

Contents

I Thesis overview	1
II Theory introduction	3
1 Theoretical introduction	5
1.1 Standard model	5
1.2 Theory of pp collisions	5
1.3 Proton structure	5
2 Physics of W and Z bosons in pp collisions	7
3 Methodology of Cross-Section Measurement	9
3.1 Fiducial volume definition	9
3.2 Results combination	10
3.3 PDF fits	10
III Experimental setup	11
4 The Large Hadron Collider	13
5 The ATLAS experiment	15
5.1 Inner detector	15
5.2 Calorimeters	15
5.2.1 Forward calorimeters	15
5.3 Muon system	15
5.4 The magnet system	15
5.5 Trigger system	15
5.6 Atlas software	15
6 Event reconstruction	17
6.1 Tracks and vertexes	17
6.2 Electron reconstruction and identification	17
6.3 Muon reconstruction and identification	17
6.4 Missing transverse energy reconstruction	17
6.4.1 Standard Missing Transverse Energy reconstruction	17
6.4.2 Reconstruction of Missing Transverse Energy from hadronic recoil	18
7 Monte Carlo simulation	21
7.1 ATLAS chain of MC production	21
7.2 Event generators	22

7.3	Simulation in Geant4	23
8	Frozen Showers	25
8.1	Problem description	27
8.2	Generation and use in simulation	30
8.2.1	Libraries tuning	31
8.3	Machine learning based bin finding procedure	32
8.3.1	Machine learning introduction	32
8.3.2	Model description	34
8.3.3	Interpretation of results	34
8.4	Validation of the new libraries	37
8.5	Plans for a future developments	37
9	Data and Monte-Carlo samples	39
9.1	Data sample	39
9.2	Monte-Carlo samples	39
IV	The Measurement	41
10	Selection	43
10.1	Data quality cuts	43
10.2	Lepton quality cuts	44
10.3	Boson selection	44
10.4	Cut flow	44
11	Monte Carlo corrections	45
11.1	Lepton efficiency corrections	45
11.1.1	Muon Trigger SF	47
11.2	Electron energy scale and resolution	49
11.3	Muon momentum correction	50
12	Hadronic recoil calibration	53
12.1	Introduction	53
12.2	Hadronic recoil resolution correction	55
12.2.1	$\sum E_T$ distribution correction	56
12.2.2	Resolution corrections using Z events	61
12.3	Hadronic recoil bias correction	63
12.3.1	Bias determination from M_T^W distribution	63
12.3.2	Bias determination using u_{\parallel} distribution	64
12.3.3	Sytematic uncertainty estimation	67
12.4	Summary on hadronic recoil systematics	67
13	Background estimation	69
13.1	QCD background estimation	70
13.1.1	Template selection	71
13.1.2	Methodology of the template sample normalization	71
13.1.3	Systematic Uncertainty from the Multi-jet Background Estimation	73

13.2 Background-subtracted W and Z candidate events	75
14 Control distributions	77
15 Uncertainties in the cross-section measurement	87
15.1 Methods of uncertainties propagation	87
15.2 Experimental systematic uncertainties	88
15.2.1 Electron energy scale and resolution	88
15.2.2 Muon energy scale and resolution	88
15.2.3 Muon and electron efficiency toy Monte-Carlo	88
15.3 Theoretical uncertainty	90
15.4 Correlation between uncertainties	91
15.4.1 Toy MC correlations	91
15.4.2 Correlations between PDF's eigenvectors	91
V Results	93
16 Results of the Cross Section Measurement	95
16.1 Combined results	95
16.2 Comparation with Theoretical Predictions	95
17 PDF fits results	97
18 Summary	99

1

Part I

2

Thesis overview

3

Part II

4

Theory introduction

3

⁵ Chapter **1**

⁶ **Theoretical introduction**

⁷ **1.1 Standard model**

⁸ **1.2 Theory of pp collisions**

⁹ **1.3 Proton structure**

10 Chapter **2**

11 **Physics of W and Z bosons in pp collisions**

12 Chapter **3**

13 **Methodology of Cross-Section Measurement**

The production of W and Z bosons is measured in a fiducial volume using the equation:

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

14 where

- 15 • N is the number of candidates measured in a data
16 • B is the number of background events
17 • L_{int} is the integrated luminosity corresponding to a run selections and trigger requirements
18 • Efficiency correction factor C_W .

The efficiency scale factor C_W is calculated from MC and defined as :

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

19 where $N_{MC,rec}$ are sums of weights of events after simulation, reconstruction and selection, $N_{MC,gen,cut}$
20 - are taken on the generator level after fiducial cuts.

The total W and Z cross-sections are calculated as:

$$\sigma_{tot} = \sigma_{W/Z} \cdot BR(W/Z \rightarrow l\nu/l\bar{l}) = \frac{\sigma_{fid}}{A_{W/Z}} = \frac{N - B}{A_{W/Z} C_{W/Z} L_{int}}, \quad (3.3)$$

where the acceptance $A_{W/Z}$ is used to extrapolate the cross-section measurement in the fiducial volume to the full kinematic region. It is derived from MC and affected only by the theoretical uncertainties. It is obtained as follows:

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

21 where $N_{MC,gen,all}$ are the sum of weights of all generated MC events

22 **3.1 Fiducial volume definition**

In order to combine results fiducial region is defined commonly for a different flavors of the analysis:

$$W \rightarrow l\nu : p_T^l > 20 GeV, |\eta^l| < 2.5, p_T^\nu < 20 GeV, M_T^W > 40 GeV \quad (3.5)$$

$$Z \rightarrow ll : p_T^l > 20\text{GeV}, |\eta^l| < 2.5, 66\text{GeV} < m_{ll} < 116\text{GeV} \quad (3.6)$$

²³ Additionally, for a future combination, 13 TeV analysis fiducial phase space have been used:

$$W \rightarrow l\nu : p_T^l > 20\text{GeV}, |\eta^l| < 2.5, p_T^\nu < 20\text{GeV}, M_T^W > 40\text{GeV} \quad (3.7)$$

$$Z \rightarrow ll : p_T^l > 20\text{GeV}, |\eta^l| < 2.5, 66\text{GeV} < m_{ll} < 116\text{GeV} \quad (3.8)$$

²⁴ **3.2 Results combination**

²⁵ **3.3 PDF fits**

26

Part III

27

Experimental setup

²⁸

Chapter

4

²⁹

The Large Hadron Collider

30 Chapter **5**

31 **The ATLAS experiment**

32

33

34 5.1 Inner detector	15
35 5.2 Calorimeters	15
36 5.2.1 Forward calorimeters	15
37 5.3 Muon system	15
38 5.4 The magnet system	15
39 5.5 Trigger system	15
40 5.6 Atlas software	15

41

42

43

44 **5.1 Inner detector**

45 **5.2 Calorimeters**

46 **5.2.1 Forward calorimeters**

47 **5.3 Muon system**

48 **5.4 The magnet system**

49 **5.5 Trigger system**

50 **5.6 Atlas software**

52

Event reconstruction

53	6.1 Tracks and vertexes	17
54	6.2 Electron reconstruction and identification	17
55	6.3 Muon reconstruction and identification	17
56	6.4 Missing transverse energy reconstruction	17
57	6.4.1 Standard Missing Transverse Energy reconstruction	17
58	6.4.2 Reconstruction of Missing Transverse Energy from hadronic recoil . .	18

64

6.1 Tracks and vertexes

65

6.2 Electron reconstruction and identification

66

6.3 Muon reconstruction and identification

67

6.4 Missing transverse energy reconstruction

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

73

6.4.1 Standard Missing Transverse Energy reconstruction

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

$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{miss,calo}} + E_{x(y)}^{\text{miss,cryo}} + E_{x(y)}^{\text{miss,muon}}. \quad (6.1)$$

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

$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{miss},e} + E_{x(y)}^{\text{miss},\gamma} + E_{x(y)}^{\text{miss},\tau} + E_{x(y)}^{\text{miss,jets}} + E_{x(y)}^{\text{miss,SoftTerm}} + E_{x(y)}^{\text{miss},\mu}. \quad (6.2)$$

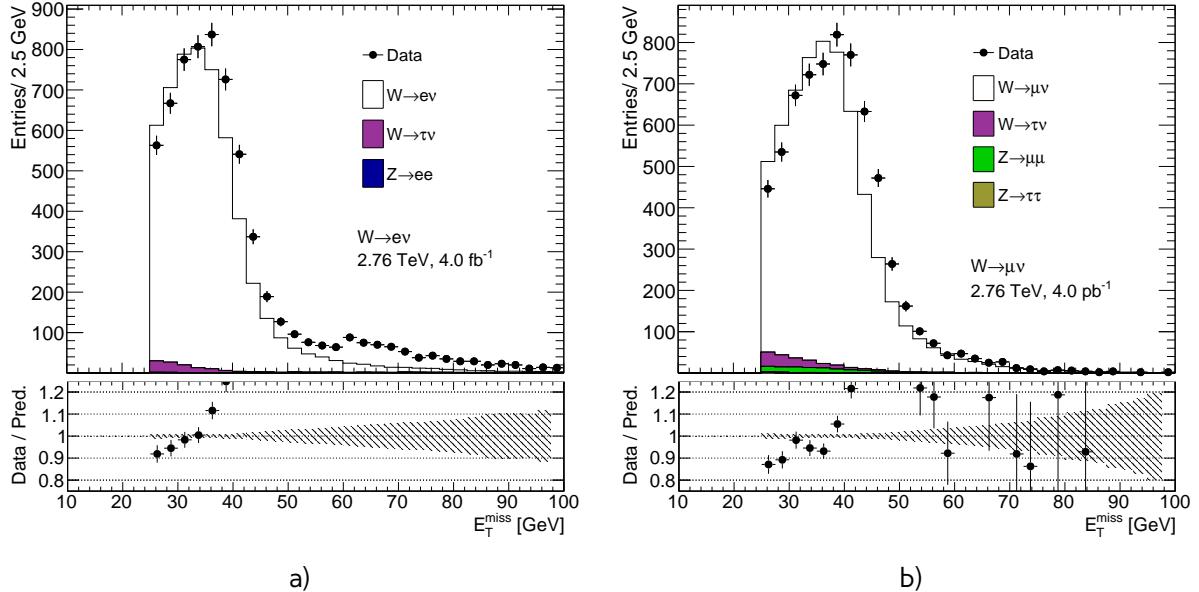


Fig. 6.1: Missing transverse energy distribution for a) the $W \rightarrow ev$ selection and b) the $W \rightarrow \mu\nu$ selection from Chap. 10. E_T^{miss} calculated using the standard ATLAS algorithm. The expected contributions from all backgrounds are estimated with Monte Carlo simulations, except for QCD background that is not included. All Monte-Carlo corrections from Chap. 11 are applied. There are visible discrepancies between data and MC, that cannot be explained by the contribution of QCD background, which is expected mainly in the low E_T^{miss} region (Sec. 13.1).

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. Since pileup has a significant effect on a E_T^{miss} performance several methods of pileup suppression are used [?].

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 figured out, that there are big discrepancies between E_T^{miss} distributions for data and MC simulation, as shown on a Fig. 6.1, 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) contributions from the QCD background, which is expected mainly in the low E_T^{miss} region (see Sec. 13.1).

6.4.2 Reconstruction of Missing Transverse Energy from hadronic recoil

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

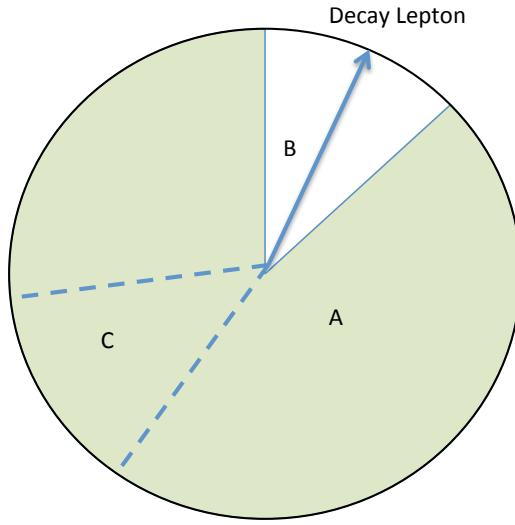


Fig. 6.2: Definition of different zones in the calculation of the cluster-based hadronic recoil. Zone B is excluded from hadron recoil calculation because it contains decay lepton. To describe properly overall acitivity it is replaced by the zone C, rotated in the direction of B. Zone A corresponds to the rest of the calorimeter [?].

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

$$\vec{P}_T^W = \vec{P}_T^l + \vec{P}_T^\nu = \sum \vec{P}_T^{ISRquarks,gluon}, \quad (6.3)$$

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

$$E_T^{miss} = -P_T^\nu = -HR + P_T^l \quad (6.4)$$

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:

$$HR = \sum_{i=0}^{N_{topo}} \vec{p}_T^{topo} \quad (6.5)$$

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

$$\sum E_T = \sum_{i=0}^{N_{topo}} E_T^{topo} \quad (6.6)$$

⁹⁰ To avoid double counting of lepton energy losses in the calorimeter, the clusters inside a cone with
⁹¹ radius $dR = 0.2$ around the lepton direction are excluded from this calculation. To compensate for the
⁹² subtracted soft activity from the cone, a replacement cone is added (Fig. 6.2). This cone is defined

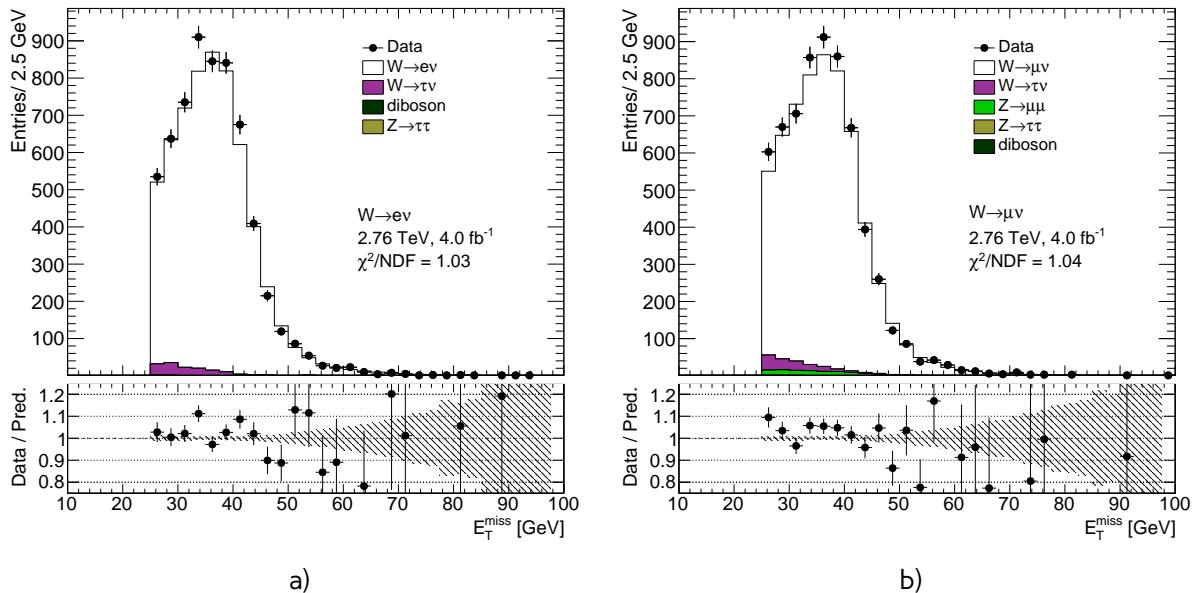


Fig. 6.3: Missing transverse energy distribution for a) the $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection from Chap. 10. E_T^{miss} calculated using the hadronic recoil algorithm. The expected contributions from all backgrounds are estimated with Monte Carlo simulations, except for QCD background that is not included. All Monte-Carlo corrections from Chap. 11 are applied.

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

Fig. 6.3 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 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 analysis.

101

Monte Carlo simulation

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

108

7.1 ATLAS chain of MC production

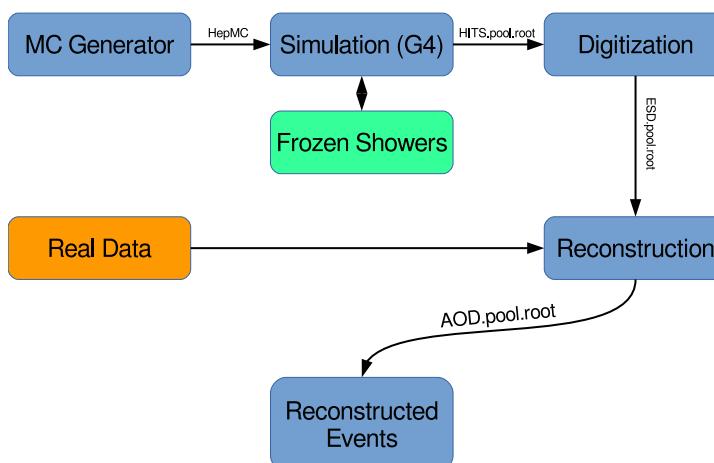


Fig. 7.1: Diagram of the ATLAS MC production chain

109 Monte Carlo method allows to perform different analysis, generate predictions for comparisons
 110 with data, study detector or selection algorithms performance. All of these applications require
 111 accurate MC predictions. Simulation software expects to use precise physics models for sampling
 112 and have large enough statistics, to exclude statistical uncertainties (usually 5 times more, than
 113 expected in a data). ATLAS simulation software is integrated into Athena framework and usually used
 114 during large production of events. Simulation chain is generally divided into 4 main steps (Figure 7.1):

115 **Event generation** Simulation of hard interaction and a resulting high-energy particles parameters.
 116 This step is independent of ATLAS detector geometry.

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

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

121 **Reconstruction** Production of the Analysis Object Data (AOD) files, which are containing sufficient
 122 information for physics analysis. This stage is identical for both data and MC

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

128

7.2 Event generators

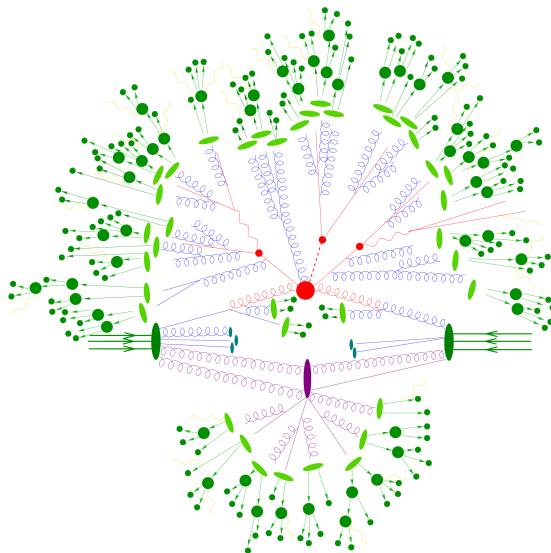


Fig. 7.2: Schematic view of a $t\bar{t}H$ event produced in a pp-collision: the hard scattering is shown as a red blob with the solid and dashed lines as the resulting three particles. Independently happening multi-particle interactions are indicated by the violet blob. Parton showers are shown with curly lines. Hadronization yields hadrons as shown in light green, while the final state particle are dark green.

129 The outcome of the hard interaction could be simple scattering of the hadron elementary con-
 130 stituents, their annihilation into new resonances or a combination of two. This can lead to a final
 131 state with a large particles multiplicity. The main goal of event generator is to provide a complete
 132 picture of this final states: description of the particle types and momentia on event-by-event ba-
 133 sis. The factorisation theorem [?] allows to make event generation in independent stages, which are
 134 dominated by different dynamics. Schematic plan of simulation of ttbar event is shown in Figure ??:

135 **Modelling of hard subprocess** Hard subprocess happens at the smallest times and distances,
 136 where the colliding partons are considered free. Process of interest is simulated by selecting

137 production channels and calculating corresponding matrix elements (ME) in the desired level of
 138 accuracy in perturbation theory . Most of the generators have leading order or next to leading
 139 order ME in α_s .

140 **Parton showering** Quarks and gluons from hard process can radiate secondary quarks and gluons,
 141 resulting in the dozens of additional partons associated with the event. This process is
 142 calculated as step-by-step evolution of momentum transfer scales from highest (hard subprocess),
 143 to the lowest (around 1 GeV). There is a possibility of double counting between showers
 144 and hard subprocess. This can be avoided using matching approaches, for which higher order
 145 corrections to ME are integrated with parton showers, or merging strategy, where jet resolution
 146 scale is used as a threshold between matrix elements and parton showers.

147 **Hadronisation** Final stable color-neutral particles, what can be detected in experiment, are formed
 148 during hadronisation. This occurs at larger nonperturbative scales and usually implemented
 149 using different phenomenological models.

150 **Modelling underlying event** Parallel to the main process other collisions of partons can occur,
 151 called underlying event. These additional interactions can produce partons which contribute to
 152 the final state. This is one of the least understood aspects of hadronic collisions.

153 The current analysis uses samples generated with the following generators:
 154 Powheg [?] Powheg is a generator with NLO ME [?], that can be interfaced with other generators (such
 155 as Pythia or Herwig) for higher precision of showering.
 156 Pythia [?] Pythia is a general purpose generator for hadronic, hadron-lepton and leptonic collisions.
 157 It can model initial and final state showers, hadronisation and decays, underlying event (via
 158 multi-parton interactions). Pythia contains a library with around 240 processes with LO ME. It
 159 uses Lund String model [?] for hadronisation.

160 Herwig [?] Herwig is a LO general purpose event generator for simulation lepton-lepton, hadron-
 161 lepton and hadron-hadron collisions. The main difference between Pythia and Herwig is that
 162 it uses angular ordering in the parton showers and also models the hadronisation step based
 163 on the cluster fragmentation

164 Sherpa [?] Sherpa is an event generator, that uses tree-level leading order matrix element for a hard
 165 scattering and featuring its own implementation of parton shower and hadronisation models.

166 Photos [?] Program used for generation of QED radiative corrections in W and Z decays.

167 Tauola [?] Generator, used to describe leptonic and semi-leptonic τ -decays.

168 7.3 Simulation in Geant4

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

173 Geant4 can simulate complex detector structures with sensitive detector material and corresponding
 174 infrastructure. It can also calculate basic properties of materials, like radiation and interaction

length. For detector Geant4 stores "hits" information - snapshots of physical interactions. In Geant4 events and particles are simulated separately and each particle is moved in steps. Size of each step is chosen to preserve both CPU performance and required precision.

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

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

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

There are two main methods used at ATLAS:

- Parametrisation of calorimeter cells response. Spacial energy response is simulated using longitudinal and lateral energy profiles.
- Frozen Showers. This technique will be described more detailed in Chap. 8

197

Frozen Showers

198	8.1	Problem description	27
199	8.2	Generation and use in simulation	30
200	8.2.1	Libraries tuning	31
201	8.3	Machine learning based bin finding procedure	32
202	8.3.1	Machine learning introduction	32
203	8.3.2	Model description	34
204	8.3.3	Interpretation of results	34
205	8.4	Validation of the new libraries	37
206	8.5	Plans for a future developments	37

211 As it was mentioned in a previous chapter, fast simulation techniques are the essential part of
 212 Monte-Carlo production at ATLAS experiment. Typical time for a simulation of $t\bar{t}$ event is around 1
 213 minute, and most of the time is spend on a simulation of particle interaction in calorimeters. This
 214 motivates a development of fast calorimetry techniques.

215 Frozen showers is currently the main fast calorimeter simulation approach used at ATLAS exper-
 216 iment. In this chapter we will discuss main principles,difficulties and current developments in this
 217 method.

218 Frozen shower method uses pre-simulated "frozen" showers instead of the full simulation. This
 219 is allowing to reduce time spend on a simulation of a large amount of low energy sub showers.
 220 This method is allowing to have a 25% speedup in simulation. It is required to have an in advance
 221 generation of a libraries for each detector and particle used in this method. Each shower from a
 222 library stores it's lateral and transverse size and list of the all energy depositions inside sensitive
 223 material (hits) with information about their energy, position and time. During a simulation, if energy
 224 of secondary electron falls below cutoff energy it is replaced by a shower from a library, as shown
 225 on a Fig. 8.1. Main parameters used in ATLAS simulation are summarized in a Tab. 8.1.

Table 8.1: 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 range	$E_\gamma < 10 \text{ MeV}$, $E_e < 1000 \text{ MeV}$, $T_n < 100 \text{ MeV}$

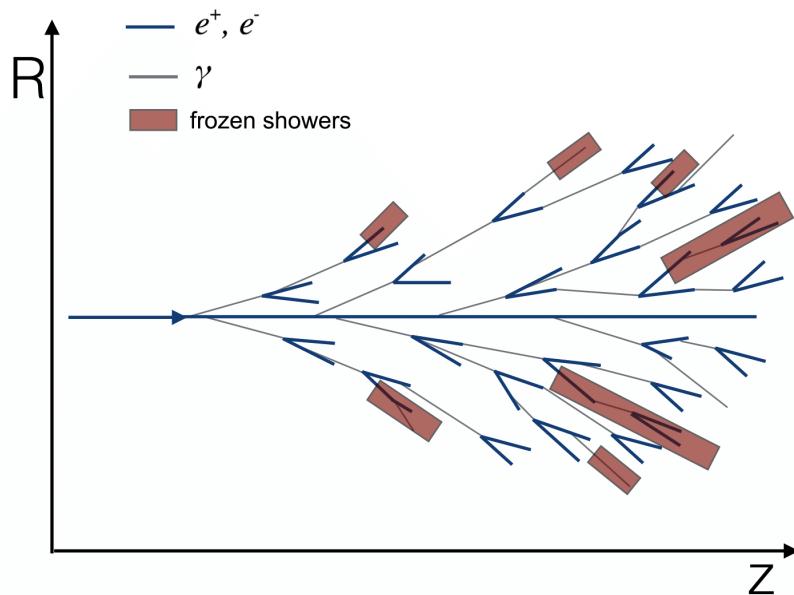


Fig. 8.1: Diagram showing the shower substitution of the low-energy particle, during the high-energy particle simulation. Some of the showers from a particles, substituted by frozen showers method marked by a red squares

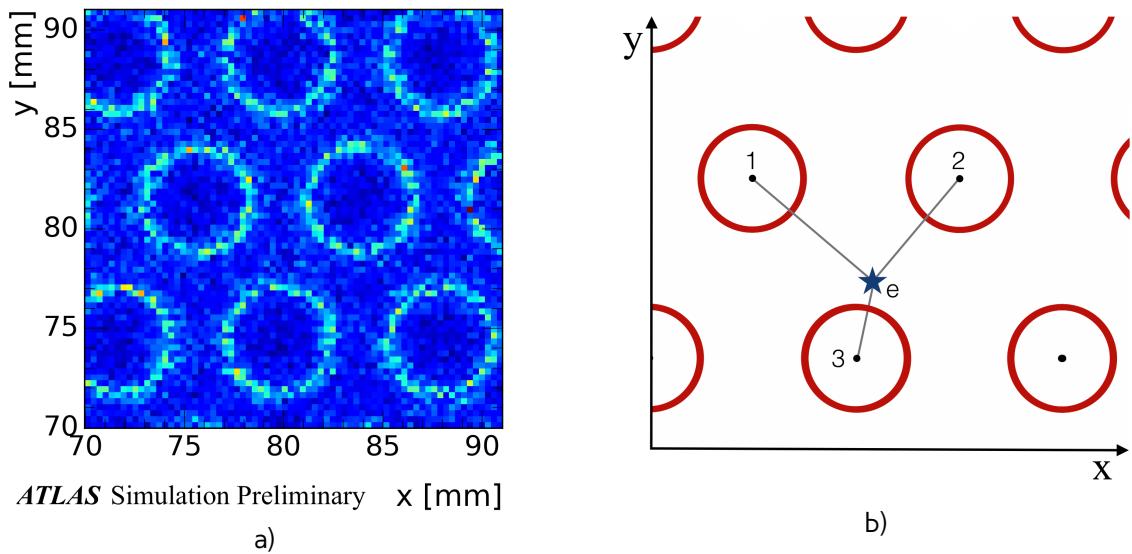


Fig. 8.2: a) Shower energy response histogram in the x vs y plane for electrons, generated with uniformly distributed x and y and energy less than 1 GeV. Light circles are corresponding to a showers, started inside a LAr gaps with on average higher energy response, while the dark parts are corresponding to dead material respectively with smaller sum of the "hits" energy. b) Distance to a closest rod center scheme, where $d_{\text{rod}} = \min(d(1,e), d(2,e), d(3,e))$. Rod centers and liquid argon gaps are shown by black and red circles respectively.

227 8.1 Problem description

228 Fast simulation of forward calorimeters (FCAL) is a complicated task due to its complex structure.
 229 As it was mentioned in a Sec. 5.2.1 FCAL consists of hexagonal absorber cells with anode tube and
 230 cathod rod in the cell center and liquid argon in the gap between rod and tube. In order to simulate
 231 resolution of high energetic electrons, good fast simulation technique should take this feature of
 232 large amount of non-uniformly distributed sensitive material.

Resolution of electron inside calorimeter can be written as:

$$\frac{\sigma}{E} \approx \frac{1}{\sqrt{E}} \oplus \frac{1}{E} \oplus const, \quad (8.1)$$

233 where symbol \oplus indicates a quadratic sum. The first term is 'stochastic term', which includes intrinsic shower fluctuations, second takes into account readout noise effects and pile-up fluctuations.
 234 Constant term derives from non-uniformities in a detector, causing large fluctuation of the energy
 235 loss. Resolution of high-energy electrons is mostly dominated by the constant term.

Fluctuations due to a detector design are visible in a simulation of small energy electrons, generated inside a different points in forward calorimeter. Shower energy E^{shower} distribution in the x vs y plane is showed on Fig. 8.2, where shower energy is defined as:

$$E^{shower} = \sum E_i^{hits}, \quad (8.2)$$

237 where E_i^{hits} is an energy of i-th shower deposit inside sensitive material. Periodic structure resembles
 238 the calorimeter design, where light circles are corresponding to gaps with liquid argon. It could be
 239 reduced to a 1-d problem by introducing d_{rod} distance to a closest rod center.

240 Typical electron substituted by a frozen showers coming from a simulation of high energy electrons
 241 have a relatively small energy (Fig. 8.3 a)). Mean number of depositions in a sensitive material in a
 242 "frozen" shower is around 5 and this value rises with electron energy (Fig. 8.3 b). Fig. 8.4 presents a
 243 distribution showers for electrons with energy below 1 GeV coming from initial electrons with energy
 244 1 TeV in the distance to a closest rod center vs shower energy plane. Liquid argon gap is marked by
 245 a red lines. There is a clear difference in a showers energies between electrons born in a sensitive
 246 and dead material. Difference in a shower properties are also visible for number of hits (Fig. 8.5 a)
 247 and standard deviation energy of hits in shower (Fig. 8.5 b) distributions. Size of this differences
 248 depends on a electron energies and higher for a smaller energies (Fig. 8.6 a) and less significant for
 249 a higher energies (Fig. 8.6 b). This fact combined with energy distribution states an importance of
 250 proper simulation non-uniformities for showers coming from a small energy electrons.

251 On another hand, use of the frozen showers in a small energy region can be suboptimal because
 252 of the small number of energy depositions in a shower. For electrons with energy less than 30 MeV
 253 90% of the showers has zero depositions and just 0.5% of showers are having more than 1 hits (Fig.
 254 8.7). Below this energy single energy spot model have showed better performance in simulation.

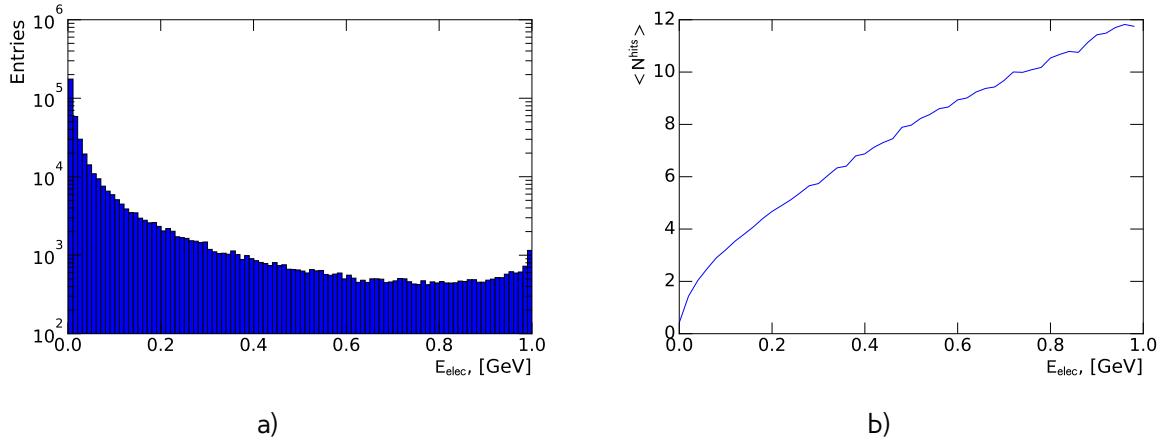


Fig. 8.3: Distribution a) electron energies and b) mean number of hits in a shower vs energy of electron for electrons with energy less than 1 GeV coming from initial electron with energy 1 TeV.

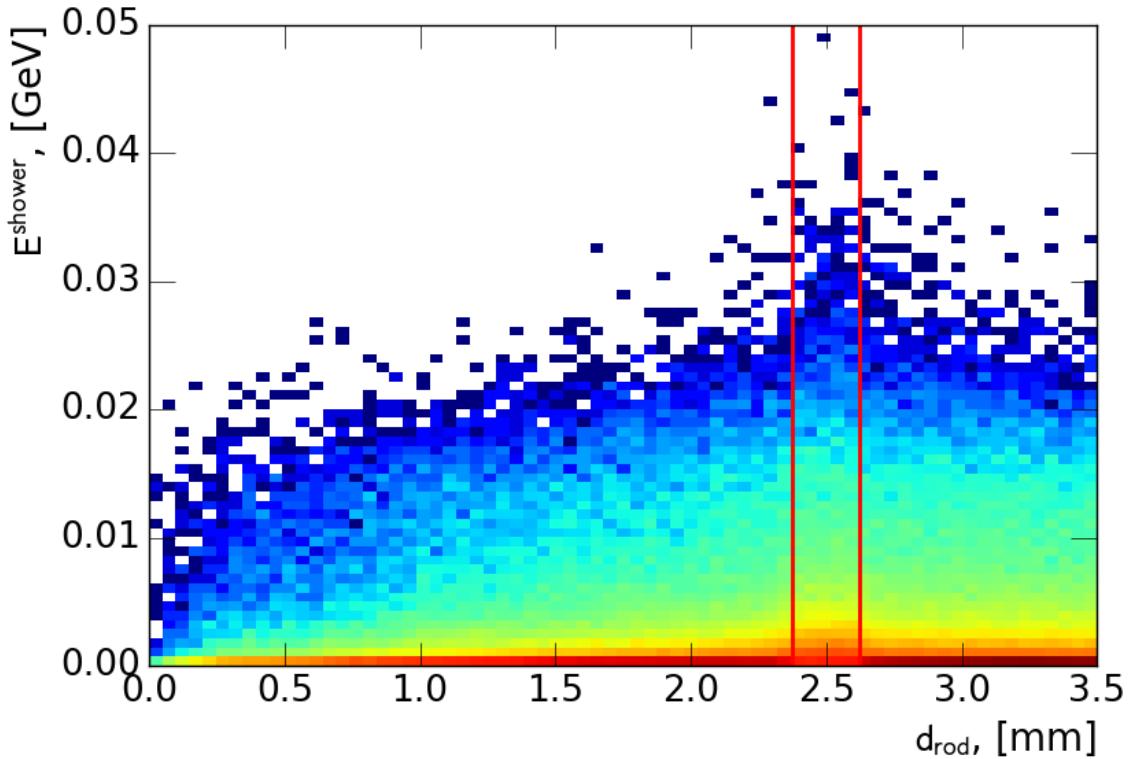


Fig. 8.4: Distribution of electron showers for electrons with energy less than 1 GeV coming from initial electron with energy 1 TeV in distance to a closest rod center vs shower energy plane. Position of a liquid argon gap is noted by a red lines. There is visible difference in shower properties between showers inside and outside of the liquid argon gaps

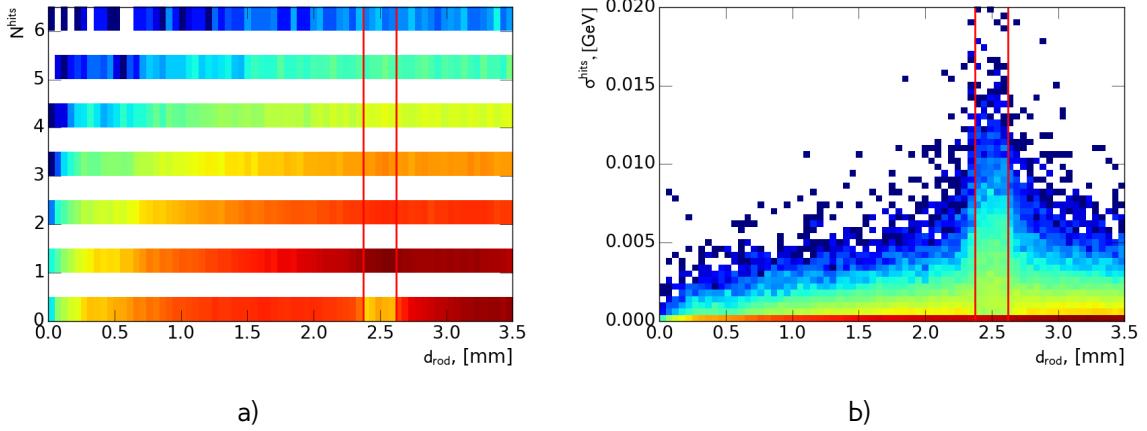


Fig. 8.5: Distribution of electron showers for electrons with energy less than 1 GeV coming from initial electron with energy 1 TeV in distance to a closest rod center vs a) number of hits in a shower plane and b) standard deviation of hits in a shower energy. Position of a liquid argon gap is noted by a red lines. There is visible difference in shower properties between showers inside and outside of the liquid argon gaps

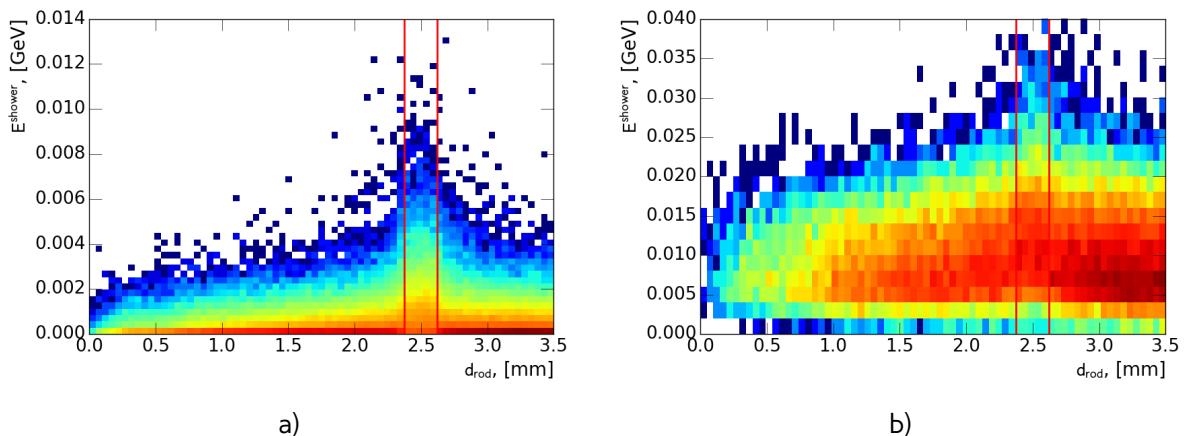


Fig. 8.6: Distribution of electron showers for electrons with energy a) less than 100 MeV and b) higher than 300 GeV coming from initial electron with energy 1 TeV in distance to a closest rod center vs shower energy plane. Position of a liquid argon gap is noted by a red lines. Size of the difference in a shower properties depends on the energy of the electrons and higher for smaller energies

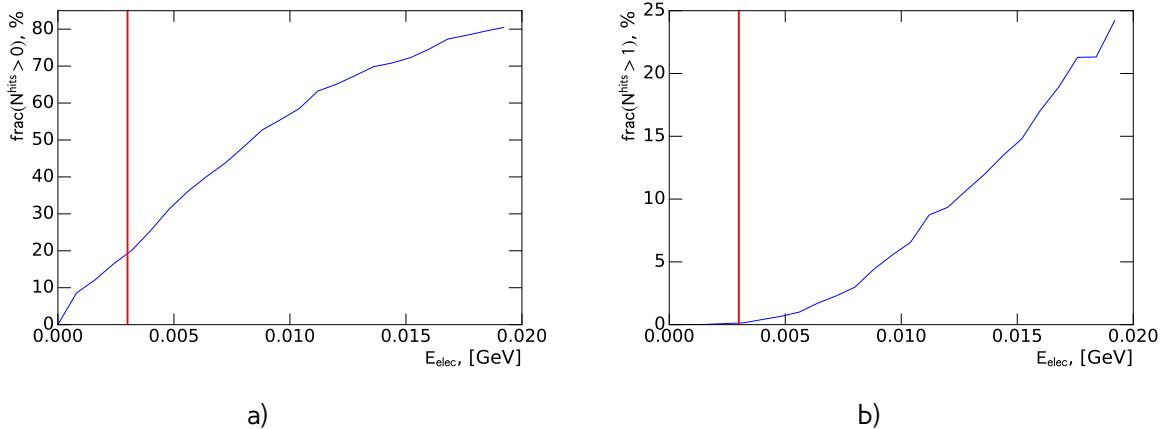


Fig. 8.7: Distribution of fraction of showers with a)at least 1 b) at least 2 depositions inside sensitive material depending on a initial electron energy. Red line denotes 30 MeV lower limit for a frozen showers method.

8.2 Generation and use in simulation

As it was mentioned in introduction frozen showers method consist of the 2 stages: generation of libraries and use in simulation. Generation needs to be repeated for each significant change in a physics processes description in Geant4 or in a description of detector. Showers are stored in a library in pseudorapidity and distance to a closest rod center bins, while energy remains unbinned. Distance binning was introduced to describe fluctuations from Sec. 8.1. Position of the liquid argon gap bin is corresponding to a real gap position.

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

268 Additionally, in order to save disc space as well as a memory consumption, hit information is
269 compressed. This compression is done in a two steps:

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

Truncation hits whose energies are below the fraction f of the total energy sum of all hits, are truncated. The energy of remaining hits is rescaled back to preserve the total deposited energy.

During simulation, if an energy of a particle falls below cut-off energy, the particle algorithm examines resulting shower containment. It checks that particle is far from the edges of calorimeter, so what shower will be by 90% inside calorimeter. This depends also on a energy of particle, because shower sizes are growing with energy. The algorithm searches for a shower with the closes energy in a corresponding pseudorapidity and distance bin. Shower is a rotated in a direction of particle. In

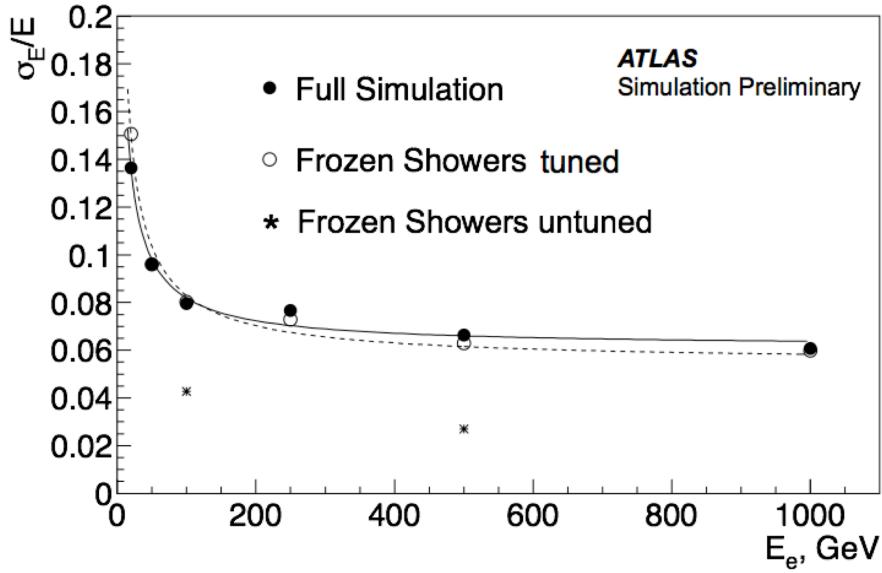


Fig. 8.8: Electron resolution for full simulation, tuned and untuned frozen showers. Electrons simulated with frozen showers libraries before tuning (star points) have twice smaller resolution, than an electrons from full simulation (circles). Tuning (black dots) is allowing to gain better agreement with full simulation.

order to correct differences in energy, each hit in a shower is scaled as:

$$E_{hit}^{new} = E_{hit} \cdot \frac{E_{part}}{E_{part,lib}}, \quad (8.3)$$

where E_{hit} is the energy of the hit, E_{part} is the energy of the particle and $E_{part,lib}$ is the energy of the particle from a library. Then particle is removed and substituted by a resulting shower. Later, reconstruction algorithm uses hits from a frozen shower as a usual energy depositions in a sensitive material.

8.2.1 Libraries tuning

Frozen showers method gives a good agreement with a full simulation for a shower shape variables(Sec. 8.4), however resolution of the reconstructed electrons is around 2 times smaller(Fig. 8.8), than in a full simulation. It can be interpreted as a lack of the showers from liquid argon gap in a simulation. Most of this effect is coming from an electron libraries. This means that this libraries are requiring additional reconstruction-based tuning after generation.

Usual tuning consists of 2-step manual procedure:

Changing bin width On this stage width of the liquid argon bin is enlarged. This is causing larger size of the fluctuations, that leads to a higher resolution and a mean energy of reconstructed electrons

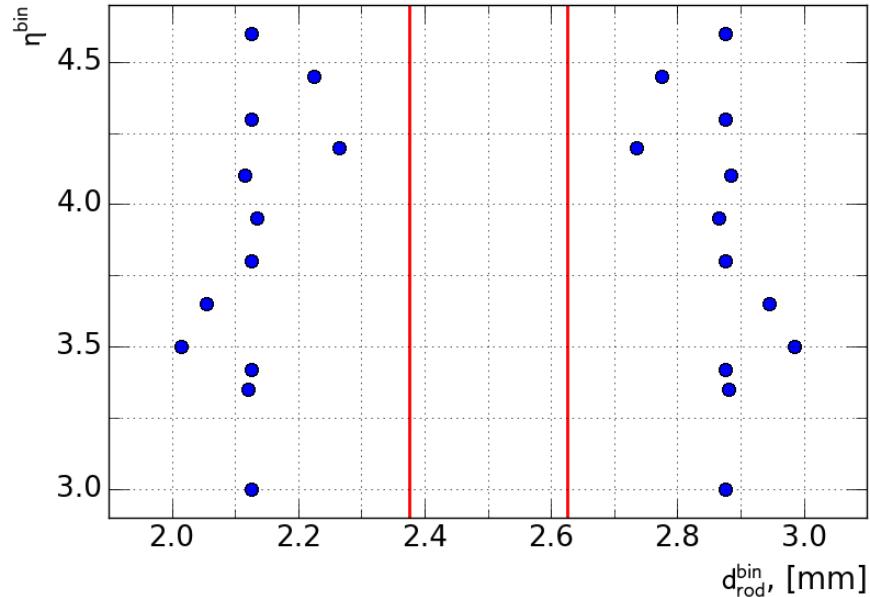


Fig. 8.9: Position of gap bins for different η bins in old libraries after tuning. Dots are corresponding to a limits of each bin. Red lines are denoting original position of bins, that are corresponding to a position of a liquid argon gap in the calorimeter.

289 **Shower energy scaling** In order to correct introduced shift in a mean energy shower energy is
290 reduced by rescaling all of the hits in a shower.

291 It is repeated iteratively in each pseudorapidity bin separately till the desired agreement is obtained.
292 The resulting bin positions are shown on a Fig. 8.9. This method allows to have a relatively good
293 agreement with full simulation (black dots on Fig. 8.8). However, it is necessary to repeat this
294 procedure for each new libraries generation and requires significant tuning effort, that makes it not
295 optimal.

296 8.3 Machine learning based bin finding procedure

297 Since frozen showers are planned to be used in Run-2 Monte-Carlo production, there is a need
298 for a more automatic procedure of library generation with proper electrons resolution. One of the
299 possible ways is to choose different position of liquid argon bin during libraries generation using
300 machine learning tools. In this section automatic bin finding procedure will be discussed.

301 8.3.1 Machine learning introduction

302 Machine learning is a set of algorithms, what allows computers to learn and give a predictions without
303 being specifically programmed. This is a modern field of computer science, that is wildly used in a
304 different fields like computer vision, natural language processing, data science etc. There are two
305 main types of machine learning algorithms: supervised, where example of desired output is given by
306 the "teacher" and the goal is to learn a general rule, that maps inputs to outputs and unsupervised

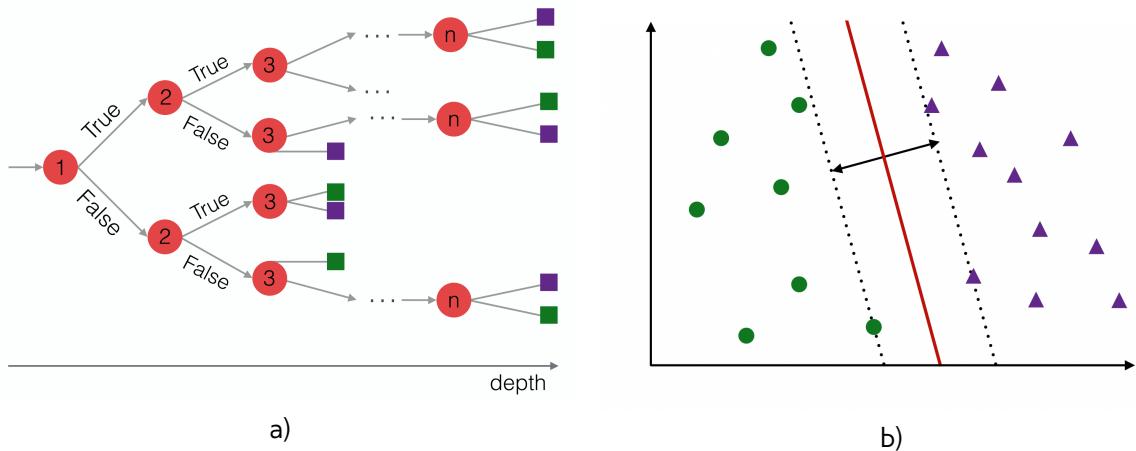


Fig. 8.10: Schematic representation of machine learning algorithms, used in the analysis for a classification. Green figures are representing first class of the events, whereas violet ones belong to a second class. a) Representation of a binary decision tree structure: red circles are corresponding to a node, that are split in the respect to the one of the features. Squares represent leafs, where all of the events are classified to a certain class. Depth of the tree is calculated as a number of edges from the node to the trees first node. b) Representation of the SVM algorithm. Dividing hyperplane is shown by a solid line. The dashed lines represent the maximum margin boundaries

307 learning, then there are no labels given to algorithm, and algorithms is discovering hidden patterns
 308 in data. Initial data parameters of interest, that are used in algorithm to learn are called features.

309 Machine learning algorithms can be used for solving a classification problem, where each event
 310 should be identified to one of the specified classes. Since the first introduction of the machine
 311 learning classifiyng algorithms called perceptron by a Rosenblatt [ref](#) many different algorithms
 312 have been invented. In this analysis decision trees and support vector machines have been used.

313 **Binary decision trees**

314 Binary decision trees are one of the most commonly used machine learning algorithms for a clas-
 315 sification problems in a particle physics. Schematic representation of this algorithm is shown on a
 316 Fig. 8.10 a. Each node is corresponding to the one of the internal input variables and connected to
 317 two branches, that are split in the respect to the a variable. The first node is called a root node.
 318 Depth of the tree is a number of edges from the node to the tree's root node. Leaf node represents
 319 classification or decision. For each node the best feature is selected and it's cut value are obtained
 320 by calculating each possible variation in the feature set and then ranking them. One of the main
 321 advantages of the decision trees is a simplicity of visualization and interpretation.

322 **Support vector machines**

Support vector machines (SVM) is a supervised machine learning algorithm which can be used for classification problems. In this algorithm each event is represented in a p-dimensional parameter space. Classification is performed by finding a the hyper-plane that differentiate the two classes

with the largest separation (Fig. 8.10 b). The hyperplane can be written as the set of points \vec{x} in a parameter space satisfying:

$$\vec{w} \cdot \vec{x} - b = 0, \quad (8.4)$$

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 margin hyperplanes are described by equations:

$$\vec{w} \cdot \vec{x} - b = 1, \quad (8.5)$$

$$\vec{w} \cdot \vec{x} - b = -1, \quad (8.6)$$

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

Because we want to prevent each point to fall into the margin, we following constrain should be satisfied:

$$\vec{w} \cdot \vec{x} - b \geq 1 \text{ where } y_i = 1, \quad (8.7)$$

$$\vec{w} \cdot \vec{x} - b \leq -1 \text{ where } y_i = -1, \quad (8.8)$$

These equations can be rewritten as:

$$y_i(\vec{w} \cdot \vec{x} - b) \geq 1 \quad (8.9)$$

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

$$K_{rbf}(\vec{x}_i, \vec{x}_j) = e^{-\gamma|\vec{x}_i - \vec{x}_j|^2}, \quad \gamma > 0, \quad (8.10)$$

330 where parameter γ adjusts the width of the kernel.

331 8.3.2 Model description

As it was mentioned in a previous sections, modules in FCAL are consisting of different types of material and showers started inside dead material are usually having smaller energies, than a sensitive material once. However, electron, created in a dead material, can cross a liquid argon gap and give a hit there (Fig. 8.13). It was decided to treat this electrons together with electrons created in sensitive material, and call a caused showers as a sensitive material showers. Oppositely, showers that did not crossed a liquid argon gap, are called a dead material showers. This model leads to a bigger gap width by a definition.

339 8.3.3 Interpretation of results

340 . As it was mentioned before, that there are two type of material used in a FCAL. Showers within them
341 are giving different response, what is affecting overall reconstructed electron energy resolution.

At the first generations distance bin have been corresponding to LAr gap or dead material positions. During tuning bin with LAr was enlarged to gain a better agreement with full simulation. So, one of the basic ideas to improve frozen showers performance is to change a size of LAr gap in a library generation.

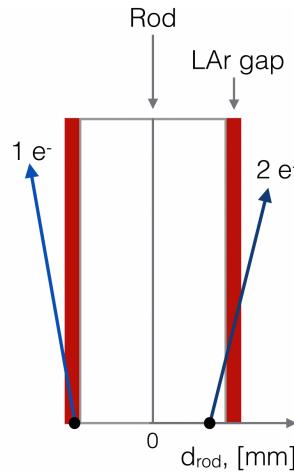


Fig. 8.11: Schematic representation of the model. Electron 1 is created in a liquid argon gap. Electron 2 is created near liquid argon gap and crosses it. This causing a smearing of sensitive material showers distribution. Electrons created in a sensitive material tend to create more energetic showers, than electrons from a dead material. However, electrons, shown on this scheme, may give similar shower and therefore be not distinguishable.

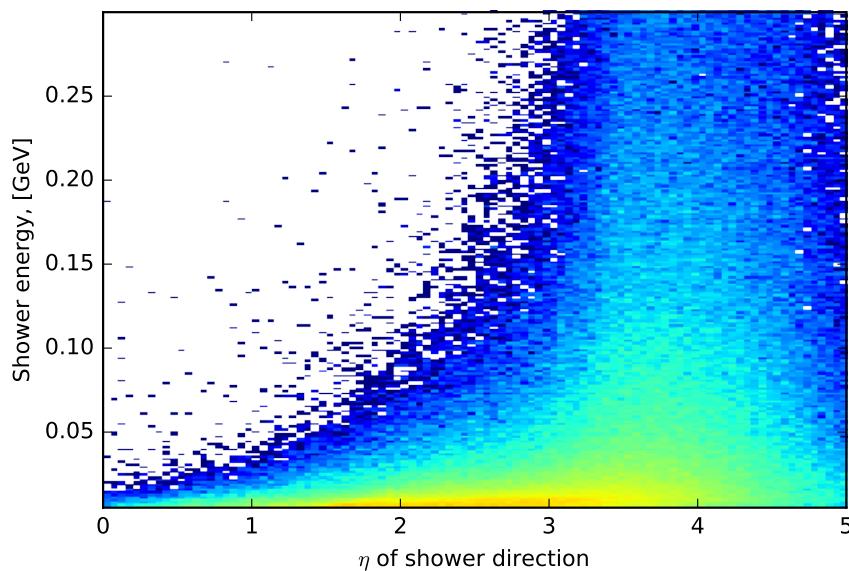


Fig. 8.12: Distribution of showers used in production of 1000 GeV electrons on shower energy vs $\eta_{momentum}$ plane.

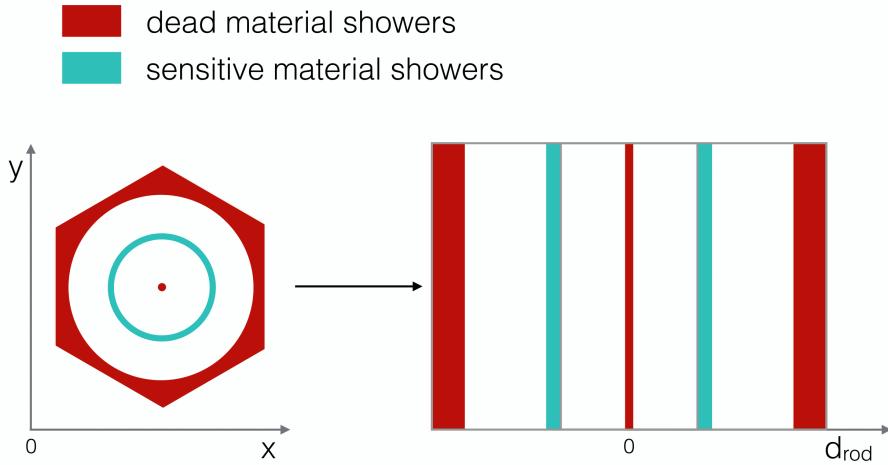


Fig. 8.13: Schematic representation of preselected data for a first classifier in x vs y plane and distance plane. Electrons, created near rod center and on the borders of the module have low probability to cross the sensitive material, while oppositely all of the electrons created inside liquid argon gap are sensitive material showers.

346 It was decided to treat showers, that have been born near LAr gap and crossed it on a radiation
 347 length, in a same way with showers in sensitive material gap, and call them sensitive material showers.
 348 Oppositely, showers, that haven't crossed LAr gap, are called dead showers. This model leads to a
 349 bigger gap width by a definition. One of the possible ways to find this bin position automatically is
 350 to use machine learning tools.

351 Machine Learning

352 From a geometrical point of view, one of the main parameter is a direction of the shower. Eta
 353 momentum distribution is showed on a Figure 8.12 . Most of the showers are collinear to an electron
 354 direction. Because of this it was decided to use as a training sample simulation results for electrons
 355 with energies less than 1 GeV and momentum uniformly distributed between eta 3.0 and 4.0. This
 356 allowing to study equally low and high energy showers equally.

357 From our definition of 2 classes of showers, it is simple to construct a pre-labelled training sample.
 358 This is done by reducing initial sample and taking showers near rod center and inside liquid argon
 359 gap. Output of this classifier, that was trained on with sample with shower features, such as energy
 360 response and number of hits, than can be used to expand our labels to a full distance range. Then
 361 it can be used as an input to a second classifier, which will separate two types of showers using
 362 particle parameters, such as energy and distance to a rod center. For a first step decision trees have
 363 showed good classification efficiency (around 97%). For a second classifier support vector machines
 364 have been used. This method is trying to reconstruct a hyperplane, that is dividing two classes.
 365 Outputs of both of this classifiers are shown on a figure . New gap position is determined using
 366 borders of hyperplane. This procedure is giving expected from the initial model results. Gap is wider,
 367 than and original one. It is also getting bigger with bigger energy, because of the radiation length
 368 growth. Validation results for two different eta bins are shown on figure a) and b). In a bin this new
 369 binning is performing better, than original one without any additional tuning. Unfortunately this is not
 370 true for all of the bins, as we can see on a figure b). This eta bin have showed worst performance

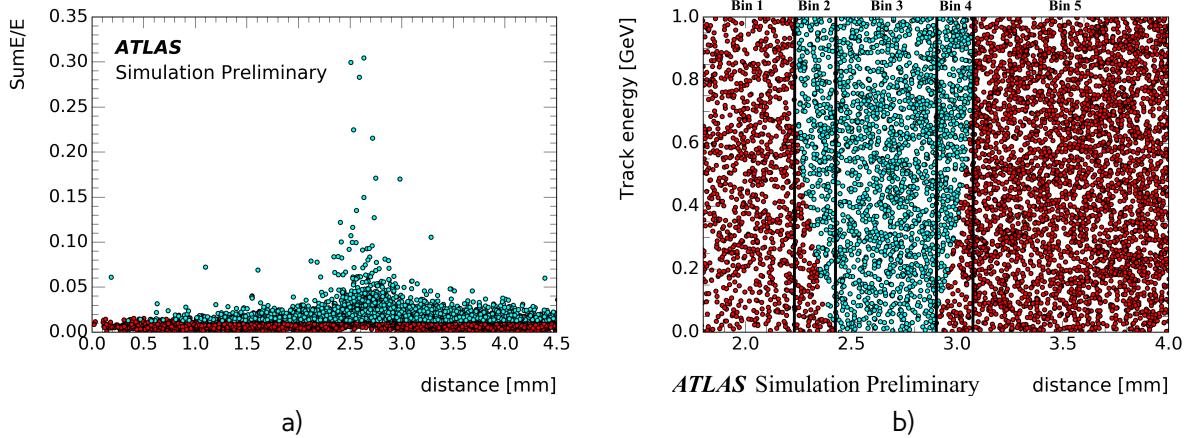


Fig. 8.14: Results of machine learning for a) first classifier b) second classifier. Cyan dots are corresponding to sensitive material showers, red - dead material showers

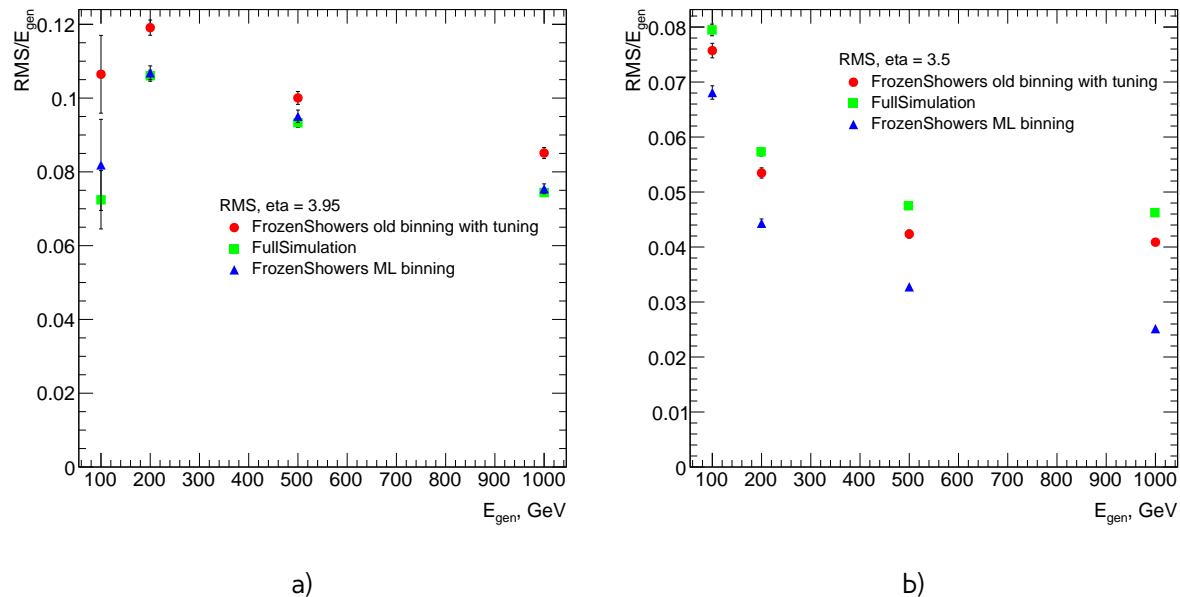


Fig. 8.15: Resolution of reconstructed electrons for full simulation, new libraries with ML binning and old tuned libraries with original binning for a) $\eta = 3.95$ b) $\eta=3.5$

371 for a new binning, but it is performing still better, than original binning without tuning.
 372 This binning was used in a production of new libraries for Monte Carlo in a Run-2. It is planned to
 373 use more precise training sample for a future iterations of this procedure for improving performance
 374 of outlying eta bins.

375 8.4 Validation of the new libraries

376 8.5 Plans for a future developments

377

Chapter 9

378 Data and Monte-Carlo samples

379 9.1 Data sample

380 9.2 Monte-Carlo samples

381

Part IV

382

The Measurement

383 Chapter **10**

384 **Selection**

385 Selection criteria is the set of requirements, that is applied both on data and MC. Analysis depends
 386 on a selection, that can separate process of interest (signal) from other processes. For $pp \rightarrow W \rightarrow$
 387 $ev/\mu\nu$ and $pp \rightarrow Z/\gamma^* \rightarrow ee/\mu\mu$ selection criteria can be divided into 3 groups: data quality, lepton
 388 and boson cuts. In this chapter all of them will be discussed and a cut flow presented.

389 **10.1 Data quality cuts**

Table 10.1: Selection criteria

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

390 Data taking conditions are important in the analysis because of the possible biases. In order
 391 to preserve high data quality some events must be rejected. The reason may be unstable beam
 392 conditions, disabled parts of the detector or events with too many noisy cells. Number of runs, than
 393 can be used for the analysis are stored in a so-called Good Run List (GRL), which in the addition to
 394 the run numbers contains information about luminosity blocks.

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

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

402 **10.2 Lepton quality cuts**

403 Both analyses use similar selection criteria, applied on a leptons. The leptons must satisfy require-
404 ment $P_T^l > 20$ GeV and match to the event trigger.

405 Electron candidates are required to be within pseudorapidity range $|\eta^l| < 2.47$. The electron candi-
406 dates found within the transition region between the barrel and the endcap electromagnetic calorime-
407 ters, $1.37 < |\eta^l| < 1.52$, are removed. Additionally, for a better multijet background rejection medium
408 identification and $P_T^{cone,20} < 0.1$ criterias are applied. The object quality (OQ) cut applied in order to
409 remove events from runs where there dead front end boards in the calorimeter.

410 Muons have to satisfy the following criteria: they should be reconstructed by a staco algorithm in
411 a muon spectrometer and fall within range $|\eta^l| < 2.5$. Set of medium requirements is applied. They
412 must also satisfy $P_T^{cone,20} < 0.1$ isolation criteria.

413 **10.3 Boson selection**

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

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

421 The full set of cuts is summarized in a Tab. 10.1.

422 **10.4 Cut flow**

423 The effect of each selection can be studied using cut-flows, which show the number of events
424 passing each set of cuts in a sequential order. Cut flows for W and Z analysis are shown in a Tab.
425 ?? and Tab. ?? respectively.

426

Chapter 11

427 Monte Carlo corrections

428 Monte Carlo plays important role in the cross-section measurement. It is constantly being improved,
 429 in order to obtain a better precision in data description. Part of these corrections have been described
 430 in Chap. 7. Unfortunately, not everything can be taken into account during simulation itself. This
 431 leads to differences between data and Monte Carlo, that need to be accounted for. There are two
 432 possible methods to correct Monte Carlo without regenerating it. First one is to apply event weights,
 433 so that each MC event can contribute by a non one entry to a histogram. This procedure called
 434 event reweighting. Second one is MC smearing. It uses random numbers to alter the reconstructed
 435 4-vectors. This chapter describes all additional corrections, that have been applied on MC samples in
 436 this analysis. All of these correction are introducing additional systematic error, that will be discussed
 437 in the Chap. 15.

438 11.1 Lepton efficiency corrections

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

- 440 • The reconstruction efficiency ϵ_{rec} is a probability to reconstruct lepton as a lepton of this
 flavor.
- 442 • The identification efficiency $\epsilon_{id|rec}$ is the probability that a reconstructed lepton survives iden-
 tification requirements.
- 444 • The trigger efficiency $\epsilon_{trig|rec,id}$ is the probability, that lepton satisfy trigger requirements.

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

$$\epsilon_{total} = \epsilon_{rec} \times \epsilon_{id|rec} \times \epsilon_{trig|rec,id} \quad (11.1)$$

445 All these efficiencies are measured using tag-and-probe method in $Z \rightarrow ll$ decays. One of the
 446 leptons from Z boson, called "probe", is initially selected with all of the cuts, except the one under
 447 study. Second one, called "probe", satisfies more tighter selection with some additional cut, e.g.
 448 trigger matching.

449 The reconstruction efficiency is associated with the algorithm used in the event reconstruction
 450 process. This is causing differences between electrons and muons efficiencies. In the electron case
 451 the reconstruction efficiency is depicted as a probability to reconstruct an electron which has de-
 452 posited its energy in electromagnetic calorimeter cluster as an electron candidate.

Muon reconstruction efficiency is given by:

$$\epsilon_{reco,muon} = \epsilon_{reco,muon|ID} \cdot \epsilon_{ID} \approx \epsilon_{reco,muon|ID} \cdot \epsilon_{ID|MS}, \quad (11.2)$$

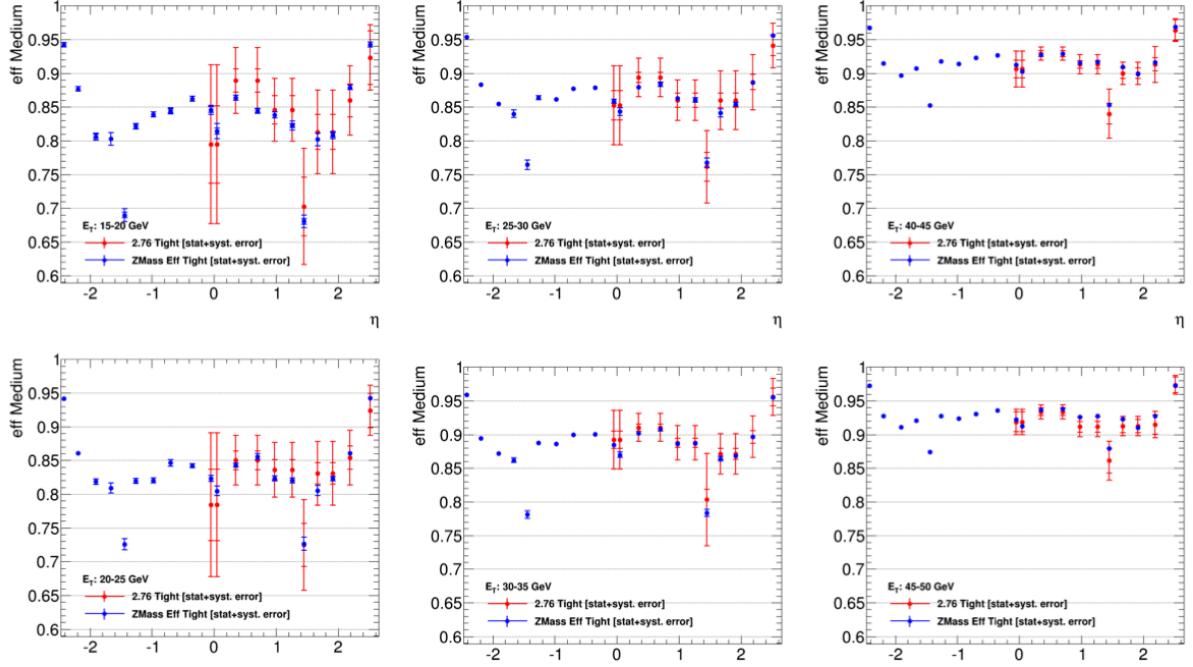


Fig. 11.1: Comparison of electron efficiencies as calculated for 8TeV (blue points) and 2.76TeV (red points) for MC simulation. Efficiencies are shown as a function of pseudorapidity (η) for different electron E_T bins. Both statistical and systematic uncertainties are shown.

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

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

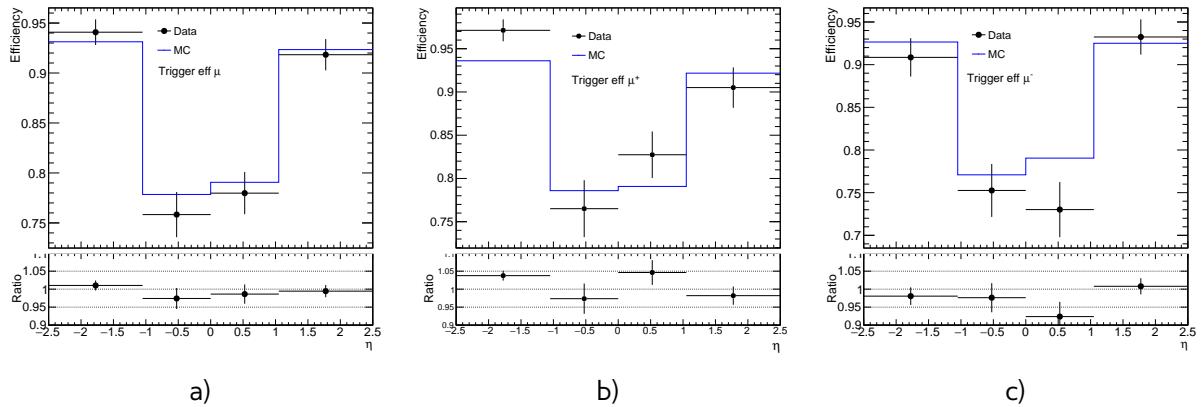
$$SF_{reco,id,trig} = \frac{\epsilon_{reco,id,trig}^{data}}{\epsilon_{reco,id,trig}^{MC}} \quad (11.3)$$

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

462 It is possible however to use scale factors for 8 TeV 2012 data. The main difference between these
 463 data samples are center of mass energy and the pile-up conditions (10 in 2012 and less than 1 in 2013).
 464 This effects have been studied on a $Z \rightarrow ee$ sample. Fig. 11.1 shows the scale factors for different
 465 P_T^l ranges as a function of η^l . The differences in the scale factors are negligible and fully covered
 466 by the statistical errors. This justifies the usage of 8 TeV scaling factors with increased uncertainty
 467 (that is considered to be fully statistical) in the analysis at 2.76 TeV.

Table 11.1: Muon trigger scale factors

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


 Fig. 11.2: Trigger scale efficiencies distribution for a) μ b) μ^+ c) μ^- as a function of pseudorapidity

11.1.1 Muon Trigger SF

Unfortunately, single muon trigger haven't 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 tje scale factors in P_T^l and η^l bins.

Since the P_T^μ cut is significantly higher, than the trigger threshold, efficiency P_T^μ dependency can be considered flat. On the another hand, η dependence are expected. Binning in η is motivated by a detector construction. 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 have been also studied.

Trigger efficiencies for data and MC in η bins are shown in a Fig. 11.2. Total scale factors are presented in Tab. ???. Scale factors for μ^+ and μ^- are more than 3σ away from each other, that is a clear indicator of a charge dependency.

Effect of applying different scale factors on muon for W analysis is shown on Fig. 11.3 - 11.5. Best agreement with data is achieved by applying single bin scale factor. This motivates a choice of single bin charge dependent scale factor for this analysis.

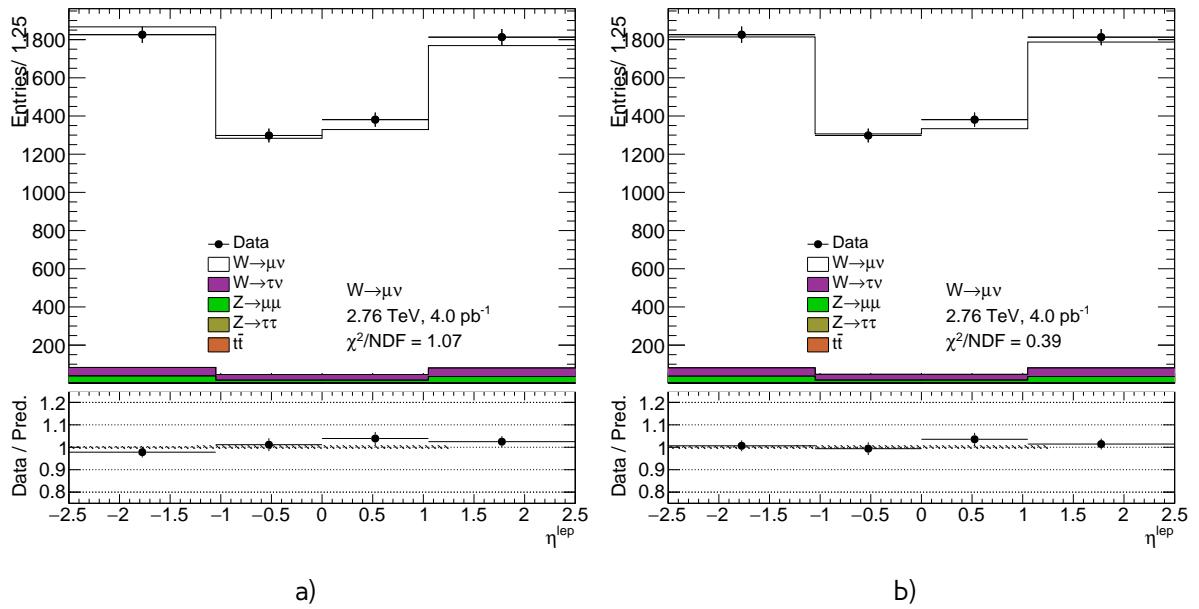


Fig. 11.3: Muon pseudorapidity distribution from the $W \rightarrow \mu\nu$ selection with a) binned b) not binned charge dependent trigger scale factor applied

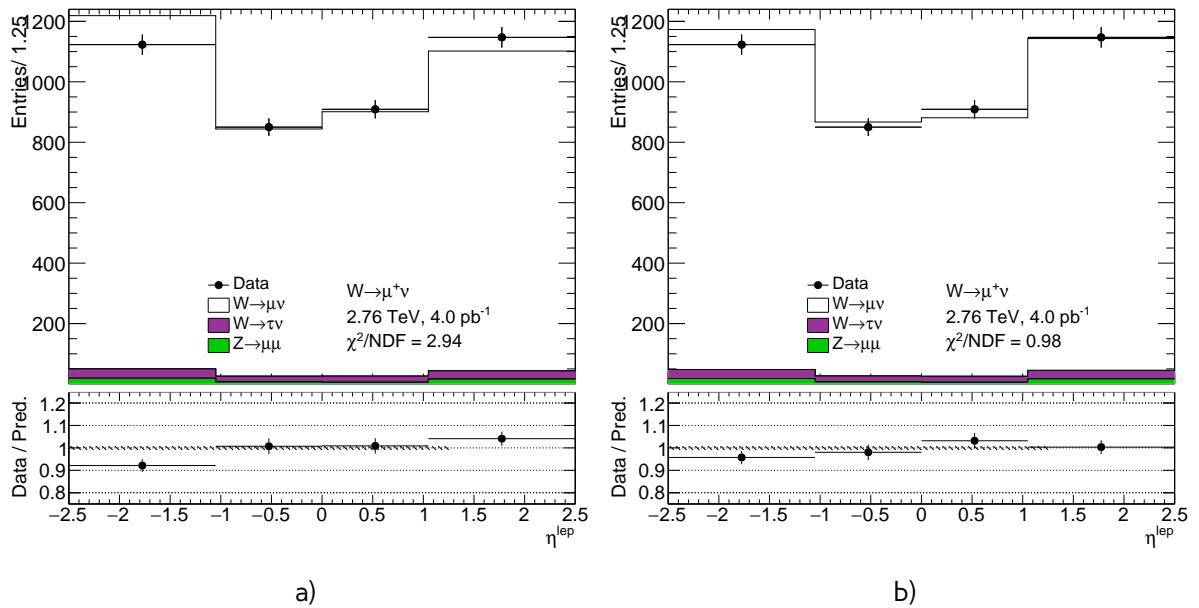


Fig. 11.4: Muon pseudorapidity distribution from the $W \rightarrow \mu^+\nu$ selection with a) binned b) not binned charge dependent trigger scale factor applied

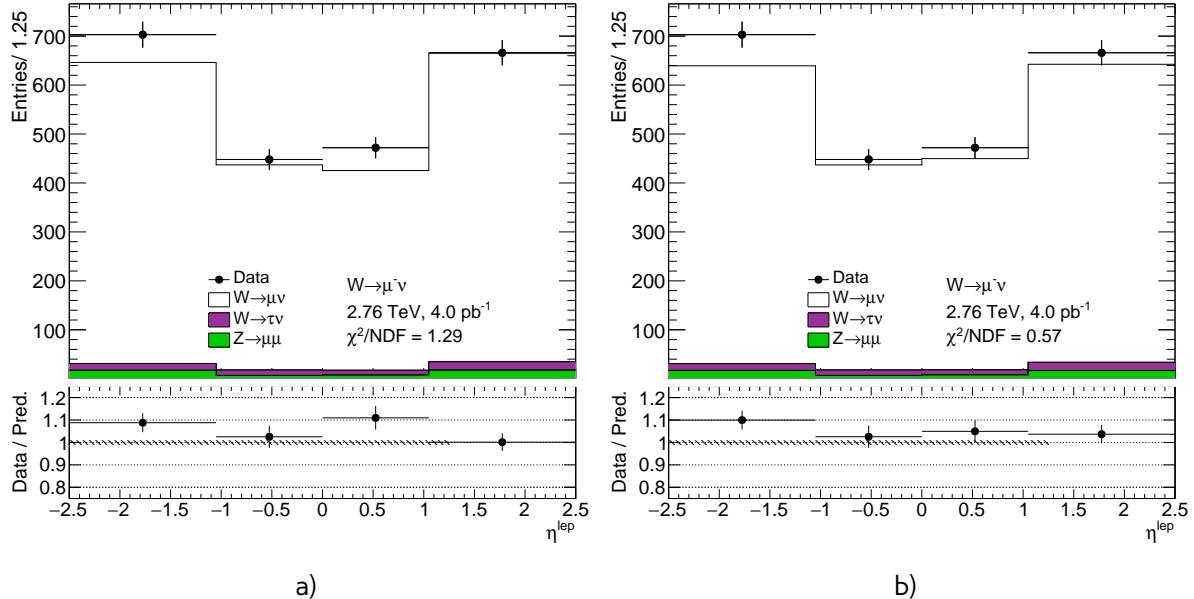


Fig. 11.5: Muon pseudorapidity distribution from the $W \rightarrow \mu^- \nu$ selection with a) binned b) not binned charge dependent trigger scale factor applied

11.2 Electron energy scale and resolution

The reconstructed electron clusters energy tend to be shifted in comparison to the true energy of the initial electron. The correction of this shift is done in both data and MC as a 3 step procedure:

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

The energy shift is parameterised, as:

$$E^{\text{data}} = E^{\text{MC}}(1 + \alpha_i), \quad (11.4)$$

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:

$$m_{i,j}^{\text{data}} = m_{i,j}^{\text{MC}}(1 + \alpha_{i,j}), \quad \alpha_{i,j} \sim \frac{\alpha_i + \alpha_j}{2}, \quad (11.5)$$

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

It is also needed to correct the difference in the electron resolution. It can be described by Eq. 8.1. It is assumed, that sampling and noise terms are well modeled by MC and the main difference is

coming from a constant term. So, the electron resolution correction can be written as:

$$\frac{\sigma_E}{E}^{Data}_i = \frac{\sigma_E}{E}^{MC}_i \oplus c_i \quad (11.6)$$

where c_i is an η dependent relative resolution correction. Similarly to an energy scale correction it is possible to derive resolution correction factor by a comparison of $m_{i,j}^{data}$ and $m_{i,j}^{MC}$ distributions. α_i and c_i correction values are obtained via the χ^2 fit on a invariant mass electrons for data and MC. The resulting energy scale is applied on a data, while the resolution is corrected in MC. The resulting scale is validated using a $J/\psi \rightarrow ee$ and $Z \rightarrow ee\gamma$ samples.

11.3 Muon momentum correction

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

$$\frac{\sigma_{Det}(p_T)}{p_T} = \frac{r_0^{Det}(\eta, \phi)}{p_T} \oplus r_1^{Det}(\eta, \phi) \oplus r_2^{Det}(\eta, \phi) \cdot p_T. \quad (11.7)$$

The first term origins from the fluctuations of the energy loss in the transversed material. The second term is coming from the magnetic field inhomogenities and the local displacements. Third term describes the intrinsic resolution effects.

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

$$p_T^{data} = p_T^{MC} + s_0^{Det}(\eta, \phi) + s_1^{Det}(\eta, \phi) \cdot p_T^{MC}, \quad (11.8)$$

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

This leads to a total correction formula:

$$p_T^{Cor,Det} = \frac{p_T^{MC,Det} + \sum_{n=0}^1 s_n^{Det}(\eta, \phi) (p_T^{MC,Det})^n}{1 + \sum_{m=0}^2 \Delta r_m^{Det}(\eta, \phi) (p_T^{MC,Det})^{m-1} g_m}, \quad (11.9)$$

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

The correction factors are extracted using events with $Z \rightarrow \mu\mu$ candidates fulfilling the requirement of two combined muons. For the correction invariant mass distributions $m_{\mu\mu}^ID$ and $m_{\mu\mu}^{MS}$ are considered individually within a specific $\eta - \phi$ region of fit. 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 with addition of the fit variable:

$$\rho = \frac{p_T^{MS} - p_T^{ID}}{p_T^{ID}}, \quad (11.10)$$

which represents the p_T imbalance between ID and MS.

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

$$p_T^{Cor,CB} = f \cdot p_T^{Cor, ID} + (1 - f) \cdot p_T^{Cor, MS}, \quad (11.11)$$

510 where the weight f is derived from the MC.

511 Chapter **12**

512 **Hadronic recoil calibration**

514	12.1 Introduction	53
515	12.2 Hadronic recoil resolution correction	55
516	12.2.1 $\sum E_T$ distribution correction	56
517	12.2.2 Resolution corrections using Z events	61
518	12.3 Hadronic recoil bias correction	63
519	12.3.1 Bias determination from M_T^W distribution	63
520	12.3.2 Bias determination using u_{\parallel} distribution	64
521	12.3.3 Systematic uncertainty estimation	67
522	12.4 Summary on hadronic recoil systematics	67
523		
524		
525		
526		

527 E_T^{miss} affects significantly the W boson measurement, so it is important to have a solid understand-
528 ing of possible sources differences in the hadron recoil reconstruction in data and Monte Carlo.
529 This chapter describes the procedure of calibration bias and resolution mismodelling in a hadron
530 recoil, that was adapted for 2.76 TeV data.

531 **12.1 Introduction**

532 This analysis uses a standard hadronic recoil calibration procedure, described in [?], that was modified
533 and adapted for the low statistics 2.76 TeV case. The standard procedure consists of the 3 main steps.

534 The first step in a hadronic recoil calibration procedure aims to correct differences in a pile-up
535 modeling in the event. Additional interactions can have a significant effect on E_T^{miss} and $\sum E_T$
536 distributions. These discrepancies are usually corrected by reweighting average number of interactions
537 per bunch crossing in MC to match the data. However, ATLAS simulation is suited for high pile-
538 up runs, so this quantity is modelled discretely in case of 2.76 TeV analysis (Fig. 12.1), what makes
539 the precise reweighting impossible. However, since the mean number is below 1, effect of these
540 discrepancies on E_T^{miss} distributions can be neglected.

541 On the second and third step possible discrepancies in a resolution and scale of hadronic recoil
542 respectively are corrected. The hadronic recoil algorithm performance can be studied in MC through
543 the projection of $\vec{H}R$ on the direction of the transverse momentum of the vector boson, as shown in
544 Fig. 12.2. This projection can be divided into perpendicular u_{\perp} and parallel u_{\parallel} components as follows:

$$u_{\parallel} = \vec{v}_{xy} \cdot \vec{H}R \quad (12.1)$$

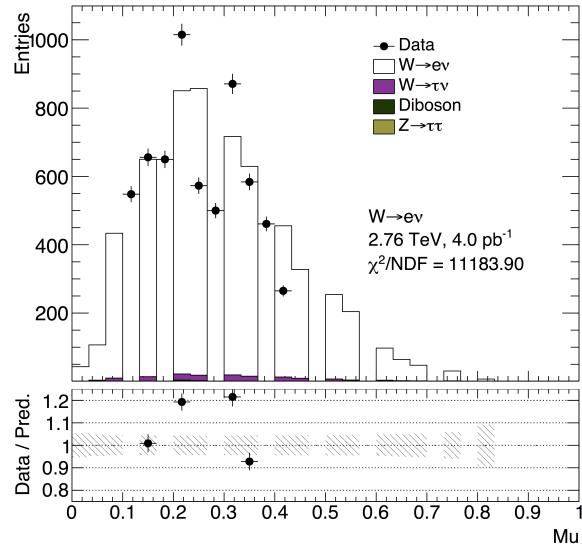


Fig. 12.1: Mean number of interactions per bunch crossing from the $W \rightarrow e\nu$ selection. MC modelled pileup discretely, that makes a standard data to MC reweighting procedure not feasible.

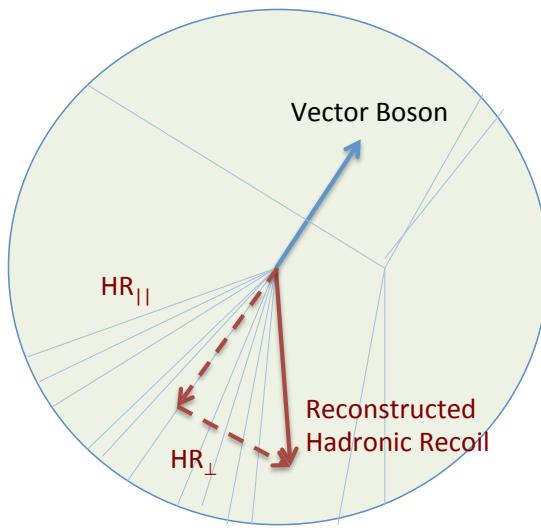


Fig. 12.2: Parallel and perpendicular projections of the hadronic recoil with respect to the transverse momentum of the vector boson [?]

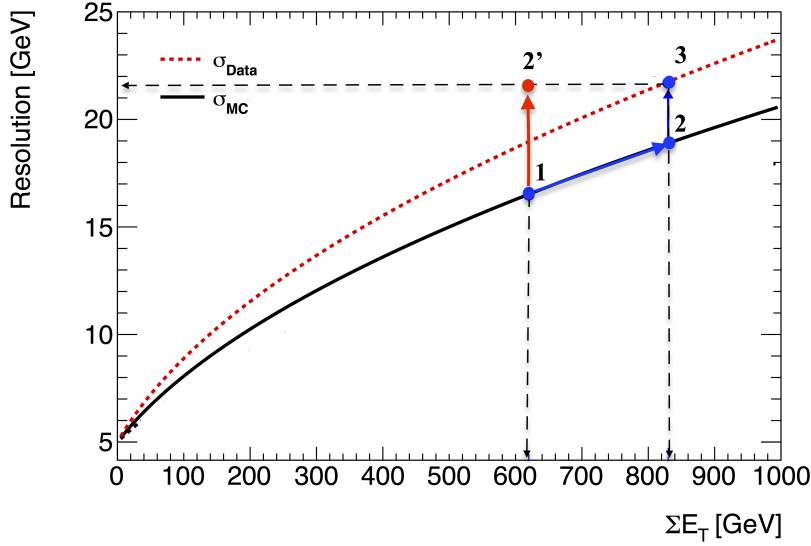


Fig. 12.3: Schematic view of the correction procedure: this figure illustrates the resolution of u_{\perp} as a function of $\sum E_T$. The dotted curve represents data resolution (σ_{data}), solid black is a nominal MC (σ_{MC}). Blue line from point 1 to point 2 corresponds to a $\sum E_T$ correction discussed in Sec.12.2.1. Red line from point 1 to point 2' corresponds to a direct correction of resolution mismodelling discussed in Sec. 12.2.2. Modified from [?]

$$u_{\perp} = v_x \cdot HR_y - v_y \cdot HR_x, \quad (12.2)$$

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 calorimeter resolution is causing relatively wide distributions for these projections. The parallel component u_{\parallel} is sensitive to a possible bias in the hadron 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 event, hadronic activity, boson P_T^{bos} etc.

It is convenient to use Z boson decays for a hadron 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 data and MC. However, small size of the Z sample in 2.76TeV 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 possible bias from P_T^W mismodelling these calibration constants can be derived through data vs MC comparison of P_T^W independent distributions (such as M_T^W). The combined Z and W boson determination procedure has been used.

12.2 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

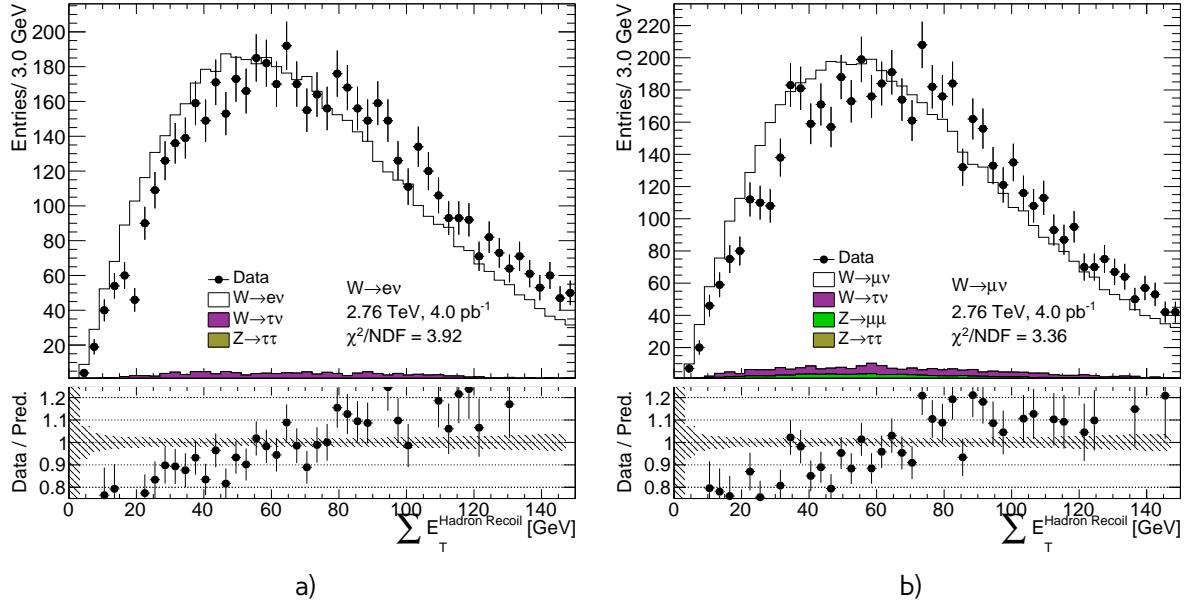


Fig. 12.4: Event activity $\sum E_T$ distribution from a) the $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection. There is a clear sign of the event activity mismodelling in both channels, that should be corrected.

560 differences between the data and Monte Carlo E_T^{miss} resolutions (Fig. 12.3). There are two ways of
 561 resolution correction in the 2.76 TeV data:

- 562 • As a two step procedure, shown as path 1-2-3 in Fig. 12.3. The first step is to correct sumet dis-
 563 tribution to match the data using reweighting of the events. Remaining differences in resolution
 564 are corrected at the second step. This method is discussed Sec. 12.2.1.
 565 • The second order effects on E_T^{miss} coming from $\sum E_T$ modelling are neglected and the resolution
 566 differences between data and MC corrected directly. This procedure matched to the path 1-2'
 567 in Fig. 12.3 and described in Sec. 12.2.2.

568 12.2.1 $\sum E_T$ distribution correction

569 The distributions of the event activity $\sum E_T$ are shown in a Fig. 12.4. There is a visible shift between
 570 data and MC distribution for both W boson channels. Standard procedure, used in M_T^W measurement
 571 at 7 TeV uses a smirnov transformation of $\sum E_T$ distributions in Z events. Unfortunately, size of the
 572 Z sample is not sufficient for this procedure (Fig. 12.5). This motivates a choice of $\sum E_T$ reweighting
 573 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. 12.6, so in order to avoid possible 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 reweighting constants are calculated as:

$$SF_{channel} = \frac{\sum E_T^{data,selection}}{\sum E_T^{MC,no\ cuts}}, \quad (12.3)$$

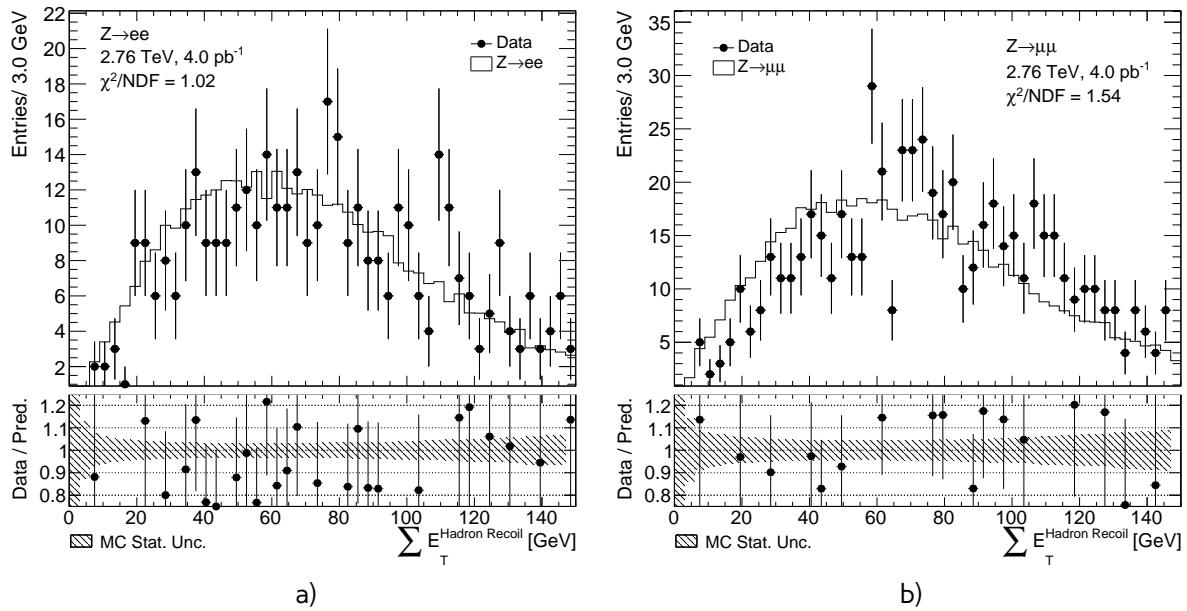


Fig. 12.5: Event activity $\sum E_T$ distribution from a) the $Z \rightarrow ee$ selection and b) the $Z \rightarrow \mu\mu$ selection. Size of the Z sample in 2.76 TeV data is insufficient indicate to the mismodelling of the event activity.

574 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
 575 to reduce systematic error from this value, a combination of $W \rightarrow ev$ and $W \rightarrow \mu\nu$ events is
 576 used. Second term $\sum E_T^{MC, no\ cuts}$ is $\sum E_T$ distribution in MC before any selection. Scale factors are
 577 determined separately for each signal MC for W boson decays, in order to leave the total number of
 578 events in MC after correction untouched. Transverse boson momentum binning is chosen so what
 579 there is an approximate equality in total number of events between bins. The total number of $P_T^{W, rec}$
 580 bins is 6. Example of correction factors for two different $P_T^{W, rec}$ bins are shown on a Fig.12.7. Resulting
 581 reweighting constants for $W \rightarrow ev$ and $W \rightarrow \mu\nu$ MC samples are shown in Fig. 12.8. This method
 582 allows to leave the reconstructed transverse momentum of the boson untouched and introduces a
 583 small change in a truth boson spectrum, as shown on Fig. 12.10.

584 Systematic error estimation

585 Systematic error on this reweighting can be estimated using the approximation of the ratio inside
 586 each $P_T^{W,rec}$ bin by a polynomial degree 2 or 1. This method also allows to neglect effect of the data
 587 fluctuations, especially for high $\sum E_T$ regions, that could be seen on a Fig. 12.8. Because of the low
 588 statistics for $\sum E_T > 220$ GeV ratio in the last bins have been set to 1 and this region haven't been
 589 included in the polynomial fit. Total reweighting constants obtained from this procedure are shown
 590 on a Fig. 12.11.

591 Statistical error estimation

Statistical error on the $\sum E_T$ reweighting is estimated using Toy MC method, described in Chap.15 from the polynomial order 2 approximation. The fitted parameters of the polynomials are varied inside each $p_T^{W,rec}$ bin within their fit uncertainties as in Eq. 15.2. Because of the possible correlations

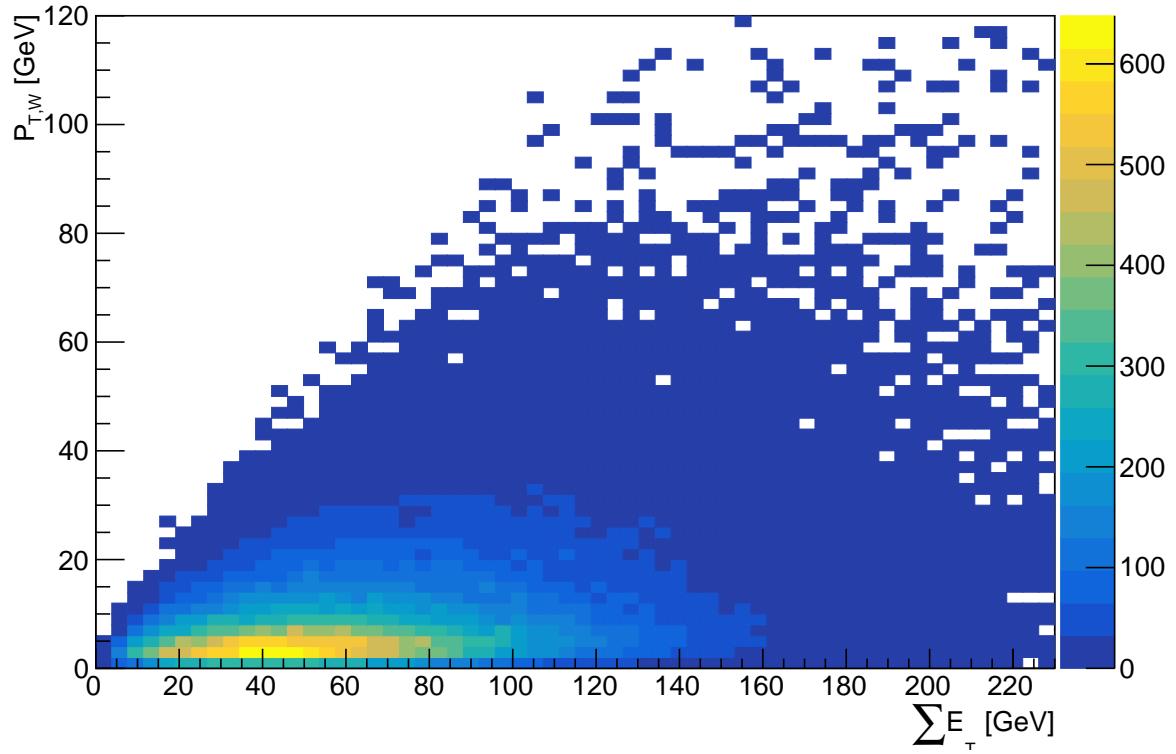


Fig. 12.6: Distribution of event activity $\sum E_T$ vs truth transverse momentum of the W boson P_T^W *truth* in the $W^+ \rightarrow e\nu$ MC sample.

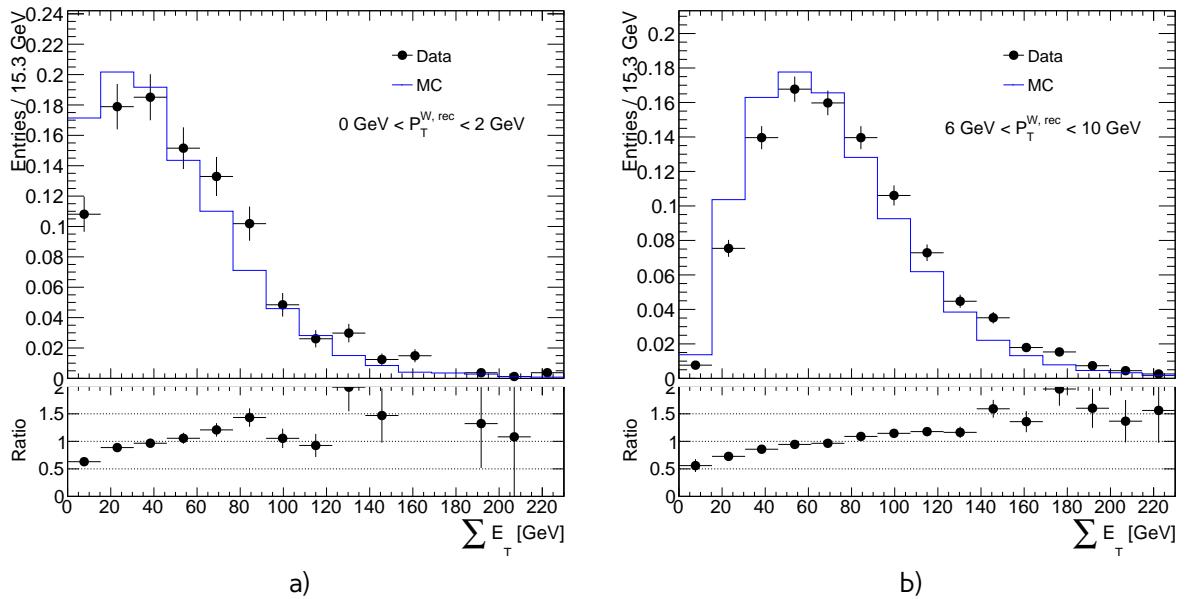


Fig. 12.7: Distribution of $\sum E_T$ for the different $p_T^{W,rec}$ bins for $W^+ \rightarrow e\nu$ MC sample

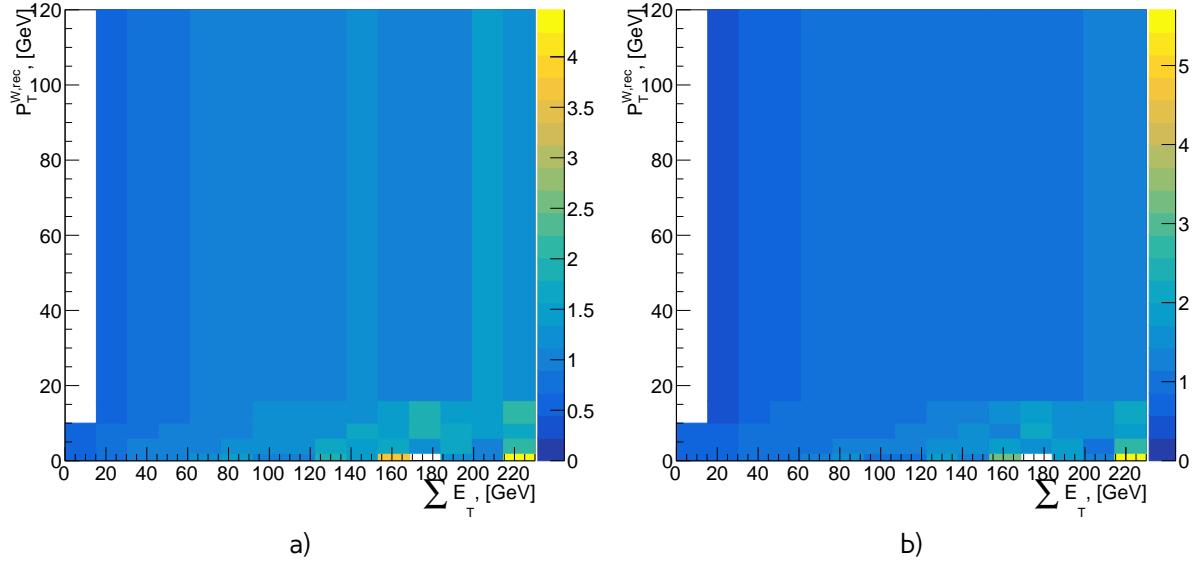


Fig. 12.8: Distribution of $\sum E_T$ reweighting constants derived for a) $W^+ \rightarrow e\nu$ and b) $W^+ \rightarrow \mu\nu$ MC sample.

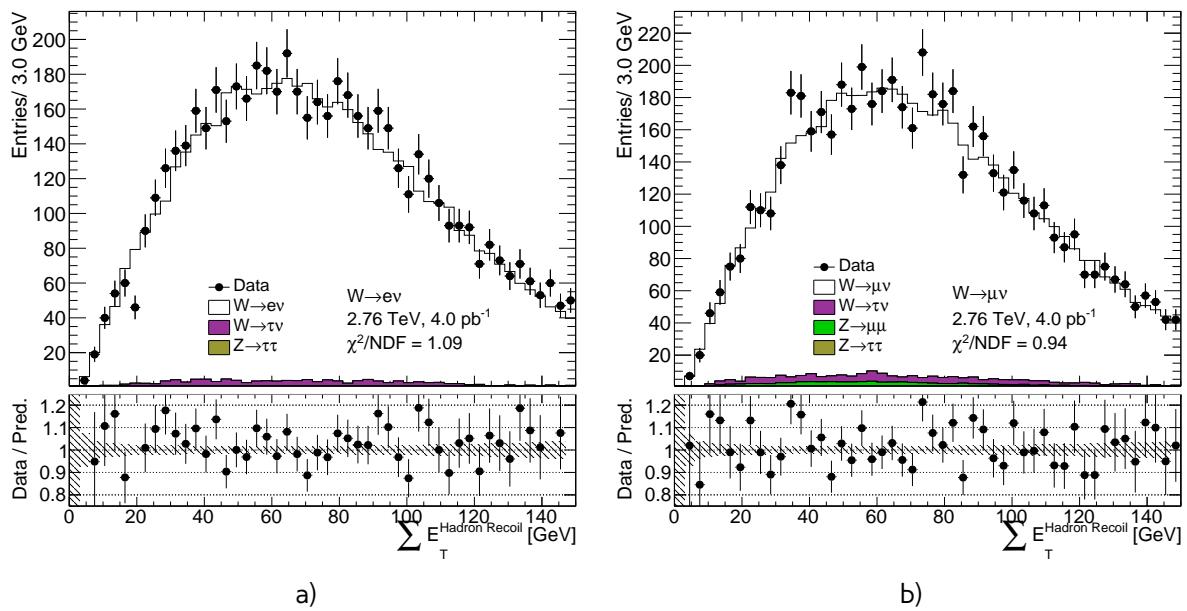


Fig. 12.9: Event activity $\sum E_T$ distribution from a) the $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection after $\sum E_T$ correction.

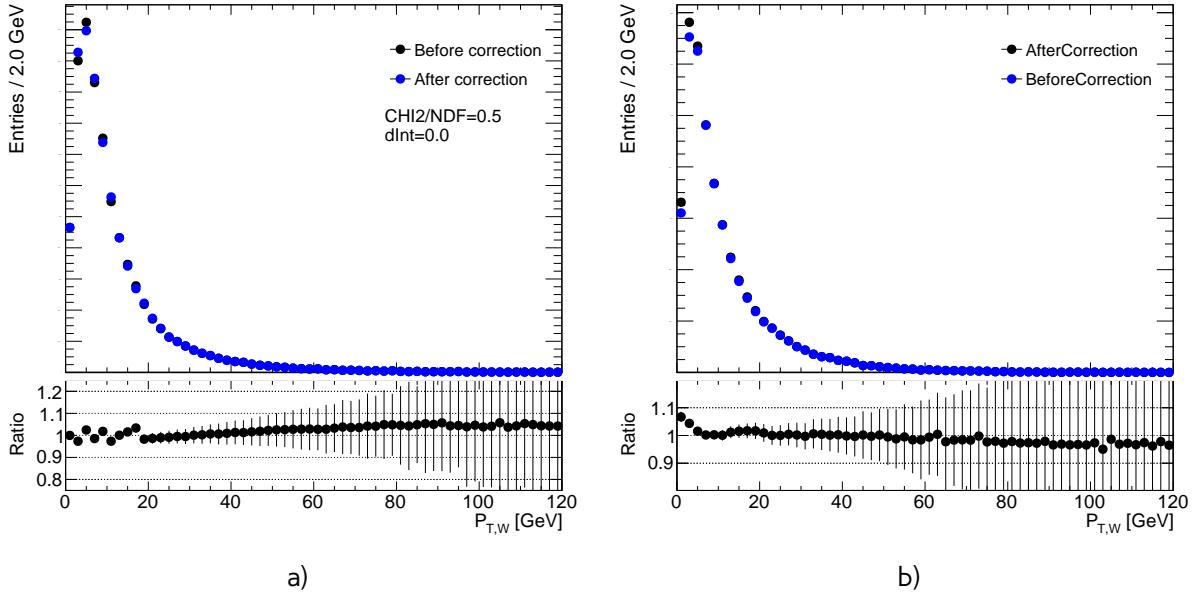


Fig. 12.10: Effect of the $\sum E_T$ reweighting on a) reconstructed transverse momentum of the boson and b) truth transverse momentum of the boson in $W^+ \rightarrow e\nu$ MC sample.

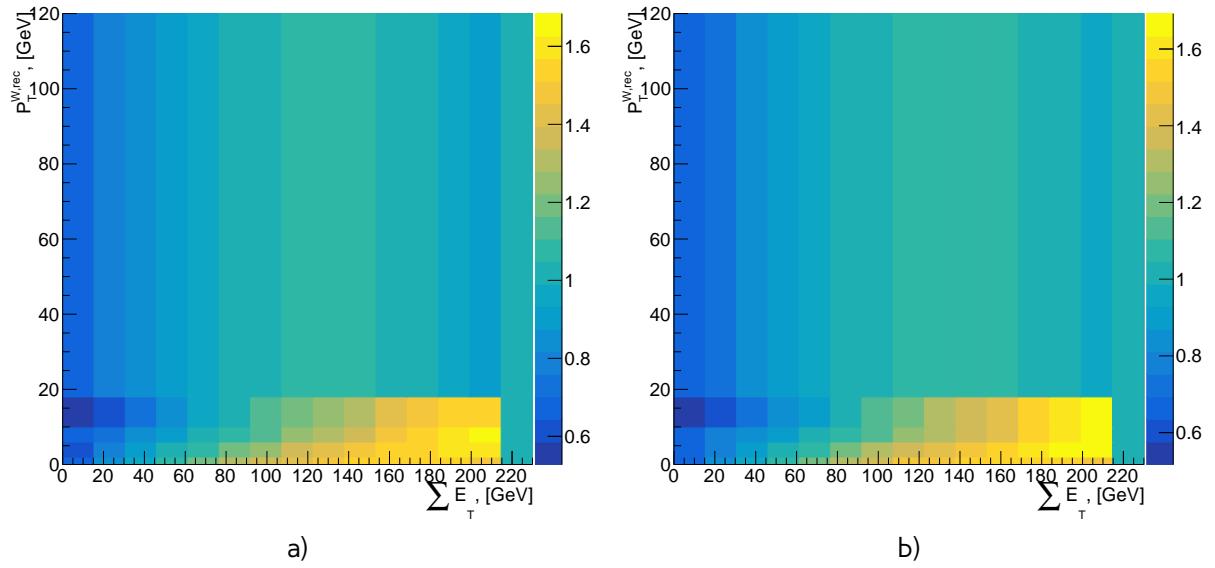


Fig. 12.11: Distribution of $\sum E_T$ reweighting constants derived for a) $W^+ \rightarrow e\nu$ and b) $W^+ \rightarrow \mu\nu$ MC sample using polynomial order 2 approximation.

Table 12.1: Effect of $\sum E_T$ correction on a C_W for a different channels and methods

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

between fitted parameters, the multivariate Gaussian distribution have been used, that is calculated as:

$$p(x; \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right), \quad (12.4)$$

where $\mu \in \mathbf{R}^n$ are the obtained from the fit paremeters and Σ is $n \times n$ covariance matrix of these parameters. In case of the polynomial order two n=3. For statistic error determination total number of 25 toys have been used. Total error is calculated using Eq.15.3.

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

Overall effect of the $\sum E_T$ correction for different methods is summarised in Tab. 12.1. Statistical error, estimated using Toy MC method is negligible. Sign of the effect differs for different W channels, that cannot be explained by a systematic error coming from the method. This effect also cannot be explained from a physical point of view, so it was decided not to use this corrections in a final analysis.

12.2.2 Resolution corrections using Z events

Another way to check resolution effects is to study u_\perp and $u_\parallel - p_T^Z$ distributions in events containing Z boson. This correction assumes, that any resolution mismodelling reflects discrepancies in the $\sum E_T$ distribution, while the difference in the resolution at a given $\sum E_T$ is subleading. There is a clear difference in an RMS of these distributions between data and MC, that cannot be accounted for a statistical uncertainties in the data (Fig. 12.12 and Fig. 12.13). The difference in resolutions is consistent for u_\perp and $u_\parallel - p_T^Z$ distributions, but depends on an analysis channel and isolation + identification criteria.

The resolution is corrected by smearing with a Gaussian distribution each component of a hadronic recoil:

$$u'_\parallel = u_\parallel + \text{Gaus}(0, d\sigma) \quad (12.5)$$

$$u'_\perp = u_\perp + \text{Gaus}(0, d\sigma), \quad (12.6)$$

where $d\sigma$ is a difference in a resoultions calculated as:

$$d\sigma = \sqrt{\sigma_{data}^2 - \sigma_{MC}^2} \quad (12.7)$$

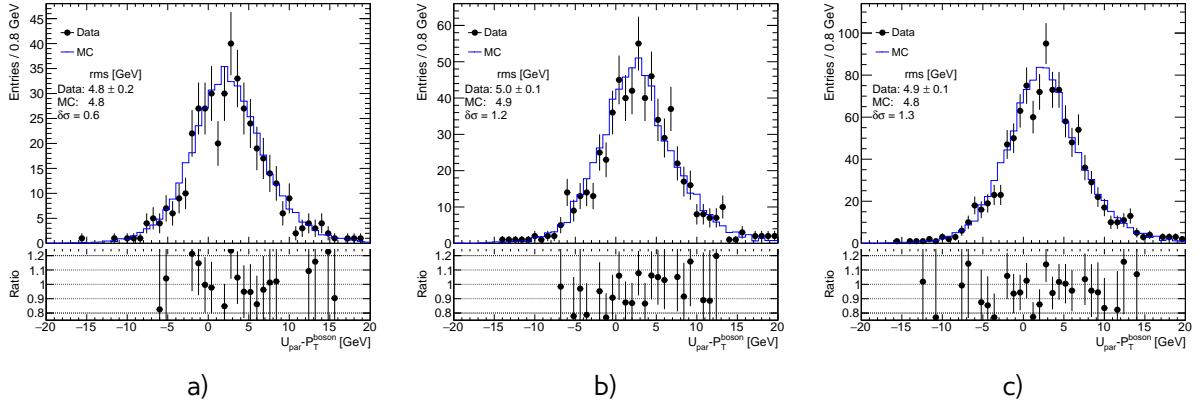


Fig. 12.12: Parallel hadronic recoil component distribution from a) the $Z \rightarrow ee$ selection b) $Z \rightarrow \mu\mu$ selection and c) $Z \rightarrow ll$ selection. The expected contribution from signal is estimated with Monte Carlo simulation, other background sources are considered negligible.

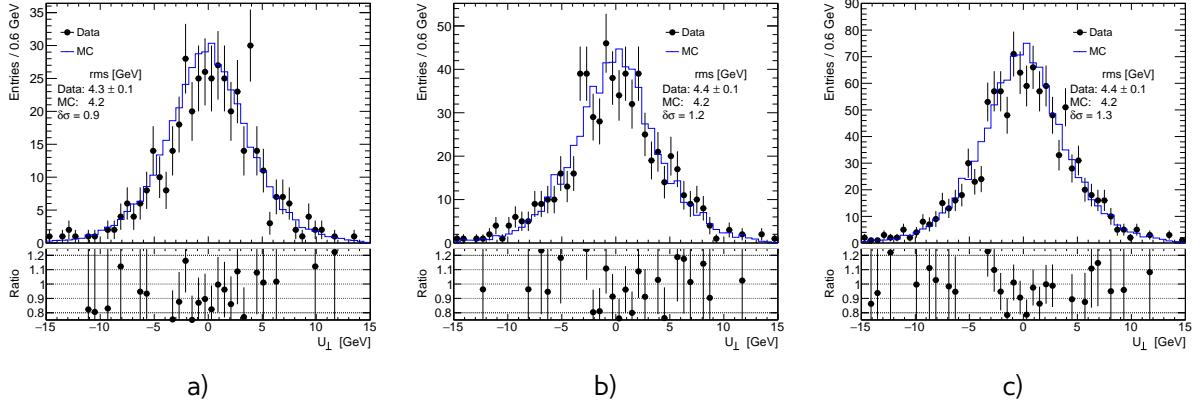


Fig. 12.13: Perpendicular hadronic recoil component distribution from a) the $Z \rightarrow ee$ selection b) $Z \rightarrow \mu\mu$ selection and c) $Z \rightarrow ll$ selection. The expected contribution from signal is estimated with Monte Carlo simulation, other background sources are considered negligible.

Table 12.2: Effect of smearing correction on a C_W for a different channels

Channel	δC_W	error
$W^+ \rightarrow e^+\nu$	-0.20%	0.04%
$W^- \rightarrow e^-\nu$	-0.11%	0.06%
$W^+ \rightarrow \mu^+\nu$	-0.16%	0.04%
$W^- \rightarrow \mu^-\nu$	-0.12%	0.07%

611 Systematic error of this $d\sigma$ is taken as an statistical error for σ_{data} . Due to a random nature of this
 612 correction, effect is not stable for a small $d\sigma$. Systematic error have been estimated by repeating
 613 correction on the same sample 40 times. Overall systematics effect is below 0.2% for each channel
 614 (Tab. 12.2).

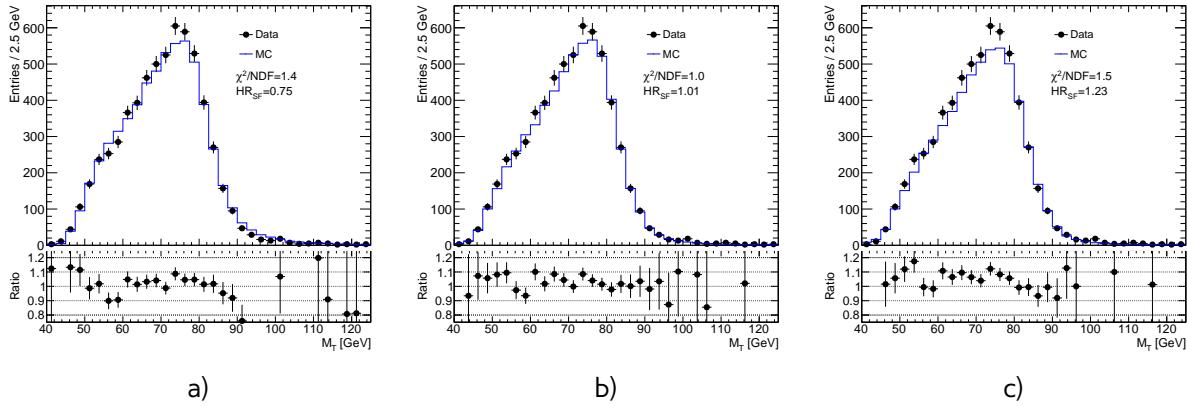


Fig. 12.14: Mass transverse distribution from the $W \rightarrow ev$ selection for different hadronic recoil scales: a) $HR_{SF}=0.75$ b) $HR_{SF}=1.1$ c) $HR_{SF}=1.23$. The expected contributions from signals and backgrounds are estimated with Monte Carlo simulation, except for a QCD background, that is not included.

615 12.3 Hadronic recoil bias correction

As it was mentioned before, it is possible to use both Z and W boson samples for the hadronic recoil bias determination. Correction factor HR_{SF} is applied as:

$$u_{\parallel}^{MC,cor} = u_{\parallel}^{MC} \cdot HR_{SF} \quad (12.8)$$

and can be obtained by scanning the impact of the scaling factor on the data to MC agreement of the distributions that are dominated by the recoil scale uncertainties. The best correction factor and its error is obtained from the fit of χ^2 distributions with the function:

$$\chi^2 = \frac{(HR_{SF} - sf_{best})^2}{\sigma_{sf}^2} + \chi_0^2, \quad (12.9)$$

616 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
617 χ^2 in a minimum.

618 This procedure can be used in both Z and W boson samples.

619 12.3.1 Bias determination from M_T^W distribution

620 Since the W boson transverse momentum cannot be measured in two different ways in order to
621 provide the reference for a hadronic recoil scale, determination of the hadronic recoil bias should
622 use the distributions, that are not sensitive to the truth P_T^W spectrum. One of the optimal choices
623 is the M_T^W distribution. The transverse mass distribution for a different scale choices is shown on a
624 Fig. 12.14. The expected contributions from signals and backgrounds are estimated with Monte Carlo
625 simulation, except for a multijet background, because its shape and number of events depends on
626 a hadronic recoil scale and thus needs to be recalculated for each value of HR_{SF} .

One of the possible ways to determine the correction factor is to use a difference in the mean of the transverse mass distributions in data and MC (Fig. 12.15). Statistical error on a correction factor is considered a dominating one and estimated as an standard error of mean $\sigma(\langle M_T^W \rangle)$, that is

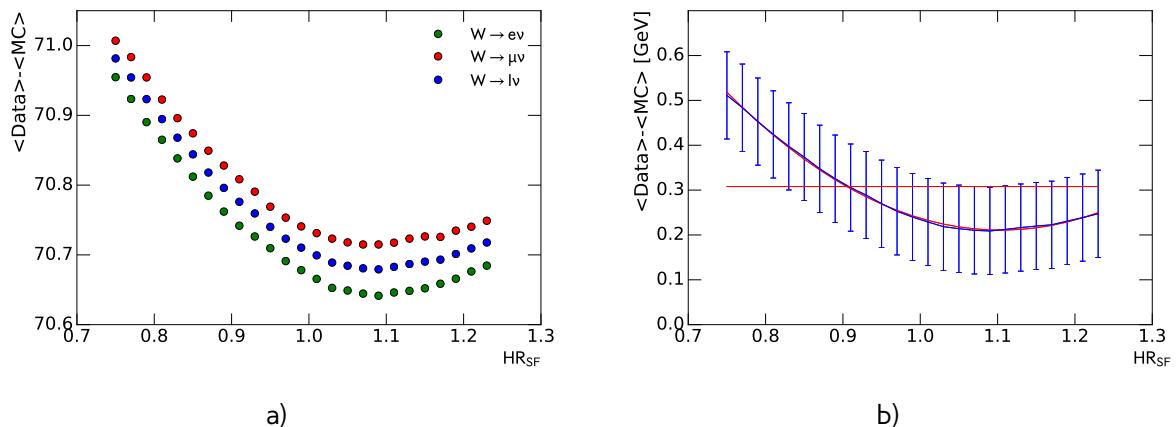


Fig. 12.15: a) Distribution of difference in a mean transverse momentum $\langle M_T^W \rangle$ between data and MC as a function of hadronic recoil scale HR_{SF} for different W boson channels. b) Distribution of difference in a mean transverse momentum $\langle M_T^W \rangle$ between data and MC as a function of hadronic recoil scale HR_{SF} for combined $W \rightarrow l\nu$ selection. Errors for each point are calculated as a standard error of mean. Below red line are values of HR_{SF} within 1σ uncertainty. The expected contributions from signals and backgrounds are estimated with Monte Carlo simulation, except for a QCD background, that is not included.

calculated as:

$$\sigma \left(< M_T^W > \right) = \frac{\sigma(M_T^W)}{\sqrt{N}}, \quad (12.10)$$

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

Distribution of χ^2 for different W selections is shown in a Fig. 12.16 a). Because of the possible mismodelling of the tail M_T^W distribution events with $M_T^W > 100$ GeV are not included in a χ^2 calculation. There is a visible peak in the χ^2 distribution for events from $W \rightarrow e\nu$ selection, coming from the missing QCD background. Hadronic recoil bias parameters are determined through the fit of χ^2 distribution in combined $W \rightarrow l\nu$ channel with the function from Eq. 12.9. The resulting bias is $HR_{SF} = 1.02 \pm 0.06(\text{stat})$.

Additionally, a cut on M_T^W lower value may be used to reduce the multijet background contamination. The M_T^W range introduces a source of systematic uncertainty in hadronic recoil scale determination. It is estimated by repeating the fit for the different M_T^W lower and upper values, as shown in Fig. 12.17. Fit range systematic error is 0.03, that is resulting in overall result for this method $HR_{SF} = 1.02 \pm 0.07$.

12.3.2 Bias determination using u_{\parallel} distribution

Similarly to the W channel, scale correction in the Z sample can be determined from the HR_{SF} scan in the $\frac{u_{\parallel}}{p_T^{\parallel}}$ distribution, as shown on a Fig. 12.18. Since there is no choice of the range and dependency on P_T^{bos} modeling, there is just one source of uncertainty.

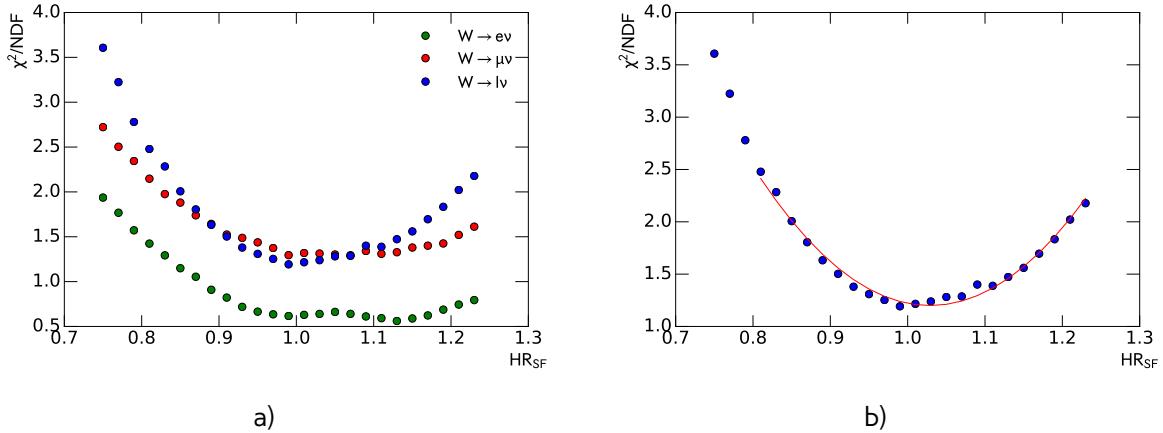


Fig. 12.16: a) Distribution of χ^2 between data and MC for transverse momentum $< M_T^W >$ as a function of hadronic recoil scale HR_{SF} for different W boson channels. b) Distribution of χ^2 between data and MC for transverse momentum $< M_T^W >$ as a function of hadronic recoil scale HR_{SF} for combined $W \rightarrow l\nu$ selection. Fit results are shown by a red line. The expected contributions from signals and backgrounds are estimated with Monte Carlo simulation, except for a QCD background, that is not included.

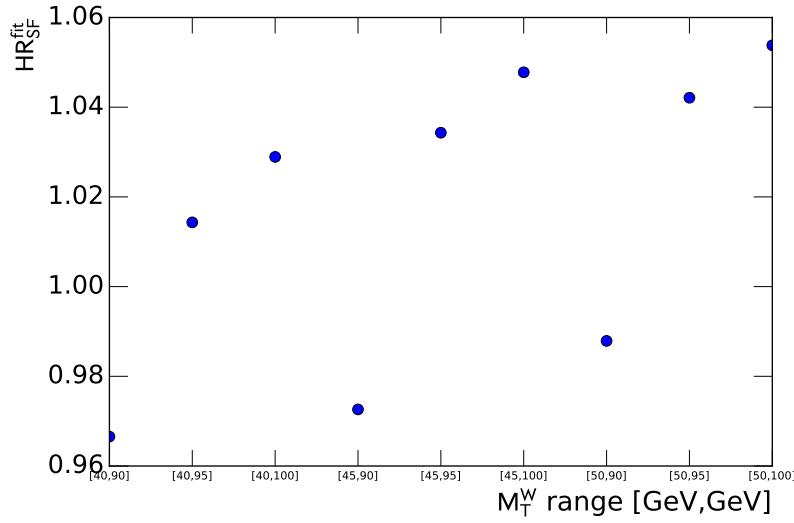


Fig. 12.17: Values of best hadron recoil biases obtained from the fit for events from combined $W \rightarrow l\nu$ selection as a function of fit range. The expected contributions from signals and backgrounds are estimated with Monte Carlo simulation, except for a QCD background, that is not included.

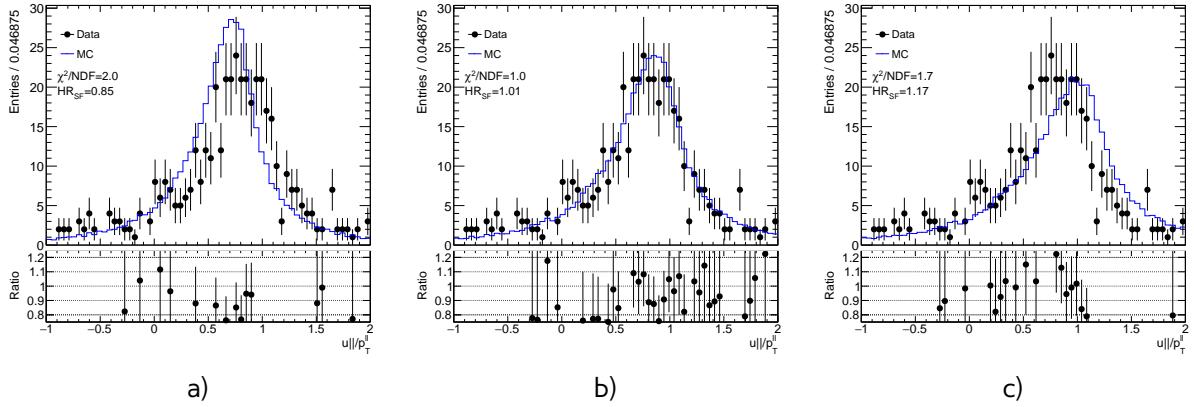


Fig. 12.18: Parallel hadronic recoil component u_{\parallel} from the $Z \rightarrow ee$ selection for different hadronic recoil scales: a) $HR_{SF}=0.75$ b) $HR_{SF}=1.1$ c) $HR_{SF}=1.23$. The expected contribution from signal is estimated with Monte Carlo simulation, other background sources are considered negligible.

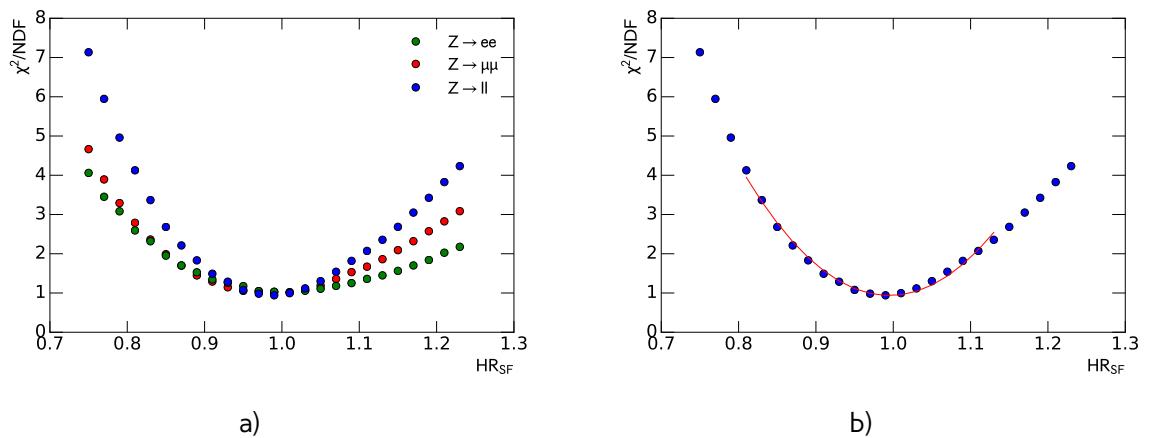


Fig. 12.19: a) Distribution of χ^2 between data and MC for transverse momentum $< M_T^W >$ as a function of hadronic recoil scale HR_{SF} for different W boson channels. b) Distribution of χ^2 between data and MC for transverse momentum $< M_T^W >$ as a function of hadronic recoil scale HR_{SF} for combined $W \rightarrow l\nu$ 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.

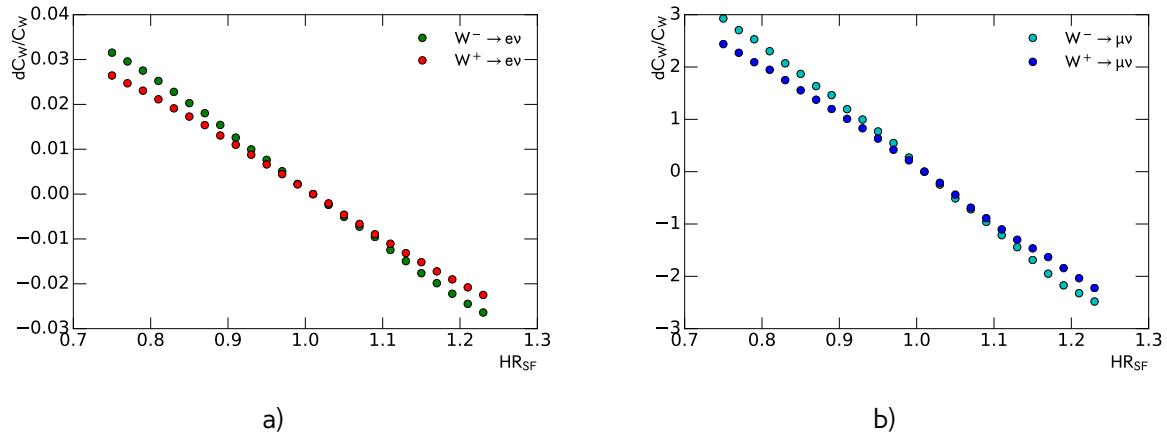


Fig. 12.20: Effect on a C_W for a different $d\sigma$ for a) $W \rightarrow e\nu$ b) $W \rightarrow \mu\nu$ channel

Table 12.3: 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

645 12.3.3 Systematic uncertainty estimation

646 Results on a hadron scale factors and its errors are shown in a Table 12.3. The results are consistent
647 within 1 sigma.

648 12.4 Summary on hadronic recoil systematics

649

Chapter 13

Background estimation

651

13.1 QCD background estimation	70
13.1.1 Template selection	71
13.1.2 Methodology of the template sample normalization	71
13.1.3 Systematic Uncertainty from the Multi-jet Background Estimation	73
13.2 Background-subtracted W and Z candidate events	75

After the event selection described in a chapter 10 the background contribution is around 4% for W-analysis and 0.2% for Z analysis (which is with this statistics is negligible). Main backgrounds for W analysis are coming from:

- Processes with τ lepton, misidentified as a electron or muon + missing energy from neutrino
- Z decays with one missing lepton.
- QCD processes. In electron channel this is mostly coming from jets faking electrons, while in a muon channel it consists mostly 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 13.1. QCD background is estimated using data driven method.

Table 13.1: Background processes with their associated cross sections and uncertainties. The quoted cross sections are used to normalise estimates of expected number of events

Process	$\sigma \cdot BR$ [pb]	Order
$W^+ \rightarrow l\nu$	2116(1.9)	NNLO
$W^- \rightarrow l\nu$	1267(1.0)	NNLO
$Z \rightarrow ll$	303(0.2)	NNLO
$Z \rightarrow \tau\tau$	303	LO
$t\bar{t}$	7.41	LO
WW	0.6	LO
ZZ	0.7	LO
WZ	0.2	LO

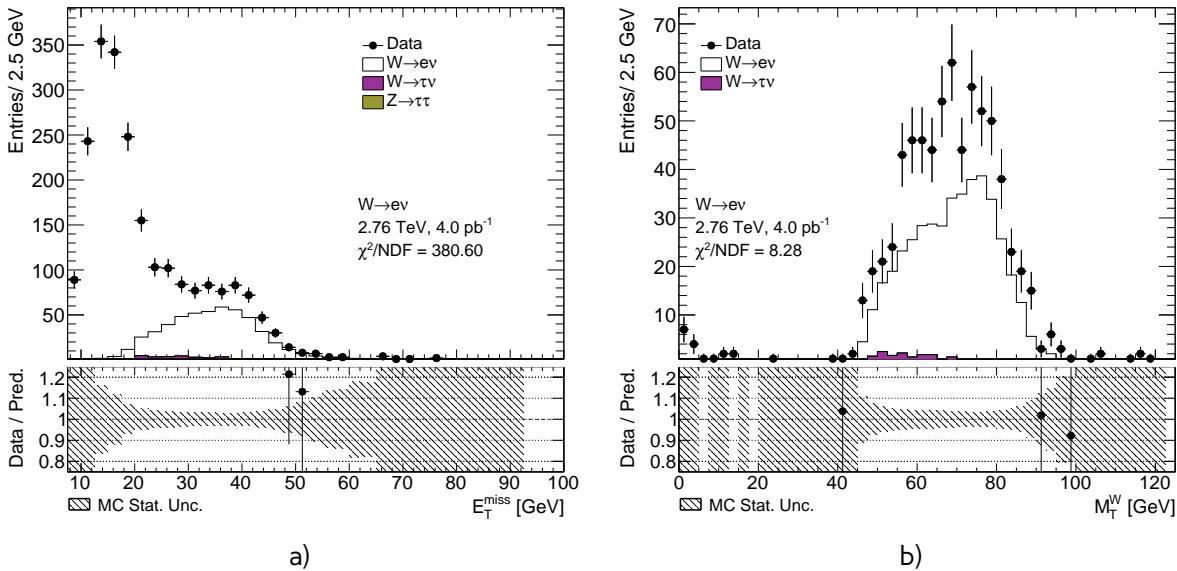


Fig. 13.1: Distribution for a) E_T^{miss} b) M_T^W template selection for $W \rightarrow e\nu$ events

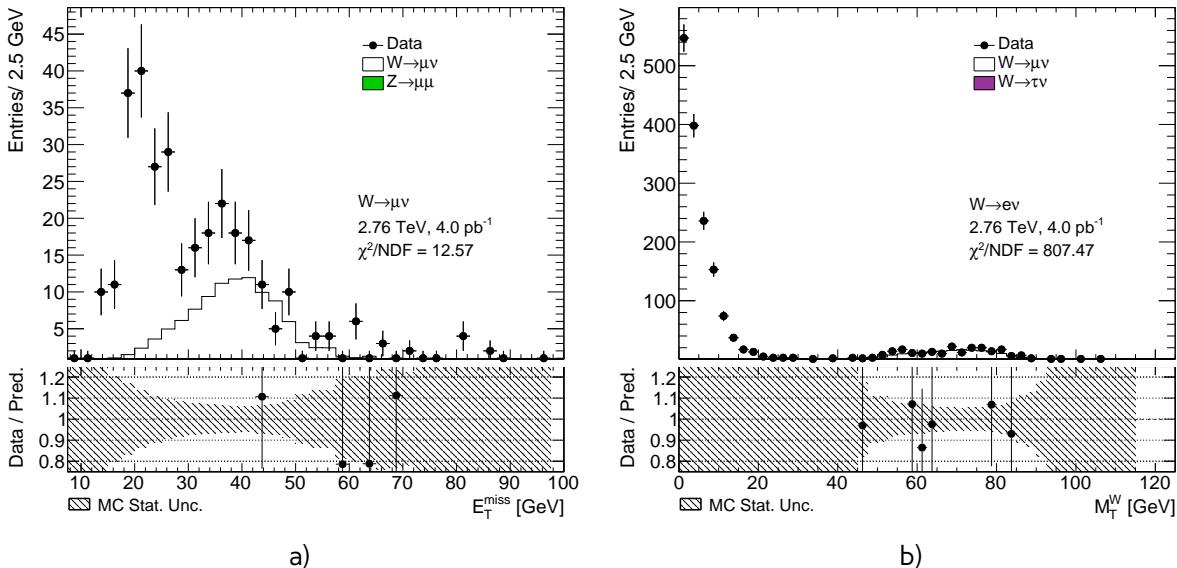


Fig. 13.2: Distribution for a) E_T^{miss} b) M_T^W template selection for $W \rightarrow \mu\nu$ events

13.1 QCD background estimation

There is a small probability, that 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 the 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 of the QCD background in a Z sample is negligible, so it is estimated just for a $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ processes.

Data driven method allows to have model independent predictions with small statistical uncertainty. This method is using QCD enriched region, where signal events are suppressed. This is usually done

680 by reversing identification or isolation criteria. It is assumed, that shape of the qcd background is
 681 staying the same in the signal region. Normalization can be derived in a control region through the
 682 template fit.

683 This section describes method of QCD background determination, that have been used in 2.76
 684 TeV data.

685 13.1.1 Template selection

686 A study have been performed to determine appropriate template selection. Because of the origins
 687 of the QCD backgrounds, missing transverse energy E_T^{miss} should be smaller in a QCD, than in a signal
 688 region. Releasing E_T^{miss} cut is allowing to gain a bigger statistics for a QCD template. It is also possible
 689 to release M_T^W cut. Most of the multijet background event should contribute than in small M_T^W
 690 region. The template sample can have a contributions from other backgrounds (mostly coming from
 691 $W \rightarrow l\nu$). Best template selection is allowing to have big data statistics and small EWK background
 692 contribution at the same time. In order to supress signal additionally reversed ID or Isolation criteria
 693 is applied.

694 For electron flavor, template is requiring for electron candidate to fail Medium isolation criteria,
 695 but to pass loose selection. Control distribution for a different template selection in electron channel
 696 are shown on a Fig. 13.1. Released E_T^{miss} cut is allowing to have a better template statistics.

697 In a muon channel template selection build by inverting isolation criteria ($PtCone20 > 0.1$). In case
 698 of $W \rightarrow \mu\nu$ the qcd background template can be achieved by releasing M_T^W cut (Fig. 13.2).

In order to avoid double counting, EWK backgrounds are substracted from a template. The total
 number of events in the template can be defined as:

$$N_{template} = N_{data}^{bkg\ enriched} - \sum_j^{MC} N_{MC_j}^{bkg\ enriched}, \quad (13.1)$$

699 where $N_{data}^{bkg\ enriched}$ and $N_{MC_j}^{bkg\ enriched}$ are number of the events in a background enriched sample in data
 700 and MC respectivelly. The resulting template statistic is 1348 and 1509 events for $W \rightarrow e\nu$ and
 701 $W \rightarrow \mu\nu$ respectively.

702 13.1.2 Methodology of the template sample normalization

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

$$M(x) = \sum_{i=1}^{N-1} f_i F_i(x) + (1 - \sum_{i=1}^{N-1} f_i) \cdot F_{qcd}(x), \quad (13.2)$$

703 where x is a fit variable (E_T^{miss} and M_T^W for $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ respectivelly), $F_i(x)$ and $F_{qcd}(x)$
 704 are the probability density functions of MC samples and QCD background template respectivelly. Fit
 705 parameters f_i are the fractions of MC within fit region. In order to eliminate systematics, coming
 706 from cross-section uncertainty, with signal fractions are left freely and and background MC fractions
 707 are varied within 5% uncertanty.

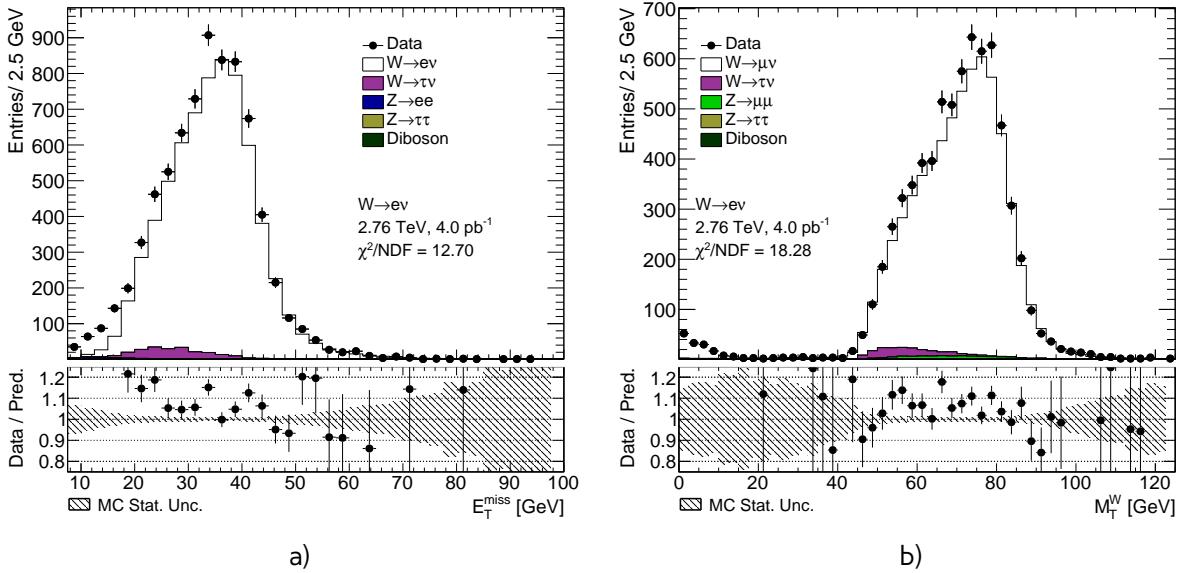


Fig. 13.3: Distributions used for multijet background estimation for a) $W \rightarrow e\nu$ b) $W \rightarrow \mu\nu$

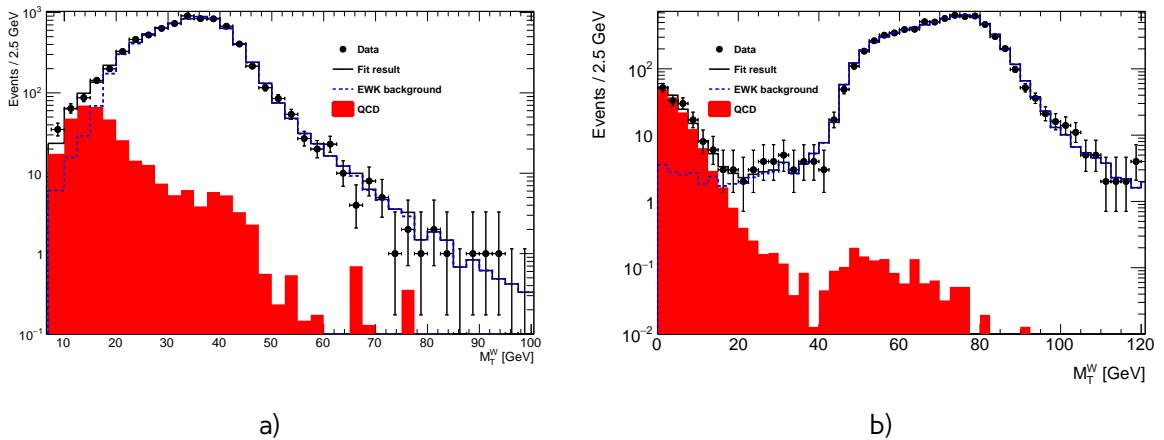


Fig. 13.4: The multijet background estimation for a) $W \rightarrow e\nu$ using reversed ID cut and released E_T^{miss} cut b) $W \rightarrow \mu\nu$ using released M_T^W cut and $bb\bar{b} + cc\bar{c}$ template

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

$$scale = \frac{(1 - \sum f_i) \cdot N_{Data}^{fit}}{N_{template}}, \quad (13.3)$$

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 uncharged W channel is used as a cross-check of the fit. The results of the fitting procedure are shown on a Fig. 13.4 .

Table 13.2: 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

712 13.1.3 Systematic Uncertainty from the Multi-jet Background Estimation

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

$$\delta_{QCD} = \sqrt{\delta_{fit\,unc}^2 + \delta_{MC}^2 + \delta_{fit\,bias}^2 + \delta_{template}^2}, \quad (13.4)$$

713 where $\delta_{fit\,unc}$ is the uncertainty for a scale from a χ^2 fit.

The second component δ_{MC} is coming from a possible mismodelling of MC in a fitted region. It can be estimated by comparison of separate 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 to have:

$$N_{QCD}^W = 0.5 \cdot N_{QCD}^{W^+} = N_{QCD}^{W^-} \quad (13.5)$$

714 Standard deviation of this 3 numbers is taken as systematic uncertainty. Since in $W \rightarrow \mu\nu$ channel
715 QCD template is fitted in a region without any additional EWK background this component is 0.

716 Third uncertainty is due to a potential bias in the template as a result of the template choise
717 and a template statistics itself. For estimation of this uncertainty different template selections have
718 been used. For $W \rightarrow e\nu$ channel different reversed isolation criteria have been tried (Fig. 13.5 a)).
719 The overall discrepancies can be considered negligible. For $W \rightarrow \mu\nu$ channel template variations
720 are estimated using fits with $b\bar{b} + c\bar{c}$ MC samples. Fig. 13.5 b) compares data template with template
721 obtained using signal selection with released M_T^W cut and template selection. Results for a different
722 template fits are presented in a Tab 13.3

723 The third component $\delta_{fit\,bias}$ is coming from an effect from arbitrary choise of bin size . This
724 error is estimated by repeating fit for a different binnings. This component is assumed negligible in
725 $W \rightarrow \mu\nu$ case, because of the small number of events.

726 Results of QCD background uncertainty estimation for $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ are shown in
727 a Tab. 13.2 and 13.3 respectively. The overall number of QCD background events is estimated as
728 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 $W^- \rightarrow \mu^-\nu$. The
729 overall fraction of QCD events is lower, than in 7 TeV data <reference to a 7 TeV paper>, what is
730 agreeing with expectations.

Table 13.3: Results of QCD background estimation for $W \rightarrow \mu\nu$ using different templates and its 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

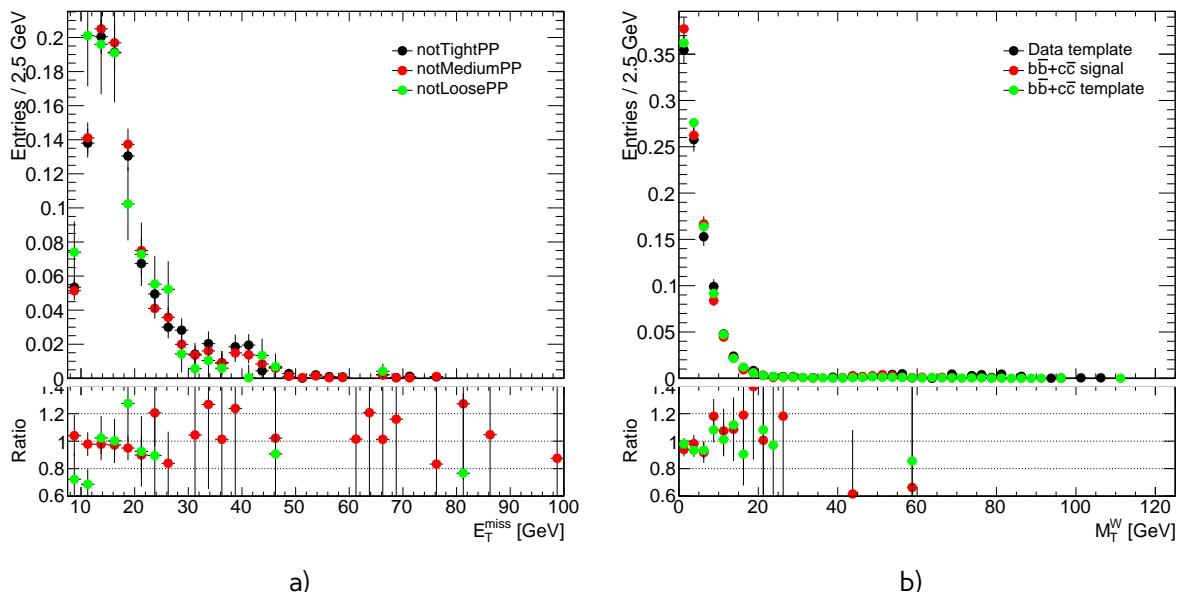


Fig. 13.5: Data and MC comparison for E_T^{miss} calculated by standard ATLAS algorithm for a) $W \rightarrow e\nu$
b) $W \rightarrow \mu\nu$ events

Table 13.4: Numbers of observed candidate events for the $W \rightarrow l\nu$ channel, electroweak (EW) plus top, and data- derived QCD background events, and background-subtracted signal events

I	Observed candidates	Background (EWK + top)	Background (Multijet)	Background-subtracted data N_W^{sig}
e^+	3914	108.1 ± 4.6	31.00 ± 13.0	$3774.9 \pm 62.6 \pm 4.6 \pm 13.0$
e^-	2209	74.2 ± 2.4	31.00 ± 13.0	$2103.8 \pm 47.0 \pm 2.4 \pm 13.0$
μ^+	4365	150.5 ± 4.8	1.50 ± 0.9	$4213.0 \pm 66.1 \pm 4.8 \pm 0.9$
μ^-	2460	105.6 ± 2.6	1.50 ± 0.9	$2352.9 \pm 49.6 \pm 2.6 \pm 0.9$

Table 13.5: Numbers of observed candidate events for the $Z \rightarrow ll$ channel, electroweak (EW) plus top and background-subtracted signal events

I	Observed candidates	Background (EWK + top)	Background-subtracted data N_Z^{sig}
e	430	1.2 ± 0.0	$428.8 \pm 20.7 \pm 0.0$
μ	646	1.5 ± 0.0	$644.5 \pm 25.4 \pm 0.0$

731 13.2 Background-subtracted W and Z candidate events

732 Tables 13.4 and 13.5 summarize the number of background events for W and Z selections respectiv-
 733 ely. Uncertainties on a number of EWK+top events are coming from a statistics, cross-section uncer-
 734 tainty (if given) and 3% of luminosity determination uncertainty. For multijet background uncertainty
 735 is coming from a method and described in a subsection 13.1.3. For the background-subtracted events
 736 the statistical uncertainty is quoted first, followed by the total systematic uncertainty, derived from
 737 the EW+top and multijet bacgrkound ones, considering the sources as uncorrelated.

Chapter 14

Control distributions

740 Kinematic distribution after all cuts (Section ??) and corrections applied on MC (Section ??), are
 741 presented in this chapter. Distributions for $W \rightarrow l\nu$ are split in charge and shown on a Figs. 14.1-
 742 14.12. Distributions for $Z \rightarrow l^+l^-$ analysis are shown on a Fig. 14.13-14.17. This plots are also showing
 743 the systematic and statistical uncertainty as a shaded band. The uncertainties are including all of
 744 the sources, described in a ??, except for uncertainty coming from shape variation due to a PDF
 745 reweighing and QCD background . All of uncorrelated sources are summed in a quadrature.

746 The expected background contributions are estimated using MC simulations, apart from QCD
 747 background.

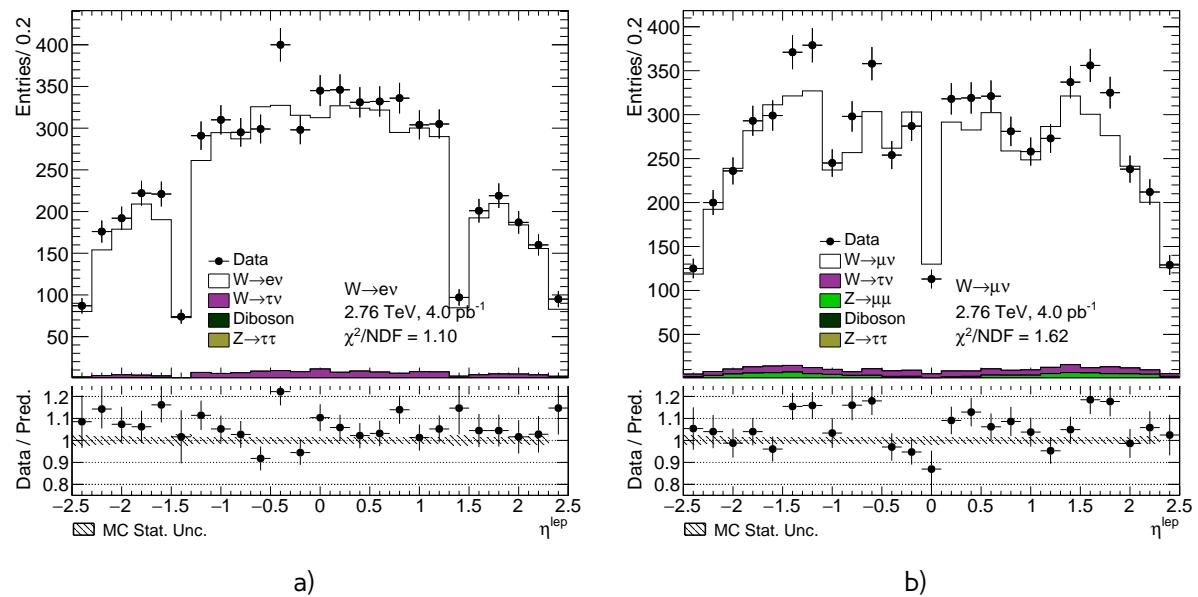


Fig. 14.1: Lepton pseudorapidity distribution from the a) $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection.

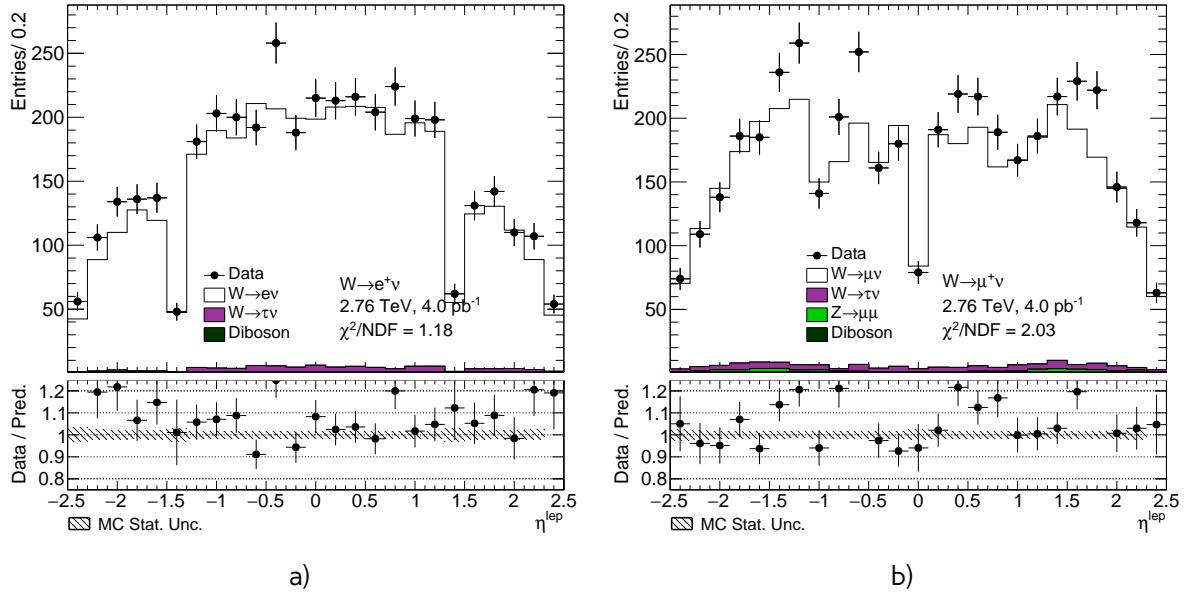


Fig. 14.2: Lepton pseudorapidity distribution from the a) $W^+ \rightarrow e^+\nu$ selection and b) the $W^+ \rightarrow \mu^+\nu$ selection.

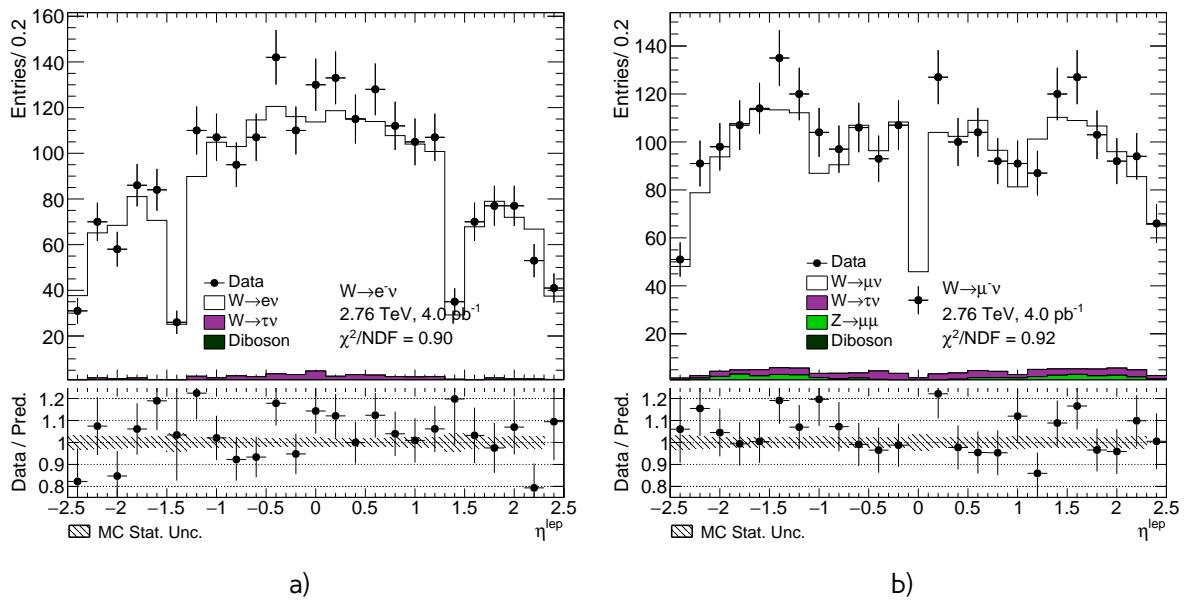


Fig. 14.3: Lepton pseudorapidity distribution from the a) $W^- \rightarrow e^-\nu$ selection and b) the $W^- \rightarrow \mu^-\nu$ selection.

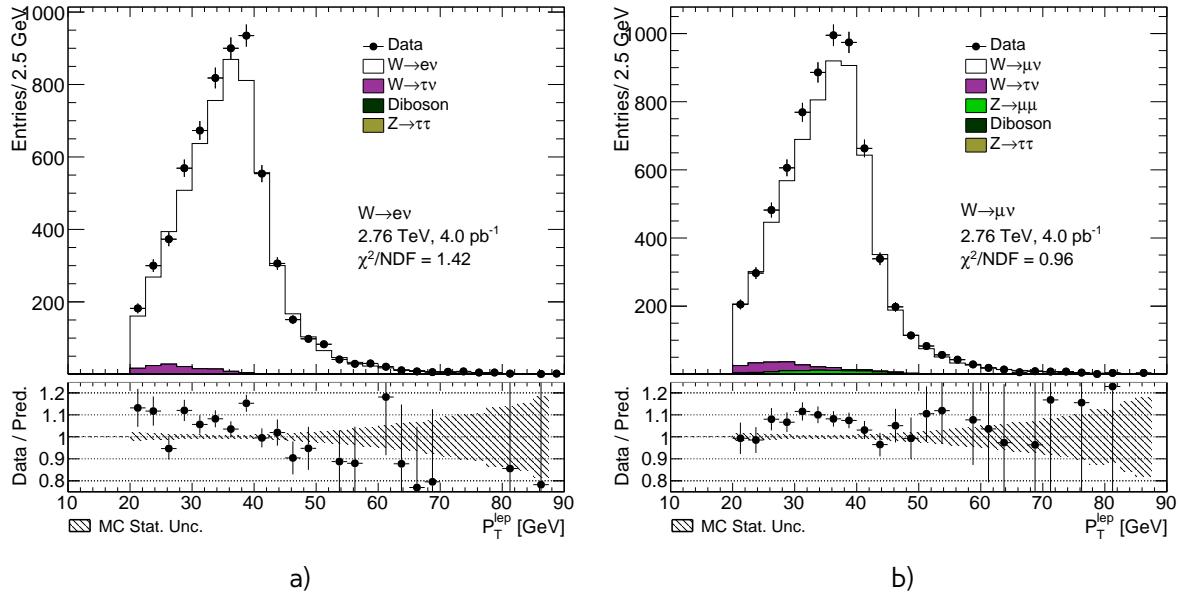


Fig. 14.4: Lepton transverse momentum distribution from the a) $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection.

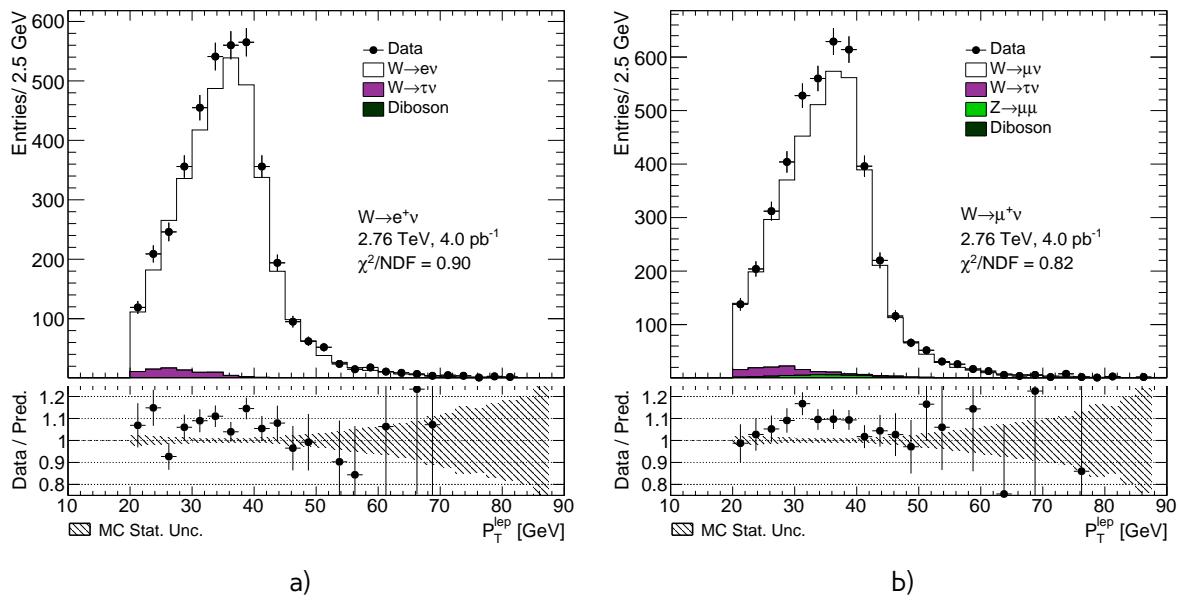


Fig. 14.5: Lepton transverse momentum distribution from the a) $W^+ \rightarrow e^+\nu$ selection and b) the $W^+ \rightarrow \mu^+\nu$ selection.

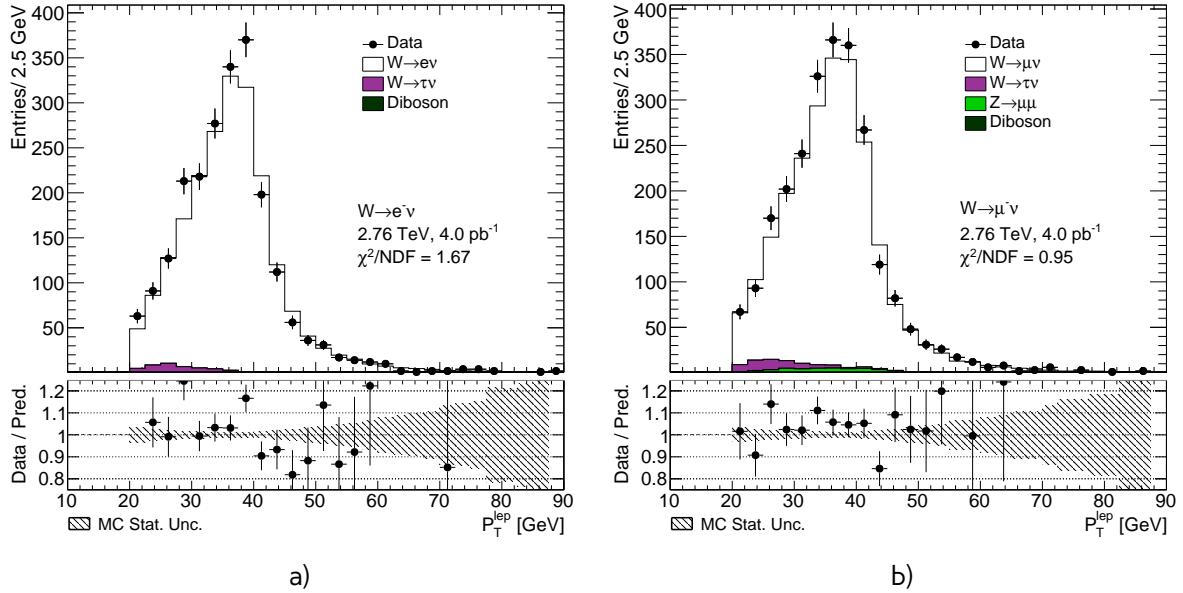


Fig. 14.6: Lepton transverse momentum distribution from the a) $W^- \rightarrow e^- \nu$ selection and b) the $W^- \rightarrow \mu^- \nu$ selection.

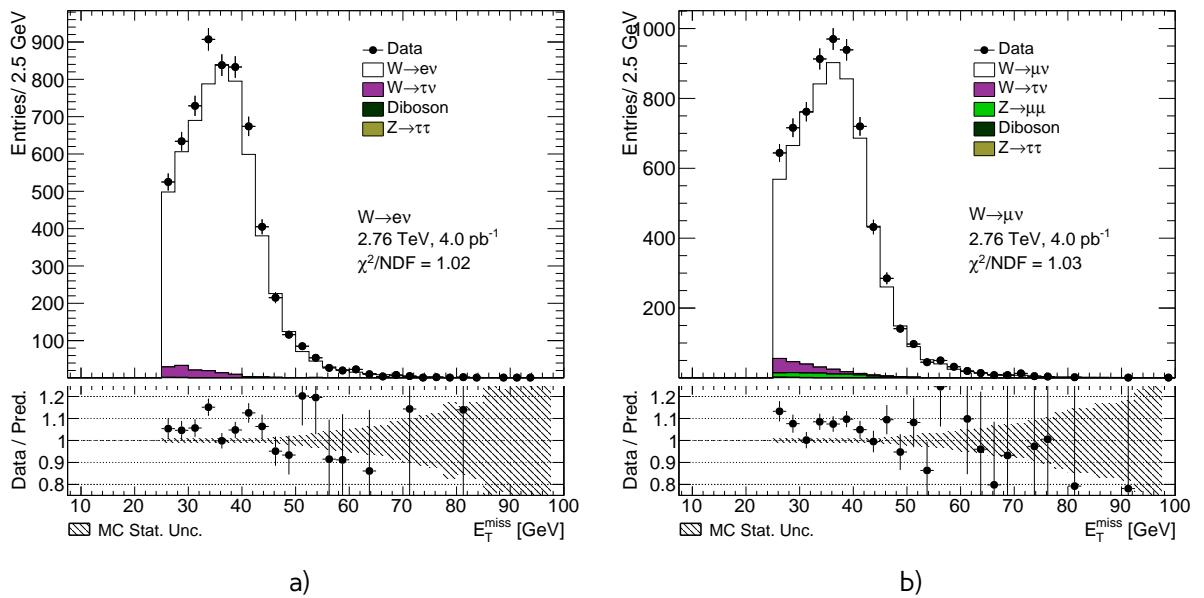


Fig. 14.7: Missing transverse energy distribution from the a) $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection.

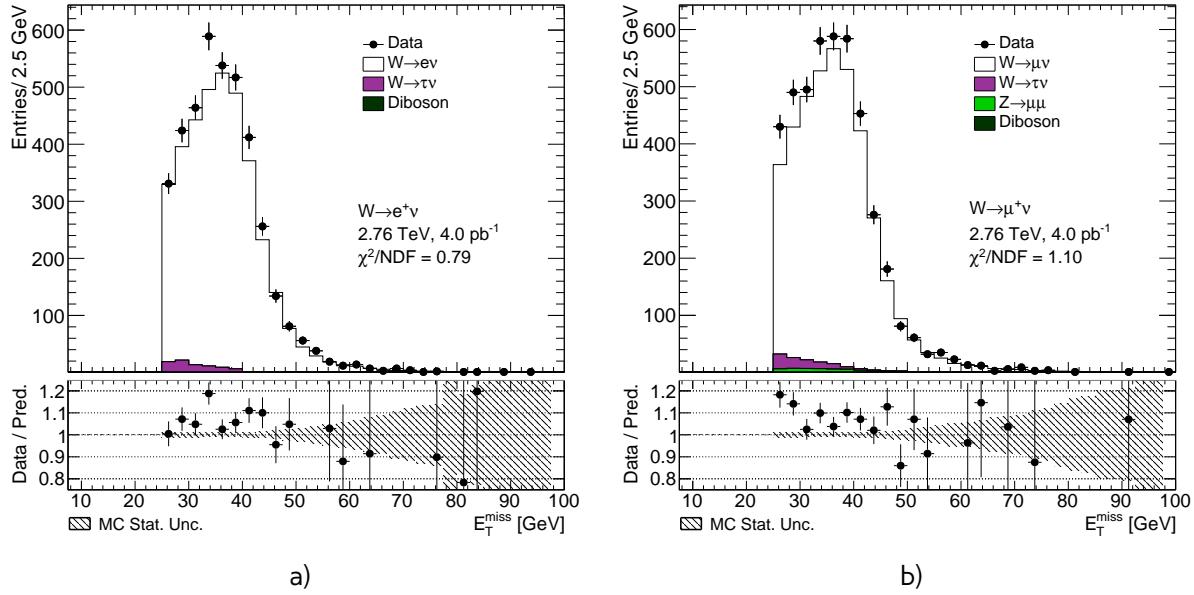


Fig. 14.8: Missing transverse energy distribution from the a) $W^+ \rightarrow e^+\nu$ selection and b) the $W^+ \rightarrow \mu^+\nu$ selection.

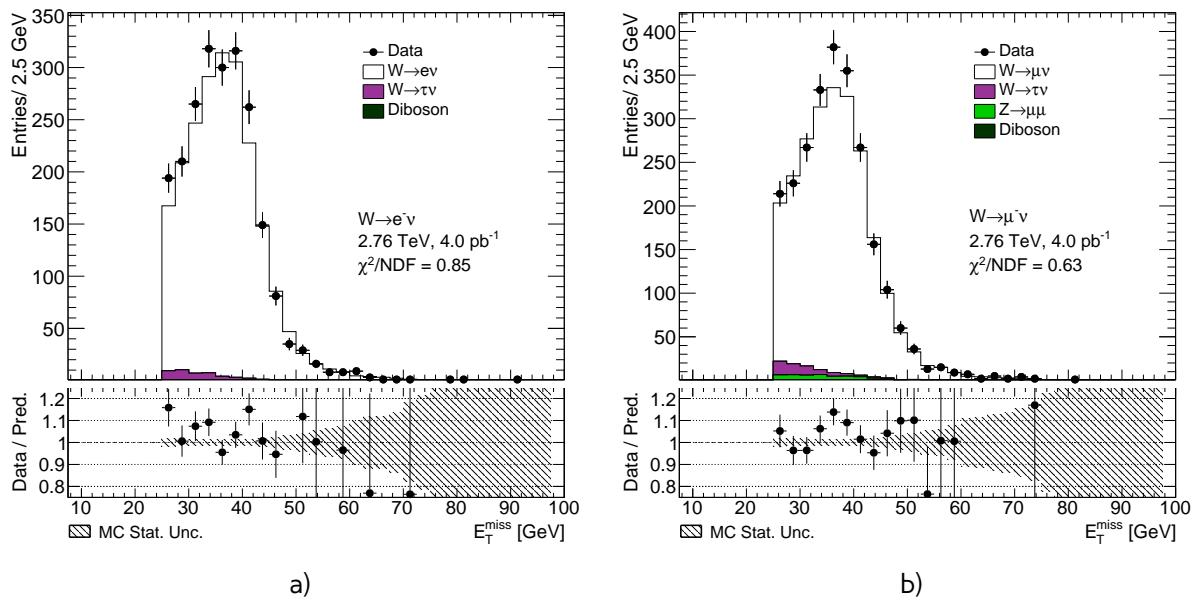


Fig. 14.9: Missing transverse energy distribution from the a) $W^- \rightarrow e^-\nu$ selection and b) the $W^- \rightarrow \mu^-\nu$ selection.

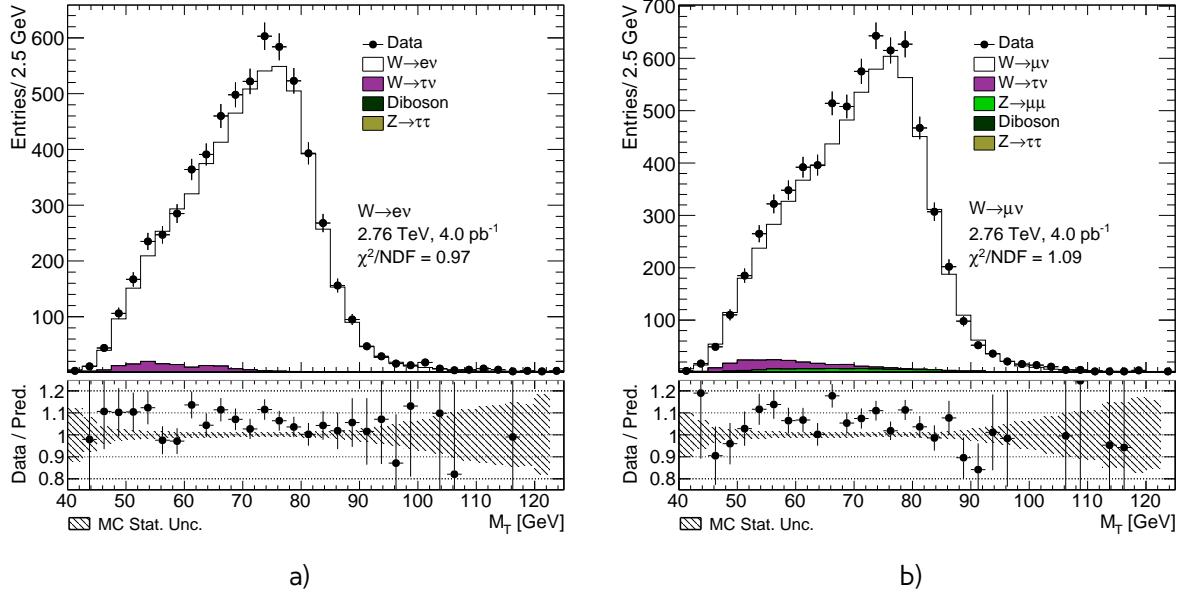


Fig. 14.10: Transverse mass distribution distribution from the a) $W \rightarrow e\nu$ selection and b) the $W \rightarrow \mu\nu$ selection.

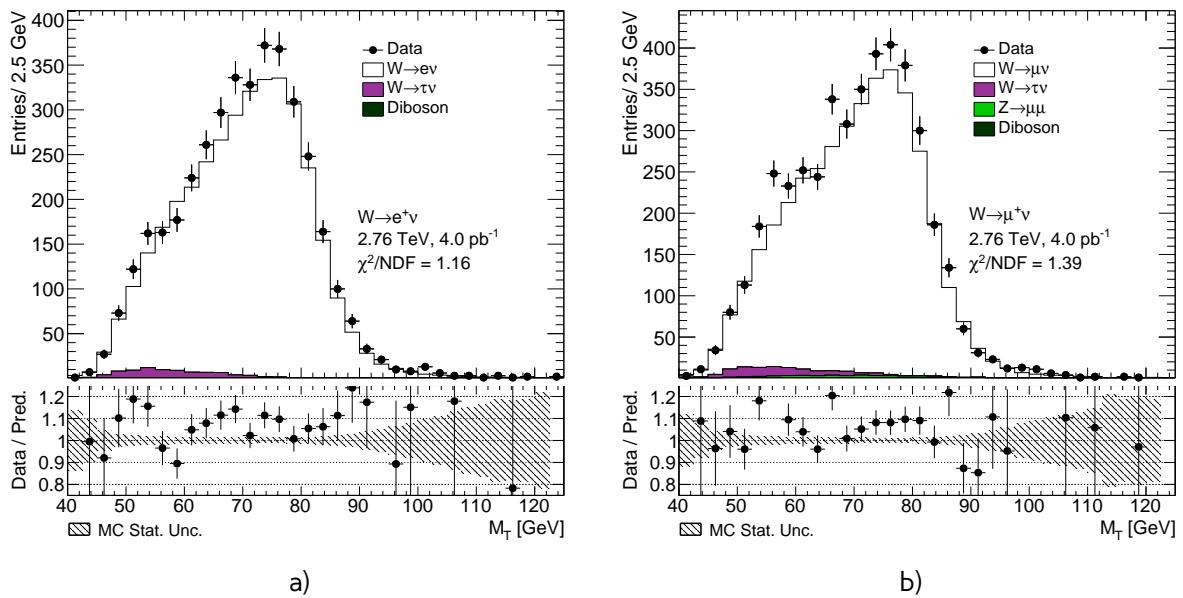


Fig. 14.11: Transverse mass distribution distribution from the a) $W^+ \rightarrow e^+\nu$ selection and b) the $W^+ \rightarrow \mu^+\nu$ selection.

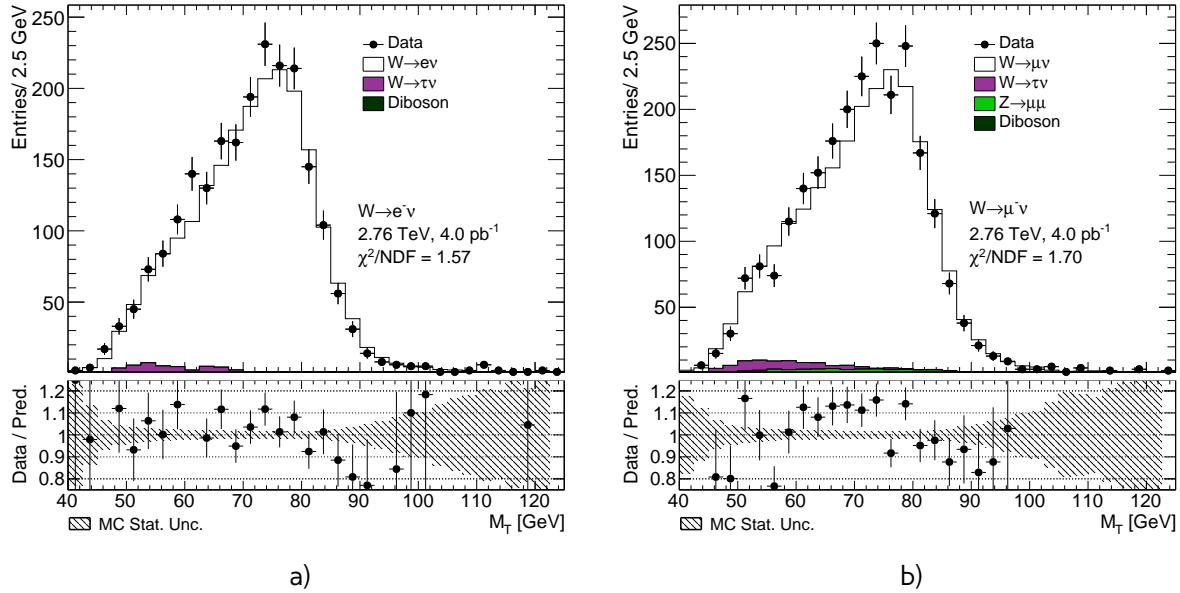


Fig. 14.12: Transverse mass distribution distribution from the a) $W^- \rightarrow e^-\nu$ selection and b) the $W^- \rightarrow \mu^-\nu$ selection.

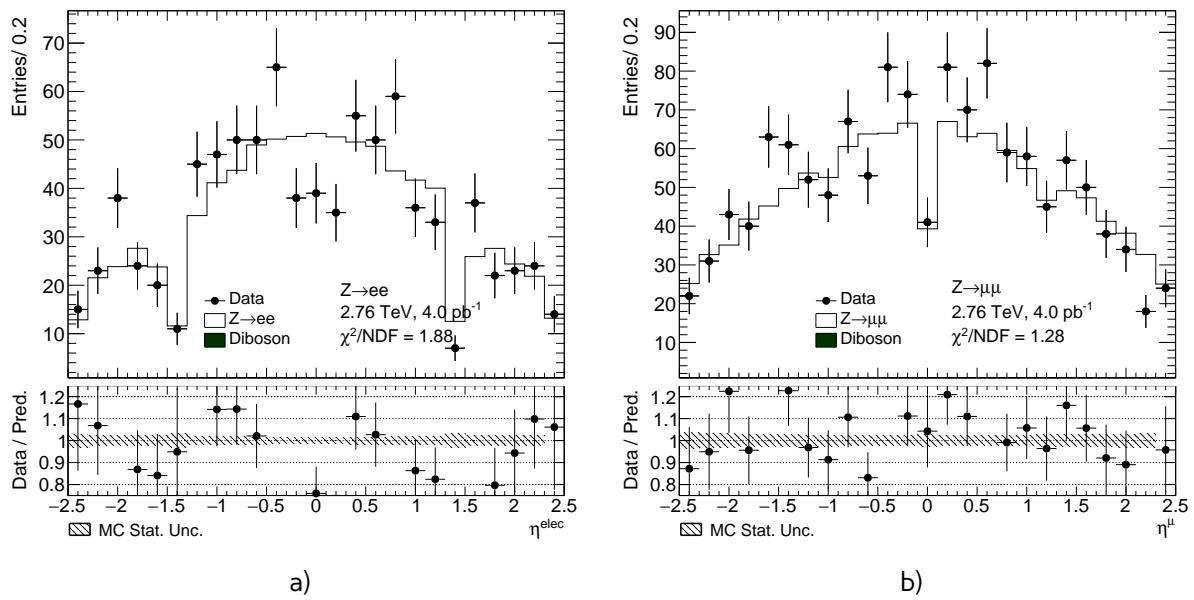


Fig. 14.13: Lepton pseudorapidity distributions from the a) $Z \rightarrow e^+e^-$ b) $Z \rightarrow \mu^+\mu^-$

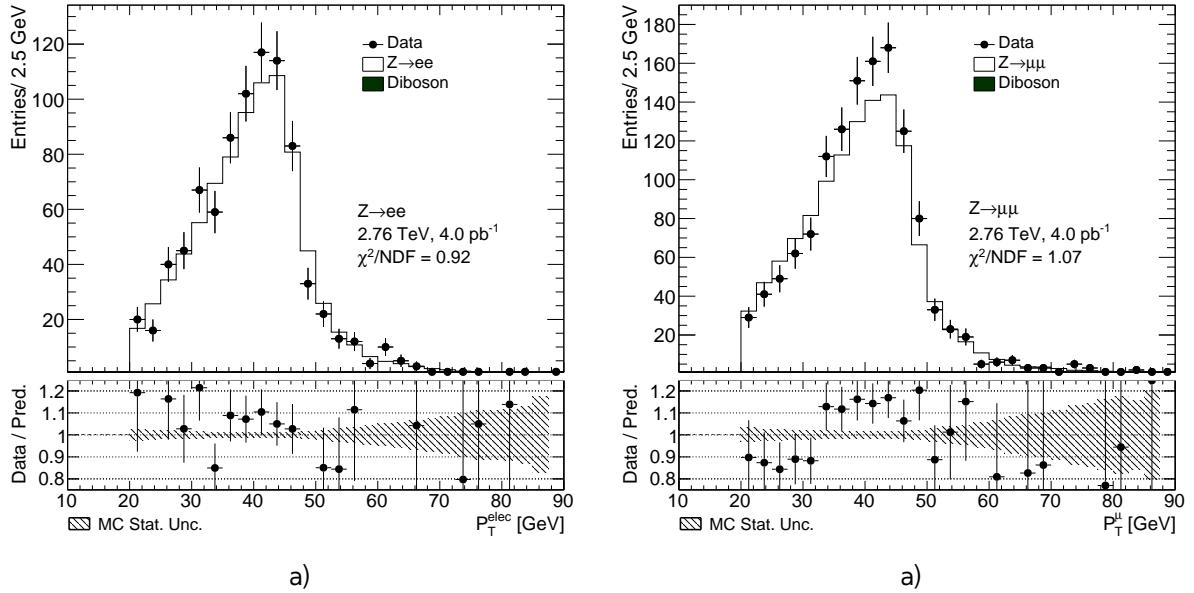


Fig. 14.14: Lepton transverse momentum distributions from the a) $Z \rightarrow e^+e^-$ b) $Z \rightarrow \mu^+\mu^-$

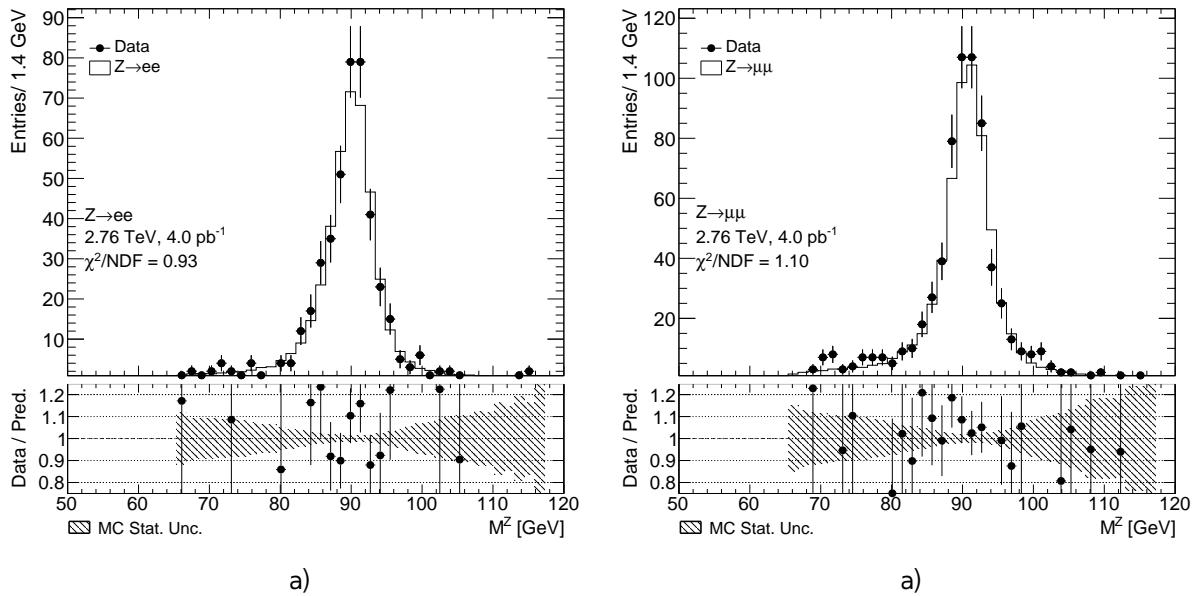


Fig. 14.15: Dilepton mass distribution distributions from the a) $Z \rightarrow e^+e^-$ b) $Z \rightarrow \mu^+\mu^-$

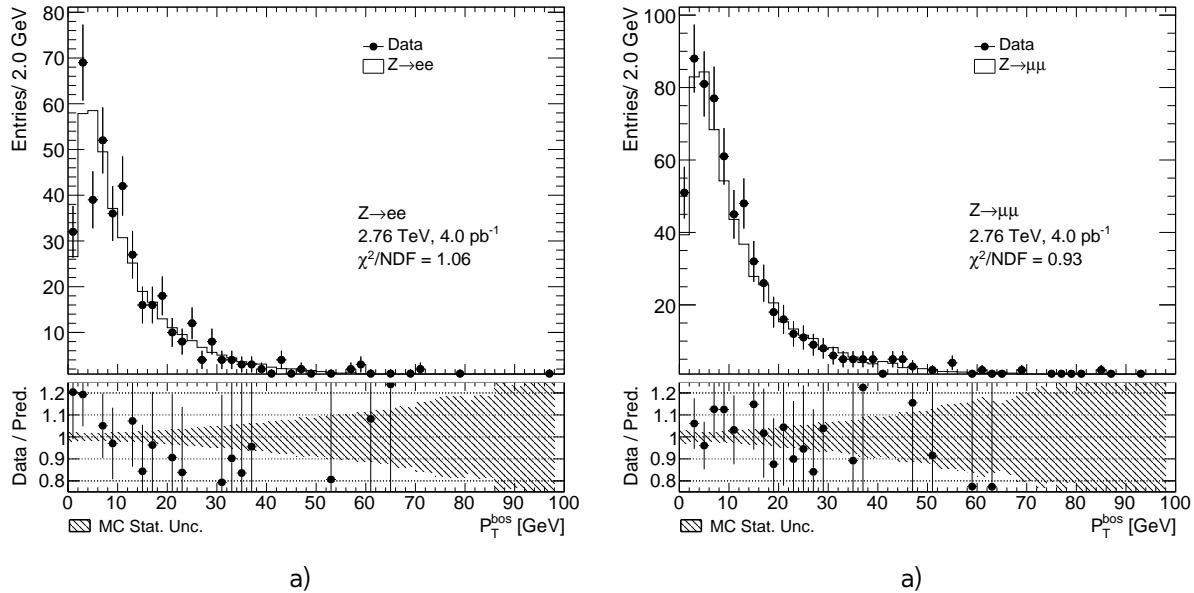


Fig. 14.16: Z boson transverse momentum distributions from the a) $Z \rightarrow e^+e^-$ b) $Z \rightarrow \mu^+\mu^-$

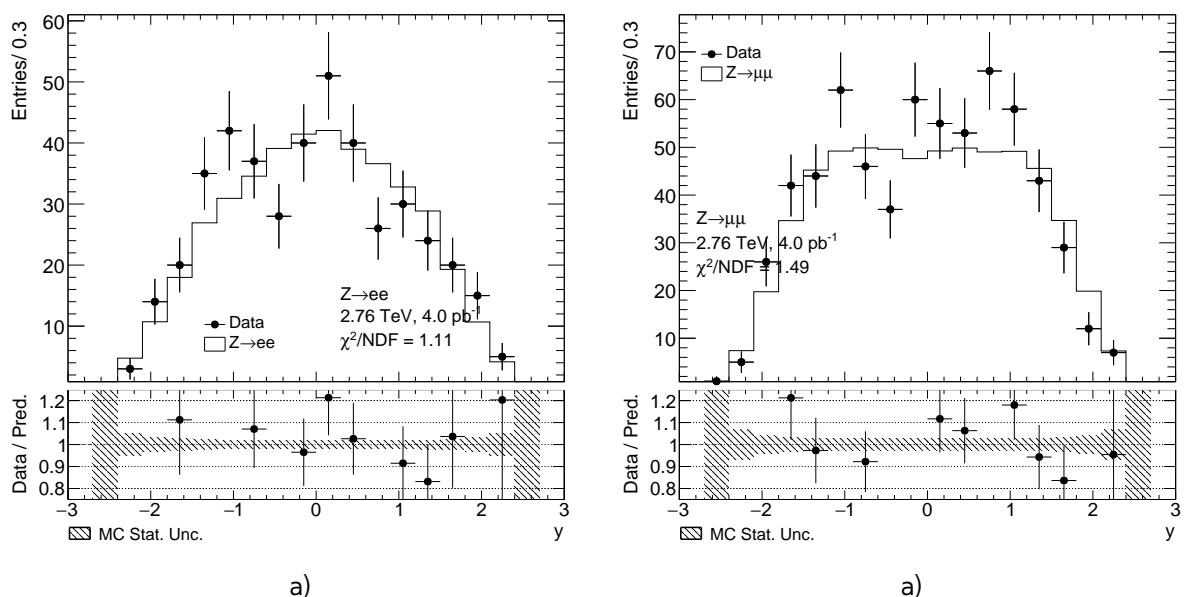


Fig. 14.17: Z boson rapidity distribution from the a) $Z \rightarrow e^+e^-$ b) $Z \rightarrow \mu^+\mu^-$

748

Chapter 15

Uncertainties in the cross-section measurement

750

751	15.1 Methods of uncertainties propagation	87
752	15.2 Experimental systematic uncertainties	88
753	15.2.1 Electron energy scale and resolution	88
754	15.2.2 Muon energy scale and resolution	88
755	15.2.3 Muon and electron efficiency toy Monte-Carlo	88
756	15.3 Theoretical uncertainty	90
757	15.4 Correlation between uncertainties	91
758	15.4.1 Toy MC correlations	91
759	15.4.2 Correlations between PDF's eigenvectors	91

760

761

762

763

Cross-section measurement relies on theoretical models and corrections, used in Monte-Carlo. Thus, their intrinsic uncertainties should be propagated to a final result. This chapter discusses main methods of uncertainties measurements and sources on $C_{W,Z}$ and $A_{W,Z}$ correction factors.

15.1 Methods of uncertainties propagation

The offset method changes a correction by a $\pm 1\sigma$ of it's 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:

$$U_i^{\text{offset}} = \frac{\sigma_i^{\text{up}} - \sigma_i^{\text{down}}}{2}, \quad (15.1)$$

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

Another method used for a uncertainties propagation is a toy MC method, that uses a pseudo experiments with modified input corrections. For a scale factors binned p_T and η uncertainties inside each bin can be divided to a 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:

$$SF_i^{\text{Toy}_n} = SF_i + \text{Gauss}(0, \Delta SF_i^{\text{uncorr+stat}}) + \sum \Delta SF_i^{\text{corr}} \cdot \text{Gauss}(0, 1), \quad (15.2)$$

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

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

$$U_i = \sqrt{\frac{\sum_{Toy_n=1}^N \sigma_i^2}{N} + \frac{\sum_{Toy_n=1}^N \sigma_i^2}{N}} \quad (15.3)$$

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

773 **15.2 Experimental systematic uncertainties**

774 Sources of experimental uncertainties, methods of estimation and their effect on a $C_{W,Z}$ are sum-
775 marized in a Tab. 15.2. Systematical errors coming from a hadron recoil calculation are discussed in
776 a Sec. 12.

777 **15.2.1 Electron energy scale and resolution**

778 Electron energy scale correction, described in Sec. 11.2 has associated uncertainties coming from
779 <reference>:

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

785 The uncertainty contribution from each component is estimated using offset method. The total
786 energy scale uncertainty is the quadratic sum of the components <reference>.

787 **15.2.2 Muon energy scale and resolution**

- 788 • MS modelling
- 789 • ID modelling
- 790 • overall scale

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

793 **15.2.3 Muon and electron efficiency toy Monte-Carlo**

794 In case of 2.76 TeV analysis scale factor errors are considered to be enlarged for a statistical and
795 uncorrelated components, so correlated error is assumed to be negligible. The toy MC experiments
796 are performed for electron reconstruction, identification and trigger scale factors and muon recon-
797 struction + identification scale factors. In the current analysis 30 pseudo-experiments are used with
798 a combined toy MC method. Plots for C_W effect on a cross-section. Correlation

Table 15.1

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	-	-	0.00	-
Muon reco+id	Toy MC	-	-	0.03	0.02	-	0.03
Muon trigger	Offset	-	-	1.01	1.01	-	0.02
Electron energy scale	Offset			-	-		-
- Statistical error	Offset			-	-		-
- Bias in method	Offset			-	-		-
- Scale uncertainty	Offset			-	-		-
- Presampler energy scale	Offset			-	-		-
- Material knowledge	Offset			-	-		-
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	-	-	Wplusmumu Wminmumu		-	Zmumu
- 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					-	-
Hadron recoil resolution	Offset					-	-
EWK + $t\bar{t}$ background							
QCD					-	-	
PDF error							
Total							
Statistics							
Luminosity							

Table 15.2

Source of uncertainty	Method	$\delta C_W/C_W(\%)$ $W^+ \rightarrow e\nu$	$\delta C_W/C_W(\%)$ $W^- \rightarrow e\nu$	$\delta C_W/C_W(\%)$ $W^+ \rightarrow \mu\nu$	$\delta C_W/C_W(\%)$ $W^+ \rightarrow \mu\nu$
Electron reconstruction	Toy MC	0.11	0.09	-	-
Electron identification	Toy MC	0.32	0.30	-	-
Electron trigger efficiency	Toy MC	0.14	0.13	-	-
Muon reco+id	Toy MC	-	-	0.03	0.02
Muon trigger	Offset	-	-	0.03	0.02
Electron energy scale	Offset	Wplusenu		Wminenu	
Electron energy resolution	Offset	0.05	0.03	-	-
Muon energy scale	Offset	-	-	Wplusmunu	
Muon energy resolution	Offset	-	-	Wplusmunu	
Hadron recoil scale	Offset	Wplusenu		Wminenu	
Hadron recoil resolution	Offset	Wplusenu		Wplusmunu	
Total					

Source of uncertainty	Method	$\delta C_Z/C_Z(\%)$ $Z \rightarrow ee$	$\delta C_Z/C_Z(\%)$ $Z \rightarrow \mu\mu$
Electron reconstruction	Toy MC	0.12	Zmumu
Electron identification	Toy MC	Zee	Zmumu
Electron trigger efficiency	Toy MC	0.00	Zmumu
Muon reco+id	Toy MC	Zee	Zmumu
Muon trigger	Offset	Zee	Zmumu
Electron energy scale	Offset	Zee	Zmumu
Electron energy resolution	Offset	0.03	Zmumu
Muon energy scale	Offset	Zee	Zmumu
Muon energy resolution	Offset	Zee	Zmumu
Total			

15.3 Theoretical uncertainty

- Theoretical uncertainties on the predictions are mostly dominated by a imperfect knowledge of the proton PDF's. They are affecting both $A_{W,Z}$ and $C_{W,Z}$. Error coming from an arbitary choice of PDF set is estimated by PDF reweighting <reference> of original MC generated using <something> to a one of the 4 pdf sets: CT10 <reference>, ATLAS-epWZ12 <reference>, abkm09 <reference> and NNPDF23 <reference>. The error is calculated as a standard deviance for all of the sets.
- Systematic uncertainty within one pdf set is evaluated using CT10 NLO set. This set contains 52 asosiated error sets, corresponding to a 90% C.L. limits along 26 eigenvectors. The resulting

PDF Set	$\delta A_W / A_W (\%)$	$\delta C_W / C_W (\%)$	$\delta A_W / A_W (\%)$	$\delta C_W / C_W (\%)$
	$W^+ \rightarrow e\nu$			$W^- \rightarrow e\nu$
CT10				
ATLAS-epWZ12				
abkm09				
NNPDF23				
	$W^+ \rightarrow \mu\nu$			$W^- \rightarrow \mu\nu$
CT10				
ATLAS-epWZ12				
abkm09				
NNPDF23				

52 variation are separately added in a quadrature as:

$$\delta_X = \frac{1}{2} \cdot \sqrt{\sum_{i=1}^N (X^+ - X^-)^2} \quad (15.4)$$

- The uncertainties arising from the choice of generator and parton showering model are considered small. They can be calculated as a difference in the acceptances $A_{W,Z}$ for MC samples, generated using same PDF set, but different models for showering and matrix element, namely Powheg + Pythia and Sherpa

15.4 Correlation between uncertainties

15.4.1 Toy MC correlations

A correlation coefficient between two observables o_1 and o_2 can be estimated as:

$$\rho_{12} = \frac{1}{\sigma(o_1)\sigma(o_2)} \cdot \frac{1}{N-1} \sum_{i=1}^N (o_1^i - \bar{o}_1)(o_2^i - \bar{o}_2) \quad (15.5)$$

Using cholesky transformation this uncertainty can be propagated to a 3 eigenvectors:

15.4.2 Correlations between PDF's eigenvectors

Correlations for $A_{W,Z}$ and $C_{W,Z}$ for a CT10nlo set have been estimated, since they could affect error on a total W boson cross section measurement and PDF fits. Given two processes X and Y (

$$\delta_{XY}^2 = \delta_X^2 + \delta_Y^2 + 2\delta_X\delta_Y\rho_{XY} \quad (15.6)$$

$$\rho_{XY} = \frac{1}{4\delta_X\delta_Y} \sum (X^+ - X^-) \cdot (Y^+ - Y^-) \quad (15.7)$$

814

Part V

815

Results

816

Chapter 16

817 Results of the Cross Section Measurement

818 16.1 Combined results

819 16.2 Comparation with Theoretical Predictions

820

Chapter 17

821

PDF fits results

822 Chapter **18**

823 **Summary**