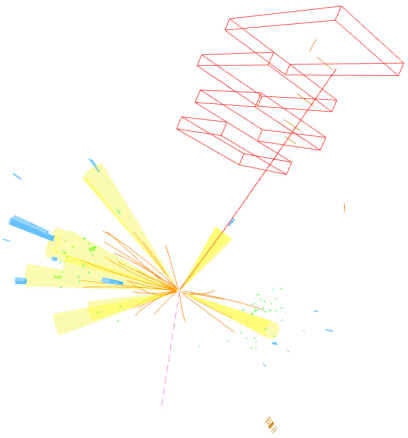
SCUOLA
NORMALE
SUPERIORE

Classe di Scienze
Corso di Perfezionamento in Fisica
A. A. 2021/2022

Precision measurements of W detected at CMS



Candidate:

Elisabetta Manca

Thesis submitted for the degree of
Doctor of Philosophy in Physics

Advisor:
Prof. Luigi Rolandi

Contents

The quest for the missing piece of the puzzle

Particle Physics is a well-established field using its settled methods of data analysis and interpretation, along with techniques of detector calibration.

At the dawn of the era of the Higgs boson discovery, with no clear sign of new Physics manifesting at the energy range covered by the Large Hadron Collider, the quest for significant deviations of the measured parameters of the Standard Model from their prediction is a mandatory plan. A fair question to ask is: to what extent pushing the limits of the achievable precision means to rethink those procedures we have been familiar with for a longtime?

Enabling increased precision in the electroweak sector

Produced copiously in the CMS and ATLAS experiments, the W bosons provide a unique opportunity to set stringent limits to the Standard Model.

A precise measurement of its kinematic distributions - rapidity and transverse momentum - has never been performed in LHC collisions due to the impossibility to fully measure the final state in leptonic decays, while hadronic decays cannot be used because of the limited control of the jet energy scale and the overwhelming background. Moreover, the holy grail of the precision measurements at the electroweak scale - a measurement of the W boson mass at 10^{-4} level - has not yet been accomplished.

Traditionally, the efforts of the collaborations have been invested in this last charming piece of the puzzle. Notwithstanding, a precise knowledge of the W rapidity, transverse momentum and polarisation is fundamental to give a fair assessment of its mass. In W leptonic decays, due to the neutrino escaping direct detection, it is not possible to define a Lorentz invariant proxy for the W mass. Thus, the estimation of the W mass is affected by the degree of knowledge of the underlying W kinematics. The classical approach has been to finely tune the simulations to resemble the measured distributions of the Z boson, used as a standard candle, to make assumptions on the unknown kinematics of the W . However, not only is the validity of this method still debated in the community, especially about the evaluation of the systematic uncertainties, but it has not been proven yet to be able to get to the ultimate precision.

The aim of the work described in this thesis has been to present a paradigm shift, introducing a change in the traditional conception of the W precision measurements. A totally new method has been devised to obtain a precise assessment of the W rapidity, transverse momentum and polarisation exploiting uniquely and exclusively the correlation between the transverse momentum and the pseudorapidity of muons emitted in W semileptonic decays.

The validity of the new procedure relies on only two properties of the W boson: it is a particle of spin

1 and when it decays, it violates completely the parity symmetry. The former allows to obtain a well-defined formula that links the W rapidity, transverse momentum and polarisation to the angles of the emitted lepton in the W center of mass. The latter allows to unfold directly the W production from the kinematics of the emitted lepton in the frame of reference of the laboratory. An *ad-hoc* simulation is used to select the W boson in all possible kinematic configurations, and for each of those, a template is produced: this is a prediction of how the kinematics of the emitted lepton in the frame of reference of the laboratory looks like given the assumed configuration of W production.

The fact that the W has spin 1 guarantees that the number of templates is finite. The idea is that the correlation between the transverse momentum and the pseudorapidity of the leptons emitted in the laboratory frame and well measured in data is given by the linear combination of the templates obtained as aforementioned. The coefficients of this linear combination can be obtained through a fit and will represent a measurement of the W kinematics to be performed simultaneously with the traditional template-based W mass extraction. Due to the correlations and degeneracy among the various templates, an extensive number of events has to be employed in order to achieve the desired level of precision.

Setting a new standard for big-data processing

An important aspect of this procedure is the study of the feasibility in terms of computing resources. In order for it to work, it is necessary to process a huge number of events (order of one billion) and to produce order one hundred thousand templates (including their systematic variations) under the form of 2-dimensional histograms. In this context, the development of new software tailored for this purpose is imperative: the traditionally used methods are not suitable for the complexity of this analysis.

Rethinking the foundations of data processing has involved the common effort of analysers and computer scientists. The close contact with the ROOT team at CERN has enabled an outstanding collaboration which set a new standard for big-data processing in the High Energy community.

A challenge in detector calibration

Once the W production has been measured using one billion W decays, the next systematic limitation for a 10^{-4} precise measurement of the W mass is the knowledge of the muon momentum scale.

Reaching this goal means having a precise knowledge of an outstanding number of details of muon reconstruction in the tracking detector. The tradition of the offline assessment of the track parameters has a long and glorious history in the CMS collaboration, which has allowed the conquer of many important milestones in the history of the experiment. But once again to break the wall of the ultimate precision, a further step is needed. In the context of this thesis, a new method has been devised for track fitting based on the Generalised Broken Line fit, in addition with a new set of procedures based on J/ψ and Z data and simulation.

Outline of this document

With the exception of Chapter 2 where the CMS detector, our experimental setup, is described, this thesis contains exclusively original work performed by the author of this thesis and close collaborators. In Chapter 1, the proposed measurements are set in the context of the electroweak precision sector and their phenomenological concepts and their feasibility at CMS are discussed. Chapter 3 includes a brief description of the new computing resources that have been set up for these measurements, introducing the context of big data in the field of high energy physics. Chapter 4 is entirely devoted to the muon momentum scale calibration, starting with an original dissertation about the assessment of track parameters at silicon detectors. Then, the producedure for the calibration is discussed and its results presented. Finally, Chapter 5 and 6 describe respectively the measurements of W differential cross sections and angular coefficients performed on data collected by CMS in 2016, and an assessment of the uncertainty on the W mass using the same dataset blinding its central value.

Chapter 1

Innovating a long standing tradition

The Standard Model of Particle Physics has an outstanding success in correctly describing physics observed at colliders. Nevertheless, it can not be the candidate for the ultimate theory: experimental evidence of Dark Matter and neutrino oscillations are some example of phenomena not finding their theoretical explanation here. Historically, electroweak fits have played a major role in predicting observables before their measurement and even more after the discovery of the Higgs boson the program of delivering more and more precise experimental determination of key quantities has become imperative for testing the Standard Model.

It is from a position of the utmost humbleness that we have to look at the exceptional work that has been carried on at LEP, SLC, Tevatron and more recently at LHC to provide valuable measurements and the superlative amount of effort infused by the theory community to improve the calculations. With this spirit, the first part of this chapter, in Sections 1.1 and 1.2 will be devoted to a brief review of the electroweak sector of the Standard Model and the global electroweak fit, with a particular emphasis on the measurement of the W mass.

Notwithstanding, there is small possibility of pursuing the ambition of continuously seeking the truth without critically analysing what has been already done, while understanding if something can be improved, at the light of the advancements in comprehension of the phenomena and of the current available technologies. Therefore, the last Section of this chapter, 1.3, will present an original review of the limiting factors in the determination of a measurement of the W mass below 10 MeV on a phenomenological point of view and also provide a new method to tackle them.

1.1 The electroweak sector of the Standard Model

The electroweak sector of the Standard Model combines the electromagnetic and weak force as elementary forces of nature at a unification scale of $v = 246$ GeV [1–4]. Four gauge bosons are present in this description: a massless photon, mediator of the electromagnetic interaction, and the massive W^+, W^- and Z bosons, mediators of the weak interaction. In order for the latter to consistently acquire mass, the Higgs mechanism is required [5–7].

Mathematically, this is formulated as a Quantum Field Theory under the symmetry $SU(2)_L \times U(1)_Y$. While this symmetry is unbroken, the fermions and the gauge bosons are massless. The Spontaneous Symmetry breaking of the electroweak sector introduced by Higgs, Brout and Englert guarantees that the W^+, W^- and Z bosons acquire mass while keeping the theory renormalisable. At the leading order

of the perturbation theory, a precise value for the predicted masses is given:

$$M_{W^\pm} = \frac{gv}{2} \tag{1.1}$$

$$M_Z = \frac{v}{2}\sqrt{g^2 + g'^2}, \tag{1.2}$$

where the coupling constants g and g' are linked to θ_W, the Weinberg angle, and to the electric charge e through the relations:

$$\tan\theta_W = \frac{g'}{g} \quad \text{and} \quad g'\cos\theta_W = e. \tag{1.3}$$

e can be defined in terms of α, the fine structure constant: $e = \sqrt{4\pi\alpha}$.

Moreover, the particular choice of the Higgs doublet generates an extra $SU(2)$ symmetry which constrains the quantity:

$$\rho \equiv \frac{M_W}{M_Z \cos\theta_W} \tag{1.4}$$

to be equal to 1 at tree level, with small corrections at NLO. Let us now concentrate on the W boson, main topic of the work presented in this document.

Electroweak corrections to the W mass

The expression at tree level for the W mass given in Eq. 1.1 can be rewritten in terms of α, the fine structure constant, G_F, the Fermi constant and θ_W:

$$M_W^2 = \frac{\pi\alpha}{\sqrt{2}G_F \sin^2\theta_W}, \tag{1.5}$$

If we take into account the electroweak corrections, i.e. radiative corrections to the leading order arising from electromagnetic and weak effects, it can be corrected in the following way [8]:

$$M_W^2 = \frac{\pi\alpha}{\sqrt{2}G_F \sin^2\theta_W}\left(1 + \Delta r\right). \tag{1.6}$$

Δr receives inputs from various contributions:

$$\Delta r = \Delta\alpha - \frac{\cos^2\theta_W}{\sin^2\theta_W}\Delta\rho + \Delta r_{rem}. \tag{1.7}$$

$\Delta\alpha$ is the correction to α due to photon vacuum polarisation and it is sensitive to the fermion masses through $\log\left(s/m_f^2\right)$, where s is the energy of the center of mass. $\Delta\rho$ is the deviation from 1 of the ρ parameter defined in Eq. 1.4.

In the leading terms this deviation comes from corrections to the W propagator with loops contain-

ing isodoublets and it is proportional to the isodoublet mass splitting $|m_{f1}^2 - m_{f2}^2|$. The most important contribution comes then from the top-bottom doublet. Finally, there is a logarithmic correction due to the Higgs boson mass.

1.2 Precision measurements as a tool for discovery

The electroweak sector of the Standard Model provides punctual predictions for the gauge boson masses. These predictions can in turn be continuously refined improving the calculations to higher perturbative orders and also considering corrections coming from Quantum Chromodynamics (QCD). On the other hand, in the history of High Energy Physics, a large number of experimental measurements of the relevant observables in the Standard Model have been carried out by various collaborations in a large number of particle acceleration facilities.

It then becomes relevant to create an efficient framework enabling for checks of consistency between predictions and measurements. A common effort in the theoretical and experimental communities with this purpose in mind is continuously evolving and has become crucial after the discovery of the Higgs boson and the determination of its mass. This historical turning point has marked the completion of the list of the Standard Model parameters from an experimental point of view and from that moment it has become possible to exploit at best the predictive power of the theory to test its internal consistency.

Global electroweak fits and analyses have started already before the discovery of the W [9, 10] and Z [11, 12] bosons by the UA1 and UA2 collaborations. The subsequent addition of the high precision measurements of the Z mass and $\sin^2\theta_{eff}$ from LEP and SLC [13] have allowed a prediction of the value of the mass of the top quark before its discovery by CDF and D0 experiments at Tevatron [14, 15]. In turn, LEP2 and Tevatron provided valuable measurements of the top mass and the W mass enabling for a prediction of the Higgs mass before its discovery [16, 17].

LHC as a machine for precision measurements

After the discovery of the Higgs boson and the experimental determination of its mass, global electroweak fits have acquired an even more crucial importance for discovering eventual hints of new Physics under the form of tensions between measurements and Standard Model predictions. At the same time, the outstanding luminosity of the LHC has opened a new era of precision measurements in the electroweak sector with an unprecedented number of events of W and Z bosons having been produced. Although the clean environment of past e^+e^- facilities such as LEP and SLC has been the most favourable for precision measurements, they have been eventually limited by statistics. On the other hand, with its many interesting events LHC is now reaching the level of precision set by e^+e^- machines in the determination of key observables of the electroweak sector, like $\sin\theta_W$ [18].

State of the art global electroweak fit

The two ingredients needed for a global fit are accurate measurements and precise calculations. The experimental measurements are used as input. At tree level all the variables in the Standard Model can be computed by three parameters, chosen to be the most precisely measured: α, the fine structure constant, known to a relative precision of $3{\cdot}10^{-10}$, G_F, the Fermi constant, known to a relative precision of $5 \cdot 10^{-7}$ and the mass of the Z , m_Z, measured with a precision of 2 MeV ($2 \cdot 10^{-5}$ relative precision) in the scan of the Z resonance in e^+e^- collisions. In order to compute higher order corrections more input is needed: the strong coupling constant α_S, the running of the fine structure constant to the Z mass $\Delta\alpha$, the mass of the Higgs bosons and the masses of the fermions, but in practice only the mass of the top quark m_t among the fermions matters, since it is orders of magnitude more massive than the others and the corrections grow with the mass value.

Global fits of the electroweak sector are possible if the input parameters are over-constrained with a prediction better than the effect of alleged new particles at one loop level. The prediction for one parameter of the Standard Model can be obtained removing the input value of that parameter and fitting all the other parameters to the theory (indirect determination). The latest version of the global electroweak fit has been updated [19] with a combination of the most precise kinematic top quark mass from ATLAS and CMS [20, 21] and W boson mass measurement from the ATLAS experiment [22], a new $\sin^2\theta_{eff}$ result from the Tevatron [23], a Higgs boson mass combination released by the ATLAS and CMS collaborations [24], and a new evaluation of the hadronic contribution to α_S at the Z pole [25]. The result of this fit is reported in Figure 1.1. The set of observables used as input in the fit includes the mass of the Higgs boson, the mass of the top quark, the masses and widths of W and Z bosons, $\sin\theta_W$, α_S at the Z pole and measurements of the asymmetries. In Figure 1.1 shows two values for each observable: the difference between the indirect determination and the experimental value (black) and the difference between the indirect determination and the global fit (orange). These differences and their uncertainty bar are rescaled by the sum in quadrature of the uncertainties of experimental measurement and the indirect determination. Also shown in blue and centred at zero there is the rescaled 1σ uncertainty of the indirect determination. When the experimental input is much more precise than the indirect determination, the width of the blue band is 1σ and the orange error bar coincides with the black one. This is the case of the top and Higgs masses, which enter in the Standard Model only at loop level, and of the Z mass, whose measurement is very precise. One can also notice the anticorrelation between the top and the Higgs mass , which compete in Δr. When instead the experimental input is much less precise than the indirect determination, the orange and the blue bands coincide. This is the case of those observables that are very difficult to measure (eg Γ_w or R_c) but this also happens when the prediction is very precise because several different observables measure essentially the same quantity, like the Asymmetries which all measure θ_W. The plot shows the long standing discrepancy between the measurements of the LR asymmetry with polarised beam at SLAC and the b-quark forward backward asymmetry at LEP. LHC-HL will possibly be in the position to contribute to solving this discrepancy. The final word will be said by the next generation of e+e- colliders [26]. These machines will be also needed to improve the experimental determination of σ^o_{had} [1] and other Z pole

[1] A new value of σ^0_{had} has been reported recently in [27] including a small correction to the LEP luminosity evaluation [28]

observables.

On a different page, a concrete opportunity to resolve an ancient tension between measurement and Standard Model prediction is offered by the measurement of m_W . It appears in Figure 1.1 that the indirect determination is more precise than the direct measurement and that they differ at the level of 2σ. For this reason, the experimental community has a strong motivation to pursue a precision measurement of m_W using the unprecedented statistics collected at the LHC experiments.

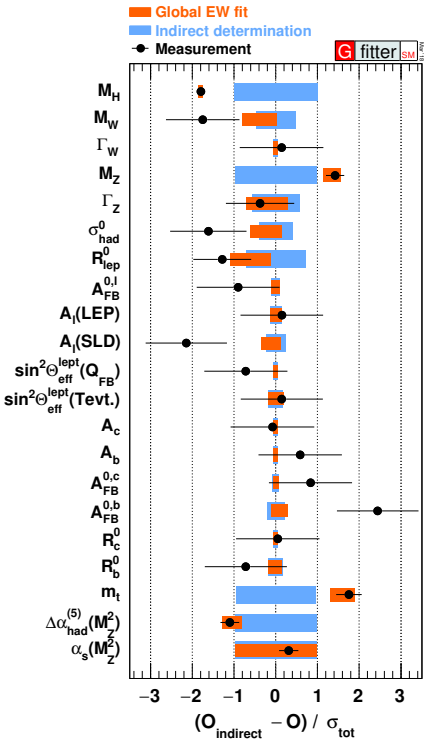


Figure 1.1: Difference of the input measurement and the indirect determination (black dots); difference of the global fit result and the indirect determination (red bands); indirect determination centred at 0 (blue bands). From ref. [19]

1.2.1 A precision measurement of the W mass

Figure 1.2 shows the experimental summary of the measurements of the W boson mass updated with the most recent result from the LHCb collaboration.

The first precision measurements of the W mass have been performed at LEP. At lepton colliders the W bosons are mainly produced in pairs through the process $e^+e^- \rightarrow W^+W^-$. There are two main methods to measure m_W at lepton colliders. The first one exploits the fact that the W^+W^- production cross section σ_{WW} is sensitive to m_W in the threshold region, $\sqrt{s} \sim 2m_W$. A measurement of σ_{WW} can then be used to measure m_W. In the second one events with two W bosons decaying to jets are selected

and an improved Bhabha cross section. Using this new value the σ^0_{had} agreement between the experimental and indirect determination moves from -1.6σ to -0.14σ.

and their invariant mass shape is constructed using the constraint from the known 4-momentum of the initial state (known at a 10^{-4} level) that fixes the jet energies, while the jet direction is measured by the detector. m_W is derived from a fit to the invariant mass.

On the other hand, at hadron colliders events with W decaying hadronically are not usable for a precision measurement of the W mass. Since the energy of the initial state is not determined, the invariant mass has to be reconstructed using the jet energy scales which is known not better than 1% level, which is not sufficient to give a competitive measurement of m_W. Moreover, due to the overwhelming background, it would be difficult to select a pure sample of W . Therefore, only events with a W decaying leptonically can be exploited for this purpose. This brings a major drawback, though: since an undetectable neutrino is present in the final state, it is not possible to reconstruct the invariant mass of each event. Then, a proxy carrying information about m_W has to be defined. The transverse momentum p_T of the charged lepton in the final state shows a jacobian peak at $m_W/2$. However, this observable is not a Lorentz invariant, its assessment will be dependent on the underlying kinematics of the W boson, which is not experimentally accessible and has to be inferred from external inputs. The limited knowledge of these inputs is the limiting factor in measuring m_W at hadron colliders and accounts for about 15 MeV in the final uncertainty in the measurement performed by ATLAS [22].

The most precise single measurement belongs to the ATLAS collaboration and has an uncertainty of 19 MeV, while the world average used as input in the global electroweak fit has an uncertainty of 13 MeV. In addition, Figure 1.2 reports the prediction from the global electroweak fit carrying an uncertainty of 7 MeV, each of the following effects contributing to an equal share: experimental and theoretical determination of the top mass, experimental Z mass, $\Delta\alpha_S$, $\Delta\alpha_{had}$, while the highest contribution, 4 MeV, is given by the calculation of m_W. In order to shed light on the existing tension in the predicted and measured values, the next goal of the experimental community is to provide a measurement with an absolute precision of below 10 MeV.

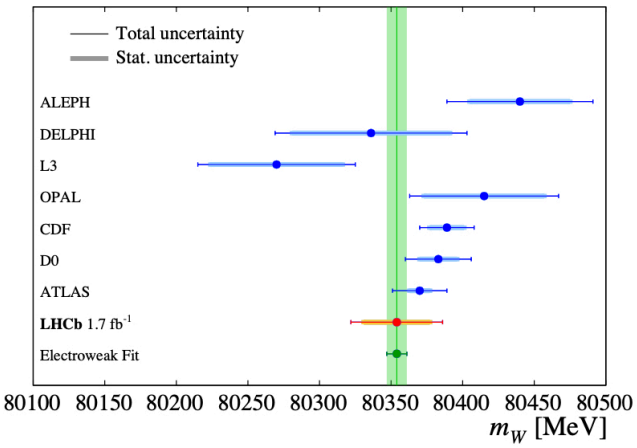


Figure 1.2: Measured value of m_W compared to those from the ALEPH [29], DELPHI [30], L3 [31], OPAL [**mWOpal**], CDF [32], D0 [33] and ATLAS [22] experiments. The current prediction of m_W from the global electroweak fit is also included. Figure from [34].

1.3 W mass measurement: an innovative proposal

In the Standard Model phenomenology W bosons are produced at hadron colliders through the annihilation of a quark and an antiquark. In proton-proton interactions the energy available in the center of mass depends on the fraction of the proton momentum carried by the partons participating to the hard scattering. Therefore, the computation of the cross section for the production of a W boson has to be weighted with the probability that the interacting partons carry a fraction x of the momentum of the proton. This probability is encoded in the Parton Distribution Functions (PDFs), which are determined experimentally, and will be extensively discussed in Section 1.3.1. 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:

$$\sigma_{pp} = \int \sum_{i,j} f_i(x_1, q^2) f_j(x_2, q^2) \hat{\sigma}_{i,j}(x_1 p_1, x_2 p_2, \mu_F, \mu_R) dx_1 dx_2, \tag{1.8}$$

where $f_i(x)$ is the PDF of parton type i at momentum xp and $\hat{\sigma}_{ij}$ represents the cross section of the interaction between partons type i, j. μ_R and μ_F are the renormalisation and factorisation scales respectively.

These scales absorb the collinear singularities and the truncation to a given perturbative order in the renormalisation procedure. Therefore, they are not to be considered observables, but rather a way to parametrise the uncertainty of a given theory calculation: the higher the perturbative order of the calculation, the smaller this residual uncertainty.

At hadron colliders m_W is measured using events in which W decay to a charged lepton and a neutrino. Due to the undetected neutrino in the final state, the measurement is done using observables in the transverse plane. The transverse momentum p_T of the charged lepton is sensitive to m_W at its maximum that is also known as Jacobian peak. However, the distribution of p_T is distorted by the W transverse momentum q_T, as will be discussed in section 1.3.2. This can be mitigated if the transverse mass m_T is introduced:

$$m_T = \sqrt{2 p_T E_T^{miss} (1 - \cos\phi)}, \tag{1.9}$$

where $\vec{E_T}^{miss} = -\sum \vec{p_T}$ over all the particles in the event is the missing transverse energy (MET) and ϕ is the angle between the lepton transverse momentum and the $\vec{E_T}^{miss}$ in the transverse plane. m_T has the advantage of depending on q_T only at second order, but the drawback of requiring the measurement of the missing transverse energy which has poor scale and resolution in presence of pileup. The transverse momentum p_T can only be measured inside the detector acceptance. This introduces an extra dependence on the PDFs that will be illustrated in Section 1.3.1. The mean value of p_T is used as a good estimator of m_W in the following discussion. We will show how its value depends on the experimentally inaccessible underlying kinematics of the produced W .

1.3.1 Dependence of p_T on PDFs

The dependence of the p_T average value on the PDFs can be illustrated in a very simplified case. We assume here that W is produced at transverse momentum $q_T = 0$ by quark-antiquark collisions and we neglect the width of the W . W bosons produced at $q_T = 0$ are transversely polarized. Firstly, let us rewrite the cross section for a resonance of mass m_W as:

$$\sigma = \sigma_0 \int dx_1 dx_2 q_1(x_1)\bar{q}_2(x_2)\delta(x_1x_2 - \frac{m_W^2}{s}) \tag{1.10}$$

where $q_1(x_1)\bar{q}_2(x_2)$ are the PDFs and s is the energy of the center of mass for the proton-proton interaction. Changing variables to W rapidity $y = \frac{1}{2}\ln\frac{x_1}{x_2}$ and $t = x_1x_2$, the Jacobian is 1 and the cross section can be rewritten as

$$\sigma = \sigma_0 \int dtdy q_1(\sqrt{t}e^y)\bar{q}_2(\sqrt{t}e^{-y})\delta(t - \frac{m_W^2}{s}) \tag{1.11}$$

and the differential cross section in one bin of rapidity is:

$$\frac{\partial\sigma}{dy} = \sigma_0 q_1(\frac{m_W}{\sqrt{s}}e^y)\bar{q}_2(\frac{m_W}{\sqrt{s}}e^{-y}) \equiv f(y) \tag{1.12}$$

We observe that Eq. 1.12 explicitly depends on the PDFs $q_1(\frac{m_W}{\sqrt{s}}e^y)$ and $\bar{q}_2(\frac{m_W}{\sqrt{s}}e^{-y})$. Let us define the pseudorapidity η of the charged lepton as $\eta = -\ln\tan\frac{\theta}{2}$, θ being the decay angle. The symbol η^L will be used for the pseudorapidity in the laboratory frame and η^0 for the pseudorapidity in the center of mass frame. Let us introduce useful relations:

$$\eta^L = y + \eta^0, \tag{1.13}$$

and

$$p_T \simeq \frac{m_W}{2}\sin\theta^* = \frac{m_W}{e^{+\eta^0} + e^{-\eta^0}}. \tag{1.14}$$

Eq. 1.13 is due to the addition property of rapidity when boosting from one reference frame to another, while Eq. 1.14 will be derived as a special case when dealing with the full calculation of the general case in section 1.3.2 where the W has finite transverse momentum.

We are now ready to compute the cross section for the detection of a charged lepton in a bin of η^L. The angular cross section for emitting a charged lepton at an angle θ^* in the W reference frame from a transversely polarized W is:

$$\frac{\partial\sigma}{\partial\cos\theta^*} = \frac{3}{8}(1 \pm \cos\theta^*)^2 = \frac{3}{2}\frac{e^{\pm 2\eta^0}}{(e^{+\eta^0} + e^{-\eta^0})^2} \tag{1.15}$$

where $\pm$ is the sign of the helicity of the W boson. The cross section as a function of η^0 is given by:

$$\frac{\partial\sigma}{\partial\eta^0} = \frac{\partial\sigma}{\partial\cos\theta^*}\frac{\partial\cos\theta^*}{\partial\eta^0} = \frac{6e^{\pm 2\eta^0}}{(e^{+\eta^0} + e^{-\eta^0})^4} \tag{1.16}$$

$$\frac{\partial^2\sigma_+}{\partial y \partial \eta^0} = \frac{\partial \sigma_+}{\partial y}\frac{\partial \sigma}{\partial \eta^0} = f^+(y)\,\frac{6e^{2\eta^0}}{(e^{+\eta^0}+e^{-\eta^0})^4}, \tag{1.17}$$

where the + sign is helicity of the W boson, the cross section to get a lepton at a given pseudorapidity bin $\Delta\eta^L$ above a given momentum cut (that is equivalent to a symmetric cut in η^0 in the interval $(-\eta_c^0, +\eta_c^0)$ is:

$$\frac{\partial\sigma_+}{\partial\eta^L} = \int_{y_{min}}^{y_{max}} dy \int_{-\eta_c^0}^{+\eta_c^0} d\eta^0 f^+(y)\frac{6e^{2\eta^0}}{(e^{+\eta^0}+e^{-\eta^0})^4}\delta(\eta^L - y - \eta^0). \tag{1.18}$$

Similarly, we can define the cross section for W of opposite helicity:

$$\frac{\partial\sigma_-}{\partial\eta^L} = \int_{y_{min}}^{y_{max}} dy \int_{-\eta_c^0}^{+\eta_c^0} d\eta^0 f^-(y)\frac{6e^{-2\eta^0}}{(e^{+\eta^0}+e^{-\eta^0})^4}\delta(\eta^L - y - \eta^0). \tag{1.19}$$

In practice we cannot select events in y and the integral is computed on the whole y distribution. Using Eq. 1.18 and 1.19 we can compute the mean value of p_T:

$$< p >= \frac{\int_{\eta_1^L}^{\eta_2^L} d\eta^L p(\frac{\partial\sigma_+}{\partial\eta^L} + \frac{\partial\sigma_-}{\partial\eta^L})}{\int_{\eta_1^L}^{\eta_2^L} d\eta^L (\frac{\partial\sigma_+}{\partial\eta^L} + \frac{\partial\sigma_-}{\partial\eta^L})}, \tag{1.20}$$

which can be assessed numerically using a set of PDFs. For this calculation we have used MMHT 2014 PDFs, described in [35]. The most precise PDFs available at present time also include NNPDF3.0 [36], [37] and CT18 [38]. In Figure 1.3a and 1.3b the results of this calculation is reported. The mean of p_T and its uncertainty respectively are computed as a function of an acceptance cut in $\Delta\eta^L$, in a range $(-\eta_c^0, +\eta_c^0) = (-0.8, +0.8)$, corresponding to $p_T \simeq 30$ GeV. We observe from Figure 1.3a that the mean value of p_T changes of around 20 MeV depending on the range in acceptance, and in Figure 1.3b we compute its uncertainty using the hessian of the PDFs. This is an evidence that due to the acceptance in η^L, only a part of the y rapidity spectrum is accepted and therefore the observables in the transverse plane acquire a dependence on the input value of y given by the PDFs. In turn, since the y spectrum is experimentally inaccessible, it is not possible to determine the relative fraction of $f^+(y)$ and $f^-(y)$, i.e. the polarisation of the W , since p_T in Eq. 1.14 only carries information about the magnitude of η^0 and not about its sign. Finally, we see how the uncertainty on the mean value of p_T vanishes while enlarging the range of acceptance. From a practical point of view, the reference value for the acceptance in η^L in ATLAS e CMS is $|\eta^L| < 2.4$.

1.3.2 Dependence on W transverse momentum and polarisation

After having discussed a simplified case in which W is produced with $q_T = 0$, we discuss the general case of finite W transverse momentum. In this situation, further polarisation states are possible for W decaying leptonically: all the nine elements of the full W spin matrix. It is possible to define a parametrisation of the W production cross section allowing to disentangle the production of the W

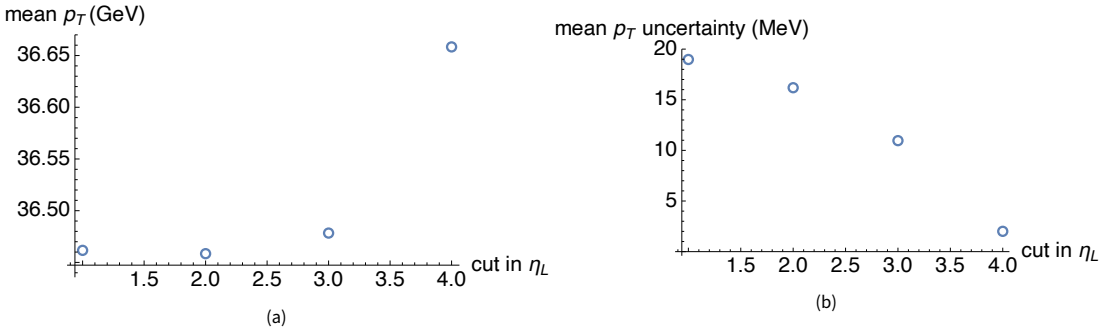


Figure 1.3: The mean of p_T and its uncertainty respectively are computed as a function of an acceptance cut in $\Delta\eta^L$. Using MMHT 2014 PDFs, from [35]

from its decay, in terms of the spherical harmonics [39]:

$$\frac{d\sigma}{dq_T^2 dy d\cos\theta^* d\phi^*} = \frac{3}{16\pi}\frac{d\sigma^{U+L}}{dq_T^2 dy}\Big[(1+\cos^2\theta^*) + \sum_{i=0}^{7} A_i(q_T, y)P_i(\cos\theta^*, \phi^*)\Big], \tag{1.21}$$

$\cos\theta^*$ and ϕ^* being the polar and azimuthal angles of the decay of the muon in a W rest frame. The angular coefficients A_i are functions of the W transverse momentum and rapidity, q_T and y, and together with the unpolarised cross section σ^{U+L}, encode the W production. On the other hand, the functions P_i, spherical harmonics of degree 0, 1, 2 only depend on muon variables and therefore carry information on the kinematics of the detectable muon.

Firstly let us choose and define a W rest frame for carrying on our calculations. The Collins-Soper [40] reference frame is convenient and widely used in literature. In this frame the axes are aligned as follows:

- The $\hat{z}$ axis bisects the angle between $\vec{p}_{p_1}$ and $-\vec{p}_{p_2}$, defined in the Collins-Soper frame, pointing in the direction of q_z^W defined in the laboratory frame, where $\vec{p}_{p_1}$ and $\vec{p}_{p_2}$ 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.

We can now compute the correlation of transverse momentum and pseudorapidity of muons decaying from W . We define the transformations to boost a W in the laboratory frame to the CS frame. The starting four-momentum of the W is $(E_W, q_x, 0, q_z)$, 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 W longitudinal momentum. Setting:

$$\beta_z \equiv -\frac{q_z}{E_W} = -\frac{q_z}{\sqrt{m_W^2 + q_z^2 + q_T^2}}, \quad \gamma_z = \frac{\sqrt{m_W^2 + q_z^2 + q_T^2}}{\sqrt{m_W^2 + q_T^2}},$$

we obtain:

$$(E_W, q_x, 0, q_z) \xrightarrow{\beta_z} (E'_W, q_x, 0, 0), \quad \text{with } E'_W = q_T^2 + m_W^2. \tag{1.22}$$

Let us introduce the W rapidity y in the followings:

$$q_z = \sqrt{m_W^2 + q_T^2}\sinh y \quad E_W = \sqrt{m_W^2 + q_T^2 + q_z^2} = \sqrt{m_W^2 + q_T^2}\cosh y \tag{1.23}$$

A second boost of β_x, γ_x is performed along x, to cancel the W transverse momentum. Setting:

$$\beta_x \equiv -\frac{q_T}{E'_W} = -\frac{q_T}{\sqrt{m_W^2 + q_T^2}}, \quad \gamma_x = \frac{\sqrt{m_W^2 + q_T^2}}{m_W}$$

we get:

$$(E'_W, q_x, 0, 0) \xrightarrow{\beta_x} (E''_W, 0, 0, 0), \quad \text{with } E''_W = m_W. \tag{1.24}$$

We now show that this set of boosts bring us to the Collins-Soper frame. We apply them to the proton beam in the laboratory frame, defined as a photon of energy E_B and momentum $(0, 0, \pm E_B)$. The momentum changes as:

$$(0, 0, \pm E_B) \to (0, 0, E_B(\pm 1 + \beta_z)),$$
$$(0, 0, E_B(\pm 1 + \beta_z)) \to E_B(1 \pm \beta_z)(\gamma_x \beta_x, 0, \pm 1),$$

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.

We can now apply this boost to a massless muon in the laboratory frame with four momentum given by $p_T(\cosh\eta, \cos\phi, \sin\phi, \sinh\eta)$, where p_T and η represent the transverse momentum and the pseudorapidity of the muon, respectively. The angle ϕ corresponds to $\phi^\mu - \phi^W$, i.e. the angle between the muon and the W in the azimuthal direction in the laboratory frame. The boost along z from Eq. 1.22 and along x from Eq. 1.24 are now applied:

$$p_T \begin{pmatrix} \cosh\eta \\ \cos\phi \\ \sin\phi \\ \sinh\eta \end{pmatrix} \xrightarrow{\beta_z} p_T \begin{pmatrix} \gamma_z(\cosh\eta + \beta_z \sinh\eta) \\ \cos\phi \\ \sin\phi \\ \gamma_z(\sinh\eta + \beta_z \cosh\eta) \end{pmatrix} \xrightarrow{\beta_x} p_T \begin{pmatrix} \gamma_x(\gamma_z(\cosh\eta + \beta_z \sinh\eta) + \beta_x \cos\phi) \\ \gamma_x(\cos\phi + \beta_x(\gamma_z(\cosh\eta + \beta_z \sinh\eta)) \\ \sin\phi \\ \gamma_z(\sinh\eta + \beta_z \cosh\eta) \end{pmatrix}. \tag{1.25}$$

The energy of the muon in the CS frame is is given by:

$$\begin{aligned} E^* =& m_W/2 = p_T\gamma_x(\gamma_z(\cosh\eta + \beta_z \sinh\eta) + \beta_x\cos\phi) \\ =& p_T\left(\frac{\sqrt{m_W^2 + q_T^2 + q_z^2}}{m_W}\cosh\eta - \frac{q_z}{m_W}\sinh\eta - \frac{q_T}{m_W}\cos\phi\right). \end{aligned} \tag{1.26}$$

And using the rapidity of the W boson, y:

$$m_W/2 = p_T\left(\frac{\sqrt{m_W^2 + q_T^2}}{m_W}\cosh(y-\eta) - \frac{q_T}{m_W}\cos\phi\right). \tag{1.27}$$

Eq. 1.27 gives the relation between p_T, η and ϕ for all possible values of y and q_T. It 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 section 1.3.1.

In analogy to section 1.3.1, we can compute the fraction of accepted events in the laboratory frame varying the kinematics of the W in the Collins-Soper frame. We can define the phase space in the Collins-Soper frame as a sphere of radius 1 that spans all possible values of $\cos\theta^*$ and ϕ^*, sampled uniformly. We fix values for y and q_T and we apply a realistic acceptance cut of $|\eta^L| < 2.4$ and $p_T > 25$ GeV and select all the couples $(\cos\theta^*, \phi^*)$ on the sphere that satisfy these conditions. In this way the discarded values will be missing from the plot and only the portions of the sphere corresponding to accepted events will be visualised. The shape and dimensions of the "holes" in the plot give us an idea of how y and q_T can distort the acceptance in the laboratory frame. In Figure 1.4 we show the results of the calculations varying y and q_T independently. In Figure 1.4a we have set $q_T = 0$ and computed the fraction of accepted events for a set of values of $y = (0, 2, 3)$. Conversely, in Figure 1.4b we have kept y fixed to 0 and used a set of values $q_T = (5, 20, 40)$ GeV. In both cases, we can can observe how the fraction of accepted values depends on the kinematics of the W boson in a non-trivial way, and in turn the mean value of p_T will be affected by this dependence. Adding a further complication in the calculation, we can imagine to sample $\cos\theta^*$ and ϕ^* according to the spherical harmonics as in Eq. 1.21. In that case the computed acceptance will also depend on the polarisation of the W .

As a last step in our analysis, we will quantify how the fraction of accepted events depends on this. To make this calculation it is sufficient to use Eq. 1.21 setting all A_i to 0 and turning them on one at a time. In Figure 1.5 we report the results of the calculations for A_0 in Figure 1.5a and A_4 in Figure 1.5b. Again, we can notice how the fraction of accepted events depends on the value of the underlying coefficients, which in turn are experimentally inaccessible and have to be given as input from the PDFs and QCD calculations.

To sum up, we have shown how a finite acceptance - which is endemic in all the particle physics

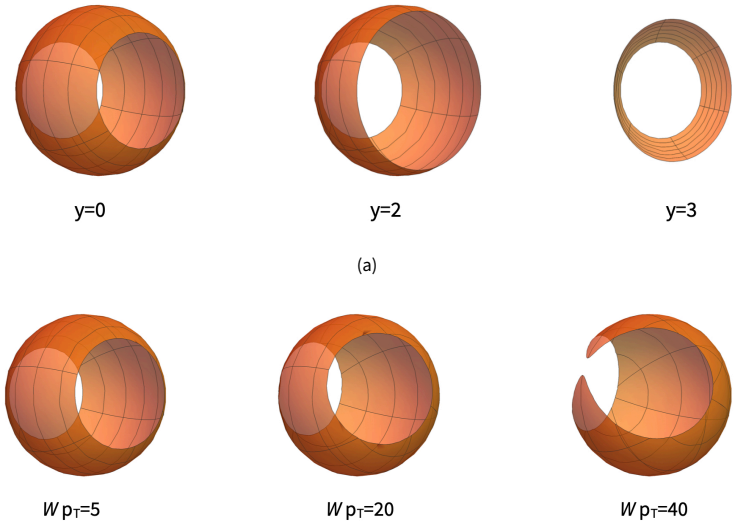


Figure 1.4: Illustration of the phase space in the laboratory at fixed W y and q_T as a sphere of radius 1 and angles θ^L and ϕ^L spanning all possible values of $\cos\theta^*$ and ϕ^*, sampled uniformly. 1.4a shows the case of fixed q_T for some values of y. 1.4b shows the case of fixed y for some values of q_T.

detectors - has an impact on the observables that are used to measure m_W, introducing a dependence on the details of the W production which are only known up to a certain level. This is the major source of uncertainty in the m_W measurement. In the ATLAS measurement, the uncertainty due to PDFs and polarisation amounts to 9 MeV, while the uncertainty on the modelling of W transverse momentum is about 8 MeV.

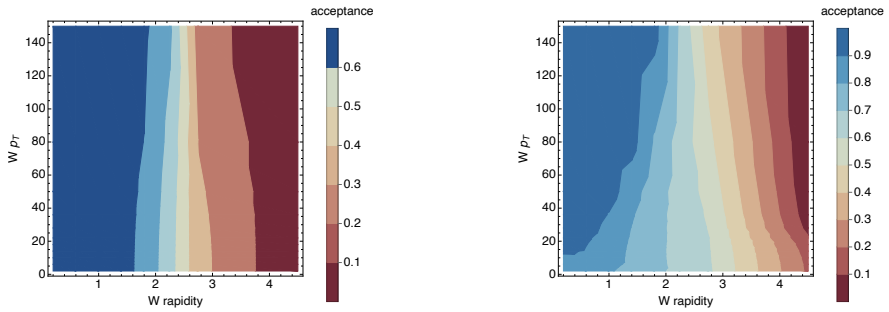


Figure 1.5: Plot of the fraction of accepted events as a function of W y and q_T and polarisation. 1.5a shows the acceptance if only the coefficient A_0 is active, while 1.5a shows the same for A_4.

W transverse momentum from state of the art calculations

At LHC the transverse momentum distribution of produced W has its bulk at around 5 GeV. This regime of energy is purely non perturbative and fixed order calculations fail in giving a reliable prediction. Moreover, it is not possible to obtain a precise measurement of this quantity. In principle, the W recoil is sensitive to q_T but in practice due to pile up its resolution is too poor to measure with a granularity better than $\simeq$ 8 GeV [41, 42]. CMS and ATLAS are studying special low pileup runs provided by LHC during 2017, with an integrated luminosity of about 200 pb^{-1}, and $\sqrt{s} = 13$ TeV, to assess the feasibility of such a measurement.

On a theoretical point of view, resummation methods have been devised to give a reliable prediction of q_T [43, 44]. At the present time, the state of the art of resummation is N^{3}LL+NNLO, where the fist term indicates the third order from the leading logarithmic term in resummation calculation. The result from [45] is shown in Figure 1.6. The uncertainty ranges between 6% and 4% for $q_T < 5$ GeV, and at is $\sim$ 2% at higher q_T.

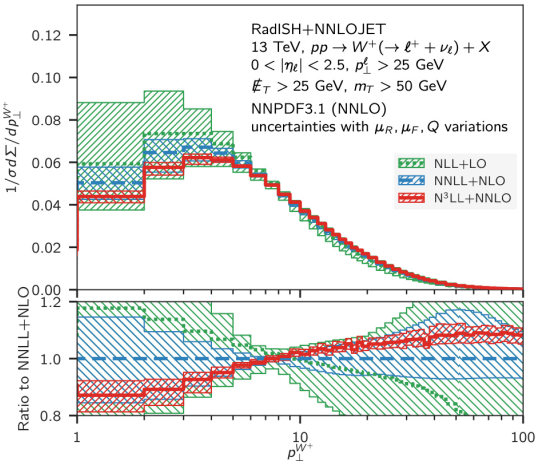


Figure 1.6: Comparison of the normalized q_T^W distribution for W^+ at $\sqrt{13}$ TeV at NNLO, NNLL+NLO and N^{3}LL+NNLO. The fiducial selection $p_T^\ell > 25$ GeV, $p_T^{\text{miss}} > 25$ GeV, $|\eta^\ell| < 2.5$, $m_T > 50$ GeV is applied (from [45]).

We now show that this prediction produces a distortion of the p_T spectrum compatible with a m_W shift much larger than 10 MeV. In Figure 1.7 the shape of the nominal p_T spectrum is compared to three relevant variations: a variation of q_T spectrum of 4% in the range $0 < q_T < 5$ GeV, compatible with the state-of-the-art uncertainty; a variation of the value of m_W shifted of $\pm$10 MeV, the target precision; a variation of the PDF according to their nominal uncertainty. Looking at the variation of the average momentum as a proxi of m_W, we can observe how the injected m_W variation produces an absolute shift of 3 MeV, while q_T of 11 MeV and The PDF uncertainty of 14 MeV.

This demonstrates that measuring m_W with the currently available inputs from theory does not guarantee the achievement of the desired precision.

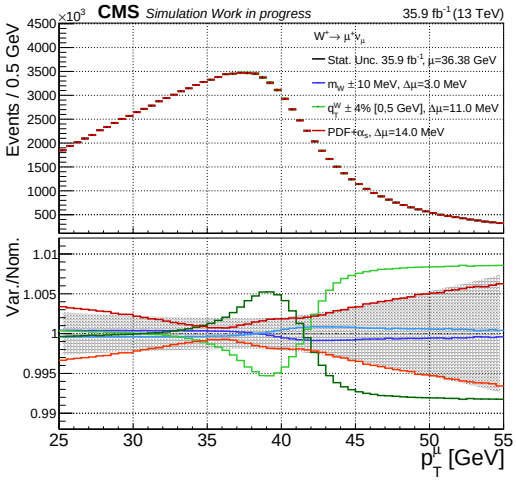


Figure 1.7: The p_T^μ spectrum from $W^+ \to \mu^+\nu$ simulated events, the effects of a ± 10 MeV m_W variation, of a flat 4% variation of the q_T^W below 5 GeV and the variation of the PDF within the current uncertainty, symmetrised around the nominal value, are shown. For each variation is also reported 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. The statistical uncertainty correspondent to a integrated luminosity of 35.9 fb^{-1} is shown as a grey band. From [46].

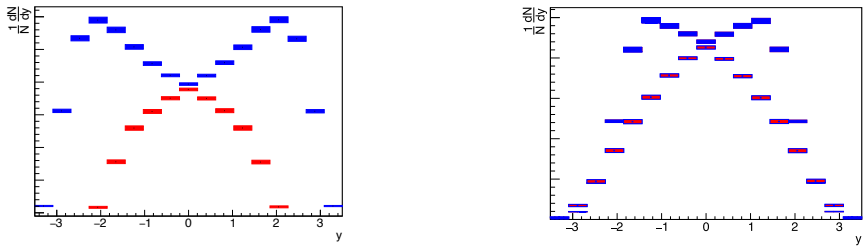
1.3.3 How to measure the W boson production

We now show how an original method devised in the context of this work allows to assess experimentally W y, q_T distributions and polarisation. This solves the issue of the dependence of the measured value of m_W from theory inputs. This section is organised as follows: in the first part we will present a procedure that allows to measure the rapidity and polarisation in the approximation in which W is produced with small transverse momentum. This is an original idea from the author of this dissertation that has been published in [47]. Then, we will extend these concepts to determine the transverse momentum of the W together with its rapidity and polarisation.

Experimental determination of W longitudinal motion

When a W is produced with small transverse momentum, its principal production mechanism involves the leading order parton subprocesses $u\bar{d} \to W^+$ and $\bar{u}d \to W^-$. Due to the valence quark content of the proton, it is likely that the W will be moving in the direction of the initial-state quark, as opposed to antiquark. Since the electroweak charged current totally violates parity, the quark must be left-handed and the antiquark right-handed. For a simple argument of angular momentum conservation [48], the direction of the spin of the W must preferencially be opposite to its direction of motion, with a small dilution occurring in case the antiquark is carrying a higher fraction of the proton momentum than the quark. This argument includes cases when a W is produced by a sea quark-antiquark pair. In this case the entity of the dilution will be higher and more W will be produced with the spin of the W along its direction of motion. In Figure 1.8 the rapidity y distribution of the W bosons for the two left and right helicities is shown. Due to the symmetry of the LHC beams they are symmetric.

A W produced in this configuration will show a very asymmetric decay: a left-handed W^+ tends to

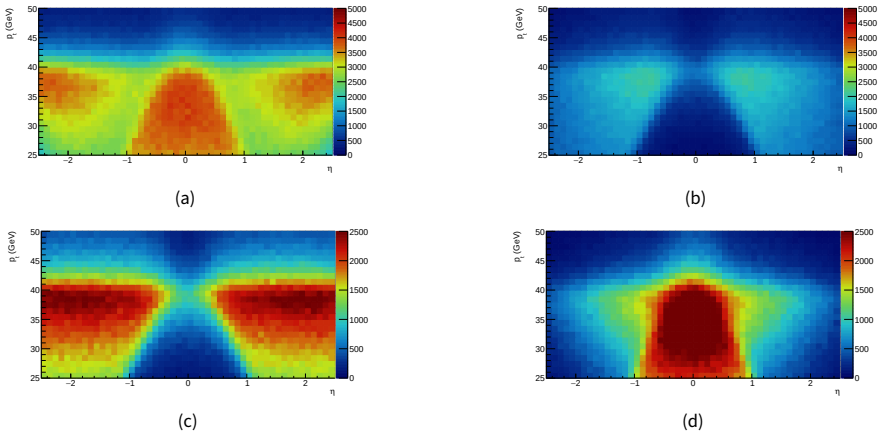


(a) (b)

Figure 1.8: W rapidity distribution for the positive (red) and negative (blue) helicity as computed using the NNPDF3.1 PDFs. a) refers to W^+ and b) refers to W^-. From [47].

decay with the left-handed neutrino forward and the right-handed charged lepton backward, while a left-handed W^- tends to decay with the left-handed charged lepton forward and the right-handed neutrino backward. Thus, the direction of the lepton momentum is strongly correlated with the direction of the spin of the W .

In [47] we have studied how the correlation of p_T and η can be exploited to access the hidden y and helicity distributions of W bosons, given these observations. Looking at Figure 1.10, it can be appreciated how the correlation plots of p_T and η show differences in shape for the two helicity states of the W . Considering W^+ events in the negative helicity state, when they are produced at positive y the pseudorapidity of leptons from their decay is typically $\eta = y - 0.5$, while when they are produced at negative y the pseudorapidity is typically $\eta = y + 0.5$. The two peaks of figure 1.8a are therefore shifted by half a unit toward $\eta = 0$ and thus there is an accumulation of events near $\eta = 0$ in the panel 1.9a. The W^+ events with positive helicity have the opposite behaviour. When produced at positive y, the typical pseudorapidity of the lepton is $\eta = y + 0.5$. Conversely, when they are produced at negative y the pseudorapidity is typically $\eta = y - 0.5$. This explains the reduction of events near $\eta = 0$ in panel 1.9b. The W^- plots can be interpreted in a similar way.



(a) (b)

(c) (d)

Figure 1.9: p_t vs η distribution for leptons from W decays. a) W^+ with negative helicity b) W^+ with positive helicity c) W^- with negative helicity d) W^- with positive helicity. From [47].

Therefore, these plots, made using observables that can be measured, can be used to discriminate the W polarisation, which is a quantity not directly detectable. Going forward, if we reproduce the plots of Figure 1.9 in a small bin of y, which is feasible using a simulation, we can exploit Eq.1.13 to discriminate y in addition to W polarisation. A practical recipe for this is thoroughly discussed in [47] and Figure 1.10 shows its final result compared with the PDFs prediction. It is argued that after taking into account the systematic effects, the statistical precision of this analysis has a substantially smaller uncertainty than the PDFs prediction.

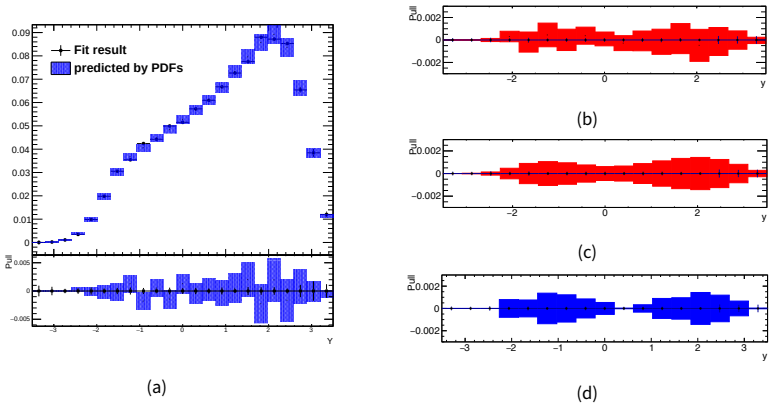


Figure 1.10: a) The result of the fit is compared to the PDFs prediction. b) Difference between fit and PDFs prediction. c) Difference between fit and PDFs prediction shifted at the same central value of the fit. d) Shape variation when modifying the q_T spectrum. From [47].

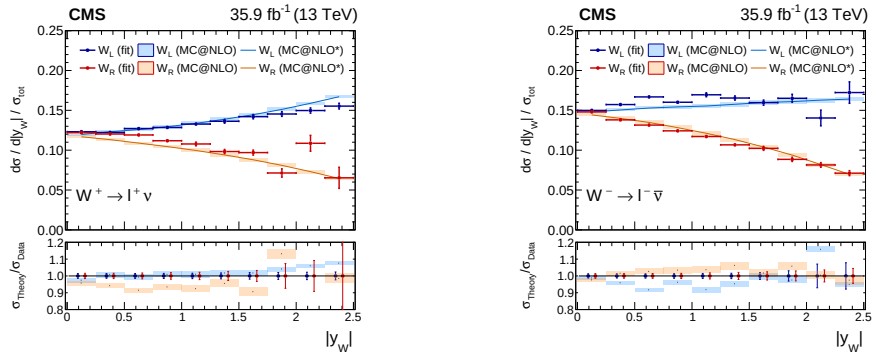
To sum up, for the first time it is argued that the correlation between the transverse momentum and the pseudo-rapidity of charged leptons emitted from a W decaying semi-leptonically can carry information about the kinematics of the original W boson. Or, in other words, that an observable at LHC experiments can be used to unfold variables, rapidity and polarisation of W , not directly available to measure.

Following these prescription, CMS has delivered a measurement of the W rapidity and helicity. It 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. In addition to [47], a small component due to the W longitudinal polarisation has been added in the fit, fixed to the simulation prediction. The results of this analysis are reported in Figure 1.11, where it can be appreciated how the final experimental uncertainty on the fitted distribution is smaller than the prediction from `MadGraph5_aMC@NLO` simulation.

Experimental determination of W longitudinal and transverse motion, and its polarisation

It comes naturally to apply the same idea discussed in [47] to the general case in which W are produced with longitudinal and transverse momentum.

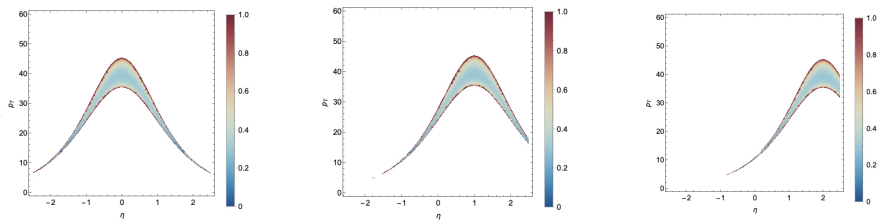
Using the formalism developed in section 1.3.2 we can plot the correlation between η and p_T in the laboratory frame for fixed values of q_T and y of the W boson and assuming a flat distribution of ϕ^* and $\cos\theta^*$ in the Collins-Soper frame. In the following we will drop the superscript L in η^L. We set as range $\eta = (-2.4, 2.4)$ and $p_T = (0, 60)$ GeV. In Figure 1.12 we show the correlation plot for $q_T = 10$



(a) (b)

Figure 1.11: 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 $W^- \to \ell^-\bar{\nu}$ (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, μ_F and μ_R scales variation. The central value of the prediction is shown as a darker line (from Ref. [49]).

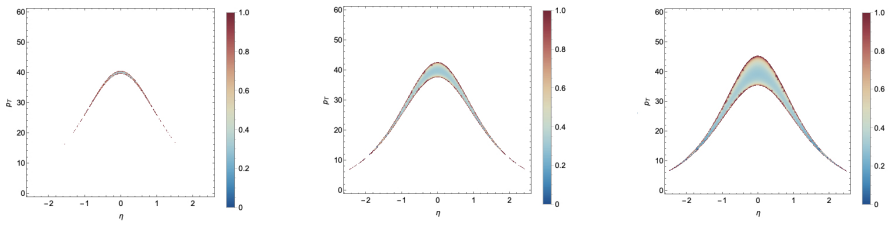
GeV and a set of y values: $(0, 1, 2)$. As expected from equation Eq. 1.27 the plot shifts from $\eta = 0$ as the absolute value of y increases. More interesting is figure Figure 1.13 that shows what happens if we keep a constant $y = 0$ and plot a set of q_T values $(1., 5., 10.)$ GeV. In this case we can observe how the width of the plot becomes thicker as the absolute value of q_T grows. This is due to the $\cos\phi$ dependence of equation Eq. 1.27 which represents the cosine of $\phi^\mu - \phi^W$, which is not experimentally detectable and that is integrated upon. When this dimension gets projected in the η and p_T plane, the plot thickens as the amplitude of the oscillation grows, as shown in Figure 1.14. This figure shows also that the population is maximal at the edges of the η, p_T contour due to the jacobian of the integration in ϕ. Moreover, given the structure of Eq. 1.27, two events with same value of $\cos\theta^*$ but opposite $\phi^\mu - \phi^W$ will be mapped in the same point of the η,p_T plane. This means that all the harmonics in Eq. 1.21 containing a function $\sin\phi$ will have a null density on the η,p_T plane. This is the case for the harmonics related to A_5, A_6, A_7.



(a) (b) (c)

Figure 1.12: Correlation plots in η and p_T plane. 1.12a: $y = 0$ and $q_T = 10$ GeV. 1.12b: $y = 1$ and $q_T = 10$ GeV. 1.12c: $y = 2$ and $q_T = 10$ GeV.

If we go further in this analysis we can add a polarisation component to the W using Eq. 1.21, turning on one A_i at a time and setting it to 1 while keeping all the others at 0. This will change the distribution in the CS plane and the population in the η, p_T plot, but not its contour. In this way we can produce for each y and q_T a peculiar set of plots reflecting the characteristics of the muon decay for each spherical



(a) (b) (c)

Figure 1.13: Correlation plots in η and p_T plane. 1.13a: $y = 0$ and $q_T = 1$ GeV. 1.13b: $y = 0$ and $q_T = 5$ GeV. 1.13c: $y = 0$ and $q_T = 10$ GeV.

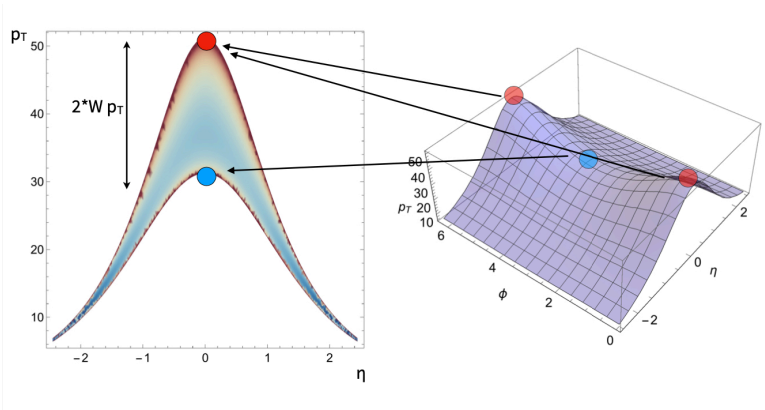


Figure 1.14: Illustration showing the effect of the projection of the angle $\phi = \phi^{\mu} - \phi^{W}$ in the η and p_T plane.

harmonic. They are shown in Figure 1.15 for the coefficients A_0, A_1, A_2, A_3, A_4. If for example we pick A_4, in Figure 1.15e, this depends on the $\cos\theta^*$ variable alone and encodes the complete violation of parity in the W decays. This can be observed in the plot as an accumulation of events in the right part of η, p_T contour. In the other plots similar features can be observed, which can allow to discriminate q_T, y and polarisation of the W if a set of templates is produced and a fit similar to the one shown in [47] is performed.

To sum up, carrying out the complete calculation for the correlation in the η and p_T plane as a function of the kinematics of a W and the muon from its decay in the CS frame, we have extended the method introduced in [47] to unfold not only the rapidity of the W and its polarisation, but also its transverse momentum. This procedure will be applied to real data collected by the CMS experiment to carry out a measurement of the distributions of W rapidity and transverse momentum and the angular coefficients in Chapter 5.

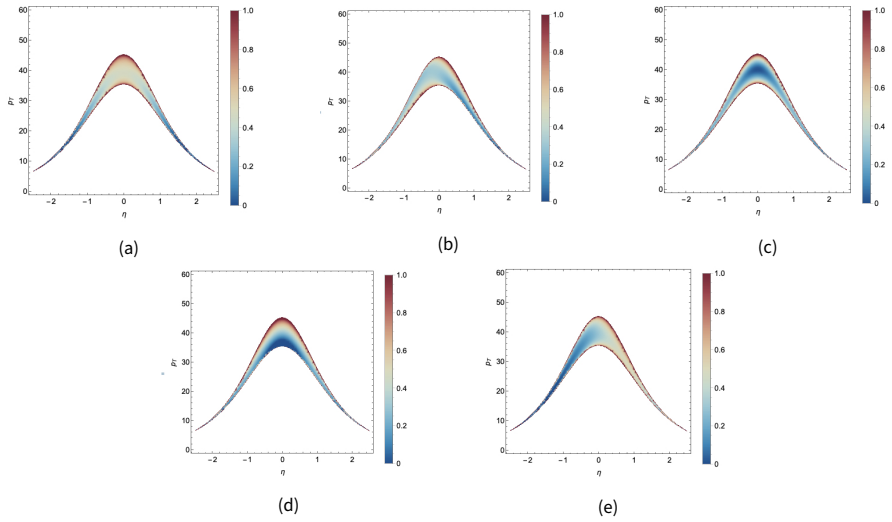


Figure 1.15: Correlation plots in η and p_T plane, where a polarisation component to the W using Eq. 1.21, turning on one A_i at a time and setting it to 1 while keeping all the others at 0. 1.15a shows A_0, 1.15b shows A_1, 1.15c shows A_2, 1.15d shows A_3, 1.15e shows A_4.

Chapter 2

The CMS experiment at LHC

This chapter contains the essential background to the facilities used in the analysis shown throughout this thesis. The first two sections give an overview on the Large Hadron Collider (LHC) and the characteristics of the physics events produced in its interaction points. The third section describes briefly the Compact Muon Solenoid (CMS) experiment, whose collected data have been employed for this work. The last sections deal with the tracking system of CMS, whose details are essential to understand the core of the following chapters, and the reconstruction of muons, which constitute an important tool in the CMS physics program.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a circular proton-proton collider located at the European Organisation for Nuclear Research (CERN).

It is contained into a circular tunnel of 27 km of circumference, inside which, in two adjacent parallel beam pipes, the proton beams circulate in opposite directions in ultra-high vacuum. The beams are guided inside the pipes by a strong magnetic field of 8.33 T provided by 1232 superconducting dipole magnets. Moreover, 392 quadrupole magnets are used to focus and steer the beam.

The protons circulating in the pipes are gathered in bunches of $\sim 10^{11}$ protons. Each bunch is accelerated crossing 8 radio frequency cavities per turn, the field of the radio frequency being synchronised with the spacing between the bunches. Since for relativistic particles the orbit grows with their energy, the magnetic field must grow synchronously with the energy gained at each stage.

Once the magnetic field reaches its maximum value, the beams are brought into collision at four points around the ring, which host the four experiments of LHC: ATLAS [50], CMS [51], LHCb [52] and ALICE [53]. The first two are multi-purpose experiments, designed to search for the Higgs boson and new particles with masses at the TeV scale. LHCb studies the properties of charm and beauty hadrons and ALICE analyses the data from relativistic heavy ion collisions to study the hadronic matter in extreme temperature and density conditions.

Since the LHC can not accelerate the protons from zero energy, various preliminary steps are necessary before the injection in the main accelerator. The full acceleration facility is composed of a linear accelerator (LINAC) and a chain of three synchrotrons: Booster, the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS) where they are accelerated up to 450 GeV. The scheme representing the full CERN acceleration complex is shown in Figure 2.1

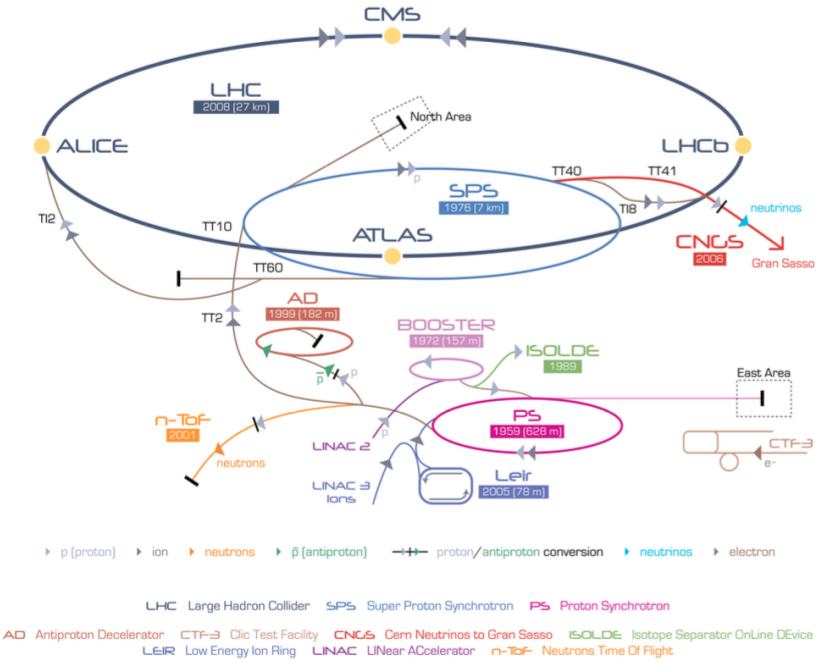


Figure 2.1: Scheme of the facilities of the CERN acceleration complex.

The LHC has been designed to collide protons at a nominal center-of-mass energy $\sqrt{s} = 14$ TeV. Another important aspect is the collision rate, which is proportional to the instantaneous luminosity $\mathcal{L}$:

$$\mathcal{L} = \frac{f k n_p^2}{4\pi\sigma_x\sigma_y}, \tag{2.1}$$

where f is the bunch revolution frequency, k the number of bunches, n_p the number of protons per bunch and σ_x, σ_y their transverse dispersion along the x and y axis. At the nominal 14 TeV LHC conditions ($\mathcal{L} = 10^{34}$ cm^{-2}s^{-1}) the parameter values are: $k = 2808$, $n_p = 1.5\times10^{11}$ and $\sigma_x = \sigma_y = 16.6$ μm (with $\sigma_z = 7.6$ cm along the beam).

The rate of production of a given process is proportional to $\mathcal{L}$ through the cross section σ. Therefore, if the process is very rare, it is necessary to maximise $\mathcal{L}$ to collect enough statistics. The integrated luminosity is defined as $L = \int \mathcal{L} dt$ and it is usually quoted to show the amount of data available for the analyses.

The work described in this thesis was carried out on data collected by CMS and therefore this chapter will give a more detailed overview of this experiment.

The LHC started its research program in spring 2010 at a center-of-mass energy of 7 TeV, and CMS collected a total integrated luminosity of 5.6 fb^{-1} with a record peak luminosity of $4.0 \cdot 10^{33}$ cm^{-2} s^{-1}. In 2012 the center-of-mass energy has been increased to 8 TeV. CMS collected 22 fb^{-1} with a record peak luminosity of $7.7 \cdot 10^{33}$ cm^{-2} s^{-1} until the beginning of 2013, when the LHC has been shut down

to prepare the Run 2, at increased center-of-mass energy and luminosity.

LHC was reactivated in 2015 running at 13 TeV energy in the center of mass till the end of Run 2 in 2018. LHC is now restarting for Run 3 after the pandemic.

2.2 Phenomenology of pp collisions

The proton-proton (pp) interaction at the intersection points of LHC is a complex phenomenon that involves the strong interaction of quarks and gluons composing the protons, also called partons. The interesting part of the collision is called hard scattering and it consists of the interaction of two partons at high transferred momentum q^2.

Along with this, the event is accompanied by several other processes: the initial and final state radiation of gluons from the quarks and the soft scattering (i.e. at low q^2) of the remnants of the protons. All these interactions together are designated underlying event. Due to colour confinement[1], the quarks created in the scatterings must combine with other quarks and antiquarks created from the vacuum and they form composite particles, hadrons. This process is known as hadronisation and it is not fully understood since it involves non-perturbative QCD. However, there are some models and parametrisations that describe it quite accurately and are used in the Monte Carlo simulations.

In addition to the underlying events, pp interactions at LHC are characterised by the presence of pileup events. The pileup is constituted of minimum bias events (i.e. events with no hard interaction) among protons in the same bunch crossing that cause the presence of many low energy particles.

Figure 2.2 summarises with an illustration the complex of pp interactions.

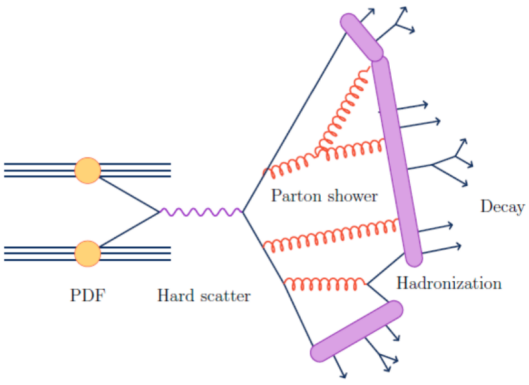


Figure 2.2: Illustration of the hard scattering process, parton shower, hadronization and decay during the generation of an event with two quarks in the final state. Figure from [54].

[1] Colour confinement is a feature of the QCD interaction that states that only non-QCD coloured objects can exist individually.

2.3 The CMS experiment

The Compact Muon Solenoid (CMS) detector is roughly cylindrical, 21.5 m long and with a radius of 7 m. It was designed around a superconducting solenoid providing a 3.8 T magnetic field. An overview of the detector is given in Figure 2.3.

The reference frame used to describe the CMS detector and the collected events has its origin in the geometrical center of the solenoid. In a right-handed Cartesian coordinate system, the x axis points to the center of the LHC ring, the y axis points upwards, perpendicular to the LHC plane and the z axis points towards the beam line. A cylindrical coordinate system is more often used. The $r = \sqrt{x^2 + y^2}$ coordinate points from the geometrical center of the cylinder outwards, and the two angles ϕ and θ are such that $\tan\phi = y/x$ and $\tan\theta = r/z$. Instead of the angle θ, the pseudorapidity, defined as:

$$\eta = -\ln\tan\frac{\theta}{2} \tag{2.2}$$

is more often used, since it is additive under boosts along the z axis.

In this chapter, it is given a brief overview of the CMS detector, with particular emphasis on the subdetectors which are used in the analysis presented in this thesis. A more complete description of the CMS detector can be found in [51].

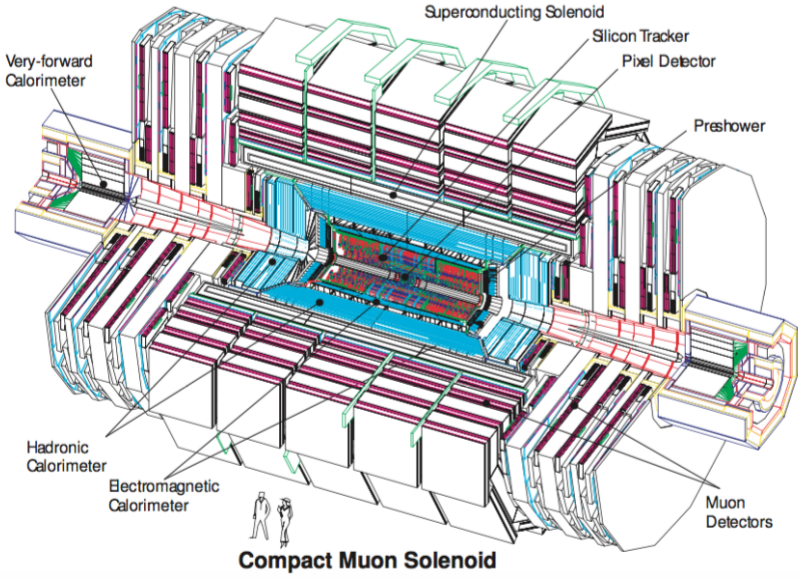


Figure 2.3: Overview of the CMS detector. Figure from [51].

2.3.1 Overview of the detector

The CMS detector consists of a cylindrical barrel up to $|\eta| < 1.2$ and the endcaps from $|\eta| = 1.2$ on. This two parts almost cover the whole solid angle, up to $|\eta| \sim 5$. The main detector is composed of

several subdetectors that form layers at increasing values of r in the barrel and increasing values of $|z|$ in the endcaps. Each layer has a different role in the detection and identification of the particles generated in the collisions.

The experiment is built around the solenoid magnet that takes the form of a cylindrical coil of superconducting cables. The whole structure is supported by a steel yoke that forms the bulk of the detector and confines the magnetic field. The tracker, the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL) are located inside the solenoid, while in the outside the magnetic flux is returned by an iron structure. Here the muon chambers are hosted, which constitute the last layer of the CMS experiment.

Magnet

The superconducting magnet of CMS has been designed to reach a 4 T field in a free bore of 6 m diameter and 12.5 m length, for a stored energy at full current of 2.6 GJ. The flux is returned through a 10k-tons yoke comprising 5 wheels and 2 endcaps, composed of three disks each. The magnet was designed to be assembled and tested in a surface hall (SX5), prior to being lowered 90 m below ground to its final position in the experimental cavern. After provisional connection to its ancillaries, the CMS Magnet has been fully and successfully tested and commissioned in SX5 during autumn 2006. Figure 2.4 shows the intensity of the magnetic inductance $\vec{B}$ for various subsystems of the CMS detector. The return flux in the barrel is between 1 and 2 T within the iron elements and typically below 0.1 T in the muon chambers (except for a peak of 0.3 T near the air gaps interspersed in the yoke wheels). Inside the volume of the solenoid bore B varies between 4 T in the central region and about 3 T in the sides. In the region occupied by the tracker, B is relatively uniform but not constant, gradually lowering in value in the endcaps.

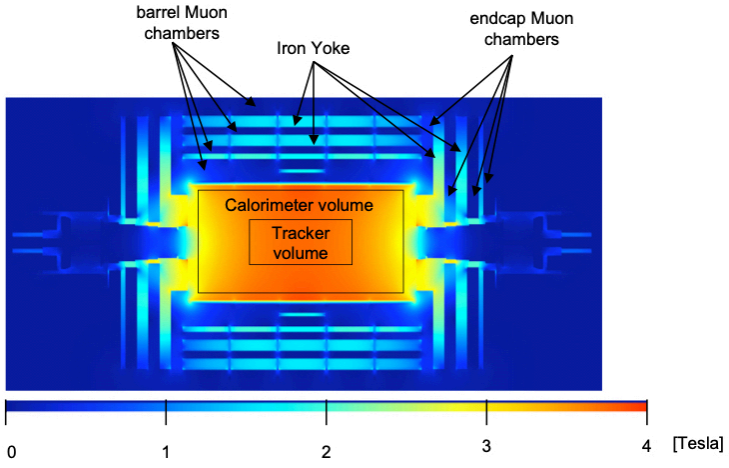


Figure 2.4: Longitudinal R-z view of CMS showing the module of the magnetic inductance B within different subsystems of the detector. Figure from [55].

Tracker

The tracker is the innermost subdetector and the closest to the interaction point. It measures 5.4 m in length and 1.1 m in radius and its coverage extends up to $|\eta| = 2.5$. It is used to observe the charged particles and measure their momentum from their curvature in the magnetic field. The tracker is equipped with silicon detectors: pixel modules in the inner region, to provide enough granularity close to the beam spot, and microstrip modules in the outer region.

A charged particle traversing the modules of the tracker knocks electrons off the valence band, thus creating electron-hole pairs. In the presence of an externally applied electric field the electrons and the holes are separated and collected by the electrodes, producing a signal proportional to the energy lost by the passing particle. In the language of tracking, a signal left in one module is called hit. Through algorithms that will be described in more details in section 2.4.2, the most probable trajectory of a charged particle in the tracker is fitted to a set of hits, allowing for the precise measurement of the momentum.

Calorimeters

The calorimeters are designed to absorb the particles to measure their energy. The ECAL is an homogeneous calorimeter, where the absorber material is the same as the sensitive one, made of lead tungstate ($PbWO_4$). It contains the electromagnetic showers of charged particles and photons and it produces light proportional to the energy of the initial particle, which is read out by photomultipliers. The ECAL has a barrel section and two endcaps for a total coverage up to $|\eta| = 3$. The thickness is 23 cm in the barrel and 22 cm in the endcaps.

The HCAL is designed to detect and absorbs hadrons. It is composed of layers of brass to stop the hadrons interleaved with tiles of plastic scintillators, whose signal is read out by photodiodes. Also the HCAL has a barrel section and two endcaps for a total coverage up to $|\eta| = 3$ and its thickness in the barrel amounts to 1.2 m.

The calorimeters are complemented in the high pseudorapidity region with two forward hadronic calorimeters (HF) which extend the coverage to $|\eta| = 5$. They consist of layers of steel and quartz read out by photomultiplier tubes.

Muon chambers

The muon system is equipped with various kinds of technologies. In the region $|\eta| < 1.3$ the muon tracking system is made of drift tubes and in the region $0.9 < |\eta| < 2.4$ of cathode strip chambers. Resistive plate chambers are installed in the region $|\eta| < 2.1$ to provide an additional trigger system. A more detailed description of the muon system is given in section 2.5.1.

2.3.2 Trigger

The trigger is an essential tool in the LHC experiments, as the amount of data produced per second is huge. It is then necessary to introduce the trigger, which makes an on-line choice on the data to be kept by the experiment on the basis of their physics content.

In CMS the triggering system is divided in two steps. The first one is the Level-1 Trigger (L1), implemented only in hardware. It exploits the information from the calorimeters and the muon chambers to identify possible interesting processes. This first step is completed in about 12.5 μs, and reduces the rate down to 100 kHz.

If an event passes the L1, it is passed to the High-Level Trigger (HLT), which exploits a simplified form of the software used to off-line analyses for partially reconstructing the event and therefore choosing the most interesting ones. The HLT reduces the event rate further down to about 1 kHz.

2.4 The tracking system of CMS

The tracker is an essential tool in a multi-purpose detector like CMS, as the reconstructed trajectories of the charged particles are a vital ingredient in the measurement of the majority of observables used in physics analyses.

2.4.1 The CMS tracker in a nutshell

The CMS tracker is entirely based on silicon detector technology. While a large number of read-out channels is desirable to increase the granularity of the system, the material used for sensors, electronics, support structures and services must be light enough in order to interact as little as possible with the particles produced in the collisions.

The tracker is composed of 1440 silicon pixel modules and 15148 silicon microstrip modules organised in layers around the interaction point. The regions very close to the interaction point have a higher density of particles and therefore high granularity pixel detectors are needed. The intermediate and outer regions, where the density of particles is reduced, are equipped with microstrip detectors. An overview of the tracker is given in Figure 2.5. In 2017 the pixel detector was upgraded to a new design with an increased number of layers [56].

Pixel detectors

The pixel modules are arranged around three concentric layers installed at radii 4.4 cm, 7.3 cm and 10.2 cm. They are completed to each side of the barrel by two endcaps consisting in 2 disks of pixel detectors, extending from 6 cm to 15 cm in radius at 34.5 cm and 46.5 cm from the nominal interaction point. This ensures the existence of three measurement points for each track over almost the whole acceptance. All the pixel sensors have a cell size of 100×150 μm^2 with the larger side along the z coordinate, and they can deliver a measurement of the hit position in the three coordinates r, ϕ and z.

The charge carriers that produce the signal in the pixel sensors are the electrons. The effect of the magnetic field in the barrel (Lorentz angle, described in [55]) and the special arrangement with slightly tilted modules in the endcaps ensures the charge signal to be spread over more than one pixel. This permits to reduce the spatial resolution to ~ 15 μm for high momentum tracks, which are about radial.

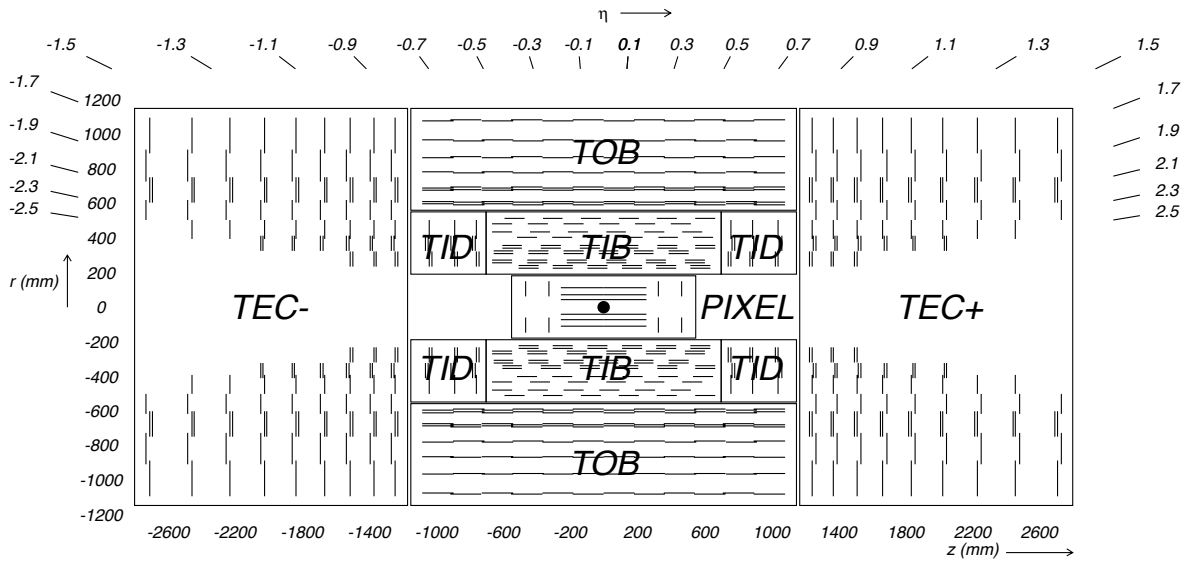


Figure 2.5: Schematic cross section of the CMS tracker in the $r - z$ plane. The strip tracker modules that provided 2D hits are shown by thin, black lines, while those allowing the reconstruction of the third coordinate are shown by closely spaced double lines. The pixel modules also give 3D hits. Within a given layer, each module is shifted slightly in r or z with respect to its neighbours to avoid gaps in the acceptance. Figure from [51].

Microstrip detectors

Further away from the interaction point the silicon strip tracker is installed. The inner silicon strip tracker is composed of the tracker inner barrel (TIB) with four layers and the tracker inner disks (TID) with three endcap layers. The outer part consists of the tracker outer barrel (TOB) composed of six layers and the tracker endcaps (TEC) composed of nine disks. Within a given layer, each module is is shifted slightly in r or z with respect to its neighbour, and the modules overlap, to avoid gaps in the acceptance.

The typical silicon strip module size is 10 cm $\times$ 5 cm with a strip pitch of 80 μm in the inner regions and 20 cm $\times$ 10 cm with a strip pitch of 140 μm in the outer regions. In the barrel, the strips are parallel to the z axis and in the endcap they are placed along the radial coordinate. The modules provide a measurement of the $r - \phi$ coordinate with a resolution of 20 - 50 μm.

To measure the z coordinate in the barrel and the r coordinate in the endcap with a precision better than the strip length some layers have an additional set of modules, tilted with respect to the original ones by a stereo angle of 100 mrad. The measurement of the third coordinate is obtained through the matching of the hits measured by the tilted modules with the ones obtained with the regular ones.

2.4.2 Algorithms for track reconstruction

The procedure of fitting the trajectory of charged particles to the hit pattern is called track reconstruction.

In a constant magnetic field, charged particles travel into the tracker on a helicoidal trajectory, de-

scribed by 5 parameters: the curvature [2] $k = q/p_T$, the azimuthal and polar angles ϕ and η, the transverse impact parameter d_{xy} and the longitudinal impact parameter d_z. The impact parameters d_{xy} and d_z are respectively the r and z coordinate of the track at the point of closest approach to the beamline. The standard algorithm used in CMS for track reconstruction is the Combinatorial Track Finder (CTF) algorithm [57], which is developed in three steps: track seeding, track finding and track fitting.

The track seeding consists in a loop on all pairs of hits compatible with some kinematical cuts. The seeding starts from the innermost pixel detectors since the high resolution on the position of the hit reduces the number of options to consider.

The track finding and fitting steps are based on a standard Kalman filter pattern recognition approach, starting with the seed parameters. This method is used to estimate the state of a dynamic system from a series of measurements with corresponding uncertainties. In the track reconstruction the state is given by the helicoidal parameters that are propagated layer by layer with an inside-out approach and fitted to the hits with a χ^2 fit. At each step the computation of the parameters is updated with local information on the magnetic field, the spatial uncertainty and the quantity of material crossed by the particle. The latter is shown in Figure 2.6 expressed in terms of fractional radiation lengths, is taken from the simulation.

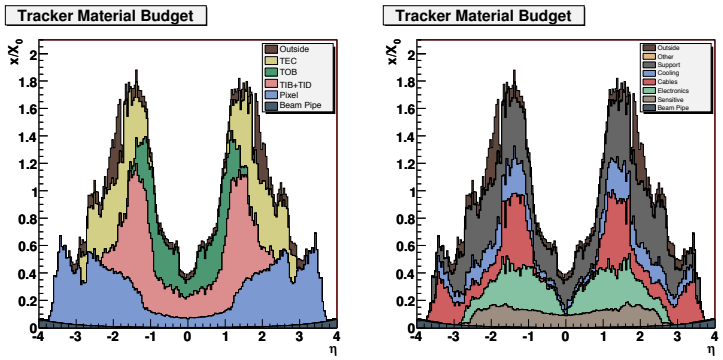
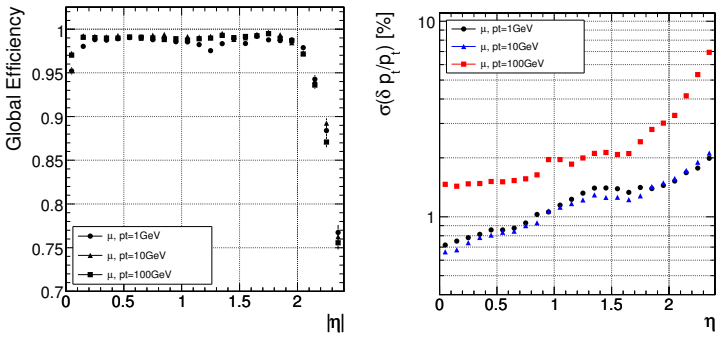


Figure 2.6: Material budget seen by a particle produced in the center of CMS and crossing the whole volume of the Tracker. The material is expressed in terms of fractional radiation lengths as a function of the particle pseudorapidity and shown for the different sub-detectors (left panel) and broken down into the functional contributions (right panel). Figure from [51].

The tracks are assigned a quality based on the χ^2 and the number of missing hits, and only the best quality tracks are kept. Once the whole information is available, the Kalman filter is re-run with an outside-in approach. The final output is the full set of helix parameters describing the trajectory at the innermost measurement point and the full covariance matrix describing the correlations of the parameters and their uncertainties. Figure 2.7 shows the expected muon reconstruction efficiencies and muon transverse momentum resolution for the CMS tracker.

In a simplified approach of a helix in a constant magnetic field and no energy loss, the bending radius of the trajectory is linked to the transverse momentum by:

[2] The transverse momentum p_T is the component of the momentum projected along the transverse plane with respect to the beamline.



(a) (b)

Figure 2.7: Global track reconstruction efficiencies (a) and resolution of transverse momentum (b) for single muons with transverse momenta of 1, 10 and 100 GeV, as computed from simulation before the startup of the experiment in the year 2008. Figures from [51].

$$p_T \text{ [GeV]} = 0.3B\rho \text{ [Tm]} \tag{2.3}$$

and to the sagitta by:

$$s = \frac{L^2}{8\rho} = \frac{0.3BL^2}{8p_T}, \tag{2.4}$$

where L is the length of the track measured on the transverse plane. The sagitta of a track of 100 GeV is about 1.4 mm, while a track of about $p_T = 0.6$ GeV curls inside the tracker.

Tracker alignment

One major challenge in tracking is represented by the alignment of the modules of the tracker. In order to properly reconstruct particle trajectories, the relative positions of the tracker components must be known to a precision better than the intrinsic resolution of the modules.

The alignment procedure is run periodically in the CMS tracker. This optimisation problem can be formulated in the context of linear least squares. The goal is to derive a set of alignment parameters, the modules position corrections, minimising through a χ^2 fit the distance between the track impact point and the related hit. In practice the reconstructed tracks are made better fit the hits by adjusting the assumed positions of the modules. To this purpose, the Generalised Broken Line Fit is introduced. This is a track refit that is able to incorporate the description of the multiple scattering in the track model, thus allowing to determine the covariance matrix of all track parameters. For this reason, it is particularly suitable to fit the tracker alignment parameters. Its original formulation [58] treats the approximate case of a tracking detector with a solenoidal magnetic field and independent bidimensional tracking in the bending and its perpendicular plane. Moreover, the planar trajectories are constructed from the measurements including the material around them as thick scatterers. In [59] this procedure is extended to give a single trajectory to bidimensional measurements and the material between adja-

cent measurements is described using up to two thin scatterers of null thickness, with the same mean and variance of the material distribution. Such thin scatterers produces no spatial shift, but a scattering angle with variance given by the expected variance θ_0^2 of the multiple scattering in that mean. At each measurement plane and thin scatterer, a local orthonormal coordinate system (u, v, w) is defined. The the w-axis is taken to be perpendicular to the sensor plane for a measurement and parallel to the track direction for a scatterer. At each thin scatterer the offset (u, v) is a fit parameter together with a curvature correction Δk, and the trajectory is constructed from the thin scatterers adding the measurements and interpolating the enclosing scatterers. The procedure consists of various steps. The first step is to adjust the absolute position and orientation of the tracker relative to the magnetic field. This is done using two alignment parameters that describe the tilts around the x and y axes. After that, the individual modules are aligned. The position of each module is parametrised using nine parameters. Three of these describe the translational shift of the module from the nominal position and three describe the rotational shift. The last three parameters describe deviations in the module geometry from a flat plane. As described in 2.4.1, within a given layer, each module overlaps slightly in r or z with its neighbours. This feature offers a further constraint on the fit of the alignment parameters, as the overlaps between the modules tightly constrain their relative position on the circumference of each barrel layer and each endcap ring. The length scale of the tracker is given by the position of the strips and pixel in a given silicon module, which is controlled to a precision of 10^{-4} by lithography.

The alignment is performed using cosmic ray data as well as data recorded during the LHC operations. The cosmic data are very useful since they impose vertical constraints, while the collision data constrain the center of the tracker. The tracker geometry is found to be very stable with time and the statistical accuracy of the alignment procedure is such that misalignment effects are small compared to the intrinsic hit resolution of the modules.

This calibration provides a correction for the major part of the alignment problem. However, a residual in the correction after the procedure is represented by so-called weak modes. These result from combinations of alignment parameters that do not change the track-hit residuals and therefore do not alter the total χ^2. A weak mode can result for instance from a twist where modules are moved coherently in ϕ by an amount proportional to the position along the z axis. Another example of weak mode is the alteration of the scale of the tracker due to deformations of the tracker shape from a perfect cylinder.

The weak modes have an impact on the momentum scale of tracks which is far below the nominal resolution. However, as extensively described in chapter 4, if the level precision needed for this observable is very high, it becomes important to correct for this effect.

2.5 Muons in CMS

Muons are very important tools in CMS. They are the only particles that can traverse the whole detectors and leave clear signatures in the muon chambers. Moreover, they can not be produced directly in soft QCD interactions that dominate the physics production at LHC and constitute a signature for more interesting events. For this reason, they are the main tool for triggering. On average, a muon in the

barrel loses 3 GeV of transverse momentum before it reaches the first muon station and another 3 GeV between the first and the last muon station. Muons with lower energy can therefore not be identified.

2.5.1 Muon system

The muon system is shown schematically in Figure 2.8. It is composed of different types of gas detectors organised in layers in the return yoke of the magnet. In the barrel, up to $|\eta| = 1.3$, the muon system consists of four layers occupied by drift tube chambers (DTs). These measure the position of the muon by converting their ionisation electrons drift time to the anode wire to a distance. In the endcaps, between $0.9 < |\eta| < 2.4$ where the flux of muons is higher, cathode strip chambers (CSCs) are used.

They are organised in four layers where closely spaced anode wires are stretched between two cathodes. The ionisation electrons drift towards the closest anode wire which provides the measurement point. The magnetic field is almost completely confined inside the steel return yoke and the trajectories are not bent within the layers of the muon system. Each layer measures the straight track and provides a vector in space called track segment. The segments are then extrapolated between the stations to reconstruct the full track.

In order to get a faster signal for triggering, resistive plate chambers (RPCs) are installed in most of the detector, up to $|\eta| = 2.1$. These are parallel plate gaseous detectors that combine an adequate position resolution with a very fast response time.

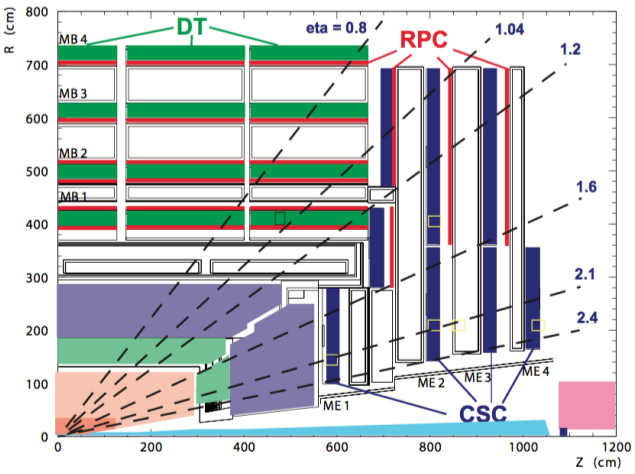


Figure 2.8: Layout of one quadrant of CMS. The four DT stations in the barrel are shown in green, the four CSC stations in the endcap in blue, and the RPC stations in red. Figure from [51].

2.5.2 Muon reconstruction

In the CMS procedure for the reconstruction of muons, tracks are first reconstructed independently in the tracker and in the muon system. The first are called tracker tracks and have the advantage of being

affected very little from the material outside the tracker itself despite the fact of being reconstructed with a low lever arm (4 Tm bending). On the other hand, the muons reconstructed exclusively in the muon chamber are called standalone muons. They are limited by multiple scattering up to very large momenta, but adding the constraint that the muons come from the beamspot it is possible to exploit the full lever arm of 12 Tm.

Afterwards, the tracks are combined following two possible approaches. The first one is to start from a standalone muon and search for a matching track, then refit into a global muon combining the hits. This gives major improvements in the measurement of the momentum for very high energy muons, above 200 GeV.

The second one is to extrapolate the tracker track to the muon system taking into account magnetic field, energy losses and multiple scattering, and search for at least a matching segment in the muon chambers. This allows the identification of low energy muons that stop inside the muon chambers. The majority of muons are reconstructed using both approaches.

Muon identification and isolation

The identification of muons is crucial in the discrimination of muons coming from decays of Z ,W and Higgs bosons from muons produced in QCD processes. In this context, a relevant variable to consider is the muon isolation. A muon is isolated if the energy flow in its vicinity is below a certain threshold. In most CMS analyses the energy flow is computed adding up linearly the transverse momenta or transverse energies measured in a cone in (η, ϕ) space, $\sqrt{\Delta\eta^2 + \Delta\phi^2} \leq R$, centered on the muon. The energy flow is corrected for the effect of pile-up. A particled is defined to be isolated if $\sqrt{\Delta\eta^2 + \Delta\phi^2} \leq 0.3$. Since the isolation depends on p_T, the relative isolation, defined as isolation divided by p_T, is commonly used in the analyses.

In addition to isolation, other criteria are used to improve the identification of muons. Three level of identification can be defined: loose, medium and tight identification. They are based on the definitions of tracker and global muon and on the quality on the track fit in order to maximise the identification efficiency. In the analysis described in this thesis, the medium identification has been used.

Muon efficiencies

Efficiencies for each step of the muon selection and reconstruction are precisely measured in CMS using the tag-and-probe method: dimuons from J/ψ and Z are selected using strict requirements on one lepton, called the tag lepton, and with relaxed selection on the second lepton, called the probe lepton. The efficiency of a set selection is defined as the fraction of probes passing the selection.

The muon efficiency in CMS is factorised in multiple independent components:

$$\varepsilon_\mu = \varepsilon_{\text{reco}} \cdot \varepsilon_{\text{track}} \cdot \varepsilon_{\text{ip+ID}} \cdot \varepsilon_{\text{iso}} \cdot \varepsilon_{\text{trig}}. \tag{2.5}$$

The tag-and-probe is performed independently for each component of ε_μ. The $\varepsilon_{\text{reco}}$ is the efficiency to reconstruct the muon in the muon system,$\varepsilon_{\text{track}}$ is the tracker track reconstruction efficiency. $\varepsilon_{\text{ip+ID}}$ is the muon ID selection efficiency combined with the impact parameter selection efficiency, ε_{iso} is

the muon isolation efficiency given a reconstructed and identified muon. $\varepsilon_{\text{trig}}$ is the trigger efficiency, given an isolated muon.

The efficiencies have been provided by the analysis working group in CMS and are taylored for the analysis that will be discussed in this thesis. They are computed in bins on muon p_T and η with a fine granularity. The ratio of each component of efficiencies between data and simulation can be used as a scale factor to weight the simulation to match the distributions in data. The scale factors are summarised in the plots in Figure 2.9. In addition to the efficiencies listed in Eq. 2.5, an additional scale factors for non-isolated muons is provided and shown in Figure 2.9f.

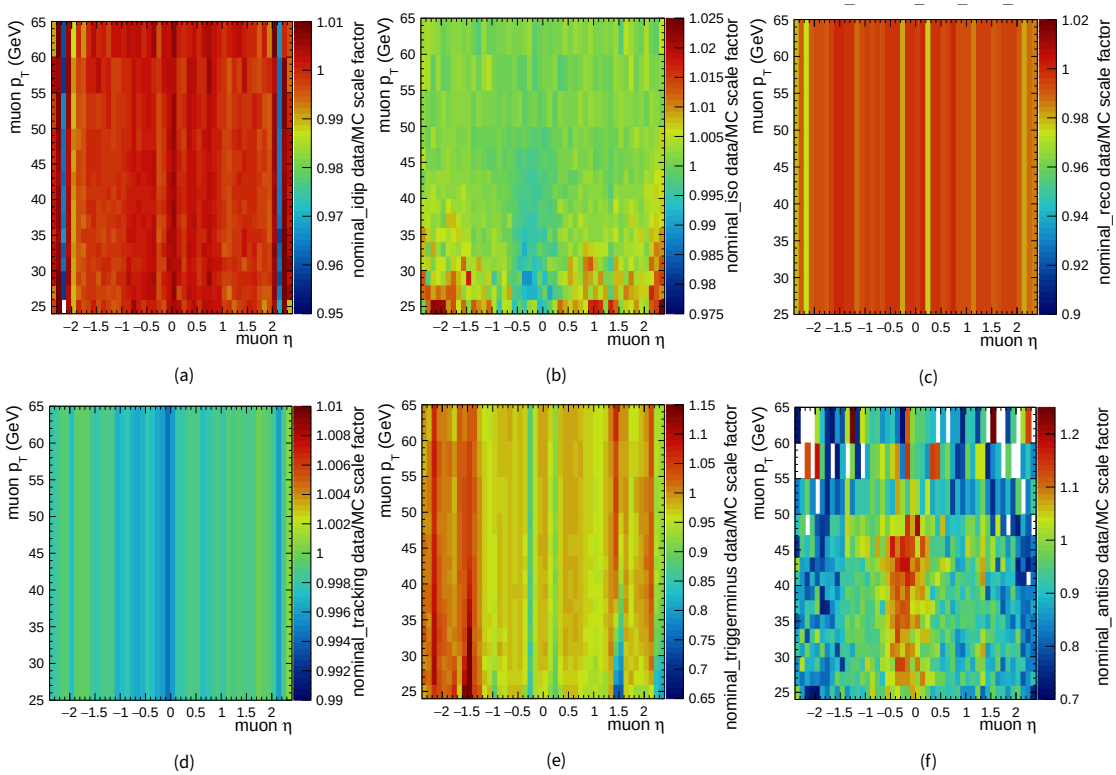


Figure 2.9: Scale factors for the efficiencies described in the text. Figure 2.9a shows $\varepsilon_{\text{ip+ID}}$, Figure 2.9b shows ε_{iso}, Figure 2.9c shows $\varepsilon_{\text{reco}}$, Figure 2.9d shows $\varepsilon_{\text{track}}$, Figure 2.9e shows $\varepsilon_{\text{trig}}$, and Figure 2.9f shows the efficiency for non-isolated muons.

Chapter 3

Tools for high performance computing for data analysis

In Chapter 1 we have discussed how the unprecedented luminosity delivered by LHC and collected by CMS has opened a new era in the sector of high precision measurements. However, the ambitious program of the measurement of the mass of W boson presented in Section 1.3 crucially relies on our ability to efficiently process a very large amount of data.

While the field of *big data* has spread out in all sectors of technology, the high energy physics community has sought a step forward with the introduction of new software that can profit from the underlying high-performance hardware while being usable by non-experts. Nowadays, multi-core hardware is commonplace and many-core architectures and computing clusters are increasingly available. The CMS collaboration has already adapted its core software to efficiently run on multiple cores. However, on the analysis side, common frameworks hardly improved with respect to the 10-year-ago state of the art [60]. The ROOT team [61] faced this challenge delivering RDataFrame, a package allowing transparent optimisations such as multi-thread data parallelism.

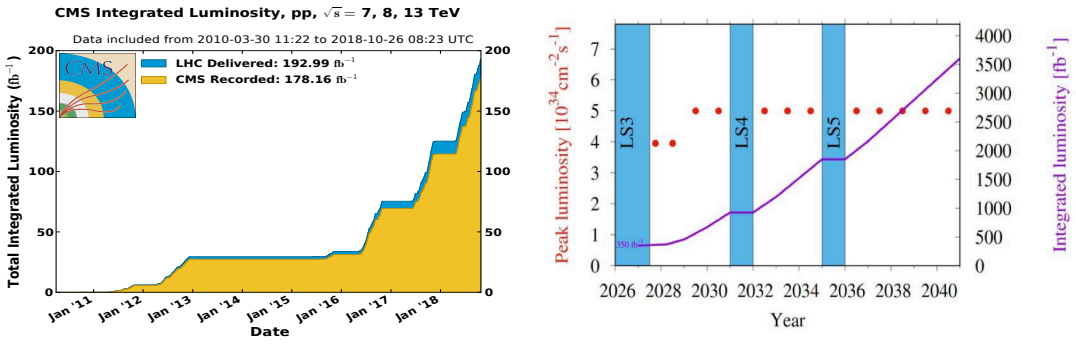
The RDataFrame's team and I worked together to adapt the existing software to the needs of large dataset precision analyses. Between the analysis conception in Section 1.3 and its practical realisation described in Chapter 5 and 6, a common effort on both sides has been necessary to develop efficient computing tools and in this sense this analysis has been taken as a 'guinea pig' for these purposes. This Chapter documents the innovative aspects that were introduced, the performance that was achieved, and the proof of concept applying it to the special cases of our analysis. It will be organised as follows: first of all the key features of RDataFrame will be briefly outlined. Then, an application to the analysis whose concept has been introduced in Section 1.3 will be presented.

The first milestone of this common effort has been presented by Enrico Guiraud and myself at one of the seminars of the Physics Department at CERN, *Using RDataFrame, ROOT's declarative analysis tool, in a CMS physics study* [62] in October 2019, this collaboration still eagerly carries on.

3.1 The need for parallelism in data analysis

The outstanding performance of LHC during Run1 and especially Run2 has allowed the experiments to collect an unprecedented amount of data which can be used to deliver high precision measurements. The plots shown in Figure 3.1 give an idea of the derivative with which the integrated luminosity in-

creased from the beginning of Run1 in 2011 to the end of Run2 in 2018 (Figure 3.1a) and how this is expected to evolve during HL-LHC from 2026 to 2040 (Figure 3.1b). Therefore, the beginning of Run2 of LHC has marked the point where a change of paradigm is needed in order to cope with the amount of data to analyse.



(a) (b)

Figure 3.1: Figure 3.1a: Delivered and recorded luminosity cumulative over all years during stable beams for pp collisions at nominal center-of-mass energy. Credits: CMS experiment. Figure 3.1b: Forecast for peak luminosity (red dots) and integrated luminosity (violet line) in the HL-LHC era with nominal HL-LHC parameters. From [63]

Data parallelism consists in parallelisation across multiple processors in parallel computing environments. The data can be distributed across independent nodes, which act on the data in parallel. This can be achieved with a number of hardware architectures, from laptops to computing clusters exploiting hundreds of cores. However, dealing with multi-threaded programming introduces a series of technical complexities that are normally beyond the standard knowledge of high energy physicists. Therefore, there is a high necessity of developing a software that allows to transparently benefit from multi-core hardware, without the need of knowing the details of its implementation.

The ROOT team has historically taken care of developing and mantaining the essential tools for high energy physics analyses. Data is distributed by the collaborations in the standardised `.root` format and ROOT provides the instruments for the input-output (I/O) operations. For this reasons, it is highly preferable to seek a transparent multi-threading solution for data analysis provided by ROOT. In fact, other solutions based on industrial tools (`numpy`, `pandas`,...) have the drawback of converting the `.root` format to a more convenient one before dealing with the actual analysis.

3.1.1 RDataFrame, a swiss-army knife for data manipulation

RDataFrame is ROOT's declarative analysis tool that has been developed with these ideas in mind. While striving for a simple programming model, it delivers modern, elegant interfaces that are easy to use correctly and hard to use incorrectly, while automatically parallelising data for the user. It has been devised to scale in servers from 1 to about 100 cores to computing clusters with hundreds and thousands of cores and it supports the most widely used programming languages in the High-Energy community: python and C++.

The design principles of RDataFrame have been to combine elements of *declarative* programming

with elements of *functional* programming. The first consists in providing high-level user interfaces that allow to easily express the action to be done but leave the decision of *how* to do it to the core software. In this way, without giving up on simplicity, it is possible to obtain a number of optimisations under the hood: transparent parallelisation, lazy evaluation and caching. *Functional* programming consists instead in allowing the users to code in terms of small reusable components. This increases thread-safety and code correctness.

From event loop to graph structure

In order to reach its goal, RDataFrame has introduced an important paradigm shift in the way a High Energy Physics analysis is conceived. Traditionally, the transformations are performed on the data on one event per time, therefore the whole data is accessed through a loop over all the events. RDataFrame has abandoned this concept in favour of a graph-style analysis: transformations are performed on *dataframe* objects through two types of actions: *Filter* applies a cut on the dataset while *Define* creates a new quantity (*Column*) starting from the existing ones. At this point, it is the user's responsibility to correctly connect the various nodes of the computation graph. Finally, it is possible to produce histograms with the desired quantities or other kinds of output. Figure 3.2 illustrates a schematic representation of RDataFrame's concept.

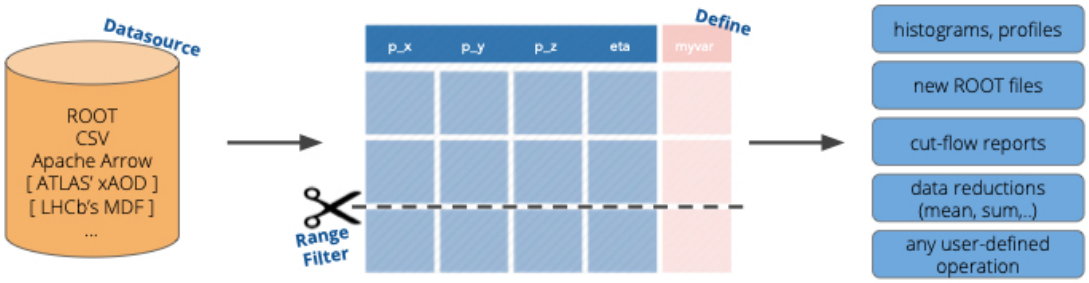


Figure 3.2: Schematic representation of RDataFrame's concept. Credits: ROOT team

3.2 Optimised high performance analyses in CMS

The introduction of RDataFrame has been a major step towards the simplification and speed up of the high energy physics analyses. However, some more work has been necessary in order to build a framework based on RDataFrame to perform a real analysis, with the required steps and complexities. The four experiments of LHC have developed independent schemes, therefore the first step is to adapt its own input dataset to the format conceived by ROOT. In this Section this step will be discussed in the context of the CMS experiment. Then, we will discuss a comprehensive analysis framework, designed to be fast, tidy, flexible, and reusable. As mentioned before, the graph structure introduced by RDataFrame is easily manageble in simple cases, but the bookkeeping of the graph nodes can become

increasingly complex in presence of various datasets and a huge number of operations on data. We will show how the designed framework solves this problem using modularity. Finally, a real analysis has to deal with the variations due to the estimation of systematic uncertainties on the observables, hence be ready to produce multiple copies of the same histogram with some inputs varied and the final output has to be organised and reducible to a format that can be fed to the subsequent steps of the analysis. This Section further describes a possible solution to this problem and shows its application to the muon momentum scale calibration (Section 4.4.2) and the measurement discussed in Section 1.3. Different ways to approach this problem are constantly developed and tested, reflecting the availability of resources in the different computing centers and the activity of the ROOT team to improve their software and standardize solutions for the average user.

3.2.1 CMS data formats

The data model employed by CMS has a tiered structure, where each subsequent format contains a more compact summary of the event data than its predecessor. The information read out from the detector is stored in the **RAW** data format which takes roughly 1 MB of disk space per event. Then, as the information is refined to build higher level objects, the size typically increases by a factor 2-3 (**RECO** tier). The subsequent tiers are created reducing the event content to facilitate the analyses and calibrations: **AOD** (Analysis Object Data) have been introduced during Run1 and reduce the **RECO** size by 85%, and **MiniAOD** [64] introduced after Run1 reduce the event size by a about a factor 10 compared to the **AOD**. The logic of reduction of the event content from **AOD** to **MiniAOD** is keeping information that is used by a large ($\simeq$ 80%) of physics analyses. At the end of Run2 CMS has decided to provide a new tier synthesising the information of **MiniAOD**: the **NanoAOD** [65], aiming to cover the needs of at least $\simeq 50 - 70\%$ of physics analyses. During Run1, the various analysis groups typically processed the **MiniAOD** to produce lighter input files for the analyses containing only the relevant pieces of information, in ROOT tree format. The advantage of introducing the **NanoAOD** format is centralising the recipes in order to save computing resources. In practice, this concept has later evolved and custom **NanoAOD** have been produced for calibrations and for analyses with special requirements, like the one we have introduced in Section 1.3 and will discuss throughly in the rest of this dissertation.

3.2.2 RDataFrame-based Analysis Framework

We present now the adopted solutions for the open issues left after the adoption of RDataFrame as backend of our analysis.

Modularity of the computing graph

In order to deal with a very complex analysis graph, we have implemented a modular analysis framework. Each module takes as input a dataframe object and returns it after a number of RDataFrame's operations. On one side, modules allow to gather multiple actions on the same logical node, and on the other the total volume of the code is substantially reduced. Modules can be implemented in both python and C++ and the presence of the mandatory methods is guaranteed through the usage of class

inheritance from a base module class. The linking of the various modules is done in python by a script that allows the user to build the desired logical graph.

The only allowed operation inside a module is the definitions of new columns (Defines). In order to ensure reusability of the same code, modules inputs are defined at the run time by the user while building the logical graph, provided that they maintain a compatible input type (boolean, Lorentz vectors,...). Filter operations are allowed only during the building of the graph in order to improve their traceability.

While building the analysis graph, the user can decide to produce a output, typically a histogram, a ROOT tree object or a cutflow at a given node. The collection of the objects produced at the various nodes is done internally in python and the structure of the graph guarantees that the code written in the modules will be transparently parallelised by RDataFrame and the data will be only looped once whatever the number of produced outputs, since RDataFrame's lazy evaluation is preserved. Figure 3.3 illustrates the concept of this analysis framework. Starting from the node 'input', the user navigates to the first node 'node 1' of the computing graph, where a series of modules containing the definition of new columns are executed. Then, the computing graph separates into two different branches, for example containing quantities relative to different systematic variations. Each branch develops into more nodes and between 'node 1' and 'node 2' some cuts are applied to the events. In correspondence of nodes 3 and 8 histograms of desired quantities are produced.

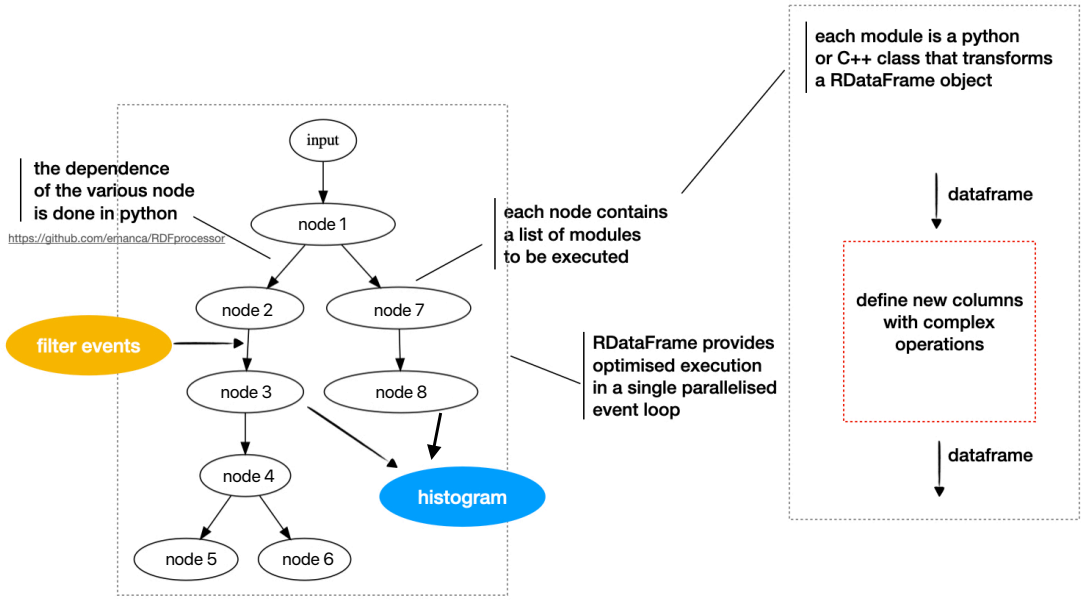


Figure 3.3: Illustrated concepts of the analysis framework discussed in the text.

Introducing multidimensionality

Both the analysis conceived in Section 1.3 and the muon momentum scale calibration in Section 4.4.2 exploit the correlations among various observables to extract variables of interest. Currently available histograms inside ROOT support up to 3 dimensions, while it would be desirable to extend this number to a generic *N*. In order to achieve this, RDataFrame has been interfaced with the recently developed Boost Histograms. Boost Histograms are built on top of the Boost library with modern C++17 and boast a number of advanced features, such as efficient use of memory, custom axis types (define how input values should map to indices) and user-defined accumulators in cells.

The interface of RDataFrame and Boost Histograms is done explicitly using RDataFrame's *Book* class, which allows to define an operation to be executed in parallel. There are two methods to fill an histogram using multi-threading. The first one is to create *n* copies of the same histogram, and let each of them to be filled with a selected chunk of data by a single thread. At the end of this process, the final histogram will be the sum of all the partial histograms. Unfortunately, when the number of bins is sufficiently large (about 1 billion), it is not affordable to keeping in virtual memory *n* copies of the same histogram and therefore it is necessary to instantiate one single histogram that can be filled by multiple threads simultaneously. This operation has to be done carefully in order to preserve thread-safety. In our implementation, this is realised keeping track of the content of each cell in the histogram using `atomic doubles`[1]. Finally, the multidimensional histograms can be produced and saved as `numpy` arrays [66].

Dealing with systematic variations

Systematic variations are an essential part of physics analyses: on a technical point of view they are taken into account as follows. The result of an analysis is extracted from one or more histograms of given observables. If some of these observables are affected by a systematic, a copy of the histograms has to be produced for each varied value of this systematic. This results in two complications: first of all producing multiple histograms at the same node has a cost in terms of the perfomance (the order of the variations for a precision analysis is about one thousand), and secondly the bookkeeping of the varied histograms becomes non trivial.

In order to overcome these limitations, we have defined a new accumulator for Boost Histograms, for both the cases of single and multiple copies in memory, which instead of returning two `doubles` (resp. `atomic doubles`), the number of events in each cell and its error, returns an `array` of `doubles` (resp. `array` of `atomic doubles`). In this way both issues are solved: only one histogram is produced at the same node and the multiple variations of the histogram will be accessed with an extra dimension of length equal to the number of systematics.

[1] An atomic operation is an operation that will always be executed without any other process being able to read or change state that is read or changed during the operation. In this context, atomic is referred to the implementation of the standard library in C++, where it represents a type that different threads can simultaneously operate on (their instances), without raising undefined behavior.

Scaling with the number of threads

To conclude this chapter, we present the performance of the current implementation. For the purpose of the test, we used a server located at INFN Pisa computing centre, bought for this analysis, equipped with a AMD EPYC 7742 processor, 128 physical cores with double socket, 2TB memory (DDR4, 3200 MHz) and a SSD-nvme disk of 54 TB. Data is stored and read from the SSD storage. Figure 3.4 shows a scaling plot for two different configurations, from real implementations in the analysis. It shows the rate in MHz as a function of the number of threads. The blue dots show the performance of filling a simple 2-dimensional histogram. In this case the level of MHz is reached already at about 16 cores and the scaling is linear up to about 64 cores. The orange dots represent instead a 7-dimensional histograms, while the green ones show the same histogram with a systematic variation. This adds an extra dimension of size 2 to the histogram. In both cases, the scaling is approximately linear for a fairly large number of cores, and the rate is well above the level of the MHz, reached already at about 32 thread. At around 48 threads, the rate saturates — as expected — as the resources used by the computer to handle the threads do not result in improvements on processing events. Eventhough the number of histogram saved in the run with the systematics (green dots) are three times the ones in the run without (orange dots), the rate of events is approximately depleted by only 20% thanks to the efficient implementation of the handling of the systematic variations. In the real analysis the number of systematic variations can be up to about 100, and this procedure guarantees that all the histograms are produced within one hour in world time.

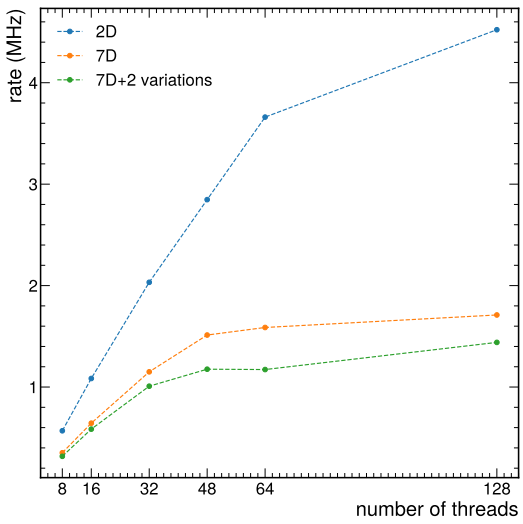


Figure 3.4: Scaling plot for two 7-dimensional histograms. The orange case contains one extra dimension to provide systematic variations.

3.2.3 Future prospects

A large number of developments have been incorporated in the RDataFrame tool and this is currently mantained by the ROOT team. After the successful demonstration of the improvement of performances and the adaptation to large scale of the NanoAOD inputs, RDataFrame is considered now a viable tool by many analyses and migrations are foreseen for the Run3 analyses.

Chapter 4

High precision calibration of the muon momentum scale

In Section 2.4 we have described the tracking system of CMS and the algorithms that are used to measure the tracks. Despite the outstanding quality of their implementations, when targeting a very high precision measurement it is necessary to carefully analyse all the detector effects and approximations that could bias the determination of the track parameters. Obtaining a calibration of the CMS tracker at the level of $2 \cdot 10^{-4}$ has not been a trivial task: an effort of years has been necessary to firstly understand the origin of the biasing effects in the determination of track parameters and then to devise the tools to remove them. This chapter will be devoted to describing the work done for this purpose in details.

Our starting point is a valid attempt documented in the final summary of a pilot measurement of the Z mass by CMS used to validate the experimental uncertainties of a future m_W measurement [67]. In this context, a procedure based on J/ψ and Υ resonances was devised employing a Kalman Filter. While the quality of the results presented in [67] was outstanding, at level of few 10^{-4}, any attempt to apply the same method to the data collected by CMS in 2016, augmented in statistics and extended up to $|\eta| < 2.4$ in acceptance, has not been satisfactory.

This has brought us to restart from scratch, almost from pen and paper, to try and figure out what was preventing us from successfully perform a calibration of the tracker of CMS. Checking all the details of the reconstruction software, we found that approximations in the various steps of track fitting could indeed lead to a bias in the measured track parameters. The calibration performed in [67] relied on a model that corrects the tracks based on effects linked to the physics of the detector: imperfections in the description of magnetic field, material budget and residual misalignment of the tracking modules. What emerged after our first studies was that the calibration was instead mixing up the physical effects with the biases induced by the approximations in the track fit. For instance, a notorious candidate in this context is the treatment of the material during the track fit. In the track model, material is placed on measurement layers, therefore a particle is only allowed to scatter at these finite locations. This is a limitation since a particle scatters in a continuous mean. Since the scattering angle has a dependence on the momentum of the track, this will introduce a bias on the scale of the reconstructed momentum, particularly in the dense regions of the tracker.

To try and single out the effect linked to the material in the track fit, in Section 4.1 we setup a simplified calculation of the track parameters with a χ^2 fit in presence of non-uniform flaws in the trajectory of whatever origin. Due to the dependence of the covariance matrix of the single measurements on

the momentum of the track, the bias on the the final track parameters will also show an explicit dependence on the momentum of the track, embedded in a non-trivial function. We discuss thoroughly the details of this model and finally we validate it using a collection of tracks reconstructed by a simulation of the CMS experiment, demonstrating our full comprehension of the phenomenon.

However, this simple exercise does not go beyond an assessment of our degree of understanding of the problem. To be able to tackle the complex matter of precising measuring the scale of the momentum of tracks, it has been necessary to devise new tools that allow to account for each local bias along the trajectory. In Section 4.2 we will document a new track fit method, the Continuous Variable Helix (CVH) Fit as an extension of the alignment procedure described in 2.4.2. The advantage of developing and using the CVH is double: on one hand, it allows to remove a series of imperfections and approximations that are present in the Kalman Filter track fit of CMS. Then, the introduction of the Geant4 propagator [68–70] allows a refined treatment of the material and multiple scattering, dumping the aforementioned undesired effects.

With this improved baseline, the physical model introduced in [67] holds again its validity. After describing briefly in Section 4.3 how the approximate description of magnetic field, material budgets and alignment can produce a bias in the track parameters, in Section 4.4 we present a procedure for extracting the correction parameters for muon momentum scale based on complex fits to the J/ψ invariant mass, as a refinement of the effort described in [67] which has been possible thanks to the recent introduction of more advanced computing resources. We conclude in section 4.7 showing the closure test and the final results.

4.1 Bias in the track parameters in a dense silicon tracking detector

We will now derive a model that allows us to parametrise the bias of the momentum scale of a track fitted with the standard CMS reconstruction software and validate it using a collection of tracks reconstructed by a simulation of the CMS tracker. This will help us understand the dependence on the material introduced by the multiple scattering in the covariance matrix of the measurements, which will naturally emerge from our calculations, but an intuitive approach to explain this effect is given in Appendix A. This simplified approach makes use of the generated particles, to which the reconstructed momentum is compared track by track and the bias singled out. When calibrating data, it is necessary to extract the corrections with a more complex procedure involving the use of the invariant mass of dimuon resonances and therefore the derived model has to be applicable to this case also.

At the end of this section we will show that although this model correctly describes the imperfections of the scale in reconstructed momentum with respect to the generated one in the simulation, it is too degenerate to be used in a simple fit of the corrections from the invariant mass of resonances in data.

Let us setup a simple calculation assuming that on the transverse plane a track can be parametrised as a parabola. Thus we have:

$$y = d_{xy} + \phi x + \frac{1}{2\rho} x^2 \tag{4.1}$$

where x and y are the coordinates along and perpendicular to the track and $\rho = p_T/0.3B$ is the radius of curvature of the track in presence of a uniform magnetic field B. The parameters d_{xy}, ϕ and ρ are found by minimising the relative χ^2. Following the procedure of the minimisation of the χ^2, the best fit track parameters $\vec{z} = \left(d'_{xy}, \phi', \rho'\right)$ and the covariance matrix of the fitted parameters V is:

$$\vec{z} = (P^T (V^{yy})^{-1} P)^{-1} P^T V^{yy\,-1} \vec{y} \tag{4.2}$$

$$V = (P^T (V^{yy})^{-1} P)^{-1} \tag{4.3}$$

where V^{yy} is the covariance matrix of the measurements y_i: $V_{ij}^{yy} =< y_i y_j >$. For this simplified case P is purely geometric and encodes the spacing of the layers:

$$P = \begin{pmatrix} 1 & x_0 & {x_0}^2 \\ 1 & x_1 & {x_1}^2 \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ 1 & x_N & {x_N}^2 \end{pmatrix} \tag{4.4}$$

We compute now explicitly the covariance matrix V^{yy} in a simple geometry. Let us assume we have N+1 layers of detectors, at positions x_i, $i = 0, .., N$ along the axis of the original direction of the particle, uniformly spaced and in a constant and uniform magnetic field B. The total length of the track in the transverse plane is L. Each layer has an intrinsic hit resolution σ and particles will undergo multiple scattering [1] when crossing them. We assume that each layer has identical radiation length $\Delta x/X_0$ .The variance of the multiple scattering angle on each layer is $\delta\theta_0^2$. The total amount of material is $(N + 1)\Delta x/X_0$ and the variance of the scattering angle of a track crossing the whole tracker is $\theta_0^2 = (N+1)\delta\theta_0^2$. The elements of the covariance matrix V^{yy} are:

$$V_{i,j}^{yy} = \sigma^2 \delta_{i,j} + \delta\theta_0^2 \left(\frac{L}{N}\right)^2 \left(\frac{i(i+1)(2i+1)}{6} + (j-i)\frac{i(i+1)}{2}\right) \tag{4.6}$$

where i runs from 0 to N and j runs from i to N.

We observe that since $\delta\theta_0 \propto 1/p$, the terms of V^{yy} depend explicitly from the momentum of the fitted track at the location where the scattering has occurred. This introduces a bias in the determination of the reconstructed momentum: smaller values of p are preferred in the fit since this maximises the uncertainty in the propagation and minimises the χ^2 value. This dependence will be relevant in

[1] The Coulomb scatterings with the nuclei of the material cause the muon trajectory in the tracker to be deflected by a small angle which is approximately Gaussian distributed with mean zero and r.m.s. given by:

$$\theta_0 \simeq \frac{13.6\ \text{MeV}}{\beta c p}\sqrt{\frac{x}{X_0}}, \tag{4.5}$$

where βc and p are respectively velocity and momentum and of the incident muon and x/X_0 is the thickness of the scattering medium in radiation lengths. For particles with η different from 0 this formula is still valid in the transverse plane, while x/X_0 still represents the total crossed material. An exhaustive description of these phenomena can be found in [71], where a complete derivation of the physical model is given.

the description of the biases in the scale of track parameters. In the following, we will make the p_T dependence in $\delta\theta_0$ explicit and use the notation $\delta = 13.6\ \text{MeV}\sqrt{\frac{\Delta x}{X_0}}$.

4.1.1 Correlation terms in track momentum resolution

Computing the matrix V explicitly for a configuration with $N+1 = 5, 7, 9, 11$ layers through Eq. 4.3 we can observe how off-diagonal terms due to the multiple scattering affect the track momentum resolution, which can be extracted from the element V_{33}:

$$\left(\frac{\sigma_{p_T}}{p_T}\right)^2 = a^2 + c^2 \cdot p_T^2 + \sum_{l=1}^{m} \frac{b_l^2}{1 + \frac{d_l^2}{p_T^2}}, \tag{4.7}$$

The sum runs up to a value $m = N/2 - 1$. Here the terms c^2 and a^2 reproduce the terms found in literature in the simplified case of very high or very low momenta. In the first case, the resolution is determined solely by the intrinsic spatial resolution [72]:

$$\left(\frac{\sigma_{p_T}}{p_T}\right)_{hit} = c \cdot p_T.$$

$$c = \frac{\sigma}{0.3B}\frac{1}{4L^2}\sqrt{\frac{720}{N+5}} \tag{4.8}$$

Conversely, in the limit of low momentum we can obtain the momentum resolution starting from the equation that links the transverse momentum with the sagitta s:

$$p_T\ [\text{GeV}] = \frac{0.3BL^2}{8s}\ [\text{T m}]. \tag{4.9}$$

The uncertainty on the sagitta is propagated to the momentum by:

$$\sigma_{p_T} = \frac{0.3BL^2}{8s^2}\sigma_s = p_T^2\frac{8}{0.3BL^2}\sigma_s. \tag{4.10}$$

For a uniformly distributed material the error on the sagitta due to multiple scattering is given by [73]:

$$\sigma_s^{ms} = \frac{L}{4\sqrt{3}}\frac{13.6\ \text{MeV}}{p_T}\sqrt{\frac{L}{X_0}}.$$

The related error on the momentum is given by eq. 4.10:

$$\sigma_{p_T}^{ms} = p_T^2\frac{8}{0.3BL^2}\frac{L}{4\sqrt{3}}\frac{13.6\ \text{MeV}}{p_T}\sqrt{\frac{L}{X_0}}, \tag{4.11}$$

thus the relative resolution is independent of momentum:

$$\left(\frac{\sigma_{p_T}}{p_T}\right)_{ms} = a,$$

where a is a constant defined by:

$$a = \frac{2 \cdot 13.6\ \text{MeV}}{\sqrt{3}\, 0.3BL}\sqrt{\frac{L}{X_0}}. \tag{4.12}$$

We see here that a^2 for uniform material is proportional to L/X_0, the material budget measured in number of radiation lengths. Moreover, the constant of proportionality contains the transverse track length L. On the other hand, additional terms b_l and d_l of Eq. 4.7 describe the correlation between the multiple scattering and the hit resolution. In the explicit case computed in Appendix B for 5 layers assuming a geometry resembling the one in CMS, the reference value of momentum for which the correlation terms become relevant is about 6 GeV, and in general this number is in a relevant momentum range for our analyses.

4.1.2 Correlation terms in track momentum scale

Let us introduce the unsigned curvature $k = 1/p_T$, which is the quantity directly measured in the track fit and thus distributed as a Gaussian. k may be biased and we use k^c to mean k after corrections.

Inserting a generic bias of the form $A' - \epsilon' k + qM'/k$ on each layer[2] and computing the best fit track parameters using Eq. 4.2 we find a complicated expression for the scale:

$$\frac{k^c}{k} = 1 + A - \epsilon k + qM/k + \sum_l^m \frac{A_l - \epsilon_l k + qM_l/k}{1 + d_l^2 k^2}, \tag{4.13}$$

where l again runs up to a value m that depends on the number of layers and in our calculations it is 1,2,3,4 respectively for 5, 7, 9, 11 layers. Non-trivial k dependencies originate from non-uniform biases along the trajectory combined with k-dependent weighting of the measurements due to the k-dependence in the covariance matrix V^{yy} from different relative contributions of the hit resolution and multiple scattering. This result is derived step by step for the simple case of a non-uniform multiplicative bias on the curvature k in Appendix B.

On a minimisation standpoint, extracting all the correlation terms from Eq. 4.13 is challenging, because there are degeneracies in the parameter space, and where the bias is small the d_l terms are poorly constrained.

4.1.3 Extracting correlation terms from track covariance matrix

Taking into account the full matrix V explicitly, besides V_{33} that we used to extract the momentum resolution, we observe that all its entries are affected by the presence of the same correlation terms.

[2] The functional form of this formula will find its justification in section 4.3

In the 3x3 covariance matrix of d_{xy}, ϕ and ρ, the elements are of the form:

$$C_{ij} = a_{ij}k^2 + c_{ij} + \sum_l \frac{g_{ijl}k^2}{1 + d_l^2 k^2}, \tag{4.14}$$

and we highlight that d_l are the same as Eq. 4.13. This is also shown in the simplified calculation reported in Appendix B. Since all the 6 elements of the covariance matrix have d_l in common, we can imagine to simultaneously fit them to extract d_l and in turn using this result to constrain the correlation terms in Eq. 4.13. For this test, we will use tracks produced with a CMS Monte Carlo simulation where the information on the generated and reconstructed momentum is available.

In practice, C_{ij} is not a constant for a given set of "identical" tracks: same momentum and same η, and has a narrow but non-trivial distribution. We can then extract the median of C_{ij} and its uncertainty in bins of η of the detector and k. Moreover, since the covariance matrix elements are all dominated by the k^2 term, extracting d_l gets more complicated. A solution to this is performing a first fit to determine the a_{ij} parameters in Eq. 4.14 and then diagonalise the matrix in the high k limit, where only the $a_{ij}k^2$ term is present. Then, a second run over the track collection is made, the diagonalisation is applied to the covariance matrix of each track and the medians recomputed. This produces a transformed matrix with k^2 removed from the off-diagonal terms. Figure 4.1 shows all the terms of the covariance matrix obtained in this second run in a bin of $2.2 < \eta < 2.3$, obtained with the above described procedure. They are plotted and fitted as a function of qk, so that the red points correspond to tracks with negative sign, and the green points to tracks with positive sign.

There are a number of features that we can observe in this plot. Since we have diagonalised in the high k limit, the diagonal terms show the eigenvalues of the a_{ij} terms. In the diagonal the eigenvalues are ordered by increasing magnitude from top-left to bottom-right, and for this reason the y axis scale of the top-left element is much smaller than the other two. As for the off-diagonal elements, with the k^2 dependence not present in the leading order all that is left is the rotated $c'_{ij} + \sum_l \frac{g'_{ijl}k^2}{1+d_l^2k^2}$ component. While, due to the diagonalisation procedure c' and g' are not guaranteed to have the same values as c and g, the denominator of the expression is left untouched. In absence of the d_l terms, this would reduce to a constant term with no dependence on k. What we observe is instead compatible with the presence of three d_l terms and the fitting function interpolates well the data points. This is shown with increased detail in Figure 4.2 where a zoomed version for one of the off-diagonal elements of the matrix is reported. The shape of the function containing the d_l terms changes regime when k^2 is much smaller or much bigger than $1/d_l^2$ and this can be appreciated in the plot as changes of slopes which are highlighted in Figure 4.2. The values found for d_l in this particular bin are $\simeq$ 5 GeV, $\simeq$ 15 GeV, $\simeq$ 30 GeV.

Another way to extract the d_l terms involves constructing a hybrid version of the analytic model described in section 4.1. When solving the equations for the track parameters and their covariance matrix, we had observed that the determinant of the inverse covariance matrix $P^T(V^{yy})^{-1}P$ contains the factors $1 + d_l^2k^2$ in the numerators. The roots of the determinants are $-1/d_l^2$, from which d_l can be extracted. We can construct the V^{yy} and P matrices for single tracks using the actual hit positions and their resolutions and multiple scattering contributions from the CMS reconstruction. Comparing the

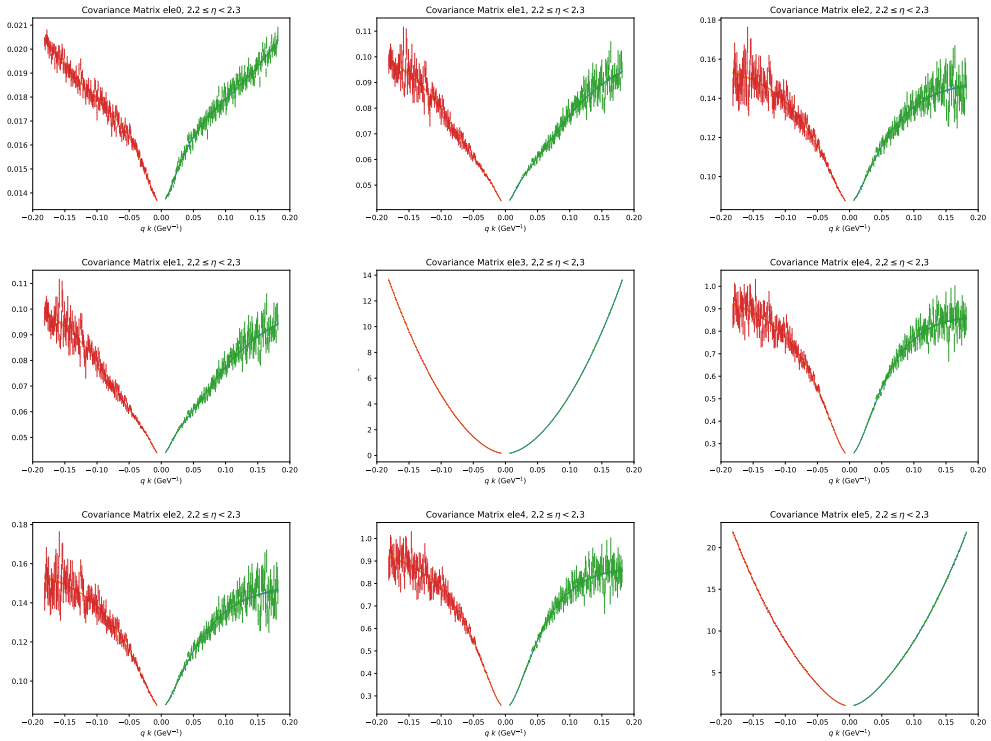


Figure 4.1: Elements of the covariance matrix of tracks reconstructed by the CMS detector, extracted with the procedure described in the text. They are plotted and fitted as a function of qk, so that the red points correspond to tracks with negative sign, and the green points to tracks with positive sign.

numbers derived with this test with the fits of the covariance matrix, we find a very good correspondence.

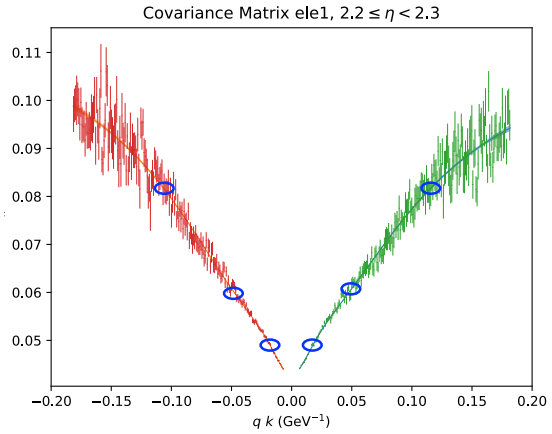


Figure 4.2: Zoomed version of one element of the covariance matrix from Figure 4.1.The red points correspond to tracks with negative sign, and the green points to tracks with positive sign. A line of the same color represents the fitting function. The points where the convexity of the fitting function changes due to d_l terms are highlighted with blue circles.

We can finally plug the extracted d_l values and fit the bias of the momentum scale. We use the first three values of d_l ordered by magnitude. For this purpose, we can compare tracks generated with known curvature and reconstructed in the CMS tracker and measure the bias in momentum scale as $\left(k_{reco} - k_{gen}\right) / k_{gen}$. In Figure 4.3 a fit to this quantity is shown in two bins of η as a function of qk and qp_T, for tracks of positive sign (green) and of negative sign (red). In all cases we observe a very good visual closure of the model with data points, even in presence of non-trivial trends, and this is a good proof of the validation of the model derived above. However, the fit parameters in this model are highly correlated and individual parameters are not well determined. Moreover, since the k_{gen} quantity is obviously not accessible in data, it is not feasible to extract such a number of terms from more complex quantities, i.e. the invariant mass of resonances. Therefore, the procedure discussed above has to be intended as a proof of principle and a validation of the model rather then a concrete method to derive the scale corrections for the momentum of tracks.

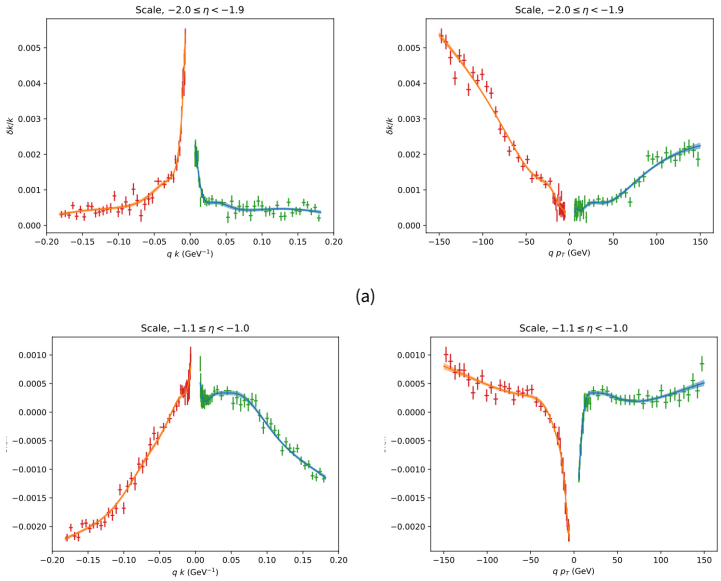


Figure 4.3: Fit to $\left(k_{reco} - k_{gen}\right) / k_{gen}$ is shown in two bins of η, for a track of positive sign (green) and one of negative sign (red), as a function of qk and qp_T.

4.2 Continuous Variable Helix fit for Generalized Global Corrections

We present now a new devised procedure that has been employed to extract the corrections to local biases due to magnetic field, energy loss and alignment for each layer. It is a generalisation of the global alignment procedure used in CMS [74] that includes layer by layer corrections to the material and magnetic field, plus a simplified residual alignment correction.

This extension of the Generalised Broken Line fit, described in Section 2.4.2, includes jacobians for the energy loss, energy loss derivative with respect to material parameter ξ, and transported curvilinear parameter gradient with respect to magnetic field magnitude. We start constructing a global χ^2 representation of the Kalman Filter based track fit, using a combination of the approach described in [75] and the CMS global alignment procedure [74], making use of the Geant4 propagator [68–70]. There is an important modification to be mentioned. With the usage of the Geant4 propagator the material is not artificially constrained into thin planes as in the nominal Global Broken Line fit, but sampled with continuity with a finer granularity then the whole distance between adjacent layers. Then, multiple scattering can cause a displacement between one measurement and another. As a consequence, the fitted parameters are 5 per layer rather than the 3 used in the nominal Global Broken Line fit, that we take to correspond to the 5 parameters in the curvilinear coordinate system. This new implementation has been called Continous Variable Helix (CVH) fit.

The advantage of using the CVH fit on top of the standard Kalman track fit of CMS is the refined treatment of material and the correction of a series of imperfections and approximations. We will show in Section 4.3 that the CVH fit allows to restore a clear baseline where the expected bias on the track momentum scale originates from the detector physics.

4.3 A model for the bias of momentum scale originating from detector physics

In this Section, we will introduce the model used in [67] that will be used to extract the corrections from the J/ψ dimuon mass in Section 4.4.

The transverse momentum p_T is measured from a fit to the trajectory in the magnetic field, as discussed in section 2.4.2. The coordinates of the trajectory are measured in the module reference frame and the track is reconstructed using information on the location of each module, on the magnetic field and on the material in the tracker. Therefore, any flaws in the knowledge of these inputs will result in a bias in the muon momentum scale that has to be parametrised and corrected. In this section we discuss the causes of the mismodelling of the magnetic field, of the alignment of the modules and of the energy loss in the tracker.

Magnetic field

The superconducting magnet has been described in 2.3.1. The map of the magnetic field in CMS had been precisely measured before the experiment was installed in the cavern. However, in order to save computing time, during the reconstruction of the tracks an approximate map is used instead of the real one. Figure 4.4 shows the ratio of the field integral, along straight lines from the origin to the last point in the tracker, computed with the real and approximate map as a function of η and ϕ. It shows variations of the order of 10^{-3} that propagate to the muon momentum scale since the transverse momentum is directly proportional to the magnetic field, as eq. 2.3 shows. Moreover, after the installation underground, the variation of the magnetic field has been measured using NMR probes. They show a

constant value since their installation in the cavern that was measured to be 0.9992 of the value measured on the surface and it is not clear how these effects propagate to the tracker volume, but it is plausible that the magnetic field is underestimated of the same amount. Therefore, the bias on the momentum scale would be of the order of $8 \cdot 10^{-4}$ which exceeds the desired precision. The map in Figure 4.4 can be applied to each track in data as a multiplicative correction to the track momentum.

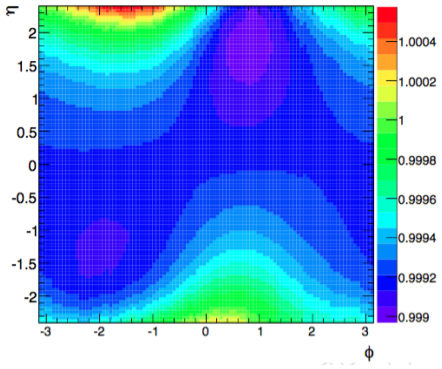


Figure 4.4: Ratio of the field integral along straight lines from the origin to the last point in the tracker, computed with the real and approximate map as a function of η and ϕ

Misalignment

The alignment procedure of the tracker is described in section 2.4.2. After this procedure the positions of the modules are known with a statistical accuracy much below the intrinsic measurement precision of the sensors. However, there is a residual bias introduced by the weak modes, extensively described in [74], that propagates to the muon momentum scale which is comparable to the level of precision we wish to achieve.

Energy loss

A muon traversing the tracker loses energy due to ionisation. This effect is described by the Bethe-Bloch formula, which models the energy loss of muons given the properties of the traversed material. The material in the tracker of CMS has a mean density of 0.2 g cm^{-3} (with the full distribution as a function of η shown in Figure 2.6) and the mean energy loss accounts for ~60 MeV for a muon of 40 GeV. As the material with which the tracker of CMS was built is far from being constant the energy loss is computed from the local information of the material in the tracker during the reconstruction of the tracks using the information from the Bethe-Bloch formula and a model of the material distribution. Therefore, an imperfect modelling of the local material budget would result in a bias in the energy loss and would in turn affect the momentum measurement at a level above the required precision.

We will now show how these three effects propagate to the momentum scale of the tracks. The unsigned curvature k is proportional to the magnetic field B through:

$$k = \frac{1}{0.3B\rho}, \tag{4.15}$$

from which it is clear that any flaws in the magnetic field map propagates to k with a multiplicative factor:

$$k^c = (1 + A) \cdot k, \tag{4.16}$$

where A is a small number. The misalignment of the modules in the tracker causes an additive bias in the muon momentum scale which is opposite in sign for positive and negative muons, as it can be understood from Figure 4.5. The correction of the curvature can thus be expressed as:

$$k^c = k + qM, \tag{4.17}$$

where q is the muon charge and M is the magnitude of the correction.

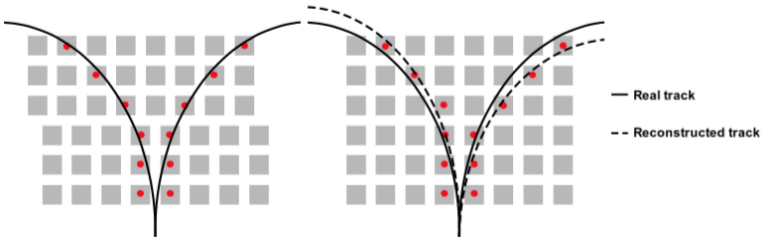


Figure 4.5: Misalignment of the tracker modules. The trajectories of the two muons in the tracker bend in opposite directions due to the magnetic field in the tracker. The actual trajectories in the misaligned tracker (left) will be reconstructed as if the detector was perfectly aligned (right). This leads to an additive term to the curvature, opposite in sign for the two muons.

Finally, any flaws in the modelling of the energy loss will result in a charge independent additive bias to the measured energy. Since in our samples the muons are ultra relativistic, the energy is equal to the module of the momentum (3D). Therefore:

$$E^c = E + \epsilon, \tag{4.18}$$

which propagates to the curvature as:

$$k = \frac{1}{E \sin\theta} \rightarrow k^c = \frac{1}{E \sin\theta + \epsilon \sin\theta} = \frac{k}{1 + k\epsilon \sin\theta}. \tag{4.19}$$

Accounting for all the effects, the corrected curvature can be expressed as:

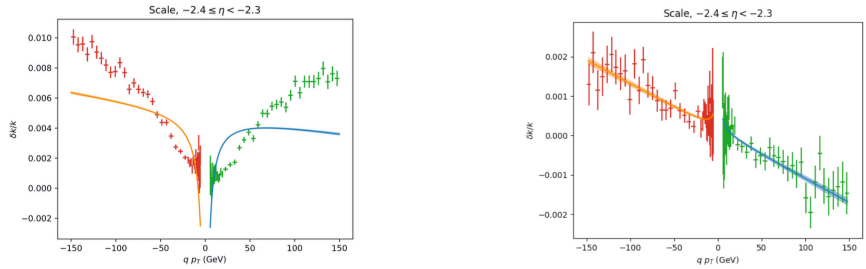
$$k^c = A \cdot k + \frac{k}{1 + k\epsilon \sin\theta} + qM, \tag{4.20}$$

and, expanding the energy loss term to first order and gathering $\sin\theta$ inside the ϵ term:

$$\frac{k^c}{k} = 1 + A - \epsilon k + qM/k, \tag{4.21}$$

where A, ϵ and M are the correction parameters for each η and ϕ bin of the detector. The values of these parameters are expected to be small according to the discussion of section 4.3. A should be less than 10^{-3}, M is expected to be less than 10^{-4} GeV and ϵ should be of the order of a few MeV. Those variations are to be compared to the typical momentum of muons from Z or W decays, around 40 GeV.

This simplified model arises from Eq. 4.13 when the biases are sufficiently uniform along the trajectory and the additional terms can be neglected. In this sense, it can be taken as the first order of a series expansion. We now show that the CVH fit restores a baseline where Eq. 4.21 well describes the biases in the momentum scale. Figure 4.6 shows how the modelling of the quantity $\left(k_{reco} - k_{gen}\right)/k_{gen}$ improves moving from the standard Kalman track fit of CMS (Figure 4.6a) to the CVH fit (Figure 4.6b). In both panels $\left(k_{reco} - k_{gen}\right)/k_{gen}$ is shown as a function of qp_T for a set of positive (green dots) and negative (red dots) tracks, obtained from a simulation of the CMS detector. The data points are overlaid with a fitting function following the model of Eq. 4.21. While in the case of Kalman fit there is a clear deviation from the simple model, when CVH fit is run the data points follow this model closely. Therefore, tracks fitted with the CVH fit on top of the standard Kalman are used for the momentum scale calibration.



(a) (b)

Figure 4.6: $\left(k_{reco} - k_{gen}\right)/k_{gen}$ is shown as a function of qp_T for a set of positive (green dots) and negative (red dots) tracks, obtained from a simulation of the CMS detector, for the standard Kalman track fit of CMS (Figure 4.6a) and for the CVH fit (Figure 4.6b). The data points are overlaid with a fitting function following the model of Eq. 4.21.

4.4 Extracting corrections from the invariant mass of dimuon resonances

Corrections to the track momentum scale and resolution are extracted from the invariant mass of J/ψ, while the Z events will be used to validate the procedure and to derive the corrections for the resolution at high p_T. The invariant dimuon mass can be expressed as a function of the curvature of the positive and negative muons as:

$$m^2 = \frac{1}{k_+}\frac{1}{k_-}\left[e^{\Delta\eta} + e^{-\Delta\eta} - 2\cos\Delta\phi\right], \tag{4.22}$$

where $\Delta\eta$ and $\Delta\phi$ are the difference between the pseudorapidity and the azimuthal angles of the two muons and k_+ and k_- their curvature. Here we assume that the bias on the dimuon mass only origi-

nates from the bias on the measured curvature, since the angles η and ϕ are measured with a much better precision. Then we can express the ratio between the reconstructed and true mass as:

$$\frac{m^2_{reco}}{m^2_{true}} = \frac{k^c_+ k^c_-}{k_+ k_-}. \tag{4.23}$$

The biasing effects are assumed to vary with the geometry of the tracker, therefore the corrections are extracted as a function of η of the detector. Since the two muons can be in a different η bin, the parameters A, M and ϵ can be different. We can then propagate the expression of the corrections on single muons Eq. 4.21 to the invariant mass as:

$$\frac{m^2_{reco}}{m^2_{true}} = \left(1 + A_+ - \epsilon_+ k + \frac{M_+}{k_+}\right)\left(1 + A_- + \epsilon_- k + \frac{M_-}{k_-}\right) \tag{4.24}$$

Instead, for the resolution we can use Eq. 4.7 and extend to the resolution on the mass:

$$\begin{aligned}\left(\frac{\sigma_m}{m}\right)^2 =& \left(\frac{\sigma_k}{k}\right)^2_+ + \left(\frac{\sigma_k}{k}\right)^2_- \\ =& a^2_+ + \frac{c^2_+}{k^2} + \frac{b^2_+}{1 + d^2_+ k^2} + a^2_- + \frac{c^2_-}{k^2} + \frac{b^2_-}{1 + d^2_- k^2},\end{aligned} \tag{4.25}$$

where we have kept only the leading order term in the correlation terms since we have shown from previous studies that this is sufficient to measure the momentum resolution in CMS to the desired level [76].

The quantity m_{true} in Eq. 4.24 is the target mass of the calibration, which has to be known at a much better level than the desired precision of the procedure. The mass of the J/ψ towards which we calibrate is known at the level of 10^{-6} [77].

4.4.1 Measured and simulated event samples

The calibration is performed on a sample of data collected by CMS in 2016 at $\sqrt{s} = 13$ TeV and luminosity 17 fb^{-1} using dimuon resonances: J/ψ and Z . Figure 4.7 shows the dimuon mass distribution in CMS collected with various dimuon triggers at 13 TeV in 2016. For the J/ψ , events are required to have passed the OR of various trigger paths, while for Z a single or double muon trigger path is employed.

For what concerns J/ψ , at analysis level we only accept events with reconstructed muon $p_T > 1.5$ GeV and $|\eta| < 2.4$ and invariant mass in the range $2.9 < m_{\mu\mu} < 3.3$ GeV, and a loose requirement on the muon identification, for about 85 million events effectively in acceptance. A sample of J/ψ simulated events has been generated with a J/ψ gun interfaced with PYTHIA8[79] and PHOTOS[80], with no pile up. These events are required to pass the same set of analysis level acceptance cuts as data, for about 32 million events effectively accepted. Figure 4.8 shows the distributions of p_T and η of the positive muon in data and simulation, normalised to their integral. We notice how the two distributions differ substantially in shape, due to the different underlying spectra of the J/ψ . This is suboptimal for this analysis and will introduce complications in the derivation of the muon momentum scale parameters, as will be discussed in Section 4.5 and in the final discussion in 4.7.

As for the Z , events are required to have a reconstructed muon $p_T > 20$ GeV and $|\eta| < 2.4$ and invariant mass in the range $75 < m_{\mu\mu} < 115$ GeV, a loose requirement on the muon identification, for about 5 million events effectively in acceptance. A sample of Z simulated events has been generated with POWHEG[81] with MiNNLO$_{PS}$[82, 83] interfaced with PYTHIA8[79] and PHOTOS[80]. These events are required to pass the same latest set of analysis level acceptance cuts as data. Figure 4.9 shows the distributions of p_T and η of the positive muon in data and simulation, normalised to their integral. These distributions have a very good agreement in shape.

4.4.2 Fit to the invariant mass of J/ψ and Z

The first step is extracting the value of $scale = m_{reco}/m_{true}$ and resolution σ from J/ψ and Z events. The invariant mass is computed in a grid $(\eta^+, \eta^-, p_T^+, p_T^-)$. For the J/$\psi$ the adopted binning in η is 48 bins in the range $|\eta| < 2.4$, while p_T^+ and p_T^- are binned with 8 bins whose edges correspond to $(2.6, 3.4, 4.4, 5.7, 7.4, 10.2, 13., 18.)$ GeV. This results in about 150k kinematic bins. As for the Z , the same binning in η is used while p_T^+ and p_T^- are binned with 4 bins corresponding to $(25, 38, 44, 48.7, 100)$ GeV, for a total of about 40k kinematic bins. The J/ψ dimuon mass is binned with 100 bins in the range $(2.9, 3.3)$ GeV, while the Z dimuon mass is binned with 100 bins in the range $(75, 115)$ GeV. Therefore, the total number of bins is about 14 million for the J/ψ and 4 million for the Z .

Computing a grid of this size is not trivial and it is performed using the high performance computing procedure described in Chapter 3. Thanks to this technology, despite the remarkable size of such a 5-dimensional histogram (4-dimensional grid plus the invariant mass dimension) the run time is reduced 5 minutes using 24 cores.

After dumping the grids, a further selection is made to remove the underpopulated regions in the phase space. Only the kinematic bins in which at least 500 events are present will be used for the analysis. This operation reduces the number of kinematic bins to about 12k for the J/ψ and 3k for the Z .

A kernel model for resonance lineshape

The invariant mass distribution in each bin of the grid is then fitted to extract $scale$ and σ. Usually, the lineshape of a dimuon resonance is modelled by a Crystalball function which describes a Gaussian core and a tail parameter to encode the energy loss due to final state radiation. However, the performance of such a model is not adequate for the requirements of this analysis and an alternative model is proposed based on the generator level information. Dimuon simulated events are selected in each kinematic bin using the reconstructed momenta. Their generated invariant masses are computed using the generator momenta after final state radiation, stored and used to build a probability distribution function of the generated mass $P(m_{gen})$. In the case of the J/ψ the distribution of m_{gen} is essentially a delta-function with a small radiative tail. In the case of the Z this is a more complex distribution including the finite width and γ^* effects. The value of the probability density function for a given mass m_{reco} and parameters $scale$ and σ is given by:

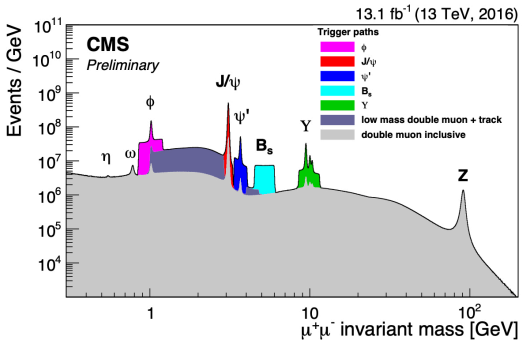
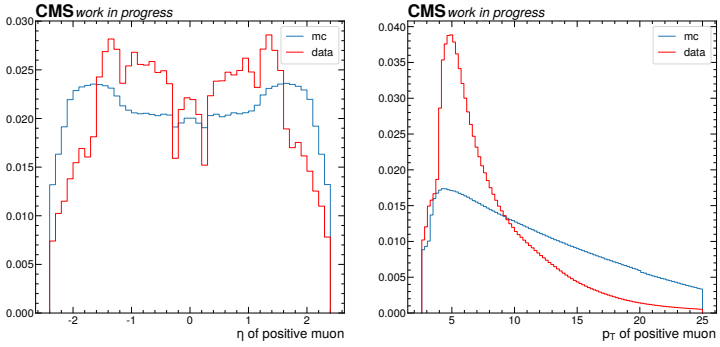
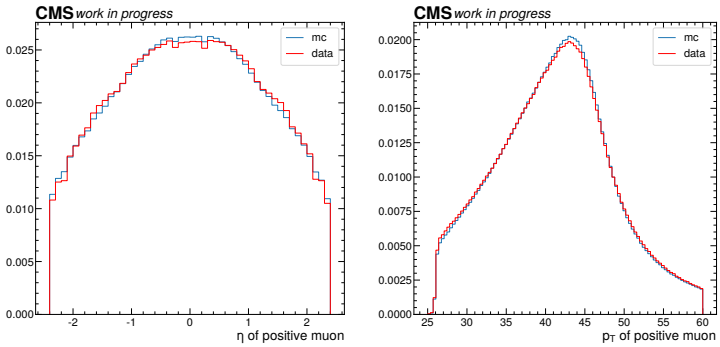


Figure 4.7: Dimuon mass distribution collected with various dimuon triggers at 13 TeV in 2016 with 13.1 fb^{-1}. The coloured paths correspond to dedicated dimuon triggers with low p_T thresholds, in specific mass windows, while the light gray continuous distribution represents events collected with a dimuon trigger with high p_T thresholds. From [78].



(a) (b)

Figure 4.8: Measured (red line) and simulated (blue line) single-muon kinematic distributions, for the positive muon of J/ψ decays: η in 4.8a and p_T in 4.8b.



(a) (b)

Figure 4.9: Measured (red line) and simulated (blue line) single-muon kinematic distributions, for the positive muon of Z decays: η in 4.9a and p_T in 4.9b.

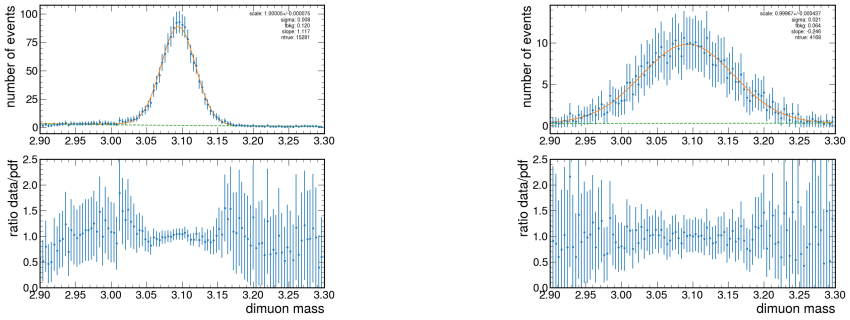
$$P_{sig}(m_{reco}) = \int_{m_{gen}} P(m_{gen}) \frac{1}{\sigma \cdot m_{gen}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{m_{reco}-scale\cdot m_{gen}}{\sigma \cdot m_{gen}}\right)^2} dm_{gen} \tag{4.26}$$

Conceptually, this model is constructed in the following way. In a given kinematic bin, the value of the content of each bin of m_{reco} is taken to be the sum of n Gaussians of unit area with mean m_{gen} and sigma σ, where n is the number of generated events which contribute to that kinematic bin. The main approximation in this model is that the resolution σ is the same for all the events with the same m_{gen}. This is one of the reasons why we have to adopt a grid with sufficiently small granularity. Moreover, this model factorises out the effects due to final state radiation and, in case of the Z , the natural width and γ^*. When fitting J/ψ data, an exponential probability density function P_{bkg} is added to take account of the background. In this case, the total likelihood is given by:

$$\mathcal{L} = -\log \sum \left(f_{bkg} P_{bkg}(m_{reco}) + \left(1 - f_{bkg}\right) P_{sig}(m_{reco})\right) \tag{4.27}$$

with the sum running on number of bins of m_{reco}. The fraction of background events in data varies from about 10% to a few percent depending on the kinematic bin.

Fits to the grid are performed simultaneously with a very fast and robust implementation based on *jax* [84] and advanced minimisation algorithms. It can also run on GPU. Figure 4.10 shows the result of this fit in J/ψ data for two bins on the grid: Figure 4.10a in the barrel and Figure 4.10b in the endcap.



(a) (b)

Figure 4.10: Fits to the invariant mass in of J/ψ with the kernel model. Blue dots represent the data point, the orange line is the sum of the kernel model and the exponential background and the dashed green line in the sole contribution from the exponential background. Figure 4.10a shows a bin in the barrel and Figure 4.10b shows a bin in the endcap.

Extraction of the parameters

After obtaining the parameters $scale$ and σ and their covariance matrix for each bin on the grid, let us indicate it with the subscript 'binned', we can minimise the difference with respect to their values computed with our model, Eq. 4.24 for the scale and 4.25 for σ and extract the relative parameters for

the corrections:

$$\chi^2 = \sum_{ibin} v^T C^{-1} v \tag{4.28}$$

where

$$v = \begin{pmatrix} \left(1 + A_+ - \epsilon_+ k + \frac{M_+}{k_+}\right)\left(1 + A_- + \epsilon_- k + \frac{M_-}{k_-}\right) \\ a_+^2 + \frac{c_+^2}{k^2} + \frac{b_+^2}{1+d_+^2 k^2} + a_-^2 + \frac{c_-^2}{k^2} + \frac{b_-^2}{1+d_-^2 k^2} \end{pmatrix}^{ibin} - \left(scale^2, \sigma^2\right)_{binned}^{ibin} \tag{4.29}$$

and C is the covariance matrix of $scale$ and σ. The sum runs over the bins in the grid. We actually minimise $scale^2$ and σ^2 because it is more convenient from a computational point of view.

The expression for the model for $scale$ and σ depends explicitly on the curvature k and the dependence on η is encoded inside the parameters. While fitting, the dependence on k is removed by explicitly computing on each bin of the grid $(\eta^+, \eta^-, p_T^+, p_T^-)$ the mean value of each term containing a function of k.

The parameter values in a given bin of η^+ can be constrained by multiple bins on the grid since in principle η^- can assume arbitrary values. In practice, the angle of separation of the two muons from a decay of a J/ψ is very narrow, so that η^- is typically contained in the range $(\eta^+ \pm \Delta\eta)$, where $\Delta\eta$ has the size of 0.2, corresponding to about 2 bins in the grid. This is not the case of the Z which is generally produced with a small momentum, thus the separation between the muons is wider and spanned across the grid.

In Figure 4.11 a summary plot for the fitted $scale$ in J/ψ simulation (Figure 4.11a) and data (Figure 4.11b) is shown. The corresponding plot of σ in J/ψ is shown for simulation (Figure 4.11c) and data (Figure 4.11d). They show a flattened version of the 4-dimensional grid, unrolled in such a way that η^+ is the external dimension, while p_T^- the internal one. Each blue dot represents the value of $scale$ or σ in a bin in this unrolled scheme. The envelope of this plot is then effectively a profile in η^+. The blue dots are overlaid by a thin red line, following closely the trend of the data points. It represents the fitted model as per Eq. 4.29. We observe that the scale in Figure 4.11a and Figure 4.11b is different: the simulation is essentially flat and needs very small corrections, while data is off up to some 10^{-3} in the endcaps. In all cases the model is able to synthesise the overflowing information contained in the grid.

Single muon fits on the simulation

In parallel to the above described procedure, a fit to parameters of the scale and resolution based on the comparison of the generated and reconstructed curvature is set up. This is useful to check the validity of the model and to measure the d parameter in the resolution, which depends only on the details of the software of the track reconstruction as throughly discussed in Section 4.1.

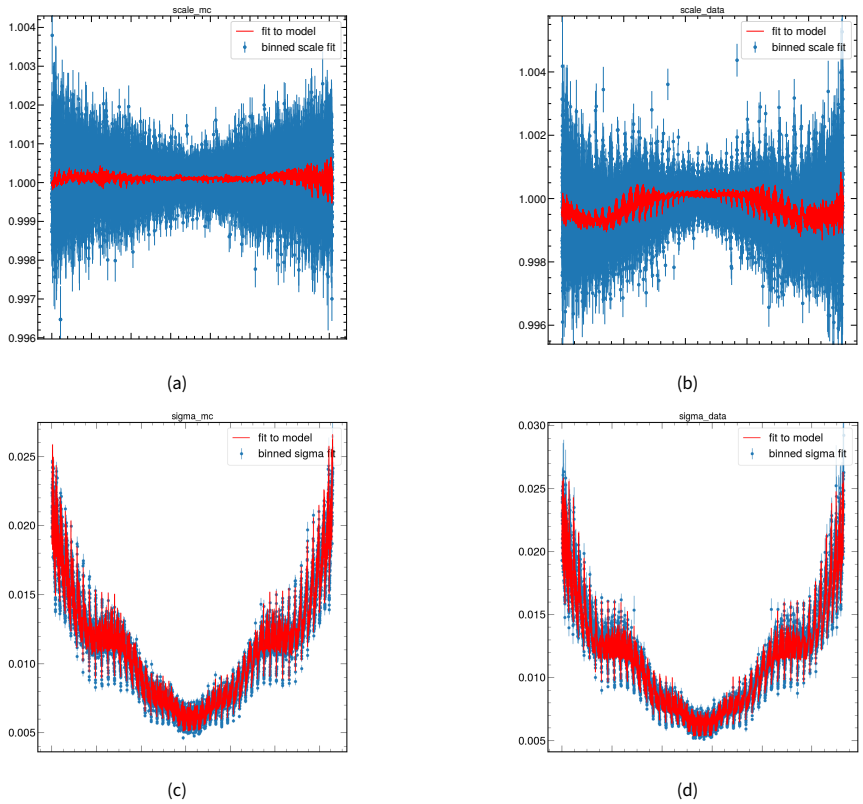
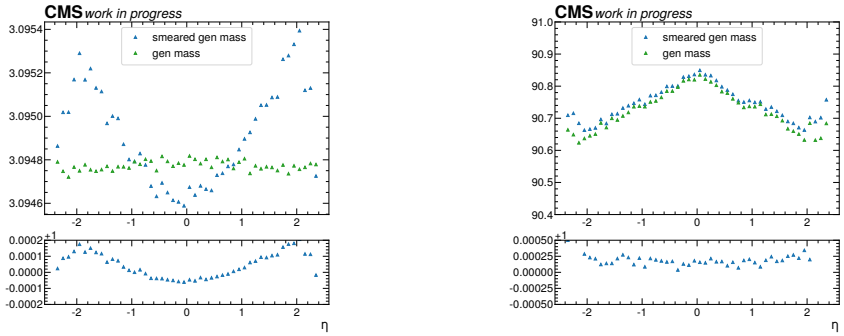


Figure 4.11: Summary plot of the values fitted using J/ψ events. Fitted $scale$ in simulation (Figure 4.11a) and data (Figure 4.11b), and fitted σ in simulation (Figure 4.11c) and data (Figure 4.11d). The blue dots represent the value of $scale$ or σ in a bin in the unrolled 4-dimensional grid. The red line is the model from Eq. 4.29 interpolating the points.

4.5 Effects due to muon momentum resolution in the calibration

A precise calibration of the muon momentum resolution is crucial when calibrating the scale at the target level of 10^{-4}. The resolution affects the average generated invariant mass when computed in the bins of the grid. Figure 4.12a shows the average value of m_{gen} (plotted in green) as a function of η of the positive muon compared with the invariant mass computed with the generated momenta smeared by the expected resolution of the reconstructed simulation (plotted in blue). We observe that the generated invariant mass, being very narrow, is flat within a relative range of 10^{-5}, but the smearing induces a variation which can be as big as about $2 \cdot 10^{-4}$, which is beyond the desired precision.

Conversely, the generated invariant mass of the Z , shown in Figure 4.12b varies with a much bigger scale than the J/ψ , having a large natural width. The reconstruction resolution adds a variation of about $3 \cdot 10^{-4}$ on top of this.



(a) (b)

Figure 4.12: Comparison of the invariant mass of generated events (green) with itself after smearing of the muon momenta with the reconstruction resolution of the simulation (blue). Figure 4.12a shows the J/ψ and 4.12b shows the Z . The bottom panels show the ratio of the two curves in the main panels.

In the J/ψ , the outlined effect will bias the derivation of the scale parameters. This is mitigated in the calibration in the following way. First of all, the resolution of the data and simulation is extracted from J/ψ events, as described in Section 4.5.1. Then, the invariant mass of generated events is smeared as in Figure 4.12a and it is used to build the kernel model of Eq. 4.26 in place of the target mass m_{gen}. In this way the effects due to the resolution in each cell of the grid are reduced.

On the other hand, a mismatch of the resolution between data and simulation can produce a non-closure that is of the same order of the desired precision. This happens for the following reason. When events are selected in a bin of a measured quantity p, the average true value of the selected event is not the center of the bin and this produces a small bias ϵ_p:

$$\epsilon_p = \sigma_p^2 \frac{1}{f(p)} \frac{\partial f}{\partial p}. \tag{4.30}$$

Here σ_p^2 is the square of the resolution of p, $f(p)$ its distribution and $\partial f / \partial p$ is the derivative of the spectrum. Because of this effect when we plot a variable as a function of p we introduce a bias that

depends on the resolution and that is more important when the underlying spectrum is steep, as in the case of the muon momentum distribution. This affects the closure of the calibration on both J/ψ and Z and a precise matching of the resolution in data and simulation together with a careful equalisation of the underlying spectra is required in order to mitigate it.

4.5.1 Measurement of the muon momentum resolution

The parameters of the muon momentum resolution are extracted in data with a simultaneous fit of the events from J/ψ and Z . In the simulation, such a simultaneous fit can not be performed since the samples for J/ψ and Z do not have the same resolution, because the first one has been produced without pile-up. Therefore, the resolution parameters in the simulation are derived from a fit of J/ψ events only. In the definition of the resolution model, Eq. 4.3, the dependence on the length of the track from the parameters has been removed to better appreciate their physical meaning. While a^2, b^2 and c^2 are left free in the minimisation, d is fixed to the value obtained in the single muon fit in the simulation which is found to be fairly constant to about 10 GeV in η and grows to about 15 GeV at $|\eta| = 1.5$ where the amount of material in the tracker is high. This is in the same ballpark as the study reported in 4.1. In Figure 4.13 the fitted values of a^2, b^2, c^2 and d are shown.

Figure 4.13a is a measurement of the material budget of the tracker, which is extracted from low p_T J/ψ . The fitted value in data is slightly higher than the one fitted from simulation. As for Figure 4.13b and 4.13c, the disagreement in data and simulation is due to the fact that the J/ψ alone in simulation does not have enough lever arm in p_T to correctly fit these terms and they are respectivaly underestimated and overestimated.

The fitted parameters in simulation are then used to smear the invariant mass of the simulated events to be used as target mass for the extraction of the correction parameters for the muon momentum scale. Since the absolute value of the resolution matters in the effect shown in Figure 4.12a, in principle the simulation should be smeared to match the resolution of data. In practice, this effect is small given the narrow width of the J/ψ , and moreover this is not feasible in the context of this analysis since the spectra of data and simulation of muon η and p_T shown in Figure 4.8 are very different. This implies that a smearing of the simulation using the extracted parameters does not provide an equalisation of data and simulation.

Not having a consistent resolution for J/ψ and Z in the simulated samples also implies that it is not possible to extract the correct parameters b^2 and c^2 to smear the Z simulation to match data. However, as seen from Figure 4.14, the Z resolution in data and simulation agrees below 5% and the residual biasing effect on the scale is expected to be small.

4.6 Calibration of muon momentum scale

The second step in the muon momentum calibration is extracting the parameters for the correction of the scale from J/ψ events. Figure 4.15 shows the results of the fit performed in the simulation (blue dots, *mc*) and data (red dots, *data*). While the correction of the simulation is rather flat, in data, the A parameter at $|\eta| \simeq 1.3$ drops with a very high derivative to about $-0.8 \cdot 10^{-4}$ in the endcaps. In this

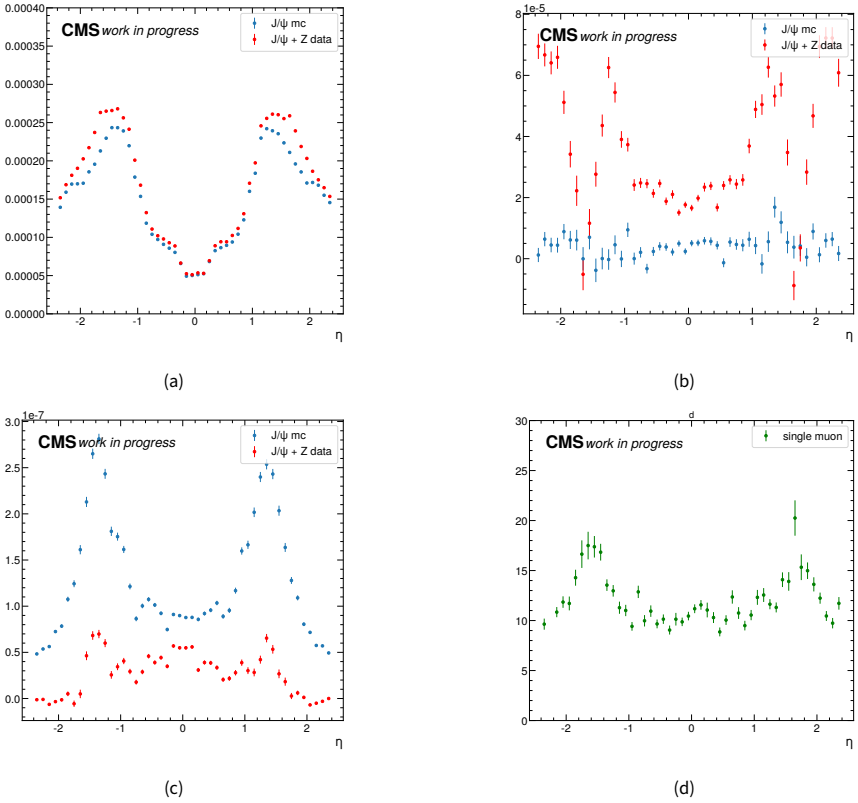


Figure 4.13: Results of the fit to J/ψ events of the muon momentum resolution. It is performed on the simulation (blue dots, *mc*) and data (red dots, *data*). 4.13a shows the a^2 parameter, 4.13b the b^2 parameter, 4.13c the c^2 parameter, and 4.13d the d parameter.

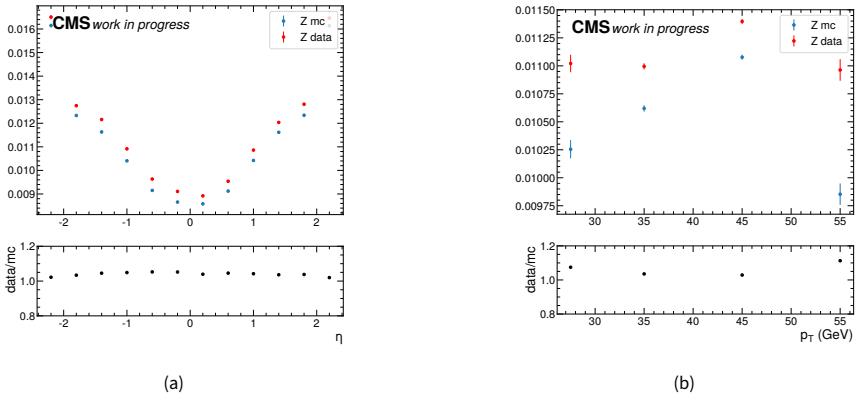
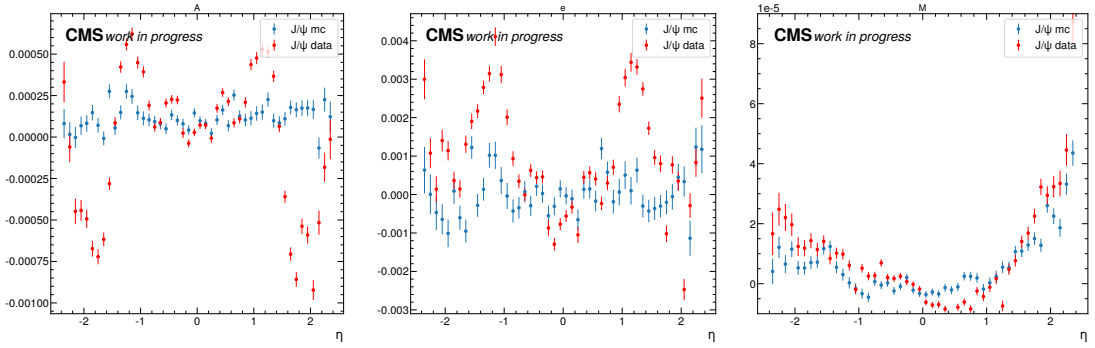


Figure 4.14: Comparison between the resolution in data (red dots) and simulation (blue dots) for Z events, a function of η (Figure 4.14a) and p_T (Figure 4.14b) of the positive muon.

transition region the magnetic field acquires a B_r component besides the B_z one as the magnetic flux has to close. Moreover, from this region moving towards the endcaps the magnetic field is affected by

the iron of the muon chambers, and a mismodelling of its quantity and shape can induce a variation with respect to the expected value of the magnetic field map.

The material correction ϵ in data peaks at about 5 MeV in correspondence of the region of the tracker where more material is present due to the cables and services. The M term encoding residual weak modes in the alignment looks very consistent to what is modelled in the simulation.



(a) (b) (c)

Figure 4.15: Results of the fit to the corrections, performed in the simulation (blue dots, *mc*) and data (red dots, *data*). 4.15a shows the A parameter, 4.15b the ϵ parameter and 4.15c the M parameter.

4.7 Closure of the momentum scale calibration

After extracting the values for the corrections in data and simulations from J/ψ , they can be applied event by event to J/ψ and Z . Then, the entire procedure of measurement of $scale$ is repeated and the agreement between data and simulation is verified before and after the corrections. This is shown in Figure 4.16. In the left panels, the difference of fitted scales in data and simulation are shown before the corrections. Conversely, in the right panels the agreement is shown after the corrections have been applied.

Overall, the agreement between data and simulation is reasonable even before applying the corrections, since the CVH fit and the application of the magnetic field map in Figure 4.4 to data events provide a good baseline. The corrections highly improve the agreement as a function of η of the positive muon in the endcaps, comparing Figure 4.16a and 4.16b, while the closure plot as a function of p_T in Figure 4.16d shows an improvement on the lowest pt bins in J/ψ and Z with respect to Figure 4.16d and equally good overall agreement elsewhere.

Since the scale corrections have been derived using the J/ψ only, the closure plots of the J/ψ are to be interpreted as a validation of the procedure. The p_T dependence of the the J/ψ after calibration shows a residual effect of the mismatch of the spectra of data and simulation, which results in a 10^{-4} non closure. This effect can be removed with a dedicated simulation.

On the other hand, the closure on the Z represents a test of the validity of the model that has been used to derive the corrections, Eq. 4.21. Being parametric in p_T, it can be precisely computed using

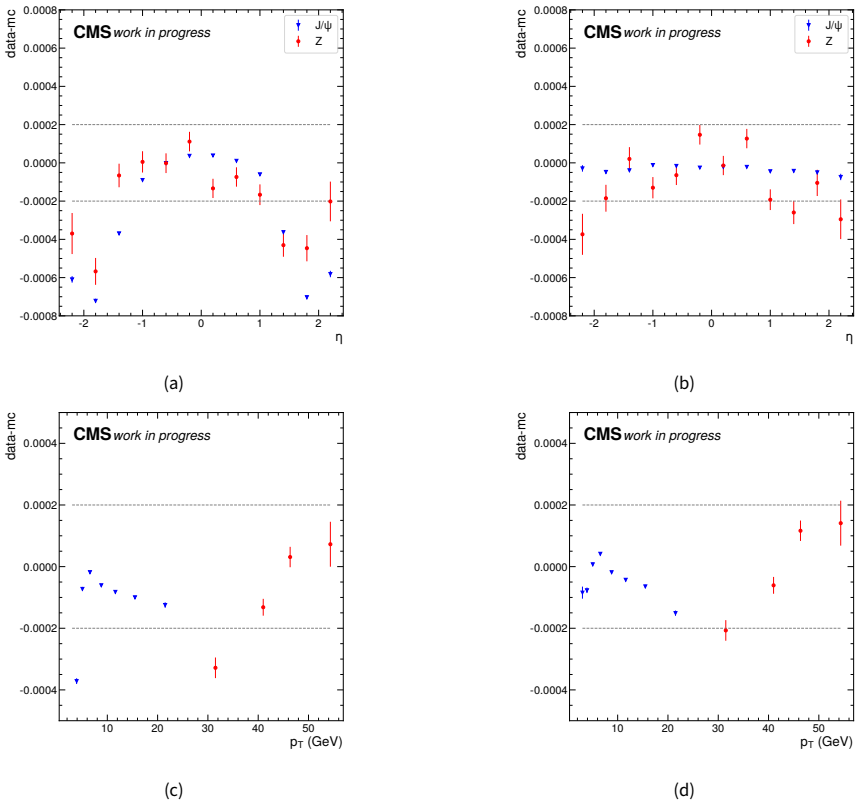


Figure 4.16: Agreement between measured scale in data and simulation before (left panels) and after (right panels) applying the corrections. This is shown as a function of η of the positive muon (top panels) and p_T of the positive muon (bottom panels).

a huge number of low p_T muons and the corrections can be extrapolated to a different phase space on Z events. The closure test computed with the Z shows a very good agreement between data and simulation, at the level of $2 \cdot 10^{-4}$ in the whole phase space. Face value, this number propagates to the final measurement of m_W to about 16 MeV. However, in practice, the analysis is performed on many bins in η and p_T of the muon and the final uncertainty on m_W is given by the combination of all the points in the phase space, which amounts to $(6 \pm 1) \cdot 10^{-5}$. In addition to this, an uncertainty will be propagated using the full covariance matrix of the parameters. This contribution is anyways expected to be small, of the order of the error bars of the plot in Figure 4.15a which is the dominant correction in the Z phase space. This procedure will be Illustrated when the fit for assessing the final uncertainty on m_W will be described, in Chapter 6.

To conclude, the muon momentum scale calibration outlined in this thesis can be improved in the future if an improved simulation of J/ψ is produced. Not only should the underlying spectra of muon η and p_T agree to a better level, but it should also be produced with pile-up, so that a combined fit of J/ψ and Z can be performed to extract the resolution parameters as it has been done for data in

Section 4.5.1. Matching the resolution of data and simulation in J/ψ and Z can remove some residual non-closure which can be of the order of 10^{-4}.

Chapter 5

Measurement of W differential cross sections

This Chapter presents a practical realisation of the analysis conceived in Section 1.3 using data collected by the CMS experiment in 2016 at $\sqrt{s} = 13$ TeV, for an integrated luminosity of 35.9 fb^{-1}. The aim of this measurement consists in the experimental assessment of the double diffential cross section of W q_T, y and the angular coefficients as a function of q_T and y.

First of all, W events are selected from data and are expected to be affected by a number of backgrounds whose entity can be reduced with the selection, but not completely removed, since their kinematics partially overlaps with the signal spectra. The main reducible backgrounds in this analysis are muons from semileptonic decays of heavy flavour mesons and Z decays to two muons, in case one muon escapes detection. Muons from W boson decays to τ are also considered a source of background. Minor reducible backgrounds, contributing to less than the percent level to the overall yield include muons from $t\bar{t}$ and single top, and diboson events: ZZ, WW and WZ. An accurate description of the samples used, their calibrations and the event selection will be given in section 5.1, while details of the treatment of the backgrounds will be extensively discussed in section 5.3.

The main step of this analysis is the construction of the signal templates, distributions of p_T and η of the muons used to unfold the cross sections of the W boson production described in Section 1.3. This is done using a generator of W events decaying to muons. Distributions of muon p_T and η are generated in bins of W q_T and y, for spherical harmonic according to Eq. 1.21. The angular coefficient decomposition holds in the hypothesis of the vectorial nature of a W boson decaying to two fermions, and while constructing the templates it is crucial to restore this regime. This allows to decouple the W production from its decay products so that the information on W production is encoded in the value of the angular coefficients and the unpolarised cross section in each bin of q_T and y. The procedure introduced in section 5.2, used to build the signal templates, is independent on the details of the generator used in the simulation, therefore it is possible to unfold the double differential cross section of q_T and y and the angular coefficients fitting the sum of signal and background templates to the distribution of muon p_T and η measured in data.

In this process, it is essential that the propagation of the generated muon inside the simulation of the CMS detector is accurate enough, so that theoretical and experimental effects can be decoupled using well calibrated templates. The fit is performed in the hypothesis that all differences between data and simulation is due to W production, and this is ensured applying systematic uncertainty of experimental origin to signal and background templates. In addition, some systematic uncertainties of theoretical origin are applied to background templates.

5.1 Measured and simulated event samples

For this analysis events have been selected using a single muon trigger, corresponding to the logical OR of multiple HLT paths. The paths used in this specific case require at least one loosely isolated muon in the event, with some requirements on the track quality, $p_T > 24$ GeV and $|\eta| < 2.5$.

Simulations of signal and background processes used throughout the analysis have been generated with different Monte Carlo generators:

- POWHEG[81] with MiNNLO$_{PS}$[82, 83] interfaced with PYTHIA8[79] and PHOTOS[80] has been used for $W \to \mu\nu$, $W \to \tau\nu$, $Z \to \mu\mu$ and $Z \to \tau\tau$. For all these processes, this package provides a description of the event kinematic accurate up to next-to-next-leading-order QCD corrections, besides to the soft or collinear radiation included in the parton shower provided by PYTHIA8, and leading-order electroweak corrections;

- POWHEG interfaced with PYTHIA8 has been used to generate $t\bar{t}$ events;

- MC@NLO[85] to generate single top events;

- PYTHIA8 for diboson events: ZZ, WW and WZ.

All Monte Carlo samples have been processed through the full simulation of the CMS experiment based on GEANT4[68] and reconstructed using the same algorithms used for data. A realistic detector alignment is included in the simulation, as well as the simulated HLT paths used to select data. Simulation of in-time and out-of-time interactions (pile-up) is also included, and events are weighted so that the number of reconstructed vertices of the simulation matches the one of the data.

During the first period of data taking in 2016, the tracker of CMS has been affected by a severe inefficiency in the strip modules. Due to the increased luminosity, the read-out chips have been experiencing deadtime after events depositing a large amount of charge (highly ionising particles) and this caused the chips to become inefficient for a number of bunch crossing. As a result, hits were lost and reconstructed tracks were shorter. This effect has finally been corrected with a tune of the "VFP" parameter. However, since the reconstruction conditions of data before and after the fix are different, the 2016 dataset has been divided in "preVFP" and "postVFP". Dedicated simulations have been provided with the inclusion of this effect, and both "preVFP" and "postVFP" simulations have been corrected with the related efficiencies for muon identification, tracking, trigger and impact parameter described in Section 2.5.2. 'preVFP" and "postVFP" will be treated in the following as two different datasets and fitted simultaneously.

Finally, the early 2016 dataset has been affected by the muon prefiring: due to a phase between the ECAL read-out and the trigger, caused by radiation damage, an event can be assigned to the previous bunch crossing (i.e. the trigger prefires). The rules of the trigger will then prevent this event to be recorded resulting in a loss of efficiency. This effect has been measured using a sample of unprefirable events and the assessed prefiring probability has been used to correct the simulation. All the simulated events are weighted to match the integrated luminosity of the data sample.

5.1.1 Event selection

W events are selected in data and simulated samples requiring exactly one identified muon in the event, with $25 < p_T < 65$ GeV, $|\eta| < 2.4$ and with small impact parameters $d_{xy} < 0.05$ cm and $d_z < 0.2$ cm. The latter is used to suppress the contributions from QCD and top, where muons decaying from b quarks can have large impact parameters. In addition, the identified muon must have fired the trigger. At this stage of the analysis no requirements are made on muon isolation and on transverse mass m_T since all the events in the m_T-isolation plane will contribute to the final fit.

5.1.2 Calibration of physics object

The selected events are then calibrated before being used for the measurement. Due to time constraints, we have not applied the calibration described in Chapter 4. Moreover, the missing transverse energy is calibrated using the Jet Energy Corrections computed by CMS.

5.2 Signal modelling

The main step of the analysis is modelling the signal using the phenomenological prescriptions discussed in section 1.3.3. The goal is obtaining a series of templates of muon p_T and η to extract the double differential cross section of W q_T, y and the angular coefficients, Eq. 1.21. Each template corresponds to one histogram of muon p_T and η in a bin of the $q_T - y$ plane[1] where all the muons have decayed according to one angular decay mode described by one spherical harmonic. This is done in two steps: first of all the simulation of the signal at generator level is used to derive the weights in each q_T and y bin needed to select one particular decay mode of the muon. Then, the weights are applied one at a time to a distribution of muon p_T and η in a bin of q_T and y.

5.2.1 Derivation of weights for angular decay modes

Weights are derived from a simulation of the signal at generator level with full acceptance. The characteristics of the POWHEG and PHOTOS generators allow to select generated muons before final state radiation: this is crucial in order to restore a regime of pure QCD where Eq. 1.21 holds. The angles θ^* and ϕ^* of these so defined preFSR muons in the Collins-Soper frame are used to compute the value of the angular coefficients in the $q_T - y$ plane. This is done using the method of momenta proposed in [86].

The momentum of $f(\theta^*, \phi^*)$, for specific values of q_T and y is defined as:

$$\langle f(\theta^*, \phi^*) \rangle = \frac{\int_{-1}^{1} d\cos\theta^* \int_0^{2\pi} d\phi^* f(\theta^*, \phi^*) d\sigma(\theta^*, \phi^*)}{\int_{-1}^{1} d\cos\theta^* \int_0^{2\pi} d\phi^* d\sigma(\theta^*, \phi^*)}, \tag{5.1}$$

[1] Since y is symmetric with respect to 0, due to the symmetry of the beams, only its absolute value will be considered in the following.

The value of each angular coefficient, A_i, can be extracted evaluating the momentum of the related spherical harmonic:

$$\begin{aligned}
\langle\frac{1}{2}(1-3\cos^2\theta^*)\rangle &= \frac{3}{20}\left(A_0-\frac{2}{3}\right), & \langle\sin\theta^*\cos\phi^*\rangle &= \frac{1}{4}A_3, & \langle\sin(2\theta^*)\sin\phi^*\rangle &= \frac{1}{5}A_6,\\
\langle\sin(2\theta^*)\cos\phi^*\rangle &= \frac{1}{5}A_1, & \langle\cos\theta^*\rangle &= \frac{1}{4}A_4, & \langle\sin\theta^*\sin\phi^*\rangle &= \frac{1}{4}A_7. \quad (5.2)\\
\langle\frac{1}{2}\sin^2\theta^*\cos(2\phi^*)\rangle &= \frac{1}{20}A_2, & \langle\sin^2\theta^*\sin(2\phi^*)\rangle &= \frac{1}{5}A_5,
\end{aligned}$$

These momenta are evaluated using the simulation, selecting events in bins of q_T and y and exchanging integrals for sums. Figure 5.1 shows the results of these computations. The values of the angular coefficients are plotted as a function of q_T and y. At leading order in α_S, W has only longitudinal momentum, $q_T = 0$ and A_4 is the only angular coefficient different from zero as a function of y. At $O(\alpha_S)$ only A_0, A_1, A_2, A_3 and A_4 are different from zero. Moreover, the Lam-Tung relation holds [87], stating that $A_0 = A_2$. At $O(\alpha_S^2)$, with which the plots in Figure 5.1 are produced, the Lam-Tung relation is violated and A_5, A_6, A_7 acquire a contribution from gluon loops, anyways they remain much smaller than the other coefficients manifesting at $O(\alpha_S)$.

5.2.2 Construction of the signal templates

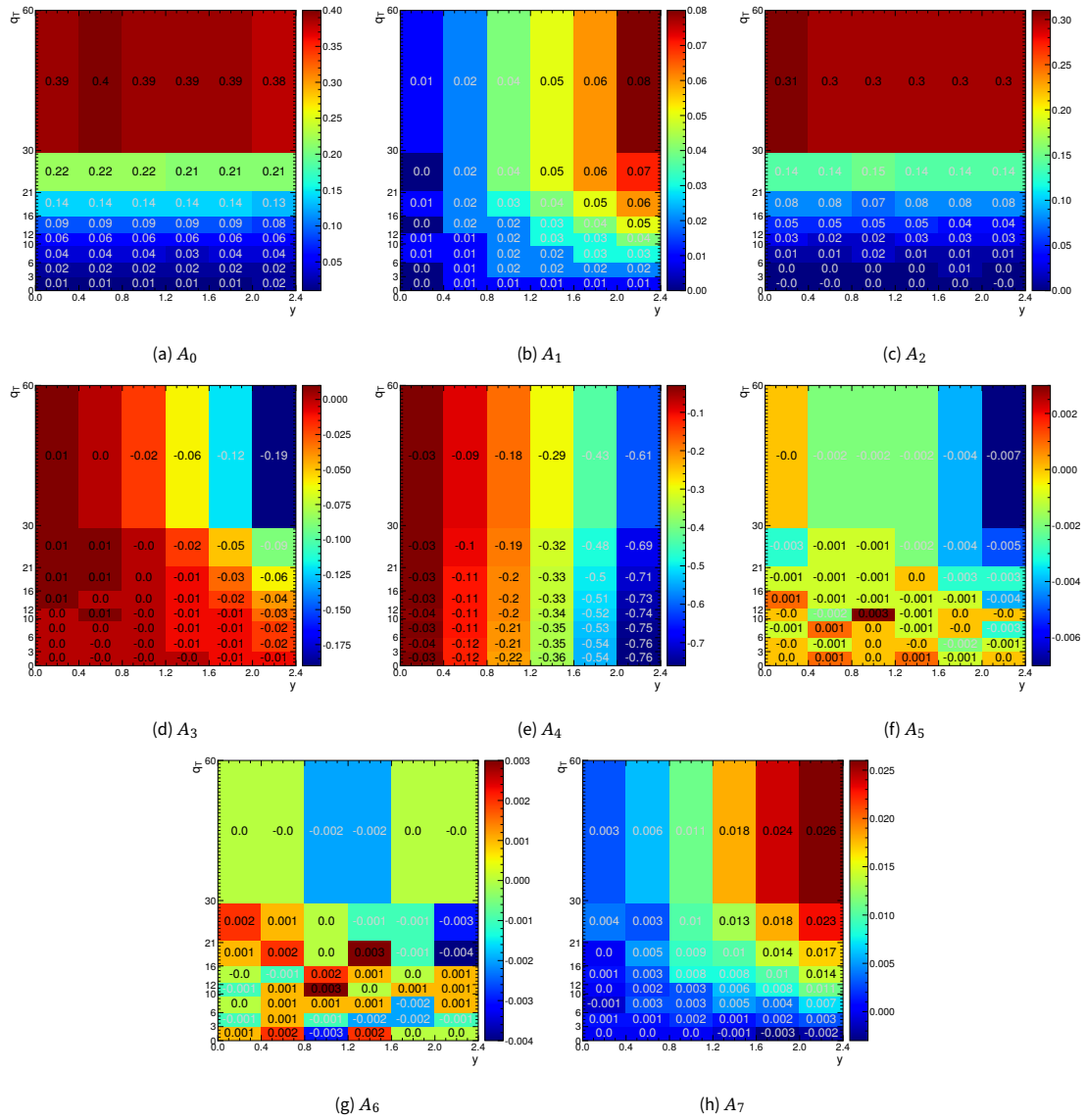
Using the values of the angular coefficients in the plane of q_T and y it is possible to construct the signal templates. Given q_T and y, there is a 2:1 relation between $\cos\theta^*$,ϕ^* and muon p_T and η, as discussed in Section 1.3. Therefore, after selecting a bin of q_T and y, each histogram of muon p_T and η can be reweighted according to:

$$w_i = \frac{A_i P_i(\cos\theta^*,\phi^*)}{\sum_{j=0}^{7} A_j P_j(\cos\theta^*,\phi^*) + 1 + \cos^2\theta^*}. \quad (5.3)$$

The function at the denominator $\sum_{j=0}^{7} A_j P_j(\cos\theta^*,\phi^*)+1+\cos^2\theta^*$ is numerically equal to the differential cross section of q_T, y, $\cos\theta^*$,ϕ^* and as a result it will remove any angular dependence of the muon from the boson matrix element, making the distribution flat in the $(\cos\theta^*,\phi^*)$ plane. Then, the numerator will morph the distribution according to the i-th spherical harmonic, normalised to the expected number of events.

After reweighting with the weights derived on the whole phase space with full acceptace, the templates of muon p_T and η are constructed for each bin of q_T and y with the full simulation of the CMS experiment using the same selection and reconstruction algorithms employed for data. The information related to the W production is encoded in the normalisation of each template, which will be fitted from data. The width of the bins of q_T and y is chosen to ensure that the dependence on the shape of the underlying distributions affecting the templates is negligible, so that the only relic dependence on the generator is the model used for final state radiation.

In order to have fully populated distributions, the signal templates are limited to the region $q_T < 60$ GeV and $|y| < 2.4$, where the acceptance is higher than 20%. A plot of the acceptance in W^+ and W^- is shown in Figure 5.2. It encodes both the effects from geometrical acceptance and reconstruction efficiency. We observe that the scale in the z axis of these plots goes up to about 0.5. The dominant



(a) A_0 (b) A_1 (c) A_2

(d) A_3 (e) A_4 (f) A_5

(g) A_6 (h) A_7

Figure 5.1: The value angular coefficients extracted from the simulation of the signal at generator level at $O(\alpha_S^2)$ as a function of q_T and y.

effect in this computation is due to the acceptance in muon p_T.

The events in the region of phase space $q_T > 60$ GeV or $|y| > 2.4$ are gathered in a single template that will be considered a source of background. Figure 5.3 shows a collection of signal templates obtained with this procedure.

5.3 Background modelling

The background sources in this analysis can be divided in two categories. The first one contains the prompt muons originating from electroweak processes (Drell-Yan, $t\bar{t}$, single top and diboson production) producing a single isolated muon in the signal region. the second category includes non-prompt muons from heavy flavour decays in QCD multijet events. These two classes of events are treated differently in the analysis. The backgrounds of electroweak origin are well modelled by the Monte Carlo simulations and can therefore be directly subtracted using this prediction. On the other hand, the contribution from non-prompt muons from QCD multijet events is not well modelled in the simulation and has to be estimated from data using the fake rate method.

5.3.1 The fake rate method

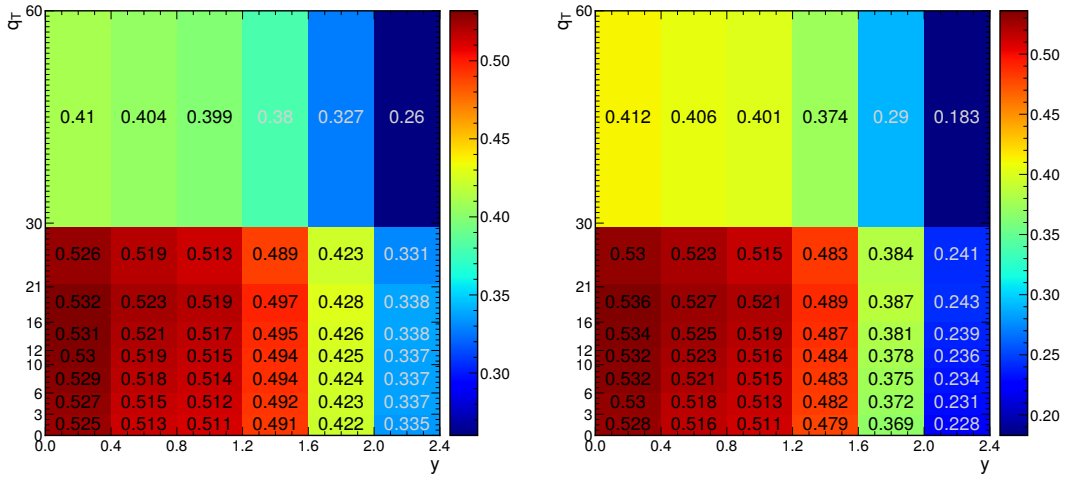
Let us consider the plane in phase space defined by the transverse mass m_T and the muon isolation. These variables are uncorrelated in QCD multijet events, therefore cuts can be applied to one of these variables without affecting the other. We define four regions in this plane:

low m_T isolated region (LI): $m_T < 40$ GeV and isolation< 0.15;

low m_T anti-isolated region (LA): $m_T < 40$ GeV and isolation> 0.15;

high m_T isolated region (HI): $m_T > 40$ GeV and isolation< 0.15;

high m_T anti-isolated region (HA): $m_T > 40$ GeV and isolation> 0.15;



(a) (b)

Figure 5.2: Signal acceptance in the plane of W q_T and y, in the range considered in the fit, estimated using the W^+ (Figure 5.2a) and W^- (Figure 5.2b).

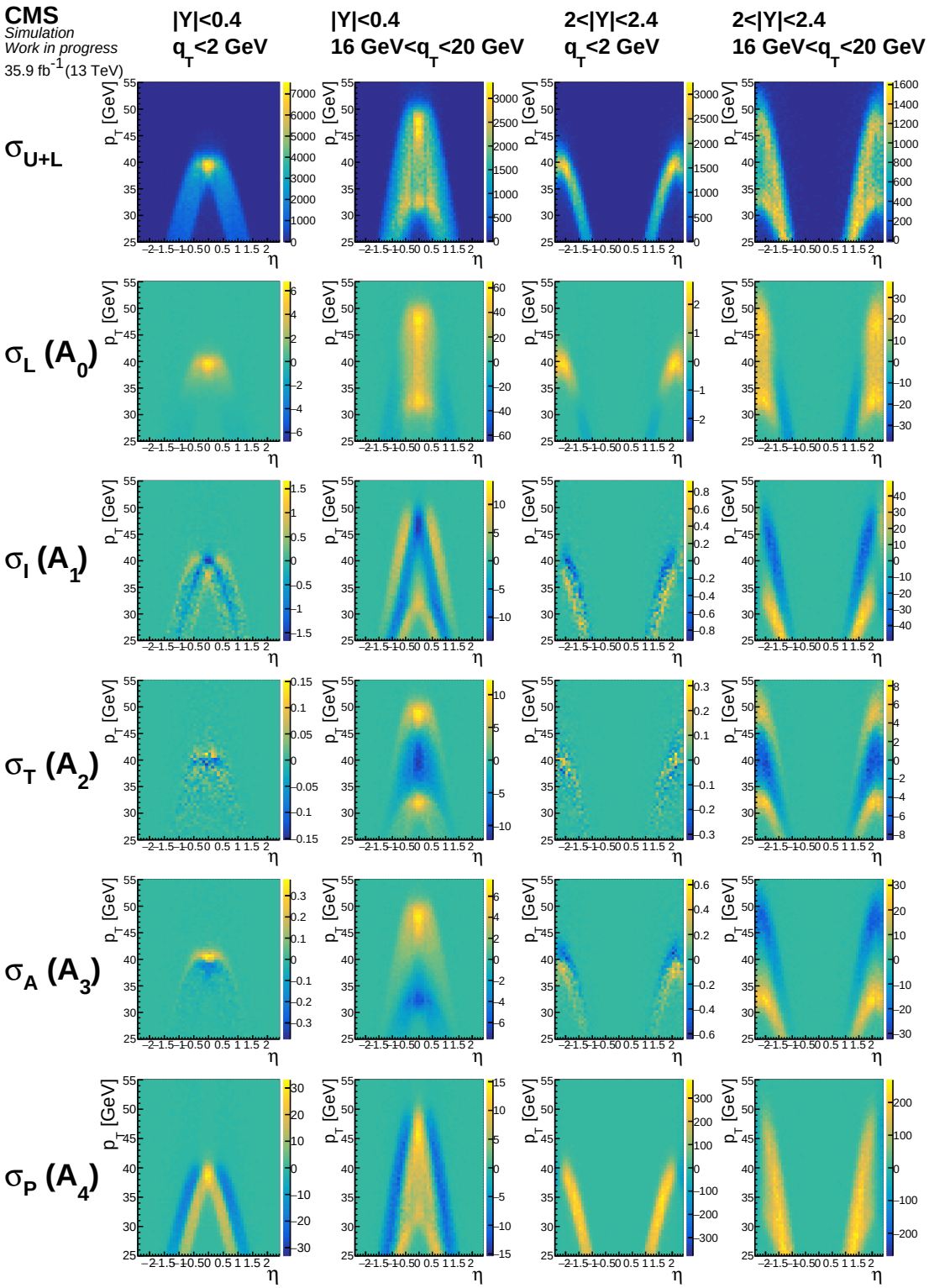


Figure 5.3: Representative examples of the signal templates for the five angular coefficients and σ^{U+L} (each row) in different bins of $|Y_W|$ and q_T^W (each column). The $W^+ \to \mu^+\nu$ sample has been used. The z axis scale is different in each plot and represents the equivalent number of events.

We will use these regions to define the fake rate as the background selection efficiency. The high m_T region is populated by signal events, while the low m_T region is mostly populated by QCD events. Since m_T and isolation are uncorrelated for QCD events, the equivalence

$$\text{fake rate} = \frac{N_{HI}}{N_{HI} + N_{HA}} = \frac{N_{LI}}{N_{LI} + N_{LI}} \tag{5.4}$$

holds. Therefore, it is possible to measure the fake rate in the low m_T region and extrapolate it to the high m_T region, where most of the interesting events are located.

5.3.2 Estimation of the fake rate

We can obtain a rough estimate of the fake rate in the LI, LA and HA regions from data subtracting the electroweak component using the prediction from Monte Carlo. Then, using Eq. 5.4, we can compute the number of QCD events in the HI region as:

$$N_{HI} = \frac{N_{LI}}{N_{LI} + N_{LI}} \cdot (N_{HI} + N_{HA}) \tag{5.5}$$

While this procedure is not meant to obtain the ultimate measurement of the number of QCD events, it allows to define a template to be used in the final fit, where the fake rate estimation will be allowed to float and to be adjusted with the correct yields of the electroweak components fitted from data.

The fit is performed in 4-dimensions, with all the four regions contributing to the minimisation. This allows to constrain the number of events due to QCD multijets in situ. This is done with the following procedure. Two sets of nuisance parameters are defined. The first set is allowed to change the normalisation of the QCD background independently in the low and high m_T region. In the second set, one nuisance parameter per $\eta - p_T$ bin is allowed to change the fake rate simultaneously in the low and high m_T region, as per Eq. 5.4. These nuisance parameters do not make any constraint on the likelihood and are freely-floating. This is done during the minimisation of the other parameters, therefore the fit will automatically adjust the electroweak component that has to be subtracted from data in Eq. 5.4. In the end of the minimisation, the fit is expected to show a good agreement between data and simulation in all the four regions.

5.3.3 Templates for backgrounds

Figure 5.4 shows the templates for the background processes obtained with the samples and selections described in Section 5.1, in the HI region. Figure 5.4f shows the template for QCD events obtained with Eq. 5.5 and used as input for the fit.

5.4 Systematic uncertainties

The predictions for the signal and background yields and kinematics are taken from Monte Carlo simulations, as described in section 5.1. They are affected by systematic uncertainties of experimental

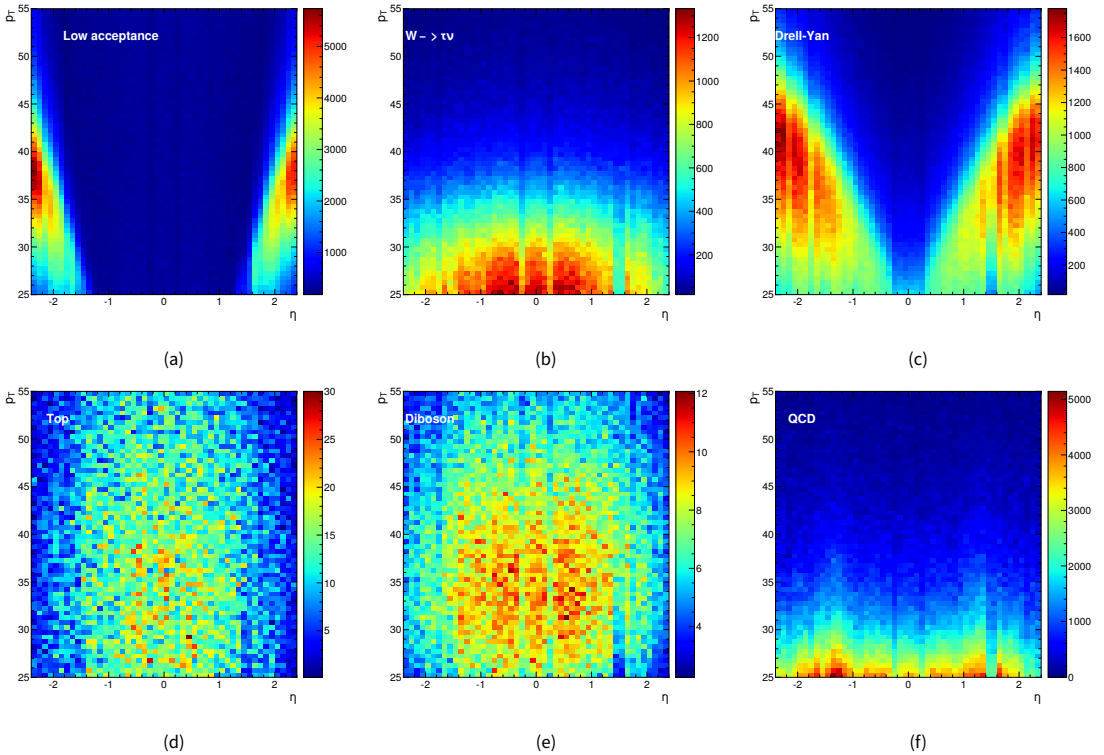


Figure 5.4: Templates for the background processes obtained with the samples and selections described in Section 5.1, in the HI region. Figure 5.4a shows the template for the signal events falling outside the fit range. Figure 5.4b, 5.4c, 5.4d, 5.4e,5.4f show respectively the templates for W to τ, Drell-Yan, top, diboson and QCD processes.

and theoretical origin, which are taken into account in the measurement. In order for the fit to work unbiasedly, it is of utmost importance that the simulation is well calibrated so that any discrepancy is only due to the underlying W kinematics to be fitted. This will be guaranteed by assigning a correct systematic to each ingredient.

5.4.1 Experimental systematic uncertainties

Jet Energy Corrections

Muon scale and resolution calibration

Efficiency Scale Factors The scale factors for efficiency described in Section 2.5.2 have statistical and systematic uncertainties. The statistical uncertainty takes into account the size of the sample used for the derivation of the efficiencies while the systematic uncertainty is obtained with an alternative fitting model. While the systematic uncertainty is taken to be fully correlated in the η-p_T plane, the statistical uncertainty is decorrelated in bins of the same granularity as the efficiency derivation in η-p_T.

Trigger prefire correction

Luminosity The CMS experiment has measured the luminosity of the dataset used in this analysis with an uncertainty of 1.2%. Since the overall normalisation of the simulated samples is based on the value of the integrated luminosity, a systematic uncertainty of 1.2% on the total event yield of each simulated sample has been considered.

5.4.2 Theoretical systematic uncertainties

While this analysis is designed to be insensitive to the W production modelled in the simulation, some background processes can still show a small dependence on it. Therefore, systematics of theoretical origin are applied as follows.

PDFs The simulated samples of W and Z have been generated using NNPDF31[37]. This set is distributed with 102 hessian variations, encoding alternative sets of PDFs and α_S. As previously discussed, the signal templates are not sensitive to the values of the PDFs used in the simulation. However, the events which fall out from the boundaries defined for the fit range in q_T and y, gathered in the low acceptance template, rely on the description of the PDFs in the simulation and therefore a systematic uncertainty due to this effect will be applied in the fit, correlated in the $\eta - p_T$ plane. The same uncertainty is applied to W to τ events and to Drell Yan events.

W and Z transverse momentum As for the PDFs, the low acceptance, W to τ and Drell Yan templates depend on the underlying boson transverse momentum. An uncertainty accounting for this effect is applied using the following recipe proposed in [49], that makes use of the renormalisation (μ_R) and factorisation (μ_F) scales. For the Drell Yan template, six alternative versions have been considered. The sample has been reweighted multiplying μ_R and/or μ_F for a factor 2 or 0.5, discarding the extremal cases to satisfy $0.5 \leq \mu_i/\mu_j \leq 2$, as described in Ref. [88]. The low acceptance template and W to τ template have been reweighted in a more refined way, allowing the variation to describe a less constrained q_T dependence. The six scale variations have been split in three q_T bins each (low [0,5 GeV], mid [5 GeV,15 GeV], high [15 GeV,∞]), and an alternative template 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 taken fully correlated in the $\eta - p_T$ plane.

Final state radiation

Electroweak background cross sections An additional uncertainty on the electroweak cross sections of the background channels has been considered. All these uncertainties are fully correlated in the $\eta - p_T$ plane. A 4% uncertainty has been considered for W to τ 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.

5.5 Global Fit

The information synthetised in the signal and background templates is condensed in a global fit aimed at unfolding the double differential cross section of q_T and y and the angular coefficients from the observed distribution of muon p_T and η. The fit is performed using an extended binned maximum likelihood fit, independently for each W boson charge. The observed muon p_T and η distribution is fitted to the sum of the templates of the signal and background processes in the four regions of the phase space in m_T and isolation. The fit is allowed to modify the freely-floating signal strength modifiers and to unfold the related cross section. Nuisance parameters associated to the systematic uncertainties discussed in section 5.4 contribute to the minimisation and can act on the freely-floating signal strength modifiers, on the fixed background strength modifiers, and to the unfolded cross sections. Details of the technical implementation of the fit will be given in Section 5.5.2.

The results are presented in terms of the unfolded cross sections and angular coefficients. The physics parametrisation of Eq. 1.21 has the disadvantage of not being linear in the fitted parameters, since the angular coefficients show the unpolarised cross section in the denominator. For this reason, the fit is performed using the decomposition in helicity cross sections presented in [39], which is instead linear in the fitted parameters. Since there exists a one to one correspondence between helicity cross sections and angular coefficients, the latter are computed after the fit using the full covariance matrix.

The adopted procedure allows to obtain the first measurement of the double differential unpolarised cross section of the W boson. Of notable interest are the integrated measurements of the W cross section in rapidity and q_T with an uncertainty which is competitive with the state of the art calculations. While the first has been recently published by CMS [49], the second is a totally new result which is important per se, and can also be used improve the measurement of m_W. This procedure also allows to measure for time the first the angular coefficients as a function of W q_T and y but in this case the precision is smaller and the purpose of including them in the fit is limited to the constrain in situ of the degrees of freedom needed for an agnostic m_W measurement. Among the angular coefficients, A_4 plays a major role since it is the only one manifesting at leading order and encodes the majority of the dependence of the W production from the PDFs. In this Chapter, results for the unpolarised cross section and for A_4 are shown as a function of q_T and y and also integrated in one variable at a time.

The fit is performed in three different configurations. The simplest one is the fit to the “Asimov” dataset, consisting in the sum of all the templates, normalised to match the data luminosity. In this dataset the signal strength modifiers are all ones by definition, and no real minimisation takes place. The central values and nuisance parameters are kept fixed to the prediction and the uncertainties are propagated to the final results. This allows to check the consistency of the fitted model and the covariance matrix and the estimation of the errors obtained with this procedure are representative of the final uncertainties in the fit to real data.

A further step in the fit validation is defining an “Asimov” dataset changing the values of the signal strength modifiers and let the minimiser assess them. In this way, the unbiasedness of the fit can be checked while factorising the experimental effects on the templates from the theoretical ones. Randomising this modified ‘Asimov” dataset with Poissonian fluctuations it is possible to show the robust-

ness of the fit on many toy experiments.

Finally, the fit can be run on real data. For each fit configuration, the results will be presented in terms of unfolded cross sections and angular coefficients.

5.5.1 Inputs to the global fit

The fit is performed on the four regions of m_T and isolation defined in section 5.3. In each region, templates for signal and background processes are computed. The templates have been derived using the analysis framework discussed in Chapter 3. For the signal, 7-dimensional histograms are used: η, p_T, charge, m_T, isolation, y and q_T, one for each helicity cross section. For the backgrounds and for observed data, 5-dimensional histograms are used: η, p_T, charge, m_T and isolation. This configuration allows to obtain all the templates for both charges and for all the four regions at the same time with an extremely good performance, less than one hour in real world time.

These tensorial histograms are then converted to `numpy` arrays and projected along the η, p_T, m_T and isolation axes and fed to the fit. Each charge is fitted individually. After the projections, we have 6 times 8 signal templates for as many bins in y and q_T for each helicity cross section. y is binned with equally-spaced bins from 0 to 2.4, while q_T is binned from 0 to 60 GeV with edges corresponding to $(0, 2, 3.7, 5., 6.3, 8, 9.8, 12, 15.2, 19, 24.5, 33.2, 60)$, chosen to equipopulate the bins. η is binned with 48 equal bins from -2.4 to 2.4 and p_T with 60 equal bins from 25 to 55 GeV. These limits are defined by the detector acceptance and efficiency and the bin size is limited by the muon momentum resolution. The bins in m_T and isolation are chosen to define the four regions of section 5.3: one bin between 0 and 40 GeV and one between 40 and 100 GeV in m_T, and for the isolation one bin between 0 and 0.15 and one between 0.15 and 1.

In total, there are 432 processes for signal, 288 of which correspond to the helicity cross sections relative to A_0, A_1, A_2, A_3, A_4 and the unpolarised cross section and are freely-floating, and the remnants, the helicity cross section relative to A_5, A_6 and A_7,ware treated as background and they are fixed to prediction since the $\eta - p_T$ distribution in data is not sensitive to them, as discussed in section 1.3. One process is dedicated to each electroweak background and two separated processes are defined for the QCD background: one for the low m_T region and one for the high m_T region, defined as described in section 5.3.

5.5.2 Technical implementation of the fit

The negative log-likelihood of the fit can be written as:

$$L = -\ln\left(\mathcal{L}(\text{data}|\boldsymbol{\mu},\boldsymbol{\theta})\right) = \sum_{i}^{\text{nbins}} \left(n_i^{\text{obs}} \ln n_i^{\text{exp}}(\boldsymbol{\mu},\boldsymbol{\theta}) + n_i^{\text{exp}}(\boldsymbol{\mu},\boldsymbol{\theta}) \right) + \frac{1}{2} \sum_{k}^{\text{nuisances}} (\theta_k - \theta_k^0)^2,$$

$$n_i^{\text{exp}}(\boldsymbol{\mu},\boldsymbol{\theta}) = \sum_{p}^{\text{processes}} \mu_p n_{i,p}^{\text{exp}} \prod_{k}^{\text{nuisances}} \kappa_{i,p,k}^{\theta_k}. \tag{5.6}$$

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} is the number of expected events per bin; $n_{i,p}^{\text{exp}}$ is 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 is the associated nuisance parameter; $\kappa_{i,p,k}$ is the size of the systematic uncertainty per bin, per process, per nuisance.

The signal strength modifiers μ_p of the signal processes are freely floating in the fit and represent the parameters of interest (POIs) of the fit. The signal strength modifiers for background processes are instead fixed to unity. The systematic uncertainties have been implemented with log-normal variations 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 $\kappa^{\theta_k} = 1$ before the fit). A unit Gaussian constraint to θ_k is added to $\mathcal{L}$ for each nuisance parameter.

The uncertainties and the covariance matrices for the POIs and the nuisance parameters are derived from the Hessian of the likelihood at the minimum:

$$\begin{aligned} V_{i,j}^{-1} &= -\frac{\partial^2 L}{\partial x_i \partial x_j}\Big|_{\vec{x}=\hat{\vec{x}}} \ , \\ \vec{x} &= \{\boldsymbol{\mu}, \boldsymbol{\theta}\}, \end{aligned} \tag{5.7}$$

where V^{-1} is the inverse of the covariance matrix of the POIs and the nuisance parameters and $\hat{\vec{x}}$ is the value of the parameters that maximises the likelihood.

The cross sections $\boldsymbol{\sigma}$ are unfolded simultaneously to the POIs fit using the predicted cross sections for each process. The fit is allowed to change the predicted cross sections through dedicated nuisance parameters.

Technically, the fit has been implemented within the framework of `Combine`, the package developed for the CMS Higgs measurements and widely used by the CMS analyses. The core of the fit, in term of minimisation and errors propagation, has been written using `TensorFlow`. Its features, and in particular the efficient calculation of the gradients with the backpropagation, allowing for a semi-analytical minimisation, guarantees the precision and robustness needed for this complex fit and it is particularly critical in the non-convex region of the likelihood, where the usual minimisers fail. Moreover, the implementation in `TensorFlow` allows the parallelisation of the fit and the optimisation of the memory consumption. The fit can run on both CPUs and GPUs.

5.5.3 Systematic uncertainties

As discussed in section 5.4, systematic uncertainties of experimental and theoretical origin affect the signal and background templates. They are implemented inside the fit as nuisance parameters of two categories: nuisance parameters affecting the normalisation of the templates and nuisance parameters affecting the shape of the templates. The normalisation uncertainties are fully correlated among all $\eta - p_T$ bins, while the correlations in the shape uncertainties are modelled depending on the systematic under examination. The summary of the systematics that have been taken into account is shown in Table 5.1.

Table 5.1: 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	Low-acc.	$N_{\text{nui.}}$
Data Luminosity	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
JES, E_U	shape	shape	shape	shape	shape	shape	2
p_T^μ scale	shape	shape	shape	shape	shape	shape	1
SF_{stat}	shape	shape	shape	shape	shape	shape	144
SF_{syst}	shape	shape	shape	shape	shape	shape	1
L1 trigger prefire	shape	shape	shape	shape	shape	shape	1
PDF	-	shape	shape	-	-	shape	100
m_W	shape	-	-	-	-	shape	1
q_T^Z (MC Scale)	-	shape	-	-	-	-	6
q_T^W (MC Scale binned in q_T^W)	-	-	shape	-	-	shape	18

5.5.4 Expected results

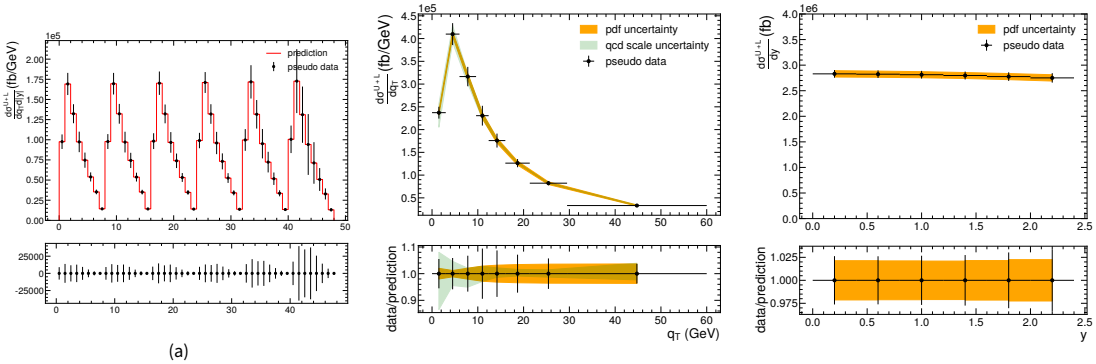
In this Section, fit results on the "Asimov" dataset are shown. In Figure 5.5a the double differential unpolarised cross section in as a function of W q_T and y, for W^+ , unrolled in one dimension. The related integrated plot in q_T as a function of y is shown in 5.5c and the integrated plot in y as a function of q_T is shown in 5.5b.

In Figure C.8 the analogue plots for A_4 are shown. Figure 5.7 and Figure 5.8 show the related plots for W^- . Since in the case of the expected results the central values are fixed to prediction, it is interesting to analyse the fit uncertainties that represent a projection of the expected uncertainty when fitting data. The uncertainties shown on data contain contributions of statistical and systematic origin. We observe from Figure 5.5b and 5.7b that the fit precision in the first bins of q_T is of order of 3%, which is better than the uncertainty band of the prediction, encoding the PDF uncertainty of NNPDF31 and the QCD scale variations of POWHEG-MiNNLO. Moreover, an additional uncertainty to the prediction has to be accounted for non-perturbative effects which are not shown in the plotted band. As for Figure 5.5c and 5.7c, the precision is of the order of 2%, which is comparable with the results of [49]. For what concerns A_4, the uncertainties are order of 200% as a function of q_T and below 40% as a function of y, less precise than the prediction.

In the analysis of the uncertainties in the fit it is interesting to understand the contribution of each group of nuisances to the total uncertainty of the double differential unpolarised cross section. One group is defined for each of the normalisation and shape uncertainty source listed in Table 5.1. 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. It is defined as:

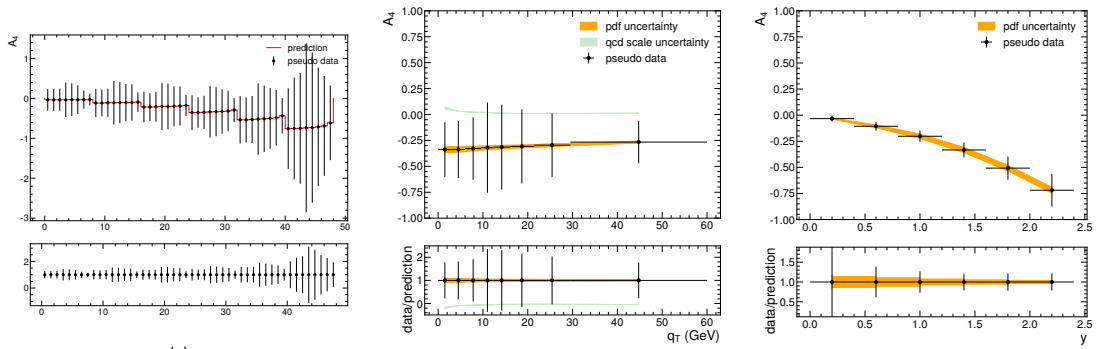
$$I_{\{\theta_k\}_G}(\mu_p) = \sqrt{C_{p,G}^T V_G^{-1} C_{p,G}}, \tag{5.8}$$

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



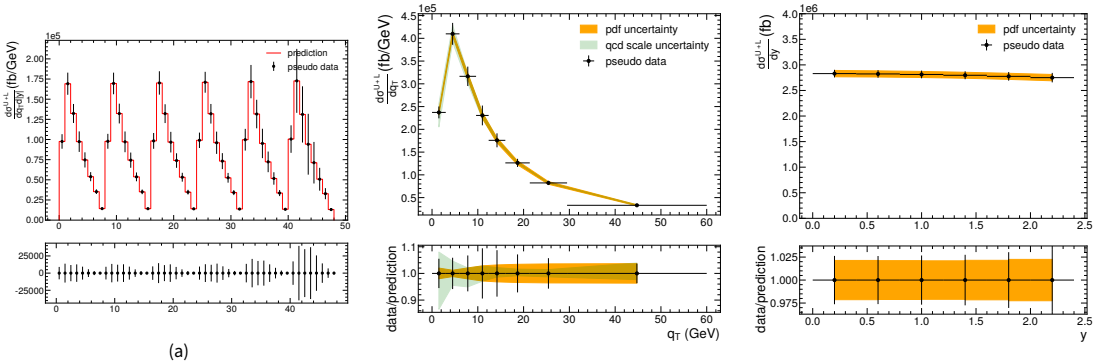
(a) (b) (c)

Figure 5.5: .



(a) (b) (c)

Figure 5.6: .



(a) (b) (c)

Figure 5.7: Placeholders for W^- .

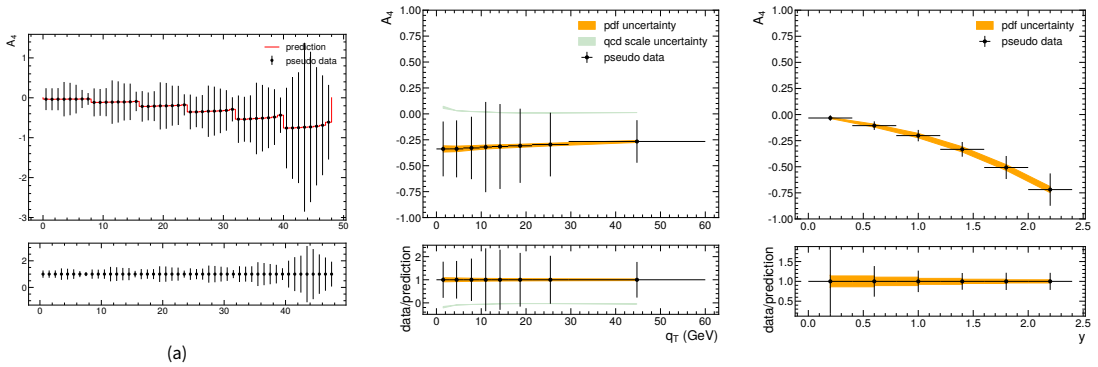
the statistical uncertainty, estimated from the covariance matrix without systematic uncertainties, but using post-fit nuisance parameters values. The uncertainty derived summing the group impacts can differ from the total uncertainty on the POIs because this neglects the correlations among different groups.

Figure 5.9a shows the relative impact due to each group of nuisance to each bin of the double differential unpolarised cross section in W^+. It can be observed that the bulk of the uncertainty is due to the statistical error and to the statistics of the simulation, varying from about 1% to 10%. These two contributions are expected to be similar since the number of events in the simulation is of the same order as the data. Among the other systematics, the uncertainty from luminosity dominates and amounts to sub-percent to about 5%.

Extending the concept of group impacts to include other POIs, it is possible to compute the contribution to the total uncertainty of the double differential unpolarised cross section from the angular coefficients. This is shown in Figure 5.9b, from which we can observe that the other angular coefficients contribute to the total uncertainty of the unpolarised cross section with similar fractions. This can be considered a breakdown of the statistical uncertainty shown in Figure 5.9a. These single fractions can be higher than the total uncertainty since, as mentioned before, the correlations among the the various coefficients are neglected to single out each contribution. Nonetheless, it is evident that the the highest contribution to the total uncertainty is given by the angular coefficients rather than the nuisance parameters in the fit.

This can be understood analysing the correlations among the fitted parameters. The correlation matrix of the unpolarised cross sections and the angular coefficients is shown for W^+ in Figure 5.10. Each block of the matrix gathers all the fitted bins in the q_T and y plane unrolled in the same logic as Figure 5.5a and C.8. The matrix for W^- looks very similar.

There are two kind of patterns that we can observe. Within the single blocks related to one coefficient or to the unpolarised cross section, there is a clear pattern of anticorrelation in the neighbouring q_T bins in the same bin of y. Milder correlation patterns are also visible between the same q_T bin of different y bins. The origin of these correlations can be traced back to the characteristics of the templates



(a) (b) (c)

Figure 5.8: Placeholders for W^-.

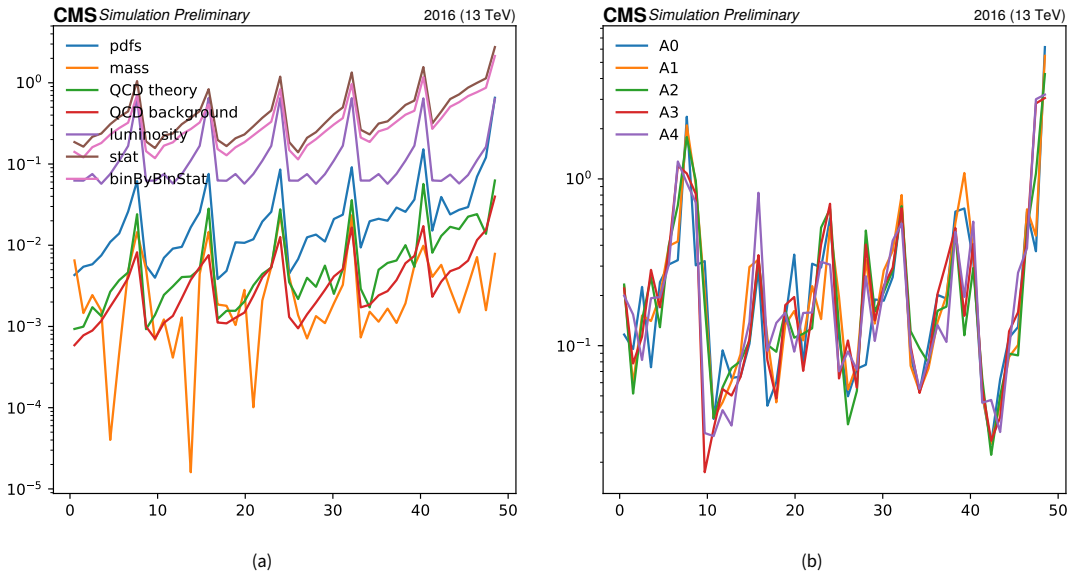
involved in the fitting procedure. Templates of nearby y and q_T bins compete for the same region in the η and p_T plane, since they change smoothly as a function of y and q_T, even overlapping partially. On the other hand we can also observe correlations among different coefficients. This kind of pattern can be understood looking at the contour of the various templates. In some regions of the η and p_T plane, templates relative to different angular coefficients can show a very similar contour, therefore the fit is not able to completely decouple the two.

The large (anti-)correlations will propagate on the fit result inflating the final uncertainties on the parameters. This explains the high impacts to the unpolarised cross section from the other angular coefficients. Moreover, the degeneracy of the templates can produce anomalous high-frequency oscillations in the predicted distributions. In practice, the correlations are spread-out across multiple bins and when integrating over either q_T or y, this effect is largely diluted and suppressed.

5.5.5 Validation on pseudo data

Before fitting real data, a further validation step can be done in order to demonstrate that the fit result has no dependence on the generator that has been used to prepare the signal templates. A "pseudo data" sample is constructed in the following way, using W^+ . The signal simulation has been reweighted to another prediction obtained using MC@NLO. Then, the backgrounds are also added, without any reweighting or modification. These pseudo data are then randomised and fitted in place of the real data against the templates built in Section 5.2. The starting values of the nuisance parameters are also randomised.

This simplified procedure allows to test the unbiasedness of the fit while decoupling the experimental systematic effects, which are identical in the templates and in the pseudo data by construction,



(a) (b)

Figure 5.9: .

from the theoretical ones. Therefore, the fit is performed keeping only the systematics of theoretical origin. The muon η and p_T distribution of the pseudo data obtained with this procedure, together with the input templates is shown in Figure 5.14, for postVFP channel. Figure 5.14a shows the unrolled distribution in the whole plane while Figure 5.14b and 5.14c show the integrated distribution, from which it can be appreciated the small deviation from 1 in the ratio panel due to the mismatched underlying W kinematics. The same distributions are displayed using post-fit values in Figure 5.15: with the W differential cross sections unfolded from data, the agreement in η and p_T is restored.

Figure 5.16 shows the result of the fit for the unpolarised cross section for one toy experiment: unrolled in Figure 5.16a, integrated in q_T as a function of y is shown in 5.16c and the integrated plot in y as a function of q_T is shown in 5.16b. The "true" value of the unpolarised cross section used to build the pseudo data is displayed with a blue line. All the data points agree within 1σ with its true value.

5.5.6 Final results

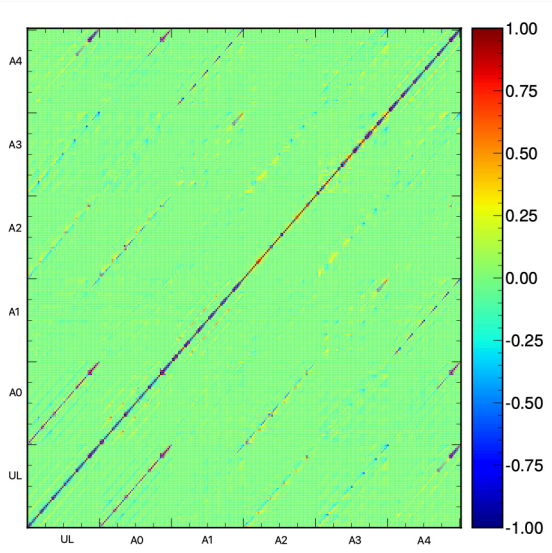
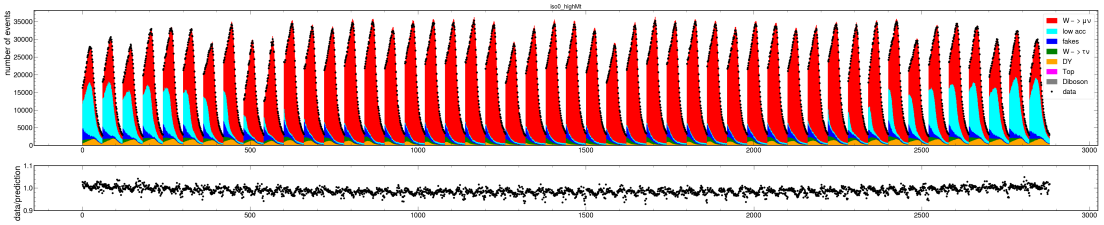
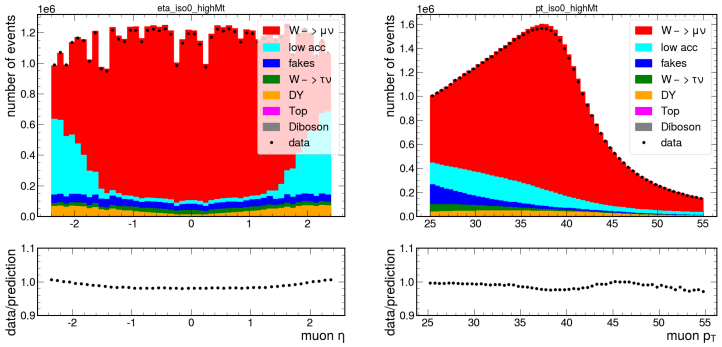


Figure 5.10

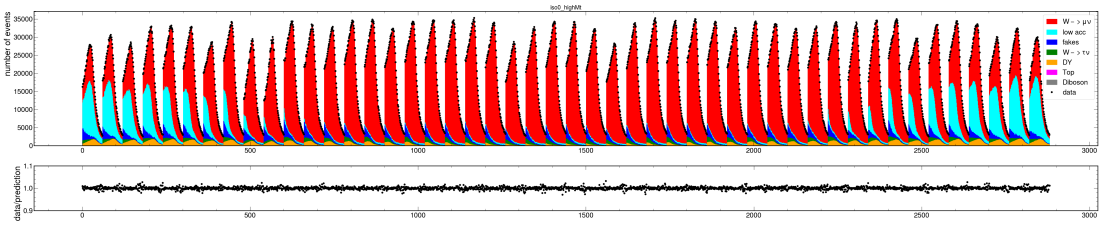


(a)

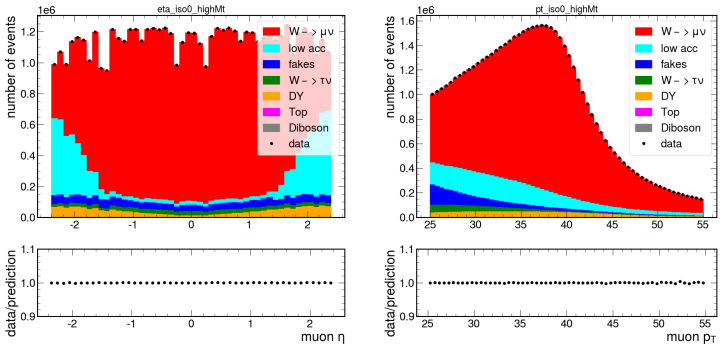


(b) (c)

Figure 5.11: .

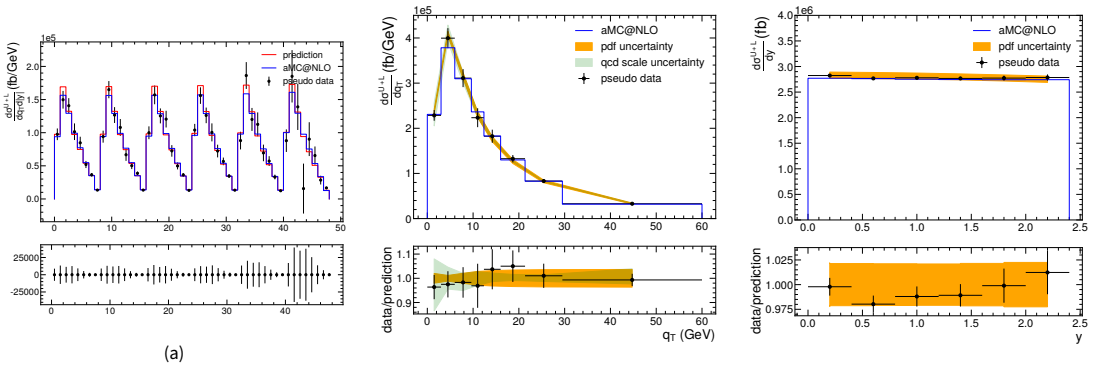


(a)



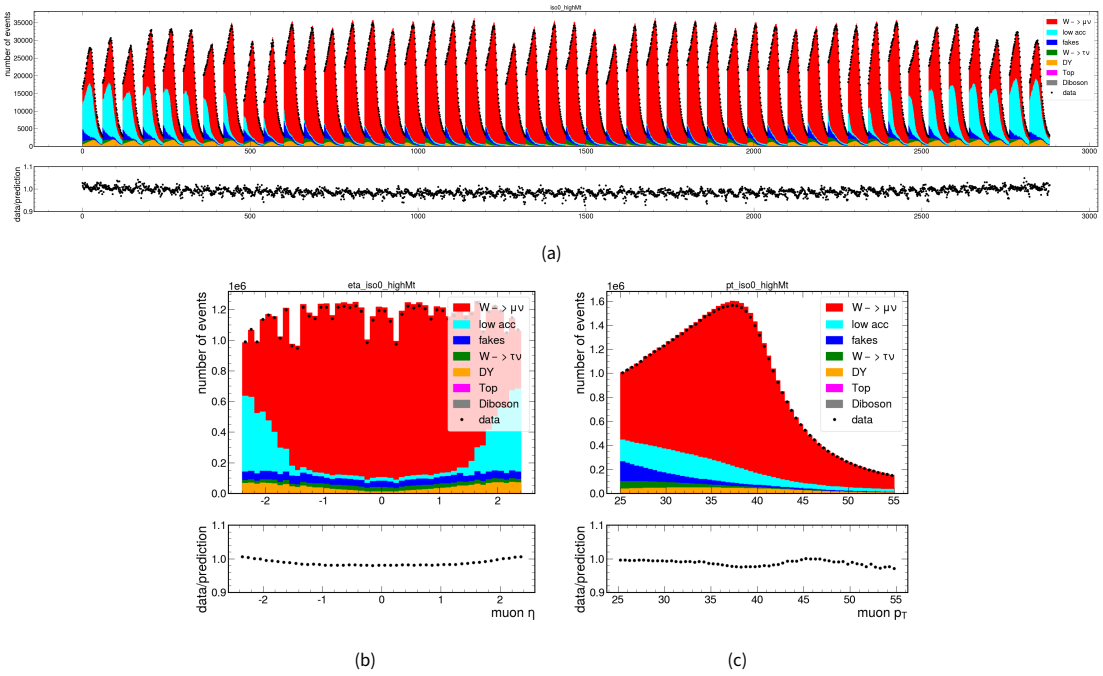
(b) (c)

Figure 5.12: .



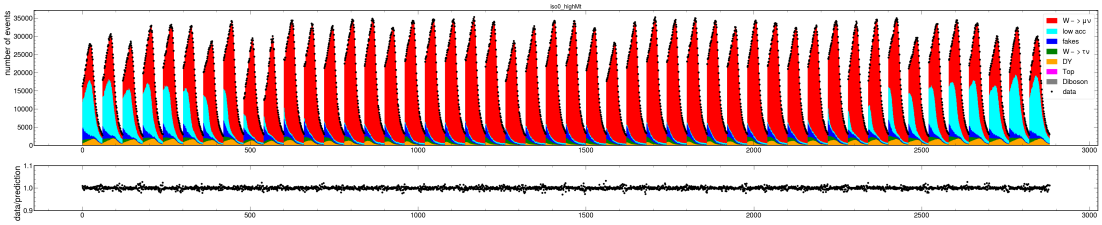
(a) (b) (c)

Figure 5.13: .

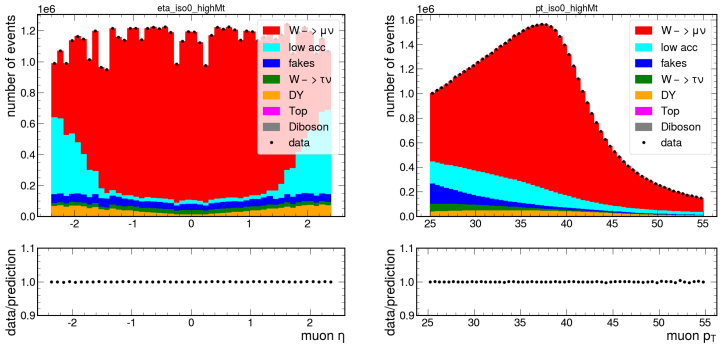


(a) (b) (c)

Figure 5.14: Placeholder for data.

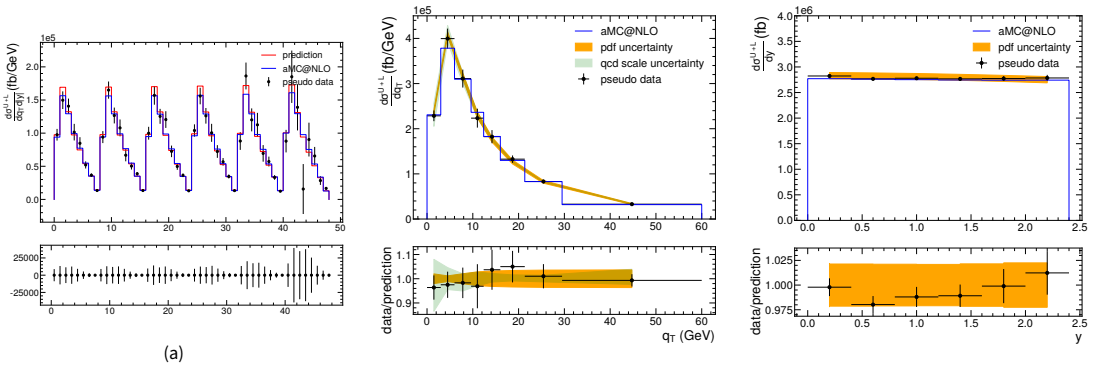


(a)



(b) (c)

Figure 5.15: Placeholder for data.



(a) (b) (c)

Figure 5.16: Placeholder for data.

Chapter 6

Assessment of the uncertainty on W mass

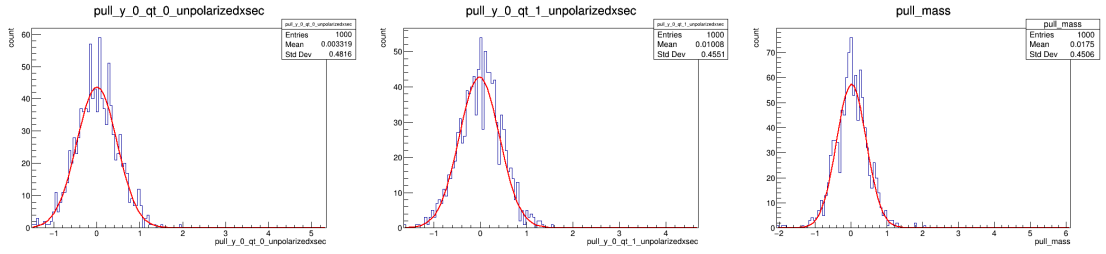
The phenomenological procedure described in Section 1.3 and its experimental realisation presented in Chapter 5 have allowed a precise measurement of the W production and especially W q_T. These measurements provide the necessary degrees of freedom to constrain the W production inside the detector acceptance when measuring the W mass.

The experimental setup described in Chapter 5 enables for a simultaneous measurement of the W differential cross sections and mass with limited technical modifications. Signal templates as a function of various mass hypotheses have already been produced to treat the W mass as a nuisance parameter. A measurement of the W mass can be performed removing its constraint on the likelihood and leaving it freely-floating in the fit. Moreover, W^+ and W^- can be fitted simultaneously to get a single value of m_W.

Before disclosing the fitted result on data, the toy study presented in Section 5.5.5 will be repeated to demonstrate that the fit is able to measure the W production and mass with no bias.

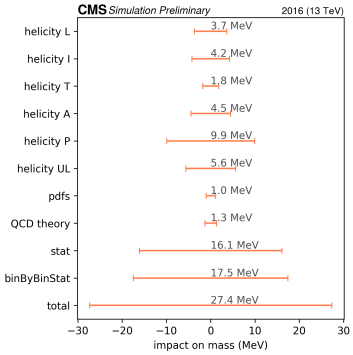
6.1 Validation on pseudo data

6.2 Final results



(a) (b) (c)

Figure 6.1: Placeholder for data.



(a)

Figure 6.2: Placeholder for data.

Appendix A

Effect of material in the track momentum resolution

Consider the tracker of length $2L$ shown in Figure A.1. There are $2N$ measuring layers equally spaced and there is no material, with the exception of a scattering layer positioned in the middle. If the material in the scattering layer is negligible, the scattering can be ignored and the momentum resolution σ_{p_T}/p_T of this tracker only contains the contribution from the hit resolution 4.8:

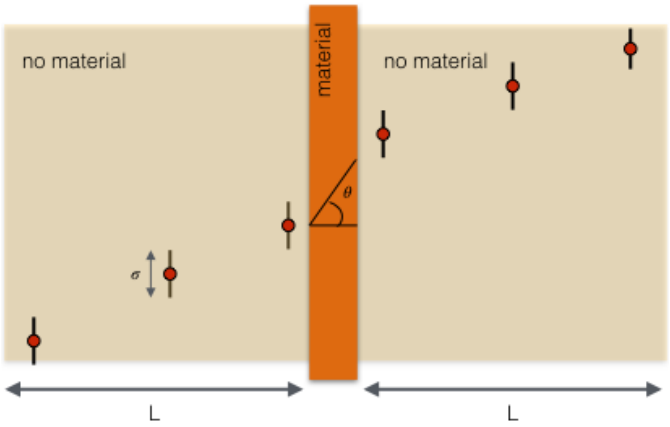


Figure A.1: Cartoon illustrating the correlation between multiple scattering and spatial hit resolution on the momentum resolution.

$$\left(\frac{\sigma_{p_T}}{p_T}\right)^2_{no\ mat} = p_T^2 \frac{\sigma}{0.3B} \frac{1}{4L^2} \sqrt{\frac{720}{2N+5}}. \tag{A.1}$$

The other limiting case is when the angle θ_0 in the scattering layer is large compared to σ/L. In this case the measurement of the momentum in the first and second half of the tracker are not correlated and the resulting momentum resolution is simply $1/\sqrt{2}$ of the momentum resolution computed with one half of the tracker:

$$\left(\frac{\sigma_{p_T}}{p_T}\right)^2_{mat} = \frac{1}{\sqrt{2}} p_T^2 \frac{\sigma}{0.3B} \frac{1}{L^2} \sqrt{\frac{720}{N+5}} \tag{A.2}$$

The ratio between the two limiting cases is:

$$\frac{\left(\frac{\sigma_{p_T}}{p_T}\right)^2_{mat}}{\left(\frac{\sigma_{p_T}}{p_T}\right)^2_{no\ mat}} = \frac{4}{\sqrt{2}}\sqrt{\frac{2N+5}{N+5}} \sim 4 \tag{A.3}$$

In the general case, when the scattering angle θ is comparable to σ/L the momentum resolution is given by an expression that contains the correlation between the scattering angle and the position resolution.

Appendix B

Explicit computation of bias in track parameters due to non-uniform flaws in the trajectory

We make an explicit calculation of the bias in track parameters due to non-uniform flaws in the trajectory in a simple case of a multiplicative bias $C \simeq 1 + A$ to the track curvature k. We assume that on the transverse plane a track can be parametrised as a parabola. In the first half of the trajectory we have:

$$y = d_{xy} + \phi x + \frac{1}{2\rho} x^2, \tag{B.1}$$

where $\rho = 1/0.3Bk$ is the radius of curvature of the track in terms of the assumed uniform magnetic field B. In the second half of the trajectory:

$$y = d_{xy} + \phi x + \frac{C}{2\rho} x^2 \tag{B.2}$$

Here x and y are the coordinates along and perpendicular to the track and the parameters d_{xy}, ϕ and ρ are found by minimising the relative χ^2. Following the procedure of the minimisation of the χ^2, the best fit track parameters $\vec{z} = \left(d'_{xy}, \phi', \rho'\right)$ and the covariance matrix of the fitted parameters V is:

$$\vec{z} = (P^T (V^{yy})^{-1} P)^{-1} P^T V^{yy\,-1} \vec{y} \tag{B.3}$$

$$V = (P^T (V^{yy})^{-1} P)^{-1} \tag{B.4}$$

where V^{yy} is the covariance matrix of the measurements y_i: $V^{yy}_{ij} =< y_i y_j >$.

For the sake of this calculation, we will assume a detector of five uniformly spaced layers. Using the general formula in Eq. 4.6 we can compute V^{yy} for this simplified case:

$$V^5_{yy} = \begin{pmatrix} \sigma^2 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{16}\delta^2 k^2 L^2 + \sigma^2 & \frac{1}{8}\delta^2 k^2 L^2 & \frac{3}{16}\delta^2 k^2 L^2 & \frac{1}{4}\delta^2 k^2 L^2 \\ 0 & \frac{1}{8}\delta^2 k^2 L^2 & \frac{5}{16}\delta^2 k^2 L^2 + \sigma^2 & \frac{1}{2}\delta^2 k^2 L^2 & \frac{11}{16}\delta^2 k^2 L^2 \\ 0 & \frac{3}{16}\delta^2 k^2 L^2 & \frac{1}{2}\delta^2 k^2 L^2 & \frac{7}{8}\delta^2 k^2 L^2 + \sigma^2 & \frac{5}{4}\delta^2 k^2 L^2 \\ 0 & \frac{1}{4}\delta^2 k^2 L^2 & \frac{11}{16}\delta^2 k^2 L^2 & \frac{5}{4}\delta^2 k^2 L^2 & \frac{15}{8}\delta^2 k^2 L^2 + \sigma^2 \end{pmatrix} \tag{B.5}$$

In the notation we are using, the scattering angle θ_0 is expressed as δk to make the dependence on the curvature k explicit. σ is the resolution of the single hit and L is the length of the track in the tranverse

plane. In this simplification, we will assume that the hit resolution and material are the same for each layer.

In case of a uniform bias C on the curvature k, i.e. choosing eq. B.2 for the whole trajectory, the values of the track parameters at the minimum of the χ^2 are:

$$\begin{aligned} d_{xy}^{unif} &= 0 \\ \phi^{unif} &= 0 \\ \rho^{unif} &= \frac{1}{0.3BCk} = \frac{\rho}{C}, \end{aligned} \tag{B.6}$$

then $k^{unif} = C \cdot k$. If the bias is only inserted in the second part of the trajectory, the values of the track parameters that minimise the χ^2 are instead:

$$\begin{aligned} d'_{xy} &= \frac{\delta^2 C 0.3Bk^3L^4 - \delta^2 0.3Bk^3L^4}{160\left(\delta^2k^2L^2 + 80\sigma^2\right)} + \frac{3\left(\delta^2 C 0.3Bk^3L^4 - \delta^2 0.3Bk^3L^4\right)}{560\left(3\delta^2k^2L^2 + 560\sigma^2\right)} - \frac{9(C-1)0.3BkL^2}{1120} \\ \phi' &= -\frac{3\left(\delta^2 C 0.3Bk^3L^3 - \delta^2 0.3Bk^3L^3\right)}{40\left(\delta^2k^2L^2 + 80\sigma^2\right)} + \frac{\delta^2 0.3Bk^3L^3 - \delta^2 C 0.3Bk^3L^3}{7\left(3\delta^2k^2L^2 + 560\sigma^2\right)} - \frac{13}{560}(C-1)0.3BkL \\ \frac{1}{2\rho'} &= \frac{\delta^2 C 0.3Bk^3L^2 - \delta^2 0.3Bk^3L^2}{7\left(3\delta^2k^2L^2 + 560\sigma^2\right)} + \frac{1}{28}(15C-1)0.3Bk \end{aligned} \tag{B.7}$$

We observe from Eq B.7 that all the track parameters are embedded in a complex expression dependent on k. In the limit $C \to 1$ all the parameters are unbiased, while in the limit $\delta \to 0$, i.e. no scattering on the layers, Eq B.7 tends to Eq. B.6 and the regime of uniform bias is restored. Therefore, the root of the extra functions of k in Eq B.7 with respect to Eq. B.6 is due to the off-diagonal terms of V^{yy} originating from multiple scattering. Studying the formula for the bias in k in Eq B.7 we can make we can make some further observations:

$$\begin{aligned} \frac{k'}{k} &= 2\delta^2L^2k^2\frac{C-1}{7\left(3\delta^2k^2L^2 + 560\sigma^2\right)} + \frac{1}{14}(15C-1) \\ &= 1 + 2\delta^2L^2k^2\frac{C-1}{7\left(3\delta^2k^2L^2 + 560\sigma^2\right)} + \frac{1}{14}(15C-1) - \frac{14}{14} \\ &= \underbrace{1 + 5\cdot 10^{-4}\cdot\delta^2\frac{L^2}{\sigma^2}k^2\frac{C-1}{\left(0.05\cdot\delta^2\frac{L^2}{\sigma^2}k^2 + 1\right)}}_{\text{non-uniform bias}} + \underbrace{\frac{15}{14}(C-1)}_{\text{uniform bias}}, \end{aligned} \tag{B.8}$$

where we have decoupled the contributions originating from uniform and non uniform biases along the trajectory. In this expression the pure numbers are linked to choice of the geometry encoded in P, while δ, L and σ are properties of the detector. To make a realistic case, in CMS we can choose $\delta \simeq 0.013$ GeV$\sqrt{0.1}$, $L \simeq 1$m and $\sigma \simeq 50\mu$m, so that:

$$\delta\frac{L}{\sigma} \simeq 80 \text{ GeV}.$$

Plugging these numbers in Eq. B.8, we can make a plot of the ratio between the non-uniform and uniform contributions as a function of k, in Figure B.1. From Eq. B.8, it is clear that the multiplicative bias C cancels in the ratio. Therefore, the relative importance of the two terms is only given by the properties of the detector under study. In the case that we have considered, the value of the non-uniform contribution with respect to the simple uniform is about 2% at 0.2 GeV^{-1}, corresponding to 5 GeV and it grows to about 1% at 1 GeV^{-1}.

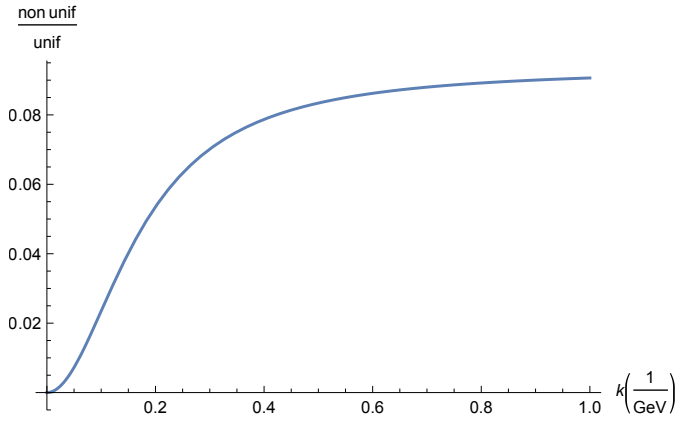


Figure B.1: Plot of the ratio between the non-uniform and uniform contributions as a function of k.

Doing a partial fraction decomposition of Eq. B.8, the bias on the curvature can be reduced to the form:

$$\sum_l \frac{g_{il}}{1 + d_l^2 k^2} \tag{B.9}$$

with the index l running over the number of layers minus 3 and i is the index of the track parameter and runs from 1 to 3. We highlight that the terms d_l do not have the index i since they are in common to all the parameters.

Let us now compute the 3x3 covariance matrix of the track parameters using Eq. B.4:

$$\begin{aligned}
V_{11} &= \frac{13\delta^2k^2L^2}{9800} - \frac{\delta^4k^4L^4}{800\left(\delta^2k^2L^2 + 80\sigma^2\right)} - \frac{9\delta^4k^4L^4}{39200\left(3\delta^2k^2L^2 + 560\sigma^2\right)} + \frac{31\sigma^2}{35} \\
V_{12} = V_{21} &= -\frac{167\delta^2k^2L}{9800} + \frac{3\delta^4k^4L^3}{200\left(\delta^2k^2L^2 + 80\sigma^2\right)} + \frac{3\delta^4k^4L^3}{490\left(3\delta^2k^2L^2 + 560\sigma^2\delta^2\right)} - \frac{108\sigma^2}{35L} \\
V_{22} &= \frac{6457\delta^2k^2}{4900} - \frac{9\delta^4k^4L^2}{50\left(\delta^2k^2L^2 + 80\sigma^2\right)} - \frac{8\delta^4k^4L^2}{49\left(3\delta^2k^2L^2 + 560\sigma^2\delta^2\right)} + \frac{696\sigma^2}{35L^2} \\
V_{23} = V_{32} &= -\frac{19\delta^2k^2}{49L} + \frac{8\delta^4k^4L}{49\left(3\delta^2k^2L^2 + 560\sigma^2\right)} - \frac{128\sigma^2}{7L^3} \\
V_{33} &= \frac{68\delta^2k^2}{49L^2} - \frac{8\delta^4k^4}{49\left(3\delta^2k^2L^2 + 560\sigma^2\right)} + \frac{128\sigma^2}{7L^4} \\
V_{13} = V_{31} &= \frac{\delta^2k^2}{490} - \frac{3\delta^4k^4L^2}{490\left(3\delta^2k^2L^2 + 560\sigma^2\right)} + \frac{16\sigma^2}{7L^2}
\end{aligned} \tag{B.10}$$

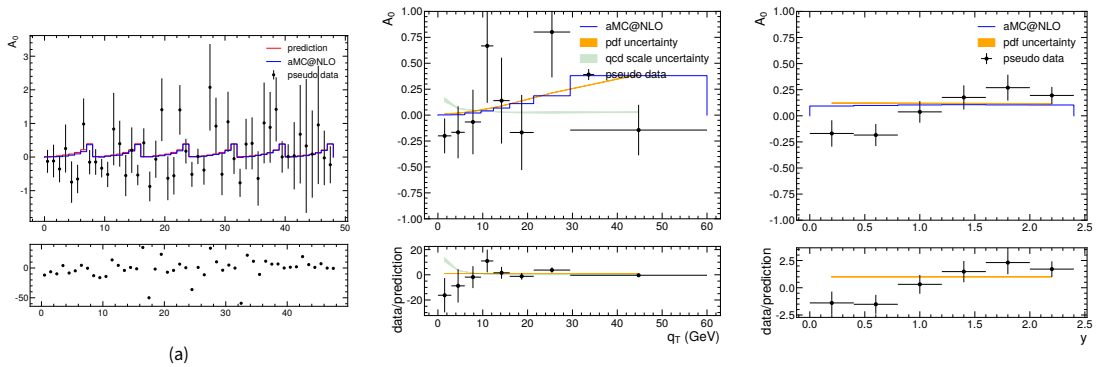
With calculation analogous to Eq. B.8, we observe that the elements of the covariance matrix can be reduced to the form:

$$C_{ij} = a_{ij}k^2 + c_{ij} + \sum_l \frac{f_{ijl}k^2}{1 + d_l^2 k^2}, \tag{B.11}$$

with the index l running over the number of layers minus 3 and i, j are the index of the matrix. We observe here that the d_l are the same numbers as in B.9.

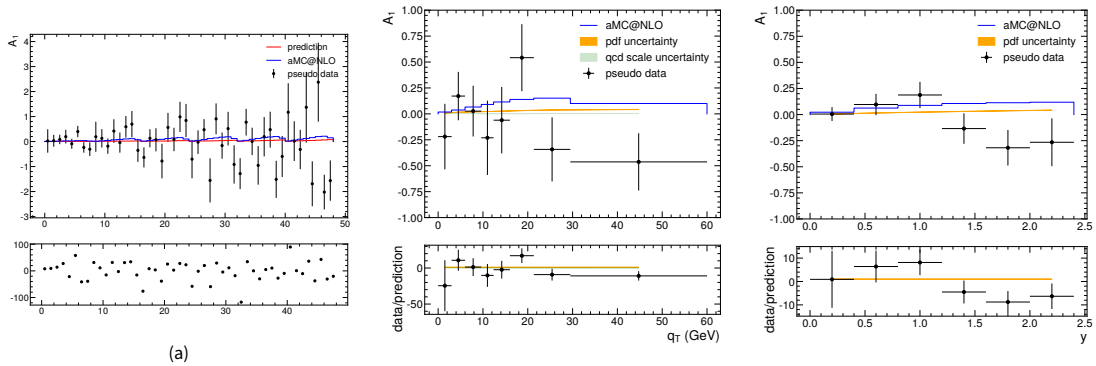
Appendix C

Additional material for differential cross sections



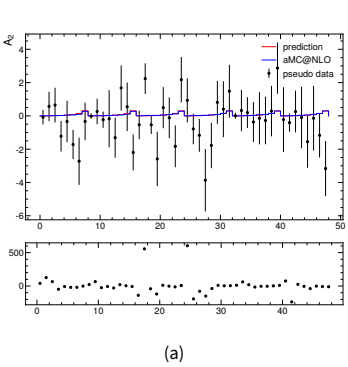
(a) (b) (c)

Figure C.1: .

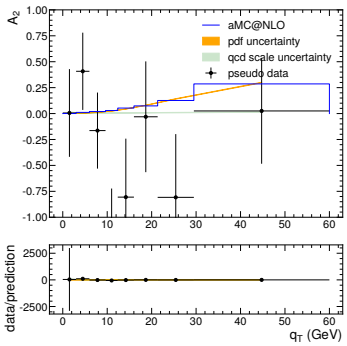


(a) (b) (c)

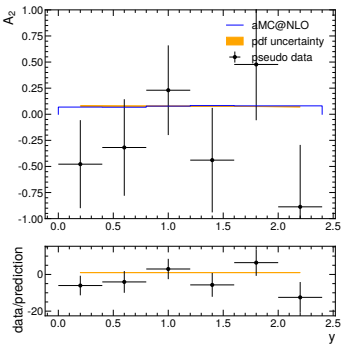
Figure C.2: .



(a)

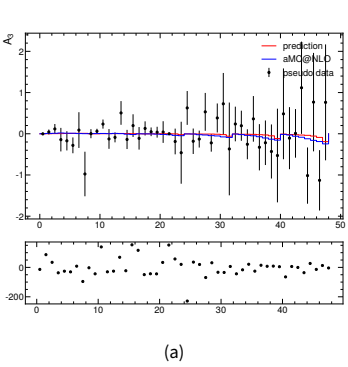


(b)

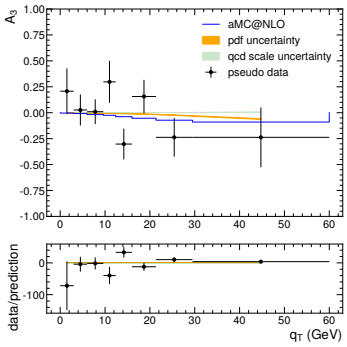


(c)

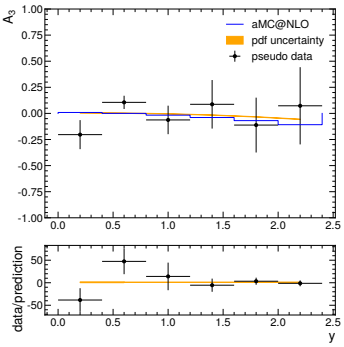
Figure C.3: .



(a)

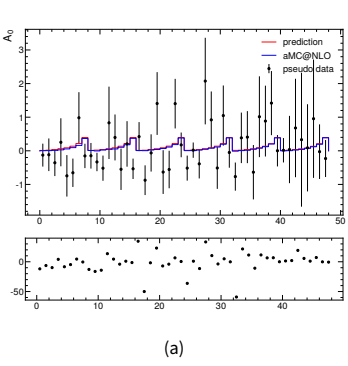


(b)

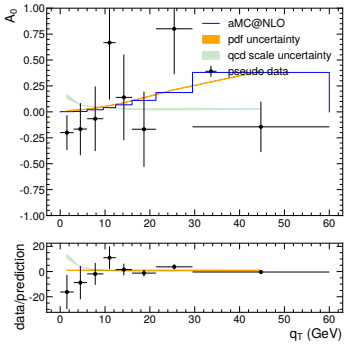


(c)

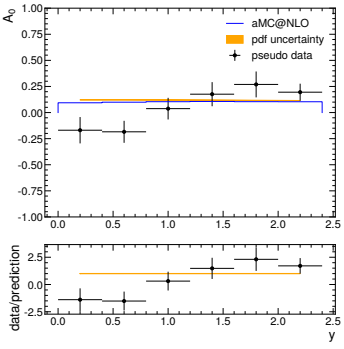
Figure C.4: .



(a)

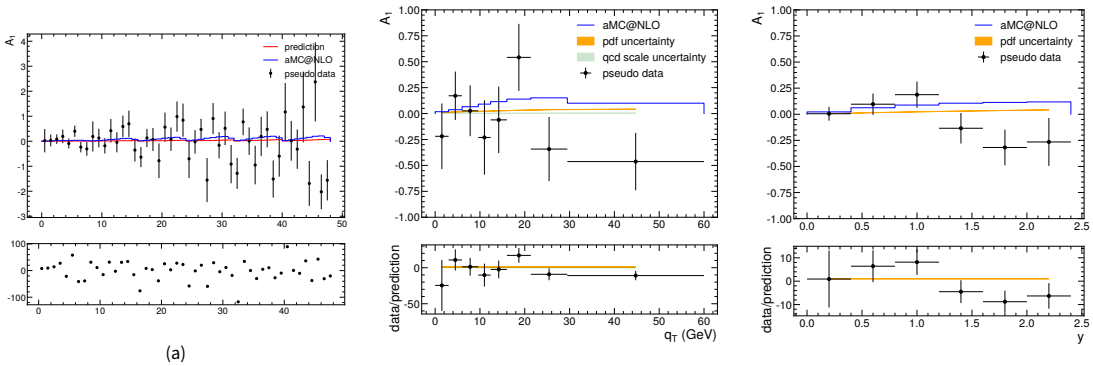


(b)



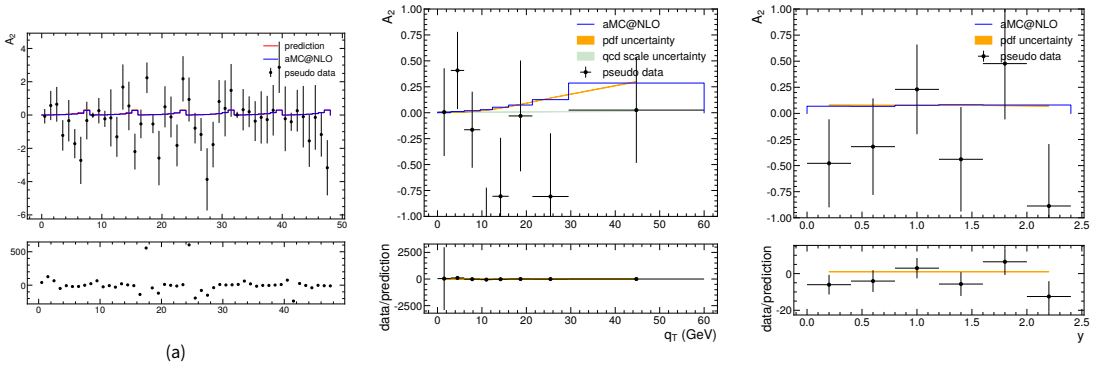
(c)

Figure C.5: .



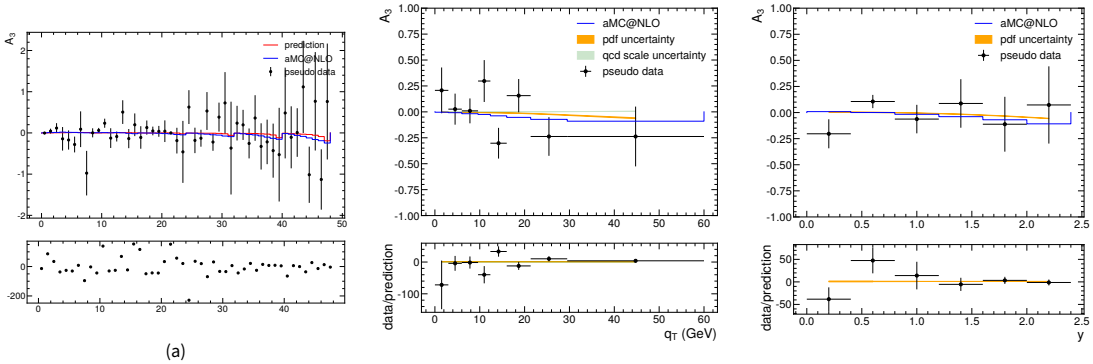
(a) (b) (c)

Figure C.6: .



(a) (b) (c)

Figure C.7: .



(a) (b) (c)

Figure C.8: .

Acknowledgements

Bibliography

[1] S. L. Glashow, "Partial-symmetries of weak interactions", Nuclear Physics **22**, 579–588 (1961) http://dx.doi.org/10.1016/0029-5582(61)90469-2.

[2] A. Salam and J. Ward, "Electromagnetic and weak interactions", Physics Letters **13**, 168–171 (1964) http://dx.doi.org/10.1016/0031-9163(64)90711-5.

[3] S. Weinberg, "A Model of Leptons", Phys. Rev. Lett. **19**, 1264–1266 (1967) 10.1103/PhysRevLett.19.1264.

[4] S. Weinberg, "Effects of a Neutral Intermediate Boson in Semileptonic Processes", Phys. Rev. D **5**, 1412–1417 (1972) 10.1103/PhysRevD.5.1412.

[5] P. Higgs, "Broken symmetries, massless particles and gauge fields", Physics Letters **12**, 132–133 (1964) http://dx.doi.org/10.1016/0031-9163(64)91136-9.

[6] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", Phys. Rev. Lett. **13**, 508–509 (1964) 10.1103/PhysRevLett.13.508.

[7] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", Phys. Rev. Lett. **13**, 321–323 (1964) 10.1103/PhysRevLett.13.321.

[8] M. Martinez et al., "Precision tests of the electroweak interaction at the Z pole", Rev. Mod. Phys. **71**, 575–629 (1999) 10.1103/RevModPhys.71.575.

[9] U. Collaboration, "Experimental observation of isolated large transverse energy electrons with associated missing energy at s=540 GeV", Physics Letters B **122**, 103–116 (1983) https://doi.org/10.1016/0370-2693(83)91177-2.

[10] UA2, "Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN anti-p p Collider", Phys. Lett. B **122**, 476–485 (1983) 10.1016/0370-2693(83)91605-2.

[11] UA1, "Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c^2 at the CERN SPS Collider", Phys. Lett. B **126**, 398–410 (1983) 10.1016/0370-2693(83)90188-0.

[12] UA2, "Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN $\bar{p}p$ Collider", Phys. Lett. B **129**, 130–140 (1983) 10.1016/0370-2693(83)90744-X.

[13] "Precision electroweak measurements on the Z resonance", Physics Reports **427**, 257–454 (2006) https://doi.org/10.1016/j.physrep.2005.12.006.

[14] "Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab", Phys. Rev. Lett. **74**, 2626–2631 (1995) 10.1103/PhysRevLett.74.2626.

[15] "Observation of the Top Quark", Phys. Rev. Lett. **74**, 2632–2637 (1995) 10.1103/PhysRevLett.74.2632.

[16] T. C. 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.

[17] T. A. 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.

[18] CMS, "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.

[19] J. Haller et al., "Update of the global electroweak fit and constraints on two-Higgs-doublet models", The European Physical Journal C **78**, 675 (2018) 10.1140/epjc/s10052-018-6131-3.

[20] ATLAS, "Measurement of the top quark mass in the $t\bar{t} \to$ lepton+jets channel from $\sqrt{s} = 8$ TeV ATLAS data and combination with previous results", Eur. Phys. J. C **79**, 290 (2019) 10.1140/epjc/s10052-019-6757-9.

[21] "Measurement of the top quark mass using proton-proton data at $\sqrt{(s)} = 7$ and 8 TeV", Phys. Rev. D **93**, 072004 (2016) 10.1103/PhysRevD.93.072004.

[22] ATLAS, "Measurement of the W-boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector", Eur. Phys. J. C **78**, [Erratum: Eur.Phys.J.C 78, 898 (2018)], 110 (2018) 10.1140/epjc/s10052-017-5475-4.

[23] CDF, D0, "Tevatron Run II combination of the effective leptonic electroweak mixing angle", Phys. Rev. D **97**, 112007 (2018) 10.1103/PhysRevD.97.112007.

[24] "Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments", Phys. Rev. Lett. **114**, 191803 (2015) 10.1103/PhysRevLett.114.191803.

[25] M. Davier et al., "Reevaluation of the hadronic vacuum polarisation contributions to the Standard Model predictions of the muon $g-2$ and $\alpha(m_Z^2)$ using newest hadronic cross-section data", Eur. Phys. J. C **77**, 827 (2017) 10.1140/epjc/s10052-017-5161-6.

[26] A. e. a. Abada, "FCC Physics Opportunities", The European Physical Journal C **79**, 474 (2019).

[27] P. Janot and S. Jadach, "Improved Bhabha cross section at LEP and the number of light neutrino species", Phys. Lett. B **803**, 135319 (2020) 10.1016/j.physletb.2020.135319.

[28] G. Voutsinas et al., "Beam-beam effects on the luminosity measurement at LEP and the number of light neutrino species", Phys. Lett. B **800**, 135068 (2020) 10.1016/j.physletb.2019.135068.

[29] 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.

[30] DELPHI, "Measurement of the Mass and Width of the W Boson in e^+e^- Collisions at $\sqrt{s}$ = 161-GeV - 209-GeV", Eur. Phys. J. C **55**, 1–38 (2008) 10.1140/epjc/s10052-008-0585-7.

[31] T. L. Collaboration, "Measurement of the mass and the width of the W boson at LEP", The European Physical Journal C - Particles and Fields **45**, 569–587 (2006).

[32] "Precise Measurement of the W-Boson Mass with the CDF II Detector", Phys. Rev. Lett. **108**, 151803 (2012) 10.1103/PhysRevLett.108.151803.

[33] "Measurement of the W Boson Mass with the D0 Detector", Phys. Rev. Lett. **108**, 151804 (2012) 10.1103/PhysRevLett.108.151804.

[34] LHCb, "Measurement of the W boson mass", (2021), arXiv:2109.01113 [hep-ex].

[35] L. A. Harland-Lang et al., "Parton distributions in the LHC era: MMHT 2014 PDFs", The European Physical Journal C **75**, 204 (2015).

[36] R. D. Ball et al., "Parton distributions for the LHC run II", Journal of High Energy Physics **2015**, 40 (2015).

[37] R. D. Ball et al., "Parton distributions from high-precision collider data", The European Physical Journal C **77**, 663 (2017).

[38] T.-J. Hou et al., "New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC", Phys. Rev. D **103**, 014013 (2021) 10.1103/PhysRevD.103.014013.

[39] 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.

[40] 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.

[41] T. C. collaboration, "Measurement of the transverse momentum spectra of weak vector bosons produced in proton-proton collisions at $$ \sqrt{s}=8 $$TeV", Journal of High Energy Physics **2017**, 96 (2017).

[42] 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.

[43] 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.

[44] S. Catani et al., "Universality of non-leading logarithmic contributions in transverse-momentum distributions", Nucl. Phys. B **596**, 299–312 (2001) https://doi.org/10.1016/S0550-3213(00)00617-9.

[45] W. Bizoń et al., "The transverse momentum spectrum of weak gauge bosons at N^{3}LL + NNLO", The European Physical Journal C **79** (2019) 10.1140/epjc/s10052-019-7324-0.

[46] V. Bertacchi, "Towards a simultaneous measurement of W boson mass and production properties with the CMS detector", Presented 23 Jul 2021 (2021), https://cds.cern.ch/record/2776894.

[47] E. Manca et al., "About the rapidity and helicity distributions of the W bosons produced at LHC", Journal of High Energy Physics **2017**, 130 (2017).

[48] Z. Bern et al., "Left-handed W bosons at the LHC", Phys. Rev. D **84**, 034008 (2011) 10.1103/PhysRevD.84.034008.

[49] 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.

[50] T. A. Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", Journal of Instrumentation **3**, S08003 (2008), http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[51] T. C. Collaboration, "The CMS experiment at the CERN LHC", Journal of Instrumentation **3**, S08004 (2008), http://stacks.iop.org/1748-0221/3/i=08/a=S08004.

[52] T. L. Collaboration, "The LHCb Detector at the LHC", Journal of Instrumentation **3**, S08005 (2008), http://stacks.iop.org/1748-0221/3/i=08/a=S08005.

[53] T. A. Collaboration, "The ALICE experiment at the CERN LHC", Journal of Instrumentation **3**, S08002 (2008), http://stacks.iop.org/1748-0221/3/i=08/a=S08002.

[54] M. Bachtis and S. Dasu, "Heavy Neutral Particle Decays to Tau Pairs in Proton Collisions at $\sqrt{s}$ = 7TeV with CMS at the CERN Large Hadron Collider", PhD thesis (CERN, 2012), http://cds.cern.ch/record/1462018.

[55] B. Mangano and L. Foa, "The CMS Tracker: contributions to hardware integration, software development and first data taking", PhD thesis (Scuola Normale Superiore, SNS, 2013), http://cds.cern.ch/record/1558329.

[56] T. T. G. of the CMS Collaboration, *The CMS Phase-1 Pixel Detector Upgrade*, 2020, arXiv:2012.14304 [physics.ins-det].

[57] W. Adam et al., *Track Reconstruction in the CMS tracker*, tech. rep. CMS-NOTE-2006-041 (CERN, Geneva, 2006), https://cds.cern.ch/record/934067.

[58] V. Blobel, "A new fast track-fit algorithm based on broken lines", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **566**, TIME 2005, 14–17 (2006) https://doi.org/10.1016/j.nima.2006.05.156.

[59] V. Blobel et al., "Fast alignment of a complex tracking detector using advanced track models", Computer Physics Communications **182**, Computer Physics Communications Special Edition for Conference on Computational Physics Trondheim, Norway, June 23-26, 2010, 1760–1763 (2011) https://doi.org/10.1016/j.cpc.2011.03.017.

[60] A. Rizzi, "The Evolution of Analysis Models for HL-LHC", EPJ Web of Conferences **245**, 11001 (2020) 10.1051/epjconf/202024511001.

[61] I. Antcheva et al., "ROOT — A C++ framework for petabyte data storage, statistical analysis and visualization", Computer Physics Communications **180**, 2499–2512 (2009).

[62] E. Manca and E. Guiraud, *Using RDataFrame, ROOT's declarative analysis tool, in a CMS physics study*, https://indico.cern.ch/event/849610/.

[63] O. e. a. Aberle, *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*, CERN Yellow Reports: Monographs (CERN, Geneva, 2020), 10.23731/CYRM-2020-0010.

[64] G. Petrucciani et al., "Mini-AOD: A New Analysis Data Format for CMS", **664**, 072052 (2015).

[65] Rizzi, Andrea et al., "A further reduction in CMS event data for analysis: the NANOAOD format", EPJ Web Conf. **214**, 06021 (2019).

[66] C. R. H. et al., "Array programming with NumPy", Nature **585**, 357–362 (2020) 10.1038/s41586-020-2649-2.

[67] CMS Collaboration, *W-like measurement of the Z boson mass using dimuon events collected in pp collisions at* $\sqrt{s} = 7$ TeV, tech. rep. (CERN, Geneva, 2016), https://cds.cern.ch/record/2139655.

[68] S. A. et al., "Geant4 simulation toolkit", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **506**, 250–303 (2003) https://doi.org/10.1016/S0168-9002(03)01368-8.

[69] A. et al., "Geant4 developments and applications", IEEE Transactions on Nuclear Science **53**, 270–278 (2006) 10.1109/TNS.2006.869826.

[70] J. A. et al., "Recent developments in Geant4", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **835**, 186–225 (2016) https://doi.org/10.1016/j.nima.2016.06.125.

[71] J. Marion and B. Zimmerman, "Multiple scattering of charged particles", Nuclear Instruments and Methods **51**, 93–101 (1967) http://dx.doi.org/10.1016/0029-554X(67)90367-9.

[72] R. Gluckstern, "Uncertainties in track momentum and direction, due to multiple scattering and measurement errors", Nuclear Instruments and Methods **24**, 381–389 (1963) https://doi.org/10.1016/0029-554X(63)90347-1.

[73] Particle Data Group, "Review of Particle Physics", Chin. Phys. **C38**, 090001 (2014) 10.1088/1674-1137/38/9/090001.

[74] T. C. collaboration, "Alignment of the CMS tracker with LHC and cosmic ray data", Journal of Instrumentation **9**, P06009–P06009 (2014) 10.1088/1748-0221/9/06/p06009.

[75] W. Hulsbergen, "The global covariance matrix of tracks fitted with a Kalman filter and an application in detector alignment", Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **600**, 471–477 (2009) https://doi.org/10.1016/j.nima.2008.11.094.

[76] E. Manca, "Validation of the muon momentum resolution in view of the W mass measurement with the CMS experiment", PhD thesis (INFN, Pisa, 2016).

[77] Particle Data Group, "Review of Particle Physics", PTEP **2020**, 083C01 (2020) 10.1093/ptep/ptaa104.

[78] "Dimuon spectrum 2016", (2016), https://cds.cern.ch/record/2212114.

[79] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", Comput. Phys. Commun. **191**, 159–177 (2015).

[80] 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).

[81] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, 2004.

[82] P. F. Monni et al., "MiNNLOPS: a new method to match NNLO QCD to parton showers", Journal of High Energy Physics **2020** (2020) 10.1007/jhep05(2020)143.

[83] P. F. Monni et al., "MiNNLO$_{\text{PS}}$: optimizing $2 \to 1$ hadronic processes", Eur. Phys. J. C **80**, 1075 (2020) 10.1140/epjc/s10052-020-08658-5.

[84] J. Bradbury et al., *JAX: composable transformations of Python+NumPy programs*, version 0.2.5, 2018, http://github.com/google/jax.

[85] 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).

[86] E. Mirkes and J. Ohnemus, "W and Z polarization effects in hadronic collisions", Phys. Rev. D **50**, 5692–5703 (1994).

[87] C. S. Lam and W.-K. Tung, "Systematic approach to inclusive lepton pair production in hadronic collisions", Phys. Rev. D **18**, 2447–2461 (1978).

[88] R. Gauld et al., "Precise predictions for the angular coefficients in Z-boson production at the LHC", J. High Energy Phys. **2017** (2017).