bins considered in the analysis. Note that the error is not the actual MC error, but it is the error associated to a data sample with the number of events expected by the MC, so: $\sigma(x) = \sqrt{\mathrm{Asim}(x)}$. The horizontal solid line does not have any physical meaning: it has been drawn just as a reference in the middle of the y-axis to help the reader.

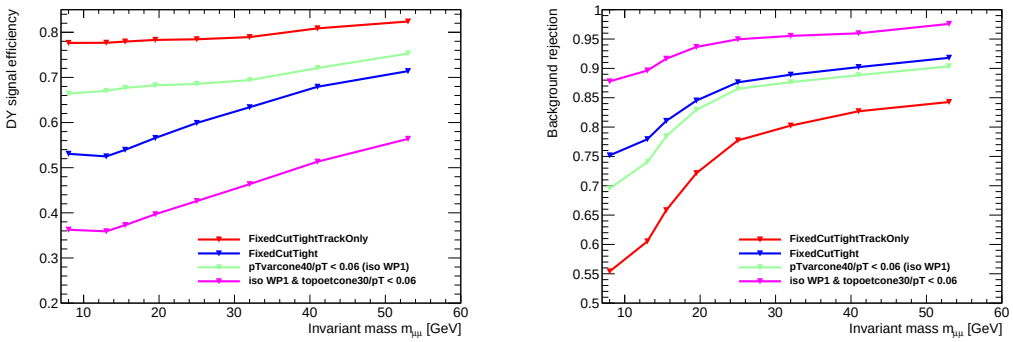


Fig. 6.28: (Left) DY signal efficiency and (Right) background rejection as a function of the invariant mass of the di-muon system for four different isolation WPs, namely FCTTO (red solid line), FCT (blue solid line), `iso WP1` (green solid line) and `iso WP2` (magenta solid line).

one for this analysis, the following two variables have been studied: the *DY signal efficiency* and the *background rejection*. They are defined as follows:

$$\epsilon_{DY} = \frac{DY_{ii}}{DY_{ii} + DY_{in} + DY_{nn}}$$
$$1 - \epsilon_{QCD} = 1 - \frac{QCD_{ii}}{QCD_{ii} + QCD_{in} + QCD_{nn}}$$

where *DY* is the number of signal events, *QCD* represents the number of QCD multijet events and the indices *in* all the events where one of the to muon of the pair is isolated and the other is not. Fig. 6.28 shows these two quantities as a function of the di-muon invariant mass for four different isolation WPs, namely FCTTO, FCT, and the two tailored ones, `iso WP1` and `iso WP2`. It can be seen how the `iso WP1` has a similar rejection power compared to FCT, but higher efficiency. The stability of this results has been studied as a function of the rapidity of the di-muon system for a given invariant mass bin and the outcome of this study for the lowest invariant mass bin (from 7 to 9 GeV) is reported in Fig. 6.29. The same conclusions as from Fig. 6.28 can be drawn: `iso WP1` has a $\sim$15% higher signal efficiency and a slightly lower rejection power (3-5%) with respect to FCT. Furthermore, the ratio between the efficiency on the DY signal and the square root of the efficiency on the $b\bar{b} + c\bar{c}$ background has been evaluated:

$$\frac{\epsilon(\mathrm{DY})}{\sqrt{\epsilon(b\bar{b} + c\bar{c})}} \tag{6.10}$$

and the outcome of this study is reported in Fig. 6.30, where it can be seen that the `iso WP1` is better than any other isolation WP in the most critical region for background subtraction ($m_{\mu\mu}$ < 25 GeV) and it is comparable with FCT for higher invariant masses. A concise summary of the efficiency for the four isolation working points considered in the lowest invariant mass bin can be found in Table 6.4.
For the two official isolation WPs (FCTTO and FCT), the statistical and systematic errors due to isolation have been assessed. Furthermore, the sum in quadrature

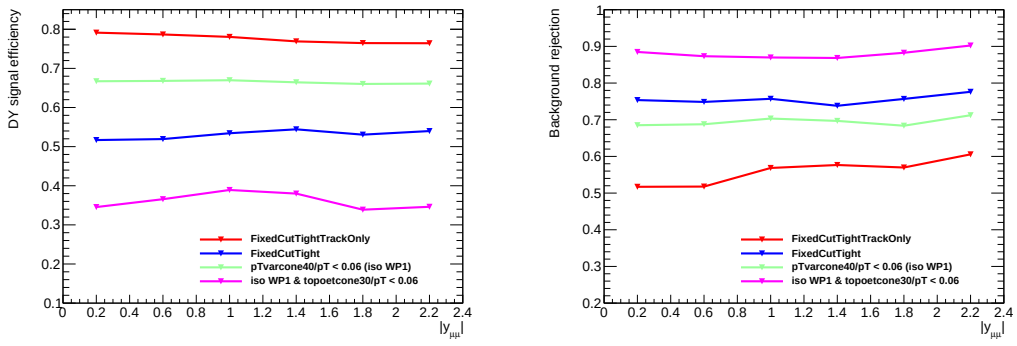


Fig. 6.29: (Left) DY signal efficiency and (Right) background rejection as a function of the rapidity of the di-muon system in the lowest invariant mass bin (from 7 to 9 GeV) for four different isolation WPs, namely FCTTO (red solid line), FCT (blue solid line), `iso WP1` (green solid line) and `iso WP2` (magenta solid line).

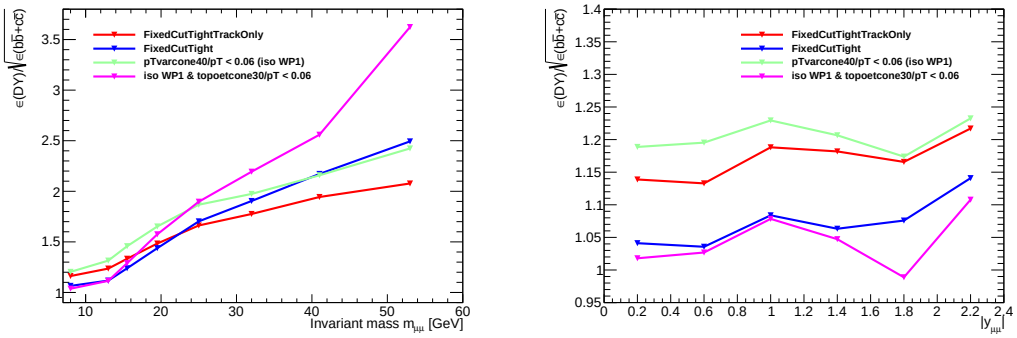


Fig. 6.30: Ratio between the efficiency on the DY signal and the square root of the efficiency on the $b\bar{b} + c\bar{c}$ background as a function of (Left) the invariant mass of the di-muon system and (Right) the di-muon system in the lowest invariant mass bin (from 7 to 9 GeV) for four different isolation WPs, namely FCTTO (red solid line), FCT (blue solid line), `iso WP1` (green solid line) and `iso WP2` (magenta solid line).

Efficiency	FCTTO	FCT	Iso WP1	Iso WP2
$\epsilon(\mathrm{DY})$	0.78	0.54	0.67	0.36
$\epsilon(\mathrm{DY})/\epsilon(b\bar{b} + c\bar{c})$	1.7	2.16	2.23	3.0

Table 6.4: Efficiency for the four isolation working points considered in the lowest invariant mass bin.

of these error has been compared with the statistical error on the DY normalisation coming from the *SuperFitter* method and it has been found that the optimal way to extract the DY normalisation is using FCT as default isolation WP, mainly because it is the one better performing in rejecting the QCD multijet background in the low mass region and it has smaller systematic uncertainties with respect to FCTTO. This study has not been conducted for the other two isolation WPs because systematic uncertainties were not available. It is visible from Fig. 6.30 and Table 6.4 that `iso WP1` has a higher signal efficiency and a comparable rejection

power with respect to FCT but it has not been used in the analysis because there were not official recommendations for its usage.
Indeed, the inclusion of this more stringent and better performing isolation WP in `Rel20.7` has been requested but our proposal was not accepted because additional manpower was needed.

6.3.5 Cut on maximum χ^2_{QCD}

The vertexing and the `iloose` isolation requirements at HLT level in our analysis trigger has been studied using the following reference trigger: `HLT_2mu6_10invm30_pt2_z10` (please refer to Sec. 6.4.2 for the description of this trigger and further discussion). It has been found that the trigger efficiency decreases for $\chi^2_{QCD} \simeq 15$ and it is not well-modelled by MC. In order to avoid this region where the data/MC agreement is poor because the trigger simulation is not perfect, a possible cut on χ^2_{QCD} has been considered.
Several cuts have been imposed and both their efficiency on the DY signal and their rejection power of the QCD background have been evaluated. In the end, it has been found that a cut $\chi^2_{QCD} < 16$ has a sufficiently large efficiency (larger than 98.5% for DY signal) and a rejection power of about 30%.

6.3.6 Improvements in the multijet background estimation

In order to maximize its discriminating power, different binning schemes for the Prob$(\chi^2_{QCD}, \text{n.d.f.} = 3)$ variable have been studied. A finer binning for lower values of Prob$(\chi^2_{QCD}, \text{n.d.f.} = 3)$ could help in discriminating the DY signal from the QCD multijet background, while wider bins in the high-Prob$(\chi^2_{QCD}, \text{n.d.f.} = 3)$ regime would avoid bins with few entries, which are difficult to be treated by the *SuperFitter* method. Three different scenarios have been tested:

- 100 fixed-width bins
- 200 fixed-width bins
- bins with variable width

Fig. 6.31 shows the comparison between the various χ^2/n.d.f. obtained with the above-mentioned methods, where it is visible that the χ^2/n.d.f. obtained with Prob$(\chi^2_{QCD}, \text{n.d.f.} = 3)$, $\chi^2_{QCD} < 16$ and adopting bins with variable width are the most stable and closest to unity.

6.3.7 The LO Pythia samples normalisation

Thanks to the `SuperFitter` method the number of QCD multijet events has been found so it was possible to scale the LO `Pythia` prediction by a fraction to exactly matched this number in each invariant mass and rapidity bins. The outcome of this study is shown in Fig. 6.32. Then, the weighted average of these scaling factor has been computed in each invariant mass bin and we applied the following factors reported in Tab. 6.5. As expected, the relative uncertainty on the

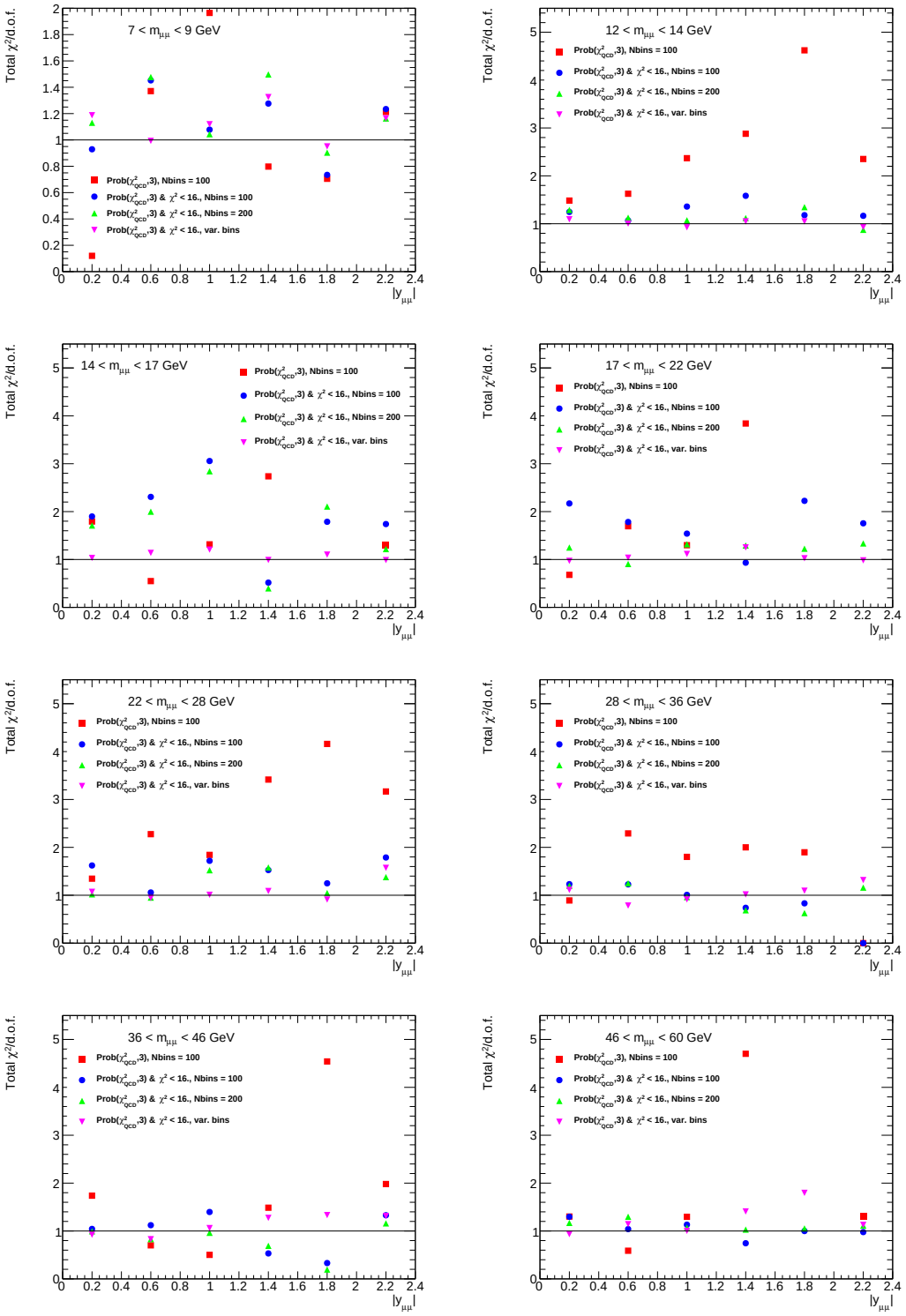


Fig. 6.31: Comparison between the χ^2/n.d.f. obtained using: Prob(χ^2_{QCD}, n.d.f. $= 3$) and 100 bins (red squares); Prob(χ^2_{QCD}, n.d.f. $= 3$). χ^2_{QCD} < 16 and 100 bins (blue circles); Prob(χ^2_{QCD}, n.d.f. $= 3$), χ^2_{QCD} < 16 and 200 bins (green triangles); Prob(χ^2_{QCD}, n.d.f. $= 3$), χ^2_{QCD} < 16 and bins with variable widths (magenta triangles).

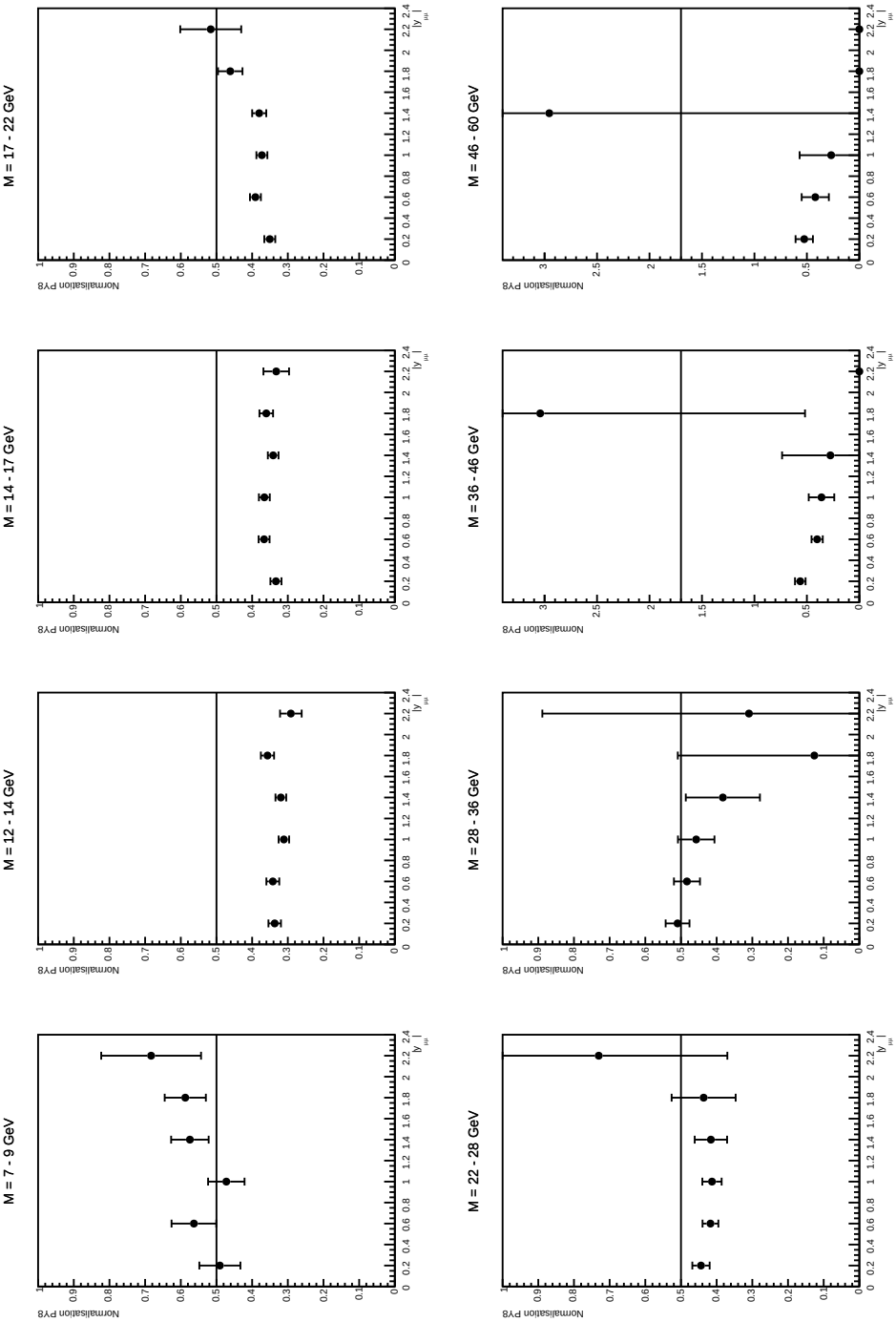


Fig. 6.32: Fractions used to scale the LO `Pythia` predictions in order to match the number of QCD multijet events found with the `SuperFitter` method. These fractions are shown as a function of the rapidity of the di-muon system for different invariant mass ranges. The vertical lines represent the statistical uncertainty. The horizontal solid line does not have any physical meaning: it has been drawn just as a reference in the middle of the y-axis to help the reader.

fractions used to scale the LO Pythia predictions gets larger when increasing the invariant mass and the rapidity. In the high-mass regime, the scaling factor applied to the MC $b\bar{b}/c\bar{c}$ sample is equal to 0, either because there are no events predicted by the simulation or the *SuperFitter* prediction for the multijet background event is null.

Mass bin [GeV]	Weighted Average	RMS
7 - 9	0.562	0.017
12 - 14	0.326	0.007
14 - 17	0.350	0.007
17 - 22	0.412	0.008
22 - 28	0.476	0.013
28 - 36	0.458	0.022
36 - 46	0.442	0.036
46 - 60	0.405	0.071

Table 6.5: Normalisation applied to the LO Pythia $b\bar{b}/c\bar{c}$ sample in each invariant mass bin with the associated error. The SF applied in each bin is the weighted average of the SFs presented in Fig. 6.32 and its error has been computed as the RMS in each bin.

6.4 Systematic uncertainties

The systematic uncertainty on the low-mass Drell-Yan cross section (σ_{DY}) comes mainly from the uncertainties on the trigger efficiencies and on the multijet background estimation. Other effects are due to inefficiencies for selecting collisions with a reconstructed primary vertex, as well as the muon momentum scale and resolution and the isolation efficiencies. The impact on the analysis is evaluated by varying scale factors within their uncertainties. The uncertainties are taken from the official tools provided by the performance groups. The total systematic uncertainty on σ_{DY} is then the sum in quadrature of the following contributions (the ATLAS nomenclature is given in brackets):

- **Muon momentum scale and resolution:** Variations in the smearing of the ID track (MUON_ID_STAT/SYS), variations in the smearing of the MS track (MUON_MS_STAT/SYS), and variations in the scale of the momentum (MUON_SCALE_STAT/SYS).
- **Muon reconstruction efficiency:** Statistical error on the scale factor (MUON_EFF_STAT), and systematic error on the scale factor (MUON_EFF_SYS). Other two systematics, due to measuring low-p_T muons (MUON_EFF_STAT/SYS_LOWPT) are added in quadrature.
- **Isolation scale factor:** Statistical error (MUON_ISO_STAT) and systematic error on the scale factor (MUON_ISO_SYS).

- **Sagitta correction:** Corrections for charge dependent local effects due to misalignments mainly in the Inner Detector and smaller local effects due to local misalignments in the Muon Spectrometer (MUON_SAGITTA_RHO and MUON_SAGITTA_RESBIAS).
- **Pileup:** The systematic uncertainty is defined as the difference between reweighting to $< \mu >$ with the nominal scale factor and scale factor varied according to the tracking-CP group recommendations, corrected for changes in the lepton identification and isolation efficiencies.
- **d_0 and z_0 smearing:** Variations in the smearing of the d_0 (TRK_RES_D0_MEAS) and z_0 track (TRK_RES_Z0_MEAS).
- **Luminosity:** A constant systematic uncertainty of 2.1% related to the luminosity is considered, as described here[12].
- **Electroweak background:** The normalisation of the DY$\rightarrow \tau\tau$, $t\bar{t}$, $\gamma\gamma \rightarrow \mu\mu$ and $\Upsilon \rightarrow \mu\mu$ MC samples has been varied by 5% up and down with respect to the nominal value.

In the above-describe list, the ATLAS label to identify the various systematics is reported. The uncorrelated component of the systematic uncertainty is labelled as *STAT*, while the correlated component is labelled as *SYS*.
The systematics uncertainties related to trigger efficiencies have not been assessed using the standard MCP tool, but they have been computed by other members of the analysis team, as reported in Sec. 6.4.1.

6.4.1 Trigger systematic uncertainties

The systematic uncertainties related to the trigger efficiency are computed merging together the uncertainties related to the *tag and probe* method done with muons from the Z boson decay and the uncertainties related to the *tag and probe* method done with muons from the J/ψ decay.
They include a statistical component and two different source of systematics. The first one (defined as '*TrigSyst1*' in the following) adds in quadrature the following contributions related to the background contamination of the samples:

- ***Z* boson *tag and probe***:
  - Cut on the $|d_0$ significance$|$ of the probe varied between 2 and 4 (nominal value: 3)
  - Variation of the invariant mass cut, from $80 < m_{\mu\mu} < 100$ GeV to $83 < m_{\mu\mu} < 97$ GeV.
- ***J*/ψ *tag and probe***:
  - Variation of minimal ΔR between tag and probe (0.24$\pm$0.02)

[12] https://twiki.cern.ch/twiki/bin/viewauth/Atlas/LuminosityForPhysics#2015_13_TeV_proton_proton_final

– Variation of sidebands from [2.7-2.9],[3.3-3.5] to [2.7-2.8],[3.4-2.5]

Furthermore, it includes a 3 per-mille uncertainty for the ID efficiency at the HLT-level in the low-p_T regime.
The latter one (defined as *'TrigSyst2'* in the following) takes into account variations of the azimuthal correlation between tag and probe: in the J/ψ region, it is done through removing the reweighted ϕ distribution of the probes, while throughout imposing a cut $|\Delta\phi - \phi| < 0.1$ between tag and probe in the Z region.
The effect of all these variations extremely small, less than 1%.

6.4.2 Vertexing and iloose requirement at HLT level

In order to evaluate the efficiency of the vertexing and the `iloose` isolation requirements at HLT level in our analysis trigger, we considered the following reference trigger: `HLT_2mu6_10invm30_pt2_z10` (we refer to this trigger as 'reference' trigger in the following), which is characterized by the following requirements:

- an invariant mass cut between 10 and 30 GeV
- the transverse momentum of the Z/γ^* boson p_T < 2 GeV
- $\Delta z_0(\mu_1, \mu_2)$ < 10 mm

This study has been conducted selecting only the Lumi Blocks (LBs) in which the analysis triggers were active. Furthermore, in order to avoid any possible threshold effect, the 12 < $m_{\mu\mu}$ < 28 GeV range has been considered in order to stay away from the mass cut applied at HLT level. The efficiency is calculated as follows:

$$\epsilon = \frac{N(\text{reference \&\& LMDY})}{N(\text{reference})} \tag{6.11}$$

where 'N(reference && LMDY)' is the number of events passing both the reference trigger and the analysis trigger and 'N(reference)' represents the events passing the reference trigger only.
Fig. 6.33 shows this efficiency as a function of Prob(χ^2_{QCD}, n.d.f. $= 3$) for both data and MC. A difference of about 2.5% between data and MC is visible for the region dominated by DY signal: Prob(χ^2_{QCD}, n.d.f. $= 3$) > 0.1. In order to evaluate a systematic uncertainty related to this correction, it has been checked this discrepancy between data and MC as a function of $|y_{\mu\mu}|$ for different invariant mass bins and the results can be found in Table 6.6, where the last column displays the SF computed as ($SF^{(max)}$ + $SF^{(min)}$)/2. Computing the RMS of the SFs it has been found a correction factor of (2.1 $\pm$ 0.3)% to be applied in every mass bin, regardless of the considered rapidity bin.

6.4.3 Possible shape differences between isolated and not isolated events

Given that the non isolated events have been used to extract the shape of the QCD background in the analysis signal region, another systematic to take into

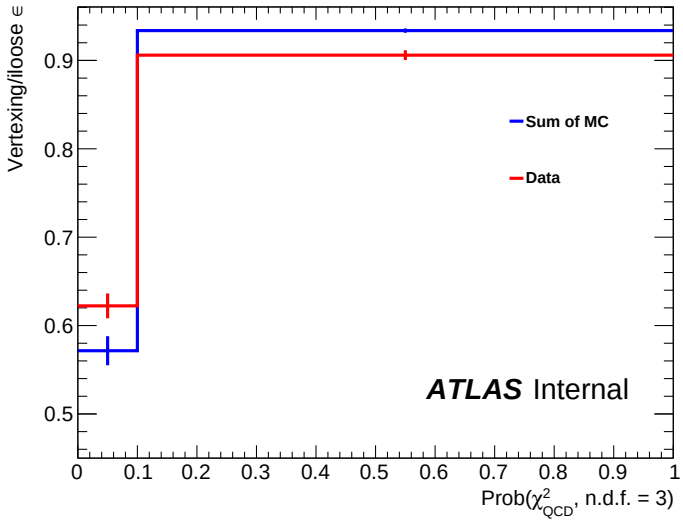


Fig. 6.33: Vertexing and isolation requirement efficiency (defined in Eq. 6.11) as a function of Prob(χ^2_{QCD}, n.d.f. $= 3$) for both data (solid red line) and MC (blue solid line).

Mass [GeV]	[0.0-0.4]	[0.4-0.8]	[0.8-1.2]	[1.2-1.6]	[1.6-2.0]	[2.0-2.4]	SF
12-14	0.968	0.994	0.986	0.976	0.968	0.970	0.981
14-17	0.978	0.979	0.983	0.982	0.974	0.979	0.979
17-22	0.979	0.977	0.972	0.981	0.975	0.967	0.976
22-28	0.989	0.963	0.971	0.962	0.952	1.004	0.979

Table 6.6: Discrepancy between data and MC as a function of $|y_{\mu\mu}|$ for fours different invariant mass bins. Last column displays the SF computed as (SF$^{(\max)}$ + SF$^{(\min)}$)/2.

account any possible difference in shape between isolated and not-isolated events has been considered. This study is fully MC-based and the ratio between isolated and not-isolated $b\bar{b}/c\bar{c}$ MC events is fitted with the following function:

$$f(x) = p_0 \cdot (1 + p_1 \cdot \exp(-\frac{x^2}{2 \cdot p_2}))$$

where x is the Prob(χ^2_{QCD}, n.d.f. $= 3$) variable. As a first step, a fit including all the invariant mass and rapidity bins has been performed, in order to extract the width of gaussian function (p_2 in the previous expression.) The results of this fit is presented in Fig. 6.34.

Then, fits in three different invariant mass regions, namely 7 to 14 GeV (with the region from 9 to 12 GeV excluded), 14 to 28 GeV and 28 to 60 GeV have been performed. In each invariant mass bin, two fits have been performed in two different rapidity regions, more precisely: 0.0 < $|y_{\mu\mu}|$ < 0.8 (first rapidity region) and 0.8 < $|y_{\mu\mu}|$ < 2.4 (second rapidity region). In these fits, the width of gaussian function has been fixed to the value found in Fig. 6.34. The result of this study is shown in Fig. 6.35. The up and down variations of the fitting functions are made in an anti-correlated way, so if p_1 was incremented by its associated error, p_2 was

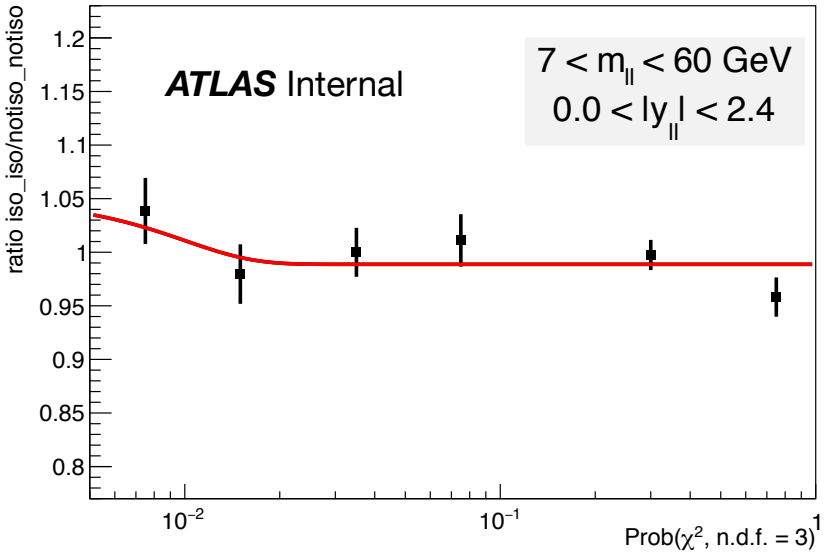


Fig. 6.34: Ratio of the isolated to not-isolated events as a function of Prob(χ^2_{QCD}, n.d.f. $= 3$) fitted with the following function: $f(x) = p_0 \cdot \left(1 + p_1 \cdot \exp(-\frac{x^2}{2 \cdot p_2})\right)$

decreased by its associated error (both p_1 and p_2 are defined in Eq. 6.4.3).
The impact of the up/down variations on the DY normalisation extracted with the *SuperFitter* method has been assessed and Fig. 6.36 shows the results of this study. The effect of these variations is tiny compared to the statistical uncertainty on the DY normalisation. In the 12 < $m_{\mu\mu}$ < 17 GeV region the effect is quite big, reaching the size of half the statistical uncertainty of the the DY normalisation. Lastly, it is visible that the values found for high invariant masses and higher rapidities may suffer from lack of statistics, as it is clearly visible also in Fig. 6.35. As previously discussed in Sec. 6.2.8, the QCD background estimation does not take into account the $W + jets$ events. This contribution can be estimated by MC simulations, but it is expected to be negligible in the considered invariant mass range, as it has been found in the previous low-mass DY measurement performed by ATLAS at 7 TeV [217].

6.4.4 Summary table

Table 6.7 - 6.8 summarises the dominant systematics contributions together with the extracted double-differential cross sections $\frac{d^2\sigma}{dm_{\mu\mu}d|y_{\mu\mu}|}$, while the impact of each source of systematic uncertainty on the extracted cross section is reported for each invariant mass bin as a function of the rapidity of the muon pair in Fig. 6.37 - 6.44. The measured fiducial double-differential cross sections will be discussed in Sec. 6.5.4.
Most of the sources of uncertainty have both correlated and uncorrelated components: the uncorrelated component is labelled as *Stat* in Fig. 6.37 - 6.44, while the

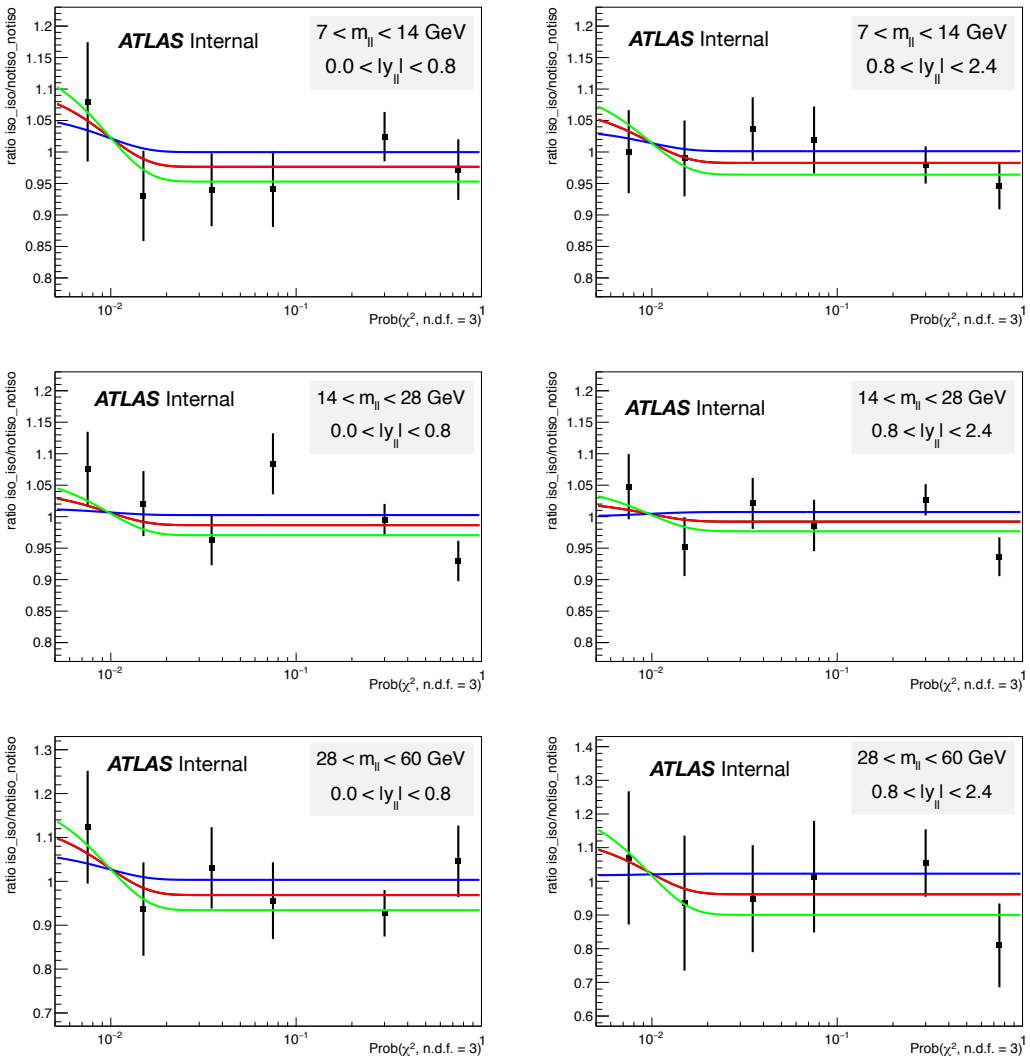


Fig. 6.35: Ratio of the isolated to not-isolated events as a function of Prob(χ^2_{QCD}, n.d.f. $= 3$) fitted with the function defined in Eq. 6.4.3, where the gaussian width has been fixed to the value found in Fig. 6.34. Also drawn are the up (solid blue line) and down (green solid line) variations of the fitting functions. The variations are made in an anti-correlated way.

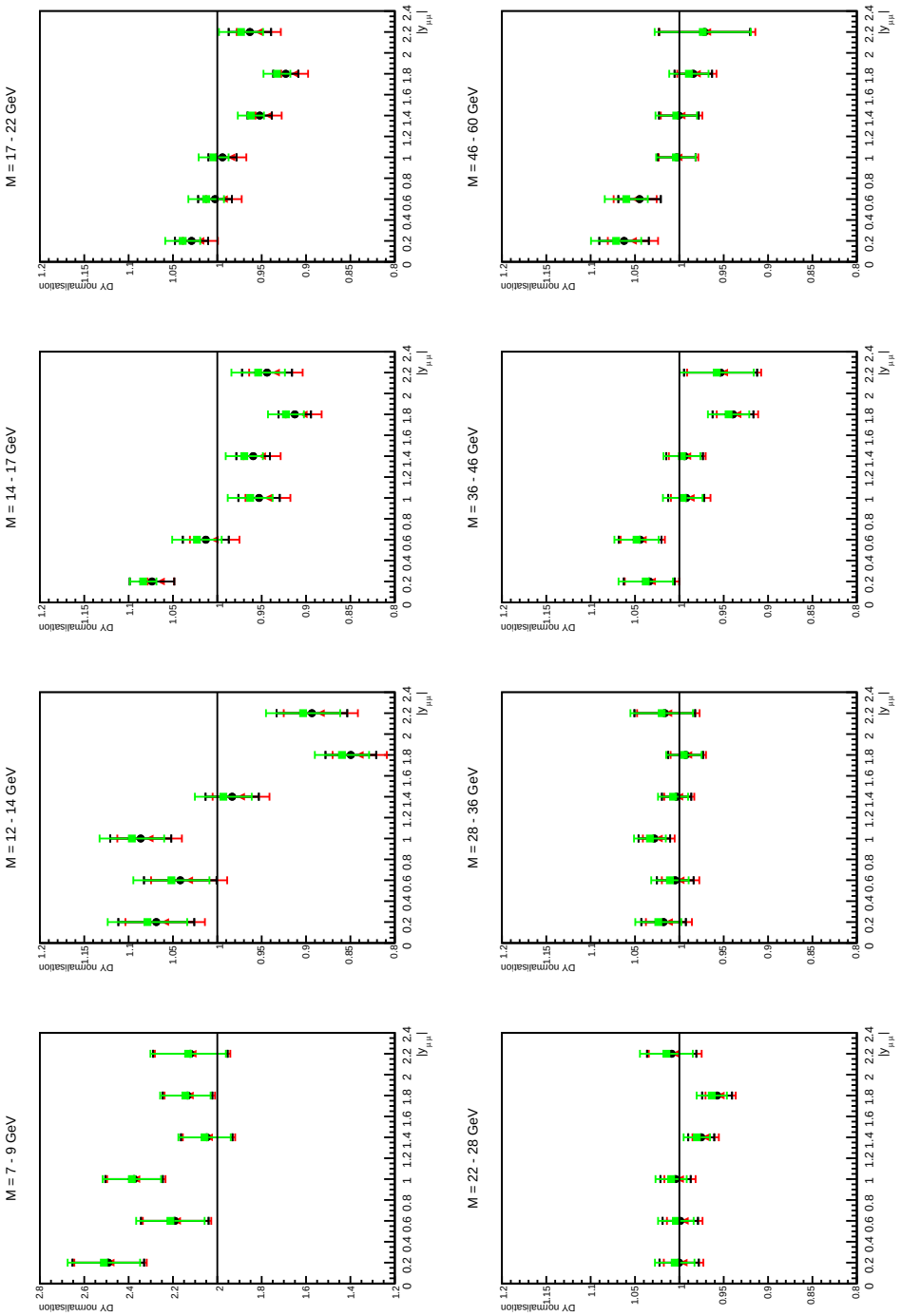


Fig. 6.36: DY normalisation computed using the *SuperFitter* method,, as described in Sec. 6.3: nominal (black dots), up shape variation (red triangles) and down shape variation (green box). These variations are done in an anti-correlated way. The horizontal solid line does not have any physical meaning: it has been drawn just as a reference in the middle of the y-axis to help the reader.

correlated component is labelled as *Syst*.

6.5 Drell-Yan Cross Section

The cross section measurements are extracted using a bin-by-bin unfolding technique due to high bin purity seen in both the one-dimensional (1D) and two-dimensionals (2D) binning schemes. The justification for the unfolding techniques used as well as the methodology for calculating the cross sections are described in this section. The two analysis cross sections that are measured are calculated using the following formulae:

$$\frac{d\sigma}{dm_{\mu\mu}} = \frac{N - B}{\mathcal{L} \cdot C_{DY} \cdot \Gamma_m} \tag{6.12}$$

$$\frac{d^2\sigma}{dm_{\mu\mu}d|y_{\mu\mu}|} = \frac{N - B}{\mathcal{L} \cdot C_{DY} \cdot \Gamma_m \cdot \Gamma_y} \tag{6.13}$$

where N represents the number of data event, B is the estimated background, $\mathcal{L}$ is the integrated luminosity of the data, C_{DY} is an unfolding correction factor (*acceptance* defined in Eq. 6.14) and Γ_m and Γ_y are the bin widths for the single differential mass measurement and the double differential mass and rapidity measurement respectively.

6.5.1 The Fiducial Volume

The cross sections measured in this analysis use the Born muon definition. The Born level defines the unfolded truth particle as a particle that has not undergone QED final state radiation (FSR). The following fiducial volume is used for the cross section measurement:

- Muon $p_T > 4.5$ GeV.
- Muon $|\eta| < 2.4$.
- Invariant mass $7 < m_{\mu\mu} < 9$ GeV or $12 < m_{\mu\mu} < 60$ GeV (with the invariant mass interval $9 < m_{\mu\mu} < 12$ excluded).

Moreover, a cut on the $p_T^Z > 8$ GeV is applied in the $7 < m_{\mu\mu} < 9$ GeV bin.
The di-muon mass range of the fiducial volume is chosen to complement measurements made of the DY cross section on the Z resonance [164] ($66 < m_{\mu\mu} < 116$ GeV).

6.5.2 Invariant mass and rapidity resolution

The resolution of the di-muon rapidity is shown in Fig. 6.45, while the di-muon mass resolution is shown in Fig. 6.46. The variables $m_{\mu\mu}^{gen}$ and $|y_{\mu\mu}^{gen}|$ are the Born generator-level values and variables $m_{\mu\mu}^{reco}$ and $|y_{\mu\mu}^{reco}|$ are the reconstructed-level quantities. The vertical bars shown in the plots represent the Root Mean Squared

Table 6.7: Summary of the statistical and the most relevant systematic uncertainties for the first four invariant mass bins (both up and down variation by 1σ). These uncertainties are listed together with the measured fiducial double-differential cross section $\frac{d^2\sigma}{dm_{\mu\mu}d|y_{\mu\mu}|}$. The fourth column reports the statistical uncertainty δ^{Stat}, while the fifth and the sixth column the total systematic uncertainty δ^{Sys} and the total uncertainty δ^{Tot} respectively. In the table just the systematic uncertainties with an effect greater than sub-percent level are reported. The luminosity uncertainty of 2.1% and the uncertainty associated to the efficiency of the vertexing and the `iloose` isolation requirements at HLT level in our analysis trigger (2.1% as well) are not shown.

$m_{\mu\mu}$ [GeV]	$\lvert y_{\mu\mu}\rvert$	$\frac{d^2\sigma}{dm_{\mu\mu}d\lvert y_{\mu\mu}\rvert}$ [pb/GeV]	δ^{Stat} [%]	δ^{Sys} [%]	δ^{Tot} [%]	$\delta^{\text{ID\_lowpT}}_{\text{unc}}$ [%]	$\delta^{\text{ID\_lowpT}}_{\text{cor}}$ [%]	$\delta^{\text{Iso}}_{\text{unc}}$ [%]	$\delta^{\text{Iso}}_{\text{corr}}$ [%]	$\delta^{\text{z0}}_{\text{unc}}$ [%]	$\delta^{\text{d0}}_{\text{unc}}$ [%]	$\delta^{\text{Trig}}_{\text{cor}}$ [%]	$\delta^{\text{TrigSyst1}}_{\text{unc}}$ [%]	$\delta^{\text{ShapeDiff}}_{\text{unc}}$ [%]
7 - 9	0.0 - 0.4	33.85	8.14	+12.88,-11.79	+15.24,-14.33	-2.40,+2.41	-1.07,+1.02	-7.76,+8.59	-4.13,+4.35	+2.70,-1.89	+0.98,-1.20	-6.24,+6.40	-1.26,+1.38	-1.35,+1.27
7 - 9	0.4 - 0.8	28.98	8.01	+12.89,-11.60	+15.18,-14.10	-1.95,+2.00	-0.75,+0.71	-7.71,+8.36	-4.08,+4.26	+3.39,-1.10	+1.34,-1.09	-6.35,+6.50	-1.28,+1.48	-1.38,+1.31
7 - 9	0.8 - 1.2	31.48	6.17	+11.52,-10.53	+13.07,-12.21	-1.51,+1.50	-0.90,+0.88	-7.08,+7.73	-3.80,+3.98	+3.15,-1.73	+1.14,-1.01	-5.03,+5.57	-1.23,+1.58	-0.88,+0.84
7 - 9	1.2 - 1.6	26.18	5.77	+10.63,-9.91	+12.10,-11.47	-1.31,+1.35	-1.15,+1.19	-7.08,+7.75	-3.79,+3.99	+2.38,-1.65	+1.41,-1.35	-3.45,+3.83	-1.19,+1.25	-1.14,+1.08
7 - 9	1.6 - 2.0	24.32	5.46	+10.63,-9.87	+11.95,-11.48	-1.34,+1.36	-1.28,+1.28	-7.23,+7.92	-3.88,+4.07	+2.15,-0.95	+1.49,-1.52	-3.23,+3.57	-1.15,+1.10	-0.85,+0.84
7 - 9	2.0 - 2.4	10.20	9.78	+11.42,-10.89	+15.04,-14.64	-1.63,+1.67	-1.34,+1.32	-7.20,+7.70	-3.76,+3.94	+3.25,-2.08	+2.70,-3.54	-4.23,+4.64	-1.37,+1.04	-0.82,+0.78
12 - 14	0.0 - 0.4	54.10	4.07	+9.24,-8.81	+10.09,-9.70	-1.85,+1.90	-0.69,+0.70	-4.70,+5.05	-2.86,+2.98	+0.83,-1.98	+0.84,-0.79	-5.33,+5.20	-1.01,+0.84	-2.12,+1.99
12 - 14	0.4 - 0.8	52.54	3.97	+8.97,-8.40	+9.81,-9.29	-1.42,+1.45	-0.53,+0.53	-4.54,+4.88	-2.79,+2.91	+1.23,-1.38	+0.83,-0.78	-5.22,+5.12	-0.93,+1.87	-1.91,+1.80
12 - 14	0.8 - 1.2	54.49	3.12	+8.62,-8.22	+9.17,-8.79	-1.14,+1.16	-0.65,+0.66	-4.67,+5.01	-2.82,+2.95	+1.66,-1.55	+0.93,-0.85	-4.43,+5.01	-0.93,+1.20	-1.26,+1.21
12 - 14	1.2 - 1.6	48.61	3.00	+8.20,-7.61	+8.73,-8.18	-1.02,+1.03	-0.93,+0.95	-4.61,+4.98	-2.83,+2.94	+2.38,-1.07	+1.11,-1.12	-3.44,+3.88	-0.85,+1.05	-1.23,+1.18
12 - 14	1.6 - 2.0	34.38	3.30	+8.19,-7.65	+8.83,-8.33	-1.03,+1.05	-1.05,+1.07	-4.52,+4.85	-2.81,+2.94	+2.48,-1.38	+1.52,-1.44	-3.34,+3.79	-0.83,+0.98	-1.34,+1.28
12 - 14	2.0 - 2.4	14.24	4.41	+8.38,-8.04	+9.48,-9.17	-1.25,+1.28	-1.03,+1.08	-4.23,+4.54	-2.67,+2.78	+2.05,-1.61	+1.86,-1.97	-4.21,+4.68	-0.93,+0.92	-0.97,+0.94
14 - 17	0.0 - 0.4	47.24	2.30	+8.54,-7.95	+8.81,-8.28	-1.56,+1.59	-0.61,+0.62	-3.63,+3.84	-2.34,+2.42	+1.92,-1.06	+0.94,-0.84	-5.32,+5.84	-1.23,+1.30	-0.68,+0.66
14 - 17	0.4 - 0.8	44.58	2.50	+8.35,-7.86	+8.72,-8.25	-1.28,+1.30	-0.57,+0.56	-3.67,+3.88	-2.34,+2.43	+1.65,-1.42	+0.88,-0.73	-5.16,+5.66	-1.09,+1.42	-0.83,+0.80
14 - 17	0.8 - 1.2	41.80	2.33	+8.05,-7.57	+8.38,-7.92	-1.03,+1.04	-0.71,+0.72	-3.70,+3.90	-2.37,+2.45	+2.08,-1.43	+1.06,-0.85	-4.53,+5.24	-1.04,+1.46	-0.79,+0.77
14 - 17	1.2 - 1.6	39.38	1.89	+7.53,-7.11	+7.77,-7.36	-0.97,+0.99	-1.00,+1.03	-3.58,+3.79	-2.31,+2.39	+2.23,-1.52	+1.26,-1.20	-3.83,+4.36	-1.07,+1.27	-0.65,+0.63
14 - 17	1.6 - 2.0	28.27	1.92	+7.61,-7.04	+7.85,-7.30	-1.02,+1.04	-1.20,+1.21	-3.36,+3.55	-2.20,+2.29	+2.53,-1.20	+1.79,-1.62	-3.85,+4.36	-1.12,+1.16	-0.58,+0.56
14 - 17	2.0 - 2.4	10.84	2.88	+8.44,-7.60	+8.91,-8.13	-1.24,+1.26	-1.18,+1.20	-3.20,+3.36	-2.07,+2.15	+3.09,-0.75	+1.89,-1.87	-4.91,+5.49	-1.32,+1.15	-0.42,+0.42
17 - 22	0.0 - 0.4	28.21	1.79	+8.30,-7.67	+8.49,-7.87	-1.30,+1.33	-0.60,+0.59	-2.81,+2.93	-1.83,+1.89	+1.97,-1.34	+1.05,-1.08	-5.65,+6.12	-1.10,+1.54	-0.56,+0.54
17 - 22	0.4 - 0.8	27.37	1.86	+8.13,-7.50	+8.34,-7.73	-1.20,+1.22	-0.65,+0.66	-2.84,+2.97	-1.82,+1.90	+2.19,-1.55	+1.11,-1.10	-5.35,+5.78	-1.06,+1.55	-0.71,+0.69
17 - 22	0.8 - 1.2	26.61	1.55	+7.77,-7.09	+7.92,-7.26	-1.02,+1.03	-0.80,+0.81	-2.77,+2.91	-1.82,+1.87	+2.49,-1.37	+1.20,-1.09	-4.80,+5.32	-1.12,+1.58	-0.50,+0.49
17 - 22	1.2 - 1.6	22.57	1.40	+7.17,-6.66	+7.30,-6.81	-0.95,+0.97	-1.03,+1.05	-2.66,+2.75	-1.73,+1.78	+2.24,-1.24	+1.52,-1.46	-4.22,+4.61	-1.18,+1.29	-0.40,+0.39
17 - 22	1.6 - 2.0	15.49	1.49	+7.25,-6.53	+7.40,-6.69	-1.00,+1.01	-1.24,+1.26	-2.44,+2.54	-1.60,+1.65	+2.72,-0.77	+1.88,-1.71	-4.10,+4.48	-1.23,+1.13	-0.37,+0.36
17 - 22	2.0 - 2.4	5.87	2.41	+8.20,-7.27	+8.55,-7.66	-1.15,+1.17	-1.16,+1.17	-2.35,+2.40	-1.55,+1.55	+3.31,-1.14	+2.21,-2.15	-5.03,+5.53	-1.20,+1.18	-0.28,+0.28

Table 6.8: Summary of the statistical and the most relevant systematic uncertainties for the first four invariant mass bins (both up and down variation by 1σ). These uncertainties are listed together with the measured fiducial double-differential cross section $\frac{d^2\sigma}{dm_{\mu\mu}d|y_{\mu\mu}|}$. The fourth column reports the statistical uncertainty δ^{Stat}, while the fifth and the sixth column the total systematic uncertainty δ^{Sys} and the total uncertainty δ^{Tot} respectively. In the table just the systematic uncertainties with an effect greater than sub-percent level are reported. The luminosity uncertainty of 2.1% and the uncertainty associated to the efficiency of the vertexing and the `iloose` isolation requirements at HLT level in our analysis trigger (2.1% as well) are not shown.

$m_{\mu\mu}$ [GeV]	$\lvert y_{\mu\mu}\rvert$	$\frac{d^2\sigma}{dm_{\mu\mu}d\lvert y_{\mu\mu}\rvert}$ [pb/GeV]	δ^{Stat} [%]	δ^{Sys} [%]	δ^{Tot} [%]	$\delta^{\mathrm{ID\_lowpT}}_{\mathrm{unc}}$ [%]	$\delta^{\mathrm{ID\_lowpT}}_{\mathrm{cor}}$ [%]	$\delta^{\mathrm{Iso}}_{\mathrm{unc}}$ [%]	$\delta^{\mathrm{Iso}}_{\mathrm{corr}}$ [%]	$\delta^{\mathrm{z_0}}_{\mathrm{unc}}$ [%]	$\delta^{\mathrm{d_0}}_{\mathrm{unc}}$ [%]	$\delta^{\mathrm{Trig}}_{\mathrm{cor}}$ [%]	$\delta^{\mathrm{TrigSyst1}}_{\mathrm{unc}}$ [%]	$\delta^{\mathrm{ShapeDiff}}_{\mathrm{unc}}$ [%]
22 - 28	0.0 - 0.4	13.95	2.18	+8.22,-7.39	+8.50,-7.70	-1.10,+1.13	-0.58,+0.58	-2.71,+2.33	-1.37,+1.47	+2.32,-1.63	+1.41,-1.18	-5.72,+6.15	-0.88,+1.65	-0.54,+0.52
22 - 28	0.4 - 0.8	13.88	1.97	+7.98,-7.13	+8.22,-7.40	-1.10,+1.12	-0.72,+0.74	-2.25,+2.34	-1.36,+1.46	+2.49,-1.19	+1.41,-1.22	-5.42,+5.78	-0.97,+1.51	-0.50,+0.49
22 - 28	0.8 - 1.2	13.28	1.66	+7.61,-6.88	+7.79,-7.08	-1.01,+1.03	-0.78,+0.81	-2.16,+2.23	-1.36,+1.40	+2.53,-1.23	+1.31,-1.36	-5.00,+5.45	-1.10,+1.56	-0.36,+0.35
22 - 28	1.2 - 1.6	10.87	1.47	+6.87,-6.31	+7.02,-6.48	-0.87,+0.90	-0.84,+0.87	-1.91,+2.04	-1.29,+1.29	+2.01,-0.97	+1.51,-1.46	-4.44,+4.75	-1.07,+1.34	-0.19,+0.19
22 - 28	1.6 - 2.0	7.34	1.72	+6.81,-6.06	+7.01,-6.30	-0.87,+0.88	-0.86,+0.90	-1.85,+1.92	-1.03,+1.17	+2.46,-0.70	+2.06,-1.81	-4.13,+4.38	-0.99,+1.07	-0.15,+0.14
22 - 28	2.0 - 2.4	2.73	2.69	+7.04,-6.43	+7.54,-6.87	-0.94,+0.95	-0.80,+0.79	-1.72,+1.83	-0.79,+1.10	+1.62,-0.36	+1.84,-1.51	-4.85,+5.22	-0.77,+1.04	-0.12,+0.11
28 - 36	0.0 - 0.4	6.93	2.52	+7.91,-7.11	+8.30,-7.55	-0.80,+0.80	-0.46,+0.49	-1.84,+2.03	-1.09,+1.18	+2.12,-1.52	+1.21,-1.15	-5.83,+5.79	-0.74,+1.14	-1.21,+1.16
28 - 36	0.4 - 0.8	6.73	2.10	+7.56,-6.88	+7.85,-7.19	-0.83,+0.87	-0.63,+0.67	-1.84,+1.95	-1.22,+1.17	+1.65,-1.04	+1.45,-1.18	-5.57,+5.62	-0.73,+1.06	-0.86,0.83
28 - 36	0.8 - 1.2	6.34	1.68	+7.15,-6.58	+7.35,-6.79	-0.78,+0.79	-0.56,+0.60	-1.75,+1.83	-1.14,+1.06	+1.58,-0.64	+1.23,-1.15	-5.19,+5.40	-0.78,+1.03	-0.83,+0.77
28 - 36	1.2 - 1.6	5.20	1.60	+6.62,-6.11	+6.81,-6.32	-0.63,+0.64	-0.37,+0.54	-1.73,+1.78	-1.05,+0.97	+1.57,-0.76	+1.60,-1.39	-4.57,+4.74	-0.70,+0.95	-0.57,+0.47
28 - 36	1.6 - 2.0	3.47	1.92	+5.79,-5.83	+ 6.10,-6.14	-0.55,+0.49	-0.57,+0.48	-1.76,+1.52	-0.96,+0.85	+0.83,-1.20	+0.95,-1.86	-4.11,+4.28	-0.53,+0.70	-0.21,+0.11
28 - 36	2.0 - 2.4	1.26	3.26	+6.17,-6.31	+6.98,-7.10	-0.48,+0.48	-0.40,+0.47	-1.66,+1.56	-0.82,+0.83	+0.34,-0.87	+0.65,-2.02	-4.63,+4.87	-0.31,+0.56	-0.46,+0.31
36 - 46	0.0 - 0.4	3.37	2.84	+7.29,-6.61	+7.82,-7.20	-0.44,+0.42	-0.42,+0.30	-1.87,+1.69	-0.94,+0.93	+2.08,-0.92	+0.95,-0.99	-5.48,+5.45	-0.51,+0.62	-0.95,+0.92
36 - 46	0.4 - 0.8	3.27	2.35	+7.19,-6.36	+ 7.56,-6.78	-0.45,+0.45	-0.22,+0.33	-1.49,+1.78	-0.87,+0.88	+2.03,-0.52	+1.26,-0.90	-5.29,+5.37	-0.53,+0.60	-0.58,+0.55
36 - 46	0.8 - 1.2	2.82	2.00	+6.52,-6.32	+6.82,-6.63	-0.46,+0.45	-0.36,+0.32	-1.64,+1.49	-0.94,+0.79	+1.08,-1.28	+1.06,-1.10	-4.92,+5.14	-0.50,+0.51	-0.24,+0.23
36 - 46	1.2 - 1.6	2.36	2.03	+5.85,-5.88	+6.19,-6.22	-0.31,+0.30	-0.20,+0.18	-1.59,+1.49	-0.74,+0.74	+0.15,-1.21	+0.19,-1.18	-4.38,+4.57	-0.33,+0.39	-0.13,+0.11
36 - 46	1.6 - 2.0	1.51	2.38	+5.60,-5.41	+6.09,-5.91	-0.20,+0.13	-0.12,+0.16	-1.52,+1.41	-0.21,+0.62	+1.23,-0.88	+1.23,-1.21	-3.82,+3.93	-0.20,+0.23	-0.43,+0.35
36 - 46	2.0 - 2.4	0.54	4.16	+7.24,-5.13	+8.35,-6.60	-0.16,+0.16	+0.54,+0.32	+0.01,+2.07	-0.37,+0.62	+1.69,+0.89	+2.28,+0.21	-4.11,+4.28	+0.09,-0.16	+0.28,+0.24
46 - 60	0.0 - 0.4	1.88	2.68	+6.32,-5.96	+6.87,-6.53	-0.25,+0.25	-0.26,+0.35	-1.46,+1.35	-0.52,+0.69	+1.06,-0.63	+0.95,-0.89	-4.82,+4.81	-0.33,+0.30	-0.59,+0.55
46 - 60	0.4 - 0.8	1.77	2.33	+5.98,-5.92	+6.42,-6.37	-0.19,+0.28	-0.26,+0.20	-1.60,+1.13	-0.65,+0.67	+0.43,-0.71	+0.65,-1.03	-4.61,+4.74	-0.34,+0.32	-0.19,+0.18
46 - 60	0.8 - 1.2	1.54	2.09	+5.88,-5.55	+6.24,-5.93	-0.22,+0.18	-0.07,+0.05	-1.34,+1.21	-0.58,+0.54	+0.93,-0.004	+1.03,-0.65	-4.32,+4.50	-0.24,+0.25	-0.01,+0.02
46 - 60	1.2 - 1.6	1.26	2.20	+5.19,-5.21	+5.64,-5.66	-0.10,+0.16	-0.10,+0.05	-1.35,+1.08	-0.59,+0.51	+0.30,-0.19	+0.51,-1.17	-3.73,+3.86	-0.19,+0.14	-0.07,-0.06
46 - 60	1.6 - 2.0	0.81	2.16	+4.78,-4.81	+5.25,-5.27	-0.08,+0.08	-0.06,+0.08	-0.84,+0.99	-0.39,+0.43	+0.05,-1.20	+0.13,-0.23	-3.32,+3.42	-0.21,+0.11	+0.001,-0.001
46 - 60	2.0 - 2.4	0.28	5.11	+4.94,-4.84	+7.11,-7.04	-0.06,+0.06	+0.14,+0.25	-1.04,+0.99	-0.51,+0.39	+0.02,+0.18	+0.20,-0.20	+3.44,+3.61	-0.12,+0.09	+0.04,+0.02

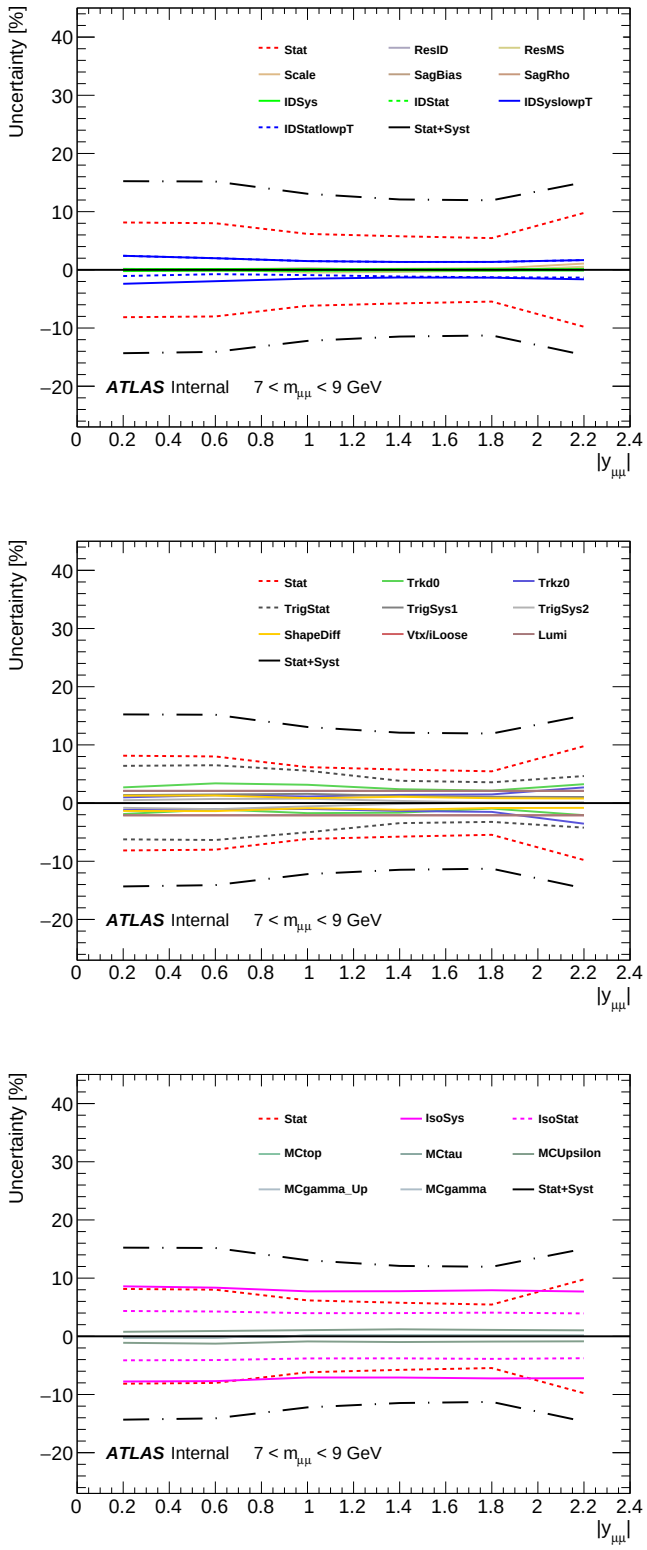


Fig. 6.37: Impact of the systematic uncertainties on the extracted cross section as a function of $|y_{\mu\mu}|$ for the $7 < m_{\mu\mu} < 9$ GeV bin. The black solid-dotted line represents the sum in quadrature of statistical and systematic uncertainties in this specific bin.

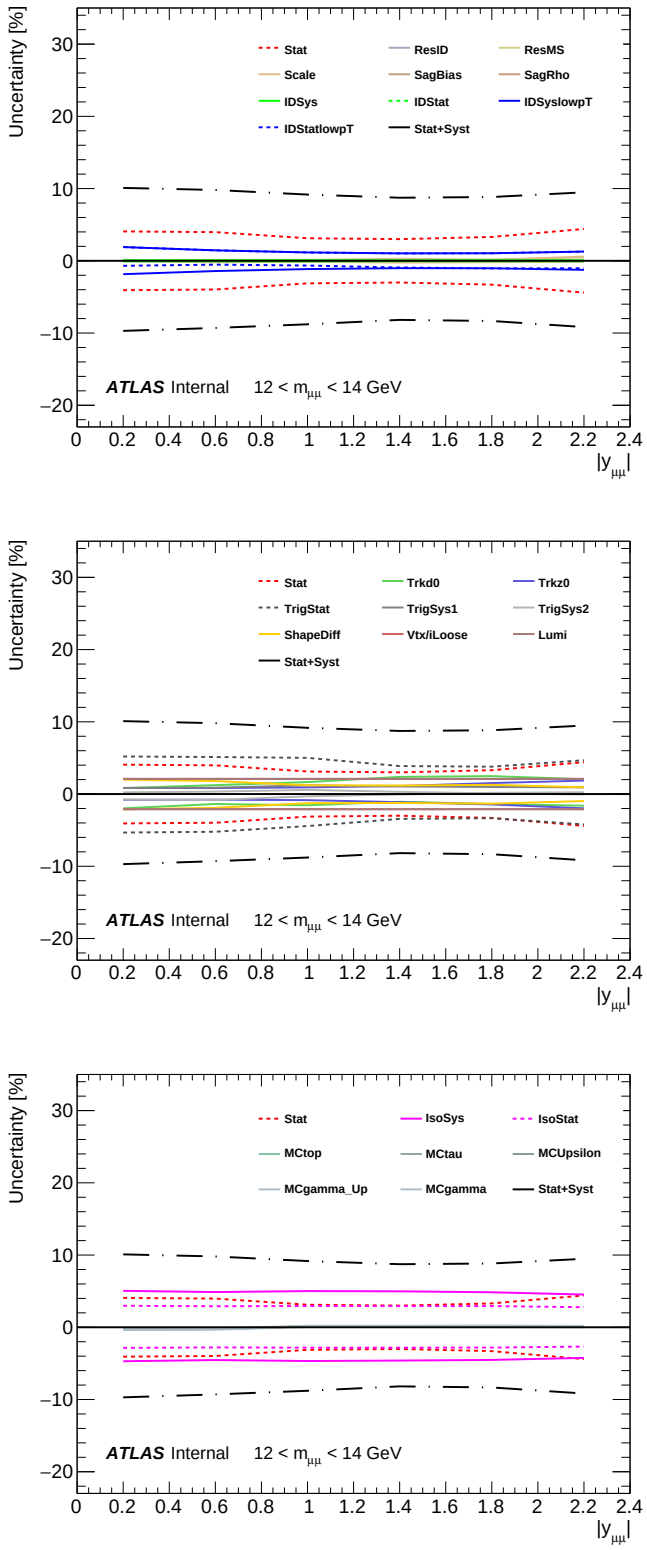


Fig. 6.38: Impact of the systematic uncertainties on the extracted cross section as a function of $|y_{\mu\mu}|$ for the $12 < m_{\mu\mu} < 14$ GeV bin. The black solid-dotted line represents the sum in quadrature of statistical and systematic uncertainties in this specific bin.

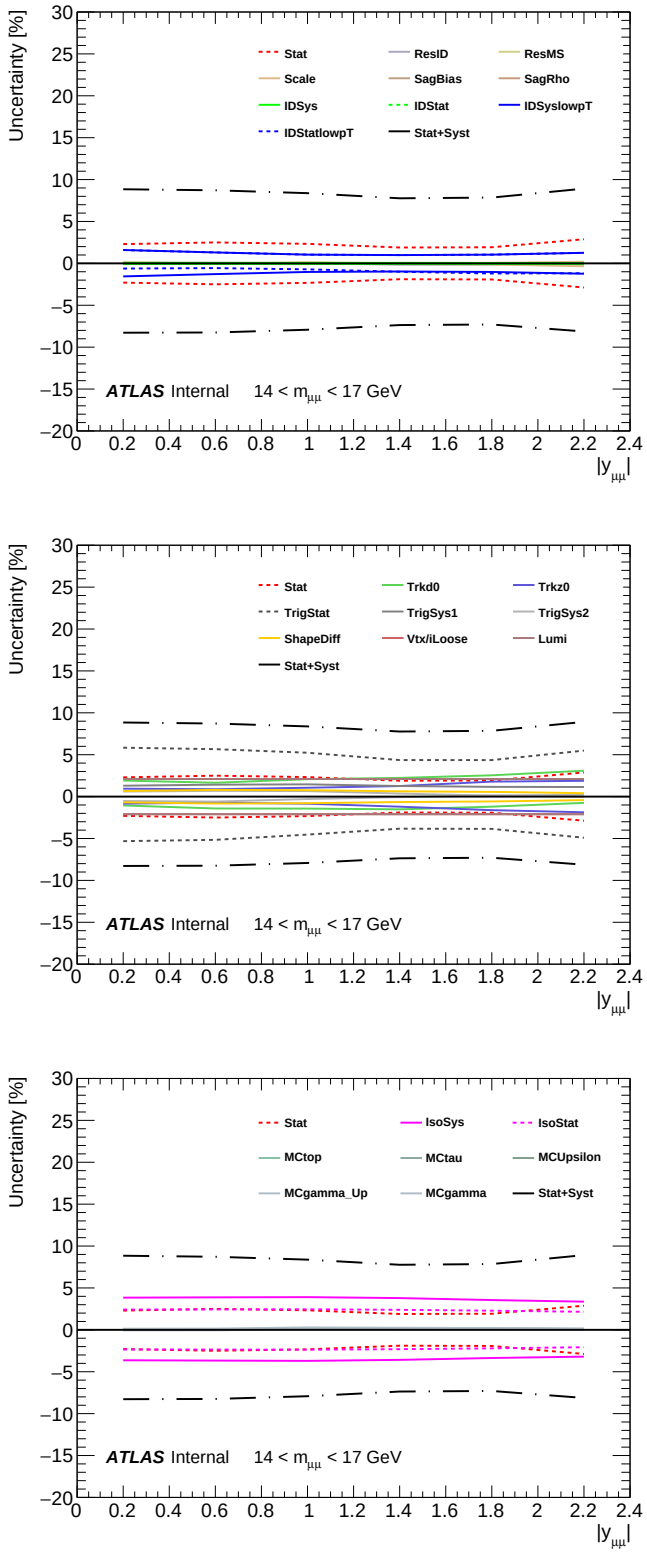


Fig. 6.39: Impact of the systematic uncertainties on the extracted cross section as a function of $|y_{\mu\mu}|$ for the $14 < m_{\mu\mu} < 17$ GeV bin. The black solid-dotted line represents the sum in quadrature of statistical and systematic uncertainties in this specific bin.

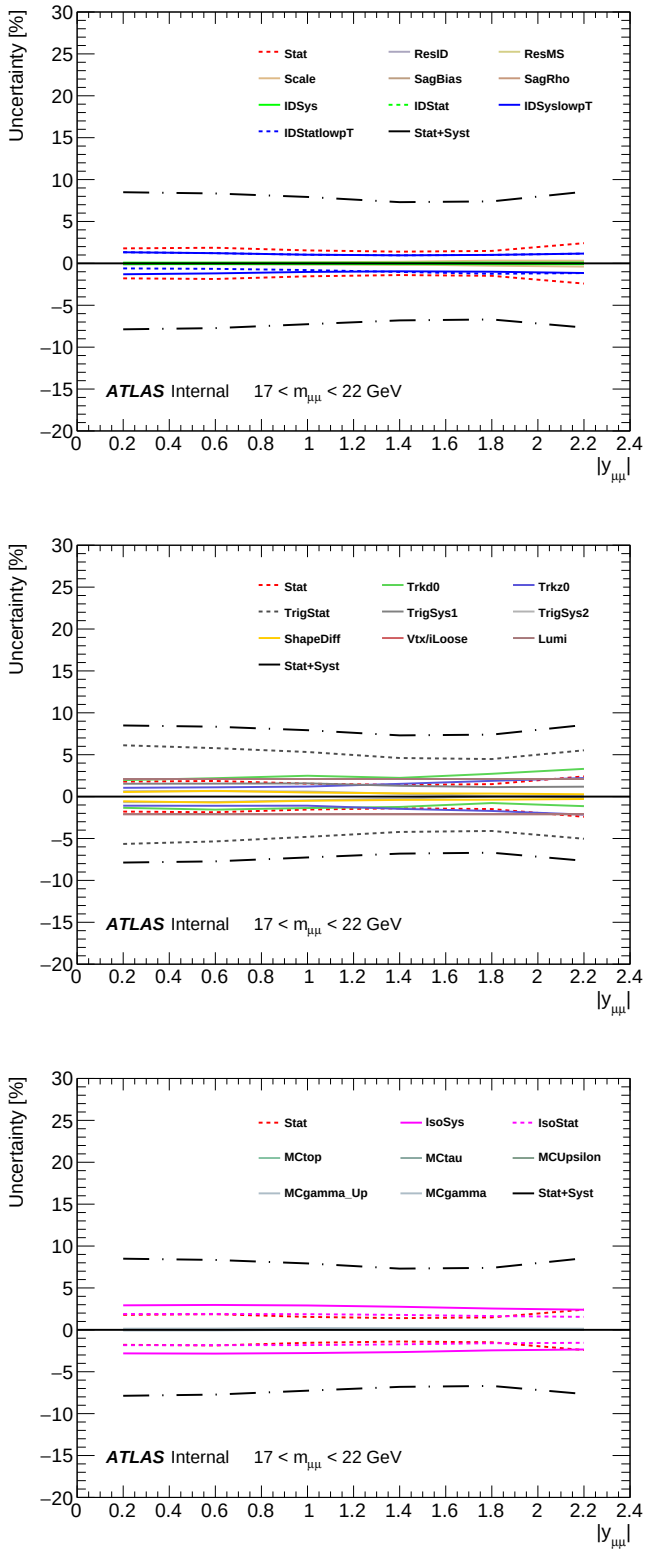


Fig. 6.40: Impact of the systematic uncertainties on the extracted cross section as a function of $|y_{\mu\mu}|$ for the $17 < m_{\mu\mu} < 22$ GeV bin. The black solid-dotted line represents the sum in quadrature of statistical and systematic uncertainties in this specific bin.

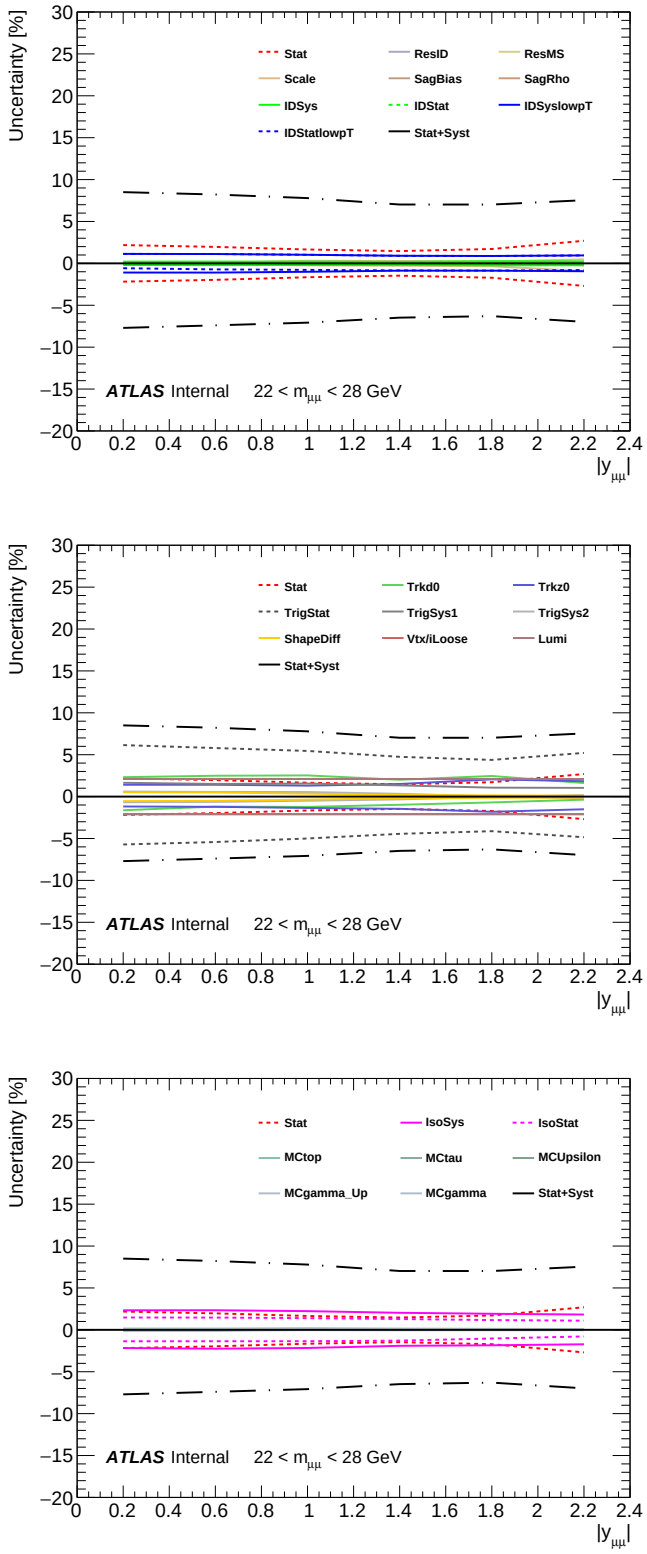


Fig. 6.41: Impact of the systematic uncertainties on the extracted cross section as a function of $|y_{\mu\mu}|$ for the $22 < m_{\mu\mu} < 28$ GeV bin. The black solid-dotted line represents the sum in quadrature of statistical and systematic uncertainties in this specific bin.

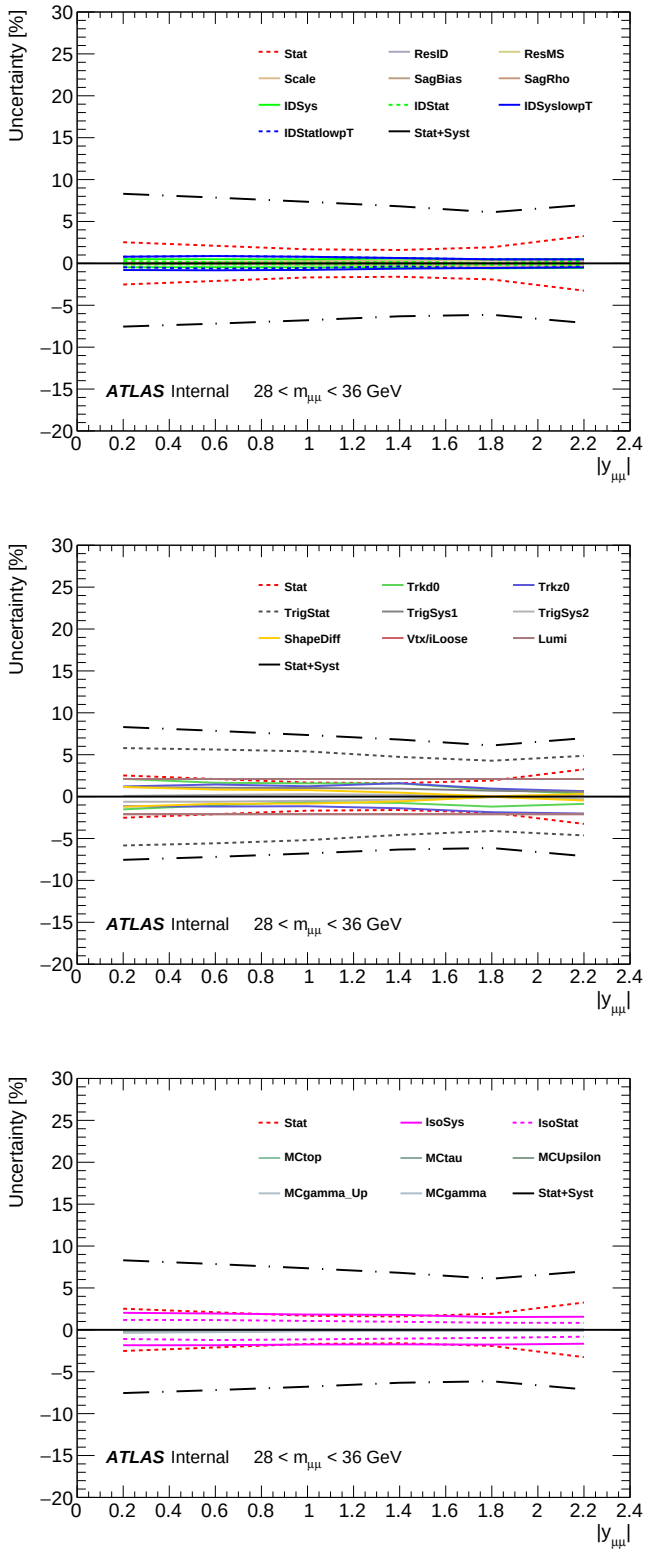


Fig. 6.42: Impact of the systematic uncertainties on the extracted cross section as a function of $|y_{\mu\mu}|$ for the $28 < m_{\mu\mu} < 36$ GeV bin. The black solid-dotted line represents the sum in quadrature of statistical and systematic uncertainties in this specific bin.

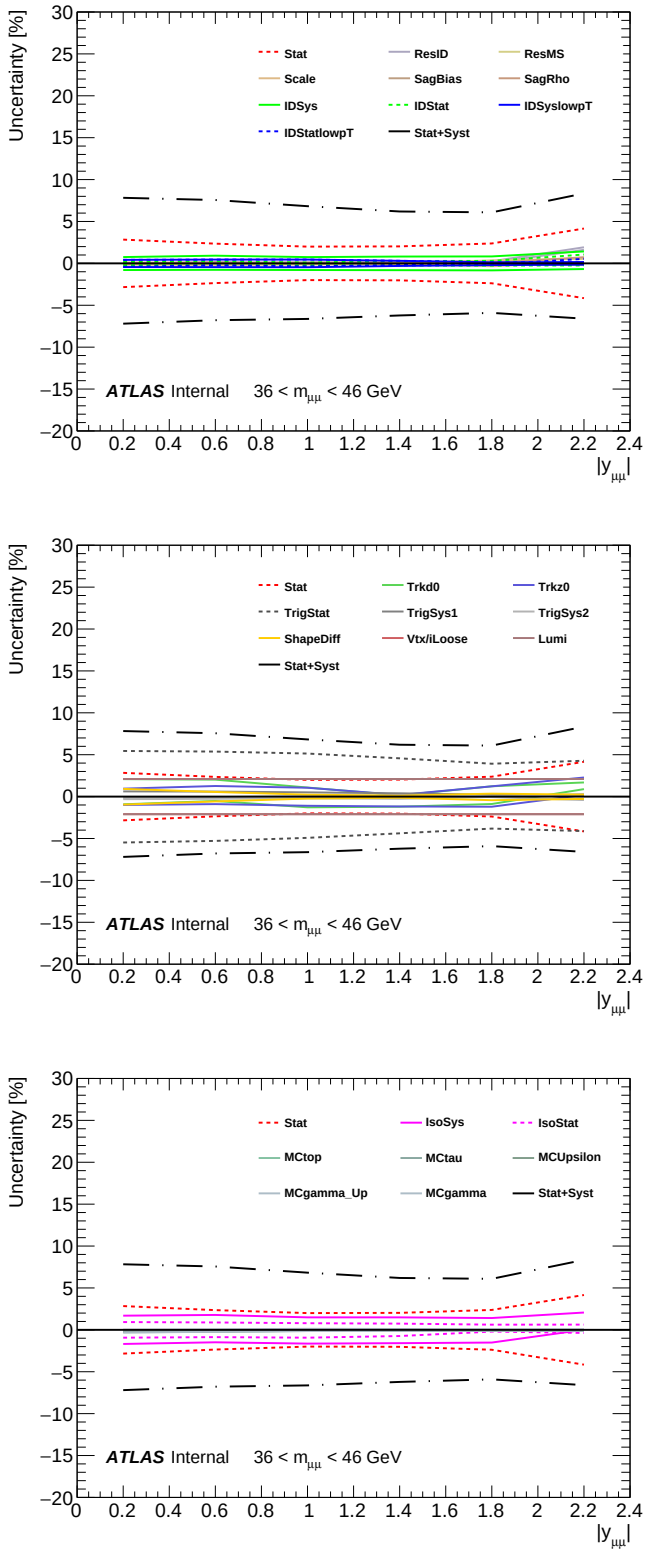


Fig. 6.43: Impact of the systematic uncertainties on the extracted cross section as a function of $|y_{\mu\mu}|$ for the $36 < m_{\mu\mu} < 46$ GeV bin. The black solid-dotted line represents the sum in quadrature of statistical and systematic uncertainties in this specific bin.

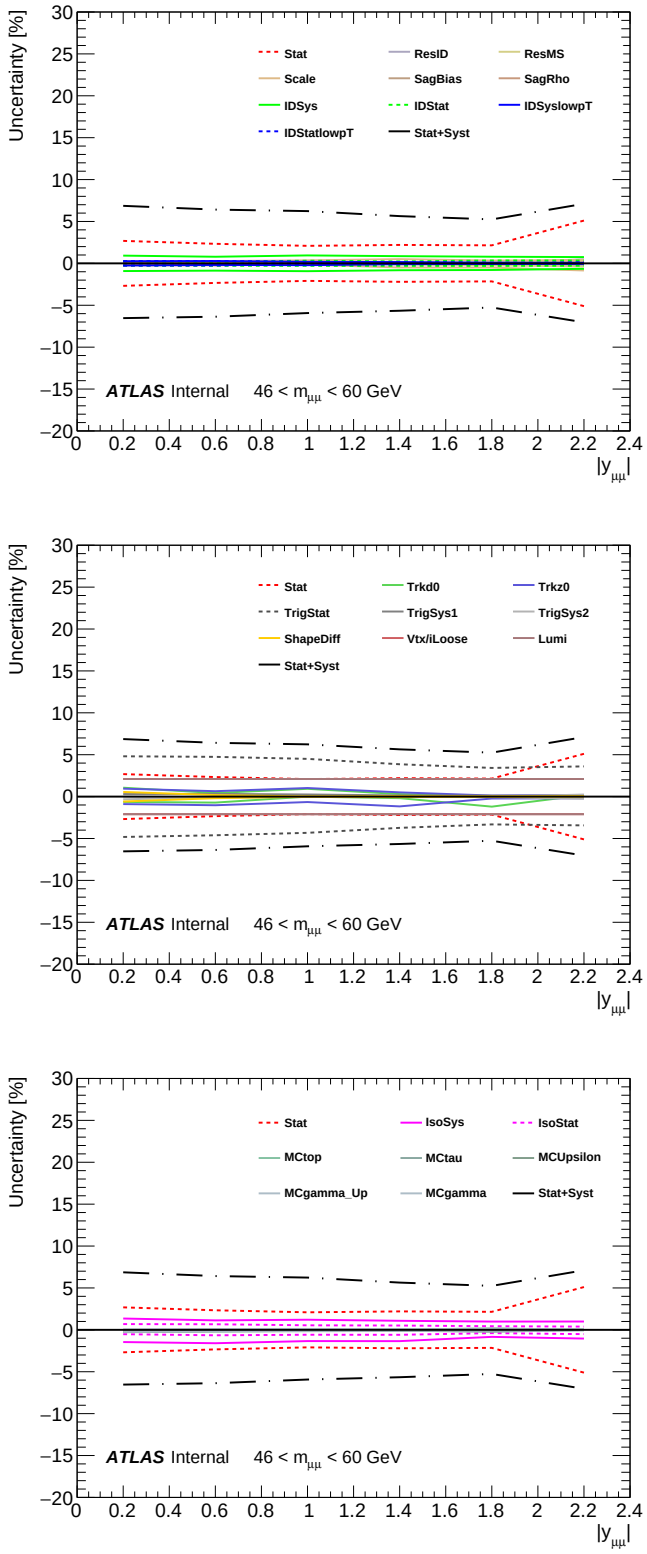


Fig. 6.44: Impact of the systematic uncertainties on the extracted cross section as a function of $|y_{\mu\mu}|$ for the $46 < m_{\mu\mu} < 60$ GeV bin. The black solid-dotted line represents the sum in quadrature of statistical and systematic uncertainties in this specific bin.

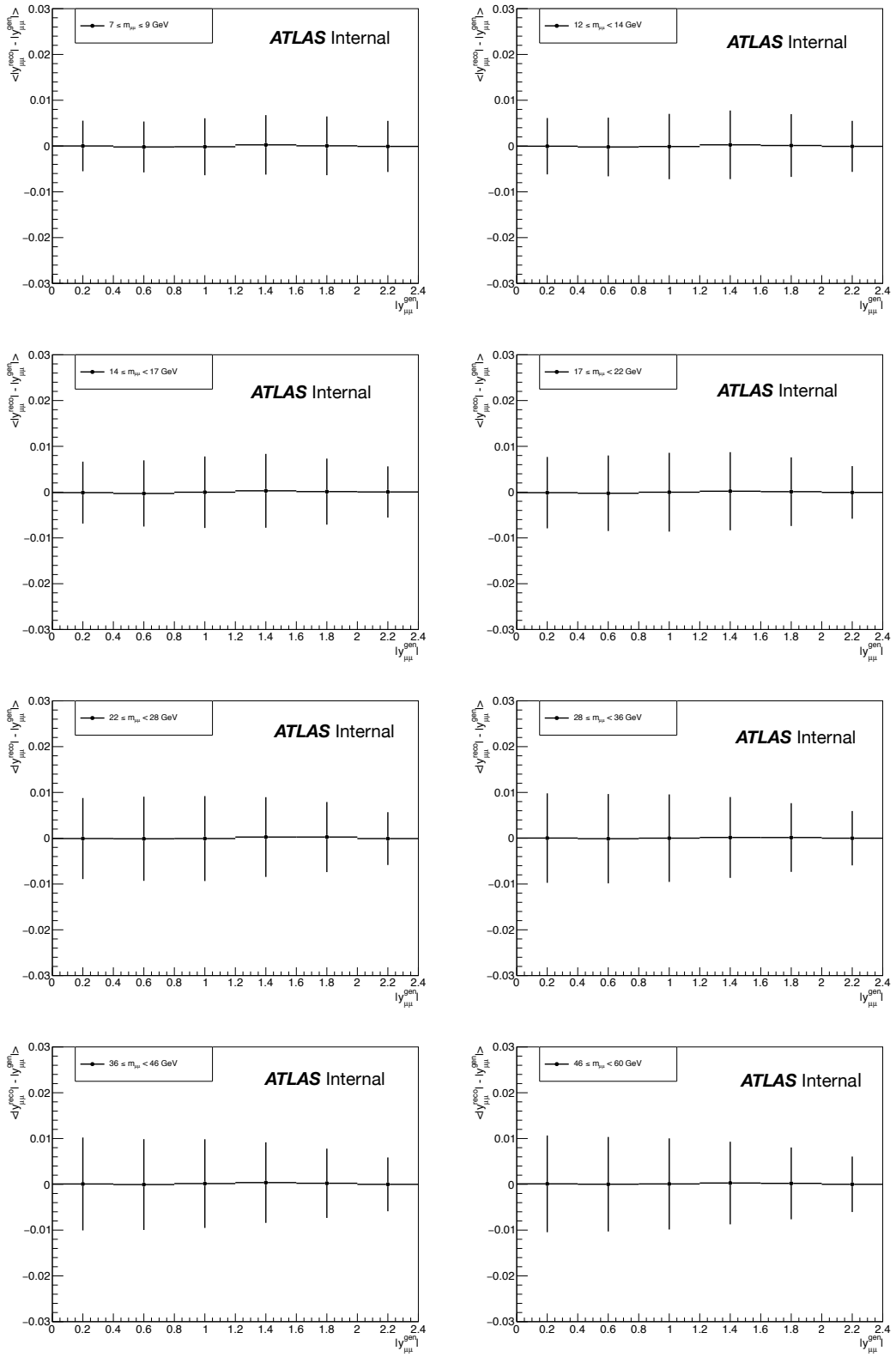


Fig. 6.45: Average ATLAS di-muon rapidity resolution as simulated by DY MC signal samples in the double differential binning scheme for mass and rapidity. The error bars represent the RMS of the entries within each bin.

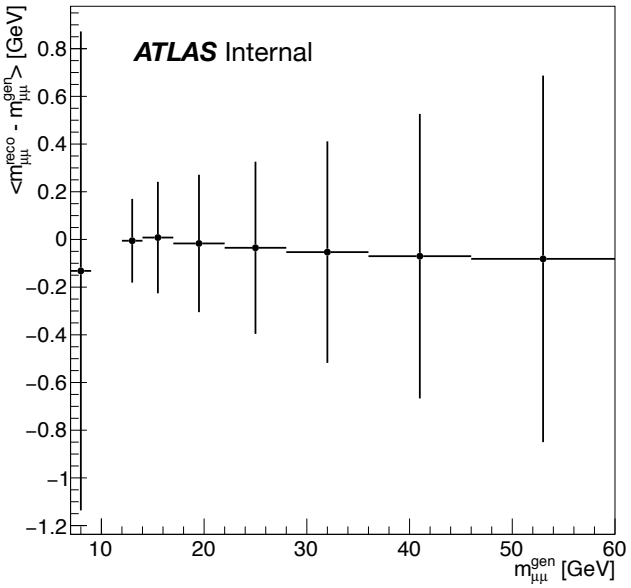


Fig. 6.46: Average ATLAS di-muon mass resolution as simulated by DY MC signal samples in the complete mass range $7 < m_{\mu\mu} < 60$ GeV.

Mass bin [GeV]	RMS (Reco - Born muons) [GeV]
7 - 9	1.004
12 - 14	0.176
14 - 17	0.234
17 - 22	0.288
22 - 28	0.361
28 - 36	0.465
36 - 46	0.597
46 - 60	0.769

Table 6.9: RMS of the invariant mass distributions for each invariant mass bin.

(RMS) of the entries within each bin. It can be seen from Fig. 6.45 that for the di-muon rapidity variables the size of RMS of the data is always at least 9 times smaller the bin width, resulting in the low bin migration seen in Fig. 6.48 (left plot). From Fig. 6.46, it can be seen that the di-muon invariant mass resolution is also much smaller than the bin size, apart from .
The mass resolution in each in invariant mass bin is listed in Table 6.9, where it is visible that the RMS does not exceed half of the bin size in every bin, apart from the 7-9 GeV one. Moreover, any bias on the reconstruction of the invariant mass is observed.

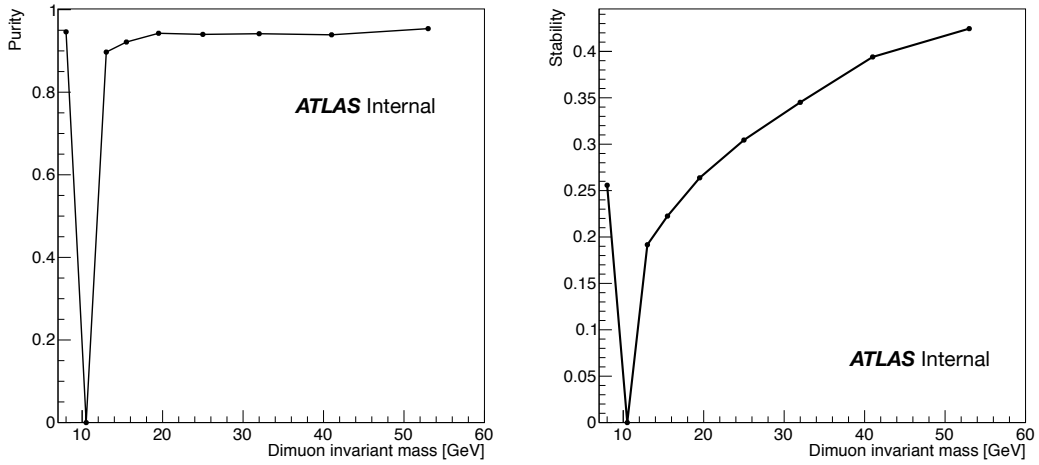


Fig. 6.47: The purity (left) and stability (right) in the 1D binning scheme using DY MC signal samples. The spike region represents the excluded 9 < $m_{\mu\mu}$ < 12 GeV region. These quantities are computed using muons at bare level[13].

6.5.3 Purity, Stability and Acceptance

In order to determine the cross section within the fiducial volume described in Sec. 6.5.1, a correction factor, C_{DY}, to take into account the differences between the reconstruction and fiducial level selections needs to be calculated. This correction accounts for any events that have migrated from their generated bins to other analysis bins. In addition, because the detector does not have a perfect detector efficiency some of the generated events will not be reconstructed at all and this needs to be accounted for as part of the selection efficiency. The C_{DY} factor is determined from signal MC only.
The fraction of events generated in a bin to those reconstructed in a bin is given by the *acceptance*, defined as follows:

$$C^i_{DY} = \frac{N^i_{rec}\ (\text{reco selection only})}{N^i_{gen}\ (\text{truth selection only})} \tag{6.14}$$

where *i* represents the bin index, N^i_{rec} and N^i_{gen} are the number of events reconstructed and generated in a given bin *i* respectively. The level of bin migration in the analysis can be assessed introducing the concepts of *purity* and *stability*. The purity describes how many events in each reconstructed bin were generated in the same bin:

$$P_i = \frac{\text{Events generated and reconstructed in a bin } i}{\text{Events reconstructed in a bin } i}, \tag{6.15}$$

while the fraction of events that have left the bin in which they were generated is described by the stability:

$$S_i = \frac{\text{Events generated and reconstructed in a bin } i}{\text{Events generated in a bin } i}. \tag{6.16}$$

[13]Particle-level kinematics after final state radiation (FSR).

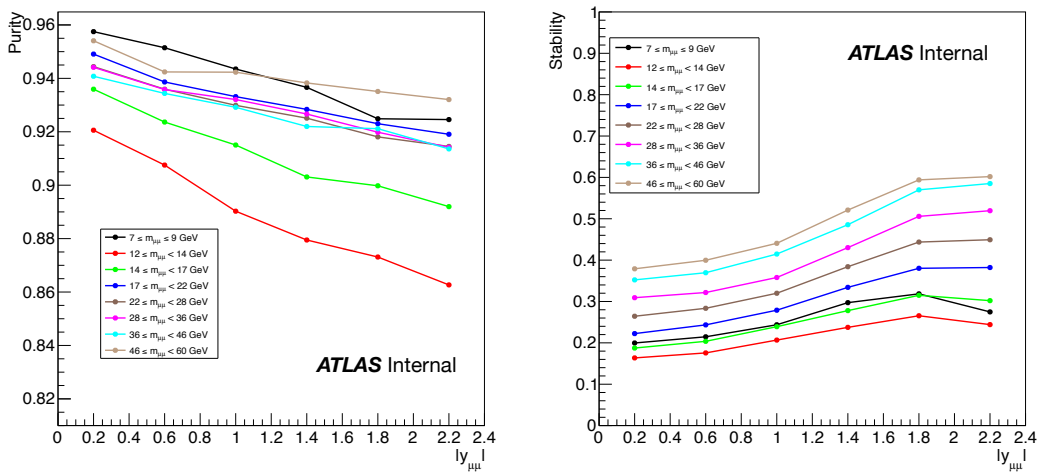


Fig. 6.48: The purity (left) and stability (right) in the 2D binning scheme using DY MC signal samples. These quantities are computed using muons at bare level.

Fig. 6.47 and Fig. 6.48 show the purity and the stability in the analysis binning schemes for the di-muon mass and di-muon rapidity respectively. From Eq. 6.15 and Eq. 6.16 it can be seen that the acceptance, defined in Eq. 6.14, can be computed from the stability and purity $C_{DY,i} = S_i / P_i$. The purity in di-muon mass is over 80% for the single-differential binning scheme. Purities higher than 90% are observed for all the of the rapidity binning, apart from the last one, the closest to the Z resonance radiative tail. All these plots are made without taking into account the prescale factor of 0.25 between the luminosities recorded by `2mu4` and `2mu6` triggers. The acceptances for both the one- and the two-dimensional distributions are shown in Fig. 6.49. Acceptances of 95% or higher have been observed in the single differential binning scheme, while the overall acceptance for the double differential distribution is between 20% and 60%. The acceptances gets larger when increasing the invariant mass and the rapidity.

6.5.4 Results

The measured fiducial double-differential cross sections $\frac{d^2\sigma}{dm_{\mu\mu}d|y_{\mu\mu}|}$ is reported in Fig. 6.50. These plots show the comparison between the lmDY cross-sections and the NLO predictions from `Powheg+Pythia`, described in Sec. 6.1. In Fig. 6.50 the extracted lmDY cross-sections is also compared to additional theoretical predictions generated with `Sherpa v2.2.1`. A more detailed description of this additional DY signal sample can be found in Ref. [2].
It can be seen that the measured cross-sections is broadly in agreement with the NLO+LL predictions from `Powheg+Pythia`, while a disagreement both in the normalisation and qualitatively in the shape of the cross section as a function of $|y_{\mu\mu}|$ is visible when comparing the lmDY data to the `Sherpa` predictions.
It needs to be taken into account that the systematic uncertainties may be overestimated, in particular the uncorrelated part between different rapidity bins. Indeed,

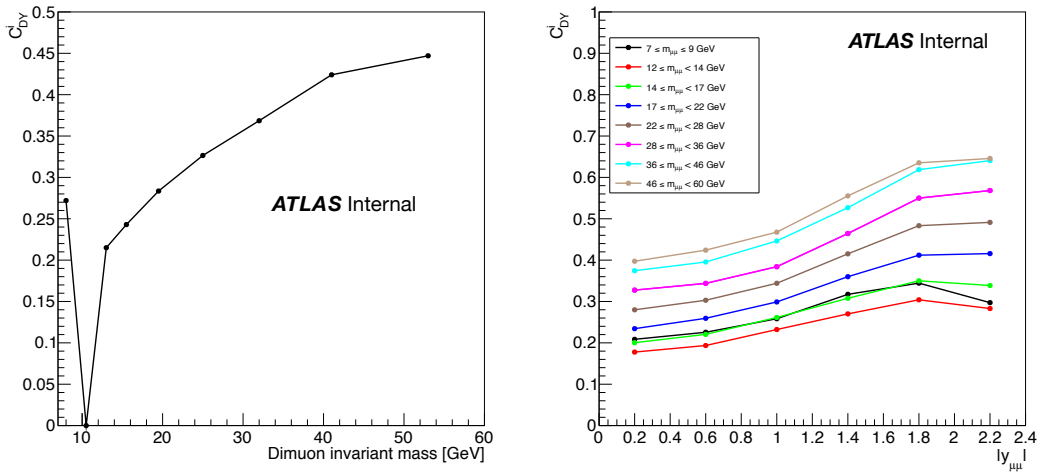


Fig. 6.49: Acceptances for the 1D binning scheme (left) and the 2D binning scheme (right) using DY MC signal samples. The spike region in the left plots represents the excluded $9 < m_{\mu\mu} < 12$ GeV region. These quantities are computed using muons at bare level.

it will be possible to reduce the uncertainty on $\delta^{\text{Iso}}_{\text{unc}}$ and $\delta^{\text{Trig}}_{\text{unc}}$ considering the different bin in the efficiency maps as uncorrelated.
The extracted lmDY cross-sections as a function of the invariant mass of the di-muon pair, namely $d\sigma/dm_{\mu\mu}$, is shown in Fig. 6.51. As already observed in Fig. 6.50, the measured cross-sections is broad in agreement with the `Powheg+Pythia` prediction, while a disagreement at 20%-30% level is visible when comparing the lmDY data to the `Sherpa` predictions. This difference could be explained by the fact that the low-mass DY cross section is extracted at Born level, but Sherpa does not have this concept and the predictions are to bare level[14].
As already discussed in Sec. 6.2.8 and as it can be seen from the plots shown in Fig. 6.23, both `Powheg+Pythia` and `Sherpa` describe the measured cross section (both single- and double-differential) poorly. This is probably due to the breakdown of the theory predictions, given that the considered phase space is very tiny and close to the cut imposed on the lepton p_T. In addition to that, this invariant mass range probes the very high-p_T tail and it could be that NLO MC predictions do not reproduce kinematic distributions so well, taking into account that the p_T of the di-muon system is much more greater than the invariant mass in this region.

6.6 Summary of Chapter 6

The ATLAS detector at the Large Hadron Collider has been used to measure the double-differential cross section of Drell-Yan production in the low-mass regime, namely between 7 and 60 GeV, with the range $9 < m_{\mu\mu} < 12$ GeV excluded. This chapter presented the results of the measurement conducted with 3.2 fb^{-1} of pp collision data collected at $\sqrt{s}$ = 13 TeV and recorded during 2015 data-taking period. The observed distributions in the di-muon final state are consistent with

[14]Particle-level kinematics after final state radiation (FSR).

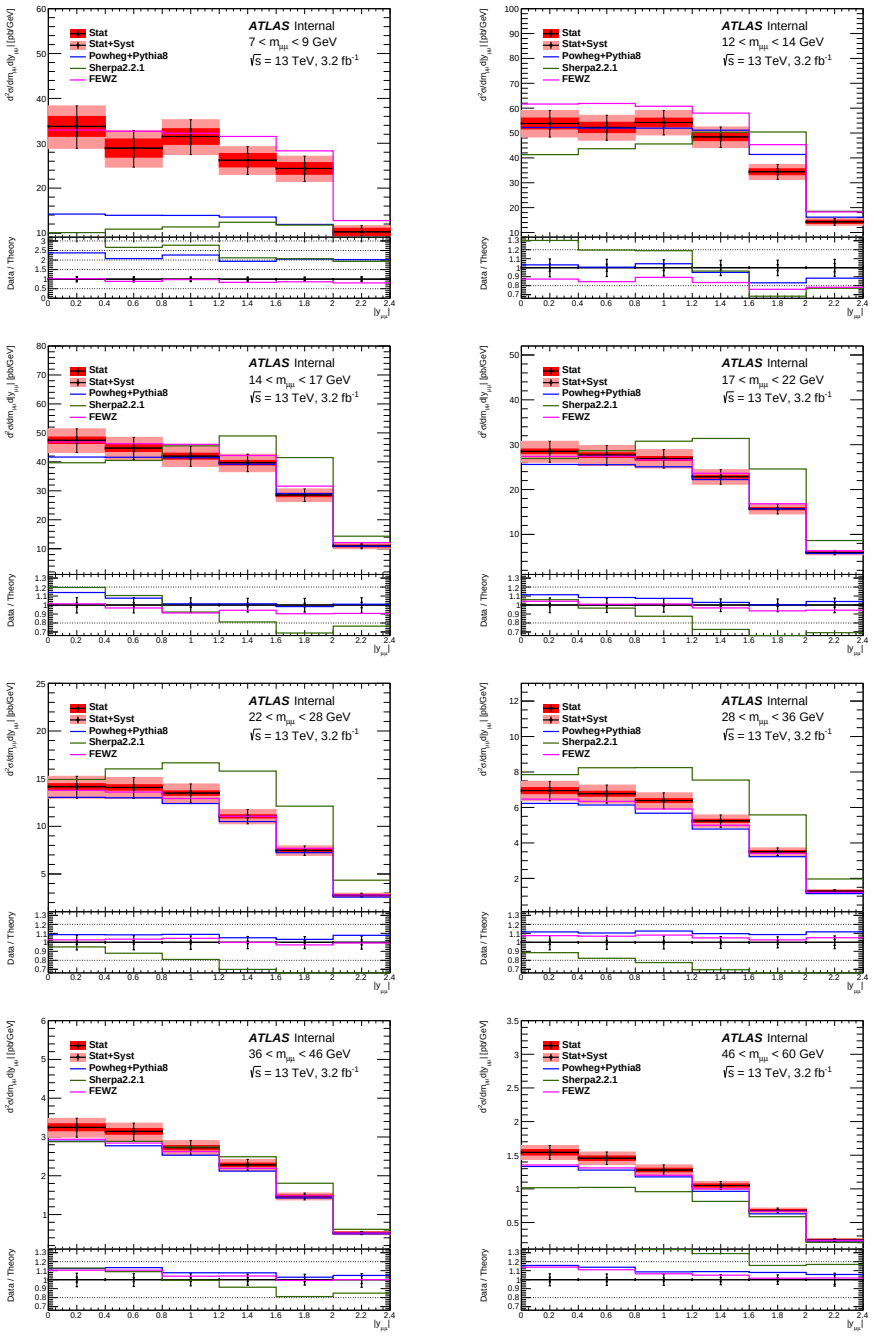


Fig. 6.50: Measured fiducial lmDY double-differential cross sections $d^2\sigma/dm_{\mu\mu}d|y_{\mu\mu}|$, defined in Eq. 6.13. The inner band represents the statistical uncertainty, while the outer band represents the sum in quadrature of statistical and systematic uncertainties, as described in Sec. 6.4.4. The solid blue line shows the theoretical predictions from `Powheg+Pythia`, while the predictions from `Sherpa` are plotted with a solid green line. The ratios of all two theoretical predictions (solid lines) to the data are shown in the lower panels. The data (solid points) are displayed at unity with the statistical (inner) and total (outer) measurement uncertainties.

the theoretical predictions provided by `Powheg+Pythia`, within systematic and statistical uncertainties. Further comparison with an additional MC signal sample

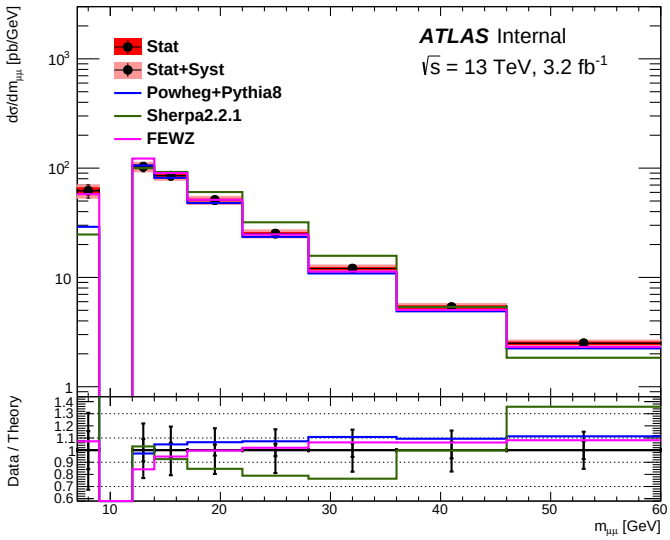


Fig. 6.51: Measured fiducial lmDY single-differential cross sections $d\sigma/dm_{\mu\mu}$, defined in Eq. 6.12. The inner band represents the statistical uncertainty, while the outer band represents the sum in quadrature of statistical and systematic uncertainties, as described in Sec. 6.4.4. The solid blue line shows the theoretical predictions from `Powheg+Pythia`, while the predictions from `Sherpa` are plotted with a solid green line. The ratios of all two theoretical predictions (solid lines) to the data are shown in the lower panels. The data (solid points) are displayed at unity with the statistical (inner) and total (outer) measurement uncertainties. The discrepancy observed in the 7-9 GeV bin is probably due to the breakdown of the theory predictions, given that the considered phase space is very tiny and close to the cut imposed on the lepton p_T. In addition to that, the invariant mass range covered in the first bin probes the very high-p_T tail and it could be that NLO MC predictions do not reproduce kinematic distributions so well, taking into account that the p_T of the di-muon system is much more greater than the invariant mass in this region.

generated with `Sherpa` have been conducted, and it shows a bigger disagreement between the measured data and the theoretical predictions.
These data will be possibly included in a full QCD fit, in order to test the need for $\ln(1/x)$ resummation in the low-x regime (as done in Ref. [142] for example). However, this will possible only once theory exists for matrix element as well as for splitting functions.

Chapter 7

Conclusion

In this thesis, three analyses covering different aspects of the Drell-Yan process have been presented. The search for a new heavy neutral gauge boson, a Z' boson, decaying in di-lepton final state is the first analysis described in this thesis. In this analysis, di-lepton events with invariant masses ranging up to 4 TeV were studied. The data in use were collected by the ATLAS experiment in the year 2015+2016, corresponding to an integrated luminosity of 36.1 fb^{-1}. No significant deviation from the Standard Model prediction is observed and several BSM theoretical models have been benchmarked. Upper exclusion limits at 95% confidence level are therefore set on the cross-section times branching ratio $\sigma \cdot B$ for a spin-1 Z' Gauge boson decaying into di-leptons. The resulting lower mass limits are 4.5 TeV for the Z'_{SSM}, 4.1 TeV for the Z'_{χ}, and 3.8 TeV for the Z'_{ψ}, along with other E_6 Z' models. Lower exclusion limits are also set on the energy scale Λ for various $qq\ell\ell$ contact interaction models, ranging between 23.5 TeV and 40.1 TeV.
In the second analysis, the double-differential cross section measurement of the process $pp \to Z/\gamma^* \to \ell^+\ell^-$ at a center of mass energy of $\sqrt{s}$ = 8 TeV is used in a PDF fit to put constraints on the photon PDF of the proton. These data allow a determination of $x\gamma(x, Q^2)$ with uncertainties at the 30% level for $0.02 \leq x \leq 0.1$. The resulting photon PDF is in agreement and exhibit smaller PDF uncertainties that the only other existing photon PDF fit from LHC data, the NNPDF3.0QED analysis, based on previous LHC Drell-Yan measurements. It is also in agreement within uncertainties with two recent theoretical calculations of the photon PDF, LUXqed and HKR16.
The last analysis presented in this thesis utilised $\sqrt{s}$ = 13 TeV data, collected in 2015 (3.2 fb^{-1}), to measure the double-differential Drell-Yan cross section $\frac{d^2\sigma}{dm_{\mu\mu}d|y_{\mu\mu}|}$ in the di-muon final state. The extracted fiducial cross sections are in broad agreement with predictions from `Powheg` interfaced with the `Pythia v8.1` parton shower program. The programs used the CT10 PDF and the AZNLO CTEQ6L1 tune for `Powheg+Pythia`. These measurements can potentially be used to test the need for $\ln(1/x)$ resummation once theory exists for matrix elements, as well as splitting functions. In this regard, some preliminary studies have been already conducted and they can be found in Ref. [200].

Bibliography

[1] ATLAS Collaboration, "Monte Carlo Generators for the Production of a W or Z/γ^* Boson in Association with Jets at ATLAS in Run 2." ATL-COM-PHYS-2015-832, Geneva, Aug 2015.

[2] ATLAS Collaboration, "ATLAS simulation of boson plus jets processes in Run 2." ATL-PHYS-PUB-2017-006, Geneva, May 2017.

[3] ATLAS Collaboration, "Simulation of top quark production for the ATLAS experiment at $\sqrt{s} = 13$ TeV." ATL-PHYS-PUB-2016-004, Geneva, Jan 2016.

[4] ATLAS Collaboration, "Studies on top-quark Monte Carlo modelling for Top2016." ATL-COM-PHYS-2016-1235, Geneva, Aug 2016.

[5] ATLAS Collaboration, "Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMC@NLO." ATL-COM-PHYS-2017-391, Geneva, Apr 2017.

[6] ATLAS Collaboration, "Determination of the parton distribution functions of the proton from ATLAS measurements of differential W and Z/γ^* and $t\bar{t}$ cross sections." ATL-COM-PHYS-2018-388, Geneva, Apr 2018.

[7] A. Djouadi, "The Anatomy of Electroweak Symmetry Breaking - Tome I: The Higgs Boson in the Standard Model," *Physics Reports*, vol. 457, no. 1, 2008.

[8] S. L. Glashow, "Partial symmetries of weak interactions," *Nuclear Physics*, vol. 22, no. 4, pp. 579–588, 1961.

[9] A. Salam and J. C. Ward, "Electromagnetic and weak interactions," *Physics Letters*, vol. 13, no. 2, pp. 168–171, 1964.

[10] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, "Advantages of the color octet gluon picture," *Physics Letters B*, vol. 47, no. 4, pp. 365–368, 1973.

[11] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," *Physical Review Letters*, vol. 13, no. 9, pp. 321–323, 1964.

[12] P. W. Higgs, "Broken symmetries, massless particles and gauge fields," *Physics Letters*, vol. 12, no. 2, pp. 132–133, 1964.

[13] P. D. Group, "Review of Particle Physics (RPP)," *Phys.Rev.*, vol. D86, p. 010001, 2012.

[14] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," *Physics Review Letters*, vol. 13, pp. 108–508, 1964.

[15] H. Yukawa, "Interaction of elementary particles - Part I," *Proc. Phys. Math. Soc. Jpn.*, vol. 17, pp. 48–57, 1935.

[16] D. J. Gross and F. Wilczek, "Ultraviolet Behavior of Nonabelian Gauge Theories," *Phys. Rev. Lett.*, vol. 30, pp. 1343–1346, 1973.

[17] R. K. Ellis, W. J. Stirling, and B. R. Webber, "QCD and collider physics," *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, vol. 8, pp. 1–435, 1996.

[18] S. Chekanov *et al.*, "Forward jet production in deep inelastic ep scattering and low-x parton dynamics at HERA," *Phys. Lett.*, vol. B632, pp. 13–26, 2006.

[19] S. Chekanov *et al.*, "Forward-jet production in deep inelastic ep scattering at HERA," *Eur. Phys. J.*, vol. C52, pp. 515–530, 2007.

[20] S. Chekanov *et al.*, "Inclusive-jet and dijet cross-sections in deep inelastic scattering at HERA," *Nucl. Phys.*, vol. B765, pp. 1–30, 2007.

[21] V. Andreev *et al.*, "Measurement of multijet production in ep collisions at high Q^2 and determination of the strong coupling α_s," *Eur. Phys. J.*, vol. C75, no. 2, p. 65, 2015.

[22] V. Andreev *et al.*, "Measurement of Jet Production Cross Sections in Deep-inelastic ep Scattering at HERA," *Eur. Phys. J.*, vol. C77, no. 4, p. 215, 2017.

[23] F. D. Aaron *et al.*, "Measurement of the Azimuthal Correlation between the most Forward Jet and the Scattered Positron in Deep-Inelastic Scattering at HERA," *Eur. Phys. J.*, vol. C72, p. 1910, 2012.

[24] F. D. Aaron *et al.*, "Combined Measurement and QCD Analysis of the Inclusive $e^{\pm}p$ Scattering Cross Sections at HERA," *JHEP*, vol. 1001, p. 109, 2010.

[25] G. Bozzi, S. Catani, G. Ferrera, D. de Florian, and M. Grazzini, "Transverse-momentum resummation: A Perturbative study of Z production at the Tevatron," *Nucl. Phys.*, vol. B815, pp. 174–197, 2009.

[26] A. V. Lipatov, M. A. Malyshev, and N. P. Zotov, "Drell-Yan lepton pair production at high energies in the k_t-factorization approach," *JHEP*, vol. 12, p. 117, 2011.

[27] V. D. Elvira, "QCD, Tevatron Results and LHC Prospects," in *Hadron collider physics. Proceedings, 19th Symposium, HCP2008, Galena, USA, May 27-31, 2008*, 2008.

[28] R. P. Feynman, "Very high-energy collisions of hadrons," *Phys. Rev. Lett.*, vol. 23, pp. 1415–1417, 1969.

[29] J. C. Collins, D. E. Soper, and G. F. Sterman, "Factorization of Hard Processes in QCD," *Adv. Ser. Direct. High Energy Phys.*, vol. 5, pp. 1–91, 1989.

[30] J. M. Campbell, J. W. Huston, and W. J. Stirling, "Hard Interactions of Quarks and Gluons: A Primer for LHC Physics," *Rept. Prog. Phys.*, vol. 70, p. 89, 2007.

[31] Y. L. Dokshitzer, "Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.," *Sov. Phys. JETP*, vol. 46, pp. 641–653, 1977. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[32] V. N. Gribov and L. N. Lipatov, "Deep inelastic e p scattering in perturbation theory," *Sov. J. Nucl. Phys.*, vol. 15, pp. 438–450, 1972. [Yad. Fiz.15,781(1972)].

[33] G. Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language," *Nucl. Phys.*, vol. B126, pp. 298–318, 1977.

[34] H. Abramowicz *et al.*, "Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data," *Eur. Phys. J.*, vol. C75, no. 12, p. 580, 2015.

[35] S. Forte, "Parton distributions at the dawn of the LHC," *Acta Phys.Polon.*, vol. B41, p. 2859, 2010.

[36] J. Pumplin *et al.*, "Uncertainties of predictions from parton distribution functions. 2. The Hessian method," *Phys. Rev.*, vol. D65, p. 014013, 2001.

[37] J. Pumplin *et al.*, "New generation of parton distributions with uncertainties from global qcd analysis," *JHEP*, vol. 07, p. 012, 2002.

[38] J. Pumplin, D. R. Stump, and W. K. Tung, "Multivariate fitting and the error matrix in global analysis of data," *Phys. Rev.*, vol. D65, p. 014011, 2001.

[39] S. D. Drell and T.-M. Yan, "Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies," *Phys. Rev. Lett.*, vol. 25, pp. 316–320, 1970. [Erratum: Phys. Rev. Lett.25,902(1970)].

[40] R. D. Field, "Applications of Quantum Chromodynamics," *AIP Conf. Proc.*, vol. 55, pp. 97–242, 1979.

[41] S. Catani and M. Grazzini, "An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC," *Phys. Rev. Lett.*, vol. 98, p. 222002, 2007.

[42] S. Catani, L. Cieri, G. Ferrera, D. de Florian, and M. Grazzini, "Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO," *Phys. Rev. Lett.*, vol. 103, p. 082001, 2009.

[43] S. Catani, L. Cieri, D. de Florian, G. Ferrera, and M. Grazzini, "Vector boson production at hadron colliders: hard-collinear coefficients at the NNLO," *Eur. Phys. J.*, vol. C72, p. 2195, 2012.

[44] S. Catani, D. de Florian, G. Ferrera, and M. Grazzini, "Vector boson production at hadron colliders: transverse-momentum resummation and leptonic decay," *JHEP*, vol. 12, p. 047, 2015.

[45] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO vector-boson production matched with shower in POWHEG," *JHEP*, vol. 07, p. 060, 2008.

[46] R. Frederix, M. K. Mandal, P. Mathews, V. Ravindran, and S. Seth, "Drell-Yan, ZZ, W^+W^- production in SM & ADD model to NLO+PS accuracy at the LHC," *Eur. Phys. J.*, vol. C74, no. 2, p. 2745, 2014.

[47] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," *JHEP*, vol. 06, p. 029, 2002.

[48] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "Parton distributions incorporating QED contributions," *Eur. Phys. J.*, vol. C39, p. 155, 2005.

[49] V. Bertone and S. Carrazza, "Combining NNPDF3.0 and NNPDF2.3QED through the APFEL evolution code," 2016.

[50] R. D. Ball *et al.*, "Parton distributions for the LHC Run II," *JHEP*, vol. 04, p. 040, 2015.

[51] C. Schmidt, J. Pumplin, D. Stump, and C. P. Yuan, "CT14QED parton distribution functions from isolated photon production in deep inelastic scattering," *Phys. Rev.*, vol. D93, no. 11, p. 114015, 2016.

[52] A. Manohar, P. Nason, G. P. Salam, and G. Zanderighi, "How bright is the proton? A precise determination of the photon parton distribution function," *Phys. Rev. Lett.*, vol. 117, no. 24, p. 242002, 2016.

[53] A. V. Manohar, P. Nason, G. P. Salam, and G. Zanderighi, "The Photon Content of the Proton," *JHEP*, vol. 12, p. 046, 2017.

[54] L. A. Harland-Lang, V. A. Khoze, and M. G. Ryskin, "Photon-initiated processes at high mass," *Phys. Rev.*, vol. D94, no. 7, p. 074008, 2016.

[55] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Physics Letters B*, vol. 716, no. 1, pp. 1–29, 2012.

[56] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Physics Letters B*, vol. 716, no. 1, pp. 30–61, 2012.

[57] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Kennedy, K. Moenig, M. Schott, and J. Stelzer, "Updated status of the global electroweak fit and constraints on new physics," *The European Physical Journal C-Particles and Fields*, vol. 72, no. 5, pp. 1–35, 2012.

[58] J. Ellis, S. Kelley, and D. V. Nanopoulos, "Constraints from gauge coupling unification on the scale of supersymmetry breaking," *Physics Letters B*, vol. 287, no. 1, pp. 95–100, 1992.

[59] J. A. Casas, J. R. Espinosa, and I. Hidalgo, "Implications for new physics from fine-tuning arguments. 1) Application to SUSY and seesaw cases," *Journal of High Energy Physics*, vol. 2004, no. 11, p. 057, 2004.

[60] P. Langacker, "The Physics of Heavy Z' Gauge Bosons," *Rev. Mod. Phys.*, vol. 81, pp. 1199–1228, 2009.

[61] D. London and J. L. Rosner, "Extra Gauge Bosons in E(6)," in *Proceedings, 23RD International Conference on High Energy Physics, JULY 16-23, 1986, Berkeley, CA*, 1986.

[62] M. Dittmar, A.-S. Nicollerat, and A. Djouadi, "Z-prime studies at the LHC: An Update," *Phys. Lett.*, vol. B583, pp. 111–120, 2004.

[63] E. Salvioni, G. Villadoro, and F. Zwirner, "Minimal Z-prime models: Present bounds and early LHC reach," *JHEP*, vol. 11, p. 068, 2009.

[64] G. Senjanovic and R. N. Mohapatra, "Exact Left-Right Symmetry and Spontaneous Violation of Parity," *Phys. Rev.*, vol. D12, p. 1502, 1975.

[65] R. N. Mohapatra and J. C. Pati, "Left-Right Gauge Symmetry and an Isoconjugate Model of CP Violation," *Phys. Rev.*, vol. D11, pp. 566–571, 1975.

[66] L. Basso, A. Belyaev, S. Moretti, and C. H. Shepherd-Themistocleous, "Phenomenology of the minimal B-L extension of the Standard model: Z' and neutrinos," *Phys. Rev.*, vol. D80, p. 055030, 2009.

[67] V. M. Abazov *et al.*, "Search for a heavy neutral gauge boson in the dielectron channel with 5.4 fb-1 of $p\bar{p}$ collisions at sqrt(s) = 1.96 TeV," *Phys. Lett.*, vol. B695, pp. 88–94, 2011.

[68] T. Aaltonen *et al.*, "Search for High Mass Resonances Decaying to Muon Pairs in $\sqrt{s} = 1.96$ TeV $p\bar{p}$ Collisions," *Phys. Rev. Lett.*, vol. 106, p. 121801, 2011.

[69] G. Abbiendi *et al.*, "Tests of the standard model and constraints on new physics from measurements of fermion pair production at 189-GeV to 209-GeV at LEP," *Eur. Phys. J.*, vol. C33, pp. 173–212, 2004.

[70] J. Abdallah *et al.*, "Measurement and interpretation of fermion-pair production at LEP energies above the Z resonance," *Eur. Phys. J.*, vol. C45, pp. 589–632, 2006.

[71] P. Achard *et al.*, "Measurement of hadron and lepton-pair production in e^+e^- collisions at $\sqrt{s}$ = 192 GeV to 208 GeV at LEP," *Eur. Phys. J.*, vol. C47, pp. 1–19, 2006.

[72] S. Schael *et al.*, "Fermion pair production in e^+e^- collisions at 189-209 GeV and constraints on physics beyond the standard model," *Eur. Phys. J.*, vol. C49, pp. 411–437, 2007.

[73] T. G. Rizzo, "$Z^{'}$ Physics at the LHC and the LHeC," in *Proceedings, 34th International Conference on High Energy Physics (ICHEP 2008): Philadelphia, Pennsylvania, July 30-August 5, 2008*, 2008.

[74] ATLAS Collaboration, "Search for high-mass dilepton resonances in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector," *Phys. Rev.*, vol. D90, no. 5, p. 052005, 2014.

[75] V. Khachatryan *et al.*, "Search for physics beyond the standard model in dilepton mass spectra in proton-proton collisions at $\sqrt{s} = 8$ TeV," *JHEP*, vol. 04, p. 025, 2015.

[76] ATLAS Collaboration, "Search for high-mass new phenomena in the dilepton final state using proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *Phys. Lett.*, vol. B761, pp. 372–392, 2016.

[77] E. Fermi, "An attempt of a theory of beta radiation. 1.," *Z. Phys.*, vol. 88, pp. 161–177, 1934.

[78] E. Eichten, K. D. Lane, and M. E. Peskin, "New Tests for Quark and Lepton Substructure," in *Beyond The Standard Model. Proceedings, 18th Recontres de Moriond, La Plagne, France, March 13-19, 1983. Vol. 2*, vol. 50, pp. 811–814, 1983. [,369(1983)].

[79] E. Eichten, I. Hinchliffe, K. D. Lane, and C. Quigg, "Super Collider Physics," *Rev. Mod. Phys.*, vol. 56, pp. 579–707, 1984. [Addendum: Rev. Mod. Phys.58,1065(1986)].

[80] P. L. Anthony *et al.*, "Precision measurement of the weak mixing angle in Moller scattering," *Phys. Rev. Lett.*, vol. 95, p. 081601, 2005.

[81] J. Abdallah *et al.*, "A Study of $b\bar{b}$ Production in e^+e^- Collisions at $\sqrt{s}$ = 130 GeV - 207 GeV," *Eur. Phys. J.*, vol. C60, pp. 1–15, 2009.

[82] F. D. Aaron *et al.*, "Search for Contact Interactions in $e^{\pm}p$ Collisions at HERA," *Phys. Lett.*, vol. B705, pp. 52–58, 2011.

[83] V. M. Abazov *et al.*, "Measurement of dijet angular distributions at $\sqrt{s}$ = 1.96 TeV and searches for quark compositeness and extra spatial dimensions," *Phys. Rev. Lett.*, vol. 103, p. 191803, 2009.

[84] A. Abulencia *et al.*, "Search for $Z' \to e^+e^-$ using dielectron mass and angular distribution.," *Phys. Rev. Lett.*, vol. 96, p. 211801, 2006.

[85] ATLAS Collaboration, "Search for contact interactions and large extra dimensions in the dilepton channel using proton-proton collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector," *Eur. Phys. J.*, vol. C74, no. 12, p. 3134, 2014.

[86] ATLAS Collaboration, "Search for new high-mass phenomena in the dilepton final state using 36 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector," *JHEP*, vol. 10, p. 182, 2017.

[87] ATLAS Collaboration, "The ATLAS experiment at the CERN Large Hadron Collider," *Journal of Instrumentation*, vol. 3, no. 08, p. S08003, 2008.

[88] CMS Collaboration, "The CMS experiment at the CERN LHC," *Journal of Instrumentation*, vol. 3, no. 08, p. S08004, 2008.

[89] LHCb Collaboration, "The LHCb experiment at the CERN LHC," *Journal of Instrumentation*, vol. 3, no. 08, p. S08005, 2008.

[90] ALICE Collaboration, "The ALICE experiment at the CERN LHC," *Journal of Instrumentation*, vol. 3, no. 08, p. S08002, 2008.

[91] T. Plehn, "Lectures on LHC Physics," *Lect. Notes Phys.*, vol. 844, pp. 1–193, 2012.

[92] ATLAS Collaboration, "Luminosity Public Results," *https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults*.

[93] ATLAS Collaboration, "Luminosity Public Results," *https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2*.

[94] ATLAS Collaboration, "Updates of the ATLAS Tracking Event Data Model (Release 13)." ATL-SOFT-PUB-2007-003, ATL-COM-SOFT-2007-008, 2007.

[95] H. Herman, "The superconducting magnet system for the ATLAS detector at CERN," *Applied Superconductivity, IEEE Transactions on*, vol. 10, no. 1, pp. 347–352, 2000.

[96] A. Yamamoto, Y. Makida, R. Ruber, Y. Doi, T. Haruyama, F. Haug, H. Ten Kate, M. Kawai, T. Kondo, Y. Kondo, *et al.*, "The ATLAS Central Solenoid," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 584, no. 1, pp. 53–74, 2008.

[97] ATLAS Collaboration, "ATLAS Barrel Toroid: Technical design report." CERN-LHCC-97-19.

[98] ATLAS Collaboration, "ATLAS Endcap Toroids: Technical design report." CERN-LHCC-97-20, 1997.

[99] ATLAS Collaboration, "ATLAS Inner Detector: Technical design report. Vol. 1." CERN-LHCC-97-16, ATLAS-TDR-4, 1997.

[100] ATLAS Collaboration, "ATLAS Inner Detector: Technical design report. Vol. 2." CERN-LHCC-97-17, 1997.

[101] ATLAS Collaboration, "ATLAS Pixel Detector: Technical design report." CERN-LHCC-98-13, 1998.

[102] ATLAS Collaboration, "ATLAS Insertable B-Layer Technical Design Report." CERN-LHCC-2010-013, ATLAS-TDR-19, 2010.

[103] S. Schramm, "The ATLAS Detector: Status and Performance in Run-II," *Int. J. Mod. Phys. Conf. Ser.*, vol. 43, p. 1660195, 2016.

[104] ATLAS Collaboration, "Calibration of the ATLAS Transition Radiation Tracker." ATLAS-CONF-2011-006, 2011.

[105] ATLAS Collaboration, "ATLAS Liquid Argon Calorimeter: Technical design report." CERN-LHCC-96-41, 1996.

[106] ATLAS Collaboration, "ATLAS Tile Calorimeter: Technical design report." CERN-LHCC-96-42, 1996.

[107] ATLAS Collaboration, "Drift time measurement in the ATLAS Liquid Argon Electromagnetic Calorimeter using cosmic muons," *The European Physical Journal C*, vol. 70, no. 3, pp. 755–785, 2010.

[108] ATLAS Collaboration, "ATLAS Calorimeter Performance Technical Design Report." CERN-LHCC-96-40, 1996.

[109] ATLAS Collaboration, "ATLAS Muon Spectrometer: Technical design report." CERN-LHCC-97-22, ATLAS-TDR-10, 1997.

[110] ATLAS Collaboration, "ATLAS Trigger Performance: Status report." CERN-LHCC-98-15, 1998.

[111] ATLAS Collaboration, "ATLAS high-level trigger, data-acquisition and controls: Technical Design Report." CERN-LHCC-2003-022, 2003.

[112] G. Pásztor, "The Upgrade of the ATLAS Electron and Photon Triggers towards LHC Run 2 and their Performance," in *Proceedings, Meeting of the APS Division of Particles and Fields (DPF 2015): Ann Arbor, Michigan, USA, 4-8 Aug 2015*, 2015.

[113] A. R. Martínez, "The Run-2 ATLAS Trigger System," in *Proceedings, 17th International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT 2016): Valparaiso, Chile, January 18-22, 2016*, vol. 762, p. 012003, 2016.

[114] J. Glatzer, "Operation of the Upgraded ATLAS Level-1 Central Trigger System," in *Proceedings, 21st International Conference on Computing in High Energy and Nuclear Physics (CHEP 2015): Okinawa, Japan, April 13-17, 2015*, vol. 664, p. 082013, 2015.

[115] E. Simioni *et al.*, "Upgrade of the ATLAS Level-1 Trigger with event topology information," in *Proceedings, 21st International Conference on Computing in High Energy and Nuclear Physics (CHEP 2015): Okinawa, Japan, April 13-17, 2015*, vol. 664, p. 082052, 2015.

[116] M. Shochet, L. Tompkins, V. Cavaliere, P. Giannetti, A. Annovi, and G. Volpi, "Fast TracKer (FTK) Technical Design Report." CERN-LHCC-2013-007, ATLAS-TDR-021, 2013.

[117] ATLAS Collaboration, "Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System." CERN-LHCC-2013-018, ATLAS-TDR-023, 2013.

[118] M. Lamanna, "The LHC computing grid project at CERN," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 534, no. 1, pp. 1–6, 2004.

[119] ATLAS Collaboration, "ATLAS computing: Technical design report." CERN-LHCC-2005-022, ATLAS-TRD-017, 2005.

[120] ATLAS Collaboration, "ATHENA: the ATLAS Common Framework," *Developer Guide version 8. CERN, Geneva*, 2004.

[121] R. Brun and F. Rademakers, "ROOT - an object oriented data analysis framework," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 389, no. 1, pp. 81–86, 1997.

[122] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, and E. Moyse, "Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)," 2007.

[123] T. Cornelissen, M. Elsing, I. Gavrilenko, W. Liebig, E. Moyse, and A. Salzburger, "The new ATLAS track reconstruction (NEWT)," in *Proceedings, 16th International Conference on Computing in High Energy and Nuclear Physics (CHEP 2007): Victoria, Canada, September 2-7, 2007*, vol. 119, p. 032014, 2008.

[124] Kalman, Rudolph Emil, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[125] ATLAS Collaboration, "Early Inner Detector Tracking Performance in the 2015 data at $\sqrt{s}$ = 13 TeV." ATL-PHYS-PUB-2015-051, Geneva, Dec 2015.

[126] G. Piacquadio, K. Prokofiev, and A. Wildauer, "Primary vertex reconstruction in the ATLAS experiment at LHC," in *Proceedings, 16th International Conference on Computing in High Energy and Nuclear Physics (CHEP 2007): Victoria, Canada, September 2-7, 2007*, vol. 119, p. 032033, 2008.

[127] A. Collaboration, "Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1," *Eur. Phys. J.*, vol. C77, p. 490, 2017.

[128] ATLAS Collaboration, "Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s}$ =13 TeV," *Eur. Phys. J.*, vol. C76, no. 5, p. 292, 2016.

[129] ATLAS Collaboration, "Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," *Eur. Phys. J. C*, vol. 75, p. 17, 2015.

[130] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_T jet clustering algorithm," *Journal of High Energy Physics*, vol. 2008, no. 04, p. 063, 2008.

[131] ATLAS Collaboration, "Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data." ATLAS-CONF-2016-024, 2016.

[132] ATLAS Collaboration, "Calorimeter clustering algorithms: description and performance," *ATLAS Note*, 2008.

[133] ATLAS Collaboration, "Expected Performance of the ATLAS Experiment: Detector, Trigger and Physics," *arXiv preprint arXiv:0901.0512*, 2008.

[134] R. Frühwirth, "A Gaussian-mixture approximation of the Bethe-Heitler model of electron energy loss by bremsstrahlung," *Computer physics communications*, vol. 154, no. 2, pp. 131–142, 2003.

[135] ATLAS Collaboration, "Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data," *The European Physical Journal C*, vol. 72, no. 3, pp. 1–46, 2012.

[136] ATLAS Collaboration, "Electron and photon energy calibration with the ATLAS detector using data collected in 2015 at $\sqrt{s}$ = 13 TeV." ATL-PHYS-PUB-2016-015, Geneva, Aug 2016.

[137] ATLAS Collaboration, "Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data," *Eur. Phys. J. C*, vol. 74, p. 3071, 2014.

[138] ATLAS Collaboration, "Electron and photon energy calibration with the ATLAS detector using 2015-2016 LHC proton-proton collision data." ATL-COM-PHYS-2018-062, Geneva, Jan 2018.

[139] S. Alekhin *et al.*, "HERAFitter," *Eur. Phys. J.*, vol. C75, no. 7, p. 304, 2015.

[140] F. Giuli *et al.*, "The photon PDF from high-mass Drell-Yan data at the LHC," *Eur. Phys. J.*, vol. C77, no. 6, p. 400, 2017.

[141] xFitter Developers' Team, "Impact of the heavy quark matching scales in PDF fits," *Eur. Phys. J.*, vol. C77, no. 12, p. 837, 2017.

[142] xFitter Developers' Team, "Impact of low-x resummation on QCD analysis of HERA data," *Eur. Phys. J.*, vol. C78, no. 8, p. 621, 2018.

[143] 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*, vol. 08, p. 009, 2016.

[144] V. Bertone, S. Carrazza, and J. Rojo, "APFEL: A PDF Evolution Library with QED corrections," *Comput.Phys.Commun.*, vol. 185, p. 1647, 2014.

[145] V. Bertone, R. Frederix, S. Frixione, J. Rojo, and M. Sutton, "aMCfast: automation of fast NLO computations for PDF fits," *JHEP*, vol. 1408, p. 166, 2014.

[146] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *JHEP*, vol. 07, p. 079, 2014.

[147] S. Chekanov *et al.*, "Measurement of prompt photons with associated jets in photoproduction at HERA," *Eur. Phys. J.*, vol. C49, pp. 511–522, 2007.

[148] R. D. Ball *et al.*, "Parton distributions with LHC data," *Nucl.Phys.*, vol. B867, p. 244, 2013.

[149] R. D. Ball *et al.*, "Parton distributions with QED corrections," *Nucl.Phys.*, vol. B877, pp. 290–320, 2013.

[150] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, *et al.*, "Reweighting and Unweighting of Parton Distributions and the LHC W lepton asymmetry data," *Nucl.Phys.*, vol. B855, pp. 608–638, 2012.

[151] R. D. Ball *et al.*, "Reweighting NNPDFs: the W lepton asymmetry," *Nucl. Phys.*, vol. B849, pp. 112–143, 2011.

[152] R. Sadykov, "Impact of QED radiative corrections on Parton Distribution Functions," 2014.

[153] M. Botje, "QCDNUM: Fast QCD Evolution and Convolution," *Comput.Phys.Commun.*, vol. 182, pp. 490–532, 2011.

[154] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "W Physics at the LHC with FEWZ 2.1," *Comput.Phys.Commun.*, vol. 184, pp. 208–214, 2013.

[155] Y. L. Dokshitzer *Sov. Phys. JETP*, vol. 46, 1977.

[156] S. Forte, E. Laenen, P. Nason, and J. Rojo, "Heavy quarks in deep-inelastic scattering," *Nucl. Phys.*, vol. B834, pp. 116–162, 2010.

[157] K. Olive *et al.*, "Review of Particle Physics," *Chin.Phys.*, vol. C38, p. 090001, 2014.

[158] L. A. Harland-Lang, V. A. Khoze, and M. G. Ryskin, "Sudakov effects in photon-initiated processes," *Phys. Lett.*, vol. B761, pp. 20–24, 2016.

[159] S. Dittmaier and M. Huber, "Radiative corrections to the neutral-current Drell-Yan process in the Standard Model and its minimal supersymmetric extension," *JHEP*, vol. 01, p. 060, 2010.

[160] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, "Parton distributions in the LHC era: MMHT 2014 PDFs," *Eur. Phys. J.*, vol. C75, no. 5, p. 204, 2015.

[161] M. Czakon, N. P. Hartland, A. Mitov, E. R. Nocera, and J. Rojo, "Pinning down the large-x gluon with NNLO top-quark pair differential distributions," 2016.

[162] F. D. Aaron *et al.*, "Inclusive Deep Inelastic Scattering at High Q^2 with Longitudinally Polarised Lepton Beams at HERA," *JHEP*, vol. 09, p. 061, 2012.

[163] ATLAS Collaboration, "Determination of the strange quark density of the proton from ATLAS measurements of the W, Z cross sections," *Phys.Rev.Lett.*, 2012.

[164] ATLAS Collaboration, "Precision measurement and interpretation of inclusive W^+, W^- and Z/γ^* production cross sections with the ATLAS detector," 2016.

[165] S. S. Wilks, "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses," *Annals Math. Statist.*, vol. 9, no. 1, pp. 60–62, 1938.

[166] L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, and J. Rojo, "Unbiased determination of the proton structure function F(2)**p with faithful uncertainty estimation," *JHEP*, vol. 03, p. 080, 2005.

[167] L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, and J. Rojo, "Neural network determination of parton distributions: The nonsinglet case," *JHEP*, vol. 03, p. 039, 2007.

[168] R. D. Ball *et al.*, "A determination of parton distributions with faithful uncertainty estimation," *Nucl. Phys.*, vol. B809, pp. 1–63, 2009.

[169] C. Schmidt, J. Pumplin, D. Stump, and C. P. Yuan, "CT14QED parton distribution functions from isolated photon production in deep inelastic scattering," *Phys. Rev.*, vol. D93, no. 11, p. 114015, 2016.

[170] A. Buckley, J. Ferrando, S. Lloyd, K. Nordstrom, B. Page, *et al.*, "LHAPDF6: parton density access in the LHC precision era," *Eur.Phys.J.*, vol. C75, p. 132, 2015.

[171] ATLAS Collaboration, "Search for new high-mass resonances in the dilepton final state using proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector." ATL-COM-PHYS-2016-1404, Geneva, Sep 2016.

[172] ATLAS Collaboration, "Search for new high-mass resonances in the dilepton final state using proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector." ATL-COM-PHYS-2016-453, Geneva, May 2016.

[173] S. Agostinelli *et al.*, "GEANT4: A Simulation toolkit," *Nucl. Instrum. Meth.*, vol. A 506, pp. 250–303, 2003.

[174] ATLAS Collaboration, "The ATLAS Simulation Infrastructure," *Eur. Phys. J.*, vol. C70, pp. 823–874, 2010.

[175] ATLAS Collaboration, "Monte Carlo Generators for the Production of a W or Z/γ^* Boson in Association with Jets at ATLAS in Run 2." ATL-PHYS-PUB-2016-003, 2016.

[176] ATLAS Collaboration, "Multi-Boson Simulation for 13 TeV ATLAS Analyses." ATL-PHYS-PUB-2016-002, Geneva, Jan 2016.

[177] 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*, vol. 06, p. 043, 2010.

[178] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.*, vol. 178, pp. 852–867, 2008.

[179] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, "New parton distributions for collider physics," *Phys. Rev.*, vol. D82, p. 074024, 2010.

[180] ATLAS Collaboration, "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector," *JHEP*, vol. 2014, p. 55, 2014.

[181] J. Pumplin *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis," *JHEP*, vol. 07, p. 012, 2002.

[182] 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.*, vol. D69, p. 094008, 2004.

[183] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, "New parton distribution functions from a global analysis of quantum chromodynamics," *Phys. Rev.*, vol. D93, no. 3, p. 033006, 2016.

[184] S. G. Bondarenko and A. A. Sapronov, "NLO EW and QCD proton-proton cross section calculations with mcsanc-v1.01," *Comput. Phys. Commun.*, vol. 184, pp. 2343–2350, 2013.

[185] T. Gleisberg *et al.*, "Event generation with SHERPA 1.1," *JHEP*, vol. 02, p. 007, 2009.

[186] T. Sjostrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *J. High Energy Phys.*, vol. 0605, p. 026, 2006.

[187] P. Z. Skands, "Tuning Monte Carlo Generators: The Perugia Tunes," *Phys. Rev.*, vol. D82, p. 074018, 2010.

[188] M. Czakon and A. Mitov, "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders," *Comput. Phys. Commun.*, vol. 185, p. 2930, 2014.

[189] "ATLAS Run 1 Pythia8 tunes," Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov 2014.

[190] Willis, C., "The LPXSignalReweightingTool: A Package for Signal Template Reweighting in High-Energy Physics Analysis." ATL-COM-PHYS-2016-912, Geneva, 2016.

[191] J. Butterworth *et al.*, "PDF4LHC recommendations for LHC Run II," *J. Phys.*, vol. G43, p. 023001, 2016.

[192] P. Motylinski, L. Harland-Lang, A. D. Martin, and R. S. Thorne, "Updates of PDFs for the 2nd LHC run," in *International Conference on High Energy Physics 2014 (ICHEP 2014) Valencia, Spain, July 2-9, 2014*, 2014.

[193] G. e. a. Cowan, "Asymptotic formulae for likelihood-based tests of new physics," *The European Physical Journal C-Particles and Fields*, vol. 71, no. 2, pp. 1–19, 2011.

[194] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke, "HistFactory: A tool for creating statistical models for use with RooFit and RooStats," Jan 2012.

[195] L. Moneta, K. Cranmer, G. Schott, and W. Verkerke, "The RooStats project," in *PoS ACAT2010*, 2010.

[196] W. Verkerke and D. P. Kirkby, "The RooFit toolkit for data modeling," no. CHEP-2003-MOLT007, 2003.

[197] G. Choudalakis, "On hypothesis testing, trials factor, hypertests and the BumpHunter," in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011*, 2011.

[198] A. Caldwell, D. Kollar, and K. Kroninger, "BAT: The Bayesian Analysis Toolkit," *Comput. Phys. Commun.*, vol. 180, pp. 2197–2209, 2009.

[199] J. Gao, L. Harland-Lang, and J. Rojo, "The Structure of the Proton in the LHC Precision Era," *Phys. Rept.*, vol. 742, pp. 1–121, 2018.

[200] R. D. Ball, V. Bertone, M. Bonvini, S. Marzani, J. Rojo, and L. Rottoli, "Parton distributions with small-x resummation: evidence for BFKL dynamics in HERA data," *Eur. Phys. J.*, vol. C78, no. 4, p. 321, 2018.

[201] ATLAS Collaboration, "Improved luminosity determination in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector at the LHC," *Eur. Phys. J.*, vol. C73, no. 8, p. 2518, 2013.

[202] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms," *JHEP*, vol. 11, p. 040, 2004.

[203] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method," *JHEP*, vol. 11, p. 070, 2007.

[204] ATLAS Collaboration, "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector," 2014.

[205] S. Frixione, P. Nason, and G. Ridolfi, "A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction," *JHEP*, vol. 09, p. 126, 2007.

[206] M. Bahr *et al.*, "Herwig++ Physics and Manual," *Eur. Phys. J.*, vol. C58, pp. 639–707, 2008.

[207] J. A. M. Vermaseren, "Two Photon Processes at Very High-Energies," *Nucl. Phys.*, vol. B229, pp. 347–371, 1983.

[208] A. Suri and D. R. Yennie, "The space-time phenomenology of photon absorption and inelastic electron scattering," *Annals Phys.*, vol. 72, p. 243, 1972.

[209] F. W. Brasse, W. Flauger, J. Gayler, S. P. Goel, R. Haidan, M. Merkwitz, and H. Wriedt, "Parametrization of the q^2 dependence of $\gamma_V p$ total cross sections in the resonance region," *Nucl. Phys.*, vol. B110, pp. 413–433, 1976.

[210] T. Sjostrand, "High-energy physics event generation with PYTHIA 5.7 and JETSET 7.4," *Comput. Phys. Commun.*, vol. 82, pp. 74–90, 1994.

[211] B. Andersson, "The Lund model," *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, vol. 7, pp. 1–471, 1997.

[212] R. Corke and T. Sjostrand, "Multiparton Interactions and Rescattering," *JHEP*, vol. 01, p. 035, 2010.

[213] N. Davidson, T. Przedzinski, and Z. Was, "PHOTOS interface in C++: Technical and Physics Documentation," *Comput. Phys. Commun.*, vol. 199, pp. 86–101, 2016.

[214] W. Buttinger, "Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data." ATL-COM-SOFT-2015-119, Geneva, May 2015.

[215] ATLAS Collaboration, "Measurement of the transverse momentum and ϕ^*_η distributions of Drell-Yan lepton pairs in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector," *Eur. Phys. J.*, vol. C76, no. 5, p. 291, 2016.

[216] R. J. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples," *Comput. Phys. Commun.*, vol. 77, pp. 219–228, 1993.

[217] ATLAS Collaboration, "Measurement of the low-mass Drell-Yan differential cross section at $\sqrt{s}$ = 7 TeV using the ATLAS detector," *JHEP*, vol. 06, p. 112, 2014.