UNIVERSITÉ PARIS DIDEROT

THÈSE

en vue d'obtenir le grade de

DOCTEUR ÈS SCIENCES

présentée par

NGUYEN Thi Hong Van

Mesure la distribution en impulsion transverse du boson W avec le detecteur ATLAS au LHC

devant être soutenue le 29 septembre 2011 devant la commission d'examen composée de

VANNUCCI François	président
OURAOU Ahmimed	directeur de thèse
BOONEKAMP Maarten	co-directeur de thèse
AURENCHE Patrick	rapporteur
FROIDEVAUX Daniel	rapporteur
SCHOTT Matthias	examinateur
HAMEL-de-MONCHENAULT Gautier	examinateur

Contents

Acknowledgements

I cordially express my gratitude to my supervisors Maarten Boonekamp and Nathalie Besson for their great help, constant arrangement and for all that I have learned from them during the whole period of my study. I am grateful to Patrick Aurenche for his continuous and strong support.

It is a great pleasure for me to thank Ursula Bassler, Bruno Mansoulie, Claude Guyot, Ahmimed Ouraou and the whole collective of the IRFU-CEA, Saclay and the university of Paris Diderot–Paris 7 for kind hospitality and excellent working conditions. I would like also to thank Marie Legendre, Anne-Isabelle Etienvre, Olivier Simard, Eve Le Menedeu, Nicolas Morange and other friends for their help and friendship during my stay in Saclay.

I'm thankful to Daniel Froidevaux, Francois Vannucci, Matthias Schott, Gautier Hamel-de-Monchenault and all members of the jury of my thesis defense for their valuable comments and discussions.

The financial and professional and technical support from CEA and CERN (including ATLAS collaboration) are highly appreciated and hereby acknowledged. In addition, it is my duty to thank the IOP, Hanoi, and the VAST-CNRS collaboration program in particle physics for great support.

I am thankful to Tran Thanh Van, Tran Kim Ngoc and Pham Xuan Yem for the warm friendship and encouragement during my stay in France.

I could not complete my study and thesis successfully without the love and support from my family, specially, from my little girl giving me much energy and courage to overcome difficult moments. I also never forget the nice time I have spent with my family in the beautiful France.

Chapter 1

Introduction

One of the prior duties at the LHC, with the first data produced, is to measure the distribution of the momentum transverse to the beam direction of electroweak gauge bosons W and Z. These measurements provide one of the most crucial tests of Standard Model, in particular, the QCD calculations for hadron colliders, which have different approaches for different ranges of the spectrum. In addition, data on W and Z production provides an important tool in searching for possible new physics.

In this thesis we describe a measurement of the W transverse momentum (p_T^W) distribution produced by proton-proton collisions at the LHC with the ATLAS detector (see Chapter 3 for a general recommendation about the LHC and the ATLAS detector). At hadron colliders, the non-zero p_T^W can be created due to the parton radiation from the initial state. Theoretically, there have been a number of methods proposed and developed to give accurate predictions for the p_T^W spectrum produced. They are summarized in Chapter 2 including theoretical background and motivation for the measurement.

Experimentally, the signature of W bosons is searched via their leptonic decay modes ($W \to \ell\nu$, $\ell = e, \mu$). Because of the neutrino in the final state which is undetectable, the p_T^W must be reconstructed as the hadronic recoil measured by summing vectorially the transverse energies deposited in the calorimeter excluding that of lepton(s). An algorithm of hadronic recoil reconstruction in W, Z events is proposed and detailed in Chapter 4. The missing transverse energy (E_{T}^{miss}) then can be calculated from the hadronic recoil by a combination with the lepton transverse momentum. This algorithm thus provides another tool for reconstructing the E_{T}^{miss} which has a number of advantages over the standard one [25].

Measuring the p_T^W spectrum using the hadronic recoil complements to measurements of the p_T^Z spectrum in which the p_T^Z is reconstructed as the momentum of the lepton pair [28]. Although the hadronic

recoil resolution is not as good as the lepton momenta resolution, the statistic of observing the $W \to \ell\nu$ events is about 10 times higher than the $Z \to \ell\ell$ events [27].

All steps and results of measuring the p_T^W spectrum are described in Chapter 5. The analysis is performed using the data produced by proton-proton collisions at $\sqrt{s} = 7$ TeV during 2010 run of the LHC, corresponding to 31 pb^{-1} of integrated luminosity. The procedure present is for the electron channel ($W \to e\nu$) but can be applied similarly for muon channel ($W \to \mu\nu$). Given the algorithm of constructing the hadronic recoil as described in Chapter 4, the hadronic recoil performance is transparent between the channels and can be calibrated from Z data. Therefore, in order to reduce the statistical uncertainty, two channels are combined to study the detector response to the recoil.

In measurements of the p_T^W spectrum using electron and muon channels, the W propagator is defined as the baseline quantity to be measured. This allows for a combination of the final results to get a better statistic. The resulting combined spectrum is then compared to theoretical predictions and measurement of the p_T^Z spectrum reconstructed from the transverse momentum of lepton pairs [28]. This part is discussed in Chapter 6. Comparing to previous p_T^W measurements, which are recently performed at Tevatron by CDF [6] and D0 [5], this measurement at LHC with higher statistics gives a better precision and allows to extend the range of p_T^W as well as apply fully the correction accounting for the bin-to-bin correlation.

Chapter 2

Theoretical background and motivation

The analysis performed in this thesis provides a measurement of transverse momentum distribution of W boson (p_T^W) produced at the LHC by proton-proton collisions at $\sqrt{s} = 7$ TeV. In order to motivate and give backgrounds for these measurements, we will give in this chapter an overview on W boson particle, and especially, theoretical calculations of p_T^W distribution. Firstly, a general summary of the Standard Model will be presented to describe how the W bosons originate and what role they play in the theory. In particular, we will talk about their participation in electroweak interactions and their production at hadron colliders. Then, we will present the p_T^W predictions at different ranges on the spectrum.

2.1 Standard Model

The Standard Model (SM) of particle physics is a quantum field theory based on $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ gauge symmetry. This theory describes the strong, the electromagnetic and the weak interactions of elementary particles in which the strong interaction is based on $SU(3)_C$ part, while the electromagnetic and weak interactions are mixed and unified into a combined interaction known as "electroweak" interaction and based on $SU(2)_L \otimes U(1)_Y$ part. The particle content of the SM consists of 12 fermions (6 quarks and 6 leptons), 4 types of "gauge" bosons including gluons (the mediators of the strong interaction), W, Z bosons (the mediators of the weak interaction) and the photon (the mediator of the electromagnetic interaction) and Higgs boson which are introduced to give mass to massive particles. These 16 particles (except the Higgs boson which has not yet been observed) and their properties such as charge, mass and spin are shown in Figure 2.1.

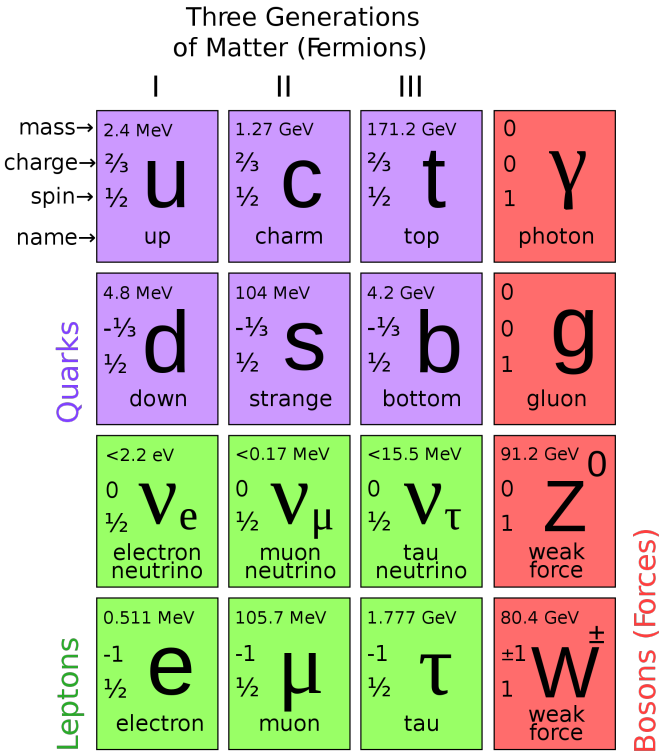
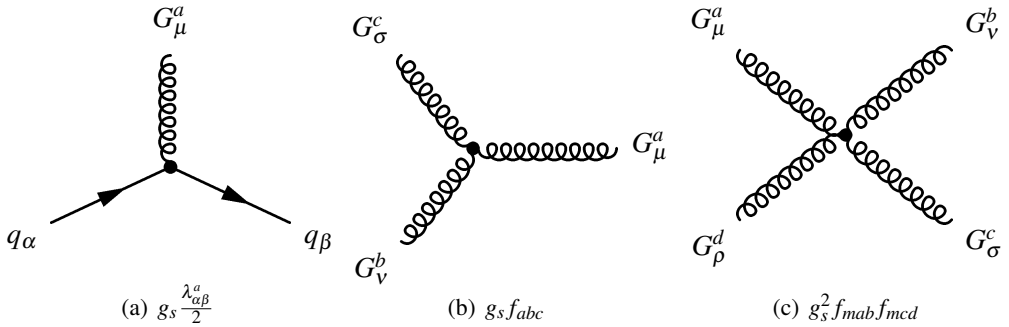


Figure 2.1: Partiles in the Standard Model and their properties [86].

2.1.1 Quantum Chromodynamics

Quantum chromodynamics (QCD) is an important part of the Standard Model. It is a quantum field theory describing the strong interaction as a result of the exchange of colored massless vector gluons between colored quarks bounded into hadrons such as protons, neutrons, pions, etc. [69] governed by the $SU(3)_C$ colored gauge symmetry [7, 53, 76]. In this theory, the colored quarks are considered as triplets in the fundamental representation of the $SU(3)_C$ group in the sense that each flavor of quark as shown in Figure 2.1 is assumed to have three degrees (colors) of freedom labeled by R, G and B (stand for "Red", "Green" and "Blue", respectively). Gluons are components of an octet in the adjoint representation of the $SU(3)_C$ group. According to this theory, the strong interaction is described by fundamental vertices shown in Figure 2.2.



(a) $g_s \frac{\lambda^a_{\alpha\beta}}{2}$ (b) $g_s f_{abc}$ (c) $g_s^2 f_{mab} f_{mcd}$

Figure 2.2: Color factors of the interaction vertices in QCD: q_α, q_β are quark fields with colored index $\alpha, \beta = R, B, G$; $\lambda^a (a = 1, 2, .., 8)$ are Gell-mann's $SU(3)$ representation matrices with constant structure f_{abc}; g_s is the strong coupling.

QCD has the two following important properties as consequences of the running coupling constant $\alpha_s = g_s^2/4\pi$ given in equation 2.1, where, Q^2 is the energy transfer squared, $n_c = 3$ is the number of colors, $n_f = 6$ is the number of quark flavors, and $\Lambda_{QCD} \sim 200\,\mathrm{MeV}/c$ is the QCD scale (taken from [75]):

$$\alpha_s(Q^2) = \frac{12\pi}{(11n_c - 2n_f)\log(Q^2/\Lambda^2_{QCD})}, \tag{2.1}$$

- *Asymptotic freedom*: equation (2.1) shows that $\alpha_s(Q^2) \to 0$ as $Q^2 \to \infty$ which implies that looking at very high energy probes or short distance, quarks and gluons appear like almost free particles. This prediction of QCD was made in 1973 by D. Politzer, D. Gross and F. Wilczek and brought

the 2004 Nobel prize in physics to the authors.

- *Confinement*: another consequence of equation (2.1) is that $\alpha_s(Q^2) \to \infty$ as $Q^2 \to \Lambda^2_{QCD}$, that is quarks are always confined inside hadrons and they can never be found as free (unbounded) states.

2.1.2 Electroweak theory

The electroweak theory proposed by G. Glashow (1961) , S. Weinberg (1967) and A. Salam (1968) is based on $SU(2)_L \times U(1)_Y$ gauge symmetry . In this model, four gauge vectors associated with $SU(2)_L$ and $U(1)_Y$ groups are denoted by W^α (α = 1, 2, 3, corresponds to 3 generators of $SU(2)$ group) and B, respectively. They, in fact, are not yet real physical gauge bosons (mass eigenstates) which are appropriate mixings of them. Hereafter, we call these mass eigenstates gauge bosons. In the beginning, the local gauge invariance requires all the gauge bosons to be massless. However, the spontaneous breaking of the gauge symmetry through the Higgs mechanism (Higgs, 1966), provides mass to three of the four gauge bosons while the rest remains massless. The mass eigenstates include charged massive weak bosons $(W^\pm = W^1 \mp iW^2)/\sqrt{2}$, physical neutral massive weak boson Z and the massless photon A in which Z and A are the orthogonal linear combinations of the neutral gauge boson W^3 and B,

$$Z = W^3\ cos\theta_w - B\ sin\theta_w; \quad A = W^3\ sin\theta_w + B\ cos\theta_w, \tag{2.2}$$

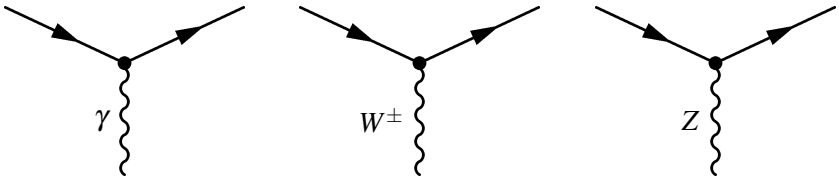
where θ_w is the electroweak mixing angle (or Weinberg angle). This angle relates the strength of the weak interactions to the electromagnetic coupling e as given in equation 2.3 by requiring the photon A to be massless

$$e = g\ sin\theta_w. \tag{2.3}$$

This mixing angle also relates the mass of the charged and neutral weak bosons as follows:

$$cos\theta_w = \frac{M_W}{M_Z}. \tag{2.4}$$

The interactions of particles in electroweak theory are described by vertices shown in Figure 2.3. As seen from this figure, all charged fermions can participate in electromagnetic interactions via photons and all fermions (including charged and neutral fermions) can participate in weak interactions via weak bosons $W^\pm$ and Z. While the charged weak bosons $W^\pm$ couple with only left-hand fermions, the neutral weak boson Z couples with both left-hand and right-hand fermions due to the electroweak mixing angle.



(a) $-iQ\gamma^{\mu}$ (b) $-i\frac{g}{\sqrt{2}}\gamma^{\mu}\frac{1-\gamma^5}{2}$ (c) $-i\frac{g}{cos\theta_w}\gamma^{\mu}\frac{c_V-c_A\gamma^5}{2}$

Figure 2.3: The gauge boson-fermion vertex factors in the electroweak theory. The factors with (γ^{μ}) are vector couplings (V) and the factors with $(\gamma^{\mu}\gamma^5)$ are axial-vector (A) couplings.

In particular, the Z coupling to a fermion is a mixture of the left-hand $V-A$ coupling of $SU(2)_L$ and the vector coupling of QED, with

$$c_V = T^3 - 2\,sin^2\theta_w Q; \quad c_A = T^3, \tag{2.5}$$

where T^3 is the third component of weak isospin, Q is the charge of fermions measured with respect to $(-e)$ where e is the electron charge. The Z couplings to the various leptons and quarks are listed in Table 2.1.

fermion	T^3	Q	c_A	c_V
ν_e, ν_μ, ν_τ	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$
e^-, μ^-, τ^-	$-\frac{1}{2}$	-1	$-\frac{1}{2}$	$-\frac{1}{2}+2sin^2\theta_w$
u, c, t	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{2}-\frac{4}{3}sin^2\theta_w$
d, s, b	$-\frac{1}{2}$	$-\frac{1}{3}$	$-\frac{1}{2}$	$-\frac{1}{2}+\frac{2}{3}sin^2\theta_w$

Table 2.1: The Z couplings in the standard electroweak theory with $sin^2\theta_w = 0.225$ [74].

The predicted branching ratios of various W and Z decay modes are listed in Table 2.2. From Figure 2.3, with a vertex factor of $-i\frac{g}{\sqrt{2}}\gamma^{\mu}\frac{1-\gamma^5}{2}$, the W boson couples with the same strength to all fermion-antifermion pairs. In Standard Model, as seen from Figure 2.1, there are 3 generations of leptons and 3 generations of quarks, and each quark flavor carries 3 colors (see Section 2.1.1). However, because the top quark is heavier than the W and Z bosons, the W^+ does not decay into $t\bar{b}$ and the Z does not decay into $t\bar{t}$. Therefore each decay mode of the W boson will have a branching ratio $\frac{1}{9}$ neglecting fermion masses. For Z boson, the branching fractions are proportional to $(c_V^2 + c_A^2)$ in which c_V and c_A are given in Table 2.1.

From Table 2.2 we see that for both W and Z, the branching fractions of decay into hadrons are

Decay mode	Branching ratio (%)
$W^+ \to \ell^+ \nu_\ell, (\ell = e, \mu, \tau)$	11.1
$W^+ \to \bar{d}u(\bar{s}c)$	33.3
$Z \to \ell^+ \ell^-, (\ell = e, \mu, \tau)$	3.4
$Z \to \nu_\ell \nu_\ell, (\ell = e, \mu, \tau)$	6.8
$Z \to q\bar{q}, (q = d, s, b)$	15.2
$Z \to q\bar{q}, (q = u, c)$	11.8

Table 2.2: Expected branching ratios of W and Z bosons decays

much higher than those into leptons. And for Z the branching fraction of decay into neutrinos is twice the branching fraction of decay into charged leptons. However, in order to reduce a huge background coming from QCD processes and to be measured more easily in detectors, the charged lepton decay modes are used as signatures of these weak bosons in experiments.

One more feature of the electroweak theory, the particle flavor is conserved (or particles are exchanged within the generations) in the interactions with photons and Z (or neutral currents). In interactions with charged weak W bosons (charged currents), the lepton flavor is conserved while the quark flavor is not. W bosons couple to

$$\begin{pmatrix} u \\ d' \end{pmatrix}, \begin{pmatrix} c \\ s' \end{pmatrix}, \begin{pmatrix} t \\ b' \end{pmatrix}, \tag{2.6}$$

where d', s' and b' are linear combinations of the quark mass eigenstates d, s and b via a unitary 3×3 mixing matrix known as Cabibbo-Kobayashi-Maskawa matrix (CKM matrix). The CKM matrix suggested by M. Kobayashima and T. Maskawa is a generalization of the Cabibbo matrix (introduced by N. Cabbibo for two generations of quarks [15]) to describe the weak interactions of three generations of quarks [62]. It is specified by three real parameters and one phase angle which helps in understanding the CP violation in electroweak interactions. The elements of the CKM matrix describe the relative probability amplitudes for transitions between quarks in interactions of charged currents with charged weak W bosons. Their magnitudes measured experimentally are given in [72].

2.2 Predictions of p_T^W distribution at hadron colliders

In the parton view, hadrons are considered as clusters of confined partons that is quarks, anti-quarks and gluons. It is thus possible to describe hadron reactions in terms of the interactions of these parton constituents. With $\sqrt{s} = 7$ TeV proton-proton collisions at the LHC, W boson can be produced in the Drell-Yan process [37] in which a quark and an antiquark annihilate to produce a lepton pair (see Figure 2.4).

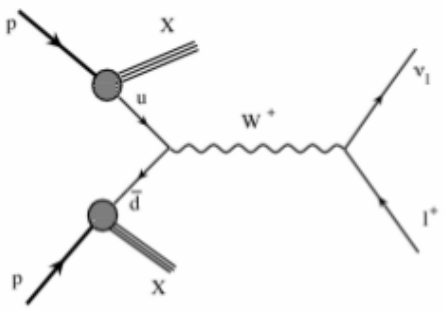


Figure 2.4: Drell-Yan process for W production in proton-proton collisions.

The Drell-Yan cross section is determined as follows by relying on the factorization theorem [54,56]:

$$\sigma_0 = \sum_{q\bar{q}} \int dx_q dx_{\bar{q}} f_q(x_q, Q^2) f_{\bar{q}}(x_{\bar{q}}, Q^2) \int \sigma_{q,\bar{q}}(x_q x_{\bar{q}} s, Q^2), \tag{2.7}$$

where $f_q, f_{\bar{q}}$ are the parton distribution functions of the annihilating quarks. Q^2 is the virtuality of the W boson. $x_q, x_{\bar{q}}$ are the momentum fractions carried by the annihilating quarks. σ describes the hard parton-parton cross section with a partonic center-of-mass energy squared $\hat{s} = x_q x_{\bar{q}} s$ (s is the center-of-mass energy squared of the collider). At the lowest order, σ is given by 2.8, where Γ_W and M_W are W width and W mass respectively.

$$\frac{d\sigma}{d\sqrt{\hat{s}}} \propto \frac{\hat{s}}{(\hat{s} - M_W^2)^2 + (\hat{s}\Gamma_W / M_W)^2}. \tag{2.8}$$

At leading order, the W is produced with zero transverse momentum as colliding partons are assumed to be collinear with the colliding beam particles [59, 83]. The transverse momentum of W is however

generated at higher orders in which p_T^W recoiling against one or more emissions of gluons and other strongly interacting particles. For example, Figure 2.5 shows subprocesses at first order in perturbative QCD in which the gluon emission is the dominant subprocess, $q\bar{q} \to Wg, gq \to Wq$ or $g\bar{q} \to W\bar{q}$.

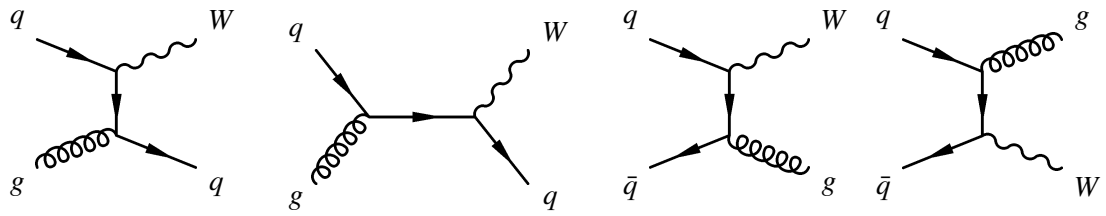


Figure 2.5: Lowest order diagrams for the production of a W and one jet at hadron colliders.

At high p_T^W, i.e. $p_T^W \sim M_W$, the p_T^W distribution can be calculated using perturbative QCD [8, 45]. In particular, for the gluon emission process as shown in Figure 2.5, the cross-section in the leading logarithm approximation is, (for more detail, see [59, 74])

$$\frac{1}{\sigma_0}\left(\frac{d\sigma}{d^2p_T^W}\right)_{O(\alpha_s)} = \frac{4\alpha_s}{3\pi^2}\frac{1}{(p_T^W)^2}\log\left(\frac{M_W^2}{(p_T^W)^2}\right), \tag{2.9}$$

where σ_0 and σ are explained as in equation 2.7. However at low p_T^W, $p_T^W < M_W$, the equation 2.9 is not reasonable as the logarithm becomes large and compensates for α_s being small. In this case, the multi-gluon emissions processes can be considered and a soft gluon resummation technique is applied to solve the problem. When p_T^W is much smaller, $p_T^W << M_W$, in which the perturbative QCD is not valid [55], it is thus needed to use non-perturbative techniques for the calculation of the p_T^W distribution. This will be discussed in more detail in subsection 2.2.1.

2.2.1 p_T^W distribution in low p_T^W region

As mentioned, in QCD, a non-zero transverse momentum p_T^W is generated by W recoiling against one or more emitted partons (see [78] and references therein). Order by order in perturbation theory, there are appearances of the logarithm, $\log(M_W^2/p_T^2)$. In particular, the order of α_s^N is followed by a series of logarithms $\log(M_W^2/(p_T^W)^2)^n, n = 0, 1, ..., 2N-1$. As p_T^W is low, the logarithms become large and divergences in which the leading logarithms, $n = 2N-1$, dominate the sum of perturbation theory. In this case, the cross section can not be calculated accurately in a fixed perturbative order but it must be summed over the orders. This technique is called resummation. In this section we will present the

main ideas of the resummation method. The formalism of this method in Drell-Yan processes has been proposed and developed by Collins, Soper and Sterman [32, 55, 57]. The resummation is performed in the Fourier conjugate of p_{T}-space, b-space, which allows to built the transverse momentum conservation condition into the formalism [42, 78].

The contributions for the small p_T^W region are from multiple gluon emission diagrams. In case of soft gluon emissions in which the W boson is balanced by just one of the emitted gluons while the rest of emitted gluons are soft, by treating similarly as in QED (see, for example, [65]), it is found that every diagram with an additional factor of α_s is accompanied by a factor $\log^2(M_W^2/(p_T^W)^2)$. Therefore summing over all these diagrams will give

$$\frac{1}{\sigma_0}\frac{d\sigma}{d^2p_T^W} = \frac{1}{\sigma_0}\left(\frac{d\sigma}{d^2p_T^W}\right)_{O(\alpha_s)} exp\left[-\frac{2\alpha_s}{3\pi}\log^2\left(\frac{M_W^2}{(p_T^W)^2}\right)\right]. \tag{2.10}$$

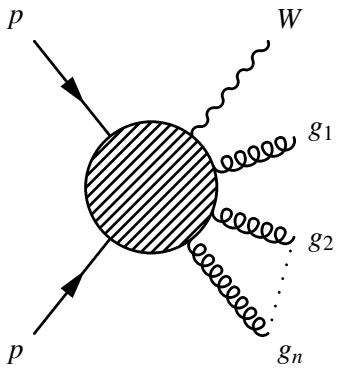


Figure 2.6: Diagram of W production with multi-gluon emissions in proton-proton collisions.

While (2.10) looks reasonable at low p_T^W region , $p_T^W < M_W$, it makes the cross section vanishes as $p_T^W \to 0$. This feature is not observed experimentally. To deal with this problem, we consider another approach which is provided by parton showers. In the low p_T^W region, it can happen that (see, for example, [42]), the dominant configurations have two or more gluons carrying larger transverse momenta which makes the p_T^W become small (the p_T^W is not balanced with the p_{T} of only one hard gluon but with a vector sum of transverse momenta of many hard gluons). These gluons are sufficient "hard" that the cross section can be calculated in perturbation theory. In this case, as in QED, the cross section for producing a W with the emission of the n gluons can be calculated by assuming that it can be factorized into the product of the cross section for emitting the gluons individually

$$\frac{1}{\sigma_0}\frac{d\sigma}{d^2p_T^W} = \sum_n \frac{1}{n!}\int \frac{d^2k_{T_1}}{\pi k_{T_1}^2}...\int \frac{d^2k_{T_n}}{\pi k_{T_n}^2}\left(\frac{4\alpha_s}{3\pi}\right)^n \log\left(\frac{M_W^2}{k_{T_1}^2}\right)...\log\left(\frac{M_W^2}{k_{T_n}^2}\right)\delta^{(2)}(k_{T_1}+...+k_{T_n}+p_T^W), \tag{2.11}$$

where $k_{Ti}, i = 1,..,n$, is transverse momentum of emitted gluons. The delta function that appears in (2.11) reflects the conservation of momenta of particles in the process. This delta function can be written as a Fourier transformation in the space of the "impact parameter" b:

$$\delta^{(2)}(k_{T_1}+...+k_{T_n}+p_T^W) = \frac{1}{2\pi}\int d^2b\ exp[i(k_{T1}+...+k_{Tn}+p_T^W)]. \tag{2.12}$$

From (2.11) and (2.12) we can get

$$\frac{1}{\sigma_0}\frac{d\sigma}{d^2p_T^W} = \frac{1}{4\pi}\int d^2b\ e^{ib.p_T^W}\chi(b) \tag{2.13}$$

by summing over multiple gluon factors, where $\chi(b) = e^{S(b)}$, $S(b)$ is known as Sudakov form factor and given by

$$S(b) = \int \frac{d^2k_T}{k_T^2}\left(\frac{4\alpha_s(k_T^2)}{3\pi}\right)\log\left(\frac{M_W^2}{k_T^2}\right)(e^{ib.k_T}-1). \tag{2.14}$$

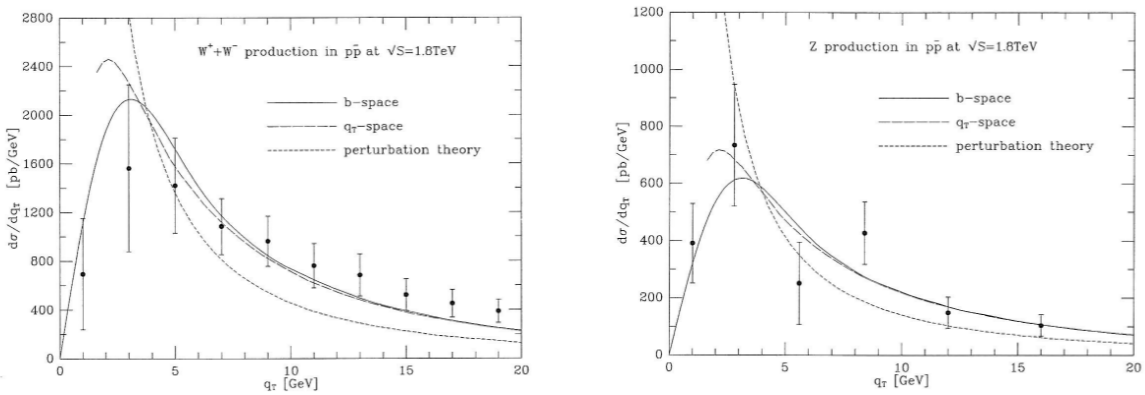
The impact parameter resummation makes sense only if the range of the integration in equation (2.14) satisfies $\Lambda^2 << k_T^2 << M_W^2$. At very low p_T^W values (below ~ 10 GeV), or at large b, the perturbative expansion is invalid and non-perturbative techniques need to be applied. These methods define a parameter

$$b_* = \frac{b}{\sqrt{1+(b/b_{max})^2}}, \tag{2.15}$$

where b_{max} is chosen such that the $\chi(b)$ function in (2.13) becomes

$$\chi(b) \to \chi(b_*)e^{-S_{NP}(b)}, \tag{2.16}$$

where S_{NP} is a non-perturbative function satisfying $S_{NP} \to 0$ as $b \to 0$ and $S_{NP} \to \infty$ as $b \to \infty$. The explicit from of this function is given in [78, 83] and references therein where the parameters to define this function are determined experimentally.



(a) $d\sigma/dp_T^W$ (b) $d\sigma/dp_T^Z$

Figure 2.7: Comparisons of various theoretical predictions with CDF data [6] of W and Z productions (taken from [78]).

Above we summarized the, so called, b-space resummation formalism. While this method succeeds in recovering a finite, positive result in the $p_{\rm T} \to 0$ limit, it still has several problems verified by Ellis and Veseli in [78]. To deal with the technical problems of the b-space method, in [78], the authors proposed a technique to resum the logarithms directly in the $p_{\rm T}$-space.

Figure 2.7 illustrates comparisons of the theoretical predictions for p_T^W and p_T^Z spectrums from b-space, $p_{\rm T}$-space formalisms, and pure perturbation theory with CDF data [6]. The b-space and $p_{\rm T}$-space results are in a good agreement describing better the data than the perturbation theory does in low $p_{\rm T}$ range. While both b-space and $p_{\rm T}$-space predictions describe quite well the Z data, they have worse agreement with the W data. However, it is clear that, the data needs a better statistic to reduce the fluctuation of the points as well as to provide a more reliable shape in order to test the model descriptions. Therefore, to investigate this region with W boson , we will need more data or maybe a different approach.

2.2.2 p_T^W distribution at high p_T^W region

With the high center-of-mass energy and the huge luminosity of proton-proton collisions at LHC, it can be produced the vector bosons having large transverse momentum at which new physics (such as new gauge bosons and extra dimensions) are expected to enhance the spectrum. It is thus needed to have an accurate calculation for W production at large p_T^W in order to distinguish the Standard Model from those new physics.

In high p_T^W region, $p_T^W > M_W/2$, the emission of hard partons becomes the dominant effect. A fixed order calculation in perturbative theory is thus valid to describe the p_T^W distribution. The fixed orders can be calculated by using matrix element convoluted with parton distribution functions as expressed in Eq. (2.7). Currently, the highest available order is the next-next-to-leading order. The $W + jet$ and W $+2$ $jets$ calculations using matrix elements were detailed in [59].

The complete calculations to NLO and NNLO soft gluon corrections for W production at high p_T^W at LHC were performed in [77]. Results from this paper found that the NLO corrections provide a large increase to the cross section but do not reduce the scale dependence relative to leading order (LO). The NNLO corrections, although they are small, significantly reduce the scale dependence thus providing a more stable theoretical prediction.

With the large amount of data produced at the LHC that allows to extend the measured p_T^W range, the theoretical predictions in this region will be compared with our measurement of the p_T^W spectrum with the ATLAS detector in Chapter 6.

2.2.3 Modeling of W bosons production by Monte Carlo event generators

In practice, especially at LHC with interactions at very high energy, multi-hadronic final states are produced. The final states may contain many particles produced by variety complicate physics processes [59, 81]. The physics is then not understood well enough to give an exact description. Instead, the computer programs are introduced to describe those physics processes basing upon a combination of analytical results and various QCD-based models. Event generators are a solution in which the processes are simulated using an event basic.

There are basically two kinds of event generators. The first one, which is so called leading order (LO) event generator, deals with leading order process with or without adding corrections of higher QCD or QED pertubation orders. The second kind, which is so called NLO event generator, deals with full NLO QCD or QED process. These event generators apply matrix element calculation for the corresponding pertubation order. The matrix element used is different depending on specific purposes of the generators.

LO generators start with the matrix element of the tree-level s-channel subprocess as given in equation 2.8. This is the first estimation to the W boson production. In order to model better the data, it is needed to take into account higher order corrections since, as mentioned, the final state in fact contains many particles. These corrections are made by adding higher order effects to the LO matrix element.

There are two kinds of corrections: QCD corrections including radiation off incoming or outgoing quarks and QED corrections containing final state radiation of leptons. As an example, the NLO diagrams for initial state radiation are displayed in Figure 2.5. There is a variety of calculations for the initial (and final) state radiation which are modeled by a number of Monte Carlo generators.

The popular LO generators such that PYTHIA [81], HERWIG [33, 41], RESBOS [9, 39, 63], ALPGEN [68] and SHERPA [44] combine a LO calculation with higher order pertubative QCD corrections. The PYTHIA and HERWIG programs merge a LO calculation with a parton shower algorithm as presented in Section 2.2.1. In particular, PYTHIA models the kinematics of soft and collinear parton radiation at all orders in which the parameters of the parton shower algorithm have been tuned to match the p_T^Z data from the Tevatron [82]. The HERWIG generator includes a more correct treatment of heavy quark decays. The RESBOS program implements resummation of leading and next-to-leading logarithms due to soft gluon radiation. The ALPGEN and SHERPA programs provide the study of multiparton hard processes in hadronic collisions resulting multi-jet final states. It is thus interesting to compare the p_T^W spectrum modeled by these programs to the highest p_{T} part of the p_T^W spectrum observed in data, which includes contributions from a W boson recoiling off multiple high-p_{T} jets.

The QED corrections for the LO subprocess includes photon radiation off the initial state partons, the final state lepton, and the W propagator. The multiple photon emissions off the final state lepton can be generated by The PHOTOS Monte Carlo simulation [11] which is often used in conjunction with other generators. The PHOTOS generator, however, does not compute initial-state radiation or radiation off the propagator. There is also a developed version of RESBOS, RESBOS-A [17] which includes final-state NLO QED corrections to W and Z production.

The NLO generator such that MC@NLO [40] combines NLO calculation with parton showering and hadronization. The MC@NLO generator produces full NLO predictions of the p_T^W spectrum, integrated with a parton shower algorithm such that the differential cross section prediction is finite for the whole range of p_T^W spectrum.

In addition, there are also NNLO, $O(\alpha_s^2)$, (currently is the highest order available for inclusive W and Z productions) programs such as FEWZ [43, 70] and DYNNLO [18]. These calculations do not include resummation or parton shower algorithms. So although the total cross section is finite, the differential cross section diverges as p_T^W approaches zero.

In Chapter 6 we make comparisons of our measurement of p_T^W spectrum produced at LHC to theoret-

ical predictions modeled by various event generators including PYTHIA, RESBOS, ALPGEN, MC@NLO, DYNNLO, and MCFM [16].

Chapter 3

ATLAS as a detector of the LHC

ATLAS (*A Toroidal LHC ApparatuS*) is one of the two general-purpose detectors, amongst all seven detectors, of the LHC (*Large Hadron Collider*) which is a new particle accelerator located at CERN, near Geneva. After over 10 years of preparation and construction the LHC released the first beam on 10 September 2008 and made the first collision in November 2009. However, before going to look at the details of the ATLAS detector let us make a brief journey to the LHC in order to understand more the construction and the functions of the ATLAS detector in the LHC.

Designed to accelerate and collide two beams of protons or heavy ions at a total center-of-mass energy of 14 TeV (2.24 microjoules) or 5.2 TeV, respectively, and installed at 50–100m underground in a circular tunnel (the former LEP tunnel) with a circumference of about 27 km (26 659 m), the LHC, costing several billions euros, has been so far the largest, highest-energy and most expensive science-technological complex ever constructed and attracting the widest collaboration in the history of the world science and technology. A view of the LHC general structure is given in Figure 3.1.

3.1 Hadron Colliders

After e^-e^+ colliders (e.g., LEP2) have reached the limit of their possibilities and when TeV linear colliders have not yet been technologically ready, while the idea of TeV muon colliders still remains an idea, hadron colliders are a natural choice for an exploration machine at the TeV scale (see, [73] and [85] and references therein). Because of their complexity and high cost, only two TeV colliders have been constructed so far: the Tevatron (in Fermilab), a 2 TeV proton-antiproton ($p\bar{p}$) collider, is the first collider reaching a TeV energy, hence the name, and the LHC has been so far the largest collider. At the

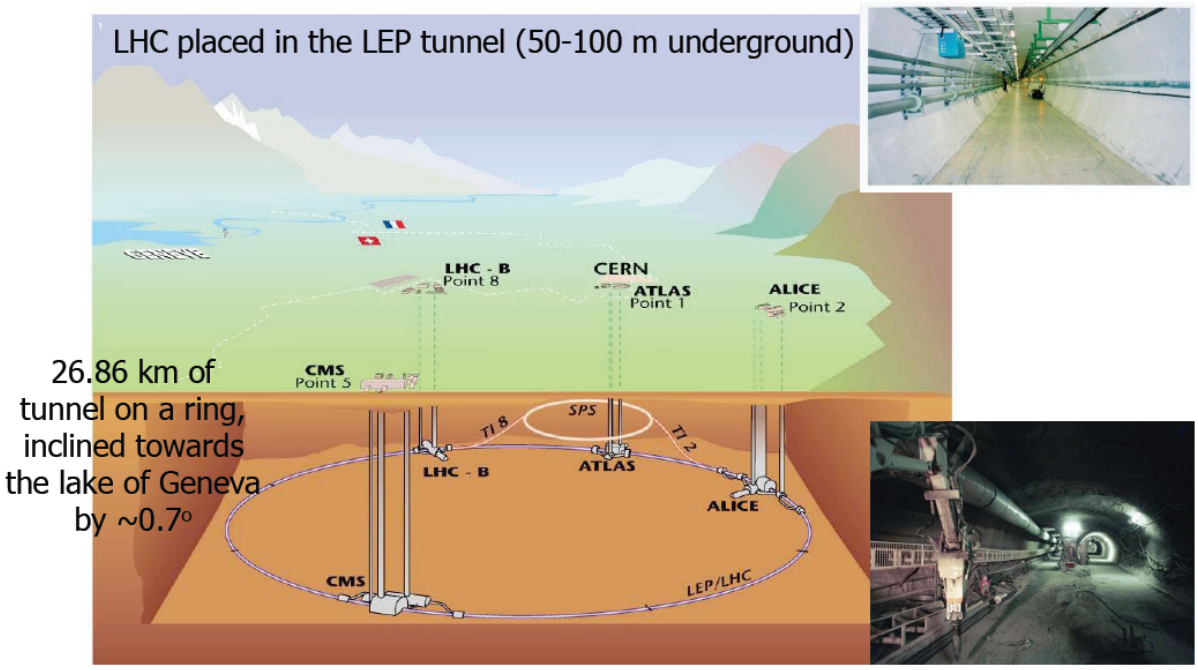


Figure 3.1: The LHC general scheme

moment the LHC is running at an energy of 7 TeV, or 3.5 TeV per beam, and, probably, it will reach the design maximal energy 14 TeV after 2012.

Compared to the e^+e^- collider, hadron colliders have the following advantages:

- Much lower synchrotron radiation and higher center-of-mass energies in head-on collisions are achieved since the proton mass is much bigger than the electron mass;
- Higher luminosity;
- Large cross-sections of hadronic reactions due to strong interactions;
- Since collison energies are higher, there are more possible channels opening up resonant productions such as $q\bar{q}$, qg, gg and WW, ZZ, WZ fusion processes.

However, hadron colliders have also disadvantages (see Ref. 6 in [85]):

- As said above, the momentum carried by the protons is shared by quarks and gluons. The interactions between protons, in fact, take place between partons. Thus, the effective center-of-mass energy of the collisions is just a fraction, about 1/6 on average, of the 14 TeV;

- There is a large cross-section of inelastic proton-proton processes at small momentum transfer with no physics interest.

3.2 General structure of the LHC

The LHC was designed and constructed to collide two beams of protons (or heavy ions). Protons first are dissociated from hydrogen atoms in the Duoplasmatron and injected into radio-frequency (RF) cavities where they are accelerated to 750 keV before being injected into the Linear Accelerator (LINAC) accelerating them to 50 MeV. Then they will be consecutively accelerated by the Proton Synchrotron Booster (PSB) and Proton Synchrotron (PS) to the energies 1.4 GeV and 25 GeV, respectively. Finally, protons are injected into the LHC after getting an energy of 450 GeV from the Super Proton Synchrotron (SPS). This acceleration process of protons is illustrated by Figure 3.2.

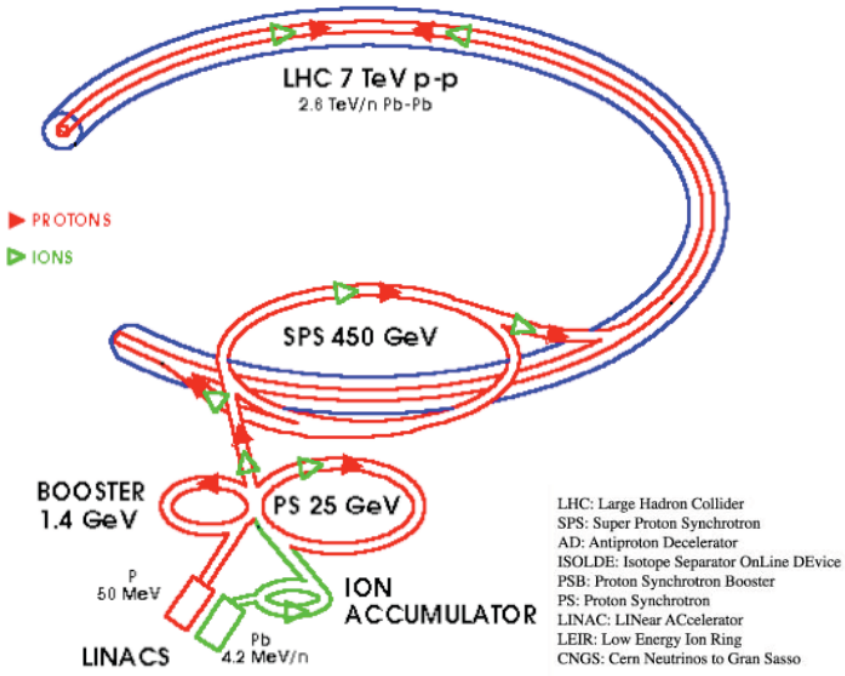


Figure 3.2: The acceleration way of the protons and heavy ions

The LHC has seven detectors, ATLAS, CMS, LHC-b, ALICE, TOTEM, LHC-f and MoEDAL installed in underground caverns at four of the eight interaction points (IP's) on the LHC ring. ATLAS located at IP1 is the biggest by size detector and it, along with CMS (*Compact Muon Solenoid*) located at IP5, is one of the two general-purpose detectors of the LHC. The five smaller and specialized detectors

are LHC-b (*LHC-beauty*) designed for B-physics, in particular, for measurements of CP-violation parameters in the processes involving B-hadrons, ALICE (*A Large Ion Collider Experiment*) at IP2 for study of ion collisions, e.g., lead-lead nuclei collisions for investigations of the quark-gluon plasma, TOTEM (*Total cross section, elastic scattering and diffraction dissociation*) at IP5 for measurements of total $p-p$ cross section, elastic scattering and diffractive processes, and LHC-f (*Large Hadron Collider forward*) at IP1 designed with the astroparticle physics purpose, and MoEDAL (*Monopole and Exotics Detector At the LHC*) sharing the IP8 cavern with LHCb and searching for the massive stable (or pseudo-stable) particles, such as magnetic monopoles or dyons, which may be produced at the LHC.

As the LHC, a proton-proton collider, accelerates and collide particles with the same charge, two beam lines for two opposite accelerations must be installed in the tunnel. The particle beams are accelerated in an ultrahigh vacuum (10^{-10} *Torr*). To circulate in the tunnel they are bent by 1232 powerful superconducting dipole magnets and focused by 392 other quadrupole magnets. To ensure their normal work these 1624 superconducting magnets with niobium-titanium coils (NbTi) must be kept at temperature 1.9K (with 96 tonnes of liquid helium). In order to bend 7 TeV beams around the LHC ring (with a bending radius of 2804 m) the magnets should have a field of about 8.33 Tesla corresponding to a current of about 11 850 A.

Protons circulate in the ring not continuously but in portions called *"bunches"*. Here for an ilustration we take the design case (i.e., not the present real situation). Then, there are 2808 (effective number) bunches for a proton beam at full intensity and the density of protons per bunch (at interacting point) is about 10^{11} protons/16 microns (1/5 the width of a human hair). Protons at the maximal speed make 11245 turns (crosses) per second and each bunch crossing (about 7.5 m long) takes place inside detector for 25 *ns* of bunch spacing or totally, a beam would make 2808×11245= about 32 millions crosses per second, i.e., its crossing rate is 31.6 MHz which is considerably lower than the nominal one, 40 MHz, when all 3564 (nominal number) bunches can be filled in a beam.

However, not all collision events but only inelastic events (or shortly, events) can be detected by a detector. The probability of a process to occur is measured by the corresponding interaction cross section. Therefore, the interaction cross section (denoted by σ_{int}), thus, the collision energy, is an important parameter of a collider like the LHC. The event rate (R), however, depends not only on the collision energy (more precisely, it grows with energy) but also another characteristic parameter of the collider called the luminosity (L):

$$R = L \times \sigma_{int}. \tag{3.1}$$

The latter is proportional to the beam intensities (n_1 and n_2), the number of bunches per beam (k) and the beam collision frequency (f) and inversely proportional to the beam transverse sizes at the interaction point (S_x and S_y):

$$L = kf\frac{n_1 n_2}{4\pi S_x S_y}. \tag{3.2}$$

Since the colliding beams are almost identical we can assume that $n_1 = n_2 \equiv N$, i.e., they have the same intensity. In this case (3.2) becomes

$$L = A \times N^2, \quad A \equiv \frac{kf}{4\pi S_x S_y}. \tag{3.3}$$

To observe new interesting processes which usually have very small cross sections it is necessary to have a luminosity large enough. Typically, for the processes to be searched for by the LHC the design luminosity is $10^{34}\ cm^{-2}s^{-1}$. At the design energy, 14 TeV, the total $p-p$ cross section is about 110 mb which can be broken down to 60 mb, 40 mb and 12 mb for inelastic, elastic and single diffractive events, respectively. Therefore, for the nominal luminosity $10^{34}\ cm^{-2}s^{-1}$, about 600 million events per second, or about 20 events per crossing, can be detected by a detector.

However, the beam intensity and the luminosity are not constants but decreasing functions of time: $N = N(t)$ and $L = L(t)$. Their decays are described by the formula [61]

$$dN(t) = -L(t).\sigma_{tot}.n_{ip}.dt, \tag{3.4}$$

where σ_{tot} is the total $p-p$ cross section and n_{ip} is the number of interaction points. Assuming A is a constant, i.e., the beam parameters hiden in A do not change, during operation, the solution of (3.4) is

$$N(t) = \frac{N_0}{1 + N_0.A.\sigma_{tot}.n_{ip}.t}, \tag{3.5}$$

where $N_0 \equiv N(0)$ is the initial beam intensity. Denoting the initial luminosity by $L_0 \equiv A.N_0^2$ we have the following functions of the beam intensity and the beam luminosity

$$N(t) = \frac{N_0}{1 + t/\tau} \tag{3.6}$$

and

$$L(t) = \frac{L_0}{(1 + t/\tau)^2}, \tag{3.7}$$

where

$$\tau \equiv \frac{N_0}{k.L_0.\sigma_{tot}.n_{ip}} \tag{3.8}$$

is the (initial) decay time of the beam intensity. Because all bunches in a beam are treated equally, it is not difficult to see that the intensity for a bunch decays as $N(t)$ in (3.6). The beam lifetime and the luminosity lifetime are then given by the formulas

$$\tau_b = (e-1)\tau, \tag{3.9}$$

and

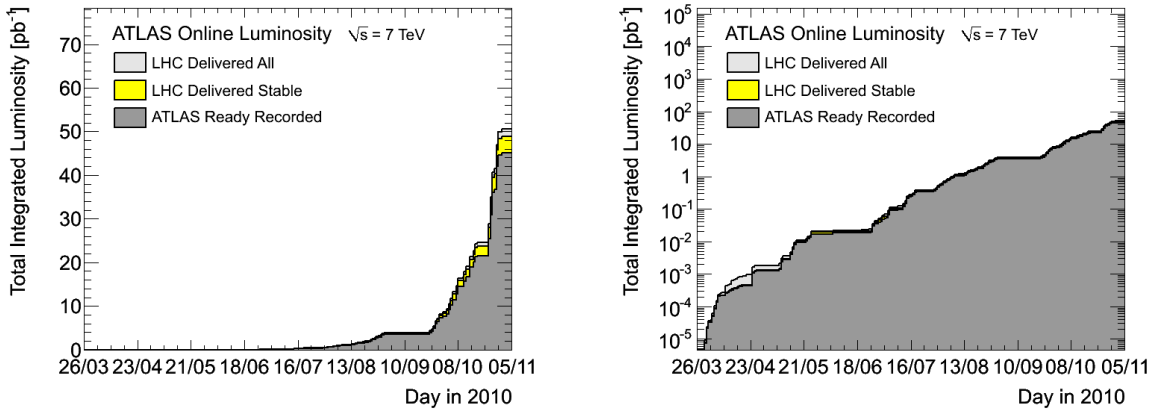
$$\tau_l = (\sqrt{e}-1)\tau, \tag{3.10}$$

respectively. With $L_0 = 10^{34}\ cm^{-2}s^{-1}$ and $N_0 = 10^{11}$, as given above, the beam lifetime and the luminosity lifetime are about $150/n_{ip}$ hours and $60/n_{ip}$ hours, respectively. For one interaction point (i.e., $n_{ip} = 1$) the beam and the luminosity lifetimes are about $150h$ and $60h$, respectively. The lifetimes become shorter for more interaction points. There are, however, other factors (such as Toucheck effect, different scatterings, etc.) reducing these lifetimes. The knowledge of the latter is very important for the machine operation (to know approximately when to stop a circulating but already weakened beam and start a new beam) [61].

During 2010, the data period used for the analysis performed in this thesis, the LHC ran with an energy per beam of 3.5 TeV; a peak instantaneous Luminosity of $2 \times 10^{32} cm^{-2}s^{-1}$; a number of bunches per beam of 368. The total integrated luminosity delivered by the LHC for the whole period is 48 pb^{-1} in which 45 pb^{-1} were recorded by ATLAS with an efficiency of 94%. The delivered and recorded integrated luminosity over the periods are displayed in Figure 3.3.

3.3 Coordinates for the LHC

There are several coordinates systems used in the LHC. The Cartesian system has an origin at the nominal collision point, the positive direction of the x-axis points to the center of the LHC ring, the positive direction of the y-axis is upward and the z-axis laying along the beam line has a direction to form with the x- and y- axes a right-handed Cartesian coordinate system. Then the polar coordinate system is defined by choosing the polar angle θ with respect to the z-axis, the azimuthal angle ϕ around this axis, and the radial coordinate $\rho = \sqrt{x^2+y^2}$.



(a) Linear scale

(b) Log scale

Figure 3.3: Integrated luminosity delivered by the LHC and recorded by the ATLAS detector during 2010.

In (high energy) hadron colliders, the parton model is used. We do not perform calculations directly for protons but for partons (quarks, gluons,...). It makes difference as the partonic center-of-mass (CM) frame is not the same as the hadronic CM frame. For the analysis of the pp collisions, the momentum vector is usually given in the coordinates (p_T, η, ϕ), where the first one, p_T, is the *transverse momenum*, the second one, η, is called *pseudo-rapidity* (see the next subsection) and the last one, ϕ, is the above mentioned *azimuthal angle* around the beam axis. Besides $1/p_T$, η and ϕ, a trajectory of a charged particle can be parameterized by two more parameters, d_0 and z_0, where d_0 is the transverse distance to the beam axis at the point of closest approach (of the track), while is z_0 is the distance from the origin along the z-axis at the point of closest approach.

3.3.1 Transverse momentum and the azimuthal angle

Because the ambiguous motion between the parton CM frame and the hadron laboratory frame is along the longitudinal beam direction ($\vec{z}$), it is thus convenient to write the phase space element in the cylindrical coordinate as

$$\frac{d^3\vec{p}}{E} = dp_x dp_y \frac{dp_z}{E} = p_T dp_T d\phi \frac{dp_z}{E}, \tag{3.11}$$

where ϕ is the azimuthal angle about the $\vec{z}$ axis, and

$$p_T = \sqrt{p_x^2 + p_y^2} = p \sin\theta \tag{3.12}$$

is the transverse momentum. In pp collider, an unbalance in p_z can come from the incoming partons themselves. So, the only quantity that one may require to equal zero after the interaction is the transverse momentum p_T. Events with production of undetectable particles (such as neutrinos or particles predicted by SUSY) are characterized by the existence of missing p_T.

3.3.2 Rapidity and pseudo-rapidity

The rapidity of a particle of momentum p^μ is defined to be

$$y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z}. \tag{3.13}$$

We define also the transverse energy

$$E_T = \sqrt{p_T^2 + m^2}, \tag{3.14}$$

$$\Rightarrow E^2 = \vec{p}^2 + m^2 = p_T^2 + p_z^2 + m^2 = E_T^2 + p_z^2. \tag{3.15}$$

From (3.13) and (3.15) we have

$$\begin{aligned} \cosh^2 y &\equiv \frac{e^{2y} + e^{-2y} + 2}{4} \\ &= \frac{1}{4}\left(\frac{E+p_z}{E-p_z} + \frac{E-p_z}{E+p_z}\right) + \frac{1}{2} \\ &= \frac{E^2}{E^2 - p_z^2} = \frac{E^2}{E_T^2}, \end{aligned} \tag{3.16}$$

thus (3.15) and (3.16) give

$$\begin{cases} \cosh y = \frac{E}{E_T}, \\ \sinh y = \frac{p_z}{E_T}. \end{cases} \tag{3.17}$$

Using (3.17) and expressing p_x and p_y in terms of the transverse momentum p_T and azimuthal angle ϕ we represent a particle four-momentum as

$$p^\mu = (E_T \cosh y, p_T \sin\phi, p_T \cos\phi, E_T \sinh y). \tag{3.18}$$

The phase space element can be given

$$\frac{d^3\vec{p}}{E} = p_T dp_T d\phi dy = E_T dE_T d\phi dy. \tag{3.19}$$

The rapidity is not Lorentz invariant but changes just by a constant under a relativistic boost. Let us consider the rapidity in a boosted frame (say, in the parton CM frame), and perform the Lorentz transformation,

$$\begin{aligned} y' &= \frac{1}{2}\ln\frac{E'+p'_z}{E'-p'_z} \\ &= \frac{1}{2}\ln\frac{(1-\beta_0)(E+p_z)}{(1+\beta_0)(E-p_z)} \\ &= y-y_0, \end{aligned} \tag{3.20}$$

where β_0 is the relative velocity of the boosted frame with its rapidity

$$y_0 = \frac{1}{2}\ln\frac{1-\beta_0}{1+\beta_0}. \tag{3.21}$$

In the massless limit, $E \simeq |\vec{p}|$, we get

$$y \to \frac{1}{2}\ln\frac{1+\cos\theta}{1-\cos\theta} = \ln\cot\frac{\theta}{2} \equiv \eta, \tag{3.22}$$

where η is the pseudo-rapidity, which has one-to-one correspondence with the scattering polar angle $\pi \geq \theta \geq 0$ for $-\infty < \eta < +\infty$, for instance, when $\theta = \pi/2$, η takes the value $\eta = 0$. We see in the massless limit that the rapidity is equivalent to the pseudo-rapidity.

3.4 The ATLAS Detector

The ATLAS is a general purpose pp detector of the LHC. Consisting of many components designed to detect different types of particle tracks produced from collisions, the ATLAS detector has four major parts (counted from the center outwards): the inner detector, the calorimeters, the muon spectrometer and the magnet systems. The inner detector tracks charged particles precisely, the calorimeters measure the energy of particles, and the muon spectrometer determines tracks of muons. All are surrounded by two magnet systems designed to bend charged particles in the inner detector and the muon spectrometer. If the strength of the magnetic field is known, the particle momenta can be also measured. Particles like neutrinos which do not interact with the detector can be inferred via an estimation of missing transverse energy. An overall ATLAS lay-out is shown in Figure 3.4 and traces leaved by various types of particles in detector components are illustrated in Figure 3.5.

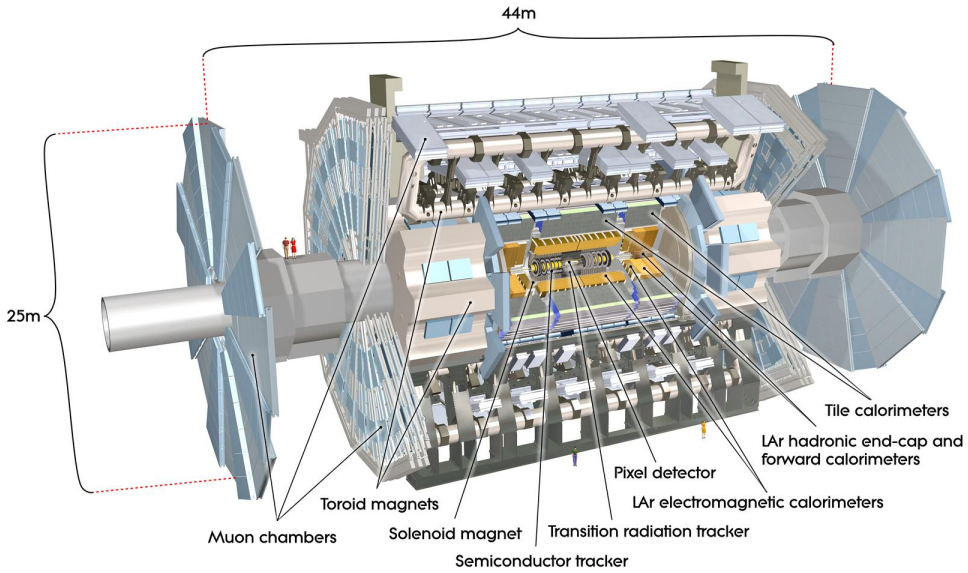


Figure 3.4: The ATLAS detector layout [19].

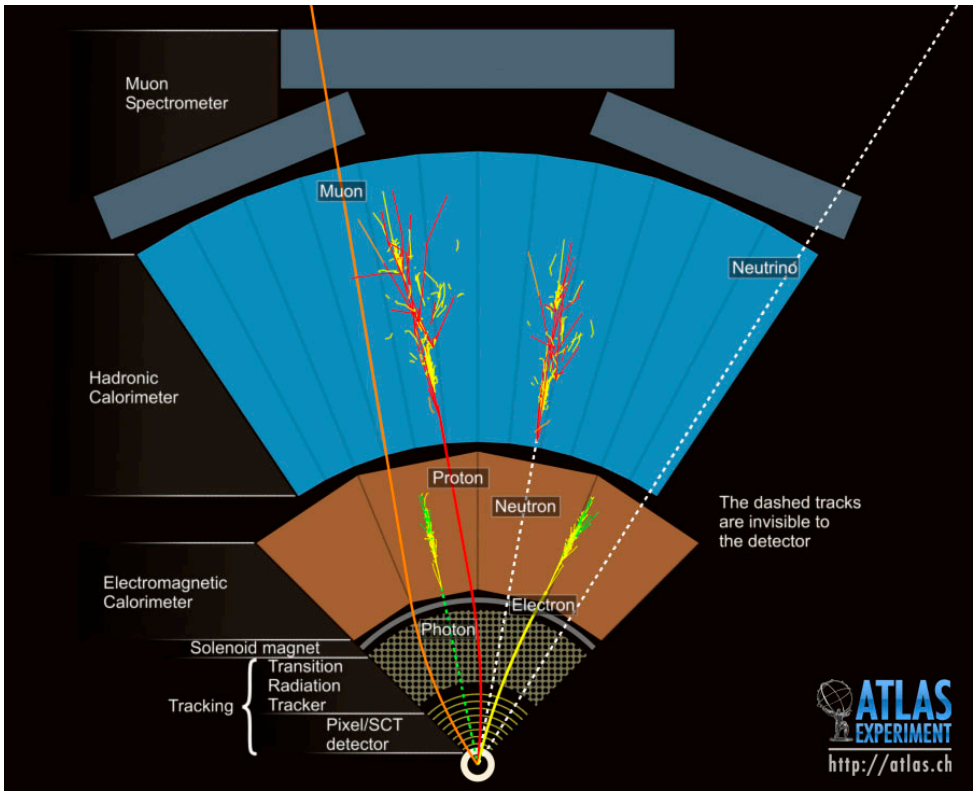


Figure 3.5: Traces of various types of particles leaved in components of the ATLAS detector [10].

3.4.1 Inner detector

The basic function of the inner detector (see Figure 3.6), which is a silicon detector, is to track charged particles, trajectories of which are bent by a 2 T solenoidal magnetic field, by detecting their interaction with matter at discrete points so that detailed information about particles of different types including their identities, momenta and their energy loss can be determined. The starting points of tracks can give useful information for identifying particles. This is a "displaced vertex" problem. For example, if tracks do not seem to point back to the original proton-proton collider point, this may be a sign that the particles came from some decays such as those of the bottom quarks.

With a radius of 1.2 m beginning a few centimeters from the proton beam axis and a length of 7 m along the beam pipe, the inner detector has three parts: the pixel detector, the semi-conductor tracker (SCT) and the transition radiation tracker (TRT) presented as in Figure 3.6 below. The pixel and the semi-conductor tracker cover the pseudo-rapidity range $|\eta| < 2.5$, and the transition radiation tracker covers the range $|\eta| < 2.0$.

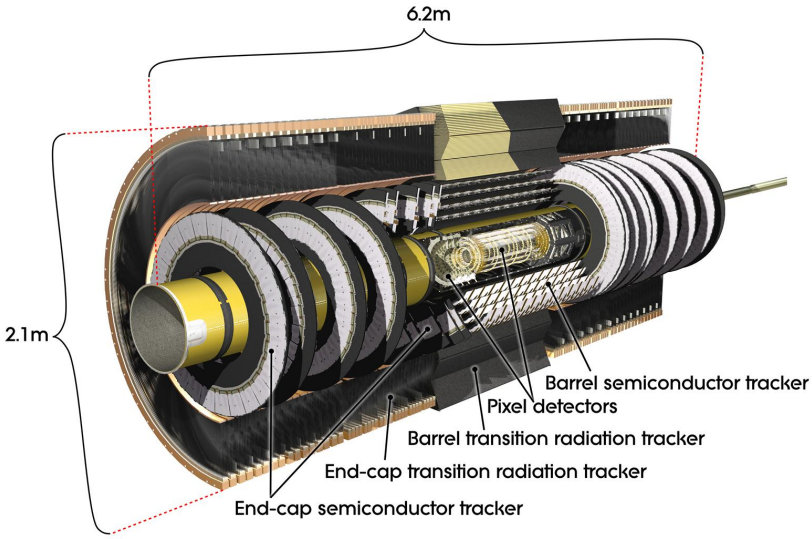


Figure 3.6: The inner detector layout [19]

The pixel detector being a very high-granular and high-precision semiconductor based tracker, is the innermost part of the detector providing very high resolution three space-point measurements adjacent to the beam pipe for vertexing of displaced decays such as those from B hadrons in the environment of the very high particle multiplicity of LHC collisions. It contains three layers and three disks on each end-

cap. The detecting material is made of silicon with a thickness of 250 μm. In total, the pixel detector has 1744 modules. Each module contains 16 readout chips and other electronic components. The smallest unit that can be read out is a pixel, (50 in each 400 μm); there are roughly 47,000 pixels per module. This size of pixel design guarantees extremely precise tracking very close to the interaction point.

The semi-conductor tracker (SCT) providing eight space-point measurements for determining particle vertices and trajectories, is the middle component of the inner detector which has four double layers of silicon strips. It works in a similar way, therefore, it has a similar design to the pixel detector but with long, narrow strips rather than small pixels. The SCT is the most critical part of the inner detector for basic tracking in the plane perpendicular to the beam.

The transition radiation tracker (TRT), the outermost component of the inner detector, providing approximately 36 measurements of particle trajectories in the bending direction and the electron identification capability, is a combination of a straw tracker and a transition radiation detector. It contains many small straws (351,000 in total). Each straw, with a diameter of 4 μm and a length of 144 cm, is filled with gas that becomes ionized when a charged particle passes through. The ions produce a current in a high-voltage wire running through the center of a straw. In addtion, TRT contains alternating materials causing charged particles to produce transition radiation and leave much stronger signals in each straw. Since the amount of transition radiation produced is greatest in highly relativistic particles, particle with many very strong signals can be identified as electrons since electron is the lightest charged particle.

The inner detectors have been shown a very good performance with a similar behavior to the expectations. The tracking performance of the inner detector can be determined from reconstructions of resonances of hadrons such as K_s^0, Φ, D mesons and Ω, Ξ, Λ baryons. As examples, the reconstructions of mass of K_s^0 decaying into a pair of pions and J/ψ decaying into a pair of muons are illustrated in Figure 3.7 and 3.8, respectively. These studies allow to determine the momentum scale and the detector resolution. The precision of momentum scale was measured at per mill level in low transverse momentum region and at per cent level in higher transverse momentum region (up to 100 GeV) [21].

3.4.2 Calorimeters

The calorimeters (see Figure 3.9) are used to measure energies and positions of electrons, photons, hadrons (detected as jets) and missing transverse energy. They are situated outside the solenoidal magnet

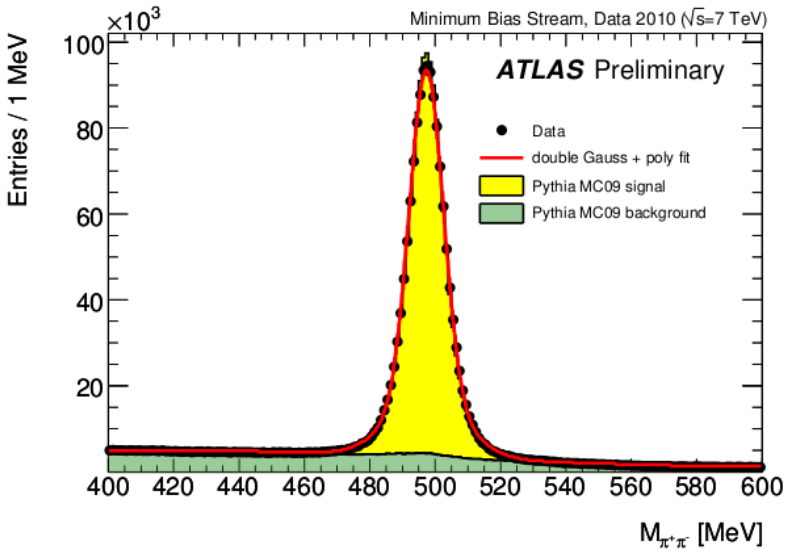


Figure 3.7: Distributions of $\pi^+\pi^-$ invariant mass (both tracks are in $|\eta| < 1.2$ region) in data (the black circles) and Monte Carlo simulation (the histogram normalized to data) at K_s^0 mass range. The red line is the line-shape function fitted to data [24].

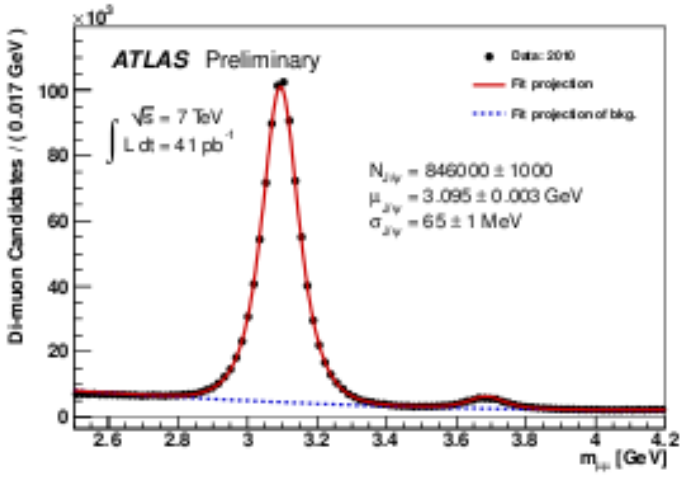


Figure 3.8: Distributions of $\mu^+\mu^-$ invariant mass in data at the J/ψ and $\psi(2S)$ mass ranges. The red line is the line-shape function fitted to data [21].

that surrounds the inner detector. They consist of an inner barrel cylinder and two end-caps. Both the barrel and the end-caps are split in an inner electromagnetic layer and a hadronic layer. The barrel part of the electromagnetic (EM) calorimeter is an "accordion" Liquid Argon (LAr) calorimeter. In the end-caps, the LAr technology is also used for the hadronic calorimeter. The calorimeters have the following main components [64]: the EM barrel (EMB) covering a pseudo-rapidity range of $|\eta| < 1.475$, the EM end-cap covering a range of $1.4 < |\eta| < 3.2$, the tile scintillator hadronic barrel covering a range of $|\eta| < 1.7$, the hadronic end-cap (HEC) covering a range of $1.5 < |\eta| < 3.2$, and the forward calorimeter (FCAL) covering a range of $3.1 < |\eta| < 4.9$. The FCAL calorimeter provides both electromagnetic and hadronic energy measurement.

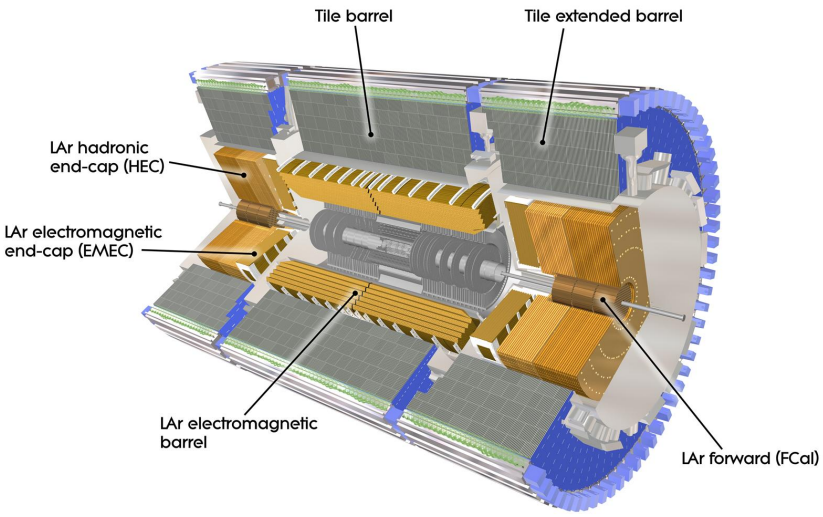


Figure 3.9: The ATLAS Calorimeters.

Electromagnetic calorimeter

The EM calorimeter absorbs and, thus, determines the energy of particles interacting electromagnetically such as electrons, photons. The energy-absorbing materials are lead and stainless steel, with a cryostat around the EM calorimeter to keep it sufficiently cool (see Figure 3.10). This calorimeter are segmented longitudinally into three layers named "strip", "middle" and "back". The cell size of $\triangle\eta \times \triangle\phi$ is 0.025 $\times$ 0.025 in the middle layer , 0.003 $\times$ 0.1 in the trips of the EMB and larger at higher range of $|\eta|$. In addition, there is also a presampler (PS) layer covering $|\eta| < 1.8$. The role of this layer is to improve the

energy measurement for particles that start showering before entering the calorimeter.

The EM energy scale and the performance of the electromagnetic calorimeter can be determined by reconstructing the invariant mass of di-photon and di-electron. With 2010 data produced by proton-proton collision at $\sqrt{s} = 7$ TeV, the Z boson decaying into a pair of electrons was used to determine the EM energy scale. The $Z \to ee$ events are selected with two opposite charged electrons having transverse momentum greater than 20 GeV and an invariant mass falling on the range of Z line-shape. Figure 3.11 displays a very good agreement of e^+e^- invariant mass spectrum around the Z mass in Monte Carlo and in the data calibrated by correcting the electron energy. The corrections of electron energy scale and resolution are also applied for this analysis of p_T^W spectrum measurement and will be discussed more detail in Section 5.7 of Chapter 5.

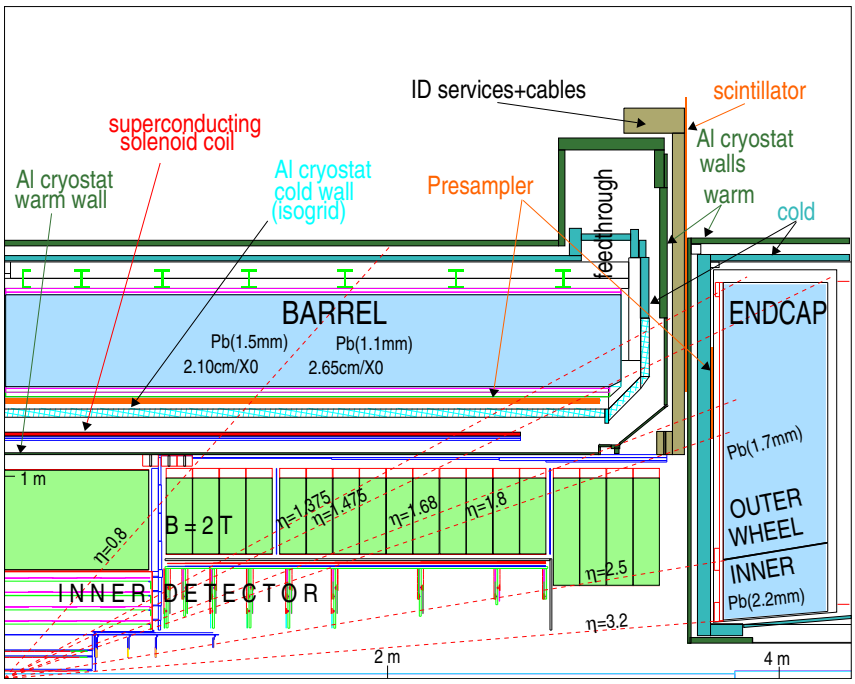


Figure 3.10: Longitudinal view of a quadrant of the EM Calorimeter [22].

Hadronic Calorimeter

The hadronic calorimeter (HCAL) absorbs and, thus, determines the energy of particles (protons, neutrons, etc.) that pass through the EM calorimeter but participate in the strong interaction. These particles are hadrons. The energy-absorbing material is steel, with scintillating tiles that sample the energy de-

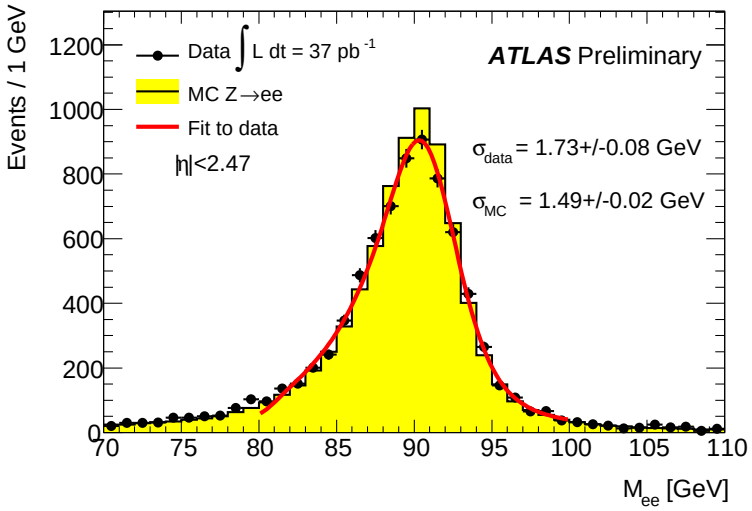
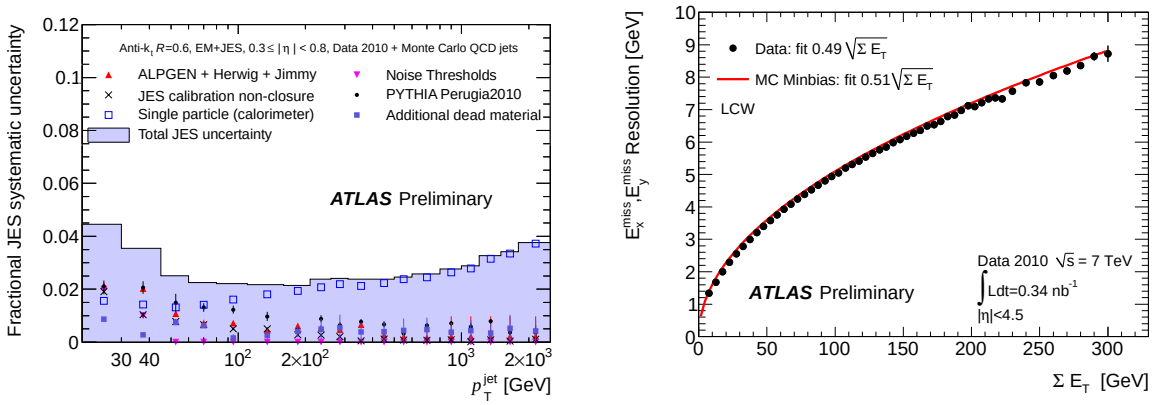


Figure 3.11: The distribution of e^+e^- invariant mass at Z mass range in data (black points) and Monte Carlo simulation (the histogram normalized to data entries).

posited (a reason for the name of "Hadron Tile Calorimeter"). The tile calorimeter is the main part of the HCAL with 8 meters in diameter and covers 12 meters along the beam axis. It is divided into three layers while the HEC component of HCAL is divided into four layers. The cell sizes in the HCAL are larger than in the electromagnetic calorimeter. They range from 0.1×0.1 to 0.2×0.2.

Performance of HCAL can be determined by agreements of data and Monte Carlo expectations for jets and missing transverse energy (E_{T}^{miss}) quantities. In ATLAS with 2010 data, jets were reconstructed by the anti-kt algorithm [20]. The fractional systematic uncertainty of jet energy scale as a function of jet transverse momentum (p_{T}^{jet}) in the barrel region is shown in left plot of Figure 3.12. It is found to be less than 3% in the range of p_{T}^{jet} between 50 and 1000 GeV. E_{T}^{miss} is a sensitive variable to evaluate the performance of the calorimeter in terms of noise dead cells, mis-calibration and mis-identification of particles, etc. It is found that the E_{T}^{miss} measured in data is in a good agreement with the simulation. For example, the E_{T}^{miss} resolution measured in data and simulated by Monte Carlo are compatible as illustrated in the right plot of Figure 3.12.

In ATLAS, two algorithms are used to reconstruct particles based on calorimeter information. These algorithms are detailed in [64]. The first one, which is called "sliding-window" algorithm, clusters calorimeter cells within fixed-size rectangles. With fixed-size clusters built, this algorithm is an efficient tool allowing to identify electron, photon and tau lepton and to reconstruct precisely electromagnetic



(a) Jet systematic uncertainty [20]

(b) E_{T}^{miss} resolution [25]

Figure 3.12: Fractional systematic uncertainty of jet energy scale as a function of jet transverse momentum and comparison of E_{T}^{miss} resolution as a function of total transverse energy in data and simulation.

showers and jets from tau lepton decay. The second one, which is called "topological" algorithm, clusters around a seed cell neighboring cells without fixing number of cells as long as the cell energy is significant compared to noise. This algorithms is thus an efficient tool for jets and missing transverse energy reconstruction where clusters with noise suppressed are built with a large number of cells.

3.4.3 Muon Spectrometer

The muon spectrometer (MS) layout is illustrated in Figure 3.13. It is an extremely large tracking system, beginning from a radius of 4.25 m around the calorimeters out to the full radius of the detector (11 m). With about one million readout channels, and a total area of its layers of 12,000 m^2, the muon spectrometer can provide particle tracking up to $|\eta| = 2.7$ and triggering up to $|\eta| = 2.4$. Its function is to accurately measure the momenta of muons which in principle pass through other components of the detector. Similarly to the inner detector, here, muons are curved by magnetic field so that their momentum can be measured. In addition, the muon spectrometer can also identify muons.

The performance of muon spectrometer can be evaluated from reconstructions of the invariant mass of well known particles decaying into two charged muons. Figure 3.14 shows the resonances obtained by reconstructing the invariant mass of two opposite charge muons originating from a common vertex and having transverse momentum (p_{T}) greater than 15 GeV. In practice, the J/ψ resonance is used to study the MS performance in the low p_{T} region while the Z resonance is used to study the MS performance in

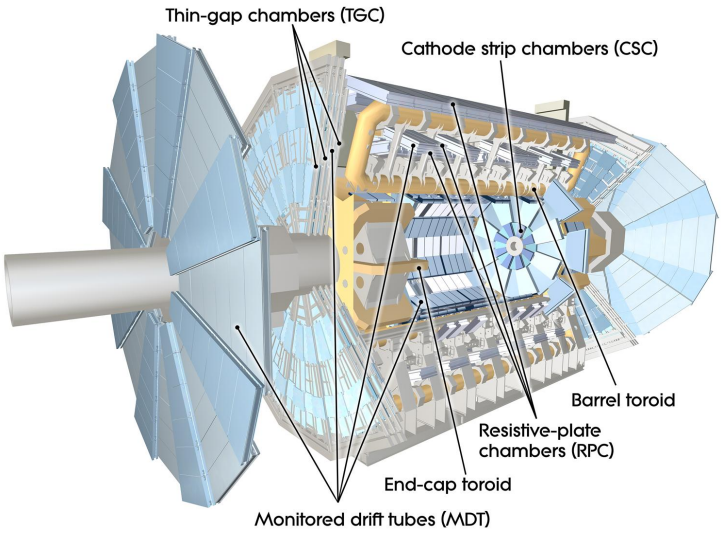


Figure 3.13: Muon detector spectrometer layout [19].

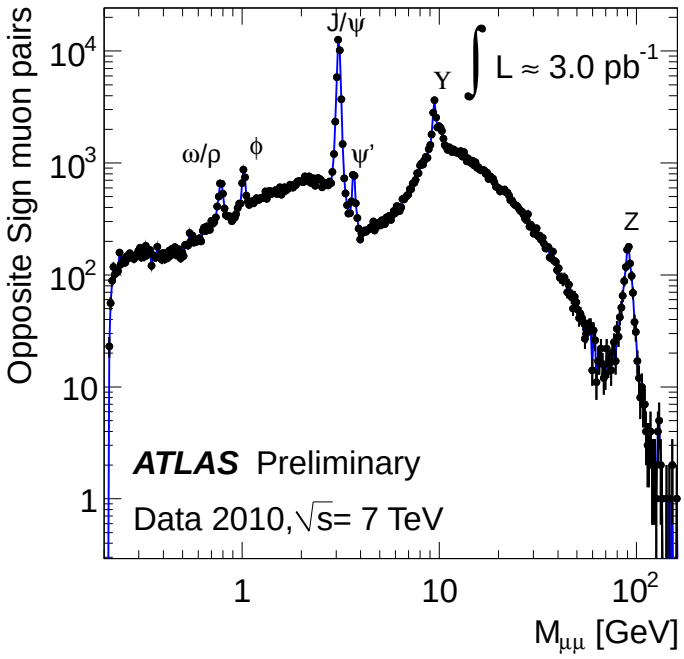


Figure 3.14: Distribution of $\mu^+\mu^-$ invariant mass reconstructed with two opposite charge muons having $p_{\mathrm{T}} > 15$ GeV and originating from a common vertex.

high p_{T} region.

3.4.4 Magnet system

The magnet system (Figure 3.13) is used to bend charged particles so that their momenta can be measured. It contains two parts: the inner solenoid and the outer toroidal.

The inner solenoid produces a 2 Tesla magnetic field surrounding the inner detector. This strong field can bend even very energetic particles to curve enough so that their momenta can be determined.

The outer toroidal magnetic field is produced by eight very large air-core super-conducting barrel loops and two end-caps, all situated outside the calorimeters and within the muon system. This magnetic field is 26 metres long and 20 metres in diameter.

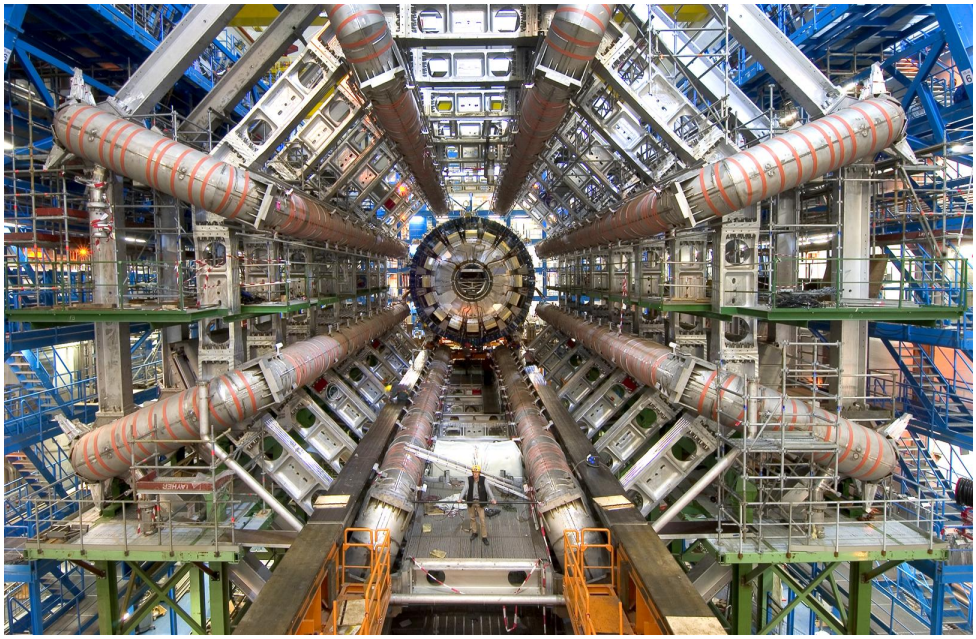


Figure 3.15: The ATLAS Toroid [19].

Chapter 4

E_{T}^{miss} from Hadronic Recoil in W, Z events

4.1 Introduction

We consider W and Z bosons signatures through their leptonic decay modes since these signatures have low backgrounds. In particular, the W boson decays into a charged lepton and a neutrino associated to the charged lepton ($W \to \ell\nu$) and the Z boson decays into two opposite charged leptons ($Z \to \ell\ell$) where charged leptons are electrons or muons. While charged leptons can be detected by the detector, the neutrino cannot be "seen" by the detector but can be measured as missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$on the transverse plane that is orthogonal to the beam direction.

In this chapter, we propose an algorithm for the reconstruction of missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ in $W \to \ell\nu$ events, starting from the reconstruction of the hadronic recoil and followed by a combination with charged leptons reconstructed as high transverse momentum objects [2]. The hadronic recoil originates from parton shower. It contains "hard" jets that balances the boson transverse momentum and "soft" components that includes particles produced by spectator quarks (known as "underlying event") and other proton proton collisions in the same bunch crossing, electronic noise, and residual energy in the detector from previous bunch crossings (known as "pileup") [4] (see also [30] for further descriptions of hadronic recoil components). In this algorithm, the hadronic recoil is considered as a fundamental quantity to be calibrated separately from the rest of the event, allowing transparent comparisons between different processes.

The hadronic recoil is calculated from topo-clusters deposited in the calorimeters where the topo-clusters are reconstructed by topological algorithm described in Section 3.4.2. In the hadronic recoil

calculation, the double counting of calorimetric energy is avoided by excluding some cone(s) around the lepton(s), and correcting the small bias induced by this exclusion by filling the cone with an estimate of the expected underlying event energy in this cone.

The cone size parameter allows a check of the stability of the performance of the method and allows to estimate the systematic uncertainty. The cone size affects, for example, the fraction of bremsstrahlung photons that are excluded from the hadronic recoil calculation, and the amount of underlying event energy needed to compensate for the excluded energy. These are non-trivial effects that need to be monitored precisely.

Another advantage over the standard method of calculating E_{T}^{miss} (described in [25] and [31]) is that the leptons can be used with optimal reconstruction and calibration. In particular, the muon momentum combination of the Inner Detector and the Spectrometer can be used, and the muon energy deposit in the calorimeters does not affect the results of the algorithm.

In this chapter, the algorithm is performed on simulation to study the dependence of the hadronic recoil response on the event dynamics, and determine the resolution correction to cross-section measurements, templates for W mass measurements.

This chapter is organized as follows. In Sections 4.2 and 4.3, all necessary concepts as well as notations, data and event selection used are explained. In Section 4.4, we recommend a method to calculate $E_{\mathrm{T}}^{\mathrm{miss}}$ in $W \to e\nu$ events from the reconstructed electron and the hadronic recoil. Following this calculation, in $W \to e\nu$ events, the resolution correction for $E_{\mathrm{T}}^{\mathrm{miss}}$ can be deduced from resolution corrections of the electron and the hadronic recoil. A resolution correction procedure for the recoil in $W \to e\nu$ events is presented in Section 4.5, where, first of all, to validate the method, we try to use the detector response obtained from simulation in $W \to e\nu$ events. In section 4.6, we apply the method to W mass and W cross-section measurements. In section 4.7, we show the agreement between the recoil response in $W \to e\nu$ events and in $Z \to ee$ events that allows for a data-driven detector response to the hadronic recoil in W events.

4.2 Conventions and definitions

In the $W \to \ell\nu$ events, the final state contains a lepton and a hadronic recoil reconstructed from the rest of the event. The energy and momentum of these objects (neglecting mass) are denoted as follow:

- Lepton: E^{ℓ}, $\overrightarrow{p}^{\ell}$.

- Cluster: E^{clus}, $\overrightarrow{p}^{clus}$. "Clusters" are considered as topo-clusters reconstructed from topological algorithm (described in Section 3.4.2) excluding the topo-clusters that match to leptons.

- Hadronic recoil:

$$\overrightarrow{p_{\mathrm{T}}}^{R} = \sum_{i=1}^{N_{clus}} \overrightarrow{p_{\mathrm{T}}}^{clus_i}. \tag{4.1}$$

  The *hadronic recoil* here is actually the "raw" hadronic recoil which is defined as the vector sum of momenta of all clusters on the transverse plane. Our "final" hadronic recoil calculated is corrected for the underlying energy in the cone used to remove electron clusters in the hadronic recoil calculation (will be discussed in Section 4.4.2).

- Hadronic activity:

$$\Sigma E_{\mathrm{T}} = \sum_{i=1}^{N_{clus}} E_{\mathrm{T}}^{clus_i}. \tag{4.2}$$

  The *hadronic activity* ΣE_{T} is the scalar sum of transverse energies of all clusters.

- Missing transverse energy and momentum from the hadronic recoil:

$$\overrightarrow{\not{p}_{\mathrm{T}}} = -\left(\overrightarrow{p_{\mathrm{T}}}^{\ell} + \sum \overrightarrow{p_{\mathrm{T}}}^{R}\right). \tag{4.3}$$

  Our missing transverse energy and momentum (denoted as $\not{E}_{\mathrm{T}}$ and $\overrightarrow{\not{p}_{\mathrm{T}}}$) are calculated from the hadronic recoil by using the transverse momentum imbalance of particles in the final state of $W \to \ell\nu$ events as given in formula 4.3.

- Missing transverse energy from standard calculation $\not{E}_T^{std}$ in ATLAS [25]:

$$\not{E}_{x(y)}^{std} = E_{x(y)}^{miss,calo} + E_{x(y)}^{miss,muon}, \tag{4.4}$$

  where $E_{x(y)}^{miss,calo}$ (calorimeter term) is x or y component of the total energy deposited in the calorimeters and $E_{x(y)}^{miss,muon}$ (muon term) is x or y component of the energy of muons measured in range of pseudo-rapidity $|\eta| < 2.7$.

  The calorimeter term is calculated using calorimeter information within range of pseudo-rapidity $|\eta| < 4.5$ from cell-based algorithm:

$$E_x^{miss,calo} = -\sum_{i=1}^{N_{cell}} E_i sin(\theta_i) cos(\phi_i); \quad E_y^{miss,calo} = -\sum_{i=1}^{N_{cell}} E_i sin(\theta_i) sin(\phi_i), \tag{4.5}$$

where E_i, θ_i and ϕ_i are the cell energy, the polar angle and the azimuthal angle respectively. N_{cell} is the number of cells inside three-dimensional topo-clusters.

The muon term includes contributions from isolated and non-isolated muons [26].

In this chapter, the $\not{E}_T^{std}$ from the standard calculation is used as a baseline for a cross-check of the performances of the $\not{E}_{\mathrm{T}}$ calculated by the proposed method.

- Transverse mass of W boson:

$$m_T = \sqrt{2p_{\mathrm{T}}^{e}p_{\mathrm{T}}^{\nu}(1 - cos(\phi^{e} - \phi^{\nu}))}. \tag{4.6}$$

- **Bias and resolution**: *bias* and *resolution* of a quantity are defined as the mean and the RMS respectively of the distribution of the difference between that quantity from reconstruction and from truth.

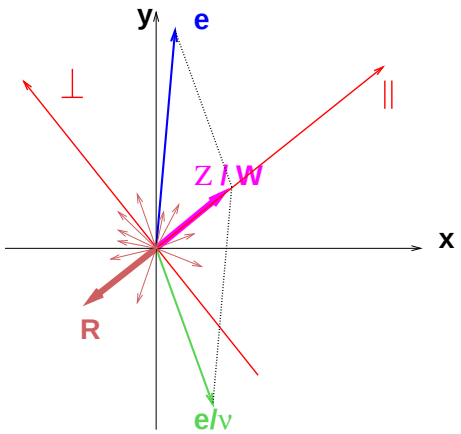


Figure 4.1: Definition of parallel and perpendicular axes in W and Z events.

The corresponding energy and momentum on the transverse plane are denoted by E_{T}, p_{T} and similarly for those on the x, y, z axes and on parallel ($\parallel$), perpendicular ($\perp$) axes (as expressed in Figure 4.1). The parallel and perpendicular axes are defined to be orthogonal to each other on the transverse plane in which the parallel axis is chosen to be the direction of true boson vector. The component of the hadronic recoil is maximal on the parallel axis while minimal on the perpendicular axis.

4.3 Data and event selection

This study relies on fully simulated samples of $W \to e\nu$ containing around 100k events and $Z \to ee$ containing around 200k events. The samples are generated using PYTHIA, and require at least one lepton within the detector acceptance.

In $W \to e\nu$ events we select exactly one electron in the final state passing "medium" requirements [2] with $E_T^e > 20$ GeV and $|\eta^e| < 1.37$ or $1.52 < |\eta^e| < 2.47$. And in the $Z \to ee$ events, the final state is selected to have exactly two opposite charge electrons passing medium requirements with $E_T^e > 20$ GeV and $|\eta^e| < 1.37$ or $1.52 < |\eta^e| < 2.47$.

Finally, since the aim of this chapter is limited to the technical validation of the proposed method, we do not consider any backgrounds and triggers.

4.4 $E_{\mathrm{T}}^{\mathrm{miss}}$ from the hadronic recoil in $W \to e\nu$ events

Here we outline a method of studying E_{T}^{miss} based on the hadronic recoil in W events as given in equation 4.3. To demonstrate the method, all plots as well as numbers in this section are obtained in $W \to e\nu$ events.

In this method of calculating E_{T}^{miss} from the hadronic recoil, we first discuss the need of an exclusion cone, and its impact on the performance in Subsection 4.4.1. Then in Subsection 4.4.2 we discuss the effect of replacing the energy. Finally in Subsection 4.4.3 we discuss advantages an limits of the proposed method in a comparison with the standard one.

4.4.1 Avoiding electron double counting

If the hadronic recoil is computed by all topo-clusters deposited in the calorimeters, $\overrightarrow{p_{\mathrm{T}}}^e$ of the electron will be counted twice in equation 4.3. The effect of this is shown in Figure 4.2 where the red line is the distribution of $(\not{E}_{\mathrm{T}}{}^{reco} - \not{E}_{\mathrm{T}}{}^{true})$, the difference between the reconstruction and the truth of $\not{E}_{\mathrm{T}}$ calculated from equation 4.3, and the blue line is that of $\not{E}_{\mathrm{T}}{}^{std}$ used as a criteria to evaluate the performance of $\not{E}_{\mathrm{T}}$. The blue line having an almost zero mean value as expected while the red line having a mean value of about 33 GeV comes from the fact that the electron is counted in the sum of all topo-clusters. Thus the electron is counted twice in the formula 4.3 creating a big bias in $\not{E}_{\mathrm{T}}$ calculation. This is called the "double counting" problem. To solve this problem we exclude all clusters in the cone $\triangle R$ around the

electron cluster.

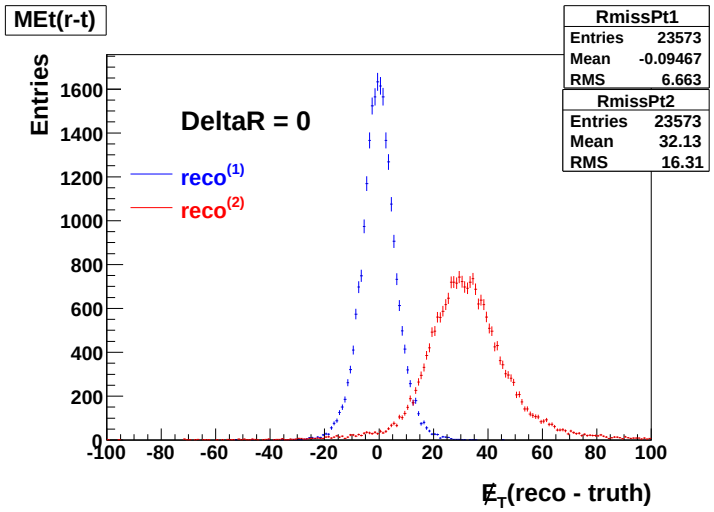


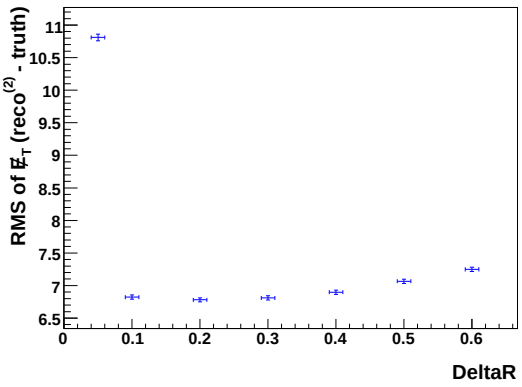
Figure 4.2: Distribution of the difference from the true $\not{E}_{\mathrm{T}}$ of the reconstructed $\not{E}_{\mathrm{T}}$ ("$reco^{(2)}$") and the reconstructed $\not{E}_T^{std}$ ("$reco^{(1)}$").

The variation of the resolution and the bias of $\not{E}_{\mathrm{T}}$ as a function of the cone size $\triangle R$ are shown in Figure 4.3. Figure 4.3(a) indicates that, with a cone size $\triangle R = 0.05$ the double counting problem is still present while with a cone size above 0.1, the electrons are completely excluded from the clusters but the $\not{E}_{\mathrm{T}}$ resolution degrades when more and more information is removed from the recoil calculation. Figure 4.3(b) shows that the larger the cone size $\triangle R$ the more underestimated the hadronic recoil. This is a consequence of removing the underlying energy within the cone $\triangle R$.

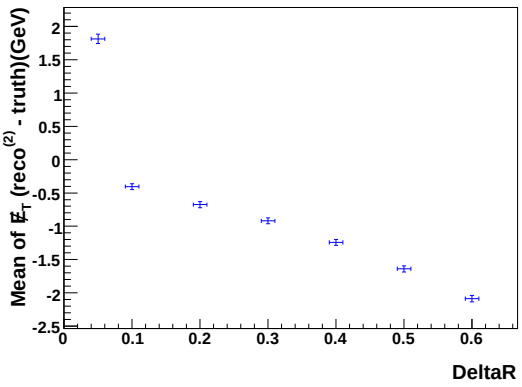
So the procedure of calculating the hadronic recoil which is graphically presented in Figure 4.4 includes three steps: first summing vectorially $\overrightarrow{p_{\mathrm{T}}}$ of all topo-clusters deposited in calorimeters; then removing a big enough cone of clusters around the electron cluster to avoid the double counting problem; and finally compensating the underlying event excluded in the cone to correct for the recoil energy. The recoil energy correction is especially needed for the recoil behavior in Z and W to be similar since the number of exclusion cones which is equal to the number of leptons in the final states are different in W and Z samples. In the next section we will discuss in detail how to replace this energy.

4.4.2 Recoil energy replacement

To recover the recoil energy and thus $\not{E}_{\mathrm{T}}$ we need to compensate the underlying energy excluded.



(a) Dependence of $\not\!\!E_T$ resolution on $\triangle R$. When the size of the cone is big enough (from 0.1), the resolution of the $\not\!\!E_T$ is stable saying that the electron is excluded completely.



(b) Dependence of $\not\!\!E_T$ scale bias on $\triangle R$.

Figure 4.3: Dependence of $\not\!\!E_T$ resolution (4.3(a)), and $\not\!\!E_T$ scale bias (4.3(b)) on $\triangle R$.

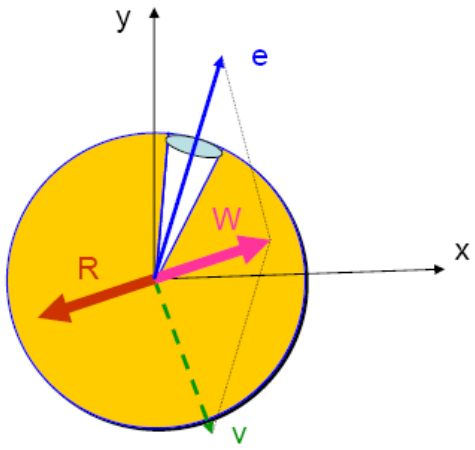
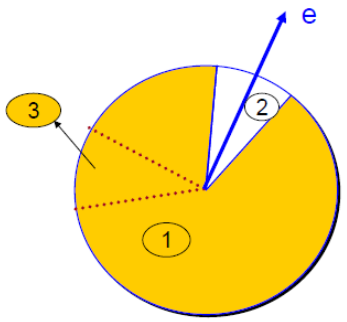


Figure 4.4: Graphical representation of the hadronic recoil calculation, on the example of the $W \rightarrow l\nu$ final state.

For each event, our hadronic recoil contains two components: the hadronic recoil $\overrightarrow{p_T}^{R_1}$ in zone 1 of clusters outside the cone of removed clusters around the electron (zone 2) and the recoil $\overrightarrow{p_T}^{R_2}$ recovered in zone 2,

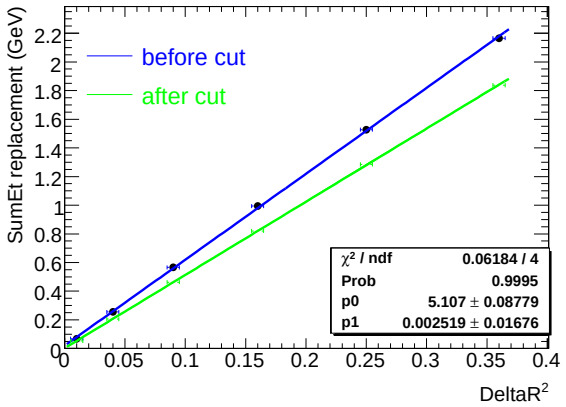
$$\overrightarrow{p_T}^{R} = \overrightarrow{p_T}^{R1} + \overrightarrow{p_T}^{R_2}, \tag{4.7}$$



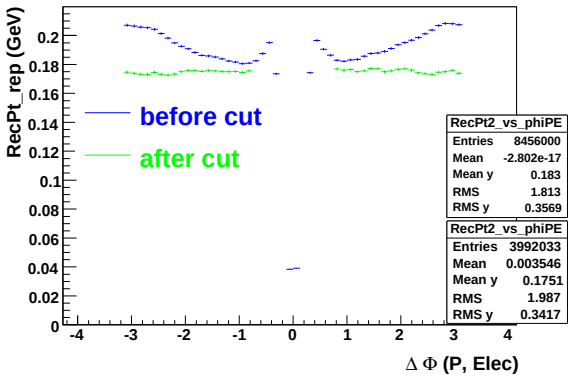
where $\overrightarrow{p_T}^{R_1}$ is a vector sum of transverse momenta of all clusters in zone 1

$$\overrightarrow{p_T}^{R_1} = \sum_{i=1}^{Nclus1} \overrightarrow{p_T}^{clus_i}, \tag{4.8}$$

and $\overrightarrow{p_T}^{R_2}$ is determined by the energy measured in a cone of the same size $\triangle R$ (named zone 3) around a point P which is outside the exclusion cone and at the η^e of the electron.



(a)



(b)

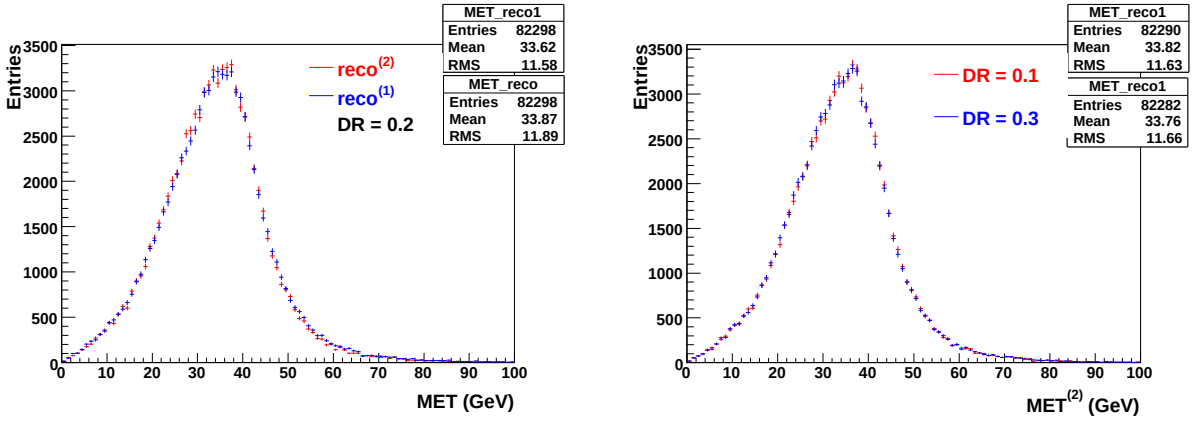
Figure 4.5: 4.5(a) Hadronic activity ΣE_T in zone 3 versus the surface of the zone. The blue line corresponds to the case of choosing the placement of zone 3 randomly in zone 1 while the green line is drawn when choosing the placement of zone 3 which is separated in $\triangle\phi$ from both electron and hadronic recoil of zone 1; 4.5(b) p_T of the recoil in zone 3 versus $\triangle\phi(P,Elec)$ in the case of $\triangle R = 0.2$.

With the idea of replacing the underlying energy, the ΣE_T of clusters in zone 3 versus $(\triangle R)^2$, the surface of the zone, should be linear as expressed in Figure 4.5(a). There, the blue line is obtained when the point is taken randomly in zone 1. This line doesn't pass the origin of coordinate system implying that when the surface is zero the hadronic activity in zone 3 is not zero but over-estimated by about 10% of the energy replacement. To explain this we plot the recoil p_T in zone 3 versus the $\triangle\phi(P,Elec)$

between the point P and the electron when $\triangle R = 0.2$ as shown in Figure 4.5(b). The blue line in Figure 4.5(b) shows that the recoil energy in zone 3 depends on the position of the point defining this zone in distance from positions of the electron and the recoil $\overrightarrow{p_{\mathrm{T}}}^{R_1}$. The recoil energy in zone 3 increases when the point is near the electron (because of the Bremsstrahlung) and the recoil $\overrightarrow{p_{\mathrm{T}}}^{R_1}$.

To avoid these effects we choose a placement of zone 3 which is separated in $\triangle\phi$ from both electron and hadronic recoil of zone 1. The result of this is plotted by the green lines in Figure 4.5. In Figure 4.5(a), the green line is linear and passes the origin of the coordinate system as expected. In Figure 4.5(b), where the point is chosen randomly further away from the electron than $\pi/4$ and the recoil $\overrightarrow{p_{\mathrm{T}}}^{R_1}$ further than $\pi/3$ in the case of $\triangle R = 0.2$, the recoil energy (about 180 MeV on average) in zone 3 is stable. The difference between the blue line and the green line in the range of $\triangle\phi(P, Elec) < \pi/3$ indicates a contamination of the recoil by the Bremsstrahlung energy that needs to be corrected for. The size of this effect is about 40 MeV.

After choosing a good position of the point to define zone 3, we calculate the vector $\overrightarrow{p_{\mathrm{T}}}$ of the recoil in this cone then rotate it to the place of zone 2 around the electron by the angle between the point and the electron to obtain the recoil replacement $\overrightarrow{p_{\mathrm{T}}}^{R_2}$.



(a) (b)

Figure 4.6: 4.6(a) The agreement between the distribution of $\not{E}_{\mathrm{T}}$ calculated from the recoil ("$reco^{(2)}$") and the distribution of $\not{E}_{\mathrm{T}}{}^{std}$ ("$reco^{(1)}$") when $\triangle R = 0.2$; 4.6(b) The stability of the $\not{E}_{\mathrm{T}}$ calculated from the recoil when changing the size of exclusion cone $\triangle R$, the red line: $\triangle R = 0.1$ and the blue line: $\triangle R = 0.3$.

The $\not{E}_{\mathrm{T}}$ is recalculated from the recoil by using formula 4.3 event by event. The results of its dis-

tribution are shown in Figure 4.6 where Figure 4.6(a) shows the agreement of the distribution of the $\not{E}_{\mathrm{T}}$ based on the recoil and that of the $\not{E}_{\mathrm{T}}{}^{std}$ in example of cone size $\triangle R = 0.2$ and Figure 4.6(b) exhibits the stability of the $\not{E}_{\mathrm{T}}$ distribution with respect to different $\triangle R$. However $\triangle R = 0.2$ gives the best resolution for $\not{E}_{\mathrm{T}}$ as seen in Figure 4.3(a). In addition, $\triangle R = 0.2$ is an expectation cone size which is associated with the size of sliding window cluster. Therefore in applications we will use this exclusion cone size.

We use this method of reconstructing $\not{E}_{\mathrm{T}}$ based on the hadronic recoil which is so-called "hadronic recoil" algorithm rather than use the $\not{E}_{\mathrm{T}}{}^{std}$ since it still gives the similar performance with the $\not{E}_{\mathrm{T}}{}^{std}$ while it has a good separation between lepton and recoil so it is transparent between μ and e channels. With the energy replacement done, it gives a better resolution for hadronic recoil as well as E_{T}^{miss}. Additionally, it is possible to calibrate from Z to W events as the recoil behaviors in Z and W are similar. More advantages of the method will be discussed in the next section.

4.4.3 Advantages and limitations of the method

Advantages

The fundamental quantity in our algorithm is the hadronic recoil. We therefore compare directly the performance of the hadronic recoil from the proposed algorithm to that from the standard calculation. In the following examples, we will see our hadronic recoil has an improvement in bias and resolution compared to the standard one.

Firstly, in Figure 4.7, the correlation between reconstructed and true p_T^W measured as the hadronic recoil is displayed as obtained from both calculations, in the $W \to e\nu$ channel. The hadronic recoil algorithm displays many outliers, *i.e.* events where the response is clearly off-diagonal.

Secondly, we display the impact of the lepton-induced calorimetric energy, and of the underlying event energy substitution, in the $W \to \mu\nu$ channel. As explained above, in case lepton-induced calorimetric energy is present in the event, it should properly be subtracted from the estimation of the recoil. If this is not done, the recoil is biased towards the direction of the lepton. As a result, the measured value of $p_\perp(W)$ is maximally biased when the lepton and W directions are aligned; the bias is 0 when their directions are orthogonal. As can be seen from the left plot in Figure 4.8, the size of the effect in the muon channel is about 2 GeV (the red line). Replacing the standard calculation by the $\overrightarrow{p_{\mathrm{T}}}{}^{R1}$ which is the raw recoil in our calculation , the bias reduces to a few hundred MeV (the green line). It becomes identically 0 when the underlying correction is applied (the blue line). Correspondingly, the overall resolution

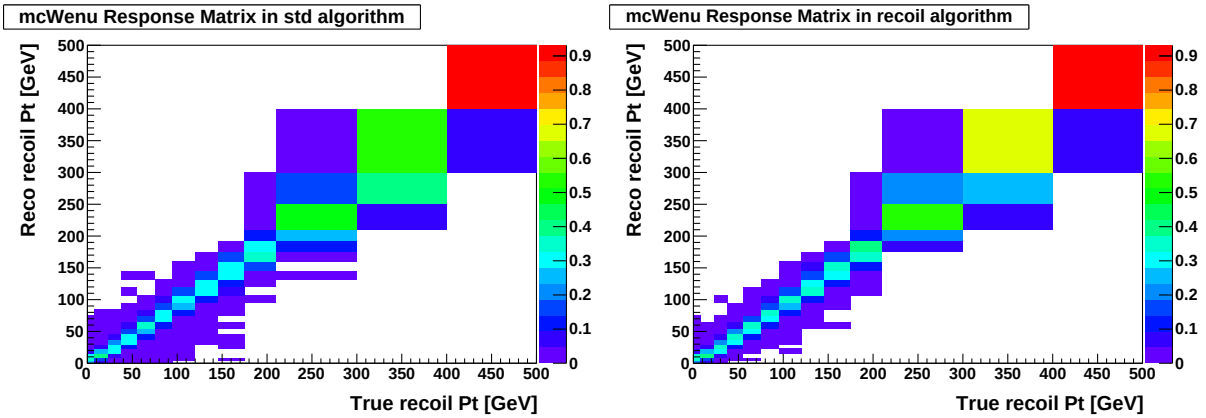


Figure 4.7: Correlation histogram for $p_T^{reco}(W)$ vs. $p_T^{true}(W)$. Left : standard algorithm. Right : hadronic recoil algorithm.

is significantly improved as shown on the right plot in Figure 4.8.

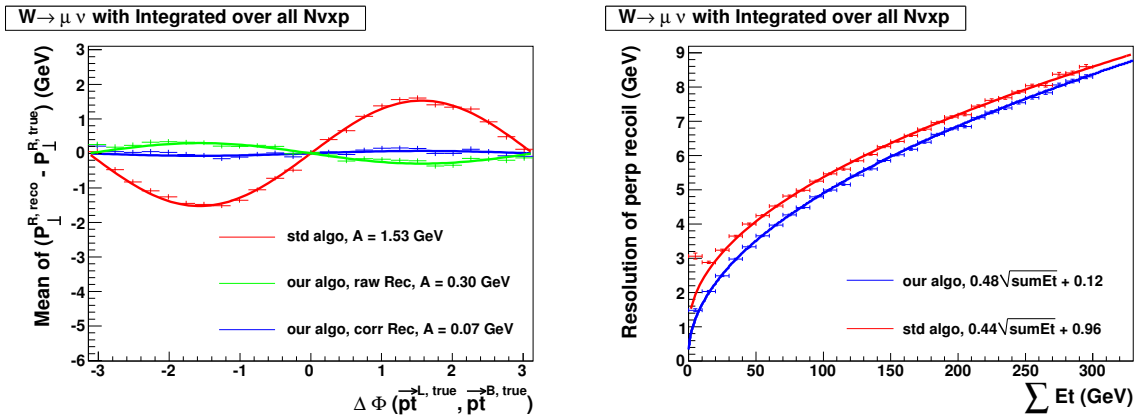


Figure 4.8: Left : average value of $p_\perp^{reco}(W) - p_\perp^{true}(W)$, as a function of the angle between the lepton and the W boson transverse momenta, for the standard algorithm, and for the hadronic recoil algorithm. Right : resolution obtained from the standard algorithm, and from the hadronic recoil algorithm.

Limitations

The first limitation of the hadronic recoil algorithm is that in its present form, the algorithm implicitly assumes isolated leptons, as each associated cone is replaced by an underlying event estimate, sampled from a quiet region in the same event. When applied to jet events, where identified leptons are mostly not isolated, this procedures biases the event topology, replacing part of the jet by underlying event. Hence,

in the next chapter of measuring the p_T^W in data, the event selection and background subtraction still rely on the standard $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation; after these steps, the transverse momentum spectrum is reconstructed and unfolded from the hadronic recoil calculation described in the previous sections.

The second limitation of the hadronic recoil algorithm is that in events where a hard, large-angle photon is produced in the W decay and reconstructed, this photon is attributed to the recoil. As it is difficult to reliably identify FSR photons among the hadronic activity, this is an intrinsic limitation, which we mention here for reference and will be discussed in the next chapter.

4.5 Resolution corrections

Recoil response is difficult to model, both from detector and physics point of view [4]. When we compare simulation to data, we will first have to improve the simulation to better match the data, and then only perform the physics measurement. The first step consists of these resolution corrections. In this chapter, we consider the MC truth as our simulation to be corrected; and the full simulation as our data. Example of measurements where this can be applied are cross-section and W mass.

4.5.1 Resolution corrections for the recoil in $W \to e\nu$ events

To deduce resolution corrections for $\not{E}_{\mathrm{T}}$ and m_T quantities from the hadronic recoil, we need to correct the resolution of the vector recoil $\overrightarrow{p_{\mathrm{T}}}^R$ or of its components resolved along axes. The plot in Figure 4.9 shows that it is possible to make independent resolution corrections for two components of the recoil along parallel ($p_{\parallel}^R$) axis and perpendicular axis ($p_{\perp}^R$) since they have no correlation.

Resolution corrections for some MC truth quantity can be done by adding (event by event) to the truth quantity a random number which has a Gaussian form

$$smeared\ \ truth = truth + Rand.Gaus(\Delta(\alpha), \sigma(\beta)). \tag{4.9}$$

This Gaussian random number is governed by two parameters (so-called *"smearing parameter"*): the first one, denoted as Δ, is the scale bias of data comparing to simulation depending on some variable(s) called α, and, the second one, denoted as σ, is the resolution of the difference between data and simulation depending on some variable(s) β. The variables, on which these parameters depend, are chosen among true quantities since we are correcting the MC truth. The effects of these parameters on a true

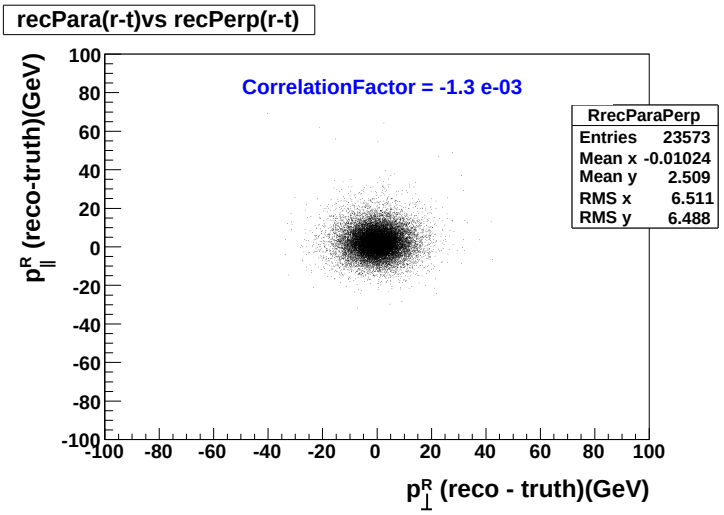


Figure 4.9: Correlation histogram for $p_{\parallel}^{R}(reco-truth)$ vs. $p_{\perp}^{R}$ (reco - truth).

distribution were studied in [36]. The quantity obtained after this correction, so-called *"smeared truth"*, is expected to have the same behavior as that in data.

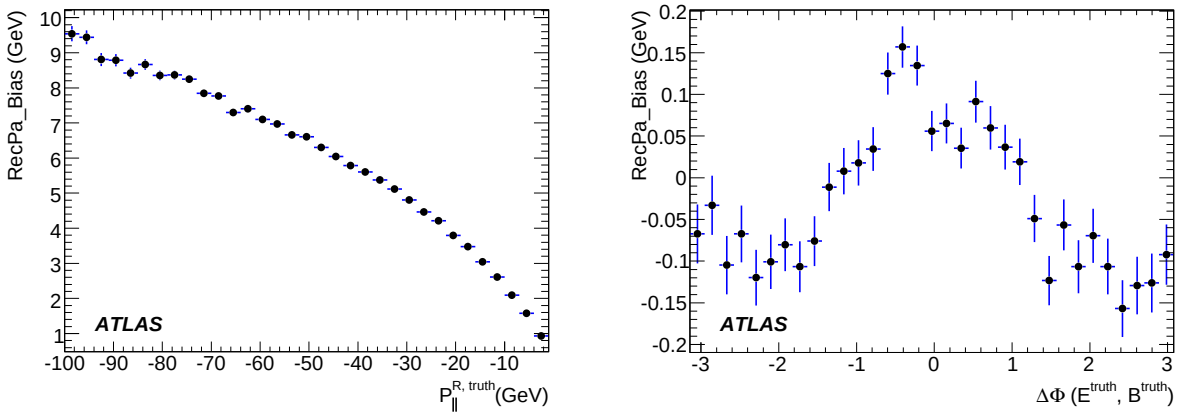
An important step of the resolution corrections procedure is to find sensitive variable(s) for these parameters.

Scale bias and resolution of the recoil

The recoil bias along parallel axis versus $p_{\parallel}^{R,truth}$ is shown in Figure 4.10(a). The origin of this bias is unseen hadrons. After correcting event by event for this, we still observe a bias versus $\triangle\phi(\overrightarrow{p_{\mathrm{T}}}^{e,truth}, \overrightarrow{p_{\mathrm{T}}}^{W,truth})$, defined as the opening angle between vector momenta of true electron and true W boson on the transverse plane, as shown in Figure 4.10(b). This is a result of the imperfect energy replacement procedure and the Bremsstrahlung energy contamination of the recoil discussed in section 4.4.2.

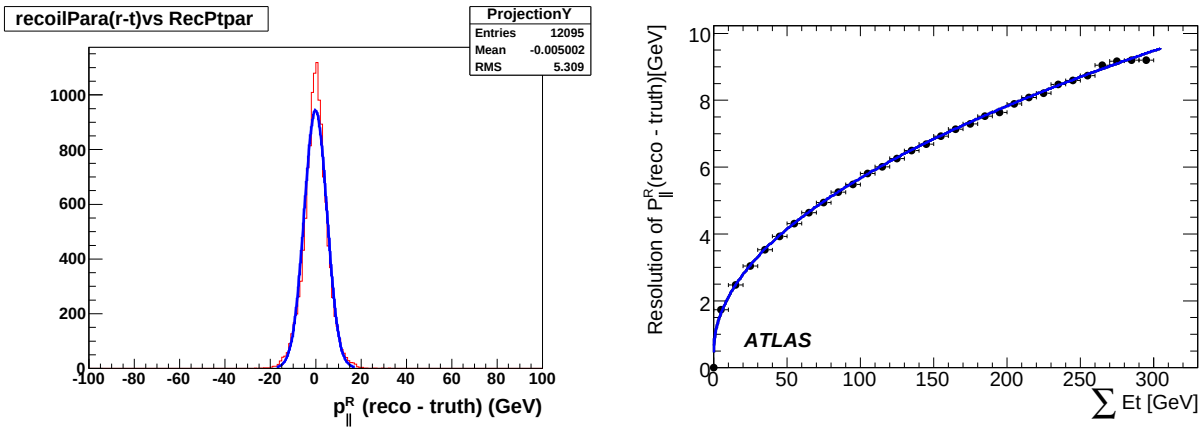
It is obvious that $p_{\parallel}^{R,truth}$ and $\triangle\phi(\overrightarrow{p_{\mathrm{T}}}^{e,true}, \overrightarrow{p_{\mathrm{T}}}^{W,true})$ are sensitive variables for the bias of parallel recoil. Additional to the $p_{\parallel}^{R,truth}$ variable, the $\triangle\phi(\overrightarrow{p_{\mathrm{T}}}^{e,true}, \overrightarrow{p_{\mathrm{T}}}^{W,true})$ variable is needed for the recoil resolution correction to obtain a better result in correcting $\not{E}_{\mathrm{T}}$ which is recomputed from the recoil and the electron as mentioned before.

The resolution behavior of the recoil along the parallel axis as shown in Figure 4.11(b) is the sigma of Gaussian fit of $(p_{\parallel}^{R,reco} - p_{\parallel}^{R,truth} - bias)$ slice distribution in bins of *hadronic activity* ΣE_{T} (Figure 4.11(a) shows an example of the fit in a bin of ΣE_{T}). In this correction procedure, the Gaussian fits will



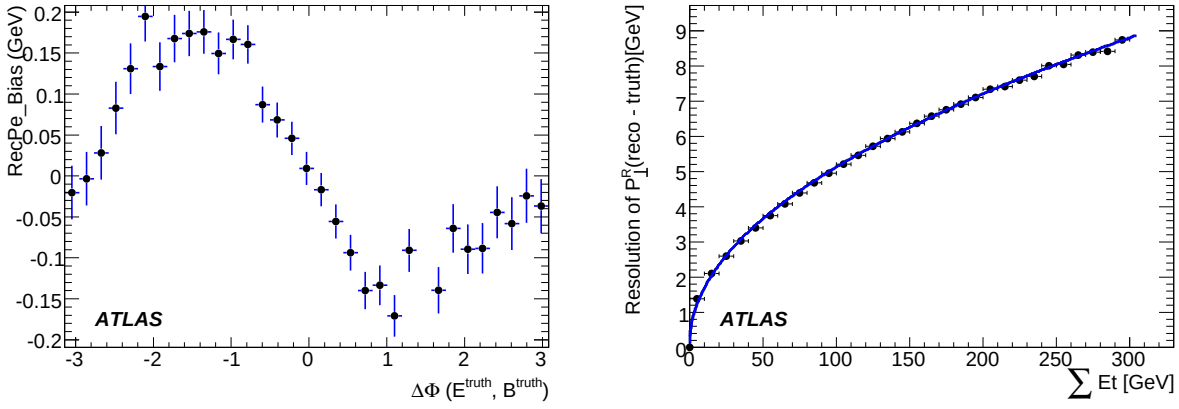
(a) (b)

Figure 4.10: 4.10(a) Bias of the recoil along parallel axis versus $p_{\parallel}^{R,truth}$ in $W \to e\nu$ events; 4.10(b) Bias of the recoil along parallel axis versus $\triangle\phi$ between directions of the true electron and the true boson on the transverse plane in $W \to e\nu$ events after correcting the bias in bins of $p_{\parallel}^{R,truth}$ (as shown in 4.10(a)).



(a) (b)

Figure 4.11: 4.11(a) Gaussian fit of $p_{\parallel}^{R}(reco - truth)$ slice at a bin of ΣE_{T} after correcting the bias; 4.11(b) Resolution behavior of the recoil along parallel axis in bins of ΣE_T in $W \to e\nu$ events.



(a) (b)

Figure 4.12: (4.12(a)) Bias of the perpendicular recoil versus $\triangle\phi$ between the true electron and the true boson; (4.12(b)) Resolution of the perpendicular recoil versus $\sum E_T$ in W events.

give better results since the smearing is chosen to have a Gaussian form. In addition, here, the resolution is extracted after correcting the scale bias in order to separate the effect of the scale bias and that of the resolution (see [36] for more detail explanation).

For smearing parameters of the recoil along perpendicular axis, we can find similar variables as for the parallel recoil one expects. The *hadronic activity* ΣE_{T} is still a good variable for parameterizing the resolution as shown in Figure 4.12(b). However, as a consequence of choosing the parallel axis to be the direction of the true boson vector $\overrightarrow{p_{\mathrm{T}}}^{B,truth}$ on the transverse plane, the perpendicular component of the truth recoil $p_{\perp}^{R,truth} = 0$, so there is only one sensitive variable, that is $\triangle\phi(\overrightarrow{p_{\mathrm{T}}}^{e,true}, \overrightarrow{p_{\mathrm{T}}}^{W,true})$, to parameterize the bias of the perpendicular recoil.

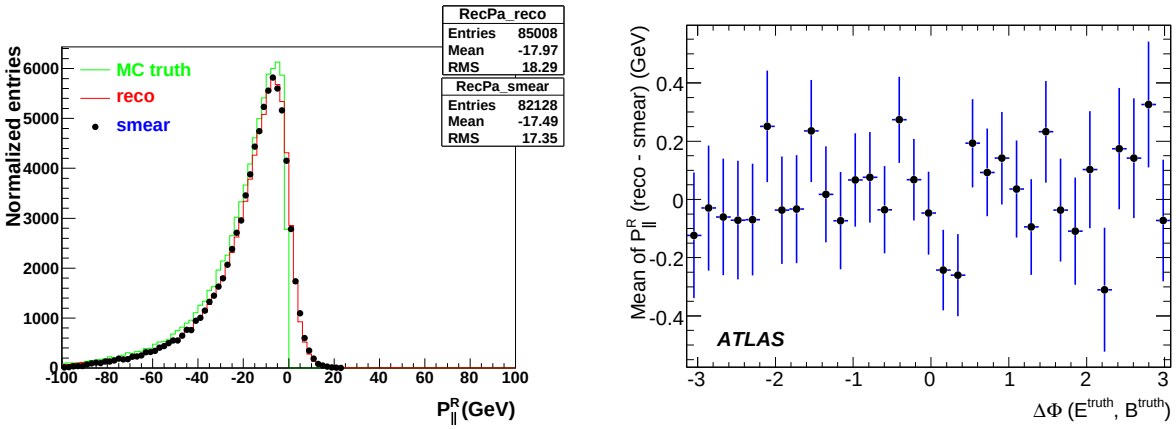
Results of correcting the recoil

Once sensitive variables to explore the scale bias and resolution are found, the smeared quantities can be obtained as follow, for example, with the recoil along parallel axis:

$$p_{\parallel}^{R,smear} = p_{\parallel}^{R,truth} + Rand.Gaus\left(\Delta(p_{\parallel}^{R,truth}, \triangle\phi(\overrightarrow{p_{\mathrm{T}}}^{e,true}, \overrightarrow{p_{\mathrm{T}}}^{W,true})), \sigma(\Sigma E_{\mathrm{T}})\right) \tag{4.10}$$

Here we model the dependence of the scale bias on two variables. The formula 4.10 is applied event by event and the smeared distribution of the parallel recoil obtained is shown in Figure 4.13(a). To evaluate the compatibility between the smeared parallel recoil $p_{\parallel}^{R,smear}$ obtained after the correction and the reconstructed parallel recoil $p_{\parallel}^{R,reco}$, we plot the mean of the difference between them ($p_{\parallel}^{R,reco}$ –

$p_{\parallel}^{R,smear}$) in bins of $\triangle\phi(\overrightarrow{p_{\mathrm{T}}}^{e,true}, \overrightarrow{p_{\mathrm{T}}}^{W,true})$ as shown in Figure 4.13(b). The plot shows a bias of $21 \pm$ 302 (MeV) implying that the resolution correction of parallel recoil has no bias within the statistical uncertainty. Or on the other way we can conclude that the $\not{p}_{\parallel}^{smear}$ distribution is smeared from the true distribution to be in a good agreement with the reconstructed distribution $\not{p}_{\parallel}^{reco}$ as expected.



(a) (b)

Figure 4.13: 4.13(a) Distributions of the true, reconstructed and smeared recoil along the parallel axis; 4.13(b) Evaluation of the agreement between the smeared and reconstructed recoil along the parallel axis: $\sim 21 \pm 302$ (MeV).

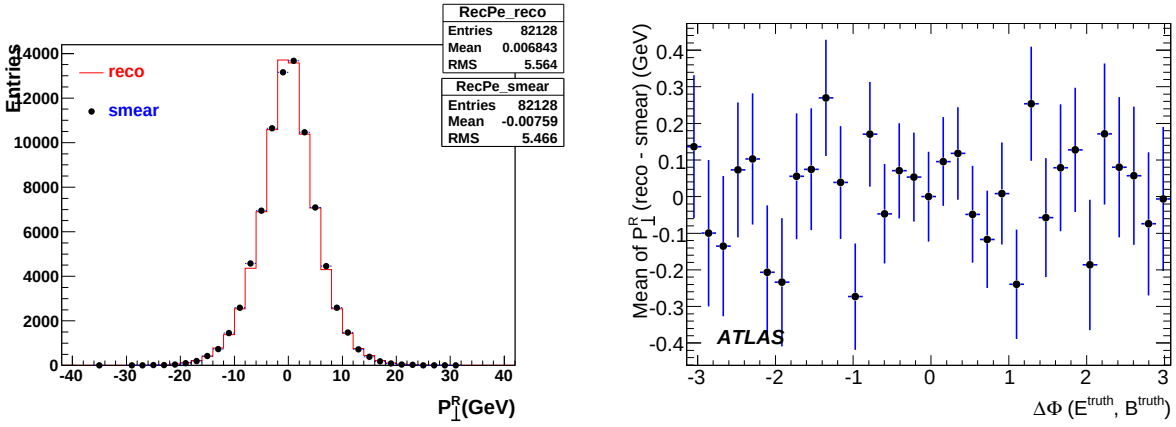
For the recoil along perpendicular axis, similarly, applying the formula 4.11 event by event to obtain the smeared distribution we can get the result as shown in Figure 4.14. The bias of resolution correction for the perpendicular recoil evaluated by Figure 4.14(b) is about 10 ± 285 (MeV).

$$p_{\perp}^{R,smear} = p_{\perp}^{R,truth} + Rand.Gaus\left(\Delta(p_{\perp}^{R,truth}), \sigma(\Sigma E_{\mathrm{T}})\right), \tag{4.11}$$

4.5.2 Resolution corrections for $\not{E}_{\mathrm{T}}$ and m_T

Resolution corrections of $\not{E}_{\mathrm{T}}$ can be computed from those of the recoil and the electron by using equation 4.3 and similarly for m_T by using equation 4.6. In this chapter, to see only effects of the recoil resolution corrections, we use directly the reconstructed electron instead of smeared electron when calculating smeared $\not{E}_{\mathrm{T}}$ and smeared m_T.

Figure 4.15(a) shows the agreement between the reconstructed and the smeared $\not{E}_{\mathrm{T}}$ spectra. This agreement is evaluated by plotting the distribution of bin by bin χ values defined as the difference be-



(a) (b)

Figure 4.14: 4.14(a) Distributions of the reconstructed and smeared recoil along the perpendicular axis (the true recoil along this axis is zero); 4.14(b) Evaluation of the agreement between them: $\sim 10 \pm 285$ (MeV).

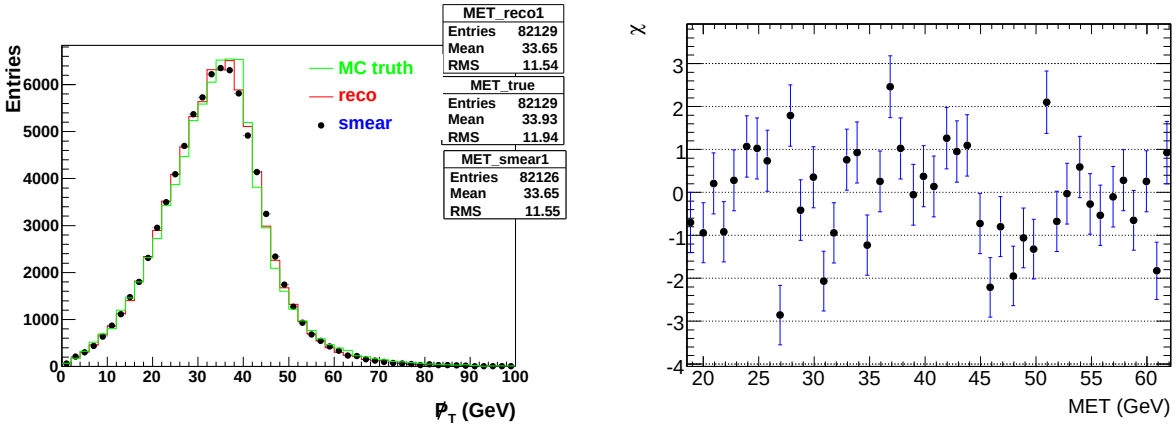
tween the reconstructed and the smeared $\not{E}_{\mathrm{T}}$ histograms divided by the square root of the sum of reconstructed and smeared uncertainty squared. Similarly, the result for the m_T distribution is shown in Figure 4.16. χ plots shown in Figures 4.15(b) and 4.16(b) indicate a small bias of $\not{E}_{\mathrm{T}}{}^{smear}$ and m_T^{smear} spectra obtained after resolution corrections comparing to those in data.

These results obtained in W events validate a good method for smearing and reconstructing E_{T}^{miss} and m_T based on the recoil. To achieve realistic results, the detector response to the recoil can be extracted from Z events.

Once the smeared distributions of $\not{E}_{\mathrm{T}}$ and m_T are in good agreement with the reconstructed distributions, we can reproduce the data by correcting the MC truth to measure precisely the W mass and W cross-section which is discussed in more details in the next section.

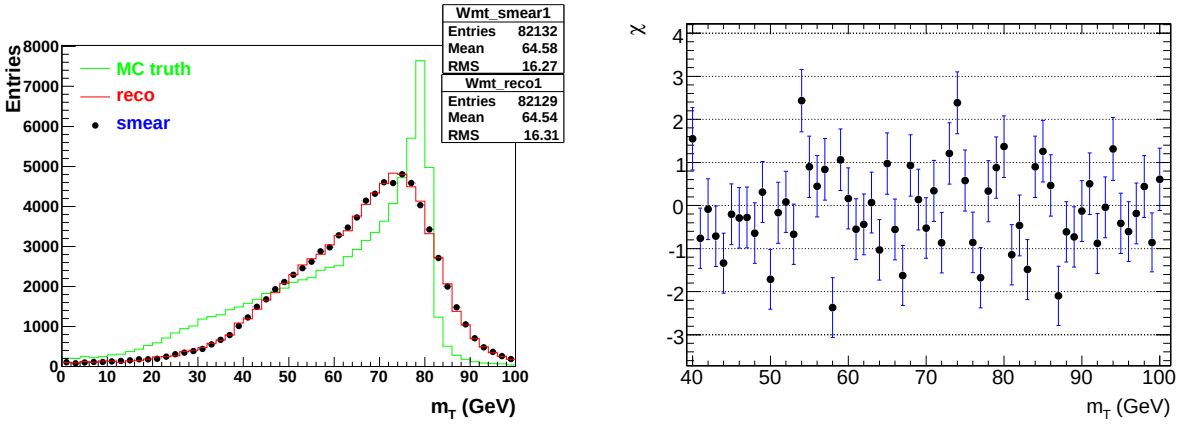
4.6 Applications of resolution corrections method

We apply the method to two measurements: W cross-section and W mass.



(a) (b)

Figure 4.15: 4.15(a) Distributions of the true, reconstructed and smeared $\not\!\!E_T$; 4.15(b) Bin by bin χ values distribution to evaluate the agreement between the smeared and reconstructed $\not\!\!E_T$ distributions.



(a) (b)

Figure 4.16: 4.16(a) Distributions of the true, reconstructed and smeared m_T; 4.16(b) Bin by bin χ values distribution to evaluate the agreement between the smeared and reconstructed m_T distributions.

4.6.1 W cross-section measurement

Experimentally, the total cross section σ_W of W production, measured from its leptonic decay for example, is determined as follows (backgrounds are neglected here):

$$\sigma_W = \frac{1}{BR(W \to \ell\nu)} \frac{1}{\int \mathscr{L} dt} \frac{N^{obs}}{A_W}, \tag{4.12}$$

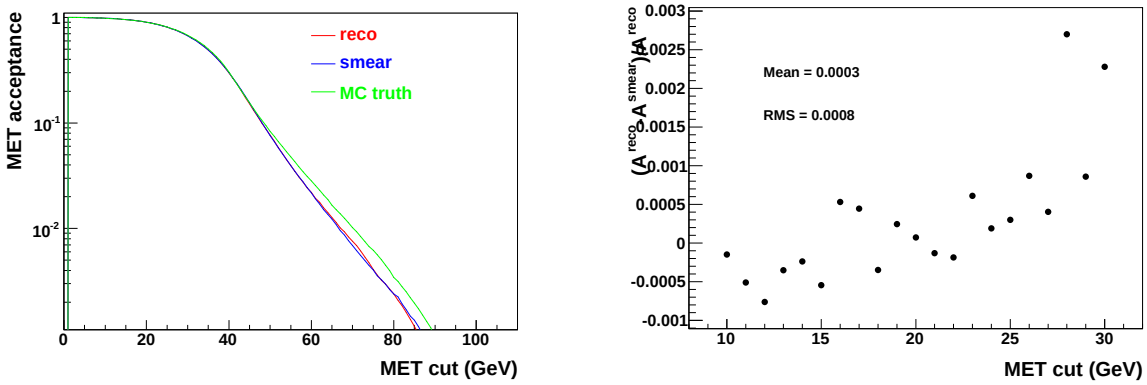
where $BR(W \to \ell\nu)$ is the branching ratio for the leptonic decay of W boson; N^{obs} is the number of observed signal events; $\int \mathscr{L} dt$ is the integrated luminosity and A_W is the acceptance of the signal, defined as the fraction of reconstructed signal events remaining after the selection cuts.

From Equation 4.12, to measure the W cross-section, we need to know the acceptance of leptons and E_{T}^{miss}. Using the resolution corrections method to reproduce the data we can calculate the acceptances of E_{T}^{miss} and lepton with some selection cuts. In addition, since the precision of W cross-section measurement depends on that of acceptance calculations, the uncertainty of resolution corrections will decide the precision of the W cross-section. Thus, to examine the resolution correction method mentioned in the previous section, we plot the acceptance of the reconstructed E_{T}^{miss} depending on various cut values of $\not{E}_{\mathrm{T}}{}^{reco}$ in comparing to that of smeared E_{T}^{miss}. As seen in Figure 4.17(a), the acceptance of $\not{E}_{\mathrm{T}}{}^{reco}$ at given cuts (varying from 1 GeV to 100 GeV) has almost the same value as that of $\not{E}_{\mathrm{T}}{}^{smear}$. The agreement between them is evaluated by $\left(A^{reco} - A^{smear}\right)/A^{reco}$ quantity. For example we plot the distribution of this quantity with cuts on $\not{E}_{\mathrm{T}}$ varying (step 1 GeV) from 10 GeV to 30 GeV in Figure 4.17(b) which shows a very good agreement (bias less than $(0.03 \pm 0.08)\%$) between the acceptance of $\not{E}_{\mathrm{T}}{}^{reco}$ and that of $\not{E}_{\mathrm{T}}{}^{smear}$.

Very good results (bias less than $(-0.08 \pm 0.09)\%$ with the cuts between 20 GeV and 60 GeV on m_T) are also obtained when doing similarly for the W transverse mass quantity (m_T) plotted in Figure 4.18, again validate the resolution corrections method.

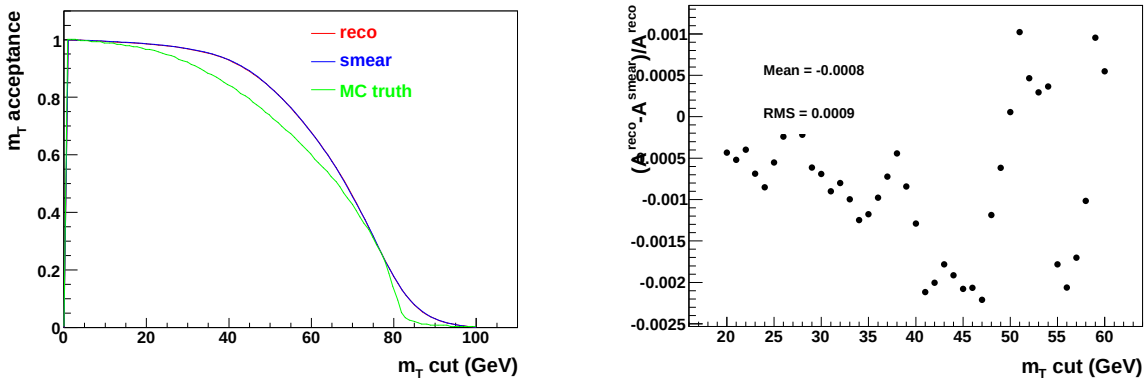
4.6.2 W mass measurement

The W mass can be measured by using the leptonic decay channel of W with the final state including a lepton and a neutrino where the neutrino is detected as missing transverse energy. Since the E_{T}^{miss} is determined only on the transverse plane, the W mass cannot be constructed as the invariant mass of the lepton and the neutrino. The W mass is thus measured by using other kinematic variables such as m_T, p_{T}^{ℓ} or E_{T}^{miss} with the template method as described in [71].



(a) (b)

Figure 4.17: 4.17(a) Acceptance of MC truth (green line) and that of reconstructed (red line), smeared (blue line) E_{T}^{miss} with the cuts varying from 1 GeV to 100 GeV; 4.17(b) $\left(A^{reco} - A^{smear}\right)/A^{reco}$ values at different cuts on $\not{E}_{\mathrm{T}}$ varying with step 1 GeV from 10 GeV to 30 GeV.



(a) (b)

Figure 4.18: 4.18(a) Acceptance of MC truth (green line) and that of reconstructed (red line), smeared (blue line) m_T with the cuts varying from 1 GeV to 100 GeV; 4.18(b) $\left(A^{reco} - A^{smear}\right)/A^{reco}$ values at different cuts on m_T varying with step 1 GeV from 20 GeV to 60 GeV.

The templates method

The idea of template method is to produce a series of distributions of some quantity corresponding to various input W mass values. These template distributions are compared to the data distribution to find the best fit and the input W mass associated with this best fit template will be the W mass measured.

We make the template distributions by correcting the MC truth distribution as presented in Section 4.5. To create different template distributions with different W mass values (called template masses) we re-weight event by event the smeared distribution with a *weight* determined by the Breit Wigner (BW) probability density function:

$$weight = \frac{BW(m, M_W^{temp}, \Gamma^{temp})}{BW(m, M_W^{truth}, \Gamma^{truth})}, \tag{4.13}$$

where m is the MC true W mass in a given event; M_W^{truth} and Γ^{truth} is respectively the mean and the width of the W mass spectrum of the MC truth while M_W^{temp} and Γ^{temp} are those of the templates.

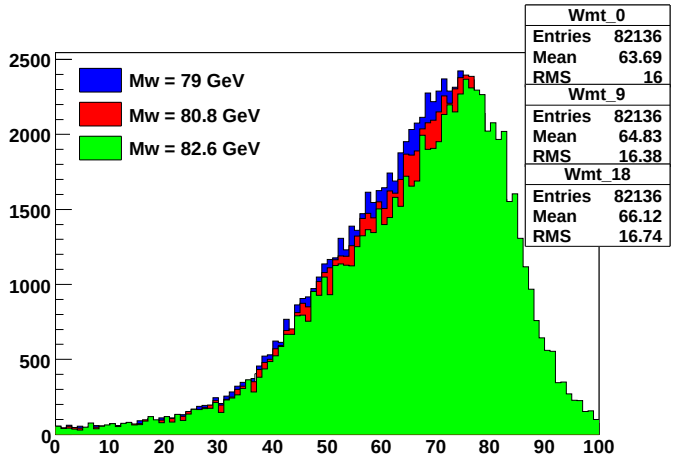


Figure 4.19: template distributions of W transverse mass corresponding to the template masses: 79 GeV (blue), 80.8 GeV (red) and 82.6 GeV (green).

For example, we create 3 templates of m_T distribution corresponding to 3 input W mass values of 79 GeV, 80.8 GeV and 82.6 GeV as plotted in Figure 4.19 by smearing the MC truth ($M_W^{truth} = 80.39$ GeV and $\Gamma^{truth} = 2$ GeV) with the weights given by equation 4.13. Once the templates are generated, they are compared to the data distribution by using the χ^2 method as explained in [71]. We plot a graph of χ^2 values versus the corresponding template masses then fit it as a function of $\chi^2(M_W^{temp})$ having a parabola form with a minimum point. The value of M_W^{temp} at this minimum point is the W mass measured M_W^{data}

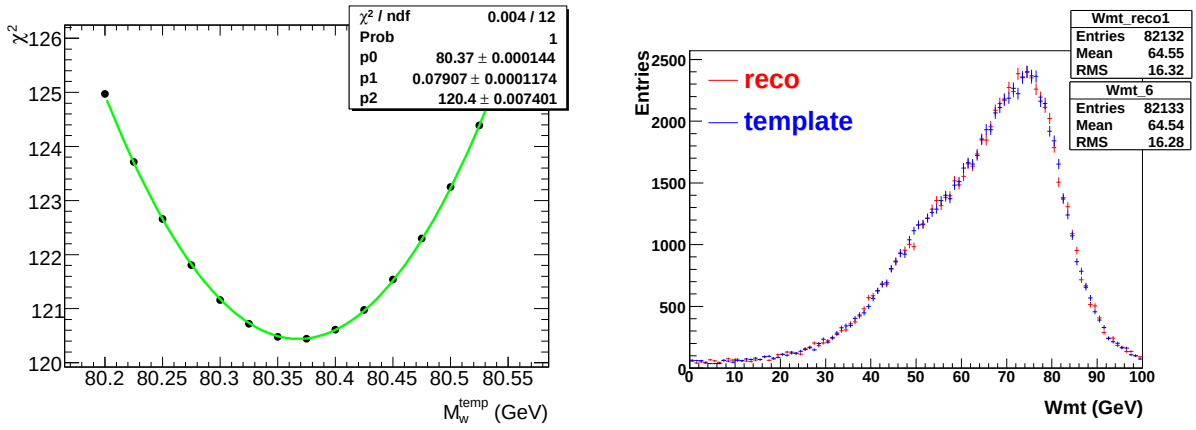
and the template corresponding to this W mass has the best fit to the data with a χ^2 value equal to the χ^2_{min}.

W mass measurement with m_T templates

Starting from the MC truth ($M_W^{truth} = 80.39$ GeV; $\Gamma^{truth} = 2$ GeV), templates are generated with a series of input W mass values with 25 MeV steps. Then, the χ^2 versus input mass curve is fitted using the following parabola function:

$$y = \frac{(x - p_0)^2}{p_1^2} + p_2. \tag{4.14}$$

With this fit function, p_0, p_1 and p_2 are alternately the W mass measured M_W^{data}, the uncertainty and the χ^2_{min}. We measure a W mass of 80.37 ± 0.079 (GeV) with a statistical uncertainty of 79 MeV. Obviously, this mass value is unbiased within the statistical uncertainty. This corresponds to a minimum χ^2 value of $\chi^2_{min}/dof = 120.4/100$ (see Figure 4.20(a)). Figure 4.20(b) shows the reconstructed m_T distribution as well as the template which is closest to the best fit.



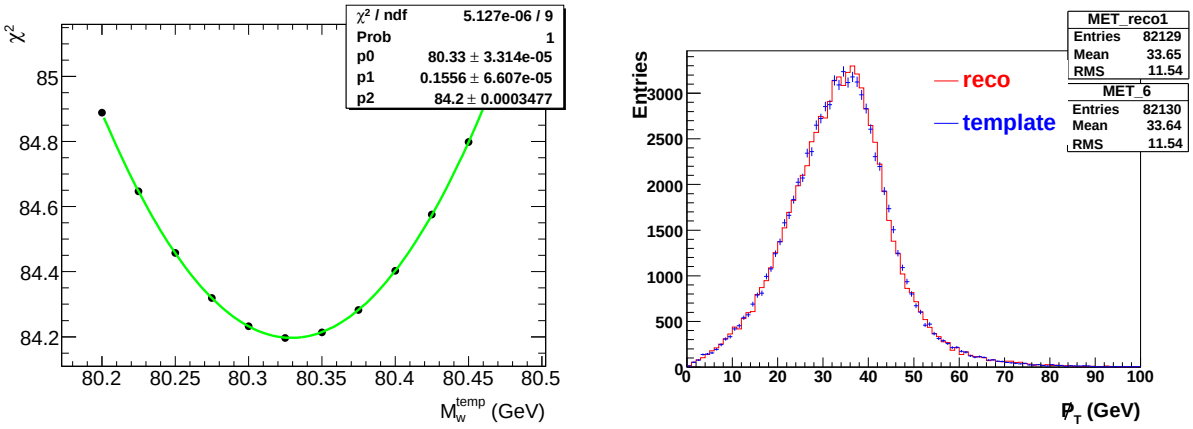
(a) (b)

Figure 4.20: 4.20(a) Parabola fit of the χ^2 values depending on various input W mass values in m_T templates; 4.20(b) Reconstructed m_T distribution and the template closest to the best fit.

W mass measurement with $\not{E}_T$ templates

Similarly to the m_T templates, we can make the E_{T}^{miss} templates to measure the W mass. The results are shown in Figure 4.21 where the W mass is measured to be 80.33 ± 0.156 (GeV) with the statistical

uncertainty of 156 MeV and $\chi^2_{min}/dof = 84.2/70$. These results are compatible (within the statistical uncertainty) with those obtained in the case of m_T templates.



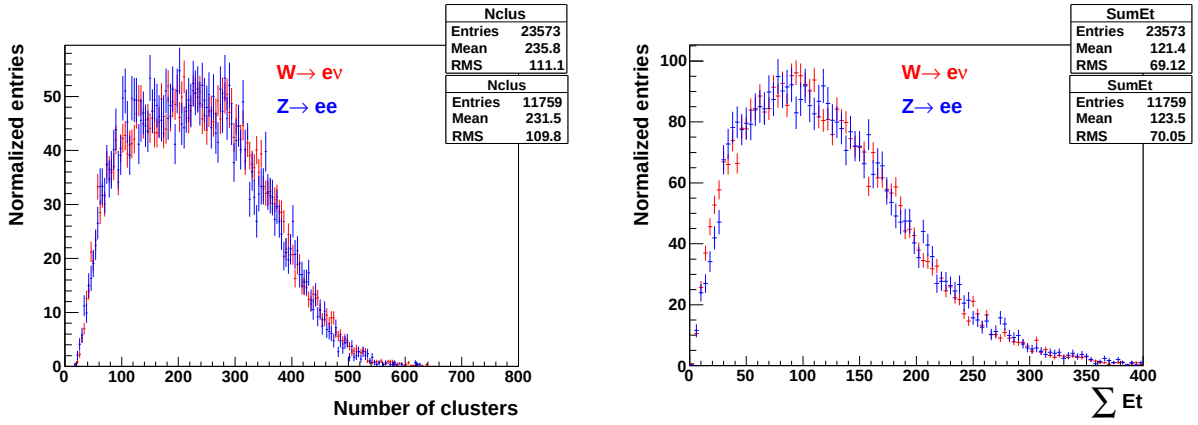
(a) (b)

Figure 4.21: 4.21(a) Parabola fit of the χ^2 values depending on various input W masses in E_{T}^{miss} templates; 4.21(b) reconstructed $\not{E}_{\mathrm{T}}$ distribution and the template closest to the best fit.

4.7 Calibration of hadronic recoil response in Z data

As mentioned in Section 4.5, *hadronic recoil* $\overrightarrow{p_{\mathrm{T}}}^{R}$ and the *hadronic activity* ΣE_{T} are good variables to parameterize the detector response to the hadronic recoil. Therefore, if the performance of these quantities are similar in W events and Z events, we can transfer the behaviors of the recoil response calibrated in Z events to W events. This way will give realistic results without relying on MC truth since in the Z events, two well-measured leptons are used to calibrate the true recoil [4].

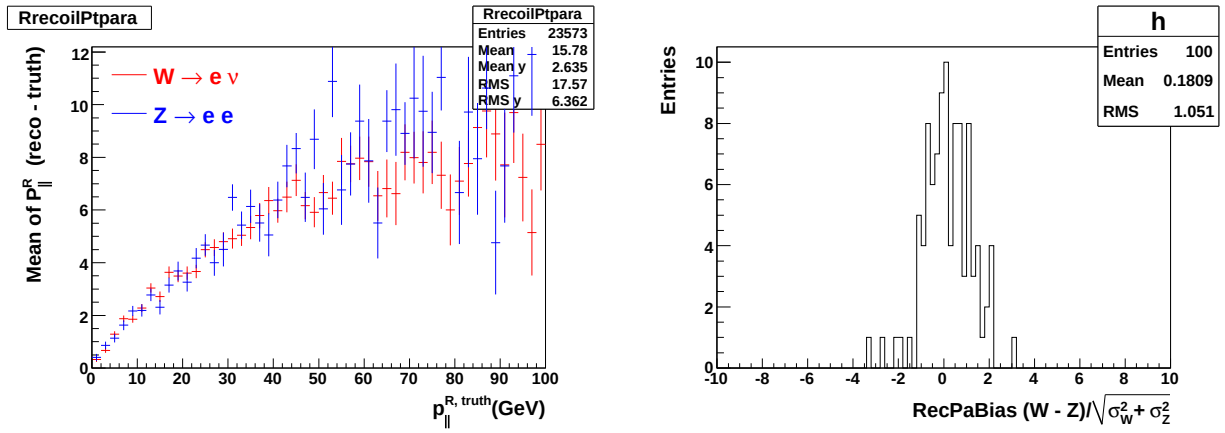
Similar to the W events, the hadronic recoil in Z events is reconstructed from topo-clusters excluding topo-clusters associated to electrons. The similar distributions of the number of clusters (fig. 4.22(a)) and those of *hadronic activity* (fig. 4.22(b)) in W and Z events imply that the energy scale of clusters is compatible in these samples. Since the number of clusters and the energy scale of clusters are compatible, the hadronic recoil performances in both samples are compatible. Therefore, as shown in Figure 4.23, the scale bias is compatible for the parallel recoil in W and Z events. In addition, since the hadronic activity ΣE_{T} is similar in W and Z, the resolution behavior of the recoil is thus similar as seen in Figure 4.24. The "pull" distributions with mean ~ 0 and RMS ~ 1 in Figures 4.23(b) and 4.24(b) show quantitatively that the recoil bias and the recoil resolution in W and Z events are statistically compatible.



(a) Number of clusters in $W \to e\nu$ and in $Z \to ee$ events

(b) Hadronic activity ΣE_{T} in $W \to e\nu$ and in $Z \to ee$ events

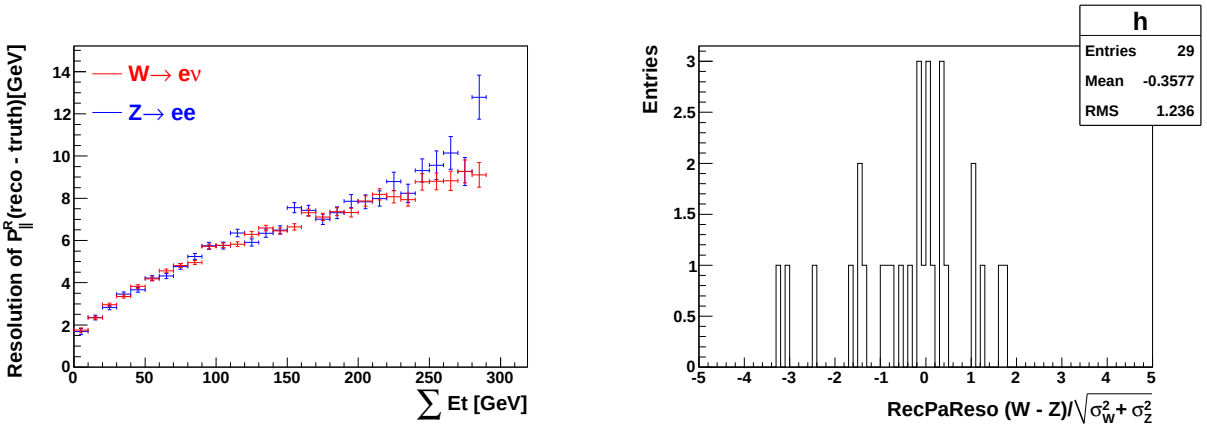
Figure 4.22: Comparing number of clusters and ΣE_{T} distributions in $W \to e\nu$ and $Z \to ee$ events.



(a) Scale bias of recoil along parallel axis in $W \to e\nu$ and $Z \to ee$ events

(b) Distribution of the normalized difference between parallel recoil bias in $W \to e\nu$ events and in $Z \to ee$ events at given bins of $p_{\parallel}^{R,truth}$.

Figure 4.23: Scale bias of the recoil along parallel axis in $W \to e\nu$ and $Z \to ee$ events.



(a) Resolution of the recoil along parallel axis in $W \to e\nu$ and in $Z \to ee$ events

(b) Distribution of the normalized difference between parallel recoil resolution in $W \to e\nu$ events and in $Z \to ee$ events at given bins of ΣE_{T}.

Figure 4.24: Resolution of recoil resolved along axis in $W \to e\nu$ and in $Z \to ee$ events.

The method of correcting recoil in W events is validated as presented in the previous section. Once the scale bias and the resolution behaviors of the recoil in W and Z events are similar, the recoil can be corrected successfully by parameterizing the recoil response in Z data.

4.8 Summary

We presented an alternative algorithm for E_{T}^{miss} computation in W events starting from the hadronic recoil. This algorithm has a number of advantages over the standard one such as improvement in bias and resolution and transparent between channels. This method is especially suited for events with leptons. Additionally, in this method, the free parameter which is the cone size used to exclude lepton clusters, allows studying the stability of results and specific correction to be applied to the energy replacement and Bremsstrahlung effect.

The hadronic recoil calculation presented here is for electron channel but it can be applied similarly for the muon channel since this calculation, by the construction, is independent on the leptons in the final state. To be reconstructed as a fundamental object, this hadronic recoil is a good input for electroweak boson measurements while the standard algorithm has several deficiencies. In the standard algorithm, firstly, the default inputs in terms of reconstruction objects (electrons, muons, jets, ...) do not necessarily

match the present analysis selections. Secondly, the hadronic recoil is indeed determined from $E_{\mathrm{T}}^{\mathrm{miss}}$ variables. There is no dedicated calculation of the hadronic recoil, only $E_{\mathrm{T}}^{\mathrm{miss}}$ is provided. This implies that no particular attention is paid to the interpretation of lepton-induced calorimetric energy deposits. These are correctly included in the overall energy imbalance, but no decision is made concerning their origin, *i.e.* whether they are related to p_T^W (the hadronic recoil) or to the W decay particles.

Determining the sensitive parameters governing the recoil response, we could simulate the response with very small bias: the acceptance bias is less than 0.03% in W cross-section measurement and the W mass bias is much smaller than the statistic sensitivity of our exercise (about 80 MeV).

Finally, we showed that Z events could be used to calibrate the recoil response. Further work in this direction is developed and first applied for the p_T^W spectrum measurement with 2010 data at LHC which will be presented in Chapter 5.

Chapter 5

Measurement of p_T^W spectrum with ATLAS data produced by proton-proton collisions at $\sqrt{s} = 7$ TeV

5.1 Introduction

This chapter describes a measurement of the transverse momentum spectrum of W boson with electron channel decay using the 2010 ATLAS data set produced at the LHC by proton-proton collisions at $\sqrt{s} = 7$ TeV . This data corresponds to 31 pb^{-1} of integrated luminosity. The work in this chapter is also reported partially in [29].

In this analysis, the background-subtracted reconstructed p_T^W distribution is unfolded to the propagator, or "Born-level", p_{T} distribution. The detector response for the unfolding is obtained from Z data by the data-driven method (as discussed in Section 4.7 of Chapter 4) which is performed in electron and muon channels combined to get more statistic. The unfolded p_T^W distribution is then corrected for the detector efficiency to get the final result.

The analysis proceeds along the following steps : event selection; background subtraction; unfolding of the transverse momentum spectrum; correction to the fiducial cross section, differential in the unfolded transverse momentum spectrum; and normalization of the result. These steps are mathematically

formulated below. The background-subtracted spectrum is:

$$S_{reco} = N - B, \tag{5.1}$$

$$\delta S_{reco} = \delta N \oplus \delta B, \tag{5.2}$$

where N and B are the observed event count and the background estimation, δN is the statistical uncertainty, and δB the uncertainty on the background estimation. All quantities at this stage are binned in the reconstruction-level W boson transverse momentum. The background-subtracted transverse momentum spectrum is unfolded under the transformation given at (5.3) by introducing the response matrix R,

$$S_{unf} = R^{-1}\, S_{reco}, \tag{5.3}$$

$$\delta S_{unf} = R^{-1}(\delta N) \oplus R^{-1}(\delta B) \oplus (\delta R^{-1}) S_{reco}. \tag{5.4}$$

The response matrix is determined from the simulation, and incorporates parametrized data-driven corrections determined from the Z control sample (see Section 5.5). Its uncertainty, δR, accounts for finite Monte Carlo statistics, for bias induced by the parametric model and the choice of the input physics model, and for the precision of the data driven corrections. Also here, R^{-1} is not strictly the inverse of the response matrix R. Rather, $R^{-1}\, X$ represents the unfolding of X using the response matrix. After the application of unfolding, all quantities are binned in the truth-level W transverse momentum. Both reconstruction-level and truth-level binning of p_T^W for this analysis are given in Section 5.5.1.

Statistical uncertainties are propagated through the unfolding procedure by mean of Poisson fluctuations applied to the input distributions. In particular, $R^{-1}(\delta N)$ is obtained by generating pseudo-experiments, varying the bin contents around the actual observation within the expected uncertainties, repeating the unfolding step and recording the spread around the central result. The statistical component of $(\delta R^{-1}) S_{reco}$ is obtained similarly, but randomizing the bin contents of R.

The unfolded spectrum is corrected to a fiducial cross section each bin as follows:

$$\sigma(p_T^W) = \frac{S_{unf}(p_T^W)}{C_W(p_T^W)\, L}, \tag{5.5}$$

$$\frac{\delta\sigma}{\sigma}(p_T^W) = \frac{\delta S_{unf}}{S_{unf}}(p_T^W) \oplus \frac{\delta C_W}{C_W}(p_T^W) \oplus \frac{\delta L}{L} \tag{5.6}$$

$\sigma(p_T^W)$ is the cross section averaged over the corresponding p_T^W bin, and $C_W(p_T^W)$ is a factor correcting for event loss and lepton migration effects. This factor is computed from the simulation, as the number of events passing all reconstruction-level cuts, divided by the number of events passing the kinematic cuts at

the generation level. Since it is to be applied after having unfolded the transverse momentum spectrum, it is computed bin by bin in the generator-level p_T^W, and includes data driven smearing corrections for the lepton resolution, and scaling factors for the lepton efficiency:

$$C_W(p_T^W) = \frac{\sum_i \Theta_i \, \Phi(p_T^W) \, SF_i}{\sum_i \Theta_i' \, \Phi(p_T^W)}, \tag{5.7}$$

where the sum is over all events in the simulation sample. Θ is 1 for an event passing the reconstruction-level cuts after applying the relevant smearing corrections, and 0 otherwise; SF is an event weight calculated from the trigger and reconstruction scale factors, estimated for each event as a function of the lepton reconstruction-level kinematics. Θ_i' is 1 for events passing the kinematic cuts at the generator level, and 0 otherwise. $\Phi(p_T^W)$ is 1 if the event falls in the p_T^W bin under consideration, and 0 otherwise. The uncertainty on C_W accounts for generator bias as above, finite MC statistics, and the smearing correction and scale factor uncertainties.

Finally, the resulting spectrum is normalized to unit integral. This step allows to significantly reduce the uncertainties that are strongly correlated across the spectrum.

Following above strategy of the analysis, this chapter is organized as follows. First, the data and Monte Carlo samples are described in Section 5.2, followed by the criteria used to select $W \to e\nu$ candidate events, in Section 5.3. Section 5.4 gives the background calculations. The construction of the response matrix relating the true p_T^W, $p_T^{W,\text{true}}$, to the reconstructed p_T^W, $p_T^{W,\text{reco}}$, from a combination of Z and W Monte Carlo and data is described in Section 5.5. The procedure which uses the response matrix to unfold the $p_T^{W,\text{reco}}$ distribution back to a measurement of the $p_T^{W,\text{true}}$ distribution is described in Section 5.6. The correction of the measured $p_T^{W,\text{true}}$ distribution for the $W \to e\nu$ reconstruction efficiency is given in Section 5.7. The treatment of statistical and systematic uncertainties is detailed in Section 5.8, including the propagation of uncertainties through the unfolding procedure and uncertainties arising from the unfolding procedure itself. Finally, Section 5.9 shows the results of the analysis.

5.2 Data and Simulation Samples

This analysis is based on approximately 31.43 pb^{-1} of data collected from September 25th to October 29th, 2010 (periods G2 to I2). There are about 105k $W \to e\nu$ candidates selected from these data periods. The choice of dataset was determined as a compromise between maximizing the available luminosity while maintaining stable trigger and reconstruction conditions throughout the sample. In addition, this analysis used simulated samples (given in Section 5.2.3) to model background contaminations, correct measured values for efficiency and acceptance effects, and determine detector response. In order to better match data, the simulated samples are corrected by event re-weighting based on the number of reconstructed vertices and applying electron momentum smearing and scaling. This section describes the requirements placed upon the collected data, the simulated samples that were used, and the corrections that were applied to them.

5.2.1 Data Quality Requirements and Integrated Luminosity

All selected events must pass certain requirements for the proper functioning of the detector, as described by a Good Runs List (GRL). In particular, the common W/Z GRL was used, which requires stable beam conditions at $\sqrt{s} = 7$ TeV as well as good working conditions for all parts of the detector and trigger used including solenoid, toroid, calorimeters, inner detector, L1 trigger hardware, and luminosity monitors.

For the $W \to e\nu$ analysis, the data used corresponds to an integrated luminosity of about 31.43 pb^{-1} with an uncertainty of 3.4% [3] as measured using van der Meer scans. Data periods A-G1 are excluded from this analysis, at a cost of 3.7 pb^{-1} . This choice is made in order to have a consistent data sample between the electron and muon channels in which the muon trigger conditions are stable. This ensures that the running conditions and calorimeter response are completely identical, enabling the combination of the response matrices for the two channels as illustrated in Section 5.5.

The data was processed using Atlas Production release 15.6.13.2 and analyzed from D3PDs produced by Standard Model W/Z group production.

5.2.2 Electron Trigger Requirements

Candidate $W \to e\nu$ events are selected from data passing the `EF_e15_medium` trigger. This trigger is used consistently for periods G2-I2. In the Monte Carlo, this trigger is not available in the configuration implemented in the HLT, so the trigger `L1_EM14` is used instead, and the data / Monte Carlo scale factors

used to correct the trigger efficiency are defined with this trigger for the denominator.

5.2.3 Simulation Samples

Simulation samples used in this analysis are listed in Table 5.1 including their production cross section multiplied by the relevant branching ratios and their total number of events. Wherever possible, we use simulated data samples that included a simulation-based model of additional interactions known as "pileup". The number of reconstructed primary vertices is used as a proxy for the number of interactions. Events from samples with pileup were re-weighted so that the distribution of the number of primary vertices would match the one observed in data.

Signal Monte Carlo, $W \to e\nu$ sample, is used to derive the data driven response matrix, to calculate efficiency corrections, and to estimate systematic uncertainties. $Z \to ee$ Monte Carlo sample is used in the production of the data driven response matrix and, along with $W \to \tau\nu$, $Z \to \tau\tau$, and $t\bar{t}$ Monte Carlo samples, to estimate the electroweak background. QCD multi-jet Monte Carlo modeled by "JF17" sample, is used as part of the QCD background estimate to extrapolate isolation efficiencies of QCD events from the control to the signal region. The MADGRAPH samples used include the photon phase space $p_T^{\gamma} > 10$ GeV and $\Delta R(\gamma, \ell) > 0.5$, and the cross section quoted in Table 5.1 correspond to that phase space.

Monte Carlo corrections

In order to model better detector conditions, it is necessary to correct the Monte Carlo to match the observed data. Firstly, since the data was collected over periods of increasing instantaneous luminosity, the pileup observed is different over periods and is not the same as modeled by the Monte Carlo as shown in Figure 5.1(a). The Monte Carlo is thus needed to reweight according to the number of primary vertices to correctly match the level of activity observed in the data. For the analysis with electron decay channel, the number of reconstructed primary vertices must be measured having at least 3 tracks to reduce beams Halo. In addition, the number of primary vertices is counted in events which pass the GRL requirement, pass the trigger requirement, and have at least an electron with $E_T > 18$ GeV passing the Robust Medium [27] selection. The distributions of number of reconstructed vertices as well as primary vertex weight obtained for $W \to e\nu$ sample in individual data periods are shown in Figure 5.1.

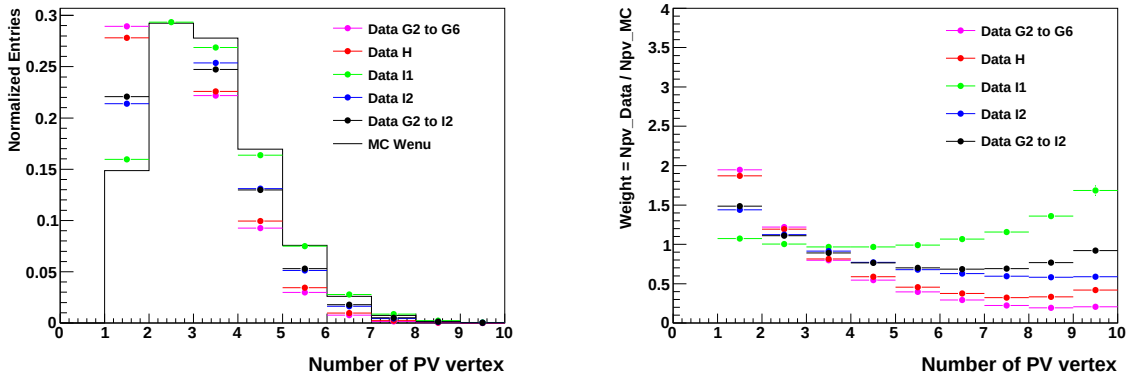
Because an identical procedure was used to model pileup in all of the simulated data samples, the

Process	Generator	Cross-section (nb)	N_{evt}(x10^6)
$W \to e\nu$	PYTHIA	10.46	6.5
$Z \to ee$	PYTHIA	0.989	1.0
$W \to \tau\nu$	PYTHIA	10.46	1.0
$Z \to \tau\tau$	PYTHIA	0.989	0.10
$(W^+ \to e^+\nu)\gamma$ (without pileup)	MADGRAPH+PYTHIA	0.028	0.05
$(W^- \to e^-\bar{\nu})\gamma$ (without pileup)	MADGRAPH+PYTHIA	0.019	0.05
$(W^+ \to \tau^+\nu)\gamma$ (without pileup)	MADGRAPH+PYTHIA	0.025	0.05
$(W^- \to \tau^-\bar{\nu})\gamma$ (without pileup)	MADGRAPH+PYTHIA	0.017	0.05
$t\bar{t}$ (without pileup)	MC@NLO	0.165	1.0
$t \to (W \to e\nu)b$ (t-channel)	MC@NLO	0.0071	0.01
$t \to (W \to \tau\nu)b$ (t-channel)	MC@NLO	0.0071	0.01
"JF17"	PYTHIA	1.15 $\times 10^6$	10.

Table 5.1: Simulated data samples used in the analysis. All samples include pileup except where noted.

primary vertex weight applied for $W \to e\nu$ Monte Carlo as listed in Table 5.2 is used for all simulated pile-up samples. The weight for an event with a given number of vertices is defined as the ratio of the fraction of events with that number of vertices in data to the fraction with that number in simulation.

Additionally, Monte Carlo samples are corrected with trigger and reconstruction scale factor and the reconstructed lepton p_T are scaled and smeared according to the results of resolution studies using the Z mass constraint, as described in Section 5.7.



(a) Number of primary vertices distribution

(b) Primary vertex weight

Figure 5.1: 5.1(a): Distributions of number of primary vertices in $W \to e\nu$ Monte Carlo sample and in different periods of data; 5.2(b): Weights according to number of primary vertices used to correct for the pileup simulation in $W \to e\nu$ Monte Carlo.

Npv	G2 to G6	H	I1	I2	G2 to I2
1	1.9463 ± 0.0047	1.8704 ± 0.0043	1.0730 ± 0.0025	1.4395 ± 0.0040	1.4846 ± 0.0026
2	1.2204 ± 0.0025	1.1937 ± 0.0023	1.0038 ± 0.0017	1.1218 ± 0.0024	1.1089 ± 0.0014
3	0.7979 ± 0.0019	0.8132 ± 0.0019	0.9664 ± 0.0017	0.9126 ± 0.0022	0.8901 ± 0.0012
4	0.5457 ± 0.0020	0.5870 ± 0.0020	0.9658 ± 0.0021	0.7724 ± 0.0025	0.7658 ± 0.0014
5	0.3957 ± 0.0025	0.4550 ± 0.0025	0.9897 ± 0.0033	0.6775 ± 0.0035	0.7003 ± 0.0020
6	0.2914 ± 0.0036	0.3764 ± 0.0039	1.0673 ± 0.0059	0.6286 ± 0.0057	0.6855 ± 0.0033
7	0.2216 ± 0.0058	0.3229 ± 0.0066	1.1558 ± 0.0116	0.5940 ± 0.0103	0.6900 ± 0.0062
8	0.1933 ± 0.0110	0.3319 ± 0.0137	1.3595 ± 0.0265	0.5801 ± 0.0208	0.7669 ± 0.0138
9	0.2065 ± 0.0254	0.4186 ± 0.0347	1.6832 ± 0.0692	0.5895 ± 0.0468	0.9222 ± 0.0355
10	0.2805 ± 0.0708	0.3949 ± 0.0795	2.0110 ± 0.1877	0.8782 ± 0.1395	1.1154 ± 0.0977

Table 5.2: Primary vertex weights and their statistical uncertainties applied for $W \to e\nu$ Monte Carlo sample in individual data periods and in the full data period.

5.3 Event Selection

$W \to e\nu$ candidates are selected with one high quality electron having E_T (p_T) > 20 GeV, $E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV, and $m_T^W > 40$ GeV. The details of the quality requirements are given in the following subsections.

5.3.1 Reconstruction level Event Selection

The selection of $W \to e\nu$ candidates can be divided into three categories: (a) collision-like event selection, which requires the reconstruction of a primary vertex (b) high E_T electron selection, and (c) $W \to e\nu$ selection, which places requirements on missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ and W transverse mass m_T. These cuts are summarized in Table 5.3.

Collision-like Event	
Trigger	`EF_e15_medium`
Jet Cleaning	Cleaning Cuts (see [50])
OTX Cleaning	Object Quality Checks (see [49])
Primary Vertex	At least one PV with $N_{\mathrm{trk}} \geq 3$
High E_T electron	
Electron Kinematics	$E_T > 20$ GeV
	$\lvert\eta_e\rvert < 1.37$ or $1.52 < \lvert\eta_e\rvert < 2.47$
Electron Quality	Robust Tight [27]
$W \to e\nu$	
Missing E_T	$E_{\mathrm{T}}^{\mathrm{miss}} > 25$ GeV
Transverse Mass	$M_T > 40$ GeV

Table 5.3: Selection of W candidates for electron channel analysis

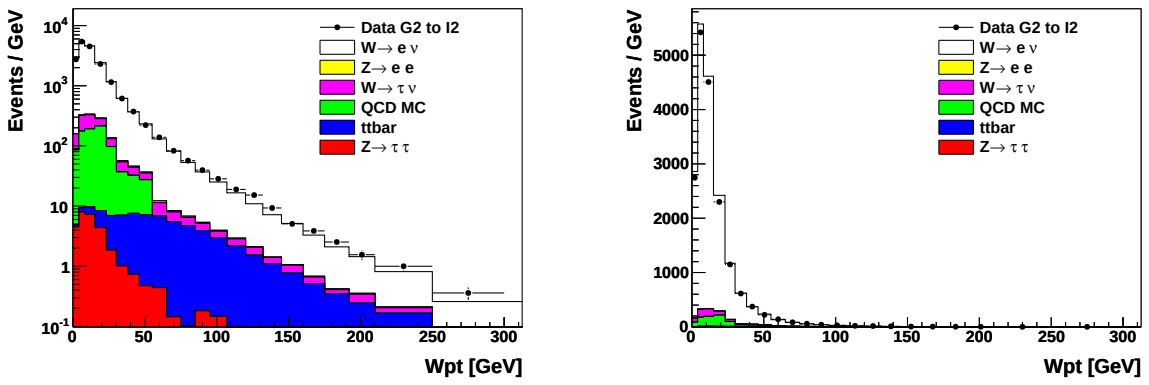
The "collision-like" event selection begins with the trigger requirement detailed in Section 5.2.2. In order to prevent the inclusion of events containing large noise spikes in the calorimeter, which could bias the $E_{\mathrm{T}}^{\mathrm{miss}}$, cuts are applied on the quality, timing, and energy distribution of jets in the event, as described in [50]. Also, the event is required to have at least a primary vertex reconstructed from at least three tracks.

The electron selection has been discussed and motivated in detail in Refs. [27] and [2]. For this

analysis, we select events containing at least one electron passing the "Robust Tight" requirements [27] with $E_T > 20$ GeV and $|\eta^e| < 2.47$, excluding the transition region, $1.37 < |\eta^e| < 1.52$, between the barrel and end-cap calorimeters.

The $W \to e\nu$ candidate selection is completed by the $E_{\mathrm{T}}^{\mathrm{miss}}$ and m_T thresholds, which require that the event is kinematically consistent with the decay of a W boson into an electron and a neutrino. The $E_{\mathrm{T}}^{\mathrm{miss}}$ definition used is `MET_LocHadTopo`. This $E_{\mathrm{T}}^{\mathrm{miss}}$ definition is chosen in order to maintain a consistent candidate sample with the W asymmetry analysis, and because it is better-suited for use in the QCD multi-jet background estimates (see Section 5.4). In periods G2-I2, a total of 104,904 $W \to e\nu$ candidates satisfy all of the event selection described here.

The Table 5.4 lists the cut flow for data as well as for signal and backgrounds expectations. The number of events of simulation samples listed are normalized to the integrated luminosity of 31.43 pb^{-1} by using theoretical cross section predictions given in Table 5.1. After all steps of event selection, the reconstructed p_T^W distribution observed in data in a comparison with that predicted by the Monte Carlo simulation with the same integrated luminosity are illustrated in Figure 5.2.



(a) Log scale (b) Linear scale

Figure 5.2: Observed and predicted p_T^W distributions after all selection cuts. Total number of events in each Monte Carlo sample is normalized to the data luminosity of 31.43 pb^{-1}.

5.3.2 Control plots after the event selection

This section presents event distributions observed in comparisons with predictions after the transverse mass cut. At this stage, the QCD background is obtained from the simulation, which explains the imperfect agreement. Figure 5.3 illustrates the lepton transverse momentum and pseudo-rapidity, missing

Cutflow	Data	$W \to e\nu$	$W \to \tau\nu$	$Z \to ee$	$Z \to \tau\tau$	ttbar	st	JF17
Presel	1743816	176438.8	19565.0	23873.4	4688.5	1074.1	214.0	22811639.5
Medium	1735775	161699.9	11194.3	22839.3	2624.8	886.7	170.6	4181674.9
Zveto	1726316	158170.7	10934.3	12532.1	2407.3	760.9	159.8	3997946.7
Rtight	395515	126385.9	7652.7	9972.8	1608.8	607.7	127.6	732222.3
MET	111780	100011.6	3973.0	286.5	475.5	521.0	105.7	23702.7
m_T	104904	98291.6	3082.7	243.3	187.0	411.6	88.4	6149.9
PV reweighted:	104904	98464.8	3073.3	226.8	181.3	411.6	88.5	5162.9

Table 5.4: $W \to e\nu$ cutflow for data periods G2 to I2 and Monte Carlo normalized to $\int L = 31.43$ pb^{-1}. The "Presel" cut is a pre-selection cut done by requiring events satisfying the collision-like event selection and having at least one Robust Loose electron [27] with $E_{\mathrm{T}} > 20$ GeV and $|\eta^e| < 1.37$ or $1.52 < |\eta^e| < 2.47$. The "Medium" cut requires events having at least one Robust Medium electron. The "Zveto" cut vetoes events having more than one Robust Medium electrons. "Rtight" cut requires events with a Robust Tight electron. "MET" and "m_T" are the cuts on missing transverse energy and W transverse mass. "PV reweighted" is the step applying the event weight according to the number of primary vertices for Monte Carlo samples.

transverse energy, and transverse mass distributions in $W \to e\nu$ events.

We will return to the W boson transverse momentum distribution, including data driven background estimates where relevant, at the end of Section 5.4.

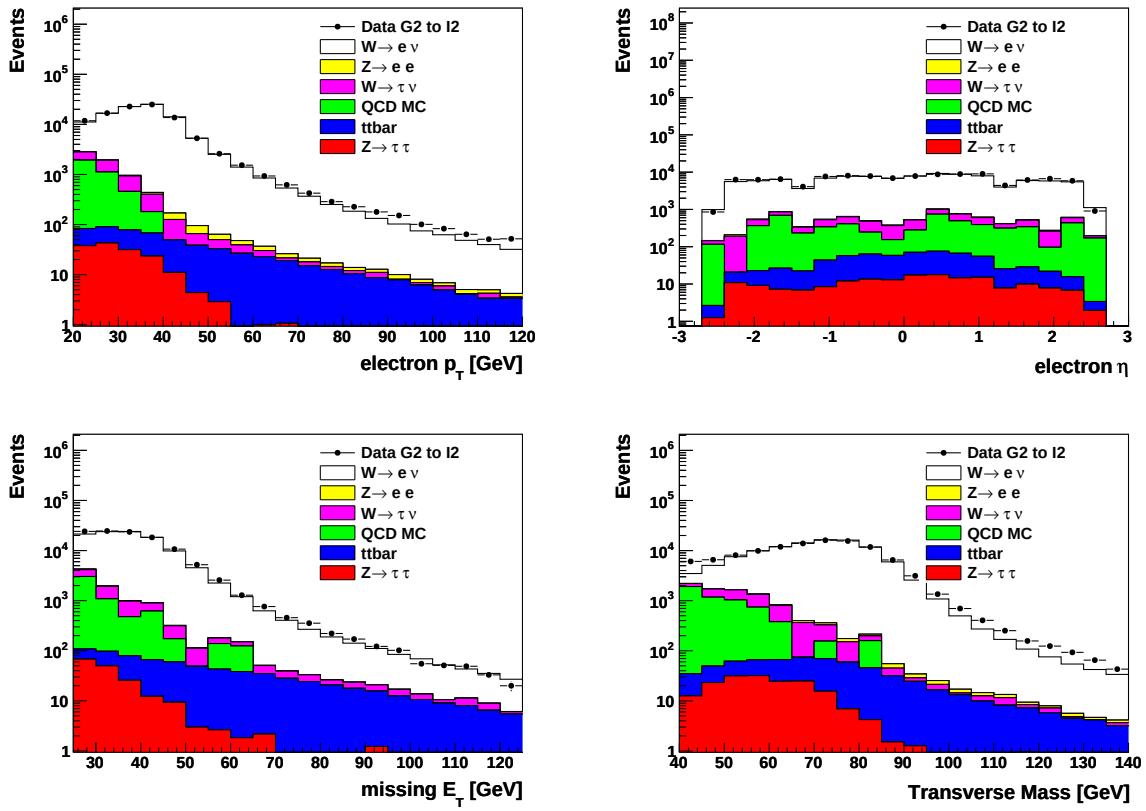


Figure 5.3: Lepton transverse momentum and pseudorapidity, missing transverse energy, and transverse mass distributions after the event selection. The jet background is as predicted by the Monte Carlo.

5.3.3 Generator-level Event Selection

The efficiency correction described in Section 5.7 is defined in terms of W events generated in the fiducial volume of the detector. This is the subset of W events passing the following requirements at generator level which are the same as for muon channel (in order to combine both channels for the final results which will be presented in Section 6.1):

1. $p_T^{\ell} > 20$ GeV
2. $|\eta^e| < 2.4$

3. $p_T^{\nu} > 25$ GeV

4. $m_T^W > 40$ GeV

These requirements approximate the reconstruction-level selection defined in the previous sections. Though the generators do include initial and final state QED effects, the generator-level kinematics above are taken at the Born level. This effectively amounts to subtracting the QED corrections. Correspondingly, the W boson momentum is equal the sum of the Born level lepton kinematics, and $p_T(W)$ refers to the boson propagator momentum. Variations around this convention are discussed in Section 5.10.

5.4 Background Calculations

Background subtraction is performed by estimating the contamination of the W candidate sample in bins of p_T^W, as reconstructed using the hadronic recoil algorithm described in Chapter 4. Electroweak, $t\bar{t}$ and single top background shapes and acceptances are determined using Monte Carlo simulation. The total number of events in each Monte Carlo sample is normalized to the data luminosity of 31.43 pb^{-1} by using theoretical cross section predictions as listed in Table 5.1. Both the shape and normalization of the QCD background are estimated from data. In the procedure of QCD backgrounds estimation, the $E_{\mathrm{T}}^{\mathrm{miss}}$ from standard definition [25], `MET_LocHadTopo`, is used as a control variable to separate signal from background. As with all of the background estimates, this procedure is performed in bins of p_T^W calculated from hadronic recoil algorithm to get the QCD background shape.

5.4.1 Backgrounds from W, Z, $t\bar{t}$ and single top production

In $W \to e\nu$ selection, non-signal W and Z decays, along with isolated leptons produced by $t\bar{t}$ and single top events, make up roughly 4% of the total selected sample. In particular, the largest backgrounds are from $W \to \tau\nu$ in which the τ decays into the electron, $Z \to ee$ where a single electron is lost, $Z \to \tau\tau$, and $t\bar{t}$ and single top events involving at least one semi-leptonic decay. The $W \to \tau\nu$, $Z \to ee$ and $Z \to \tau\tau$ backgrounds are collectively referred to as the "electroweak" background; $t\bar{t}$ and single top backgrounds are merged and referred to as "top" background in which the single top background represents about 20% of the total top background.

The shape of these backgrounds is estimated from Monte Carlo simulation by applying the appropriate selection described in Section 5.3 and applying the corrections (re-weighting according to primary vertex and smearing the resolution of electron energy) described in Section 5.2.3. The total number of events in each background sample is determined by scaling the sample according to its theoretical cross section as given in Table 5.1 and the measured luminosity. The separate contributions from the top and electroweak background is shown in Figure 5.4, and the resulting total number of electroweak and top events per GeV per bin in p_T^W is shown in Figure 5.5.

Systematic uncertainties on the electroweak and top backgrounds estimate include uncertainties on the scale factors used to correct the electron reconstruction efficiencies, a 3% uncertainty due to the variation of selection efficiency and acceptance on the PDF set used, taken from Ref. [66], as well as an uncertainty introduced by electron resolution smearing, and an uncertainty on the theoretical cross

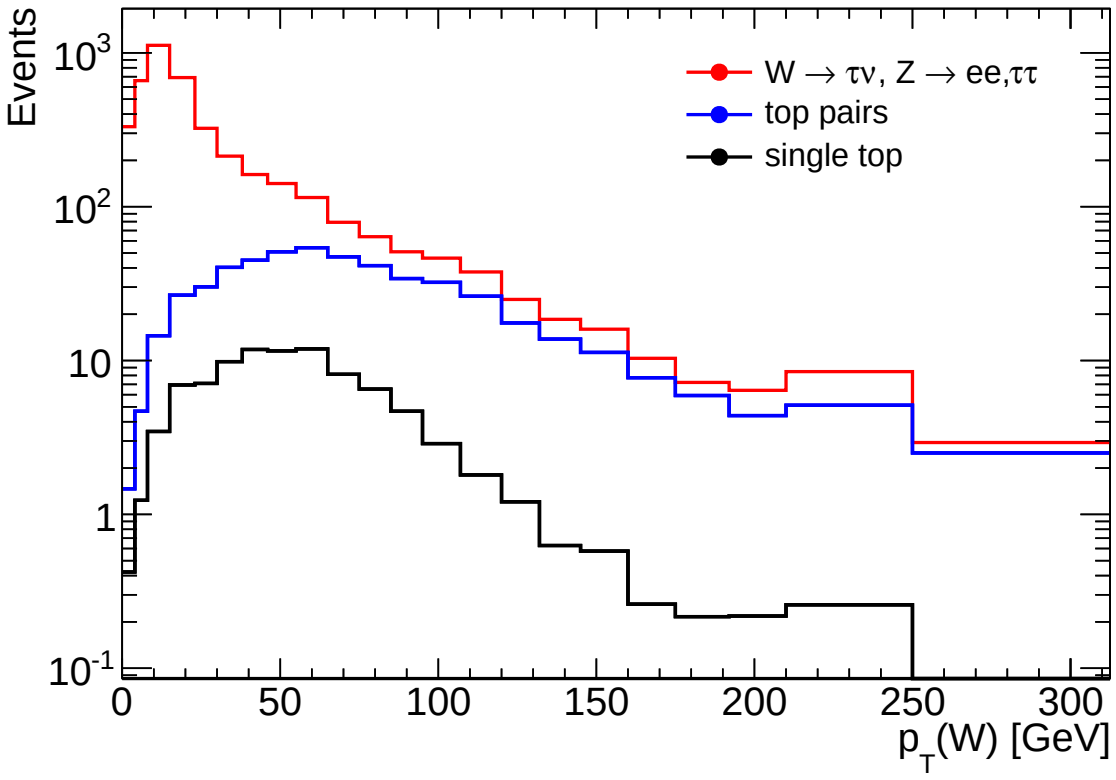


Figure 5.4: Stacked electroweak, $t\bar{t}$ and single top backgrounds in the electron channel.

section used to normalized these background samples. The scale factor uncertainties are computed as described in Section 5.7.1. The smearing error is taken to be 100% of the difference between the background with smearing applied, and without smearing applied. The uncertainty on the cross section is taken to be 5% and completely correlated for all of the W and Z backgrounds. For the top background, the uncertainties on the cross section are taken to be 6% for $t\bar{t}$ and 10% for single top, fully correlated to each other and uncorrelated to the W and Z backgrounds. Finally, the overall estimate is assigned an 3.4% uncertainty due to the uncertainty on the total luminosity. The statistic uncertainties on all electroweak and top backgrounds are treated fully uncorrelated between backgrounds. The resulting uncertainties on the total electroweak and top backgrounds in bins of p_T^W, are shown in Figure 5.5. The total uncertainty is obtained by summing in quadrature all sources of uncertainties mentioned.

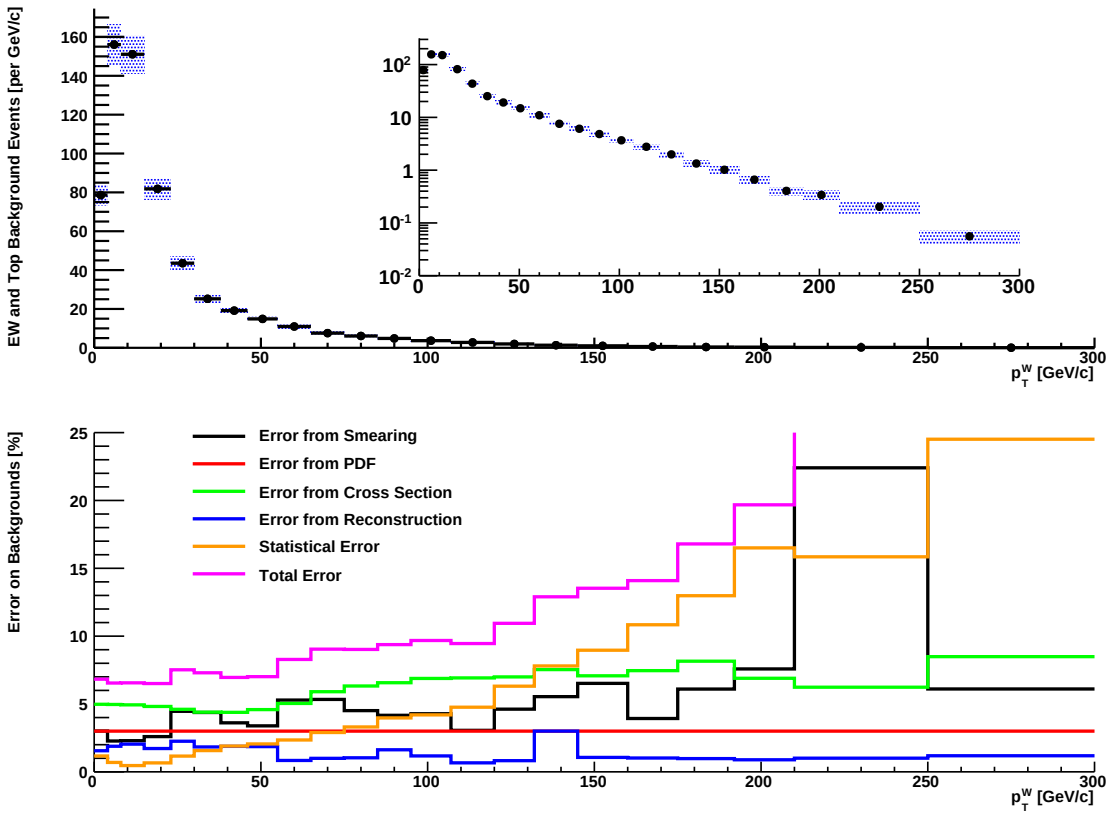


Figure 5.5: Estimated strength of all electroweak and top backgrounds and uncertainties, as a function of p_T^W. Note that an addition 3.4% uncertainty for the luminosity estimate is applied, but not shown here.

5.4.2 QCD background estimation

The sources of QCD background in $W \to e\nu$ selection include electrons from heavy-quark decays, electrons from conversions and hadrons faking electrons where QCD jet objects (i.e. quarks, gluons) are misidentified as high energy electrons [23]. At LHC, the rate of producing QCD jets events is very large. The QCD jets events is thus the dominant background component for $W \to e\nu$ analysis. In addition, the jet production cross-section and fragmentation properties are not well understood. Therefore the QCD jets production and the fake rate is difficult to be modeled precisely by current Monte Carlo and induce a significant uncertainty on the magnitude of this background. For these reasons, it is needed to develop methods to estimate QCD background from the data. These techniques are so-called "data-driven" estimations.

In this analysis, we use the data-driven method of template fit to the $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum where the $E_{\mathrm{T}}^{\mathrm{miss}}$ is

determined from standard definition [25]. This technique uses the prediction for the shape of the signal and background distributions in the form of two histogram templates, one for the QCD background and one for everything else which is so-called "signal" template, including the $W \to e\nu$ signal and all electroweak and $t\bar{t}$ backgrounds. By using the $E_{\mathrm{T}}^{\mathrm{miss}}$ as a discriminating variable, we perform a binned maximum likelihood fit to the data accounting for the finite Poissonian statistics of the templates [12]. This is the same technique as used for the W lepton charge asymmetry measurement [13], however, the fit is performed in bins of p_T^W reconstructed from the hadronic recoil algorithms as described in Chapter 4 to provide the shape of the QCD background for this analysis.

The signal template is obtained from the combination of $W \to e\nu$, $W \to \tau\nu$, $Z \to ee$, $Z \to \tau\tau$ and $t\bar{t}$ Monte Carlo samples described in Section 5.2. These samples are selected by applying the full event selection listed in Table 5.3 except for the $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold. The signal template is used to estimate, based on Monte Carlo simulation, the contributions from $W \to e\nu$ signal, electroweak and $t\bar{t}$ backgrounds to the $W \to e\nu$ data. The relative normalization of the various contributions is set by their respective theoretical cross sections. The statistical uncertainty on these samples is included as an error on the template, but their systematic uncertainties as mentioned in Section 5.4.1 are neglected.

The QCD background template is taken from a data control sample selected by applying the full $W \to e\nu$ candidate selection as listed in Table 5.3, but the $E_{\mathrm{T}}^{\mathrm{miss}}$ threshold is skipped and some of the requirements for "tight" electron identification are reversed. In particular, the electron candidate must pass the "loose" identification requirements and the track quality requirements from the "medium" electron category. However, the electron candidate must fail at least one of the remaining "medium" or "tight" requirements. These remaining requirements refer mainly to the signature of the electromagnetic shower in the innermost compartment of the calorimeter (the "strips") and the TRT high threshold hits. No requirements to reject conversions were applied. To veto the contribution from $W \to e\nu$ signal, events with isolated electrons are rejected by applying a cut on isolation variable which is track-based or calorimeter-based isolation in a cone of 0.3 normalized to the electron transverse energy. The track-based isolation can be measured by summing up transverse momentum of the tracks surrounding the electron cluster while the calorimeter-based isolation can be measured by summing up transverse energy of calorimeter cells in the chosen cone. This modified $W \to e\nu$ selection was tested in the QCD dijet Monte Carlo sample JF17 as given in Table 5.1 to confirm that it does not bias the $E_{\mathrm{T}}^{\mathrm{miss}}$ template.

Figure 5.7 shows the fit to the template in three different bins of p_T^W in the low-, medium-, and high-

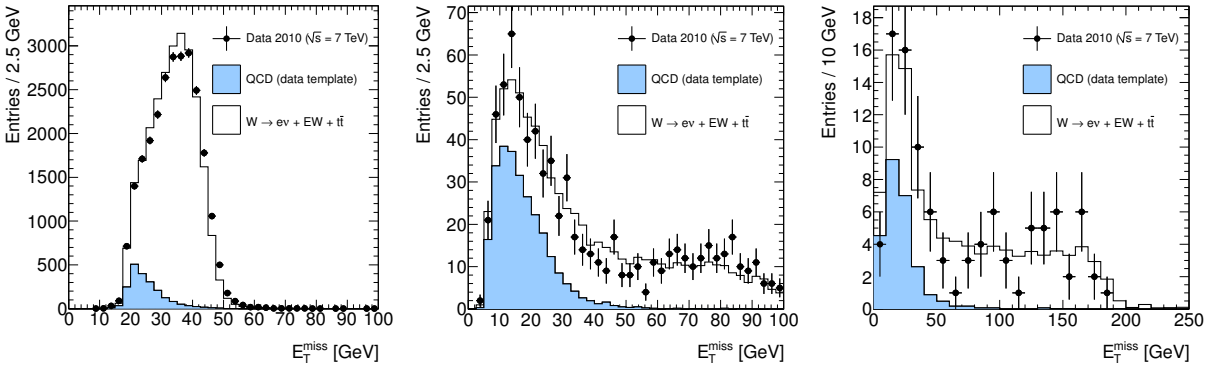


Figure 5.6: Template fits of the QCD background in the electron channel. Three ranges are displayed : $4 < p_T^W < 8$ GeV; $85 < p_T^W < 95$ GeV; $170 < p_T^W < 210$ GeV.

p_T parts of the spectrum. The fit quality remains generally good even as the $E_{\mathrm{T}}^{\mathrm{miss}}$ spectrum shifts and the number of events in the data being fit and in the templates drops off with increasing p_T^W. The shift in the signal templates is the result of the neutrino momentum increasing with increasing p_T^W.

The systematic uncertainty on the background is determined by varying the requirements used to generate the QCD background $E_{\mathrm{T}}^{\mathrm{miss}}$ template. The template selection itself was varied by applying the following requirements one at a time, instead of simultaneously: track isolation; calorimeter isolation; failure of at least one "medium" or "tight" cut; finally, a track veto was applied. The experimental uncertainty was assessed by varying the electron candidate energy within its uncertainty of 3%. The physics model dependence of the fit was estimated by extracting the W signal templates from `MC@NLO` instead of our baseline generator `PYTHIA`.

Table 5.5 gives the results of the QCD background estimates, along with the statistical and systematic uncertainties for each bin. The results are also shown as Figure 5.7.

5.4.3 Treatment of the Wγ process

In the present measurement, the inclusive W process is considered as the signal to be measured. In particular, events containing hard, isolated photons in the final state are not addressed explicitly. This choice is justified below.

We first assess the quality of the `PYTHIA + PHOTOS` model of this process. We consider the $W\gamma$ phase space region defined by $p_T(\gamma) > 10$ GeV, and $\Delta R(l,\gamma) > 0.5$, and compare the `PYTHIA` predictions with an explicit matrix element computation as obtained from `MadGraph`. The lepton and photon distributions

W pt bin	N_{QCD}	total systematic (%)
0 - 4 GeV	345±40	38.0
4 - 8 GeV	870±46	25.4
8 - 15 GeV	1440±39	14.0
15 - 23 GeV	1215±23	8.7
23 - 30 GeV	719±16	4.7
30 - 38 GeV	510±12	5.9
38 - 46 GeV	298±9	4.9
46 - 55 GeV	201±7	10.6
55 - 65 GeV	138±6	11.3
65 - 75 GeV	90±5	8.7
75 - 85 GeV	59±4	12.8
85 - 95 GeV	42±4	21.4
95 - 107 GeV	27±4	32.2
107 - 120 GeV	26±4	27.2
120 - 132 GeV	24±4	44.7
132 - 145 GeV	14±3	28.7
145 - 175 GeV	11±3	24.8
175 - 210 GeV	22±4	12.6
210 - 300 GeV	7±2	100.8

Table 5.5: QCD background vs. $p_T(W)$, in the electron channel.

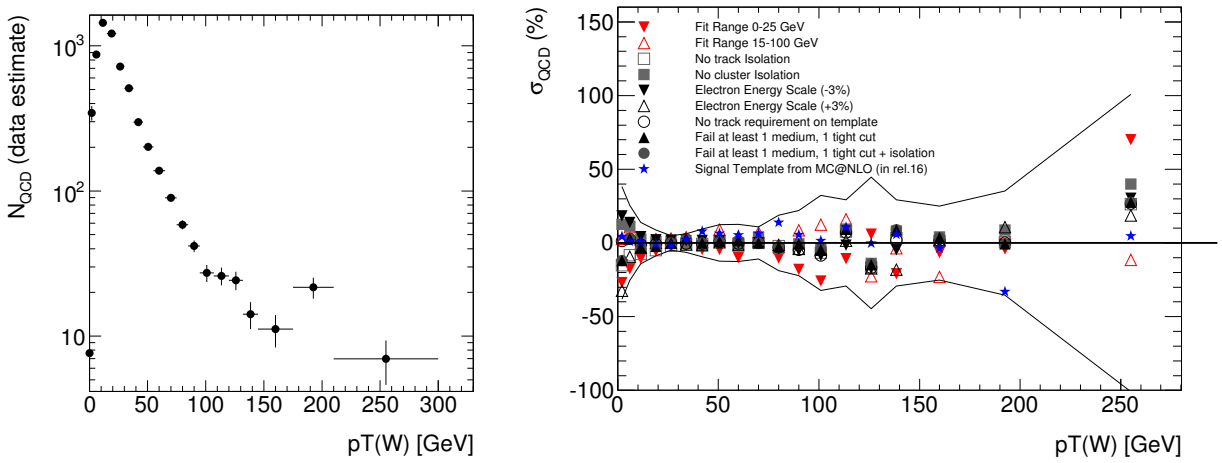


Figure 5.7: Estimated QCD background in the electron channel, and evaluation of the associated uncertainty.

obtained in both cases are significantly different, as can be seen in Figure 5.8. In particular, the lepton transverse momentum spectrum is much harder in MadGraph, and the lepton-photon opening angle is larger in this case.

On the other hand, the fiducial selection efficiency remains very similar in both cases, as shown in Figure 5.9, indicating the fiducial cross section measurement is insensitive to these differences. The cross sections for this process are also very similar: the fraction of PYTHIA events satisfying the $W\gamma$ signal definition is 0.5% which, normalizing to the inclusive cross sections given in Table 5.1 gives $\sigma_{W\gamma}^{\texttt{PYTHIA}} = 50$ pb. The MadGraph prediction is $\sigma_{W\gamma}^{\texttt{MadGraph}} = 58$ pb, agreeing within ∼15%.

We conclude form this study that the rate of events with hard photons from the W production vertex is correctly predicted by PYTHIA, and that the event selection for this event category is under control. From the above numbers and the residual differences between the two generators, we can assign a 0.1% systematic to this part of the signal model, independently of p_T^W. The normalized observable being measured, $1/\sigma d\sigma/dp_T^W$, is not affected by this uncertainty.

5.4.4 $p_T(W)$ distribution with data-driven QCD background

We present here the improved p_T^W distributions, replacing the Monte Carlo QCD background prediction by the measurements described above. As can be seen from Figure 5.10, very satisfactory agreement is obtained.

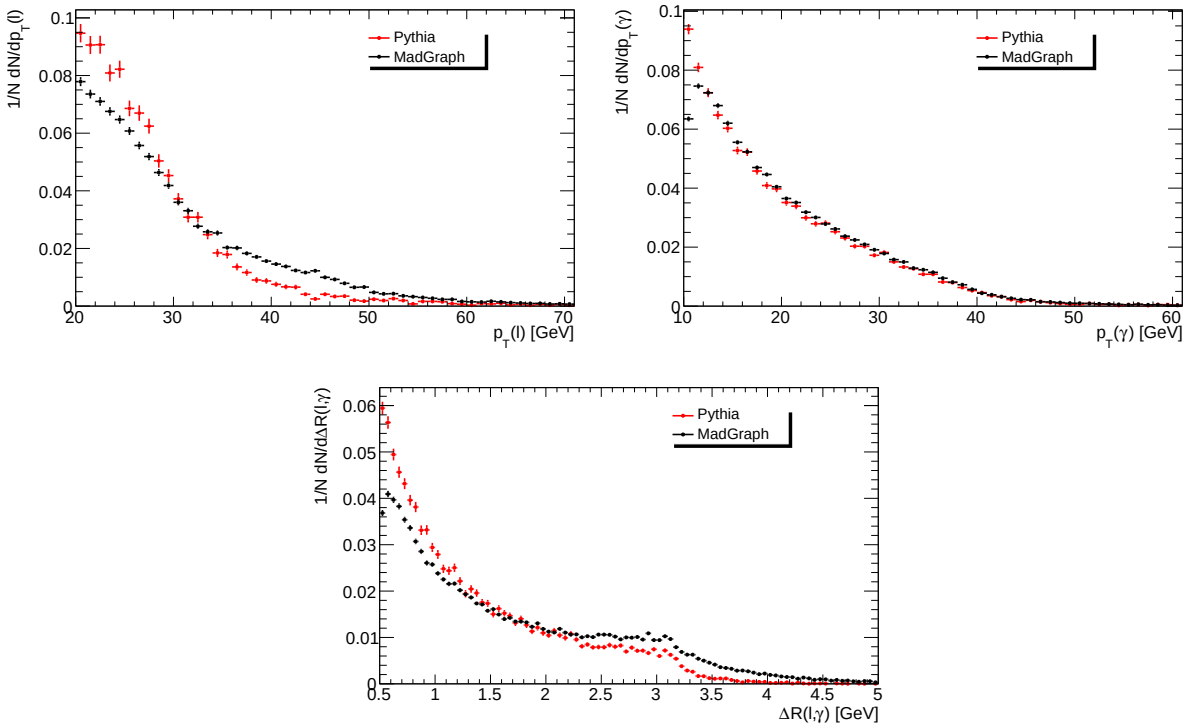


Figure 5.8: Final state, truth-level lepton and photon distributions, as obtained from the PYTHIA and MadGraph generators.

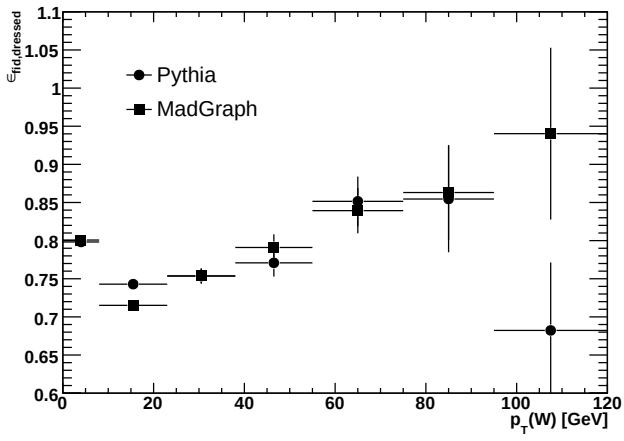


Figure 5.9: Fiducial selection efficiency as a function of the W boson transverse momentum, as obtained from the PYTHIA and MadGraph generators.

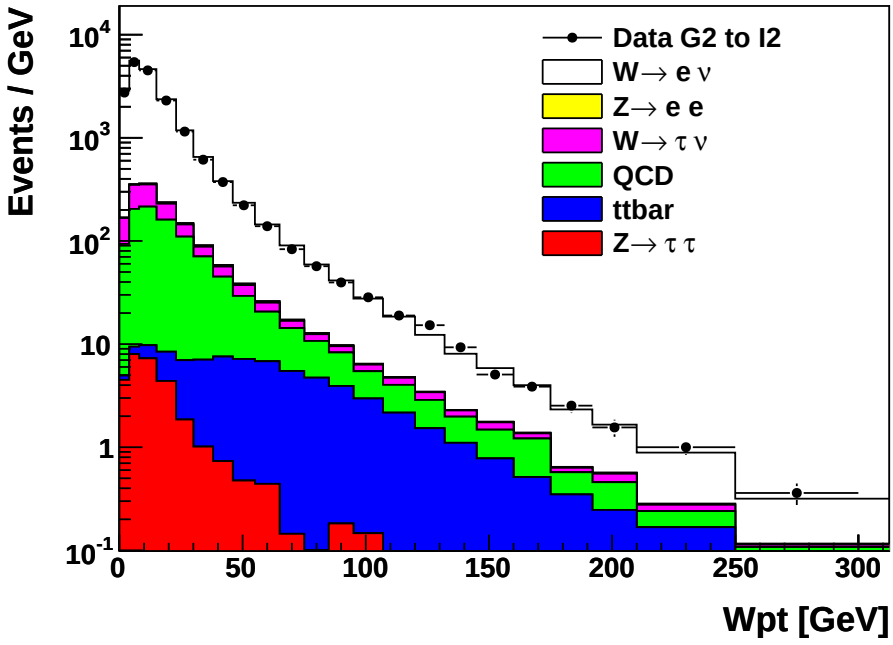


Figure 5.10: Observed and predicted $p_T(W)$ distribution in $W \to e\nu$ events, including data-driven jet background determination.

5.5 Definition of the Response Matrix

The response matrix used in the unfolding as described in Section 5.6 expresses the transformation between the true quantity to be measured and the observed quantity. In this analysis of p_T^W measurement, the response matrices are defined as $p_T^{W,\text{true}}$, $p_T^{W,\text{reco}}$correlation histograms in which the p_T^W is calculated from the hadronic recoil algorithm given in Chapter 4 , with the binning of each axis is defined in Section 5.5.1.

To define the response matrix for this analysis, we proceed in three steps as detailed below. First estimations of the response matrices are obtained from the Monte Carlo signal samples. The response is then parametrized, in order to smooth the matrices, compensating for the limited size of the samples, and preparing for subsequent data-driven corrections to the response parameters. This parametrization comes at the cost of a potential bias, to be evaluated. Corrections to the response parameters are then estimated from Z events, defining our baseline response matrices.

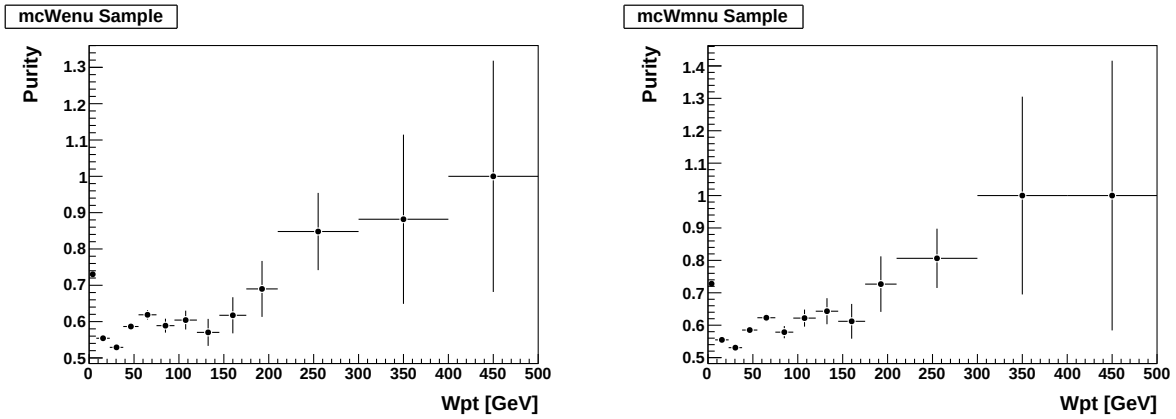
5.5.1 Binning

Exploiting the hadronic recoil algorithm described in Chapter 4 and its performance, we define the binning of our unfolded measurement to minimize the bin size while maintaining stability in the unfolding. We define a slightly finer binning for the reconstruction-level p_T^W distribution, in order to provide more information to the unfolding procedure while keeping sufficient statistics for reliable background subtraction. We use the following bin boundaries, in GeV:

- Reconstruction-level, $p_T^{W,\text{reco}}$, distribution : 0, 4, 8, 15, 23, 30, 38, 46, 55, 65, 75, 85, 95, 107, 120, 132, 145, 160, 175, 192, 210, 250,300, 400, 500.
- Truth-level (unfolded), $p_T^{W,\text{true}}$ distribution : 0, 8, 23, 38, 55, 75, 95, 120, 145, 175, 210, 300, 400, 500.

While the binning of the final result is considered in the limit of 300 GeV, the binning on both reconstruction and truth level applied during the unfolding process are extended to 500 GeV in order to avoid edge effects due to migration outside of the range of response matrices.

Figure 5.11 displays the expected purity of the measurement bins in electron and muon channel, as calculated from the PYTHIA signal samples. The purity is defined as the fraction of events with $p_T^{W,\text{reco}}$falling in the same bin as $p_T^{W,\text{true}}$.

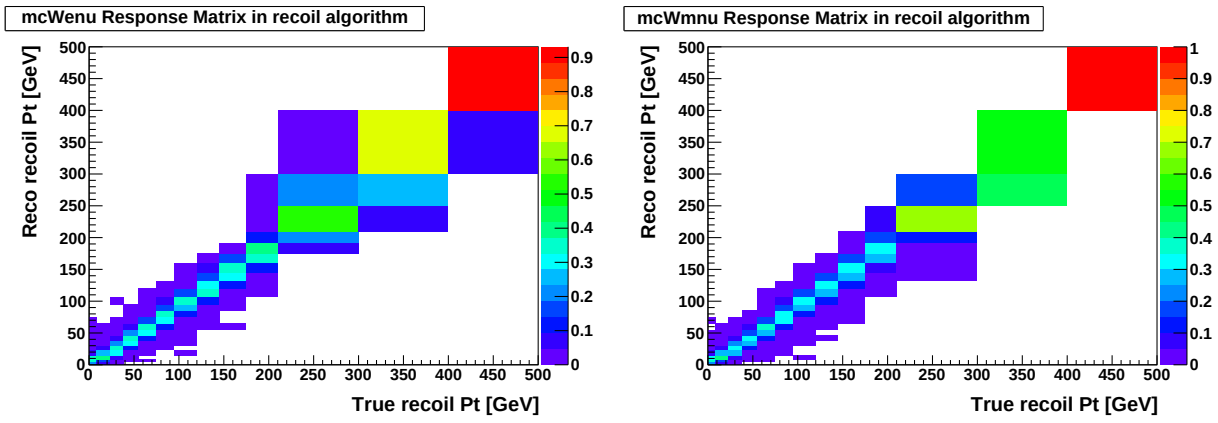


(a) Electron channel (b) Muon channel

Figure 5.11: Purity in electron and muon channels.

5.5.2 Monte Carlo based response matrices

A first estimate of the response matrices is obtained from simply plotting the correlation histograms obtained from the Monte Carlo, populated by the simulated events surviving the event selections. The response matrices for electron and muon channels as obtained from the PYTHIA signal samples are displayed in Figure 5.12.



(a) Electron channel (b) Muon channel

Figure 5.12: Monte Carlo based response matrices in the electron and muon channels.

5.5.3 Response parametrization

The two dynamical variables that affect the response to p_T^{true}, the true hadronic recoil, are the hadronic activity of the event, ΣE_T, and p_T^{true} itself. Other event characteristics, such as the lepton kinematics, are not relevant when using the hadronic recoil calculation defined in Chapter 4.

Following the parametrization procedure as described in Chapter 4, we parametrize the hadronic recoil response in two components which are parallel and perpendicular to the direction of the true transverse momentum of W boson. At each point in (ΣE_T, p_T^{true}) space, the response on both components is gaussian to a very good approximation. This property is justified in appendix A, and can be seen as resulting from the central limit theorem. The gaussian response is defined by a bias Δ, essentially function of p_T^{true}, and resolution terms $\sigma_\parallel$ and $\sigma_\perp$, function of both $p_T^{\text{true}}$and ΣE_T, so that our parametrization reads:

$$p_\parallel(p_T^{\text{true}}, \Sigma E_T) = p_T^{true} + G(\Delta_W^{MC}(p_T^{\text{true}}), \sigma_W^{MC}(p_T^{\text{true}}, \Sigma E_T)); \quad (5.8)$$

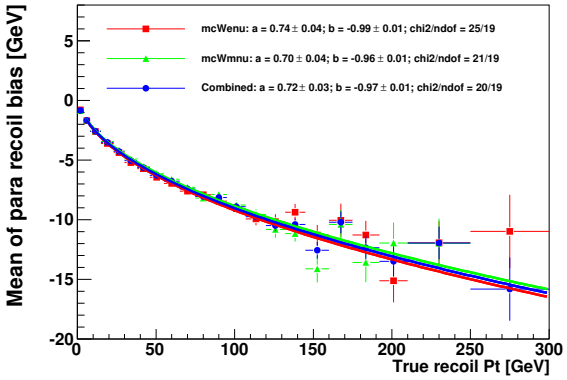
$$p_\perp(p_T^{\text{true}}, \Sigma E_T) = 0 + G(0, \sigma_W^{MC}(p_T^{\text{true}}, \Sigma E_T)). \quad (5.9)$$

where G is the gaussian probability density function. This way of parameterization is called "smearing" technic in which two components of the hadronic recoil are reconstructed by smearing the truth, given their bias and resolution respecting to the truth. Performing (5.8) and (5.9) event by event to get the $p_\parallel$ and the $p_\perp$, the hadronic recoil is then reconstructed by $p_T^{reco} = \sqrt{p_\parallel^2 + p_\perp^2}$. The distribution of the hadronic recoil obtained from this parametrization is expected to be the same as that reconstructed from the hadronic recoil calculation given in Chapter 4.

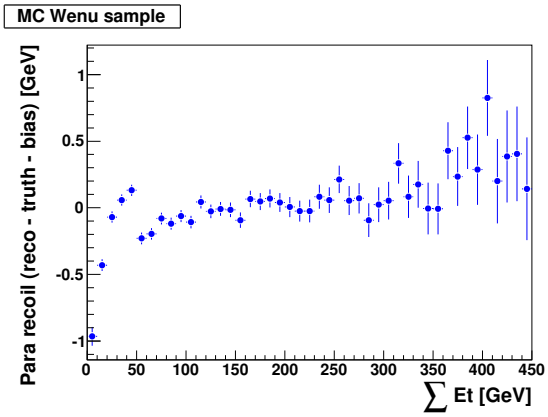
Fits to the bias Δ as a function of p_T^{true} are shown in Figure 5.13(a). As advertised, the bias along ΣE_T is negligible and we neglect this component (see Figures 5.13(b) and 5.13(c)). As can also be seen, $\Delta_W = 0$ when p_T^{true}=0, and since $p_\perp^{true}$ is identically 0 we assume no bias along this direction.

Fits to $\sigma_\parallel$ and $\sigma_\perp$ are displayed in Figure 5.14, showing good agreement between the channels as expected from the hadronic recoil algorithm. This allows for a combination of fits obtained from individual channels. The dependence of $\sigma_\parallel$ and $\sigma_\perp$ as a function of p_T^{true} is illustrated in Figure 5.15. Along the parallel projection, where the influence of the recoil jet is maximal, a dependence vs. p_T^{true} is observed as expected. Along the perpendicular projection, no such dependence is seen. Hence, we parametrize $\sigma_\parallel$ as a function of ΣE_T separately in three different regions of p_T^{true}: $p_T^{\text{true}} \leq 8$ GeV, $8 < p_T^{\text{true}} \leq 23$ GeV, $p_T^{\text{true}} > 23$ GeV.

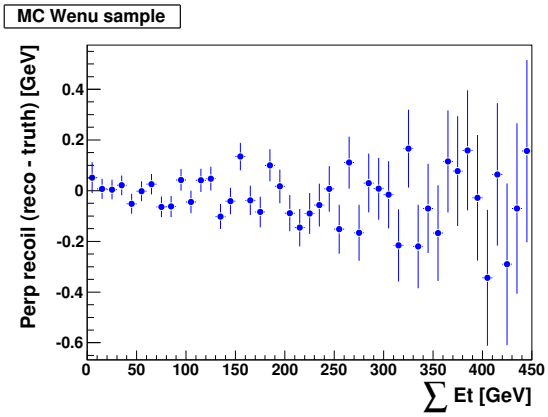
The quality of the fits and the parameterization procedure are checked by comparing the distributions of the parallel and perpendicular recoil obtained by the smearing technique and those calculated from the hadronic recoil algorithm. A very good agreement between them as shown in Figures 5.16 (for the parallel component of the recoil) and 5.17 (for the perpendicular component of the recoil) estimated in $W \to e\nu$ Monte Carlo sample can validate the proposed parameterization method.



(a) Fits to bias of parallel recoil as a function of $a + b.\sqrt{p_T^{\text{true}}}$.

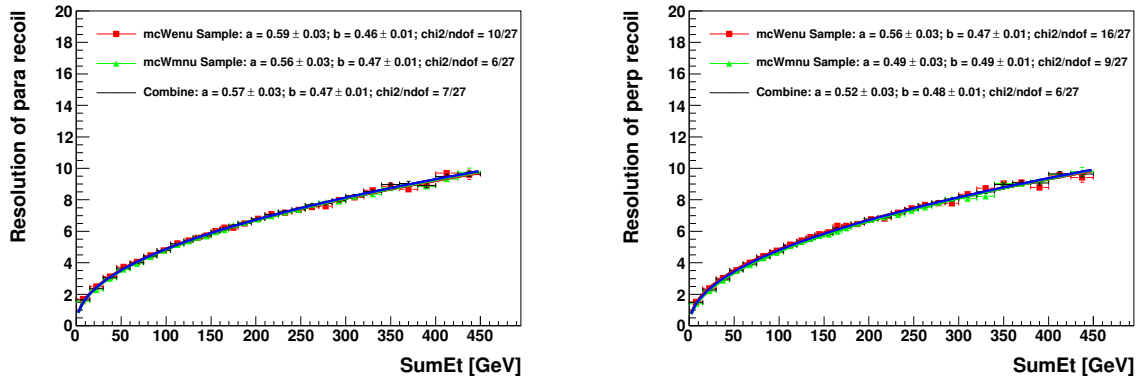


(b) Averaged bias of parallel recoil in bins of ΣE_{T} after correcting the bias shown in Figure 5.13(a), which was observed along the true recoil p_T^{true}.



(c) Averaged bias of perpendicular recoil in bins of ΣE_{T} .

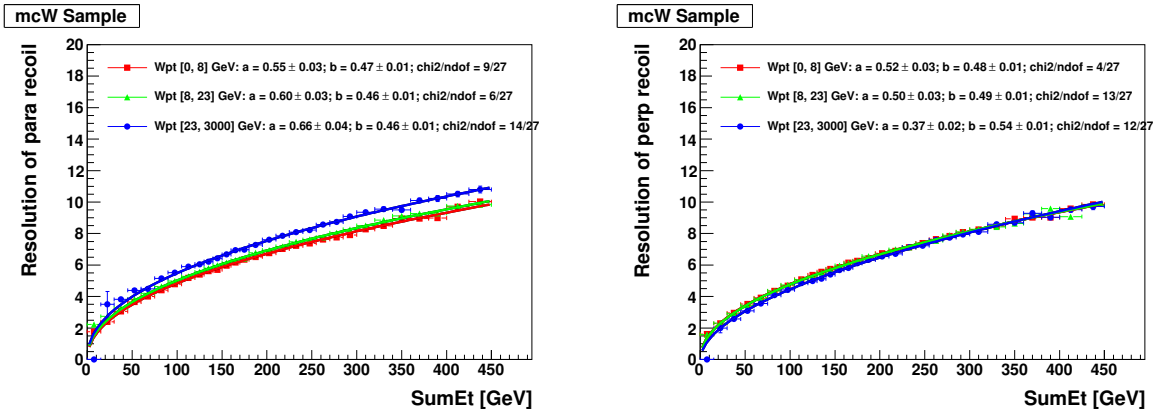
Figure 5.13: Fits to the recoil bias in simulated W events (in red: electron channel, in green: muon channel and in blue: combined channel). The bias is defined as the mean of the distribution of reconstructed minus true hadronic recoil .



(a) Resolution of parallel recoil.

(b) Resolution of perpendicular recoil.

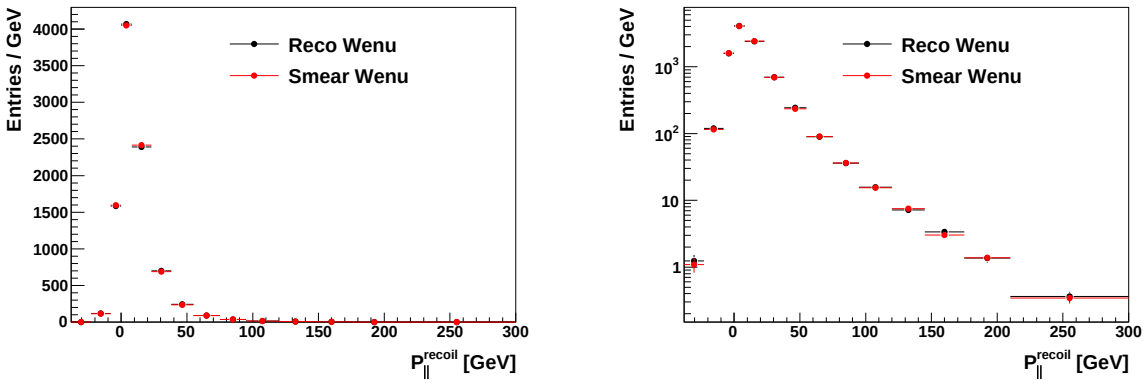
Figure 5.14: Fits to the recoil resolution in simulated W events (electron channel in red, muon channel in green and combined channel in blue), averaged over p_T^{true}, as a function of $a+b\sqrt{(\Sigma E_{\text{T}})}$. The resolution is the RMS of the bias distribution. "Para recoil" ($\sigma_{\parallel}^{W,MC}$) refers to the component of the recoil along the direction of the boson, and "perp recoil" ($\sigma_{\perp}^{W,MC}$) refers to the component perpendicual to the boson direction.



(a) Resolution of parallel recoil.

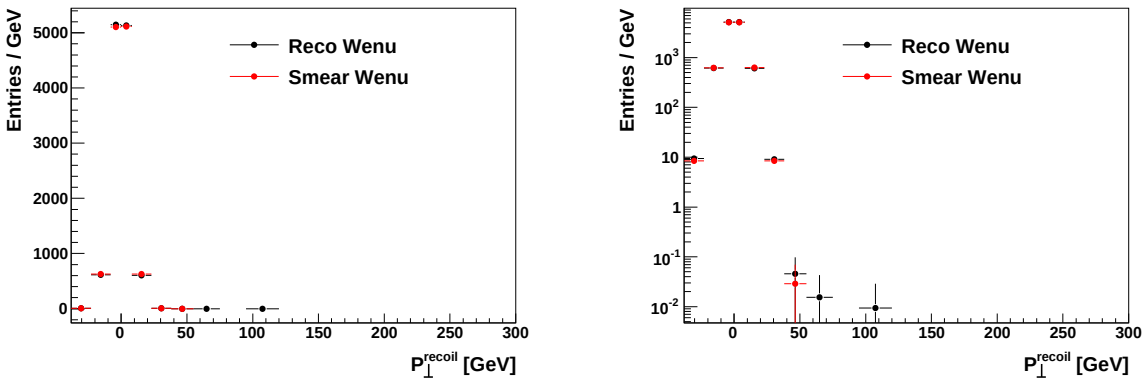
(b) Resolution of perpendicular recoil.

Figure 5.15: Dependence of the parallel and perpendicular recoil resolution curves on p_T^{true} in simulated W events. In red: $0 < p_T^{\text{true}} \leq 8$ GeV; in green: $8 < p_T^{\text{true}} \leq 23$ GeV; in blue: $p_T^{\text{true}} > 23$ GeV. Fit function: $f(x) = a+b\sqrt{x}$



(a) Linear scale. (b) Log scale.

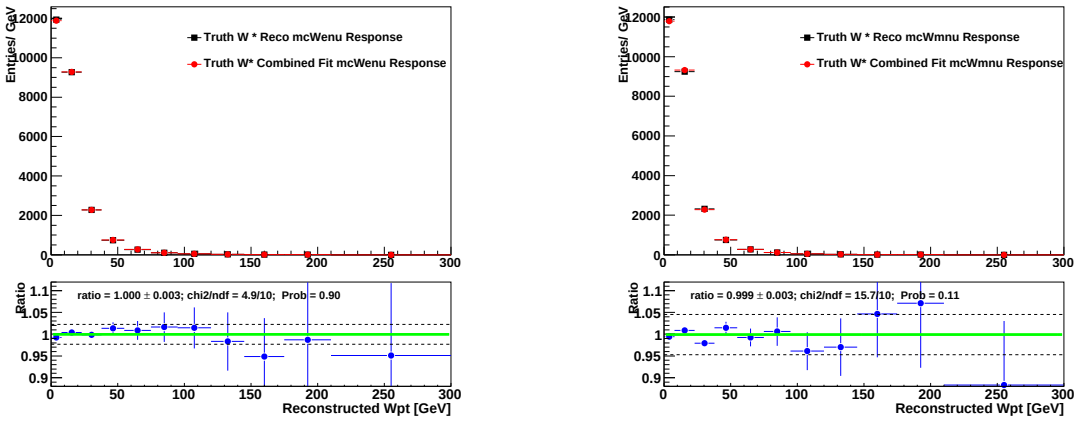
Figure 5.16: Comparison of the distributions of the parallel recoil obtained from the smearing technic (red) and reconstructed from the hadronic recoil algorithm (black).



(a) Linear scale. (b) Log scale.

Figure 5.17: Comparison of the distributions of the perpendicular recoil obtained from the smearing technique (red) and reconstructed from the hadronic recoil algorithm (black).

The parametrized response matrices is determined from the correlation between the true p_T^W and the p_T^W obtained by the smearing technic. This parametrized response matrix is expected to be similar with the nominal one determined in Section 5.5.2. Therefore, in order to have another check for the quality of the parametrization played in the aspect of response matrix, we compare the p_T^W distributions obtained by folding the truth-level p_T^W distribution by this parametrized response matrix and by the nominal response matrix. The result of this closure test is displayed in Figure 5.18. A systematic uncertainty related to this procedure is assigned to the measurement, as described in Section 5.8.



(a) Electron channel.

(b) Muon channel.

Figure 5.18: Comparison of the distributions of $p_T^{W,\text{true}}$ folded by the nominal (black) and by the parametrized response matrices (red).

5.5.4 Data driven corrections and baseline response matrices

The response parameters corrected is tuned from Z data as follow:

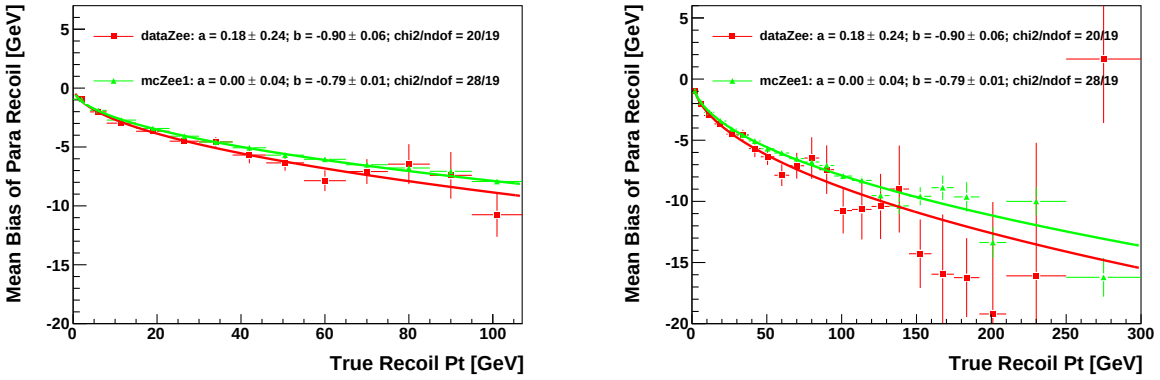
$$\Delta_W^{cor} = \Delta_W^{MC} + (\Delta_Z^{data} - \Delta_Z^{MC}), \tag{5.10}$$

$$\sigma_{\parallel,\perp}^{W,cor} = \sigma_{\parallel,\perp}^{W,MC} + (\sigma_{\parallel,\perp}^{Z,data} - \sigma_{\parallel,\perp}^{Z,MC}). \tag{5.11}$$

The bias and resolution parameters measured in Z events are extracted from the momentum imbalance between the measured hadronic recoil and dilepton transverse momentum. Fits to these parameters in Z Monte Carlo and data are displayed in Figures 5.19 and 5.20.

In the calculation of the response parameters, events from Z data, Z Monte Carlo, and W Monte Carlo are all reweighted to match the ΣE_T distribution observed in the W data. This is done to correct

for the different ΣE_T distributions expected for W and Z production, as well as any inaccuracies in the modeling of the ΣE_T distribution in the Monte Carlo. The ΣE_T distributions used to calculate the event weights as well as ΣE_T event weights are shown in Figure 5.21.



(a) Bias fits in low range of $p_{\mathrm{T}}^{\ell^+\ell^-}$

(b) Bias fits in full range of $p_{\mathrm{T}}^{\ell^+\ell^-}$

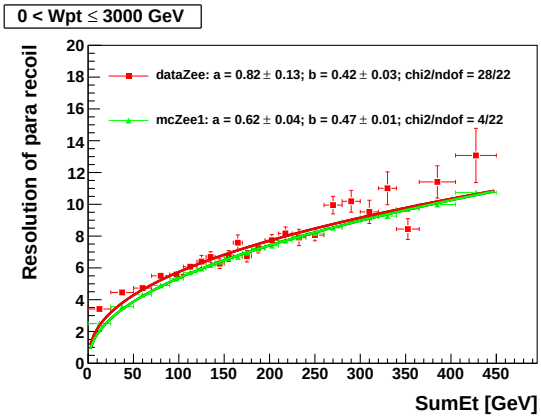
Figure 5.19: Fits to bias of parallel recoil in data (red) and simulation (green) of Z events: $\Delta_Z = a + b.\sqrt{p_{\mathrm{T}}^{\ell^+\ell^-}}$. The bias ($\Delta_Z$) is defined as the mean of the distribution of reconstructed minus true hadronic recoil where the true hadronic recoil is against the dilepton transverse momentum, $\vec{p}_{\mathrm{T}}{}^{true} = -\vec{p}_{\mathrm{T}}{}^{\ell^+\ell^-}$.

5.5.5 Corresponding Systematic Uncertainties

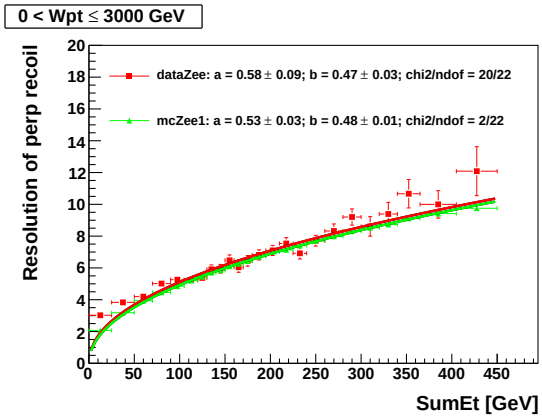
The systematic uncertainties corresponding to this procedure are the following:

- the assumed functional form of the response parametrization, and its statistical precision as obtained from our finite-size samples;
- the assumption that the data driven correction can be read off from data - Monte Carlo discrepancies in Z events;
- the statistical precision of the parameter fits in Z data and Monte Carlo.

These will be evaluated in Section 5.8.

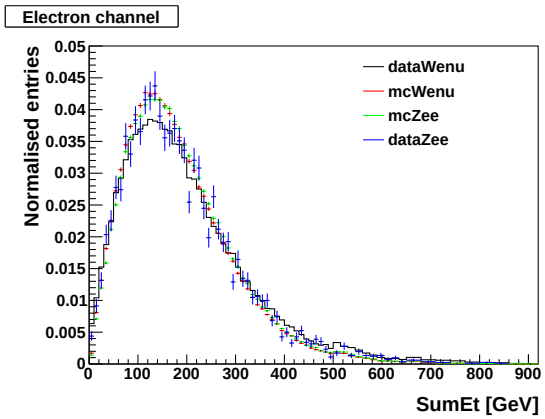


(a) Resolution of parallel recoil

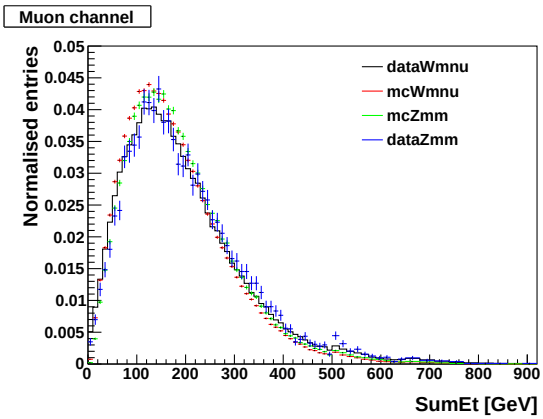


(b) Resolution of perpendicular recoil

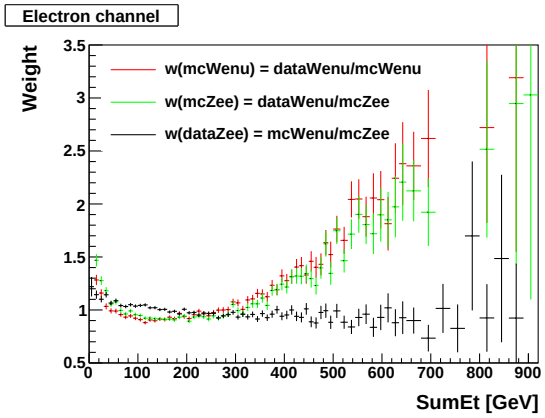
Figure 5.20: Fits to resolution of recoil along parallel and perpendicular projections in data (red) and simulation (green) of Z events: $\sigma^Z_{\parallel,\perp} = a + b\sqrt{(\Sigma E_{\mathrm{T}})}$. The resolution ($\sigma^Z$) is defined as the width of the distribution of reconstructed minus true hadronic recoil where the true hadronic recoil is against the dilepton transverse momentum. "Para recoil" ($\sigma^Z_{\parallel}$) and "perp recoil" ($\sigma^Z_{\perp}$) refer to the component of the recoil along and perpendicular to the direction of the dilepton transverse momentum, respectively.



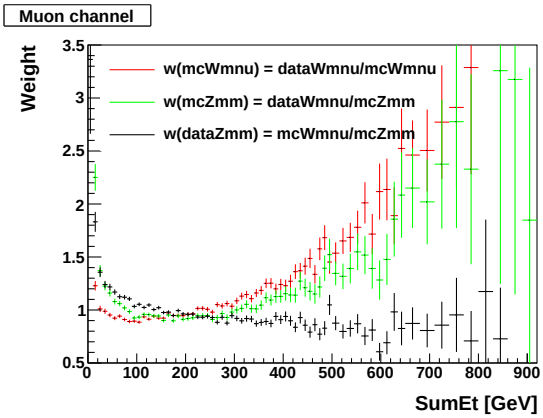
(a) ΣE_{T} distributions in electron channel.



(b) ΣE_{T} distributions in muon channel.



(c) ΣE_{T} event weights in electron channel.



(d) ΣE_{T} event weights in muon channel.

Figure 5.21: ΣE_T distributions ((*a*) and (*b*)) in W Monte carlo (red), Z Monte Carlo (green) and Z data (blue) versus that in W data (black) and ΣE_T event weight ((*c*) and (*d*)) for W Monte Carlo, Z Monte Carlo and Z data.

5.6 Unfolding Methods

In experiments, the distribution of the measured observable is smeared and distorted from that of the corresponding true physical quantities due to detector effects, such as limited acceptance, imperfect efficiency, and finite resolution. Mathematically, given some true variable x (to be determined in an experiment) distributed according to its probability density function (p.d.f.) $f(x)$, can not be measured perfectly due to both experimental and statistical uncertainties. As a result, instead of measuring x one typically measures a related variable y distributed according to a different p.d.f $g(y)$. The relation between $f(x)$ and $g(y)$ can be expressed as a convolution of the true distribution $f(x)$ with a kernel $A(y,x)$ such that,

$$g(y) = \int A(y,x) f(x) \mathrm{d}x. \tag{5.12}$$

(5.12) is known as the first kind of Fredholm integral equation. The kernel $A(y,x)$ is the response function describing the detector effects on the measurement. In practice, a known background contribution estimated from data and simulation has to be added to the right hand side of equation (5.12). But in this section with the purpose of presenting the methodology, to be simple, this contribution is ignored. From this equation, the true distribution $f(x)$ can be determined from knowing the response function $A(y,x)$ and measured distribution $g(y)$. This procedure of determining the true distribution $f(x)$ is usually called "unfolding" or "deconvolution" or "unsmearing".

For the numerical solution of equation 5.12, the distributions can be presented by histograms and the resolution function can be presented by a matrix in order to have a finite set of parameters. Equation 5.12 then becomes having a matrix form

$$\mathbf{y} = \mathbf{A}\mathbf{x}. \tag{5.13}$$

The vector $\mathbf{y}$ presents a histogram with n bins of measured quantity while the vector $\mathbf{x}$ presents a histogram with m bins of the true quantity to be measured. The transformation from $\mathbf{x}$ to $\mathbf{y}$ is performed by the so-called response matrix $\mathbf{A}$ with a dimension of $n \times m$. The elements of the response matrix a_{ij} can be considered as the probability for a true value x_j to be measured as a value y_i.

Assuming that the response matrix $\mathbf{A}$ depends on only detector effects and is known, the equation (5.13) can be solved for $\mathbf{x}$, given the vector $\mathbf{y}$. The detector response can be modeled by using Monte

Carlo simulation of underlying physics where the interactions of predicted particles with the matter content of detector are simulated. In addition, it may be possible to use control regions of the data itself to create a fully or partially data-driven response matrix. For our case of p_T^W measurement, the determination of the response matrix is described in details in Section 5.5.

Actually, given the response matrix, there are two methods of determining the underlying true distribution from a measured spectrum:

- **Smearing**: by varying the parameters of the underlying physics model, then smearing them with the detector model, one can obtain a series of smeared distributions. These smeared distributions are compared to the measured data and the best fit obtained would determine the true measured distribution.

- **Unfolding**: using MC determined detector smearing, one can invert the smearing effects of the measured data to produce directly a measured true distribution, i.e. look for a solution of the form $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$.

The smearing method is preferable to use in case the main goal of the measurement is to compare with an existing theory since it is mathematically more simple than the unfolding method. However, the unfolding method is needed in order to easily compare results with other experiments or with theoretical models which may be developed in the future. In addition, for experiments with complex underlying theories and intricate detectors, smearing may be extremely computationally intensive. For these reasons, we have chosen to unfold our results.

There are many unfolding procedures in which a few of them aimed for high energy physics are well described in [14, 35, 52, 67]. In addition, a formal discussion of the method of statistical and computational inverse problems can be found in [60]. In this section, We will describe three of these methods including the bin-by-bin unfolding method, unfolding based on regularized matrix inversion method and Bayesian unfolding method based on Bayes' theorem.

5.6.1 Notation

In this section, we use the following notation:

- One dimensional histograms or vectors are denoted by small letters (e.g. $\mathbf{x}$, $\mathbf{y}$)

- Two dimensional histograms or matrices are denoted by capital letters (e.g. $\mathbf{A}$)

- Bold letters indicate a vector or matrix (e.g. $\mathbf{x}$, $\mathbf{A}$)
- Regular (non-bold) letters indicate a scalar (e.g. τ)
- Regular letters with indices indicate vector or matrix elements (e.g. x_j, A_{ij})
- Summation is **NOT** implied over repeated indices. Summation is explicit (e.g. $\sum_j A_{ij}x_j$)
- Covariance matrices are indicated by a bold and capital $\mathbf{V}$ followed by the variable in parentheses or as a subscript (e.g. for vector $\mathbf{y}$ the covariance matrix is $\mathbf{V}(\mathbf{y})$ or $\mathbf{V}_y$)

5.6.2 Bin-by-bin unfolding

Bin-by-bin is the simplest of three unfolding methods presented here. In this method, the true histogram $\mathbf{x}^{\mathbf{data}}$ to be measured in data is unfolded bin by bin from the observed data histogram $\mathbf{y}^{\mathbf{data}}$ as follows:

$$x_i^{data} = y_i^{data}/C_i. \tag{5.14}$$

where i is the bin number and C_i is a correction factor describing the ratio of observed to true events in bin i of histograms. This correction factor is calculated from Monte Carlo simulation for the signal process as given by this formula,

$$C_i = y_i^{sim}/x_i^{sim}. \tag{5.15}$$

Here y_i^{sim} and x_i^{sim} are the contents of bin i of observed and true histograms respectively in a given simulation sample.

This method has number of advantages. First it is conceptually very simple and it can neglect the unavoidable correlations between adjacent bins. However, it has also a number of disadvantages. The most important one is that all of the corrections are within bin, as reflected by the name of the method. In addition, it cannot take into account large migrations of events from a bin to the others. Thus this method can give a reliable result only if the amount of migration is negligible and if the standard deviation of the smearing is smaller than the bin size so that the bin purity is high enough. The reason is that when bin purities are low, the unfolded result is highly biased to the control sample used to produce the bin-by-bin corrections.

In this measurement, the bin-by-bin unfolding is not used as it is expected that the bin-to-bin migration will be a large effect.

5.6.3 Response Matrix Inversion

Difficulties with Matrix Inversion

To illustrate the difficulties of matrix inversion, we will discuss a simple example which is similar to the one given in [14]. Assuming that the response matrix $\mathbf{A}$ is a square $n \times n$ matrix. This assumption is just a case of the response matrix form where it is reversible to examine problems of unfolding with matrix inversion. But in general, the unfolding techniques do not require the response matrix to be symmetric. In this case, a direct solution $\hat{\mathbf{x}}$ can be computed as

$$\hat{\mathbf{x}} = \mathbf{A}^{-1}\mathbf{y} \qquad \text{with covariance matrix:} \quad \mathbf{V}(\hat{\mathbf{x}}) = \mathbf{A}^{-1}\mathbf{V}(\mathbf{y})(\mathbf{A}^{-1})^T. \tag{5.16}$$

Where $\mathbf{V}(\mathbf{y})$ is the covariance matrix of the measured data. This data covariance matrix signifies the amount by which the data $\mathbf{y}$ may differ from the expectation of the smeared true distribution $\mathbf{Ax}$. This solution has nice statistical properties, in that the expectation E of the solution is unbiased because

$$E[\hat{\mathbf{x}}] = \mathbf{A}^{-1}E[\mathbf{y}] = \mathbf{A}^{-1}\mathbf{A}E[\mathbf{x}] = E[\mathbf{x}]. \tag{5.17}$$

Although the solution $\hat{\mathbf{x}}$ may be unbiased, this solution is in practice not satisfactory when the response matrix has large off-diagonal elements (this can happen for example if the bin size is too small compared to the measurement resolution). In this case of large bin-to-bin migration, this inversion procedure can introduce large oscillations in the solution and cause the covariance matrix of the solution to become extremely large (essentially to account or these oscillations). Explanations with illustrations by figures for this problem is also detailed in [34].

Mathematically, to understand the source of these oscillations, we can look at the equation for the solution $\hat{\mathbf{x}}$ as given at (5.16). In this example, $\mathbf{A}$ is a symmetric matrix and can be diagonalized by using this transformation,

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^T, \tag{5.18}$$

where U is an orthogonal matrix satisfying $\mathbf{U}\mathbf{U}^T = \mathbf{1}$ (with columns which are the eigenvectors $\mathbf{u}_j$ of the matrix $\mathbf{A}$), and $\mathbf{D}$ is diagonal with entries that are the eigenvalues λ_j of $\mathbf{A}$ (in decreasing order). We can now rotate equation (5.13) with the multiplication of $\mathbf{U}^T$

$$\mathbf{U}^T\mathbf{y} = \mathbf{U}^T\mathbf{A}\mathbf{x} = \mathbf{U}^T\mathbf{U}\mathbf{D}\mathbf{U}^T\mathbf{x} = \mathbf{D}\mathbf{U}^T\mathbf{x}$$

By defining $\mathbf{c} = \mathbf{U}^T\mathbf{y}$ and $\mathbf{b} = \mathbf{U}^T\mathbf{x}$, we arrive at the new equation with diagonal response matrix,

$$\mathbf{D}\mathbf{b} = \mathbf{c} \tag{5.19}$$

From equation (5.19), we can see that each of the elements of $\mathbf{b}$ and $\mathbf{c}$ transform independently of other elements and satisfy the simple equation $c_j = \lambda_j b_j$. Inversion of $\mathbf{D}$ to solve for the elements b_j is now paramount to dividing each element c_j by the corresponding eigenvalue λ_j, i.e. $b_j = c_j/\lambda_j$. The final step to unfolding would proceed by multiplying $\mathbf{b}$ with $\mathbf{U}$,

$$\mathbf{x} = \mathbf{U}\mathbf{b} = \sum_j b_j \mathbf{u}_j = \sum_j \frac{c_j}{\lambda_j}\mathbf{u}_j. \tag{5.20}$$

By looking at the solution as obtained in (5.20), we can easily explain for the problem mentioned. The c_j elements come from the transformation $\mathbf{U}^T\mathbf{y}$ and thus contain all of the statistical fluctuations of the original data. When the eigenvalues λ_j are very small (much smaller than 1), the coefficients b_j become very large. Eventually, for very small eigenvalues λ_j, the solution $\mathbf{x}$ will be dominated by one or by few of the coefficients b_j with small eigenvalues and large statistical errors. And this leads to a completely unsatisfactory of solution $\mathbf{x}$.

Regularized Matrix Inversion

The difficulties of matrix inversion as indicated by the above example in Section 5.6.3 are known as an ill-posed problem with unfolding. There are several techniques to solve this problem but in this section we only concentrate on one method called Tikhonov regularization [84], which is nicely reviewed in [51]. In fact, the solutions described in the following sections have been implemented within the ROOT data analysis framework classes TUnfold and TUnfoldSys [79, 80].

In this section, the response matrix $\mathbf{A}$ is not any more required to be a square matrix but rather than a form of size $n \times m$, where $n \geq m$ is assumed. With this form of the response matrix, the measured histogram $\mathbf{y}$ can be of size n and the true histogram which is the solution of unfolding method $\mathbf{x}$ can be of size m. This choice of number of bins in the measured histogram to be greater than number of bins in the true histogram is to perform better the regularization when choosing parameters and to reduce correlations between bins of the unfolded solution [14].

Finding regularized solutions to inversion problems involves a trade-off between the size (or amount) of regularization and the quality of the unfolding fit to the data. We thus can pose this matrix inversion

problem as least-squares minimization problem where,

$$\chi^2 = (\mathbf{Ax}-\mathbf{y})^T\mathbf{V}_y^{-1}(\mathbf{Ax}-\mathbf{y}) + \tau^2(\mathbf{Lx})^T(\mathbf{Lx}). \tag{5.21}$$

Here τ is the regularization parameter which must be chosen by the user (or an optimization method for finding the "best" τ must be chosen, as discussed in Section 5.6.3), and $\mathbf{L}$ is so called Tikhonov matrix defining a regularization condition on the unfolded solution $\mathbf{x}$. The χ^2 in equation (5.21) contains two parts. The first one, $(\mathbf{Ax}-\mathbf{y})^T\mathbf{V}_y^{-1}(\mathbf{Ax}-\mathbf{y})$, is the standard χ^2 measuring the residual between our data $\mathbf{y}$ and the unfolded solution $\mathbf{x}$ taking into account both the response matrix $\mathbf{A}$ and the errors on the data $\mathbf{V}_y$. The second one, $(\mathbf{Lx})^T(\mathbf{Lx})$, can be though of as the "size" of the regularization condition. The parameter τ then determines the relative importance of the regularization and the residual of the fit.

In this method, the unfolded $\mathbf{x}$ can be solved by requiring a minimization of the χ^2 given at (5.21) respecting to $\mathbf{x}$

$$\begin{aligned}\frac{d\chi^2}{d\mathbf{x}} = 0 &= (\mathbf{A})^T\mathbf{V}_y^{-1}(\mathbf{Ax}-\mathbf{y}) + (\mathbf{Ax}-\mathbf{y})^T\mathbf{V}_y^{-1}(\mathbf{A}) + \tau^2(\mathbf{L})^T(\mathbf{Lx}) + \tau^2(\mathbf{Lx})^T(\mathbf{L}) \\ &= ((\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{A}+\tau^2\mathbf{L}^T\mathbf{L})\mathbf{x} - \mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{y}) + ((\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{A}+\tau^2\mathbf{L}^T\mathbf{L})\mathbf{x} - \mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{y})^T \\ &\rightarrow \qquad (\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{A}+\tau^2\mathbf{L}^T\mathbf{L})\mathbf{x} = \mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{y}. \end{aligned} \tag{5.22}$$

From (5.22), in priciple, for a given value of τ we can get a corresponding unfolded solution

$$\mathbf{x}_\tau = (\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{A}+\tau^2\mathbf{L}^T\mathbf{L})^{-1}\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{y}. \tag{5.23}$$

The covariance matrix for the solution can be then calculated as follows:

$$\mathbf{V}(\mathbf{x}_\tau) = \mathbf{V}_x = (\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{A}+\tau^2\mathbf{L}^T\mathbf{L})^{-1}\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{A}(\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{A}+\tau^2\mathbf{L}^T\mathbf{L})^{-1}. \tag{5.24}$$

Equations (5.23) and (5.24) are the general solutions to our unfolding problem, and will be used for the computational determination of our solutions and uncertainties using the TUnfoldSys software in ROOT [79, 80]. In the next section we will discuss these solutions in some particular cases of parameter τ and matrix $\mathbf{L}$.

Impact of Regularization

A simple example to see an impact of regularization is to consider a case where **L** is the identity matrix, **L** = **I**. In this case the equation (5.23) becomes

$$(\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{A} + \tau^2\mathbf{I})\mathbf{x} = \mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{y}, \tag{5.25}$$

and $\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{A}$ should have a symmetric form to be consistent in dimension with the identity matrix as expressed in equation (5.25). With a symmetric form, we can diagonalize $\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{A}$ by decomposing it to $\mathbf{U}\Lambda\mathbf{U}^T$ where **U** is the matrix of eigenvectors of $\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{A}$ and Λ the diagonal matrix of associated eigenvalues. Also here we define $\mathbf{z} = \mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{y}$ then we will have

$$(\Lambda + \tau^2\mathbf{I})\mathbf{U}^T\mathbf{x} = \mathbf{U}^T\mathbf{z}$$

As before, defining $\mathbf{b} = \mathbf{U}^T\mathbf{x}$ and $\mathbf{c} = \mathbf{U}^T\mathbf{z}$, we arrive at the diagonal equation,

$$(\Lambda + \tau^2\mathbf{I})\mathbf{b} = \mathbf{c} \qquad \rightarrow \qquad b_i = \frac{c_i}{\lambda_i + \tau^2}$$

where λ_i is an eigenvalue associated to the eigenvector $\mathbf{u}_i$ of the matrix $\mathbf{A}^T\mathbf{V}_y^{-1}\mathbf{A}$. Now left multiplying **b** by **U**, we find,

$$\mathbf{x}_\tau = \mathbf{U}\mathbf{b} = \sum_j b_j\mathbf{u}_j = \sum_j \frac{c_j}{\lambda_j + \tau^2}\mathbf{u}_j \tag{5.26}$$

The role of the regularization parameter can be seen clearly from the equation (5.26). Similarly with the case before where there is no contribution of the regularization, we know that the coefficients c_j contain statistical fluctuations of the measured histogram. Several coefficients c_j will have statistical errors making their values consistent with zero. However, now these coefficients are no longer simply divided by the eigenvalue but the eigenvalue plus the regularization parameter squared . So these coefficients are not amplified when the eigenvalues are small if we choose the τ parameters big enough. In other words we can say, the regularization parameter τ serves to damp the importance of small eigenvalues and thus save the solution from being dominated by statistical noise. This damping behavior is also found for more general regularization conditions **L**. However, if the regularization is too large the solution will only minimize the regularization condition and will not fit the data. For this reason one must choose an

optimal value of τ to keep both roles for the unfolded solution : damping statistical fluctuations and fit with the data. In the next sections we will discuss more about this.

Regularization Conditions

In the example in the previous section, the matrix $\mathbf{L}$ in equation (5.23) has a particular form which is the identity matrix. In principle, this matrix can have any form but for this purpose, there are typically three kinds of this matrix (as expressed below) which give constraints on the size, derivative (first derivative) or curvature (second derivative) of the solution. These choices of the matrix $\mathbf{L}$ are considered as regularization conditions.

$$\mathbf{L}_{size} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \qquad \mathbf{L}_{derivative} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}$$

$$\mathbf{L}_{curvature} = \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 1 \end{pmatrix}$$

As mentioned, each regularization condition imposes a different constraint on the solution. The size condition constrains the overall solution to not grow large in any bin; the derivative condition constrains that the bin-to-bin differences of the solution to be small; and the curvature condition constrains the solution to be smooth. In practice, people usually examine all three conditions and choose the condition which will give of the smallest expected unfolding bias.

Regularization Parameter Optimization

Once a condition $\mathbf{L}$ has been chosen, the solution as given in equation (5.23) depends on the regularization parameter τ. While there are several algorithms proposed to choose an optimal value of τ, we present here the "L-Curve" algorithm [51]. The aim of this algorithm is to balance the residual against the regularization in the χ^2 function given in equation (5.21). To do this, in particular, we plot the two quantities of the χ^2 against each other, i.e. as a curve. Defining $||\mathbf{A}\mathbf{x}_\tau - \mathbf{y}||_V = (\mathbf{A}\mathbf{x}_\tau - \mathbf{y})^T \mathbf{V}_y^{-1} (\mathbf{A}\mathbf{x}_\tau - \mathbf{y})$ and $||\mathbf{L}\mathbf{x}_\tau|| = (\mathbf{L}\mathbf{x}_\tau)^T (\mathbf{L}\mathbf{x}_\tau)$, we have the curve

$$\Big(\log||\mathbf{A}\mathbf{x}_\tau - \mathbf{y}||_V, \quad \log||\mathbf{L}\mathbf{x}_\tau||\Big) \tag{5.27}$$

parameterized by the regularization parameter τ. A generic L-Curve showing the general features of the curve can be seen in Figure 5.22.

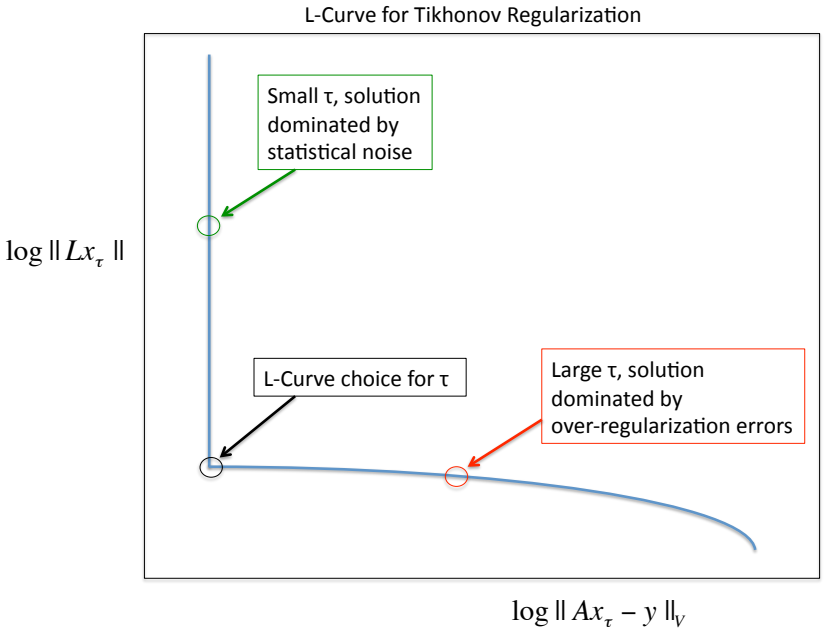


Figure 5.22: The Generic L-Curve.

Figure 5.22 shows that there are essentially three regions of the L-Curve. One is vertical region where decreasing τ, the regularization size $||\mathbf{L}\mathbf{x}_\tau||$ increase while the residual $||\mathbf{A}\mathbf{x}_\tau - \mathbf{y}||_V$ does not change. In this region, the solution is dominated by the statistical noise of the data. Another one is horizontal region, where the residual increases while the regularization size almost does not change when increasing τ.

Here, the solution is over-regulated and will show large differences with the data. In the elbow of the curve, there is an optimal point at which the solution is neither over or under regulated. This region of the L-Curve has the maximum curvature κ of the curve. Thus, in the L-curve algorithm, the point of maximum curvature κ_{max} [51] will correspond to the value of τ_L to be chosen.

Disadvantages of the matrix inversion method

Although the response matrix inversion algorithm based on the regularization method using L-Curve optimization has number of advantages as discussed, we don't choose it for our analysis because of the following reasons:

- The regularization constraints typically penalize large curvature or size changes in neighboring bins in order to damp oscillations. In the case of this analysis, the distribution is rapidly falling, with a quickly rising peak in the first two bins. Many times the χ^2 optimization penalized this expected feature, which resulted in large biases in the final result.
- Both the χ^2 minimization and the L-Curve optimization are very computationally expensive. As the uncertainty propagation framework of this analysis described in Section 5.8 is based heavily upon the use of toy Monte Carlo and pseudo-experiments, the added computational burden of this technique made things considerably more difficult.
- The L-Curve optimization algorithm has difficulties evaluating optimal regularization parameters in the case of smooth L-Curve distributions as observed in this analysis. Because the L-Curves were found to be very smooth, the error on the optimal value of the regularization parameter was large, resulting in a large overall error.
- This method only works in dealing with one dimensional observables.

5.6.4 Bayesian Unfolding

Bayesian unfolding method is our choice to perform the unfolding for this analysis. This method is an iterative unfolding algorithm that uses Bayes theorem. It is another approach of unfolding and has number of advantages with respect to other unfolding methods [38] such as it can be applied to multidimensional variables; it can take into account any kind of smearing and migration from the true values to the observed ones; it does not require matrix inversion and it is implemented in the RooUnfold package, which

makes using it ROOT very simple and fast. With a fast calculation, it makes uncertainty calculations via toy Monte Carlo and pseudo-experiments, as described in Section 5.8, computationally feasible. In addition, as an iterative method, it thus contains only one parameter which is the number of iterations. This parameter is not only easily understood but also easy to search and optimize over it since good solutions occupy a relatively small phase space in this parameter.

Mathematically, in the Bayesian unfolding method, the probability distribution of true events can be determined as follows (given the observed data and the response matrix):

$$P(\mathbf{x} \mid \mathbf{y}, \mathbf{A}, I), \tag{5.28}$$

where, as before, $\mathbf{x}$ is the true distribution, $\mathbf{y}$ is the measured distribution, $\mathbf{A}$ is the response matrix, and the new variable, I, is the underlying assumptions of the analysis, which are usually left implicit.

From Bayes theorem, equation (5.28) can be written as

$$P(\mathbf{x} \mid \mathbf{y}, \mathbf{A}, I) \propto P(\mathbf{y} \mid \mathbf{x}, \mathbf{A}, I) \cdot P(\mathbf{x} \mid I) \tag{5.29}$$

where the first term on the right is the likelihood of the observed data, and the second term is the prior on the underlying truth distribution.

By looking at equation (5.29) one can see that the unfolded value is strongly influenced by the underlying prior distribution. This can cause a problem in the sense that, in practice, a Monte Carlo is usually used to produce a prior, the unfolded value will be biased towards the truth distribution of the Monte Carlo used. On the other hand, this method can meet a similar problem related to the measurement uncertainties due to statistical fluctuations in the finite measured sample as seen in the regularized matrix inversion. To solve these problems, Bayesian unfolding uses an iterative method in which the number of iterations is requireed to be a regularization parameter.

The iterative procedure proceeds as follows. The first iteration uses the supplied Monte Carlo truth distribution as the prior distribution to solve equation (5.29). Then for next iterations, the result from the previous iteration is used as the prior. The more iterations that are run, the closer to the Monte Carlo truth distribution the unfolded distribution becomes. However if the number of iterations performed is too large, the unfolded distribution will converge toward a distribution which shows strong fluctuations around the true one. The reason of this is explained more detail in [38]. Thus, the number of iterations serves to balance the strength of the bias with the size of the oscillations, exactly as the regularization

parameter did in the regularized matrix inversion technique described in Section 5.6.3. In general, the number of iterations taken is small (roughly 10), as otherwise the statistical errors become very large.

5.7 Reconstruction Efficiency Correction

Following the unfolding which does not deal with the acceptance, it is necessary to correct for the effects of varying event selection efficiency as a function of $p_T^{W,\text{true}}$. To do this, signal Monte Carlo is used to calculate the event selection efficiency correction $C_W(p_T^{W,\text{true}})$ in bins of $p_T^{W,\text{true}}$. This efficiency corrects the distribution from the $p_T^{W,\text{true}}$ spectrum of events passing all of our selection at reconstruction level to the $p_T^{W,\text{true}}$ spectrum of events passing the kinematic selection at truth level. The event reconstruction efficiency is corrected for differences between data and simulation in trigger efficiency, reconstruction efficiency, and lepton momentum resolution. Finally, a systematic uncertainty is assessed based upon the uncertainties on the correction factors applied, dependence on the generator used to model the signal, and uncertainties due to lepton p_{T} and η.

5.7.1 $W \to e\nu$ Event Reconstruction Efficiency

The event selection efficiency correction C_W^i is calculated for each true p_T^W bin i using equation (5.30). It is the ratio of the sum of reconstruction weight, namely w_{reco}, over all events passing reconstruction selections as described in Section 5.3.1 to the sum of true weight, namely w_{true}, over all events generated in the fiducial volume of the detector as given in Section 5.3.3. In this analysis, equation (5.30) is performed in bins of $p_T^{W,\text{true}}$ to calculate bin by bin in $p_T^{W,\text{true}}$the selection efficiency. The true weight w_{true} is simply the generator weight which is equal to 1 for PYTHIA generator. The reconstruction weight w_{reco} is defined in equation (5.31) where SF_{reco} is the reconstruction scale factor, the ratio between the electron identification efficiency measured in data and predicted by the Monte Carlo. Here, only the data/MC scale factor of electron identification efficiency is considered since the electron reconstruction and trigger efficiencies can be reliably taken from Monte Carlo. All these are described in details on the Egamma twiki page [46] and documentations quoted there but will be summarized briefly in the following subsections.

$$C_W^i = \frac{\sum\limits_{\text{events, bin } i} w_{reco}}{\sum\limits_{\text{events, bin } i} w_{true}}, \tag{5.30}$$

where

$$w_{true} = w_{gen} \quad \text{and} \quad w_{reco} = w_{true} \times SF_{reco}(\eta^e), \tag{5.31}$$

and the sum in both the numerator and denominator is over all events generated in $p_T^{W,\text{true}}$ bin i.

The resulting central values of selection efficiency in bins of $p_T^{W,\text{true}}$ are shown in top half of Figure 5.23. Here the generator used is PYTHIA. Monte Carlo sample is smeared by applying resolution correction for electron energy to have the same resolution as in data (more details about this is described in Section 5.7.1). Uncertainties on the selection efficiency are also studied as shown in bottom half of Figure 5.23. These uncertainties are due to the following sources:

- Uncertainty on the reconstruction scale factor SF_{reco}: by varying the SF_{reco} a sigma away from its central value, an uncertainty of about 2% on the efficiency correction is observed.
- The difference between with and without smearing MC sample according to electron energy resolution correction.
- Generator: re-weighting PYTHIA to RESBOS via the ratio of the truth distributions. The effects of this on the efficiency factor are very small.
- MC statistical uncertainty: it is assigned from the number of simulated events used to calculate the correction factor. This error becomes significant at higher value of p_T^W where there are fewer simulated events.

PDF uncertainties on the efficiency correction are not explicitly calculated here. In the recent measurement of Z transverse momentum distribution at ATLAS [28], this uncertainty is found to be 0.1% across the spectrum. Due to the large correlation of W and Z production under PDF variations, a similar figure is expected here and can safely be neglected in comparison with other uncertainties.

Those uncertainties are treated as uncorrelated so the total uncertainty on efficiency correction as shown by the black line in Figure 5.23 is a quadratic sum over all individual uncertainties mentioned above.

Reconstruction, trigger and identification efficiencies of electron

In this analysis, electron candidates which are isolated and with high p_{T} (above 20 GeV) are passed robuster Tight cut [48]. The electron reconstruction is thus based on clusters reconstructed in the electromagnetic calorimeter, in association with a matched Inner Detector track. The reconstruction efficiency was studied by using W, Z and J/ψ tag and probe method. Results from those studies which are mentioned here [46] indicate that the reconstruction efficiency in Monte Carlo and in data agree within a

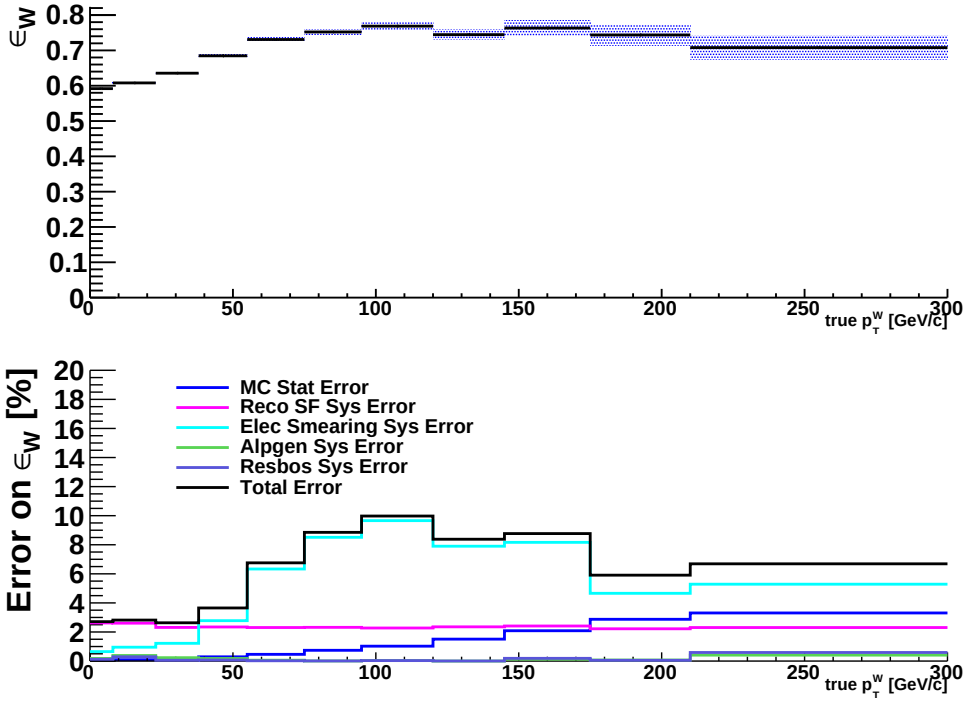


Figure 5.23: Event selection efficiency C_W, and its uncertainties, as a function of $p_T^{W,\text{true}}$in electron channel.

systematic uncertainty of 1.5% so no data/MC scale factor according to this efficiency measurement needs to be applied.

The trigger efficiencies on plateau normalized to offline cuts used in this analysis as listed in table 5.6 are above 99% with an associated systematic <1% and the data/MC scale factor is unity within statistical uncertainty. The trigger scale factor is thus not included in equation (5.31) for the weight of w_{reco}.

	Offline Medium	Offline Tight
Data	98.67 ± 0.10	99.03 ± 0.09
MC	99.235 ± 0.006	99.538 ± 0.005
Scale Factor	0.9943 ± 0.0010	0.9950 ± 0.0009

Table 5.6: Electron trigger efficiencies in Monte Carlo and Data and Data/MC scale factor [46].

The identification efficiency was also measured with tag and probe method using W and Z. For this analysis, measurements of efficiencies and Data/MC scale factor are performed in 8 bins of η^e with a range of E_{T}^e between 20 and 50 GeV. Results of scale factors as well as their errors including statistical

and systematic uncertainties for robuster Tight electron are listed in Table 5.7.

Eta bin	$[-2.47,-2.01]$	$[-2.01,-1.52]$	$[-1.37,-0.8]$	$[-0.8,0]$
SF tight	0.925 ± 0.034	0.995 ± 0.024	1.006 ± 0.021	0.982 ± 0.018
Eta bin	$[0,0.8]$	$[0.8,1.37]$	$[1.52,2.01]$	$[2.01,2.47]$
SF tight	0.987 ± 0.018	1.033 ± 0.025	1.028 ± 0.045	0.936 ± 0.034

Table 5.7: Data/MC scale factor of identification efficiency for robuster Tight electron [46].

Electron energy resolution

Studies of the e^+e^- invariant mass distribution [47] in $Z \to ee$ events show that the Monte Carlo does not model correctly the electron energy resolution (see Figure 5.24). So a smearing procedure is applied to Monte Carlo to correct for this discrepancy. In particular, the resolution of electron energy in Monte Carlo is corrected by smearing event by event the electron energy as described in formula (5.32) and (5.33). There, E and E_{corr} are respectively the electron energy before and after the correction; S and C are sampling and constant terms and ΔS and ΔC are their systematic uncertainties. The sampling and constant terms are measured in data by using formula (5.34) with an upper limit on the constant term as given by equation (5.35). Their values and uncertainties are listed in Table 5.8.

$$E_{corr} = E + gRandom \to Gaus(0, \Delta\sigma), \tag{5.32}$$

with

$$\Delta\sigma = \sqrt{\left(S.(1+\Delta S).\sqrt{E}\right)^2 + (C.(1+\Delta C).E)^2 - \left(S.\sqrt{E}\right)^2 - (C.E)^2}, \tag{5.33}$$

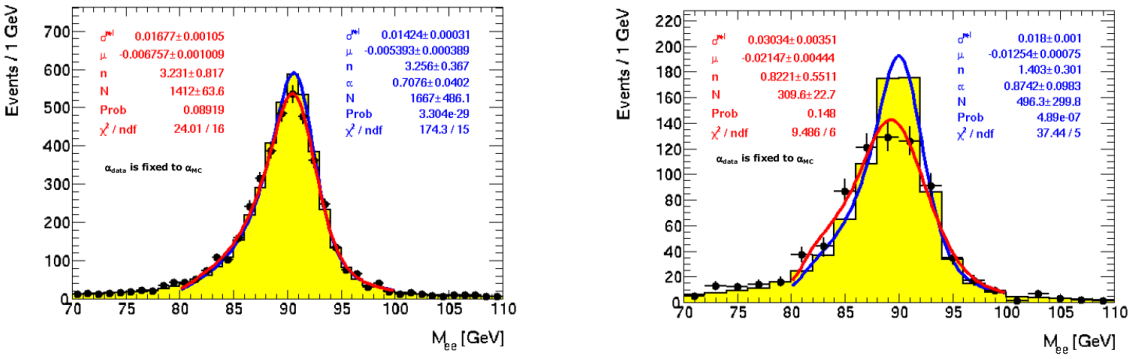
Eta bin	$]0.0,1.4[$	$[1.4,2.5[$
S	0.1 ± 0.2	0.1 ± 0.2
C	0.007 ± 1.0	0.007 ± 4.0

Table 5.8: Sampling and constant terms and their systematic uncertainties in Barrel and Endcap.

After the smearing correction, the e^+e^- invariant mass distributions in $Z \to ee$ Monte Carlo are in a very good agreement with those in data as shown in Figure 5.25.

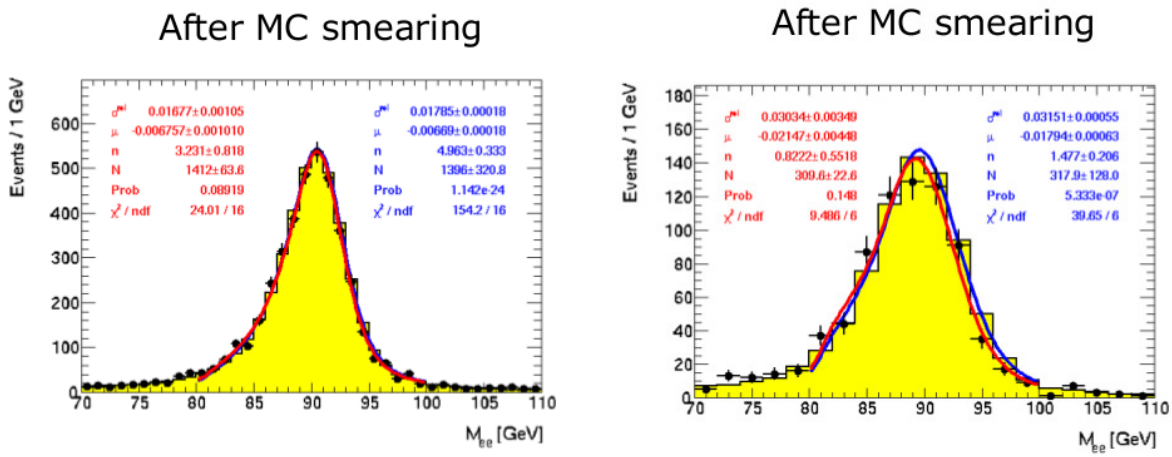
$$\frac{\sigma}{E} = \frac{S}{\sqrt{E}} \oplus C, \tag{5.34}$$

$$C \leq \sqrt{2}.\sqrt{\left(\frac{\sigma^{data}}{E}\right)^2 - \left(\frac{\sigma^{MC}}{E}\right)^2}. \tag{5.35}$$



(a) Barrel (b) End-cap

Figure 5.24: e^+e^- invariant mass distributions in $Z \to ee$ data (the red curve) and Monte Carlo (the blue curve) before the smearing correction.



(a) Barrel (b) End-cap

Figure 5.25: e^+e^- invariant mass distributions in $Z \to ee$ data (the red curve) and $Z \to ee$ Monte Carlo (the blue curve) after the smearing correction.

5.8 Calculation of Uncertainties

Procedures of uncertainties calculation presented here are the same as described in [29]. In this measurement, unfolding and normalization introduce non-negligible correlations between bins in the final result. Therefore, the uncertainty propagation framework in this analysis is based around the production and propagation of covariance matrices. In particular, the uncertainties on the background estimation described in section 5.4 are converted to covariance matrices after unfolding using pseudo-experiments. The uncertainties on the response matrix and those due to bias of the unfolding algorithm are also reported as covariance matrices and combined with the background covariance matrices. Finally, the uncertainties on the efficiency correction described in section 5.7 are combined, and the final, normalized result is presented with a covariance matrix for both the statistical and systematic uncertainties. The statistical uncertainty of the result is also determined using pseudo-experiments, as the unfolding introduces non-trivial statistical effects. This section describes the details of the calculation of the covariance matrices, which are listed in Table 5.9, and their propagation to the final answer.

Source	Covariance Name	Evaluation Method
QCD Bkg Subtraction	V_{QCD}	Pseudoexperiments
WZ Bkg Subtraction	V_{WZ}	Pseudoexperiments
top Bkg Subtraction	V_{top}	Pseudoexperiments
SumET reweighting	V_{sumet}	On/Off comparison
Lepton resolution and scale	$V_{smearing}$	On/Off comparison
Generator dependence	V_{truth}	Comparison with RESBOS
Fit parameterization	V_{fit}	Comparison with Generated
Fit statistics	$V_{fitstats}$	Pseudoexperiments
Unfolding Bias	V_{bias}	Closure Test
Unfolding Statistical Errors	$V_{stat}^{unfolding}$	Pseudoexperiments
Efficiency correction	V_{eff}	Analytic Propagation

Table 5.9: Sources of uncertainty.

5.8.1 Statistical Formulation

A number of uncertainty calculations in this analysis use pseudo-experiments, in which either the response matrix or the reconstructed signal Monte Carlo is fluctuated and a covariance matrix is produced. The statistical formulation of this process is described here. Consider a histogram, $\mathbf{x}$, with n bins, whose elements are $\mathbf{x} = x_1, \ldots, x_n$. For each pseudo-experiment, a new histogram, $\mathbf{x^j}$, with elements $\mathbf{x^j} = x_1^j, \ldots, x_n^j$ is measured. Three quantities are built from this set of p measurements.

First, the average histogram, $\mathbf{E}$ is defined by

$$E_i = \frac{1}{p} \sum_j x_i^j \tag{5.36}$$

Next, the covariance matrix $\mathbf{V}$ is defined by

$$V_{ij} = \frac{1}{(p-1)} \sum_k \left[\left(x_i^k - E_i \right) \left(x_j^k - E_j \right) \right] \tag{5.37}$$

The covariance matrix has a number of nice properties. In particular, the diagonal elements are the square of the standard deviation of that bin. Ignoring correlations between bins, then, one may plot the square root of the diagonals to represent the uncertainty. This also means that for uncorrelated uncertainties, summing in quadrature is equivalent to adding covariance matrices. However, because the covariance matrices include the absolute variance of each bin, and there are large variations in the number of entries per bin in this analysis, it is difficult to visual uncertainties using covariance matrices. For example, if the first bin has 10,000 events with an uncertainty of 1%, while the last bin has 10 entries with an uncertainty of 10%, the diagonal element of the covariance matrix for the first bin will be roughly 10,000 times larger than that for the last bin.

To allow easier visualization, a third quantity, the correlation matrix, ρ, is also calculated:

$$\rho_{ij} = \frac{V_{ij}}{\sqrt{V_{ii} V_{jj}}} \tag{5.38}$$

The correlation matrix has the nice property that the diagonals are, by construction, 1, and the off diagonals are between -1 and 1. Entries which have correlations near 1 tend to vary up and down together, while those with correlations near -1 tend to vary oppositely, and entries with correlations near 0 tend to be independent.

5.8.2 Background subtraction uncertainties

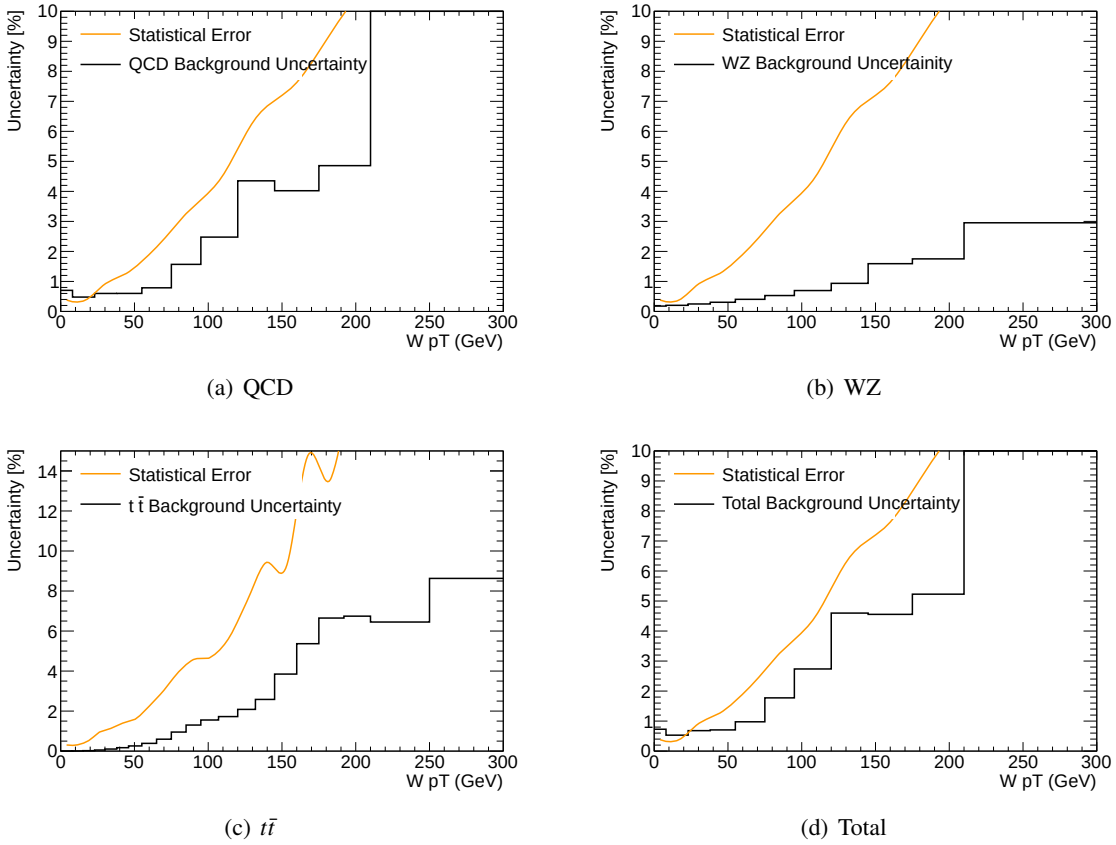
The background uncertainties described in section 5.4 are converted to a final covariance matrix in a two step process. First, the background estimations are varied within their uncertainties over 1000 trials. Next, the data is background subtracted using the estimated background for each trial, and unfolded. The resulting covariance matrix is used as the uncertainty for the background subtraction.

The backgrounds are broken into three groups: W/Z style backgrounds ($W \to \tau\nu$, $Z \to \ell\ell$, and $Z \to \tau\tau$), $t\bar{t}$ background, and the QCD background. The uncertainties are broken into two types: shape errors, which are considered 100% uncorrelated bin-to-bin, and scale errors, which are considered 100% correlated bin-to-bin. For the electroweak backgrounds, the shape errors are taken to be the statistical error, the PDF error, and the scale factor errors, while the scale error is taken to be the cross section error. For the QCD, the total error is taken as a shape error.

For each pseudo-experimental trial of the background variations, the background estimate is produced as follows. First, a luminosity is picked according to a gaussian distribution centered around the nominal value with a width of 3.4%. The W/Z and $t\bar{t}$ backgrounds are scaled according to this luminosity, while the QCD is left unchanged. Next, for the W/Z and $t\bar{t}$ backgrounds, a scale value is picked from a gaussian distribution centered around one, with a width given by the scale error, and the background group is scaled by that amount. Note that this technique accounts for the fact that the cross section uncertainties on the W/Z style backgrounds are all correlated, while the cross section uncertainty on the $t\bar{t}$ is uncorrelated with those results. Finally, for each background group, the estimate is changed bin by bin according to the shape error. In order to allow visualization of the uncertainty due to each source, each trial actually produces four background estimates: one where only the W/Z backgrounds are varied, one where only the $t\bar{t}$ background is varied, one where only the QCD is varied, and one where all three are varied simultaneously. As there are small correlations between the samples (due mainly to the luminosity appearing in both WZ and $t\bar{t}$), only the final estimate, where all three groups are varied at once and labeled V_{BKG}, is used in the final analysis.

The pseudo-experiment variations of the backgrounds then are translated into covariance matrices. For each trial, the estimated background for that variation is subtracted from the observed data, and the result is unfolded using the nominal response matrix. The uncertainty covariance matrices for the background subtraction, which are termed V_{WZ}, $V_{t\bar{t}}$, and V_{QCD}, are calculated over the pseudo-experiments according to equation (5.37). The resulting diagonals of the covariance matrix, for each background var-

ied separately, along with all three varied at once, is shown in Figure 5.26, while the correlation matrices are shown in Figure 5.27 .



(a) QCD (b) WZ

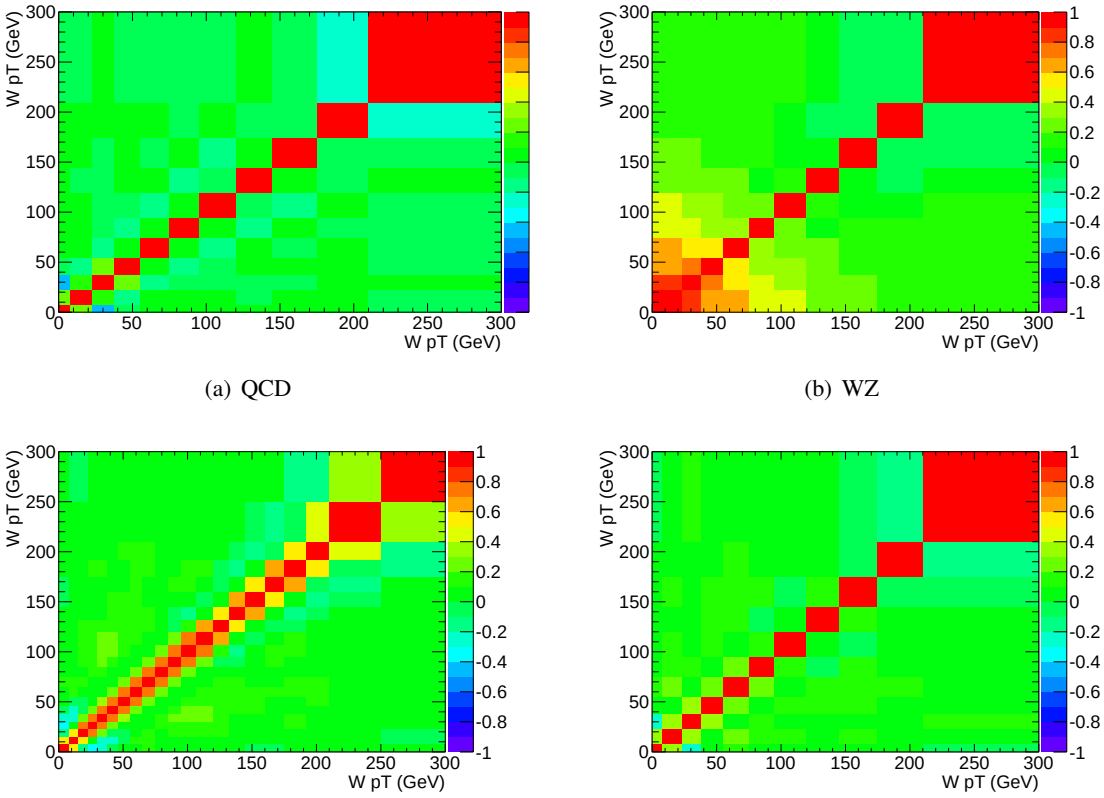
(c) $t\bar{t}$ (d) Total

Figure 5.26: Diagonal elements of background uncertainty covariance matrices for the electron channel broken down by sample, and for the total background.

5.8.3 Statistical uncertainty and unfolding bias

Two sources of unfolding uncertainty are considered in this analysis. First, the unfolding algorithm tends to enlarge the statistical errors on the sample, as well as produce correlations between bins. This statistical uncertainty is calculated by poisson fluctuating the Monte Carlo 100 times and comparing with the nominal truth distribution. Second, there is a small but non-negligible systematic uncertainty due to inherent bias in the unfolding algorithm. This is calculated by folding and then unfolding the same sample with the nominal response matrix.

To calculate the statistical uncertainty following unfolding, which is used as the final statistical uncertainty for the result, pseudo-experiments involving poisson fluctuating the Monte Carlo is used. For



(a) QCD (b) WZ

(c) $t\bar{t}$ (d) Total

Figure 5.27: Correlation matrices for the background uncertainty of the electron channel broken down by sample, and for the total background. In all cases, the colors range from -1 to 1.

each trial, the reconstructed signal Monte Carlo is scaled to the number of events observed in data after background subtraction, and each bin value is sampled according to a poisson distribution with parameter equal to the number of entries in the bin after scaling. The fluctuated reconstruction level Monte Carlo is then unfolded using the nominal response matrix, and the answer is compared with the (unfluctuated) Monte Carlo truth distribution for selected events. The resulting difference is used to fill the covariance matrix for the statistical uncertainty after unfolding, termed, $V_{stat}^{unfolding}$.

The other source of uncertainty in the unfolding process is the bias of the algorithm. This is understood by dividing the Monte Carlo into two samples: a sample for building the response matrix, and a sample for evaluating the systematic. The reconstructed Monte Carlo distribution of the second sample is then unfolded using a response matrix built from filling the truth and reconstruction values from the first sample. The result is then compared with the truth distribution in the second sample for selected events, and the full bin-by-bin difference is taken as the systematic uncertainties. A diagonal covariance

matrix, V_{bias}, is built using these uncertainties. The diagonal elements of the unfolding statistical and systematic covariance matrices are shown in Fig. 5.28, and the correlation matrices for the unfolding statistical uncertainty are shown in Fig. 5.29.

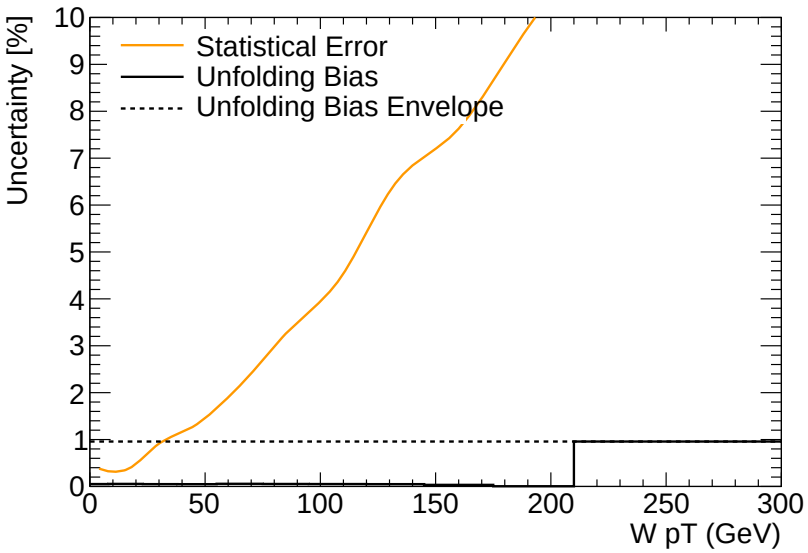


Figure 5.28: Diagonal elements of the covariance matrix of the electron channel for systematic and statistical uncertainties due to unfolding.

5.8.4 Response matrix uncertainties

There are two types of uncertainties on the response matrix determination which are considered here. The first type is due to fluctuations in the formulation of the response due to the statistics of the data and Monte Carlo samples. Covariance matrices for these types of uncertainties are calculated by sampling from the fitting functions described in Section 5.5, producing many copies of the response matrix varied within error, and then unfolding many times. This process is repeated for 1000 trials. The second type is due to uncertainties in the formulation of the response matrix from the reweighting procedure described in Section 5.5. Covariance matrices for these types are produced by folding the sample with the nominal response matrix, and then unfolding with the response matrix under study. As each contribution is independent, the resulting covariance matrix for the uncertainty of the response matrix is taken as the sum of the covariance matrices produced in these studies.

Fluctuations in the fitted results of the response matrix formulation are propagated to the systematic covariance matrix by performing pseudo-experiments in which the fit parameters are varied within their

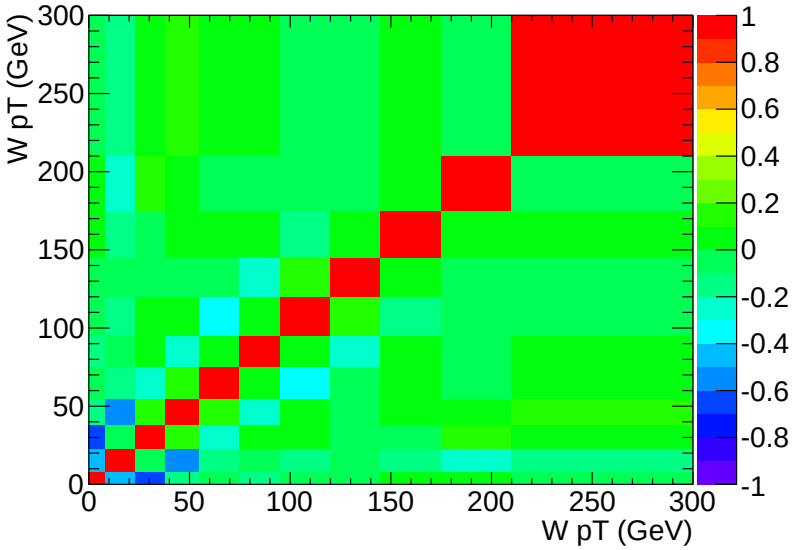


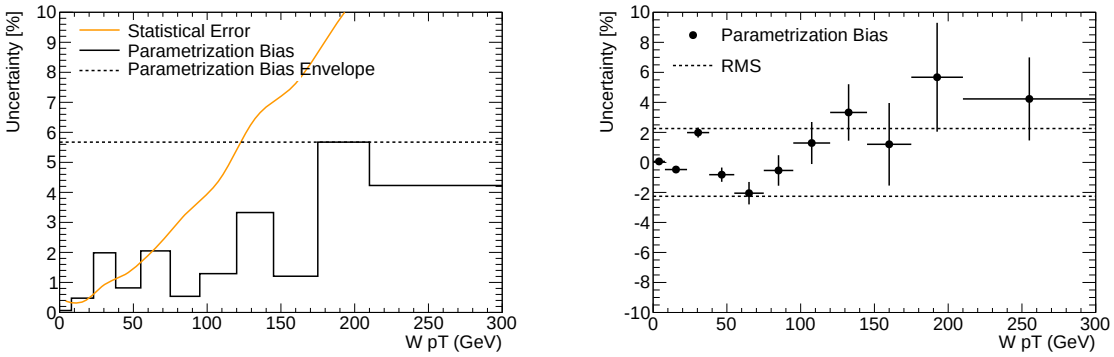
Figure 5.29: Correlation matrix for the unfolding statistical uncertainty for the electron channel.

allowed errors. First, the nominal fits are performed on the Z data, the W Monte Carlo, and the Z Monte Carlo, and the parameter errors and covariances are recorded. For each pseudo-experiment, the fit parameters for all three data samples are sampled from the fit errors using Cholesky decomposition (see [1] for the ROOT implementation used in this analysis) in order to correctly account for the correlations between the fit parameters. The filling of the response matrix described in Section 5.5 is then repeated with the sampled fit parameters, and a new response matrix is produced. The background subtracted data is then unfolded with each produced response matrix, and the systematic covariance matrix, called $V_{fitstat}$ is produced from the trials.

Systematic uncertainties for the reweighting and fitting procedure described in Section 5.5 are produced by comparing the results of unfolding signal Monte Carlo with the the various weightings removed with the nominal unfolding. Four types of systematics are considered. PYTHIA is used to form the nominal response matrix, so it is necessary to determine the effects of the underlying truth distribution used to sample the fitted response matrix. To determine this, the signal Monte Carlo is unfolded with the data driven response matrix sampled using PYTHIA Monte Carlo and another sampled using RESBOS Monte Carlo. Similarly, the effects of sumet reweighting and lepton resolution smearing and scaling are probed by comparing the results of unfolding the signal Monte Carlo with sumet reweighting and lepton momentum smearing turned off with the results of unfolding with the nominal response matrix. Finally, the effects of the fitting procedure itself are considered by comparing the results of unfolding with a

response matrix built by fitting the signal Monte Carlo with a response matrix built by simply filling the matrix with the truth and reconstructed values of the Monte Carlo. In all cases, the maximum bin-by-bin difference, shown as a dashed line, is taken as the systematic, and a diagonal covariance matrix is built using those values. This results in four systematic covariance matrices, termed V_{sumet}, $V_{smearing}$, V_{truth}, and V_{fit}.

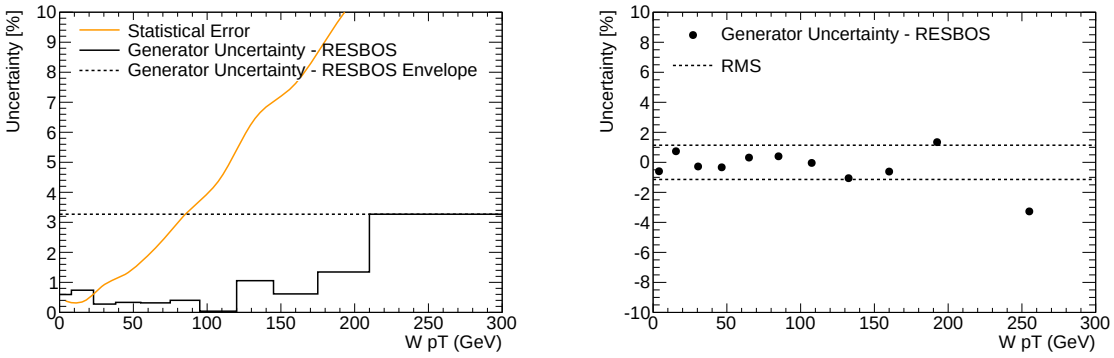
The diagonal elements of the reweighting and fitting covariance matrices are shown in Figures 5.30-5.33 The correlation matrices for the fit paramaterization uncertainty and the fit statistical uncertainty are shown in Figures 5.34.



(a) Bias from a single exercise

(b) Averaged bias from 100 pseudo-experiments

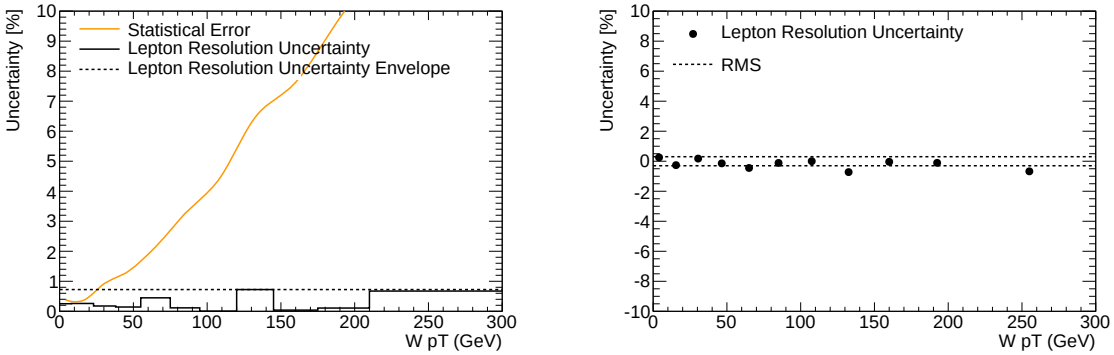
Figure 5.30: Parametrization bias in the electron channel, obtained from folding with the MC response matrix, and unfolding with the parametrized response matrix.



(a) Bias from a single exercise

(b) Average bias from 100 pseudo-experiments

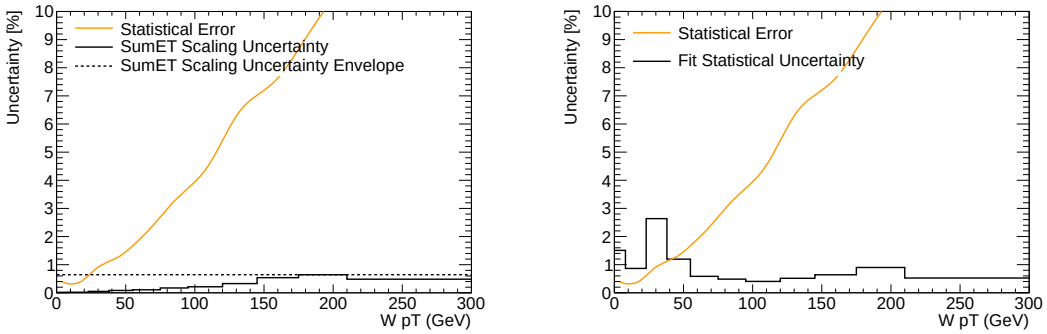
Figure 5.31: Generator bias in the electron channel, obtained from folding with the PYTHIA response matrix, and unfolding with the RESBOS response matrix.



(a) Bias from a single exercise

(b) Averaged bias from 100 pseudo-experiments

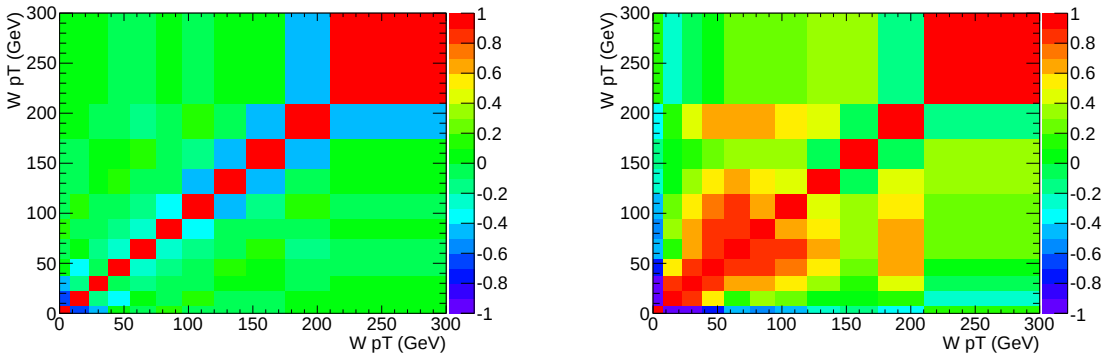
Figure 5.32: Systematic uncertainty due to the electron scale and resolution uncertainty, entering *via* the Z-based data driven corrections.



(a) SumET reweighting

(b) Parametrization

Figure 5.33: Uncertainty due to the SumET reweighting precision and to the parametrization precision, in the electron channel. The uncertainty is obtained from 200 random configurations, and the given uncertainty is the resulting spread on the unfolded result.



(a) Fit parametrization uncertainty

(b) Fit statistical uncertainty

Figure 5.34: Correlation matrix for the fit paramaterization uncertainty and the fit statistical uncertainty for the electron channel.

5.8.5 Efficiency Calculation

The efficiency correction uncertainties are typically small compared with other uncertainties, and the efficiency corrections themselves are relatively flat. Thus, we consider uncertainties on the efficiency to be bin-by-bin uncorrelated, and there use a diagonal covariance matrix to represent its uncertainties.

5.8.6 Propagation and Normalization

In general, the propagation of uncertainties using covariance matrices is accomplished via simple addition if the uncertainties are uncorrelated among each other. However, for more complicated operations like unfolding and efficiency correction, the propagation of the matrices becomes more complicated. The precise steps used to propagate the statistical and systematic uncertainties are described here.

The observed data is initially assigned a systematic covariance matrix of all zeroes, and a diagonal statistical covariance matrix with entries equal to the number of entries in each bin of the data. After background subtraction, the statistical covariance matrix is unchanged, while the systematic covariance matrix is taken to be a diagonal covariance matrix with the simple errors described in Section 5.4. These uncertainties are used only for visualizing the results at this stage, and then are removed for the unfolding step. Following unfolding, the covariance matrices described in the proceeding section are used. The statistical covariance matrix of the data is taken from the results of the process described in section 5.8.3, while the systematic covariance matrix is taken to be the sum of the covariance matrices described in Sections 5.8.2, 5.8.4, and 5.8.3.

Mathematically, this may be written as:

$$V_{stat} = V_{stat}^{unfolding}$$

where V_{stat} is the statistical covariance matrix of the data, and $V_{stat}^{unfolding}$ is the statistical covariance due to unfolding that is determined using pseudo-experiments as described in Sec. 5.8.3 and

$$V_{syst} = V_{BKG} + V_{bias} + V_{sumet} + V_{smearing} + V_{truth} + V_{fit} + V_{fitstat}$$

where V_{syst} is the systematic covariance matrix of the data, V_{QCD}, $V_{W/Z}$, and $V_{t\bar{t}}$ are the covariance matrices of background uncertainties, V_{bias} is the covariance matrix of the unfolding bias, and V_{sumet}, $V_{smearing}$, V_{truth}, V_{fit}, $V_{fitstat}$ are the various response matrix covariance matrices.

The efficiency correction step requires a division, which results in an amplification of the covariance matrix. To simplify this calculation, as discussed, the covariance matrix for the efficiency correction is assumed to be diagonal. The statistical covariance matrix is updated as follow:

$$V_{stat}^{i,j} \to \frac{V_{stat}^{i,j}}{\varepsilon_i \varepsilon_j}$$

where $V_{stat}^{i,j}$ is the ij-th entry of the data statistical covariance matrix and ε_i is the value of the efficiency correction in the i-th bin. Similarly, the systematic covariance matrix is updated by:

$$V_{syst}^{i,j} \to \frac{V_{syst}^{i,j}}{\varepsilon_i \varepsilon_j} + \left(\frac{\sigma_i}{\varepsilon_i}\right)^2 \left(\frac{x_i}{\varepsilon_i}\right)^2$$

where x_i is the unfolded measurement (before efficiency correction) in the i-th bin, and σ_i is the uncertainty on the i-th bin of the efficiency correction. The first term is the original covariance matrix rescaled for the efficiency correction so that it stays in proportion with the central value. The second term is the added covariance matrix for the contribution of the uncertainties on the efficiency correction. It is diagonal because there are no correlations between the bins for the efficiency correction.

Following efficiency correction, the last step is normalization. Typically, scaling by a number simply multiplies the covariance matrix by that number. However, in the case of normalization, the operation itself introduces correlations between bins. For example, in a two binned histogram, fluctuations up in the first bin result in smaller results for the second bin after normalization. To account for this effect, the following normalization procedure is used in this analysis:

Consider a histogram with bins $x_1, \ldots, x_n$, and a covariance matrix $\mathbf{V}$. The transformation $\mathbf{x} \to \mathbf{y}$ transforms the covariance matrix by

$$\mathbf{V} \to \mathbf{A}\mathbf{V}\mathbf{A}^{\mathbf{T}} \tag{5.39}$$

where $A_{ij} = \left[\frac{\delta y_i}{\delta x_j}\right]$. Let $N = \sum_i x_i$ be the total number of events and $N_i = N - x_i$ be the sum of events not in the i-th bin. Then, the normalization transformation is

$$x_i \to \frac{x_i}{N_i + x_i}$$

which yields

$$A_{ii} = \frac{N_i}{N^2}$$

$$A_{ij} = \frac{-x_i}{N^2}(i \neq j)$$

Both the final statistical and systematic covariance matrices are transformed according to equation 5.39 using the above values.

5.9 Results

Following event selection, described in Sections 5.3.1, the resulting p_T^W distribution is background subtracted using the background estimation described in Sections 5.4.1 and 5.4.2. The results of this are shown in Figure 5.35 which is compared with reconstruction level PYTHIA signal Monte Carlo. The associated uncertainties shown here are described in Section 5.8.2.

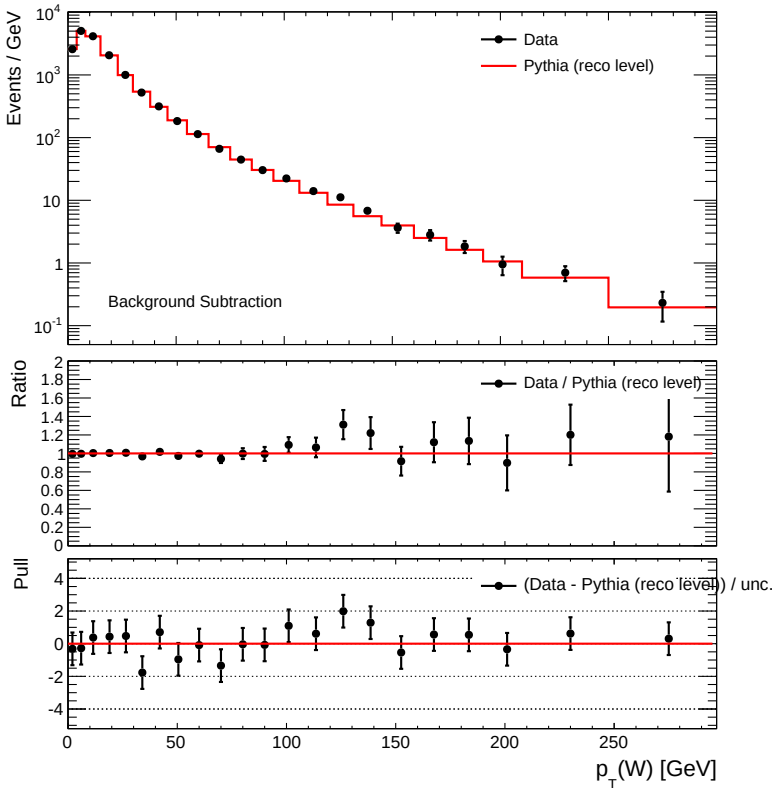


Figure 5.35: $W \to e\nu$ sample after background subtraction compared with reconstruction level PYTHIA signal Monte Carlo.

The observed spectrum is then unfolded using iterative Bayesian unfolding, described in section 5.6.4 using 3 iterations, and the channel combined data driven response matrix described in section 5.5. The uncertainties were calculated according to the methods described in section 5.8. The results of the unfolding, compared with the truth level distribution of events passing the event selection from PYTHIA

Monte Carlo are shown in Figure 5.36. The associated statistical and systematic correlation matrices are shown in Figure 5.37.

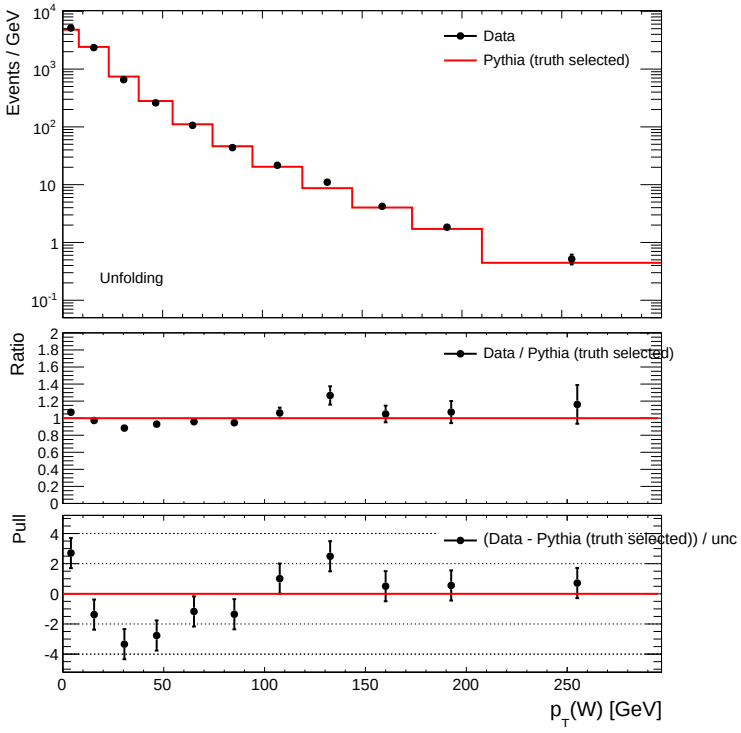
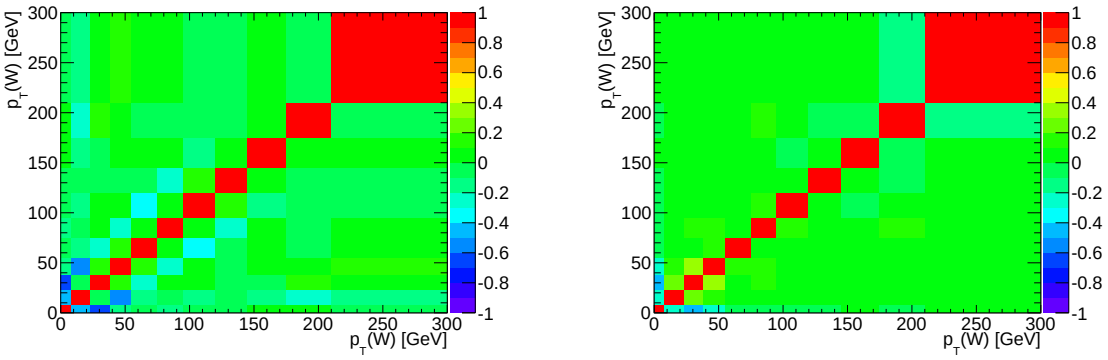


Figure 5.36: $W \to e\nu$ sample after Bayesian unfolding compared with truth level PYTHIA signal Monte Carlo from events passing event selection.

The unfolded spectrum is then corrected to the fiducial volume via a Monte Carlo based efficiency correction described in Section 5.7. The results of this correction are shown in Figure 5.38 and is compared with the truth level distribution of events in the PYTHIA signal Monte Carlo that fall within the fiducial volume defined in Section 5.3.3. The statistical and systematic correlation matrices are shown in Figure 5.39.

The final step is to normalize the distribution. This is done according to the procedure described in Section 5.8.6. The results of the normalization are shown in Figure 5.40. The final statistical and systematic correlation matrices are shown in Figure 5.41.



(a) Statistical Correlation Matrix

(b) Systematic Correlation Matrix

Figure 5.37: Statistical and systematic correlation matrices following bayesian unfolding.

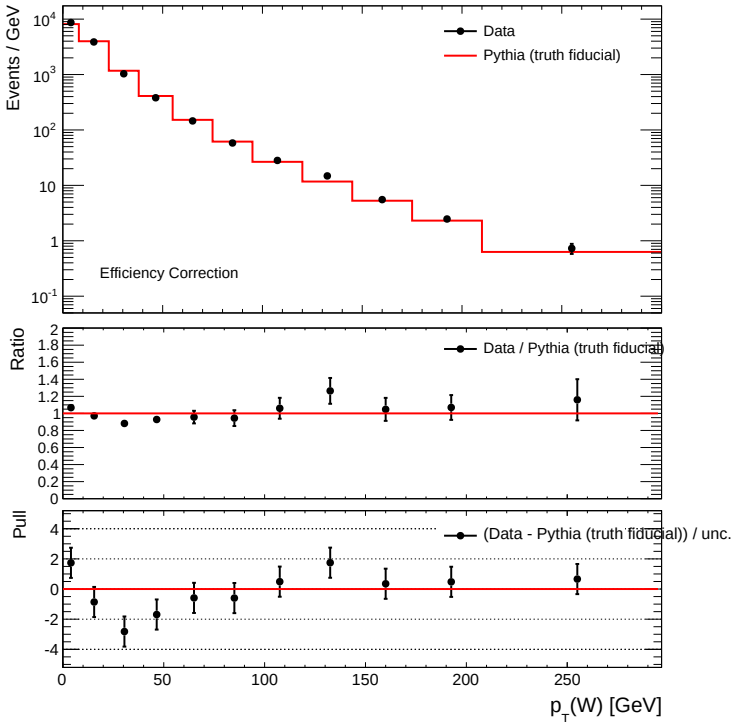
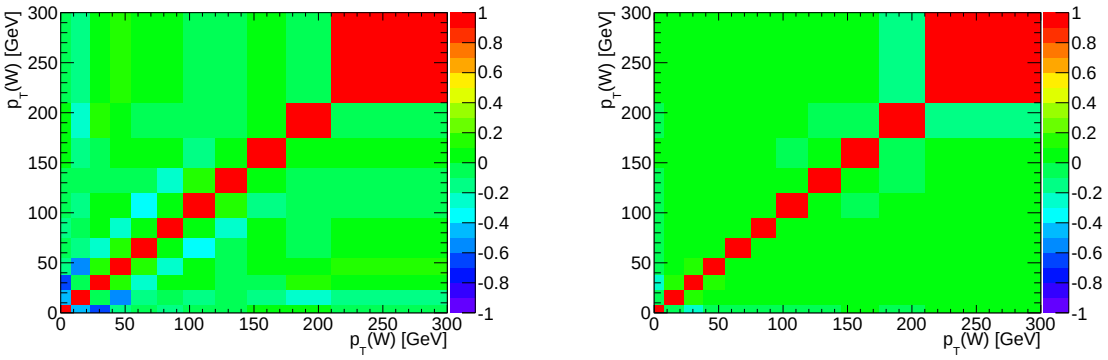


Figure 5.38: $W \to e\nu$ sample after efficiency correction compared with truth level PYTHIA signal Monte Carlo from events within the fiducial volume.



(a) Statistical Correlation Matrix (b) Systematic Correlation Matrix

Figure 5.39: Statistical and systematic correlation matrices following efficiency correction.

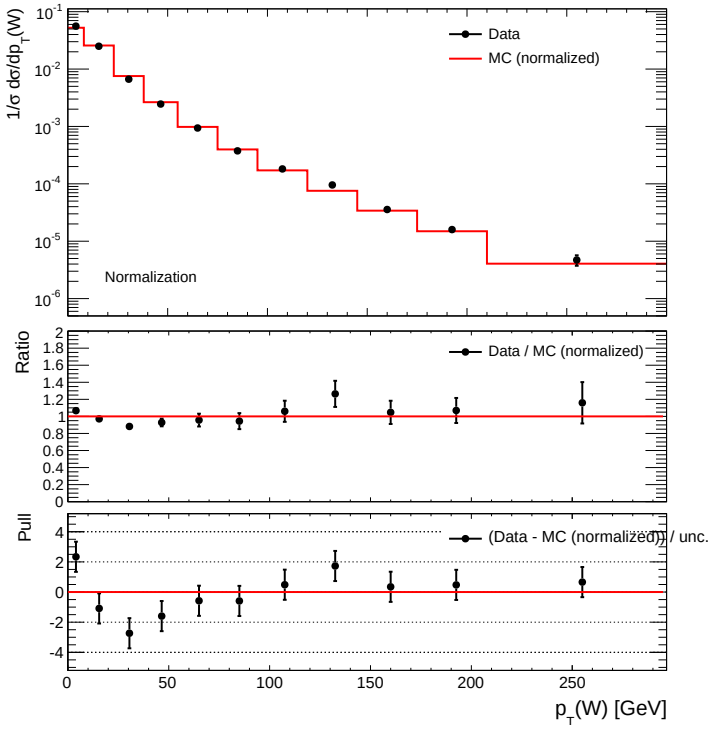
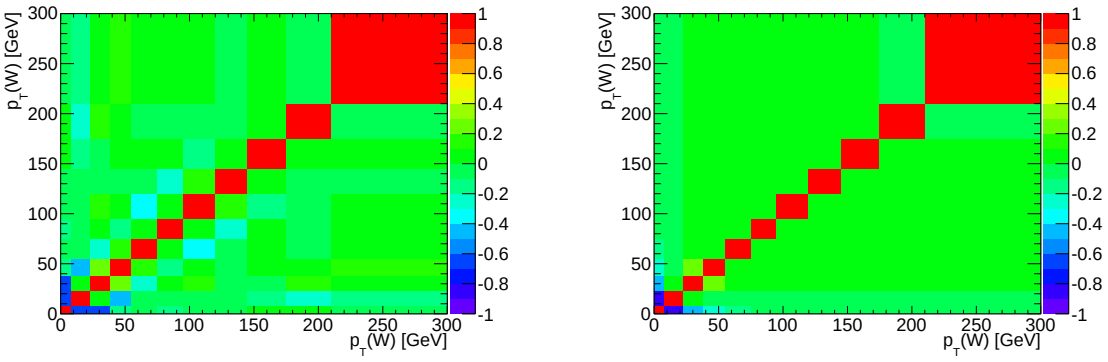


Figure 5.40: $W \to e\nu$ sample after normalization compared with truth level PYTHIA signal Monte Carlo from events within the fiducial volume.



(a) Statistical Correlation Matrix

(b) Systematic Correlation Matrix

Figure 5.41: Statistical and systematic correlation matrices following normalization.

5.10 Translation to alternate signal definitions

Measuring the p_T^W from leptonic decay (into electron or muon), in its baseline form, the measurement is corrected to the Born level lepton kinematics, and p_T^W corresponds, in the unfolded measurement, to the W boson propagator kinematics. The measured spectrum is a differential fiducial cross section, normalized to unity.

We present here variations around these definitions, and how the measured spectrum is transformed accordingly. In all cases, the analysis modifications are performed at the truth level, correcting to a different reference signal, and lead to corrections factors which are illustrated below.

There are two categories of alternatives. The first category, described in Sections 5.10.1 and 5.10.2, redefines the W boson transverse momentum from final state kinematics, which are affected by channel-dependent QED corrections, rather than from the propagator. This has the advantage of expressing the measurement in terms of physical quantities, rather than a virtual intermediate state. Its disadvantage is that the Born level kinematics are universal, and allow to combine the electron and muon channel measurements.

In the second category, we vary the definition of the phase space to which the differential cross section is corrected. In the baseline analysis, the W signal is considered inclusively, and the fiducial phase space is defined by $p_T(l) > 20$ GeV, $|\eta^l| < 2.4$, $p_T(\nu) > 25GeV$, and $M_T > 40$ GeV. In the variations, we study an alternate definition of the signal, defined by excluding the events that contain a hard, wide-angle photon related to the hard vertex and treating this component as a background; we also study the impact of correcting the cross section to the full leptonic phase space before normalizing the obtained distribution.

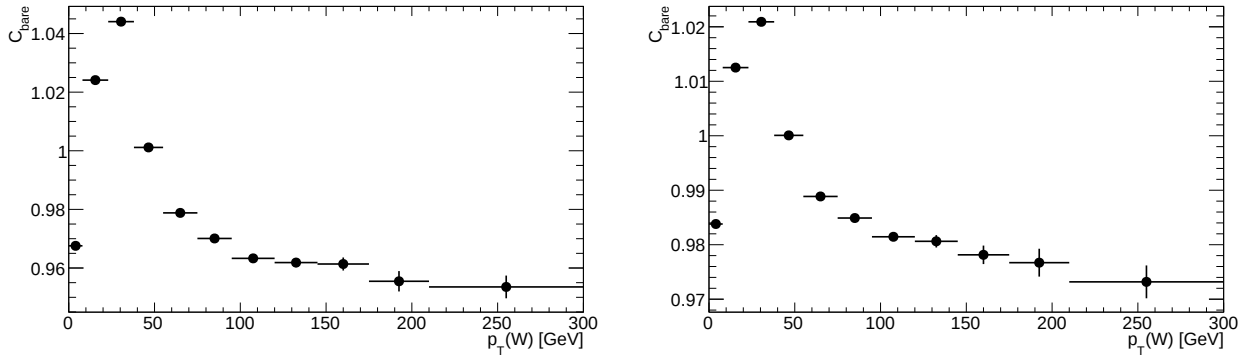
5.10.1 p_T^W from bare charged leptons

In this study, we unfold to the W boson transverse distribution computed from bare leptons (i.e. "after" FSR radiation) :

$$p_T(W) = p_T(l_{\text{Born}}) + p_T(\nu_{\text{Born}}) \;\rightarrow\; p_T(W_{\text{bare}}) = p_T(l_{\text{bare}}) + p_T(\nu_{\text{bare}})$$

Due to the electron-muon mass difference, the radiation pattern differs in both channels. As can be seen from Figure 5.42, this definition results in a spectrum that differs from the baseline by up to 4% in the electron channel, and up to 2% in the muon channel. The deviation is maximal and positive for

$p_T(W)$ 30 GeV, and negative in the low- and high-$p_T(W)$ regions. This can be applied to each channel separately, and results in distributions that are different by nature.



(a) Electron channel (b) Muon channel

Figure 5.42: $p_T(W)$-dependent correction factors converting the baseline measurement into the W boson transverse momentum distribution defined from bare leptons.

5.10.2 $p_T(W)$ from dressed charged leptons

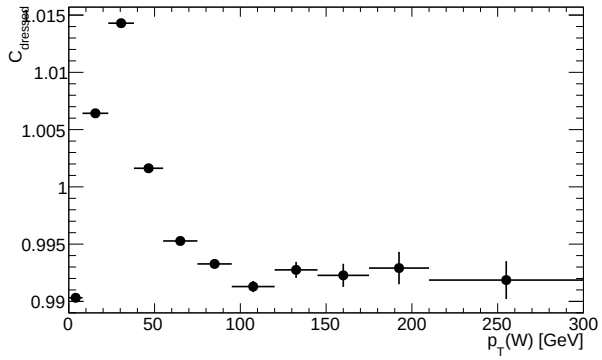
The W boson transverse distribution is now computed from dressed leptons, by which we mean that any photons emitted from the hard vertex, and satisfying $\Delta R(\gamma, l) < 0.2$, are added to the charged lepton momentum:

$$p_T(W) = \rightarrow \; p_T(W_{\text{dressed}}) = p_T(l_{\text{dressed}}) + p_T(\nu_{\text{bare}})$$

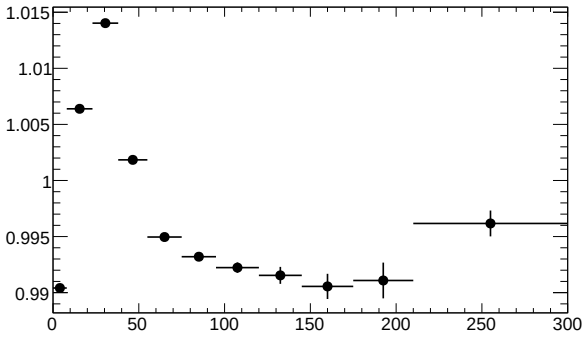
This definition is closest to the experimental definition of $p_T(W)$, described in Chapter 4. The correction factors are illustrated in Figure 5.43. In this case, the correction factors in electron and muon channels agree within the statistical uncertainty of the simulated sample, so they can be applied to the combined measurement.

5.10.3 Excluding $W\gamma$ events from the signal definition

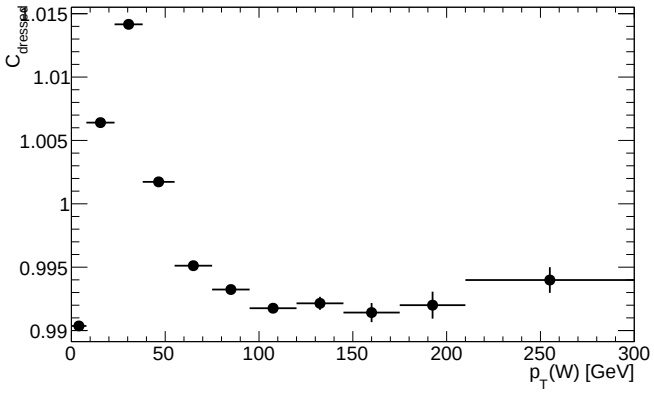
In the baseline measurement, we correct to the truth-level fiducial volume defined in terms of Born kinematics, so that all FSR photons are included in the lepton momentum. However, in the measured p_T^W, hard photons emitted at large angle are not recombined to the charged lepton, and contribute to the reconstruction level $p_T(W)$ distribution. Monte Carlo may then be used to correct for this effect, but this



(a) Electron channel



(b) Muon channel



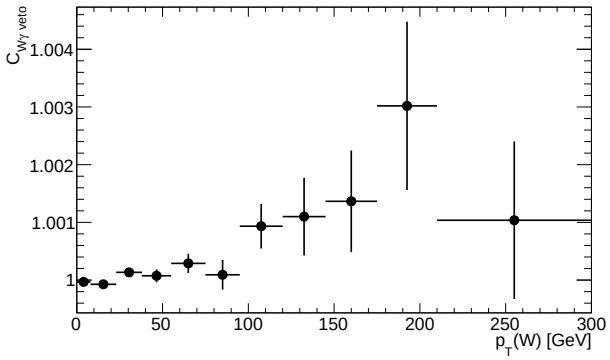
(c) Average

Figure 5.43: $p_T(W)$-dependent correction factors converting the baseline measurement into the W boson transverse momentum distribution defined from dressed charged leptons.

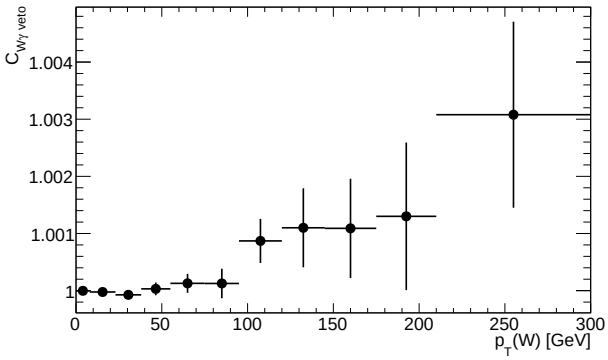
then relies on a proper QED model for this region. The present exercise evaluates the effect of excluding from the signal definition any event satisfying, at the truth level:

$$p_T(\gamma) > 10\,\text{GeV}, \Delta R(\gamma, l) > 0.2$$

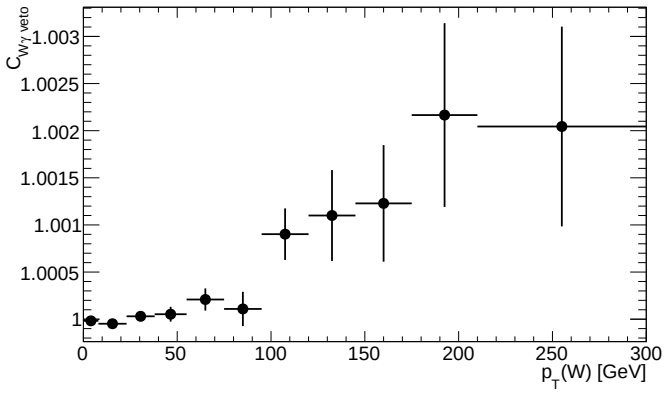
and perform the unfolding to the corresponding reference. Figure 5.44 shows the impact of this change, as the ratio of measurement results as based on this alternate definition and the baseline measurement. After normalization of the distribution, the impact on its shape is smaller than 0.1% for $p_T(W) < 100$ GeV, and elsewhere smaller than 0.3%. As in the previous section, the effect is identical in both channels, the correction factors can be applied on the combined measurement.



(a) Electron channel



(b) Muon channel



(c) Average

Figure 5.44: $p_T(W)$-dependent correction factors converting the baseline measurement into the W boson transverse momentum distribution defined by excluding events containing hard, large angle photons from the signal definition.

5.10.4 Normalized differential 4π cross section

Finally, we present the impact on the measured shape of correcting, bin by bin in p_T^W, to the total lepton phase space instead of the fiducial cross section. The effect is large, up to 20% at $p_T^W \simeq 80$ GeV. The correction factors as a function of p_T^W are shown in Figure 5.45.

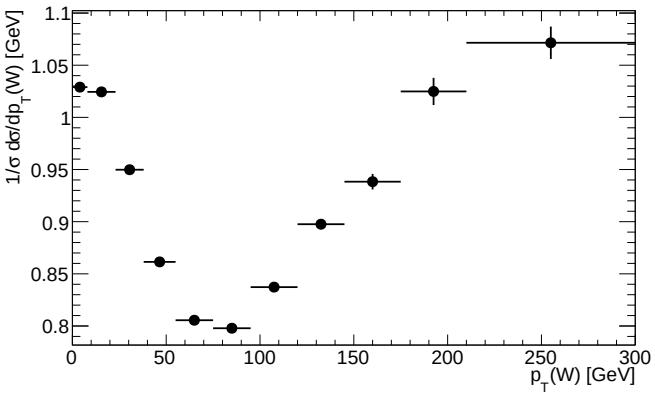


Figure 5.45: $p_T(W)$-dependent correction factors converting the baseline measurement into the W boson transverse momentum distribution defined from correcting, in each bin of $p_T(W)$, to the total lepton phase space.

5.11 Summary

We have described a measurement of the true transverse momentum spectrum of W boson produced by $\sqrt{s} = 7$ TeV proton-proton collisions at the LHC with the ATLAS detector. The data used corresponds to an integrated luminosity of 31.43 pb^{-1} . The work presented in this chapter was for $W \to e\nu$ channel (the $W \to \mu\nu$ channel was only used to study and calibrate the hadronic recoil response). A similar analysis using $W \to \mu\nu$ channel can be proceeded as described in [29]. Both measurements in individual channels can be combined to get more statistic for the final result which will be presented in Chapter 6.

The p_T^W was reconstructed as the hadronic recoil calculated by the hadronic recoil algorithm proposed in Chapter 4. After expected backgrounds from $Z \to ee$, $W \to \tau\nu$, $t\bar{t}$, and QCD multijet events were subtracted, the resulting spectrum was unfolded using a Bayesian unfolding algorithm.

Unfolding requires the definition of a response matrix which relates the true to the reconstructed p_T^W for events passing all of the W selection. For this analysis, the response matrix was constructed by parameterizing the bias and resolution of $p_T^{W,\text{reco}}$ with respect to $p_T^{W,\text{true}}$ as a function of ΣE_T and $p_T^{W,\text{true}}$. This parameterization of the response can then be fitted to the hadronic recoil observed in the Z data, using dilepton transverse momentum $p_{\text{T}}^{\ell^+\ell^-}$ measured from lepton pair as the true p_T^Z. After corrections for differences between Z and W kinematics, most notably the slight difference in the ΣE_T distributions, these fits can be used to model the W response with minimal dependence on the detector simulation.

The unfolded spectrum is the $p_T^{W,\text{true}}$ spectrum of all events passing the event selection. Using simulated signal events to model the detector response, this was corrected back to the $p_T^{W,\text{true}}$ spectrum of all events produced in the fiducial volume of the analysis, that is, all events produced with kinematic features such that they are likely to be reconstructed and pass the event selection.

Uncertainties on the unfolded, corrected spectrum were calculated as covariance matrices to account for bin-to-bin correlations introduced by the unfolding. The statistical uncertainty on the unfolded spectrum was evaluated through pseudo-experiments modeling the propagation through the unfolding procedure of possible statistical fluctuations in the selected data sample. Systematic uncertainties were evaluated for the the background normalization and shape, the modeling of the calorimeter response through the response matrix, the bias introduced by the unfolding procedure itself, and uncertainties on the efficiency correction.

Chapter 6

Channel combination and comparisons to theoretical predictions and p_T^Z measurement

The measurement of p_T^W spectrum performed in electron channel as presented in Chapter 5 can be applied similarly for muon channel. With the result obtained for muon channel as described in [29], we can combine the results for electron and muon channels into a single result to have more statistic. The resulting distribution combined is then compared to current Monte Carlo generators in order to test the theoretical predictions for p_T^W spectrum. In addition, the combined result is compared to the p_T^Z measurement in which the p_T^Z is reconstructed as the transverse momentum of lepton pair in the final state as described in [28].

In this chapter, we present firstly in Section 6.1 a procedure for channels combination. The combined result in comparisons with theoretical predictions and p_T^Z measurement are given in Section 6.2.

6.1 Combination procedure

The electron and muon results are combined at the level of the fiducial cross section. By defining a common fiducial volume for the two channels, the cross sections may be corrected to this common phase space, which allows the results to be combined. The uncertainties on the two channels are sorted according to whether they are correlated between the two channels or not, and a joint covariance matrix

describing the uncertainty on both measurements is constructed. Using this covariance matrix, we define a chisquared between the two measurements and a common underlying distribution. This chisquared is minimized to find the best estimate of the common underlying distribution, which is the combined measurement.

In this section, we first show the general pattern of the solution, and then turn to a detailed discussion of the combination procedure.

6.1.1 χ^2 definition and least squares solution

The χ^2 to be minized is defined as:

$$\chi^2 = (\mathbf{X} - \bar{\mathbf{X}})^{\mathbf{T}} \mathbf{C}^{-\mathbf{1}} (\mathbf{X} - \bar{\mathbf{X}}) \tag{6.1}$$

where X is the joint histogram of measured N-bin distributions in the electron and muon channels, i.e the $2N$-sized vector $\mathbf{X} = \{X_1^e, ..., X_N^e; X_1^\mu, ..., X_N^\mu\}$, $\bar{\mathbf{X}} = \{\bar{X}_1, ..., \bar{X}_N; \bar{X}_1, ..., \bar{X}_N\}$ is the vector of averages to be determined, and $\mathbf{C}^{-\mathbf{1}}$ is the complete, $2N \times 2N$ covariance matrix :

$$\mathbf{C} = \begin{pmatrix} C^e & C^{e\mu} \\ C^{e\mu} & C^\mu \end{pmatrix}. \tag{6.2}$$

The $2N \times 2N$ matrix C is constructed from 4 blocks of $N \times N$ matrices specified by channel-index. Two diagonal blocks C^e and C^μ are the covariance matrices determining uncertainties in individual channels (electron and muon, respectively) in which all sources and calculations of uncertainty considered are described in Section 5.9. Off-diagonal block $C^{e\mu}$ is the covariance matrix reflecting the sources of uncertainty that correlate both channels.

The χ^2 minimization and combined uncertainty calculation is performed in two ways: numerically, using the program `MINUIT` [58], and analytically. The analytical solution, following the discussion in the Statistics section of the PDG [72], is

$$\bar{\mathbf{X}} = (\mathbf{H}^{\mathbf{T}} \mathbf{C}^{-\mathbf{1}} \mathbf{H})^{-\mathbf{1}} \mathbf{H}^{\mathbf{T}} \mathbf{C}^{-\mathbf{1}} \mathbf{X}, \tag{6.3}$$

where $\mathbf{H}$ is a $2N \times N$ matrix specifying the structure of the eqation system:

$$\mathbf{H} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}. \tag{6.4}$$

Finally, the combined covariance matrix is:

$$\bar{\mathbf{C}} = (\mathbf{H^T C^{-1} H})^{-1}. \tag{6.5}$$

Results presented in this chapter are derived using the analytical solution.

6.1.2 Construction of the two-channel covariance matrix

The full covariance matrix of equation (6.2) is constructed as follows. Each source of uncertainty is categorized as being correlated or uncorrelated between the two channels, and correlated or uncorrelated between bins. All of the uncertainties considered for this measurement are assigned to one of these four categories in Table 6.1. For each source of uncertainty, a $2N \times 2N$ covariance matrix is constructed using equation (6.2). Then the resulting set of $2N \times 2N$ matrices is summed.

For a given source of uncertainty, the $2N \times 2N$ covariance given in equation (6.2) is constructed from 4 blocks of $N \times N$ matrices in which two diagonal blocks C^e and C^μ, covariance matrices from electron and muon measurements, respectively, are known (see Section 5.8). The only component which needs to be determined is the off-diagonal block $C^{e\mu}$, covariance matrix representing the uncertainty that correlates between channels. As discussed in Section 5.8 of Chapter 5, the $N \times N$ covariance matrices for uncertainties that are not correlated between bins are diagonal, and those for uncertainties that are correlated between bins contain off-diagonal elements. Thus, similarly, if the source is uncorrelated between the channels, the off-diagonal blocks $C^{e\mu}$ have zero entries, and if the source is correlated between the channels, the off-diagonal blocks $C^{e\mu}$ have nonzero entries. We therefore have two cases of determining the joint covariant matrix corresponding to a given source of uncertainty.

If the source of uncertainty is uncorrelated between the channels, the $2N \times 2N$ covariance matrix contains only the diagonal blocks,

$$\mathbf{C} = \begin{pmatrix} C^e & 0 \\ 0 & C^\mu \end{pmatrix}. \tag{6.6}$$

If the source of uncertainty is correlated between the channels, the diagonal blocks are still filled by copying the covariance matrices from the individual channels. The off-diagonal blocks are not zero and filled using the assumption that if the channels are correlated, they are 100% correlated. This is a reasonable assumption since in this analysis, as mentioned in Section 5.8, the sources of uncertainty that are correlated between the channels are the response matrix uncertainties and some of the background uncertainties, such as the luminosity and cross sections used to normalize the electroweak and top backgrounds. If a particular uncertainty is 100% correlated between the channels, then the correlations between bins, as represented by the correlation matrix for that uncertainty, is identical. This can be seen, for example, in the correlation matrices for the fit parameterization uncertainty and the fit statistical uncertainty shown in Figures 6.1 and 6.2 for the electron and muon channels respectively. Although the correlation matrices are the same, the overall normalization of the two covariance matrices may be different.

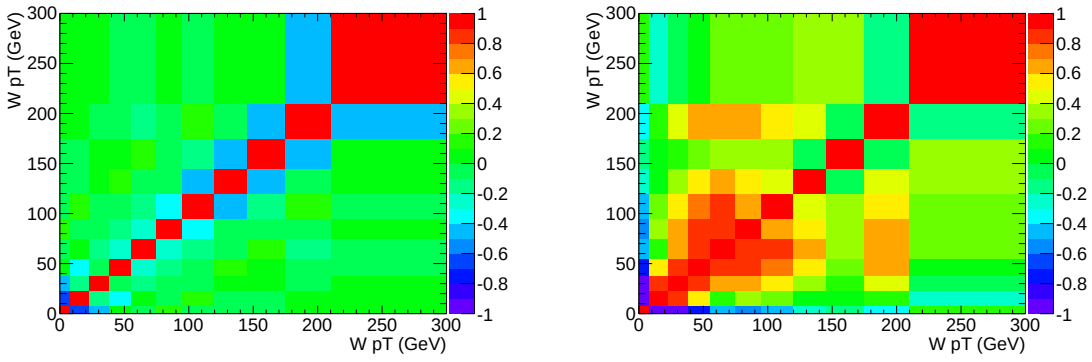


Figure 6.1: Correlation matrix for the fit parameterization uncertainty (left) and the fit statistical uncertainty (right) for the electron channel.

Therefore, to calculate the off-diagonal blocks of the joint $2N \times 2N$ covariance matrix, we factorize out the correlation matrix from the overall magnitude of the covariance. For example, for one covariance matrix, one can write (as in equation (5.38)):

$$V_{ij} = \sqrt{V_{ii}V_{jj}} \times \rho_{ij}\ . \tag{6.7}$$

For the ideal case where we in fact have identical correlation matrices for the two channels, we can then

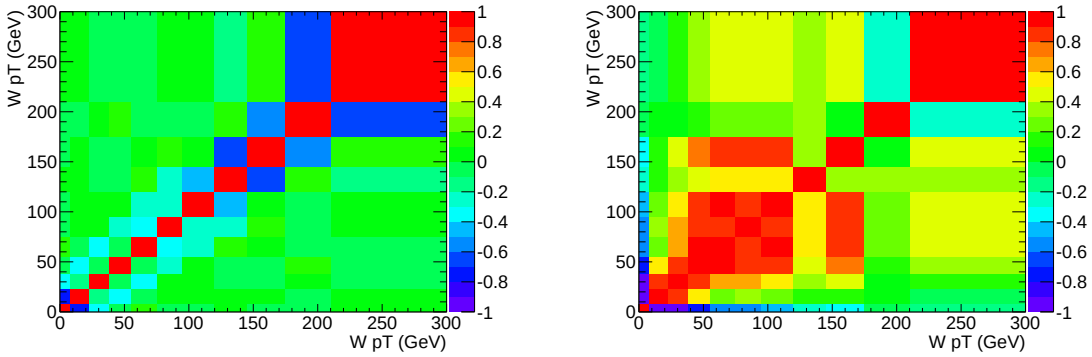


Figure 6.2: Correlation matrix for the fit parameterization uncertainty (left) and the fit statistical uncertainty (right) for the muon channel.

just take the geometrical average of the magnitudes:

$$C_{ij}^{e\mu}(\text{ideal}) = \sqrt{\sqrt{C_{ii}^{e}C_{jj}^{e}}\sqrt{C_{ii}^{\mu}C_{jj}^{\mu}}} \times \rho_{ij} \ . \tag{6.8}$$

The geometrical average reduces to the expected answer in simple cases such as having 0 or 100% correlation between bins. In reality, though, we don't get completely identical correlation matrices for the two channels. To smooth out differences between the channels in bins with small correlations and low statistics, while preserving the important sign information about the correlations, we take the geometrical average of the covariance matrices. The resulting formula is:

$$C_{ij}^{e\mu} = \sqrt{\sqrt{C_{ii}^{e}C_{jj}^{e}}\sqrt{C_{ii}^{\mu}C_{jj}^{\mu}}} \times \frac{1}{2}\left(\rho_{ij}^{e} + \rho_{ij}^{\mu}\right) \ . \tag{6.9}$$

Note also that the efficiency correction performed at the end of the analysis chain sculpts the covariance matrices but does not modify the correlation matrices.

For the uncertainties on the electroweak and top backgrounds that are correlated between the bins and the channels, to properly account for these correlations, the covariance matrices resulting from the luminosity and scale (cross section) variations must each be held separate from each other and the other variations, passed through unfolding separately, and combined only at the end. This differs from the procedure described in Sec. 5.4 but should result in a consistent covariance matrix.

Correlated among			
bins	channels	both	neither
Statistical (Unfolded)			
QCD MJ Bkg Shape		W/Z + top lumi scale	
top Bkg Shape		top Bkg Scale	
W/Z Bkg Shape		W/Z Bkg Scale	
	Unfolding Bias		
	Generator dependence		Lepton resolution and scale
	Fit Parameterization	Fit Statistics	
	SumET reweighting		
			Efficiency Correction

Table 6.1: Categorization of uncertainties by correlation pattern.

6.2 Combined p_T^W spectrum and comparisons to theoretical predictions and p_T^Z measurement

The common fiducial volume used for both electron and muon channel is described in Section 5.3.3 of Chapter 5. The fiducial cross sections ($d\sigma/dp_T^W$) binned in $p_T^{W,\text{true}}$ for the electron and muon channels are combined using the procedure described in Section 6.1. These inputs are compared in Figure 6.3 with a very good agreement evaluated by the ratio from them as shown in the bottom part of the figure. The normalized, fiducial cross sections $(1/\sigma)(d\sigma/dp_T^W)$ measured in individual channels, for different definitions of lepton kinematics, are listed in Table 6.2.

The combined result of p_T^W distribution presented here is normalized to the unity and the bin contents are divided by the bin width. This result is then compared to theoretical predictions and measurement of p_T^Z reconstructed from lepton pair. Firstly, the combined, normalized p_T^W distribution or fiducial cross section, $(1/\sigma)d\sigma/dp_T^W$, are compared with the prediction from RESBOS considered as a reference for this measurement (see Figure 6.4).

In case all uncorrelated uncertainties are considered, the combined result corresponds to a minimal χ^2 defined in equation (6.1) of $\chi^2/ndof = 9.56/13$. When the correlations between the channels are taken into account, as currently modeled, the combined result gives a minimum of $\chi^2/ndof = 12.33/13$. Here,

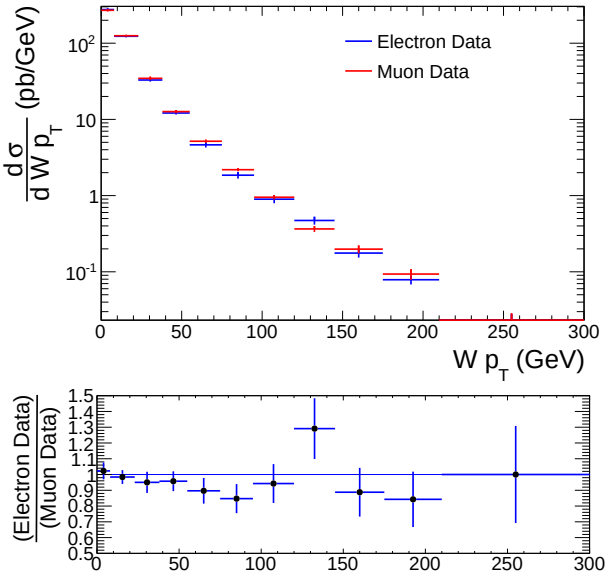


Figure 6.3: Electron and muon channel fiducial cross sections binned in $p_T^{W,\text{true}}$. The top panel shows the cross sections overlaid in log scale, and the bottom panel is the ratio between them.

in Figure 6.4, the error bars shown are the square roots of the diagonal elements of the total covariance matrix and include both the statistical and the systematic uncertainties. Particularly, the contributions of the various sources of uncertainty to the individual channel results, and to the combination result are displayed in Figure 6.5. In the lower range of p_T^W, below $\sim$ 50 GeV in electron channel and below $\sim$ 120 GeV in muon channel, the total uncertainty on individual channel measurements is between 5% and 8% at each bin of the spectrum. The larger uncertainty on the electron measurement in the p_T^W range between 50 and 150 GeV is due to the dominant of the uncertainty on the electron efficiency. The uncertainty on both electron and muon measurements increase up to 17% at higher p_T^W range, $p_T^W < 200$ GeV. In the combined measurement, the uncertainty is improved as expected. The total uncertainty on the combined result is about 4% in the range of $p_T^W < 100$ GeV. It increases slightly up to 10% with increasing p_T^W until 200 GeV. In all cases, the uncertainty increases dramatically at the last bin of the spectrum because of the low statistic. The detail combined results in all bins of $p_T^{W,\text{true}}$ are listed in Table 6.3.

There has been, recently, two measurements of p_T^W spectrum performed at Tevatron by CDF [6] and D0 [5] collaborations. The results of these measurements with comparisons to theoretical predictions are shown in Figure 6.6. The CDF result uses 2496 $W \to e\nu$ candidates selected from data corresponding to 4.1 pb^{-1}. The reconstructed p_T^W spectrum is corrected back to the parameterized true p_T^W distribution

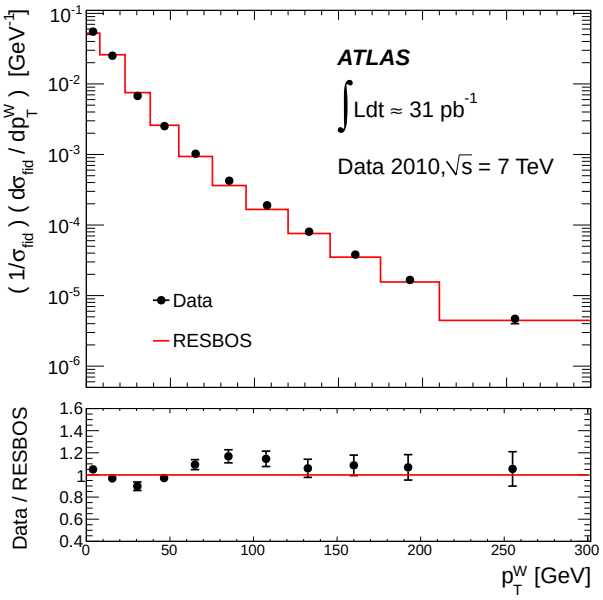


Figure 6.4: Normalized differential cross section obtained from the combined electron and muon measurements, compared to the RESBOS prediction.

on a bin-by-bin basis. The D0 collaboration measured the p_T^W spectrum using 7132 $W \to e\nu$ candidates selected from data corresponding to 12.4 pb^{-1} of integrated luminosity. The reconstructed p_T^W spectrum was compared with the theoretical prediction which had been smeared by an empirical model of the detector resolution and good agreement observed within the uncertainties. Both of these results are limited by the number of candidate events used, and by the partial unfolding which does not take into account bin-to-bin correlations.

Obviously, this p_T^W spectrum measurement at LHC, with more statistic (about 100,000 events for each channel) and fully correction accounting for bin-to-bin correlations using the data-driven response tuned from Z events, is the first real measurement resulting in a better precision than two earlier measurements at Tevatron. In addition, larger statistic allows us to extend the range of p_T^W so that the measurement becomes more robust in testing the QCD calculations at higher region of the spectrum.

In order to test QCD calculations for the p_T^W spectrum produced at LHC as described in Chapter 2, we compare the normalized differential cross section $(1/\sigma)d\sigma/dp_T^W$ obtained from the combination measurement to theoretical predictions modeled by various event generators including PYTHIA, RESBOS, SHERPA, ALPGEN + HERWIG, POWHEG + PYTHIA and MC@NLO. In addition, the combined result is compared with perturbative QCD calculations from DYNNLO and MCFM. All these comparisons are illustrated as the ratio of $(1/\sigma)d\sigma/dp_T^W$ from the predictions and measurement to RESBOS prediction

shown in Figure 6.7.

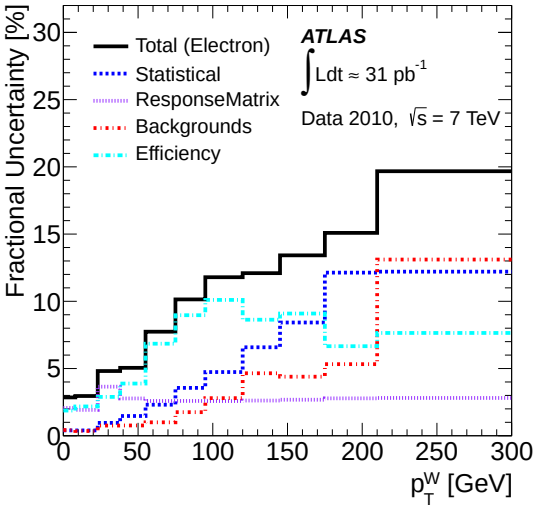
Overall the p_T^W spectrum, the predictions from PYTHIA, RESBOS, SHERPA and ALPGEN + HERWIG describe the measurement within 20%, especially, PYTHIA and RESBOS tuned from the Tevatron provide the best predictions for the LHC data. MC@NLO and POWHEG + PYTHIA, the NLO event generators interfaced with parton shower algorithms, give a reasonable description of the data for $p_T^W < 38$ GeV, but both underestimate the data for $p_T^W > 38$ GeV, with a deficit gradually increasing up to 40% at high p_T^W.

The p_T^W distribution from the DYNNLO and MCFM predictions are similar at the same order of $\mathcal{O}(\alpha_s)$ and $\mathcal{O}(\alpha_s^2)$ which are, respectively, the leading order and next to leading order calculation for W + 1 parton production. These predictions do not include resummation effects so they are not expected to predict the data well at low p_T^W because of the diverging prediction for vanishing p_T^W as discussed in Chapter 2. Therefore, the lowest bin ($p_T^W < 8$ GeV) is omitted from Figure 6.7. At higher range of p_T^W, the $\mathcal{O}(\alpha_s)$ order from these calculations underestimate the data about 30% at $p_T^W \sim 23$ GeV and get closer to data with increasing p_T^W. The $\mathcal{O}(\alpha_s)$ prediction from FEWZ is not shown in Figure 6.7 but is in agreement with those from DYNNLO and MCFM calculations at the same order. The $\mathcal{O}(\alpha_s^2)$ predictions, which are wihtin 15% of data for the entire range of p_T^W, describe better the data than those at $\mathcal{O}(\alpha_s)$.

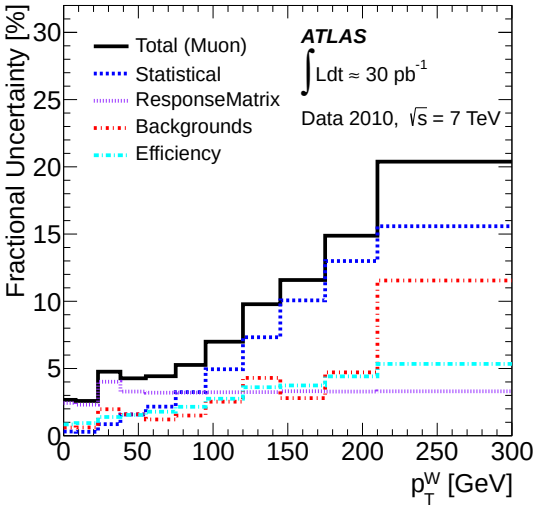
Finally, we compare the combined result to the measurement of $(1/\sigma)d\sigma/dp_T^Z$ described in [28]. The combined measurement of $(1/\sigma)d\sigma/dp_T^Z$, in comparisons with predictions from various Monte Carlo generators is displayed in Figure 6.8. In the high p_T^Z range, all predictions except for the MC@NLO are in agreement with the measurement within uncertainties. In the low p_T^Z range, FEWZ shows the largest deviation from the data in particular below $\sim$ 20 GeV while PYTHIA and RESBOS still provide good predictions for the measurement. This behavior of data/MC comparison in p_T^Z spectrum measurement is compatible with that in p_T^W spectrum measurement. In particular, the ratio of data to the RESBOS prediction from the two measurements overlaid in Figure 6.9 shows the similar trends, though, the p_T^W measurement with 10 times more statistic gives more reliable results in a better agreement with the theoretical prediction at high p_{T} range. This implies that W and Z productions differ from the generators in the same way, describing the p_T^W and p_T^Z spectrum. This confirms that the QCD mechanisms are universal between the two productions.

p_T^W Bin	$(1/\sigma)(d\sigma/p_T^W)$ (GeV^{-1})											
[GeV]	$W \to e\nu$				uncert. (%)		$W \to \mu\nu$				uncert. (%)	
	propag.	dressed	bare	p	stat.	syst.	propag.	dressed	bare	p	stat.	syst.
$0-8$	5.60	5.55	5.42	10^{-2}	0.4	2.8	5.44	5.39	5.35	10^{-2}	0.4	2.6
$8-23$	2.50	2.52	2.56	10^{-2}	0.4	2.9	2.52	2.54	2.55	10^{-2}	0.3	2.6
$23-38$	6.66	6.76	6.96	10^{-3}	0.9	4.7	6.96	7.06	7.11	10^{-3}	0.8	4.7
$38-55$	2.46	2.46	2.46	10^{-3}	1.3	4.8	2.55	2.55	2.55	10^{-3}	1.3	4.0
$55-75$	9.39	9.35	9.19	10^{-4}	2.0	7.4	1.04	1.04	1.03	10^{-3}	2.0	3.9
$75-95$	3.75	3.73	3.64	10^{-4}	3.4	9.5	4.40	4.37	4.34	10^{-4}	3.3	4.1
$95-120$	1.82	1.80	1.75	10^{-4}	4.1	10.8	1.92	1.90	1.88	10^{-4}	4.4	4.9
$120-145$	9.56	9.49	9.19	10^{-5}	6.0	10.1	7.35	7.29	7.21	10^{-5}	7.5	6.4
$145-175$	3.57	3.54	3.43	10^{-5}	7.9	10.4	3.99	3.96	3.91	10^{-5}	11.0	5.8
$175-210$	1.59	1.58	1.52	10^{-5}	10.0	8.9	1.88	1.86	1.84	10^{-5}	14.7	7.4
$210-300$	4.71	4.67	4.49	10^{-6}	12.2	15.5	4.68	4.66	4.55	10^{-6}	17.9	13.1

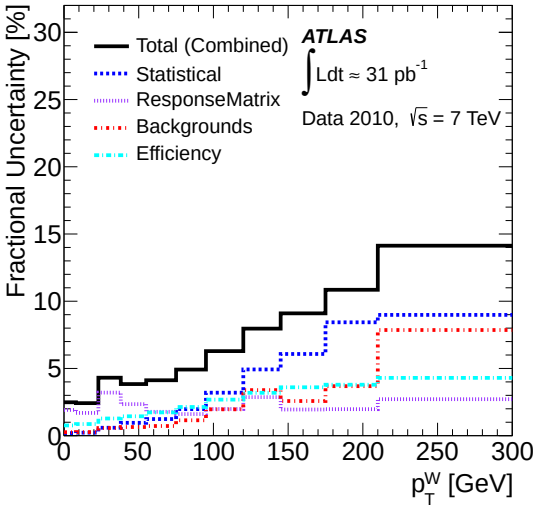
Table 6.2: The normalized, differential cross section $(1\sigma)(d\sigma/dp_T^W)$, measured in $W \to e\nu$ and $W \to \mu\nu$ events, for different definitions of p_T^W. The Born-level definition ("propag."), the analysis baseline, ignores the leptons and takes the W momentum from the propagator. The "dressed" and "bare" definitions of p_T^W are calculated using the momenta of the leptons from the W decay. In the "dressed" case, the charged lepton momentum includes the momenta of photons radiated within a cone of $\Delta R = 0.2$ centered around the lepton. In the "bare" case, the charged lepton momentum after all QED radiation is used. The factor p is the power of ten to be multiplied by each of the three cross section numbers for each channel. It has been factorized out for legibility.



(a) Electron Channel Measurement



(b) Muon Channel Measurement

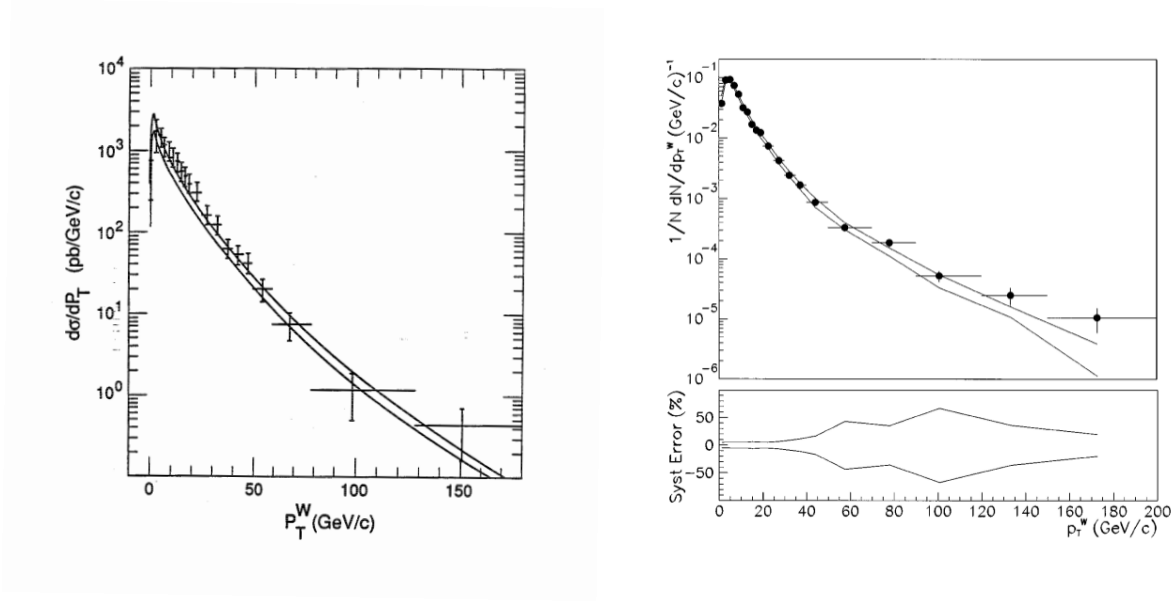


(c) Combined Measurement

Figure 6.5: Fractional uncertainties on the normalized $(1/\sigma)d\sigma/dp_T^W$ results for the electron, muon, and combined data, with the contributions from different categories shown.

p_T^W Bin [GeV]	$(1/\sigma)(d\sigma/dp_T^W)$ (GeV^{-1})	ResponseMatrix uncert. (%)	Backgrounds uncert. (%)	Efficiency uncert. (%)	Statistical uncert. (%)	Total uncert. (%)
$0-8$	$5.510\cdot10^{-2}$	1.91	0.26	0.76	0.22	2.48
$8-23$	$2.512\cdot10^{-2}$	1.69	0.28	0.87	0.24	2.42
$23-38$	$6.766\cdot10^{-3}$	3.20	0.57	1.28	0.57	4.31
$38-55$	$2.523\cdot10^{-3}$	2.34	0.65	1.44	0.84	3.78
$55-75$	$1.025\cdot10^{-3}$	1.78	0.74	1.74	1.19	4.09
$75-95$	$4.263\cdot10^{-4}$	1.61	1.15	2.13	1.91	4.94
$95-120$	$1.896\cdot10^{-4}$	1.98	1.94	2.67	2.68	5.99
$120-145$	$7.985\cdot10^{-5}$	2.84	3.30	3.16	4.78	7.91
$145-175$	$3.710\cdot10^{-5}$	1.98	2.66	3.66	5.72	9.31
$175-210$	$1.692\cdot10^{-5}$	2.00	3.72	3.84	7.75	10.56
$210-300$	$4.803\cdot10^{-6}$	2.69	7.81	4.26	9.28	14.40

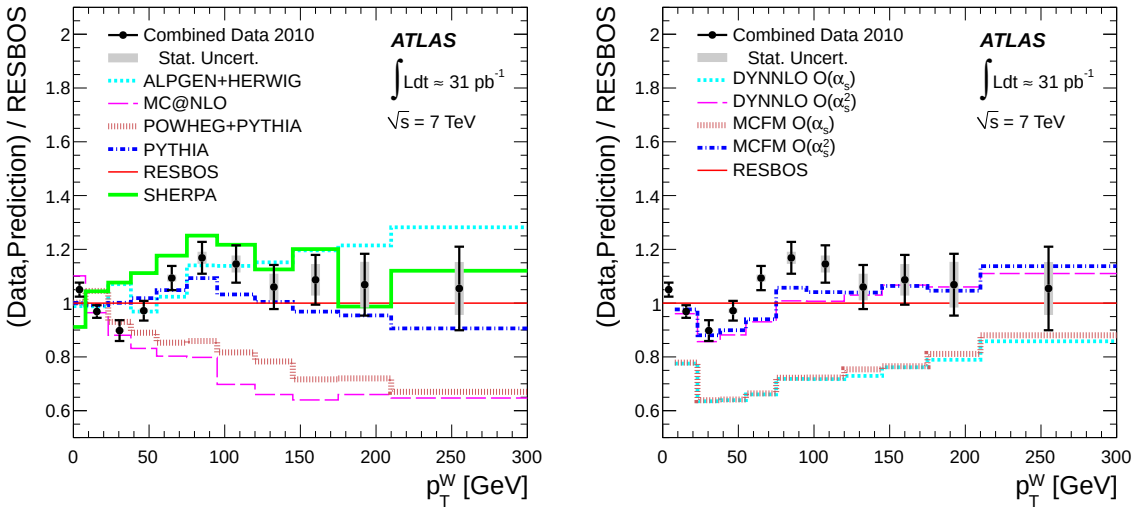
Table 6.3: Measured p_T^W using combined electron and muon data, with all uncertainties shown by source.



(a) CDF result [6]

(b) D0 result [5]

Figure 6.6: Measurement of p_T^W spectrum at Tevatron by CDF and D0 with comparisons to the theoretical predictions.



(a) Comparisons to various MC generators

(b) Comparisons to fixed orders pQCD

Figure 6.7: Ratios of combined measurement (using electron and muon decay modes) and various predictions to RESBOS for $(1/\sigma)(d\sigma/dp_T^W)$: (a) theoretical predictions modeled by various Monte Carlo generators; (b) perturbative QCD calculations at $\mathcal{O}(\alpha_s)$ and $\mathcal{O}(\alpha_s^2)$ from DYNNLO and MCFM.

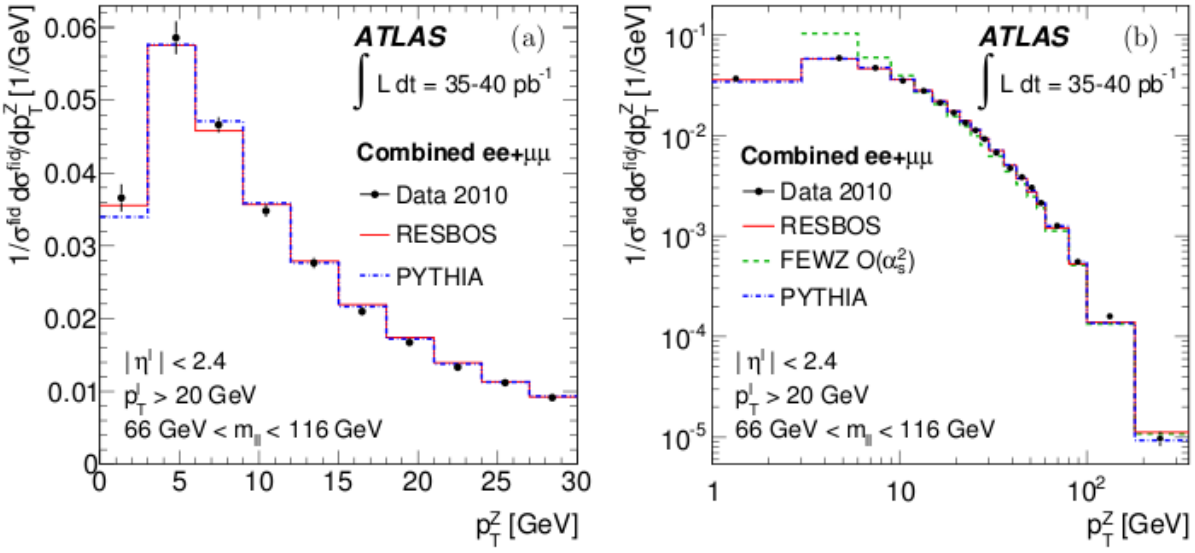


Figure 6.8: Combined measurement of p_T^Z spectrum in electron and muon channels, with comparisons to theoretical predictions modeled by various generators [28].

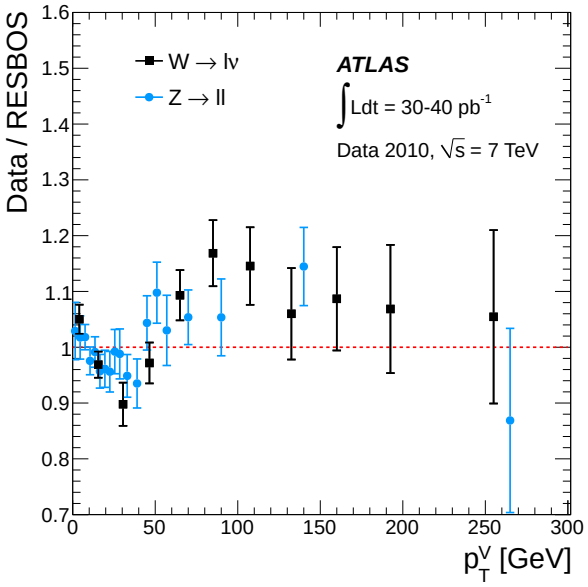


Figure 6.9: The ratio of $(1/\sigma)d\sigma/dp_T^W$ as measured in the combined electron and muon data to the RESBOS prediction, overlaid with the ratio of $(1/\sigma)d\sigma/dp_T^Z$ measured as described in [28] to the RESBOS prediction.

Chapter 7

Conclusions

The differential cross section as a function of the transverse momentum of W bosons has been measured for the range between 0 and 300 GeV in $W \to \ell\nu$ events. The data used, with approximately 31 pb^{-1} integrated luminosity, was collected during 2010 run of the LHC at $\sqrt{s} = 7$ TeV of pp collisions, and recorded by the ATLAS detector. This is the first complete measurement with taking into account the bin-to-bin migrations, performing data-driven calibrations and improving uncertainties compared to the previous measurements.

The measurement is compared to predictions modeled by different Monte Carlo event generators and calculated from fix-order perturbative QCD by different programs. The predictions from event generators such as ALPGEN + HERWIG, PYTHIA, RESBOS, and SHERPA match the data within 20% over the entire p_T^W range. For the next to leading order generators interfaced with parton shower algorithms, MC@NLO provides the closest description of the data for $p_T^W < 38$ GeV, but MC@NLO and POWHEG + PYTHIA both underestimate the data at higher p_T^W. Fixed-order perturbative QCD predictions from the DYNNLO and MCFM programs agree very well with each other. They predict fewer events at high p_T^W at $\mathcal{O}(\alpha_s)$ but the agreement with the measured distribution is significantly improved by the $\mathcal{O}(\alpha_s^2)$ calculations.

In spite of using a very different technique from the measurement of the differential cross section as a function of p_T^Z, this measurement displays similar behaviors in comparisons with predictions. This provides the first confirmation for the expected university of strong interaction effects between W and Z productions. Also, this gives additional confidence in the measurement.

With the data-driven calibrations performed in this measurement, the dominant uncertainty sources can be constrained with more data. For example, with the data available from the 2011 run now in

progress of the LHC, future measurements should not only be improved statistical and systematic uncertainties but also be able to measure the p_T^W spectrum to at least double the current range in p_T^W. In addition, it should also be possible to measure the ratios of the W to Z and W^+ to W^- differential cross sections as functions of the boson p_{T}, which will further test the predictions of QCD.

Appendix A

Response Gaussian Fits

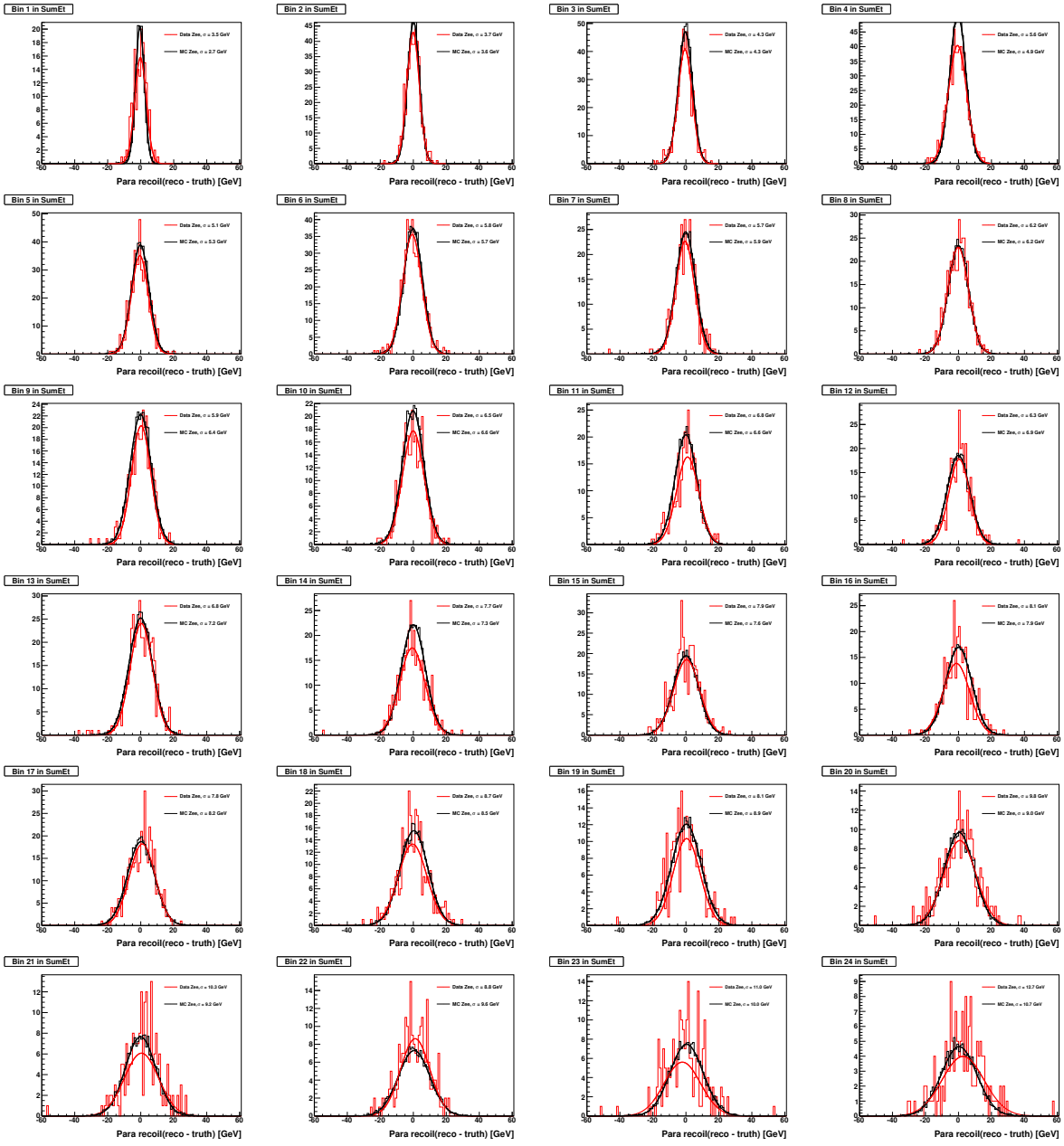


Figure A.1: Gaussian fits of $(p_{\parallel}^{reco} - p_{\parallel}^{true})$ slice in bins of ΣE_T in Z data and Z Monte Carlo.

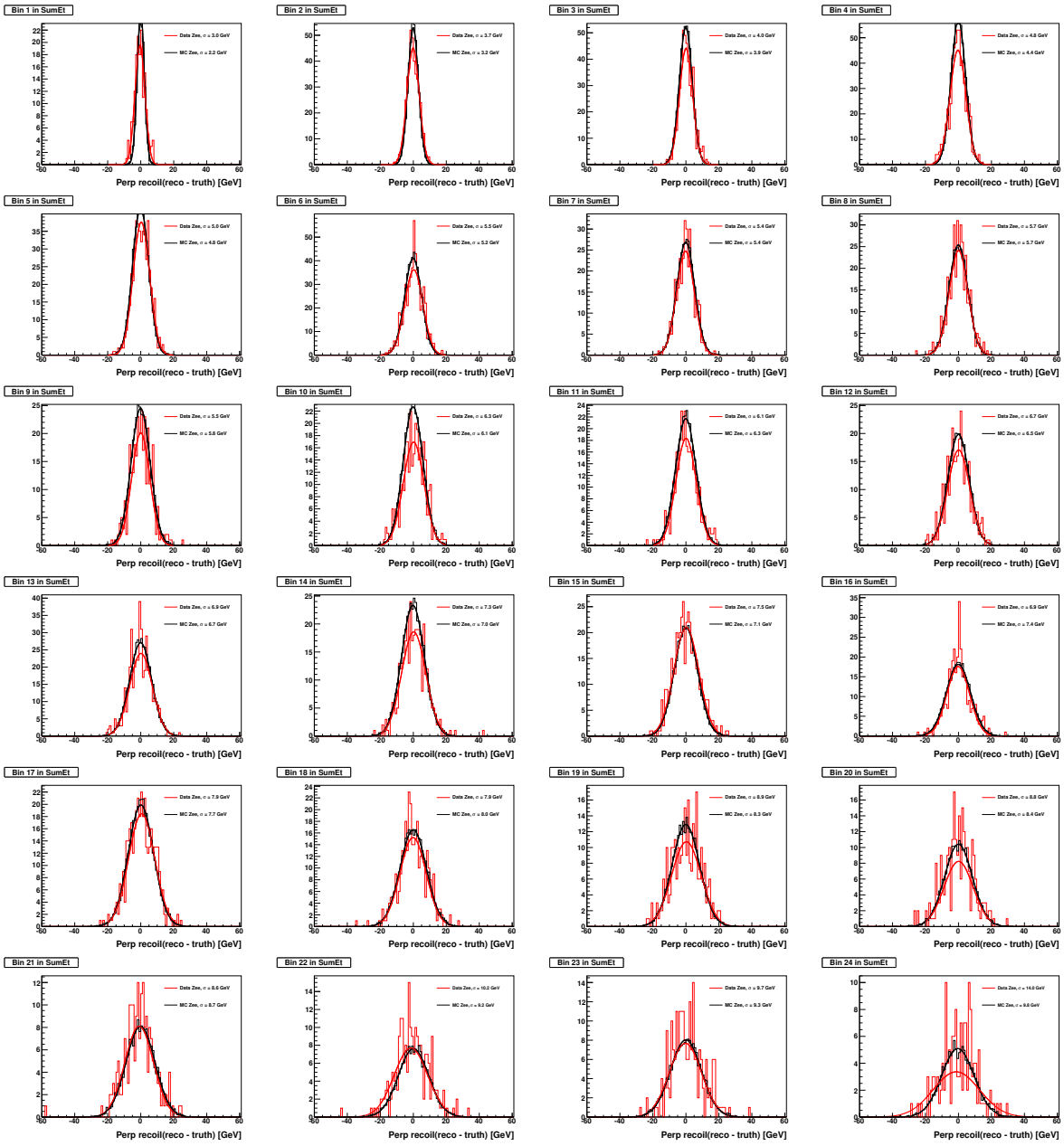


Figure A.2: Gaussian fits of $(p_{\perp}^{reco} - p_{\perp}^{true})$ slice in bins of ΣE_T in Z data and Z Monte Carlo.

List of Figures

List of Tables

Bibliography

[1] Class description for tdecompchol. http://root.cern.ch/root/html/TDecompChol.html.

[2] G. Aad et al. "Electron and Photon Reconstruction and Identification in ATLAS: Expected Performance at High Energy and Results at $\sqrt{s} = 900$ GeV". ATLAS internal note: ATLAS-CONF-2010-005.

[3] G. Aad et al. "Luminosity Determination Using the ATLAS Detector". ATLAS-CONF-2010-060.

[4] V.M. Abazov et al. "A novel method for modeling the recoil in W boson events at hadron colliders". arXiv: 0907.3713v1 [hep-ex], 2009.

[5] B. Abbott et al. "Measurement of the shape of the transverse momentum distribution of W bosons produced in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV". *Phys. Rev. Lett.*, 80:5498–5503, 1998.

[6] F. Abe et al. "Measurement of the pT(W) distribution in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV". *Phys. Rev. Lett.*, 66:2951–2955, 1991.

[7] E.S. Abers and B. W. Lee. *Phys. Rep.*, 9C:1–141, 1973.

[8] P. Arnold and M. Reno. *Nucl. Phys. B*, 319:37, 1989.

[9] C. Balazs and C. P. Yuan. " Soft gluon effects on lepton pairs at hadron colliders ". *Phys. Rev. D*, 56:5558–5583, 1997.

[10] Interactions Image Bank. " Illustration of particle detection in the subsystems of the ATLAS detector ". http://www.interactions.org/cms/?pid=2100&image no=CE0155.

[11] E. Barberio and Z. Was. . *Comput. Phys. Commun.*, 79:291, 1994.

[12] Roger J. Barlow and Christine Beeston. “Fitting using finite Monte Carlo samples ”. *Comput. Phys. Commun.*, 77:219–228, 1993.

[13] M. Bellomo et al. “Measurement of $W \to \ell\nu$ charge asymmetry in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector ”. ATLAS internal note: ATL-COM-PHYS-2010-1032.

[14] V. Blobel. “ An Unfolding Method For High Energy Physics Experiments ”. arXiv: 0208.022 [hep-ex].

[15] N. Cabbibo. “ Unitary Symmetry and Leptonic Decays ”. *Phys. Rev. Lett.*, 10:531–533, 1963.

[16] John M. Campbell, R. Keith Ellis, and David L. Rainwater. “ Next-to-leading order QCD predictions for W + 2jet and Z + 2jet production at the CERN LHC ”. *Phys. Rev. D*, 68:094021, 2003.

[17] Qing-Hong Cao and C.-P. Yuan. Resbos-a: an upgrade of resbos. http://hep.pa.msu.edu/resum/grids/resbosa/resbosa-grid.html.

[18] Stefano Catani, Leandro Cieri, Giancarlo Ferrera, Daniel de Florian, and Massimiliano Grazzini. “Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO ”. *Phys. Rev. Lett.*, 103:082001, 2009.

[19] ATLAS Experiment @ 2011 CERN. ATLAS Photos. http://www.atlas.ch/photos/index.html.

[20] ATLAS Collaboration. “ Jet energy scale and its systematic uncertainty in proton-proton collision at $\sqrt{s} = 7$ TeV with ATLAS 2010 data ”. ATLAS-CONF-2011-032.

[21] ATLAS Collaboration. “ The ATLAS detector performance ”. ATL-PHYS-PROC-2011-046.

[22] ATLAS Collaboration. “ATLAS Calorimeter Performance”. CERN-LHCC-96-40.

[23] ATLAS Collaboration. “Expected performance of the ATLAS experiment : detector, trigger and physics”. CERN-OPEN-2008-020.

[24] ATLAS Collaboration. “Kinematic distributions of K_s^0 and Λ^0 decays in collision data at $\sqrt{s} = 7$ TeV”. ATLAS-CONF-2010-033.

[25] ATLAS Collaboration. “Performance of the missing transverse energy reconstruction and calibration in proton-proton collisions at $\sqrt{s}$ of 7 TeV with the ATLAS detector”. ATLAS-CONF-2010-057.

[26] ATLAS Collaboration. "Readiness of ATLAS Tile Calorimeter for LHC collisions". arXiv: 1007.5423 [physics.ins-det], 2010.

[27] ATLAS collaboration. " Measurement of $W \to l\nu$ and $Z/\gamma^* \to ll$ production cross sections in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector ". *JHEP*, 2012:060, 2010; arXiv:1010.2130v1 [hep-ex].

[28] ATLAS Collaboration. "Measurement of the transverse momentum distribution of Z/γ^* bosons in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". *Phys. Lett. B*, 705:415–434, 2011.

[29] ATLAS Collaboration. "Measurement of the transverse momentum distribution of W bosons in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". *Phys. Rev. D*, 85:012005, 2012.

[30] D0 Collaboration. "A Measurement of the W Boson Mass with 1 fb^{-1} of D0 Run II Data". D0 Note 5893-CONF, 2009.

[31] The ATLAS Collaboration. "Performance of missing transverse momentum reconstruction in proton-proton collisions at 7 TeV with ATLAS". *Eur.Phys. J. C*, 72:1844, 2010; arXiv:1108.5602v2 [hep-ex].

[32] J. C. Collins and D.E. Soper. *Nucl. Phys. B*, 193:381, 1981.

[33] G. Corcella et al. "HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes)". *JHEP*, 01:010, 2001.

[34] G. Cowan. *"Statistical Data Analysis"*. Oxford University Press, 1998.

[35] G. D'Agostini. "Improved Iterative Bayesian Unfolding". arXiv: 1010.0632 [physics.data-an], 2010.

[36] E. Dobson. "Detector response and impact on acceptance calculations for $W \to e\nu$ and $Z \to ee$ cross sections and their ratio". ATLAS internal note: ATL-PHYS-INT-2009-102, 2009.

[37] S.D. Drell and T.M. Yan. *Phys. Rev. Lett.*, 25:316, 1970.

[38] G. DAgostini. "A Multidimensional unfolding method based on Bayes theorem". *Nucl. Instr. Meth.A*, 362:487, 1995.

[39] F. Landry, R. Brock, P. M. Nadolsky, and C. P. Yuan. *Phys. Rev. D*, 67:073016, 2003; arXiv: 0212159 [hep-ph].

[40] Stefano Frixione and Bryan R. Webber. "Matching NLO QCD computations and parton shower simulations". *JHEP*, 06:029, 2002.

[41] G. Marchesini and B.R. Webber. *Nucl. Phys. B*, 310:461, 1988.

[42] G. Parisiand and R. Petronzio. *Nucl. Phys. B*, 154:427, 1979.

[43] Ryan Gavin, Ye Li, Frank Petriello, and Seth Quackenbush. "FEWZ 2.0: A code for hadronic Z production at next-to- next-to-leading order". arXiv: 1011.3540 [hep-ph], 2010.

[44] T. Gleisberg et al. JHEP 02, 007 (2009), arXiv:0811.4622 [hep-ph].

[45] R. J. Gonsalves et al. *Phys. Lett. B*, 252:663, 1990.

[46] Egamma Group. Efficiency measurements. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EfficiencyMeasurements.

[47] Egamma Group. Energy resolution recommendations. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/EnergyScaleResolutionRecommendations.

[48] Egamma Group. isem identification. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ IsEMIdentification.

[49] Egamma Group. Object quality maps usage. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/OQMapsUsage.

[50] Jet/EtMiss Group. How to clean jets. https://twiki.cern.ch/twiki/bin/view/AtlasProtected/HowToCleanJets.

[51] P. C. Hansen. *The L-Curve and its Use in the Numerical Treatment of Inverse Problems*, pages 119–142. WIT Press, 2001.

[52] A. Hocker and V. Kartvelishvili. "SVD Approach to Data Unfolding". arXiv: 9509307v2 [hep-ph], 1995.

[53] J. Iliopoulos. *An Introduction to Gauge Theories*. CERN Yellow Rep. 76-11, 1976.

[54] J. C. Collins and D. E. Soper. *Ann. Rev. Nucl. Part. Sci.*, 37:383, 1987.

[55] J. C. Collins, D.E. Soper and G. Sterman. "Transverse momentum distribution in Drell-Yan pair and W and Z boson production". *Ref. TH.3923 - CERN*, 1984.

[56] J. C. Collins, D.E. Soper and G. Sterman. *Nucl. Phys. B*, 261:104, 1985.

[57] J. C. Collins, D.E. Soper and G. Sterman. *Nucl. Phys. B*, 250:199, 1985.

[58] F. James. Minuit Function Minimization and Error Analysis Reference Manual. 1994. CERN Program Library Long Writeup D506.

[59] J.M. Campbell, J.W. Huston and W.J. Stirling. "Hard Interactions of Quarks and Gluons: a Primer for LHC Physics". arXiv: 0611148 [hep-ph].

[60] J. Kaipio and E. Somersalo. *Statistical and Computational Inverse problems*. Springer, 2005.

[61] O. Kepka. *QCD et diffraction dans l'experience ATLAS au LHC*. PhD thesis, Universite Paris XI et De L'universite Charles, Prague, 2009.

[62] M. Kobayashi and T. Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". *Prog. Theor. Phys.*, 49:652–657, 1973.

[63] G. A. Ladinsky and C. P. Yuan. . *Phys. Rev. D*, 50:4239, 1994.

[64] W. Lampl et al. "Calorimeter Clustering Algorithms: Description and Performance". ATLAS-LARG-PUB-2008-002.

[65] E. Lifshitz and E. Pitayevski. *Relativistic Quantum Mechanics*. Oxford: Pergamon, 1976.

[66] M. Boonekamp, D. Froidevaux, and M. Schott. "Total inclusive W and Z boson cross-section measurements, cross-section ratios and combinations in the electron and muon decay channels at 7 TeV based on 300 nb^{-1}". ATL-COM-PHYS-2010-703, 2010.

[67] B. Malaescu. "An Iterative, Dynamically Stabilized Method of Data Unfolding". arXiv:0907.3791[physics.data-an].

[68] Michelangelo L. Mangano, Mauro Moretti, Fulvio Piccinini, Roberto Pittau, and Antonio D. Polosa. "ALPGEN, a generator for hard multiparton processes in hadronic collisions". *JHEP*, 07:001, 2003; arXiv: 0206293 [hep-ph].

[69] W. Marciano and H. Pagels. Quantum chromodynamics. *Phys. Rep.*, 36C:137, 1978.

[70] Kirill Melnikov and Frank Petriello. "Electroweak gauge boson production at hadron colliders through O(alpha(s)**2)". *Phys. Rev. D*, 74:114017, 2006; arXiv: 0609070 [hep-ph].

[71] N. Besson, M. Boonekamp, E. Klinkby, S. Mehlhase and T. Petersen. "Re-evaluation of the LHC potential for the measurement of m_W". arXiv: 0805.2093 [hep-ex], 2008.

[72] K. Nakamura et al. "Review of Particle Physics: CKM Quark-Mixing Matrix". *J. Phys. G*, 37:075021, 2010.

[73] M. Nessi. *LHC accelerator and experiments*. 2007 ICTP summer school on particle physics, Abdus Salam ICTP, Trieste, 11-22 June 2007, 2007.

[74] P. D. B. Collins, A. D. Martin. *Hadron Interactions*. Adam Hilger Ltd, Bristol (published in association with the University of Sussex Press), 1984.

[75] M.E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Westview Press, 1995.

[76] H. D. Polizer. *Phys. Rev. Lett.*, 30:1346, 1973.

[77] R. J. Gonsalves, N. Kidonakis and A. S. Vera. "W Production at Large Transverse Momentum at the CERN Large Hadron Collider". *Phys. Rev. Lett.*, 95:222001, 2005.

[78] R. K. Ellis and S. Veseli. "W and Z transverse momentum distributions: resummation in q_T-space". *Fermilab-PUB-97/207-T*, 1997.

[79] S. Schmitt. TUnfold Class. http://root.cern.ch/root/html/TUnfold.html. Unfolding software, part of the ROOT data analysis framework.

[80] S. Schmitt. TUnfoldSys Class. http://root.cern.ch/root/html/TUnfoldSys.html. Unfolding with Systematic Errors software, part of the ROOT data analysis framework.

[81] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.

[82] P. Z. Skands. *Phys. Rev. D*, 82:074018, 2010; arXiv:1011.3540 [hep-ph].

[83] R. M. Thurman-Keup, A. V. Kotwal, M. Tecchio, and A. Byon-Wagner. "*W* boson physics at hadron colliders". *Rev. Mod. Phys.*, 73:267–306, 2001.

[84] A. N. Tikhonov. "On the Solution Of Improperly Posed Problems and the Method of Regularization". *Sov. Math*, 5:1035, 1963.

[85] Nguyen Thi Hong Van. *Discovering the standard model Higgs at the LHC*. Diploma thesis, Abdus Salam ICTP, Triest, Italy, 2007.

[86] Wikipedia. Standard model of elementary particles. http://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.