

UNIVERSITÀ DI PISA



DIPARTIMENTO DI FISICA "E.FERMI"

Corso di Laurea Magistrale in Fisica

Hadronic recoil in the W boson production at LHC for a W mass measurement with the CMS experiment

Candidato:
Olmo Cerri

Relatore:
prof. Luigi Rolandi

ANNO ACCADEMICO 2016/2017



Audentes Fortuna iuvat

*Alle persone che mi hanno aiutato a completare questo percorso.
Che possa restare accanto a voi e ricambiare ciò che ho ricevuto.*

Contents

Introduction	1
1 The Standard Theory of particles	3
1.1 Fields and Symmetries	3
1.2 Spontaneous symmetry breaking and Brout – Englert – Higgs mechanism	7
1.2.1 Gauge boson masses	7
1.3 Precision test of the ElectroWeak sector	8
1.3.1 Higher order prediction of W mass	9
1.3.2 The global electroweak fit	10
2 The Compact Muon Solenoid experiment	13
2.1 The Large Hadron Collider	13
2.1.1 Proton-Proton collision physics	16
2.2 The CMS detector	18
2.2.1 The layers structure	20
2.2.2 Trigger	22
2.3 Physics object reconstruction	22
2.3.1 Particle Flow algorithm	23
2.4 Simulation	24
3 The W boson mass measurement	27
3.1 From discovery to LEP	28
3.2 W mass measurement in modern hadron collider	31
3.2.1 Analysis strategy	31
3.2.2 Main uncertainties	34
3.2.3 Recent efforts summary	36

4	The hadronic recoil in the W mass measurement	39
4.1	Production Mechanism	40
4.1.1	Comparison between Z and W	41
4.2	Definition and characterization of hadronic recoil	42
4.2.1	Detector effects in measuring the recoil	43
4.3	Samples definition	44
4.3.1	Events selection	45
4.4	Hadronic recoil effect on transverse mass	48
4.4.1	W mass sensitivity and boson transverse momentum dependence	50
4.5	A precise measurement of the recoil	52
4.5.1	The recoils as a 2D object	52
4.5.2	Resolution effects	55
5	A new recoil definition	59
5.1	Recoil definitions	60
5.1.1	An improved recoil	61
5.2	A multivariate semi-parametric regression	62
5.2.0.1	Neural Networks: a generic function	62
5.2.0.2	Semi-parametric regression	64
5.2.1	Application to recoil definition	65
5.2.2	Details of the trained regression	68
5.3	Goodness of the definition	71
5.3.1	Regression - network convergence check	72
5.3.1.1	Sum of pdfs closure	73
5.3.2	Goodness event-by-event of recoil estimators	76
6	Systematic uncertainties	79
6.1	Systematic effects and factorization scheme	80
6.2	W mass fitting procedure	81
6.3	Effect of boson kinematic mismodeling	83
6.3.1	Estimating the systematic uncertainty	84
6.3.2	Recoil definitions ranking	88
6.4	Systematic on the recoil mismodeling	90
6.4.1	Differences between data and simulations	91
6.5	A multi-dimensional morphing procedure	92
6.5.1	Procedure for a n -dimensional k -conditional morphing	94

6.6	Recoil modeling effect and calibration	96
6.6.1	Procedure check	98
6.6.2	Systematic assessment before calibration	99
6.6.3	Systematic assessment after calibration	100
6.7	Porting the calibration from Z to W	102
6.7.1	Generated instead of reconstructed boson kinematic variables	102
6.7.2	Intrinsic MC differences between Z and W	103
7	Conclusions	105
7.1	Future developments	107
A	DNN training details	111
	Bibliography	113
	Acknowledgments	127

Introduction

After the discovery of the Higgs boson in 2012, the standard model of particle physics (SM) has been further validated and all its free parameters established. Using the full set of SM parameters, it is thus possible to predict with increasing precision relations among observables which can be verified in experiments. The agreement between measurements and theoretical predictions is a severe consistency test for the model: possible deviations of measured values from those predictions would be a clear sign of new physics beyond the standard model.

Specifically, precise determination of the W boson mass is of great importance in this testing procedure. From the time of its discovery in 1983, the W boson has been studied and its mass determined in both hadron and lepton colliders.

At the state of the art, the W boson mass has a smaller uncertainty in the SM prediction of 80.360 ± 0.007 GeV than in the measured value of 80.385 ± 0.015 GeV. A measurement of the W mass with an uncertainty smaller than 7 MeV might be a breakthrough and may result in a direct evidence of SM inconsistency. For this reason, such measurement is currently studied at the LHC. The CMS experiment is planning to deliver a measurement of W mass within the next years. At hadron colliders, production of on-shell W bosons is tagged by the high transverse momentum (p_T) charged lepton from its decay in the leptonic final state, the only one suitable for a precise mass measurement. In practice, W mass is extracted from data via a fit to several distributions, which need to be understood and optimized up to an unprecedented level of precision.

In the first chapter of this work, an overall picture of the theoretical basis is presented. Starting from the foundations of the Standard Model, Higgs mechanism and electroweak symmetry breaking are introduced, focusing on their role of providing SM gauge boson with masses. The important facts of electroweak precision test are also introduced in the last part of the first chapter.

After an overview of the Large Hadron Collider (LHC), which is currently operating at CERN, the second part of this work describes the Compact Muon Solenoid (CMS) experiment, aimed to explore in depth particle physics up to the TeV scale: the main

features of the subdetectors are briefly described, together with the reconstruction algorithms; focus has been put mostly on those features of interest for W mass physics. The third chapter is devoted to discuss the past and the on going efforts for the W boson mass measurement.

The original work developed during the thesis is fully discussed in chapters four, five and six. Two are the main objectives: to deepen the knowledge of the variables used in the W mass measurement, with particular attention on the event-by-event experimental estimator of the boson transverse momentum; to define and calibrate an experimental definition of the recoiling system to the W , suitable for the real measurement process at the CMS experiment. Transverse momentum of the recoil and of the boson are two faces of the same coin: both are crucial in the extraction of the mass value.

Events in which the W boson decays into a muon and a neutrino are considered. The W mass is extracted from the distributions of the modulus of the lepton transverse momentum (p_μ) and of the transverse mass (M_T), a scalar quantity function of the lepton momentum and the recoil, which is the vectorial sum of the momenta of all reconstructed particles excluding the lepton. With the purpose to maximize the performance in terms of systematic uncertainty on the final measurement, a new experimental definition of the recoil, based on machine learning algorithms, is discussed. As support to the remarkable impact of my work, there are reported cross checks and performances of the new definition, in terms of resolution and uncertainty improvement for a W mass measurement.

In the last chapter, recoil-related systematic uncertainties on the W mass measurement are presented. Furthermore, it is reported a new method, based on multi-dimensional morphing, used to calibrated the Monte Carlo simulation using collision data. The systematic uncertainties of the W mass measurement before and after this calibration are studied.

Finally, conclusions summarize the main results, underlining the importance of the work, and suggesting possible future developments.

Document notation

Throughout this document the classical notation of high energy physics will be used. In particular, the value of the speed of light and the Plank constant are set to $c = 1$ and $\hbar = 1$, so that masses, energies and momenta are all expressed in electron volts (eV).

All the angles are represented in the space $[-\pi, \pi]$.

Chapter 1

The Standard Theory of particles

The “Standard Model” (SM) is the theory that summarizes the knowledge we have about fundamental components of matter and their interactions; it describes a wide range of phenomena and has been tested experimentally with a remarkable accuracy, even if there are some important problems to which it gives no satisfactory answer. The theory has emerged gradually in almost a century of theoretical and experimental investigations and it is now fully established. I will describe it from the logical point of view. Quantum field theory, the framework in which SM is formulated, is a vast subject and even a short but comprehensive resume of it would be too long for this thesis, so I will only introduce the key concepts useful later on.

This compilation chapter is thus devoted to describe the SM accurately enough to serve as a theoretical background to the main subject of this thesis. A complete reference can be found in the original papers, such as [1], and in the classic textbooks, such as [2]. Additionally, given the main focus of this thesis, all mathematical details are left over.

1.1 Fields and Symmetries

The standard model is based on the existence of 17 different elementary fields which can be grouped into three categories.

The Matter category, which includes 12 fermionic fields with spin $1/2$. This category is split into 6 quarks and 6 leptons, organized into 3 generations all with a similar content but with generally growing masses. Each generation consists of one *up* quark with electric charge $2/3$ (u , c or t), one *down* quark with electric charge $1/3$ (d , s or b) and a lepton with electric charge 1 (e , μ or τ), with its associated charge-less neutrino

1.1. FIELDS AND SYMMETRIES

(ν_e , ν_μ or ν_τ). The neutrinos were first introduced in the SM as massless particles, but experiments have shown that they have a small but non-zero mass. This can be resolved, for example, adding 3 heavy right handed neutrinos. Moreover, each field in this category has an associated anti-particle, with same mass and opposite quantum numbers.

The gauge boson category comprehends 4 bosonic fields with spin 1. Each of these particles is associated with one fundamental interaction: electric force is mediated by the photon (γ), strong force by the gluons (g) and weak interactions by the $W^\pm$ and Z^0 bosons.

Finally, the Higgs sector, which has been recently observed [3], consists of only 1 neutral complex scalar field: the Higgs boson (H).

A summary chart of the SM is shown in Figure 1.1.

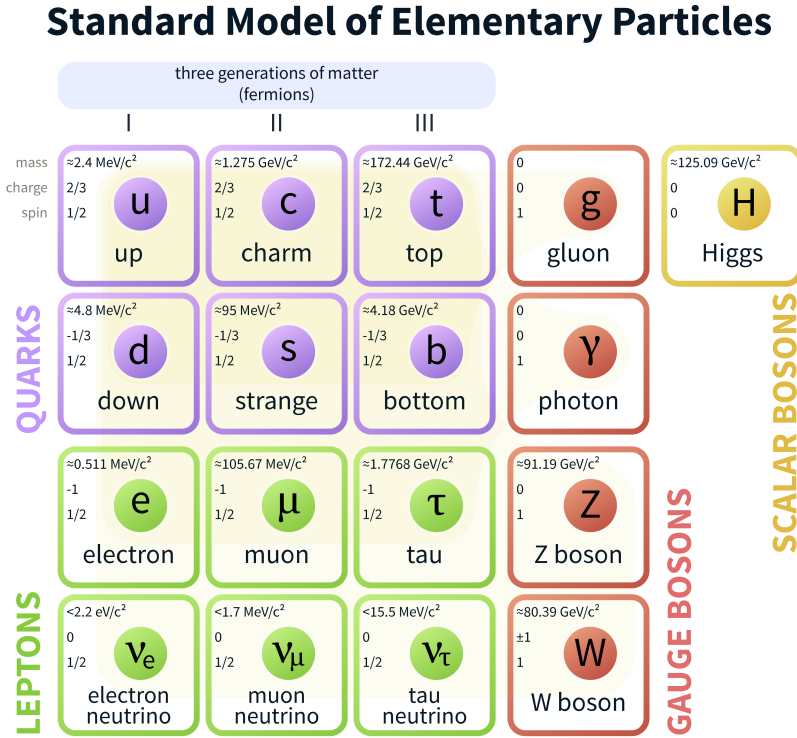


Figure 1.1: Fundamental fields presented in the SM [4]. The 3 matter generations are shown in the first 3 columns, while gauge bosons and scalar boson are shown respectively in the 4th and 5th column.

The SM is a renormalizable theory coherent with special relativity, based on gauge theories [5]. In this framework, vector bosons can arise only as gauge fields of an assumed internal local symmetry, where internal means that it does not touch the space-time degrees of freedom of a particle and local means that the parameters of this

1.1. FIELDS AND SYMMETRIES

transformation are functions of the space-time position.

An arbitrary gauge field theory [6] is characterized by a certain group of symmetries $\mathcal{G}$, with associated generators T_a satisfying the algebra

$$[T_a, T_b] = if_{abc}T_c ,$$

and the representation r_ψ and r_ϕ under $\mathcal{G}$ of fermionic and scalar fields.

Given these assumptions, the Lagrangian of the theory can be built to satisfy the symmetries simply replacing the derivatives with covariant derivatives of the form

$$D_\mu = \partial_\mu - igT^a A_\mu^a$$

where A_μ^a are the gauge vector fields and g is the charge of the field under the interaction induced by the symmetry. It has to be noticed that, if the group is simple, there is just one charge g , otherwise $\mathcal{G}$ can be the product of many groups $\mathcal{G}_i$ each with its own charge g_i . In general, for any simple non-abelian factor of $\mathcal{G}$ there is a single arbitrary constant, whereas the charges under any $U(1)$ -factor are all arbitrary.

Restricting ourselves to local operators of mass dimension at most four, the required invariances force the Lagrangian $\mathcal{L}_\mathcal{G}$ to be of the form

$$\mathcal{L}_\mathcal{G} = \mathcal{L}_{min} - \left(\frac{1}{2} \psi^T M \psi + \phi \psi^T \lambda_1 \psi + \phi^\dagger \psi^T \lambda_2 \psi + \text{h.c.} \right) + V(\phi) \quad (1.1)$$

where ψ and ϕ are the fermionic and scalar fields, M and $\lambda_{1,2}$ are in general (sometimes necessarily vanishing) matrices so that invariance under $\mathcal{G}$ is guaranteed and, finally,

$$\begin{aligned} \mathcal{L}_{min} &= -\frac{1}{4} F_{\mu\nu}^a F_{\mu\nu}^a + i \bar{\psi} \not{D} \psi + |D_\mu \phi|^2 \\ F_{\mu\nu}^a &= \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g f_{abc} A_\mu^b A_\nu^c \end{aligned}$$

The gauge group of the SM is the direct product

$$\mathcal{G}_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y.$$

As Noether's theorem asserts, each symmetry corresponds to a conserved quantity. The SM global symmetries, induced by $\mathcal{G}_{SM}$, have as associated quantity the color, the weak isospin and the hypercharge.

The Quantum Chromo-Dynamics (QCD) arises from the $SU(3)$ symmetry and describes the strong interaction among quarks and gluons. This non-abelian group has

1.1. FIELDS AND SYMMETRIES

8 generators and thus in the theory 8 massless self-interacting vector bosons emerge. The quarks fill the simplest non trivial representation for the fermions (3 and $\bar{3}$), whereas leptons and the Higgs boson are singlets under this group.

At low energies, the interactions among quarks are very strong and no perturbative treatment can be done. As an experimental fact the *confinement* happens, that is all free particles are colorless: mesons are indeed in the scalar $\bar{q}^a q^a$ state, whereas barions are in the anti-symmetric one $\epsilon^{abc} q^a q^b q^c$.

In the ultraviolet regime ($\Lambda \gg 200\text{MeV}$), the behavior changes dramatically because vacuum polarization for this theory is such that renormalized charge decreases with decreasing distance or increasing energy. This phenomenon, called asymptotic freedom, allows a perturbative treatment of strong interactions at high energies.

The electromagnetic and weak interactions are usually treated together as they emerge from the breaking of a $SU(2) \otimes U(1)$ symmetry, as it is described in sec. 1.2. This two groups bring respectively 3 self-interacting ($\vec{W}$ bosons) and 1 non self-interacting (B boson) massless vector bosons.

The left-handed fermions and the Higgs boson fill the non trivial doublet state, whereas the right-handed ones are singlet under this group.

$U(1)_Y$ is the only abelian group of the SM and it allows for arbitrary values of the particle observed electric charges.

Table 1.1: Fields content of the SM from a gauge point of view, the three generations of fermions have not been distinguished since they have exactly the same gauge interactions.

Particle	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
$L = \begin{pmatrix} \nu \\ e_L \end{pmatrix}$	1	2	$-1/2$
e_R	1	1	-1
$Q = \begin{pmatrix} u_L \\ d_L \end{pmatrix}$	3	2	$+1/6$
u_R	3	1	$+2/3$
d_R	3	1	$-1/3$
H	1	2	1
g	8	1	0
$\vec{W}$	1	3	0
B	1	1	0

The fields content and the symmetries of SM here presented (summarized in tab. 1.1) are still unsatisfactory because, above all, it fails to give a clear explanation to the observed mass spectrum.

1.2 Spontaneous symmetry breaking and Brout – Englert – Higgs mechanism

Given the SM group of symmetry, all particles must be massless since there are no gauge-invariant mass terms which can be added to give directly mass to fermions and gauge bosons. As a result, the gauge symmetry must be broken and, to leave unharmed the predictive power of the theory, only spontaneous symmetry breaking is allowed (SSB).

The SSB is a feature of systems in which a symmetric Lagrangian can produce a physics which is not symmetric: in QFT this can happen when the vacuum state of the theory is not symmetric and thus all the observables, which are vacuum expectation values of some functions of the fields, are not symmetric too. In order for the gauge bosons to acquire mass, it is necessary to introduce in the theory a way to make possible a realization of the vacuum states that break the $SU(2)_L$ electroweak symmetry. This mechanism has been proposed by Higgs [7], Brout and Englert [8]. A similar procedure can produce masses for the fermions; the description of this is out of the purpose of this thesis but can be found in [2].

1.2.1 Gauge boson masses

As it has been shown in the previous section, to build a SM Lagrangian, which is invariant under $SU(2)_L \otimes U(1)_Y$, it is necessary to introduce three fields W_μ^a transforming as a triplet of $SU(2)$ and with an associate coupling constant g and a field associated with $U(1)$ symmetry B_μ with coupling constant g' . It can be chosen, then, the scalar sector of the SM to be made of a scalar field Φ , which is doublet under $SU(2)_L$ and has hypercharge 1; its covariant derivative can be written as

$$D_\mu \Phi = \partial_\mu \Phi - i\frac{1}{2}g\vec{W}_\mu \cdot \vec{\sigma}\Phi - i\frac{1}{2}g'B_\mu \Phi$$

where σ stands for the standard Pauli matrices.

The most general renormalizable scalar potential for Φ that can be plugged in eq. 1.1 assumes the form

$$V(\Phi) = m^2\Phi^\dagger\Phi + \lambda(\Phi^\dagger\Phi)^2.$$

If $m^2 < 0$ and $\lambda > 0$, then it has a series of degenerate minima when $\Phi^\dagger\Phi = -m^2/\lambda$. This property brings the scalar doublet to acquire a non-zero vacuum expectation value

(VEV) which in the unitary gauge has the form

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$$

The choice of one particular value breaks the SM gauge symmetry $\mathcal{G}_{SM}$ into $SU(3)_C \otimes U(1)_{em}$ and sets the perturbative expansion to be around the minimum, using a scalar field H , the Higgs boson. Given this, the quadratic terms in v , coming from $|D_\mu \Phi|$, provides masses to the gauge bosons which result mixed in the manifest mass eigenstates

$$\begin{aligned} W^\pm &= \frac{1}{\sqrt{2}}(W^1 \mp W^2) & m_W &= g \frac{v}{2} \\ Z^0 &= \frac{1}{\sqrt{g^2 + g'^2}}(gW^3 - g'B) & m_Z &= \frac{v}{2}\sqrt{g^2 + g'^2} \\ A &= \frac{1}{\sqrt{g^2 + g'^2}}(gW^3 + g'B) & m_A &= 0 \end{aligned} \quad (1.2)$$

where all the masses are expressed in the tree level form, A is the photon field and $W^\pm$ and Z^0 are the weak interaction mediators. It has to be noticed that of the four degrees of freedom of Φ only one is left as a dynamical field, the neutral scalar field H , whereas the others have been *eaten up* by the weak gauge bosons, now massive, to satisfy the need of an additional (longitudinal) polarization.

1.3 Precision test of the ElectroWeak sector

Upon writing the most general Lagrangian without neutrinos, one finds that the dynamics depend on 19 parameters, whose numerical values are established by experiment. They can be summarized as: 9 fermion masses or Yukawa couplings, 4 CKM matrix parameters, 2 QCD parameters (coupling and CP angle), 2 electroweak gauge couplings and 2 Higgs potential parameters. With neutrinos, 7 more parameters are needed, 3 masses and 4 PMNS matrix parameters, for a total of 26 parameters. On one hand, the neutrino parameter values are still uncertain, on the other hand, the other 19 parameters of the SM, after the discovery of the Higgs boson by the CMS and ATLAS experiments [9], have been experimentally well constrained.

At the same time, a consistent progress has been made in the theoretical calculation of observables which depends on the measured values of the SM parameters. It is then possible to exploit the predictive power of the theory to set more stringent limits to

1.3. PRECISION TEST OF THE ELECTROWEAK SECTOR

know observables and test the SM for internal consistency. This comparison can be realized through precision measurement and global fit, such as electroweak fit. Using accurate measurements and calculations, this procedure returns predictions for key observables in the SM whose precision can even exceed the direct measurement, as in the case of the W mass.

1.3.1 Higher order prediction of W mass

The expression at tree level for the W mass given in eq. 1.2 can be rewritten in terms of the fine structure constant $\alpha = \frac{1}{4\pi} \frac{g^2 g'^2}{g^2 + g'^2}$, the Fermi constant $G_F = \frac{1}{v^2 \sqrt{2}}$ and the weak angle $\theta_W = \tan^{-1} \frac{g'}{g}$, in the following form :

$$M_W^2 = \frac{\pi \alpha}{\sqrt{2} G_F \sin^2 \theta_W}$$

It is particularly convenient to use these parameters because very well experimentally measured: α and G_F are well measured in the low energy regimes of, respectively, electron gyromagnetic ratio ($g - 2$) and muon lifetime; θ_W is precisely extracted from measurements at the Z -pole ($Q \sim 91 \text{ GeV}$). Of course, it must be taken into account that for the first two constants the running from the measurement energy to the electroweak scale is sizable. This reduces the effectiveness of the extraordinary precision reachable for low energy measurements.

If we consider the electroweak corrections, i.e. radiative corrections to the leading order arising from electromagnetic and weak effects, the tree level mass can be corrected [10] by a factor $M_W(1 + \Delta r)$. The main contribution to Δr comes from corrections to the W propagator with loops containing doublets of weak isospin or the Higgs boson (fig. 1.2). Corrections from fermion isodoublets are proportional to the isodoublet



Figure 1.2: Most significant corrections to the W propagator at one-loop level. Left: fermion isodoublet correction. Right: correction from the emission of a Higgs boson.

mass split $|m_u^2 - m_d^2|$ and hence are mostly coming from the top-bottom pair, whereas corrections from Higgs boson loops are logarithmic with the Higgs mass.

1.3.2 The global electroweak fit

For a long time, global fits of the electroweak sector of the SM have been used to exploit measurements of electroweak precision observables at colliders and accurate theoretical predictions at multi-loop level, to constrain free parameters of the SM. Today, all fundamental SM parameters entering these fits are experimentally determined. Global fits are then used as powerful tools to assess the validity of the theory and to constrain scenarios for new physics. They providing accurate prediction which may even have smaller uncertainties with respect to the experimental values.

At the state of art, after the measurements of the Higgs properties, the electroweak free parameters are over constrained by the large number of independent experimental observations. Therefore, it is possible to use the SM theory to accurately predict relation among these observables [11]: a simultaneous fit to all the observables assuming the SM converges at a global minimum value of $\hat{\chi}_{red}^2 = 17.8/14$, which has been found to have a p-value of 0.2, and thus being non significant to reject the theory.

Besides, precise prediction of a SM parameters can be extracted from the fit by removing that parameter from the experimental inputs and then taking as estimator the value of that parameter which minimizes the fit χ^2 and its uncertainty. A table with the result for this kind of procedure for some parameters is reported in tab. 1.2.

Comparing the input and output numbers for the m_W parameter, it is clear that,

Table 1.2: Results of the electroweak global fit as reported in [11]. The second column contains the world average experimental value used as input in the fit; the third column contains the result of the fit to all the observables except for the one considered.

Parameter	Experimental value	Fit result
M_H [GeV]	125.14 ± 0.24	93^{+25}_{-21}
M_W [GeV]	80.385 ± 0.015	80.358 ± 0.008
M_Z [GeV]	91.1875 ± 0.0021	91.200 ± 0.011
m_t [GeV]	173.34 ± 0.76	$177.0^{+2.3}_{-2.4}$

differently from other parameters, the precision of experimental measurements is worse than the one of the prediction. This is one of the the main reasons for which a better measurement of the W mass is needed: improving the uncertainty on the experimental value will enhance our ability in resolving the discrepancy between data and theory.

A further important consistency test of the SM is the simultaneous indirect determination of m_t and m_W . The discrepancy between fit results and experimental values can here be properly estimated, being the top loop one of the most important corrections to the W mass.

1.3. PRECISION TEST OF THE ELECTROWEAK SECTOR

As shown in fig. 1.3, this discrepancy is not significant yet, but a shrinkage of the uncertainty band of the m_W may reveal a sizable disagreement.

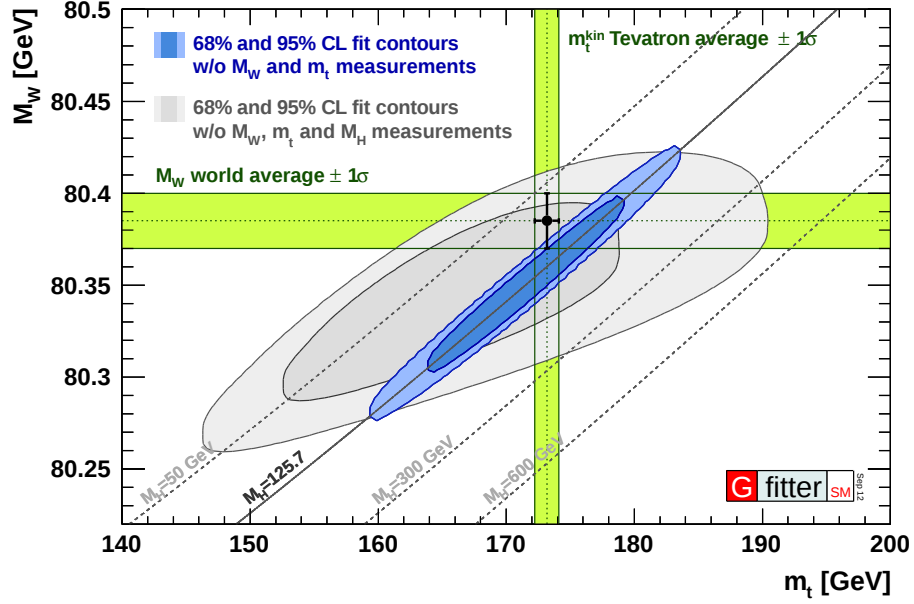


Figure 1.3: From [11]. 68% and 95% confidence level (CL) contours in the $m_t - m_W$ plane for the fit including m_H (blue) and excluding m_H (grey). In both cases the direct measurements of m_W and m_t were excluded from the fit. The values of the direct measurements are shown as green bands with their one standard deviations. The dashed diagonal lines show the SM prediction for m_W as function of m_t for different assumptions of m_H .

Chapter 2

The Compact Muon Solenoid experiment

This compilation chapter describes briefly the experimental facilities used to produce and collect the data used in the analysis presented in this thesis. The aim of this chapter is both to summarize the work done to deliver analyzable experimental data and to describe the available tools and work environment.

The first section gives an overview of the particle accelerator which has provided high energy proton beams, the Large Hadron Collider (LHC), and the characteristics of the collisions (*events*) produced in its interaction points. The second section describes briefly the Compact Muon Solenoid (CMS) detector, which collected the data used for this work. The last section deals with the CMS experiment software, summarizing the algorithms and techniques adopted to reconstruct the particles produced in the collisions. The procedures used to simulate the event production and reconstruction is described in the last part of the chapter. It constitutes a crucial part for the W mass measurement and specifically for the work presented in this thesis.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [12] is a circular particles accelerator. It is the world's largest and most powerful particle collider ever built and it takes its name from the particles (proton and ions) which are there accelerated and brought into collision. It was built between 1998 and 2008 by the European Organization for Nuclear Research (CERN), whose core mission is to provide particle beams for physics research, in collaboration with hundreds of Universities and laboratories all over the world. The

2.1. THE LARGE HADRON COLLIDER

purpose of the LHC is to produce data in order to answer fundamental open questions in physics, concerning the basic laws that govern the interactions among elementary particles: it provides experimental condition suitable to explore high energy phenomena up to the Tera Electron Volt (TeV) scale.

The LHC succeeded the Tevatron in the role of the leading and most powerful particle accelerator. The Tevatron [13] was a circular proton-antiproton ($p\bar{p}$) collider operating at the Fermilab (Batavia, IL) until 2011 and it reached a center of mass energy $\sqrt{s} \sim 2$ TeV.

The LHC is located near the CERN site on the Swiss-French border in a tunnel, the same used for the Large Electron Positron (LEP) collider [14], which is 27 km long and about 100 meters underground.

The LHC is made of two adjacent parallel beam pipes [15], where the two proton or ion beams circulate in opposite directions in ultra-high vacuum; the beams are guided inside the pipes by a strong magnetic field - up to 8.33 T - provided by 1232 superconducting dipole magnets; 392 quadrupole magnets focus and steer the beam, while radio-frequency (RF) cavities accelerate it. Approximately 96 tons of superfluid helium-4 are needed to keep the magnets, made of copper-clad niobium-titanium, at their operating temperature of 1.9 K.

Before entering the LHC, protons need to be grouped in bunches of about 10^{11} particles and accelerated to a minimum energy of 450 GeV, because of magnetic field constraints. This is achieved through a chain of 4 accelerators present at the CERN site (fig. 2.1): a linear accelerator (LINAC), a Booster ring (PSB), the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS) which directly injects the LHC. The LHC design energy of $\sqrt{s} = 14$ TeV has been dictated by the radius of the LEP tunnel and the maximum available bending magnetic field: the accelerated particles orbit radius is the limiting factor. However, since protons are composite particles, the energy available in the physical process of partons scattering is smaller than 14 TeV. For example, in a collision among valence quarks - there are 3 of them in a proton carrying in total about $1/2$ of the energy - $\frac{1}{6}\sqrt{s}$ is a rough estimate of the available energy, so the real physics reach of LHC is few TeV. Even if W and Z bosons are mainly produced with a valence quark and a sea quark, which has a softer spectrum, the available energy is far above the electroweak scale of few hundreds GeV. Thus, W and Z bosons are copiously produced.

One of the main improvement of the LHC is the significant increase in luminosity with respect to Tevatron. A high luminosity is necessary to produce a great amount of collisions for both exploring rare processes and fighting statistical uncertainties. This

interactions (called *minimum bias*) happens at every bunch crossing. That slightly reduces the number of protons in the bunches and degrades the overall beam quality. In addition, protons can interact with the residual gas inside the beam pipe and the bunches themselves spread because of intra beam interactions. Combining all these effects with operating errors and limited efficiency of correction and focalization systems, the luminosity lifetime of the beam is 15 hours, but only about 10 hours are typically used for physics.

The amount of data available for the analyses is usually quoted by the integrated luminosity, defined as $L = \int \mathcal{L} dt$ and usually measured in inverse picobarn (pb^{-1}).

Collisions between beams are induced in four points around the accelerator, each of them hosts one of the four LHC experiments: ATLAS [17], CMS [18], LHCb [19] and ALICE [20].

The work described in this thesis uses data from the proton-proton (pp) collisions recorded by the CMS experiment (fig. 2.2) and therefore in the following sections a detailed overview of this experiment will be given.

The LHC started the physics run in spring 2010 - this period is called Run 1 - at a center of mass energy of 7 TeV, and CMS collected about an integrated luminosity of 6 fb^{-1} . In 2012 the center of mass energy has been increased to 8 TeV, and CMS collected about 22 fb^{-1} . At the beginning of 2013 the LHC has been shut down to prepare the Run 2, at increased center of mass energy and luminosity.

Since the LHC reactivation in early 2015, CMS has collected about 50 fb^{-1} at a center of mass energy of 13 TeV.

LHC Run 2 will last until the end of 2018, when a long shut down of two years will allow upgrades in preparation of the Run 3 starting in 2021, which will double the integrated luminosity. In 2023, LHC will stop for a major upgrade: the High luminosity-LHC (HL-LHC [22]), which has a project luminosity a factor 10 higher than LHC.

2.1.1 Proton-Proton collision physics

This subsection briefly discusses general physics concepts of proton-proton collisions happening at LHC. At very low transferred energies (q^2), protons behave like point-like particles, but for q^2 above $\sim 10 \text{ GeV}$ a radically different picture appears. In this regime, proton can be treated as a bunch of loosely bound point-like particles called *partons*, each of them bringing a certain fraction x of the proton energy. Therefore, the production cross section of a given final state from a pp collision is the sum of the cross sections of all possible interactions among partons that can produce the final state,

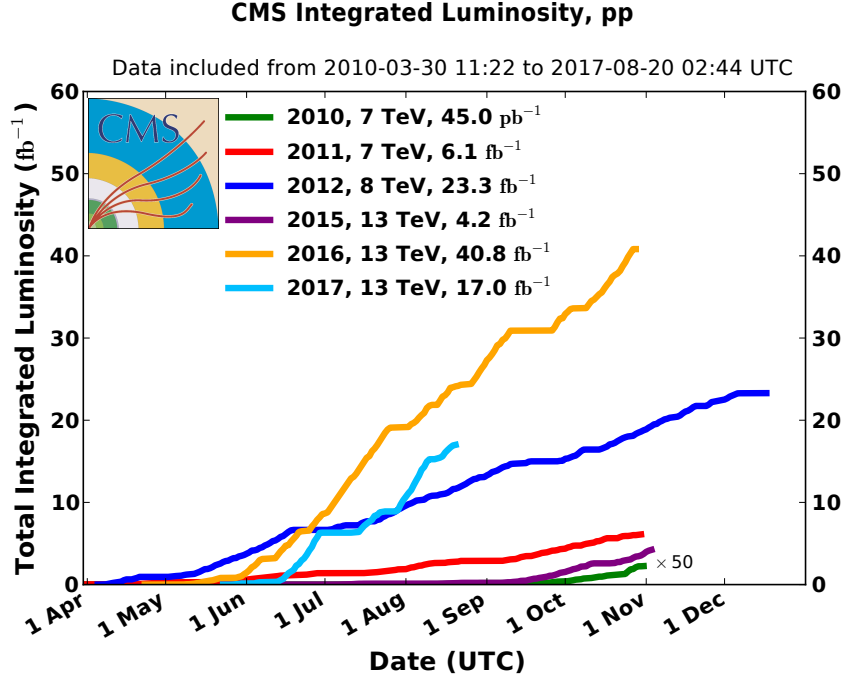


Figure 2.2: Cumulative luminosity versus day delivered to CMS during stable beams and for p - p collisions. This is shown for 2010 (green), 2011 (red), 2012 (blue), 2015 (purple), 2016 (orange) and 2017 (light blue) data-taking. Figure from [21].

weighted for their probability:

$$\sigma(pp \rightarrow X|s) = \sum_{i,j} \int dx_i dx_j f_i(x_i) f_j(x_j) \hat{\sigma}(i + j \rightarrow X | \hat{s} = x_i x_j s)$$

where i and j run on the partons of the two protons, X is the final state produced with cross section $\hat{\sigma}$, f_i is the probability density function for the parton i to have a certain fraction of the proton energy x_i , and $\hat{s}$ is the partonic center of mass energy.

These probability functions are known as Parton Distribution Functions (PDFs) and have to be determined experimentally. The knowledge of the PDFs is today one of the major theoretical systematics affecting the LHC precision measurements.

Along with main interaction, called *hard scattering*, the event is accompanied by several other processes happening between the other partons in the protons. Given the nature of the strong interaction, initial and final state radiation of gluons is very common and the soft scattering (i.e. at low q^2) of the remnants of the protons must always be taken into account. All these interactions together are designated underlying event. Due to color confinement, produced gluons can not be present in the final state and must convert into quarks and anti-quarks, which must combine among them to

2.2. THE CMS DETECTOR

form *colorless* composite particles, the hadrons. This process, known as *hadronization*, is not calculable since it involves non-perturbative QCD. However, phenomenological models, with parameters tuned to reproduce the experimental results, can be used to produce Monte Carlo simulations.

Finally, when two bunches cross each other and two protons have an interesting interaction, it is very likely that other pairs of protons from the same bunches have a soft QCD scattering. This happens because the cross section for a soft QCD process in a proton-proton collision is several order of magnitude larger than the process of interests (e.g. ~ 100 mb of QCD interactions versus ~ 100 nb for electroweak processes). At the CMS interaction point, the number of these additional non-interesting *pp* collisions, usually called *pile-up* (PU), varies in the different runs. However, the number of PU interactions per bunch crossing has always been of order 10: it increases with luminosity reaching the value of about 30 in the highest luminosity runs.

Pile-up produces a large number of soft particles which have not to be confused with the particles from the hard scattering and subsequent hadronization. Unfortunately, the presence of pile-up is the direct consequence of the trade off between the requirements of high luminosity and the experimental ability to reconstruct interesting events.

2.2 The CMS detector

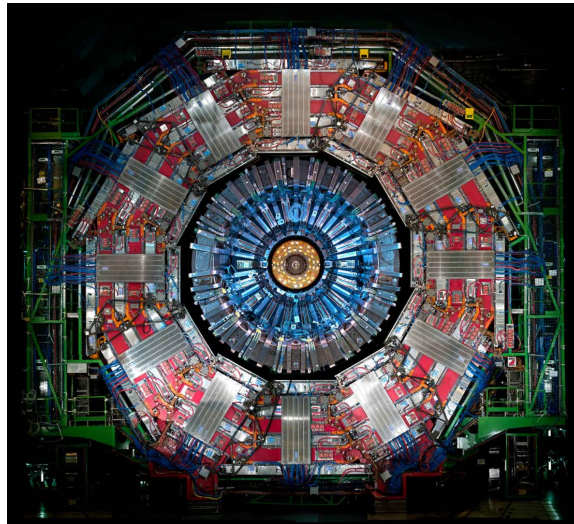


Figure 2.3: Picture of CMS in its location at point 5 of LHC. This section of the detector shows in full the 15 m of the outer diameter. (Image: Michael Hoch/Maximilien Brice)

The Compact Muon Solenoid (CMS) experiment is one of two large general-purpose

2.2. THE CMS DETECTOR

particle physics detectors built at LHC. The goal of the CMS experiment is to investigate a wide range of physics, including the measurement of the Higgs boson properties and of the other SM particles, searches for new physics as dark matter or Supersymmetry and measurements QCD properties in heavy ions collision.

The CMS detector [23] has a cylindrical shape, centered at the nominal point where the LHC beams collide. It consists of a central cylindrical part called barrel and two external parts called endcaps, placed as basis of the cylinder. CMS impressive dimensions are: about 20 m length overall and about 7 m of outer radius. It is equipped with 3.8 T superconducting solenoidal coil coaxial with the detector and the beams. A picture of the detector as well as an overview of the CMS subdetectors is present in fig. 2.3 and fig. 2.4.

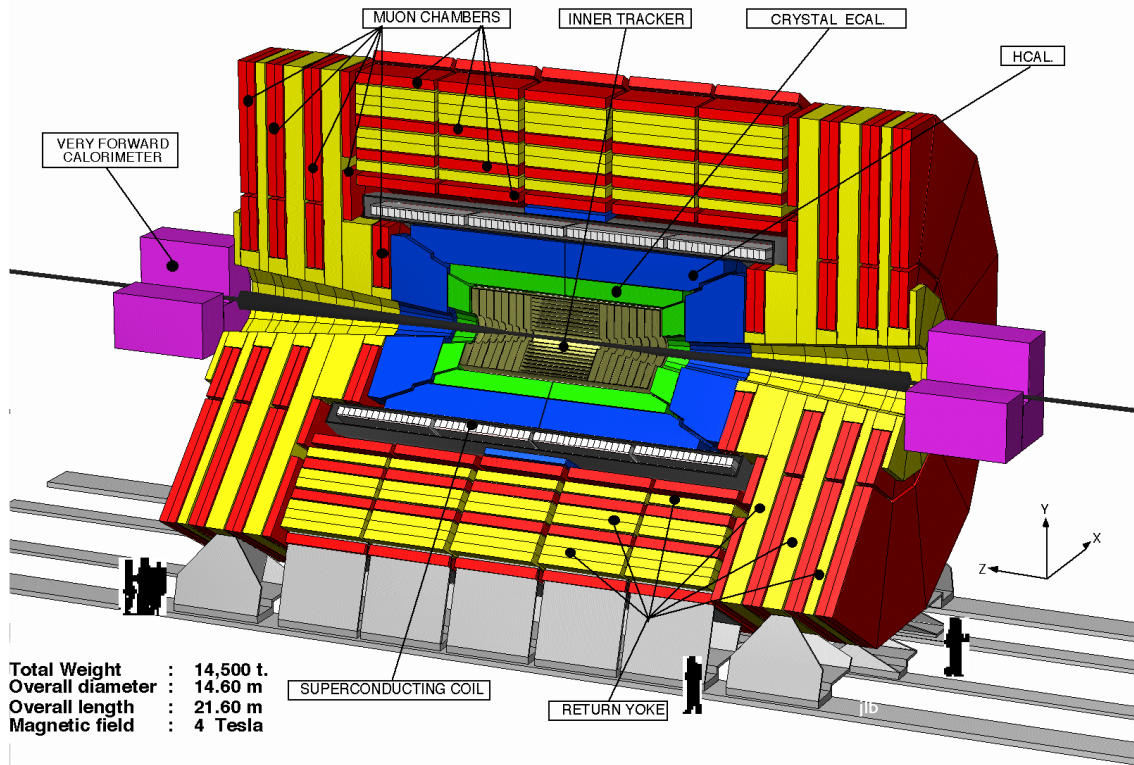


Figure 2.4: Longitudinal view of the CMS detector. Picture from <http://www-collider.physics.ucla.edu/cms/>.

The standard reference frame used in CMS is a right-handed Cartesian coordinate system with its origin in the geometrical center of the solenoid. The $\hat{x}$ axis points to the center of the LHC ring, the $\hat{y}$ axis points upwards, and the $\hat{z}$ axis points towards the beam line. A cylindrical coordinate system is more often used, described by the $r = \sqrt{x^2 + y^2}$ coordinate pointing from the axis of the cylinder outwards, and two angles

defined by $\tan(\phi) = y/x$ and $\tan(\theta) = r/z$. Instead of the angle θ , the pseudorapidity,

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

is more often used, since it is additive under boosts along the $\hat{z}$ axis and corresponds to the rapidity for mass-less particles. The rapidity is defined as

$$Y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right).$$

where E is the particle energy and p_z the particle momentum along the beam axis. The plane identified by the $\hat{x}$ and $\hat{y}$ directions is referred to as *transverse plane*.

2.2.1 The layers structure

Subdetectors are organized in a cylindrical nested layer structure (fig. 2.5) which, starting from the interaction point and moving outward, includes the silicon tracker, the electromagnetic calorimeter (ECAL) and the hadron calorimeter (HCAL) before reaching the coil. Outside the coil, the muon detection system is embedded in the iron yoke of the magnet. In the barrel, layers are cylindrical surfaces coaxial with the beam whereas in the endcaps they are disks perpendicular to the beam.

Tracker The inner most subdetector is the tracker. It is made of silicon pixels and strips which record a signal proportional to the energy lost by the charged particles passing through them. Each channel of the readout electronics is used as flag for a silicon region (strip or pixel) and the registered signal is converted to a digital logic indicating if that particular area has been hit. For this reason, the signal from the tracker modules is often referred as *hit*.

The tracker typical resolution of the order of $10 - 50 \mu\text{m}$, allows, together with the bending power of the coil, to measure the tracks transverse momentum with percent precision.

The tracker acceptance is limited at the value of $|\eta| < 2.5$, where the more forward endcaps modules are placed.

ECAL The electromagnetic calorimeter is situated outside the tracker. It is a homogeneous calorimeter, segmented only in the $\eta - \phi$ plane, designed to absorb particles like electrons and photons. ECAL is composed by scintillating PbWO_4 crystals of ap-

2.2. THE CMS DETECTOR

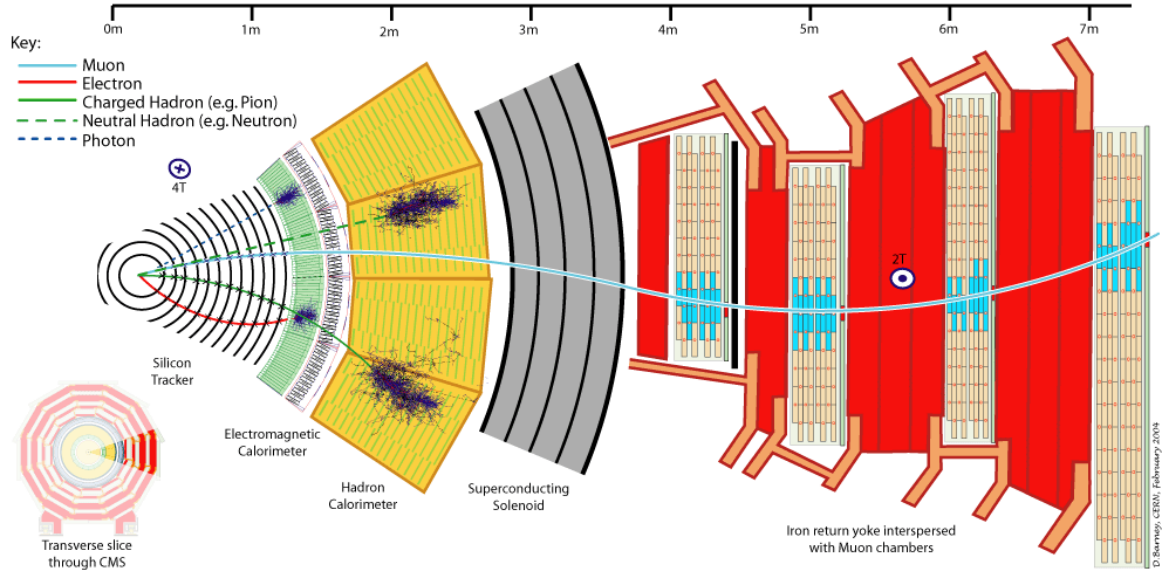


Figure 2.5: A transverse slice of the CMS detector, with the qualitative experimental signatures of the different particles. Picture from [24].

proximate dimensions $3 \text{ cm} \times 3 \text{ cm} \times 20 \text{ cm}$, with an excellent resolution for energies ranging from 1 GeV to the TeV range. Its coverage is up to $|\eta| = 3$. In the endcaps a preshower detector (lead and silicon strips) is used to improve the π^0 rejection and to help identifying the interaction vertex.

ECAL energy measurement resolution has a dependence from energy itself but typical values are of the order of 2%.

HCAL and HF CMS measures hadronic energy through a sampling calorimeter made of layers of brass and scintillator plates. HCAL can detect hadrons up to $|\eta| = 3$, so an additional calorimeter (HF) in quartz and scintillating fibers is installed to extend the acceptance to $|\eta| = 5$. Its thickness in the barrel amounts to 1.2 m and is limited by the requirement to fit inside the coil.

The hadrons energy resolution obtained using this detector is about 10-20% for high energy jets.

Muon system The muon tracking system is the outermost subdetector and is made by drift tubes (barrel) and cathode strip chambers (endcaps) inside the return yoke of the magnet. Muons with energies below some few hundred GeV lose energy in matter only through ionization, so they can easily traverse all the CMS detector with a energy loss of a few GeV.

Muons are the only particle surviving the whole detector. So, they are the cleanest signals available at LHC and this system is crucial for a reliable identification.

2.2.2 Trigger

The amount of data produced per second from proton interaction is much higher than the available bandwidth of the storing chain. A trigger system, which makes an on-line selection of the data to be kept on the physics content basis, is then crucial. The CMS triggering system is divided in two steps. The first one, called Level-1 Trigger (L1), is implemented in hardware and exploits the information from calorimeters and muon chambers. This first step is completed in about $1\ \mu\text{s}$, and reduces the event rate from 40 MHz to 100 kHz. An event accepted by L1 is passed to the High-Level Trigger (HLT) which, through a simplified version of the software used for offline analyses, partially reconstructs the event and takes decisions based on high level physical quantities. The HLT reduces the event rate further down to a few 100 Hz.

2.3 Physics object reconstruction

Once the trigger gives the green light to store the event on disk, a huge amount of offline work is done to give physical meaning to all the registered raw electronic information from the detector. This crucial phase is called reconstruction and is of course driven by the expected behavior of particles passing through the detector.

Produced in the beams interaction region, particles first enter the tracker, whose information is used to reconstruct charged-particle trajectories (tracks) and origins (vertexes). This happens through a fitting procedure, called tracking, to the signals (hits) recorded in the sensitive layers. Practically, it requires to find the parameters of the track which minimize the distance from the measured points fixed by the hits. During this step, a sizable part of the computing time is used to perform the pattern recognition (i.e. group together the hits coming from the same particle) and noise rejection.

To fully exploit the performances of the CMS tracker, a very precise alignment proce-

2.3. PHYSICS OBJECT RECONSTRUCTION

ture is needed: the position of each module, strip and pixel has to be known with a much better accuracy than the typical module resolution.

Since the tracker is immersed in a magnetic field parallel to the detector axis, trajectories are bent in the transverse plane and that allows the electric charges and momenta of charged particles to be measured.

After the tracker, particles moving outward pass through the ECAL. There, electrons and photons are absorbed and the corresponding electromagnetic showers are detected as clusters of energy recorded in neighboring cells. From that, the energy and the direction of absorbed particles are determined. Charged and neutral hadrons may initiate a hadronic shower in the ECAL as well, but they are mainly absorbed in the hadron calorimeter HCAL. The corresponding clusters are used to estimate their energies and directions. Finally, muons and neutrinos traverse the calorimeters with little or no interactions. While muons produce additional hits in muon system and can be easily identified, neutrinos escape undetected. The only way to partially recover the information from the neutrinos is using the conservation laws; it is possible to deduce their momentum by the sum of all the other particles momenta in the event. Ideally, in each vertex the sum of all momenta should be zero, given the symmetrical proton initial state, and if the sum of measured particles does not close to zero the remaining momentum can be assigned to the undetected ones. Unfortunately, given the proton collisions phenomenology, the kinematics can be closed only in the transverse plane and the computed quantity is called missing transverse energy, defined as:

$$\vec{\cancel{E}}_T = 0 - \sum_i \vec{p}_T^{(i)}$$

where $\vec{p}_T = \vec{p} - (\vec{p} \cdot \hat{z})\hat{z}$ is the transverse momentum and the index i runs on all the produced visible particles.

2.3.1 Particle Flow algorithm

A significantly improved event description can be achieved by correlating the basic elements from all detector layers (tracks and clusters) to identify each final-state particle. It is also important to combine the corresponding measurements to reconstruct the particle properties on the basis of this identification. This approach, called particle-flow (PF) reconstruction [25], requires, in order to have sizable improvements on performances, a highly-segmented tracker, a fine-grained electromagnetic calorimeter, a hermetic hadron calorimeter, a strong magnetic field and an excellent muon spectrom-

eter. The CMS apparatus has these properties and it is therefore the perfect place to apply this technique.

The PF algorithm proceeds in four steps. First, a bottom up approach builds PF elements from detector signal and, secondly, groups them into PF blocks. Then, PF candidate particles are separated within a given PF block and, finally, a post processing is performed, in order to reject or re-identify particles obviously fake or misidentified. The first step is clustering detector cells in sets called PF elements: calorimeter clusters (preshower, ECAL and HCAL) and charged particle tracks (tracker and, if matching, muon system). The second step is essentially a linking algorithm that connects the information from different sub-detectors. After a sorting in the $\phi - \eta$ plane, nearest neighbors pairs of PF elements are considered and a distance, aimed at quantifying the quality of the link, is defined. The specific conditions required to link two elements depend on their nature. For example, a track and a calorimeter cluster are linked if the extrapolation of the track ends is within the cluster area. At the end of this step, PF blocks are produced as sets of PF elements. Different PF blocks may be associated each other either by a direct link or by an indirect link through common elements. The third step, performed in each PF block, is the identification and reconstruction sequence which proceeds in the following order. First, muon candidates are identified and reconstructed and the corresponding PF elements are removed from the PF block. The electron identification and reconstruction follows, with the aim of collecting the energy of all bremsstrahlung photons; energetic and isolated photons, converted or unconverted, are identified in the same step. The corresponding tracks and ECAL or preshower clusters are excluded from further consideration. The remaining elements in the block are then subjected to a cross-identification of charged hadrons, neutral hadrons, and photons, arising from parton fragmentation, hadronization, and decays in jets.

After the post-processing, at the end of this procedure, all information is expressed in form of PF candidates with one of the following label: e , μ , γ , $\pi^\pm$ and k_0 .

2.4 Simulation

Simulation of physical events and detector effect is one of the most used tool in modern analysis. Given the good level of description that simulation programs have reached, they provide a clean and fully controlled environment in which developing new techniques and even understanding the physics underlying real observation. On top of this, simulations are the only way to compare data and predictions in the harsh environment

2.4. SIMULATION

of hadron colliders.

The CMS collaboration has its own dedicated simulation group which provides software for the event generation, detector simulation and signal digitization with high trustworthiness to the real performances.

Simulation of physical processes proceeds through three steps: first, the physics of the hard scattering (i.e. interesting interaction) is simulated through a Monte Carlo program (i.e. POWHEG [26], MadGraph [27]) and saved in some standardized format; next, a second Monte Carlo program (i.e. PYTHIA [28]) takes care of the hadronic environment, colors lines joining and hadronization; finally, events are passed through the detector simulation (based on GEANT4 [29]), which is also in charge of simulating particle decays, happening outside a small region of few microns which is instead already been treated by the previous MC. Interactions with the whole detector, readout electronics and trigger process are also simulated in the last step.

The final product of this procedure is thus a file with information content and format equivalent to real stored data, except for additional MC truth information such as undetected particles and true generated quantities. The MC produced in this way is then passed to the reconstruction chain and can be used in the same way of real data.

Chapter 3

The W boson mass measurement

The W boson is the charged carrier of the ElectroWeak interactions and, as discussed in 1.3, the measurement of its properties is one of the crucial test to prove the validity of the SM. Before discussing in details my contribution to the W mass measurement, this compilation chapter presents an overview of the related literature.

Starting from the W discovery at CERN and the first precision measurements of its mass at e^+e^- collider, the core sections will discuss the approach to these measurements at modern hadron colliders. First, it is summarized the world average leading measurement made at Tevatron and, then, the efforts made at LHC by ATLAS and CMS to further improve the current precision.

At the state of art [30], as shown in fig. 3.1, the world average of $m_W = 80385 \pm 15$ MeV is largely dominated by the Tevatron combination which has alone a relative uncertainty of $\Delta M/M \sim 2 \cdot 10^{-4}$ (19 MeV). Recently the ATLAS collaboration has submitted for publication a measurement of W mass made with LHC data [31] but it has not been included in the world average yet.

According to the SM, the W boson directly couples to each possible quark-antiquark pair made of one *up-like* and one *down-like* quark. This coupling is governed by the CKM matrix [2], for a total measured branching fraction into hadrons of about 67%. The remaining total width is almost equally divided among the three possible pairs of a charged lepton and its associated neutrino ($\tau\nu_\tau$, $\mu\nu_\mu$ and $e\nu_e$), each having a branching fraction of about 10%.

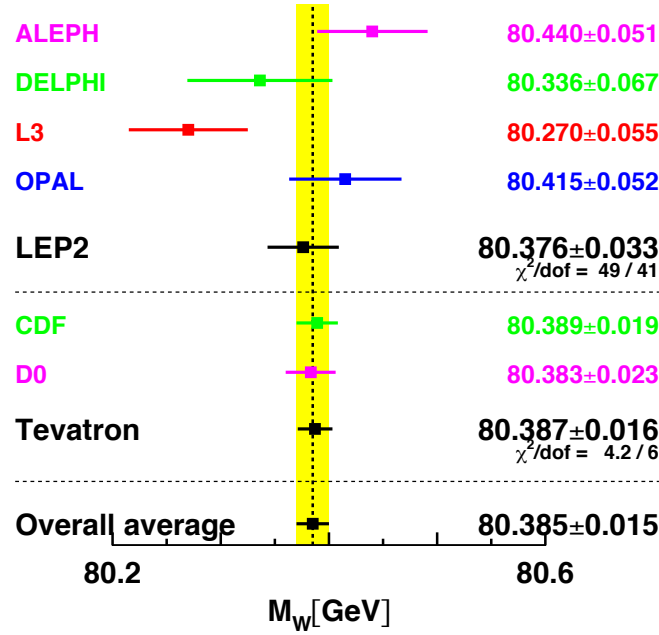


Figure 3.1: Results and world average from the past W mass measurements from [30].

3.1 From discovery to LEP

In the 60s a fully consistent theory of electroweak interactions [32] was already existing. Nevertheless, it was not before the 70s that an experimental proof of this theory was obtained. In June 1976 it was proposed to modify the CERN SPS accelerator [33], from a one-beam accelerator into a two-beam $p\bar{p}$ collider. The goal was to reach the energy of 450 GeV and explore the production of W and Z bosons. The search lead was very fruitful and, on the 20th January 1983, Carlo Rubbia announced, in a packed CERN auditorium, the first experimental observation of a signature compatible with the production of a W boson.

The official paper [34] reported the observation of six events from the UA1 detector characterized by an isolated electron, with large transverse energy, and the presence of missing energy. These events were interpreted as a W boson production through a $q\bar{q}$ annihilation, and its subsequent decay into a lepton pair, being the process $q\bar{q} \rightarrow W^\pm \rightarrow e^\pm \bar{\nu}_e^{(-)}$. Assuming that all the missing transverse energy was coming from a single neutrino $\vec{E}_T = \vec{p}_T^{(\nu)}$, a limit of $m_W > 73$ GeV (90% C.L.) was set fitting the transverse mass variable;

$$m_T^2 = 2p_T^{(e)}p_T^{(\nu)}(1 - \cos \Delta\phi_{e-\nu}) \quad (3.1)$$

3.1. FROM DISCOVERY TO LEP

where $p_T^{(\nu)}$ ($p_T^{(e)}$) is the magnitude of the transverse momentum of the neutrino (electron) and $\Delta\phi_{e-\nu}$ is the azimuthal angle between the electron and the neutrino in the transverse plane. Furthermore, theoretical guidelines and further fit to data with additional reasonable assumptions were predicting the W mass to be around 80 GeV with an uncertainty of about 5 GeV.

The Z boson discovery followed soon after [35].

After the electroweak vector bosons discovery at the SPS, the CERN cutting edge research was reoriented to study the properties of these particles with a new lepton collider: the LEP. There, W bosons were mainly produced in pairs through the reaction $e^+e^- \rightarrow W^+W^-$ (fig. 3.2).

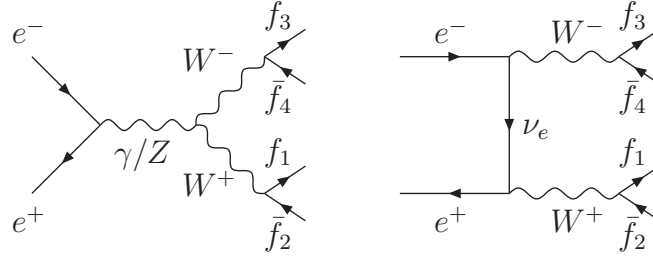


Figure 3.2: Main tree level W boson production diagrams at e^+e^- collider.

There are two main methods to measure m_W at lepton colliders. The first one exploits the fact that the W^+W^- production cross section σ_{WW} is sensitive to m_W in the threshold region, $\sqrt{s} \sim 2m_W$, and can thus be used to measure it. Fig. 3.3 shows the W pair production cross section as a function of the center of mass energy for various values of m_W . Near the energy threshold, the three curves, representing different m_W hypothesis, are separate enough to allow for a determination of m_W from σ_{WW} . From 1996 to 2000, LEP ran at center-of-mass energy above the WW production threshold. With a subset of the data collected of $\sim 10 \text{ pb}^{-1}$ at $\sqrt{s} = 161 \text{ GeV}$, a first measurement of the m_W has been delivered using the scan of σ_{WW} :

$$m_W = 80400 \pm 200 \text{ (stat)} \pm 70 \text{ (sys)} \pm 30 \text{ (} E_{beam} \text{)} \text{ MeV.}$$

The statistical uncertainty is dominant, whereas the systematic uncertainty is mainly driven by the modeling of fragmentation and hadronization. The latter factor has a large effect on the fully hadronic decay channel and contributes to the final uncertainty for about 50 MeV.

The second method to measure m_W consists of the fitting of the reconstructed invariant mass distribution line-shape. In the much cleaner leptonic environment, with

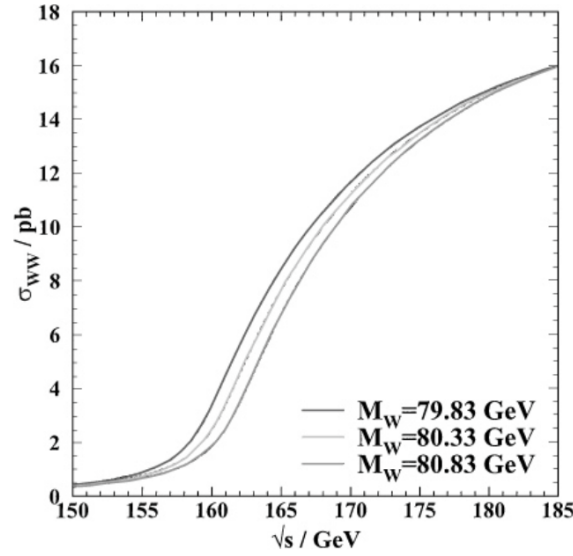


Figure 3.3: Cross section for the process $e^+e^- \rightarrow W^+W^-$ as a function of the center of mass energy for various m_W hypothesis [36].

beams made of point-like weakly interactive particles, it is possible to have a event-by-event estimation of the invariant mass in $WW \rightarrow q\bar{q}q\bar{q}$ decays. In addition, also in $WW \rightarrow q\bar{q}l\nu$ events it is possible to completely close the 3D kinematics and measure the invariant mass.

LEP published its final measurement [36] with the full dataset of about 700 pb^{-1} , using a fit to the invariant mass shape as the most sensitive observable to m_W .

In that measurement, fully leptonic decays $WW \rightarrow ll\nu\nu$ were not exploited. Given the presence of 2 neutrinos in the final state, the momentum conservation is not enough to resolve the two bosons in that events and therefore the invariant mass can not be computed. In this case, the energy spectrum of the reconstructed charged lepton could have been used, since it peaks at $m_W/2$ and thus is sensitive to m_W .

The final result is:

$$m_W = 80375 \pm 25 \text{ (stat)} \pm 22 \text{ (sys)} \text{ MeV}.$$

The most important contribution to the systematics (tab. 3.1) is given by the modeling of the hadronization and fragmentation of quarks together with the color reconnection. The effect of a similar mismodeling is to change the color flow between the four final state quarks and to alter the hadronization process, causing momentum transfer between the two W 's.

3.2. W MASS MEASUREMENT IN MODERN HADRON COLLIDER

Source	Uncertainties on m_W [MeV]		
	$q\bar{q}l\nu$	$q\bar{q}q\bar{q}$	Combination
ISR/FSR	8	5	7
Hadronization	13	19	14
Detector	10	8	10
Beam energy spread	9	9	9
Color re-connection	-	35	8
Other	3	11	4
Total systematic	21	44	22
Statistical	30	40	25
Total	36	59	33

Table 3.1: Summary of the uncertainties in the combined LEP II measurement of m_W based on direct mass reconstruction in semileptonic and fully hadronic channels [36].

3.2 W mass measurement in modern hadron collider

At modern hadron colliders, the approach for a precise measurement of the W mass needs to be revised with respect to both lepton and SPS-like colliders. First of all, at lepton colliders the energy scale is given by the energy of the beams, that is known with good precision. This is not the case of hadron collider, where the energy scale is given by the detector and requires careful calibration. Secondly, the amount of data available allows a drastic reduction of statistical uncertainties and a consequent need for a deeper understanding of subtle detector and theoretical systematics. Finally, because of the crowded QCD-dominated environment of hadron colliders, with the addition of a sizable pile-up at LHC, it is necessary a particular attention in the treatment of the interesting collision identification and missing transverse energy determination.

As discussed in the following, in these measurements the use of Z decays plays a central role in facing both of the two previous points.

This section focuses on the procedure used to measure W mass at Tevatron and LHC and ends with a summary of the main experimental efforts pursued.

3.2.1 Analysis strategy

At hadron colliders W and Z bosons are mainly produced through a quark pair annihilation (tree-level process). At $p\bar{p}$ colliders, such as Tevatron, the hard scattering is mainly among $u\bar{d}$ or $\bar{u}d$ valence quarks, while at the LHC it is mainly between a u or d valence quark and sea antiquark. Smaller contribution from other flavor sea quarks are also present. The production cross section at LO requires two inputs: the amplitude

of the electroweak process and the PDFs of the proton. Since it involves colored fields, the amplitude computation can be calculated perturbatively only in the EW part. A more complicate approach is needed for the QCD corrections, especially for soft initial state radiation (ISR) which modifies the momentum of the produced bosons. The EW bosons production cross section at proton colliders is of order 100 nb, which, with the integrated luminosity for Tevatron and LHC, gives an expected number of produced events of respectively millions and hundreds of millions.

A direct measurement of the invariant mass of the decays products is nevertheless very little effective. The only events with a fully visible final state are the ones in which the W decays into a pair of quarks. There, the hadronization process of the quarks generates two jets of particles with a total momentum of order 40 GeV, half of the mass of the W . Unfortunately, these events are very hard to identify in the hadron collider environment because of the huge background. Indeed, the QCD induced proton-proton scattering produces a great quantity of di-jet events which are very similar to the W boson signal. When the W boson has a small p_T , the di-jet final state is completely submerged by the QCD background: the two jets coming from the W decay can be easily confused with the other jets produced in the bunch crossing. A slightly better situation is obtained when the W has a very large transverse momentum (about 0.5 TeV). In this case the two jets are merged together and are reconstructed as a single *fat jet*. However, a high p_T selection has a very low efficiency and the gain in the signal over background ratio is not worth. Fig. 3.4 shows the invariant mass spectrum of reconstructed fat jet with transverse momentum between 0.5 TeV and 0.6 TeV. Looking at the fitted contribution from Z (green) and W (red), it is clear that the two contribution can hardly be separated and that the invariant mass spectrum is very little sensitive to m_W . In addition, with the current techniques, it is impossible to calibrate the jet energy scale at the precision level required for a competitive W mass measurement: the percent level is hardly reached even in very particular regions of the phase space.

For the reasons stated above, the W mass measurement is usually done selecting leptonic events, where the W decays in $l\nu$, with l being an electron or a muon. These events manifest as a single high p_T lepton, which is a clear and easily recognizable signature. In this channel the background is at the percent level and the main contributions come from: Z events in which one of the leptons is lost; top quark decaying to a b quark through a W boson; and QCD events, in which a high p_T lepton from heavy flavor decay is accidentally isolated.

In the W leptonic final state it is not possible to measure event-by-event the invariant

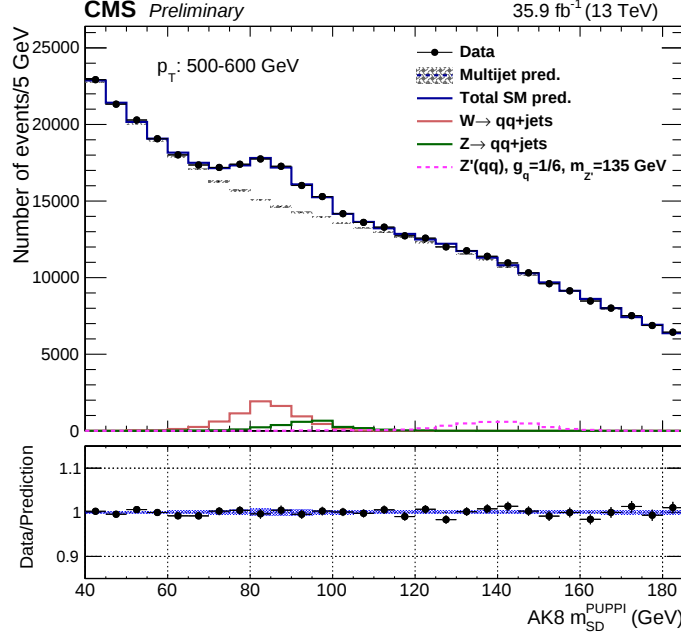


Figure 3.4: Di-jet mass distribution for the p_T category 500-600 GeV. Data are shown as the black points. The QCD background prediction, including uncertainties, is shown in the gray boxes. Contributions from the W , Z , and a hypothetical Z' signal at a mass of 135 GeV are indicated as well. Figure from [37].

mass since the kinematics can be successfully closed only in the transverse plane. So, other observables correlated with the m_W parameter have to be considered. The m_W parameter value is then extracted comparing data and predictions made with different mass value hypotheses. The leading observables usually chosen are the transverse mass (eq. 3.1 and fig. 3.5 right) and the p_T of the detected charged lepton (fig. 3.5 left). Basically, events are selected requiring a isolated high p_T lepton¹ and the presence of missing transverse energy. In each event, m_T is then computed using the full missing transverse energy, as if it all belonged to the neutrino from the W decay. This assumption has no influence as long as the same procedure is followed in data and prediction.

No explicit formula for m_T and p_μ is able to describe the observed distribution at a satisfactory level of precision, since the physics is convolved with detector effects. So, the only way in which m_W can be extracted is through a *template fit*. This means that, exploiting Monte Carlo generators and detector simulations, for a fixed hypothesis of the value of m_W a large number of events is simulated and reconstructed exactly like real data. Comparing the distributions of real data and MC simulation for several hypotheses, the measured value of m_W is determined as that of the hypothesis which

¹In the following, p_μ is used to indicate the transverse momentum of the charged lepton.

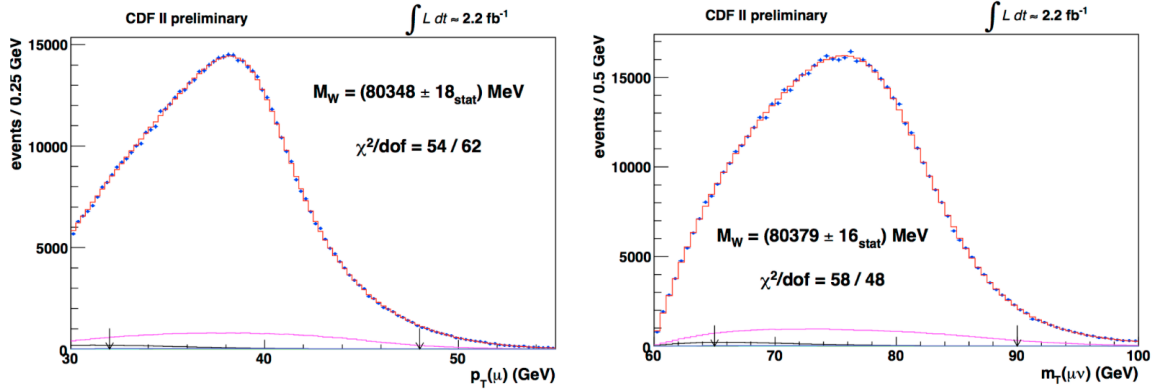


Figure 3.5: Transverse mass and muon p_T spectrum distribution fitted by the CDF collaboration [38] for the measurement of the W mass.

better describes data.

It is important to underline that in practice the generation and detector simulation happen only once and different mass hypothesis are obtained assigning a proper weight to each event, i.e. re-weighting the distributions.

Clearly this approach deeply relies on accurate simulations: eventual discrepancies between real data and templates, induced by imperfections in the simulation process, may induce systematic effects in the determination of m_W . A powerful tool to increase the agreement between data and MC and check for systematic uncertainties is the usage of Z events. When the Z boson decays into a charged lepton pair, the event is very similar to a $W \rightarrow l\nu$ where, however, the neutrino is *visible*. In addition, these events have a production mechanism similar to W events, except for a slightly higher contribution from heavy flavors, and the decay products phase space which differs for about the ratio of the two boson masses (m_W/m_Z). Nevertheless, differently from W events, Z events can be fully reconstructed and all the boson properties measured event-by-event since both leptons in the final state are detected. For this reason Z events play a crucial role in the MC calibration: boson kinematic can be compared in data and simulations, allowing for the determination of corrections aimed at increasing the agreement with data also in W events.

3.2.2 Main uncertainties

While statistical uncertainty was leading at SPS and LEP, at Tevatron the statistical uncertainties were almost of the same order than systematics. On the other hand, given the LHC luminosity delivery, it will be sub-leading for the CMS and ATLAS measurement (unless a different method is adopted).

Then, what should concern more in a high precision W mass measurement is the treatment of systematic uncertainties.

The first important component of systematic uncertainties is given by an imperfect depiction of the physical objects, as are in data, by the Monte Carlo simulation. The two main sources in this category are the lepton scale and the missing energy modeling. The lepton momentum scale needs to be equal in data and MC, up to the desired or better level of precision of the final W mass measurement. The template fitting, indeed, compares quantities which scale as p_μ . For muons, the calibration, able to reach a precision of 10^{-4} , is performed by comparing, in data and MC, the mass distribution of narrow resonances decaying into two muons, such as J/Ψ and Υ . At this point, the agreement is improved fitting corrections to the MC reconstructed muons in order to take into account possible discrepancies such as magnetic field in-homogeneity, mis-modelling of the energy loss in track reconstruction due to different detector material budget and residual tracker misalignment. A closure test on Z is usually performed given the close momentum range of its muons to the ones produced by W decays. On the other hand, the missing energy might present discrepancy between data and Monte Carlo due to approximate modeling of fragmentation and reconstruction of particles produced in the bunch crossing together with the W . Main factors are the simulation of the underlying event and of the QCD radiation, the simulation of detector effects (acceptance, efficiency, etc.) and the simulation of the pile-up. The calibration of these features deeply relies on the Z events: differences between Z data and Z MC are used as a standard candle and the corrections derived are then applied to the W . The assumption behind this procedure is that, at fixed kinematics of the boson, only the decay products are different between W and Z .

Another relevant source of systematic error is the background estimation. A detailed description of the background in single isolated lepton events at LHC can be found in [39]. Typically, in the W mass measurement a few percent fraction of the fitted histograms entries are not coming from W decays. The background not removed using kinematic cuts is usually estimated with a simulation: normalization and shapes of the predicted distribution produce sub-leading systematic.

Finally, one of the most important contribution to the W mass total uncertainties is given by the systematic on the physical modeling of the W production. Templates shapes are heavily dependent on the simulation of the W p_T , rapidity and polarization distributions. The first one is mostly influenced by simulation of the QCD initial state radiation, whereas the second and the third are dominated by the knowledge of the parton distribution functions.

A minor role is played by QED corrections which enter in the EW contributions to the propagators and in the final state radiation.

3.2.3 Recent efforts summary

Considered the topic of this thesis, a particular focus is given to the treatment of the recoil (i.e. missing energy); a breakdown of the error is reported too. For a complete reference, see the original papers [40], [31] and [41].

Tevatron results and ATLAS first attempt

The Tevatron average is dominated by the measurement of the CDF collaboration. They define the hadronic recoil as $\vec{u} = \sum_i E_i \sin(\theta_i) \hat{n}_i$, where the sum is performed over calorimeter towers, with energy E_i , polar angle θ_i , and transverse directions specified by unit vectors $\hat{n}_i$. The sum excludes towers containing energy deposits from the charged lepton(s). The transverse momentum of the neutrino is inferred as $\vec{p}_\nu = \vec{\cancel{E}}_T = -\vec{p}_\mu - \vec{u}$ and is then used to compute the transverse mass. The $\vec{u}$ resolution due to the underlying event is modeled using data triggered on inelastic $p\bar{p}$ interactions (random bunch crossings). The p_T imbalance in Z events between the reconstructed muon pair momentum ($\vec{p}_{ll}$) and $\vec{u}$ is used to tune the recoil model, which also includes the response to the initial state QCD radiation and its resolution. Cross-checks of the recoil model show good agreement between W boson data and simulation.

The more recent analysis from ATLAS essentially followed the same approach for the recoil measurement used by CDF. The missing energy is computed by the sum of the charged lepton and all calorimeter clusters, often called Calo MET. Closure tests on the recoil calibration are made, on Z events, comparing MC and data distribution of global event variables, such as $\sum E_T$, and the magnitude of the recoil perpendicular and parallel to the boson. A qualitative check consisted in the ratio plot of the two, MC and data, distributions being compatible with 1.

An interesting comparison of the uncertainties breakdown in these two measurements is discussed in [42], whose summary table is reported in tab. 3.2. For both CDF and ATLAS measurements, the uncertainties due to experimental calibrations are sub-leading with respect to the physical modeling, being the knowledge of the parton distribution function the biggest uncertainty. It also has to be noticed that the uncertainty on the recoil calibration for ATLAS transverse mass fit is sizable and makes m_T contribution to the m_W combined fit practically negligible with the respect to the p_μ one.

3.2. W MASS MEASUREMENT IN MODERN HADRON COLLIDER

	CDF		ATLAS	
	p_μ	M_T	p_μ	M_T
Statistical	16	15	7.2	9.6
Lepton Scale and Resolution	7	7	6.5	6.5
Recoil Scale and Resolution	5.5	6	2.5	13
Backgrounds	4	3.5	4.6	8.3
PDFs	9	10	9	10.2
W p_T model / QCD	9	3	8.3	9.6
Photon Radiation/EWK	4	4	5.7	3.4

Table 3.2: Break down of the uncertainties in the W mass measurement in CDF and ATLAS analyses separately for the transverse mass and lepton momentum observable fit. All figures in MeV.

CMS W -like

The CMS collaboration has not yet performed a full W mass measurement and efforts are ongoing to deliver one in the next months. Nevertheless, the most recent result is a measurement of the Z mass using the same analysis work flow that can be used on the W .

For this purpose, one muon is artificially removed from triggered $Z \rightarrow \mu\mu$ events, making them equivalent to W ones. The measurement requires essentially all the same technicalities of the W mass, except that no porting of the MC calibration needs to be done, since it is performed directly on Z events.

The experimental estimator for the missing transverse energy used in that analysis is the so called TK MET: it is built using the TK recoil instead of the calorimeter clusters sum. TK recoil is defined as the sum of all the reconstructed tracks momenta from the Z vertex, except the two muons. This estimator, that has almost no PU dependence, is however neglecting all the neutral particles, which, on average, account for about the 30% of the total energy.

The calibration of the TK MET is obtained transforming the distribution of the TK recoil in MC in order to be the same as the one in data. This transformation has been performed separately for the parallel and perpendicular recoil component to the reconstructed Z (i.e. sum of the two muons) flight direction and conditional to the reconstructed Z p_T .

A summary of the uncertainties, reported in tab. 3.3, underlines how the statistical error has been dominant since only the 7 TeV integrated luminosity has been used. The main systematic uncertainties quoted are the modeling of the QED final state radiation and the muon calibration.

Source of uncertainty	p_μ	M_T	$\cancel{E}_T$
Statistical	40	36	46
Muon calibration	14	13	14
Recoil calibration	0	9	13
PDFs	6	5	5
QED radiation	22	23	24
MC sample size	7	6	8
Other syst.	24	25	27

Table 3.3: Break down of the uncertainties in the Z mass measurement using a W -like approach performed by the CMS collaboration. Results for events tagged as positive charged W are reported separately for the three observable used. All figures are reported in MeV. Table from [41].

Chapter 4

The hadronic recoil in the W mass measurement

This work focuses its attention more on the particles produced in association with the boson than on the undetected neutrino. In this view, the concept of missing energy leaves the primary role to the hadronic recoil, being the process under study $pp \rightarrow W + X$, where X indicates the recoil and with $W \rightarrow l\nu$. Ideally, the recoil is defined as the vectorial sum of the momenta of all the particles from the pp collision, excluding the lepton. The information collected and reconstructed for each triggered event includes the lepton momentum and the momenta of all other long-lived primary or secondary particles that are produced in the acceptance of the detector. The neutrino escapes undetected.

In this framework, the transverse mass can be rewritten in a way that makes it dependent only on the muon and on the transverse momentum of the recoil which is, by definition, equal and opposite to the one of the produced boson. This chapter, the first based on original work, groups the definition and the understanding about the recoil that have been developed during the thesis. It starts explaining the recoil production mechanism, before analyzing the effects produced by the detector on its measurement. The core section goes through the impact of the recoil on the W mass analysis and treats the relevant aspects to be taken into account for the reduction of the uncertainties.

All the plots presented in this and the following chapters have been originally produced by the authors of this work and [43] for the purpose of the master thesis work.

4.1 Production Mechanism

The presence of a measurable recoil is mainly given by the fact that the W boson, and similarly the Z boson, is produced with a non zero transverse momentum. The tree level production mechanism of both bosons involves one parton from each of the incoming protons and then, at this order, no p_T is produced. The two interacting partons have essentially only longitudinal momentum and their momenta are generally unbalanced. The produced W bosons have a random boost along the beam direction. The boson transverse momentum is indeed generated by higher order corrections: the interaction among quarks and gluons, particularly gluon radiation, gives a finite p_T to the mother particles of the boson and then to the boson itself. In order to satisfy the conservation of momentum, also the remaining hadronic system has to gain the same amount of transverse momentum in the opposite direction, practically recoiling.

As shown in fig. 4.1, the W p_T spectrum peaks at about 3 GeV, with an average of about 13 GeV. It presents also sizable tails towards high value which are, however, not

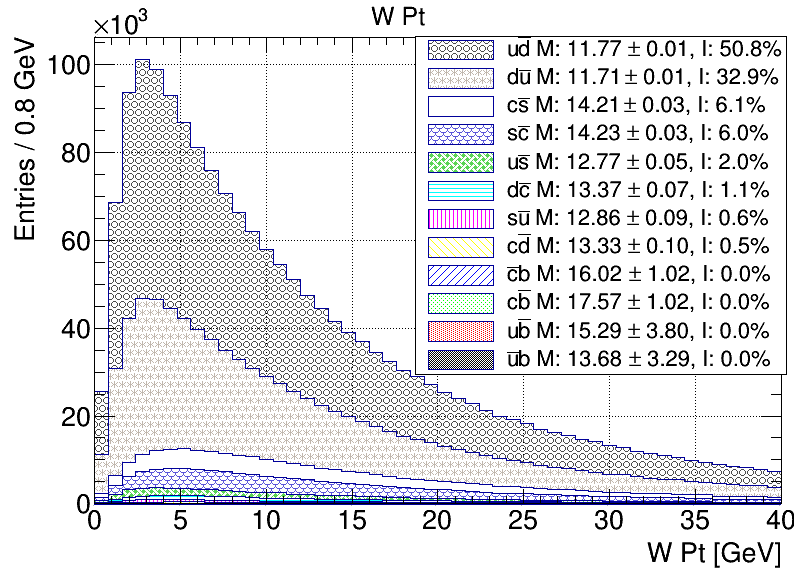


Figure 4.1: Spectrum of W boson transverse momentum from Pythia 8 standalone generation, both signs summed together. A stack histogram of the different production flavors shows the negligible impact of the b quark contribution. In the legend M stands for the mean and I for the relative fraction of each sub-sample.

relevant for the W mass since only events coming from the soft ($\lesssim 40$ GeV) part of the spectrum are used (discussed sec. 4.4). In the soft regime, and especially close to $p_T^W \approx 0$, the p_T generation process is not perturbative and resummation technique [44] is strictly required for a proper prediction.

4.1. PRODUCTION MECHANISM

The spectrum is reported in fig. 4.1 as a stack plot of the contributions of the quark pair producing the boson. Since this is a MC generator level study, the quark pair which has produced the boson is defined as the two particles which precede the W in the MC history.

In addition, the legend reports for each production mechanism the mean of the spectrum (M) and relative fraction (I), computed as the ratio between the number produced by a given quarks combination and the total number of W produced. The main contributions are from first generation quarks, $u\bar{d}$ (W^+) and $d\bar{u}$ (W^-). The relative abundance of the different two charge signs is due to the proton valence quarks composition. Second generation quark pairs contribute for the same amount of 6% in both charge signs, whereas CKM matrix off-diagonal couplings between first and second generation are at the percent level. Finally, b quark contribution is below the per mill level.

It is also interesting to notice that the p_T spectrum of W produced by heavy flavours is harder, as discussed in [45].

4.1.1 Comparison between Z and W

Fig. 4.2 shows a comparison between the p_T spectra of Z and W . The spectra are normalized to their integral and are shown for all the possible production mechanisms. The two bosons have a similar contribution in the leading component, where they are

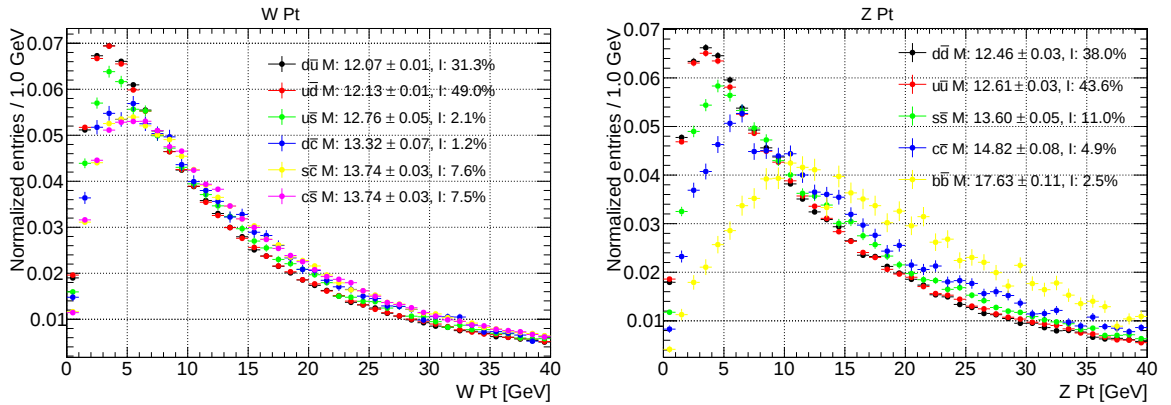


Figure 4.2: Comparison between W (left) and Z (right) transverse momentum spectra at generator level, made with Pythia 8. The distribution for the various production flavors is shown normalized to 1 for each possible pair of the quarks annihilating into the boson. I and M have the same meaning as in fig .4.1.

produced by light quarks. Differences still arise in the heavy flavor contribution, which is much more prominent for the Z boson and has a sizable harder spectrum.

As discussed in [46], a better variable to compare the two spectra is p_T/M . When this variable is used the two spectra agree better than 10%.

4.2 Definition and characterization of hadronic recoil

One of the main new conceptual points of this work is to change the focus from the missing transverse energy to the recoil. Formally, the recoil $\mathbf{h}$ is a four-vector defined as the sum of the four-vectors of all the particles from W vertex, excluding the neutrino and the charged lepton. In principle, it does not include detector effects in its definition but, given all the practical constraints, this object is almost impossible to measure. That is the reason why it is important to identify an effective experimental definition.

The recoil does not depend on the lepton variables, this fact is important to properly factorize all the systematic uncertainties, as it will be explained in the last chapter. The MET, on the contrary, does depend on the lepton.

According to the given recoil definition, the kinematic is fully closed in the 3D space,

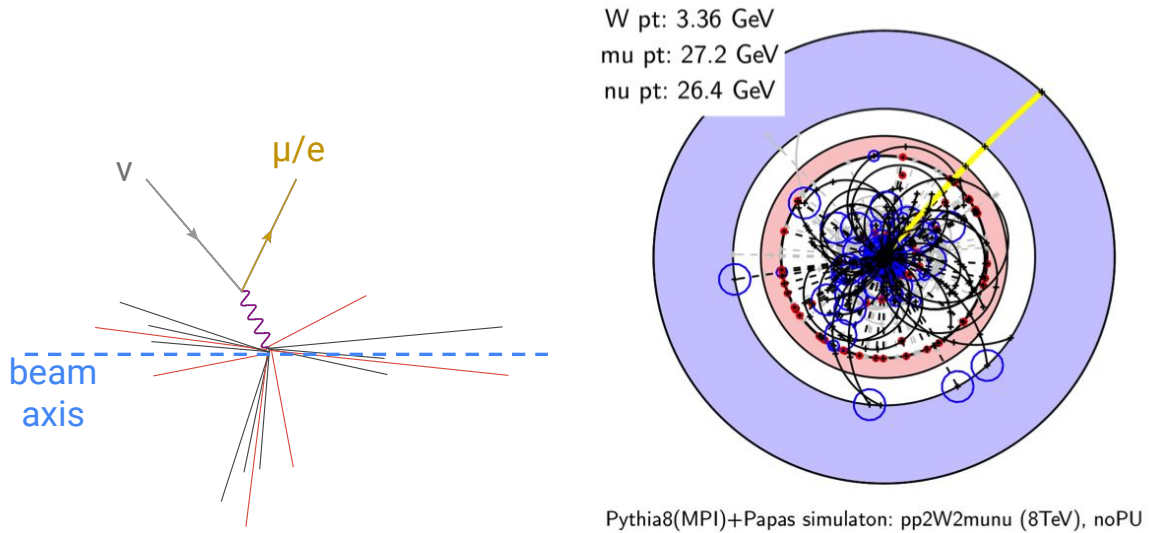


Figure 4.3: Left: Sketch of the production of a W boson from a pp collision. The boson decay products may result almost opposite to some of the recoil particles but still an important part of the momentum is lost in the forward region.

Right: Display of W production event generated with Pythia 8 and passed through a fast simulation of the CMS detector made with PAPAS.

but for the purpose of our study is enough to consider that the transverse kinematic is

closed. The transverse component ($x - y$ plane) of the recoil $\vec{h}$ satisfies:

$$\vec{h} = \sum_{i \neq \mu, \nu} \vec{p}_T^{(i)} = -\vec{p}_T^{(W)}$$

Fig. 4.3 shows the display of a simulated $W \rightarrow \mu\nu$ event: qualitatively, the underlying event is made of about 100 soft particles with energy below few GeV, whose transverse momenta summed together gives the recoil vector. This composition makes it very similar to the *minimum bias* event and, therefore, also to the particles production from *pile-up* vertexes, which happens in the same bunch crossing.

The particles entering in the recoil are far from jet regime and almost spherically distributed in the detector. This makes very hard to recognize which neutral particles belong to the W vertex: no physical handles are present to distinguish the ones coming from the primary vertex and those coming from secondary PU vertexes.

4.2.1 Detector effects in measuring the recoil

Detector effects are really dominant in measuring the recoil and only a fraction of the particles belonging to the recoil is detected. As illustrated in fig. 4.4, a very large fraction of particles is lost because under detection threshold or because outside the forward acceptance. The left plot shows how the number of particles considerably

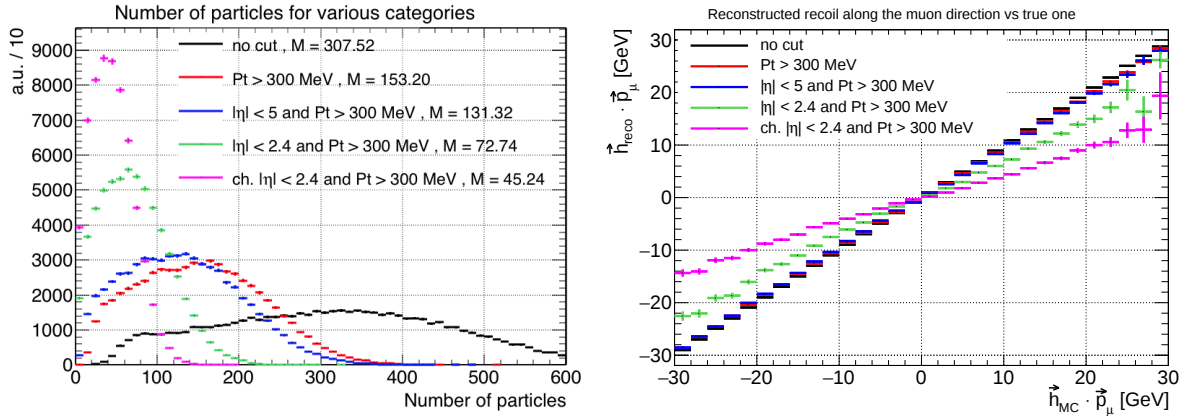


Figure 4.4: Left: Number of particles generated for a set of different requirements reported in the legend. In particular the "ch." in the pink histogram legend stands for the fact that only charged particles are included.

The mean of the distribution is also reported as M .

Right: Profile plot from the 2D distribution of the recoil projection on the muon flight direction when true value (x axis) and particle (y axis) information is used. The plot is shown for the same cuts present on the left one.

The plots report the information from a Pythia 8.2 at generator level.

drops, compared to the amount of all generated particles, when a requirement on the minimum p_T is set (qualitatively similar to detection threshold). A second major drop is then observed when also a requirement on the maximum η is added, emulating the CMS forward acceptance.

A similar conclusion can be drawn from the right plot, where it is shown the profile plot of the recoil projection along the muon direction. On the $\hat{y}$ axis is reported the value built using different particles selections and on the $\hat{x}$ axis the true MC recoil. The same cuts presented for the number of particles are compared, the main displacement from the diagonal happens when the $|\eta| < 2.4$ requirement is set.

For completeness, it is important to mention that, when $|\eta| < 2.4$ is required, the p_T cut is irrelevant: in the plots of fig. 4.4 the situation with only the $|\eta| < 2.4$ cut would have overlapped the green curve.

The effect of the p_T threshold is smaller than the forward acceptance one. On one hand, even if the number of particles drops considerably when cutting at $p_T < 300$ MeV, the recoil is still well reconstructed: black, red and blue lines are coincident in the right plot. On the other hand, forward acceptance really limits the detectable information: it is not the beam pipe ($|\eta| < 5$, blue) the most relevant limits but the tracker and calorimeters acceptance ($|\eta| < 2.4$, green). This happens because the soft and very forward particles have a small p_T .

As a consequence of the poor amount of information recorded, both angle and magnitude of the recoil generally have a resolution order of magnitudes worse than the one for compact objects, such as leptons or photons.

The angle is reconstructed with the precision worse than 1 rad and the magnitude with a relative uncertainty of order 1, as will be discussed in sec.5.3.2.

4.3 Samples definition

A non negligible part, in terms of time, of the thesis work has been spent in acquiring the technical abilities to generate simulated events and produce root *ntuples*, which are the analyzable file format of CMS data, from raw data files.

Different samples have been used in this thesis:

- Z and W MC simulations generated with a standalone version of Pythia 8.2 have been used for all the generator level studies presented up to now.
- CMS data collected in 2012 at 8 TeV: about 257 M events collected with the single lepton trigger (HLT_IsoMu24), corresponding to about 20 fb^{-1} .

4.3. SAMPLES DEFINITION

- Monte Carlo samples produced centrally by CMS: 56 M of $pp \rightarrow W(l\nu) + \text{Jets}$ and 30 M of $pp \rightarrow \gamma^*/Z(ll) + \text{Jets}$ simulated events at 8 TeV center of mass. The generation chain used Madgraph + Pythia 6 + official CMS software for the detector simulation.

In the following, these two samples will be simply recalled as W and Z MC.

4.3.1 Events selection

First of all, both muons and electrons could have been suitable for the discussions presented in this thesis but, for sake of simplicity, only the former have been considered. In some generator studies - as those described in the previous sections - explicit cuts are applied to mimic the acceptance of the detector.

From now on, the selection used on data and the fully simulated MC samples is described.

The primary vertex is defined as the vertex in a bunch crossing with the higher value of the scalar sum of the transverse energy of the associated tracks.

All the events are required to pass the HLT_IsoMu24 trigger, which requires an isolated muon with at least 24 GeV of transverse momentum.

Furthermore, dedicated cuts to select W and Z events used in the W mass analysis are applied. Only the muons with tight ID¹ and with a distance on the $\hat{z}$ axis smaller than 0.1 cm from the PV are considered.

W events - The W mass analysis is run separately for the two signs of charge and, then, two distinguished samples are defined out of the W events. An event is selected to belong to a certain W sign sample (+ or -) if there is exactly one muon satisfying the following:

- Reconstructed charge of the given sign.
- $p_T \geq 30$ GeV and $|\eta| < 2.1$, in order to avoid turn on curves of the trigger and of the acceptance.
- relative isolation smaller than 0.5. This is defined as

$$\text{relIso} = \frac{\sum_{\text{tracks from PV}} \vec{p}_T + \max(0, \sum_{\text{neutral PF cand.}} \vec{p}_T - 0.5 \cdot \sum_{\text{tracks not PV}} \vec{p}_T)}{p_T^\mu}$$

¹For more info see [47]

4.3. SAMPLES DEFINITION

This is meant to reduce the background from muons produced from QCD interactions.

In addition, the PF MET - defined as minus the vector sum of the transverse momenta of all the PF candidates of a given bunch crossing - is required to be larger than 30 GeV. The above selection is meant to require the presence of a high p_T neutrino, reducing the QCD background.

From the total W sample (56M), when the above conditions with the plus charge sign are required, about 2.1 M of events are remaining (3.7% of the total). The efficiency (selected over produced) for $W^+ \rightarrow \mu^+ \nu$ events is about 20%.

Z events - As discussed before, Z events are crucial in the W mass analysis since they can be used as W - *like* events, where the boson transverse momentum can be measured from both the recoil and the sum of the leptons reconstructed momenta. In order to maximize the analogy, it is convenient to have a selection for these events as much similar as possible to the one used for W events.

First of all, a pair of opposite sign muons with invariant mass between 70 GeV and 110 GeV is required. If more than one pair satisfies this requirement, only the one with invariant mass nearer to $M_Z = 91.19$ GeV is considered. At this point, each event can be W - *like* in two ways: removing the positive muon and having the negative one passing the W selection or vice versa. To resolve this ambiguity, two samples of Z events are defined: $Z+$ and $Z-$.

Given for example the charge sign $+$, an event is retained in the $Z+$ sample if:

- The positive muon satisfies the requirements in the bullet points for the W , but with all the GeV quantities scaled by M_Z/M_W .
- The negative muon has: $p_T > 10$ GeV, $|\eta| < 2.4$, and $\text{relIso} < 0.5$. This requirement is necessary to reduce the combinatorial background given by real Z events, where one muon is out of acceptance and is, instead, replaced by a fake or a QCD muon.
- The PF MET, computed removing the negative muon from the event, exceed $30 \text{ GeV} \cdot M_Z/M_W$.

A similar procedure can be applied for the $Z-$.

From the total Z sample (30M), when the $Z+$ conditions are required, about 1.8 M of events survive (6%).

4.3. SAMPLES DEFINITION

From the CMS data sample, 4.2M of events are remaining after the Z + selections.

Reconstructed recoil components - Once removed the muon(s) from the $W(Z)$, all other particles in the event belong either to the pile-up or to the recoil.

Charged tracks are divided into two categories, tracks from PV and tracks not from PV, according to their distance from it. Studying the distribution of this distance in simulated events without PU vertexes (fig. 4.5), it has been decided to apply the following cut to select tracks from PV (and the opposite for tracks not from PV):

$$\frac{d_x^2 + d_y^2}{0.03^2 \text{ cm}^2} + \frac{d_z^2}{0.05^2 \text{ cm}^2} \leq 1$$

where $d_{\hat{n}}$ is the distance on the $\hat{n}$ axis between the reconstructed PV and the track closest approach point to PV.

Once the charged PF candidates are grouped in the two above exclusive categories,

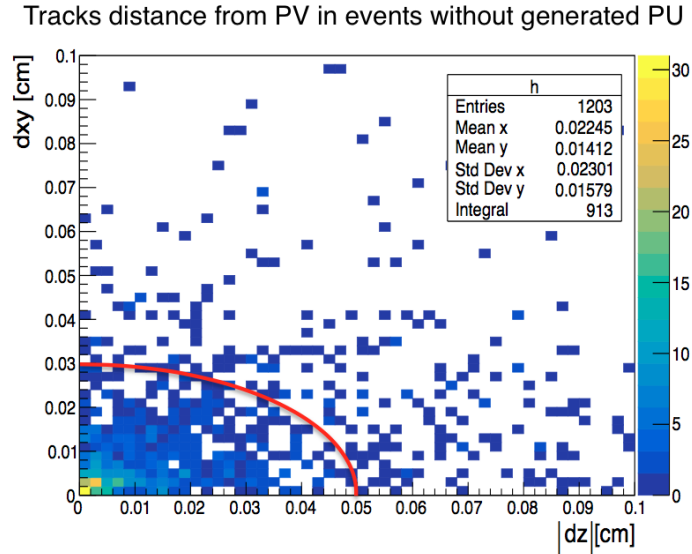


Figure 4.5: Distribution of the distance of charged tracks from the primary vertex in the transverse and longitudinal plane obtained from a subsample of the CMS simulation in events without generated pile-up vertexes. The red ellipses retains 90% of the tracks.

only neutral PF candidates are left. Since no vertex assignment is possible, they are all grouped together.

Track recoil (TK) indicates the 2D recoil vector built summing all the transverse momenta of tracks from PV and its symbol is $\vec{h}_{TK}$. The sum of the remaining tracks (i.e. tracks not from PV) is instead labeled as $\vec{h}_{npv}$ and, finally, the sum of all the

neutral PF candidates transverse momenta is called $\vec{h}_{nt}$.

The PF recoil is defined as the sum of the three above quantities:

$$\vec{h}_{PF} = \vec{h}_{TK} + \vec{h}_{npv} + \vec{h}_{nt}.$$

The reason why tracks not from PV are added in the PF recoil is that, on average, they cancel the contribution from neutral not from PV, which can not be distinguished on event-by-event bases.

It is also useful to define $\vec{h}_{ntnpv} = \vec{h}_{npv} + \vec{h}_{nt}$.

4.4 Hadronic recoil effect on transverse mass

In the W mass analysis the recoil is used in the event selection and in the calculation of the transverse mass. In the selection, the recoil is used to eliminate the events with large boson transverse momentum. However, the spectrum peaks at low transverse momentum and the fraction of events rejected by this cut have a negligible dependence on the recoil definition. In addition, given the poor resolution on the recoil, there is a long tail of events at large transverse momentum that pass the selection. Practically, the cut on the boson spectrum is smooth and applied in a flat and low-populated region.

The role of the recoil in the calculation of the transverse mass is crucial. Starting from the definition in eq. 3.1, with some algebra and using the relation in the transverse plane $\vec{p}_T^{(W)} = \vec{p}_T^{(\nu)} + \vec{p}_T^{(\mu)} = -\vec{h}$, the transverse mass formula can be rewritten as:

$$M_T^2 = 2p_\mu \left| \vec{p}_\mu + \vec{h} \right| + 2p_\mu^2 + 2\vec{p}_\mu \cdot \vec{h} \quad (4.1)$$

where it is clearly shown that the transverse mass depends only on $\vec{p}_\mu$ and $\vec{h}$.

Expanding eq. 4.1 at the first order in h/p_μ , it is obtained that

$$M_T^2 \approx 2p_\mu \left(2p_\mu + \hat{p}_\mu \cdot \vec{h} \right)$$

This formula shows that the recoil enters in the transverse mass computation scaled with p_μ . The requirement on its scale calibration is more relaxed than the muon one. Even if the recoil is relatively better measured at high transverse momentum, the region where the W p_T is soft ($\lesssim 40$ GeV) is more suitable for W mass measurement. In that region, indeed, muon p_T , typically about $m_W/2 \approx 40$ GeV, is leading with

4.4. HADRONIC RECOIL EFFECT ON TRANSVERSE MASS

the respect to hadronic recoil, which is about an order of magnitude smaller, since the W p_T spectrum peaks at 4 GeV. This allows the recoil to have a more relaxed requirements on the data/MC calibration. In particular, to have the same performances in the transverse mass, the muon, whose scale can be calibrated up to 10^{-4} , has to be calibrated about a factor $p_\mu/h \sim 10$ better than the recoil, which even in the best case scenario is hard to calibrate up to 10^{-3} .

Additional interest in the m_T first order expansion consists in the fact that only the projection of the recoil on the muon component is relevant at first order; thus, when the approximation holds, the recoil can be treated as a 1D object instead as a 2D vector. Unfortunately, a similar approach will strongly correlate the recoil and the muon, which is a downside at moment of systematic uncertainty evaluation. During the development of the work, the usage of just the first order expression for M_T has been tested and then dropped for the above reason.

In fig. 4.6 several transverse mass distributions from W^+ events are shown. They

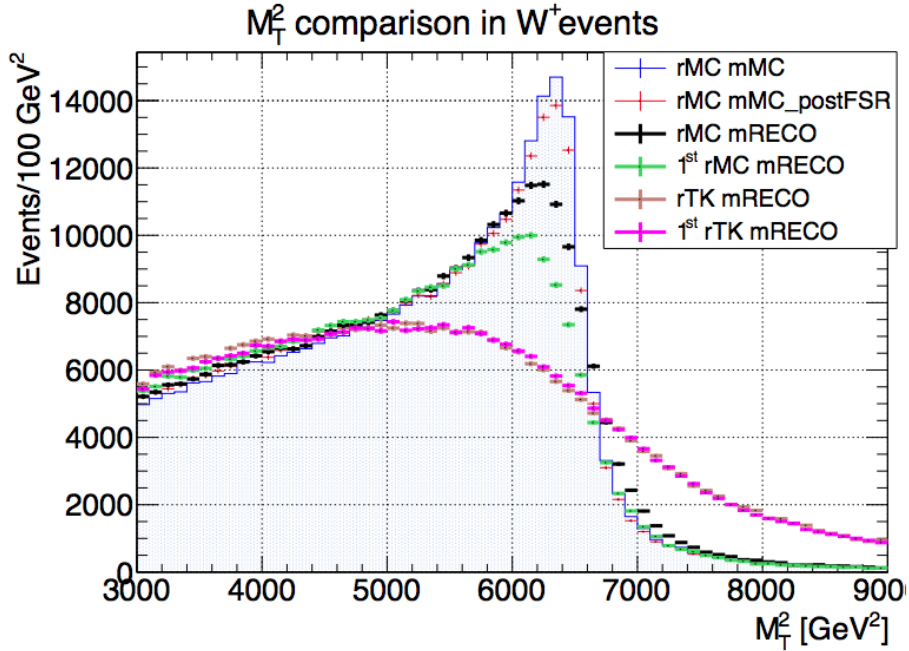


Figure 4.6: Transverse mass distributions from W MC for different kinds of variables used. For each legend line, two codes of the form $yXXX$ are reported: y stands for the kind of variable (m for muon and r for recoil) and XXX stands for the level of the variable used (MC means Monte Carlo truth, $RECO$ for reconstructed after detector simulation). The label 1^{st} is added when the first order approximation is used for M_T .

have been computed using variables at different level of generation and reconstruction. The first thing that emerges is that, even if muons are well reconstructed in CMS, the experimental effect on their detection is not negligible. Indeed, sizable differences

can be appreciated among the first top three distributions: the blue filled one is built using MC truth (i.e. the muon as it is radiated from the W); the red one is made considering the muon momentum after possible photon emission (FSR) and, finally, the black one uses the CMS reconstructed muon momentum. All these three curves have been computed using the recoil information from the MC truth, i.e. there is no experimental uncertainty on this quantity. The brown points (almost hidden by purple ones) show the distribution built using the reconstructed TK recoil: comparing it with the black ones, it is possible to gauge how much experimental effects on the recoil are dominating the shape.

Finally, green and purple points represent the distribution of the transverse mass when the first order approximation is used. Comparing green points with black ones and purple with brown, it is clear that the effect of the approximation is sizable when the recoil is well measured and negligible when the recoil is not very well measured, as it is in the real data.

4.4.1 W mass sensitivity and boson transverse momentum dependence

The precision of the W mass measurement depends on the accuracy in the measurement of $\vec{p}_\mu$ and $\vec{h}$ and on the agreement between data and simulation on the W production model. In particular, the agreement of the W transverse momentum spectrum in real data and simulation in MC, is a crucial detail. This can be appreciated from the plots of fig. 4.7. The left plot shows the transverse mass, the right one the p_μ distribution.

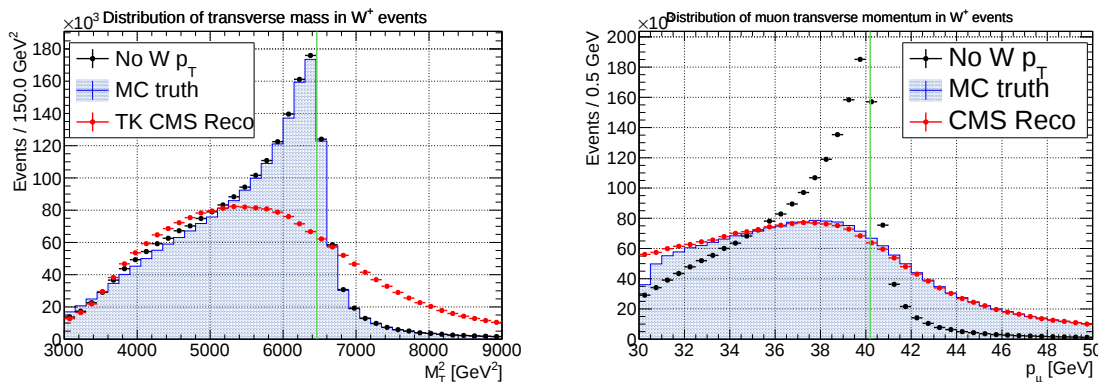


Figure 4.7: Distribution of the transverse mass (left) and muon transverse momentum (right) in W^+ from W MC. Both distributions are presented in three cases: the W is produced always with $p_T = 0$ (black), the MC truth information is used (blue) and the reconstructed information is used (red). The green line indicates the value of the W mass.

4.4. HADRONIC RECOIL EFFECT ON TRANSVERSE MASS

The black histograms are made at MC truth level, artificially removing the W p_T : in each event the muon is boosted back in the W center of mass frame and then boosted again in the lab frame, but using a transformation where the transverse component of the W momentum is set to 0.

The two black distributions have the same physics content: when $h = 0$, then $M_T^2 = 4p_\mu^2$. Another jacobian, in the phase space this time, creates the peak with the sharp edge, which is clearly correlated to the W mass (green line). This peak is due to events where the muon is emitted at $\eta = 0$ in the W rest frame (i.e. perpendicular to the beam axis as seen in the W reference frame). When this happens, the M_T coincides with the invariant mass and the muon p_T has the maximum allowed value of $M_W/2$. In both p_μ and M_T^2 distributions, the small tail at large values is due to events in the upper tail of the W Bright-Wigner distribution ($\Gamma_W = 2$ GeV). On the contrary, the large tail at small values is due to events where the muon from the W decay is emitted large $|\eta|$.

When the W p_T distribution is switched on, the two variables p_μ and M_T^2 have a different physics content and distribution. The latter is represented by the blue filled histogram in both plots. The p_μ distribution gets smeared while the transverse mass is almost unchanged. In [48] it is reported an extensive discussion about the tiny dependence of the transverse mass variable from the W p_T spectrum.

The red histograms are, finally, obtained at detector level, when acceptance and resolution are included and the TK recoil is used as a proxy of W $\vec{p}_T$. The muon p_T distribution changes only slightly, since the muon momentum resolution is very good compared to the distribution width; instead, the M_T^2 distribution is mainly modified by detector effects entering in the recoil measurement. The peak is not visible anymore.

These plots are very illustrative since they underline pros and cons of the two variables used to measure the W boson mass. The muon momentum spectrum is only mildly affected by the detector effects, however it has a strong dependence on the knowledge and simulation of the W p_T spectrum. That will result in a higher systematic uncertainty, which must be assessed performing variations of the simulated spectrum inside an error band given by independent knowledge (experiment and theory) of it. On the other hand, it is relatively easy to address detector resolution and calibration effects since a single, well measured object is involved.

On the contrary, transverse mass estimator has almost no dependence on the W p_T spectrum but it suffers from the strong dependence on the recoil: since the recoil is poorly measured, its calibration and the tuning of simulation to the data are challenging. The dependence of the transverse mass from W p_T is mitigated by the recoil,

which brings information about the W p_T on event-by-event basis; the event-by-event correlation between muon and recoil retains the peak and improves the systematics.

For the above reasons, the W mass measurement has been done fitting simultaneously the M_T and p_T spectrum. The weight of one or the other distribution on the fit final result depends on the knowledge of the W p_T spectrum used in the simulation and on the accuracy of the recoil measurement.

The bottom line is that a good W mass measurement needs a measurement of the recoil on an event-by-event basis able to retain the correlation with the muon, which is otherwise washed by the detector effect and resolution.

Finally, it is important to add that, even if M_T and M_T^2 have the same information content, M_T^2 is preferred: in fact, it is easier to be understood because the recoil enters linearly in its calculation.

4.5 A precise measurement of the recoil

A precise measurement of the recoil presents a series of technical challenges: it is a non-local object, since the particles forming the recoil cover in each event a sizable part of the detector acceptance; it involves the information from both calorimeters and tracking; it is a vector quantity and there is no uniquely better definition since the performances heavily depend on the application.

4.5.1 The recoils as a 2D object

One of the biggest difficulties in facing the recoil experimental challenge is that it is indeed a 2D object. Dealing with a 2D object not only doubles the efforts required, but also introduces ambiguities in the choice of the frame of reference. Even the concept of resolution is well defined only as the vector quantity

$$\Delta\vec{h} = \vec{h}_{\text{true}} - \vec{h}_{\text{measured}}$$

and there is not a clear 1D estimator that can be quoted as estimator of the goodness of the measurement. As shown in the following, often the magnitude $|\Delta\vec{h}|$ is of the same order of $\vec{h}_{\text{true}}$. In these cases, care has to be taken before applying concepts that are usually used in calibration of physical quantities.

In previous approaches, it has always been considered important to use estimators which have on average the same magnitude (modulus of the recoil vector) of the true

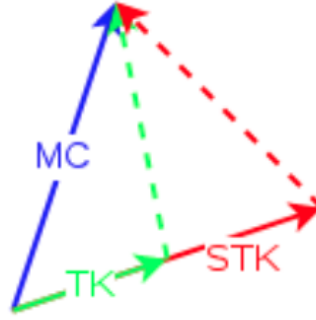


Figure 4.8: Sketch exemplifying a situation where the angular mismatch between the real quantity (MC, blue) and the reconstructed quantity (TK, green) makes the reconstructed quantity, scaled in order to have the correct magnitude, (STK, red) a worse estimator (in the $|\Delta\vec{h}|$ figure of merit) than the non-scaled one. This is often the case for the recoil measurement in the W mass events.

recoil: this is what is meant by the jargon *scale corrected*. A simple sketch (fig. 4.8) is enough to illustrate how this requirement can reduce the accuracy of the estimator when the angular direction of the recoil is badly measured as it is in the case of W mass phase space.

The problem can be split in angular and scale corrections, but always considering the resolution on the angle a focal point. The effect of getting a wrong angle is comparable to, or even larger than, getting a wrong scale. Differently from what happens for muons or electron, where the angle is very well measured and the dominant error is by far the one on the magnitude of the momentum, a precise measurement of the recoil needs to be corrected both in scale and in angle. To understand this point, a series of plot are shown in fig. 4.9. There, it is reported the transverse mass distribution in different W p_T (i.e. true recoil) regimes, the p_T grows from left to right and from top to bottom. For each regime, four different M_T are reported, according to the different component used for the calculation of the recoil: MC truth information for angle and magnitude (blue); the MC truth for the magnitude and TK reconstructed for the angle (black), MC truth for the angle and TK reconstructed for the magnitude (green), only TK reconstructed information (red). Taking the blue distribution as reference of the best available information, it is clear that, in the soft W p_T regime, scale and angle have qualitatively the same importance, since green and black plots are very near. This is due to the fact that p_μ is much larger than the recoil, which is almost negligible in the calculation of M_T . In the most populated region ($4 \text{ GeV} < W \text{ } p_T < 10 \text{ GeV}$), instead, the transverse mass computation with an algorithm based on TK gets remarkably improved when the MC truth information for the angle is inserted. On

4.5. A PRECISE MEASUREMENT OF THE RECOIL

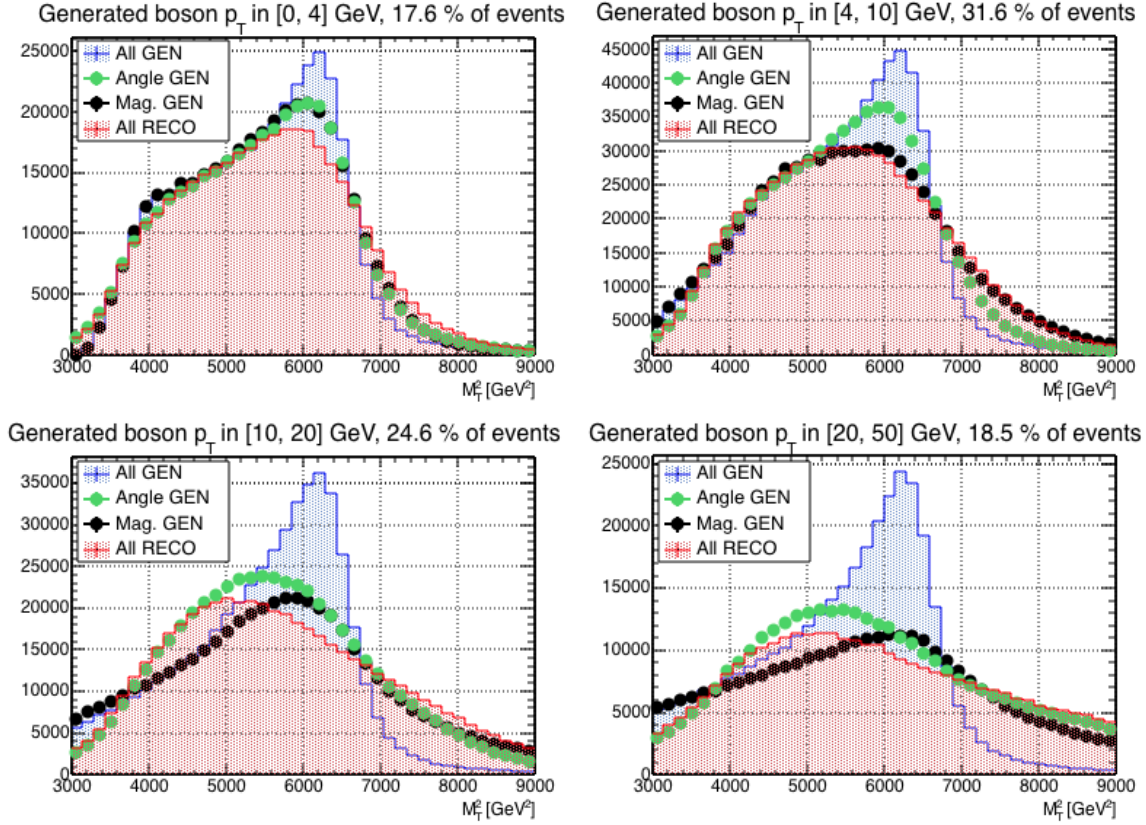


Figure 4.9: Transverse mass distribution for W events in different generated boson p_T regions (i.e. MC truth recoil). The fraction of events belonging to that region is reported as well. In each p_T region, different colors indicate different recoil variables used to build M_T : blue MC truth angle and magnitude; green MC truth angle and reco TK magnitude; black reco TK angle and MC truth magnitude; red reco TK angle and magnitude. Reconstructed muon is always used.

the contrary, no improvement is visible when the MC truth information is used only for the magnitude. In the bulk of the spectrum, it is more convenient to improve the recoil angle measurement accuracy rather than trying to perfectly correct the scale. In the harder part of the spectrum, approaching the jet regime, TK measures the angle correctly but has a small response: on average the charged tracks contribute only to 60% of the recoil. The scale is recovered when the true magnitude is used instead of the TK magnitude.

In summary, the fact that the recoil is a 2D object cannot be neglected and the relative importance of the angular and magnitude component of the recoil uncertainty changes depending on the magnitude of the true recoil.

4.5.2 Resolution effects

In order to understand how important the detector effects are, it is useful to fix a given experimental definition of the recoil. Fig. 4.10 reports the distribution of $\Delta \vec{h} = \vec{h}_{MC} - \vec{h}_{TK}$ in two different frames of reference: the CMS experiment x - y transverse frame and the frame identified by the unit vector parallel and orthogonal to the boson flight direction.

The shape of the 2D distributions in the above cases presents great differences. It

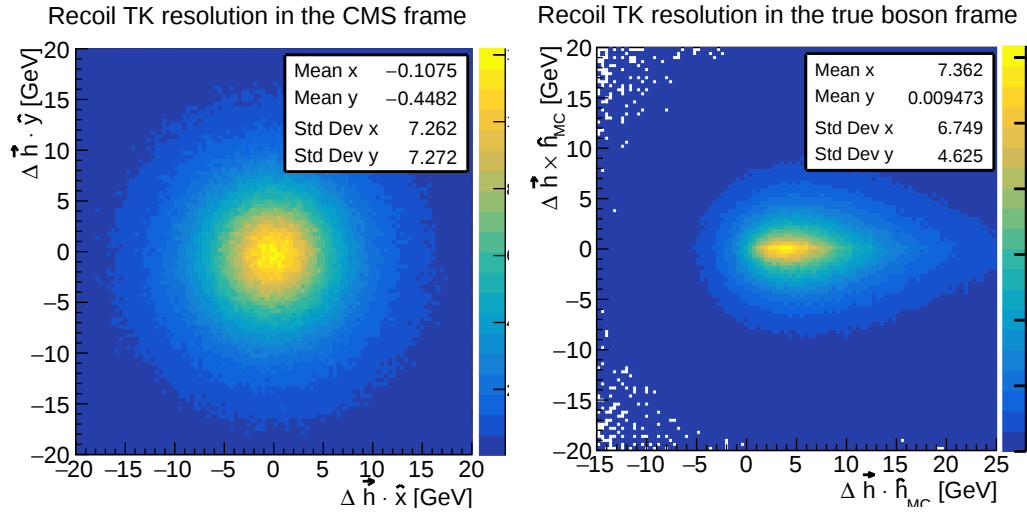


Figure 4.10: Distribution of the TK recoil resolution in the CMS frame (left) and in the true produced W boson flight direction frame (right).

underlines the large dependence on the frame of reference choice that the recoil and its resolution have. In the CMS frame, which is event-independent, the resolution is almost circular and small discrepancies of its mean from 0 are due to relative displacement of the beam and detector axis. Given the symmetry, it is possible to quote 7 GeV as the resolution on the recoil but, however, it is hard to correlate this number with some effects on the transverse mass.

The right plot, showing the resolution in an event-dependent frame, presents a completely different situation where the circular symmetry is lost. It is clear that the reconstructed momentum of particles coming from the W vertex has different variances for parallel and orthogonal component.

Looking at both these plots, it is also possible to see, considering the true recoil (i.e. W boson p_T) spectrum, how sizable is the detector effect. Being the resolution of a reasonable experimental definition, such as TK recoil, of order 7 GeV, it is very hard to promptly reconstruct the W p_T spectrum, peaking at 4 GeV.

4.5. A PRECISE MEASUREMENT OF THE RECOIL

A different and frame-independent way of studying the resolution is to look at the distribution of the magnitude $|\Delta\vec{h}|$ (fig. 4.11 left) and the distribution of the difference ΔM_T^2 between MC and reconstructed transverse mass (fig. 4.11 right). The TK recoil

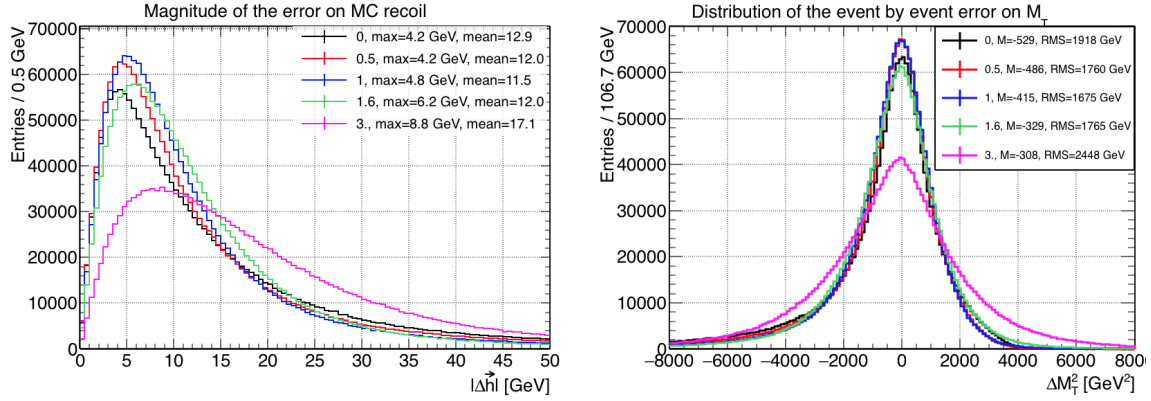


Figure 4.11: Left: Distribution of the magnitude of the recoil resolution for different scale factor applied to the TK recoil.

Right: Distribution of the error on the transverse mass for the same scale factors used on the left.

is compared for different scale factors applied to its magnitude, including 0. In that particular case, the distribution of $|\Delta\vec{h}|$ is the W p_T spectrum. From the reported plots, it is very hard to understand if there is a scale factor clearly performing better than the others. The plots indicate very clearly that the resolution is comparable to the recoil magnitude.

The plot on the right shows that the choice of the scale have a sizable effect on the transverse mass resolution. The scale factor 0, corresponding to neglecting the recoil and using $M_T^2 \approx 4p_\mu^2$, leads to a resolution compatible to the factor 1.6, which is the value that recovers on average the recoil scale. The best scale factor for this figure of merit seems to be between 0.5 and 1, whereas exceptionally large scale factor rapidly degrades the resolution.

The bottom line is that a finite recoil resolution produces a broader M_T^2 distribution since its width will roughly be the squared sum of the $M_T^2(MC)$ distribution and of ΔM_T^2 :

$$\sigma_{M_T^2}^2 \approx \sigma_{M_T^2(MC)}^2 + \sigma_{\Delta M_T^2}^2$$

The transverse mass distribution built using the true MC recoil (blue histogram of fig. 4.7 left) has a RMS of 1014 GeV², whereas the RMS of ΔM_T^2 distribution has a RMS of 1675 GeV² when TK recoil is used. A larger width, with the consequence of a fainter peak, is related, with an increase of the statistical uncertainty on the W

4.5. A PRECISE MEASUREMENT OF THE RECOIL

mass (at fixed number of events) but it does not necessarily imply an increase of the systematic uncertainty: a broader recoil estimator can be preferable if it improves the systematic.

It is important to say that often the recoil component orthogonal to the boson $u_{\perp} = \vec{h} \times \hat{p}_W$ has been used to evaluate the resolution. The understanding developed in this work suggests that, on the contrary, this is not a good estimator since it convolves many physical quantities (i.e. angle and magnitude). In addition, the 2D resolution is projected in a very particular frame and this implies that scaling the recoil directly scales the resolution. The latter point may bring some oddities: in the case of recoil estimator $\vec{h} = 0$, the resolution would be perfectly 0, even if the recoil in each event is wrong by 100%.

Without any doubts, looking at 1D distribution is easier for comparison and is a faster way to have a qualitative feeling of what is happening. In our case a rough but clear estimate is $|\Delta\vec{h}|$, which does not depend on a specific frame.

Chapter 5

A new recoil definition

This core chapter of the thesis presents the original work done in order to improve the experimental definition of the recoil. The goal is to find the best way of combining the information detected in each event to obtain an estimator of the value of the true recoil vector $\vec{h}$.

Practically, the definition is a mathematical function $f(x)$ of a certain number of observable features x of the event. The function f maps the phase space of x into a 2-dimensional space ($\sim \mathbb{R}^2$): the vectors belonging to that space are interpreted as the recoil in a given reference frame.

The recoil computed with this function is used for the main purpose of evaluating the transverse mass. Put differently, it is an effective tuning of the way in which the information in each event is used. The final goal is to maximize the performances in the W mass measurement, both in terms of statistical and systematic uncertainty.

While the proper treatment of uncertainties is left for the last chapter, this one will focus just on the definition. It begins with a brief introduction which contextualizes previous definitions in the "function-as-definition framework" and then explains the idea of the improvement. The main sections contain the technicalities of how multivariate analysis can be used for improving the recoil definition. The dissertation shows in full the application details and explains the work necessary to obtain an event-by-event prediction of the probability density function (pdf) of the true recoil $\vec{h}$. The jargon used to indicate this procedure is *regression*. The pdf, provided in each event by the *regression*, is not a point estimation of the recoil but it contains all the information that can be extracted from the observables x measured in the event.. Eventually, functions of the pdf features (e.g. the median) are used for the derivation of a event-by-event recoil estimator, practically consisting in a definition.

In the last section of the chapter, regression performances and closure tests are shown.

5.1 Recoil definitions

The W boson mass is measured comparing the transverse mass distribution in data with those obtained in the simulation. Practically the comparison is between a histogram, representing the data transverse mass distribution, and several histograms, called templates, built using different mass hypothesis. This will be discussed in detail in sec. 6.2.

One important thing to remember before analyzing different definitions is that no intrinsic bias can be brought in the m_W fit by a particular definition. As long as the same procedure is applied in data and in MC, bias can arise only from differences between the samples. Any definition is in principle acceptable and, at worst, sub-optimal. Indeed, a definition can be a bad estimator of the true recoil: the transverse mass computed with it might be broader or more dependent on W p_T but, if the MC is an accurate description of the data, still no intrinsic difference is present between fitted data distributions and the MC templates.

A first and trivial example of the recoil definition is the identical null function, which maps every event into $\vec{h} = 0$. This very odd definition induces a relative error on the recoil estimation of 100%, since $\Delta\vec{h} = \vec{h}_{\text{true}}$, but still no particular bias arises. Specifically the m_T templates made with this definition need no data-MC calibration of the recoil (no systematic error due to the recoil modeling). In this case, of course, no information is added to p_μ and thus no mitigation on the limited W p_T knowledge occurs.

The state of the art of popular recoil definition among the CMS collaboration presents two alternatives (already introduced in sec. 4.3.1): the track recoil (TK) and the particle flow recoil (PF).

The TK recoil can be seen as a definition of the recoil which takes two input observables of the event $x = \{h_{TK}^x, h_{TK}^y\}$, where h_{TK}^x and h_{TK}^y are respectively the $\hat{x}$ and $\hat{y}$ projections of $\vec{h}_{TK}$. The function $f(x)$ applied is trivially the identity. This definition, which has been used for the W – like measurement at CMS [41], is characterized by an high purity, given by the tracker good discrimination power between PV and not PV tracks, and by the consequent absence of dependence from PU and neutral cluster modeling. On the other hand, it just combines in a straight forward way the information from tracks with a limited efficiency: only the four-vectors total sum is considered and shape or single track information are totally lost. In addition, reconstructed tracks

(i.e. charged particles) contribute for about the 60% of the total energy released by the recoil in the detector. The remaining 40% due to neutrals is neglected.

Similarly to TK recoil, PF recoil is a definition which takes $x = \{h_{PF}^x, h_{PF}^y\}$ and uses the identity as f . In this case, the full information of the detector is used but again with a limited performance: no distinction is done between PV and not PV and only the sum of four-vectors is considered. Since each pile-up vertex in the bunch crossing has a null total sum of the particle transverse momenta, PF recoil in principle measures the W p_T . However, detector effects make this property valid only on average and the result is an estimator with a huge variance: for each vertex not all the particles are successfully reconstructed and the effect of these missing particles is that of adding a random component to the the W p_T in the PF recoil.

5.1.1 An improved recoil

In this work a new definition of the recoil is presented. There are two new key points which this definition has been developed for: using as much present information as possible and combining it in a way which maximizes the event-by-event knowledge of the true recoil.

The idea is to use all the PF candidates, muon excluded, and to extend the list of variables used beyond the single sum of the four-momenta. For instance, this *improved PF recoil* uses, as explicit input observable x , the sum of four-momenta of different groups of particles (e.g. TK, ntnpv), the shape information of the same groups, a selection of single particle information (e.g. the pt of the most energetic track) and the global event features (e.g. number of tracks). Shape variables includes parameter which can describe the distribution of particle in the detector, such as the ratio between the vector and scalar sum of the tracks momentum. A long list of variables has been considered and the ones really used are reported later.

For what is about the function $f(x)$, the goal is not to fix an explicit formula but, instead, to optimize a non explicit formula. The optimization is done in order to have the best performance on the W mass extraction from the transverse mass fit. Practically, this goal is pursued predicting $P(\vec{h}|x)$ in each event, which is the probability for the true recoil to assume a certain value given the set of x observed. Once that the full pdf is known, the choice of the point estimator for the recoil to be used in the analysis is still free: any estimator calculable from the pdf is acceptable.

It has to be remarked that if the set of observable x completely determines the recoil (e.g. if $\vec{h}_{\text{true}} \in x$, just as extreme case), than $P(\vec{h}|x)$ is a delta function centered on the

true value. In the opposite case, when all the variables x are completely uncorrelated with $\vec{h}$, the pdf prediction is flat for the recoil direction and has the shape of W p_T spectrum for the magnitude. A more reasonable and real case lays in the middle, when the information from x can constrain the value of $\vec{h}$ without determining it completely. A peaky shape as pdf prediction is expected, with a maximum placed at the most probable value.

Finally, it is important to underline that additional variables, which are loosely or not correlated with the recoil, do not hurt, in principle, the predictive power of the regression. They are just useless. In practice this is not true, since a higher number of dimensions in the input phase space requires bigger statistical samples and more computational power. Thus, it is desirable to use only features which bring a meaningful amount of information on the recoil.

5.2 A multivariate semi-parametric regression

The above mentioned goal is pursued through the implementation of a semi-parametric regression based on deep dense neural networks (DNN) which fits the functional form of $P(\vec{h}|x)$. Before passing to the specific application on the recoil, a recap of what is meant in this work for DNN and semi-parametric regression is provided.

5.2.0.1 Neural Networks: a generic function

Similarly to the other multivariate analysis methods (i.e. BDTs, SVM, etc.), Neural Networks are generic function depending on a usually large number of parameters. The explicit form of the function and the number of parameters are usually addressed as *architecture* of the network. While the architecture has to be fixed *a priori*, the parameters value is usually (but not exclusively) set during a phase called training or fitting. The values for the parameters are chosen in order to minimize a function of the network output, called *loss function*. This function, used as figure of merit, often coincides with the likelihood of a fit.

For the purpose of this work only deep dense neural networks (DNN) are used (fig. 5.1). They are made with a hierarchical structure of one input layer, one or more subsequent hidden layers - more hidden layers make the network *deep* - and one output layer. Each neuron takes as input the output of all the neurons of the previous layer. *Dense* refers indeed to the fact that a neuron is connected to all the neurons of the previous and of the following layer. The number of neurons in each layer is arbitrary and the input

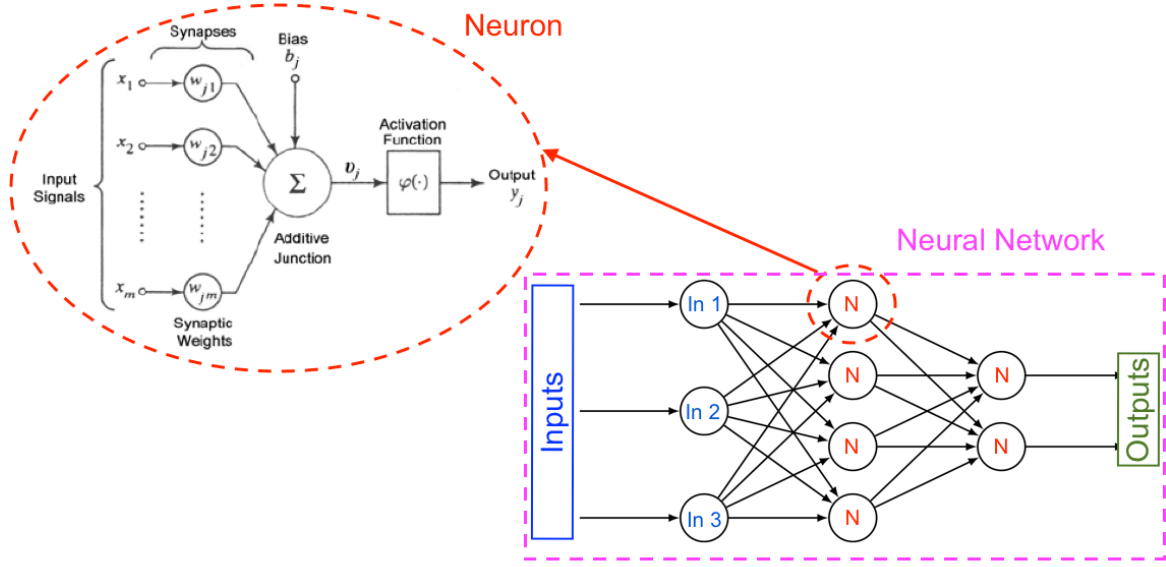


Figure 5.1: Scheme of a dense fully-connected neural network with 3 input and 2 outputs.

neurons do not have inputs but just pass the given information to the subsequent layer. Each neuron, excluded the input ones, performs a linear combination of the inputs, applies a function to the combination value and then returns as output the function result. In formula:

$$\text{Output } N^i = \phi_L \left(\sum_j w_j^i x_j + b^i \right)$$

where b^i and w_j^i are tunable parameters of the neuron, the sum on j is performed on the output of all the neurons of the previous layer and ϕ_L is a function (called *activation function*) fixed by the architecture and equal for all the neurons in the same layer.

The intrinsic symmetry between neurons of the same layer is broken by the randomness of the training phase. At the beginning of this phase, the value of each parameter is randomly set and, during the optimization, is often changed in order to reduce the value of the loss function. A gradient descent optimization algorithm [49] is usually implemented in its stochastic form to speed up the network convergence.

Formally, it can be demonstrated that, with a sufficient number of neurons, a DNN can approximate an arbitrary function with an arbitrary degree of accuracy. For the purpose of this work, it is important to know that the DNN form is flexible enough to describe a generic function in a high dimensional space with a relatively small number of parameters. All DNNs developed in this thesis are based on the python library [50], chosen because it supports the usage of GPUs to speed up the computation.

During the developmental phase of the work other architectures and frameworks have been considered (i.e. series of polynomial, convolutional NN, BDTs, ...) but they are not here reported since not relevant for the final work.

5.2.0.2 Semi-parametric regression

Perform a semi-parametric regression means to fit a set of data $\{y, x\}_i$, where i runs on the observations, in order to describe $P(y|x)$ with a pre-determined model $f(y; \alpha(x; k))$: f is a fixed functional form and α its parameters; each parameter $\alpha(x; k)$ is a fixed functional form of the observable x and of some parameter k ; the parameters k are the degrees of freedom of fit and are chosen in order to describe $P(y|x)$ at best.

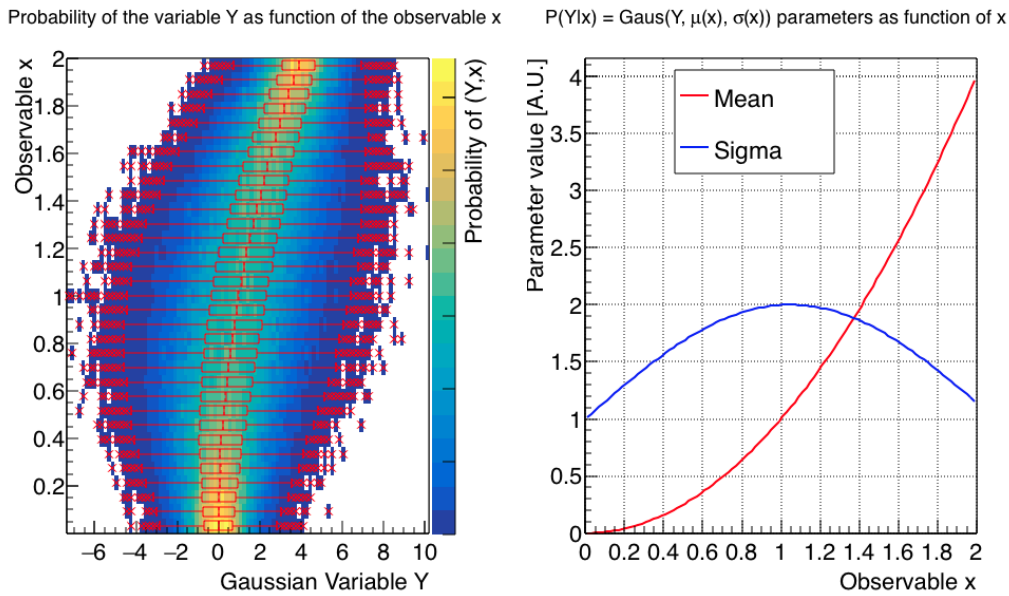


Figure 5.2: Example of a semi-parametric regression.

Left: 2D distribution of the variable y and the observable x with the so called candle plot over-imposed: red box center corresponds to the mean and red box half-width to the RMS. Horizontal red lines ends at the 90% of the distribution and the small red crosses shows the events which stay outside.

Right: Fitted dependence of the parameters μ and σ of the y Gaussian distribution as a function of the observable x .

In fig. 5.2 is reported an example of semi-parametric regression to Gaussian distributed points y as a function of one observable x . The distribution is fitted with a Gaussian, whose mean and sigma depend on x . So $\alpha = \{\mu(x; k_\mu), \sigma(x; k_\sigma)\}$ and each α_i being a fixed function form. In this example the following form is used: $\alpha_i = \sum_{n=0}^3 k_{i,n} x^n$.

The eight parameters k are then chosen minimizing the likelihood function:

$$f(y, x; k) = \mathcal{L} = \frac{1}{\sigma(x; \vec{k}_\sigma) \sqrt{2\pi}} \exp \left(-\frac{(y - \mu(x; k_\mu))^2}{2\sigma^2(x; k_\sigma)} \right)$$

The negative log likelyhood is also often called *loss function*.

At the end of the fitting phase, when all the k parameters are freezed, $f(y; x)$ becomes a function of the x which is supposed to describe the conditional pdf of y given x : $f(y; x) = P(y | x)$.

5.2.1 Application to recoil definition

Setting up a *regression* of $P(\vec{h}|x)$ requires to find a sufficient elastic functional form f that, opportunely varying the parameters α , can describe $P(\vec{h}|x)$ in the whole input phase space (i.e. for each value of x). The DNN is then used as the function which gives the values of the parameters α as function of the observables x . This means that the x are the input of the DNN, the α are the output and the free parameters in the fit (coincident with the DNN training) are the w_j^i and b^i parameters of each neuron i . On one hand, it is a reasonable task to properly define two 1D functional form which can describe the pdf of the recoil component in a given frame. On the other hand, it is much harder to find a function with high input space dimension (dimension of the x space, order 10) and output space dimension equal to the number of the parameter of the 1D functions above (typically 4-6): for this reason, a DNN is needed for the latter task.

Once again, it is important to underline that the sample fitted (called training sample), according to which the DNN parameters are chosen, is relevant only for the performances of the recoil definition, but does not bring any intrinsic bias on the m_W measurement. Of course, the more the training sample is similar to W events in data the better: when the predicted pdf in the W events is more accurate, the true recoil value is better estimated.

In this work, the training is performed on the W MC sample where the true value of $\vec{h}$ is known and the events are fairly similar to the data. Assuming that MC describe perfectly data¹, a bad convergence of the regression implies that the recoil is not perfectly estimated in data and MC in the same way, so no difference between data distribution and templates is induced because of the recoil definition.

¹If that is not the case a systematic uncertainty arises. This is extensively treated in sec.6.4

Recoil correction parametrization

Even if DNNs are a powerful tool, it is convenient for a matter of computational power and stability to have a model f with a small number of parameters. Furthermore, since f has to describe the pdf in the whole space, the DNN training is easier if $P(y | x)$ has a shape with small variation across the x phase space. When this happens, the α parameters have a modest variation, which makes the DNN task easier, thus, a simpler architecture is sufficient and the training convergence faster. Given that an explicit form for f is needed, it is better when the pdf is similar to reasonable function, such as a gaussian-like distribution or in general a shape with one peak. It is very inconvenient to use as a fitting function a very complex form such a sum of several Gaussian or even more weird forms with 10 or more parameters. A peaky shape is expected since the fit targets the true recoil value so, if the chosen x brings some information on the recoil, the probability in a given event should clump near the true value. On top of this, if $\alpha(x, k)$ has a very strange form, which means that the parameters are changing fast in the event phase space, the fit will have convergence problems because it can easily get stuck in local minimum.

For what is above, instead of fitting directly the recoil value, it is more convenient to define some corrections and, then, fit them. Starting from an easy definition of the recoil (i.e. TK or PF), the correction are meant to retrieve the true value of the recoil. Being the recoil a 2D object, there is no obvious or unique parametrization for this correction. A possibility is the $e_1 - e_2$ or *scale-angle* correction which uses the

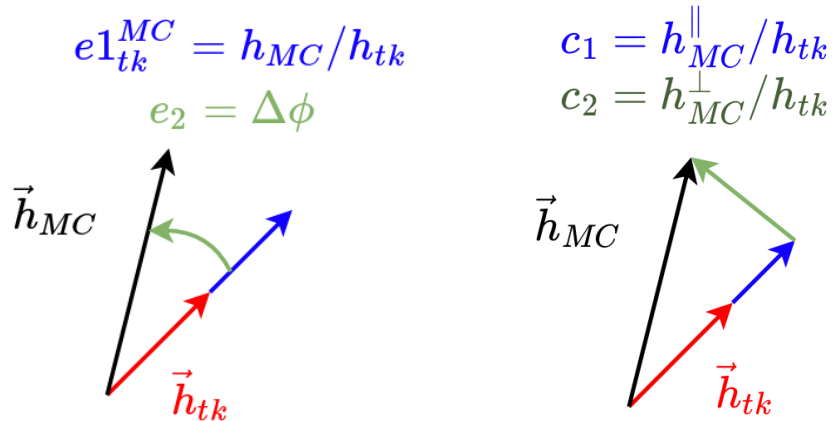


Figure 5.3: Sketches illustrating the definition of the scale-angle correction coefficient (left) and of the scalar recoil frame oriented correction coefficient (right).

ratio between true and TK recoil magnitude and their angular distance. A different possibility is the $c_1 - c_2$ correction which uses scalar correction for the recoil component

in the TK recoil frame (i.e. the two-dimensional reference frame where the TK vector is along the $\hat{x}$ axis). In fig. 5.3 sketches illustrating these two parametrizations are shown. Other corrections have been taken into account but for sake of brevity are not here reported.

The freedom in the parametrization choice needs to be exploited to simplify the task of finding the 1D model functions: different parametrizations have different shapes and it is better to choose the one which can be better described by a simple function. The goal is to fit separately two uncorrelated correction coefficients, one for each component of the chosen reference frame. In addition, some correction coefficient can present singularities, which has to be avoided. Looking at fig. 5.4, it is clear that the

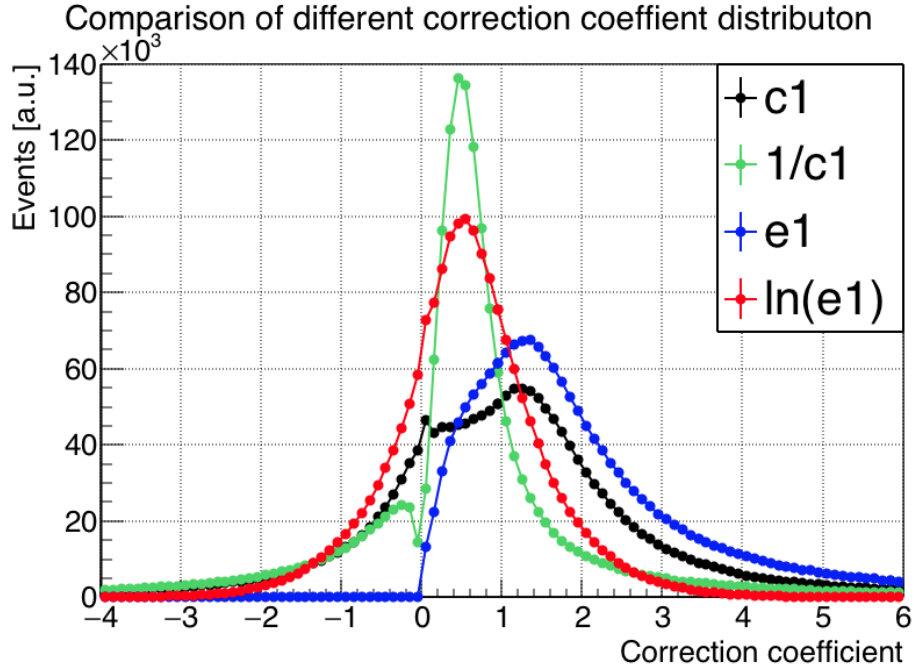


Figure 5.4: Distribution of correction coefficient from different recoil parametrization schemes.

distribution of different corrections may be really different. In particular, c_1 or $1/c_1$ present singularities and multiple peaks due to the kinematics, making them almost impossible to fit. On the contrary, the e_1 coefficient presents a regular shape (blue), whose boundary at 0 can be further removed by applying a logarithmic transformation (red).

The above reasoning fixes the choice of the parametrization scheme but still leaves a freedom in choosing which basic recoil definition should be the starting point. As previously discussed, the importance of correcting the angle is sizable but the information that can be exploited for finding the right correction is limited. Given the cylindrical

5.2. A MULTIVARIATE SEMI-PARAMETRIC REGRESSION

symmetry of the detector and the invariance under rotation of the production mechanism, the only variables that can improve the knowledge of the angle are the angular distance between the three different components of the recoil: tracks from PV; tracks not from PV; and neutrals.

For this reason, it has been used TK recoil as baseline since its angular information requires slightly less correction than the PF recoil one (fig. 5.5).

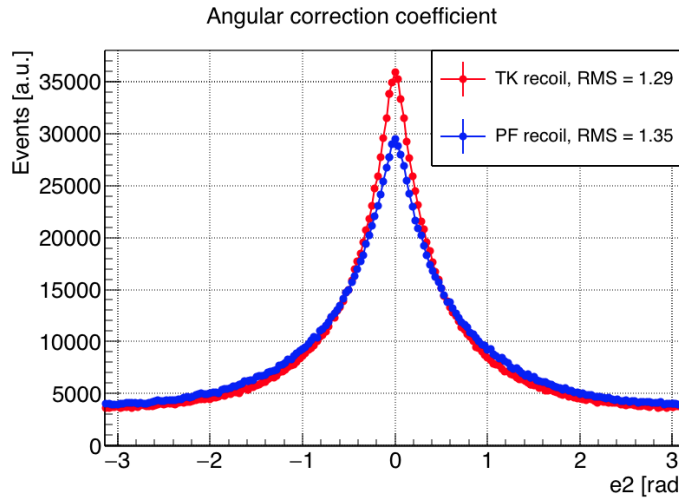


Figure 5.5: Angular distance between the experimental measured recoil and true recoil in the PF and TK case. This corresponds to the $e2$ correction coefficient.

5.2.2 Details of the trained regression

The regression trained has as target the conditional distribution of

$$P(y|x) = P(\ln(e1)|x_1) \times P(e2|x_2)$$

being $e1 = h_{MC}/h_{tk}$ and $e2 = \phi_{MC} - \phi_{tk}$ the *scale-angle* correction coefficient for the TK recoil, x_1 and x_2 the variables for which the probability of $e1$ and $e2$ is conditional and $x = x_1 \cup x_2$.

A complete list of the observable used as input of the DNN regression is presented in tab. 5.1. The observable input space of the x_1 variables comprehends the ones with a tick in the $\ln(e1)$ column, the x_2 , instead, have a tick in the $e2$ column.

The function used to fit $P(\ln(e1)|x_1)$ is a Gaussian with two different exponential

5.2. A MULTIVARIATE SEMI-PARAMETRIC REGRESSION

Symbol	Formula	Notes	ln(e1)	e2
h_{tk}	$ \vec{h}_{TK} $		✓	✓
ϕ_{tk}		Azimuthal angle of TK recoil	✓	✓
R_{ntnpv}	$\ln(h_{ntnpv}/h_{TK})$		✓	✓
R_m	$\ln(m_{TK}/h_{TK})$	m_{TK} is the invariant mass of the TK recoil	✓	✓
R_{lTK}	$\ln(p_T^{\text{leading TK}}/h_{TK})$	leading TK is track with highest p_T	✓	✓
R_{lnt}	$\ln(p_T^{\text{leading nt}}/h_{TK})$	leading nt is neutral cluster with highest p_T	✓	✓
S_{TK}	$\frac{\sum_{i \in TK} \vec{p}_T^{(i)}}{\sum_{i \in TK} p_T^{(i)}}$	Vector over scalar sum of particles $\vec{p}_T$ belonging to TK recoil	✓	✓
S_{nt}	„	Same as above but with neutral clusters	✓	✓
S_{PF}	„	Same as above but with all PF candidates	✓	✓
	$\cos(\phi_{tk} - \phi_{ntnpv})$		✓	
	$ \Delta\phi_{TK-ntnpv} $	Modulus of angle between TK and ntnpv recoil		✓
	$\varsigma \cdot \Delta\phi_{TK-PF}$	Angle between TK and PF recoil		✓
	$\varsigma \cdot \Delta\phi_{TK-\text{leading TK}}$	Angle between TK and leading track		✓
N_{TK}		# of tracks from PV	✓	✓
N_{vtx}		# of reconstructed vertexes	✓	✓

Table 5.1: Event observables used as input of the DNN for the semi-parametric regression. Definition of TK, PF and ntnpv can be found at the end of sec. 4.3.1; The sign of the angles is defined as $\varsigma = \text{sgn}(\Delta\phi_{TK-ntnpv})$.

tails matched with c_1 continuity (example in fig. 5.6 left):

$$f_{lne1}(y|\vec{\alpha}_{lne1}(x_1)) = \begin{cases} \frac{1}{N}e^{\alpha_1^2/2}e^{\alpha_1 t}, & \text{if } t < -\alpha_1 \\ \frac{1}{N}e^{t^2/2}, & \text{if } -\alpha_1 \leq t \leq \alpha_2 \\ \frac{1}{N}e^{\alpha_2^2/2}e^{-\alpha_2 t}, & \text{if } t > \alpha_2 \end{cases}$$

where $t = \frac{y - \mu}{\sigma}$, $\vec{\alpha}_{lne1} = \{\mu, \sigma, \alpha_1, \alpha_2\}$ is the vector of the model parameters and

$$N = N(\sigma, \alpha_1, \alpha_2) = \sqrt{\frac{\pi}{2}}\sigma \left[\operatorname{erf}\left(\frac{\alpha_2}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{\alpha_1}{\sqrt{2}}\right) \right] + \frac{e^{-\frac{\alpha_2^2}{2}}}{\alpha_2}\sigma + \frac{e^{-\frac{\alpha_1^2}{2}}}{\alpha_1}\sigma$$

is the normalization factor.

The model (f_{e2}) used to fit $P(e2|x_2)$ is a Gaussian with two exponential tails, as before, plus a constant term (example in fig. 5.6 right): it has five free parameters and the normalization factor depends on the constant term too. The constant term is necessary to account for the fact that $e2$ does not go to 0 at the boundaries $[-\pi, \pi]$ of its domain. This behavior is only effectively treated with the constant term, is due to the circularity of the variable: in principle, $e2$ is continuous in $-\pi][\pi$, but the fitting function it is not. It has been verified that this inconsistency introduces discrepancies which can be neglected.

The regression fitting has been performed separately for the two correction coefficients and, for each parameter α of the fitting functions, a DNN has been trained. So, the fitting of $P(\ln(e1)|x_1)$ with f_{lne1} has then involved four DNN trained simultaneously, whose parameters has been determined minimizing the likelihood:

$$\text{loss} = \ln \mathcal{L} = \sum_i \ln (f_{lne1}(\ln(e1^i) | \vec{\alpha}_{lne1}(x_1^i)))$$

where i runs on the fitted sample events, $e1^i$ is the value of the correction coefficient in the i -th event and x_1^i is the set of observable of the i -th event. A similar procedure has been performed for the fitting of $e2$.

Further details on the training can be found in appendix A.

At the end of the fitting phase, two functions are obtained, able to predict $P(\ln(e1) | x)$ and $P(e2 | x)$ event-by-event as a function of the observables x . However, to extract the pdf of the recoil, it is necessary to take care of the jacobian factor, introduced when applying the corrections. First, the predicted (i.e. expected) pdf of the magnitude and of the angle of the true recoil is computed applying the transformation

$$h_{exp} = e^{\ln(e1)} \cdot h_{TK} \quad \text{and} \quad \phi_{exp} = e2 + \phi_{TK}$$

and applying the proper jacobian. Then, two event-by-event estimators of the recoil are built: a vector $\vec{h}_{pk}$ made with the peak of the h_{exp} predicted distribution and the peak of the ϕ_{exp} predicted distribution; a vector $\vec{h}_{mnpk}$ made with the mean of the

h_{exp} predicted distribution and the peak of the ϕ_{exp} predicted distribution. For all the previous four estimator, the explicit form in terms of the DNN outputs $\vec{\alpha}_{lne1}$ and $\vec{\alpha}_{e2}$ has been derived.

Given the finite size of the interval in which ϕ_{exp} takes value, the tails do not have an important weight since they are necessarily truncated at π and $-\pi$: the mean and the peak of the e_2 and ϕ_{exp} are then very similar and no difference can be appreciated between them.

The two event-by-event estimators $\vec{h}_{pk}$ (PK recoil) and $\vec{h}_{mnpk}$ (MNPK recoil) are not a complete set of possible definitions but are both reasonable and simple.

5.3 Goodness of the definition

At this point, a fair question is if the regression is able to describe the probability it has been trained for and how the extracted definitions of the recoil perform, in comparison to ones used before (i.e. PF and TK recoil). Unfortunately, there is no simple unique answer and the previous questions have to be explored in several aspects.

First of all, it is necessary to check if the regression has converged. Once again, there is no bias in the W mass analysis if the convergence is not good, as long as the same definition is applied both on data and MC. However, this question is important since only a good convergence may assure an optimal usage of the input information.

Secondly, it is important to check how the extracted definition behaves on physical observables, such as the scalar resolution $|\Delta\vec{h}|$. As the previous point, this one is relevant not only for the W mass analysis but also for all other possible applications. However, a particular focus is given to the figure of merit more important for W mass. For instance, it is possible to check the $\Delta M_T^2 = M_T^2(MC) - M_T^2(RECO)$ distribution which is correlated with the statistical uncertainty in the W mass fit (see sec. 4.5.2).

Finally, it is crucial to check the performance in recovering the event-by-event correlation between the muon and the recoil. This corresponds to the mitigation power of the systematic uncertainty due to the knowledge of the W p_T spectrum. Despite the efforts dedicated, it is very hard to estimate, with a simple figure of merit, the recovered correlation and no immediate way of determining it has been individuated. The brute force way followed is performing the W mass fit with different W p_T spectra in pseudo-data and templates (emulating the DATA/MC discrepancies), but this is discussed in the last chapter.

In the following of this section, first, the proper convergence of the regression is tested through two procedures: a local check consisting in the comparison of predicted

and real distributions; a particular global check called *sum of pdfs*.

Then, the performances of pk and $mnpk$ estimators are compared with PF and TK recoil using several figure of merits: residual corrections and resolution estimators.

5.3.1 Regression - network convergence check

In order to check if the regression, and thus the fit of the DNN parameters, had converged well, it would have been necessary to check - in each single corner of the phase space - if $P(\ln(e1), e2|x)_{true}$ is well described by

$$P(\ln(e1), e2|x)_{predicted} \sim f_{lne1}(\ln(e1); \vec{\alpha}_{e1}(x_1)) \times f_{e2}(e2; \vec{\alpha}_{e2}(x_2))$$

and, eventually, perform a statistical test. In practice this is impossible given the high number of dimensions of the x space: no reasonable statics will never be enough to do a fine binning in ~ 20 variables and perform a significant statistical test on it.

It is however possible to approximate the above investigation by binning in only one or two variables. For meaningful results, it is convenient to choose the variables more correlated with the recoil true value, so that a reasonable division of the phase space is obtained without losing too much statistic, binning in almost useless variables. This procedure has been extensively done in the early stages of the work and one more time after the training of the final regression. In fig. 5.6, the distribution of $\ln(e1)$ and of $e2$, in an illustrative bin of the x space, are shown to exemplify the procedure. It is important to choose the bins boundaries in order to have a reasonable statistics of non less than 1000 events, so that the procedure has some statistical meaning even in the very important part of the tails. In each bin, it is then compared the distribution of the correction coefficients (blue) with the pdf predicted by functions f_{lne1} and f_{e2} , whose parameters are set to the average of the predicted values on the events populating the bin (red). Even if this is clearly an approximation, the function obtained in this way can be used for a qualitative comparison. It is also interesting to report the result of a dedicated fit (dashed black), performed in each bin, at the correction coefficient distribution with a function of the same form of f_{lne1} or f_{e2} . The parameters of this fitting function are set maximizing the likelihood.

Even if this is just a rough figure of merit, a good agreement has been found in all the plots checked. This has been considered a good qualitative proof that the functions f_{lne1} and f_{e2} can provide a satisfactory description of the conditional probability.

5.3. GOODNESS OF THE DEFINITION

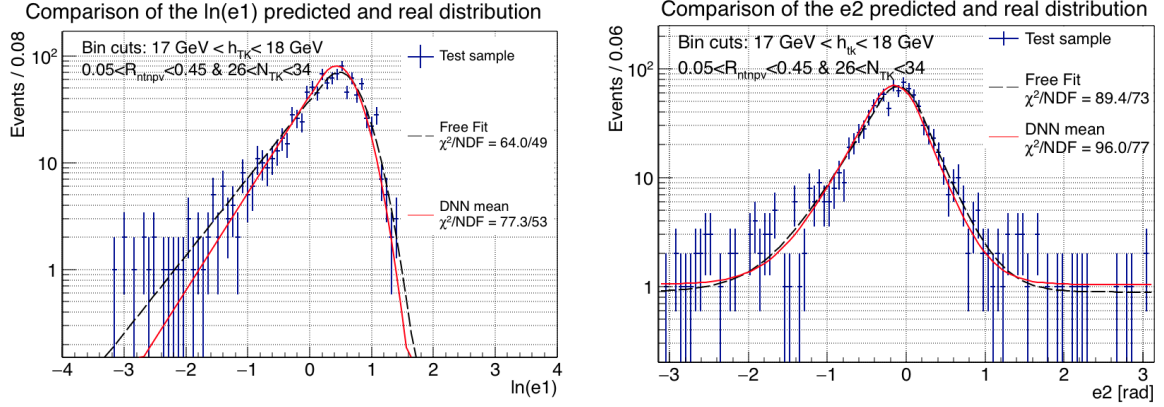


Figure 5.6: Distribution of the correction coefficient $\ln(e1)$ (left) and $e2$ (right) in a W MC sample identical but statistically independent from the one used for the regression training. Specific cuts, reported on the top of both plots, have been applied to check the regression performance in this illustrative bin. The total integral of the histograms is 9000, which contains in the $4 \cdot 10^{-4}$ of the W MC sample. The red line presents the regression average prediction in this bin. The black dashed line is a fit to the presented distribution performed with the same function used by the regression.

5.3.1.1 Sum of pdfs closure

A powerful tool to check the global convergence of the regression is to compare the sum of the predicted probability in each single event with the integrated distribution. This procedure, called *sum of pdfs plot*, is performed on the testing sample; a sample equal to the one used for training but statistically independent.

To explain how these plots are built, it is useful to consider again the example of the Gaussian variable y presented in fig. 5.2. There, the variable of which the probability is predicted by the regression, in each event, is y . The procedure to build the sum of pdfs plot is hereafter explained following the sketch in fig. 5.7: first, a histogram of the integrated distribution of the y variable in all the events is built (black points); then, for each event the probability of y given the observed x is predicted through the regression, converted to an histogram and normalized to unity (bottom blue distributions); finally, all the unity histograms are summed together and compared with the integrated distribution (right of the equal sign).

The convergence of this plot is, formally, a necessary but not a sufficient condition for the convergence of the regression. However, since no bias is induced on the W mass analysis in case of bad convergence of the regression, the goal for which this plot is used is to check for a general and qualitative agreement. In addition, the agreement between the integrated distribution and the sum of pdfs depends also on the goodness of the model. if the fitting function f is not a good model for $P(y|x)$, the agreement

5.3. GOODNESS OF THE DEFINITION

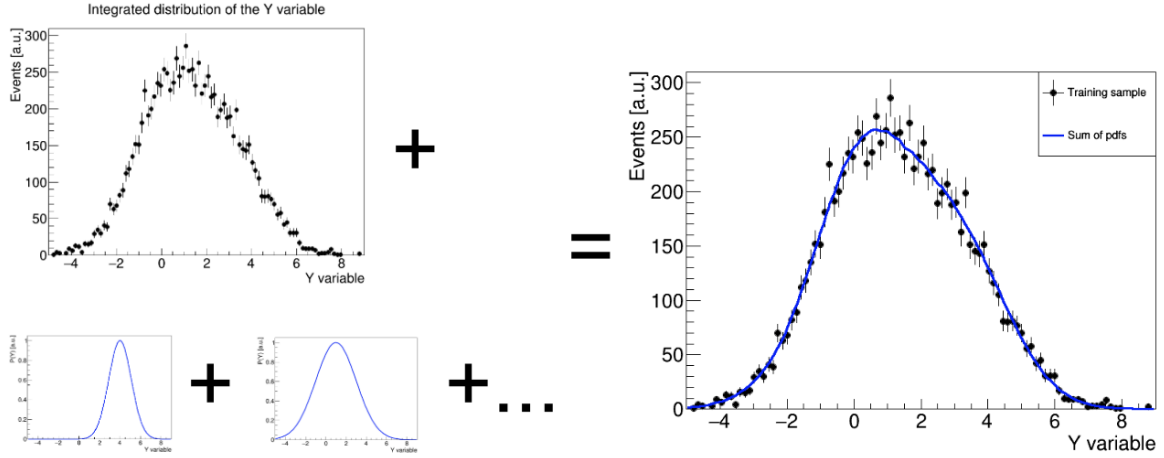


Figure 5.7: Sketch illustrating the procedure to build the sum of pdfs plots. Integrated distribution of the y variable is compared with the y probability predicted by the regression in each event summed altogether.

is not satisfactory even when the fitting has converged to its best value.

Application on the regression This approach has been extensively used to check the performance of the DNN regression on the recoil. In fig. 5.8 the sum of pdfs plots for the two fitted coefficients in the recoil regression are shown. The agreement

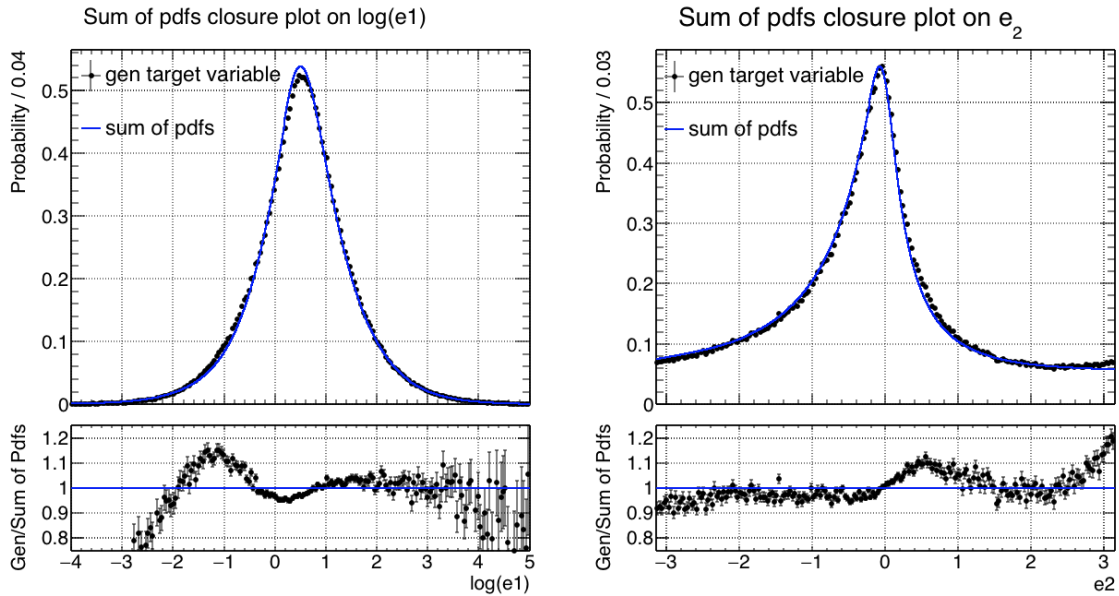


Figure 5.8: Sum of pdfs plots for the two correction coefficients which are targeted by the regression: $\ln(e_1)$ on the left and e_2 on the right.

is qualitatively very good and the ratio plot below each distribution shows that the

5.3. GOODNESS OF THE DEFINITION

description of the integrated pdf of both correction coefficients is at the 10% level. It has to be noticed that, on the right plot, the asymmetry is given by the fact that - for technical reasons - the correction signs are always taken in the direction of the angle $\Delta\phi_{tk-ntnpv}$.

Furthermore, the same procedure of sum of pdfs plots can be applied to the predicted distribution of recoil. Fig. 5.9 shows a very good agreement in both the recoil magnitude (left) and angular (right) distribution. In the left plot, three different distributions of

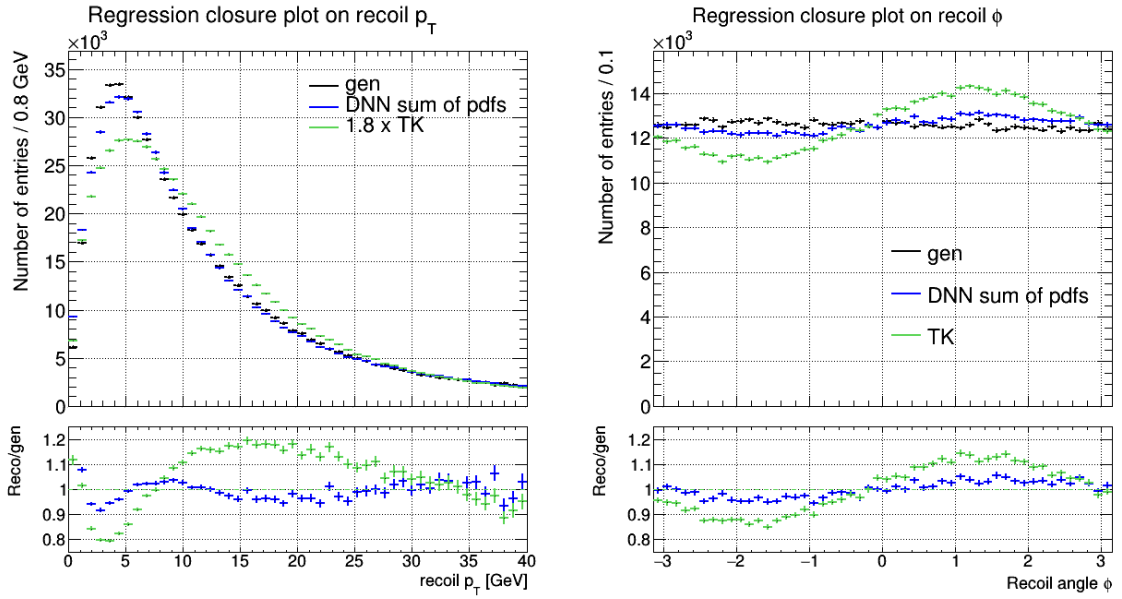


Figure 5.9: Sum of pdfs plots for the recoil in the scale-angle frame. In the plots it is presented a comparison between generator level distribution (black) for which the regression has been trained, sum of pdfs from DNN (blue). It is also reported the TK distribution of the recoil angle and the magnitude distribution of a TK recoil whose modulus has been scaled for 1.8.

the recoil magnitude (i.e. bonson p_T) are shown. As can be seen from the underlying ratio plot, the generator level distribution (black) is reproduced by the sum of pdfs from DNN (blue) with a order 5% agreement, whereas the TK recoil scaled for 1.8 - factor which should recover the correct scale on average - is worse than 10%.

A similar figure is presented for the angular distribution. The TK recoil is affected by the detector axis displacement with respect to the beam axis: it produces the periodical wiggle in the green curve of the right plot in fig. 5.9. Once again, the sum of pdf plot made with the DNN regression shows a better agreement than TK recoil. The increased flatness in the angular distribution is a sign of the successful deconvolution of detector effects.

The sum of pdfs figure of merit has also been checked in different bases for the recoil (e.g. $x - y$), but no significant difference has been found.

5.3.2 Goodness event-by-event of recoil estimators

The above subsection demonstrated how the regression has reasonably converged. That is important to be sure that all the information in the input variables has been properly exploited, crucial to have meaningful event-by-event estimators. It is now important to check how that estimators behave in estimating the recoil.

For this goal, the distribution of the residual corrections is studied for different recoil estimators, using events from the W MC sample. The residual correction on the magnitude is defined as the logarithm of the ratio between the estimated and true recoil modulus ($\ln(h^{\text{est}}/h^{\text{ref}})$). The residual correction on the angle is defined as relative angle between the estimated and true recoil ($\phi_{\text{ref}} - \phi_{\text{est}}$). In fig. 5.10 the residual correction

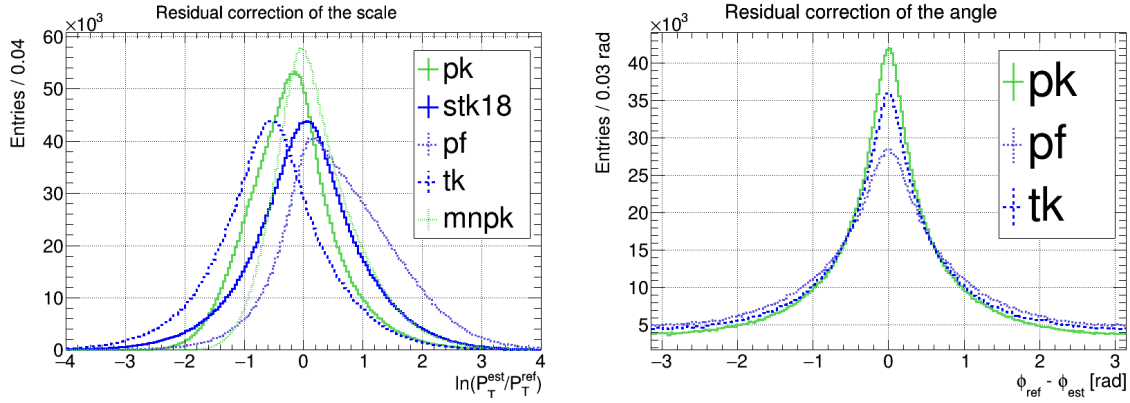


Figure 5.10: Distribution of the residual correction needed event-by-event to recover the MC information for the TK recoil, STK18 (TK with magnitude scaled by 1.8), PF recoil and the two point-like estimators produced with the DNN regression. In the x axis title, the superscript est stands for estimator (i.e. pk, TK, ..), whereas the ref stands for reference (i.e. MC).

on magnitude and angle of a given experimental estimator of the recoil is presented. The left plot, showing the distribution of the ratio between reconstructed and true MC recoil magnitude, addresses for a significant improvement of the two DNN regression estimators (PK and MNPK in green) with respect to the classical estimators (TK, STK18 and PF in blue). In particular, the improvement on the magnitude can be quantified mentioning that the mnpk distribution (dotted green) has a RMS of 0.73, whereas the STK18 has a RMS of 0.96.

Similar conclusions are suggested by the right plot, in which it is clear that the angle estimated with the PK recoil (peak of the predicted distribution, the one used also in *mnpk*) has a better resolution than the TK or PF recoil one. The pk distribution has a RMS of 1.24 rad, whereas the TK one of 1.32 rad.

In addition, the distributions of the scalar resolution $|\Delta\vec{h}|$ and of the event-by-event

5.3. GOODNESS OF THE DEFINITION

error on the transverse mass are reported in fig. 5.11. The same experimental estimators of the recoil used before are compared. The left plot shows that, as the maximum

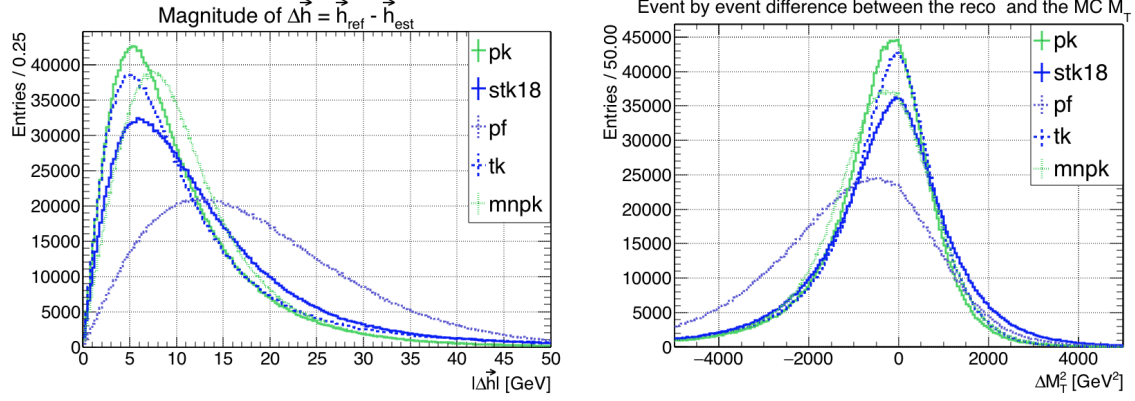


Figure 5.11: Resolution plot for different experimental recoil estimators. The magnitude of the 2D resolution on the recoil is reported on the left and the difference between the transverse mass computed with MC truth variables and reconstructed variables is reported on the right.

likelihood principle suggests, the estimator built using the peak of the predicted distributions for $P(\vec{h}|x)$ has very good performances: its distribution has higher peak, smaller tails and thus a smaller average. The mean has a value of 10.1 GeV for pk, 11.1 GeV for mnpk, 11.3 GeV for TK and 12.4 GeV for STK18.

As discussed in the previous chapter, scaling the TK recoil to recover on average the right modulus makes the estimator less performing and, indeed, the solid blue line of STK18 had a smaller peak.

A similar conclusion is addressed by the right plot which shows the ΔM_T^2 distributions. The PK recoil is slightly better but still comparable with the TK one. The mnpk recoil however presents a clearly broader distribution and is than expected to have a larger statistical uncertainty in the W mass fit.

Finally, it is interesting to compare directly the recoil magnitude distribution for the different estimators. Fig. 5.12 reports the generated spectrum in red, the classical recoil definitions in blue and the regression derived estimators in green. As well as for sake of comparison, that plot is important to underline that the event-by-event estimators do not have to reproduce the integrated spectrum but have to give a reasonable estimation of the recoil in a precise event. The peaky form of green distributions derives from the fact that the information passed to the regression is not enough to completely specify the recoil event-by-event. The estimators built are still affected by the generated recoil distribution. Indeed, the MNPK recoil is peaked around the mean of the red distribution and the PK recoil around its peak.

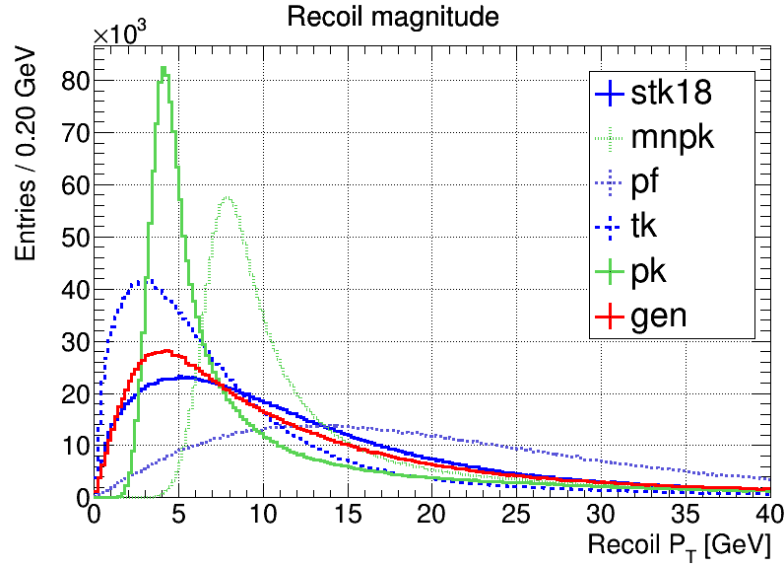


Figure 5.12: Recoil magnitude distribution for the different estimators. In red the true generated distribution, in blue the classical recoil definitions and in green the estimators derived from the regression.

The bottom line of all these control checks is that the new recoil definitions, obtained with the semi-parametric regression technique, are improving the resolution on magnitude and angle. The improvement is instead less sizable on recoil and transverse mass resolution. As a conclusion, these plots are not enough to judge the various recoil definitions: the performances have to be evaluated also in terms of systematic uncertainty on the m_W fit.

Chapter 6

Systematic uncertainties

The crucial point of the whole W mass analysis is the proper evaluation of systematic uncertainties. In this last chapter, the original work done in order to assess that problem is presented with the purpose of, first, quantifying the differences between the new and old definitions of the recoil and, then, showing a possible procedure which assesses the systematic uncertainties related to the recoil in a future W mass measurement.

The dissertation starts with an introduction of how systematic uncertainties arise in the measurement and what is the framework which allows to factorize part of them. The second section describes the procedure used to fit the W mass in this work.

In the third section, the importance of the model assumed for the boson production is discussed. Several definitions of the recoil are compared and ranked based on their sensitivity to the assumed W p_T spectrum. The value of that estimated systematics, obtained when different definitions of the recoil are used, is considered as a figure of merit to compare and rank the definitions themselves.

The remaining part of the chapter discusses the systematic uncertainties induced by imperfect description of the recoil in the simulation and how it is possible to modify the MC in order to increase the agreement with data. In sec. 6.4, recoil distribution in Z events is compared between data and simulation. In sec. 6.5 it is introduced a novel morphing technique which is used to correct the recoil in simulated events in which a Z boson decays into a muons pair. Data and simulated events with the same Z kinematics - as measured with the leptons - are particularly interesting since any difference here can only be due to modeling of the recoil. Given the similarity of W and Z production, is then possible to improve the simulation of W events, porting the corrections derived for the Z simulation to the W simulation.

In sec. 6.6 the necessary modifications to the Z simulation are derived and quantitative tests are presented to show the accuracy of the procedure. Finally, in sec. 6.7 possible problems in porting the corrections derived using the Z boson to the W simulation are discussed.

6.1 Systematic effects and factorization scheme

In the W mass measurement, systematic uncertainties arise from differences between data and simulation used to produce templates to fit the data distributions.

In the following, we assume that differences between data and simulation in the selection process produce negligible systematic effects in the distributions of interest. Eq. 4.1 shows that biases on the transverse mass distribution can happen only if the pdf $P(p_\mu, h, \Delta\phi(\vec{p}_\mu, \vec{h})|m_W)$ is different in data and simulation, due to imperfections in the simulation. In this pdf, muon and recoil variables are correlated through the boson kinematics. For this reason it is useful to split the variables of interest in the following groups:

- z : $\{p_T^{(W)}, \phi_W, p_z^{(W)}, M\}$, which identify the boson kinematic.
- x : $\{h_{tk}, N_{tk}, \dots\}$, which are the input variables to the recoil definition.
- y : $\{h, \Delta\phi_h\}$, which re-parametrize the recoil.
- t : $\{p_\mu, \Delta\phi_\mu\}$, which identify the muon kinematic.

where $\Delta\phi_v = \Delta\phi(\vec{v}, \vec{p}_W) = \phi_v - \phi_W$ stands for the angle in the transverse plane between the vector $\vec{v}$ and the flight direction of the boson.

With these definitions, the pdf of the transverse mass distribution can be split in the convolution of three pdfs:

$$P(M_T) \sim P(t, y) = P(t | z) \cdot P(y | z) \cdot P(z) \quad (6.1)$$

For a fixed boson kinematics z , recoil and muon can be treated as independent. This fact remarkably simplifies the estimation of systematic uncertainties and allows a factorization of muon and recoil.

The agreement of the boson kinematic distributions $P(z)$ between data and MC has to be fixed *a priori* and specifically at generator level, developing better calculations. However, the relevance of these discrepancies for the W mass analysis depends on the

ability to measure the recoil. As shown in fig. 4.7, the transverse mass distribution built with MC truth level variables has almost no dependence on the W p_T spectrum: a good definition of the recoil mitigates the requirement on the agreement of the W p_T spectrum between data and MC. In particular, in the next section the systematic uncertainty given by the accuracy of the simulation of the p_T spectrum is explored and used as figure of merit to rank the different definitions of the recoil.

In all the following discussions, two reasonable assumptions are considered: i) $P(y | z)$ does not depend on the boson mass M and azimuthal angle; ii) $P(y | z)$ is the same for W and Z events. When these holds, it is possible to improve the data-MC agreement learning the correction to be applied in Z events - where z variables can be measured - and then apply it on W MC.

Finally, given the goal of this chapter, in the following all the effects of $P(t | z)$ are neglected. At the same time, all the procedure presented are independent from $P(t | z)$, in the working hypothesis that it will be properly studied for the final m_W measurement.

Given the precision target of the W mass measurement, comparing qualitatively the transverse mass distribution when known discrepancy are injected is not enough. Even sub-percent differences may induce bias in the m_W parameter extraction of few MeV. To estimate the bias induced on the W mass a proper fit has been performed.

6.2 W mass fitting procedure

The fitting procedure here presented is not the one used for a CMS measurement. However, it is a good and fast approximation of that fit and is satisfying for the purpose of systematic uncertainty evaluation.

The fit is based on a template fitting procedure. It consists in the comparison between a histogram, representing the data or the pseudo-data transverse mass distribution, and several template histograms built using different mass hypothesis. Often in this thesis, pseudo-data and templates are obtained from two MC which differ for a known (i.e. injected) difference. In the MC used to build templates, different mass hypothesis distributions are obtained, starting from the unmodified one generated with $m_W = m_W^{(PDG)} = 80.385$ GeV, through a re-weighting procedure. Several sets of weights are created, one for each different masses considered. The weights for a given mass shift are computed considering the Breit-Wigner formula for resonances production: the idea is to assign to each event a weight w , which is the ratio of production probability in the desired m_W hypothesis divided by production probability in the $m_W^{(PDG)}$

6.2. W MASS FITTING PROCEDURE

hypothesis. In the approximation that all the factors in the production cross section do not change, except the denominator of the Bright-Wigner:

$$w = \frac{\sigma_{new}}{\sigma_{PDG}} \approx \frac{(M^2 - M_{PDG}^2)^2 + \Gamma_W^2 M_{PDG}^2}{(M^2 - M_{new}^2)^2 + \Gamma_W^2 M_{new}^2}$$

where $M_{new} = M_{PDG} + \Delta M$, being ΔM the shift in the W mass parameter and $\Gamma_W = 2.085 \text{ GeV}$ the boson natural width. This expression of the weight is just an approximate form because it does not take into account the relation between the mass and other parameters such as Γ_W , the couplings and so on. However, it holds under the assumption that $\Delta M \ll \Gamma_W$, which is verified in the presented situation since there is no reason to explore the region $\Delta M \gg 100 \text{ MeV}$.

The best fit value of the ΔM parameter is then chosen as the value whose relative hypothesis template better describes pseudo-data.

There are two crucial points on which this procedure stands: the definition of the histograms and the choice of the decision test used to indicate the best ΔM hypothesis. The former has been observed to have a sizable impact on the ΔM best fit value, but a negligible effect on the qualitative comparison between different recoils. This means that event selection and histogram feature will play a role in the exact determination of m_W but also that, for the scopes of this work, this procedure is stable enough to compare the recoil definitions.

It is not useful to exploit the full W mass spectrum since the m_W information is brought mainly by events in which the recoil is small and the muon is at $\eta \sim 0$ in the boson reference frame. Hence, the additional event selection applied on W events, on top of the one described in 4.3.1, is $33.4 \text{ GeV} < p_\mu < 52.7 \text{ GeV}$ and $h_{TK} < 17.6$: the additional efficiency is of 80%. This selection, resembling the selection applied in literature for the m_W measurement, has been developed for historical reasons on Z and then scaled for $M_W^{(PDG)}/M_Z^{(PDG)}$ to be applied on W . In addition, the histograms definition is variable independent: the transverse mass distributions, built with the different recoil definition, are taken only between the 25% and 90% quantile and split into 50 bins. The quantile cut is performed because the tail of the spectrum are very sensitive to the production mechanism and little sensitive to the boson mass.

Pseudo-data are scaled such that the number of events, after the selections declared in sec. 4.3.1, is 25M: this number represents the expected sample size when the whole CMS statistics is used. Templates histograms are then forced to the same normalization of pseudo-data.

The decision test, chosen to pick the best ΔM hypothesis, is the likelihood ratio test.

In practice, for computational reasons, it has been effectively approximated with a χ^2 scan. A similar approach holds when the bin content is large enough to approximate the Poisson distribution with a Gaussian one. This is the case for the transverse mass histograms since they are filled with about 1M of events divided almost evenly into 50 bins and no tails are present because are removed by the cut on quantiles (example in fig. 6.3 right).

For each mass hypothesis template, the likelihood to the data is quantified by the χ^2 value, computed as

$$\chi^2 = \sum_{i=1}^{50} \frac{(n_i^T - n_i^D)^2}{n_i^T}$$

where i runs on the bin number and $n_i^T(n_i^D)$ refers to the template (data) histogram i -th bin content.

The χ^2 values obtained in the different mass hypothesis are then fitted with a parabola $f(\Delta M) = a \cdot \Delta M^2 + b \cdot \Delta M + c$. The best fit value of mass displacement is determined as the minimum of the parabola $\Delta M^* = -\frac{b}{2a}$ and the expected statistical uncertainty as the value in which the X^2 is 1 larger than its minimum, $\sigma_{stat} = 1/\sqrt{a}$. The uncertainty on these two parameters, given by the parabola fit uncertainty, has been checked to be always negligible.

One of the main goal for which this fit is performed in this thesis is to have a comparison between different experimental definitions of the recoil, so the fitting procedure is often performed in parallel for different choices of the recoil estimator (i.e. TK, PF, mnpk, ...).

Since templates and pseudo-data are always chosen in a way in which they are statistically dependent, it has to be remarked that the cited σ_{stat} is not a value representing the variation of ΔM^* in this fit. However, since in all the histograms used $\sqrt{n_i}/n_i \ll 1$, σ_{stat} is a good estimation of the statistical uncertainty on m_W , when the MC statistics used to build the templates is much higher than the one present in data (i.e. can be neglected) and the data sample is of 25M (after the selections in sec.4.3.1).

6.3 Effect of boson kinematic mismodeling

This section is dedicated to the study of the systematic error induced on the W mass measurement by the differences in the boson p_T spectrum between data and MC. For a visual comparison of data-MC differences, a sample of Z decaying in two leptons is used as a proxy of the W . Fig. 6.1 shows the spectrum of reconstructed Z bosons (i.e.

the sum of four-vectors of the reconstructed muon pair) in data and MC. It is clear

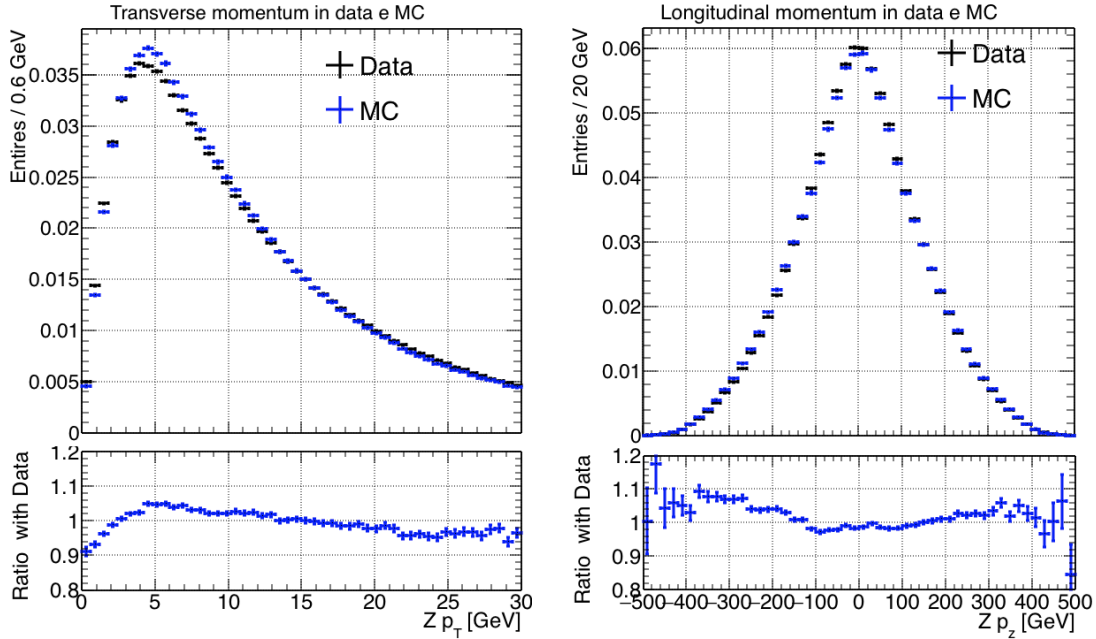


Figure 6.1: Comparison between reconstructed Z kinematic distributions in data and MC: the transverse momentum is shown on the left and the longitudinal momentum is shown on the right.

that discrepancies of order 10% are present. Even if this is not the ultimate MC that will be used for the measurement, it is useful to consider the present situation as a possible and conservative scenario.

In the following subsection, a sample of W MC is used to compare the effectiveness of different recoil definition in mitigating the transverse mass dependence on the W p_T spectrum.

6.3.1 Estimating the systematic uncertainty

To estimate the bias induced on the W mass from the differences in the boson p_T spectrum, a fitting procedure, as described in sec. 6.2, has been performed. First of all, it has to be noticed that, considered the present literature (see sec. 3.2.3), the expected size of the effect on m_W is of $O(10 \text{ MeV})$. This makes it sub-leading to the expected statistical uncertainty, which given the available MC sample is expected to be about 25 MeV. The statistical uncertainty would have been a limiting factor to investigate the systematic of interest and then a workaround has been adopted to isolate the W p_T effect. The idea is to fit a W MC sample with templates made by a copy of itself. In this way the effect of the finite statistics in pseudo-data and templates is canceled by

6.3. EFFECT OF BOSON KINEMATIC MISMODELING

the fully statistical dependence of the two samples and the limited amount of available MC is enough to have a meaningful result.

In order to evaluate the effect of the boson p_T , pseudo-data have been re-weighted to artificially modify the p_T spectrum. Different sizes and shapes of the re-weighting have been tried: fig. 6.2 reports, in the central and in the left plot, the ratio between the distribution of pseudo-data W p_T , before and after the re-weighting, for two different scenarios. The size of the re-weighting is quantified using a percentage value: that is

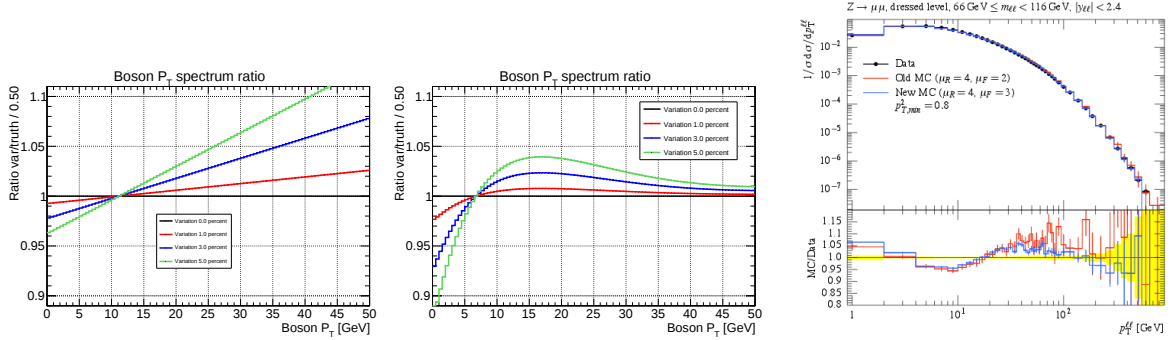


Figure 6.2: Left and center: Ratio between the pseudo-data p_T spectrum before and after the re-weighting. A linear discrepancy of different size has been tested with the left plot, whereas a power law exponentially dumped has been tested in the central plot. Right: For sake of comparison, it is here reported the ratio of the p_T distribution form Z data and a pythia only MC tested by the CMS W mass analysis group.

because the weights are first computed with a given shape and an arbitrary size and then scaled by the reported percent factor before being applied. The percent value has been chosen in a way that resembles the maximum size of the effect.

Even if the artificial modification may seems sizable, it induces a very tiny variation of the p_T spectrum and an even smaller variation on the transverse mass template. Fig. 6.3 shows the distribution of the boson transverse momentum and transverse mass for the linear re-weighting of the pseudo-data presented in the left plot of fig. 6.2. It is clear that the effect on W mass can not be appreciated just looking at ratio plots but only a proper fit can give a quantitative estimation of the differences induced by the re-weighting.

Once that pseudo-data have been obtained following the above procedure, an unmodified copy of the same MC - i.e. exactly the same, statistically fully dependent - is used to build templates. In fig. 6.4 the χ^2 scan for the STK18¹ recoil is reported for four different linear re-weighting sizes. The full circle represents the χ^2 value of the template made in the hypothesis of the ΔM value reported on the $\hat{x}$ axis, whereas

¹TK recoil with magnitude scaled by a factor 1.8.

6.3. EFFECT OF BOSON KINEMATIC MISMODELING

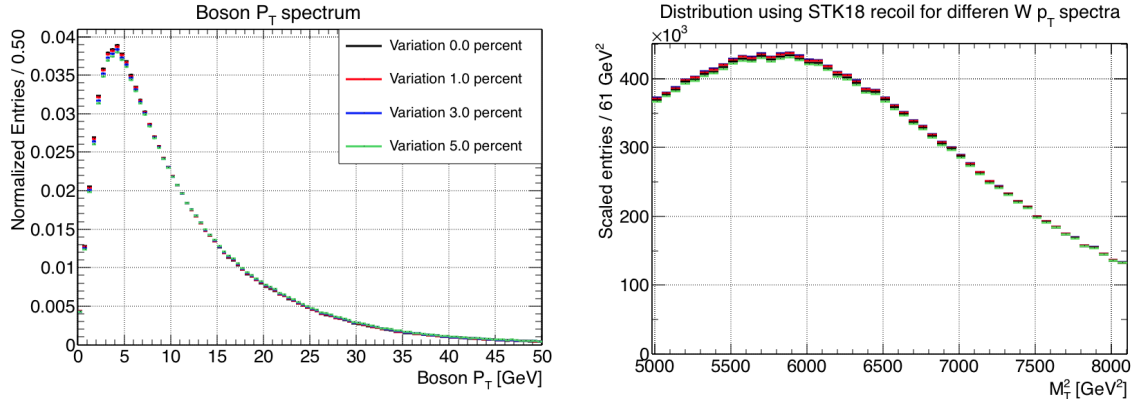


Figure 6.3: Effects of the linear re-weighting procedure on pseudo-data distributions of boson momentum (left) and transverse mass (right) are reported for different re-weight size. The left plot is normalized in the full range, whereas the right plot content is scaled by a factor 10 in order to resemble the expected CMS statistics.

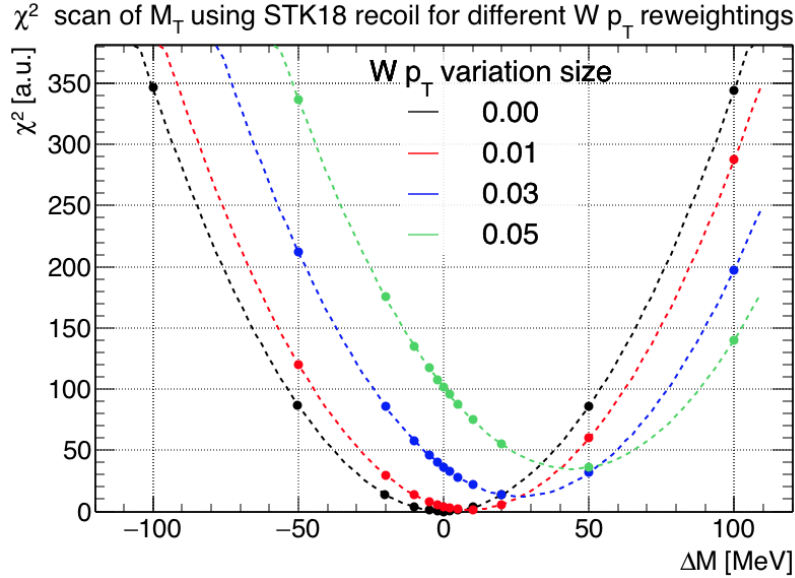


Figure 6.4: Value of the χ^2 for the mass templates in the range $-100 \text{ MeV} < \Delta M < 100 \text{ MeV}$, when compared to pseudo-data re-weighted in order to have a linear variation of the $W p_T$ spectrum, as shown in fig. 6.2 left.

different colors refer to the values of the coefficient in front of the re-weighting: the higher is the number in the legend, the more the spectrum is modified. The colored dashed curves reported are the best fit parabola, which clearly give a good description of the points. At the end of this procedure, a set of $\{\Delta M^*, \sigma_{stat}\}$ values is obtained, one for each $W p_T$ variation size.

The fitting procedure, presented above using STK18 as example, is repeated for transverse mass built using different definitions of the recoil. Fig. 6.5 shows the trans-

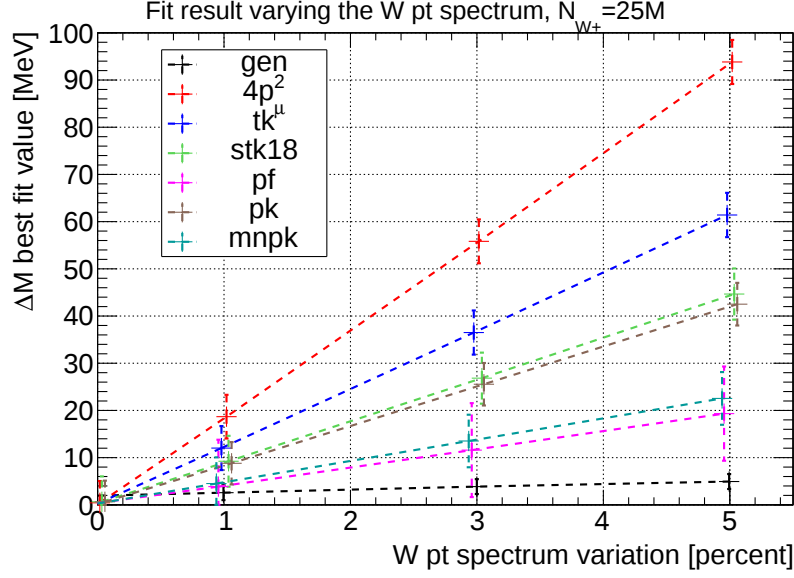


Figure 6.5: Transverse mass fit results for different estimators, obtained fitting half of the W MC sample. The abscissa reports the best fit value ΔM^* obtained in each scan at different W p_T re-weighting sizes. The error reported on it represents the σ_{stat} value. The best fit line for each set of points is reported dashed.

verse mass fit results: it reports on the abscissa ΔM^* obtained in each scan - with the σ_{stat} value as its error - and on the ordinate the re-weighting size. Given the small size of the W p_T variations, it is possible to fit the point coming from the same recoil definition with a straight line. What is more, the line is not even spoiled by the statistical fluctuation since templates and pseudo-data are statistically dependent. An important cross check is that, when no re-weighting is applied, the χ^2 at its minimum is 0 and this happens for $\Delta M = 0$. This fact is reflected in the minimum position of the black parabola in fig. 6.4 and in the 0 $\hat{y}$ -axis interception of all the lines in fig. 6.5. It is also interesting to notice that the σ_{stat} value does not significantly depend on the W p_T re-weighting: all the error bars for a given color are equal (still fig. 6.5).

At this point, the slope of the fitted line can be used as parameter to quantify the dependence of the W mass analysis on the W p_T spectrum knowledge. In particular, the value of the slope is a proxy to assess the recovered correlation between recoil and muon.

As anticipated before, the transverse mass distribution is not completely independent from the W p_T , even when the true MC recoil is used. As the slope of the black line of fig. 6.5 suggests, the bias on m_W in this case is very small compared to the one for more realistic definitions. A comparison between the performances of the recoil definitions can be already foreseen, but a deeper discussion is presented in the next subsection.

6.3.2 Recoil definitions ranking

The different recoil definitions can be compared in terms of goodness in mitigating the systematic uncertainty on the W mass, through a comparison of the slopes discussed above. They express an expected bias on the fit in terms of MeV each percent of variation. Even if the output number can not be taken as the effective bias present in a real measurement (the bias depends on the MC and on the fitting procedure that will be eventually used), the comparison of these slopes can be taken as a qualitative way to rank the different definitions. However, it is important to consider simultaneously the expected statistical uncertainty too. Indeed, it is possible to build variables less dependent on W p_T , but also with a small dependence on m_W and, thus, less effective in the measurement. A trivial and extreme example of this behavior is to fit, instead of transverse mass, a variable not correlated with the W (e.g. fitting the distribution of the number of PU vertexes): it may be not dependent on W p_T , but the χ^2 will be completely flat. A similar effect is also obtained when the recoil is badly measured and the random component summed to the W p_T is bigger than the muon momentum. For each recoil variable used, in fig. 6.6 the bias slope on the abscissa and the expected statistical uncertainty on the ordinate are reported, being the latter estimated as the σ_{stat} when no re-weighting is applied.

A subtle point in the whole procedure is that, even if fully statistical dependent samples are used for pseudo-data and templates, a statistical uncertainty in the determination of the ΔM^* slope and σ_{stat} is present due to the finite MC sample used. To evaluate this effect, the W MC sample has been divided randomly into 100 sub-samples. The value of the ΔM^* slope and σ_{stat} is then obtained for each sub-sample, building distributions of these two observables for each recoil definition used. What is then reported in fig. 6.6 is the mean of those distributions, as central value, and a scaled unbiased estimator of the variance, as error box. ΔM^* slope and σ_{stat} are treated as uncorrelated and the error bar is given by the expression

$$\hat{s}_x = \frac{1}{\sqrt{N}} \frac{\left(\frac{N-3}{2}\right)!}{\left(\frac{N-2}{2}\right)! \sqrt{2}} \sqrt{\sum_{i=1}^N (x_i - \bar{x})^2}$$

where $N = 100$ is the number of sub-samples and x is either the ΔM^* slope or σ_{stat} .

The ranking presented in fig. 6.6 can be interpreted as the fact that the recoil definitions, developed on top of the DNN recoil regression, are behaving well in the mitigation of the systematic uncertainty due to the knowledge of the W p_T spectrum. On one hand, when the mnpk recoil is used, the fit presents almost the same statisti-

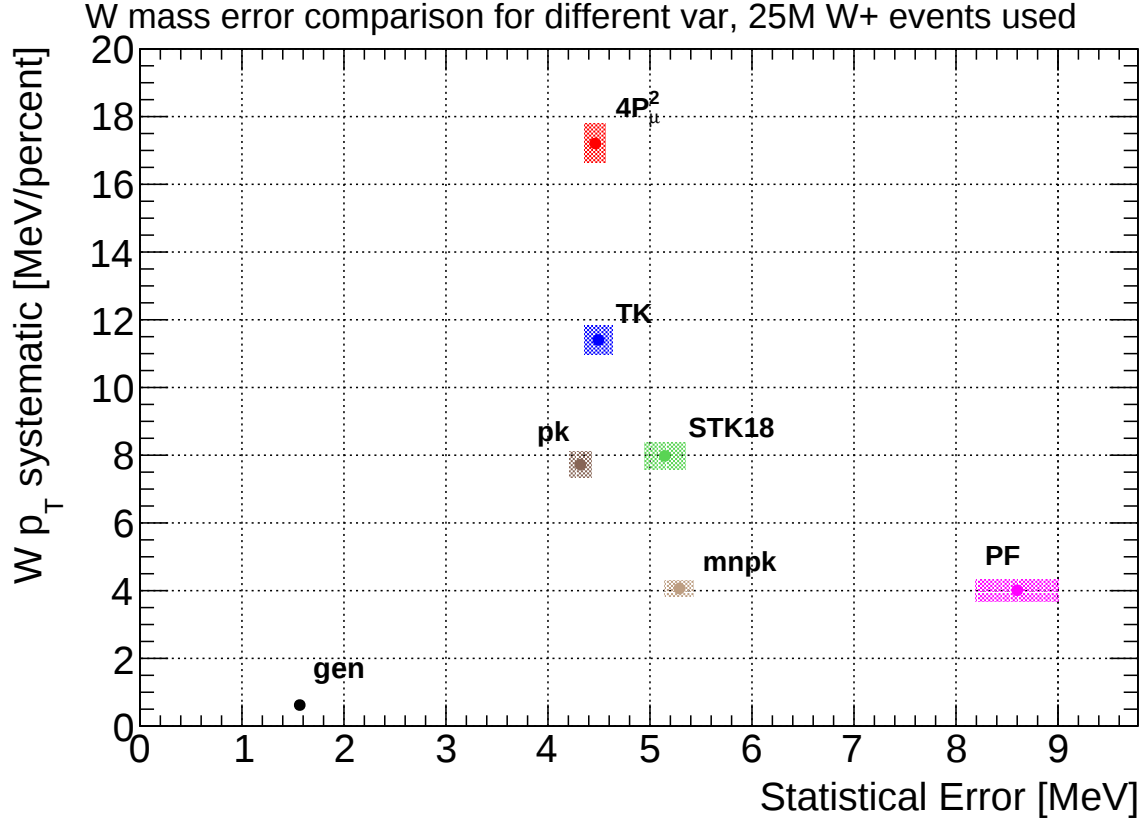


Figure 6.6: The variables ranking in terms of statistical and $W p_T$ systematic uncertainty. The error box reported represents the statistical uncertainty due to the finite size of the MC sample used.

cal uncertainty as when TK recoil is used, but with $1/3$ of systematics. On the other hand, when mnpk is used, the systematics arising is almost the same as when the PF recoil is used, but with about $1/2$ of statistic uncertainty. Even if almost arbitrary $W p_T$ variation have been used, they have a reasonable shape and are of the same order of magnitude of the expected ones. The results are presented only for linear re-weighting of the $W p_T$ spectrum, but other shapes have been tried and no significant difference has emerged.

Once again, it is useful to keep in mind that the ranking here presented is just a qualitative comparison between different experimental recoil definition; a proper systematic treatment needs to be further developed for a future W mass measurement.

6.4 Systematic on the recoil mismodeling

The previous section has shown the ranking of different recoil definitions based on their sensitivity to the assumed W p_T spectrum. Another criterion is the dependence of the estimated recoil value on the accuracy of the MC simulation. Put differently, it considers how sizable are the data-MC discrepancies in the distribution of the estimated recoil, when the same boson momentum is fixed. For example, it is assumed that the MC overestimates the tracks momentum by 10% and underestimates the neutrals momentum by 50%. Then, the recoil definitions behave as follows: the TK recoil will be on average 10% bigger in MC than in data, whereas PF recoil will be about 40% smaller. The usage of PF will induce bigger differences on the transverse mass distribution and, thus, a bigger bias in the fit.

As a second example, it is assumed now that MC describes perfectly the sum of the tracks momentum but overestimates their number, which is directly proportional to the boson momentum. In this case, TK recoil behaves the same in data and MC and no bias is present on m_W . The MNPK recoil, instead, does not behave the same and its usage brings a bias in the m_W fit. Indeed, for a fixed number of tracks, the regression prediction is the same and so is the $mnpk$ recoil. However, the true boson value is different in data and MC. Given the factorization scheme of eq. 6.1 and assuming that $P(z)$ will not exactly compensate the differences in $P(z | x)$, a difference in $P(x | z)$ is induced. As discussed later, this reflects directly into differences in $P(y | z)$, which bring differences on the transverse mass distribution and, eventually, a bias in the m_W fit.

The remaining sections go through the systematic uncertainties induced by the recoil modeling and report a trial of MC calibration.

The W mass measurement reports a bias if $P(y | z)$ is not equal in data and MC. Since $y = f(x)$,

$$P(y) = P(x) \frac{1}{f'(x)} = \frac{P(x | z)P(z)}{f'(x)}.$$

So, assuming a perfect description of $P(z)$, the discrepancy between DATA and MC in $P(y)$ arises from a discrepancy in $P(x | z)$; such discrepancy is present and is due to mismodeling that happens in hadronization, fragmentation and detector simulations. Even if the source of the recoil disagreement comes from the x space, it is more convenient to treat directly the x -induced differences on the y space. That is because, first of all, it reduces the space dimensions and, secondly, only the y variables enter in the W mass measurement. Discrepancies among the x not affecting the y can be neglected.

The core idea is to evaluate how much Z data and MC differ from each other and to fit corrections of the y variables, conditional to the z variables, in order to remove this difference. Here, and from now on, for z is meant only a subset of the previously defined variables made of $\{p_T, p_z\}$.

In this reduction of the z space, an assumption is introduced: the y variables probability has to be independent from the mass of the boson and from the boson azimuthal angle. The bottom line is that this should derive from the cylindrical symmetry of both production and detector. If such assumption is not verified, the procedure is not complete and further morphing has to be developed to correct the remaining disagreement.

This procedure of fitting corrections can be performed only on Z , since z variables are not measured in W data. However, given the similarities in the production mechanism between Z and W , the correction derived on the Z sample can be ported on the W one. Indeed, both W and Z are produced by a quark-antiquark annihilation and no main differences in the recoil, at fixed boson momentum, are expected.

For this and 6.6 whole sections, only Z data and Z MC are used.

6.4.1 Differences between data and simulations

Differences between data and MC are present in the x variables, but they are very hard to investigate. In fig. 6.7 the ratio plots between data and MC distribution, in two example variables, are reported for different bins of the Z boson kinematics. The bins are defined by cuts in the two variables $p_T^{(Z)}$ and $p_z^{(Z)}$: two cuts in each variable are set and then combined to form the four possible regions, indicated in the plots by the four different colors. Even if the bin size is relatively big (see the plot's legend and compare it with fig. 6.1), the statistics available is not enough to clearly identify differences and, on top of this, it is even harder to foresee what is the effect on the rest of the W mass analysis.

As discussed before, looking directly at discrepancies between y variables is more effective: those are the variables which should be really distributed in the same way both in data and MC. Before discussing quantitatively the recoil modeling effect, it is useful to introduce the correction procedure which will be used to improve the agreement between data and MC. It is called *multi-dimensional morphing* and it practically consists in the derivation of a function which changes the values of the recoil scale and angle in each MC event, in order to have a distribution of $P(y | z)$ similar to the one in data.

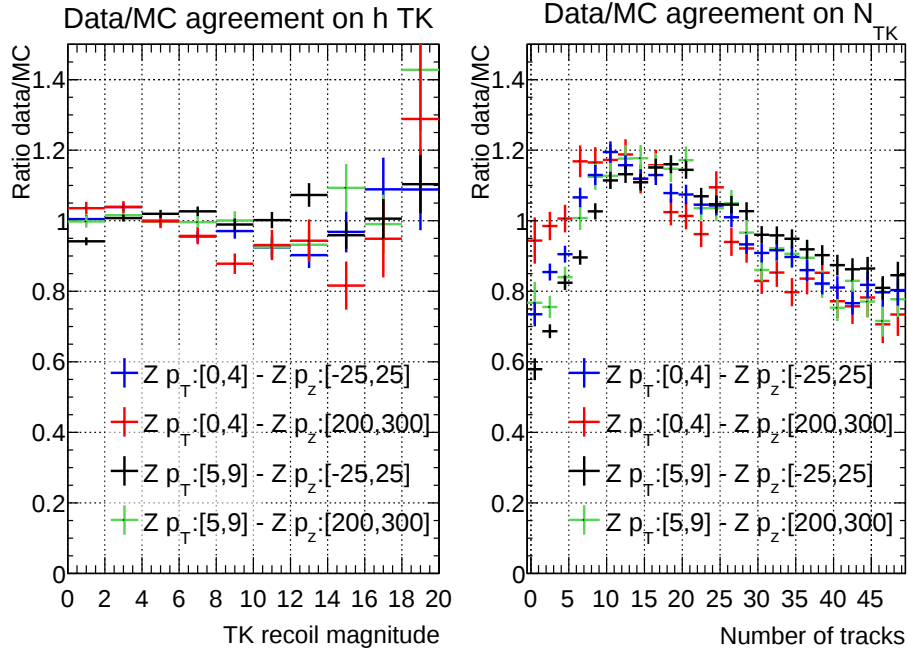


Figure 6.7: Ratio plot of the distribution of two example x variables (TK recoil magnitude on the left and number of tracks on the right) in four different bins of the boson kinematic phase space.

6.5 A multi-dimensional morphing procedure

The interest in improving the data-MC agreement goes far beyond the W mass analysis. In high energy physics this is a recurrent problem and its applications range from precision measurement to searches.

The way to pursue the improvement presented in this section consists in the derivation of a function, called *morphing function*, which transforms the variable of each event in order to restore a general agreement in the desired distribution(s). The explicit form of that function is fixed by a machine learning algorithm and its parameters by a dedicated training.

One of the main technical achievements of this thesis, after the convergence of the DNN semi-parametric regression, is the development of a general purpose multi-dimensional morphing, where multi-dimensional indicates that more than one variable is simultaneously changed in each event. Since the purpose is more general than the single W mass, the procedure is discussed in a conceptual framework.

Given two sets of variables² $\{y, z\} \sim f(y, z)$ and $\{y', z'\} \sim g(y', z')$, both with values

²The symbol $\sim$ means "distributed as". This implies that both f and g are functions $\mathcal{R}^n \times \mathbb{R}^m \rightarrow [0, 1]$ and are normalized to 1: they are two probability density functions.

in $\mathbb{R}^n \times \mathbb{R}^m$, having a morphing procedure means to determine a function $M : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that the conditional probability $P(M(y) | z = z_0)$ is equal to $P(y' | z' = z_0)$ for every acceptable values of z_0 . The new y variables obtained after the application of the morphing function $\bar{y} = M(y)$ are called morphed variables. In other words, the morphed variables $\{M(y), z\} \sim h(M(y), z)$ are such that $g(y, z)|_{z=z_0}$ and $h(y, z)|_{z=z_0}$ have the same shape (i.e. different just for a normalization factor) for each fixed value of z_0 .

In short, morphing a variable into another means to change the first one in order to have the same distribution of the second one, potentially in a way which is conditional to some other variables.

That is exactly what we are interested in for the MC calibration of the W mass analysis. Indeed, we want to check the differences in $P(y | z)$ between data and MC of Z events, learn how to modify the MC to better describe data and finally apply the learned correction to the W MC. In particular, the recoil variables have to be morphed conditionally to the boson transverse and longitudinal momentum. The morphing function is the correction that we need to apply to MC in order to make it more similar to the data.

Quantile mapping

The problem of defining a morphing function that maps a variable into another is usually faced, in the 1D case, by a quantile mapping procedure. Given two variables $x \sim f(x)$ and $y \sim g(y)$, the morphing identifies the function

$$M(x) = G^{-1}(F(x))$$

as the one such that the transformed variable $x' = M(x)$ is distributed as $x' \sim g(x')$, where capital letters F and G indicate the cumulative distribution function (CDF). Applying the function M , as above defined, means essentially to map a value of the x variable with the one having the same CDF value in the y distribution, as illustrated in fig. 6.8.

This procedure is well defined only in 1D because there is no unique definition of the CDF in higher dimensions. However, it is possible to extend the quantile mapping concept in order to reach the goal of a multi-dimensional morphing.

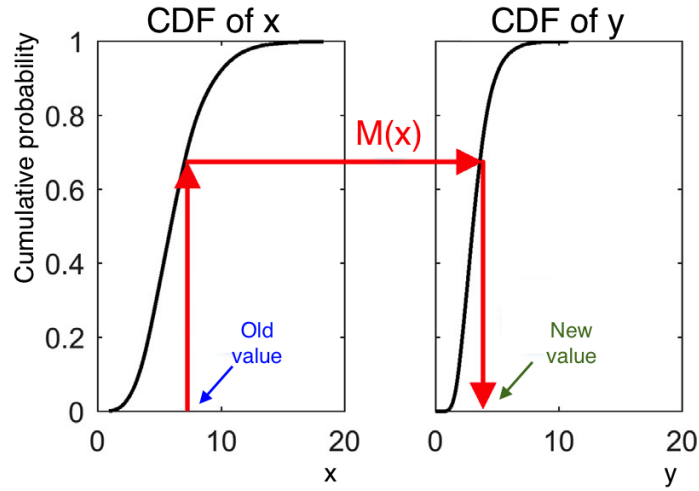


Figure 6.8: Sketch of the quantile mapping procedure.

6.5.1 Procedure for a n -dimensional k -conditional morphing

Before describing the procedure, it is useful to recap what a quantile regression is and how it works. After that, the dissertation will go through the details of the quantile morphing procedure developed in this work.

Quantile regression

Quantile regression [51] is a type of regression analysis used in order to estimate the quantiles value. In particular, it is used to fit a function $f_\tau(x)$, which describes the value of the τ -th quantile of the distribution of a y variable, conditionally to a set of variables x .

Since in principle the number of dimensions of x can be arbitrarily high, a machine learning algorithm is used to describe the functional form of $f_\tau(x)$. Specifically, a BDT [52] is used in the following, even if approaches with DNN have been tried too. Practically, if the goal is to fit the value of the τ -th quantile of $P(y | x)$ using a function defined with a BDT, it is necessary to choose the BDT parameter in a way in which the following loss function is minimized:

$$\text{loss} = E_y[\rho_\tau(y - u(x))]$$

where τ is a fixed quantile, $u(x)$ is the BDT prediction, E_y is the expectation value on the y sample, the function $\rho_\tau(z) = z(\tau - \mathbb{I}_{(z < 0)})$ and $\mathbb{I}$ is the indicator function. A straightforward informal demonstration can be obtained directly minimizing the loss

function

$$\frac{\partial}{\partial u} (E_y[\rho_\tau(y - u)]) \big|_{u=q} = -(\tau - 1) \int_{-\infty}^q f(y) dy - \tau \int_q^{+\infty} f(y) dy = 0$$

which can be re-written as

$$(1 - \tau)F_y(q) - \tau(1 - F_y(q)) = 0$$

and that is solved by the value of q which satisfies $F_y(q) = \tau$.

When the BDT training (i.e. fitting) phase is over, the function $f_\tau(x)$ can be used to predict the position of the τ -th quantile of $P(y | x)$. A specific training for each value of τ is needed, since the τ value is fixed to a given value in the loss function.

The morphing procedure

To achieve a morphing procedure on a multi-dimensional distribution of a set of variables y and conditional to the set of variables x , the idea is to merge the quantile mapping with the quantile regression. Quantile mapping is applied on the quantile regression output. In short, the method to map the full space proceed recursively:

- The morphing for the first variable $y_1 \in y$ is fitted conditionally to the x space.
 - The morphing is performed with a quantile mapping, where the quantile values are provided by the quantile regression.
- The morphed set of $\bar{y}_1$ is computed.
- The morphing for the second variable $y_2 \in y$ is fitted conditionally to the $\{x, \bar{y}_1\}$ space.
- The morphed set of $\bar{y}_2$ is computed.
- The morphing for the third variable $y_3 \in y$ is fitted conditionally to the $\{x, \bar{y}_1, \bar{y}_2\}$ space.
- ...

The choice of the variables order is free and the final result is always to have a morphing of $P(y | x)$.

A more detailed description follows: the notation used refers to sec. 6.5.

The starting point consists in training two sets of quantile regressions: one set for the

quantiles of the conditional distribution of the y_i ; one set for the target associated y'_i variable. For example, nine quantile regressions are trained for y_1 and nine for y'_1 , addressing the values of $\tau = 0.1, 0.2, \dots, 0.9$. For each value of τ and for each variable a BDT is trained. When the BDTs training phase is over, 18 functions are calculable: $q_{1,\tau_i}(x)$ with $i = 1, \dots, 9$, one for each value of the quantile of $P(y_1 | x)$; $q'_{1,\tau_i}(x)$ with $i = 1, \dots, 9$, one for each value of the quantile of $P(y'_1 | x)$.

The morphed values of y_1 , called $\bar{y}_1$, are then computed for a fixed value of $\{y_1, x\}$: i) the points in a 2D *old-new* space are built pairing the values of the quantiles of y and y'

$$\{\{q_{1,\tau_1}(x), q'_{1,\tau_1}(x)\}, \dots, \{q_{1,\tau_9}(x), q'_{1,\tau_9}(x)\}\};$$

ii) a function M_1 is built interpolating with monotonic piecewise cubic Hermite polynomial the above set. Outside the first and the last predicted quantile, linear extrapolation is used; iii) $\bar{y}_1$ is computed as $\bar{y}_1 = M_1(y_1)$.

The computation is performed for all the events (i.e. for all the $\{y_1, x\}_i$) and the full set of $\bar{y}_1$ is obtained.

Now, to take care of the $n - 1$ dimensions, which still need to be morphed, the same procedure is repeated, adding the already morphed variable(s) $\bar{y}$ together with the conditional variables x .

For example, a second dimension y_2 is chosen and the probability $P(y_2 | x, \bar{y}_1)$ is morphed to be the same as $P(y'_2 | x, y'_1)$. This procedure is repeated until all the $\bar{y}$ have been computed. All the morphing function can be joined to obtain the function M which transforms the y in

$$\bar{y} = M(y, x) = \{\bar{y}_1, \bar{y}_2, \bar{y}_3, \dots\} = \{M_1(y_1, x), M_2(y_2, x, \bar{y}_1), M_3(y_3, x, \bar{y}_1, \bar{y}_2), \dots\}$$

The obtained $\bar{y}$ has the same x -conditional probability density function as the y' .

It has to be remarked that the order in which the y variables are morphed does affect the explicit form of M , but does not affect the results.

6.6 Recoil modeling effect and calibration

A slightly better qualitative view, with respect to comparing the x variables, can be given by the comparison between the distributions of the y variables. In fig. 6.9, four different plots are reported showing the comparison between data and MC in the two y variables, for the TK recoil definition and for the mnpk definition. The recoil

6.6. RECOIL MODELING EFFECT AND CALIBRATION

distribution after the morphing of y , conditionally to z , is also reported. It has been fitted in order to morph the Z MC into Z data. The three different colors identify the

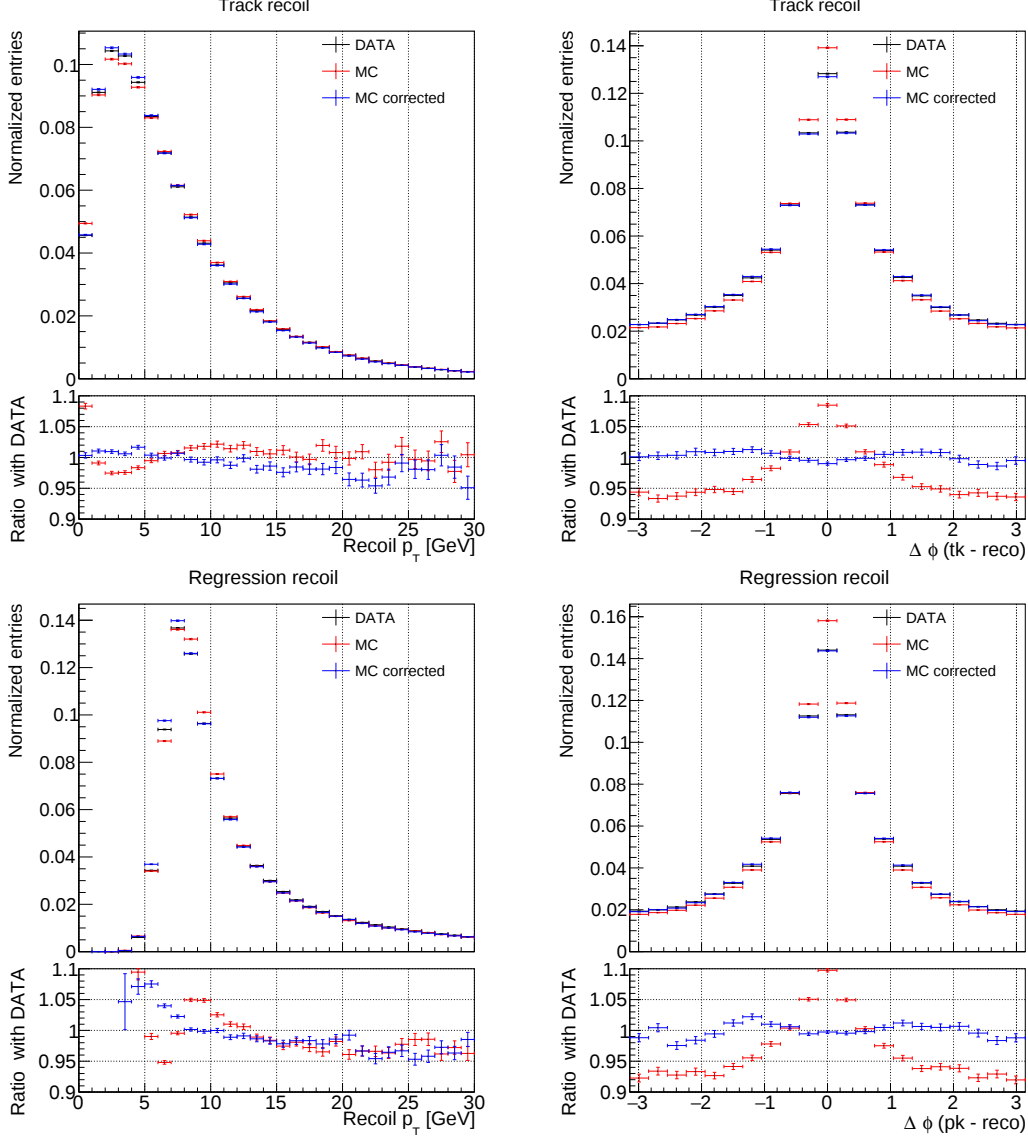


Figure 6.9: Recoil variable distributions for the TK (top) and mnpk (bottom) recoil definitions. For each variable the three different colors identify the data distribution (black), the bare MC distribution (red) and the MC distribution after the morphing procedure (blue). Ratio plots are also shown below each plot between the data and the other two distributions.

distributions of data, bare MC and MC after that a morphing procedure, as the one explained above, has been applied to recover the differences in the y variables. As it is clear from the ratio plots, the agreement for the TK variables is slightly better and this is of course due to the fact that less factors enter in the definition and, thus, less discrepancies can be induced. In addition, it is clear that the morphing procedure has

remarkably improved the agreement between data and MC, lowering the ratio to the percent level. A particular feature interesting to underline is the disagreement in the lower part of the recoil magnitude spectrum for *mnpr* recoil (bottom left plot, around 5 GeV). This effect can be reduced by increasing the number of quantiles in that region: the wiggles are due to the approximation introduced by the morphing interpolation or extrapolation.

Even if from these plots it can be seen that the corrections to the recoil are performing well, it is hard to understand if the agreement level is good enough for the W mass analysis and at what bias size can induce a similar situation on the m_W parameter.

A proper way to quantify the data-MC agreement for what matters the W mass measurement is to perform the fit for the m_W extraction. For this purpose, a machinery similar to the one used in sub. 6.2 is used. Specifically, it has to be remarked the importance of using statistical dependent samples to create pseudo-data and templates, since the bias remaining after the morphing will hopefully be very small and the finite statistics effect might be large enough to cover it.

In the following, all the morphings have been performed on the two $y = \{h_r, \Delta\phi_r\}$ variables (r being *TK*, *PF* or *mnpr*), conditionally to $p_T^{(Z)}$ and $p_z^{(Z)}$.

6.6.1 Procedure check

In order to verify the whole morphing plus fitting procedure, a preliminary closure test has been performed using Z MC only. The full sample has been randomly split into three, nominally MC0, MC1 and MC2. Then, two morphing functions have been fitted: $M_{0 \rightarrow 1}$, which should transform the MC0 into MC1; $M_{0 \rightarrow 2}$, which should transform the MC0 into MC2. These two functions have been applied on MC0, practically obtaining three copies of MC0: the original unmodified one $MC0_0$; $MC0_1$ ($MC0_2$) obtained when $M_{0 \rightarrow 1}$ ($M_{0 \rightarrow 2}$) is applied. The three $MC0_i$ are fully statistical dependent and the only differences are introduced by the morphing function. Indeed, the three copies are identical except for the y variables, which show differences introduced by the morphing. These differences belong to two types: random fluctuations, given by the statistical independence of the samples used to fit the morphing functions; and, potentially, bias effects introduced during the morphing fitting and application, due to imperfect convergence, interpolation or extrapolation.

To quantify the differences among the copies of MC0, three fits, similar to the ones used in subsec. 6.2, have been performed. The fitted distribution is the one of the transverse mass computed using the *mnpr* recoil. The bottom line is to use a MC0

6.6. RECOIL MODELING EFFECT AND CALIBRATION

copy as pseudo-data and a different MC0 copy to build the templates. For each possible combination of MC0 copies, the sample is further split in 100 sub-samples and the fit performed in each of them. The bias distribution on the mass resulting from the fit is reported in fig. 6.10 for all the combinations. It is clear that, the expected

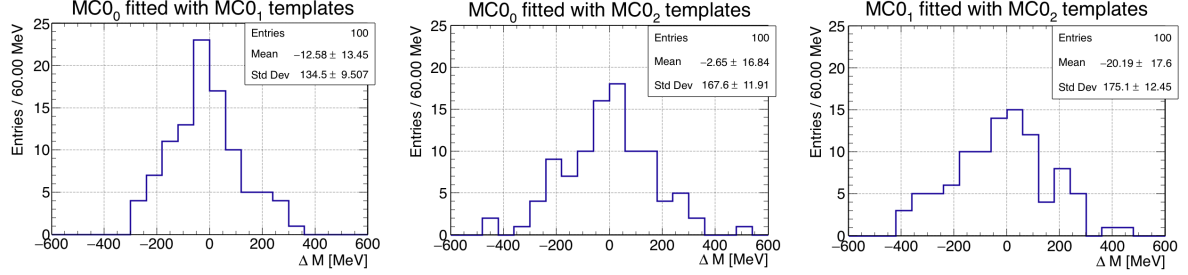


Figure 6.10: Distributions of the bias obtained from the fit to the transverse mass, computed with the *mn*pk recoil of 100 sub-samples for each combination of the MC0 copies. In the left and central plot, MC0₀ is used as pseudo-data and respectively MC0₁ and MC0₂ as templates. In the right plot, MC0₁ is used as pseudo-data and MC0₂ as template.

result for a positive output of the test, is to have all the three distributions with a mean compatible with 0. This is the case, since the three obtained distribution means are (referring to the plots from left to right): -12.6 ± 13.5 MeV, -2.7 ± 16.8 MeV and -20.2 ± 17.6 MeV.

The uncertainty on the mean, of order 20 MeV, is due to the limited amount of MC available. The size of this uncertainty - of the same order for all the fits that will be presented in the following - fixes the level of precision at which the calibration can be tested today. Larger MC samples will be needed for the final measurement.

6.6.2 Systematic assessment before calibration

A morphing function between Z data and Z MC has been derived for three different sets of y variables: TK, PF and *mn*pk. The morphing has been fitted conditionally to the reconstructed $p_T^{(Z)}$ and $p_z^{(Z)}$ of the di-muon system. These morphing functions are meant to be the MC calibration for the corresponding recoil. They have been applied to the Z MC obtaining a *data-like* copy with all the same variables, except the recoil ones given, instead, by the morphing.

To assess the systematic present when no morphing is applied, it is possible to perform a m_W fit to pseudo-data made with the *data-like* copy of the Z MC, with templates made with the unmodified copy (i.e. bare Z MC). Once again, pseudo-data and templates are statistically dependent and the only differences are introduced by the morphing procedure. As before, the samples are split in 100 sub-samples and the mean, with its

uncertainty, is used as estimator of the bias. Tab. 6.1 reports the results in the case in

Bias on Z mass before MC recoil calibration		
Recoil estimator used in M_T^2	Average bias [MeV]	Average bias error [MeV]
mnpk	140	14
PF	-80	24
TK	28	11

Table 6.1: Average mass bias and its uncertainty obtained in the fit between data-like Z MC and Z MC. The result obtained when different experimental recoil estimators are used is reported.

which the three different recoil definitions are used. As expected, the more information a definition uses, the higher is the bias. Specifically, none of the variables reports a bias compatible with 0 and all of them have a average value which is much higher than the precision goal of the W mass analysis.

This is the reason why the morphing is really necessary and a recoil calibration is of crucial importance.

6.6.3 Systematic assessment after calibration

Even if the morphing between data and MC developed above has been used to asses the systematics when the bare MC is used to fit data, actually it is the correction to be applied to the recoil on MC in order to better describe data. In principle, the same morphing function used to obtain the mnpk recoil in the *data-like* Z MC can be applied to the W MC and then use the latter to build templates for the fitting of W data. Given this, it is interesting to understand at what level the morphing has converged and what is the residual bias on the mass after the calibration.

For this reason, a further morphing has been fitted between the original Z MC and the data-like one. A third copy of Z MC, called data-like 2, is then obtained applying the latter morphing function to the original Z MC. A sketch summarizing the procedure followed to obtain the data-like 2 Z MC is shown in fig. 6.11. The bottom line is that the differences between data-like and data-like 2 should be of the same order of the differences between data-like Z MC and Z real data. Under this assumption, a quantitative estimation of the residual bias after the morphing can be obtained performing a mass fit between pseudo-data, made with the data-like MC, and templates, made with the data-like 2 MC. It has been followed the same procedure, explained in the previous paragraph, of splitting in sub-samples and taking the averages as bias estimator. The results, reported in tab. 6.2, have been computed only in the case in which TK or mnpk

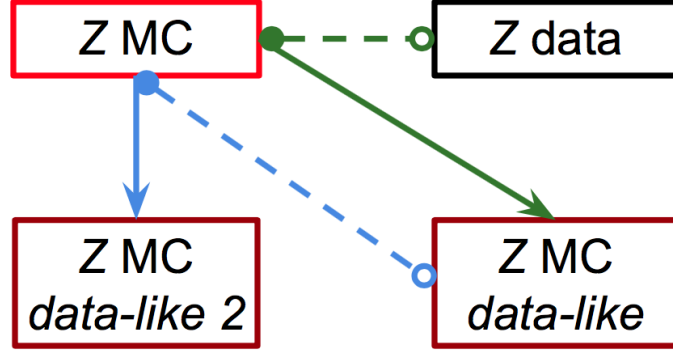


Figure 6.11: Sketch explaining the morphings trained and applied to obtain the data-like 2 Z MC. The dashed lines, connecting two samples, indicate that a morphing has been trained to modify the sample on the full circle side, in order to have the same $P(y | z)$ pdf of the sample on the empty circle side. The full lines represent the application of the morphing with the same color: the sample indicated by the arrowhead is produced by the application of the fitted morphing to the sample from which the arrow starts. Different colors indicate different morphings.

Bias on Z mass after MC recoil calibration

Recoil estimator used in M_T^2	Average bias [MeV]	Average bias error [MeV]
mnpc	-11.2	9.8
TK	-14.4	7.6

Table 6.2: Average mass bias and its uncertainty obtained in the fit between data-like 2 Z MC and data-like Z MC. It is reported the result obtained when different experimental recoil estimator are used.

recoil is used.

In particular, it is interesting to notice that in both cases the uncertainty on the residual bias of order 10 MeV makes the result compatible with 0 within 2 RMS, meaning that the size of our MC sample does not allow us to identify any relevant effect remaining after the morphing. Since it is important for a W mass measurement, aiming to a competitive precision, to investigate uncertainties well beyond 10 MeV, it will be necessary to repeat similar studies when a definitive version of the MC will be individuated. In addition, since the bias uncertainty scales approximately with the square root of the MC sample size, a MC sample at least 9 times larger is desirable for future developments.

Anyway, it is interesting to remark how the morphing procedure developed in this work is able to correct the recoil distribution in the MC, in order to better describe real data, conditionally to the boson kinematics. The agreement between the target

and corrected distribution, estimated indirectly with the auxiliary data-like 2 Z MC copy, has been tested up to the equivalent level on m_W of 10 MeV.

6.7 Porting the calibration from Z to W

The assumptions that make this procedure meaningful and complete for the W mass are multiple.

The MC calibration procedure heavily relies on the fact that, at fixed boson kinematics, the recoil behaves in the same way for Z and W . In particular, the assumption under which the morphing derived on the Z samples has the same performances on the W one is even stronger. Considering the way in which the MC to data morphing functions is obtained, it must be true that the distribution of $P(y | z)$ is the same in Z and W , both in MC and data. More explicitly: $P(y | z)$ has to be the same between Z and W , in both data and MC:

$$P(y | z)_{MC}^Z = P(y | z)_{MC}^W \quad \text{and} \quad P(y | z)_{data}^Z = P(y | z)_{data}^W$$

If this does not hold, the corrections derived on the Z are not effective in improving the data-MC agreement on the W .

Given the experimental constraints, it is impossible to fully check the validity of this hypothesis, because in the W data the boson momentum can be reconstructed only through the recoil since the neutrino is not detected, hence the z variable are not present at reconstruction level in that sample.

In addition, the MC/data calibrating morphing between Z MC and Z data has been obtained conditionally to the reconstructed z variables. A workaround is needed to apply the morphing on the W MC sample.

6.7.1 Generated instead of reconstructed boson kinematic variables

An effective and rough way of facing the problem that no reconstruction level variables are available for the W events, is to simply use, instead, the generator level ones. So, even if the morphing for the y variables has been fitted on the Z sample, conditionally to the reconstructed z variables, it is instead applied passing the MC truth variables to the interpolation algorithm.

To test the effectiveness and quantify the bias introduced by this approach, the Z MC is

6.7. PORTING THE CALIBRATION FROM Z TO W

once more used. Starting from the original copy of the Z MC, a new *data-like gen* copy is obtained replacing the recoil variables with the ones computed applying the data-MC morphing function in the way described above. As usual, to quantify the effect on the W mass extraction, a fit to the transverse mass spectrum is performed, using pseudo-data made with the data-like MC and templates made with the new *data-like gen* MC. Results of the bias obtained fitting the 100 sub-samples are reported in fig. 6.12, when the TK or mnpk recoil is used. In both cases the mean is compatible with zero within

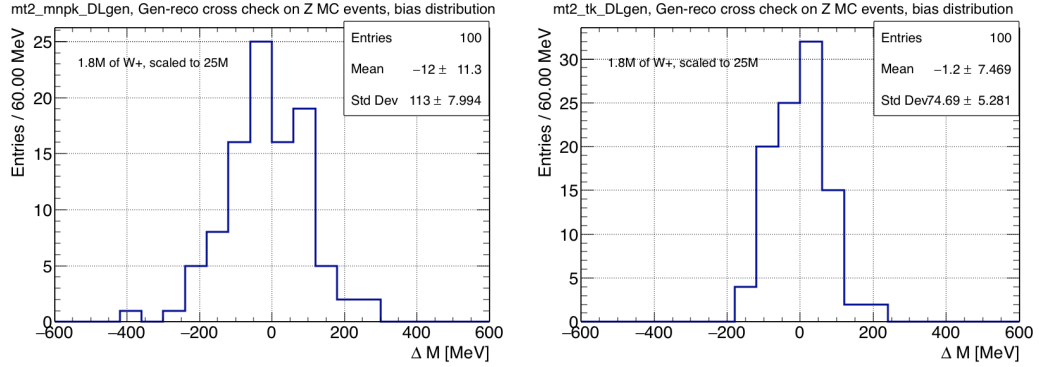


Figure 6.12: Distributions of the bias obtained from the fit of 100 sub-samples of pseudo-data made with standard data-like Z MC and templates made with data-like Z MC applied using the generator level Z boson kinematics. On the left mnpk recoil is used, whereas on the right TK recoil is used.

one or two standard deviations. It is then possible to draw the conclusion that, given the level of accuracy reachable with the current statistics, training the morphing on reconstructed level boson kinematics variables and applying it on generator level one produce a bias of order 10 MeV or less.

A different workaround to this gen-reco problem is to artificially smear the MC truth level neutrino momentum in order to obtain a *detected neutrino* with a behavior similar to muons. It can then be used to simulate a reconstructed W from the dilepton pair which should be equivalent to the reconstructed Z . For sake of time, this alternative approach has not been tested.

6.7.2 Intrinsic MC differences between Z and W

This section discusses the differences of the recoil in Z and W at fixed boson kinematics, assumed null in the previous sections. Differences can be induced, for example, by the different flavors of the quarks which annihilate into the boson. For instance, as discussed in sec.4.1.1, a larger contribution to the Z production is given by the b and c quarks.

6.7. PORTING THE CALIBRATION FROM Z TO W

To quantify these differences two morphings have been trained: one to morph the y variables (conditionally to the usual p_T and p_z) of the Z MC into the W MC; the other one vice versa. Through their application, a copy of the Z MC W -like and a copy of the W MC Z -like are obtained.

Given the small size of the differences, it is impossible to quantify, or even identify, some effects by looking at the ratio plot of the recoil distributions before and after the morphing.

For this reason, the effect has been quantified performing the mass fit on the transverse mass spectrum in a way similar to what has been done in the previous paragraphs. The original unmodified copy of the MC has been used to build pseudo-data, whereas

Bias on the mass obtained in the cross sample morphing test			
Pseudo-data	Templates	Average bias [MeV]	Average bias error [MeV]
bare Z MC	W -like Z MC	1.9	11.6
bare W MC	Z -like W MC	-36.2	9.4

Table 6.3: Average mass bias and its uncertainty obtained in the W/Z morphing cross checks. Both fits are performed to the transverse mass distribution computed using the $mnpk$ recoil. The samples are divided into 100 sub-samples and the average of the biases obtained is reported.

the W -like (for Z MC) and the Z -like (for W MC) copies have been used to build templates. It has been explored only the case in which $mnpk$ is used. The average bias value obtained in the 100 sub-samples is reported in tab. 6.3.

The output of this double cross check has not been fully understood yet. On one hand, the average on the Z sample is largely compatible with 0 but, on the other hand, the average on the W sample is three standard deviations away from it. Even if further studies are needed to better understand this discrepancy, it can be stated that the assumption that $P(y | z)$ is equal in Z and W MC holds at least until the level of few tens of equivalent MeV on m_W .

Finally, it has to be remarked that no other quantitative approaches to estimate the goodness of the above assumption have been found in literature, even if both ATLAS and CDF measurement explicitly use it. Given the absence of possible comparison, it is hard to fully trust this last result or to set an expected level of precision required. Without any doubt, a deeper investigation of this phenomenon is crucial in order to properly understand the systematic uncertainties of the W mass measurement. Preliminary results are encouraging, but further efforts are needed.

Chapter 7

Conclusions

The W mass measurement is one of the more interesting challenges that the physics at the energy frontiers is facing. The LHC is nowadays providing a good experimental environment to explore new and more accurate approaches to this topic. The CMS experiment intends to exploit this opportunity and is indeed setting up all the tools for a W mass measurement of unprecedented precision.

Considering the amount of data gathered, the systematic uncertainty will represent the limiting factor. From the experimental point of view, one of the most fertile fields for original and innovative contributions is the study of the recoil system, which is produced in association with the W boson. The original work presented in this thesis has vastly explored this topic.

Studies conducted at generator level on the W and Z boson production mechanism set the basis of a deeper understanding of the recoil origins. A solid definition of the recoil has been justifiably preferred with respect to the previous used MET concept. The investigation of the recoil phenomenology has remarked that the CMS acceptance limits deeply affect the detection of a sizable part of the particles constituting the recoil.

In the W mass analysis the recoil enters mainly in the computation of the transverse mass. The distribution of this variable presents, at MC level, a sizable peak due to the correlation between the muon from the W decay and the recoil. The correlation, and then the peak, is slightly reduced by the muon experimental reconstruction. On the contrary, the experimental effects, which limit the precision of a recoil measurement, considerably sculpt the transverse mass shape.

The main reason for using the transverse mass in the W mass analysis consists in the fact that its distribution has a little dependence from the boson transverse momentum

spectrum. The better the recoil is measured, the smaller is this dependence and the smaller is the required accuracy of the MC simulation. Being a 2D vector, the recoil presents a series of technical difficulties. One of the most important is that a precise measurement of both the magnitude and the angle is crucial. In addition, there is not a unique figure of merit which can be used to evaluate the recoil measurement performances. A good estimator is $|\Delta\vec{h}|$, which is often of the same size as the recoil magnitude itself. Finally, also the difference between the true and the reconstructed transverse mass (ΔM_T^2) is interesting, since it is correlated to the statistical uncertainty on the W mass fitted value.

One of the main achievements of this thesis is the development of a novel approach to the recoil measurement. Using a semi-parametric regression based in dense neural networks, the probability distribution function of the recoil vector is predicted, as a function of the observed features of the event. The trustworthiness of this prediction has been tested and found at the satisfactory level of few percent.

Two point estimators of the recoil vector have been derived using the mean (MNPK) and the most probable value (PK) of the recoil pdf predicted by the regression. A comparison with previously used recoil estimators (TK and PF) has underlined the superior quality of the approach presented in this thesis: in particular, both magnitude and angle of the recoil are measured with a precision of few percent better. This is reflected in an improvement in both $|\Delta\vec{h}|$ and ΔM_T^2 figures of merit.

Recoil related systematic uncertainties arising in the transverse mass fit, for the W mass measurement, are discussed. Those are due to discrepancies between the data and the MC simulation used in the template fitting procedure. In particular, since the transverse mass depends only on three scalar variables, their joint probability has to be studied: the muon transverse momentum; the recoil transverse momentum and their relative angle on the transverse plane. Given the W boson role, all the correlation between recoil and muon passes through it. As a result, the systematic uncertainties (i.e. the differences between data and MC) can be divided into three independent sources: the muon simulation accuracy at fixed boson kinematics; the recoil simulation accuracy at fixed boson kinematics; and the boson kinematics prediction.

The latter is studied in this work. The systematic uncertainty induced by the knowledge of the W p_T spectrum has been estimated in several scenarios. The value obtained when different definitions of the recoil are used has been computed and used to rank the definitions. The new definitions developed in this work demonstrate to be a substantial improvement. In particular, when the MNPK recoil is used, the fit presents $1/3$ of the systematic with respect to the approach used by CMS in its last public result. This

number can be used as a benchmark but not as a proper estimation of the uncertainty that will be present in future CMS measurements. The MC which will be used and the final fitting strategy need still to be defined.

A further substantial part of the work consists in the development of a novel method to increase the agreement between data and MC. Based on the systematic factorization scheme mentioned above, it is called *quantile morphing* and uses a quantile mapping on top of a quantile regression to modify, event-by-event, the recoil in MC. The major improvement is that the corrections are derived in a way to obtain the agreement in the 2D angle-magnitude recoil distribution and conditionally to the boson longitudinal and transverse momentum. This method has been applied on Z events since the boson properties are available in both MC and data, of course at reconstruction level.

The bias induced by the recoil modeling on the Z mass measurement, when a W -like approach is used, has been computed. When the bare MC is used, it has been estimated to be several tens of MeV. After the recoil correction, it has been verified to be compatible with 0 up to the level of 10 MeV, being the available MC sample size the limiting factor.

Finally, the porting of the MC calibration from Z to W sample is discussed. Two main points are faced: the morphing is derived on Z as a function of the boson reconstructed variables, but there is no such feature in W events; the conditional probability of the recoil, given the boson kinematic, has to be the same in Z and W for the calibration to be effective. A solution to the former is studied. The usage of generator level boson kinematic variables instead of the reconstructed level ones has been tested to induce a bias on the W mass of order 10 MeV or less. The latter point has been investigated with a cross check of Z events morphed to W ones and viceversa. The bias induced on the mass extraction has been estimated to be at the level of few tens of MeV. A similar quantitative test has never been performed before, hence no similar studies are present in literature. Preliminary results are encouraging, but further efforts are needed.

7.1 Future developments

This section is devoted to discuss possible or natural developments to the work exposed in this document.

In order to improve the studies presented in this thesis, a larger MC sample is needed. The reached level of 10 MeV in the systematic uncertainty estimation is not satisfactory for a precise W mass measurement. The morphing procedure discussed above has to be tested up to few MeV and possible further improvements have to be

investigated in order to increase the data/MC agreement. When a new and bigger MC sample will be available, it will be interesting to repeat the analysis presented in the last chapter and possibly develop it to the level of the proper W mass measurement. Furthermore, it is of crucial interest to deepen the study about the discrepancies, between W and Z MC, in the recoil distribution at fixed boson momentum. The studies performed in this work suggest that discrepancies of $O(10 \text{ MeV})$ may be present. Unfortunately, no similar numerical estimation is present in literature and so further trials are required to know if those discrepancies are real or morphing induced. A very fast test of repeating the $W - Z$ cross check selecting only the light flavor production has been tried but no significant results have been obtained. The physics inducing the discrepancies still needs to be properly individuated and, possibly, the induced bias computed with a few MeV precision.

An additional interesting point about systematics to be further explored is the understanding of the induced uncertainty from the W p_T spectrum knowledge. Even if the fitting procedure presented in this work is a very effective solution, it requires a non-negligible amount of time and is a very high level figure of merit. It is important to isolate the feature of the recoil which mitigates the dependence of the transverse mass variable from the boson p_T , so that it can be explicitly improved. During the thesis, a sizable effort has been dedicated to the study of this topic but no crucial result has been obtained.

The recoil regression is also a fertile field for future studies. Remaining in the current framework, a study to optimized the DNN hyper-parameters and architecture is of sure interest and can bring to order 10% improvement in a limited amount of time. For instance, the effect of additional variable should be carefully studied. Inspired by the procedure developed for the multi-dimensional quantile morphing, it is interesting to test a regression where one of the two variables has as input the predicted value of the other one. In this way, possible correlation among the variables can be properly treated. Trials addressing the usage of a different recoil architecture are also required for maximizing the regression performances.

A great impact on the W mass analysis can instead be made by a better usage of the information coming from the regression in each event. For example, an estimator of the predicted recoil pdf width (variance, FWHM, ...) contains information about the accuracy on the recoil value point prediction; it would be interesting to investigate the effects of exploiting such information in the fitting procedure. A possibility in this direction is to use the inverse of the predicted recoil pdf variance as weight of each event.

Besides the small improvements quoted above, a major framework revolution can also be studied. Machine learning techniques are today far more developed than simple dense neural networks. In this thesis, the regression is fed with high level features (S_{TK} , R_{ntnpv} , ...) because the DNN architecture is suitable for problems with a fixed number of inputs and whose information is meaningful standalone. To explain with a ground level example, using singularly all the tracks in the event is not possible with DNNs because their number varies every event. In addition, single tracks are meaningless and the information stands in the event as a whole. However, more sophisticated approaches, such as convolutional neural networks (CNN) or long-short time memory networks (LSTM), are able to deal with different problems. Specifically, in the early stages of the thesis, it has been tried to convert the event into a 2D $\eta - \phi$ color map of the detected particles. The map has then be fed to a CNN targeting, as the DNN discussed in the thesis body, the recoil pdf. A similar approach is more suitable to exploit the full information detected, starting also from the low level features. A full development of this approach has not been fulfilled for time requirements but, given the extraordinary capabilities that CNN has shown in a broad class of application, it might be interesting to invest few months on this topic.

The LSTM approach has never been tried for recoil measurement but is widely used by the CERN community ([53] and [54]). A possible way to apply it to the recoil is to feed the network with all the detected particles ordered with a physics induced metrics, such as angular distance or transverse momentum.

Two possibilities are here suggested, but a broader spectrum of approaches is available and already implemented in common Python or C++ libraries.

Finally, a direct application of the DNN regression can deliver a standalone measurement of the W p_T spectrum, unfolded by detector effects. Looking at the sum of pdfs plots shown in sec.5.3.1.1, it is clear that a regression can be optimized in order to describe the generator level boson p_T spectrum. Using the sum of pdfs is effectively equivalent to perform unbinned unfolding of the detector effects. Now, no experimental measurement of the W p_T in the soft range with a fine binning has never been published. Nevertheless, a similar result would be of great interest for the theoretical community which needs it to improve the computation used for MC simulations. An unbinned measurement of the W p_T can then represent an important measurement to be delivered in less than one year. No trials have been performed in this direction during the thesis. Even if the DNN regression can be applied straightforward, a precise and demanding study on the systematic uncertainty is required to finalize the measurement idea here exposed.

Appendix A

DNN training details

As explained in sec.5.2.2, for each parameter of the fitting functions a DNN has been trained: 4 in parallel for $\ln(e1)$ and 5 in parallel for $e2$. The training consisted in fitting the conditional probability of the correction coefficient in a W MC sample, selected according to the criteria explained in sec. 4.3.1: 70% of events have been used as training sample and 30% as testing sample. The loss function used is respectively the negative logarithm of the $f_{\ln e1}$ and f_{e2} functions, opportunely normalized to 1 (crucial for the meaningfulness of the procedure). The architecture chosen for all the DNNs has three hidden layers, an input layer where all the x are passed and an output layer made by a single neuron. The hidden layers have, from the nearer to the input to the output, respectively 100, 50 and 8 neurons. The activation function chosen for all the hidden layers is a LeakyRelu:

$$\phi(x) = \begin{cases} ax, & \text{if } x < 0 \\ x, & \text{if } x > 0 \end{cases}$$

where the parameter a has been fixed to the value of 0.2. The output neurons which give values to the α_1 , α_2 and σ parameters of both correction coefficients have an exponential activation function so that their output is necessary positive. The μ parameter output neuron of $\ln(e1)$ has a linear activation function, whereas the one for $e2$ has a sigmoid scaled to take values in $[-\pi, \pi]$.

The starting values of bias parameters b_i are initialized according to a *Glorot Uniform*, while a *Glorot Normal* is used for the weights w_j^i starting values. The total number of free parameters for each DNN is of order 22k.

The fitting procedure, which takes about five hours on a standard laptop using the keras library [50], consists of 100 epochs with a batch size of 200 events. The mini-

mizing algorithm which has been chosen is adam [55] with a decay of learning rate of $1 \cdot 10^{-6}$ and a clip norm set to 1.

Furthermore two callbacks are set during the training: *early stopping* on the loss function of the testing sample with patient 8 and a *reduce learning rate on plateau* of 0.3 with patient 3 on the loss function of the training sample.

The details of the minimizing algorithm choice and hyper parameter (DNN architecture, training parameters, ...) optimization have been explored with several training trails but are not here reported.

Bibliography

- [1] Sheldon L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579 – 588, 1961.
- [2] F. Halzen and A. D. Martin. *Quarks and leptons: an introductory course in modern particle physics*. Wiley, New York, USA, 1984.
- [3] Georges Aad, T Abajyan, B Abbott, J Abdallah, S Abdel Khalek, AA Abdelalim, O Abdinov, R Aben, B Abi, M Abolins, et al. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, 2012.
- [4] Wikipedia. Standard model — wikipedia, the free encyclopedia, 2017. [Online; accessed 3-July-2017].
- [5] Michael Edward Peskin. *An introduction to quantum field theory*. Westview press, 1995.
- [6] Riccardo Barbieri. *Lectures on the ElectroWeak Interactions*. Edizioni della Normale.
- [7] P.W. Higgs. Broken symmetries, massless particles and gauge fields. *Physics Letters*, 12(2):132 – 133, 1964.
- [8] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.
- [9] Georges Aad, B Abbott, J Abdallah, O Abdinov, R Aben, M Abolins, OS AbouZeid, H Abramowicz, H Abreu, R Abreu, et al. Combined measurement of the higgs boson mass in p p collisions at $s = 7$ and 8 tev with the atlas and cms experiments. *Physical review letters*, 114(19):191803, 2015.

- [10] M. Martinez, R. Miquel, L. Rolandi, and R. Tenchini. Precision tests of the electroweak interaction at the Z pole. *Rev. Mod. Phys.*, 71:575–629, Apr 1999.
- [11] M. Baak, J. Cúth, J. Haller, A. Hoecker, R. Kogler, K. Mönig, M. Schott, and J. Stelzer. The global electroweak fit at NNLO and prospects for the LHC and ILC. *The European Physical Journal C*, 74(9):1–14, 2014.
- [12] Lyndon Evans and Philip Bryant. Lhc machine. *Journal of Instrumentation*, 3(08):S08001, 2008.
- [13] Stephen Holmes, Ronald S Moore, and Vladimir Shiltsev. Overview of the tevatron collider complex: goals, operations and performance. *Journal of Instrumentation*, 6(08):T08001, 2011.
- [14] Ralph Aβmann, Mike Lamont, Steve Myers, et al. A brief history of the lep collider. *Nuclear Physics B-Proceedings Supplements*, 109(2-3):17–31, 2002.
- [15] Lyndon Evans and Philip Bryant. LHC Machine. *JINST*, 3:S08001, 2008.
- [16] CERN. Te-epc-lpc in lhc. <http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/lhc/general.stm>, 2017.
- [17] The ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, 2008.
- [18] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, 2008.
- [19] The LHCb Collaboration. The LHCb Detector at the LHC. *Journal of Instrumentation*, 3(08):S08005, 2008.
- [20] The ALICE Collaboration. The ALICE experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08002, 2008.
- [21] The CMS collaboration. Cms luminosity - public results. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>, 2017.
- [22] G Apollinari, I Béjar Alonso, Oliver Brüning, M Lamont, and Lucio Rossi. High-luminosity large hadron collider (hl-lhc): Preliminary design report. Technical report, Fermi National Accelerator Laboratory (FNAL), Batavia, IL (United States), 2015.

- [23] The CMS Collaboration. The cms experiment at the cern lh. *Journal of Instrumentation*, 3(08):S08004, 2008.
- [24] Siona Ruth Davis. Interactive Slice of the CMS detector. Aug 2016.
- [25] Patrick Janot et al. Particle-flow reconstruction and global event description with the CMS detector. 2017.
- [26] Paolo Nason. A new method for combining nlo qcd with shower monte carlo algorithms. *Journal of High Energy Physics*, 2004(11):040, 2004.
- [27] Johan Alwall, Pavel Demin, Simon De Visscher, Rikkert Frederix, Michel Herquet, Fabio Maltoni, Tilman Plehn, David L Rainwater, and Tim Stelzer. Madgraph/madevent v4: the new web generation. *Journal of High Energy Physics*, 2007(09):028, 2007.
- [28] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. A brief introduction to PYTHIA 8.1 . *Computer Physics Communications*, 178(11):852 – 867, 2008.
- [29] Sea Agostinelli, John Allison, K al Amako, J Apostolakis, H Araujo, P Arce, M Asai, D Axen, S Banerjee, G Barrand, et al. Geant4—a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.
- [30] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.
- [31] Morad Aaboud et al. Measurement of the W -boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. 2017.
- [32] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.
- [33] C Rubbia, P McIntyre, and D Cline. Producing massive neutral intermediate vector bosons with existing accelerators. In *Proceedings of the International Neutrino Conference Aachen 1976*, pages 683–687. Springer, 1977.
- [34] G. Arnison et al. Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $s^{*(1/2)} = 540\text{-GeV}$. *Phys. Lett.*, 122B:103–116, 1983. [611(1983)].

- [35] G. Arnison et al. Experimental Observation of Lepton Pairs of Invariant Mass Around $95\text{-GeV}/c^{*2}$ at the CERN SPS Collider. *Phys. Lett.*, 126B:398–410, 1983.
- [36] Douglas A. Glenzinski and Ulrich Heintz. Precision measurements of the W boson mass. *Annual Review of Nuclear and Particle Science*, 50(1):207–248, 2000.
- [37] CMS Collaboration. Search for light vector resonances decaying to a quark pair produced in association with a jet in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$. 2017.
- [38] T. at al. Aaltonen. Precise measurement of the w -boson mass with the cdf ii detector. *Phys. Rev. Lett.*, 108:151803, Apr 2012.
- [39] The CMS Collaboration. Measurement of the differential cross section and charge asymmetry for inclusive $\mathrm{p}\mathrm{p}\rightarrow\mathrm{p}\mathrm{p}\mathrm{W}^{\pm}\mathrm{X}$ production at $\sqrt{s}=8\text{ tev}$. *The European Physical Journal C*, 76(8):469, 2016.
- [40] T. Aaltonen et al. Precise measurement of the W -boson mass with the CDF II detector. *Phys. Rev. Lett.*, 108:151803, 2012.
- [41] W-like measurement of the Z boson mass using dimuon events collected in pp collisions at $\sqrt{s} = 7\text{ TeV}$. Technical Report CMS-PAS-SMP-14-007, CERN, Geneva, 2016.
- [42] Gigi Rolandi. Analysis and discussion of the recent W mass measurements. In *52nd Rencontres de Moriond on EW Interactions and Unified Theories (Moriond EW 2017) La Thuile, Italy, March 18-25, 2017*, 2017.
- [43] Nicolo Foppiani. Definition and calibration of the hadronic recoil in view of the measurement of the W boson mass with the CMS experiment. *Final thesis for the Master Degree in Physics*, 2017.
- [44] Stefano Catani, Daniel de Florian, and Massimiliano Grazzini. Universality of non-leading logarithmic contributions in transverse-momentum distributions. this work was supported in part by the eu fourth framework programme “training and mobility of researchers”, network “quantum chromodynamics and the deep structure of elementary particles”, contract fmrx-ct98-0194 (dg 12 - miht). *Nuclear Physics B*, 596(1):299 – 312, 2001.

- [45] G. Altarelli, G. Parisi, and R. Petronzio. Transverse momentum in drell-yan processes. *Physics Letters B*, 76(3):351 – 355, 1978.
- [46] Guido Altarelli, G Parisi, and Roberto Petronzio. Transverse momentum of muon pairs produced in hadronic collisions. *Physics Letters B*, 76(3):356–360, 1978.
- [47] The CMS Collaboration. Baseline muon selections for Run-I. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGwideMuonIdTightMuon>.
- [48] J Smith, WL Van Neerven, and Jos AM Vermaseren. Transverse mass and width of the w boson. *Physical Review Letters*, 50(22):1738, 1983.
- [49] Sebastian Ruder. An overview of gradient descent optimization algorithms. *CoRR*, abs/1609.04747, 2016.
- [50] François Chollet et al. Keras. <https://github.com/fchollet/keras>, 2017.
- [51] Roger Koenker. *Quantile regression*. Number 38. Cambridge university press, 2005.
- [52] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in python: Gbr documentation. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [53] Maciej Wielgosz, Andrzej Skoczeń, and Matej Mertik. Using lstm recurrent neural networks for detecting anomalous behavior of lhc superconducting magnets. *arXiv preprint arXiv:1611.06241*, 2016.
- [54] Daniel Guest, Julian Collado, Pierre Baldi, Shih-Chieh Hsu, Gregor Urban, and Daniel Whiteson. Jet flavor classification in high-energy physics with deep neural networks. *Physical Review D*, 94(11):112002, 2016.
- [55] Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

List of figures

1.1	Fundamental fields presented in the SM [4]. The 3 matter generations are shown in the first 3 columns, while gauge bosons and scalar boson are shown respectively in the 4 th and 5 th column.	4
1.2	Most significative corrections to the W propagator at one-loop level. Left: fermion isodoublet correction. Right: correction from the emission of a Higgs boson.	9
1.3	From [11]. 68% and 95% confidence level (CL) contours in the $m_t - m_W$ plane for the fit including m_H (blue) and excluding m_H (grey). In both cases the direct measurements of m_W and m_t were excluded from the fit. The values of the direct measurements are shown as green bands with their one standard deviations. The dashed diagonal lines show the SM prediction for m_W as function of m_t for different assumptions of m_H . .	11
2.1	Scheme of the facilities of the CERN acceleration complex from [16]. . .	15
2.2	Cumulative luminosity versus day delivered to CMS during stable beams and for p-p collisions. This is shown for 2010 (green), 2011 (red), 2012 (blue), 2015 (purple), 2016 (orange) and 2017 (light blue) data-taking. Figure from [21].	17
2.3	Picture of CMS in its location at point 5 of LHC. This section of the detector shows in full the 15 m of the outer diameter. (Image: Michael Hoch/Maximilien Brice)	18
2.4	Longitudinal view of the CMS detector. Picture from http://www-collider.physics.ucla.edu/cms/	19
2.5	A transverse slice of the CMS detector, with the qualitative experimental signatures of the different particles. Picture from [24].	21
3.1	Results and world average from the past W mass measurements from [30].	28
3.2	Main tree level W boson production diagrams at e^+e^- collider.	29

3.3	Cross section for the process $e^+e^- \rightarrow W^+W^-$ as a function of the center of mass energy for various m_W hypothesis [36].	30
3.4	Di-jet mass distribution for the p_T category 500-600 GeV. Data are shown as the black points. The QCD background prediction, including uncertainties, is shown in the gray boxes. Contributions from the W , Z , and a hypothetical Z' signal at a mass of 135 GeV are indicated as well. Figure from [37].	33
3.5	Transverse mass and muon p_T spectrum distribution fitted by the CDF collaboration [38] for the measurement of the W mass.	34
4.1	Spectrum of W boson transverse momentum from Pythia 8 standalone generation, both signs summed together. A stack histogram of the different production flavors shows the negligible impact of the b quark contribution. In the legend M stands for the mean and I for the relative fraction of each sub-sample.	40
4.2	Comparison between W (left) and Z (right) transverse momentum spectra at generator level, made with Pythia 8. The distribution for the various production flavors is shown normalized to 1 for each possible pair of the quarks annihilating into the boson. I and M have the same meaning as in fig .4.1.	41
4.3	Left: Sketch of the production of a W boson from a pp collision. The boson decay products may result almost opposite to some of the recoil particles but still an important part of the momentum is lost in the forward region. Right: Display of W production event generated with Pythia 8 and passed through a fast simulation of the CMS detector made with PAPAS.	42
4.4	Left: Number of particles generated for a set of different requirements reported in the legend. In particular the "ch." in the pink histogram legend stands for the fact that only charged particles are included. The mean of the distribution is also reported as M . Right: Profile plot from the 2D distribution of the recoil projection on the muon flight direction when true value (x axis) and particle (y axis) information is used. The plot is shown for the same cuts present on the left one. The plots report the information from a Pythia 8.2 at generator level. . . .	43

4.5	Distribution of the distance of charged tracks from the primary vertex in the transverse and longitudinal plane obtained from a subsample of the CMS simulation in events without generated pile-up vertexes. The red ellipses retains 90% of the tracks.	47
4.6	Transverse mass distributions from W MC for different kinds of variables used. For each legend line, two codes of the form yXXX are reported: y stands for the kind of variable (m for muon and r for recoil) and XXX stands for the level of the variable used (MC means Monte Carlo truth, RECO for reconstructed after detector simulation). The label 1 st is added when the first order approximation is used for M_T	49
4.7	Distribution of the transverse mass (left) and muon transverse momentum (right) in W^+ from W MC. Both distributions are presented in three cases: the W is produced always with $p_T = 0$ (black), the MC truth information is used (blue) and the reconstructed information is used (red). The green line indicates the value of the W mass.	50
4.8	Sketch exemplifying a situation where the angular mismatch between the real quantity (MC, blue) and the reconstructed quantity (TK, green) makes the reconstructed quantity, scaled in order to have the correct magnitude, (STK, red) a worse estimator (in the $ \Delta\vec{h} $ figure of merit) than the non-scaled one. This is often the case for the recoil measurement in the W mass events.	53
4.9	Transverse mass distribution for W events in different generated boson p_T regions (i.e. MC truth recoil). The fraction of events belonging to that region is reported as well. In each p_T region, different colors indicate different recoil variables used to build M_T : blue MC truth angle and magnitude; green MC truth angle and reco TK magnitude; black reco TK angle and MC truth magnitude; red reco TK angle and magnitude. Reconstructed muon is always used.	54
4.10	Distribution of the TK recoil resolution in the CMS frame (left) and in the true produced W boson flight direction frame (right).	55
4.11	Left: Distribution of the magnitude of the recoil resolution for different scale factor applied to the TK recoil. Right: Distribution of the error on the transverse mass for the same scale factors used on the left. . . .	56
5.1	Scheme of a dense fully-connected neural network with 3 input and 2 outputs.	63

5.2	Example of a semi-parametric regression. Left: 2D distribution of the variable y and the observable x with the so called candle plot overimposed: red box center corresponds to the mean and red box half-width to the RMS. Horizontal red lines ends at the 90% of the distribution and the small red crosses shows the events which stay outside. Right: Fitted dependence of the parameters μ and σ of the y Gaussian distribution as a function of the observable x	64
5.3	Sketches illustrating the definition of the scale-angle correction coefficient (left) and of the scalar recoil frame oriented correction coefficient (right).	66
5.4	Distribution of correction coefficient from different recoil parametrization schemes.	67
5.5	Angular distance between the experimental measured recoil and true recoil in the PF and TK case. This corresponds to the $e2$ correction coefficient.	68
5.6	Distribution of the correction coefficient $\ln(e1)$ (left) and $e2$ (right) in a W MC sample identical but statistically independent from the one used for the regression training. Specific cuts, reported on the top of both plots, have been applied to check the regression performance in this illustrative bin. The total integral of the histograms is 9000, which contains in the $4 \cdot 10^{-4}$ of the W MC sample. The red line presents the regression average prediction in this bin. The black dashed line is a fit to the presented distribution performed with the same function used by the regression.	73
5.7	Sketch illustrating the procedure to build the sum of pdfs plots. Integrated distribution of the y variable is compared with the y probability predicted by the regression in each event summed altogether.	74
5.8	Sum of pdfs plots for the two correction coefficients which are targeted by the regression: $\ln(e1)$ on the left and $e2$ on the right.	74
5.9	Sum of pdfs plots for the recoil in the scale-angle frame. In the plots it is presented a comparison between generator level distribution (black) for which the regression has been trained, sum of pdfs from DNN (blue). It is also reported the TK distribution of the recoil angle and the magnitude distribution of a TK recoil whose modulus has been scaled for 1.8. . . .	75

5.10	Distribution of the residual correction needed event-by-event to recover the MC information for the TK recoil, STK18 (TK with magnitude scaled by 1.8), PF recoil and the two point-like estimators produced with the DNN regression. In the x axis title, the superscript est stands for estimator (i.e. pk, TK, ..), whereas the ref stands for reference (i.e. MC).	76
5.11	Resolution plot for different experimental recoil estimators. The magnitude of the 2D resolution on the recoil is reported on the left and the difference between the transverse mass computed with MC truth variables and reconstructed variables is reported on the right.	77
5.12	Recoil magnitude distribution for the different estimators. In red the true generated distribution, in blue the classical recoil definitions and in green the estimators derived from the regression.	78
6.1	Comparison between reconstructed Z kinematic distributions in data and MC: the transverse momentum is shown on the left and the longitudinal momentum is shown on the right.	84
6.2	Left and center: Ratio between the pseudo-data p_T spectrum before and after the re-weighting. A linear discrepancy of different size has been tested with the left plot, whereas a power law exponentially dumped has been tested in the central plot. Right: For sake of comparison, it is here reported the ratio of the p_T distribution form Z data and a pythia only MC tested by the CMS W mass analysis group.	85
6.3	Effects of the linear re-weighting procedure on pseudo-data distributions of boson momentum (left) and transverse mass (right) are reported for different re-weight size. The left plot is normalized in the full range, whereas the right plot content is scaled by a factor 10 in order to resemble the expected CMS statistics.	86
6.4	Value of the χ^2 for the mass templates in the range $-100 \text{ MeV} < \Delta M < 100 \text{ MeV}$, when compared to pseudo-data re-weighted in order to have a linear variation of the W p_T spectrum, as shown in fig. 6.2 left.	86
6.5	Transverse mass fit results for different estimators, obtained fitting half of the W MC sample. The abscissa reports the best fit value ΔM^* obtained in each scan at different W p_T re-weighting sizes. The error reported on it represents the σ_{stat} value. The best fit line for each set of points is reported dashed.	87

6.6	The variables ranking in terms of statistical and W p_T systematic uncertainty. The error box reported represents the statistical uncertainty due to the finite size of the MC sample used.	89
6.7	Ratio plot of the distribution of two example x variables (TK recoil magnitude on the left and number of tracks on the right) in four different bins of the boson kinematic phase space.	92
6.8	Sketch of the quantile mapping procedure.	94
6.9	Recoil variable distributions for the TK (top) and mnpk (bottom) recoil definitions. For each variable the three different colors identify the data distribution (black), the bare MC distribution (red) and the MC distribution after the morphing procedure (blue). Ratio plots are also shown below each plot between the data and the other two distributions. . . .	97
6.10	Distributions of the bias obtained from the fit to the transverse mass, computed with the mnpk recoil of 100 sub-samples for each combination of the MC0 copies. In the left and central plot, $MC0_0$ is used as pseudo-data and respectively $MC0_1$ and $MC0_2$ as templates. In the right plot, $MC0_1$ is used as pseudo-data and $MC0_2$ as template.	99
6.11	Sketch explaining the morphings trained and applied to obtain the data-like $2 Z$ MC. The dashed lines, connecting two samples, indicate that a morphing has been trained to modify the sample on the full circle side, in order to have the same $P(y z)$ pdf of the sample on the empty circle side. The full lines represent the application of the morphing with the same color: the sample indicated by the arrowhead is produced by the application of the fitted morphing to the sample from which the arrow starts. Different colors indicate different morphings.	101
6.12	Distributions of the bias obtained from the fit of 100 sub-samples of pseudo-data made with standard data-like Z MC and templates made with data-like Z MC applied using the generator level Z boson kinematics. On the left mnpk recoil is used, whereas on the right TK recoil is used.	103

List of tables

1.1	Fields content of the SM from a gauge point of view, the three generations of fermions have not been distinguished since they have exactly the same gauge interactions.	6
1.2	Results of the electroweak global fit as reported in [11]. The second column contains the world average experimental value used as input in the fit; the third column contains the result of the fit to all the observables except for the one considered.	10
3.1	Summary of the uncertainties in the combined LEP II measurement of m_W based on direct mass reconstruction in semileptonic and fully hadronic channels [36].	31
3.2	Break down of the uncertainties in the W mass measurement in CDF and ATLAS analyses separately for the transverse mass and lepton momentum observable fit. All figures in MeV.	37
3.3	Break down of the uncertainties in the Z mass measurement using a W-like approach performed by the CMS collaboration. Results for events tagged as positive charged W are reported separately for the three observable used. All figures are reported in MeV. Table from [41].	38
5.1	Event observables used as input of the DNN for the semi-parametric regression. Definition of TK, PF and ntnpv can be found at the end of sec. 4.3.1; The sign of the angles is defined as $\varsigma = \text{sgn}(\Delta\phi_{TK-ntnpv})$. . .	69
6.1	Average mass bias and its uncertainty obtained in the fit between data-like Z MC and Z MC. The result obtained when different experimental recoil estimators are used is reported.	100

LIST OF TABLES

6.2	Average mass bias and its uncertainty obtained in the fit between data-like 2 Z MC and data-like Z MC. It is reported the result obtained when different experimental recoil estimator are used.	101
6.3	Average mass bias and its uncertainty obtained in the W/Z morphing cross checks. Both fits are performed to the transverse mass distribution computed using the mnpk recoil. The samples are divided into 100 sub-samples and the average of the biases obtained is reported.	104

Acknowledgments

I would like to thank Scuola Normale Superiore, for the support provided during all these years: all the means are always been more than enough to overcome the limits I faced. Now, I feel part of a collaborative and supportive environment, full stimulating professors and brilliant colleagues, which made me grown as a physicist, as a scientist, as a man.

It is also fair to mention the CERN, for its hospitality during the completion of part of this work.

Nevertheless, the results reported in this thesis would not have been possible without the support of my supervisor Gigi Rolandi which guided me throughout the work. It is a pleasure and a honor to discuss with you, not only as a professor, but also as a mentor.

A special thanks goes to Nicolo. The great experience of sharing not only an intense year of thesis work but the full university path has been unique: the team work has always allowed one plus one to be four. I hope that we will have the chance to join our efforts once again in the future.

In addition, I would like to express my deepest gratitude to all the people that advised me with fruitful comments and suggestions. In particular, I would like to mention Maurizio Pierini, Patrick Janot, Josh Benedavid and Luca Perrozzi: the long hours spent in discussing ideas and problems are the foundation of this thesis. It is also remarkable the contribution of the whole W mass group: Elisabetta, Marc and all the others. Finally, it is with immense pleasure that I thank Marianna for the extraordinary support, patience and motivation. From day one to the thesis submission, you played the crucial role of being always by my side. May the life give back all that you deserve.

Con particolare coinvolgimento, vorrei inoltre aggiungere una parte in Italiano per ringraziare tutte le persone che, anche senza un aiuto diretto sul lavoro mi hanno permesso di raggiungere i miei obiettivi, spronandomi nelle mie decisioni e aiutandomi a realizzare le mie idee.

Fino dalla prima infanzia, la crescita della mia curiosità è stata trattata con particolare cura: avvincenti stimoli e la guida al ragionamento, scientifico senza sapere che lo fosse oserei dire, hanno fondato le basi della persona che sono oggi. Per questo, oltre che per gli innumerevoli aiuti, vorrei ringraziare la mia Famiglia: Babbo, Mamma, Gaia e anche Nonna Paola, Nonna Quirina e Nonno Massimo. Questo traguardo è, in parte, anche vostro.

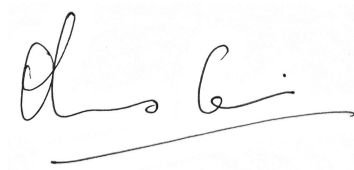
Altro essenziale ringraziamento va ai miei amici, tutti, da quelli che hanno condiviso con me pochi mesi fino a quelli di cui ho da sempre ricordo. Il mio percorso universitario non sarebbe stato neanche comparabile in assenza delle persone che ho incontrato. Tengo particolarmente a citare Fede, Marco e Nicolo: nella famiglia della normale, non esiterei a chiamarvi fratelli. Similmente ricordo con piacere come mai nessuno, specialmente degli studenti più avanti, abbia mancato di rispondere a una mia richiesta di chiarimento: farò del mio meglio per diffondere questo atteggiamento che mi avete insegnato.

Oggi, il futuro mi chiama altrove.

Tuttavia, le fondamenta della mia persona sono, e resteranno, profondamente radicate a questo periodo, forgiate da persone, eventi e dedizioni.

Semel normalista, semper normalista.

Con gratitudine e affetto,

A handwritten signature in black ink, consisting of the letters 'D' and 'G' followed by a horizontal line, likely representing 'D. G.'.