

Measurement of W/Z cross-section at 2.76 TeV in 2013 ATLAS experiment data

**Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Humboldt Universität**

vorgelegt von
Ksenia Gasnikova
aus Tchelyabinsk

Berlin
2017

Abstract

A measurement of the total $pp \rightarrow W^\pm/Z \rightarrow ll$ cross-section in electron and muon channel. For this analysis $4pb^{-1}$ of data collected at center-of-mass energy $\sqrt{s} = 2.76$ TeV in 2013 with ATLAS detector at LHC have been used.

In this measurement several sources of systematic uncertainties have been estimated. The background contribution have been estimated using both simulation (for electroweak and $t\bar{t}$ backgrounds) and data-driven methods (for QCD background). The results for separate electron and muon W^\pm and Z cross-sections are combined with the respect of the correlated uncertainties between channels. The combined cross-sections have been used for calculation of W/Z and W^+/W^- the cross-section ratios.

The resulting cross-sections and cross-section ratios have been compared and showed a good agreement with different next-to-leading and next-to-next-to-leading order QCD calculations. The overall uncertainty of these measurements is of the order of the PDF uncertainties, that makes them applicable for PDF fits.

The studies of the simulation of electrons inside forward calorimeter have been performed. The new method of binning finding procedure have been developed, that allowed to reduce time, spend on the generation of new libraries for the fast simulation method, called Frozen Showers.

Kurzfassung

Contents

| | |
|--|-----------|
| 1 Thesis organization | 1 |
| I Theory introduction | 3 |
| 2 Theoretical introduction | 5 |
| 2.1 Standard model | 5 |
| 2.2 Proton structure | 8 |
| 2.3 Physics of W and Z bosons in pp collisions | 10 |
| 3 Methodology of the measurement | 15 |
| 3.1 Cross-Section methodology | 15 |
| 3.1.1 Fiducial phase-space definition | 16 |
| 3.1.2 Extrapolation to the 13 TeV fiducial phase-space | 17 |
| 3.1.3 The $W^{+/-}$ boson cross-section calculation | 17 |
| 3.2 Ratios calculation | 18 |
| 3.3 Combination of electron and muon cross-section measurements | 18 |
| 3.4 PDF fits | 19 |
| II Experimental setup | 23 |
| 4 The Large Hadron Collider (LHC) and the ATLAS experiment. | 25 |
| 4.1 The LHC and accelerator complex | 25 |
| 4.2 The ATLAS experiment | 27 |
| 4.2.1 Coordinates and kinematic variables | 28 |
| 4.2.2 Inner detector | 28 |
| 4.2.3 Calorimeters | 29 |
| 4.2.4 Muon spectrometer | 33 |
| 4.2.5 Trigger system | 33 |
| 4.3 Luminosity measurement | 34 |
| 5 Event reconstruction | 37 |
| 5.1 Tracks and vertexes | 37 |
| 5.2 Electron reconstruction and identification | 38 |
| 5.2.1 Central electrons reconstruction | 38 |
| 5.2.2 Forward electrons reconstruction | 38 |
| 5.2.3 Electron identification | 39 |
| 5.3 Muon reconstruction and identification | 39 |
| 5.4 Missing transverse energy reconstruction | 40 |
| 5.4.1 Standard missing transverse energy reconstruction | 40 |
| 5.4.2 Reconstruction of missing transverse energy from hadronic recoil | 41 |

| | |
|--|-----------|
| 6 Monte-Carlo simulation | 45 |
| 6.1 ATLAS chain of Monte-Carlo production | 45 |
| 6.2 Event generators | 46 |
| 6.3 Simulation in Geant4 | 48 |
| 7 Frozen Showers | 49 |
| 7.1 Introduction | 49 |
| 7.2 Properties of electron showers in FCAL | 50 |
| 7.3 Generation and use in simulation | 55 |
| 7.3.1 Tuning of libraries- | 55 |
| 7.4 Machine learning based bin finding procedure | 57 |
| 7.4.1 Machine learning introduction | 57 |
| 7.4.2 Electron shower categorization | 59 |
| 7.4.3 Interpretation of results | 63 |
| 7.4.4 Reconstructed electron energy | 63 |
| 7.4.5 Plans for the future | 63 |
| 7.5 Validation of the new libraries | 67 |
| 8 Data and Monte-Carlo samples | 69 |
| 8.1 Data sample | 69 |
| 8.2 Monte-Carlo samples | 70 |
| III The Measurement | 71 |
| 9 Selection | 73 |
| 9.1 Data quality cuts | 73 |
| 9.2 Lepton quality cuts | 74 |
| 9.3 Boson selection | 75 |
| 9.4 Cut flow | 75 |
| 10 Monte-Carlo corrections | 77 |
| 10.1 Lepton efficiency corrections | 77 |
| 10.1.1 Muon Trigger SF | 79 |
| 10.2 Electron energy scale and resolution | 82 |
| 10.3 Muon momentum correction | 82 |
| 11 Hadronic recoil calibration | 85 |
| 11.1 Introduction | 85 |
| 11.2 Hadronic recoil resolution correction | 88 |
| 11.2.1 Event activity correction | 88 |
| 11.2.2 Resolution corrections using Z events | 94 |
| 11.3 Hadronic recoil bias correction | 97 |
| 11.3.1 Bias determination from M_T^W distribution | 98 |
| 11.3.2 Bias determination using u_{\parallel} distribution | 100 |
| 11.3.3 Systematic uncertainty estimation | 101 |
| 11.4 Summary on hadronic recoil systematics | 101 |

| | |
|--|------------|
| 12 Background estimation | 103 |
| 12.1 QCD background estimation | 104 |
| 12.1.1 Template selection | 105 |
| 12.1.2 Methodology of the template sample normalization | 105 |
| 12.1.3 Systematic uncertainty from the multi-jet background estimation | 107 |
| 12.2 Background-subtracted W and Z candidate events | 109 |
| 13 Control distributions | 111 |
| 14 Uncertainties in the cross-section measurement | 121 |
| 14.1 Methods of uncertainties propagation | 121 |
| 14.2 Systematic uncertainties on $C_{W,Z}$ | 122 |
| 14.2.1 Electron energy scale and resolution | 122 |
| 14.2.2 Muon energy scale and resolution | 122 |
| 14.2.3 Muon and electron efficiency SF | 123 |
| 14.2.4 Theoretical uncertainty | 123 |
| 14.3 Statistical uncertainty in data | 123 |
| 14.4 Theoretical uncertainty on $A_{W/Z}$ and $E_{W/Z}$ factors | 125 |
| 14.5 Correlation between uncertainties | 125 |
| 14.5.1 Treatment of partially correlated uncertainties | 126 |
| 15 Results of the cross-section measurement | 129 |
| 15.1 Cross-sections results | 129 |
| 15.1.1 Lepton universality test | 129 |
| 15.2 Combined results | 131 |
| 15.2.1 Cross-sections ratios | 132 |
| 15.3 Effect on PDF distributions | 137 |
| 16 Summary | 143 |
| A Cholesky transformation results | 145 |
| B Covariance matrixes | 147 |
| C Additional comparisons with theoretical predictions | 149 |
| D Additional PDF profiling plots | 153 |
| Bibliography | 157 |

1 Chapter 1

2 Thesis organization

3 This thesis presents the measurement of $W \rightarrow l\nu$ and $Z \rightarrow ll$ cross-sections in electron and muon
4 channels using 2.76 TeV data collected by ATLAS experiment.

5 The thesis is organized in three parts. The theoretical basis is described in part 1. The experimental
6 input and software organization is explained in the part 2. The cross-section measurement performed
7 by the author is described in part 3. The results and its interpretation via parton density functions
8 is presented in a final part.

9 The work presented was performed within the ATLAS collaboration. All plots in the thesis were
10 produced by the author, unless it is referenced otherwise.

11 The theoretical input as presented in part 1 consist of the following chapters:

12 **Chapter 2 Theoretical introduction**, contains brief overview of the current status of Standard
13 Model, the proton structure and theory of W and Z bosons in pp collisions. The cross-sections
14 predictions at NNLO order are presented.

15 **Chapter 3 Methodology**, describes procedures of the cross-section measurement and calculation
16 of their ratios and also methods of PDF extractions using, in addition to other measurement
17 the cross-sections provided by this analysis.

18 The experimental setup described in the part 2 in the following chapters:

19 **Chapter 4 The LHC and ATLAS experiment** gives an overview of the LHC accelerator complex
20 and its experiments. The ATLAS detector, used to collect data for this analysis, is shortly
21 discussed.

22 **Chapter 5 Event reconstruction** contains the detailed description of the event reconstruction.
23 The study of missing transverse energy reconstruction algorithm, done by author, is presented.

24 **Chapter 6 Monte-Carlo** provides an information of Monte Carlo production steps and a short
25 description of generators used in this analysis.

26 **Chapter 7 Frozen showers** gives a description of Frozen Showers method used for fast Monte
27 Carlo simulation. The machine learning procedure for optimization of this method, made by
28 author, is presented.

29 **Chapter 8 Data and Monte Carlo samples** describes data and Monte Carlo samples, used in the
30 analysis.

31 The following chapters present work, done by the author, unless stated otherwise:

- 32 **Chapter 9 Event selection** gives a set of selection criteria used to derive $W \rightarrow l\nu$ and $Z \rightarrow ll$ in
33 collected data
- 34 **Chapter 10 Monte-Carlo corrections** presents the correction, applied to Monte Carlo in order to
35 gain a better data vs Monte Carlo agreement. The correction factors have been derived by
36 the performance group, except for muon trigger scale factors, which are determined by the
37 author.
- 38 **Chapter 11 Hadronic recoil calibration** describes a method of missing transverse energy calibra-
39 tion in 2.76 GeV data and methods of the corresponding uncertainty determination.
- 40 **Chapter 12 Background estimation** provides a description of main backgrounds, that can pass
41 the selection criteria and techniques of their contribution estimation
- 42 **Chapter 13 Control distributions** shows the comparison of different data and Monte Carlo sim-
43 ulation distributions for all analyses, covered in this thesis.
- 44 **Chapter 14 Uncertainties of the cross-section measurements** presents main sources of ex-
45 perimental and theoretical uncertainties and gives methods of their propagation to the final
46 cross-sections and their ratios.
- 47 **Chapter 15 Results of the cross-section measurements** presents the results of the cross-
48 section measurement for $W \rightarrow l\nu$ and $Z \rightarrow ll$ in electron and muon channels separately.
49 The results have been used to test the lepton universality, predicted in the Standard Model.
50 Also, the combined cross-sections, their ratios are shown. The effect of these measurements
51 on the PDF distributions is estimated.

52

Part I

53

Theory introduction

Chapter 2

55 Theoretical introduction

| | |
|--|----|
| 58 2.1 Standard model | 5 |
| 59 2.2 Proton structure | 8 |
| 60 2.3 Physics of W and Z bosons in pp collisions | 10 |

62 This thesis presents the measurement of a Standard Model (SM) process, so the first section of
 63 this chapter gives a small overview of the Standard Model. Since this thesis uses data collected in
 64 proton-proton collisions, the proton structure (Sec. 2.2) and physics of W and Z bosons in proton
 65 collisions (Sec. 2.3) are discussed. Additionally, the predictions for $W \rightarrow l\nu$ and $Z \rightarrow ll$ cross-sections
 66 are presented in the last section.

67 A variety of sources were used for a preparation of this chapter, including [1, 2].

70 2.1 Standard model

71 The Standard Model (SM) is the theory, that explains the fundamental elementary particles
 72 and their interactions. It provides our best understanding of the particle physics and unifies the
 73 quantum mechanics, special relativity and a field theory. It was postulated by Weinberg-Salam in
 74 mid-1970 [3–5] and was successfully tested for the last 40 years. Despite the fact, that there are some
 75 unexplained phenomena in SM, such as dark matter and gravitation, it describes almost all laboratory
 76 data. The summary of all standard model cross-section measurements at ATLAS experiment is given
 77 in Fig. 2.1. The results are agreeing with SM over several orders of magnitude and no significant
 78 deviation from SM has been found yet.

79 The SM postulates two types of fundamental particles: fermions and boson. Graphical representa-
 80 tion of particles in SM with their masses and quantum numbers is shown in Fig. 2.2. The fermions are
 81 the spin 1/2 particles that form the matter. They can be further divided to leptons and quarks. Lep-
 82 tons can interact just electromagnetically and weakly, while quarks undergo strong and electroweak
 83 interactions. Both groups are divided into 3 separate generations.

84 The bosons are the carriers of the fundamental forces with the integer spin. Strong interactions
 85 are mediated by 8 massless gluons. The massless photons are carrying the electromagnetic interac-
 86 tions, while the W and Z bosons are responsible for the weak forces. The last SM boson, observed
 87 in experimental data is the Higgs boson [8, 9]. It is associated with Yukawa interactions, that are
 88 responsible for the fermion masses.

89 The SM is a *non-abelian gauge theory*, that means, that this theory is based on an invariant under
 90 local and global transformations (called *symmetries*) Lagrangian. From the Noether theorem [10, 11] it

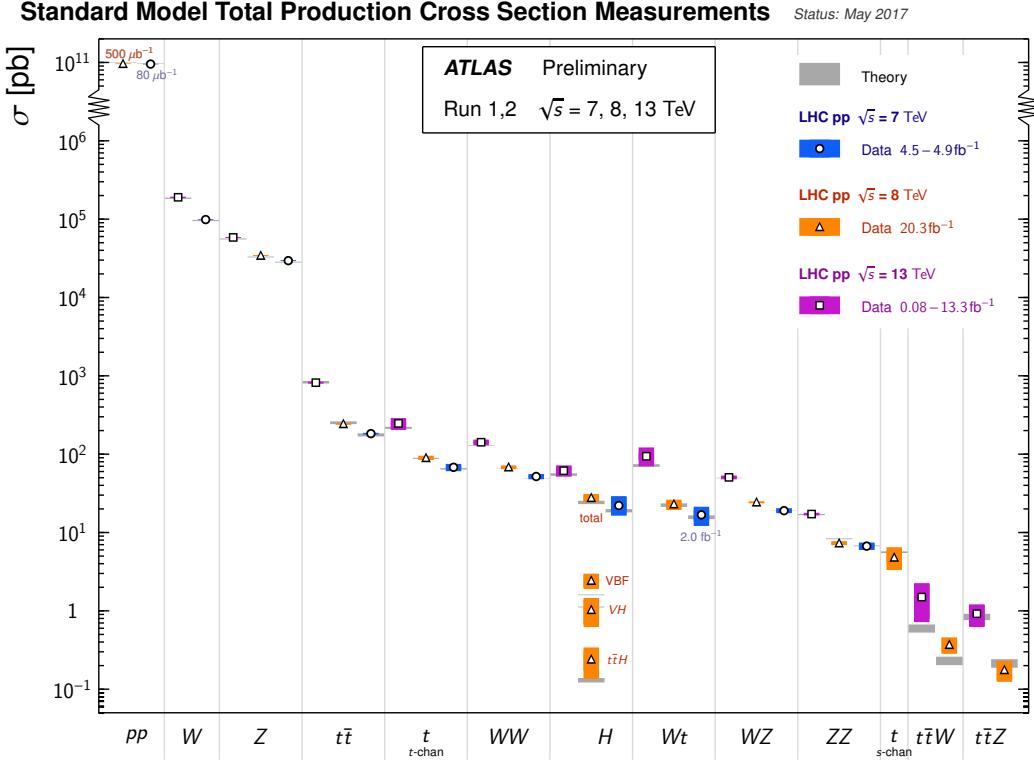


Fig. 2.1: Summary of several Standard Model total production cross section measurements, corrected for leptonic branching fractions, compared to the corresponding theoretical expectations. All theoretical expectations were calculated at NLO or higher. The luminosity used for each measurement is indicated close to the data point [6]

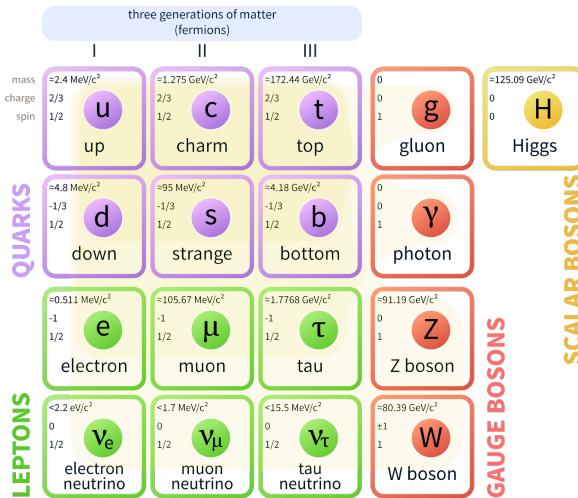


Fig. 2.2: All fundamental particles of the Standard Model with three generations of matter, the gauge and Higgs bosons [7]

is known, that each symmetry is connected to at least 1 conserved quantity. The symmetry group of SM is:

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

where Y a hypercharge, L - a left-handed helicity and C a color charge - are the conserved values of the corresponding symmetry group. The $SU(2)_L \times U(1)_Y$ symmetry stands for quantum electrodynamics (QED), while the $SU(3)_C$ corresponds to the theory of strong interactions - Quantum Chromodynamics (QCD) [12-14].

The QED postulates three massless vector fields in $SU(2)_L$ - the isospin triplet of vector fields $W_\mu^1, W_\mu^2, W_\mu^3$ with the coupling constant g and a single gauge field B_μ in $U(1)_L$ group with coupling strength g' . The actual γ and the massive Z and W bosons are produced due to the spontaneous breaking of the electroweak gauge symmetry as:

$$\begin{pmatrix} \gamma \\ Z^0 \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B \\ W^3 \end{pmatrix} \quad (2.2)$$

$$W^\pm = \frac{W^1 \pm iW^2}{\sqrt{2}}, \quad (2.3)$$

where θ_W is the electroweak mixing angle which value is not predicted in the theory. The masses of Z and W bosons are connected by the relation:

$$M_W^2 = \cos^2 \theta_W M_Z^2 \quad (2.4)$$

The QCD Lagrangian has only one parameter - a *strong coupling constant* α_s . QCD theory uses a concept similar to QED, however it is more complicated because of the color charge of gluons, that allows them to interact with each other. The inclusion of the gluon-gluon interaction causes a *ultraviolet* (UV) divergences in the corresponding integrals. The renormalization procedure [15] allows to move them into the coupling constant by introducing the arbitrary renormalization scale μ_R . A typical scale for a physics process corresponds to the momentum transferred Q^2 .

The QCD does not predict the value of α_s , however it can predict its evolution with renormalization scale using the renormalization group equation (RGE):

$$\beta(\alpha_s) = \mu^2 \frac{d\alpha_s(\mu^2)}{d\mu^2} = -b_0 \alpha_s^2(\mu) - b_1 \alpha_s^3(\mu) - \dots, \quad (2.5)$$

where the first coefficient is:

$$b_0 = \frac{33 - 2f}{12\pi}, \quad (2.6)$$

where $f=3$ is the number of flavors. This leads to energy dependence of the coupling constant showed in Fig. 2.3 together with results, obtained in experiments. The coupling constant increases with smaller scales, but becomes small for higher Q^2 . As a consequence, the quarks have a property of *asymptotic freedom* and *confinement*, meaning that they cannot be observed as free particles, but are almost not interacting in a bound state, like a proton.

The full physical quantities do not depend on renormalization scale, however, the calculation on any perturbative order is a function of μ_R . The cross-section of the partonic interaction (or the *partonic cross-section*) $\sigma_{ab \rightarrow X}$ can be expressed perturbatively in orders of α_s as:

$$\sigma_{ab \rightarrow X} = \hat{\sigma}_0 + \alpha_s(\mu_R) \hat{\sigma}_0 + \alpha_s^2(\mu_R) \hat{\sigma}_2 + \dots, \quad (2.7)$$

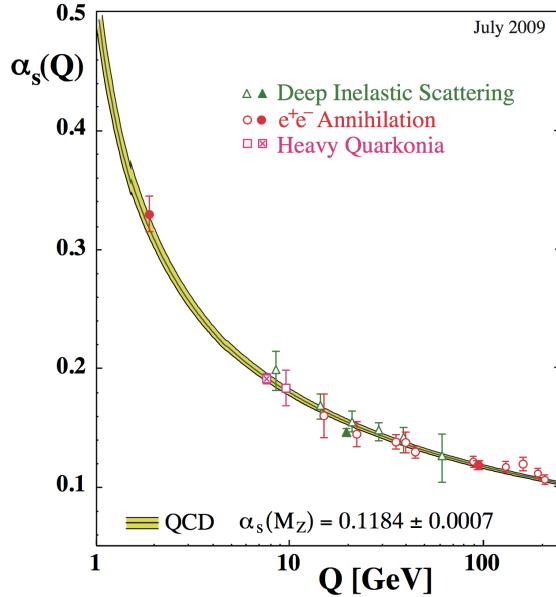


Fig. 2.3: The coupling α_s of the strong interaction as a function of energy scale [16]

where σ_i are the i-th order contributions to the final cross-section. The cross-section, calculated at lowest order of the expansion is called the *leading order* (LO) cross-section. The calculation using the expansion of α_s up to the i-th order is called (next to)ⁱ order (NⁱLO) cross-section, where "next-to" (N) is repeated i times. The inclusion of higher order corrections allows to reduce the dependency on the renormalization scale.

2.2 Proton structure

The cross-section of individual quark-quark interactions is predicted in QCD at the different orders of α_s . However, since the quarks haven't been observed in a free state, the test of QCD predictions is possible only using the experiments with hadrons (e.g. proton beams), so the internal quark composition in the hadron should be taken into account.

Richard Feynman proposed model of the proton structure called a parton model of the hadrons in 1969 [17]. In this model it is assumed, that any hadron can be treated as a composition of point-like constituents called partons. In the high momentum scattering the soft interaction between partons can be neglected, and therefore they can be treated as quasi free in collision. In this approximation of incoherent scattering a total cross-section for process in a hadron-hadron interaction could be written as a sum of all partonic cross-sections:

$$\sigma = \int dx_1 dx_2 f_1^{(P_1)}(x_1) f_2^{(P_2)}(x_2) \hat{\sigma}(x_1 x_2 s), \quad (2.8)$$

where:

- $f_1^{(P_1)}(x_1)$ and $f_2^{(P_2)}(x_2)$ are the parton distribution functions (PDF) for both colliding hadrons. It describes the probability to find parton i in hadron j with a fraction of longitudinal momentum x_i .

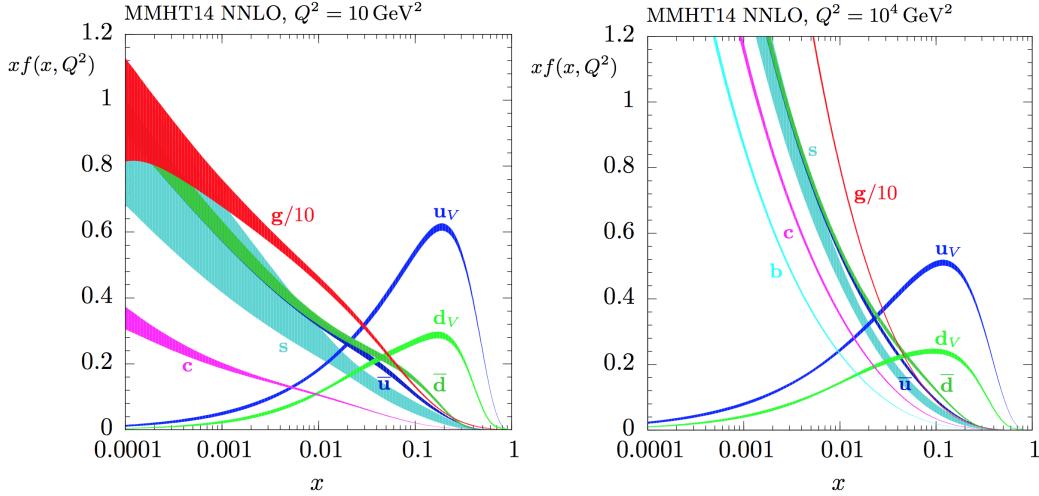


Fig. 2.4: The MMHT2014 NNLO PDFs predictions at $Q^2 = 10 \text{ GeV}^2$ (left) and $Q^2 = 10^4 \text{ GeV}^2$ (right), with associated 68% confidence-level uncertainty bands. [19]

- 118 • $\hat{\sigma}(x_1 x_2 s)$ is the partonic cross-section for a given scattering process, calculated in QCD.

119 The partons, which determine the quantum numbers of hadron, are called *valence quarks* (u and
 120 d quarks in case of the proton). However, due to the fluctuations, the infinite number of quark pairs
 121 of $u\bar{u}$, $d\bar{d}$, $c\bar{c}$ etc with low momentum could be created. These quarks are called *sea-quarks*. Due
 122 to the conservation of the total momentum and the flavor of a proton, the following sum rules are
 123 applicable for proton PDFs:

$$\begin{aligned} 124 \quad & \int_0^1 dx \sum_i x f_i^{(p)}(x) = 1 \\ 125 \quad & \int_0^1 dx (f_u^{(p)}(x) - f_{\bar{u}}^{(p)}(x)) = 1 \\ 126 \quad & \int_0^1 dx (f_d^{(p)}(x) - f_{\bar{d}}^{(p)}(x)) = 1 \\ 127 \quad & \int_0^1 dx (f_s^{(p)}(x) - f_{\bar{s}}^{(p)}(x)) = 1 \\ 128 \quad & , \end{aligned}$$

129 where index i runs over the quark flavors.

130 For the partonic cross-sections, the soft emmission of real and virtual gluons causes a collinear
 131 singularities that do not cancel. However, it is possible to include intial state emisions below a given
 132 scale into non-perturbative parton distribution functions. The cutoff parameter in this procedure is
 133 called a *factorisation scale* μ_F . This definition of PDFs is universal, i.e. does not depend on a physics
 134 process. Similarly to renormalization scale, the factorization scale μ_F is not a physical quantity and
 135 the total partonic cross-section should be independent independent of it. The typical choice of the
 136 factorization scale is $\mu_R \approx \mu_F \approx Q^2$.

137 These distributions cannot be calculated pertrubatively in QCD and therefore need to be obtained
 138 from the data. However, it is possible to predict the evolution of PDFs with factorisation scale in
 139 pertrubative order using the DLGAP evolution equations [18]. An example of quark and gluon PDFs
 140 predicted by MMHT2015 NNLO [19] for different scales are shown in Fig. 2.4. The PDFs of valence u
 141 and d quarks are peaking at $x=1/3$, while the sea quark and gluon densities are rising at low x. The
 142 contribution of the sea quarks becomes larger with increasing Q^2 .

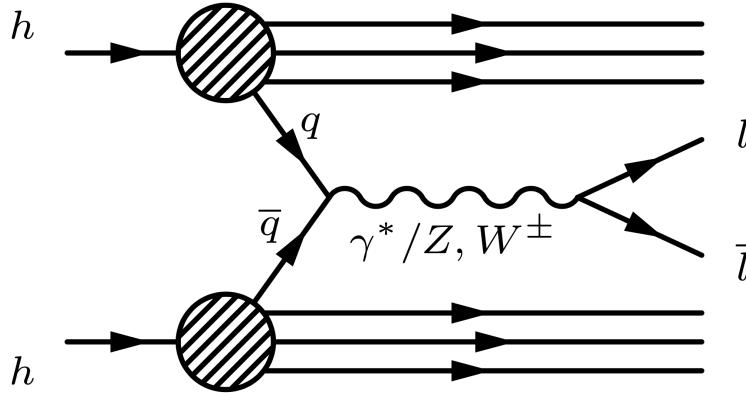


Fig. 2.5: A schematic representation of the W/Z production via Drell-Yan process in a hadron collider.

2.3 Physics of W and Z bosons in pp collisions

As it was mentioned the W and Z bosons are the vector bosons in Standard Model. They have been predicted by the Glasgow, Weinberg, Salam un 1960's and discovered in 1983 by UA1 and UA2 at CERN $p\bar{p}$ collider [20–23]. These particles are mediating weak interactions and decaying almost immidiately ($t \approx 10^{-25}$ s).

The leading algorithm of their production is a Drell-Yan mechanism, shown schematically in Fig. 2.5. The Drell-Yan process is the process of production of lepton pair with 4-momentum p_1 and p_2 and large invariant mass $M_{ll} = \sqrt{(p_1^l + p_2^l)^2}$ in quark-antiquark annihilation. A simpliest example of this process is the production of the virtual photon $q\bar{q} \rightarrow \gamma^* \rightarrow l^+l^-$. The corresponding cross-section can be calculated from fundamental QED as:

$$\hat{\sigma}(q\bar{q} \rightarrow \gamma^* \rightarrow e^+e^-) = \frac{4\pi\alpha}{3\hat{s}} \frac{1}{N} Q_q^2, \quad (2.9)$$

where Q_q is the quark charge, \hat{s} is the centre of mass energy of quark-antiquark system and $1/N=1/3$ is the overall color factor, coming from the fact, that quark and antiquark colors should match each other in order to create a colour-singlet in a final state.

In analogy to this process, the on-shell production of W and Z bosons with Drell-Yan process can be calculated as:

$$\hat{\sigma}^{q\bar{q}' \rightarrow W} = \frac{\pi}{3} \sqrt{2} G_F M_W^2 |V_{q\bar{q}'}|^2 \delta(\hat{s} - M_W^2), \quad (2.10)$$

$$\hat{\sigma}^{q\bar{q}' \rightarrow Z} = \frac{\pi}{3} \sqrt{2} G_F M_Z^2 (v_q^2 + a_q^2) \delta(\hat{s} - M_Z^2), \quad (2.11)$$

where the q and q' are the different quarks. The $V_{q\bar{q}'}$ is the appropriate Cabibbo-Kobayashi-Maskawa matrix element, what describes a strength of flavor changing weak decays, and v_q (a_q) the vector (axial vector) coupling of the Z to the quarks. For the full production cross-section, the partonic spectrum of colliding hadrons should be considered. For 2 quarks, carrying fractions x_1 and x_2 of the protons momenta, the momentum transfer Q^2 can be written as:

$$Q^2 = (x_1 p_1 + x_2 p_2) \approx x_1 x_2 \sqrt{s} = M_{Z,W,\gamma^*}^2, \quad (2.12)$$

where the parton masses have been neglected in the calculation, s in the Eq. 2.12 is the center-of-

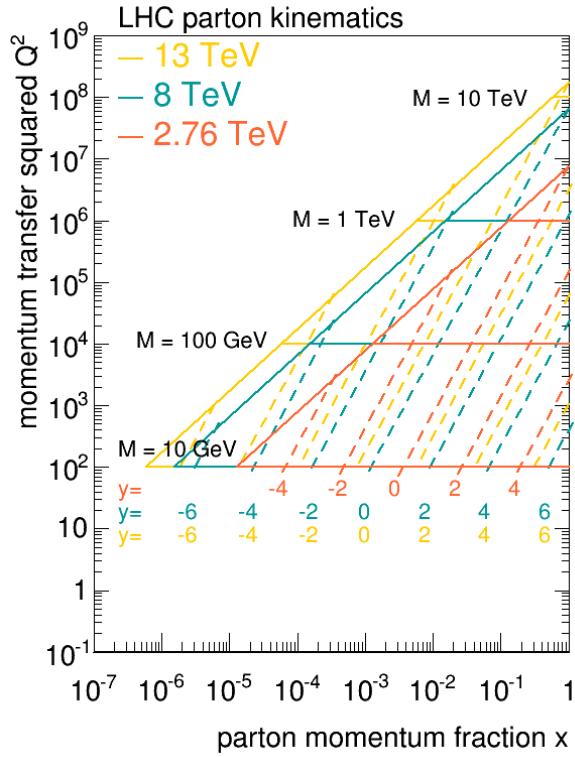


Fig. 2.6: The LHC parton kinematics in $Q^2 \times x$ plane. The LHC experiment limits are shown at $\sqrt{s} = 13$ TeV, 8 TeV and 2.76 TeV.

mass energy of 2 hadrons:

$$s = (p_1 + p_2)^2. \quad (2.13)$$

Another Lorentz invariant, used to describe the parton scattering is the rapidity of the boson, defined as:

$$y = \frac{1}{2} \log \frac{E + P_z}{E - P_z}, \quad (2.14)$$

where E is the energy of the boson and P_z is the z component of the momentum. This quantity can be connected to the momentum fraction carried by initial partons in a leading order approximation as:

$$x_{1,2} = \frac{M_{Z,W,\gamma} e^{\pm y}}{\sqrt{s}}. \quad (2.15)$$

The maximum accessible rapidity range for the production of the bosons can be determined from the center-of-mass energy and the mass of the boson as:

$$|y_{bos}^{max}| = \ln \frac{\sqrt{s}}{M_{bos}}. \quad (2.16)$$

151 The full parton kinematic phase-space, accessible at LHC for a different center-of-mass energies
 152 is shown in Fig. 2.6. The W and Z cross-sections at 2.76 TeV correspond to $Q^2 \approx 10^4 \text{GeV}^2$ and
 153 therefore can probe the ranges of partons fractions $x^W > 8 \cdot 10^{-4}$ and $x^Z > 1 \cdot 10^{-3}$ for W and
 154 Z bosons respectively. The W^+ production depends mainly on the u and \bar{d} distributions (because

Table 2.1: Branching ratios of the different W and Z decay modes [25]. Invisible denotes the Z decays with a neutrino-antineutrino pair as a final state. The predicted values are estimated with $\sin^2\theta_W = 0.23$.

| Boson | Decay mode | Measured branching ratio | SM prediction |
|-------|----------------|--------------------------|---------------|
| W | $e\nu_e$ | $(10.71 \pm 0.16)\%$ | |
| | $\mu\nu_\mu$ | $(10.63 \pm 0.15)\%$ | 11.1% |
| | $\tau\nu_\tau$ | $(11.38 \pm 0.21)\%$ | |
| | hadrons | $(67.41 \pm 0.27)\%$ | 66.7% |
| Z | e^+e^- | $(3.363 \pm 0.004)\%$ | |
| | $\mu^+\mu^-$ | $(3.366 \pm 0.007)\%$ | 3.4 % |
| | $\tau^+\tau^-$ | $(3.3658 \pm 0.008)\%$ | |
| | invisible | $(20.00 \pm 0.06)\%$ | 20.5% |
| | hadrons | $(69.91 \pm 0.06)\%$ | 69.2% |

of the leading order process $u\bar{d} \rightarrow W^+$), the W^- oppositely to \bar{u} and d distributions. The Z boson production cross-section is mostly sensitive to the valence quark distributions.

The Drell-Yan process contributes to around 65% of the total production cross-section [24], with both valence-sea and sea-sea quarks interactions included. The dominant higher order cross-section correction is the interaction of a quarks with a gluon, that occurs in approximately 20% of the events.

Due to the small time of life of W and Z bosons, the production is instantly followed by a decay. The probability of a certain decay mode is described by the *branching ratio*, $\text{BR}(X \rightarrow a + b)$, that is a fraction of a partial decay rate of the decay mode of interest and the total decay rate of the boson. The different decay modes are summarized in Tab. 2.1. The W and Z bosons can decay hadronically with production of fermion-antifermion pair for all fermions, except for top quark, in case of the W boson, since its mass exceeds the mass of W. This mode of the decay is the dominant one because of the 3 possible color states for each quark. In a leptonic decay of the W boson the lepton plus corresponding same flavor neutrino/antineutrino pair is produced. The leptonic decays of Z create a lepton-antilepton pair. The visible fraction of Z bosons decaying into leptons is smaller, compared to W, because of the invisible mode of Z decay, where neutrino-antineutrino pair is produced.

Due to experimental difficulty to measure the hadronic decays, the W/Z cross-sections are usually measured through their leptonic decays. The expected NNLO production cross-sections times their branching ratios estimated using FEWZ program [26] and CT14nnlo [27], at 2.76 TeV are:

$$\sigma_{W^+ \rightarrow l\nu}^{\text{NNLO}} = 2114^{+8}_{-11}(\text{scale})^{+57}_{-59}(\text{PDF})[\text{pb}] \quad (2.17)$$

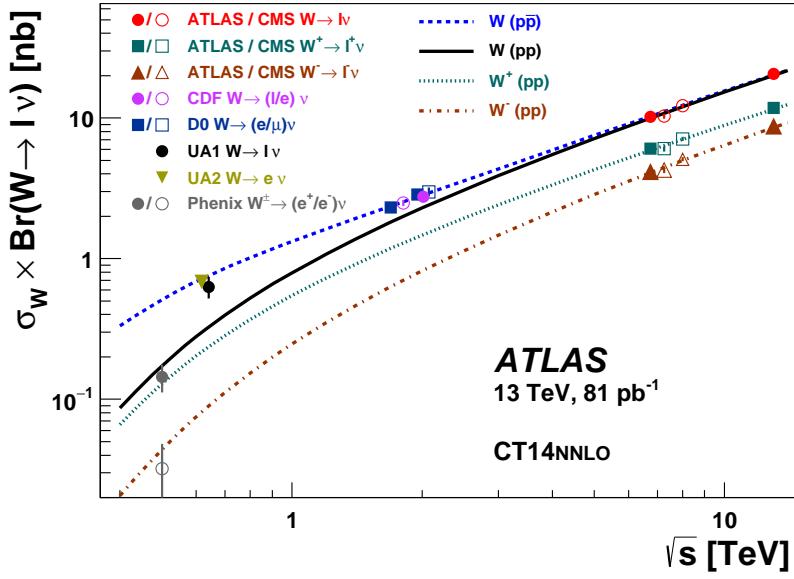
$$\sigma_{W^- \rightarrow l\nu}^{\text{NNLO}} = 1265^{+5}_{-6}(\text{scale})^{+32}_{-38}(\text{PDF})[\text{pb}] \quad (2.18)$$

and

$$\sigma_{Z^- \rightarrow ll}^{\text{NNLO}} = 304^{+1}_{-1}(\text{scale})^{+7}_{-7}(\text{PDF})[\text{pb}], \quad (2.19)$$

where the first uncertainty comes from the uncertainty of the scale $Q^2 = M_{\text{bos}}^2$. The second uncertainty arises from the imperfect knowledge of proton PDFs. The difference between W^+ and W^- cross-sections (called charge asymmetry) is due to the higher probability of finding a u_v quark rather than d_v quark in protons.

The NNLO predictions of W and Z cross-sections in pp and $p\bar{p}$ collisions and results obtained at



a)

Fig. 2.7: The measured values of $\sigma_{W \rightarrow l\nu}$ for W^+ , W^- and their sum compared to the theoretical predictions based on NNLO QCD calculations. The ATLAS results are shown for the combined electron-muon channel only. The predictions and previous measurements are shown for both proton-proton and proton-antiproton colliders as a function of \sqrt{s} . The data points at the various energies are staggered to improve readability. All data points are displayed with their total uncertainty. The calculations were performed with the program FEWZ using the CT14nnlo parton density function parametrisation. The theoretical uncertainties on the cross-section predictions are not shown. Taken from [28]

175 different experiments are shown in Fig. 2.7- 2.8. It can be seen, that the W and Z cross-sections
 176 have not been measured so far in region around 3 TeV in proton-proton collisions.

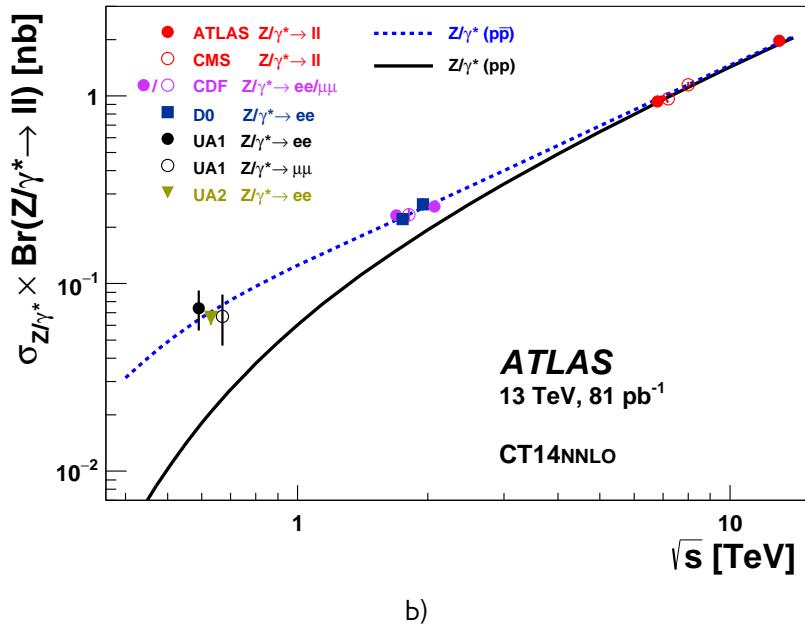


Fig. 2.8: The measured values of $\sigma_{Z \rightarrow ll}$ for Z-boson compared to the theoretical predictions based on NNLO QCD calculations. The ATLAS results are shown for the combined electron-muon channel only. The predictions and previous measurements are shown for both proton-proton and proton-antiproton colliders as a function of \sqrt{s} . The data points at the various energies are staggered to improve readability. All data points are displayed with their total uncertainty. The calculations were performed with the program FEWZ using the CT14nnlo parton density function parametrisation. The theoretical uncertainties on the cross-section predictions are not shown. Taken from [28]

177 Chapter **3**

178 **Methodology of the measurement**

179

| | | |
|-----|--|----|
| 180 | 3.1 Cross-Section methodology | 15 |
| 181 | 3.1.1 Fiducial phase-space definition | 16 |
| 182 | 3.1.2 Extrapolation to the 13 TeV fiducial phase-space | 17 |
| 183 | 3.1.3 The $W^{+/-}$ boson cross-section calculation | 17 |
| 184 | 3.2 Ratios calculation | 18 |
| 185 | 3.3 Combination of electron and muon cross-section measurements | 18 |
| 186 | 3.4 PDF fits | 19 |
| 187 | | |
| 188 | | |
| 189 | | |

190 This thesis presents various measurements on 2.76 TeV data: cross-section measurement for
191 different lepton flavors, combination of the cross-sections and interpretation of the obtained cross-
192 sections in terms of pdf distributions.

193 In this chapter a methodology of the measurements is given. First section gives a description of
194 cross-section measurement in the different phase-space regions. The second chapter shows the
195 methodology of linear averaging of the results, used for combination of electron and muon analyses.
196 In the last section the pdf fit procedure is described.

197 **3.1 Cross-Section methodology**

198 Due to the limited geometrical detector acceptance and reconstruction efficiency the cross-section
199 of a process of interest cannot be measured in the full phase-space. The fiducial cross-section, i.e.
200 the cross-section limited to selection region, is measured, as:

$$\sigma_{W/Z}^{fid} = \frac{N^{W/Z} - B^{W/Z}}{C_{W/Z} L_{int}} = \frac{N_{sig}^{W/Z}}{C_{W/Z} L_{int}}, \quad (3.1)$$

201 where:

- 202 • N is the number of candidates measured in the data;
203 • B is the number of background events;
204 • $N_{sig}^{W/Z} = N^{W/Z} - B^{W/Z}$ is the number of the signal events;
205 • L_{int} is the integrated luminosity calculated for the given run selection and trigger requirements;

- 204 • $C_{W/Z}$ is a correction factor for an experimental selection and resolution effects.

The correction factor $C_{W/Z}$ is calculated from Monte Carlo (MC) for each process and each decay channel separately and defined as :

$$C_{W/Z} = \frac{N_{MC,rec}}{N_{MC,gen,cut}}, \quad (3.2)$$

205 where $N_{MC,rec}$ is a sum of weights of events after simulation, reconstruction and selection, $N_{MC,gen,cut}$
 206 - is taken on the generator level after fiducial cuts. This correction factor includes efficiencies
 207 for trigger, reconstruction and identification (see Sec.10.1).

Later, this measurement can be extrapolated to the full phase-space using MC information:

$$\sigma_{W/Z}^{tot} = \frac{\sigma_{W/Z}^{fid}}{A_{W/Z}} = \frac{N_{sig}^{W/Z}}{A_{W/Z} C_{W/Z} L_{int}}, \quad (3.3)$$

208 where $\sigma_{W/Z}^{tot}$ is the total inclusive production cross-section of the W or Z bosons, and $A_{W/Z}$ is the
 209 acceptance.

The $A_{W/Z}$ factor is determined from Monte Carlo simulation as:

$$A_{W/Z} = \frac{N_{MC,gen,cut}}{N_{MC,gen,all}}, \quad (3.4)$$

210 where $N_{MC,gen,all}$ are the sum of weights of all generated MC events. Both $A_{W/Z}$ and $C_{W/Z}$ are defined
 211 at the "born level", i.e. before the decay leptons emit photons via QED final state radiation.

212 3.1.1 Fiducial phase-space definition

213 The fiducial volume definition corresponds to the analysis selection described in Chap. 9. For W boson
 214 measurement it is:

- 215 • $P_T^l > 20$ GeV
 216 • $|\eta^l| < 2.5$
 217 • $P_T^\nu > 25$ GeV
 218 • $M_T^W > 40$ GeV

where P_T^l and P_T^ν are the charged lepton and neutrino transverse momentums respectively, η^l is a lepton pseudo-rapidity, and M_T^W is the transverse mass, defined as:

$$M_T^W = \sqrt{2P_T^l \cdot P_T^\nu [1 - \cos(\phi_l - \phi_\nu)]}, \quad (3.5)$$

219 where ϕ_l is an azimuthal angle of the charged lepton and ϕ_ν is an azimuthal angle of neutrino.

220 For the Z boson measurement fiducial phase-space is defined as:

- 221 • $P_T^l > 20$ GeV
 222 • $|\eta^l| < 2.5$
 223 • 66 GeV $< M_Z < 116$ GeV

224 where M_Z is a di-lepton invariant mass.

225 The differences between analysis selection for electrons and muons are neglected in this definition.

226 **3.1.2 Extrapolation to the 13 TeV fiducial phase-space**

The 13 TeV inclusive cross-section measurement uses stricter set of analysis criteria and therefore another definition of the fiducial phase space. However, it is possible to extrapolate cross-section to a new fiducial phase-space as:

$$\sigma_{W/Z}^{fid,13} = \frac{\sigma_{W/Z}^{fid}}{E_{W/Z}}, \quad (3.6)$$

where $E_{W/Z}$ is an additional extrapolation factor:

$$E_{W/Z} = \frac{N_{MC,gen,cut}}{N_{MC,gen,cut^{new}}}, \quad (3.7)$$

227 where $N_{MC,gen,cut^{new}}$ is the sum of the weights of Monte-Carlo events after the new set of cuts on the
228 generator level. This extrapolation allows to directly calculate the cross-sections ratios for a different
229 \sqrt{s} for correlated uncertainties cancellation.

230 The 13 TeV fiducial phase-space is defined as:

231 • $P_T^l > 25$ GeV

232 • $|\eta^l| < 2.5$

233 • $P_T^\nu > 25$ GeV

234 • $M_T^W > 50$ GeV

235 for the W decays and

236 • $P_T^l > 25$ GeV

237 • $|\eta^l| < 2.5$

238 • 66 GeV $< m_Z < 116$ GeV

239 for the Z measurements.

240 **3.1.3 The $W^{+/-}$ boson cross-section calculation**

In this analysis the cross-section for the production of W bosons is calculated indirectly by combining W^+ and W^- cross-sections in the following way:

$$\sigma_W = \sigma_{W^+} + \sigma_{W^-} = \frac{1}{L_{int}} \cdot \left(\frac{N_{sig}^{W^+}}{A_{W^+} C_{W^+}} + \frac{N_{sig}^{W^-}}{A_{W^-} C_{W^-}} \right). \quad (3.8)$$

The absolute uncertainty of this measurement is calculated from uncertainties of W^+ and W^- cross-sections as:

$$(\delta X_W)^2 = (\delta X_{W^+})^2 + (\delta X_{W^-})^2 + 2 \cdot \rho_{W^+W^-}^X \delta X_{W^+} \delta X_{W^-}, \quad (3.9)$$

241 where δX is a systematic component on the cross-section and $\rho_{W^+W^-}^X$ is a correlation between W^+
242 and W^- for this component, which will be estimated in Chap. 14.

243 3.2 Ratios calculation

Ratio calculation is a powerful tool for testing the predictions of the theory, because of the cancellation of the correlated uncertainties. The ratio for two cross-section measurements σ_i and σ_j is calculated as:

$$R_{i,j} = \frac{\sigma_i}{\sigma_j} = \frac{\frac{N_{sig}^i}{A_{W/Z}^i C_{W/Z}^i L_{int}}}{\frac{N_{sig}^j}{A_{W/Z}^j C_{W/Z}^j L_{int}}} = \frac{N_{sig}^i}{N_{sig}^j} \cdot \frac{A_{W/Z}^j}{A_{W/Z}^i} \cdot \frac{C_{W/Z}^i}{C_{W/Z}^j} = \frac{N_{sig}^i}{N_{sig}^j} \cdot A_{i/j} C_{i/j}, \quad (3.10)$$

244 what means that this value does not depend on the integrated luminosity and its uncertainty.

The relative uncertainty can therefore be obtained taking into account correlation between two measurements as:

$$\left(\frac{\delta R}{R}\right)^2 = \left(\frac{\delta N_{sig}^i}{N_{sig}^i}\right)^2 + \left(\frac{\delta N_{sig}^j}{N_{sig}^j}\right)^2 + \left(\frac{\delta A_{i/j}}{A_{i/j}}\right)^2 + \left(\frac{\delta C_{i/j}}{C_{i/j}}\right)^2. \quad (3.11)$$

Uncertainties on first two terms are considered to be uncorrelated, while other are derived using the following error propagation formula:

$$\left(\frac{\delta X_{i/j}}{X_{i/j}}\right)^2 = \left(\frac{\delta X_i}{X_i}\right)^2 + \left(\frac{\delta X_j}{X_j}\right)^2 - 2\rho_{ij} \frac{\delta X_i}{X_i} \frac{\delta X_j}{X_j}, \quad (3.12)$$

245 where X is a systematic component on the A or C and ρ_{ij} is a correlation between two estimates.

246 The estimation of correlation parameters will be discussed in Chap. 14.

247 3.3 Combination of electron and muon cross-section measurements

249 The standard model predicts the same branching ratios for the leptonic decays of W and Z bosons.
 250 Therefore, it is possible to combine the measurements in electron and muon channel into one
 251 cross-section by averaging them. This analysis uses the standard tool for the averaging of the mea-
 252 surements called Havereger, that was originally developed for the HERA data [29]. This program uses
 253 the method of linear averaging, that is described below.

The probability density function for a "true" value m to get a value μ with uncertainty Δ in measurement is:

$$P(m) = \frac{1}{\sqrt{2\pi\Delta}} \exp\left(-\frac{(m-\mu)^2}{2\Delta^2}\right), \quad (3.13)$$

where it is assumed, that the uncertainty has a Gaussian shape. The corresponding χ^2 function is:

$$\chi^2(m) = \frac{(m-\mu)^2}{\Delta^2}. \quad (3.14)$$

In case of N statistically independent measurements, the probability density function is proportional to:

$$P(m) \propto \prod_{i=0}^N \exp\left(-\frac{(m-\mu_i)^2}{2\Delta_i^2}\right), \quad (3.15)$$

which corresponds to the χ^2_{sum} :

$$\chi^2_{sum}(m) = \sum_i \chi_i^2 = \sum_i \frac{(m - \mu_i)^2}{\Delta_i^2}, \quad (3.16)$$

that can be rewritten in the form of the Eq. 3.14:

$$\chi^2_{sum}(m) = \frac{(m - \mu_{ave})^2}{\Delta_{ave}^2} + \chi_0^2, \quad (3.17)$$

where μ_{ave} and Δ_{ave} are the average value and its uncertainty respectively. These values are found by minimizing χ^2_{sum} with respect to m . The value χ_0^2 indicate a consistency of the measurements and should be $\chi_0^2/N \approx 1$.

Systematic uncertainties can be treated as a result of an experiment (e.g. measurement of the calibration) and added to the χ^2 as:

$$\chi^2_{syst}(\alpha) = \frac{\alpha - \alpha_0}{\Delta_\alpha^2} \equiv b^2, \quad (3.18)$$

where α is the "true" parameter, what is measured at the experiment as α_0 with uncertainty Δ_α . The nuisance parameter b corresponds to a coherent change of the measurements $\mu_i \rightarrow \mu_i + \mathbf{b}\mathbf{F}_i$.

Using this nuisance parameters representation, Eq. 3.17 can be rewritten in a more general way:

$$\chi^2_{sum}(\mathbf{m}, \mathbf{b}) = \sum_i \frac{(m - \mu_i - \sum_j \Gamma_i^j b_j)^2}{\Delta_i^2} + \sum_j b_j^2, \quad (3.19)$$

where

- i runs over all experiments used in averaging;
- \mathbf{b} is the vector of nuisance parameters b_j corresponding to each source of systematic uncertainty;
- Γ_i^j is the absolute correlated systematic uncertainty;
- Δ_i^2 is the uncorrelated (statistical) uncertainty.

3.4 PDF fits

In this thesis the xFitter program [30] has been used to determine the parton distribution functions. In this section the used fit formalism is presented in details, such as parametrisation of the PDFs at the starting scale, definition of χ^2 and the treatment of the experimental uncertainties.

The parton distribution functions can be parametrized differently at the starting scale Q_0^2 . Standard form, adapted by the groups, providing the PDF distributions, uses a simple polynomial for interpolation between low and high x regions:

$$xf(x; Q_0^2) = Ax^\beta(1-x)^\gamma P_i(x), \quad (3.20)$$

where P_i is a polynomial of some order. In this analysis a standard HERA style parametrisation is used, so its implementation will be discussed in details

271 There are, in principle, 11 different partons to consider, however heavy parton (c and b quarks)
 272 distributions can be determined perturbatively, that leaves at least 7 independent combinations.
 273 The parameterized PDFs at HERA are the valence distributions xu_v and xd_v , the gluon distribution
 274 xg , and the u-type and d-type sea $x\bar{U}$ and $x\bar{D}$ respectively, where:

$$275 \quad xu_v = xu - x\bar{u}, \quad xd_v = xd - x\bar{d}, \quad x\bar{U} = x\bar{u}, \quad x\bar{D} = x\bar{d} + x\bar{s}.$$

276 For the strange quark distributions it is assumed, that $xs = x\bar{s} = f_s x\bar{D}$ at Q_0^2 , where $f_s = 0.31$ - is
 277 a strange fraction chosen to match the experimental data.

The following functional form is used for parametrization:

$$xf(x; Q_0^2) = Ax^B(1-x)^C(1+Dx+Ex^2), \quad (3.21)$$

278 where A_{uv} , A_{dv} and A_g are constrained by the number sum-rules and the momentum sum-rule, and
 279 the parameters $B_{\bar{U}}$ and $B_{\bar{D}}$ are set equal, so there is a single parameter for the sea distributions.

Similarly to averaging procedure, the nuisance parameters representation of the systematic uncertainties is used:

$$\mu_i = m_i(\mathbf{p}) + r_i \sigma_i + \sum_{\alpha=1}^{N_{syst}} \Gamma_{\alpha}^i b_{\alpha}, \quad (3.22)$$

280 where:

- 281 • $m_i(\mathbf{p})$ is the true value, that depends on set of parameters $\mathbf{p} = (p_1, p_2, \dots)$;
- 282 • μ_i is the value, observed in the experiment;
- 283 • σ_i are the statistical and systematic uncertainties;
- 284 • Γ_{α}^i is the sensitivity of i-th measurement to the correlated systematic source α ;
- 285 • r_i are the normal random variables;
- 286 • b_{α} are the nuisance parameters.

The simple parametrisation form of χ^2 is defined in Eq. 3.19. This equation can be rewritten using a Poisson distribution for the statistical error:

$$\chi_{sum}^2(\mathbf{m}, \mathbf{b}) = \sum_i \frac{(m - \mu_i - \sum_j \Gamma_j^i b_j)^2}{\delta_{i,stat}^2 \mu^i m^i \prod_{\alpha} \exp(-\gamma_{\alpha}^i b_{\alpha})} + \sum_{\alpha} b_{\alpha}^2, \quad (3.23)$$

287 The PDF uncertainties are estimated by varying the data points within their statistical and sys-
 288 tematic uncertainties using a Monte-Carlo technique. For each data set a QCD fit is performed to
 289 extract the PDF set. Typical number of data sets is $N > 100$. The pdf uncertainties are estimated using
 290 the values and RMS of the replicas.

The effect of adding new data to the PDF determination can be evaluated using the profiling procedure [31]. It is preformed using the following representation of the χ^2 function with both theoretical and experimental uncertainties included:

$$\chi^2(\mathbf{m}, \mathbf{b}_{exp}, \mathbf{b}_{th}) = \sum_i \frac{(m - \mu_i - \sum_{\alpha} \Gamma_{\alpha}^{exp} b_{\alpha,exp} + \sum_{\beta} \Gamma_{\beta}^{th} b_{\beta,th})^2}{\delta_{i,stat}^2 \mu^i m^i \prod_{\alpha} \exp(-\gamma_{\alpha}^i b_{\alpha})} + \sum_{\alpha} b_{\alpha,exp}^2 + \sum_{\beta} b_{\beta,th}^2, \quad (3.24)$$

291 where \mathbf{b}_{exp} and \mathbf{b}_{th} are the vectors of correlated experimental and theoretical uncertainties re-
292 spectively. The sensitivity of the measurements is splitted into two experimental and theoretical
293 components $\Gamma_{\alpha}^{\text{exp}}$ and $\Gamma_{\beta}^{\text{th}}$. This χ^2 function can be represented as a system of linear equations, that
294 are minimized iteratively, allowing to determine shifted PDF parameters. The minimum value of χ^2
295 provides a compatibility of the measurement.

296

Part II

297

Experimental setup

298 Chapter **4**

299 **The Large Hadron Collider (LHC) and the**
300 **ATLAS experiment.**

301

302

| | |
|--|----|
| 303 4.1 The LHC and accelerator complex | 25 |
| 304 4.2 The ATLAS experiment | 27 |
| 305 4.2.1 Coordinates and kinematic variables | 28 |
| 306 4.2.2 Inner detector | 28 |
| 307 4.2.3 Calorimeters | 29 |
| 308 4.2.4 Muon spectrometer | 33 |
| 309 4.2.5 Trigger system | 33 |
| 310 4.3 Luminosity measurement | 34 |

311

314 Particle physics requires to build a large machine for probing it theories. This a modern field of the
315 physics, that pushes a limits of engineering. In this section an overview of the experimental setup,
316 used to collect a data for this analysis will be given. The first section introduces an cutting edge of
317 particle accelerators: LHC. In the second section an ATLAS detector will be described.

318 **4.1 The LHC and accelerator complex**

319 Large hadron collider is the largest accelerator in the world. It was build in Geneva, Switzerland and
320 started its operation with the first collisions in 2009. It lies in a tunnel 27 kilometers in circumference.
321 It operates with pp and Pb-Pb beams with centre-of-mass energies up to 14 GeV. It was designed
322 to make a precise studies of Standard Model predictions and to search for a new physics beyond
323 standard model, such as supersymmetry. The heavy ion program serves a purpose of studying
324 matter properties and a quark-gluon plasma.

325 Beams acceleration consists of several stages [32], as it shown in Fig. 4.1. The beam source is
326 a hydrogen gas or lead source in case of the heavy ion. An electrical current is used to remove
327 the electrons from each atom, and then the ion begin its ride through the only linear accelerator
328 in the chain LINAC2, that accelerates protons up to 50 GeV or LINAC3 with up to 3.2 GeV/nucleon.
329 The proton beams from linac are ejected into the PS booster, where they are further accelerated
330 to 1.4 GeV. During a heavy ion operations nucleons from LINAC3 are injected into the Low Energy
331 Ion Ring (LEIR) and are accelerated up to 72.2 MeV/nucleon before injecting into PS. The last steps
332 before ejecting beams to the LHC are the rings of the Proton Synchrotron (PS) and Super Proton
333 Synchrotron (SPC) that accelerates protons to 25 and 450 GeV respectively. The bunch structure of

334 the beam is formed at PS step and has a nominal pattern of 39 groups of 72 bunches with time
 335 25-50 ns time spacing.

336 In LHC beams are accelerated up to 7 TeV (yet 6.5 TeV achieved). The beams are circulating
 337 in opposite directions inside one of the 2 beam pipes, that are 6.3 cm in diameter. In order to
 338 bend beam trajectory, the pipes are surrounded by a 1232 superconducting dipole magnets. The
 339 superconducting cavities are used to accelerate the protons and maintain beam constant energy
 340 during the operation time.

341 As most of the circular colliders, LHC has several experiments installed in the regions, where
 342 beams are intersecting, that allows them to collect data in parallel. The main experiments are:

343 **ALICE** A large ion Collider Experiment [33] - a dedicated heavy ion detector, build to in the physics
 344 of strongly interacting matter, where a new phase of mater (quark-gluon plasma) is expected.

345 **ATLAS** A Toroidal LHC ApparatuS [34] is a largest particle detector build. It is a general purpose
 346 detector, that is used to study QCD and a Standard Model predictions and searches for a new
 347 physics. A detailed description of this detector is given in Sec. 4.2.

348 **CMS** the Compact Muon [35] Solenoid is an another multiple purpose detector at LHC, build with
 349 the different technologies in the respect to ATLAS.

350 **LHCb** the Large Hadron Collider [36] beauty is specialized for measurement of heavy (charm and
 351 bottom) quark properties, that allows to study the parameters of CP violation.

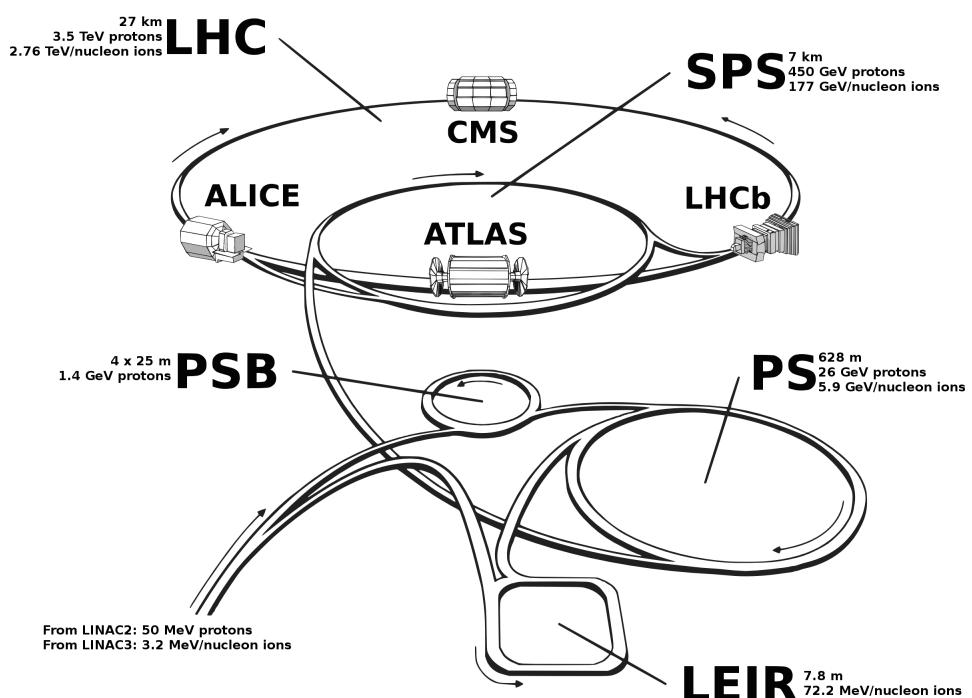


Fig. 4.1: The LHC acceleration complex [37]

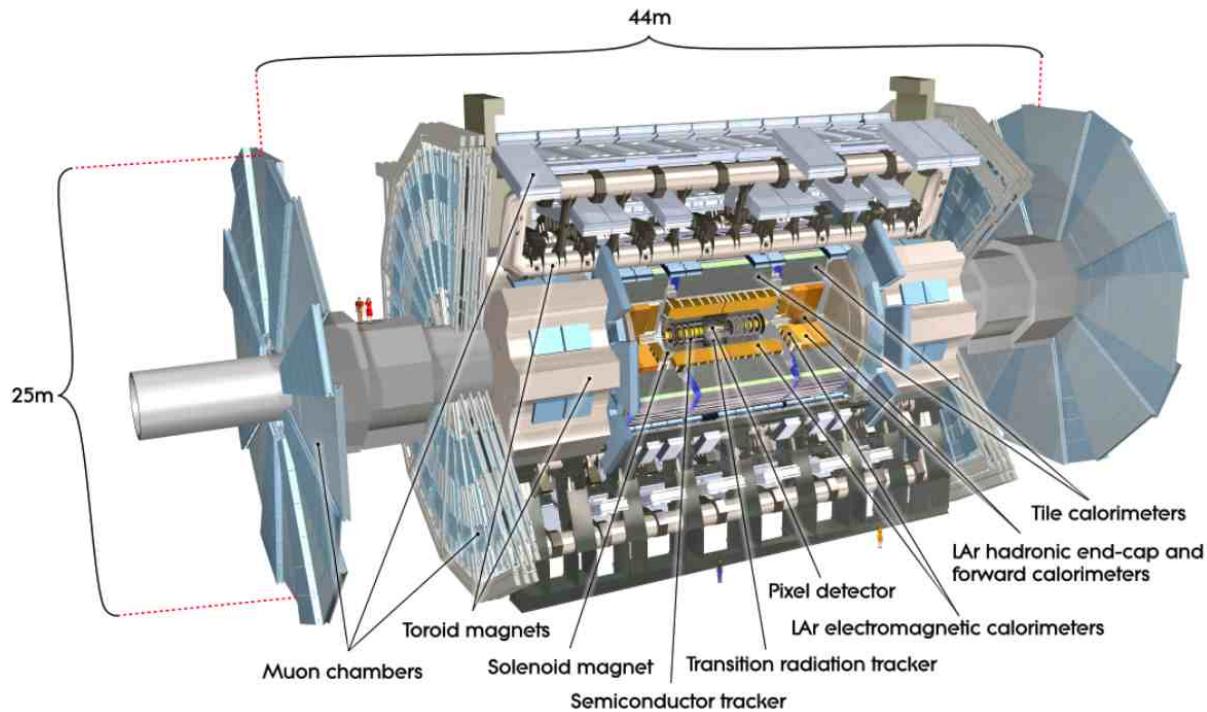


Fig. 4.2: The ATLAS detector [34]

4.2 The ATLAS experiment

ATLAS is a multi-purpose detector. It is used to perform different types of analysis. It is supposed to operate for 2 decades.

The physics goals and signatures of the particles are putting a strict set of requirements on ATLAS detector. Heavy particles, born in pp collisions are expected to decay almost immediately. Thus their properties are measured only indirectly through their decay products, that should be detected and identified. Good identification efficiency of photons, electrons and hadrons is achieved by the system of calorimeters. Because muons are minimum ionizing particles, they go further, than most of the particles, the special outer detector system is build. B-hadrons travel to the distance around 1 mm, producing a secondary vertex, so the vertex detectors near interaction point should be able to detect this. Neutrino from the decays are leaving the detector without interacting with it. However, it is possible to measure them through the energy imbalance in the detector. This means, that detector should cover a hermetically closed area near interaction point, so that no information is missed. Radiation hardness for all of the detectors and electronic is required because of the LHC harsh environment.

The ATLAS detector can be divided into 3 main subdetectors, each serving its own purpose

- The Inner Detector (ID), that is used for tracking and precise measurement of particles momentum
- The Calorimetry system that is used to measure the energies of electrons, photons and hadrons and to identify them
- Muon system designed to detect muons and measure their parameters.

373 Each subdetector can be divided into 3 parts: the barrel region near the interaction point and
374 end-cap in the forward directions. The magnet system is used for tracking of charged particles
375 for measurements of momentum and charge. The ATLAS detector has 4 large superconducting
376 magnet: solenoid, build around inner detector and 3 toroids (end-cap and forward) used for the
377 muon spectrometer.

378 4.2.1 Coordinates and kinematic variables

379 The detector shape motivates the choice of the coordinate system. It is natural to choose z axis
380 to be aligned with the beam, with the start in interaction point, while leaving x and y axis to be
381 perpendicular to it. Because of detector symmetry along the beam z axis, the cylindrical coordinates
382 are often used, with the radial distance $r = \sqrt{x^2 + y^2}$ and the polar θ and azimuth ϕ angles.

 The direction of the particles can be quantified via rapidity:

$$y = \ln \sqrt{\frac{E + p_z}{E - p_z}}, \quad (4.1)$$

where E is the energy of the particle and p_z is z component of the momentum. In the limit of the vanishing masses, this quantity is converges to another widely used variable called pseudorapidity:

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right]. \quad (4.2)$$

383 It is preferred over the polar angle, because the differences in pseudorapidity is the Lorentz invariant
384 along the boost in the beam direction.

 The spacial distance between two Lorentz vectors is defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}. \quad (4.3)$$

 The transverse momentum is defined as:

$$p_T = \sqrt{p_x^2 + p_y^2}, \quad (4.4)$$

385 where p_x and p_y are the x and y components of the particle momentum. Because the incoming
386 protons are aligned along the z axis, the total transverse momentum (together with non-interacting
387 particles) in the detector should be fixed to 0.

388 4.2.2 Inner detector

389 The Inner Detector (ID) is the closest to the interaction point detector system. It used for recon-
390 struction of charged particles tracks and vertexes. Approximately 1000 particles are emerging during
391 one collision within Inner Detector (ID) acceptance ($|\eta| < 2.5$). In order to achieve a good momentum
392 and verex resolution it is required to use a high granularity detector. The layout of ID is shown in
393 Fig. 4.3.

394 The inner detector consists of 3 sub-detectors:

- 395 • The precise reconstruction of vertexes with spacial resolution of 1 mm achieved because
396 of the pixel detectors, placed close to the interaction point. The pixel detector consists of
397 approximately 80.4 million readout channel placed in 3 barrel and 3 disk layers at the end of

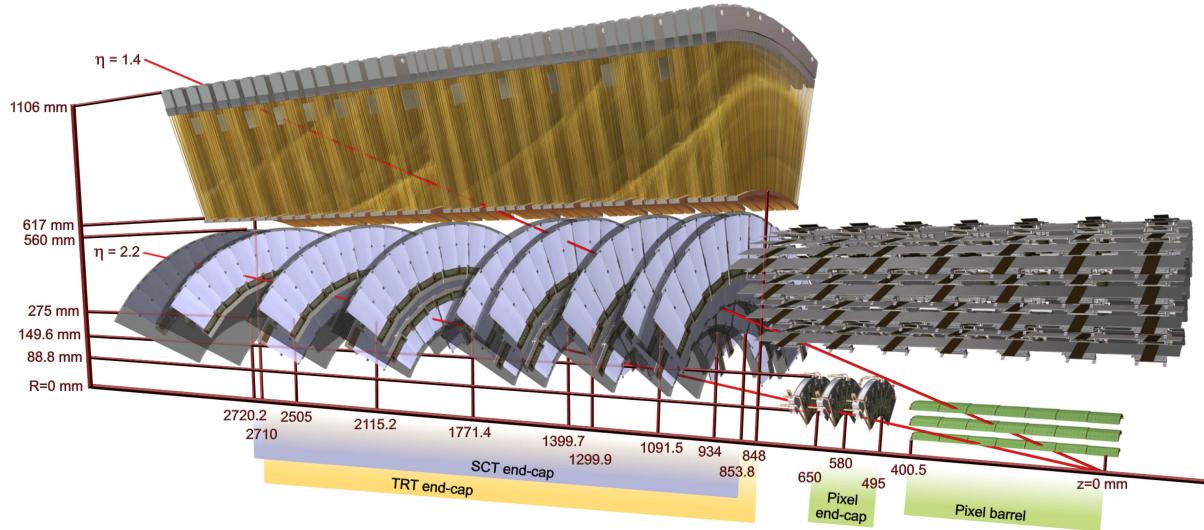


Fig. 4.3: Cut-away view of the ATLAS Inner Detector. Drawing is showing the sensors and structural elements traversed by two charged tracks of 10GeV pT in the end-cap inner detector [38]

each barrel region. Each pixel module made of the silicon layer and s lower layer of electronics. Charged particle, passing through the module, creates a movement of electron-hole pair, that causes a current in the readout electronics.

- The silicon strip detector (SCT) consists of 6.3 million readout channel. It gives a significant contribution on a measurement of charged particle momentum, because of the large number of hits particle produce in the detector. It works in a similar way with pixel detector, however has a smaller precision. Each track crosses around 8 strip layers.
- Transition radiation tracker (TRT) consists of the straw tubes and provides the largest number of hits (~ 36 per track). Each straw is a polimide drift tube 4 mm in a diameter. The tubes are placed parallel to the beam in the barrel region and radially in the wheels.

The combination of the precision measurements near the interaction point and big amount of hits at larger distances allows atlas to have a good precision of the coordinates measurements.

4.2.3 Calorimeters

The calorimeter system is used to measure position and energy of particles from their deposits in the material. The general structure of ATLAS calorimeter is shown in Fig. 4.4. The calorimetric system consists of barrel ($|\eta| < 3.2$) and two end-cap parts ($3.1 < |\eta| < 4.9$). The central part used of the high precision measurements, while the end-cap part has a coarser granularity and mostly used for jet reconstruction and E_T^{miss} measurements.

Particles, entering the calorimeter, produce a cascade of the secondary particles called a particle shower. Each shower is registered by the set of smallest structures of calorimeter providing the response, called cells. Cell structure, called shower shape, differs for different types of particles and used for identification. Each calorimeter also has a dead material, that is used to absorb particles and does not produce any signal response. The full particle energy is reconstructed from the ratio between amount of energy absorbed in dead and active material.

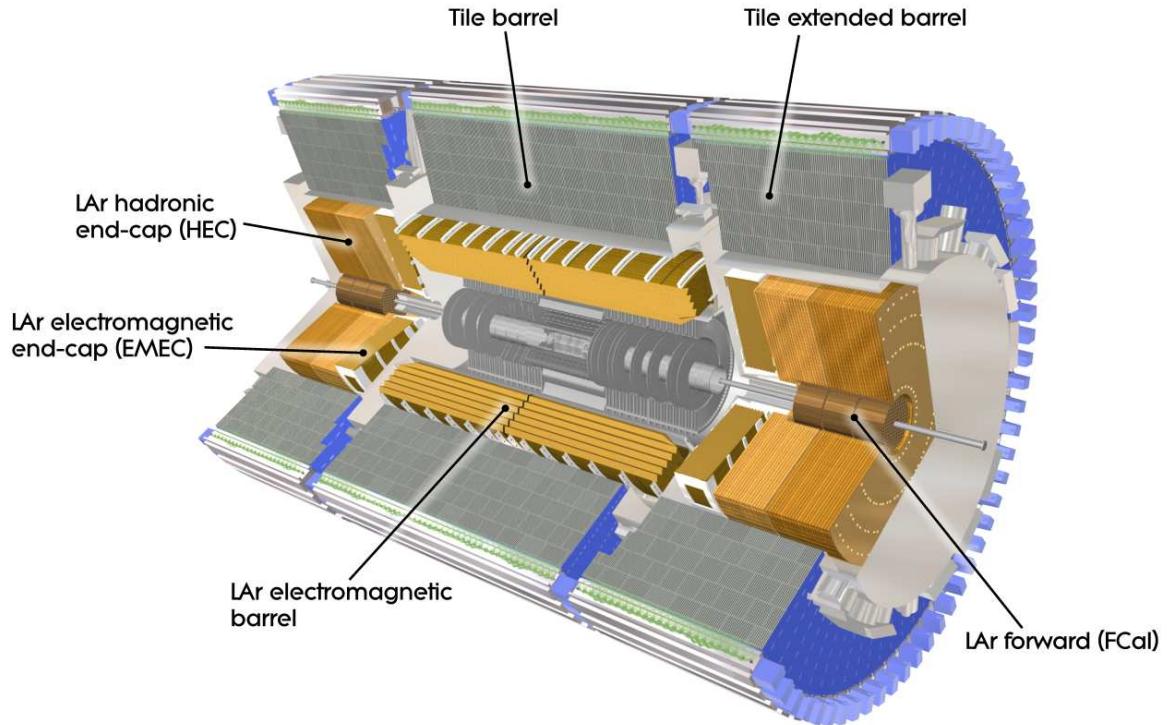


Fig. 4.4: Cut-away view of the ATLAS calorimeter system [38]

422 In order to measure the energy of the particle properly, shower should be fully contained by the
 423 calorimeter. Since depth of the shower, caused by electromagnetic particle is significantly smaller,
 424 than depth of the hadronic shower, calorimeters are divided into two types: electromagnetic (EM)
 425 and hadronic. EM calorimeters are placed closer to the interaction point and has smaller amount of
 426 the dead material, compared to the hadronic calorimeters.

427 The central region calorimeters are required to have high granularity for combination with ID
 428 information and a precision measurement of photons and electrons. Forward part has aa smaller
 429 granularity and mostly used for jet reconstruction and missing transverse energy measurements.

430 **Electromagnetic calorimeter**

431 The main purpose of electromagnetic calorimeter is to measure energies of electrons and photons.
 432 The EM showers starts from initial high-energy electron and photon entering the calorimeter. High-
 433 energy photons are loosing their energy via production of electron-positron pairs, while electrons
 434 and positrons are emitting photons via Bremsstrallung. These two processes continues till the photon
 435 reaches the pair production threshold.

436 The EM calorimeter consists of a barrel part (EM barrel = EMB) and two symmetric end-caps (EM
 437 end-cap = EMEC), that cover a range of pseudorapidity $|\eta| < 1.475$ and $1.5 < |\eta| < 3.2$ respectively.
 438 These calorimeters have an accordion stucture, as shown in Fig. 4.5. This geometry allows to have a
 439 full coverage in ϕ coordinate. It consists of the layers of lead/steel interplaced with liquid argon, that
 440 acts as a sensitive material, and electronics.

441 There are four samplings in EMB calorimeter, as showed in Fig. 4.5:

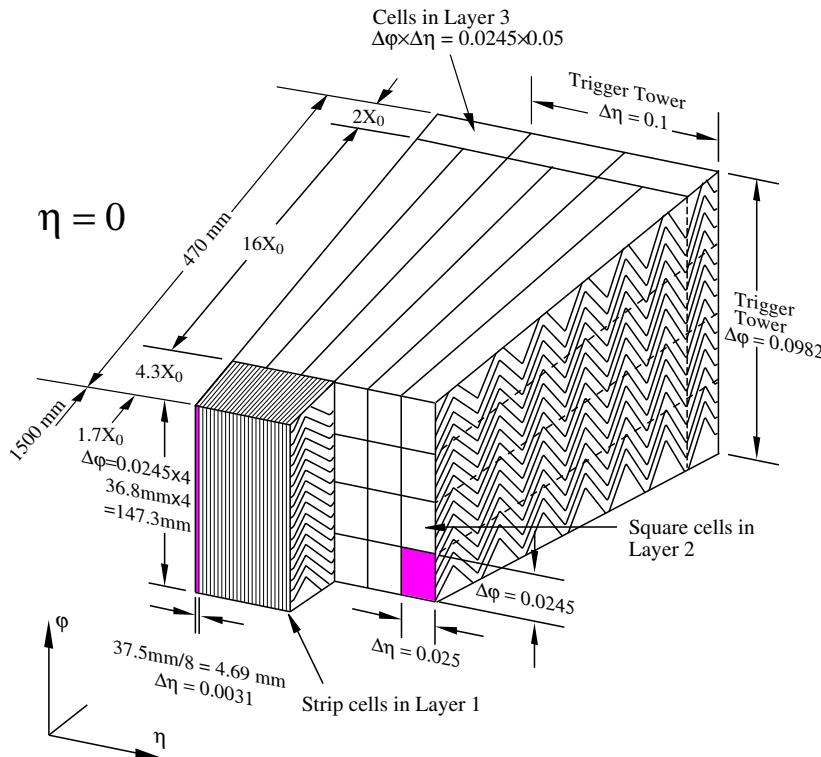


Fig. 4.5: The sketch of EMB module. The granularity in η and ϕ of the cells of each of the three layers and of the trigger towers is also shown [38]

442 **presampler** A single layer of LAr without dead material in front of it. It allows to correct for the
443 energy losses in front of the calorimeter.

444 **1st sampling** The first layer has a fine segmentation in η with thin η strips with size $\Delta\eta \times \Delta\phi =$
445 0.0031×0.098 . Because of the resolution this layer provides an information for γ and π^0
446 separation.

447 **2nd sampling** The majority of the energy is deposited in the second sampling layer. It consists of
448 the square cells with size $\Delta\eta \times \Delta\phi = 0.0245 \times 0.0245$.

449 **3rd sampling** Just the highest energy electrons are reaching the third layer, therefore it has a bigger
450 cell size.

451 Each wheel of the EMEC calorimeter consists of the 2 co-axial wheels: Inner Wheel (IW) and Outer
452 Wheel (OW). Each endcap wheel is divided into 8 wedge-shaped modules. In the central region $1.5 <$
453 $|\eta| < 2.5$, the EMEC calorimeter consists of 3 layers with granularity $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$.

454 **Hadronic calorimeters**

455 The mechanism of hadronic shower development differs from the electromagnetic one. The main
456 physical processes, that are determining the shower development are: hadron production, nuclear
457 deexcitation and pion and muon decays. It usually takes longer to develop a hadronic shower, than
458 an EM one.

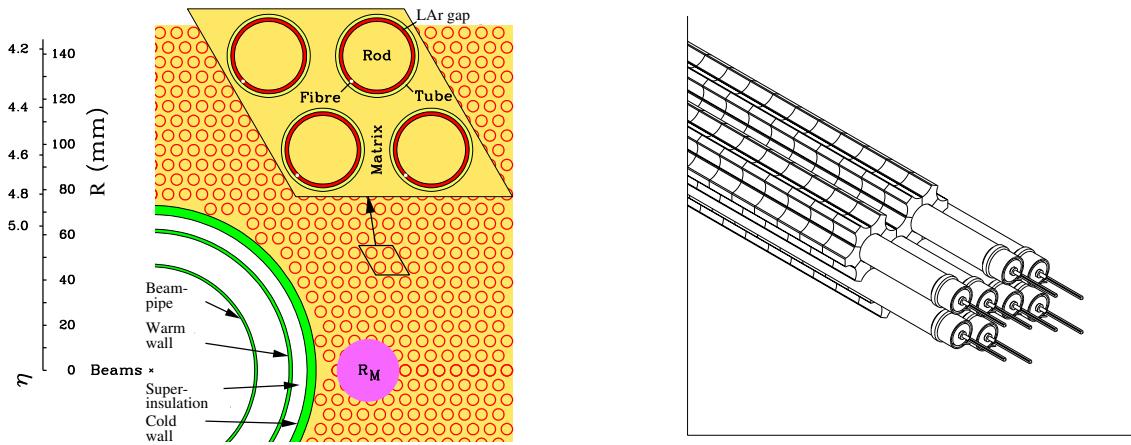


Fig. 4.6: Electrode structure of FCAL1 with the matrix of copper plates and copper tubes and rods with the LAr gap for electrodes [38]

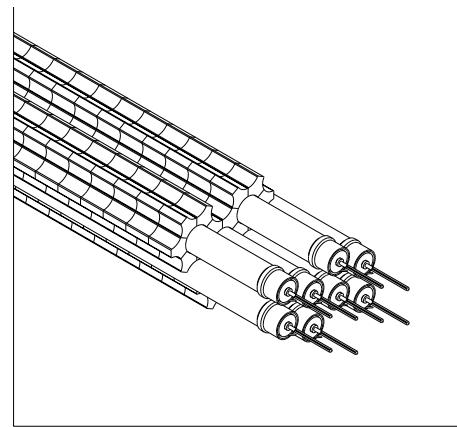


Fig. 4.7: View of the FCAL hadronic module absorber matrix, including a set of tungsten rods and copper tubes surrounded by 1 cm long tungsten slugs [38]

459 The ATLAS hadronic calorimeter consists of the tile and liquid argon Hadronic End-cap Calorimeter
460 (HEC). The forward part of the hadronic calorimeter will be discussed separately.

461 The tile calorimeter placed right after the EMEC and covers a pseudorapidity range up to $|\eta| < 1.0$ in
462 the barrel region and $0.8 < |\eta| < 1.7$ in the 2 end-caps. It is a sampling calorimeter with steel acting as
463 a dead material and scintillator ties for a sensitive material. The readout from scintillator performed
464 using the wavelength shifting fibers. The readout cells are build by grouping the fibers into the
465 photomultiplier. Granularity of the detector

466 The HEC calorimeter uses a liquid argon as a sensitive material and shares the same LAr cryostat
467 with EMEC. The copper-plate are acting as an absorbers with a flat-plate design. The size of the cell
468 in HEC is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ for $|\eta| < 2.5$ and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for forward region $|\eta| > 2.5$

469 Forward calorimeter

470 The forward calorimeter (FCAL) is placed in the same cryostat with EMEC and covers the range of
471 pseudorapidity $3.1 < |\eta| < 4.9$. It is placed 4.7 m away from the interaction point and imposed for a
472 high fluxed. This motivates the choice of detector design with small amount of the sensitive material.
473 The FCAL module consists of the co-axial copper rod and anode tube, separated by the wire around
474 the rod. The LAr fills the gap between rod and anode. Small size of the gaps allows to have faster
475 signal transfer together with avoiding signal degradation caused by distortion of the electric field in the
476 gap (some reference). The structure of one of the FCAL calorimeters is shown in Fig. 4.6.

477 The FCAL is divided into 3 modules: one electromagnetic FCAL1 and 2 hadronic FCAL2 and FCAL3.
478 Parameters of these modules are summarized in Tab. 4.1. In hadronic modules it was decided to use
479 a tungsten instead of the copper in order to optimize the high absorption length. These modules
480 a similar to the FCAL1, except for the use of tungsten rods instead of the copper rods. The space
481 between the end-plates and tubes is filled with tungsten slugs, as shown in Fig. 4.7.

482 The electrodes are forming the readout cells from the group of four, six and nine for FCAL1, FCAL2
483 and FCAL3 respectively. The granularity of the detector is about $\Delta\eta \times \Delta\phi \approx 0.2 \times 0.2$

Table 4.1: Table of parameters for the three FCAL modules

| Calorimeter | Type | Absorber | Gap width (μm) | Number of electrodes | Number of readout channels |
|-------------|-----------------|----------|-----------------------|----------------------|----------------------------|
| FCAL1 | electromagnetic | coper | 250 | 12 260 | 1008 |
| FCAL2 | Hadronic | tungsten | 375 | 10 200 | 500 |
| FCAL2 | Hadronic | tungsten | 500 | 8 224 | 254 |

4.2.4 Muon spectrometer

The muon trajectories are already measured in the ID, however for a high P_T muons it could be difficult to make a precise determination of charge and momentum. The Muon Spectrometer (MS) provides an information at a much larger scales to measure the bending of the trajectory because of the magnetic field. The MS is placed in the most outer part of the ATLAS detector, behind the calorimeters. The amount of the material in front of the MS is adjusted so, that it can be assumed, that all of the particles entering the it are muons.

The muon spectrometer covers the area up to $|\eta| < 2.7$ and allows to trigger on these particles in the range $|\eta| < 2.4$. The precision tracking is performed by the Monitored Drift Tubes (MDT). The MDT consists of 8 layers of the drift tubes and allowing to have a resolution of $80 \mu m$ per tube or $35 \mu m$ per chamber. In addition, in forward region $2.0 < |\eta| < 2.7$, the Cathode-Strip Chambers (CSC) are used. The CSC are the multiwire proportional chambers and giving a resolution $40 \mu m$ in the bending plane and 5 mm in the transverse plane.

The trigger system in muon spectrometer is composed from the fast detectors namely Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC) in the barrel($|\eta| < 1.05$) and end-cap ($1.05 < |\eta| < 2.4$) regions respectively.

4.2.5 Trigger system

Processing and storing of the events is a difficult task at LHC. The collision frequency at LHC is 40 MHz. The full readout information of 1 second of operation requires 1 Tb of storage space. However, the events of interest (such as production of bosons) makes just a small fraction out of these events. The trigger system used for reducing the information stored, while leaving "interesting" events untouched. It covers a range of pseudorapidity up to 2.5.

The trigger system can be divided into 3 levels of selection:

Level-1 The first level trigger should have a high operation speed, so it uses reduced-granularity information from Resistive Plate Chambers (RPC) and Thin-Gap Chambers (TGC) and calorimeter systems. It searches for leptonic and hadronic signatures (or large total transverse energy) in the detector. This trigger allows to reduce a rate, that can be handled by a readout electronics (75 kHz).

Level-2 The second level trigger analyses in more details Regions-of-Interest (RoI's) identified by Level-1 trigger. It uses an information on RoI's such as energy and a position of clusters to further reduce the data transferred. The Level-2 event rate is below 3.5 kHz, with processing time around 30 ms in average.

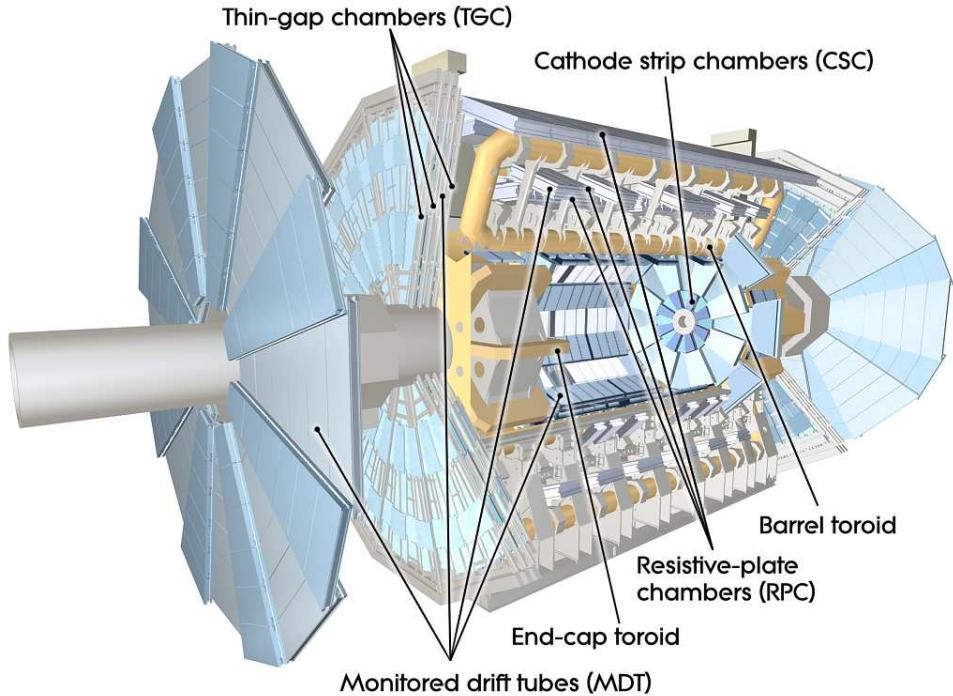


Fig. 4.8: Cut-away view of the ATLAS muon system. [38]

516 **High-Level Trigger (HLT)** The last level selection is performed offline on large farm of CPUs. It
 517 analyses a full information from detectors to refine a trigger selections. The additional in-
 518 formation from tracking allows to improve particle identification and distinguish electrons and
 519 photons. About 200 events per second are left after the HLT and transmitted to the permanent
 520 storage.

521 The data acquisition system (DAQ) receives information from the readout electronics at the L1
 522 trigger rate and transfers the data to L2. After passing the L2 selection criteria the event is built
 523 and transmitted to HLT.

524 **4.3 Luminosity measurement**

One of the main components, characterizing the collider is the instantaneous luminosity $\mathcal{L}(t)$ delivered, that is defined as a proportional factor between the cross-section σ_p and number of interactions per second $\frac{dR}{dt}$, as :

$$\frac{dR}{dt} = \mathcal{L}(t) \times \sigma_p (cm^{-2}s^{-1}). \quad (4.5)$$

525 This value is a relativistic invariant and independent on physical reaction.

In case of LHC, that performs head-on collisions of particle bunches, it could be calculated as per beam value:

$$\mathcal{L} = \frac{N_p^2 N_b f_{rev}}{4\pi\sigma_x\sigma_y} F, \quad (4.6)$$

526 where N_p is the number of protons per beam, N_b - number of bunches, f_{rev} is the revolution
 527 frequency, σ_x and σ_y are the horizontal and vertical beam profile widths. The factor F is coming from
 528 the beam crossing angle. In 2012, at 8 TeV centre-of-mass energy, the LHC machine was able to
 529 reach an instantaneous luminosity $7.7 \times 10^{33} [\text{cm}^{-2}\text{s}^{-1}]$.

Luminosity can be also measured from the interaction rate μ_{vis} in a detector of some process as:

$$L = \frac{\mu_{\text{vis}} n_b f_r}{\sigma_{\text{vis}}}, \quad (4.7)$$

530 where σ_{vis} is the visible interaction cross-section.

531 The ATLAS experiment uses several detectors to measure the recorded luminosity. The Beam
 532 Condition Monitor (BCM) monitors a beam parameters close to the interaction point and allows to
 533 measure bunch intensities. In the forward region a special detector for a luminosity measurements
 534 is placed: the LUCID (LUminosity measurement using Cerenkov Integrating Detector) detector, that
 535 detects the inelastic scattering. The beam profile and the visible interaction cross-section σ_{vis} are
 536 measured during so-called van-der-Meer scans [39].

Chapter 5

538 Event reconstruction

| | | |
|-----|--|----|
| 541 | 5.1 Tracks and vertexes | 37 |
| 542 | 5.2 Electron reconstruction and identification | 38 |
| 543 | 5.2.1 Central electrons reconstruction | 38 |
| 544 | 5.2.2 Forward electrons reconstruction | 38 |
| 545 | 5.2.3 Electron identification | 39 |
| 546 | 5.3 Muon reconstruction and identification | 39 |
| 547 | 5.4 Missing transverse energy reconstruction | 40 |
| 548 | 5.4.1 Standard missing transverse energy reconstruction | 40 |
| 549 | 5.4.2 Reconstruction of missing transverse energy from hadronic recoil | 41 |

553 The raw detector information could not be used in the physics analysis. It is required to have a
 554 separate process of interpretation of electronics signals, called reconstruction. It is used to determine
 555 the particles, born in the collision, their momentia, direction and vertexes, they are coming from. In
 556 this chapter the reconstruction and identification of the objects at ATLAS experiment, used in the
 557 analysis will be described.

558 It should be noted, that the missing transverse energy reconstruction stands separately, since
 559 the standard procedure was not applicable for 2.76 TeV data and a different approach have been
 560 adapted.

561 **5.1 Tracks and vertexes**

562 The tracks and vertexes are reconstructed from the ID information [40]. The reconstruction can
 563 be divided into 2 steps. On the first step the inside-out algorithm is used for the pixel and silicon
 564 detector. Tracks are reconstructed from the random 3-points seed in the detector and then adds
 565 the new points, while moving to the interaction point using the Kalman filter [41]. The Kalman
 566 filter is the iterative algorithm that provides the best estimate based on projection of previous and
 567 current measurement. Ambiguities in the track candidates are resolved and tracks are extended in TRT.

568 At the second step, algorithm searches for the segments, reconstructed in TRT and then extends
 569 them into the silicon detector. The tracks in TRT with no silicon extension are referred as TRT-
 570 standalone tracks.

571 Vertexes are reconstructed using the iterative vertex finding algorithm. The vertex starts from
 572 the z-position at the beamline of one of the tracks. The χ^2 fit is performed on that seed and

573 nearby tracks. The tracks, that are displaced for more, than 7σ are treated as a separate vertex.
574 The procedure is repeated till no new vertexes are found. During the reconstruction vertexes are
575 required to contain at least 2 tracks, but the requirement of the 3 tracks could give more robustness.
576 From the vertex candidates, the vertex with highest sum of the transverse momenta of the outgoing
577 tracks is defined as a primary vertex.

578 **5.2 Electron reconstruction and identification**

579 Due to the detector design electrons are divided into the 2 groups: central and forward. For the
580 central electrons ($|\eta| < 2.5$) there is a tracking information available. Presence of the ID track allows
581 to perform the precise reconstruction and identification. On another hand, for the forward ($|\eta| > 2.5$)
582 electrons could be reconstructed using just the calorimeter information, so the different algorithm
583 is used. In this section the identification of the central electrons and reconstruction for both central
584 and forward electrons will be discussed.

585 **5.2.1 Central electrons reconstruction**

586 The central electron reconstruction starts from the clusters in the EM calorimeter. On the first step
587 the calorimeter is divided by the grid with the cell size $\Delta\eta \times \Delta\phi = 0.25 \times 0.25$. The EM calorimeter
588 clusters are formed from the cells with total transverse energy in all layers above 2.5 GeV using the
589 sliding window algorithm [42] with size 3×5 cells. The position of the cluster is determined from its
590 barycenter.

591 On the second step track with $P_T > 0.5$ GeV are extrapolated to the middle layer of the EM
592 calorimeter. A track and a cluster are considered matched to each other if the distance between
593 track and cluster is $|\Delta\eta| < 0.5$ GeV. In order to take into account effect of the bremsstrahlung losses
594 the azimuthal distance is allowed within $\Delta\phi < 0.1$ on the side where the extrapolated track bends as
595 it traverses the solenoidal magnetic field.

596 An electron is considered to be reconstructed if at least 1 track is matched to EM cluster. In case if
597 there are several tracks passing the requirements, the tracks with silicon hits are given the priority,
598 an the match with the smallest distance $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. In case if there is no track matched,
599 the cluster is treated as a photon candidate.

600 After the track matching the cluster size is optimised. The cluster size is enlarged to 3×7 and
601 5×5 in barrel and end-cap EM respectively. The total reconstructed electron energy is determined
602 from the corrected cluster energy, estimated energy deposit in the material in front of EM and
603 the estimated energy deposits outside of the cluster and calorimeter. The absolute energy scale
604 determination is described in Sec.??

605 **5.2.2 Forward electrons reconstruction**

606 Since there is no tracking in the forward region ($2.5 < |\eta| < 4.9$), the electron could be reconstructed
607 using just the information from EMEC and FCAL detectors. Opposite to the central electrons with
608 fixed size of the cell, for forward electrons the topological clustering algorithm is used [43]. The
609 main principle of this algorithm is that cells with energy higher than expected noise, are merged
610 together iteratively. The average noise in the cell is obtained in the calibration runs and includes a
611 contribution from pile-up. The cluster starts from the cell with significant energy and then expanded
612 by neighborhood cells. If 2 clusters are sharing 1 neighborhood cell, they are merged together. The

613 threshold, defined as $t = \frac{E_{cell}}{E_{noise}}$, is 4 and 2 for the starting the cell and expanding neighborhood
614 respectively.

615 The energy of the electrons is defined as the sum of the energies, taking into account for energy
616 losses in passive material in front of calorimeter. The direction of the electron is defined as the
617 barycentre of the cluster cells.

618 5.2.3 Electron identification

619 The application of additional criteria on reconstructed electrons allows to get better purity of the
620 sample and exlude objects, that can be misidentified as electrons, such as: jets and electrons from
621 the photon conversion.

622 The identification of the central electron is based on a sequential cuts, on calorimeter infromation,
623 tracking and combined variables. There are 3 sets of selection criteria, used for a physics analyses,
624 designed in hierarchical way at to provide incresied background rejection with cost of decreasing
625 identification efficiency. They are:

626 **Loose** The loose identification criteria uses the shower shape variables in first and second layer of
627 EM calorimeter and the fraction of energy, deposited in hadronic calorimeter. There are also
628 additional requirements on electron track and track-cluster matching.

629 **Medium** The medium selection is made out of loose identification with adding the information from
630 the 3-rd level of EM calorimeter, transverse impact parameter d_0 and TRT (to reject charged
631 hadron background) if available. Additionally the measured hit in the innermost layer of pixel
632 is required to discriminate against the photon conversions. These requirements are allowing to
633 increase the background-rejection power by an order of magnitude, compared to loose.

634 **Tight** The tight selection uses the full information of the particle identification tools available. In
635 addition to medium criteria, it puts stricter requirements on track quality, on ration of EM
636 cluster energy to the track momentum and veto on reconstructed photon conversion vertices
637 assosiated with the cluster. The overall strength of background rejection is 2 times higher, that
638 for a medium selection.

639 It should be noted, that neither of these criteria requires no additional tracks close to the identi-
640 fied electrons. The optimization of these requirements (called isolation requirements) is left for the
641 dedicated analysis.

642 5.3 Muon reconstruction and identification

643 The ATLAS experiment uses the information from ID and muon spectrometer for a precise re-
644 construction of the muons. Energy measurments in calorimeter can also be used for the muon
645 identification. The muons, based on the information, available from these detectors, can be divided
646 into different types:

647 **Combined (CB)** Muons with track both in ID and MS, that could be matched to each other. This
648 is the main type of the muons.

649 **Segment-tagged (ST)** Muons with track in the ID and at least one local track segment in the MDT
650 or CSC chambers. This type of muons could be used for the small P_T muons or in the reduced
651 MS acceptance region.

652 **Stand-Alone (SA)** These are the muons, that are crossing at least 2 layers of MS chambers, but
653 have no reconstructed track in the ID. The parameters of the track are determined using the
654 extrapolation to the primary vertex, taking into account the estimated energy loss in the
655 detector in front of MS. These muons are mainly used to extend the acceptance up to $|\eta| < 2.7$,
656 where there is no ID information.

657 **Calorimeter-tagged (CaloTag)** Muons, that have a track in the calorimeter, that can be associated
658 with the minimum ionizing particle.

659 The ID track, used in muon reconstruction, should satisfy additional requirements:

- 660 • at least 1 pixel hit
- 661 • at least 2 SCT hits
- 662 • at most 2 active pixel or SCT hits, that are transversed by the track, but have no hit.
- 663 • in the region of full TRT acceptance ($0.1 < |\eta| < 1.9$) at least 9 TRT hits.

664 The muons are reconstructed in MS in two steps: first the local segments within one layer are
665 combined and then the segments are combined in a full track. The reconstruction of
666 the MS and combined ID-MS track can be done using one of the two independent reconstruction
667 procedures [44], called Staco and Muid.

668 The Muid algorithm performs full track refit using the parameters from ID and MS [45]. For the
669 staco algorithm the reconstruction of the track in MS starts from the segment from the outers
670 station. The segments from middle and inner layers are iteratively added till the full track is ob-
671 tained. The matching between ID and MS sub-detectors performed via statistical combination of the
672 parameters in ID and MS using the corresponding covariance matrices [46]. Staco algorithm is the
673 algorithm, used in this analysis.

674 5.4 Missing transverse energy reconstruction

675 ATLAS detector has almost 4π coverage. This allows to calculate imbalance of energies inside
676 calorimeter, especially transversal part of it called E_T^{miss} . In W-analyses E_T^{miss} is used as a proxy
677 for neutrino from a $W \rightarrow l\nu$ decay. It leaves detector without interacting with it and that causes
678 large energy imbalance in an eventA. In this section two methods of E_T^{miss} reconstruction and the
679 reasons for using non-standard one will be discussed.

680 5.4.1 Standard missing transverse energy reconstruction

Standard reconstruction of E_T^{miss} at ATLAS experiment [47] uses transverse energy deposits in the calorimeter, energy losses in cryostat and reconstructed muons for a calculation:

$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{miss,calo}} + E_{x(y)}^{\text{miss,cryo}} + E_{x(y)}^{\text{miss,muon}}. \quad (5.1)$$

The calorimeter term is using information from reconstructed physics objects for calibration of the cell response. The total transverse energy in calorimeter is defined as:

$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{miss},e} + E_{x(y)}^{\text{miss},\gamma} + E_{x(y)}^{\text{miss},\tau} + E_{x(y)}^{\text{miss,jets}} + E_{x(y)}^{\text{miss,SoftTerm}} + E_{x(y)}^{\text{miss},\mu}. \quad (5.2)$$

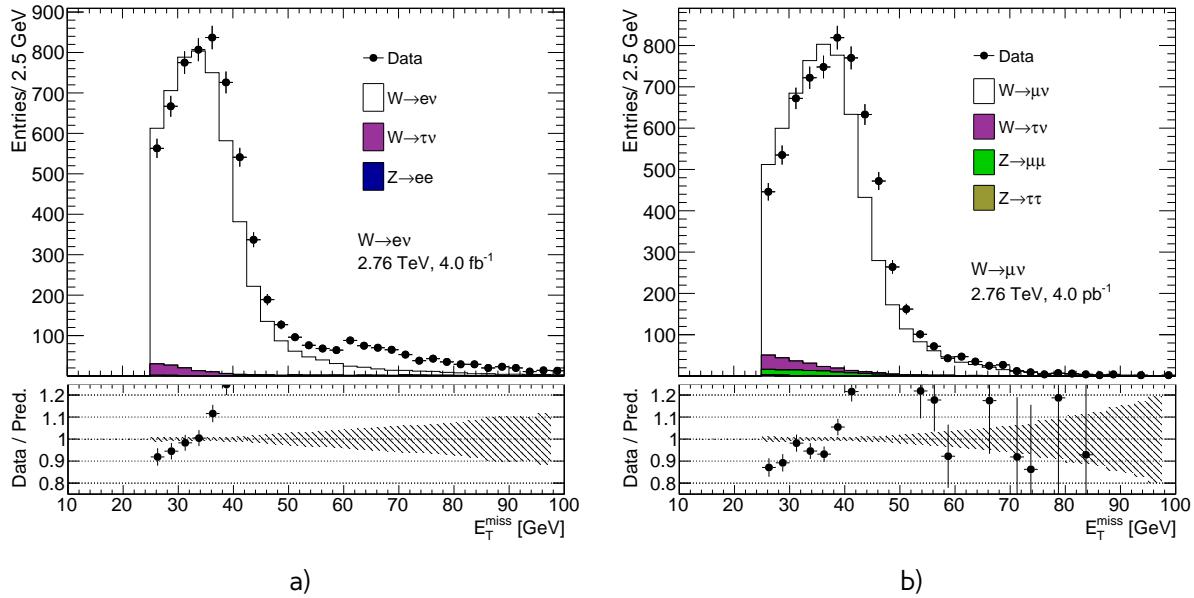


Fig. 5.1: Missing transverse energy distribution for a) the $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection from Chap. 9. E_T^{miss} calculated using the standard ATLAS algorithm. The expected contributions from all backgrounds are estimated with Monte Carlo simulations, except for QCD background that is not included. All Monte-Carlo corrections from Chap. 10 are applied. There are visible discrepancies between data and MC, that cannot be explained by the contribution of the QCD background, which is expected mainly in the low E_T^{miss} region (Sec. 12.1).

where each term is calculated as a negative sum of the calibrated reconstructed objects, projected onto the x and y directions. Each jet with energy $P_T > 20$ GeV is corrected for a pile-up and a jet energy scale is applied. Soft term is calculated from topoclusters and tracks, that are not associated with high-pt objects. To avoid double counting, muon energy loss in the calorimeter is subtracted from E_T^{miss} . The E_T^{miss} muon term is calculated from the momenta of muons measured in a range of pseudorapidity $|\eta| < 2.7$. Since pileup has a significant effect on the E_T^{miss} performance several methods of pileup suppression are used [48].

The runs at 2.76 TeV are characterized by a low pileup (mean number of interaction per bunch crossing < 1.0), so the usage of a procedure optimized for high pileup 8 TeV runs may not be optimal. It was examined and found out, that there are big discrepancies between the E_T^{miss} distributions for data and MC simulation, as shown in Fig. 5.1, where the missing transverse energy for data is compared to signal and background MC predictions.

The differences are visible in both electron and muon channels and cannot be explained by the (missing on the control plots) contribution from the QCD background, which is expected mainly in the low E_T^{miss} region (see Sec. 12.1).

5.4.2 Reconstruction of missing transverse energy from hadronic recoil

A different way of E_T^{miss} calculation was developed for W and Z decays by the W mass measurements group [50]. This procedure is based on a requirement of a balance in the transverse momentum of

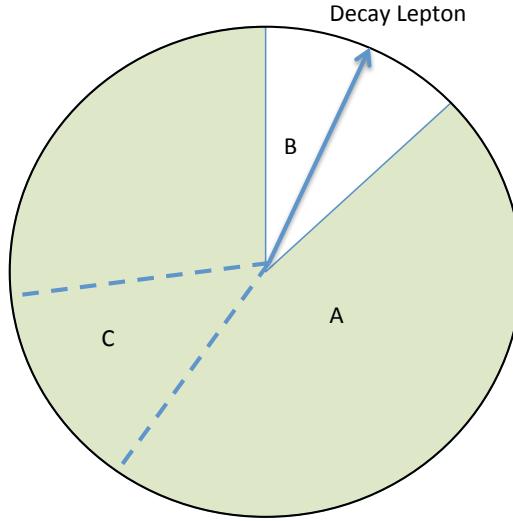


Fig. 5.2: Definition of different zones in the calculation of the cluster-based hadronic recoil. Zone B is excluded from hadronic recoil calculation because it contains decay lepton. To describe properly the overall acitivity it is replaced by the zone C, rotated in the direction of B. Zone A corresponds to the rest of the calorimeter [49].

a W-boson and the initial (quark-gluon) state radiation:

$$\vec{P}_T^W = \vec{P}_T^l + \vec{P}_T^\nu = \sum \vec{P}_T^{ISRquarks,gluon}, \quad (5.3)$$

where $\sum \vec{P}_T^{ISRquarks,gluon}$ is a transverse momentum of partons from the initial state radiation, also called hadronic recoil (HR), \vec{P}_T^l and \vec{P}_T^ν are the transverse momenta of lepton and neutrino respectively. Therefore, E_T^{miss} can be determined as:

$$E_T^{miss} = -P_T^\nu = -HR + P_T^l \quad (5.4)$$

This procedure assumes, that recoil arises from one single leading jet, and the rest is coming from a soft hadronic activity. The hadronic recoil is computed as a vector sum of calorimeter clusters:

$$HR = \sum_{i=0}^{N_{topo}} \vec{p}_T^{topo} \quad (5.5)$$

while a scalar sum of all transverse energy contributions corresponds to the hadronic activity in the event:

$$\sum E_T = \sum_{i=0}^{N_{topo}} E_T^{topo} \quad (5.6)$$

⁶⁹⁷ To avoid double counting of lepton energy losses in the calorimeter, the clusters inside a cone with
⁶⁹⁸ a radius $dR = 0.2$ around the lepton direction are excluded from this calculation. To compensate
⁶⁹⁹ for the subtracted soft activity from the cone, a replacement cone is added (Fig. 5.2). This cone is

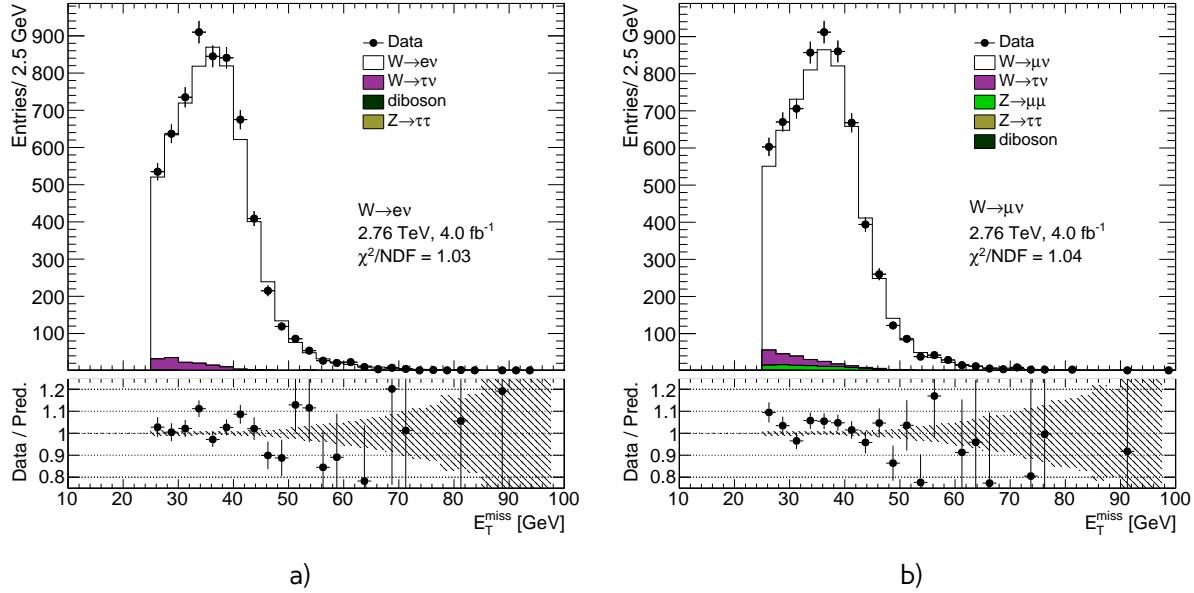


Fig. 5.3: Missing transverse energy distribution for a) the $W \rightarrow ev$ selection and b) the $W \rightarrow \mu\nu$ selection from Chap. 9. E_T^{miss} calculated using the hadronic recoil algorithm. The expected contributions from all backgrounds are estimated with Monte Carlo simulations, except for QCD background that is not included. All Monte-Carlo corrections from Chap. 10 are applied.

defined as a cone at the same pseudorapidity, but at a different ϕ . It should be far from any other lepton and hadronic recoil direction. The cone is then rotated to the original lepton direction. This definition does not take into account the jet reconstruction aspects.

Fig. 5.3 shows the control plots for the distributions of missing transverse energy calculated using the hadronic recoil procedure. In both electron and muon channels the agreement between data and MC simulation is much better than in the case of the standard procedure described in a previous chapter. It was decided to use hadronic recoil E_T^{miss} reconstruction method in 2.76 TeV data analysis.

708

Monte-Carlo simulation

| | | |
|-----|---|----|
| 6.1 | ATLAS chain of Monte-Carlo production | 45 |
| 6.2 | Event generators | 46 |
| 6.3 | Simulation in Geant4 | 48 |

The Monte Carlo (MC) method was invented by scientists working on the atomic bomb in the 1940s. Its core idea is to use random samples of parameters or inputs to explore the behavior of a complex system or process. Nowadays, MC experiments are essential part of research in both theoretical and experimental particle physics. This chapter gives an overview of the ATLAS experiment simulation scheme, the simulation methods and the software used. Also, techniques for fast simulation will be discussed.

723

6.1 ATLAS chain of Monte-Carlo production

724 Monte Carlo method allows to perform different analyses, generate predictions for comparisons with
 725 data, study the the detector or the selection algorithms performance. All of these applications require
 726 accurate MC predictions. The simulation software implements precise physics models and uses
 727 statistics large enough , to exclude statistical uncertainties (usually 5 times more, than expected in a
 728 data). ATLAS simulation software is integrated into Athena framework used at ATLAS experiment [51].

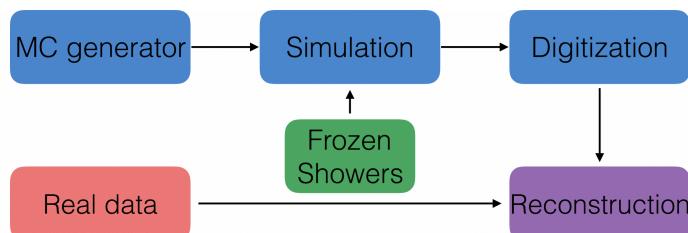


Fig. 6.1: Diagram of the ATLAS MC production chain. Stages in blue are completely related to Monte-Carlo production. The Frozen Showers technique for fast simulation will be explained in Chap. 7. Data sample collection is decribed in Sec. 4.2. Reconstruction is common stage for data and MC and described in Chap. 5. After the reconstruction events are going to the analysis chain.

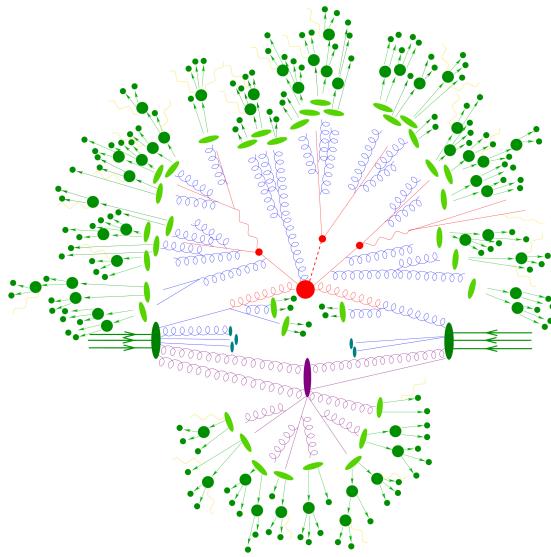


Fig. 6.2: Schematic view of a $t\bar{t}H$ event produced in a pp-collision: the hard scattering is shown as a red blob with the solid and dashed lines as the resulting three particles. Independently happening multi-particle interactions are indicated by the violet blob. Parton showers are shown with curly lines. Hadronization yields hadrons as shown in light green, while the final state particle are dark green. [52]

729 Simulation chain is generally divided into 4 main steps (Figure 6.1):

730 **Event generation** Simulation of hard interaction, parton evolution and hadronisation. This step is
731 independent of the ATLAS detector geometry.

732 **Simulation** Simulation of energy depositions ("hits") which are produced by final state particles.

733 **Digitalization** Simulation of detector response using "hits" information: first inputs to the read out
734 drivers (ROD's), called "digits" are constructed, then ROD functionality is emulated. Detector
735 noise effects are added at this stage.

736 **Reconstruction** Production of the Analysis Object Data (AOD) files, which are containing the in-
737 formation needed for physics analysis. This stage is identical for both data and MC

738 Additionally, the pileup effects are added to MC by overlaying the simulation of the hard interactions
739 with the simulation of soft inelastic scatterings. This scheme allows to use computing resources more
740 efficiently, than with a single-step simulation and simplifies software validation, since it is possible
741 to reuse files from previous stages. In the following sections event generation and simulation will be
742 described in more details.

743 6.2 Event generators

744 The outcome of the hard interaction could be a simple scattering of the hadron elementary con-
745 stituents, their annihilation into new resonances or a combination of two. This can lead to a final
746 state with a large particle multiplicity. The main goal of the event generator is to provide a complete

747 picture of the final state: description of the particle types and momentia on the event-by-event
748 basis. The factorisation theorem [53] allows to make event generation in independent stages, which
749 are dominated by different dynamics. Schematic plan of simulation of a $t\bar{t}H$ event is shown in Figure
750 6.2:

751 **Modelling of hard subprocess** Hard subprocess happens at the smallest times and distances,
752 where the colliding partons are considered free. The process of interest is simulated by selecting
753 production channels and calculating corresponding matrix elements (ME) at a fixed order
754 of the strong coupling constant and including randomly chosen momenta of the incoming par-
755 tons, which are based on the parton distribution functions (PDF). Most of the generators have
756 leading (LO) order or next to leading order (NLO) in α_s .

757 **Parton showering** Quarks and gluons from hard process can radiate secondary quarks and gluons,
758 resulting in dozens of additional partons associated with the event. This process is calculated
759 as step-by-step evolution of momentum transfer scales from highest (hard subprocess), to the
760 lowest (around 1 GeV), where the perturbative calculations are not valid. There is a possibility of
761 double counting between showers and hard subprocess. This can be avoided using matching
762 approaches, for which higher order corrections to ME are integrated with parton showers, or
763 merging strategy, there jet resolution scale is used as a threshold between matrix elements
764 and parton showers.

765 **Hadronisation** Final, stable, color-neutral particles, which can be detected in an experiment, are
766 formed during hadronisation. This occurs at larger nonperturbative scales and is usually im-
767 plemented using different phenomenological models.

768 **Modelling underlying event** Parallel to the main process other collisions of partons can occur.
769 They are called underlying event. These additional interactions can produce partons which
770 contribute to the final state. This is one of the least understood aspects of hadronic collisions.

771 The current analysis uses samples generated with the following generators:

772 Powheg [54] Powheg is Monte-Carlo, which calculates the matrix element (ME) at the NLO level [55],
773 that can be interfaced to other generators (such as Pythia or Herwig) to get higher precision
774 of showering.

775 Pythia [56] Pythia is a general purpose generator for hadronic, hadron-lepton and leptonic collisions.
776 It can model ME, initial and final state showers, hadronisation and decays, underlying event
777 (via multi parton interactions). Pythia contains a library with around 240 processes with LO ME.
778 It uses Lund String model [57] for hadronisation.

779 Herwig [58] Herwig is a LO general purpose event generator for simulation of lepton-lepton, hadron-
780 lepton and hadron-hadron collisions. The main difference between Pythia and Herwig is that
781 Herwig uses angular ordering in the parton showers and models the hadronisation step using
782 the cluster fragmentation

783 Sherpa [59] Sherpa is an event generator, that uses tree-level leading order matrix element for
784 a hard scattering and features its own implementation of parton shower and hadronisation
785 models.

786 Photos [60] Photos is a program used for generation of QED radiative corrections. It is linked to
787 multipurpose generators.

788 Tauola [61] Tauola is a generator, used to describe leptonic and semi-leptonic τ -decays. It is also
789 linked to multipurpose generators.

790 6.3 Simulation in Geant4

791 After event generation, simulation software is used to provide hardware response for final state
792 particles. The main method used by ATLAS experiment, referred to as *Full Simulation*, makes use of
793 the Geant4 [62]. Geant4 is C++ based toolkit for the simulation of the passage of particles through
794 matter. It is used in a wide range of experiments in high energy and nuclear physics.

795 Geant4 can simulate complex detector structures with sensitive detector material and correspond-
796 ing infrastructure. It can also calculate basic properties of materials, like radiation and interaction
797 length. Geant4 stores "hits" information - snapshots of physical interactions. In Geant4 events and
798 particles are simulated separately and each particle is moved in steps. Size of each step is chosen
799 to preserve both CPU performance and required precision.

800 Physics interactions are treated as a set of discrete processes. They could be handled either for
801 particle at rest, or its along step, the maximum value of which depends on physics process, or
802 after it. Geant4 package has different models and approximations for hadronic and electromagnetic
803 processes. Some of them are approximate and computationally fast. It allows to choose a set of
804 the models, called physics list, depending on particular requirements. There are several reference
805 physics lists, that are validated for each new release of Geant4 software. ATLAS experiment uses
806 one of these lists.

807 Most of the computing resources are taken by a mass MC production, required for each data
808 taking periods. Uncertainties of some of Run-I analyses are dominated by available MC statistics. It
809 is possible to improve in CPU usage by tuning physics list or replacing a complex magnetic field maps
810 by a parametrisation. Also there are long-term developments for multi-threading and vectorisation
811 of the code.

812 Run-2 has a higher pileup and luminosity, so even more MC events are needed. This means that
813 fast and accurate simulation approach is essential. During the simulation largest time is spent on
814 calorimeters. This is the motivation for development of fast calorimetry techniques.

815 There are two main methods used at ATLAS:

- 816 • Parametrisation of the calorimeter cells response. Spacial energy response is simulated using
817 longitudinal and lateral energy profiles.
- 818 • Frozen Showers. This technique will be described in more detailed in Chap. 7

820

Frozen Showers

| | | | |
|-----|-------|--|----|
| 821 | 7.1 | Introduction | 49 |
| 822 | 7.2 | Properties of electron showers in FCAL | 50 |
| 823 | 7.3 | Generation and use in simulation | 55 |
| 824 | 7.3.1 | Tuning of libraries- | 55 |
| 825 | 7.4 | Machine learning based bin finding procedure | 57 |
| 826 | 7.4.1 | Machine learning introduction | 57 |
| 827 | 7.4.2 | Electron shower categorization | 59 |
| 828 | 7.4.3 | Interpretation of results | 63 |
| 829 | 7.4.4 | Reconstructed electron energy | 63 |
| 830 | 7.4.5 | Plans for the future | 63 |
| 831 | 7.5 | Validation of the new libraries | 67 |
| 832 | | | |
| 833 | | | |
| 834 | | | |
| 835 | | | |
| 836 | | | |

837 As it was mentioned in Chap. 6, fast simulation techniques are an essential part of the Monte-
 838 Carlo production in the ATLAS experiment. The typical time needed for simulating a $t\bar{t}$ event is
 839 around 1 minute, and most it is spent on the simulation of particle interaction in the calorimeters.
 840 This motivates the development of fast calorimetry techniques, allowing to describe calorimeter
 841 response.

842 In this chapter a Frozen Showers method for the forward calorimeter simulation is described. The
 843 first section gives a small introduction to the method. In the next section properties of the electron
 844 shower in the FCAL are showed. In the Sec. 7.3 the use of method is explained. The Sec. 7.4 introduces
 845 a new method of finding the bin size for a libraries. Finally, in the last section, the validation of Frozen
 846 Showers is showed.

847

7.1 Introduction

848 Frozen showers is currently the main fast calorimeter simulation approach used at ATLAS experiment
 849 [63]. This method uses pre-simulated "frozen" showers. This allows to reduce the time spent on a
 850 simulation of a large amount of low energy sub showers. This method gives a 25% speedup in
 851 simulation. It is required to have in advance generated libraries for each detector and particle used
 852 in this method.

853 For each shower in the library its lateral and transverse size and a list of the all energy deposition
 854 inside the sensitive material (hits) with information about their energy, position and time are stored.
 855 During simulation, if the energy of a secondary electron falls below the cut-off energy it is replaced

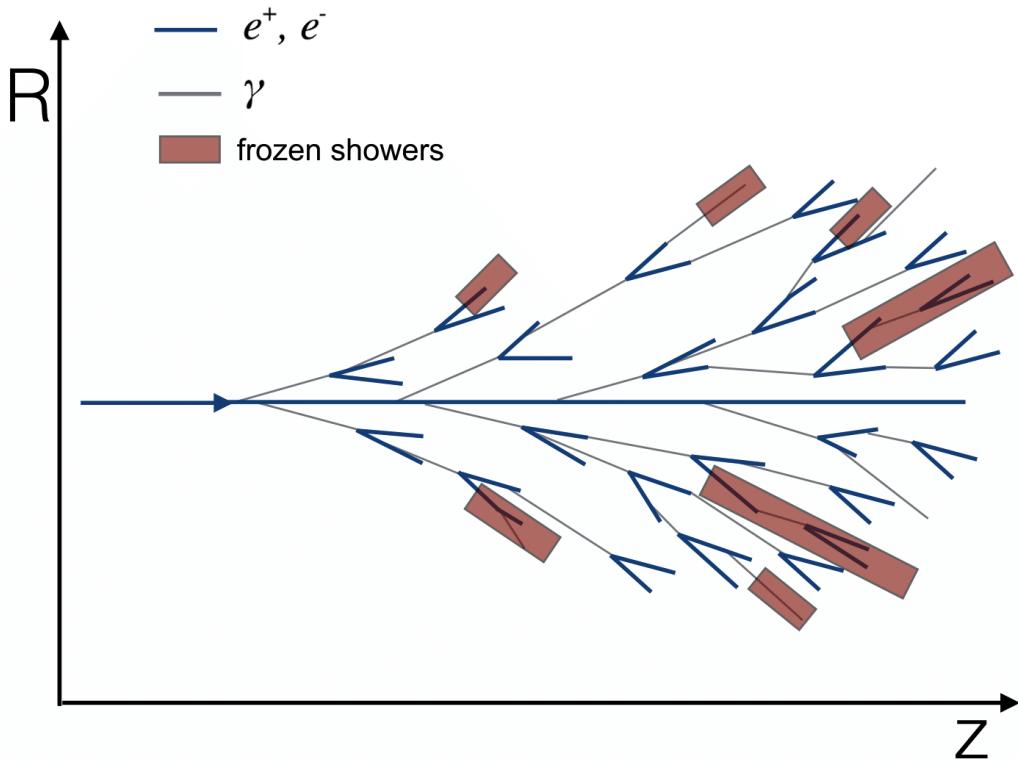


Fig. 7.1: Diagram showing the shower substitution of the low-energy electron, during the high-energy particle simulation. Some of the showers from a particles, substituted by frozen showers method marked by a red squares

Table 7.1: Main parameters used for the frozen shower libraries

| The general frozen showers parameters | |
|---------------------------------------|--|
| Detectors used | FCAL1, FCAL2 |
| Type of the particle | photons, electrons, neutrons |
| Energy cut-off | $E_\gamma < 10 \text{ MeV}$, $E_e < 1000 \text{ MeV}$, $E_n < 100 \text{ MeV}$ |

by a shower from a library, as shown in Fig. 7.1. Main parameters used in ATLAS simulation are summarized in a Tab. 7.1, where FCAL1 and FCAL2 are the first two forward calorimeters (see Sec. 4.2.3) and $E_\gamma < 10$, $E_e < 1000$, $T_n < 100$ are the maximum energies of photons, electrons and neutrons used in the method.

Since currently the Frozen Showers method is used only for FCAL, this chapter will fully concentrate on optimisation of Frozen Showers in forward calorimeter.

7.2 Properties of electron showers in FCAL

The fast simulation of the forward calorimeters is a complicated task due to their complex structure. As it was mentioned in Sec. 4.2.3 FCAL consists of hexagonal absorber cells with an anode tube and cathod rod in the cell center and the liquid argon in the gap between rod and tube. In order to

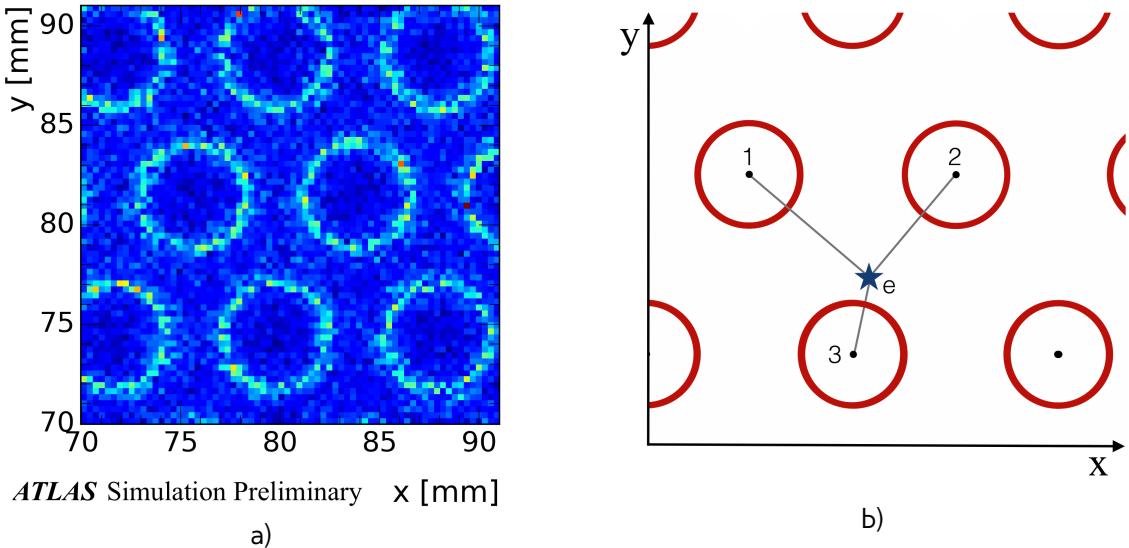


Fig. 7.2: a) Shower energy response histogram in the transverse (x vs y) plane for electrons, generated with uniformly distributed x and y and total energy less than 1 GeV. Light circles correspond to showers, starting inside the LAr gaps with on average higher energy response, while the dark parts correspond to dead material with smaller sum of the "hits" energy respectively. b) Distance to the closest rod center scheme $d_{rod} = \min(d(1,e), d(2,e), d(3,e))$, where 1,2,3 are the positions of the rod centers and e is the position of initial electron. The rod centers and liquid argon gaps are shown by black dots and red circles respectively.

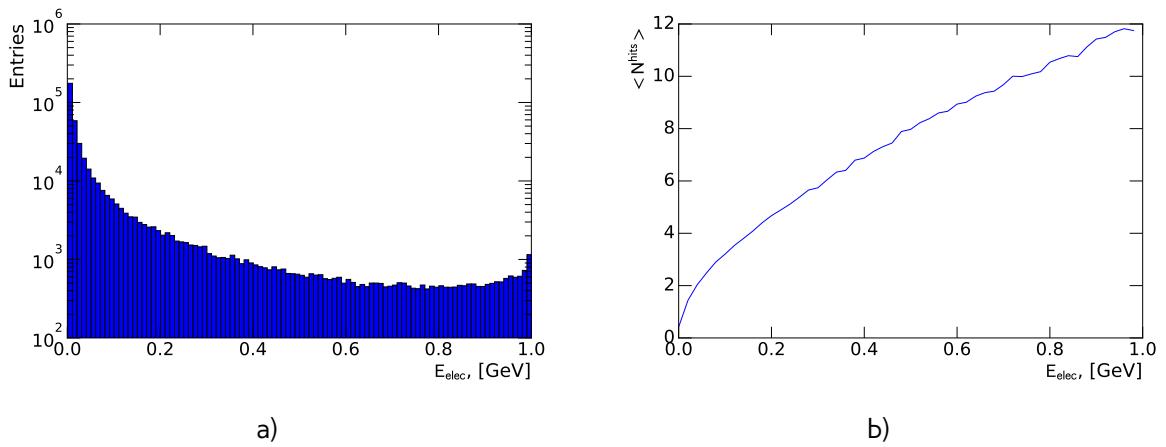


Fig. 7.3: Distribution of the a) electron energies and b) mean number of hits in the shower vs energy of electron for electrons with energy less than 1 GeV originating from 1 TeV initial electron.

866 simulate the resolution of high-energy electrons, an efficient fast simulation technique should take
 867 into account this large amount of non-uniformly distributed sensitive material.

The electron energy resolution for this calorimeter can be written as:

$$\frac{\sigma}{E} \approx \frac{1}{\sqrt{E}} \oplus \frac{1}{E} \oplus const, \quad (7.1)$$

868 where \oplus indicates the quadratic sum. The first term is the 'stochastic term', which includes intrinsic
 869 shower fluctuations, the second one takes into account readout noise effects and pile-up fluctuations
 870 of the energy loss. The energy resolution of high-energy electrons is mostly dominated by the con-
 871 stant term.

Fluctuations due to the detector design are visible in the simulation of low energy electrons,
 generated inside different points in the forward calorimeter. The shower energy E^{shower} distribution
 in the transverse x vs y plane is shown on Fig. 7.2 a). The shower energy is defined as:

$$E^{shower} = \sum E_i^{hits}, \quad (7.2)$$

873 where E_i^{hits} is the energy of the i-th hit in shower shower inside the sensitive material. The periodic
 874 structure resembles the calorimeter design, where the light circles correspond to gaps with liquid
 875 argon, which are acting as sensitive material. Introduction of distance to the closest rod center,
 876 calculated as shown in Fig. 7.2 b) allows to catch this periodic structure.

877 A typical electron substituted by frozen shower coming from simulation of high-energy electrons
 878 has a relatively small energy (Fig. 7.3 a). The mean number of deposits in the sensitive material in a
 879 "frozen" shower is around 5 and this value rises with the electron energy (Fig. 7.3 b). Fig. 7.4 presents
 880 the distribution of the distance to a closest rod center vs shower energy for showers from electrons
 881 with energy below 1 GeV originating from initial electrons with an energy of 1 TeV. The liquid argon
 882 gap is marked by red lines. There is visible peak in showers energies for the region around liquid
 883 argon gap. The similar structure is also visible in a number of hits (Fig. 7.5 a) and the standard
 884 deviation of energy of the hits in the shower (Fig. 7.5 b) distributions. The magnitude of the peak
 885 depend on the electron energy and is higher for the low energies (Fig. 7.6 a) and less significant for
 886 higher energies (Fig. 7.6 b). This fact combined with energy distribution states the importance of a
 887 proper simulation of non-uniformities for showers coming from a low energy electrons.

888 On the another hand, the use of the frozen showers in too low energy region can be suboptimal
 889 because of the small number of energy depositions in a shower. For electrons with energies below
 890 30 MeV 90% of the showers have no depositions and only 0.5% of showers have more than 1 hit
 891 (Fig. 7.7). It was figured out, that below this energy, the substitution of the electron by the single hit
 892 with electron energy have showed a faster speed of simulation.

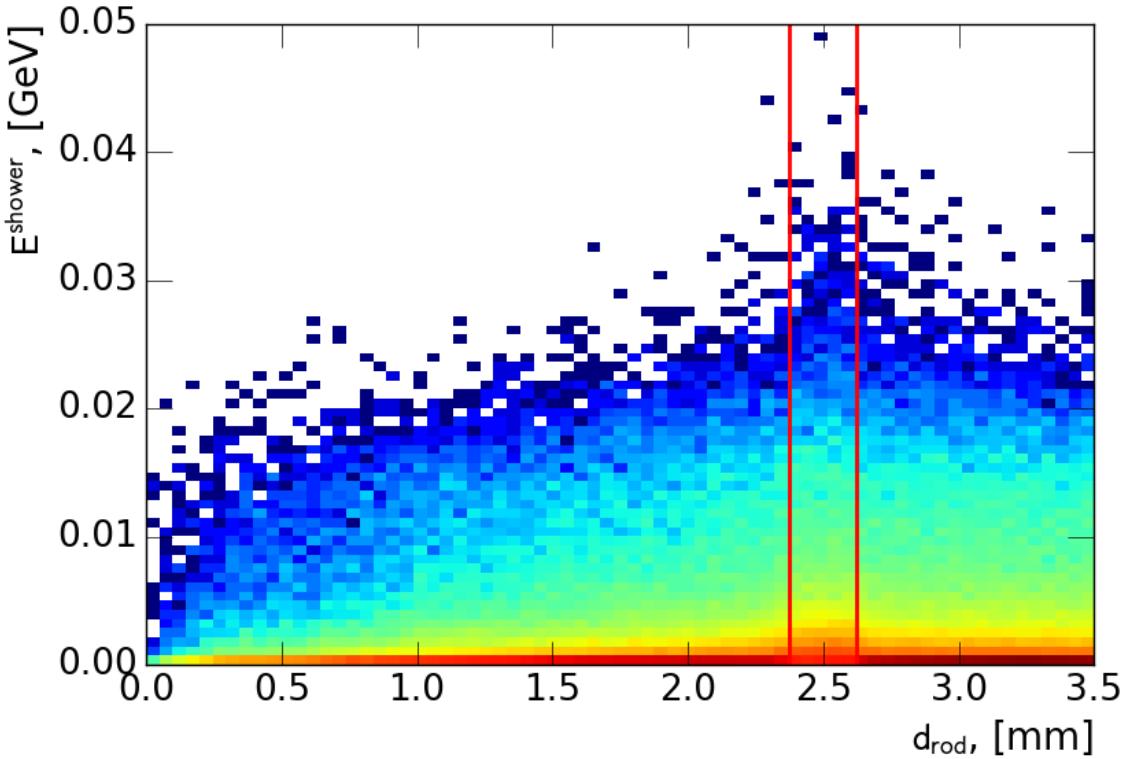


Fig. 7.4: Distribution of distance to the closest rod center vs shower energy for electron showers created by electrons with energy less than 1 GeV originating from the initial electrons with energy 1 TeV in distance to the closest rod center vs shower energy plane. Position of the liquid argon gap is noted by a red lines.

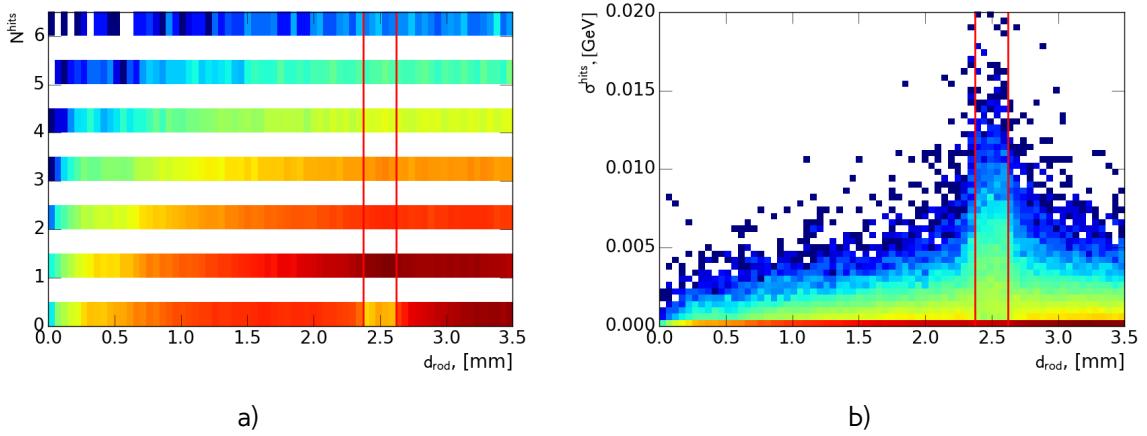


Fig. 7.5: Distribution of distance to a closest rod center vs a) number of hits in a shower plane and b) standard deviation of hits in a shower energy of electron showers created by electrons with energy less than 1 GeV coming from initial electron with energy 1 TeV. Position of a liquid argon gap is noted by a red lines.

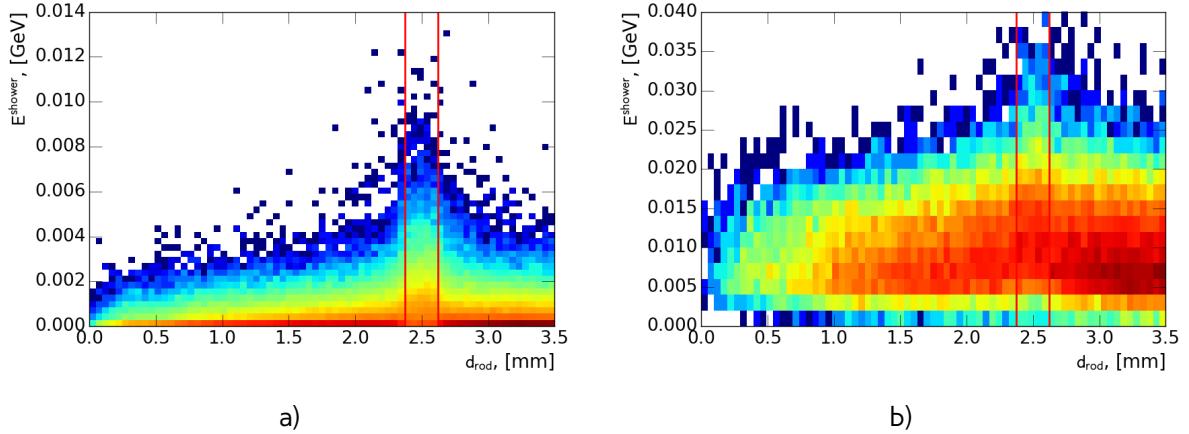


Fig. 7.6: Distribution of distance to a closest rod center vs shower energy for electron showers created by electrons with energy a) less than 100 MeV and b) higher than 300 GeV coming from initial electron with energy 1 TeV in plane. Position of a liquid argon gap is noted by red lines.

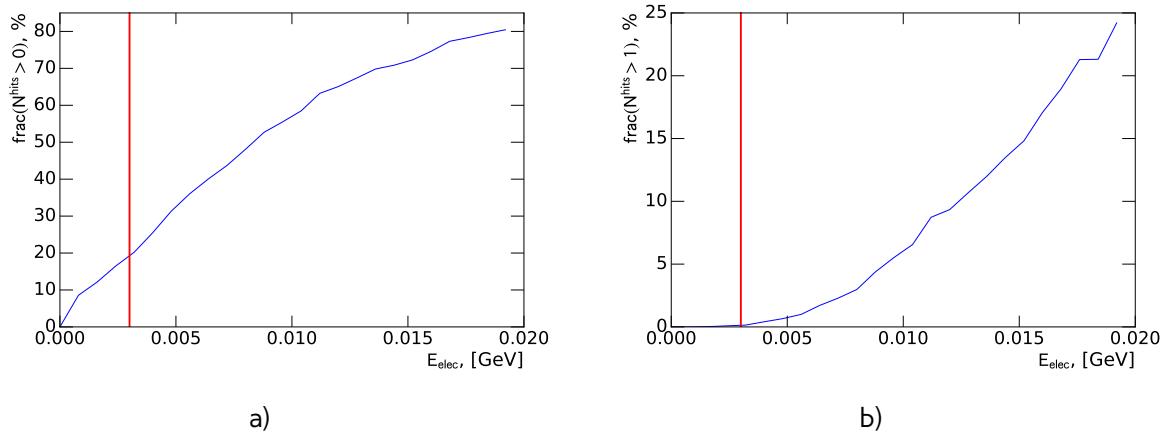


Fig. 7.7: Fraction of showers with a) at least 1 b) at least 2 depositions inside the sensitive material depending on the initial electron energy. The red line denotes the 30 MeV limit for the frozen showers method.

893 7.3 Generation and use in simulation

894 As it was mentioned in the introduction, the frozen showers method consist of two stages: generation
 895 of libraries and the use in simulation. The generation needs to be repeated for each significant
 896 change in the physics processes description of Geant4 or in the description of the detector. Showers
 897 are stored in a library in bins of pseudorapidity and distance to the closest rod center, while the
 898 energy remains unbinned. There is special *liquid argon bin* for distance with position and width,
 899 that corresponds to the parameters of liquid argon gap, that was introduced in order to catch non-
 900 uniformities form Sec. 7.2.

901 In order to obtain a proper energy distribution for the generation of the library, particles originating
 902 from SM process ($t\bar{t}$ or high energy electrons) are usually used. For each particle eligible for frozen
 903 showers use parameters are saved for a later use. On a second stage, these particles are propagated
 904 through the calorimeter using full ATLAS simulation infrastructure. Each hit is saved as a shower
 905 inside the library in a corresponding pseudorapidity and distance bin.

906 Additionally, in order to save disk space as well as a memory consumption, the hit information is
 907 compressed. This compression is performed in two steps:

908 **Hit merging** If the distance between any two hits is smaller, than a given parameter R_{min} , then
 909 these hits are merged into one deposit at the energy weighted center of them. This process
 910 is done iteratively.

911 **Truncation** Hits which energies are below a fixed fraction f of the total energy sum of all hits, are
 912 truncated. The energy of the remaining hits is rescaled in order to preserve the total deposited
 913 energy.

During simulation, if the energy of a particle falls below a cut-off energy, the FS algorithm examines the resulting shower. It checks whether the particle is far from the edges of the calorimeter, such that the shower is by 90% contained inside the calorimeter. This depends also on the energy of the particle, since the shower sizes are increasingly growing with energy. The algorithm searches for a shower with the closest energy in the corresponding pseudorapidity and distance bins. The shower is rotated in the direction of the particle. In order to correct for the differences in the energy, each hit in the shower is scaled as:

$$E_{hit}^{new} = E_{hit} \cdot \frac{E_{part}}{E_{part,lib}}, \quad (7.3)$$

914 where E_{hit} is the original energy of the hit, E_{part} is the energy of the particle and $E_{part,lib}$ is the energy
 915 of the particle from the library. Afterwards particle is substituted by the resulting shower. Later, the
 916 reconstruction algorithm uses these hits from the frozen shower as usual energy deposits in the
 917 sensitive material.

918 7.3.1 Tuning of libraries-

919 The good simulation method is required to be consistent with full simulation on all possible recon-
 920 stucted objects. In case of Frozen Showers in forward calorimeter, the electron energy resolution is
 921 the most problematic value, since the resolution of the reconstructed electrons is around 2 times
 922 smaller(Fig. 7.8), than in the full simulation. Using the Eq. 7.1, this behavior can be interpreted as a
 923 smaller size of fluctuations for fast simulation and therefore lack of the high-energy showers origi-
 924 nating in sensitive material. This problem can be solved by tuning the parameters of library in order
 925 to match the full simulation.

926 The tuning consists of a 2-step manual procedure:

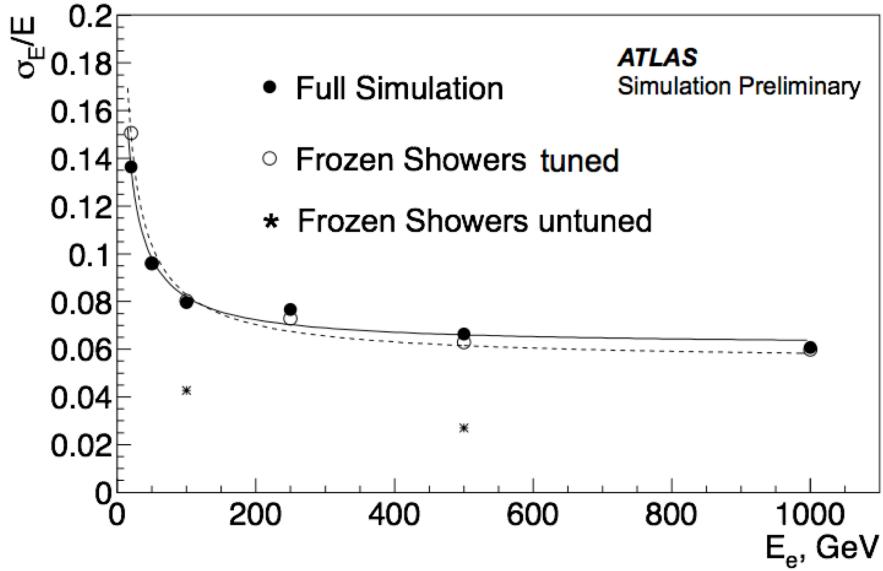


Fig. 7.8: Electron resolutions for full simulation(black dots), tuned(white circles) and untuned(star points) frozen showers. Electrons simulated with frozen showers libraries before tuning have twice smaller resolution, than electrons from full simulation. Tuning allows to gain better agreement with full simulation.

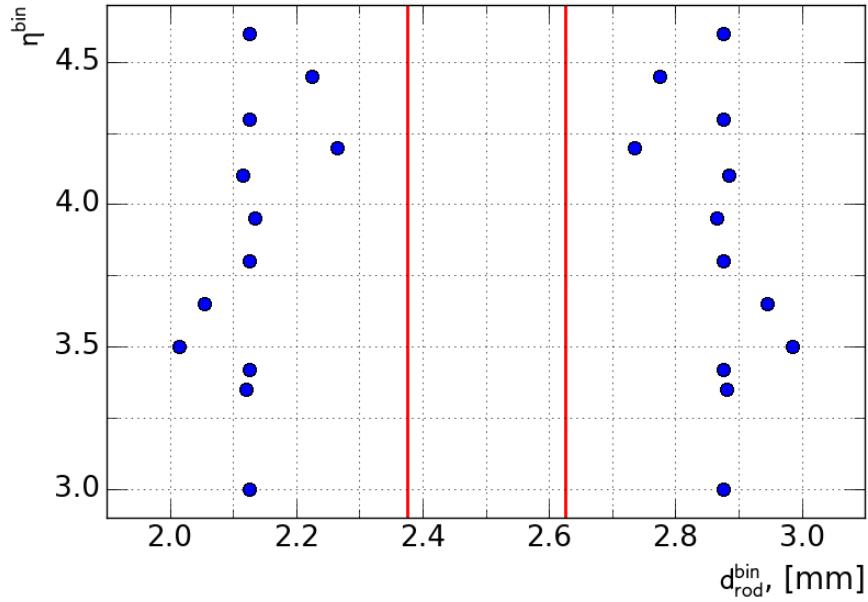


Fig. 7.9: Position of gap bins for different η bins in the previous libraries after tuning. The dots correspond to the limits of each bin. The red lines are denoting the original position of the bins, which correspond to the position of the liquid argon gap in the calorimeter.

927 **Changing bin width** At this stage position of the liquid argon bin is moved, so what a bin width
928 is enlarged. This causes a higher number of showers with higher response in simulation and
929 therefore higher fluctuations. This procedure causes a higher resolution and a mean energy
930 of reconstructed electrons.

931 **Shower energy scaling** In order to correct for the introduced shift in the energy scale, the shower
932 energy is reduced by rescaling all the hits in the shower to the shift in the mean reconstructed
933 energy.

934 It is repeated iteratively in each pseudorapidity bin separately until the desired agreement is ob-
935 tained. The resulting bin positions are shown on a Fig. 7.9. This method yields a relatively good
936 agreement with full simulation (black dots on Fig. 7.8). However, it is necessary to repeat this pro-
937 cedure for each new library generation and this requires a significant tuning effort, which makes it
938 not optimal.

939 7.4 Machine learning based bin finding procedure

940 Since frozen showers were planned to be used in the Run-2 Monte Carlo production, there was a
941 need for a more automatic procedure of library generation with proper electron resolution. One of
942 the possible ways is to choose the different positions of liquid argon bins during library generation
943 using machine learning tools. In this section a newly developed automatic bin finding procedure will
944 be discussed.

945 7.4.1 Machine learning introduction

946 Machine learning is a set of algorithms, which allows algorithms to learn and improve from experience
947 without being explicitly programmed. This is a modern field of computer science, that is used in
948 different fields like computer vision, natural language processing, data science etc. There are two
949 main types of machine learning algorithms: *supervised*, where an example of the desired output is
950 provided by the "supervisor" and the goal is to learn a general rule, that maps inputs to outputs,
951 and *unsupervised* learning, where no labels are given to the algorithm, which discovers the hidden
952 patterns in the data [64]. The initial data parameters of interest, which are used in the algorithm to
953 "learn" are called *features*.

954 Machine learning algorithms can be used for solving a classification problem, where each event
955 should be identified to one of the specified classes. Since the first introduction of machine learning
956 classifying algorithms called perceptron by Rosenblatt [65], many different algorithms have been
957 invented. In this analysis, decision trees and support vector machines implemented in Scikit-Learn
958 python package [66] are used.

959 Binary decision trees

960 *Binary decision trees* [67], called also single decision trees, are one of the most commonly used
961 machine learning algorithms for a classification problems in particle physics. It can be represented
962 as a set of sequential cuts on input variables. Scheme of this algorithm is shown in Fig. 7.10 a). Red
963 circles show the nodes of the tree. Each node corresponds to one of the internal input variables
964 and connects to two branches, that are split in respect to the value. The first node is called a
965 root node. The depth of the tree is the number of branches from the node to the tree's root node.
966 The tree ends with squares, called leaf nodes, where all events are classified to a certain class. Leaf

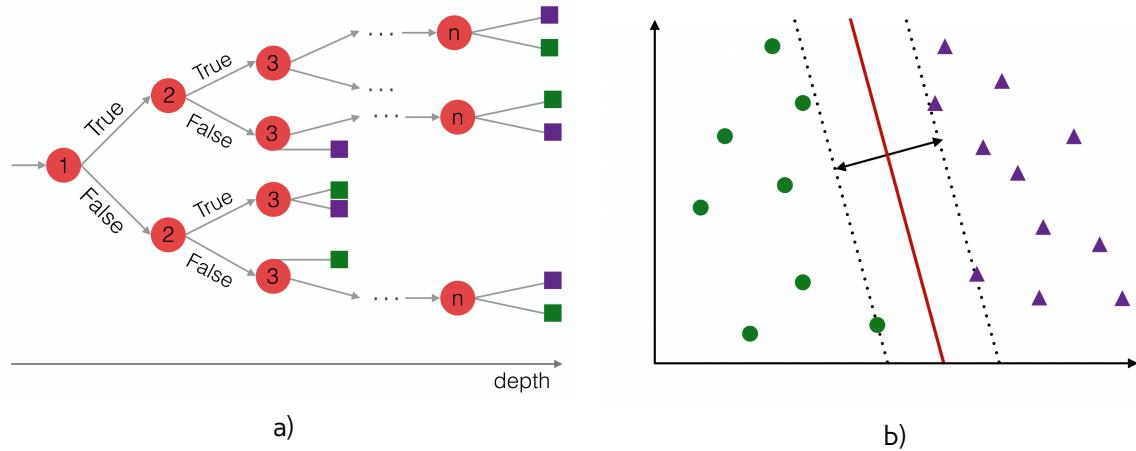


Fig. 7.10: Schematic representation of machine learning algorithms, used in this analysis for the classification of showers. Green figures represent the first class of events, whereas violet ones belong to a second class. a) Representation of a binary decision tree structure: red circles correspond to nodes, which are split with respect to the selected feature. Squares represent leafs, where events classified to a certain class. The depth of the tree is calculated as a maximum number of edges from the node or leaf to the root node. b) Representation of the SVM algorithm. The dividing hyperplane is shown by the solid line. The dashed lines represent the maximum margin boundaries.

967 node represents classification or decision. The tree, where each node has at most 2 children called
 968 binary decision tree.

The tree is build using the variable called Shannon entropy [68], what is similar to the entropy in physics:

$$S = - \sum_{i=1}^N p_i \log_2 p_i, \quad (7.4)$$

where p_i is the probability to find event of class i . Each split in a variable should decrease the entropy of the system. The information gain is defined as the difference in entropy after the split:

$$IG(Q) = S_0 - \sum_{i=1}^2 S_i, \quad (7.5)$$

969 where S_0 is the initial entropy, without new node, S_i is the entropy of the one of the 2 node children.
 970 The node with the highest information gain is taken. One of the main advantages of the decision
 971 trees its simplicity of visualization and interpretation.

972 Support vector machines

Support vector machines (SVM) is a supervised machine learning algorithm which can be used for classification problems [69]. In this algorithm each event is represented in a p -dimensional parameter space. The classification is performed by finding hyper-plane that differentiates given two classes with the largest possible separation (Fig. 7.10 b). The hyperplane can be described with the set of

points \vec{x} in the parameter space satisfying:

$$\vec{w} \cdot \vec{x} - b = 0, \quad (7.6)$$

where \vec{w} is the normal vector to the hyperplane and the parameter $\frac{b}{\|\vec{w}\|}$ determines the offset of the hyperplane from the origin along the normal vector \vec{w} .

The maximum margin boundaries are described by equations:

$$\vec{w} \cdot \vec{x} - b = 1, \quad (7.7)$$

$$\vec{w} \cdot \vec{x} - b = -1, \quad (7.8)$$

where $\frac{2}{\|\vec{w}\|}$ is the distance between these 2 hyperplanes, such that the planes with the maximum margin between them should have minimum $\|\vec{w}\|$.

In order to prevent each point from falling into the margin, the following constrain should be satisfied:

$$\vec{w} \cdot \vec{x} - b \geq 1 \text{ where } y_i = 1, \quad (7.9)$$

$$\vec{w} \cdot \vec{x} - b \leq -1 \text{ where } y_i = -1, \quad (7.10)$$

where y_i represents the class of the i-th event, that can be either 1 or -1. These equations can be rewritten as:

$$y_i(\vec{w} \cdot \vec{x} - b) \geq 1 \quad (7.11)$$

It is also possible to construct a non-linear classifier by replacing the dot-product with a different kernel function. In this thesis, a radial basis function (RBF) kernel is used:

$$K_{rbf}(\vec{x}_i, \vec{x}_j) = e^{-\gamma|\vec{x}_i - \vec{x}_j|^2} \quad \gamma > 0, \quad (7.12)$$

where the parameter γ adjusts the width of the kernel.

7.4.2 Electron shower categorization

As it was mentioned in the previous sections, the FCAL modules consist of different types of material such that showers starting inside the dead material are usually having lower energies, than those started in sensitive material. However, the validation (Fig. 7.8) can be interpreted as an implication that there are high-energy showers outside the liquid argon gap. It could be explained by the fact, that electrons, created in a dead material, can cross a liquid argon gap and give a hit there as shown in Fig. 7.11 (electron 2). These electrons would be indistinguishable from electrons created directly in the sensitive material (electron 1 in Fig. 7.11).

Due to this similarity, electrons created outside the sensitive material can be treated together, labelling such showers *sensitive material showers*. Showers that did not cross a liquid argon gap, are called dead material showers. This means, that in this model a real liquid argon gap can be substituted by the "effective" one, with larger width.

The width of the effective liquid argon gap, from the definition, depends on the following parameters:

Electron energy The gap should get wider with higher energy of the initial electron, because of the growth of the mean free path with energy.

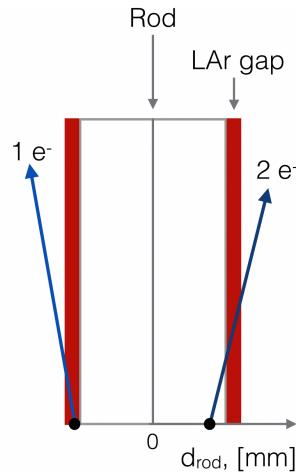


Fig. 7.11: Schematic representation of the model of shower creation in the FCAL. Electron 1 is created in a liquid argon gap. Electron 2 is created near a liquid argon gap and crosses it. This causes a smearing of sensitive material showers distribution. Electrons created in the sensitive material tend to create more energetic showers, than electrons from the dead material. However, electrons shown on this scheme may give similar showers and therefore may not be distinguishable.

997 **Direction of the electron** Electrons aligned collinearly with the liquid argon gap will have smaller
998 probability to cross it. This probability will grow with the angle reaching its maximum at 90°

999 The bin finding is performed as a 2-step procedure: on the first stage the first classifier distinguish
1000 the showers based on their simulated parameters and on the second step the second classifier aims
1001 to produce a hyperplane in d_{rod} , E of initial electron phase-space. This classifiers and the training
1002 sample used will be discussed in details in the following subsections.

1003 **Training sample**

1004 Real distributions of parameters of electrons, used in simulation, have a complicated structure and
1005 depend on the physics processes simulated. Machine learning could identify these dependencies
1006 instead of the needed ones. This is why a simplified data is needed as a training sample for ma-
1007 chine learning. The training sample was produced by simulation of electrons directly in the forward
1008 calorimeters. In order to treat equally high and low energy electron initial showers, a uniform energy
1009 distribution is used.

1010 Fig. 7.12 shows the distribution of the shower direction ($\eta^{direction}$) vs electron energy for electrons
1011 coming from simulation of 1 TeV electron. Most of the showers have direction in η range between
1012 3.0 and 5.0, which corresponds to position of the FCAL. It was figured out, that the direction of the
1013 shower is highly correlated with the position of the electron. Therefore, electrons were generated
1014 uniformly in η between 3.0 and 5.0.

1015 **First classifier**

1016 The first classifier aims to categorize all showers by means their parameters. An supervised learning
1017 algorithm is used on artificially reduced training sample, where the labels can be put automatically.
1018 The obtained classifier is later expand the classification to the full training sample.

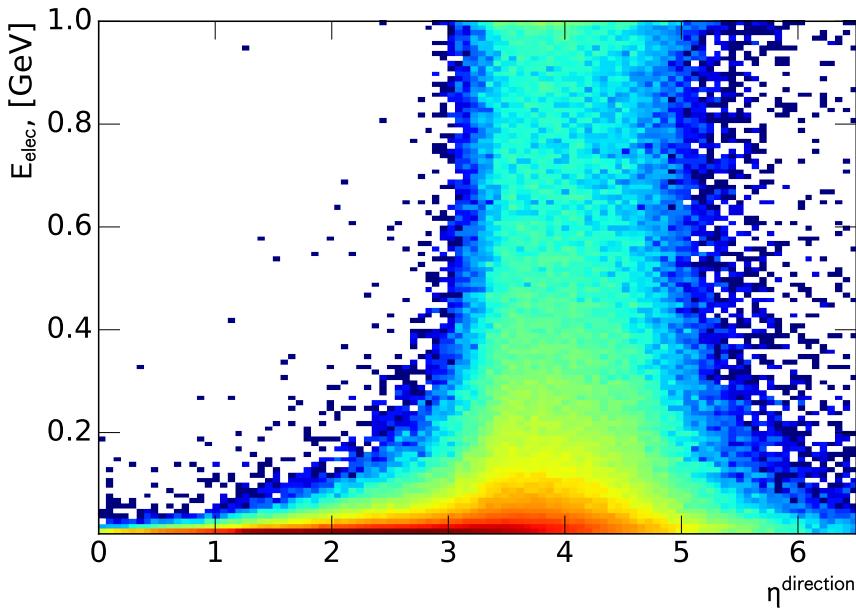


Fig. 7.12: Distribution of sub-showers energy vs direction of shower $\eta_{momentum}$ for showers from the production of 1000 GeV electrons.

1019 Pre-labeling is done using the definitions of sensitive and dead material showers based on the
 1020 distance to a closest rod center (Fig. 7.13). Showers starting in a liquid argon gap are defined as
 1021 sensitive material showers, while showers coming from electrons born near the rod center and on
 1022 the edges of the cells can be labeled as dead material showers, since there is a small probability for
 1023 the electron, which caused the shower, to reach the liquid argon gap.

1024 For this classifier a simple decision trees have been chosen, since it has showed a good classifi-
 1025 cation efficiency on the reduced training sample. Different input parameters have been tested using
 1026 their variance, and it was figured out, that the best set of the differentiating parameters is:

- 1027 • Shower energy, defined as the sum of all sensitive material hits energies in the shower;
- 1028 • Maximum hit fraction. This quantity is calculated as the energy of the most energetic hit
 1029 divided by the shower energy;
- 1030 • RMS of the hits, calculated as the standard deviation of the hits energies in the shower.

1031 The classification efficiency of the obtained binary search tree on the reduced sample with the
 1032 depth = 2 is 97%. The expanded to the full phase-space results are shown in Fig. 7.14 a).

1033 **Second classifier**

1034 The second classifier uses predictions of the first classifier as input label.

1035 It reconstructs a best dividing hyperplane between two types of shower using a support vector
 1036 machines. It uses as input the truth parameters of the electron, e.g. energy of the initial electron and
 1037 its distance to the closest rod center. Different kernels have been tested and the best predictions

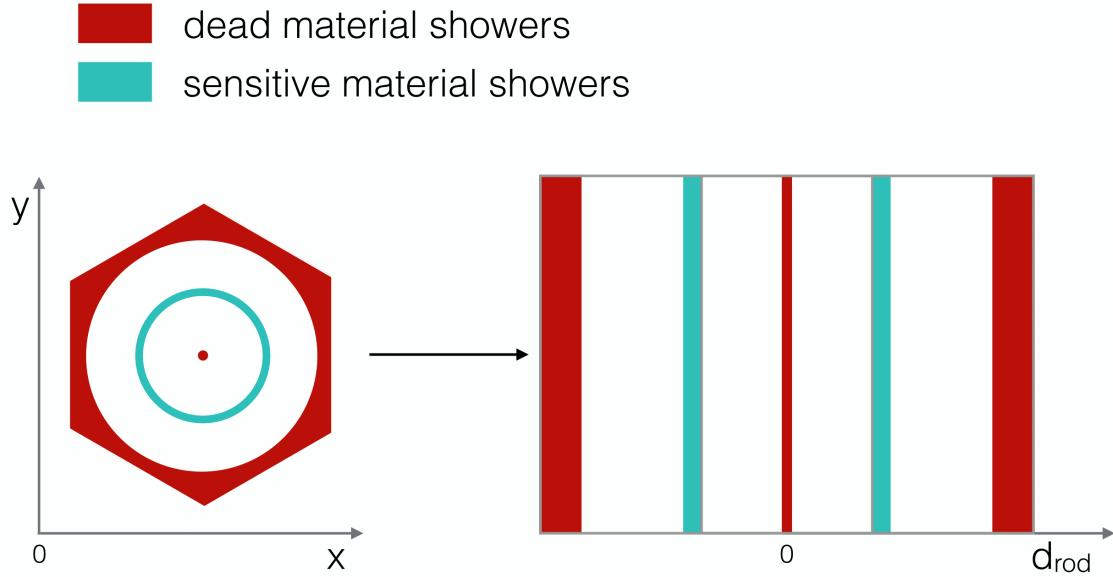


Fig. 7.13: Schematic representation of the preselected data for the first classifier in the x-y (left) and distance(right) plane. Electrons, created near the rod center and on the borders of the module have low probability to cross the sensitive material, while those created inside the liquid argon gap are considered as sensitive material showers.

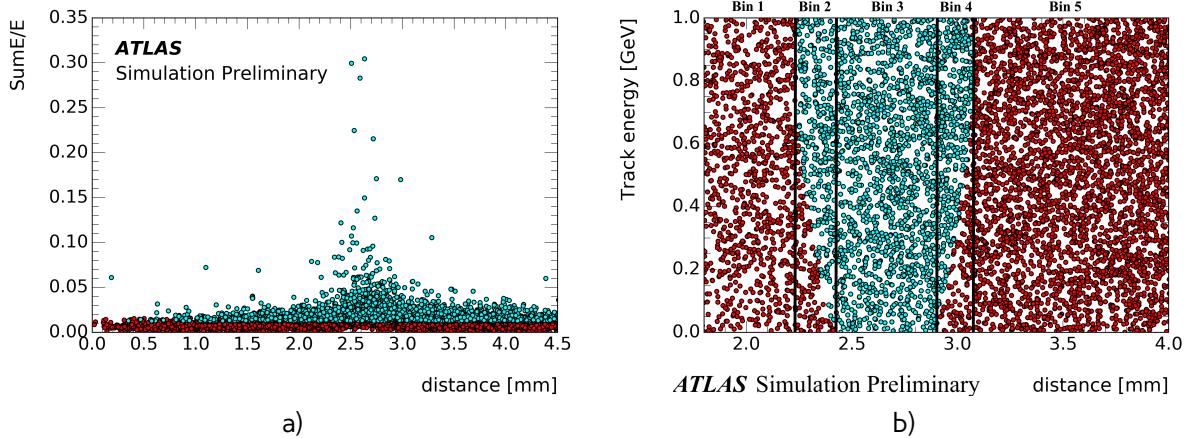


Fig. 7.14: Results of machine learning algorithm classification for a) first classifier b) second classifier. Cyan dots correspond to sensitive material showers, red to dead material showers. The black lines in Fig. b correspond to the resulting bin positions.

1038 have been obtained using RBF kernel (Eq. 7.12). Assuming, that $\eta^{momentum} \approx \eta^{position}$ a classification is
 1039 performed in each $\eta^{distance}$ bin used in the library.

1040 An example of the classifier output is shown in a Fig. 7.14 b). The obtained gap positions are wider,
 1041 than the original ones, as it is expected from the model. Variation of the obtained parameters have
 1042 been found small, so the mean of the parameters has been as an input for the binning.

1043 7.4.3 Interpretation of results

1044 Since full new regeneration of libraries and the validation of reconstructed variables is time-consuming
 1045 procedure, a Monte Carlo method has been developed for a cross-check of the classifiers and its in-
 1046 terpretation. It uses pseudorapidity $\eta^{position}$, energy of electron and distance to the closest rod center
 1047 from the data as a reference for the random generator. This simulation allows to compare the shower
 1048 energies and shower energies divided by the energy of the initial electron (SumE/E) distributions with
 1049 the distributions coming from the full simulation, which are considered as a reference.

1050 The resulting hyperplane from second classifier can be translated to the bin positions in the
 1051 different ways. Several interpretations of the bin positions have been tested and the best one is
 1052 shown in Fig. 7.14 b) with black lines. The best SumE/E agreeemen have been achieved using three
 1053 bins in liquid argon position instead of the only one. The central bin, according to the classifier,
 1054 contains just sensitive material showers events, while the other two there is a mixture of dead
 1055 and sensitive material showers. The obtained positions of the liquid argon bins are wider, than the
 1056 nominal ones for both FCAL1 and FCAL2.

1057 Comparison of SumE/E distributions using toy MC on old libraries (Fig. 7.15 a) and the libraries
 1058 with the new binning (Fig. 7.15 b) has shown, that we could expect a better performance on the
 1059 reconstructed values for the new binning.

1060 7.4.4 Reconstructed electron energy

1061 Since the resolution of the single electrons can have a significant difference, the energies of recon-
 1062 structed electrons are validated before the mass validation of the different groups. Measurement of
 1063 the shift in the mean energy between full and fast simulation allows to correct the scale for frozen
 1064 showers.

1065 Validation is performed for the following electron energies: 100 GeV, 200 GeV, 500 GeV and 1000
 1066 GeV and within the η directions that are corresponding to the 12 η bins of the library. The resolution
 1067 is calculated as RMS of all reconstructed energies for the certain energy and η . The results of the
 1068 electron resolution validation for the new machine learning based binning and old "tuned" libraries
 1069 is shown in Fig. 7.16. The new methods gives a better or comparable resolution agreement than an
 1070 old libraries. However, there are 2 bins, there new method resolution is significantly worse, than old
 1071 one (3.5 and 4.3). This means, that this method still needs to be improved. The possible ways of its
 1072 improvement are discussed in Sec. 7.4.5.

1073 In a meanwhile it was decided to use a combination of the new and old libraries. The mean shift
 1074 is corrected as described in Sec. 7.3.1 and showed in Fig. 7.17. The remaining differences between
 1075 full and fast simulation are considered negligible.

1076 7.4.5 Plans for the future

1077 The validation have showed a good agreement between full simulation and fast simulation for most of
 1078 the bins, however, because not all of the bins are performing equally well, the additional modifications

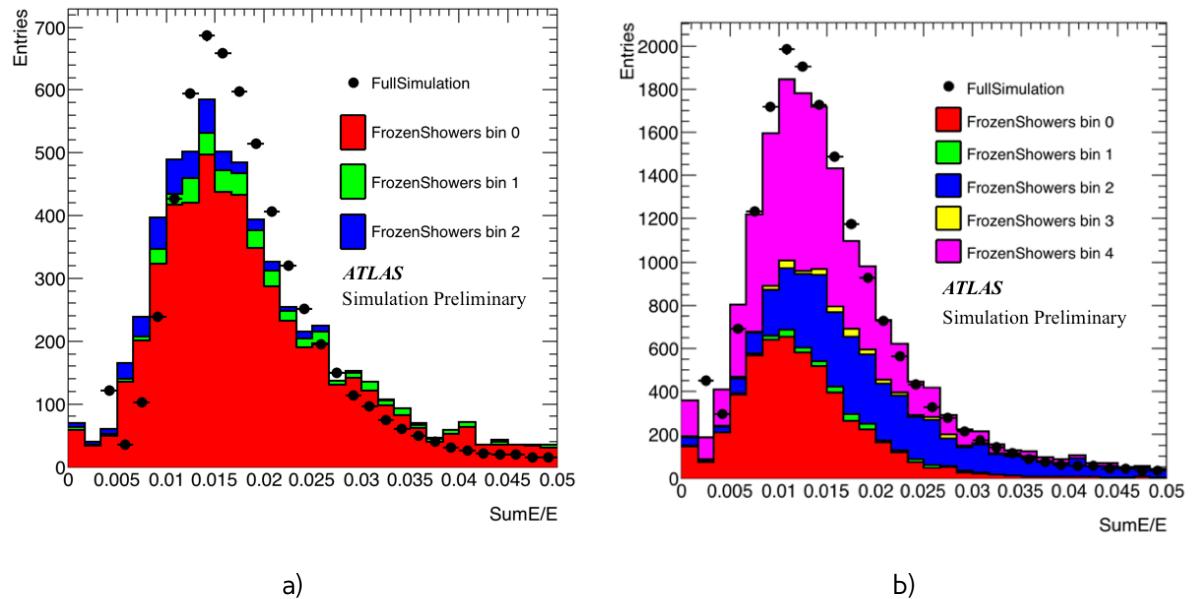


Fig. 7.15: Comparison of the distributions of shower energy divided by the energy of the initial electron between full simulation and toy MC using libraries for liquid argon gap bins and 2 closest to them bins for a) old "tuned" libraries with 1 liquid argon gap bin b) new libraries using 3 liquid argon gap bins. There are still remaining differences between full simulation and toy MC, but the new machine learning binning gives a better agreement with full simulation.

1079 of the algorithm are needed. The following modifications have been investigated and planned to be
1080 performed in the nearest future:

- 1081 • Procedure with η dependent bin size. Currently all of the bins have the same size and position
1082 of the liquid argon bin. However, because of the outlying bins, the procedure should be modified
1083 and determine the bin position separately for each bin.
- 1084 • Use of the closer to real case for training sample. The problem of electron resolution could be
1085 also caused by too simplified models, used to train on. It is planned to repeat the procedure
1086 for training sample with distributions, closer to the nominal ones.
- 1087 • Adding the direction of the shower binned variable. Since there is a complicated dependency
1088 between position of the electron and its direction (especially in small energy region), the addi-
1089 tional binning could solve the remaining problems with electron energy resolution.

7.4 Machine learning based bin finding procedure

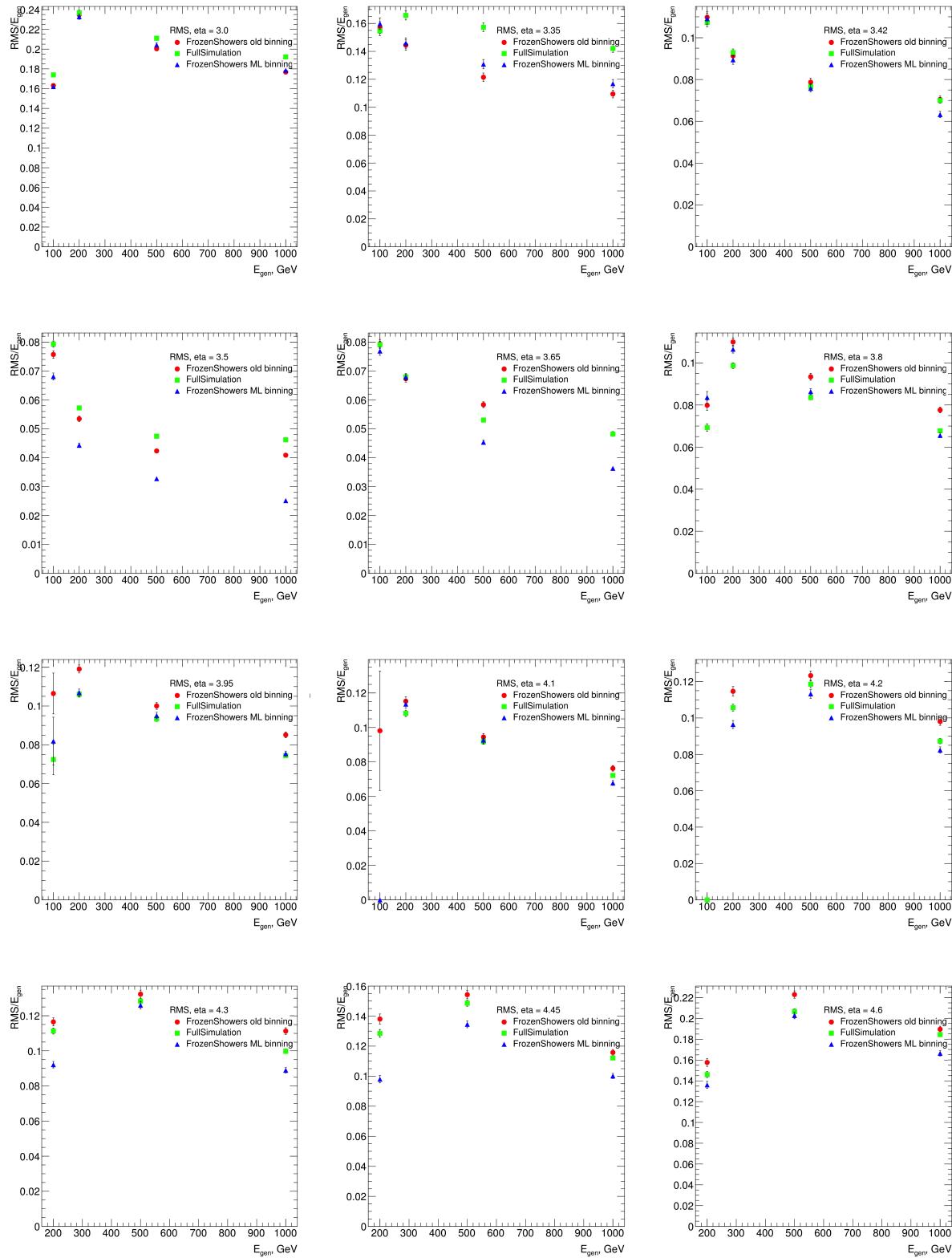


Fig. 7.16: Energy resolution of reconstructed electrons for full simulation, new libraries with ML binning and old tuned libraries with original binning for different η bins

Frozen Showers

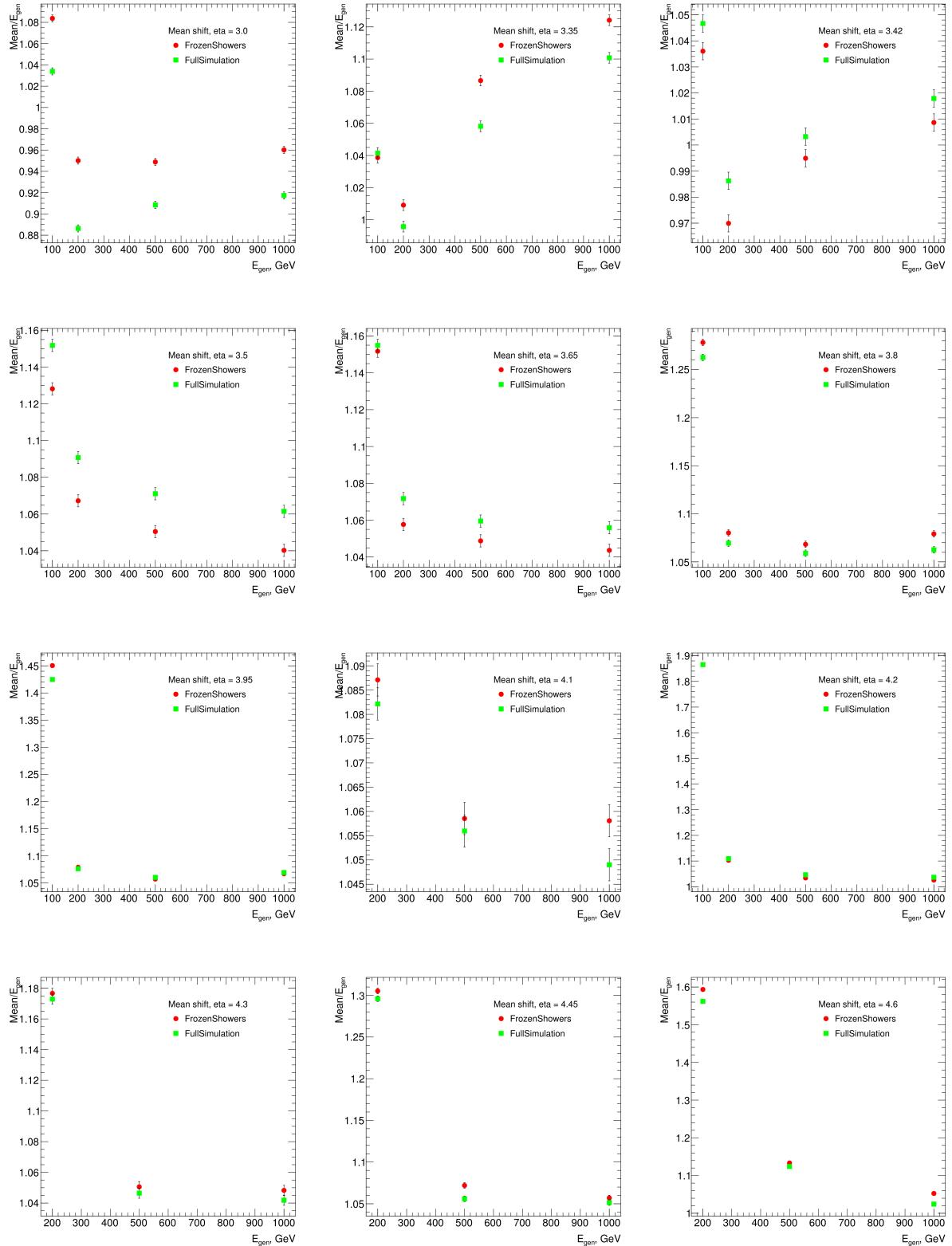


Fig. 7.17: Shift of the reconstructed energy to the truth energy of electrons for full simulation, new FS libraries with ML binning and full simulation for different η bins

1090 **7.5 Validation of the new libraries**

1091 The good fast simulation method should work equally good on all types of reconstructed objects,
1092 this is why for each new frozen showers libraries production an new iteration of mass validation is
1093 performed. The validation is done separately for each object by the different groups by comparing
1094 the distibutions obtained from full and fast simulation

1095 Frozen showers have been validated on following objects and showed a good agreement:

- 1096 • Z bosons from $Z \rightarrow ee$ sample with one central and one forward electrons(Fig. 7.19 a). The
1097 resolution of Z-mass peak is dominated by the resolution of the central electron, so Z boson
1098 is mostly sensitive just to the mean energy of the forward electrons. There is visible shift in
1099 the mass distribution between data and Monte-Carlo, that however, is within tolerable region.
- 1100 • Jets form two jet events. The validation have showed a good agreement for all of the variables.
1101 The distribution of the jet response (Fig. 7.19 b) showed, that Frozen Showers method does
1102 not change the jet scale.
- 1103 • Topo clusters from $t\bar{t}$ events.
- 1104 • Forward electrons. The forward electrons validation have showed, that usage of the frozen
1105 showers is not changing the η and E_T distributions of the forward electrons. Studies of forward
1106 electrons resolution have been performed separately and will be discussed in the previous
1107 section.

1108 The total agreement between full and fast simulation for different objects makes a Frozen Showers
1109 method applicable for a MC production in 2016.

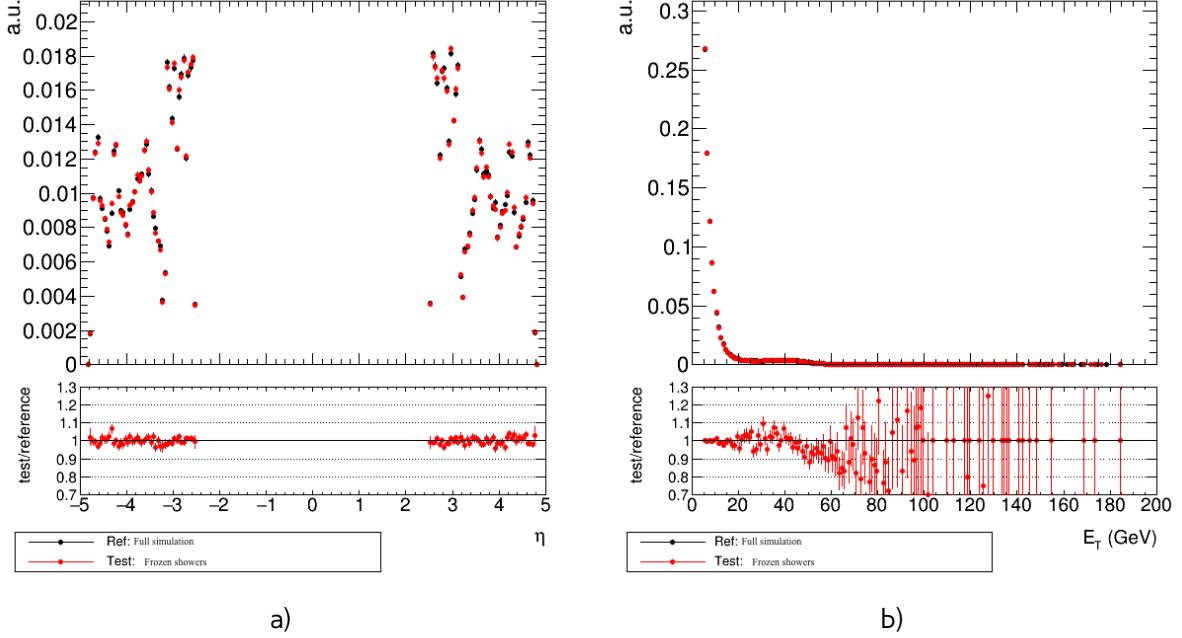


Fig. 7.18: Results of validation of the frozen showers library on forward electrons. Comparison between full simulation and fast simulation using frozen showers in forward electron events for a) pseudorapidity and b) electron transverse energy distributions. Modified from [70].

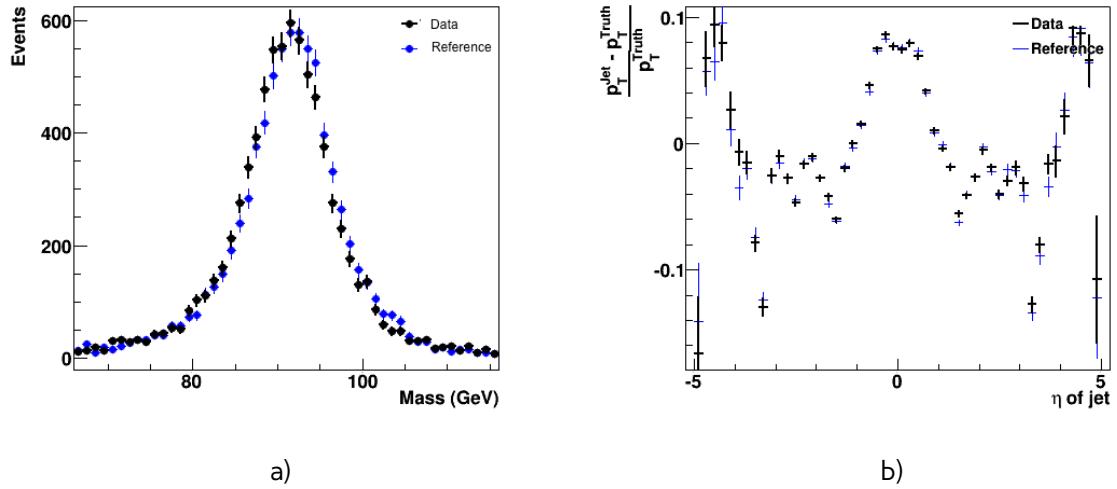


Fig. 7.19: Results of validation of the frozen showers library on jets and $Z \rightarrow ee$ sample. Comparison between full simulation and fast simulation using frozen showers for a) mass of the dilepton pair in $Z \rightarrow ee$ events (modified from [71]) b) jets response vs pseudorapidity distribution (modified from [72]).

Chapter 8

Data and Monte-Carlo samples

8.1 Data sample

The data used in this analysis was collected in pp collisions with centre-of-mass energy 2.76 TeV during the Run-1 of the LHC operation using the ATLAS detector. This was a dedicated 2013 run for a heavy-ion group with unusual for LHC low pile-up. The mean number of interactions per bunch crossing, as shown in Fig. 8.1 a) was lower, than 1.

The this run ATLAS collected 4.45 pb^{-1} of data (Fig. 8.1 b). However, not all of the data is applicable for a precise physics analysis, so the set of additional data quality (DQ) cuts was applied. It uses an information about sub-detectors, that could be disabled during data-taking. These information is stored in so called Good Run List (GRL). Total luminosity, used in the analysis is 4.0 pb^{-1} with the uncertainty 3.1% [73].

This value, using the NNLO predictions from Eq. 2.17-2.19 could be translated to the expected number of events in data as:

$$N_{\text{exp},\text{bos}} = \sigma_{\text{bos}} \cdot \mathcal{L}, \quad (8.1)$$

what leads to the numbers $1.3 \cdot 10^4$ and $1.2 \cdot 10^3$ events per channel for W and Z bosons respectively. However, due to the limited detector acceptance, reconstruction and analysis selection efficiency, not all of the boson would be visible in analyses.

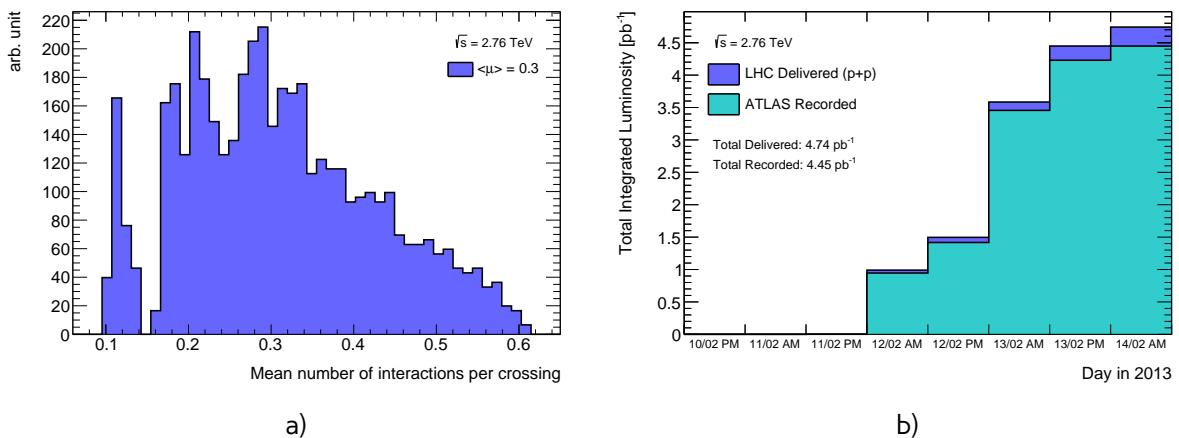


Fig. 8.1: a) Mean number of interactions per bunch crossing.

b) Cumulative luminosity versus day delivered to (dark blue), and recorded by ATLAS (light blue) during stable beams and for pp collisions at 2.76 TeV centre-of-mass energy in 2013

Table 8.1: Monte-Carlo samples, used to simulate various signals and backgrounds.

| Process | Generator | N_{events} |
|---------------------------|----------------|------------------|
| Signal MC | | |
| $W^+ \rightarrow e\nu$ | Powheg+Pythia8 | $2.0 \cdot 10^5$ |
| $W^+ \rightarrow \mu\nu$ | Powheg+Pythia8 | $2.0 \cdot 10^5$ |
| $W^- \rightarrow e\nu$ | Powheg+Pythia8 | $1.2 \cdot 10^5$ |
| $W^- \rightarrow \mu\nu$ | Powheg+Pythia8 | $1.2 \cdot 10^5$ |
| $Z \rightarrow ee$ | Powheg+Pythia8 | $9.5 \cdot 10^4$ |
| $Z \rightarrow \mu\mu$ | Powheg+Pythia8 | $7.0 \cdot 10^4$ |
| Background MC | | |
| $W^+ \rightarrow \tau\nu$ | Powheg+Pythia8 | $5.0 \cdot 10^4$ |
| $W^- \rightarrow \tau\nu$ | Powheg+Pythia8 | $3.0 \cdot 10^4$ |
| $Z \rightarrow \tau\tau$ | Powheg+Pythia8 | $2.0 \cdot 10^4$ |
| $t\bar{t}$ | Powheg+Pythia6 | $5.0 \cdot 10^3$ |
| WW | Herwig | $1.0 \cdot 10^3$ |
| ZZ | Herwig | $5.0 \cdot 10^3$ |
| WZ | Herwig | $1.0 \cdot 10^3$ |
| $b\bar{b}$ | Pythia8 | $5.0 \cdot 10^5$ |
| $c\bar{c}$ | Pythia8 | $4.8 \cdot 10^5$ |
| Additional MC | | |
| $W \rightarrow e\nu$ | Sherpa | $4.9 \cdot 10^5$ |
| $W \rightarrow \mu\nu$ | Sherpa | $4.8 \cdot 10^5$ |

8.2 Monte-Carlo samples

The Monte-Carlo was used to simulate both signal and background processes. A summary of the MC samples used in analysis is given in Tab. 8.1. The primary signals MC is generated using Powheg generator using CT10 [74] PDFs interfaced with Pythia8 for parton showering using AU2 [75] tune. There is also additional signal MC sample for W-analyses, generated using Sherpa with CT10 PDFs. This sample is used for studies of the effect of matrix elements on final cross-section (see Chap. 14).

The background sources are used to estimate fraction of the background events. The more detailed explanation of the background sources could be found in Chap. 12. The $W \rightarrow \tau\nu$ and $Z \rightarrow \tau\nu$ are generated similarly to signal MC with Powheg+Pythia8 generator with CT10 PDF and AU2 tune. Events with diboson decays are generated by Herwig with CTEQ6L1 [76] PDF set and using AUET2 [77] tune. The $t\bar{t}$ sample generated using Powheg generator interfaced with Pythia6. The additional $b\bar{b}$ and $c\bar{c}$ samples have been generated for QCD background determination and cross-check (Sec. 12.1) and generated using Pythia8 with AU2 tune and CTEQ6L1 PDF set.

1139

Part III

1140

The Measurement

1141

Chapter 9

Selection

1143 Selection criteria is a set of requirements, that are applied both on data and MC. The selection
 1144 criteria should separate process of interest (signal) from other processes. For $pp \rightarrow W \rightarrow ev/\mu\nu$
 1145 and $pp \rightarrow Z/\gamma^* \rightarrow ee/\mu\mu$ selection criteria can be divided into 3 groups: data quality, lepton
 1146 and boson cuts. The full set of cuts is summarized in a Tab. 9.1. In this chapter all of them will be
 1147 discussed and a cut flow presented.

9.1 Data quality cuts

1148 Data taking conditions are important in the analysis because of the possible biases. In order to pre-
 1149 serve high data quality some events must be rejected. The reason may be unstable beam conditions,
 1150 disabled parts of the detector or events with too many noisy cells. Numbers of runs, that can be
 1151

Table 9.1: Selection criteria

| Data quality event selection | |
|--|--|
| Single lepton trigger | |
| Good Run List | |
| Reject events with LAr errors | |
| Number of tracks at primary vertex ≥ 3 | |
| Lepton selection | |
| Electron Selection | Muon Selection |
| EF_e15_loose1 $P_T^l > 20\text{GeV}$ $ \eta^l < 2.47$ excluding $1.37 < \eta^l < 1.52$ OQ cut Medium electron identification $P_{T,cone,20}^{cone} < 0.1$ | EF_mu10 $P_T^l > 20\text{GeV}$ $ \eta^l < 2.5$ staco reconstruction chain combined muon $P_{T,cone,20}^{cone} < 0.1$ |
| W boson selection | Z boson selection |
| $E_T^{miss} > 25\text{ GeV}$ $M_T^W > 40\text{ GeV}$ | $66 < M^Z < 116\text{ GeV}$ |

Table 9.2: Number of W boson candidates in data and signal MC, remaining after each major requirement. The signal MC is normalised to the NNLO cross-section shown in Tab. 12.1

| Requirement | Number of candidates | | | |
|------------------|---|-----------|----------|-----------|
| | Data | signal MC | Data | signal MC |
| | $W^+ \rightarrow e\nu$ $W^+ \rightarrow \mu\nu$ | | | |
| No selection | 27491394 | 8354 | 27491394 | 8354 |
| Event selection | 333054 | 6044 | 26475069 | 8226 |
| Lepton selection | 15075 | 4315 | 11466 | 4782 |
| Boson selection | 3914 | 3544 | 4365 | 3936 |
| | $W^- \rightarrow e\nu$ $W^- \rightarrow \mu\nu$ | | | |
| No selection | 27491394 | 5002 | 27491394 | 5002 |
| Event selection | 333054 | 3139 | 26475069 | 4916 |
| Lepton selection | 15075 | 2306 | 11466 | 2595 |
| Boson selection | 2209 | 1941 | 2460 | 2164 |
| | $Z \rightarrow ee$ $Z \rightarrow \mu\mu$ | | | |
| No selection | 27491394 | 1196 | 27491394 | 1196 |
| Event selection | 333054 | 1051 | 445817 | 1059 |
| Lepton selection | 459 | 430 | 698 | 621 |
| Boson selection | 430 | 418 | 646 | 603 |

used in the analysis are stored in the so-called Good Run List (GRL), which in the addition to the run numbers contains also information about luminosity blocks.

Events, where LAr calorimeter was malfunctioning are excluded by LAr quality criteria. Furthermore events are required to have at least one primary vertex from a hard scattering with at least 2 tracks, that are reconstructed from this vertex.

Online selection of events is based on a single lepton trigger, depending on a lepton flavor. For electron analysis EF_e15_loose1 trigger is used, which records electrons with $E_T > 15$ GeV. This trigger uses additional "loose" isolation requirements to exclude jets, that are misidentified as electrons. In the muon channel the lowest single lepton trigger is EF_mu10. It records events containing muons with $E_T > 10$ GeV.

9.2 Lepton quality cuts

Both analyses use similar selection criteria, applied on leptons. The leptons must satisfy requirement $P_T^l > 20$ GeV and match to the event trigger.

Electron candidates are required to be within pseudorapidity range $|\eta'| < 2.47$. The electron candidates found within the transition region between the barrel and the endcap electromagnetic calorimeters, $1.37 < |\eta'| < 1.52$, are removed. Additionally, for a better multijet background rejection medium identification criteria is applied. The object quality (OQ) cut is applied in order to remove events from runs with dead front end boards in the calorimeter.

Muons have to satisfy the following criteria: they should be reconstructed by a staco algorithm and fall within range $|\eta'| < 2.5$. Additionally, for the better exclusion of background, it is required for

1172 muon to be combined.

1173 In order to have better background rejection the *isolation* criteria $P_T^{cone,20} < 0.1$ is applied on both
1174 electrons and muons. This criteria uses the information about the tracks in ID that are fall within
1175 the region $\Delta R = \sqrt{\eta^2 + \phi^2} = 20$ of the selected lepton track. It excludes the events, there sum all
1176 of these tracks momenta (except for the selected lepton) is greater than $0.2 \cdot P_T^l$.

1177 9.3 Boson selection

1178 The events, containing W boson candidates are required to have exactly one lepton, fulfilling the
1179 lepton selection. Missing transverse energy, used as a proxy for a neutrino from W decay is required
1180 to be $E_T^{miss} > 25\text{GeV}$. The transverse mass, calculated from the lepton and missing transverse energy
1181 as in Eq. 3.5 has to be bigger than 40 GeV ($M_T^W > 40\text{ GeV}$).

1182 The reconstructed lepton pair in case of Z boson analysis is required to have the invariant mass
1183 between 66 and 116 GeV. Both upper and bottom limits allow to exclude regions with high background
1184 contamination.

1185 9.4 Cut flow

1186 The effect of each selection can be studied using cut-flows, which show the number of events
1187 passing each set of cuts in a sequential order. Cut flows for W and Z analysis are shown in a Tab.
1188 9.2.

1189 Chapter 10

1190 Monte-Carlo corrections

| | | |
|------|---|----|
| 1191 | 10.1 Lepton efficiency corrections | 77 |
| 1192 | 10.1.1 Muon Trigger SF | 79 |
| 1193 | 10.2 Electron energy scale and resolution | 82 |
| 1194 | 10.3 Muon momentum correction | 82 |
| 1195 | | |
| 1196 | | |
| 1197 | | |
| 1198 | | |

1200 Monte Carlo plays important role in the cross-section measurement. It is constantly being im-
1201 proved, in order to obtain a better precision in data description. Part of these corrections have been
1202 described in Chap. 6. Unfortunately, not everything can be taken into account during simulation
1203 itself. This leads to differences between data and Monte Carlo, that need to be accounted for. There
1204 are two possible methods to correct Monte Carlo without regenerating it. First one is to apply event
1205 weights, so that each MC event can contribute by a non one entry to a histogram. This procedure
1206 is called event reweighting. Second one is a MC smearing. It uses random numbers to alter the
1207 reconstructed 4-vectors.

1208 This chapter describes all additional corrections, that have been applied on MC samples in this
1209 analysis. All of these correction are introducing additional systematic error, that will be discussed in
1210 the Chap. 14.

1211 **10.1 Lepton efficiency corrections**

1212 The efficiency of lepton detection at ATLAS detector can be divided into three components:

- 1213 • The reconstruction efficiency ϵ_{rec} is a probability to reconstruct lepton as a lepton of this
1214 flavor.
- 1215 • The identification efficiency $\epsilon_{id|rec}$ is a probability that a reconstructed lepton survives identi-
1216 fication requirements.
- 1217 • The trigger efficiency $\epsilon_{trig|rec,id}$ is a probability, that the lepton satisfies the trigger requirements.

The full efficiency of a single lepton can be written as:

$$\epsilon_{total} = \epsilon_{rec} \times \epsilon_{id|rec} \times \epsilon_{trig|rec,id} \quad (10.1)$$

1218 All these efficiencies are measured using tag-and-probe method in $Z \rightarrow ll$ decays. One of the
1219 leptons from Z boson, called "probe", is initially selected with all of the cuts, except the one under

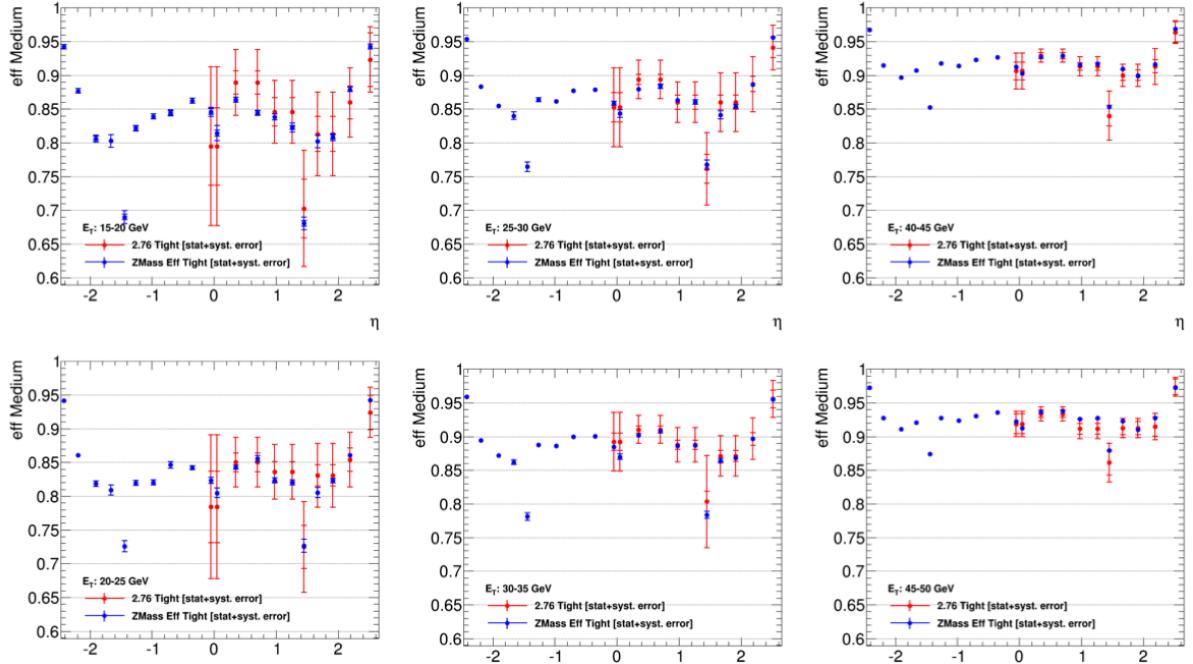


Fig. 10.1: Comparison of electron efficiencies as calculated for 8TeV (blue points) and 2.76TeV (red points) for MC simulation. Efficiencies are shown as a function of pseudorapidity (η) for different electron E_T bins. Both statistical and systematic uncertainties are shown. [78]

study. Second one, called "probe", satisfies tighter selection with some additional cut, e.g. trigger matching.

Considering the event selection for W and Z events, the single lepton efficiencies can be translated to the corresponding efficiencies for the detection of the lepton in W and Z bosons:

$$\epsilon_{total}^W = \epsilon_{rec} \times \epsilon_{id|rec} \times \epsilon_{trig|rec,id} \quad (10.2)$$

$$\epsilon_{total}^Z = \epsilon_{rec} \times \epsilon_{id|rec} \times (1 + (1 - \epsilon_{trig|rec,id})^2) \quad (10.3)$$

The reconstruction efficiency is associated with the algorithm used in the event reconstruction process. This is causing differences between electrons and muons efficiencies. In the electron case the reconstruction efficiency is understood as a probability to reconstruct an electron which has deposited its energy in electromagnetic calorimeter cluster as an electron candidate.

The muon reconstruction efficiency is given by:

$$\epsilon_{reco,muon} = \epsilon_{reco,muon|ID} \cdot \epsilon_{ID} \approx \epsilon_{reco,muon|ID} \cdot \epsilon_{ID|MS}, \quad (10.4)$$

where $\epsilon_{reco,muon|ID}$ is a conditional probability that a muon reconstructed in ID is also reconstructed using MS as a combined muon, and ϵ_{ID} is a probability that the muon is reconstructed as an ID track. This quantity ϵ_{ID} cannot be measured directly in the data and therefore is replaced by $\epsilon_{ID|MS}$ - probability that muon reconstructed in MS is also reconstructed in ID, that can be measured by the tag-and-probe method.

Table 10.1: Muon trigger scale factors

| | SF | SF stat.error |
|---------|-------|---------------|
| μ | 0.988 | 0.011 |
| μ^+ | 1.012 | 0.015 |
| μ^- | 0.964 | 0.015 |

Simulation samples are corrected to match data efficiencies by a scale-factor :

$$SF_{reco,id,trig} = \frac{\epsilon_{reco,id,trig}^{data}}{\epsilon_{reco,id,trig}^{MC}}. \quad (10.5)$$

The scale factors are calculated in P_T^l and η^l bins and have associated statistical and systematic uncertainty components. The statistical component is connected to a size of $Z \rightarrow ll$, which in our case is around 500 events per each lepton flavor. This makes statistical error the dominant one and means that precise calculation of scaling factors based on this data is difficult.

It is possible, however, to use scale factors for 8 TeV 2012 data [78]. The main difference between these data samples are center of mass energy and the pile-up conditions (10 in 2012 and less than 1 in 2013). The effects of these differences have been studied by the electron performance group at ATLAS using $Z \rightarrow ee$ MC sample. Fig. 10.1 shows the scale factors for different P_T^l ranges as a function of η^l , for the MC data produced at 2.76 TeV and 8 TeV centre-of-mass energy. The differences in the scale factors are negligible and fully covered by the statistical errors. This justifies the usage of 8 TeV scaling factors with increased uncertainty (that is considered to be fully statistical) in the analysis at 2.76 TeV.

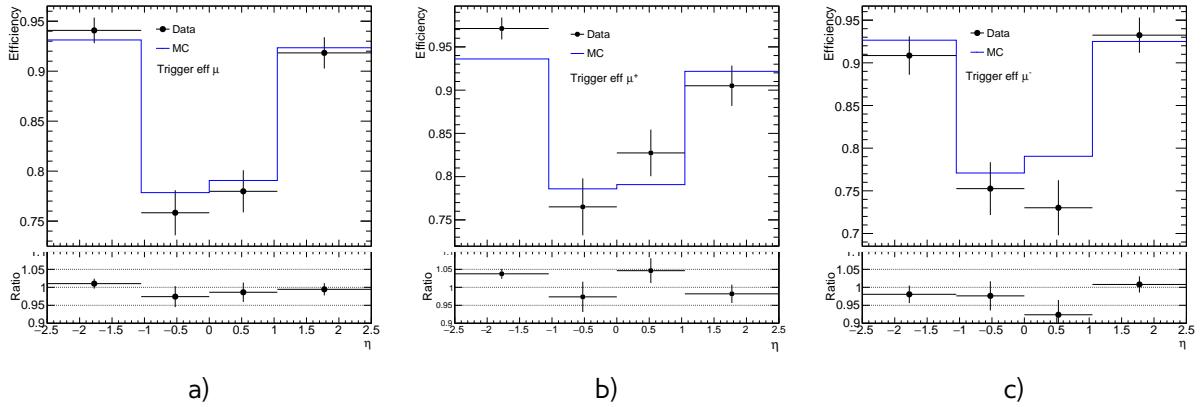
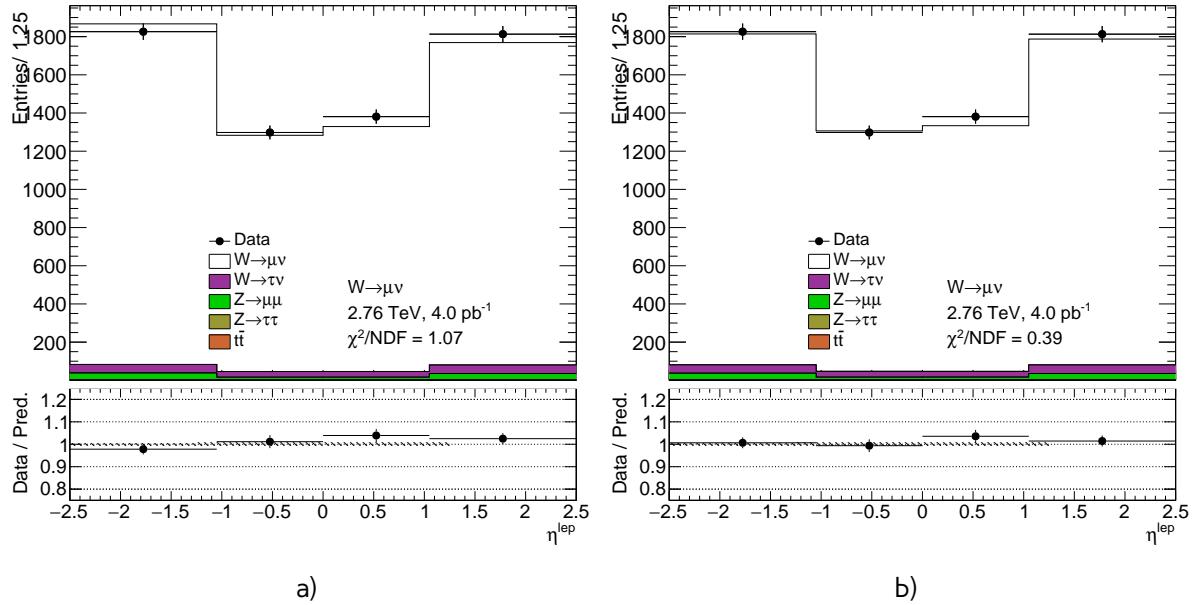
10.1.1 Muon Trigger SF

Unfortunately, single muon trigger has not been present in the 2012 data, so muon trigger scale factor had to be derived from the 2.76 TeV data. The size of the Z sample is not big enough to make the scale factors in P_T^l and η^l bins.

Since the P_T^μ cut is significantly higher, than the trigger threshold, the trigger efficiency in P_T^μ can be considered flat. On the another hand, the η dependence are expected. Binning in η is motivated by the detector construction: $|\eta| < 1.05$ corresponds to a barrel part of the muon spectrometer, while $1.05 < |\eta| < 2.5$ is an end-cap MS (see Sec. 4.2.4). The muon trajectory is bend in a magnetic field. That can lead to small differences in a trigger efficiencies for different muon charges. Possible charge dependency of the scale factors also have been studied.

Total scale factors are presented in Tab. 10.1. Scale factors for μ^+ and μ^- are more than 3σ away from each other, that is a clear indicator of a charge dependency. Trigger efficiencies for data and MC in η bins are shown in a Fig. 10.2.

Effect of applying different scale factors on muon for W analysis is shown on Fig. 10.3 - 10.5. Best agreement with data is achieved by applying single bin scale factor. This, together with the fact that difference in a efficiencies is smaller than 1σ for most of the η bins motivates a choice of single bin charge dependent scale factor for this analysis.


 Fig. 10.2: Trigger scale efficiencies distribution for a) μ b) μ^+ c) μ^- as a function of pseudorapidity

 Fig. 10.3: Muon pseudorapidity distribution from the $W \rightarrow \mu\nu$ selection with a) binned b) not binned charge dependent trigger scale factor applied

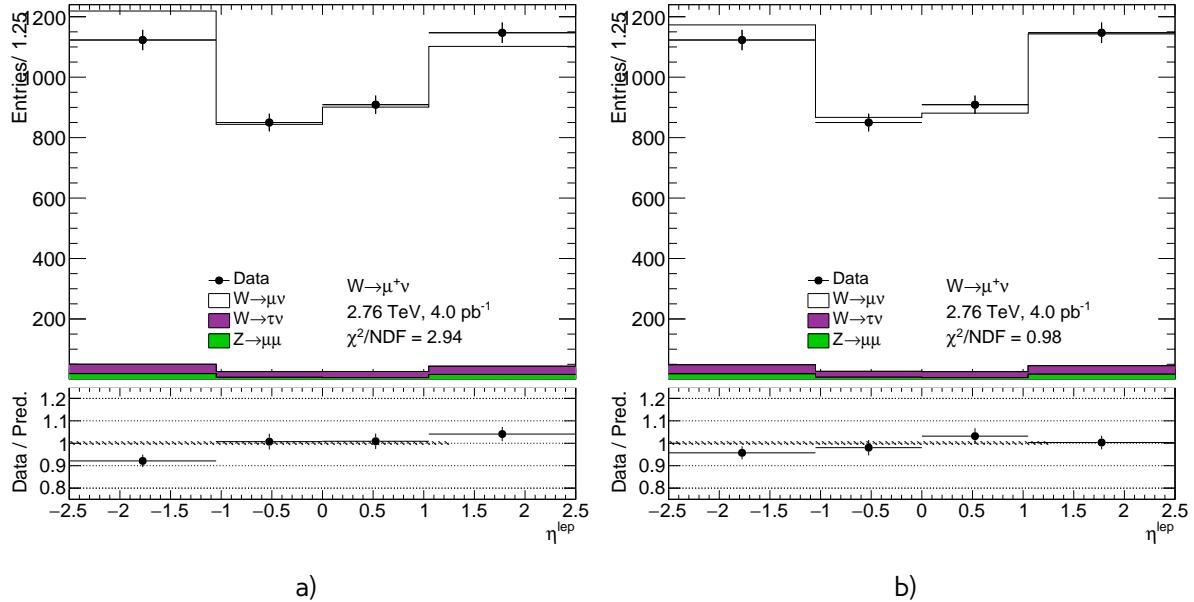


Fig. 10.4: Muon pseudorapidity distribution from the $W \rightarrow \mu^+\nu$ selection with a) binned b) not binned charge dependent trigger scale factor applied

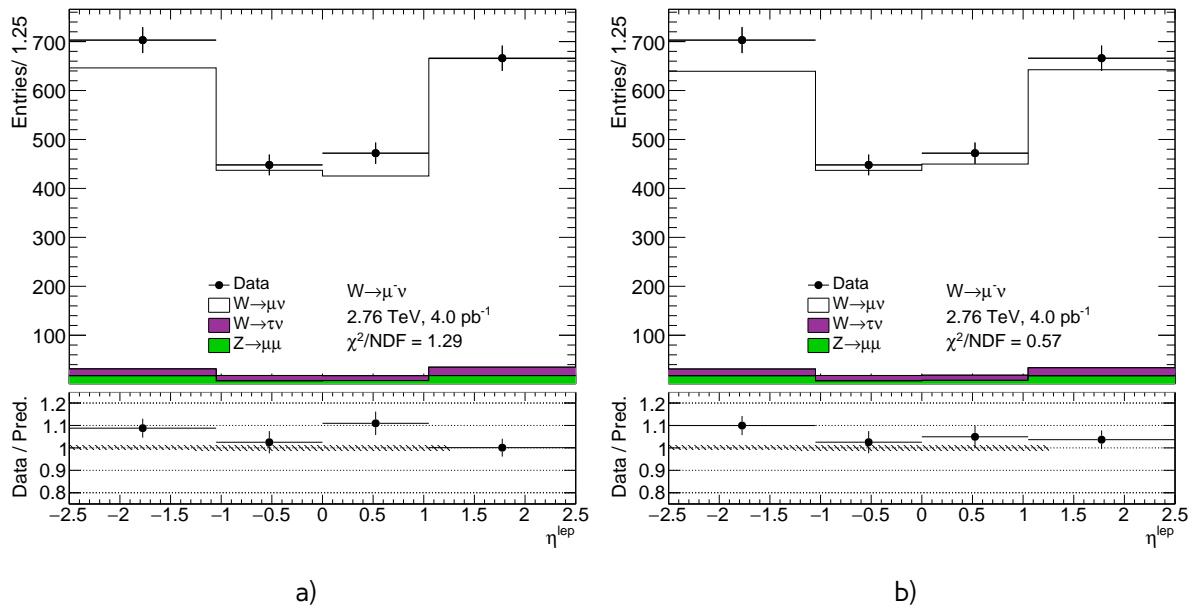


Fig. 10.5: Muon pseudorapidity distribution from the $W \rightarrow \mu^-\nu$ selection with a) binned b) not binned charge dependent trigger scale factor applied

1260 10.2 Electron energy scale and resolution

1261 The energy of the reconstructed electron clusters tends to be shifted in comparison to the true
 1262 energy of the initial electron. The correction of this shift is done in both data and MC as a 3 step
 1263 procedure [79]:

- 1264 • Electronic calibration, that transfers a raw signal from a readout to a cluster energy deposit.
- 1265 • MC based calibration. It corrects the effects of the energy loss in the material in front of the
 1266 calorimeter and the leakage into the hadronic calorimeter. This calibration is applied on both
 1267 data and MC.
- 1268 • Correction of the calorimeter cell response in the data. This allows to get the right response in
 1269 non-optimal HV-regions and to exclude biases in the calorimeter electronics reconstruction.

The energy shift is parameterized, as:

$$E^{data} = E^{MC}(1 + \alpha_i), \quad (10.6)$$

where E^{data} and E^{MC} are the energies in data and simulation, respectively and α_i is a mean shift in a given bin i in η . The effect of this miscalibration on a reconstructed mass of Z boson neglecting second order terms is:

$$m_{i,j}^{data} = m_{i,j}^{MC}(1 + \alpha_{i,j}), \quad \alpha_{i,j} \sim \frac{\alpha_i + \alpha_j}{2}, \quad (10.7)$$

1270 where $m_{i,j}^{data}$ and $m_{i,j}^{MC}$ are the reconstructed mass of the Z boson in i and j bins in η for data and
 1271 MC respectively.

Additionally the difference in the electron resolution has to be corrected. The dependency of electron resolution on its energy was described by Eq. 7.1. It is assumed, that the sampling and the noise terms are well modeled by the MC simulation and the main difference is coming from a constant term. So, the electron resolution correction can be written as:

$$\frac{\sigma_E}{E_i}^{Data} = \frac{\sigma_E}{E_i}^{MC} \oplus c_i \quad (10.8)$$

1272 where c_i is an η dependent relative resolution correction. Similarly to an energy scale correction it
 1273 is possible to derive resolution correction factor by a comparison of $m_{i,j}^{data}$ and $m_{i,j}^{MC}$ distributions.

1274 The correction values α_i and c_i are obtained via the χ^2 fit of an invariant mass of electron pairs in
 1275 data and MC. The resulting energy scale is applied to the data, while the resolution is corrected in
 1276 MC. The resulting scale is validated using the $J/\psi \rightarrow ee$ and $Z/\gamma \rightarrow ee$ samples.

1277 10.3 Muon momentum correction

The muon momentum resolution depends on η , ϕ and p_T of the muon [38]. There is an empirical formula to describe it inside the detector (ID or MS) [80]:

$$\frac{\sigma_{Det}(p_T)}{p_T} = \frac{r_0^{Det}(\eta, \phi)}{p_T} \oplus r_1^{Det}(\eta, \phi) \oplus r_2^{Det}(\eta, \phi) \cdot p_T. \quad (10.9)$$

1278 The first term origins from the fluctuations of the energy loss in the traversed material. The second
 1279 r_1^{Det} term is coming from the magnetic field inhomogenities and the local displacements. The third
 1280 term r_2^{Det} describes the intrinsic resolution effects.

Similarly to electrons, the overall energy scale shift between data and MC is parameterised as:

$$p_T^{data} = p_T^{MC} + s_0^{Det}(\eta, \phi) + s_1^{Det}(\eta, \phi) \cdot p_T^{MC}, \quad (10.10)$$

1281 where $s_0^{Det}(\eta, \phi)$ is coming from the imperfect knowledge of energy losses for muons passing through
 1282 detector.

This leads to a total correction formula:

$$p_T^{Cor, Det} = \frac{p_T^{MC, Det} + \sum_{n=0}^1 s_n^{Det}(\eta, \phi) (p_T^{MC, Det})^n}{1 + \sum_{m=0}^2 \Delta r_m^{Det}(\eta, \phi) (p_T^{MC, Det})^{m-1} g_m}, \quad (10.11)$$

1283 where g_m are normally distributed random variables with the mean 0 and the width 1. Due to a small
 1284 amount of material between the interaction point and the ID, $\Delta r_0^{ID}(\eta, \phi)$ and $s_0^{ID}(\eta, \phi)$ are set to 0.
 1285 The misalignment effect of the MS is corrected on a simulation level by adding a random smearing
 1286 to the alignment constants. This allows to set $\Delta r_2^{MS}(\eta, \phi)$ to 0 during the fit.

The correction factors are extracted using events with $Z \rightarrow \mu\mu$, where muon candidates fulfill the combined muon criteria described in Sec.5.3. For the correction, the invariant mass distributions $m_{\mu\mu}^{ID}$ and $m_{\mu\mu}^{MS}$ are considered individually within a specific $\eta - \phi$ region of the fit. The combined muon parameters are used to obtain angles η and ϕ . The correction extraction is performed first for the ID and then for the MS. Additionally, the following fit variable is considered:

$$\rho = \frac{p_T^{MS} - p_T^{ID}}{p_T^{ID}}, \quad (10.12)$$

1287 which represents the p_T imbalance between the ID and the MS.

In a second step corrections are propagated to the combined momentum, using a weight average:

$$p_T^{Cor, CB} = f \cdot p_T^{Cor, ID} + (1 - f) \cdot p_T^{Cor, MS}, \quad (10.13)$$

1288 where the weight f is derived from the MC.

1289 Chapter 11

1290 Hadronic recoil calibration

| | | |
|------|--|-----|
| 1291 | 11.1 Introduction | 85 |
| 1292 | 11.2 Hadronic recoil resolution correction | 88 |
| 1293 | 11.2.1 Event activity correction | 88 |
| 1294 | 11.2.2 Resolution corrections using Z events | 94 |
| 1295 | 11.3 Hadronic recoil bias correction | 97 |
| 1296 | 11.3.1 Bias determination from M_T^W distribution | 98 |
| 1297 | 11.3.2 Bias determination using u_{\parallel} distribution | 100 |
| 1298 | 11.3.3 Systematic uncertainty estimation | 101 |
| 1299 | 11.4 Summary on hadronic recoil systematics | 101 |
| 1300 | | |
| 1301 | | |
| 1302 | | |
| 1303 | | |
| 1304 | | |

1305 As it was mentioned in Sec. 5.4 due to the discrepancies between data and Monte-Carlo, this
1306 analysis uses a hadronic recoil based algorithm for the missing transverse energy reconstruction.
1307 The missing transverse energy E_T^{miss} affects significantly the W boson measurement, so it is important
1308 to have a solid understanding of possible sources of differences in the hadronic recoil reconstruction
1309 in the data and the Monte-Carlo.

1310 This chapter describes the procedure of the calibration bias and the resolution mismodelling
1311 corrections for a hadronic recoil, that was adapted for the 2.76 TeV data.

1312 11.1 Introduction

1313 This analysis uses a standard hadronic recoil calibration procedure, described in [81], that was mod-
1314 ified and adapted for the low statistics 2.76 TeV case. The standard procedure consist of the 3 main
1315 steps.

1316 The first step in the hadronic recoil calibration procedure aim to correct differences in the pile-up
1317 modeling in the event. Additional interactions can have a significant effect on E_T^{miss} and $\sum E_T$ dis-
1318 tributions. These discrepancies are usually corrected by reweighting average number of interactions
1319 per bunch crossing in MC to match the data. However, the ATLAS simulation is adjusted for high
1320 pile-up runs, so this quantity is modeled discrete in case of 2.76 TeV analysis (Fig. 11.1), what makes
1321 the precise reweighting impossible. However, since the mean number is below 1, effect of these
1322 discrepancies on E_T^{miss} distributions can be neglected.

1323 In the second and the third steps possible discrepancies in the resolution and the scale of the
hadronic recoil are corrected. The hadronic recoil algorithm performance can be studied in MC
through the projection of $\vec{H}\vec{R}$ on the direction of the transverse momentum of the vector boson, as

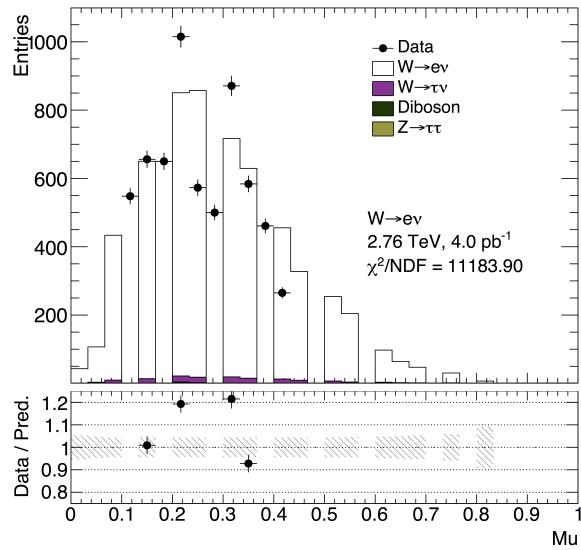


Fig. 11.1: Mean number of interactions per bunch crossing in $W \rightarrow e\nu$ events. In MC the pileup is modeled in a few bins only, that makes the application of the standard data to MC reweighting procedure not feasible.

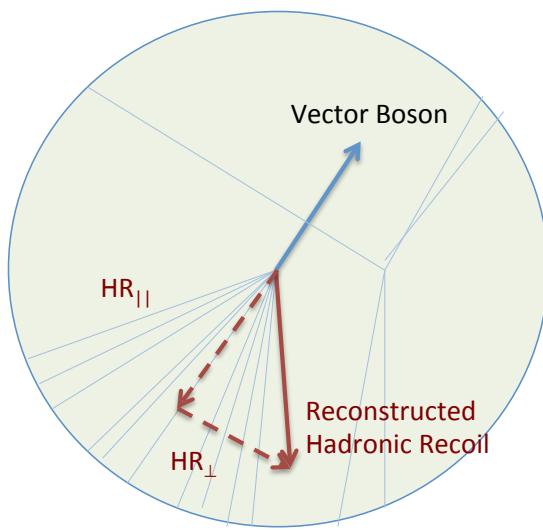


Fig. 11.2: Parallel and perpendicular projections of the hadronic recoil with respect to the transverse momentum of the vector boson [49]

shown in Fig. 11.2. This projection can be divided into perpendicular u_{\perp} and parallel u_{\parallel} components as follows:

$$u_{\parallel} = \vec{v}_{xy} \cdot \vec{HR} \quad (11.1)$$

$$u_{\perp} = v_x \cdot HR_y - v_y \cdot HR_x, \quad (11.2)$$

where \vec{v}_{xy} is a unitary vector along the transverse component of a vector boson momentum and v_x and v_y are its projections on x and y axis respectively. In the case of the true kinematics $u_{\parallel} = -p_T^{bos}$ and $u_{\perp} = 0$. However the limited calorimeter resolution is causing relatively wide distributions for these projections. The parallel component u_{\parallel} is sensitive to a possible bias in the hadronic recoil, while the perpendicular u_{\perp} can be used for determination of the resolution discrepancies. The mean and the width of these distributions can depend on different variables, such as a mean number of interactions in an event, hadronic activity, boson P_T^{bos} etc.

It is convenient to use Z boson decays for a hadronic recoil calibration, since its transverse momentum P_T^Z can be determined not only from the hadronic recoil, but also from its decay products. The P_T^Z resolution coming from a lepton reconstruction is 3-4 times more precise, than the one extracted from a hadronic recoil. This allows to treat leptonically reconstructed P_T^Z as a reference P_T of the boson and to compare directly u_{\perp} and u_{\parallel} in the data and the MC. However, a small size of the Z sample in the 2.76 TeV data leads to a high statistics error for these distributions.

The calibration constants can also be derived from W boson decays. In order to exclude a possible bias from the P_T^W mismodelling these calibration constants are derived through the data vs MC comparison of P_T^W independent distributions (such as M_T^W).

In this analysis a combined procedure based on Z and W bosons decays has been used for a hadronic recoil calibration.

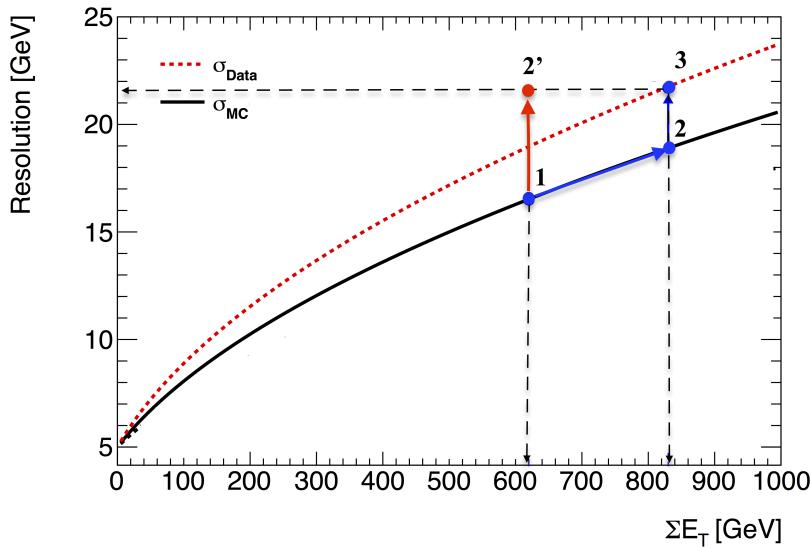


Fig. 11.3: Schematic view of the correction procedure: this figure illustrates the resolution of u_{\perp} as a function of event activity $\sum E_T$. The dotted curve represents data resolution (σ_{data}), the solid black one depicts a nominal MC resolution (σ_{MC}). Blue line from point 1 to point 2 corresponds to the $\sum E_T$ correction discussed in Sec. 11.2.1. The red line from point 1 to point 2' corresponds to a direct correction of the resolution mismodelling discussed in Sec. 11.2.2. Modified from [81]

11.2 Hadronic recoil resolution correction

The event activity plays an important role in the E_T^{miss} reconstruction. Since $\sum E_T$ and the hadronic recoil resolution values are correlated, the possible mismodelling of the event activity can lead to differences between the data and the Monte-Carlo E_T^{miss} resolutions. There are two ways to correct the resolution in the 2.76 TeV data (Fig. 11.3):

- A two step procedure, shown as path 1-2-3 in Fig. 11.3. The first step is to correct $\sum E_T$ distribution to match the data using reweighting of the events. Remaining differences in resolution are corrected at the second step. This method is discussed in Sec. 11.2.1.
- The second order effects on E_T^{miss} coming from $\sum E_T$ modelling are neglected and the resolution differences between data and MC corrected directly. This procedure is shown as the path 1-2' in Fig. 11.3 and described in Sec. 11.2.2.

Both of these methods have been implemented and will be described next.

11.2.1 Event activity correction

The distributions of the event activity $\sum E_T$ are shown in Fig. 11.4. There is a visible shift between data and the MC distribution for both W boson channels. The standard procedure, used in the M_T^W measurement at 7 TeV uses a Smirnov transformation of MC, determined from the $\sum E_T$ and P_T^{bos} distributions in Z events [81]. Distribution of the event activity in the Z-boson events are shown in Fig. 11.5. Since the value of χ^2/NDF is around 1, it is clear that size of the Z sample is not

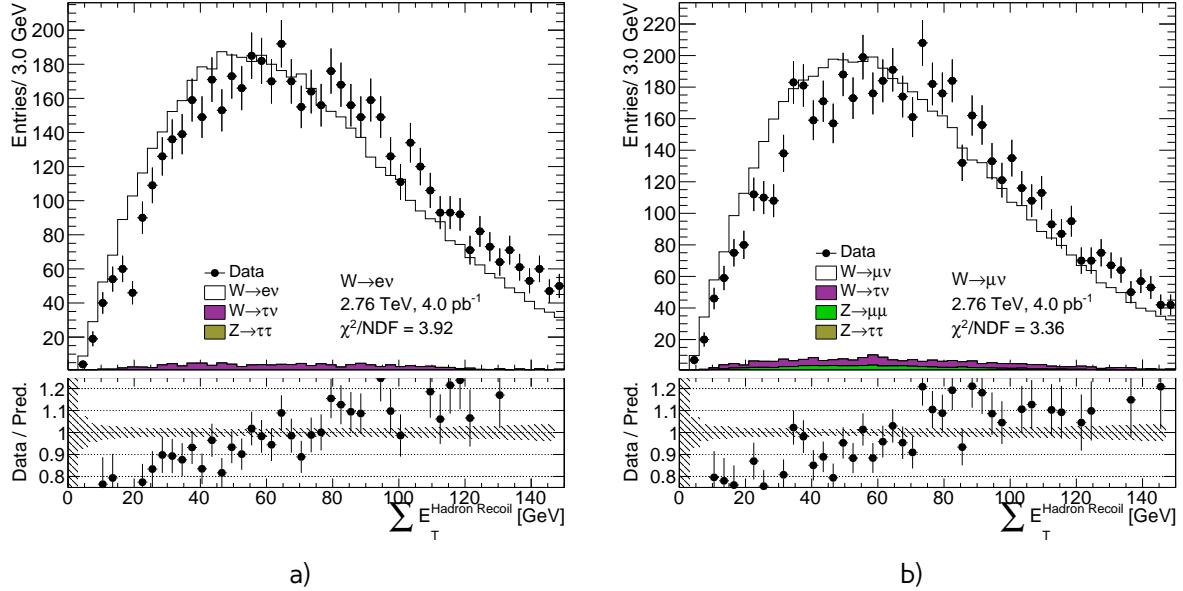


Fig. 11.4: Event activity $\sum E_T$ distribution from a) the $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection. There is a clear sign of the event activity mismodelling in both channels, that should be corrected.

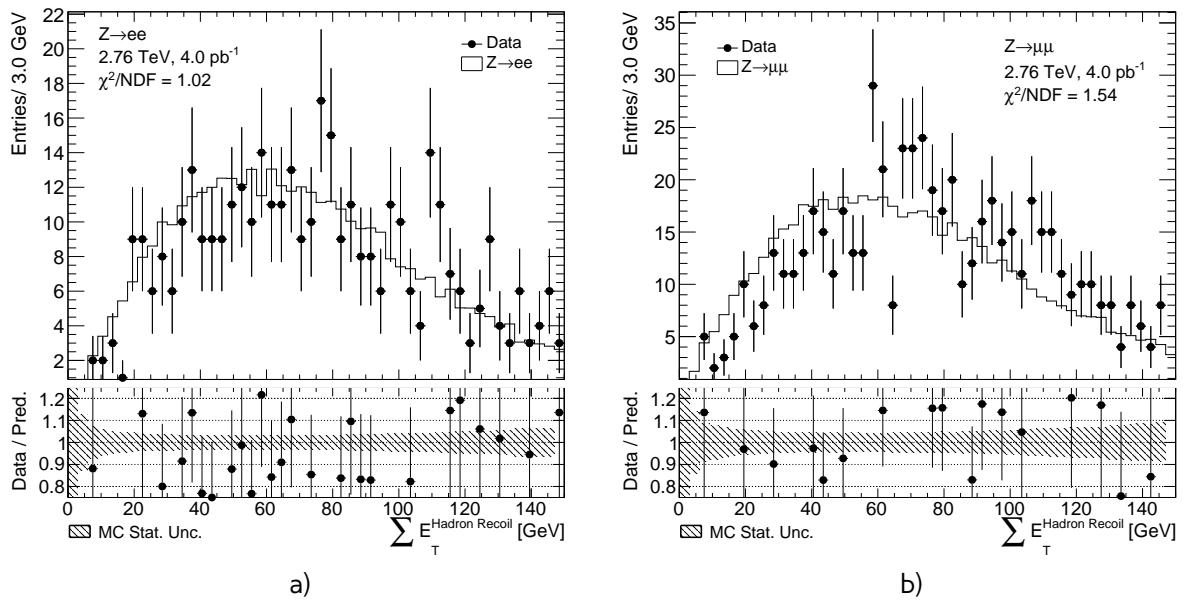


Fig. 11.5: Event activity $\sum E_T$ distribution from a) the $Z \rightarrow ee$ selection and b) the $Z \rightarrow \mu\mu$ selection. The size of the Z sample in 2.76 TeV data is insufficient to correct the mismodelling of the event activity.

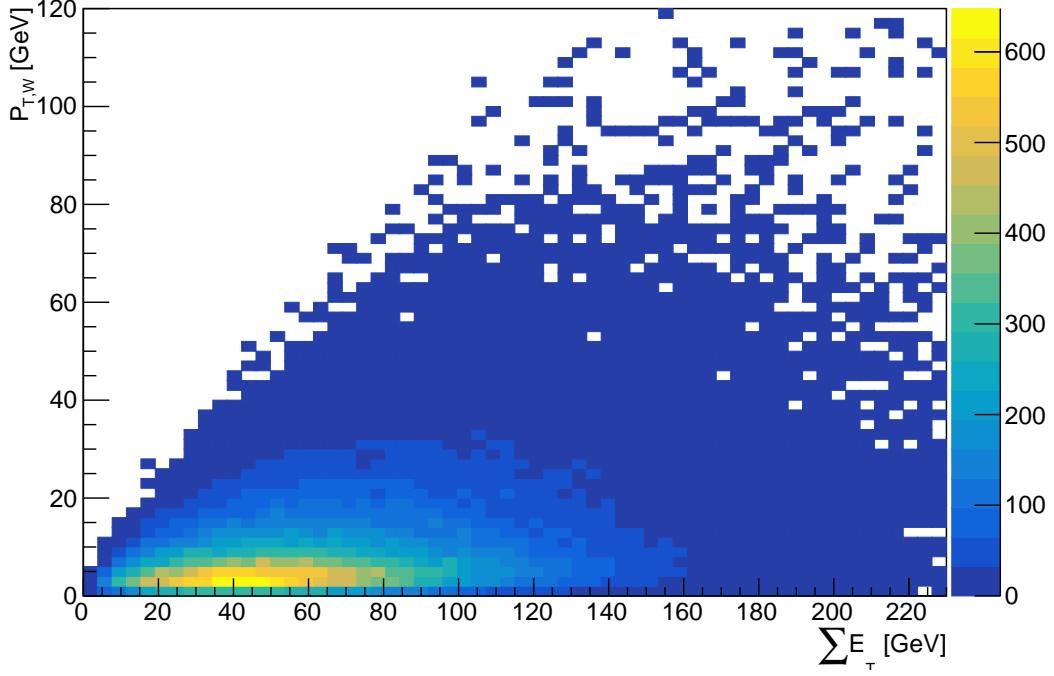


Fig. 11.6: Distribution of event activity $\sum E_T$ vs truth transverse momentum of the W boson P_T^W *truth* in the $W^+ \rightarrow e\nu$ MC sample.

1359 sufficient for this procedure for both $\sum E_T$ and P_T^{bos} distributions. This motivates a choice of the
 1360 $\sum E_T$ reweighting constants determination from the W-boson sample.

The event activity $\sum E_T$ is correlated to the truth transverse momentum of the boson, as shown in Fig. 11.6, so in order to avoid a bias from changing P_T^W spectrum, reweighting constants are derived in bins of reconstructed boson momentum $P_T^{W,rec}$. Inside each $P_T^{W,rec}$ bin the reweighting constants are calculated as:

$$SF^{channel} = \frac{\sum E_T^{data,selection}}{\sum E_T^{MC,no\ cuts}}, \quad (11.3)$$

1361 where $\sum E_T^{data,selection}$ is a $\sum E_T$ distribution inside a given $P_T^{W,rec}$ after the full event selection. In order
 1362 to reduce systematic uncertainty from this value, a combination of $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ events
 1363 is used.

1364 Second term $\sum E_T^{MC,no\ cuts}$ stands for $\sum E_T$ distribution in MC before any selection. The scale factors
 1365 are determined separately for each signal MC for W boson decays, in order to leave the total number
 1366 of events in the the MC after the correction unchanged. Transverse boson momentum binning is
 1367 chosen so that there is approximately the same number of events per bin. The total number of
 1368 $P_T^{W,rec}$ bins is 6. The scale factor are applied as a reconstructed weight on MC.

1369 The correction factors for two example $P_T^{W,rec}$ bins are shown in Fig.11.7. Resulting reweighting
 1370 constants for $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ MC samples are shown in Fig. 11.8. This method allows to
 1371 leave the reconstructed transverse momentum of the boson nearly unmodified and introduces only
 1372 a small change in a the truth boson spectrum, as shown in Fig. 11.10.

1373 There are two possible sources of the uncertainties of this correction: systematical, coming from

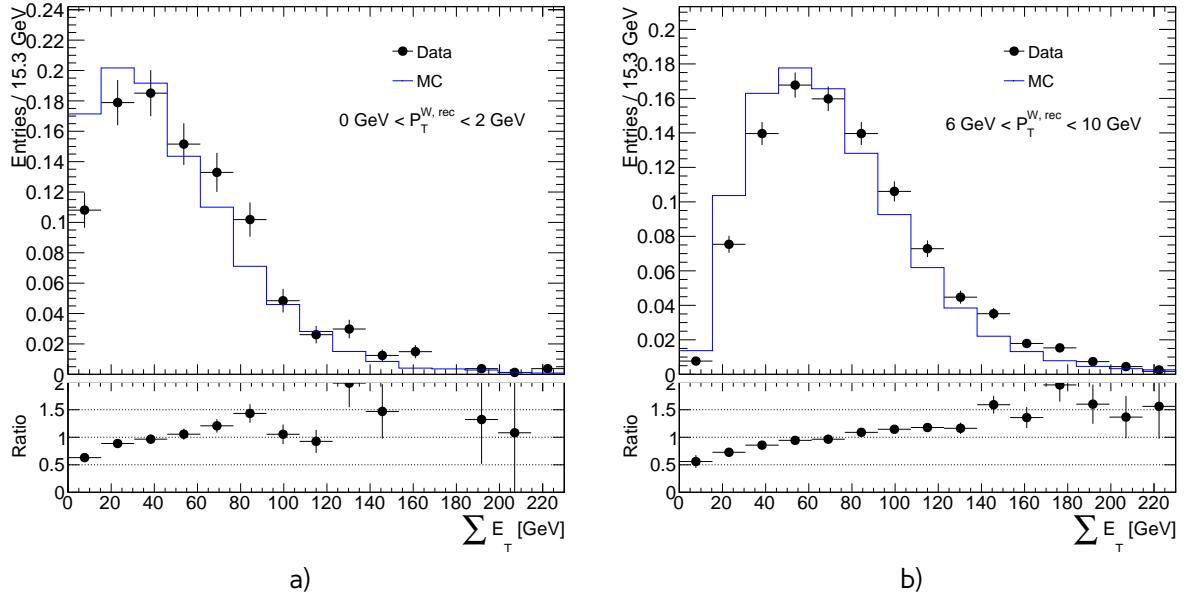


Fig. 11.7: Distribution of $\sum E_T$ for the different $p_T^{W,\text{rec}}$ bins for $W^+ \rightarrow e\nu$ MC sample

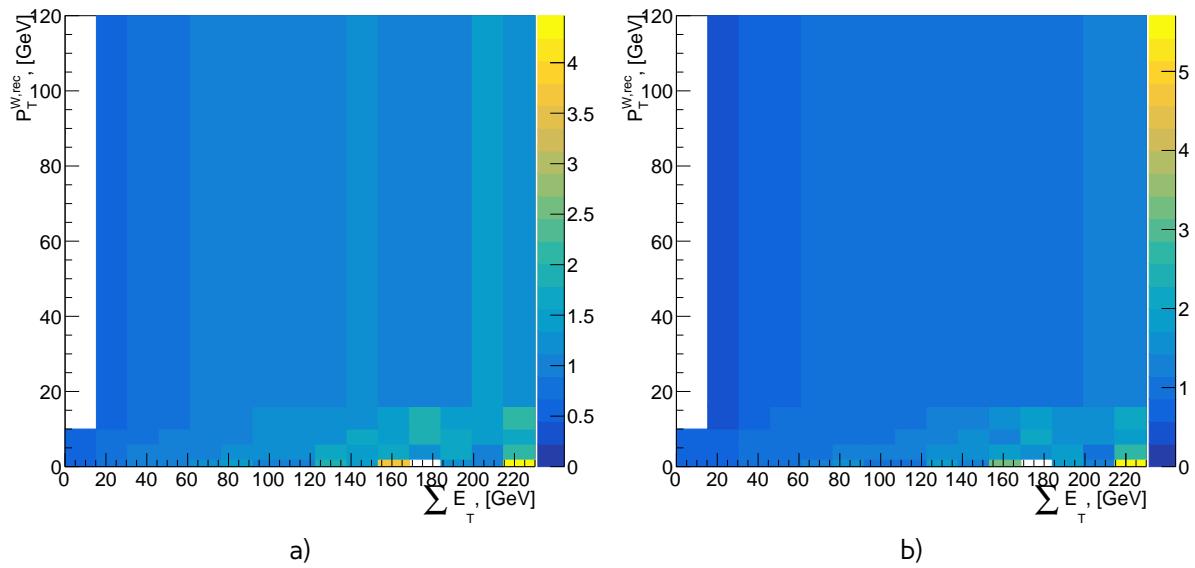


Fig. 11.8: Distribution of $\sum E_T$ reweighting constants derived for a) $W^+ \rightarrow e\nu$ and b) $W^+ \rightarrow \mu\nu$ MC sample.

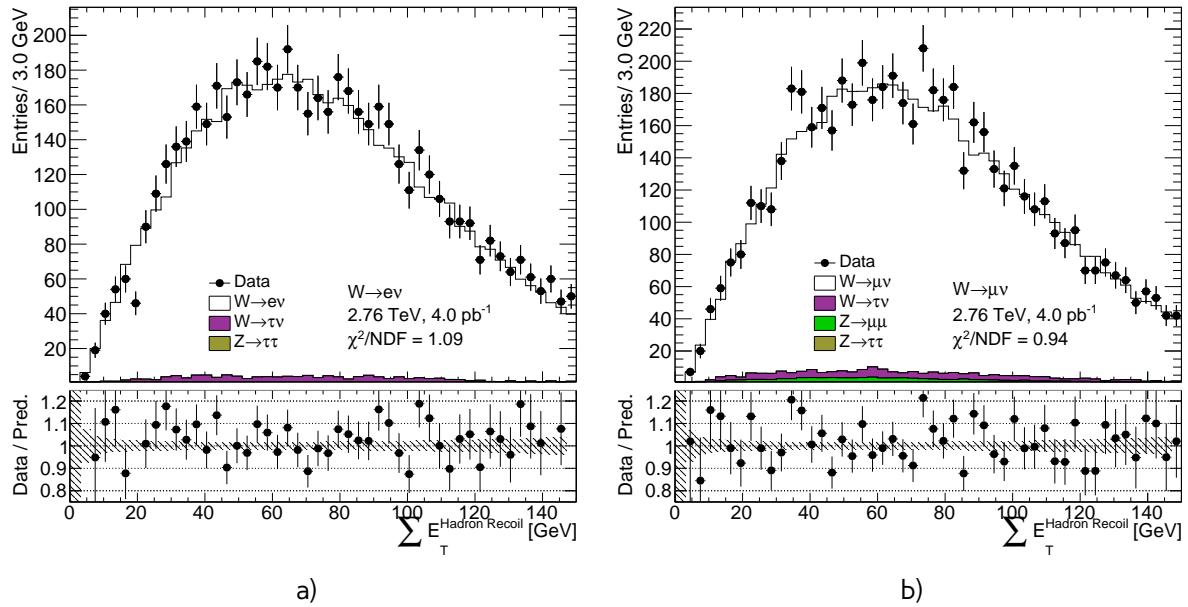


Fig. 11.9: Event activity $\sum E_T$ distribution from a) the $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection after $\sum E_T$ correction.

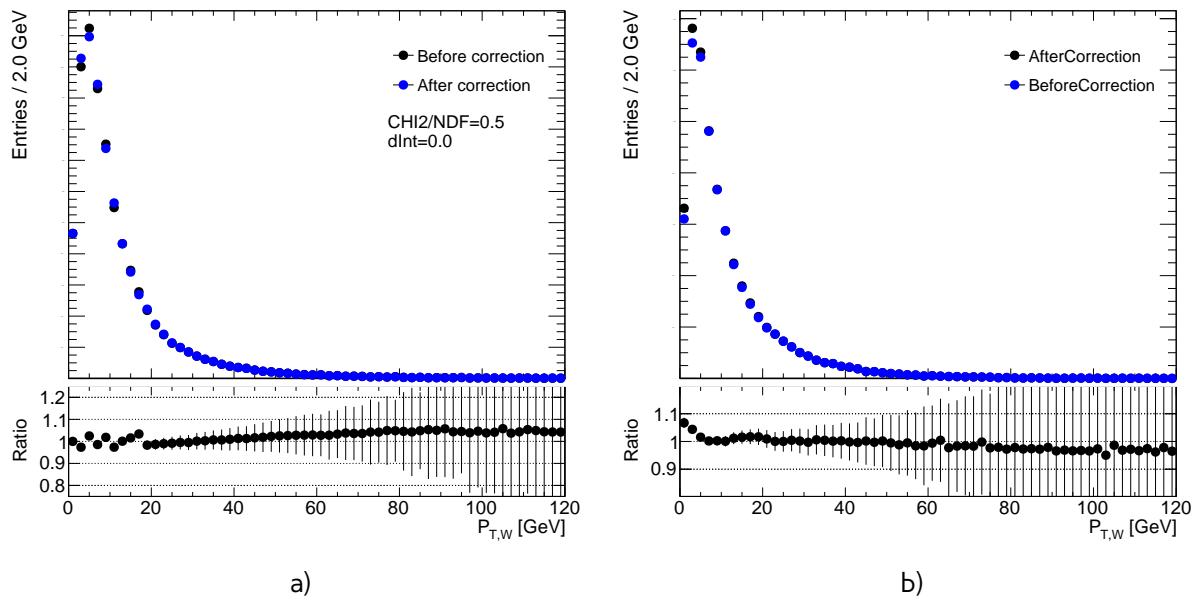


Fig. 11.10: Effect of the $\sum E_T$ reweighting on a) reconstructed transverse momentum of the boson and b) truth transverse momentum of the boson in $W^+ \rightarrow e\nu$ MC sample.

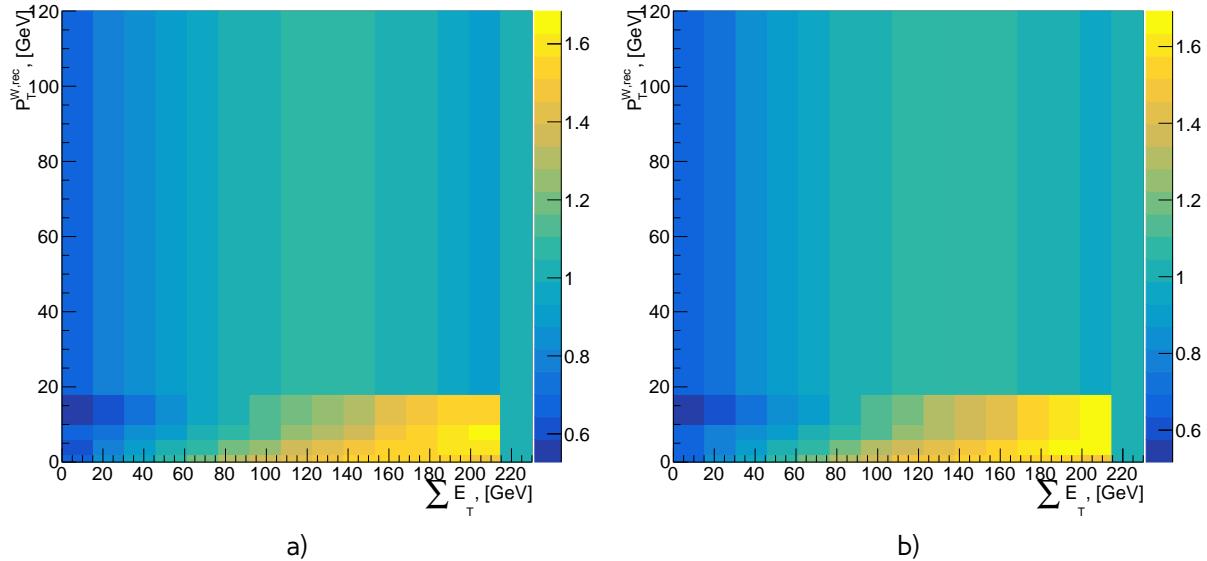


Fig. 11.11: Distribution of $\sum E_T$ reweighting constants derived for a) $W^+ \rightarrow ev$ and b) $W^+ \rightarrow \mu\nu$ MC sample using polynomial order 2 approximation.

the method itself and a statistical, coming from the limited data statistic. The methods of their determination and an effect of the correction on C_W factors will be discussed.

1376 Systematic error estimation

Systematic error on this reweighting can be estimated approximating the data to MC ratio as a function of $\sum E_T$ inside each $P_T^{W,rec}$ bin with a polynomial degree 2 or 1. This method allows also to neglect effects of the data fluctuations, especially for the high $\sum E_T$ regions with low statistics, as it could be seen in Fig. 11.8. Because of the low statistics for $\sum E_T > 220$ GeV the ratio in the last bins has been set to 1 and this region haven't been included in the polynomial fit. The total reweighting constants obtained from this procedure are shown in Fig. 11.11.

1383 Statistical error estimation

1384 Statistical error on the $\sum E_T$ reweighting is estimated using Toy MC method, described in Chap.14
 1385 from the polynomial order 2 approximation, since the uncertainty of the fit parameters obtained
 1386 from the fit includes statistical error of data points. The fit parameters of the polynomials are varied
 1387 inside each $p_T^{W,rec}$ bin within their fit uncertainties as in Eq. 14.3.

Because of the possible correlations between the fit parameters, a multivariate gaussian distribution has been used. It is calculated as:

$$p(x; \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right), \quad (11.4)$$

where $\mu \in \mathbf{R}^n$ is obtained from the fit parameters vector and Σ is $n \times n$ covariance matrix of these parameters, $(x - \mu)^T$ is a transpose of the vector $(x - \mu)$. In case of the polynomial order two $n=3$.

Table 11.1: Effect of $\sum E_T$ correction on C_W for different channels and methods

| Channel | δC_W no approximation | δC_W polynomial order 2 | δC_W polynomial order 1 | δC_W Toy MC |
|-----------------------------|----------------------------------|------------------------------------|------------------------------------|------------------------|
| $W^+ \rightarrow e^+ \nu$ | 0.48% | 0.39% | 0.31% | 0.03% |
| $W^- \rightarrow e^- \nu$ | 0.49% | 0.33% | 0.22% | 0.03% |
| $W^+ \rightarrow \mu^+ \nu$ | -0.27% | -0.20% | -0.28% | 0.03% |
| $W^- \rightarrow \mu^- \nu$ | -0.29% | -0.21% | -0.27% | 0.03% |

1390 For statistic error determination total number of 25 toys have been used. Total error is calculated
 1391 using Eq.14.4.

1392 Effect of the $\sum E_T$ correction on cross-section

1393 The effect of the $\sum E_T$ correction on cross-section is estimated by applying different correction
 1394 factors on MC. The error is estimated by calculating difference in C_W using On/Off method (see
 1395 Chap. 14). The overall effect of the $\sum E_T$ correction for different methods is summarized in Tab. 11.1.
 1396 Statistical error, estimated using Toy MC method is negligible. The systematic error is calculated as
 1397 a difference between C_W for a two methods and is considered to be small, compared to the overall
 1398 effect.

1399 The sign of the effect differs for different W channels, that cannot be explained by a systematic
 1400 error coming from the method or a data statistics. This effect also cannot be explained from a
 1401 physical point of view, since we expect a similar errors for both lepton flavors, so it was decided not
 1402 to use this corrections in a final analysis.

1403 11.2.2 Resolution corrections using Z events

1404 Another way to check resolution effects is to study u_\perp and $u_\parallel - p_T^Z$ distributions in events containing
 1405 Z boson. This correction assumes, that any resolution mismodelling reflects discrepancies in the $\sum E_T$
 1406 distribution, while the difference in the resolution at a given $\sum E_T$ is subleading.

Difference in hadronic recoil resolutions $d\sigma$ between the data and the MC can be quantified by value:

$$d\sigma = \sqrt{\sigma_{data}^2 - \sigma_{MC}^2}, \quad (11.5)$$

where σ_{data} and σ_{MC} are the RMS of these distributions. This value is affected by the statistical uncertainty of data standard deviation, that in case of the distributions close to normal can be calculated as [82]:

$$\sigma(\sigma_{data}) = \frac{\sigma_{data}}{\sqrt{2N}}, \quad (11.6)$$

1407 where N is the number of entries in histogram. The distributions of u_\perp and $u_\parallel - p_T^Z$ in $Z \rightarrow ee$,
 1408 $Z \rightarrow \mu\mu$, $Z \rightarrow ll$ events are shown in Fig. 11.12. The typical resolution uncertainty for data is
 1409 around 0.1 GeV for all distributions, while the the difference in resolution is 1.0 GeV and higher, that
 1410 is a clear indication of mismodelling of hadronic recoil resolution in the Monte-Carlo. The overall
 1411 difference in resolutions is consistent between u_\perp and $u_\parallel - p_T^Z$ distributions, however it depends

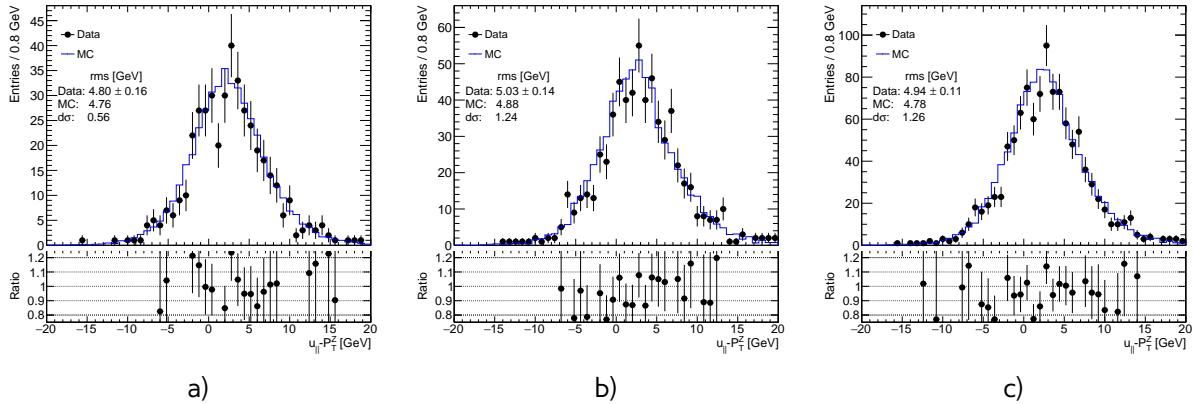


Fig. 11.12: Parallel hadronic recoil component distribution from a) the $Z \rightarrow ee$ selection b) $Z \rightarrow \mu\mu$ selection and c) $Z \rightarrow ll$ selection. The expected contribution from signal is estimated with Monte Carlo simulation, other background sources are considered negligible.

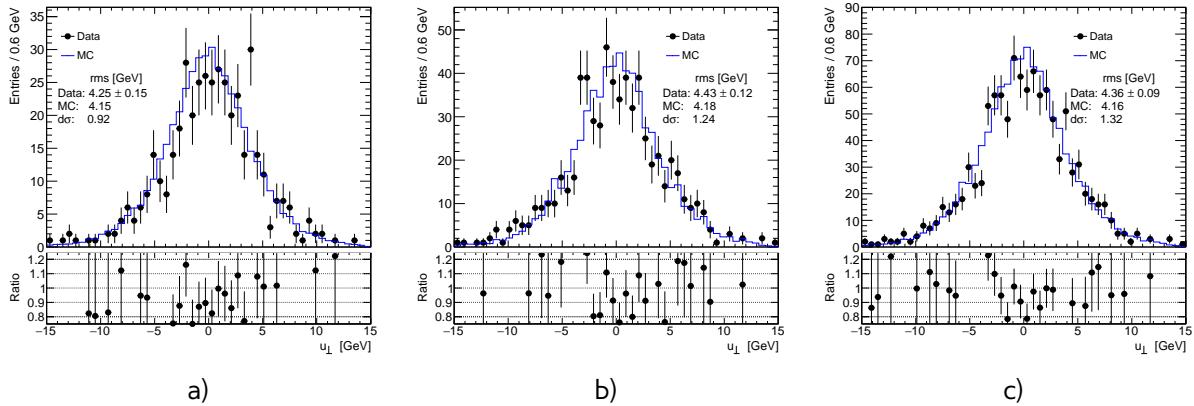


Fig. 11.13: Perpendicular hadronic recoil component distribution from a) the $Z \rightarrow ee$ selection b) $Z \rightarrow \mu\mu$ selection and c) $Z \rightarrow ll$ selection. The expected contribution from signal is estimated with Monte Carlo simulation, other background sources are considered negligible.

1412 on a lepton flavor, isolation and identification criteria. For a correction it was decided to choose the
 1413 value, obtained from combined $Z \rightarrow ll$ channel, where $d\sigma = 1.3$ GeV.

The resolution is corrected by a smearing, using a Gaussian distribution, of each component of the hadronic recoil in Monte-Carlo:

$$u'_{||} = u_{||} + Gaus(0, d\sigma) \quad (11.7)$$

$$u'_{\perp} = u_{\perp} + Gaus(0, d\sigma), \quad (11.8)$$

1414 Effect of the smearing correction on cross-section

1415 Effect of smearing correction is estimated using On/Off method (Chap. 14) on a C_W factor. Scan in
 1416 a big range of the parameter $d\sigma$ up to 2.0 GeV (Fig. 11.14) have showed, that C_W becomes smaller
 1417 with growth of the smearing parameter $d\sigma$. However, due to the random nature of the correction,
 1418 the C_W fluctuates within the mean value.

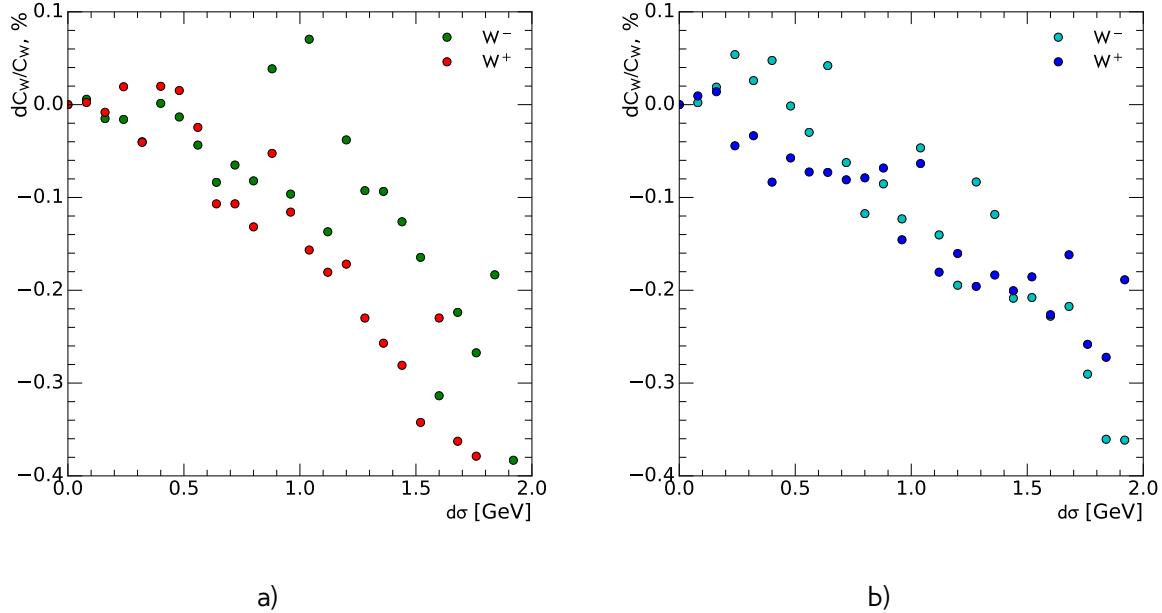


Fig. 11.14: Effect on a C_W from hadronic recoil resolution correction with different $d\sigma$ for a) $W \rightarrow e\nu$
 b) $W \rightarrow \mu\nu$ channel.

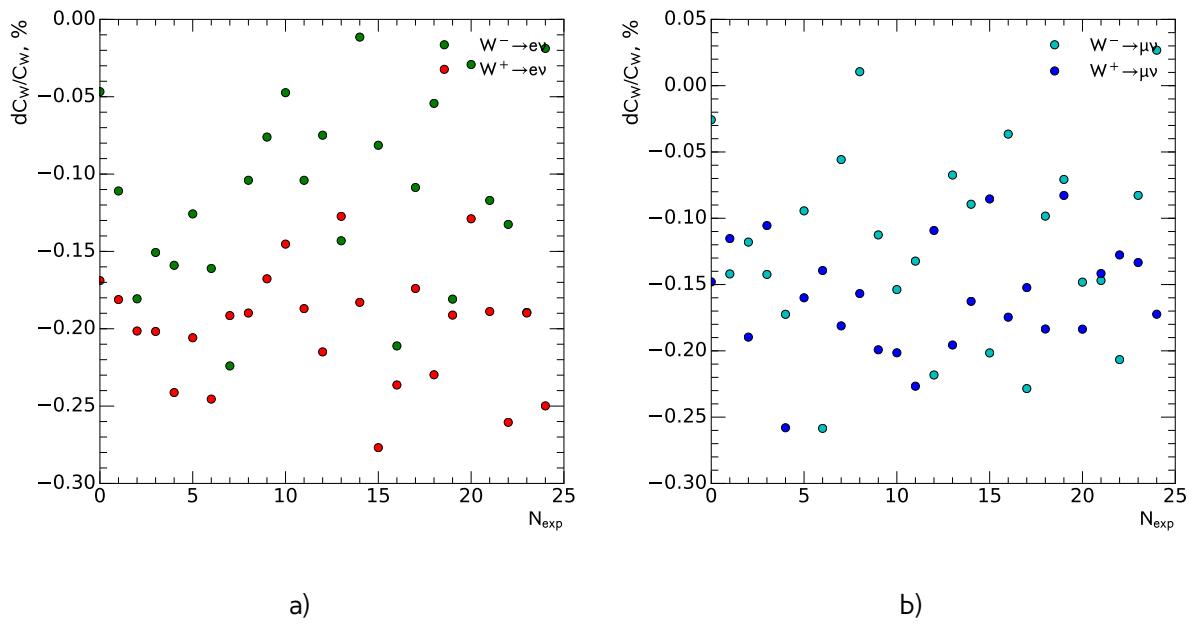


Fig. 11.15: Effect on a C_W from hadronic recoil resolution correction with $d\sigma = 1.3$ GeV for a) $W \rightarrow e\nu$
 b) $W \rightarrow \mu\nu$ channel for repeated experiments. The overall systematic uncertainty of this
 correction is estimated as the mean value of dC_W .

Table 11.2: Effect of smearing correction on a C_W for a different channels. The statistical error (noted *stat.err.*) of the mean value is estimated using the Eq. 11.11

| Channel | $\delta C_W \pm \text{stat,err.}$ | rms |
|-----------------------------|-----------------------------------|-------|
| $W^+ \rightarrow e^+ \nu$ | -0.20±0.01% | 0.04% |
| $W^- \rightarrow e^- \nu$ | -0.11±0.01% | 0.06% |
| $W^+ \rightarrow \mu^+ \nu$ | -0.16±0.01% | 0.04% |
| $W^- \rightarrow \mu^- \nu$ | -0.12±0.01% | 0.07% |

Systematic error have been estimated by repeating correction on the same sample 25 times (Fig. 11.15). Table 11.2 presents the mean effect on C_W together with the rms of the distribution. Overall systematic effect is below 0.2% for each analysis channel, that makes it negligible compared to the statistics uncertainty in W samples (Chap. 14).

11.3 Hadronic recoil bias correction

As it was mentioned before, the hadronic recoil value in Monte Carlo could be shifted in a respect to data because of the mismodelling of underlying event, calorimeter cluster responses, etc. Since the value of hadronic recoil affects the E_T^{miss} distribution this discrepancy should be corrected. It could be done by applying the correction factor HR_{SF} on a hadronic recoil in Monte-Carlo sample as:

$$u_{\parallel}^{\text{cor}} = u_{\parallel} \cdot HR_{SF}, \quad (11.9)$$

where u_{\parallel} is a parallel component in the respect to the true boson direction of hadronic recoil.

The procedure of hadronic recoil bias determination uses a parameter scan through the wide range of the possible HR_{SF} values. It is assumed, that the "real" value of the bias is corresponding

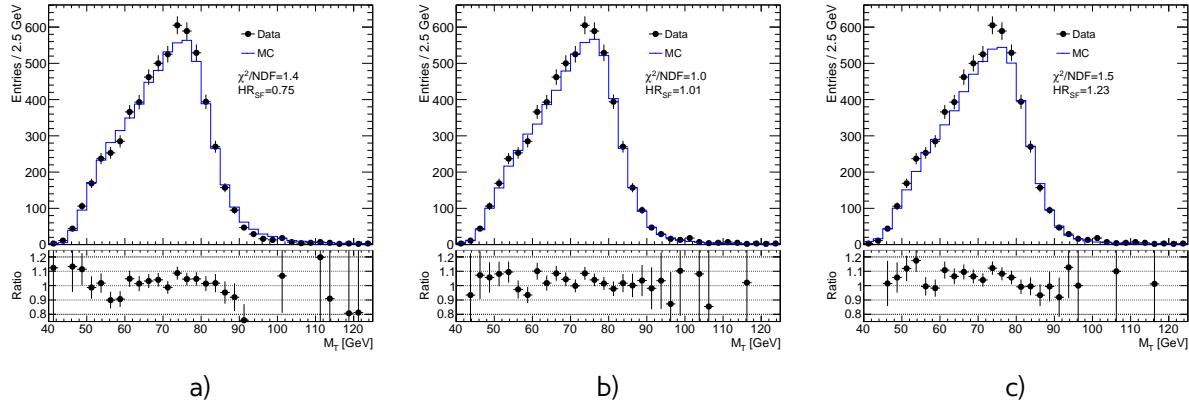


Fig. 11.16: Transverse mass distribution from the $W \rightarrow e\nu$ selection for different hadronic recoil scales: a) $HR_{SF}=0.75$ b) $HR_{SF}=1.1$ c) $HR_{SF}=1.23$. The expected contributions from signals and backgrounds are estimated with Monte Carlo simulation, except for a QCD background, that is not included.

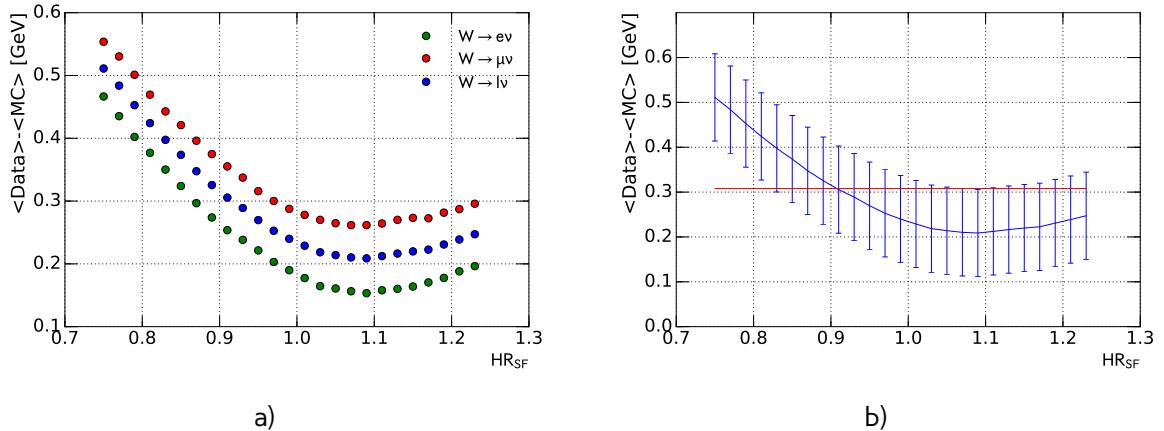


Fig. 11.17: Distribution of a difference in a mean transverse mass $\langle M_T^W \rangle$ between data and MC as a function of the hadronic recoil scale HR_{SF} a) for different W boson channels and b) for combined $W \rightarrow l\nu$ selection. Errors for each point are calculated as a standard error of mean (Eq. 11.11). Below red line is the 68% CL on the best HR_{SF} correction factor. The expected contributions from signals and backgrounds are estimated with Monte Carlo simulation, except for a QCD background, that is not included.

to a best agreement between data and MC and therefore can be obtained through the fit of χ^2 of some distribution as:

$$\chi^2 = \frac{(HR_{SF} - sf_{best})^2}{\sigma_{sf}^2} + \chi_0^2, \quad (11.10)$$

where sf_{best} is the best scale factor, σ_{sf} is a statistical error of this parameter and χ_0^2 is a value of χ^2 in a minimum.

In the following sections methods of hadronic recoil bias determination using W and Z events will be discussed.

11.3.1 Bias determination from M_T^W distribution

Since the W boson transverse momentum cannot be measured in two different ways in order to provide the reference for a hadronic recoil scale, determination of the hadronic recoil bias should use the distributions, that are not sensitive to the true P_T^W spectrum, to exclude the effect of possible P_T^W mismodelling in MC. One of the optimal choices is the M_T^W distribution.

The transverse mass distribution for a different correction parameters HR_{SF} is shown on a Fig. 11.16. The expected contributions from signal and backgrounds are estimated with Monte Carlo simulation, except for a multijet background, because its shape and number of events depends on a hadronic recoil scale and thus needs to be recalculated for each value of HR_{SF} .

One of the possible methods to determine the correction factor is to use a difference in the mean of the transverse mass distributions in data and MC (Fig. 11.17). Statistical error on a correction factor is considered a dominating one and estimated as a standard error of a mean $\sigma(\langle M_T^W \rangle)$, calculated as:

$$\sigma\left(\langle M_T^W \rangle\right) = \frac{\sigma(M_T^W)}{\sqrt{N}}, \quad (11.11)$$

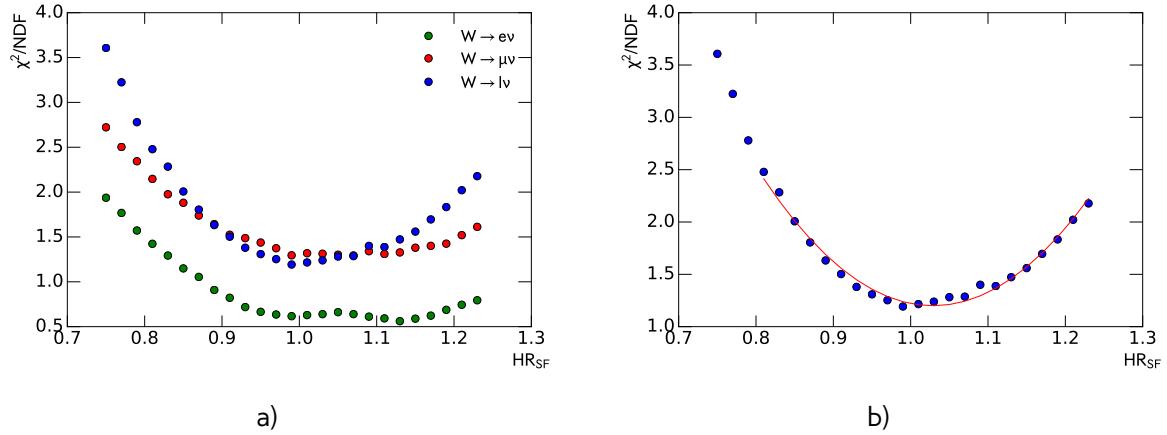


Fig. 11.18: Distribution of χ^2/NDF between data and MC for transverse mass $\langle M_T^W \rangle$ as a function of hadronic recoil scale HR_{SF} a) for different W boson channels. b) for combined $W \rightarrow l\nu$ selection. Fit result is shown by the red line. The expected contributions from signal and backgrounds are estimated with Monte Carlo simulation, except for a QCD background, that is not included.

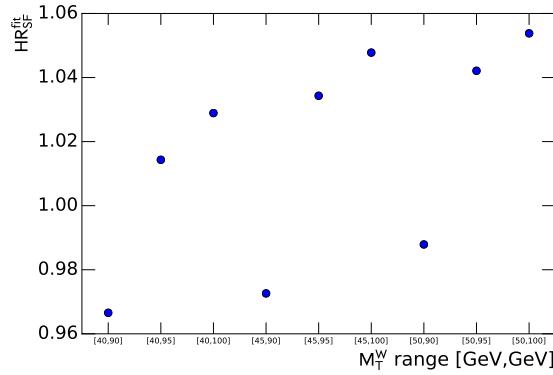


Fig. 11.19: Values of hadronic recoil biases obtained from the fit for events from combined $W \rightarrow l\nu$ selection as a function of fit range. The expected contributions from signals and backgrounds are estimated with Monte Carlo simulation, except for a QCD background, that is not included.

1438 where $\sigma(M_T^W)$ is a standard deviation of M_T^W distribution and N is a total number of events used.
 1439 The minimal difference is obtained at $HR_{SF} = 1.1 \pm 0.2$. The precision of this method is low, and it
 1440 is mainly used as a cross-check for other methods.

1441 Distribution of χ^2 for a scan of possible values of HR_{SF} for different W channels is shown in
 1442 a Fig. 11.18 a). Because of a possible mismodelling of the tail M_T^W distribution, events with $M_T^W >$
 1443 100 GeV are not included in a χ^2 calculation. There is a small peak visible in the χ^2 distribution for
 1444 events from $W \rightarrow e\nu$ selection, that can be assumed to come from the missing QCD background
 1445 contributions. Hadronic recoil bias parameters are determined through the fit of χ^2 distribution in
 1446 combined $W \rightarrow l\nu$ channel using the function from Eq. 11.10. The resulting bias is $HR_{SF} = 1.02$,
 1447 with the statistical error 0.06.

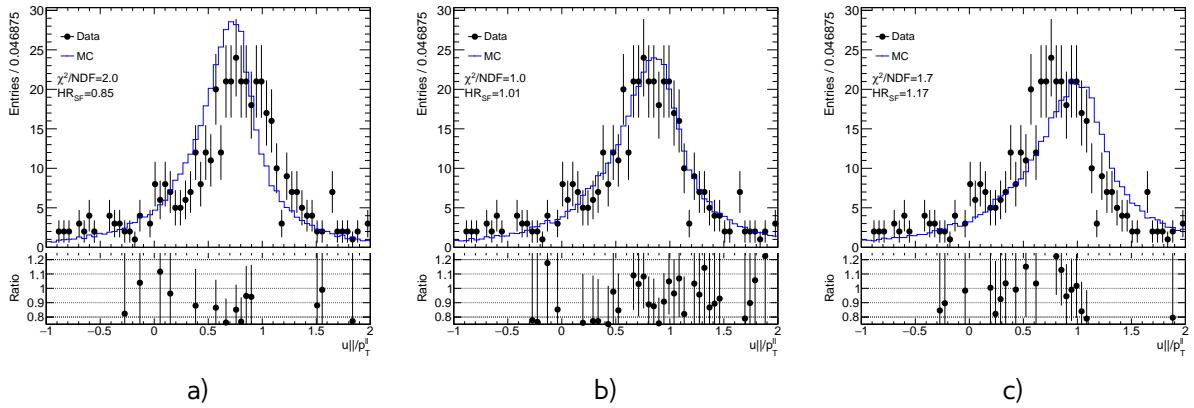


Fig. 11.20: Parallel hadronic recoil component u_{\parallel} from the $Z \rightarrow ee$ selection for different hadronic recoil scales: a) $HR_{SF}=0.75$ b) $HR_{SF}=1.1$ c) $HR_{SF}=1.23$. The expected contribution from signal is estimated with Monte Carlo simulation, other background sources are considered negligible.

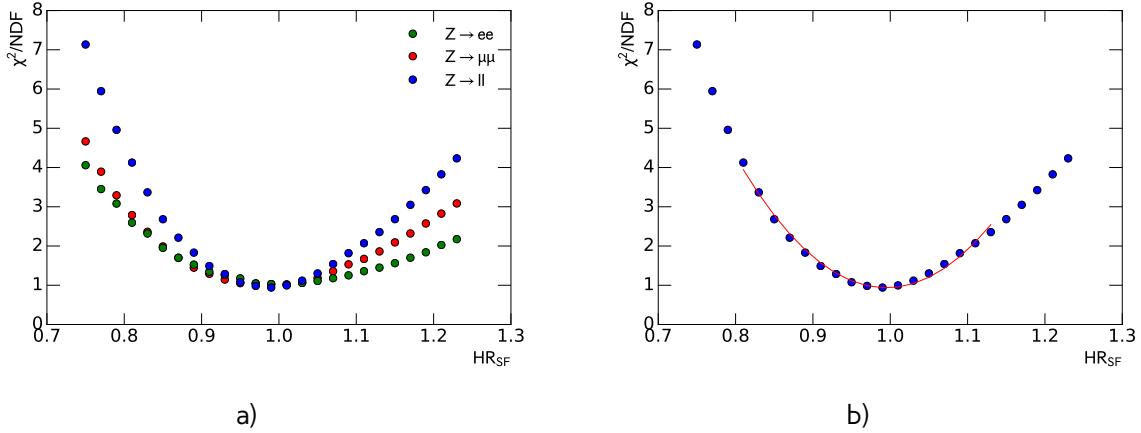


Fig. 11.21: a) Distribution of χ^2 between data and MC for $\frac{u_{\parallel}}{p_T^2}$ distribution as a function of hadronic recoil scale HR_{SF} for a) different Z boson channels. b) for combined $Z \rightarrow ll$ selection. Fit results are shown by a red line. The expected contribution from signal is estimated with Monte Carlo simulation, other background sources are considered negligible.

1448 Additionally, a cut on M_T^W lower value may be used to reduce the multijet background contamination.
 1449 The M_T^W range introduces a source of the systematic uncertainty in the hadronic recoil scale
 1450 determination. It is estimated by repeating the fit for different M_T^W lower and upper values, as shown
 1451 in Fig. 11.19. Fit range systematic error calculated as an RMS of the obtained values and is 0.03. The
 1452 final result for this method is $HR_{SF} = 1.02 \pm 0.07$.

1453 11.3.2 Bias determination using u_{\parallel} distribution

1454 Similarly to the W channel, the scale correction in the Z sample can be determined from the HR_{SF}
 1455 scan of the $\frac{u_{\parallel}}{p_T^2}$ distribution, as shown in Fig. 11.20. All of the backgrounds sources are considered
 1456 negligible in this case. Results of the χ^2 test for data and MC in different channel are shown in Fig.

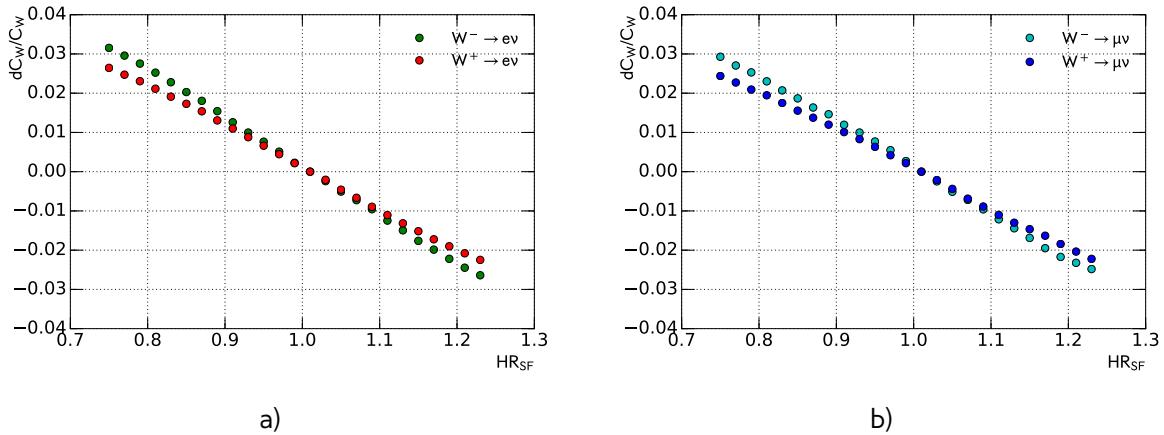


Fig. 11.22: Effect on a C_W for a different $d\sigma$ for a) $W \rightarrow e\nu$ b) $W \rightarrow \mu\nu$ channel

1457 11.21. A fit of the combined $Z \rightarrow ll$ distribution gives the most precise estimation of the hadronic
 1458 recoil bias $HR_{SF} = 1.00 \pm 0.01$. Since there is no choice of the range and dependency on P_T^{bos}
 1459 modeling, there is just one statistical source of uncertainty.

1460 11.3.3 Systematic uncertainty estimation

Table 11.3: Hadronic recoil bias determination results and errors for different methods.

| Method | SF | error |
|------------------------|------|-------|
| Mean M_T^W | 1.10 | 0.2 |
| $M_T^W \chi^2$ | 1.01 | 0.07 |
| $u_{\parallel} \chi^2$ | 1.00 | 0.014 |

1461 Results on a hadronic scale factors and its errors are shown in a Table 11.3. The results are con-
 1462 sistent within one sigma. As a final result it was decided to choose HR_{SF} determined from $Z \rightarrow ll$
 1463 selection as the are established with smallest uncertainty/ Scale factors extracted with other meth-
 1464 ods are used as a cross-check.

1465 Effect of the hadronic recoil bias correction for different bias scale factors presented in Fig. 11.22.
1466 Systematic error, coming from the bias correction is estimated using offset method (see Chap. 14).

1467 11.4 Summary on hadronic recoil systematics

Because of the problems with data vs MC comparison it was decided to use a hadronic recoil algorithm of E_T^{miss} reconstruction. Because of the differences in operation conditions the calibration of hadronic recoil must be determined directly from 2.76 TeV data. The limited statistics of the Z sample does not allow to use the standard procedure, used for the M_T^W measurement at 7 TeV, so the new methodology was developed.

1473 The hadronic recoil calibration can be divided into two parts: the correction of resolution and the
1474 bias correction. The hadronic recoil resolution have been corrected using the following methods:

Table 11.4: Hadronic recoil bias systematics for different W boson channels.

| Systematic source | $W^+ \rightarrow e^+\nu$ | $W^- \rightarrow e^-\nu$ | $W^+ \rightarrow \mu^+\nu$ | $W^- \rightarrow \mu^-\nu$ |
|----------------------------|--------------------------|--------------------------|----------------------------|----------------------------|
| Hadronic recoil resolution | -0.2% | -0.11% | -0.16% | -0.12% |
| Hadronic recoil scale | 0.21% | 0.20% | 0.23% | 0.24% |

- Event activity correction through the reweighting of $\sum E_T$ distribution. Different methods of the data/MC ratio parametrization have been developed and showed the consistent result. However, this method gives a unphysical difference between electron and muon channels, that cannot be accounted for the data statistics, so it was decided to drop this method.
- Smearing correction of the hadronic recoil. This method uses the Z sample to determine the difference in resolutions of the hadronic recoil components. The overall effect of these correction was estimated by repeating the smearing 25 times and consistent between electron and muon channels

The bias of hadronic recoil was estimated on W and Z events using 3 methods:

- Difference in the mean of the M_T^W distributions in data and MC. This method gives the highest uncertainty and used as a cross-check for other results
- Through the scans of the hadronic recoil scale effect on χ^2 in data vs MC M_T^W distributions. Error on this method is dominated by the statistics
- Through the scans of the hadronic recoil scale effect on χ^2 in data vs MC $\frac{u_{||}}{p_T^H}$ distributions. Despite the small size of the Z boson sample, this distribution has the biggest sensitivity to the hadronic recoil scale. It was decided to use this result and its error as a final result.

The results are agreeing between channels within 1 sigma.

The corresponding error sources for the hadronic recoil calibration have been estimated and summarised in the Tab. 11.4. The overall error on E_T^{miss} is around 0.3 % for all W-boson channels and can be considered a subdominant.

1495

Chapter 12

Background estimation

1497

1498

| | |
|--|------------|
| 12.1 QCD background estimation | 104 |
| 12.1.1 Template selection | 105 |
| 12.1.2 Methodology of the template sample normalization | 105 |
| 12.1.3 Systematic uncertainty from the multi-jet background estimation | 107 |
| 12.2 Background-subtracted W and Z candidate events | 109 |

1499

1500

1501

1502

1503

1504

1505

After the event selection described in Chap. 9 the background contribution is around 4% for W-analysis and 0.2% for Z analysis (which with this statistics is negligible). Main backgrounds for W analysis are coming from:

- Processes with τ lepton, misidentified as an electron or muon + missing energy from neutrino
- Z decays with one missing lepton.
- QCD processes. In electron channel these are mostly jets faking electrons, while in the muon channel it consists mainly of a real muons produced in decays of heavy-flavor mesons.

Most of the backgrounds are estimated using MC. They are normalized using highest cross-section order available. The total list of simulated backgrounds and its cross-section is shown in a Table 12.1. The QCD background is estimated using data driven method.

Table 12.1: Background processes with their associated cross sections and uncertainties (if given). The quoted cross sections are used to normalise estimates of expected number of events

| Process | $\sigma \cdot BR(\pm unc.)$ [pb] | Order |
|------------------------|----------------------------------|-------|
| $W^+ \rightarrow l\nu$ | 2116(± 41) | NNLO |
| $W^- \rightarrow l\nu$ | 1267(± 24) | NNLO |
| $Z \rightarrow ll$ | 303(± 6) | NNLO |
| $t\bar{t}$ | 7.41 | LO |
| WW | 0.6 | LO |
| ZZ | 0.7 | LO |
| WZ | 0.2 | LO |

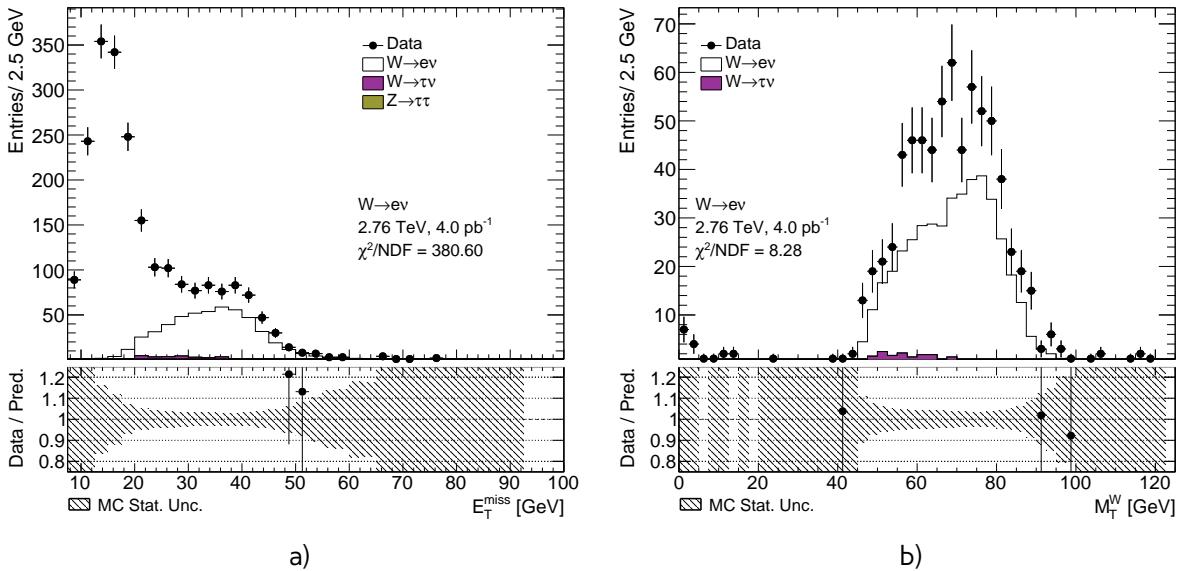


Fig. 12.1: Distribution for a) missing transverse energy E_T^{miss} b)mass transverse M_T^W from the QCD template selection for $W \rightarrow e\nu$ events

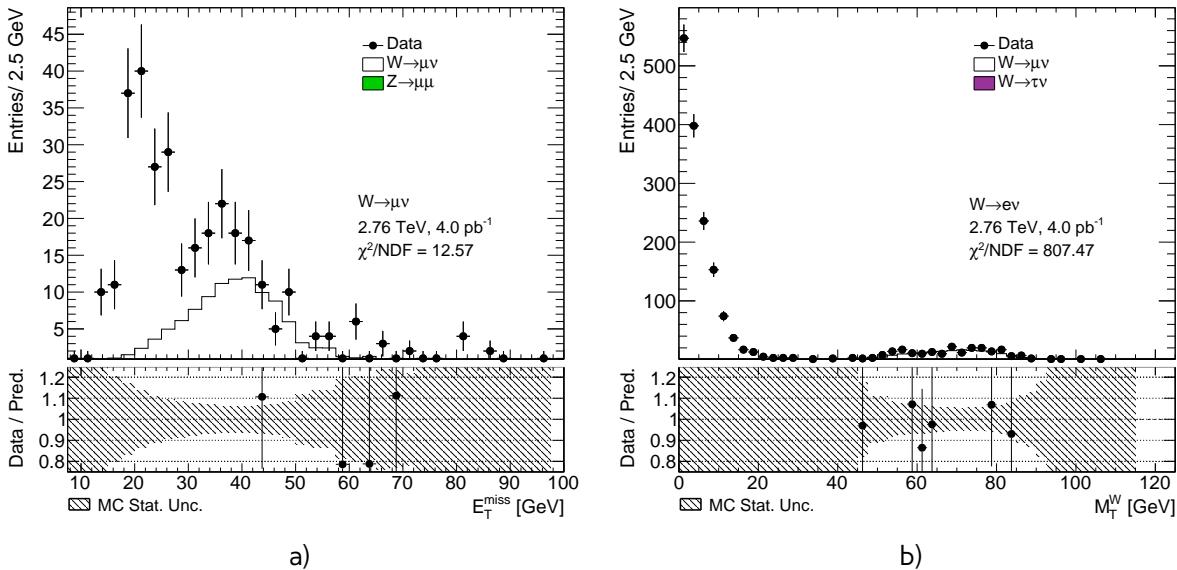


Fig. 12.2: Distribution for a) missing transverse energy E_T^{miss} b)mass transverse M_T^W from the QCD template selection for $W \rightarrow \mu\nu$ events

12.1 QCD background estimation

There is a small probability, that a jet can fake W -boson decay with isolated lepton and E_T^{miss} through the energy mismeasurement in the event. Event selection is suppressing this type of background, but not fully eliminating it. Due to a large jet production cross-section and complex composition, generation of MC events becomes impractical. This is why data driven technique for QCD background estimation have been used. In our case contribution from the QCD background in the Z sample is negligible(Fig. 13.15), so it is estimated just for $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ processes.

1524 Data driven method allows to have model independent predictions with small statistical uncertainty.
 1525 This method is using *QCD* enriched region, where signal events are suppressed. This is usually done
 1526 by reversing identification or isolation criteria. It is assumed, that shape of the *QCD* background
 1527 stays the same in the signal region. Normalization can be derived in a control region through the
 1528 template fit.

1529 This section describes method of *QCD* background determination, that have been used in 2.76
 1530 TeV data.

1531 12.1.1 Template selection

1532 A study have been performed to determine appropriate template selection. Because of the origin
 1533 of the *QCD* backgrounds, missing transverse energy E_T^{miss} should be smaller in background sample,
 1534 that in a signal region. Releasing E_T^{miss} cut allows to gain a bigger statistics for a *QCD* template.
 1535 Another possibility is to relax the transverse mass M_T^W cut. Most of the multijet background event
 1536 should contribute to the small M_T^W region. The template sample can contain also contributions from
 1537 other backgrounds (mostly coming from $W \rightarrow l\nu$). The best template selection should allow for big
 1538 data statistics and small electroweak contributions at the same time. In order to suppress the signal
 1539 additionally reversed ID or isolation criteria is applied.

1540 In electron channel, the template selection requires an electron candidate to fail medium identifi-
 1541 cation criteria, but to pass loose selection. Control distributions for a different template selection in
 1542 electron channel are shown on a Fig. 12.1. Relaxed E_T^{miss} cut allows to gain bigger template statistics.

1543 In a muon channel template selection is build by inverting isolation criteria ($P_T^{cone,20} > 0.1$). In case of
 1544 $W \rightarrow \mu\nu$ the *QCD* background template the best statistics is achieved by relaxing mass transverse
 1545 M_T^W cut (Fig. 12.2).

In order to avoid double counting, electroweak processes (i.e. signal and backgrounds) are sub-
 stracted from a template. The total number of events in the template can be defined as:

$$N_{template} = N_{data}^{bkg\ enriched} - \sum_j^{MC} N_{MC_j}^{bkg\ enriched}, \quad (12.1)$$

1546 where $N_{data}^{bkg\ enriched}$ and $N_{MC_j}^{bkg\ enriched}$ are numbers of the events in a background enriched sample in data
 1547 and different MC samples. The resulting template statistic is 1348 and 1509 events for $W \rightarrow e\nu$ and
 1548 $W \rightarrow \mu\nu$ respectively.

1549 12.1.2 Methodology of the template sample normalization

The normalization is found through the χ^2 fit of the template and backgrounds to the data. The
 following composite model has been used for estimation:

$$M(x) = \sum_{i=1}^{N-1} f_i F_i(x) + (1 - \sum_{i=1}^{N-1} f_i) \cdot F_{QCD}(x), \quad (12.2)$$

1550 where index i goes over the MC samples, x is a fit variable (E_T^{miss} or M_T^W), $F_i(x)$ and $F_{QCD}(x)$ are
 1551 the probability density functions of MC samples and *QCD* background template respectively. Fit
 1552 parameters f_i are the fractions of MC events within the fit region. In order to eliminate systematics,
 1553 coming from the cross-section uncertainty, the signal fractions are left as free parameters of fit and
 1554 and the background MC fractions are allowed to be varied within 5% uncertainty.

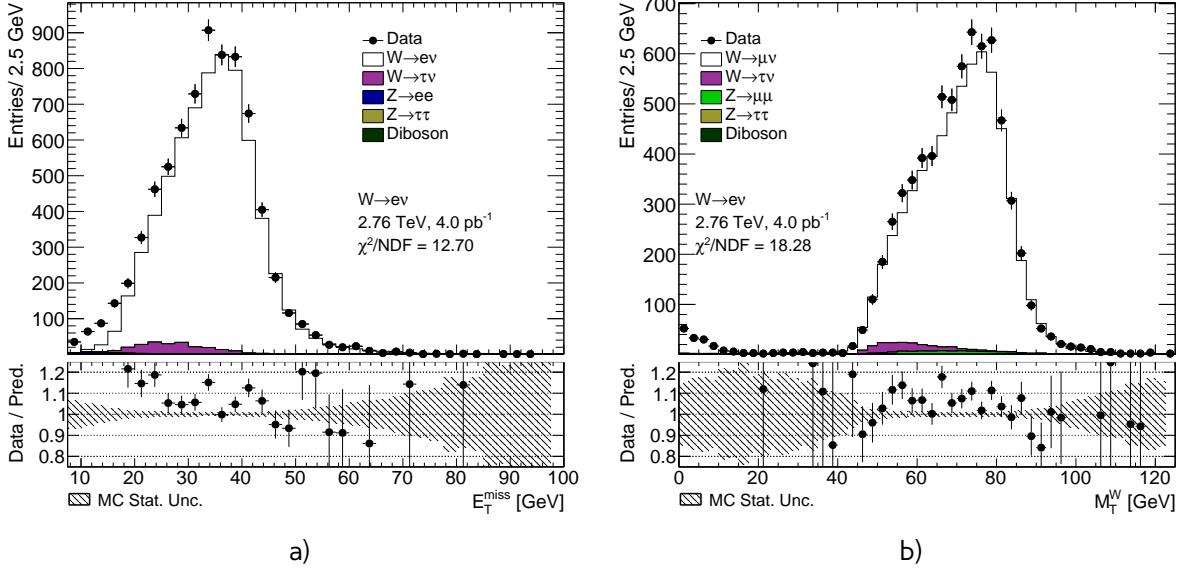


Fig. 12.3: Distributions used for multijet background estimation for a) $W \rightarrow e\nu$ b) $W \rightarrow \mu\nu$

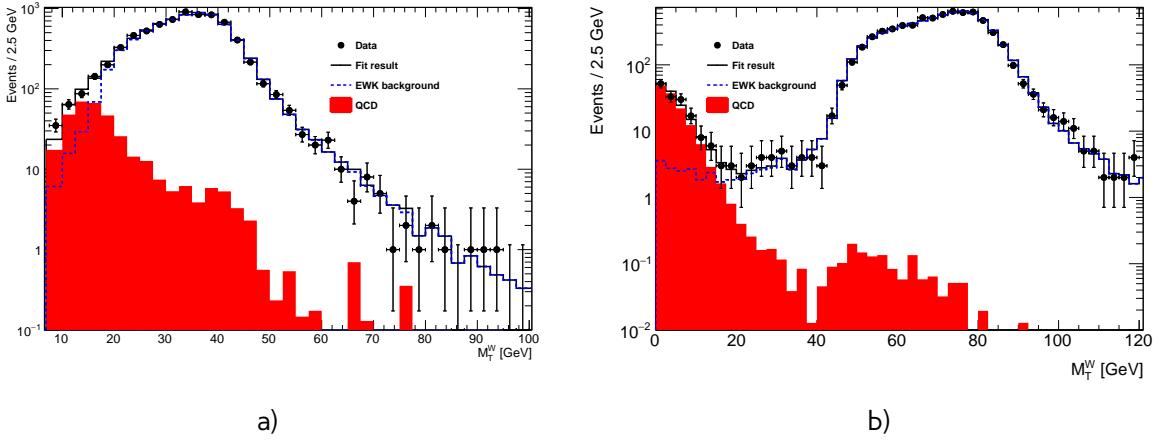


Fig. 12.4: The multijet background estimation for a) $W \rightarrow e\nu$ using reversed ID cut and released E_T^{miss} cut b) $W \rightarrow \mu\nu$ using released M_T^W cut and $bb\bar{b} + cc\bar{c}$ template

Normalisation scale of the QCD events is calculated from the obtained fit parameters as:

$$scale = \frac{(1 - \sum f_i) \cdot N_{Data}^{fit}}{N_{template}}, \quad (12.3)$$

where $\sum f_i$ is a sum of all fractions in the fit, N_{Data}^{fit} is a number of data events in a fit histogram and $N_{template}$ is a number of event in a template. The fit is performed separately for W^+ and W^- . Additionally, fit in total W channel is used as a cross-check of the fit. The results of the fitting procedure are shown on a Fig. 12.4 .

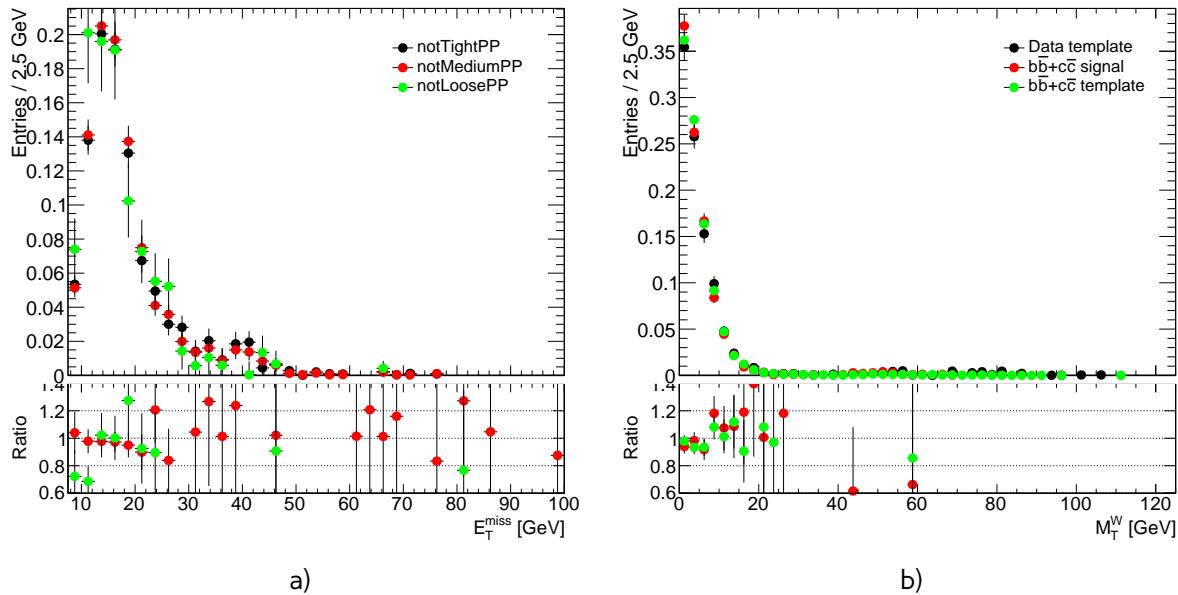


Fig. 12.5: Data and MC comparison for E_T^{miss} calculated by standard ATLAS algorithm for a) $W \rightarrow e\nu$
b) $W \rightarrow \mu\nu$ events

12.1.3 Systematic uncertainty from the multi-jet background estimation

The uncertainty of the multi-jet background estimation can be divided into 3 main components:

$$\delta_{\text{QCD}} = \sqrt{\delta_{\text{fit unc}}^2 + \delta_{\text{MC}}^2 + \delta_{\text{fit bias}}^2 + \delta_{\text{template}}^2}, \quad (12.4)$$

where $\delta_{\text{fit unc}}$ is the uncertainty for a scale from a χ^2 fit. The meaning of other components is explained below

The second component δ_{MC} is coming from a possible mismodelling of MC in a fitted region. It can be estimated by comparison of fit results for W , W^+ and W^- . Number of multijet background events should not depend on a charge of the analysis, so it is expected:

$$N_{\text{QCD}}^W = 0.5 \cdot N_{\text{QCD}}^{W^+} = N_{\text{QCD}}^{W^-} \quad (12.5)$$

Standard deviation of these 3 numbers is taken as systematic uncertainty. Since in $W \rightarrow \mu\nu$ channel the QCD template normalization is derived from the fit in small M_T^W region, where electroweak contributions negligible and data statistics is high, this systematic source is equal to 0.

The third component $\delta_{\text{fit bias}}$ is coming from an effect of an arbitrary choice of bin size. This error is estimated by repeating the fit for a different binnings. This component is assumed negligible in $W \rightarrow \mu\nu$ case.

The uncertainty δ_{template} is due to a potential bias in the template as a result of the template choice and a template statistics itself. For estimation of this uncertainty different template selections have been used. For $W \rightarrow e\nu$ channel different reversed isolation criteria have been tried (Fig. 12.5 a)). The overall discrepancies can be considered negligible. For $W \rightarrow \mu\nu$ channel template variations are estimated using fits with $b\bar{b} + c\bar{c}$ MC samples. Fig. 12.5 b) compares data template with template obtained using signal selection with released M_T^W cut and template selection. Results for a different

Table 12.2: Results of QCD background estimation for $W \rightarrow e\nu$ and corresponding error

| Charge | N_{QCD} | $\delta N_{fit\,unc}$ | δN_{MC} | $\delta N_{fit\,bias}$ |
|--------|-----------|-----------------------|-----------------|------------------------|
| W^+ | 38.3 | 7.0 | 7.0 | 5.0 |
| W^- | 21.5 | 0.7 | -9.4 | 4.0 |
| W | 66.1 | 21.2 | 4.2 | 10. |
| Total | 31.0 | 6.1 | 8.6 | 4.7 |

 Table 12.3: Results of QCD background estimation for $W \rightarrow \mu\nu$ using different templates and it's fit error

| Charge | N_{QCD} data template | N_{QCD} $b\bar{b} + c\bar{c}$ template selection | N_{QCD} $b\bar{b} + c\bar{c}$ signal selection |
|-------------------|----------------------------|---|---|
| W^+ | 2.48 | 0.73 | 1.34 |
| W^- | 2.48 | 0.73 | 1.35 |
| W | 4.97 | 1.47 | 2.70 |
| Total per channel | 2.48 | 0.73 | 1.35 |
| Fit error | 0.60 | 0.73 | 0.19 |

¹⁵⁷⁴ template fits are presented in Tab 12.3

¹⁵⁷⁵ Results of the QCD background uncertainty estimation for $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ are shown
¹⁵⁷⁶ in Tab. 12.2 and 12.3 respectively. The overall number of QCD background events is estimated as
¹⁵⁷⁷ 31.00 ± 13.0 for $W^+ \rightarrow e^+\nu$ and $W^- \rightarrow e^-\nu$ and 1.50 ± 0.9 for $W^+ \rightarrow \mu^+\nu$ and $W^- \rightarrow \mu^-\nu$.
¹⁵⁷⁸ The overall fraction of the QCD events is lower, than in 13 TeV [28], and 7 TeV [83] data, what is in
¹⁵⁷⁹ agreement with expectations.

 Table 12.4: Number of observed candidate events for the $W \rightarrow l\nu$ channel, electroweak (EWK) and top, and data-driven QCD background events, and background-subtracted signal events

| I | Observed candidates | Background (EWK + top) | Background (Multijet) | Background-subtracted data N_W^{sig} |
|---------|---------------------|------------------------|-----------------------|--|
| W boson | | | | |
| e^+ | 3914 | 108.1 ± 5.7 | 31.00 ± 13.0 | $3774.9 \pm 62.6 \pm 5.7 \pm 13.0$ |
| e^- | 2209 | 74.2 ± 3.3 | 31.00 ± 13.0 | $2103.8 \pm 47.0 \pm 3.3 \pm 13.0$ |
| μ^+ | 4365 | 152.4 ± 6.7 | 1.50 ± 0.9 | $4211.1 \pm 66.1 \pm 6.7 \pm 0.9$ |
| μ^- | 2460 | 107.5 ± 4.2 | 1.50 ± 0.9 | $2351.0 \pm 49.6 \pm 4.2 \pm 0.9$ |
| Z boson | | | | |
| e | 430 | 1.3 ± 0.0 | - | $428.7 \pm 20.7 \pm 0.0$ |
| μ | 646 | 1.6 ± 0.1 | - | $644.4 \pm 25.4 \pm 0.1$ |

1580 12.2 Background-subtracted W and Z candidate events

1581 Tables 12.4 summarize the number of background events for W and Z selections. Uncertainties on
1582 a number of EWK+top events are coming from a statistics, cross-section uncertainty (if given) and
1583 3% of luminosity determination uncertainty. For multijet background uncertainty is coming from a
1584 method and described in a subsection 12.1.3. For the background-subtracted events the statistical
1585 uncertainty is quoted first, followed by the total systematic uncertainty, derived from the EWK+top
1586 and multijet background ones, considering the sources as uncorrelated.

Chapter 13

Control distributions

1589 Kinematic distributions after all cuts (Section 9) and corrections applied on MC (Section 10), are
 1590 presented in this chapter. Distributions for $W \rightarrow l\nu$ are split in charge and shown on a Figs. 13.1-
 1591 13.12. Distributions for $Z \rightarrow l^+l^-$ analysis are shown on a Fig. 13.13-13.17.

1592 These plots also showing the systematic and statistical uncertainty as a shaded band. The un-
 1593 certainties are including all sources, described in a 14, except for uncertainties coming from shape
 1594 variation due to a PDF reweighting and QCD background and luminosity. All uncorrelated uncertainty
 1595 sources are summed in quadrature. The expected background contributions are estimated using MC
 1596 simulations, apart from QCD background, which is found with a data driven method, as explained in
 1597 a previous chapter.

1598 Good overall agreement between data and MC is observed.

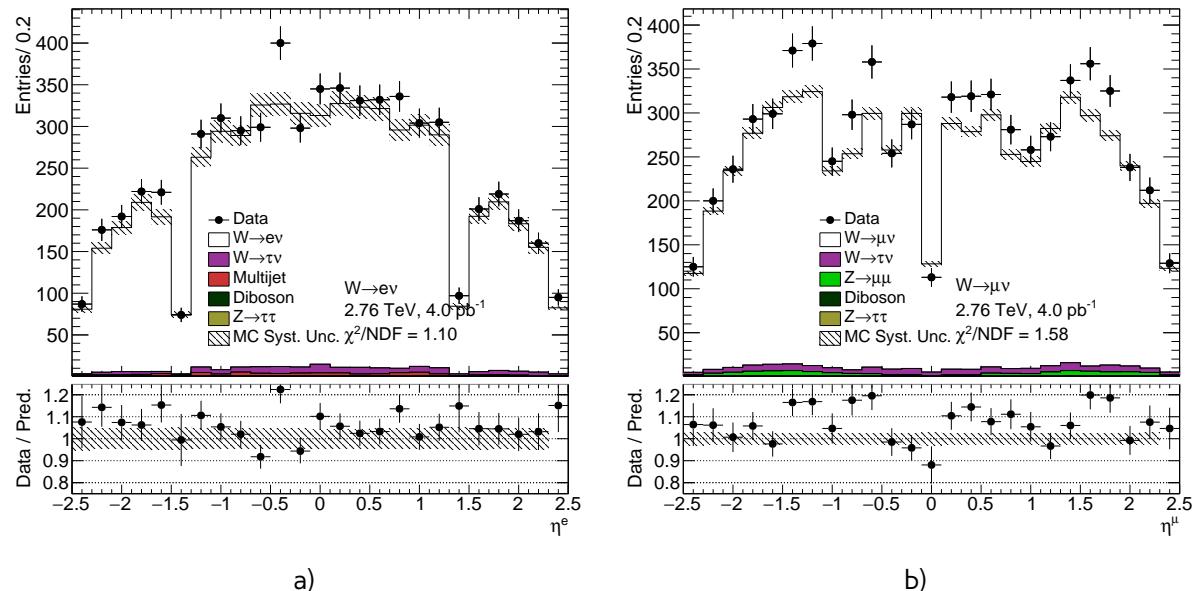


Fig. 13.1: Lepton pseudorapidity distribution from the a) $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection.

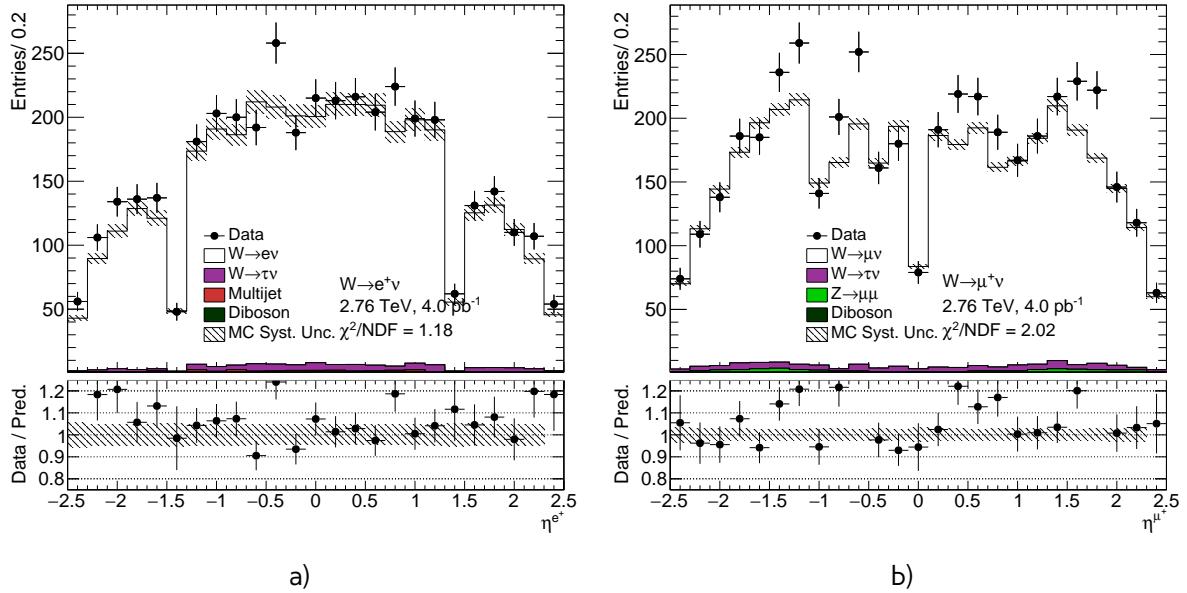


Fig. 13.2: Lepton pseudorapidity distribution from the a) $W^+ \rightarrow e^+\nu$ selection and b) the $W^+ \rightarrow \mu^+\nu$ selection.

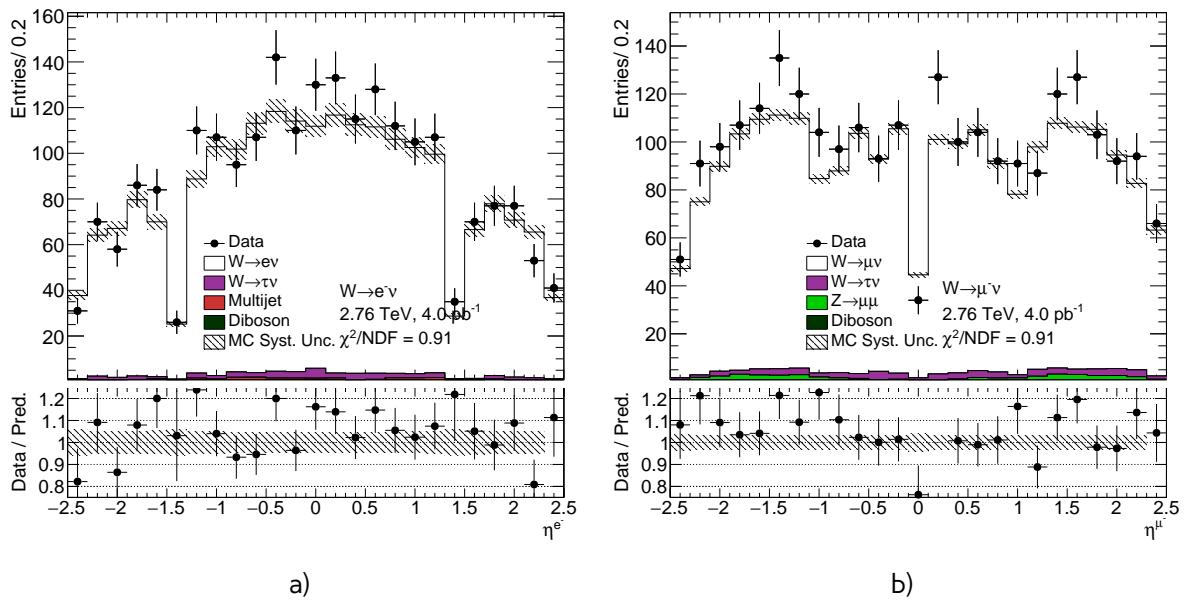


Fig. 13.3: Lepton pseudorapidity distribution from the a) $W^- \rightarrow e^-\nu$ selection and b) the $W^- \rightarrow \mu^-\nu$ selection.

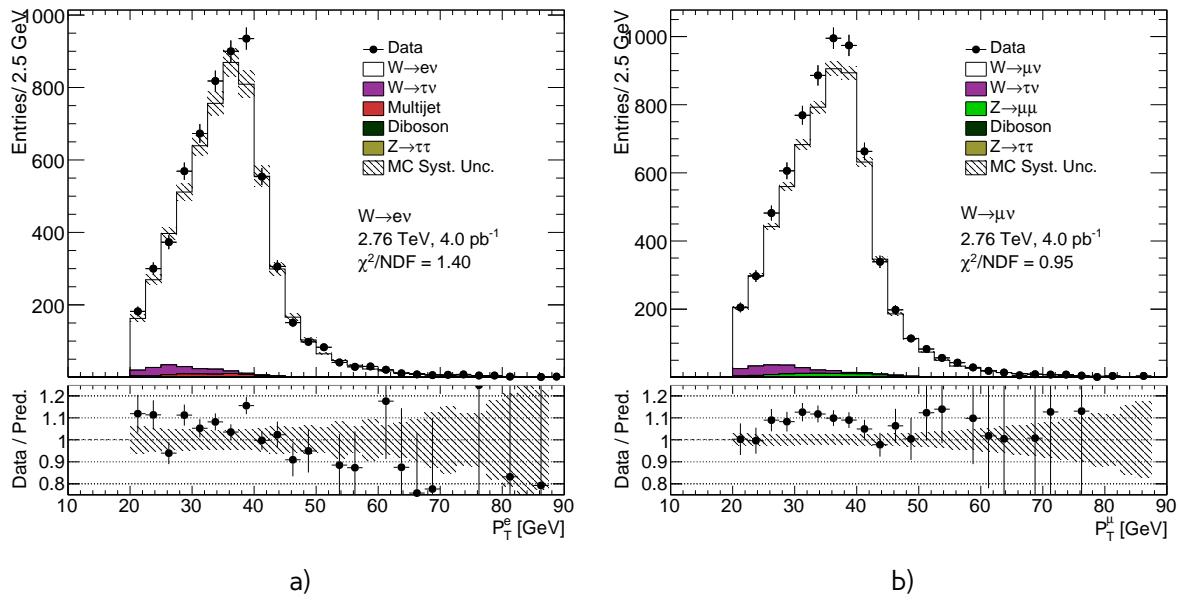


Fig. 13.4: Lepton transverse momentum distribution from the a) $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection.

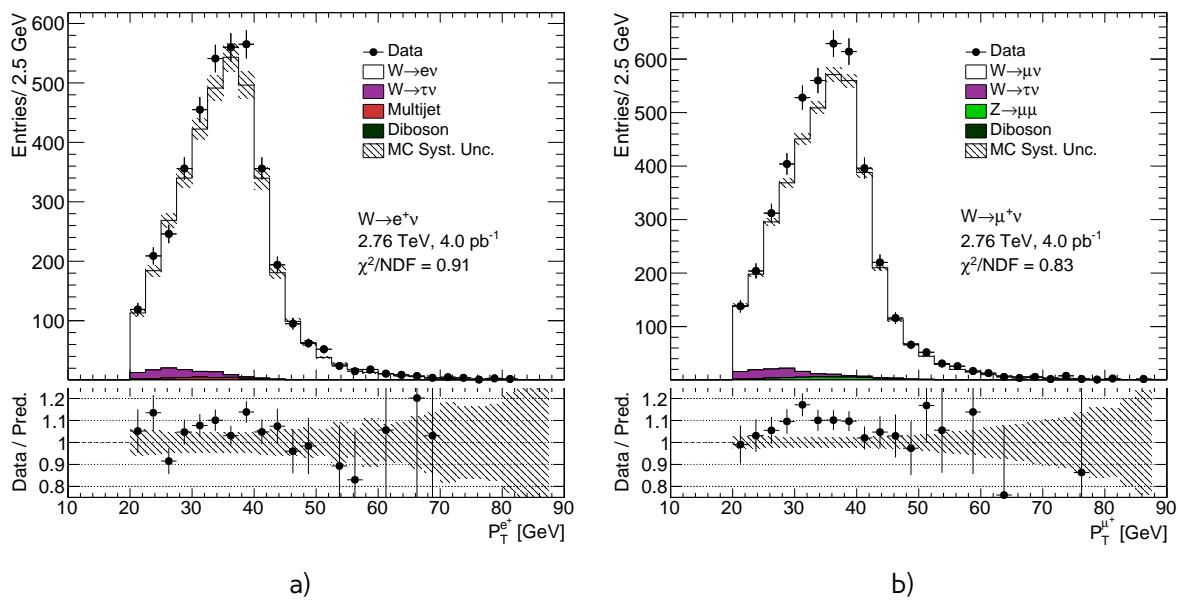


Fig. 13.5: Lepton transverse momentum distribution from the a) $W^+ \rightarrow e^+\nu$ selection and b) the $W^+ \rightarrow \mu^+\nu$ selection.

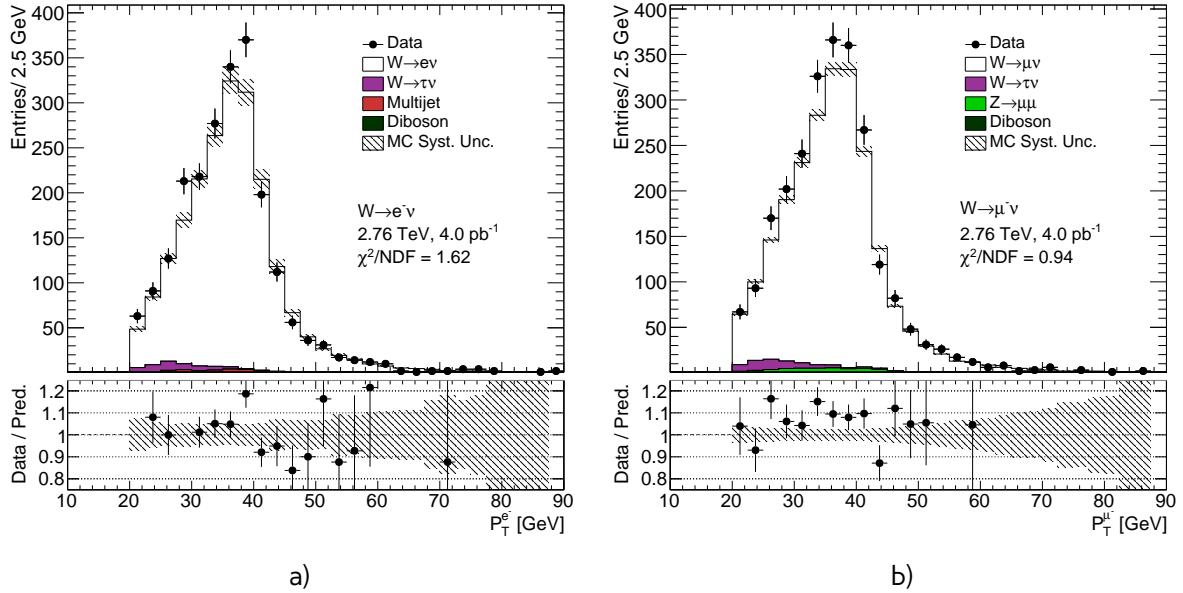


Fig. 13.6: Lepton transverse momentum distribution from the a) $W^- \rightarrow e^- \nu$ selection and b) the $W^- \rightarrow \mu^- \nu$ selection.

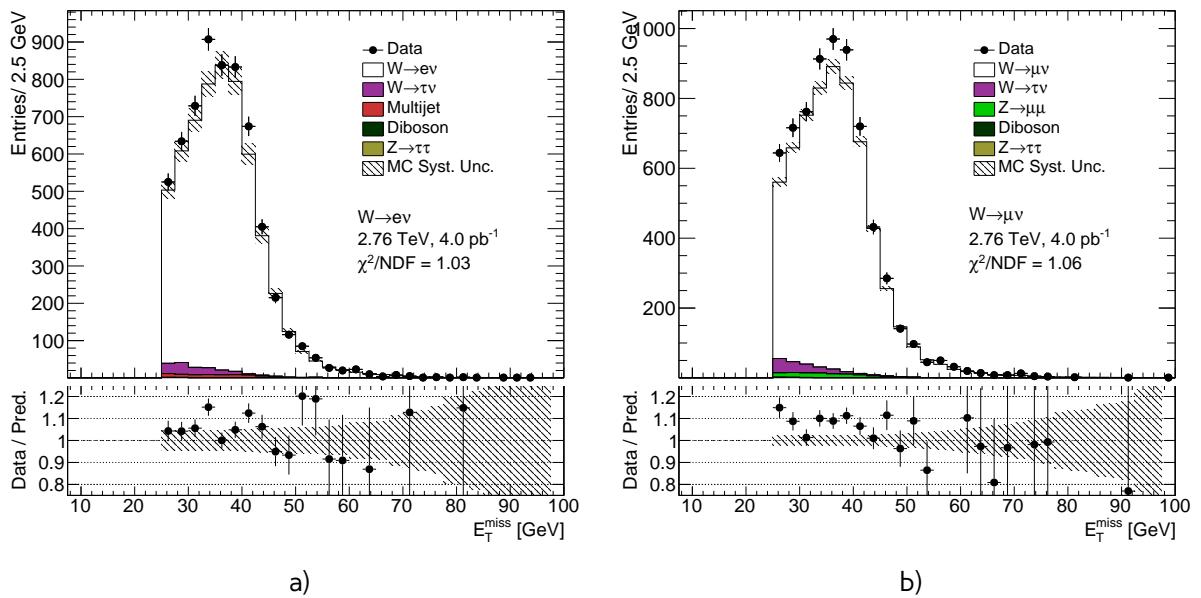


Fig. 13.7: Missing transverse energy distribution from the a) $W \rightarrow e \nu$ selection and b) the $W \rightarrow \mu \nu$ selection.

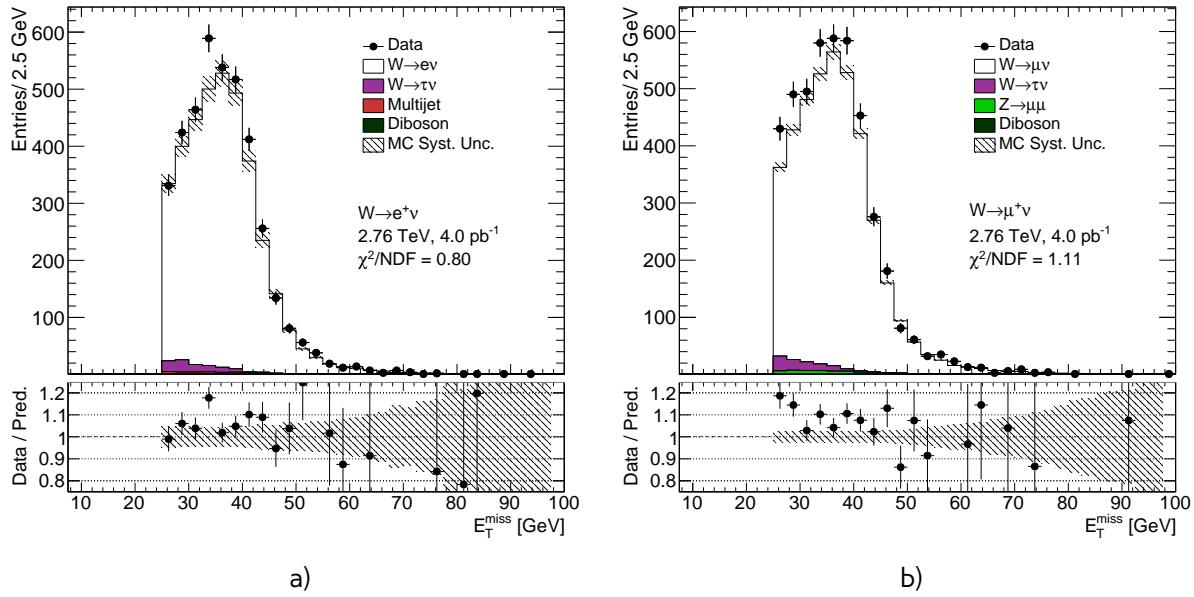


Fig. 13.8: Missing transverse energy distribution from the a) $W^+ \rightarrow e^+\nu$ selection and b) the $W^+ \rightarrow \mu^+\nu$ selection.

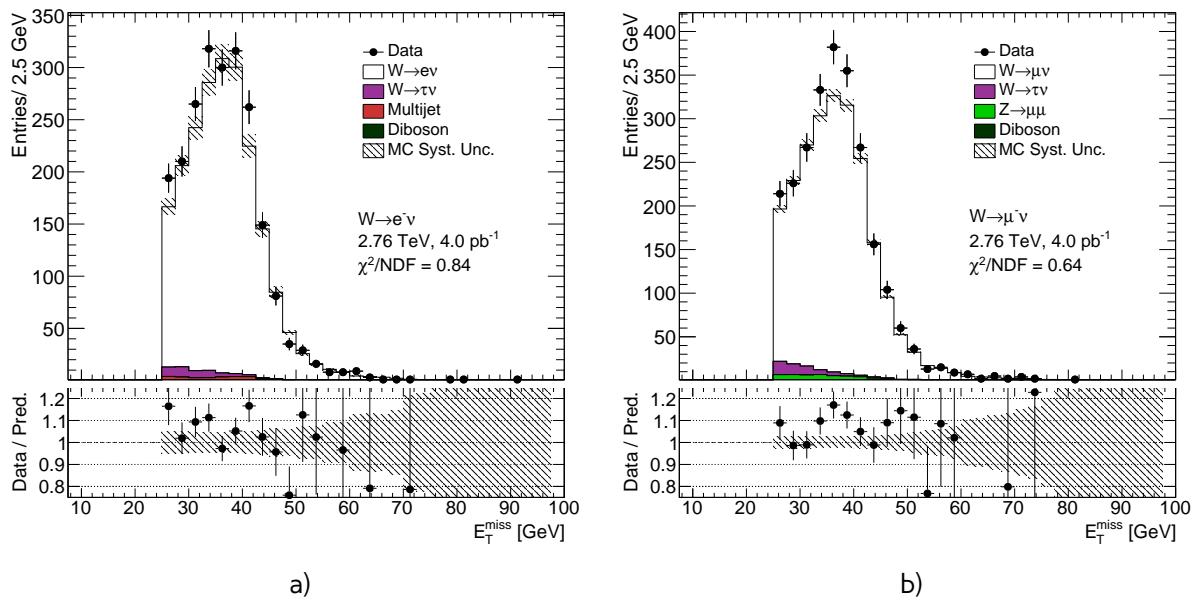


Fig. 13.9: Missing transverse energy distribution from the a) $W^- \rightarrow e^-\nu$ selection and b) the $W^- \rightarrow \mu^-\nu$ selection.

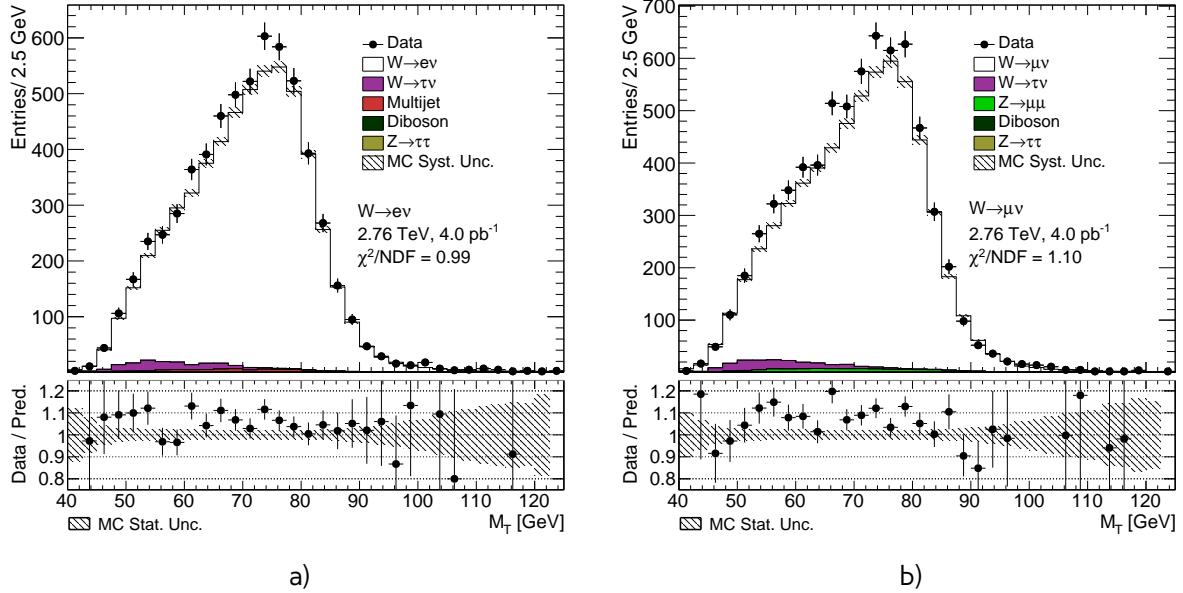


Fig. 13.10: Transverse mass distribution distribution from the a) $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection.

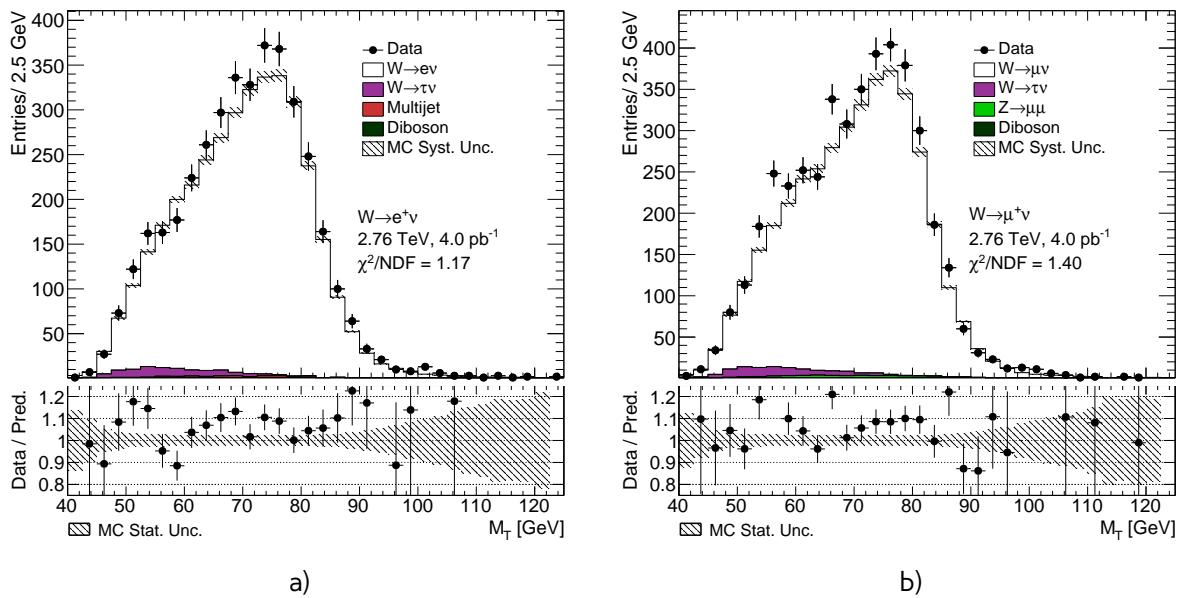


Fig. 13.11: Transverse mass distribution distribution from the a) $W^+ \rightarrow e^+\nu$ selection and b) the $W^+ \rightarrow \mu^+\nu$ selection.

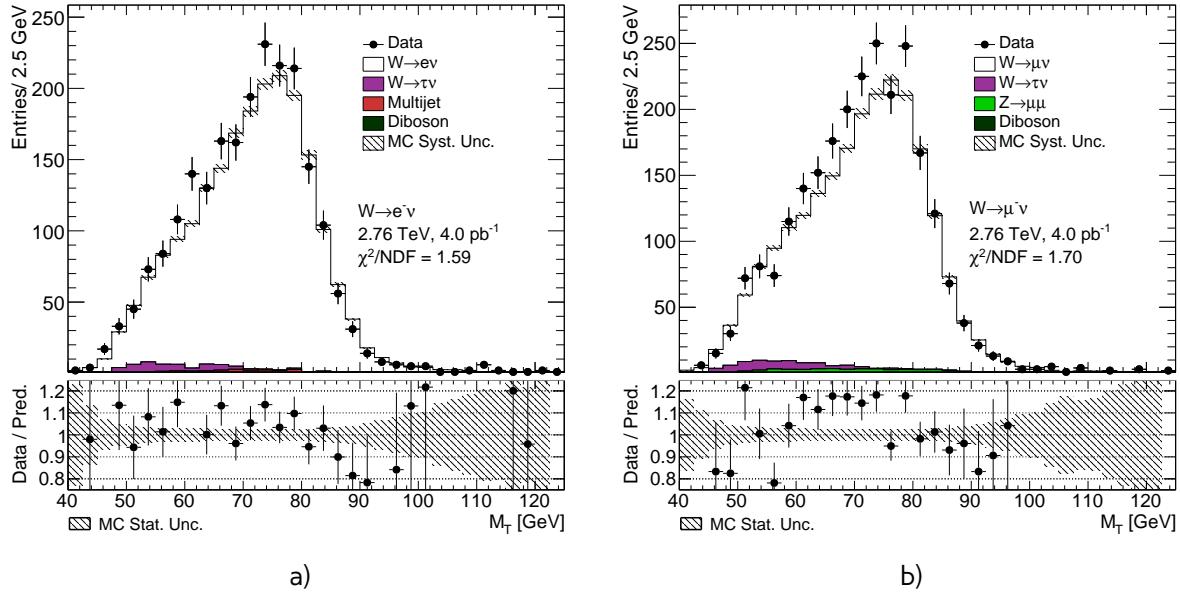


Fig. 13.12: Transverse mass distribution distribution from the a) $W^- \rightarrow e^-\nu$ selection and b) the $W^- \rightarrow \mu^-\nu$ selection.

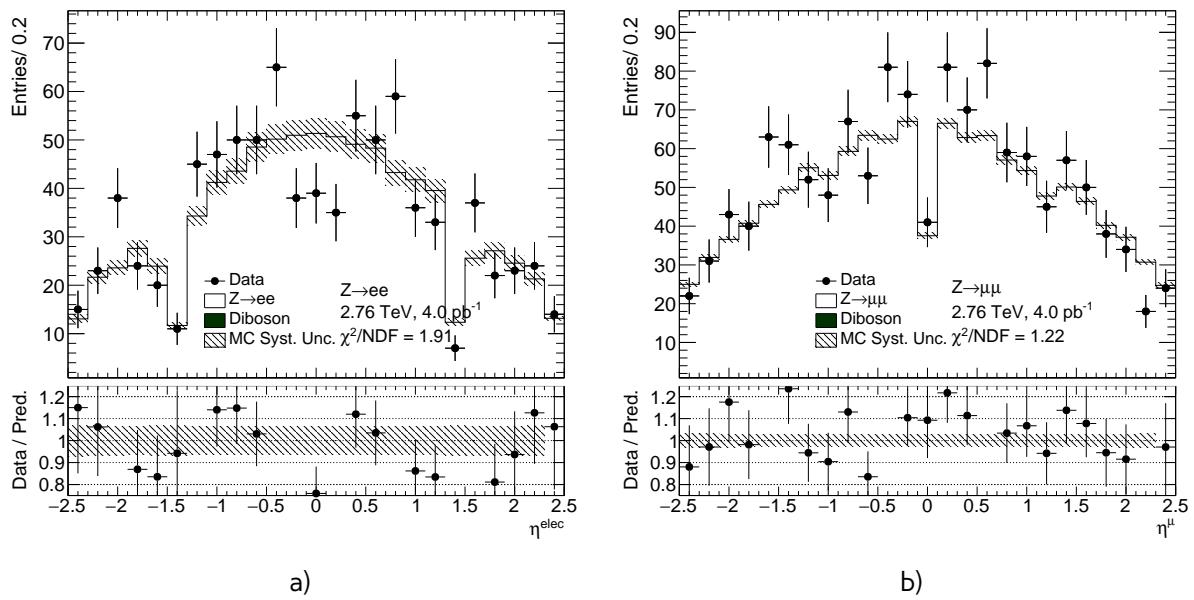


Fig. 13.13: Lepton pseudorapidity distributions from the a) $Z \rightarrow e^+e^-$ b) $Z \rightarrow \mu^+\mu^-$.

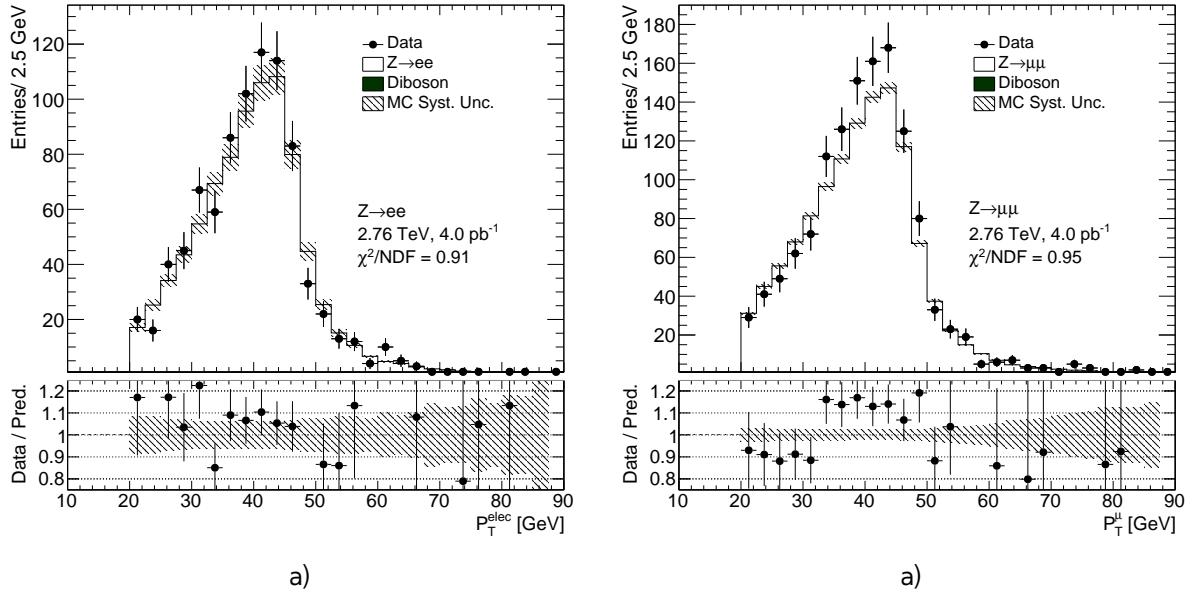


Fig. 13.14: Lepton transverse momentum distributions from the a) $Z \rightarrow e^+e^-$ b) $Z \rightarrow \mu^+\mu^-$.

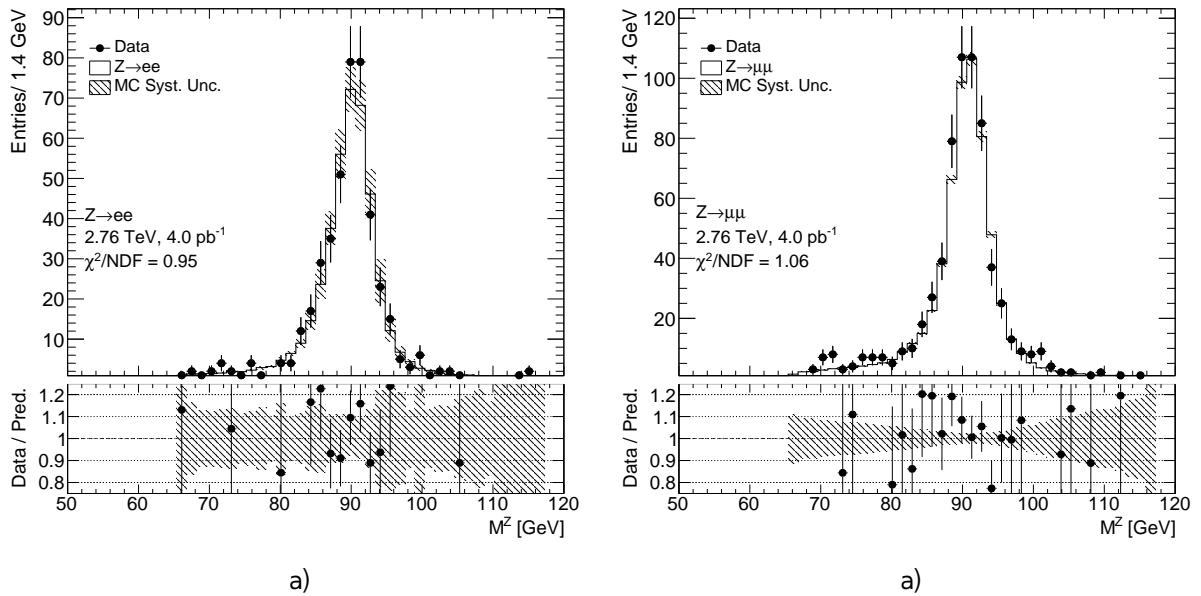


Fig. 13.15: Dilepton mass distribution distributions from the a) $Z \rightarrow e^+e^-$ b) $Z \rightarrow \mu^+\mu^-$.

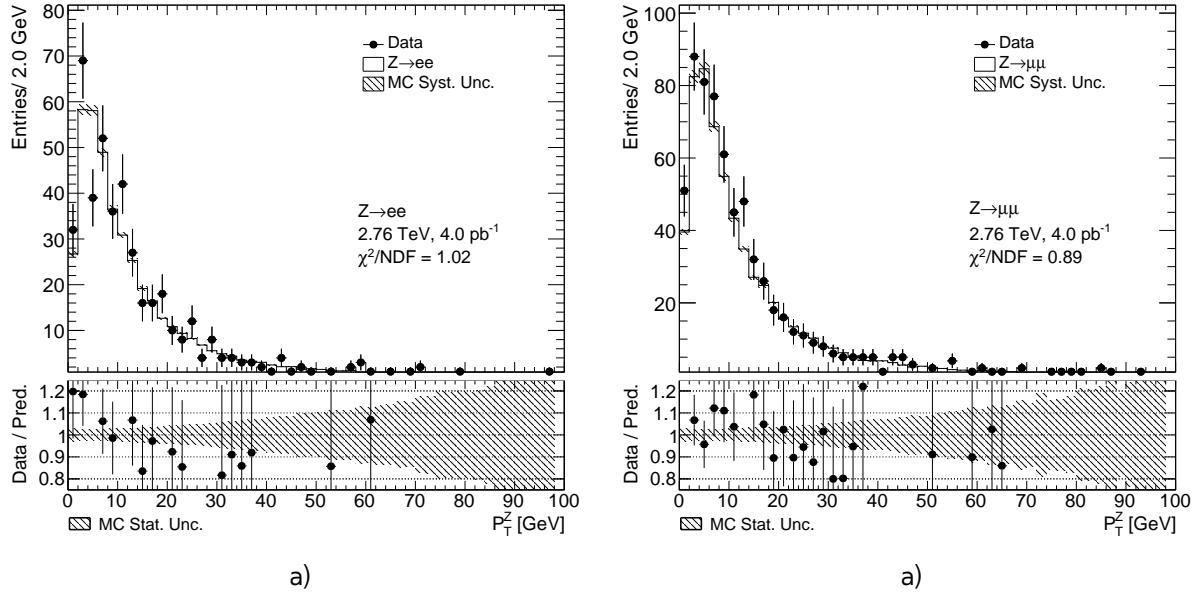


Fig. 13.16: Z boson transverse momentum distributions from the a) $Z \rightarrow e^+e^-$ b) $Z \rightarrow \mu^+\mu^-$.

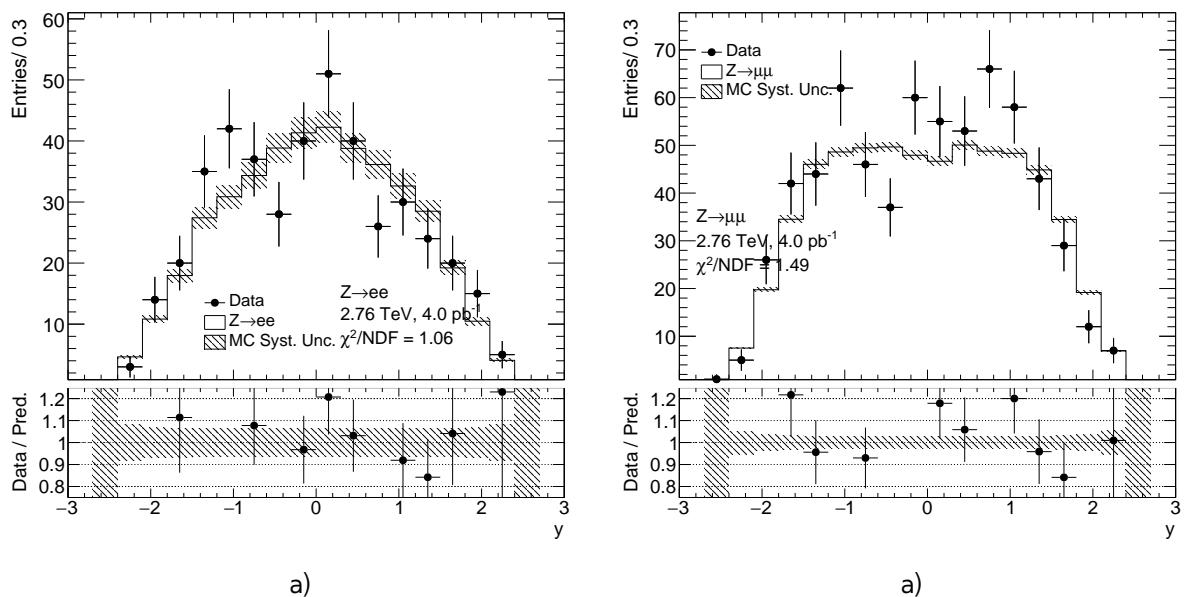


Fig. 13.17: Z boson rapidity distribution from the a) $Z \rightarrow e^+e^-$ b) $Z \rightarrow \mu^+\mu^-$.

1599

Chapter 14

Uncertainties in the cross-section measurement

1601

| | | |
|------|--|-----|
| 1602 | 14.1 Methods of uncertainties propagation | 121 |
| 1603 | 14.2 Systematic uncertainties on $C_{W,Z}$ | 122 |
| 1604 | 14.2.1 Electron energy scale and resolution | 122 |
| 1605 | 14.2.2 Muon energy scale and resolution | 122 |
| 1606 | 14.2.3 Muon and electron efficiency SF | 123 |
| 1607 | 14.2.4 Theoretical uncertainty | 123 |
| 1608 | 14.3 Statistical uncertainty in data | 123 |
| 1609 | 14.4 Theoretical uncertainty on $A_{W/Z}$ and $E_{W/Z}$ factors | 125 |
| 1610 | 14.5 Correlation between uncertainties | 125 |
| 1611 | 14.5.1 Treatment of partially correlated uncertainties | 126 |

1612

1613 Cross-section measurement relies on theoretical models and corrections, used in Monte-Carlo.
 1614 Thus, their intrinsic uncertainties should be propagated to a final result. This chapter discusses
 1615 main methods of uncertainties measurements. The sources of systematic uncertainties on $C_{W,Z}$ are
 1616 discussed in Sec. 14.2, the statistical uncertainty is described in Sec. 14.3 and the theoretical compo-
 1617 nents of $A_{W,Z}$ and $E_{W,Z}$ extrapolation factors are shown in Sec. 14.4. The background uncertainties
 1618 have been previously discussed in Chap.12. In the last section the correlation between uncertainties
 1619 for different analyses is presented.

14.1 Methods of uncertainties propagation

All sources of systematic uncertainties are propagated in these analyses using one of the main methods: Offset, On/Off or Toy Monte-Carlo. The offset method changes a correction by a $\pm 1\sigma$ of its systematic uncertainty. The contribution of each correction's uncertainty on the observable (e.g. $C_{W,Z}$, $A_{W,Z}$ or a cross-section) is taken as a symmetric approximation:

$$U_i^{\text{offset}} = \frac{\sigma_i^{\text{up}} - \sigma_i^{\text{down}}}{2}, \quad (14.1)$$

1624 where $\sigma_i^{\text{up}(\text{down})}$ - the change in a observable due to the shift of the correction on σ up (or down).

For On/Off method the contribution of each correction is estimated with (σ^{on}) and without(σ^{off}) correction applied. A systematic error can be estimated as:

$$U^{\text{on/off}} = \sigma^{\text{on}} - \sigma^{\text{off}}. \quad (14.2)$$

The a toy MC [84], uses pseudo experiments with modified input corrections. For scale factors binned in p_T and η , the uncertainties inside each bin can be divided to correlated and uncorrelated systematic components and statistical error. For each pseudo-experiment, a table of new scale factors is filled, where inside each bin a scale factor is randomly varied as:

$$SF_i^{Toy_n} = SF_i + \text{Gauss}(0, \Delta SF_i^{uncorr+stat}) + \sum \Delta SF_i^{corr} \cdot \text{Gauss}(0, 1), \quad (14.3)$$

where $SF_i^{Toy_n}$ is a new scale factor in i-th bin, $\Delta SF_i^{uncorr+stat}$ - is the quadratic sum of uncorrelated and statistical errors and ΔSF_i^{corr} is a correlated error.

The overall effect on an observable is calculated as a standard deviation of the values in pseudo-experiments:

$$U_i = \sqrt{\frac{\sum_{Toy_n=1}^N \sigma_i^2}{N} - \left(\frac{\sum_{Toy_n=1}^N \sigma_i}{N} \right)^2} \quad (14.4)$$

The number N of pseudo experiments should be sufficiently large to avoid possible bias in the uncertainty estimation.

14.2 Systematic uncertainties on $C_{W,Z}$

Sources of experimental uncertainties, methods of estimation and their effect on $C_{W,Z}$ are summarized in a Tab. 14.1. Systematic errors coming from the hadronic recoil calculation are discussed in Sec. 11.

14.2.1 Electron energy scale and resolution

Electron energy scale correction, described in Sec. 10.2 has associated uncertainties coming from [85]:

- Statistical component of the scale uncertainty
- Uncertainty from the possible bias of the calibration method
- Scale uncertainty from the choice of generator
- Uncertainty from the presampler energy scale
- Imperfect knowledge of the material in front of EM calorimeter.

The uncertainty contribution from each component is estimated using offset method. The total energy scale uncertainty is the quadratic sum of the components [86].

14.2.2 Muon energy scale and resolution

Systematic uncertainties coming from muon momentum corrections described in Sec. 10.3 can be divided into 3 major independent categories:

- variations of the smearing of MS track
- variation of the smearing of ID track

- 1648 • overall scale uncertainty
- 1649 The uncertainty contribution from each component is estimated using offset method. The total
1650 energy scale uncertainty is the quadratic sum of the components.

1651 14.2.3 Muon and electron efficiency SF

- 1652 The systematic uncertainty coming from efficiency scale factors is estimated using the toy MC
1653 method for reconstruction, identification and trigger scale factors for electron analyses and recon-
1654 struction and identification for muon analyses. Since the muon trigger scale factors are unbinned at
1655 2.76 TeV, the offset method for the corresponding uncertainty estimation is used.
1656 In case of the analysis of W/Z production at 2.76 TeV the scale factor errors are taken from 8 TeV
1657 with significantly enlarged statistical error (see Chap. 10), so correlated and uncorrelated errors are
1658 considered to be negligible. In the current analysis 30 pseudo-experiments are used with a combined
1659 toy MC method.

1660 14.2.4 Theoretical uncertainty

1661 The theoretical uncertainty considered to be coming from imperfect knowledge of parton functions
1662 and is calculated as:

- 1663 • Error coming from an arbitrary choice of PDF set is estimated by PDF reweighting [87] of original
1664 MC generated using CT10 PDF set to one of the 4 pdf sets: ATLAS-epWZ12 [88], abkm09 [89]
1665 and NNPDF23 [90]. The error is calculated as a maximum deviation between the C-factor
1666 calculated using CT10 and C-factor from the different PDF set.
- Systematic uncertainty within one pdf set is evaluated using CT10 NLO set. This set contains
52 associated error sets, corresponding to a 90% C.L. limits along 26 eigenvectors. The resulting
52 variation are separately added in a quadrature as:

$$\delta_X = \frac{1}{2} \cdot \sqrt{\sum_{i=1}^N (X^+ - X^-)^2}, \quad (14.5)$$

1667 where the sum goes over N=26 eigenvectors. The X^+ and X^- are the up and down variations
1668 along one eigenvector.

1669 14.3 Statistical uncertainty in data

The statistical uncertainty is coming from the limited statistics in data and MC. It is calculated
separately for each analysis using the total number of events N observed after selection:

$$\delta N = \sqrt{N} \quad (14.6)$$

- 1670 The statistical uncertainty in data is a dominant one for all of the analyses.

Table 14.1: Summary of the different terms contributing to the uncertainty on C factors for all analyses

| | | $W^+ \rightarrow e\nu$ | $W^- \rightarrow e\nu$ | $W^+ \rightarrow \mu\nu$ | $W^+ \rightarrow \mu\nu$ | $Z \rightarrow ee$ | $Z \rightarrow \mu\mu$ |
|------------------------------|--------|--|--|--|--|--|--|
| Value C | | 0.6801 | 0.6799 | 0.7458 | 0.7491 | 0.5575 | 0.8020 |
| Source of uncertainty | Method | $\delta C_W/C_W(\%)$ $W^+ \rightarrow e\nu$ | $\delta C_W/C_W(\%)$ $W^- \rightarrow e\nu$ | $\delta C_W/C_W(\%)$ $W^+ \rightarrow \mu\nu$ | $\delta C_W/C_W(\%)$ $W^+ \rightarrow \mu\nu$ | $\delta C_Z/C_Z(\%)$ $Z \rightarrow ee$ | $\delta C_Z/C_Z(\%)$ $Z \rightarrow \mu\mu$ |
| Electron reconstruction | Toy MC | 0.11 | 0.09 | - | - | 0.12 | - |
| Electron identification | Toy MC | 0.32 | 0.30 | - | - | 0.54 | - |
| Electron trigger efficiency | Toy MC | 0.14 | 0.13 | - | - | 0.001 | - |
| Muon reco+id | Toy MC | - | - | 0.03 | 0.02 | - | 0.03 |
| Electron energy scale | Offset | 0.44 | 0.43 | - | - | 0.34 | - |
| - Statistical error | Offset | 0.05 | 0.04 | - | - | 0.04 | - |
| - Bias in method | Offset | 0.41 | 0.40 | - | - | 0.31 | - |
| - Scale uncertainty | Offset | 0.0 | 0.0 | - | - | 0.0 | - |
| - Low Pt | Offset | 0.0 | 0.0 | - | - | 0.003 | - |
| - Presampler energy scale | Offset | 0.04 | 0.03 | - | - | 0.04 | - |
| - Material knowledge | Offset | 0.14 | 0.13 | - | - | 0.14 | - |
| Electron energy resolution | Offset | 0.05 | 0.03 | - | - | 0.03 | - |
| Muon energy scale | Offset | - | - | 0.05 | 0.05 | - | 0.03 |
| Muon energy resolution total | Offset | - | - | 0.02 | 0.01 | - | 0.02 |
| - Muon ID energy scale | Offset | - | - | 0.02 | 0.01 | - | 0.01 |
| - Muon MS energy scale | Offset | - | - | 0.01 | 0.00 | - | 0.01 |
| Hadron recoil scale | Offset | 0.15 | 0.17 | 0.15 | 0.19 | - | - |
| Hadron recoil resolution | On/Off | 0.20 | 0.11 | 0.16 | 0.12 | - | - |
| MC statistics | | 0.24 | 0.31 | 0.24 | 0.34 | 0.30 | 0.43 |
| Theory | | 0.31 | 0.23 | 0.13 | 0.26 | 0.11 | 0.06 |
| Total | | 0.69 | 0.76 | 1.12 | 1.15 | 0.72 | 0.43 |
| | | Backgrounds | | | | | |
| EWK + $t\bar{t}$ background | Offset | 0.15 | 0.16 | 0.16 | 0.18 | 0.01 | 0.01 |
| QCD | Offset | 0.34 | 0.62 | 0.02 | 0.04 | - | - |
| | | Data statistics | | | | | |
| | | 1.60 | 2.13 | 1.51 | 2.02 | 4.82 | 3.93 |

Table 14.2: Acceptance values (A) and extrapolation values (E) and their relative uncertainties in percent for W and Z production in electron and muon channels. The various components of the uncertainties are defined in the text. The total uncertainties (δA_{tot} and δE_{tot}) are obtained as the quadratic sum of the three parts.

| | A | $\delta A_{err}^{pdf} (\%)$ | $\delta A_{sets}^{pdf} (\%)$ | $\delta A_{hs+ps} (\%)$ | $\delta A_{tot} (\%)$ |
|-------|------|-----------------------------|------------------------------|-------------------------|-----------------------|
| W^+ | 0.62 | 0.20 | 0.95 | 0.9 | 0.9 |
| W^- | 0.57 | 0.64 | 1.04 | 0.9 | 0.9 |
| Z | 0.63 | 0.40 | 1.37 | 0.9 | 0.9 |

14.4 Theoretical uncertainty on $A_{W/Z}$ and $E_{W/Z}$ factors

The effect of theoretical uncertainties must be considered for extrapolated cross-sections, through its effect on extrapolation factors $A_{W,Z}$, $E_{W,Z}$ factors. Since they are estimated using just the generator-level information, the uncertainties are purely theoretical. The A factors obtained and their theoretical uncertainties are summarized in Tab. 14.2. Main sources of uncertainties are:

- Error coming from an arbitrary choice of PDF set and systematic error within one pdf set is estimated in the same way, as for C_W (see Sec. 14.2.4). These sources are considered to be independent and added in a quadrature.
- The uncertainties arising from the choice of generator and parton shower model δA_{hs+ps} . They can be calculated as a difference in the acceptance $A_{W,Z}$ for MC samples, generated using the same PDF set, but different models for showering and matrix element, namely Powheg + Pythia and Sherpa. The systematic error obtained for W channels is 0.9% for A_W . It is consistent with 13 TeV and 7 TeV measurements.

Because of the lack of simulation samples for $Z \rightarrow ll$ using Sherpa generator a systematic uncertainty for Z is estimated using the fact, that for 13 TeV analysis and 8 TeV systematic errors coming from that source have been the same for A-factors in W^+ , W^- and Z.

The $E_{W,Z}$ factors, used for extrapolation to the 13 TeV fiducial phase-space are 1.09, 1.07, 1.11 for W^+ , W^- and Z bosons respectively. Because these values are close to 1, the theoretical uncertainties for these factors are considered negligible, compared to the experimental uncertainties on $C_{W,Z}$.

14.5 Correlation between uncertainties

In order to calculate W/Z ratios and combine channels different channels of the analysis it is crucial to take into account correlations between different channels. In this section the assumptions about correlations between channels will be discussed.

The theoretical uncertainties on A and C factors, except for the uncertainty within 1 pdf set are considered to be fully correlated between all analysis. The uncertainties within 1 pdf set are considered to be partially correlated (see Sec. 14.5.1). The systematic uncertainties from electroweak background sources are treated as uncorrelated between W and Z channels and 100% correlated for different W and Z channels.

1699 In addition to the electroweak background uncertainty the following systematic sources are con-
 1700 sidered to be fully correlated between $W^+ \rightarrow e\nu$, $W^- \rightarrow e\nu$, $W^+ \rightarrow \mu\nu$ and $W^- \rightarrow \mu\nu$:

- 1701 • QCD background
 1702 • Hadronic recoil scale
 1703 • Hadronic recoil resolution

1704 In addition to the mentioned systematics, the following uncertainties are considered 100% corre-
 1705 lated in electron analyses:

- 1706 • Electron energy scale
 1707 • Electron resolution

1708 and in muon analyses:

- 1709 • Muon energy scale
 1710 • Muon resolution
 1711 • Muon trigger efficiency

1712 The uncertainties, estimated using toy MC method are considered partially correlated and co-
 1713 variances between the analyses are estimated in the following section. The PDF uncertainties are
 1714 considered to be fully correlated for all of the analyses. The statistical uncertainty of MC is consid-
 1715 ered to be fully uncorrelated for all of the analyses. The total correlation coefficients matrices for
 1716 different analyses could be found in Appendix B

1717 **14.5.1 Treatment of partially correlated uncertainties**

1718 The following uncertainties are considered to be partially correlated between Z , W^+ and W^- analyses
 1719 for an observables o_X and o_Y :

- 1720 • Electron trigger efficiency
 1721 • Electron resolution efficiency
 1722 • Electron identification efficiency
 1723 • Muon reconstruction + identification efficiency

For each source of uncertainty a correlation coefficients between analysis X and Y can be estimated
 as:

$$\rho_{XY} = \frac{1}{\sigma(o_X)\sigma(o_Y)} \cdot \frac{1}{N} \sum_{i=1}^N (o_X^i - \bar{o}_X)(o_Y^i - \bar{o}_Y) = \frac{C_{XY}}{\sigma(o_X)\sigma(o_Y)}, \quad (14.7)$$

1724 where \bar{o}_X and \bar{o}_Y are the mean values of o_X and o_Y respectively, $\sigma(o_X)$ and $\sigma(o_Y)$ are the uncertainties
 1725 and i runs over the number of experiments N. C_{XY} denotes elements of the covariance matrix.
 1726 Resulting correlation matrices for each toy MC systematic source are shown in Fig. 14.1.

In order to include these covariances in further calculations, the Cholesky decomposition [91] have
 been performed. Using it, the covariance matrix C is re-written as:

$$C = L \cdot L^T, \quad (14.8)$$

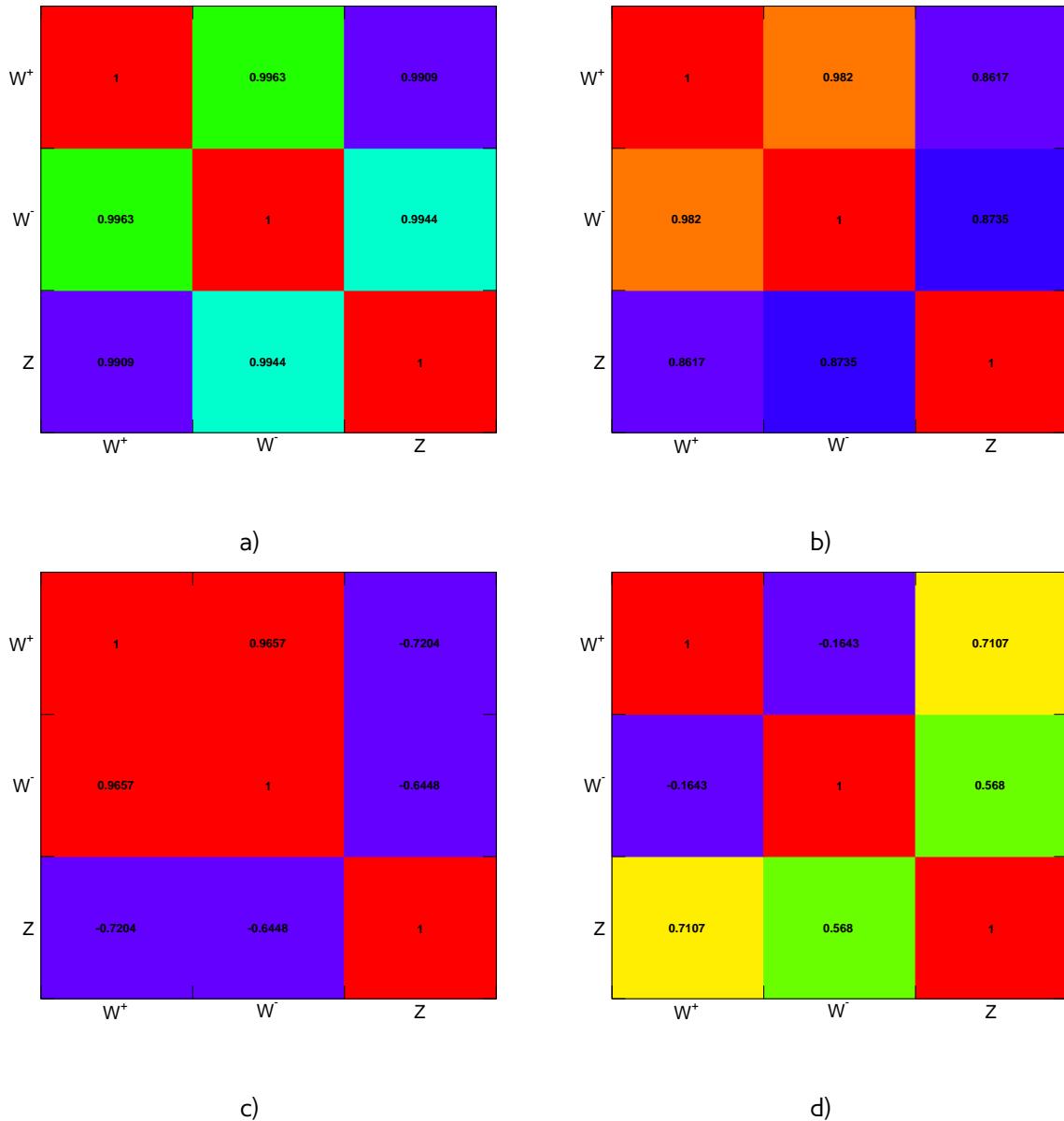


Fig. 14.1: Correlation coefficients ρ_{XY} among C_Z , C_{W^+} and C_{W^-} for a) electron reconstruction, b) electron identification, c) electron trigger and d) muon trigger scale factor uncertainties.

1727 where L is a lower triangular matrix, and L^T is a transpose of L . This decomposition is always unique
1728 because the covariance matrix has a positive definitive.

Rows of the matrix L are corresponding to an 3 systematic error vectors, that are fully correlated between W^+ , W^- and Z analyses. The quadratic sum of each row is corresponding to an total systematic uncertainty:

$$\sigma_X^{tot} = \sum_i^3 L_{iX}^2 \quad (14.9)$$

1729 The results of the Cholesky decomposition could be found in Appendix A.

1730 Chapter **15**

1731 **Results of the cross-section measurement**

| | | |
|------|--|-----|
| 1732 | 15.1 Cross-sections results | 129 |
| 1733 | 15.1.1 Lepton universality test | 129 |
| 1734 | 15.2 Combined results | 131 |
| 1735 | 15.2.1 Cross-sections ratios | 132 |
| 1736 | 15.3 Effect on PDF distributions | 137 |

1742 In this chapter results of the W/Z cross-section measurements will be discussed. In Sec. 15.1 the
1743 cross-sections measured in each lepton flavors are presented. These results are used to test lepton
1744 universality in 2.76 TeV data.

1745 Section 15.2 describes a results obtained for combined W and Z cross-sections. Additionally, in
1746 the subsection the cross-section ratios are presented. In the final section, the interpretation of the
1747 results using the the PDF profiling method is presented.

1748 **15.1 Cross-sections results**

1749 The cross-sections are calculated separately for each boson and lepton flavor as described in Chap. 3.
1750 The sources of uncertainties on the measurements are discussed in Chap. 14. Table 15.1 summarizes
1751 the obtained $W^{+/-}$ and Z cross-sections obtained from 2.76 TeV data in fiducial and extrapolated
1752 full and 13 TeV regions. The uncertainty coming from luminosity (labeled lumi) is a dominant one for
1753 this measurement. The statistical uncertainty is labeled stat. and is the second dominant uncertainty
1754 for the W and Z bosons cross-section. For W boson in electron channel the systematic uncertainty
1755 is around 1%, that makes it significantly lower, than a statistical uncertainty. For W boson in a muon
1756 channel cross-sections the overall systematical uncertainty is higher because of the trigger scale
1757 factors, and is around 1.2%, that makes it comparable with the statistical uncertainty. The results
1758 are consistent between lepton flavors.

1759 **15.1.1 Lepton universality test**

1760 Because of the lepton universality of the Standard Model, the results, obtained in electron and muon
1761 channel are expected to agree with each other. The 2.76 TeV data could be used to test it via fiducial
1762 cross-section ratios of e and μ branching fractions:

1763
$$R_W = \frac{\sigma_W^\mu}{\sigma_W^e} = \frac{BR(W \rightarrow \mu\nu)}{BR(W \rightarrow e\nu)} = 1.012 \pm 0.026(\text{sys.}) \pm 0.019(\text{stat.}),$$

Table 15.1: Results on a fiducial σ^{fid} and total cross-section measurement for W^+ , W^- and Z bosons in electron and muon channels. The cross-sections are shown with their statistical, systematical and luminosity uncertainties (and extrapolation error for total cross-section) quoted in that order

| | value \pm stat \pm syst \pm lumi (\pm ext) | value \pm stat \pm syst \pm lumi (\pm ext) |
|-----------------------|---|---|
| | W^+ | |
| | $W^+ \rightarrow e\nu$ | $W^+ \rightarrow \mu\nu$ |
| σ_W^{fid} [pb] | $1406.0 \pm 23.3 \pm 9.7 \pm 43.6$ | $1430.3 \pm 22.4 \pm 16.2 \pm 44.3$ |
| σ_W^{tot} [pb] | $2253.5 \pm 37.3 \pm 15.5 \pm 69.9 \pm 24.8$ | $2292.6 \pm 36.0 \pm 25.9 \pm 71.1 \pm 27.1$ |
| σ_W^{13} [pb] | $1293.9 \pm 21.4 \pm 8.9 \pm 40.1$ | $1316.7 \pm 20.7 \pm 14.9 \pm 40.8$ |
| | W^- | |
| | $W^- \rightarrow e\nu$ | $W^- \rightarrow \mu\nu$ |
| σ_W^{fid} [pb] | $783.7 \pm 17.5 \pm 5.3 \pm 24.3$ | $794.9 \pm 16.8 \pm 9.2 \pm 24.6$ |
| σ_W^{tot} [pb] | $1373.0 \pm 30.7 \pm 9.3 \pm 42.6 \pm 17.4$ | $1393.9 \pm 29.4 \pm 16.1 \pm 43.2 \pm 32.4$ |
| σ_W^{13} [pb] | $733.8 \pm 16.4 \pm 5.0 \pm 22.7$ | $743.7 \pm 15.7 \pm 8.6 \pm 23.1$ |
| | Z | |
| | $Z \rightarrow ee$ | $Z \rightarrow \mu\mu$ |
| σ_Z^{fid} [pb] | $194.8 \pm 9.4 \pm 1.4 \pm 6.0$ | $203.5 \pm 8.0 \pm 0.9 \pm 6.3$ |
| σ_Z^{tot} [pb] | $310.4 \pm 15.0 \pm 2.2 \pm 9.6 \pm 6.2$ | $322.9 \pm 12.7 \pm 1.4 \pm 10.0 \pm 5.9$ |
| σ_Z^{13} [pb] | $176.1 \pm 8.5 \pm 1.3 \pm 5.5$ | $183.2 \pm 7.2 \pm 0.8 \pm 5.7$ |

where the W cross-sections are calculated in fiducial region following the prescription from Sec. 3.1.3:

$$\begin{aligned}\sigma_W^{fid}(W \rightarrow e\nu) &= 2190.0 \pm 29.1(stat.) \pm 7.8(sys.) \pm 65.7(lumi.) [pb] \\ \sigma_W^{fid}(W \rightarrow \mu\nu) &= 2217.0 \pm 27.9(stat.) \pm 50.6(sys.) \pm 66.5(lumi.) [pb]\end{aligned}$$

This result agrees within the uncertainty with the world average of 0.991 ± 0.018 [25].

Similarly, this ratio can be measured in a Z boson decays as:

$$R_Z = \frac{\sigma_Z^\mu}{\sigma_Z^e} = \frac{BR(Z \rightarrow \mu\mu)}{BR(Z \rightarrow ee)} = 1.046 \pm 0.004(sys.) \pm 0.065(stat.)$$

This ratio value is statistics dominated. The world average for a corresponding value is a 1.0009 ± 0.0028 [25].

Comparison of the R_W and R_Z with the respect of the correlated systematic uncertainties with the world average is shown in Fig. 15.1. The correlation matrix for R_W and R_Z could be found in appendix. The ellipse angle is obtained from the correlation matrix using the eigenvector decomposition and corresponds to the angle between x axis and the one of the 2 eigenvectors. The obtained values are agreeing withing the systematic uncertainty with the standard model expectations and world averages.

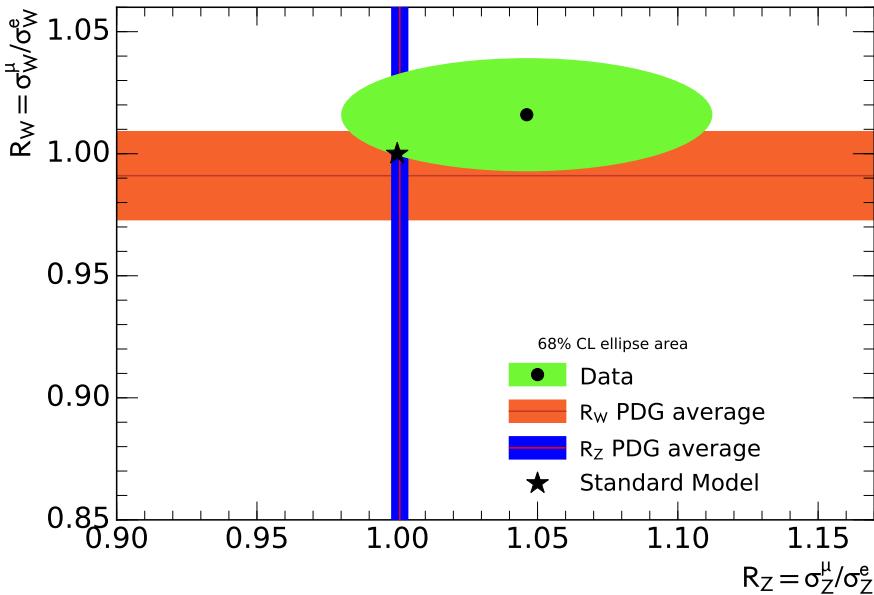


Fig. 15.1: The correlated measurement of the electron-to-muon fiducial cross-section ratios in the W and the Z channels. The vertical (horizontal) band represents the uncertainty of the corresponding Z (W) branching fractions based on the current world average data. The green ellipse illustrates the 68% CL for the correlated measurement of R_W and R_Z .

15.2 Combined results

Since the results between channels are agreeing well, it is possible to perform averaging as described in Sec. 3.3. The combination is done for fiducial region, and then combined cross-sections are extrapolated to the full and new 13 TeV regions. Systematic uncertainties for the averaging are taken from Chap. 14 with the respect of correlated uncertainties. The common luminosity uncertainty is excluded from the combination process.

The resulting cross-sections are summarized in Tab. 15.2. The combination procedure allows to significantly reduce statistical uncertainty of the measurement compared to the separate cross-sections. The systematic uncertainty is also visibly reduced, because most of the sources are uncorrelated for different flavors of the analysis. The combination yields a good $\chi^2/NDF \approx 1/3$ indicating good agreement between measurements. The W cross-section is calculated from the combined W^+ and W^- cross-sections and is also in a good agreement with separate channels (Fig. 15.3 b).

The theoretical predictions are obtained at NLO and NNLO level of precision. The NLO calculations are performed using the MCFM generator [92], interfaced for a faster calculation with APPLGRID [93], that provides a x vs Q^2 grid for a calculation and convolution with the given PDF set. The NNLO predictions have been provided from the FEWZ program [26].

The comparison of NLO and NNLO predictions for CT14nnlo [27] pdf set for the cross-sections in fiducial region with the experimental data for separate lepton and combined channels are shown in Fig. 15.2 and Fig. 15.3. The NLO and NNLO cross-sections are agreeing with each other and experimental data withing the PDF uncertainty. The NLO cross-sections are visibly smaller, than the obtained experimental results and have a higher uncertainty. The NNLO predictions have better agreement

Table 15.2: Results on a fiducial σ^{fid} and total cross-section measurement for W^+ , W^- and Z bosons in electron and muon channels. The cross-sections are shown with their statistical, systematical and luminosity uncertainties (and extrapolation error for total cross-section) quoted in that order

| | value \pm stat \pm syst \pm lumi (\pm ext) | value \pm stat \pm syst \pm lumi (\pm ext) |
|-----------------------|---|---|
| | | $W^{+/-}$ |
| | $W^+ \rightarrow l\nu$ | $W^- \rightarrow l\nu$ |
| σ_W^{fid} [pb] | $1417.2 \pm 16.2 \pm 6.7 \pm 43.9$ | $789.1 \pm 12.1 \pm 4.5 \pm 24.5$ |
| σ_W^{tot} [pb] | $2271.2 \pm 25.9 \pm 10.7 \pm 70.4 \pm 0.9$ | $1382.0 \pm 21.2 \pm 7.8 \pm 42.8 \pm 0.9$ |
| σ_W^{13} [pb] | $1303.8 \pm 14.9 \pm 6.1 \pm 40.4$ | $738.9 \pm 11.3 \pm 4.2 \pm 22.9$ |
| | | $W \rightarrow l\nu$ |
| σ_W^{fid} [pb] | | $2206.3 \pm 20.2 \pm 10.6 \pm 68.4$ |
| σ_W^{tot} [pb] | | $5817.1 \pm 53.3 \pm 27.9 \pm 180.3 \pm 0.9$ |
| σ_W^{13} [pb] | | $1846.9 \pm 16.9 \pm 8.8 \pm 57.3$ |
| | | $Z \rightarrow ll$ |
| σ_Z^{fid} [pb] | | $200.4 \pm 6.1 \pm 0.3 \pm 6.2$ |
| σ_Z^{tot} [pb] | | $319.1 \pm 9.8 \pm 0.5 \pm 9.9 \pm 0.9$ |
| σ_Z^{13} [pb] | | $181.2 \pm 5.5 \pm 0.3 \pm 5.6$ |

1799 and a similar uncertainty.

1800 Additionally, the obtained W^+ , W^- and Z cross-sections in a combined channel are compared
1801 to the NNLO predictions for a following PDF sets: ABM12nlo [94], CT14nnlo [27], MMHTnnlo [19], AT-
1802 LASepWZ12 [88], NNPDF3.0 [90] and HERApdf2.0nnlo [95] in Fig. 15.4-15.6. The best overall agreement
1803 is achieved in a NNPDF3.0 pdf set. Additional NNLO comparisons for full and 13 TeV regions could
1804 be found in the appendix.

1805 15.2.1 Cross-sections ratios

1806 Measurement of the ratios is a powerful tool to test PDF predictions, since it allows to cancel the
1807 biggest uncertainty coming from luminosity and reduce other sources of systematic uncertainties.
1808 The ratios of W and Z cross-sections have been measured for 7 TeV [83] and 13 TeV [28] analyses.
1809 The measurement of the ratio R_{W^+/W^-} is sensitive to the u_v , d_v valence quarks distributions, while
1810 the ratio $R_{W/Z}$ can put a constraints on the strange quark distributions. The ratios for W/Z cross
1811 section have been calculated in a fiducial region following the prescription from Sec. 3.2 for the
1812 electron channel analyses:

$$\begin{aligned} R_{W/Z}^e &= 11.231 \pm 0.563 \text{ (stat.)} \pm 0.058 \text{ (sys.)} \\ R_{W^+/Z}^e &= 7.210 \pm 0.369 \text{ (stat.)} \pm 0.034 \text{ (sys.)} \\ R_{W^-/Z}^e &= 4.021 \pm 0.214 \text{ (stat.)} \pm 0.030 \text{ (sys.)} \\ R_{W^+/W^-}^e &= 1.793 \pm 0.050 \text{ (stat.)} \pm 0.004 \text{ (sys.)} \end{aligned}$$

1817 and muon channel analyses:

$$R_{W/Z}^\mu = 10.907 \pm 0.452 \text{ (stat.)} \pm 0.150 \text{ (sys.)}$$

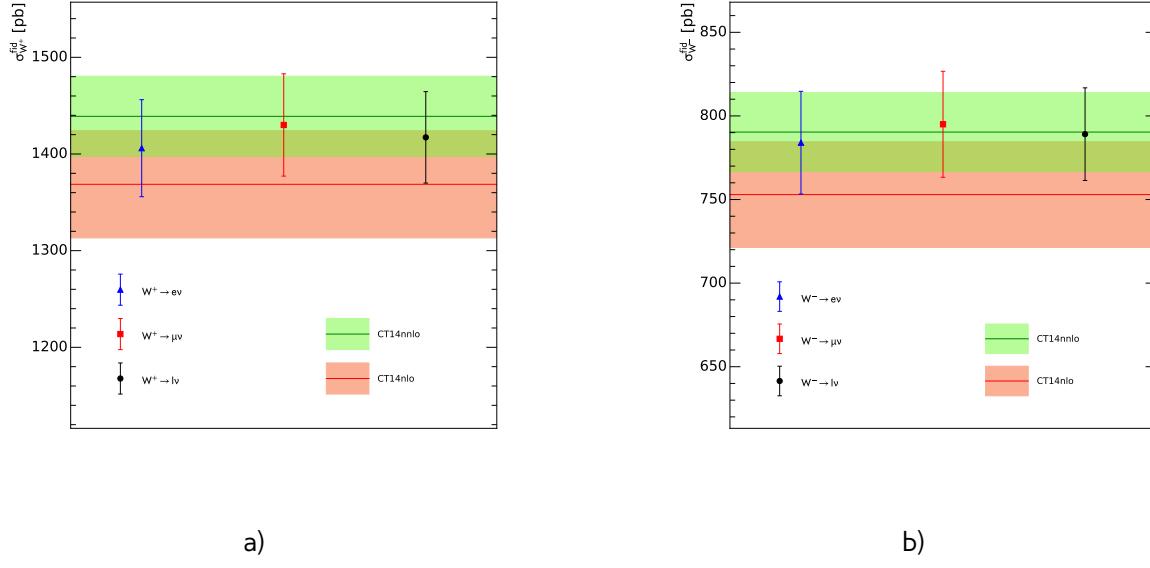


Fig. 15.2: The NLO and NNLO theoretical predictions calculated using the CT14nnlo PDF set compared to the measured fiducial cross-sections as given in Tab. 15.1 and Tab. 15.2 for a) $\sigma_{W^+}^{fid}$ and b) $\sigma_{W^-}^{fid}$. The blue and red dots are corresponding to the electron and muon channels respectively, while black dots are representing the combined channel. The NLO and NNLO predictions are presented by the red and green lines with error-bands respectively.

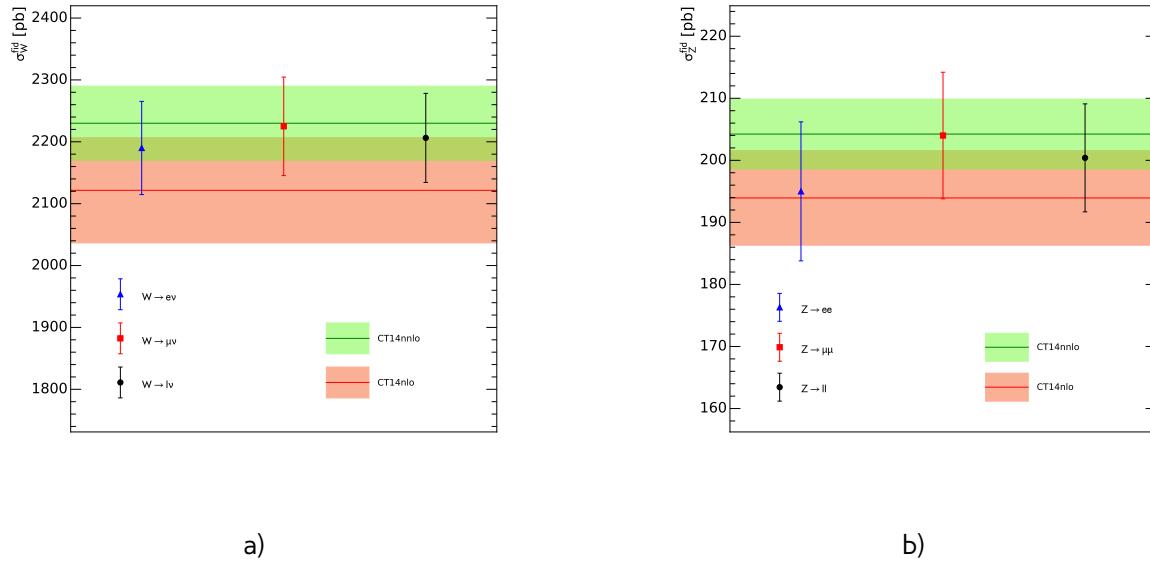


Fig. 15.3: The NLO and NNLO theoretical predictions calculated using the CT14nnlo PDF set compared to the measured fiducial cross-sections as given in Tab. 15.1 and Tab. 15.2 for a) σ_W^{fid} and b) σ_Z^{fid} . The blue and red dots are corresponding to the electron and muon channels respectively, while black dots are representing the combined channel. The NLO and NNLO predictions are presented by the red and green lines with error-bands respectively.

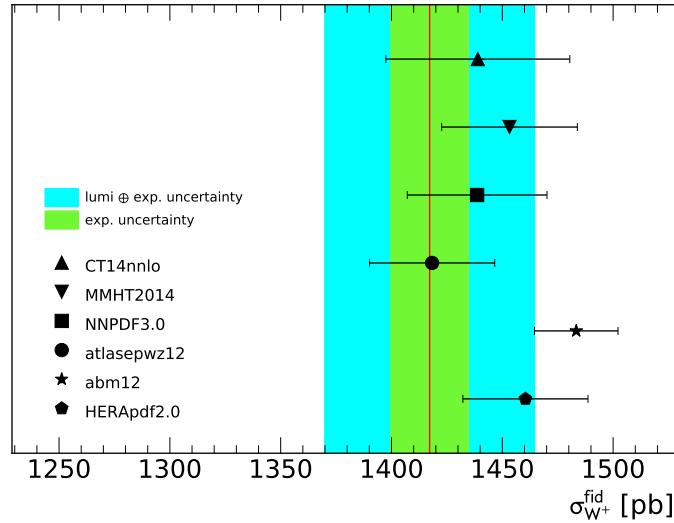


Fig. 15.4: NNLO predictions for the fiducial cross-section $\sigma_{W^+}^{fid}$ in pb for the six PDFs CT14nnlo, MMHT2014, NNPDF3.0, ATLASepWZ12, abm12, HERApdf2.0 compared to the measured fiducial cross-section as given in Tab. 15.2. The green (cyan) band corresponds to the experimental uncertainty without (with) the luminosity uncertainty. The theory predictions are given with the corresponding PDF uncertainties shown as error bands.

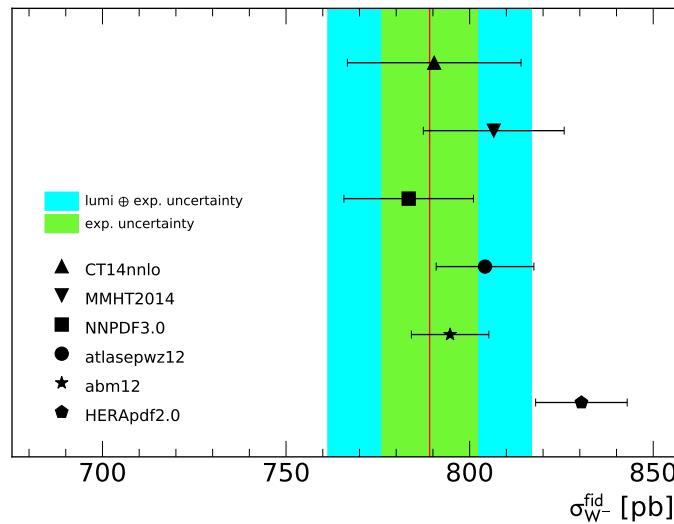


Fig. 15.5: NNLO predictions for the fiducial cross-section $\sigma_{W^-}^{fid}$ in pb for the six PDFs CT14nnlo, MMHT2014, NNPDF3.0, ATLASepWZ12, abm12, HERApdf2.0 compared to the measured fiducial cross-section as given in Tab. 15.2. The green (cyan) band corresponds to the experimental uncertainty without (with) the luminosity uncertainty. The theory predictions are given with the corresponding PDF uncertainties shown as error bands.

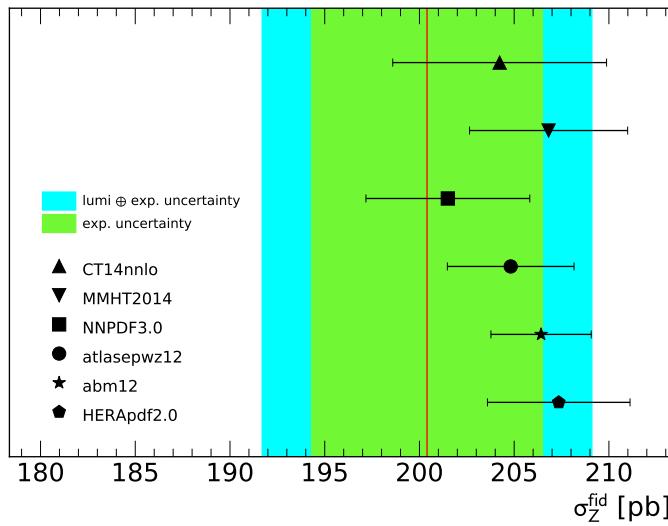


Fig. 15.6: Predictions for the fiducial cross-section σ_Z^{fid} in pb for the six PDFs CT14nnlo, MMHT2014, NNPDF3.0, ATLASepWZ12, abm12, HERApdf2.0 compared to the measured fiducial cross-section as given in Tab. 15.2. The green (cyan) band corresponds to the experimental uncertainty without (with) the luminosity uncertainty. The theory predictions are given with the corresponding PDF uncertainties shown as error bands.

$$\begin{aligned} R_{W+/Z}^\mu &= 7.010 \pm 0.298 \text{ (stat.)} \pm 0.097 \text{ (sys.)} \\ R_{W-/Z}^\mu &= 3.897 \pm 0.174 \text{ (stat.)} \pm 0.057 \text{ (sys.)} \\ R_{W+/W^-}^\mu &= 1.799 \pm 0.047 \text{ (stat.)} \pm 0.003 \text{ (sys.)} \end{aligned}$$

The obtained results are in a good agreement with the combined channel results, that have a reduced uncertainty:

$$\begin{aligned} R_{W/Z} &= 11.010 \pm 0.351 \text{ (stat.)} \pm 0.052 \text{ (sys.)} \\ R_{W+/Z} &= 7.072 \pm 0.231 \text{ (stat.)} \pm 0.033 \text{ (sys.)} \\ R_{W-/Z} &= 3.938 \pm 0.135 \text{ (stat.)} \pm 0.022 \text{ (sys.)} \\ R_{W+/W^-} &= 1.796 \pm 0.034 \text{ (stat.)} \pm 0.002 \text{ (sys.)} \end{aligned}$$

The resulting ratios are statistically dominated. Because of the good agreement between the ratios in a different channels, the ratios of combined results are used to compare with NLO predictions (Fig. 15.7 and Fig. 15.8) for the six PDFs CT14nnlo, MMHT2014, NNPDF3.0, ATLASepWZ12, abm12, HERApdf2.0. Because of the higher statistics for the W-bosons for the ratio R_{W+/W^-} the resulting error is compatible with PDF uncertainties. The ratios to the Z cross-sections $R_{W/Z}$, $R_{W+/Z}$ and $R_{W-/Z}$ are significantly less sensitive because of the large statistical uncertainty of the Z cross-sections. The best agreement with the predictions is achieved by the MMHT pdf.

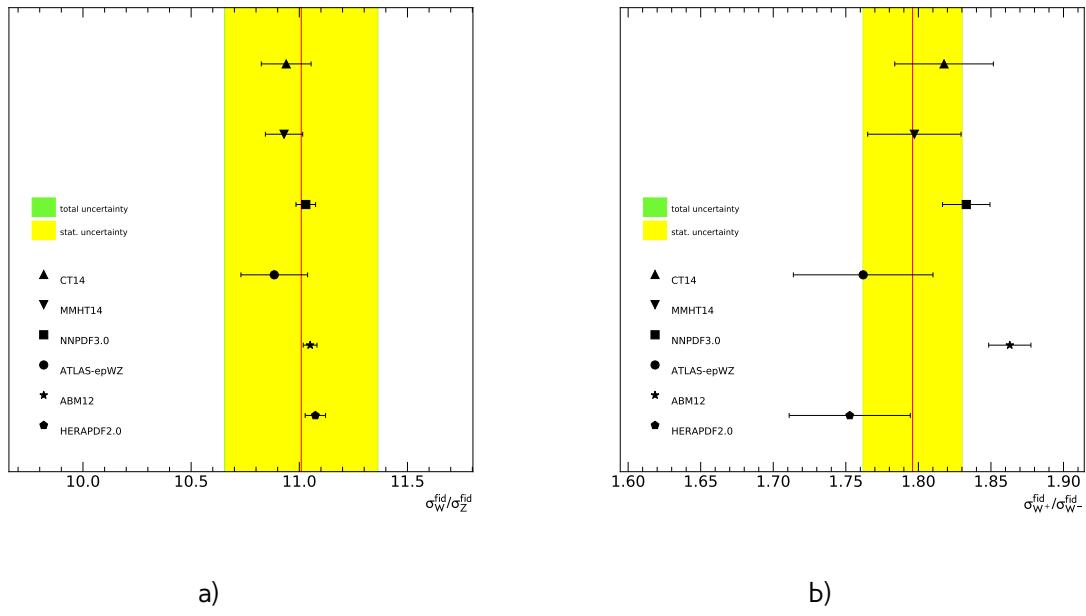


Fig. 15.7: Ratio of the a) W to Z and b) W^+ to W^- production fiducial cross-sections compared to predictions based on the six pdf sets: CT14nnlo, MMHT2014, NNPDF3.0, ATLASepWZ12, abm12, HERApdf2.0. The yellow band corresponds to the statistical uncertainty, while the systematical uncertainty is considered to be negligible. The theory predictions are given with the corresponding PDF uncertainties shown as error bands.

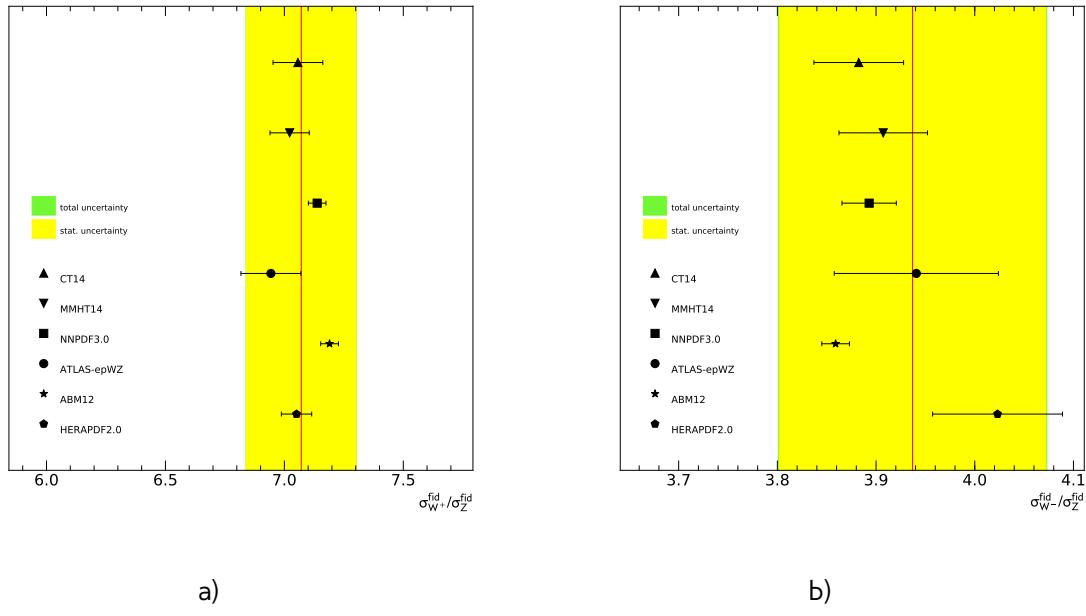


Fig. 15.8: Ratio of the a) W^+ to Z and b) W^- to Z production fiducial cross-sections compared to predictions based on the six pdf sets: CT14nnlo, MMHT2014, NNPDF3.0, ATLASepWZ12, abm12, HERApdf2.0. The yellow band corresponds to the statistical uncertainty, while the systematical uncertainty is considered to be negligible. The theory predictions are given with the corresponding PDF uncertainties shown as error bands.

1835 15.3 Effect on PDF distributions

1836 The effect of inclusion of obtained cross-sections in PDF set have been estimated using the profiling
 1837 method, described in Sec. 3.4. As a reference, it was decided to use CT14 pdf set, because of its
 1838 relatively good agreement with data for both NLO and NNLO predictions. The profiling was performed
 1839 at NLO order, however it is also possible to make NNLO profiling with use of the NNLO K-factors.

1840 As it was mentioned in Chap. ?? the measurement at 2.76 TeV is mostly sensitive to the valence
 1841 u_v and d_v , light-sea \bar{d} and \bar{u} quark distributions. The inclusion of 2.76 TeV cross-sections in PDF
 1842 can introduce both the reduction of the uncertainties of the PDF distributions and shift in the
 1843 distributions.

1844 The sensitivity of this mesurement on the PDF uncertainties have been studied by adding in
 1845 the PDF set the W and Z cross-sections scaled to match the theoretical predictions. This method
 1846 does not introduce any shift in distributions, however, it makes a reduction of uncertainties more
 1847 visible. The resulting distributions are shown in Fig. 15.9 at the initial scale $Q^2 = 1.9 \text{ GeV}^2$. This
 1848 measurement allows to significantly reduce the uncertainties on \bar{u} and \bar{d} distributions. There is also
 1849 a visible reduction of the uncertainties in the small x region for the valence quarks. Due to the limited
 1850 statistics, the measurement can not reduce the uncertainties on the strange quark distributions. As
 1851 it was expected, the W and Z cross-sections are not sensitive to the gluon density. It is possible
 1852 to reduce the uncertainties on PDF distribution even more with inclusion of 5, 7 and 13 TeV W and
 1853 Z cross-sections, because of the large number of correlated uncertainties for a different energy
 1854 measurements (especially luminosity).

1855 The full profiling results are shown in Fig. 15.10- 15.11. The initial value of $\chi^2/NDF=1.2/3$ shows
 1856 the good agreement with theoretical predictions, however the fit allows to reduce it to the down
 1857 to the value $\chi^2/NDF=0.8/3$. The profiling procedure introduces a shift in u_v , d_v , \bar{u} , \bar{d} , s quark
 1858 distributions within 1 sigma. The gluon distribution is left untouched. The additional PDF profiling
 1859 plots can be found in Appendix D.

1860 The predictions of the obtained profiled CT14 set are shown in Fig. 15.12- 15.13. The cross-sections,
 1861 predicted by the profiled CT14 set are in a better agreement, however not perfect, agreement with
 1862 the data, compared to the original set. There is also a small reduction of the uncertainties.

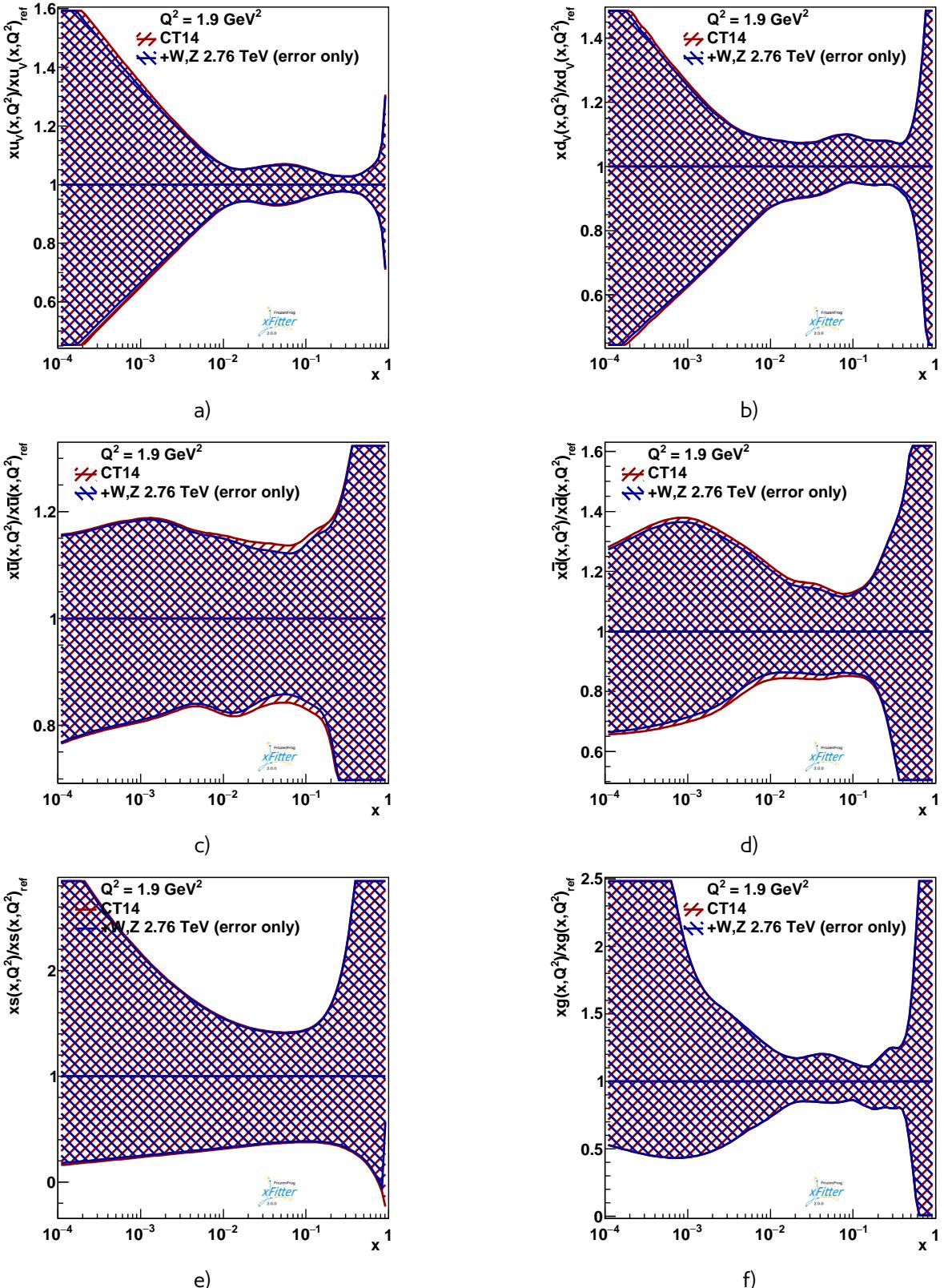


Fig. 15.9: The relative experimental uncertainties for the quark and gluon densities as a function of x at scale $Q^2 = 1.9 \text{ GeV}^2$. The red band denotes the reference NLO PDF distributions from CT14 pdf set. The impact of addition of the new W,Z cross-sections at 2.76 TeV on the PDF set uncertainties is shown by the blue boundaries.

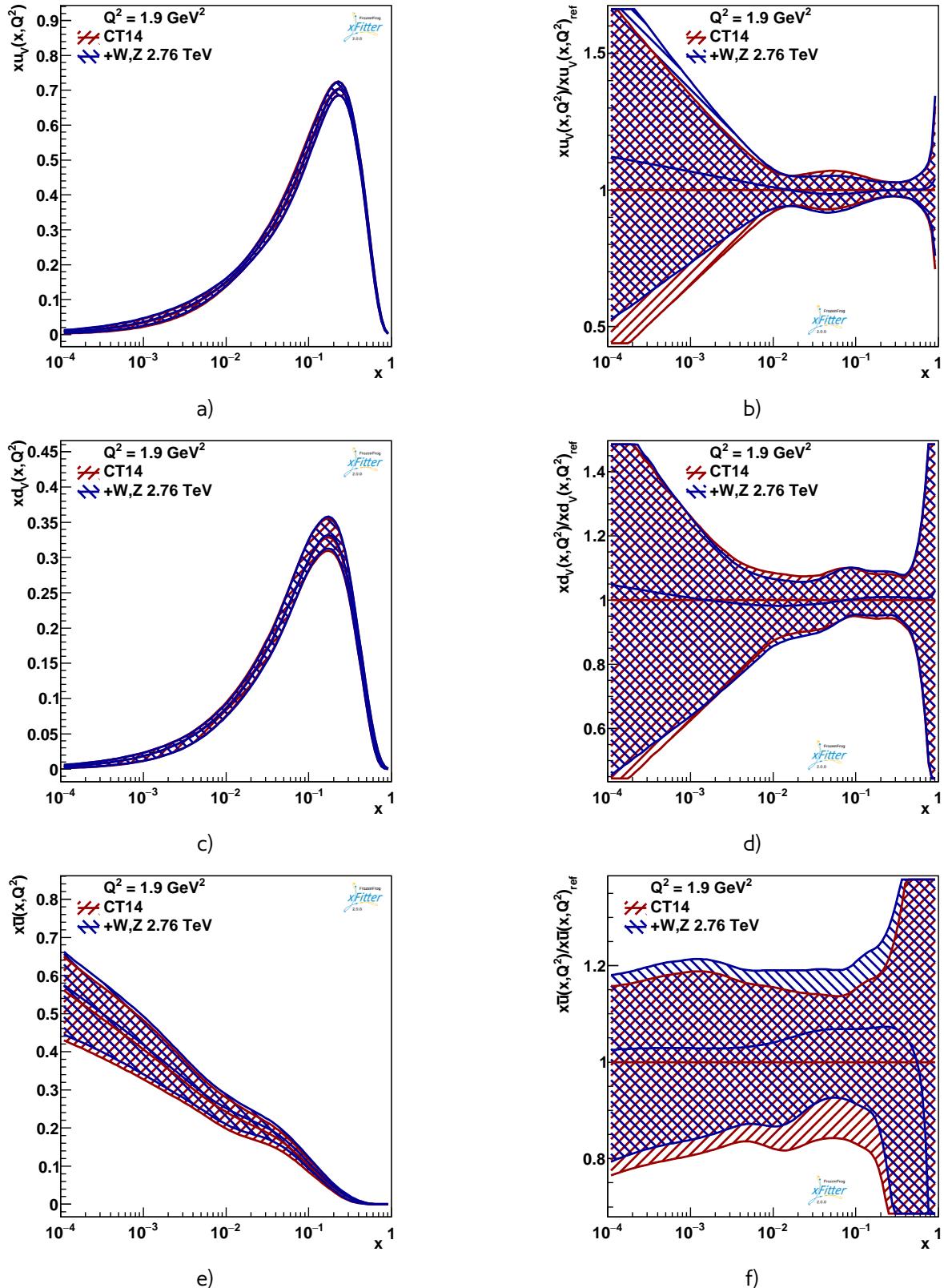


Fig. 15.10: The absolute and relative distributions for the u_v , d_v , \bar{u} quark densities as a function of x at scale $Q^2 = 1.9 \text{ GeV}^2$ with the experimental uncertainties. The red band denotes the reference NLO PDF distributions from CT14 pdf set. The impact of addition of the new W,Z cross-sections at 2.76 TeV on the PDF set is shown by the blue boundaries.

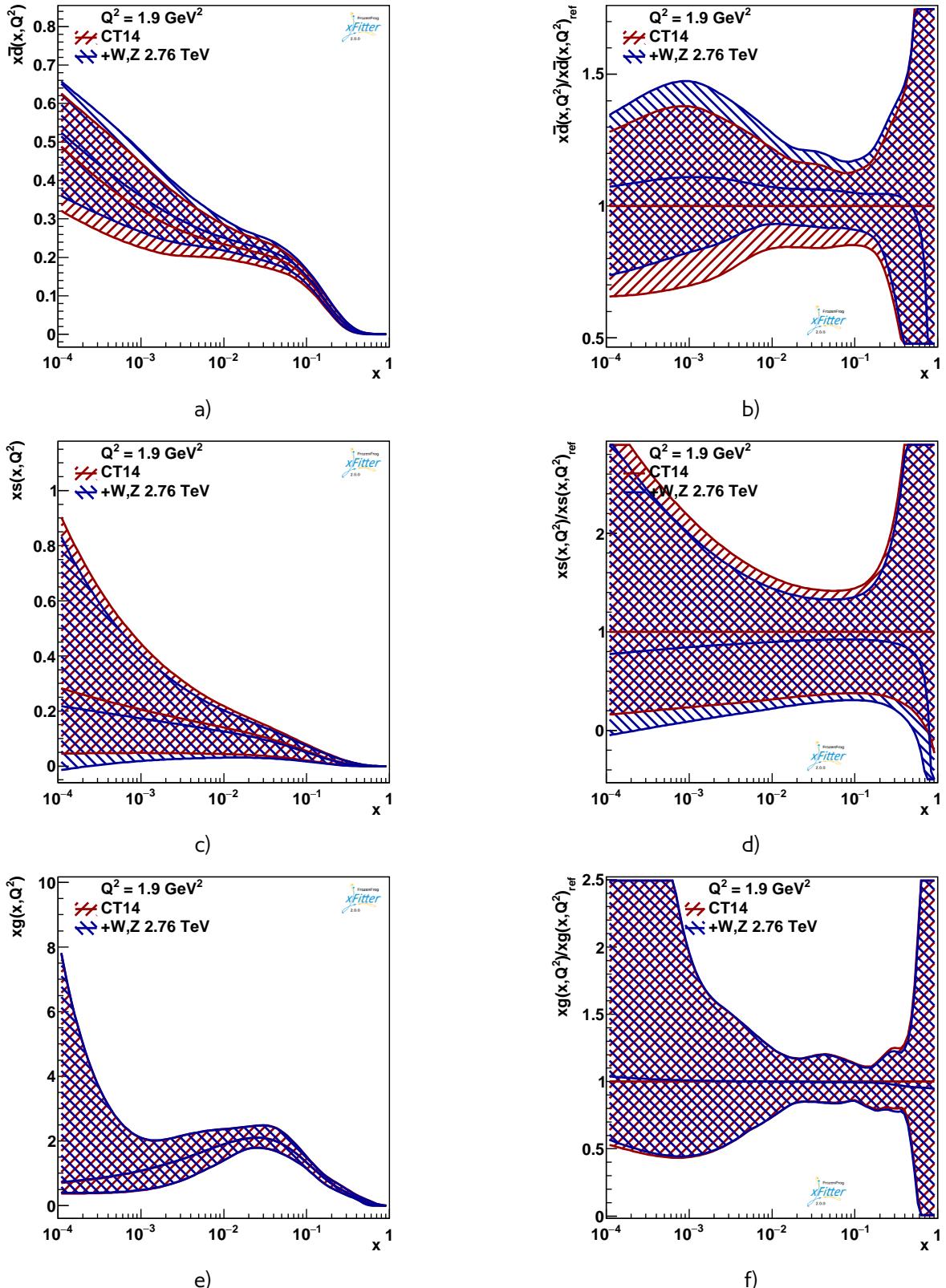


Fig. 15.11: The absolute and relative distributions for the \bar{d} and s quark and gluon densities as a function of x at scale $Q^2 = 1.9 \text{ GeV}^2$ with the experimental uncertainties. The red band denotes the reference NLO PDF distributions from CT14 pdf set. The impact of addition of the new W,Z cross-sections at 2.76 TeV on the PDF set is shown by the blue boundaries.

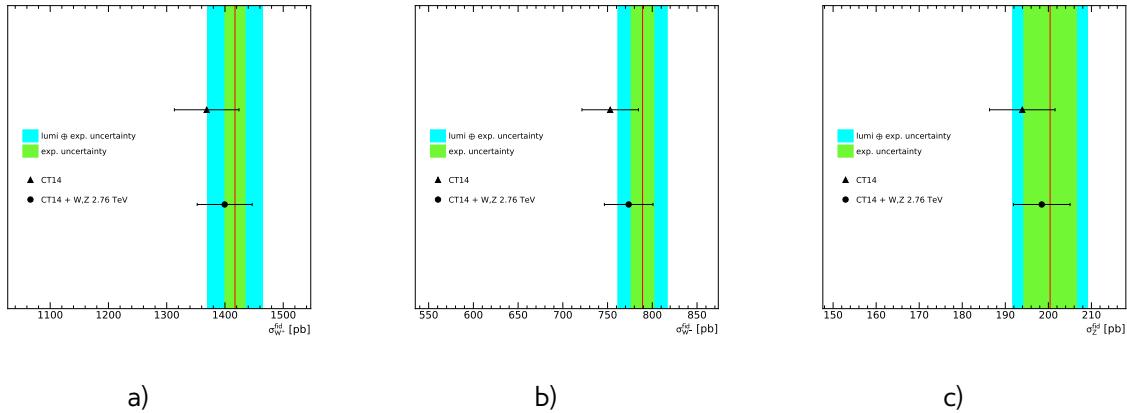


Fig. 15.12: The obtained a) W^+ b) W^- c) Z production fiducial cross-sections in combined channel compared to NLO predictions based on original and profiled CT14 PDF set.

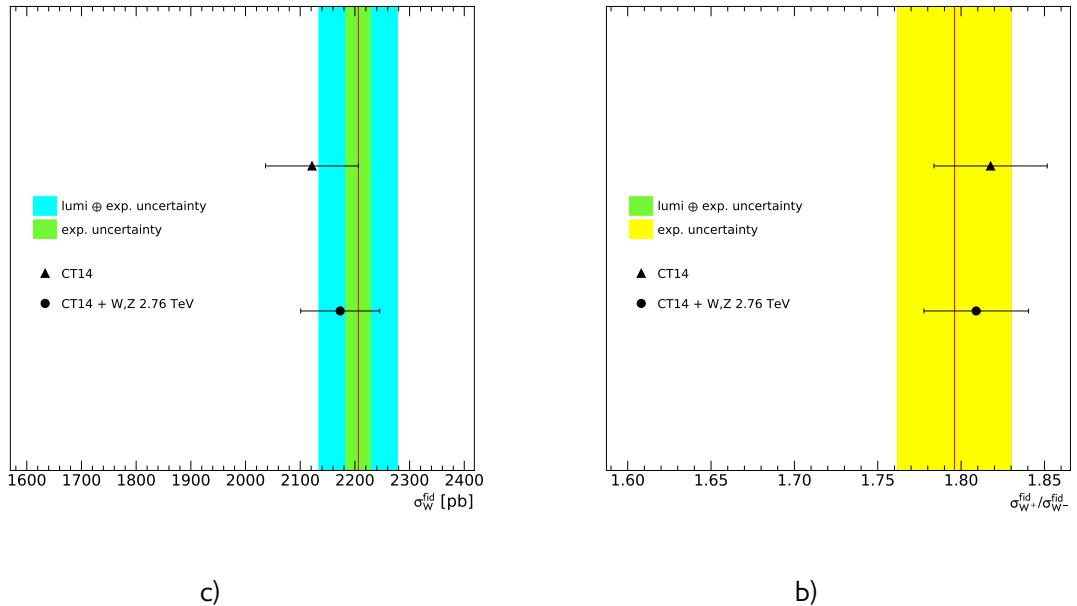


Fig. 15.13: The obtained a) W and b) ratio of the W^+ to W^- production fiducial cross-sections in combined channel compared to NLO predictions based on original and profiled CT14 PDF set.

1863

Chapter 16

1864

Summary

1865 The measurement of $W \rightarrow l\nu$ and $Z \rightarrow ll$ cross-sections in electron and muon channels using
1866 4pb^{-1} ATLAS data at center-of-mass energy 2.76 TeV have been performed. The obtained data is
1867 used (assuming the lepton universality) for calculation of the combined cross-section and ratios of
1868 the W and Z cross-sections.

1869 The uncertainty of these measurements is dominated by the limited statistics. Several sources
1870 of the systematic uncertainties were studied as well. The studies of the hadronic recoil calibration
1871 have been performed. The corresponding uncertainties have small, but not negligible contribution
1872 to the systematic uncertainties of the W analyses.

1873 The contribution from the background processes have been estimated using both MC simulation
1874 and data-driven method.

1875 The measured cross-section is compared to the both NLO and NNLO predictions for the different
1876 PDF set. The results are agreeing well with the theoretical predictions.

1877 The measured data have been used for the PDF fits.

Appendix A

1879 Cholesky transformation results

1880 In this Appendix results of Cholesky decomposition for partially correlated uncertainties (see Sec. 14.5.1)
 1881 are presented. The resulting uncorrelated systematic sources are shown in Fig. A.1 in source vs the
 1882 analysis plane. The content of the cell corresponds to the value of $\delta C/C(\%)$. The resulted systematic
 1883 uncertainties are included in averaging process and PDF profiling.

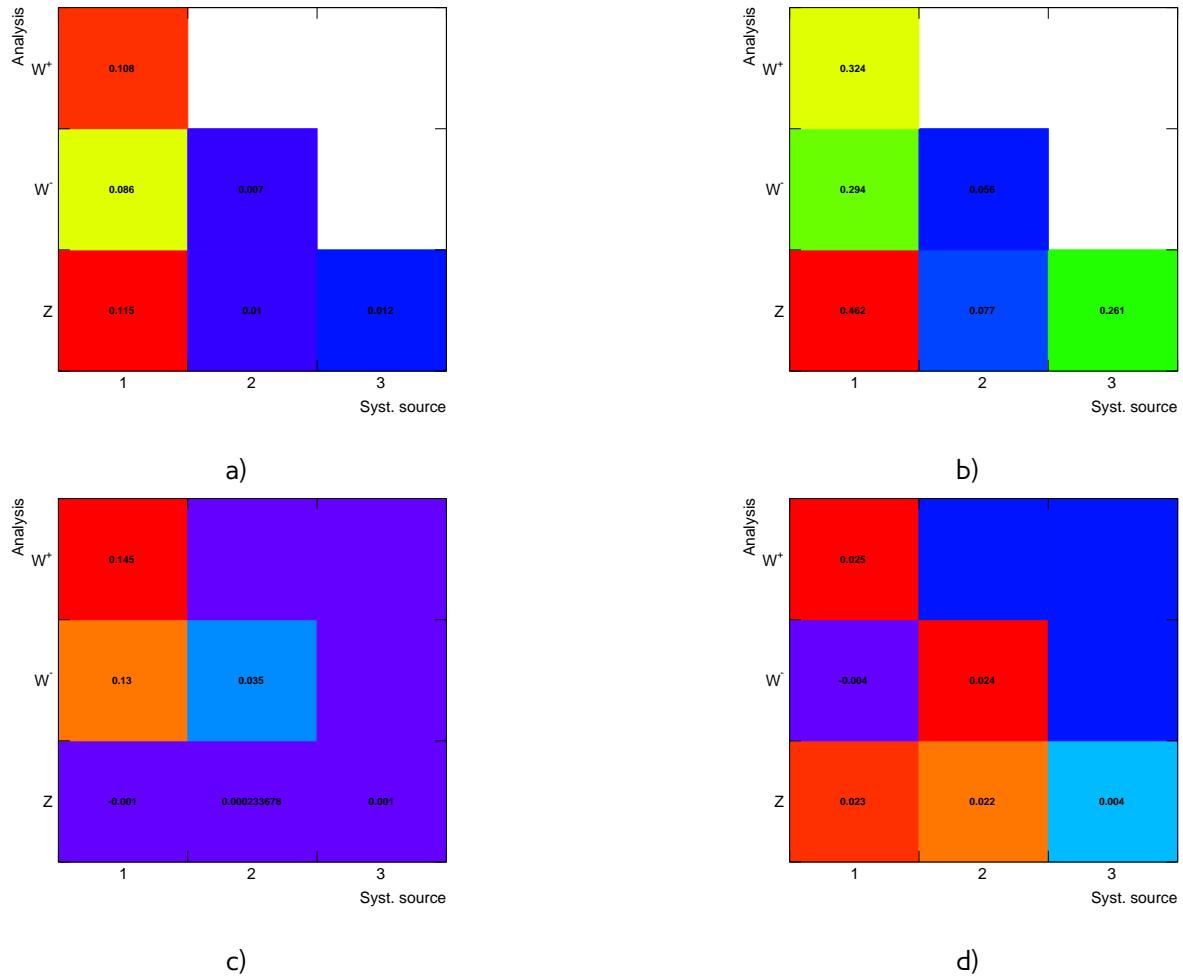


Fig. A.1: Results of Cholesky decomposition for correlated uncertainties for a) electron reconstruction, b) electron identification, c) electron trigger and d) muon trigger scale factor uncertainties estimated using Toy MC method.

Appendix B

1885 Covariance matrixes

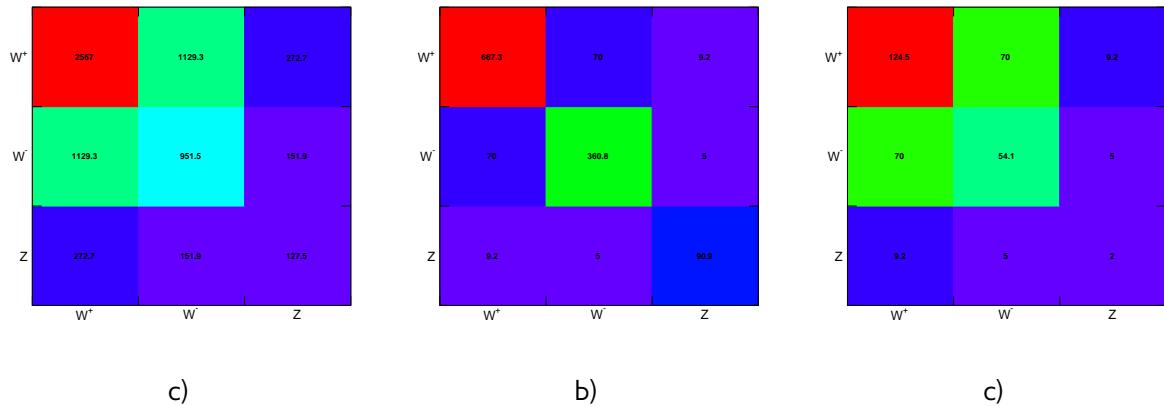


Fig. B.1: Covariance matrix for the measurements of Z , W^+ and W^- cross sections for electron channel in fiducial region a)for all uncertainty b)for all but luminosity uncertainty c)for all but luminosity and statistical uncertainty.

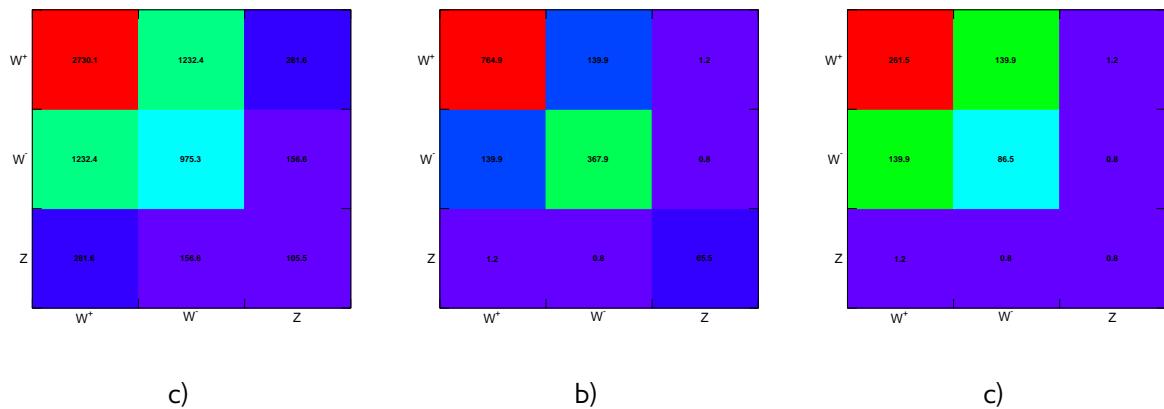


Fig. B.2: Covariance matrix for the measurements of Z , W^+ and W^- cross sections for muon channel in fiducial region a)for all uncertainty b)for all but luminosity uncertainty c)for all but luminosity and statistical uncertainty.

Covariance matrixes

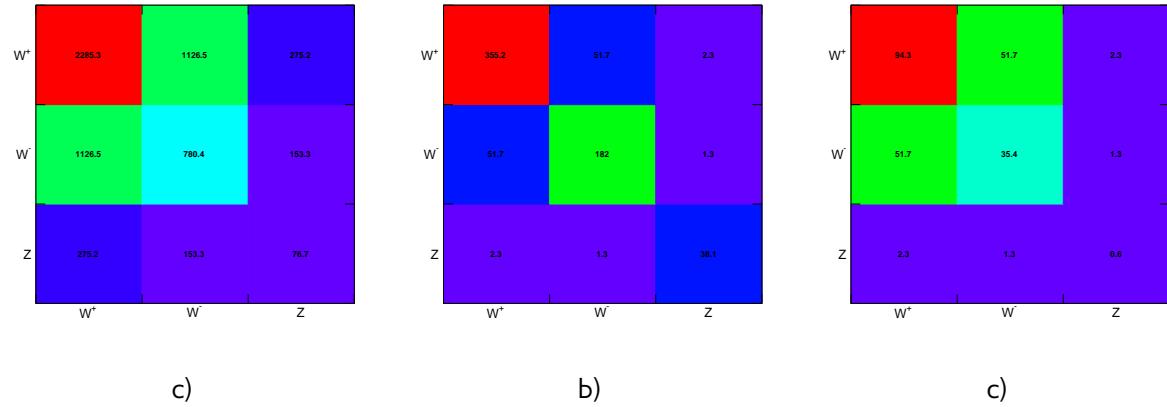


Fig. B.3: Covariance matrix for the measurements of Z , W^+ and W^- cross sections for combined channel in fiducial region a)for all uncertainty b)for all but luminosity uncertainty c)for all but luminosity and statistical uncertainty.

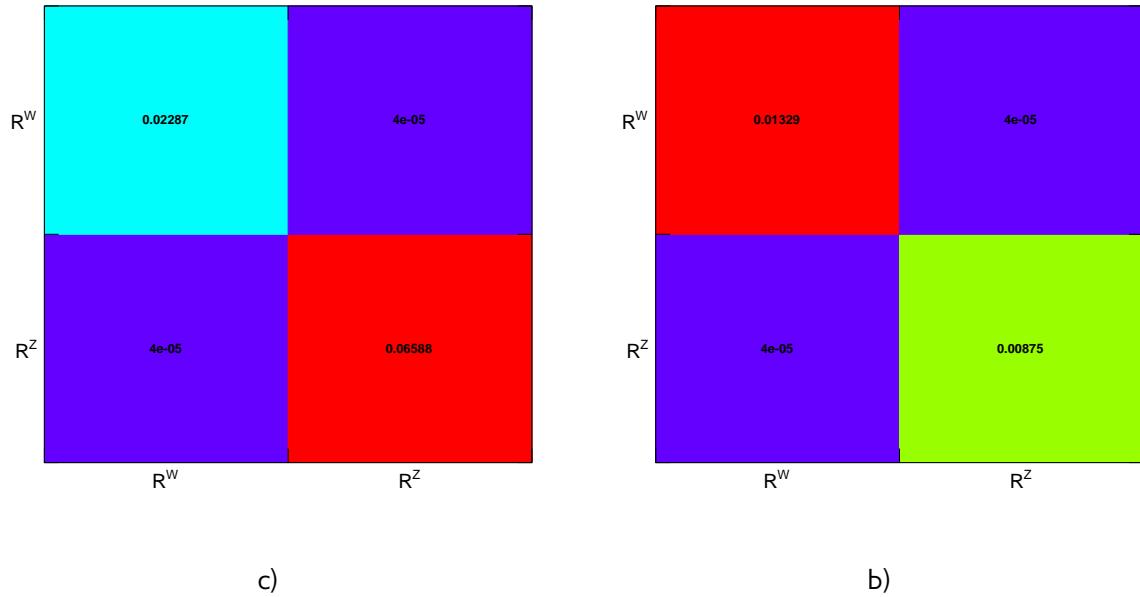


Fig. B.4: Covariance matrix for the measurements of R_W and R_Z ratios from Sec. 15.1.1 for electron channel in fiducial region a)for all uncertainty b)for all but statistical uncertainty.

Appendix C

Additional comparisons with theoretical predictions

The cross-section results for combined channel in fiducial regions from have been compared to NNLO predictions for different PDF sets in Sec. 15.2. This appendix presents the comparison of cross-section in full and extrapolated to 13 TeV regions with NNLO predictions (Fig. C.1-Fig. C.2). The agreement between predictions and results in full region is worse, than for fiducial and extrapolated regions, however it is still within 2σ of uncertainty.

Additionally, the comparison for NLO predictions in fiducial region is presented in Fig. C.3- C.4.

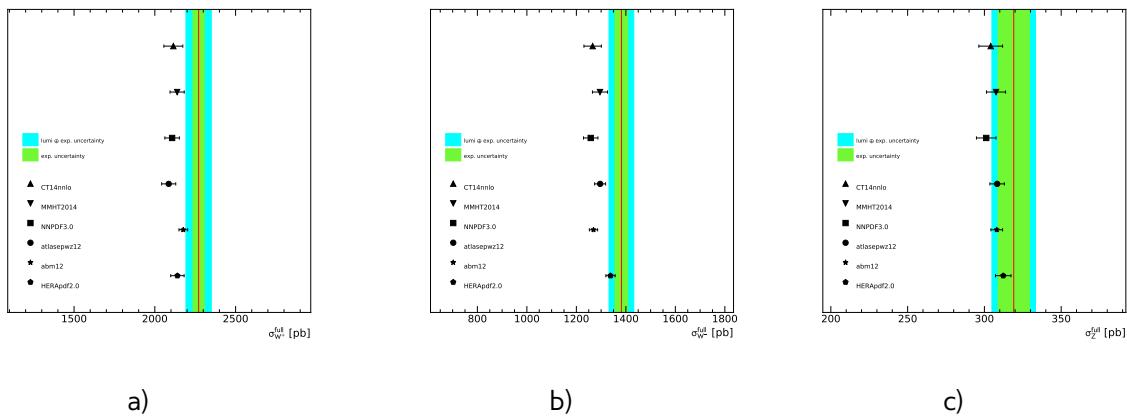


Fig. C.1: The NNLO predictions for the full cross-section a) $\sigma_{W^+}^{full}$ b) $\sigma_{W^-}^{full}$ c) σ_Z^{full} in pb for the six PDFs CT14nnlo, MMHT2014, NNPDF3.0, ATLASepWZ12, abm12, HERApdf2.0 compared to the measured cross-section as given in Tab. 15.2. The green (cyan) band corresponds to the experimental uncertainty without (with) the luminosity uncertainty. The theory predictions are given with the corresponding PDF uncertainties shown as error bands.

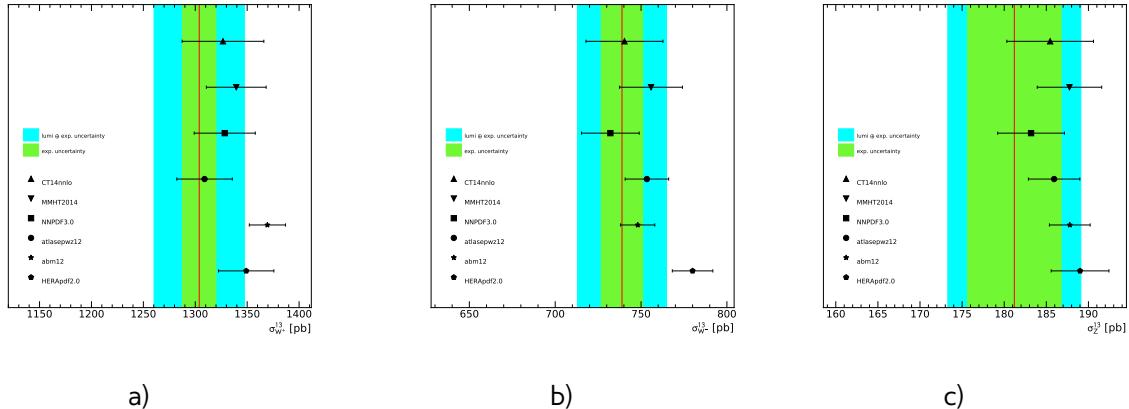


Fig. C.2: The NNLO predictions for the extrapolated to the 13 TeV cross-section a) $\sigma_{W^+}^{13}$ b) $\sigma_{W^-}^{13}$ c) σ_Z^{13} in pb for the six PDFs CT14nnlo, MMHT2014, NNPDF3.0, ATLASepWZ12, abm12, HERA-pdf2.0 compared to the measured cross-section as given in Tab. 15.2. The green (cyan) band corresponds to the experimental uncertainty without (with) the luminosity uncertainty. The theory predictions are given with the corresponding PDF uncertainties shown as error bands.

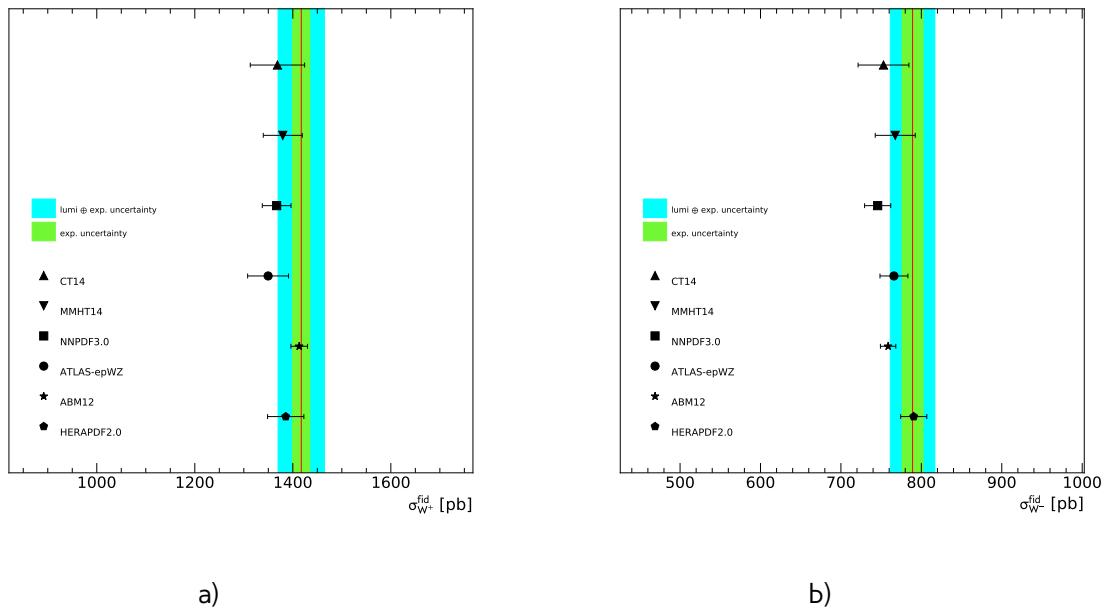


Fig. C.3: The NLO predictions for the fiducial cross-section a) $\sigma_{W^+}^{fid}$ b) $\sigma_{W^-}^{fid}$ in pb for the six PDFs CT14nnlo, MMHT2014, NNPDF3.0, ATLASepWZ12, abm12, HERApdf2.0 compared to the measured fiducial cross-section as given in Tab. 15.2. The green (cyan) band corresponds to the experimental uncertainty without (with) the luminosity uncertainty. The theory predictions are given with the corresponding PDF uncertainties shown as error bands.

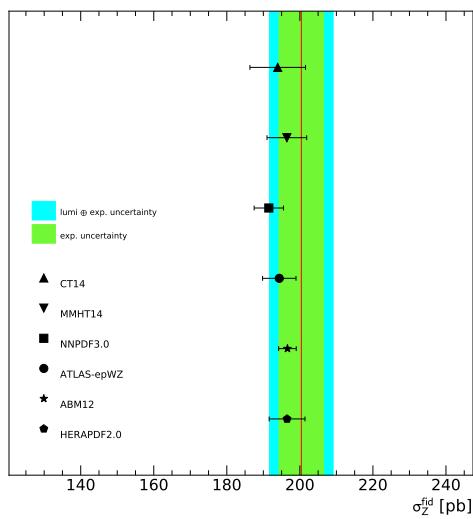


Fig. C.4: The NLO predictions for the fiducial cross-section σ_Z^{fid} in pb for the six PDFs CT14nnlo, MMHT2014, NNPDF3.0, ATLASepWZ12, abm12, HERApdf2.0 compared to the measured fiducial cross-section as given in Tab. 15.2. The green (cyan) band corresponds to the experimental uncertainty without (with) the luminosity uncertainty. The theory predictions are given with the corresponding PDF uncertainties shown as error bands.

Appendix D

1896 Additional PDF profiling plots

1897 The results of PDF profiling have been showed in Sec. 15.3. In this Appendix the effect on valence
 1898 quarks ratio d_v/u_v (Fig. D.1 and difference in sea u and d quarks $\bar{d} - \bar{u}$ (Fig. D.2) is shown. The effect
 1899 of inclusion of the new data at the scale of the measurement $Q^2 \approx M_W^2$ is shown in Fig. D.3- D.5.

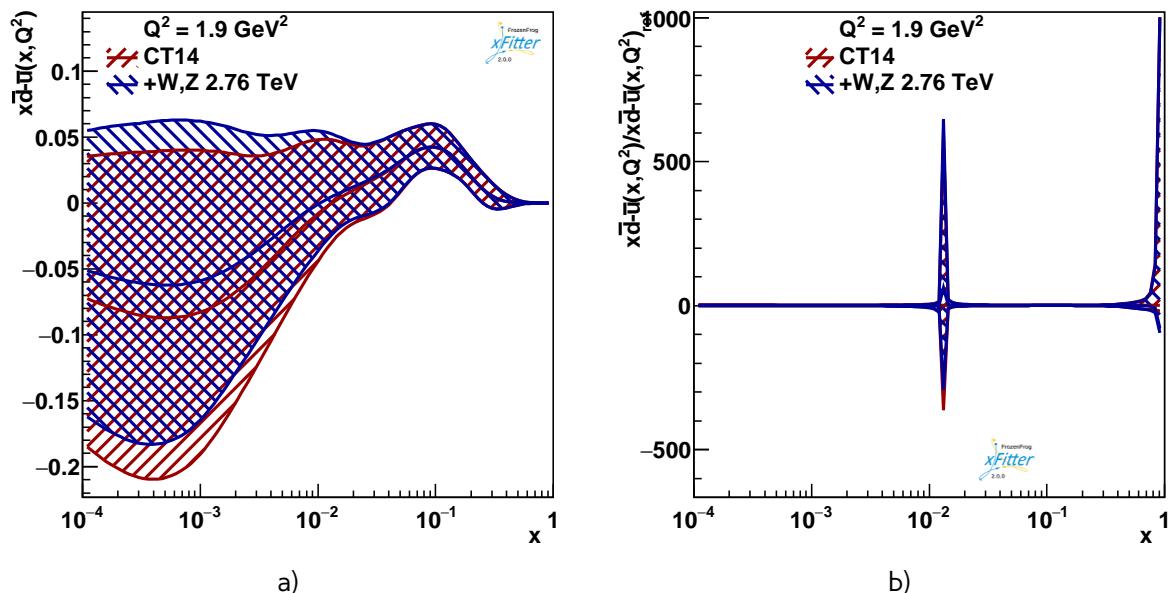


Fig. D.1: The a) absolute and b) relative distributions for the $\bar{d} - \bar{u}$ quark densities as a function of x at scale $Q^2 = 1.9$ GeV 2 with the experimental uncertainties. The red band denotes the reference NLO PDF distributions from CT14 pdf set. The impact of addition of the new W,Z cross-sections at 2.76 TeV on the PDF set is shown by the blue boundaries.

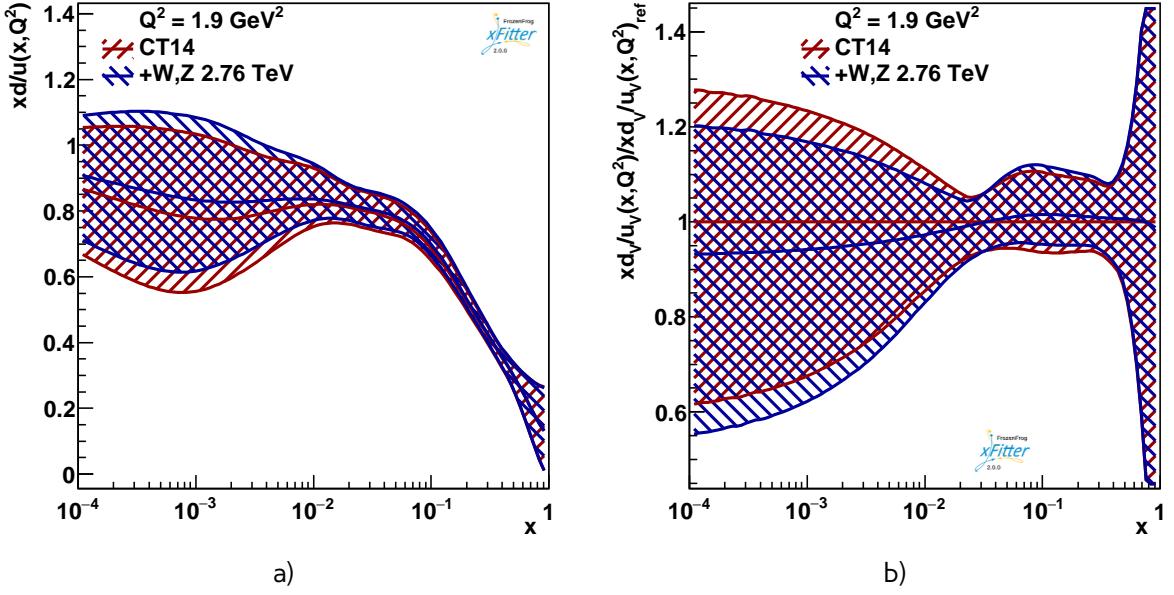


Fig. D.2: The a) absolute and b) relative distributions for the u_v/d_v quark densities as a function of x at scale $Q^2 = 1.9 \text{ GeV}^2$ with the experimental uncertainties. The red band denotes the reference NLO PDF distributions from CT14 pdf set. The impact of addition of the new W,Z cross-sections at 2.76 TeV on the PDF set is shown by the blue boundaries

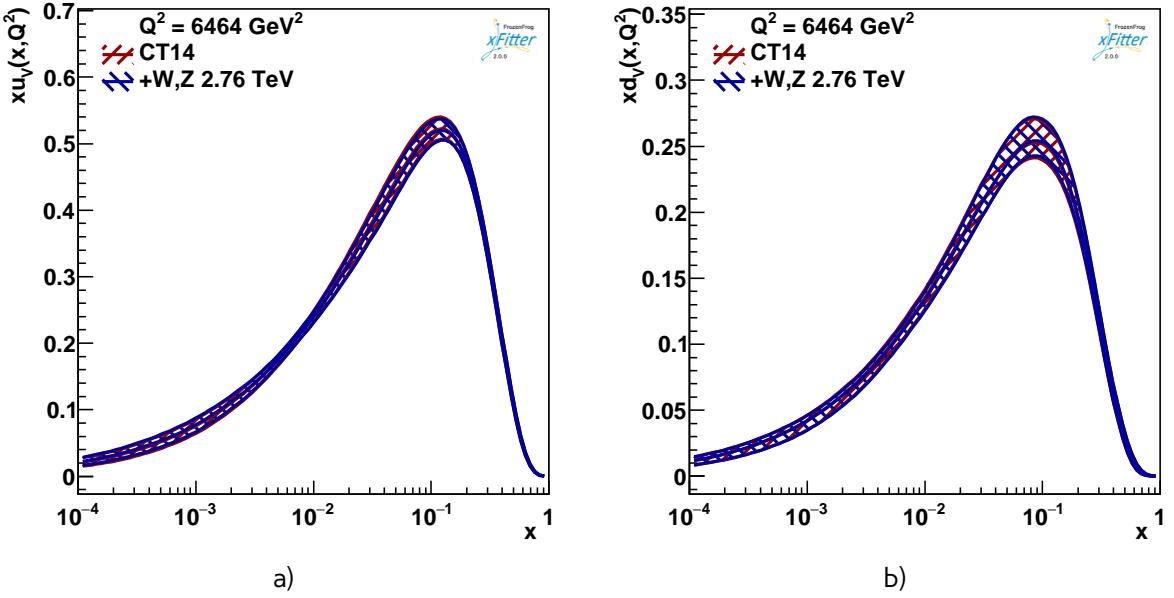
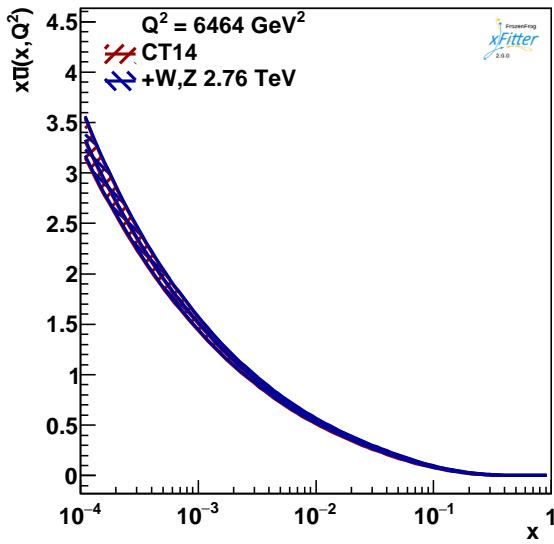
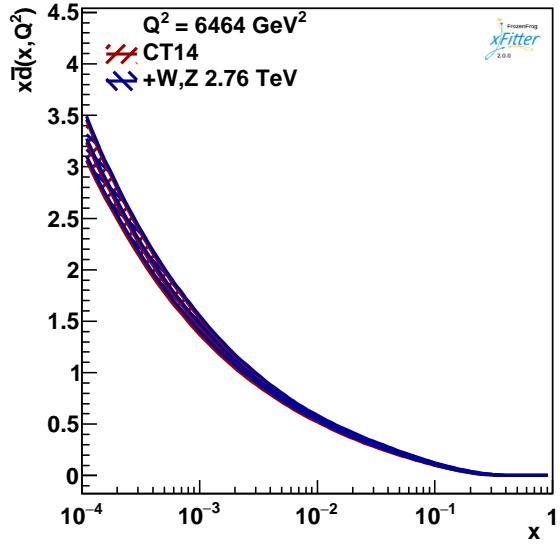


Fig. D.3: The absolute for the a) u_v and b) d_v quark densities as a function of x at scale $Q^2 = M_W^2$ with the experimental uncertainties. The red band denotes the reference NLO PDF distributions from CT14 pdf set. The impact of addition of the new W,Z cross-sections at 2.76 TeV on the PDF set is shown by the blue boundaries.

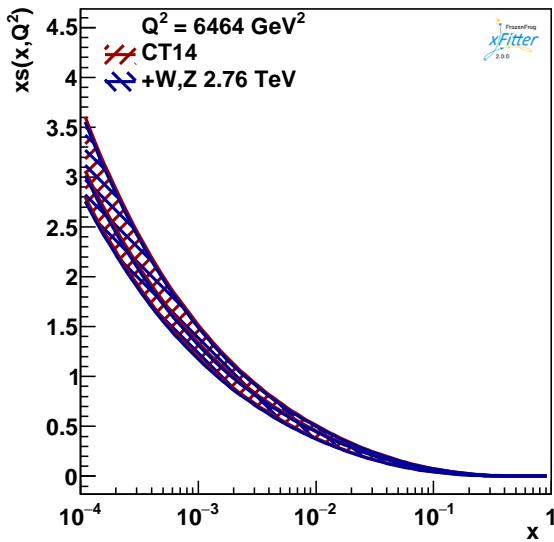


c)

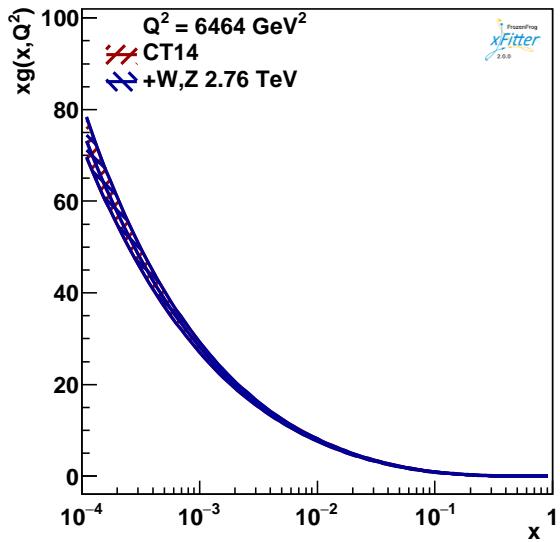


d)

Fig. D.4: The absolute for the a) \bar{u} and b) \bar{d} quark densities as a function of x at scale $Q^2 = M_W^2$ with the experimental uncertainties. The red band denotes the reference NLO PDF distributions from CT14 pdf set. The impact of addition of the new W,Z cross-sections at 2.76 TeV on the PDF set is shown by the blue boundaries.



e)



f)

Fig. D.5: The absolute for the a) s quark and b) gluon densities as a function of x at scale $Q^2 = M_W^2$ with the experimental uncertainties. The red band denotes the reference NLO PDF distributions from CT14 pdf set. The impact of addition of the new W,Z cross-sections at 2.76 TeV on the PDF set is shown by the blue boundaries.

1900 Bibliography

- 1901 [1] J. M. Campbell, J. W. Huston, and W. J. Stirling. Hard interactions of quarks and gluons: a primer
1902 for lhc physics. 2006.
- 1903 [2] L.B. Okun. *Leptons and Quarks*. North-Holland Personal Library. Elsevier Science, 2013.
- 1904 [3] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579–588,
1905 feb 1961.
- 1906 [4] Steven Weinberg. A model of leptons. *Physical Review Letters*, 19(21):1264–1266, nov 1967.
- 1907 [5] Abdus Salam. Renormalizability of gauge theories. *Physical Review*, 127(1):331–334, jul 1962.
- 1908 [6] Standard Model public results.
- 1909 [7] Office of Science Fermilab MissMj. Standard Model of Elementary Particles.
- 1910 [8] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs
1911 boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.
- 1912 [9] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS
1913 experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.
- 1914 [10] E. Noether. Invarianten beliebiger Differentialausdrücke. *Nachr. Ges. Wiss. Göttingen, Math.-Phys.*
1915 *Kl.*, 1918:37–44, 1918.
- 1916 [11] Emmy Noether. Invariante Variationsprobleme. *Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl.*,
1917 1918:235–257, 1918.
- 1918 [12] H. Fritzsch, Murray Gell-Mann, and H. Leutwyler. Advantages of the Color Octet Gluon Picture.
1919 *Phys. Lett.*, B47:365–368, 1973.
- 1920 [13] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Physical*
1921 *Review Letters*, 30(26):1343–1346, jun 1973.
- 1922 [14] H. David Politzer. Reliable perturbative results for strong interactions? *Physical Review Letters*,
1923 30(26):1346–1349, jun 1973.
- 1924 [15] John C. Collins. Renormalization: general theory, 2006.
- 1925 [16] Siegfried Bethke. The 2009 world average of α_s . 2009.
- 1926 [17] Richard P. Feynman. Very high-energy collisions of hadrons. *Physical Review Letters*, 23(24):1415–
1927 1417, dec 1969.
- 1928 [18] V. N. Gribov and L. N. Lipatov. Deep inelastic e p scattering in perturbation theory. *Sov. J. Nucl.*
1929 *Phys.*, 15:438–450, 1972. [Yad. Fiz.15,781(1972)].
- 1930 [19] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne. Parton distributions in the lhc
1931 era: Mmht 2014 pdfs. 2014.

Bibliography

- 1932 [20] Experimental observation of isolated large transverse energy electrons with associated missing
1933 energy at $\sqrt{s} = 540\text{GeV}$. *Phys. Lett. B*, 122(CERN-EP-83-13):103–116. 31 p, Jan 1983.
- 1934 [21] M. Banner et al. Observation of Single Isolated Electrons of High Transverse Momentum in
1935 Events with Missing Transverse Energy at the CERN anti-p p Collider. *Phys. Lett.*, 122B:476–485,
1936 1983.
- 1937 [22] G. Arnison et al. Observation of Muonic Z0 Decay at the anti-p p Collider. *Phys. Lett.*, B147:241–
1938 248, 1984.
- 1939 [23] P. Bagnaia et al. Evidence for $Z0 \rightarrow e^+ e^-$ at the CERN anti-p p Collider. *Phys. Lett.*, B129:130–140,
1940 1983.
- 1941 [24] Richard Keith Ellis, William James Stirling, and Bryan R Webber. *QCD and Collider Physics*. Cam-
1942 bridge monographs on particle physics, nuclear physics, and cosmology. Cambridge Univ. Press,
1943 Cambridge, 2003. Photography by S. Vascotto.
- 1944 [25] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C40:100001, 2016.
- 1945 [26] Ryan Gavin, Ye Li, Frank Petriello, and Seth Quackenbush. Fewz 2.0: A code for hadronic z
1946 production at next-to-next-to-leading order. 2010.
- 1947 [27] Sayipjamal Dulat, Tie Jiun Hou, Jun Gao, Marco Guzzi, Joey Huston, Pavel Nadolsky, Jon Pumplin,
1948 Carl Schmidt, Daniel Stump, and C. P. Yuan. New parton distribution functions from a global
1949 analysis of quantum chromodynamics. 2015.
- 1950 [28] ATLAS Collaboration. Measurement of w^+ , w^- and z/γ^* production cross sections in pp colli-
1951 sions at $\sqrt{s}=13$ tev with the ATLAS detector. *Physics Letters B*, 759:601–621, aug 2016.
- 1952 [29] A. Glazov. Averaging of DIS cross section data. *AIP Conf. Proc.*, 792:237–240, 2005. [,237(2005)].
- 1953 [30] S. Alekhin et al. Herafitter, open source qcd fit project, 2014.
- 1954 [31] Hannu Paukkunen and Pia Zurita. Pdf reweighting in the hessian matrix approach. 2014.
- 1955 [32] Lyndon Evans and Philip Bryant. LHC Machine. *JINST*, 3:S08001, 2008.
- 1956 [33] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008.
- 1957 [34] James T. Shank. The ATLAS Detector: Status and Results from Cosmic Rays. In *Particles and*
1958 *fields. Proceedings, Meeting of the Division of the American Physical Society, DPF 2009, Detroit, USA,*
1959 *July 26-31, 2009*, 2009.
- 1960 [35] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.
- 1961 [36] A. Augusto Alves, Jr. et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.
- 1962 [37] J.J. Goodson. *Search for Supersymmetry in States with Large Missing Transverse Momentum and*
1963 *Three Leptons including a Z-Boson*. PhD thesis, Stony Brook University, May 2012. Presented 17
1964 Apr 2012.
- 1965 [38] ATLAS collaboration. The ATLAS experiment at the CERN Large Hadron Collider. *JINST*, 3, 2008.
- 1966 [39] S. van der Meer. Calibration of the Effective Beam Height in the ISR. 1968.

- [40] T Cornelissen, M Elsing, S Fleischmann, W Liebig, E Moyse, and A Salzburger. Concepts, Design and Implementation of the ATLAS New Tracking (NEWT). Technical Report ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar 2007.
- [41] R. Fruhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Meth.*, A262:444–450, 1987.
- [42] W Lampl, S Laplace, D Lelas, P Loch, H Ma, S Menke, S Rajagopalan, D Rousseau, S Snyder, and G Unal. Calorimeter Clustering Algorithms: Description and Performance. Technical Report ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008.
- [43] W Lampl, S Laplace, D Lelas, P Loch, H Ma, S Menke, S Rajagopalan, D Rousseau, S Snyder, and G Unal. Calorimeter Clustering Algorithms: Description and Performance. Technical Report ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008.
- [44] G. Aad et al. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics. 2009.
- [45] Th Lagouri et al. A Muon Identification and Combined Reconstruction Procedure for the ATLAS Detector at the LHC at CERN. *IEEE Trans. Nucl. Sci.*, 51:3030–3033, 2004.
- [46] S. Hassani, L. Chevalier, E. Lancon, J. F. Laporte, R. Nicolaïdou, and A. Ouraou. A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC using the (MUONBOY, STACO, MuTag) reconstruction packages. *Nucl. Instrum. Meth.*, A572:77–79, 2007.
- [47] ATLAS collaboration. Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s}=7$ tev with ATLAS. *The European Physical Journal C*, 72(1), jan 2012.
- [48] ATLAS collaboration. Pile-up Suppression in Missing Transverse Momentum Reconstruction in the ATLAS Experiment in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. Technical Report ATLAS-CONF-2014-019, CERN, Geneva, May 2014.
- [49] Matthias Schott, Giovanni Siragusa, Jakub Cuth, and Tai-hua Lin. Measurement of m_W at 7 TeV: Reconstruction of the hadronic recoil. Technical Report ATL-COM-PHYS-2014-1435, CERN, Geneva, Nov 2014.
- [50] T H V Nguyen, N Besson, and M Boonekamp. ETmiss from the reconstruction and calibration of the hadronic recoil. Technical Report ATL-PHYS-INT-2010-039, CERN, Geneva, Apr 2010.
- [51] P Calafiura, W Lavrijsen, C Leggett, M Marino, and D Quarrie. The Athena Control Framework in Production, New Developments and Lessons Learned. 2005.
- [52] Montecarlogenerators – sherpa – hepforge. <https://sherpa.hepforge.org/trac/wiki/MonteCarloGenerators>.
- [53] S.D. Drell and Tung-Mow Yan. Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies. *Phys.Rev.Lett.*, 25:316–320, 1970.
- [54] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 0411:040, 2004.
- [55] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. NLO vector-boson production matched with shower in POWHEG. *JHEP*, 0807:060, 2008.

Bibliography

- 2005 [56] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006.
- 2007 [57] Bo Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand. Parton Fragmentation and String Dynamics. *Phys.Rept.*, 97:31–145, 1983.
- 2009 [58] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M.H. Seymour, and B.R. Webber. HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes). *JHEP*, 01:010, 2001.
- 2012 [59] T. Gleisberg, S. Hoeche, F. Krauss, M. Schoenherr, S. Schumann, F. Siegert, and J. Winter. Event generation with sherpa 1.1. *JHEP* 0902:007, 2009, 2008.
- 2014 [60] Elisabetta Barberio and Zbigniew Was. Photos - a universal monte carlo for qed radiative corrections: version 2.0. *Computer Physics Communications*, 79(2):291 – 308, 1994.
- 2016 [61] Zbigniew Was. TAUOLA for simulation of tau decay and production: perspectives for precision low energy and LHC applications. *Nucl.Phys.Proc.Suppl.*, 218:249–255, 2011.
- 2018 [62] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl.Instrum.Meth.*, A506:250–303, 2003.
- 2019 [63] E. Barberio et al. Fast simulation of electromagnetic showers in the ATLAS calorimeter: Frozen showers. *J. Phys. Conf. Ser.*, 160:012082, 2009.
- 2021 [64] Tom M. Mitchell. *Machine Learning*. McGraw-Hill Education, 1997.
- 2022 [65] C. Van Der Malsburg. Frank rosenblatt: Principles of neurodynamics: Perceptrons and the theory of brain mechanisms. In *Brain Theory*, pages 245–248. Springer Berlin Heidelberg, 1986.
- 2024 [66] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- 2028 [67] Leo Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. *Classification and Regression Trees*. Statistics/Probability Series. Wadsworth Publishing Company, Belmont, California, U.S.A., 1984.
- 2030 [68] Sriram Vajapeyam. Understanding shannon’s entropy metric for information, 2014.
- 2031 [69] V. Vapnik and A. Lerner. Pattern Recognition using Generalized Portrait Method. *Automation and Remote Control*, 24, 1963.
- 2033 [70] Validation of the frozen showers libraries on forward electrons results. http://atlas-computing.web.cern.ch/atlas-computing/links/PhysValDir/Egamma/electrons/10_02_2017/task1/.
- 2036 [71] Validation of the frozen showers libraries on zee sample with 1 electron in a forward region. http://atlas-computing.web.cern.ch/atlas-computing/links/PhysValDir/SMZee/01-12-2016/task1_norm/Zee/FWDZee/index.html.
- 2039 [72] Validation of the frozen showers libraries on jets sample. https://atlas-computing.web.cern.ch/atlas-computing/links/PhysValDir/JetEtMiss/jet_17-02-10_task1/JetsETmiss/Jets/AntiKt10LCTopoJets/index.html.

- 2042 [73] ATLAS Collaboration. Measurements of the nuclear modification factor for jets in pb+pb collisions at $\sqrt{s_{\text{NN}}} = 2.76$ tev with the atlas detector. 2014.
- 2043
- 2044 [74] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C.-P. Yuan. New parton distributions for collider physics. *Phys. Rev. D*, 82:074024, Oct 2010.
- 2045
- 2046 [75] Summary of ATLAS Pythia 8 tunes. Technical Report ATL-PHYS-PUB-2012-003, CERN, Geneva, Aug 2012.
- 2047
- 2048 [76] Jonathan Pumplin, Daniel Robert Stump, Joey Huston, Hung-Liang Lai, Pavel Nadolsky, and Wu-Ki Tung. New generation of parton distributions with uncertainties from global QCD analysis. *Journal of High Energy Physics*, 2002(07):012–012, jul 2002.
- 2049
- 2050
- 2051 [77] First tuning of HERWIG/JIMMY to ATLAS data. Technical Report ATL-PHYS-PUB-2010-014, CERN, Geneva, Oct 2010.
- 2052
- 2053 [78] Sarah Heim and Kristin Lohwasser. Status of electron id and other recommendations for run-2. <https://indico.cern.ch/event/363200/session/6/contribution/46/material/slides/0.pdf>.
- 2054
- 2055
- 2056 [79] ATLAS Collaboration. Electron and photon energy calibration with the atlas detector using lhc run 1 data. 2014.
- 2057
- 2058 [80] ATLAS Collaboration. Measurement of the muon reconstruction performance of the atlas detector using 2011 and 2012 lhc proton-proton collision data. 2014.
- 2059
- 2060 [81] Aleksandra Dimitrijevska, Nenad Vranjes, Matthias Schott, and Maarten Boonekamp. Measurement of m_W at 7 TeV: Hadronic recoil corrections. Technical Report ATL-COM-PHYS-2015-344, CERN, Geneva, Apr 2015.
- 2061
- 2062
- 2063 [82] Kendall, Maurice, and Alan Stuart. *The advanced theory of statistics*. London: Griffin, 1969.
- 2064
- 2065 [83] Measurement of the $w \rightarrow l\nu$ and $z/\gamma \rightarrow ll$ production cross sections in proton-proton collisions at $\sqrt{s} = 7$ tev with the ATLAS detector. *Journal of High Energy Physics*, 2010(12), dec 2010.
- 2066
- 2067 [84] Paolo Saracco, Maria Grazia Pia, and Matej Batic. Theoretical grounds for the propagation of uncertainties in monte carlo particle transport. 2014.
- 2068
- 2069 [85] ATLAS Collaboration. Electron performance measurements with the atlas detector using the 2010 lhc proton-proton collision data. *Eur. Phys. J. C72 (2012) 1909*, 2011.
- 2070
- 2071 [86] A. e/gamma Combined Performance Group. Energy scale and resolution recommendations. <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EGammaCalibrationGE020>.
- 2072
- 2073 [87] M. R. Whalley, D. Bourilkov, and R. C. Group. The les houches accord pdfs (lhapdf) and lhaglue, 2005.
- 2074
- 2075 [88] Determination of the strange-quark density of the proton from ATLAS measurements of the $w \rightarrow e\nu$ and $w \rightarrow ee$ cross sections. *Physical Review Letters*, 109(1), jul 2012.
- 2076
- 2077 [89] S. Alekhin, J. Blümlein, S. Klein, and S. Moch. The 3-, 4-, and 5-flavor nnlo parton from deep-inelastic-scattering data and at hadron colliders. 2009.

Bibliography

- 2078 [90] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Christopher S. Deans, Luigi Del Debbio, Ste-
2079 fano Forte, Alberto Guffanti, Nathan P. Hartland, Jose I. Latorre, Juan Rojo, and Maria Ubiali.
2080 Parton distributions with lhc data. 2012.
- 2081 [91] Bradley Dickinson. Estimation of partial correlation matrices using cholesky decomposition. In
2082 *1978 IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes*.
2083 IEEE, jan 1978.
- 2084 [92] John M. Campbell and R.K. Ellis. MCFM for the Tevatron and the LHC. *Nucl.Phys.Proc.Suppl.*,
2085 205-206:10–15, 2010.
- 2086 [93] Tancredi Carli, Dan Clements, Amanda Cooper-Sarkar, Claire Gwenlan, Gavin P. Salam, Frank
2087 Siegert, Pavel Starovoitov, and Mark Sutton. A posteriori inclusion of parton density functions
2088 in nlo qcd final-state calculations at hadron colliders: The applgrid project. 2009.
- 2089 [94] S. Alekhin, J. Bluemlein, and S. Moch. The abm parton distributions tuned to lhc data. 2013.
- 2090 [95] H1 and ZEUS Collaborations. Combination of measurements of inclusive deep inelastic $e^\pm p$
2091 scattering cross sections and qcd analysis of hera data, 2015.