Classe di Scienze
Corso di Perfezionamento in Fisica
XXXIII ciclo

Tesi di Perfezionamento

Towards a simultaneous measurement of W boson mass and production properties with the CMS detector

---

Candidato
Valerio Bertacchi

Relatore
Prof. Luigi Rolandi

Anno Accademico 2020-2021

Se mi rilasso collasso

(Bandabardò)

Contents

Introduction

The Compact Muon Solenoid (CMS, [1]) is one of the four main detectors which operate at the Large Hadron Collider (LHC, [2]) at CERN, Geneva. LHC is a hadron-hadron circular collider, with a design center of mass energy for pp collisions of $\sqrt{s} = 14$ TeV, and it is currently the only operating machine at the TeV scale in the world. CMS is a *general purpose* experiment, build to efficiently collect and analyse the LHC collisions and to cover an extensive physic program, exploring different sectors of the Standard Model (SM), the current theory which best describes the fundamental forces and the subatomic particles.

Thanks to its high luminosity, the LHC allows to collect an unprecedented amount of events compared to previous and concurrent colliders, in particular at electroweak energy scale and above. In these years a strong effort has been performed to obtain excellent control of the machine and detector effects. A parallel effort has been made to improve the theoretical understanding to finally constrain and reduce the systematic uncertainties which affect the investigated phenomena. This has turned the LHC from a discovery into a precision machine, a tool to improve the sensitivity of precision test of the SM and to constrain Beyond the Standard Model (BSM) theories.

A closer look at the theoretical and experimental framework of SM measurements at LHC is reported in Chapter 1, while the LHC itself and the CMS experiment are described in detail in Chapter 2.

The electroweak precision observables, combined in the electroweak fit, are one of the fields of major interest. The W boson mass (m_W) is one of the most critical electroweak parameters and its measurement is part of the scientific program of the CMS Collaboration. The combination of m_W measurement from ATLAS [3] and Tevatron [4, 5] has a precision of 12 MeV and it is limited by systematic uncertainty. The m_W theoretical prediction from the electroweak fit has an uncertainty of 7 MeV and there is a tension of 1.5σ with respect to the experimental result. One of the goals of CMS for Run 2 is to perform a measurement of m_W with precision at the level of the theoretical prediction.

The W-mass is measured using the leptonic decays of the W boson and in particular the $W^{\pm} \to \mu^{\pm}\nu$ channel, because of the excellent muon reconstruction performance of CMS, in terms of efficiency, momentum calibration and background discrimination power. Nevertheless, the final state neutrino cannot be reconstructed and it leaves the signature of missing transverse energy (p_T^{miss}). Thus the invariant mass of the boson can not be directly inferred. The p_T^{miss} has a poor scale and resolution compared to the requirement of the measurement. Thus transverse plane variables, like muon transverse momentum p_T^{μ}, must be used to recover the m_W information.

The muon is the only reconstructed particle in the final state and so the control of the momentum scale calibration is one of the central ingredients to reach the required precision. To deliver a precision on the muon scale at 10^{-4} level, the Z boson mass must be used as a standard candle. The momentum distributions of the colliding partons (Parton Distribution Functions, PDFs)

produce a distortion on the Z boson lineshape, shifting in particular the mode of the invariant mass distribution. This effect can in principle produce a bias in the momentum scale evaluation and has been studied in detail in Ref. [6] and reported in the Chapter 3 of this thesis.

When m_W is extracted with a template fit from p_T^μ distribution, the limited knowledge of the W production model (transverse momentum q_T^W , rapidity Y_W and polarization) induces a systematic uncertainty the value of m_W. In particular, the finite η^μ acceptance of the CMS detector introduces a dependence of the charged lepton transverse momentum distribution on the W boson rapidity [7]. The Y_W spectrum and W polarization are determined by the PDFs of the incoming protons, and the induced systematic uncertainty on m_W (the so-called "PDF uncertainty") is quoted $\mathcal{O}(10)$ MeV in the recent ATLAS, CDF and D0 papers [3–5]. A precise measurement of the Y_W spectrum and of the W polarization can constrain the PDF uncertainty, as the proof of principle of Ref. [8] shows. A template fit can extract the Y_W and helicity (i.e. polarization) distribution from the charged lepton kinematics alone. The q_T^W is an additional source of systematic uncertainty. It is typically measured using the hadronic recoil i.e. the vectorial sum of all the particle momenta in the event but the charged lepton. However, an excellent precision can not be achieved because of the large uncertainties both from the theory and the experimental side [9, 10].

The m_W measurements performed in the past have *assumed* the W boson production properties as external input and they have treated them as systematic uncertainties for the mass extraction. This is a mandatory choice when the analysis does not have enough statistical power to afford a multi-differential measurement of q_T^W , Y_W and polarization. With an increased statistic the W production properties can be *measured*, fixing the value of m_W as the previously measured value. Then, the production properties measurement can be used as a *constraint* to the systematic uncertainties of the m_W direct measurement.

The former approach has been used in the Tevatron and ATLAS measurements. Both Tevatron and Run 1 of LHC did not have the required statistical power to measure the W boson production properties. The latter approach instead motivated the recent CMS activities. CMS has performed a measurement of W boson rapidity and helicity distributions [11] following the method of Ref. [8]. They extracted the Y_W distribution for the three helicity state (left, right, longitudinal) and the W charge asymmetry from the direct measurement of the lepton double-differential cross section $\frac{d\sigma}{d\eta^\ell dp_T^\ell}$. The W boson mass state-of-the-art at CMS is presented in Chapter 4, together with the historical background of the previous measurement and future developments.

The content of this thesis goes beyond the previously described perspective, further exploiting the idea of Ref. [8]. The statistics collected by CMS during Run 2 of LHC allows to extract the value of m_W *simultaneously* with q_T^W , Y_W and polarization spectra, to finally obtain a measurement of the mass with extremely low systematic uncertainty coming from the W production mechanism.

The vectorial nature of the W boson brings to eight angular coefficients to entirely describe the polarization in its decays. The W boson differential cross section can be expressed as a function of these angular coefficients $A_i(q_T^W, Y_W)$:

$$\frac{d\sigma}{dq_{T,W}^2 dY_W d\cos\theta_\mu d\phi_\mu} = \frac{3}{16\pi}\frac{d\sigma^{U+L}}{dq_{T,W}^2 dY_W}\Big[(1+\cos^2\theta_\mu) + \sum_{i=0}^{7} A_i P_i(\cos\theta_\mu, \phi_\mu)\Big], \tag{0.1}$$

where σ^{U+L} denotes the production cross section for unpolarized W bosons, the P_i are known functions of $\cos\theta_\mu, \phi_\mu$ for a given A_i and the angles θ_μ, ϕ_μ are measured in the dilepton system rest frame [12]. The W rapidity, transverse momentum and angular coefficients can be extracted from the muon kinematics with a $\eta^\mu \times p_T^\mu$ template fit in bins of Y_W, q_T^W for each P_i, unfolding the underlying boson distribution from the charged lepton kinematics alone.

The analysis framework is challenging due to the high level of complexity of the template fit (which has order of 10^4 templates) and the large dataset involved for signal and background estimation. A custom framework, based on `RDataFrame` `ROOT` package [13] has been developed, which runs on NanoAOD ntuples, the state-of-the-art compressed CMS data format [14].

Despite the $W^{\pm} \to \mu^{\pm}\nu$ decay channel having very high purity, two possible sources of background must be taken into account. The first is composed of electroweak processes which fall in the signal kinematic acceptance because they mimic the single muon signature (Z boson, Di-boson and top quark decays with one reconstructed muon). Their background yield is about 5% of the selected events and can be subtracted using the Monte Carlo (MC) because of the accurate description of these processes in the simulation. The second and most relevant source of background, accounting for about 7% of the selected events, are the energetic muons from multi-jet production. This QCD background is reduced by selecting isolated muons and restricting to high transverse mass region only (transverse mass $m_T = 2|p_T^{\ell}|p_T^{\nu}|\sqrt{1-\cos\phi_T^{\ell\nu}}$). However, the QCD muons from heavy flavour decays spectra partially overlap also this kinematic selection. A data-driven background estimation approach, called *fake rate* method, has been chosen to assess the QCD yield in the signal region because the MC description of multi-jet production spectra has limited precision. The idea of this method is to assess the probability of a QCD muon falling inside the isolated signal region, measuring it in a QCD enriched region at low m_T. The independence from m_T of the isolation efficiency is the only assumption of this method and therefore it provides a precise estimation of the QCD yield in the signal region without adding large systematic uncertainties.

The description of the main analysis of this thesis is presented in Chapter 5. The measurement of the background of the analysis is reported in Chapter 6, while the description of the fit framework and the results is in Chapter 7 and Chapter 8. The technical details of the experimental framework are given in Appendix D.

Unfortunately, at the time of this work, the required calibrations are not available for the entire Run 2 dataset and only the data and MC samples corresponding to 2016 data taking can be used. With this reduced sample, the results are therefore presented as a proof of feasibility for what concerns the simultaneous fit to the mass. The results can be instead interpreted as an additional and more refined simultaneous measurement of W boson transverse momentum, rapidity and polarization, fixing m_W to the current measured value. The final result on Y_W, q_T^W and polarization has not been extracted from data, but the analysis has been performed on the CMS simulation, keeping the real parameters of interest blinded.

A relevant contribution to the CMS reconstruction software is also described in this thesis. This is not directly related to W boson production properties measurement and is therefore reported in Appendix A, but this development required a non-negligible fraction of the PhD time and involved CMS reconstruction software and machine learning wide studies that globally improves the tracking performance of the CMS experiment. Tracking in high-density environment, such as high energy jet cores ($p_T^{\text{jet}} \gtrsim 0.5$ TeV) is particularly challenging. The CMS tracking algorithm is based on a Combinatorial Kalman Filter (CKF) which uses track-seed built from the pixel detectors hits. The CKF does multiple iterations with different optimized track-seeds to cover the entire kinematic range. The high track density of the core of the high energy jet produces merged clusters on the pixel detector layers which are difficult to split and assign to the correct track. This problem has been addressed in the past with a dedicated CKF iteration and cluster splitting algorithms, working layer by layer, followed by a pattern recognition step where a high number of candidate tracks are tested [15]. Nevertheless, the quality of the seeds is this region was still suboptimal, degrading the final tracking performance. Modern Deep Learning techniques can be used to better handle the high energy jet seeding: the basic idea is to use an Artificial Neural

Network to obtain the track-seeds parameters of the list of track inside the jet core directly from the unclustered pixel information. The critical step, the cluster splitting, is skipped with this approach and the information of four pixel detector layers is combined to predict the track parameters. The developed algorithm is a Convolutional Neural Network [16], which takes as input the pixel hit position and charge in four windows aligned with the jet axis and it returns as prediction the track-seeds parameters of a variable number of seeds. The network has shown very promising performances [17], and will be integrated into the CMS reconstruction for Run 3 of LHC. Moreover, the good performance of DeepCore gives the chance of an extension DeepCore approach in the other steps of pattern recognition where there is still room for improvement.

Notation

In this thesis, the typical particle physics notation will be adopted. In particular $\hbar = c = 1$, when not otherwise specified, where c is the speed of light and $\hbar$ the reduced Plank constant, $e \simeq 1.602 \cdot 10^{-19}$ C is the absolute value of the electron charge. With this convention the mass, momentum and energy are measured in electronvolt (eV), while time and length in eV^{-1}. The vectors are indicated in bold.

Chapter 1

The Standard Model and the electroweak precision tests

The Standard Model of particle physics is currently our best theory to describe the microscopic world in term of elementary particles and fundamental forces i.e. interactions between these particles. In this Chapter a brief summary of the SM is given, focusing then on the electroweak sector. The experimental status and challenges concerning the SM understanding are described, also in this case with attention for the electroweak precision tests and in particular for what concerns the W boson mass. The aim of this Chapter is to provide the theoretical framework needed for the W boson mass measurement, for a more exhaustive description and discussion on the SM see for instance Ref. [18–20].

1.1 Introduction to the Standard Model of particle physics

The SM is a quantum field theory that aims to describe the electromagnetic, weak and strong interactions and the elementary particles affected by these interactions. The theory is renormalizable since the ultra-violet divergences which typically arise from loop diagrams in the SM can be cancelled with a proper redefinition of the parameters of the Lagrangian itself.

The matter content of the SM is realized with fermions spin-$\frac{1}{2}$ fields, divided in two categories: quarks and leptons. The elementary fermions which interact with strong force are called quarks. They exist in 6 *flavours*, organized in 3 doublets or *generations*. Each doublet contains an up-type and a down-type quark, with electric charge $+\frac{2}{3}e$ and a $-\frac{1}{3}e$ respectively, where e is the absolute value of the electron charge. The same pattern is repeated for the antiquarks, with inverted charge and flavour quantum numbers.

The elementary fermions which do not interact with strong force are called leptons. The leptons also exist in 3 generations and they are organized in doublets. An electrically charged (with charge $-e$) massive lepton and a neutral massless lepton called neutrino are present in each doublet. The antimatter content of the lepton sector has the same scheme, with inverted quantum numbers.

The SM is a non-abelian gauge theory, therefore the Langrangian is invariant under local (i.e. space-time dependent) transformations according to its gauge symmetry group, and the generators of this group are not commutative. These symmetries reflect in redundancies of the field

description of the theory. Each gauge symmetry requires the addition of a proper number of gauge fields to the theory, according to the number of generators of the symmetry group. These gauge fields can be physically interpreted as force carriers which are exchanged between the interacting matter fields, and they are called *gauge bosons*. The SM symmetry group is:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y. \tag{1.1}$$

In particular, the quantum chromodynamics (QCD) theory, which describes the strong interactions, implements the $SU(3)_C$ symmetry. This group has 8 generators, which produce 8 gluon fields G^a. The QCD has 3 charges called *colours* and the quark fields are triplets under $SU(3)_C$, therefore they can exist in 3 states with a different colour. The antiquarks have also inverted colour charges. On the other hand, the leptons are singlet under $SU(3)_C$ i.e. they do not have a colour charge. The gluons themselves have one colour and one anticolour charge, and therefore the gluons self-interactions are possible.

The electroweak interaction implements the $SU(2)_L \otimes U(1)_Y$ symmetry. The group $SU(2)_L$ has three generators which reflect in the boson fields W^i. The left-handed chirality fermions are represented as $SU(2)_L$ doublet, while the right-handed chirality fermions are $SU(2)_L$ singlets. In particular, the quarks are described with the doublet $Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}$ and two singlets u_R, d_R, while the leptons are described with one doublet $L_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$ and one singlet e_R. This scheme is repeated for the three generations of quark and leptons, and the chiral behaviour of the antiparticles is inverted. The right-handed neutrino singlet ν_R is not included in the SM, but recent discoveries may suggest its existence. Further discussion about that will follow in Sec. 1.1.1.

The group $U(1)_Y$ has one generator which produces the boson field B and the related charge is called *hypercharge*. The summary of the relevant quantum numbers for the SM fields are reported in table 1.1.

Table 1.1. *Representations of $SU(3)_C$ and $SU(2)_L$ and Q, T_3 and Y quantum numbers for the SM matter and gauge fields, where $Q = T_3 + Y/2$; only a single fermion generation is shown, since they have the same quantum numbers.*

Field	$SU(3)_C$	$SU(2)_L$	Q	T_3	Y
$L_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix}$	**1**	**2**	$\begin{pmatrix} 0 \\ -1 \end{pmatrix}$	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$	-1
e_R	**1**	**1**	-1	0	-2
$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}$	**3**	**2**	$\begin{pmatrix} +2/3 \\ -1/3 \end{pmatrix}$	$\begin{pmatrix} +1/2 \\ -1/2 \end{pmatrix}$	$+1/3$
u_R	**3**	**1**	$+2/3$	0	$+4/3$
d_R	**3**	**1**	$-1/3$	0	$-2/3$
g	adj. **8**	**1**	0	0	0
B	**1**	**1**	0	0	0
W	**1**	**3**	± 1	± 1	0
H	**1**	**2**	0	$-1/2$	$+1$

The Lagrangian of the standard model[1] can be organized in the following form:

$$\mathcal{L}_{SM} = \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}} + \mathcal{L}_{\text{H}} + \mathcal{L}_{\text{Yuk}}. \tag{1.2}$$

[1] $\mathcal{L}$ is a Lagragian density, with the dimension of eV^4, despite it is called Lagrangian in the entire chapter for simplicity, defined by $\mathcal{S} = \int \mathscr{L} dt = \int \mathcal{L} d^4x$, while $\mathcal{S}$ is the dimensionless action and $\mathscr{L}$ is the true Lagrangian.

The $\mathcal{L}_{\text{gauge}}$ term includes the kinetic term and the self-interaction of the gauge fields G^C_μ, B_μ, W^i_μ associated to $SU(3)_C$ colour, $U(1)_Y$ and $SU(2)_L$ left symmetries:

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}\sum_{C=1}^{8}\mathcal{G}^C_{\mu\nu}\mathcal{G}^{\mu\nu,C} - \frac{1}{4}\mathcal{B}_{\mu\nu}\mathcal{B}^{\mu\nu} - \frac{1}{4}\sum_{i=1}^{3}\mathcal{W}^i_{\mu\nu}\mathcal{W}^{\mu\nu,i}, \tag{1.3}$$

where i runs over the 3 $SU(2)_L$ fields, while C runs over the $N^2_{\text{colour}} - 1 = 8$ $SU(3)$ gluon fields, and the fields are expressed via the field strength tensors defined as:

$$\begin{aligned}
\mathcal{G}^C_{\mu\nu} &= \partial_\mu G^C_\nu - \partial_\nu G^C_\mu - g_s f_{CAB} G^A_\mu G^B_\nu \\
\mathcal{B}_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu \\
\mathcal{W}^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - g\varepsilon_{ijk} W^j_\mu W^k_\nu,
\end{aligned}$$

The g_s and g are the strong and weak coupling constants, respectively. The f_{CAB}, ε_{ijk} are the structure constant of $SU(3)$ and $SU(2)$ groups (the latter are simply the totally antisymmetric Levi-Civita tensor). They allow to define the commutation rules of the groups, given the infinitesimal generators t^C, Y, T^i for $SU(3)$, $U(1)$ and $SU(2)$ respectively:

$$[t^A, t^B] = i f_{ABC} t^C, \quad [T^i, T^j] = i\varepsilon_{ijk} T^k.$$

Identifying the Y and T^i as the weak hypercharge and the weak isospin is now possible to define the electric charge, in units of e, as $Q = T_3 + Y/2$, where T_3 is the third component of the weak isospin.

The dynamical term $\mathcal{L}_{\text{QCD}}$ describes the interaction between the gluons and the quark fields, accordingly to QCD:

$$\mathcal{L}_{\text{QCD}} = \sum_{q=1}^{6}\sum_{a,b=1}^{3}\sum_{C=1}^{8}\bar{\psi}_{q,a} i\Big(\frac{1}{2} i\gamma^\mu g_s \lambda^C_{ab} G^C_\mu\Big)\psi_{q,b}. \tag{1.5}$$

The $\psi_{f,a}$ are the quark field of flavour q and colour a, the $\lambda^C = 2t^C$ are the Gell-Mann matrices, γ^μ are the Dirac matrices. From g_s the $\alpha_s \equiv \frac{g_s^2}{4\pi}$ can be defined. Note that $\mathcal{L}_{\text{QCD}}$ should also include the term $+\gamma^\mu\partial_\mu\delta_{ab}$ inside the parenthesis, to be a consistent description of QCD, but this term will be included in Eq. 1.6, and thus is omitted here to make Eq.1.2 consistent.

The dynamical term $\mathcal{L}_{\text{EW}}$ describes the interaction between the $SU(2)_L \otimes U(1)_Y$ gauge bosons and the fermions fields.

$$\mathcal{L}_{\text{EW}} = \sum_{g=1}^{3}\sum_{f=\{Q_L, L_L, u_R, u_L, e_R\}} i\bar{\psi}_{f,g}\gamma^\mu D_\mu \psi_{f,g}, \tag{1.6}$$

where g runs on the 3 generations, f runs on the various fermions classes and the covariant derivative is defined as:

$$D_\mu = \partial_\mu + \frac{1}{2} i g' Y B_\mu + \frac{1}{2} i g \sum_{i=1}^{3}\tau^i W^i_\mu,$$

where i runs on the $SU(2)_L$ generators, $\tau^i = 2T^i$ are the Pauli matrices, g' is an additional coupling of the electroweak sector. Note that for all the $SU(2)_L$ singlets $\tau^i W^i \psi_R = 0$. It is also useful to remind that given a generic fermion field ψ the left- and right-handed chiral projection can be obtained by:

$$\psi_{L,R} = \Big(\frac{1 \mp \gamma^5}{2}\Big)\psi, \quad \bar{\psi}_{L,R} = \bar{\psi}\Big(\frac{1 \pm \gamma^5}{2}\Big).$$

The SM observed phenomenology is not properly reproducible with the component of the Lagrangian described until now, $\mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{QCD}} + \mathcal{L}_{\text{EW}}$. The electroweak sector requires 3 massive vector bosons, two charged and one neutral ($W^{\pm}$ and Z) and one massless neutral boson (γ) to implement the observed currents. Moreover, also the observed quarks and charged leptons are massive.

This result is obtained with the addition of the two terms $\mathcal{L}_H$ and $\mathcal{L}_{\text{Yuk}}$, which however require the existence of an additional complex scalar boson field, called the Higgs boson. The $\mathcal{L}_H$ describes the interaction of the Higgs boson field with the gauge fields and with itself:

$$\mathcal{L}_{\text{H}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi), \quad \text{with} \quad V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \tag{1.7}$$

where $\phi = \begin{pmatrix} \phi_+ \\ \phi_0 \end{pmatrix}$ is the Higgs boson field, a complex doublet of $SU(2)_L$ with four degrees of freedom. The Higgs potential $V(\phi)$, with the choice of $\lambda > 0$ and $\mu^2 < 0$, has a degenerate minimum for

$$\phi^\dagger \phi = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}.$$

This means that ϕ assumes a non-zero vacuum expectation value (VEV):

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

inducing a spontaneous breaking of the $SU(2)_L \otimes U(1)_Y$ symmetry (EWSB) with the new vacuum condition. Nevertheless, the subgroup generated $Q = T_3 + \frac{1}{2}Y$ is not broken, realizing the $U(1)_{em}$ symmetry of the SM i.e. the electric charge conservation. From the 3 broken generators, the same number of Goldstone bosons are produced. However, due to the degeneracy of the ground state, a proper gauge transformation (unitary gauge) can reabsorb these 3 degrees of freedom into the mixed broken generators of $SU(2)_L \otimes U(1)_Y$, which now acquire mass as a longitudinal component. The remaining degree of freedom is the physical Higgs boson, a scalar boson, neutral under $U(1)_{em}$ and a singlet of $SU(3)_C$.

After the EWSB the boson fields can be rewritten in the unitary gauge as:

$$A_\mu \equiv B_\mu \cos\theta_W + W^3_\mu \sin\theta_W, \tag{1.8a}$$

$$W^\pm_\mu \equiv \frac{1}{\sqrt{2}}(W^1_\mu \pm i W^2_\mu), \tag{1.8b}$$

$$Z^\mu \equiv -B_\mu \sin\theta_W + W^3_\mu \cos\theta_W, \tag{1.8c}$$

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{1.8d}$$

which are respectively the photon (γ), electroweak charged ($W^\pm$), neutral (Z), and Higgs boson field, while the weak mixing angle is defined as:

$$\theta_W \equiv \tan^{-1}\left(\frac{g'}{g}\right) \quad \text{or equivalently} \quad \cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}}.$$

Defining the absolute value of the electric charge $e = g \sin\theta_W$ and the fine-structure constant $\alpha \equiv \frac{g^2 g'^2}{4\pi(g^2+g'^2)} = \frac{1}{4\pi} g^2 \sin^2\theta_W$ the entire set of boson mass can be expressed as (at tree level):

$$\begin{aligned} m_Z &= \frac{v}{2}\sqrt{g^2 + g'^2} = \frac{ev}{2\sin\theta_W \cos\theta_W}, & m_{W^\pm} &= \frac{v}{2} g = \frac{ev}{2\sin\theta_W} = m_Z \cos\theta_W, \\ m_\gamma &= 0, & m_H &= \sqrt{2\lambda} v = 2|\mu|. \end{aligned} \tag{1.9}$$

The interactions between the fermions fields and the Higgs field are described with the Yukawa term of the Lagrangian:

$$\mathcal{L}_{\text{Yuk}} = -y^d_{ij}\bar{Q}_{L_i}\phi d_{R_j} - y^u_{ij}\bar{Q}_{L_i}\tilde{\phi}u_{R_j} - y^e_{ij}\overline{L}_{L_i}\phi e_{R_j} + h.c., \tag{1.10}$$

where $\tilde{\phi} = i\sigma_2\phi^*$ and y^f_{ij} yukawa coupling matrix for f =up, down, or charged lepton fermions, and the indices i, j run on the 3 generations. Once the electroweak symmetry breaking occurs, the y^f_{ij} can be diagonalized using the fermion mass eigenstates basis $y^f_{ij} \to y^f_i$, and all the fermions acquire a mass $m^f_i = \frac{1}{\sqrt{2}}y^f_i v$.

After the Yukawa matrices diagonalization, the quark electroweak doublets and singlets Q_L, d_R should be replaced by $Q'_L = \begin{pmatrix} u \\ d' \end{pmatrix}$, d'_R, where $d'^i_{L,R} \equiv \sum_j V^{ij} d^j_{L,R}$ and V is the 3×3 Cabibbo-Kobayashi-Maskawa unitary matrix (CKM). Thus, from now on Q_L, d_R will be the mass eigenstates, Q'_L, d'_R the $SU(2)_L$ ones.

The EWSB via the Higgs field and the consequent production of the mass terms in the Lagrangian is called the *Higgs mechanism*. Finally, it is possible to isolate the electroweak interaction of the fermions rewriting the Lagrangian as follow:

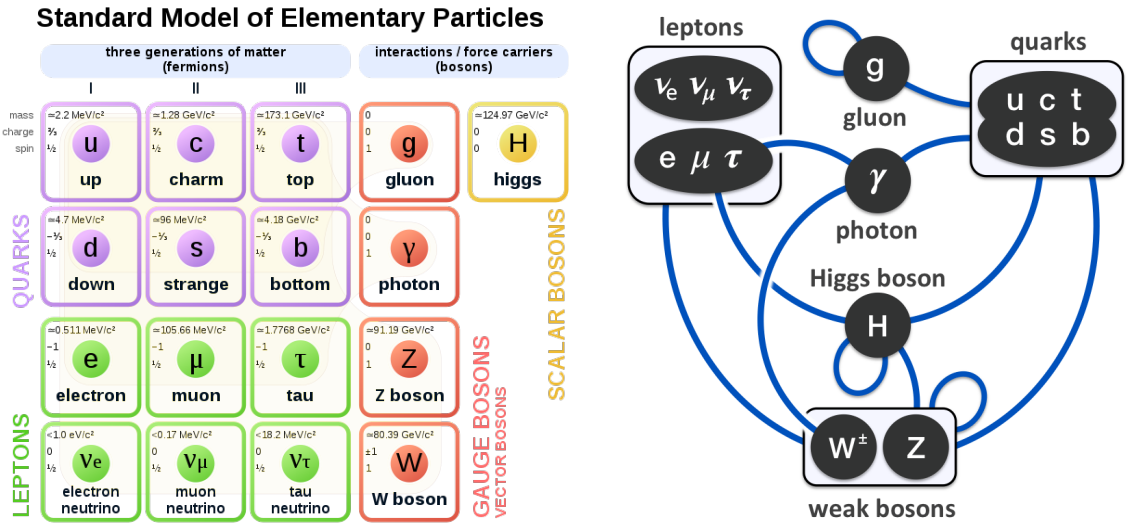
$$\begin{aligned}
\mathcal{L}_f = &\sum_f \bar{\psi}_f\Big(i\partial_\mu\gamma^\mu - m_f - \frac{m_f H}{v}\Big)\psi_f \\
&- \frac{g}{2\sqrt{2}}\sum_{g=1}^{3}\sum_{d=\{Q'_L, L_L\}} \bar{\psi}_{g,d}\gamma^\mu\big(\frac{1-\gamma^5}{2}\big)(t^+W^+_\mu + t^-W^-)\psi_{g,d} \\
&- e\sum_f Q_f\bar{\psi}_f\gamma^\mu\psi_f A_\mu \\
&- \frac{g}{2\cos\theta_W}\sum_f \bar{\psi}\gamma^\mu(g^f_V - g^f_A\gamma^5)\psi_f Z_\mu,
\end{aligned} \tag{1.11}$$

where f runs on all the fermion SM fields (quarks and leptons of all the generations), g runs on the 3 generations, d runs on the $SU(2)$ doublets. The first line represents the diagonalized Yukawa coupling of the Higgs field with the fermions, the second line the coupling with the charged weak bosons (where $t^\pm$ are the weak isospin raising and lowering operators), the third line describes the coupling with the photon filed (i.e. the QED), and the last line the coupling with the weak neutral boson (where $g^f_V \equiv T_3(f) - 2Q_f\sin^2\theta_W$, $g^f_A \equiv T_3(f)$ are the vector and axial coupling, respectively).

The table of all the elementary particles of the SM, their relations and coupling is reported in Fig. 1.1.

1.1.1 Phenomenology of the Standard Model

After the EWSB, the electroweak sector observed phenomenology is well described by the SM. The electromagnetic interaction occurs via the photon neutral current while the massive weak boson has been observed with $m_Z \simeq 91.2$ GeV, $m_W = 80.4$ GeV and $\sin^2\theta_W \simeq 0.23$. The EWSB scale v has been predicted using the Fermi Constant $G_F \equiv \frac{\sqrt{2}}{8}\frac{g^2}{m^2_W}$, measured from muon decay time with a relative precision of 10^{-7}. Indeed $v = (\sqrt{2}G_F)^{-1/2} \simeq 246$ GeV, which represents the only absolute energy scale of the SM and it is also called the electroweak scale. Moreover, the Higgs boson has been experimentally observed [23, 24], and its mass is $m_H \simeq 125$ GeV.



(a) (b)

Figure 1.1. *(a): Summary table of the elementary particles of the Standard Model, with mass, electric charge, and spin also shown for each particle (from Ref. [21]). (b): Diagram of the coupling between the particles of the Standard Model (from Ref. [22]).*

This result has two relevant consequences. Given the observed m_H, the SM lies in the stability region (i.e. the vacuum minimum of the EWSB is a global minimum), very close to the metastability region (i.e. the EWSB minimum is local and the global minimum is placed at higher energy). This condition holds up to the m_{plank} scale ($m_{\text{plank}} \equiv \sqrt{\hbar c/G} \sim 10^{19}$ GeV, where G is the gravitational constant, is the energy scale at which gravity effects are expected at quantum scale) [25]. This implies that there is no requirement of additional particles at higher energy, from a SM consistency point of view.

On the other hand, the electroweak scale, and so m_H, is so small compared to m_{plank} (the only other absolute scale of the model, which parametrizes our ignorance), that a strong fine-tuning of the SM is needed to make m_H stable when the radiative corrections are taken into account. The Higgs boson mass, considering the radiative correction, should be expressed as:

$$m_H^2 = (m_H^0)^2 + (\delta m_H)^2, \quad \text{with: } m_H^0 = \sqrt{2\lambda}v, \quad (\delta m_H)^2 \sim \left(\sum_i {m_i}^2 + \lambda\right)\Lambda^2,$$

where m_H^0 is the tree-level mass while δm_H represent the higher order corrections to the mass, i runs over the Higgs field coupling and Λ is the energy cut-off used in the loop evaluation. The value of Λ is completely arbitrary and constrained by the range of validity of the SM only. Considering the SM stable up to m_{plank}, Λ can be set up to 10^{19} GeV, making $(\delta m_H)^2/(m_H)^2 \gg 1$. It means that a very fine-tuned cancellation is required between m_H^0 and δm_H to realize the observed m_H. This is the so-called *naturalness* [26] or *hierarchy* problem: a fine-tuning of the free parameters of a model is a quite unlikely situation, and it may suggest the existence of Beyond the Standard Model (BSM) physics in the energy range $v < E_{BSM} < m_{\text{plank}}$.

With this assumption, or the mass (SM coupling) of these BSM particles is too high (low) and they have not been observed yet, or the observed boson is not the Higgs predicted by the SM. However, the properties of the observed scalar particle with mass ∼125 GeV have been investigated in details in these years and the agreement with the SM Higgs boson is excellent. In particular, the coupling of the Higgs boson with the weak gauge field and the heavier fermion

fields have been observed showing good consistency with the SM prediction. The state-of-the-art of CMS experiment measurements is shown in Fig. 1.2.

For what concerns the electroweak quark sector a rich phenomenology emerges. The different Yukawa couplings of the quarks produce a not-trivial mass scheme whose eigenstates are different with respect to the electroweak one. Since the flavour quantum number is assigned to the mass eigenstates Q_L, u_R, d_R, but the rotated Q'_L, u_R, d'_R are involved in the weak interaction vertices, the weak interaction does not conserve the flavour of the particles and flavour mixing occurs. On the other hand, the flavour of the quarks is conserved by strong and electromagnetic interactions. Moreover, the CKM matrix includes a complex phase that allows $\mathcal{CP}$ violating terms. An exhaustive review of both theoretical framework and experimental results in the flavour sector and $\mathcal{CP}$ violation can be found in Ref. [27].

The lepton sector has a much simpler phenomenology, due to the vanishing neutrino masses and the exclusion of ν_R from the SM. The 3 lepton generations are identical copies that differ only for the Yukawa coupling (*lepton universality*). This induces 3 accidental $U(1)_{e,\mu,\tau}$ symmetries to the SM. The related conserved quantities are the *lepton family numbers* L_e, L_μ, L_τ, quantum numbers conserved in all the SM interactions. Each lepton has the family number of its generation equal to 1, and the other numbers equal to 0. The quarks (and the bosons) do not carry lepton numbers. The lepton number L is defined as the sum of the 3 family numbers, and it is also conserved in the SM, protected by the relative $U(1)_\ell$ symmetry.

However, neutrino oscillations have been observed i.e. the change of neutrino lepton flavour in the propagation (see Ref. [18] for details). This phenomenon can be explained by assigning masses to the ν_L, and implementing a mixing matrix between the mass and the electroweak eigenstates. The mass term can be generated with different extensions of the SM. A possible approach postulates the existence of sterile (i.e. not SM interacting) ν_R, which can plug the neutrinos in the EWSB and generate their masses via the Yukawa couplings of the Higgs mechanism. This approach treats the neutrinos as Dirac fermions, like the rest of the SM fermions, and breaks the $U(1)_{e,\mu,\tau}$ symmetries leaving $U(1)_\ell$ unbroken. However, this approach does not explain why the neutrino masses i.e. the Yukawa couplings are so small ($< \text{eV}$) compared to the relative up-component of the $SU(2)_L$ doublets. Another possible approach is to treat the neutrinos as Majorana particles i.e. considering ν and $\bar{\nu}$ the same particle. In this case, the masses are directly generated by the Majorana operator, without involving the EWSB, and breaking also $U(1)_\ell$. Assuming Majorana neutrinos the ν_R can both included as very high mass particles, or excluded. Since the neutrino mass information collected by the experiments is not complete, a final answer about the neutrino nature and its inclusion in the SM cannot be given yet.

The EWSB affects the strong sector producing the quark mass scheme, which enters in $\mathcal{L}_{QCD}$ and produces the vast phenomenology of the quark strong bound states. In contrast, the gluon fields are singlet under $SU(2)_L$ and do not couple with the Higgs boson. The strong interaction has two critical features to describe its phenomenology. The quarks cannot be observed free because of the *colour confinement*, which forbids the observation of coloured particles. On the other hand, the value of the strong coupling constant α_s decreases with the increase of the transferred momentum q^2 of the interaction, therefore the QCD is an *asymptotically free* theory. The latter property allows the quarks to directly interact at high energy, while the colour confinement produces tight-bounded colourless observable states, the hadrons. The hadrons are subdivided in mesons (formed by a quark and an anti-quark) and baryons (formed by three quarks or anti-quarks). An additional accidental symmetry, $U(1)_B$, arises from the QCD sector. The conserved quantity is the *baryon number* $N_B \equiv N_q - N_{\bar{q}}$, which imply that the lightest baryon (the proton) must be stable in the SM.

The $\mathcal{CP}$ violation should be present also in the strong sector [18]. A natural term

$$L_\theta = \sum_{a=1}^{8} -\bar{\theta}\frac{\alpha_s}{8\pi}G^{\mu\nu,C}\tilde{G}^C_{\mu\nu}$$

is expected in $\mathcal{L}_{\text{QCD}}$, where $\tilde{G}^C_{\mu\nu} = \varepsilon_{ijkl}G^C_{kl}/2$, and $\bar{\theta}$ is the effective θ after the mass diagonalization. However experimental constraints, mainly based on neutron dipole moment, set $|\bar{\theta}| \lesssim 10^{-10}$, thus there is no evidence of $\mathcal{CP}$ violation in the QCD sector. This very unnatural value is another example of fine-tuning in the SM and it is called *strong $\mathcal{CP}$ problem*. As previously discussed for the m_H radiative correction, it can hide some symmetries which require $\theta = 0$. In particular, a Peccei-Quinn symmetry can explain $\theta = 0$: a $U(1)_{PC}$ symmetry is added, spontaneously broken by the not vanishing VEV of a related scalar field. The axion is the resulting Goldston boson, eventually reabsorbed by the gauge boson, whose field ϕ_A cancels out the strong $\mathcal{C}P$ violating term L_θ. The axion is expected to be a very light scalar boson, which can weakly couple with the SM fields, but no experimental evidence has been found so far.

The SM does not conserve any of the simpler discrete symmetry: the charge conjugation $\mathcal{C}$, which inverts the internal quantum number of the field, the parity $\mathcal{P}$ which inverts the spatial and momenta coordinate, and the time reversal $\mathcal{T}$ which inverts the time arrow, are violated by the SM Lagrangian as well as the combined $\mathcal{CP}$ in the electroweak sector. However, the combination $\mathcal{CPT}$ is a symmetry of the SM.

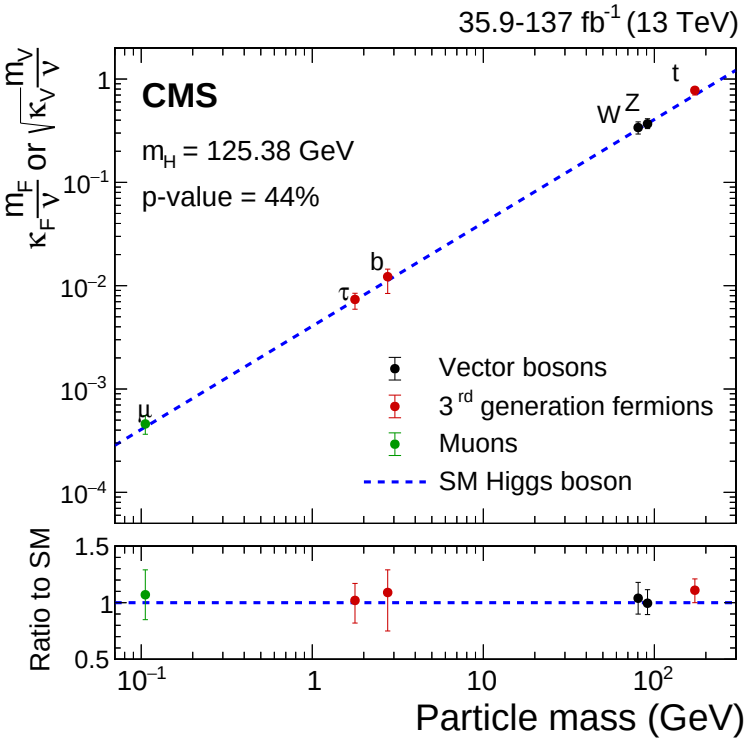


Figure 1.2. *Best fit estimates of the Higgs reduced coupling strength modifier of fermions and weak bosons, compared with the SM prediction. The uncertainty bars represent the 68% CL intervals of the measured parameters. The reduced coupling strength modifier are defined $\sqrt{\kappa_V}m_V/v$ and $\kappa_F m_F/v$ for weak bosons and fermions respectively, since the Yukawa coupling is proportional to m_V^2 and m_F respectively; the $k_i = 1$ are expected in the SM (from Ref. [28]).*

1.1.2 LHC motivation and Standard Model challenges

The LHC has been built as a discovery machine: in the early 2000s the high energy physics community focus was on the search for the Higgs boson. This was the last Standard Model building block not directly observed by previous experiments. The Tevatron [29] and Large Electron-Positron collider (LEP) [30] experiments nevertheless provided a strong constraint on Higgs boson mass window [31–36]. Moreover in the SM, the WW boson scattering cross section will violate unitarity at ∼1 TeV if the Higgs-exchange diagrams are not included in the perturbative calculation. Therefore the LHC has been built protected by a *No-Lose theorem*: either the Higgs boson would be discovered or new physics would appear investigating the TeV energy scale.

The discovery of a scalar boson resonance at $m_H \sim 125$ GeV was announced together by CMS [23] and ATLAS [24] experiments, in 2012. The new particle was compatible with the expectation for the Standard Model Higgs boson. This discovery and the measured value of the Higgs boson mass changes the following path of the particle physics community. Since the '50s the accelerators were the leading experimental tools for particle physics to reach new discoveries, because of the increasing energy and statistical requirement of the investigated phenomena, with rich and proficient feedback between the theoretical and experimental sides of the community. The framework of the Standard Model of particle physics was built in this period, and its prediction leads all the discoveries of the last decades of the 20th century. These predictions ended with the Higgs boson discovery. As previously discussed the measured m_H value implies a (meta)stability of the SM up to the Plank scale and this in particular means that after 2012 there is no more the "guarantee" of the discovery of new particles.

The SM had, and still has, a strong predictive power, being able to anticipate by years the experimental discoveries with accurate predictions. However, cosmological and astrophysical measurements reveal phenomena that have no explanation in the SM, extrapolating to the proper space and energy scale. An example of a macroscopic effect is the *dark matter* i.e. gravitation anomalies in astrophysical objects that can be described assuming the presence of a non-electromagnetic matter [18]. Another example is the *baryon asymmetry* i.e. the asymmetry in the observable universe between the baryonic matter and antimatter, with a magnitude that cannot be traced back to SM predicted $\mathcal{CP}$ violation [18]. These macroscopic effects could have fundamental sources at the microscopic level which are not included in particles and interactions of the SM.

The neutrinos oscillation can be seen as a successful example of facing these SM limits, despite the puzzle is not completely solved yet. A phenomenon, the neutrinos oscillation, has been observed (in a not-accelerator frame), and the theory developed extensions of the SM to include the observed effect. Given the extended theory dedicated searches are performed to identify the proper extension between the proposed solutions.

The tensions which the model itself shows in several precision observables between the measurements and the theoretical predictions are another hint of BSM physics. Multiple discrepancies arise from the heavy flavour sector [37], which investigate the SM processes related to the heavier quark generations, where the low energy QCD effects are smaller and the SM prediction can be very precise. These tensions can be hints of BSM physics, but a statistical fluctuation or a systematic effect of the measurements cannot be excluded at current experimental significance.

The third reason for further investigation of high energy physics phenomena is the already discussed fine-tuning, which appears both in the hierarchy problem and in the strong $\mathcal{CP}$ problem. These phenomena suggest *strongly unnatural* SM parameters, which can hide some additional symmetries that actually constrain the fine-tuned parameters to be exactly the specific observed value. Such a discovery hints that the current Standard Model is a sector of a wider unknown

theory. Besides that, many other questions regarding why the SM has exactly the observed structure can be risen, like why 3 generations are present or why so different Yukawa coupling are observed. But there is the chance that this kind of questions will receive an answer as soon as the immanent problems discussed before will be solved.

Over the years the theory developed a large number of models to describe and explain these effects. From the experimental point of view, these models are continuously tested, thus the space of the allowed parameters is constrained and new exclusion limits are set. Therefore the current limitations are the experimental data capability to constrain the theoretical models, and so a big effort is done to extend the investigable phase space of the experiments. This is pursued with two approaches: the *energy frontier* and the *intensity frontier*. In the former, the investigable energy range is extended, with direct searches of BSM phenomena not accessible at lower energy. This is typically done with hadron colliders, where the center-of-mass energy can be more easily increased with larger rings and more powerful magnets, equipped with multi-purpose detectors to proficient study the new available kinematic range. On the other hand in the intensity frontier, precision measurements are performed in the known kinematic range, looking for rare (i.e. low cross section) effects and tiny deviation from SM expectation. This approach allows to constrain BSM models at higher energy scales not directly accessible and is typically pursued with lepton collider with multi-purpose detectors or with more dedicated experiments, with ad hoc features to investigate processes of particular interest.

The LHC holds the world energy record, and it was developed and optimized to study an unprecedented energy range. The CMS and ATLAS experiments were designed primarily to look for direct searches of new particles. Nevertheless, they have progressively turned into (also) precision experiments, thanks to the high luminosity and the excellent calibration of the detectors achieved after years of operations. However, with the *caveat* of the neutrino sector, no BSM physics has been observed so far.

1.2 Electroweak precision tests

The electroweak sector offers both very precise theoretical predictions and clear experimental signatures, which helps to investigate this class of processes. A simultaneous fit of the SM theoretical prediction to the measurements in the electroweak sector can be performed to predict missing parameters which are not directly measured or to check the internal consistency of the model. This procedure is called *electroweak global fit* [18, 38, 39].

The SM, as described in Sec. 1.1, has 18 free parameters. Nine parameters come from the Yukawa coupling of 6 quarks and 3 charged leptons. Four parameters are encoded in the CKM matrix (3 mixing angles and the complex phase). The remaining 5 parameters describe the electroweak sector, and different combinations can be chosen. For instance g, g', g_s, μ, λ is a set which use only Lagrangian-level parameters. Since the latter is not directly accessible from an experimental point of view, a more common set is instead: α, G_F, α_S, m_H, m_Z.

The electroweak precision observables (EWPO) are quantities derived from the aforementioned SM free parameters, like the boson masses or width, the mixing angles, the asymmetry in boson decays or the boson branching ratios. They can be directly measured and used as an input for the electroweak global fit. The likelihood function of the fit is built with the measured EWPO and their correspondent theoretical expression, plus some additional theoretical constraint from model assumptions or prior knowledge. The fit returns both the fitted value of the input EWPO and the best-fit value of the SM parameters, predicted by the combined information of the fit.

The electroweak global fit has been used to successfully predict the SM parameters without the direct experimental information both in the case of top quark and Higgs boson, until their discovery. Both m_t and m_H enter in the theoretical expression of several EWPO, and thus the fit is capable to constrain them. After the discovery of the Higgs boson, all the fundamental parameters of the SM are experimentally measured and the fit can be used with different purposes:

- Predict all the EWPO, given the full experimental and theoretical information, obtaining a test of internal consistency of the SM and the best-fit value of the SM parameters.
- Extend the likelihood of the electroweak global fit with some BSM assumption, which modifies the EWPO dependence from the parameters, and thus test the model with the measurements of all the EWPO. This is typically implemented by adding effective field theory (EFT) operators to the SM Lagrangian, which parametrize the BSM effects. These additional parameters can be consequently extracted from the EWPO resulting from the extended electroweak global fit.
- Remove one or more EWPO measurements and predict them in an unbiased way, to allow a comparison between the best-fit value and the direct measurement of the chosen EWPO. This feasibility of this method, called *indirect determination*, relies on the overconstraint of the electroweak sector given by the large number of the available correlated EWPO.

The latter approach is the most interesting usage for the purposes of this thesis since the agreement of the W boson mass direct measurement with the theoretical prediction and the global SM experimental picture can be tested in this way. Further discussions follow in the next section.

The electroweak global fit has been performed by several groups, with different choices and technical implementation. In this thesis, the results of the `GFitter` group are reported. More details can be found in their last publication, on Ref. [40], and in the previous works [41–44]. In this framework the fit uses as input 22 EWPO (listed in Fig. 1.3), which include the state-of-the-art experimental and theoretical results. Moreover, the value of G_F, since it is measured at low energy with very high precision, is used but its value is fixed in the fit. On the other hand, the measurement of $\alpha_s(m_Z^2)$ is not included in the likelihood, leaving the parameter free, since it brings a negligible improvement of the overall precision.

Several EWPO are calculated taking into account of the radiative correction, at next-to-leading-order (NLO) or next-to-next-to-leading-order (NNLO), and for instance the tree level expressions of Eq. 1.9 do not hold. In particular α must be evolved from $q^2 \simeq 0$, where it is precisely measured, to the $s = m_Z^2$ where most of the EWPO are computed. The running of α can be described as:

$$\alpha(s) = \frac{\alpha(0)}{1 - \Delta\alpha(s)}, \quad \text{where } \Delta\alpha(s) \simeq \Delta\alpha_\ell(s) + \Delta\alpha(s)^{(5)}_{\text{had}}(s) + \Delta\alpha_t(s), \tag{1.12}$$

and the 3 components of $\Delta\alpha_\ell(s)$ are the lepton, five-quark hadronic and top loop corrections, respectively. Since the experimental uncertainty is ruled by $\Delta\alpha(s)^{(5)}_{\text{had}}$, it is used as EWPO in the electroweak global fit instead of α. $\Delta\alpha(m_Z^2)^{(5)}_{\text{had}}$, despite the much lower precision compared to $\alpha(0)$ (about $3 \cdot 10^{-3}$ versus 10^{-10}) is not a fundamental parameter of the SM, and it used only as the current best description of the evolution of α at the Z mass.

The electroweak global fit encodes also 10 theoretical uncertainties in the likelihood. These uncertainties are treated as nuisance parameters of the fit, on which a Gaussian constraint is applied. More details are provided in Ref. [41].

The free parameters of the fit are m_H, m_Z, m_c, m_b, m_t, $\Delta\alpha(m_Z^2)^{(5)}_{\text{had}}$, α_s and the 10 theoretical nuisances. The full list of EWPO is estimated with individual profile likelihood ratio scans. The

χ^2 of the fit is 18.6 for 15 degrees of freedom, corresponding to a p-value of 0.23, demonstrating thus a strong agreement with the SM description. The results for the complete list of EWPO is shown in Fig. 1.3, while the numerical results for the free parameters are reported in table 1.2.

Table 1.2. *Input values, electroweak global fit result and indirect determination of the free parameters of the fit; the input value of α_S is omitted since is not used in the fit (from Ref. [40]).*

Parameter	Input value	Global fit result	Indirect determination result
m_H [GeV]	125.1 ± 0.2	125.1 ± 0.2	90^{+21}_{-18}
m_Z [GeV]	91.1875 ± 0.0021	91.1882 ± 0.0020	91.2013 ± 0.0095
m_c [GeV]	$1.27^{+0.07}_{-0.11}$	$1.27^{+0.07}_{-0.11}$	-
m_b [GeV]	$4.20^{+17}_{-0.07}$	$4.20^{+17}_{-0.07}$	-
m_t [GeV]	172.47 ± 0.68	172.83 ± 0.65	176.4 ± 2.1
$\Delta\alpha(m_Z^2)^{(5)}_{\rm had}$	$(2760 \pm 9) \cdot 10^{-5}$	$(2758 \pm 9) \cdot 10^{-5}$	$(2713 \pm 39) \cdot 10^{-5}$
α_s	-	0.1194 ± 0.0029	0.1194 ± 0.0029

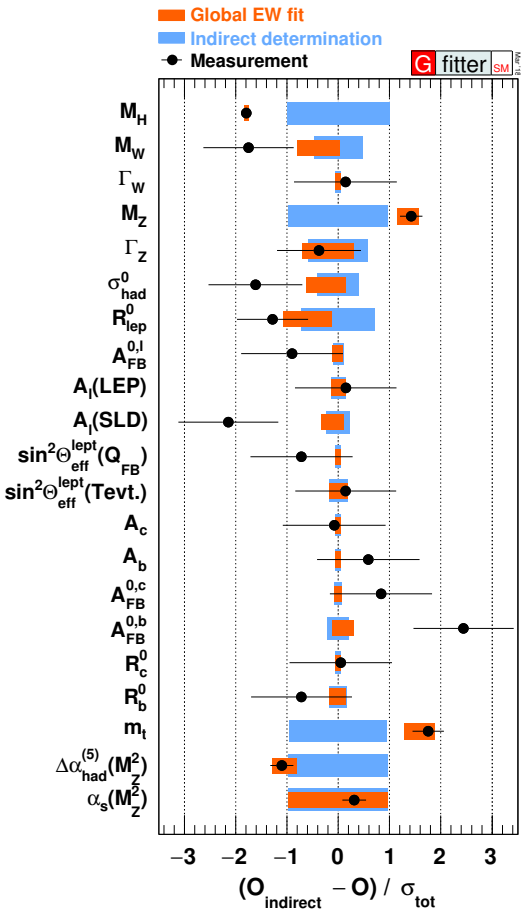


Figure 1.3. *Difference of the input measurement and the indirect determination (black dots); difference of the global fit result and the indirect determination (red bands); indirect determination centred at 0 (blue bands). Each value is rescaled by the uncertainty of the two components of the difference added in quadrature and it is shown for each EWPO included in the electroweak global fit and for the free parameters (from Ref. [40], where also the description of the variables can be found).*

1.2.1 W mass measurement motivation

The W boson mass m_W is one of the EWPO included in the electroweak global fit. The experimental input used in the fit is:

$$m_W^{\text{exp}} = 80.379 \pm 0.013 \text{ GeV}, \tag{1.13}$$

from the combined measurement of LEP experiments, Tevatron experiments and ATLAS [18]. The different measurement approaches and the sources of uncertainties will be described in detail in Chapter 4.

The tree level expression of m_W is reported in Eq. 1.9. Taking into account of the radiative corrections, m_W can be expressed as [18]:

$$m_W^2\left(1 - \frac{m_W^2}{m_Z^2}\right) = \frac{\pi\alpha}{\sqrt{2}G_F}(1 + \Delta r) \quad \text{i.e.} \quad m_W = \sqrt{\frac{\pi\alpha}{\sqrt{2}G_F}}\frac{1}{\sin\theta_W\sqrt{1-\Delta r}}, \tag{1.14}$$

where the the radiative corrections has been encoded in the parameter Δr. In the one-loop calculation it can be parametrized as:

$$\Delta r = \Delta\alpha - \Delta\rho\frac{1}{\tan^2\theta_W} + \Delta r_{\text{rem}}, \quad \text{where } \Delta\rho = \frac{3G_F m_t^2}{8\sqrt{2}\pi^2}. \tag{1.15}$$

The first term is due to the α running (from Eq. 1.12), the second one is the dominant term due to the top quark loops. The last term Δr_{rem} includes the remaining corrections, like the m_H-dependent logarithmic terms and the higher order corrections ($\mathcal{O}(\alpha\alpha_s)$, $\mathcal{O}(\alpha\alpha_s^2)\ldots$).

The electroweak global fit estimation of m_W, combining the full experimental and theoretical information, is $m_W^{\text{comb}} = 80.359 \pm 0.006$ GeV. However the indirect determination is much more relevant, since it represents the unbiased SM prediction of the mass:

$$m_W^{\text{pred}} = 80.354 \pm 0.007 \text{ GeV}. \tag{1.16}$$

The value is obtained with a profile likelihood ratio scan, and the uncertainty is estimated as $\Delta\chi^2 = 1$ from the minimum, as shown in Fig. 1.4. The uncertainty can be broken down in the different sources, from the fit input:

source	total	m_t	m_t^{theo}	m_Z	α_S	$\Delta\alpha(m_Z^2)^{(5)}_{\text{had}}$	m_H	m_W^{theo}
value [MeV]	7	2.7	3	2.6	2.6	2.4	1	4.

The largest component arises from the nuisance parameter related to the theoretical uncertainty on m_W. The latter emerges from not completely known higher order corrections with 3-loops: $\mathcal{O}(\alpha^2\alpha_s)$, (α^3), (α_s^3). Further details about the theoretical uncertainties can be found in Ref. [45].

The indirect determination m_W^{pred} shows a tension of about 1.5σ with the measured value. This must be interpreted as good compatibility between the global SM prediction and the observed value of the mass, since the discrepancy is too small to allow BSM interpretations. However, the fit results highlight the necessity to improve the precision of the experimental measurement of m_W, since m_W^{pred} is about a factor 2 more precise than m_W^{exp}.

Moreover, a more precise experimental input of m_W^{exp} also affects the global picture of several EWPO. For instance, the m_t uncertainty is ruled by the m_W knowledge. Because of this correlation, an important SM test is the simultaneous indirect determination of m_W and m_t i.e. removing both measurements from the likelihood of the global fit and then predict their values.

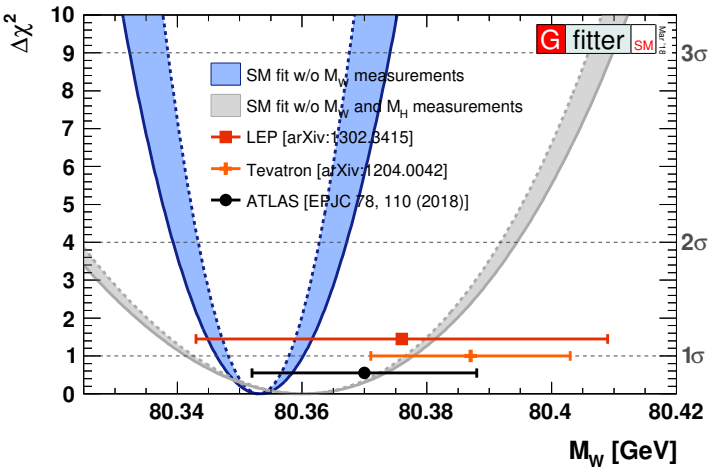


Figure 1.4. *Scan of the χ^2 of the as a function of m_W, in case m_W (blue band) or m_W and m_H (grey band) are removed from the likelihood for the indirect determination of m_W . The dotted lines represent the results excluding the theoretical uncertainties, in both cases. On the right axis the equivalent σ are also shown. The measurement from LEP, Tevatron and ATLAS are also superimposed to the χ^2. These results are taken from* `GFitter`*, updated with 2018 experimental and theory input [40].*

The result is reported in Fig. 1.5(a), and it shows a good agreement with the experimental values of m_W^{exp} and m_t^{exp}, with a discrepancy below 2σ.

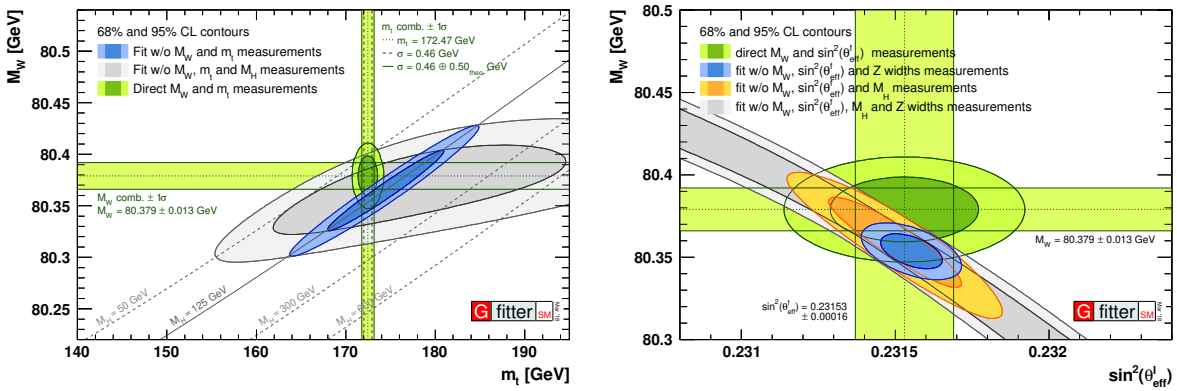
The sine of the effective leptonic weak mixing angle is another EWPO with a relevant correlation with m_W. It is defined as:

$$\sin^2\theta_{\text{eff}}^{\ell} \equiv \frac{1}{4Q_\ell}\left(1 - \text{Re}\frac{g_V^\ell}{g_A^\ell}\right) = (1+\Delta\kappa)(1-\frac{m_W^2}{m_Z^2}) = (1+\Delta\kappa)\sin^2\theta_W,$$

where the vector and axial couplings $g_{V,A}^\ell$ have been defined in Eq. 1.11 and $\Delta\kappa$ contains the corrections to Z-lepton vertex form factor. Like m_W, also $\sin^2\theta_{\text{eff}}^{\ell}$ is determined with higher precision in the electroweak fit compared to the direct measurement. From theoretical side, the most relevant uncertainties arise from m_t and $\Delta\alpha(m_Z^2)_{\text{had}}^{(5)}$. The indirect determination of $\sin^2\theta_{\text{eff}}^{\ell}$ simultaneously to m_W is reported in Fig. 1.5(b).

Finally, the precision of the electroweak fit predictions allows to efficiently exploit the fit to constrain BSM theories, as previously mentioned. Also in the new physics searches a more precise measurement of m_W and the correlated EWPOs can improve the exclusion limits of BSM models. For instance, a discussion about the electroweak fit constraint to two-Higgs-doublet models (2HDM) can be found in Ref. [40], while a report about the future electroweak measurements impact can be found in Ref. [41].

By way of example, between the large number of SM possible extensions, is interesting to report a brief discussion about the *oblique parameters* [46]. The effect of electroweak precision observables on the BSM is often described in the framework of the oblique parameters S, T, U. In this parametrization, the contribution of new physics is considered in electroweak radiative correction only i.e. correction to the bosons self-energies. Direct vertex or box corrections to the SM Lagrangian are not included, therefore in this class of models, the new physics is only weekly coupled with SM fermions. The gauge group is still $SU(2)_L \times U(1)_Y$ and no additional electroweak gauge boson are considered. Moreover, the radiative corrections are considered flavour universal (thus an additional parameter is required to describe $Z \to b\bar{b}$ vertex, receiving large top corrections). Finally, new physics is expected to have an energy scale much larger than



(a) (b)

Figure 1.5. *Contours at 68% and 95% CL from χ^2 scans as a function of m_W and m_t (a) or m_W and $\sin^2\theta_{eff}$ (b), excluding them from the likelihood. m_H can be included in the fit (blue ellipse) or excluded (grey and yellow ellipse) . In the latter case are highlighted some values of the m_H scan. The experimental measurement of m_W and m_t are shown as green bands. These results are taken from* **GFitter**, *updated with 2018 experimental and theory input [40].*

m_Z. With these assumptions, the new physics will be encoded in 6 corrections to the γ, Z, W self couplings. The S, T and U parameters and defined subtracting the SM component to the electroweak boson radiative correction making the BSM parameters $\mathcal{O}(1)$, and the remaining 3 corrections are reabsorbed in the renormalization of α, G_F and m_Z. The SM prediction is $S = T = U = 0$. This can be done by choosing a reference point of the SM, in term of m_H and m_t (which represent the leading contributions to the SM radiative corrections). The effect of non-vanishing oblique parameters can be summarized as:

- T is related to the difference between W and Z self-energies at low energies ($q^2 \simeq 0$), implying a vector $SU(2)$-breaking. It is constrained by Γ_Z measurement mainly.
- S ($S + U$) is related to the difference between the Z (W) self-energy at high energy $q^2 = m_{Z,W}$ and low energy $q^2 \simeq 0$, implying a axial $SU(2)$-breaking. S is constrained by m_Z measurement.
- U produces an additional correction to Δr, from Eq. 1.15, thus is strongly related to m_W and Γ_W.

The W boson mass can be rewritten, taking into account of the oblique correction, as:

$$m_W^2 = \frac{m_{W_{\rm SM}}^2}{1 - G_F m_{W_{\rm SM}}^2 (S+U)/2\sqrt{2}\pi}.$$

A deviation of these parameters from zero can be assessed by several new physics models, considering multiplets of heavy fermions, Majorana particles, warped extra dimensions, composite Higgs sectors, etc. An exhaustive review of the current model under investigation can be found in Ref. [18], while the accurate definition of the oblique parameters can be found in Ref. [42].

The oblique parameters are fully compatible with the SM expectations combining the full information of the electroweak fit [40] :

$$S = 0.001 \pm 0.11, \quad T = 0.09 \pm 0.14 \quad U = -0.02 \pm 0.11.$$

However large correlations are present (+0.92 between S and T, −0.68 between S and U, −0.87 between T and U), and since U is expected to be small in several BSM scenarios, it can be set to zero. The result of this choice is reported in Fig. 1.6, and also in this case the parameters are consistent with the SM hypothesis and they are strongly correlated. If only one of the parameters is left free in the fit, fixing the other two at 0, the parameter shows a discrepancy between 1.6σ and 1.9σ with the SM, reflecting the observed deviation of m_W.

In conclusion, the electroweak sector will strongly benefit from a more precise measurement of m_W. It can be used to test the consistency of the SM, improve the knowledge of the other correlated parameters or explore and constrain new physics scenarios.

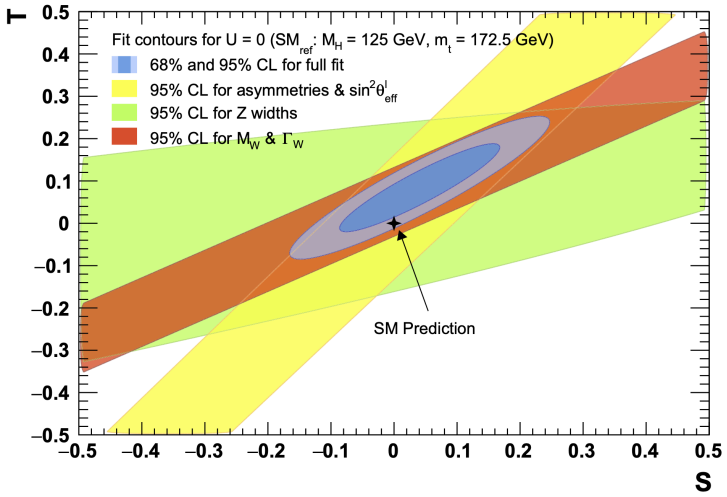


Figure 1.6. *Electroweak fit prediction of the oblique parameter S and T, fixing $U = 0$ with the full fit (blue ellipse) or removing specific observables from the likelihood (yellow, green, orange bands). These results are taken from* `GFitter`*, updated with 2018 experimental and theory input [40].*

Chapter 2

The Compact Muon Solenoid experiment at LHC

The Compact Muon Solenoid (CMS) experiment is a multi-purpose detector for the Large Hadron Collider (LHC) at CERN. In this chapter, both the detector and the accelerator are described. An overview of the entire system is presented with more focus on the features relevant for the analysis discussed in this thesis.

2.1 The Large Hardon Collider

The LHC [2] is a circular collider, placed at CERN (Geneva), hosted in the 26.7 km long LEP tunnel. It is designed to accelerate and collide protons or heavy nuclei (Lead and Xenon) with a design center-of-mass energy per nucleon of $\sqrt{s} = 14$ TeV in proton-proton (pp) collisions and $\sqrt{s} = 5.52$ TeV for Pb-Pb and Xe-Xe. The LHC construction ended in 2008 and the first physics run, the *Run 1*, took place between 2010 and 2013. After the *Long Shutdown 1*, an upgrade period of the machine, the *Run 2* took place between 2015 and 2018, followed by the *Long Shutdown 2* for further upgrades. The *Run 3* operations of LHC will start in 2021. During the Run 1, the beam energy per nucleon was 3.5 TeV (4 TeV in 2012) and 6.5 TeV during the Run 2, while a beam energy per nucleon of 7.0 TeV is expected for the Run 3.

The LHC is the final acceleration ring while the rest of the CERN accelerator complex acts as injector. The protons are first injected in the *Linear Accelerator 2* (Linac2), then passed to the *Proton Synchrotron Booster* (PS booster), then to the PS ring and finally to the *Super Proton Synchrotron* (SPS) before the injection to the LHC [47]. All these steps are required to gradually ramp up the beam energy and ensure the stability of the beam. In addition, they create the required beam structures, described later on. The LHC is composed of two rings, with four interaction points (IPs) where the experiments are placed in the respective caverns: ATLAS [48], ALICE [49], CMS and LHCb [50]. Two additional smaller experiments are placed on the ring, LHCf [51] and TOTEM [52], which share the cavern with ATLAS and CMS, respectively. At the interaction point regions the LHC beams are flanked, with straight sections of 528 m for experimental and utility purposes. A diagram of the CERN accelerator complex is shown in Fig. 2.1.

The acceleration is realized with 8 radiofrequency cavities per beams, which operate at 400 MHz delivering 2 MV each and an accelerating field of 5 MV/m. The cavities operate at a temperature

of 4.5 K in a superconductive state. The focusing and the bending of the proton bunches are realized with a set of superconductive dipoles and quadrupoles magnets, which operates at 1.9 K with superfluid helium cooling. They provide a magnetic field of about 8 T. The two beampipes are located inside the superconductive magnets and the ultrahigh vacuum is achieved with a cryopump system. Inside the beampipe the pressure is about 10^{-7} Pa (along the tunnel) or 10^{-9} Pa (at IPs), and the operation temperature is 5-20 K. The heating of the beampipe occurs because of several effects, like synchrotron radiation and dissipation due to image currents.

The LHC beams are bunched with a spacing of 25 ns (equivalent to 7.5 m, travelling at the speed of light), with $1.15 \cdot 10^{11}$ protons per bunch. The LHC is filled with up to 2556 bunches per beam, and the bunches have a cross section of 16×16 μm^2 at IPs and a longitudinal dimension of about 7.5 cm. These parameters are optimized constantly and change between the runs. A summary of the LHC parameters across years is reported in Table 2.1, from Ref. [53].

The bunch parameters are the key ingredients for the *luminosity* of a collider defined as $\mathcal{L} = \frac{1}{\sigma}\frac{dN}{dt}$, where $\frac{dN}{dt}$ is the event rate of a given process with cross section σ. For a circular collider with Gaussian transverse cross section bunch it can be written as:

$$\mathcal{L} = \frac{f N_1 N_2}{4\pi\sigma_x\sigma_y}k = \frac{f N_1 N_2 \gamma}{4\pi\beta^*\varepsilon}k,$$

where the latter formula is expressed as a function of betatron function at IP β^* and normalized emittance ε. N_1 and N_2 are number of protons per bunch, $\gamma = E_p/m_p$, $f = 40$ MHz is the bunch crossing rate, σ_x and σ_y are the transverse dimensions of the bunch, and $k \lesssim 1$ is a reduction factor due to the crossing angle between beams at IP. The design peak luminosity for LHC is $\mathcal{L} = 10^{34}$ cm^{-2}s^{-1}, but in 2017 $\mathcal{L}$ reached up to 2×10^{34} cm^{-2}s^{-1} due to the tuning of the LHC beam parameters. The *integrated luminosity* is defined as $\mathcal{L}_{\text{int}} = \int \mathcal{L}dt = N/\sigma$. Fig. 2.2 shows the peak luminosity and the integrated luminosity of LHC delivered to CMS experiments during the Run 1 and Run 2. The integrated luminosity is about 30 fb^{-1} and 163 fb^{-1} for Run 1 and Run 2, respectively.

More details about LHC can be found in Ref. [2].

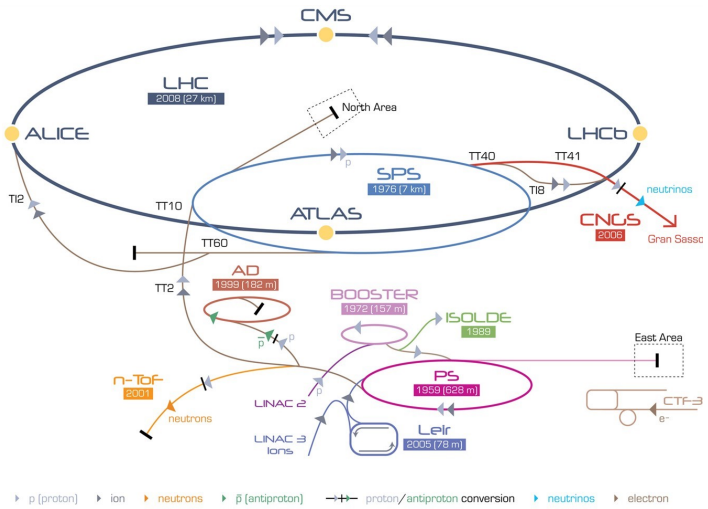


Figure 2.1. *The CERN accelerator complex, the grey arrows follow the path for pp collisions (from [47]).*

Table 2.1. *LHC parameters per year (from [53]).*

	desing	2012	2016	2017	2018
$\mathcal{L}_{\text{peak}}$ [10^{34} cm^{-2}s^{-1}]	1	0.75	1.4	2.05	2.01
Beam energy [TeV]	7	4	6.5	6.5	6.5
Bunch spacing [ns]	25	50	25	25	25
Proton per bunch [10^{11}]	1.15	1.65	1.1	1.15	1.15
Number of bunches per beam	2808	1374	2220	2556	2556
β^* CMS/ATLAS [cm]	55	60	40	40/30	30-25
Crossing angle [mrad]	285	290	370-280	300-240	320-260
ε [mm]	3.75	2.5	2.2	2.2	2
Average pile-up $\langle\mu\rangle$	~ 20	~ 21	~ 27	~ 38	~ 37

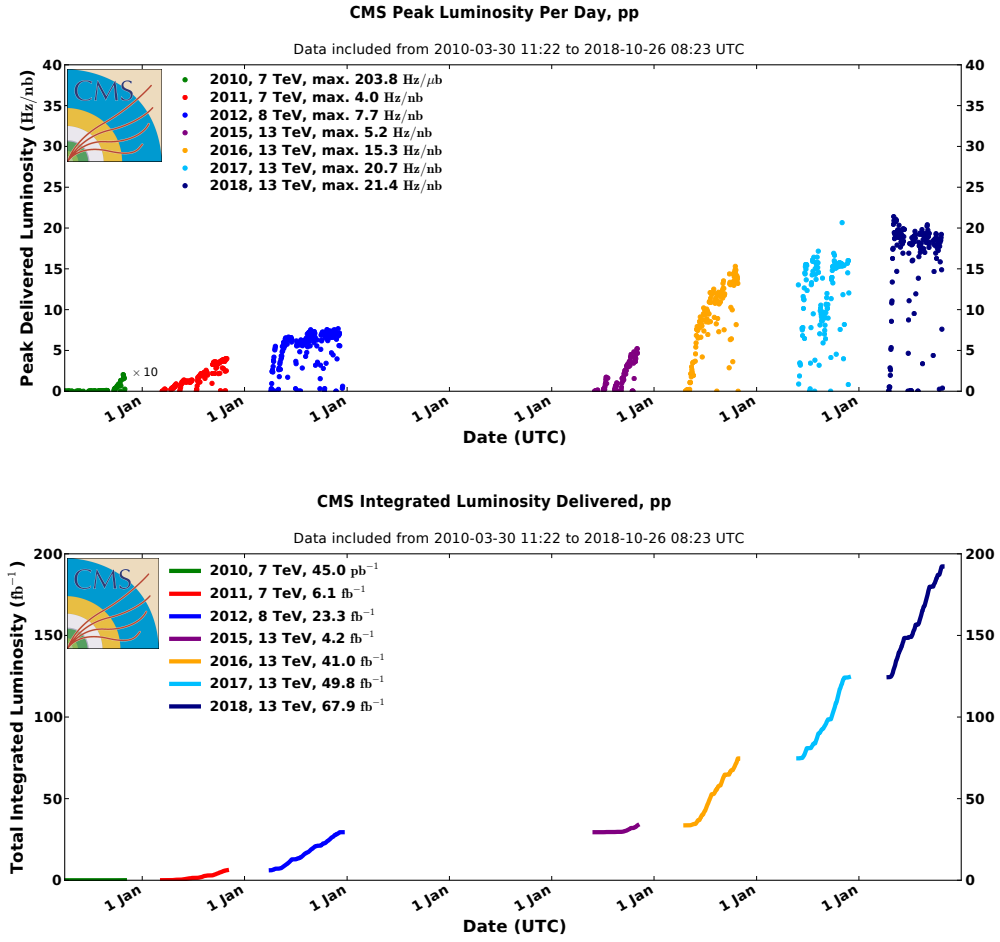


Figure 2.2. *LHC luminosity (upper plot) and integrated luminosity (lower plot) delivered to CMS experiment between 2010 and 2018 in pp collisions (from [54]).*

2.1.1 LHC *pp* event properties

The *pp* interaction can be described in the context of the parton model at LHC center-of-mass energy. In this framework, the hadrons are loosely-bounded states of the constituents, i.e. the quarks and the gluons, called *partons* in general. At large transferred momentum

($|q^2| \gtrsim$ 10 GeV^2, where $q^2 = (p_1 - p_2)^2$ and p_1, p_2 are the 4-momenta of the two interacting partons) the process is called *hard scattering*. It can be treated as an interaction of an isolated pair of partons, with center-of-mass energy $\sqrt{\hat{s}} = \sqrt{sx_1x_2}$, where x_1 and x_2 are the fractions of the proton momentum carried by the interacting partons, described by the *parton distribution functions* (PDFs). The total hard scattering cross section is the sum of all possible interactions between partons, while the center-of-mass of the initial state is unknown and changes on event by event basis:

$$\sigma_{pp} = \int \sum_{i,j} f_i(x_1, q^2) f_j(x_2, q^2) \hat{\sigma}_{i,j}(x_1 p_1, x_2 p_2, \mu_F, \mu_R) dx_1 dx_2, \tag{2.1}$$

where $f_i(x)$ is the PDF of parton type i at momentum xp and $\hat{\sigma}_{ij}$ represent the cross section between partons type i, j, μ_R and μ_F are the renormalization and factorization scales respectively. The latter represent the arbitrary scales that separate the perturbative and non-perturbative dynamics. They arise from fixed-order calculation, but the cross section must not depend on these parameters at all orders. The PDFs, as a function of x and $Q^2 = -q^2$, and the cross sections for several processes are reported in Fig. 2.3.

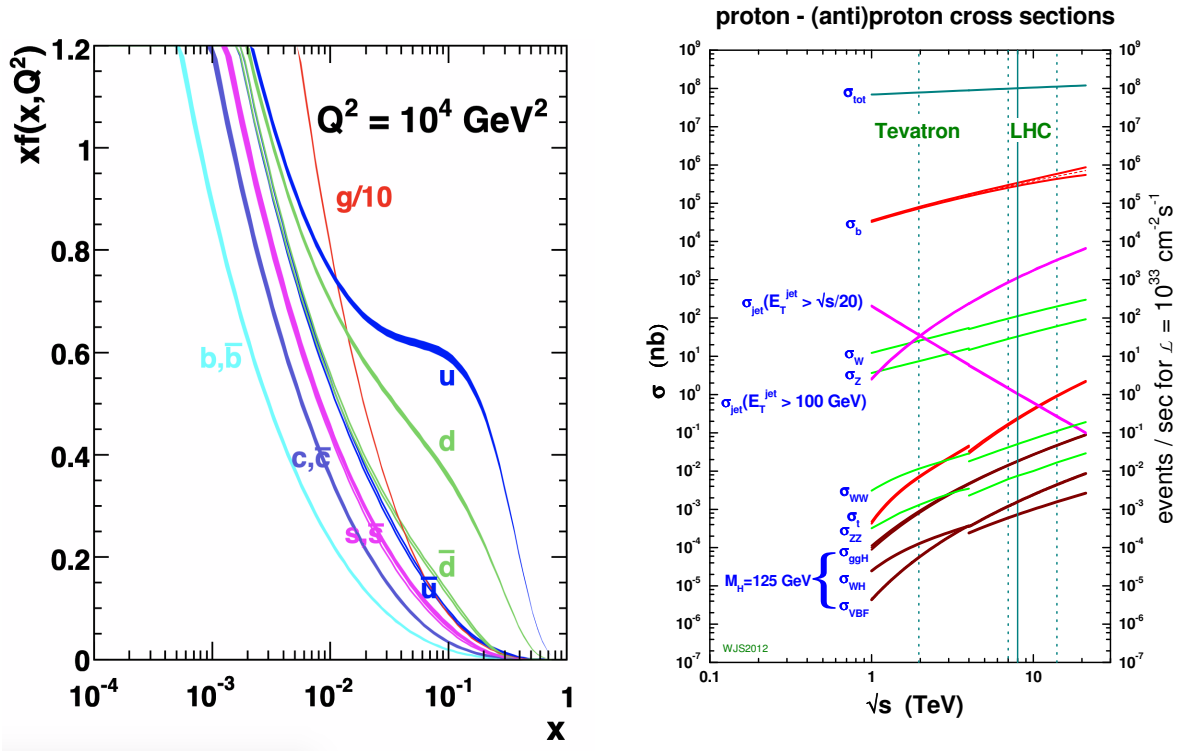
Along with the hard scattering process, there is the interaction between the proton remnant, called the *underlying event*, which involves low q^2 interaction between partons, initial and final state radiation (ISR and FSR, respectively). These interactions are typically strong processes between quark and gluons, which hadronize to have colourless final states.

The hard scattering is quite rare and most of the interactions are *soft scattering*, with low q^2. Therefore it is quite likely that a hard scattering process happens together with additional soft scattering processes, between different protons of the same bunch. A *primary vertex* can be associated with each scattering process, ideally representing the position of the interaction of the primary pair of particles. The soft scattering can be classified as *elastic*, *single-diffractive* or *double-diffractive* if both, one or none of the involved protons are preserved in the final state, respectively. These additional interactions called Pile-Up (PU) produce an additional background for the hard scattering event, which must be properly identified and removed. For instance see Fig. 2.4(a), where a high-PU event is shown. The mean PU $\langle\mu\rangle$ is defined as the mean number of inelastic interactions per event. An increase of the luminosity leads to the increase of the PU rate, as shown in Fig. 2.4(b), therefore a trade-off exists in the experiment design between the high luminosity (i.e. the higher rate of interesting events) and the ability to distinguish between soft scatting and hard scattering processes.

An additional source of background in LHC events is the *Machine Induced Background* (MIB), subdivided into two main sources [55, 56]:

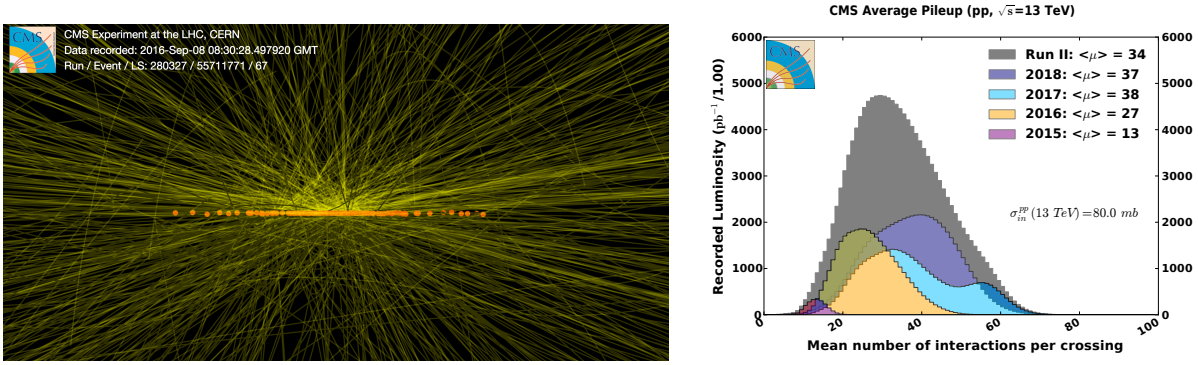
- **Beam gas** particles produced in the interaction of the beam protons with residual gas molecules in the beampipe, produced close to the IP (local beam gas) or produced far from the IPs but scattered with small angles (global beam gas). This is the main source of MIB particles.
- **Beam halo:** particles produced in hadron and electromagnetic showers from the interaction of beam protons with the collimators. The beam halo is also composed of particles produced by Touschek effect (intra-beam Coulomb scattering), synchrotron radiation and long-range interactions [57]. These sources are subleading in terms of background particles.

The MIB is the main source of beam losses of LHC together with the effects of the pp interactions at IPs. The combination of these effects produce progressive degradation of luminosity during the run and therefore the beam is refilled each 10-20 hours.



(a) (b)

Figure 2.3. *(a): NLO PDF,expressed as $xf(x,q^2)$ for various parton flavours as a function of x at $q = 100$ GeV (from Ref. [58]). (b): NLO and NNLO cross section of SM processes using MSTW2008 scheme, as a function of $\sqrt{s}$, for $p\bar{p}$ and pp collisions; the total hadronic cross section only is based on a parametrization of the PDG; the discontinuity at 4 TeV is due to the switch from $p\bar{p}$ to the pp collision at that energy; the width represents the uncertainty in the extrapolation (from Ref. [59]).*



(a) (b)

Figure 2.4. *(a): Longitudinal view of the IP of CMS during an event recorded in 2016; the orange dots and the green lines are the reconstructed primary vertices and the reconstructed trajectories of charged particles (tracks), respectively (from Ref. [60]). (b): PU distribution along years during Run 1 and Run 2; the distributions and $\langle\mu\rangle$ are defined counting the inelastic interactions only (from Ref. [54]).*

2.2 The Compact Muon Solenoid detector

The Compact Muon Solenoid (CMS) [1, 61] is the detector located at the fifth access point of LHC (*Point 5*), in Cessy (France). CMS is a general purpose detector, optimized to efficiently collect and reconstruct the *pp* events at LHC. In particular, the deliveries from the collider are: thousands of neutral and charged particles per event in the energy range between a few MeV to hundreds of GeV in a wide angular range; an event by event separation of 25 ns with the presence of tens of PU interaction in the proximity of the hard scattering vertices; the presence of very dense environment in collimated regions (with tens of tracks in 0.2 sr) due to hadronic jets; a radiation dose up to $\sim 10^5$ Gy or $5 \cdot 10^{13}$ $\mathrm{HEH_{eq}/cm^{-2}}$ (Equivalent High-Energy Hadrons fluence) per year [62].

CMS is approximately a cylinder with a radius 7 m, length 21 m and a total weight of about 14000 tons. The central feature is a superconducting solenoid, providing the bending magnetic field inside the solenoid. CMS is equipped with a tracking system that allows for efficient charged particle reconstruction and a calorimeter system for neutral particle identification and energy measurements. A muon detector system, composed of gas chambers embedded in the steel yoke of the solenoid, is placed outside of the superconducting coil. All the subdetectors are realized with different geometries for the central region (*barrel*) and for the forward/backward regions (*endcaps*), to extend the angular acceptance of the experiment and maximize the hermeticity of the detector, taking into account the LHC geometrical constraint. CMS integrates a high efficiency and low bias hardware and software trigger to cope with the *pp* soft scattering rate and the LHC backgrounds. The trigger system together with the excellent time resolution of the detectors allows fulfilling the bunch-crossing time constraint. CMS is designed to be also able to reconstruct Pb-Pb and Xe-Xe events during the heavy ion runs of LHC.

The structure of the CMS detector is shown in Fig. 2.5, and the main features of the subsystem are listed in table 2.2. A detailed description of CMS detector can be found in Ref. [1, 61].

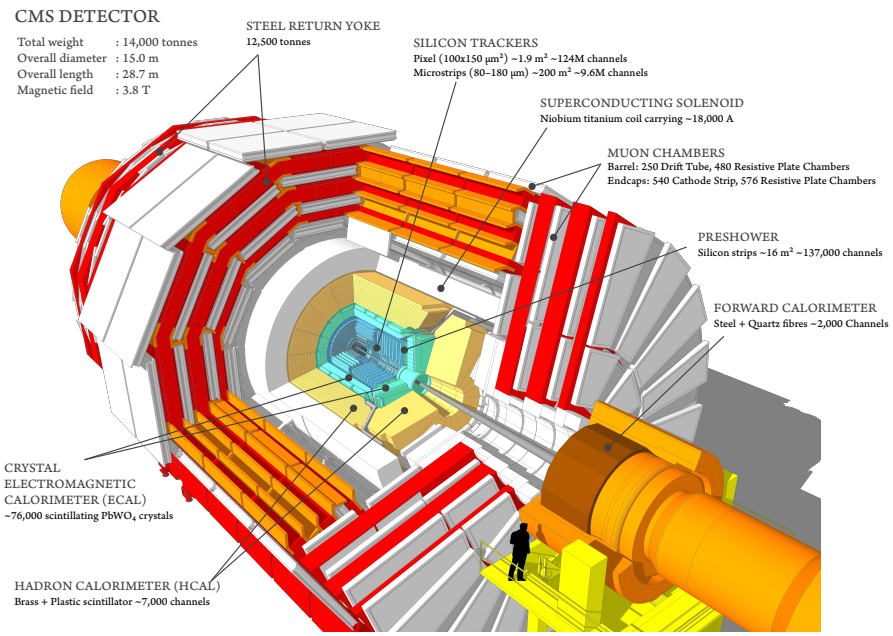


Figure 2.5. *Cross section of the CMS detector in a 3D rendering, the subdetectors and their main features are highlighted (from [63]).*

Table 2.2. *Summary of the detector components of CMS, in 2016 configuration.*

System	Subsystem	Technology	Configuration	$N_{channels}$	$\|\eta\|$ acceptance
Pixel	BPIX	n^+-in-n silicon pixel	pixel size: 100×150 μm^2; 3 layers at r =44, 73, 102 mm	48M	$[0, 1.5]$
	FPIX	n^+-in-n silicon pixel	pixel size: 100×150 μm^2; 2 disks at $\|z\|$ =345, 465 mm	18M	$[1.5, 2.5]$
Tracker	TIB	p-in-n silicon strip	strip pitch: 80 μm (L1,L2), 120 μm (L3,L4); 4 layers at r =255,339, 418.5, 498 mm	9.3M	$[0, 1.5]$
	TID	p-in-n silicon strip	strip pitch: 100-141 μm; 3 disk at $\|z\|$ =800-900 mm		$[1.5, 2.5]$
	TOB	p-in-n silicon strip	strip pitch: 183 μm (L1-L4), 122 μm (L5,L6); 6 layers at r =608, 692, 780, 868, 965, 1080 mm		$[0, 1.5]$
	TEC	p-in-n silicon strip	strip pitch: 97-184 μm; 9 disk at $\|z\|$ =1240-2800 mm		$[1.5, 2.5]$
ECAL	EB	homogeneous calo. ($PbWO_4$ crystals)	$25X_0$ and $1\lambda_I$	61.2k	$[0, 1.48]$
	EE	homogeneous calo. ($PbWO_4$ crystals)	$26X_0$ and $1\lambda_I$	$2 \cdot 7234$	$[1.48, 3]$
	PS	sampling calo. (lead-silicon stirp)	2 strip layers alternated with 2 absorber layers; $3X_0$	137.7k	$[1.6, 2.6]$
HCAL	HB	sampling calo. (brass-scintillator)	16 alternated layers of scintillator and absorber; $6\lambda_I$	9k	$[0, 1.3]$
	HE	sampling calo. (brass-scintillator)	16 alternated layers of scintillator and absorber; $6\lambda_I$		$[1.3, 3]$
	HF	Cherenkov calo. (steel-quartz fibers)	grooved absorber interlayered with quartz fibers		$[3, 5]$
	HO	sampling calo. (scintillator)	2 layers (central ring), 1 layer ($\pm 1, \pm 2$ rings), using the coil as absorber		$[0, 1.3]$
Muon system	MB	DT	section: 4.2×1.3; 4 layers	195k	$[0, 1.2]$
	ME	CSC	strip pitch: 8-16 mm; 4 layers	500k	$[0.9, 2.4]$
	RB	RPC	avalanche mode; pitch: 1 cm; 6 layers	192k	$[0, 1.2]$
	RE	RPC	avalanche mode; pitch: 1 cm; 4 layers		$[0.9, 1.6]$

2.2.1 Coordinate systems

Different standard coordinate systems are used in CMS, depending on the specific convenience.

– The Cartesian reference frame has the origin at the nominal IP, the x-axis points to the center of LHC ring, the y-axis points upward orthogonally to LHC plane and the z-axis

along the beam line, to realize a right-handed frame.

– The polar reference frame has the origin at the nominal IP, with $r = \sqrt{x^2 + y^2}$ pointing outwards, $\tan(\theta) = r/z$ (i.e. measured from positive z axis, with $\theta = \pi/2$ for any vector orthogonal to the z axis), $\tan(\phi) = y/x$ (i.e. measured from positive x axis, with $\phi = \pi/2$ in the positive y direction).

– The physics-wise reference frame uses the previously described r, ϕ, but it uses the *pseudorapidity*

$$\eta = -\ln\big[\tan(\theta/2)\big]$$

instead of the θ angle. The former is convenient because of its relation with the *rapidity* of a particle with respect to the beam axis:

$$Y = \frac{1}{2}\ln\Big[\frac{E + p_z}{E - p_z}\Big],$$

where E is the energy and p_z is the component of the momentum along the beam axis. In the ultra-relativistic case where $|\mathbf{p}| \gg m$, $\eta \simeq Y$ holds. The variable Y is additive for boosts along the beam axis, thus the $\Delta Y \simeq \Delta\eta$ can be measured in the laboratory frame ignoring p_z, usually not known with sufficient precision in a hadron collider environment. Moreover, the production of soft QCD hadrons is almost flat in Y at LHC. In this reference frame it is convenient to express the momentum in the basis $\mathbf{p} = (p_T, \eta, \phi)$ and derive $|\mathbf{p}| = p_T \cosh(\eta)$, where the *transverse momentum* p_T is defined as the projection of the momentum of the particle on the plane orthogonal to the beam axis (the transverse plane), $\mathbf{p}_T \equiv (p_x, p_y, 0)$. The p_T of a particle is also invariant under boost along z.

2.2.2 The silicon tracker

The silicon tracker is the innermost subdetector of CMS [64]. It is composed of several subsystems, as shown in Fig. 2.6. The entire tracker is a cylinder, 6.5 m long, with a radius of 1.2 m. This subdetector is designed to provide an efficient and precise measurement of charged particle trajectories (tracks) and of the position of primary and secondary interaction vertices. The entire system works in the 3.8 T magnetic field provided by the superconducting solenoid, which guarantees the bending of the charged particles. The trajectory of a particle is extracted connecting the *hits* i.e. the processed electrical signal correspondent to a punctual interaction of a charged particle with the silicon sensor. The layered structure of the silicon tracker allows determining the 3D path of the particles, while the bending field allows measuring the momentum and the charge of the particles.

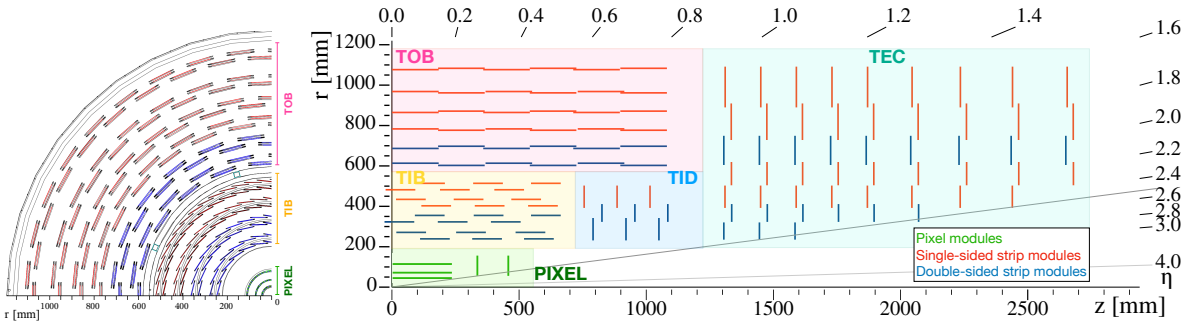
The environment of LHC collisions demands challenging requirement to fulfil these tasks. The material budget of the tracker should be as low as possible to minimize the distortions of the trajectory of the particles due to multiple scattering and to avoid biases in the energy measurement of the calorimeter caused by the interactions of electrons and photons. The tracker modules must guarantee radiation hardness able to sustain years of operation in LHC environment conditions. The high charged particles multiplicity of LHC events imposes a high granularity to identify individual tracks and vertices. The number of channels and the consequent readout electronics timing performance must be tuned to sustain the 40 MHz bunch crossing frequency to collect and store data efficiently. These requirements impose a trade-off between several design and technology aspects and CMS choose different solutions for the various subsystems, optimized to guarantee the optimal physics performances.

The pixel tracker is the innermost subsystem. It is composed of 3 layers in the barrel region (BPIX) and 2 disks for each endcap (FPIX), which extend a 3-hits coverage up to $\eta = 2.5$. The radii of the barrel layers are 44, 73, 102 mm, the distances of the disks from the IP are 345, 465 mm. The pixels size is 100×150 μm^2 in both regions. The active silicon area of the entire system about 1 m^2, with about 66 million pixels. The n$^+$-in-n pixel concept [65, 66] guarantees the radiation hardness of the detector. In this design, the pixel consists of n$^+$ implants in a high resistance n-substrate, while the pn-junction is located on the backside of the sensor.

The strip tracker is composed of 4 subsystems. The Tracker Inner Barrel (TIB) and the Tracker Inner Disks (TID) cover the region $r < 55$ cm and $|z| < 118$ cm, with four strip detector layers and tree strip detector disks, respectively. The Tracker Outer Barrel (TOB) and the Tracker EndCaps (TEC) cover the region 55 cm $< r <$ 120 cm, $z < 118$ cm and $r < 120$ cm, 118 cm $< z <$ 250 cm, with 6 layers and 9 disks, respectively. The strip pitch varies from 80 μm in the inner layers to 180 μm in the outer layers. On the two innermost layers of the TIB and TOB and in the 1,2 (1,2,5) modules of the TOB (TEC) an additional strip sensor is placed back-to-back to the main one, with a *stereo* angle of 0.1 rad, to allow 3D hit reconstruction. The sensors of the strip tracker are of p-in-n type [67, 68]. Globally the strip tracker has an active silicon area of 198 m^2 and 9.3 million strips.

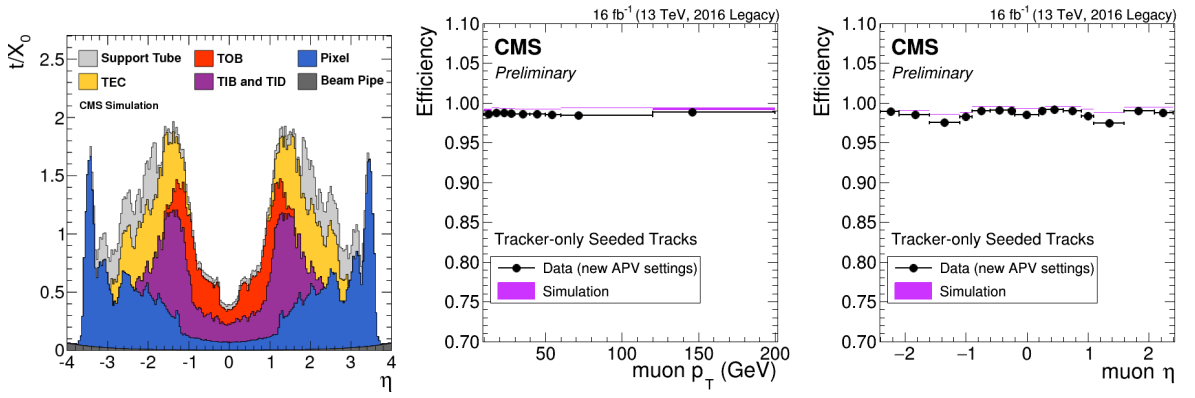
The high granularity of the tracker keeps the occupancy, defined as the number of simultaneously activated channels over the total number of channels of a given detector, at 10^{-5}-10^{-4} level in the pixel and at 10^{-3} level in the strip layers. The pixel detector provides a single hit resolution of about 10 μm in $r - \phi$ and 20-40 μm in z direction. The material budget of the silicon tracker, as shown in Fig. 2.7(a), is estimated from CMS Monte Carlo (MC) simulation with an accuracy better than 10% [69]. These features allow obtaining excellent track reconstruction performance in a wide kinematic range.

Concerning the muon reconstruction performances, the measurements of the efficiency on the data collected by CMS in 2016 is reported in Fig. 2.7(b) and 2.7(c), for tracker-only reconstruction. The muon reconstruction is one of the key ingredients of the W boson mass measurement, and the silicon tracker provides an efficiency close to 100% also in high PU scenario in a wide η and p_T range (from 7-800 MeV to 200 GeV).



(a) (b)

Figure 2.6. *Scheme of the silicon tracker in the r-ϕ view (a) and in the r-z view (b). In both cases only a quadrant of the detector is shown. The various subsystem, and the three classes of modules are highlighted in different colors. In the figure (a) the TID, TEC and FPIX modules are not shown.*



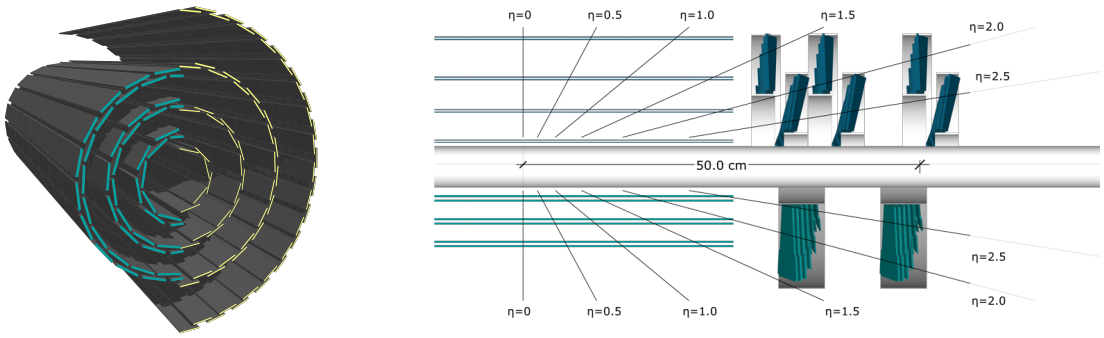
(a) (b) (c)

Figure 2.7. *(a): Material budget of CMS silicon tracker expressed in radiation length as a function of η (from [70]). Tracking efficiency as a function of p_T of the muon (b) or of η of the muon (c) for the tracker-only reconstruction, compared on data collected in 2016 and the CMS simulation (from Ref. [71]).*

2.2.2.1 Pixel detector *Phase 1* upgrade

The current pixel detector installed in CMS is different compared to the described one. The performance of the latter degraded due to the radiation damage after the Run 1 and first Run 2 year dose. Therefore it was replaced with an upgraded version in 2016 Technical Stop, usually referred to as *Phase 1* pixel detector. The latter has been used in the 2017 and 2018 data taking period. The main new feature compared to the pixel detector described in the previous section is the presence of a fourth layer in the barrel and a third disk in the endcap. Therefore the current radii of the barrel layers are 29, 68, 109, 160 mm, while the distances of the disks from the IP are 291, 396, 516 mm. The new configuration ensures better performance in the track-seeding, due to the additional redundancy in the pixel hit information. The Phase 1 pixel detector provides indeed 4 tracking hit coverage up to $|\eta| = 2.5$.

This new detector has not been used in the data or simulation reconstruction discussed in this thesis but has been used for the tracking studies of Appendix A. In Fig. 2.8 the new layer scheme is shown compared to the previous one. More details of this upgrade are provided in Ref. [72].



(a) (b)

Figure 2.8. *The compared layer configuration between the original and the Phase 1 pixel detector in a 3D rendering or the barrel (a) and in the r-z plane for both the barrel and the endcap (b) (from Ref. [72]).*

2.2.3 The electromagnetic calorimeter

The Electromagnetic Calorimeter (ECAL) is the innermost of the calorimeter systems of CMS [73]. The role of the calorimeter is the measurement of the energy of the particle in a destructive way i.e. stopping and absorbing the particles to collect the energy released in the detector inclusively. ECAL is optimized to maximize the probability of electromagnetic interaction with the detector material to efficiently measure the energy of photons and electrons. ECAL also has a primary role in the hardware trigger of CMS, due to its fast response and readout.

ECAL is a homogeneous calorimeter composed of $PbWO_4$ crystals. It is subdivided into a barrel section, which covers up to $|\eta| < 1.48$ and contains 61200 crystals, and two endcap sections which extend the acceptance to $|\eta| < 3$ and include additional 7234 crystals. In addition, a sampling calorimeter of lead and silicon strip detector, called Preshower, is placed in front of each endcap section to improve the detection performance in the $1.6 < |\eta| < 2.6$ range. The scheme of the ECAL structure is shown in Fig. 2.9(a).

The $PbWO_4$ crystals are fully active material, providing 26 (25) radiation lengths X_0 in barrel (endcap) region and generating the scintillation light upon irradiation. The $PbWO_4$ is an extremely dense material with a radiation length of 0.89 cm, resulting in 23 (22) cm long crystals for the barrel (endcap). This is a crucial feature to allow ECAL to be placed inside the volume surrounded by the superconductive solenoid. The crystals transverse size matches the Molière radius of 2.2 cm, providing a granularity 0.0174×0.0174 in the $\eta \times \phi$ plane. This granularity guarantees the capability to distinguish single particle deposits as close as 5 cm. These features guarantee that on average the 98% of 1 TeV electromagnetic shower is included in ECAL crystals. Moreover, the ECAL material provides about 1 nuclear interaction length (λ_I), which make the hadrons start showering in the ECAL materials.

The scintillation light is collected by silicon avalanche photodiodes in the barrel, mounted in pairs at the end of each crystal. Vacuum phototriodes are used instead in the endcap, glued to the back of each crystal [1, 74].

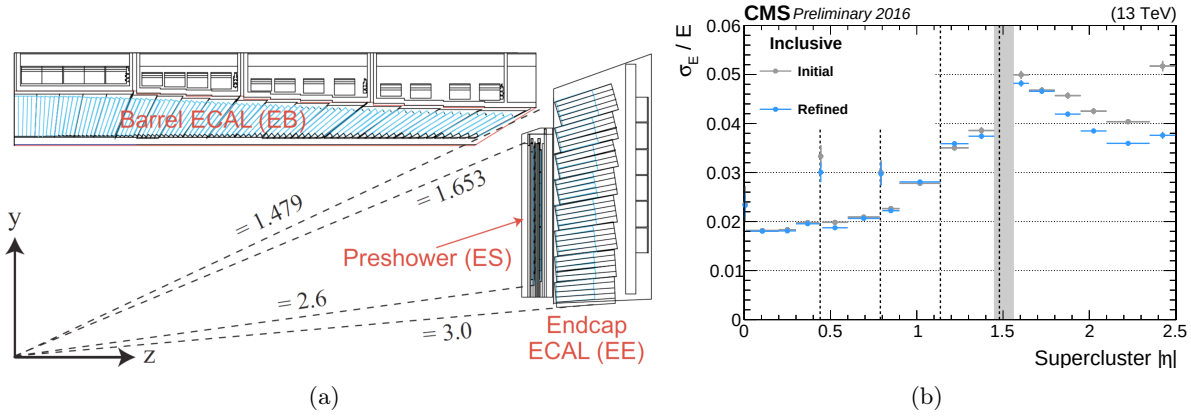
The energy resolution of ECAL can be parametrized as:

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E/\text{GeV}}} \oplus \frac{12\%}{E/\text{GeV}} \oplus 0.3\%,$$

where the $\sim 1/E$ term is due to electronic noise, independent from the energy, the $\sim 1/\sqrt{E}$ term is the stochastic term due to the intrinsic fluctuation of the shower (a Poissonian process) and the constant term is due to the energy scale calibration. An example of the energy resolution performance is given in Fig. 2.9(b).

The ECAL response and thus the performance strongly depends on the transparency of the crystals, which suffers from radiation damage. However, the degradation due to radiation is not permanent and the crystals recover transparency during the shutdowns of the machine. The response is therefore dynamically corrected with frequent monitoring of the transparency condition. The main monitoring tool is a laser system optically connected to each crystal which allows to measure a calibration signal every 40 minutes. In addition, the symmetry of the system allows equalizing the response of the detector using well known physical processes (see Sec.2.3.4).

The preshower detector is a sampling calorimeter instrumented with a lead layer as a radiator to initiate the electromagnetic showers and a silicon strip sensor placed behind the radiator [75]. The total thickness of the system is 20 cm, and provides an additional $3X_0$ in its η acceptance. The task of this detector is mainly to improve the π^0 identification in the endcaps region. Moreover, the electron identification also benefits from the presence of the preshower.



(a) (b)

Figure 2.9. *(a) $r-z$ schematic view of ECAL, highlighted in light blue the crystals (from [76]). (b) Energy resolution for electrons from $Z \to ee$ decays function of $|\eta|$, measured in data collected in 2016 operations; the resolution is affected by the material in front of the crystals and is degraded close to module ends, indicated with vertical lines (from Ref. [77]).*

2.2.4 The hadron calorimeter

The neutral hadron energy measurement is the primary task of the Hadron Calorimeter (HCAL) since it is the single subdetector in CMS capable of performing efficiently this task. Moreover, HCAL contributes more in general to charged hadron energy measurement and identification. In addition, like for ECAL, the HCAL information is used in the hardware trigger of CMS. HCAL is composed of four subsystems. The Barrel Hadron calorimeter (HB) is placed outside the ECAL barrel and covers up to $|\eta| < 1.3$. The two Endcaps Hadron calorimeter (HE) extend the acceptance to $|\eta| < 3$ and are placed outside the ECAL endcap. The Forward Hadron calorimeter (HF) covers up to $|\eta| < 5$ and is placed 11 m far from the IP, outside the Muon detector and the magnetic field return yokes. The Outer Hadron calorimeter HO is located in the barrel region, but outside the superconducting coil.

HB and HE are sampling calorimeters, made out of alternate layers of brass plates and scintillators, the former as absorber material, and the latter to produce the detectable signal. They guarantee 6 λ_I both in the barrel and endcaps region. HF is instead a Cherenkov calorimeter, instrumented to steel interlayered with quartz-fibers, which collect the Cherenkov light produced in the absorber [78]. HO is made of plastic scintillators and uses the solenoid magnet material as the absorber. At least 12 (10) λ_I are guaranteed in the barrel (endcap), taking into account all the material between the IP and the HO detector, i.e. the silicon tracker, ECAL, HCAL and the magnet. This feature allows HO to collect the far tails of the hadronic showers, and ensure the hermeticity of the entire detector.

The HCAL granularity is coarser than ECAL and a single HCAL module matches with a matrix of 5×5 ECAL crystals (0.087×0.087 radians in the $\eta \times \phi$ plane), which allows an easier matching and identification for the shower initiated in ECAL. The HCAL energy resolution, measured with a pion test beam, taking into account the showering in ECAL as well can be written as:

$$\frac{\sigma_E}{E} = \frac{110\%}{\sqrt{E/\text{GeV}}} \oplus 9\%.$$

Despite this feature is not excellent taken as a standalone performance, the role of HCAL in the entire subdetectors ensemble must be considered. HCAL actual task is to guarantee the hermeticity and to help the particle identification of all the produced particles. The poor resolution

of HCAL therefore affects only the neural hadrons, which represent only $\simeq 10\%$ of the hadronic jet energy.

2.2.5 The superconducting magnet

The CMS magnet [79] is a cylinder of length 12.5 m and diameter of 6 m, which provides a magnetic field $\mathbf{B} = 3.8$ T along z direction in the internal volume. The 220-tons cold mass store energy 2.6 GJ at full current and operate in a superconducting state at 4.5 K. The conductor is NbTi reinforced with an aluminium alloy, and the coil is realized with a 4-layer winding. The solenoid is a thin coil with a radial extent $\Delta R/R = 0.1$. The $\mathbf{B}$ flux is returned using 5 barrel wheels plus 3 disks per endcap iron yoke, where the muon system is embedded.

In Fig. 2.10 is reported a scheme of the $\mathbf{B}$ field lines, simulated with `TOSCA` software [80], in the CMS cavern. The magnetic field has been extensively measured directly, with proper probes, and indirectly using cosmic rays reconstruction inside CMS detector [81].

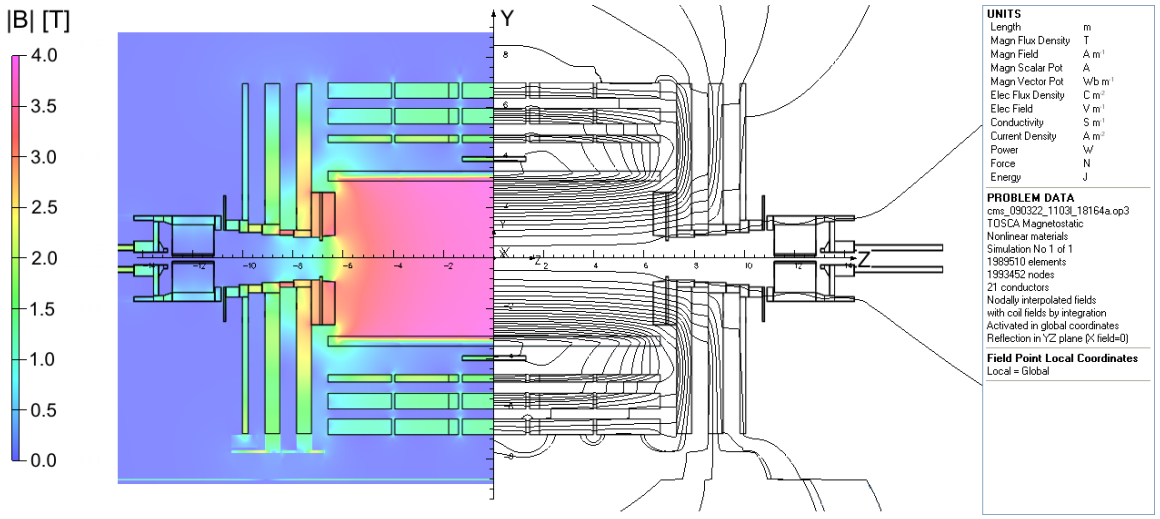


Figure 2.10. *Map of* $|\mathbf{B}|$ *(left) and its field lines (right) in the r-z plane, predicted with* `TOSCA` *(from Ref. [81]).*

2.2.6 The muon system

The muon system [82] is the outermost subdetector of CMS. It is placed outside the superconducting coil, embedded in its iron return yoke. The muon system instruments with gas detector layers (or *stations*) the space between the iron layers in the barrel ($|\eta| < 1.2$) and endcap ($0.9 < |\eta| < 2.4$) regions. The energetic muons are the only detectable particles that escape the inner regions of CMS. The measurement of the momentum, position and timing of the muons is the primary task of the muon system. The muon system is used mainly for hardware trigger and fast high-p_T tracking.

The silicon tracker remains the primary tool to measure the momentum of the muons up to $p_T \lesssim 200$ GeV, and in this momentum range the muon system is used only to *identify* the muon tracks. At higher p_T the tracker p_T resolution degrades and the lever arm of the muon system stations starts to be important for accurate measurement of the muon p_T. Below 3 GeV the bending of the muon tracks prevents the muon to escape the internal region of the detector. In the p_T range 3-6 GeV, the muons reach the muon system but are stopped inside the return yoke iron, therefore the muon system information can be only partially recovered.

The barrel region is instrumented with Drift Tubes (DT) stations. They are organized in 4 layers in r, in 12 ϕ sections, and each layer is divided into 5 modules (one per return yoke wheel) at fixed r. The DTs have a section of 4.2 cm $\times$ 1.3 cm with a conducting wire in the centre, toward which the ionization electrons drift providing the position measurement. The DTs are not affected by the magnetic field which is confined in the iron layers, hence, particle propagation is linear inside the DT layers. The DTs provide 100 μm resolution in r-ϕ and 150 μm resolution in z.

The endcaps region is instrumented with Cathode Strip Chambers (CSC), organized in 4 layers in z for each endcap. The CSC are used in this η region because of their better performance with higher particle rates and in presence of magnetic fields. The strips are placed radially, while the anode wires along ϕ, providing a measurement in the $r-\phi$ plane with a resolution of 75-150 μm and a z resolution of about 200 μm.

In addition to DT and CSC, both the barrel and the endcaps are instrumented with Resistive Plate Chambers (RPC.). The RPC operate in avalanche mode, and the readout strip pitch is coarse, about 1 cm, with the strip oriented in z (r) direction in the barrel (endcaps). Their primary goal is the time measurement, for muon triggering and bunch crossing identification. The RPCs time resolution is about 3 ns, to be compared with the 25 ns of the bunch crossing rate. They are installed up to $|\eta| = 1.6$, because the RPCs cannot sustain a high radiation environment. To cover this gap in the forward region, the first layer of Triple Gas Electron Multiplier detectors (GEM) has been installed in the Long Shutdown 2. The GEM are capable to increase the redundancy of the muon system and accomplish the task of RPC in the $1.6 < |\eta| < 2.4$ region [83].

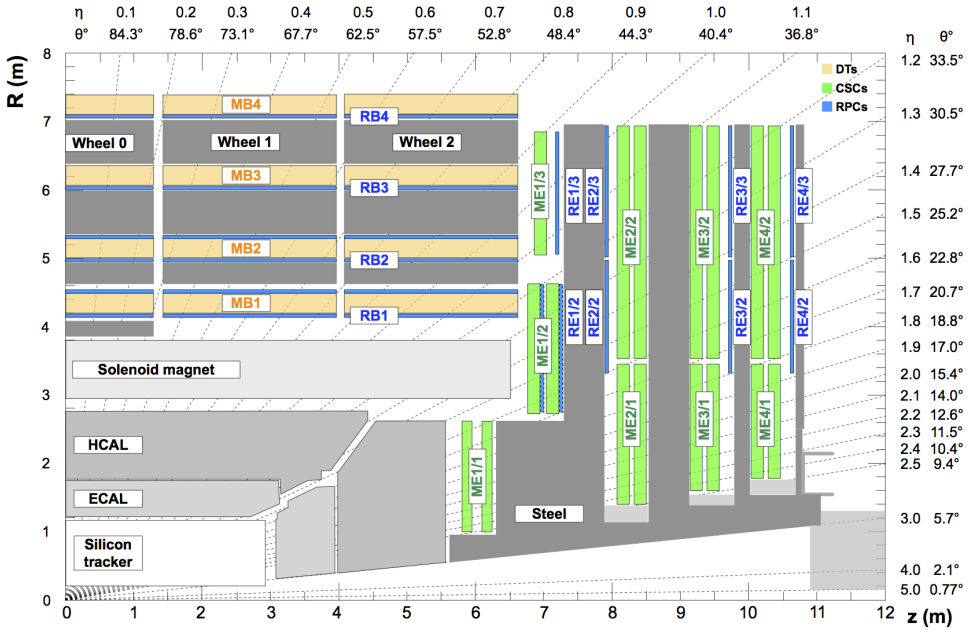


Figure 2.11. *$r-z$ cross-section of a quadrant of the CMS detector, with the Run 2 configuration. The locations of the various muon stations and the iron yoke wheels and disks are highlighted. The DTs stations are labelled MB and the CSCs stations are labeled ME. The RPCs stations are mounted in both the barrel and endcaps of CMS, where they are labelled RB and RE, respectively (from Ref. [84]).*

2.2.7 The trigger system

The LHC bunch crossing frequency is 40 MHz, and the full event information from CMS readout is about 1 MB of data per event. The allowed readout bandwidth is about 2 GB/s, therefore a trigger system must select the events to reduce the rate to about 1 kHz. The pp cross section at

LHC center-of-mass energies is ~ 100 mb, while the interesting events for CMS physics program have a cross section several orders of magnitude lower (for instance the electroweak cross section is order of tens of nb, see Fig. 2.3(b)). A data reduction algorithm is therefore needed. In conclusion, only the events which pass the trigger selection will be fully readout and stored permanently.

The CMS trigger system is subdivided in two stages: the *Level 1 trigger* (L1), and the *High Level Trigger*. The former is a hardware trigger with short latency, which strongly reduces the event rate and uses the information of the faster subsystem only, the latter is a more refined software selection algorithm that uses the full CMS information.

2.2.7.1 Level 1 Trigger

The L1 is realized with programmable hardware reducing the input rate of 40 MHz to 100 kHz. L1 processes the information from ECAL, HCAL and the muon system since they are the only subsystem whose information can be readout and reconstructed at 40 MHz. The L1 exploits a synchronous pipeline where the detector information of up to 160 bunch crossings can be stored. This approach has a latency of 3.2 μs.

The L1 implements different algorithms, called *L1 bits* or *seeds*, to decide if an event needs to be discarded or not. The seeds are grouped in various types according to the detector information used in the specific seed. For each type different thresholds can be set, depending on the value of physical quantities that are used in the selection algorithm.

The L1 hardware has been strongly upgraded during the Long Shutdown 1, to face the higher luminosity of the Run 2. In this process, the primary goal was to obtain the same efficiency in the selection of interesting events, despite the higher collision and pile-up rate.

2.2.7.2 High Level Trigger

The HLT is a software selection algorithm with uses a farm with more than $3 \cdot 10^4$ CPUs to process in parallel the events selected by the L1 and to reduce the event rate from 100 kHz to 1 kHz. The output of HLT is permanently stored on disk. The CPUs of the farm run a reconstruction software similar to the CMS software used in the standard *offline* reconstruction. The differences are only in some highly time consuming procedures which are optimized and simplified for the HLT.

The HLT implements the reconstruction of the events in modules dedicated to specific physical objects reconstruction (jets, leptons, photons...), paired with proper selection filters. A single list of modules and filters is called *HLT path*. A path can be fed by a single or multiple L1 seeds. In the selection process, the general approach is to discard an event as soon as it does not satisfy the filter requirement, without executing the subsequent modules if there is already enough information to make a decision. This method aims to avoid CPU wastage by running unnecessary reconstruction. To further optimize the CPU time, usually the more sophisticated (and time consuming) algorithms are run near the end of the path.

A *HLT menu* is a list of the various paths, and represents the logical OR of the filters of the included paths. The output of the HLT menu is grouped in primary datasets based on similar topology and event properties. A single primary dataset can be filled with events from different HLT paths.

An example of the performance of HLT selection and reconstruction is reported in Fig. 2.12, using multiple di-muon paths to select the events and reconstruct the di-muon invariant mass.

2.2.7.3 The trigger rates

Some L1 seeds and HLT paths with loose thresholds need a *prescaling* after the procedures i.e. a random selection not based on physical quantities. This is needed to fulfil trigger rate constraints of L1 or HLT.

The use of a synchronous pipeline in the L1 and a CPU farm that works in parallel in the HLT makes possible the tuning of thresholds or prescales factors. During a single fill of LHC the luminosity gradually decreases because of the beam losses and the effects *pp* collisions at IP. Since the L1 trigger rate is limited by the peak rate, because of the fixed pipeline dimension, the thresholds and the prescale factors are relaxed during the fill to optimize the bandwidth and keep the L1 trigger rate close to the maximum value of 100 kHz. In the HLT, because of the parallel processing, the rate is limited on average to 1 kHz. Therefore in the initial part of the fill the thresholds and the prescale factors are kept above 1 kHz, then with the decrease of the luminosity, the trigger rate also decreases keeping the average below 1 kHz.

This approach maximizes the number of collected events, keeping the trigger rates constant fill by fill. Discrepancies from a constant trend are used to identify possible malfunctioning of the system.

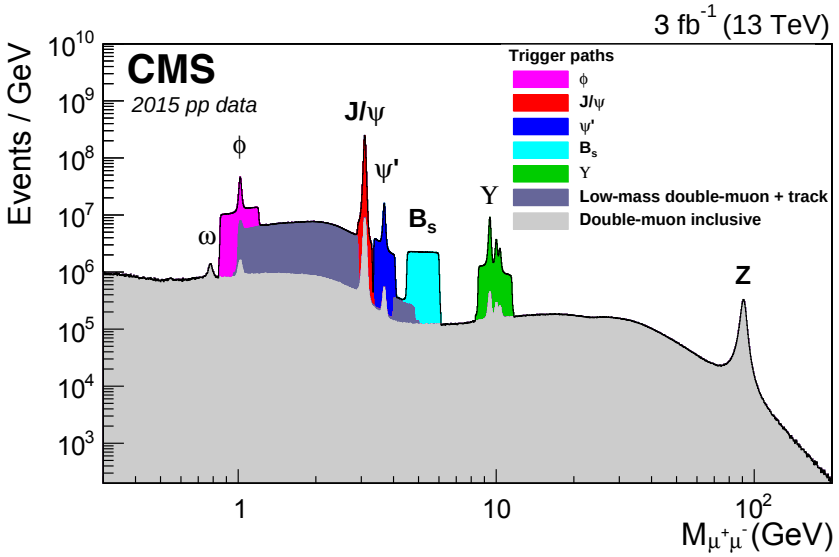


Figure 2.12. *The di-muon invariant mass distribution reconstructed by the CMS HLT, from data collected in 2015 operation. The data collected with inclusive double-muon trigger algorithm events are shown as well as triggers dedicated to selecting resonances at low masses (from Ref. [84]).*

2.3 The event reconstruction of CMS

After the multi-level trigger event selection, the signals from subdetectors of CMS are stored permanently on disk. At a later stage, the *CMS offline reconstruction* software is used to build high-level observables using the detector information, for physics analysis purposes. The reconstruction proceeds in multiple steps, operating in modules for specific tasks, which are collectively known as `CMSSW` [85].

The fist first stage of the event reconstruction uses the information of the different subdetectors independently. The result of this first stage are different collections of physical objects (charged particle tracks from the tracker, photon and electron from ECAL...), typically with lower quality

with respect to the final products of the entire reconstruction. In a second stage of the reconstruction, the information produced by the various subdetectors is combined in the *particle flow* reconstruction.

This section does not cover exhaustively the CMS reconstruction, but describes mainly the relevant aspects for the W mass and properties analysis.

2.3.1 Tracking and vertexing

The first reconstruction step in the silicon tracker is called *local reconstruction*, and it builds the clusters and then the hits starting from the detectors signals. Afterwards, the tracker reconstruction targets the charged particle track reconstruction (or *tracking*). The input of the tracking step are the individual hits from pixel and strip detectors, which represent the best estimate of the position of the charged particles as it crosses the tracker layers. The primary task of the tracking is to link the hits produced from the same particle to reconstruct the trajectory of all the particles in the tracker acceptance. Due to the presence of the solenoidal magnetic field of CMS, almost constant in the tracker volume, the charged particles have almost an helix trajectory with radius $R[\mathrm{m}] = p_T[\mathrm{GeV}]/(0.3B[\mathrm{T}])$. The tracking task is thus the measurement of the five parameters which describe the helix for each trackable particle.

Due to the high granularity of the CMS tracker, $\mathcal{O}(10^2 - 10^4)$ hits lie on the silicon layers in a single event. The combinatorial burden is therefore not affordable with brute force approaches. For exhaustive reports of the techniques developed to overcome this challenging task see for instance Ref. [86–88].

The CMS track reconstruction algorithm, called *Combinatorial Track Finder* (CTF), is based on the combinatorial Kalman filter (CKF) [88–91]. The two main ideas of this approach are: to perform pattern recognition and track fitting in the same framework and to manage the high level of complexity of the events (i.e. the combinatorial burden) with multiple passes of the same reconstruction sequence. The first iterations look for tracks that are easier to find and then the associated hits are removed. Then the following iterations look for more difficult kinematic regions (low or very high p_T, displaced vertex, high η ...), but each iteration searches in a less dense environment because of the removed hits. This approach is called *iterative tracking* and each pass proceeds in four steps:

1. **Seed generation**. The proto-tracks, called seeds, are built with the use of few hits (from 2 to 4) from specific layers of the tracker. This rough estimation of the track parameters is used as the starting point for the second step. The minimum requirements to obtain an estimate of the tracks parameters are three points, as two 3D hits with the vertex constraint or three 3D hits. An outward seeding has been chosen because the pixel layers provide 3D information and the occupancy is lower than in the strip detector layers. For each iteration, a set of *seeding layers* and a *tracking region* are defined. The seeding layers are the detectors where the seed hits are searched (a pair, triplet or quadruplet of tracker layers), the tracking regions are the kinematic or geometric selections applied on the hits to identify the seeds in the phase space of the region of interest. Since 2017, if the seeding layers are pixel triplets or quadruplets layers a Cellular Automaton [92] is used to produce the seed list instead of the tracking region constraint.

2. **Pattern recognition**. The CKF builds the track candidates starting from the track-seeds parameters, and extrapolating the track to the outer layers looking for compatible hits. Several possible extrapolations are considered and the track parameters are then updated exploring multiple hypotheses of hit association. The hypotheses are selected or discarded

relying on the χ^2 of the candidate. The CKF *navigates* outward until the tracker edge is reached or too many missing hits are present. The navigation is done taking into account the material effects (multiple scattering, energy losses). If the track candidate has enough hits the extrapolation is also repeated inward to recover precision in the seeding region, linking the hits found in the extrapolation through the seeding layers to the track candidate. If the fraction of hits shared between two tracks is above a given threshold the track with fewer hits is discarded.

3. **Fitting**. The fit of the trajectory is performed using the Kalman filter and smoother [93], moving outward with the Runge-Kutta propagator [94] which takes into account both the material effects and the inhomogeneities of the magnetic field. Then the fit is repeated outside-in, to further improve the performance.

4. **Quality flagging**. The track candidates are flagged with different tags depending on their quality (based on the number of hits, χ^2, track parameters ...), or discarded if the quality results too low. Three quality classes are realized: *loose*, *tight* and *high-purity* [95].

After the selection, the track collections from the various iterations are merged in a single collection, called *general tracks*. More details about the track reconstruction can be found in Ref. [95].

In Fig. 2.13 the performance of tracking estimated on CMS simulation are shown. The efficiency on simulation is defined as: $\varepsilon = N_{\text{assoc}}/N_{\text{sim}}$, where N_{sim} is the number of simulated tracks and N_{assoc} is the number of simulated tracks associated to a reconstructed one. The simulated track flagged as "associated" if a reconstructed track with at least 75% of its hits coming from the single simulated track is found. Fig. 2.13(a) shows the tracking efficiency as a function of simulated track p_T highlighting the contribution of each seeding iteration. In the shown case, $t\bar{t}$ events with high purity and some fiducial selection applied, the efficiency saturates at 90% in the range 1 GeV $< p_T <$ 100 GeV. At higher p_T the global tracking performances degrade, because of strong interaction of the pions with the detector material and because a relevant fraction of tracks of this topology is produced in the core of energetic jets, in which the dense environment makes the seeding inefficient (more details about this effect are given in Appendix A). The degradation of the performance below $p_T \sim 1$ GeV is due to multiple scattering and energy loss effects. Fig. 2.13(b) shows the tracking efficiency, on the same sample, as a function of η, comparing the performances of the 2016 and 2017 configuration. Another quantity to assess the tracking performance is the fake rate is defined as: $R_{\text{fake}} = N_{\text{not assoc}}/N_{\text{reco}}$, where N_{reco} is the number of reconstructed tracks and $N_{\text{not assoc}}$ is the number of reconstructed tracks *not* associated to a simulated one. Fig. 2.13(c) and 2.13(d) show the fake rate, on the previously discussed sample, as a function of p_T and η, comparing the performances of 2016 and 2017 configuration. The clear improvement of the efficiency and reduction of the fake rate is due to the new pixel detector and the cellular automaton used in the tracking.

Despite the lower performances, the 2016 case must be taken as a reference for the W mass and properties measurement, given the data set used in the analysis discussed in this thesis. However, in the case of muon events, relevant for the sake of the W mass and properties measurement, the tracking efficiency is close to unity and the fake rate is close to zero in all the kinematic range considered. Quantitative results are discussed in sec. 2.3.2.

The tracker reconstruction is also devoted to the primary vertex reconstruction. The reconstructed tracks with higher quality and low displacement from the beam spot are clustered and a fit is performed to estimate the primary vertex position with high precision. CMS vertexing adopts a deterministic annealing algorithm [97] for clustering and an adaptive vertex fitter for the refitting [98]. The core idea is to use an approach similar to the Kalman filter, where the

tracks have the role of the hits in the tracking, introducing annealing in the iterations to cope with the presence of a non-negligible fake rate. The quality of the vertex strongly depends on the number of tracks used in the clustering and its resolution is typically in the range of $10-50$ μm. Multiple vertices can be reconstructed for each event, because of pile-up presence. The primary vertex is chosen as the one with the higher $\sum_{\text{tracks}} p_T^2$.

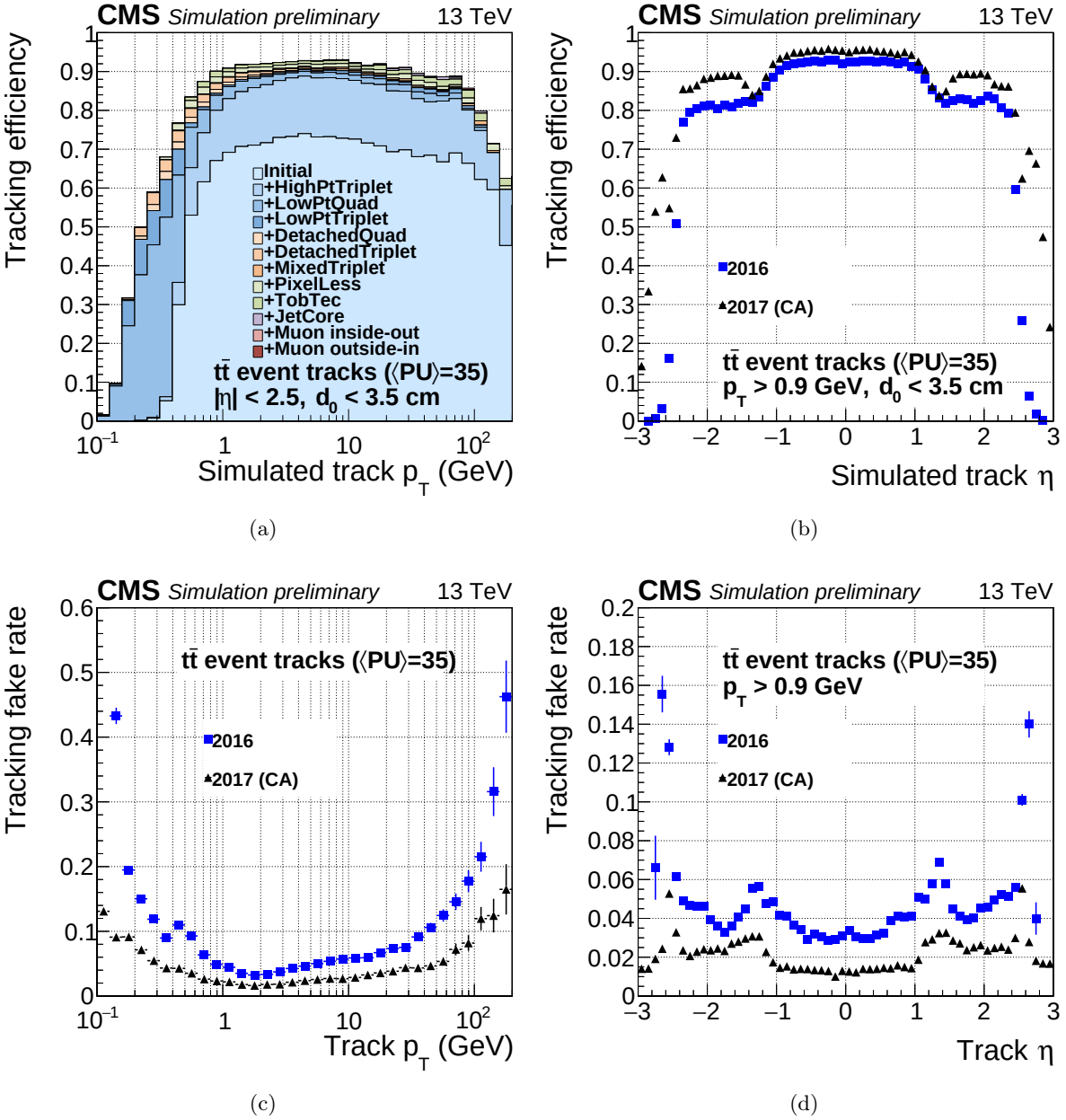


Figure 2.13. *(a): Tracking efficiency as a function of track p_T for 2017 configuration; the contribution to the total efficiency of the various iteration of the CTF are shown in different colors. (b): Tracking efficiency as a function of track η, comparing 2016 and 2017 configuration. (c,d): Fake rate as a function of p_T and η, comparing 2016 and 2017 configuration. In all the cases the efficiency is estimated using CMS simulation, with simulated $t\bar{t}$ events with impact parameter $d_0^{track} < 3.5$ cm, $\langle PU \rangle = 35$ and $|\eta^{track}| < 2.5$ (a,c) or $p_T > 0.9$ GeV (b,d) (from Ref. [96]).*

2.3.2 Particle Flow reconstruction

The Particle Flow (PF) algorithm [99], exploits the combined information of the various subdetector of CMS to improve the reconstruction performance of each detector and to give a global event description. It provides high level objects, called *PF candidates*, which can be used at the analysis level. The PF approach can be proficiently exploited at CMS thanks to some non-trivial features of the detector. Firstly, the high granularity of both the tracking system and the calorimeters allows having a high-resolution picture of the events, with high separation power between different particles. Secondly, the configuration of the detectors, with most of the subsystems inside the magnet volume strongly reduces the material placed in front of the calorimeters and thus allows a better geometrical linking of the tracks with the calorimeter energy deposits. Lastly, the hermeticity of CMS provides an almost complete picture of the event. The former feature guarantees to link information from one particle only, reducing the incorrect links. The two latter features allow keeping the efficiency of the linking for a single particle high, without information loss.

In the first step, the PF algorithm recollects the detector information, for each detector, to obtain *PF elements*: inner tracks, muon tracks, and the different kind of calorimeter clusters (ECAL, HCAL, preshower). This step is essentially a reorganization of the single detectors reconstruction objects in the same common PF framework, as PF elements.

The second step of the PF is the grouping of the PF elements into *PF blocks*. Two PF elements are assigned to the same PF block if they are linked by the algorithm. The links are established under specific condition depending on the nature of the PF elements under consideration. Only nearest neighbours in $\eta - \phi$ plane are tested to limit the combinatorial burden (and so the computation time). A distance is defined for each pair of linked elements, which assess the quality of the link. An overview of the designed links is given below:

- An inner track is linked to a calorimeter cluster if the extrapolation of the track in the ECAL, HCAL or preshower volume matches the position of a calorimeter cluster (in $\eta - \phi$ plane for the barrel and in $x - y$ plane for the endcap). The algorithm takes into account the gaps in the calorimeter or low-momentum material effects. The distance is defined using the difference between the center of the cluster and the extrapolation point on the proper plane. If multiple tracks are linked to the same ECAL cluster only the best-quality link is kept. The link of an ECAL cluster compatible to the tangent trajectory of tracks is considered to properly recover the bremsstrahlung photons from electrons. The possibility of $\gamma \rightarrow e^+e^-$ conversion in the tracker material of these bremsstrahlung photons is also considered in the link-finding.

- A calorimeter-to-calorimeter cluster link is established if the higher granularity calorimeter (ECAL or preshower) cluster lies in the envelope of the less granular calorimeter (HCAL or ECAL, respectively) cluster. The distance is assessed from the cluster centers on the proper plane ($\eta - \phi$ or $x - y$, depending on η). In case of multiple linking to the same higher granularity calorimeter cluster, the best-quality link is kept. The link of ECAL-to-ECAL cluster is considered in the case of shared cells between two clusters.

- Tracks are linked together through a common secondary displaced vertex, to take into account nuclear interaction. In this case, at least three tracks are required, with the incoming track reconstructed with the hits between the primary and the secondary vertex.

- An inner track is linked to a muon track testing both the extrapolation from the inner tracker to the muon system and the extrapolation from the muon system to the inner tracker.

In the third step of the PF algorithm, each PF block is analysed separately and the PF candidates are identified within each PF block. Once a candidate is reconstructed and identified the associated elements are removed from the PF block. The search proceeds in the following order:

- **Muons.** A muon is identified by a track in the inner tracker and in the muon system without particular activity in the calorimeters. A link from the inner tracker to the muon system (*tracker muon*) or from the muon system to the tracker (*global muon*) is sufficient to identify a PF candidate if the object is isolated i.e. if the energy in the calorimeter and the p_T of the tracks in the proximity of the candidate is lower than 10% of muon p_T. If the link is found inside a jet a strict identification criterion must be applied to avoid charged hadron misidentification, requiring additional hits in the muon system and a calorimeter deposit compatible with a muon crossing. More details on Muon object are given in Sec. 2.3.3.

- **Electrons and isolated photons.** An electron is identified by a track in the inner tracker and an electromagnetic shower in ECAL, without a shower in HCAL. A photon is identified as an electromagnetic shower in ECAL without an associated track in the inner tracker or a shower in HCAL. For a candidate electron the links with ECAL cluster compatible with bremsstrahlung photons, with possible e^+e^- conversion, are taken into account. A dedicated track reconstruction, called Gaussian-Sum Filter (GSF, see Ref. [100]) is performed for the electrons using the combined information of the tracker and the calorimeter deposits of the bremsstrahlung photon. Isolated photon PF candidates are identified simultaneously to the electrons, checking the absence of associated tracks. Additional quality requirements are set to avoid misidentification.

- **Hadrons and non-isolated photons.** First, the non-isolated photons are identified, because of their larger abundance in the jets. A non-isolated photon PF candidate is built from a link between an ECAL deposit without a correspondent track or a link between an ECAL deposit without a HCAL deposit, if the link is in the tracker acceptance or not, respectively. Then the neutral hadrons are identified from a link between a HCAL deposit without a correspondent track or a link between an ECAL deposit with a HCAL deposit, if the link is in the tracker acceptance or not, respectively. In all the discussed cases the energy of the particles is measured from the calorimeter deposits only. As the final step, the charged hadrons are identified from a link between the inner tracker and an HCAL deposit. If a significant excess of energy compared to the track p_T is found, it is assigned to additional neutral hadrons. The measurement of the energy for the charged hadrons is computed with a weighted average of the track momenta and the cluster energy.

- **Nuclear interactions.** The tracks link to secondary vertices helps to remove from the PF blocks additional elements produced by the nuclear interaction, improving the energy measurement of the entire event. The secondary charged particles are replaced by a single charged hadron with $\pi^\pm$ mass and momentum compatible with the sum of the measured ones. In the case of proper measurement of the incoming primary track, the energy imbalance between the primary and the secondary particles is used to extract the undetected energy in the nuclear interaction.

The fourth step of the PF is the global post-processing, where the global properties of the reconstructed event are studied to find possible anomalies which are signature of particle misidentification or wrong reconstruction. A typical figure of merit is the missing transverse momentum, p_T^{miss}, which represents the imbalance of the momentum of visible particles in the transverse plane. In case of extremely large p_T^{miss} a misreconstruction is further investigated. Typically a misidentification of the leading p_T particle in the event, often a muon track, can produce large

p_T^{miss}. Different hypotheses are tested in this case (misidentified charged hadron, cosmic ray, wrong momentum measurement, displaced semileptonic decays of b hadrons...), to refine the reconstruction and eventually reduce p_T^{miss}.

The result of the PF algorithm is a list of PF candidates, with per particle energy measurement and a particle identification label in the following list: $\mu^{\pm}$, $e^{\pm}$, γ isolated, γ not isolated, h^0 (n or K_L), $h^{\pm}$ ($\pi^{\pm}$, $K^{\pm}$, $p^{\pm}$). In Fig. 2.14 a sector of CMS detector is shown with the typical signatures of the classes of particles identified by the particle flow.

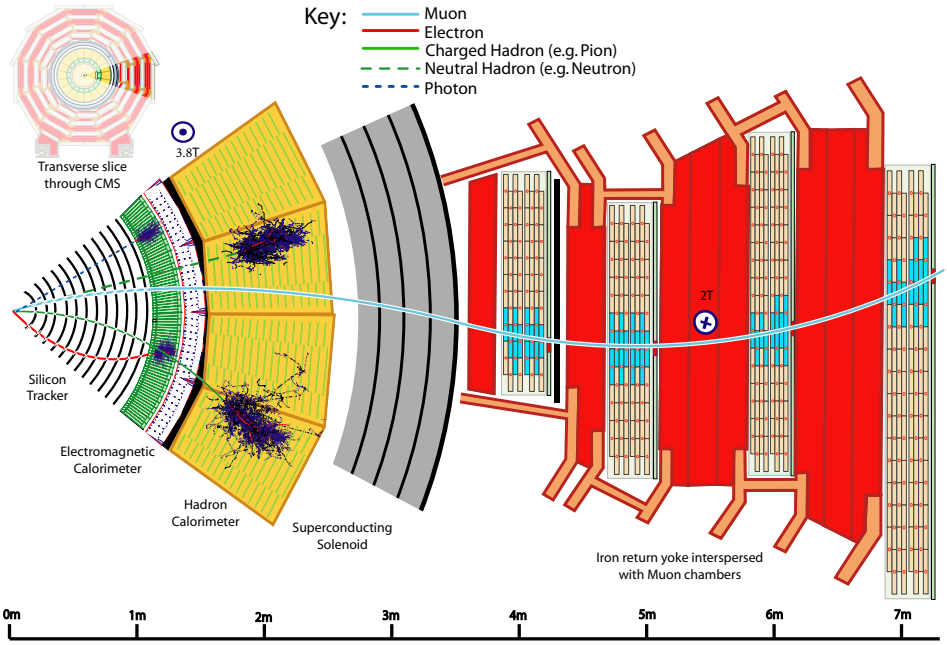


Figure 2.14. *Sketch of a transverse plane section of CMS detector; highlighted the different kind of interaction for the detectable particles (note: the muon and the charged pion are positively charged, the electron is negatively charged); the average magnetic field and the distance from IP is also shown (from Ref. [99]).*

2.3.3 Muons

In this section, some highlights about the muon reconstruction and its properties are given, since they have a central role in the W mass and properties measurement. The muon reconstruction is performed prior to the PF algorithm and it already combines the information of the tracker with the muon system. The PF algorithm further refines the reconstructed muons, exploiting also the calorimeter information. A quality flag is finally assigned to each muon, accordingly to the measured properties and the algorithm(s) used for its reconstruction.

A relevant variable used in the PF muon reconstruction is the *isolation* of the muon particle. This property estimates the activity in the spatial region surrounding the particle. Conceptually the isolation of a given lepton defined as:

$$\text{Iso}_{\text{ideal}} \equiv \sum_{x^{\pm}} p_T + \sum_{\gamma} p_T + \sum_{h^0} p_T,$$

where sum run over all the charged particles ($x^{\pm}$), photons (γ) and neutral hadrons (h^0) produced from the primary vertex in a cone aligned with the direction of the particle of inter-

est. The aperture of the cone depends on the specific selection and it is measured in term of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, to be longitudinal-boost invariant.

From an experimental point of view, the p_T of the neutral particles is measured only in the calorimeters, and it is mixed with the contribution from the particles from the pile-up vertices particles. Therefore in the PF algorithm the isolation is defined as:

$$\mathrm{Iso_{PF}} = \mathrm{Iso}_{\delta\beta} \equiv \sum_{x^{\pm}_{\mathrm{PV}}} p_T + \max\big(0, \sum_{\gamma} p_T + \sum_{h^0} p_T - \frac{1}{2}\sum_{x^{\pm}_{\mathrm{PU}}} p_T\big), \tag{2.2}$$

where the first sum runs over the charged particles associated with the primary vertex ($x^{\pm}_{\mathrm{PV}}$), the γ and h^0 sums run over the PF-candidates. The last term is an estimation of the neutral pile-up contribution and the sum runs over the charged particles associate with pile-up vertices. This is needed because the neutral particles do not have an associated vertex, and therefore the PF can not distinguish neutral hadrons and photons coming from pile-up vertices from those coming from the primary vertex. The factor 1/2 is motivated by the isospin symmetry, thus about 1/3 of the hadron components of a jet are expected to be neutral and 2/3 charged. In all the cases only the particles found in a cone or $\Delta R = 0.3$ or 0.4 (accordingly to the required selection) with respect to the lepton direction are considered. This type of correction to the isolation is called the $\delta\beta$ correction. Note that this definition subtracts the PU contribution on average, while more refined definitions are possible (see Ref. [101] for details) but are not used in the analysis described in this thesis. Because of the p_T-dependency of the isolation variable the *relative isolation*, $\mathrm{RelIso}^{\mu}_{\mathrm{PF}} \equiv \mathrm{Iso_{PF}}/p^{\mu}_T$, is typically used at analysis level.

Beyond to the PF muon collection, described in the previous section, muons are classified into 3 additional collections, based on the reconstruction algorithm used:

- **Standalone muons:** tracks obtained from the muon system information only, without using the inner tracker.
- **Global muons:** standalone muons geometrically matched to a track from the inner tracker, propagating the standalone muon to the inner tracking, and check the compatibility of the track parameters.
- **Tracker muons:** inner tracker tracks extrapolated outward to the muon system, and matched with hit or segments found by the muon system. For each track with $p_T > 0.5$ GeV, $p > 2.5$ GeV this extrapolation is tested.

The momentum scale calibration is a central step in the muon reconstruction. The *scale* is the transfer function between an estimated quantity and the "true" one. This scale typically is biased, and it must be calibrated to allow a precise and accurate use of the estimated quantity. The calibration consists of a measurement of this bias and a proper correction of the quantity. The scale correction can be different between data and simulation, according to the details of the latter. Thus the scale must be separately measured and corrected both in data and simulation.

The muon momentum scale and resolution is measured with different methods, accordingly to the kinematic region of interest and the required granularity of the calibration. At low and intermediate p_T ($\lesssim$ 100 GeV), the pp collisions data are directly used to correct the muon momentum scale and assess the calibration. Two methods have been developed by CMS and are reported in the next two sections. All the described methods exploit a well-known resonance as a standard candle for the muon momentum.

Additional methods have been developed by CMS to calibrate muon momentum scale and resolution above 100 GeV. The first method uses the cosmic ray muons data to measure the muon

resolution. It selects cosmic muons which cross the pixel detector and then compare the two legs of the cosmic muon track. The second method uses the muons from pp collisions to measure the muon momentum scale at high p_T with the *generalized endpoint method* [84].

2.3.3.1 Momentum scale and resolution: the Rochester corrections

This method, called *Rochester corrections* [102, 103], is based on the proof of concept described in Ref. [104]. It exploits the $Z \to \mu\mu$ events as a standard candle to calibrate the momentum scale of the muon. A $Z \to \mu\mu$ MC sample is generated and reconstructed using the MC-truth information, with the purpose of realizing an unbiased perfectly-aligned reference sample. Then, the calibration proceeds in two steps.

In the first step, a correction is derived comparing the mean of $1/p_T$ distribution, $\langle 1/p_T \rangle$, of the reference sample with the observed $\langle 1/p_T \rangle$, as a function of η^μ, p_T^μ and the charge. This correction is iterated until the agreement with the reference sample is reached. The correction of $1/p_T$ is separated in two parts, one multiplicative $\sim C_1(\eta, \phi)1/p_T$ dependent on the mismodelling of the magnetic field, and one additive $\sim qC_2(\eta, \phi)$ dependent on the misalignment of the detector. An additional multiplicative parameter is added to match the Z peak position fitting the generated spectrum with a Breit-Wigner function. This correction is derived independently for data and MC (comparing both with the reference sample) since the momentum scale correction are in principle different.

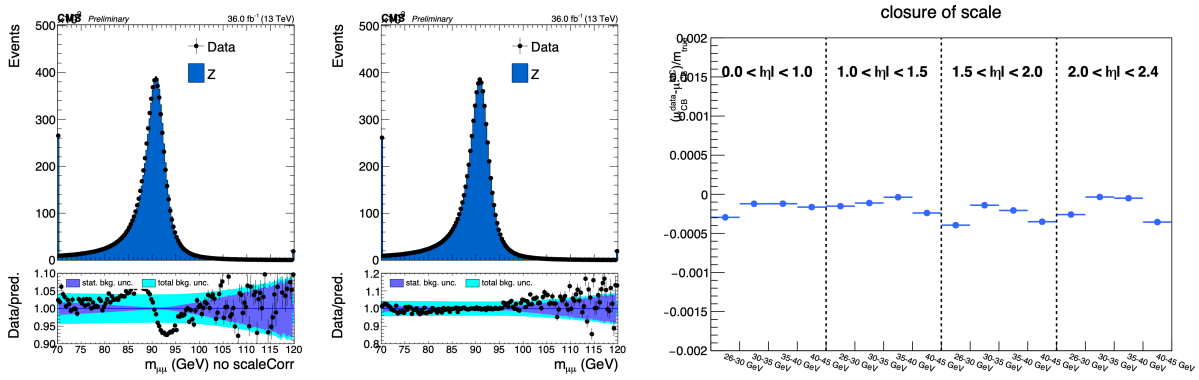
Also the muon p_T resolution is corrected at this stage, matching the invariant mass $m_{\mu\mu}$ between data and MC. This is performed fitting some parameters which shape the muon momentum resolution from $1/p_T^{\text{smear}} = \frac{1}{p_T} + R_1 \cdot G(1, R_2)$, where $G(\mu, \sigma)$ is a normal distribution of mean μ and variance σ.

All the biases are removed from the MC after the application of the corrections from the first step. However, some residual discrepancies between data and MC can be present to mismodelling of trigger and reconstruction efficiencies, used to identify $Z \to \mu\mu$ events. These results in different invariant mass $\langle m_{\mu\mu} \rangle$ as a function of ϕ and η, between data and MC. The previously defined reference sample and the $\Delta M_Z = m_Z^{\text{measured}} - m_Z^{\text{expected}}$ variables are used to correct these discrepancies. In particular the $\langle 1/p_T \rangle$ is corrected for an additional factor $1 + 2\Delta M^Z / m_Z^{\text{data}}$, iterating the correction until the compatibility with 0 of $\Delta M^Z / m_Z^{\text{data}}$ is reached.

In conclusion, the applied Rochester corrections are in the form:

$$1/p_T^{\text{RC}} = \kappa(\eta, \phi)\frac{1}{p_T} + q\lambda(\eta, \phi), \quad \sigma_{p_T}^{\text{RC}} = \sigma_{p_T}/\kappa_{\text{res}}(|\eta|). \tag{2.3}$$

The effect of this correction is shown in Fig. 2.15, from the studies performed in the analysis from Ref. [11]. The muon momentum scale correction are about 0.2% (0.3%) in the barrel (endcap), while the resolution is measured to be 1-2% (5-8%) up to 100 GeV. The residual not-closure of muon momentum scale is below $5 \cdot 10^{-4}$ in the acceptance region.



(a) (b)

Figure 2.15. *(a): Reconstructed invariant mass $m_{\mu\mu}$ with a $Z \to \mu\mu$ event selection, before (left) and after(right) the Rochester corrections application to the p_T^{μ}. (b): Residual muon momentum scale not-closure after the application of the Rochester correction, as a function of p_T^{μ} and $|\eta^{\mu}|$. From Ref. [11, 105].*

2.3.3.2 Momentum scale and resolution: Kalman filter method

This method [106, 107] uses a physics-motivated model to correct $1/p_T$:

$$\frac{1}{p_T^{\text{corr}}} = (A-1)\frac{1}{p_T} + qM + \frac{1/p_T}{1+\epsilon \sin\theta/p_T}, \tag{2.4}$$

where q and θ are the charge and the polar angle of the muon, respectively. The parameters A, M and ϵ model the physical effects which can distort the curvature. The mismodelling of the magnetic field is encoded the multiplicative term $A = (A_1 + A_2\eta^2)$. The residual misalignment of the detector results in the charge-dependent additive term M, function of ϕ and estimated in bins of η. The mismodelling of the material budget and consequent energy loss is described with ϵ, in bins of η.

The calibration has been performed, as previously described, using J/ψ and $\Upsilon(1S)$ resonances. It has been implemented with a Kalman filter [93], estimating $m_{\mu\mu}$ and propagating the uncertainty of the tracks on event by event basis. Then, the parameter of Eq. 2.4 has been estimated comparing the reference and the reconstructed $m_{\mu\mu}$, with proper uncertainty propagation (where the reference mass has been estimated as the mean value of the simulated events after FSR). The performance of the method, validated on $Z \to \mu\mu$ sample are reported in Fig. 2.16(right). This method allows to obtain a scale precision below $2 \cdot 10^{-4}$.

The muon momentum resolution has been calibrated, based on the model derived in Ref. [108]:

$$\left(\frac{\sigma_{p_T}}{p_T}\right)^2 = a^2 + c^2 p_T^2 + \frac{b^2}{1 + d^2/p_T^2}. \tag{2.5}$$

The model takes into account of the effect of multiple scattering ($\sigma_{\text{ms}} \sim a p_T$), the effect of intrinsic hit resolution ($\sigma_{\text{hit}} \sim c p_T^2$) and their correlation (term function of b and d). The calibration is provided in bins of η^{μ}. The momentum resolution measured on data or in simulation as a function of p_T^{μ} and η^{μ} is shown Fig. 2.16(left).

The described method has been successfully exploited in the analysis of Ref. [107] and is still under development to improve its performance.

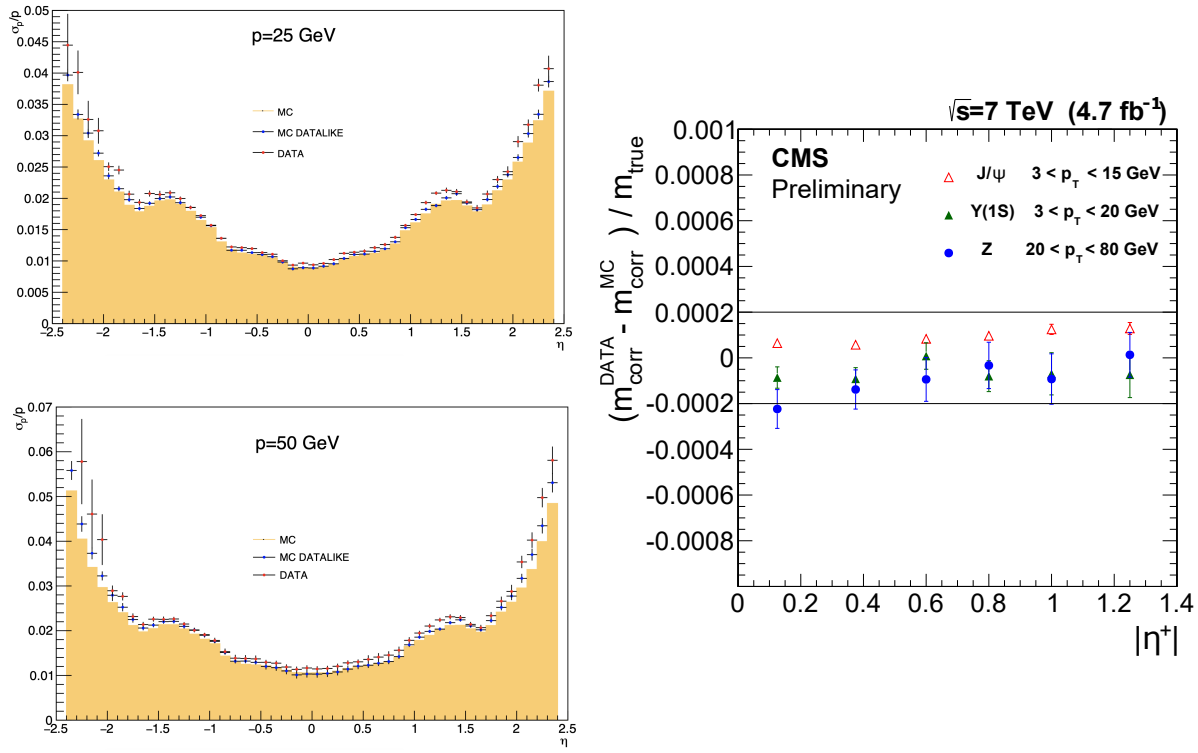


Figure 2.16. *Muon traverse momentum resolution as a function of η^{μ} for $p_T^{\mu} = 25$ GeV (upper left) and $p_T^{\mu} = 50$ GeV (lower left), measured on data, on a simulated J/ψ and Z sample (Datalike) and on simulated single muon sample (MC), from Ref. [108], where more details about the measurement method are provided. (Right plot): closure between data and MC of the muon scale calibration from Ref. [107] after applying the correction estimated on J/ψ and Υ samples, as a function of $|\eta|$ of the positive muon.*

2.3.3.3 Identification and Isolation criteria

Additional identification (ID) criteria are added to the muon collection to improve its purity [84].

Loose ID. It aims to suppress the misidentification of hadrons as muons and require the muon to be a PF muon and also a tracker or global muon.

Medium ID. It aims to select prompt and heavy flavour decays muons and it requires hits in 80% of the silicon layers and high compatibility with muon segment in the muon system with additional χ^2_{fit} quality selection; these cuts are tuned to obtain an efficiency of 99.5% on W and Z events.

Tight ID. It aims to suppress muon decays in flight and hadron punch-through (hadron showers which reach the muon system) and requires the muon to be reconstructed as tracker and global muon, with hits on at least 6 silicon layers, at least a pixel hit, two muon station compatibility, high χ^2_{fit} quality, and a transverse (longitudinal) impact parameter $d_{xy} < 0.2$ cm ($d_z < 0.5$ cm).

Soft ID. It aims to select low-p_T muons (for B-physics and quarkonia analysis) and requires the muon to be a tracker muon with high purity tracking flag, with hits on at least 6 silicon layers, at least a pixel hit, tight muon system segment matching.

High momentum ID. It is optimized for muons with $p_T > 200$ GeV and requires the same conditions of the tight muon ID, but χ^2 request is relaxed, to take into account inefficiencies due to muon radiation in the detector material.

Moreover, two muons isolation working points are defined. The isolation is typically used to distinguish between prompt muons production and secondary muons produced inside hadronic jets. For PF muon the *tight isolation* working point requires $\text{RelIso}_{\text{PF}} < 0.15$, while *loose isolation* working point requires $\text{RelIso}_{\text{PF}} < 0.25$. In both cases, the isolation definition uses $\Delta R < 0.4$. If the isolation is evaluated with reconstructed tracks only, without PF information, the $\Delta R < 0.3$ is used and $\text{RelIso}_{\text{PF}} < 0.05$ ($\text{RelIso}_{\text{PF}} < 0.10$) is required for Tight (Loose) Isolation working point. These working points are defined to have efficiencies of 95% and 98% for Tight and Loose Isolation respectively, tuned on simulated tight muons with $p_T > 20$ GeV from $Z \to \mu\mu$ decays.

2.3.3.4 Efficiencies: Tag-and-probe method

The efficiency of the reconstruction algorithm and of the selections described in the previous section must be measured with high precision to fully exploit the muon objects as a tool for physics analysis. The efficiency reported, for instance, in Fig 2.13 are measured using the CMS MC simulation i.e. comparing the reconstructed results with the MC truth. Nevertheless, it is mandatory to measure the efficiencies directly on data, to properly apply the described selection without adding a bias from the MC simulation.

The efficiencies for muon reconstruction and selection are measured using the *tag-and-probe* method. The purpose of the method is to assess the efficiency of a given selection in the most unbiased way. The logic is the following. Di-lepton decays from a very well know resonance are selected, applying only a very loose invariant mass selection. In general $Z \to \mu\mu$ events or $J/\psi \to \mu\mu$ events are chosen for this study. The events are selected with strict requirement on one lepton, called the *tag* lepton, and with relaxed selection on the second lepton, called the *probe* lepton. The ε_X of the selection X is quoted as the fraction of probes which overcome also the selection X, called *passing probes* (while the probes which not pass X are called *failing probes*). In this way, the tag selection is completely decoupled from ε_X measurement. It is used together with the invariant mass selection only to enrich the dilepton sample under study of "real" leptons, suppressing the background. Therefore the tag selection can be also correlated with X. The probe selection instead is the baseline on which ε_X is measured and must be chosen accordingly to the purpose. The number of passing and failing probes are measured from a fit to the invariant mass distribution, which is done simultaneously to the background fit. In this way the residual background yield is subtracted.

The muon efficiency at CMS is computed factorizing it in multiple independent components in such a way that [84, 109]:

$$\varepsilon_\mu = \varepsilon_{\text{tracking}} \cdot \varepsilon_{\text{reco+ID}} \cdot \varepsilon_{\text{iso}} \cdot \varepsilon_{\text{trig}}. \tag{2.6}$$

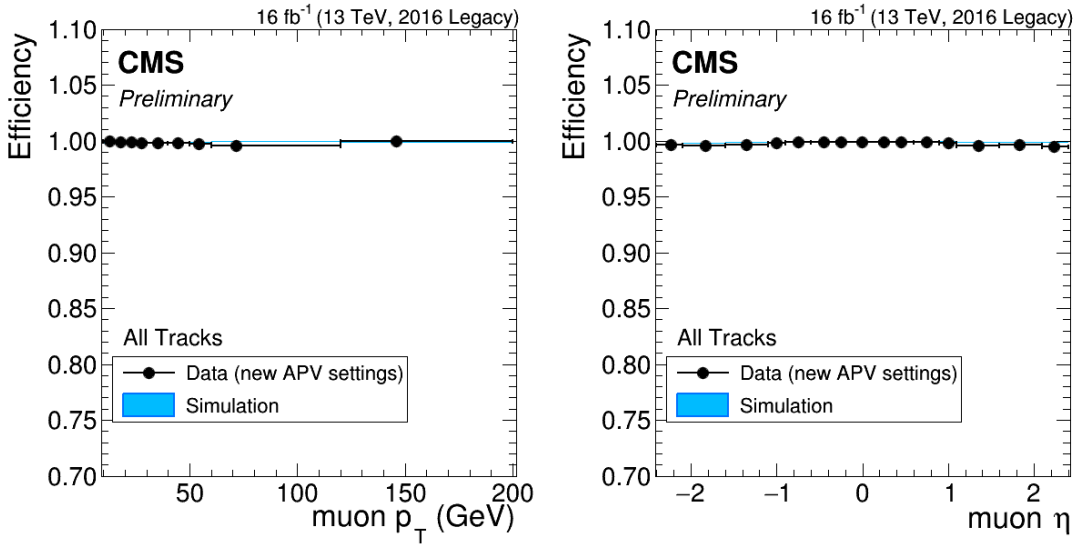
The tag-and-probe is performed independently for each component of ε_μ. The $\varepsilon_{\text{tracking}}$ is the tracker track reconstruction efficiency, and its measurement will be described in the next paragraph. $\varepsilon_{\text{reco+ID}} \sim N_{\text{reco+ID}}/N_{\text{tracking}}$ is the muon ID selection efficiency combined to the efficiency to reconstruct the muon in the muon system, given a tracked muon. $\varepsilon_{\text{iso}} \sim N_{\text{iso}}/N_{\text{reco+ID}}$ is the muon isolation efficiency given a reconstructed and identified muon. $\varepsilon_{\text{trig}} \sim N_{\text{trigger}}/N_{\text{iso}}$ is the trigger efficiency, given an isolated muon.

The tree latter efficiency measurements depend on the details of the analysis, in terms of tag and probe definition and related uncertainties. Therefore they will be described in detail in the next section from the specific case of W rapidity and helicity measurement [11] . For the tracking efficiency, the procedure is more standard [110] and it will be discussed here.

To measure the tracking efficiency dimuon pairs are selected with an invariant mass between 70 and 115 GeV. The tag muon is required to satisfy the tight ID, with $p_T > 29$ GeV, tight isolation working point and geometrical matching with a single muon trigger object. The probe

is defined as a standalone muon. The passing probe criterium is the matching with an inner tracker track, with minimal quality criteria ($\Delta R < 0.3$). The very loose matching requirement results in a matching efficiency very close to 100%. On the other hand, the efficiency must be corrected to take into account fake reconstruction and matching. A simultaneous fit to $Z \to \mu\mu$ and background yield to the observed lineshape is performed to achieve this.

The results of the efficiency measurements with this approach for data collected in 2016 and the corresponding MC Simulation sample are shown in Fig. 2.17. The efficiency is above 99% in the entire kinematic range of p_T^μ, η^μ. For this reason, the tracking inefficiency for muons is typically neglected and no uncertainty is assigned to it. These results can be compared with Fig. 2.7 in which only the information of the silicon tracker is used to produce the track collection. The strong improvement is therefore due to the muon system contributing to the track seeding.



(a) (b)

Figure 2.17. *Tracking efficiency as a function of p_T (a) and η (b) measured with the tag-and-probe method. The result obtained with Madgraph DY simulation is also shown. The tracks are produced by the complete reconstruction algorithm of CMS (from [111]).*

2.3.3.5 Efficiencies: from W rapidity and helicity measurement

In this section is reported the efficiencies measurement performed in Ref. [11], which will be directly exploited in the W mass and properties analysis. The usual definition of muon efficiency follows the Eq. 2.6. However $\varepsilon_{\text{tracking}}$ has been considered unitary, while $\varepsilon_{\text{reco+ID}}$ and ε_{iso} has been estimated in a single step since the former is very close to unity. Thus, the following efficiencies have been measured:

- **Selection efficiency** ε_{sel}: the probability to pass the isolation and identification criteria, given a reconstructed (PF) muon.
- **Trigger efficiency** $\varepsilon_{\text{trig}}$: the probability to pass the trigger requirement, given a isolated and identified PF muon.

In the same fashion of Eq. 2.6 the two steps are defined in a factorized way, to be able to obtain $\varepsilon_\mu = \varepsilon_{\text{sel}} \cdot \varepsilon_{\text{trig}}$. With the efficiency measurement on data and on MC, the data-to-MC Scale

Factors are defined:

$$\mathrm{SF} = \mathrm{SF}_{\mathrm{sel}} \cdot \mathrm{SF}_{\mathrm{trig}}, \quad \text{with: } \mathrm{SF}_{\mathrm{sel}} \equiv \frac{\varepsilon_{\mathrm{sel}}^{\mathrm{data}}}{\varepsilon_{\mathrm{sel}}^{\mathrm{MC}}}, \quad \mathrm{SF}_{\mathrm{trig}} \equiv \frac{\varepsilon_{\mathrm{trig}}^{\mathrm{data}}}{\varepsilon_{\mathrm{trig}}^{\mathrm{MC}}}. \tag{2.7}$$

The total scale factor SF is applied as a weight to the MC samples to correct for different efficiency between data and simulation.

The efficiency measurement has been performed with the tag-and-probe technique, using high-purity $Z \to \mu\mu$ events. The event selection requires opposite charge and same flavour leptons and the dilepton invariant mass between 70 GeV $< m_{\mu\mu} <$ 110 GeV. The tag is identified with very tight isolation and identification requirements. The probe selection is defined based on the efficiency under estimation. For the $\varepsilon_{\mathrm{sel}}$ the probe requirement is to be a PF muon, while the passing probe must pass the isolation and identification criteria of the event selection (see sec. 5.3 for the details). For the $\varepsilon_{\mathrm{trig}}$ the probe requirement is to be an isolated and identified PF muon, while the passing probe must pass the trigger requirements.

The MC efficiencies estimation is performed by counting the failing and passing probes in the $m_{\mu\mu}$ window since the purity of the MC sample is 100%. The measurement of the efficiency on data must take into account the presence of background events and subtract them to avoid bias in the efficiency estimation. Thus, a fit to the Z invariant mass distribution is performed to estimate the number of passing and failing probe, modelling both the signal and the background contributions. The signal is fitted with a binned template derived from the simulation, convolved with a Gaussian (with floating scale and width) to describe the detector resolution. The background is modelled with an exponential function.

A relevant discrepancy between the two muon charges has been observed in the data efficiency. Therefore the $\varepsilon_{\mathrm{trig}}$ has been measured separately for the two charges, while $\varepsilon_{\mathrm{sel}}$ not. The tag-and-probe measurement has been performed in coarse bin of p_T^μ and fine bins of η^μ, for $|\eta^\mu| < 2.4$ and 25 GeV $< p_T^\mu <$ 55 GeV. The η^μ granularity is crucial for rapidity and angular coefficient measurements, while the choice of coarse p_T^μ bin is driven by the need to keep the statistical uncertainty low. The bins have a width of $\Delta\eta = 0.1$, while the Δp_T ranges between 1.5 GeV and 5 GeV. The efficiencies in bins of η^μ and p_T^μ are reported in Fig. 2.18.

Because of the coarse p_T^μ bins the efficiency can sensibly vary in nearby bins, therefore a smoothing procedure has been implemented to avoid artificial jumps in the scale factors. The efficiencies have been fitted with an error function in each bin of η^μ:

$$\varepsilon_{\mathrm{step,\ kind}}^{\eta,q}(p_T) = p_0 \mathrm{erf}\left(\frac{p_T - p_1}{p_2}\right), \quad \mathrm{erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy,$$

where step $= \{\mathrm{sel}, \mathrm{trig}\}$, kind $= \{\mathrm{data}, \mathrm{MC}\}$, η and q indicate the η-charge bin under analysis and p_0, p_1, p_2 are the free parameters of the fit.

Finally the scale factors have been defined as the ratio of the smoothed efficiency on data and the smoothed efficiency on the MC, in bin of η (and charge for $\varepsilon_{\mathrm{trig}}$) for the two efficiencies separately, according to Eq. 2.7. The SFs are evaluated for each η^μ bin and continuous in p_T^μ (for practical reasons they will be actually binned with a very fine binning, $\Delta p_T = 0.2$ GeV, in their application), for each muon charge.

The $\mathrm{SF}_{\mathrm{sel}}$ has been also smoothed in η. The smoothing has been performed splitting each η bin into 3 sub-bins, leaving the central unchanged and reweight the two lateral sub-bins according to a linear interpolation between the central sub-bin and the central sub-bin of the neighbour η bins. These procedure was not applied for $\mathrm{SF}_{\mathrm{trig}}$ to preserve its relevant, not-trivial η trend.

The smoothing procedure impact has been estimated by observing the ratio of the SF with and without smoothing on the $p_T \times \eta$ plane. The discrepancy is below the 2% level for all the bins.

A closure test on $Z \to \mu\mu$ events has been performed. The distribution of p_T^μ and η^μ of data and MC with and without the $\mathrm{SF_{sel}}$ application have been compared. Both the leading and subleading muons have been used. The $\mathrm{SF_{trig}}$ has not been applied since the $\varepsilon_{\mathrm{tirg}}$ using a single muon trigger in a dimuon sample is expected to be very close to 100%. The comparison shows a major improvement in the closure between the data and MC distributions. The residual discrepancy is below the $\sim 3\%$ level in most of the acceptance range. Only in the $|\eta^\mu| > 2$ and $p_T^\mu < 30$ GeV regions a discrepancy $\sim 4-5\%$ is observed, probably due to trigger turn-on effects. The $\eta^\mu \times p_T^\mu$ maps of the $\mathrm{SF_{sel}}$ and $\mathrm{SF_{trig}}$ for the two charges is shown in Fig. 2.19. The systematic uncertainties related to the SFs are described in Sec. 5.4. Additional details and plots about the SF can be found in Ref. [105].

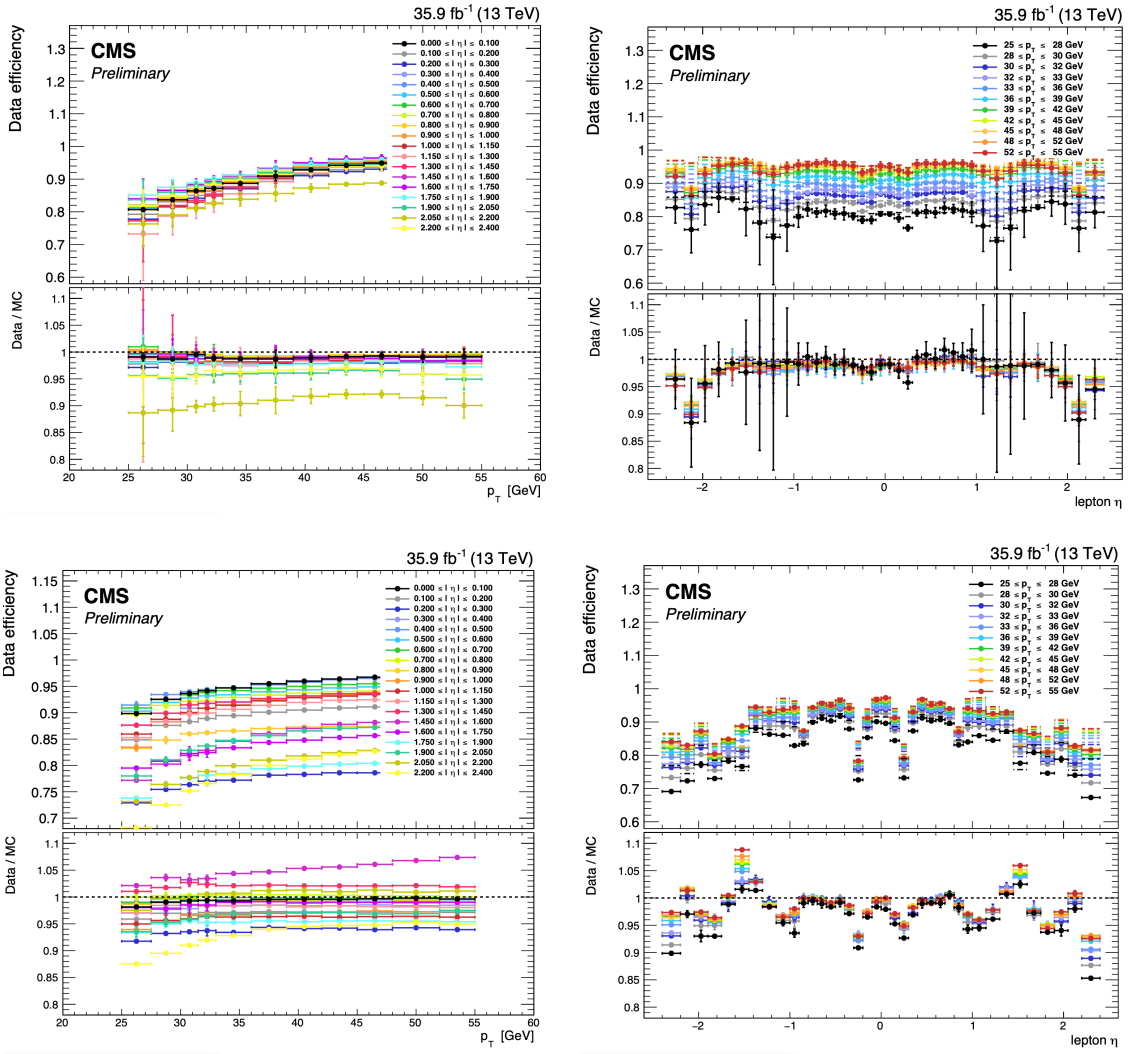


Figure 2.18. *Values of the selection (upper plots) and trigger-μ^+ (lower plots) efficiencies measured on data as a function of p_T^μ (left plots) or η^μ (right plots) before the application of the smoothing. The SFs are reported in the panels below each plot (from Ref. [105]).*

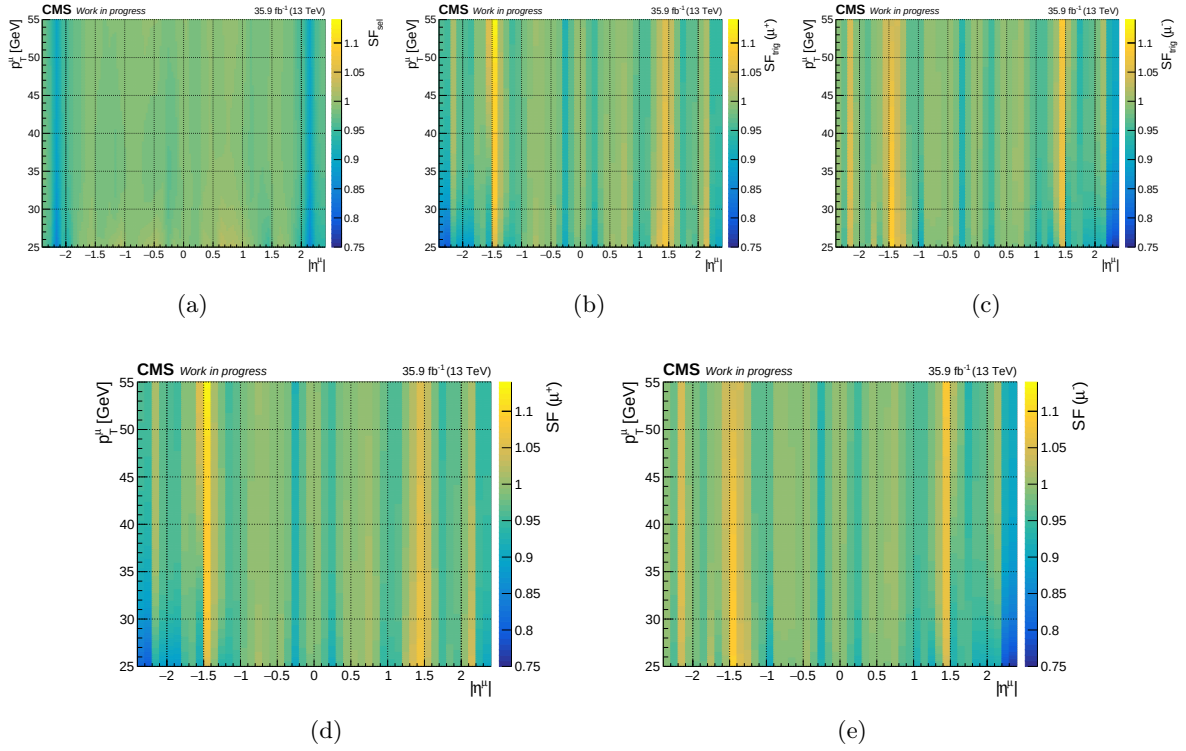


Figure 2.19. *Values of the selection scale factor (a) and trigger scale factor for μ^+ and μ^- (b and c, respectively) on the $\eta^\mu \times p_T^\mu$ plane. This is the result of the ratio of data and MC efficiency p_T-smoothed (trigger) or p_T and η-smoothed (selection). The total SF, as defined in Eq. 2.7 in (d) and (e) for μ^+ and μ^-, respectively.*

2.3.4 Jets and missing transverse momentum

In this section, the jet and missing transverse momentum objects are described. They are not directly used in the W mass and production properties analysis as variables of interest, but the applied selection relies on the missing transverse momentum.

2.3.4.1 Jets

In high energy pp collisions, free light quark and gluons are products with the highest cross section. At high q^2, α_s is small and thus the prompt products are loosely-bound (the asymptotic freedom of QCD) and the probability of gluon radiation is high. As the energy density decreases, far from the IP region, α_s value increases and the QCD tends to bound together quarks and gluons. Due to the QCD colour confinement only colourless particles can be observed free. Therefore the quarks and gluons must hadronize into bound states without colour charge. The result of these processes are sprays of collimated hadrons in conical structures that are globally called hadronic jets and are the footprint of the QCD.

The jets do not have a unique definition but are the result of the clustering process of the final state particles and signals collected in the detector. The jet definition must be collinear safe i.e. the result must be unchanged in case of collinear gluon emission (splitting of a particle in two particles with the same direction and a fraction of momentum each). The jet definition must be also infrared safe i.e. the result must be unchanged in case of soft gluon(s) emission (addition of extra particles with $p \to 0$). These requirements make the result predictive, otherwise the obtained jet can not be produced from theoretical calculations. This happens because the σ_{QCD}

is dominated by InfraRed and Collinear (IRC) contributions, but they cancel out in the cross section calculation. Moreover, the jet definition should be resilient to hadronization, i.e. robust against not-calculable dynamics, to obtain stable results in similar conditions.

The jets at CMS are built from PF objects accordingly to the anti-k_T clustering algorithm [112], which satisfy all these properties. The anti-k_T is part of a class of algorithms, IRC safe, which differ only from the value of a parameter, κ. These algorithms loop over the particles of the event, evaluating for each pair:

$$\begin{cases} d_{ij} = \min\left(p_{T,i}^{2\kappa}, p_{T,j}^{2\kappa}\right) \cdot \frac{\Delta_{ij}^2}{\Delta R^2} \\ d_{iB} = p_{T,i}^{2\kappa} \end{cases}$$

where B means "beam", ΔR is the aperture of the cone, a fixed parameter of the algorithm, $\Delta_{ij} = \sqrt{(Y_i - Y_j)^2 + (\phi_i - \phi_j)^2}$ and $\kappa = 1, 0, -1$ for the k_T [113], Cambridge-Aachen [112] and anti-k_T algorithm respectively. After every single evaluation if $d_{ij} < d_{IB}$ the i and j objects are clustered, otherwise i is removed from the list of objects to cluster and added to the list of the jets. The clustering stops when no more objects remain. In this algorithm, ΔR governs the geometry, while κ the power of the dynamics.

The anti-k_T choice produces reasonable jet from a physical point of view: soft objects tend to cluster with harder one, in the case of a hard object in a radius $2\Delta R$ a perfect conical shape is obtained, if two hard objects are close they can merge if $\Delta_{ij} < \Delta R$, o they can split with not-conical shape if $\Delta R < \Delta_{ij} < 2\Delta R$. At CMS $\Delta R = 0.4$ has been chosen in Run 2 for standard jets, while $\Delta R = 0.8$ is used for heavy objects initiated jets. In CMS reconstruction the `FastJet` package implementation [114] of anti-k_T is used.

The PF jets have an angular resolution of $\Delta\eta, \Delta\phi = 0.3$ (0.01) at $p_T = 20$ GeV (1 TeV), which is a strong improvement with respect to the standalone calorimeter jet (up to a factor 3-4). The Jet Energy Scale (JES) requires a post-processing step with the application to the *Jet Energy Corrections* (JEC) to be properly used. The JEC are applied both to data and simulation, to obtain a more accurate and precise estimation of the measured jet energy and to calibrate the MC on data.

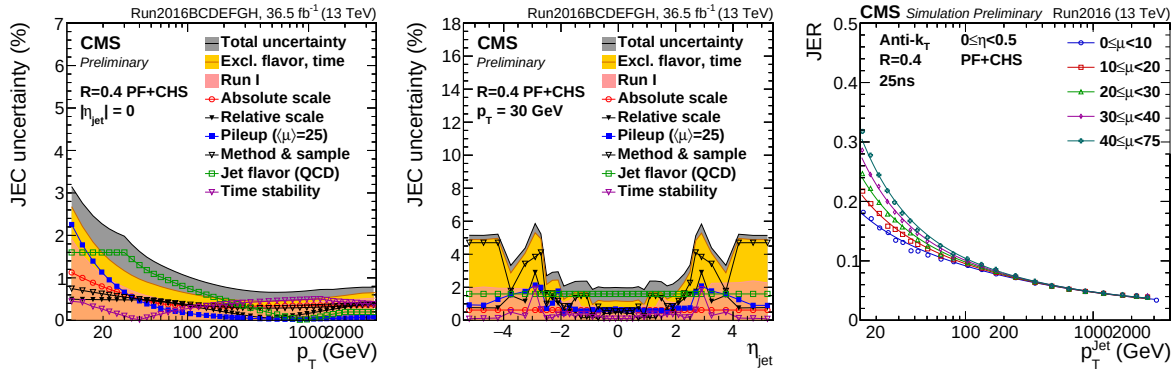
The JEC consist in a reweighing of the jet momentum $p^{\text{corr}} = Cp^{\text{raw}}$, where a factorized approach is chosen [115, 116]:

$$C = C_{\text{offset}}(p_T^{\text{raw}}, \eta) \cdot C_{\text{response}}(p_T', \eta) \cdot C_{\text{rel}}(\eta) \cdot C_{\text{abs}}(p_T'') \cdot C_{FT}(p_T''') \tag{2.8}$$

- **Offset corrections C_{offset}:** due to excess of energy for the PU or electronic noise. The PU tracks are subtracted and then the per-event median energy density $\langle\rho\rangle$ is used correct the p_T^{raw}, function of p_T^{raw} and η.
- **Response correction C_{response}:** due to simulation imperfection detector response. It is estimated as the ratio between the generator level p_T response and the reconstructed one, as a function of η and $p_T\prime = C_{\text{offset}}p_T^{\text{raw}}$.
- **Relative residual correction C_{rel}:** due to inhomogeneities in the response of the detector, which must be equalized. It is measured using the imbalance in p_T response in dijet events, using the barrel region as a reference. It is estimated as a function of η. It is applied to data only.
- **Absolute residual correction C_{abs}:** absolute calibration of the energy scale. It is estimated as the ratio between the response of a reference known object and the jet response in Z/γ + jet events. It is estimated as a function of $p_T'' = C_{\text{response}}C_{\text{rel}}p_T\prime$. It is applied to data only.

– **Flavour and time stability corrections C_{FT}:** due to different response between gluon and quark jet, and time instability of the response. The flavour correction is evaluated from the simulation, the time correction from data in different periods, as a function of $p_T''' = C_{\text{abs}} p_T''$.

These corrections are measured every year of data taking because they are strongly dependent on the detector and event environment conditions. The sum of the various uncertainty related to the JEC is shown in Fig. 2.20(a), 2.20(b). The Jet Energy Resolution (JER) can be also properly calibrated, comparing the energy resolution in data and in the simulation. The result is shown in Fig. 2.20(c).



(a) (b) (c)

Figure 2.20. *JEC uncertainty separated in the various components, as a function of p_T (a) or η (b). Corrected JER as a function of p_T (c), for various PU scenarios.*

2.3.4.2 Missing transverse momentum

The *missing transverse momentum* or MET is defined as [117]:

$$\mathbf{p}_T^{\text{miss}} \equiv -\sum \mathbf{p}_T,$$

where the sum runs over all the PF candidates of the event. Since $\sum p_T = 0$ expected, for symmetry of the system, in case of fully reconstructed events a $p_T^{\text{miss}} > 0$ is the signature of invisible final states, like neutrinos. The $\mathbf{p}_T^{\text{miss}}$ variable is very sensitive to the details of the reconstruction and identification, in terms of resolution, measurements, detector inefficiencies or artefacts, and the amount of PU impact.

Since the PF jets are the objects with the worst p_T resolution, the JEC are propagated to the $\mathbf{p}_T^{\text{miss}}$ to improve its performance. The $\mathbf{p}_T^{\text{miss}}$ is therefore redefined as:

$$\mathbf{p}_T^{\text{miss}} \equiv -\sum_{\text{PF}} \mathbf{p}_T - \sum_{\text{jet}} (\mathbf{p}_T^{\text{corr}} - \mathbf{p}_T), \tag{2.9}$$

where $\mathbf{p}_T^{\text{corr}}$ is the same of JEC-corrected jet $\mathbf{p}_T$. Only jets with $p_T > 15$ GeV are included in the sum, to reduce the impact of PU jets. Moreover, jets with more than 90% of energy associated to ECAL are excluded from the sum, to remove the overlap with electrons and photons. In addition, the momentum of PF muons which overlaps with the jet are subtracted from the p_T^{jet}.

After the calibration, the JER uncertainties on p_T^{miss} ranges between 5-20% . The JES uncertainty on p_T^{miss} is less than 3% in the tracker acceptance and below 12% outside. The leptons

and photons uncertainties are $\lesssim 1\%$. Another source of uncertainty on p_T^{miss} arise from the *unclustered energy*, E_U, defined as the sum of PF candidates energy not associated with $\mu^{\pm}$, $e^{\pm}$, γ, hadronic $\tau^{\pm}$ decays or jets. The E_U uncertainty is estimated based on the resolution of each PF candidate. The leading contribution to the E_U uncertainty is given by the neutral hadrons candidates ($\sigma_{UE} \sim \min(0.25, 0.8/p_T[\text{GeV}] \oplus 0.05)$) and the candidates reconstructed in the HF ($\sigma_{UE} \sim 1/p_T[\text{GeV}] \oplus 0.05$).

Anomalous high p_T^{miss} can be often traced back to detector or reconstruction effects. Therefore some specific filters have been applied to the events to reject the ones which are most probably sources of anomalous high p_T^{miss}. Three kinds of filters have been implemented:

- **Calorimeter filters:** these filters look for noisy crystals and signal shapes, or anomalous activities in ECAL and HCAL. They can operate as an event filter or as a noise channel filter.
- **Beam halo filters:** these filters remove events with energy deposits along constant ϕ only, due to beam halo particles.
- **Reconstruction filters:** these filters remove events with poorly reconstructed high p_T muons or charged hadrons.

The effect of the described filters is shown in Fig. 2.21(a), where an excess of high p_T^{miss} events can be seen when no filtering is applied.

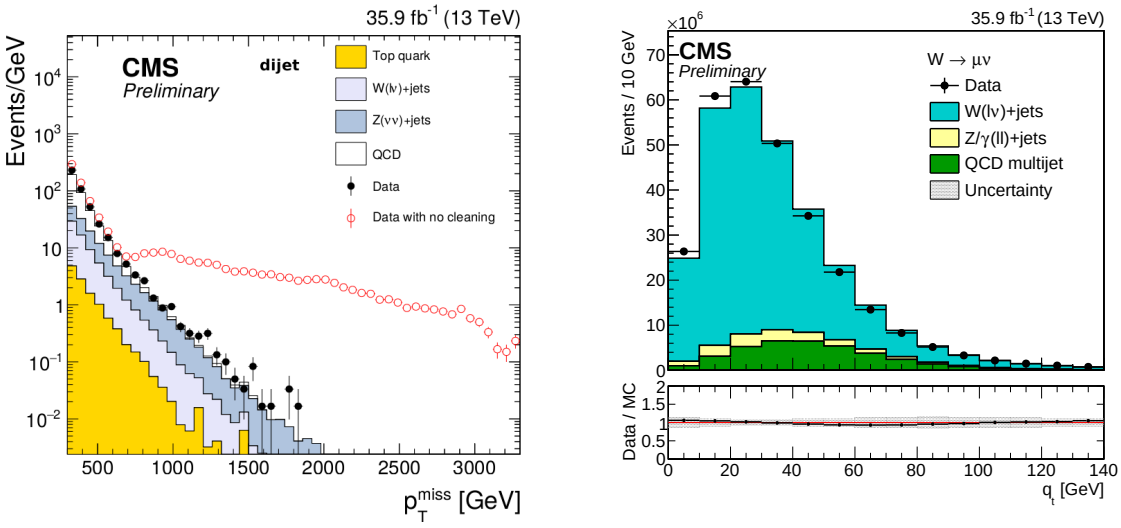
After calibration, the filtering and the response correction, the p_T^{miss} resolution can measured directly on $Z \to \mu\mu$ or $Z \to ee$ events, where no genuine p_T^{miss} is expected. The absolution resolution is between 20 and 30 GeV.

The significance of the p_T^{miss} is defined to assess the reliability of the measured quantity, and distinguish between genuine p_T^{miss} and spurious measurement. The significance quantifies the degrees of compatibility of p_T^{miss} with zero on a event by event basis. It is defined with the log-likelihood ratio:

$$\mathcal{S} \equiv 2\ln\Big(\frac{\mathcal{L}(\mathbf{p}_{T,\text{true}}^{\text{miss}} = \mathbf{p}_{T,\text{obs}}^{\text{miss}})}{\mathcal{L}(\mathbf{p}_{T,\text{true}}^{\text{miss}} = 0)}\Big).$$

The $\mathcal{L}(\mathbf{p}_{T,\text{true}}^{\text{miss}})$ is modelled as a Gaussian distribution and therefore the significance can be written as $\mathcal{S} = (\mathbf{p}_{T,\text{obs}}^{\text{miss}})^T V^{-1} \mathbf{p}_{T,\text{obs}}^{\text{miss}}$, and it is distributed as a χ^2 variable with two degrees of freedom, under the hypothesis $\mathbf{p}_{T,\text{true}}^{\text{miss}} = 0$.

An example of the performance of the $\mathbf{p}_T^{\text{miss}}$ is shown in Fig. 2.21(b). Events with a single tight muon are selected, requiring $p_T^{\mu} > 25$ GeV, rejecting events with additional leptons with $p_T > 10$ GeV or with a b-tagged jet. The result is a sample strongly enriched in W+jets events. The backgrounds are electroweak processes with similar signature and multi-jet production. The formers are estimated from CMS simulation, the latter are measured from data, using a method equivalent to the one described in Chapter 6. Given that $\mathbf{q}_T^W = \mathbf{p}_T^{\mu} + \mathbf{p}_T^{\nu}$, the W boson transverse momentum can be measured using the p_T^{μ} and the p_T^{miss}, with the assumption that the latter is a genuine measurement of the p_T^{ν}. The results show good agreement between data and the simulation, and a non negligible impact of the jet-related uncertainties. Note that the observed distribution is given by the convolution with the q_T^W spectrum and the p_T^{miss} resolution, which produces the peaking structure at $q_T^W \sim 30$ GeV, while the physical q_T^W spectrum would peak at $\simeq 4$ GeV.



(a) (b)

Figure 2.21. *(a): The p_T^{miss} distribution with and without anomalous p_T^{miss} event filters applied, compared to the CMS simulation, using the 2016 data set. (b): The q_T^W distribution measured using the p_T^{miss}, compared to the CMS simulation, using the 2016 data set; in the lower pad the ratio Data/MC is shown, with the JER, JES, and E_U uncertainty added in quadrature shown in the band. From Ref. [117]*

2.3.5 Event simulation

The event simulation is a central ingredient of the high energy physics analysis. The entire chain of physical processes which allow to perform a physics analysis at CMS are simulated: the hard interaction between partons, dynamics of the secondary processes, interaction with the detector material and electrical response of the detector. The combination of these simulation steps allows to develop, optimize and test the algorithms of event collection, reconstruction and analysis. Prior to physics analysis, the simulation of low-level variables must be calibrated on the observed detector response in real data. The muon momentum scale calibration or the JEC discussed in previous sections are examples of this calibration process.

The first step of the simulation process is the use of matrix elements generators which encode the information of the hard scattering. These generators are evaluated with fixed perturbative order accuracy in α and α_s. Often QCD resummation techniques are also used to improve the description in particular phase space regions. The experimental information from PDFs is encoded in the generators with specific libraries. The two common Monte Carlo generators used are `POWHEG` [118] and `MadGraph` [119], while a common PDFs library is `NNPDF` [120].

The second step of the simulation is the parton showering and hadronization process to produce colourless final state particles. Then the decay of unstable particles has been simulated, accordingly to experimental data. These processes are simulated by another Monte Carlo simulation, `PYTHIA` [121].

The final steps passage of final state particles through the different subdetectors of CMS. The interactions of the particles with the detector material and the detector response are generated using `GEANT4` [122] software. In Fig. 2.22 a sketch of the entire simulation chain, excluding the interaction with the detector material, is shown.

A simulated event, following the entire simulation chain, is typically called Monte Carlo (MC) event. Because of the tuning of the simulation process on the experimental condition (LHC

conditions, CMS description, state-of-the-art theory implementation, experimental inputs) is essential to have a dedicated *MC campaign* for each analysed data sample, which requires specific calibration before use.

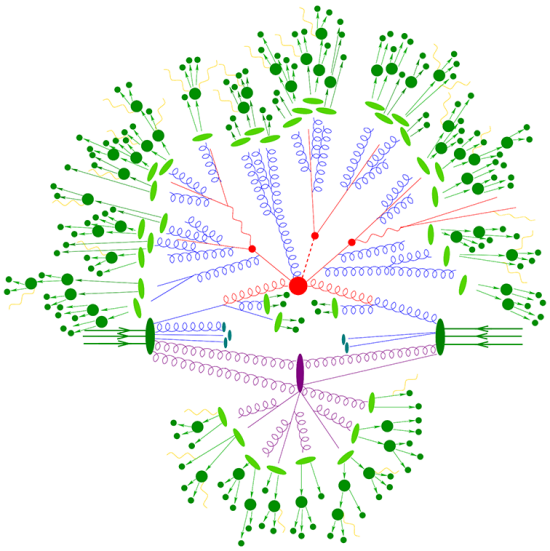


Figure 2.22. *A sketch of the simulation of a pp event. The protons are represented in dark green incoming from the left and right. The red blob and lines represent the hard processes, while the ISR and FSR are drawn in blue. A soft interaction (like a PU vertex) is also present, in purple. The lightgreen blobs represent the hadronization, while the dark green lines and blobs are the hadron decay products. The photon radiation is drawn in yellow. (from [123])*

Chapter 3

The impact of PDFs on the Z and W boson lineshape

The use of Z boson mass as a momentum scale reference is one of the central assumptions of the muon momentum scale calibration method. The Z boson mass is estimated from the di-muon invariant mass distribution, however the underlying parton luminosities can induce non-trivial effects on the invariant mass lineshape. Therefore the use of Z mass distribution as a standard candle *for the momentum scale has been further investigated to avoid inducing a bias in the calibration.*

The content of this Chapter is almost fully extracted from the paper in Ref. [6], which has been conceived, developed and published during this thesis. In this Chapter, the motivation of the study is first described in detail, followed by a phenomenological tree-level model description. Finally, a MC simulation of Drell-Yan production is used to validate the tree-level model and extend the study to the full phase-space. Additional details are provided in Appendix B.

3.1 Motivations

The Z boson mass was extracted at LEP from a fit to the lineshape, i.e. from the cross section of the process $e^+e^- \to \gamma, Z \to f\overline{f}$ measured with different beam energies spanning center-of-mass energies near $\sqrt{s} \sim M_Z$. At LEP, the lineshape was distorted by the initial state radiation of the colliding electrons, which was theoretically very well understood, resulting in a systematic uncertainty of less than 0.1 MeV on the Z mass. The Z mass was ultimately measured with an uncertainty of 2.1 MeV [124] where the largest contribution was the precision of the beam energy calibration.

The experimental situation at LHC differs by at least two aspects. Firstly, proton-proton collisions at a center-of-mass-energy $\sqrt{s}$ result in partonic $f\overline{f}$ collisions with a broad distribution of partonic center-of-mass-energies $\sqrt{\hat{s}} = \sqrt{x_1 x_2 s}$, where x_1 and x_2 are the fractions of the proton momentum carried by the interacting partons, as described in the empirical Parton Distribution Functions (PDFs). Thus the center of mass of the initial state is unknown and changes on an event by event basis covering all relevant energies. The lineshape of the boson can be measured using the decay product kinematics, like the invariant mass of the $\mu^+\mu^-$ pairs in Z decays.

Secondly, at the LHC the gauge bosons are produced at a rate that is several orders of magnitude larger than at LEP one. The ATLAS and CMS experiment have already collected about 400 million W and 40 million Z bosons each during the Run 2 operations of the LHC. By the end of Run 3, 10^8 Z leptonic decays will be available and a factor of 10 more at the end of the High Luminosity LHC program. These numbers must be compared to the 10 million Z bosons produced at LEP in all decay channels. This unprecedented number of W and Z bosons produced at the LHC offers new opportunities for precise measurements, but it also forces consideration of sources of systematic uncertainty which may have been neglected so far, e.g. those related to the modelling of the virtuality of the gauge bosons [7].

A thorough assessment of the systematic uncertainty on the lineshape is crucial because of the se of Z as a standard candle for the muon momentum calibration. In addition, this information may be useful for a new, more precise, measurement of the Z mass in case the momentum scale of the detector could be calibrated to a relative uncertainty of 10^{-5} using independent experimental information, like the J/ψ mass which is known with a relative uncertainty of 10^{-6} [18], and assuming that final state radiation effects can be understood to this level of precision.

At first order, the distribution of the virtuality Q of a gauge boson (V) originates from the convolution of a relativistic Breit-Wigner with the partonic luminosity function, see e.g. Ref. [125]. The latter is a function of the dimensionless parameter $\tau = Q^2/s$. The non-trivial dependence of the partonic luminosity on τ implies a distortion of the lineshape compared to a pure Breit-Wigner. Given the narrowness of the electroweak gauge boson width Γ_V, this effect can be treated, in first approximation, as a shift Δ_V of the mode of the distribution compared to M_V. The limited knowledge of the PDFs introduces an uncertainty on Δ_V, which contributes directly to the model uncertainty in the extraction of M_V from the dilepton mass distribution.

The goal of the study described in this chapter is to assess the size of this shift and of its uncertainty due to the limited knowledge of the PDFs. This shift can be regarded as a proxy of the systematic uncertainty on the Z mass extracted from the fit to the dilepton mass distribution.

The W boson lineshape is also distorted by the same effect. However, this is mostly of academic interest since the invariant mass of the leptonic final state cannot be measured in W decays due to the presence of the neutrino. As will be discussed in Sec. 4.3, the traditional measurement of the W mass at hadron colliders uses non-Lorentz invariant quantities (e.g. the transverse mass or the lepton transverse momentum) whose distributions have a dependence on the PDFs which induces a systematic uncertainty much larger than the effect discussed in this Chapter [126]. However, the larger than ever number of W decays collected by the LHC opens up possibilities for novel measurements that might be less sensitive to the PDFs [127, 128]. Likewise, the increased coverage in rapidity planned by the upgrades of the LHC experiments may have a strong impact in reducing the aforementioned PDF uncertainty [129, 130]. For the above reasons, the results of this study may become relevant also for a future W mass measurement at LHC.

3.2 Tree level study

A simplified model of Drell-Yan production is first considered based on a minimal set of tree-level diagrams. This approximation amounts to considering just one Feynman diagram per quark-antiquark pair, as illustrated for W boson in the upper-left Fig. 3.1. Besides accounting already for the bulk of the total cross section (about 80% for a 20 GeV threshold on the transverse momentum of the extra parton at $\sqrt{s} = 13$ TeV), these diagrams are also expected to be the most sensitive to the PDF-dependent shift under study. Indeed, they are the only $2 \to 1$ diagrams contributing to the amplitude, whereas higher-order diagrams are at least $2 \to 2$. The leading

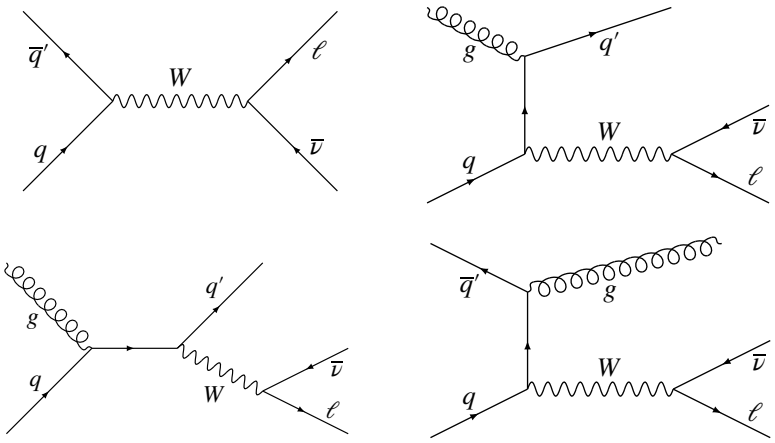


Figure 3.1. *Feynman diagrams of W boson production, the tree level diagram with $q_T^W = 0$ (upper left), the Wg associate production NLO diagram (lower right), and the two gluon Compton NLO diagrams (lower left and right). The leptonic decay of the W boson is also shown.*

examples of these diagrams are shown in the upper right and lower diagrams of Fig. 3.1 for W boson production. As such, they include additional invariants besides the transferred momentum Q. The existence of these extra scales is expected to dilute the sensitivity of the lineshape to the details of the PDFs. This assumption will be validated by a MC analysis of $pp \to V + X$ production discussed in Sec. 3.3. In the following, the `NNPDF3.0` [120] PDF set will be used to evaluate the PDFs relevant for W and Z production in proton-proton collisions at $\sqrt{s} = 13$ TeV.

Within the tree-level approximation, the double-differential cross section for $pp \to V(\to \ell\ell') + X$, as a function of the quark momentum fractions $x_{1,2}$, is given by

$$\frac{d^2\sigma_V}{dx_1 dx_2} = \frac{1}{N_C} \sum_{ij} \left[f_i(x_1) f_j(x_2) + f_i(x_2) f_j(x_1) \right] \frac{16\pi\Gamma_V^2 \mathrm{BR}_{V\to q_i\overline{q}_j} \mathrm{BR}_{V\to\ell\ell'}}{(x_1 x_2 s - M_V^2)^2 + M_V^2\Gamma_V^2}, \tag{3.1}$$

where N_C is the number of QCD colours, M_V and Γ_V are the mass and width of the resonance, $\mathrm{BR}_{V\to ab}$ are the relevant branching fractions, and the sum at the right-hand side runs over the different combinations of quark flavours contributing to the process under study. For simplicity, the scale-dependence is omitted from the quark PDF $f_i(x)$. In Eq. 3.1 a relativistic Breit-Wigner function with fixed width has been assumed, which is also the functional form used for the MC simulation discussed in Sec. 3.3. The opportunity of using a running width scheme [18] has been studied as well. However, since the interesting feature in Eq. 3.1 concerns the core of the distribution, where $Q \sim M_V$, the results would not change because the two schemes differ only when $|Q - M_V| \gg \Gamma_V$.

In order to express the double-differential distribution of Eq. 3.1 as a function of Q and y the canonical transformation:

$$y = \frac{1}{2} \ln \frac{x_1}{x_2}, \quad Q^2 = x_1 x_2 s, \tag{3.2}$$

is performed. By combining Eq. 3.1 and 3.2, the single-differential distribution $d\sigma_V/dQ$, conditional on the rapidity y, is obtained:

$$\begin{aligned} \frac{d\sigma_V}{dQ}(Q \mid y) &= \left(\frac{d\sigma_V}{dy}\right)^{-1} \frac{1}{N_C} \sum_{ij} \frac{8Q}{s} \left[f_i(\bar{x}_1) f_j(\bar{x}_2) + f_i(\bar{x}_2) f_j(\bar{x}_1) \right] \frac{16\pi\Gamma_V^2 \mathrm{BR}_{V\to q_i\overline{q}_j} \mathrm{B}_{V\to\ell\ell'}}{(Q^2 - M_V^2)^2 + M_V^2\Gamma_V^2} \\ &\equiv \sum_{ij} C_{ij}^V \left[f_i(\bar{x}_1) f_j(\bar{x}_2) + f_i(\bar{x}_2) f_j(\bar{x}_1) \right] \frac{Q}{(Q^2 - M_V^2)^2 + M_V^2\Gamma_V^2}, \end{aligned} \tag{3.3}$$

where $\bar{x}_{1,2} = \sqrt{\tau}e^{\pm y}$ and the constants C_{ij}^V include terms that depend on y but not on Q. The fact that $\Gamma_V/M_V \ll 1$ and that f_i are smooth functions in the relevant range of Bjorken x values ($10^{-3} \lesssim x \lesssim 10^{-1}$) can be exploited to perform a Taylor expansion of the right-hand side of Eq. 3.3 around $Q = M_V$:

$$\frac{d\sigma_V}{dQ}(Q \mid y) \sim \frac{Q}{(Q^2 - M_V^2)^2 + M_V^2\Gamma_V^2} \sum_{ij} |V_{ij}|^2 \left(F^{ij} + F^{ji}\right) \times \tag{3.4}$$

$$\left[1 + \underbrace{\frac{\sum_{ij} |V_{ij}|^2 \left(F^{ij}H^{ij} + F^{ji}H^{ji}\right)}{\sum_{ij} |V_{ij}|^2 \left(F^{ij} + F^{ji}\right)}}_{H_V} \left(\frac{Q}{M_V} - 1\right) + \underbrace{\frac{\sum_{ij} |V_{ij}|^2 \left(F^{ij}K^{ij} + F^{ji}K^{ji}\right)}{\sum_{ij} |V_{ij}|^2 \left(F^{ij} + F^{ji}\right)}}_{K_V} \left(\frac{Q}{M_V} - 1\right)^2\right]$$

where the flavour-dependent terms have been factored out of $\mathrm{BR}_{V\to q_i\bar{q}_j}$ in the form of the square of the V matrix elements. The latter should be interpreted as the usual CKM matrix for the case of W production, and as $\left(T_i^3 - 2Q_i \sin^2\theta_W\right)\delta_{ij}$ for Z production, where T_i^3 and Q_i are the weak isospin and electric charge of quark i, respectively. In Eq. 3.4, the following auxiliary functions of y have been introduced:

$$\begin{aligned} F^{ij} &= \left[f_i(\bar{x}_1) f_j(\bar{x}_2)\right]_{Q=M_V} \\ H^{ij} &= \left[\frac{f_i'(\bar{x}_1)}{f_i(\bar{x}_1)}\bar{x}_1 + \frac{f_j'(\bar{x}_2)}{f_j(\bar{x}_2)}\bar{x}_2\right]_{Q=M_V} \\ K^{ij} &= \frac{1}{2}\left[\frac{f_i''(\bar{x}_1)}{f_i(\bar{x}_1)}\bar{x}_1^2 + \frac{f_j''(\bar{x}_2)}{f_j(\bar{x}_2)}\bar{x}_2^2 + 2\bar{x}_1\bar{x}_2\frac{f_i'(\bar{x}_1)f_j'(\bar{x}_2)}{f_i(\bar{x}_1)f_j(\bar{x}_2)}\right]_{Q=M_V} \end{aligned} \tag{3.5}$$

where f' (f'') are the first (second) order derivatives[1] of the PDFs with respect to x. The constants H_V and K_V defined in Eq. 3.4 represent the appropriate average of the auxiliary functions of Eq. 3.5 over the flavour space. The validity of the Taylor expansion of Eq. 3.4 has been assessed by comparing the lineshape from Eq. 3.3 and 3.4 at different values of y. The relative difference between the two is found to be below 0.5% for Q in a range of ± 2 GeV around M_V.

In Eq. 3.4, the contribution of the PDFs to the lineshape is fully encoded in the constants H_V and K_V. The mode Q_0 of the lineshape can be readily calculated from Eq. 3.4:

$$Q_0 \approx M_V - \frac{\Gamma_V^2(H_V + 1)M_V}{2\left[\Gamma_V^2(H_V + K_V) - 4M_V^2\right]} \approx M_V + \frac{\Gamma_V^2}{8M_V}(H_V + 1), \tag{3.6}$$

where the approximation $\Gamma_V^2(H_V + K_V) \ll 4M_V^2$ can be justified *a posteriori*. The quantity

$$\Delta_V \equiv \frac{\Gamma_V^2}{8M_V}(H_V + 1) \tag{3.7}$$

represents the displacement of the mode Q_0 from M_V. Part of it is simply due to the Jacobian factor from the transformation of Eq. 3.2, and does not depend on the PDFs. The right-hand side of Eq. 3.6 depends on K_V only at higher order in Γ_V/M_V because it enters as the coefficient of a quadratic correction to the Breit-Wigner functions, which is symmetric around M_V. Both H_V and K_V are functions of y and M_V, albeit the dependence on the latter is negligible in the range of experimental uncertainty on M_Z (~ 2 MeV) and M_W (~ 12 MeV) compared to the PDF uncertainties on Δ_V.

[1] All the derivatives are evaluated as $f' = \frac{1}{\Delta x}\left[f(x + \Delta x/2) - f(x - \Delta x/2)\right]$, with a step of $\Delta x = 10^{-5}$. This step has been varied in $\Delta x \in [10^{-6}, 10^{-4}]$ and the result is found to be stable in this range.

The shift Δ_V determined from Eq. 3.7 is plotted in Fig. 3.2 as a function of the boson rapidity y for Z and $W^{\pm}$ production. The error bars correspond to the RMS of the distribution obtained by sampling 100 replicas in the `NNPDF30_nlo_nf_5_pdfas` set from the LHAPDF library [131]. A comparison with `NNPDF31_nlo_pdfas`, which includes Drell-Yan measurements from the 8 TeV run of the LHC [132], has been performed as well, showing consistent results both in the central value and in the uncertainty. Numerical values are reported in Table 3.1 for three representative values of y. A negative shift with typical size $|\Delta_V| \sim 13$ MeV is observed in the central region $|y| \lesssim 3$, steeply increasing at larger rapidity values. This behaviour can be understood qualitatively in terms of the valence quark density xu_V and xd_V, which are typical benchmarks in PDF fits [120]. Indeed, for any derivable and positive-definite function f, it holds that

$$\frac{xf'}{f} = \frac{1}{f}(xf)' - 1. \tag{3.8}$$

The left-hand side of Eq. 3.8 is of the same form of the terms that appear in the definition of H_V (see the second line of Eq. 3.5). The valence quark densities feature a local maximum at $x \sim 10^{-1}$, which corresponds to $|y| \sim 3$ at $Q \sim 90$ GeV. By identifying f in Eq. 3.8 with u_V or d_V one can easily see that the terms $(xf)'/f$ vanish around $|y| \sim 3$, thus giving the smallest shift, whereas they steeply decrease at higher rapidity values since $f \to 0$ and $(xf)'$ becomes negative. The relative PDF uncertainty on Δ_V is found to be in the $\tilde{5}\%$ ballpark, ranging from 0.3 MeV at $|y| \sim 0$ to 1 MeV at $|y| \sim 3.5$.

A further inspection of Fig. 3.2 shows that in the case of the Z boson, the shift happens to lie between the shifts for the W boson of opposite charge, as an effect of the different partonic favours probed by the gauge bosons.

Finally, the approximation which leads to Eq. 3.7 has been numerically validated. The resulting K_V values range between 5 and 15, depending on y, with a relative uncertainty below 3%.

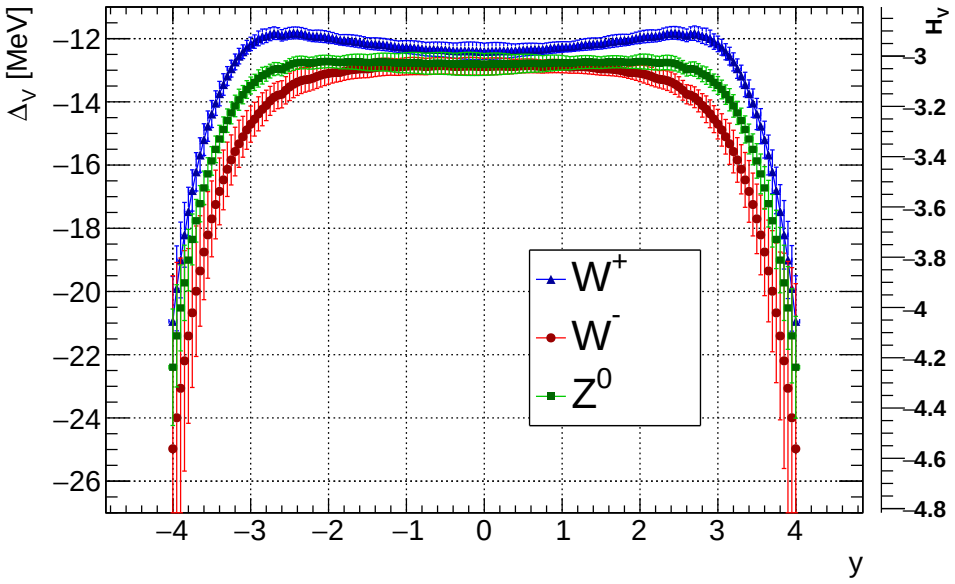


Figure 3.2. *The shift Δ_V and the coefficient H_V as a function of y averaged over the various quark flavours that enter the tree-level production of $W^{\pm}$ and Z in pp collisions at $\sqrt{s} = 13$ TeV. On the right the equivalent scale for the correspondent H_V.*

Table 3.1. *Numerical values of the shift Δ_V in pp collision at $\sqrt{s} = 13$ TeV, for three selected values of $|y|$ with their PDF uncertainty.*

$\lvert y\rvert$	Z [MeV]	W^+ [MeV]	W^- [MeV]
0.0	-12.8 ± 0.3	-12.4 ± 0.3	-12.9 ± 0.3
2.0	-12.7 ± 0.2	-11.9 ± 0.2	-13.1 ± 0.3
3.5	-15.9 ± 0.5	-14.4 ± 0.5	-17.7 ± 1.1

3.3 Monte Carlo simulation study

The tree-level calculation of Sec. 3.2 has been validated by using a MC simulation of $pp \to V + X$ production. Besides corroborating the tree level model, the MC analysis will also allow us to extend the study to the full phase-space, which includes the contribution of other diagrams. Given the similarity between neutral- and charged-current Drell-Yan production, and the observation that W boson production, split by charge, can serve as a good proxy also for the Z boson (see Fig. 3.2), and given the much larger MC sample available, the analysis has been restricted hereafter to the special case $V = W^{\pm}$ About 8×10^7 events in the final state $W^{\pm} \to \mu^{\pm}\nu_{\mu}$ have been generated using the `MG5_aMC@NLO` [119] program interfaced with `Pythia8` [121] [2] The dilepton mass is reconstructed using the muon momentum before QED final state radiation. The MC simulation is NLO accurate for observables inclusive in additional QCD radiation and it assumes $M_W^{\mathrm{MC}} = 80.419$ GeV and $\Gamma_W^{\mathrm{MC}} = 2.047$ GeV. As already anticipated in Sec. 3.2, the tree-level prediction is expected to be reproduced in the limit $q_{\mathrm{T}} \to 0$, where q_{T} is the transverse momentum of the boson. Indeed, in this regime the relative contribution of the tree-level $2 \to 1$ diagrams, which produce the boson at rest in the transverse plane, is enhanced compared to higher-order $2 \to 2$ diagrams. In contrast, a reduction of the shift in the large q_{T} region is expected, where gluon-initiated diagrams dominate, thus reducing the sensitivity of the boson virtuality on the partonic luminosity.

3.3.1 Fit to the MC sample

The analytical study of Sec. 3.2 shows that the shift in the MC sample has to be extracted from a statistical analysis of the dilepton mass distribution $d\sigma_W^{\mathrm{MC}}/dQ$. A crucial part of this task is to choose the correct functional form for $d\sigma_W^{\mathrm{MC}}/dQ$, capable of modelling the lineshape without introducing a bias in the estimator of Δ_V. Motivated by the tree-level study, an *ansatz* function of the same form of Eq. 3.4 has been chosen:

$$\frac{d\sigma_V^{\mathrm{MC}}}{dQ}(Q \,|\, y) = A \frac{Q^{\alpha}}{(Q^2 - M^2)^2 + M^2\Gamma^2} \left[1 + H\left(\frac{Q}{M} - 1\right) + K\left(\frac{Q}{M} - 1\right)^2\right]. \tag{3.9}$$

The choice $\alpha = 1$ defines the baseline function, which will be referred to as the *modified Breit-Wigner*. Indeed, this functional form, which is identical to Eq. 3.4 from the tree-level study, will be explicitly validated by checking that the estimator of M_W and Γ_W is consistent with the input values of the MC simulation M_W^{MC} and Γ_W^{MC}. As further validation of this choice, two alternative instances of the parametric family of functions in Eq. 3.9 has been considered. The first is obtained by the choice $\alpha = H = K = 0$, which reduces to a Breit-Wigner. This function is formally incorrect to model the dilepton mass distribution since it does not account

[2] Note that, despite the similar generator configuration the sample adopted here is different from the one described in Chapter 5.

for the Jacobian factor proportional to Q. However, it is a useful benchmark since it is symmetric around $Q = M_W$ so that the mass estimator also matches the mode. The second alternative function is obtained by choosing $\alpha = 1$ and $H = K = 0$. This function, which will be referred to as *Breit-Wigner with Jacobian*, would be correct in the absence of the PDF distortion to the Q distribution. It peaks at $Q \approx M_W + \frac{\Gamma^2}{8M_W}$, which is always larger than M_W.

Three statistical analyses of the simulated events have been performed. The first analysis is inclusive in the phase-space of the W boson and allows us to benchmark the different fit functions with the largest possible statistical precision. The second analysis is differential in the boson rapidity and is expected to reproduce, at least qualitatively, the y-dependence from the tree-level model, as shown in Fig. 3.2. However, the comparison can only be approximate, since the latter predicts the transverse momentum q_T to be identically zero, whereas the MC simulation generates a physical spectrum of transverse momenta. The third analysis is performed in bins of q_T, and is inclusive in y. It allows us to both validate the tree-level calculation of Sec. 3.2 by extrapolating to $q_T \to 0$, and to study the dilution effect at larger values of q_T. The fit parameters of Eq. 3.9 are determined by minimizing a χ^2 constructed using the event counts in each bin of the histogram in the range $[79, 82]$ GeV and the value of the fit function at the center of the bin.

Figure 3.3 shows the result of the three functional fits for the inclusive analysis for the W^+ sample. The W^- sample shows very similar results for both the inclusive and the differential analyses, which are thus omitted for brevity. The quality of the fit improves dramatically when using the modified Breit-Wigner, with a reduced χ^2 of about 1.0 compared to 1.6 and 3.9 for the alternative functions. The best-fit value of M_W when using the baseline function is consistent

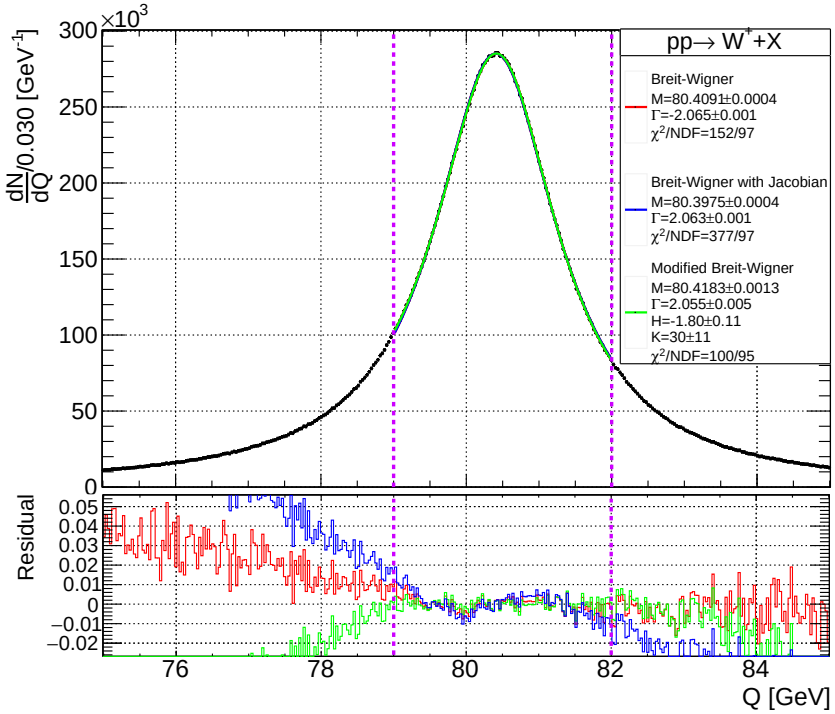
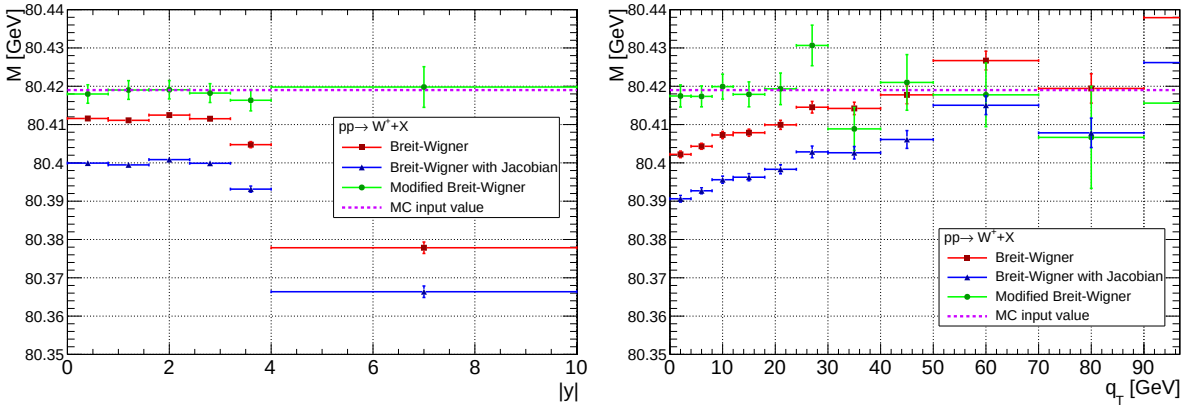


Figure 3.3. *The dilepton mass distribution for the inclusive W^+ sample. The result of the fit using the Breit-Wigner (red), Breit-Wigner with Jacobian (blue), and modified Breit Wigner (green) are superimposed to the distributions. In the bottom pad, the residuals between the fitted function and the histogram are shown. Only events in the range $[79, 82]$ GeV (marked by the vertical dashed lines) are used in the fit. A similar result is obtained for the W^- sample.*

with the MC input within 1σ (1 ± 1 MeV), whereas the alternative functions depart from it by -9.9 ± 0.4 MeV and -21.5 ± 0.4 MeV, respectively. Likewise, the best-fit value of Γ_W is consistent with the MC input values within 1.5σ (8 ± 5 MeV) for the baseline function, while it departs from it by 18 ± 1 MeV and 16 ± 1 MeV for the alternative functions. The uncertainty is larger in the first case because of the correlation with additional parameters (H, K). The residual discrepancy on Γ_W when using the modified Breit-Wigner is ascribed to higher order terms in the power expansion, not included in Eq. 3.9. In order to validate this assumption, a toy MC has been used to verify that such a discrepancy is indeed consistent with the neglected terms that mostly enter through the tails of the distribution, whereas M, H and K are seen to be robust.

The best-fit values of M_W from the differential analysis in the W boson rapidity are reported in Fig. 3.4(a) for the W^+ sample. The Breit-Wigner fit underestimates M_W all over the rapidity spectrum, as also observed in the inclusive analysis. The same applies to the Breit-Wigner with Jacobian function. For the latter, the discrepancy is even more pronounced. Indeed, the Jacobian factor contributes via a positive bias to the peak position. By neglecting the PDF term, which pulls in the opposite direction, the estimator of M_W is thus shifted to even lower values compared to M_W. The modified Breit-Wigner function correctly reproduces the input value M_W^{MC} in all bins of $|y|$, including the high $|y|$ regimes, where the alternative functions perform rather poorly.

Finally, the results of the analysis differential in the W boson transverse momentum are shown in Fig. 3.4(b). The modified Breit-Wigner is seen to correctly reproduce the input mass value for all bins of q_{T}, whereas the two alternative functions disagree, especially at low transverse momenta. For q_{T} in excess of about 40 GeV the statistical error is too large to discriminate among the models.



(a) (b)

Figure 3.4. *The best-fit value of M_W using the Breit-Wigner (red), Breit-Wigner with Jacobian (blue), and modified Breit Wigner (green), in bins of $|y|$ (a) and q_{T} (b), for the simulated W^+ sample. The dotted line corresponds to the input value of M_W^{MC}. A similar result is obtained for the W^- sample.*

3.3.2 Extraction of Δ_W

Since the fit reproduces well the true values for M_W and Γ_W, we fix the values of M and Γ to the MC input values in Eq. 3.6 and repeat the fit with A, H, K as the only free parameters.

For the inclusive sample (Fig. 3.3) the value of the shift is found to be:

$$\text{(Full phase-space)} \qquad \begin{aligned} \Delta_{W^+} &= -5.4 \pm 0.2\,(\text{stat.})\ \text{MeV} \pm 0.1\,(\text{PDF})\ \text{MeV}, \\ \Delta_{W^-} &= -5.8 \pm 0.2\,(\text{stat.})\ \text{MeV} \pm 0.1\,(\text{PDF})\ \text{MeV}. \end{aligned} \tag{3.10}$$

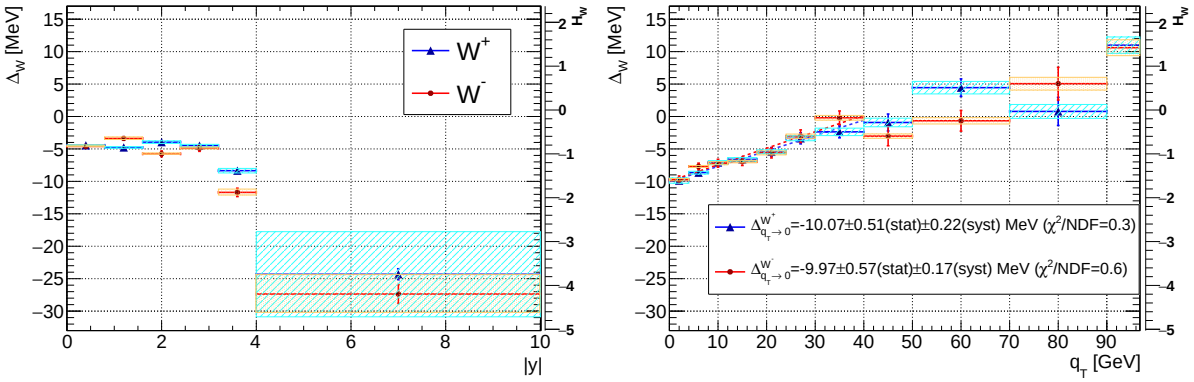
The first uncertainty is statistical-only while the second is the estimation of the systematic uncertainty from the PDFs. The latter is estimated from the RMS of the shifts determined using the first 100 replicas, as described in Sec. 3.2. In these fits the parameters M and Γ have been left free since the uncertainty on the PDFs would be otherwise over-constrained by the imposed knowledge on the mass and the width of the resonance.

The fitted values of Δ_W for the differential analyses are shown in Fig. 3.5 in bins of $|y|$ and q_{T}, separately for W^+ and W^-. The variation of Δ_W with the boson rapidity is shown in Fig. 3.5(a). It agrees well with the tree-level expectation of a flat shift in the central rapidity region followed by a rapid decrease at larger rapidity values. However, the shift in the central region is found to be smaller by a factor of about two, like for the inclusive results. Such a difference has been interpreted as the result of the dilution from higher-order diagrams. Indeed, in the limit $q_{\mathrm{T}} \to 0$, the measured shift gets closer to the tree-level result as shown by Fig. 3.5(b), while it vanishes for q_{T} in excess of about 40 GeV. A simple linear extrapolation to $q_{\mathrm{T}} \to 0$ yields limiting values of:

$$(q_T \to 0 \text{ extrapolation}) \qquad \begin{aligned} \Delta_{W^+} &= -10.1 \pm 0.5\,(\text{stat.}) \pm 0.2\,(\text{PDF})\ \text{MeV}, \\ \Delta_{W^-} &= -10.0 \pm 0.6\,(\text{stat.}) \pm 0.2\,(\text{PDF})\ \text{MeV}. \end{aligned} \tag{3.11}$$

Although reasonably close to the tree-level expectation, this result still disagrees with it by roughly 30%. This residual difference is interpreted as a pure next-to-leading-order correction to the leading-order prediction, stemming from collinear gluon emission and from gluon-initiated diagrams which contribute to the small-q_{T} regime. The relative PDF uncertainty is found to agree reasonably well with the expectation from the tree-level model averaged over the W boson rapidity.

As a cross-check of this result, the fit has been repeated varying the range symmetrically by ±10%. The results for Δ_W are stable, with a maximum discrepancy of 5%, which is within the uncertainty of the parameter. The fit has been also repeated after changing the renormalization



(a) (b)

Figure 3.5. *The shift $\Delta_{W^\pm}$ in bins of the W boson rapidity y (a) and transverse momentum q_{T} (b). For the latter, a linear fit in the range $[0, 40]$ MeV is performed to extrapolate the result to $q_{\mathrm{T}} \to 0$. The shaded boxes correspond to the PDF systematic uncertainty, as described in the text. On the right side of each plot, the equivalent scale for the H_W parameter is reported.*

and factorization scales in the matrix elements of the MC simulation by factors of 0.5 and 2, respectively. The results are again found to be stable within the PDF uncertainty.

Due to the agreement between the MC studies above and the tree level model of Sec. 3.2, the results for Δ_W are expected to be valid also for the Z boson case with $\Delta_{W^-} < \Delta_Z < \Delta_{W^+}$ and a similar PDF-related uncertainty.

3.4 Conclusions

In this study, the impact of the PDFs on the lineshape of gauge bosons at the LHC has been investigated. Given the narrow width of the electroweak gauge bosons, the PDF impact can be treated, to a first approximation, as a shift Δ_V of the mode of the dilepton mass spectrum from the boson mass M_V. The origin of such a shift can be traced back to the dependence of the partonic luminosity on the virtuality Q of the gauge boson. It has been first studied analytically by using a tree level model of Drell-Yan production and then validated by a statistical analysis of a MC simulated sample. The tree-level calculation agrees reasonably well with the MC study in the phase-space where the two are expected to be comparable. The results of this study prove that the PDF uncertainty on Δ_V is below one MeV all over the phase space relevant for the muon scale calibration.

The muon momentum scale calibration has a relative precision $\sigma^{\mu}_{p_T}/p^{\mu}_T \sim \mathcal{O}(10^{-3} - 10^{-4})$. This produces an uncertainty at ∼MeV level, for the sake of m_W measurement. Therefore, the systematic uncertainty related to the PDFs impact on the Z lineshape is negligible compared to the typical systematic uncertainty of the calibration. Once the muon scale calibration will reach the 10^{-5} level, the PDFs impact on the Z lineshape should be considered.

Chapter 4

The W boson mass measurement

The W boson has been discovered almost forty years ago. Since then several techniques have been developed to measure the W boson production properties, adapting them to the different experimental environments. A historical overview of the discovery and the measurement at LEP is presented in the first part of this chapter. Then, the state-of-the-art is summarized with a description of the traditional measurement strategy at hadron colliders, from the legacy of Tevatron and ATLAS work. The limiting systematic uncertainties will be highlighted and a discussion about the proposed improvements is reported. Finally, the CMS efforts pursued in the last years and the ongoing activities to reach the target precision of 10 MeV are presented in the final part of the chapter.

4.1 The discovery at S$p\overline{p}$S

The model of the electroweak interaction, including the prediction of the W boson, was developed in the '60s, with a major contribution from Glashow, Weinberg and Salam [133–135]. The $W^{\pm}$ and Z bosons were experimentally observed at SPS collider with the experiments UA1 and UA2, which announced the discovery in 1983 [136, 137]. The SPS [138] was initially designed to run as a proton accelerator to provide a high energy proton beam to fixed target experiments, and it was capable to accelerate protons up to 450 GeV. Since the center-of-mass energy of SPS was too low to significantly produce W and Z bosons, the SPS was modified to run as a proton-antiproton collider (S$p\overline{p}$S), with the two beams sharing the same beampipe. The S$p\overline{p}$S was designed to meet the requirement of $Z \to ee$ production and it operated with $\sqrt{s} \sim 500-600$ GeV and $\mathcal{L} \sim 10^{27-29}$ cm^{-2}s^{-1}.

The W discovery paper of UA1 (UA2) reported the observation of 6 (4) events with large transverse energy produced by an isolated electron and large missing energy angularly separated from the electron. They have been interpreted as a $q\overline{q}' \to W^{\pm}$ production and consequent $W^{\pm} \to e^{\pm}\nu_e$ decay. With this assumption, m_W has been estimated through a fit to transverse mass m_T distribution, with additional constraints from p_T^e distribution and some theoretical guidelines. The first results were provided with an uncertainty of 5-10 GeV. The final results of 1985 with a statistic of 290 (251) W candidates for UA1 (UA2) collaboration were [139]:

$$m_W = 82.7 \pm 1_{\text{stat}} \pm 2.7_{\text{syst}} \text{ GeV (UA1)} \qquad m_W = 80.2 \pm 0.8_{\text{stat}} \pm 1.3_{\text{syst}} \text{ GeV (UA2)}, \quad (4.1)$$

where the first error is the statistical one and the second arises from the calorimeter energy scale uncertainty. An example of the m_T distribution used to fit m_W by UA1 is reported in Fig. 4.2(a).

4.2 The W mass measurement at LEP

LEP was an electron-positron collider whose center-of-mass energy ranged from 90 to 200 GeV. Its purpose was the study of the electroweak sector with unprecedented precision. The first run at lower energy was focused on physics close to the Z peak, and the LEP results still represent the state-of-the-art in this sector. The second run of LEP scanned higher energies ($\sqrt{s} = 130 - 209$ GeV) and was used to further explore the electroweak sector and set exclusion limits on Higgs boson mass. The W boson mass and properties were investigated through the $e^+e^- \to W^+W^-$ production by the four LEP experiments (ALEPH [140], OPAL [141], DELPHI [142], L3 [143]). The two tree level diagrams which contribute to this process are shown in Fig.4.1.

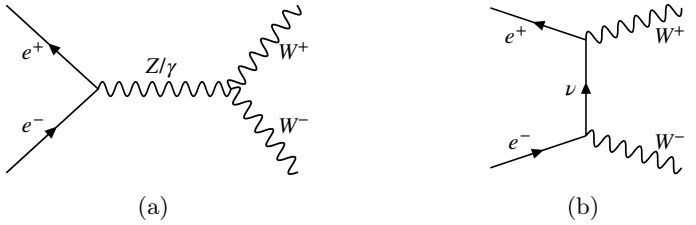


Figure 4.1. *Tree level Feynman diagrams for WW production at e^+e^- colliders: s-channel with neural vector boson propagator (a), t-channel with neutrino propagator (b).*

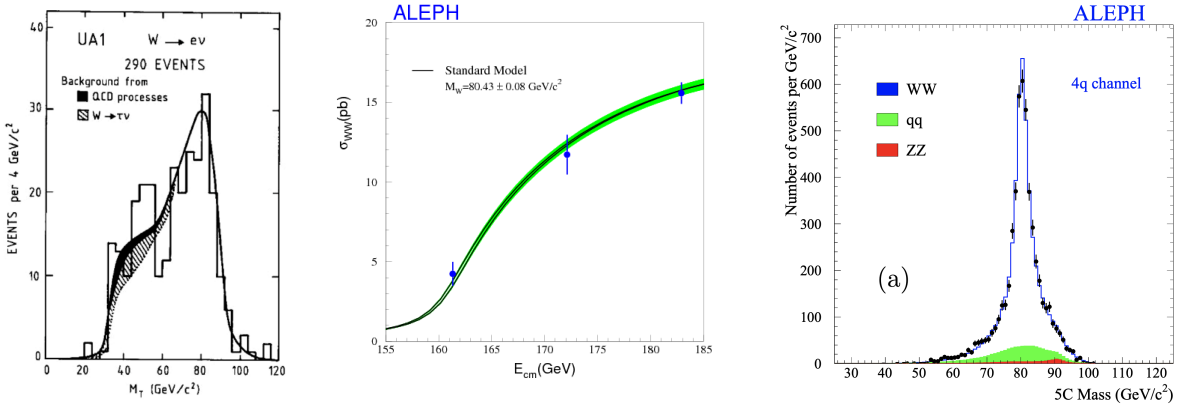
Two methods have been used to measure m_W at LEP. The first approach, also called m_W^{thr} measurement, exploited the threshold behaviour of the $e^+e^- \to W^+W^-$ production cross section σ_{WW} at $\sqrt{s} = 2M_W$, fitting the shape of the cross section as a function of $\sqrt{s}$. Close to the threshold the $\sigma_{WW} \propto \beta = \sqrt{1 - 4m_W^2/s}$, neglecting the finite width of the boson Γ_W and radiative correction. Thus, the mass can be extracted from the shape. The effect of Γ_W, higher order QCD and electroweak corrections and interaction between final states have been taken into account as 2% of theoretical uncertainty on the σ_{WW}. However only a small fraction of LEP data have been collected close to the threshold energy, and therefore the measurement has been statistically limited. The combination of the m_W^{thr} measurements from the four LEP experiments assessed 200 MeV of experimental uncertainty (including statistical and about 70 MeV of systematical uncertainty) and 30 MeV of uncertainty related to LEP beam energy, using an integrated luminosity of about 10 pb^{-1} for each experiment. The $\sigma_{WW}(\sqrt{s})$ measurement performed by ALEPH experiment is reported in Fig. 4.2(b) to show the turn-on behaviour.

In the second approach, m_W has been directly fitted from the invariant mass distribution of $WW \to q\bar{q}q\bar{q}$ or $WW \to q\bar{q}l\bar{\nu}$. In the former process, four hadronic jets have been reconstructed in the final state. The well know initial state can be used as a constraint on the jets four momenta, and the measurement does not rely on the jet energy scale to reconstruct the invariant mass, but only on the resolution in jet momentum direction. In the semileptonic channel, the charged lepton momentum scale and resolution are measured with high precision, while the momentum of the neutrino is inferred from the conservation of the four-momentum of the initial state. An integrated luminosity of about 750 pb^{-1} for each experiment was available for this measurement, which allowed to strongly reduce the statistical uncertainty. The combined m_W^{direct} measurements of the four experiments assessed 25 MeV of statistical uncertainty. The combined measurement included also 22 MeV of systematic uncertainties, which arose from hadronization and colour reconnection effect between the final state, detector effects, LEP beam energy uncertainty, initial and final state radiation. In particular, the hadronization and reconnection degraded the impact of the purely hadronic channel. An example of the invariant mass fit is shown in Fig. 4.2(c).

The LEP W boson mass measurement, combining threshold and direct measurement is:

$$m_W^{\mathrm{LEP}} = 80.376 \pm 0.025_{\mathrm{stat}} \pm 0.022_{\mathrm{syst}}\ \mathrm{GeV}, \tag{4.2}$$

with a total uncertainty 33 MeV. In this result, m_W^{thr} produces only a minor improvement, due to the limited statistic used. The complete discussion of this combination can be found at Ref. [144].



(a) (b) (c)

Figure 4.2. *(a): The m_T distribution of $W \to e\nu$ events collected by UA1 between 1982 and 1985, with shaded areas for background estimation and solid line for the signal fit result (from Ref. [139]). (b): Cross section as a function of $\sqrt{s}$ from the events collected by ALEPH, compared to the SM prediction with 2% of theoretical uncertainty (green band) and ±80 MeV of m_W uncertainty (solid lines); the world average m_W value is used as central value (from Ref. [145]). (c): Mass distribution from $WW \to q\bar{q}q\bar{q}$ channel from the events collected by ALEPH, with the background as shaded area and signal+background MC blue line, encoding the fitted value of m_W (from Ref. [146]).*

4.3 The W mass measurement at hadron colliders

The collision environment of modern hadron colliders is substantially different from the LEP. This leads to a radically different approach to W boson mass measurement.

First, in hadron-hadron collisions at center-of-mass $\sqrt{s}$, the energy of initial state $\sqrt{\hat{s}} = \sqrt{x_1 x_2 s}$ is not directly accessible. The momentum fraction of the colliding protons x_1 and x_2 change on event by event basis and are described by the PDFs of the incoming hadrons. Therefore the energy scale calibration must be provided by the detector i.e. by the final state information, and not by the beam energy calibration as at LEP. The procedure has limited precision and introduces a systematic uncertainty in the mass extraction.

Secondly, the electroweak processes represent only a small fraction of the total hadron-hadron collision cross section. The most common process is instead the strong interaction between partons, which reflects at detector level in the hadronic jet production. These interactions are a huge background that must be properly suppressed to analyse the electroweak events of interest. Moreover, the additional processes which occur beside the hard interaction between partons (PU-vertices, MIB, underlying event, as described in Sec 2.1.1) produce additional particles, which make the precise and accurate reconstruction of electroweak events more complex.

On the other hand, the higher luminosity of hadron colliders allows collecting much larger samples. If the higher luminosity is coupled with an efficient and low background trigger and reconstruction software this results in higher statistical power in the electroweak observables extraction.

4.3.1 Analysis strategy

The W boson has a branching ratio $\mathrm{BR}_{W\to q\overline{q}'} \simeq 67\%$, and a $\mathrm{BR}_{W\to \ell\overline{\nu}_\ell} \simeq 10.8\%$ for each lepton family. Therefore, the hadronic decay channel seems to be favoured from a statistical point of view. Unfortunately, this decay channel is affected by all the disadvantages described in previous paragraphs. The $W \to q\overline{q}'$ decays result in di-jet signature, with the invariant mass of the two jets peaking at m_W value, with a momentum of about 40 GeV each. This final state has a large background from multi-jet hadronic production, which does not allow to properly trigger and identify the W resonance. This effect can be mitigated by choosing events where the W boson is produced with high transverse momentum ($p_T^{\mathrm{jet}} \gtrsim 300 - 500$ GeV). In this case, the boson resonance results in a boosted topology, where the two jets are merged in one single *fat jet* in which the substructure of the two original jets can be identified. The p_T^{jet} selection strongly suppresses the multi-jet background, but also considerably reduces the available statistic events spoiling the advantage of the hadronic channel. Moreover, the energy scale and resolution for di-jet events is not sufficient to allow accurate extraction of m_W from the di-jet invariant mass spectrum. In particular, the hadronic jet scale calibration is at per cent level, while a $10^{-3} - 10^{-4}$ energy scale calibration is needed to be competitive with LEP measurement, and the poor resolution does not allow to distinguish between Z and W resonances. Currently, the hadronic channel is therefore a not feasible approach to precisely measure m_W at hadron colliders. However, some studies for future measurements have been performed. For instance, the possibility to reach about 30 MeV precision on $m_W - m_Z$ observable in boosted di-jet topology is discussed in Ref. [147], assuming an integrated luminosity of about 3 ab^{-1} and making some assumptions on the future improvement of the theoretical and experimental source of systematic uncertainties. The purpose of this approach is to measure m_W with a set of systematic uncertainties mostly orthogonal to the ones which affects the m_W measurement using leptonic decays.

The $W \to l\nu$ channel, restricting to $l = e, \mu$, produces a very clear signature with an energetic isolated lepton and a considerable amount of missing transverse momentum due to the escaping neutrino. The $W \to \tau\nu$ channel is not considered, due to the multiple neutrinos in the final state (plus hadrons, in case of semileptonic τ decays), which introduce large uncertainties in the reconstruction.

The background processes are due to energetic leptons from heavy flavour decays, misidentified leptons or electroweak decays with similar signatures, like $Z \to \ell\ell$ with one of the leptons not reconstructed. A high signal-over-noise-ratio (~ 10) can be achieved with a proper kinematic selection that suppresses these background sources.

Due to the presence of the neutrino in the final state the kinematic of the event cannot be closed, and the invariant mass of the W cannot be reconstructed on event by event basis. Nevertheless is still possible to exploit the conservation of the momentum on the transverse plane and extract the m_W information from transverse variables only.

At first order in perturbation theory the W boson is produced with no transverse momentum, via the upper-left Feynman diagram of Fig. 3.1 only. In this simplified case, the production and decay chain is: $q\overline{q}' \to W \to \ell\overline{\nu}$, and proceeds through a single s-channel Feynman diagram. With this assumption the lepton transverse momentum p_T^ℓ depends on the η of the decay only, and reaches its maximum value at $p_T^\ell(\eta = 0) = m_W/2$. A jacobian peak emerges in the p_T^ℓ distribution, correspondent to $m_W/2$, as shown in Fig. 4.4(a), and m_W can be extracted from this feature. The p_T^ℓ distribution has a small tail at $p_T^\ell > m_W/2$ due to the W boson finite width.

Taking into account NLO correction, the W boson is produced associated with gluons or in gluon Compton scattering, as shown in upper-right, lower-left and lower-right Feynman diagram of Fig. 3.1. These production modes provide a finite transverse momentum (q_T^W) to W boson.

The q_T^W spectrum at LHC is reported in Fig. 4.3(a), obtained from the CMS simulation. The most probable value is at $q_T^W \simeq 4$ GeV, but the spectrum has a long tail which reaches tens of GeV setting the mean at $\simeq 20$ GeV. The q_T^W spectrum produces a smearing in the p_T^ℓ distribution spoiling the jacobian peak, as shown in Fig. 4.4(a). The q_T^W spectrum is not known with the required precision (both from a theoretical and experimental side) to be able to precisely unfold the original peak. Therefore additional variables sensitive to m_W must be used.

The momentum conservation on the transverse plane, considering the non-vanishing q_T^W and expressed in term of final state objects, implies:

$$\mathbf{q}_T^W = \mathbf{p}_T^\ell + \mathbf{p}_T^\nu \equiv -\mathbf{h}, \tag{4.3}$$

where $\mathbf{h}$ is defined as the transverse plane projection of the recoil of the rest of the event, which balances the q_T^W on the transverse plane, while p_T^ν is the neutrino transverse momentum. Experimentally the entire missing transverse momentum of the event is assigned to the neutrino from W boson decay, thus $\mathbf{p}_T^\nu = \mathbf{p}_T^{\text{miss}}$. Note that the recoil $\mathbf{h}$ and $\mathbf{p}_T^\ell$ are measured independently, while the $\mathbf{p}_T^{\text{miss}}$ and thus $\mathbf{p}_T^\nu$ are inferred from the relation 4.3. The longitudinal component of the neutrino momentum is not accessible, since the longitudinal momentum of the initial state partons is unknown.

With the use of the recoil the transverse mass m_T variable can be built. The W boson mass can be expressed as:

$$m_W = \sqrt{2p_T^\ell p_T^\nu\big(\cosh(\Delta\eta) - \cos(\Delta\phi)\big)},$$

where $\Delta\eta$ and $\Delta\phi$ represent the difference in pseudorapidity and angle on the transverse plane between the lepton and the neutrino. A projection of the previous relation on the transverse plane ($\eta = 0$) leads to the definition of transverse mass of the W boson:

$$m_T = \sqrt{2p_T^\ell, p_T^\nu\big(1 - \cos(\Delta\phi)\big)}. \tag{4.4}$$

Since p_T^ν is measured from the recoil, m_T can be also expressed as:

$$m_T = \sqrt{2\big(p_T^\ell|\mathbf{p}_T^\ell + \mathbf{h}| + (p_T^\ell)^2 + \mathbf{p}_T^\ell \cdot \mathbf{h}\big)}. \tag{4.5}$$

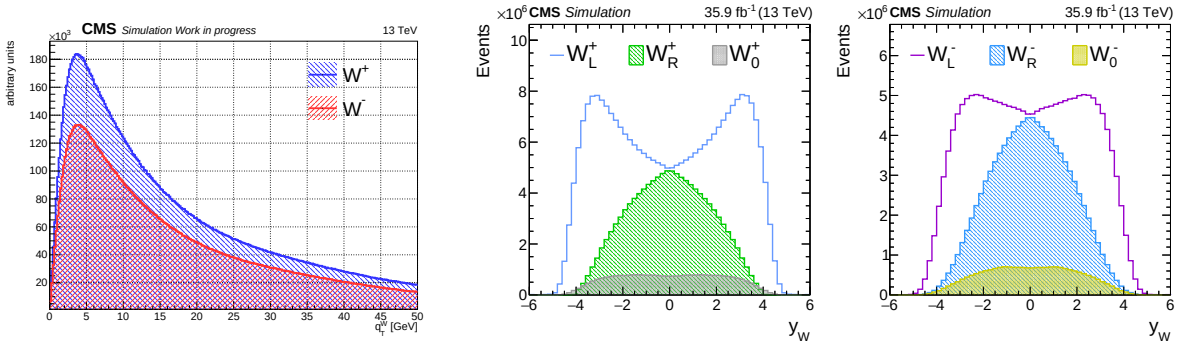
The distribution of m_T for $W \to \mu\nu$ events is shown in Fig 4.4(b), comparing the vanishing q_T^W case with the realistic one. The m_T distribution has an endpoint at m_W value, and $m_T = 2p_T^\ell$ in the $q_T^W = 0$ limit . Like the p_T^ℓ also the m_T distribution has a small tail at $m_T^\mu > m_W$ due to the W boson finite width. The m_T distribution has essentially no dependence from the W transverse momentum compared to p_T^ℓ. The jacobian peak at $m_T = m_W$ is not spoiled by q_T^W, allowing to extract the mass value from m_T distribution. This is justified since p_T^ℓ depends linearly on the W velocity on the transverse plane β_T, while m_T has only β_T^2 dependence.

Until now the distributions that have been considered are at generator level, without taking into account the experimental reconstruction effects. The lepton energy resolution (typically few per cent) and the lepton energy scale (calibrated at $10^{-3} - 10^{-4}$ level) affect only mildly the lepton p_T^ℓ distribution, and the above description holds. On the other hand, m_T uses the recoil as a central ingredient. Given the typical q_T^W spectrum, the recoil is far from the jet regime and consist of low energy hadrons. The estimation of q_T^W using h has limited accuracy and is affected by large uncertainties. The reconstruction effects on m_T are large, as shown in Fig 4.4(b), and the jacobian peak is also spoiled in this variable.

Since the invariant mass is not accessible on event by event basis, and the related distributions like m_T and p_T^μ, do not offer a clear handle to extract m_W, the *template fit* approach is typically

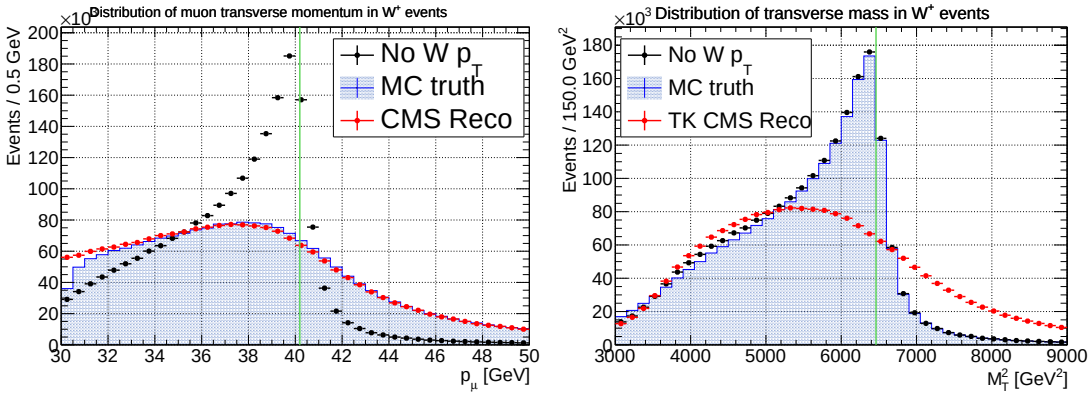
used at hadron colliders. It consists of the simulation of a large number of events to reproduce the relevant distribution (m_T and p_T^ℓ) with the same reconstruction procedure used on real data. Several replicas of these distributions are reproduced starting with different m_W hypothesis, called templates. Then, the distribution measured from real data is compared to these templates performing a likelihood scan. m_W is extracted interpolating the templates to identify the value of m_W which maximizes the agreement with data, combining the p_T^ℓ and m_T distribution information.

This approach strongly relies on an accurate description of the template variables in the simulation. A disagreement between data and simulation reflects in a modification of the shapes of the templates and thus a systematic bias to the value of the mass extracted from the fit. The calibration of the simulation is therefore a central ingredient of the analysis and is often performed using data-driven methods. A common procedure is to use Z boson events to precisely calibrate the simulation. These aspects will be discussed with more detail in Sec. 4.3.3, 4.3.4, for Tevatron and ATLAS m_W measurements, presenting the different choices. In the next section an overview of systematic uncertainties affecting the m_W measurement with the presented method is given.



(a) (b) (c)

Figure 4.3. *Simulated W boson transverse momentum (a) and rapidity distribution (b,c, from Ref. [11]), separated in the three helicity states. The CMS experiment simulation for $\sqrt{s} = 13$ TeV LHC pp collisions has been used.*



(a) (b)

Figure 4.4. *Simulated p_T^μ (a) and m_T^2 (b) distributions in case of vanishing q_T^W, realistic q_T^W spectrum but exact knowledge of recoil and p_T^μ, and after the reconstruction effects. The green line correspond to the value of m_W, located at jacobian peak edge before the smearing due to q_T^W or reconstruction effects. The The CMS experiment simulation for $\sqrt{s} = 13$ TeV LHC pp collisions has been used (from Ref. [9, 10]).*

4.3.2 Systematic uncertainties

The m_W measurement at hadron colliders has a large statistic available and a high signal to background ratio. Therefore the systematic uncertainties are typically the limiting factor to reach the target precision. They can be divided in two main categories. First, the theoretical uncertainties derive from a lack of knowledge in the description of the W boson production mechanism (q_T^W, Y_W, polarization) and higher order correction (mainly FSR), and they reflect in a limited accuracy of the hard processes encoded in the simulation. Second, the experimental uncertainties, which arise from non-perfect calibration of the physical objects used in the extraction of m_W (the recoil and the lepton momentum) and from the background estimation. The summary of the systematic uncertainties for Tevatron and ATLAS m_W measurements is reported in table 4.1.

Table 4.1. *Summary of systematic uncertainties for m_W measurement at hadron collider. All the values are reported in MeV (from [107, 148]).*

Experiment	CDF		ATLAS	
Source	p_T^ℓ	m_T	p_T^ℓ	m_T
Statistical	16	15	7.2	9.6
W transverse momentum	9	3	8.3	9.6
PDFs (W rapidity, polarizazion)	9	10	9	10.2
Higher order corrections	4	4	5.7	3.4
Lepton momentum calibration (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

4.3.2.1 W production mechanism (rapidity, transverse momentum, polarization)

The modelling of the W boson production represents one of the most important systematic uncertainties for the m_W measurement. It is discussed here in term of the factorized distribution of transverse momentum q_T^W, rapidity Y_W and W boson polarization, since this is the main approach followed by the previous m_W measurements at hadron colliders.

The q_T^W spectrum emerges beyond the tree level description of the W production cross section, as previously discussed. The source of q_T^W are the low-energy products of hadronization, which form the recoil. The theoretical description in this low-q_T regime cannot rely on fixed-order perturbative calculation, due to the presence of logarithmic terms $\sim \ln(m_W/q_T^W)$. They must be dealt with resummation techniques [149, 150], which can treat this kinematic region, and they are affected by large uncertainties. Currently, the state-of-the-art theoretical prediction is N^{3}LL+NNLO, where the fist term indicates the third order from the leading logarithmic term in resummation calculation. The result from Ref. [151] is shown in Fig. 4.5. The uncertainty ranges between 6% and 4% for $q_T^W < 5$ GeV, and at is $\sim 2\%$ at higher q_T^W.

Since the Z boson transverse momentum (q_T^Z) measurement can be performed with higher precision, it is typically used as an indirect measurement of q_T^W, with a proper extrapolation to W boson phase space. The q_T^Z measurement can be also used to validate the calculation of the theoretical prediction. In both cases, large uncertainties arise from the porting from Z to W spectrum. The explicit case of ATLAS and CDF will be discussed in the next sections. For the sake of m_W measurement, the q_T^W uncertainties mainly affect the p_T^ℓ distribution.

The W boson rapidity distribution Y_W induces a systematic uncertainty on m_W because of the finite acceptance of the measurement. An angular selection on the accepted events induces a

cutoff in the Y_W distribution. This selection sculpts the p_T^ℓ and m_T distributions used for the m_W extraction, as described in details in Ref. [7]. The Y_W distribution is strongly dependent on the polarization of the W boson which also differs for the two boson charges. This effect is shown in Fig. 4.3 as the Y_W distribution for the two charges, separated in the left, right and longitudinal helicity states, where the helicity of the W boson is the projection of the W boson spin in the direction of the W momentum. The polarization is defined from the momentum of the initial state partons i.e. from the PDFs of the proton. In summary, the lack of knowledge on the PDFs distributions induces an uncertainty on Y_W and polarization, which can affect the p_T^ℓ and m_T distributions in case of finite detection acceptance. In the (unrealistic) case of complete Y_W acceptance the p_T^ℓ and m_T systematic uncertainty from the PDFs will be negligible.

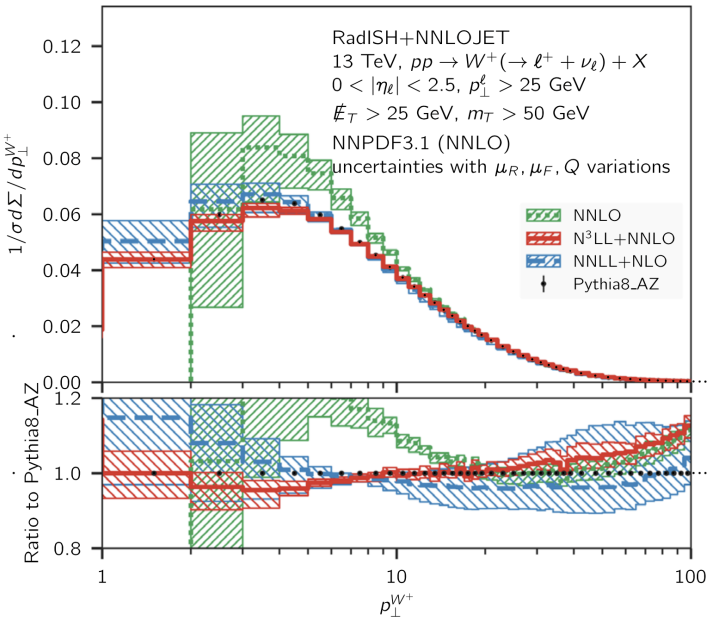


Figure 4.5. *Comparison of the normalized q_T^W distribution for W^+ at $\sqrt{13}$ TeV at NNLO, NNLL+NLO and N^{3}LL+NNLO, and the* `PYTHIA AZ tune`*. The fiducial selection $p_T^\ell > 25$ GeV, $p_T^{miss} > 25$ GeV, $|\eta^\ell| < 2.5$, $m_T > 50$ GeV is applied (from Ref. [151]).*

4.3.2.2 Higher order corrections

The uncertainty in modelling higher order QED and QCD correction directly reflects in the simulation used to build the template for the m_W extraction. The dominant correction is the QED final state radiation (FSR), but it can be carefully simulated limiting its uncertainty. Other sources of uncertainties are the pure weak corrections (virtual loop and box diagrams), the final state emission of lepton pairs ($\gamma^* \to \ell^+\ell^-$) and the interference between the initial state radiation (ISR) and the FSR.

4.3.2.3 Lepton momentum calibration

The p_T^ℓ is the main reconstructed object for the m_W extraction and should be carefully calibrated to reduce the systematic uncertainty to m_W. Any discrepancy in p_T^ℓ between the data and the simulation directly reflects in the templates, and thus produces a bias in m_W.

The main aspect is the p_T^ℓ scale calibration, which should be under control at 10^{-4} level to obtain an uncertainty below 10 MeV. The general strategies have been outlined in Sec. 2.3.3.1

and Sec. 2.3.3.2. The approach will be discussed in the specific case of the Tevatron and ATLAS measurements, differently for electron and muons, in the next sections.

The p_T^ℓ resolution must be also corrected to properly match the resolution on data and the one on the simulation. A discrepancy in the resolutions modifies the shapes of the quantity of interest for the p_T^ℓ scale calibration differently in data and simulation. This may cause a bias in the calibration procedure and thus in the measurement.

Finally, in the same fashion, the lepton selection efficiency must be measured in data and simulation as a function of η^ℓ and p_T^ℓ and a proper correction must be applied to the simulation taking into account the observed discrepancies. This is typically performed with tag-and-probe efficiency measurement and subsequent extraction of Data/MC efficiency scale factors, as described in Sec. 2.3.3.4 and Sec. 2.3.3.5.

4.3.2.4 Recoil

The recoil is the central ingredient for an accurate description of the m_T variable. Any discrepancy between data and simulation in the QCD fragmentation and reconstruction of recoil particles reflects in a systematic uncertainty to m_T and thus m_W. Multiple effects contribute to the simulation of the recoil, like proper modelling of QCD radiation as well as the underlying event and the pile-up interactions and the simulation of the detector effects (acceptance, efficiencies, etc.).

The calibration of these effects is performed using Z events also in this case, directly comparing data and MC. Then, a proper extrapolation to W phase space must be done. The specific porting choices will be discussed in the Tevatron and ATLAS measurements.

4.3.2.5 Background estimation

The background contribution to $W \to \ell\nu$ reconstruction channel is below 10% level, but it must be properly modelled and subtracted to reach the target precision. The background events from electroweak processes which mimic the W decays can be directly subtracted estimating their yield from the simulation, because the simulation of the electroweak process guarantees sufficient precision. These processes include $Z \to \ell\ell$ decays where one lepton is not reconstructed, diboson events, top quark decays, $W \to \tau\nu$ decays.

Another background source is composed of the energetic leptons from multijet production mimicking the prompt lepton kinematics or leptons from a misidentified jet. These sources are called *QCD background* and cannot be estimated precisely from the simulation, due to the large uncertainties of the multijet production description. Data-driven methods are therefore usually exploited to obtain a robust estimation of the QCD background.

4.3.3 The W mass measurement at Tevatron

The Tevatron was the proton-antiproton collider, which operated at Fermilab (Chicago, IL) between 1983 and 2011. The Tevatron had a center-of-mass energy $\sqrt{s} = 1.96$ TeV and hosted the two experiments CDF and D0. The major discoveries of these experiments were the direct observation of top quark and the first observation of B_s oscillations.

The Tevatron experiments were able to perform the first high-statistic W boson mass measurement at hadron colliders. The combined measurement [152] of CDF and D0 collaborations exploits an integrated luminosity of 2.2 fb^{-1} from CDF and 4.3 fb^{-1} from D0. The CDF measurement [4] uses both the $W \to \mu\nu$ and $W \to e\nu$ decay channels, with about $4.7 \cdot 10^5$ and

$6.2 \cdot 10^5$ events, respectively. The D0 measurement [5] uses instead only the $16.8 \cdot 10^5$ $W \to e\nu$ events.

The measurement strategy was the template fit to m_W using p_T^ℓ and m_T distributions, as described in the previous sections. In addition, also the missing transverse momentum p_T^{miss} distribution is used in the same fitting framework. In the combination of the 3 quantities the most relevant was m_T, due its robustness against q_T^W (for CDF the relative weights were 0.53, 0.31 and 0.16 for m_T, p_T^ℓ and p_T^{miss}, respectively). An example of the data and fitted distribution is shown in Fig. 4.6(a).

In $p\overline{p}$ collisions the W^+ (W^-) is produced mainly from valence quark $u\overline{d}$ ($\overline{u}d$). Therefore the W^+ and W^- are produced in the same amount, but W^+ (W^-) is boosted on average in the p ($\overline{p}$) direction, because the u quark carries on average higher momentum compared to d quark. These effects produce two specular rapidity distributions for the two charges, not centered at $Y_W = 0$.

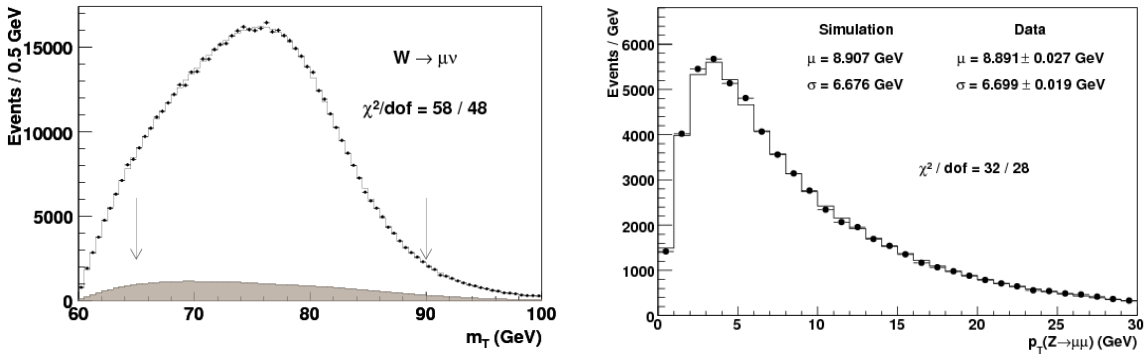
The Tevatron experiments assessed the rapidity/PDF uncertainty from the variation of the fitted m_W value, varying the PDF information encoded in the MC. The PDF set used in this procedure are `CTEQ6.6`,`CTEQ6.1` and `MSTW2008` [58, 153, 154]. For what concerns the q_T^W distribution, the Tevatron experiments fit the measured q_T^Z distribution with a resummed calculation prediction obtaining a disagreement below 2%, as shown in Fig. 4.6(b). Then, they exploited the same function to predict the q_T^W, with a proper propagation of the uncertainties of the model.

The D0 experiment calibrated the lepton momentum scale on $Z \to ee$ sample, while CDF used the J/ψ and $\Upsilon \to \mu\mu$ decays to calibrate the central tracker scale and then to port the calibration to the calorimeter using E/p ratio in $W \to e\nu$ decays. Furthermore, CDF used Z events to cross-check the model. In both cases, the limiting factor was the size of the Z, J/ψ and Υ samples.

The Tevatron experiments estimated the recoil distribution using the calorimeter information. First, they calibrated the recoil model on Z data. Then, they propagated the recoil model uncertainty on m_W, extracting the value of the mass for different recoil hypothesis within the calibration uncertainty.

The combined result from the various CDF and D0 measurements is:

$$m_W = 80.387 \pm 0.016 \text{ GeV}. \tag{4.6}$$



(a) (b)

Figure 4.6. *(a): Distribution of m_T reconstructed from $W \to \mu\nu$ decays for data and simulation including the fitted maximum likelihood value of m_W with the arrows indicate the fit range. (b): Distribution of q_T^Z in data and simulation, fitted to model the q_T^W spectrum and assess the respective systematic uncertainty. Both figure from CDF measurement, using the 2.2 fb^{-1} of $W \to \mu\nu$ events collected in Tevatron runs (from Ref. [4])*

Additional discussions on the combination are reported in Ref. [152].

4.3.4 The W mass measurement at ATLAS

The ATLAS experiment performed a measurement of m_W using a luminosity 4.6 fb^{-1} from the Run 1 of LHC, with $\sqrt{s} = 7$ TeV. The sample consists in $7.8 \cdot 10^6$ $W \to \mu\nu$ and $5.9 \cdot 10^6$ $W \to e\nu$ candidates. Like the Tevatron experiments, the ATLAS collaboration pursued the usual strategy, performing a template fit to m_W using p_T^ℓ and m_T distributions. Thanks to the much larger available statistics the fit has been performed subdividing the sample in $|\eta^\ell|$ categories.

In the pp collisions the W^+ is produced mainly from a u valence quark and a $\overline{d}$ sea quark, while the W^- is produced from a valence d quark and sea $\overline{u}$ quark. Because of the proton composition, W^+ and W^- yield and the q_T^W distribution are expected to be different between the two charges, differently from the Tevatron case. The Y_W distribution is expected to be symmetric, due to the symmetry of the two beams. Because of the higher $\sqrt{s}$ of the collisions of LHC, the fraction of heavy quark initiated collision is higher (approximately 25% of the inclusive cross section is initiated from the second generation) compared to the Tevatron case (5%), with a strong impact on the q_T^W and Y_W distributions.

To model the production mechanism the differential cross section has been described factorized in four terms:

$$\frac{d\sigma}{dp_1 dp_2} = \left[\frac{d\sigma(m)}{dm}\right] \times \left[\frac{d\sigma(Y)}{dY}\right] \times \left[\frac{d\sigma(q_T, Y)}{dq_T dY}\left(\frac{d\sigma(Y)}{dY}\right)^{-1}\right] \times \left[(1+\cos^2\theta + \sum_{i=0}^{7} A_i(q_T, Y) P_i(\cos\theta, \phi)\right], \tag{4.7}$$

where p_1, p_2 are the four momenta of the ℓ, ν, while the rest of the variables follow the same nomenclature of Eq. 0.1. The first term has been modelled as a relativistic Breit-Wigner, taking into account electroweak corrections. The Y_W distribution and the angular coefficients (second and fourth terms) have been described using fixed order perturbative QCD NNLO prediction as a function of q_T and Y of the boson. ATLAS adopted the `CT10nnlo` PDF set [155] to model the systematic uncertainties related to these therms. The PDF choice has been validated using the ATLAS measurements of rapidity (Z and W) and angular coefficients (Z only) from Ref. [156, 157], which show a satisfactory agreement with the NNLO prediction.

The q_T^W spectrum, the third term of Eq. 4.7, has been modelled using the `PYTHIA 8` Parton Shower with the `AZ tune` [158]. This tune has been calibrated on the measurement of the q_T^Z distribution [159], where the disagreement has been reduced below 1%. The choice has been validated directly on W boson data using the distribution of the recoil component parallel to lepton direction $u_{||}^\ell$, which is sensitive to q_T^W. A relevant distribution of this validation is shown in Fig. 4.7(a). ATLAS assessed several sources of systematic uncertainty in the porting q_T^Z to q_T^W, including the accuracy of Z measurement for the `AZ tune`, the mass of c and b quarks, the factorization scale of the parton shower (decorrelating the heavy quark variation and correlating the light quarks) and the parton shower PDF variation on the fixed-order prediction. The latter represents the leading uncertainty and has been applied simultaneously in Y_W, q_T^W and polarization, showing strong anti-correlation between the two boson charges. The `CTEQ6L1 LO` set [160] has been used in this context. Resummed predictions have been tried (`RESBOS`, `DYRes`, `POWHEG MiNLO+PYTHIA` [161–165]), but they predict harder q_T^W spectrum which shows a discrepancy not covered by the uncertainty assessed by ATLAS using the `AZ tune`. The discrepancies are shown in Fig. 4.7(b). These features have not been fully understood and since the same resummed calculations have been successfully exploited on Tevatron data, these discrepancies are still an open question. These issues will strongly benefit from a direct measurement of the q_T^W distribution, in particular in the low momentum region.

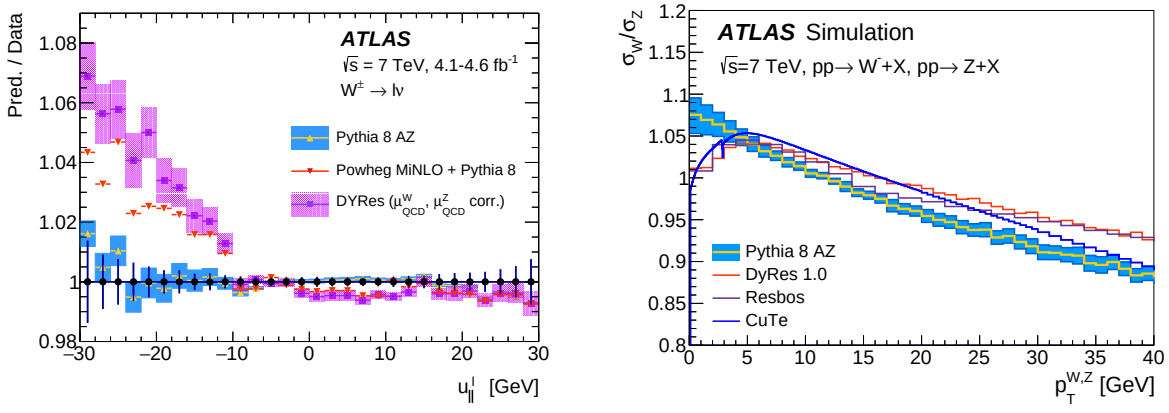
The recoil was reconstructed using the calorimeter information only, excluding the clusters associated with leptons. ATLAS calibrated the recoil in two steps. First, the overall event activity has been corrected separately for W and Z events, then the Z sample has been used to correct residual discrepancies between the simulation and data, as done in Tevatron measurements. The systematic uncertainties arose from both steps, and have been estimated by comparing the effects of different hypotheses in the calibration procedures. In the combined fit, the relative weight of m_T and p_T^ℓ were 0.14 and 0.86, respectively. Thus, contrary to the Tevatron case, at LHC the higher pile-up degrades the recoil scale and resolution and the improvement given by m_T is limited.

The lepton momentum scale and resolution has been calibrated using $Z \to \ell\ell$ events. The calibration and the efficiency scale factors estimation has been performed as a function of $1/p_T^\mu$, in bins of η, as described in Sec. 4.3.2.3. In addition, the biases in the sagitta reconstruction have been corrected directly on data, for the muon channel only. The main source of systematic uncertainty was the extrapolation from the Z phase space to the W, with a scale residual uncertainty between $5 \cdot 10^{-5}$ and $2 \cdot 10^{-4}$, depending on $|\eta|$.

The final result from ATLAS is:

$$m_W = 80370 \pm 7_{\text{stat}} \pm 11_{\text{exp.syst}} \pm 14_{\text{mod.syst}} \text{ MeV} = 80370 \pm 19 \text{ MeV}, \tag{4.8}$$

where the the first uncertainty corresponds to the statistical uncertainty, the second takes into account the experimental systematic uncertainties and the third one comes from the physics-modelling. More detail about the ATLAS measurement can be found in Ref. [3].



(a) (b)

Figure 4.7. *(a): Data-driven validation of the q_T^W modelling, as prediction over data ratio of $u_{||}^\ell$ for different simulation configuration, where the error bars of data represent the experimental uncertainty, while the simulation uncertainties reflect the q_T^W systematic uncertainty; only the* `PYTHIA 8 AZ tune` *shows good agreement in the entire spectrum, while a large disagreement is present for the resummed calculation the region correlated to high q_T^W. (b): Comparison of q_T^W distribution normalized to Z one, for* `PYTHIA AZ tune` *and several MC resummation programs, for W^- sample; the uncertainty of* `PYTHIA` *takes into account μ_F, m_c and LO PDF; a clear discrepancy is observed below 5 GeV. Both figure from ATLAS measurement, using an integrated luminosity of 4.6 fb^{-1} from the LHC run 1 (from Ref. [3]).*

4.4 CMS ongoing activities and motivation for a W production properties measurement

The state-of-the-art of W boson mass measurement is summarized in Fig. 4.8, where the single measurements and their combination are reported. For additional comparisons about Tevatron and ATLAS measurement see, for instance, Ref. [148, 166].

Currently, the CMS collaboration is working toward a m_W measurement following the traditional strategy, like Tevatron experiments or ATLAS, using the data collected during the LHC Run 2. The ATLAS measurement reveals that the use of m_T brings mild advantages at LHC, given the high PU scenario. In Run 2 conditions, the effect is even more pronounced and the CMS measurement will be based completely on p_T^μ. Some additional details about the issue related to the use of the recoil and thus m_T can be found in Ref. [9, 10].

Multiple efforts are pursued by CMS to reduce the systematic uncertainty of the measurement and to reach the 10 MeV target precision. The next sections will outline the achieved results and ongoing activities.

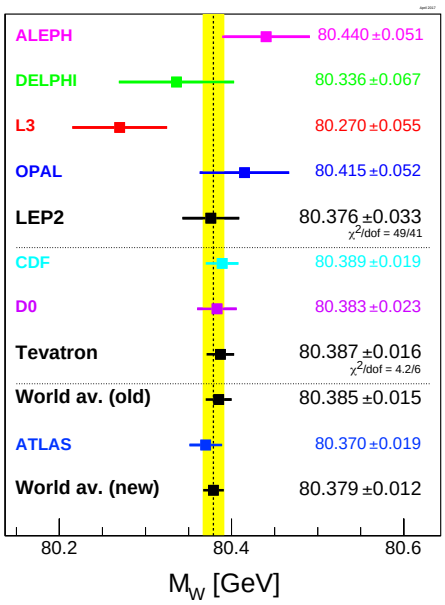


Figure 4.8. *Measurement of the W boson mass by the LEP and Tevatron experiments and ATLAS, performed by the Particle Data Group [18]. Note that the PDG reports an uncertainty slightly smaller than the GFitter result reported in Eq. 1.13, which uses a more conservative correlation scheme (the PDG assumed the PDF uncertainties fully correlated between Tevatron and LHC, the GFitter group did not. More details can be found in Ref. [18, 40]).*

4.4.1 The W transverse momentum and low-PU Run measurement

The modelling of the q_T^W has been one of the most delicate steps of the ATLAS measurement and despite the constant development from the theory side, the theoretical prediction cannot provide the required precision. The state-of-the-art prediction uncertainty, with combined resummation and perturbative techniques as reported in Fig 4.5, produces a distortion in the p_T^μ spectrum compatible with a m_W shift much larger than 10 MeV.

This can be understood from the Fig. 4.9. In this figure the nominal MC p_T^μ spectrum is compared to 3 relevant varied versions. In particular: the q_T^W spectrum has been shifted of 4%

in the $0 < q_T^W < 5$ GeV, compatible with the state-of-the-art uncertainty; the value of m_W has shifted of ± 10 MeV, the target precision; the PDFs have been varied according to their nominal uncertainty. All the variation are performed keeping the total normalization, which has no impact on m_W, constant. A 10 MeV variation on m_W reflects in a 0.05%-0.1% variation on the p_T^μ spectrum, while the q_T^W and PDF uncertainties produce larger variations. However, the relevant feature is the shape of the variation close to the jacobian peak, where the p_T^μ spectrum is more sensitive to m_W. The shift of the mean value of the p_T^μ spectrum, $\Delta\mu$, can be used as a proxy for the expected m_W uncertainty. The injected m_W variation produces $\Delta\mu = 3$ MeV, thus is reasonable to consider a scale factor of about 3 between this quantity and the mass uncertainty. In this framework the q_T^W limitation is manifest, producing a $\Delta\mu = 11$ MeV. The PDF uncertainty produces a shift $\Delta\mu = 14$ MeV. To validate this simple estimation, the ATLAS uncertainty on q_T^W (1%) has been injected in the p_T^μ spectrum. It results in a $\Delta\mu = 3$ MeV i.e. about 9 MeV uncertainty on m_W. This is consistent with the ATLAS result. However this estimation is strongly dependent of p_T^μ range used to assess the $\Delta\mu$. Thus, these values must be considered only an order of magnitude estimation to show how the production model can mimic a m_W shift.

In conclusion, a direct measurement of q_T^W is strongly required. This measurement must be able to achieve a 1-2% precision in the low q_T^W region, with a binning below 5 GeV, to significantly

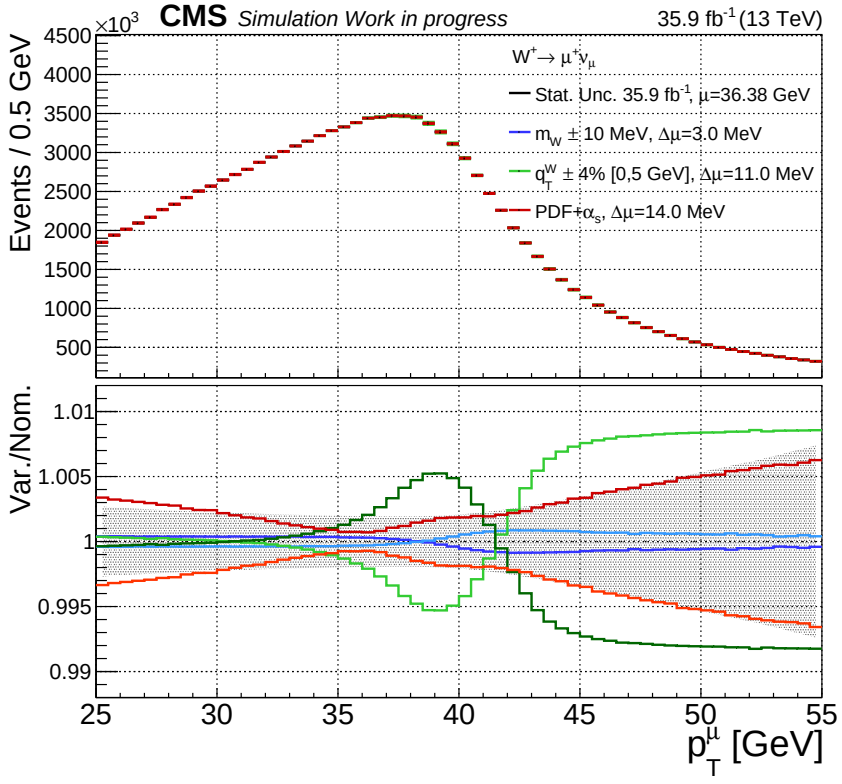


Figure 4.9. *The p_T^μ spectrum from $W^+ \to \mu^+\nu$ simulated events, the effects of a ± 10 MeV m_W variation, of a flat 4% variation of the q^W below 5 GeV and the variation of the PDF within the current uncertainty, symmetrized around the nominal value, are shown. All the variation are normalized to the nominal one, to highlight the shape discrepancy only. For each variation is also reported the absolute shift ($\Delta\mu$) in the mean value of the spectrum compared to the nominal one. The PDF + α_s value corresponds to the sum of squares of the shift induced by the 60 Hessian eigenvalues (which is larger than the shift induced by the sum of squares of the variations, shown graphically in the plot). The statistical uncertainty correspondent to a integrated luminosity of 35.9 fb^{-1} is shown as a gray band.*

improve the related systematic uncertainty on m_W. Both CMS [167] and ATLAS [168] performed a measurement of q_T^W with the LHC Run 1 data at $\sqrt{s} = 8$ TeV and 7 TeV respectively. In these measurements a precision of 2-3% in the region below 10 GeV is achieved, thanks to the low pile-up condition of Run 1 ($\langle\mu\rangle \simeq 2-4$ in the exploited samples). However, the region below $q_T^W = 7.5\,(8)$ GeV has been measured with a single bin in CMS (ATLAS) measurements.

LHC provided a low pile-up run during 2017, with an integrated luminosity of about 200 pb^{-1}, $\langle\mu\rangle \simeq 3$ and $\sqrt{s} = 13$ TeV [169]. These data will be exploited by CMS and ATLAS to perform the measurement of the q_T^W, taking advantage of the relatively clean environment to improve the precision of the analysis.

4.4.2 The W rapidity and helicity measurement at CMS

The measurement of the W rapidity and helicity distributions [11] has been pursued by CMS collaboration in the prospect of a future m_W measurement to reduce the PDF uncertainty. The PDF can be constrained by several measurements, like differential Drell-Yan cross section, charge asymmetries measurements, W boson and charm quark associated production [170–175]. The required precision has not been reached yet. An innovative method has been proposed in Ref. [8] to directly measure the W boson rapidity distribution with the capability to perform a constrain *in situ* of the PDF. This approach has been adopted by CMS in W rapidity and helicity measurement.

Let us first consider the simplified tree level case, where the W is produced without q_T^W. In this condition the W boson is produced uniquely via $q\bar{q}'$ in s-channel (upper-left Feynman diagram of Fig. 3.1), and the helicity of the W boson is determined by the partons spin and the partons relative longitudinal momentum p_z only. Due to the $V-A$ coupling of the weak interaction, the W spin is parallel to the quark spin, which is in the left helicity state. Typically $p_z^{\text{valence}} > p_z^{\text{sea}}$, thus the W flight direction will follow the quark one. Therefore most of W bosons show left polarization at LHC, independently of the boson charge. This behaviour leads to the rapidity distributions of Fig. 4.3(b) and Fig. 4.3(c), where the different helicity states are highlighted.

The immediate consequence of this polarization is in the distribution of the decay leptons. Considering the $W \to \ell\nu$ decays, again, because of the $V-A$ coupling, most of W^+ (W^-) tends to emit the charged lepton in the opposite (same) direction compared to the p_z^W direction. For W with right helicity, the behaviour is reverted. In the decay, the behaviour is different for the two boson charges because of the strong constraint on the neutrino helicity, while the polarized production depends only on the PDF. The scheme of the described process is summarized in the diagram of Fig. 4.10.

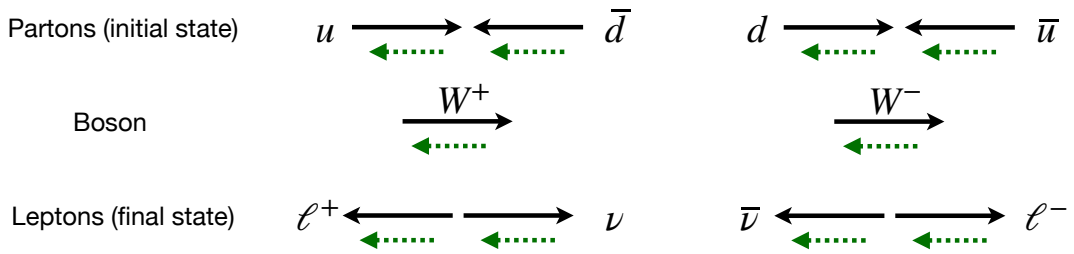


Figure 4.10. *Scheme of the momentum (p_z, black arrows) and spin (s_z, green dashed arrows) alignment at partons, boson and lepton final state, assuming the most common production and decay channel, at $q_T^W \simeq 0$, as described in the text.*

The introduction of a the realistic q_T^W spectrum produces the longitudinal polarization of W

boson, but for the experimental range of interest ($q_T^W \lesssim 30-50$ GeV) it remains subdominant as shown in Fig. 4.3, and the described scheme holds. The process can be described with the differential W boson production cross section:

$$\frac{d\sigma}{dq_T^W dY_W d\cos\theta^*} \propto \frac{3}{8}(1\mp\cos\theta^*)^2 f_L(q_T^W, Y_W) + \frac{3}{8}(1\pm\cos\theta^*)^2 f_R(q_T^W, Y_W) + \frac{3}{4}\sin^2\theta^* f_0(q_T^W, Y_W), \tag{4.9}$$

where θ^* is the polar decay angle of the lepton (charged for W^-, neutrino for W^+) in the dilepton system rest frame [12] and the upper (lower) sign correspond to W^+ (W^-) boson. The coefficients f_L, f_R, f_0 are the helicity fractions (left, right and longitudinal, respectively) which are defined to be $f_i > 0$ and $f_L + f_R + f_0 = 1$ and they are function of q_T^W and Y_W.

The different helicity states reflect in different, very characteristic distributions on the $\eta^\mu \times p_T^\mu$ plane, as shown in Fig. 4.11. This allows to perform a template fit to Y_W distribution in the three helicity states using templates in $\eta^\mu \times p_T^\mu$ distribution produced in bins of Y_W. The q_T^W dependence of the helicity fractions has been integrated and cannot be predicted with this fit.

The measurement has been performed on CMS data collected during 2016, equivalent to an integrated luminosity of 35.9 fb^{-1}, in $W \to \mu\nu$ and $W \to e\nu$ channels. This analysis had a double purpose for the m_W measurement. The fit is able to predict the Y_W and polarization distribution, which can be useful to check and improve the description of these variables encoded in the simulation. The result, in term of unpolarized and polarized Y_W spectrum, is reported in Fig. 4.12 and Fig. 4.13, respectively. In addition the analysis is able to set a direct constraint on the PDF set used to describe the systematic variation of the W rapidity distributions. This constraint can be exploited in a future m_W measurement.

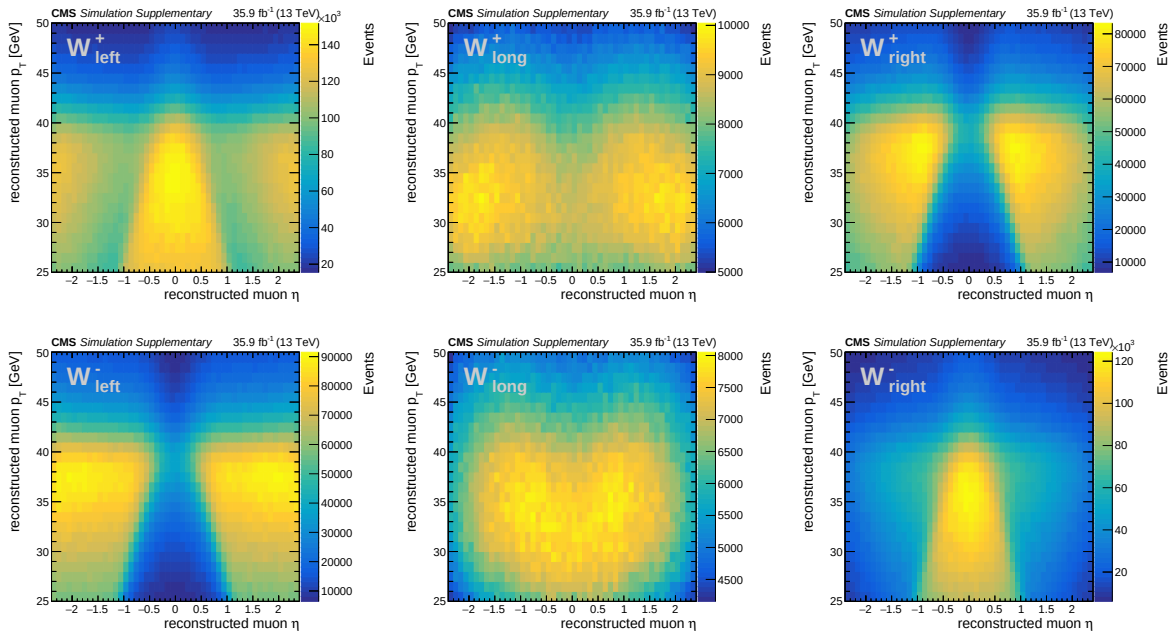
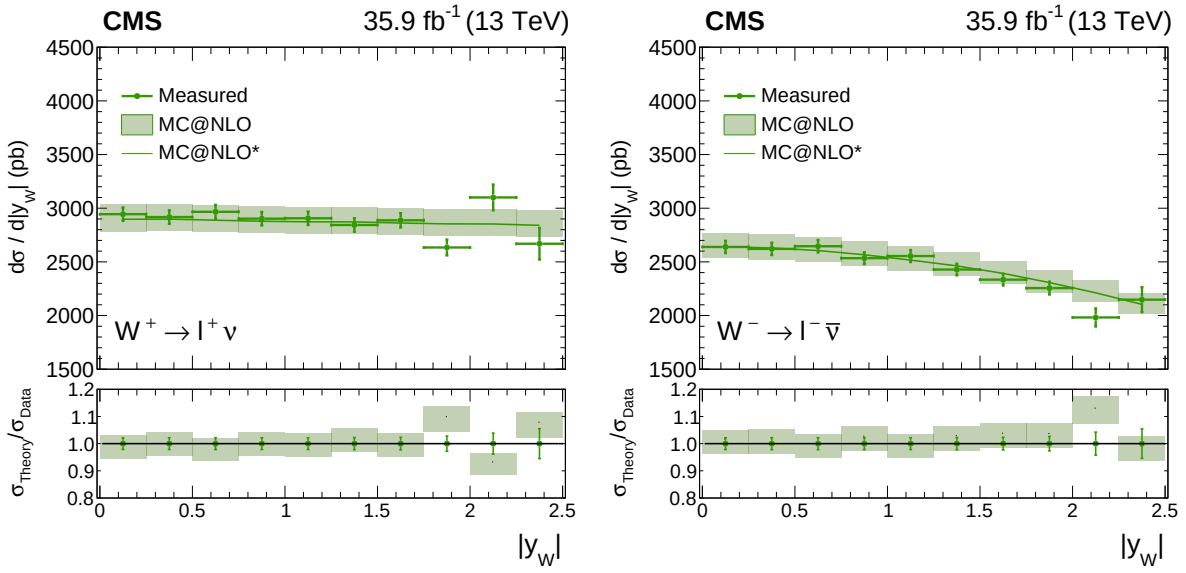
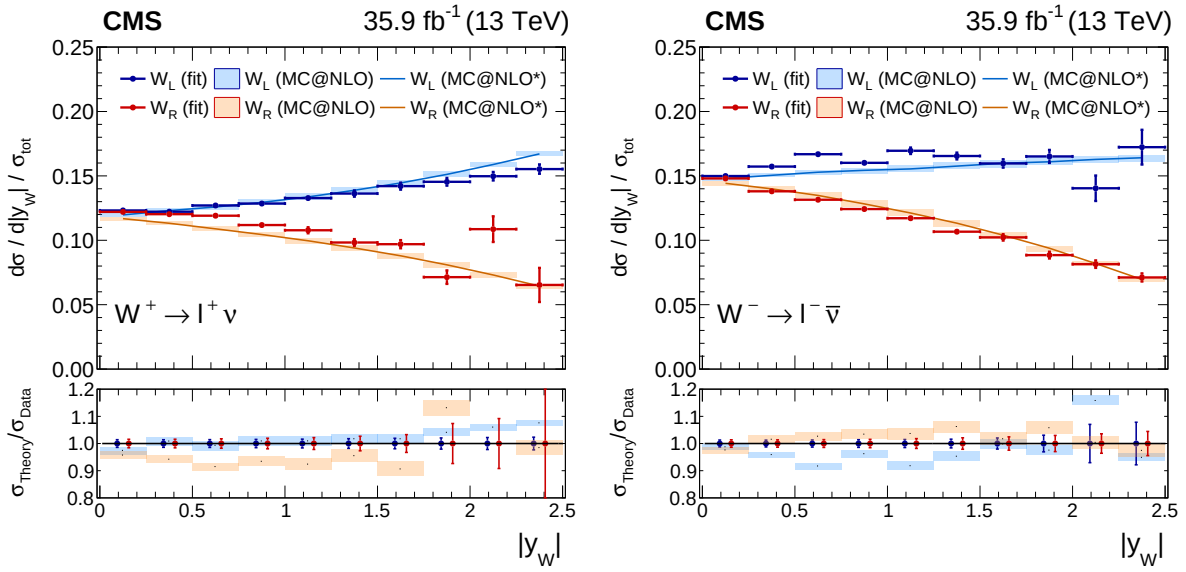


Figure 4.11. *Distributions of $p_T^\mu \times \eta^\mu$ from $W \to \mu\nu$ decays for W^+ (top) and W^- (bottom) in the left, longitudinal and right polarization states (from left to right). The η features of these templates can be qualitatively interpreted using $\eta^\mu \simeq Y_W + \eta^0$, where η^0 is the pseudorapidity of the muon in the W reference frame, with mode ± 0.5 (for $W^{\pm}$ or left/right). Events from CMS simulation and reconstruction are used in these distributions (From Ref. [105]).*



(a) (b)

Figure 4.12. *Measured absolute cross section as a function of Y_W, for $W^+ \to \ell^+\nu$ (a) and $W^- \to \ell^-\overline{\nu}$ (b), $\ell = \mu, e$. The prediction of the* `MadGraph5_aMC@NLO` *simulation is also shown as comparison. The simulation uncertainty include the PDF, α_S, MC μ_F and μ_R scales variation. The MC central value, where the q_T^W spectrum is weighted by the ratio of measured and predicted spectrum for Z production, is shown as a darker line (from Ref. [11]).*



(a) (b)

Figure 4.13. *Measured normalized cross section as a function of Y_W for left-handed and right-handed helicity states, for $W^+ \to \ell^+\nu$ (a) and $W^- \to \ell^-\overline{\nu}$ (b), $\ell = \mu, e$. The prediction of the* `MadGraph5_aMC@NLO` *simulation is also shown as comparison. The simulation uncertainty include the PDF, α_S, MC μ_F and μ_R scales variation. The MC central value, where the q_T^W spectrum is weighted by the ratio of measured and predicted spectrum for Z production, is shown as a darker line (from Ref. [11]).*

4.5 Conclusions

In the previous sections it has been outlined that the precision on muon scale calibration and the W production mechanism is the limiting factor to measure m_W with a precision below 10 MeV. For the former, an effective and robust method is currently under development by CMS to reach a relative precision of 10^{-4}. For the latter, more precise measurements are required. The precision on the q_T^W must be lowered to reach the 1-2% level also in the region below 5 GeV. The m_W measurement would also strongly benefit by a further reduction PDF uncertainty, and a factor 2-3 is desirable.

Chapter 5

The W boson mass and production properties analysis

In this chapter, an innovative method to fit simultaneously the W boson mass and production properties is presented. Then the details of the analysis implementation are described, in terms of exploited samples and adopted calibration and selection. Finally, the systematic uncertainties of the analysis are discussed.

5.1 Theoretical foundation

The Collins-Soper (CS) reference frame [176] is conveniently introduced to properly describe the W boson leptonic decays for the purposes of this analysis. The graphical scheme of the CS frame is reported in Fig. 5.1. The CS frame is defined as the di-lepton rest frame where the axes are aligned as follows:

- The $\hat{z}$ axis bisects the angle between $\mathbf{p}_{p_1}$ and $-\mathbf{p}_{p_2}$, pointing in the direction of q_z^W in the laboratory frame, where $\mathbf{p}_{p_1}$ and $\mathbf{p}_{p_2}$ are the momenta of the two colliding protons p_1 and p_2, and q_z^W is the component of the boson momentum aligned with the beam axis.

- The $\hat{x}$ axis is defined to be orthogonal to $\hat{z}$ and to lie in the colliding protons plane, pointing in the direction of q_T^W in the laboratory frame.

- The $\hat{y}$ axis is defined as the normal vector to the plane defined by the two colliding proton momenta, to form a right-handed Cartesian coordinate system.

The angles θ^* and ϕ^* are the polar and azimuthal angle of the charged lepton in the CS frame i.e. θ^* is the angle between the $\hat{z}$ axis and the charged lepton direction, while ϕ^* is the angle between the $\hat{x}$ axis and the lepton direction projected on the $\hat{x} - \hat{y}$ plane. The variables in the CS frame are indicated with the * symbol hereinafter.

The description will be restricted to $\ell = \mu$, since the analysis has been performed on the $W \to \mu\nu$ channel only, without any loss of generalities for the phenomenological discussion. Moreover, the lepton is considered before the emission of the final state radiation (a so-called *pre-FSR* lepton), if not specified otherwise.

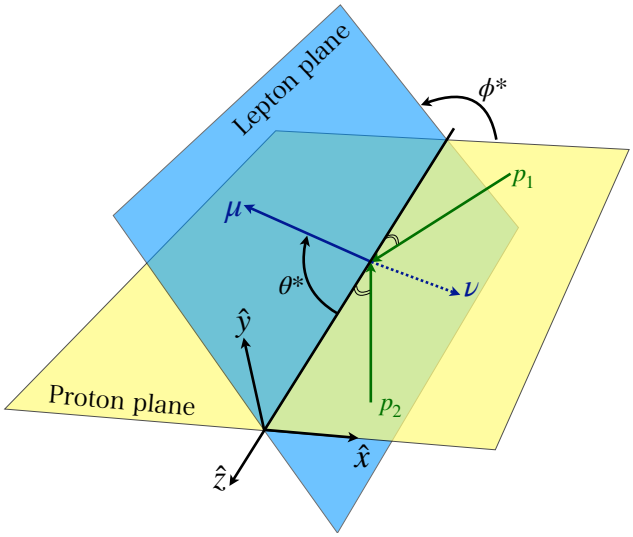


Figure 5.1. *Diagram of the Collins-Soper frame, in yellow the plane spanned by the protons momenta, in blue the plane spanned by the leptons momenta, highlighted the direction of protons and leptons on these planes, the ϕ^* and θ^* muon angles and the axes $\hat{x}, \hat{y}, \hat{z}$ of the frame.*

The differential production cross section of the W boson can be described using 8 angular coefficients [12]:

$$\frac{d\sigma}{dq^2_{T,W}dY_W d\cos\theta^*_\mu d\phi^*_\mu} = \frac{3}{16\pi}\frac{d\sigma^{U+L}}{dq^2_{T,W}dY_W}\Big[(1+\cos^2\theta^*_\mu) + \sum_{i=0}^{7} A_i P_i(\cos\theta^*_\mu, \phi^*_\mu)\Big]. \tag{5.1}$$

The $\cos\theta^*$, ϕ^* are the decay angles of the charged lepton in the CS frame (the index μ will be dropped without ambiguity hereinafter). The eight A_i are the angular coefficients, which depends of the W boson charge, Y_W and q_T^W. The σ^{U+L} denotes the production cross section for unpolarized W bosons. The P_i are angular functions which arise from the decomposition of the differential cross section in terms of spherical harmonics of degree 0, 1, 2:

$$\begin{aligned}
P_0 &= \frac{1}{2}(1-3\cos^2\theta^*), & P_3 &= \sin\theta^*\cos\phi^*, & P_6 &= \sin(2\theta^*)\sin\phi^*,\\
P_1 &= \sin(2\theta^*)\cos\phi^*, & P_4 &= \cos\theta^*, & P_7 &= \sin\theta^*\sin\phi^*,\\
P_2 &= \frac{1}{2}\sin^2\theta^*\cos(2\phi^*), & P_5 &= \sin^2\theta^*\sin(2\phi^*), & P_{U+L} &= 1+\cos^2\theta^*,
\end{aligned} \tag{5.2}$$

where the ninth term, P_{U+L}, has been included for completeness. This description arises from the vectorial nature of the W boson, which allows the decomposition in the aforementioned nine components. Another equivalent rewriting is in terms of helicity cross sections [12]:

$$\frac{d\sigma}{dq^2_{T,W}dY_W d\cos\theta^*_\mu d\phi^*_\mu} = \frac{3}{16\pi}\sum_\alpha \frac{d\sigma_\alpha}{dq^2_{T,W}dY_W} P_\alpha(\cos\theta^*_\mu, \phi^*_\mu), \tag{5.3}$$

with $\alpha \in \{U+L, L, I, T, A, P, 7, 8, 9\}$. The σ_α are defined from the following relations:

$$\begin{aligned}
A_0 &\equiv 2\frac{d\sigma^L}{d\sigma^{U+L}}, & A_3 &\equiv 4\sqrt{2}\frac{d\sigma^A}{d\sigma^{U+L}}, & A_6 &\equiv 4\sqrt{2}\frac{d\sigma^8}{d\sigma^{U+L}},\\
A_1 &\equiv 2\sqrt{2}\frac{d\sigma^I}{d\sigma^{U+L}}, & A_4 &\equiv 2\frac{d\sigma^P}{d\sigma^{U+L}}, & A_7 &\equiv 4\sqrt{2}\frac{d\sigma^9}{d\sigma^{U+L}},\\
A_2 &\equiv 4\frac{d\sigma^T}{d\sigma^{U+L}}, & A_5 &\equiv 2\frac{d\sigma^7}{d\sigma^{U+L}},
\end{aligned} \tag{5.4}$$

and the P_α are the polynomials of Eq. 5.2[1]. The latter description highlights the physical decomposition of the differential cross section as a sum of polarized cross sections. On the other hand, the angular coefficients have the advantage to be smooth functions of Y_W and q_T^W. The description of Eq. 5.1, in term of A_i, will be used in the rest of the analysis, unless specified otherwise.

The decomposition of Eq. 5.1 allows to isolate the kinematics of the W boson production from the decay kinematics. The production is completely encoded in $A_i(Y_W, q_T^W)$ and $\sigma^{U+L}(Y_W, q_T^W)$, while the decay kinematics is completely described by the P_i angular function. According to Eq. 5.1, given m_W, Y_W and q_T^W, the A_i are fixed and the momentum distribution of the lepton in the CS frame and in the laboratory is completely determined. The impact of the proton PDFs is encoded in the values of $A_i(Y_W, q_T^W)$ and $\sigma^{U+L}(Y_W, q_T^W)$, decoupled from lepton decay kinematics. Therefore, by measuring the A_i distributions it is possible to constraint the PDFs and their systematic impact on m_W measurement.

In the simplified case of $q_T^W = 0$, the only non-vanishing angular coefficient is A_4. If the W boson is produced with non-zero q_T^W, the eight A_i are generically different from zero. However at $\mathcal{O}(\alpha_s)$, only A_0, A_1, A_2, A_3 and A_4 are non-zero. In addition, the Lam-Tung relation holds [177], which implies $A_0 = A_2$. The A_5, A_6, A_7 are not-zero only taking into account of $\mathcal{O}(\alpha_s^2)$ gluon loops. Also in the full $\mathcal{O}(\alpha_s^2)$ calculation, the magnitude of A_5, A_6, A_7 remains much smaller than the $\mathcal{O}(\alpha_s)$ coefficients, apart from A_1 (whose magnitude is smaller than other $\mathcal{O}(\alpha_s)$ coefficients, and it is compatible with $\mathcal{O}(\alpha_s^2)$ terms). At $\mathcal{O}(\alpha_s^2)$ the Lam-Tung relation is broken and $A_0 \neq A_2$. Moreover the σ^i related to $\mathcal{O}(\alpha_s^2)$ coefficients are odd under parity and time reversal transformations, while the $\mathcal{O}(\alpha_s)$ terms are parity-even.

Note that the Eq. 4.9 can be recovered integrating Eq. 5.1 over ϕ^* with the relations:

$$f_L = \frac{1}{4}(2 - A_0 \mp A_4), \qquad f_R = \frac{1}{4}(2 - A_0 \pm A_4), \qquad f_0 = \frac{1}{2}A_0,$$

thus $f_L - f_R = \mp\frac{1}{2}A_4$ can be directly extracted, where the upper (lower) sign corresponds to W^+ (W^-) boson.

5.2 Analysis strategy

The correlation between the polarization of W boson and the consequent charged lepton kinematic distributions can be further exploited starting from the proof of concept of Ref. [8]. The A_i distributions and the unpolarized cross section, differential in Y_W, q_T^W can be extracted from the muon kinematics alone using the decomposition of Eq. 5.1. This is a natural extension of the W rapidity and helicity measurement of Ref. [11] with two major changes. First, the ϕ^*-dependence of the polarization has been taken into account, leading to the eight A_i plus σ^{U+L} description, instead of the 3 helicity fractions. Second, the q_T^W trend of the angular coefficients (and more in general the q_T^W spectrum via σ^{U+L}) has been included in the fit. In a non-vanishing q_T^W regime, the ϕ^*-dependence is important for a correct description of the acceptance of the analysis.

The measurement has been performed decomposing the inclusive $\eta^\mu \times p_T^\mu$ distribution in bins of $|Y_W|$, q_T^W and for each A_i. The resulting $(\eta^\mu \times p_T^\mu)(|Y_W|, q_T^W, A_i)$ distributions are called *templates*. The templates are produced using the CMS MC simulation. In this procedure the contribution of the background sources, which are present in the measured $\eta^\mu \times p_T^\mu$ distribution,

[1] The σ_α has been introduced with the index names found in literature. However from now on for clarity we refer to them with the same indices of the angular coefficients, given the bijective correspondence $A_i \leftrightarrow \sigma_\alpha$: $U+L \leftrightarrow U+L$, $0 \leftrightarrow L$, $1 \leftrightarrow I$, $2 \leftrightarrow T$, $3 \leftrightarrow A$, $4 \leftrightarrow P$, $5 \leftrightarrow 7$, $6 \leftrightarrow 8$, $7 \leftrightarrow 9$.

must be taken into account as well with additional templates. Once the templates have been produced, the $\eta^{\mu} \times p_T^{\mu}$ distribution measured on data has been fitted with the sum of signal and background templates. From the fitted values of the signal strength for each template is it possible to unfold the desired distributions, since the single template provides the information of the signal yield in the single bin of $|Y_W|$, q_T^W and A_i. The fit is able to extract A_0, A_1, A_2, A_3, A_4, and σ^{U+L}, but not the last three A_i, as discussed in Sec. 5.2.2.

5.2.1 Template building

The fit has been performed separately for W^+ and W^- channels since the expected distributions are different due to the different PDFs which modulate the production mechanism of the two boson charges. The Y_W distribution is symmetric with respect to $Y_W = 0$, due to the symmetry of the beams. Therefore the templates have been produced folding Y_W distribution into $|Y_W|$, and binning the latter.

The direct information of Y_W, q_T^W, and polarization distributions described in the simulation must not be used to construct the template distributions decomposed in proper bins. Such a choice would bias the extracted distribution with the information encoded in the simulation.

The strategy is to select simulated W events in a narrow bin of $|Y_W|$ and q_T^W, at generator level with full acceptance. Then the distribution $\cos\theta^* \times \phi^*$ is evaluated in CS frame. This distribution is reweighted to remove any angular dependence of muon from the boson matrix element, making it flat. After that, the distribution is reweighted again according to a specific angular function P_i to obtain the desired template, in the given bin of $|Y_W|$ and q_T^W. The easier approach to make the $\cos\theta^* \times \phi^*$ distribution flat would be to bin it and reweight each bin (i, j) by $1/N_{\text{events}}^{i,j}$. This approach is easy and robust, however it requires an explicit binning choice, which degrades the statistical power of the sample.

A more convenient solution to implement this strategy is to use the *momenta method* proposed in Ref. [178]. This approach is more complex, but avoids a binning in the $\cos\theta^* \times \phi^*$ plane, applying an event-by-event reweighing, and exploits the full statistical power of the sample. The moment of $f(\theta^*, \phi^*)$, integrated over a specific range of q_T^W and Y_W is defined as:

$$\langle f(\theta^*, \phi^*) \rangle = \frac{\int_{-1}^{1} d\cos\theta^* \int_0^{2\pi} d\phi^* f(\theta^*, \phi^*) d\sigma(\theta^*, \phi^*)}{\int_{-1}^{1} d\cos\theta^* \int_0^{2\pi} d\phi^* d\sigma(\theta^*, \phi^*)}, \tag{5.5}$$

which represent the angular weighted average of $f(\theta^*, \phi^*)$. The A_i can be extracted evaluating the correspondent momenta, relying on the orthogonality of the P_i harmonic functions :

$$\begin{aligned}
\langle \frac{1}{2}(1 - 3\cos^2\theta^*) \rangle &= \frac{3}{20}\Big(A_0 - \frac{2}{3}\Big), & \langle \sin\theta^* \cos\phi^* \rangle &= \frac{1}{4} A_3, & \langle \sin(2\theta^*) \sin\phi^* \rangle &= \frac{1}{5} A_6, \\
\langle \sin(2\theta^*) \cos\phi^* \rangle &= \frac{1}{5} A_1, & \langle \cos\theta^* \rangle &= \frac{1}{4} A_4, & \langle \sin\theta^* \sin\phi^* \rangle &= \frac{1}{4} A_7. \\
\langle \frac{1}{2} \sin^2\theta^* \cos(2\phi^*) \rangle &= \frac{1}{20} A_2, & \langle \sin^2\theta^* \sin(2\phi^*) \rangle &= \frac{1}{5} A_5,
\end{aligned} \tag{5.6}$$

These integrals can be evaluated using the inclusive cross section, ϕ^* and θ^* from MC truth information. Thus, the normalization weight can be defined as:

$$w_{\text{tot}} = \frac{1}{\sum_{i=0}^{8} A_i P_i(\cos\theta^*, \phi^*)}, \tag{5.7}$$

where $A_8 \equiv 1.0$, $P_8 \equiv 1 + \cos^2\theta$ is used to include σ^{U+L}. Reweighting the simulated events, selected from a narrow bin of Y_W and q_T^W for w_{tot} on event-by-event basis makes the $\cos\theta^* \times \phi^*$ distribution flat by construction. In fact, the weight w_{tot} is exactly the value of Eq. 5.1, given a certain value of $|Y_W|$, q_T^W, $\cos\theta^*$, ϕ^*. Therefore, if the selected events are reweighted for $P_i(\cos\theta^*, \phi^*) \cdot w_{\text{tot}}$, the $\cos\theta^* \times \phi^*$ distribution will describe the harmonic i. The normalization of this distribution represents the value of A_i in this bin of $|Y_W|$, q_T^W. In this step, the $|Y_W|$ and q_T^W bin ranges, which define the selected MC events to build the template, should be narrow enough to allow to neglect variation of the A_i within the bin.

After the event-by-event-reweighting for $P_i(\cos\theta^*, \phi^*) \cdot w_{\text{tot}}$, the $\eta^\mu \times p_T^\mu$ templates are built using the CMS simulation by applying the reconstruction and kinematic selections, for each $|Y_W|$, q_T^W bin and for each P_i. This step relies on the simulation of the muon propagation and interaction with the detector material only, and it is independent from the W boson kinematics. The only additional theoretical assumption is a proper description of the final state radiation and it will be discussed in Sec. 5.4.2.3.

Is relevant to stress again that the details of the Y_W, q_T^W and polarization encoded in the MC have not been used in the template building procedure. In principle a simulation of $W \to \mu\nu$ events which produces non-physical spectra (eg. flat) can be also used, since the shape within a $Y_W \times q_T^W \times A_i$ is made flat and then reshaped as P_i, while the normalization of each bin is freely floating and adjusted by the fit.

5.2.2 Features of Templates

The explicit transformation from the CS frame variables to the laboratory frame p_T^μ, η^μ is:

$$
\begin{aligned}
p_T^\mu &= \frac{1}{2}\sqrt{(E_T^W \cos\phi^* \sin\theta^* - q_T)^2 + (m_W \sin\phi^* \sin\theta^*)^2}, \quad \text{where } E_T^W = \sqrt{m_W^2 + (q_T^W)^2}, \\
\eta^\mu &= -\ln\left[\frac{p_T^\mu}{E_\mu + p_z^\mu}\right] = -\ln\left[\frac{\frac{1}{2}\sqrt{(E_T \cos\phi^* \sin\theta^* - q_T^W)^2 + (m_W \sin\phi^* \sin\theta^*)^2}}{\cosh y(1 - \tanh y)(\frac{E_T^W}{m_W} - \frac{q_T^W}{m_W}\cos\phi^* \sin\theta^* + \cos\phi^*)}\right].
\end{aligned}
\tag{5.8}
$$

The derivation is provided in Appendix C. The Eq. 5.8 contains only even ϕ^* terms, which means that ϕ^* and $-\phi^*$ are not distinguishable on the $\eta^\mu \times p_T^\mu$ of the templates. Therefore the laboratory variables (p_T^μ, η^μ) and the CS frame variables $(\cos\theta^*, \phi^*)$ has a 1-to-1 relation, modulo the $\phi^* \leftrightarrow -\phi^*$ folding.

This degeneracy would be broken by the distribution in the laboratory frame of :

$$
\phi = \phi^\mu - \phi^W. \tag{5.9}
$$

However, templates are integrated in this variable, since ϕ cannot be measured, due to the ignorance in the W boson momentum direction. In principle, the ϕ distribution can be measured including a proxy of the direction of q_T^W with the use of the recoil (i.e. p_T^{miss}). However, the precision of the recoil in the low q_T^W region (which contain the bulk of the distribution, the most relevant for this analysis) is poor, and is not reliable to build 3-D templates $\eta^\mu \times p_T^\mu \times \phi$.

This result has a strong implication in the W boson production properties measurement. The harmonic function of A_5, A_6 and A_7 are odd in ϕ^*, therefore their contributions to the templates statistically cancel out and they cannot be measured with the proposed $\eta^\mu \times p_T^\mu$ template fit. This is not a limitation for the measurement of the W production properties, since the W production and the detector acceptance are symmetric in ϕ. The integration in ϕ ensures that the acceptance for $+\phi^*$ and $-\phi^*$ are equal for any $\cos\theta^*$. The cancellation is then exact. Therefore no systematic uncertainty arises from the not measured angular coefficients.

The feature of the templates can be now described as a function of the laboratory variables. In the approximation of $\Gamma_W = 0$, $m_\mu = 0$, it is possible to describe the relation between p_T^μ, η^μ, and ϕ (from Eq. 5.9) in the laboratory frame, as a function of the W boson kinematics (m_W,Y_W and q_T^W):

$$p_T^\mu = \frac{m_W^2/2}{E_T^W \cosh(Y_W - \eta^\mu) - q_T^W \cos\phi}. \tag{5.10}$$

It is useful to analyse this relation as a 1-D function $p_T^\mu(\eta^\mu; m_W, Y_W, q_T^W, \phi)$ to describe the templates. In Fig. 5.2 some examples of the analytic expression of Eq. 5.10 are shown. This function has a Gaussian-like shape, with mode at $\eta^\mu = Y_W$, where $p_T^{\mu,\max} = m_W/2$ at $q_T^W = 0$. More in general m_W determines the value of the curve, while a variation in Y_W results in a rigid shift of the curve toward higher or smaller η^μ. The ϕ produces a modulation of the shape similar to the mass variation. The magnitude of this modification is zero for $q_T^W = 0$ (i.e. there is no ϕ dependence), while increases for higher q_T^W. In particular, due to the $\cos\phi$ dependence, ϕ and $-\phi$ reflects in the same value of p_T^μ. At fixed value of ϕ, the q_T^W reduces the width and increases the peak value of the curve. The introduction of $\Gamma_W \neq 0$ smears m_W with a Breit-Wigner distribution and $\sqrt{s} \neq m_W$ produces a shift of the peak of the curve which mimic the q_T^W effect.

The described features are shown in Fig. 5.3, as the distribution $\eta^\mu \times p_T^\mu$ of simulated events at generator level in different bins of Y_W and q_T^W, at fixed m_W value, integrated over ϕ. The ϕ integration produces the observed band. The effect of reconstruction is minimal and does not spoil the main features. This aspect is crucial, since represents in the sensibility of the templates to the W boson kinematics. Most important, the reconstruction limits the acceptance, mainly with a lower cut on the p_T^μ (due to the trigger threshold and background suppression requirements) and an upper cut on $|\eta^\mu|$ (due to the geometrical acceptance of the detector). The acceptance in the $\cos\theta^* \times \phi^*$ plane depends in a non trivial way on the laboratory acceptance and on the W boson kinematics and polarization. This is the source of the PDF dependence on the m_W measurement using p_T^μ and m_T template fit.

5.2.3 Regularization of the angular coefficients

The angular coefficients are expected to be smooth functions of $|Y_W|$ and q_T^W, from the legacy of previous measurements. For instance, CDF [179], CMS [180] and ATLAS [157] previously performed differential measurements of rapidity distribution, polarization and angular coefficients of $Z \to \ell\ell$ decays on Tevatron and LHC-Run 1 data. The Tevatron Collaborations also performed several measurements of W polarization and angular coefficient distributions [181–183] (while CMS and ATLAS performed only the measurement of the helicity fractions [184, 185], so far).

In addition there are some specific behaviours at $q_T^W = 0$ and $Y_W = 0$ which can be derived from the theoretical description of the angular coefficients or symmetries of the measurement framework. Specifically, all the angular coefficients but A_4 are expected to vanish at $q_T^W = 0$, recovering the LO pure V-A behaviour. On the other hand A_4 is expected to be not-null at $q_T^W = 0$. In addition at $Y_W = 0$, $\theta^* = \theta_{\text{lab}}$. The cross section must be symmetrical in θ i.e. even with respect to $\theta_{\text{lab}} = \pi/2$, due to the symmetry of the beams of LHC. At $Y_W = 0$ harmonic functions $P_1(\theta^* = \theta_{\text{lab}}, \phi^*)$ and $P_4(\theta^* = \theta_{\text{lab}}, \phi^*)$ are odd with respect to $\theta = \pi/2$. Thus $A_1(Y_W = 0)$ and $A_4(Y_W = 0)$ must vanish.

These trends can be converted in constraints on the template fit to angular coefficient. The hypothesis of smooth function can be converted in a polynomial description of the $Y_W \times q_T^W$ distributions. This results in a regularization of the angular coefficient distributions, which can strongly improve the precision of the measurement. The regularization can be implemented with different approaches, and a specific implementation will be presented in Sec 7.5.

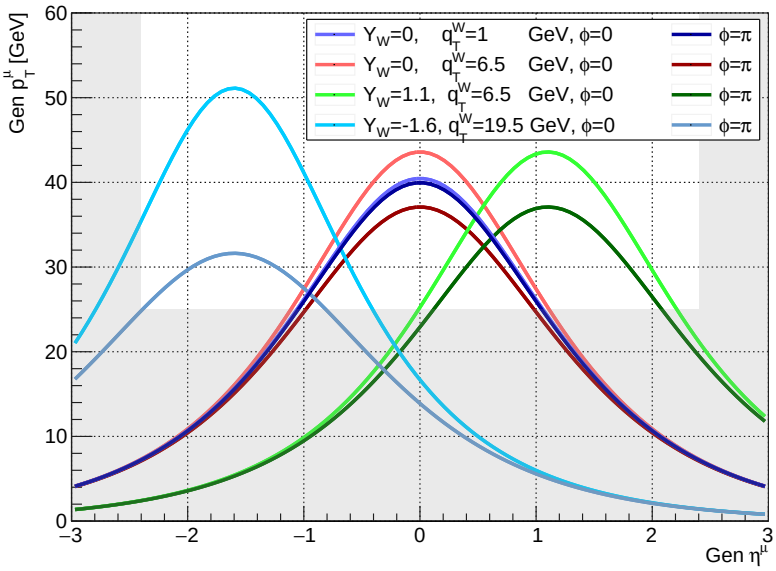


Figure 5.2. *Analytic expression of $p_T^\mu(\eta^\mu; m_W, Y_W, q_T^W, \phi)$ from Eq. 5.10 for different values of the parameters. For each Y_W, q_T^W combination, ϕ modulates the function between two extremal value that are shown. The shaded grey area delimits the region out of the acceptance (in the approximation $p_T^{gen} = p_T^{reco}$,$\eta^{gen} = \eta^{reco}$).*

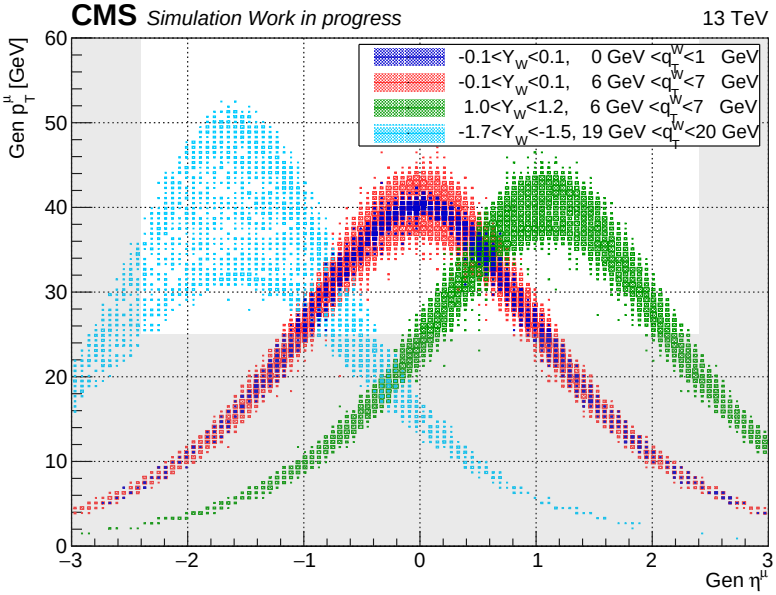


Figure 5.3. *Templates distribution at generator level, for different bins of Y^W and q_T^W, integrating above the polarization states and ϕ. The parameters of the functions of Fig. 5.2 corresponds to the bins shown here. A sample of $W^+ \to \mu^+\nu$ simulated events has been used, without applying acceptance of reconstruction selection. The color intensity is proportional to the event content. Higher $|Y_W|$ corresponds to a η^μ shift of the distribution, while higher q_T^W to a larger spread due to the larger ϕ^μ modulation. The ϕ trend reflects in higher population at the edge of the bands ($\phi = 0, \pi$). The shaded grey area delimits the region out of the acceptance (in the approximation $p_T^{gen} = p_T^{reco}$,$\eta^{gen} = \eta^{reco}$).*

The hypothesis of smooth function is valid on the angular coefficients distribution A_i, but not on the helicity cross sections σ_i. Specifically the σ^{U+L} cannot be regularized as a function of q_T^W, to avoid to set artificial constraints on the transverse momentum spectrum of the W boson.

The expected distributions of the angular coefficients as a function of q_T^W and Y_W will be shown in Figs. 7.7 to 7.11 together with the fit results.

5.2.4 Analysis requirements

Given the analysis strategy outlined in the previous sections, it is relevant to summarize the underlying assumptions. The description of the W boson production in term of 9 helicity cross sections of the Eq. 5.1, relies on the vectorial nature of the primary particle, and the fermionic nature of the decay products. Therefore the description is valid for each $A \to BC$ decays where the particle A has spin 1 and particles B and C have spin $\frac{1}{2}$.

Both the parton-level hadronic component, which contains the PDF information, and the electroweak lepton component of the matrix element are fully encoded in the values of the angular coefficients and σ^{U+L}. No assumption on the production or decay physics is encoded in the decomposition of Eq. 5.1 and the consequent strategy, but the spin of the particles. Since the value of spin the exploited particles ($W^{\pm}$, $\mu^{\pm}$, ν) is not debated , no systematic uncertainties have been considered for the production model.

The templates rely on the value of m_W encoded in the simulation. This is an additional assumption that requires a dedicated systematic uncertainty evaluation. However, in Sec. 5.2.5 will be discussed the possibility to fit simultaneously m_W and the W production mechanism, relaxing this assumption.

The Γ_W dependence of the templates has been not considered in the analysis. It produces an effect similar to q_T^W, but the experimental uncertainty on Γ_W is 42 MeV [18], much smaller than the adopted q_T^W binning. Thus, the Γ_W dependence can be neglected.

Another required assumption is the availability of a MC simulation that is able to describe the propagation of the muon and its detection with sufficient accuracy. In particular the MC must include a precise description of the FSR of the muon. A proper systematic uncertainty must be assigned to that.

The analysis requires 3 major ingredients from the experimental side. The first is the muon scale and resolution calibration, which must be sufficiently accurate and precise to fulfil the analysis target precision. The second is a measurement of the selection and reconstruction efficiency of the events in data and MC simulation. The third is a precise and robust background estimation method. Systematic uncertainties must be assigned to each of these 3 ingredients, and taken into account performing the fit.

Finally, the regularization discussed in Sec. 5.2.3 represent an additional assumption which must be implemented avoiding to bias the final result. The regularization is an optional feature which can be used or not, and is not a mandatory ingredient of the measurement. The analysis will be discussed in the next chapters without the regularization hypothesis. This extra feature will be introduced in Sec. 7.5 only.

5.2.5 Simultaneous fit to m_W

The $\sigma^{U+L}(Y_W, q_T^W)$, the $A_i(Y_W, q_T^W)$ and the consequent muon distributions naturally depend on the value of m_W. In the Eq. 5.1 m_W has been hidden and this choice reflects at analysis level in relying on m_W value encoded in the simulation to describe the W boson properties. However,

the Eq. 5.1 can be rewritten as:

$$\frac{d\sigma}{dq^2_{T,W}dY_W d\Omega^* dm_W} = \frac{3}{16\pi}\frac{d\sigma^{U+L}}{dq^2_{T,W}dY_W dm_W}\Big[F(\Omega^*)\Big], \tag{5.11}$$

where $d\Omega^* \equiv d\cos\theta^*_\mu d\phi^*_\mu$ and $F(\Omega^*)$ is a compact form for the content of the square bracket of Eq. 5.1. In the given relation the dependence of m_W will affect the $A_i = A_i(m_W, Y_W, q^W_T)$, $\sigma^{U+L} = \sigma^{U+L}(m_W, Y_W, q^W_T)$ only. Moreover the muon $d\Omega^*$ distributions are not affected by the value of m_W in the CS frame, since they are described by the spherical harmonics only, while the muon distributions in the laboratory frame are sensitive to m_W. For this reasons the templates can be built in bins of Y_W, q^W_T and m_W.

The m_W value can be directly extracted simultaneously to the W boson production properties with the same fitting approach, if the m_W dependence is taken into account using Eq. 5.11. If m_W is treated as a nuisance parameter of the fit, the measurement can be used as a constraint of systematic uncertainties related to the W production mechanism. If m_W is fitted simultaneously to the W boson production properties, the fit can be used for the determination of m_W itself. Since the analysis strategy is the same in both cases, the m_W dependence has been neglected in the description for simplicity. The relevant difference will be highlighted when needed. The entire analysis strategy is valid both in case m_W is fixed to the MC value or in case it is extracted from the fit.

Obviously, in the case of this extended simultaneous fit, no hypothesis on the value of m_W are needed to build the templates, but several versions of the entire set of templates will be produced with different values of m_W. A preliminary version of this approach and the related results will be presented in Sec. 8.1.

5.3 Samples and event selection

The analysis has been performed exploiting the state-of-the-art compressed CMS data format, called NanoAOD [14]. This data format contains much less information compared to the dataset containing the CMS full-reconstruction information, and it is intended for physics analysis only. On the other hand, they are very lightweight and easy to access also outside of the `CMSSW` framework. The W mass and properties analysis performs a measurement differential in several variables. It means that about 400 histograms must be produced and managed for each step of the analysis. These numbers increase by a factor of 100 considering the variations needed to study the systematic uncertainties. These constraints require a fast and flexible system that is able to strongly parallelize the tasks. The adopted solution was the development of a framework based on `RDataFrame` [13]. Additional details about the NanoAOD and the *W-properties* framework and `RDataFrame` are provided in Appendix D.

In this section, the samples (data and MC simulation) used in the analysis are described together with the selections applied to define the signal region.

5.3.1 Data samples

This analysis is based on the data sample recorded by CMS during 2016 at $\sqrt{s} = 13$ TeV. The sample corresponds to an integrated luminosity of 35.9 fb^{-1}, measured with an uncertainty of 2.5% as reported in Ref. [186]. Proper functioning of all the CMS sub-systems is guaranteed in all the data-taking runs which form the used data sample, ensuring CMS standard high quality of the data. The analysis has been performed separately for W^+ and W^-, therefore two different samples are considered.

The data sample is taken from the *SingleMuon* Primary Dataset, which corresponds to the logical `OR` of multiple HLT paths. Specifically, the exploited paths are:

`HLT_IsoMu24_v* OR HLT_IsoTkMu24_v*.`

These paths require at least one isolated muon in the event, with some requirement on the track quality, $p_T > 24$ GeV and $|\eta| < 2.5$. Different triggers are used to collect the Z events needed for momentum scale calibration and efficiency measurement, but these ancillary measurements have not been performed directly in this work and the respective samples are mentioned in the proper sections.

The data sample is subdivided in different eras, enumerated from B to H (while era A is used for detector commissioning phase, preliminary to physics analysis), which correspond to slightly different hardware, software and trigger conditions to accommodate the data taking conditions. The integrated luminosities for each era are reported in table 5.1. The events in the selected sample have been reconstructed with the most updated reconstruction software of CMS at the time of the writing of this thesis. Additional information about the details of the reconstruction and the sample used are provided in Appendix D.

Table 5.1. *Integrated luminosity for each run era in 2016 data taking period.*

Era	B	C	D	E	F	G	H	Total
$\mathcal{L}_{\text{int}}$ [fb^{-1}]	5.75	2.57	4.24	4.03	3.11	7.58	8.65	35.93

5.3.2 Monte Carlo samples

The analysis exploits several MC samples to model the signal and background events. The source and the description of the background channels are given in Chapter 6, while here only the details of the simulated samples are given.

The $W \to \mu\nu$ sample has been generated at NLO accuracy in perturbative QCD with the `MadGraph5_aMC@NLO` generator [119]. Likewise also $Z \to \mu\mu$ and $W \to \tau\nu$ samples have been generated with `MadGraph5_aMC@NLO` at NLO accuracy, while diboson (WW, WZ, ZZ) and $t\bar{t}$ samples have been generated at LO accuracy. The single-top sample has been generated with `POWHEG2.0` generator [118] at NLO accuracy. All processes have been interfaced with `PYTHIA 8.226` parton shower [121] with `CUETP8M1` tune [187], for proper modelling of hadronization and underlying event description. The `NNPDF 3.0` package [120], accurate at NLO in perturbative QCD, has been used to describe the PDF in all the processes. The `Geant4` package [122] interfaced with `CMSSW` has been used to simulate the propagation in the material and the detector response. A summary of the MC samples and the simulation details is given in table 5.2.

After the generation, the MC samples have been reconstructed with exactly the same algorithm as the data sample. Therefore, they result in a completely consistent format compared to data. However, the additional information from the MC-truth is preserved, for further usage. It is usually called the *Gen level* information, to be compared to the *Reco level* (i.e. the information after the standard reconstruction, also present on data).

5.3.3 Physics objects calibration

The calibration of the physics objects has been performed prior to the event selection.

Muon momentum. The main object of interest in this analysis is the muon momentum. Thus, the muon momentum scale and resolution have been calibrated applying the Rochester corrections, as described in Sec. 2.3.3.1. Therefore in the later stage of the analysis, the p_T^μ must be intended as the Rochester-corrected-p_T^μ resulting from this calibration procedure. The effect of this correction is shown in Fig. 2.15, from W rapidity and helicity measurement, since the same sample with the same selection has been used in this analysis. No dedicated studies have been performed on top of these corrections.

MET. The p_T^{miss} has been calibrated with the JEC, previously introduced in Sec. 2.3.4. The JEC have been applied according to CMS Jet-MET physics object group (POG) recommendation as reported in Ref. [188]. However, the impact of p_T^{miss} is limited, since m_T is the only object which relies on p_T^{miss} and it will be used only to refine the event selection, as discussed later. The use of the particle-flow based p_T^{miss} (defined in Eq. 2.9) in the m_T definition is the only relevant choice for the sake of this analysis.

The systematic uncertainties related to the p_T^μ and p_T^{miss} calibration are described in Sec. 5.4.

Table 5.2. *MC samples used in the analysis, with the information about the generator, the encoded production cross section, and the equivalent integrated luminosity.*

Process	σ [pb]	$\mathcal{L}_{\text{int}}^{\text{eq}}$ [fb^{-1}]	generator
$W(\to \ell\nu)$+jets	61526.70	4.7	`MadGraph_aMC@NLO`
$Z/\gamma^*(\to \ell\ell)$, $m_{\ell\ell} > 50$ GeV	6025.20	9.1	`MadGraph_aMC@NLO`
$Z/\gamma^*(\to \ell\ell)$, 10 GeV $< m_{\ell\ell} < 50$ GeV	1093.00	29.1	`MadGraph_aMC@NLO`
$t\bar{t}(\ell)$	182.00	623.6	`Madgraph, LO`
$t\bar{t}(\ell\ell)$	95.02	319.2	`Madgraph, LO`
t (t-channel)	136.20	493.3	`POWHEG, NLO`
$\bar{t}$ (t-channel)	80.95	479.4	`POWHEG, NLO`
top (s-channel)	3.68	105.5	`POWHEG, NLO`
tW	35.60	195.3	`POWHEG, NLO`
WW	115.00	69.4	`Madgraph, LO`
WZ	47.13	84.8	`Madgraph, LO`
ZZ	16.50	59.9	`Madgraph, LO`

5.3.4 Event selection

After the physics objects calibration, the event selection of the analysis is performed. It consists of a list of selection criteria on global event and muon properties. They are applied to data and MC samples exactly in the same way.

The selection procedure has been performed in order to build two separate samples for data and for each MC sample, based on the muon charge, since the W boson mass and properties measurement is performed separately for the two W boson charges. The same selection procedure has been applied to the two charges, excluding the muon charge identification. Due to the excellent performance of muon reconstruction of CMS and the p_T^μ range under analysis, the probability of wrong muon charge assignment is negligible.

The selection criteria, both for data and MC samples, are:

– The event must pass the `HLT_IsoMu24_v*` or `HLT_IsoTkMu24_v*` trigger paths.

– The distance between the primary vertex of the hard scattering process and the nominal IP position must be smaller than 24 cm along the beam axis and smaller than 2 cm in the transverse plane. In addition, the vertex fit must have more than 4 degrees of freedom. The primary vertex is chosen among the PU vertices as the one with the largest $\sum_{\text{tracks}} p_T^2$ where the sum runs over the tracks associated with the vertex.

– Exactly one muon must be present in the event. The muon must pass the following requirements:

  - Pass the medium muon ID (as defined in Sec. 2.3.3), which can be summarized as: global muon, more than 80% of valid hits in the tracker, high compatibility with the segments in the muon system, low material effects, high track quality in term of χ^2.
  - Low transverse and longitudinal impact parameters: $d_{xy} < 0.05$ cm, $d_z < 0.2$ cm. This selection improves the QCD and top background suppression, in which semileptonic b decays can produce muons with a large impact parameter.
  - Fiducial kinematic region: $p_T^\mu > 20$ GeV, $|\eta^\mu| < 2.4$. This selection also improves the background suppression and defines a region where the calibration can be performed with high precision.

– If the event has additional leptons (electron or muons), it is vetoed. The muons are identified with the loose muon ID and $p_T^\mu > 10$ GeV. The electron are identified as: GSF-electrons which pass the veto-electron ID [189], $p_T^e > 10$ GeV, RelIso$^e < 0.3$, $d_{xy}^e < 0.05$ cm (0.1 cm), $d_z^e < 0.1$ cm (0.2 cm) for $|\eta| < 1.479$ ($|\eta| > 1.479$).

– The event must pass MET-filters, which refine the p_T^{miss} description according to the Jet-MET POG recommendations [190].

– The muon must satisfy 25 GeV $< p_T^\mu < 55$ GeV. This selection must be explicitly applied since the efficiency Scale Factors are provided in this p_T^μ range only. The lower limit discards a region where the background contribution is high (from QCD and $W \to \tau\nu$). The higher region of p_T^μ spectrum does not substantially improve the sensitivity of the analysis, because of the low rate of this region.

– The event must have $m_T > 40$ GeV and RelIso < 0.15 (without ambiguity RelIso$^\mu_{\text{PF}}$ will be indicated as RelIso hereinafter). This selection is needed to strongly improve the signal purity, reducing in particular the QCD background, ad described in detail in Chapter 6.

5.3.5 Monte Carlo reweighting

The reconstructed MC samples do not reproduce *exactly* the real data distribution, despite being processed by the same reconstruction algorithm. There are in fact multiple sources of discrepancy. First, the generators used in the simulation are able to produce the physics processes with limited accuracy and precision. Thus the limited knowledge encoded in the generators produces not-fully accurate MC samples. Second, the detector simulation also has a limited accuracy and precision. The knowledge of several aspects of the detector and of the reconstruction is not exact, in particular in some regions of the phase space. Examples are the material budget, the magnetic field, the detector alignment, the trigger simulation, the reconstruction with high PU or in dense environment, local failure/inefficiency of the subdetectors, time-dependent (eg. radiation-driven) effects not fully under control, etc. Third, the event selection procedure can affect differently data and MC, since the finite accuracy of the MC in the description of the selection variables.

The W boson mass and properties measurement relies on the simulated distribution of the templates to match the data distribution. The fit will adjust the signal strength of the templates to match the data predicting the production model and the mass of the W boson. The underlying assumption is that any discrepancy between data and MC arise from the W physics (i.e. if the MC encodes the exact W boson production model, the sum of the templates would exactly reproduce the data yield, without the adjustment of the fit). Therefore any discrepancy between data and MC arising from experimental or theoretical assumptions must be corrected.

This is typically done with a calibration procedure, which proceeds in two steps. First, the discrepancy between data and simulation is measured using a different data sample or a different phase space region. Second, the MC sample is reweighted accordingly to the observed discrepancy, extrapolating to the signal region/sample. After the reweighing, data and MC match by definition in the region/sample where the calibration is performed, while possible residual discrepancy can be present in the measurement region, due to inaccuracy in the extrapolation. A systematic uncertainty must be assessed to the calibration procedure and residual discrepancy.

The MC samples have been reweighted according to several calibration procedures:

Boson q_T and Y. The q_T and Y spectra of $Z \to \mu\mu$ sample have been reweighted according to the direct measurement [191] performed by CMS on the same sample used for this analysis. The result of Ref. [191] is reported in Fig. 5.4. The reweighting has been performed keeping the normalization of the distribution of q_T^Z and Y_Z constant in the two steps separately. Thus:

$$w_Y = \frac{(d\sigma_Z/dY)^{\text{meas}}}{(d\sigma_Z/dY)^{\text{MC}}} \cdot \frac{\int (d\sigma_Z/dY)^{\text{MC}} dY}{\int (d\sigma_Z/dY)^{\text{meas}} dY}, \quad w_{q_T} = \frac{(d\sigma_Z/dq_T)^{\text{meas}}}{(d\sigma_Z/dq_T)^{\text{MC}}} \cdot \frac{\int (d\sigma_Z/dq_T)^{\text{MC}} dq_T}{\int (d\sigma_Z/dq_T)^{\text{meas}} dq_T}$$

Since it is a direct measurement the reweighting procedure ensures consistency for the $Z \to \mu\mu$ background. The discrepancy between data and MC, estimated on η^μ and p_T^μ spectra with the W event selection, is reduced when this weight is applied to Z MC sample.
Also the $q_T^{W,\text{MC}}$ spectrum has been reweighted for w_{q_T}, justified from the fact both q_T^W and q_T^Z data/MC discrepancies mainly arise from the incompleteness of the NLO description at low q_T. This additional reweighting improves the agreement between data and MC in the relevant observables (p_T^μ). Since the q_T^W is measured in this analysis the q_T^W reweighting affects only the QCD background estimation. Numerically w_{q_T} is between 1% and 10% and w_Y is 1%-2%, as can be seen from right panels of Fig. 5.4.

Pile-up. The simulated in-time (same bunch crossing) and out-of-time (adjacent bunch crossings) PU scenarios are typically different from the one observed in the data. Therefore all the MC samples are reweighted with w_{PU} to match the data pile-up distribution, as a function of the number of PU interactions of the event, according to the indication of the Luminosity POG of CMS. The number of vertices and the average energy density ρ distributions from $Z \to \ell\ell$ events are used to validate the reweighting. Some residual discrepancy is observed, because of the vertex reconstruction efficiency mismodelling in the simulation. More details about the PU reweighting can be found in Ref. [11]. Numerically w_{PU} is between 0.8 and 1.2.

Reconstruction and selection efficiencies. The reconstruction and selection efficiencies have been aligned between data and MC exploiting the efficiency measurement of Ref. [11]. The W boson mass and properties analysis exploits exactly the same sample and selection of Ref. [11]. The procedure is described in detail in Sec. 2.3.3.5. The result is the application of a set of efficiency Scale Factors (SFs) to the MC samples, function of η^μ, p_T^μ and the charge, defined as $\text{SF}_i = \varepsilon_i^{\text{data}}/\varepsilon_i^{\text{MC}}$, where ε_i^j represent the measured efficiency of the selection i on the sample j. The total efficiency scale factor weight applied is: $w_{\text{SF}} = \text{SF}_{\text{sel}}\text{SF}_{\text{trig}}$. The values of the weight are shown in Fig. 2.19 and are between 0.75 and 1.2.

L1 Trigger prefire. In 2016 the shape of the digitized pulse from ECAL started to drift toward negative values, because of the transparency losses due to radiation damage. The effect is strongly dependent on η, and particularly relevant at $|\eta| > 2$. The phase between the detector readout and the trigger has not been corrected for the described drift, resulting in a possible wrong assignment of the bunch crossing (BX). If an event is assigned to the previous bunch crossing (i.e. the trigger prefires) it can be vetoed by the trigger rules, which automatically discard the two consecutive BX after a trigger fire, producing a loss of efficiency. This effect can be measured using a sample of un-prefirable events and performing a tag-and-probe efficiency measurement with $Z \to ee$ events. The tag has been chosen to be an electron (with correct BX) matched with a HLT object, while the probe has been chosen a L1 trigger object only.
A prefire probability $p_{\text{BX-1}}$ has been measured with this approach in Ref. [11], estimating the scale factor $\text{SF}_{\text{prefire}} = 1 - p_{\text{BX-1}}$, as a function of η. In this analysis, the electron trigger is not directly used, but m_T and RelIso are used as selection variables and they exploit also the calorimeter measurements. Moreover, the prefire can be also triggered by jets present in muon-triggered events. Therefore the $w_{\text{prefire}} = \text{SF}_{\text{prefire}}$ has been applied to all the events of all samples as recommended by the Jet-MET POG of CMS. Numerically w_{prefire} is between 0.1% and 1%.

Luminosity. This is not a calibration of the MC but a standard procedure that is applied whenever a MC sample must be compared to data. Since the total number of events generated in the MC sample is arbitrary and it is not related to the expected yield, the MC must be rescaled to match the expected yield on data, according to the integrated luminosity of the data sample and the expected cross section of the process (σ_i), reported in table 5.2. A luminosity weight is applied to each sample, on event by event basis:

$$w_{\mathcal{L}}^{i,j} = \frac{\mathcal{L}_{\text{int}}^{\text{data}} \sigma_i w_j}{\sum_j w_j},$$

where i denotes the MC sample, while w_j is the generator weight of the event j of the sample i. Is this framework it is useful to define the MC equivalent luminosity i.e. the luminosity of a sample with the same statistical power of the MC one. It is defined as: $\mathcal{L}_{eq} = N_{\text{eff}}/\sigma_i$, where $N_{\text{eff}} = (\sum_j w_j)^2 / \sum_j w_j^2$. The values of L_{eq} of the adopted MC samples are reported in table 5.2.

All the reweighting procedures are intended on event by event basis, unless specified otherwise. All the weights applied have been estimated on different sample or phase space region, therefore they are not affected by the previously required selection. All the MC events which passed the selection described in Sec. 5.3.4 have been reweighted. The total applied weight is:

$$w_{\text{tot}} = (w_Y \cdot w_{q_T}) \cdot w_{\text{PU}} \cdot w_{\text{prefire}} \cdot w_{\text{SF}} \cdot w_{\mathcal{L}}. \tag{5.12}$$

5.4 Systematic uncertainties

The procedures described in the previous sections to select and calibrate the samples are affected by several sources of systematic uncertainty. These uncertainties affect the templates and thus propagate to the fit results. They can be grouped in two categories depending on their source:

- **Experimental systematic uncertainties:** induced by the modeling and the definition of the experimental quantities used in the reconstruction, selection and calibration of the events.
- **Theoretical systematic uncertainties:** induced by the theoretical input of the MC simulation, from which the used variables have dependencies.

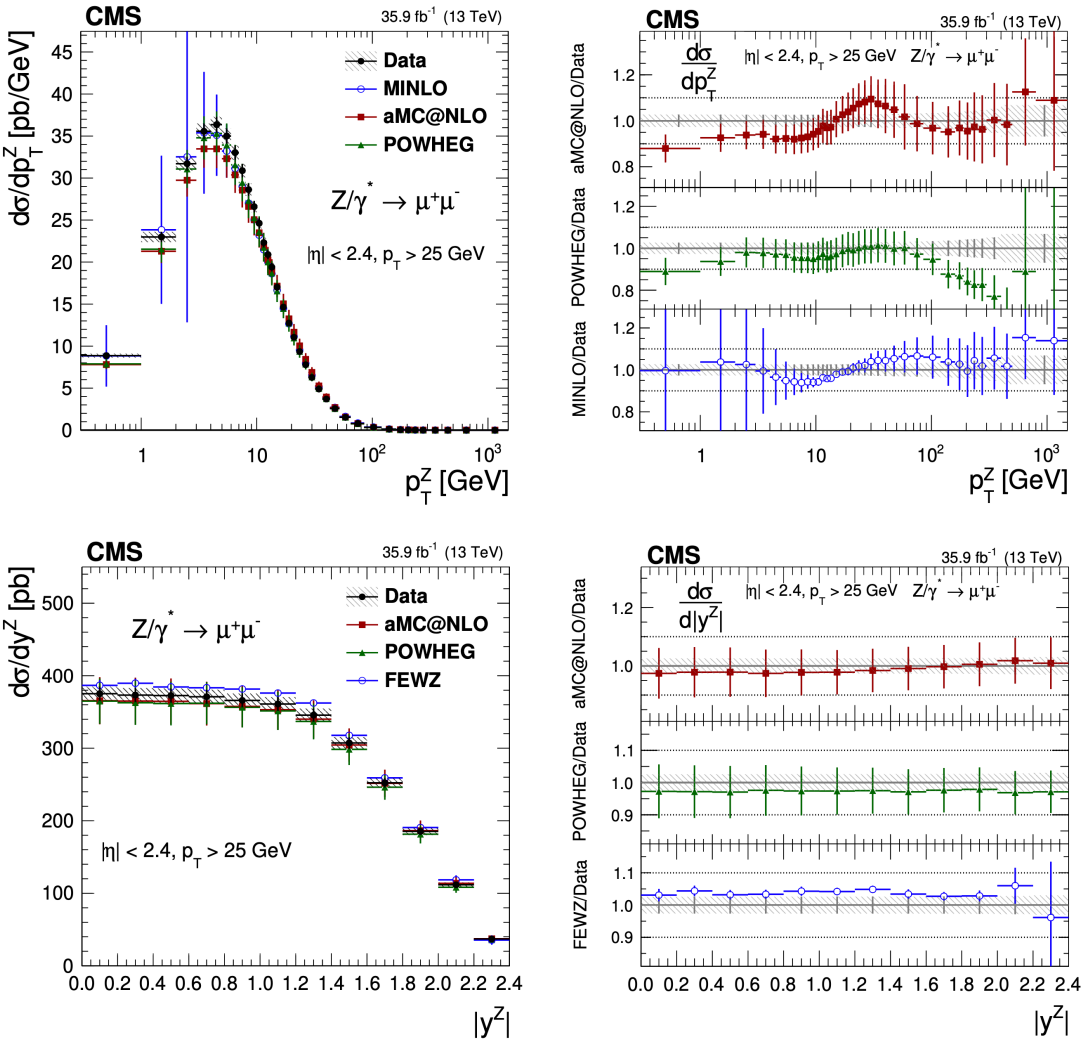


Figure 5.4. *Measured q_T^Z (top) and Y_Z (bottom) distribution, compared to different simulation predictions. In the right pads the ratios of predictions to data are shown, where the shaded areas correspond to the experimental uncertainty. The weights applied to analysis are equivalent to the* `aMC@NLO` *ratios (from Ref. [191]).*

5.4.1 Experimental systematic uncertainties

Whenever a selection based on experimental quantities is applied both on data and MC simulation a bias can be induced, if the quantity is not exactly aligned between the data and the MC. Some events can in fact fall inside or outside the acceptance of the selection differently in data and MC, producing a bias.

These systematic uncertainties result in variation of the correspondent corrections or variables, typically shifting the nominal value of a variable up and down of the same quantity. These up/down-varied versions of the correction or variables are used instead of the nominal value in the analysis to propagate the systematic uncertainty to the templates. This is common procedure to manage the nuisance parameters of the fit, as will discussed in detail in Chapter 7.

In the next paragraphs, the full list of experimental systematic uncertainties relevant for this analysis is described.

5.4.1.1 MET-related

The Jet Energy Corrections (JEC) have been applied to p_T^{miss} and the Jet Energy Resolution (JER) has been calibrated according to standard CMS reconstruction. Two up/down p_T^{miss}-related variations have been taken into account, following the Jet-MET POG indication. The first is related to the JES calibration, the second to the systematic uncertainty in the estimation of the unclustered energy when the jets and the p_T^{miss} objects are built during reconstruction [117, 192]. These variations produce alternative versions of p_T^{miss} and affect mainly m_T and thus the definition of the signal region. Both variations are completely correlated on the $\eta^\mu \times p_T^\mu$ plane, for each sign of W charge. The described JEC variations take into account the pile-up reweighting uncertainties.

5.4.1.2 Muon momentum calibration

The Rochester corrections are provided together with a set of systematic variations which take into account the q_T^Z modelling uncertainty, the electroweak effects on the Z lineshape and the effects of different $m_{\mu\mu}$ acceptance range. In addition a set of 100 replicas of the data and MC samples are provided to estimate the statistical uncertainty due to the limited Z sample used in the calibration. All the variations are provided in fine bins of $\eta^\mu \times p_T^\mu$.

The use of the Rochester corrections is only a temporary solution. They do not provide a scale calibration precision at 10^{-4} level, and therefore a more refined method is currently under development for future measurements. In this perspective, for practical reasons a less detailed scheme has been exploited in this proof of feasibility for the Rochester correction systematic variations.

A single up/down variation of the p_T^μ has been considered, which represents the envelope of the entire set of variation assuming a full correlation on $\eta^\mu \times p_T^\mu$ plane. This solution is clearly a strong simplification, which does not allow to trace back the physical origin of the systematic uncertainty and it smoothens the effect of the variation. However, it should reproduce the proper order of magnitude of the systematic uncertainty related to the muon scale calibration. This systematic uncertainty affects directly p_T^μ in the binning of the templates, and indirectly m_T and RelIso variables used in the selection.

5.4.1.3 Efficiency scale factors

The angular coefficients distributions are especially sensitive to the muon efficiencies. The A_i are defined as ratios between cross sections, therefore the uncertainties typical of the differential cross section measurement (like luminosity) cancel out. On the other hand, the efficiencies are not completely correlated and typically produce residual uncertainties also in the ratios.

The SFs measured in Ref. [11] are provided with the variations which describe the systematic uncertainty of the efficiencies measurement. Two kinds of systematic uncertainties have been considered.

First, a variation of the data efficiencies has been obtained changing the signal and the background models implemented in the fit. This variation has been propagated to the SFs and produces a variation almost flat in p_T^μ, but $|\eta^\mu|$-dependent. This variation is fully correlated in η^μ and p_T^μ. The trend has been modelled in 3 large bin of $|\eta^\mu|$, inclusive in p_T^μ, in which the variation has been considered flat. The magnitude of the variation is 0.2% for $0 < |\eta^\mu| < 1$, 0.4% for $1 < |\eta^\mu| < 1.5$, for 1.4% in $1.5 < |\eta^\mu| < 2.4$.

Second, a systematic uncertainty due to the finite statistic used to estimate the SFs has been considered. This variation is uncorrelated in η^μ, because of its statistical source, while it is

fully correlated in p_T^μ due to the smoothing procedure. Only the SF with the higher statistical uncertainty has been considered (i.e. the $\mathrm{SF}_{\mathrm{trig}}$). For each η^μ bin the fit has been repeated and the covariance matrix of the parameter has been diagonalized separately for data and MC. The 3 eigenvalues of the diagonalized matrix have been varied by $\pm 1\sigma$, independently. The values of the SF have been recalculated in the 3 cases, summing in quadrature the variation given by the $\varepsilon^{\mathrm{data}}$ and ε^{MC}. The results are 3 different values of the $\mathrm{SF}(p_T)$, for each bin of η. The magnitude of the variation has been inflated by a conservative factor $\sqrt{2}$ to take into account also of the statistical uncertainty of $\mathrm{SF}_{\mathrm{sel}}$. The magnitude of the variations is shown in Fig. 5.5, in the full $\eta^\mu \times p_T^\mu$ plane.

In conclusion the SF systematic uncertainty has been modelled with one up/down variation fully correlated both in η^μ and p_T^μ and 3 up/down variations fully correlated in p_T^μ and uncorrelated in η^μ. These variations are applied as a global weight to all the MC samples, affecting all the variables under analysis.

5.4.1.4 L1 Trigger prefire correction

An up/down variation of the prefire weight has been considered, to take into account the uncertainties arising from the prefire probability estimation. This variation has been provided by the Jet-MET POG of CMS.

5.4.1.5 Lepton veto

The selection described in Sec. 5.3.4 includes a veto on extra leptons in the event. This selection suppresses especially the $Z \to \ell\ell$ events, and a discrepancy between the efficiency of this selection between data and MC is possible. Therefore a systematic uncertainty on the $Z \to \ell\ell$ event yield has been considered. The magnitude of this uncertainty has been assessed as 2% following the choice of Ref. [11].

5.4.1.6 Data luminosity

The Luminosity POG of CMS measured the luminosity with a 2.5% precision [186]. Since the overall normalization of the MC samples is based on the value of the integrated luminosity, a systematic uncertainty of 2.5% on the total event yield of each MC sample has been considered.

5.4.2 Theoretical systematic uncertainties

Despite the analysis has been designed to extract the W production model without theoretical assumption, some residual effects must be considered. Some sources can be actually measured or constrained by the measurement itself (m_W or the PDFs). Some other minor contributions arise instead from the background contributions and their estimation.

5.4.2.1 PDFs and α_s

The shapes of the templates constructed with the procedure described in Sec. 5.2.1 are independent from the PDF, for a given bin of q_T^W, Y_W and for a given angular coefficient. The normalization of the templates will be predicted by the simultaneous fit, therefore the input value is arbitrary. The initial value of the normalization has been chosen to reproduce the $A_i(q_T^W, Y_W)$ encoded in the MC. Changing the PDF set in the MC, this normalization will change, but no bias is expected from the PDF, since the normalizations are left freely floating in the fit.

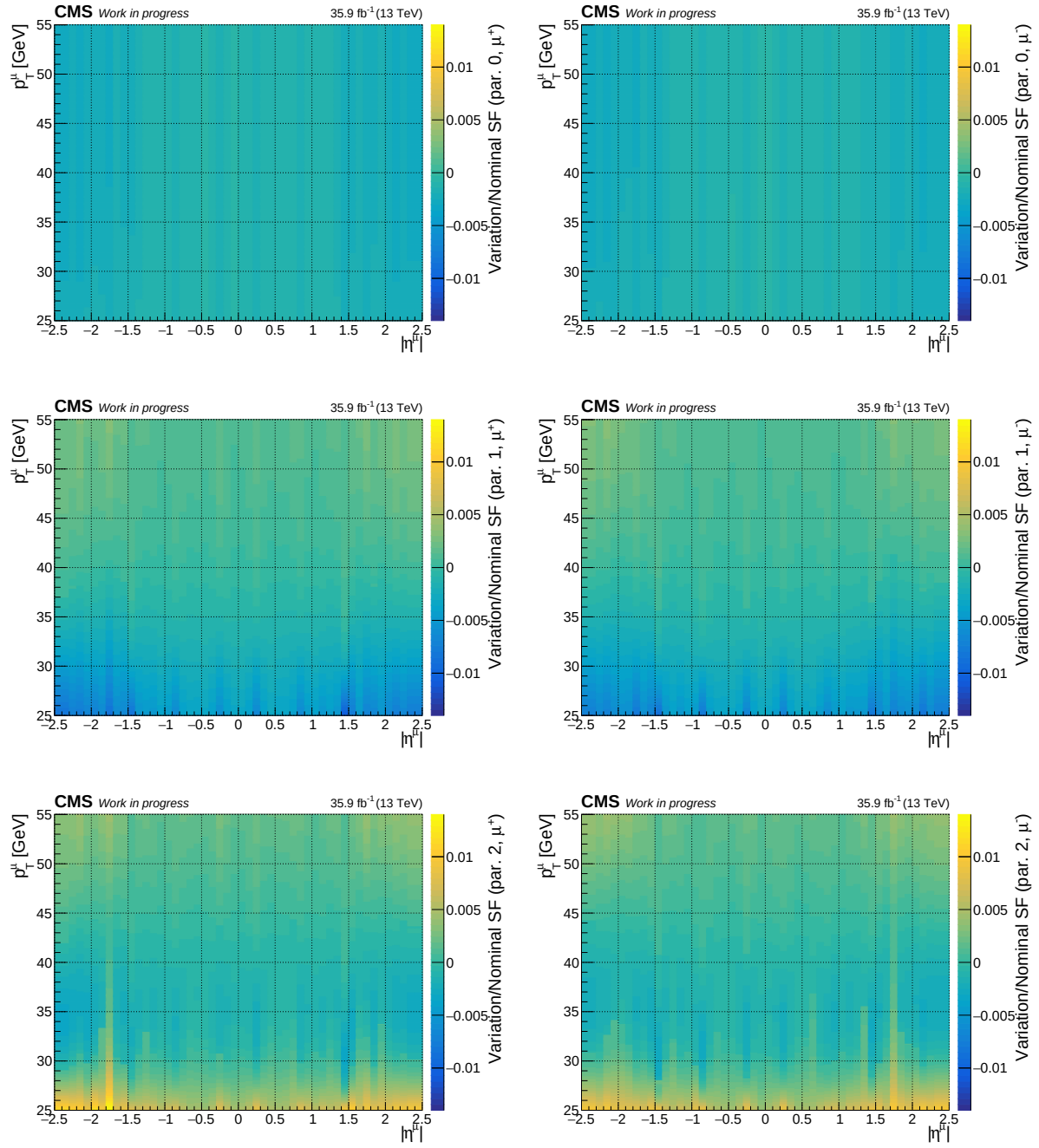


Figure 5.5. *Systematic uncertainties of the SFs due to the finite statistic of the adopted sample, expressed as $SF_{varied}/SF_{nominal}$ on $\eta^{\mu} \times p_T^{\mu}$ plane, separated in the contributions of the 3 parameter of the smoothing error function, inflated with the factor $\sqrt{2}$ as described in the text (μ^+ on the left column, μ^- on the right column).*

However, there are some effects that can produce a systematic uncertainty related to the PDF. The electroweak background templates are computed using the MC and are their shape and normalization varies according to the PDF choice. Also the QCD background, despite has been measured from data, uses the $W \to \mu\nu$ MC in the estimation procedure (see Sec. 6), therefore it has a PDF-dependence.

The fit is performed on a limited range in $|Y_W| \times q_T^W$ plane. There are $W \to \mu\nu$ events produced outside this range which fall in the acceptance region of the $\eta^\mu \times p_T^\mu$ of the templates. Their yield depends on the W production mechanism, ruled by the PDF in a x-region in which they are not measured by the fit. A PDF systematic uncertainty must be estimated also for this background source.

A secondary effect is given by the finite width of the $|Y_W| \times q_T^W$ bins. The shape of the templates is independent from the PDF only for a fixed value of Y_W and q_T^W, while the width of the bins produces an inevitable integration, producing a PDF shape dependence. However this effect is very mild and its impact can be tuned optimizing the $Y_W \times q_T^W$ binning.

The PDF uncertainty has been assessed by reweighing the MC with different PDF sets. The alternative sets have been provided from `NNPDF3.0` group [120] as 100 variations of each PDF, called replicas. A replica of a given PDF is a random sampling from the uncertainty band of each parameter of the functions which parameterize the PDF, i.e. statistically equivalent to the nominal set used in the analysis. These replicas have been converted in 60 eigenvectors in a Hessian representation. This description allows considering the 60 eigenvectors directly as nuisance parameters in the fit framework. The method from Ref. [193] has been exploited, which allows to convert the replicas into Hessian eigenvectors in an unbiased way, with a negligible information loss.

The α_s uncertainty is described with an up/down variation of its nominal value. Since it affects the observable distribution in the same way as the PDF it is grouped together with the 60 Hessian eigenvectors. The central value is $\alpha_s(m_Z) = 0.1180$, and has been shifted up and down of $\Delta\alpha_s(m_Z) = \pm 0.0015$.

The PDF and α_s systematic uncertainties are global weights applied to the event, therefore the induced variation is fully correlated in the $\eta^\mu \times p_T^\mu$ plane. A different variation is obtained for each Hessian eigenvector, plus the two extra α_s variations. These variations have been considered for the $W \to \mu\nu$ sample (signal templates and QCD background estimation) and for $W \to \tau\nu$ and $Z \to \mu\mu$ background samples.

5.4.2.2 W boson mass

m_W is a single parameter which correlates each variable of the analysis. The template distributions depend on the value of m_W encoded in the MC, as shown in Eq. 5.10 and 5.8. The A_i themselves encode a m_W dependence, as shown in Eq. 5.11. However the single distribution dependence is very mild, and the m_W sensitivity arise only by the simultaneous fit of the multi-differential distribution.

Two alternative sets of templates have been generated with a different m_W value, both for signal and background channels dependent on m_W. The central value encoded in the MC is $m_W = 80.419$ GeV, while the alternative m_W has been chosen with a conservative shift of ± 50 MeV using a Breit-Wigner assumption with the fixed nominal width $\Gamma_W = 2.0476$ GeV for the invariant mass distribution at generator level. In detail, defining:

$$\mathrm{BW}(m; M, \Gamma) = \frac{1}{(m^2 - M^2)^2 + M^2\Gamma^2} \qquad \text{the weights are: } w_{m_W^\pm} \equiv \frac{\mathrm{BW}(m; M \pm \Delta M, \Gamma)}{\mathrm{BW}(m; M, \Gamma)},$$

where $\Delta M = 50$ MeV. Note that this choice is completely equivalent to an introduction of a third dimension in the binning of the templates, with only 3 large bins. This approach, extending the number of m_W bins, should be used if m_W is fitted together to the production properties as discussed in Sec. 5.2.5

The use of pure Breit-Wigner to model the m_W distribution relies on the assumption of negligible PDF contribution on the invariant mass distribution. This assumption has been studied in detail in Chapter 3 and no additional systematic uncertainties are needed.

5.4.2.3 Final state radiation

The procedure of template building is independent of the simulation of the W boson kinematics. However, it relies on the description of the muon momentum in term of kinematics and propagation. From the theoretical side, this implies proper modelling of FSR, as discussed in Sec. 5.2.4. The available MC samples do not include enough information to perform a dedicated study of FSR at the time of this work and the uncertainty has been neglected.

The CMS W rapidity and helicity analysis assessed the FSR systematic uncertainty producing alternative $\eta^\mu \times p_T^\mu$ templates with a dedicated reweighing based on different modelling of FSR [11]. In particular, `PHOTOS 3.56` [194] has been considered instead of `PYTHIA` as a parton shower. The weights for the alternative templates are built with the ratio between `PHOTOS` and `PYTHIA`, binned in $r_{\text{FSR}} \equiv p_T^{\ell,\text{bare}}/p_T^{\ell,\text{dress}}$ (where $p_T^{\ell,\text{bare}}$ is the final state p_T^μ, while the radiation within a cone of $\Delta R = 0.1$ are added to the $p_T^{\ell,\text{bare}}$ to obtain $p_T^{\ell,\text{dress}}$). The induced systematic uncertainty is subleading both in measurement of unpolarized cross section as a function of Y_W or $A_4(Y_W)$, with a relative impact at 0.1%-3% level.

5.4.2.4 W and Z boson transverse momentum

The limited knowledge of q_T^W cannot bias the signal templates, due to the adopted procedure. However, some of the background templates are built relying on the W MC information. Therefore an uncertainty related to q_T^W and q_T^Z must be taken into account. To model the $q_T^{Z,W}$ uncertainty the renormalization scale μ_R and factorization scale μ_F of MC have been exploited.

Six alternative MC versions have been considered for the $Z \to \mu\mu$ sample. The sample has been reweighted multiplying the μ_R and/or μ_F for a factor 2 or 0.5, discarding the extremal case to satisfy $0.5 \leq \mu_i/\mu_j \leq 2$, as described in Ref. [195]. In detail the six variation are the following, as fraction of nominal value (μ_R, μ_F): (0.5,0.5), (0.5,1), (1,0.5), (1,2), (2,1), (2,2).

The $W \to \mu\nu$ sample used for QCD background estimation and $W \to \tau\nu$ sample has been reweighed in a more refined way, allowing the variation to describe a less constrained q_T^W dependence. The six aforementioned MC scale variations have been split in tree q_T^W bins each (low [0,5 GeV], mid [5 GeV,15 GeV], high [15 GeV,∞]), and an alternative reweighting is produced reweighting only the events in the considered bins for the 6 MC scale variations, for a total of 18 variations.

In both cases, these systematic variations are fully correlated in the $\eta^\mu \times p_T^\mu$ plane.

5.4.2.5 Electroweak background cross sections

An additional uncertainty on the electroweak cross sections of the background channels has been considered, following the approach of Ref. [11]. All these uncertainties are fully correlated in the $\eta^\mu \times p_T^\mu$ plane. A 4% uncertainty has been considered for $W \to \tau\nu$ to take into account the residual uncertainty related to the lower lepton momentum. A 6% and 16% uncertainty has been considered for top and diboson channels, respectively, to take into account the theoretical uncertainties related to the encoded cross section.

Chapter 6

Background estimation

The W boson mass and properties measurement has exploited a decay channel and an event selection which allow for a very high purity of the signal. Nevertheless, some background events are expected in the signal region. The muons of these background events can be produced both by electroweak sources or hadron decays. The modelling and the measurement on data of these background sources are discussed in this chapter.

6.1 Background sources generalities

The data sample collected by CMS in the 2016 data taking period, with event selection described in Sec. 5.3, results in a sample of $133 \cdot 10^6$ W^+ and $104 \cdot 10^6$ W^- candidates. This sample is mainly composed of W boson events, but some background events are expected. These backgrounds are not completely removed with the applied selection because their kinematic spectra partially overlap with the W boson signal. The yields of the background sources are summarized in table 6.1 and they can be subdivided in two categories:

- **Electroweak:** these backgrounds are produced by prompt muons from electroweak processes (like Drell-Yan, top quark decays, diboson production), which mimic the signature of the $W^{\pm} \to \mu^{\pm}\nu_{\mu}$ decay producing a single isolated muon in the signal kinematic region.

- **QCD:** these backgrounds are produced by non-prompt muons, loosely isolated by chance. The multijet production is the main source of this class of events, where the non-prompt muons are produced mainly in semileptonic decays of heavy-flavour mesons inside the jets.

The two classes of backgrounds have been treated with different approaches, which are described in the following sections together with the induced systematic uncertainties and their impact on the rest of the analysis.

6.2 Electroweak backgrounds

The electroweak backgrounds are produced by the following processes:

Table 6.1. *Estimated signal, background yields and background-to-signal ratio in the Signal region. The events are selected and weighted according to the signal selection described in Sec. 5.3. The electroweak processes are estimated directly from the MC simulation, while the QCD yield is estimated with the data-driven method described in this chapter.*

	Events/10^6		Bkg/Signal	
	W^+	W^-	W^+	W^-
Data	132.68	104.13	-	-
$W \to \mu\nu_\mu$	117.94	90.13	-	-
$W \to \tau\nu_\tau$	3.32	2.75	2.82%	3.05%
Drell-Yan ($Z/\gamma*$)	5.75	5.13	4.87%	5.69%
Top	0.77	0.72	0.65%	0.80%
Diboson	0.12	0.11	0.10%	0.12%
QCD	7.27	6.86	6.17%	7.61%

– **Neutral Drell-Yan**: $Z/\gamma^* \to \mu\mu$ decays can mimic a signal event when one of the two final state muons is not reconstructed because of identification inefficiency or because it falls out of acceptance (they are indicated as Drell-Yan events, hereinafter).

– **Top decays:** single top or $t\bar{t}$ production, with a single muon reconstructed in the final state (for instance in the decay $t \to bW(\to \mu\nu_\mu)$).

– **Diboson decays:** production of weak boson pairs (WW, WZ, ZZ), with a single muon reconstructed in the final state. This channel is considered a background despite a W boson can be actually promptly produced because of different kinematic properties compared to the signal.

– τ **decays:** $W \to \tau\nu_\tau$ or $Z \to \tau\tau$ production, with the subsequent decay $\tau \to \mu\nu_\mu\nu_\tau$. The p_T^μ spectrum from τ decays is softer compared to the signal, the m_W dependence is diluted from the extra intermediate state, and the spin correlation is diluted due to additional neutrino missing information. In conclusion, like the diboson processes, the τ channel has been considered, despite the prompt W production.

In all these processes the energetic muons are produced promptly and they can fall in the kinematic region of the signal. Thus, they cannot be removed with a kinematic selection only, but they must be statistically subtracted from the selected sample. The knowledge of the production yields of these processes and their kinematics is described with enough precision in the MC simulation of CMS, therefore they can be directly subtracted using the MC prediction in the signal region with small systematic uncertainty.

6.3 QCD backgrounds

Energetic muons are often produced in events with multiple jets in the final state. They are typically generated by heavy flavour meson decays in the jet environment and their yields exponentially decrease at high m_T. Moreover, the isolation requirement of the signal selection strongly suppresses the QCD muon yield. However, these muons can be produced at large angle with respect to the jet direction, according to the meson fragmentation function. Therefore there is a non-vanishing probability to have an high$-m_T$ event, with an isolated muon in the final state which falls in the signal region.

The MC simulation of the multijet production and of the parton showering is not accurate enough for the purpose of this measurement. In addition, this background source is composed of events in the kinematic tail of the QCD distribution generated by the Monte Carlo, therefore the statistical precision is very low (the use of the kinematic tail of a MC sample typically implies large luminosity weight $w_{\mathcal{L}}$ and thus an amplification of the statistical fluctuation of the few generated events). In conclusion, the simulation is not reliable for the QCD background evaluation and it has not been used in any part of the analysis.

In this section, a data-driven method to estimate the QCD background yield with the use of data and Electroweak MC only is described. This method, called *fake rate method* or *ABCD method* is often used to make a data-driven background estimation.

6.3.1 The *ABCD* method

Let us study a data sample composed by a signal (EWK) and a background (QCD) populations. Consider now two variables called v_1 and v_2, with some discriminative power of signal and background. In particular examine the simplified case where applying two orthogonal selection, $v_1 > v_1^*$ and $v_2 < v_2^*$, the signal efficiency is 100% (but not the purity). It is possible to define now four regions inverting the two selections, called A ($v_1 < v_1^*$, $v_2 > v_2^*$), B ($v_1 > v_1^*$, $v_2 > v_2^*$), C ($v_1 < v_1^*$, $v_2 < v_2^*$) and D ($v_1 > v_1^*$, $v_2 < v_2^*$), where D is the Signal region. These regions are displayed in Fig. 6.1. The event yields in the region i will be indicated as N_i. The purpose of the method is to estimate the background yield in D region.

These variables v_1, v_2 are chosen to be uncorrelated for the background, in such a way that the selection efficiency of the variable v_2 is independent from v_1, for the background yield. This background selection efficiency is typically called *fake rate* (f) and the aforementioned uncorrelation implies that:

$$f \equiv \frac{N_D}{N_D + N_B}\Big|_{\text{QCD}} = \frac{N_C}{N_A + N_C}\Big|_{\text{QCD}}, \tag{6.1}$$

which is called the *ABCD hypothesis*. Since the entire EWK component is in Signal region D and there is no EWK contamination in A, B, C regions, the left member of Eq. 6.1 can be directly evaluated from the total observed yield on data i.e.:

$$f = \frac{N_C}{N_A + N_C}\Big|_{\text{QCD}} = \frac{N_C}{N_A + N_C}\Big|_{\text{data}}. \tag{6.2}$$

Combining now the two previous equations:

$$N_D^{\text{QCD}} = f(N_D + N_B)_{\text{QCD}} \Leftrightarrow N_D^{QCD} = \frac{f}{1-f} N_B^{\text{data}} = \frac{N_C \cdot N_B}{N_A}\Big|_{\text{data}}. \tag{6.3}$$

Therefore in this simplified case, after the measurement of f in A and C regions, the QCD yields in D region can be obtained applying the weight $\frac{f}{1-f}$ to the observed events in B region.

In real data the EWK component populates all four regions, despite most of EWK events falling in D region. It results in the *prompt rate* p, defined as the v_2 selection efficiency of EWK events in B, D regions, lower than 1 (while in the Eq. 6.3 $p = 1$ artificially). The prompt rate can be measured from Monte Carlo simulation, because it is dependent only on the EWK processes which are well simulated. On the other hand, the fake rate must be measured form data, removing the contamination from EWK events in the A and C regions. In conclusion, they are measured as:

$$\begin{cases} p = \frac{N_D}{N_D+N_B}\Big|_{\text{EWK}} \\ \\ f = \frac{N_C^{\text{data}} - N_C^{\text{EWK}}}{(N_A+N_C)_{\text{data}} - (N_A+N_C)_{\text{EWK}}}. \end{cases} \tag{6.4}$$

After the measurement of f and p, the QCD yields in region D can be estimated as:

$$\begin{aligned} N_D^{\mathrm{QCD}} &= f(N_D + N_B)_{\mathrm{QCD}} \\ &= f(N_D + N_B)_{\mathrm{data}} - f(N_B + N_B)_{\mathrm{EWK}} \\ &= f(N_D + N_B)_{\mathrm{data}} - \frac{f}{p} N_D^{\mathrm{EWK}} \\ &= f(N_D + N_B)_{\mathrm{data}} - \frac{f}{p}(N_D^{\mathrm{data}} - N_D^{\mathrm{QCD}}), \\ N_D^{\mathrm{QCD}}\left(\frac{p-f}{p}\right) &= f\left[N_B^{\mathrm{data}} + N_D^{\mathrm{data}}\left(\frac{p-1}{p}\right)\right], \\ N_D^{\mathrm{QCD}} &= \frac{f}{p-f}\left[pN_B^{\mathrm{data}} - (1-p)N_D^{\mathrm{data}}\right]. \end{aligned} \tag{6.5}$$

In conclusion, QCD event yields in D region can be obtained reweighing the data events in D or B region according to the Eq. 6.5: if an event falls in B region the weight $\frac{f}{p-f}p$ is assigned, if it falls in D region the weight $-\frac{f}{p-f}(1-p)$ is assigned instead.

6.3.2 QCD background measurement

In the QCD background measurement, the transverse mass (m_T) has been chosen as v_1 variable and the relative isolation of the muon (RelIso) has been chosen as v_2. The missing transverse energy p_T^{miss} is a possible alternative to m_T, but the latter has more discrimination power between signal and background events.

The m_T cut and the RelIso cut are relaxed from the event selection described in Sec. 5.3.4. The sample selected in this way is composed of single-muon events with a very loose selection, called "Base-selection" (a summary of the Base-selection is listed in table 6.2). Considering the bidimensional distribution of RelIso of the muon versus m_T of the events, the phase space is divided in the aforementioned four regions: low transverse mass isolated and non-isolated (Sideband (C) and SidebandAiso (A) respectively) and high transverse mass isolated and non-isolated (Signal (D) and SignalAiso (B) respectively).

The RelIso = 0.15 value has been chosen as a discriminator between isolated and non-isolated region because the same value has been used in the SFs measurement, providing the proper calibration of the events of the isolated region. Moreover, RelIso = 0.15 threshold results in a background isolation efficiency in the sideband (A and C region) close to 50%, the optimal value to minimize statistical errors. The RelIso < 0.15 selection has a high efficiency on data, about 87%, applied on Base-selection.

The $m_T = 40$ GeV value has been chosen as the lower limit of signal regions to be a value high enough to obtain a strong QCD yields suppression. On the other hand, the value $m_T = 30$ GeV value has been chosen as the upper limit of sideband regions because of the lower derivative of the m_T distribution at that point (compared for instance to $m_T = 40$ GeV), which reduces the dependence from systematic uncertainties related to the m_T cut (see further discussion in Sec. 6.3.3). The $m_T > 40$ GeV selection, applied on the top of Base-selection has an efficiency of about 81% on data.

The four region regions are displayed in Fig. 6.1 on the top of data with the Base-selection applied, while the numerical value of the cuts which define the regions are listed in table 6.2. The combined efficiency of m_T and RelIso selection which define the Signal region has an efficiency of about 75% on data.

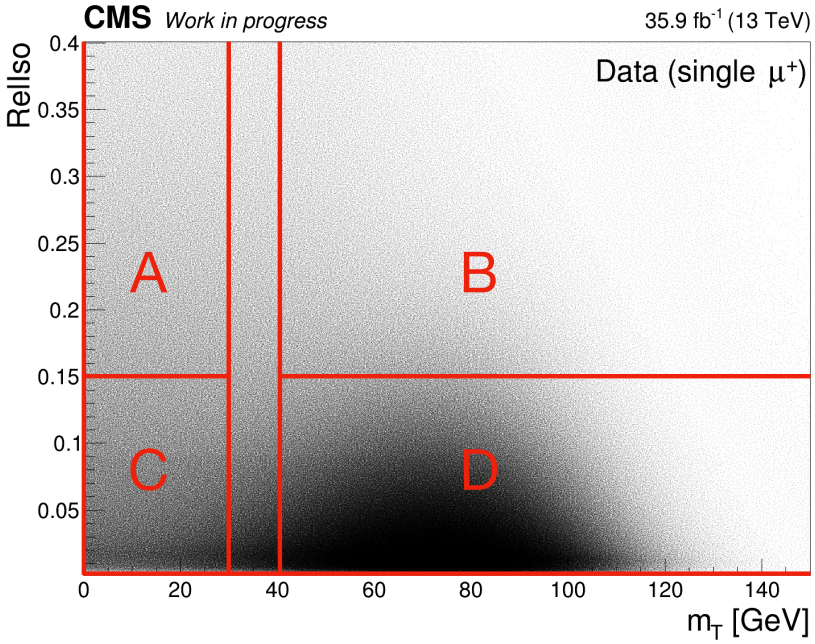


Figure 6.1. *Event distribution in the W^+ data sample on the plane RelIso $\times$ m_T with the Base-selection applied. The four regions of ABCD method are highlighted, the isolation cut is set to RelIso=0.15, the transverse mass cuts are set to $m_T = 30$ GeV and $m_T = 40$ GeV.*

Table 6.2. *List of selection applied to the regions used in the ABCD Method. "Base-selection" region represent the preliminary very loose selection*

Region	Selection applied
Base-selection	HLT_Iso(tk)Mu24, $PV_z < 24$ cm, $PV_{xy} < 2$ cm, MediumID,$\lvert d^{\mu}_{xy}\rvert < 0.05$ cm, $\lvert d^{\mu}_{z}\rvert < 0.2$ cm, MET-filters, $N_{\text{veto e}} = 0$, $\lvert\eta^{\mu}\rvert < 2.4$, 25 GeV$< p^{\mu}_{T} <$ 55 GeV
SidebandAiso (A)	Base-selection, $m_T < 30$ GeV, RelIso>0.15
SignalAiso (B)	Base-selection, $m_T > 40$ GeV, RelIso>0.15
Sideband (C)	Base-selection, $m_T < 30$ GeV, RelIso<0.15
Signal (D)	Base-selection, $m_T > 40$ GeV, RelIso<0.15

The sum of MC $W \to \mu\nu$ events and the MC electroweak background events are indicated as EWK sample in the framework of the QCD background measurement. This is analogous to the case discussed in Sec. 6.3.1, where the data sample is the sum of EWK and QCD only.

In this specific case, ABCD hypothesis of Eq. 6.1 is equivalent to consider the isolation efficiency in the QCD events independent from the transverse mass of the event. This is reasonable because the isolation of the muon from heavy flavour meson decays is expected to have little correlation with a variable like m_T, which has a physical meaning only in W events. Despite this intuition, this assumption has been validated and discussed in Sec. 6.3.3. In particular, this assumption implies that the fake rate f, the isolation efficiency in QCD events in the Signal region, can be measured in the Sideband regions (A and C) and then applied in the Signal regions (B and D).

The estimation of the QCD yields in the Signal region has been performed in two steps, following the approach of the ABCD method:

1. Measurement of the fake and prompt rate.

2. Reweighing of the data sample in Signal and SignalAiso regions according to the Eq. 6.5, to produce the QCD background templates required for the fit to the W production properties.

The entire background analysis has been performed separately for the W^+ and W^- samples. Most of the intermediate steps are shown for W^+ sample only, but they are very similar also for the W^- sample case.

The Base-selection includes the reweighting procedures described in Sec. 5.3.5 for all MC samples, but the efficiency SF. The latter is applied in the isolated regions (C, D) only. In the non-isolated regions no efficiency SFs are applied, because $\mathrm{SF}_{\mathrm{sel}}$ includes the muon isolation selection component. Therefore SFs for non-isolated kinematic regions cannot be reliable. The QCD yields have been estimated by applying the SF in non-isolated regions as well to the isolated ones, to test the robustness of this approach, and the discrepancy between the two methods is below the statistical uncertainty.

6.3.2.1 Fake rate measurement

To perform the measurement of the fake rate the EWK and data samples have been binned

- in η^μ, between $\eta^\mu = -2.4$ and $\eta^\mu = 2.4$ with constant width of $\Delta\eta = 0.1$,
- in p_T^μ bins, between $p_T^\mu = 25$ GeV and $p_T^\mu = 55$ GeV with constant width of $\Delta p_T = 0.5$ GeV,

for a total of 48×60 $\eta^\mu \times p_T^\mu$ bins. The selection of SidebandAiso A region and Sideband C region are applied for each bin. Then f has been measured for each bin accordingly to Eq. 6.4.

Afterwards, the f values have been fitted linearly in p_T^μ, to obtain smoother results in the reweighing procedure. The functional form used is

$$f(p_T) = m(p_T - 25[\mathrm{GeV}]) + q, \tag{6.6}$$

where m, q are the fitted parameters and the 25 GeV shift has been applied to reduce the correlation between the two parameters.

The linear fit has been chosen as the simplest which is able to reproduce $f(p_T)$, in term of χ^2. The reduced χ^2 of the fits are compatible with 1 within the uncertainties. More complex functions have been tested, but the additional parameters are found to be compatible with 0 without significant variation of the central value. The fit has been performed independently for each η^μ bin, therefore the outcome of the fake rate measurement are $2 \times N_{\eta\ \mathrm{bins}}$ parameters and their covariance matrices.

6.3.2.2 Prompt rate estimation

To perform the measurement of the prompt rate, the EWK sample has been binned with the same bins discussed before. The selection of SignalAiso B region and Signal D region are applied for each bin, and p has been measured for each bin accordingly to Eq. 6.4.

Then the p values have been fitted in p_T^μ with an error function, to obtain smoother results in the reweighing procedure. The functional form used is:

$$p(p_T) = A\frac{2}{\sqrt{\pi}}\int_0^{Bp_T+C} e^{-t^2}dt \tag{6.7}$$

where A, B, C are the fitted parameters. The reduced χ^2 of the error function fits are in good agreement with 1. This functional form has been chosen since the prompt rate is the isolation

efficiency measured in the signal regions on the MC, a quantity that is typically modelled with a turn-on function like the one in Eq. 6.7. Moreover, the error function has been also used to smoothen the trigger and selection efficiency used to build SF_{trig} and SF_{sel}. Furthermore, several alternative functions have been tested: simpler models (like 1,2,3,4 degrees polynomials) produce significantly worse results in term of χ^2, more complex models (like an error function multiplied with a polynomial) do not produce significant improvement of χ^2 and the additional parameters result compatible with 0.

As for the fake rate fit mentioned in the previous section, the prompt rate fit has been performed independently for each η^μ bin and the outcome are $3 \times N_{\eta \text{ bins}}$ parameters and their covariance matrices.

In Fig. 6.2(a) the fake rate and the prompt rate distributions in a single η^μ bin are shown together with the fitting functions. In Fig. 6.2(b) the χ^2 of the described fits are also shown.

6.3.2.3 QCD estimation

The QCD yields can be estimated subsequently to the measurement of the fake and prompt rate, according to Eq. 6.5. The proper weights are applied to data sample events in Signal D or SignalAiso B region, separated by change. The weights are evaluated on event by event basis, in bins of η^μ, with the same binning used in the fake and prompt rate, using the $m(\eta), q(\eta), A(\eta), B(\eta), C(\eta)$ parameters from Eq. 6.6 and 6.7 and the event p_T^μ.

With this reweighing approach, every variable of interest can be estimated for QCD events. For example in Fig. 6.2(c) the QCD yields in a single bin of η^μ, function of p_T^μ, are shown.

6.3.3 QCD background systematic uncertainties

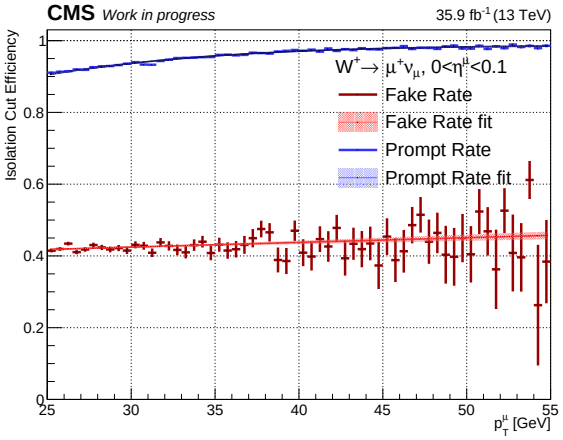
The QCD background estimation described in the previous sections is affected by several sources of systematic uncertainty. They can be grouped in two categories depending on their source:

- **Input-related systematic uncertainties:** induced by the modelling and the definition of the experimental quantities and the theoretical assumption underlying the adopted variables to estimate the QCD yields.
- **Strategy systematic uncertainties:** induced by the QCD yields estimation method itself and independent from the input variable of the method.

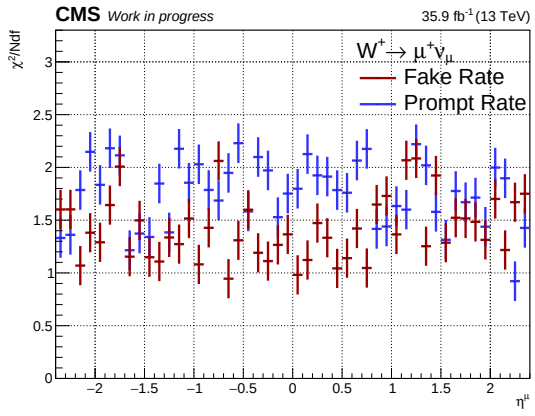
6.3.3.1 Input-related systematic uncertainties

The QCD yield estimation is strongly dependent on the accuracy of the MC simulation of the electroweak processes. They enter directly in a large subtraction in Sideband regions for fake rate measurement, and indirectly with a factor $(1 - p)$ as a weight to Signal D region events, very sensitive to small changes of prompt rate p value. Several high-level experimental variables have been used in the QCD yield estimation with the ABCD method: m_T, RelIso, η^μ, p_T^μ and the additional variables which enter in the Base-selection listed in table 6.2. In principle, all the systematic uncertainties discussed in Sec. 5.4, from the experimental and theoretical side, must be taken into account. For experimental systematic uncertainty, it means that a set of variations of the nominal variables must be considered. For the theoretical systematic uncertainty, the proper weight must be applied to the events resulting again in a set of variations of the considered variables.

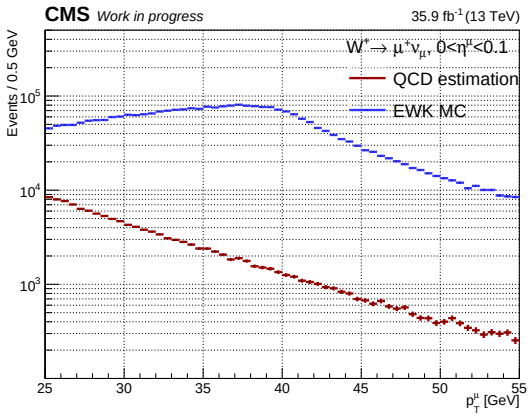
To estimate the impact of each variation on QCD yield estimation, the variable (weight) variation under study has been applied, keeping all the other variables (weights) to the nominal value. The



(a)



(b)



(c)

Figure 6.2. *(a): transverse momentum distribution of fake and prompt rate in* $0 < \eta < 0.1$ *bin in* W^+ *sample, the shaded red and blue band represent the uncertainties from the respective fits. (b): reduced* χ^2 *from the linear fit to the fake rate and error function fit to the prompt rate for the entire set of* η^μ *bins for* W^+ *sample. (c): QCD events transverse momentum distribution in Signal region estimated for* W^+ *sample, as comparison also the EWK component from MC is shown.*

result is a varied version or A,B,C,D yields for the EWK sample (while the data are not affected by the systematic variations by definition), in each bin of p_T^μ and η^μ. The fake rate and prompt rate are measured again as described in Sec. 6.3.2 with this alternative sample, producing a set of varied parameters. With this set, the usual reweighing of the data sample can be performed to obtain the desired QCD distribution, which can be compared to the nominal QCD distribution to assess the impact of the variation under study.

The following variations have been propagated to the QCD yields estimation:

– **MET:** JES up/down and the unclustered energy up/down have been considered, for all the MC samples, fully correlated in the $\eta^\mu \times p_T^\mu$ plane.

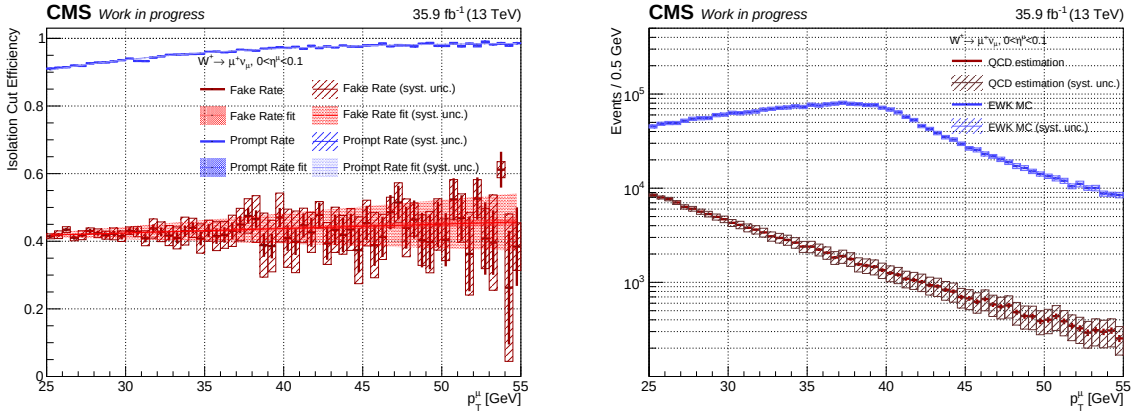
– **Muon Momentum Scale:** Rochester corrected p_T^μ up/down have been considered, for all the MC samples, fully correlated in the $\eta^\mu \times p_T^\mu$ plane.

– **Efficiency Scale Factors:** the set of variation provided by Ref. [11] have been considered

i.e. one up/down variation fully correlated in the $\eta^\mu \times p_T^\mu$ plane and 3 up/down variations uncorrelated in η^μ, for all the MC samples.

- **L1 Trigger Prefire:** one up/down variation have been considered, for all the MC samples, fully correlated in the $\eta^\mu \times p_T^\mu$ plane.
- **Lepton Veto:** one up/down variation of the Z event yield of 2% has been considered, flat in the entire phase space.
- **Data Luminosity:** one up/down variation of 2.5% has been considered for all the samples, flat in the entire phase space.
- **PDF and α_s:** 60 PDF variation and one α_s up/down variation have been considered, for $W \to \mu\nu$, Z, and $W \to \tau\nu$ samples, fully correlated in the $\eta^\mu \times p_T^\mu$ plane.
- W **boson mass:** a 50 MeV up/down variation of m_W has been considered for the signal sample only, applying the weight as described in Sec. 5.4.
- W **boson** q_T**:** 18 q_T^W uncorrelated MC scale variation, as described in Sec. 5.4, has been considered for $W \to \mu\nu$ and $W \to \tau\nu$ samples. 6 MC scale variation, as described in Sec. 5.4, has been considered for Z sample only.
- **Electroweak cross sections:** an up/down variation of 4%, 6% and 16% has been considered for $W \to \tau\nu$, top and diboson samples, respectively, flat in the entire phase space.

In Fig. 6.3 the effect of the systematic uncertainties on fake rate and on the QCD yields estimation is shown, in a single η^μ bin. The systematic uncertainty bands are obtained from independent fits performed on each variation. The band is built summing in quadrature the half-difference up-down for each systematic uncertainty (or the difference between nominal and the varied values, for PDF and MC Scales variations). The same result in the entire range of η^μ is shown in Fig. 6.4.



(a) (b)

Figure 6.3. *(a): transverse momentum distribution of fake and prompt rate in $0 < \eta^\mu < 0.1$ bin in W^+ sample, the dashed box represents the systematic uncertainty band for each p_T^μ bin (pre-fit), the central dark shaded red and blue band represent the statistical uncertainties from the respective fits, while the light shaded red and blue band represent the systematic uncertainty band after the fit. (c): QCD events transverse momentum distribution in Signal region estimated for W^+ sample compared with the EWK MC, in $0 < \eta^\mu < 0.1$ bin in W^+ sample; the systematic uncertainties band is shown both for QCD yields estimation and for EWK MC as dashed boxes.*

The fake rate is lower in the central region and about 50% higher in the large η^{μ} region, with a smooth modulation in between.

The fitted value of the fake rate m (slope) and q (offset) parameters from Eq. 6.6 for each bin of η^{μ} and for the two boson charges are shown in Fig. 6.5. The two parameters have a very low correlation and the agreement between the two boson charges is good. The dominant source of systematic uncertainty arises from the slope, while the relative spread of the offset is much smaller. The systematic uncertainty band has been estimated as previously described for the fake rate itself. The η^{μ} trend of the fake rate is produced by the modulation of the offset parameter, while the slope is stable.

In Fig. 6.6 the relative shift of each class of systematic uncertainty is shown in a single η^{μ} bin. The systematic uncertainty relative shift has a strong p_T^{μ} dependence on fake rate, which reflects at QCD p_T^{μ} spectrum level. The dominant sources of uncertainty are the p_T^{miss}-related variations, which reach up to 15%(30%) discrepancy at the fake rate (QCD yields estimation) level. The q_T^W and PDF uncertainties are the main subleading systematic variations together with luminosity and reach a 6% level in the fake rate estimation. However, on the QCD p_T^{μ} spectrum they are at the same level of the statistical uncertainty with a shift up to 15%. All the other systematic variations are below the statistical uncertainty. The relative impact of the variation on the prompt rate is at 0.1%-0.01% level, but it can not be neglected because of the dependence from the fake rate on the QCD yields in Eq. 6.5, and it contributes to the uncertainty shown in Fig. 6.3(b). Fig. 6.7 reports the same result extended in the entire range of η^{μ}. The trend of the systematic uncertainty is reasonably flat in η^{μ} and there are no particular features neither in the fake rate, prompt rate nor in the QCD yields. The W^- sample shows similar trends for all the discussed results.

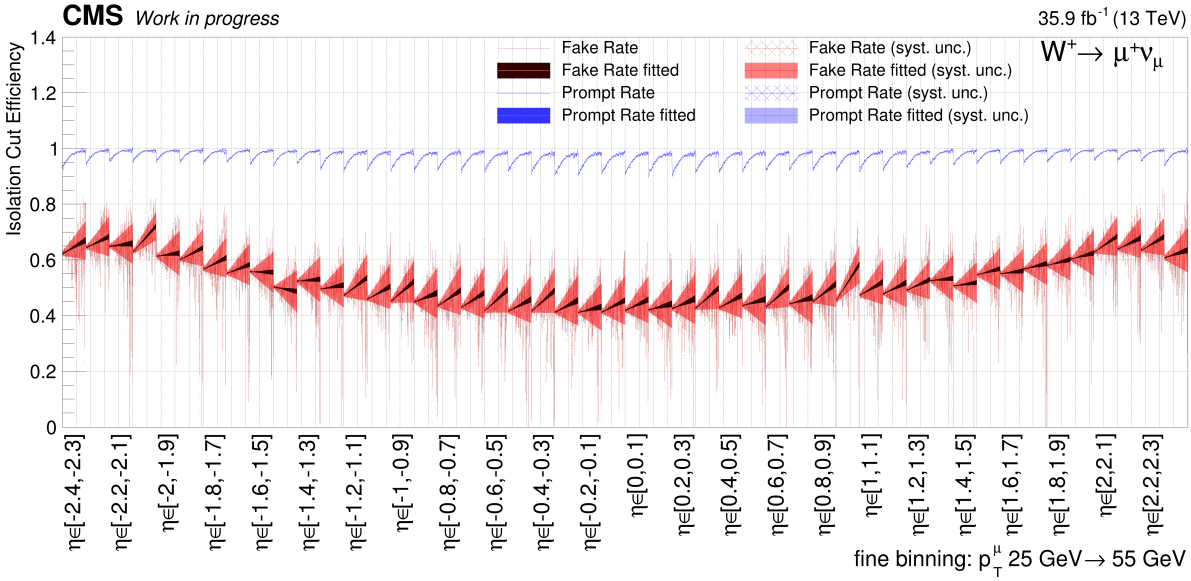
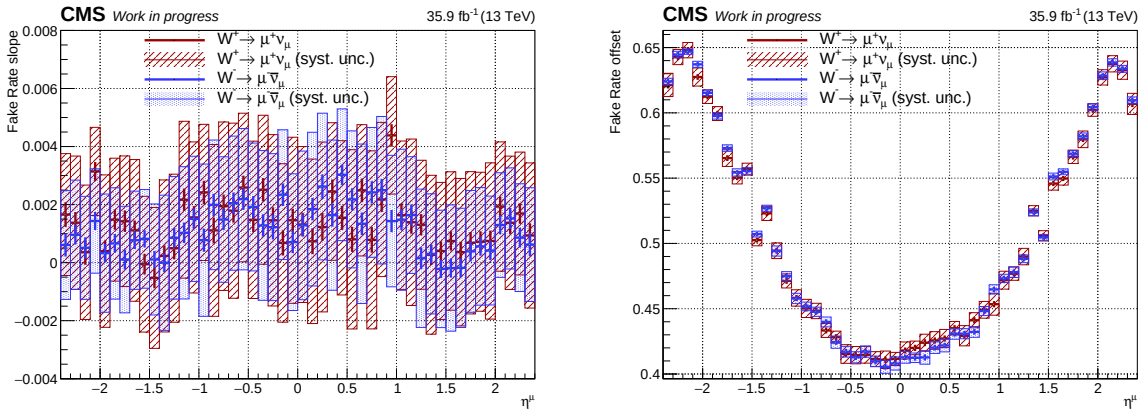
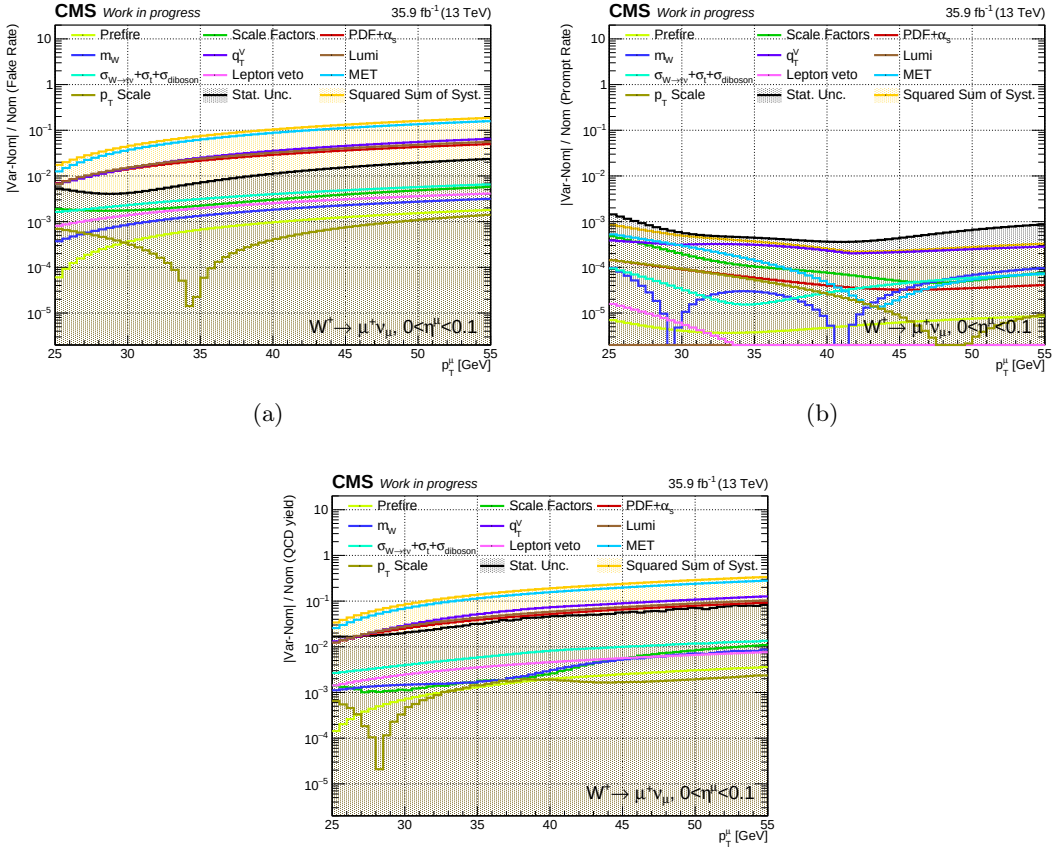


Figure 6.4. *Transverse momentum distribution of fake rate and prompt rate for each η^{μ} bin. The systematic uncertainty band is shown in lighter colors, while the nominal value of fake and prompt rate is shaded in darker color within its statistical uncertainty.*



(a) (b)

Figure 6.5. *Value of the slope (a) and offset (b) fake rate parameters as a function of η^{μ} for W^{+} and W^{-} samples. The shaded bands show the spread of parameters given by the systematic uncertainties*



(a) (b)

(c)

Figure 6.6. *Relative shift of the systematic uncertainties, subdivided in classes, on the fake rate (a), prompt rate (b) and QCD yields (c) in $0 < \eta^{\mu} < 0.1$ bin in W^{+} sample. The shift is evaluated as the sum in quadrature of the relative discrepancy |up-down|/nominal (or |variation-nominal|/nominal for PDF and MC Scales variations).*

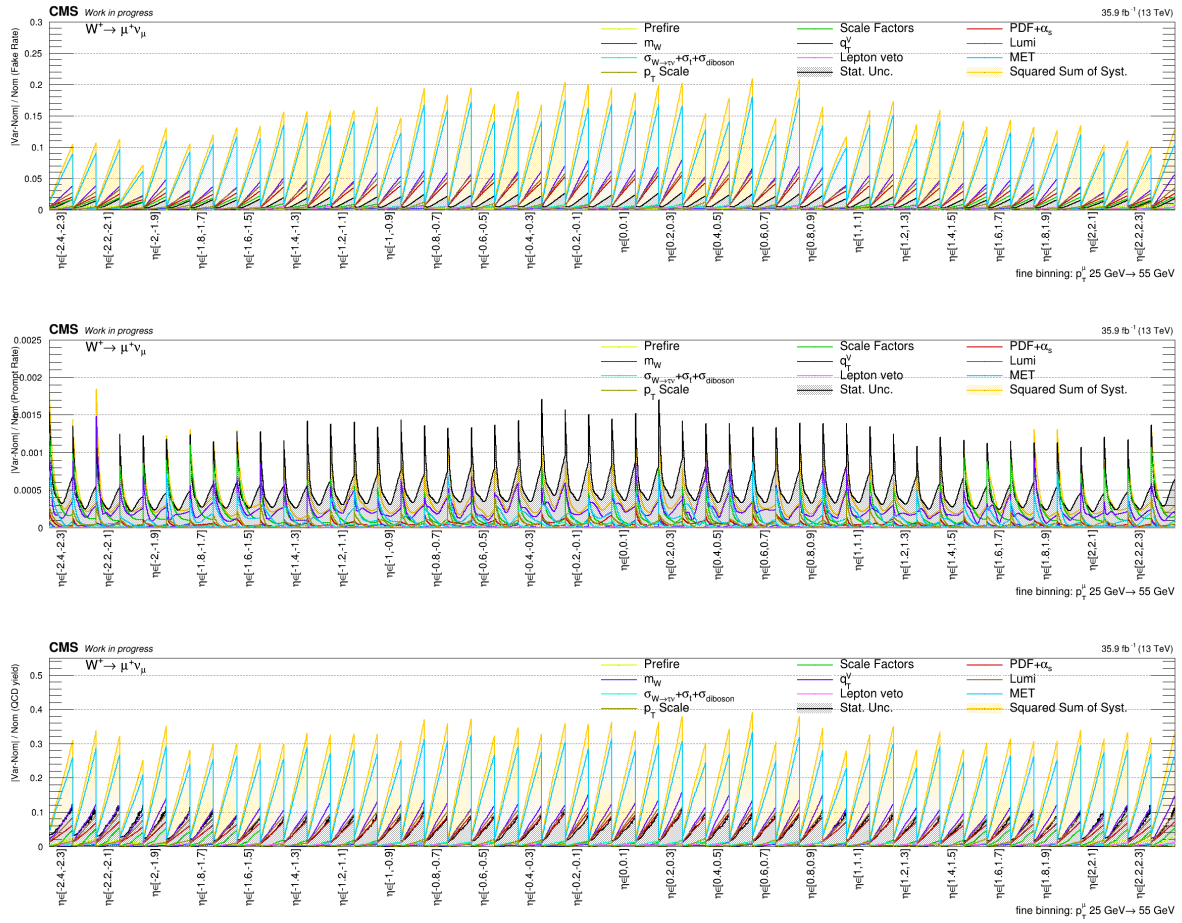


Figure 6.7. *Relative shift of the systematic uncertainties, subdivided in classes, on the fake rate (upper plot), prompt rate (central plot) and QCD yields (lower plot) in W^+ sample, unrolled in η^μ, p_T^μ. The shift is evaluated as the sum in quadrature of the relative discrepancy |up-down|/nominal (or |variation-nominal|/nominal for PDF and MC Scales variations).*

6.3.3.2 Strategy systematic uncertainties - Fit

Assumptions made in the procedure of QCD yields estimation can induce an additional systematic uncertainty. The robustness of the fake rate p_T^μ linear fit has been tested as described in Sec. 6.3.2.1, and thus no additional systematic uncertainty has been considered. The robustness of the error function p_T^μ fit to the prompt rate has been discussed in Sec. 6.3.2.2, and no additional systematic uncertainty has been considered.

6.3.3.3 Strategy systematic uncertainties - ABCD method

In the ABCD method description both the RelIso cut and m_T cut are arbitrary and therefore they potentially induce systematic uncertainties. The RelIso selection directly affects the Signal Region, therefore a change in QCD yields estimation is expected. The value of the RelIso cut is motivated by the calibration of the efficiency SFs already discussed in Sec. 6.3.2. The same argument can be used for $m_T = 40$ GeV cut, but the QCD yields estimation in the Signal region must be independent of the $m_T = 30$ GeV cut, which defines Sideband A and C regions. This assumption, which reflects in Eq. 6.1, and has been investigated in detail.

The non-correlation between RelIso and m_T can be directly verified using the MC simulation of QCD. As previously discussed this MC is not reliable for an estimation of the QCD yield in the Signal region or to assess the accurate value of the fake rate. However, the isolation efficiency and m_T are non-correlated variables in multijet production, from a physical point of view. Therefore the isolation efficiency trend as a function of m_T can be studied on the QCD MC to check this non-correlation and ensure that the ABCD assumption is not neglecting physical features of the QCD trend in m_T. The QCD MC events have been selected with the Base-selection and then the fake rate is estimated as

$$f_{\text{MC QCD}}(m_T) = \frac{N_{\text{RelIso}<0.15}}{N_{\text{tot}}}(m_T).$$

The value of the fake rate is shown in Fig. 6.8 in different ranges of η^μ and p_T^μ, and integrated in the entire $\eta^\mu \times p^\mu$ phase space. The constant trend is confirmed, despite the low statistical power of the QCD MC produces large fluctuations. To overcome this issue $f(m_T)$ has been fitted with a constant or linear function. The slope of the linear function results compatible with 0, without relevant χ^2 improvement. In Fig. 6.8 the fit of the integrated $f(m_T)$ in the two cases is shown.

The possibility to verify the fake rate constant behaviour in m_T directly on data has been also investigated. It cannot be directly verified in Signal Region because the expected QCD yields are very low and the required subtraction of the prompt rate produces not affordable uncertainties. Therefore the sideband regions (A, C) has been sliced in 3 m_T sub-region to evaluate the fake rate from data with the usual procedure and to investigate a possible m_T-trend between the 3 slices. Unfortunately, the data statistical power in this region is not sufficient to perform a study with the required level of prevision and the proper $\eta^\mu \times p_T^\mu$ binning. However extrapolating the value of the fake rate measured in the 3 sideband bins in the signal region the discrepancies are below 5-10% and localized in the high-p_T region only, probably due to the instability of the p_T^μ

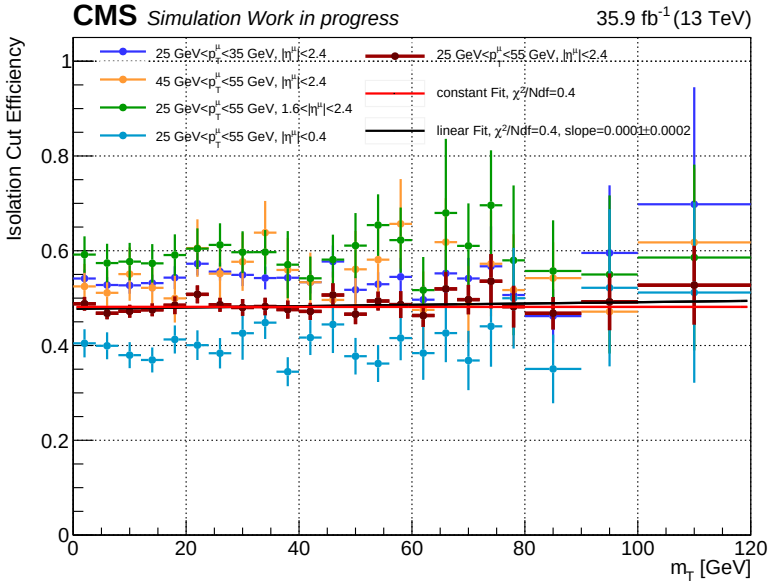


Figure 6.8. *Fake rate estimated from the QCD MC to validate the ABCD hypothesis and corroborate the assumption of no correlation between the isolation efficiency and m_T. The Base-selection is applied as pre-selection to the simulated events, and then additional selection on p_T^μ and η^μ has been applied to highlight the behaviour in different phase-space regions. The result of a constant and linear fit on the integrated $f(m_T)$ is also shown.*

fit with a low statistic. This discrepancy is largely covered by the already considered systematic uncertainties.

In conclusion, no additional systematic uncertainties have been assigned to the extrapolation between the sideband region and the signal region.

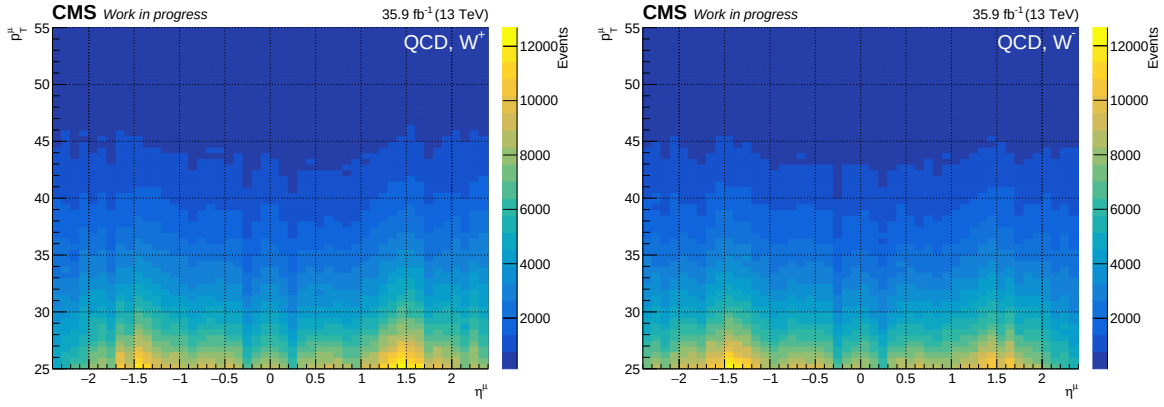
6.4 Background templates

The ABCD method described in the previous section provides a set of parameters that allow to estimate the QCD yields distribution of each variable of interest. The set is realized with m, q parameters of linear fit to $f(p_T^\mu)$ and A, B, C parameters of the error function fit to $p(p_T^\mu)$. The parameters are provided with the covariance matrix of their statistical uncertainties, due to the finite statistics of the data and EWK sample used in their estimation. The parameters are provided in bins of η^μ for W^+ and W^- samples separately.

Given the fake and prompt rate parameters set it is possible to build the QCD background template for W^+ and W^- samples. The data events have been selected following the procedure described in Sec. 5.3 but the RelIso selection. Then the events have been reweighted accordingly to the formula of Eq. 6.5, based on the RelIso of the muon of the event. Finally, a $\eta^\mu \times p_T^\mu$ template has been filled. The chosen binning is aligned with the binning used for the background parameter estimation i.e. 48 η^μ bins between $-2.4 < \eta^\mu < 2.4$, with $\Delta\eta^\mu = 0.1$ and 60 p_T^μ bins between 25 GeV $< p_T^\mu <$ 55 GeV, with $\Delta p_T^\mu = 0.5$ GeV. The resulting templates are shown in Fig. 6.9. The two charges have a good agreement, as shown in Fig. 6.10, where the QCD background p_T^μ distributions of the two charges are compared, integrating the templates in η^μ in the considered range. The discrepancies arise both from inaccuracies of EWK MC subtraction, whose SF are different for the two charges, and from the charge asymmetry of QCD background distributions.

Alternative variations of the parameter set are also provided for each experimental and theoretical systematic in addition to the nominal set of parameters. The complete list and the η^μ correlation of these variations are described in Sec. 6.3.3 in detail. The summary of the systematic uncertainty magnitude on QCD yields is provided in table 6.3.

These variations allow to produce alternative versions of the QCD events distribution, with the same scheme of the MC Signal and EWK backgrounds samples. With this approach, alternative templates can be provided to the template fit to W production properties to properly take into



(a) (b)

Figure 6.9. *QCD background templates for W^+ (a) and W^- (b) selection.*

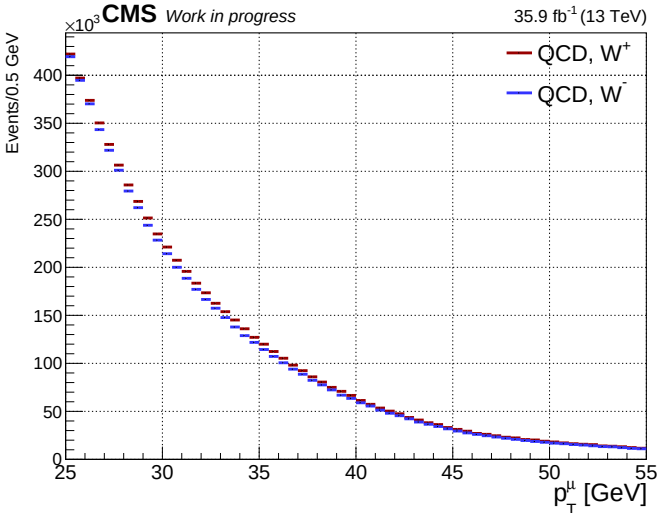


Figure 6.10. *p_T^μ projection of QCD background templates, integrating on η^μ considered range, comparing the two charges. The systematic uncertainties are not shown.*

Table 6.3. *Summary of the systematic uncertainty which affect the background estimation and the relative induced variation; the statistical uncertainty is reported for comparison.*

Assumption	Approach	yield variation (%)				
		QCD	$Z/\gamma*$	$W \to \tau\nu_\tau$	Top	Diboson
Exp. - p_T^{miss}	input variation	2-30	1	0.5	0.5	0.5
Exp. - p_T^μ scale	input variation	0.1-2	0.01	0.01	0.01	0.01
Exp. - Efficiency SF	input variation	0.1-5	0.5-1	0.5-1	0.5-1	0.5-1
Exp. - L1 Trigger Prefire	input variation	0.1-2	0.1-0.5	0.1-0.5	0.5-1	0.1-0.5
Exp. - Lepton Veto	input variation	0.1-1	2	-	-	-
Exp. - Luminosity	input variation	1-10	2.5	2.5	2.5	2.5
Theo. - PDF, α_s	input variation	1-10	2	2	-	-
Theo. - m_W	input variation	0.1-1	-	-	-	-
Theo. - $q_T^{W,Z}$	input variation	2-15	5	3	-	-
Theo. - σ_t, σ_τ, σ_{diboson}	input variation	0.1-1	-	4	6	16
$f(M_T) = \text{const}$	extrapolation study	0-10	-	-	-	-
$f(p_T)$ =linear	function variation	0	-	-	-	-
$p(p_T)$ =erf	function variation	0	-	-	-	-
finite data statistics	uncertainity propagation	2-10	1-5	2	6	20

account the systematic uncertainties of the fit. No additional dedicated templates have been produced to propagate the statistical uncertainty of the fake and prompt rate to the QCD yield, since this is a subleading source of systematic uncertainty and if considered would require a full decorrelation in $\eta^\mu \times p_T^\mu$ plane.

The electroweak background templates, built directly from MC, will be shown in Chapter. 7. The variation of the electroweak background yields induced by statistical and systematic uncertainties are reported in table 6.3 together with the QCD background yield variations. The treatment of the systematic uncertainties in the template fit will be discussed in detail in Chapter 7.

The closure between data and MC in the Signal Region has been checked using the QCD yields estimated from this measurement, the electroweak background estimation from MC and the

Signal sample MC. This step is needed to understand the agreement between the MC description that will be used to build the templates of the W mass and properties fit and the data distribution preliminary to the fit itself. Figure 6.11 shows the p_T^μ, η^μ and m_T distributions of data and stacked MC samples for W^+ and W^- selections. The agreement is satisfactory overall. From the η^μ distribution it can be seen that the MC and data disagreement is below 3% and almost flat in η^μ. In the p_T^μ distribution there is a clear linear trend, with up to 5% of disagreement. In the m_T distribution, the disagreement is about 10% at higher m_T values. The systematic uncertainties cover this disagreement in the full η^μ and m_T range, while in the low p_T^μ region the discrepancy in the ratio is significantly different from 1. The p_T^μ trend of the discrepancy may suggest a mismodelling of the q_T^W spectrum of the $W \to \mu\nu$ MC, and is not correlated with the QCD yield.

The systematic uncertainty band breakdown is shown in Fig. 6.12, for p_T^μ, η^μ and m_T distributions. The relative impacts of the various classes of systematic uncertainties are similar to the QCD events distributions from Fig. 6.7. The p_T^μ and η^μ trend of the systematic impact is almost flat, dominated by PDF uncertainty, as expected. Also the luminosity produces a relevant contribution, as expected for a differential measurement. On the other hand, the p_T^{miss} uncertainty is the leading systematic uncertainty in the high-m_T region, where the low statistics makes the ratio more sensitive to the missing energy calibration effects. The relative variation induced by the QCD background systematic uncertainties only is shown in Fig. 6.12(d), suppressing the variation induced by the EWK backgrounds and the Signal. In this case the shift is at $7 \cdot 10^{-3}$ level, dominated by p_T^{miss} and q_T^W uncertainties. This highlights the small impact of the QCD background systematic uncertainty on the cumulative distribution, thanks to the precise and accurate description derived in this chapter.

6.5 Closure test

The procedure of QCD background estimation can be validated by evaluating the QCD background yield in the Sideband region C. In this case, the fake rate will be estimated with the usual procedure, while the prompt rate will be evaluated from EWK MC in A and C regions. By performing this estimation the same events are used to measure the fake rate and to predict the QCD yield using the relation of Eq. 6.5 (with the replacement $N_B^{\text{data}} \to N_A^{\text{data}}, N_D^{\text{data}} \to N_C^{\text{data}}$), therefore the values close mathematically if the procedure is consistent. However, the fake and prompt rates have been smoothed with a linear fit or with an error function fit, respectively. Therefore some discrepancy between data and the sum of all the MC are expected.

The result of this closure test is shown in Fig. 6.13. Since the Sideband region is enriched in QCD is it possible to better see the impact of the fake rate features on the muon variables spectra. The closure shows a very good agreement, with discrepancy below the 1% level in the in m_T and η^μ distributions. Discrepancies at 2%-4% level are observed in the p_T^μ spectrum at mid-high p_T^μ induced by the smoothing.

An additional conservative 5% flat systematic uncertainty on the QCD yield has been considered to cope with residual discrepancies. This in particular will cover also the possible bias due to the extrapolation from Sideband to Signal Region.

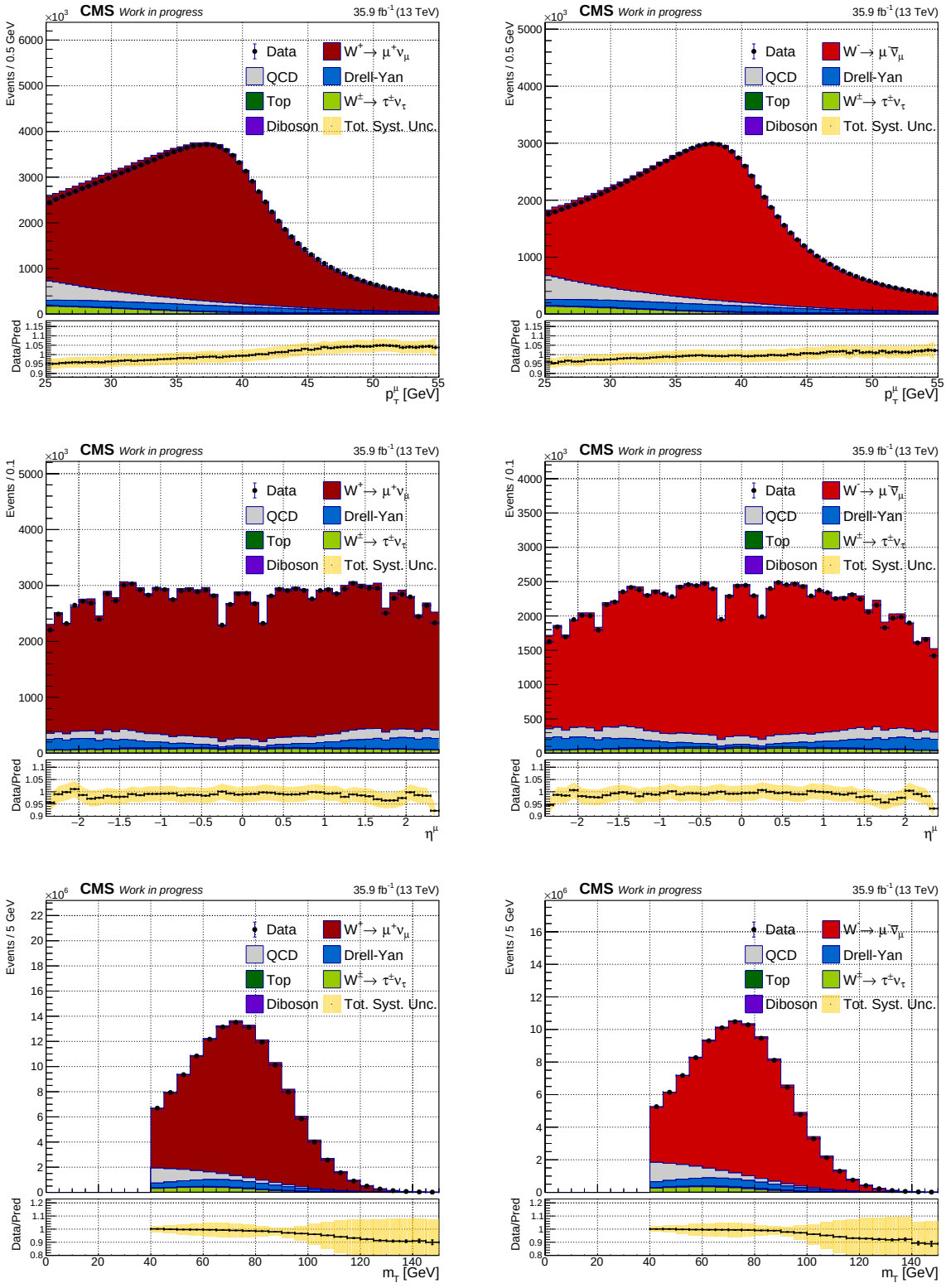


Figure 6.11. *Pre-fit distributions of transverse momentum p_T^μ (Upper plots), η^μ (central plots), and transverse mass m_T (lower plots), with Signal region selection for W^+ (left plots) and W^- (right plots), for data and stacked MC samples, together with the QCD yields estimation from data. In the panel below each distribution the ratio Data/MC prediction is shown, with the highlighted systematic uncertainty band.*

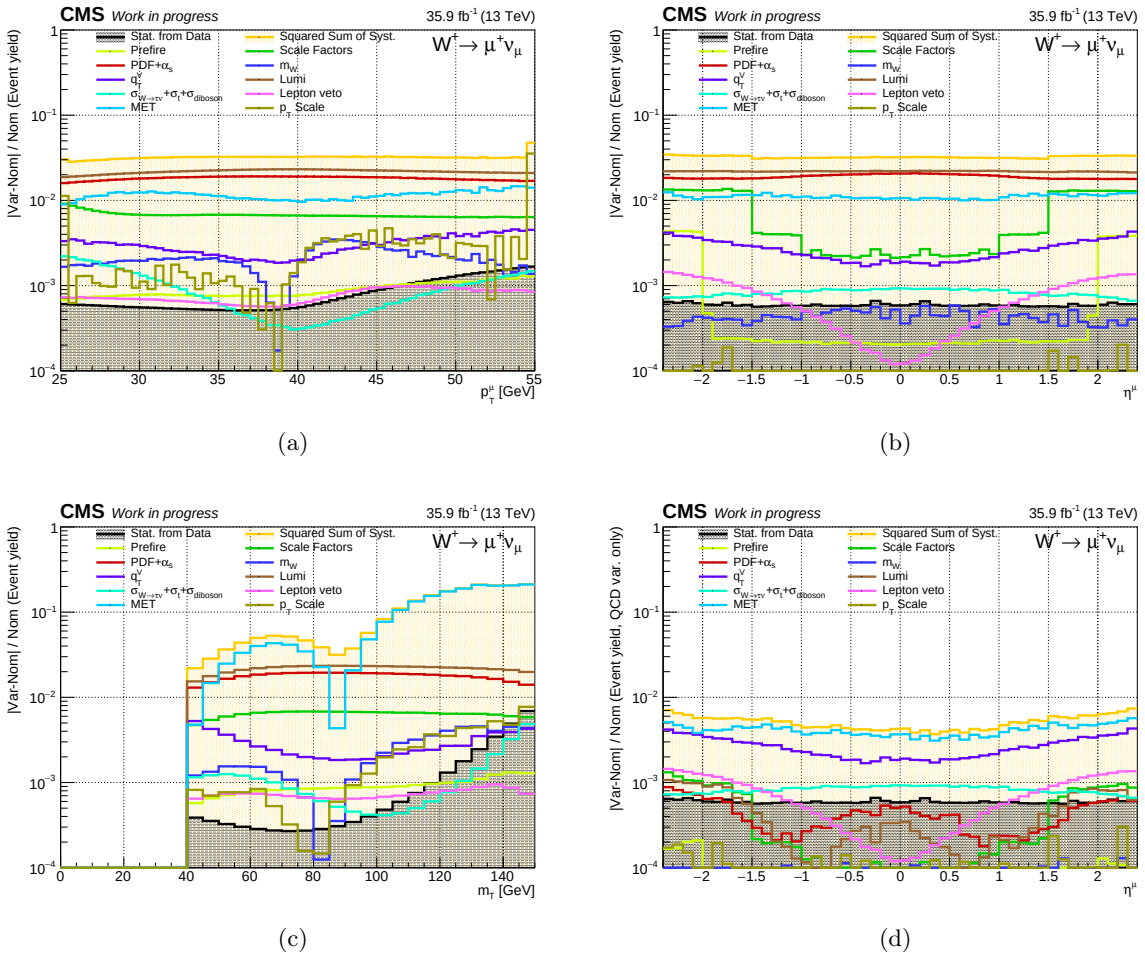


Figure 6.12. *Relative shift of the systematic uncertainties in the p_T^μ, η^μ and m_T distributions of the stacked MC sample, subdivided in classes, with W^+ Signal selection. The shift is evaluated as the sum in quadrature of the relative discrepancy /up-down//nominal (or /variation-nominal//nominal for PDF and MC Scales variations). In (d) only the variation induced by the QCD background yield is taken into account to evaluate the shift.*

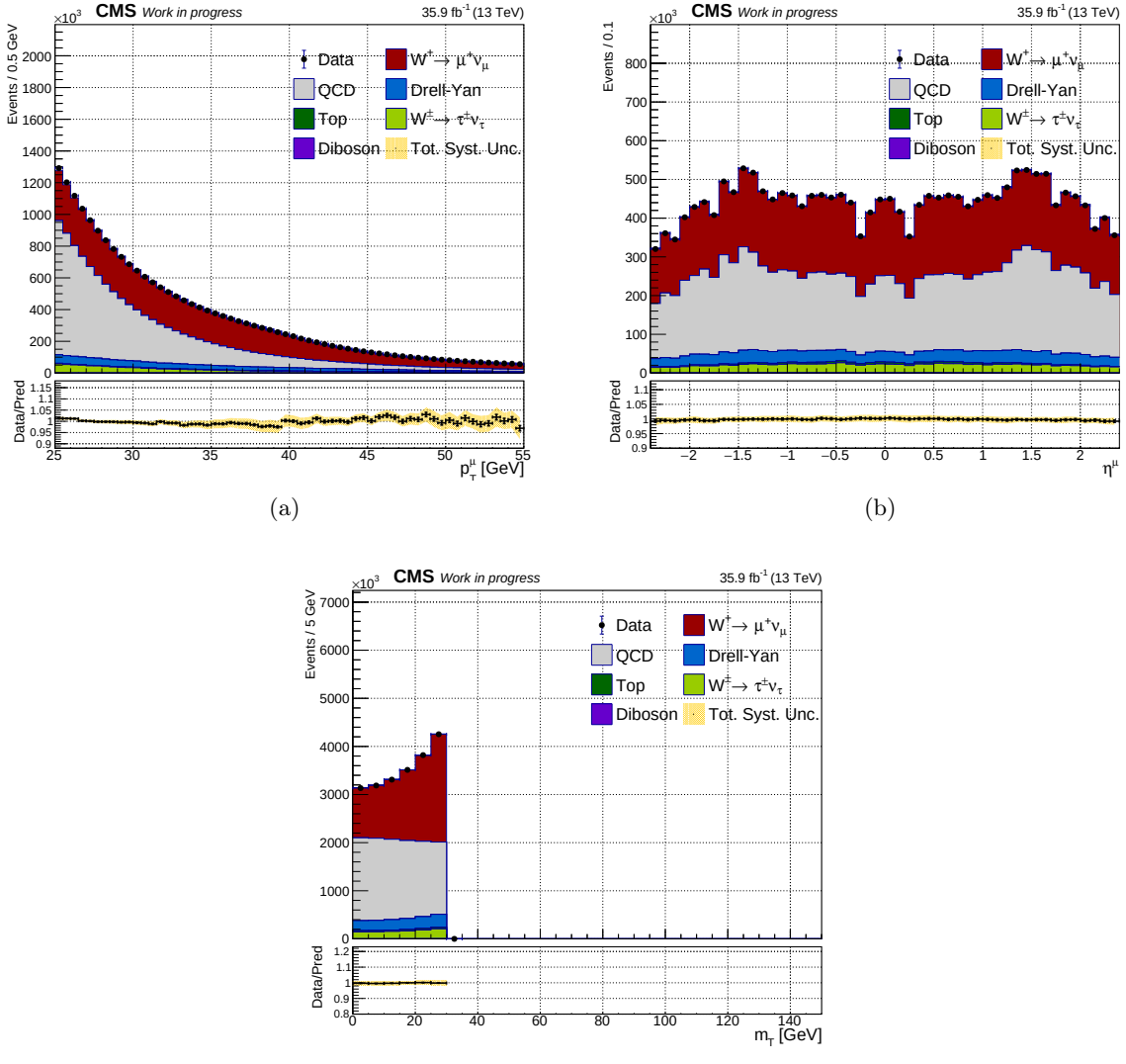


Figure 6.13. *Closure test of the QCD background estimation: as distributions of transverse momentum p_T^{μ} (a), η^{μ} (b), and transverse mass m_T (c), with Sideband region selection for W^+ sample, for data and stacked MC samples, together with the QCD yields estimation from data. In the panel below each distribution the ratio Data/MC prediction is shown, with the highlighted systematic uncertainty band.*

Chapter 7

Fit to the W boson production properties

In the first part of this chapter, the details of the template fit are provided. Then, the description of the signal and background templates together with the implementation of their systematic uncertainty variations is discussed. Finally, the results of the template fit are presented, in terms of single and double-differential unpolarized cross sections and angular coefficient distributions. In the last part of the chapter, an overview of the polynomial regularization study is introduced, with some preliminary results of this implementation on W production properties template fit. All the results are presented for simulated data only.

7.1 Fit description and setup

The aim of the fit is to unfold the W boson production cross section, differential in $|Y_W|$, q_T^W and for each A_i $(i = 0, \dots 4)$. The cross section has been extracted using an extended binned maximum likelihood fit. The fit has been performed independently for each boson charge.

The observed data[1] $\eta^\mu \times p_T^\mu$ distribution is fitted to the sum of the templates of the considered processes. The processes are:

- The signal processes i.e. one process per bin of $|Y_W|$, q_T^W and for each helicity cross section correspondents to angular coefficients σ_{UL}, $A_0, \dots A_4$, containing the events from $W^\pm \to \mu^\pm \nu$ decays
- The electroweak background processes.
- The QCD background process.
- An additional background process for the $W \to \mu\nu$ events produced outside the considered $|Y_W| \times q_T^W$ acceptance, but which fall inside the $\eta^\mu \times p_T^\mu$ acceptance of the templates (called *low-acceptance* process).

The templates of the aforementioned processes will be described in detail in Sec. 7.2.

[1] the fit has not been performed on real data in this thesis, but "data" is used instead of "pseudo-data generated from the MC" for convenience.

The negative log-likelihood function of the fit is:

$$L = -\ln\left(\mathcal{L}(\text{data}|\boldsymbol{\mu},\boldsymbol{\theta})\right) = \sum_{i}^{p_T^\mu,\eta^\mu \text{ bins}} \left(n_i^{\text{obs}} \ln n_i^{\text{exp}}(\boldsymbol{\mu},\boldsymbol{\theta}) + n_i^{\text{exp}}(\boldsymbol{\mu},\boldsymbol{\theta})\right) + \frac{1}{2}\sum_{k}^{\text{nuisances}} (\theta_k - \theta_k^0)^2,$$

$$n_i^{\text{exp}}(\boldsymbol{\mu},\boldsymbol{\theta}) = \sum_{p}^{\text{processes}} \mu_p n_{i,p}^{\text{exp}} \prod_{k}^{\text{nuisances}} \kappa_{i,p,k}^{\theta_k}. \tag{7.1}$$

The index i runs over the bins of the templates and n_i^{obs} is the number of observed events in each bin, assuming independent Poisson distributions. The index p runs over the processes, n_i^{exp} are the number of expected events per bin; $n_{i,p}^{\text{exp}}$ are the number of expected events per bin per process; μ_p is the signal strength modifier per signal process. The index k runs over the systematic uncertainties; θ_k are the associated nuisance parameters; $\kappa_{i,p,k}$ is the size of the systematic uncertainty per bin, per process, per nuisance.

The Parameters Of Interest (POIs) of the fit are the signal strength multipliers μ_p of the signal processes, formally defined by the second line of Eq. 7.1 and freely floating in the fit. The signal strength multipliers of the background processes are fixed to unity. The nuisance parameters θ_k describe the effect of the systematic uncertainties which can modify $n_i^{\text{exp}}(\boldsymbol{\mu},\boldsymbol{\theta})$. The systematic uncertainties have been implemented with log-normal variation of of the yield $n_{i,p}^{\text{exp}}$, with mean 0 and width equal to $\ln \kappa_{i,p,k}$. With this choice, the nuisances result in a multiplicative factor κ^{θ_k} on the event yields (with $\kappa^{\theta_k} = 1$ at pre-fit level). A unit Gaussian constraint to θ_k is added to $\mathcal{L}$ for each nuisance parameter. Additional details about the systematic uncertainty description can be found in Sec. 7.3. The parameter θ_k^0 is zero for the fit on data, but a non-vanishing value of θ_k^0 will be used in toy MC experiments, as described in Sec. 7.4.

The uncertainties and the covariance matrices for the POIs and the nuisance parameters have been obtained from the Hessian of the likelihood at minimum:

$$V_{i,j}^{-1} = -\frac{\partial^2 L}{\partial x_i \partial x_j}\Big|_{\mathbf{x}=\hat{\mathbf{x}}} \quad , \quad \mathbf{x} = \{\boldsymbol{\mu},\boldsymbol{\theta}\}, \tag{7.2}$$

where V^{-1} is the inverse of the covariance matrix of the POIs and the nuisance parameters and $\hat{\mathbf{x}}$ are the maximum likelihood estimator (MLE) of the parameters. This is calculated with high numerical precision thanks to the backpropagation exploited by `TensorFlow` [196]. The Eq. 7.2 relies on $-2L = \chi^2$ with Gaussian approximation. This approximation has been explicitly verified in the context of Ref. [11], where the `Combine-TensorFlow` fitting framework had been exploited for the first time.

The cross sections $\boldsymbol{\sigma}$ have been unfolded simultaneously to the POIs extractions. Dedicated processes have been defined for the cross section of each process σ_p. These processes, called *masked channel*, do not enter in the likelihood, but are needed to properly include the dependence of the predicted cross section on the nuisance parameters. The predicted value of the cross sections can be evaluated as $\hat{\mu}_p \sigma_p(\hat{\boldsymbol{\theta}})$, where $\hat{\boldsymbol{\theta}}$ is the predicted value of the nuisances and includes only the relevant nuisance parameters (the theoretical systematic uncertainties), $\hat{\mu}$ is the signal strength multiplier predicted from the fit. Note that $\sigma_p \neq \hat{\mu}_p \sigma^{\text{exp}}$, since the nuisance parameters can modify the value of the cross section. The fit allows also to extract also the post-fit covariance matrix for $\boldsymbol{\sigma}$, given the covariance matrix of $\boldsymbol{\mu}$ and $\boldsymbol{\theta}$ and the Jacobian of the transformation, internally constructed exploiting the masked channels dependence on $\boldsymbol{\theta}$.

The fit predicts signal strengths, which are directly related to the number of events per bin per process and they are related to the cross sections using the masked channels. Thus, the decomposition in terms of helicity cross sections of Eq. 5.3, and not the decomposition in term

of angular coefficients, is implicitly used in this implementation. The description in terms of the helicity cross sections is more robust since all the σ_i have the same dimensions, can be linearly summed, and can be easily related to the number of events.

The desired description of Eq. 5.1, in term of A_i, can be recovered. With the same approach used for the $\boldsymbol{\sigma}$, it is possible to exact other derived quantities, simultaneously to the POIs, with the proper covariance matrix. In particular, the angular coefficients are derived from the predicted cross sections for each process. The single-differential distributions $d\sigma/d|Y_W|$ and $d\sigma/dq_T^W$ for each helicity cross section (and angular coefficient) have been extracted as well, integrating the double-differential predicted cross sections in the range used in the measurement. The effect of the nuisance parameters is properly propagated to the corresponding covariance matrices.

The described fit setup and parameterization follow the prescriptions of the LHC Higgs Combination Working Group [197]. The fit has been implemented within the framework of `Combine`, the `CMSSW` package developed for the CMS Higgs measurements and widely used by the CMS analysis community [198]. The core of the fit, in term of minimization and errors propagation, has been rewritten with the use of `TensorFlow` software [196]. In the described configuration the fit has to manage a very large number of events (see table 6.1), a total of 2880 bins and 402 processes, with 396 POIs and 242 nuisance parameters which must be simultaneously fitted, for each boson charge. The standard `Combine` tools cannot provide the required precision and numerical stability because of this complexity. On the other hand `TensorFlow` exploits an efficient calculation of the gradients with the backpropagation approach, which allow for a semi-analytical minimization procedure. This feature guarantees the precision and the robustness needed for the fit and is particularly critical in the non-convex region of the likelihood, where the usual minimizers fail. Moreover, `TensorFlow` implementation allows to parallelize the fit procedure and optimize the memory consumption. A sparse tensor representation has been used for this purpose. The fit can natively run on CPU or on GPU.

7.2 Signal and Background templates

The $\eta^\mu \times p_T^\mu$ signal templates have been built with the procedure described in Sec. 5.2.1, from the simulated $W^\pm \to \mu^\pm \nu$ events. A total of 396 templates for each W boson charge has been built, binning the events in $|Y_W| \times q_T^W \times A_i$ with the following bins:

- 6 equal-width bins of $|Y_W|$, between $|Y_W| = 0$ and $|Y_W| = 2.4$, with $\Delta Y_W = 0.4$,
- 11 bins of q_T^W, between $q_T^W = 0$ GeV and $q_T^W = 60$ GeV, with bin edges corresponding to $[0, 2, 6, 8, 10, 12, 16, 20, 26, 36, 60]$ (in GeV),
- 5 angular coefficients and the unpolarized cross section: A_0, A_1, A_2, A_3, A_4, σ_{U+L}.

The templates are constructed with the helicity cross sections σ_i, despite being called A_i. Referring to a template of the given process p (i.e. given bin of $|Y_W|$, q_T^W) as T_p, from the Eq. 5.3:

$$T_p(\eta^\mu, p_T^\mu) = \frac{3}{16\pi} \mathcal{L} \sigma_p P_p \left(\cos\theta^*(\eta^\mu, p_T^\mu), \phi^*(\eta^\mu, p_T^\mu)\right), \tag{7.3}$$

where σ_p are the helicity cross sections, $\mathcal{L}$ is the data luminosity, P_p are the harmonic angular functions from Eq. 5.2. The factor $\frac{3}{16\pi}$ is needed to obtain $n_{tot} = \sum_i \sum_p T_p$, where i runs on the template bins, p on the signal processes of the selected $(|Y_W|, q_T^W)$ bin and n_{tot} is the total number of expected signal events in the $(|Y_W|, q_T^W)$ bin. At fixed $|Y_W|$ and q_T^W, the event content

of a (η^μ, p_T^μ) bin is function of $(\cos\theta^*, \phi^*)$ only. With this definition the expected number of events in some bins of some T_p's can be negative.

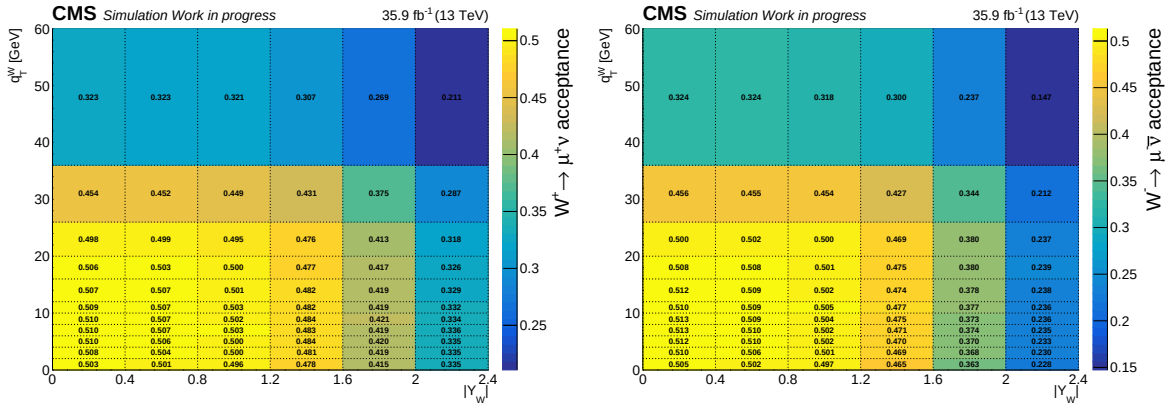
For each template, η^μ ranges between -2.4 and 2.4, with 48 equal-width bins of $\Delta\eta^\mu = 0.1$, while p_T^μ ranges between 25 GeV and 55 GeV, with 60 equal-width bins of $\Delta p_T^\mu = 0.5$ GeV. The limits are defined by the selection requirements (trigger and CMS kinematic acceptance, background suppression, etc. described earlier in Sec. 5.3). The η^μ and p_T^μ binning width are defined to be as fine as technically possible, to catch the features of the muon-differential distributions, and the width is limited by the muon momentum resolution.

The edges of the region ($|Y_W| = 2.4$, $q_T^W = 60$ GeV) have been defined in such a way that the $\eta^\mu \times p_T^\mu$ acceptance is higher than 20% for each of the considered bins. Referring to Fig. 5.3, it can be seen that beyond $|Y_W| = 2.4$ about half of the distribution accepted at $Y_W = 0$ falls outside of the acceptance. Moreover, outside the chosen acceptance, the signal over background ratio decreases rapidly, degrading the prediction power of the measurement. The signal acceptance estimated from MC simulation is shown in Fig. 7.1. The acceptance is evaluated as the ratio between the reconstructed signal yield, summing over all the signal templates, and the generator level yield, for each $|Y_W|$, q_T^W bin. Thus it is encoding both the geometrical acceptance and the reconstruction efficiency effects. The acceptance information is not explicitly used to correct the fit prediction, since the masked channels allow to directly relate the fit μ_p prediction with the σ_p at generator level in the unfolding procedure.

The bin widths in $|Y_W|$, q_T^W have been chosen to have about equally-populated bins. The q_T^W binning has been reduced to $\Delta q_T^W = 2$ GeV in the low q_T^W region, to catch the fine details of the spectrum. The limitation, excluding the statistical uncertainty, is given by the W boson finite width ($\Gamma_W \simeq 2$ GeV) since Γ_W smears the templates distribution similar to a variation of q_T^W of about $\Gamma_W/2$.

Some examples of the signal templates are reported in Fig. 7.2. The change of the shapes, as a function of the $|Y_W|$, q_T^W and A_i, is the key feature that allows the extraction of the W boson production properties from the template fit.

The electroweak background templates, for the W^+ channel, are reported in Fig. 7.3. One template for each process has been generated, directly using the simulated distribution from



(a) (b)

Figure 7.1. *Signal acceptance in the $|Y_W| \times q_T^W$ ranges considered in the fit, estimated using the W^+ (left) and W^- (right) MC samples, as the sum of all the signal templates over the generator level $W^\pm$ yield.*

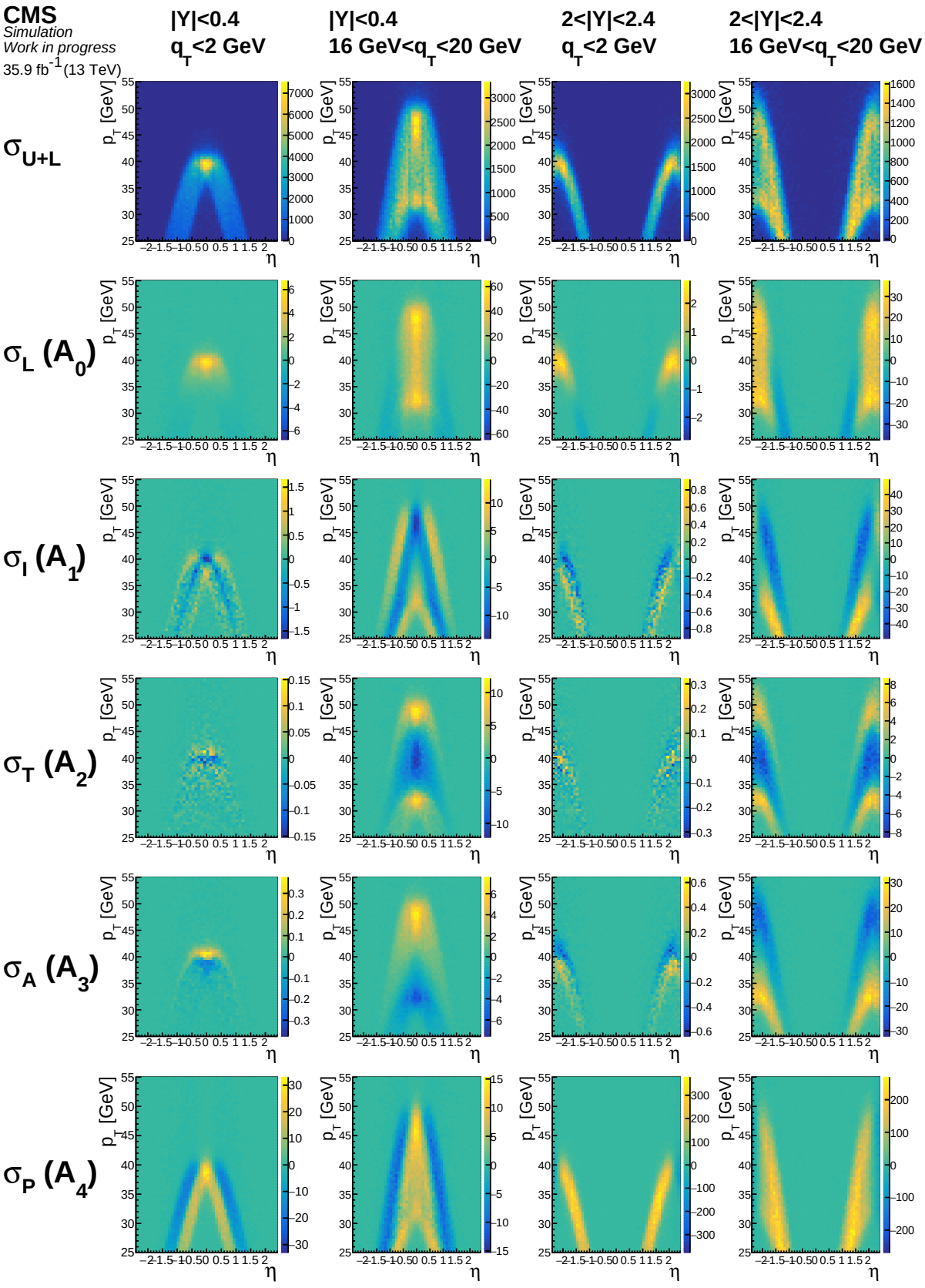


Figure 7.2. *Representative examples of the signal templates for the five angular coefficients and σ^{U+L} (each row) in different bins of $|Y_W|$ and q_T^W (each column). The $W^+ \to \mu^+\nu$ sample has been used. The value of the templates is coherent with Eq. 7.3. The z axis scale is different in each plot and represents the equivalent number of events.*

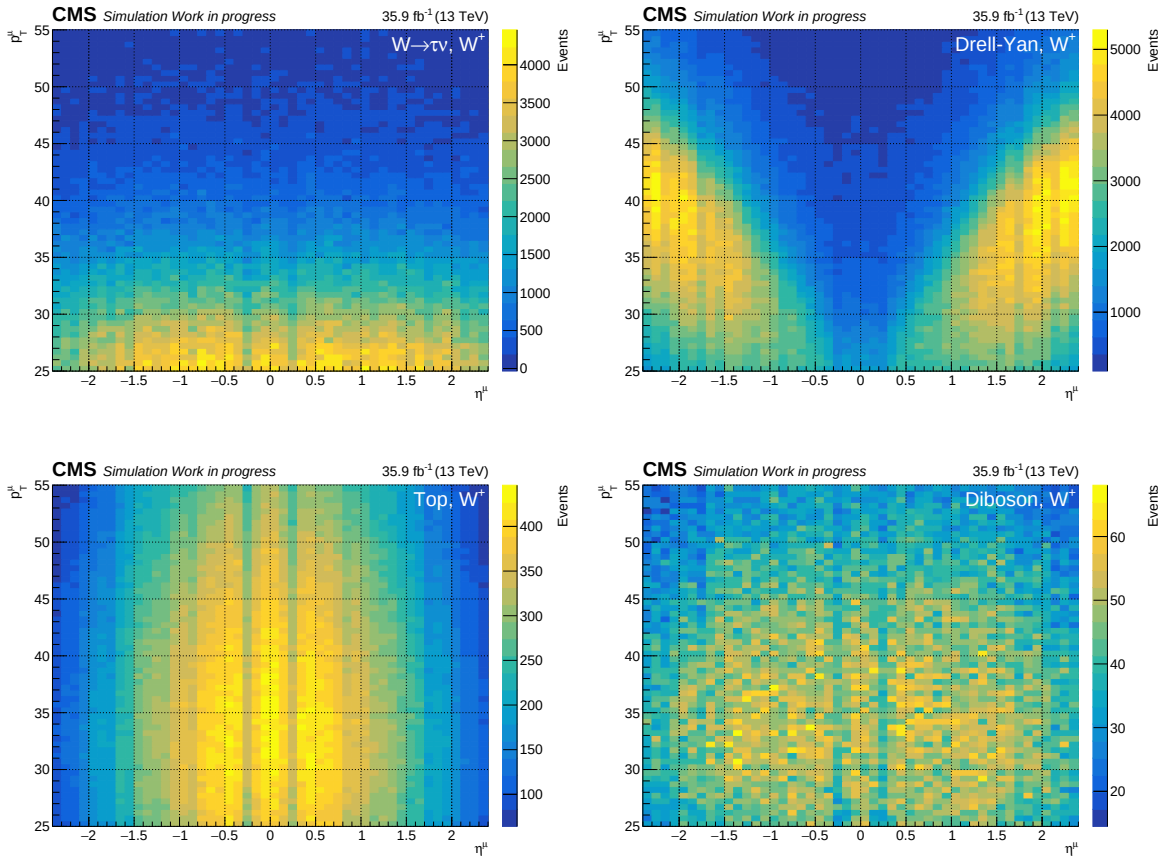
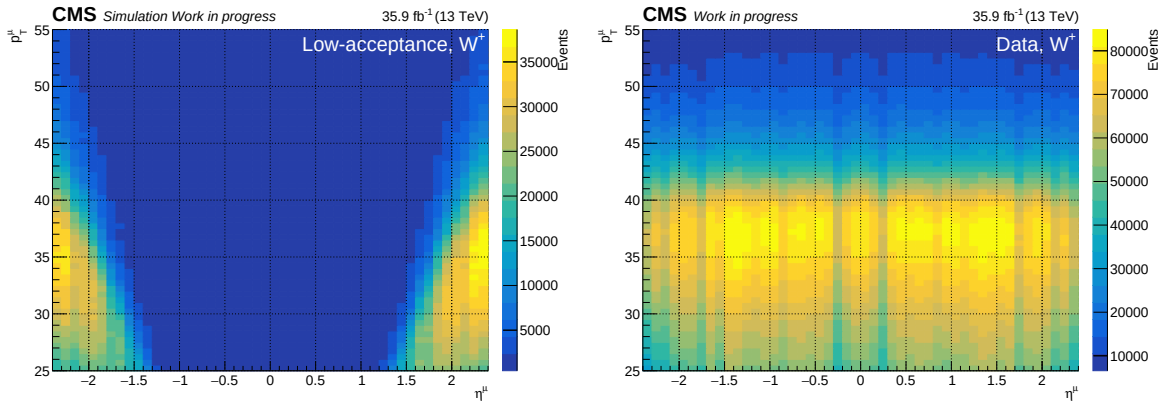


Figure 7.3. *Electroweak background templates, for W^+ channel. The value of the templates is coherent with Eq. 7.3. The vertical scale is different in each plot. The total number of events in each template is (in million of events): $N_{W\to\tau\nu} = 3.32\ M$, $N_{Z/\gamma^*} = 5.75\ M$, $N_t = 0.77\ M$, $N_{diboson} = 0.12\ M$*

.



(a) (b)

Figure 7.4. *Low-acceptance (a) and Data (b) templates, for W^+ channel. The value of the templates is coherent with Eq. 7.3. The vertical scale is different in each plot. The total number of events in the templates is (in million of events): $N_{low\text{-}acc} = 15.48\ M$, $N_{data} = 132.68\ M$.*

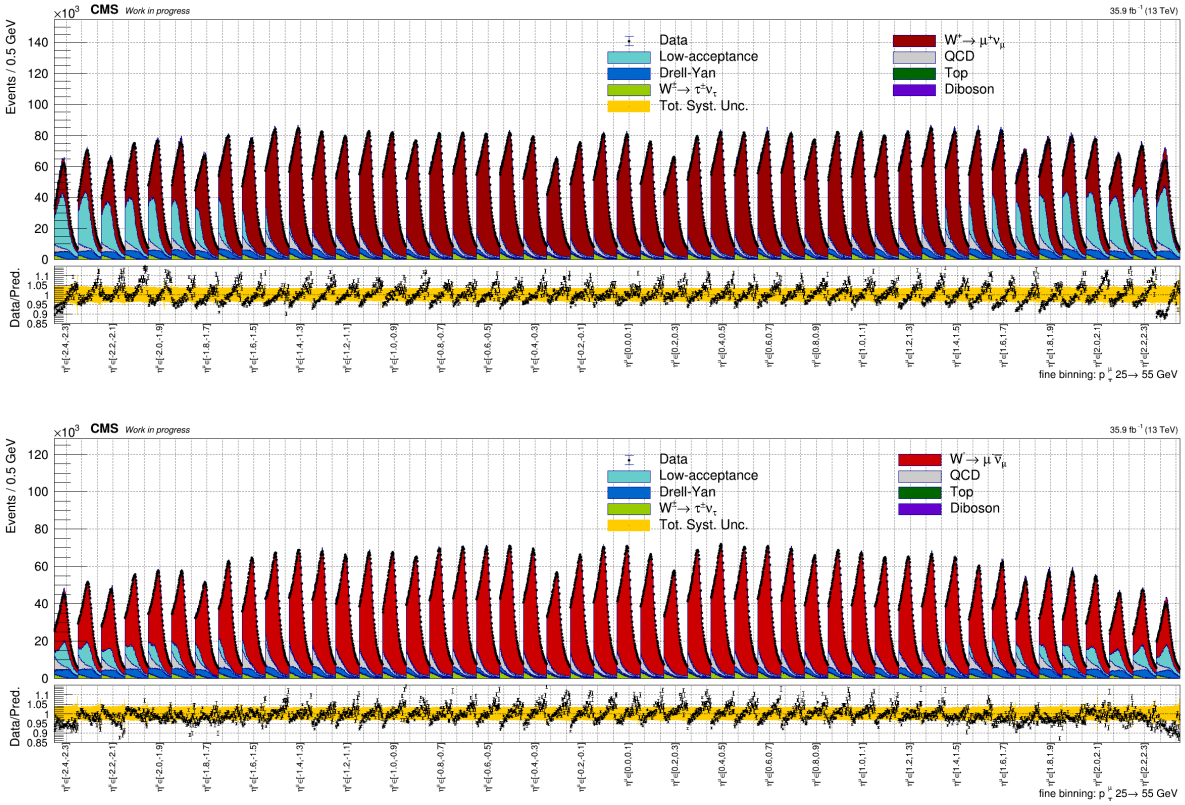


Figure 7.5. *Stacked distribution of the unrolled templates for W^+ (upper plot) and W^- (lower plot) channels, provided in input to the fit, together with real data distributions (despite not being used in the fit). The $W^\pm \to \mu^\pm\nu$ process include the sum of all the signal processes. The statistical and entire set of systematic uncertainties are included in the band of the lower panels.*

MC. The same $\eta^\mu \times p_T^\mu$ binning of the signal templates has been used. The processes considered are: $W \to \tau\nu$, neutral Drell-Yan (Z/γ^*), top decays, diboson decays. These processes have been described in Chapter 6. The QCD-background templates have been already reported in Sec. 6.4, with the same binning.

The low-acceptance template is generated using the $W \to \mu\nu$ MC, with the signal selection, but requiring $|Y_W| > 2.4$ or $q_T^W > 60$ GeV. The latter is reported in Fig. 7.4(a) for W^+ channel.

The data template for W^+ channel is shown in Fig 7.4(b). In this case the real data are shown, despite not being used in the fit. The data template allows a direct comparison between the signal and backgrounds event yields on $p_T^\mu \times \eta^\mu$ plane.

The stacked distribution of the processes provided to the fit is shown in Fig. 7.5 together with the real data distribution (despite not being used in the fit). The distribution is unrolled in $\eta^\mu(p_T^\mu)$ bins in the considered range. The distribution of Fig. 6.11 is obtained integrating this distribution in η^μ or p_T^μ. The low-acceptance yield is different in the two boson charge, because of the spin correlation in W decays. The W^+ tends to produce the muon backwards compared to its flight direction, increasing the acceptance of muons produced by a high-Y W^+, while W^- produces the opposite effect. In the bottom panel, the band of the uncertainties is shown. These include all the variations discussed in Sec. 5.4 and the statistical uncertainty. Also the ratio between data and the sum of the templates is shown. Technically, the signal and background templates are converted in this unrolled format to be processed by the fit.

7.3 Systematic uncertainties

The systematic uncertainties which affect the templates are described in Sec. 5.4. They have been implemented in the fit as nuisance parameters, tracking their correlation across each bin of each template. The systematic uncertainties are subdivided in two categories: the systematic uncertainties which affect the normalization of the templates only, and the ones which modify the shape of the templates. The summary of the considered uncertainties is reported in table 7.1.

Table 7.1. *Systematic uncertainties considered in the fit for each source and each process. A log-normal nuisance parameter is applied where the explicit value is shown, and it represents the value of the κ parameter. A shape nuisance parameter is applied where "shape" is reported. The total number of nuisance parameters per source is also reported.*

Nuisance	Signal	Z/γ^*	$W \to \tau\nu$	Top	Diboson	QCD	Low-acc.	$N_{\text{nui.}}$
Lepton veto	-	2%	-	-	-	-	-	1
Data Luminosity	2.5%	2.5%	2.5%	2.5%	2.5%	2.5%	2.5%	1
$\sigma_{W\to\tau\nu}$	-	-	4%	-	-	-	-	1
σ_t	-	-	-	6%	-	-	-	1
σ_{diboson}	-	-	-	-	16%	-	-	1
QCD normalization	-	-	-	-	-	5%	-	1
JES, E_U	shape	shape	shape	shape	shape	shape	shape	2
p_T^μ scale	shape	shape	shape	shape	shape	shape	shape	1
SF_{stat}	shape	shape	shape	shape	shape	shape	shape	144
SF_{syst}	shape	shape	shape	shape	shape	shape	shape	1
L1 trigger prefire	shape	shape	shape	shape	shape	shape	shape	1
Luminosity on fake rate	-	-	-	-	-	shape	-	1
PDF	shape	shape	shape	-	-	shape	shape	60
α_s	shape	shape	shape	-	-	shape	shape	1
m_W	shape	-	-	-	-	shape	shape	1
q_T^Z (MC Scale)	-	shape	-	-	-	-	-	6
q_T^W (MC Scale binned in q_T^W)	-	-	shape	-	-	shape	shape	18

7.3.1 Normalization uncertainties

These uncertainties are fully correlated among all p_T^μ and η^μ bins of each templates and they can change the overall normalization of each process only. They are implemented with log-normal priors. The log-normal distribution for a parameter θ is defined as:

$$p(\theta) = \frac{1}{\sqrt{2\pi}\ln\kappa}\frac{1}{\theta}\exp\left[-\frac{(\ln\theta - \ln\theta_0)^2}{2(\ln\kappa)^2}\right], \tag{7.4}$$

which means that the variable $\ln\theta$ has a normal distribution with mean $\ln\theta_0$ and standard deviation $\ln\kappa$. With this formulation, relative uncertainty on the parameter θ is represented by $\kappa - 1 = \Delta\theta/\theta$. An observable N, estimated by $\hat{N}$, with a given log-normal uncertainty κ is parameterized as $N = \hat{N}\kappa^\theta$, with θ distributed as a Gaussian with null mean and unit standard deviation [197]. Referring to Eq. 7.1, the nominal yield of the template is described when $\theta_k = 0$. For $\theta = \pm 1\sigma$ the nominal yield is multiplied or divided by κ, respectively.

The following normalization uncertainties have been considered in the fit:

– **Lepton veto.** A $\kappa - 1 = 2\%$ uncertainty has been considered for the Drell-Yan background process only (from Ref. [11]).

– **Data luminosity.** A $\kappa - 1 = 2.5\%$ uncertainty has been considered for all the signal and background processes [186].

– **Electroweak cross sections.** A $\kappa - 1 = 4\%, 6\%, 16\%$ uncertainties have been considered for the $W \to \tau\nu$, top and diboson processes, respectively (from Ref. [11]).

– **QCD background.** A $\kappa - 1 = 5\%$ uncertainty has been considered for the QCD background channel, to cope with the discrepancy described in Sec. 6.4.

7.3.2 Shape uncertainties

These nuisance parameters associated with the shape uncertainties parametrize a modification of the shape of the $\eta^\mu \times p_T^\mu$ template of the related channels. Two additional templates have been provided beside the nominal one for each uncertainty (i.e. for each nuisance parameter θ), corresponding to the up and down variations of the given uncertainty. They represent the $\pm 1\sigma$ variation of θ. The central and the varied templates are vertically interpolated using a spline function $f(\theta)$, where $f(-1)$, $f(0)$, $f(1)$ correspond to down, nominal and up template, respectively [198]. The nuisance parameter θ is encoded with a Gaussian constraint in the likelihood of the fit, likewise the log-normal uncertainties.

The following shape uncertainties have been considered in the fit:

– **MET-Related.** The JES and unclustered energy uncertainties have been considered for all the processes. The two uncertainties are both fully correlated in the $\eta^\mu \times p_T^\mu$ plane, for a total of 2 nuisance parameters.

– **Muon momentum calibration.** The Rochester correction uncertainty has been considered for all the processes. The simplified-fully correlated shapes have been used, thus a single nuisance parameter has been considered.

– **Efficiency scale factors.** An uncertainty fully correlated in η^μ and p_T^μ plus 3 uncertainties fully correlated in p_T^μ and uncorrelated η^μ have been considered for all processes. In conclusion, 147 nuisance parameters are provided.

– **L1 Trigger prefire correction.** A single uncertainty, fully correlated in the $\eta^\mu \times p_T^\mu$ plane has been considered, for all the processes.

– **Luminosity on fake rate.** The 2.5% luminosity uncertainty produces a non-flat variation of the fake rate and thus of the QCD background templates. Therefore an additional shape uncertainty has been considered for the QCD background process only, using the luminosity-varied version of the fake and prompt rate.

– **PDF and α_s.** The 60 Hessians-eigenvector plus one α_s shape variations have been considered for Signal processes and Drell-Yan, $W \to \tau\nu$, QCD and low-acceptance backgrounds. The uncertainties are fully correlated in the $\eta^\mu \times p_T^\mu$ plane each, for a total of 61 nuisance parameters.

– **W boson mass.** The shapes correspondent to 50 MeV m_W variation has been considered for Signal, low-acceptance and QCD background processes. The uncertainty is fully correlated in the $\eta^\mu \times p_T^\mu$ plane thus a single nuisance parameter has been used.

– **W and Z boson transverse momentum.** The shapes corresponding to the 6 MC scale variations have been considered for the Drell-Yan process. The shapes corresponding to the q_T-decorrelated MC scale variation have been considered for $W \to \tau\nu$, low-acceptance and QCD backgrounds processes, as described in Sec. 5.4.2.4. The total number of nuisance parameters is 24.

– **Monte Carlo finite statistic.** The likelihood description properly takes into account the data statistical uncertainties. However, the use of finite-size MC samples requires to consider also the statistical fluctuation for each process, independently in each bin. In principle, a nuisance parameter should be assigned for each process for each bin. However, it is possible to use the Barlow-Beeston-lite approach [199], relying on the assumption of reasonable-high statistic for each process for each bin. In this case, Gaussian uncertainties can be considered for each process, and thus assign a single nuisance parameter to the resulting Gaussian uncertainty on the total yield of each bin. A unit Gaussian constraint is also added to the likelihood, as usual. This shape uncertainty is referred as bin-by-bin (BBB) uncertainty. In this analysis, a limited size MC sample has been used, which has about 1/9th of the statistical power of the data sample of 2016 data taking period (35.9 fb^{-1}). This work is a proof of feasibility of the future analysis, and only pseudo-data built from the same parent distributions used to build the templates have been used. Therefore the BBB will be disabled (i.e. the related nuisances are removed from the likelihood) in the entire analysis. In principle, this could lead to bias in the uncertainty estimations, because the uncertainties are estimated from the Hessian of the fit, whose precision can be affected by the MC finite size and without the BBB this effect is not properly corrected. However, a test with and without BBB has been performed and no relevant discrepancies in the uncertainties have been observed.

7.4 Angular coefficients and unpolarized cross section estimate

The fit has been performed on the so-called *Asimov dataset* with a number of expected events corresponding to an integrated luminosity of 35.9 fb^{-1}, equivalent to the 2016 data taking period sample. This pseudo-data sample is defined by the sum of the signal and all the backgrounds. It is obtained adding all the templates considered in the fit, computed with all the nuisance parameters constrained to the nominal value.

The fit to Asimov pseudo-data is performed by fixing the signal strengths modifiers $\mu_p = 1$ in the fit for all the processes considered. The nuisance parameters are constrained to the nominal value in the likelihood setting the $\theta_0^k = 0$ parameter for all the k (from Eq. 7.1). A fit to the Asimov pseudo-data returns the expected values of the POIs, by definition, and is important to assess the consistency of the fitting model. The covariance matrix and the constraints on the nuisance parameters returned by the fit are a reliable expectation of the results which can be obtained fitting an equivalent data sample. The fit to the Asimov pseudo-data can be also performed with μ_p left freely floating, testing the consistency of the minimization. In this case, an unbiased fit must return $\mu_p = 1$ for all the considered processes, where the minimum of the likelihood lies.

The results of the fit to the Asimov dataset have been validated with *toy MC experiments*. In this case, the Asimov pseudo-data distribution is randomized according to Poissonian fluctuations. The systematic uncertainty effect is taken into account randomizing the θ_0 parameters in the likelihood according to the nuisance distributions. The μ_p are left freely floating in the likelihood. Possible bias in the fit prediction and the uncertainties can be estimated generating a large sample

of toy MC experiments. For each toy the fit is repeated. The mean of the predictions on each POI must be compatible with the expected value i.e. the Asimov dataset fit prediction. The variance of the toys predictions must be compatible with the expected uncertainty from the fit to the Asimov dataset.

With this purpose the distribution of pulls can be built for each POI using the prediction from the toys. For a given POI X, the pull is defined as:

$$\text{pull} = \frac{X_{\text{pred}} - X_{\text{exp}}}{\sigma_{X_{\text{pred}}}},$$

where X_{pred}, $\sigma_{X_{\text{pred}}}$ indicate the predicted POI of the single toy and its uncertainty, respectively, and X_{exp} the POI value from the fit to the Asimov dataset. The pull distribution is expected to be a Gaussian, centered at 0 with standard deviation equal to 1. Discrepancy from this expectation indicates a possible bias in the fit. The Asimov pseudo-data represents the asymptotic behaviour of the toys distributions for $N_{\text{toys}} \to \infty$, and the toys allow only to test the consistency of the fit procedure, while the uncertainties on the POIs, their correlation and the impact of the nuisance parameters are fully encoded in the Asimov dataset fit result. Additional details about the Asimov dataset can be found in Ref. [200].

The results are presented in terms of the unfolded cross sections and angular coefficients. Therefore, they can be directly compared with the generator-level distribution, to theoretical predictions or other experiments results.

7.4.1 Double-differential measurement

The result of the fit on the Asimov dataset are presented in Figs. 7.6 to 7.11 for σ_{U+L}, $A_0, \dots A_4$. The fitted double-differential distributions in $|Y_W|$ and q_T^W of σ_{U+L} and of the angular coefficients are shown. Since the POIs of the fit are the signal strength of the helicity cross sections, the angular coefficients are obtained from the latter using Eq. 5.4, with the proper propagation of the fit covariance matrix. With the same technique, the single differential distribution has been obtained. For each helicity cross section, the fitted values have been integrated in q_T^W or in $|Y_W|$ and the angular coefficient is computed using Eq. 5.4. The results of this procedure are also shown.

The fitted results are compared with the expected values from MC simulation. The expected uncertainties from the PDF and α_S and from the MC scales μ_R and μ_F variations are also shown. The PDF uncertainty arises from the quadrature sum of the 60 Hessian eigenvalues variations, coherently with the PDF systematic uncertainty description. The α_S and MC scales uncertainties have been estimated like the corresponding systematic uncertainties, as described in Sec. 5.4. In particular μ_R and μ_F have been varied independently in the numerator and denominator of the angular coefficients. For all the observables but A_4, the limiting factor is represented by the MC scales. The expected precision on A_4 is dominated by the PDF uncertainty, since A_4 is completely determined by the PDF, being the only LO angular coefficient.

The central values of the fit results are identically equal to the expected values because the fit has been performed on the Asimov dataset. The uncertainty on the double-differential distributions is dominated by the statistical uncertainty and it is larger than the simulation uncertainty. The impact of the systematic uncertainties will be discussed in Sec. 7.4.3. The σ^{U+L} distribution is the only observable properly resolved, despite predicted uncertainty ranges between 5% and 10%, with peaks at 20% at high $|Y_W|$ (to be compared with the 6% of MC scale uncertainty and 2% of PDF uncertainty).

The size of the uncertainty includes the large effect of the correlation among the different bins of $|Y_W|$, q_T^W. In Fig. 7.12 the sectors of the correlation matrix corresponding to σ^{U+L} and each

angular coefficients are shown (note that the widths of the bins are proportional to the widths of the correspondent q_T^W bin, for example the large correlation-1-square on the diagonal is the 36 GeV $< q_T^W <$ 60 GeV bin). In all the angular coefficients an anti-correlation pattern is visible between neighbour q_T^W bins in a single bin of $|Y_W|$. This anti-correlation can be present only between the nearest neighbours (eg. for A_3 correlation matrix) or spread between multiple neighbours (eg. A_0). In the former case the anti-correlation is very strong (60%-90%), while in the latter it is moderate (20%-40%). Milder correlation patterns are also visible between the same q_T^W bin of different $|Y_W|$ bins. These correlation patterns increase the uncertainty on the distributions of the angular coefficients. The equivalent correlation matrices for W^- sample are reported in Appendix E.

The effect of the correlation can be quantified comparing the fit uncertainty with the Poissonian uncertainty of each bin, estimated as $1/\sqrt{n}$, where n is the number of events in the bin. This result is shown in Fig. 7.13, showing how the correlations increase the uncertainty by a factor of about 10^2. The statistical component of the fit uncertainty is highlighted, to exclude the effect of systematic uncertainties in the correlation effect.

The source of these correlation patterns must be traced back to the templates. The templates of nearby $|Y_W|$ and q_T^W bins compete for the same $\eta^\mu \times p_T^\mu$ regions, since the templates change smoothly as a function of the boson variables, with overlaps between them. This can be partially seen directly from Fig. 7.3, where the overlap is visible also for quite far templates in $|Y_W|$ and q_T^W. This is expected from the features described phenomenologically in Sec. 5.2.2 and Fig. 5.3, where large overlaps are also visible in the polarization-integrated distributions. This overlap is also present between different angular coefficients at the same value of $|Y_W|$, q_T^W (A_0 and σ_{U+L} for instance, referring to Fig. 7.2).

If the overlapping templates are only the nearest neighbours the anti-correlations is very high but limited between adjacent bins. If there are multiple overlapping templates the anti-correlation is spread between several bins. These features of the template limit the capability of the fit to distinguish between the overlapping bins. This "degeneracy" may produce a fluctuation in neighbour bins in $|Y_W|$, q_T^W, or between different angular coefficients in the same $|Y_W|$ and q_T^W bin, resulting in an anomalous high-frequency oscillation of the predicted distribution. The Asimov fit, forcing μ to be one avoids this behaviour, but it can be observed using toy MC experiments, as will be discussed in Sec. 7.4.4. This feature is typical of unfolding problems. It has been already predicted in the proof-of-concept paper of Ref. [8] and observed in W rapidity measurement [11]. It represents one of the main technical limiting factors for the precision of this analysis. Mitigation strategies will be discussed in Sec. 7.5.

7.4.2 Single-differential measurement

The single-differential distributions have smaller statistical uncertainty compared to the double-differential case. They are shown in Figs. 7.6 to 7.11, as $A_i(|Y_W|)$ and $A_i(q_T^W)$. However they are still predicted with larger uncertainties compared to the simulated ones, and they are compatible with zero in most of the phase space. $A_4(|Y_W|)$ spectrum is the only angular coefficient significantly different from zero (Fig. 7.11, central row), with uncertainties below 50% in most of the range, but still less precise than the theoretical expected value.

The predicted $d\sigma^{U+L}/d|Y_W|$ has a precision of 2%, which is comparable with the expected PDF uncertainty, with a small increase in the high $|Y_W|$ region. This prediction is an important consistency check of the analysis, since it reproduces the measurement of the W rapidity spectrum from Ref. [11], shown in Fig. 4.12, with comparable precision. The two results are not *identical* since the selections applied is different. In particular, the events with $q_T^W > 60$ GeV are excluded

from the result presented here, differently from Ref. [11].

The predicted $d\sigma^{U+L}/dq_T^W$ has a precision compatible or better than the expected values, between 8% below 2 GeV, 6% below 10 GeV and 4% at higher q_T^W. This is particularly relevant, since it demonstrates the capability of the analysis to constrain the q_T^W spectrum, also in the very low q_T^W region. It should be noted that MC scales band underestimates the full theory uncertainty below $q_T^W \sim 10$ GeV and the N^{3}LL+NNLO prediction shown in Fig. 4.5 must be considered, as discussed in Sec. 4.3.2.1.

The covariance matrices for the differential results in $|Y_W|$ and for the differential measurement in q_T^W are shown in Fig. 7.14. These correlation matrices allow to study the interplay between different angular coefficients in a simplified environment compared to the double differential case. The q_T^W differential, integrated in $|Y_W|$ (right figure), shows clear correlations between the various angular coefficients. The most relevant ones are the anti-correlation between A_0 and A_4 and σ^{U+L}, which are instead positively correlated between them, and the anti-correlation between A_4 and A_1. These (anti)correlations can be understood in term of competing templates of the involved processes.

The neighbours q_T^W bins anti-correlation pattern is reduced but still present, in particular for A_3. The sensitivity of A_3 (i.e. σ_A) to this effect can be understood referring to Fig. 7.2 (A_3, $|Y_W| < 0.4$, 16 GeV $< q_T^W <$ 20 GeV). $\sigma_A(p_T^\mu)$ has a single peak and a single valley, differently from other σ_i (consider, for instance, σ_T). The increase of q_T^W produces a increase (decrease) of σ_A at higher (lower) p_T^μ. This trend produces an almost complete overlap between adjacent q_T^W bins, while the double-peak shape of the other σ_i reduces the overlap.

The $|Y_W|$ differential correlation matrix, integrated in q_T^W, does not show any neighbours bins anti-correlation pattern. This means that, with the q_T^W integration, the templates result better separated in $|Y_W|$. However, the anti-correlation between different angular coefficient is still present. Moreover, a positive correlation within all the $|Y_W|$ bins of σ^{U+L} is observed. The source of this correlation is the luminosity nuisance parameter, which is fully correlated among the bins and represent the dominant source of uncertainty in $d\sigma^{U+L}/d|Y_W|$. The impact of luminosity will be discussed in more detail in Sec. 7.4.3.

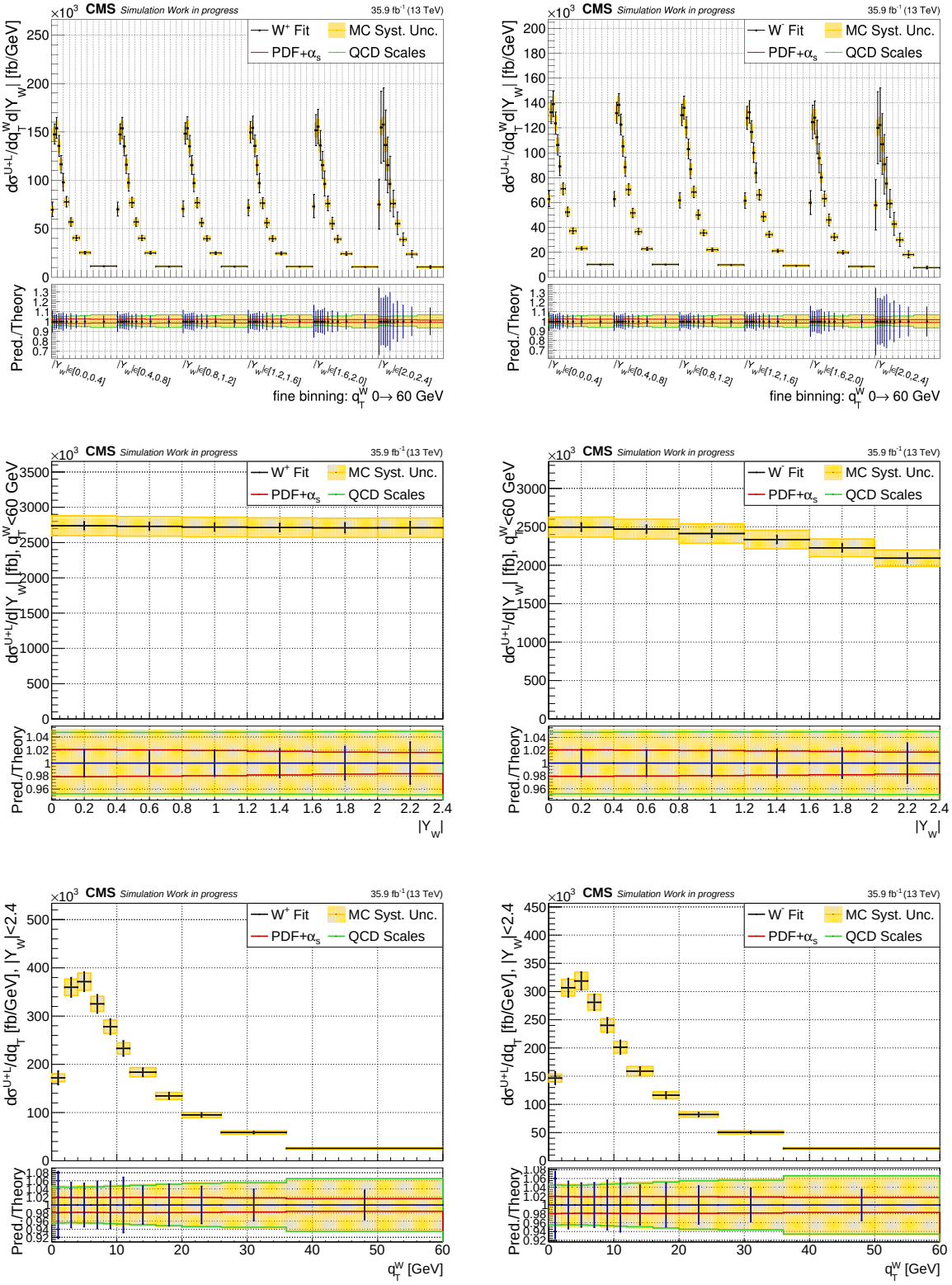


Figure 7.6. *Fitted σ^{U+L}, double differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central), differential in q_T^W (lower) for W^+ (left) and W^- (right) samples. The fit has been performed on the Asimov dataset. The yellow shaded area corresponds to the expected theory uncertainty, summing in quadrature the PDF and the MC scale contributions. In the panel below each distribution the ratio between the fitted value and the expected one is shown, together with the band of the theory uncertainties, highlighting the two sources. The BBB uncertainty has not been considered.*

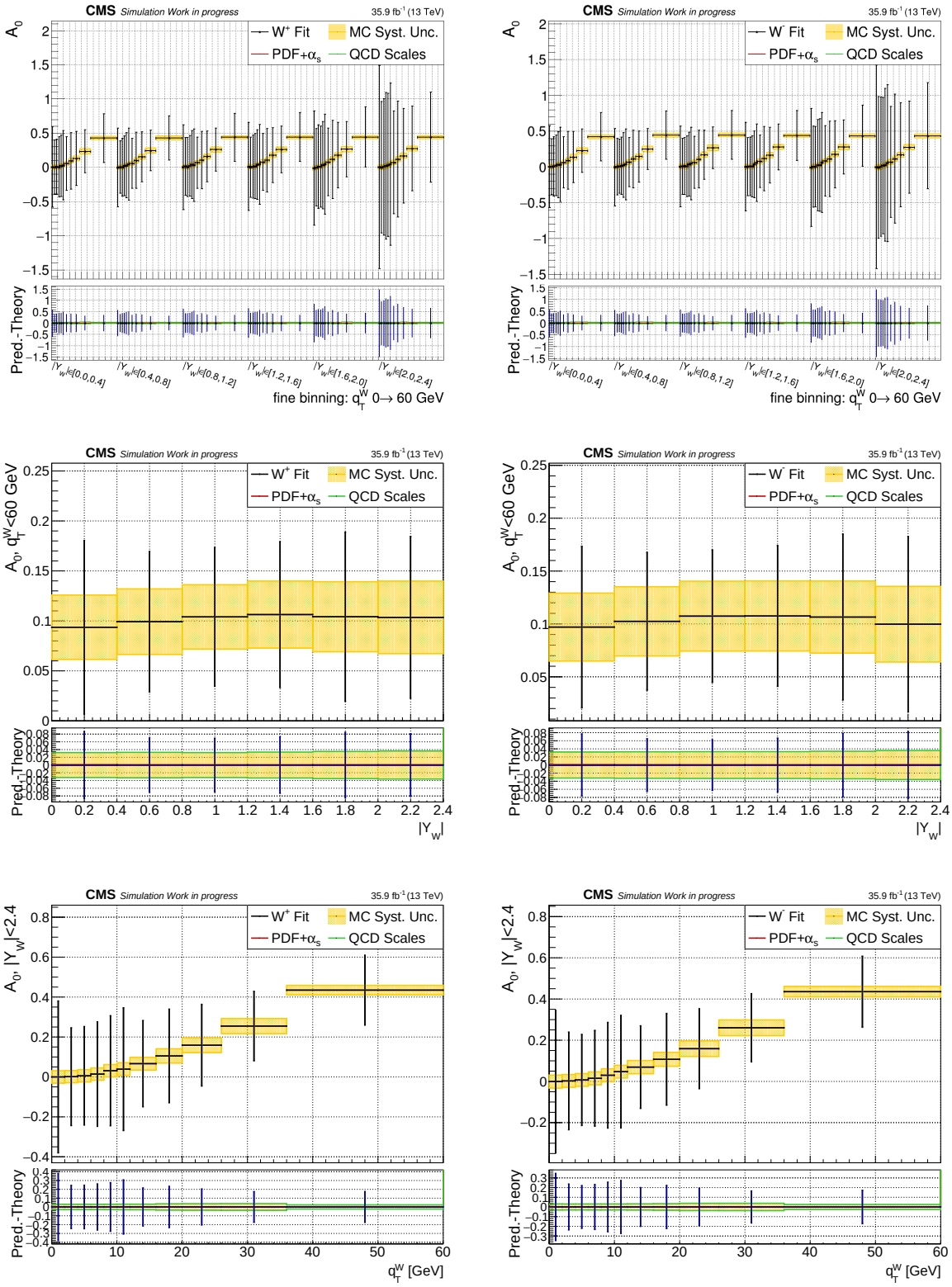


Figure 7.7. *Fitted A_0, double differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central), differential in q_T^W (lower) for W^+ (left) and W^- (right) samples. The fit has been performed on the Asimov dataset. The yellow shaded area corresponds to the expected theory uncertainty, summing in quadrature the PDF and the MC scale contributions. In the panel below each distribution the difference between the fitted value and the expected one is shown, together with the band of the theory uncertainties, highlighting the two sources. The fit uncertainties in this panel are removed in the double differential result for clarity. The BBB uncertainty has not been considered.*

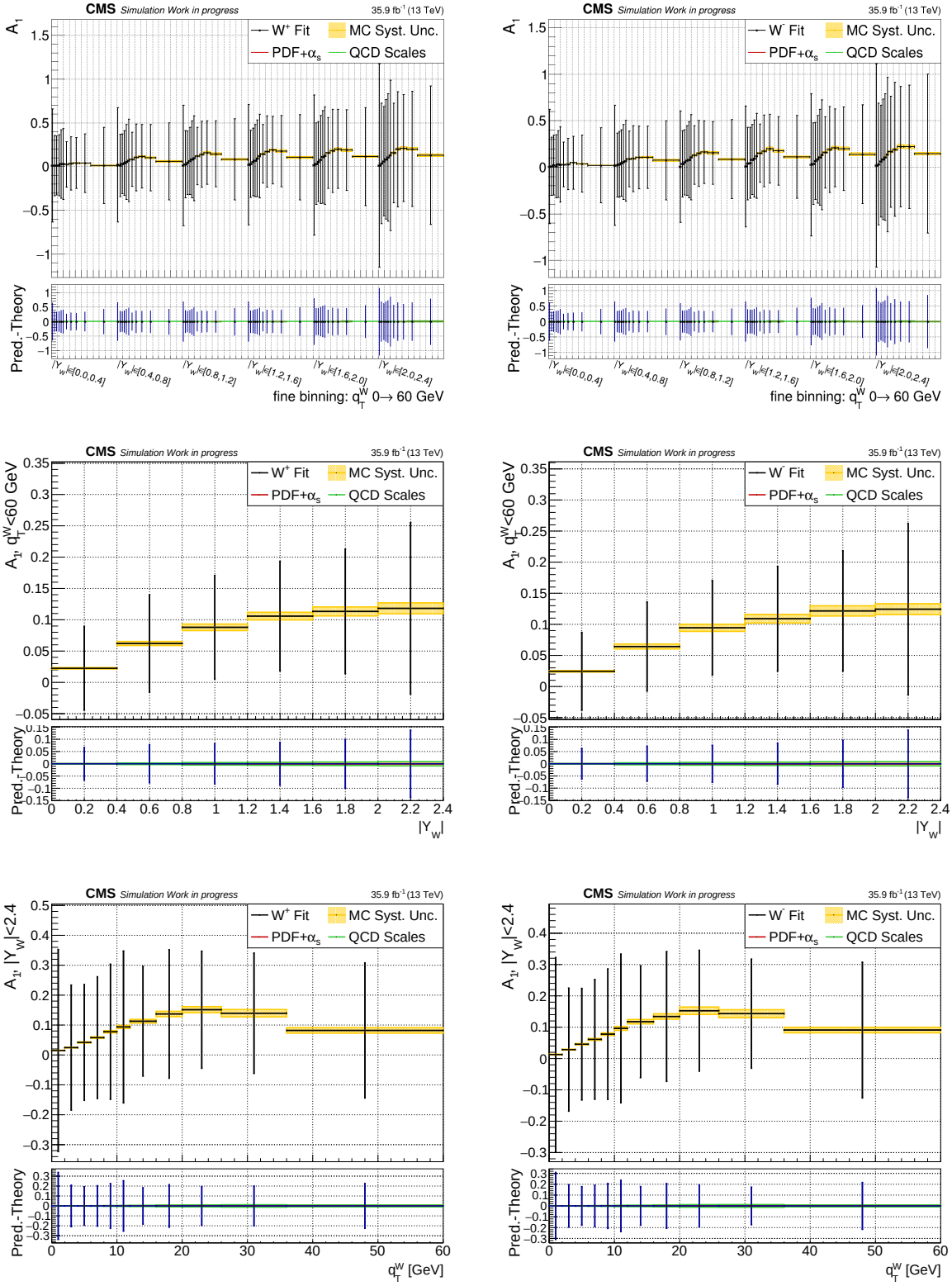


Figure 7.8. *Fitted A_1, double differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central), differential in q_T^W (lower) for W^+ (left) and W^- (right) samples. The fit has been performed on the Asimov dataset. The yellow shaded area corresponds to the expected theory uncertainty, summing in quadrature the PDF and the MC scale contributions. In the panel below each distribution the difference between the fitted value and the expected one is shown, together with the band of the theory uncertainties, highlighting the two sources. The fit uncertainties in this panel are removed in the double differential result for clarity. The BBB uncertainty has not been considered.*

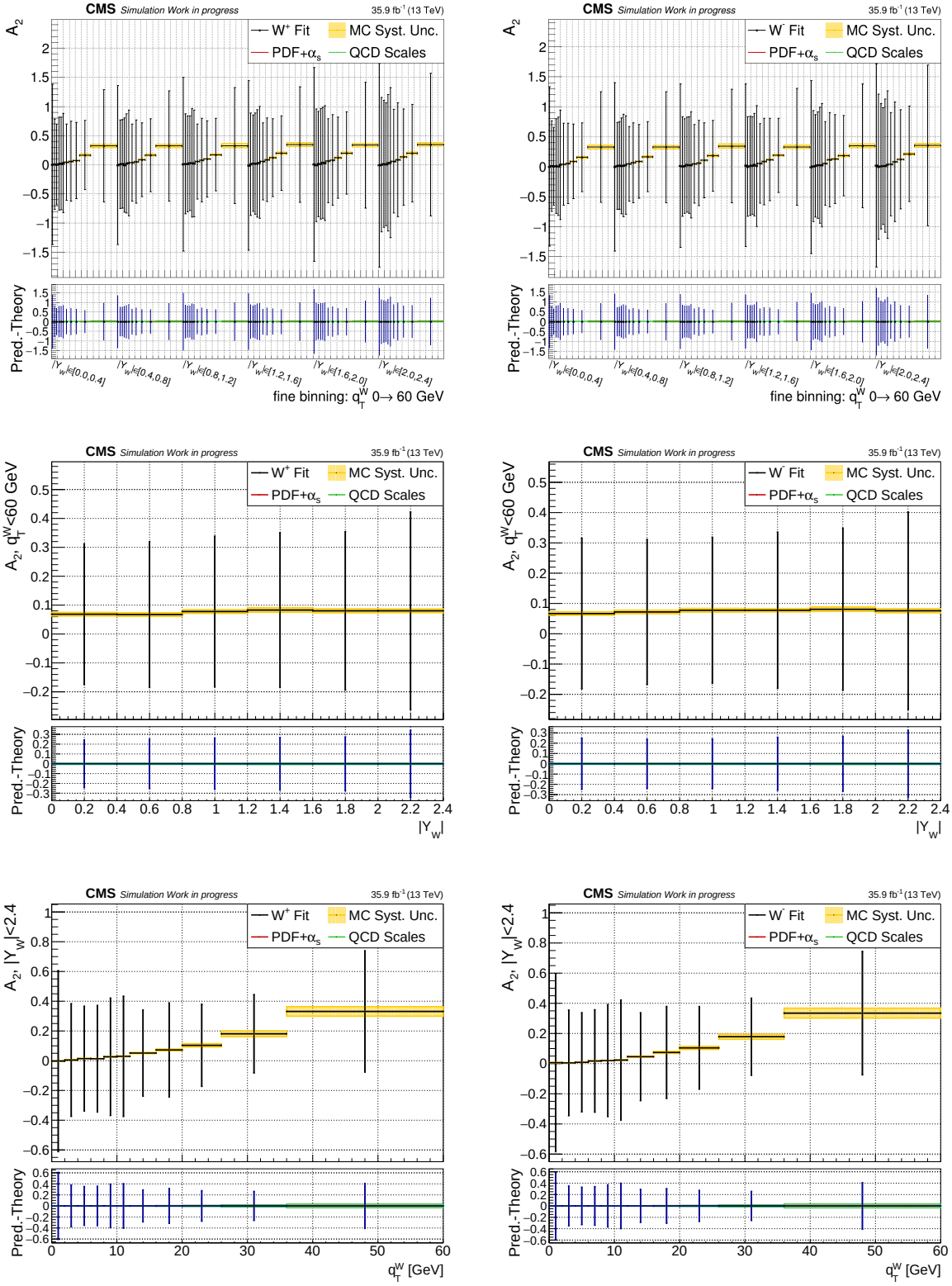


Figure 7.9. *Fitted A_2, double differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central), differential in q_T^W (lower) for W^+ (left) and W^- (right) samples. The fit has been performed on the Asimov dataset. The yellow shaded area corresponds to the expected theory uncertainty, summing in quadrature the PDF and the MC scale contributions. In the panel below each distribution the difference between the fitted value and the expected one is shown, together with the band of the theory uncertainties, highlighting the two sources. The fit uncertainties in this panel are removed in the double differential result for clarity. The BBB uncertainty has not been considered.*

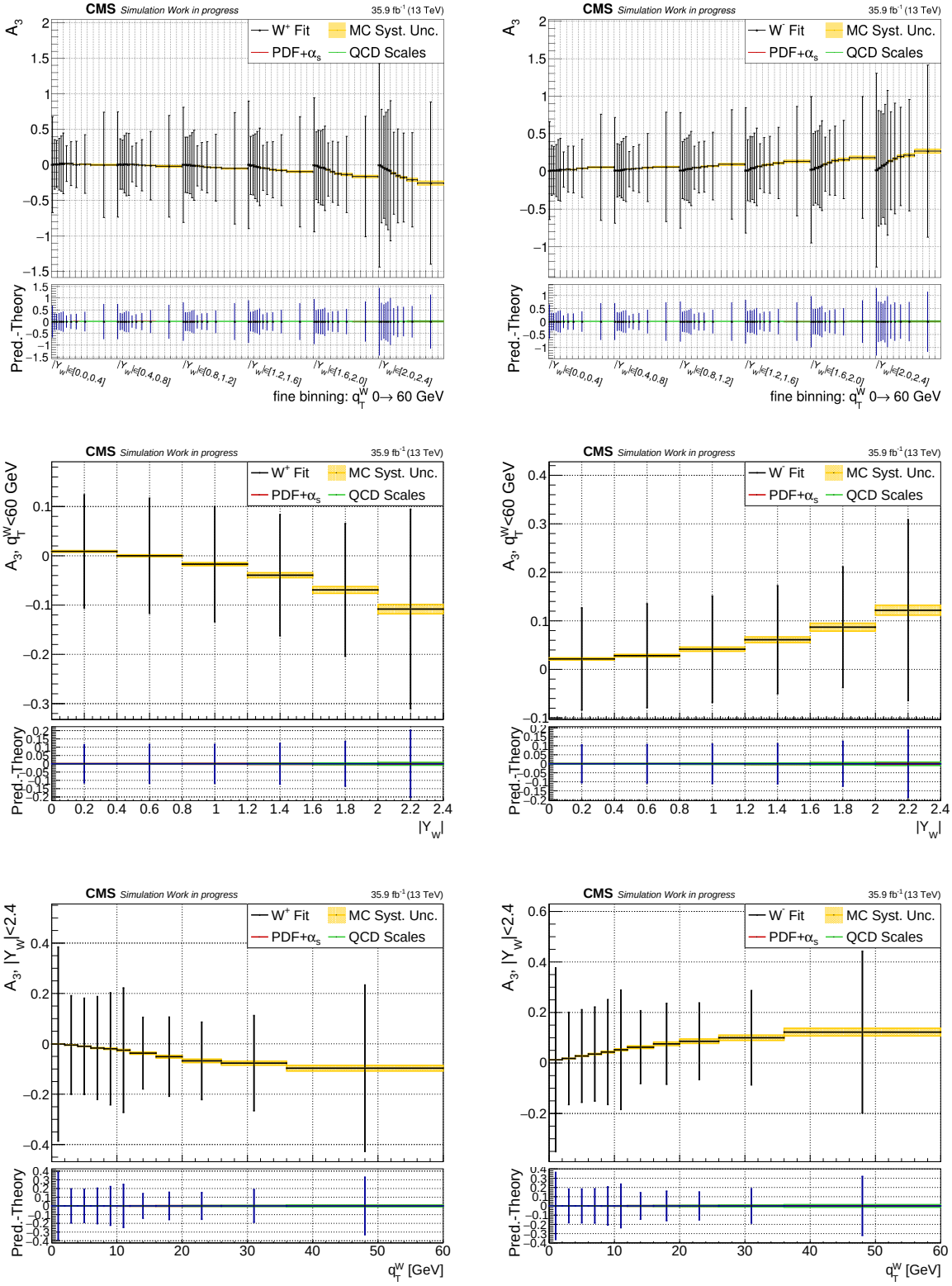


Figure 7.10. *Fitted A_3, double differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central), differential in q_T^W (lower) for W^+ (left) and W^- (right) samples. The fit has been performed on the Asimov dataset. The yellow shaded area corresponds to the expected theory uncertainty, summing in quadrature the PDF and the MC scale contributions. In the panel below each distribution the difference between the fitted value and the expected one is shown, together with the band of the theory uncertainties, highlighting the two sources. The fit uncertainties in this panel are removed in the double differential result for clarity. The BBB uncertainty has not been considered.*

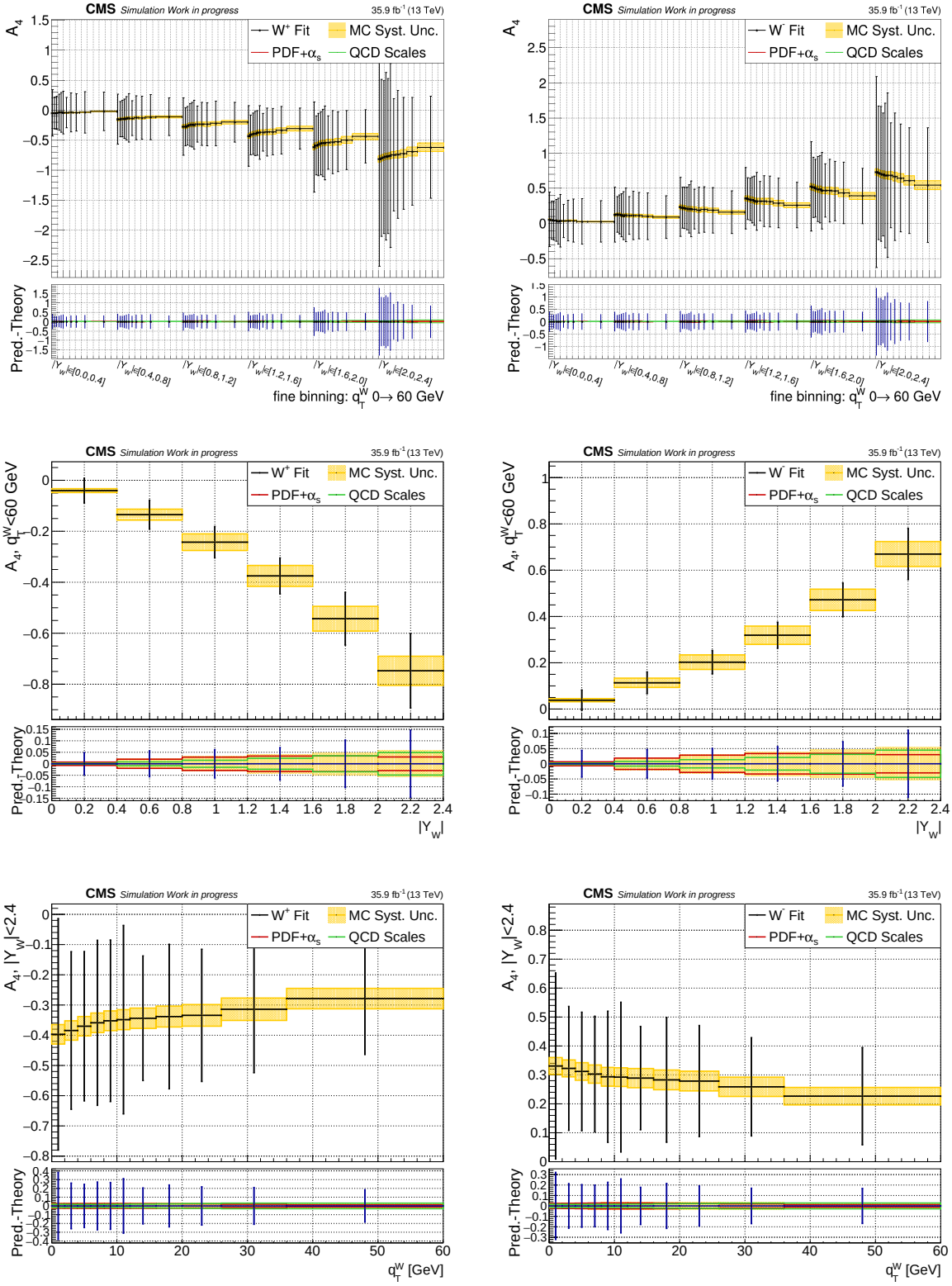


Figure 7.11. *Fitted A_4, double differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central), differential in q_T^W (lower) for W^+ (left) and W^- (right) samples. The fit has been performed on the Asimov dataset. The yellow shaded area corresponds to the expected theory uncertainty, summing in quadrature the PDF and the MC scale contributions. In the panel below each distribution is shown the difference between the fitted value and the expected one, together with the band of the theory uncertainties, highlighting the two sources. The fit uncertainties in this panel are removed in the double differential result for clarity. The BBB uncertainty has not been considered.*

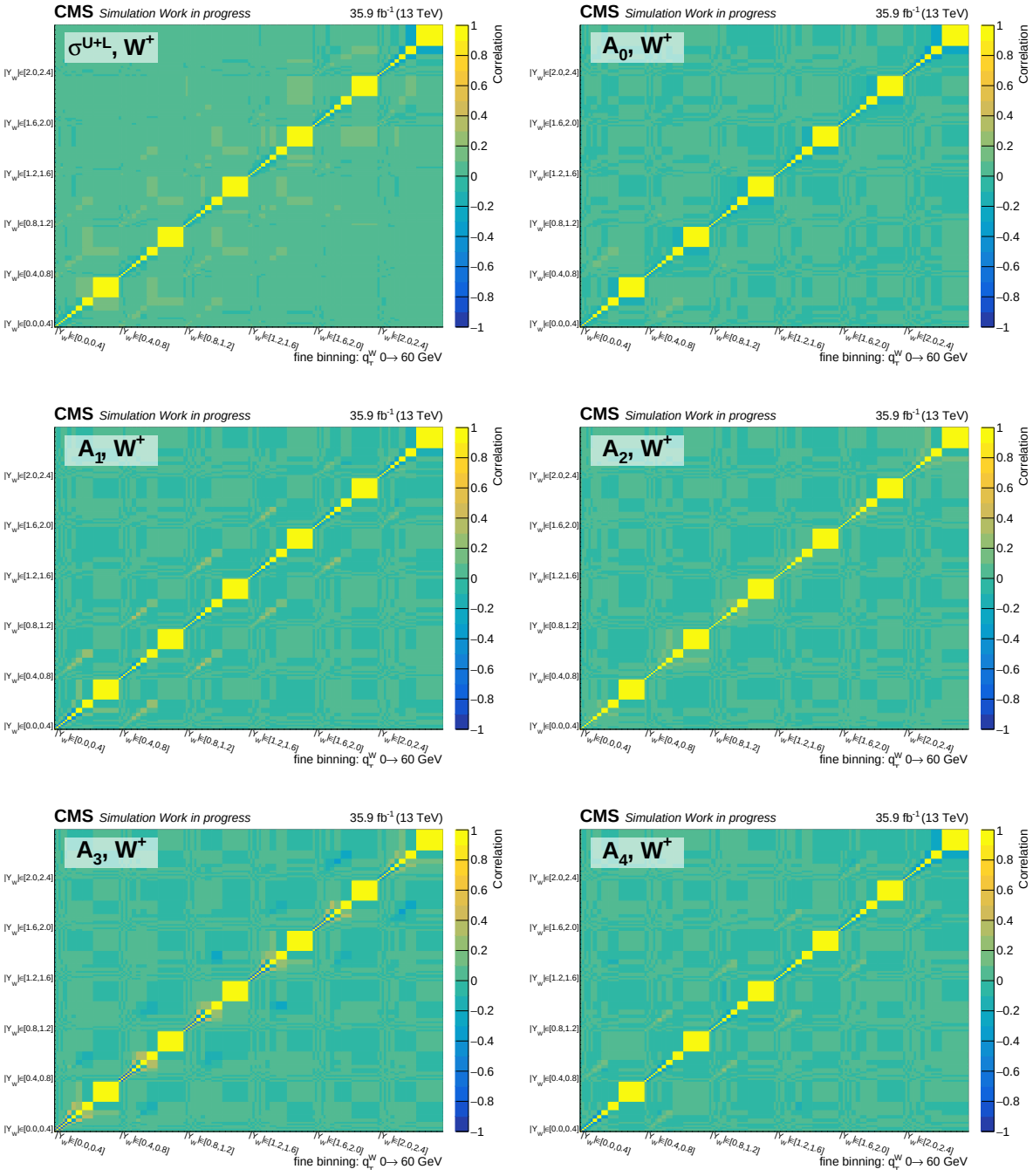
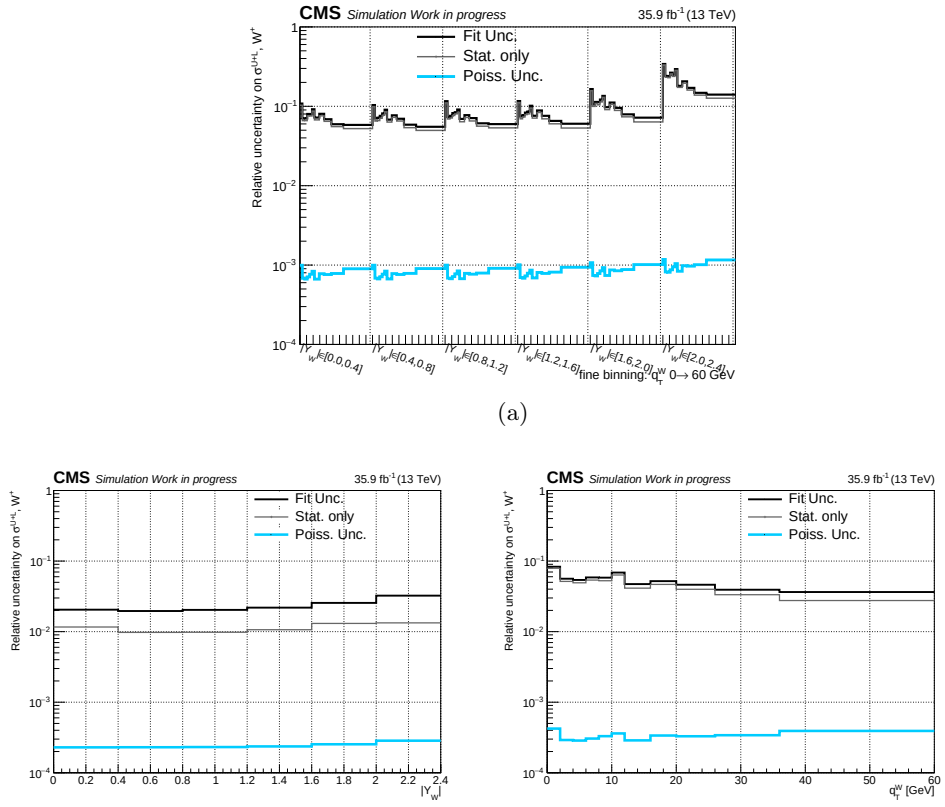


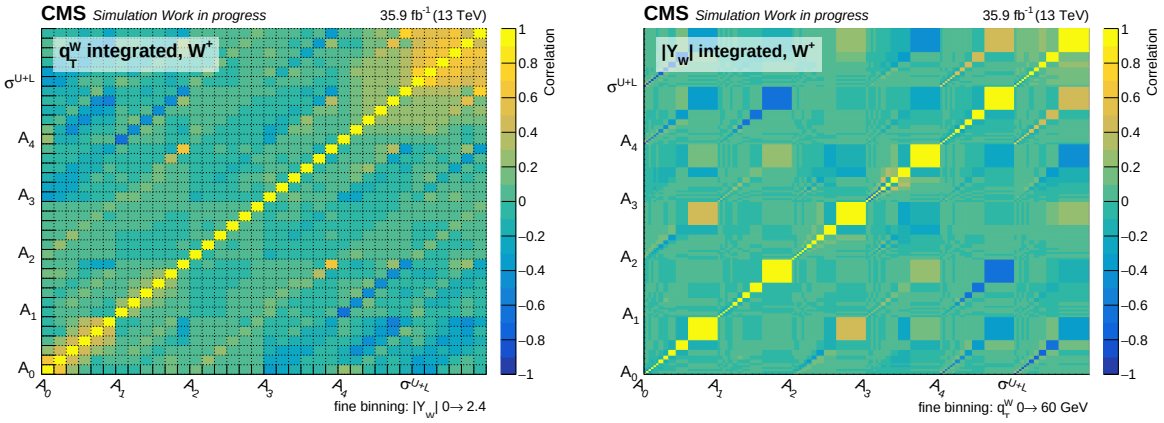
Figure 7.12. *Different sectors of the correlation matrix of the fit performed on the Asimov dataset for the double-differential σ^{U+L} and the five angular coefficients distributions, for the W^+ sample. Each matrix sector contains the correlation between all the $|Y_W|$, q_T^W processes for a given angular coefficient. The correlation between the processes of different angular coefficients and the correlations with the nuisance parameters are not shown. The BBB uncertainty has not been considered. The widths of the bins of the matrices are proportional to the width of the corresponding q_T^W bins.*



(a)

(b) (c)

Figure 7.13. *Relative uncertainty on σ^{U+L}, double differential in $|Y_W|$ and q_T^W (a), differential in $|Y_W|$ (b), differential in q_T^W (c) for W^+ sample. The Poissonian line is the expected $1/\sqrt{n}$ uncertainty in absence of systematic uncertainties and correlations. The statistical uncertainty is the fit uncertainty without considering the effect of the nuisance parameters.*



(a) (b)

Figure 7.14. *Full A_i, σ^{U+L} correlation matrix for the single differential measurement in $|Y_W|$ (a) and for the single-differential measurement in q_T^W (b), performed on the Asimov dataset for the W^+ samples. The correlation between the A_i, σ^{U+L} and the nuisance parameters are not shown. The BBB uncertainty has not been considered. The widths of the bins of the q_T^W-differential matrix is proportional to the width of the corresponding q_T^W bins.*

7.4.3 Impact of systematic uncertainties

The discussed uncertainties on the POIs include both the statistical uncertainty and the effect of the systematic uncertainties considered in the analysis. It is relevant to investigate how the latter impact the POIs and if the related nuisance parameters are constrained by the fit. The effect of systematic uncertainties on the POIs of the analysis is presented in term of *impacts*. The impact of the nuisance parameter θ_k on the POI μ_p is defined as:

$$I_{\theta_k}(\mu_p) = \frac{V_{p,k}}{\sigma_k}, \tag{7.5}$$

where $V_{p,k}$ is the covariance of μ_p and θ_k, and σ_k is the post-fit uncertainty of θ_k. In the limit of Gaussian uncertainties $I_{\theta_k}(\mu_p)$ corresponds to the shift $\Delta\mu_p$ induced on μ_p as θ_k is fixed to $\theta_k \pm 1\sigma_k$ and all the other parameters are profiled as usual.

Because of the large number of nuisance parameters considered in the fit, it is worth subdividing the nuisance parameters into groups, considering that the nuisance parameters are strongly correlated within a group and weakly correlated between the different groups. One group is defined for each of the normalization and shape uncertainty sources listed in Sec. 7.3. The groups are mutually exclusive and cover all the systematic uncertainties considered in the analysis. The impact of a group is the result of the combined variation of all the nuisance parameters within the group. Formally:

$$I_{\{\theta_k\}_G}(\mu_p) = \sqrt{C^T_{p,G} V_G^{-1} C_{p,G}}, \tag{7.6}$$

where V_G is the subset of the covariance matrix correspondent to the nuisances of the group G and $C_{p,G}$ is the covariance matrix between μ_P and the nuisances of the group. An additional group is defined for the statistical uncertainty, estimated from the covariance matrix without systematic uncertainties, but using post-fit nuisance parameters values. Note that the sum of the group impacts can be smaller than the total uncertainty on the POIs because of the correlation between different groups.

The same definitions can be used to define the impact of the nuisance parameters on any quantity derived from the POIs. In Fig. 7.15 the impacts of the nuisance parameters on some representative quantities are shown i.e. σ^{U+L} and A_4, single and double-differential distributions, from the fit to the Asimov dataset. They can be interpreted as the relative or absolute uncertainties for σ^{U+L} or A_4, respectively, subdivided in the various sources. The entire set of impacts for all the angular coefficients and both W charges is reported in Appendix E.

The $d\sigma^{U+L}/d|Y_W|$ is limited by the data luminosity uncertainty. For all the other quantities the leading source of uncertainty is the statistical one, in the entire phase space, as previously anticipated. The total precision is actually above the 2.5% level for $d\sigma^{U+L}/dq_T^W$ and $d\sigma^{U+L}/d|Y_W|dq_T^W$, while in the angular coefficients the cancellation of the normalization uncertainty like luminosity occurs, thus the luminosity impact is subleading.

The other relevant systematic uncertainty is the statistical uncertainty of the efficiency scale factors, both for σ^{U+L} and angular coefficients. This uncertainty is at 1-2% level in the single differential σ^{U+L} distributions, and 2%-8% in the double differential one. The absolute impact of the scale factors follows the statistical uncertainty trend in A_4, increasing in the high Y_W region, and it represents the main source of systematic uncertainty.

Finally, also the PDF uncertainty and the q_T^V uncertainty contribute at 1% (10%) level in double differential σ^{U+L} (A_4) uncertainties. The trend of these impacts in the double differential results also follows the statistical uncertainty trend, despite there is no intuitive physical meaning to this behaviour. However, it must be noted that their value is 1-2 orders of magnitude lower than the leading uncertainty. In this condition, this feature is possibly a numerical artefact due to the

method used to calculate the impacts. The statistical uncertainty is the dominant component of the covariance matrix. The inversion of the latter (from Eq. 7.6) can produce artificial trends ruled by the leading eigenvalues. Therefore, the result in this condition must be considered only as an order of magnitude estimation.

Because of this interplay, to analyze the impact of the subleading uncertainties, the single differential cases must be considered. In the latter cases, the impact of statistical uncertainty compared to the subleading uncertainties is smaller and thus the impact evaluation is more robust.

The template fit also has the capability to constrain the nuisance parameters, thanks to their correlation with the measured quantities. The post-fit unrolled $\eta^{\mu}(p_T^{\mu})$ distribution can be used to see the overall effect on the systematic uncertainties. The post-fit distributions are shown in Fig. 7.16. The "data" process reported in this figure are not real, but represent exactly the sum of all the expected processes, since the fit has been performed on the Asimov dataset. Thanks to the fit the systematic uncertainties band has been reduced from the 5% of the pre-fit to 1%-0.1% of the post-fit.

The constraint on the single nuisance parameter can be also studied. The likelihood encodes the dependence between the POIs and the nuisance parameters as the multiplicative factors $\kappa_{i,p,k}^{\theta}$, where $\theta = \pm 1$ produces the expected event template variation. Thus, the pre-fit value of θ_k can be represented as a null value with unit uncertainty. The constraint produced by the fit to the nuisance θ_k can be expressed as the post-fit shift compared to zero, and the post-fit uncertainty compared to 1. The results from the Asimov dataset fit are shown in Fig. 7.17 for the entire set of nuisance parameters. The shift of the central value is omitted since in the fit to Asimov dataset it is below 10^{-10}, also if μ_p are left freely floating.

The following features are observed:

- The fit is capable to constrain the PDF nuisances at 90% level, with some specific eigenvector constrained up to 70%-60%. This is coherent with the result reported in Ref. [11].
- The p_T^{miss}-related nuisances are extremely constrained, at 10%-5% level. These nuisances are related to the POIs only via the m_T selection applied to the events. However, it has been shown in Fig. 6.11 and Fig. 6.12 how this selection produces a 1% variation of the $p_T^{\mu} \times \eta^{\mu}$ spectrum on average, with a non-trivial trend in p_T^{μ}, and a variation up to 10% in the high-m_T region. The template fit is capable to predict the event yield distribution with higher precision (1%-0.1%, from Fig.7.16), thus it constrains the single nuisance parameter which control the size of the shape variation. This constraint to the p_T^{miss} uncertainty cannot be considered valid in general, but only in this particular region of the phase space, because of the specific event selection applied. If nuisances are decorrelated in p_T^{μ} the constraint will be strongly reduced (70%-80%).
- The p_T^{μ} momentum scale nuisance parameter is extremely constrained at 3% level. The reason is the same of the p_T^{miss} nuisance parameters, but in this case the constraint is an artefact of the simplified p_T^{μ} scale description, which has been encoded as a single nuisance parameter, fully correlated in the phase space. A more refined description, with a set of nuisance parameters uncorrelated in $\eta^{\mu} \times p_T^{\mu}$ will not produce this effect.
- The nuisance parameter related to the luminosity is constrained at 60% level. The signal processes have not the capability to constrain luminosity, because the cross section is predicted from the fit. However the low-acceptance and $Z \to \mu\mu$ background templates are exploited using fixed signal strength parameter, and they do not have dedicated nuisances for their normalization, but their variation is described via several shape nuisance parameters. Thus, they can partially reduce the luminosity nuisance parameter uncertainty. This hypothesis has been tested removing the two background process from the fit to the Asimov

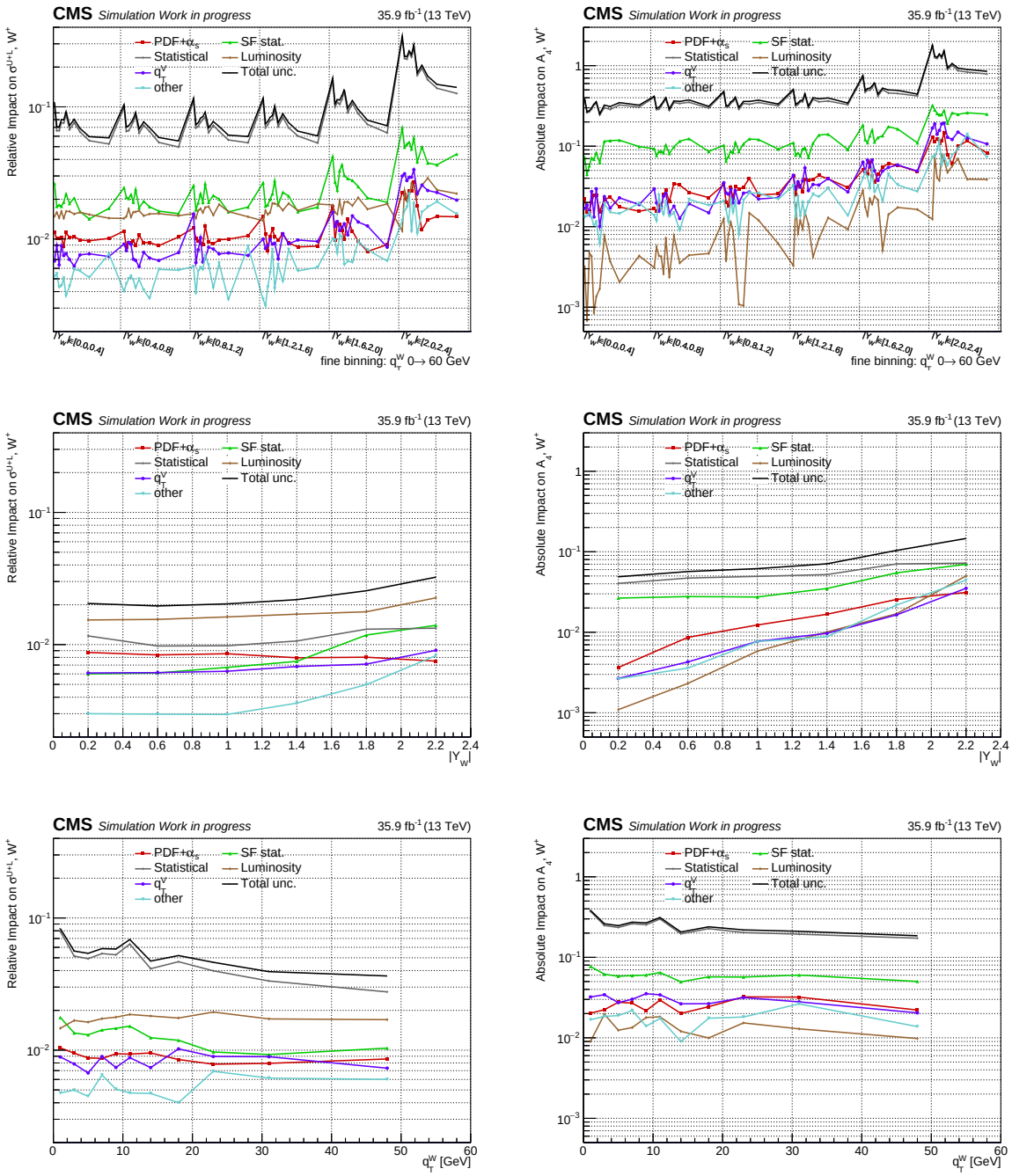


Figure 7.15. *Relative impacts on σ^{U+L} (left) and absolute impact on A_4 (right) double-differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central) and differential in q_T^W (lower). The impacts are evaluated for each group and then divided by the quantity itself. The group "other" includes the lepton veto, the electroweak background cross sections, the QCD normalization, the MET-related systematics, the p_T^μ scale, the systematic uncertainty on the SF, the L1 Trigger prefire, m_W. The BBB uncertainty has not been considered coherently to the fit results.*

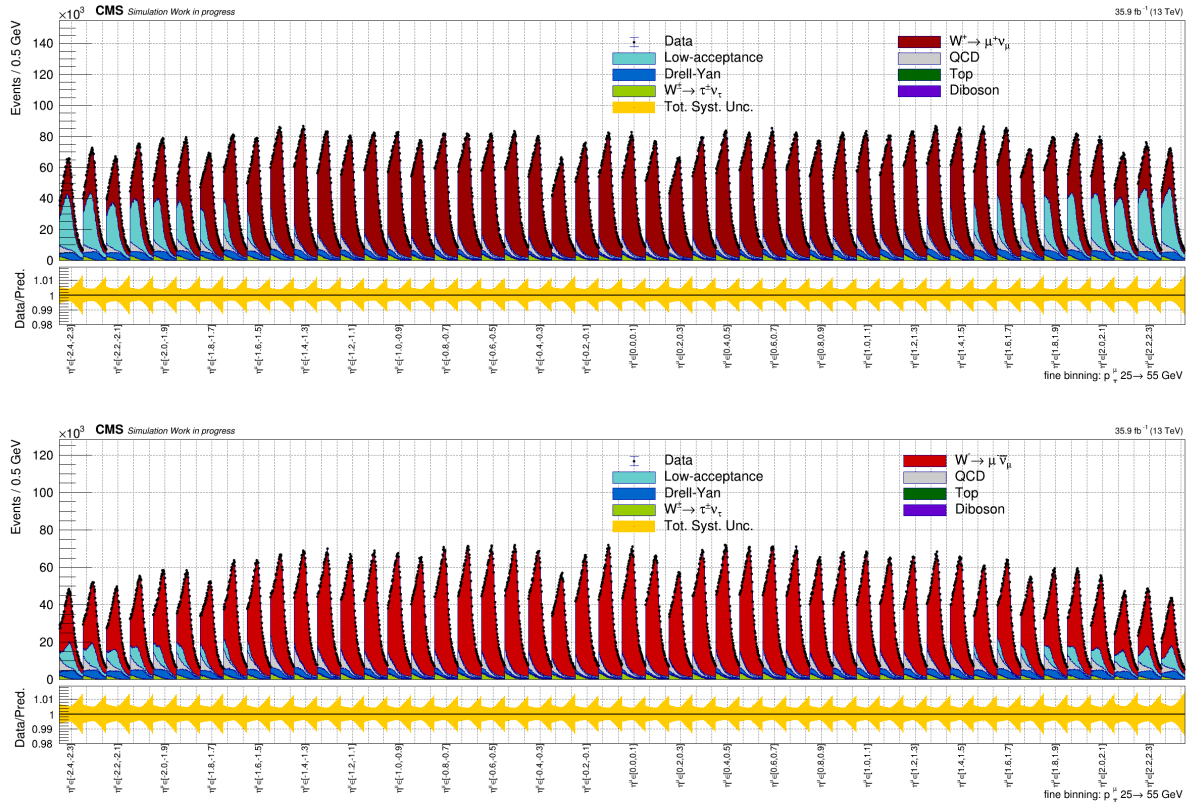


Figure 7.16. *Stacked distribution of the unrolled templates for W^+ (upper) and W^- (lower) channel, from the output of the fit to the Asimov dataset. The data process corresponds to a MC equivalent to the sum of all the expected processes. The $W^\pm \to \mu^\pm\nu$ process includes the sum of all the signal processes. The post-fit uncertainties are included in the band in the panels below each distribution.*

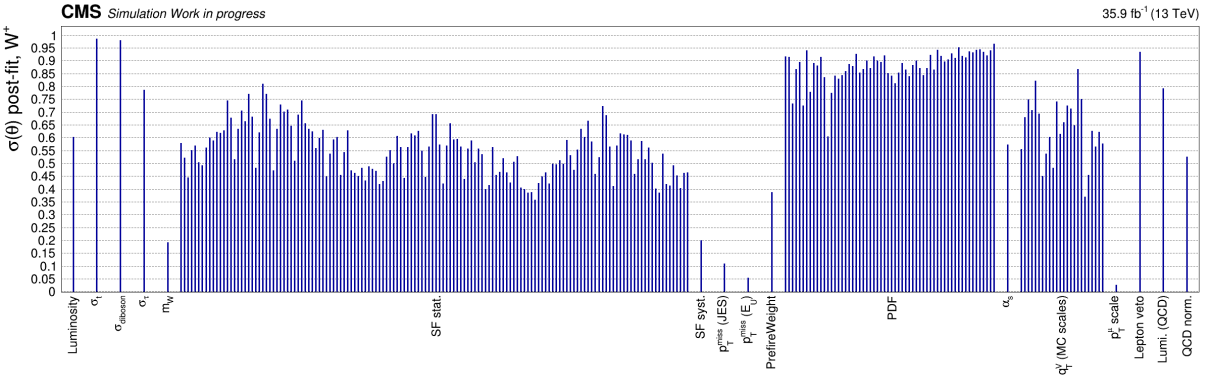


Figure 7.17. *Post-fit nuisance parameters uncertainties, to be compared to $\sigma(\theta_{prefit}) = 1$, $\theta_{prefit} = 0$. The shift in the central values is not shown, since the fit has been performed on the Asimov dataset.*

dataset, and the constraint to luminosity nuisances relaxes from 0.58 to 0.92.

– The fit is able to constrain the nuisances related to the efficiency scale factor, at 50%-70%, due to the data-MC comparison implicit in the template fit.

– Several background-related nuisances are constrained, in particular the ones related to the most relevant sources (QCD background and the $Z \to \mu\mu$). The background nuisances are constrained both in shape via the MC scales, and normalization.

– the m_W nuisance is constrained at 20% level, revealing the sensitivity of the analysis to the m_W value and thus to the possible constraint that can be set. This aspect will be discussed in Sec. 8.1.

7.4.4 Toy MC exercise

The results of the fit to the Asimov dataset has been validated using toy MC experiments, as described in Sec. 7.4. Because of the long time required by the fit (several hours fully exploiting the parallelization), a single toy has been fitted. The toy is produced with a randomization of the Asimov pseudo-data distribution of the $W^+ \to \mu^+\nu$ sample, and thus it is statistically equivalent to the fit to the Asimov dataset. It can be considered representative of the performance of the fit when it is performed starting with POIs far from the global minimum, differently from the fit to the Asimov dataset case.

The χ^2 of the fit is 428.2 with 396 degrees of freedom, corresponding to a p-value of 0.1275. Thus, the agreement between data and the prediction is acceptable. Some examples of the unfolded distributions are reported in Fig. 7.18, for σ^{U+L} and A_4. Both in double and single differential case the fit to the toy behaves as expected from the fit to the Asimov dataset. The uncertainties are compatible with the ones discussed on fit the Asimov dataset. The predicted central values discrepancy is compatible with the fit the Asimov dataset uncertainty, within the toy fit uncertainty. Despite μ_p are freely floating and far from the minimum, the expected [8, 11] high-frequency oscillation of the predicted central values does not appear as a clear up-down shift of the neighbour bins. Because the overlap between the templates occurs between neighbour bins in multiple dimensions ($|Y_W|$, q_T^W) and between multiple adjacent bins, only the uncertainties inflated by the (anti)correlation pattern are observed. Only the analysis of the correlation matrix, as discussed in Sec. 7.4.1, allows to observe a clear pattern.

A sample of 10^3 toy MC experiments has been used to obtain a statistically significant validation. However, because of the time requirement of the fit, all the nuisance parameters, but m_W, have been removed from the likelihood of the fit. This is a strong simplification, which produces reduced uncertainties. On one hand, the result can be considered significant since the statistical uncertainty is the dominant source of uncertainty, as shown and discussed in Sec. 7.4.3. On the other hand, the removal of the nuisance parameters can produce biases, since a non-zero θ_k predicted by the fit can modify the n_i^{exp} (from Eq. 7.1). In conclusion, this result must be considered only to add statistical robustness to the single-toy complete fit result of Fig. 7.18, but not as general performance of the fit. With this *caveat*, the pull distribution has been built for each POI from the 10^3 toy sample.

The mean value and the variance of the pull distributions is shown in Fig. 7.19. The values are the result of a Gaussian fit, for each bin of the double and single differential σ^{U+L} and A_i distributions. The mean value of the pull distributions is compatible with 0 and the variance is compatible with 1, within the expected uncertainties in most of the phase space.

7.5 Regularization

In the discussion of the fit results, it has been highlighted how the uncertainties are inflated by the competition between multiple neighbour templates in the same region of the $\eta^\mu \times p_T^\mu$ phase space. To mitigate the effect, the typical approach is to add a regularization term to the likelihood, which breaks the (partial) degeneracy of the minimum and allow to recover a physical behaviour in the predicted distributions.

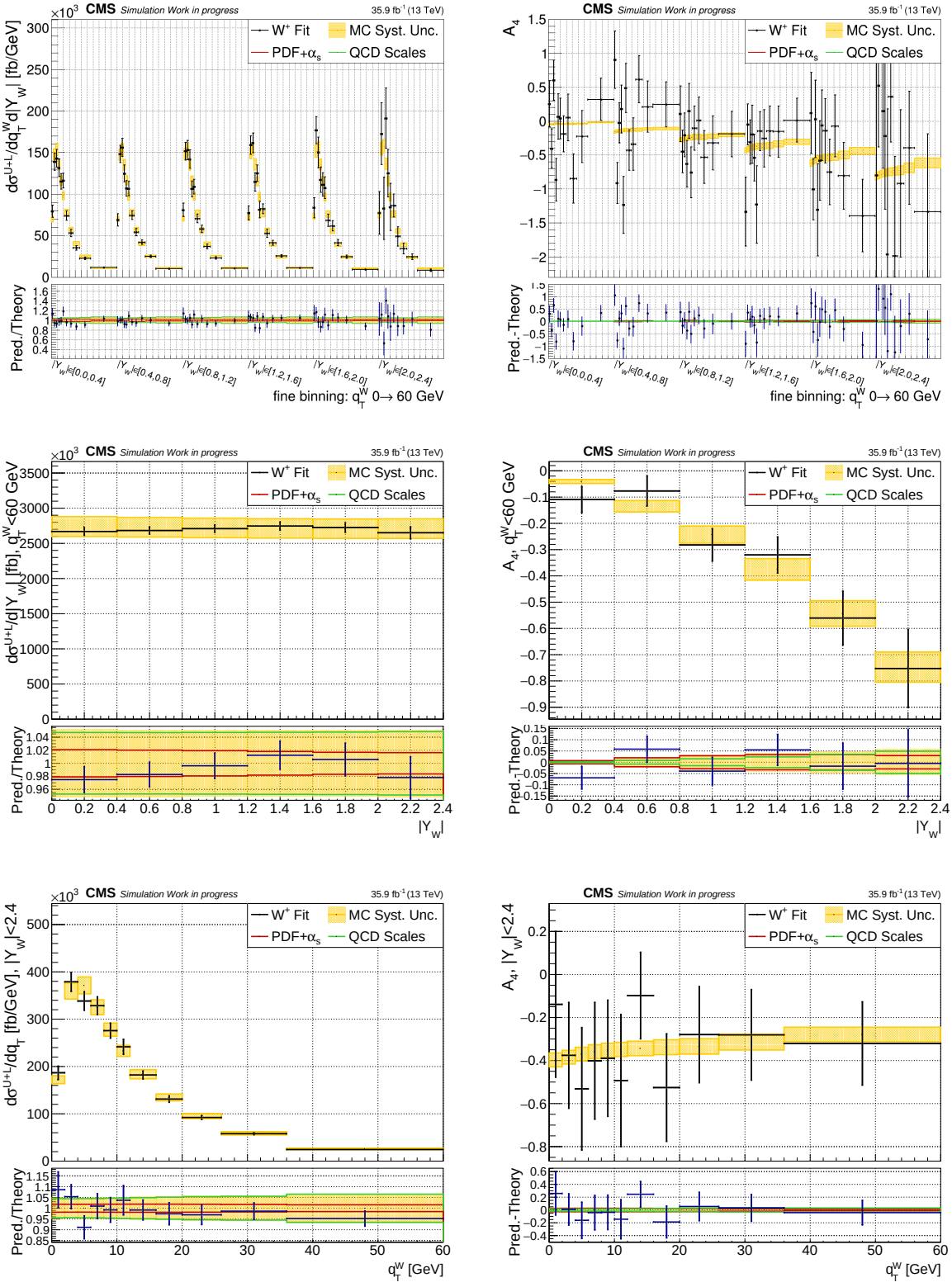


Figure 7.18. *Fitted distribution of σ^{U+L} (left) and A_4 (right) double-differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central) and differential in q_T^W (lower). The template fit has been performed on a single toy MC experiment. In the panel below each distribution the difference between the fitted value and the expected one is shown, together with the band of the theory uncertainties, highlighting the two sources. The template fit results in this panel are removed in the double differential result on A_4 for clarity. The BBB uncertainty has been not considered*

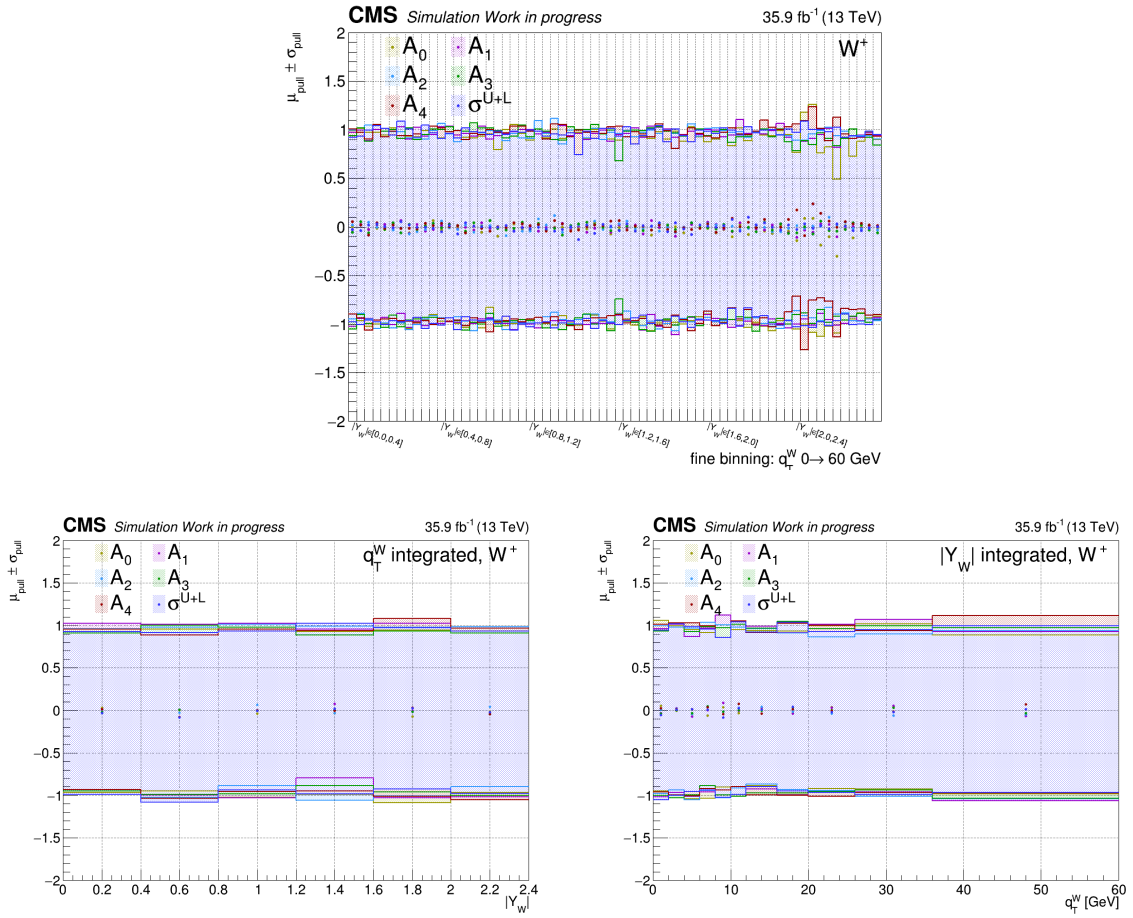


Figure 7.19. *Mean value and variance of the pull distributions of* 10^3 *toy MC experiments for* σ^{U+L} *and five* A_i*, for each* $|Y_W|$*,* q_T^W *bin, in the double differential case (upper plot), differential in* $|Y_W|$ *(left plot) or differential in* q_T^W *(right plot). The fit has been performed on the* W^+ *sample, removing all the nuisance parameters from the likelihood but* m_W*.*

In Sec. 5.2.3 it has already been discussed how, independent from the correlation pattern issue, a regularization can improve the precision of the template fit without producing a bias in the results, thanks to the smooth behaviour of the angular coefficients. The most natural choice is to regularize them with the hypothesis of a fixed-order polynomial behaviour, as a function of $|Y_W|$ and q_T^W. The implemented regularization scheme has been called post-fit linear fit. In this approach, the regularization is performed after the template fit. Therefore, the template fit is performed without regularization, and then an additional fit is performed on the first fit outcome. This second fit uses an explicit polynomial parameterization for each angular coefficient, linear in the fit parameters:

$$A_i(|Y_W|, q_T^W) \sim p_i(|Y_W|, q_T^W) \equiv \left(\sum_{y=0}^{n_Y} K_{i,y} c_{i,y} |Y_W|^y \right) \times \left(\sum_{q=0}^{n_{q_T}} K_{i,q} c_{i,q} (q_T^W)^q \right), \tag{7.7}$$

where the polynomial coefficients $c_{i,j}$ are the parameters predicted by the polynomial fit, n_Y (n_{q_T}) is the number of Y_W (q_T^W) bins of the fit, and $K_{i,j}$ can be 0 or 1, and represents a mask that defines which degrees must be fitted ($K_{i,j} = 1$) or should be 0 ($K_{i,j} = 0$).

The 5 angular coefficients are fitted simultaneously, minimizing the χ^2:

$$\chi^2 = (\mathbf{A} - \mathbf{p})^T V^{-1} (\mathbf{A} - \mathbf{p}), \tag{7.8}$$

where $\mathbf{A}$ is the vector which contains the unrolled values of the 5 angular coefficients in bin of Y_W, q_T^W and V is the covariance matrix as predicted from the template fit, and $\mathbf{p}$ is the vector of the polynomials from Eq. 7.7, evaluated for each value of Y_W, q_T^W.

The assumption of smooth behaviour is valid for the angular coefficients only, thus σ^{U+L} has not been regularized but it is included in the fitting model together with the nuisance parameters of the template fit, adding additional terms $c_{y,q}\sigma_{y,q}^{U+L}$, $c_k\theta_k$, where $c_{y,q}$, c_k are additional polynomial fit parameters, and $\sigma_{y,q}^{U+L}$, θ_k the unpolarized cross section and nuisance parameters predicted the template fit results, respectively. The covariance matrix V from Eq. 7.8 also includes the corresponding terms.

7.5.1 Post-fit regularization results

The post-fit linear fit regularization has been applied to the predicted distribution from the template fit to the Asimov dataset. The full covariance matrix of the fit has been used to take into account of correlation between the 5 angular coefficients, σ^{U+L} and the nuisance parameters. The results of the 2-D polynomial fit have been integrated in q_T^W or $|Y_W|$ to obtain the single-differential distributions, with the proper propagation of the covariance matrix. The fit and the uncertainty propagation procedures have been performed using a minimizer developed in `JAX` [201]. `JAX` is a python package which implements automatic differentiation and allows precise and accurate evaluation of gradient and hessian matrices.

The $K_{i,y}$ and $K_{i,q}$ masks from Eq. 7.7, which defines the polynomial models have been chosen in the following way:

- all the angular coefficients have been fitted considering up to $(q_T^W)^3$ terms
- A_0, A_1, A_2 and A_3 have been fitted considering up to $(Y_W)^2$ terms, for A_4 fit also $(Y_W)^3$ terms have been considered.
- The coefficient $c_{q=0}$ is set to 0 for A_0, A_1, A_2 and A_3.
- The coefficient $c_{y=0}$ is set to 0 for A_1 and A_4.

This parametrization includes the physics-motivates constraints at $Y_W = 0$ and/or $q_T^W = 0$, from the symmetries discussed in Sec. 5.2.3. The maximum degree considered for each angular coefficient has been chosen to obtain numerically stable and robust results, without biasing the central value of the predicted A_i distributions, compared to the template fit results to the Asimov dataset.

The results are presented in Fig. 7.20 for selected distributions (σ^{U+L} and A_4), representative of the general performance of the fit. The full results for all the angular coefficients and both W charges are reported in Appendix E. The post-fit regularization produces a negligible bias on the angular coefficient. The induced bias is below 0.02 level on the angular coefficient (concentrated in the high Y_W, q_T^W region) and below 0.4% on σ^{U+L} in the double-differential prediction. This bias is reduced by an order of magnitude in the single-differential case. Since the regularization has been performed on the result of the template fit to the Asimov dataset, its impact is observed mainly as an uncertainty reduction. In Fig. 7.21 the total uncertainty before and after the regularization are compared. The post-fit linear fit is able to strongly reduce the fit uncertainties both on the observables directly fitted, like A_4, and on the observables correlated to that, like

σ^{U+L}. The value of the reduction depends on the specific observable and region of the phase space. Focusing on the single-differential distribution a reduction factor 2-3 is achieved in most of the phase space. Only the high-q_T^W region is not strongly constrained.

Particularly relevant is the constraint provided to $d\sigma^{U+L}/dq_T^W$, which allows to reach a 3% precision in the entire considered range. Such a precise description of the q_T^W spectrum in the region below 5 GeV can strongly reduce the m_W uncertainties related to the production model. Note that no direct assumption has been done on the shape of the $d\sigma^{U+L}/d|Y_W|dq_T^W$, but this result is derived only from the correlation between the angular coefficients and σ^{U+L}.

To avoid any strict polynomial assumption also on the angular coefficient different regularization scheme can be adopted. In Appendix F a regularization approach currently under development is presented, which allow to perform the regularization simultaneously to the template fit with a proper modification of the likelihood. In that framework a more refined optimization method for the polynomial parametrization and the choice of the $K_{i,y}$ and $K_{i,q}$ masks is also presented.

7.6 Summary of the fit results

The presented measurement of the angular coefficients is limited by the data statistical uncertainty. The latter does not allow to resolve the angular coefficients in the double-differential or single differential cases. Only A_4 integrated in q_T^W can be resolved, but with an uncertainty larger than the PDF theory uncertainty. On the other hand, the single-differential σ^{U+L} distributions i.e. the rapidity and transverse momentum distribution of the W boson can be measured with a 2% and 8-4% precision, respectively.

The presented approach demonstrates the capability to convert a systematically limited measurement into a statistically limited measurement. The systematic uncertainties on the measurement of the angular coefficients and double-differential unpolarized cross sections are strongly subleading and related to experimental features which can be improved (like the scale factors measurement, or q_T^W spectrum modelling in QCD background measurement) or which does not affect the shape of the differential spectra (like data luminosity). The first real limitation is the PDF residual uncertainty. The template fit approach constrains the PDF uncertainty at 1%-2% level on the unpolarized cross section and at 10^{-2} level on the angular coefficients.

This precision matches the requirement discussed in Sec. 4.4.1 to obtain a systematic uncertainty on m_W due to q_T^W below 10 MeV. For what concerns the PDF uncertainty on m_W it is not trivial to propagate the uncertainty from Y_W and A_i to m_W. In Sec. 8.2 an explicit test to assess this uncertainty will be discussed.

The statistical limitation of the measurement can be faced with two possible approaches. First, the presented measurement exploits only a small fraction of the CMS data. The statistical uncertainty can be reduced of a factor 2 exploiting the full data sample collected during the Run 2, corresponding to an integrated luminosity of 163 fb^{-1}, compared to the 35.9 fb^{-1} adopted in this proof of feasibility. Also the residual PDF systematic uncertainty is expected to improve with an increased luminosity. Second, a regularization scheme can be adopted as mentioned in Sec. 5.2.3. This second approach improves the angular coefficient distributions description and reduces the impact of the correlation terms. Preliminary results showed how the use of a polynomial regularization for the angular coefficient distributions reduces the uncertainties of a factor 2-3.

Is worth to mention that also with the current statistical limitation the W rapidity measurement is capable to confirm the result of Ref. [11] and perform the first measurement of $d\sigma/dq_T^W$ at $\sqrt{s} = 13$ TeV, with 8%-4% precision. This is further reduced to 3% if the polynomial regularization is used.

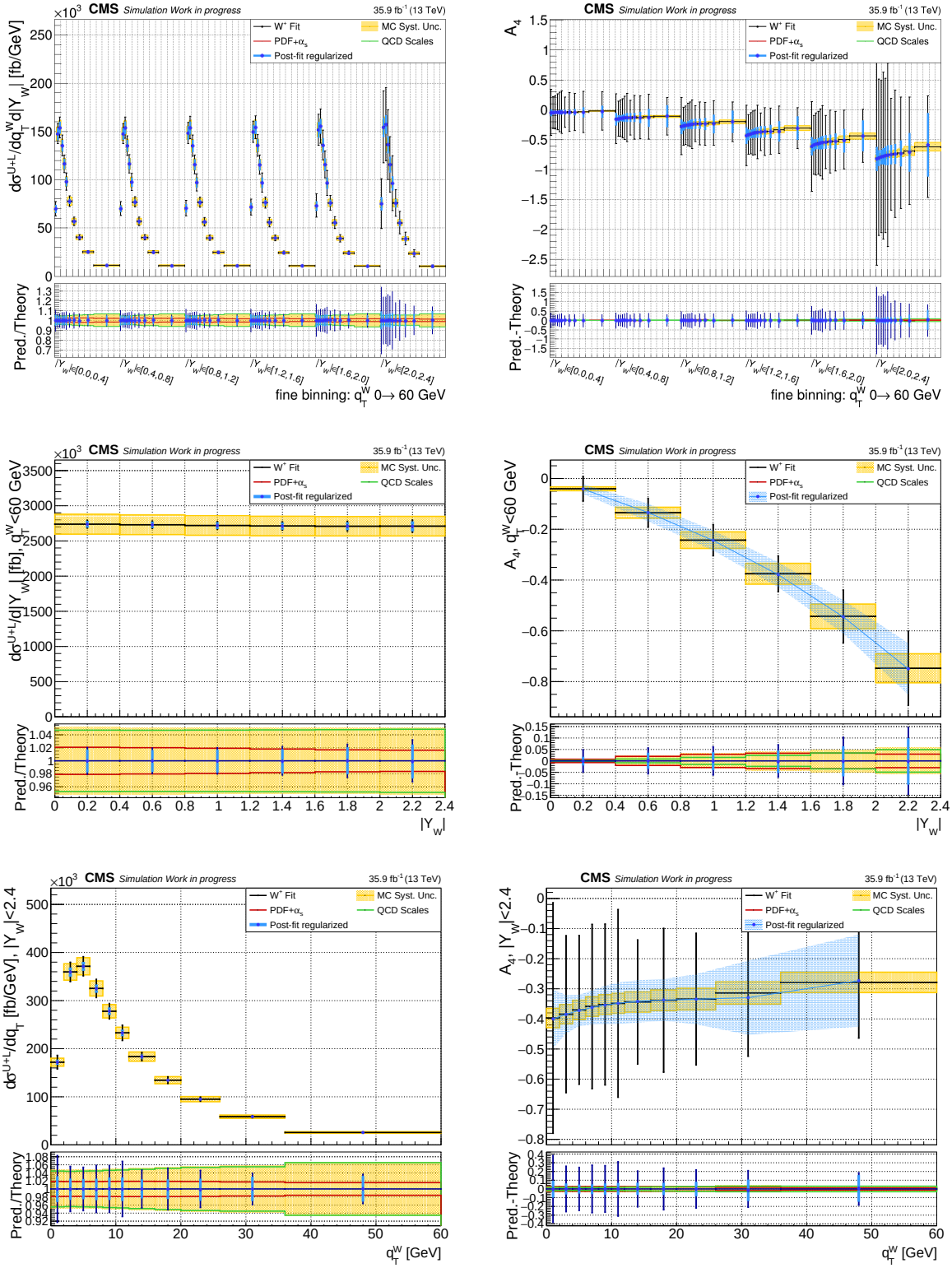


Figure 7.20. *Fitted distribution of σ^{U+L} (left) and A_4 (right) double-differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central) and differential in q_T^W (lower). The template fit has been performed on the Asimov dataset. The post-fit regularization has been performed after the template fit and is shown as a blue band in the single-differential case or only with the uncertainty bands in the double-differential case and in σ^{U+L}. In the panel below each distribution the ratio (difference) between the fitted value and the expected one is shown for σ^{U+L} (A_4), together with the band of the theory uncertainties, highlighting the two sources. The ratio (difference) between the post-fit regularization and the template fit results is shown. The template fit uncertainties in this panel are removed in the double differential result for clarity. The BBB uncertainty has been not considered*

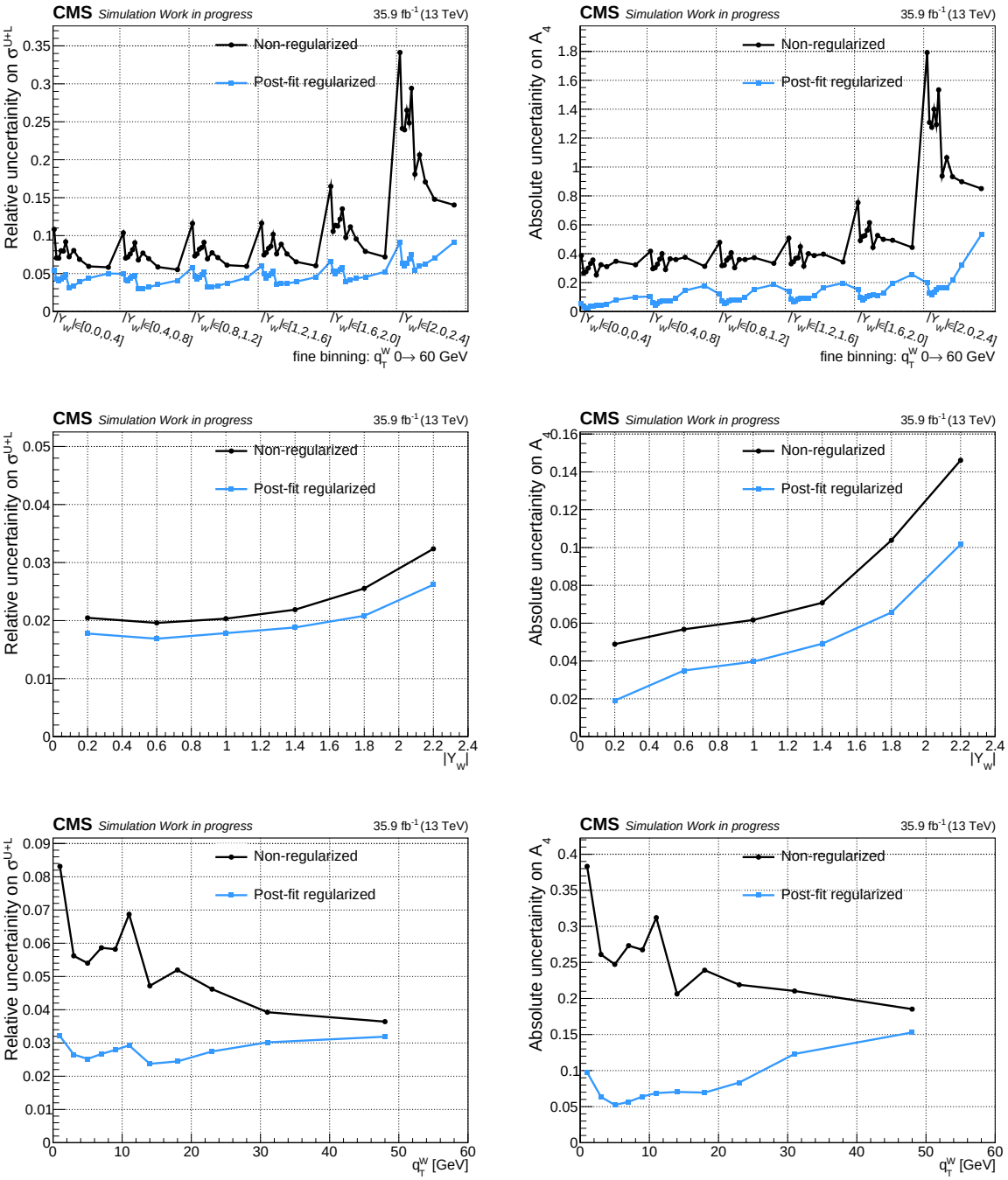


Figure 7.21. *Relative uncertainty from the fit on σ^{U+L} (left) and absolute uncertainty on A_4 (right) in double-differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central) and differential in q_T^W (lower) cases. The template fit has been performed on the Asimov dataset in all the cases. The uncertainty has been estimated directly from* `Combine-TensorFlow` *fit in the non-regularized case. The uncertainty of the Post-fit regularized case has been estimated with the* `JAX` *fit on the top of the* `Combine-TensorFlow` *non-regularized fit. The BBB uncertainty has not been considered.*

Chapter 8

Fit to the W boson mass

In this chapter a preliminary simultaneous fit to m_W and the W production properties is presented.

8.1 Fit to m_W

The possibility to simultaneously fit m_W and the W boson production properties has been discussed in Sec. 5.2.5. The current analysis has not been designed for the purpose of performing a fit to m_W. It would require the introduction of several replicas of the entire set of templates produced with different value of m_W, to cope with the additional differential dimension of $\frac{d\sigma}{dq_T^W dY_W dm_W}$.

Nevertheless, the current analysis implementation has the capability to perform a simplified m_W prediction exploiting the m_W nuisance parameter. The Gaussian constraint related to the m_W nuisance parameter θ_{m_W} has been removed from the likelihood, allowing m_W to be freely floating. Then, the fit is repeated. This is equivalent to have only 3 templates differential in m_W i.e. the nominal one and the up/down template, which correspond to a 50 MeV shift of m_W. The description with a single nuisance parameter correlated for all the channel is consistent because of the physical interpretation of m_W parameter.

The fit has been performed on the Asimov data sample, without the BBB uncertainty, coherently with the rest of the analysis. No regularization has been applied to this fit. The result of this fit has a double purpose. Firstly, if the μ_p and the θ_k are left free is possible to estimate the possible bias induced by the fit on the m_W value. Secondly, the uncertainty on m_W can be estimated from the post-fit uncertainty on θ_{m_W}. The contribution of the different source of uncertainties can be estimated by studying the impacts $I_{\theta_k}(m_W)$, with the same interpretation of the impact on the POI $I_{\theta_k}(\mu)$. The results are reported separately for W^+ and W^-, because the fit is performed separately for the two charges.

The value of m_W from the fit to the Asimov dataset is:

$$\begin{aligned} m_{W^+} &= m_W^{\text{nom}} \pm 9.7 \text{ MeV} \pm \Delta m_{W^+}, \\ m_{W^-} &= m_W^{\text{nom}} \pm 9.9 \text{ MeV} \pm \Delta m_{W^-}, \end{aligned} \tag{8.1}$$

where $m_W^{\text{nom}} = 80419$ MeV and the numerical value is the total uncertainty estimated by the fit, which contains the statistical and the systematical component. $\Delta m_{W^\pm}$ is the residual uncertainties not included in the performed fit, discussed in the next paragraphs. The breakdown of

the included uncertainties is shown in Fig. 8.1. The bias induced by the fit on the nominal m_W value is much smaller than 0.1 MeV.

The result of Eq. 8.1 demonstrates the capability of the template fit to extract simultaneously m_W and the production properties. This approach automatically takes into account of the W boson production-related systematic uncertainties, via the combined covariance matrix. The results of Eq. 8.1 is not claiming that the current analysis is capable of performing a measurement of m_W with a precision below 10 MeV. Also excluding the limited MC size, several central ingredients have been neglected and they are included in the factor $\Delta m_W^{\pm}$. In particular a complete description of momentum scale uncertainty and the systematic uncertainty related to FSR will likely produce a relevant impact on m_W.

The current result demonstrates that it is already possible to reduce the impact of the W production model below 10 MeV level, with a simultaneous fit to m_W. In this approach the systematic uncertainty related to the W production model is converted into statistical uncertainty on m_W.

Given the aforementioned prescriptions, Fig. 8.1 shows how the residual PDF and q_T^W uncertainties are below the MeV level. The uncertainty of q_T^V is due to the background processes. The higher events yield in the low-acceptance background in the W^+ sample produces the larger observed impact compared to W^-. The q_T^W uncertainty related to the signal is instead included in the statistical component, since the q_T^W is fitted simultaneously to the mass and no dedicated nuisance parameters are needed. The source of PDF uncertainty is partially the low-acceptance background process and partially the residual dependence of signal processes on the PDFs. Finally, the efficiency SFs are the leading systematic uncertainty in this condition, highlighting again the relevance to improve the precision in the scale factor measurement.

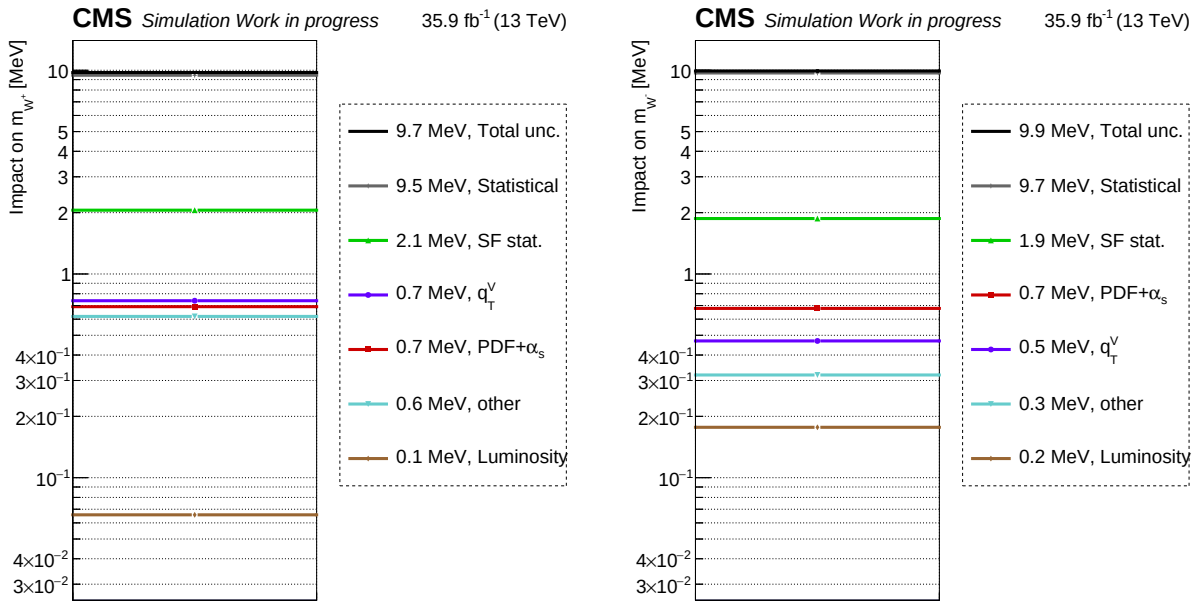


Figure 8.1. *Breakdown of the uncertainties on m_W, for W^+ (left) and W^- (right) fit to the Asimov dataset. The values have been estimated for each group of uncertainties as described in Sec. 7.4.3. The fit has been performed on the Asimov dataset, fitting m_W simultaneously to $A_0, \dots A_4$ and σ^{U+L} double-differential cross sections. The group "other" includes the lepton veto, the electroweak background cross sections, the QCD normalization, the MET-related systematics, the p_T^{μ} scale, the systematic uncertainty on the SF, the L1 Trigger prefire. The BBB uncertainty has not been considered.*

In the Chapter 7 has been outlined how the template fit without regularization is capable to predict the $d\sigma^{U+L}/dq_T^W$ with 4%-8% precision. In Sec 4.4.1 has been discussed how this precision is not sufficient to constraint the related systematic uncertainty on m_W below 10 MeV. The result of Eq. 8.1 seems to contradict this statement. However, the $d\sigma^{U+L}/dq_T^W$ precision is inflated by the correlation pattern of the fit. The simultaneous prediction of m_W is robust against this effect because m_W is only mildly affected by the correlation pattern. To support this statement two tests have been performed.

First, the post-fit regularization fit has been run on the result of the simultaneous fit to m_W and W production properties. The predicted m_W uncertainty is 8.8 MeV, with a bias of -0.02 MeV. As expected the improvement compared to Eq. 8.1 is small, because the latter is not strongly affected by the correlation pattern.

Second, the template fit to the W production properties has been repeated fixing $A_0(q_T^W, Y_W)$ to the nominal value from MC. A_0 is the coefficients which higher (anti)correlation with σ^{U+L}, in particular in the low q_T^W region, from Fig. 7.14. The uncertainly on $d\sigma^{U+L}/dq_T^W$ is reduced of a factor 2. On the other hand, the m_W uncertainty is almost unchanged (9.5 MeV, to be compared to Eq. 8.1). This shows the correlation effect on the σ^{U+L} uncertainty and demonstrates the low dependence of m_W uncertainty on this effect.

8.2 Constraint on PDF uncertainty

Instead of fitting m_W simultaneously to the W boson production properties, the same fitting framework allows to propagate the PDF uncertainty, constrained by the fit of Chapter 7, on m_W, mimicking a standard template fit to m_W using the $\eta^\mu \times p_T^\mu$ distribution. A fit to m_W fixing all the POIs has been performed on the Asimov dataset. This fit, fixing the μ_p to the nominal value, is assuming an external knowledge of the angular coefficients, Y_W and q_T^W distributions. The nuisance parameters are treated as usual. The numerical results are reported for W^+ sample. The W^- sample shows similar results.

The template fit to m_W, fixing all the POIs and the PDF nuisance parameters to their pre-fit values and uncertainties predicts a systematic uncertainty due to the PDF on m_W of 12.7 MeV. This is compatible to the expectations [3].

The template fit to m_W, fixing all the POIs and allowing the PDF nuisance parameters to be constrained, predicts a PDF uncertainty of 3.0 MeV on m_W. This is equivalent to the procedure performed in Ref. [11] to assess the PDF constraint, but in this framework it is possible to propagate the constrained uncertainty directly to m_W.

The template fit to m_W, fixing all the POIs predicts a statistical uncertainty of 2.0 MeV on m_W. This is in good agreement with a scaling of the 10.1 MeV of statistical uncertainty estimated by ATLAS m_W measurement [3] to the number of W boson candidate exploited in this work ($132.7 \cdot 10^6$ muons, to be compared to $7.8 \cdot 10^6$ muons exploited by ATLAS). This is an additional cross check of the predicted uncertainties.

The fit does not have the nuisance parameters to describe the q_T^W uncertainty on the signal templates, because in the main analysis the q_T^W is fitted. Therefore the method described in this section cannot be used to estimate the constraint on q_T^W uncertainty for a traditional m_W measurement.

Conclusions

The large amount of events delivered by LHC and the excellent detector calibration achieved turned CMS from discovery to a precision experiment. In the framework of the W boson mass measurement, new methods can be designed to face the limiting systematic uncertainties and target a precision of 10 MeV.

In this work, a novel approach to measure the W boson production properties has been outlined. A template fit to unfold the single and double-differential unpolarized cross section σ^{U+L} and 5 angular coefficients $(A_0, \dots A_4)$ distributions, as a function of q_T^W and $|Y_W|$, has been developed. The quantities have been extracted from a fit to $\eta^\mu \times p_T^\mu$ templates built with the reconstructed muon variables only, from $W \to \mu\nu$ events. The fit relies on the sole assumption of a spin 1 W boson and an accurate description of the muon decay in the MC simulation, and thus a proper modelling of the FSR.

The method has been tested and validated on the CMS MC simulation. It demonstrates in particular the capability to measure the W boson transverse momentum spectra $d\sigma^{U+L}/dq_T^W$ with a granularity of 2 GeV, with 8% precision in below 2 GeV, 6% below 10 GeV, decreasing to 4% at 60 GeV.

A regularization scheme has been implemented to reduce the uncertainties. Preliminary results with a polynomial regularization show that the uncertainties on $d\sigma^{U+L}/dq_T^W$ can be reduced at 3% level in the entire spectrum.

The measurement is currently designed to be performed on the CMS data sample corresponding to the 2016 data taking period. The predictions are currently limited by the statistical uncertainty. The use of the entire Run 2 data sample can reduce the precision on $d\sigma^{U+L}/dq_T^W$ to 1-2%. Beyond the statistical precision, the main shape systematic uncertainties are the limited precision on the efficiency scale factors, the uncertainties related to the modelling of q_T^W spectrum in QCD background estimation and the residual dependence on the PDF. The scale factors are the leading source of systematic uncertainty and a more refined measurement can improve the overall result of the W production properties measurement.

The background measurement developed in this thesis is delivered with a 5% precision at low p_T^μ to 30% precision at high p_T^μ. The limiting systematic uncertainties arise from the p_T^{miss} calibration and the modelling of q_T^W spectrum in the $W \to \mu\nu$ MC sample.

In this work it has been also demonstrated, in a simplified fit with only 3 mass templates, how the same fitting framework has the capability to simultaneous fit m_W and the W production properties, reducing the systematic uncertainty related to production properties well below 10 MeV. With this method, a 10 MeV precision on m_W can probably be achieved already with the 2016 data sample only.

In the muon scale calibration, the Z boson lineshape is used as a standard candle for the muon momentum. A dedicated study has been performed to assess the impact of the PDFs on the Z lineshape. In this study, it has been demonstrated that the uncertainty on the distortion of

the lineshape due to the lack of knowledge on the PDF is below the MeV level. Therefore it is negligible at the target precision for W mass and production properties measurement.

As a final remark, this analysis has to manage a non-standard data and MC sample size and the production of order of 10^4 templates for signal, background and systematic uncertainties assessment. This leads to the the development of a dedicated and high-performance analysis and fitting framework, to meet the precision, flexibility and speed requirement of the analysis. State-of-the-art tools, like RDataFrame and TensorFlow have been used for this purpose.

Currently CMS is working to perform the measurement with the strategy outlined in this thesis, exploiting a larger size MC sample. The target is a simultaneous measurement of m_W and the W production properties. The new MC will allow the assessment of FSR uncertainty. More refined QCD estimation approach to avoid the residual q_T^W uncertainty are under development. An improved p_T^μ scale and resolution calibration, to target 10^{-4} relative precision, will be exploited. Finally W^+ and W^- channels will be combined and fitted simultaneously, reducing the impact of the shared nuisance parameters.

Appendix A

DeepCore: Convolutional Neural Network for high energy jet tracking

Tracking in high-density environments, such as the core of TeV jets, is particularly challenging both because combinatorics quickly diverge and because tracks may not leave anymore individual hits but rather large clusters of merged signals in the innermost tracking detectors. In CMS, this problem has been addressed in the past with cluster splitting algorithms, working layer by layer, followed by a pattern recognition step where a high number of candidate tracks are tested. Modern Deep Learning techniques can be used to better handle the problem by correlating information on multiple layers and directly providing proto-tracks without the need for an explicit cluster splitting algorithm. Preliminary results will be presented with ideas on how to further improve the algorithms.

This appendix reports the activity performed in the CMS tracking POG related to the development of a neural network, called DeepCore, to face the tracking in dense environment. The majority of the content is extracted from the Ref. [17], from the work presented at Connecting the Dots conference in 2019. In the final part of the appendix, a discussion about the current and future development is reported.

A.1 Motivation

The events with high-energy ($p_T^{\text{jet}} \gtrsim 0.5$ TeV) jets emission are part of a rich physics program at LHC, both for the new physics searches and the Standard Model (SM) physics. The boosted environment is exploited in the analysis which involves high-mass SM objects, like vector bosons, Higgs bosons or top quarks [202–204]. The track reconstruction inside the jets is a fundamental step for all the analyses which want to investigate the composition of the jets, looking for substructures and specific particle signatures. In the CMS experiment, the full reconstruction of the event relies on the particle flow algorithm [99], which combine the information of the subdetectors to assign to each reconstructed object a particle tag. The silicon tracker information is one of the blocks of the particle flow, and an improvement in tracking gives large benefits to the entire event reconstruction of CMS. More information about the particle flow algorithm and the tracking of CMS has been provided in Sec. 2.3 or can be found in Ref. [95].

The number of charged particle tracks and their spatial density inside the jets grow with the energy of the jet. The rule-of-thumb of jet physics is that the bulk of the particles are contained in a cone of $\Delta R = 2M/p_T^{\text{jet}}$, where M is the mass of the prompt particle and $\Delta R^2 = \Delta\phi^2 + \Delta\eta^2$. The presence of high-mass particle inside a jet (t, W, Z, H) produces characteristic substructures (sub-jets) inside a fat-jet. At low energy, the presence of high-mass particle inside a jet keeps high the ratio $E_{\text{massive objects}}/E_{\text{jet}}$, and the substructures inside the jet result isolated. At jet energy above several hundreds of GeV, the contribution of the massive particle becomes irrelevant and the available energy increases the fragmentation of the hadronization process.

In the framework of the CMS *Combinatorial Track Finder* (CFT), a dedicated iteration for high energy jet has been added for the LHC Run 2 because the tracking performance in the jet core (i.e. the central region) resulted lower than the average [15]. This iteration, called *jetCore* has been added as the last of the iterative tracking and searches seeds only in a cone of $\Delta R < 0.1$ around the jet axis if $E_T^{\text{jet}} > 100$ GeV. The jet axis is defined from calorimeter deposit only. The seeds are built with pairs of hits on the pixel detector and/or in the internal strip detector barrel, compatible with $p_T > 10$ GeV. In addition, the CKF tests a larger number of candidates in the jet core cone region (~ 50 against the standard 5). The tracks in the jet core, due to the high density, often leave on the pixel layers large merged cluster and not individual hits. A dedicated cluster splitter has been developed to face the merged clusters, using a modified *k-means* [205] based algorithm, which exploits the jet axis information to predict the cluster shape and charge for a single particle cluster or a multiple-particle merged cluster.

The performances of the jetCore iteration (called from now standard jetCore) are shown in Figure A.1, where the tracking efficiency as a function of ΔR between the track and the jet axis is shown in different scenarios. The jetCore iteration improves the efficiency, but the ideal splitting scenario reveals that there are still rooms of improvement.

However, the tracking efficiency still degrades in the jet core also with the ideal splitting, point out that the inefficiency is not due to the merged cluster only. Therefore has been decided

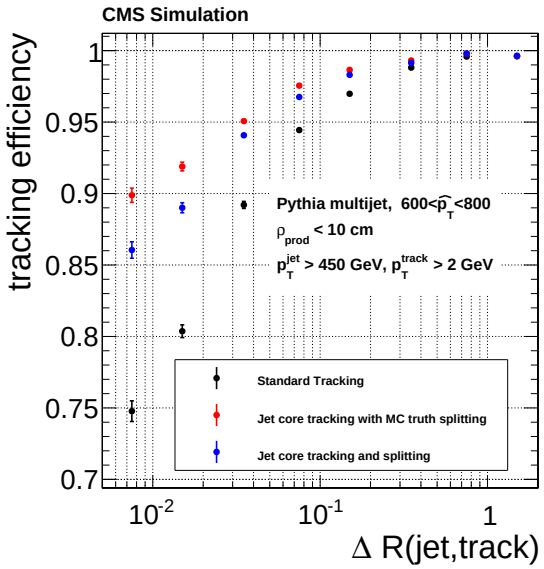


Figure A.1. *Tracking efficiency as a function of ΔR between the track and the jet axis in the scenario without jetCore (black), with the standard jetCore (blue) and with a jetCore iteration with ideal splitting of the merged cluster (red). Efficiency evaluated on a multijet sample generated with* `PYTHIA`*, with $600 < \hat{p}_T < 800$ GeV, $p_T^{jet} > 450$ GeV, where $\hat{p}_T$ is the transferred momentum. Only the tracks with $p_T^{track} > 2$ GeV with production vertex < 10 cm from beam line are included.*

to change the approach and develop a new version of the jetCore seeding algorithm avoiding an explicit splitting step and using the combined information of multiple pixel detector layers to produce a new list of jetCore seeds, instead of focusing to improve a layer-by-layer cluster splitting. An Artificial Neural Network, called *DeepCore*, has been developed, trained and tested in CMS reconstruction software to cope with this task. The description of DeepCore together with its performance is presented in the following sections.

Since DeepCore has been developed for the Run 3 reconstruction the Phase 1 upgraded pixel detector simulation has been used, as described in Sec. 2.2.2.1. Additional information can be found in Ref. [72].

A.2 Description of DeepCore network

In this section, the general strategy on which the novel high energy jet seeding algorithm is based is presented. First, the the details of DeepCore are described. Then, the integration of the NN in the CMS CTF is outlined. DeepCore is currently developed for barrel region only, therefore the pixel detector is simply made of 4 cylindrical layers in this framework. The entire discussion is restricted to $|\eta| \lesssim 1.4$, if not specified otherwise.

A.2.1 The strategy

The purpose of the seeding algorithm is to produce a list of track-seed i.e. sets of proto-track parameters for the interested tracking region. The primary goal of the algorithm is to find additional seeds in the jet core region, recovering seeding efficiency and lowering the rate of the fake track. This result can be reached by producing better quality seeds in terms of track parameters. The secondary goal is to lower the time consumption of jetCore iteration, currently one of the most expensive of the CTF (mainly because of the large number of explored candidates to recover efficiency).

The seeding algorithm directly produces the list of seeds (i.e. track parameters) from the raw pixel detector information, without any clustering algorithm on the top. A good candidate to reproduce the function:

$$f : \{\text{raw pixel information}\} \longrightarrow \{\text{list of track seeds}\}$$

is an Artificial Neural Network (NN). With *raw pixel* information from now on it is referred to individual pixel charge and position, without any clustering algorithm, but the default charge calibration and zero-suppression algorithms applied. In particular, explicit clustering algorithm and the consequent cluster splitting step has been skipped since the previous cluster splitting algorithm resulted suboptimal.

In the wide field of NN a Convolutional Neural Network (CNN) has been used to face the problem. The CNNs [16] are one of the most natural choices with a 2D-picture input, like the pixel detector information. Each node of the input layer of the NN is a single pixel of the 2D pixel detector layer, considering a input window of $n \times n$ pixels. The window it is called *input feature map* (IFM), in the CNN language. Each node of the first hidden layer is connected to few nodes of the IFM, typically a matrix of $m \times m$ nodes, $m < n$. The value of the node is evaluated as the scalar product between the values of the $m \times m$ nodes of the IFM and the values of a trainable matrix of $m \times m$ weights, called *filter*. The resulting 2D matrix of the values of the nodes of the hidden layer is called *output feature map* (OFM). The same filter is applied to the various $m \times m$ regions of the IFM to obtain the full OFM. It can be seen as the

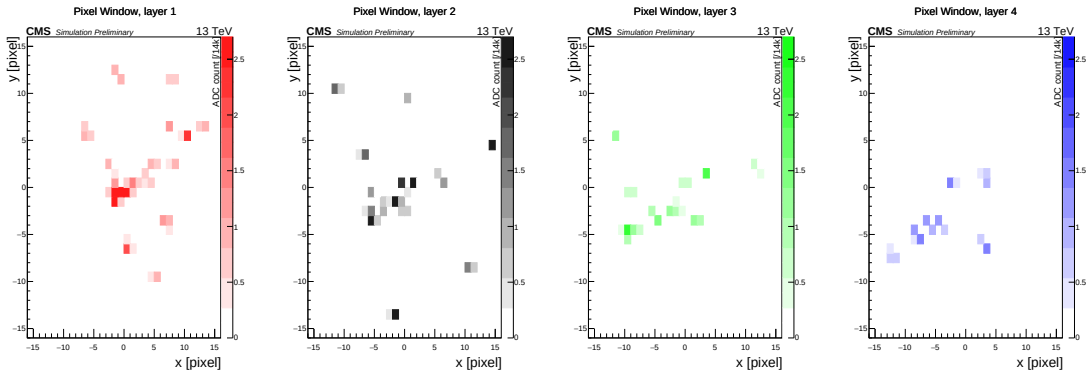
filter which is swept along the IFM, identifying specific features and encoding their presence in the OFM. This reflects the possibility to have the same feature in different position of the IFM. The CNN uses multiple filters to find different features, analyzing multiple times the IFM, and producing an OFM for each filter. Additional convolutional hidden layer are added after the first one, interpreting the OFM of the first hidden layer as the IFM of the next convolutional layer. The relevant parameters are the number of filters (how many kinds of features are expected), the dimension of the filters (how many pixels are needed to identify a feature) and the number of convolution layer (the complexity of the features). The number of filters and their dimension change at each convolutional layer, according to the target features at each step.

In the tracking environment, the pixel detector layers can be interpreted as *RGB channels* of the same 2D picture (i.e. as an additional dimension). The inputs are fixed-size windows of pixels (the jet core regions). The features inside the filters are the track patterns on the 4 layers. Thus, the dimension of the filters must be large enough to include the track hits on the 4 layers. The network is realized with convolutional layers only. A 2D-picture output allows to be completely independent on the number of tracks in the layer but only to the mean occupancy. In addition, the network can be rescaled for different window size or different tracker geometry without changing the architecture but few hyper-parameters only. Another relevant feature of the convolutional approach is that all the seeds are predicted at the same time, and not removing the correspondent hits in a sequential way. This approach has been previously used in Ref. [206] to identify a variable number of targets in videos with real-time detection.

A.2.2 DeepCore Neural Network

Training Input. The input of the network are 4 pixel *maps* centred on the merged clusters. A cluster is flagged as *merged* if its charge and shape are compatible with multiple particles. The procedure to build them is: for each jet with $p_T > 1$ TeV the interception between the jet axis from the calorimeter information and the first layer of pixel detector is found, then it is opened a cone of $\Delta R = 0.1$ and all the merged clusters inside the cone on the layer 1 are found. If the crossed pixel detector module is inactive the list of the merged cluster on the next layer, layer 2, is used. Then, for each merged cluster a 30×30 pixels window is opened in each of the 4 layers, using as a center the interception between the layer and the direction defined by the primary vertex (PV) and the merged cluster. The jet axis is also added as an additional direction to open the 4 windows. For each of the direction, for each window, the x, y and charge information of the hits inside the windows is stored. The use of charge and shape to identify the merged clusters is used for the training input only and does not bias the CNN with respect to an MC-truth merged cluster because of the large overlap between windows.
The charge information is normalized to a fixed value (14000, the mean value of the charge deposition in a pixel), to obtain input values of order 1, easier to handle for a NN. Each training input is made of the 4 windows, called pixel maps, thus for each jet multiple overlapping inputs are produced. In addition also the jet η and jet p_T are added to the input, as a unique information for the input using the calorimeter-jet information, because the shape of the cluster depends on the energy and the crossing angles of the particles. In Figure A.2 an example of the 4 pixel maps input is shown.

Training Target. For each input, the target of the network is made of three copies of a *Track Crossing Points (TCP) Map* and a *Track-Parameters Map*. Each copy of the two Maps is a pair of 30×30 matrices. For each pixel of layer 2 input map, if a track crosses that pixel, 1 will be stored in the correspondent pixel (1-pixel) of the first TCP Map, 0 will be stored in the pixel otherwise (0-pixel).
For each 1-pixel of the TCP Map, the track parameters of the track are stored in the correspon-



(a) (b) (c) (d)

Figure A.2. *Example of of the pixel maps used as input for the DeepCore neural network: the maps shows a windows on the 4 pixel detector layer of CMS, aligned to the jet direction. The ADC counts are divided by 14000.*

dent pixel of the Track-Parameters Map. The track parameters are stored in local coordinate: Δx and Δy with respect to the center of the pixel, $\Delta\eta$ and $\Delta\phi$ with respect to the merged cluster-PV direction and the p_T of the track. The track parameters are also stored for the pixels in a radius of 2 pixels with respect to each TCP, with the local pixel reference (these pixels has been called Near to track Crossing Points, NCP). The rest of the pixels of the Track-Parameters Map are filled with 0. The second and third copies (called Overlap 2 and 3 Maps) are filled to take into account multiple tracks which cross the same pixel: if another track crosses a TCP another 1 will be is stored in the TCP Map-Overlap 2 with the relative filling for the Track-Parameters Map-Overlap 2. The same for the Overlap 3 Maps in case of three tracks in the same pixel.
In Figure A.3(a) this complex target is shown graphically.

Architecture. The architecture of the network is completely convolutional. It is schematically shown in Figure A.3(b). The inputs feed five 2D convolutional layers with reducing filter size and number, then the network is split in two trunks: four 2D convolutional layers to produce the Track-Parameters Maps and four 2D convolutional layers for the TCP Maps. The total number of parameters of the network is 77373. The activation functions are ReLU for all the layer but the last TCP Maps layer, where Sigmoid is used (the Sigmoid is recommended for limited range output and binary losses, see training details later on). They are defined as follow:

$$\text{ReLU} = \max(x, 0), \quad \text{Sigmoid}(x) = \frac{1}{1 + e^{-x}}.$$

Prediction. The Prediction of the network has the same structure of the Target i.e. three 30×30 TCP Maps and three Track-Parameters Maps. The TCP Maps will contain values between 0 and 1 for each pixel thus can be interpreted as a probability that a track cross that pixel. The Track-Parameters Maps contains instead the 5 parameters for the TCP and NCP pixels in local coordinate.

Training details. The NN has been trained with a large sample of inputs, for which also the relative target information is given. During the training the network must predict the target given the input only, then it must compare the prediction with the true target. The comparison proceeds with a given metric i.e. the *loss function*, which defines the grade of accuracy of the prediction. Two losses, one for each target, has been used to train DeepCore. A weighted *Binary*

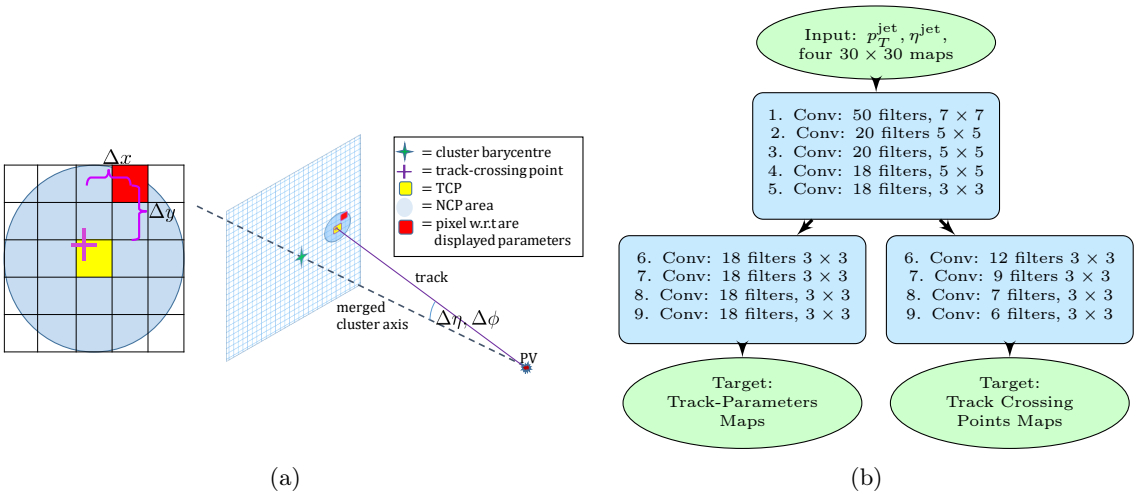
Cross Entropy has been used for the TCP Maps i.e

$$\mathcal{L}_{TCP} = -\frac{1}{N}\sum_{i=1}^{N} w_i \Big[10 \cdot y_i^{\text{true}} \ln(y_i^{\text{pred}}) + (1 - y_i^{\text{true}}) \ln(1 - y_i^{\text{pred}})\Big],$$

where the sum run over the TCP (i.e. $30 \times 30 \times 3_{\text{overlap}}$ "pixel") and the w_i is a weight of the given pixel, while $y_i^{\text{true,pred}}$ are the target and predicted value of the i-pixel, respectively. The TCP and NCP pixel weight is set to 1, while for the other pixels is set to 0.01. In addition the cross entropy is biased towards positive values, with the factor 10 for the positive TPC (this is almost equivalent to provide weight 10 to TCP, weight 1 to NCP and weight 0.01 to the rest). Both weights and bias are needed to avoid a vanishing TCP Map prediction because of the sparse target, reducing the false negatives.
A clipped mean square error has been used for all the parameters:

$$\mathcal{L}_{\text{par}} = \frac{\sum_{p \in TCP,NCP} \min[(p^{\text{pred}} - p^{\text{true}})^2, 25]}{N_{TCP+NCP}},$$

where the sum runs over the 5 parameters, on the TCP and NCP pixels only. The $p^{\text{true,pred}}$ are the target and predicted value of the parameter, respectively. The clipping is needed to avoid large tails in the prediction which enlarge artificially the loss.
The training sample is composed of 22 million input (about 2 million jets) plus two million used for validation and it is composed of multijet events with the transfer $\hat{p}_T$ between 1.8 and 2.4 TeV. The jets are required to have $p_T^{\text{jet}} > 1$ TeV and $|\eta^{\text{jet}}| < 1.4$, while only the tracks with $p_T > 1$ GeV have been used to build the targets. The batch size (the number of input analysed for each prediction) is 32, which is the largest possible given the available computation power. The chosen optimizer is `Adam` [207], the learning rate has been changed during the 246 epochs of training, gradually from $2 \cdot 10^{-4}$ to 10^{-7}, and in each epoch the entire training sample is explored.



(a) (b)

Figure A.3. *On the left (a) an example of the Target for a single track: on the left the TCP (in yellow), the track parameters are stored for all the pixels inside the shaded blue area, the red pixel is the one with respect of which are evaluated the parameters. The Overlap Maps are not shown. On the right (b) the architecture of DeepCore.*

A.2.3 Integration of DeepCore in CMS reconstruction

The training of DeepCore has been performed outside of `CMSSW` on GPU and then the final weights have been permanently stored and given to `CMSSW`. DeepCore has been developed with `Keras` library [208] both for the training and the prediction inside `CMSSW`. DeepCore has been integrated into the jetCore iteration of CMS reconstruction: standard jetCore seeding has been disabled and the following algorithm is the replacement.

The cluster list in a cone of $\Delta R = 0.1$ with respect to the jet axis is identified for each calorimeter jet with $p_T > 300$ GeV. Each cluster defines a new direction on which a DeepCore Input is built (the 4 pixel maps and the $p_T^{\text{jet}}, |\eta^{\text{jet}}|$). The input is defined for all the cluster, and not only for the merged ones, to recover as much efficiency as possible at the seeding level. The standard duplicate remover will take care to remove overlapped tracks in the following steps of reconstruction. The input is given to DeepCore NN which returns the prediction given the weights of the training. The list of actual seeds is made from DeepCore prediction with the sets of 5 track parameters of the *most probable* pixels. *Most probable* is defined as TCP output greater than 0.85, 0.75, 065 for the three Overlaps or greater than 0.5, 0.4, 0.3 in case the layer 2 is missing for the given input (because of inactive modules). The threshold is lowered in the latter case because the target of TCP is built on layer 2, thus it is crucial in the prediction. In addition to the standard duplicate remover, the list of seeds is cleaned from duplicates: if two seeds have $\Delta x, \Delta y < 50$ μm, $\Delta\eta, \Delta\phi < 0.002$ the one from the lower value of TCP is removed from the list. The uncertainty on the parameters is fixed for all the seeds: $\sigma_{1/p_T} = 0.15$ GeV^{-1}, $\sigma_\eta = \sigma_\phi = 0.01$, $\sigma_{xy} = \sigma_z = 44$ μm, without off-diagonal terms, based on the performance of the prediction of DeepCore (see next section).

A.3 Preliminary performances of DeepCore

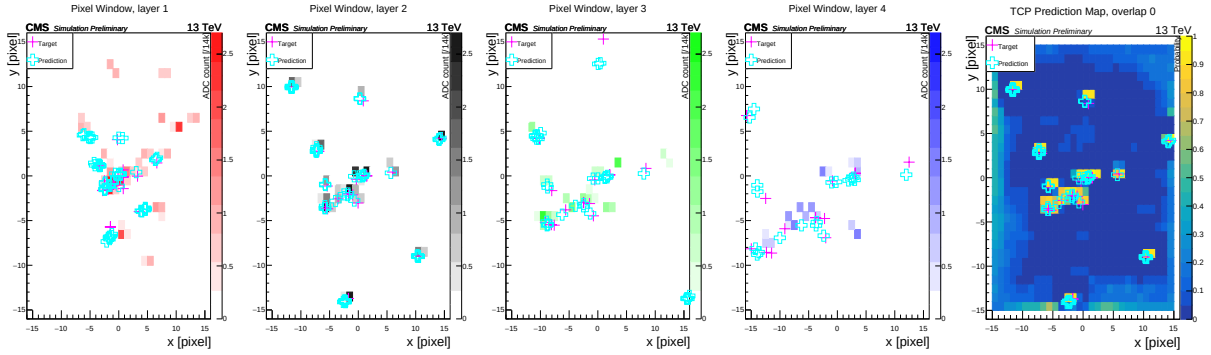
The behaviour of DeepCore can be checked during the training with an "event display", developed for optimization studies externally from `CMSSW`. The same event of Figure A.2 is shown in Figure A.4, together with TCP Map, the target and the track-parameter prediction of the *most probable* hits only, at the end of the training. The TCP map, the target and track-parameter prediction are directly provided by the NN on the layer 2 and are propagated using a linear extrapolation to the layers 1, 3 and 4 (thus neglecting the p_T target and prediction). The event display has only a qualitative interpretation, but it reveals an almost full efficiency and an accuracy of 1-2 pixels also with the used linear propagation, with an affordable level of duplication. The duplicate remover is not run in the event display. The Figure A.5 shows an example of the quantitative validation of the training performance, in term of residual of η parameter between the prediction and the target. The null average bias, the 1.4% spread and the strong correlation with the target show that DeepCore is able to predict the parameters given the pixel input.

DeepCore has been validated integrated in the CMS reconstruction on $2 \cdot 10^4$ multijet events with the trasfer $\hat{p}_T$ between 1.8 and 2.4 TeV. The jets are required to have $p_T^{\text{jet}} > 1$ TeV and $|\eta^{\text{jet}}| < 1.4$ and on the simulated tracks has been applied the typical CMS selection $|\eta| < 2.5$, $r_{\text{prod}} < 3$ cm, $|z_{\text{prod}}| < 30$ cm, $p_T > 0.9$ GeV.

The results for the final tracking performances are shown in Figure A.6 in a stacked plot with highlighted the contribution of the various iterations of the CTF in the jet core region. The tracking efficiency is defined as $\varepsilon = N_{\text{assoc}}/N_{\text{sim}}$, where N_{sim} is the number of simulated tracks and N_{assoc} is the number of simulated tracks associated to a reconstructed one. The fake rate is defined as $R_F = N_{\text{not assoc}}/N_{\text{reco}}$, where N_{reco} is the number of reconstructed tracks and $N_{\text{not assoc}}$ is the number reconstructed tracks not associated to a simulated one. A reconstructed track is flagged as "associated" if the χ^2 between its parameters and the simulated is lower than 25.

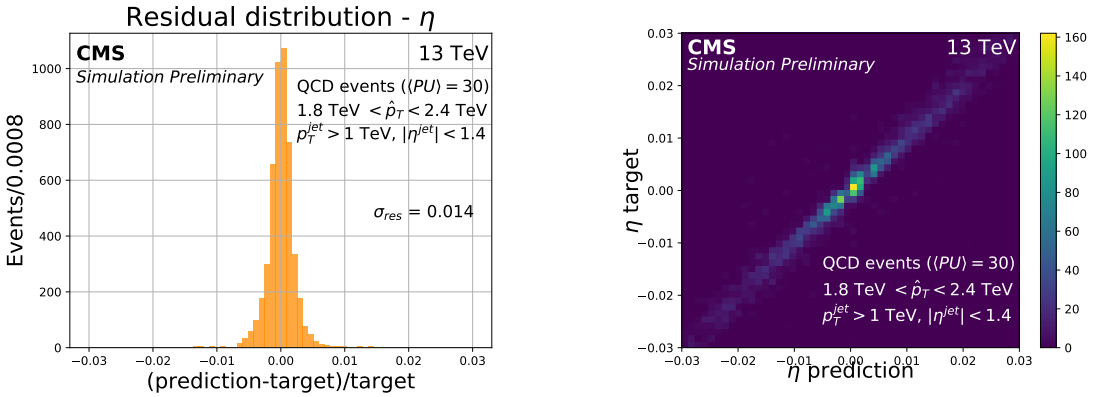
This definition replaces the usual CMS one (based on the fraction of true hits used) for these validation studies, because DeepCore seeding is without pixel hits and with the usual association it will be negatively biased.

The improvement given by DeepCore to CMS reconstruction is better shown in Figure A.7, where the performance with the standard jetCore algorithm and the one with DeepCore are compared. Also the tracking performances obtained producing the seed for the jetCore iteration using the simulated track information is shown (*MC truth seeding*), for which the seeding efficiency is 100% and the fake rate 0% by definition. DeepCore is able to reproduce the perfect seeding efficiency with degradation below 1%, flat in ΔR. On the other hand, all the fake tracks produced by the standard jetCore are avoided, reducing the seeding fake rate below 5%. In particular, the good purity of DeepCore seeds lowers the fakes below the rate without the jetCore iteration because DeepCore is able to correctly reconstruct tracks reconstructed as fakes by different iteration in the low ΔR region.



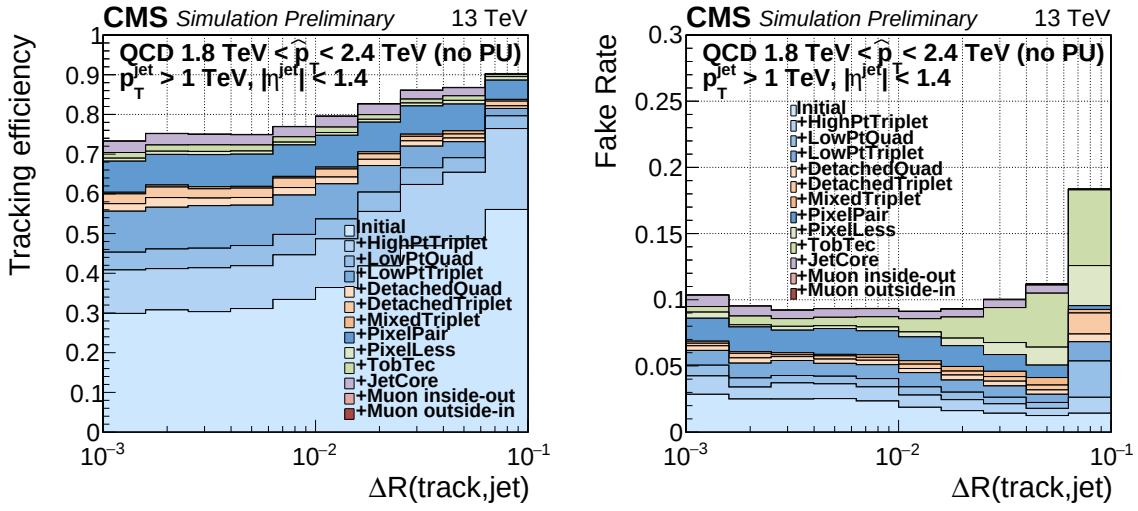
(a) (b) (c) (d) (e)

Figure A.4. *Example of the pixel maps used as input. On the top are also shown the crosses of the crossing point of the target (simulated) tracks and the correspondent prediction of DeepCore for the most probable hits. The prediction is produced on layer 2 and propagated linearly on the other layers. The most right figure is the map of the predicted crossing point on the window of layer 2, expressed as probability, with the crosses of the predictions and the targets. The linear propagation is used in the event display only, in seed production the predicted p_T is used.*



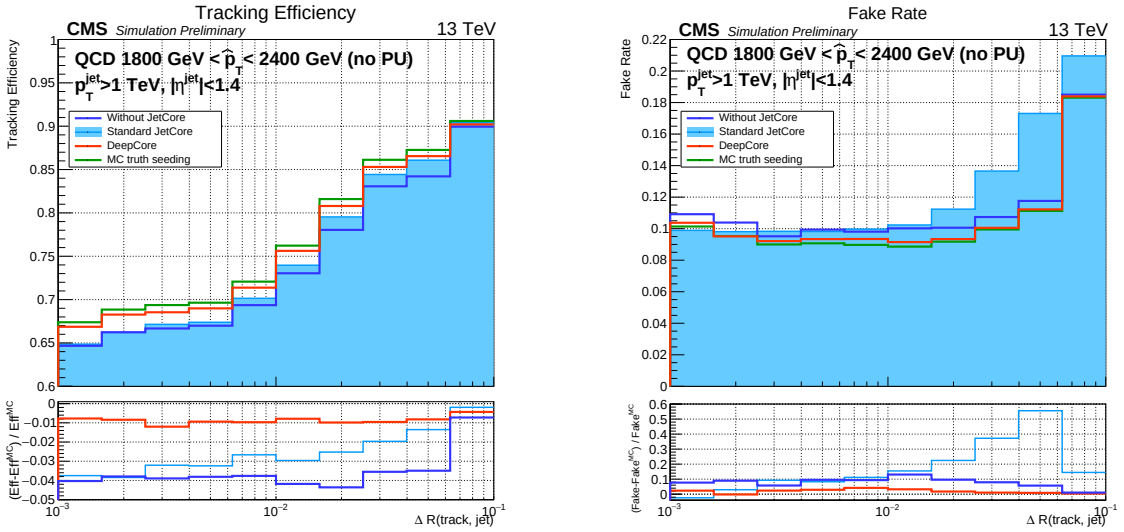
(a) (b)

Figure A.5. *On the left (a) the residual between the seed η parameter predicted by DeepCore and the target (simulated) track η parameter. On the right (b) the correlation between prediction of DeepCore and target parameters shown with seed η parameter predicted against the simulated track η parameter.*



(a) (b)

Figure A.6. *Tracking efficiency (left figure) and fake rate (right figure) in the jet core region ($\Delta R < 0.1$, between the reconstructed jet axis and the simulated track direction). The contribution of the different iterations of the CKF are shown as stacked histograms. The DeepCore algorithm is used in the iteration dedicated to the cores of the jets [jetCore (purple)]. In the efficiency the shared reconstructed tracks (duplicated) between various iterations are not removed.*



(a) (b)

Figure A.7. *Tracking efficiency (left figure) and fake rate (right figure) in the jet core region ($\Delta R < 0.1$, between the reconstructed jet axis and the simulated track direction). The light blue filled histogram is obtained with the standard CMS tracking algorithm. The dark blue histogram is obtained removing the CKF iteration dedicated to the jet cores. The red histogram is obtained using the DeepCore in the seeding for the iteration dedicated to the jet cores. The green histogram is obtained producing the seed for the jetCore iteration using the MC truth seeding. In the lower pads are shown the differences between various tracking efficiencies (fake rates) and the MC truth seeding one, divided by the MC truth seeding efficiency (fake rate).*

A.3.1 Summary of the 2019 status

The CNNs have been shown to be a valid approach to perform seeding for track reconstruction in a dense environment. The DeepCore algorithm, developed and validated with the CMS tracker in the central region, shows better performance than the standard seeding algorithm in such a dense environment: it almost cancels the seeding inefficiencies, reduces the fake rate up to 60% and also the seeding time is reduced.

The described performance has been estimated on `CMSSW` in the software release `10_5_0`.

A.4 Current status and future development

In 2020 DeepCore, has been fully integrated in `CMSSW` with the purpose of enabling it for the LHC Run 3 restart.

At the time of the integration the performance of DeepCore has been confirmed, in term of efficiency increase and fake rate reduction. Moreover, a detailed timing consumption study has been performed. Overall the timing consumption of the jetCore iteration will be reduced by a factor 0.7, if DeepCore is enabled. In particular, the seeding time is strongly increased, but the track building step timing is reduced by a factor 2. This is due to the improved quality of the DeepCore seeds and consequent lower fake candidates. The full breakdown of the timing analysis is reported in Fig. A.8(a). The described performance has been estimated on `CMSSW` in the software release `11_3_0`, on the CMS Run 3 2021 or 2024 data taking conditions

However the potential of DeepCore is not currently fully exploited, and several upgrades have been studied and planned for the next future. They are briefly outlined in the next sections. All the described result are very preliminary and can change in the future.

A.4.1 Updated training

The validation of DeepCore performance done in 2020 has been performed without repeating the training of the neural network. It means that a training performed on the 2017 condition has been used to do prediction in 2021-2024 conditions. Suboptimal results are expected in this case, since DeepCore has the capability to learn the details of the detector, like alignment or local inefficiency, which are spoiled if it is used in different detector conditions.

The re-training gives also the opportunity to improve the tuning of the hyperparameter of the NN, like the learning rate, batch size, and the details of the architecture. Specific failure scenarios, like a completely deactivated layer impact, can be also studied in this framework.

The plan is therefore to repeat the training of DeepCore in CMS Run 3 conditions. Is under discussion the advantage of repeat the training of DeepCore regularly (once a year) to maintain optimal reconstruction performance.

A.4.2 Improved Rebuilding of the Seeding Region

The CTF performs a back-propagation from the outer layers toward the inner pixel layer, after the first combinatorial Kalman Filter outward propagation, as discussed in Sec 2.3.1. This backward propagation repeats the CKF propagation using the entire set of hits associated with the track candidate as a starting point. In the default configuration, the *Rebuilding of the Seeding Region* (RSR) is performed as a preliminary step to the back-propagation. This step links the hits found in the seeding layers to the hit collection of the track candidate. In this way,

the back-propagation can propagate inwards through the seeding layers. The back-propagation improves the track parameters precision since the information contained in the seeding layers will be exploited with full precision, thanks to the accuracy of the track candidate parameters. While the first outward propagation the seed precision is limited, due to time and computation constraint. Moreover, additional hits can be found in layers closer to the IP than the seeding layers (in the case of strip-seeded iteration). If the RSR is disabled the back-propagation will stop at the last valid hit, before the outermost seeding layers. This spoils most of the advantage of the back-propagation.

The current implementation of DeepCore uses the default configuration of the back-propagation. However, this choice is suboptimal, since in the RSR the usual cluster splitter is used to find the hits in the pixel detector layers. The cluster splitter performances are relatively low in the jet core regions, and thus the RSR results inefficient. With this procedure, the entire advantage of the DeepCore seeds is spoiled for the purpose of the back-propagation.

The DeepCore seeds has a seed parameter resolution 5-10 times better than the standard seed of jetCore iteration. Has been considered for instance the transverse impact parameter d_{xy} on the high-p_T jet sample used for the previous validation. The standard jetCore provides $\sigma(d_{xy})^{\text{seed}}_{\text{standard}} \simeq 1300$ μm, while $\sigma(d_{xy})^{\text{seed}}_{\text{DeepCore}} \simeq 200$ μm. The residual distribution is shown in Fig. A.8(b).

This improvement in the seed parameters completely vanish at the track parameter level. The same comparison performed on on the track d_{xy} provides $\sigma(d_{xy})^{\text{seed}}_{\text{standard}} \simeq \sigma(d_{xy})^{\text{seed}}_{\text{DeepCore}} \simeq 30$ μm.

The seed information is actually used only to build the track for the outward propagation. The track parameters precision will be ruled by the track precision, given the lever arm of the entire tracker. However, the effect of the RSR is not negligible and if disabled the $\sigma(d_{xy})$ increase of factor 2. Therefore the key point is to properly exploit the information of the seeding layers in RSR step.

The planned solution is to link to the track candidate the seed parameters of DeepCore seeds converted in a 5D hit format. This 5D hit can be interpreted by the CKF as an additional hit placed on the layer 2 of the pixel detector, with the full precision of DeepCore seed. However, this solution has not been implemented yet and its quantitative impact on the track parameters cannot be predicted.

A.4.3 Impact of DeepCore on b-tagging performances

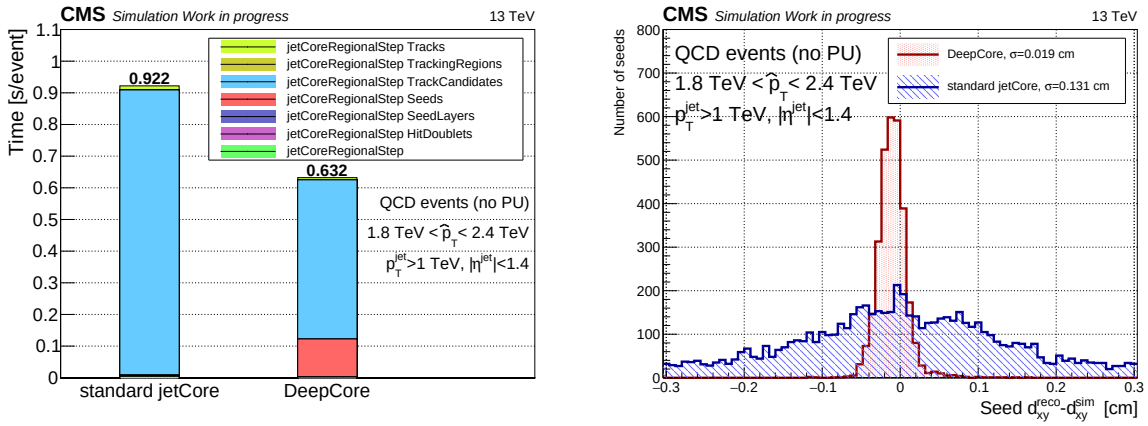
If a track parameter improvement is expected this will reflects also in higher level performance. In particular, an improvement of the impact parameter of the tracks will improve the b-tagging performance of CMS algorithms, which strongly relies on the impact parameter of the tracks.

As outlined in the previous section, at the current stage DeepCore does not improve the impact parameter, thus no improvement in b-tagging performance has been observed. Therefore also for this purpose, the dedicated RSR reconstruction must be implemented.

A.4.4 Endcap extension

Currently, DeepCore has been implemented in the barrel region only. The architecture of the neural network and the prediction in `CMSSW` is based on 4 pixel layers, parallels to the beam axis. If a track with $|\eta| > 1.4$ is passed to DeepCore the input will be partially empty and the prediction will be inefficient or unsuccessful.

However, the standard jetCore algorithm works up to $|\eta| = 2.5$ i.e. up to the geometrical limit



(a) (b)

Figure A.8. *Timing (a) and seed residual (b) comparisons between standard jetCore and DeepCore reconstruction, validated on* $2 \cdot 10^4$ *multijet events, with* $1.8\ TeV < \hat{p}_T < 2.4\ TeV$, $|\eta^{jet}| < 1.4$, $p_T^{jet} > 1\ TeV$, $|\eta^{track}| < 2.5$, $r_{prod}^{track} < 3\ cm$, $|z_{prod}^{track}| < 30\ cm$, $p_T^{track} > 0.9\ GeV$. *In the timing comparison only the relevant modules are shown*

of the pixel detector. The impact on the total tracking efficiency of the standard jetCore in the endcap regions is quite small also in high energy jets sample. The standard jetCore iteration is able to increase the tracking efficiency of about 2% in the endcap region, to be compared to the 15% increase in the barrel region, using the sample adopted for the previous validations.

This lower impact is expected since the layer distance from the IP is larger in the endcap region compared to the barrel. Therefore the jet tracks span larger regions on the pixel detector modules, at a given aperture angle. Moreover, in the endcap region, the p_T^{jet} cannot be used to select and analyse the jet of interest, since the longitudinal momentum of the jet contributes significantly to the aperture and thus the presence of high track density. The jet energy E_{jet} must be used instead. However, the p_T^{jet} must be higher enough to cross the pixel detector disk. In conclusion, the available phase space is limited by the energy constraint of LHC and thus only a small number of jets are expected to have dense cores in the pixel detector endcaps.

Two approaches have been developed to extend DeepCore in the endcap region. Both have different advantages and disadvantages and the final decision on which will be integrated has not taken yet:

Single architecture. The first approach consists of the addition of 3 more layers to the input of the NN, correspondent to the 3 disks of the endcap pixel. The target is still built on the layer 2. In this configuration 7 pixel map are passed to DeepCore as input, independently from the η of the jet. In this configuration, if a barrel jet is passed to the NN the maps 1, 2, 3, 4 will be filled, while the 5, 6, 7 will be empty. On the other hand in case of an endcap jet the map 1, 2, 5, 6, 7 will be filled, while the map 3, 4 will be empty. The NN should learn the geometry of the tracker and thus the position in the real space and thus in the track parameter space automatically, given the presence or the absence of pixels on certain input layers. This is a challenging task since the convolutional "projective" hypothesis below DeepCore is partially spoiled by the geometrical rotation of the endcap disks compared to the barrel layers. However, the p_T^{jet} and η^{jet} from the calorimeter information, are provided to DeepCore as well to help the NN in this task. A strong limitation of the described approach is the use of the layer 2 for the target. This choice limits the acceptance to $|\eta| < 2.1$, and leaves uncovered the region $2.1 < |\eta| < 2.5$. Since the target

and thus the seed must be built on a physical detector surface such a choice cannot be avoided, if a single architecture will be used for the barrel and the endcap. A sketch of this approach is shown in Fig. A.9

Different architectures. The second approach abandons the latter assumption. Two different trainings with different neural networks architecture are used for the barrel and the endcap. The DeepCore-barrel is the already discussed one. The DeepCore endcap has a 2-layers plus 3-disks input but the target is built on the first disks, allowing to extend the acceptance to $1.4 < |\eta| < 2.5$. Note that with this solution the difficulties of the geometrical rotation of the provided input hold, and the advantages compared to a 3-disks input only has not been studied yet. The described approach requires two separate training, with differently prepared training samples, for the two regions. Then a switcher must be inserted in the `CMSSW` integration base on the $|\eta|$ of the jet. The possibility to develop three versions of DeepCore, purely barrel, endcap and the overlap region between the two, can be also investigated.

Currently, only the first option, with a single architecture for the entire phase space, has been partially tested. Additional difficulties arise from the training sample identification. Since the interesting events are rare (high energy (TeV) jets, enough large p_T, large η) a dedicated sample has been produced. This samples contains jets of mixed flavour with a flat spectrum in energy up to the LHC kinematic limit, 1 TeV $< E_{\text{jet}} <$ 7 TeV. However, the jet core aperture angle is larger than in the barrel case and the 30×30 pixel windows of the input are not sufficient to properly contain the jet core in the endcap disks. Enlarging the windows sizes would increase the computation burden of the NN. These issues suggest that the different architectures option can

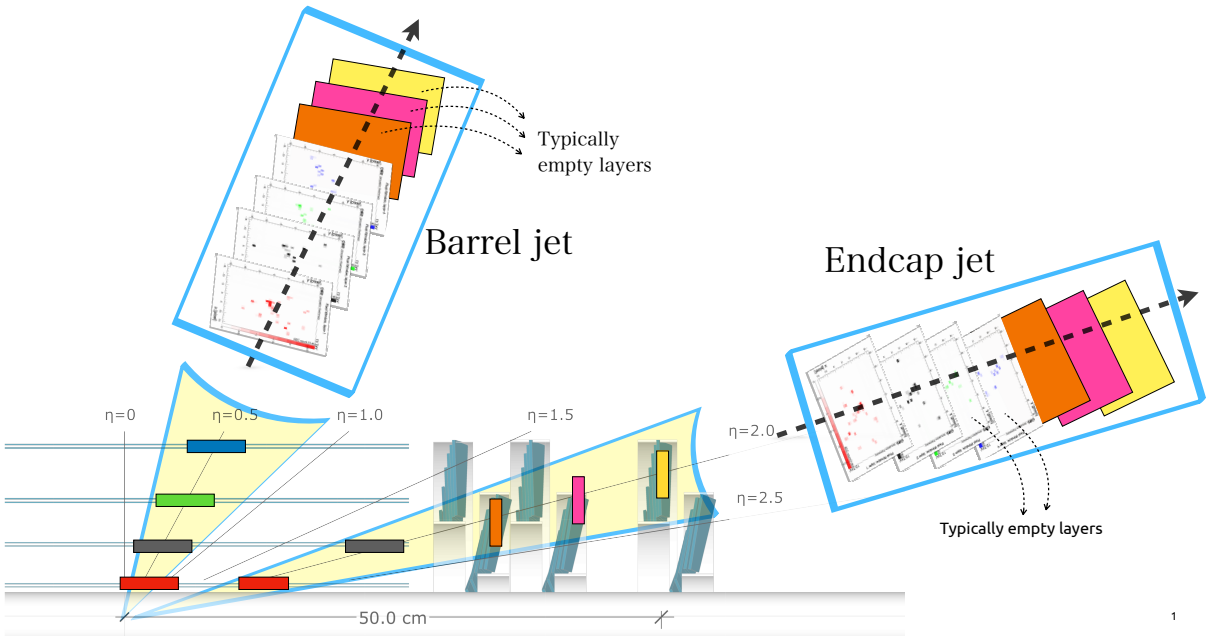


Figure A.9. *Sketch of the single architecture approach. Two jets are shown, one in the barrel and one in the endcap region, superimposed to the pixel detector diagram. The solid boxes on the pixel detector layers represents the DeepCore windows, also reported in the two cases in the large light-blue boxes, with a consistent color scheme. The windows 5,6,7 are empty in the barrel jet, while the windows 3,4 are empty in the endcap jet. The Endcap jet is close to the geometrical acceptance, ruled by the layer 2, where the target is built. The windows are not in scale with the detector diagram.*

be the optimal solution, allowing to have different window size in the barrel and in the endcap, increasing the computation time only in a limited number of events, only where is strongly needed.

Currently, the DeepCore-endcap is not fully functional, and only the DeepCore-barrel has been integrated in `CMSSW`. Therefore the worst-case scenario is to adopt a hybrid solution, enabling DeepCore in the barrel region and the standard jetCore algorithm in the endcap region.

A.4.5 Strip extension

The good performance of the DeepCore algorithm suggests studying the impact of applying such an approach to pattern recognition as well. DeepCore recovers almost completely the seeding inefficiency. However, the tracking efficiency is still 70%-90%. The residual inefficiency must be addressed to the subsequent step of track reconstruction. A possible idea is to extend the DeepCore approach to the strip tracker propagation to improve the tracking efficiency in the jet core regions.

No studies have been performed yet on a realistic implementation of this approach. The possibility that different approaches can be more feasible or effective cannot be excluded at this level and should be considered before starting to develop this approach.

Appendix B

Additional material on the impact of PDFs on bosons lineshape

Ref. [6] as well as Chapter 3 contain the main results of the study on the impact of the PDFs on the vector bosons lineshapes. In this appendix, the additional material is reported, in term of explicit calculations and plots supporting the already discussed results.

B.1 Beyond the muon scale calibration

The study discussed in Chapter 3 is needed not only to ensure the consistency of the use of the Z lineshape as standard candle for the muon scale calibration, but also for future, more precise measurement of Z boson mass. For instance, by analyzing the dileptonic Z decays of Run 3 of the LHC a statistical-only precision of about 10^{-5} GeV on M_Z might be achievable. This level of precision would demand a control of the dilepton mass lineshape at the sub-MeV level.

B.2 Extended calculation of the PDFs impact

First, the transformation from Eq. 3.1, $\frac{d^2\sigma_V}{dx_1 dx_2}$, to Eq. 3.3, $\frac{d\sigma_V}{dQ}(Q|y)$, can be made explicit with the change of variables:

$$\begin{cases} Y = \frac{1}{2}\ln\frac{x_1}{x_2} \\ Q^2 = s x_1 x_2 \end{cases} \Rightarrow \begin{cases} x_1 = \frac{Q^2}{\sqrt{s}} e^{Y} \\ x_2 = \frac{Q^2}{\sqrt{s}} e^{-Y} \end{cases}$$

thus:

$$\frac{d^2\sigma}{dQ^2 dY} = \frac{4}{s}\frac{d\sigma}{dx_1 d_2}, \text{ where } \frac{4}{s} = \det\begin{pmatrix} \frac{dx_1}{dY} & \frac{dx_1}{dQ^2} \\ \frac{dx_2}{dY} & \frac{dx_2}{dQ^2} \end{pmatrix} \text{ is the Jacobian of the transformation.}$$

Then the additional change from $Q^2 \to Q$, expressing the cross section conditional in $Y = y$:

$$\frac{d^2\sigma(Q,Y)}{dQ^2 dY}(Q|Y) = \frac{1}{2Q}\left(\frac{d\sigma}{dY}\right)_{Y=y}\frac{d\sigma}{dQ}(Q|Y).$$

Second, the Taylor expansion to derive the Eq. 3.4 can be clarified. Let us consider for the moment a single parton flavour pair i, j. The function $F^{i,j}(Q)$, defined in Eq. 3.5, can be

expanded around $Q = M_V$:

$$F^{i,j}(Q) \simeq F^{i,j}(M_V) + (Q - M_V)\frac{dF^{i,j}(Q)}{dQ}\Big|_{Q=M_V} + (Q - M_V)^2 \frac{1}{2}\frac{d^2F^{i,j}(Q)}{dQ^2}\Big|_{Q=M_V}, \tag{B.1}$$

where:

$$\frac{dF^{i,j}(Q)}{dQ}\Big|_{Q=M_V} = \frac{f_1(\bar{x}_1)f_2(\bar{x}_2)}{Q}\Big|_{Q=M_V} H^{i,j},$$
$$\frac{1}{2}\frac{d^2F^{i,j}(Q)}{dQ^2}\Big|_{Q=M_V} = \frac{f_1(\bar{x}_1)f_2(\bar{x}_2)}{Q^2}\Big|_{Q=M_V} K^{i,j}.$$

Now, the Eq. B.1 can be substituted in Eq. 3.3. Then, the Eq. 3.4 is directly obtained using the $F^{i,j}$, $H^{i,j}$ and $K^{i,j}$ and properly taking into account the sum over all the parton flavours and the CKM matrix element.

Finally, is worth to explicit the shift parameter H_V as a function of the PDFs:

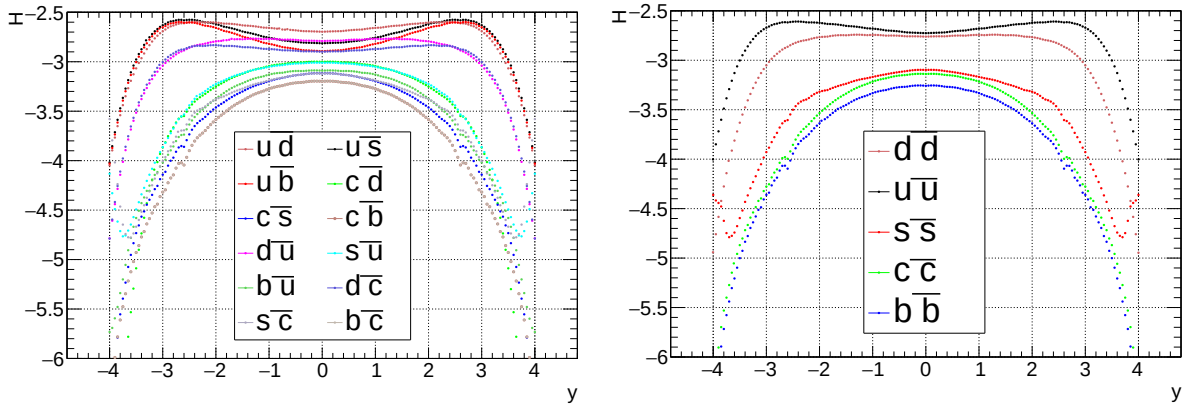
$$H = \frac{\sum_{\text{flavour } ij}\left[V_{ij}^{CKM}(f_i(\bar{x}_1)f_j(\bar{x}_2)\left(\frac{f_i'(\bar{x}_1)}{f_j(\bar{x}_1)}\bar{x}_1 + \frac{f_j'(\bar{x}_2)}{f_i(\bar{x}_2)}\bar{x}_2\right) + f_i(\bar{x}_2)f_j(\bar{x}_1)\left(\frac{f_j'(\bar{x}_1)}{f_i(\bar{x}_1)}\bar{x}_1 + \frac{f_i'(\bar{x}_2)}{f_j(\bar{x}_2)}\bar{x}_2\right)\right]}{\sum_{\text{flavour } ij}\left[V_{ij}^{CKM}(f_i(\bar{x}_1)f_j(\bar{x}_2) + f_i(\bar{x}_2)f_j(\bar{x}_1)\right]}\Big|_{Q=M_V} \tag{B.2}$$

B.3 Additional details on the numerical evaluation of the shift

The H^{ij} as a function of y is shown in Fig. B.1 for W and Z parton flavour pairs. It represents the flavour-breakdown of the $\Delta_V(y)$ reported in Fig. 3.2. The trend in the combined $\Delta_V(y)$ is ruled by the valence quark combination.

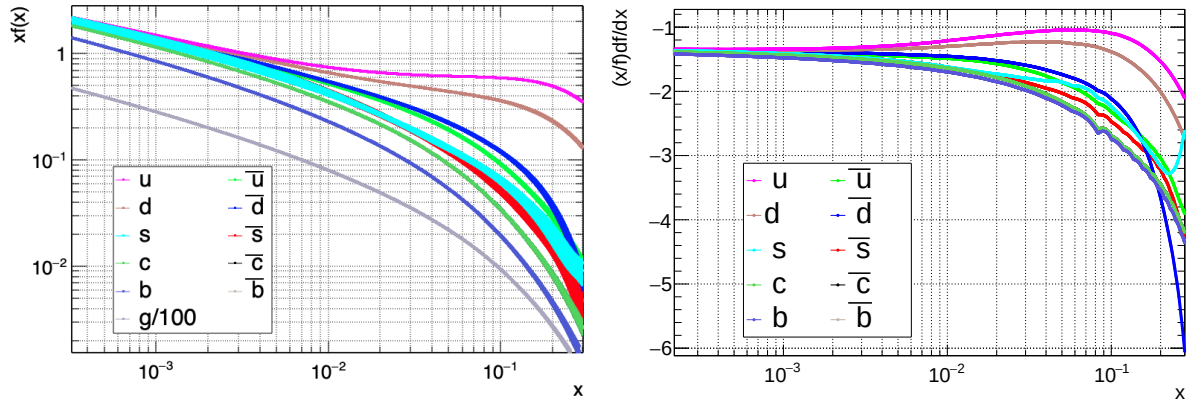
It is also interesting to study the x dependence of the PDF induced shift. As reference in Fig. B.2(a) are reported the PDF, as $xf(x)$, in the relevant x range. The width of the PDF curves represents the considered uncertainties estimated with the replicas. In the studies reported in Chapter 3, has been considered vector boson with $|y_V| < 3.4$. With this selection the available x range is $2 \cdot 10^{-4} \lesssim x \lesssim 0.18$. The value of $xf'(x)/f$ is reported in Fig. B.2(b) as a function of x, for each parton flavour. This quantity, from Eq. 3.5, represents the single-parton flavour contribution to H. It is relatively flat at low x, while starts to strongly increase in absolute value for $x > 0.1$, in the region where the PDFs are close to 0.

The valence quark PDFs induce a shift closer to 0, compared to the sea ones, but they represent the leading contribution since are the most present, especially at high x.



(a) (b)

Figure B.1. *First order shift parameter H as a function of y_V, for the relevant parton flavours pairs, for $W^{\pm}$ (a) and Z (b) cases.*



(a) (b)

Figure B.2. *PDFs as $xf(x)$ (a) and the $xf'(x)/f(x)$ function, the kernel of the first order shift parameter H (b), for each parton flavour, as a function of x in the relevant range.*

B.4 Robustness test: a study of H as a function of the injected M_V

The Eq. 3.4 has been exploited to study the shift Δ_V as a function of y. The latter equation for $\frac{d\sigma}{dQ}$ is the result of an expansion around $Q = M_V$. However, M_V is a quantity with an experimental and theoretical uncertainty much larger than the one quoted for Δ_V, therefore a study to check the robustness of $\Delta_V(y)$ as a function of the value of M_V used in the expansion is needed.

The value of Δ_V has been estimated as a function of y in the range [-4, 4] and M_V in range [78 GeV, 82 GeV] i.e. the fit range adopted in Sec. 3.3.1). As the nominal case, H is evaluated from the PDFs set for each parton flavour pair and averaged, as a function of M_V e y. At fixed y, the variation of H along M_V is less than 1% in the considered range. Thus, this variation is completely covered by the PDF replicas uncertainty. Therefore the assumption of the expansion around the nominal value of M_V is fully consistent.

B.5 Robustness test: a study of the quality of the lineshape approximation

In Chapter 3 has been mentioned that the disagreement between the Eq. 3.4 and Eq. 3.1 lineshape is 0.5% in $M_V \pm 2$ GeV range. The relative discrepancy between the complete lineshape (Eq. 3.1, using the full PDF set information) and the first-order in Q/M_V approximation (Eq. 3.4 setting $K = 0$) is shown Fig.B.3 in a wider range of Q. The discrepancy is shown as:

$$\left(\frac{d\sigma}{dQ}|_{\text{complete}} - \frac{d\sigma}{dQ}|_{\text{1st-order}} \right) / \frac{d\sigma}{dQ}|_{\text{complete}}.$$

and estimated in bins of $|y|$. It shows clearly a parabolic trend which can be corrected by the introduction of the second order term. The discrepancy starts to be of the same magnitude of the PDF uncertainty at the edge of the shown range and at high $|y|$ only.

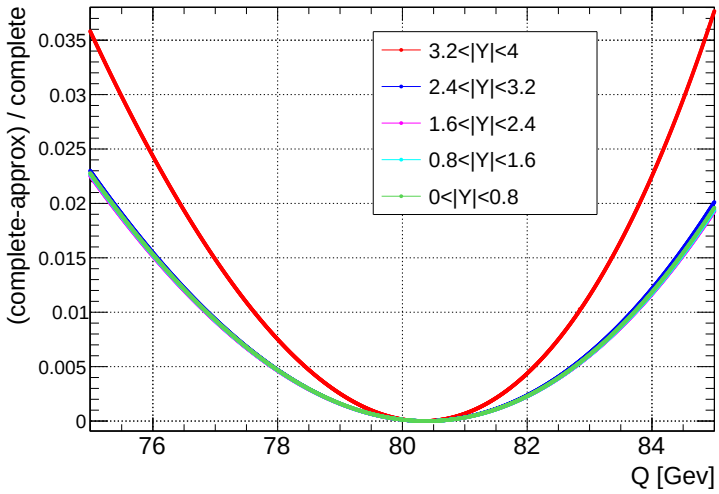


Figure B.3. *Discrepancy between the $d\sigma/dQ$ with the complete information and the first order in Q/M approximation, in bin of $|Y|$. The central bins of are almost overlapped.*

Appendix C

Derivation of templates analytic expression

The purpose of this appendix is to derive the results of Eq. 5.10 and Eq. 5.8. The former describes the relation between the muon and the W boson variables, thus encodes the features of the template distribution. The latter the describe the relation between $(\cos\theta^*, \phi^*)$ *in the CS frame and* (p_T^μ, η^μ) *in the laboratory frame.*

First, the Lorentz transformation between the laboratory frame and the CS frame must be defined. A W boson in the laboratory frame is considered. Without loss of generalities the four-momenta of the W boson will be $(E_W, q_x, 0, q_z)$, with $E_W^2 = q_x^2 + q_z^2 + m_W^2$ i.e. the x axis is aligned with the q_T direction. With this choice $q_T = q_x$.

A first boost of β_z, γ_z is performed along z to cancel the longitudinal momentum of the W boson. Choosing:

$$\beta_z \equiv -\frac{q_z}{E_W} = \frac{q_z}{\sqrt{m_W^2 + q_z^2 + q_T^2}}, \quad \gamma_z = \frac{\sqrt{m_W^2 + q_z^2 + q_T^2}}{\sqrt{m_W^2 + q_T^2}},$$

is obtained:

$$(E_W, q_x, 0, q_z) \xrightarrow{\beta_z} (E'_W, q_x, 0, 0), \quad \text{with } E'_W = q_T^2 + m_W^2. \tag{C.1}$$

A second boost of β_x, γ_x is performed along x, to reach the W rest frame. Choosing:

$$\beta_x \equiv -\frac{q_T}{E'_W} = -\frac{q_T}{\sqrt{m_W^2 + q_T^2}}, \quad \gamma_x = \frac{\sqrt{m_W^2 + q_T^2}}{m_W}$$

is obtained:

$$(E'_W, q_x, 0, 0) \xrightarrow{\beta_x} (E''_W, 0, 0, 0), \quad \text{with } E''_W = m_W^2. \tag{C.2}$$

If this pair of boosts is applied to the proton beam in the laboratory frame, neglecting their masses:

$$(0, 0, \pm E_B) \to (0, 0, E_B(\pm 1 + \beta_z)\,),$$
$$(0, 0, E_B(\pm 1 + \beta_z)\,) \to E_B(1 \pm \beta_z)(\gamma_x \beta_x, 0, \pm 1),$$

and the angle between the beam direction the z axis in the final reference frame is $\frac{q_x}{m_W}$ for both beams. Given the choice of x axis, this reference frame follows the definition of the CS frame.

A muon in the laboratory frame has a four momentum equivalent to $p_T(\cosh\eta, \cos\phi, \sin\phi, \sinh\eta)$, where p_T and η are the transverse momentum and the pseudorapidity of the muon, respectively. The angle $\phi = \phi^{\mu} - \phi^{W}$, but given the previous q_T definition, $\phi = \phi^{\mu}$ (obviously this is not true in general, but ϕ is the quantity which contains the physical information). The muon mass is neglected in the entire discussion.

The boost along z from Eq. C.1 and along x from Eq. C.2 are now applied:

$$p_T \begin{pmatrix} \cosh\eta \\ \cos\phi \\ \sin\phi \\ \sinh\eta \end{pmatrix} \xrightarrow{\beta_z} p_T \begin{pmatrix} \gamma_z(\cosh\eta + \beta_z \sinh\eta) \\ \cos\phi \\ \sin\phi \\ \gamma_z(\sinh\eta + \beta_z \cosh\eta) \end{pmatrix} \xrightarrow{\beta_x} p_T \begin{pmatrix} \gamma_x(\gamma_z(\cosh\eta + \beta_z \sinh\eta) + \beta_x \cos\phi) \\ \gamma_x(\cos\phi + \beta_x(\gamma_z(\cosh\eta + \beta_z \sinh\eta)) \\ \sin\phi \\ \gamma_z(\sinh\eta + \beta_z \cosh\eta) \end{pmatrix}. \tag{C.3}$$

The energy of the muon in the CS frame is $E^* = m_W/2$ and is possible to express the energy of the muon in the CS frame as:

$$\begin{aligned} E^* =& m_W/2 = p_T\gamma_x(\gamma_z(\cosh\eta + \beta_z \sinh\eta) + \beta_x \cos\phi) \\ =& p_T \frac{\sqrt{m_W^2 + q_T^2}}{m_W} \left(\frac{\sqrt{m_W^2 + q_T^2 + q_z^2}}{\sqrt{m_W^2 + q_T^2}} (\cosh\eta - \frac{q_z}{\sqrt{m_W^2 + q_T^2 + q_z^2}} \sinh\eta) - \frac{q_T}{\sqrt{m_W^2 + q_T^2}} \cos\phi \right) \\ =& p_T \left(\frac{\sqrt{m_W^2 + q_T^2 + q_z^2}}{m_W} \cosh\eta - \frac{q_z}{m_W} \sinh\eta - \frac{q_T}{m_W} \cos\phi \right). \end{aligned} \tag{C.4}$$

Now is possible to express q_Z as a function the W boson rapidity Y with the relations:

$$\begin{aligned} q_z =& \sqrt{m_W^2 + q_T^2} \sinh Y \\ \sqrt{m_W^2 + q_T^2 + q_z^2} =& \sqrt{m_W^2 + q_T^2} \cosh Y. \end{aligned}$$

The relation of Eq. C.4 can be rewritten as:

$$\begin{aligned} E^* = m_W/2 =& p_T \left(\frac{\sqrt{m_W^2 + q_T^2} \cosh Y}{m_W} \cosh\eta - \frac{\sqrt{m_W^2 + q_T^2} \sinh Y}{m_W} \sinh\eta - \frac{q_T}{m_W} \cos\phi \right) \\ =& p_T \left(\frac{\sqrt{m_W^2 + q_T^2}}{m_W} (\cosh Y \cosh\eta - \sinh Y \sinh\eta) - \frac{q_T}{m_W} \cos\phi \right) \\ =& p_T \left(\frac{\sqrt{m_W^2 + q_T^2}}{m_W} \cosh(Y - \eta) - \frac{q_T}{m_W} \cos\phi \right). \end{aligned} \tag{C.5}$$

Finally the target formula of Eq. 5.10 is obtained reverting the Eq. C.5. Is worth considering the simplified case at $q_T = 0$, which leads to:

$$p_T = \frac{m_W/2}{\cosh(Y - \eta)},$$

where the relation is independent from ϕ and it brings to the phenomenology discussed originally in Ref. [8].

A muon in the CS frame must be initially considered to derive Eq. 5.8. Like in the previous case the muon mass is neglected in the entire discussion. The four momentum of the muon is $E^*(1, \sin\theta^*\cos\phi^*, \sin\theta^*\sin\phi^*, \cos\theta^*)$. The use of the revered Lorentz transformation of the previous gives :

$$E^* \begin{pmatrix} 1 \\ \sin\theta^*\cos\phi^* \\ \sin\theta^*\sin\phi^* \\ \cos\theta^* \end{pmatrix} \overset{-\beta_x}{\to} E^* \begin{pmatrix} \gamma_x + \gamma_x\beta_x\cos\phi^*\sin\theta^* \\ \gamma_x\cos\phi^*\sin\theta^* + \gamma_x\beta_x \\ \sin\phi^*\sin\theta^* \\ \cos\theta^* \end{pmatrix} \overset{-\beta_z}{\to} E^* \begin{pmatrix} \gamma_z(\gamma_x + \gamma_x\beta_x\cos\phi^*\sin\theta^*) + \gamma_z\beta_z\cos\theta^* \\ \gamma_x\cos\phi^*\sin\theta^* + \gamma_x\beta_x \\ \sin\phi^*\sin\theta^* \\ \gamma_z\cos\theta^* + \gamma_z\beta_z(\gamma_x + \gamma_x\beta_x\cos\phi^*\sin\theta^*) \end{pmatrix}. \tag{C.6}$$

The variables of the muon in the laboratory frame now can be written as a function of the CS angles:

$$p_T^2 = \frac{m_W^2}{4}\left((\gamma_x\cos\phi^*\sin\theta^* + \gamma_x\beta_x)^2 + \sin^2\phi^*\sin^2\theta^*\right), \tag{C.7}$$

$$\begin{aligned} \tan(\frac{\theta}{2}) &= \frac{\sin\theta}{1+\cos\theta} = \frac{p_T/E}{1+p_z/E} = \frac{p_T}{E+p_z} \\ &= \frac{\sqrt{(\gamma_x\cos\phi^*\sin\theta^* + \gamma_x\beta_x)^2 + \sin^2\phi^*\sin^2\theta^*}}{\gamma_z(\gamma_x + \gamma_x\beta_x\cos\phi^*\sin\theta^*) + \gamma_z\beta_z\cos\theta^* + \gamma_z\cos\theta^* + \gamma_z\beta_z(\gamma_x + \gamma_x\beta_x\cos\phi^*\sin\theta^*)} \\ &= \frac{\sqrt{(\gamma_x\cos\phi^*\sin\theta^* + \gamma_x\beta_x)^2 + \sin^2\phi^*\sin^2\theta^*}}{\gamma_z(1+\beta_z)(\gamma_x + \gamma_x\beta_x\cos\phi^*\sin\theta^* + \cos\theta^*)}, \end{aligned} \tag{C.8}$$

$$\cos\phi = \frac{p_x}{p_T} = \frac{\gamma_x\cos\phi^*\sin\theta^* + \gamma_x\beta_x}{\sqrt{(\gamma_x\cos\phi^*\sin\theta^* + \gamma_x\beta_x)^2 + \sin^2\phi^*\sin^2\theta^*}}. \tag{C.9}$$

These equations are exactly the relations of Eq. 5.8.

Finally, is also interesting to derive the approximation at low q_T^W, expanding at first order in q_T/m_W. In this case, the boost is $\gamma_x \simeq 1$ and $\beta_x \simeq q_T/m_W$ and thus:

$$\begin{aligned} \cos\phi &\simeq \cos\phi^* + \frac{\beta_x\sin^2\phi^*}{\sin\theta^*} \\ p_T^2 &\simeq \frac{m_W}{2}\sin\theta^* + \frac{q_T}{2}\cos\phi^* \end{aligned}$$

Appendix D

Technical tools: data format, reconstruction and analysis framework

The technical details of the W boson mass and properties analysis are collected and described in this appendix. The information about the CMS samples and reconstruction and the data format are reported in the first part. Then, the main features and the advantages of the custom analysis framework are described. Finally, the analysis workflow is summarized.

D.1 The CMS data formats

The CMS data flow is organized in multiple steps, with different purposes. The following datasets are centrally produced by CMS, both for data and simulated events:

- **Raw.** This dataset contains the digitization of the detector information collected by CMS in LHC collisions.
- **Reco.** This dataset contains the reconstructed physics objects and the processed hits and clusters information. The Raw data are processed immediately after the data taking (Prompt-Reco dataset) and reprocessed in a later stage (Re-Reco dataset). The average size is 1-3 MB/event.
- **AOD.** The Analysis Object Data format is produced from Reco dataset for physics analysis purpose. This is typically done a few times per year. The average event size is 480 kB/event.
- **MiniAOD.** This a reduced-size data format, developed for Run 2 analysis, starting from the AOD dataset. The event size is reduced to 30-100 kB/event, reaching globally 10% of the size of the AOD dataset [209]. The MiniAOD are designed to cope with most of the CMS physics analysis, and they can be produced in 1-2 days from the AOD dataset. This step is typically done on a monthly basis. The MiniAOD contains the high-level reconstructed physics objects (leptons, photons, jet, p_T^{miss}, PF candidates, etc.) as well as trigger information. For the MC samples, only the essential information about the generated particles is stored. No event selection is performed, but only a reduction of the stored information for each event.
- **NanoAOD.** This further reduced-size data format has been developed starting from MiniAOD dataset. The event size is strongly reduced, and reaches 1-3 kB/event. Despite the reduced sored information the NanoAOD has been designed to be exploited by

a large fraction of the CMS physics analysis (at least 50%) [14]. The Reco, AOD and MiniAOD data format store the information as EDM (Event Data Model), fully accessible thought `CMSSW` only. The NanoAOD are instead "flat" ntuples, accessible directly with `ROOT` software. The purpose of the NanoAOD is to provide a flexible data format, which requires limited input/output (I/O) resources and is potentially usable directly at analysis level.
These features have been obtained avoiding to store the information of the single particles, the detector information and most of generator level information, precomputing the particle ID flags, limiting the physics object collections size and optimizing the numerical precision of the stored objects. Finally, the variables which can be recomputed from the existing information are not stored. This is particularly relevant for the physics object variations typically required for the systematic uncertainties estimation. Thus, they are not directly stored, but must be evaluated in later analysis stages. On the other hand, the generator weights, which cannot be recomputed, are stored in the NanoAOD. No events selection occurs also in the NanoAOD case.
The NanoAOD data format is still in development together with the NanoAOD-Tools, the dedicated framework to post-process the NanoAOD and realize the analysis ntuples with the full computed information. More details are provided in Sec. D.3.

D.2 Details of samples and reconstruction

A simulated sample is typically identified by the generator event content and configuration, the production MC campaign (which represents the production period), the CMS *Global Tag* or GT (which represents the detector configuration) and the dataset format.

The NanoAOD-v6 data and MC samples have been used in the analysis described in this thesis. The runs from 273158 to 284044 has been used for the data samples. The MC samples have been produced by the CMS collaboration in October 2019, starting from the MiniAOD-v3 samples produced in the `RunIISummer16` campaign. The GT of the samples is `102X_dataRun2_v12` (data) and `102X_mcRun2_asymptotic_v7` (MC). The `CMSSW_10_2_18` software release has been used for the reconstruction of the data and MC samples.

The $W \to \ell\nu$ sample has been privately produced for this analysis. The numerical precision of the centrally produced sample on the generator level quantities was not sufficient for the purpose of the analysis. Thus, the sample has been reproduced with the same generator and configuration (GT, `CMSSW` release) of the central one, but with increased precision.

The full list of the samples used in the analysis is reported in table D.1. The QCD samples have been used for the validation test described in Sec. 6.3.3.3 only. Additional details about the samples and the reconstruction can be found at the Ref. [210] and [211].

Table D.1. *MC samples used in the analysis, with the full unique name containing the generator, campaign, GT and dataset information.*

Process	Sample
$W(\to \ell\nu)$+jets	privately produced
$Z(\to \ell\ell)$	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext2-v1/NANOAODSIM
$Z(\to \ell\ell)$,m10-50	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
$t\bar{t}(\ell)$	/TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
$t\bar{t}(\ell\ell)$	/TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
t (t-channel)	/ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
$\bar{t}$ (t-channel)	/ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
top (s-channel)	/ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
tW	/ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
	/ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
WW	/WW_TuneCUETP8M1_13TeV-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/WW_TuneCUETP8M1_13TeV-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
WZ	/WZ_TuneCUETP8M1_13TeV-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/WZ_TuneCUETP8M1_13TeV-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
ZZ	/ZZ_TuneCUETP8M1_13TeV-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/ZZ_TuneCUETP8M1_13TeV-pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
QCD	/QCD_Pt-1000toInf_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-1000toInf_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
	/QCD_Pt-120to170_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-120to170_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_backup_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-15to20_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-170to300_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-170to300_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
	/QCD_Pt-170to300_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_backup_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-20to30_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-300to470_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-300to470_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
	/QCD_Pt-300to470_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext2-v1/NANOAODSIM
	/QCD_Pt-30to50_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-470to600_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-470to600_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
	/QCD_Pt-470to600_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext2-v1/NANOAODSIM
	/QCD_Pt-50to80_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-600to800_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-600to800_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
	/QCD_Pt-600to800_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_backup_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-800to1000_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-800to1000_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM
	/QCD_Pt-800to1000_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext2-v1/NANOAODSIM
	/QCD_Pt-80to120_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7-v1/NANOAODSIM
	/QCD_Pt-80to120_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8/RunIISummer16NanoAODv6-PUMoriond17_Nano25Oct2019_102X_mcRun2_asymptotic_v7_ext1-v1/NANOAODSIM

D.3 Post-processor: from NanoAOD to analysis ntuples

The NanoAOD are the optimal format for the W mass and properties analysis. This analysis adopted a very large MC sample, required for a precise modelling of the templates of the fit. The total size of the data and MC samples is about 500 GB in the NanoAOD format. The dataset size is reduced of about a factor 20 compared to the equivalent MiniAOD samples. This allows for much more flexibility in term of disk usage and I/O performance. Moreover, the NanoAOD format provides a faster and lighter procedure to build the analysis ntuples, compared to MiniAOD. This allows to quickly reproduce the ntuples from the central datasets multiple times, if necessary.

The NanoAOD contains enough information for the purposes of the W mass and properties analysis, however a post-processing step is required to evaluate the variables needed for the analysis. As previously mentioned the NanoAOD-Tools framework is provided by CMS for this purpose. In this framework is possible to develop specific modules to derive high-level variables or to evaluate the variations needed for the systematic uncertainties estimation. In addition, it can be used to apply some preliminary skimming to the samples and obtain lighter and higher purity ntuples.

In particular, in the post-processing step the m_T variable and the CS reference frame variables have been estimated. The entire set of experimental variations of the variables has been evaluated at this level, and in particular the p_T^{μ} variation (with the Rochester correction guidance) and the p_T^{miss} variations (according to the indication of the Jet-MET POG of CMS).

The events have been skimmed applying at this level a part of the analysis selection described in Sec. 5.3.4. In particular, the requirements on the primary vertex and the muon ID, impact parameter and fiducial region selection have been applied. The remaining selection requirements have been instead applied at analysis level. The subdivision between this preselection at post-processing level and the selection at analysis level have no physical meaning, but has been performed for practical reasons. The selection requirement applied at analysis level has been tuned in the development of the analysis. Thus it was necessary to quickly change the selection without reproducing the ntuples. On the other hand, the requirements applied with the post-processor were expected to be stable and not under optimization.

After these post-processing procedure the analysis ntuples have been realized keeping or dropping the relevant branches from the `ROOT` events tree contained in the NanoAOD files.

D.4 RDataFrame W-properties analysis framework

The W mass and properties analysis had to face several challenging tasks from the analysis framework and data management side. It exploited very large data and MC samples. The analysis has been performed deferentially in multiple variables (q_T^W, Y^W, A_i, p_T^{μ}, η^{μ}, etc.) requiring the use of a large number of multidimensional histograms. The analysis had also to take into account a large number of different copies of these histograms to describe the full set of the considered systematic variations.

The adopted data and MC samples are composed of $\mathcal{O}(10^8)$ events. Let us consider for instance the use of $p_T^{\mu} \times \eta^{\mu} \times \text{charge(W)}$ 3D histograms. With the binning adopted in the analysis about 400 histograms of $5 \cdot 10^3$ bin each, must be produced for the nominal version. This number must be multiplied for a factor 100, considering the systematic uncertainty variation for a total of $4 \cdot 10^4$ histograms. The background templates and additional histograms needed for the analysis must be also considered. This large amount of events and histograms must be managed and

processed with different purposes in the analysis workflow.

In conclusion, to face this analysis scenario, an analysis framework must include the following features:

- It must be fast, with a smart handling of the I/O, and easily parallelizable
- It must be tidy, flexible and customizable to allow quick change of the details of the workflow during the optimization of the analysis.
- It worth do develop a framework which can be reused for future, more accurate analysis (eg. the m_W measurement with full Run 2 data) or also in different analysis, thanks to the flexibility. The difficulties of this analysis will be faced by much more analyses in the future of CMS.

The adopted solution has been the `ROOT` package `RDataFrame` (RDF). The relevant features of this package will be outlined in the following paragraphs, but this section has not the purpose to be a comprehensive description of RDF.

The RDF package offers a declarative framework with several of the required features built-in. First, RDF is designed to be completely parallelizable, thus this implementation can be adopted without any additional effort, in a clear and transparent way. Second, RDF allows to avoid multiple iterations of the events-loop for histograms building. This is realized thanks to a *lazy actions* system. A lazy action is executed only when the object it returns is accessed for the first time (and not at the first call of the action). Adopting this feature the entire set of histograms can be booked preliminary, and then filled in a single events-loop when the histograms are written on the output file. Obviously, this guarantees a strong reduction of the processing time. Third, the RDF workflow can be organized in modules with a graph structure. In the initial phase, the RDF object is built and the *Columns* (i.e. the variables) of the frameworks are defined. Then, different modules can apply transformations on the RDF object, *Defining* additional columns (higher-level variables) or applying *Filters* (selection on the defined columns). The output objects (eg. the histograms) are booked within the modules. The various modules can be placed sequentially in the analysis workflow, exploiting the previous defined and transformed columns, or the workflow can branch in different paths, producing a parallel development of the graph.

The RDF package is in constant development within the active and vibrant `ROOT` developer community. The W-properties framework has been developed in close touch with the `ROOT`'s developer, allowing proficient feedback and optimization. Actually, the development of RDF and W-properties framework proceeds in parallel. Currently, the latter has been already moved to a more advanced and performing version, compared to the one used to produce the results of this thesis.

The W-properties framework on the top of RDF, mainly in `C++` language, with a `python` interface. The framework exploits the modularity allowed by RDF. Each module performs a different task and is typically called several times in different configurations (according to different samples, kinematic selection or systematic variation). The python interface is also devoted to linking the various modules. The framework exploits the RDF native parallelization. A study of the scaling of framework performance as a function of the number of cores has been performed and the results are reported in Fig. D.1. The study shows how a factor 500 in the number of histograms results in a reduction of a factor 4 in the event processing frequency, from 1600 kHz to 500 kHz at the optimal number of cores (128). The reduction of the processing frequency at higher number of cores has been studied and is due to technical `ROOT` limitations. The test has been performed on the CMS-Pisa server, a machine equipped with a AMD EPYC 7742 processor, 256 cores, 2TB memory (DDR4, 3200 MHz) and a SSD-nvme disk of 54 TB.

Additional details about RDF can be found at Ref. [13, 212], while the W-properties analysis framework [213] has been presented in [214, 215].

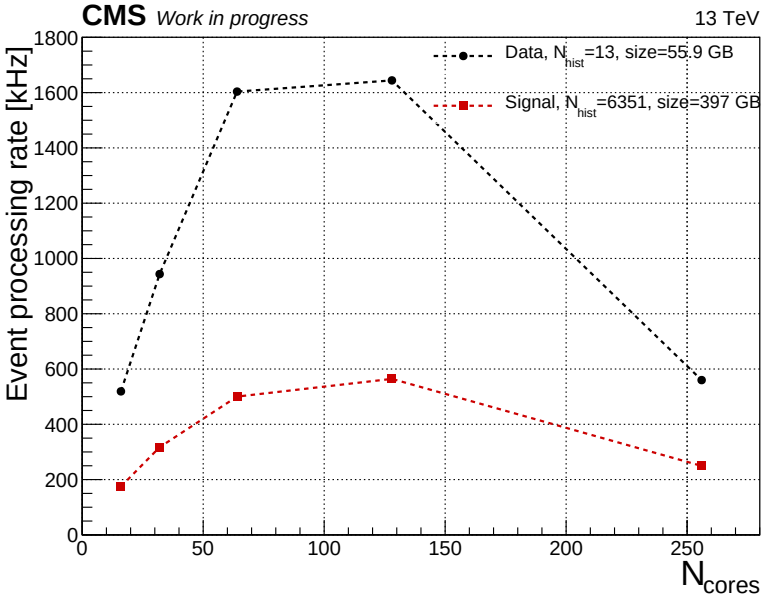


Figure D.1. *Event processing frequency in the W-properites RDF-based framework. Test performed on CMS-Pisa server (technical specifics in the text). A run on data sample and on Signal sample ($W \to \ell\nu$) are compared.*

D.5 Summary of the analysis workflow

The full analysis workflow is presented in Fig. D.2. First, the NanoAOD are skimmed in the NanoAOD-tools post-processor, to obtain the analysis ntuples. Then, the ntuples are plugged in the RDF W-properties framework with the additional input (q_T^Z and Y_Z measured spectra and W helicity efficiency scale factors). Multiple workflows proceed in parallel inside the RDF framework. In the background workflow, the input histograms for the background measurement are prepared. Then the background analysis is performed (outside the RDF Framework), and the background parameters (i.e. the fake and prompt rate parameters in bins of η^μ) are passed to the RDF framework module which produces the background templates. At this level is possible to produce the background validation plots, from the background analysis interface. In the data workflow the data distributions are built. In the signal workflow, the Generator-level distributions (Y_W, q_T^W, angular coefficients A_i and m_W) are estimated and the signal templates are built. At this level is possible to produce the pre-fit plots, where the data and the MC is compared preliminary to the fitting step. From the signal workflow, the generator-level distributions are permanently stored and provided to the regularization analysis. This step does not provide an output objects, but return only the optimal parameterization which will be used in the fit. However, also in this step is possible to produce dedicated validation plots. The templates, combining data, signal and background workflow output, and the generator level distributions are provided to the `Combine-TensorFlow` fit. It returns the final result of the analysis, from which a detailed set of validation and post-fit plots are produced.

The analysis is presented in this thesis has been performed using 128 cores, given the scaling profile of Fig. D.1, running on the server of CMS-Pisa. All the steps of the analysis are paral-

lelizable (partially or completely). The entire analysis workflow, from the post-processed ntuples to the final results takes about 2-3 hours and can be performed in a completely automatic mode, from a single configuration script.

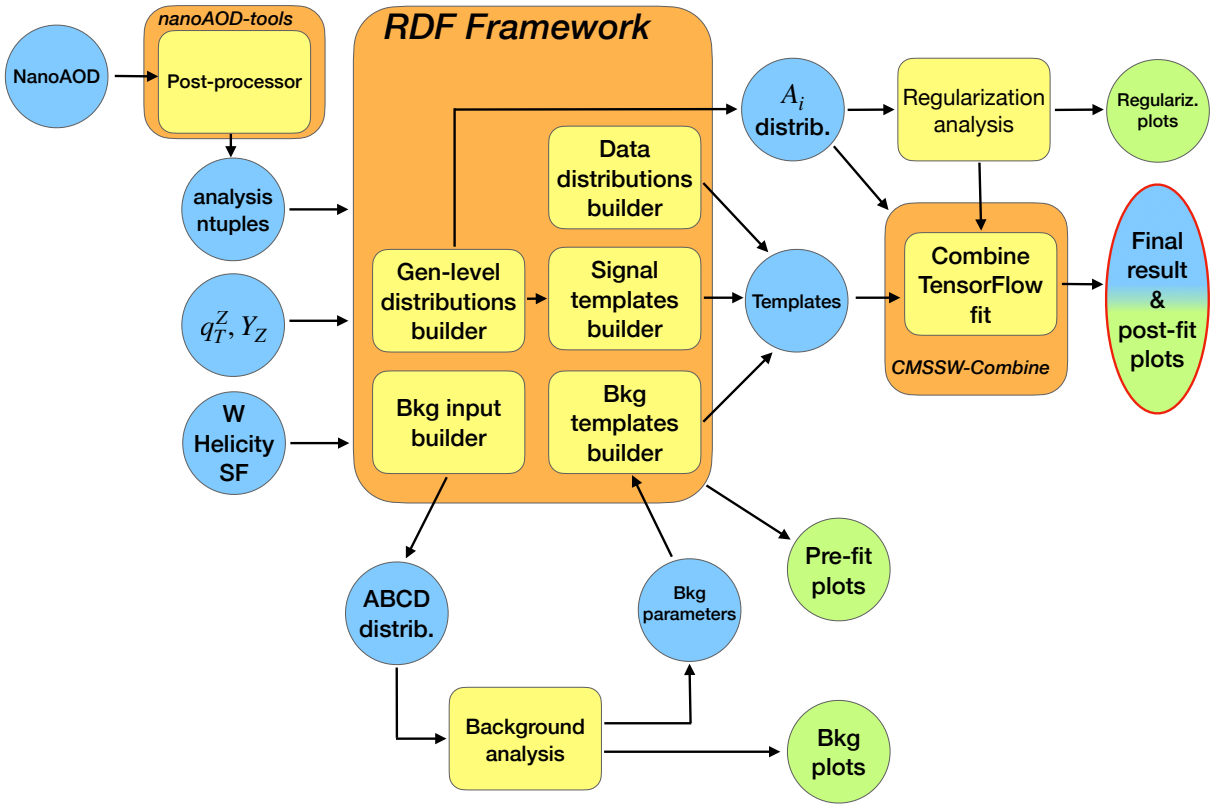


Figure D.2. *Full W mass and properties workflow. In the yellow boxes are represented the analysis steps, in the blue circles the input/output objects, in the green circles the validation plots. The analysis steps performed within the NanoAOD-tools, RDF or CMSSW-Combine environments are surrounded in larger orange boxes.*

Appendix E

Additional template fit plots

In this appendix additional plots from the fit results discussed in Chapter 7 are reported for completeness.

The following plots are shown:

- Impact of the systematic uncertainties on σ^{U+L} and A_0, A_1, A_2, A_3, A_4, in Figs. E.1 to E.6, respectively. The double and single-differential cases, for W^+, W^-, are shown in each figure.
- Correlation matrices of σ^{U+L} and A_0, A_1, A_2, A_3, A_4, double and single-differential cases, for W^-, in Fig. E.7. This is complementary to W^+ case shown of Fig. 7.12 and 7.14.
- Fit to the Asimov dataset results with the post-fit linear fit regularization applied, for σ^{U+L} and A_0, A_1, A_2, A_3, A_4, in Figs. E.8 to E.13. The double and single-differential case, for W^+, W^- is shown in each figure.

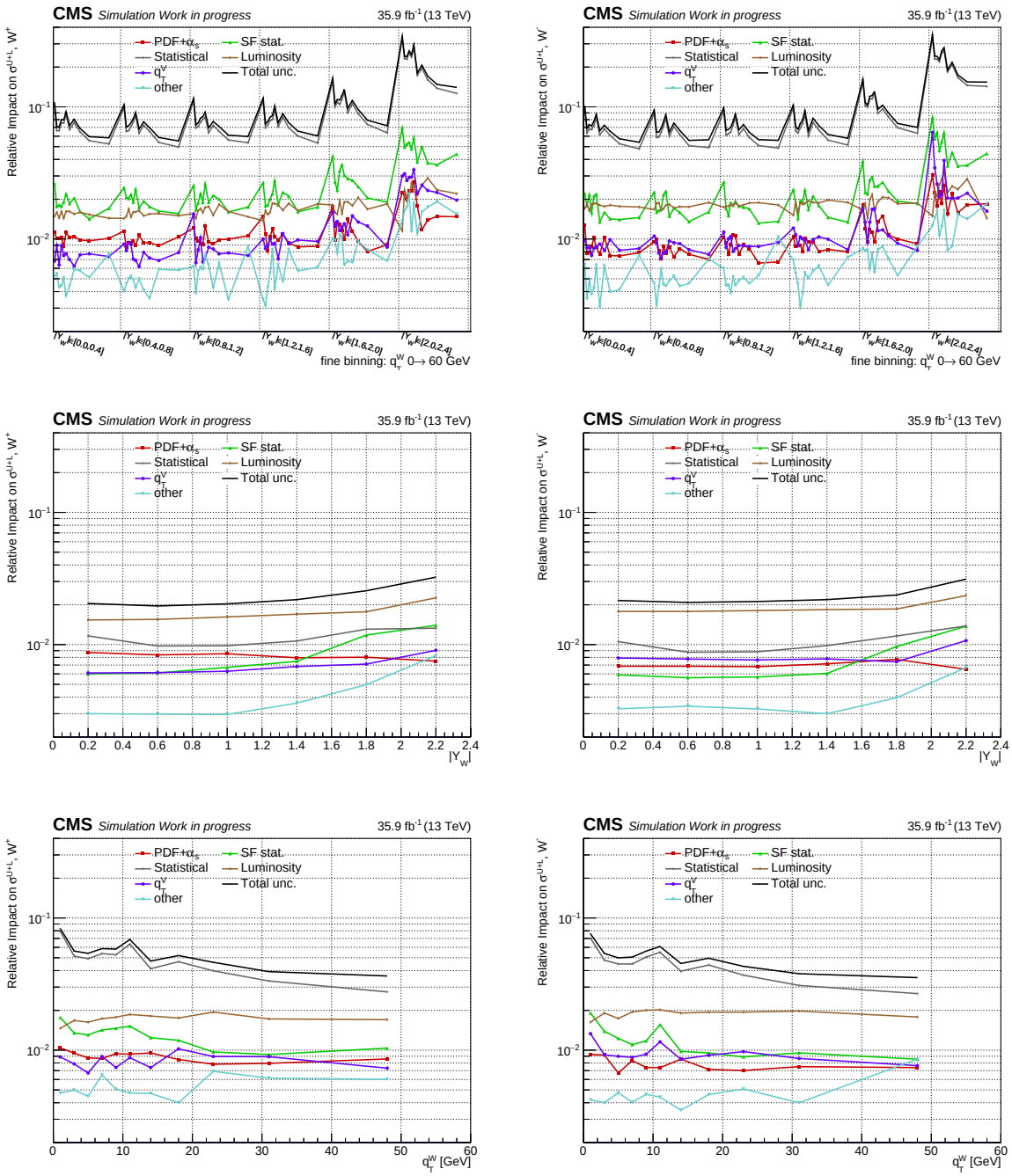


Figure E.1. *Relative impacts on σ^{U+L} double-differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central) and differential in q_T^W (lower), for W^+ (left) and W^+ (right) samples. The impacts are evaluated for each group and then divided for the quantity itself. The group "other" includes the lepton veto, the electroweak background cross sections, the QCD normalization, the MET-related systematics, the p_T^μ scale, the systematic uncertainty on the SF, the L1 Trigger prefire, m_W. The BBB uncertainty has not been considered coherently to the fit results.*

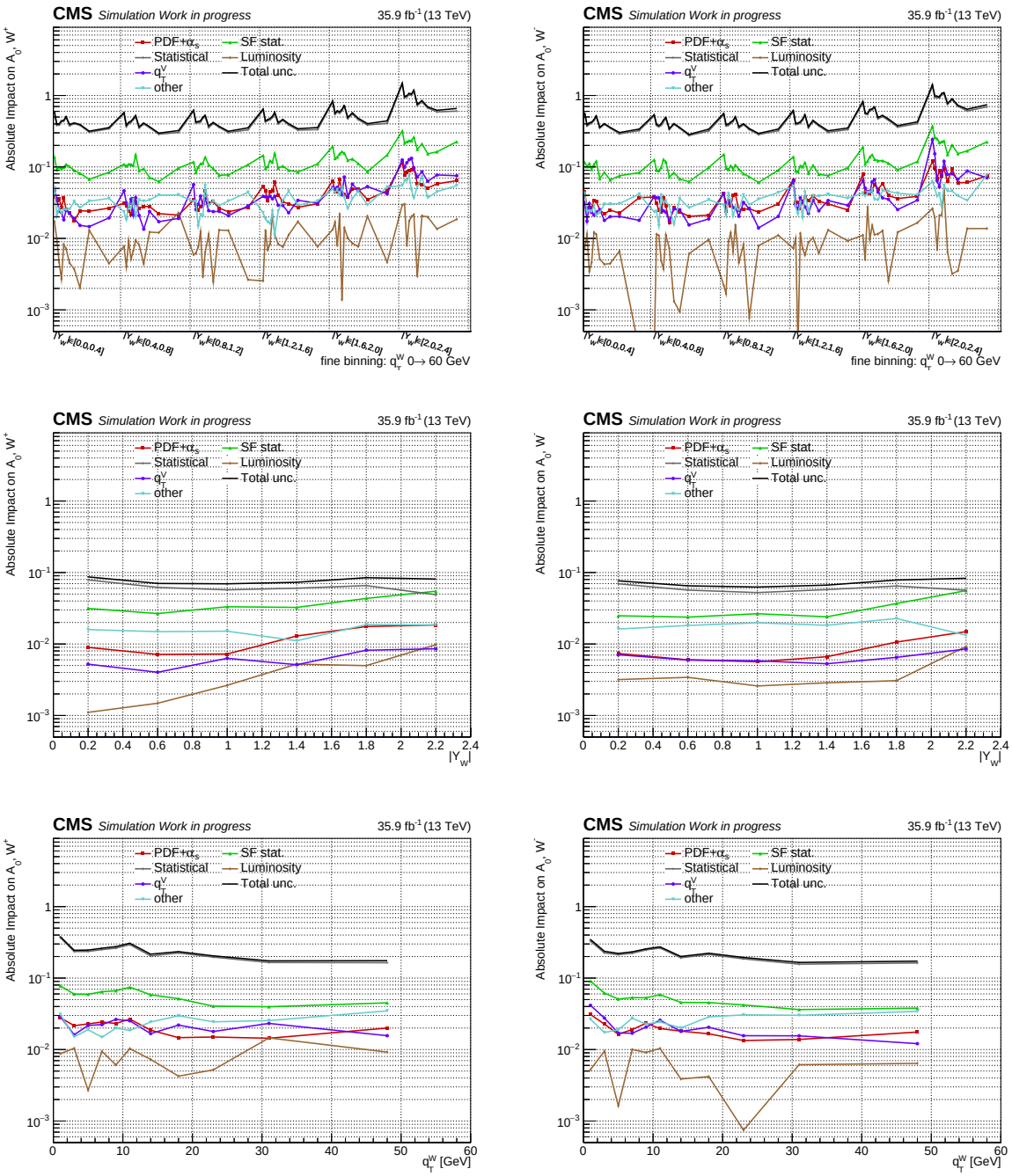


Figure E.2. *Absolute impacts on A_0 double-differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central) and differential in q_T^W (lower), for W^+ (left) and W^+ (right) samples. The impacts are evaluated for each group and then divided for the quantity itself. The group "other" includes the lepton veto, the electroweak background cross sections, the QCD normalization, the MET-related systematics, the p_T^μ scale, the systematic uncertainty on the SF, the L1 Trigger prefire, m_W. The BBB uncertainty has not been considered coherently to the fit results.*

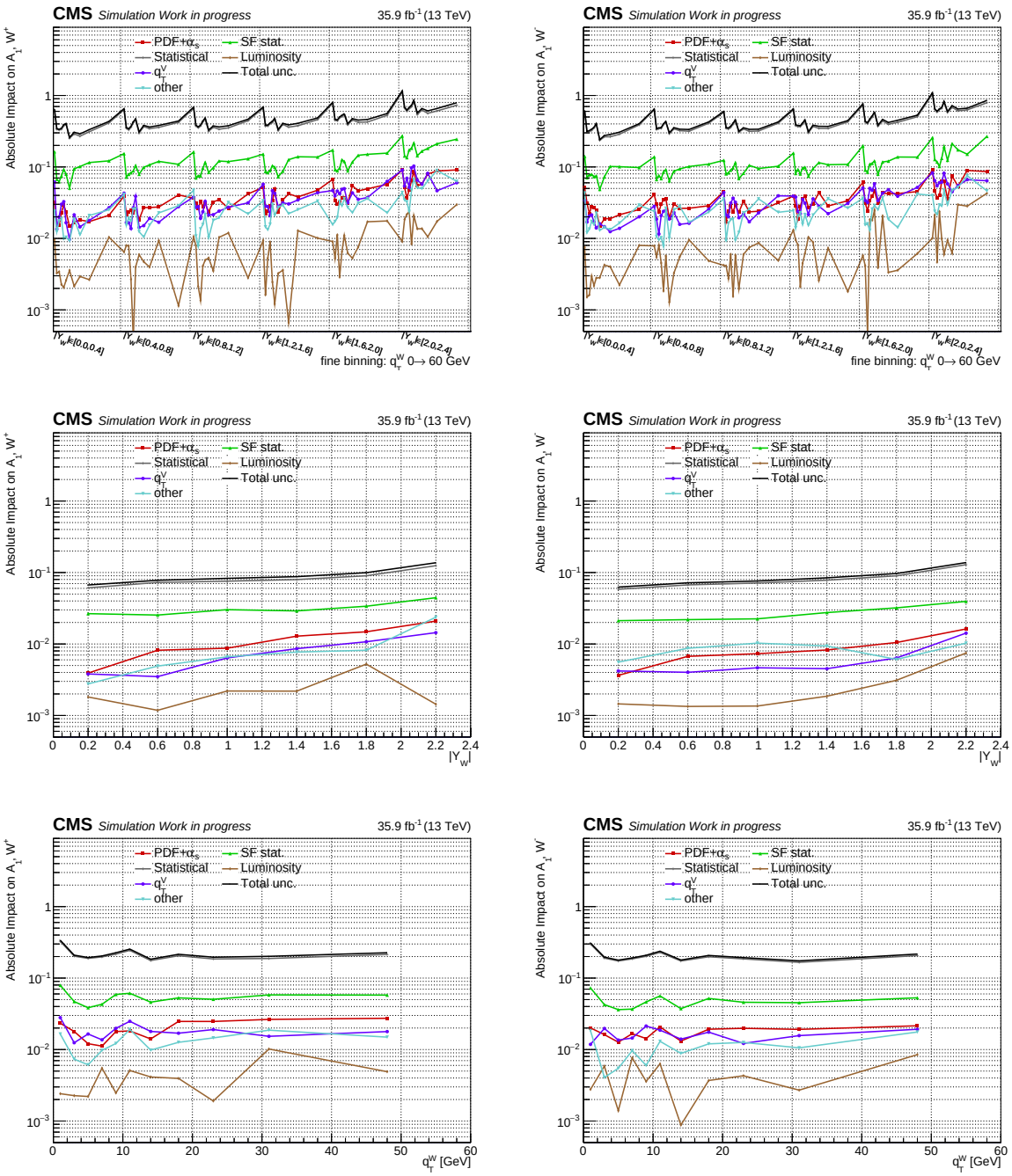


Figure E.3. *Absolute impacts on A_1 double-differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central) and differential in q_T^W (lower), for W^+ (left) and W^+ (right) samples. The impacts are evaluated for each group and then divided for the quantity itself. The group "other" includes the lepton veto, the electroweak background cross sections, the QCD normalization, the MET-related systematics, the p_T^μ scale, the systematic uncertainty on the SF, the L1 Trigger prefire, m_W. The BBB uncertainty has not been considered coherently to the fit results.*

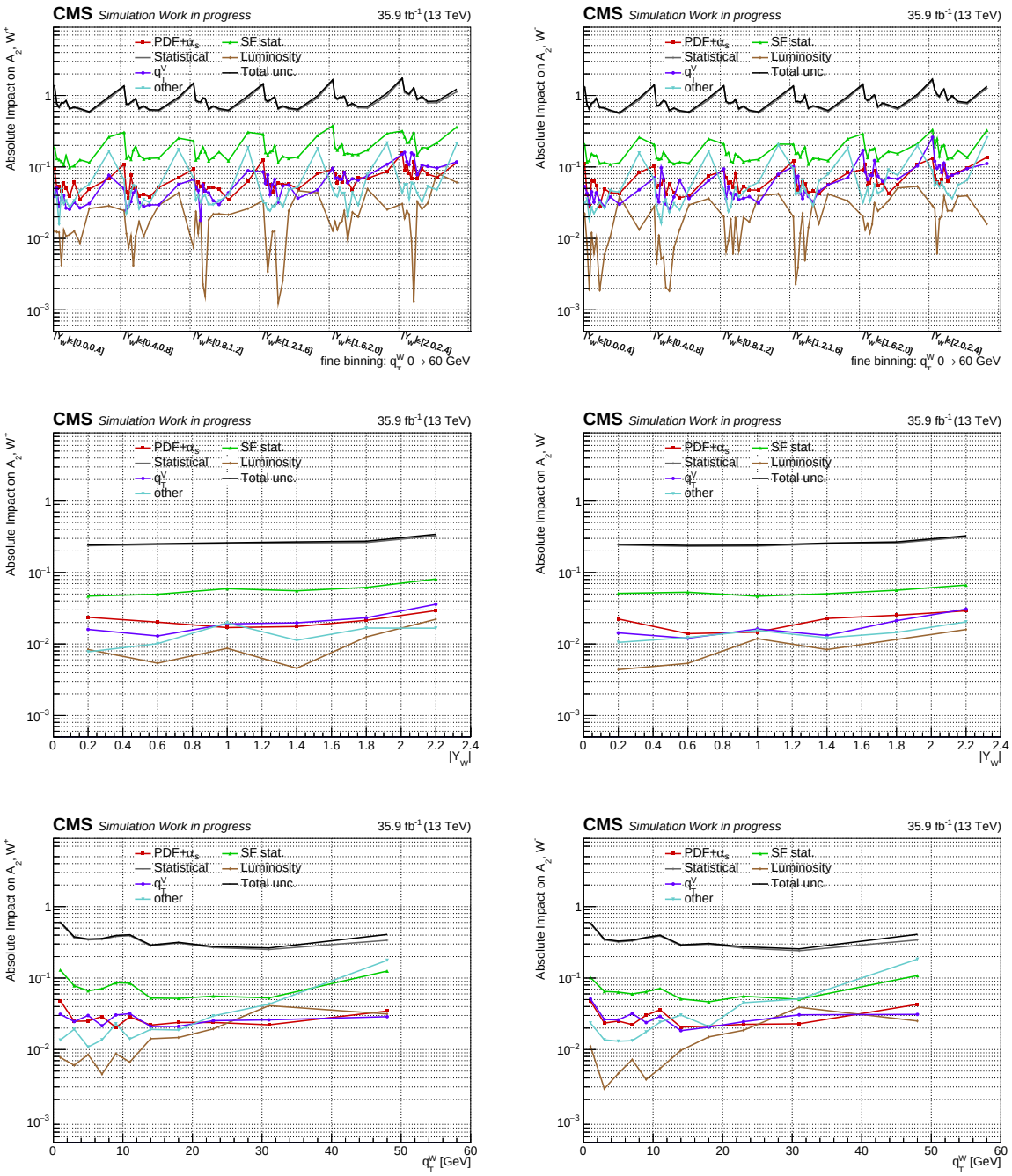


Figure E.4. *Absolute impacts on A_2 double-differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central) and differential in q_T^W (lower), for W^+ (left) and W^+ (right) samples. The impacts are evaluated for each group and then divided for the quantity itself. The group "other" includes the lepton veto, the electroweak background cross sections, the QCD normalization, the MET-related systematics, the p_T^μ scale, the systematic uncertainty on the SF, the L1 Trigger prefire, m_W. The BBB uncertainty has not been considered coherently to the fit results.*

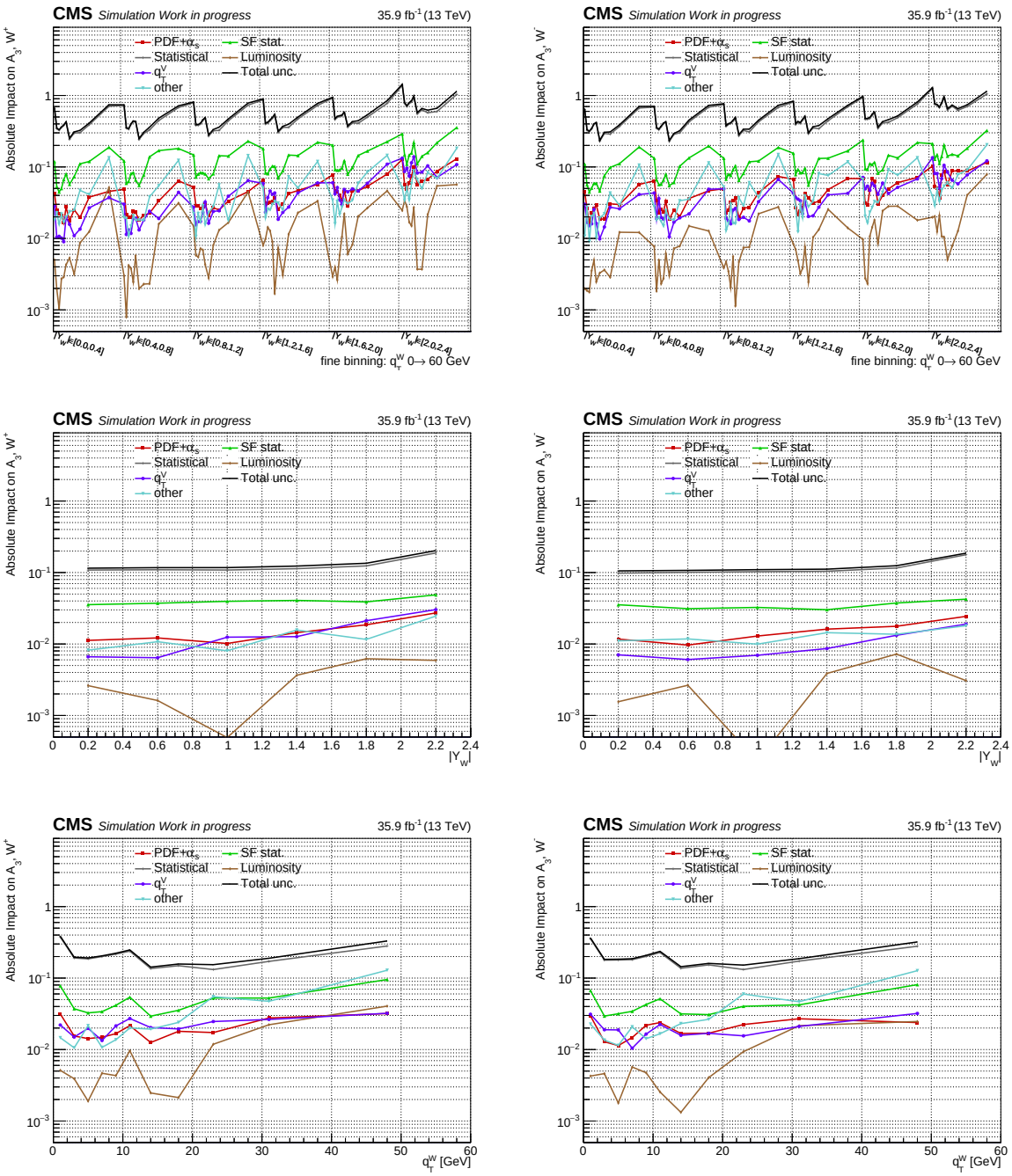


Figure E.5. *Absolute impacts on A_3 double-differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central) and differential in q_T^W (lower), for W^+ (left) and W^+ (right) samples. The impacts are evaluated for each group and then divided for the quantity itself. The group "other" includes the lepton veto, the electroweak background cross sections, the QCD normalization, the MET-related systematics, the p_T^μ scale, the systematic uncertainty on the SF, the L1 Trigger prefire, m_W. The BBB uncertainty has not been considered coherently to the fit results.*

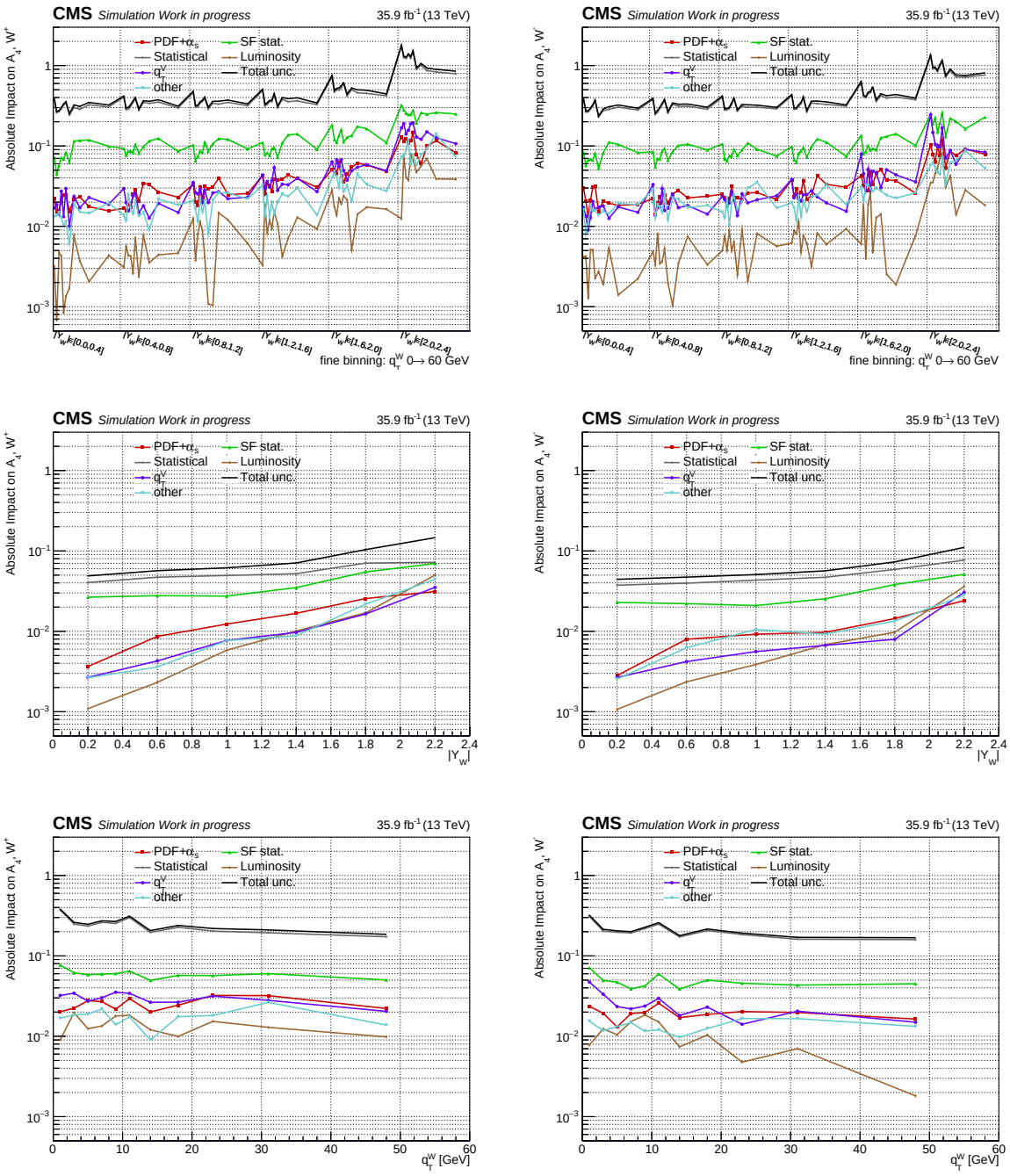


Figure E.6. *Absolute impacts on A_0 double-differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central) and differential in q_T^W (lower), for W^+ (left) and W^+ (right) samples. The impacts are evaluated for each group and then divided for the quantity itself. The group "other" includes the lepton veto, the electroweak background cross sections, the QCD normalization, the MET-related systematics, the p_T^μ scale, the systematic uncertainty on the SF, the L1 Trigger prefire, m_W. The BBB uncertainty has not been considered coherently to the fit results.*

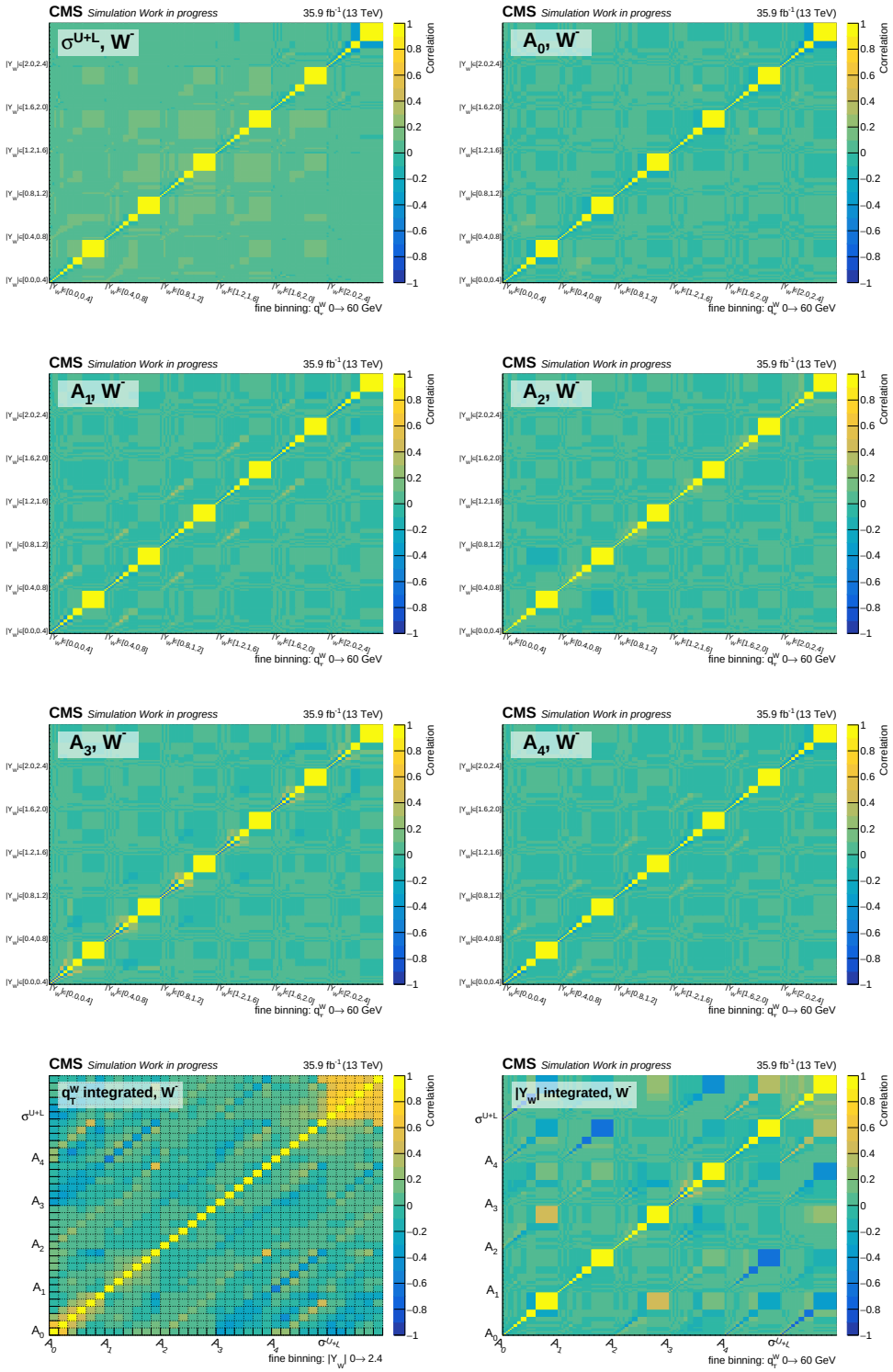


Figure E.7. *(6 upper plots): Different sectors of the correlation matrix for the double-differential σ^{U+L} and the five angular coefficients distributions; each matrix sector contains the correlation between all the $|Y_W|$, q_T^W processes for a given angular coefficient; the correlation between the processes of different angular coefficients and the correlations with the nuisance parameters are not shown. (Two lower plots): Full A_i, σ^{U+L} correlation matrix for the single differential measurement in $|Y_W|$ and for the single-differential measurement in q_T^W. The correlation between the A_i, σ^{U+L} and the nuisance parameters are not shown. In both cases the fit has been performed on the Asimov dataset for the W^- sample. The widths of the bins of the double differential and the q_T^W-differential matrices are proportional to the width of the correspondent q_T^W bins. The BBB uncertainty has been not considered.*

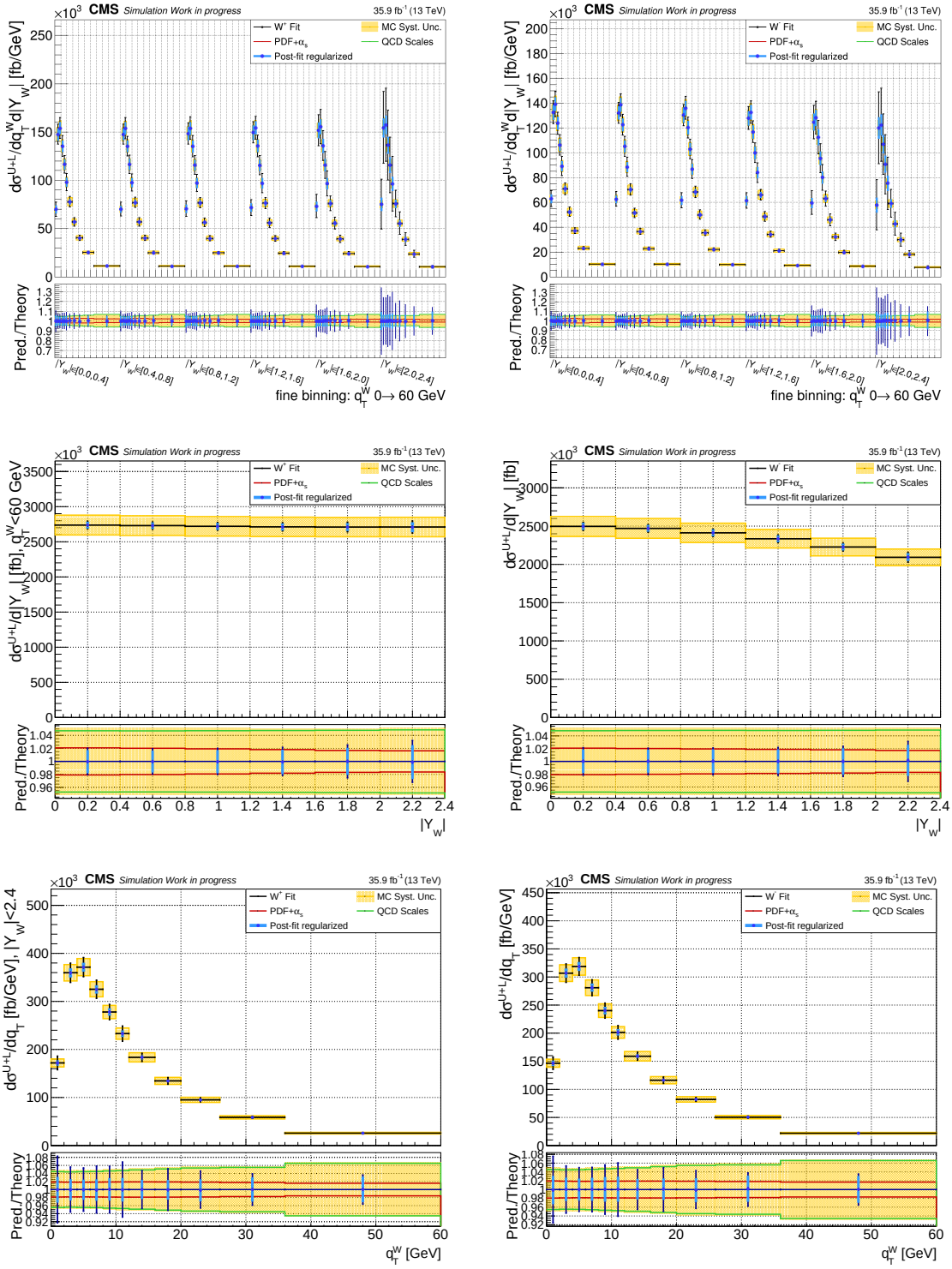


Figure E.8. *Fitted σ^{U+L}, double differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central), differential in q_T^W (lower) for W^+ (left) and W^- (right) samples. The fit has been performed on the Asimov dataset. The yellow shaded area corresponds to the expected theory uncertainty, summing in quadrature the PDF and the MC scale contributions. The post-fit regularization has been performed after the template fit and is shown as the blue uncertainty bands. In the panel below each distribution the ratio between the fitted value and the expected one is shown, together with the band of the theory uncertainties, highlighting the two sources. The ratio between the post-fit regularization and the template fit results is shown. The BBB uncertainty has been not considered.*

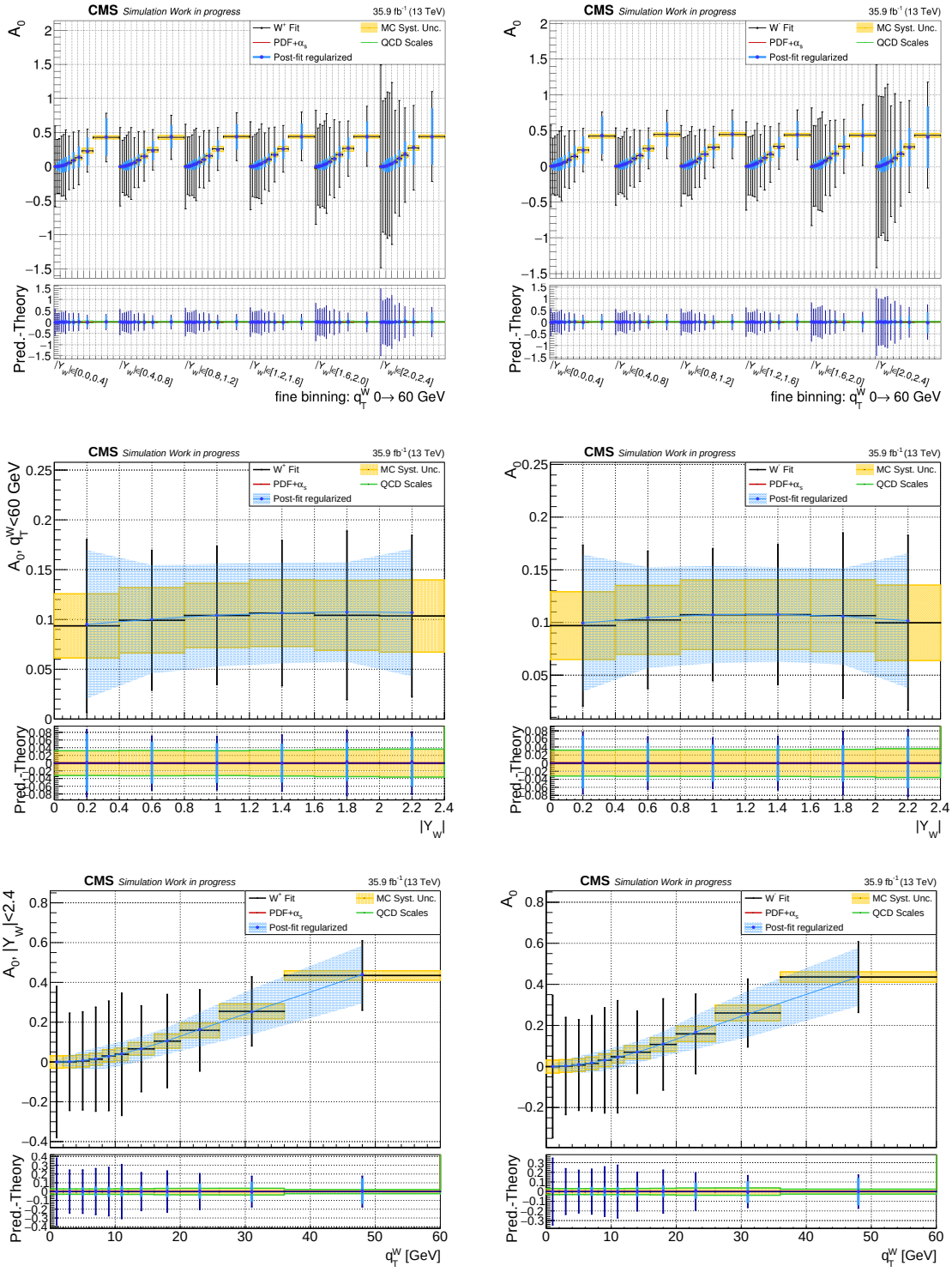


Figure E.9. *Fitted A_0, double differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central), differential in q_T^W (lower) for W^+ (left) and W^- (right) samples. The fit has been performed on the Asimov dataset. The yellow shaded area corresponds to the expected theory uncertainty, summing in quadrature the PDF and the MC scale contributions. The post-fit regularization has been performed after the template fit and is shown as a blue band in the single-differential case or only with the uncertainty bands in the double-differential case. In the panel below each distribution the difference between the fitted value and the expected one is shown, together with the band of the theory uncertainties, highlighting the two sources. The fit uncertainties in this panel are removed in the double differential result for clarity. The difference between the post-fit regularization and the template fit results is shown. The BBB uncertainty has been not considered.*

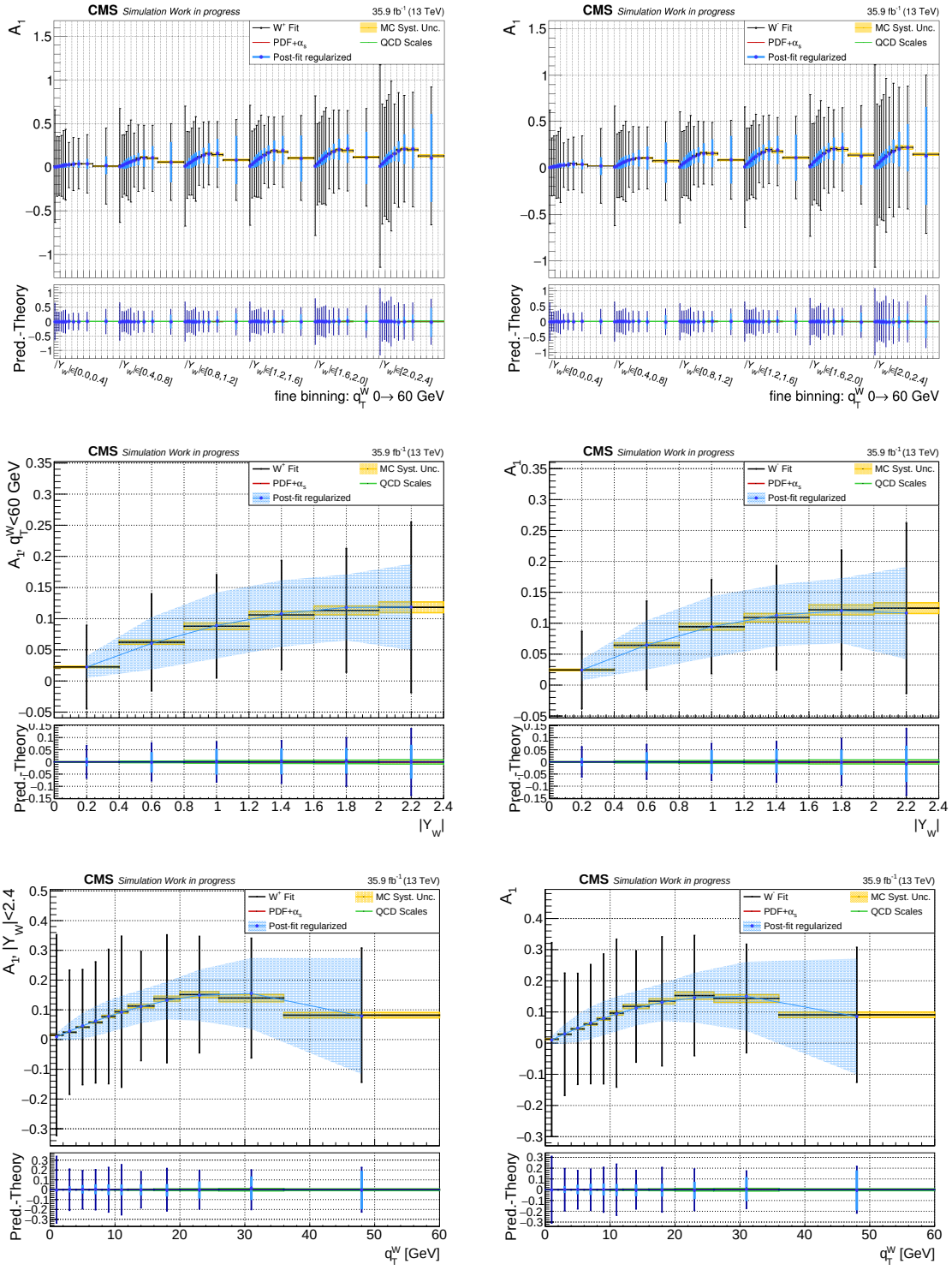


Figure E.10. *Fitted A_1, double differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central), differential in q_T^W (lower) for W^+ (left) and W^- (right) samples. The fit has been performed on the Asimov dataset. The yellow shaded area corresponds to the expected theory uncertainty, summing in quadrature the PDF and the MC scale contributions. The post-fit regularization has been performed after the template fit and is shown as a blue band in the single-differential case or only with the uncertainty bands in the double-differential case. In the panel below each distribution the difference between the fitted value and the expected one is shown, together with the band of the theory uncertainties, highlighting the two sources. The fit uncertainties in this panel are removed in the double differential result for clarity. The difference between the post-fit regularization and the template fit results is shown. The BBB uncertainty has been not considered.*

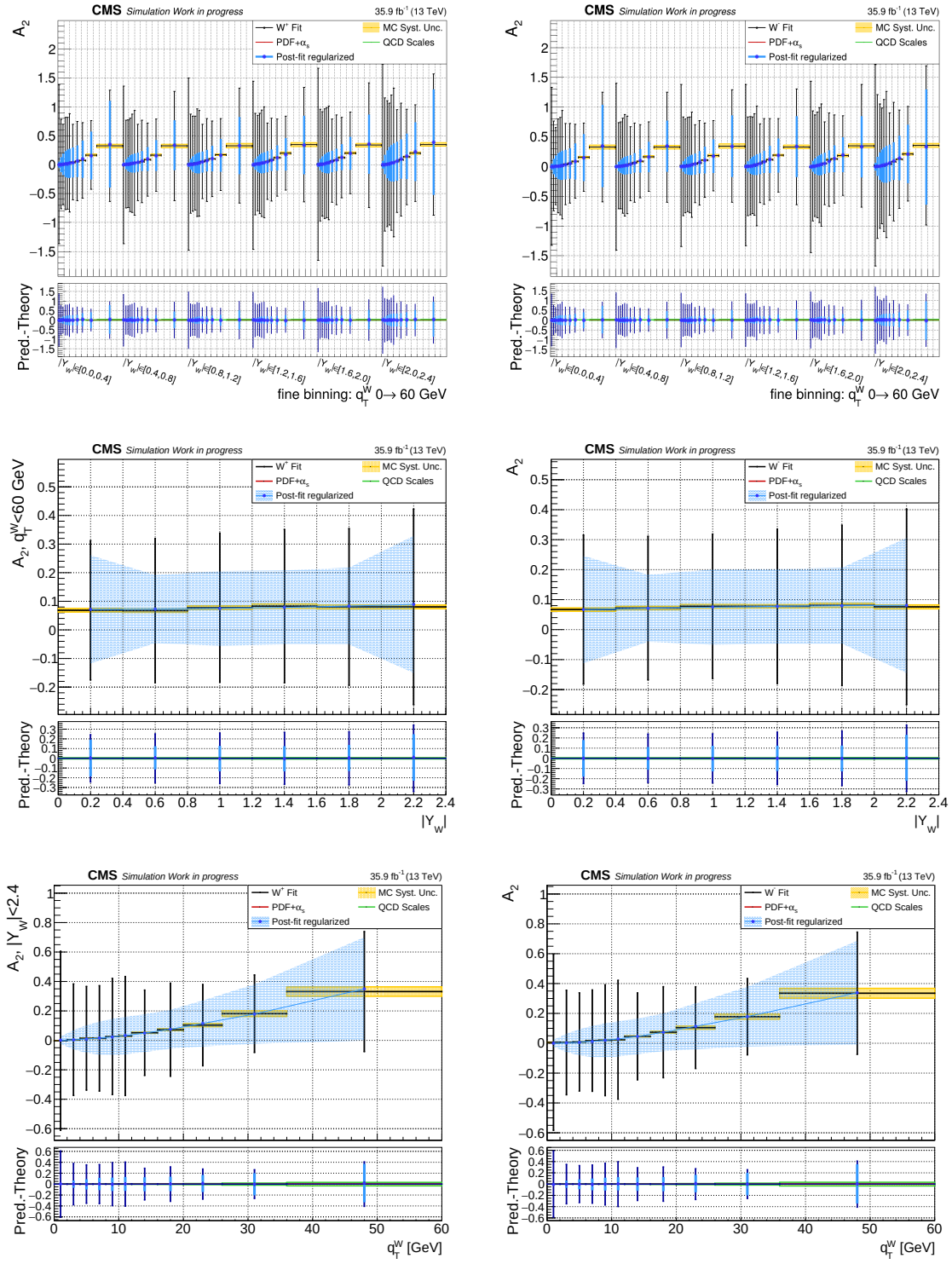


Figure E.11. *Fitted A_2, double differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central), differential in q_T^W (lower) for W^+ (left) and W^- (right) samples. The fit has been performed on the Asimov dataset. The yellow shaded area corresponds to the expected theory uncertainty, summing in quadrature the PDF and the MC scale contributions. The post-fit regularization has been performed after the template fit and is shown as a blue band in the single-differential case or only with the uncertainty bands in the double-differential case. In the panel below each distribution the difference between the fitted value and the expected one is shown, together with the band of the theory uncertainties, highlighting the two sources. The fit uncertainties in this panel are removed in the double differential result for clarity. The difference between the post-fit regularization and the template fit results is shown. The BBB uncertainty has been not considered.*

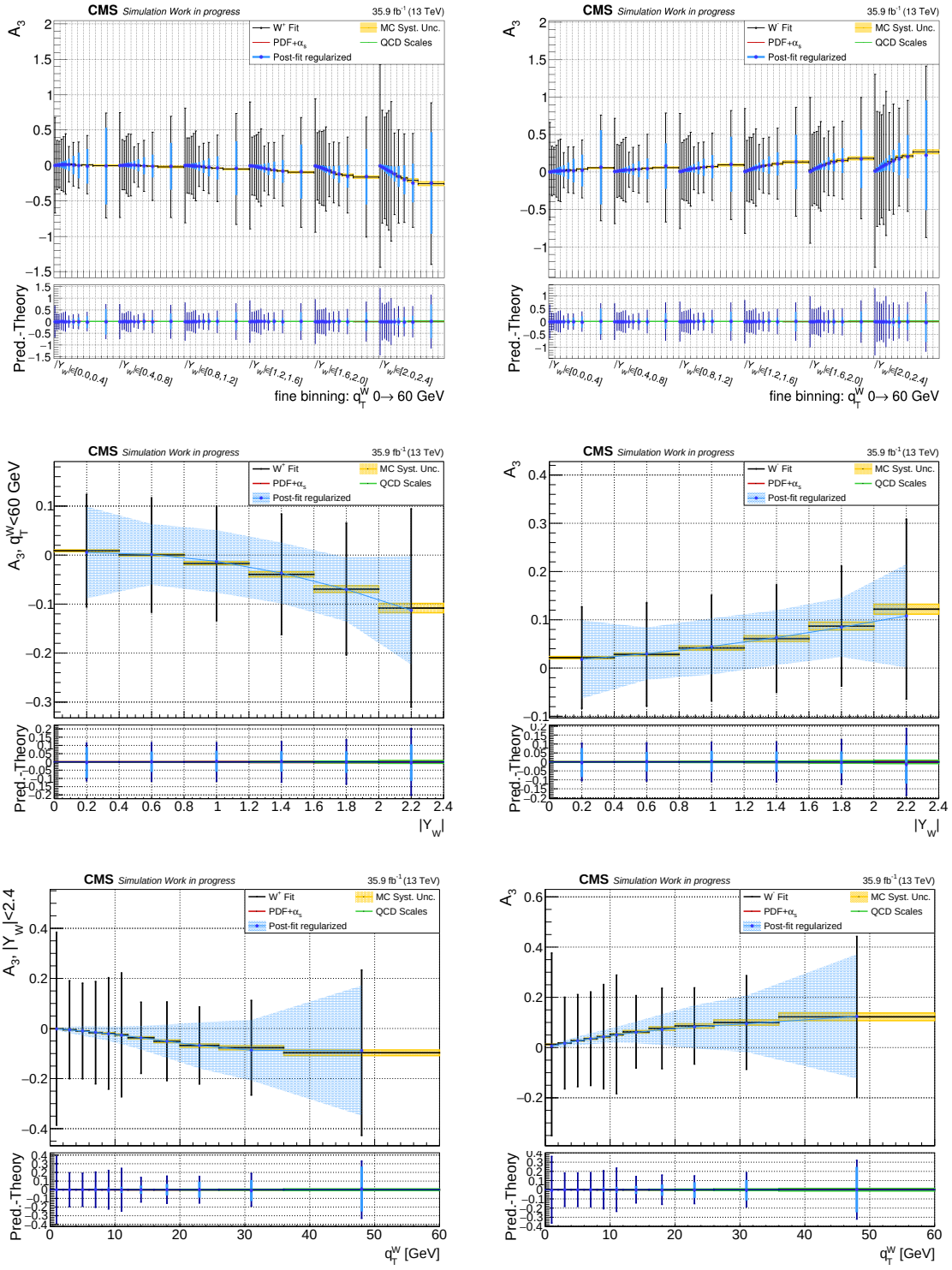


Figure E.12. *Fitted A_3, double differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central), differential in q_T^W (lower) for W^+ (left) and W^- (right) samples. The fit has been performed on the Asimov dataset. The yellow shaded area corresponds to the expected theory uncertainty, summing in quadrature the PDF and the MC scale contributions. The post-fit regularization has been performed after the template fit and is shown as a blue band in the single-differential case or only with the uncertainty bands in the double-differential case. In the panel below each distribution the difference between the fitted value and the expected one is shown, together with the band of the theory uncertainties, highlighting the two sources. The fit uncertainties in this panel are removed in the double differential result for clarity. The difference between the post-fit regularization and the template fit results is shown. The BBB uncertainty has been not considered.*

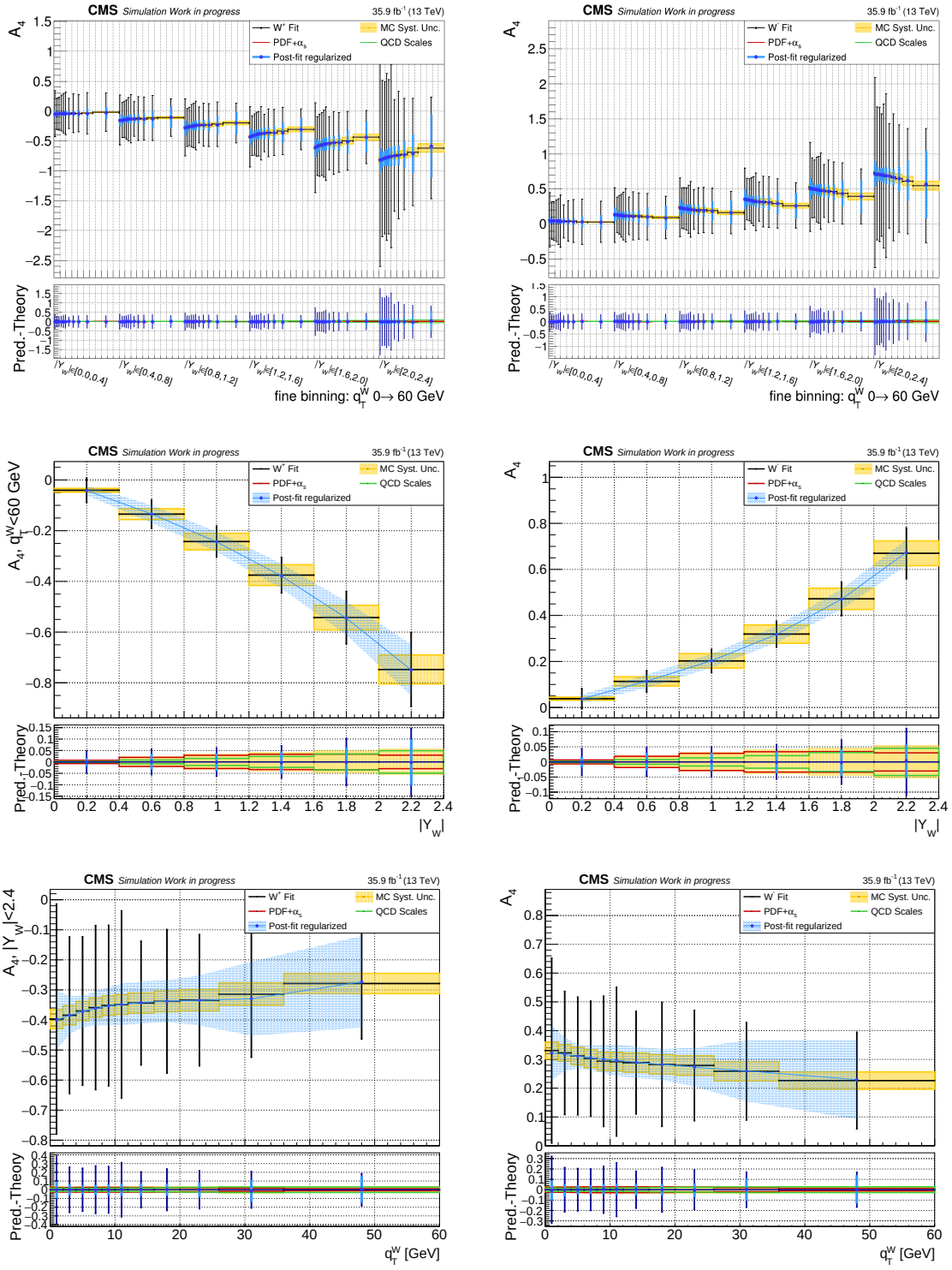


Figure E.13. *Fitted A_4, double differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central), differential in q_T^W (lower) for W^+ (left) and W^- (right) samples. The fit has been performed on the Asimov dataset. The yellow shaded area corresponds to the expected theory uncertainty, summing in quadrature the PDF and the MC scale contributions. The post-fit regularization has been performed after the template fit and is shown as a blue band in the single-differential case or only with the uncertainty bands in the double-differential case. In the panel below each distribution the difference between the fitted value and the expected one is shown, together with the band of the theory uncertainties, highlighting the two sources. The fit uncertainties in this panel are removed in the double differential result for clarity. The difference between the post-fit regularization and the template fit results is shown. The BBB uncertainty has been not considered.*

Appendix F

Simultaneous Gaussian Constraint regularization

In this appendix an alternative regularization scheme is presented. This implementation is currently under development and its performance are currently suboptimal.

In this approach, the regularization is performed simultaneously to the template fit. The likelihood function is modified to damp the high-order degrees of freedom with the following prescription. Given the finite $|Y_W|$ and q_T^W binning of the analysis, the $n_Y \times n_{q_T}$-degree polynomial which passes from all the bin centers always exists, where n_Y and n_{q_T} are the number of bins in the two variables. If some specific coefficient of this hypothetical polynomial will be constrained to zero, the respective order will be damped in the angular coefficient distributions. This has been implemented in the likelihood adding to L (Eq. 7.1) the term

$$L_\tau = \tau \sum_{i=0}^{4} \sum_{j=0}^{n_Y \times n_{q_T}} c_{i,j}^2 K_{i,j}, \tag{F.1}$$

where i runs over the angular coefficients $A_0 \dots A_4$, and j runs over the coefficients of the respective maximum-order polynomial (i.e. from 0 to $n_Y \times n_{q_T}$), τ is the strength of the Gaussian constraint, which is a single parameter for all the regularization scheme, $c_{i,j}$ are the polynomial coefficients for A_i of polynomial order j, and $K_{i,j}$ can be 0 or 1, and represent a mask that defines which degrees must be constrained to 0 ($K_{i,j} = 1$) or left free ($K_{i,j} = 1$). The assumption of smooth behaviour is valid for the angular coefficients only, thus σ^{U+L} has not been regularized but it is used to internally build the angular coefficient from the signal strength and the masked channels. The mask $K_{i,j}$ presented here has an opposite meaning compared to the $K_{i,j}$ of the post-fit linear fit regularization (Eq. 7.7): in the simultaneous Gaussian constraint it selects the degrees which must be damped, in the post-fit linear fit it selects the degrees which must be considered.

Compared to post-fit linear fit regularization, the two approaches have different advantages. The Gaussian constraint reflects more conservative assumptions, since is not imposing an exact polynomial behaviour to the distributions of the angular coefficients, but allow to modulate the strength of this assumption with τ parameter. In this framework, the polynomial linear fit corresponds to $\tau \to \infty$. On the other hand, the post-fit linear fit is a simpler and more robust procedure, which allow a more immediate interpretation.

F.1 Optimization of the parametrization

The optimal choice of the $K_{i,j}$ mask, which defined which degrees of freedom will be allowed and which constrained, requires a dedicated study. The MC accuracy in the angular coefficient $|Y_W|$ and q_T^W distribution is sufficient to discriminate between the possible polynomial parametrization. Thus, this study has been performed on $W \to \mu\nu$ MC sample.

The $A_i(Y_W, q_T^W)$ distribution has been constructed with the same range adopted in the template fit, but with a finer, equal-size binning in q_T^W i.e. 6 equal size bins in $|Y_W|$ and 30 equal-size bins in q_T^W. Then, each distribution has been fitted independently with different test models:

$$p_{a,b}(|Y_W|, q_T^W) = \left(\sum_{y=0}^{a} c_y |Y_W|^y\right) \times \left(\sum_{q=0}^{b} c_q (q_T^W)^q\right), \tag{F.2}$$

for the maximum degrees (a, b) spanning all the combinations from $\{a = 0, b = 0\}$ to $\{a = n_Y, b = n_{q_T}\}$. The fits have been performed independently for each angular coefficient. Some specific constraints has been set for all the tested fit models, according to the symmetries discussed in Sec. 5.2.3:

- The coefficient $c_{q=0}$ is set to 0 for A_0, A_1, A_2 and A_3 fits.
- The coefficient $c_{y=0}$ is set to 0 for A_1 and A_4 fits.

The χ^2 of the fit cannot be used directly to choose the optimal model for each angular coefficient. Excluding pathological cases, the χ^2 continue to decrease, with the increase of the complexity of the model, to the limit of $\chi^2 = 0$ for $\{a = n_Y, b = n_{q_T}\}$ fit. However, this improvement can be not statistically significant. Therefore the optimal fit model has been chosen as the most simple polynomial model whose χ^2 will not be significantly improved increasing the complexity of the model.

To decide if a certain model describes the angular coefficient distribution significantly better than another, the F-test has been used, which is briefly described here. Consider two models with a number of parameters p and q, respectively. The two models are nested, meaning that $p < q$ and the second model is an extension of the first. The two models are used to describe a distribution of n elements. The fit to the n elements with the two models produces the chi-squares χ_p^2, χ_q^2. The test statistic can be built as:

$$F = \frac{(\chi_p^2 - \chi_q^2)/(q-p)}{\chi_q^2/(n-q)}. \tag{F.3}$$

The null hypothesis H_0 is that the model with q parameters does not improve the description compared to the model with p parameters. If H_0 is true, the test statistic F follows a F-distribution $F(q-p, n-q)$, with $q-p$ and $n-q$ degrees of freedom. Thus, a p-value is built from the cumulative distribution of $F(q-p, n-q)$ as $p = 1 - \mathrm{CDF}(F_{\mathrm{obs}})$. If $p < \alpha$, H_0 is rejected.

In the performed tests $\alpha = 0.05$ has been chosen. Considering the class of model from Eq. F.2, the hypothesis of nested models is not valid for all the a, b combinations, because of the two possible dimensions of increasing complexity. Therefore a search algorithm has been developed which starts from $\{a = 1, b = 1\}$ and then tests all the possible increased complexity model up to the maximum complexity considered model. If the null hypothesis is rejected for a certain combination, the algorithm moves to this model as baseline, and restart the comparison with more complex models. The search proceeds in multiple search-paths, since the models are not nested. Each path stops when the null hypothesis is accepted for all the available increased

complexity models. When multiple ending-points are found, from different search-paths they are directly compared with the F-test. This is always possible, except in the case of the same number of degrees of freedom between the two models. In this case, the model with a lower q_T maximum degree is chosen (this use-case has never appeared in the performed search). For computation reason, the maximum-complexity model tested has been $\{a = 6, b = 7\}$. However, it has been never reached by the search algorithm.

The chosen models from the search algorithm are shown in table F.1. The robustness of the optimal models has been tested repeating the search on variations of the simulated angular coefficient distributions. The PDF and MC scales variations have been applied, according to the usual theory systematic uncertainties considered in the analysis. The chosen models have been confirmed. However, the fit parameter are sensitive to the applied variation. This is relevant to ensure that the sensitivity of the angular coefficient distributions on the theory uncertainties, and thus their capability to constraint the correspondent nuisance parameters, is not spoiled.

The reduced χ^2 for all the tested a, b combinations is shown in the left column of Fig F.1, for the five angular coefficients. The optimal models are highlighted with a red box. The trend of the reduced χ^2 on the $|Y_W| \times q_T^W$ plane and the selected models well reflect the decision to identify the simplest model whose χ^2 does not improve significantly increasing the complexity. In the central column of Fig F.1 the pull distributions between the optimal polynomial fit and the MC distribution of the five angular coefficients are shown. These pull distributions allow to see that the agreement is good in the entire plane and no local bias is induced. Summing all the bins of the $|Y_W| \times q_T^W$ plane, the mean and the variance of the pull distribution can be obtained. The former must be compatible with 0 and the latter with 1 for a good pull distribution. The two values, extracted with a Gaussian fit to the pull distribution, are also reported on the plots and show acceptable agreement.

However in the post-fit linear fit implementation, the optimal models have not been used (for technical reasons discussed in Sec. 7.5.1), but more simplified models have been adopted. This set of models is highlighted in the left column of Fig. F.1 with a green box, for each angular coefficient. The respective pull distributions are shown in the right column of Fig. F.1. In this case, the agreement is worst compared to the optimal model, revealing the sub-optimal choice. The implications of these features will be discussed in Sec. 7.5.1.

Table F.1. *Optimal model for angular coefficient post-fit polynomial regularization, from the search algortihm based on the F-test.*

	W^+		W^-	
	max $\|Y_W\|$ deg.	max q_T^W deg.	max $\|Y_W\|$ deg.	max q_T^W deg.
A_0	2	3	2	4
A_1	2	5	2	5
A_2	1	4	1	3
A_3	2	4	2	4
A_4	3	5	3	4

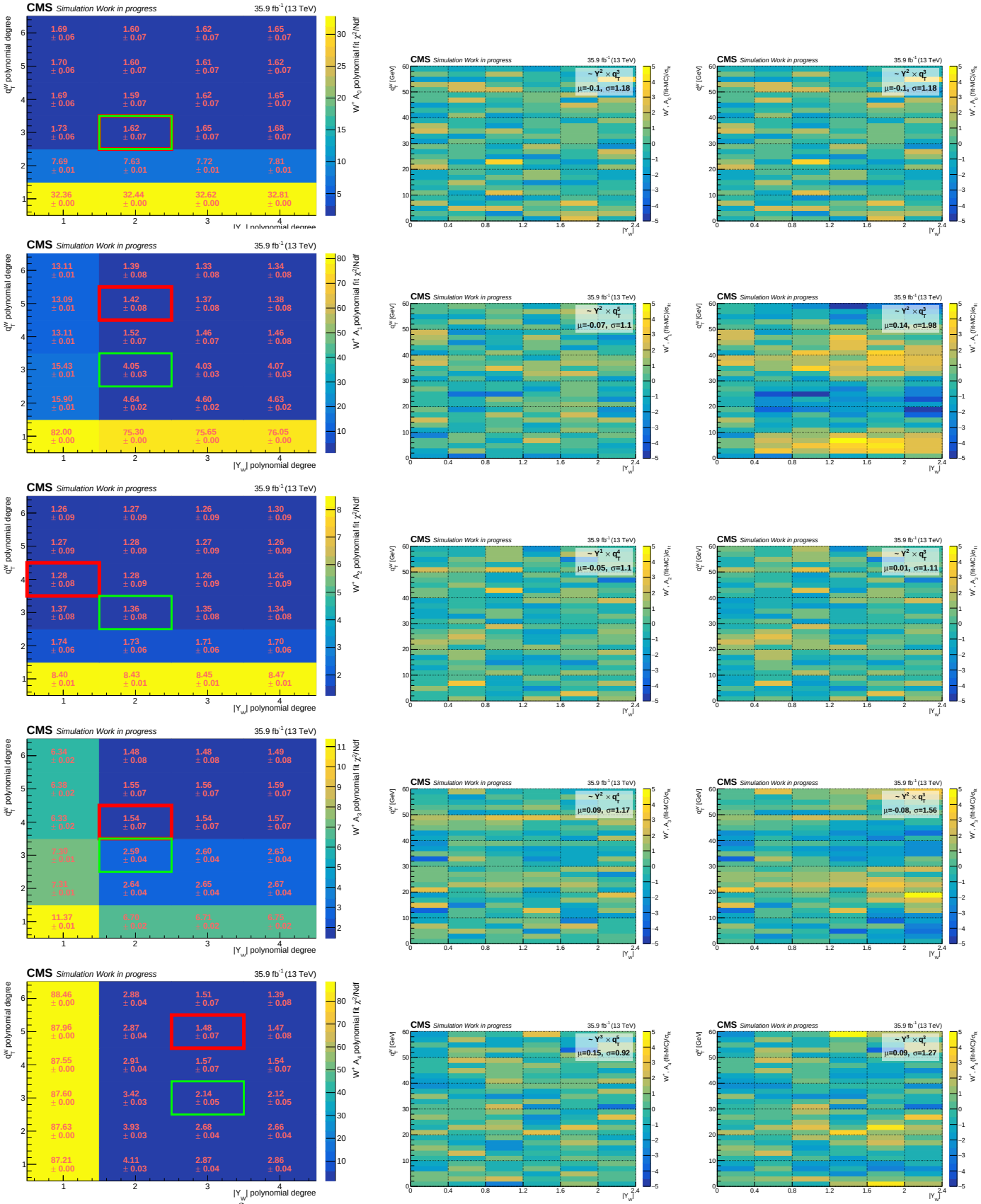


Figure F.1. *(left) χ^2/Ndf of the fit to $W \to \mu\nu$ MC distributions of A_0 (upper),... A_4 (lower), for different polynomial models. The optimal (suboptimal) model is highlighted with a red (green) box and the central (right) column contains the respective pulls between the fit and the MC. The mean and variance of a Gaussian fit to the pull distribution is shown.*

F.2 Preliminary results

The Simultaneous Gaussian constraint regularization is performed adding to the template fit likelihood the terms of Eq. F.1. The $K_{i,j}$ mask has been defined according to optimal models of table F.1. The global regularization strength parameter τ has been set to 100. This value is a trade off between a robust but not regularized fit ($\tau \to 0$) and a less robust minimization but properly regularized ($\tau \to \infty$). A very high τ makes the minimization less robust because forces the likelihood to the polynomial model increasing the complexity of the simultaneous determination of POI and nuisance parameters. Since the simultaneous Gaussian constraint regularization is still an under development tool, there is large room for improvement in the optimization of the τ parameter.

The fit has been performed on the Asimov dataset. The central values are not biased by the regularization and are fully consistent with the expected values (with a relative discrepancy below 10^{-6}). Since the results are graphically very similar to the already presented Figs. 7.6 to 7.11 they are not shown.

However, is interesting to discuss the predicted uncertainties in the not-regularized case compared to the Simultaneous Gaussian constraint regularized fit and the post-fit regularized fit. This comparison is presented in Fig. F.2 for a selected σ^{U+L} and A_4. The post-fit linear fit produces a stronger reduction of the uncertainty in the entire phase space. The Gaussian constraint regularization produces a small uncertainty reduction only in where the uncertainty is larger and in particular in the low-q_T^W region, both in double and single-differential case. This behaviour reflects the hypotheses of the two implementations. The post-fit linear fit is assuming a strict fixed-order polynomial behaviour for the angular coefficients. The simultaneous Gaussian constraint is only damping high-order degrees when the simultaneous determination of the POIs take advantage of that (in terms of negative log-likelihood minimization). The simultaneous Gaussian constraint has the purpose to reduce the uncertainty inflated by the correlation scheme of the template fit. Therefore, it is exploiting a minimal constraint to the template fit to optimally perform the unfolding, avoiding to assume additional hypotheses. On the other hand, a strict polynomial description is independent of the correlation pattern issue, and actually represents a physical hypothesis on the described distributions. This hypothesis should be considered in presence of bias of the central values, as discussed for the post-fit linear fit regularization results.

The simultaneous Gaussian constraint regularization is an approach currently under development and optimization, and improvement can be expected in the future, compared to the presented result. However, it already demonstrates the capability to perform this kind of simultaneous regularization with some small advantages in term of uncertainties reduction, without the strict polynomial assumption.

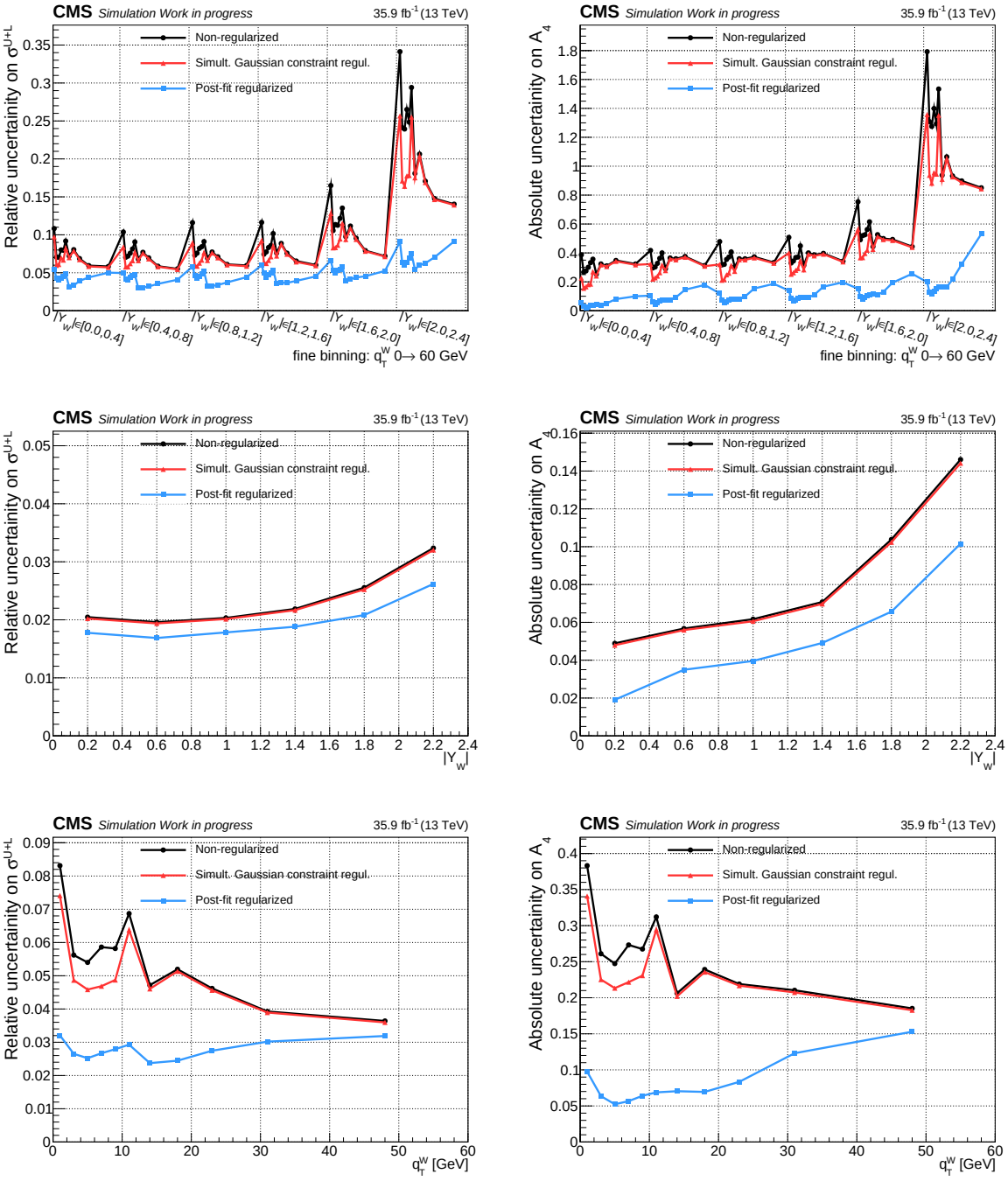


Figure F.2. *Relative uncertainty from the fit on σ^{U+L} (left) and absolute uncertainty on A_4 (right) in double-differential in $|Y_W|$ and q_T^W (upper), differential in $|Y_W|$ (central) and differential in q_T^W (lower) cases. The template fit has been performed on the Asimov dataset in all the cases. The uncertainty have been estimated directly from* `Combine-TensorFlow` *fit in the non-regularized case and in the simultaneous Gaussian constraint regularized case. The uncertainty of the Post-fit regularized case has been estimated with the* `JAX` *fit on the top of the* `Combine-TensorFlow` *non-regularized fit. The BBB uncertainty has been not considered.*

Acknowledgements

Italian version

Tre anni sono lunghi, e dovrei quindi ringraziare molte persone che direttamente o indirettamente hanno contribuito alla mia cresita, umana e professionale, in questo periodo. Mi limiterò però unicamente alle persone direttamente coinvolte nel mio lavoro di tesi, per poter essere conciso.

In primo luogo ringrazio profondamente il mio relatore, Gigi Rolandi, per avermi fatto da guida in questi anni di ricerca nei momenti belli e in quelli difficili. Quando ho iniziato questo dottorato mi sono approacciato a un nuovo esperimento e a un nuovo tipo di fisica. Sono davvero felice di aver potuto avere come riferimento una persona che ha sempre trovato il tempo, e la voglia, di fare lunghe o brevi chiacchierate di fisica, riuscirendo ad essere una miniera di spunti e idee su come affrontare i problemi in modo diverso. Gli sono profondamente grato di aver sempre avuto la pazienza di ascoltare i miei dubbi e trovare le parole giuste per insegnarmi qualcosa di più. Spesso si finisce a concentrarsi troppo sul singolo risultato da raggiungere, e ci si dimentica di tutta la galassia di dettagli che lo circondano. Gigi in questi anni è stata la persona che mi faceva allargare lo sguardo, puntandomi a quell'aspetto o quell'insidia che nascondevano la giusta chiave di lettura. Questo spesso mi ha fatto spesso pensare che sono quei dettagli il motivo per cui è così divertente fare questo lavoro.

Il secondo ringraziamento va ai membri giovani del gruppo W-mass Pisa, Suvankar Roy Chowdhury e Elisabetta Manca. Abbiamo lavorato insieme notte e giorno, in lockdown chiusi in casa o in giro per il mondo, fianco a fianco nell'ufficio 115 o sparsi dalla Sardegna all'India. Arrabbiati, allegri, stufi o soddisfatti. E' stata tosta, ma è stato bello plasmare questa analisi insieme a voi.

Devo poi chiudere il cerchio del gruppo W-mass Pisa ringraziando Lorenzo Bianchini, per tutti i pazienti suggerimenti datemi negli anni e in particolare per avermi seguito nella pubblicazione del mio primo articolo. L'ultimo e particolarmente sentito ringraziamento legato a W-mass va a Josh Bendavid, che con un occhio pazzesco è riuscito, specialmente in questi ultimi mesi, a darmi una mano a risolvere giorno dopo giorno gli infiniti problemi incontrati in questa analisi.

Vorrei poi ringraziare le persone che mi hanno permesso di proseguire il mio percorso come sviluppatore di software di tracking. Il primo è Andrea Rizzi, che mi ha introdotto al codice di CMS e alle reti neurali, è ha sempre saputo vedere la strada giusta dove lanciarmi. La seconda è Mia Tosi, che ha sempre supportato e spronato le mie attività nel tracking POG di CMS. Entrambi li voglio ringraziare anche per gli spritz o i bicchieri di vino che ho potuto condividere con loro, tra una cena sociale o un aperitivo dopo qualche meeting.

Vorrei inoltre ringraziare gli altri membri di CMS-Pisa con cui ho preso infiniti caffè alle macchinette dell'INFN, Leonardo Giannini, Giulio Mandorli, Shubhi Parolia, Negin Shafiei, Tommaso Boccali, Silvio Donato e Paolo Azzurri. Quest'ultimo lo vorrei anche ringraziare per le lunghe chiacchierate, tra fisica, aneddoti, pettegolezzi e consigli, fatte tra a quelle macchinette e i nostri uffici.

Un ultimo ringraziamento vorrei farlo a Francesco Forti che, nonostante il dottorato mi abbia portato distante da Belle II, ha sempre trovato il tempo per fare due chiacchiere con me, di fisica o di altro, o darmi una dritta al momento giusto.

English version

Tree years is a long time, and I have to thank several people which directly or indirectly contributed to my personal and professional growth. However I will just acknowledge the people directly involved in my thesis work, allowing me to be concise.

First of all, I sincerely thank my supervisor, Gigi Rolandi, to guides me in these years of research, in the good moments and in the difficult ones. When I started this PhD I joined a new experiment and I was introduced to a new kind of physics. I was very lucky to have as a supervisor a person which always finds the time, and the will, to have long or short discussions about physics, being a great source of ideas to face the issues with a different approach. I am sincerely grateful to him to always find the patience to listen to my doubts and find the best words to teach me something more. Often we focus too much to reach the single result, and we forget about the entire universe of details which surround the result. Gigi, in these years, has been the person who helped me to broaden my vision, point me to this aspect or that issue which hides the best interpretation. This often teaches me that are these details which make this work so interesting.

The second thank must go to the young members of the W-mass Pisa group, Suvankar Roy Chowdhury and Elisabetta Manca. We worked together night and day, in lockdown or around the world, side by side in office 115 o spread between Sardinia and India. We were angry, happy, annoyed or satisfied. It was tough, but it was great to build this analysis with you.

I have to complete of W-mass Pisa group by thanking Lorenzo Bianchini, for all the patient suggestions of these years and in particular to have followed me in the publication of my first journal paper. The last heartfelt thanks related to W-mass is for Josh Bendavid, that helped me with a very keen eye to solve the never-ending issues encountered in this analysis, especially in these last months.

I would like also to thank the people which allow me to follow my record as tracking software developer. The first is Andrea Rizzi, which introduced me to the CMS code and to the neutral networks, and has always thrown me towards the the best way to proceed. The second is Mia Tosi, which always supported me and encouraged my activities in CMS tracking POG. I would like also to thank them both also for the spritz or the glasses of wine which I shared with them, between a social dinner or a drink after a meeting.

I would like to thank also the other members of CMS-Pisa group with which I shared countless coffee break at the coffee machine of INFN, Leonardo Giannini, Giulio Mandorli, Shubhi Parolia, Negin Shafiei,Tommaso Boccali, Silvio Donato and Paolo Azzurri. A great thanks to the latter also for the long chats, between physics, stories, rumours and suggestions, had in front of the coffee machine.

A last thank must go to Francesco Forti, which always found the time to talk with me about physics or about something else, to give me a tip at the right moment, despite my PhD led me far from Belle II.

Bibliography

[1] CMS Collaboration, "The CMS Experiment at the CERN LHC", JINST **3**, S08004 (2008) 10.1088/1748-0221/3/08/S08004.

[2] O. S. Brüning et al., *LHC Design Report*, CERN Yellow Reports: Monographs (CERN, Geneva, 2004), http://cds.cern.ch/record/782076.

[3] ATLAS Collaboration, "Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", Eur. Phys. J. C **78**, 110 (2018) 10.1140/epjc/s10052-017-5475-4.

[4] CDF Collaboration, "Precise measurement of the W-boson mass with the Collider Detector at Fermilab", Phys. Rev. D **89**, 072003 (2014) 10.1103/PhysRevD.89.072003.

[5] D0 Collaboration, "Measurement of the W boson mass with the D0 detector", Phys. Rev. D **89**, 012005 (2014) 10.1103/PhysRevD.89.012005.

[6] V. Bertacchi et al., "Impact of the PDFs on the Z and W lineshapes at LHC", Eur. Phys. J. C **80** (2020) 10.1140/epjc/s10052-020-7892-z.

[7] W. J. Stirling and A. D. Martin, "Parton distribution uncertainty in the measurement of M_W in proton-antiproton collisions", Physics Letters B **237**, 551–558 (1990) 10.1016/0370-2693(90)91223-X.

[8] E. Manca et al., "About the rapidity and helicity distributions of the W bosons produced at LHC", J. High Energy Phys. **2017**, 130 (2017) 10.1007/JHEP12(2017)130.

[9] N. Foppiani, "Definition and calibration of the hadronic recoil in view of the measurement of the W boson mass with the CMS experiment", MA thesis (Pisa U., 2017), https://cds.cern.ch/record/2281312.

[10] O. Cerri, "Hadronic recoil in the W boson production at LHC for a W mass measurement with the CMS experiment", MA thesis (Pisa U., 2017), https://cds.cern.ch/record/2285935.

[11] CMS Collaboration, "Measurements of the W boson rapidity, helicity, double-differential cross sections, and charge asymmetry in pp collisions at $\sqrt{s} = 13$ TeV", Phys. Rev. D **102** (2020) 10.1103/physrevd.102.092012.

[12] E. Mirkes, "Angular decay distribution of leptons from W-bosons at NLO in hadronic collisions", Nucl. Phys. B **387**, 3–85 (1992) 10.1016/0550-3213(92)90046-E.

[13] E. Guiraud et al., *TDataFrame: functional chains for ROOT data analyses*, 2017, 10.5281/zenodo.260230.

[14] CMS Collaboration, "A further reduction in CMS event data for analysis: the NANOAOD format", EPJ Web of Conferences **214**, 06021 (2019) 10.1051/epjconf/201921406021.

[15] CMS Collaboration, *High pt jets tracking*, (2014) https://twiki.cern.ch/twiki/bin/view/CMSPublic/HighPtTrackingDP.

[16] V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning", (2016), arXiv:1603.07285.

[17] V. Bertacchi (CMS Collaboration), *DeepCore: Convolutional Neural Network for high p_T jet tracking*, PROC-CTD19-020, 2019, arXiv:1910.08058 [physics.ins-det].

[18] Particle Data Group, "Review of Particle Physics", PTEP **2020**, 083C01 (2020) 10.1093/ptep/ptaa104.

[19] L. Maiani and L. Rolandi, *The standard theory of particle physics: essays to celebrate CERN's 60th anniversary*, Advanced series on directions in high energy physics (World Scientific, Hackensack, NJ, 2016), 10.1142/9878.

[20] C. W. Fabjan et al., *Particle physics reference library* (Springer, Cham, 2020), 10.1007/978-3-030-38207-0.

[21] Cush, Public domain, via Wikimedia Commons, *Standard Model of Elementary Particles*, https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.

[22] Eric Drexler, Public domain, via Wikimedia Commons, *Elementary particle interactions in the Standard Model*, https://commons.wikimedia.org/wiki/File:Elementary_particle_interactions_in_the_Standard_Model.png.

[23] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", Physics Letters B **716**, 30–61 (2012) http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[24] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", Physics Letters B **716**, 1–29 (2012) http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[25] D. Buttazzo et al., "Investigating the near-criticality of the Higgs boson", J. High Energy Phys. **2013** (2013) 10.1007/jhep12(2013)089.

[26] G. F. Giudice, *Naturalness after LHC8*, 2013, arXiv:1307.7879 [hep-ph].

[27] BaBar Collaboration and Belle Collaboration, "The Physics of the B Factories", Eur. Phys. J. C **74**, 3026 (2014) 10.1140/epjc/s10052-014-3026-9.

[28] CMS Collaboration, "Evidence for Higgs boson decay to a pair of muons", J. High Energy Phys. **01**, 148 (2021) 10.1007/jhep01(2021)148.

[29] Fermilab, *Design Report Tevatron 1 project*, tech. rep. FERMILAB-DESIGN-1984-01 (1984), https://cds.cern.ch/record/1478620.

[30] CERN, *LEP design report*, Copies shelved as reports in LEP, PS and SPS libraries (CERN, Geneva, 1984), https://cds.cern.ch/record/102083.

[31] ALEPH Collaboration, CDF Collaboration, D0 Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, SLD Collaboration, LEP Electroweak Working Group, Tevatron Electroweak Working Group, SLD electroweak and heavy flavour groups, *Precision Electroweak Measurements and Constraints on the Standard Model*, tech. rep. (2010), arXiv:1012.2367, http://cds.cern.ch/record/1313716.

[32] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, The LEP Working Group for Higgs Boson Searches, "Search for the Standard Model Higgs boson at LEP", Physics Letters B **565**, 61–75 (2003) https://doi.org/10.1016/S0370-2693(03)00614-2.

[33] CDF Collaboration and D0 Collaboration, "Combination of Tevatron Searches for the Standard Model Higgs Boson in the W^+W^- Decay Mode", Phys. Rev. Lett. **104**, 061802 (2010) 10.1103/PhysRevLett.104.061802.

[34] CDF Collaboration, "Combined Search for the Standard Model Higgs Boson Decaying to a $b\bar{b}$ Pair Using the Full CDF Data Set", Phys. Rev. Lett. **109** (2012) 10.1103/physrevlett.109.111802.

[35] CDF Collaboration and D0 Collaboration, "Evidence for a Particle Produced in Association with Weak Bosons and Decaying to a Bottom-Antibottom Quark Pair in Higgs Boson Searches at the Tevatron", Phys. Rev. Lett. **109**, 071804 (2012) 10.1103/PhysRevLett.109.071804.

[36] D0 Collaboration, "Combined Search for the Standard Model Higgs Boson Decaying to $b\bar{b}$ Using the D0 Run II Data Set", Phys. Rev. Lett. **109**, 121802 (2012) 10.1103/PhysRevLett.109.121802.

[37] E. Graverini, "Flavour anomalies: a review", J. Phys. Conf. Ser. **1137**, 012025 (2019) 10.1088/1742-6596/1137/1/012025.

[38] GFitter group, *Gfitter. A Generic Fitter Project for HEP Model Testing*, http://project-gfitter.web.cern.ch/project-gfitter/.

[39] HEPfit group, *HEPfit*, https://hepfit.roma1.infn.it/index.html.

[40] J. Haller et al., "Update of the global electroweak fit and constraints on two-Higgs-doublet models", Eur. Phys. J. C **78** (2018) 10.1140/epjc/s10052-018-6131-3.

[41] M. Baak et al., "The global electroweak fit at NNLO and prospects for the LHC and ILC", Eur. Phys. J. C **74** (2014) 10.1140/epjc/s10052-014-3046-5.

[42] M. Baak et al., "Updated status of the global electroweak fit and constraints on new physics", Eur. Phys. J. C **72** (2012) 10.1140/epjc/s10052-012-2003-4.

[43] M. Baak et al., "The electroweak fit of the standard model after the discovery of a new boson at the LHC", Eur. Phys. J. C **72** (2012) 10.1140/epjc/s10052-012-2205-9.

[44] H. Flächer et al., "Revisiting the global electroweak fit of the Standard Model and beyond with Gfitter", Eur. Phys. J. C **60**, 543–583 (2009) 10.1140/epjc/s10052-009-0966-6.

[45] M. Awramik et al., "Precise prediction for the W-boson mass in the standard model", Phys. Rev. D **69** (2004) 10.1103/physrevd.69.053006.

[46] M. E. Peskin and T. Takeuchi, "Estimation of oblique electroweak corrections", Phys. Rev. D **46**, 381–409 (1992) 10.1103/PhysRevD.46.381.

[47] CERN Bulletin, *The particle suppliers. Les fournisseurs de particules*, BUL-NA-2010-077, 2010, https://cds.cern.ch/record/1255151.

[48] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", JINST **3**, S08003 (2008) 10.1088/1748-0221/3/08/S08003.

[49] ALICE Collaboration, "The ALICE experiment at the CERN LHC", JINST **3**, S08002 (2008) 10.1088/1748-0221/3/08/S08002.

[50] LHCb Collaboration, "The LHCb Detector at the LHC", JINST **3**, S08005 (2008) 10.1088/1748-0221/3/08/S08005.

[51] LHCf Collaboration, "The LHCf detector at the CERN Large Hadron Collider", JINST **3**, S08006 (2008) 10.1088/1748-0221/3/08/S08006.

[52] TOTEM Collaboration, "The TOTEM experiment at the CERN Large Hadron Collider", JINST **3**, S08007 (2008) 10.1088/1748-0221/3/08/S08007.

[53] J. Wenninger, "LHC status and performance", PoS **CHARGED2018**, 001. 9 p (2018) 10.22323/1.339.0001.

[54] CMS Collaboration, *CMS Luminosity - Public Results*, (2018) https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[55] R. Bruce et al., "Sources of machine-induced background in the ATLAS and CMS detectors at the CERN Large Hadron Collider", Nucl. Instrum. Methods Phys. Res., A **729**, 825–840. 22 p (2014) 10.1016/j.nima.2013.03.058.

[56] ATLAS Collaboration, *Beam-gas Background Observations at LHC*, tech. rep. ATL-DAPR-PROC-2017-002 (CERN, Geneva, 2017), 10.18429/JACoW-IPAC2017-TUPVA032.

[57] R. Assmann et al., "Equilibrium beam distribution and halo in the LHC", in 8th European Particle Accelerator Conference (EPAC 2002), CERN-LHC-PROJECT-REPORT-592 (), https://cds.cern.ch/record/569470.

[58] A. D. Martin et al., "Parton distributions for the LHC", Eur. Phys. J. C **63**, 189–285 (2009) 10.1140/epjc/s10052-009-1072-5.

[59] J. Stirling, *Parton Luminosity and Cross Section Plots*, http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html.

[60] CMS Collaboration and T. Mc Cauley, "Collision events recorded by CMS in 2016", 2017, https://cds.cern.ch/record/2241144.

[61] CMS Collaboration, *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*, Technical Design Report CMS (CERN, Geneva, 2006), https://cds.cern.ch/record/922757.

[62] R. García Alía et al., "LHC and HL-LHC: Present and Future Radiation Environment in the High-Luminosity Collision Points and RHA Implications", IEEE Transactions on Nuclear Science **65**, 448–456 (2018) 10.1109/TNS.2017.2776107.

[63] CMS Collaboration, "Cutaway diagrams of CMS detector", (2019), https://cds.cern.ch/record/2665537.

[64] CMS Collaboration, *The CMS tracker system project: Technical Design Report*, Technical Design Report CMS (CERN, Geneva, 1997), http://cds.cern.ch/record/368412.

[65] K. Arndt et al., "Silicon sensors development for the CMS pixel system", Nucl. Instrum. Methods Phys. Res. A **511**, Proceedings of the 11th International Workshop on Vertex Detectors, 106–111 (2003) https://doi.org/10.1016/S0168-9002(03)01773-X.

[66] G. Bolla et al., "Design and test of pixel sensors for the CMS experiment", Nucl. Instrum. Methods Phys. Res. A **461**, 8th Pisa Meeting on Advanced Detectors, 182–184 (2001) https://doi.org/10.1016/S0168-9002(00)01201-8.

[67] CMS Collaboration, *Sensor Design for the CMS Silicon Strip Tracker*, tech. rep. CMS-NOTE-2003-020 (CERN, Geneva, 2003), http://cds.cern.ch/record/687861.

[68] J.-L. Agram et al., "The silicon sensors for the Compact Muon Solenoid tracker—design and qualification procedure", Nucl. Instrum. Methods Phys. Res. A **517**, 77–93 (2004) https://doi.org/10.1016/j.nima.2003.08.175.

[69] CMS Collaboration, *Studies of Tracker Material*, tech. rep. CMS-PAS-TRK-10-003 (2010), http://cds.cern.ch/record/1279138.

[70] CMS Collaboration, *Tracker Material Budget plots*, (2020) https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackerMaterialBudgetplots.

[71] CMS Collaboration, *CMS Tracking POG Performance in Run-2 Legacy data*, (2020) https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackingPOGResultsRun2Legacy#Muon_tracking_performance_in_AN3.

[72] CMS Collaboration, *The Phase-1 Upgrade of the CMS Pixel Detector*, tech. rep. CMS-CR-2017-135. 06 (CERN, Geneva, 2017), 10.1088/1748-0221/12/07/C07009.

[73] CMS Collaboration, *The CMS electromagnetic calorimeter project: Technical Design Report*, Technical Design Report CMS (CERN, Geneva, 1997), https://cds.cern.ch/record/349375.

[74] K. W. Bell et al., "Vacuum phototriodes for the CMS electromagnetic calorimeter endcap", IEEE Transactions on Nuclear Science **51**, 2284–2287 (2004) 10.1109/TNS.2004.836053.

[75] E. Tournefier (CMS Collaboration), "The preshower detector of CMS at LHC", Nucl. Instrum. Methods Phys. Res. A **461**, 8th Pisa Meeting on Advanced Detectors, 355–360 (2001) https://doi.org/10.1016/S0168-9002(00)01243-2.

[76] CMS Collaboration, *The CMS ECAL performance with examples*, tech. rep. CMS-CR-2013-430 (CERN, Geneva, 2013), 10.1088/1748-0221/9/02/C02008.

[77] CMS Collaboration, *ECAL 2016 refined calibration and Run2 summary plots*, tech. rep. CMS-DP-2020-021 (2020), https://cds.cern.ch/record/2717925.

[78] CMS-HCAL Collaboration, "Design, performance, and calibration of CMS forward calorimeter wedges", Eur. Phys. J. C **53**, 139–166 (2008) 10.1140/epjc/s10052-007-0459-4.

[79] CMS Collaboration, *The CMS magnet project: Technical Design Report*, Technical Design Report CMS (CERN, Geneva, 1997), https://cds.cern.ch/record/331056.

[80] Vector Fileds Ltd. Oxfort U.K., *TOSCA/OPERA-3d Software*, www.vectorfields.com.

[81] CMS Collaboration, "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays", JINST **5**, T03021. 35 p (2009) 10.1088/1748-0221/5/03/T03021.

[82] CMS Collaboration, *The CMS muon project: Technical Design Report*, Technical Design Report CMS CERN-LHCC-97-032 (CERN, Geneva, 1997), https://cds.cern.ch/record/343814.

[83] D. Abbaneo et al., "Status report of the upgrade of the CMS muon system with Triple-GEM detectors", Nucl. Instrum. Methods Phys. Res. A **824**, Frontier Detectors for Frontier Physics: Proceedings of the 13th Pisa Meeting on Advanced Detectors, 521–525 (2016) https://doi.org/10.1016/j.nima.2015.11.125.

[84] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", JINST **13**, P06015. 53 p (2018) 10.1088/1748-0221/13/06/P06015.

[85] CMS Collaboration, *CMSSW*, http://cms-sw.github.io/.

[86] R. Fruehwirth et al., *Data Analysis Techniques for High-Energy Physics*, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology (Cambridge University Press, Cambridge, 2000).

[87] R. Fruhwirth and A. Strandlie, "Pattern recognition and estimation methods for track and vertex reconstruction", in 13th International Workshop on Advanced Computing and Analysis Techniques in Physics Research, PoS(ACAT2010)003 (2010), https://pos.sissa.it/093/003/.

[88] A. Strandlie and R. Fruhwirth, "Track and vertex reconstruction: From classical to adaptive methods", Rev. Mod. Phys. **82**, 1419 (2010) 10.1103/RevModPhys.82.1419.

[89] P. Billoir, "Progressive track recognition with a Kalman like fitting procedure", Comput. Phys. Commun. **57**, 390 (1989) 10.1016/0010-4655(89)90249-X.

[90] P. Billoir and S. Qian, "Simultaneous pattern recognition and track fitting by the Kalman filtering method", Nucl. Instrum. Meth. **A294**, 219 (1990) 10.1016/0168-9002(90)91835-Y.

[91] R. Mankel, "A Concurrent track evolution algorithm for pattern recognition in the HERA-B main tracking system", Nucl. Instrum. Meth. **A395**, 169 (1997) 10.1016/S0168-9002(97)00705-5.

[92] F. Pantaleo et al., "New Track Seeding Techniques for the CMS Experiment", PhD thesis (2017), http://cds.cern.ch/record/2293435.

[93] R. Frühwirth, "Application of Kalman filtering to track and vertex fitting", Nucl. Instrum. Meth. **A262**, 444 (1987) 10.1016/0168-9002(87)90887-4.

[94] J. R. Cash and A. H. Karp, "A Variable Order Runge-Kutta Method for Initial Value Problems with Rapidly Varying Right-hand Sides", ACM Trans. Math. Softw. **16**, 201 (1990) 10.1145/79505.79507.

[95] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", JINST **9**, P10009 (2014) 10.1088/1748-0221/9/10/P10009.

[96] CMS Collaboration, *2017 tracking performance plots*, tech. rep. CMS-DP-2017-015 (2017), https://cds.cern.ch/record/2290524.

[97] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems", Proceedings of the IEEE **86**, 2210–2239 (1998) 10.1109/5.726788.

[98] W. Waltenberger et al., "Adaptive vertex fitting", J. Phys. G **34**, N343–N356 (2007) 10.1088/0954-3899/34/12/n01.

[99] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", JINST **12**, P10003–P10003 (2017) 10.1088/1748-0221/12/10/p10003.

[100] W. Adam et al., "Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC", J. Phys. G **31**, N9–N20 (2005) 10.1088/0954-3899/31/9/n01.

[101] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", JINST **15**, P09018–P09018 (2020) 10.1088/1748-0221/15/09/p09018.

[102] A. Bodek et al., "Misaligment and Muon Scale Corrections Extracted from 2011A $Z \to \mu\mu$ Sample", 2012, https://www-cdf.fnal.gov/~jyhan/cms_momscl/rochcor_cmsnote.pdf.

[103] CMS Collaboration, *Rochester Corrections*, (2020) https://twiki.cern.ch/twiki/bin/viewauth/CMS/RochcorMuon.

[104] A. Bodek et al., "Extracting muon momentum scale corrections for hadron collider experiments", Eur. Phys. J. C **72**, 2194 (2012) 10.1140/epjc/s10052-012-2194-8.

[105] M. Cipriani, "Measurement of the helicity of the W boson with the CMS experiment", PhD thesis (Rome U., 2019).

[106] CMS Collaboration, *Accurate muon momentum scale and resolution in view of the W mass measurement*, https://twiki.cern.ch/twiki/bin/view/CMS/MuonScaleResolKalman.

[107] CMS Collaboration, *W-like measurement of the Z boson mass using dimuon events collected in pp collisions at* $\sqrt{s} = 7$ TeV, tech. rep. CMS-PAS-SMP-14-007 (CERN, Geneva, 2016), http://cds.cern.ch/record/2139655.

[108] E. Manca, "Validation of the muon momentum resolution in view of the W mass measurement with the CMS experiment", MA thesis (Pisa U., 2016), https://cds.cern.ch/record/2233647.

[109] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV", JINST **7**, P10002–P10002 (2012) 10.1088/1748-0221/7/10/p10002.

[110] CMS Collaboration, *Measurement of Tracking Efficiency*, tech. rep. CMS-PAS-TRK-10-002 (CERN, Geneva, 2010), http://cds.cern.ch/record/1279139.

[111] CMS Collaboration, *Muon tracking performance in the CMS Run-2 Legacy data using the tag-and-probe technique*, tech. rep. CMS-DP-2020-035 (2020), https://cds.cern.ch/record/2724492.

[112] M. Cacciari et al., "The anti-ktjet clustering algorithm", J. High Energy Phys. **2008**, 063–063 (2008) 10.1088/1126-6708/2008/04/063.

[113] S. Catani et al., "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions", Nucl. Phys. B **406**, 187–224 (1993) 10.1016/0550-3213(93)90166-M.

[114] M. Cacciari et al., "FastJet user manual", Eur. Phys. J. C **72** (2012) 10.1140/epjc/s10052-012-1896-2.

[115] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", JINST **12**, P02014–P02014 (2017) 10.1088/1748-0221/12/02/p02014.

[116] CMS Collaboration, *Introduction to Jet Energy Corrections at CMS*, https://twiki.cern.ch/twiki/bin/view/CMS/IntroToJEC.

[117] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector", JINST **14**, P07004–P07004 (2019) 10.1088/1748-0221/14/07/p07004.

[118] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", J. High Energy Phys. **11**, 040 (2004) 10.1088/1126-6708/2004/11/040.

[119] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", J. High Energy Phys. **07**, 079 (2014) 10.1007/JHEP07(2014)079.

[120] NNPDF Collaboration, "Parton distributions for the LHC run II", J. High Energy Phys. **2015**, 40 (2015) 10.1007/JHEP04(2015)040.

[121] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", Comput. Phys. Commun. **191**, 159–177 (2015) 10.1016/j.cpc.2015.01.024.

[122] S. Agostinelli et al., "Geant4—a simulation toolkit", Nucl. Instrum. Methods Phys. Res. A **506**, 250–303 (2003) https://doi.org/10.1016/S0168-9002(03)01368-8.

[123] SLAC, *Simulations*, https://theory.slac.stanford.edu/our-research/simulations.

[124] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, SLD Collaboration, LEP Electroweak Working Group, SLD Electroweak Group, SLD Heavy Flavour Group, "Precision electroweak measurements on the Z resonance", Phys. Rep. **427**, 257–454 (2006) https://doi.org/10.1016/j.physrep.2005.12.006.

[125] R. K. Ellis et al., *QCD and collider physics*, Cambridge monographs on particle physics, nuclear physics, and cosmology (Cambridge University Press, Cambridge, 2003), 10.1017/CBO9780511628788.

[126] G. Bozzi et al., "Impact of the parton distribution function uncertainties on the measurement of the W boson mass at the Tevatron and the LHC", Phys. Rev. D **83**, 113008 (2011) 10.1103/PhysRevD.83.113008.

[127] L. Bianchini and G. Rolandi, "A critical point in the distribution of lepton energies from the decay of a spin-1 resonance", J. High Energy Phys. **2019**, 44 (2019) 10.1007/JHEP05(2019)044.

[128] E. Bagnaschi and A. Vicini, "A new look at the estimation of the PDF uncertainties in the determination of electroweak parameters at hadron colliders", (2019), arXiv:1910.04726 [hep-ph], https://arxiv.org/abs/1910.04726.

[129] G. Bozzi et al., "Prospects for improving the LHC W boson mass measurement with forward muons", Eur. Phys. J. C **75**, 601 (2015) 10.1140/epjc/s10052-015-3810-1.

[130] R. Abdul Khalek et al., "Towards Ultimate Parton Distributions at the High-Luminosity LHC", Eur. Phys. J. C **78**, 962 (2018) 10.1140/epjc/s10052-018-6448-y.

[131] A. Buckley et al., "LHAPDF6: parton density access in the LHC precision era", Eur. Phys. J. C **75**, 132 (2015) 10.1140/epjc/s10052-015-3318-8.

[132] R. D. Ball et al., "Parton distributions from high-precision collider data", Eur. Phys. J. C **77**, 663 (2017) 10.1140/epjc/s10052-017-5199-5.

[133] S. L. Glashow, "Partial-symmetries of weak interactions", Nucl. Phys. **22**, 579–588 (1961) https://doi.org/10.1016/0029-5582(61)90469-2.

[134] S. Weinberg, "A Model of Leptons", Phys. Rev. Lett. **19**, 1264–1266 (1967) 10.1103/PhysRevLett.19.1264.

[135] A. Salam, "Weak and Electromagnetic Interactions", Conf. Proc. C **680519**, 367–377 (1968) 10.1142/9789812795915_0034.

[136] UA1 Collaboration, "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV", Physics Letters B **122**, 103–116 (1983) https://doi.org/10.1016/0370-2693(83)91177-2.

[137] UA2 Collaboration, "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider", Physics Letters B **122**, 476–485 (1983) https://doi.org/10.1016/0370-2693(83)91605-2.

[138] CERN, *The 300 GeV programme*, French version available : Le programme 300 GeV (CERN, Geneva, 1972), https://cds.cern.ch/record/104068.

[139] L. Di Lella and C. Rubbia, "The Discovery of the W and Z Particles", Adv. Ser. Dir. High Energy Phys. **23**, 137–163 (2015) 10.1142/9789814644150_0006.

[140] ALEPH Collaboration, *ALEPH: technical report 1983* (CERN, Geneva, 1983), https://cds.cern.ch/record/300680.

[141] OPAL Collaboration, *The OPAL detector: technical proposal*, tech. rep. CERN-LEPC-83-4. LEPC-P-3 (1983), https://cds.cern.ch/record/300669.

[142] DELPHI Collaboration, *DELPHI: technical proposal*, tech. rep. CERN-LEPC-83-3. DELPHI-83-66. LEPC-P-2 (1983), https://cds.cern.ch/record/300668.

[143] L3 Collaboration, *Technical proposal: L3*, tech. rep. CERN-LEPC-83-5. LEPC-P-4 (1983), https://cds.cern.ch/record/297266.

[144] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, LEP Electroweak Working Group, "Electroweak measurements in electron–positron collisions at W-boson-pair energies at LEP", Phys. Rep. **532**, Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP, 119–244 (2013) https://doi.org/10.1016/j.physrep.2013.07.004.

[145] ALEPH Collaboration, "Measurement of W pair production in e^+e^- collisions at 183 GeV", Phys. Lett. B **453**, 107–120 (1999) 10.1016/S0370-2693(99)00304-4.

[146] ALEPH Collaboration, "Measurement of the W boson mass and width in e^+e^- collisions at LEP", Eur. Phys. J. C **47**, 309–335 (2006) 10.1140/epjc/s2006-02576-8.

[147] M. Freytsis et al., "Prospects for a Measurement of the W Boson Mass in the All-Jets Final State at Hadron Colliders", J. High Energy Phys. **02**, 003 (2019) 10.1007/JHEP02(2019)003.

[148] G. Rolandi, "Analysis and discussion of the recent W mass measurements", in 52nd Rencontres de Moriond on EW Interactions and Unified Theories (2017), pp. 79–83, arXiv:1705.04962 [hep-ex].

[149] J. Collins et al., "Transverse momentum distribution in Drell-Yan pair and W and Z boson production", Nucl. Phys. B **250**, 199–224 (1985) https://doi.org/10.1016/0550-3213(85)90479-1.

[150] S. Catani et al., "Universality of non-leading logarithmic contributions in transverse-momentum distributions", Nucl. Phys. B **596**, 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)., 299–312 (2001) https://doi.org/10.1016/S0550-3213(00)00617-9.

[151] W. Bizoń et al., "The transverse momentum spectrum of weak gauge bosons at N^{3}LL+NNLO", Eur. Phys. J. C **79** (2019) 10.1140/epjc/s10052-019-7324-0.

[152] CDF Collaboration and D0 Collaboration, "Combination of CDF and D0 W-Boson mass measurements", Phys. Rev. D **88**, 052018 (2013) 10.1103/PhysRevD.88.052018.

[153] P. M. Nadolsky et al., "Implications of CTEQ global analysis for collider observables", Phys. Rev. D **78**, 013004 (2008) 10.1103/PhysRevD.78.013004.

[154] D. Stump et al., "Inclusive jet production, parton distributions, and the search for new physics", J. High Energy Phys. **2003**, 046–046 (2003) 10.1088/1126-6708/2003/10/046.

[155] H.-L. Lai et al., "New parton distributions for collider physics", Phys. Rev. D **82**, 074024 (2010) 10.1103/PhysRevD.82.074024.

[156] ATLAS Collaboration, "Precision measurement and interpretation of inclusive W^+ , W^- and Z/γ^* production cross sections with the ATLAS detector", Eur. Phys. J. C **77**, 367 (2017) 10.1140/epjc/s10052-017-4911-9.

[157] ATLAS Collaboration, "Measurement of the angular coefficients in Z-boson events using electron and muon pairs from data taken at $\sqrt{s} = 8$ TeV with the ATLAS detector", J. High Energy Phys. **2016** (2016) 10.1007/jhep08(2016)159.

[158] ATLAS Collaboration, *Summary of ATLAS Pythia 8 tunes*, tech. rep. ATL-PHYS-PUB-2012-003 (CERN, Geneva, 2012), https://cds.cern.ch/record/1474107.

[159] ATLAS Collaboration, "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", J. High Energy Phys. **2014** (2014) 10.1007/jhep09(2014)145.

[160] J. Pumplin et al., "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis", J. High Energy Phys. **2002**, 012–012 (2002) 10.1088/1126-6708/2002/07/012.

[161] G. A. Ladinsky and C.-P. Yuan, "Nonperturbative regime in QCD resummation for gauge boson production at hadron colliders", Phys. Rev. D **50**, R4239–R4243 (1994) 10.1103/physrevd.50.r4239.

[162] C. Balázs and C.-P. Yuan, "Soft gluon effects on lepton pairs at hadron colliders", Phys. Rev. D **56**, 5558–5583 (1997) 10.1103/physrevd.56.5558.

[163] S. Catani et al., "Vector boson production at hadron colliders: transverse-momentum resummation and leptonic decay", J. High Energy Phys. **2015**, 1–47 (2015) 10.1007/jhep12(2015)047.

[164] K. Hamilton et al., "MINLO: multi-scale improved NLO", J. High Energy Phys. **2012** (2012) 10.1007/jhep10(2012)155.

[165] K. Hamilton et al., "Merging H/W/Z + 0 and 1 jet at NLO with no merging scale: a path to parton shower + NNLO matching", J. High Energy Phys. **2013** (2013) 10.1007/jhep05(2013)082.

[166] ATLAS Collaboration and CMS Collaboration, "W boson mass measurement", edited by E. Auge et al., 71–77 (2017), https://cds.cern.ch/record/2264497/files/?docname=ATL-PHYS-PROC-2017-051&version=all.

[167] CMS Collaboration, "Measurement of the transverse momentum spectra of weak vector bosons produced in proton-proton collisions at $\sqrt{s} = 8$ TeV", J. High Energy Phys. **02**, 096 (2017) 10.1007/JHEP02(2017)096.

[168] ATLAS Collaboration, "Measurement of the transverse momentum distribution of W bosons in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", Phys. Rev. D **85**, 012005 (2012) 10.1103/PhysRevD.85.012005.

[169] CMS Collaboration, *CMS luminosity measurement for the 2017 data-taking period at* $\sqrt{s} =$ 13 TeV, tech. rep. CMS-PAS-LUM-17-004 (CERN, Geneva, 2018), https://cds.cern.ch/record/2621960.

[170] CMS Collaboration, "Measurements of differential Z boson production cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV", J. High Energy Phys. **2019** (2019) 10.1007/jhep12(2019)061.

[171] CMS Collaboration, "Measurement of the muon charge asymmetry in inclusive $pp \to W + X$ production at $\sqrt{s} = 7$ TeV and an improved determination of light parton distribution functions", Phys. Rev. D **90**, 032004 (2014) 10.1103/PhysRevD.90.032004.

[172] CMS Collaboration, "Measurement of the differential cross section and charge asymmetry for inclusive pp $\to \mathrm{W}^{\pm} + X$ production at $\sqrt{s} = 8$ TeV", Eur. Phys. J. C **76**, 469 (2016) 10.1140/epjc/s10052-016-4293-4.

[173] CMS Collaboration, "Measurement of the Electron Charge Asymmetry in Inclusive W Production in pp Collisions at $\sqrt{s} = 7$ TeV", Phys. Rev. Lett. **109**, 111806 (2012) 10.1103/PhysRevLett.109.111806.

[174] CMS Collaboration, "Measurement of the weak mixing angle using the forward-backward asymmetry of Drell-Yan events in pp collisions at 8 TeV", Eur. Phys. J. C **78**, 701 (2018) 10.1140/epjc/s10052-018-6148-7.

[175] CMS Collaboration, "Measurement of associated production of a W boson and a charm quark in proton-proton collisions at $\sqrt{s} = 13$ TeV", Eur. Phys. J. C **79**, 269 (2019) 10.1140/epjc/s10052-019-6752-1.

[176] J. C. Collins and D. E. Soper, "Angular distribution of dileptons in high-energy hadron collisions", Phys. Rev. D **16**, 2219–2225 (1977) 10.1103/PhysRevD.16.2219.

[177] C. S. Lam and W.-K. Tung, "Systematic approach to inclusive lepton pair production in hadronic collisions", Phys. Rev. D **18**, 2447–2461 (1978) 10.1103/PhysRevD.18.2447.

[178] E. Mirkes and J. Ohnemus, "W and Z polarization effects in hadronic collisions", Phys. Rev. D **50**, 5692–5703 (1994) 10.1103/PhysRevD.50.5692.

[179] CDF Collaboration, "First Measurement of the Angular Coefficients of Drell-Yan e^+e^- Pairs in the Z Mass Region from $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV", Phys. Rev. Lett. **106**, 241801 (2011) 10.1103/PhysRevLett.106.241801.

[180] CMS Collaboration, "Measurement of the Z boson differential cross section in transverse momentum and rapidity in proton–proton collisions at 8 TeV", Physics Letters B **749**, 187–209 (2015) 10.1016/j.physletb.2015.07.065.

[181] CDF Collaboration, "Measurement of the azimuthal angle distribution of leptons from W boson decays as a function of the W transverse momentum in pp¯collisions at $\sqrt{s} =$ 1.8 TeV", Phys. Rev. D **73** (2006) 10.1103/physrevd.73.052002.

[182] CDF Collaboration, "Measurement of the polar-angle distribution of leptons from W boson decay as a function of the W transverse momentum in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$TeV", Phys. Rev. D **70**, 032004 (2004) 10.1103/PhysRevD.70.032004.

[183] DØ Collaboration, "Measurement of the angular distribution of electrons from $W \to e\nu$ decays observed in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV", Phys. Rev. D **63**, 072001 (2001) 10.1103/PhysRevD.63.072001.

[184] CMS Collaboration, "Measurement of the Polarization of W Bosons with Large Transverse Momenta in $W+$jets Events at the LHC", Phys. Rev. Lett. **107**, 021802 (2011) 10.1103/PhysRevLett.107.021802.

[185] ATLAS Collaboration, "Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector", Eur. Phys. J. C **79** (2019) 10.1140/epjc/s10052-019-7027-6.

[186] CMS Collaboration, *CMS Luminosity Measurements for the 2016 Data Taking Period*, tech. rep. CMS-PAS-LUM-17-001 (CERN, Geneva, 2017), https://cds.cern.ch/record/2257069.

[187] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", Eur. Phys. J. C **76** (2016) 10.1140/epjc/s10052-016-3988-x.

[188] CMS Collaboration, *Recommended Jet Energy Corrections and Uncertainties For Data and MC*, https://twiki.cern.ch/twiki/bin/view/CMS/JECDataMC.

[189] CMS Collaboration, *Cut Based Electron ID for Run 2*, https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2.

[190] CMS Collaboration, *MET Filter Recommendations for Run II*, https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2.

[191] CMS Collaboration, "Measurements of differential Z boson production cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV", J. High Energy Phys. **2019** (2019) 10.1007/jhep12(2019)061.

[192] CMS Collaboration, *Pileup mitigation at CMS in 13 TeV data*, tech. rep. arXiv:2003.00503. CMS-JME-18-001-003 (CERN, Geneva, 2020), https://cds.cern.ch/record/2711736.

[193] S. Carrazza et al., "An unbiased Hessian representation for Monte Carlo PDFs", Eur. Phys. J. C **75** (2015) 10.1140/epjc/s10052-015-3590-7.

[194] P. Golonka and Z. Was, "PHOTOS Monte Carlo: a precision tool for QED correctionsin Z and W decays", Eur. Phys. J. C **45**, 97–107 (2006) 10.1140/epjc/s2005-02396-4.

[195] R. Gauld et al., "Precise predictions for the angular coefficients in Z-boson production at the LHC", J. High Energy Phys. **2017** (2017) 10.1007/jhep11(2017)003.

[196] Martìn Abadi et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, Software available from tensorflow.org, 2015, https://www.tensorflow.org/.

[197] ATLAS Collaboration, CMS Collaboration, LHC Higgs Combination Group, *Procedure for the LHC Higgs boson search combination in Summer 2011*, tech. rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11 (CERN, Geneva, 2011), https://cds.cern.ch/record/1379837.

[198] CMS Collaboration, *Combine*, https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/.

[199] R. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples", Comput. Phys. Commun. **77**, 219–228 (1993) https://doi.org/10.1016/0010-4655(93)90005-W.

[200] G. Cowan et al., "Asymptotic formulae for likelihood-based tests of new physics", Eur. Phys. J. C **71** (2011) 10.1140/epjc/s10052-011-1554-0.

[201] J. Bradbury et al., *JAX: composable transformations of Python+NumPy programs*, version 0.2.5, 2018, http://github.com/google/jax.

[202] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", JINST **13**, P05011 (2018) 10.1088/1748-0221/13/05/P05011.

[203] CMS Collaboration, "Measurement of jet substructure observables in $t\bar{t}$ events from proton-proton collisions at $\sqrt{s} = 13$ TeV", Phys. Rev. D **98**, 092014 (2018) 10.1103/PhysRevD.98.092014.

[204] L. Asquith et al., "Jet Substructure at the Large Hadron Collider : Experimental Review", (2018), arXiv:1803.06991 [hep-ex].

[205] J. MacQueen, "Some methods for classification and analysis of multivariate observations", in Proc. Fifth Berkeley Symp. on Math. Statist. and Prob., Vol. 1 (1967), p. 281, https://projecteuclid.org/euclid.bsmsp/1200512992.

[206] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement", (2018), arXiv:1804.02767.

[207] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization", (2014), arXiv:1412.6980.

[208] F. Chollet et al., *Keras*, https://keras.io, 2015.

[209] CMS Collaboration, "Mini-AOD: A New Analysis Data Format for CMS", J. Phys. Conf. Ser. **664**, 072052 (2015) 10.1088/1742-6596/664/7/072052.

[210] CMS Collaboration, *NanoAODv6*, https://gitlab.cern.ch/cms-nanoAOD/nanoaod-doc/-/wikis/Releases/NanoAODv6.

[211] CMS Collaboration, *PdmV Data Reprocessing NanoAODv6*, https://twiki.cern.ch/twiki/bin/viewauth/CMS/PdmVDataReprocessingNanoAODv6.

[212] ROOT team, *RDataFrame Class Reference*, https://root.cern/doc/master/classROOT_1_1RDataFrame.html.

[213] E. Manca et al., *W-properties analysis framework GitHub repository*, Version used in this work: https://github.com/vberta/wproperties-analysis/tree/vberta_v1.0.0_rebased, State-of-the-art: https://github.com/emanca/wproperties-analysis.

[214] E. Manca and E. Guiraud, *Using RDataFrame, ROOT's declarative analysis tool, in a CMS physics study*, https://indico.cern.ch/event/849610/.

[215] S. Roy Chowdhury et al., *RDataFrame for m_W measurement*, https://indico.cern.ch/event/957050/contributions/4063885/attachments/2122924/3573562/wmassPresentationxpog.pdf.