SCUOLA
NORMALE
SUPERIORE

Classe di Scienze

Corso di Perfezionamento (Ph.D.) in
Fisica

XXXVII ciclo

High-precision measurement of the W boson mass with the CMS experiment

Settore Scientifico Disciplinare **PHYS-01/A**

Candidato
dr. Davide Bruschini

Relatore	Supervisione interna
Prof. Lorenzo Bianchini	Prof. Michael Joseph Morello
	Ph.D. coordinator
	Prof. Vittorio Giovannetti

Anno accademico 2024–2025

Contents

Chapter 1

Introduction

The measurement of the mass of the W boson (m_W) provides one of the most stringent tests of the ability of the Standard Model of Particle Physics to describe the interactions between the fundamental constituents of matter. This thesis describes the m_W measurement performed by the CMS experiment [1] at the Large Hadron Collider (LHC), which has been communicated in September 2024: the result of this measurement is in agreement both with the Standard Model and with the world average of all the other direct measurements [2], which has an uncertainty of 13.3 MeV. The CMS measurement has an uncertainty of 9.9 MeV, close to the relative precision of 10^{-4} from the Standard Model indirect prediction. This uncertainty is compatible with the most precise measurement from a single experiment, as performed by CDF [3], which is, however, in strong disagreement with the Standard Model. The CMS measurement is the result of an effort that started more than ten years ago, making it long awaited, especially after the latest CDF result.

Measurements of the W boson mass at hadron colliders are performed by analyzing the leptonic decays of W bosons, in which a charged lepton (muon or electron) and a neutrino, which escapes detection, are produced. A proxy for the presence of a neutrino and its kinematic properties is the missing transverse momentum p_T^{miss}, determined from the kinematic imbalance of all the particles measured in the event. The measurement then relies on the precise reconstruction of the kinematic properties of the charged lepton in the decay, which is also often combined with p_T^{miss} to determine the transverse mass m_T. The redundancy of observables sensitive to m_W, albeit correlated, can be exploited to reduce the impact of theoretical or experimental uncertainties. Most measurements rely on the reconstructed distributions of m_T and of the lepton transverse momentum p_T^ℓ used in template fits, in which p_T^ℓ and m_T templates are built from Monte Carlo simulation for discrete values of m_W, which is then determined by matching the simulated templates with the measured p_T^ℓ and m_T distribution. The precision of this approach is limited by the theoretical modelling of the production mechanism of the W boson, which affects particularly the p_T^ℓ distribution, and the experimental precision with which they are calibrated, affecting both m_T and p_T^ℓ.

The CMS measurement employs an innovative approach to determine the theoretical modelling of the production mechanism of the W boson. It improves upon the approach traditionally followed at hadron colliders which exploits the similarity between W and Z production to tune the W boson production mechanism, as the kinematic properties of the Z bosons can be precisely measured, and then theoretical predictions are used to port the $Z-$based constraints to the modelling of the W boson production. This approach is limited by the assumptions made in this porting procedure, which might not take into account non-trivial differences between the production of Z and W bosons. Indeed, different assumptions on the theoretical uncertainties lead to different predictions on the total uncertainty in the extraction of m_W. The measurement performed by CMS is based on precise template fits of the two-dimensional $p_T - \eta$ distribution of muons from the W boson decays, where η is related to the direction of flight the muon in the laboratory, since it has been proved that the theoretical model can be constrained together with m_W without using external data from Z decays. The most precise theoretical calculations currently available are employed in this analysis, together with a model for their systematic uncertainties based on the introduction of nuisance parameters that can be determined simultaneously with m_W, without relying on external measurements. For this to work, and to have a total competitive uncertainty, CMS relied on the largest dataset ever employed in any m_W analysis, for a total number of 1.2×10^8 analyzed data events and $O(4 \times 10^9)$ simulated events, requiring a complete overhaul of the software usually employed in most analyses in order to make this measurement feasible.

From the experimental point of view, this measurement relies on the precise reconstruction and identification of muons. The lepton transverse momenta need to be determined with a precision better than 10^{-4} to reach the precision required: this required implementing a new software to refit the trajectory of tracks, from which the transverse momentum is determined, leading to a significant reduction in the bias in the reconstructed muon momenta, and the derivation of dedicated corrections to account for the remaining biases. These corrections rely on a physics-driven model that takes into account different physical sources of bias, where accuracy and robustness have been thoroughly validated. The analysis also relies on the precise knowledge of the muon reconstruction and identification efficiencies, which are the probabilities that a muon is successfully reconstructed passing a set of quality criteria, as a function of $p_T - \eta$, to avoid differences in the templates that could introduce artificial biases in the extraction of m_W. Muon reconstruction and identification efficiencies have been measured with an unprecedented level of granularity, similar to that of the analysis, to be able to extract as much information as possible from the $p_T - \eta$ plane. Biases in the standard muon reconstruction in CMS, which can have a non-negligible impact on the determination of the muon reconstruction and identification, were identified and fixed also in the standard CMS muon reconstruction algorithm. The procedure to extract corrections on muon reconstruction and identification efficiencies, based on the tag-and-probe procedure using $Z \to \mu\mu$ events, was thoroughly validated to identify sources of bias of physical nature, particularly those related to the production mechanism of the Z boson and its differences with respect to the W boson, and to the recomputation of the quantities used to define the quality criteria required in this analysis with respect to what is done in CMS by default. This is the aspect of the analysis in which I was mainly involved and is extensively documented in this thesis. The uncertainties on the muon reconstruction and identification efficiencies result in a contribution of 3.0 MeV to the total m_W uncertainty. This work was carried out with the financial support of the European Research Council, under the Horizon 2020 research and innovation programme of the European Union (Grant agreement N. 101001205) [4].

Besides providing a fundamental input to the last measurement of the mass of the W boson, during my Ph.D. I contributed to the CMS Tracking Physics Object Group (POG), by performing detailed studies of tracking performances using Run 3 data collected by CMS: this work is essential for understanding the level of agreement between data and simulation, supporting the tuning of the different algorithms that enter in the reconstruction of the tracks, debugging of tracker conditions (alignment, bad components, etc.), and commissioning of the track reconstruction software as it evolves, and to establish the quality of CMS data-taking and the prompt identification of problems in the detector or the reconstruction.

This thesis is structured as follows: chapter 2 provides a general overview of the Standard Model of Particle Physics, summarizes the electroweak precision tests, and describes how the measurement of m_W can be considered a stringent test of the Standard Model. Chapter 3 describes the Large Hadron Collider and the experimental apparatus of CMS. Chapter 4 describes how *tracking*, i.e. the reconstruction of the trajectories (tracks) of charges particles produced in proton-proton collisions, is performed in CMS, with some details about the differences for data collected in Run 2, relevant for the measurement of m_W, and data collected in Run 3, exploring some of the latest upgrades to the tracking software. This chapter also contains the results of my work performed in collaboration with the Tracking POG, described in particular in section 4.2. Chapter 5 describes the reconstruction and identification of muons and their kinematic properties in CMS, with particular focus on the identification criteria to select high purity muons employed by CMS analyses in general, and for the m_W measurement in particular. The work I performed to validate the tag-and-probe procedure used to extract efficiency corrections is described in section 5.2. Chapter 6 provides an historical overview of the discovery of the W boson and the first measurements of its mass, providing a detailed description of the most precise measurements at lepton and hadron colliders, with a brief outlook on the current projections on the results achievable at FCC. Chapter 7 describes in detail the m_W measurement by CMS. Section 7.4 complements the report on my contributions to the m_W measurement started in section 5.2, summarizing the efficiency corrections and their validation using $Z \rightarrow \mu\mu$ events. Chapter 8 explores possible directions and improvements for future measurements of m_W.

Chapter 2

The W boson in the Standard Model

This chapter describes the Standard Model of Particle Physics, our current best theory of the interaction between the fundamental constituents of matter. The electroweak precision observables and their importance in testing the Standard Model are discussed, together with a summary of the experimental tests and their agreement with the predictions from the Standard Model. Particular care is dedicated to the Standard Model prediction of the mass of the W boson and the comparison with the direct measurements, in the context of possible extensions of the Standard Model.

2.1 The Standard Model

The Standard Model of Particle Physics is a quantum field theory that describes the interactions between the fundamental constituents of matter, which can be divided into two main groups: fermions, with half-integer spin, and bosons, with integer spin. The Standard Model can describe three different forces: the strong force, which binds particles together inside nucleons (protons and neutrons) and is responsible for the formation of atomic nuclei, the electromagnetic force, which governs interactions between charged particles and is responsible for the formation of atoms and molecules, and the weak force, responsible for the decay of many known particles, both fundamental and composite, as well as the interaction of neutrinos.

Table 2.1 provides a summary of key properties of fermions. The elementary fermions that interact via the strong force are known as *quarks*. These are grouped into three doublets (or generations), with each doublet consisting of an up-type and a down-type quark. Up-type quarks carry a charge of $+\frac{2}{3}e$, while down-type quarks have a charge of $-\frac{1}{3}e$, where e is the magnitude of the charge of the electron. The doublets differ in the masses of their up- and down-type quarks, as shown in the table. The *anti-quarks*, which are the antiparticles of quarks, follow the same structure but have inverted quantum numbers. According to the CPT theorem [5], anti-quarks have the same masses as their corresponding quarks.

Fermions that do not interact via the strong force are called *leptons*. Like quarks, leptons are organized into three doublets. Each doublet contains an up-type neutral fermion, called a *neutrino* (denoted by ν), and a down-type charged fermion with a charge of $-e$. While neutrinos were initially thought to be massless, evidence from neutrino oscillation experiments [6] suggests that at least some neutrinos have mass, with the most stringent limit from direct measurements of < 0.45 eV [7] on the mass of the electron neutrino. The antimatter counterparts of leptons follow the same structure but with inverted quantum numbers. According to the CPT theorem, they have the same mass and lifetime as the corresponding particles. The table displays the mass and some quantum numbers of the particles, including charge, isospin ($\mathtt{I}$), spin-parity ($\mathtt{J}^{\mathrm{P}}$), $z-$component of the isospin ($\mathtt{I}_{\mathrm{z}}$), strangeness for quarks, charge and lifetime for leptons.

The Standard Model (SM) is a non-abelian gauge theory. This means that the SM Lagrangian is invariant under local transformations, i.e. those that depend on space-time coordinates, as dictated by its gauge symmetry group, and the generators of this group are not commutative. Each gauge symmetry requires

Quarks

Name	Mass	Charge	I(J^P)	I$_z$	Strangeness
u	2.16 ± 0.07 MeV	$+\frac{2}{3}$e	$\frac{1}{2}(\frac{1}{2}^+)$	$+\frac{1}{2}$	0
d	4.70 ± 0.07 MeV	$-\frac{1}{3}$e	$\frac{1}{2}(\frac{1}{2}^+)$	$-\frac{1}{2}$	0
s	93.5 ± 0.8 MeV	$-\frac{1}{3}$e	$0(\frac{1}{2}^+)$	0	-1
c	1.2730 ± 0.0046 GeV	$+\frac{2}{3}$e	$0(\frac{1}{2}^+)$	0	0
b	4.183 ± 0.007 GeV	$-\frac{1}{3}$e	$0(\frac{1}{2}^+)$	0	0
t	172.57 ± 0.29 GeV	$+\frac{2}{3}$e	$0(\frac{1}{2}^+)$	0	0

Leptons

Name	Mass	Charge	τ
e	$0.51099895000 \pm 0.00000000015$ MeV	-e	$> 6.6 \times 10^{28}$ y
μ	$105.6583755 \pm 0.0000023$ MeV	-e	$(2.1969811 \pm 0.0000022) \times 10^{-6}$ s
τ	(1776.93 ± 0.09) MeV	-e	$(290.3 \pm 0.5) \times 10^{-15}$ s

Table 2.1: Table summarizing some of the properties of the Standard Model fermions. The masses for u, d, s are the so-called "current quark masses". The t-quark mass is extracted from direct measurements, following the approach in [8]. From [9].

the introduction of a specific number of gauge fields, corresponding to the number of generators in the symmetry group. These gauge fields are physically interpreted as force carriers, mediating interactions between matter fields, and are known as *gauge bosons*. The Standard Model symmetry group is:

$$SU(3)_C \bigotimes SU(2)_L \bigotimes U(1)_Y \tag{2.1}$$

The strong force is described by the $SU(3)_C$ symmetry group, the electromagnetic force by the $U(1)_{em}$ group (different from the $U(1)_Y$ group in equation 2.1, but with the same algebraic properties), and the weak force by the $SU(2)_L$ group. The symmetry induced by the $SU(2)_L$ group is actually broken: the electromagnetic and weak are unified in one theory, called electroweak theory, with $SU(2)_L \bigotimes U(1)_Y$ as its symmetry group.

In the Standard Model the interactions are described in terms of quantum field theory, which uses the Lagrangian of a system to derive the equations of motion of the fields.

Gauge Bosons

Name	Mass	Charge	Lifetime/Γ	
γ (photon)	$< 1 \times 10^{-18}$ eV	0	Stable	I(J^{PC})=0,1(1^{--})
g (gluon)	$< O(1)$ MeV[*]	0	Stable	I(J^P) = 0(1^-)
W	80.3692 ± 0.0133 GeV	$\pm e$	2.085 ± 0.042 GeV	J=1
Z	91.1880 ± 0.0020 GeV	0	2.4955 ± 0.0023 GeV	J=1

Higgs Boson

Name	Mass	Charge	Lifetime/Γ	
H^0	125.20 ± 0.11 GeV	0	$3.7^{+1.9}_{-1.4}$ MeV[**]	J=0

Table 2.2: Table summarizing some of the properties of the bosons. The upper limit quoted in [*] is the one derived from high-energy experiments as described in [10]. The value in [**] assumes equal on-shell and off-shell effective couplings. From [9].

In the context of the path integral formalism, the scattering amplitude from initial state $|\phi_I\rangle$ at time $t = 0$ to final state $|\phi_F\rangle$ at time $t = T$ is obtained as:

$$\langle\phi_F|e^{iHT}|\phi_I\rangle = \int \mathcal{D}\phi \exp\left(i\frac{S}{\hbar}\right) \tag{2.2}$$

where H is the Hamiltonian of the system, $\mathcal{D}\phi$ is the integration measure for all possible "paths" the system can take at each infinitesimal step and S is the action of the system:

$$S[\phi] = \int \mathrm{d}^4x \mathcal{L}(\phi, \delta\phi) \tag{2.3}$$

which is the integral of the Lagrangian density $\mathcal{L}$ over the space-time volume for a generic quantum field ϕ and its derivatives $\delta\phi$. The Lagrangian density can be divided as:

$$\mathcal{L} = \mathcal{L}_0 + \mathcal{L}_{int}, \tag{2.4}$$

where $\mathcal{L}_0$ is the Lagrangian density for a free quantum field, and $\mathcal{L}_{int}$ is the part of the Lagrangian density that describes the interactions between particles.

Using the formulas in 2.2 and 2.3 is analytically unfeasible for the Standard Model. What is usually done is to expand the calculation of the scattering amplitude as a perturbative series in the interaction part:

$$\begin{aligned}
\langle\phi_F|e^{iHT}|\phi_I\rangle &= \int \mathcal{D}\phi \exp\left(i\int \mathrm{d}^4x(\mathcal{L}_0 + \mathcal{L}_{int})\right) \\
&= \int \mathcal{D}\phi \exp\left(i\int \mathrm{d}^4x\mathcal{L}_0\right) \exp\left(i\int \mathrm{d}^4x\mathcal{L}_{int}\right) \\
&= \int \mathcal{D}\phi \exp\left(i\int \mathrm{d}^4x\mathcal{L}_0\right) \left(1 + g\mathcal{M}_{FI}^{(1)} + g^2\mathcal{M}_{FI}^{(2)} + ...\right)
\end{aligned} \tag{2.5}$$

where the amplitude is broken down into terms proportional to the first power of the coupling constant g, the second power of g, etc. If $g \ll 1$, the lowest order terms are the most important, and the higher orders enter in the calculation when an increase in precision is desired, while if $g \approx 1$ all the terms of the series have the same weight, and the perturbative approach no longer holds. While performing higher-order calculations, unphysical infinities appear due to the presence of "loops". These infinities are dealt with using the so-called *renormalization techniques*. The Standard Model has been shown to be a mathematically renormalizable theory [11, 12]. When only the lowest order in the theoretical calculation is used, the computation is said to be at *leading order*, also referred to as tree-level calculation. As contributions with higher orders of precision are included, it is said that the calculation is at *next-to-leading-order* (NLO), *next-to-next-to-leading-order* (NNLO), and so on. As explained in the following, the Standard Model describes multiple interactions, characterized by different coupling constants: when performing theoretical calculations at a given order one must keep all contributions from the different interactions with similar weight.

Defining $\phi(x^\mu)$ the field of a fermion, the Lagrangian of the free fermion is:

$$\mathcal{L} = i\overline{\phi(x^\mu)}\gamma^\mu\partial_\mu\phi(x^\mu) - m\overline{\phi(x^\mu)}\phi(x^\mu) \tag{2.6}$$

where γ^μ are the Dirac matrices:

$$\gamma^0 = \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -I \end{bmatrix}, \qquad \gamma^i = \begin{bmatrix} \mathbf{0} & \sigma^i \\ -\sigma^i & \mathbf{0} \end{bmatrix}, \tag{2.7}$$

where I is the 2×2 identity matrix, **0** is the 2×2 null matrix, and σ_i are the Pauli matrices:

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \qquad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.8}$$

The auxiliary γ^5 matrix is added as it separates nominal left and right chiral representations in electroweak theory:

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \tag{2.9}$$

Gauge symmetries are introduced by modifying the field of the fermion:

$$\begin{aligned} \phi'(x^\mu) &= e^{i\alpha_a T^a}\phi(x^\mu) \\ \overline{\phi'(x^\mu)} &= \overline{\phi(x^\mu)}e^{-i\alpha_a T^a} \end{aligned} \tag{2.10}$$

with α_a the parameter of the transformation and T^a the generators of the transformation. To conserve gauge symmetry, the Lagrangian must remain invariant under transformations of the symmetry group. What has been described so far implements global gauge symmetries. However, the Standard Model exhibits invariance under *local* gauge symmetries, where the transformation parameters α_a are arbitrary functions of the space-time coordinate x^μ. To maintain the invariance of the Lagrangian under these local gauge symmetries, it becomes necessary to modify the partial derivatives with respect to x^μ in the following way:

$$\partial^\mu \rightarrow D^\mu = \partial^\mu + igT^a W_a^\mu(x^\mu) \tag{2.11}$$

where g is the coupling constant for the particular gauge symmetry that is applied to the Lagrangian, T^a are the generators of the transformation, $W_a^\mu(x^\mu)$ are the fields corresponding to gauge bosons, introduced to maintain the gauge symmetry. The fields corresponding to the gauge bosons must transform according to the gauge symmetry, and the Einstein notation is used to sum over $T^a W_a^\mu(x^\mu)$. The sum runs over the total number of gauge fields related to the symmetry group.

Electromagnetism, as already mentioned, is described by the $U(1)_{em}$ symmetry group. This acts on the fields in the following way:

$$\phi'(x^\mu) = e^{iq\alpha(x^\mu)}\phi(x^\mu) \tag{2.12}$$

where q is the electric charge. One new gauge field, $A^\mu(x^\mu)$, is introduced, and it transforms as:

$$A'^\mu = A^\mu + \partial^\mu\alpha(x^\mu) \tag{2.13}$$

The kinetic energy of this new Gauge boson is:

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \text{ with } F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{2.14}$$

to be added to the Lagrangian of the fermion.

In weak theory, fermions are paired in doublets:

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, \begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \begin{pmatrix} t \\ b' \end{pmatrix}_L \tag{2.15}$$

where d', s' and b' are linear combinations of d, s and b. The coefficients of this linear combination are the elements of the Cabibbo-Kobayashi-Maskawa matrix V_{CKM}[13, 14, 15], commonly expressed using the Chau-Keung parameterization [16]:

$$\begin{aligned} V_{CKM} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{bmatrix} \begin{bmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{bmatrix} \begin{bmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{bmatrix} \end{aligned} \tag{2.16}$$

where $s_{ij} = \sin\theta_{ij}$, $c_{ij} = \cos\theta_{ij}$, and δ is the phase responsible for all the phenomena in flavor-changing processes in the SM that are not invariant under combined charge-conjugation and parity transformations (CP violation). The angles θ_{ij} can be chosen to lie in the first quadrant, so s_{ij}, $c_{ij} \geq 0$.

The absolute values of the V_{CKM} elements, obtained from fits to data from several experiments [9] are:

$$|V_{CKM}| = \begin{bmatrix} 0.97435 \pm 0.00016 & 0.22501 \pm 0.00068 & 0.003732^{+0.000090}_{-0.000085} \\ 0.22487 \pm 0.00068 & 0.97349 \pm 0.00016 & 0.04183^{+0.00079}_{-0.00069} \\ 0.00858^{+0.00019}_{-0.00017} & 0.04111^{+0.00077}_{-0.00068} & 0.999118^{+0.000029}_{-0.000034} \end{bmatrix} \tag{2.17}$$

from which the values in eq. 2.16 are determined to be:

$$\begin{aligned} \sin\theta_{12} &= 0.22501 \pm 0.00068 & \sin\theta_{13} &= 0.003732^{+0.000090}_{-0.000085} \\ \sin\theta_{23} &= 0.04183^{+0.00079}_{-0.00069} & \delta &= 1.147 \pm 0.026. \end{aligned}$$

Particles within each doublet share the same weak isospin $\mathrm{I_W}$, which is the quantity conserved by the $SU(2)_L$ symmetry. This conservation allows particles to change type only within their respective doublets. Particles within the same doublet only differ, as far as weak interactions are concerned, by the $z-$component of the weak isospin: particles in the first row have $\mathrm{I}_z^{\mathrm{W}} = \frac{1}{2}$, while those in the second row have $\mathrm{I}_z^{\mathrm{W}} = -\frac{1}{2}$. Importantly, this arrangement applies exclusively to particles with *left chirality*, or *left-handed particles*, hence the L subscript in the $SU(2)_L$ symmetry group. In contrast, particles with *right chirality*, or *right-handed particles*, are represented as singlets:

$$\nu_e^R, e_R^-, \nu_\mu^R, \mu_R^-, \nu_\tau^R, \tau_R^-, u_R, d'_R, c_R, s'_R, t_R, b'_R. \tag{2.18}$$

Under the transformations generated by the $SU(2)_L$ group, the fields transform in the following way:

$$\psi'(x^\mu) = \begin{pmatrix} u'_q(x^\mu) \\ d'_q(x^\mu) \end{pmatrix} = e^{i\sum_{i=1}^{3} \alpha_i(x^\mu)T^i(x^\mu)}\psi(x^\mu) \tag{2.19}$$

where $T^a(x^\mu)$ are the three generators of the $SU(2)_L$ group. This results in the introduction of three new gauge fields $W_a^\mu(x^\mu)$, with kinematic energy term:

$$\mathcal{L}_{\text{kin}} = -\frac{1}{4}F^{a,\mu\nu}F^a_{\mu\nu}, \quad F^{a,\mu\nu} = \partial^\mu W^{a,\nu} - \partial^\nu W^{a,\mu} - g\epsilon^{abc}W^{b\mu}W^{c\nu} \tag{2.20}$$

where ϵ^{abc} is the Levi-Civita tensor and the Einstein notation is used. The kinematic term for $SU(2)_L$ is similar to that of $U(1)$, but with an additional term arising from the different commutation relations between these groups. While $U(1)_{em}$ (and also $U(1)_Y$) is an Abelian group (its gauge symmetry generators commute), $SU(2)_L$ is non-Abelian. This non-Abelian nature results in self-interactions among the $SU(2)_L$ vector bosons. The field and generators are described using the basis T^+, T^-, and T^3, where $T^\pm = \frac{T^1 \pm iT^2}{2}$. This basis gives rise to two charged particles, $W^\pm = \frac{W^1 \mp iW^2}{\sqrt{2}}$. The choice of this basis is significant because the $T^\pm$ operators facilitate movement up and down within the $SU(2)_L$ doublets, while T^3 leaves the particle unchanged.

As mentioned above, electromagnetism and weak theory are unified in electroweak theory [17, 18, 19], described by the symmetry group $SU(2)_L \times U(1)_Y$. The $U(1)_Y$ symmetry has the same structure as $U(1)_{em}$, but the associated conserved charge is the *hypercharge* Y. This is the same for particles in the same doublet, while they still have different electric charge. In this case the Gauge fields B^μ, associated to $U(1)_Y$, and W^3, associated to $SU(2)_L$, are mixed:

$$\begin{pmatrix} Z^\mu \\ A^\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_3^\mu \\ B^\mu \end{pmatrix} \tag{2.21}$$

This unification preserves the charged currents (interactions involving charged vector bosons) while the neutral current also involves particles with right chirality and anti-particles with left chirality. Consequently, the couplings involve both chiralities, expressed in terms of c_L and c_R at the coupling vertices. This interaction can also be described using vectorial and axial components. Left-handed particles exhibit pure $V - A$ interactions, while right-handed particles have pure $V + A$ interactions. As a result, the interaction vertices are decomposed into these components, characterized by an axial component a_f and a vectorial component v_f for each fermion type. Figure 2.1 shows the couplings of fermions to vector bosons described by $SU(2)_L \times U(1)_Y$. The vertices also show the coupling constants and the vertex structure in each case. The relation between the electric charge Q, hypercharge Y and third component of the weak isospin T_3 is $Q = T_3 + Y/2$, which together with equation 2.21 gives the decomposition of the coupling into v_f and a_f, as summarized in table 2.3. For leptons, a consequence of the above discussion is that the interactions described by $SU(2)_L \times U(1)_Y$ involving neutral bosons do not change the flavour of the fermions involved (this is actually true also for quarks), while those involving charged bosons change fermion flavour only within the same doublet for leptons. This is called *lepton flavour conservation*. Furthermore, all of the lepton doublets couple in the same way as of $SU(2)_L \times U(1)_Y$, and the interactions are identical for all the doublets. This is called *lepton flavour universality*.

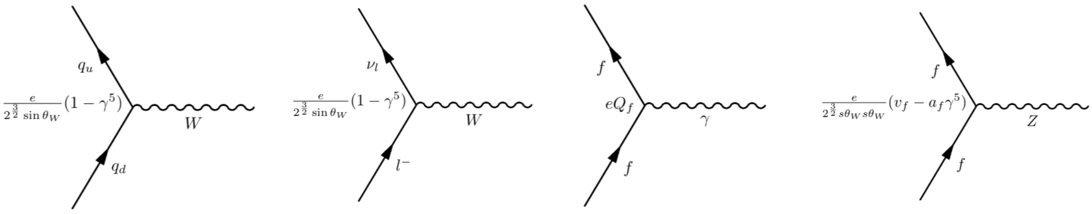


Figure 2.1: Couplings of fermions to electroweak vector bosons.

	u, c, t	d, s, b	ν_e, ν_μ, ν_τ	e^-, μ^-, τ^-
v_f	$\frac{1}{2} - \frac{4}{3}\sin^2\theta_W$	$-\frac{1}{2} + \frac{2}{3}\sin^2\theta_W$	$\frac{1}{2}$	$-\frac{1}{2} + 2\sin^2\theta_W$
a_f	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$

Table 2.3: Table summarizing v_f and a_f for each of the couplings between fermions and Z.

Quantum Chromodynamics (QCD) describes the strong force, employing the $SU(3)_C$ symmetry group. This group features eight generators, each corresponding to a gluon field G^a. QCD introduces the concept of *colour charge*, of which there are three types. Quarks transform as triplets under $SU(3)_C$, therefore can exist in three different colour states. Antiquarks carry complementary anticolour charges. In contrast, leptons are $SU(3)_C$ singlets, meaning they do not participate in strong interactions and are colourless. Gluons, the force carriers of QCD, uniquely possess both a colour and an anticolour charge. This property enables gluons to interact not only with quarks but also with each other, resulting in self-interactions between gluons. The QCD Lagrangian, for a single quark, is:

$$\mathcal{L} = i\overline{\psi_i(x_i^\mu)}\left[\gamma_\mu (D^\mu)_{ij} - m\right]\psi_j(x_j^\mu) - \frac{1}{4}\mathcal{G}^a_{\mu\nu}\mathcal{G}_a^{\mu\nu} \tag{2.22}$$

where $(D_\mu)_{ij} = \partial_\mu \delta_{ij} - ig_S (T_a)_{ij} G^a_\mu$, with g_S the strong coupling constant, G^a_μ the gluon fields, and T_a are the $SU(3)_C$ generators in the fundamental representation. An explicit representation of these generators is given by $T_a = \lambda_a/2$, where the λ_a are the Gell-Mann matrices [20], and the sum over a runs on $N_C - 1 = 8$ gluon fields (Einstein notation is used, the sums over i and j run on the three colours of the quarks). The gluon field strength tensor $\mathcal{G}^a_{\mu\nu}$ is given by:

$$\mathcal{G}^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_S f^{abc} G^b_\mu G^c_\nu \tag{2.23}$$

where f^{abc} are the structure constants of $SU(3)_C$ [20]. From g_S $\alpha_S \equiv \frac{g_S^2}{4\pi}$ can be defined. Quantum chromodynamics exhibits two crucial properties that affect its phenomenology:

- *Colour confinement*. As two colour charges are pulled apart, the force between them remains constant, unlike the other forces, which instead diminish with distance. This unique property leads to an increase in energy as the separation grows. Eventually, this energy becomes sufficient to spontaneously create a new quark–antiquark pair. This process results in the formation of two separate *hadrons*, i.e. combinations of quarks and gluons that are colour singlets under $SU(3)_C$, from the original one, instead of having the initial colour charges separately.
- *Asymptotic freedom*. The strength of the interactions between quarks and gluons, which can be described by α_S, decreases as the energy scale of the interaction increases. This allows the quarks to interact directly at high energy. This is opposite to what happens for electromagnetism, where $\alpha = \frac{e^2}{4\pi}$ increases as the interaction energy increases.

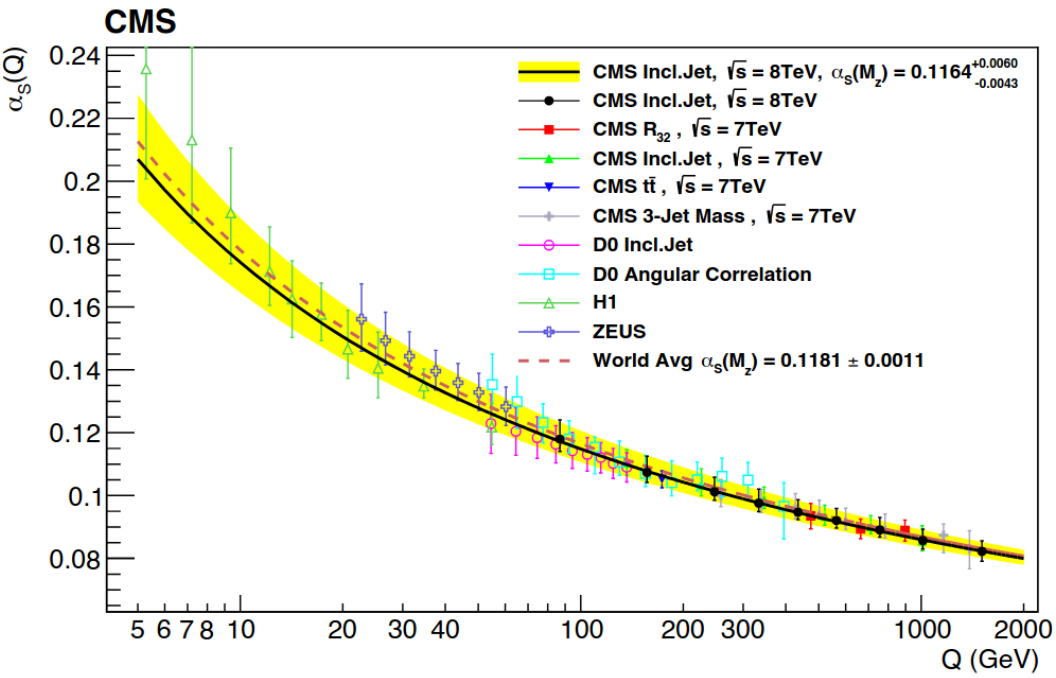


Figure 2.2: Strong coupling constant ad a function of the energy scale Q (from [21]).

As a consequence of these two properties, QCD can be treated as a perturbative theory only at very high energies, while non-perturbative QCD (such as lattice QCD [22]) is required to give accurate predictions at low energies. In particular, figure 2.2 shows the strong coupling constant as a function of the energy scale Q. In particular, it can be seen that α_s increases steadily moving towards lower energies, and it formally diverges at $\Lambda_{\mathrm{QCD}} \approx 100 - 300$ MeV. Colour confinement implies that quarks and gluons cannot be observed as freely propagating particles. When these are produced in particle collisions, such as proton-proton collisions, the *hadronization* process occurs, i.e. the production of hadrons as a result of the colour confinement, resulting in a shower of colour singlets. These showers, called *jets*, are those that are reconstructed in high-energy physics experiments.

The electroweak sector requires 3 massive vector bosons, two charged and one neutral ($W^{\pm}$ and Z), and one massless neutral boson (γ) to implement the observed currents. Moreover, also the observed quarks and charged leptons are massive. The presence of the mass terms in the Lagrangian breaks the $SU(3)_C \bigotimes SU(2)_L \bigotimes U(1)_Y$ symmetry. This is restored by adding two terms, $\mathcal{L}_H$ and $\mathcal{L}_{\text{Yukawa}}$, to the Standard Model Lagrangian, which require the introduction of the complex scalar field corresponding to the Higgs field. $\mathcal{L}_H$ describes the interaction of the Higgs boson field with the gauge fields and with itself:

$$\mathcal{L}_H = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi), \qquad \text{where} \quad V(\phi) = \mu^2\phi^\dagger\phi + \lambda\left(\phi^\dagger\phi\right)^2 \tag{2.24}$$

where $\phi = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix}$ is the Higgs field, complex doublet under $SU(2)_L$, resulting in four degrees of freedom. The Higgs potential has a degenerate minimum (assuming $\mu^2 < 0$ and $\lambda > 0$) for:

$$\phi^\dagger\phi = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2} \tag{2.25}$$

and ϕ assumes a non-zero vacuum expectation value (VEV):

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v\end{pmatrix} \tag{2.26}$$

This induces the spontaneous breaking of the $SU(2)_L \times U(1)_Y$ symmetry (Electroweak Symmetry Breaking, EWSB), but still keeping the $U(1)_{em}$ symmetry (conservation of electric charge). Three of the four degrees of freedom of the Higgs field give rise to three Goldstone bosons, corresponding to excitations in directions which do not change the Higgs potential. These three Goldstone bosons can be reabsorbed through an appropriate gauge transformation, known as *unitary gauge* transformation, so that the degrees of freedom corresponding to the Goldstone bosons are replaced by those related to the longitudinal polarization of three of the generators of $SU(2)_L \times U(1)_Y$, and the corresponding particles are now the massive bosons of the Standard Model. The remaining degree of freedom results is the physical Higgs boson, a scalar, uncharged, colourless boson. The Higgs potential is approximately quadratic for excitations corresponding to this degree of freedom.

The masses of the electroweak bosons and the Higgs boson are related to the VEV, charge of the electron, and $\sin^2 2\theta_W$ (at lowest order in perturbation theory) through the relation:

$$\begin{aligned} m_Z &= \frac{ev}{2\sin\theta_W\cos\theta_W}, & m_W^{\pm} &= \frac{ev}{2\sin\theta_W} = m_Z\cos\theta_W, \\ m_\gamma &= 0, & m_H &= \sqrt{2\lambda}v = 2|\mu| \end{aligned}$$

The Yukawa term $\mathcal{L}_{\text{Yukawa}}$ describes the interaction between fermions and the Higgs field:

$$\mathcal{L}_{\text{Yukawa}} = -y^d_{ij}\overline{Q}_{L_i}\phi d_{R_j} - iy^u_{ij}\overline{Q}_{L_i}\sigma_2\phi^* u_{R_j} - y^e_{ij}\overline{L}_{L_i}\phi e_{R_j} + h.c. \tag{2.27}$$

where $Q_{L_i} = \begin{pmatrix}u^i_L\\ d^i_L\end{pmatrix}$, $L_{L_i} = \begin{pmatrix}\nu^i\\ e^i\end{pmatrix}$, y^f_{ij} is the Yukawa coupling matrix for $f =$ up, down, or charged lepton fermions, and the indices i, j run on the 3 generations. After the electroweak symmetry breaking, the y^f_{ij} matrix can be diagonalized using the fermion mass eigenstates basis, and all the fermions acquire a mass

$m_i^f = \frac{1}{\sqrt{2}} y_i^f v$ (y_i^f being the eigenvalues of y_{ij}^f). To reproduce the phenomenology of the Standard Model, after EWSB $Q_{L_i} = \begin{pmatrix} u_L^i \\ d_L^i \end{pmatrix}$ should be replaced with $Q'_{L_i} = \begin{pmatrix} u_L^i \\ d_L'^i \end{pmatrix}$ and d_R with d'_R, where $d_{L,R}$ are the same linear combinations as in equation 2.15.

The electroweak Lagrangian is fully reproduced after EWSB, and an additional term describing the interaction between fermions, bosons and Higgs boson is added to the total Lagrangian:

$$\mathcal{L}_{add} = -\sum_f \overline{\psi}_f \frac{m_f H}{v} \psi_f + \frac{1}{2} \frac{e^2}{\cos^2\theta_W} v W_\mu^- W^{+\mu} H + \frac{1}{4} \frac{e^2}{\cos^2\theta_W} W_\mu^- W^{+\mu} HH, \tag{2.28}$$

where H is the field corresponding to the Higgs boson[1]. Table 2.2 summarizes some of the properties of bosons [9].

2.2 Electroweak precision tests

In the context of the Standard Model it is possible to calculate with very high precision several different quantities, which are compared to the most precise experimental measurements. Differences between the predicted and measured values of these observables could be hints of Physics Beyond the Standard Model. As of this thesis, no evidence of Physics Beyond the Standard Model has been found from these tests. However, the Standard Model appears to be incomplete, as there are phenomena that it cannot currently explain, some of which will be described in the following.

It has been already mentioned that in the Standard Model the neutrinos are predicted to be massless, but the phenomenon of neutrino oscillations has been first observed by the Super-Kamiokande experiment [6], and this requires that at least some of the neutrinos have mass. The KATRIN experiment has put an upper limit on the mass of the electron neutrino of < 0.45 eV [7], and measurements on neutrino oscillations provide values on the difference of the squared masses ΔM^2 between neutrino generations of $O(7\times10^{-5})$ eV^2 between electron neutrino ν_e and ν_μ neutrino and $|\Delta M^2_{32}| = O(10^{-3})$ eV^2 (only the absolute value is known in the case) [9]. If neutrinos acquire mass through the Higgs mechanism, they would have to couple very weakly with the Higgs boson, see equation 2.28, where it can be seen that the coupling is proportional to the mass (for fermions). This means that the Yukawa coupling the Higgs field should be very small. Another possibility is that neutrinos are actually Majorana particles, meaning that ν and $\overline{\nu}$ are the same particle. In this case, the masses are generated directly by the Majorana operator, without involving the EWSB, and this would also break the conservation of lepton flavour.

The Standard Model is not currently able to explain the existence of *dark matter* [23] and *dark energy* [24] and the way in which they interact. These two ingredients are crucial to explain several cosmological observations regarding the matter content of galaxies and the evolution of the universe. Ordinary matter, i.e. the one that has been observed by experiments and that the Standard Model is able to describe, is estimated to account for approximately 5% of the total matter content of the universe, *dark matter* is estimated to account for approximately 27% of the total matter content, with the remainder being *dark energy*. This means that the Standard Model is able to explain the existence and interactions of a relatively small fraction of the universe.

All of the cosmological observations show that the known universe is prevalently composed of matter, with only a small component of antimatter. One could expect that equal amounts of matter and antimatter were produced after the Big Bang, which should reflect in the same amount of matter and antimatter observed today. It was discussed in the previous chapter that the Standard Model, particularly the electroweak sector, is not symmetric with respect to combined charge-conjugation and parity transformations,

---

[1] After EWSB, the Higgs field becomes $\phi(x^\mu) = \begin{pmatrix} 0 \\ v + H(x^\mu) \end{pmatrix}$.

which could in principle lead to an asymmetry between matter and antimatter. However, the CP violation in the Standard Model does not appear to be sufficient to explain the actual amount of matter compared to antimatter [25], so that the Standard Model appears to be incomplete also in this respect.

Another open problem is the so called *naturalness* or *hierarchy problem* [26]. Using the relationships between the fundamental parameters predicted by the Standard Model, the vacuum expectation value v can be determined using the measured value of the Fermi constant G_F to be $v \approx 246$ GeV, also referred to as electroweak scale. The Standard Model appears to be complete (apart from the open problems already mentioned) at least for energy scales up to the M_{Planck} scale ($M_{\text{Planck}} \equiv \sqrt{\hbar c/G} \sim 10^{19}$ GeV, where G is the gravitational constant), the energy scale at which the effects of quantum gravity can no longer be ignored in other fundamental interactions [27]. However, taking into account radiative corrections, the mass of the Higgs boson m_H can be expressed as:

$$m_H^2 = (m_H^0)^2 + (\delta m_H)^2, \qquad (\delta m_H)^2 \sim \left(\sum_i m_i + \lambda\right) \Lambda^2, \tag{2.29}$$

where $m_H^0 = \sqrt{2\lambda}v$ is the Higgs mass at lowest order in perturbation theory, while δm_H are the corrections from higher-order interactions, i runs over the particles coupling to the Higgs and Λ is an arbitrary scale. If $\Lambda = O(M_{\text{Planck}})$ it can be seen that $(\delta M_H)^2 \gg (M_H)^2$: a very finely tuned cancellation must occur between m_H^0 and δm_H to keep the variations in the value of m_H below, for example, 1 TeV. Another solution could be the presence of new particles that appear at intermediate energies that modify the relationship in equation 2.29, potentially protecting the predicted value of m_H from radiative corrections.

All of these phenomena could lead to the existence of new particles that couple to the ones in the Standard Model. However, as already mentioned, all the comparisons between the theoretical predictions and the experimental measurements are in agreement. This could mean that the coupling between these new particles and the Standard Model could be very small, or that the masses of these new particles can be higher than those that can be produced, for example, at the LHC. Indirect evidence of these particles could still be provided if they entered in the radiative corrections of quantities that can be predicted and measured very precisely. An example of these tests are the electroweak precision tests.

The theoretical predictions for the electroweak precision observables (EWPO) are simultaneously fitted to several experimental measurements, in so-called *global electroweak fits* [9, 28, 29]. This can be performed for the following reasons:

- Test the internal consistency of the Standard Model. The experimental information for all the EWPO is included and the Standard Model parameters are fit simultaneously.
- Test BSM models. The Lagrangian of the Standard Model is extended to include terms that parameterize the BSM contributions, typically through the addition of effective field theory (EFT) operators, which modify the EWPO dependence from the parameters, and thus test the model with the measurements of all the EWPO. A similar approach was used before the discoveries of the top quark and the Higgs boson to predict their masses, assuming the Standard Model relations.
- Provide an *indirect determination* of a given EWPO. This is similar to the first approach, except that the experimental measurement of the EWPO under study is removed, so that a *prediction* that is completely independent can be derived in the context of the Standard Model.

The Standard Model has 19 free parameters: 9 are the Yukawa couplings of the charged fermions, 4 are the parameters of the CKM matrix, 1 is the QCD vacuum angle to account for a possible CP-violation in QCD [30], the remaining 5 describe the electroweak sector. At lowest order in perturbation theory, all of the EWPO can be predicted in terms of three quantities, chosen to be those measured with the highest precision: α, the fine structure constant, known to a relative precision of 1.6×10^{-10} [31], G_F, the Fermi constant, known to a relative precision of 5×10^{-7} [32], and the mass of the Z boson, m_Z, measured with a precision of 2 MeV (2×10^{-5} relative precision) in the scan of the Z resonance in e^+e^- collisions at LEP [33].

To compute higher-order corrections, additional inputs are required: the strong coupling constant α_S, the running of the fine-structure constant at the Z boson mass $\Delta\alpha$, the mass of the Higgs boson, and the masses of the fermions. In practice, only the top quark mass m_t plays a significant role, as it is significantly larger (at least two orders of magnitude) than the masses of other fermions, and the corrections increase with mass.

The latest global electroweak fit results obtained by the GFITTER group are contained in [34] and shown in figure 2.3. The fit incorporates a set of measured observables, including the Higgs boson mass, the top quark mass, the masses and widths of W and Z bosons, the weak mixing angle $\sin\theta_W$, the strong coupling constant α_S at the Z pole, and measurements of various asymmetries. Several theoretical observables have been computed at next-to-next-to-leading order (NNLO), and they include the full two-loop EW contributions for Z boson production [35, 36], so the theoretical uncertainties are now smaller than the experimental ones in all observables entering the EW fit. The experimental inputs in the fit include the updated version of the mass of the W boson m_W performed by ATLAS [37], by the LHCb collaboration [38], the measurement of the Z pole observables by LEP [33] corrected for beam-beam effects on the luminosity measurement [39], a combination of the most precise kinematic top quark mass from ATLAS and CMS [40, 41], a combination of several measurements of $\sin^2\theta^l_{eff}$ performed at LEP [33] and at hadron colliders [42, 43, 44, 45]. In addition to the electroweak precision data from LEP and SLD [33], other experimental inputs to the fit are the hadronic contribution to the electromagnetic coupling strength [46], m_H [47] and the masses of the b and c quarks [48].

Figure 2.3 shows the pulls, defined as the difference between the quantity measured experimentally and the fitted value from the global fit divided by the experimental uncertainty. The fit converges on a minimum χ^2 value of 13.8 for 15 dof, corresponding to a p-value of 0.55. The largest contribution to the χ^2 originates from the forward-backward asymmetry from b quarks, A_{FB}, which shows a deviation of 2.4σ from the value obtained in the global fit. The leptonic left-right asymmetry A_ℓ from SLD has a deviation of -2.1σ, while that derived from LEP has a deviation of 0.1σ. This is a result of the well-known discrepancy between the measurements of the LR asymmetry with polarized beams at SLAC and the b-quark forward-backward asymmetry at LEP [33]. The best-fit value of the strong coupling strength at the mass of the Z boson, $a_S(m_Z^2)$, is 0.1196 ± 0.0029, to be compared to the PDG value from 2022 of 0.1179 ± 0.0009 [48], not included in the fit, resulting in a pull value of 2.0σ. The pull of m_W also decreased with respect to the previous GFITTER result, that is, from -0.8σ to -0.5σ. The GFITTER also provides an indirect determination of m_W, which is 80.354 ± 0.007 GeV.

Another set of results is the one obtained by the Particle Data Group in 2024 [9]. Although most results are in agreement with the ones shown above, there are a few differences: the PDG fit includes inputs from low-energy parity-violation data and the muon magnetic moment, the set of input data, the implementation of radiative corrections, and the fitting tools used are different. The experimental inputs to the fit are: for m_t the combination of CDF and D0 measurements at Tevatron in [49], the Run 1 data from LHC mentioned above and the Run 2 measurements [50, 51, 52, 53, 54, 55], for m_H [56, 57], for Γ_H [58, 59], for m_W the measurement already mentioned from ATLAS [37] and the combination of measurements in [2], for Γ_W [60, 37, 61], $\mathcal{B} \to$hadrons [60, 62], for the weak charges of the electron [63], the proton [64], cesium [65, 66] and thallium [67, 68], for the weak mixing angle extracted from eDIS [69], $\nu_\mu(\bar{\nu}_\mu) - e$ scattering [70, 71, 72], the τ lifetime, and the μ anomalous magnetic moment [73], and the principal Z pole observables at LEP 1 [33]. These are shown in tables 2.4 and 2.5, together with the result from the global electroweak fit and, when available, the indirect prediction removing the experimental information on the observables shown.

The agreement is generally very good. The global electroweak fit describes the data well, with a very good χ^2/dof of $49.5/47$. The probability of a larger χ^2 is 37%, and only $g_\mu - 2$ is currently showing a larger (3.2σ) conflict. In addition, A_0^{LR} (SLD) from hadronic final states, $A_{FB}^{(0,b)}$ (LEP 1), Γ_W (ATLAS) and Q_W (Cs) deviate at the 2σ level. A_b extracted from $A_{FB}^{(0,b)}$ $A_b = 0.885 \pm 0.017$, is 2.9σ below the global

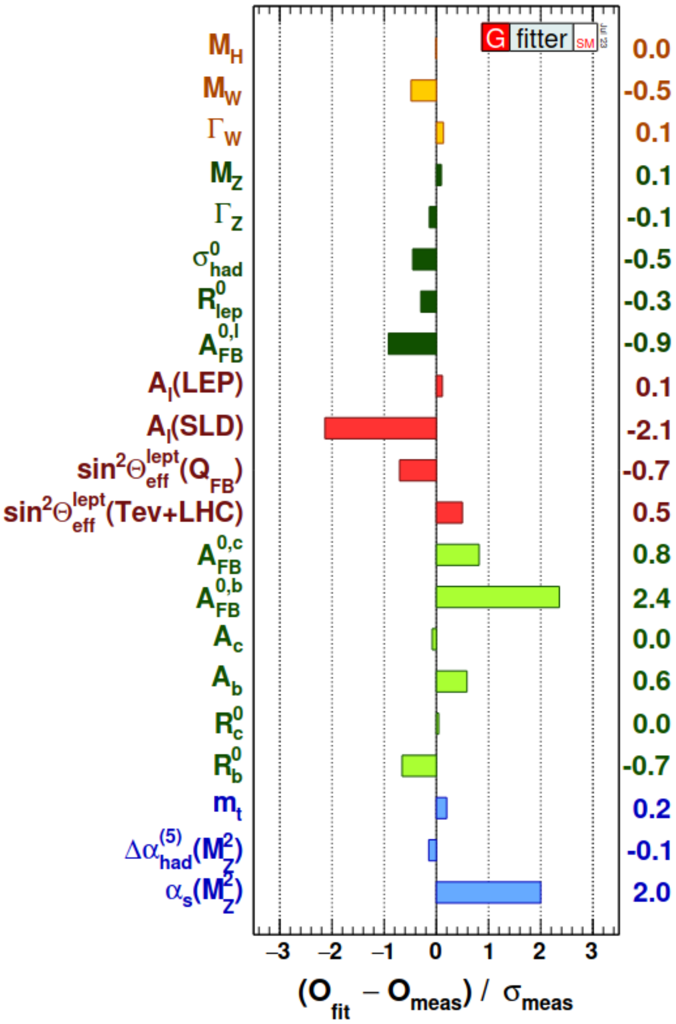


Figure 2.3: Pull values of the global electroweak fit described in [34], defined as deviations between measurements and predictions evaluated at the best-fit point, divided by the experimental uncertainties.

electroweak fit and also 1.4σ below $A_b = 0.923 \pm 0.020$ obtained from $A^{FB}_{LR}(b)$ at SLD. The left-right asymmetry, $A^0_{LR} = 0.15138 \pm 0.00216$ [76], from hadronic decays at SLD, differs by 2.3σ from the value from the global fit of 0.1463 ± 0.0003. The combined value of $A_\ell = 0.1513 \pm 0.0021$ from SLD (using lepton-family universality and including correlations) is also 2.4σ above the value from the global, but there is experimental agreement between this SLD value and the LEP 1 value, $A_\ell = 0.1481 \pm 0.0027$, obtained from a fit to $A^{(0,\ell)}_{FB}$, $A_e(P\mathcal{T})$, and $A_\tau(P_\tau)$.

2.3 The SM prediction for m_W

The mass of the W boson m_W depends on the other parameters of the Standard Model according to equation 2.30:

$$m_W^2\left(1-\frac{m_W^2}{m_Z^2}\right)=\frac{\pi\alpha(m_Z)}{\sqrt{2}G_F(1-\Delta r)}, \tag{2.30}$$

Quantity	Value	Electroweak Fit	Pull	Indirect Prediction
m_t [GeV]	172.61 ± 0.58	172.85 ± 0.55	-0.4	175.2 ± 1.8
m_H [GeV]	125.10 ± 0.09	125.10 ± 0.09	0.0	97^{+18}_{-16}
Γ_H [MeV]	3.5 ± 1.5	4.09 ± 0.05	-0.4	
m_W [GeV]	80.376 ± 0.033[60]	80.356 ± 0.005	0.6	80.353 ± 0.006
	80.355 ± 0.016[37]		-0.1	
	80.347 ± 0.033[38]		-0.3	
	80.372 ± 0.026[74]		0.6	
	$[80.432 \pm 0.016]$[3]			
Γ_W [GeV]	2.195 ± 0.083[60]	2.089 ± 0.001	1.3	
	2.198 ± 0.049[37]		2.2	
	2.059 ± 0.049[61]		-0.6	
$\mathcal{B}(W \to$hadrons$)$	0.6736 ± 0.0018	0.6751 ± 0.0001	-0.8	
$g_V^{\nu e}$	-0.040 ± 0.015	-0.0395 ± 0.0001	0.0	
$g_A^{\nu e}$	-0.507 ± 0.014	-0.5063	0.0	
$Q_W(e)$	-0.0403 ± 0.0053	-0.0469 ± 0.0002	1.3	
$Q_W(p)$	0.0719 ± 0.0045	0.0705 ± 0.0002	0.3	
$Q_W(\mathrm{Cs})$	-72.41 ± 0.42	-73.26 ± 0.01	2.0	
$Q_W(\mathrm{Th})$	-116.4 ± 3.6	-116.93 ± 0.01	0.1	
$\hat{s}_Z^2(\mathrm{eDIS})$	0.2299 ± 0.0043	0.23129 ± 0.00004	-0.3	
τ_τ [fs]	290.75 ± 0.36	288.59 ± 2.31	0.9	
$\frac{1}{2}\left(g_\mu - 2 - \frac{\alpha}{\pi}\right)$	$(4510.86 \pm 0.35) \times 10^{-9}$	$(4509.73 \pm 0.03) \times 10^{-9}$	3.2	

Table 2.4: Non-Z pole observables, compared with the global electroweak fit value (and the pull) and, when provided, the indirect prediction obtained removing the experimental information on the observable. The values of m_W and Γ_W from [37] reported here are obtained from the simultaneous fit to extract both the mass and width of the W boson. $\mathcal{B}(W \to$hadrons$)$ is a combination of [60, 62]. The world average results for $g_{V,A}^{\nu e}$ are dominated by the CHARM II results [70]. The τ_τ value is the τ lifetime world average computed by combining the direct measurements with values derived from the leptonic branching ratios [75]. From [9].

where

$$\Delta r = -\frac{3G_F m_t^2}{8\sqrt{2}\pi^2}\frac{\cos^2\theta_W}{\sin^2\theta_W} + \frac{11 G_F m_W^2}{24\sqrt{2}\pi^2}\log\frac{m_H^2}{m_W^2} + \Delta r_{\mathrm{rem}}. \tag{2.31}$$

It can be seen that at lowest order in perturbation theory, i.e. with $\Delta r = 0$, m_W is a function of only G_F, α and m_Z. Δr encapsulates the contribution of radiative corrections: the first term, the dominant one, is the contribution from the top quark mass m_t, the second one contains the dependence on the mass of the Higgs boson m_H, the third one encapsulates the additional corrections (e.g. the running of α, terms of order $O(\alpha\alpha_S), O(\alpha\alpha_S^2)$, etc.).

In the previous section it was discussed how the EWPO are used to test the internal consistency of the Standard Model, in particular the latest results from the GFITTER group and Particle Data Group were discussed. When the experimental information on m_W is not included in the electroweak fit, m_W is predicted to be:

$$\begin{aligned} m_W &= 80354 \pm 7 \text{ MeV [Gfitter]} \\ m_W &= 80353 \pm 6 \text{ MeV [PDG]} \end{aligned} \tag{2.32}$$

while the most recent combination of experimental measurements [2] is:

$$m_W = 80369.2 \pm 13.3 \text{ MeV} \tag{2.33}$$

which differs from the electroweak fit predictions at about 1σ. It can be seen that the experimental uncertainty on m_W is about twice as large as the one from the indirect measurement. BSM physics can contribute

Quantity	Value	Electroweak Fit	Pull
m_Z [GeV]	91.1876 ± 0.0021[33]	91.1884 ± 0.0019	-0.4
	91.192 ± 0.007[3]		0.6
Γ_Z [GeV]	2.4955 ± 0.0023	2.4940 ± 0.0009	0.7
σ_{had} [nb]	41.481 ± 0.033	41.481 ± 0.009	0.0
R_e	20.804 ± 0.050	20.736 ± 0.010	1.4
R_μ	20.784 ± 0.034	20.736 ± 0.010	1.4
R_τ	20.764 ± 0.045	20.781 ± 0.010	-0.4
R_b	0.21629 ± 0.00066	0.21583 ± 0.00002	0.7
R_c	0.1721 ± 0.0030	0.17221 ± 0.00003	0.0
$A_{FB}^{(0,e)}$	0.0145 ± 0.0025	0.01606 ± 0.00006	-0.6
$A_{FB}^{(0,\mu)}$	0.0169 ± 0.0013		0.6
$A_{FB}^{(0,\tau)}$	0.0188 ± 0.0017		1.6
$A_{FB}^{(0,b)}$	0.0996 ± 0.0016	0.1026 ± 0.0002	-1.8
$A_{FB}^{(0,c)}$	0.0707 ± 0.0035	0.0732 ± 0.0002	-0.7
$A_{FB}^{(0,s)}$	0.0976 ± 0.0114	0.1027 ± 0.0002	-0.4
$\bar{s}_l^2$	0.2324 ± 0.0012[33]	0.23161 ± 0.00004	0.7
	0.23148 ± 0.00033[45]		-0.4
	0.23145 ± 0.00028		-0.6
A_e	0.15138 ± 0.00216[76]	0.1463 ± 0.0003	2.3
	0.1544 ± 0.0060[77]		1.3
	0.1498 ± 0.0049[33]		0.7
A_μ	0.142 ± 0.015		-0.3
A_τ	0.136 ± 0.015[77]		-0.7
	0.1439 ± 0.0043[33]		-0.6
	0.144 ± 0.015[78]		-0.2
A_b	0.923 ± 0.020	0.9347	-0.6
A_c	0.670 ± 0.027	0.6674 ± 0.0001	0.1
A_s	0.895 ± 0.091	0.9356	-0.4

Table 2.5: Z pole observables, compared with the global electroweak fit value (and the pull). The third value of $\bar{s}_l^2$ is a combination of [79, 42, 43, 80, 44]. From [9].

to m_W through additional radiative corrections in equation 2.31, so reaching a level of precision similar, if not better, to the theoretical predictions is crucial to test the Standard Model.

The combination of m_W measurements in [2] does not include the reanalysis of $\sqrt{s} = 7$ TeV data published by ATLAS in 2024 [37] of $m_W = 80366.5 \pm 15.9$ MeV, in agreement with the most recent average of measurements, and does not include the measurement by the CDF Collaboration [3] $m_W = 80433.5 \pm 9.4$ MeV. The CDF measurement was considered in combination with the other measurement and removed from the combination due to poor agreement with the other experimental measurements: the combined $p-$value of all m_W measurements is 0.5% when the CDF measurement is included. This $p-$value increases to 91% when the CDF measurement is removed. A measurement of m_W from the CMS experiment might shed light on this disagreement. If the result from CDF is confirmed, this could be a hint towards the presence of BSM physics.

The value of m_W is very important as it also affects the determination of other EWPO. For example, the uncertainty on the indirect prediction of m_t is directly affected by m_W, and, prior to its discovery, m_W was used to provide a confidence interval of the mass of the Higgs boson m_H. In these cases, global electroweak fits can be used to provide simultaneous predictions on pairs of EWPOs that are compared with the direct measurements. This is shown in figure 2.4, where the simultaneous prediction on m_H and m_t is shown on the left, while the simultaneous prediction on m_W and m_t is shown on the right, together with the experimental measurements. These are shown with several sources of uncertainty, indicated in the figures. It can be seen that the agreement between the experimental measurements and indirect predictions is good, except for the CDF II measurement of m_W.

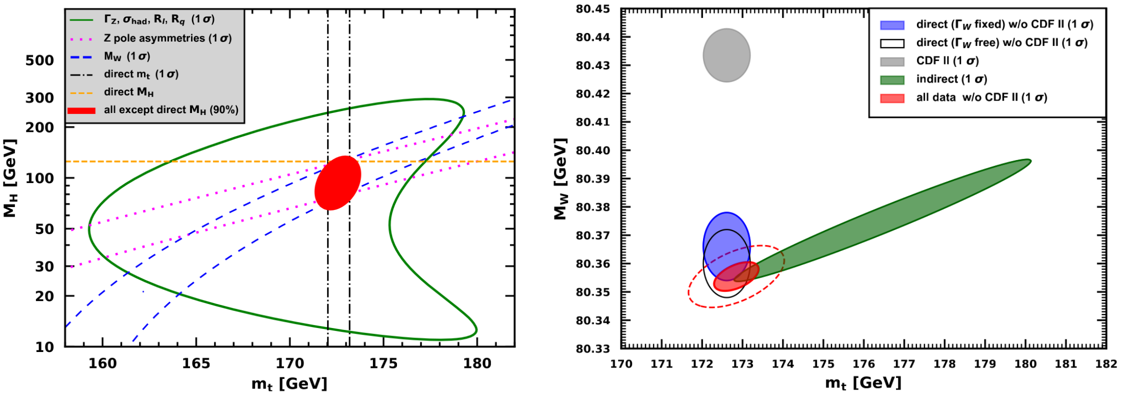


Figure 2.4: Left: fit result and one-standard-deviation uncertainties in m_H as a function of m_t for various inputs, and the 90% CL region allowed by all data. Right: One-standard-deviation regions in m_W as a function of m_t for the indirect data from the fit (green), the direct data for m_t and m_W without the CDF II measurement (blue and black contour, respectively, for two differently treatments of the W boson width), and the combination thereof (red). For the combination, the 90% CL region is also shown as a red dashed contour. The grey region uses only the m_W value from CDF II. From [9]

As mentioned above, global electroweak fits can be exploited to constrain BSM models. One possible strategy to model the presence of BSM physics is to include oblique parameters S, T, and U [81]. In this parameterization, the contribution of new physics is considered only in the electroweak radiative corrections, which are assumed to be flavour universal. Direct vertex or box corrections to the SM Lagrangian are not included, and therefore the new physics is only weakly coupled with SM fermions. The symmetry group is still $SU(2)_L \times U(1)_Y$. It is assumed that the BSM physics described by this parameterization appears only at energy scales significantly higher than the electroweak scale. The new physics will be encoded in six corrections to the γ, Z, W self-couplings. The parameters S, T, and U are defined by subtracting the SM component from the radiative corrections of the electroweak boson, and the remaining three corrections are reabsorbed in the renormalization of α, G_F, and m_Z. The SM prediction is $S = T = U = 0$. The effect of non-vanishing oblique parameters is the following:

- T is proportional to the difference between the W and Z self-energies at $Q^2 = 0$ (i.e., vector SU(2)-breaking), mainly constrained by precision measurements of Γ_Z
- S (S + U) is associated with the difference between the Z (W) self-energy at $Q^2 = M^2_{Z,W}$ and $Q^2 = 0$ (axial $SU(2)$-breaking). S is mainly constrained by precision measurements of m_Z.
- U modifies directly the radiative corrections on m_W, and is therefore strongly related to m_W and Γ_W.

m_W can be rewritten as:

$$m_W^2 = \frac{m^2_{W,SM}}{1 - G_F m^2_{W,SM}(S+U)/2\sqrt{2}\pi} \tag{2.34}$$

The results of the global electroweak fit performed by the Particle Data Group [9] are shown in figure 2.5, where it can be seen that the values are in agreement with the Standard Model prediction of $S = T = U = 0$. The figure also includes the oblique parameters obtained when the CDF II measurement is included, and a deviation from the Standard Model prediction can be observed.

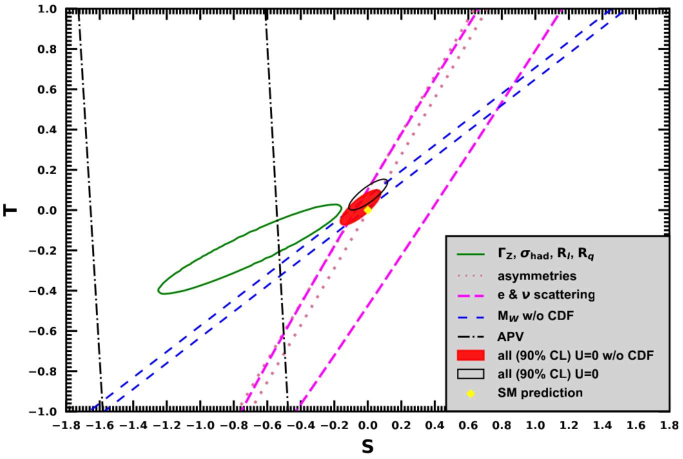


Figure 2.5: 1σ constraints on S and T (for U = 0) from various inputs combined with m_Z. S and T represent the contributions of new physics only. Data sets not involving m_W or Γ_W are insensitive to U. The yellow dot indicates the Standard Model values $S = T = 0$. From [9].

An alternative approach to probe the existence of BSM Physics is to extend the Standard Model Lagrangian in the framework of Effective Field Theories (SMEFT). In SMEFT, the leading effects of the BSM physics above the EW scale are parametrized with a set of dimension-6 (6D) operators:

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i}{\Lambda^2} O_i \tag{2.35}$$

where $\mathcal{L}_{\text{SM}}$ defines the SM Lagrangian and c_i and Λ denote dimensionless Wilson coefficients and the cutoff scale of the BSM physics, respectively. This approach was employed in [82] to explain the discrepancy between the CDF measurement and the Standard Model prediction on m_W. In particular, they considered only the key 6D operators that are closely related to the W boson physics. By fitting to the data, they showed that an upward shift in m_W is driven by the operator $O_T = \frac{1}{2}(H^\dagger \overleftrightarrow{D}_\mu H)^2$ with a coefficient $c_T(\text{TeV}/\Lambda)^2 \gtrsim 0.01$. This suggests that the new physics scale favored by the CDF data should be multiple TeV for tree-level effects and sub TeV for loop-level effects.

Chapter 3

The CMS experiment at the LHC

The first part of this chapter contains a description of the Large Hadron Collider (LHC) at CERN. The second part of this chapter is devoted to the Compact Muon Solenoid (CMS) experiment at CERN, with particular attention to the functioning of its subdetectors.

3.1 The LHC

The Large Hadron Collider (LHC) [83] is a circular proton-proton collider located at the European Organization for Nuclear Research (CERN). It is hosted in the 26.7 km long LEP tunnel and it is designed to accelerate and perform head-on proton-proton (pp), nuclei (heavy nuclei such as Lead and Xenon, but also lighter ones such as Oxygen) and proton-nuclei collisions. The design center-of-mass energy in proton-proton collisions is $\sqrt{s} = 14$ TeV and the design instantaneous luminosity is 10^{34} cm^{-2} s^{-1}. The LHC construction ended in 2008 and the first physics run, Run 1, took place between 2010 and 2013, with a beam energy of 3.5 TeV (4 TeV in 2012) in pp collisions. After the Long Shutdown 1, in which several upgrades were implemented on the LHC and its experiments, the Run 2 started in 2015 and ended in 2018. The beam energy in pp collisions was increased to 6.5 TeV. Further upgrades were implemented during the Long Shutdown 2, and Run 3, the current physics period, started in 2021, and will end in 2026. In 2022 the proton beam energy was increased to 6.8 TeV, with an unprecedented total center-of-mass energy of 13.6 TeV. After Run 3 the LHC will undergo a major upgrade, the High-Luminosity Large Hadron Collider (HL-LHC) [84], with the aim of achieving instantaneous luminosities a factor of 5 to 7.5 larger than the LHC nominal value, and the center-of-mass energy should be increased to 14 TeV.

The Large Hadron Collider is the latest addition to the CERN accelerator complex [85], figure 3.1. First, protons are injected in the Linear Accelerator 2 (Linac2), then they are passed to the Proton Synchrotron Booster (PS booster), then to the PS ring, and finally to the Super Proton Synchrotron (SPS), where they reach a beam energy of 450 GeV, before the injection to the LHC, where they are accelerated to the final energy and where the proton beams finally collide. The LHC is composed of two rings, with four interaction points (IPs), which host the four experiments of the LHC: ATLAS [86], CMS [87], ALICE [88] and LHCb [89]. The first two are general purpose experiments, covering the largest range of physics possible, and being responsible for the discovery of the Higgs boson in 2012, 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.

The particles are accelerated by 8 radiofrequency cavities per beam operating at a frequency of 400 MHz and providing an accelerating field of 5 MV/m. The bending of the beams is realized with 1232 superconducting dipole magnets that provide a magnetic field of 8.33 T. A total of 392 quadrupole magnets provide the focusing of the beams and additional steering. The two rings of the LHC are located inside the superconductive magnets and the ultrahigh vacuum is achieved with a cryopump system.

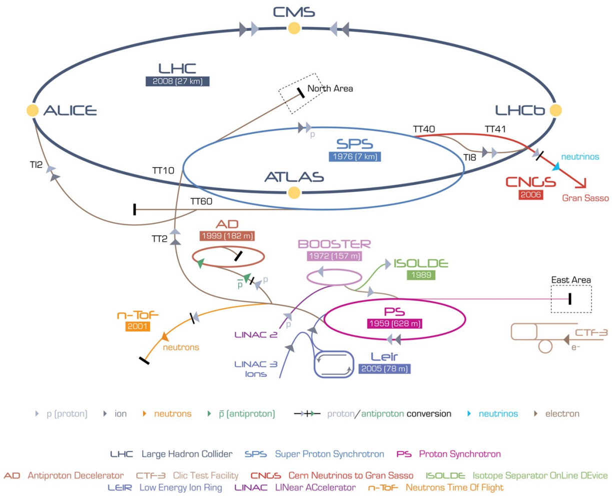


Figure 3.1: The CERN accelerator complex, the grey arrows follow the path for pp collisions (from [85]).

The LHC beams are bunched with a spacing of 25 ns, with $O(10^{11})$ protons per bunch. The LHC was designed to have 2808 circulating bunches, each with a roughly Gaussian-shaped distribution in the transverse plane with $\sigma_x = \sigma_y = 16\ \mu$m at IPs and a longitudinal dimension of about 7.5 cm. These parameters are fundamental as they determine the collision rate R for a specific process:

$$R = \mathcal{L} \times \sigma \tag{3.1}$$

where σ is the cross section of the specific process and $\mathcal{L}$ is the luminosity. The luminosity (in the case of head-on collisions) is related to the beam parameters through the following:

$$\mathcal{L} = \frac{f N_1 N_2 k}{4\pi\sigma_x\sigma_y} \tag{3.2}$$

where $f = 40$ MHz is the bunch-crossing rate, N_1 and N_2 are the number of protons per bunch for each beam, σ_x and σ_y are the transverse dimensions of the bunches and k is the number of bunches. The integrated luminosity $\mathcal{L}_{int} = \int \mathcal{L}dt = N/\sigma$, where N is the number of events for a given process characterized by the cross-section σ, is usually quoted to show the amount of data available for the analyses. Figure 3.2 shows the peak and integrated luminosities for each year of data taking until 2024. In 2017 LHC was able to deliver an instantaneous luminosity which doubled the design one, and delivered even higher luminosities in the later years. The total integrated luminosity delivered to CMS (not all of this used for physics analyses) is roughly 30 fb^{-1} for Run 1, roughly 163 fb^{-1} for Run 2 and roughly 196 fb^{-1} for Run 3 as of 2024 (122 fb^{-1} were delivered in 2024 alone).

Protons are composite objects. When a proton-proton collision occurs, it is the quarks and gluons, also referred to as partons, that collide. The proton-proton interaction is described in terms of the parton model: at large transferred momentum $q^2 = (p_1 - p_2)^2$, where p_1 and p_2 are the four-momenta of the colliding partons inside the protons, the process is called *hard scattering* (for this regime $|q^2| \gtrsim 10$ GeV^2). The process can be interpreted as the collision of the two isolated partons, each of which carries a fraction x_i of the total proton momentum. The center-of-mass energy of this collision is $\sqrt{\hat{s}} = \sqrt{x_1 x_2 s}$. Each of the partons is characterized by a parton density function, or PDF, which quantifies the fraction of the total proton momentum a parton carries. Figure 3.3 on the left shows the PDFs for each of the partons as a function of x and as a function of $Q^2 = -q^2$. The total hard-scattering cross section for a proton-proton collision can therefore be written as:

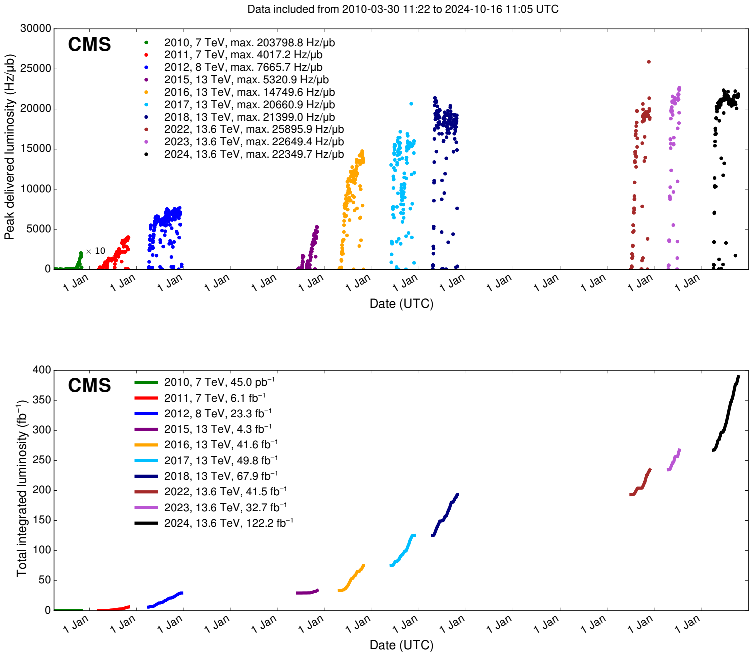


Figure 3.2: LHC luminosity (upper plot) and integrated luminosity (lower plot) delivered to CMS experiment between 2010 and 2024 in pp collisions (from [90])

$$\sigma_{pp} = \sum_{i,j} \int \int f_i(x_1, Q^2 = \mu_F^2) f_j(x_2, Q^2 = \mu_F^2) \hat{\sigma}_{i,j}(x_1 p_1, x_2 p_2, \mu_F, \mu_R) dx_1 dx_2 \tag{3.3}$$

where $\hat{\sigma}_{i,j}$ is the cross section for the interactions of partons i and j, f_i and f_j are the parton density functions, μ_F and μ_R are the factorization and renormalization scales, which arise from fixed-order calculations, but the cross section must not depend on these parameters at all orders. The integral runs over all the range of x for both partons (between 0 and 1) and the sum runs over all the possible pairs of partons. Figure 3.3 on the right shows the pp cross-sections for some processes as a function of the center-of-mass energy. The interaction of the proton remnant is called *underlying event*, which involves low q^2 interaction between partons, initial and final state radiation. Due to colour confinement, these objects must combine to create colourless bound states, the so-called *hadrons*. This is a complex process, since it requires non-perturbative QCD, and is described in terms of dedicated simulations.

Given the high luminosity at the LHC, multiple proton-proton collisions occur when two bunches collide. In addition to the hard scattering (and the underlying event), there is a (large) number of collisions with lower transferred momentum q^2, they are usually referred to as minimum bias events, and a lot of low-momentum particles are produced. The additional collisions are referred to as pileup (PU) collisions, which are often a source of background in the event and must be properly described and controlled, as they have a large impact on some the reconstructed quantities in the event. Figure 3.4 shows, on the left, an example of a high-pileup proton-proton collision at the LHC. In orange, the reconstructed vertices can be seen, and in yellow some of the charged particles passing a dedicated selection are shown. Figure 3.4 on the right shows the distribution of the number of interactions per bunch crossing for all the years of data taking up to 2024. The increase in instantaneous luminosity can be clearly seen, as the distributions are systematically shifted to the right.

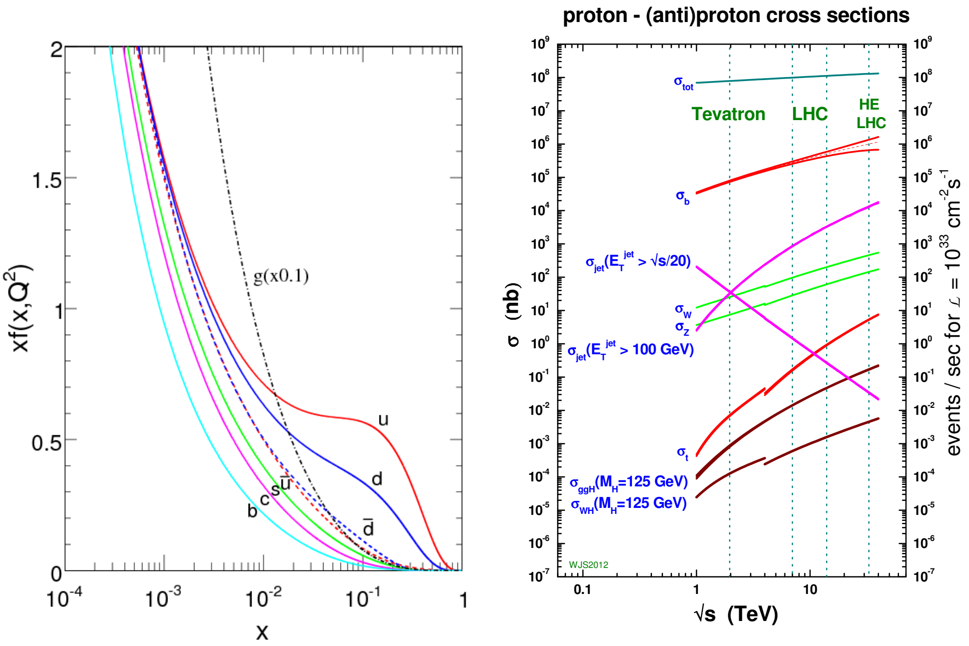


Figure 3.3: Left: figure showing $x \times \mathrm{PDF}(x)$ for the various proton constituents (partons) (from [91]). Right: Standard Model process cross sections at hadron colliders as a function of centre-of-mass energy (from https://twiki.cern.ch/twiki/bin/view/LHCPhysics/HiggsEuropeanStrategy)

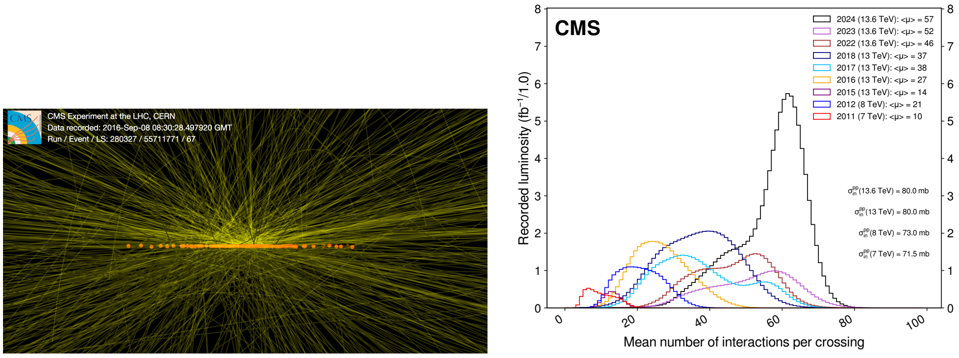


Figure 3.4: Left: a high-pileup event with 86 reconstructed vertices, here given by the orange circles (from https://cds.cern.ch/record/2241144). Right: distribution of the average number of interactions per crossing (pileup) for pp collisions in 2011 (red), 2012 (blue), 2015 (purple), 2016 (orange), 2017 (light blue), 2018 (navy blue), and 2022 (brown). The overall mean values and the minimum bias cross sections are also shown (from [90]).

3.2 The CMS experiment

The CMS experiment [87] is one of the four experiments at the Large Hadron Collider. It is one of two general-purpose detectors at the Large Hadron Collider (LHC) at CERN, alongside ATLAS [86]. CMS played a crucial role in discovering the Higgs boson in 2012, confirming the Higgs mechanism that explains electroweak symmetry breaking in the Standard Model. As a general-purpose experiment, CMS conducts a wide range of research, including studying Higgs boson properties, conducting precision tests of the Standard Model, and exploring phenomena like extra dimensions, supersymmetry, and dark matter.

The CMS experiment uses a cylindrical coordinate system (r, ϕ, z) with the $z-$axis being directed along the magnetic field and beam directions, r being the distance from the $z-$axis and ϕ the angle between the $x-$axis, which points towards the center of the LHC. The particle kinematics is described using pseudorapidity (η), defined as $-\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$, where θ is the angle between the direction of flight of the particle and z. This is an approximation in the case the energy is much larger than the mass of particles, which is always the case for the particles that are detected by the CMS system, of the rapidity y:

$$y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z}, \tag{3.4}$$

where E is the energy and p_z is the momentum of the particle along the $z-$axis. Distributions computed in this variable are invariant under Lorentz transformations along the $z-$axis, and this variable is used since the total momentum along the $z-$axis is not known, as partons are colliding inside the protons.

The detector weighs 14,000 tonnes and is 15 meters high and 21 meters long. It is divided into barrel ($|\eta| \leq 1.2$) and endcap ($|\eta| > 1.2$) regions. At its core is a 13-meter-long, 6-meter-diameter superconducting solenoid generating a 3.80 T magnetic field and whose return field is enough to saturate 1.5 m of iron. The detector layout, shown in figure 3.5, includes:

- Inner tracking system: composed of 3 (4, since 2017) layers of silicon pixel detectors and 10 silicon microstrip layers with a coverage up to $|\eta| < 2.5$.
- Electromagnetic Calorimeter (ECAL): Made of $PbWO_4$ crystals, coverage up to $|\eta| < 3$.
- Hadronic Calorimeter (HCAL): Composed of brass and scintillator plates, coverage $|\eta| < 3$.
- Forward calorimeter: Covering $3 < |\eta| < 5$.
- Four muon stations integrated into the iron return yoke comprise Drift Tubes (DTs), Cathode Strip Chambers (CSCs), and Resistive Plate Chambers (RPCs), covering $|\eta| < 2.4$.

All the CMS detectors, except for the muon stations and the forward calorimeters, are housed inside of the bore of the magnet.

Each of the subdetectors is described in more detail in the following subsections.

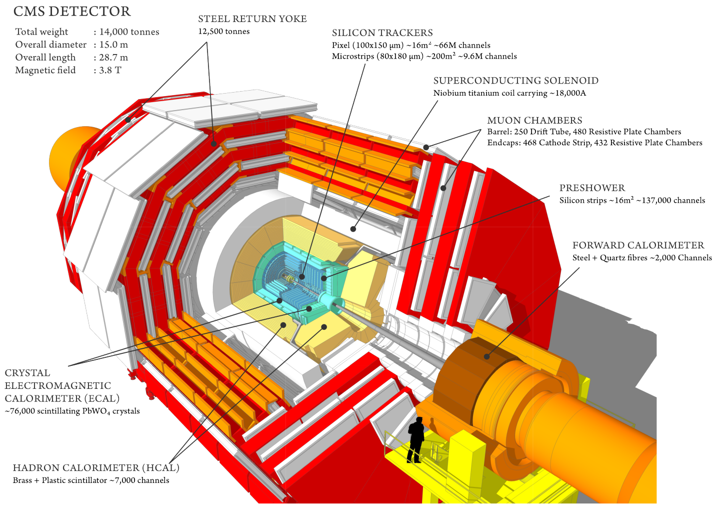


Figure 3.5: Sectional view of the CMS detector. Reference: https://cms.cern/book/export/html/1190

3.2.1 Detectors

CMS tracker

Figure 3.6 ([92, 93]) shows a layout of the CMS silicon tracker. The entire tracker is a cylinder, 6.5 m long, with a radius of 1.2 m. It is designed to reconstruct the trajectories of charged particles, referred to as tracks, from the interaction of the charged particles with the silicon modules of the tracker (hits). The reconstructed tracks are combined to reconstruct interaction vertices, both from the proton-proton interactions, the so-called *primary vertices*, and from decays of light particles, the so-called *secondary vertices*.

The Large Hadron Collider (LHC) environment presents significant challenges for particle tracking systems. Due to the high pileup, the detectors must have a high granularity to identify individual tracks and vertices. To match the 40 MHz bunch crossing frequency at the LHC, the number of channels and the consequent readout electronics must operate at high speeds and allow for efficient and fast data storing. At the same time, the total material budget of the tracker, i.e. the amount of material the tracks cross as they go through the tracker, must be minimized to reduce trajectory distortions caused by multiple scattering, while the individual modules need exceptional radiation resistance.

Closest to the interaction region the hits are measured by Silicon Pixel Detectors. Until 2016, these were arranged in three cylindrical layers in the barrel region, positioned at increasing radial distances r, two concentric layers on each side of the interaction region, placed at increasing distances along the $z-$axis. The dimensions of the pixels are 100×150 μm^2 and 285 μm thickness and have very low occupancy. These pixel detectors provide true 3D hits with a resolution of approximately 10 μm in the $r - \phi$ direction and $20 - 40$ μm in the z direction. The pixel detectors was replaced during 2016 Technical Stop due to the radiation damage after Run 1 and the beginning of Run 2, and the current version is referred to as *Phase-1* pixel detector. A fourth layer in the barrel and a third disk in the endcap were added compared to the original design, with both the first layer in the barrel and the first disk in the endcap brought closer to the interaction point. The new configuration ensures better performance in the track-seeding, due to the additional redundancy in the pixel hit information. The Phase-1 pixel detector provides 4 tracking hit coverage up to $|\eta| = 2.5$. Both pixel layouts are shown in the first figure of figure 3.6. The first layer of the barrel pixel was replaced again prior the to beginning of Run 3 to cope with the radiation damage from Run 2.

Outside the Pixel detectors the charged tracks trajectories are measured using Silicon Microstrip Detectors. They are arranged in four subsystems: the Tracker Inner Barrel (TIB) and the Tracker Inner Disks (TID) cover the region $r < 55$ cm and $|z| < 118$ cm, with four strip detector layers and tree strip detector disks, respectively. They provide hit position measurements in $r - \phi$ with a resolution of $13 - 38$ μm. The Tracker Outer Barrel (TOB) and the Tracker Endcaps (TEC) cover the region 55 cm $< r < 120$ cm, $z < 118$ cm and $r < 120$ cm, 118 cm $< z < 250$ cm, with 6 layers and 9 disks, respectively. The hit resolution for TOB and TEC is approximately $18 - 43$ μm. To enable 3D hit reconstruction, additional strip sensors are placed back-to-back with the main sensors, at a stereo angle of 100 mrad, in the two innermost layers of TIB and TOB, modules 1, 2 of the TOB, and modules 1, 2, 5 of the TEC.

Electromagnetic calorimeter

The Electromagnetic calorimeter (ECAL) [94], shown in figure 3.7, is a homogeneous calorimeter made of about 75000 lead-tungstate ($PbWO_4$) crystals. The primary function of the calorimeter is to measure the energy of particles in a destructive manner by stopping and absorbing them, thereby collecting the total energy they release within the detector. The Electromagnetic Calorimeter (ECAL) is specifically optimized to enhance the likelihood of electromagnetic interactions with its material, enabling precise energy measurements for photons and electrons. In addition, ECAL plays a crucial role in the CMS hardware trigger system, thanks to its rapid response time and fast readout capabilities.

The ECAL consists of a barrel (ECAL Barrel, EB) whose coverage extends to $|\eta| < 1.48$ and two endcaps

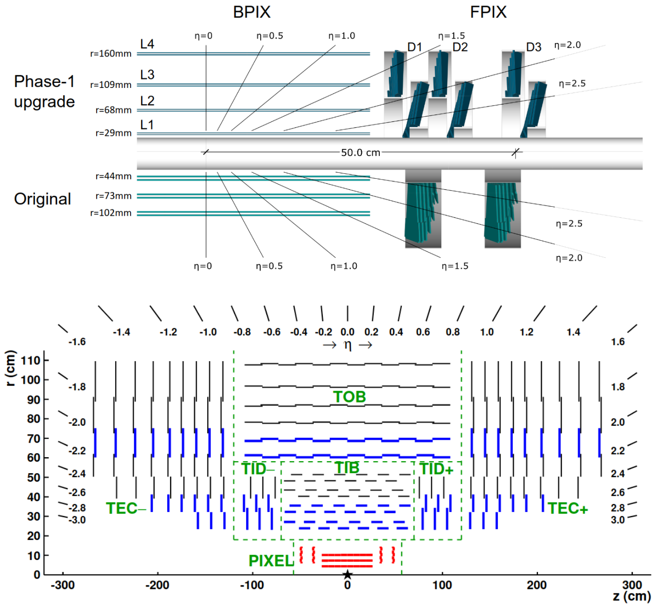


Figure 3.6: Layout of the CMS tracker. The first figure, the layout of the pixel subdetector, is from [93], the second figure, the overall layout of the tracker, is from [92].

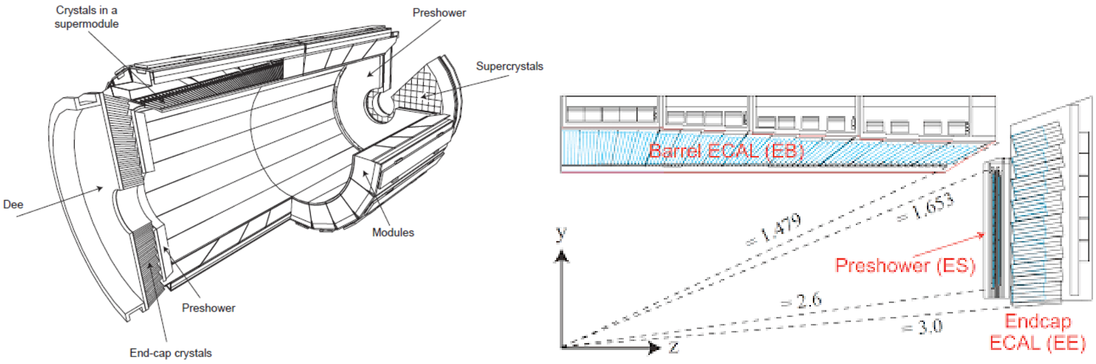


Figure 3.7: Left: Schematic layout of the CMS electromagnetic calorimeter. Right: Geometric view of one quarter of the ECAL.

(ECAL Endcap, EE) covering up to $|\eta| < 3$. In the region $1.6 < |\eta| < 2.6$ also features the presence of a lead/silicon-strip pre-shower detector. The depth of the detector is $26\ X_0$[1] in the barrel region and $3 + 25$ X_0 in the pre-shower and endcap region, and the transverse dimension is equal to the Molière radius (2.2 cm). The ECAL material also provides about 1 nuclear interaction length (λ_0), so that the hadrons start showering in the ECAL material. The scintillation light is collected by silicon avalanche photodiodes in the barrel, mounted in pairs at the end of each crystal. Vacuum phototriodes are used instead in the endcap, glued to the back of each crystal. This configuration allows for comprehensive electromagnetic energy measurements across a wide range of pseudorapidities, enhancing the ability of the detector to accurately reconstruct and measure particles produced in collisions.

The energy resolution can be parameterized as:

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{a}{\sqrt{E}}\right)^2 + \left(\frac{b}{E}\right)^2 + c^2 \tag{3.5}$$

---

[1] PbWO_4 $X_0 = 0.89$ cm

where a is the stochastic term due to the fluctuations of the shower produced when a particle crosses the calorimeter, b is the electronic noise term, c is the calibration term. Measurements on test beams [94] have yielded the results in table 3.1.

Contribution	EB	EE
Stochastic term (a)	2.7% GeV$^{\frac{1}{2}}$	5.7% GeV$^{\frac{1}{2}}$
Noise term (b) for low luminosity	155 MeV	205 MeV
Noise term (b) for high luminosity	210 MeV	245 MeV
Calibration term (c)	0.55%	0.55%

Table 3.1: ECAL resolution fit results [94]

Hadronic calorimeter

The hadronic calorimeter (HCAL) [95] is designed to measure the energies of hadrons. The energy of nuclear hadrons can only be measured by HCAL, while the information from the CMS tracker is used to measure the kinematic properties of charged hadrons together with the HCAL information, which is also used to trigger events. Figure 3.8 shows the layout of the CMS hadronic calorimeter.

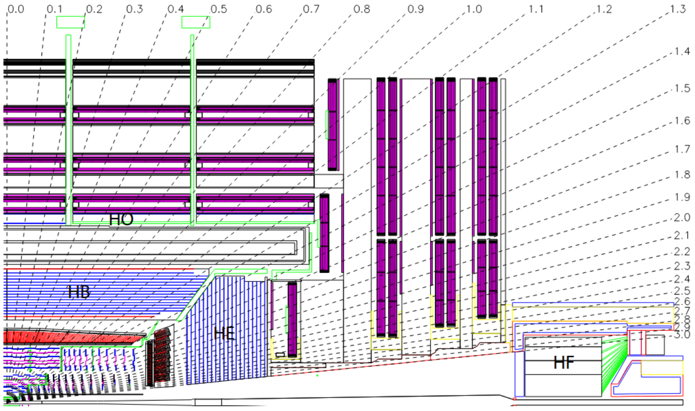


Figure 3.8: Longitudinal view of one quarter of the CMS hadronic calorimeter. Reference: [87].

The HCAL system consists of four main components: the *Barrel Hadronic Calorimeter* (HB), the *Endcap Hadron Calorimeter* (HE), the *Very Forward Calorimeter* (HF) and the *Outer Barrel* (HO). HB is a sampling calorimeter made of plastic scintillators interleaved with brass plates surrounding the ECAL barrel, with a coverage of $|\eta| < 1.3$ and a depth of $5\ \lambda_0$, whereas HE has a coverage of $|\eta| < 3.0$ and a depth of $10\ \lambda_0$. HF is a Cherenkov calorimeter, instrumented with steel interlayered with quartz fibers, with a coverage of $|\eta| < 5.0$. HO, made of scintillators, is included in the design because the depth of the barrel region alone is not sufficient to fully contain hadronic showers. This additional component ensures a more complete energy measurement for hadrons in the barrel region. The granularity of the HCAL is coarser than that of the ECAL and a single HCAL module matches a matrix of 5×5 ECAL crystals. The resolution for the hadronic calorimeters can be parameterized as:

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{a}{\sqrt{E/\mathrm{GeV}}}\right)^2 + b^2 \tag{3.6}$$

where in the barrel region $a \approx 65\%$ and $b \approx 5\%$, up to the very forward region in which $a \approx 100\%$ and $b \approx 9\%$.

Superconducting magnet

The CMS magnet system [96] consists of a cylindrical solenoid 12.5 m long and 6 m in diameter, producing a magnetic field of 3.80 T along the $z-$axis within its inner volume. This 220-ton cold mass, operating in a superconducting state at 4.5 K, stores 2.6 GJ of energy at full current. The magnet conductor is made of NbTi strengthened with an aluminum alloy, wound in four layers to form a thin coil with a radial extent ratio ($\Delta R/R$) of 0.1.

The magnetic flux return path comprises five barrel wheels and three disks at each endcap, constructed from iron and housing the muon detection system. Figure 3.9 depicts the intensity of the magnetic field across various CMS detector subsystems. In the barrel return yoke, the field strength ranges from 1 to 2 T within the iron components, dropping below 0.1 T in the muon chambers. The magnetic field within the internal volume of the solenoid varies from 3.8 T in the central region to about 3 T at the extremities.

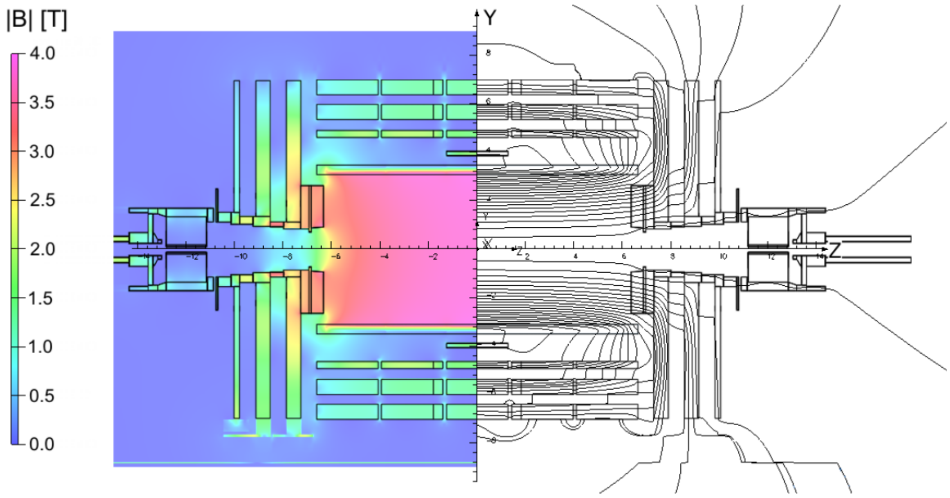


Figure 3.9: Absolute value of the magnetic field $|B|$ (left) and field lines (right) predicted on a longitudinal section of the CMS detector, for the underground model at a central magnetic flux density of 3.8 T (from [97]).

Muon system

Figure 3.10 shows a section of the Muon Spectrometer [98]. Energetic muons are the only detectable particles that traverse the entire CMS detector. The primary function of the muon system is to measure the momenta, positions, and timing of the muons, while also serving for hardware triggering and rapid high-p_T tracking. For muons with $p_T \lesssim 200$ GeV, the silicon tracker already provides all the information for precise muon momentum measurements, and the muon system is primarily used for identification, while the information from the muon system becomes relevant above 200 GeV. Muons below 3 GeV are confined within the inner region of the detector because of magnetic bending. Between $3-6$ GeV, muons reach the muon system but are absorbed by the iron return yoke, allowing only partial information recovery from the muon system.

The muon system consists of four stations integrated into the iron return yoke of the magnetic field. It employs drift tubes (DTs) in the barrel region ($|\eta| < 1.3$) and cathode strip chambers (CSCs) in the endcap region ($0.9 < |\eta| < 2.4$). In addition, Resistive Plate Chambers (RPCs) are incorporated throughout the region $|\eta| < 1.6$, with Gas Electron Multipliers (GEMs) installed from $1.6 < |\eta| < 2.4$, complementing the DT and CSC detectors.

Drift tubes measure the drift time for electron ionization, initiated by fast moving charged particles, to reach the wire, converting this time into a distance. In each station, they are arranged in four radial layers, divided into 12 ϕ sections. The magnetic field, confined to iron layers, does not affect the DTs, so particle propagation can be considered linear inside the DT layers. Each of the muon stations measures the positions of the muons with a resolution of ≈ 150 μm in the $z-$direction and ≈ 100 μm in the $r - \phi$ direction. The direction is measured with a resolution of ≈ 1 mrad.

Cathode strip chambers (CSCs) are used in the endcaps instead of drift tubes because of the highly variable magnetic field in this region and the much higher punch-through flux (particles that penetrate all detectors before reaching the muon system). In CSCs, measurements are taken using anode wires stretched between two cathodes, one of which is segmented into strips perpendicular to the wires. The anode wires detect the avalanche caused by charged particles to measure one coordinate, while the induced charge distribution on the cathode strips provides the perpendicular coordinate. Each endcap station consists of six layers of CSCs, achieving a hit resolution of $75 - 150$ μm in the $r - \phi$ coordinate and 250 μm in the z coordinate.

Resistive Plate Chambers (RPCs) complement DTs and CSCs in both the barrel and endcap regions. RPCs primarily provide timing measurements for muon triggering and bunch crossing identification, with a resolution of ≈ 3 ns. Due to high radiation at high $|\eta|$, they only cover up to $|\eta| < 1.6$, with Gas Electron Multiplier (GEM) detectors installed during Long Shutdown 2 [99] to cover the region $1.6 < |\eta| < 2.4$.

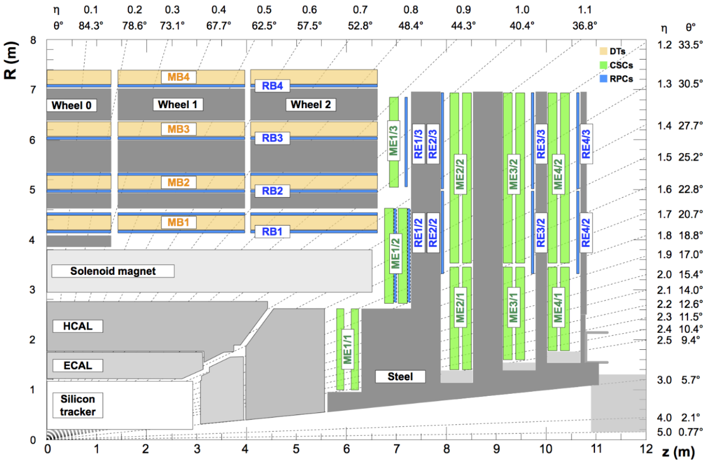


Figure 3.10: An $R - z$ cross-section of a quadrant of the CMS detector, with the beam axis (z) horizontal and radius (R) vertical. The interaction point is at the bottom left. Muon stations and steel flux-return disks (dark areas) are shown. Drift tubes (DTs) are labeled MB, cathode strip chambers (CSCs) are ME, and resistive plate chambers (RPCs) are RB in the barrel and RE in the endcaps. From [98].

Particle Flow algorithm

The *Particle Flow* (PF) algorithm [100] uses the information from all the CMS detectors to provide a global description of the events, exploiting the different kinds of interactions with the detectors. In the first step, the PF algorithm gathers information from each detector to create *PF elements*, which include inner tracks, muon tracks, and various calorimeter clusters (ECAL, HCAL, preshower). In the second step, these PF elements are grouped into *PF blocks*. Two PF elements are assigned to the same PF block if the algorithm establishes a link between them. These links are created based on specific criteria depending on the types of PF elements involved. To reduce computational complexity, only nearest neighbors in the $\eta - \phi$ plane are considered. For each linked pair, a distance metric is defined to evaluate the quality of the connection. Each block is then analyzed separately to identify the different kinds of *PF candidates*, starting from muons, electrons and isolated photons, hadrons and non-isolated photons and nuclear interactions.

The CMS trigger system

At the LHC, proton collisions occur every 25 ns, equivalent to a rate of 40 MHz. When two bunches cross, 25-50 collisions are produced on average, resulting in about O(10^9) interactions per second. However, only a fraction of these interactions is valuable for the CMS physics program. The Trigger System is specifically designed to identify and select these interesting events for further analysis.

The reduction of the data rate occurs in two steps:

- *Level-1 (L1)* [101]: programmable electronics that processes information from calorimeters and muon system, and selects events on the basis of specific physics objects in the events (such as muons, electrons, photons or jets). It reduces the event rate from 40 Mhz to about 100 khz. The L1 exploits a synchronous pipeline where the detector information of up to 160 bunch crossings can be stored. This approach has a latency of 3.2 μs. The L1 hardware has been strongly upgraded during the Long Shutdown 1, to face the higher luminosity in Run 2, with the aim of retaining the same efficiency in the selection of interesting events. Figure 3.11 shows the decomposition of the total L1 by trigger objects.

- *High-Level Trigger (HLT)* [102]: a processor farm of about 16000 CPUs that implements the CMSSW software framework [103], which is also used for offline reconstruction. This was updated in Run 3 to a farm composed of 25600 CPU cores and 400 GPUs in total [104], with an increasing implementation of algorithms that can be run on both CPUs and heterogeneous architectures. Full detector information is available at the HLT. To meet Level-1 (L1) timing requirements, event processing occurs in stages, with events being discarded or retained as soon as sufficient information is available. The HLT consists of "paths" - sets of algorithms and filters designed to select specific physics objects. These paths can focus on individual physics objects such as photons, electrons, or muons, or construct and evaluate more complex objects. Due to the limited bandwidth of the HLT trigger, some of the trigger paths have a prescale factor of N: this means that only a fraction of $1/N$ of all the triggered events are recorded. These prescale factors depend on the physics processes probed by the trigger paths (and to the cross-sections of these processes). Figure 3.12 shows the decomposition of the total HLT rate in 2018 by CMS Physics Analysis Groups (PAGs), grouping different analyses depending on the sector of the Standard Model being probed (e.g. electroweak physics, top physics, $b-$physics), and by CMS Physics Object Groups (POGs), related to specific objects or detectors.

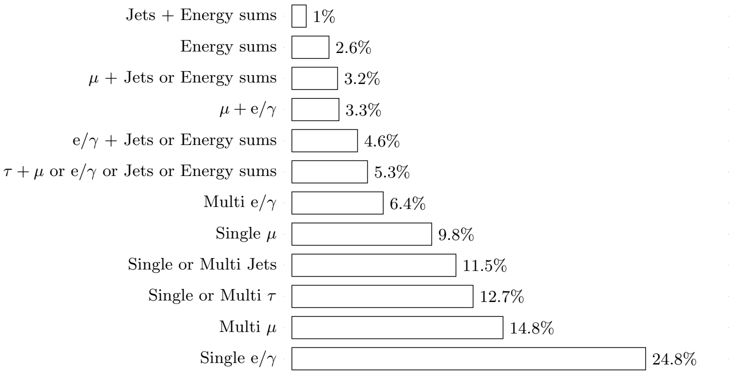


Figure 3.11: Decomposition of total L1 rate by triggering objects. From [105].

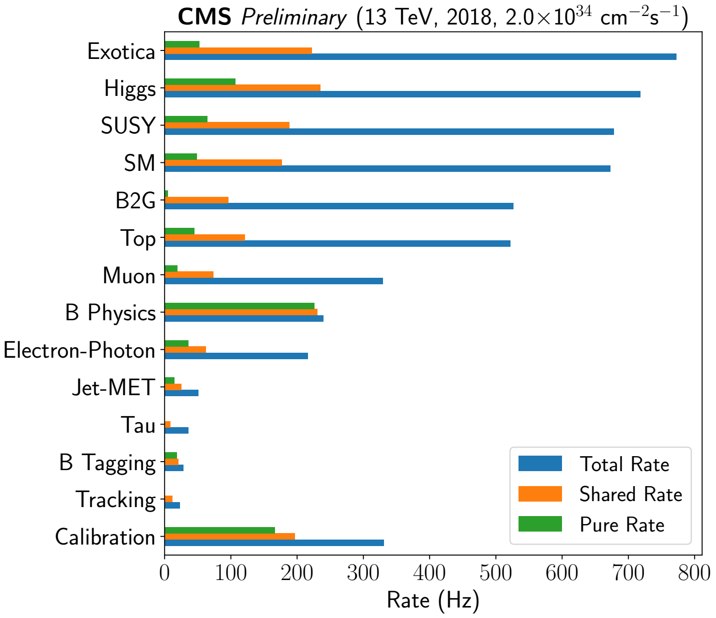


Figure 3.12: Decomposition of total HLT rate by POGs and PAGs, as of 2018. From [106].

Chapter 4

Tracking in CMS

The mass of the W boson is measured by the CMS experiment by reconstructing the kinematics of the muons produced in their decays with high precision. This reconstruction relies heavily on the tracking performances of the CMS tracker. The first part of the chapter describes how tracking, i.e. the reconstruction of charged particle trajectories, is performed in CMS. The second part of this chapter contains a summary of my original contributions to the Tracking POG. I contributed by performing detailed studies to assess the tracking performance during several years of Run 3: these studies are crucial to ensure high-quality data to be used by the whole CMS collaboration.

4.1 Tracking

Charged particles interact with the modules of the CMS silicon tracker, producing signals that are processed by the so-called *local reconstruction*, where the charge clusters and then the hits are built. Individual hits from pixel and strip sub-detectors are provided as input to the reconstruction of the trajectories of all the charged particles produced in the event, in a process referred to as *tracking*. Charged particles have almost an helix trajectory, due to the approximately uniform magnetic field provided by the CMS solenoid (figure 3.9), with a curvature radius $R[m] = p_T[\mathrm{GeV}]/(0.3B[\mathrm{T}])$, where B is the magnetic field expressed in Tesla and p_T is in GeV, assuming that the electric charge of the particles is $\pm e$. The five parameters needed to describe the helix trajectory (the curvature q/p_T, the η and ϕ angles, the transverse d_{xy} and longitudinal d_{z} impact parameter at the point of closest approach with respect to the beam line) for each charged particle in the event are computed as a result of the tracking step.

The algorithm used to build the tracks in the CMS Tracker is the *Combinatorial Track Finder* [92]: tracks are reconstructed through multiple iterations of the *Combinatorial Kalman Filter* (CKF) [107, 108]. The reason is that O($10^2 - 10^4$) hits are built in each event, resulting in a combinatorial burden not affordable in terms of computing resources if all possible hit combinations were considered. Therefore, the CMS tracking algorithm uses the Kalman filter for pattern recognition and builds a list of compatible hits across different layers, and then the track fitting is run using this information. The first iterations are designed to reconstruct tracks that are easier to find, while in the latter iterations tracks from more complex topologies are reconstructed. At the end of each iteration, the tracks are classified according to different quality criteria. According to these quality criteria, the hits associated with the reconstructed tracks may be masked from the following iterations, reducing the total number of hits and therefore the combinatorial burden.

Each iteration consists of three steps:

- *Seed generation*: Seed tracks are built using only the information from few hits (2 to 4), obtaining a rough estimate of the track parameters, used as input for the second step of the tracking sequence. At least three hits are needed to provide an estimate of the track parameters, so that at least three 3D hits (a triplet) or pairs of 3D hits (doublets) and the information on the beamspot are required. Since 2017, when a new pixel layer was added to the tracker (see subsection 3.2.1), also combinations of four pixel hits (quadruplets) are used in the seeding step. For each iteration, a set of *seeding layers* and a *tracking region* are defined: the seeding layers are the layers in which the next compatible hits are searched, while the tracking region is the region of interest determined by kinematic and geometric cuts (which vary depending on the iteration), defining the window for the pattern recognition step. Since 2017, if the seeding layers are pixel triplets or quadruplets layers a Cellular Automaton [109] is used to produce the seeds, instead of the constraints on the tracking region.

- *Pattern recognition*: The seeds are propagated by the CKF to the next layers of the tracker, looking for compatible hits, and updating the track parameters with this new information. If multiple compatible hits are found in the next layer, separate track candidates are created for each combination and propagated independently. The propagation continues until the tracker edge is reached or too many missing hits are present, taking into account the tracker conditions as a function of time, so that tracks are not penalized for missing hits associated to detector channels known to be faulty. Pattern recognition is performed inside-out, but another search is performed backward starting from the outermost hit to improve the hit-collection efficiency after the track is completed.
- *Track fitting*: for each of the track candidates, a collection of hits and an estimate of the helix parameter is returned by the pattern recognition step. The tracks are then refitted using the Combinatorial Kalman Filter to avoid biases coming from the seeding and pattern recognition step (e.g., the beam spot requirement). The Runge-Kutta propagator [110] is used to move through the tracker, as it takes into account material effects and inhomogeneities in the magnetic field. After this step, a cleaning procedure to remove hits spuriously associated to each track is performed, resulting in a better χ^2, p_T and impact parameter resolution.
- *Quality flagging*: in order to keep a high efficiency, large search windows and loose requirements are used during pattern recognition, which results in a non-negligible fake rate (up to some 10%). Therefore, tracks are categories according to specific selection criteria (based on the normalised χ^2 of the track, the number of hits, the distance of closest approach to the primary vertex, etc.). Three quality classes are realized: *loose*, *tight* and *high-purity* [92].

The CMS tracking algorithm produces a collection of tracks called *general tracks*. The hits associated to high-purity tracks are removed from the next iterations, while this is not the case for the low-purity ones, and it can occur that the same track is reconstructed in multiple iterations. A merging step is then performed after all the iterations are performed, selecting only the best tracks among those that share a large number of hits.

At the beginning of Run 3, the mkFit algorithm [111] was implemented in some iterations of the tracking reconstruction sequence. The mkFit algorithm is a new implementation of the traditional Kalman filter approach for charged particle tracking, with the aim of exploiting parallelization and vectorization in multi-core CPU architectures. This algorithm is being developed to cope with the high-pileup environment of the HL-LHC, resulting in a much higher number of charged particles, and therefore reconstructed tracks, resulting in a faster computational throughput. Figure 4.1 shows the comparison of the tracking time for the different steps of the CMS track reconstruction sequence between the standard CKF algorithm used in Run 2 and mkFit, considering all iterations and only the iterations for which mkFit is implemented. A significant reduction in the track building step, i.e. the pattern recognition step, can be observed in both cases, with a reduction in the timing of about 30% when all tracks are considered.

Figure 4.2 summarizes the different iterations implemented in the CMS tracking sequence, with a short description of the hit combinations used in the seeding step (pixel quadruplets, triplets, tracking inside jets, etc.), as of Run 3. The figure also shows what are the iterations for which the mkFit algorithm has been implemented. The distinction between *All Tracks* track candidates and *Tracked-only Seeded* track candidates is shown, as the latest two iterations target muon reconstruction. Unlike the other iterations,

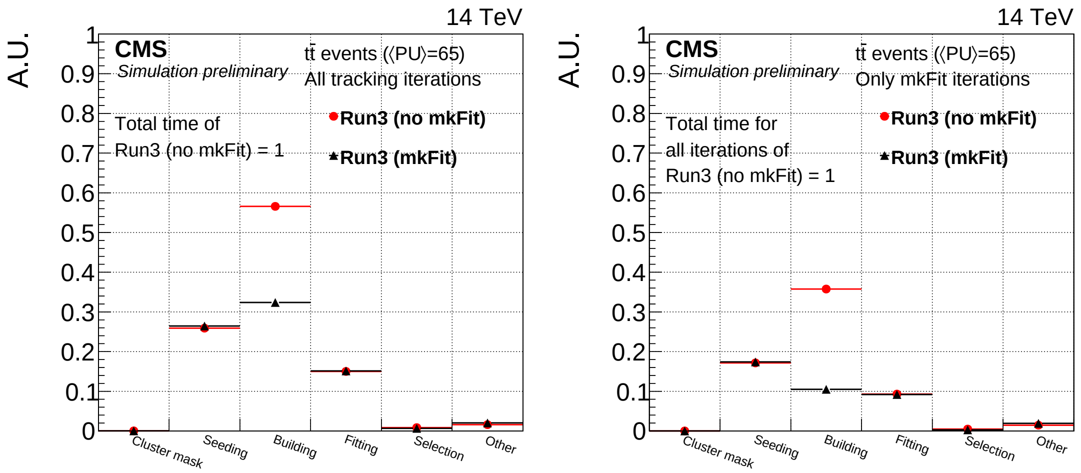
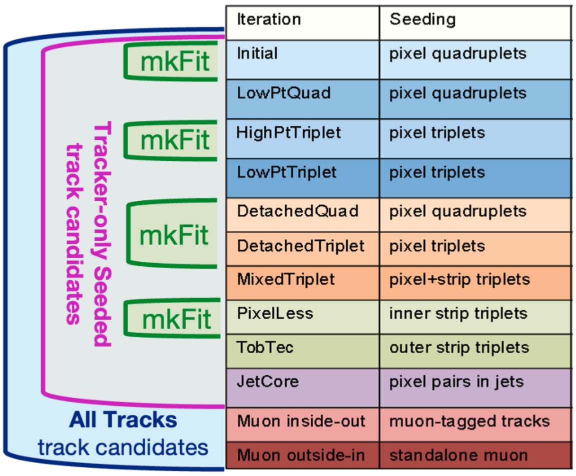


Figure 4.1: Comparison of tracking times for the different steps of the CMS track reconstruction sequence between the standard CKF algorithm used in Run 2 and mkFit, considering all iterations (on the left) and only the iterations for which mkFit is implemented (on the right). From [112].



Iteration	Seeding
Initial	pixel quadruplets
LowPtQuad	pixel quadruplets
HighPtTriplet	pixel triplets
LowPtTriplet	pixel triplets
DetachedQuad	pixel quadruplets
DetachedTriplet	pixel triplets
MixedTriplet	pixel+strip triplets
PixelLess	inner strip triplets
TobTec	outer strip triplets
JetCore	pixel pairs in jets
Muon inside-out	muon-tagged tracks
Muon outside-in	standalone muon

Figure 4.2: Table with the different iterations used in the CMS tracking algorithm as of Run 3. The table contains the information used for the seeding step in each iteration. The table also indicates the kind of track candidates reconstructed in the different iterations and which iterations implement the mkFit algorithm.

these two also use the information from the CMS muon spectrometer: the first iteration performs muon reconstruction inside-out, with the second doing an outside-in pass.

Figure 4.3 shows the track reconstruction efficiency, that is, the probability of reconstructing a track belonging to a charged particle, and the fake rate, that is, the probability of reconstructing tracks that are not associated with any real charged particle, as a function of p_T and η, for 2016 and 2017, based on simulation. The first plot shows the contributions of each of the iterations to the total efficiency. The year 2016 is the reference for the measurement of the W boson mass, with 2017 shown to illustrate the changes in performance due to the new pixel detector and updates to tracking software. In this simulation all general tracks are considered: for these it can be seen that the efficiency has a plateau at around 90% for $1\ \mathrm{GeV} < p_T < 100\ \mathrm{GeV}$, at lower p_T the performance is degraded due to energy loss and at higher p_T due to the strong interactions of pions with the silicon and because a significant fraction of these tracks are produced inside jets, an environment that makes the seeding inefficient. The performance is mostly flat in η in the barrel and decreases at higher η. In spite of the fact that the tracker acceptance is $|\eta| < 2.5$, tracks with $|\eta|$ up to 3 can be reconstructed in the silicon tracker: this is because the interaction region has a longitudinal extension that can reach approximately 20 cm, and the extension of the pixel detector is 50 cm, so that tracks produced in vertices which are not at the center of the pixel detector can still be reconstructed even if their η would appear to be outside of the nominal tracker acceptance. For 2017 it can be seen that the reconstruction efficiency is higher and that the fake rate is lower, especially at higher values of $|\eta|$. This is due to modifications in the seeding to take into account the new geometry and the implementation of the Cellular Automaton based reconstruction, the overall increase in the number of iterations for 2017, and the possibility to apply tighter cuts to reduce the fake rate.

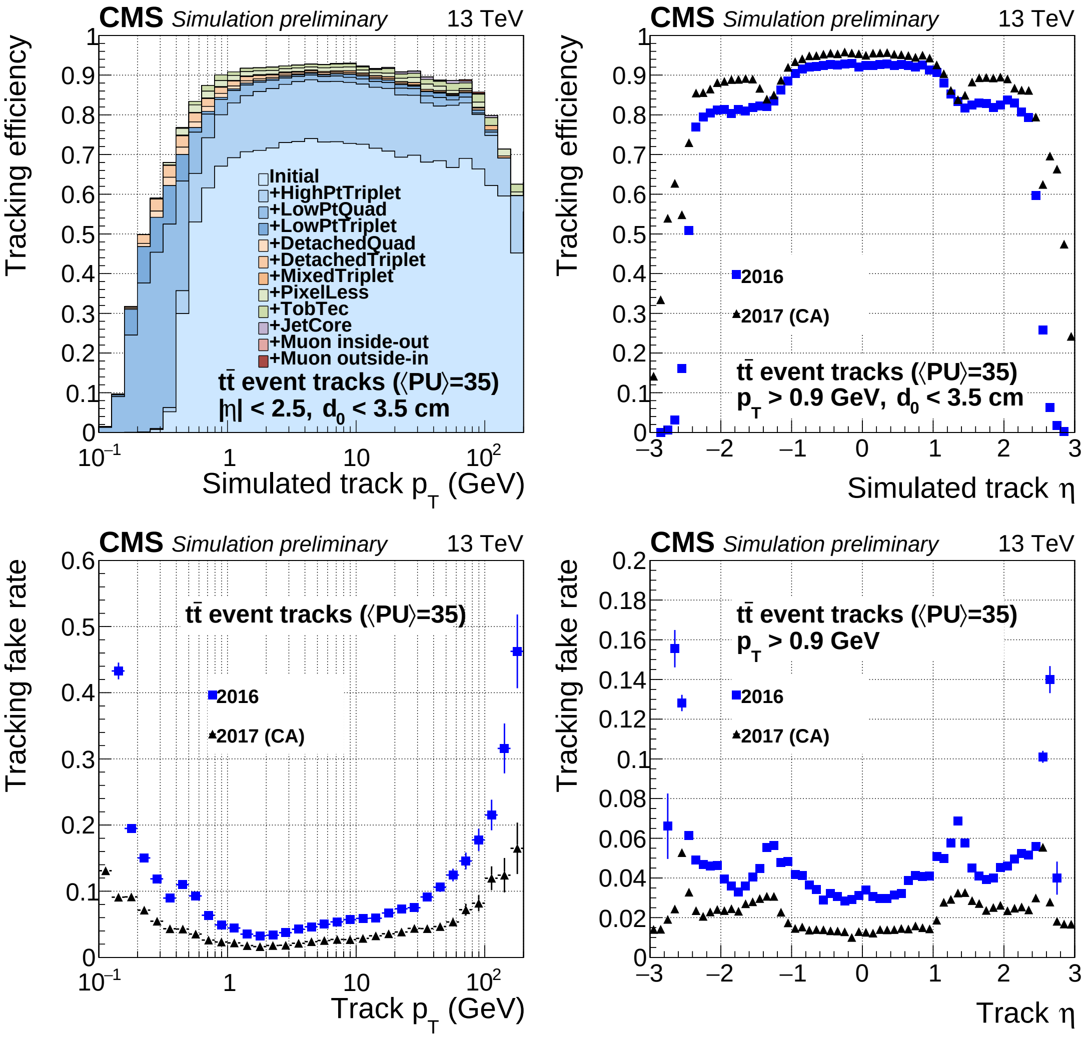


Figure 4.3: Tracking efficiency as a function of track p_T for the 2017 reconstruction, with a breakdown of the different contributions of the various iterations to the total efficiency shown in different colours (top left). Tracking efficiency as a function of track η (top right), and fake rate as function of p_T (bottom left) and η (bottom right), comparing 2016 and 2017 configuration. The efficiencies and fake rates are estimated from simulated $t\bar{t}$ events with $< \mathrm{PU} >= 35$, for high-purity tracks that pass the selection indicated in the figures. From [113].

Figure 4.4 shows the same distributions as in figure 4.3, but in this case the simulation is run at a center-of-mass energy of 14 TeV, and $< \mathrm{PU} >= 65$, against the $< \mathrm{PU} >= 35$ used for figure 4.3, more in line with the average pileup values of Run 3. A comparison between the standard algorithm used in Run 2 based on CKF and the new algorithm based on mkFit is shown in the figures. Regarding the efficiencies, the level of performance is similar to that of 2017 for both algorithms, with a possible slight degradation due to the higher pileup in the simulation. However, it can be noticed that the fake rate is lower when mkFit is used in the pattern recognition step.

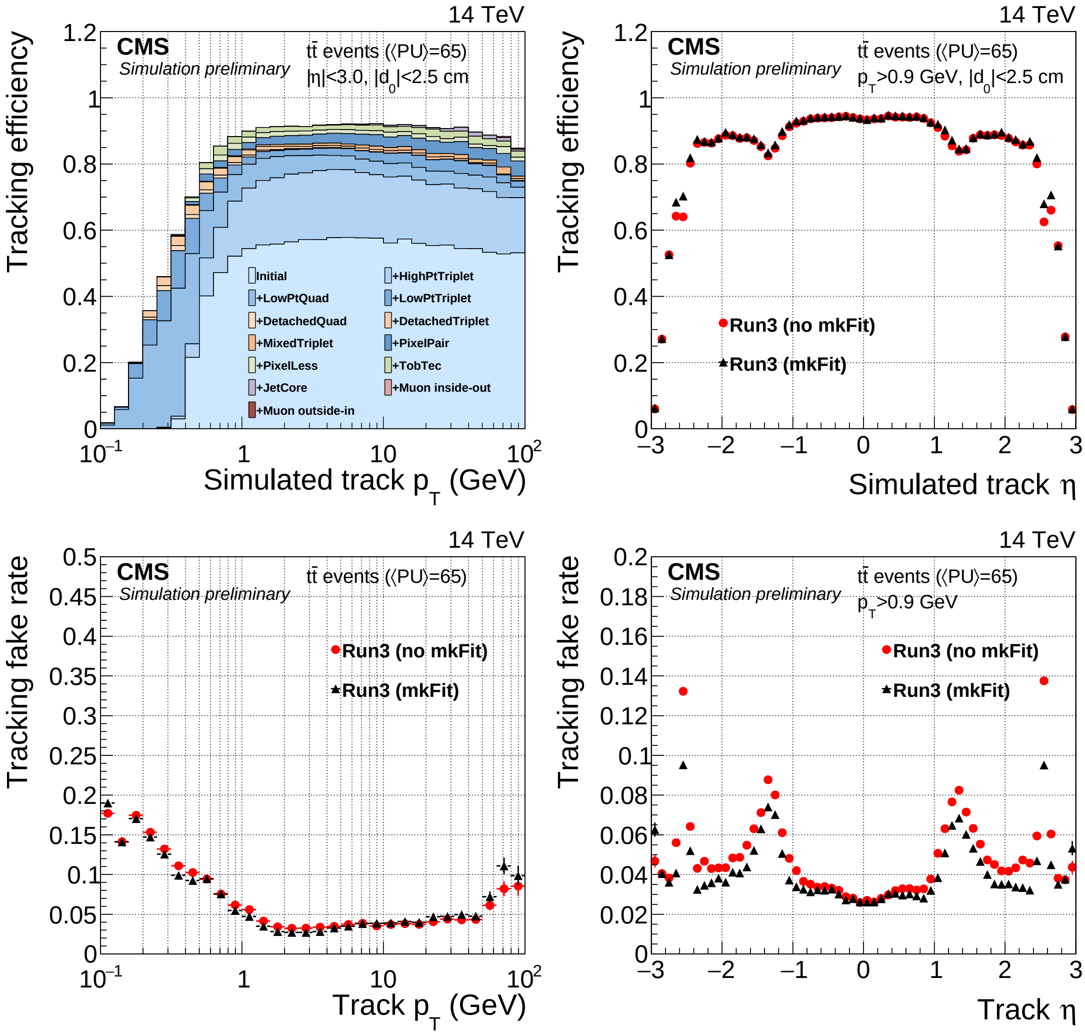


Figure 4.4: Tracking efficiency as a function of track p_T for the Run 3 configuration, with a breakdown of the different contributions of the various iterations to the total efficiency shown in different colours (top left). Tracking efficiency as a function of track η (top right), and fake rate as function of p_T (bottom left) and η (bottom right), comparing the track reconstruction performed using the standard CKF algorithm used in Run 2 to the new mkFit algorithm (all iterations, also those for which mkFit was not implemented, are considered). The efficiencies and fake rates are estimated from simulated $t\bar{t}$ events with $< \mathrm{PU} >= 65$, for high-purity tracks that pass the selection indicated in the figures. From [112].

What has been described so far is the *offline tracking*, i.e. the reconstruction of tracks after the selected events have been stored for reconstruction and analysis. At the High-Level Trigger, tracks are reconstructed to select that events have to be stored for future analysis using a similar approach as in offline tracking, using an iterative approach also in this case.

For Run 2, the general track reconstruction in the HLT consists of three iterations: the first two require four consecutive hits in the pixel detector to seed the tracking, targeting first high p_T and then lower p_T

tracks and use the full volume of the pixel detector. The third iteration relaxes the requirement on the number of hits in the track seeds to three and is restricted to the vicinity of jet candidates identified from calorimeter information and the tracks reconstructed in the previous iterations. An additional iteration was added during the technical stop between 2017 and 2018: this is because several issues with the Phase-1 pixel detector were identified leading to a non-negligible fraction of non-active pixel modules in each event. This was mostly due to the failure of DCDC converters used in the powering of the detector. During the technical stop the pixel detector was equipped with new converters and the initial performance was restored, but a new iteration was added to mitigate the effects of a potential re-occurrence of these issues [114].

For Run 3 this was replaced by one iteration that requires at least three hits in the pixel detector with $p_T > 0.3$ GeV. This single iteration is seeded by pixel tracks reconstructed by the Patatrack algorithm, which can be offloaded to GPUs [115, 116]. An additional iteration that requires at least two pixel hits, run after the first one is completed and the hits used in the previous iteration are masked, was added at the beginning of 2024 to mitigate the effects of the failure in the readout electronics of some modules in the third and fourth layer of the Barrel Pixel detector observed in 2023 [117].

After track reconstruction is completed, the reconstruction of *primary vertices*, i.e. the vertices associated to proton-proton collisions in a bunch crossing, is performed. High-quality tracks compatible with the primary interaction region are selected according to the significance on the transverse impact parameter with respect to the beam spot position, the number of pixel and strip hits associated to the track and normalized χ^2. These are then clustered according to their $z-$coordinate at the point of closest approach to the beam spot position. The clustering algorithm is called *deterministic annealing* [118], which finds the minimum for a problem with many degrees of freedom. The vertex positions are fitted using the *adaptive vertex fitter* [119], which estimates the x, y and z coordinates of the vertices, weighting the tracks according to their probability of belonging to a specific vertex. The vertex corresponding to the interesting collision, and not to pileup, is determined by clustering the tracks belonging to a given vertex into jets using the anti-k_T algorithm used as default in jet reconstruction [120]. The vertex with the highest sum of p_T^2 of the jets returned by the clustering procedure just described is selected as the main primary vertex.

4.2 Validation of tracking performances in Run 3

This section describes the work I performed in the context of the CMS Tracking Physics Object Group (Tracking POG). The work consisted in performing data/Monte Carlo comparisons to assess the tracking performance in Run 3, as a function of time. This work is essential for:

- understanding the level of agreement between data and MC, and supporting the tuning of the different ingredients that enter in data and simulation reconstruction, like tuning of MC generators (one big change in Run 3 with respect to Run 2 is that for the first time proton-proton collisions occur at a center-of-mass energy of 13.6 TeV);
- debugging of tracker conditions, for example what is the quality of alignment (how precisely the position of the silicon modules is known), or the presence of malfunctioning silicon modules;
- commissioning of the tracking software, as already mentioned the new mkFit algorithm was introduced in Run 3.

This work is very important to establish the quality of data-taking (for general use in CMS) and the prompt identification of problems in the detector or reconstruction. Part of this work has been presented at several conferences and is contained in CMS Detector Performance notes [121, 122, 123].

Results were derived using ZeroBias events, which are events triggered using only the information on the beam-beam collisions (no information from the CMS detectors is used), but also from events compatible with $Z \to e^+e^-$ and $Z \to \mu^+\mu^-$ decays. In the following, only the results on ZeroBias events will be shown. Only *general tracks* with $p_T > 1$ GeV that pass the high-purity flag [92] are included in these comparisons.

The first set of results is a comparison between the standard Run 2 track reconstruction based on CKF and the new reconstruction algorithm based on mkFit using data from proton-proton collisions at a center-of-mass energy of 900 GeV in 2021. In this case, no simulation is used. The aim of this studies is to check the performance of the new mkFit algorithm compared to the Run 2 baseline for the same events. Figure 4.5 shows the η and p_T distributions of the tracks, with p_T shown only for the tracks that have $|\eta| < 2.4$. The agreement for the η distribution is at the percent level, reaching $O(5\%)$ at high $|\eta|$, while the agreement for the p_T distribution is at the percent level for $p_T < 5$ GeV, while it worsens at higher values. In particular, for values at $p_T = 65$ GeV a sudden increase is observed: this was due to a bug in the initial versions of mkFit, which affected the reconstruction of tracks at high $|\eta|$. This increase is not seen if only tracks within $|\eta| < 2.4$ are considered. This is one of the issues that has been found in the context of these comparisons.

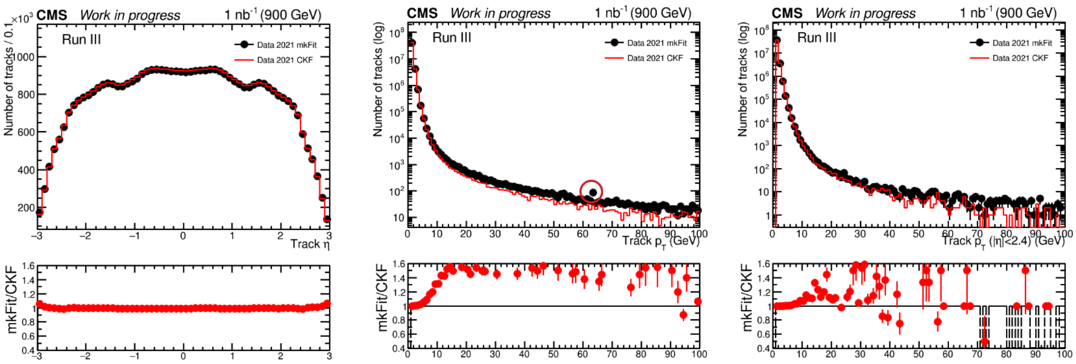


Figure 4.5: η (on the left) and p_T (in the middle, on the right only for $|\eta| < 2.4$) distributions for *general tracks* with $p_T > 1$ GeV for data collected in 2021 at a center-of-mass energy of 900 GeV, obtained both reconstructing tracks with the CKF algorithm and the mkFit algorithm. In the middle panel the bin at 65 GeV is highlighted: this was due to a bug fixed in the later versions of mkFit. From [124].

Figure 4.6 shows the distributions of the distance of closest approach with respect to the main primary vertex, the significance of the impact parameter (abbreviated as SIP) in two dimensions and in three dimensions for the tracks in the events. These variables are sensitive to the quality of alignment of the tracker, and are used in physics analyses such as those for b-physics, b/tau tagging, among other applications.

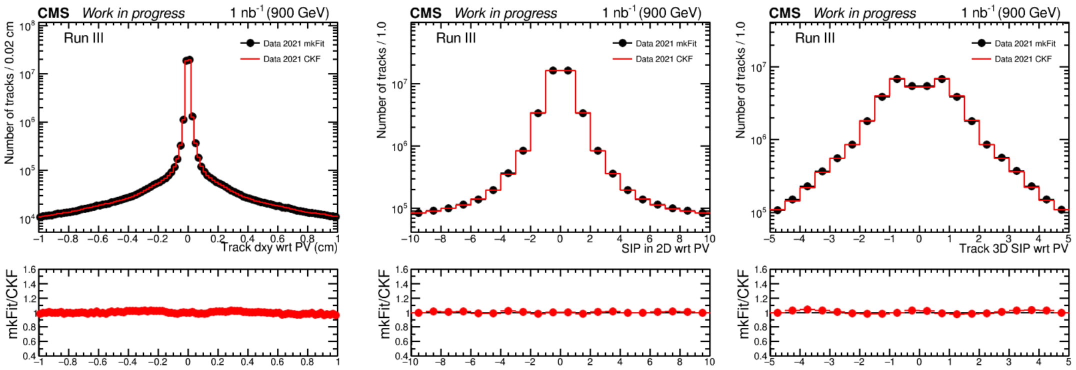


Figure 4.6: Distributions for the distance of closest approach (on the left), the significance of the impact parameter (abbreviated as SIP) in two dimensions (in the middle) and in three dimensions (on the right) for *general tracks* in ZeroBias events with $p_T > 1$ GeV for data collected in 2021 at a center-of-mass energy of 900 GeV, obtained both reconstructing tracks with the CKF algorithm and the mkFit algorithm. From [124].

Figure 4.7 shows the distribution for the number of hits in the pixel subdetector and in the strip subdetector. Regarding the pixel subdetector, the agreement on valid hits is at the 20% level, showing an increase in the number of tracks with 4 hits and a reduction in the number of tracks with fewer hits. This seems to agree well with the reduction in fake rate observed in figure 4.4: as the number of hits per track increases, the probability that these can be randomly assigned to a fake track decreases. Regarding the strip subdetector, the distribution of the number of hits is shifted towards lower values when mkFit is used: in later versions of mkFit several updates were implemented to increase the length of the tracks (and therefore increase the number of hits). Changes in the configuration of mkFit can modify these distributions and also the physics properties of the reconstructed objects.

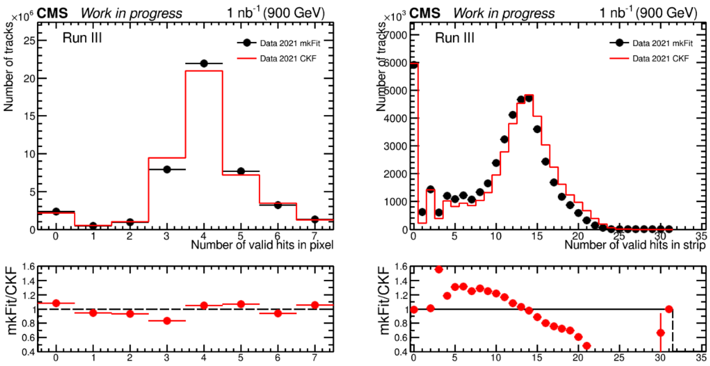


Figure 4.7: Distribution for the number of hits in the pixel subdetector (on the left) and in the strip subdetector (on the right) for *general tracks* in ZeroBias events with $p_T > 1$ GeV for data collected in 2021 at a center-of-mass energy of 900 GeV, obtained both reconstructing tracks with the CKF algorithm and the mkFit algorithm. From [124].

The second set of results is a series of data/simulation comparisons for ZeroBias events collected at a center-of-mass energy of 13.6 TeV collected during 2022. The tracks considered here pass the same selection as in the comparisons for 2021 data. These results are contained in [121]. Prior to the start of Run 3, the first layer of the barrel pixel detector was replaced to cope with the high pileup environment at the LHC. Indeed, effects related to radiation damage of this component were observed during 2022.

Figure 4.8 shows the distribution of the p_T and η of the tracks. Regarding the p_T distribution, the agreement between data and simulation is at the 10% level, while for η it can be observed that the distribution in simulation is wider than that in data. This can potentially be attributed to the insufficient tuning of the MC generator used in the simulation.

Figure 4.9 shows the distribution of the distance of closest approach to the main primary vertex for the tracks in ZeroBias events, for the periods of time indicated in the figures. Each of the figures also indicates the total integrated luminosity delivered since the installation of the new Barrel Pixel layer 1. It can be observed that the data/MC agreement worsens over time, indicating the aging of Barrel Pixel layer 1 due to accumulated irradiation. In the last period of data-taking shown in figure 4.9 the high-voltages of the Barrel Pixel layer 1 were increased to cope with the effects of radiation damage, and updates in the alignment were implemented, which could explain the better level of agreement observed.

Data collected during 2022 was reprocessed to include updates in the conditions of the detector after data-taking was completed. In particular, updates in the pixel local reconstruction and the alignment were implemented for the events shown in figure 4.10 (this updates were also extended to the rest of 2022 data), where it can be seen that the data/MC agreement on the SIP3D distribution improves significantly for the reprocessed events.

The third set of results is for ZeroBias events in 2023 [122]. The figures in 4.11 show the ϕ distribution for the two periods of time indicated. In the second period, readout problems were observed in layer 3 and 4 of the barrel pixel tracker both in the same $\phi-$sector on the negative half, with track coverage within

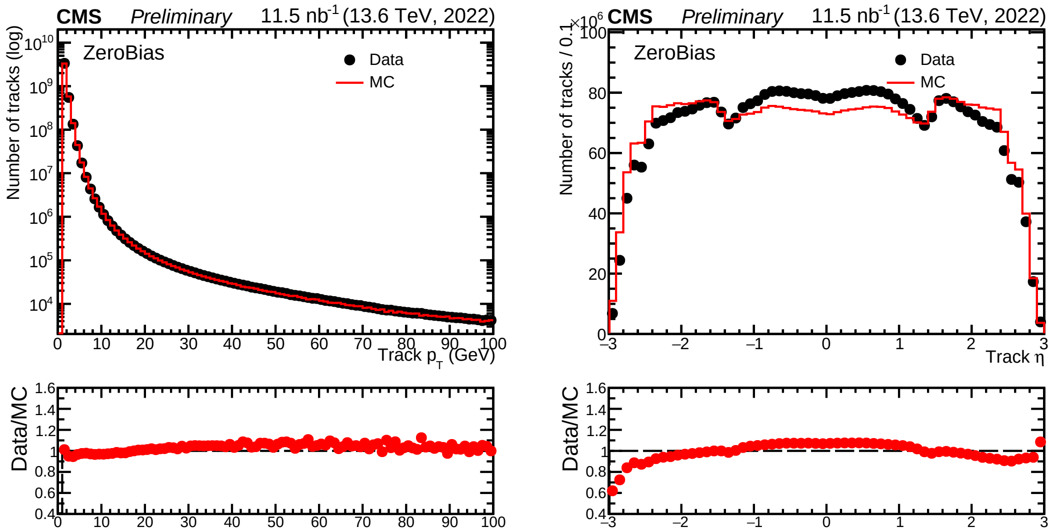


Figure 4.8: p_T and pseudorapidity (η) distributions of the tracks with $p_T > 1$ GeV from ZeroBias events collected in 2022 at a center-of-mass energy of 13.6 TeV. From [121].

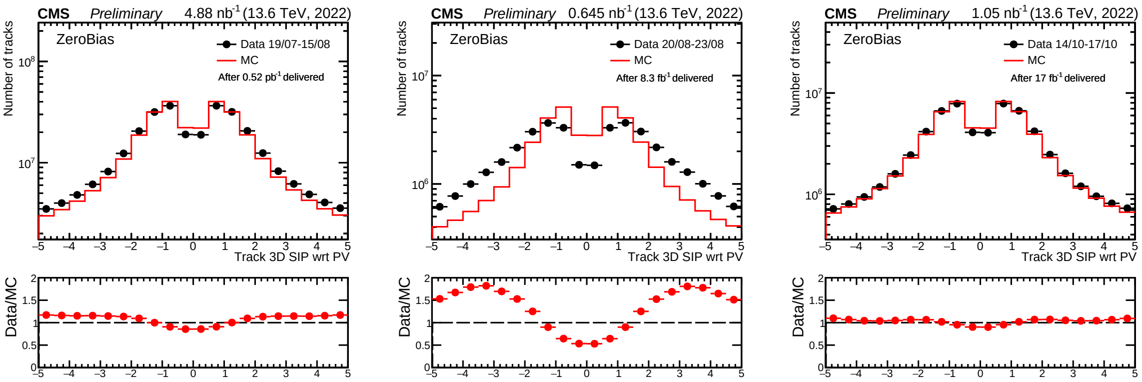


Figure 4.9: Distribution of the distance of closest approach to the main primary vertex of tracks with $p_T > 1$ GeV from ZeroBias events collected in 2022 at a center-of-mass energy of 13.6 TeV, for the periods of time indicated in the figures. From [121].

$-1.5 < \eta < -0.2$ and $-1.1 < \phi < -0.9$. The Monte Carlo sample used for the second period of time seems to reproduce the failure at the 5% level in ϕ.

Figure 4.12 shows the distribution of the distance of closest approach for the periods of time indicated in the figures, with a dedicated comparison for the tracks affected by the failure of the readout electronics in Barrel Pixel layers 3 and 4 for the two periods of time together. In general, it can be observed that the distribution in data is broader than in simulation, indicating that the quality of alignment is better in the latter. When only tracks with $-1.5 < \eta < -0.2$ and $-1.1 < \phi < -0.9$ are considered, the distributions are broader after the failure has occurred, both for data and for simulation, indicating that the resolution has worsened due to the lack of information from layers 3 and 4. The distributions in data are, however, still broader than the corresponding ones in simulation.

The last results is for ZeroBias events collected in 2024 [123]. During the 2024 data-taking, some channels in 1st disk of FPix were blacklisted, corresponding to a track coverage within $\eta < -1.4$ and $\phi > 2.5$, referred to as FPix hole. Figure 4.13 shows the distributions of the distance of closest approach and of the significance of 3D impact parameter with respect to the primary vertex, right before and right after the FPix failure, for the tracks in the FPix hole. It can be seen that the distributions in data are wider after the FPix

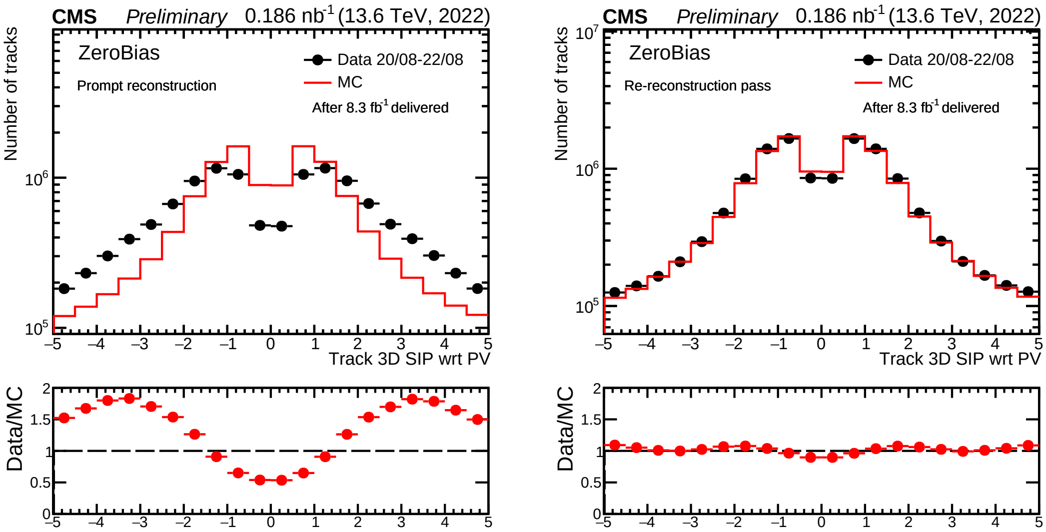


Figure 4.10: Distributions of the significance of 3D impact parameters with respect to the main primary vertex of tracks with $p_T > 1$ GeV from ZeroBias events collected in 2022 at a center-of-mass energy of 13.6 TeV. Only the events for which the re-reconstruction, which includes updates to pixel local reconstruction and the alignment of the tracker, are considered here. From [121].

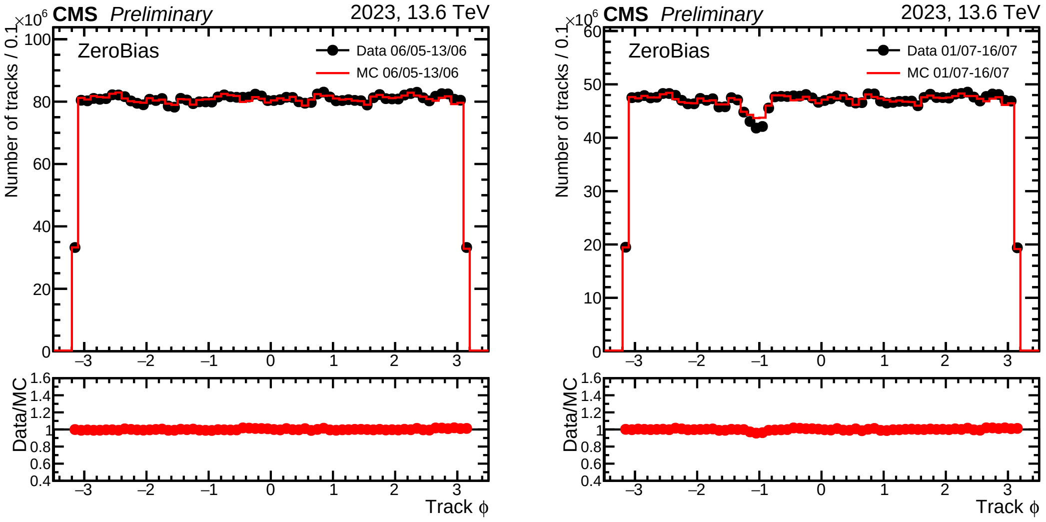


Figure 4.11: Distributions of the azimuthal angle ϕ of the tracks with $p_T > 1$ GeV from ZeroBias events collected in 2023 at a center-of-mass energy of 13.6 TeV, in the periods of time indicated in the figure. From [122].

failure, indicating a worsening of the resolution due to the lack of information in the first disk of FPix. The distributions are compared to a Monte Carlo simulation that models the effects from the FPix failure. It can be seen that the agreement is better for the data before the FPix failure: this can be due to the fact that quality of alignment is higher in simulation than in data. This might compensate for the loss of information from the first FPix disk, reducing its impact on track reconstruction, while on data the effects from the poorer quality of alignment are compensated from the extra information. This is not the case after the FPix failure, where the extra information is also missing.

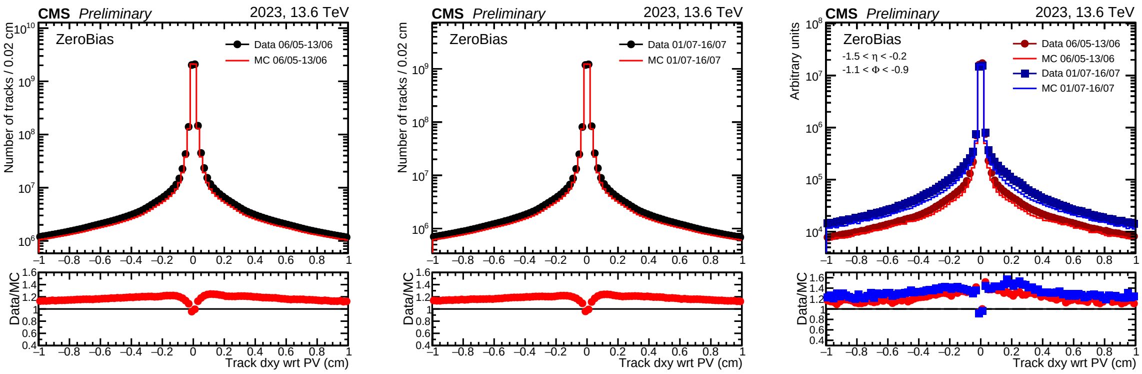


Figure 4.12: Distributions of the distance of closest approach to the primary vertex for the tracks with $p_T > 1$ GeV from ZeroBias events collected in 2023 at a center-of-mass energy of 13.6 TeV: for the first (left) and the second (middle) period of 2023 in the full geometric range, with a comparison of the two periods in the pixel tracker region affected by the readout failure described in the text, $-1.5 < \eta < -0.2$ and $-1.1 < \phi < -0.9$ (right). From [122].

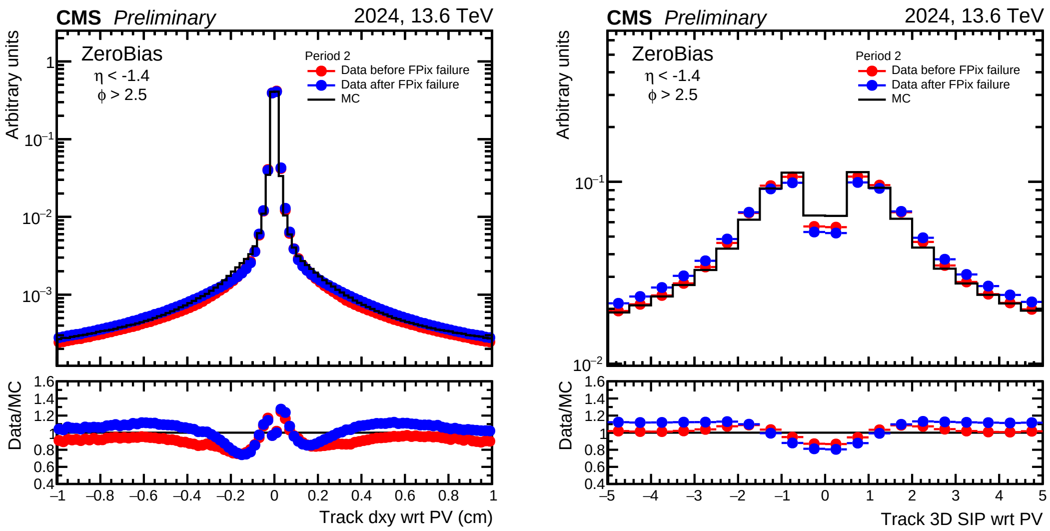


Figure 4.13: Left: Distributions of the distance of closest approach to the primary vertex. Right: Distributions of the significance of 3D impact parameter, abbreviated as 3DSIP, with respect to the primary vertex. In both cases, the tracks in the FPix hole are considered. The data distributions are shown right before and right after the FPix failure indicated in the text, and are compared to a simulation in which the effect from the FPix failure is modelled. From [123].

Chapter 5

Muon reconstruction and identification in CMS

This chapter provides a detailed description of the muon reconstruction and identification in CMS, and its importance in relation to the measurement of the W boson mass. The second part of the chapter (section 5.2) contains original studies that I performed in the context of the W boson mass measurement regarding the validation of the procedure employed to measure the muon reconstruction and identification efficiencies. As a result of this work, the default CMSSW framework was modified to fix bugs in the central reconstruction. It was also modified to introduce physical variables, such as a new definition of isolation, which are more suitable than those centrally computed for measurements of the W boson mass with a target uncertainty of 10 MeV.

5.1 Muon reconstruction in CMS

High-p_T muons are very important for the CMS physics program. They are the only particles that can travel through the whole CMS detector, as can be seen in figure 5.1, and leave a clear signature in the muon chambers. They cannot be easily produced in the soft QCD interaction that dominate the proton-proton collisions at the LHC, providing a clear signature of electroweak processes described by the Standard Model, or potentially serving as a probe for Beyond the Standard Model interactions. Because of this, they are used by the trigger system to select interesting events.

The measurement of the mass of the W boson by the CMS experiment presented in this thesis, described in more detail in section 7, revolves around the reconstruction of decays of the W boson into a muon and a neutrino, and the mass is inferred from the reconstruction of the transverse momentum of the muon p_T^μ. Figure 5.2 shows the distribution of the p_T of muons from decays observed in CMS in 2016. The impact on the p_T spectrum a mass variation of 10 MeV is also shown for comparison. This is the order of uncertainty that CMS aims at reaching, for the reasons discussed in section 2.3. It can be observed that the muons from W boson decays have $p_T \approx 40$ GeV, and that a precision in the muon reconstruction and identification better than 10^{-3} needs to be achieved to meet the goal of a 10 MeV uncertainty.

Muon reconstruction in CMS occurs independently in the CMS tracker and the muon chambers [98]. Muons reconstructed by the CMS tracker benefit from the reduced amount of material, which is designed to reconstruct muons with low p_T (down to 0.5 GeV, but with total momentum $p > 2.5$ GeV), but its performance is limited by the relatively short lever arm (4 Tm), while reconstruction in the muon chambers benefits from the larger lever arm when the constraint on the beamspot is used (12 Tm). Muons reconstructed by the tracker only are referred to as *tracker muons*, and include the information from only a few segments in the muon chambers for identification purposes (the geometrical matching between the track and at least segment in the DT and CSC is required). Muons reconstructed in the muon chambers are referred to as

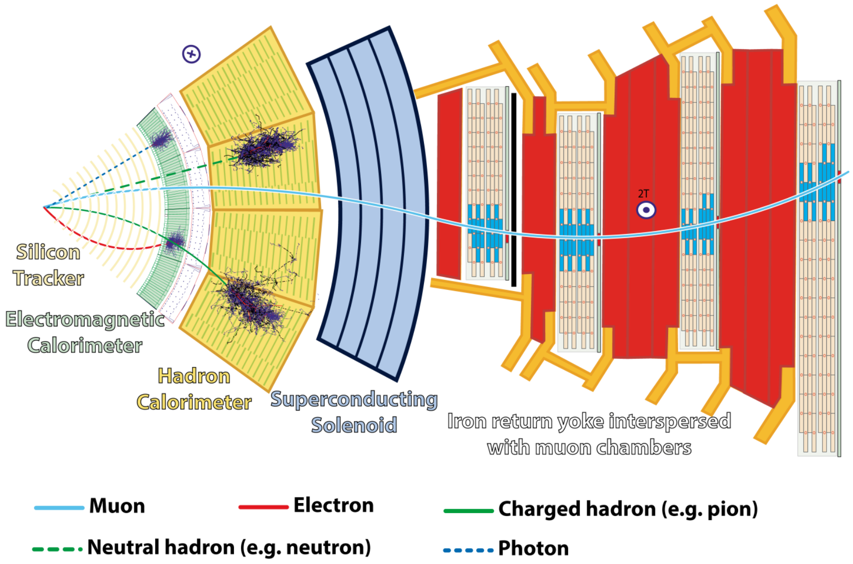


Figure 5.1: A cross-section of the CMS detector. The parts of the detector in which particles typically deposit their energy are also depicted. From [125].

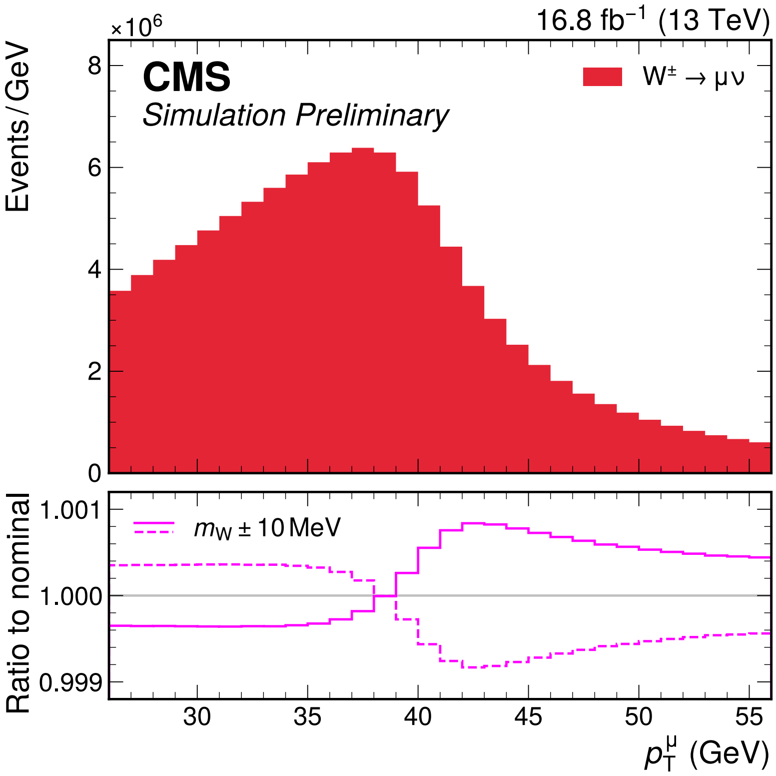


Figure 5.2: Distribution of the p_T^μ of the muon from the W boson decays. The bottom panel shows the impact of a 10 MeV variation on the p_T^μ distribution. From [1].

standalone muons: reconstruction begins with segments identified in each muon station, which are subsequently combined. In the barrel region, the magnetic field is almost entirely confined within the return yoke, so it can be assumed that muons propagate in a straight line inside the chambers. Based on this, the muon trajectory is reconstructed independently in the $r - \phi$ and $r - z$ planes. The two-dimensional measurements, each providing a coordinate and an angle, are then merged. In contrast, in the endcap region, the two-dimensional hits from all six layers of the endcap stations are combined to form track segments. Segment combination proceeds by propagating the muon track from one segment to the next, accounting for the magnetic field and multiple scattering effects in the iron. Finally, parameter estimates are combined if they are consistent with each other. Muons can be then reconstructed starting from standalone muons and propagating them backwards to find a matching *general track* within $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 1$, and a track fit using all the hits and segments is run to obtain the kinematic properties. These muons are referred to as *global muons*. Figure 5.3 shows the p_T resolution as a function of p_T using only the information from the inner tracker (in red) and including also the information from the muon spectrometer (in black). For p_T up to $O(100)$ GeV, the resolution is very similar between both reconstructions, with an improvement in resolution for $p_T > 200$ GeV when the information from the muon spectrometer is included. Therefore, global muons benefit from the highest possible muon momentum resolution for the range of p_T of interest for the m_W analysis performed by CMS, and also from the identification capabilities of the muon spectrometer.

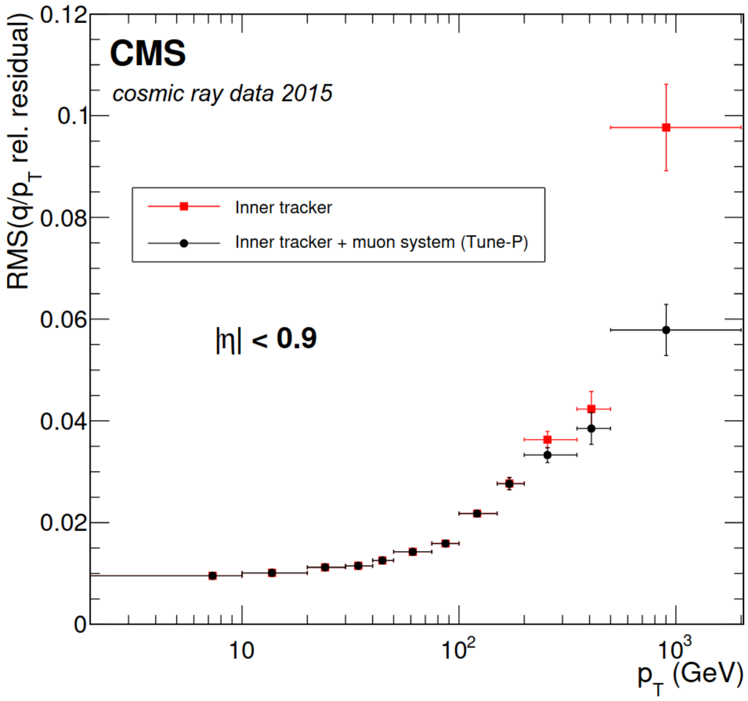


Figure 5.3: p_T resolution as a function of p_T for muons reconstructed using only the inner tracker information (in red) and those using also the information from the muon spectrometer (in black). From [98].

Muons considered in the m_W analysis are required to pass additional criteria to discriminate between muons produced in electroweak processes, such as decays of W, Z, Higgs decays, which can be usually modeled by Monte Carlo simulation with high precision, and those produced in QCD interactions, which are instead prefentially estimated with data-driven techniques. Background estimation is described in more detail in section 7.7. Isolation is a criterion that provides high discrimination power with respect to QCD events. It is defined as the sum of the p_T of all the charged hadrons, neutral hadrons, and photons in a cone centered around the muon direction of flight with radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} < 0.4$. The particles used to define muon isolation are reconstructed by the Particle Flow algorithm [100] described in 3.2.1. The isolation variable is then computed as:

$$\mathrm{Iso}_{\mathrm{PF}} = \sum_{\mathrm{PF\ char.\ had.\ from\ PV}} p_T + \max\Big(\sum_{\mathrm{PF\ neutr.\ had.}} p_T + \sum_{\mathrm{PF\ photons}} p_T - 0.5\times \sum_{\mathrm{PF\ char.\ had.\ from\ PU}} p_T, 0\Big) \quad (5.1)$$

where the different sums run over all the PF candidates within the isolation cone. For charged hadrons, the information from the inner tracker is available, and the vertex from which the track originated can be identified, but this is not the case for the neutral particles, so all neutral particles in the event need to be included. The PU contribution to the neutral component is estimated using the charged component from PU vertices, then subtracted from the neutral component, as shown in the last term in equation 5.1, the factor 0.5 was validated from simulation [100]. Most commonly, a cut on the relative isolation $\mathrm{RelIso_{PF}}$, $\mathrm{Iso_{PF}}$ in equation 5.1 divided by the p_T of the muon, is used, as muons from QCD have a softer p_T spectrum compared to those from electroweak processes. A cut on $\mathrm{RelIso_{PF}} < 0.15$ is used in this analysis, which returns a fraction of ≈95% for signal muons.

In addition to isolation, other quality criteria are defined to increase the purity of the muon selection:

- *Loose ID*: it requires a PF muon to be either a tracker or global muon, with the aim of reducing the misidentification of charged hadrons by ensuring that global or tracker muons deposit at most a small fraction of their energy in the calorimeters.
- *Medium ID*: it requires hits in 80% of the silicon layers and high compatibility with muon segment in the muon system with additional χ^2_{fit} quality selection. It aims at selecting prompt and heavy flavor decay muons.
- *Tight ID*: it requires the muon to be reconstructed both as tracker and global muon, with hits on at least 6 silicon layers, at least a pixel hit, two muon station compatibility, high χ^2_{fit} quality, and a transverse (longitudinal) impact parameter $|d_{\mathrm{xy}}| < 0.2$ cm ($|d_{\mathrm{z}}| < 0.5$ cm). Its aim is to suppress muons from decays in flight and hadronic punch-through (hadronic showers that are not fully contained in the hadronic calorimeters and reach the muon spectrometer).
- *Soft ID*: It aims to select low-p_T muons and requires the muon to be a tracker muon with high purity flag, at least 6 hits in the silicon tracker, one of which in the pixel detector, and tight muon system segment matching.
- *High-momentum ID*: It is optimized for muons with $p_T > 200$ GeV and requires the same conditions of the tight muon ID, but χ^2 request is relaxed, to take into account inefficiencies due to muon radiation in the detector material.

The analysis presented in this thesis uses the medium ID with the explicit additional requirement that the transverse impact parameter with respect to the beamspot ($|d_{\mathrm{xyBS}}|$) to be less than 0.05 cm to suppress muons from decays in flight. The reason why the tight ID is not used is because of the requirements on the impact parameters are applied with respect to the main primary vertex, selected as described in the previous chapter. The probability that this vertex is the one in which a W or Z boson is produced is shown in figure 5.4, as a function of the transverse momentum of the boson (p_T^V). It can be seen that, despite the fact that these probabilities are of $O(1)$, especially for the W boson this differs from 1 at the percent level, especially at low p_T^V.

For the Z boson, this difference is at the per-mille level: Z events are used to correct for muon reconstruction and identification efficiencies between data and simulation using the tag-and-probe procedure described in the following section. The corrections are then applied to W events assuming universality between W and tag-and-probe events. By using simulated events, it has been observed that this assumption introduces a difference that is not recovered by the tag-and-probe procedure, leading to a residual effect which is non-negligible for the target precision of this analysis.

This difference in vertex association probabilities has also an effect on $\mathrm{RelIso_{PF}}$, as the same selection algorithm is used in the default CMS reconstruction to separate between charged hadrons from the primary vertex of interest and those from pileup vertices. For this analysis, the default CMSSW framework has been modified to compute a "vertex-agnostic" definition of isolation, in which the PF charged hadrons are split

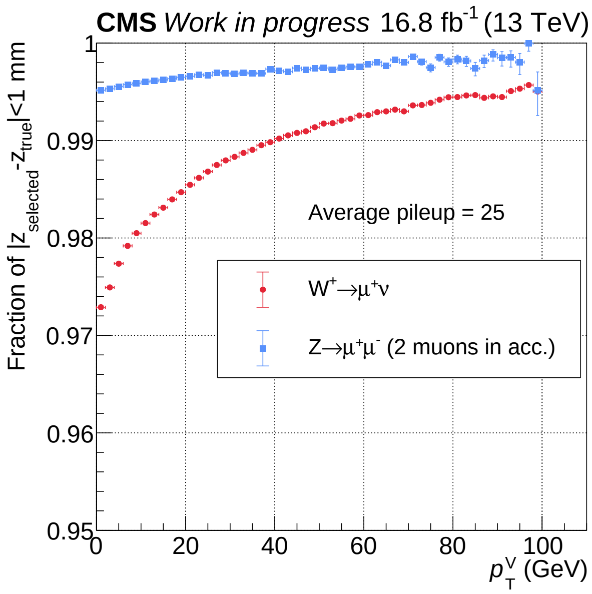


Figure 5.4: Fraction of the events in which the distance along the beam axis between the reconstructed vertex chosen as primary vertex by the standard CMS finding algorithm and the true vertex for the W boson (Z boson) events is less than 1 mm is shown in red (blue) as a function of the boson p_T. The average pileup for the data used in the analysis is indicated in the figure.

based on the longitudinal distance ΔZ with respect to the muon: charged hadrons within $|\Delta Z| < 2$ mm are considered as coming from the same vertex as the muon, while the others are considered as pileup particles. This is described in detail in section 5.2.4.

5.2 Precision studies for m_W

5.2.1 Validation of tag-and-probe efficiencies

It is crucial that the muon reconstruction and identification are well under control and reproduced by the Monte Carlo simulations, otherwise this might lead to negligible differences in the templates built to extract m_W, biasing the measured value. The muon reconstruction and identification efficiencies, i.e. the probability that a muon is reconstructed by the CMS detectors using the described reconstruction algorithms and that it passes a specified set of identification criteria, are corrected in simulation to match those observed in data.

Efficiencies in data are measured using the tag-and-probe method [126]: dimuons from a resonance, the Z boson in this case, are selected using strict requirements on one lepton, called the "tag" lepton, and with relaxed selection on the second lepton, called the "probe" lepton. This gives a well-defined shape that can be easily fitted in data, which helps separating from background processes. The probe muons are divided into "passing" and "failing" according to whether or not they pass the selection for which one wants to measure the efficiency. Invariant mass distributions obtained from the tag muon and passing probe and the tag muon and the failing probes are fitted to extract the number of muons that pass or fail the selection. The efficiency is then derived as:

$$\epsilon = \frac{\#\ \text{passing probes}}{\#\ \text{passing probes} + \#\ \text{failing probes}} \tag{5.2}$$

This is done in multiple reconstruction and identification steps:

$$\epsilon_{\text{tot}} = \epsilon_{\text{RECO}} \times \epsilon_{\text{Tracking}} \times \epsilon_{\text{IDIP}} \times \epsilon_{\text{Trigger}} \times \epsilon_{\text{Isolation}},$$

where:

- RECO is the efficiency corresponding to the reconstruction in the muon spectrometer, this is the probability that a muon is reconstructed as a standalone muon;
- Tracking is the probability that a standalone muon is also reconstructed as a global muon (it is in principle similar to just requiring that a track is present in the silicon tracker, but the fact the global muon reconstruction needs to be successful introduces some differences);
- IDIP is the probability that the muon passes the medium ID and has transverse impact parameter with respect to the beam spot $|d_{\text{xyBS}}| < 0.05$ cm;
- Trigger is the probability that the muon is matched to an object returned by the `HLT_IsoMu_24` or `HLT_IsoTkMu_24` trigger paths, both of which return an isolated muon with $p_T > 24$ GeV, but the first performs the muon reconstruction outside-in starting from the muon spectrometer, while the other performs the reconstruction inside-out starting from the silicon tracker;
- Isolation is the probability that the muon has $\text{RelIso}_{\text{PF}} < 0.15$.

Each step is evaluated on top of the previous steps, so that the overall muon selection is tighter at each step and the probability is properly factorized. This is done in steps because, as will be described in section 7.7, the isolation variable is used for background estimation by using control regions defined by reversing the isolation selection, while keeping the rest of the selection the same, but also to ensure control over the reconstruction in the various subdetectors (RECO and Tracking steps) and the impact of the identification requirements (IDIP, Trigger, and also Isolation). In addition to this, since we are using global muons, we need to evaluate the reconstruction efficiency for both the muon spectrometer and the inner silicon tracker, but there are no other detectors in which muons are reconstructed, so we need to use muons reconstructed in the inner tracker and muons reconstructed in the muon spectrometer alternatively to define the numerator and the denominator in the efficiency definition.

The efficiencies are measured as a function of η and p_T, with an unprecedented number of bins, especially in η: this is because, as will be outlined in chapter 7, the addition of the η information imposes significant *in-situ* constraints on many of the systematic uncertainties that have historically limited the precision on m_W.

The ideal definition of efficiency in the analysis is:

$$\epsilon_{\text{tot}} = P(\text{passes RECO, Tracking, IDIP, Trigger, Isolation}|\text{muon}) \tag{5.3}$$

which in simulation would correspond to just computing the fraction of muons that pass the total selection. What is done in tag-and-probe is instead different:

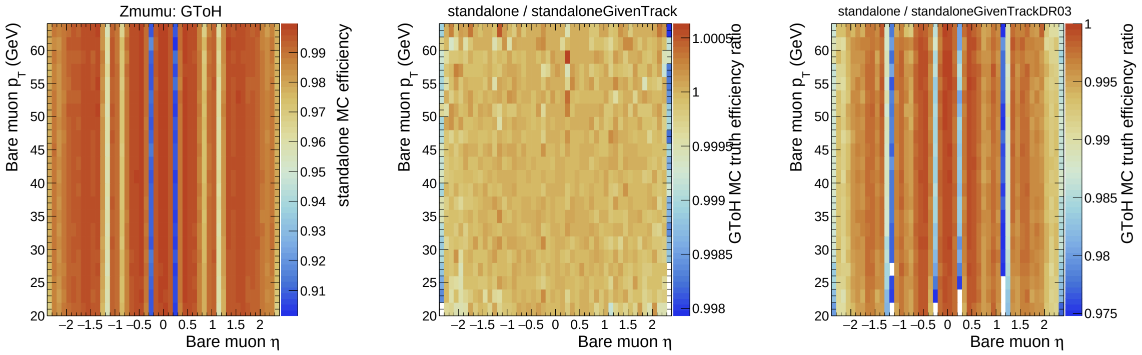
$$\begin{aligned}\epsilon_{\text{tot}} = {} & P(\text{is Isolated}|\text{passes Trigger}) \times P(\text{passes Trigger}|\text{passes IDIP}) \times P(\text{passes IDIP}|\text{is global}) \times \\ & P(\text{is global}|\text{standalone}) \times P(\text{matched to standalone muon}|\text{track}),\end{aligned} \tag{5.4}$$

and this might introduce biases in the corrections that are applied to simulations to reproduce the data.

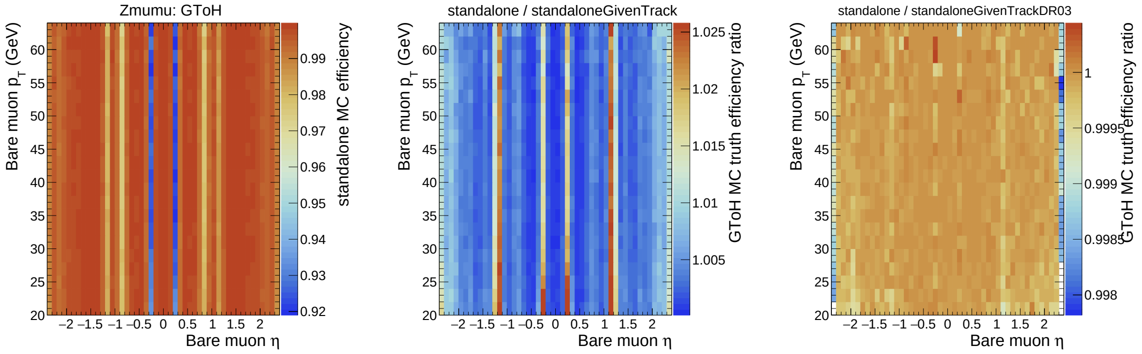
Apart from the definitions in equations 5.3 and 5.4, another difference arises from the fact the requirement of a selection on tag muon, which is not used to compute the efficiencies, modifies the overall phase space of events that are used to compute the efficiencies, and this introduces an additional source of biases if not taken into account. We will see that this is especially true for the Trigger and Isolation steps.

The validity of the tag-and-probe procedure was validated by comparing the efficiencies computed as in equation 5.3 with the efficiencies computed as in equation 5.4, using simulated $Z \to \mu\mu$ events. First a set of comparisons that uses the tag-and-probe definition of efficiencies but that does not actually require the presence of a tag muons is made, then tag requirement is introduced. The matching between generator and reconstructed muon is done by minimizing $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ and requiring it to be less than 0.1, although in some comparisons the case in which ΔR is extended to 0.3 is studied.

Figure 5.5 shows the RECO efficiency, i.e. the probability to reconstruct a standalone (SA) muon matched to a muon after the emission of final-state radiation (postFSR muon) at generator-level, as a function of the generator muon $p_T - \eta$, which is referred to as MC truth efficiency. This is compared to the efficiency of reconstructing a SA muon given a tracker-seeded track, that is a *general track* that is not seeded by the iterations that use the information from the muon spectrometer, which is the same definition used with by tag-and-probe. These efficiencies and ratios are computed using different ΔR for the matching between standalone muon and generator muon and between standalone muon and track: the radius for generator-level matching is indicated in figure 5.5 as a subscript (for example in $P(\mathrm{SA}|\mathrm{track})_{0.1}$ 0.1 refers to the ΔR to match the generator-level muon to the track in the denominator), while the radius to define the matching to the standalone muon is explicitly written. Looking at the ratio plot, it can be seen that the disagreement is within 0.1%, when using consistent ΔR values between objects, while the ΔR distance for generator-level matching is not relevant. The plots in figure 5.5 demonstrate that the approximation of RECO efficiency as $P(\mathrm{SA}|\mathrm{track})$ is well posed and correctly implemented.



(a) $P(\mathrm{SA}|\mathrm{gen})_{0.1}$ (b) $P(\mathrm{SA}|\mathrm{gen})_{0.1}/P(\mathrm{SA}|\mathrm{track})_{0.1}$, $\Delta R_{\mathrm{SA,track}} = 0.1$ (c) $P(\mathrm{SA}|\mathrm{gen})_{0.1}/P(\mathrm{SA}|\mathrm{track})_{0.1}$, $\Delta R_{\mathrm{SA,track}} = 0.3$



(d) $P(\mathrm{SA}|\mathrm{gen})_{0.3}$ (e) $P(\mathrm{SA}|\mathrm{gen})_{0.3}/P(\mathrm{SA}|\mathrm{track})_{0.3}$, $\Delta R_{\mathrm{SA,track}} = 0.1$ (f) $P(\mathrm{SA}|\mathrm{gen})_{0.3}/P(\mathrm{SA}|\mathrm{track})_{0.3}$, $\Delta R_{\mathrm{SA,track}} = 0.3$

Figure 5.5: RECO efficiency from MC truth with tag-and-probe definition, namely the probability $P(\mathrm{SA}|\mathrm{gen})$ to find a standalone (SA) muon matched to a postFSR muon (left), and ratios between $P(\mathrm{SA}|\mathrm{gen})$ and the efficiency $P(\mathrm{SA}|\mathrm{track})$ to reconstruct a standalone (SA) muon matched to a tracker-seeded track with $\Delta R_{\mathrm{SA,track}} = 0.1$ (middle) or 0.3 (right). Plots on top (bottom) have a gen matching cone of 0.1 (0.3). Good closure is obtained when using consistent $\Delta R_{\mathrm{SA,gen}}$ and $\Delta R_{\mathrm{SA,track}}$.

The effect of the choice of ΔR is clearly visible around the η bins corresponding to the holes of the muon spectrometer (figure 3.10): here the resolution is ΔR is worse than in the rest of the detector. This has an impact on the definition of global muons: the standard reconstruction of global muons requires only a very loose matching cone $\Delta R = 1.0$ between the standalone object and the track used to form the global muon candidate, but in many cases these are much closer. In particular, figure 5.6 shows the fraction of global muons that have ΔR between the standalone muon and track in the silicon detector (inner track) to be greater than 0.3: it can be seen that this percentage increases in the holes of the muon spectrometer, particularly around $|\eta| = 1.2$ where this fraction reaches the percent level, but smaller peaks at 0.5% can be also seen at $|\eta| = 0.2$. This ΔR variable is essentially the same as the one between the standalone muon and the tracker-seeded track in the RECO efficiency definition. The value $\Delta R < 0.3$ is chosen in the RECO efficiency definition as it should be loose enough to make resolution effects in ΔR negligible in most of the η space, but not too loose to match objects that are not related to the muon under consideration. In order to have proper efficiency corrections, global muons are considered in the analysis only if ΔR between inner track and standalone muon is less than 0.3.

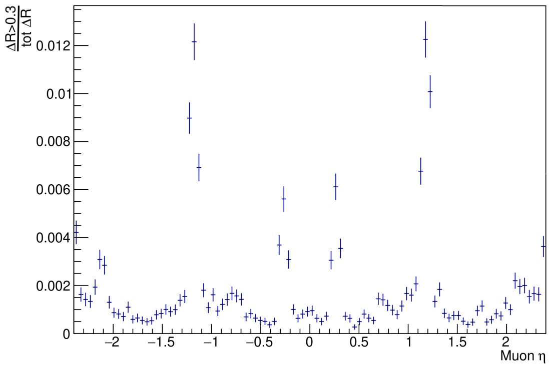


Figure 5.6: Fraction of muons, as a function of the postFSR muon η, where the ΔR between the inner and outer muon track is larger than 0.3.

With this new definition of global muons, the same procedure to compare MC truth efficiencies and tag-and-probe efficiencies is done also for the $\mathrm{Tracking}$ efficiency. In this case, the MC truth efficiency is simply:

$$\epsilon_{\mathrm{Tracking}}^{\mathrm{MC\ truth}} = P(\mathrm{global\ muon}|\mathrm{generator\ level\ muon}), \tag{5.5}$$

while the tag-and-probe efficiency is defined as:

$$\epsilon_{\mathrm{Tracking}}^{\mathrm{TnP}} = P(\mathrm{global\ muon}|\mathrm{SA\ muon}) \times P(\mathrm{SA\ muon}|\mathrm{track}). \tag{5.6}$$

For the MC truth efficiency the matching between postFSR muon and global muon is done using the inner tracker information, while for the tag-and-probe efficiency two matchings are performed separately, as the efficiency is the product of two steps: for $P(\mathrm{SA\ muon}|\mathrm{track})$ the matching between track and generator muon is performed as in the RECO step, while for $P(\mathrm{global\ muon}|\mathrm{SA\ muon})$ the matching between standalone muon and generator muon is performed using the standalone η and ϕ and $\Delta R_{\mathrm{postFSR,SA\ muon}} < 0.3$, while $\Delta R_{\mathrm{postFSR,track}} < 0.1$. The generator-level matching for the tag-and-probe efficiencies is bijective: for each generator-level muon, only one reconstructed muon can be matched (the one with the smallest ΔR), and vice-versa. This is because in this case the generator-level matching ensures that only the muons from the $Z \to \mu\mu$ decays are considered: when the actual tag-and-probe efficiencies are derived, the requirement on the presence of a tag muon (not required here yet) and the fact that the distribution of the invariant mass of the tag-and-probe pair is the one for $Z \to \mu\mu$ decays should be enough to ensure that is the case. The result of this comparison is shown in figure 5.7. Here the efficiencies are shown as a function of the postFSR muon p_T and η, for both MC truth and tag-and-probe, to avoid confusion from resolution effects. When efficiency corrections are derived with the tag-and-probe procedure, they are always given as a function of the inner track η and p_T for all steps except for $\mathrm{Tracking}$, where the standalone muon variables are used. It can be seen that the agreement is better than 0.1% except for $|\eta| \approx 0$, where the disagreement reaches 1%, and $|\eta| > 2$, where the disagreement reaches 0.5%.

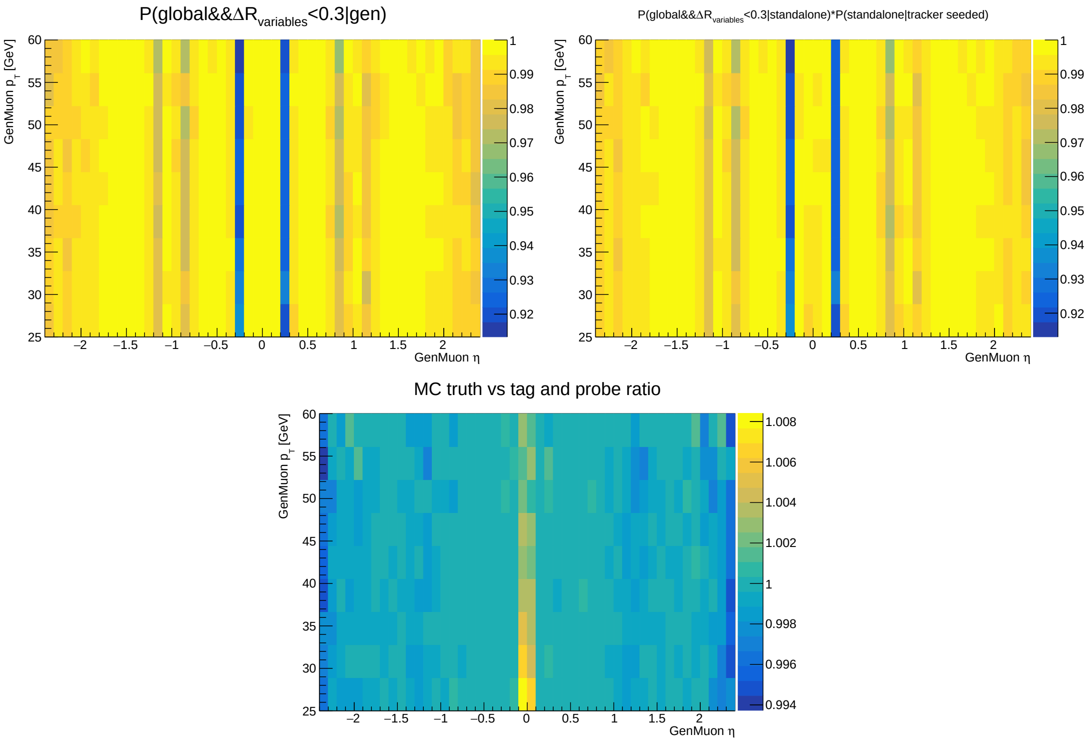


Figure 5.7: MC truth and tag-and-probe efficiency for the $\mathrm{Tracking}$ step, and ratio between the two, as a function of the postFSR muon η and p_T. More details are provided in the text.

The possibility for this disagreement to arise from the presence of muons that are copies of other existing muons (henceforth referred to as clone muons) was studied. In the tag-and-probe procedure, all possible combinations of muons found in the events are used to form pairs. The population of probes is augmented by the presence of clones, which might then lead to an overestimation or underestimation of the efficiencies depending on the properties of these clone muons. Figure 5.8 shows the fraction of global muons that have another standalone muon within $\Delta R < 0.3$. It can be seen that this percentage reaches the percent level around $|\eta| = 0$ and 2% at $|\eta| > 2$.

First, the additional muons around $\eta = 0$ were studies. For these, it was observed that they had a number of hits in the muon spectrometer identical to that of the global muon, and that the p_T derived using only the muon spectrometer was very similar, within resolution uncertainties (for the muon spectrometer when used alone, these are $O(5-10\%)$ in the barrel). It was also seen that these are generally very close to the original standalone muon, as it can be seen on the left of figure 5.9, so we concluded that these are not genuine muons, but clone muons. The right plot of figure 5.9 shows the distribution of the ΔR between the global muon and the closest standalone muon for $0.5 < \eta < 0.7$, where no clones should be present according to figure 5.8.

To actually verify that the disagreement is due to such clone muons, global muon efficiencies were derived for muons that have ΔR between the global muon and the closest standalone muon not used in the global muon reconstruction to be greater than 0.1, a procedure that we called ”clone cleaning”, both in the MC truth and tag-and-probe definitions. The results are shown in figure 5.10: it can be seen that the

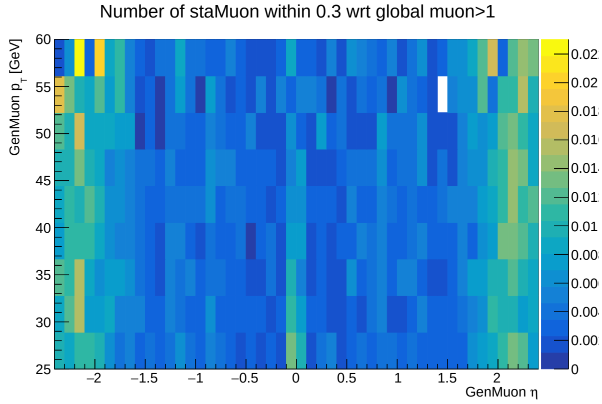


Figure 5.8: Fraction of global muons with another standalone muon within $\Delta R < 0.3$.

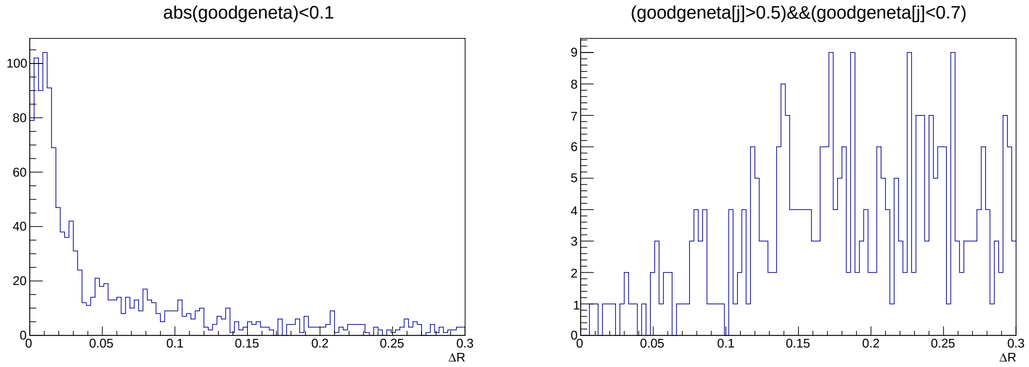


Figure 5.9: Distribution of ΔR between the global muon and the closest standalone muon not associated to the global muon, for $\eta = 0$, where the clones are expected to be found, and for $0.5 < \eta < 0.7$, where no clones should be present.

disagreement at $\eta = 0$ is reduced.

Careful studies of the standard muon reconstruction allowed us to find the cause of the presence of these clone muons, and it was concluded that this was due to a bug in the central reconstruction of standalone muons implemented in CMSSW. The bug-fixes to the central CMSSW code are described in section 5.2.2.

The additional muons at $|\eta| > 2$ were also thoroughly studied to understand their origin. Figure 5.11 shows the distribution of the standalone muon p_T of the clone muon versus the difference between the *best track* p_T and the standalone p_T. On the right of the same figure, the case in which the additional standalone muon is also a global muon is shown. The *best track* is the muon track (between standalone, inner track, global track) that is determined in the default muon reconstruction to best describe the muon properties and/or to have the highest quality in reconstruction. When only the standalone track is present, this is chosen as best track. Three components of the distribution can be observed:

- A peak at (0,0) can be observed in both plots: we associate this to genuine muons from pileup. This component, which we define as muons with $p_T < 6$ GeV, amounts to 51% of all events in the left plot.
- A structure at $\Delta p_T \approx 0$ can be observed in the plot on the left but not on the right: we associate this to muons that are standalone but not global. This component amounts to 27% of all events in the left plot.
- A diagonal with negative slope (meaning that the standalone p_T is larger than the best track p_T);

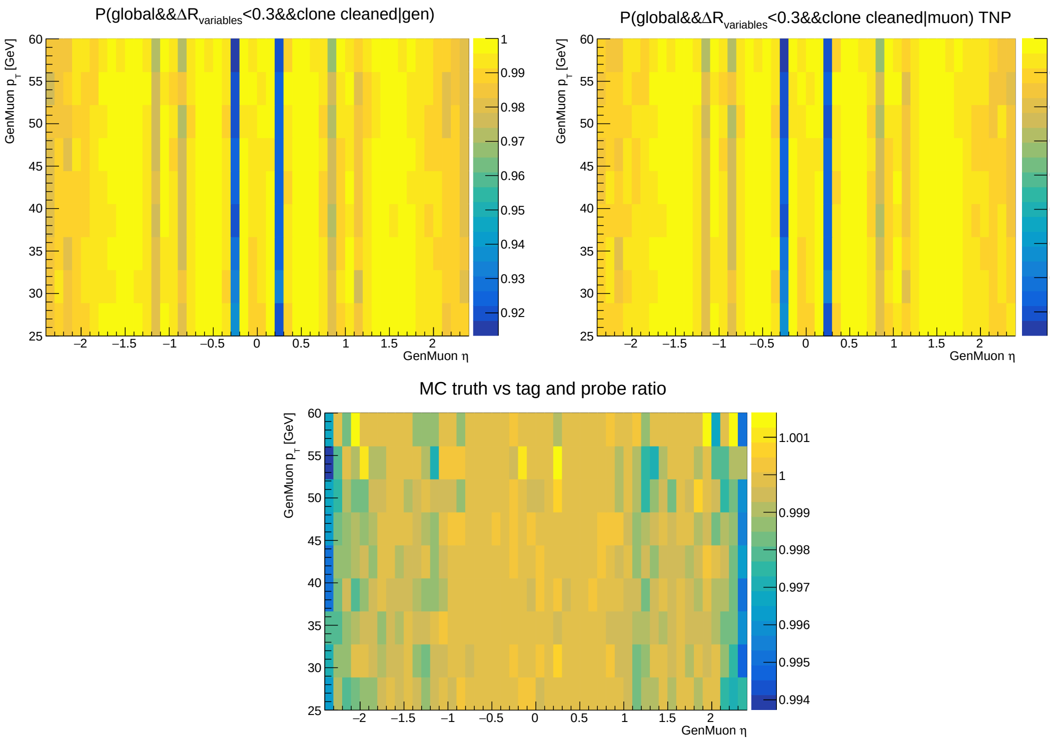


Figure 5.10: MC truth and tag-and-probe efficiency for the $\mathrm{Tracking}$ step when only global muons that have ΔR between the global muon and the closest standalone muon not used in the global muon reconstruction to be greater than 0.1, and ratio between the two, as a function of the postFSR muon η and p_T.

we identify this component as standalone muons that are improperly matched to pileup tracks. This component amounts to the remaining 22% of events.

Figure 5.12 shows the scatter plot of the standalone muon p_T for the global muon and the additional muon on the left and the scatter plot of the ΔR between global muon and the additional muon versus the Δp_T between them on the right. From the first figure it can be seen that the values seem to be highly correlated (points with clone standalone $p_T < 6$ GeV are excluded to suppress tracks from pileup), and this can also be seen in the plot on the right, with only a mild dependence on ΔR up to 0.3, with the large spread in p_T potentially due to the genuine muon spectrometer resolution in the endcaps, where it can reach 30%. From this, we could seemingly conclude that, apart from the pileup contamination already discussed, the additional standalone muons are clone muons.

In order to exclude global muons from pileup, either as a result of an incorrect association between standalone muon track and inner track or as a result of a genuine pileup track reconstructed as a muon, we further restrict our global muon selection to have inner track p_T, which corresponds to the best track p_T for global muons, to be greater than 15 GeV and standalone muon $p_T > 15$ GeV. The impact on the corrections from raising this cut to 10 GeV was found to be negligible. The closure after this additional selection requirement is shown in figure 5.13, where it can be seen that the disagreement in the ratio goes down to 0.1%. For the RECO step, only standalone muons with $p_T > 15$ GeV are considered, and this will be the case also in the actual tag-and-probe derivation of efficiency corrections.

The presence of additional muons can still have an impact on the actual efficiency corrections, as in that case all the muon candidates need to be considered, while in these studies only the candidate that is closest to the generator-level muon is considered. For the tag-and-probe, the matching is done using the standalone muon information, unlike for the MC truth, for which the matching is done using the inner tracker information, so in principle the generator-level muons could be matched to genuine or clone muons equally likely, as is the case for clones at $\eta = 0$. The left of figure 5.14 shows the difference between the ΔR between inner and outer track of the global muon and ΔR between the inner track of the global muon and the clone standalone muon[1]. It can be observed that the clone muon is actually further away from the inner track of the global muon than the corresponding standalone muon associated to the global muon itself, explaining why one is able to match the generator-level muon to the correct global muon using either the inner tracker or muon spectrometer information even in the presence of clones. The same figure, on the right, shows the difference in the number of CSC hits for the standalone muon and the clone muon: it can be seen that the global muon has a higher number of hits in the majority of cases. Inspecting the actual hits also revealed that at least some segments are in common between the genuine and additional muon.

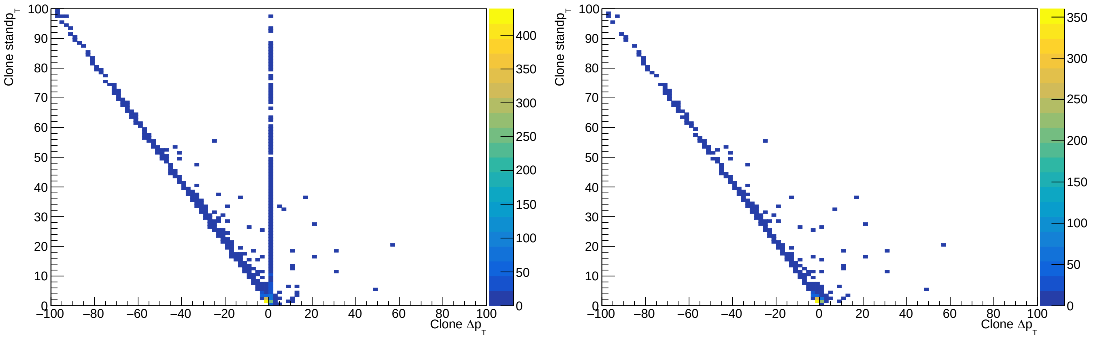


Figure 5.11: Left: 2-D distribution of the clone standalone p_T vs the difference between the clone *best track* p_T and the clone standalone p_T for muons with $|\eta| > 2$. Right: same distribution but including only the clone muons that are also promoted to global muons for muons $|\eta| > 2$.

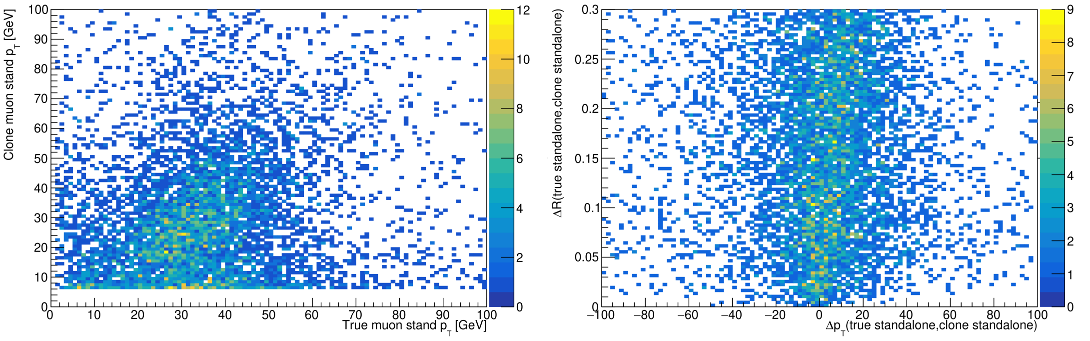


Figure 5.12: Left: 2-D distribution of the clone standalone p_T vs the true standalone p_T for muons with $|\eta| > 2$. Right: 2-D distribution of ΔR(clone standalone, true standalone) vs Δp_T(clone standalone muon, true standalone muon) for muons $|\eta| > 2$.

The matching to generator-level muons in the actual efficiency derivations via tag-and-probe method is done so that all the reconstructed muons that have $\Delta R < 0.1$, or $\Delta R < 0.3$ for the $\mathrm{Tracking}$ step, with respect to a generator-level muon are considered, so that clones are also present in the Monte Carlo efficiencies. In this way, the presence of clones should cancel out between data and simulation, if they are in the same amount. The cause for the additional muons at $|\eta| > 2$ has not been found as of the m_W measurement presented in [1] and in this thesis. In principle, one could use the number of CSC hits to disambiguate between genuine muons and clone muons. Figure 5.15 shows the same efficiencies as in figure 5.13, but selecting the standalone muon with the highest number of hits among muons within a

---

[1] To clarify, this is a difference between two different ΔR, i.e. $\Delta\Delta R$.

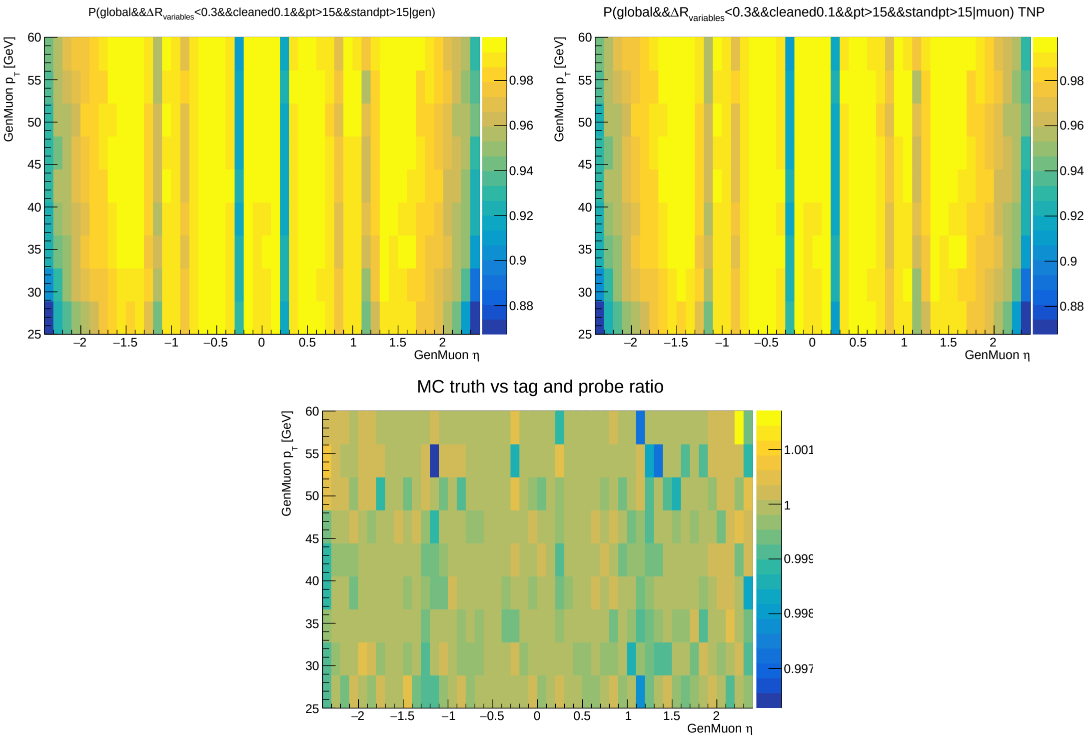


Figure 5.13: Global efficiency for muons with $p_T > 15$ GeV and standalone $p_T > 15$ GeV for MC truth (on the left), tag-and-probe (on the right) and ratio between them (in the bottom).

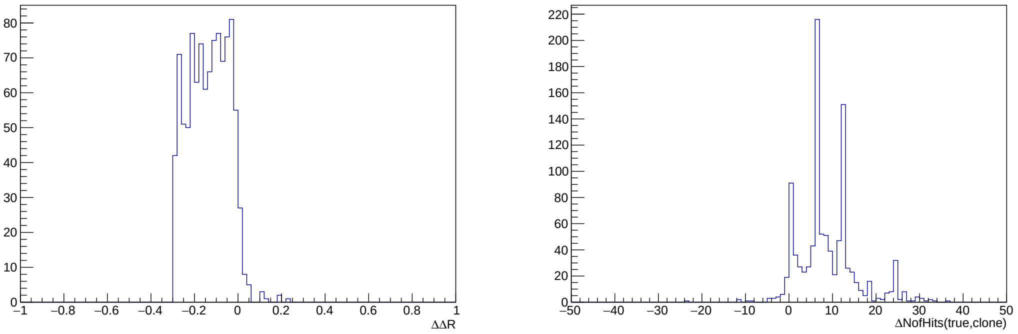


Figure 5.14: Left: difference of the ΔR between internal and external track for true global muon and ΔR between internal track of global muon and standalone clone for global muons with $p_T > 15$ GeV, standalone $p_T > 15$ GeV and $|\eta| > 2$. Right: difference in number of CSC hits for true standalone and clone standalone for standalone $p_T > 15$ GeV and $|\eta| > 2$.

cone of $\Delta R < 1.0$: it can be seen that the agreement is $O(0.1\%)$. However, no attempt at removing or disambiguating these muons has been performed in the analysis, because at this stage the cause for these muons has not been unambiguously identified. The impact of neglecting the presence of these muons in the derivation of the $\mathrm{Tracking}$ efficiency corrections, in the assumption that these are clones, was studied as described in section 7.4, but it was found to have a negligible impact on the extraction of the mass of the W boson.

Figures 5.16, 5.17, 5.18 show the same comparisons between MC truth and tag-and-probe efficiencies for the IDIP, $\mathrm{Trigger}$ and $\mathrm{Isolation}$ steps. The matching to the generator-level muon is done using the inner tracker information. It can be seen that for all of them the agreement is at the per-mille level: this

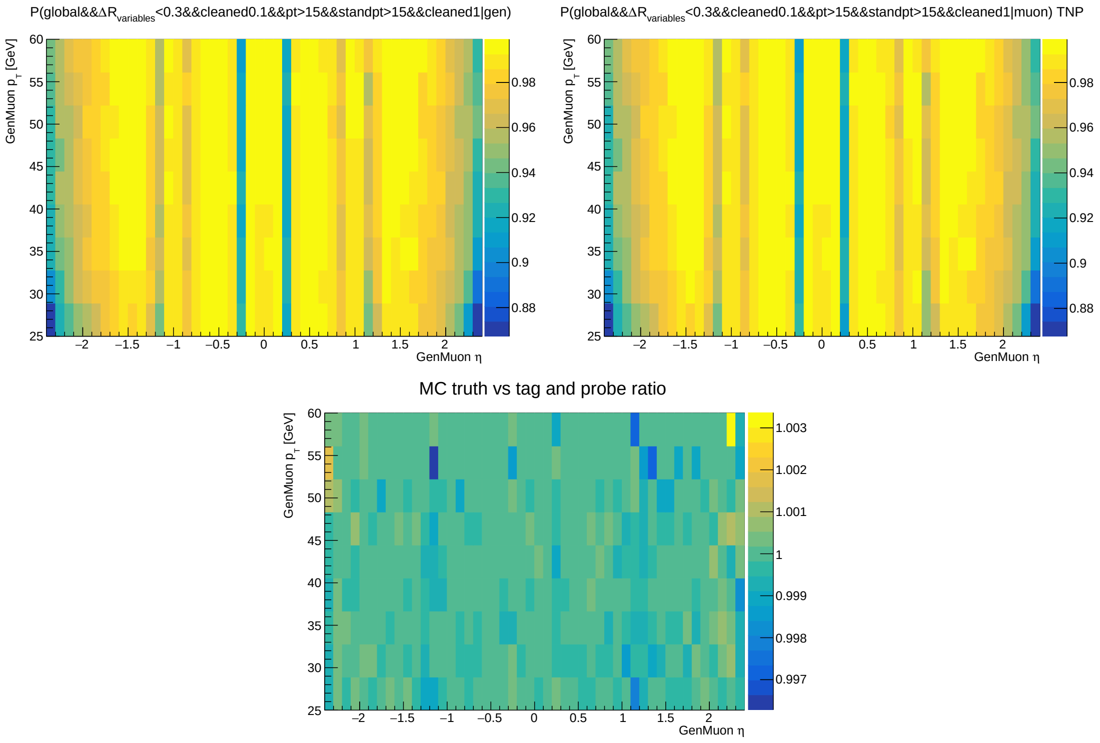


Figure 5.15: Global efficiency for muons with $p_T > 15$ GeV and standalone $p_T > 15$ GeV, selecting the standalone muon with the highest number of hits among muons within a cone of $\Delta R < 1.0$, for MC truth (on the left), tag-and-probe (on the right) and ratio between them (in the bottom).

could be expected because the additional selection is applied directly to the global muon, both for MC truth and for tag-and-probe, and here the tag muon is still not explicitly required.

5.2.2 Bug-fixes to CMS default reconstruction

The flow chart in figure 5.19 shows the standard muon reconstruction chain in CMS, as implemented in the default CMSSW software framework.

- First, segments reconstructed in the external muon chambers are fed into the `StandAloneMuonProducer` module, which combines them to produce standalone muons. The combination can also use the beamspot constraints for improved precision: standalone muons reconstructed in this way are stored in the “standAloneMuons:UpdatedAtVtx” collection, they are mostly duplicate muons from “standAloneMuons” (which do not use the beamspot constraints). However, the reconstruction including also the beamspot constraint is not always successful, and therefore the ”standAloneMuons:UpdatedAtVtx” is only a subset of the ”standAloneMuons” collection.
- The `GlobalMuonProducer` takes both of these collections as input to produce global muons, and implements a disambiguation logic to avoid duplicates.
- `MuonIdProducer` is responsible for merging tracks, standalone muons, and global muons into a single output collection. This producer also returns tracker muons starting from general tracks including also the information from the ”standAloneMuons:UpdatedAtVtx” collection.

The original implementation of this logic in the CMSSW central framework had two problems:

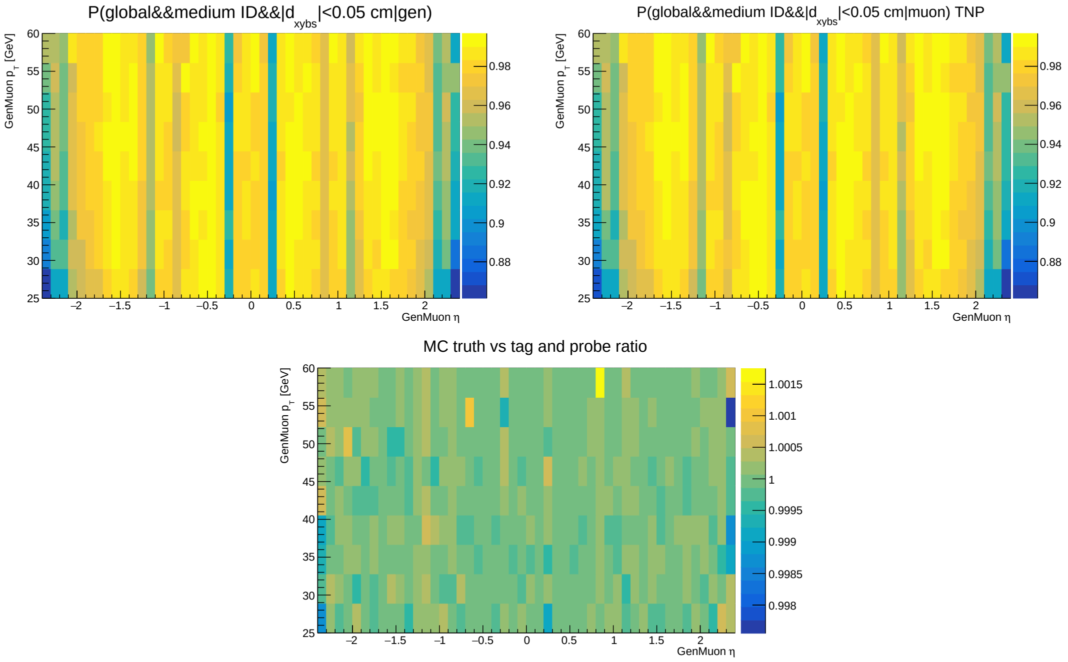


Figure 5.16: Muon reconstruction and identification efficiency up to the IDIP step for MC truth (on the left), tag-and-probe (on the right) and ratio between them (in the bottom). The titles in the first two plots have been shortened for readability, but the full selection described in the text is applied.

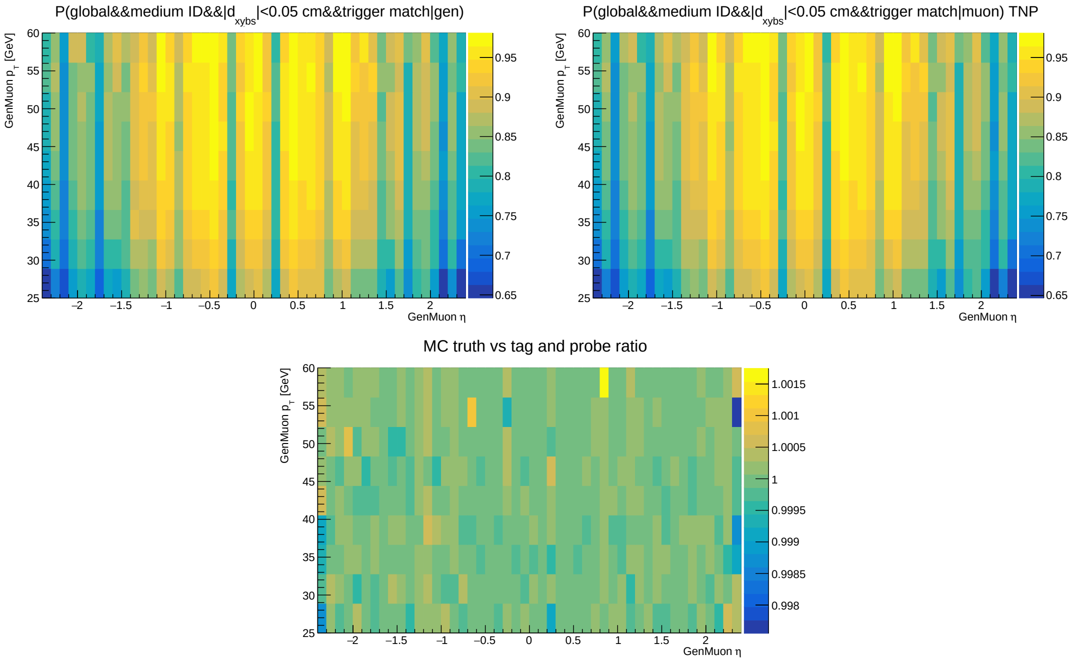


Figure 5.17: Muon reconstruction and identification efficiency up to the $\mathrm{Trigger}$ step for MC truth (on the left), tag-and-probe (on the right) and ratio between them (in the bottom). The titles in the first two plots have been shortened for readability, but the full selection described in the text is applied.

- Standalone muon tracks from the “standAloneMuons” collection for which the “UpdatedAtVtx” reconstruction fails and do not become global muons are not stored in the final “muons” collection. This leads to a bias in the derivation of the RECO and $\mathrm{Tracking}$ efficiencies: for $\mathrm{Tracking}$, the impli-

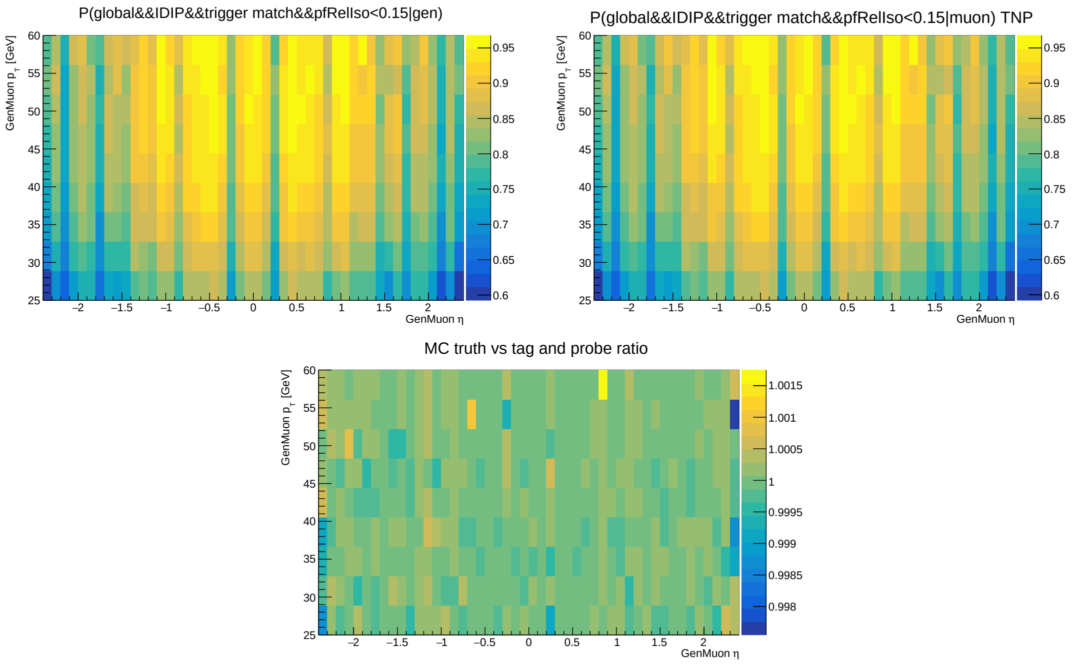


Figure 5.18: Muon reconstruction and identification efficiency up to the $\mathrm{Isolation}$ step for MC truth (on the left), tag-and-probe (on the right) and ratio between them (in the bottom). The full selection described in the text is applied.

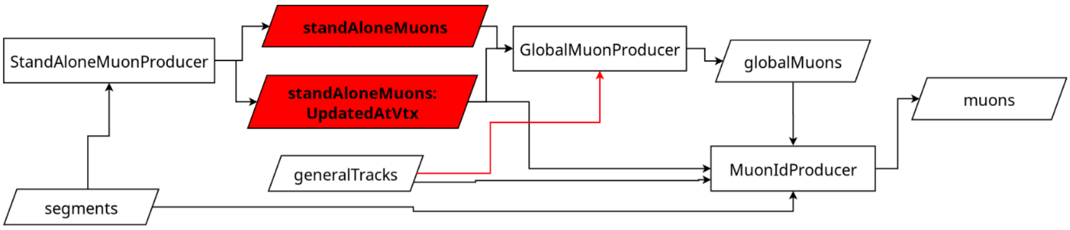


Figure 5.19: Flow chart illustrating the muon reconstruction in CMS. See text for details.

cation is that the efficiency $P(\mathrm{global}|\mathrm{standalone})$ is overestimated, since a fraction of standalone muons that did not become global muons is missing.

- `GlobalMuonProducer` and `MuonIdProducer` use different sets of standalone muons, the former takes as input a combination of ”standAloneMuons” and ”standAloneMuons:UpdatedAtVtx” and implements a disambiguation logic, while the latter only receives the ”standAloneMuons:UpdatedAtVtx” collection as input. The internal duplicate removal in `MuonIdProducer` can fail to remove cases where the global muon was built with an outer track from “standAloneMuons” and the same standalone muon was imported from “standAloneMuons:UpdatedAtVtx”, causing the presence of clone/duplicate standalone muons in the final ”muons” collection.

The disambiguation logic in `GlobalMuonProducer` uses the outer track from “standAloneMuons:UpdatedAtVtx” if it exists, otherwise it falls back to the outer track from “standAloneMuons”, unless the track from “standAloneMuons:UpdatedAtVtx” and the track from ”standAloneMuons” have opposite signs of η: in this case the track from ”standAloneMuons” is used. This is more likely to occur when the muons are at $\eta = 0$, explaining the presence of these clones. For this analysis, we created a new collection of ”MergedStandAloneMuons”, which takes both “standAloneMuons:UpdatedAtVtx” and “standAloneMuons” collections as input and stores all the standalone muons with the disambiguation logic from `GlobalMuonProducer`, addressing both problems. We store all the necessary information, in the form of unique identifiers assigned to each of the objects reconstructed in CMSSW, to match the ”MergedStandAloneMuons” objects to the global muons in the ”muons” collection. ”MergedStandAloneMuons” is the standalone muon collection that is used to derive the RECO and $\mathrm{Tracking}$ (as the denominator in equation 5.2) efficiency corrections to data. It was proven that replicating the comparison in figure 5.13 using this collection returns the same level of agreement, but without the ”clone cleaning”, thus recovering a small loss of efficiency at $\eta = 0$. This issue has also now been corrected in the central CMSSW reconstruction and is implemented for the central production, for example, of samples for Run 3 data-taking analysis.

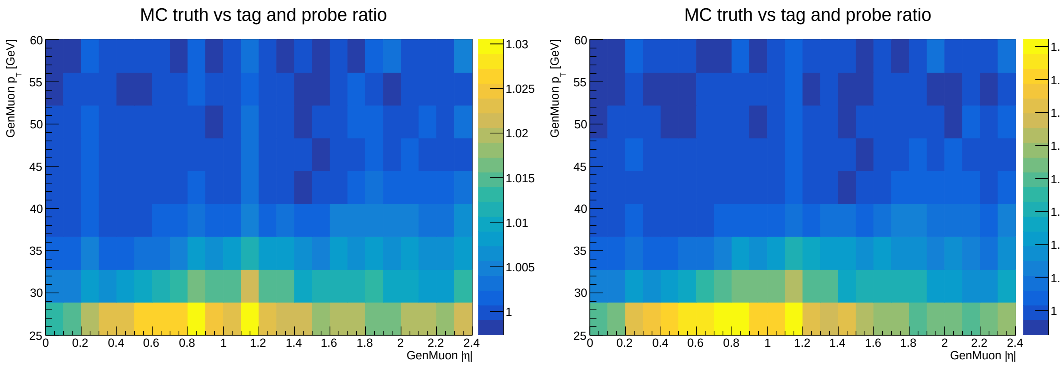
This bug-fix does not address the presence of additional muons at $|\eta| > 2$, for which, as already mentioned, no particular action has been taken.

5.2.3 Recoil-dependence

The same set of comparisons between MC truth and tag-and-probe is repeated with the requirement on the presence of a tag muon. The selection on the tag muon is:

- $p_T > 25$ GeV and $|\eta| < 2.4$;
- matched to either a trigger object returned by the `HLT_IsoMu24` or the `HLT_IsoTkMu24` trigger menus;
- passes medium ID, $|d_{\mathrm{xyBS}}| < 0.05$ cm (i.e. the IDIP requirement), and is a global muon;
- isolation as defined in equation 5.1 is less than 15% of the muon p_T (i.e. the $\mathrm{Isolation}$ step).

The level of agreement, within the statistical uncertainties of the sample, is similar to that of the previous comparisons for all steps except for trigger and isolation, for which the ratios between MC truth and tag-and-probe are shown[2] in figure 5.20, where the disagreement between the two definitions reaches 4%.



(a) $P(\mathrm{Trigger}, \mathrm{IDIP}, \mathrm{Tracking}, \mathrm{RECO}|\mathrm{gen})$ (b) $P(\mathrm{Isolation}, \mathrm{Trigger}, \mathrm{IDIP}, \mathrm{Tracking}, \mathrm{RECO}|\mathrm{gen})$

Figure 5.20: Ratio of MC truth and tag-and-probe efficiencies for $\mathrm{Trigger}$ (on the left) and $\mathrm{Isolation}$ (on the right). In this case, the selection on the tag muon outlined in the text is applied.

Figure 5.21 shows the muon identification and reconstruction efficiency up to $\mathrm{Isolation}$ step, but without requiring that the probe muon is matched to a trigger object. The efficiencies are higher at low p_T

---

[2] Efficiencies here are shown as a function of $|\eta|$ to reduce statistical fluctuations in this studies. The total statistics has been increased anyway with respect to the studies without the tag muon.

compared, for example, to figure 5.18, because of the lack of requirement on the Trigger matching, but the disagreement between MC truth and tag-and-probe is similar, indicating that the effect is mostly due to isolation. The trigger selection used in this analysis, as well as in a large number of CMS analyses with at least one muon, requires the presence of one muon that also passes some isolation criteria, explaining why this behaviour is already visible in the Trigger step.

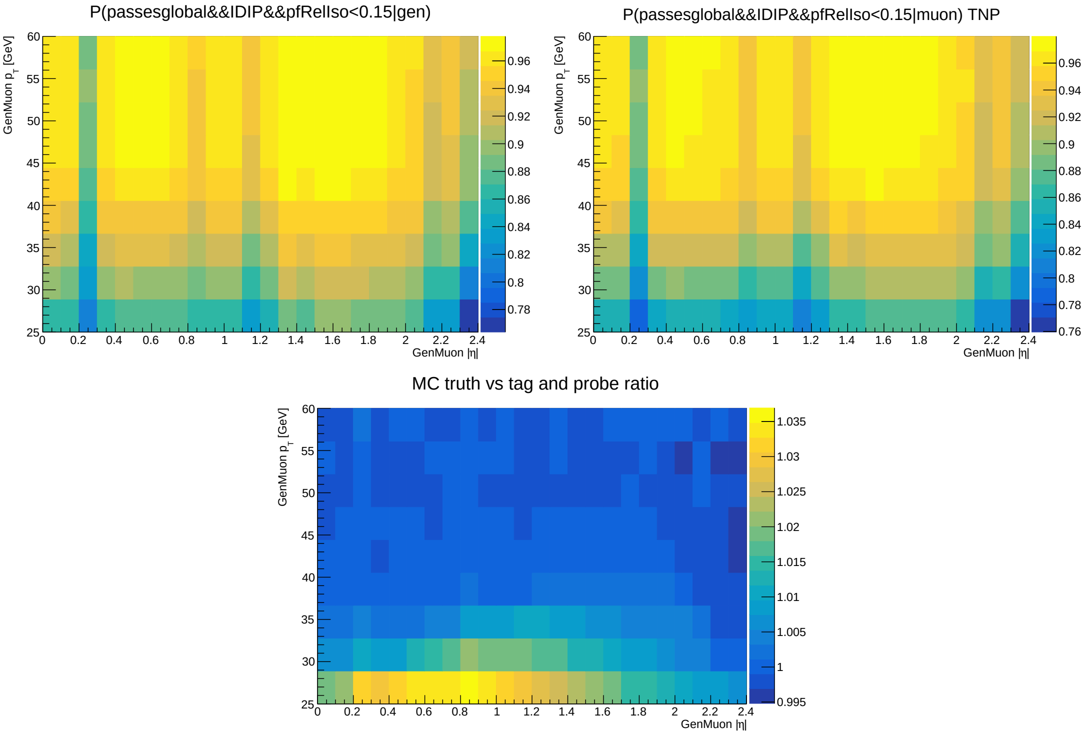


Figure 5.21: Muon reconstruction and identification efficiency up to the Isolation step without the Trigger requirement for MC truth (on the left), tag-and-probe (on the right) and ratio between them (in the bottom). In this case, the selection on the tag muon outlined in the text is applied.

The efficiency of the isolation requirement depends on the boson recoil and its proximity to the lepton. This is schematically shown in figure 5.22: if the muon is in the shadow of the hadronic recoil of the boson, a large number of particles is most likely inside the cone used to define the isolation variable, as opposed to when this is not the case. A sketch is shown for a tag-and-probe event, but this is also valid for W decays, where only muon is present. The variable that is traditionally used to describe the dependence of the isolation efficiency on the boson recoil is [38]:

$$u_T = \frac{\vec{p}_T^{\,\mu} \cdot \vec{p}_T^{\,\mathrm{boson}}}{p_T^{\mu}}, \tag{5.7}$$

which is the projection of the p_T of the boson on the direction of flight of the muon. The expected dependence of isolation and trigger efficiency on u_T in simulation is shown in figure 5.23 for $W^+ \to \mu^+\nu_\mu$ events that pass the full selection as described in chapter 7. The u_T distribution has a peak at 0, mostly corresponding to events in which the boson p_T is soft, and is skewed towards positive values, mostly correlated with high values of p_T, particularly above the Jacobian peak. For these events, the trigger and isolation efficiency decreases for negative u_T values, where the muon tends to be in the shadow of the hadronic recoil. The efficiency increases slightly for $u_T > 0$ up to a value of $u_T \approx 25$ GeV: again, this is due to the correlation between p_T and u_T and because the analysis requires a cut on the relative isolation, implying that the isolation requirement is more relaxed for high p_T. The decrease for $u_T > 25$ GeV is due to a more intense hadronic activity as the boson p_T increases. The figure also shows the efficiency corrections applied in the analysis, to show their magnitude and dependence on u_T. Notice that the figure also shows an alternative version of the corrections that is derived in tag-and-probe without taking the u_T dependence into account, to highlight that, in our case, it is necessary to explicitly correct for u_T.

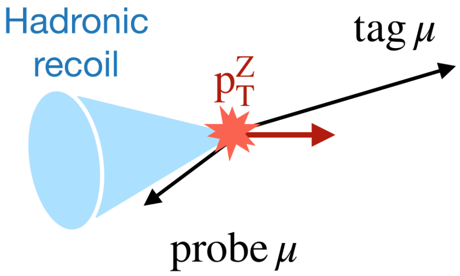


Figure 5.22: Sketch showing a $Z \to \mu\mu$ decay and the Z boson recoil. The muons from the decay are indicated as tag and probe.

The requirement on the tag muon modifies the total phase space in which the isolation and trigger efficiencies are computed, modifying the magnitude of the boson recoil and the angle between recoil and lepton, thus also affecting the u_T distribution and the total isolation and trigger efficiencies as a consequence.

In the following, a series of studies are presented to check whether the u_T variable is sufficient to capture the full dependence of isolation on recoil and angle. They consist in computing the MC truth and tag-and-probe $\mathrm{Isolation}$ efficiencies, without the trigger matching requirement, as a function of three variables: $p_T - \eta - u_T$. If isolation is only dependent on $p_T - \eta - u_T$, MC truth and tag-and-probe efficiencies computed in bins of the three variables should be the same. Figure 5.24 shows the non-closure between MC truth and tag-and-probe in four bins of u_T: there is a disagreement at the percent level for low p_T at negative (-16 GeV$< u_T < -14$ GeV) and positive (6 GeV$< u_T < 8$ GeV and 14 GeV$< u_T < 16$ GeV) values of u_T, indicating that part of the dependence is not fully captured by the u_T variable.

Dependence on final-state radiation

The behaviour described above applies to hadronic activity, but another source of bias that might be introduced by the request of a tag muon is related to the total final-state radiation: for example, two muons might arise from a boson with low p_T, but then one them radiates, resulting in a change in the p_T of the pair. The request on the tag muon might change the contribution from final-state radiation. Final-state radiation can be assumed to be mostly collinear with the muon, making it likely to be within the isolation cone. The isolation requirement in the High-Level Trigger and the offline requirement exclude photons that are within a radius of $\Delta R < 0.01$, but there might be FSR photons that fall outside this cone but are still within the total isolation cone, which can carry higher p_T. Figure 5.25 shows the distribution of $\Delta p_T(\mathrm{preFSR{,}postFSR})$, a variable directly related to final-state radiation in the assumption of collinear emission, for 0 GeV$< u_T < 5$ GeV for a low p_T bin (25 GeV$< p_T < 30$ GeV) and a high p_T bin ($p_T > 40$ GeV). For this particular u_T bin the amount of FSR radiation between MC truth and tag-and-probe is quite different at low p_T, while it is mostly the same for high p_T muons, which should be less affected by request of a tag muon. However, this behaviour might be dependent on u_T.

Figure 5.26 shows the isolation efficiency as a function of $\Delta p_T(\mathrm{preFSR{,}postFSR})$ for the same bins of u_T and p_T as in figure 5.25. A clear dependence on the amount of final-state radiation can be seen, and is similar between low and high p_T. This behaviour is found in other u_T bins as well, indicating that it is likely due to final-state radiation inside the isolation cone but outside of the inner cone of radius $\Delta R < 0.01$.

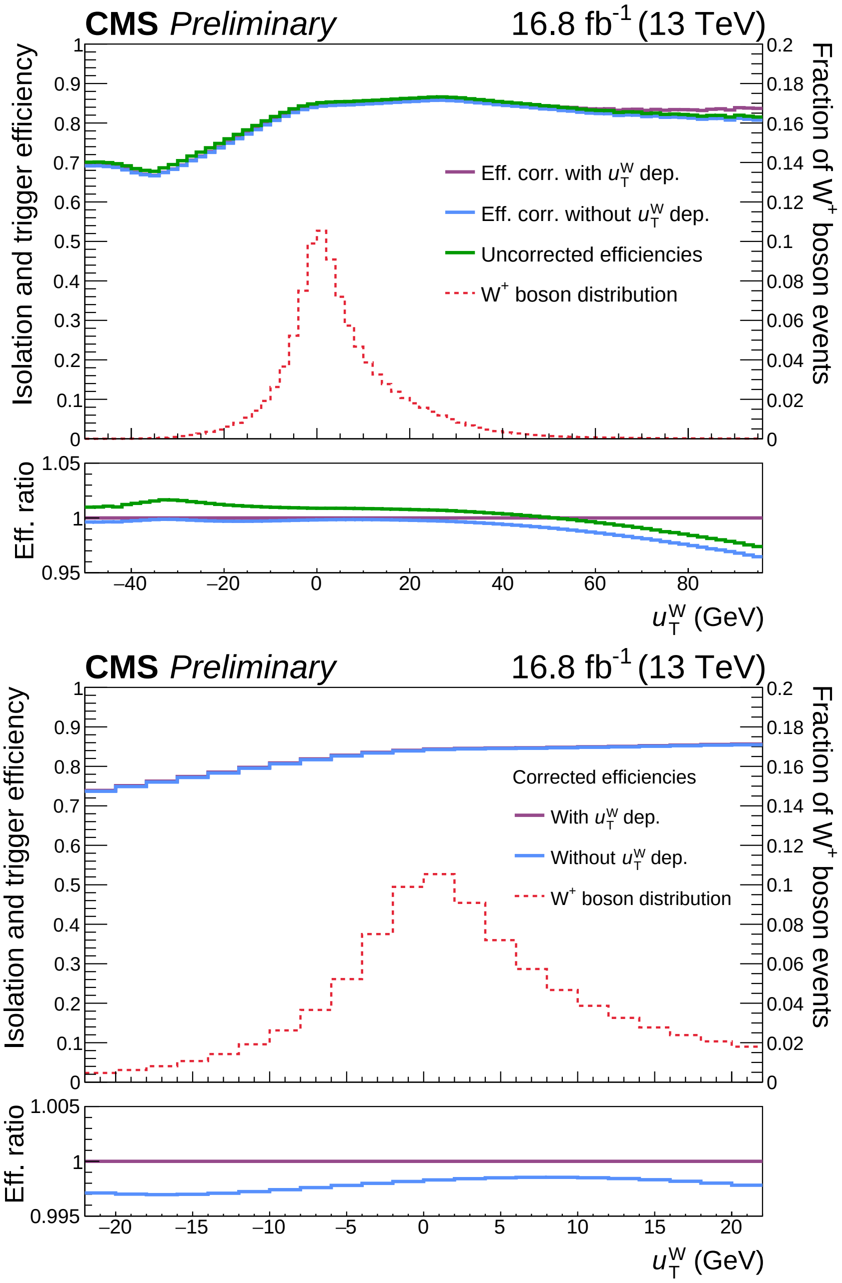


Figure 5.23: Isolation and trigger efficiency for muons from the W boson decays as a function of the projection of the p_T of the W boson along the direction of the muon (u_T^W). The event selection is the same as in the analysis. The green line shows the efficiency without corrections extracted from data, the blue line shows the efficiency corrected only as a function of the p_T and η of the muon, the purple line shows the efficiency corrected as a function of p_T, η and u_T. The u_T^W distribution is plotted in red. In the bottom figure shows the efficiency and u_T^W distribution for $|u_T^W| < 22$ GeV.

In order to separate the effects of final-state radiation from those due to hadronic activity, the same comparison between tag-and-probe and MC truth was repeated by subtracting the amount of final-state radiation, expressed in terms of $\Delta p_T(\mathrm{preFSR{,}postFSR})$, from the relative isolation, both differentially and integrated over u_T. Figure 5.27 shows the non-closure before and after the final-state radiation subtraction from the isolation. The p_T binning for this comparison has been increased. The non-closure with this final-state radiation subtraction is now approximately half of that observed for the standard isolation. The same

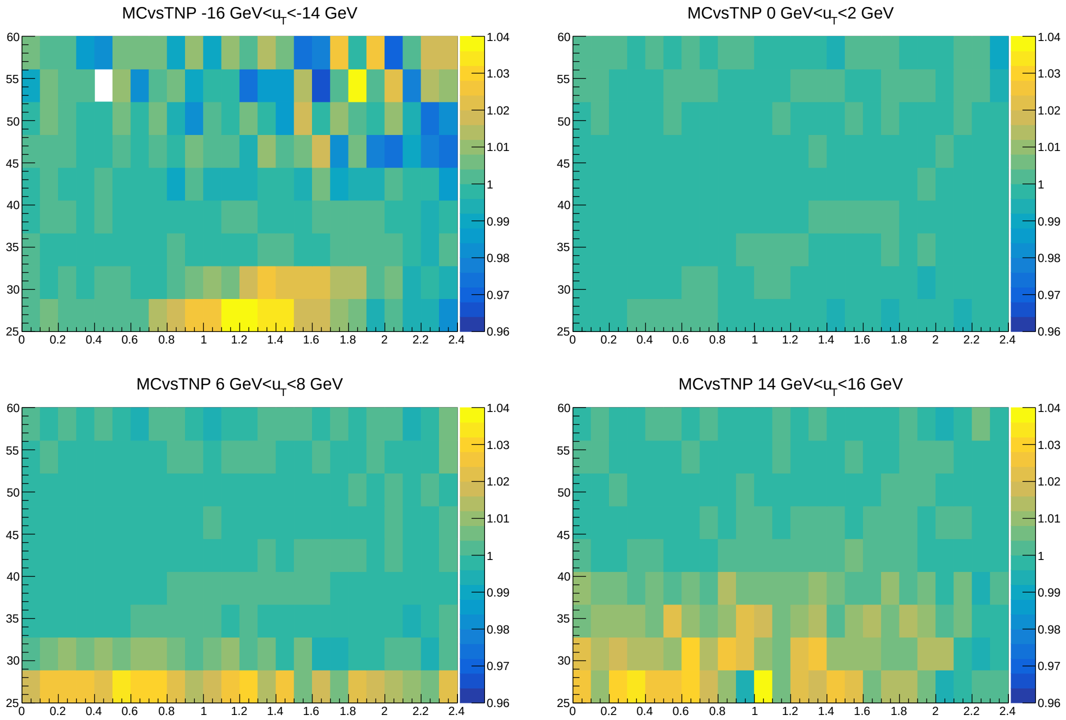


Figure 5.24: Comparison of the isolation efficiency between MC truth and tag-and-probe in four u_T bins.

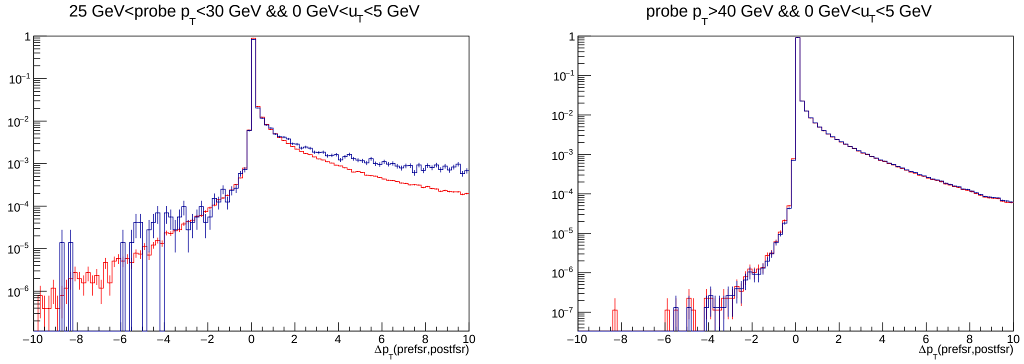


Figure 5.25: Δp_T(preFSR,postFSR) for probe muon for 0 GeV$< u_T < 5$ GeV for two p_T bins. In red MC truth, in blue tag-and-probe.

was also done differentially as a function of u_T. Figure 5.28 shows this non-closure in the same u_T bins as in figure 5.24: the non-closure appears to be smaller overall, but there is still some residual difference, particularly for positive values of u_T.

The same comparisons were repeated using preFSR variables and without removing the final-state radiation from isolation: a given postFSR p_T bin is filled with events from various preFSR p_T that then emit radiation, and this could introduce a difference in the efficiencies as a function of u_T, especially at postFSR p_T lower than the tag threshold. But one would expect the FSR emission process to be independent from the requirement on a tag muon. Figure 5.29 shows the non-closure between MC truth and tag-and-probe where the final-state radiation is subtracted as described above (similarly to the right of figure 5.27) and the non-closure between MC truth and tag-and-probe when the efficiencies are binned as a function of preFSR p_T and $|\eta|$ and no final-state radiation subtraction is done. It can be seen that the two non-closures are similar. Figure 5.30 shows the same as in figure 5.29 but differentially in u_T for the same bins as in figure

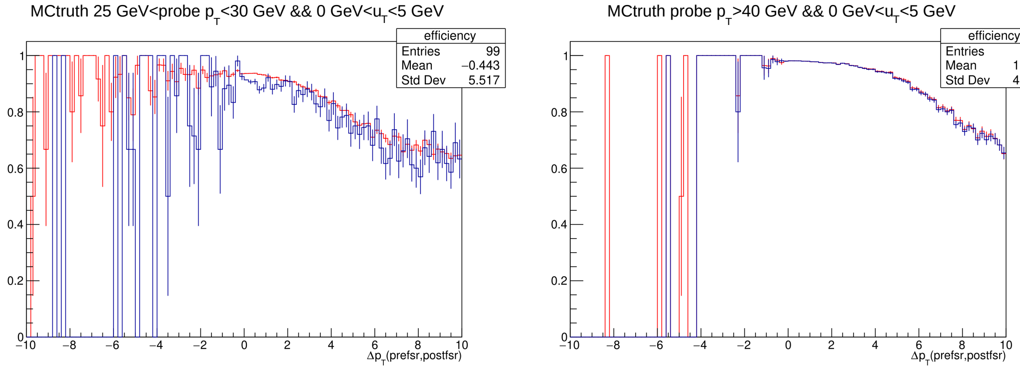


Figure 5.26: Isolation efficiency as a function of Δp_T(preFSR,postFSR) for probe muon for 0 GeV$< u_T < 5$ GeV for two p_T bins. In red MC truth, in blue tag-and-probe.

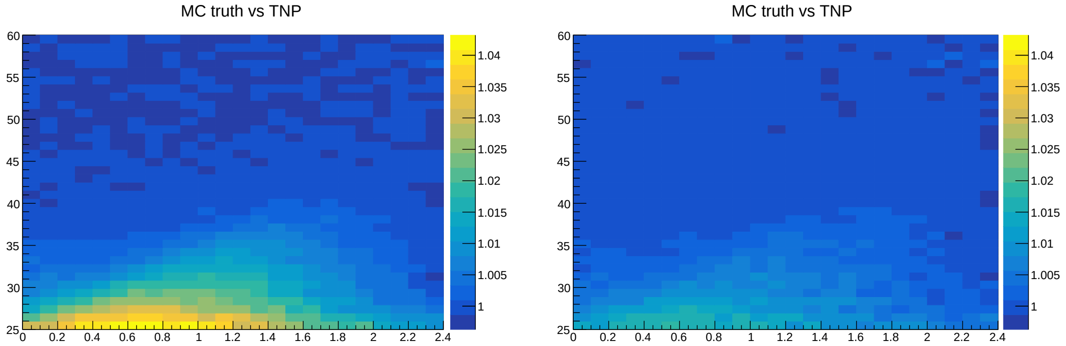


Figure 5.27: Non-closure between MC truth and tag-and-probe for P(pfRelIso04_all<0.15|idip) before and after the final state radiation subtraction from the isolation, as described in the text.

5.28. The preFSR variables are used also in the definition of u_T, showing that the non-closures are similar also in this case.

After verifying the impact of final-state radiation on the isolation variable, the possibility of using variables different from u_T was investigated. This was done with the aim of capturing the dependence on the boson recoil and direction of isolation and trigger efficiencies in a more complete way. The isolation efficiency was evaluated as a function of the variables that describe the vector boson production mechanism and the subsequent decay into leptons: q_T^V, y^V, and the Collins-Soper angles $\cos\theta^*$, ϕ^*.

The two variables that were first studied at first are q_T^V and $\Delta\phi(q_T, \text{probe } p_T)$ (in the lab frame): u_T is a function of both variables and sensitive to their underlying distributions. For example, a given u_T bin can be filled from events having different values of q_T^V and $\Delta\phi(q_T^V, \text{probe } p_T)$, resulting in different isolation efficiencies. PreFSR variables were used to disentangle the effects of final radiation from QCD hadronization. Figure 5.31 shows the isolation efficiency as a function of q_T and $\Delta\phi(q_T^V, \text{probe } p_T)$ for probe p_T in the range $[25{,}28.9]$ GeV and $|\eta| < 0.1$: on the top left the MC truth efficiencies are shown, the tag-and-probe efficiencies are shown on the right. On the bottom of the same figure, the ratio between the two is shown. Overall, no trend in the $q_T - \Delta\phi(q_T^V, \text{probe } p_T)$ plane seems to be present. However, in the tag-and-probe method efficiencies cannot be evaluated at low p_T when the probe is parallel to the boson and q_T^V is larger than 5 GeV, a configuration that is kinematically forbidden, as shown in figure 5.32.

As an additional check, the tag-and-probe efficiencies were evaluated by reweighting the events with the MC truth efficiencies as a function of p_T, η, q_T^V and $\Delta\phi(q_T^V, \text{probe } p_T)$. If q_T^V and $\Delta\phi(q_T^V, \text{probe } p_T)$ were sufficient to capture isolation dependence, the only difference between MC truth and tag-and-probe should be related to the different probability distributions. The residual non-closure is shown in figure 5.33

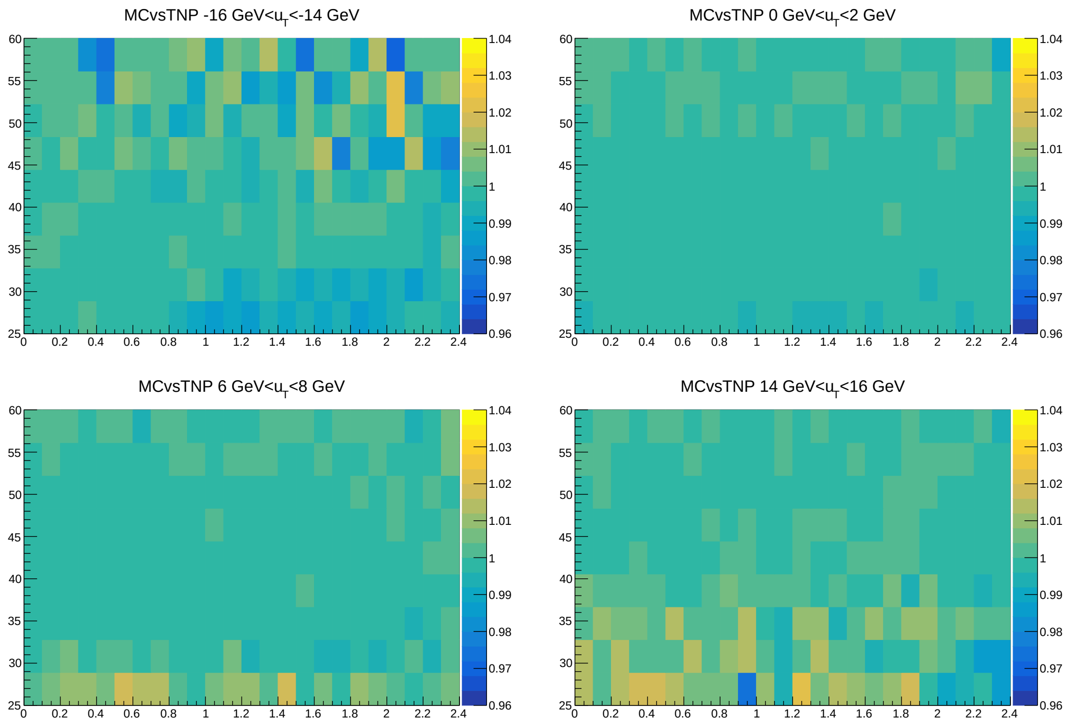


Figure 5.28: Comparison between the isolation efficiency in MC truth and tag-and-probe in four u_T bins but with final state radiation subtracted from the isolation variable as described in the text.

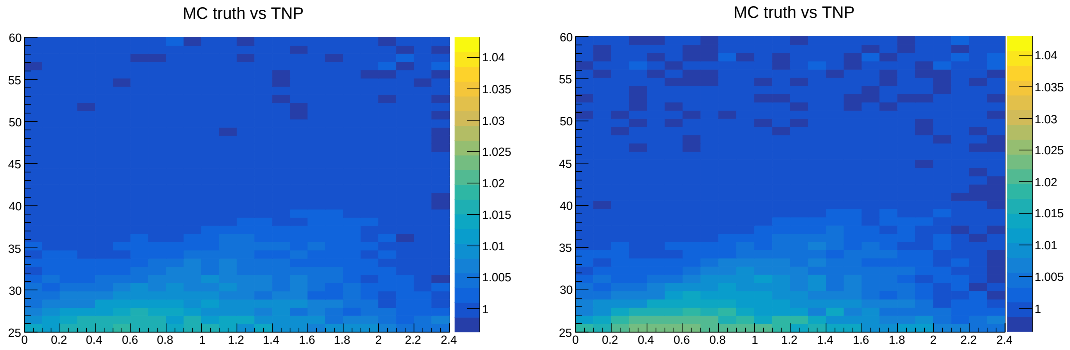


Figure 5.29: Non-closure between MC truth and tag-and-probe for P(pfRelIso04_all¡0.15|idip) after the final state radiation subtraction from the isolation (on the left) and when preFSR variables are used and no FSR subtraction is performed (on the right).

for $P(\mathrm{Isolation}|\mathrm{IDIP})$. The disagreement is reduced at most to 0.5%. For values of $p_T > 40$ GeV the agreement is at $O(0.1\%)$. However, because part of the phase space cannot be probed by this approach, $\mathrm{Isolation}$ and $\mathrm{Trigger}$ Scale Factors are extracted only as a function of $p_T - \eta - u_T$. Figure 5.34 shows the non-closure on $W^+ \to \mu^+\nu$ decays compared to Z with tag-and-probe selection: in the first figure, no reweighting is applied, in the second figure, the u_T distribution for tag-and-probe events is corrected to match the one of W^+ decays for each $p_T - \eta$ bin, and on the right the same is done as a function of preFSR variables. The u_T reweighting improves the agreement, especially when preFSR variables are considered (in this case the disagreement is atmost $O(0.5\%)$).

In the computation of Scale Factors, where the ratio of efficiencies between data and MC is taken, at least part of the residual non-closure should cancel out, including the effect of final-state radiation.

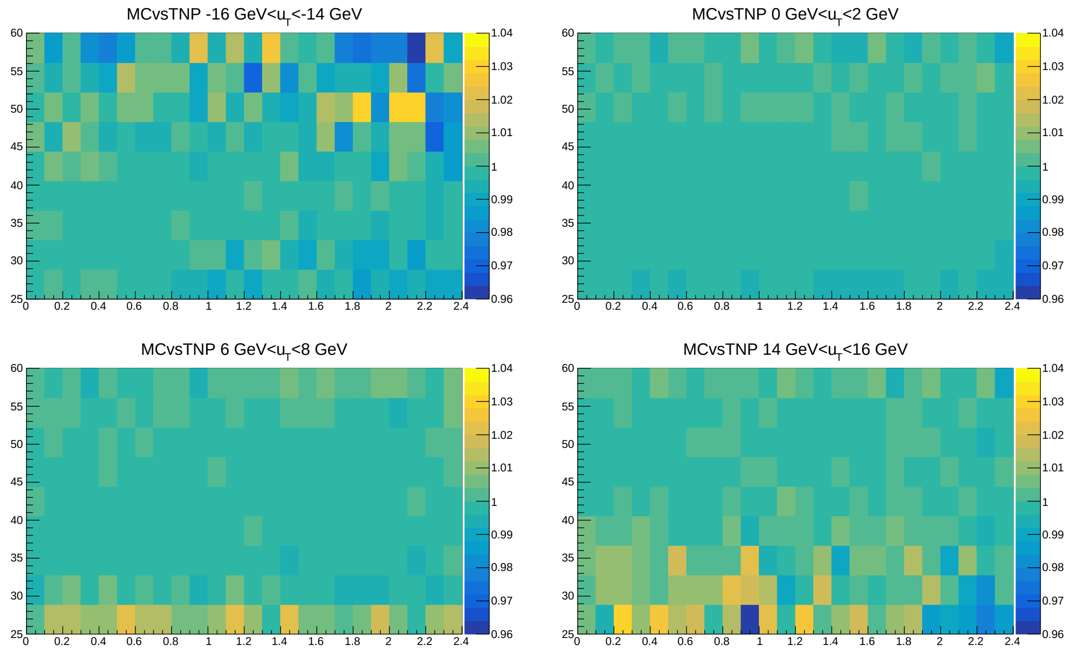


Figure 5.30: Comparison between the isolation efficiency in MC truth and tag-and-probe in four u_T bins but using preFSR variables as described in the text.

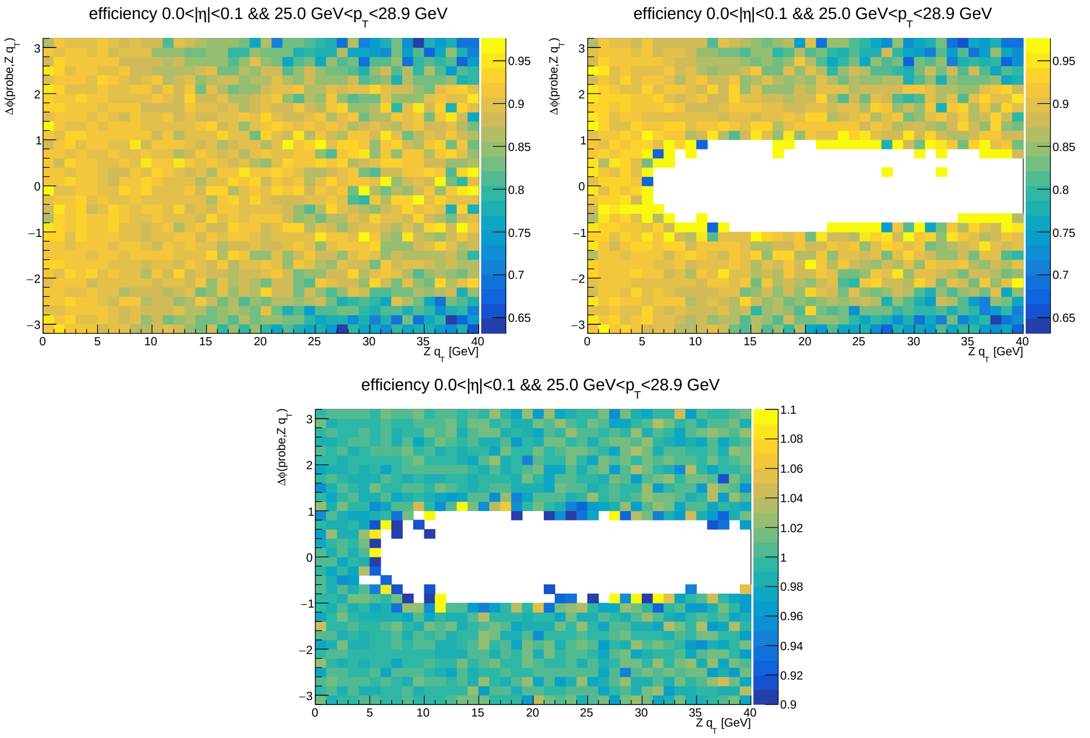


Figure 5.31: Efficiencies as a function of q_T and $\Delta\phi(q_T^V, \text{probe } p_T)$ for probe $25 < p_T < 28.9$ GeV and $|\eta| < 0.1$: on the top left the MC truth efficiencies are shown, on the top right the tag-and-probe efficiencies are shown, on the bottom the ratio between the two is shown.

The possibility of using only charged PF hadrons to compute the isolation has been considered, and it seems to given a better agreement between MC truth and tag-and-probe efficiencies, with a similar level

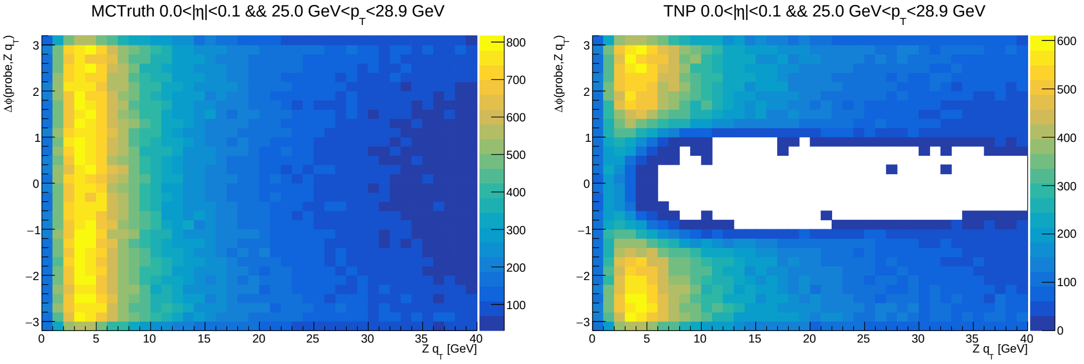


Figure 5.32: $q_T - \Delta\phi(q_T^V, \text{probe } p_T)$ spectrum for probe $25 < p_T < 28.9$ GeV and $|\eta| < 0.1$. On the left: MC truth. On the right: tag-and-probe.

of agreement as in right of figure 5.34, but this would still not be available for the trigger, so that biases from final-state radiation would not be removed, unless the isolation threshold used in the $\mathrm{Trigger}$ step is tight enough to be completely inside that of the trigger. It is not guaranteed, however, that a definition of the isolation variable that only uses the charged PF hadrons is as good at separating between signal (the electroweak processes) and background (the nonprompt muons) as the definition used by default, so other difficulties might arise.

To conclude, for this analysis all studies support the choice of deriving the $\mathrm{Trigger}$ and $\mathrm{Isolation}$ efficiencies as a function of $p_T - \eta - u_T$ to capture their dependence on the hadronic recoil.

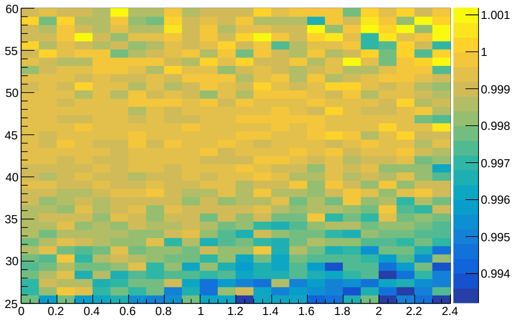


Figure 5.33: Ratio of tag-and-probe and MC truth efficiencies where the latter were reweighted to match the q_T and $\Delta\phi(q_T^V, \text{probe } p_T)$ distributions observed in tag-and-probe for each p_T and η bin.

5.2.4 Vertex-agnostic isolation

As already mentioned, charged PF candidates used in the computation of isolation in 5.1 are split into pileup and non-pileup using the information from the main primary vertex, selected as described in 4.1. In figure 5.4 it can be seen that the probability that the main primary vertex is the $W \to \mu\nu$ or $Z \to \mu\mu$ production vertex is different for Z and W, which can result in differences in isolation efficiency that might not be fully accounted for by the tag-and-probe procedure. This analysis uses a definition of isolation in which the pileup component is determined using $\Delta Z_{\text{Muon, charged PF candidate}}$, where Z is the coordinate of point of closest approach along the beamline. A similar logic is already implemented in the `HLT_IsoMu24` and `HLT_IsoTkMu24` trigger paths used in the analysis, where PF candidates belonging to the same vertex of the muon are those with $|\Delta Z_{\text{Muon, charged PF candidate}}| < 2$ mm, as used in the analysis.

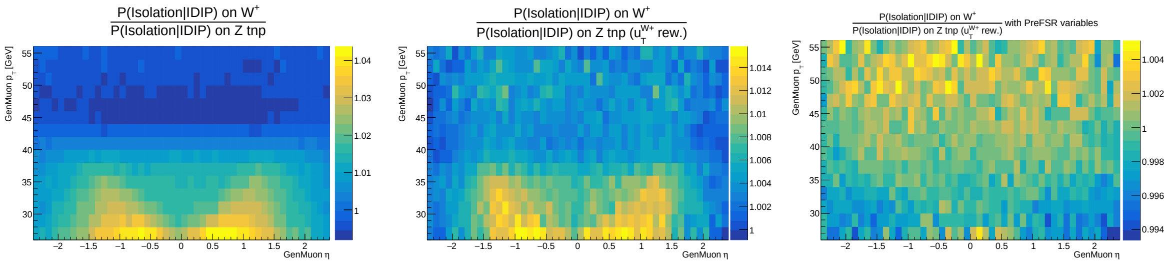


Figure 5.34: Ratio of isolation efficiencies for $W^+ \to \mu^+\nu$ decays and Z with tag-and-probe selection for positive charge without reweighting in u_T (left), reweighting the Z efficiencies for the $u_T^{W^+}$ distribution with postFSR (center) and preFSR (right) variables, used also for plotting. Note the different Z scales in each of the plots.

The CMSSW software framework was modified so that a new isolation variable can be computed using this logic. In the following, a number of studies to validate the new isolation variable are presented. In particular, for most CMS analyses only a small fraction of the total information in the event is stored in the top-level nTuples, amounting to about 2 kbytes per event in the NanoAOD data-tier, which is the lightest data-tier currently available. In this particular data-tier only composite variables are stored (the standard physics objects such as muons, but also other variables such as isolation, the ID bits described in 5.1, MET, etc.) but not the information necessary to recompute them, the latter one only available in lower-level data-tiers, such as MiniAODs and AODs, for which the event size is larger and the objects are stored as C++ classes. In addition, information in NanoAOD is typically stored with lower precision (single precision floating point numbers are used, and typically with a reduced number of digits). In this analysis the NanoAOD data-tier is used, but additional information had to be included, for example to take into account the presence of clone muons described previously. They have been computed from either MiniAOD or AOD[3], increasing the total size per event is 4 kbytes.

The goal is, therefore, to understand what is the minimum amount of information needed to properly compute the isolation variable without information on the main primary vertex, that is, what collections of particles are required and in what data-tier they are stored, but still verifying that it behaves as expected. This is done by verifying that the individual components of isolation defined in 5.1 are in good agreement between the calculations, using a sample of $Z \to \mu\mu$ events considering global muons with $p_T > 25$ GeV that pass the IDIP selection, for which the change in isolation definition should result in a negligible effect.

The exact list of particles used in the default computation of isolation is as follows:

- charged PF hadrons within $\Delta R < 0.4$ from the muon, with an internal cut of $\Delta R > 10^{-4}$ to remove overlap;
- neutral PF hadrons with $p_T > 500$ MeV, within $0.01 < \Delta R < 0.4$ from the muon;
- photons with $p_T > 500$ MeV, within $0.01 < \Delta R < 0.4$ from the muon;
- charged PF hadrons used in the subtraction of the pileup component within $0.01 < \Delta R < 0.4$ from the muon.

It was verified that the neutral PF hadron component and the photon component of isolation are in very good agreement between the definitions, as shown in figure 5.35, and in principle there is no need to recompute them, as they are not affected by the choice of vertex, but it was still useful to have as a cross check.

The biggest differences are in the charged PF components, as shown in figure 5.36. It can be seen that both distributions peak at 1, but there are also large differences, especially for the charged component

---

[3] In CMS, the NanoAOD data-tier is processed from the input in MiniAOD, and MiniAOD is processed from the input in AOD

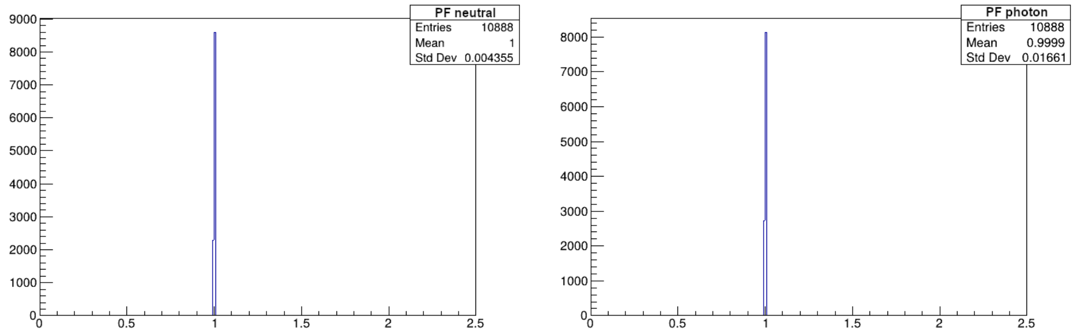


Figure 5.35: Ratio between the standard and recomputed isolation, for the neutral PF hadrons (left) and photon (right) components.

from the muon vertex, which result in 0.3% effect on efficiency (integrated over phase space). We further investigated this to see whether this might be due to the limited amount of information available in MiniAOD or to genuine differences introduced by the new definition. To assess this hypothesis, we reproduced the splitting into charged and PU components as it is done in the standard computation, i.e. dividing the PF candidates according to the minimum $|\Delta Z_{\text{PF candidate, vertex}}|$. The results of this comparison are shown in figure 5.37, where it can be seen that the agreement is very good, indicating that it is possible to recompute the isolation variable using the information from MiniAOD with a high-level of agreement.

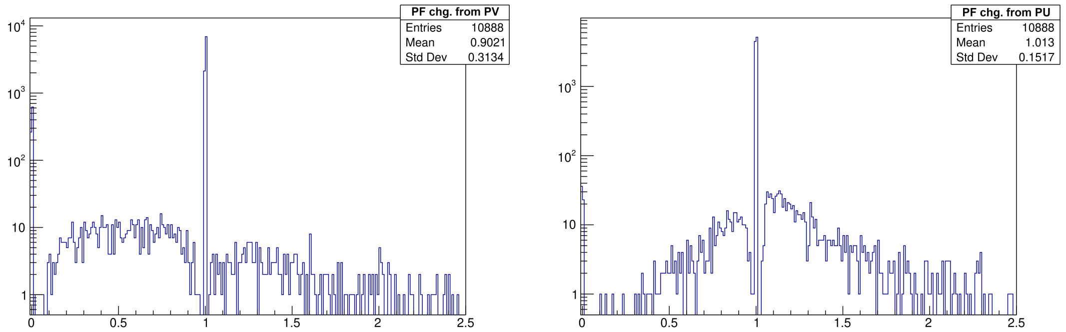


Figure 5.36: Ratio between the standard and recomputed isolation, for the charged (left) and PU (right) components.

The differences observed in figure 5.36 are possibly due to resolution effects (a track that belongs to the same vertex as the muon might be considered as pileup if $|\Delta Z| > 2$ mm and vice-versa) or to the actual value of ΔZ used to determine the pileup and non-pileup component, which has not been optimized, but inherited from what is done at the trigger level. Let us consider a numerical example: considering the distribution of $z-$coordinates the primary vertices as a Gaussian of standard deviation 3.5 cm (typical at CMS) and a number of 30 vertices per bunch-crossing, the probability to have at least one vertex within $|\Delta Z| < 2$ mm from one particular vertex among the 30 is about 60%. The typical vertex resolution is $O(20)$ μm, so it is likely that part of the pileup component is being considered as from the same vertex as the muon, a contamination which could be reduced by tuning the ΔZ requirement. However, this pileup contamination is the same for $Z \to \mu\mu$ and $W \to \mu\nu$ events, so that its effects on the isolation variable should already be taken into account when efficiency corrections are derived via the tag-and-probe method. In addition to this, to compute how much is the contamination from pileup on the isolation variable, the probability that pileup tracks randomly fall into the isolation cone must also be taken into account, potentially reducing

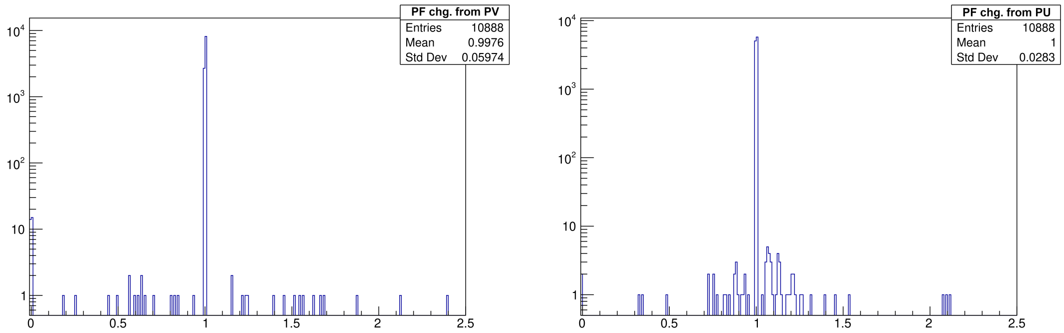


Figure 5.37: Ratio between the standard and recomputed isolation, where the latter is computed at high-level reproducing the default splitting into charged and PU components, for the charged (left) and PU (right) components (the standard isolation variable is computed using low-level variables).

the pileup contamination even in the event of close primary-vertices. This could allow us to afford a more relaxed ΔZ cut. Figure 5.38 shows the ratio between the new definition of isolation and the standard one as a function of p_T and η on the same Z boson events: the efficiencies are about 1% lower at low p_T and almost the same for higher p_T. This tells us that the new definition of isolation behaves similarly to the standard one, with only a slight loss of identification efficiency.

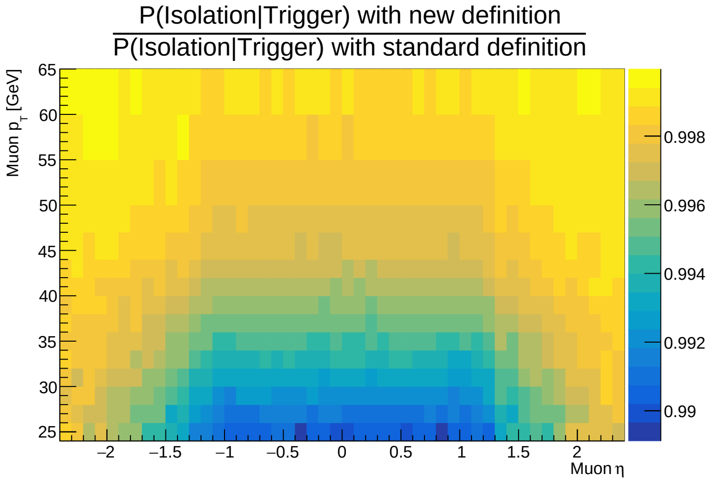


Figure 5.38: Ratio of the isolation efficiencies for the new isolation definition and the standard computation as a function of the muon $p_T - \eta$.

Therefore, this is the isolation definition that is actually used in the analysis, given its robustness with respect to the choice of main primary vertex.

Chapter 6

Measuring the mass of the W boson

A historical summary of the discovery of the W boson, and the first measurements of its mass, are presented in the first part of this chapter. In the second part, the most precise m_W measurements at lepton and hadron colliders are described in detail, showing the experimental techniques employed by the various experiments and their precision, together with a description of the limiting sources of uncertainty.

6.1 The discovery

The $W^{\pm}$ and Z bosons were discovered at the SPS collider by the experiments UA1 and UA2 in 1983 [127, 128]. SPS [129] was a fixed-target collider in which protons circulated at energies up 450 GeV. It was modified so that proton and antiproton bunches could circulate in the same beampipe and collide at center-of-mass energies of $\sqrt{s} \approx 500 - 600$ GeV, with enough energy to produce W and Z bosons: in the context of the Standard Model, at the time of discovery m_W and m_Z were predicted to be in the ranges from 60 to 80 GeV and 75 and 92 GeV, respectively [130].

The W boson was discovered analyzing events with one high-p_T isolated electron, which is expected to have an energy distribution that peaks at $\approx m_W/2$ (the Jacobian peak), and large missing transverse energy with large angular separation from the electron: UA1 (UA2) reported the observation of 6 (4) of such events that have been interpreted as $q\bar{q}' \rightarrow W^{\pm} \rightarrow e^{\pm}\overset{(-)}{\nu}_e$. Under this assumption, m_W can be estimated by a best fit to the distribution of the transverse momentum of the lepton of the decay (p_T) or of the transverse mass (m_T), i.e. the invariant mass considering only the components of the energy and momentum in the plane transverse to the beam axis:

$$m_T = \sqrt{2 p_T^{\ell} p_T^{\text{miss}} (1 - \cos \Delta\phi)}, \tag{6.1}$$

where p_T^{miss} is the absolute value of the negative vectorial sum of transverse momenta of all the reconstructed objects in the event:

$$\vec{p}_T^{\,\text{miss}} = - \sum_{\text{i=0}}^{\text{N. of particles}} \vec{p}_T^{\,i} \tag{6.2}$$

and $\Delta\phi$ is the angle between p_T^{ℓ} and p_T^{miss}. In this case $\ell = e$, but it can also be a muon. In the m_W fits performed by the UA1 and UA2 experiments, p_T^{miss} is measured from the sum of the transverse components of vectors associated with the individual calorimeter cells, with directions from the event vertex to the center of each of the cells and magnitude equal to the energy deposition in each cell. It will be explained in more detail in section 6.3 why the transverse mass is used to extract m_W. The first m_W measurement, published in the same papers announcing the discovery of W boson, were derived from fits to the p_T^{e} or p_T^{miss} distributions and had uncertainties in the range of $5 - 10$ GeV, where constraints from the distribution of the momentum imbalance, as probed by $|\vec{p}_T^{\,e} - \vec{p}_T^{\,\text{miss}}|$, and some theoretical assumptions based on QCD in the modelling of the distributions from which m_W was extracted were used to improve the accuracy. The final results from both UA1 and UA2, which were derived from fits to the transverse mass distribution, were published in 1985 [130], with a total number of 290 (251) $W^\pm \to e^\pm \overset{(-)}{\nu_e}$ candidates for UA1 (UA2) collaboration. These are:

$$m_W = 82.1 \pm 1.0_{\text{stat}} \pm 2.7_{\text{syst}}\ \text{GeV} \tag{6.3}$$

for UA1, and:

$$m_W = 80.2 \pm 0.8_{\text{stat}} \pm 1.3_{\text{syst}}\ \text{GeV} \tag{6.4}$$

for UA2. In both cases, the first error is statistical, and the second one is systematic, mostly stemming from the calorimeter energy scale uncertainty. During the last three years of collider operation (1988–90), UA2 collected large samples of $W^\pm \to e^\pm \overset{(-)}{\nu_e}$ and $Z \to e^+e^-$ decay events, which allowed them to perform an independent measurement of m_W:

$$m_W = 80.84 \pm 0.22\ \text{GeV}, \tag{6.5}$$

where the error is only statistical, and the most precise measurement of the ratio of m_W and m_Z, in which the systematic uncertainties from the calorimeter energy scale calibration cancel out to a large extent, obtaining:

$$\frac{m_W}{m_Z} = 0.8813 \pm 0.0036_{\text{stat}} \pm 0.0019_{\text{syst}}, \tag{6.6}$$

where the systematic uncertainty takes into account calorimeter non-linearities. By combining this result with the most precise measurement of m_Z performed at the LEP collider in 1991 ($m_Z = 91.175 \pm 0.021$ GeV) an estimate of:

$$m_W = 80.35 \pm 0.33 \pm 0.17\ \text{GeV} \tag{6.7}$$

was obtained.

Figure 6.1 shows the transverse mass distributions obtained using the total statistics available from the UA1 and UA2 experiments.

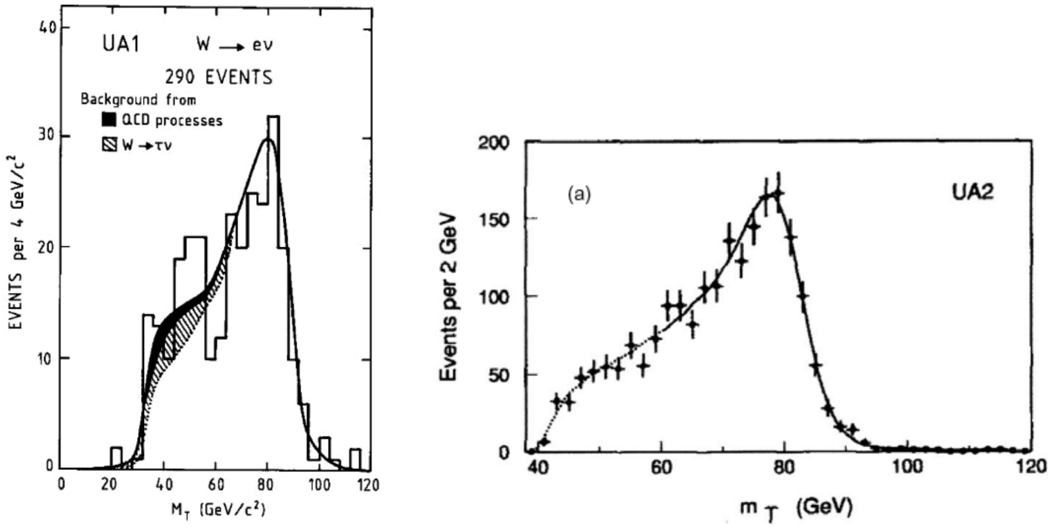


Figure 6.1: Transverse mass distribution for $W \to e\nu$ decays reconstructed by UA1 (left) and UA2 (right). The curve is the best fit to the experimental distribution using m_W as a free parameter. The backgrounds are also indicated for the UA1 measurement. From [130].

6.2 Electron-positron colliders

6.2.1 LEP [60]

The most precise measurements of the W boson mass at electron-positron colliders were made at the Large Electron-Positron (LEP) collider. LEP operated at center-of-mass energies around the $Z-$peak from 1989 to 1995, providing the most precise measurement of the mass of the Z boson currently available:

$$m_Z = 91.1876 \pm 0.0021 \text{ GeV} \tag{6.8}$$

and several other precision electroweak precision measurements, as discussed in 2.2, some of which are still among the most precise to date. The LEP accelerator was then upgraded to run at increasing energies in the range $\sqrt{s} = 130-209$ GeV, to test the electroweak sector of the Standard Model even further, such as to probe the existence of the Higgs boson. LEP stopped operating in 2000. In this second phase, referred to as LEP 2, the production of pairs of W bosons occurred through $e^+e^- \to W^+W^-$. The Feynman diagrams at lowest order in perturbation theory are shown in figure 6.2. Several properties of the W boson, among which its mass, were measured by the four LEP experiments: ALEPH [131], DELPHI [132], L3 [133], OPAL [134].

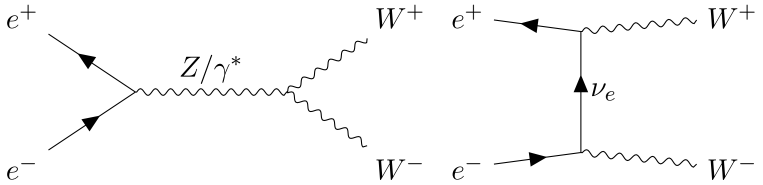


Figure 6.2: Tree level Feynman diagrams for $e^+e^- \to W^+W^-$: s-channel (left), t-channel (right).

In the earlier parts of LEP 2, m_W was extracted from the measured $e^+e^- \to W^+W^-$ production cross-section σ_{WW} as a function of the center-of-mass energy: in particular, for $\sqrt{s} \approx 2m_W$, the production cross-section $\sigma_{WW} \propto \beta = \sqrt{1 - \frac{4m_W^2}{s}}$, due to kinematics. The finite width of the W boson, Γ_W, and final-state radiation modify this relationship only slightly, and are taken into account in the theoretical uncertainties (which have an impact of 2% on the cross-sections). The center-of-mass energy where the cross-section is most sensitive to m_W is at $\sqrt{s} = 161$ GeV, determined from Monte Carlo simulations, but data at $172 - 183$ GeV were also analyzed. Only 3% of the total LEP 2 data was taken around these energies, amounting to a total integrated luminosity of about 10 pb^{-1} available to each experiment. These measurements were limited by statistics: the combination of all LEP measurements has an uncertainty that is slightly larger than 200 MeV, and the contribution from the uncertainty in the LEP center-of-mass energy is 30 MeV. Systematic uncertainties from hadronization and fragmentation effects in hadronically decaying W bosons, radiative corrections, final-state interactions are all negligible compared to the statistical uncertainty. Figure 6.3 shows examples of the measured WW production cross-section from ALEPH and OPAL.

The most precise m_W measurements at LEP 2 were obtained by reconstructing the decay products of the W bosons. The final state in which both W bosons decay leptonically, i.e. into a lepton and a lepton neutrino, lacks the theoretical uncertainties on hadronization or colour reconnection, which limit the precision with which m_W can be extracted from semi-leptonic (in which only one W boson decays leptonically) or fully hadronic decays (in which both W bosons decay into quarks). However, is limited by the low cross-section. OPAL determines a value of:

$$m_W(\ell\nu_\ell\ell\nu_\ell) = 80.41 \pm 0.41_{\text{stat.}} \pm 0.13_{\text{syst.}} \text{ GeV}, \tag{6.9}$$

which can still be used as a validation of the results from the other channels.

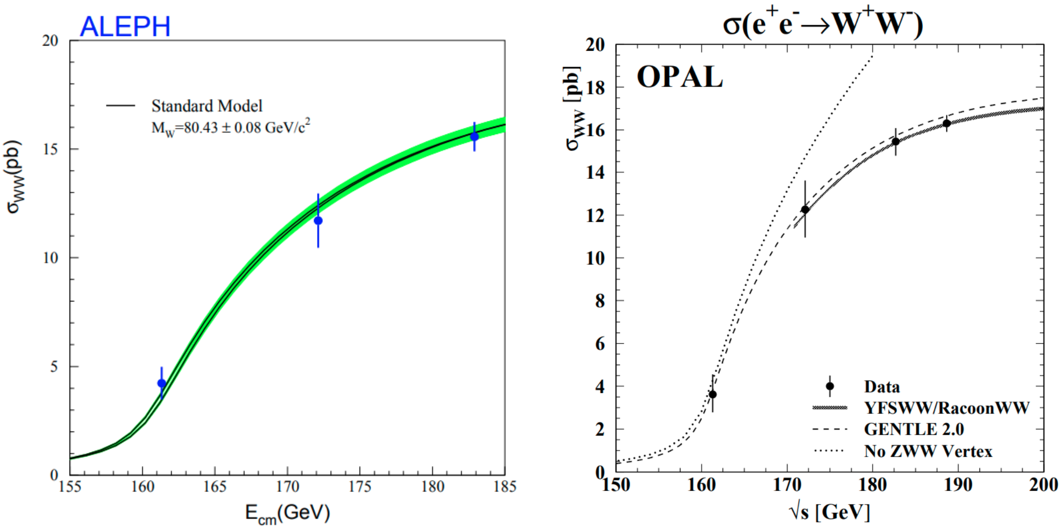


Figure 6.3: Measured σ_{WW} cross-sections as a function of the center-of-mass energy for ALEPH [135] (left) and OPAL [136] (right). The figure from ALEPH also shows the impact of a variation of $\Delta m_W = \pm 80$ MeV (the solid lines in black), and the impact on the cross-sections from the theoretical uncertainties (in green). The figure from OPAL shows the impact from an alternative theoretical calculation and from excluding the ZWW vertex from the σ_{WW} calculation.

$W^+W^- \to q\bar{q}q\bar{q}$ decays are reconstructed from hadronic jets, formed from measured particle tracks and energy depositions in the calorimeters. Different jest clustering algorithms are used, leading to different event topologies, with up to 5 jets in the final state. The detection of initial-state (ISR) or final-state photon radiation (FSR) can also be detected by the electromagnetic calorimeters of the experiments, and this information can be used to improve the mass resolutions of the candidate W bosons. Semi-leptonic W pair decays are reconstructed as a pair of hadronic jets, possibly with a third jet from gluon radiation, and an isolated lepton. In the leptonic $W \to e\nu_e$ and $W \to \mu\nu_\mu$ decays, the missing momentum vector is calculated applying total momentum conservation and is assigned to the momentum of the neutrino, so that the masses of both W decays can be reconstructed, while this procedure cannot be applied to cases in which one of the W bosons decays into a tau and a neutrino, so only the hadronically decaying W contains useful information about m_W.

The mass resolution of the di-jet system is mainly determined by the resolution of jet energy measurements, which is typically $\Delta E/E \approx 60\% - 80\%/\sqrt{E/\mathrm{GeV}}$, resulting in a di-jet mass resolution for W boson decays in the order of $8-9$ GeV. This can be substantially improved by adjusting the measured parameters of the jets and leptons to satisfy the constraints of energy and momentum conservation, compatibly with the measurement uncertainties of each of the reconstructed quantities. This method neglects the effects from ISR if not detected in the event, but this was found to have an impact that is smaller than the intrinsic jet energy resolution. In the $q\bar{q}q\bar{q}$ case, the improved kinematic reconstruction is referred to as a 4C fit, because there are four energy and momentum constraints, while in the $q\bar{q}e\nu$ and $q\bar{q}\mu\nu_\mu$ the three momentum components of the neutrino have to be determined, eliminating three of the constraints, and this reconstruction is referred to as 1C fit. A further improvement can be obtained by imposing that the masses of the two W bosons are equal, leading to a 5C or 2C fit. Although the assumption of equal mass between the two W boson candidates is not valid on an event-by-event basis, it is valid on average. m_W is then extracted from template fits to the invariant mass distributions obtained in the 5C or 2C fits, depending on the final state.

Each of the experiments had an integrated luminosity of about 750 pb^{-1} available for these measurements. Table 6.1 summarizes the measured value of m_W by the four LEP experiments: they are in agreement with each other and with the independent extraction from the σ_{WW} measurements and with the OPAL result that employs events in which both W bosons decay to leptons, which is then included in the semileptonic channel. ALEPH and OPAL have the smallest uncertainties.

Table 6.2 summarizes the sources of uncertainties and their impact on m_W, both for each of the $q\bar{q}\ell\nu_\ell$ and $q\bar{q}q\bar{q}$ channels separately and for their combination, combining the four LEP experiments together. The total uncertainty on m_W is mostly dominated by the semileptonic channel, as the precision in the fully-hadronic channel is limited by the modelling of colour reconnection and Bose-Einstein correlation uncertainties: the products of the two W decays, which are colour singlets, might have a significant space–time overlap, as the separation of their decay vertices, $\tau_W \sim \frac{1}{\Gamma_W} \approx 0.1$ fm, is smaller than the characteristic distance scales of ≈ 1 fm of the QCD interactions, which leads to a reorganisation of the colour flow between the decay products of the two W bosons, referred to as colour reconnection. Bose–Einstein correlations, instead, refer to the enhanced probability of producing pairs (or multiplets) of identical mesons close together in phase space, as a result of quantum-mechanical interference originating from the symmetry of the amplitude of the particle production process under exchange of identical mesons. Both of these effects have a direct impact on the reconstruction of the masses of the initial W boson, by affecting some of the constraints imposed in the fits (for example, the equal mass assumption), resulting in non-negligible variations in the templates used to in the 5C fits to extract m_W. These effects are absent in the semileptonic channel. The main systematic uncertainties common to both final states arise from:

- the uncertainty on the LEP energy of the beams, fully correlated across final states, roughly corresponding to rescaling the beam energy uncertainties by $2m_W/\sqrt{s}$ (for LEP 2 the beam energies have uncertainties in the $24 - 27$ MeV range);
- the calibration of the scale of the detectors, mostly correlated between final states;
- hadronization, mostly impacting the number of jets and their properties, larger for the fully hadronic channel;
- initial and final state radiation, which has an impact on the four-momentum constraints, resulting in a larger impact on the semileptonic channel, because these are used to determine the four-momentum of the neutrino.

Experiment	$W^+W^- \to q\bar{q}\ell\nu_\ell$ m_W (GeV)	$W^+W^- \to q\bar{q}q\bar{q}$ m_W (GeV)	Combined m_W (GeV)
ALEPH	80.429 ± 0.059	80.477 ± 0.082	80.444 ± 0.051
DELPHI	80.339 ± 0.076	80.310 ± 0.101	80.330 ± 0.064
L3	80.217 ± 0.071	80.324 ± 0.090	80.254 ± 0.058
OPAL	80.449 ± 0.062	80.353 ± 0.081	80.415 ± 0.052

Table 6.1: W mass measurements from direct reconstruction ($\sqrt{s} = 172 - 209$ GeV), for the semileptonic and fully-hadronic channels, and their combination. The $W^+W^- \to q\bar{q}\ell\nu_\ell$ results from the OPAL collaboration include mass information from the $W^+W^- \to \ell\nu_\ell\ell\nu_\ell$ channel.

Figure 6.4 shows two examples of invariant mass templates obtained by the ALEPH experiment [137], for the 5C fit and the 2C fit.

6.2.2 Future colliders: FCC [138]

The Future Circular Collider (FCC) [139, 140, 141] is the next collider designed to be hosted at CERN, and it is foreseen to be built after the High-Luminosity phase of the Large Hadron Collider (HL-LHC). It will consist of a 91 km tunnel located beneath the French departments of Haute-Savoie and Ain, and the Swiss

Source of uncertainty	Impact on m_W (MeV)		
	$q\bar{q}\ell\nu_\ell$	$q\bar{q}q\bar{q}$	Combined
ISR/FSR	8	5	7
Hadronisation	13	19	14
Detector effects	10	8	9
LEP energy	9	9	9
Colour reconnection	–	35	8
Bose-Einstein correlations	–	7	2
Other	3	10	3
Statistical	30	40	25
Statistical in absence of systematics	30	31	22

Table 6.2: Impacts on m_W from several sources of uncertainty for the combined LEP W mass from the direct reconstruction method. Detector effects include uncertainties in the jet and lepton energy scales and resolution. The 'Other' category refers to errors related to: simulation statistics, background estimation, four-fermion treatment, fitting method and event selection. Large correlated uncertainties, mainly from colour reconnection and Bose-Einstein correlations, lead to a reduced weight of fully-hadronic measurements contributing to the average result.

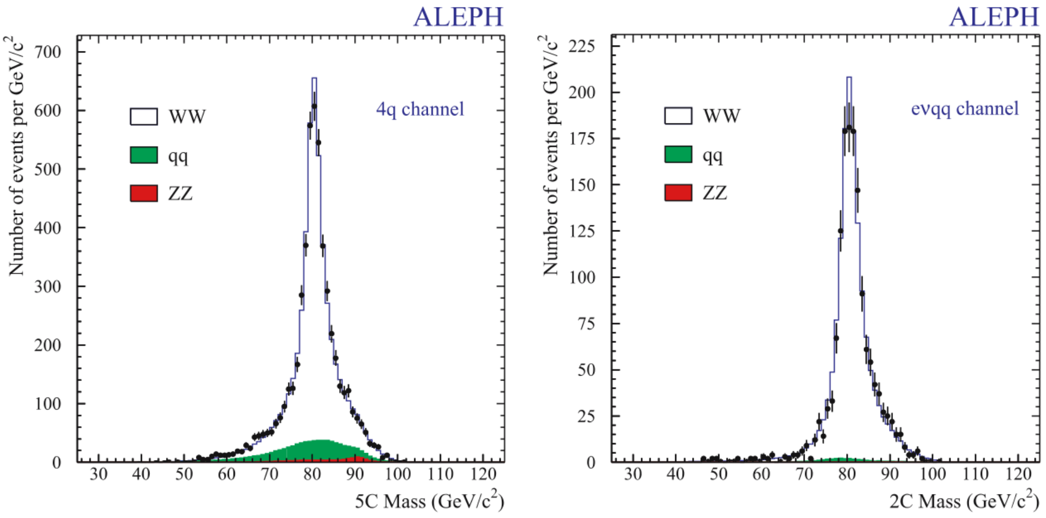


Figure 6.4: Examples of 5C (left) and 2C (right) invariant mass distributions from the ALEPH experiment. Non-WW backgrounds are also shown in the figures. From [137].

canton of Geneva, that will use the CERN accelerator complex to accelerate particles. In the first phase, expected to start around 2045, electron-positron collisions will occur at center-of-mass energies around the $Z-$peak, near the WW pair production threshold, and up to approximately twice the mass of the top quark. In the second phase, expected to start around 2070, proton-proton collisions will occur with center-of-mass energies up to $\sqrt{s} = 100\,$TeV, allowing the experiments to test the Standard Model at energies much higher than the current energy frontier.

Precision measurements of m_W using electron-positron collisions [138] are expected to be performed at center-of-mass energies around the W pair production threshold, mostly from the scan of the production cross-section as indicated for the LEP experiments, but with a much higher amount of statistics, for an expected integrated luminosity available to the experiments (4 are currently expected) of 19 ab^{-1}, for a total of 10^8 WW events. The expected statistical uncertainty, which was the limiting source of uncertainty for the LEP 2 measurements of m_W from the cross-section scan, is expected to be 0.25 MeV, and the expected systematic uncertainty is 0.3 MeV, limited by the precision with which the beam energy is measured using resonant depolarization [142, 143].

Figure 6.5 shows the expected limits on the oblique parameters (assuming that no deviation from the Standard Model is observed) including the results from m_W and other electroweak precision observables at FCC. The expected performance of linear colliders such as ILC [144, 145, 146, 147, 148] and CLIC [149]) are also shown.

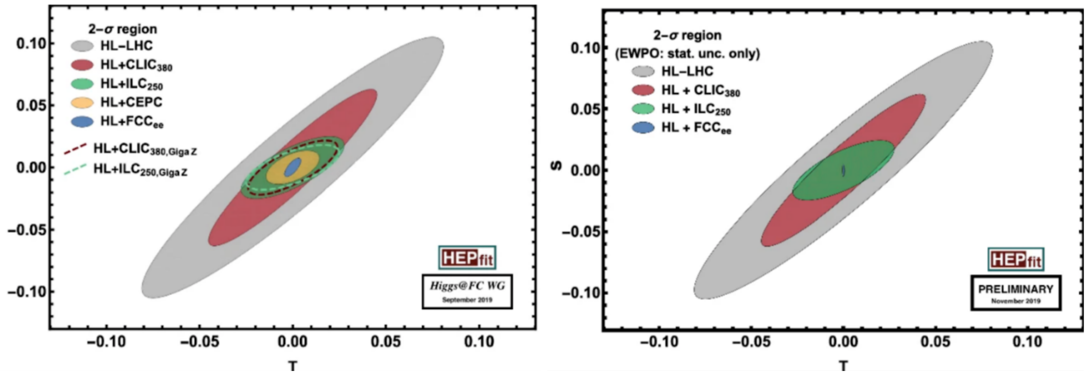


Figure 6.5: Expected uncertainties for the S and T parameters for various colliders, when they are used for e^+e^- collisions. The figure on the left shows the limits when the total uncertainties on each of the measurements to determine S and T are considered, while the figure on the right shows the same limits when only the statistical uncertainties are considered. From [138].

Other estimates, such as the one in reference [150] also include the predictions on the uncertainty on m_W extracted from the reconstruction of the final state of the $e^+e^- \to W^+W^-$ production and subsequent W boson decay, indicating that a similar uncertainty could be achieved, but this requires particular care in the calibration of the detector energy scale.

6.3 Hadron Colliders

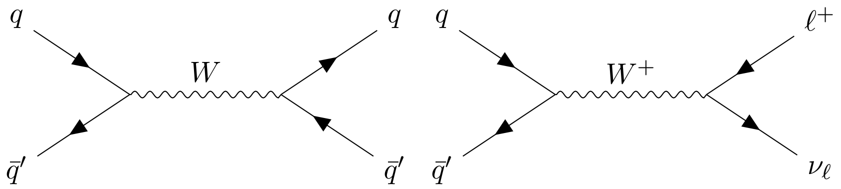


Figure 6.6: Tree level Feynman diagrams for $q\bar{q}' \to W \to q\bar{q}'$ (left) and $q\bar{q}' \to W^+ \to \ell^+\nu_\ell$ (right).

At hadron colliders, the mass of the W boson is inferred from the reconstructed final state in W boson decays. Figure 6.6 shows representative Feynman diagrams for the production of W bosons at hadron colliders and subsequent decay: on the left, the decay into hadrons is shown, on the right, the one into leptons. As discussed in section 3.1, protons are composite objects, so it is actually a pair of partons, i.e. quarks or gluons, that take part in the proton-(anti)proton collisions. Partons only carry a fraction of the total four-momentum of the protons. The probability density of carrying a fraction x of the total proton momentum is described by the *parton density functions* (PDFs). As a consequence, the total four-momentum of the initial state is not on an event-by-event basis. In particular, the total momentum along the longitudinal direction is not known.

As the protons in the beams have negligible motion in the transverse plane at the interaction point, it can be assumed that the total momentum of all the final-state particles produced in the collision, including both the hard-scattering interaction in which W bosons are produced and the interactions involving the proton remnants, in the transverse plane is zero.

Events in which the W boson decays into hadrons could be fully reconstructed, but the typical precision with which the jet energy scale is determined is not sufficient to provide competitive m_W measurements (typically at the percent level, to be compared with the target of $10^{-3} - 10^{-4}$ to be competitive with LEP). Furthermore, one cannot make use of kinematic constraints to improve the resolution on the invariant mass, unlike at lepton colliders, because of the missing information about the momenta in the longitudinal direction. In addition to this, partons are coloured, resulting in a large hadronic activity and number of jets reconstructed in the event.

Studies to assess the precision that could be obtained in future measurements from hadronic decays have been conducted. For example, reference [151] discusses the possibility to reach about 30 MeV on the difference between m_W and m_Z in boosted di-jet topologies, where the W and Z bosons are considered only if they have $p_T > 300 - 500\,$GeV, so that the background from QCD is largely reduced and the W and Z bosons are reconstructed as one jet with radius $\Delta R = 0.8$, while normally jets are reconstructed with a jet radius of $\Delta R = 0.4$. The integrated luminosity assumed in these studies is about 3 ab^{-1}, the one that should be achieved at the HL-LHC, and assumptions are made on the future improvement of the theoretical and experimental source of systematic uncertainties. This measurement should provide a value of m_W affected by different sources of systematic uncertainties compared to those that affect the other m_W measurements at hadron colliders.

Some studies in the context of rare W bosons decays into hadrons and photons, such as $W^\pm \to \pi^\pm\gamma$, have also been investigated [152, 153]: although their main aim is to measure the branching ratios into these decay channels, which could be enhanced by the presence of physics beyond the Standard Model, they could be also used as independent measurements of m_W, with the advantage that the final state is relatively clean and can be reconstructed very precisely.

Therefore, only events in which the W boson decays into leptons are considered, where the lepton is either an electron or a muon. The kinematic quantities used to extract m_W are those that involve only the transverse plane: transverse momentum of the lepton p_T^ℓ and the transverse mass m_T defined in equation 6.1. Sometimes also p_T^{miss} (see equation 6.2) is used either directly to extract m_W or as a cross-check of the results obtained with p_T^ℓ or m_T. The W boson mass is typically extracted from template fits to the p_T or m_T (or p_T^{miss}) distributions, also simultaneously, split by lepton charge. p_T^ℓ can be reconstructed with high precision, the scale, i.e. the average lepton momentum reconstructed by the detectors as a function of the real lepton momentum, and the resolution, i.e. the root mean square of the event-by-event difference between the real lepton momentum and the reconstructed lepton momentum, can be calibrated very precisely.

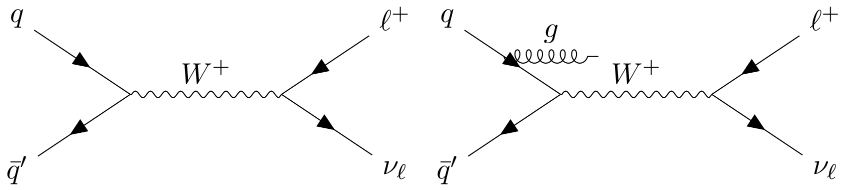


Figure 6.7: Tree level diagram (left) and an example of next-to-leading order (right) Feynman diagram for $q\bar{q}' \to W^+ \to \ell^+\nu_\ell$.

However, the p_T^ℓ distribution is sensitive to the intrinsic momentum q_T^W of the W boson. Figure 6.7 shows two examples of Feynman diagrams: on the left, the Feynman diagram corresponding to the lowest order of perturbation theory is shown, on the right, an illustrative next-to-leading order Feynman diagram. At lowest order in perturbation theory, $q_T^W = 0$, and p_T^ℓ depends only on the decay angle with respect to the beam line of the lepton, in the center-of-mass frame of the W boson. This results in a well distinct shape of the p_T^ℓ distribution, with a peak at $p_T^\ell = m_W/2$ corresponding to the Jacobian peak of the distribution, and a small tail above $m_W/2$ due to the finite width of the W boson. The possibility of having a $q_T^W > 0$, for example as induced by gluon radiation on the right of figure 6.7, results in a higher percentage of events with $p_T^\ell > m_W/2$ and a reduced percentage of events with $p_T^\ell > m_W/2$.

A precise modelling of p_T^W has historically been a limiting source of systematic uncertainty in the determination of m_W, as the uncertainties in the determination of q_T^W result in variations on p_T^ℓ that can resemble m_W variations. The left of figure 6.8 shows the state-of-the-art q_T^W spectrum for the theoretical models indicated in the figure. It can be seen that it is distributed predominantly at low q_T^W values, with a peak around $q_T^W = 5$ GeV: in this region, dominated by low-energy emissions of quarks and gluons, which then hadronize to form the recoil, the q_T^W computation cannot be performed using fixed-order perturbation theory, due to the presence of logarithmic terms $\approx \ln(m_W/q_T^W)$ that are dealt with using resummation techniques [154, 155]. This leads to uncertainties that can be large for a m_W measurement targeting an uncertainty of 10 MeV, as this requires that the uncertainty in the q_T^W modelling for values of $q_T^W \lesssim 5$ GeV be less than 1%, and of comparable magnitude in q_T^W range of interest for the analysis ($q_T^W \lesssim 40$ GeV).

The right of figure 6.8 shows the ratio between $q_T^{W^+}$ and q_T^Z: experiments have measured the p_T^{ll} spectrum in $Z \to \ell\ell$ events, as it can be reconstructed very precisely from the reconstruction of both final-state muons, and then multiplied it by the q_T^W/q_T^Z ratio, with the p_T^V uncertainties expressed in terms of the μ_R and μ_F scales decorrelated in bins of p_T^W and polarization. However, the QCD scales do not capture the non-perturbative effects, and there is a large dependence on the final uncertainty in q_T^W on the degree of correlation between W and Z. In particular, the complexity of predicting the p_T^W/p_T^Z ratio lies in the different initial-state parton composition, and in the differences in the heavy-flavour mass effects, the collinear and transverse non-perturbative dynamics of the partons. In addition to this, as already stated, these are not physical parameters, and therefore they do not have a defined physical and statistical meaning, with an undefined impact on the precision with which m_W is measured.

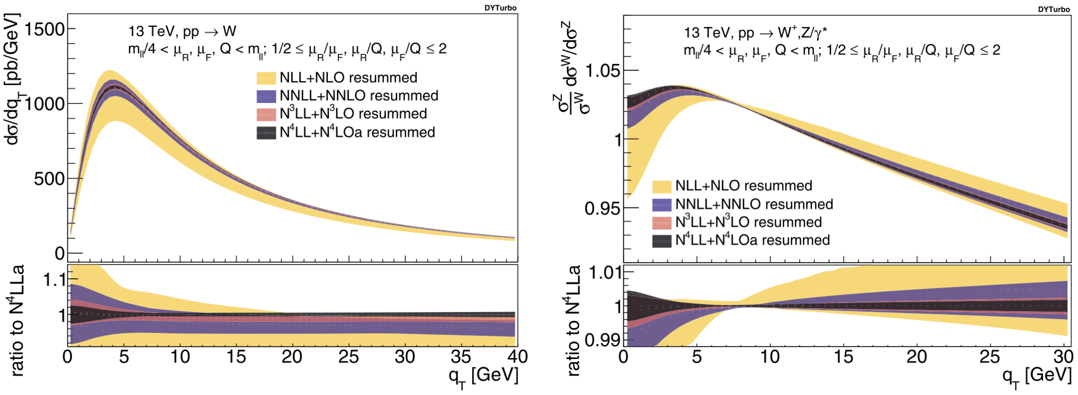


Figure 6.8: q_T^W distribution for the combination of W^+ and W^- (on the left) and ratio between the normalized $q_T^{W^+}$ and q_T^Z distributions (on the right) at $\sqrt{s} = 13$ TeV at calculated at NLL+NLO, at NNLL+NNLO, at N^{3}LL+N^{3}LO and N^{4}LL+N^{4}LOa computed using DYTurbo [156]. From [157].

The transverse mass is less dependent on p_T^W, as it is, at least theoretically, the invariant mass computed using only the momenta and energies projected in the transverse plane, and is invariant under Lorentz transformations in the transverse plane at first order in p_T^W/Q (where $Q \approx m_W$). However, the resolution in m_T is worsened by the resolution in the reconstruction of p_T^{miss}, resulting in generally worse experimental resolutions compared to p_T.

Parton density functions are another source of uncertainty that has historically limited the precision with which m_W can be measured at hadron colliders: there is a direct relation between the rapidity and polarization of the W boson and the $p_T - \eta$ distribution of the lepton from the decay. The rapidity of the lepton y_ℓ, which can be approximated very well by η, is related to the rapidity of the W boson y_W through:

$$y_\ell = \sinh^{-1}\left(\frac{m_W}{2p_T^\ell}\cos\theta^*\right) + y_W, \tag{6.10}$$

where $\cos\theta^*$ is the cosine of the polar angle of the lepton in the Collins-Soper frame of the W boson. In the limit where $q_T^W \to 0$, this becomes:

$$y_\ell = y_\ell^0 + y_W, \tag{6.11}$$

where y_ℓ^0 is the rapidity of the lepton in the center-of-mass frame of the W boson. From this, it can be inferred that changes in y_W have an effect on the $\eta = y_\ell$ distribution, introducing effects due to the limited acceptance of the detectors. Regarding the polarization, the approximation that production at lowest order in perturbation theory occurs can be made, as in the left of figure 6.7, and in this case the rapidity of the W boson is determined only by the partonic spin and the vectorial sum of longitudinal momenta of the partons (p_z). Due to the $V - A$ coupling between fermions and W bosons of the weak interaction, the spin of the W boson is aligned with that of the quark, which has negative helicity. It can be seen from figure 3.3 that valence quarks typically carry a larger fraction of the proton four-momentum compared to sea quarks, so that the W boson has negative helicity as the quark. In the decays, W^+ (W^-) tend to emit the charged lepton in the opposite (same) direction compared to the p_z^W direction in the majority of cases. The behaviour is different for the two charges. For inverted polarization of $W^\pm$, one gets the opposite behaviour in the decay. Figure 6.9 [158] summarizes the relationship between spins and p_z of the partons and those of the leptons in the decay in this case. Taking into account the q_T^W distribution, one introduces also the longitudinal polarization of the W, which is, however, still subdominant for the majority of W decays.

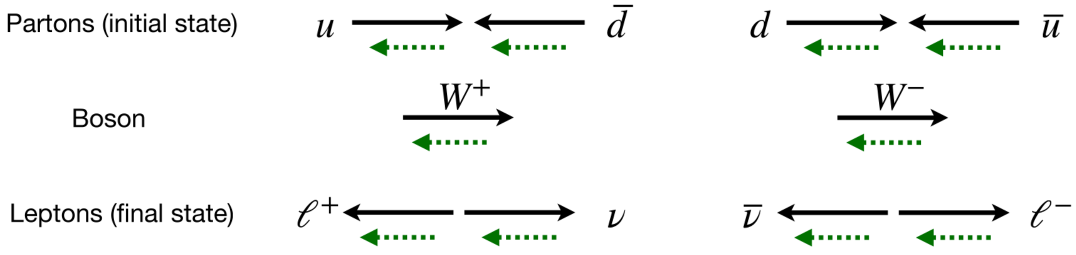


Figure 6.9: Scheme of the momentum (p_z, black arrows) and spin (s_z, green dashed arrows) alignment for partons, boson and lepton final state, assuming the most common production and decay channel, at $q_T^W = 0$. From [158].

The polarization of the W boson is then related to the decay lepton kinematics through:

$$\frac{d\sigma}{d\cos\theta^*} \propto \frac{3}{8}(1 \mp \cos\theta^*)^2 f_L(q_T^W, y_W) + \frac{3}{8}(1 \pm \cos\theta^*)^2 f_R(q_T^W, y_W) + \frac{3}{4}\sin^2\theta^* f_0(q_T^W, y_W), \tag{6.12}$$

where θ^* is the decay angle of the lepton (charged for W^-, neutrino for W^+) with respect to the $z-$axis in the rest frame of the $W^\pm$, and the upper (lower) sign correspond to W^+ (W^-) boson. The coefficients f_L, f_R, f_0 are defined to be $f_i > 0$ and $f_L + f_R + f_0 = 1$ and are functions of q_T^W and y_W: they quantify the probability that a W boson is produced in the left, right or longitudinal polarization. Figure 6.10 [159] shows the different shapes in the $p_T - \eta$ plane of the lepton as a function of the polarization and the charge of the W. In particular, the difference between W^+ and W^- reflects the $V - A$ structure in the decay, as seen in figure 6.9 (and this reflects in the plus or minus sign in equation 6.12).

Equation 6.12 is a particular projection of [160]:

$$\begin{aligned}\frac{d\sigma}{dq_T^W\,dy^W\,dM\,d\cos\theta^*d\phi^*} = &\frac{3}{16\pi}\frac{d\sigma^{\text{unpol.}}}{dq_T^W\,dydM}\times \\ &\{(1+\cos^2\theta^*)+A_0\frac{1}{2}(1-3\cos^2\theta^*)+A_1\sin 2\theta^*\cos\phi^*+ \\ &A_2\frac{1}{2}\sin^2\theta^*\cos 2\phi+A_3\sin\theta^*\cos\phi^*+A_4\cos\theta^*+ \\ &A_5\sin^2\theta^*\sin 2\phi^*+A_6\sin 2\theta^*\sin\phi^*+A_7\sin\theta^*\sin\phi^*\}\end{aligned} \tag{6.13}$$

valid at all orders in QCD, neglecting electroweak corrections. All coefficients are functions of p_T^W and y^W. When integrating over ϕ^*, only the terms multiplying A_0 and A_4 are nonzero, and the f_i coefficients are related to the A_i coefficients through:

$$A_0 \propto f_0 \qquad\qquad A_4 \propto -\frac{q}{e}(f_L - f_R), \tag{6.14}$$

where q is the charge of the W boson.

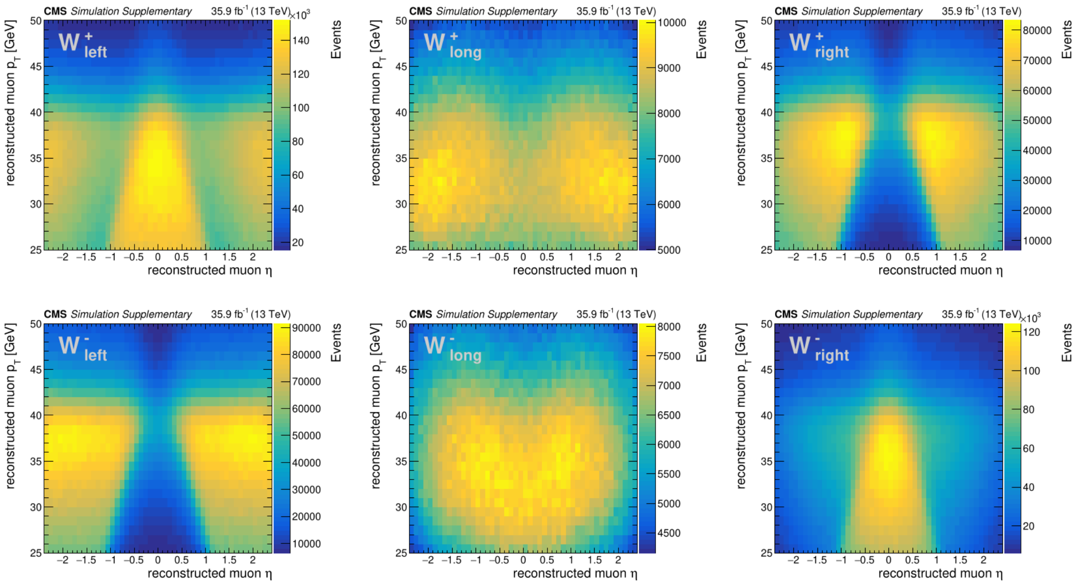


Figure 6.10: Distributions of $p_T^\mu - \eta^\mu$ from $W \to \mu\nu$ decays for W^+ (top) and W^- (bottom) in the left, longitudinal and right polarization states (from left to right). From [159].

Therefore, variations in the f_i (or A_i) coefficients have a direct impact on the p_T and m_T templates used to extract m_W. Other theoretical systematic uncertainties are related to final-state radiation from QED (FSR) and its interference with initial-state radiation for QED, and the pure weak corrections, all of which can be carefully modeled.

Experimental systematic uncertainties are related to:

- lepton reconstruction and identification and calibration of the momentum (or energy) scale, which needs to be known at the level of 10^{-4} for a 10 MeV uncertainty in m_W;
- muon momentum resolution;
- the calibration of the hadronic recoil, important to measure the transverse mass m_T precisely;

- the estimation of the backgrounds from QCD processes, which typically cannot be modeled by Monte Carlo simulation, unlike most of the other electroweak processes.

These uncertainties are specific to each experiment and different strategies being used.

In the following subsections, the most precise m_W measurements at Tevatron and LHC, before the publication of the CMS results, are discussed.

6.3.1 Tevatron

The Tevatron [161] was a circular proton-antiproton collider located at Fermilab (Chicago, IL). It operated at a center-of-mass energy $\sqrt{s} = 1.96$ TeV, between 1983 and 2011. Data were collected by the CDF [162, 163] and D0 [164, 165] experiments, which announced the first direct observation of the top quark in 1995 and the first observation of B_s oscillations in 2006, among many other scientific achievements.

D0

The most precise measurement of the W boson mass by the D0 experiment [74, 166] was performed based on a sample of 1.68×10^6 of $W^\pm \to e^\pm \overset{(-)}{\nu_e}$ events, selected by requiring:

- the presence of an electron with $p_T > 25$ GeV that passes shower shape and isolation requirements and has $|\eta| < 1.05$;
- $p_T^{\text{miss}} > 25$ GeV, where p_T^{miss} is defined as the negative vector sum of the energy deposits in the calorimeters taking into account their positions in the detectors;
- p_T^{recoil} defined by $\vec{p}_T^{\text{recoil}} = -\vec{p}_T^{e} - \vec{p}_T^{\text{miss}}$ to be less than 15 GeV;
- 50 GeV$< m_T < 200$ GeV.

The integrated luminosity of the sample used is 4.3 fb^{-1}. The backgrounds are $Z \to e^+e^-$ events in which an electron escapes detection, QCD events in which a jet is misidentified as an electron with p_T^{miss} arising from misreconstruction, and $W \to \tau\nu \to e\nu\nu\nu$ events, estimated using data-driven techniques or Monte Carlo generators as described in [167].

The RESBOS [168] event generator, combined with PHOTOS [169] is used to simulate the kinematics of the production and decay of W and Z boson, providing a next-to-leading order accuracy in the perturbative part of the calculation and including next-to-next-to-leading logarithm resummation of soft gluons, including also the generation of up to two final-state radiation photons. The non-perturbative part of the p_T^V calculation was tuned using Z data from previous data-taking runs at the Tevatron [170, 171]. The PDF set used in the generation is CTEQ6.6 [172]. The predictions are compared with those computed using the WGRAD [173] and ZGRAD [174, 175] event generators, which provide a more complete treatment of electroweak corrections at the one-radiated photon level, in order to assess the uncertainty on the m_W measurement due to quantum electrodynamics (QED) corrections. A fast, parameterized Monte Carlo (MC) simulation (FASTMC) is used to simulate electron identification efficiencies and the energy response and resolutions of the electron and recoil system.

$Z \to e^+e^-$ events are used to calibrate the response and resolution of the electromagnetic calorimeter, and also the hadronic recoil, using the m_Z result from LEP [33]. For the electromagnetic calorimeter response, the functional form

$$R_{\text{EM}} = F_{\eta-eq}(\eta) \times F_{\text{HV-loss}}(L, \eta) \times (\alpha(E_0 - \bar{E}_0) + \beta + \bar{E}_0) \tag{6.15}$$

is assumed, where $F_{\text{HV-loss}}(L, \eta)$ takes into account the dependence on the instantaneous luminosity L and is derived from ZeroBias events, $F_{\eta-eq}(\eta)$ takes into account the differences in response as a function of η, and the α and β are the scale and offset terms, with $\bar{E}_0 = 43$ GeV, a constant introduced to reduce the correlation between α and β. $F_{\eta-eq}(\eta)$ is determined from $Z \to e^+e^-$ invariant mass fits in $|\eta|$ categories, and α and β are determined from $Z \to e^+e^-$ invariant mass fits that are run separately in bins of instantaneous luminosity. The electromagnetic calorimeter resolution is parameterized as:

$$\frac{\sigma_{\mathrm{EM}}}{E_0} = \sqrt{C^2_{\mathrm{EM}} + \frac{S^2_{\mathrm{EM}}}{E_0} + \frac{N^2_{EM}}{E^2_0}}, \tag{6.16}$$

in which C_{EM}, S_{EM} and N_{EM} correspond to the constant, sampling and noise terms, respectively, derived taking into account their dependence on η. The hadronic response (resolution) is calibrated using the mean (width) of the η_{imb} distribution in $Z \to e^+e^-$ events in bins of p_T^{ee}, where:

$$\eta_{\mathrm{imb}} = (\vec{p}_T^{\,ee} + \vec{p}_T^{\,\mathrm{recoil}}) \cdot \hat{\eta}, \tag{6.17}$$

where $\hat{\eta}$ is the bisector of the electron and positron directions. A parameterization as a function of p_T^{ee}, which takes into account different contributions to the hadronic energy (e.g. the hadronic system recoiling from the W boson, the effects of low-energy products from spectator parton collisions and other beam collisions, final-state radiation), is used to give a smooth set of corrections.

The left of figure 6.11 shows the $Z \to ee$ invariant mass distribution for data and simulation after the calibrations are applied. Good agreement was found.

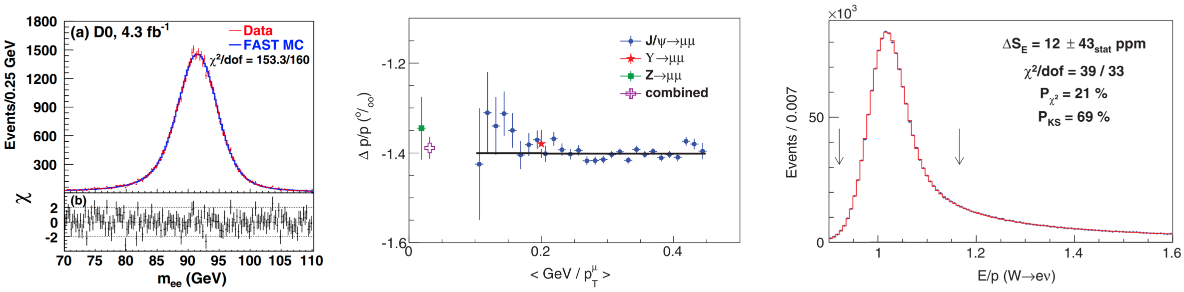


Figure 6.11: Left: is used to calibration the electromagnetic energy in $Z \to e^+e^-$ data and from the FASTMC and (b) the χ values, where $\chi_i = [N_i - (\mathrm{FASTMC}_i)]/\sigma_i$ for each bin in the distribution, N_i and $\mathtt{FASTMC}_i$ are the data and FASTMC template yields in bin i, respectively, and σ_i is the statistical uncertainty in bin i for the D0 measurement (from [74]). Middle: Fractional deviation of momentum $\Delta p/p$ derived from fits to the $J/\Psi \to \mu\mu$ resonance peak as a function of $\langle 1/p_T \rangle$ (blue circles) for the CDF measurement. A linear fit to the points, shown in black, has a slope consistent with zero (17 ± 34 keV). The corresponding values of $\Delta p/p$ extracted from fits to the $\Upsilon \to \mu\mu$ and $Z \to \mu\mu$ resonance peaks, and their combination (applied to the lepton tracks in W data) are also shown (from [3]). Right: Distribution of E/p for the $W \to e\nu_e$ data (points) and the best-fit simulation (histogram) including the small background from hadrons misreconstructed as electrons for the CDF measurement. The arrows indicate the fitting range used for the electron energy calibration (from [3]).

$Z \to ee$ events are also used to correct for the electron reconstruction and identification efficiencies, using the standard tag-and-probe procedure already described in 5.2.1.

The W boson mass is extracted from binned-likelihood fits to the p_T^e, m_T and p_T^{miss} distributions. Figure 6.12 shows the distributions used to extract m_W for data and simulation. The results and fitting ranges used in these fits are summarized in table 6.3. The systematic uncertainties are summarized in table 6.4. They are also grouped into experimental and theoretical uncertainties. It can be seen that the largest contribution on the experimental uncertainty comes from the electromagnetic calorimeter scale and resolution calibration, regardless of the variable used, while a large contribution from the recoil calibration can be observed for p_T^{miss}. Concerning theoretical uncertainties, the largest uncertainties come from the PDFs, while the uncertainties from the boson p_T are at most 5 MeV, stemming from the propagation of the uncertainties in the parameters used to model the p_T^V spectrum [166].

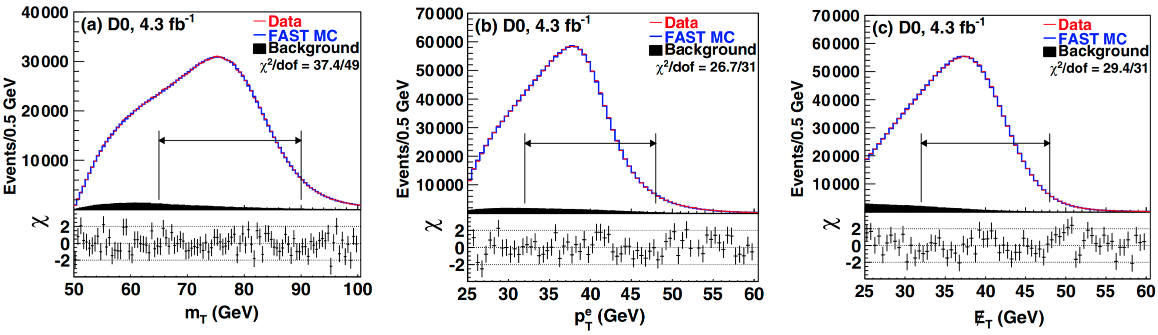


Figure 6.12: The m_T (left), p_T^e (middle), and (c) p_T^{miss} distributions for data and FASTMC simulation with backgrounds. The χ values are shown below each distribution. The fit ranges are indicated by the double ended horizontal arrows. From [74].

Variable	Fit Range (GeV)	m_W (MeV)	χ^2/ndof
m_T	$65 < m_T < 90$	80371 ± 13	$37.4/49$
p_T^e	$32 < p_T^e < 48$	80343 ± 14	$26.7/31$
p_T^{miss}	$32 < p_T^{\text{miss}} < 48$	80355 ± 15	$29.4/31$

Table 6.3: Results from the fits to data from the D0 m_W measurement. The uncertainty is solely due to the statistics of the W boson sample. From [74].

The measurement results in table 6.3 are combined to give:

$$\begin{aligned} m_W^{\text{D0}} &= 80.367 \pm 0.013_{\text{stat}} \pm 0.022_{\text{syst}} \text{ GeV} \\ &= 80.367 \pm 0.026 \text{ GeV}, \end{aligned} \tag{6.18}$$

where most of the uncertainty and central value is driven by the fit to the m_T distribution.

This is then combined with an earlier m_W measurement performed using 1 fb^{-1} of data, corresponding to a sample of 4.99×10^5 events [167], to obtain:

$$\begin{aligned} m_W^{\text{D0,combined}} &= 80.375 \pm 0.011_{\text{stat}} \pm 0.020_{\text{syst}} \text{ GeV} \\ &= 80.375 \pm 0.023 \text{ GeV}, \end{aligned} \tag{6.19}$$

which is in agreement with the Standard Model prediction.

CDF

The latest m_W measurement by CDF [3] was performed on 8.8 fb^{-1} of data, on a dataset of approximately 4 million selected W boson decays, analysing both decays to (anti-)muon and (anti-)electron. Data were selected by a trigger that requires the presence of a (anti-)muon or (anti-)electron with $p_T > 18$ GeV. The muon candidates must have $p_T > 30$ GeV, and pass tight reconstruction quality and energy deposition criteria. The electron candidates must have an electromagnetic calorimeter deposition of 30 GeV matched to a track with $p_T > 18$ GeV, with requirements on the shape of the calorimeter shower and the matching between energy deposition and reconstructed track p_T ($E_T/p_T < 1.6$). In both cases, leptons must have $|\eta| < 1$. Events with a second lepton of the same flavour as the one used in the analysis are rejected to suppress the $Z \to \ell\ell$ background. Events are further required to pass 30 GeV$< p_T^\ell < 55$ GeV,

Source	Impact on m_W (MeV)		
	m_T	p_T^e	$p_T^{\rm miss}$
Experimental subtotal	18	20	24
Electron energy calibration	16	17	16
Electron resolution model	2	2	3
Electron shower modelling	4	6	7
Electron energy loss model	4	4	4
Hadronic recoil model	5	6	14
Electron efficiencies	1	3	5
Backgrounds	2	2	2
Production subtotal	13	14	17
PDF	11	11	14
QED	7	7	9
Boson p_T	2	5	2
Total	22	24	29

Table 6.4: Table summarizing the several sources of systematic uncertainty for the D0 m_W measurement, grouped also into experimental and theoretical (production). From [74].

30 GeV $< p_T^{\rm miss} < 55$ GeV, where the definition of $p_T^{\rm miss}$ is the same as for D0, $p_T^{\rm recoil} < 15$ GeV and 60 GeV$< m_T < 100$ GeV.

The W boson mass is extracted from template fits to p_T^ℓ, m_T and $p_T^{\rm miss}$, as in the D0 analysis. The templates for the W and Z process are simulated using ResBos [168], using the data in [176, 170] and the $p_T^{\ell\ell}$ spectrum from $Z \to \ell\ell$ events from the same 8.8 fb^{-1} data used in the analysis to tune the non-perturbative part of the p_T^V modelling. These calculations are compared with those from DYQT [177, 178], which provides a next-to-next-to-leading order accuracy in the perturbative part of the calculation and including next-to-next-to-leading logarithm resummation of soft gluons. The measured p_T^W spectrum, where $p_T^W = p_T^{\rm recoil}$, is used to significantly constrain the uncertainties from the renormalization μ_R, factorization μ_F and resummation Q scales on the p_T^W/p_T^Z ratio, increasing the precision with which m_W can be measured. The radiation of final-state photons is implemented through Photos [169], which is calibrated to the more accurate predictions from HORACE [179, 180]. The PDF set used to generate the events and extract m_W is NNPDF3.1 [181], while CT18 [182] and MMHT2014 [183] have been used as alternative sets: the m_W values extracted using these sets are compatible with NNPDF3.1 within the PDF systematic uncertainties (3.9 MeV). The simulations are then interfaced with a custom simulation that includes a detailed calculation of the lepton and photon interactions in the detector, as well as models describing their individual position measurements within the COT (the CDF time-projection chamber used in track reconstruction), with the relative alignment of the wires determined from data as described in [3].

The muon momentum scale and resolution are calibrated from invariant mass fits as a function of the p_T of the muons to the J/Ψ and $Y(1S)$ resonances. Corrections for small inaccuracies in the magnetic field map, the COT position measurements, and the modelling of the energy loss by particles traversing the detector are derived and applied. A mismodelling of the energy loss would lead to a bias linear in $\langle 1/p_T\rangle$ of the two muons, and no such bias is observed after the corrections are applied, as shown in the middle of figure 6.11. The $\Delta p/p$ corrections extracted from the individual J/Ψ and $Y(1S)$ invariant mass fits are found to be consistent with each other, so these results are combined. After these corrections are derived, $Z \to \mu\mu$ events are used to fit m_Z from the invariant mass distribution, as a cross-check of the calibration procedure, obtaining:

$$m_Z^{\rm CDF}(\mu\mu) = 91192.0 \pm 6.4_{\textsf{stat}} \pm 4.0_{\textsf{syst}}\ \text{MeV}, \tag{6.20}$$

compatible with the measurement from LEP, so that also the calibration constant using the Z data is added to the total calibration. The calibration on the muon scale is the used to calibrate the electromagnetic energy scale using the peak of the E/p distribution for the W, shown on the right of figure 6.11, and Z events. This is done as a function of the electron E_T. The tail of the E/p distribution is used to correct for the amount of radiative material traversed in the tracking volume. A fit to the m_{ee} distribution from $Z \to ee$ events is used to extract m_Z, as a cross-check of the calibration procedure, obtaining:

$$m_Z^{\mathrm{CDF}}(ee) = 91194.3 \pm 13.8_{\text{stat}} \pm 7.6_{\text{syst}}\ \text{MeV}, \tag{6.21}$$

also consistent with the LEP measurement, so that also these measurement is also used in the calibration of the energy scale. The width of the E/p resolution is used to calibrate the electromagnetic energy resolution.

The hadronic recoil scale and resolution are calibrated taking into account different sources of contamination: the contributions from the spectator partons in the proton and antiproton, as well as the additional $p\bar{p}$ interactions in the same collider bunch crossing, which significantly degrade the p_T^{recoil} resolution, are measured from ZeroBias events. The imbalance of p_T^{recoil} and p_T^Z, projected along the parallel and orthogonal directions with respect to p_T^Z, in $Z \to ee$ events, is used to measure the calorimeter response to, and resolution of, the initial-state QCD radiation accompanying boson production. The definition of p_T^{recoil} used in the CDF analysis does not include the contribution from charged electrons, which then needs to be carefully determined not to bias its distribution: this is done using W boson data.

The lepton selection criteria in the CDF analysis are degraded by the presence of nearby energy associated with the hadronic recoil. The tag-and-probe method is used to derive the selection efficiencies as a function of u_T, and then an analytic parameterization is used, and this is done separately for electrons and muons. The efficiency for reconstructing electrons depends on η due to the trigger requirement in the analysis, as the presence of a calorimeter deposit matched to a track is required to select the events: this efficiency is measured using W boson events collected with a trigger with no track requirement, and is directly sensitive to the geometrical structure of the time-projection chamber used by CDF.

The W boson samples contain a small contamination of background events arising from QCD jet production with a hadron misidentified as a lepton, $Z \to \ell\ell$ decays with only one reconstructed lepton, $W \to \tau\nu \to \ell\nu\nu\nu$, pion and kaon decays in flight to muons (DIF), and cosmic-ray muons. The jet, DIF, and cosmic-ray backgrounds are estimated using data-driven techniques, while the $Z \to \ell\ell$ and $W \to \tau\nu$ backgrounds are estimated from simulation.

Figure 6.13 shows examples of templates for m_T, p_T^ℓ, and p_T^{miss}. Table 6.5 summarizes the measured values and fitting ranges for m_W in the various fits: they are all consistent with each other. The combined value:

$$\begin{aligned} m_W^{\mathrm{CDF}} &= 80433.5 \pm 6.4_{\text{stat}} \pm 6.9_{\text{syst}}\ \text{MeV} \\ &= 80433.5 \pm 9.4\ \text{MeV} \end{aligned} \tag{6.22}$$

is also shown, and it disagrees with the Standard Model at the 7σ level. The m_T, p_T^ℓ, and p_T^{miss} fits in the electron (muon) channel contribute weights of 30.0% (34.2%), 6.7% (18.7%), and 0.9% (9.5%), respectively. Table 6.6 summarizes the impact on m_W from several sources of uncertainties: from the theoretical point-of-view, the leading systematic uncertainty comes from the PDFs, as was the case for D0, but this is significantly smaller, as the NNPDF3.1 PDF set includes the wealth of data collected at the LHC, unlike CTEQ6.6, and QED final-state radiation. The impact from p_T^W modelling is around 2 MeV. From the experimental point-of-view, the leading uncertainties are related to the scale calibration (3.0 MeV), which are mostly independent for the electron and muon channel, so their impact on the combination of m_W is reduced. The experimental uncertainties from the background estimation have an impact of 3.3 MeV, and the impact from the hadronic recoil is around 2.0 MeV (combining both the scale and resolution).

6.3.2 LHC

The main differences between LHC and Tevatron are due to the fact that the first is a proton-proton collider with center-of-mass energies that are higher than those at the Tevatron, which has an impact on the

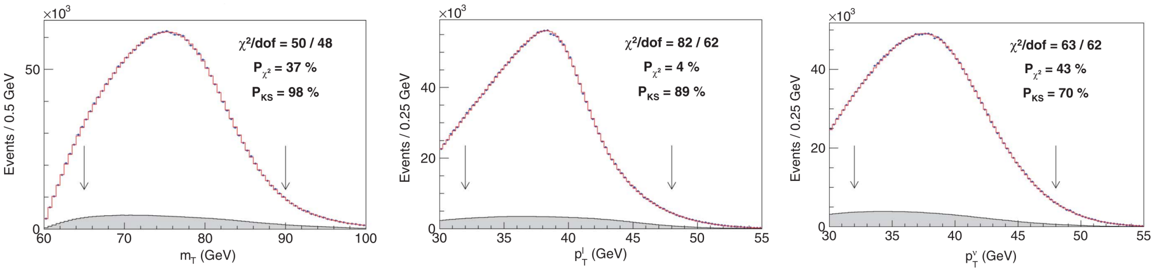


Figure 6.13: Distributions for m_T (left), p_T (middle), and p_T^{miss} (right) for the muon channel. The arrows indicate the fitting range. From [3].

Distribution	W boson mass (MeV)	χ^2/dof
$m_T(e,\nu)$	$80429.1 \pm 10.3_{\mathsf{stat}} \pm 8.5_{\mathsf{syst}}$	$39/48$
p_T^e	$80411.4 \pm 10.7_{\mathsf{stat}} \pm 11.8_{\mathsf{syst}}$	$83/62$
$p_T^\nu(e)$	$80426.3 \pm 14.5_{\mathsf{stat}} \pm 11.7_{\mathsf{syst}}$	$69/62$
$m_T(\mu,\nu)$	$80446.1 \pm 9.2_{\mathsf{stat}} \pm 7.3_{\mathsf{syst}}$	$50/48$
p_T^μ	$80428.2 \pm 9.6_{\mathsf{stat}} \pm 10.3_{\mathsf{syst}}$	$82/62$
$p_T^\nu(\mu)$	$80428.9 \pm 13.1_{\mathsf{stat}} \pm 10.9_{\mathsf{syst}}$	$63/62$
Combined	$80433.5 \pm 6.4_{\mathsf{stat}} \pm 6.9_{\mathsf{syst}}$	$7.4/5$

Table 6.5: Individual fit results and uncertainties for the m_W. From [3].

cross-section for $W^\pm$ production: W^+ and W^- have the same production cross-section at the Tevatron, since the collisions are charge symmetric, and a net asymmetry is observed in the η^ℓ distributions, since the W^+ (W^-) tend to be produced prevalently in the direction of the proton (anti-proton), while the W^+ has a larger production cross-section compared to the W^-, but the final-state lepton η^ℓ distributions are symmetric across the beamline. The larger center-of-mass energies result in larger production cross-sections compared to Tevatron, as can be seen in figure 3.3, and the luminosity is larger, so that more $W^\pm$ events are available.

The fact that the LHC is a proton-proton collider and that the center-of-mass energies are higher than at the Tevatron results in a larger impact on m_W from the PDF uncertainties: at the Tevatron, the valence quarks are more relevant for the $W^\pm$ production, while at the LHC the contribution from sea quarks, including the s, c and b quarks, becomes non-negligible, as smaller values of x (see section 3.1) are being probed, for which the PDFs have larger uncertainties.

ATLAS

ATLAS [37] performed a m_W analysis on data collected in 2011 at a center-of-mass energy of 7 TeV. They considered W boson decays to (anti-)muons and (anti-)electrons, with an integrated luminosity of 4.1 fb^{-1} and 4.6 fb^{-1} for the muon and electron channel, respectively. Their most precise result is a reanalysis performed on the same data used for the m_W measurement presented in 2017 [184].

Events were selected by triggers that require a muon with $p_T > 18$ GeV and $|\eta| < 2.4$ or an electron with $p_T > 20$ GeV, increased to 22 GeV in later data-taking periods to cope with increased instantaneous luminosity, and $|\eta| < 2.4$, excluding the range $1.2 < |\eta| < 1.8$. An isolation requirement is applied in both the electron and muon cases to suppress the backgrounds from QCD. A further selection is applied on the events:

- events must have exactly one lepton with $p_T^\ell > 30$ GeV;
- the leptons must be matched to the trigger objects;

Source	Impact on m_W (MeV)
Lepton energy scale	3.0
Lepton energy resolution	1.2
Recoil energy scale	1.2
Recoil energy resolution	1.8
Lepton efficiency	0.4
Lepton veto	1.2
Backgrounds	3.3
p_T^Z model	1.8
p_T^W/p_T^Z model	1.3
PDF	3.9
QED radiation	2.7
Statistics	6.4
Total	9.4

Table 6.6: Uncertainties on the combined m_W result. From [3].

- $p_T^{\mathrm{miss}} > 30$ GeV and $p_T^{\mathrm{recoil}} < 30$ GeV, where p_T^{miss} and p_T^{recoil} are defined from the calorimeter deposits excluding those with $\Delta R < 0.2$ with respect to muons or electrons;
- $m_T > 60$ GeV.

Approximately 5.89×10^6 candidate events are selected in the muon channel and 7.84×10^6 events in the electron channel.

The W boson mass is extracted from fits to the p_T^ℓ and m_T spectra separately for the charges and different $|\eta^\ell|$ categories. The new reanalysis of the ATLAS data uses profile likelihood fits, in which the sources of systematic uncertainties are implemented as nuisance parameters that are extracted together with m_W to improve the agreement with the observed data, reducing their impact on m_W. This is the same procedure the CMS used to the extract m_W, and will be described in more detail in 7.9. Their previous result, as well as the CDF and D0 results, was extracted by minimizing the χ^2 computed between data and Monte Carlo as a function of m_W, with the systematic uncertainties defined as shifts from the nominal m_W value from new minimum χ^2 fits in which the several source of uncertainty are shifted to their 1σ values. The difference between the nominal m_W value and the value extracted for each of the systematic uncertainties is added in quadrature to determine the total uncertainty.

The W (both for signal and background) and Z processes are simulated using the POWHEG Monte Carlo generator (v1/r1556) [185, 186, 187] interfaced to PYTHIA8 (v8.170) [188, 189] for the modelling of parton shower, hadronization and underlying event. Parton shower and underlying event parameters are set according to the AZNLO tune, after tuning to the $pp \to Z/\gamma^* \to \mu\mu$ distribution observed on 7 TeV data [190]. The predicted p_T^W was validated using data collected in conditions of low pileup (the average number of interactions per bunch-crossing is 2, compared to 9 in the dataset used for the ATLAS measurement) at $\sqrt{s} = 5.02$ TeV and $\sqrt{s} = 13$ TeV [191]. The angular coefficients A_i in equation 6.13 are corrected as a function of $p_T^{W,Z}$ and $y^{W,Z}$ to match the predictions from DYNNLO [192] to reach next-to-next-to-leading order precision in QCD, and they were validated using Z data collected at 8 TeV [193], with residual differences propagated as uncertainties. The CT10 PDF set [194] is used for the hard process, and the CTEQ6L1 PDF set [195] is used in the parton shower. The stability of m_W against several PDF sets, including a number of more up-to-date sets, was tested as described in [37]. The reference m_W value reported by the ATLAS collaboration is extracted using the CT18 [182] PDF set, which was found to give the most conservative PDF uncertainties and is able to cover the m_W values extracted from the other PDF sets considered in most cases. The POWHEG+PYTHIA simulation is interfaced with PHOTOS (v2.154) [169] to simulate QED final-state radiation. Next-to-leading order electroweak corrections are implemented through WINHAC [196], and the differences in final-state distributions with and without the electroweak corrections are used to define a source of systematic uncertainty, with others including the interference between ISR and FSR QED corrections (IFI), pure weak corrections due to virtual-loop and box diagrams, and final-state emission of lepton pairs.

The backgrounds in this analysis are from W decays into tau, which then decays into a muon and neutrinos, Z events with one undetected muon, top-quark pair and single-top-quark, Gauge-boson pair production (WW, WZ, ZZ), and QCD multijet and heavy-flavour decays. The electroweak backgrounds are modeled from simulations, while the QCD multijet and heavy-flavour decays are estimated using data-driven techniques. The response of the ATLAS detector is simulated using GEANT4 [197, 198].

The muon momentum scale and resolution are corrected from fits to invariant mass distributions for $Z \to \mu\mu$ events, as a function of η and ϕ of the muon. Corrections for biases in the radial direction, which originate from detector movements along the particle trajectory and are independent of the charge, and in the sagitta, which typically originate from curl distortions or linear twists of the detector around the $z-$axis and are charge-dependent, are applied according to the following formula:

$$\begin{aligned} p_T^{\mathrm{MC,corr}} &= p_T^{\mathrm{MC}} \times [1+\alpha(\eta,\phi)] \times [1+\beta_{\mathrm{curv}}(\eta) \times G(0,1) \times p_T^{\mathrm{MC}}] \\ p_T^{\mathrm{data,corr}} &= \frac{p_T^{\mathrm{data}}}{1+q\times\delta(\eta,\phi)\times p_T^{\mathrm{data}}}, \end{aligned} \tag{6.23}$$

where $p_T^{\mathrm{data,MC}}$ is the uncorrected muon p_T in data or simulation, $G(0,1)$ is normally distributed with mean equal to 0 and standard deviation equal to 1, $\alpha(\eta,\phi)$ is the muon momentum scale correction, $\beta_{\mathrm{curv}}(\eta)$ is the resolution correction, $\delta(\eta,\phi)$ is the sagitta correction. No correction for the energy loss, relevant at p_T values that are significantly lower than the range of muons in the analysis, is applied. The muon momentum scale α corrections are extrapolated from the $Z \to \mu\mu$ momentum range to the $W \to \mu\nu$ momentum range and the residual non-closure is evaluated as a function of $1/p_T$ in $|\eta|$ bins. The left of figure 6.14 shows the results of the extrapolation of the corrections from Z to W events: $\delta\alpha$ is compatible with one within two standard deviations, so that no residual correction is needed, and these results are implemented as systematic uncertainties in the fit. Corrections for the sagitta are derived both from $Z \to \mu\mu$ events and from $W \to e\nu_e$ events, where the sagitta is corrected from the E/p distributions, which has the advantage that the corrections are extracted for muons already in the range for the W analysis. The right of figure 6.14 shows the sagitta corrections as a function fo η^ℓ for both methods: it can be seen that they are in agreement, so they are combined.

The response of the electromagnetic calorimeter is calibrated using $Z \to ee$ events, after a set of intermediate calibrations (using $Z \to \mu\mu$ events and studying the shower profiles for electrons and photons), from fits to the dielectron invariant mass distributions, from which both the scale and resolution are extracted. $W \to e\nu_e$ are used to derive a set of residual corrections to cover for differences between Z and W processes, for example related to the different ϕ_ℓ (the corrections from the Z events are averaged over ϕ_ℓ). Figure 6.15 shows the relative energy scale as a function of ϕ_e for two $|\eta|$ bins.

The muon reconstruction, trigger, and isolation efficiencies are corrected for differences between data and simulation using the standard tag-and-probe method, as described in 5.2.1, as a function of p_T, η and charge of the probe lepton. For isolation, it was verified that its dependence on u_T is well reproduced by Monte Carlo simulation, so corrections are integrated over it. Figure 6.16 shows the efficiency Scale Factors averaged of η (on the left), and the η distribution after the application of efficiency Scale Factors for $Z \to \mu\mu$ events.

Electron efficiency corrections are determined using the tag-and-probe method using samples of $W \to e\nu_e$ (the tag here is the presence of missing transverse energy), $Z \to ee$, and $J/\Psi \to ee$ events, and measured separately for electron reconstruction, identification and trigger efficiencies, as a function of the electron η and p_T, with additional dedicated corrections for the trigger to correct for front-end board failures in the electromagnetic calorimeter, in a region corresponding approximately to $0 < \eta < 1.475$ and $-0.9 < \phi < -0.5$, that affected about 20% of the data used in the analysis.

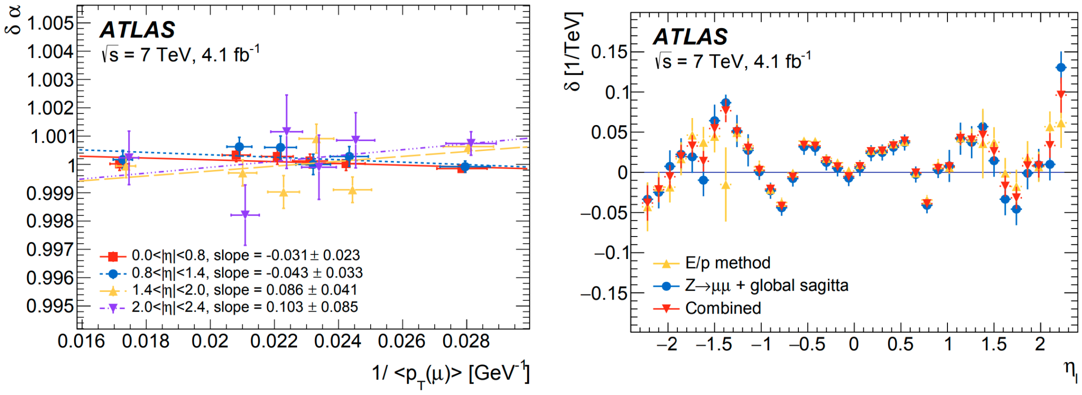


Figure 6.14: Left: residual muon momentum scale corrections as a function of muon $1/p_T$ in four η bins. Fits using a linear function that parameterises the extrapolation of the muon momentum scale correction from Z to W events is superimposed. Right: Sagitta bias, δ, as a function of η^ℓ. The results are obtained with the $Z \to \mu\mu$ and E/p methods and the combination of the two. The two measurements are combined assuming they are uncorrelated. From [184].

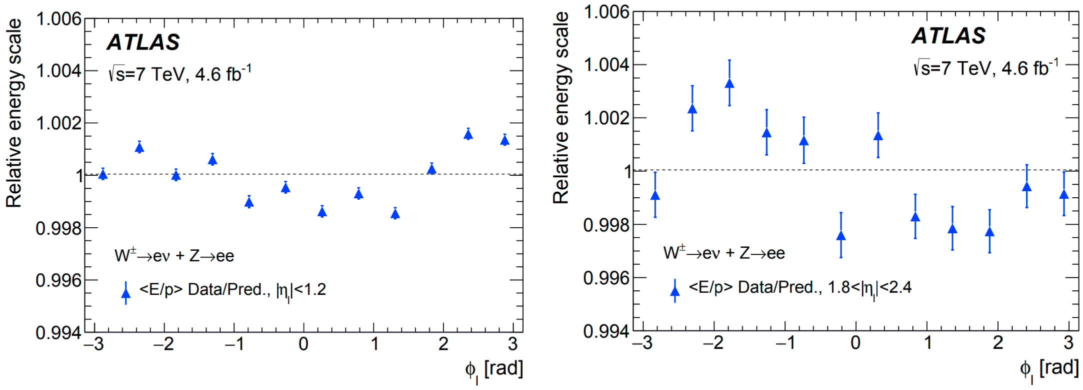


Figure 6.15: Azimuthal variation of the data-to-prediction ratio of $\langle E/p \rangle$ in W and Z events, for electrons in $|\eta_e| < 1.2$ (left) and $1.8 < |\eta_e| < 2.4$. The electron energy calibration based on $Z \to ee$ events is applied, and the track momentum is corrected for the momentum scale, resolution and sagitta bias. From [184].

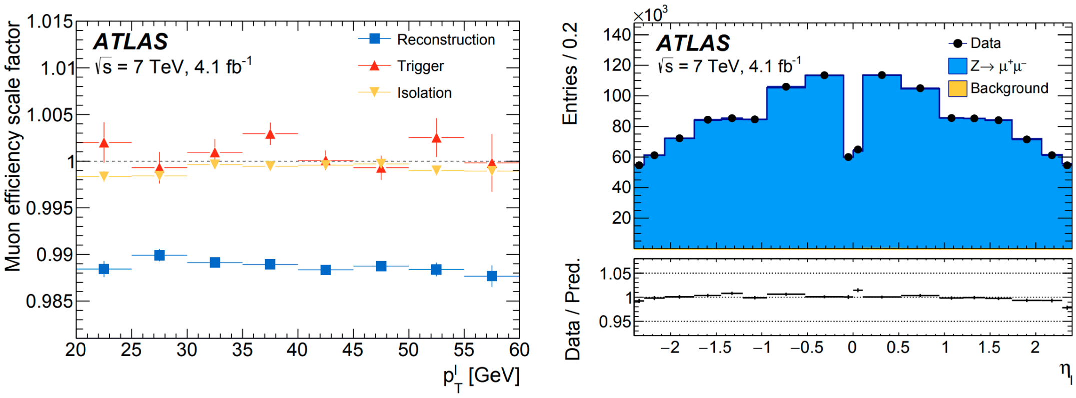


Figure 6.16: Left: reconstruction, trigger, and isolation Scale Factors for muons. Right: η_μ distribution from $Z \to \mu\mu$ events. From [184]

The hadronic recoil is calibrated using Z as a function of p_T^{ll} to reweigh the $\sum E_T^*$ distribution in simulation to match the one observed in data, after correcting for differences in the pileup distribution between data and Monte Carlo, and applied to the W events as a function of p_T^W at generator level, with a procedure to take into account possible differences in $\sum E_T^*$ response between W and Z events. After this set of corrections, the scale and resolution of the hadronic recoil are calibrated from Z events from the projections of $\vec{p}_T^{\,\mathrm{recoil}}$ on directions parallel and perpendicular to $\vec{p}_T^{\,ll}$, and are then applied as a function of p_T^W at generator level.

Figure 6.17 shows examples of p_T distributions for the electron and muon channels, for data and Monte Carlo before and after the fit. The post-fit agreement is at the per mille level. The value of m_W obtained by fitting simultaneously the p_T^ℓ and m_T spectra for both charges and for all η categories is:

$$\begin{aligned} m_W^{\mathrm{ATLAS}} &= 80366.5 \pm 9.8_{\mathsf{stat}} \pm 12.5_{\mathsf{syst}} \text{ MeV} \\ &= 80366.5 \pm 15.9 \text{ MeV} \end{aligned} \tag{6.24}$$

in agreement with the Standard Model prediction. Out of the several PDF sets tested, the maximum variation with respect to the nominal result was -20.9 MeV obtained when the NNPDF4.0 [199] set is used.

Table 6.7 summarizes the various sources of uncertainty in the analysis. The results are mostly driven by the p_T^ℓ fit, the relative weights of p_T^ℓ and m_T are 86% and 14% when the CT18 set is used, with a slightly larger contribution on the total uncertainty from the systematic variations, mostly dominated by the muon and electron calibrations, and the PDFs. For m_T, a large contribution to the uncertainty comes from the calibration of the hadronic recoil, which spoils the precision with which the W boson mass can be measured. This is largely due to the pileup at the LHC during Run 1, with an average number of interactions per bunch-crossing equal to 9.

The breakdown of uncertainties in table 6.7, the one presented in the final paper [37], relies on the so-called "global" impacts presented in [200]. Fluctuations in data, Monte Carlo and the nuisance parameters are, in this case, correlated, resulting in an increase of what is defined as statistical uncertainty, especially if the data are able to constrain the nuisance parameters, reducing their uncertainties compared to the ones provided as input to the fit. The impact of the systematic uncertainties is evaluated from the fluctuations of the external measurements corresponding to the nuisance parameters within their uncertainties.

An alternative definition of the breakdown of uncertainties is the "nominal" breakdown of the uncertainties by CMS presented in 7.10: in this approach, the statistical uncertainty is defined freezing the nuisance parameters at the post-fit values. The impact of the individual sources of systematic uncertainty is evaluated by repeating the fits to extract m_W with the nuisance parameters corresponding to the source under consideration fixed to their post-fit values, computing the corresponding total uncertainty, and then quadratically subtracting this uncertainty from the total uncertainty when all the parameters are allowed to float. In this case, the impact of the systematic uncertainties is larger than the one using the "global" impacts definition, while the statistical uncertainties are smaller.

In practice, in both cases only one fit is run, and the impacts are computed directly from the covariance matrix: defining the correlation matrix elements as $\rho_{\mu\theta}$, where μ is the parameter of interest (m_W) and θ is the nuisance parameter associated to a given source of systematic uncertainty, the global impacts are defined as:

$$\mathrm{I}_g = \sigma_\mu \rho_{\mu\theta} \sigma_\theta, \tag{6.25}$$

where σ_μ is the post-fit standard deviation on the parameter of interest, σ_θ is the post-fit uncertainty on the nuisance parameter, normalized in units where a 1σ deviation in the nuisance parameter corresponds to $\theta = 1$, while $\theta = 0$ corresponds to the nominal parameter. The statistical uncertainty is then defined as the difference in quadrature between the total and the quadratic sum of all the systematic uncertainty. The nominal impacts are defined as:

$$\mathrm{I}_N = \sigma_\mu \rho_{\mu\theta}, \tag{6.26}$$

which is equivalent to the definition of impacts defined in the text (in the asymptotic assumption). For the global impacts, these definitions are also equivalent to repeating the fits shifting the nuisance parameters by one pre-fit standard deviation, as described in [200].

The definition based on global impacts quantifies correctly the impact from the finite size of data (and Monte Carlo), as it captures the fluctuations of m_W if the fit is repeated on completely independent data (and Monte Carlo) of the same nature, while this is not the case for the "nominal" definition that does not take into account possible variations of the nuisance parameters, i.e. the fluctuations of the external measurements. In addition to this, the uncertainties defined from the "global" impacts can be summed in quadrature to obtain the total, unlike those obtained from the "nominal" ones. However, the global impacts definition does not properly capture the decrease in the uncertainty when a source is removed (for fixed amount of statistics) in a trivial way: while it is possible to estimate the effect of removing one uncertainty in the global impact method by repeating the fit without that uncertainty, it is not possible to obtain such an estimate simply neglecting the impact for a given nuisance or group of nuisances (i.e., by removing the impact for the given nuisance from table 6.7), due to the post-fit correlations between uncertainties.

The comparison between impacts can be useful to understand what sources of uncertainty are constrained from the fit, as is done in the m_W measurement by CMS.

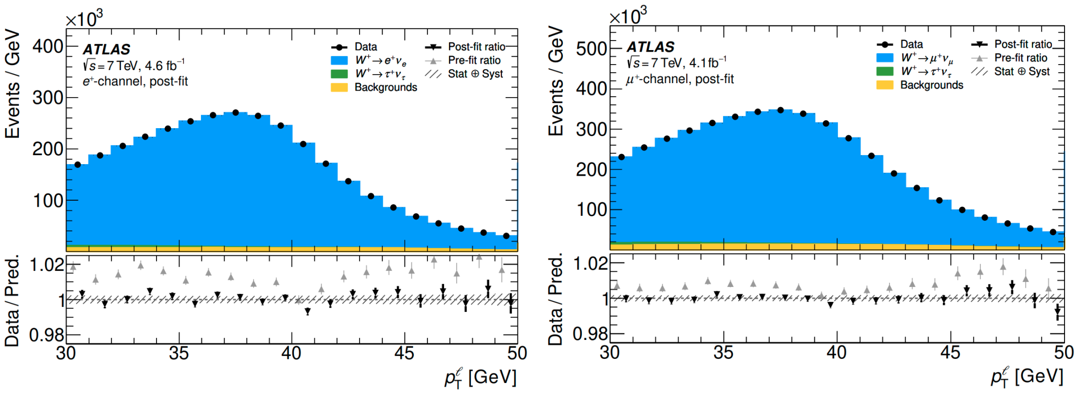


Figure 6.17: Post-fit and pre-fit p_T distributions for e^+ (left) and μ^+ (right). From [37].

Unc. [MeV]	Total	Stat.	Syst.	PDF	A_i	Backg.	EW	e	μ	p_T^{recoil}	Lumi.	Γ_W	PS
p_T^ℓ	16.2	11.1	11.8	4.9	3.5	1.7	5.6	5.9	5.4	0.9	1.1	0.1	1.5
m_T	24.4	11.4	21.6	11.7	4.7	4.1	4.9	6.7	6.0	11.4	2.5	0.2	7.0
Combined	15.9	9.8	12.5	5.7	3.7	2.0	5.4	6.0	5.4	2.3	1.3	0.1	2.3

Table 6.7: Summary of the sources of uncertainty for the ATLAS measurement, with the decomposition as described in [200]. The statistical uncertainty refers to both data and Monte Carlo. e and μ refer to the uncertainty from scale and resolution calibration and reconstruction and identification efficiencies. PS refers to the uncertainties in the PYTHIA8 tuning. From [37].

LHCb

The LHCb experiment is a single arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing b or c quarks. Unlike CMS and ATLAS (or the experiments at Tevatron), the LHCb subdetectors do not surround the interaction point, meaning that p_T^{miss}, and therefore m_T, cannot be reconstructed, so m_W is extracted only from the p_T information, in particular the template fits to extract m_W are performed on the q/p_T distribution, where q is the electric charge. However, the pseudorapidity range of LHCb allows them to cover a phase space of the proton-proton collisions that cannot be easily probed by ATLAS and CMS, being sensitive to a different $x-$range of the proton parton density functions. This allows LHCb to provide a cross-check for the other m_W measurements, which are performed in the central pseudorapidity range, and also to exploit anti-correlations in the PDF uncertainties that can reduce their overall impact on m_W, as shown in figure 6.18.

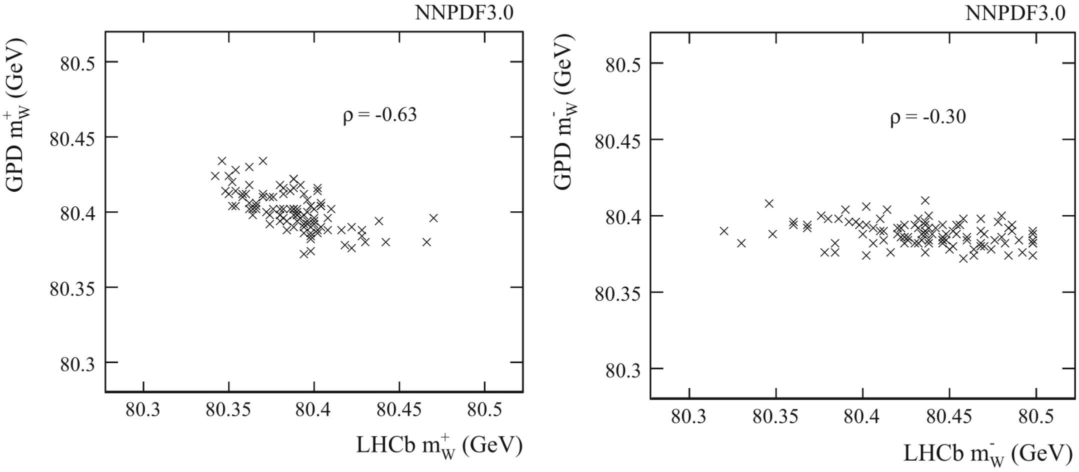


Figure 6.18: The fitted m_W for general purpose detectors (GPD) versus LHCb for each variation of the NNPDF3.0 set, and for (left) W^+ and (right) W^-. The anti-correlation ρ is superimposed in the plots. From [201].

LHCb focused on W boson decays into a (anti-)muon and a neutrino in their m_W measurement [38]. The LHCb analysis uses a sample of pp collected at $\sqrt{s} = 13$ TeV in 2016, corresponding to an integrated luminosity of 1.7 fb^{-1}. Events are selected by a trigger that requires a muon with $p_T > 6$ GeV, and the isolation variable defined in equation 5.1 within a radius of $\Delta R < 0.4$ is required to be less than 4 GeV (the absolute isolation is used here), and with tight identification requirement, for example on the impact parameter with respect to the primary vertex. Muons are required to have $2.2 < \eta < 4.4$, and events with at least two muons with $p_T > 25$ GeV are discarded in order to suppress the $Z \to \mu\mu$ background. This corresponds to roughly 2.4 million $W \to \mu\nu$ candidates in the range $28 < p_T < 52$ GeV, selected as the fitting range in the analysis. All signal and background processes are simulated using an LHCb specific tune [202] of Pythia version 8.186 [203]. The interaction of the generated particles with the detector is implemented through GEANT4 as described in [204].

The muon momentum scale and resolution are calibrated from the J/Ψ, Υ and Z resonances, on top of the calibration performed online [205, 206]. First, corrections for charge-dependent biases, related to the alignment of the tracker, are extracted from fits to the *pseudomass* distribution:

$$\mathcal{M}^{\pm} = \sqrt{2p^{\pm}p_T^{\pm}\frac{p^{\mp}}{p_T^{\mp}}(1-\cos\theta)}, \tag{6.27}$$

where $p^{\pm}$ and $p_T^{\pm}$ are the momenta and transverse momenta of $\mu^{\pm}$, respectively, and θ is the opening angle between the two muons. The pseudomass is not sensitive to the magnitude of the $\mu^{\mp}$ momentum, but only to alignment biases on $\mu^{\pm}$. The pseudomass is an approximation of the dimuon invariant mass in the limit of low p_T^Z. This is enforced by requiring that $\phi^* < 0.05$, where:

$$\phi^* = \frac{\tan((\pi - \Delta\phi)/2)}{\cosh(\Delta\eta/2)} \approx \frac{p_T^Z}{m_Z}. \tag{6.28}$$

Corrections are extracted as a function of η and ϕ, and the detectors which are crossed by the muons, and are shown in figure 6.19 for data and simulation. The momenta of the simulated muons are then

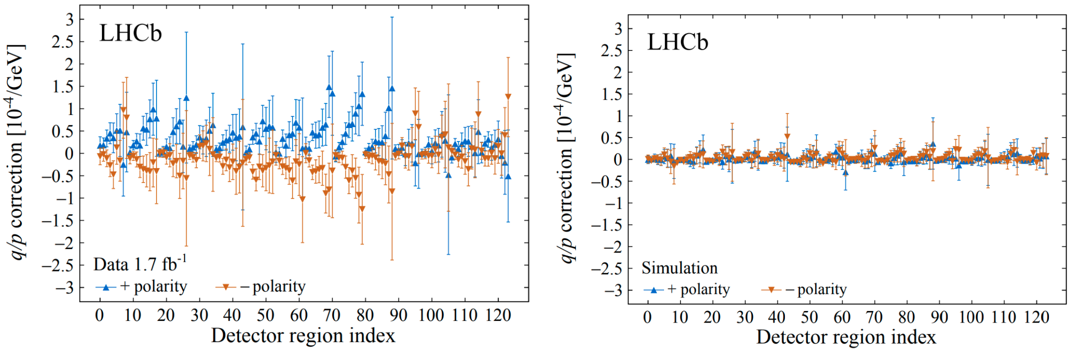


Figure 6.19: Curvature corrections as a function of the detector region index (depends on η, ϕ and tracking detector, as described in the text) for (left) data and (right) simulation. The periodic pattern corresponds to a dependence on η that repeats in the intervals of the azimuthal angle. From [38].

smeared to match those in the data according to:

$$\frac{q}{p_{\text{corr.}}} = \frac{q}{p \times G(1+\alpha, \sigma_{\text{MS}})} + G\left(\delta, \frac{\sigma_\delta}{\cosh \eta}\right), \tag{6.29}$$

where $G(a, b)$ represents a random number sampled from a Gaussian distribution with mean a and width b. The σ_{MS} and σ_δ parameters correspond to multiple scattering and curvature measurement contributions to the resolution, respectively. The parameters are determined in several η categories from a simultaneous fit of $J/\psi \to \mu\mu$, $\Upsilon(1S) \to \mu\mu$ and $Z \to \mu\mu$ candidates in data and simulation. Figure 6.20 shows the impact of these corrections.

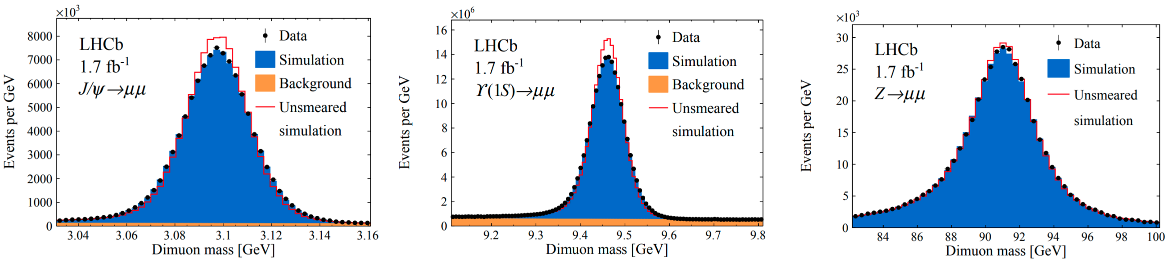


Figure 6.20: Dimuon mass distributions for selected J/ψ, $\Upsilon(1S)$ and Z boson candidates. The data are compared with the fit model. The red histogram corresponds to the model before the application of the smearing. From [38].

Efficiency corrections are extracted for trigger, identification, tracking and isolation using the tag-and-probe method on $Z \to \mu\mu$ and $J/\Psi \to \mu\mu$ events (for the trigger). Corrections for trigger, identification and tracking are extracted in bins of η, ϕ and p_T, while the isolation corrections are extracted as function of η and u_T. The efficiency corrections are shown in figure 6.21 for the trigger (on the left) and for isolation (on the right).

The QCD background, corresponding mostly to in-flight decays of pions and kaons into muons, is modeled using a sample of high-p_T tracks, selected by dedicated triggers without muon identification requirements, with the W boson selection requirements applied with the muon identification requirement inverted. The probability of an unstable hadron of mass m, lifetime τ, and momentum p to decay within a detector of length d is:

$$1 - \exp\left(-\frac{md}{\tau p}\right) \approx \frac{md}{\tau p}, \tag{6.30}$$

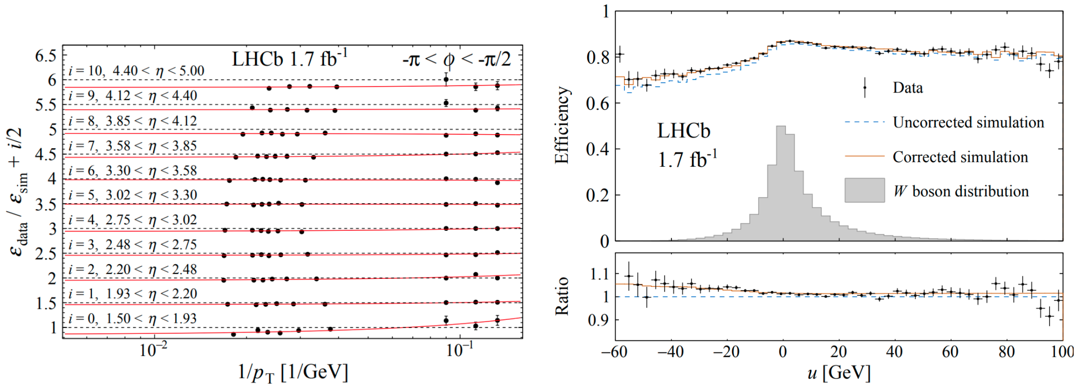


Figure 6.21: Left: trigger efficiency Scale Factors for one ϕ bin as a function of η and $1/p_T$. Right: isolation efficiency as a function of u_T (u in the plot), with and without the corrections, the W boson u_T distribution is superimposed. From [38].

Therefore, the in-flight decay background can be modeled by the data with weights of $1/p$, and the normalization is extracted together with m_W.

Different models are considered for the W and Z boson production, and weights to correct for the angular coefficients and QED final-state radiation are applied. The angular coefficients are corrected as a function of p_T^V, y^V and mass to the predictions from DYTurbo [156]. It was observed that the uncertainties on A_3 result in a large uncertainty on m_W: a freely floating nuisance fully correlated between W^+ and W^- is included in the fit. Several models for the p_T^V distribution are considered: the nominal model is chosen as the one that provides the best agreement for the p_T^Z distribution between data and Monte Carlo, after the parameters of the model are tuned from fitting the p_T^Z distribution up to 30 GeV. The final choice is Powheg [185, 186, 187] interfaced with Pythia8 [203]. For higher values, a comparison between data and Monte Carlo is performed and the difference between the p_T^Z prediction and data is propagated as a source of uncertainty. In the fit to extract m_W, the parameters describing the p_T^W spectrum are allowed to float freely. The impact of generator-choice is evaluated by repeating the fits to extract m_W using the results from the different predictions to compute p_T^W, with the envelop of the different m_W values assigned as a systematic uncertianty. Three different PDF sets are chosen in the event generation, NNPDF3.1 [181], CT18 [182], and MSHT20 [207], and the average of the three m_W values is taken as the final m_W value, with the PDF uncertainty assigned as the average of the three PDF uncertainties, assuming that they fully correlated across sets. Different predictions from Pythia8, Herwig [208], and Pythia8 interfaced to Photos [169] are considered, with the nominal prediction taken as the logarithmic average of the three, and the envelope of the differences with respect to this average taken as sources of uncertainty. The impact of electroweak corrections is studied by comparing the templates with and without them [209, 210, 211].

The W boson mass is extracted from simultaneous likelihood template fits to the q/p_T distribution on $W \to \mu\nu$ and the ϕ^* distribution (defined in equation 6.28) in $Z \to \mu\mu$, to determine the p_T^W (and p_T^Z) parameters. The ϕ^* variable is chosen as it reduces the impact from the modelling uncertainties while retaining sensitivity to main parameters to which the p_T^V is sensitive. Apart from QCD, other electroweak backgrounds, similar to those for the ATLAS result, such as $Z \to \mu\mu$ with one reconstructed lepton, $W \to \tau\nu$, are estimated from simulation. The parameters that are extracted together with m_W are:

- the fraction of W^+ signal;
- the fraction of W^- signal;
- the fraction of QCD background;
- the value of α_s for the Z boson processes, needed for the modelling of the p_T^V distribution;

- an independent value of α_s that is shared for the W^+ and W^- signals;
- a shared value of the intrinsic transverse momentum of the partons (k_T^{intr}) for all W and Z boson processes, needed for the modelling of the p_T^V distribution;
- a value of A_3, freely floating in the fit, shared between W^+ and W^- processes. This is to reduce the impact of the uncertainties on the angular coefficients, mostly driven by A_3, on m_W.

The final result obtained by the LHCb collaboration is:

$$\begin{aligned} m_W^{\mathrm{LHCb}} &= 80354 \pm 23_{\text{stat}} \pm 10_{\text{exp}} \pm 17_{\text{theory}} \pm 9_{\text{PDF}} \text{ MeV} \\ &= 80354 \pm 32 \text{ MeV}, \end{aligned} \tag{6.31}$$

where the uncertainties are broken down in the several sources indicated in table 6.8. The PDF contributes to 9 MeV, with the other leading theoretical uncertainties stemming from the p_T^W (and p_T^Z) model, given from the envelope of the different p_T^W models, and angular coefficients (with A_3 extracted from the fit). The experimental systematic uncertainties are driven by the muon momentum scale and resolution calibration, and the determination of the efficiency corrections. The statistical and systematic uncertainty have roughly the same impact on the m_W^{LHCb} measurement, which can therefore benefit from the extra data collected during Run 2 and in the future data-taking periods and improvements on the theoretical side. The result is in agreement with the Standard Model. Figure 6.22 shows the post-fit q/p_T and ϕ^* distributions, together with the fitting ranges.

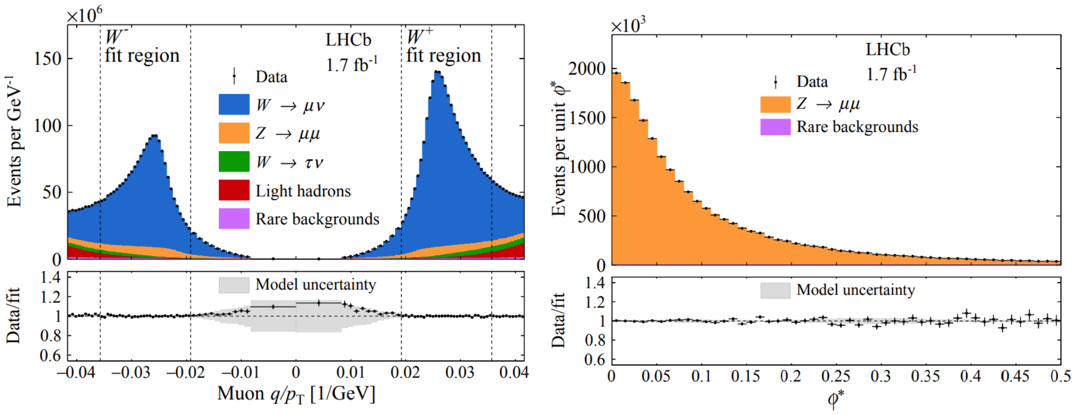


Figure 6.22: Distributions of q/p_T for $W \to \mu\nu$ events (left) and ϕ^* for $Z \to \mu\mu$ events (right). From [38].

Source	Impact [MeV]
Parton distribution functions	9
Theory (excl. PDFs) total	17
Transverse momentum model	11
Angular coefficients	10
QED FSR model	7
Additional electroweak corrections	5
Experimental total	10
Momentum scale and resolution modelling	7
Muon ID, trigger and tracking efficiency	6
Isolation efficiency	4
QCD background	2
Statistical	23
Total	32

Table 6.8: Contributions to the systematic uncertainty on m_W. Negligible contributions below 1 MeV are not listed. From [38].

Chapter 7

First measurement of the W boson mass with CMS

This chapter provides a detailed description of the W boson mass measurement performed by the CMS experiment. In first section, the innovative approach used by CMS is presented, describing how it addresses the limitations of previous measurements at hadron colliders, as seen in section 6.3. The event selection is described in section 7.2. The theoretical modelling employed by the analysis, together with the implementation of the various sources of systematic uncertainty, is described in detail in section 7.3. Section 7.4 contains my original contribution to the W boson mass measurement, which represents the continuation of the work presented in section 5.2. Section 7.5 describes the muon scale and resolution calibration. Section 7.6 describes the calibration of the hadronic recoil. Section 7.7 describes the background estimation, describing particularly the estimation of the contamination from nonprompt muon decays. Section 7.8 contains the results of various ancillary measurements performed to validate the muon scale and resolution calibration and the theoretical modelling used in the analysis. Section 7.9 describes the fit to the $p_T^\mu - \eta^\mu - q^\mu$ spectrum, describing the likelihood function that is minimized to extract m_W, together with the implementation of the systematic uncertainties. Finally, section 7.10 contains the results of the likelihood minimization and the extracted value of m_W, together with various tests to check the robustness of the result with respect to changes to the experimental and theoretical inputs.

7.1 Overview

It was seen in section 6.3 that the uncertainties on the parton density functions are traditionally a limiting source of systematic uncertainty. p_T^W is usually modelled using the information on p_T^Z and the theoretical knowledge on the p_T^W/p_T^Z ratio, with cross-checks with p_T^W data, if measured, used to constrain some of the theoretical uncertainties, but it was discussed that these do not capture the non-perturbative uncertainties entirely, and the assumption on the correlations between p_T^W and p_T^Z can introduce relatively large variations in the templates used to extract m_W.

It was shown in the measurement of the rapidity and helicity of the W boson published by CMS in 2020 [212] that the precise measurement of the two-dimensional $p_T - \eta$ distribution of the leptons from W boson decays constrains the PDF uncertainties significantly. This innovative method is used for the first time to measure m_W and simultaneously constrain *in-situ* the nuisance parameters associated to the PDFs. The same argument can actually be applied to the QCD modelling. As proven in section 7.3, the parameters that describe the p_T^W and p_T^Z theoretical modelling can be extracted together with m_W and the PDFs. Regarding the QCD modelling, an innovative model, described in section 7.3, is implemented to describe the non-perturbative uncertainties. No information is used directly from $Z \to \mu\mu$ events, either to tune the theoretical QCD model or to calibrate the lepton momentum scale, with the intention of keeping them independent from the W events, so that they can be used as a validation of the theoretical modelling and calibration procedure.

The CMS measurement [1] extracts m_W from a fit to the two-dimensional distribution of $p_T - \eta$ as function of the charge of the lepton, with a very fine binning in p_T and η: 30 uniform bins in the range 26 GeV$< p_T < 56$ GeV and 48 uniform bins in the range $-2.4 < \eta < 2.4$ separately for the two charge, for a total of 2880 bins. This is done to extract as much information as a possible on the theoretical modelling. The data sample used in the analysis is taken from the latter part of 2016 at a center-of-mass energy of $\sqrt{s} = 13$ TeV and corresponds to an integrated luminosity of 16.8 fb^{-1}, allowing us to rely on a high-statistics sample after event selection is applied. This results in significant constraints in the theoretical modelling. The sample is selected to ensure an optimal performance of the CMS detector, particularly regarding the reconstruction of charged particle tracks.

The data from the earlier part of 2016 is not used in the CMS analysis because of a reduction in the signal-to-noise ratio in the strip detectors leading to a loss of hits correlated with the increasing instantaneous luminosity during 2016. In particular, this lead to an increase in the electronic current in the APV25 readout chip, resulting in discharge times longer than the 25 ns gap between consecutive bunch crossings. This issue was fixed by adjusting the pre-amplifier feedback voltage bias (VFP), which governs the discharge speed of the pre-amplifier, in the latter parts of 2016 CMS data-taking, and the data that is used for this analysis belongs to the period referred to as `postVFP`. The VFP issue leads to a reduction in muon reconstruction efficiency as the instantenous luminosity increases: it was verified that this effect is not fully recovered by the efficiency corrections derived using the tag-and-probe procedure, as explained in subsection 7.4.2, as the data and Monte Carlo agreement became worse with time (i.e. with increasing instantaneous luminosity) after corrections were derived in different sub-periods of 2016.

In the ATLAS measurement, the precision with which m_T was measured was largely degraded due to the high pileup in the data sample used to extract m_W. The average pileup in 2016 is even higher than that in 2011 used in the ATLAS analysis, as can be seen in figure 3.4, so the measurement does not rely directly on the transverse mass m_T, which is instead used only to define the signal region to extract m_W and the control regions to estimate the nonprompt QCD background, as described in section 7.7. The analysis relies only on $W^\pm \to \mu^\pm \overset{(-)}{\nu}_\mu$ and not $W^\pm \to e^\pm \overset{(-)}{\nu}_e$ the electron energy scale and resolution could not be calibrated in time for this first measurement with a level of precision sufficient to have a competitive measurement.

Figure 7.1 shows the measured $W^+ \to \ell^+\nu_\ell$, $W^- \to \ell^-\bar{\nu}_\ell$ and $Z/\gamma^* \to \ell^+\ell^-$ cross-sections as a function of the center-of-mass energy. The theoretical predictions at N^{3}LO in QCD, computed using the MSHT20aN^{3}LO PDF set [213], are superimposed. Using these cross-sections, and the peak instantaneous luminosity for 2016 shown in figure 3.2, one expects the event rate indicated in table 7.1 for each of these processes. This table also contains the expected number of events considering the integrated luminosity of the 2016 postVFP period of 16.8 fb^{-1}. The total number of events after the analysis selection is reported in section 7.2.

Process	Event rate [s^{-1}]	Expected events for 2016 postVFP
$W^+ \to \mu^+\nu_\mu$	$\lesssim 170$	$(1.98 \pm 0.05) \times 10^8$
$W^- \to \mu^-\bar{\nu}_\mu$	$\lesssim 130$	$(1.46 \pm 0.04) \times 10^8$
$Z/\gamma^* \to \mu^+\mu^-$	$\lesssim 30$	$(3.28 \pm 0.09) \times 10^7$

Table 7.1: Approximate event rates computed using the cross-sections in figure 7.1 and the peak instantaneous luminosity shown in figure 3.2 for 2016 for each of the processes. This table also contains the expected number of events considering the integrated luminosity of the 2016 postVFP period of 16.8 fb^{-1} and the same cross-sections. The uncertainties are the sum in quadrature of the uncertainty in the measured cross-section and integrated luminosity (1.2%).

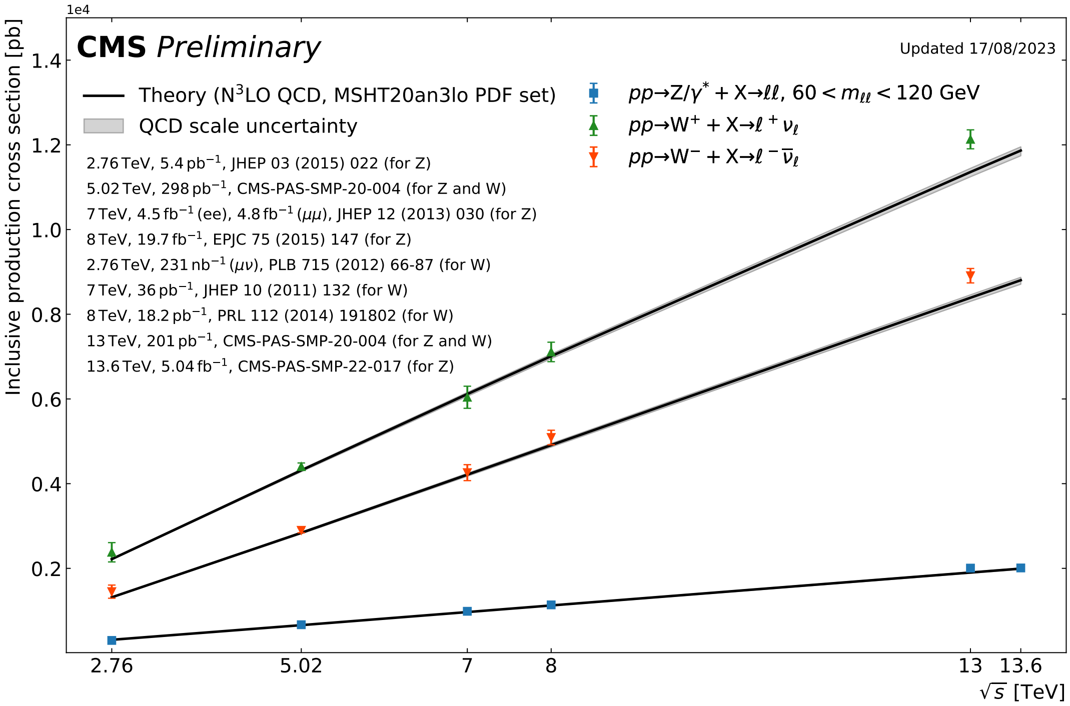


Figure 7.1: Measured $W^+ \to \ell^+\nu_\ell$, $W^- \to \ell^-\bar{\nu}_\ell$ and $Z/\gamma^* \to \ell^+\ell^-$ cross-sections as a function of the center-of-mass energy. The theoretical predictions at N^{3}LO in QCD, computed using the MSHT20aN^{3}LO PDF set [213], are superimposed. From [214].

7.2 Event selection

Events are selected by the logical `OR` of the `HLT_IsoMu24` and `HLT_IsoTkMu24` triggers, which require an isolated muon with $p_T > 24$ GeV. In the first trigger, the muon reconstruction is performed from the muon spectrometer to the inner tracker, in the second trigger the reconstruction is performed starting from the inner tracker. Then, only events with exactly one muon that passes a loose selection are selected. The loose selection on the muon corresponds to the *veto muon* selection, used to suppress the $Z \to \mu\mu$ background. The corrections for this selection and the relative uncertainties are described in detail in 7.4.9, while here the veto selection for the nominal result can be summarized as:

- global muon with all the additional requirements determined in 5.2.1:
    - $|\eta| < 2.4$;
    - tracker $p_T > 15$ GeV;
    - standalone $p_T > 15$ GeV;
    - $\Delta R < 0.3$, where ΔR is the distance between the inner track and the standalone muon track;
    - the inner track must pass the high-purity [92] selection;
    - the outer track must have at least 1 valid hit in the muon spectrometer, where the definition of valid hit is given in 7.4;
- passes the loose ID described in 5.1;
- the impact parameter with respect to the beam spot must be $|d_{\mathrm{xyBS}}| < 0.05$ cm.

The muons selected in the analysis must pass an additional set of criteria, which are the RECO, $\mathrm{Tracking}$, IDIP, $\mathrm{Trigger}$, $\mathrm{Isolation}$ efficiencies described in 5.1 and 7.4:

- RECO: the muon is a standalone muon with $p_T > 15$ GeV and at least one valid hit in the muon spectrometer;
- $\mathrm{Tracking}$: the muon is a global muon with inner track $p_T > 15$ GeV that passes the high-purity selection [92], with ΔR between the inner and outer track to be less than 0.3;
- IDIP: the muon must pass the medium ID described in 5.1 and must have $|d_{\mathrm{xyBS}}| < 0.05$ cm;
- $\mathrm{Trigger}$: the muon must be matched to a trigger object returned by the `HLT_IsoMu24` or `HLT_IsoTkMu24` trigger paths;
- $\mathrm{Isolation}$: the relative isolation defined from the Particle Flow candidates, in the vertex-agnostic implementation, as described in 5.1, within a radius of $\Delta R < 0.4$ to be less than 0.15.

The muons must also have 26 GeV$< p_T < 56$ GeV.

Events are also rejected if at least one electron with $p_T > 10$ GeV, $|\eta| < 2.4$, $|d_{\mathrm{xyBS}}| < 0.05$ cm and $|d_{\mathrm{zBS}}| < 0.2$ cm, passing a cut-based ID selection [215] is found.

The missing transverse momentum (p_T^{miss}) is computed using the `DeepMET` algorithm [216, 217], which uses machine-learning techniques. Other CMS analyses used the so-called `PFMET`, in which p_T^{miss} is defined as in 6.2 and particles entering in the definition are those reconstructed by the Particle Flow algorithm [100]. For the charged particles, only those associated with the main primary vertex in the analysis are considered, while all the neutral PF candidates are retained. The use of the `DeepMET` algorithm improves the resolution on the reconstructed p_T^{miss} significantly, particularly improving the determination of the contribution from neutral particles, improving the resolution on the transverse mass m_T, and increasing the separation between signal and background. For this analysis, the `DeepMET` algorithm is modified with respect to the default implementation so that charged particles that enter in the p_T^{miss} are those associated to the same vertex as the one associated to the muon. This is done for the same reason that motivated the definition of the vertex-agnostic isolation described in 5.2.4. Events are required to have $m_T > 40$ GeV, while m_T is not used in the fit to extract m_W.

As already mentioned, $Z \to \mu\mu$ events are used to validate the theoretical modelling and the scale and resolution calibration. The strategy to extract the theoretical modelling is validated by repeating the m_W fit to extract m_Z using the same strategy. To achieve this, one of the muons in $Z \to \mu\mu$ events is added to p_T^{miss}, and m_Z and the nuisance parameters related to theory model are extracted from fitting the $p_T - \eta$ distribution simultaneously for the positive and negative muon from the $Z \to \mu\mu$ decays. We refer to this selection as W-like fit. The p_T^Z distribution obtained from the W-like fit is then compared with the measured $p_T^{\ell\ell}$ spectrum.

The muon calibration performance is evaluated by extracting m_Z from the m_{ll} distribution. For $Z \to \mu\mu$ events, two muons with 26 GeV$< p_T < 60$ GeV and $|\eta| < 2.4$ that pass the RECO, $\mathrm{Tracking}$, IDIP and $\mathrm{Isolation}$ selections are requested. The two muons must have opposite charges and the invariant mass of their combined system must satisfy 60 GeV$< m_{\mu\mu} < 120$ GeV. For the W-like selection, the $Z \to \mu\mu$ events are split into two sets based on the event number, so that odd (even) events are used to analyze positive (negative) muons. Only the analyzed muon is required to pass the $\mathrm{Trigger}$ selection, which avoids the need to evaluate correlations in the triggering efficiency in events where both muons satisfy the trigger requirements. The m_T requirement, defined from the analyzed muon and p_T^{miss} modified to include the other muon, is increased to 45 GeV. For the dilepton selection, no m_T requirement is applied and the same event splitting and trigger requirement as in the W-like selection are imposed.

Figures 7.2 and 7.3 show the $p_T - \eta$ distribution for the positive and negative muons, respectively for the analysis to extract m_W and the W-like analysis to extract m_Z. The two-dimensional distribution is "unrolled" such that each bin on the $x-$axis represents one (p_T^μ, η^μ) cell. The gray band represents the uncertainty in the prediction, before the fit to the data.

The signal is modelled as described in 7.3, whereas the backgrounds are modelled as described in 7.7. When event generators are used to construct the distributions for certain processes, the response of the CMS detector is reproduced using a full detector simulation based on Geant4 [198].

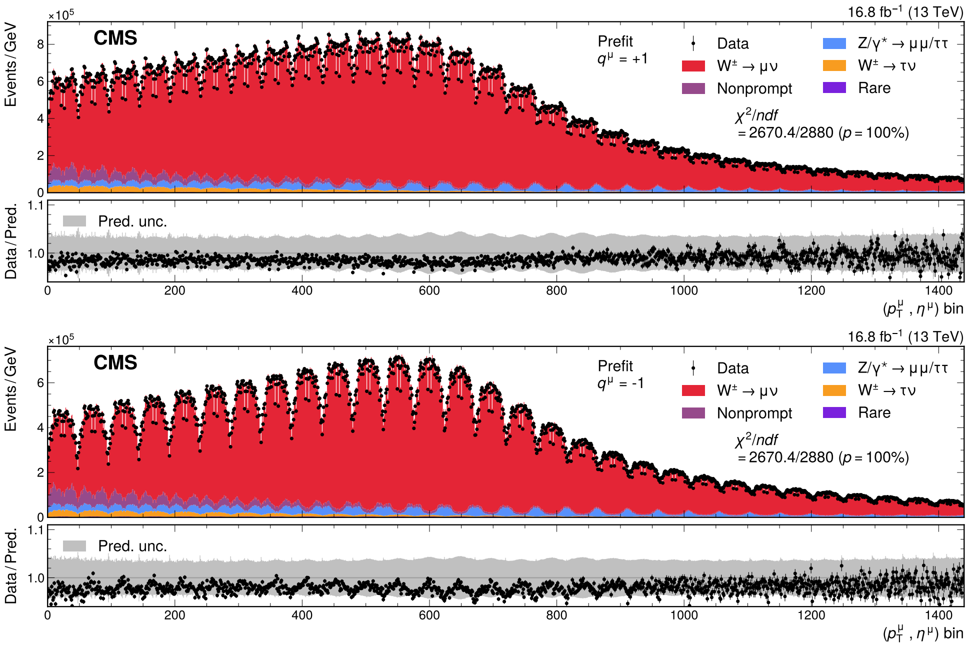


Figure 7.2: Measured and predicted (p_T^μ, η^μ) distributions used in the m_W measurement, for positively (upper) and negatively (lower) charged muons. The two-dimensional distribution is "unrolled" such that each bin on the $x-$axis represents one (p_T^μ, η^μ) cell. The gray band represents the uncertainty in the prediction, before the fit to the data. The bottom panel shows the ratio of the number of events observed in data to the nominal prediction. The vertical bars represent the statistical uncertainties in the data. From [1].

Table 7.2 contains the pre-fit number of events for the signal and background processes for the m_W and the W-like m_Z analyses. It also contains the acceptances and selection efficiencies for the m_W and W-like m_Z event selection for the signal processes. For the m_W analysis, the acceptance is defined as the fraction of events with one muon with 26 GeV$< p_T < 56$ GeV and $|\eta| < 2.4$, for the m_Z W-like analysis it is defined as the fraction of events with two muons with 26 GeV$< p_T < 60$ GeV and $|\eta| < 2.4$. This fraction is computed with respect to the total number of events with 60 GeV$< m_{\mu\mu} < 120$ GeV. The selection efficiencies are computed with respect to the acceptance selection.

7.3 Theoretical modelling

The W and Z production is simulated at next-to-next-to-leading order in QCD using the MiNNLO$_{\text{PS}}$ Wj and Zj [218, 219, 220] (rev. 3900) processes in POWHEG-BOX-V2 [185, 186, 187], interfaced with Pythia 8.240 [203] for the parton shower and hadronization, and with Photos++ 3.61 [169, 221] for final-state photon radiation. The CP5 tune [222] is used for the modelling of the underlying event, with the value of the intrinsic transverse momentum of the partons k_T set to 2.225 GeV, tuned using the $p_T^{\mu\mu}$ distribution from $Z \to \mu\mu$ events as described in [223]. Simulated samples of more than 4 billion (400 million) W (Z) boson production events have been produced, to ensure that the MC sample size is not a significant source of uncertainty

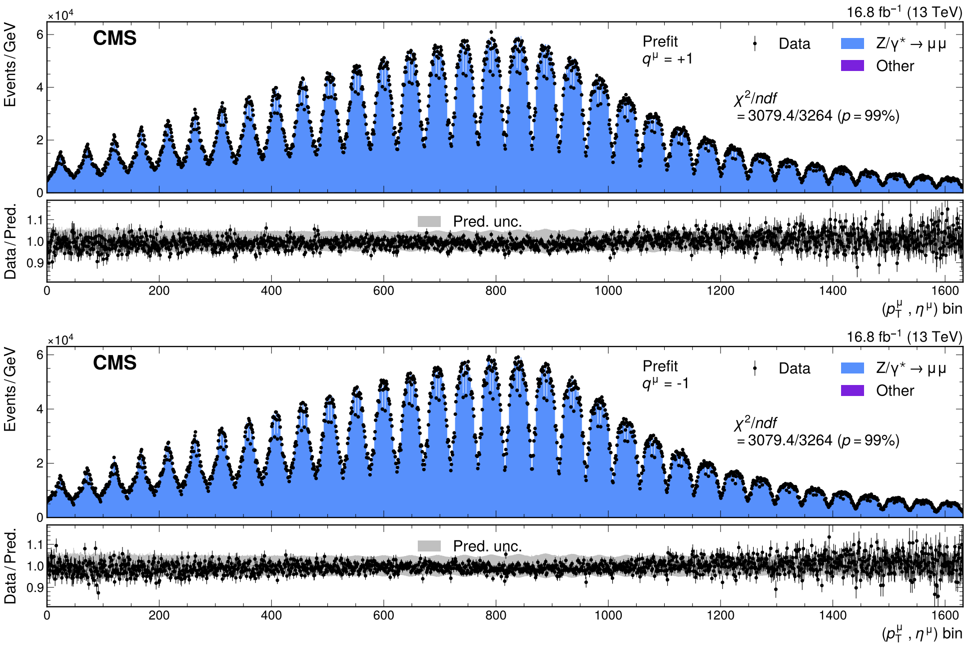


Figure 7.3: Measured and predicted p_T^μ, η^μ distributions used in the W-like m_Z measurement, for positively (upper) and negatively (lower) charged muons. The two-dimensional distribution is "unrolled" such that each bin on the $x-$axis represents one (p_T^μ, η^μ) cell. The gray band represents the uncertainty in the prediction, before the fit to the data. The bottom panel shows the ratio of the number of events observed in data to the nominal prediction. The vertical bars represent the statistical uncertainties in the data. From [1].

in the measurement. The statistical uncertainties due to the finite size of the generated samples are implemented using the Barlow–Beeston method [224], as simplified by Conway [225], with the nominal statistical uncertainty scaled up by 25% to account for the fluctuations in the alternate templates used to build the systematic variations [226]. This scaling factor has been determined using bootstrap methods [227], verifying that the uncertainties have the appropriate coverage.

7.3.1 p_T^V modelling

The generator-level p_T^V and y^V distributions in MINNLO$_{\text{PS}}$ are corrected to state-of-the-art calculations in QCD, including the resummation of logarithmically-enhanced contributions at small p_T^V and a model for non-perturbative effects also at small p_T^V. This is done with the SCETLIB code [228, 229, 230], which performs the p_T^V resummation. These predictions are matched to the next-to-next-to-leading order calculations from DYTURBO [156, 231], so that a N^{3}LL$+$NNLO accuracy is reached. After these corrections, the dependence of the p_T^V distribution on the parton shower and tune is negligible. The predictions computed using DYTURBO v1.4.0 [156, 231, 232], MATRIX+RADISH v1.0.0 [233, 234], and CuTe-MCFM v10.2 [235, 236], with a precision up to the same or higher perturbative order, were also considered for comparison. Using these alternate predictions results in expected shifts in m_W that are within the SCETLIB+DYTURBO uncertainties.

No correction from the $p_T^{\mu\mu}$ spectrum measured in $Z \to \mu\mu$ data is applied to the $Z \to \mu\mu$ or $W \to \mu\nu$ events. Instead, the correct p_T^W spectrum is inferred using the profiling procedure employed in the maximum likelihood fits used to extract m_W. This is possible thanks to the high accuracy of the theoretical predictions employed, and the use of novel techniques to model their uncertainties and correlations across phase space.

The large statistical power of the analyzed data sample allows us to determine and constrain the nui-

Number of events for m_W analysis

Process	Positive charge	Negative charge	Sum of charges
Data	65256880	51252359	116509239
$W \to \mu\nu$	$(5.9489 \pm 0.0004) \times 10^7$	$(4.6014 \pm 0.0003) \times 10^7$	$(1.05503 \pm 0.00005) \times 10^8$
$Z/\gamma^* \to \mu^+\mu^-/\tau^+\tau^-$	$(2.663 \pm 0.001) \times 10^6$	$(2.483 \pm 0.001) \times 10^6$	$(5.145 \pm 0.002) \times 10^6$
Nonprompt	2.32140×10^6	2.36251×10^6	4.68390×10^6
$W \to \tau\nu$	$(1.3722 \pm 0.0004) \times 10^6$	$(1.1246 \pm 0.0006) \times 10^6$	$(2.4967 \times 0.0007) \times 10^6$
Rare	$(4.319 \pm 0.002) \times 10^5$	$(4.007 \pm 0.002) \times 10^5$	$(8.326 \pm 0.003) \times 10^5$

Number of events for W-like m_Z analysis

Process	Positive charge	Negative charge	Sum of charges
Data	3687709	3702119	7389828
$Z/\gamma^* \to \mu^+\mu^-$	$(3.7025 \pm 0.0007) \times 10^6$	$(3.7164 \pm 0.0007) \times 10^6$	$(7.419 \pm 0.001) \times 10^6$
Other	$(1.422 \pm 0.004) \times 10^4$	$(1.419 \pm 0.004) \times 10^4$	$(2.841 \pm 0.005) \times 10^4$

Acceptance and selection efficiencies

Signal process	Acceptance		Selection efficiency	
	Positive charge	Negative charge	Positive charge	Negative charge
$W \to \mu\nu$	40%	41%	75%	76%
$Z/\gamma^* \to \mu^+\mu^-$ (W-like analysis)	33%		71%	71%

Table 7.2: Table containing the pre-fit number of events for the signal and background processes for the m_W and the W-like m_Z analyses. The uncertainty shown for the processes are only due to the finite Monte Carlo statistics (it is not shown for the nonprompt background). The table also contains the acceptances and the selection efficiencies on top of the acceptance for the m_W and m_Z W-like event selection for the signal processes.

sance parameters associated to the p_T^W modelling, considering several sources are considered. Perturbative uncertainties in the resummed predictions are evaluated using the "theory nuisance parameter" (TNP) approach [237], which exploits the known all-order perturbative structure of the resummed calculation and is implemented in SCETlib. Three perturbative components in the p_T^V resummation are considered in the formalism used in this analysis:

- the "hard function" that describes the hard virtual corrections for W and Z production;
- the "proton beam functions" that extend the PDFs to include collinear radiation;
- the "soft function" describing soft radiation.

In the TNP approach, the minimal independent set of ingredients that would be needed at the next perturbative order are identified and parameterized in terms of common nuisance parameters. The TNPs have a true, but unknown, value that can be varied according to their expected typical magnitude, so that the uncertainties due to the missing higher orders are correlated correctly across p_T^V, y^V, and m_V, and the TNPs can be consistently used in the profile maximum likelihood fit used to extract m_W to provide a reliable estimate of p_T^V. The SCETlib program implements different configurations for the TNPs, in terms of the logarithmic accuracy of the prediction and the perturbative order at which the TNPs are included. The N^{3+0}LL scheme is used for the result reported in the CMS m_W paper, where the prediction has N^{3}LL accuracy and the TNPs representing the unknown higher-order corrections are estimated from multiplicative variations of the same N^{3}LL terms. Two alternate schemes are also considered: N^{3+1}LL, which implements the full N^{4}LL perturbative structure, combining the known values for the parameters up to N^{3}LL with best estimates of the higher-order terms and their variations to define TNP variations, and N^{4+0}LL, which uses the same approach as N^{3+0}LL but with full N^{4}LL calculation. The choice of scheme has a very small impact on the results.

Traditional 7-point variations of the factorization and renormalization scales μ_F and μ_R, are implemented in DYTurbo to determine the perturbative uncertainty due to the matching to fixed-order predictions. This uncertainty is implemented through two nuisance parameters, one for all the W processes and one for all the Z processes. The variation is treated as fully correlated (for the W also between W^+ and W^-) within the two groups of processes, but uncorrelated between Z and W. The nominal value used to determine the transition between the resummation and fixed-order regimes is set to $0.5\,m_V$: an uncertainty is implemented by varying this threshold between $0.35\,m_V$ and $0.75\,m_V$. This is found to have a negligible impact on the results.

Non-perturbative effects, due for example to the residual transverse motion of the partons inside the proton, are expected to scale as $(\Lambda_{\mathrm{QCD}}/p_T^V)^2$ [238, 239], with $\Lambda_{\mathrm{QCD}} \approx 100 - 300$ MeV being the QCD vacuum expectation value, resulting in a large impact $p_T^V \leq 10$ GeV. The predictions considered here implement phenomenological models that require tuning to data to describe these non-perturbative effects. Two sources of non-perturbative effects impact p_T^V: non-perturbative corrections to the Collins-Soper rapidity anomalous dimension [238, 239], universal for W and Z boson production, and non-perturbative contributions to the beam (and soft) functions, accounting for the "intrinsic k_T" of the partons inside the protons, not universal as they can depend on the flavor and x of the interacting parton. The SCETlib program implements a corresponding non-perturbative model for both sources [228]. In the CMS m_W analysis, these parameters are loosely constrained around nominal values that correspond to minimal non-perturbative k_T variations. The Collins-Soper anomalous rapidity parameters are correlated between Z and W boson production, while the terms related to the k_T variations are treated as uncorrelated. Their best fit values are extracted from the maximum likelihood fits to the measured distributions.

The SCETlib+DYTurbo and MiNNLO$_{\mathrm{PS}}$ calculations are performed in a fixed-flavor scheme with massless quarks, and the effect of their mass is expected to become relevant for $p_T^V \approx m_{b,c}$ the masses of the $b-$ and $c-$quarks. This is partially taken into account by varying the heavy-quark thresholds using $m_{b,c}$ variations of the MSHT20 PDF set [207], while it should also be captured by the non-perturbative uncertainty model used in this analysis, since its parameters are only loosely constrained, as they are expected to be determined in the fit, and its correlation scheme between W and Z processes (these effects are expected to be different between them, as different PDFs enter into the calculation of the cross-section for these processes).

Figure 7.4 shows the $p_T^{\mu\mu}$ distribution on data and simulation, as measured in the dilepton analysis, and the p_T^{μ} distribution, in the W-like selection, from $Z \to \mu\mu$ events. The various sources of uncertainty due to the p_T^V modelling are shown, and their corresponding impact on $p_T^{\mu\mu}$ and p_T^{μ} can be seen: the non-perturbative uncertainty is most pronounced at low p_T^V and around the Jacobian peak, resulting in a potentially large impact on m_W or m_Z. Uncertainties in the resummation calculation and in the matching to fixed-order calculations are relevant up to $p_T^V \approx 40$ GeV and again around the Jacobian peak. The perturbative uncertainties in fixed-order QCD are dominant only at high p_T^V, but have a small impact on p_T^{μ} in the region sensitive to m_W or m_Z.

Section 7.8 contains the results from the W-like and dilepton fits to validate the full theoretical model and the calibration procedure. The total impact on m_W from the theoretical modelling of p_T^V is 2.0 MeV.

7.3.2 Parton density functions

Several modern PDF sets at NNLO accuracy have been considered: CT18Z [182], CT18 [182], NNPDF3.1 [181], NNPDF4.0 [199], MSHT20 [207], PDF4LHC21 [240]. The only available approximate N^{3}LO PDF set MSHT20aN^{3}LO [213] has also been considered. Figure 7.5 shows the pre-fit η^μ distributions for $W^\pm$ events for data and the various PDF sets considered: CT18Z is shown in the filled histograms, the filled band represents the uncertainty from this PDF set, while the others are shown as the different lines. The consistency among the five PDF sets and the observed data is determined by performing likelihood fits to the distribution of the η of the muon from the decay for $W^\pm$ events and the dilepton rapidity $y^{\mu\mu}$ for $Z \to \mu\mu$ events simultaneously for each PDF set under consideration. Fits are performed including all the uncertainties of the nominal m_W fit, removing the PDF+α_s or the theory uncertainties. The impact of α_s is evaluated by shifting its nominal value by $\pm$0.015 and deriving alternative PDF sets. Table 7.3 summarizes the results of

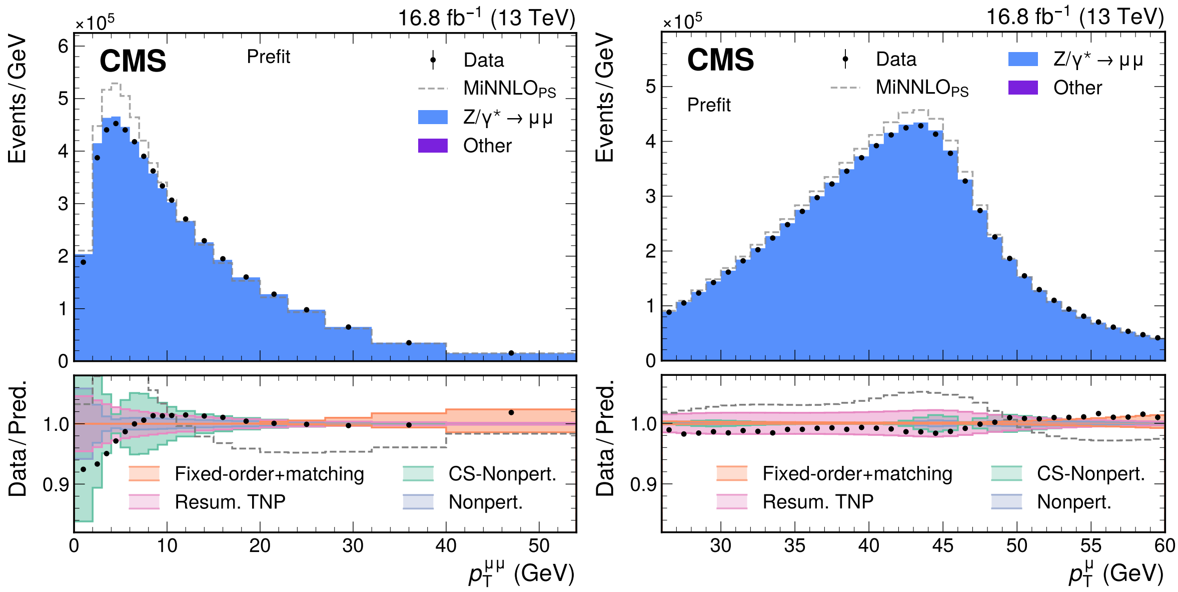


Figure 7.4: Measured and simulated $p_T^{\mu\mu}$ (left) and p_T^{μ} (right) distributions in $Z \to$ events (the events on the left pass the dilepton selection, those on the right pass the W-like selection). The uncorrected MiNNLO$_{\text{PS}}$ predictions are shown by the dashed gray line. The nominal predictions from MiNNLO$_{\text{PS}}$ with the corrections from SCETLIB+DYTURBO, as described in the text, are shown. The bottom panel shows the ratio between the number of events observed in data and that of the total nominal prediction, as well as the relative impact of variations of the predictions. Different sources of uncertainty are shown in the lower panel: the fixed-order uncertainty and the uncertainty in the resummation and fixed-order matching (orange), resummed prediction using TNPs (pink), the Collins–Soper (CS) kernel non-perturbative uncertainty (green), and other non-perturbative uncertainties (light blue). From [1].

these studies.

The agreement with data is good for all sets when all uncertainties are included and when the other theory uncertainties are removed. When the PDF and α_s uncertainties are removed, the CT18 and CT18Z tend to have significantly lower $p-$values, with the former reaching the percent level. It should be noted, however, that both these sets are among those with the largest PDF uncertainties, enough to cover for differences with respect to data. Before unblinding the result, the consistency of the various PDF sets was studied by performing bias tests purely on simulation: pseudodata are built using one PDF set, while templates to extract m_W are built using another PDF set. The uncertainties considered in the bias test are only those from the PDF set used to build the templates. This is repeated for all the possible pairs of PDF sets, also swapping between pseudodata and templates to extract m_W. This is important because also the uncertainties change when the PDF sets are chosen to extract m_W. For example, the case in which the pseudodata is built using NNPDF4.0 and template to extract m_W is built using CT18Z is different from the case in which they are swapped, because in the first case the PDF uncertainties considered when m_W is extracted are those of the CT18Z set, while in the second case they are those of NNPDF4.0.

We decide that two PDF sets are consistent if the shift in m_W is within the impact of the PDF systematic uncertainties on m_W:

$$|m_W^{\text{PDF,pseudodata}} - m_W^{\text{PDF,nominal}}| < \sigma_{m_W}^{\text{PDF,nominal}}. \tag{7.1}$$

This is not always the case: for the PDF sets that do not satisfy the condition in equation 7.1 when they are not used as pseudodata, all of the PDF variations are inflated by a scale factor, identical for all the variations, determinated as the smallest scale factor that allows the PDF set to satisfy the condition in equation 7.1 for all the pseudodata PDF sets. Table 7.4 shows the scale factors for each PDF set and their impact on m_W, both before and after scaling the PDF uncertainties as described. CT18, CT18Z, and PDF4LHC21 are not rescaled, as they are able to cover the differences in the predictions from the other PDF sets within the nominal uncertainties. The NNPDF sets, particularly NNPDF4.0, are those with the largest scale factors, but also the smallest impacts from the PDF uncertainties on m_W when these are not rescaled, compatible with the fact the uncertainties on these sets are among the smallest out of the sets considered. This can also be seen in table 7.3, where the $p-$values are lower (although still quite high) compared to the other sets when the other theory uncertainties are removed from the fit. The CT18Z set is used to derive the nominal m_W result, as it is able to describe the W and Z data within its uncertainties and does not need scale factor to cover for the alternate predictions from the other sets. The impact on m_W from changing PDF sets on m_W is tested, as described in 7.10.

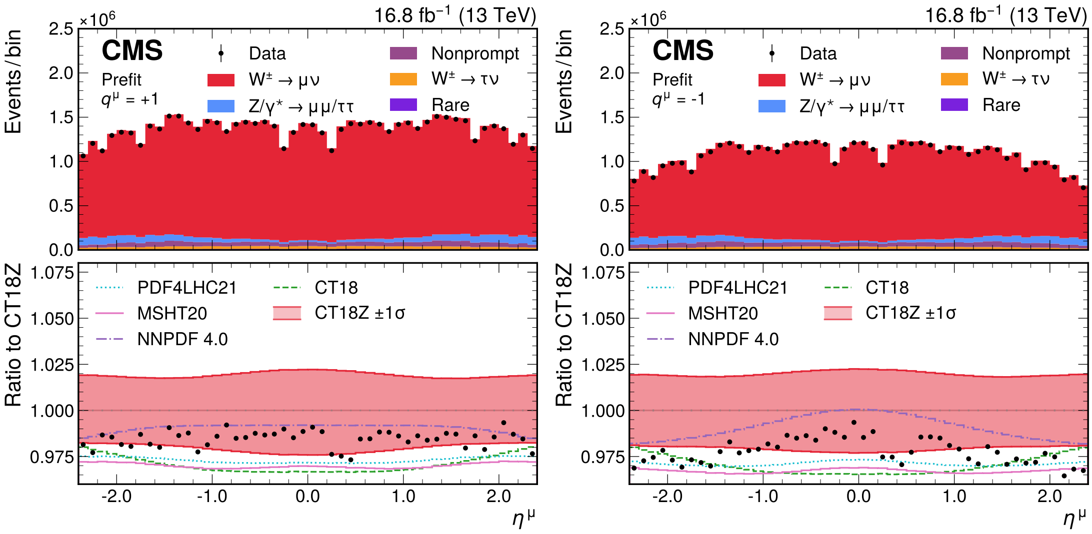


Figure 7.5: Measured and predicted η^μ distributions for positively (left) and negatively (right) charged muons. The nominal prediction, obtained with the CT18Z PDF set, is shown in filled light red. The uncertainty, evaluated as the sum of the eigenvector variation sets, is represented by the filled band in the lower panel. The nominal predictions using the PDF4LHC21, MSHT20, NNPDF4.0, and CT18 sets are also shown. The bottom panel shows the ratio between the number of events observed in data, or in variations in the predictions, and that of the nominal prediction from CT18Z. From [1].

The uncertainties in the PDFs are evaluated by propagating the Hessian eigenvectors of the CT18Z PDF set to the templates. The dependence of m_W on the PDF set was checked by repeating the fits with alternate PDF sets mentioned above, with and without the inflation of uncertainties: as described in 7.10, the extracted value of m_W has a negligible dependence on the PDF set. The total impact from the PDFs is 4.4 MeV: the inclusion of the η information in the fit allowed us to reduce their impact to this value. This can also be seen in the discussion in 7.10, where the nominal and global impact breakdown, discussed in 6.3.2, are reported.

7.3.3 Angular coefficients

The angular coefficients in equation 6.13 and their dependence on the W or Z production mechanism are modelled by MiNNLO$_{\text{PS}}$, and are accurate at NNLO in QCD. These predictions have been cross-checked with those from DYTurbo and MCFM, which have the same accuracy, and also with those from SCETlib, which are N^{3}LL accurate: compatibility within uncertainties was found in all cases. The uncertainties on the angular coefficients are evaluated by varying μ_R and μ_F in the nominal predictions. These variations

PDF set	Nominal fit		Without PDF+α_S unc.		Without theory unc.	
	χ^2/ndf	p-val. (%)	χ^2/ndf	p-val. (%)	χ^2/ndf	p-val. (%)
CT18Z	$100.7/116$	84	$125.3/116$	26	$103.8/116$	78
CT18	$100.7/116$	84	$153.2/116$	1.0	$105.7/116$	74
PDF4LHC21	$97.7/116$	89	$105.5/116$	75	$104.1/116$	78
MSHT20	$97.0/116$	90	$107.4/116$	70	$98.8/116$	87
MSHT20aN^{3}LO	$99.0/116$	87	$122.8/116$	31	$101.9/116$	82
NNPDF3.1	$99.1/116$	87	$105.5/116$	75	$115.0/116$	51
NNPDF4.0	$99.7/116$	86	$104.3/116$	77	$116.7/116$	46

Table 7.3: Goodness-of-fit test statistic for different PDF sets when fitting simultaneously the η^μ distributions for selected W^+ (W^-) events and the $y^{\mu\mu}$ distribution for $Z \to \mu\mu$ events. The saturated likelihood ratios, which are expected to follow a χ^2 distribution with ndf degrees of freedom if the model is an accurate representation of the data, and the associated p-value are both shown. The fit is performed in the nominal configuration with all uncertainties (left column), nominal configuration without PDF and α_S uncertainties (middle column), and nominal configuration without theory uncertainties (right column).

PDF set	Scale factor	Impact on m_W (MeV)	
		Original σ_{PDF}	Scaled σ_{PDF}
CT18Z	-	4.4	
CT18	-	4.6	
PDF4LHC21	-	4.1	
MSHT20	1.5	4.3	5.1
MSHT20aN^{3}LO	1.5	4.2	4.9
NNPDF3.1	3.0	3.2	5.3
NNPDF4.0	5.0	2.4	6.0

Table 7.4: Pre-fit uncertainty scaling factors required to cover the central predictions of the considered PDF sets and post-fit impact on m_W, with and without scaled PDF uncertainties.

are treated as uncorrelated among the helicity cross-sections σ_i corresponding to each of the A_i coefficients[1] and the unpolarized cross-section, and uncorrelated in p_T^V in ten bins, but correlated across y_V. One fully correlated variation across p_T^V is also included, and the contribution from this variation is subtracted from the uncorrelated components, to avoid double counting in the uncertainties. This set of variations should capture the correlations of higher-order corrections across phase space, which are not well known. The variations are uncorrelated between W and Z processes, but fully correlated between W^+ and W^-. The full difference between the angular coefficients before and after the PYTHIA 8 shower and the isotropic smearing of the colliding partons, due to the intrinsic k_T model, is taken as an additional source of systematic uncertainty, treated as fully correlated across the angular coefficients, phase space, and W and Z boson production, as this was found to induce non-negligible variations in the angular coefficients. The total impact (using the nominal definition) on m_W from these uncertainties is 3.2 MeV, dominated by the uncertainties on A_0, A_2 and A_4.

7.3.4 Electroweak corrections

QED final-state radiation (FSR) is considered at leading logarithmic (LL) accuracy by interacing MINNLO$_{\text{PS}}$ to PHOTOS++. Uncertainties are evaluated by propagating the differences in the generator post-FSR dimuon mass and p_T^μ distributions, without the corrections from PHOTOS++, with those from HORACE v3.2 [241, 242],

---

[1] The helicity cross-sections are defined as $\sigma_i = A_i \times \sigma_{UL}$.

of similar accuracy. The impact on m_W is smaller than 0.3 MeV. The impact from the modelling of QED initial-state radiation (ISR), evaluated from the differences in the distributions computed by switching it off, is negligible.

For Z processes, the uncertainties on the electroweak virtual corrections are implemented by taking the ratios of the Z boson mass, rapidity and $\cos\theta^*$ distributions computed at NLO with the Z_ew process in the POWHEG-BOX-V2 (rev. 3900) program [243, 244] including universal higher-order corrections, and those at LO.

For W processes, the ReneSANCe 1.3.11 program [245] is used to separate between the virtual and photonic contribution to the electroweak corrections, and the uncertainty on the weak virtual corrections is evaluated using the same procedure as for the Z processes. The uncertainty from the virtual EW corrections has an impact on m_W of 1.9 MeV.

7.4 Efficiency measurements

The tag-and-probe method to derive efficiencies and the different steps in which those are derived were described in section 5.2.1. After efficiencies are measured as a function of $p_T - \eta - q$ (and u_T, defined in equation 5.7, for $\mathrm{Trigger}$ and $\mathrm{Isolation}$), the ratios of them between data and Monte Carlo are computed, resulting in what we refer to as Scale Factors: this should enable the cancellation of some systematic uncertainties either related to the fits (for example, the shape used in the invariant mass fits, the invariant mass cut used, the binning in invariant mass, etc.) or to the introduction of a tag muon, not present in the $W \to \mu\nu$ decays (as for example in the case of the dependence on the boson recoil and distance in transverse plane between recoil and muon for $\mathrm{Isolation}$ and $\mathrm{Trigger}$, described in 5.2.3).

This also simplifies the implementation of analytical shapes to model the Scale Factors to lift some of the statistical fluctuations from the fits or biases that stem from individual $p_T - \eta - q$ fits: this is done using polynomial functions, and the choice of using Scale Factors instead of efficiencies requires lower degrees to properly catch the features as opposed to directly fitting efficiencies, as the Scale Factors would result in relatively small (and slowly varying) corrections on top of the efficiencies themselves. This is particularly true for $\mathrm{Trigger}$ and $\mathrm{Isolation}$ efficiencies, which typically increase rapidly up to a threshold and then plateau at high p_T, and a number of different families of curves could mimic this behaviour particularly well, but at the cost of inducing artificial variations on m_W depending on the choice of family of curves chosen. This procedure, which we refer to as *smoothing*, is described in more detail in subsection 7.4.6.

7.4.1 Tag and probe definitions

The tag muon must satisfy the following requirements:

- $p_T > 25$ GeV and $|\eta| < 2.4$;
- matched to the trigger objects returned by the $\mathtt{HLT\_IsoMu24}$ or $\mathtt{HLT\_IsoTkMu24}$;
- pass the medium ID, $|d_{\mathrm{xyBS}}| < 0.05$ cm, and is a global muon;
- relative isolation as defined in equation 5.1 less than 0.15 (evaluated within ΔR = 0.4).

Probe muons are defined differently depending on the specific step in the efficiency derivation. There are, however, common requirements that apply to multiple (if not all) steps:

- a tracker-seeded track, or the inner track of a global muon must satisfy the high-purity flag. This criterion reduces the amount of background by approximately a factor of 4 in the failing probe sample for the RECO step, allowing a more accurate measurement of the corresponding efficiency. Almost all muons selected in the entire acceptance region of the W and W-like analyses have inner tracks that already satisfy this requirement. the fraction is at least 99.5% and is 99.9% in most bins. This is also applied to all the other steps to preserve consistency.

- When selecting standalone or global muons, the outer track must have at least 1 valid hit in the muon spectrometer. The contribution of a hit/segment to the track χ^2 is required to be below a tight threshold in order to be used in the track fit in the muon standalone reconstruction, after the seeding stage. Hits with a looser threshold that are kept in the track hit collection for ID purposes, but are not used for the track fit, are called "invalid" hits. Muons that have no valid hit in the muon spectrometer can exist, and they use only the information from the seed. They are expected to have a worse resolution. This distinction between valid and invalid hits is specific to muon reconstruction, and is different from that used in the tracking described in 4.2. This criterion reduces the amount of background by approximately of factor of 2 in the failing probe sample for the $\mathrm{Tracking}$ step. In the following, these valid hits will also be referred to as "MergedStandAloneMuon_numberOfValidHits".
- When selecting global muons, the standalone track must have $p_T > 15$ GeV, for consistency with that of the tracking efficiency, as described in 5.2.1.
- A matching of $\Delta R < 0.3$ is required between the inner and outer track for global muons, for consistency with the matching applied in the RECO step between the standalone muon and the tracker-seeded track described in 5.2.1.

The tag-and-probe pair must have an invariant mass between 60 GeV $< m_{ll} < 120$ GeV for all reconstruction and identification steps, except for $\mathrm{Tracking}$ where the requirement is 50 GeV $< m_{ll} < 130$ GeV, as the probes are reconstructed using the muon spectrometer information and no information from the inner tracker is available. For $\mathrm{Trigger}$ and $\mathrm{Isolation}$, the invariant mass requirement is relaxed to 60 GeV $< m_{ll} < 200$ GeV for bins with $u_T < 0$, since the invariant mass distributions are shifted significantly to higher values for $p_T > 40$ GeV. This is because a probe emitted opposite to the direction of the boson must have p_T smaller than the tag p_T. In the case of a probe of high p_T, this results in a off-shell tag-and-probe pair with an invariant mass above the $Z-$peak. This effect becomes progressively more pronounced as u_T moves towards negative values. The tag and probe are required to have opposite charges for all steps, except for $\mathrm{Tracking}$, due to the usage of standalone muons for which the charge misidentification is not negligible (for the other steps the charge misidentification was found to be 10^{-5}).

When deriving efficiencies as a function of the charge for the $\mathrm{Tracking}$ step, the charge is inferred from the charge of the tag muon, assuming that the probe has opposite charge, allowing us to define the proper Scale Factors to apply to the muons in the analysis even if the charge misidentification for the muon spectrometer is relatively large (the charge for the muons in the analysis is defined from the inner tracker information, since the analysis uses only global muons, for which the track is present by construction).

In addition to the previous criteria, stage-dependent requirements on the probes, both in the numerator and the denominator, are applied:

- RECO:
  - all: tracker-seeded track with high-purity flag, $p_T > 24$ GeV, and $|\eta| < 2.4$. Only for this step, we require $\Delta z_{track} < 0.2$ cm, where Δz_{track} is the difference of the position of the two inner tracks for tag and probe along the beam axis. This criterion reduces the amount of failing probes from background processes by another factor of 2 on top of the high-purity flag;
  - passing: tracker-seeded track matched to a standalone muon with $p_T > 15$ GeV within a cone of $\Delta R = 0.3$.

- $\mathrm{Tracking}$
    - all: standalone muon with $p_T > 24$ GeV and $|\eta| < 2.4$, computed using the muon spectrometer;
    - passing: the standalone muon is also a global muon with inner track $p_T > 10$ GeV (to suppress cases where standalone muons are also global muons formed using a random pileup track, as described in 5.2.1) and outer track $p_T > 15$ GeV, with $\Delta R < 0.3$ between inner and outer track.
- IDIP
    - all: global muon with inner track $p_T > 24$ GeV that passes the $\mathrm{Tracking}$ selection;
    - passing: medium ID and $|d_{\mathrm{xyBS}}| < 0.05$ cm.
- $\mathrm{Trigger}$
    - all: probes that pass of all the stages up to IDIP;
    - passing: object at denominator is matched to a trigger object returned by the `HLT_IsoMu24` or `HLT_IsoTkMu24` within $\Delta R = 0.3$.
- $\mathrm{Isolation}$ for triggering muons
    - all: probes that pass all stages up to $\mathrm{Trigger}$;
    - passing: vertex-agnostic isolation to be less than 15% of the muon p_T.
- $\mathrm{Isolation}$ regardless of trigger (for W-like and Z dilepton analyses)
    - all: probes that pass all stages up to IDIP;
    - passing: vertex-agnostic isolation to be less than 15% of the muon p_T.

In the following, the binning in $p_T - \eta$ (and u_T) for the different steps are summarized:

- η: 48 bins with width of 0.1 from -2.4 to 2.4 for all working points;
- p_T: [24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 47, 50, 55, 60, 65] GeV for IDIP, $\mathrm{Trigger}$, and $\mathrm{Isolation}$;
- p_T: [24, 26, 30, 34, 38, 42, 46, 50, 55, 60, 65] GeV for RECO;
- p_T: [24, 35, 45, 55, 65] GeV for $\mathrm{Tracking}$;
- u_T: [-30, -15, -10, -5, 0, 5, 10, 15, 30, 40, 50, 60, 70, 80, 90, 100] for $\mathrm{Trigger}$, and $\mathrm{Isolation}$.

The RECO and the $\mathrm{Tracking}$ steps are those that have the looser requirements on the probe muons, and their efficiencies are also very close to 1 in most of the $p_T - \eta$ plane, resulting in failing probe templates that are mostly populated by backgrounds, making it difficult to distinguish an invariant mass peak. This results in a coarser binning compared to the other steps: this is particularly true for the $\mathrm{Tracking}$ step, in which the invariant mass distribution is reconstructed using the muon spectrometer for the probes, resulting in a worse resolution. It is also true that the $\mathrm{Tracking}$ efficiencies are expected to be very close to 1 in most of the $p_T - \eta$ plane, and to be mostly flat as a function of p_T, so that the binning adopted for this step can catch the differences between data and Monte Carlo simulation. The analysis uses muons that have p_T reconstructed using the muon spectrometer down to 15 GeV: for these muons, however, the invariant mass fits become very difficult, especially in the high $|\eta|$ regions, because of the high amount of background processes and because the invariant mass shapes for the signal start to share some similarities with the background. What is done is to propagate the $\mathrm{Tracking}$ Scale Factors fitted from 24 to 65 GeV also outside of this range. However, effects due to the p_T resolution in the muon spectrometer should be relevant only for high $|\eta|$, where the p_T resolution reaches 30% (while it is $O(5-10\%)$ in the barrel). This is visible, for example, in figure 5.13, where the probability to reconstruct a global muon decreases mostly in the high $|\eta|$ region when the standalone $p_T > 15$ GeV requirement is imposed. This can also be seen in the results presented in 7.4.10 about the presence of residual clone muons. Therefore, for most of η bins, extrapolating the $\mathrm{Tracking}$ Scale Factors fitted from 24 to 65 GeV also outside of this range should be sufficient to capture their p_T dependence, as the $\mathrm{Tracking}$ efficiencies are mostly flat as a function of p_T. When extracting the corrections for the second muon veto to suppress the $Z \to \mu\mu$ background, the fitting range in the steps relevant to the veto efficiency corrections is extended to 15 GeV, as we reject events with more than one reconstructed muon with $p_T > 15$ GeV. More details are provided in subsection 7.4.9.

The u_T binning is chosen to be narrower around the peak of the u_T distribution, as in figure 5.23, and broader in the tails. Initially, it extended for -30 to 30 GeV, but it was later extended to 100 GeV to better capture the u_T dependence of $\mathrm{Trigger}$ and $\mathrm{Isolation}$ at high values, mostly relevant in the high p_T region of the analysis, but no relevant difference was observed in the corrections. In the region between -30 and 30 GeV, an attempt at doubling the number of bins by dividing the current bins in half was made, and no significant differences were observed, and it was decided to keep the current binning as the invariant mass fits with the coarser u_T division are more stable.

7.4.2 Efficiencies and vertex position

Prior to deriving efficiency corrections, it was necessary to make sure that the beamspot, i.e. the interaction region of two colliding proton bunches, is the same between data and Monte Carlo simulation. Figure 7.6 shows how a shift in the position along the beamline ΔZ translates into a shift in the vertical position Δh in the position at the interception of a particle when it crosses a surface (for example, a layer of the tracker). It can be useful to consider a numerical example: assume that a layer of a detector is at $z = 1.0$ m, which is roughly half of the total extension of the silicon tracker from $z = 0$. If two particles originate for z = 0 with $\eta = 2.3$ for the first one and $\eta = 2.4$ for the second one, so that their different in η is the bin width used in the analysis, one gets a vertical distance between the trajectories of the two particles spanned a $z = 1$ m $\Delta H = 1.96$ cm. However, if the position along the $z-$axis of the particles is moved, for example, by $\Delta z = 10$ cm (the beam width along z is typically around 3.5 cm, see figure 7.8, with tails that can reach up to 10 cm), the trajectories of the particles will move in the vertical direction by $\Delta h = 2.03(1.83)$ cm for $\eta = 2.3(2.4)$, resulting in a shift that is as large as the bin width (if one had considered the full length of the tracker it would roughly be half of it), which could result in the particle not being reconstructed within tracker acceptance.

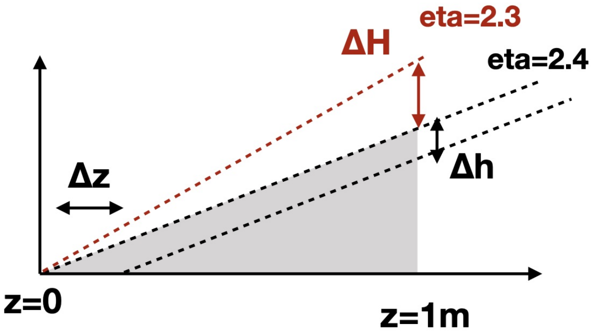
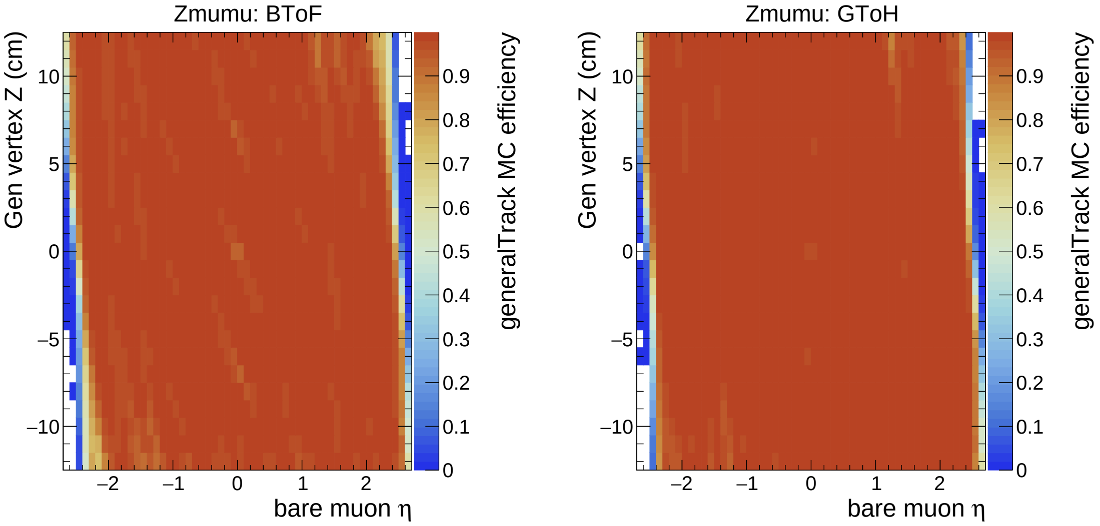


Figure 7.6: Illustrative picture showing how a shift in the vertex position, ΔZ, translates into a vertical shift ΔH in the position where a particle hits a surface.

.

The effect on the efficiency is shown in figure 7.7, where MC truth efficiency for the reconstruction of a tracker-seeded track, here considered only to illustrate the effect, as a function of η and the $z-$position, for the preVFP and postVFP period of 2016. A clear effect on the highest $|\eta|$ bins can be seen. This effect is worse for the preVFP data, which is not used in the analysis because of the saturation effects in the pre-amplifier of the APV25 readout chip, the VFP issue, already described, but it is still present also for postVFP. It was also observed that this affects the reconstruction in the muon spectrometer, although the effect is less pronounced because of its larger dimensions.



(a) General tracks, preVFP

(b) General tracks, postVFP

Figure 7.7: MC truth efficiencies for the reconstruction of tracker-seeded tracks matched to a postFSR muon, for preVFP (left) and postVFP (right) simulations, up to $|\eta| = 2.7$

.

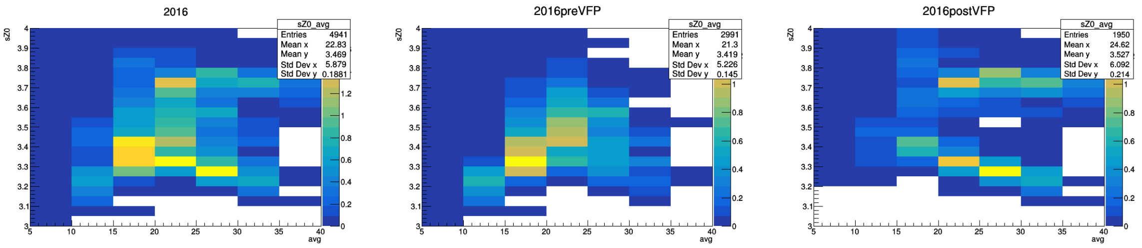
This can bias the efficiencies measured via the tag-and-probe method for high-vertex z and muon η. Two things can happen in this case:

- the muon falls outside of the fiducial tracker acceptance and is not reconstructed;
- the muon is produced within the tracker acceptance but is not measured because of inefficiencies, such as the one caused by the VFP issue.

In either situation, one of the two muons to form the tag-and-probe pair would not be reconstructed.

This also explains why the effects of the VFP issue on the tracking efficiencies cannot be taken into account by the tag-and-probe procedure: the VFP issue results in an effective reduction in the geometrical acceptance of the tracker. This is because the VFP issue results in tracks being reconstructed with fewer hits, with the consequence that it is more likely that the track reconstruction fails. This is particularly true for tracks at high $|\eta|$ and if the vertex corresponding to the $Z \to \mu\mu$ decay is significantly displaced from the center of the tracker, resulting in a reduced probability to form a tag-and-probe pair to use in the efficiency measurements for the steps in which an inner track is required. Therefore, a bias on the efficiencies that cannot be trivially recovered using the tag-and-probe method is induced.

The beamspot width along the $z-$direction has a non-trivial relation with the average pileup, shown in figure 7.8 for the whole of 2016, only preVFP and postVFP. A clear luminosity dependence is visible, as well as a two peak structure, for the inclusive 2016 profile. In the preVFP era, σ_Z increases as the average PU increases (it is decreasing with time, as the average PU decreases as a function of the time during a LHC fill), while in the postVFP era, σ_Z is decreasing as the the average PU increases and the two-fold structure



(a) 2016

(b) 2016 preVFP

(c) 2016 postVFP

Figure 7.8: Beam width along z vs average pileup

mentioned is clearly visible. These trends are compatible with the beam dynamics and the synchrotron radiation observed at the LHC.

In order to account for this time-dependent effect, Scale Factors to match the $z-$position of the vertex between data and Monte Carlo are derived as a function of the average pileup, corresponding to the conditional probability $P(z\text{-position of vertex}|N_{PU})$, and applied to the simulation for the analysis and the efficiency corrections derivation. In data, the distribution of the vertex position is obtained from the beam spot information stored as a function of the different data-taking periods. In MC, it is obtained from a fit to the true position of the vertex at generator level. In real data, the average N_{PU} is obtained from the measured instantaneous luminosity in each of periods relevant for 2016, then dividing by the effective inelastic cross-section: a value of 69.2 mb is assumed.

Figure 7.9 shows the Scale Factors that are actually applied in the analysis: on the left they are shown inclusively in the average pileup to show how they vary as a function of z, on the right they are shown double differentially (as they are applied in the analysis and the efficiency derivation). In the lowest pileup bin the weights can be much large than 1 for $z < 0$ due to both the low statistics and the actual differences in shape between data and simulation events. All the events in analysis, and the efficiency derivations, are also reweighted to match the actual distribution of average pileup observed in data (which tends to be very low for values in the first and last PU bin in figure 7.9).

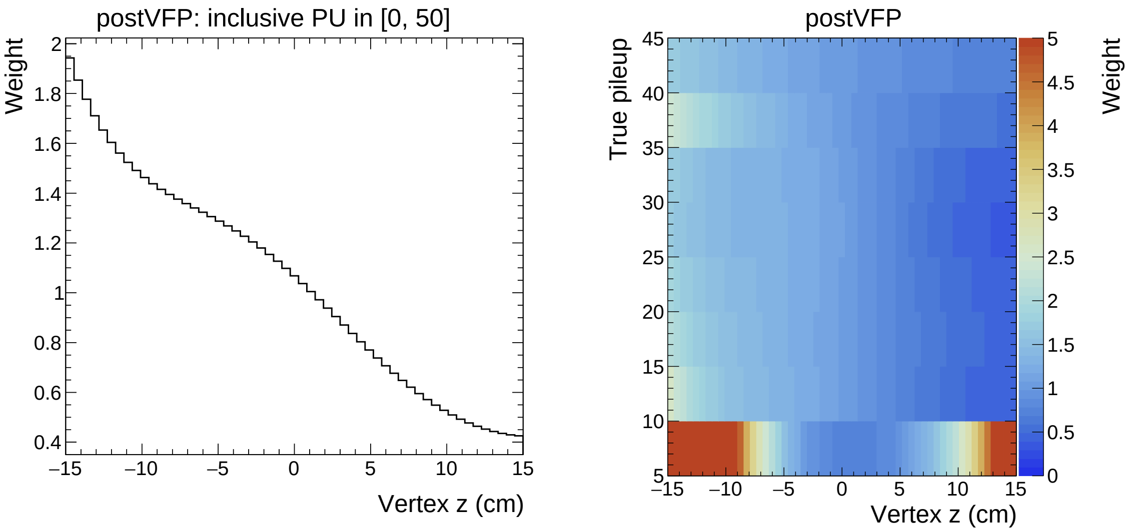


Figure 7.9: Event weights as a function of the true pileup and vertex position along the beam axis. These are used to scale simulated events to match the distribution of the beamspot z position in the corresponding data set. The plot on the left shows the weights inclusively in pileup to illustrate the average shape, the plot on the right shows the double-differential weights that are actually applied in the analysis.

7.4.3 Fits to the invariant mass distributions

The fits to extract the efficiencies in the individual $p_T - \eta - q$ (and u_T) are based on the RooFit library of ROOT [246], which relies on Minuit2 [247] for the numerical minimization. For each $p_T - \eta - q$ bin, a nominal fit, to define the central value of efficiencies and correction, and an alternate fit that employs a different signal model, to provide a systematic variation, are run. In data, the nominal fit model is the sum of a signal and background template. The signal templates are derived from the Monte Carlo sample for each $p_T - \eta - q$ (and u_T) bin, and these are the same samples on which the events were counted to define the nominal efficiency in Monte Carlo described later, convoluted with a Gaussian distribution to take into account differences in resolution between data and simulation. For the passing probe templates of all steps, the alternate signal model is the convolution of a Breit-Wigner and `RooCBExGauss` shape[2] to model the finite resolution in data, while for the failing templates:

- for the RECO and $\mathrm{Tracking}$ steps the signal model is given by the convolution of a Breit-Wigner and a Gaussian distribution;
- for IDIP the signal model is the convolution of a Breit-Wigner and a `RooCBExGauss` shape;
- for $\mathrm{Trigger}$ and $\mathrm{Isolation}$ the signal model is the convolution of a Breit-Wigner and a `RooCBExGauss` summed to a Gaussian distribution to model the second peak at invariant masses around 70 GeV (see for example figure 7.11) that was observed in some of the bins at low p_T. The peak is attributed to final-state radiation, which reduces the isolation efficiency[3].

The background model is an exponential for all the steps in the passing probe templates and for the IDIP, $\mathrm{Trigger}$ and $\mathrm{Isolation}$ steps for the failing probe templates, while for the failing probe templates of the RECO and $\mathrm{Tracking}$ steps a `RooCMS` shape, defines as:

$$\texttt{RooCMSShape} = k \times \mathrm{erf}\left((\alpha - x)\beta\right) \times \exp((\mu_{\mathrm{pole}} - x)\gamma), \tag{7.2}$$

where k is a normalization constant, $\mathrm{erf}\,(x)$ is the error function, $\exp(x)$ is the exponential. This is a probability density function that has exponential decay distribution at high mass beyond the pole position but turns over (i.e., error function) at low mass due to threshold effects. The shape of the background, together with the fact that the signal is much smaller than the backgrounds for the RECO and $\mathrm{Tracking}$ steps explains why a simpler signal model was used in the alternate fits.

For the Monte Carlo simulation, the nominal efficiency is simply given by the event counts in the passing and the failing probe templates for each $p_T - \eta - q$ (and u_T) bins, while the alternate fit model is the same as the one used for the alternate fits in data, also including the shape for the background.

In Monte Carlo, the reconstructed muons are required to be matched to postFSR muon at generator level: the matching is done requiring that ΔR between the reconstructed object and the postFSR muon is less than 0.1 for all of the tag-and-probe steps except for the $\mathrm{Tracking}$ step, in which this is relaxed[4] to 0.3. The background is actually negligible, as the gen-matching largely suppresses the cases with jets misidentified as muons in the failing probe sample, and the nominal and alternate fits in Monte Carlo are in agreement within the statistical uncertainties for most of the η bins and most of the steps. Therefore,

---

[2] A `RooCBExGauss` shape is similar to a Crystal Ball function with left or right tail, but the Gaussian core is made of two Gaussian functions connecting at the peak, with same mean but independent widths on each side

[3] This is an agreement with the discussion in subsection 5.2.3. It was also verified that this second peak is not present when an isolation definition that uses only charge PF hadrons is used, but this was not employed in analysis because the isolation requirement in the trigger at HLT also uses photons.

[4] The gen-matching here is done by matching the reconstructed objects to the generator level muons, which is different from what was done in 5.2.1. With this matching, multiple reconstructed muons that pass the ΔR requirement can be present, and this will be relevant for the studies about the presence of residual clone muons presented in 7.4.10. Since in data there is no gen-matching, this situation is more similar to it. An alternative could be to drop the gen-matching and fit also the background in the MC fits, but this permits to define a MC template for the nominal fits to data.

no systematic uncertainty from changing the fit model in Monte Carlo is applied in the analysis, only the nominal fit is used.

It is crucial that the fits converge properly to ensure that proper efficiencies and Scale Factors are derived: this might in principle require tuning the fit parameters for some of the fits. What was done is to set boundaries or add Gaussian constraints to some of the parameters in the fit and using backup background shapes for the alternate fits in the failing templates, in particular Chebychev polynomials (implemented in the `RooChebychev` class)[5] for the fits that do not converge successfully, have a covariance matrix that is not positive definite, or have a $\chi^2/\mathrm{n.d.f.}$ that is too high. This ensures that almost all of the fits converge successfully automatically, with a very little number still needing dedicated care. The numerical stability of the results and the statistical uncertainties was also assessed by fitting the efficiencies in data for each of the reconstruction and identification steps as a function of the charge splitting all of the events according to the parity of their event number. The distribution of the pulls, defined as difference between the efficiencies derived from odd and even events divided by the quadratic sum of the statistical uncertainties of the odd and even event fits, for all of the $p_T - \eta - q$, separately for each step: they were found to be compatible with 0 and with standard deviation mostly compatible with one, derived from Gaussian fits.

Table 7.5 summarizes the signal and background models for the fits to the mass distributions in data. The systematic uncertainty mostly arises from the alternate fits in data based on the different signal model. An alternate background model based on a template shape obtained directly from simulated background processes (for the prompt component) and from muon pairs with same sign from real data (for misidentified muons, i.e. fakes) has been tested, and a preliminary version of this approach has been used in the analysis as an alternate source of systematic uncertainty, only for the RECO and $\mathrm{Tracking}$ efficiencies, where background is expected to be relevant. The effect was found to be as large as the statistical and systematic uncertainties in the $\mathrm{Tracking}$ Scale Factors, and the corresponding uncertainty propagated to the measured m_W was found to be at the level of 0.5 MeV. The implementation of the statistical and systematic uncertainty is discussed in 7.4.6 and in 7.4.11.

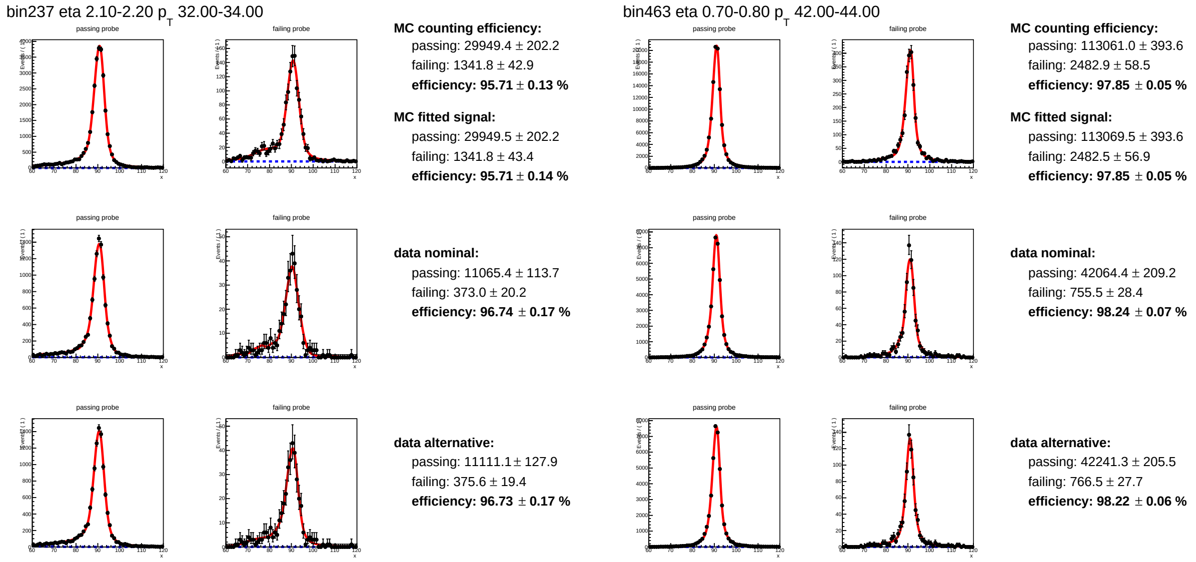
Step		Passing probes		
	Signal (nom.)	Signal (alt.)	Background	
RECO	MC $\otimes$ Gauss	BW $\otimes$ RooCBExGauss	Exp	
$\mathrm{Tracking}$	MC $\otimes$ Gauss	BW $\otimes$ RooCBExGauss	Exp	
IDIP	MC $\otimes$ Gauss	BW $\otimes$ RooCBExGauss	Exp	
$\mathrm{Trigger}$	MC $\otimes$ Gauss	BW $\otimes$ RooCBExGauss	Exp	
$\mathrm{Isolation}$	MC $\otimes$ Gauss	BW $\otimes$ RooCBExGauss	Exp	
Step		Failing probes		
	Signal (nom.)	Signal (alt.)	Background	Background (backup)
RECO	MC $\otimes$ Gauss	BW $\otimes$ Gauss	RooCMS	RooChebychev 3^{rd}
$\mathrm{Tracking}$	MC $\otimes$ Gauss	BW $\otimes$ Gauss	RooCMS	RooChebychev 4^{th}
IDIP	MC $\otimes$ Gauss	BW $\otimes$ RooCBExGauss	Exp	RooChebychev 3^{rd}
$\mathrm{Trigger}$	MC $\otimes$ Gauss	BW $\otimes$ RooCBExGauss $\oplus$ Gauss	Exp	RooChebychev 3^{rd}
$\mathrm{Isolation}$	MC $\otimes$ Gauss	BW $\otimes$ RooCBExGauss $\oplus$ Gauss	Exp	RooChebychev 3^{rd}

Table 7.5: Signal and background models for the fits to the mass distributions in data, for the various steps of the tag-and-probe procedure, for both the nominal and alternate fits. The symbol $\otimes$ denotes a convolution between PDFs, $\oplus$ represents a sum of functions, "MC" denotes the template from Monte Carlo simulation, and "BW" stands for a Breit-Wigner shape. The backup background shape for failing probes is used automatically in the fitting algorithm to recover fits that did not converge properly.

---

[5] The difference in the measured efficiency with respect to using the nominal background model was found to be negligible in most tag-and-probe bins, which was expected because the background component is small in most bins.

Examples of invariant mass fits are shown in the figures 7.10–7.15. For the $\mathrm{Tracking}$ efficiency, the invariant mass for passing pairs is computed using the using the inner tracker information also for the probe muon, resulting in a much better resolution. The aforementioned peak at 70 GeV due to final-state radiation can be seen in figure 7.11.



MC counting efficiency:
passing: 29949.4 ± 202.2
failing: 1341.8 ± 42.9
efficiency: 95.71 ± 0.13 %

MC fitted signal:
passing: 29949.5 ± 202.2
failing: 1341.8 ± 43.4
efficiency: 95.71 ± 0.14 %

data nominal:
passing: 11065.4 ± 113.7
failing: 373.0 ± 20.2
efficiency: 96.74 ± 0.17 %

data alternative:
passing: 11111.1 ± 127.9
failing: 375.6 ± 19.4
efficiency: 96.73 ± 0.17 %

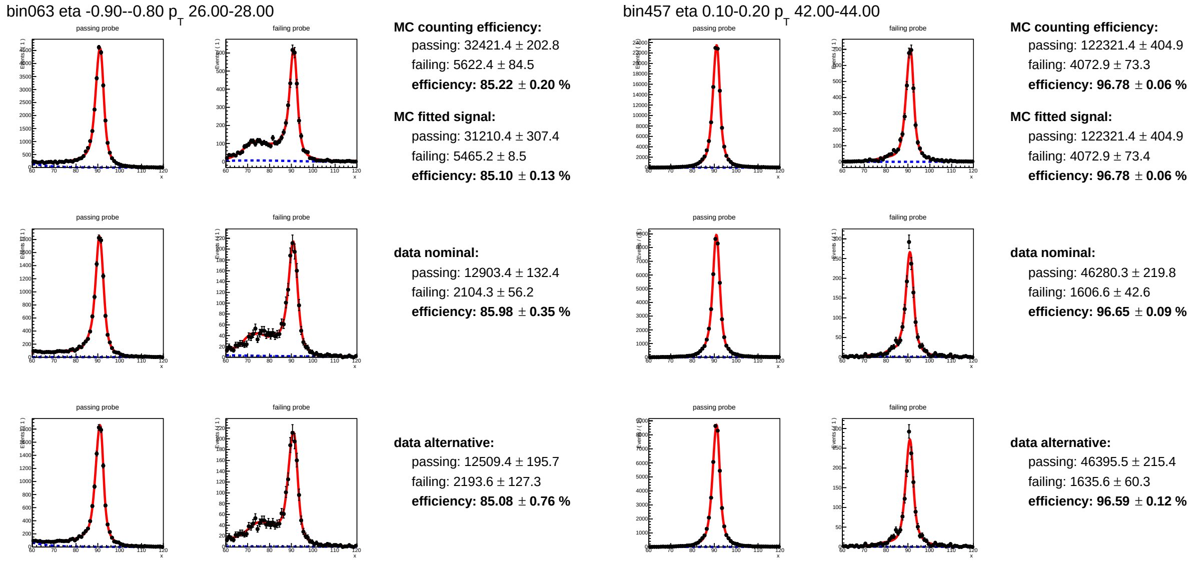
MC counting efficiency:
passing: 113061.0 ± 393.6
failing: 2482.9 ± 58.5
efficiency: 97.85 ± 0.05 %

MC fitted signal:
passing: 113069.5 ± 393.6
failing: 2482.5 ± 56.9
efficiency: 97.85 ± 0.05 %

data nominal:
passing: 42064.4 ± 209.2
failing: 755.5 ± 28.4
efficiency: 98.24 ± 0.07 %

data alternative:
passing: 42241.3 ± 205.5
failing: 766.5 ± 27.7
efficiency: 98.22 ± 0.06 %

Figure 7.10: Mass distributions for $\mathrm{Isolation}$ efficiency (on top of $\mathrm{Trigger}$) at $2.1 < \eta < 2.2$ and $32 < p_T < 34$ GeV (left), and $0.7 < \eta < 0.8$ and $42 < p_T < 44$ GeV (right).



MC counting efficiency:
passing: 32421.4 ± 202.8
failing: 5622.4 ± 84.5
efficiency: 85.22 ± 0.20 %

MC fitted signal:
passing: 31210.4 ± 307.4
failing: 5465.2 ± 8.5
efficiency: 85.10 ± 0.13 %

data nominal:
passing: 12903.4 ± 132.4
failing: 2104.3 ± 56.2
efficiency: 85.98 ± 0.35 %

data alternative:
passing: 12509.4 ± 195.7
failing: 2193.6 ± 127.3
efficiency: 85.08 ± 0.76 %

MC counting efficiency:
passing: 122321.4 ± 404.9
failing: 4072.9 ± 73.3
efficiency: 96.78 ± 0.06 %

MC fitted signal:
passing: 122321.4 ± 404.9
failing: 4072.9 ± 73.4
efficiency: 96.78 ± 0.06 %

data nominal:
passing: 46280.3 ± 219.8
failing: 1606.6 ± 42.6
efficiency: 96.65 ± 0.09 %

data alternative:
passing: 46395.5 ± 215.4
failing: 1635.6 ± 60.3
efficiency: 96.59 ± 0.12 %

Figure 7.11: Mass distributions for $\mathrm{Isolation}$ efficiency (without the requirement on the $\mathrm{Trigger}$) at $-0.9 < \eta < -0.8$ and $26 < p_T < 28$ GeV (left), and $0.1 < \eta < 0.2$ and $42 < p_T < 44$ GeV (right). The second bump at lower mass in the low p_T bin originates from final-state radiation.

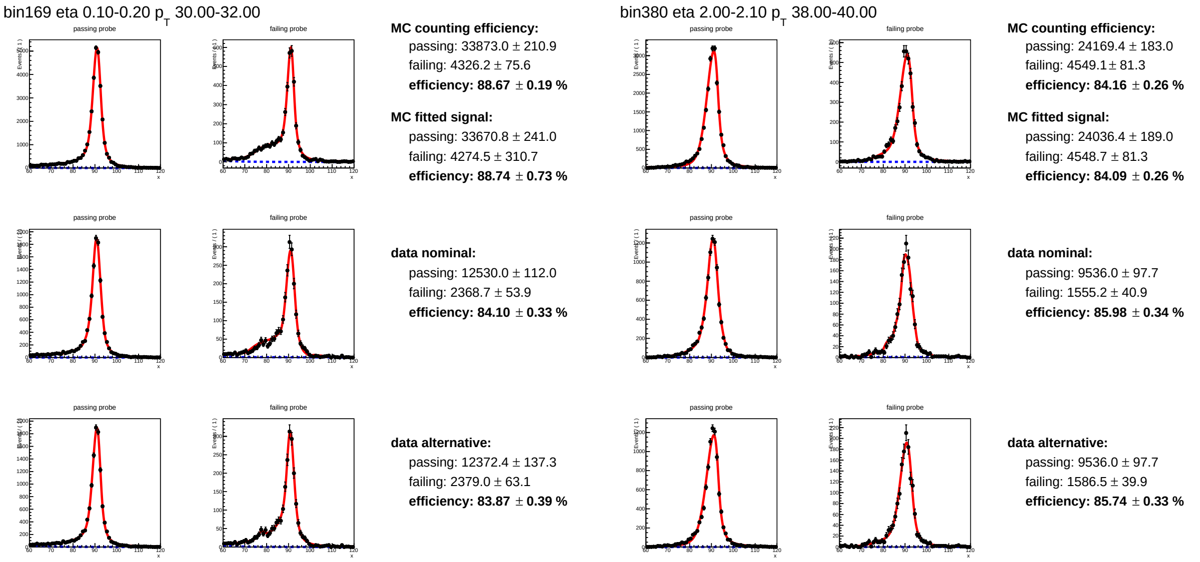


Figure 7.12: Mass distributions for $\mathrm{Trigger}$ efficiency (for positive charge) at $0.1 < \eta < 0.2$ and $30 < p_T < 32$ GeV (left), and $2.0 < \eta < 2.1$ and $38 < p_T < 40$ GeV (right).

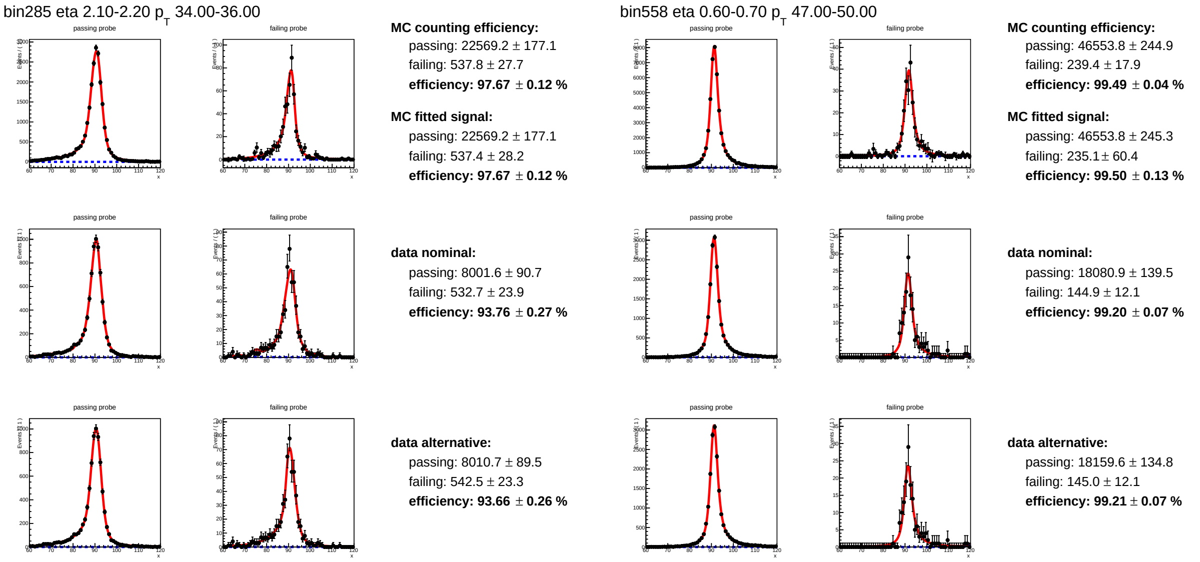
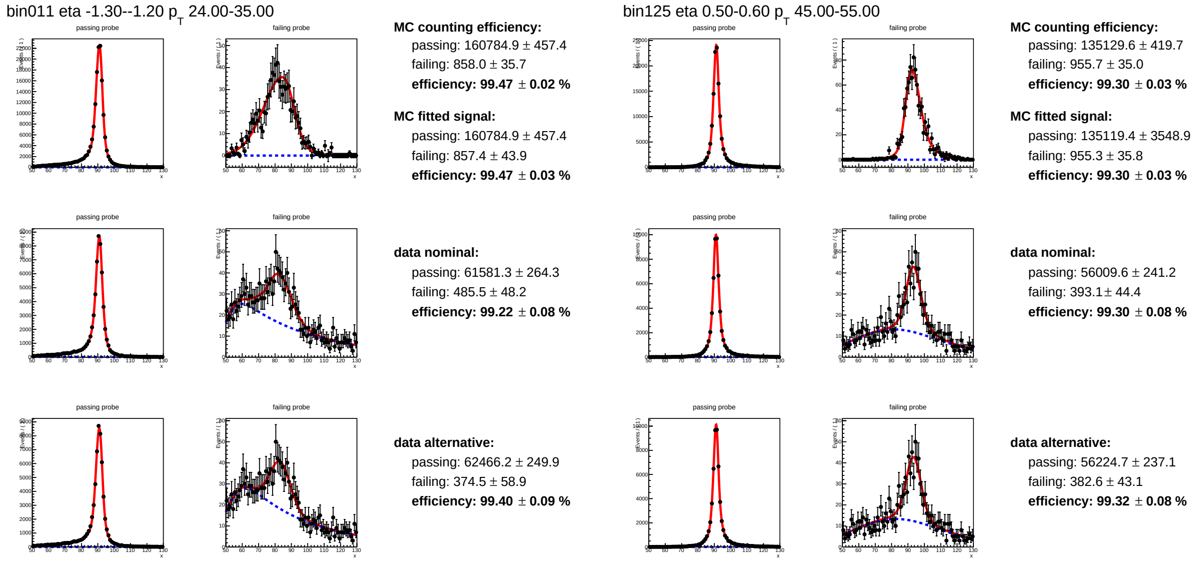


Figure 7.13: Mass distributions for IDIP (for positive charge) at $2.1 < \eta < 2.2$ and $34 < p_T < 36$ GeV (left), and $0.6 < \eta < 0.7$ and $47 < p_T < 50$ GeV (right).

7.4.4 Charge asymmetry in efficiencies and Scale Factors

Muon reconstruction and identification efficiencies, and consequently Scale Factors, are derived separately for positive and negative charge for all steps except for $\mathrm{Isolation}$ (both on top of $\mathrm{Trigger}$ and IDIP), since some charge-dependence (e.g. from alignment) can be expected a priori. Regarding the $\mathrm{Trigger}$, figure 7.16 shows the ratios of efficiencies and Scale Factors between positive and negative charge, the bottom row also shows the charge asymmetry, defined as the difference of efficiencies between charge divided by their sum. An asymmetry can clearly be observed, especially in data at high $|\eta|$ and in some specific η bins corresponding to the holes in the muon spectrometer (figure 3.10). It seems that this feature has been present since Run 1, and it might be related to imperfect alignment affecting the *Level 1* (see 3.2.1)



MC counting efficiency:
passing: 160784.9 ± 457.4
failing: 858.0 ± 35.7
efficiency: 99.47 ± 0.02 %

MC fitted signal:
passing: 160784.9 ± 457.4
failing: 857.4 ± 43.9
efficiency: 99.47 ± 0.03 %

data nominal:
passing: 61581.3 ± 264.3
failing: 485.5 ± 48.2
efficiency: 99.22 ± 0.08 %

data alternative:
passing: 62466.2 ± 249.9
failing: 374.5 ± 58.9
efficiency: 99.40 ± 0.09 %

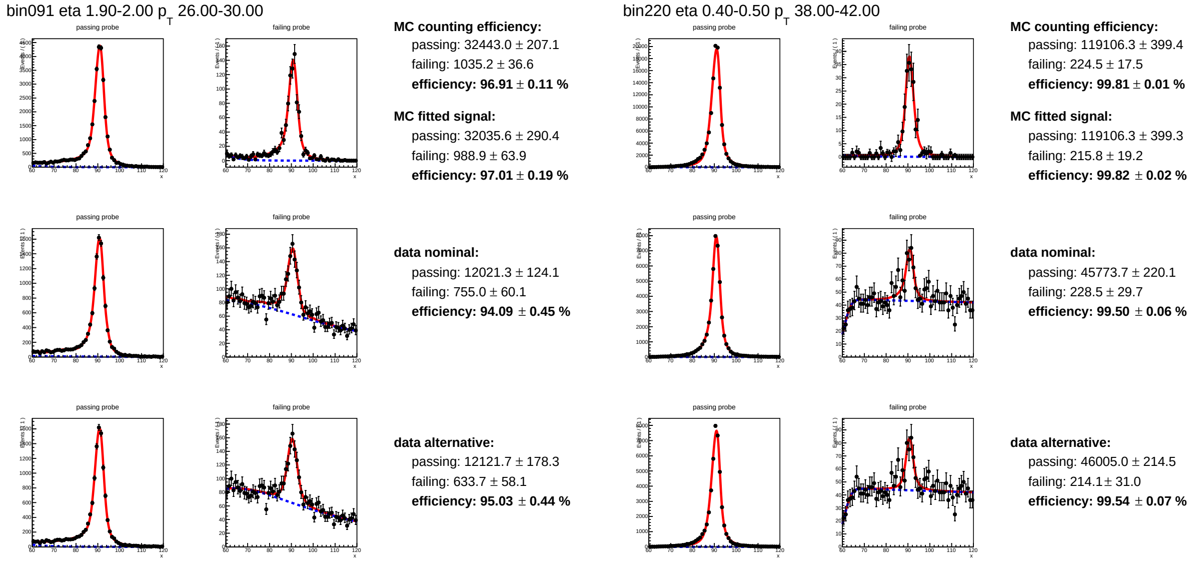
MC counting efficiency:
passing: 135129.6 ± 419.7
failing: 955.7 ± 35.0
efficiency: 99.30 ± 0.03 %

MC fitted signal:
passing: 135119.4 ± 3548.9
failing: 955.3 ± 35.8
efficiency: 99.30 ± 0.03 %

data nominal:
passing: 56009.6 ± 241.2
failing: 393.1 ± 44.4
efficiency: 99.30 ± 0.08 %

data alternative:
passing: 56224.7 ± 237.1
failing: 382.6 ± 43.1
efficiency: 99.32 ± 0.08 %

Figure 7.14: Mass distributions for $\mathrm{Tracking}$ efficiency (for positive charge) at $-1.3 < \eta < -1.2$ and $24 < p_T < 35$ GeV (left), and $0.5 < \eta < 0.6$ and $45 < p_T < 55$ GeV (right).



MC counting efficiency:
passing: 32443.0 ± 207.1
failing: 1035.2 ± 36.6
efficiency: 96.91 ± 0.11 %

MC fitted signal:
passing: 32035.6 ± 290.4
failing: 988.9 ± 63.9
efficiency: 97.01 ± 0.19 %

data nominal:
passing: 12021.3 ± 124.1
failing: 755.0 ± 60.1
efficiency: 94.09 ± 0.45 %

data alternative:
passing: 12121.7 ± 178.3
failing: 633.7 ± 58.1
efficiency: 95.03 ± 0.44 %

MC counting efficiency:
passing: 119106.3 ± 399.4
failing: 224.5 ± 17.5
efficiency: 99.81 ± 0.01 %

MC fitted signal:
passing: 119106.3 ± 399.3
failing: 215.8 ± 19.2
efficiency: 99.82 ± 0.02 %

data nominal:
passing: 45773.7 ± 220.1
failing: 228.5 ± 29.7
efficiency: 99.50 ± 0.06 %

data alternative:
passing: 46005.0 ± 214.5
failing: 214.1 ± 31.0
efficiency: 99.54 ± 0.07 %

Figure 7.15: Mass distributions for RECO efficiency (for positive charge) at $1.9 < \eta < 2.0$ and $26 < p_T < 30$ GeV (left), and $0.4 < \eta < 0.5$ and $38 < p_T < 42$ GeV (right).

muon p_T determination.

For the other steps, the asymmetry is smaller but still present. For this, we derived the distribution of the pulls, defined as the difference between the efficiencies for positive and negative probes divided by the quadrature sum of their statistical uncertainties. Figures 7.17 and 7.18 show these pulls for the IDIP and RECO steps. It can be seen that the means of the pulls distributions are shifted from 0 (corresponding to an absence of charge dependence), particularly for RECO steps. The standard deviations of the distributions for the pulls for the IDIP step are compatible with one within errors, while they are larger for the RECO step.

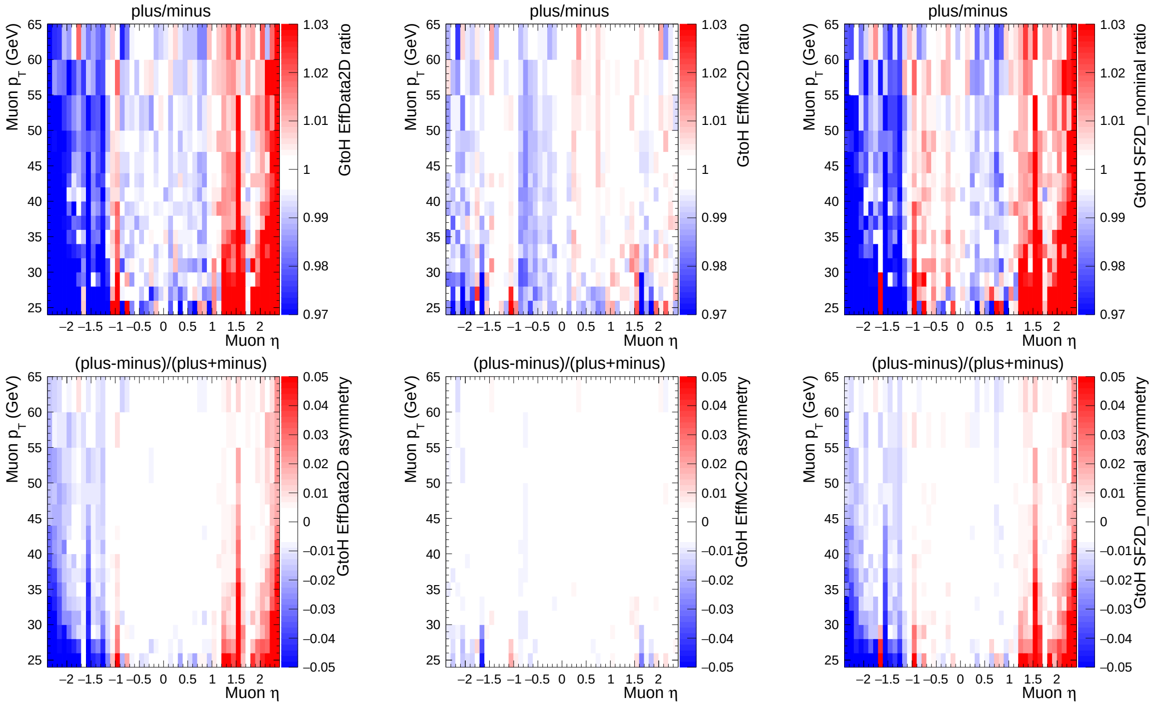


Figure 7.16: Ratio (top) and asymmetry (bottom) of $\mathrm{Trigger}$ efficiencies in data (left) or MC (middle), and Scale Factors (right) between positive and negative charges.

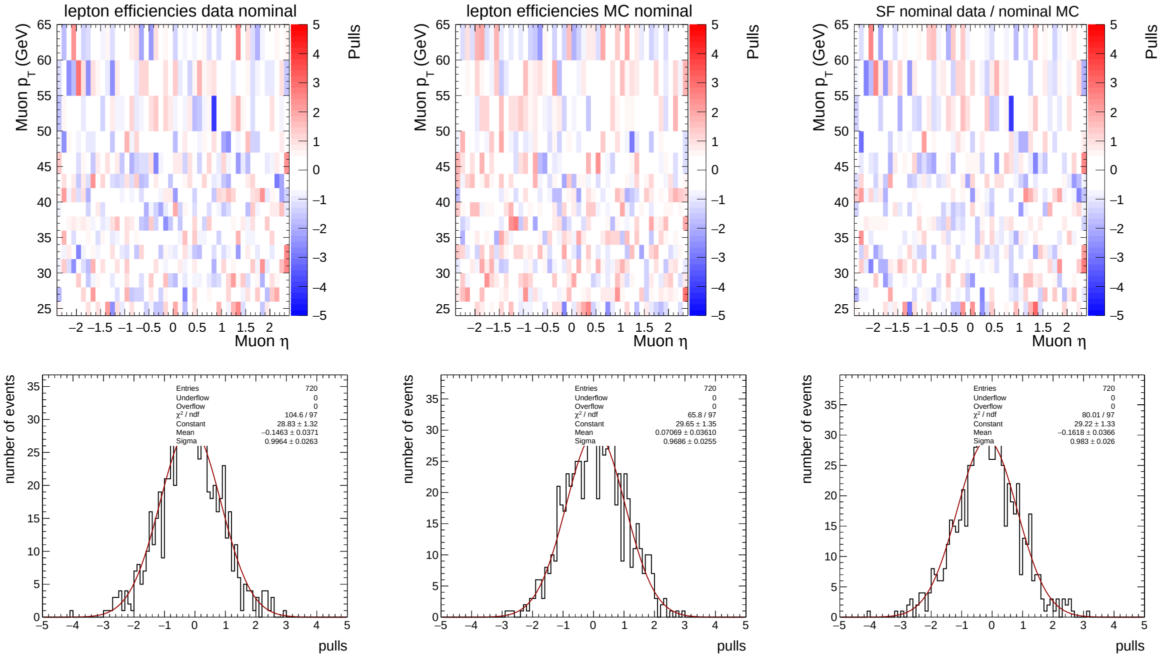


Figure 7.17: Distribution of the pulls for the evaluation of the charge asymmetry in IDIP efficiencies. Pulls are shown for data efficiency (left), MC efficiency from counting (center), and the resulting Scale Factors (right), either in 2D as a function $p_T - \eta$ (top), or in 1D (bottom). The mean and width reported in the legend for the bottom plots refer to the Gaussian used to fit these distributions.

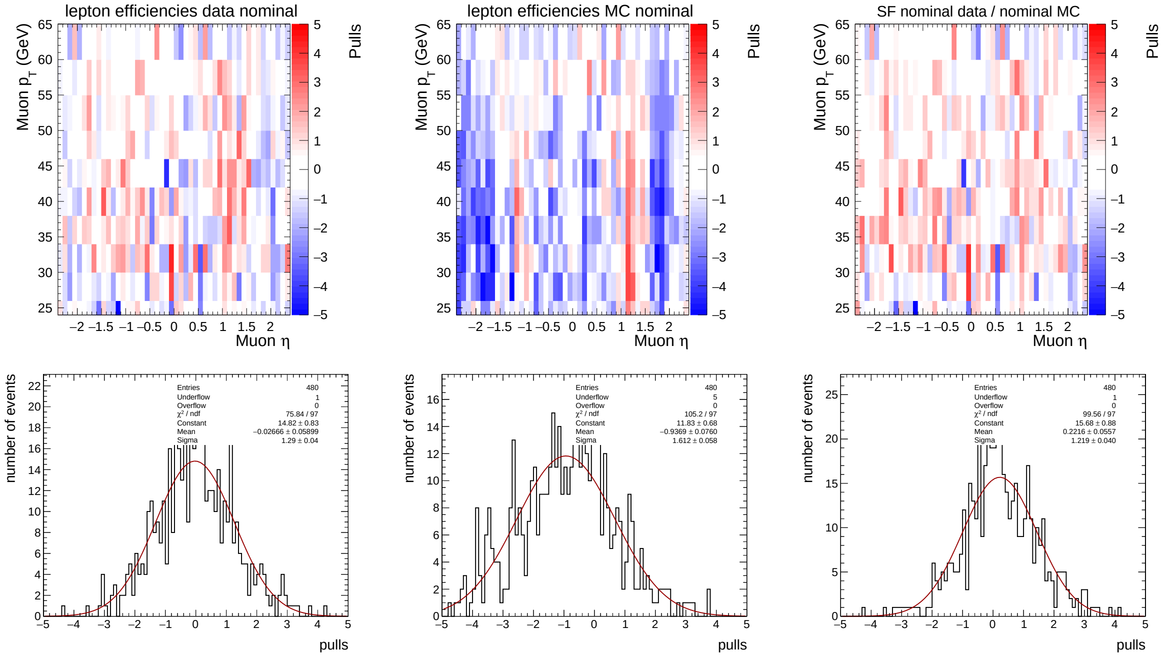


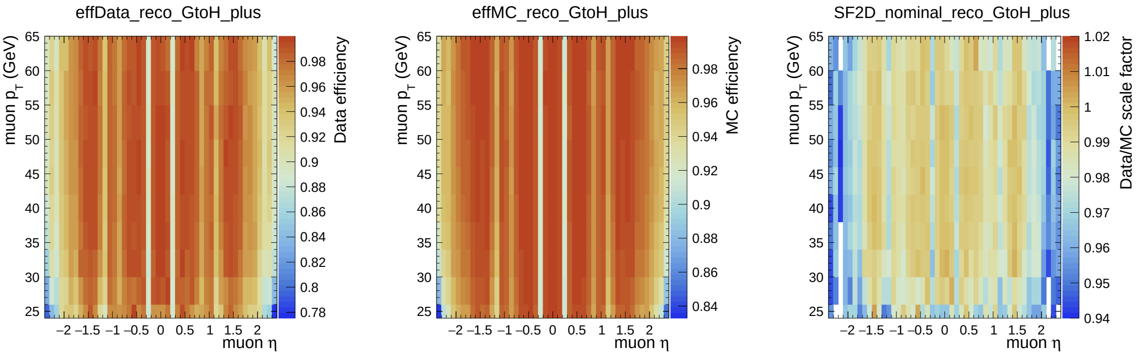
Figure 7.18: Distribution of the pulls for the evaluation of the charge asymmetry in RECO efficiencies. Pulls are shown for data efficiency (left), MC efficiency from counting (center), and the resulting Scale Factors (right), either in 2D as a function of $p_T - \eta$ (top), or in 1D (bottom). The mean and width reported in the legend for the bottom plots refer to the Gaussian used to fit these distributions.

7.4.5 Efficiencies as a function of $p_T^\mu - \eta^\mu - q^\mu$

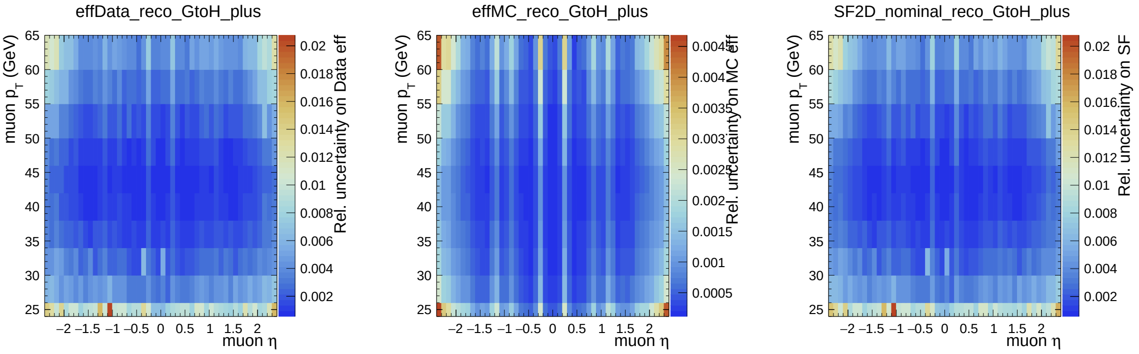
The efficiencies and Scale Factors are shown for the different steps in figures 7.19 –7.25. The bottom rows of each plot show the Scale Factors derived from the nominal and alternate fit, and their ratio. The efficiencies and Scale Factors are shown for the steps individually, unlike in subsection 5.2.1 where the validation of tag-and-probe procedure was done on the product of the steps. For the RECO step, the features due to the holes in the muon spectrometer and the requirement on standalone $p_T > 15$ GeV at high $|\eta|$ can be observed, and the Scale Factors differ from one up to the 10% level, the largest differences are observed at high $|\eta|$, particularly for low p_T. The Scale Factors derived from the alternate fit differ mostly at low p_T at the percent level, and seem to pull the nominal Scale Factors in the same direction. This factor was taken into account when implementing the systematic uncertainties from the alternate fits. The statistical uncertainties are $O(0.5\%)$ in most of the $p_T - \eta$ plane and reach a maximum of 2%, and are larger in data than in Monte Carlo.

The $\mathrm{Tracking}$ efficiencies differ from 1 by less than 1% in most of the $p_T - \eta$ plane, except for $|\eta| = 1.1$ and the highest $|\eta|$ bins, with an asymmetry in η clearly visible in data an Monte Carlo simulation, likely due to the beamspot position, as explained in 7.4.2. The Scale Factors are in agreement with 1 at the per-mille level, exhibiting an asymmetry in the highest $|\eta|$ bins. The Scale Factors exhibit a slight decrease at $O(0.1\%)$ as p_T increases for most η bins. The statistical uncertainties are mostly at $O(0.1\%)$, except for particular bins in data, for which an automated strategy to recover the fits described in 7.4.3 might have occurred, and in the highest p_T bin, where the total statistics available is reduced. The alternate fits mostly give consistent Scale Factors, with a few differences for particular $p_T - \eta$, except for the highest p_T bin. These Scale Factors are applied in the analysis as a function of the standalone p_T and η, as these are the variables used in the derivation of the Scale Factors.

The IDIP efficiencies are mostly compatible with 1 within 1% except for $|\eta| = 0.2$ and $|\eta| = 2.2$, and the largest differences between data and Monte Carlo are around $|\eta| = 2.2$, otherwise the Scale Factors



(a) μ^+ RECO eff. (Data) (b) μ^+ RECO eff. (MC) (c) μ^+ RECO SF



(d) rel. stat. unc. in eff. (data) (e) rel. stat. unc. in eff. (MC) (f) rel. stat. unc. in SF

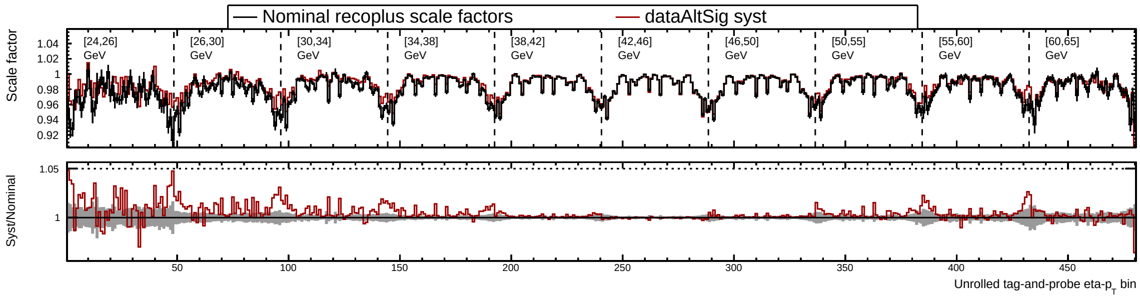


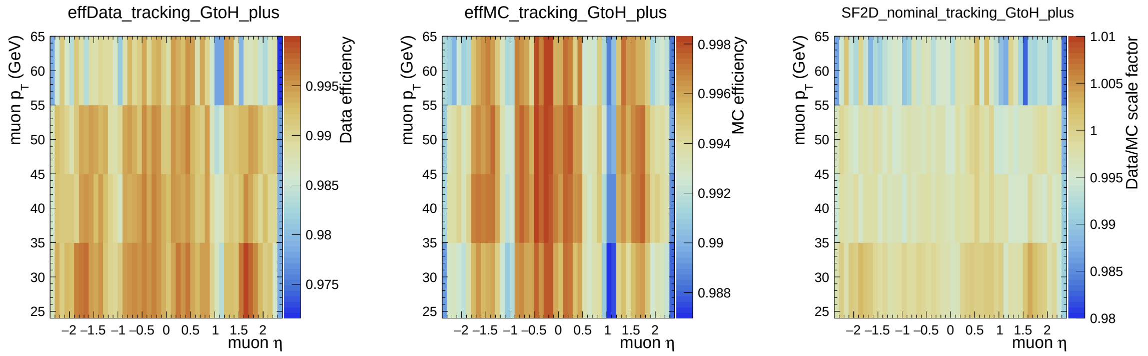
Figure 7.19: RECO efficiencies in data (left) and MC (center), and Scale Factors (right) from tag-and-probe, for positive charge. The relative statistical uncertainties for each of the quantities are shown in the middle. The bottom plot shows the unrolled Scale Factors with their statistical (gray band) and systematic (red line) uncertainties.

appear to differ from 1 at the 0.5% level. The statistical uncertainties are less than 0.1% except for $|\eta| = 2.2$, as the backgrounds in the invariant mass fits are very low and the efficiencies close to unity[6]. The Scale Factors are mostly independent of p_T and η, except for $|\eta| = 2.2$. The variations induced by the alternate signal model seem to be mostly compatible with the statistical uncertainties.

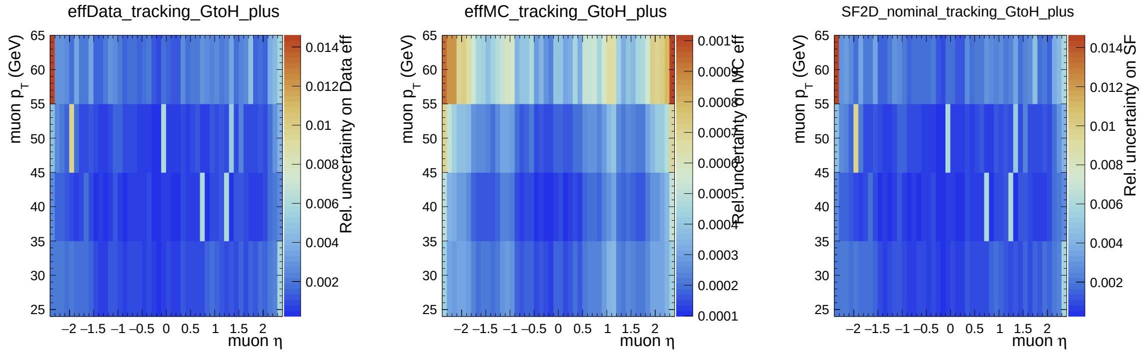
For the $\mathrm{Trigger}$ efficiencies the features related to the muon spectrometer are visible, including an asymmetry, different for the positive and the negative charge, as already discussed in 7.4.4. The effect of the muon spectrometer is more pronounced in data than in Monte Carlo simulation. The efficiencies also show the effect of the turn-on of the trigger in the lowest p_T bin. The statistical uncertainties are larger

---

[6] Assuming a binomial distribution with total probes N and p the probability that a probe passes the selection, the efficiency uncertainty is $\sqrt{\frac{p(1-p)}{N}}$; for invariant mass fits, it is $\frac{\sqrt{\sigma_p^2 N_f^2 + \sigma_f^2 N_p^2}}{N^2}$, where σ_i and N_i are the errors and counts of signal events in passing and failing templates.



(a) μ^+ Tracking eff. (data) (b) μ^+ Tracking eff. (MC) (c) μ^+ Tracking SF



(d) rel. stat. unc. in eff. (data) (e) rel. stat. unc. in eff. (MC) (f) rel. stat. unc. in SF

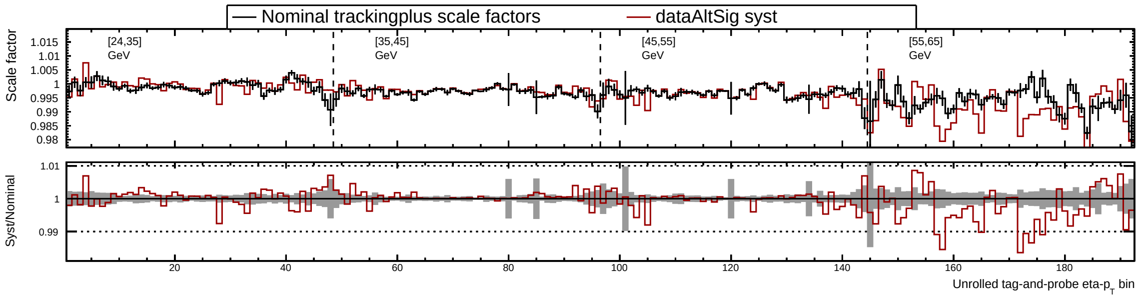
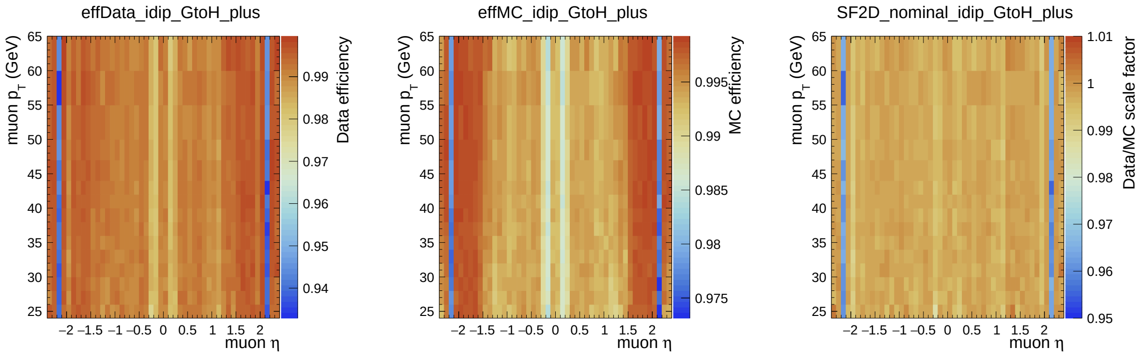


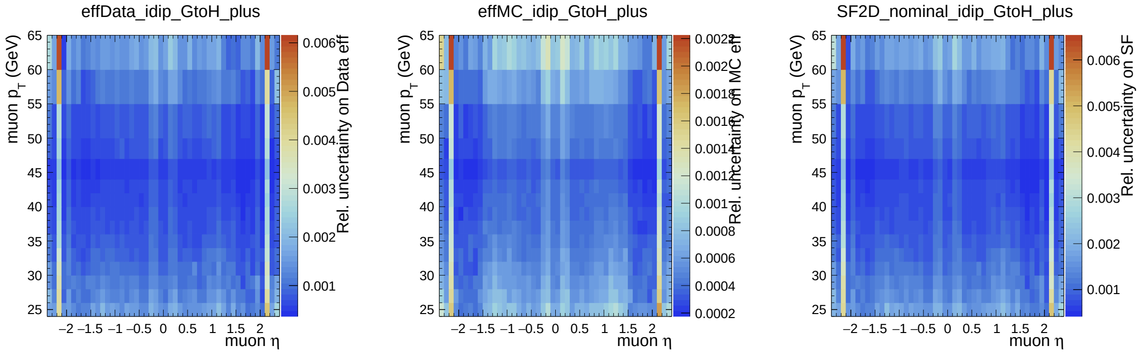
Figure 7.20: $\mathrm{Tracking}$ efficiencies in data (left) and MC (center), and Scale Factors (right) from tag-and-probe, for positive charge. The relative statistical uncertainties for each of the quantities are shown in the middle. The bottom plot shows the unrolled Scale Factors with their statistical (gray band) and systematic (red line) uncertainties.

than IDIP because the efficiencies are lower, but they are still mostly below 1%. The Scale Factors derived using the alternate fit mostly differ at low p_T, potentially because of the presence of the bump at 70 GeV due to final-state radiation, already visible in figure 7.12, which might be captured differently by the signal model derived from Monte Carlo and the analytic model, which just employs a Gaussian to model it.

The $\mathrm{Isolation}$ efficiencies are strongly dependent on p_T, as the isolation requirement is tighter for lower than for higher p_T, and in the low p_T region it also varies as a function of η: this reflects the structures of the several subdetectors to reconstruct the Particle Flow candidates that enter in the definition of isolation. For example, the electromagnetic calorimeter transitions between the barrel and the endcap around $|\eta| = 1.5$ (see figure 3.7), and the reconstruction in the endcap has a lower efficiency, resulting in fewer particles that enter the isolation cone. This region also corresponds to the hadronic barrel region, for which the same



(a) μ^+ IDIP eff. (data) (b) μ^+ IDIP eff. (MC) (c) μ^+ IDIP SF



(d) rel. stat. unc. in eff. (data) (e) rel. stat. unc. in eff. (MC) (f) rel. stat. unc. in SF

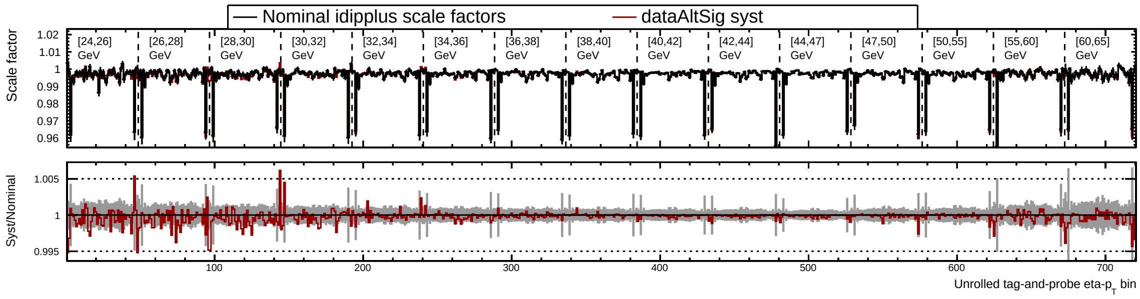


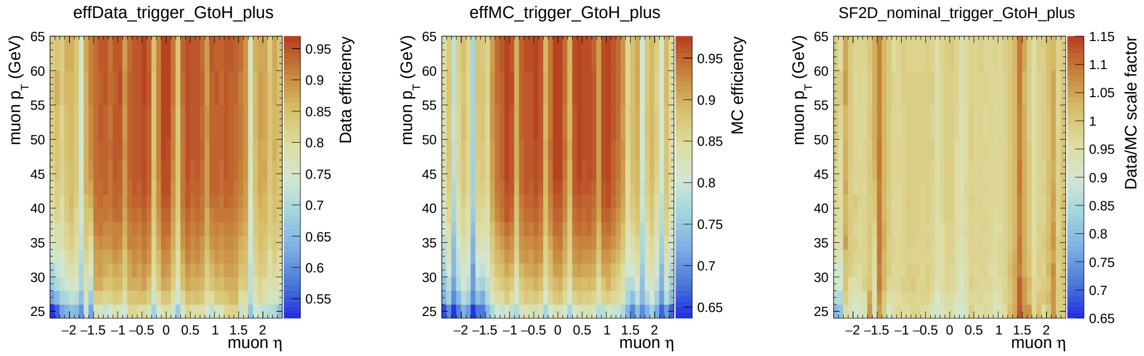
Figure 7.21: IDIP efficiencies in data (left) and MC (center), and Scale Factors (right) from tag-and-probe, for positive charge. The relative statistical uncertainties for each of the quantities are shown in the middle. The bottom plot shows the unrolled Scale Factors with their statistical (gray band) and systematic (red line) uncertainties.

argument applies[7]. These features are similar between $\mathrm{Isolation}$ with and without the trigger requirement, with the latter set of efficiencies being lower, as part of the isolation requirement is already present in the $\mathrm{Trigger}$ efficiency. The Scale Factors differ from 1 up to 2.5% (3% when the trigger requirement is not imposed), mostly at low p_T. The statistical uncertainties are mostly $O(0.1\%)$ except for the lowest p_T bins, as the statistics is lower and the efficiencies are also lower. The same considerations about the Scale Factors from the alternate fits as for the $\mathrm{Trigger}$ efficiency apply here.

The $\mathrm{Trigger}$ and $\mathrm{Isolation}$ efficiencies described here are not used directly in the analysis, as they do not take into account the u_T dependence discussed in 5.2.3, but they are used to define the systematic uncertainties from the alternate fits: the fits with the alternate signal model are not run for each of the

---

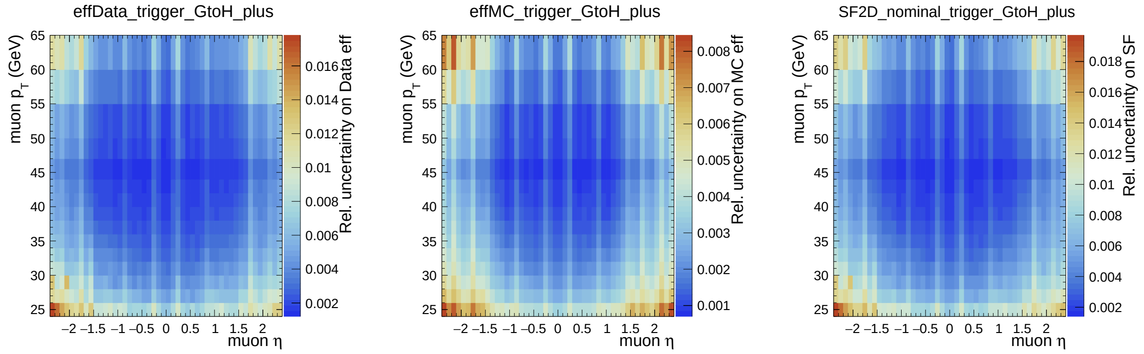
[7] Charged particles are still reconstructed using the inner tracker, but the hadronic calorimeter is used to identify them as charged hadrons.



(a) μ^+ Trigger eff. (data)

(b) μ^+ Trigger eff. (MC)

(c) μ^+ trigger SF



(d) rel. stat. unc. in eff. (data)

(e) rel. stat. unc. in eff. (MC)

(f) rel. stat. unc. in SF

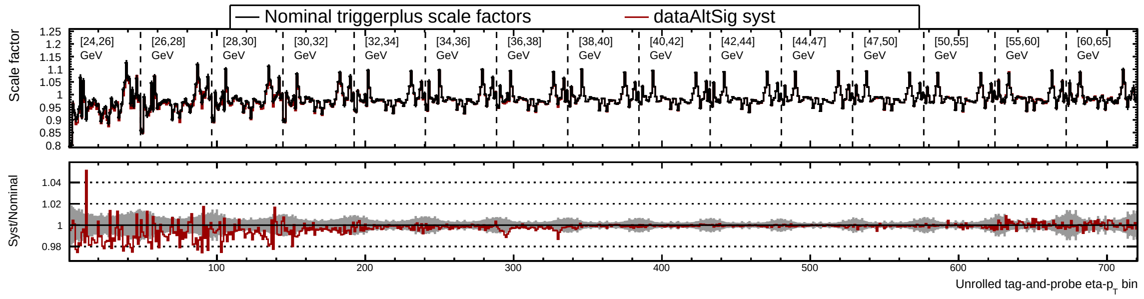
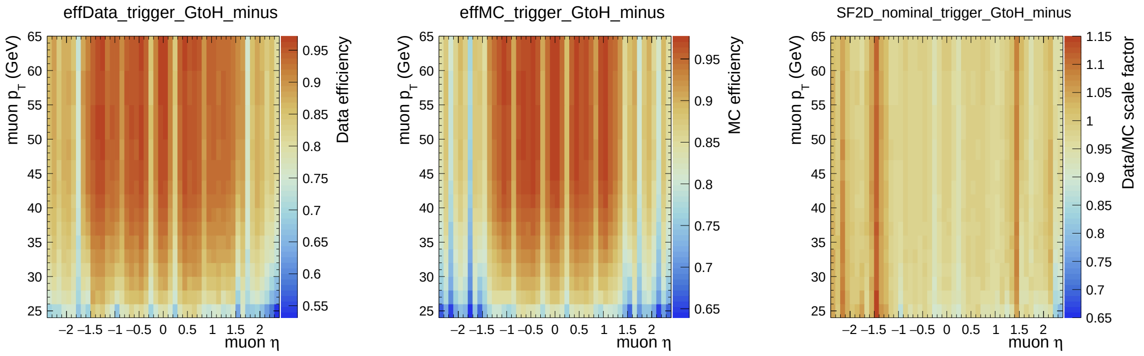


Figure 7.22: Trigger efficiencies in data (left) and MC (center), and Scale Factors (right) from tag-and-probe, for positive charge. The relative statistical uncertainties for each of the quantities are shown in the middle. The bottom plot shows the unrolled Scale Factors with their statistical (gray band) and systematic (red line) uncertainties. A clear asymmetry in η is visible in data.

$p_T - \eta - u_T$ bins, as it is very difficult to ensure the proper convergence of some of them, given the large number of fits. The ratio between Scale Factors derived from the alternate fits and those derived from the nominal fits are used to define this systematic uncertainty, as the effect should become similar when integrating in u_T. The nominal Scale Factors without u_T integration were the first to be derived, but the comparisons in figures 5.23 and 7.26 show that the Monte Carlo does not reproduce correctly the dependence of u_T of trigger and isolation (in this case, the corrections should be flat in u_T, so using Scale Factors integrated over or differential in u_T should yield the same results).

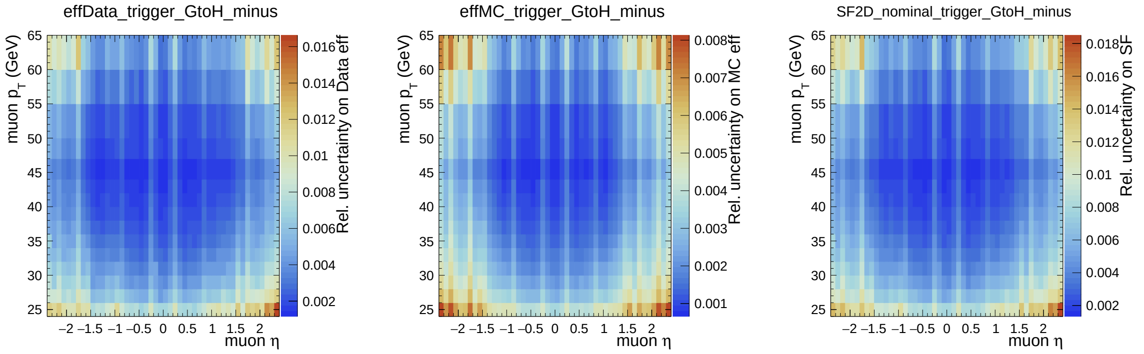
This was also cross-checked with a bias test in which pseudodata were built from simulation using the Scale Factors derived differentially as a function of u_T and the nominal templates are built using those derived without the u_T dependence. The W boson mass is then fit on the biased pseudodata. The result of this bias test was a shift in m_W of 7 MeV, inducing us to implement the dependence on u_T in the analysis, as described in 7.4.6.



(a) μ^- Trigger eff. (data)

(b) μ^- Trigger eff. (MC)

(c) μ^- Trigger SF



(d) rel. stat. unc. in eff. (data)

(e) rel. stat. unc. in eff. (MC)

(f) rel. stat. unc. in SF

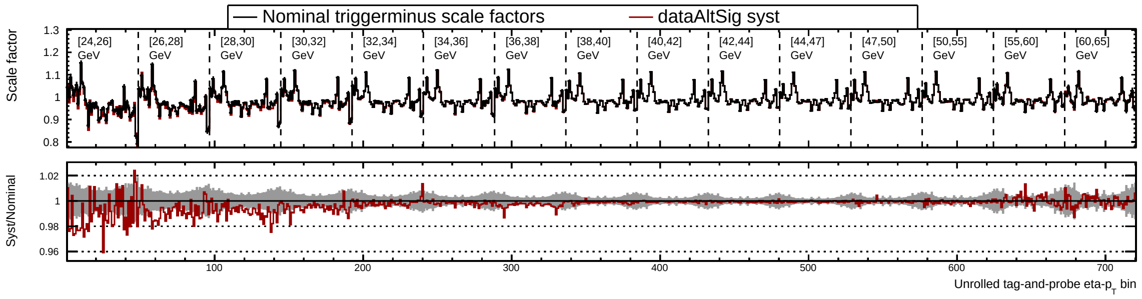
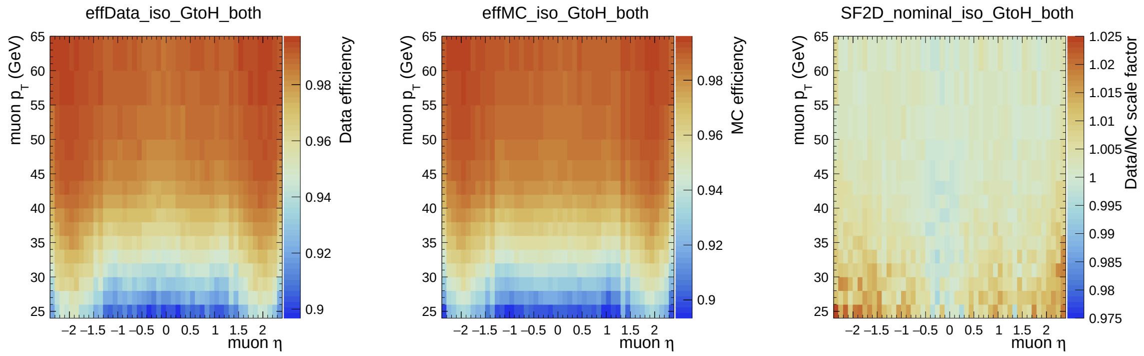


Figure 7.23: Trigger efficiencies in data (left) and MC (center), and Scale Factors (right) from tag-and-probe, for negative charge. The relative statistical uncertainties for each of the quantities are shown in the middle. The bottom plot shows the unrolled Scale Factors with their statistical (gray band) and systematic (red line) uncertainties. A clear asymmetry in η is visible in data.

7.4.6 Smoothing of Scale Factors

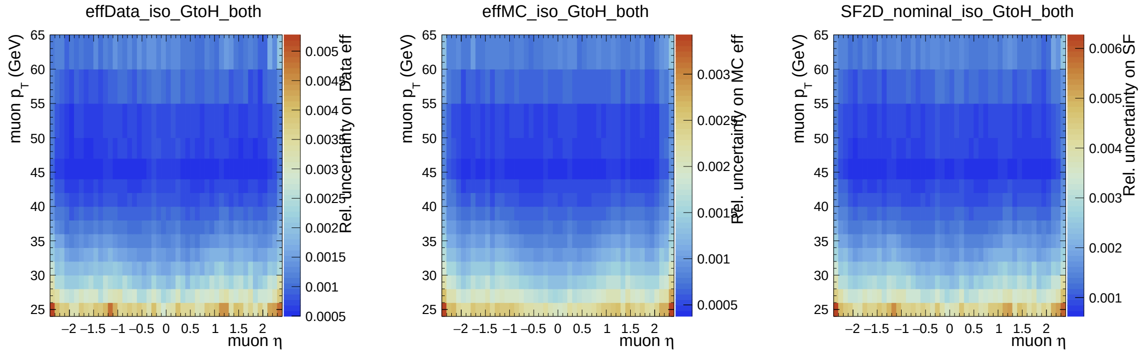
The Scale Factors derived as a function of $p_T - \eta - q$ and u_T for Trigger and Isolation are not used directly in the analysis, but are instead replaced by functions of p_T or $p_T - u_T$ for each η bin. We refer to this procedure as *smoothing*. This is done with the aim of reducing the statistical fluctuations in the Scale Factors, and therefore their impact on the mass, but also to regularize the corrections as a function of p_T, so as to reduce the impact of potential individual and sporadic fits for which the Scale Factor might deviate significantly from those in adjacent bins and to use corrections that are continuous and not binned, to avoid the discontinuities due to the finite binning. This is particularly important for the Isolation efficiency: as will be discussed in 7.7, the isolation variable is used to define a control region that is highly populated by events in which muons are produced as a result of heavy flavour decays (the so-called nonprompt background), by reversing the requirement applied to the muons in the analysis. The Isolation efficiency on top of the



(a) Isolation eff. (data)

(b) Isolation eff. (MC)

(c) Isolation SF



(d) rel. stat. unc. in eff. (data)

(e) rel. stat. unc. in eff. (MC)

(f) rel. stat. unc. in SF

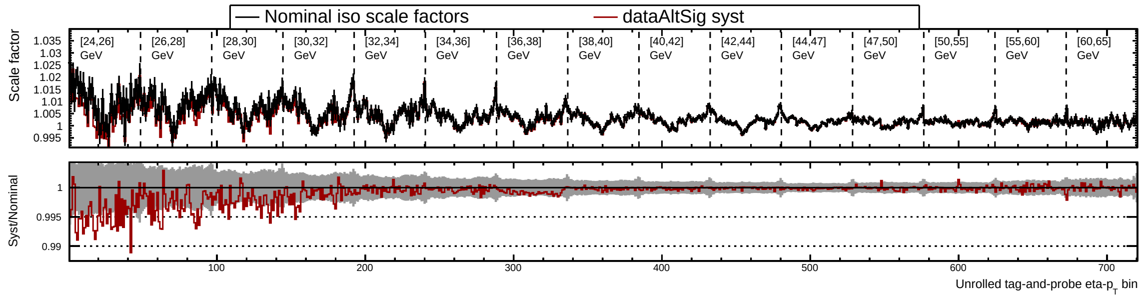


Figure 7.24: Isolation efficiencies on top of trigger in data (left) and MC (center), and Scale Factors (right) from tag-and-probe. The relative statistical uncertainties for each of the quantities are shown in the middle. The bottom plot shows the unrolled Scale Factors with their statistical (gray band) and systematic (red line) uncertainties.

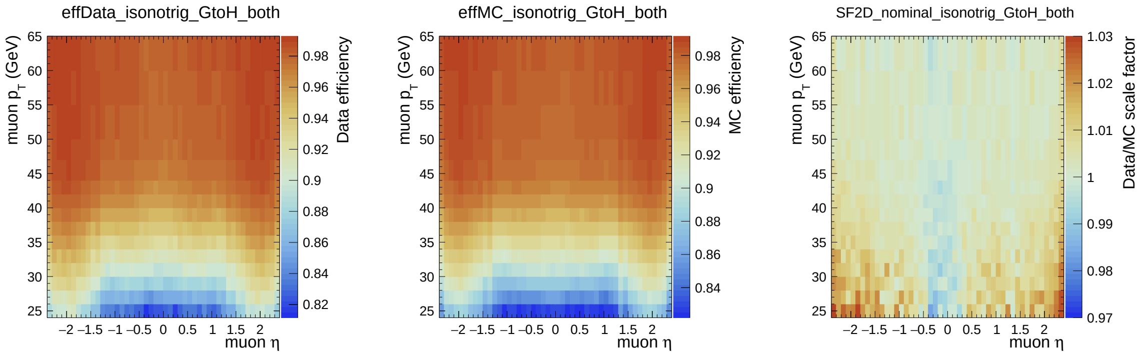
Trigger step (this is the one relevant for the nonprompt background estimate) ranges from 90% at low p_T to almost 100% at the high p_T: the fraction of events that fail this requirement is:

$$P(\text{fail Isolation}|\text{Trigger}) = 1 - P(\text{Isolation}|\text{Trigger}), \tag{7.3}$$

which then ranges from approximately 10% to being very close to 0, as the Isolation efficiency increases. Variations in $P(\text{Isolation}|\text{Trigger})$ translate into variations on $P(\text{fail Isolation}|\text{Trigger})$ which are same in absolute value but opposite sign:

$$\Delta P(\text{fail Isolation}|\text{Trigger}) = -\Delta P(\text{Isolation}|\text{Trigger}). \tag{7.4}$$

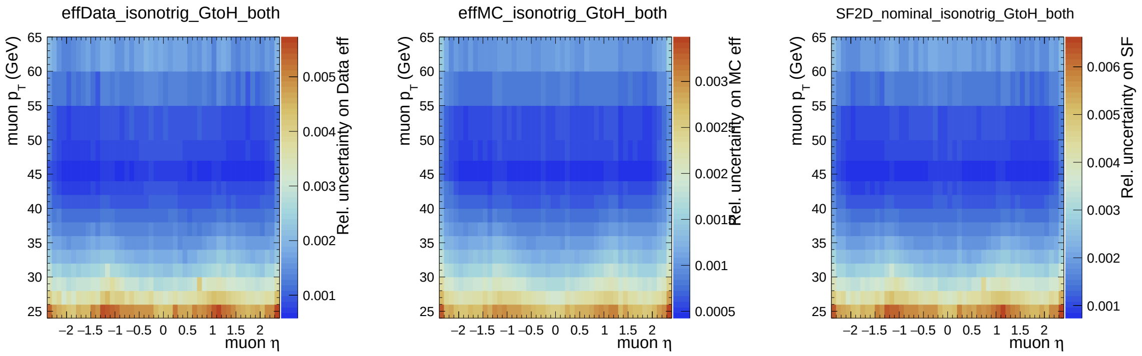
However, the relative magnitude of this variations is much larger for $P(\text{fail Isolation}|\text{Trigger})$ than for $P(\text{Isolation}|\text{Trigger})$: for example, if the Isolation efficiency in one $p_T - \eta - q$ (and u_T) bin is 0.91,



(a) Isolation without trigger eff. (data)

(b) Isolation without trigger eff. (MC)

(c) Isolation without trigger SF



(d) rel. stat. unc. in eff. (data)

(e) rel. stat. unc. in eff. (MC)

(f) rel. stat. unc. in SF

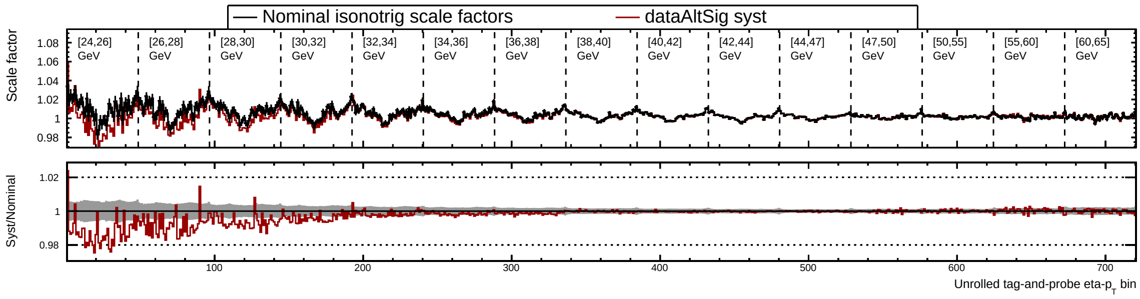


Figure 7.25: Isolation efficiencies on top of IDIP in data (left) and MC (center), and Scale Factors (right) from tag-and-probe. The relative statistical uncertainties for each of the quantities are shown in the middle. The bottom plot shows the unrolled Scale Factors with their statistical (gray band) and systematic (red line) uncertainties.

and has a statistical uncertainty of 0.01, corresponding to a relative statistical uncertainty of 0.011, the anti-Isolation efficiency is 0.09 still with a statistical uncertainty of 0.01, but corresponding to a relative statistical uncertainty of 0.11, so that any variation in the Isolation efficiency gets amplified, resulting in a large statistical uncertainty in the control region to extract the nonprompt background or unphysical discontinuities that might have an impact on the actual convergence of the fit to extract m_W.

The solution was to smoothen the Scale Factors instead of efficiencies: this requires simpler functions to properly catch the features of the efficiencies as opposed to smoothing directly the efficiencies themselves, as the Scale Factors are relatively flat, and less dependent on p_T than the efficiencies. This results in relatively small corrections applied on top of the efficiencies: this is particularly true for Trigger and Isolation efficiencies, which typically increase rapidly up to a threshold and then plateau at high p_T. In principle, a number of different families of curves could mimic the dependence on p_T of each of the effi-

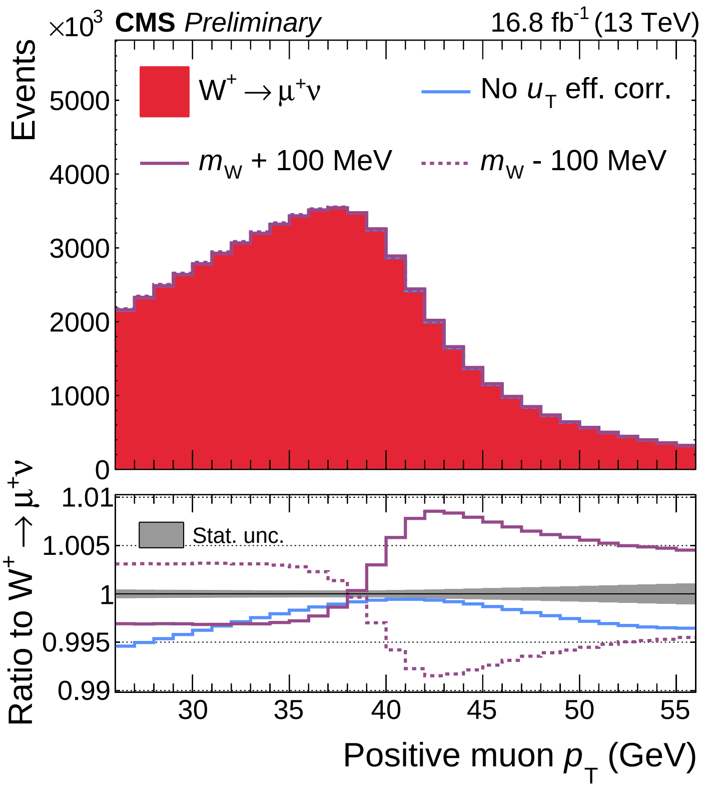


Figure 7.26: Spectrum of $W^+ \to \mu^+\nu$ decays with the full analysis selection, with the Scale Factors that take into account the dependence on u_T and those that are derived inclusively in u_T. In the ratio plot, a variation in m_W of ± 100 MeV is shown for comparison. The corrections are smoothed as described in 7.4.6.

ciency steps, but this could induce artificial variations on m_W depending on the family of curves chosen. Smoothing efficiencies and then taking their ratio might still induce non-cancellation between data and simulation due to the choice of functional form. This is less the case when the Scale Factors are smoothed directly.

The Scale Factors are smoothed as a function of p_T, and u_T for $\mathrm{Trigger}$ and $\mathrm{Isolation}$, independently for each η bin using polynomials of second or third degree depending on the step, as each step has a different binning and because the Scale Factors have a different dependence on p_T. The smoothing cannot be performed as a function of η because of the discontinuities of the CMS detector. The degree of the polynomials has been chosen to be the lowest possible that still enables to capture the p_T and u_T dependence, and it was verified that this smoothing procedure does not introduce biases in each of the η bins (that is, the fitted polynomials are in good agreement with the measured Scale Factors), but also across η bins, to avoid the possibility that the fitted curves still pass through the measured values, but induce trends that are similar in all η bins, resulting in artificial m_W shifts. Using polynomials of degree $N+1$, where N is the degree of the polynomial chosen for the smoothing, yielded similar results, so the lowest available was kept to avoid overfitting and keep the statistical uncertainties lower. The same polynomial degree is used for all the η bins in each step. The p_T range used in the analysis is 26 GeV$<p_T<56$ GeV, but the smoothing fits are run from 24 to 65 GeV to increase precision and reduce effects at the extrema of the fits from boundary conditions. The fits are performed using a TensorFlow [248] backend, and a variable transformation is performed such that the p_T range is mapped into the [0, 1] interval (the same is applied also to the u_T variable for the 2D smoothing): this should ensure good convergence of the smoothing fits and goodness of the parameter covariance matrix.

Table 7.6 summarizes the degrees of the polynomials used in the smoothing, for each step. As a function of p_T, a third-degree polynomial is used for all the steps except for $\mathrm{Tracking}$, for which a quadratic form is used, as only 4 bins in p_T are available. When $p_T - u_T$ is smoothed, the functional form is the product of a third-degree polynomial, to capture the dependence on p_T, and second-degree polynomial, to capture the dependence on u_T. Fitting the $\mathrm{Isolation}$ and $\mathrm{Trigger}$ Scale Factors simultaneously as a function of p_T and u_T allows to model the correlation between them. The same curves are also used for the smoothing of the Scale Factors obtained from the alternate fits, except for $\mathrm{Trigger}$ and $\mathrm{Isolation}$, where the u_T dependence on the alternate fits is neglected, and they are smoothed only as a function of p_T. More details about the implementation of the statistical and systematic uncertainties are provided in section 7.4.11.

Step	Number of charges	Degree in p_T	Degree in u_T
RECO	2	3	inclusive
$\mathrm{Tracking}$	2	2	inclusive
IDIP	2	3	inclusive
$\mathrm{Trigger}$	2	3	2
$\mathrm{Isolation}$	1	3	2

Table 7.6: Table summarizing the degrees of the polynomials used in the smoothing as a function of p_T and u_T.

Figure 7.27 shows two examples of smoothing, for the RECO and IDIP steps. Figures 7.28, 7.29 and 7.30 show the ratios between the binned and smoothed Scale Factors, evaluated at the bin center of the binned Scale Factors, for the RECO, $\mathrm{Tracking}$ and IDIP steps, respectively. The figures also show the pulls, defined as the difference between binned and smoothed Scale Factors divided by the statistical uncertainty of the binned Scale Factors. The pulls are shown for each $p_T - \eta$ bin independently and taken all together. No particular trends as function or p_T or η are observed, and the distribution of the pulls has a mean compatible with 0, which indicates that for these steps the smoothing is not introducing artificial biases, particularly in p_T where they could mimic a variation in m_W. The widths of the distributions are less than 1: this is likely due to the correlations between the binned and smoothed Scale Factors, as it is expected that the difference between them is smaller then the error used as input in the smoothing fits.

Figures 7.31 and 7.32 show examples of Scale Factors for the $\mathrm{Isolation}$ and $\mathrm{Trigger}$ respectively, for the η bins indicated in the figures. The figures also show the smoothed curves in the corresponding bins and the pulls.

Figures 7.33 and 7.34 show the means and the widths from Gaussian fits to the distributions of the pulls from the $p_T - u_T$ smoothing in each of the η bins for the different steps. The means are compatible with 0, indicating the smoothing function is correctly modelling the $p_T - u_T$ dependence of the Scale Factors. The width of the pulls is less than one in some of the η bins, which is likely due to the correlations between the binned Scale Factors and the smoothed functions derived from them.

When binning the $\mathrm{Trigger}$ and $\mathrm{Isolation}$ Scale Factors as a function of u_T in the tag-and-probe procedure, p_T^μ in 5.7 is the reconstructed muon p_T of the probe (using the inner tracker) and $\vec{p}_T^{\,\mathrm{boson}}$ is defined from the vectorial sum of the tag and the probe $\vec{p}_T$. A different definition is used for the processes in the analysis to evaluate the Scale Factors to apply to the templates, as $\vec{p}_T^{\,\mathrm{boson}}$ might not be properly defined or cannot be reconstructed directly (for example for the signal one of the muons is not present). For $W \to \mu\nu$ and $Z \to \mu\mu$, $\vec{p}_T^{\,\mathrm{boson}}$ is defined as the preFSR W boson p_T. A cross-check in which $\vec{p}_T^{\,\mathrm{boson}}$ is defined from the reconstructed muon and the neutrino at generator level showed that this induces modifications in the p_T spectrum that are $O(10^{-5})$[8]. For the other background processes, $\vec{p}_T^{\,\mathrm{boson}}$ is defined as the vector sum

---

[8] This could be explained from the fact that this definition differs from the one based on preFSR variables mostly because of resolution effects, which modify the overall u_T distribution. However, the corrections vary slowly enough that on average, for each u_T bin, the same Scale Factor is being applied. In addition to this, there are no requirements on u_T in the analysis, so that it is

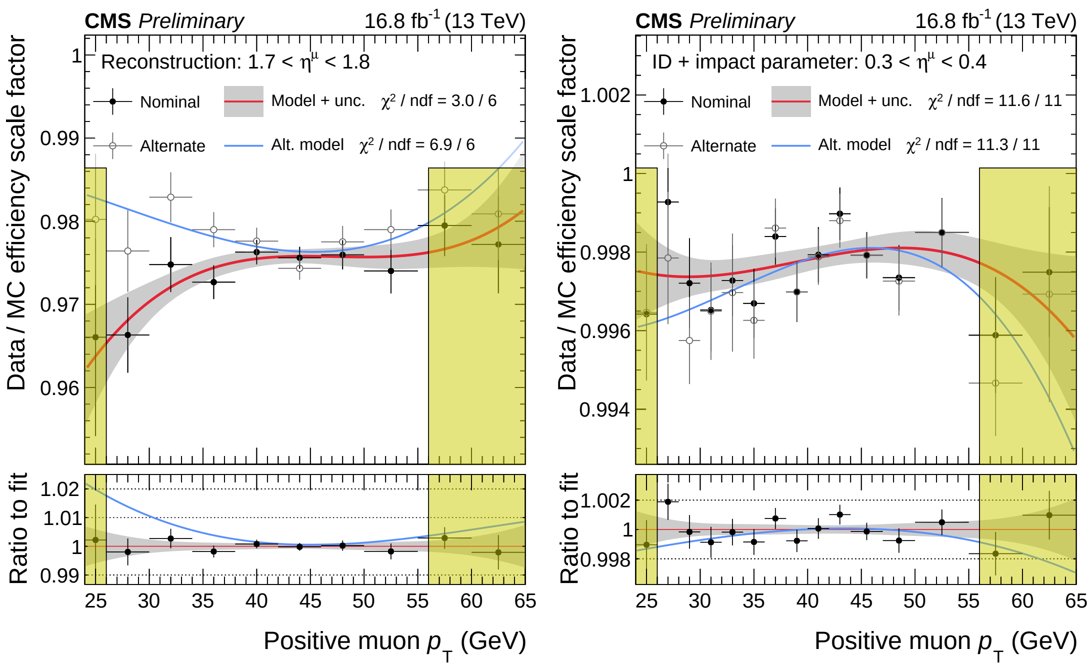


Figure 7.27: Examples of smoothing as a function of p_T, for the RECO (left) and IDIP (right) Scale Factors.

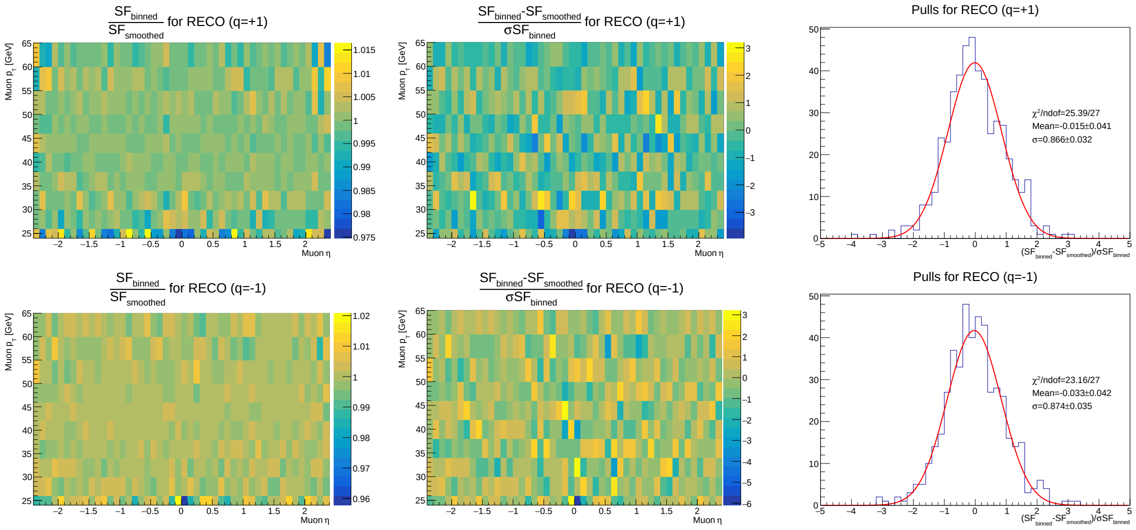


Figure 7.28: Comparison between binned and smoothed Scale Factors for the RECO efficiencies: the figures on the left shows their ratio, the figures in the middle and on the right show the pulls. The 1D distributions are provided for completeness, and fits to a Gaussian are superimposed, the mean and the width are shown (these are shown to provide a quantitative statement of the agreement between and smoothed SF, the standard deviations are smaller than one because of correlations).

of all generator level leptons and photons (paying attention not to double count muons from decays of tau leptons), to mimic the single V + jets topology. The distributions of u_T for the signal and the background processes with the highest number of events (apart from the nonprompt background), inclusively in m_T, but also separating between events that pass of fail the m_T requirement in the analysis, are shown in figure 7.35. The majority of events pass the m_T requirement, as can be inferred from the fact that the shapes for

---

integrated over all its range, regardless of what u_T definition is used.

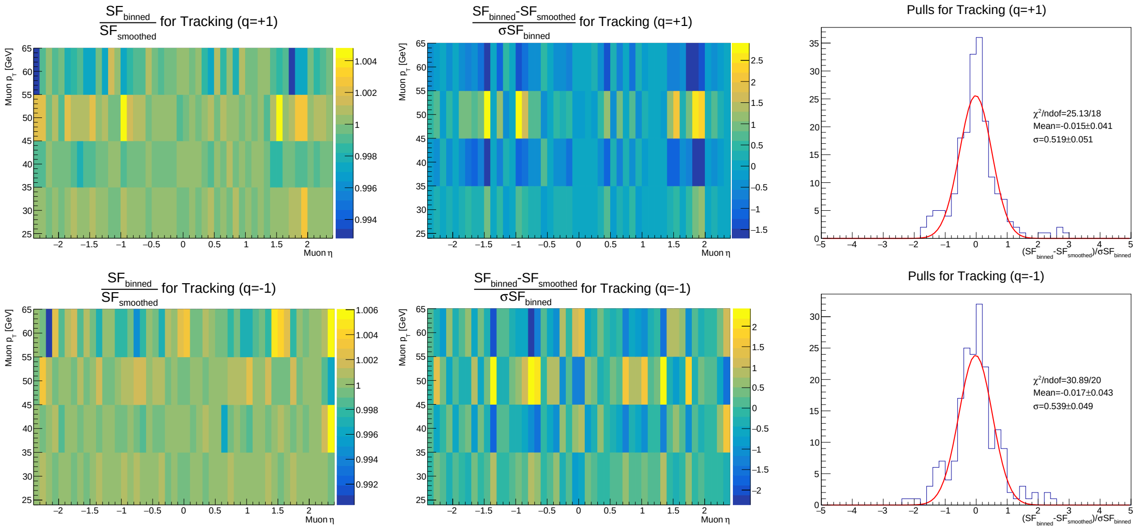


Figure 7.29: Comparison between binned and smoothed Scale Factors for the $\mathrm{Tracking}$ efficiencies: the figures on the left shows their ratio, the figures in the middle and on the right show the pulls. The 1D distributions are provided for completeness, and fits to a Gaussian are superimposed, the mean and the width are shown (these are shown to provide a quantitative statement of the agreement between and smoothed SF, the standard deviations are smaller than one because of correlations).

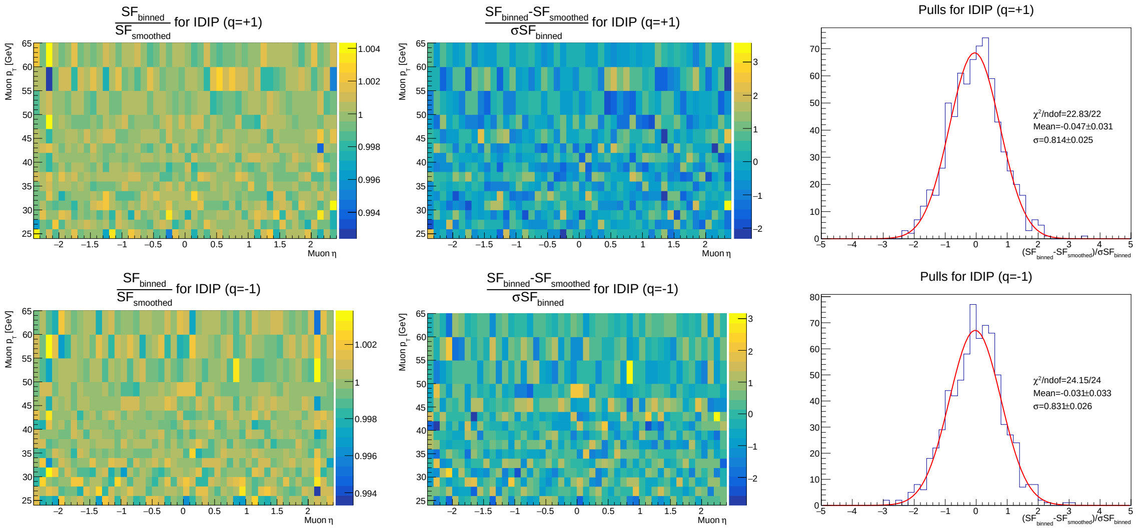


Figure 7.30: Comparison between binned and smoothed Scale Factors for the IDIP efficiencies: the figures on the left shows their ratio, the figures in the middle and on the right show the pulls. The 1D distributions are provided for completeness, and fits to a Gaussian are superimposed, the mean and the width are shown (these are shown to provide a quantitative statement of the agreement between and smoothed SF, the standard deviations are smaller than one because of correlations).

inclusive case and when the events pass the m_T are similar. However, it can be observed that those that fail m_T have significantly different u_T distributions shifted towards higher values. This could be partially due to kinematics, as these distributions are integrated in p_T and η, and reversing the m_T threshold has an effect on the p_T spectrum, for both signal and background, but also to a genuine correlation between the two. For the events that pass the m_T requirement, it can be observed that the u_T distributions for the backgrounds are shifted to higher values compared to those for W and Z events.

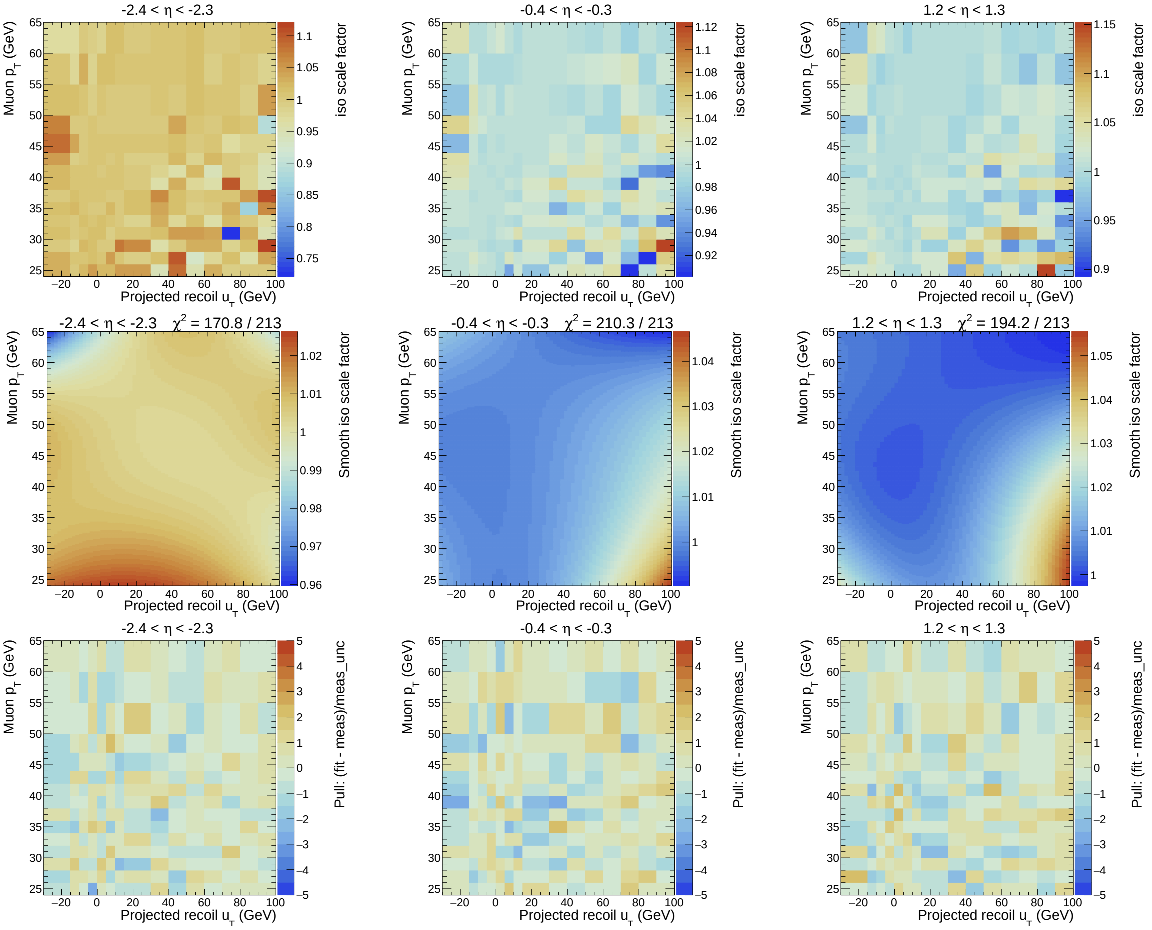


Figure 7.31: Example of $\mathrm{Isolation}$ Scale Factors as a function of $u_T - p_T$ in selected η bins. The plots in the top show the SF measurements, the plots in the middle show the Scale Factors after smoothing. The plots in the bottom show the pulls, defined as the difference between measured and smoothed SF divided by the measured SF uncertainty.

Treatment of anti-$\mathrm{Isolation}$

After smoothing the $\mathrm{Isolation}$ efficiency, the corrections for the muons that fail the isolation requirement, but are still matched to a trigger object, needed for the estimation of the nonprompt background, get a correction, as a function of $p_T - \eta - u_T$, which is:

$$\mathrm{SF}_{\text{anti-Isolation}} = \frac{1 - \mathrm{SF}_{\mathrm{Isolation}}\,\epsilon^{\mathrm{MC}}_{\mathrm{Isolation}}}{1 - \epsilon^{\mathrm{MC}}_{\mathrm{Isolation}}} \tag{7.5}$$

where $\epsilon^{\mathrm{MC}}_{\mathrm{Isolation}}$ is the $\mathrm{Isolation}$ efficiency for the Monte Carlo events and $\mathrm{SF}_{\mathrm{Isolation}}$ is the $\mathrm{Isolation}$ Scale Factor derived as described previously. In this way, the corrections derived for the events that pass the isolation requirement are propagated correctly to those that fail it, while ensuring the correct normalization for the sum of both sets of events. This also enables us to properly propagate the statistical and systematic uncertainties to the control region to estimate the nonprompt background. $\epsilon^{\mathrm{MC}}_{\mathrm{Isolation}}$ is the isolation efficiency for $W \to \mu\nu$: in principle, it should not matter what is the process for which the efficiency is computed, but, as discussed in 5.2.1, there are still some residual differences between tag-and-probe events and the processes in the analysis (both signal, W decays, and background) when reweighting the efficiencies for the different u_T distributions. However, in the region obtained by reversing the $\mathrm{Isolation}$ requirement, mostly populated by nonprompt muons, the dominant prompt contribution is from W boson

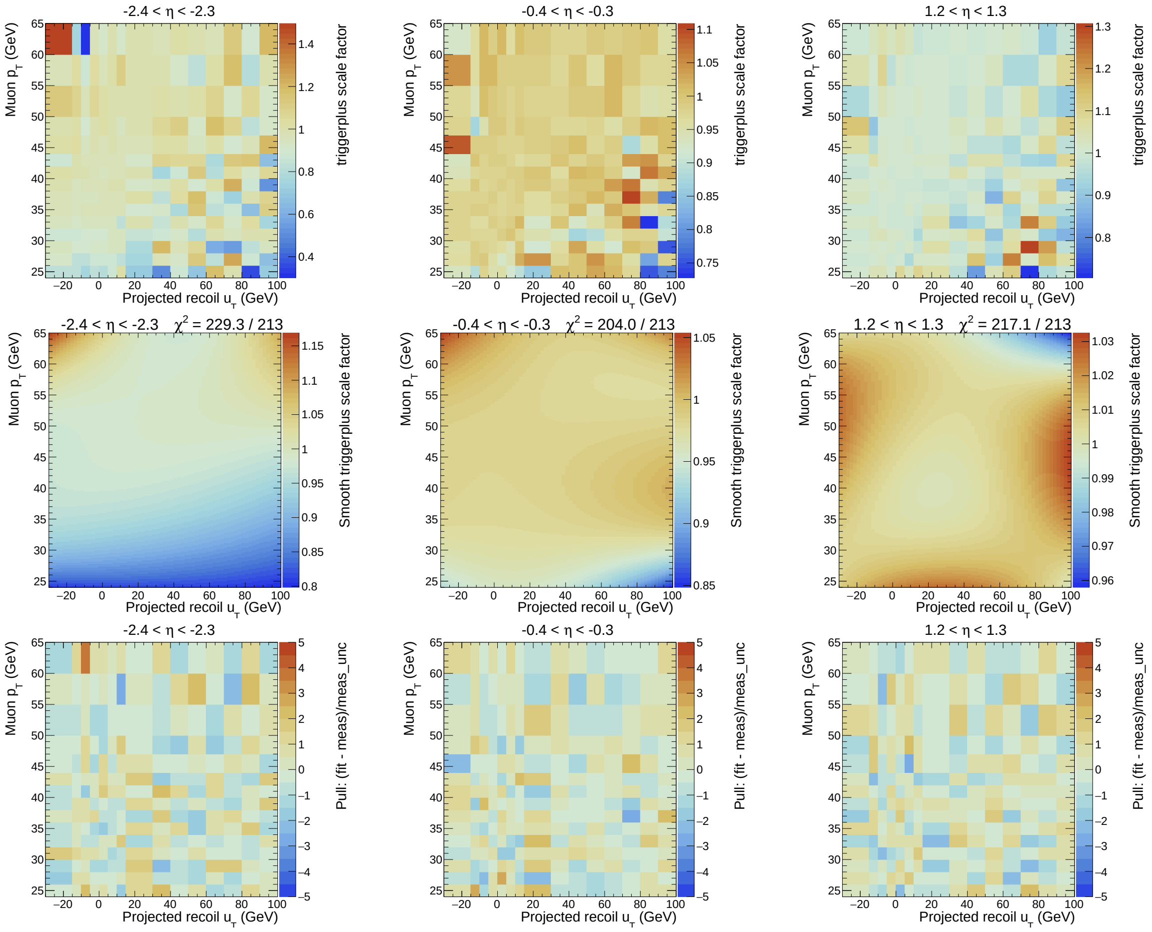
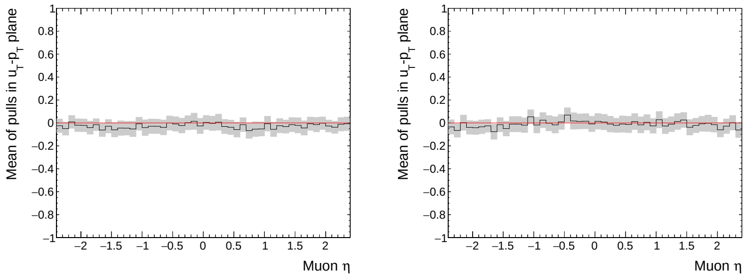


Figure 7.32: Example of $\mathrm{Trigger}$ $p_T - u_T$ Scale Factors for the η bins indicated in the figures, for positive muons. The plots in the top show the SF measurements, the plots in the middle show the Scale Factors after smoothing. The plots in the bottom show the pulls, defined as the difference between measured and smoothed SF divided by the measured SF uncertainty.

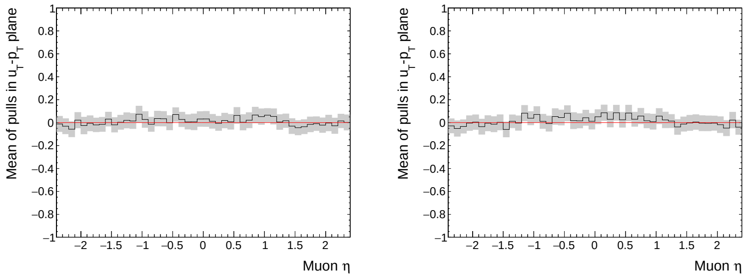
events, with only a negligible contribution from the other processes, so that the efficiencies for the W are used for all the processes when applying the anti-$\mathrm{Isolation}$ Scale Factors, also because they are the ones computed with the highest amount of statistics.

They are computed in fine bins of $u_T - p_T$ for each η bin and then smoothed with spline functions to interpolate between bin centers. The contribution to the anti-$\mathrm{Isolation}$ Scale Factors due to the finite statistics used to derive the $\mathrm{Isolation}$ efficiencies is assumed to be negligible, as the statistics in the W sample is very large. $\epsilon^{\mathrm{MC}}_{\mathrm{Isolation}}$ is computed without the m_T requirement, so that they can also used in the other control regions used in the nonprompt background estimation (see 7.7).

For completeness, figure 7.36 shows the $\mathrm{Trigger}$ MC truth efficiencies for signal events for positive muons for selected η bins. The features visible here are similar for other steps. The efficiencies are shown for -100 GeV$<\ u_T\ <\ 100$ GeV, but effectively only the values from $u_T\ >\ -30$ GeV are used, as Scale Factors are available only in this region. They are assumed to be constant outside of this range. This assumption should induce a negligible bias on the $p_T - \eta$ templates, as the majority of the events in the analysis are in this u_T range, as can be seen, for example, in figures 5.23 and 7.40.

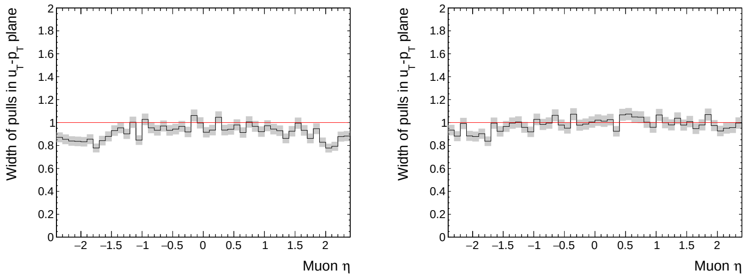


(a) SF $P(\text{Isolation}|\text{Trigger})$ (b) SF $P(\text{Isolation}|\text{IDIP})$

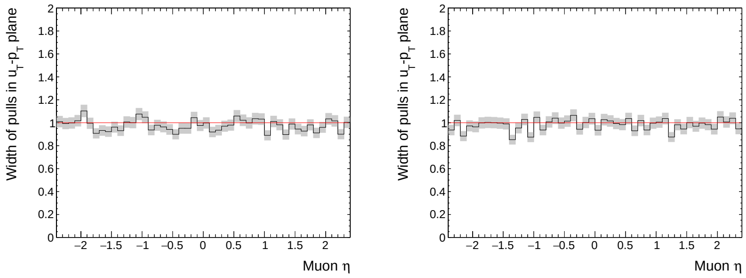


(c) SF $P(\text{Trigger}|\text{IDIP})$ for μ^+ (d) SF $P(\text{Trigger}|\text{IDIP})$ for μ^-

Figure 7.33: Mean values of the distributions of pulls from the 2D smoothing fits to SF in the $p_T - u_T$ plane as a function of η, for all the relevant SF steps.



(a) SF $P(\text{Isolation}|\text{Trigger})$ (b) SF $P(\text{Isolation}|\text{IDIP})$



(c) SF $P(\text{Trigger}|\text{IDIP})$ for μ^+ (d) SF $P(\text{Trigger}|\text{IDIP})$ for μ^-

Figure 7.34: Width values of the distributions of pulls from the 2D smoothing fits to SF in the $p_T - u_T$ plane as a function of η, for all the relevant SF steps.

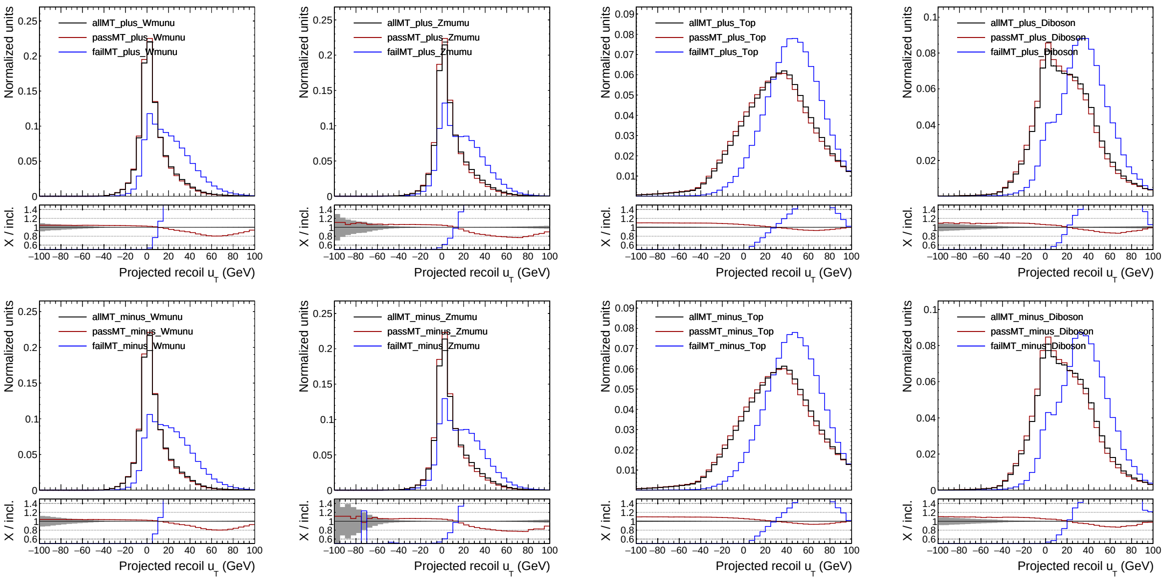


Figure 7.35: u_T distribution for signal and some background processes after the full analysis selection except the m_T selection for positive (top) and negative (bottom) muon charge, for events passing (red) or failing (blue) the m_T cut, or inclusive in m_T (black).

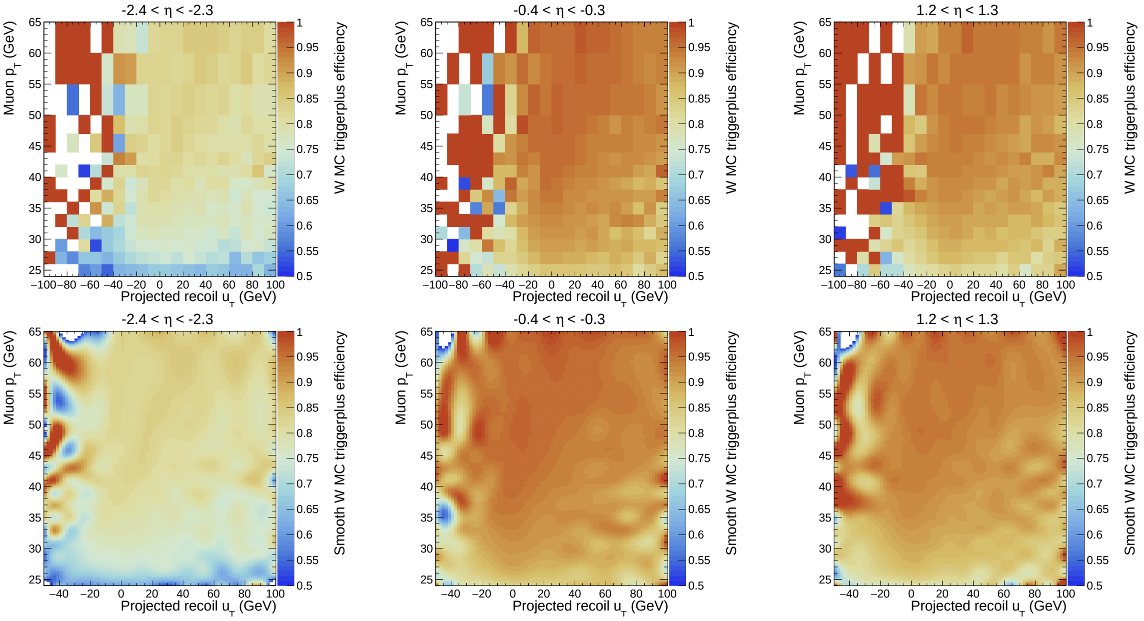


Figure 7.36: Example of MC truth $\mathrm{Trigger}$ efficiencies for signal decays as a function of $u_T - p_T$, for positive muon charge in few selected η bins. Top plots show them as derived directly from Monte Carlo in the $u_T - p_T$ plane, while bottom ones are the efficiencies as a result of the spline interpolation. Note the different u_T ranges for top and bottom plots. The effective range used in the analysis is the one defined by the Scale Factors, for $u_T > -30$ GeV.

7.4.7 Trigger prefiring

The synchronization of the CMS detector and the LHC clocks is used to implement the division of proton-proton collisions into bunch crossings (BX), so that appropriate timing windows to reconstruct the events can be defined, with the windows corresponding to collisions between different pairs of bunches. It can occur, however, that objects are assigned to the incorrect bunch crossings, because of the limited timing resolution of the detectors that reconstruct them, or to non-trivial effects leading to a de-synchronization between the CMS and the LHC clock. When this happens to L1 trigger objects, the wrong BX is accepted by the L1 trigger system: when the L1 trigger object is assigned to the previous bunch crossing (BX=-1) instead of the one in which it was actually produced (BX=0), this leads to "trigger prefiring". If an event is accepted by the trigger system, the CMS trigger rules [249, 250] prevent the following two BXs from generating a valid trigger accept signal, resulting in the correct event being lost. This is done to suppress trigger requests coming too soon after one or several triggers, as in this case subsystems may not be ready to accept additional triggers, thus potentially resulting in discarded events. This also occurs in the case of prefiring, with the difference that the BX that produced the L1 trigger object is one of the two BXs being vetoed, with the result that the event in this BX is discarded.

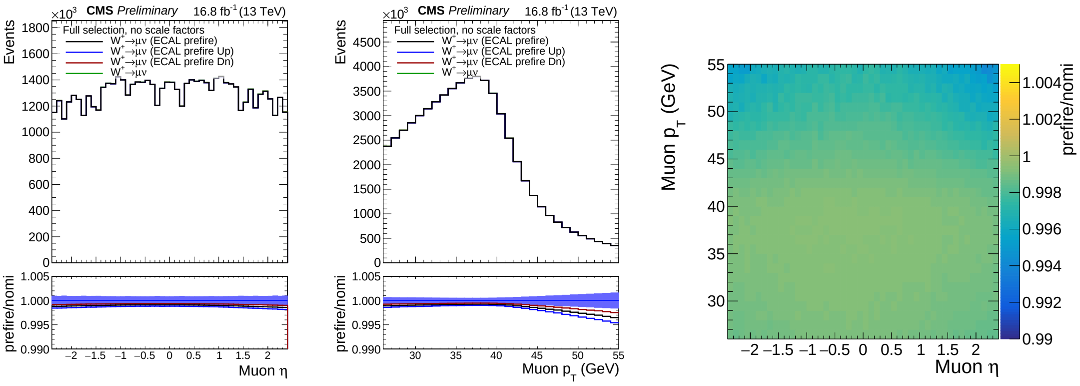
The standard tag-and-probe technique cannot be used to correct for this effect, because events are totally missing in the selected sample. Therefore, the estimation of the prefiring probability is done using a set of rare "unprefirable" events: these are events in BX=0 for which the event in BX=-3 is accepted by the Level-1 trigger, so that the events in BX=-2 and BX=-1 are vetoed. For each triggered event, all Level-1 objects decision bits are stored in a window of $\pm$2BX. The prefiring probability is then defined as the fraction of events in BX=0 that have a L1 trigger object in BX=-1. These events are discarded (as they normally would) because of the veto in BX=-2 and BX=-1 imposed by the acceptance of the event in BX=-3. This is done differentially in the phase space, and corrections are applied to the templates as event weights. Since "unprefirable" events are rare, the binnings are usually not too fine and time dependence within 2016 postVFP is not studied, the corrections are assumed to be constant.

ECAL prefiring

ECAL prefiring has been noticed since the beginning of Run 2, in the form of an increasing shift in the shapes of electronic pulses, growing with the amount of irradiation of the ECAL crystals. The timing delays in the readout electronics where corrected in 2018 to take this into account, but the 2016 postVFP data used in the m_W measurement is still affected by this. It induces a loss of events with high-energy photons or electrons, particularly for $p_T > 30\,\text{GeV}$ and $2.0 < |\eta| < 3.0$, but also for jets with a large fraction of electromagnetic energy. Maps to correct for the prefiring rate are provided by the CMS JetMET group as a function of p_T and η of electrons, photons, or jets. These maps are used to define a global per-event prefiring probability taking into account all physics objects that can lead to prefiring, which is then used to reweight the events that enter the $p_T - \eta$ templates to extract m_W. Figure 7.37 shows the effect in η and p_T for the $W \to \mu\nu$ that pass the full analysis selection. The inefficiency induced by the ECAL prefiring is mostly relevant at above the Jacobian peak, and it increases up to 0.5% as a function of p_T above the Jacobian peak, mainly in the high η region, because of the slight correlation between the muon variables and those of possible jets in the event, as well as the possible emission of FSR photons. The uncertainties in the prefiring weights are also shown: they modify the distributions mostly at high p_T by at most 0.1%.

Muon prefiring

Muon prefiring is due to the limited timing resolution of the CMS L1 muon trigger. This was limited by the muon track seeding algorithm [250], and changes were implemented throughout 2016 to the *muon track finder* (MTF) algorithm to reduce the prefiring rate. These have been implemented only starting from the H period of 2016 postVFP, the latest period of 2016 data-taking, and only in the endcap. In the late part of the H period, the changes were also extended to the rest of the detector acceptance, reducing the prefiring rate to $O(0.1\%)$. The prefiring probability is measured in bins of muon $|\eta|$ inclusively in p_T due to the limited statistics available. The p_T dependence of the prefiring probability parameterized using an error function with threshold at $22\,\text{GeV}$ with the plateau set to the prefiring probability measured above. This threshold



(a) (b) (c)

Figure 7.37: Effect of ECAL prefiring as a function of the muon η and p_T in simulated $W \to \mu\nu$ events, after the signal region selection. The left and middle plots also show the associated prefiring uncertainty band. The right plot is the two-dimensional ratio to show the application of prefiring weights. The statistical uncertainty is shown as a blue band.

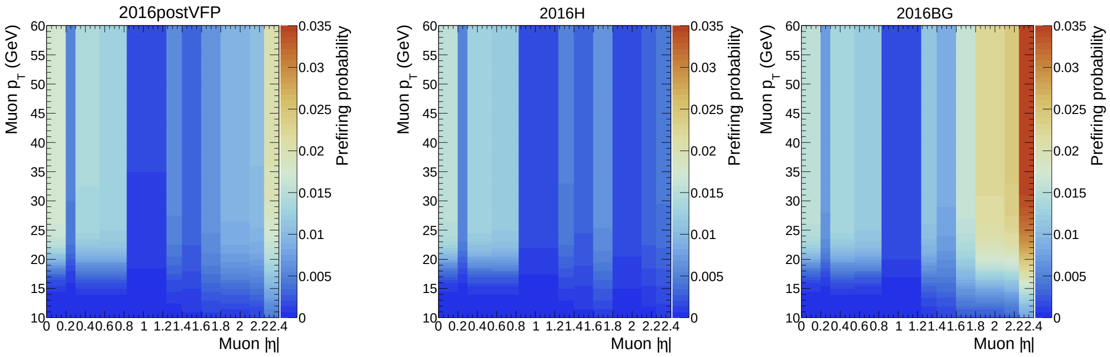
was chosen because it corresponds to the lowest unprescaled L1 bit in the trigger menu of 2016, used in the analysis. Figure 7.38 shows the prefiring probabilities as a function of $|\eta|$ and p_T for the whole of 2016 postVFP and for the H period, with figure 7.38d showing the prefiring probability for as a function of $|\eta|$ for $p_T > 25$ GeV, where the dependence on p_T is negligible. It can be observed that the prefiring probability is at the percent level, higher above $|\eta| > 1.2$ before the application of MTF update, while in the barrel the situation is very similar (2016 preVFP is also shown for completeness). The prefiring probability is negligible between $0.8 < |\eta| < 1.2$ as this is the transition region between the drift tubes and RPCs. Figure 7.38c shows the hotspot region with $1.24 < \eta < 1.60$ and $2.4 < \phi < 2.8$ which has a much larger prefiring probability, not dependent on the era: a dedicated uncertainty is assigned to cover for this, as described in 7.4.11.

The prefiring, applied as an event weight as a function of the muon $p_T - \eta$, is $1 - \epsilon_{\text{Prefiring}}$, where $\epsilon_{\text{Prefiring}}$ is that probability that at least one muon, only those that pass the loose ID are considered, induces the prefiring. The impact from muon prefiring is different for the analysis to extract m_W and the W-like and dilepton analyses, because two muons are required for the latter, and this has an impact on the Z boson rapidity distribution in the forward region, as the $|\eta|$ of the muons are correlated.

7.4.8 Validation of corrections

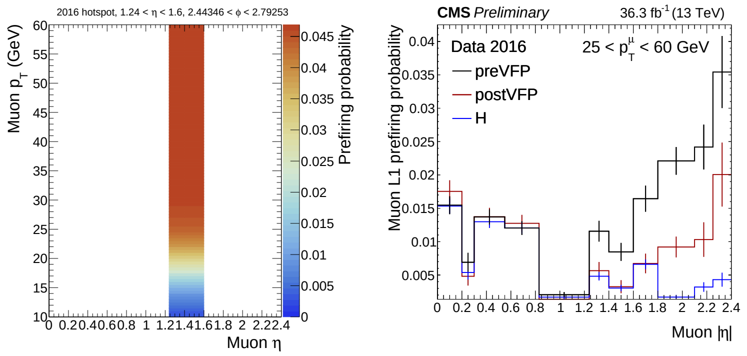
Figure 7.39 shows the η distribution for W-like events for data and Monte Carlo simulation, where the simulation is normalized to data. The bottom panel shows the ratio and the all the statistical and systematic uncertainties of the analysis as a gray band. The data/MC ratio appears to mostly flat around 1, except for the region around $\eta = 0.2$. After the publication of the analysis, it was found that this dip is related to the probability that the CVH refit, introduced in subsection 7.5, is successful for global muons, as these plots are shown using the reconstructed muon η after the full muon scale calibration and resolution, which is not available when the muon reconstruction and isolation Scale Factors and prefiring corrections are derived. The average tracker resolution is $O(1\%)$, slightly worse at high η, and the effects of muon scale calibration should be negligible the estimation of the efficiencies, especially because the p_T width is always larger than 2 GeV. It was checked that, applying the muon reconstruction and identification Scale Factors as a function of the calibrated or the uncalibrated muon p_T makes a negligible difference in the templates.

Figure 7.40 shows the results of a study conducted to validate that the $\mathrm{Trigger}$ and $\mathrm{Isolation}$ (on top of $\mathrm{Trigger}$) Scale Factors are correctly derived and capture the dependence on the boson recoil: W-like



(a) 2016 postVFP

(b) 2016 H



(c) Hotspot

(d) Prefiring versus muon $|\eta|$

Figure 7.38: Muon prefiring probabilities measured from data for different periods in 2016, as a function of muon $p_T - \eta$. The last plot in figure 7.38d shows the probability versus muon $|\eta|$ for p_T above the turn on, with error bars representing statistical uncertainties.

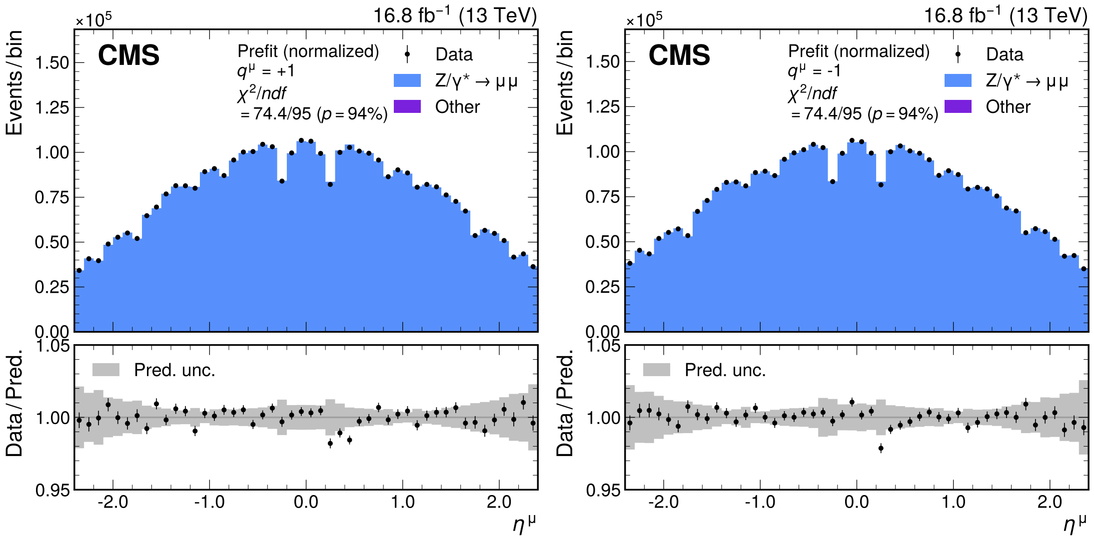


Figure 7.39: Measured and predicted η^{μ} distributions in $Z \to \mu\mu$ events with the W-like Z boson selection for positively (left) and negatively (right) charged muons. The total uncertainties (statistical and systematic) are represented by the gray bands and the normalization of the simulated spectrum is scaled to the measured distribution.

events are considered, but the triggering muon p_T is restricted between 35 and 55 GeV, while the rest of the selection is kept the same. The non-triggering muon is considered, and the $\mathrm{Trigger}$ and $\mathrm{Isolation}$ efficiencies are computed as a function of u_T, and compared against data with and without corrections. The triggering muon p_T range was restricted to perform this comparison in a phase space that is different from the one used to derive the Scale Factors, to avoid that data and Monte Carlo agree by construction. Here, u_T is defined from both the reconstructed muons, in data and Monte Carlo simulation. Background processes, estimated from Monte Carlo simulation, are subtracted from data in each u_T bin: they are found to make a difference mostly in the high u_T region on the right, with distributions that are qualitatively similar to what is shown in figure 7.35. The efficiencies are then defined as the fraction of events that pass the $\mathrm{Trigger}$ or $\mathrm{Isolation}$, depending on the step. It can be seen that the agreement is good in most of the u_T range, when the Scale Factors are applied. This gives us confidence that these corrections are derived correctly and that they capture the dependence on the production mechanism.

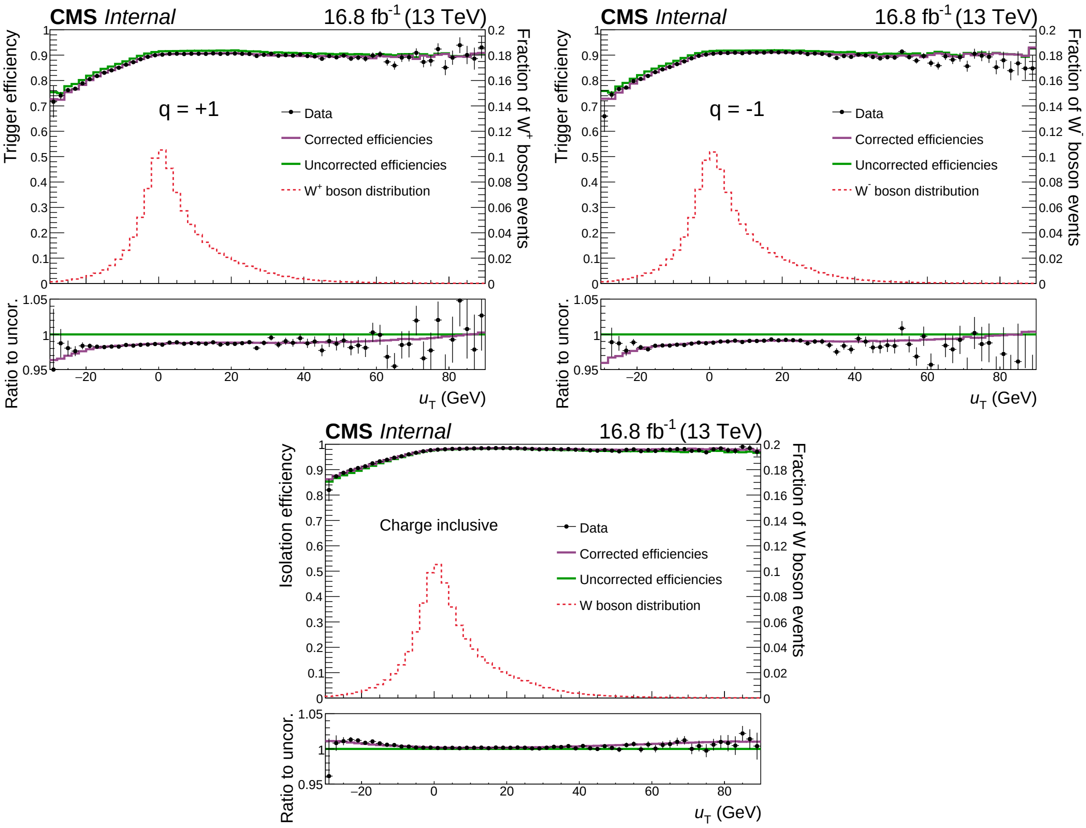


Figure 7.40: $\mathrm{Trigger}$ and $\mathrm{Isolation}$ efficiencies in data and Monte Carlo, with and without the efficiency corrections.

Another test made to validate the efficiency corrections has been to compare the u_T distribution for data and simulation on W-like events, eventually with and without efficiency corrections, but these have large uncertainties arising from the modelling of the Z boson production mechanism. They are much larger than the magnitude of the efficiency corrections, as shown on the left of figure 7.41, where the gray band is the uncertainty band from the model used in the fit to extract m_Z. In the center and on the right of this figure, the same distributions are shown after setting the Monte Carlo predictions to their post-fit values, with the post-fit uncertainties properly propagated to the u_T distribution. In the middle, the results are shown when the Monte Carlo predictions are set to the post-fit values determined from the W-like fit, while on the right the same results are shown when the Monte Carlo predictions are set to the post-fit values determined from the fit to the $p_T^{ll} - y_{ll}$ double differential distribution.

The post-fit distributions in Monte Carlo simulation are in agreement with those in data, with high $p-$value in both cases. The uncertainty bands that are shown in the plots and the $p-$values shown in the figures correctly take into account the covariance matrix of the fits. It can be observed that the $p_T^{ll} - y_{ll}$ fit is able to constrain the high u_T region significantly better than the $p_T^{\mu} - \eta^{\mu} - q^{\mu}$ fit, as p_T^{ll} is fit up to 100 GeV, but the $p_T^{\mu} - \eta^{\mu} - q^{\mu}$ fit indicates that we are able to reproduce the u_T distribution correctly.

In both sets of results, the nuisance parameters related to the efficiencies and to the production mechanism might be pulled so as to obtain the best possible agreement, even though not directly in u_T (as it is not used in either fit). This might imply that differences between data and Monte Carlo simulation due to the Scale Factors might be reabsorbed by other nuisance parameters. In spite of this, the results indicate that the u_T distributions used to apply the efficiency corrections match between data and Monte Carlo, at least after the fit. Therefore, no discrepancies in the templates from integrating over and incorrect u_T distribution should arise. In summary, all the ingredients that enter into the u_T distribution, such as efficiencies and the production mechanism, are under control.

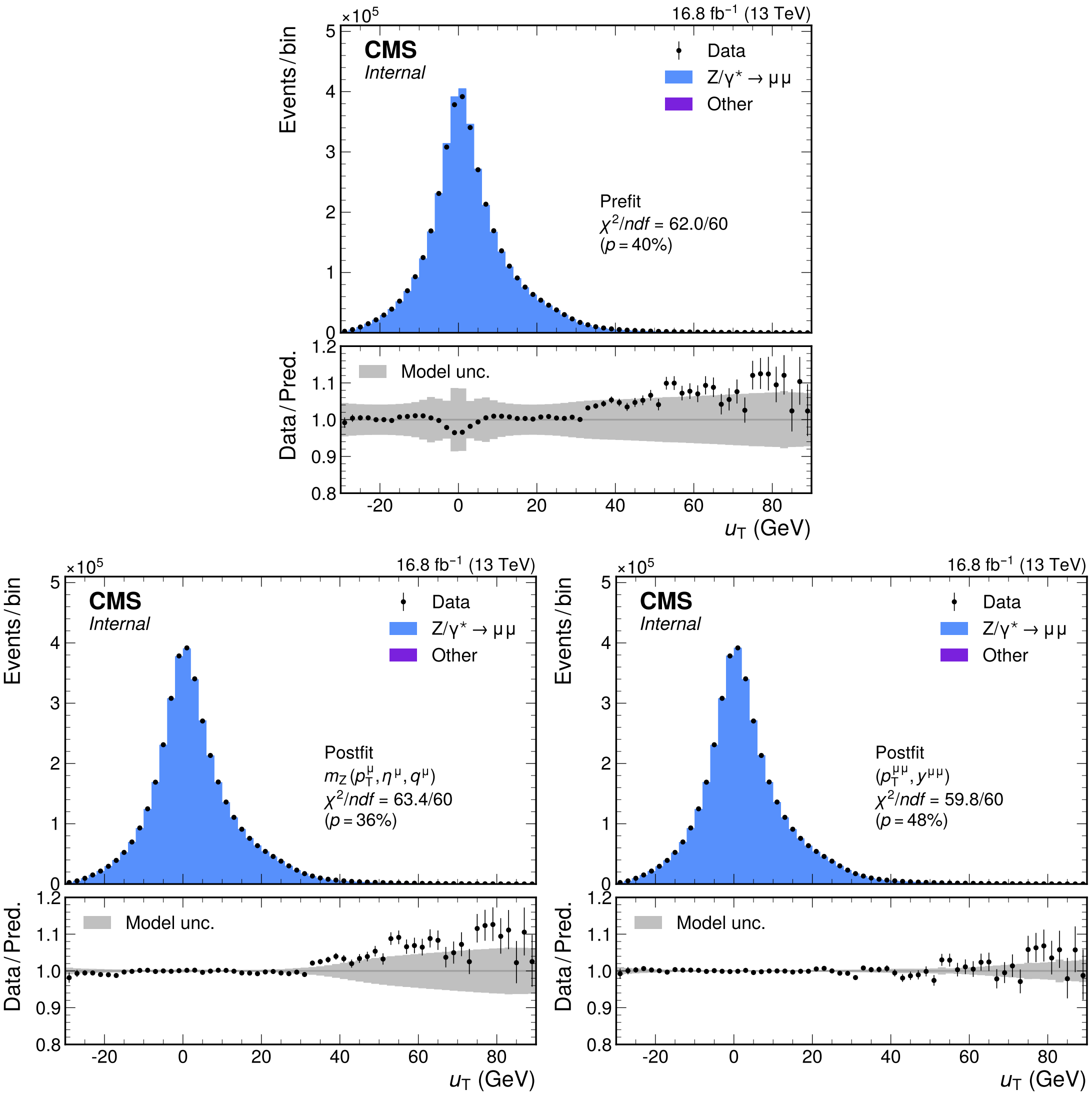


Figure 7.41: Top: pre-fit distributions of u_T for data and Monte Carlo, with ratio between them and all the uncertainties from the fit. Post-fit distributions where the model is tuned to the results of the $p_T^{\mu} - \eta^{\mu} - q^{\mu}$ fit (left) and to the results of the $p_T^{ll} - y_{ll}$ (right).

7.4.9 Second muon veto

One of the largest backgrounds in the analysis is represented by $Z \to \mu\mu$, which has a shape in p_T that is qualitatively similar to that of $W \to \mu\nu$ events, if all events with at least one muon are selected.

To suppress this background, events that have more than one reconstructed muon are excluded from the analysis. The muons considered for this purpose pass only a loose selection, with the idea of rejecting as many background events as possible, so the probability to reconstruct these "veto" muons is very close to 1.

The events that are then selected in the analysis are those that have only one reconstructed muon, either because only one muon has been genuinely produced or because only one muon was reconstructed but more were produced. For the latter case, it is very important to derive efficiency corrections carefully, as small changes in the probability to reconstruct the veto muons can result in large changes in the background yields in the analysis, as only the events for which the veto reconstruction fails are considered. This is the same argument as for $\mathrm{Isolation}$ and anti-$\mathrm{Isolation}$ as discussed in 7.4.6, also in equations 7.4 and 7.5, which lead to the implementation of smoothed efficiency corrections, with the difference that, unlike for anti-$\mathrm{Isolation}$, where nonprompt background estimation was affected by electroweak contamination in control regions, here the yields of the $Z \to \mu\mu$ background change directly in the signal region used to extract m_W.

Two veto definitions were considered for this analysis:

- global muons with $|d_{\mathrm{xyBS}}| < 0.05$ cm and that pass the loose ID, with inner and outer $p_T > 15$ GeV and $\Delta R_{\text{inner track, outer track}} < 0.3$, where the inner track passes the high-purity selection. This definition has the advantage that efficiency corrections are relatively easy to derive, as they are mostly similar to those derived for the muons in the $p_T - \eta$ plane of the analysis, but it is also less effective at rejecting background events, especially if the muons have $p_T \approx 15$ GeV;
- the logical `OR` of global and tracker muons, where the global muons pass the same selection as in the other veto definition, the tracker muons must have $p_T > 15$ GeV, must be a tracker-seeded track, must pass the high-purity selection and the loose ID and must have $|d_{\mathrm{xyBS}}| < 0.05$ cm. This definition has a higher background rejection, but the efficiency corrections are less well defined, and some differences between the derivation and the application remain.

The efficiency corrections are again derived with the tag-and-probe method, in different steps. For the first veto definition, the steps are RECO, $\mathrm{Tracking}$ and a modified version of IDIP with the loose ID instead of the medium ID, and the derivation is performed as before. The other difference is that in this case, extra bins are added, for RECO and IDIP two bins from 15 to 20 GeV and from 20 to 24 GeV are added, while for $\mathrm{Tracking}$ one extra bin from 15 to 24 GeV is added.

For the second veto definition, IDIP is defined as:

$$\mathrm{IDIP}_{\text{second veto}} = P(\text{loose ID}, |d_{\mathrm{xyBS}}| < 0.05\ \text{cm} | \text{global OR tracker}), \tag{7.6}$$

while two other steps are defined as follows:

- The first one, which we refer to as $\mathrm{T\text{-}Tracking}$ (i.e. $\mathrm{Tracking}$ for tracker-seeded tracks), is the probability that a standalone is geometrically matched to a tracker-seeded track within $\Delta R < 0.3$.
- The second one, which we refer to as $\mathrm{T\text{-}RECO}$ (i.e. RECO for tracker muons), is the probability that the tracker-seeded track becomes a tracker muon but not a global muon as in the other veto definition.

The product of these two steps is summed to the product of RECO and $\mathrm{Tracking}$ from the first veto definition, and then multiplied by the IDIP probability for this definition. The steps are defined so that double counting of events between global and tracker muons is avoided, and so that one can rely on the corrections derived for the first veto definition, which are more robustly defined: for the steps specific to the tracker muon, the geometric matching between inner and outer track that defined the probability to find an inner track cannot be applied in the veto, as the efficiency corrections for the first steps are derived as a function of the standalone muon p_T, not available for tracker muons. Both veto definitions require muons with $p_T > 15$ GeV: this is because it becomes extremely difficult to run tag-and-probe invariant mass fits below this threshold. This is the same reason why in $\mathrm{Tracking}$ the Scale Factors were only derived down to 24 GeV, as running fits in the bin from 15 to 24 GeV proved to be difficult already. However, those corrections apply to the muon used to fill the $p_T-\eta$ templates, which must have inner-track $p_T > 26$ GeV, so that the bin from 15 to 24 GeV is less important and extrapolating the Scale Factors from higher p_T might be sufficient. This is not the case for the veto, as we need to reject events with a second muon of low p_T. All the steps, in both veto definitions, are computed separately for positive and negative muons.

Smooth anti-veto[9] efficiency corrections are applied, to have smaller statistical uncertainties and to avoid some of the issue related to binned Scale Factors already discussed. Two smoothing strategies were adopted. In the first one, the Scale Factors from the individual steps are smoothed independently for each step, for data and Monte Carlo, and smooth Scale Factors are then derived by computing:

$$\mathrm{SF}^{\mathrm{first}}_{\mathrm{anti\text{-}veto}} = \frac{1 - \mathrm{SF}_{\mathrm{RECO}}\epsilon^{\mathrm{MC}}_{\mathrm{RECO}} \times \mathrm{SF}_{\mathrm{Tracking}}\epsilon^{\mathrm{MC}}_{\mathrm{Tracking}} \times \mathrm{SF}_{\mathrm{IDIP}}\epsilon^{\mathrm{MC}}_{\mathrm{IDIP}}}{1 - \epsilon^{\mathrm{MC}}_{\mathrm{RECO}} \times \epsilon^{\mathrm{MC}}_{\mathrm{Tracking}} \times \epsilon^{\mathrm{MC}}_{\mathrm{IDIP}}} \tag{7.7}$$

for the first veto definition, and:

$$\mathrm{SF}^{\mathrm{second}}_{\mathrm{anti\text{-}veto}} = \frac{1 - (\mathrm{SF}_{\mathrm{RECO}}\epsilon^{\mathrm{MC}}_{\mathrm{RECO}} \times \mathrm{SF}_{\mathrm{Tracking}}\epsilon^{\mathrm{MC}}_{\mathrm{Tracking}} + \mathrm{SF}_{\mathrm{T\text{-}RECO}}\epsilon^{\mathrm{MC}}_{\mathrm{T\text{-}RECO}} \times \mathrm{SF}_{\mathrm{T\text{-}Tracking}}\epsilon^{\mathrm{MC}}_{\mathrm{T\text{-}Tracking}}) \times \mathrm{SF}_{\mathrm{IDIP}}\epsilon^{\mathrm{MC}}_{\mathrm{IDIP}}}{1 - (\epsilon^{\mathrm{MC}}_{\mathrm{RECO}} \times \epsilon^{\mathrm{MC}}_{\mathrm{Tracking}} + \epsilon^{\mathrm{MC}}_{\mathrm{T\text{-}RECO}} \times \epsilon^{\mathrm{MC}}_{\mathrm{T\text{-}Tracking}}) \times \epsilon^{\mathrm{MC}}_{\mathrm{IDIP}}} \tag{7.8}$$

for the second definition. The other strategy is to compute binned Scale Factors from the binned efficiencies in data and Monte Carlo from equations 7.9:

$$\mathrm{SF}^{\mathrm{first}}_{\mathrm{anti\text{-}veto}} = \frac{1 - \epsilon^{\mathrm{Data}}_{\mathrm{RECO}} \times \epsilon^{\mathrm{Data}}_{\mathrm{Tracking}} \times \epsilon^{\mathrm{Data}}_{\mathrm{IDIP}}}{1 - \epsilon^{\mathrm{MC}}_{\mathrm{RECO}} \times \epsilon^{\mathrm{MC}}_{\mathrm{Tracking}} \times \epsilon^{\mathrm{MC}}_{\mathrm{IDIP}}} \tag{7.9}$$

and 7.10:

$$\mathrm{SF}^{\mathrm{second}}_{\mathrm{anti\text{-}veto}} = \frac{1 - (\epsilon^{\mathrm{Data}}_{\mathrm{RECO}} \times \epsilon^{\mathrm{Data}}_{\mathrm{Tracking}} + \epsilon^{\mathrm{Data}}_{\mathrm{T\text{-}RECO}} \times \epsilon^{\mathrm{Data}}_{\mathrm{T\text{-}Tracking}}) \times \epsilon^{\mathrm{Data}}_{\mathrm{IDIP}}}{1 - (\epsilon^{\mathrm{MC}}_{\mathrm{RECO}} \times \epsilon^{\mathrm{MC}}_{\mathrm{Tracking}} + \epsilon^{\mathrm{MC}}_{\mathrm{T\text{-}RECO}} \times \epsilon^{\mathrm{MC}}_{\mathrm{T\text{-}Tracking}}) \times \epsilon^{\mathrm{MC}}_{\mathrm{IDIP}}} \tag{7.10}$$

and then fit them with a third-degree polynomial. Since the binnings for the steps are different, a common binning is first defined taking the most granular one. Then, the efficiency product in each of these new bins is computed using the efficiency value in the corresponding bin for each step, for data and Monte Carlo, and then the Scale Factors are computed using either 7.9 or 7.10. Statistical uncertainties are calculated based on error propagation for the product, inflating uncertainties in the original p_T bins to account for the splitting of the bin widths, assuming a uniform distribution of events in the original p_T bins. Finally, these Scale Factors are smoothed using a third-degree polynomial in each η bin.

The first approach to derive anti-veto Scale Factors requires the smoothing of the separate Scale Factors in data and Monte Carlo, and then a product (and a sum) and a ratio is required, resulting in functional form which is quite complicated, with larger statistical uncertainties. The second approach requires a much simpler functional form, with lower statistical uncertainties, but in this particular case one has to deal with the different binnings for the different steps.

Both approaches to derive the anti-veto Scale Factors were tested by performing a bias test in which the nominal template is derived with one set of Scale Factors and uncertainties, which will be described

---

[9] We consider the events in which the veto reconstruction fails for one of the muons.

in the following, and the pseudodata is computed using the other set. The shift in m_W was found to be compatible with 0 within the anti-veto Scale Factor uncertainties. Figures 7.42 and 7.43 show examples of RECO, $\mathrm{Tracking}$ and IDIP Scale Factors and the corresponding anti-veto Scale Factor derived using the first approach, the results of the smoothing on the alternate fits, performed separately for each step, are also shown in the same figures.

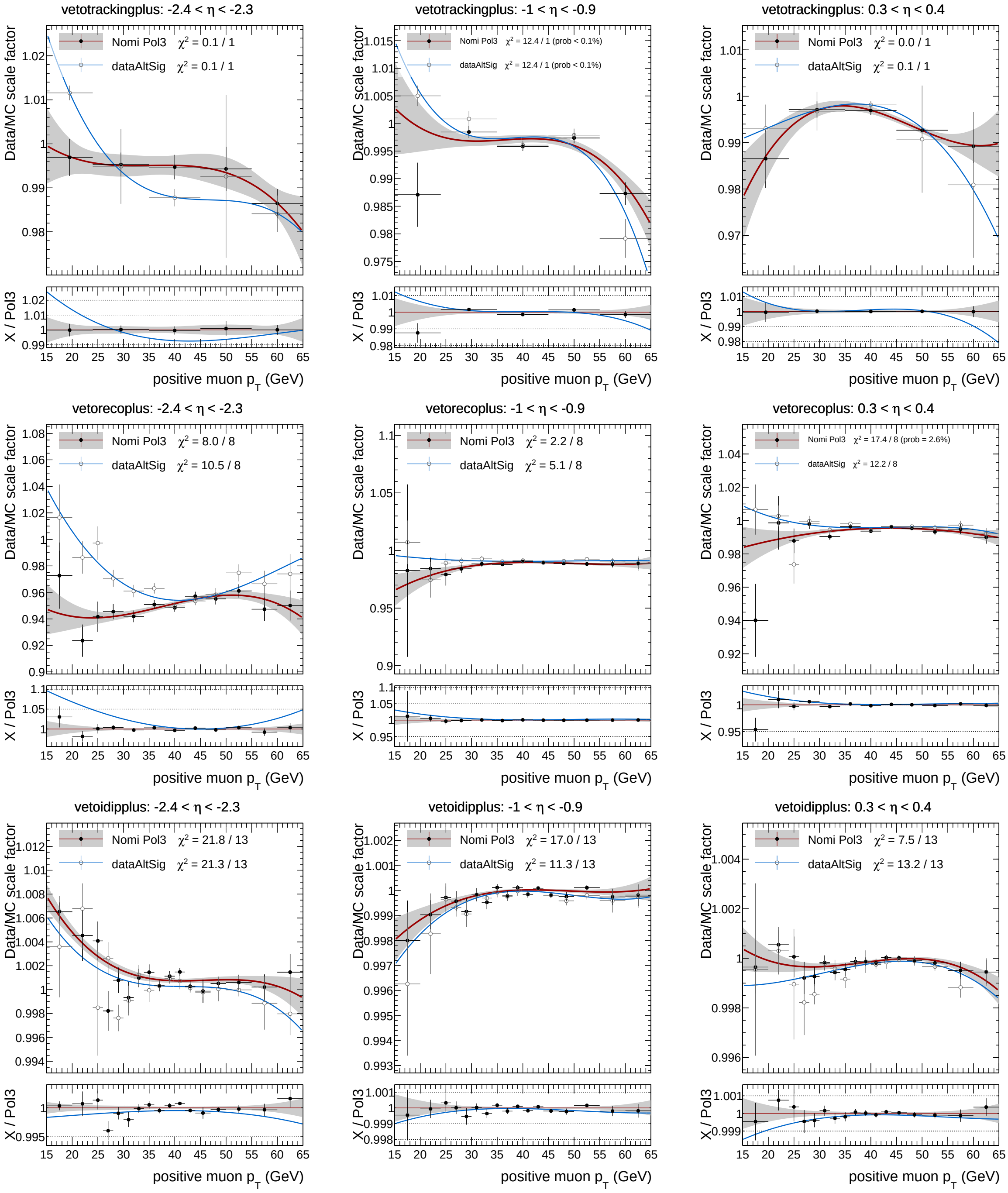


Figure 7.42: Smoothing of veto Scale Factors for global muons with positive charge, for the $\mathrm{Tracking}$, RECO, and IDIP steps from top to bottom, respectively, for the η bins indicated. The red line is the result of the smoothing on the nominal Scale Factors, while the blue one is the result on the Scale Factors obtained from the alternate invariant mass fits.

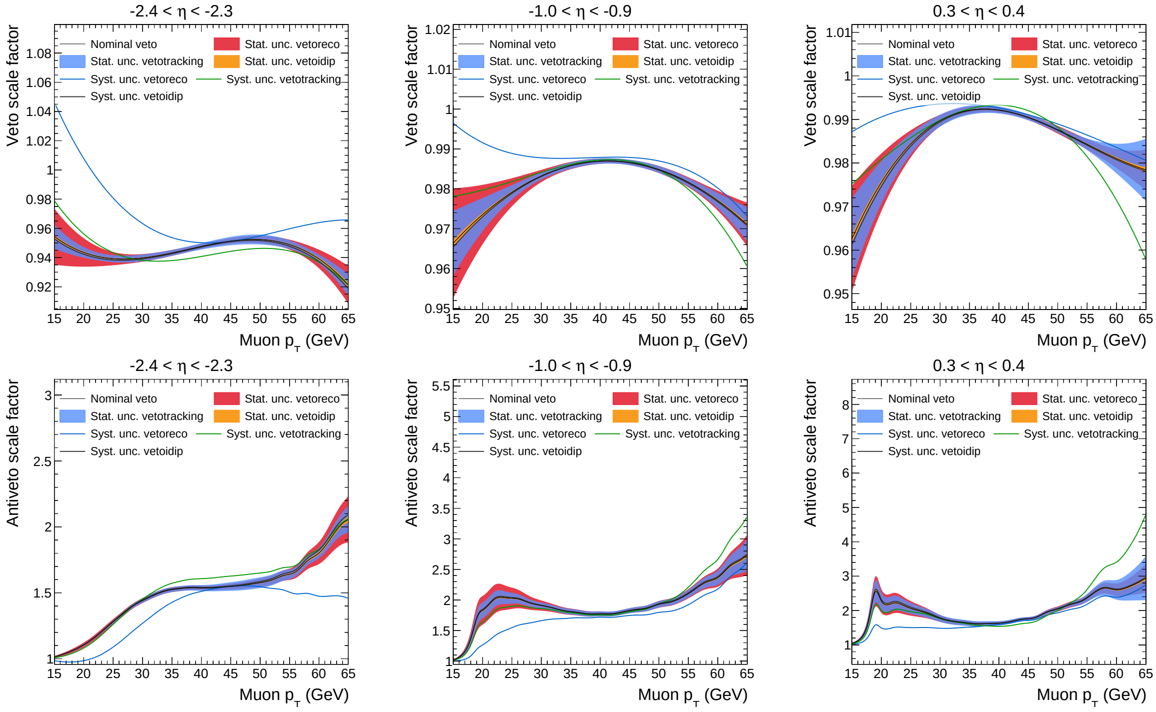


Figure 7.43: Overall veto (top) and anti-veto (bottom) Scale Factors for the η bins indicated. The solid bands represent the statistical uncertainty from each single step entering the product, while the lines are the systematic variations obtained from smoothing the Scale Factors obtained from the invariant mass fits with the alternate signal model.

These Scale Factors are applied on both the $Z \to \mu\mu$ and $Z \to \tau\tau$ events, although this background is much smaller, as a function of the generator level variables for muons with $p_T > 15$ GeV and $|\eta| < 2.4$: this allows us to properly take into account cases in which the veto muons are not reconstructed falling out of the acceptance (mostly geometrical, but also in p_T) and cases in which the veto reconstruction fails even with the muon in acceptance, both of which result in events that are selected by the signal selection criteria. The choice of the p_T definition to use to apply the veto efficiency corrections might result in different $p_T - \eta$ templates, since in some steps the efficiencies are measured as a function of muon spectrometer variables, but this should have a negligible impact as the corrections are applied to the second muon[10] and not the one used to fill the $p_T - \eta$ templates.

Figure 7.44 shows the probability that a generator-level muon fails the veto reconstruction as a function of η and p_T for both veto selections, with and without corrections. The nominal veto definition is the one in which only global muons are used, the alternate definition is the one where both global and tracker muons are used. It can be seen that the alternate definition has a much larger rejection efficiency, resulting in less events being selected, especially for $p_T \approx 15$ GeV and high $|\eta|$, mostly because in the nominal definition a cut on both the inner track and outer track $p_T > 15$ GeV is applied. The corrections decrease the rejection efficiency significantly in both cases. The anti-veto Scale Factors are found to be $O(2)$ in most $p_T - \eta$ bins, as for the examples in figure 7.43.

Figure 7.45 shows the $Z \to \mu\mu$ background after the full W analysis selection, for the two veto definitions, with all the efficiency corrections, including those for anti-veto, applied. The background yields differ

---

[10] One could have also considered only the events with both muons in the geometrical acceptance, the muon that is used to fill the $p_T - \eta$ template is weighted by the normal Scale Factors, while the other gets a weight that is: $\mathrm{SF}_{\text{anti-veto}} = \frac{1-\epsilon^{\text{Data}}_{\text{veto}}}{\epsilon^{\text{MC}}_{\text{veto}}}$ for $p_T > 15$ GeV, and $\mathrm{SF}_{\text{anti-veto}} = \frac{1}{\epsilon^{\text{MC}}_{\text{veto}}}$ below, where the corrections are evaluated with the proper p_T, but this still does not take into account cases in which one of the muons is out of the geometrical acceptance.

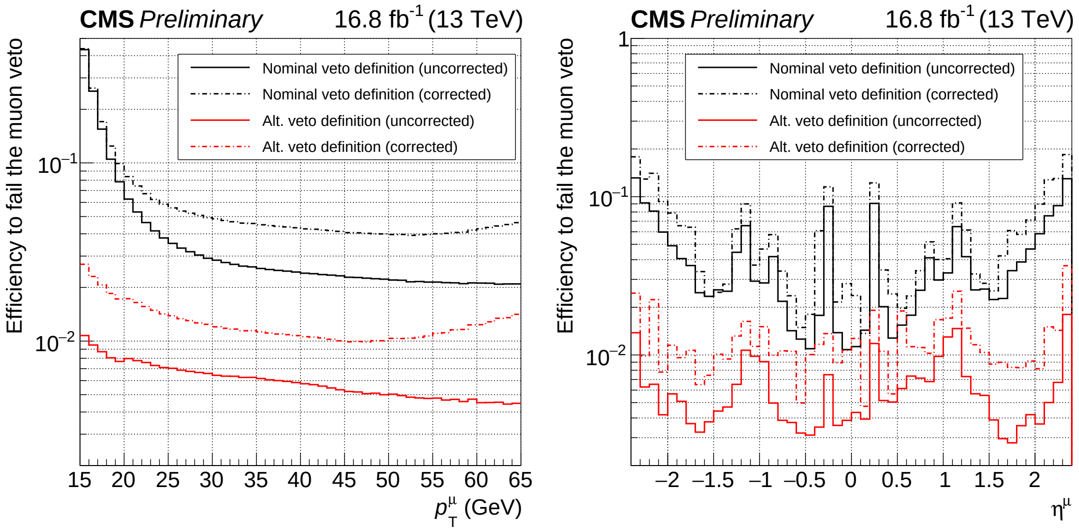


Figure 7.44: Probability that a muon fails the veto selection as a function of p_T (left) and η (right), for the default veto selection (in black) and the alternate veto selection (in red), obtained from simulation. In the former case, only global muons (which also pass the $\Delta R < 0.3$ requirement between internal and external variables, the high purity flag and have standalone $p_T > 15$ GeV) are considered, while in the latter case tracker muons with $p_T > 15$ GeV are also included. In the dashed curves the efficiency corrections derived from data are applied to simulation. The alternate veto selection is looser and implies a smaller efficiency to fail the veto, especially at low p_T because of the absence of acceptance effects for the track reconstruction in the muon spectrometer. The structure of the muon spectrometer can be seen as a function of η. The difference in the measured W boson mass from changing the veto selection is 0.05 MeV.

up to 20%, mostly at high p_T and in the central η region.

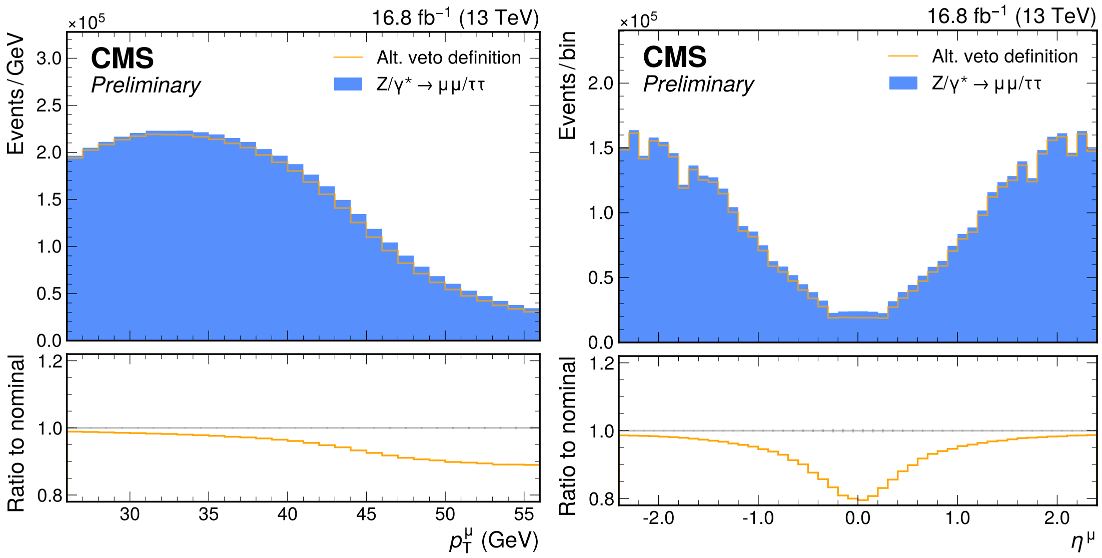


Figure 7.45: Comparison of the yields of the $Z \to \mu\mu$ background between the nominal (in blue) and alternate (the green line) veto definition, as a function of p_T (left) and η (right). All the experimental efficiency corrections are applied in both cases. The looser alternate veto selection provides a larger Z background rejection but has larger uncertainties in the measured efficiency corrections. The difference in the measured W boson mass from changing the veto selection is 0.05 MeV.

Figure 7.46 shows the impact of the anti-veto Scale Factors for the nominal and alternate veto definition as a function of p_T on the combination of the $W \to \mu\nu$ and $Z \to \mu\mu$ processes, a variation in m_W of ± 100 MeV is also shown in the ratio plot for comparison. The corrections affect mostly the p_T bins above the Jacobian peak, almost up to the 100 MeV variation for the nominal veto, while they modify the spectrum significantly less in the case of the alternate. This is mostly due to the reduced background in the latter definitions, so that the correction is applied to a smaller fraction of the total $Z \to \mu\mu$ background, rather than the magnitude of the corrections themselves. The more conservative version of the veto is used to derive m_W because we deem the anti-veto corrections to be more reliable. As a cross check, m_W was extracted using both veto definitions and only a shift of 0.05 MeV was found between the two cases, corroborating that the result seems to be very stable regardless of the veto definition.

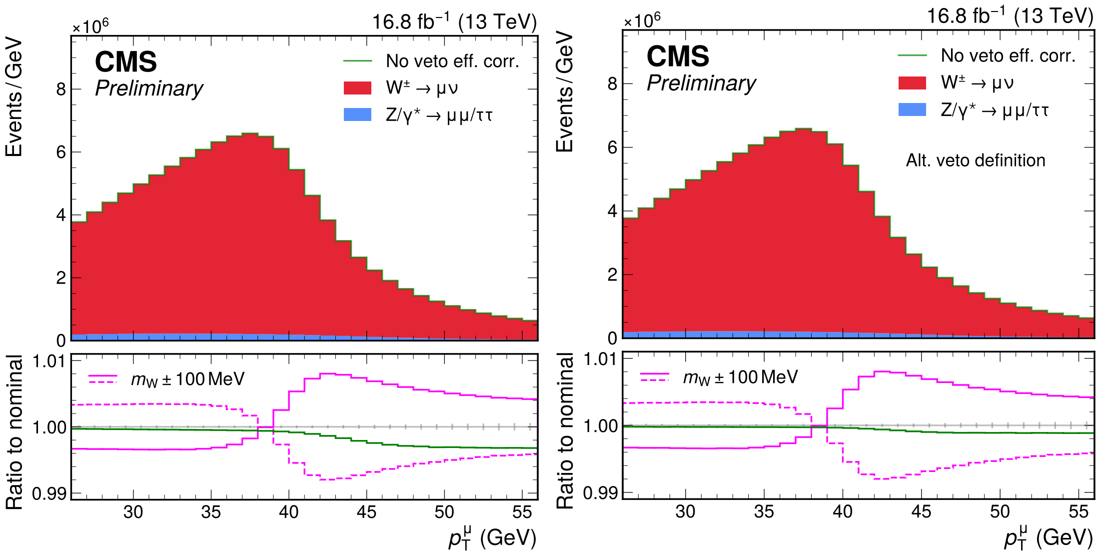


Figure 7.46: Impact of neglecting the muon veto efficiency correction on the sum of the W boson and Z boson processes, for the default veto selection used in the analysis (left) and for the alternate veto selection (right). In the former case, only global muons (which also pass the $\Delta R < 0.3$ requirement between internal and external variables, the high purity flag and have standalone $p_T > 15$ GeV) are considered, while in the latter case tracker muons with $p_T > 15$ GeV are also included. The looser alternate veto selection provides a larger Z boson background rejection but has larger uncertainties in the measured efficiency corrections. The difference in the measured W boson mass from changing the veto selection is 0.05 MeV.

As a final validation, figures 7.47 and 7.48 show comparisons between data and Monte Carlo simulation on W-like events with a modified selection, so that the non-triggering muon is required to pass the nominal veto selection and the p_T range is extended to 15 GeV$< p_T < 60$ GeV. The non-triggering muon receives veto efficiency corrections obtained from the anti-veto corrections described previously using equation 7.5 (where $\mathrm{Isolation}$ is replaced by anti-veto). The results are in agreement within the systematic uncertainties on the Scale Factors from the alternate signal model in the invariant mass fits, as can seen in figure 7.42. Figures 7.49 and 7.50 show the same comparisons for the alternate veto definition, similar conclusions can be drawn.

The systematic uncertainties stemming from the veto efficiency corrections are described in more detail in 7.4.11, are derived using a similar approach as in the case of the main efficiency corrections described in 7.4.11: for the nominal definition, where only the global muons are used, the statistical uncertainties contribute with $48 \times 2 \times 4$ systematic variations, and the systematic uncertainties contribute with $(48+1) \times 3$ systematic variations, for a total of 531 dedicated parameters.

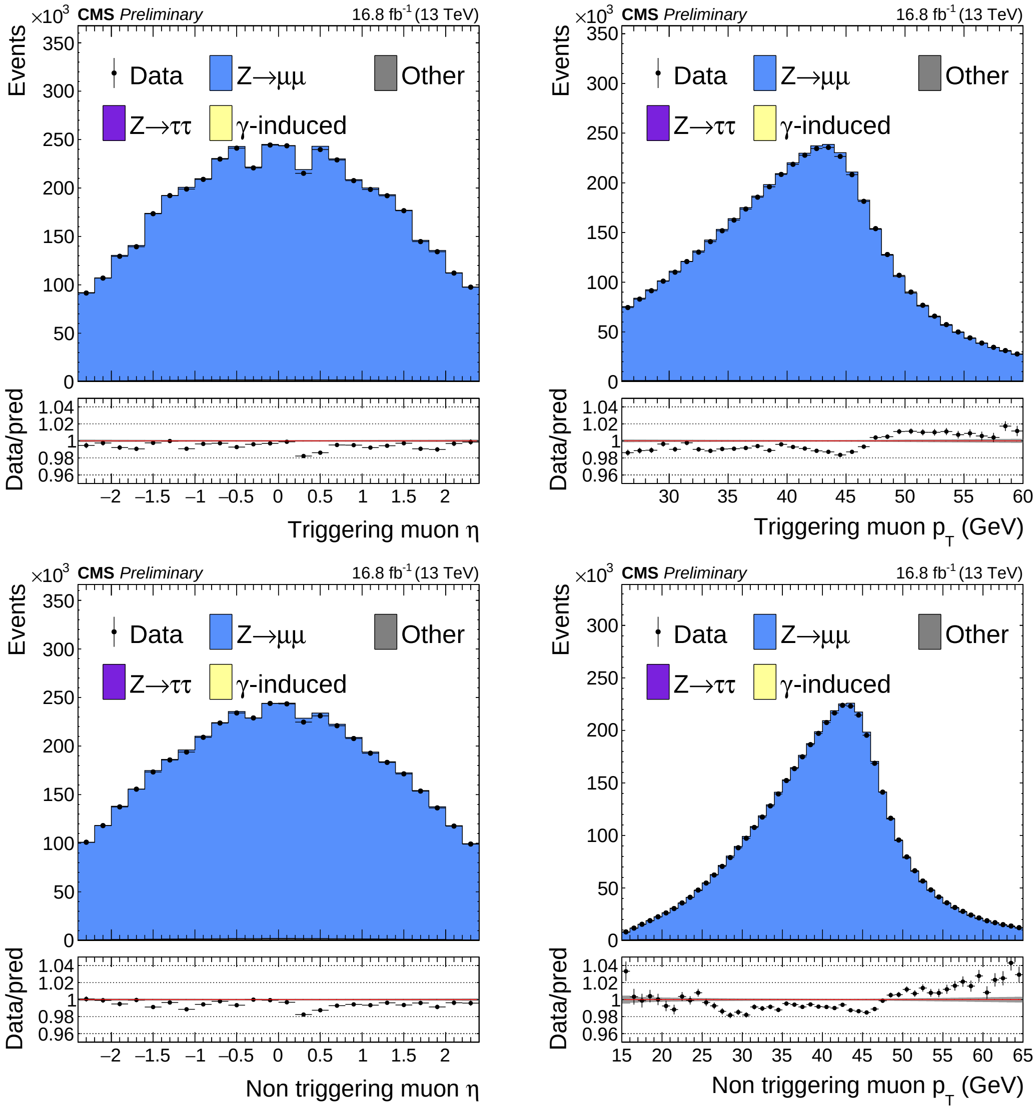


Figure 7.47: Data/MC comparison to validate veto SF for global muons in W-like events, where the non triggering muon passes the veto selection. Top plots show the triggering muon $p_T - \eta$ distributions, only the positive triggering muons are considered, while bottom ones show the same for the non triggering muon. Uncertainties in the ratio are statistical only.

7.4.10 Effect of residual clones on m_W

A procedure has been devised to take into account the presence of standalone clones, at $|\eta| > 2$ but also in the rest of the η space, similar to the tag-and-probe technique used to derive the efficiency corrections for muon reconstruction and identification efficiencies: a tag-and-probe pair is formed, where the tag muon passes the standard selection as in the other efficiency steps, while the probes are all global muons that also pass IDIP and $\mathrm{Isolation}$. The invariant mass for the pair, computed using the standalone variables for the probe muon, needs to be between 50 and 130 GeV. The passing probes are required to have an additional standalone muon with $p_T > 15$ GeV and "MergedStandAloneMuon_numberOfValidHits" ≥ 1 within a radius of $\Delta R < 0.3$. The invariant mass formed from the tag muon and the additional standalone muon must be between 50 and 130, as in the efficiency derivation for the $\mathrm{Tracking}$ step. Here, no invariant mass fit is run and the efficiencies are computed by counting the events in the passing and failing samples, assuming no contamination from background processes.

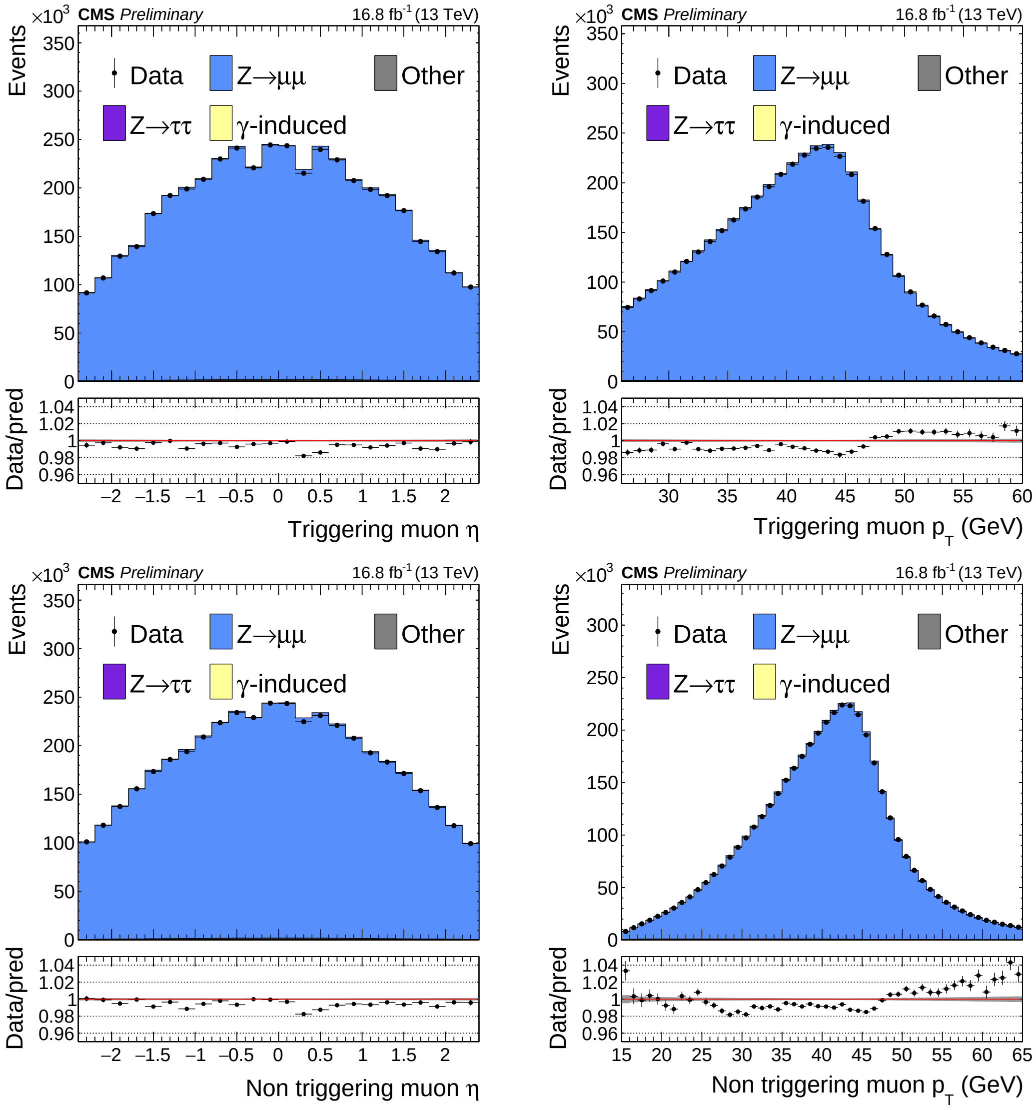


Figure 7.48: Same distributions as in figure 7.47 but negative triggering muons are considered.

In the presence of clone standalone muons, with the assumption that none of these muons becomes a global muon, the efficiency that is returned by the tag-and-probe method is related to the true efficiency in the following way:

$$
\begin{aligned}
\epsilon_{\text{measured}} &= \frac{\#\text{ passing probes}}{\#\text{ passing probes} + \#\text{ failing probes} + \#\text{ clones}} \\
&= \epsilon_{\text{true}} \frac{1}{1 + \frac{\#\text{ clones}}{\#\text{ passing probes} + \#\text{ failing probes}}} \\
&= \epsilon_{\text{true}} \frac{1}{1 + f_{\text{clones}}},
\end{aligned}
\tag{7.11}
$$

where f_{clones} indicates the fraction of clones on the total sample. This formula holds for both data and Monte Carlo simulation, so an additional correction to account for the difference in clones between data and Monte Carlo is:

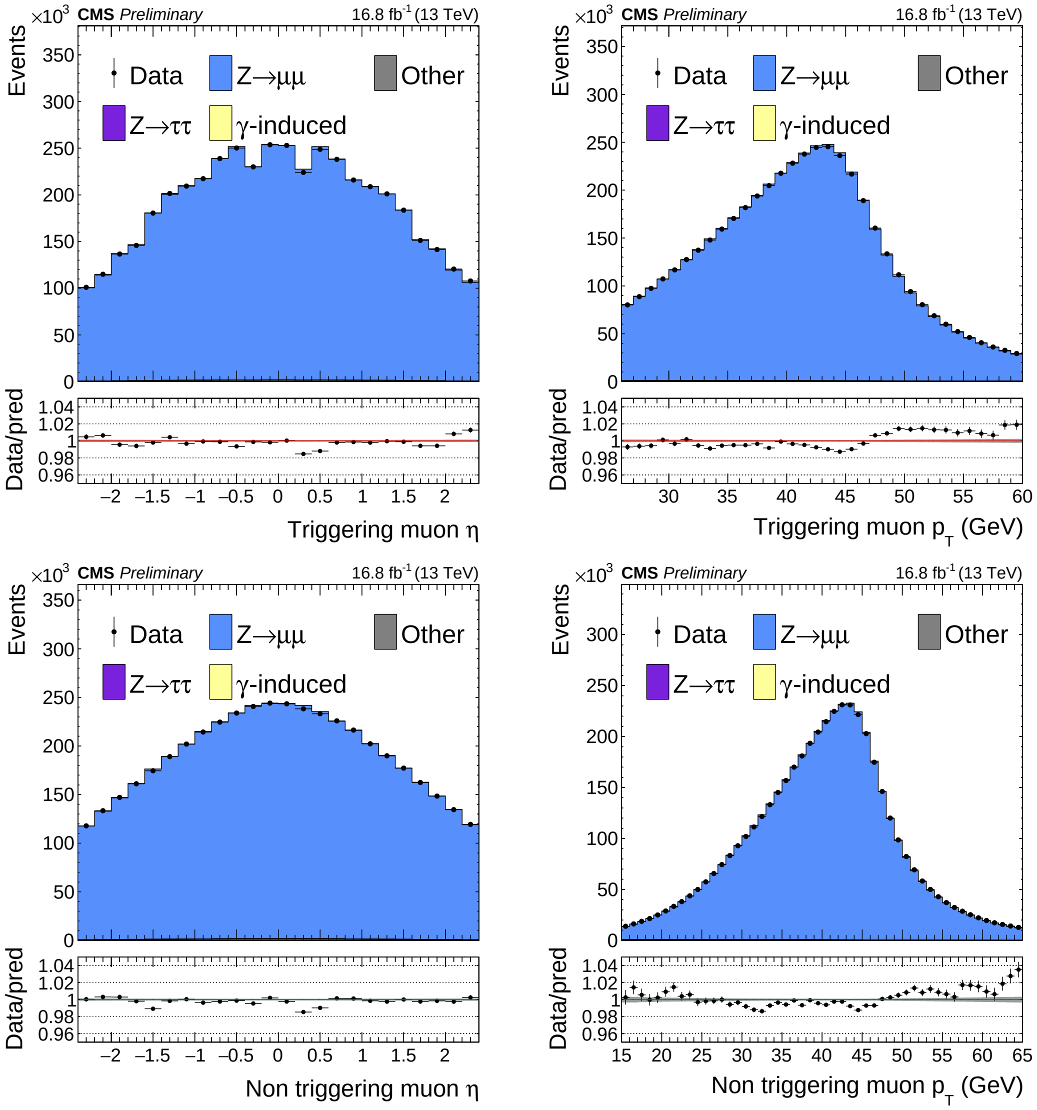


Figure 7.49: Same distributions as in figure 7.47 but for the global or tracker veto selection.

$$\mathrm{SF}_{\mathrm{clones}} = \frac{1 + f_{\text{clones in data}}}{1 + f_{\text{clones in MC}}}. \tag{7.12}$$

Figures 7.51 and 7.52 show the fraction of passing probes for data and Monte Carlo simulation and the correction that needs to be applied to the templates for the positive and negative charge, respectively.

It can be seen that the Scale Factor for $|\eta| > 2$ are consistent with unity within 0.1%, but it seems that there are differences in the central η region, mostly concentrated in certain η bins and Scale Factor seems to be mostly flat in p_T.

As in the standard tag-and-probe procedure, the variables used to compute the efficiencies and Scale Factor in the $p_T - \eta$ plane are those of the probe muon itself, which might not correspond to the p_T and η of the matched standalone muon. As an alternate approach, figures 7.53 and 7.54 show the same efficiencies and Scale Factors as in figures 7.51 and 7.52 but using the η and p_T of the candidate clone muons to fill the templates to compute the efficiencies. For real clone muons, this should not differ too much from the approach described before, but for muons that are not related to the probe and might still constitute a combinatorial background to the procedure this might differ quite significantly. For this source of contami-

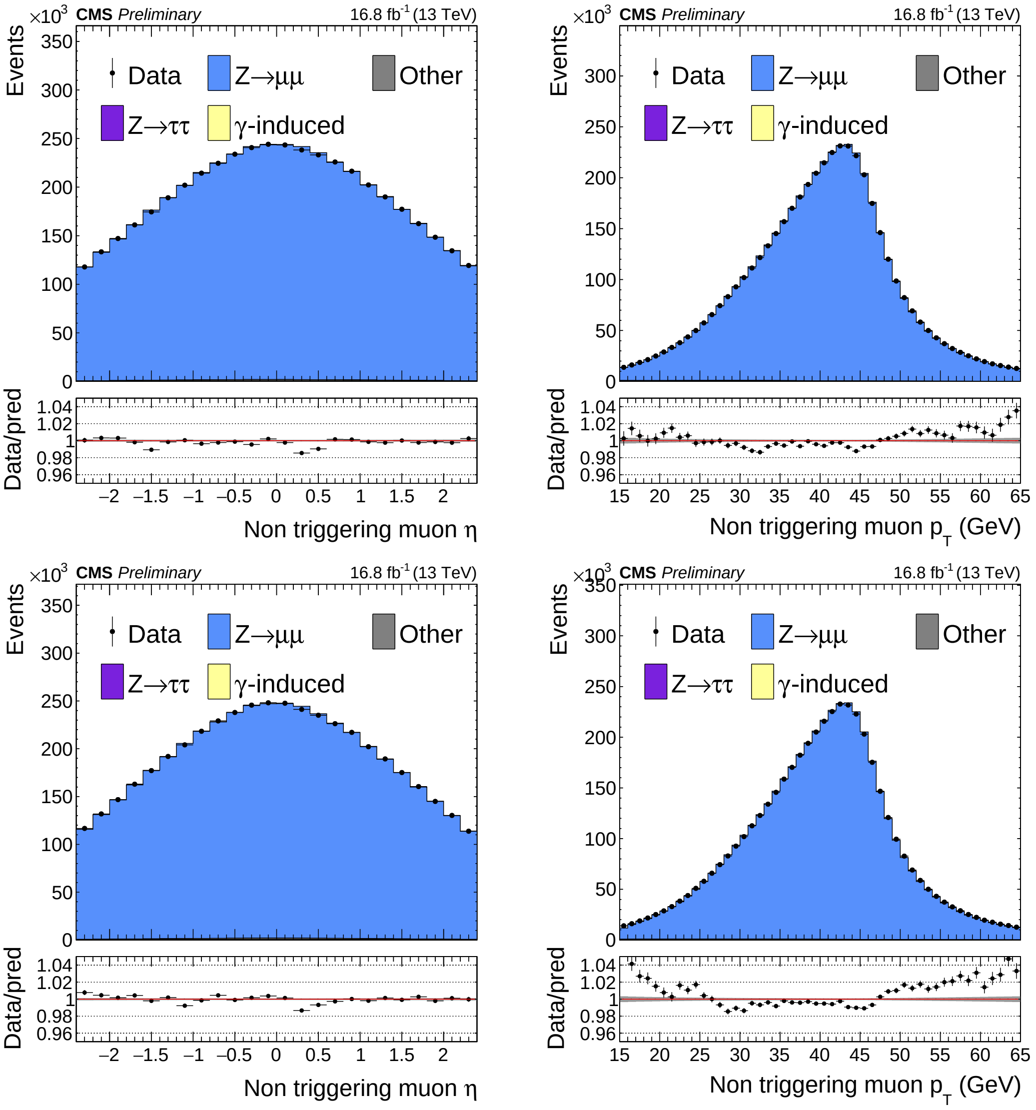


Figure 7.50: Same distributions as in figure 7.48 but for the global or tracker veto selection.

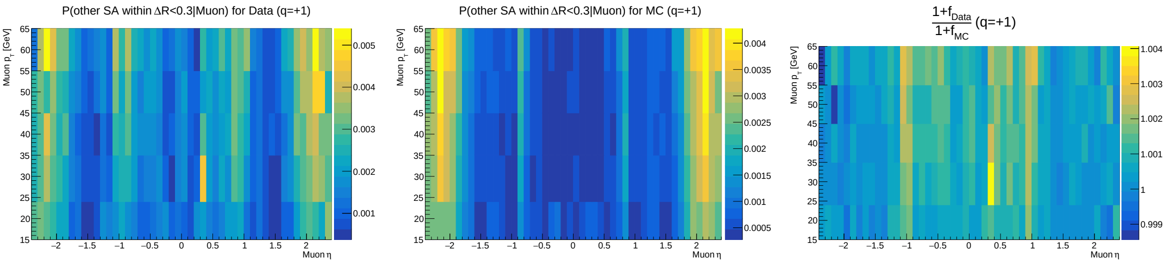


Figure 7.51: Fraction of probe muons that have an additional standalone muon with $p_T > 15$ GeV and ”MergedStandAloneMuon_numberOfValidHits” ≥ 1 within a radius of $\Delta R < 0.3$ for data (left), Monte Carlo (center) and Scale Factor to be applied to simulation as described in the text (right) for positive muon charge.

nation, the efficiency might not even be properly defined, as there would be no failing probe, with the effect

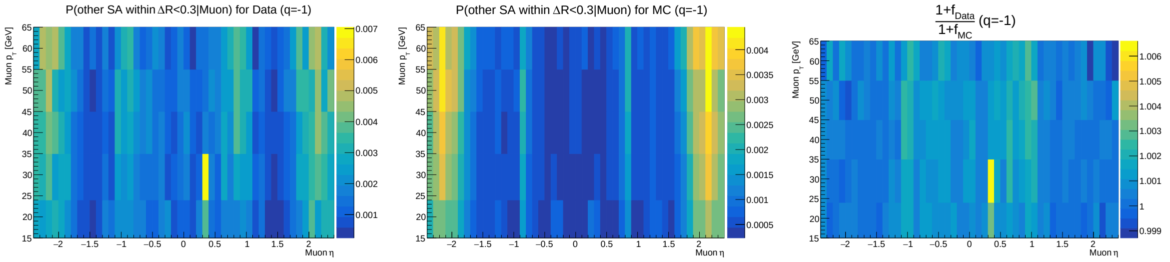


Figure 7.52: Fraction of probe muons that have an additional standalone muon with $p_T > 15$ GeV and "MergedStandAloneMuon_numberOfValidHits" ≥ 1 within a radius of $\Delta R < 0.3$ for data (left), Monte Carlo (center) and Scale Factor to be applied to simulation as described in the text (right) for negative muon charge.

of artificially increasing the fraction of clones in certain regions of the $p_T - \eta$ plane[11]. It can be seen that in this case this efficiency reaches 1% in certain bins of η in the first p_T bin, and the Scale Factors agree within 0.1% for $|\eta| > 2$, larger in the central η region. The large difference between the two approaches, especially at low p_T, might be ascribed to the presence of the combinatorial background mentioned above, for which additional clones randomly appear to be within $\Delta R < 0.3$ of the probe. The efficiencies computed in the two approaches are similar for the other p_T bins.

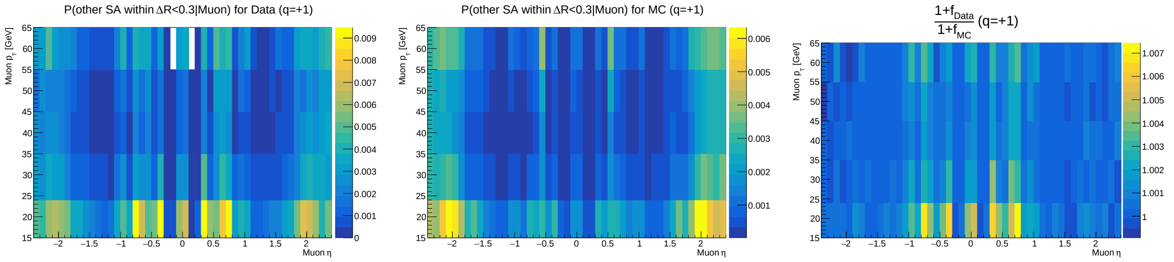


Figure 7.53: Fraction of probe muons that have an additional standalone muon with $p_T > 15$ GeV and "MergedStandAloneMuon_numberOfValidHits" ≥ 1 within a radius of $\Delta R < 0.3$ for data (left), Monte Carlo (center) and Scale Factor to be applied to simulation as described in the text (right) for positive muon charge. In this case, the variables of the candidate clone muon are used to fill the template for the passing probes.

Cross-checks were performed to verify that these muons actually arise from clones. The invariant mass distribution computed using the clone and tag variables, computed in data events, appears to be similar to that computed from the tag and standalone muon probe variables for $p_T > 24$ GeV, with some differences for lower p_T, as is shown in figure 7.55. It has been checked how many of the candidate clones are also global muons with $p_T > 10$ GeV, to check the validity of the assumption in equations 7.11 and 7.12 that none of the clones is a global muon. This is shown in figures 7.56 and 7.57, where it can be seen that these are only a negligible fraction of the total.

Additional cross-checks were made to assess the accuracy with which the fraction of clones in data and simulation are determined:

- The radius for the geometrical matching between probe and clone candidate was increased to $\Delta R < 1$. It was verified that the Scale Factors to correct for the presence of clone muons are consistent within 0.1%.

---

[11] This source of contamination could in principle be removed by running an invariant mass fit incorporating a background model (it is what is done in the $\mathrm{Tracking}$ step), but it will be demonstrated in the text that the impact on m_W is less than 1 MeV in any case.

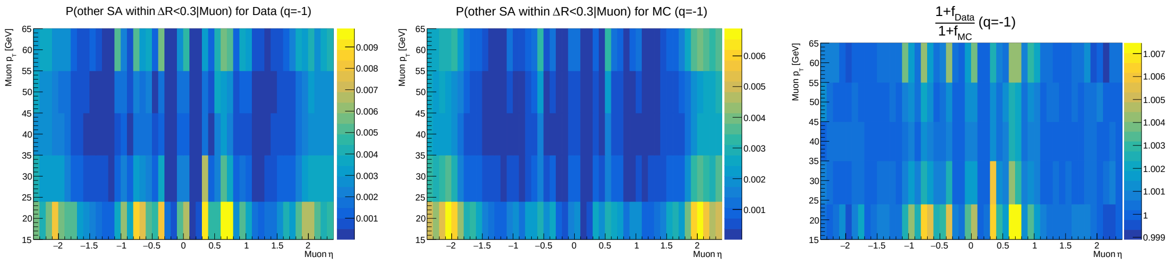


Figure 7.54: Fraction of probe muons that have an additional standalone muon with $p_T > 15$ GeV and "MergedStandAloneMuon_numberOfValidHits" ≥ 1 within a radius of $\Delta R < 0.3$ for data (left), Monte Carlo (center) and Scale Factor to be applied to simulation as described in the text (right) for negative muon charge. In this case, the variables of the candidate clone muon are used to fill the template for the passing probes.

- It was verified that the fractions of muons that has more than one standalone muon within a radius of $\Delta R < 1$ is a negligible portion of the total, so that at most one clone muon is present in the majority of cases. This could have in principle affected the selection of the clone muon to evaluate the invariant mass and to fill the $p_T - \eta$ templates.
- It was verified that the presence of clone is not due to the tag muon, by reevaluating the same fraction for muons with a clone with $\Delta R < 1$ considering only events in which $\Delta R_{\text{tag,probe}} > 0.3$.
- The introduction of a requirement on the difference in p_T between the clone candidate and genuine muon was tested. Different Δp_T values were considered, from 5 to 30 GeV. The introduction of this requirement had the effect of lowering the fraction of clones mostly above 24 GeV, making the fraction of clones more similar to the results of figures 7.53 and 7.54 for these bins.

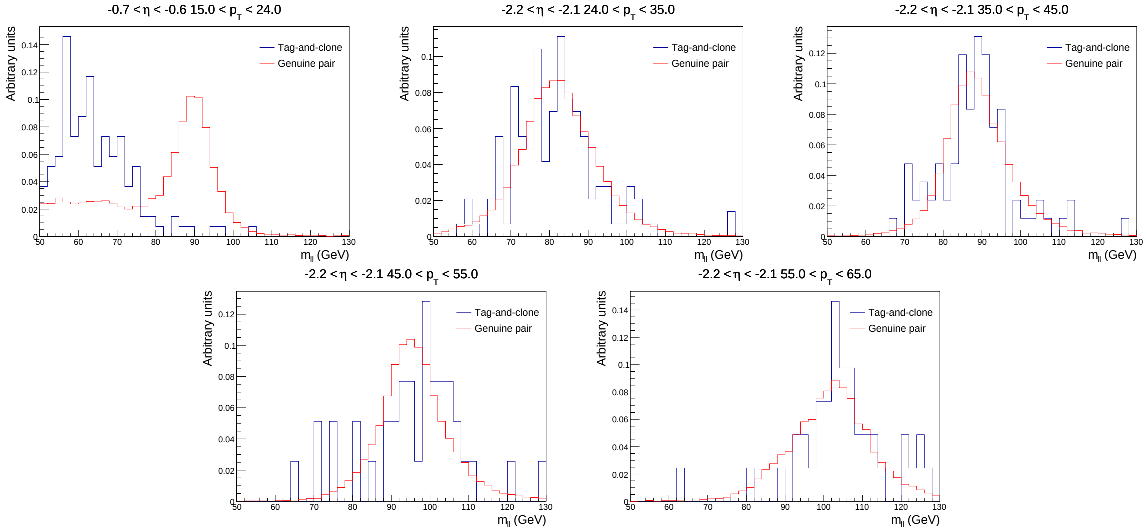


Figure 7.55: In blue, the invariant mass distribution obtained from the tag and the candidate clone variables. In red, the invariant mass distribution obtained from the tag and the standalone probe variables. These are shown for the different $p_T - \eta$ bins indicated. For the lowest p_T bin, a different η bin is shown to illustrate to source of difference in clone fraction seen in figures 7.53 and 7.54. These distributions are obtained from data events.

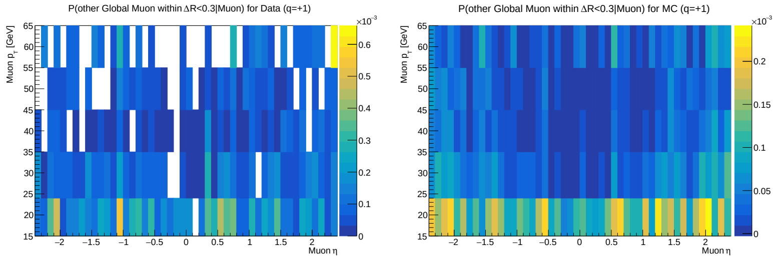


Figure 7.56: Fraction of probe muons that have an additional standalone muon with $p_T > 15$ GeV and "MergedStandAloneMuon_numberOfValidHits" ≥ 1 within a radius of $\Delta R < 0.3$ that is also a global muon with $p_T > 10$ GeV for data (left), Monte Carlo (right) for the positive charge.

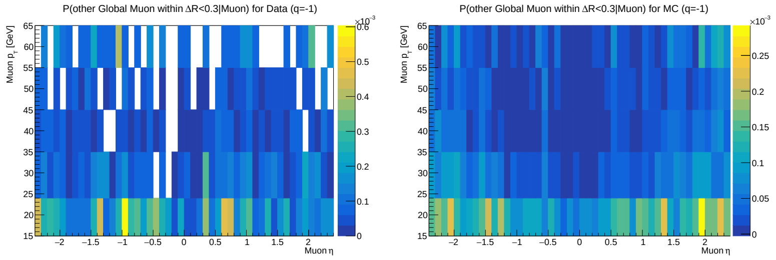


Figure 7.57: Fraction of probe muons that have an additional standalone muon with $p_T > 15$ GeV and "MergedStandAloneMuon_numberOfValidHits" ≥ 1 within a radius of $\Delta R < 0.3$ that is also a global muon with $p_T > 10$ GeV for data (left), Monte Carlo (right) for the negative charge.

The genuine fraction of candidate clones should be between those shown in figures 7.51 – 7.53: the corrections shown on the right of both sets of figures were used as additional event weights as a function of the standalone muon variables to modify the $p_T - \eta$ spectrum. With these new templates, the complete fit to extract m_W from $p_T - \eta - q$ was run, and the difference in m_W with respect to the published value is -0.57 MeV when the corrections in figures 7.51 and 7.52 are applied, and -0.20 MeV when the corrections in figures 7.53 and 7.54 are applied, so that the effect of the presence of these clones was found to be negligible. The differences between the two results might be attributed to statistical fluctuations in the corrections (which are not taken into account, but should have only a second order effect on the shift from the published value of m_W), or to migrations of clone candidates from the high to the low p_T region in the barrel. Indeed, in the barrel the muon spectrometer resolution is better, so that only a small fraction of muons used in the analysis should have standalone $p_T < 24$ GeV, meaning that the only the fraction of genuine clones affects the $p_T - \eta$ template in the analysis, leading to a smaller bias.

7.4.11 Systematic uncertainties from efficiencies

Uncertainties in the efficiency corrections are used to create alternate $p_T - \eta$ templates for signal and background processes that are then provided as input to the fit to extract m_W as described in 7.9.

The statistical uncertainties on muon reconstruction and identification Scale Factors are derived by diagonalizing the covariance matrices for the smoothing fits to derive their eigenvectors, corresponding to independent combinations of the fit parameters that define independent variations of fitted curves, and a separate nuisance parameter is assigned to each eigenvector, the eigenvalues are used to determine the magnitude of the 1σ variations. This is done separately for each η bin and charge for every step (except for isolation, for which the charge-dependence was found to be negligible).

Table 7.7 summarizes the number of nuisance parameters for each step. The number of variations for η bin for each each step corresponds to the degree of the polynomial used in the smoothing fits N plus

1, which is the number of coefficients of a polynomial of $N-$degree. For $\mathrm{Trigger}$ and $\mathrm{Isolation}$ some of these eigenvectors might actually induce variations mostly along the u_T-axis, so they are expected to have a small impact, since u_T is integrated out when the $p_T - \eta$ templates are built. The systematic uncertainties from varying the signal or background models in the invariant mass tag-and-probe fits are implemented by fitting the Scale Factors obtained using the alternate model with the same functional forms used for the nominal Scale Factors, assuming these to be correlated in p_T, so that all of the binned nominal Scale Factors are replaced by those obtained from the alternate fit. Then, a separate $p_T - \eta$ template is generated by replacing the nominal with the alternate Scale Factors independently for each step. We do this both independently and simultaneously for each η bin, so that both uncorrelated and correlated effects in $p_T - \eta$ should be captured, correlated for both charges. Statistical correlations between the nominal and alternate Scale Factors should become negligible after the smoothing of both, and no statistical uncertainties corresponding to the alternate fits is assigned.

Tables 7.8 and 7.9 summarize the number of variations due to the systematic uncertainties from varying the signal model and the background model, respectively. The systematic uncertainties on $\mathrm{Trigger}$ and $\mathrm{Isolation}$ are not derived from the fits to the $p_T - u_T$ distributions in each η bin, but from those integrated in u_T. In this case, for each η bin, and separately for the two steps, the nominal and alternate u_T-integrated Scale Factors are fitted with third-degree polynomials. Then, an alternate histogram is built by applying (as weights in the events) the ratio of the alternate and the nominal u_T-integrated smoothed Scale Factors to the $p_T - \eta$ template. Independent nuisance parameters are assigned to each of these variations.

The anti-correlation of $\mathrm{Isolation}$ uncertainties between the signal and sideband regions to estimate the nonprompt background is implemented through equation 7.5, used to define the corrections and variations to apply to the $p_T - \eta$ template in the sidebands used in the nonprompt background estimation. In this way, increases or decreases in the Scale Factors in the signal region, related to the statistical or systematic uncertainties, are properly propagated to the sidebands and, therefore, to the nonprompt background estimate.

The statistical and systematic uncertainties for the other steps and their nuisance parameters are fully correlated across all regions.

Step	Number of charges	Variations in p_T $(\times u_T)$	Number of η bins	Total nuisance parameters
RECO	2	4	48	
$\mathrm{Tracking}$	2	3	48	
IDIP	2	4	48	
$\mathrm{Trigger}$	2	4 ($\times$ 3)	48	
$\mathrm{Isolation}$	1	4 ($\times$ 3)	48	
				2784

Table 7.7: Number of nuisance parameters implemented in the analysis corresponding to the statistical uncertainty in the muon reconstruction and identification Scale Factors. The number of variations in p_T (and u_T) are equal to degree of the polynomial used in the smoothing plus one.

A few tests were run to take into other possible sources of systematic uncertainties in the fitted efficiencies, and therefore Scale Factors, other than the change in signal shape:

- raising the requirement on the minimum tag p_T from 25 GeV to 30 GeV;
- lowering the requirement on the relative PF isolation from 0.15 to 0.10.

The largest differences were observed in the $\mathrm{Trigger}$ and $\mathrm{Isolation}$ Scale Factors and efficiencies (only for $\mathrm{Isolation}$), shown in figures 7.58 and 7.59. The grey bands in the plots are the statistical uncertainties, derived after the smoothing procedure. The differences are mostly related to the change in the tag p_T

Step	Number of charges	Variations in η	Total nuisance parameters
RECO	1	48+1	
Tracking	1	48+1	
IDIP	1	48+1	
Trigger	1	48+1	
Isolation	1	48+1	
			245

Table 7.8: Number of nuisance parameters implemented in the analysis corresponding to the systematic uncertainty from varying the signal model in the muon reconstruction and identification Scale Factors. The variation fully correlated in η is also indicated.

Step	Number of charges	Variations in η	Total nuisance parameters
RECO	1	48+1	
Tracking	1	48+1	
			98

Table 7.9: Number of nuisance parameters implemented in the analysis corresponding to the systematic uncertainty from using a template driven model for the background in the muon reconstruction and identification Scale Factors. The variation fully correlated in η is also indicated.

threshold, are they are related to the dependence on the boson recoil and its proximity to the muon as already discussed in 5.2.1. In principle, this source of uncertainty is already taken into account in the derivation of the efficiency corrections thanks to the u_T binning. It can also be noticed that the Scale Factors, unlike the efficiencies, vary only mildly, with only Isolation exhibiting mild changes at low p_T at the per-mille level. No additional systematic uncertainties are therefore implemented in the fit.

Regarding the anti-veto efficiency corrections, considering the nominal result that uses only global muons in the veto definition, a procedure similar to that for the reconstruction and identification Scale Factors is employed. In the first approach, independent nuisance parameters are assigned corresponding to each or reconstruction steps that appear in equation 7.7, in the second approach, the independent nuisance parameters are assigned to the variations induced by the eigenvectors of the covariance matrix from the fits to the anti-veto Scale Factors obtained through 7.9, resulting in a smaller number of nuisance parameters. This is done separately for each η bin in both approaches, resulting in additional 1056 nuisance parameters in the first approach, and 384 for the second approach.

Systematic uncertainties due to changing the signal model in the invariant mass fits are implemented differently in the two approaches. In the first case, an alternate set of smooth Scale Factors for each of the steps in the veto selection is obtained by smoothing the binned Scale Factors obtained from the alternate invariant mass fits. Then, equation 7.7 is used to compute the anti-veto Scale Factors corresponding to the systematic variations, by replacing the nominal smooth Scale Factors with the alternate ones separately for each step at a time. This is done separately for each η bin. In the second case, an alternate set of binned anti-veto Scale Factors is computed using equation 7.9 where the binned nominal Scale Factors are replaced with the alternate ones, separately for each of the steps of the veto selection. These alternate anti-veto Scale Factors are then smoothed with a third-degree polynomial, as in the nominal case. The correlation scheme in η is the same as for the reconstruction and identification Scale Factors in both approaches, so 49 nuisance parameters are assigned for each step, resulting in a total of 147 nuisance parameters.

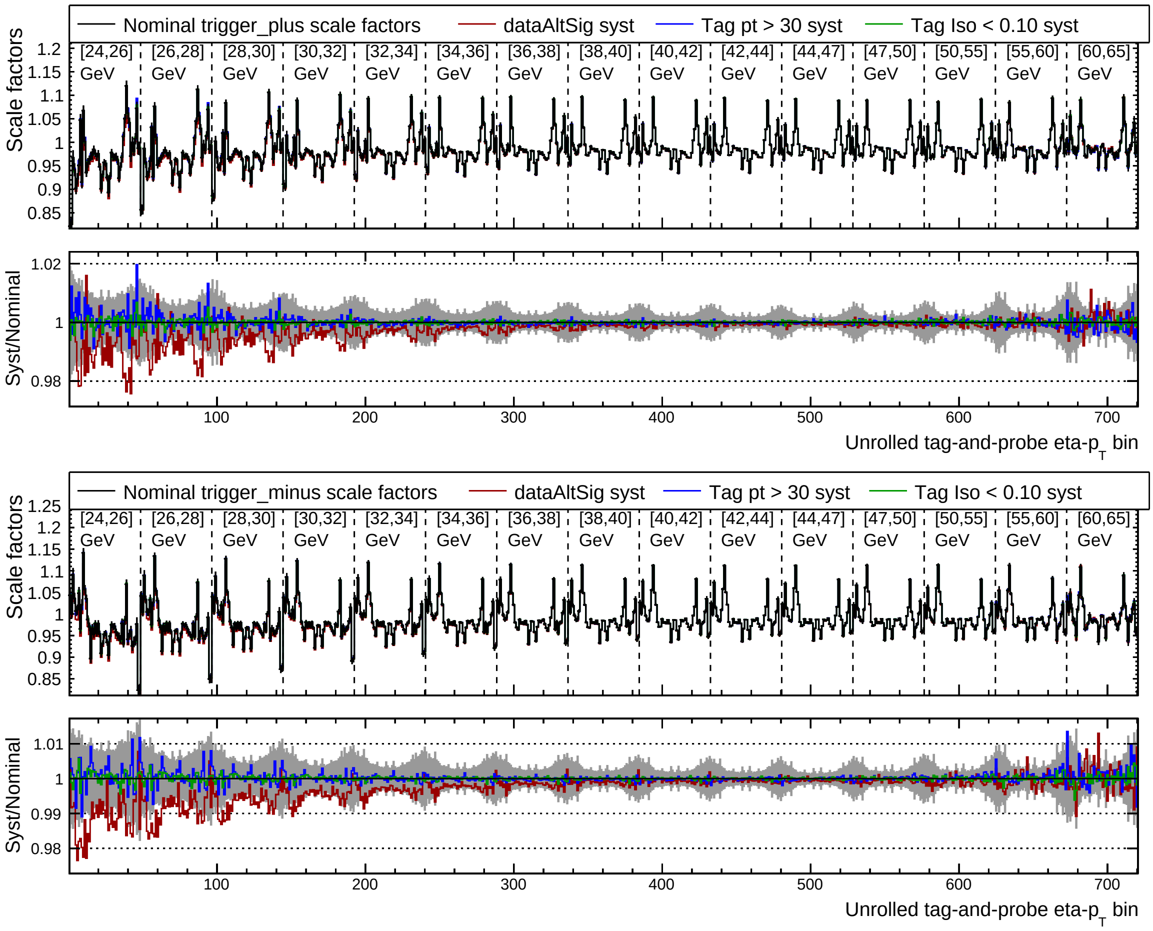


Figure 7.58: Systematic uncertainties on muon trigger Scale Factors for positive (top) or negative (bottom) probe charge, for change in signal model (red), tag p_T (blue), or tag isolation (green).

The second approach is used in the fit to data to extract m_W as it requires a lower number of systematic variations. It was checked, however, that the expected uncertainty, defined as the uncertainty on m_W obtained by fitting the sum of all the Monte Carlo processes set to their nominal parameters, from the anti-veto corrections is very similar for both approaches, slightly larger for the former, because of the increase number of nuisance parameters. No correlation is considered between the anti-veto corrections uncertainties and those from muon reconstruction and identification for simplicity: as will be seen in 7.10 the impact on the total expected uncertainty on m_W is less than 0.5 MeV, low enough so that the correlation can be neglected.

Statistical uncertainties in the ECAL prefiring efficiencies from the limited size of the sample with "unprefirable" events induce a variation of signal event yields ranging from less than 0.05% to at most 0.1% in the high p_T region, and a single nuisance parameter is assigned to it. For the muon prefiring, independent nuisance parameters corresponding to difference sources of variations of the prefiring weights are implemented. The statistical uncertainty is treated as 11 uncorrelated nuisance parameters, one for each $|\eta|$ bin used in the measurement of the prefiring probability. An additional independent nuisance parameter is implemented for the hotspot region shown in figure 7.38c. The systematic uncertainty is defined as a uniform 11% scaling of the prefiring probability in each $|\eta|$ bin, and a single nuisance parameter correlated across all bins is assigned.

The total impact from the efficiency corrections on m_W is 3.0 MeV, with approximately equal contributions from the statistical and systematic sources. The uncertainties on the corrections on the muon

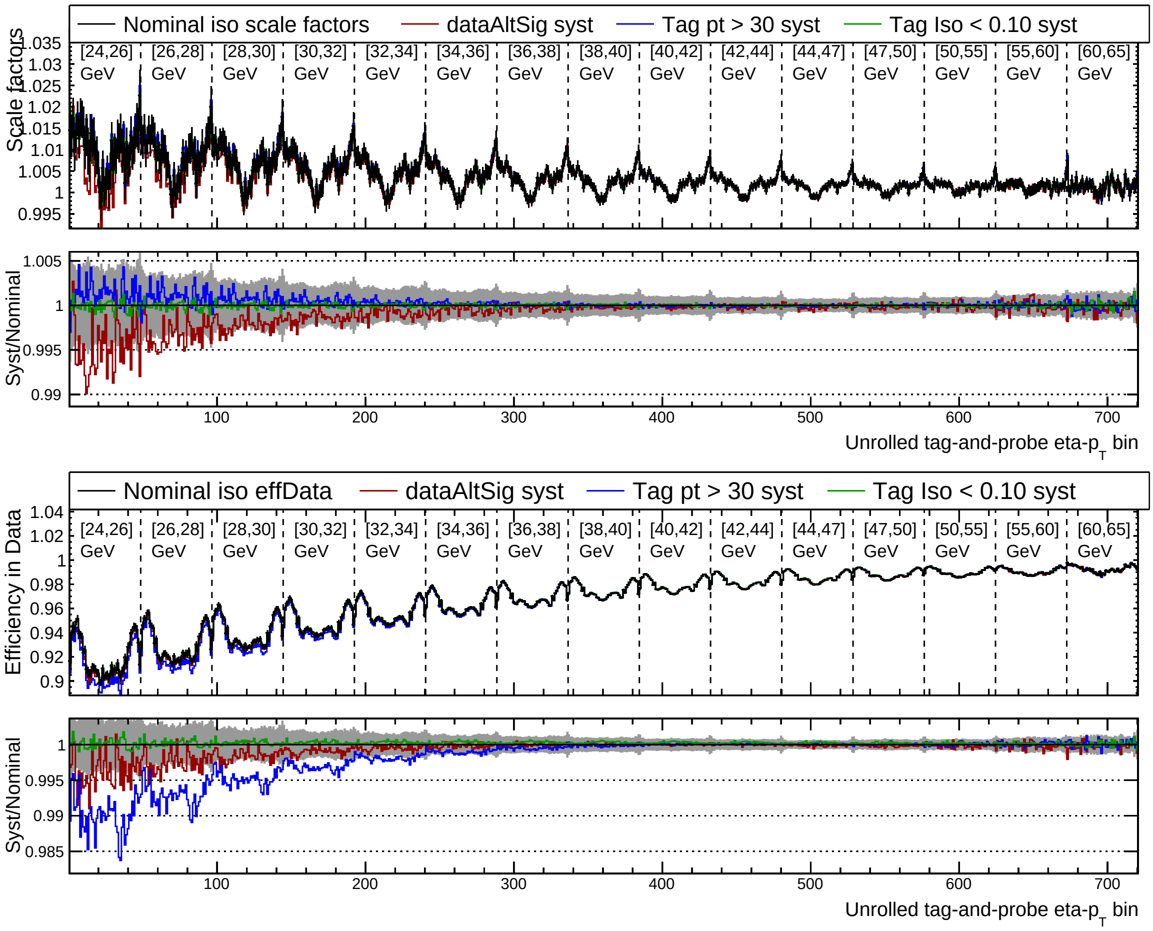


Figure 7.59: Systematic uncertainties on muon isolation Scale Factors (top) and efficiencies in data (bottom) inclusively in probe charge, for change in signal model (red), tag p_T (blue), or tag isolation (green).

veto efficiencies have an impact on m_W of less than 0.5 MeV. The prefiring corrections have an impact of 0.7 MeV.

7.5 Muon scale and resolution calibration

As discussed in section 5.1, for the p_T^μ range of interest for the m_W analysis, the muon kinematic information is extracted entirely from the information in CMS silicon tracker, while the muon spectrometer is used for muon identification. The inner track is reconstructed using the standard CMS iterative tracking algorithm based on the Kalman filter [92]. However, it was found that the Kalman filter fit introduced biases in the reconstructed p_T^μ with respect to the true p_T^μ in simulated events, of order O(1%). In standard CMS analyses, this is reduced to O(10^{-3}) by applying the *Rochester corrections* [251] to data and Monte Carlo events, which is not sufficient to reach the O(10^{-4}) precision requirement for a m_W measurement with an uncertainty of 10 MeV, so this approach is not useful for this measurement.

After the standard track reconstruction is completed, muon tracks are refitted using a so-called "continuous variable helix" (CVH) fit, a global χ^2 fit that extends the generalized broken-line fit [252, 253] to incorporate continuous energy loss and multiple scattering from finite material elements. The material model of the CMS detector includes both the initial design of the tracker material and its support structures and the information from in situ measurements [254], and the starting point for alignment is the one used in the standard CMS reconstruction [255, 256]. This information is implemented in the CVH track fit using the GEANT4e propagator [198, 257, 258]. The magnetic field is modelled using a parameterization of the detailed three-dimensional solenoidal field map based on measurements made in the ground-level assembly hall without the CMS tracker [259]. Additional quality criteria to select pixel hits and a refined parameterization of the local hit position in trapezoidal strip modules in the endcaps of the strip detector are implemented in the CVH track refit.

Muon momentum corrections are derived to take into account biases arising from:

- magnetic field model, for example due to the absence of the CMS tracker in the measurements to build the three-dimensional map;
- energy loss, for example due to differences in the amount of material or due to the Gaussian model in the description of energy loss;
- and alignment, both due to intrinsic biases of the procedure and to the so-called "weak modes" (misalignment patterns that can bias the parameter extraction from the track without having an impact the overall χ^2 of the track fits [256]).

They are corrected in a two-step procedure. First, $J/\Psi \to \mu\mu$ decays are used to derive correction parameters for the three sources of bias indicated at the level of the individual modules of the CMS tracker, imposing that the muons from the decay are produced from a common vertex and have invariant mass consistent with that of the J/Ψ meson. The parameters derived in this first step correspond to three translation and three rotation degrees of freedom for each tracker module, and additional parameters to correct for the $z-$component of the magnetic field and amount of material in the proximity of each module.

However, this set of corrections has reduced sensitivity to weak modes, residual biases resulting from limitations in the J/Ψ meson mass constraint, convolution effects from the detector resolution and for final-state radiation, and background contributions. Residual biases are then taken into account in the second step of the calibration procedure, where fits to the invariant masses of dimuon pairs of J/Ψ decays are used to derive correction factors on the measured mass that are finely binned in the space of the p_T^μ and η^μ of the muons. The signal model is based on templates from simulation, convoluted with a Gaussian whose mean and standard deviation account for the residual scale and resolution difference, and an exponential shape is added to describe the combinatorial background. These correction factors are then converted into calibration factors for each individual muon, as a function of η^μ and charge.

Indicating the curvature of the muon $k = 1/p_T$, the calibration factors to the muon momenta derived in this second step are parameterized as:

$$\frac{\delta k}{k} = A_{i\eta} - \epsilon_{i\eta} k + \frac{qM_{i\eta}}{k}, \tag{7.13}$$

where $i\eta$ is the index representing the individual η bin (the corrections are extracted in the same 48 η bins used in the analysis to measure m_W), $A_{i\eta}$ represents the correction term for differences in the magnetic field, $\epsilon_{i\eta}$ represents the correction to the energy loss, and $M_{i\eta}$ represents the bias in the track sagitta resulting from a misalignment of the tracker in the plane transverse to the magnetic field. The relationship between the fitted values of $m_{\mu\mu}(J/\Psi)$ and the coefficients in equation 7.13 is:

$$\left(\frac{m_{\mu\mu}^{\text{Reconstructed}}}{m_{\mu\mu}^{\text{True}}}\right)^2 = \left(1 + A_{i\eta^+} - \epsilon_{i\eta^+}k^+ + \frac{M_{i\eta^+}}{k^+}\right)\left(1 + A_{j\eta^-} - \epsilon_{j\eta^-}k^- - \frac{M_{j\eta^-}}{k^-}\right), \tag{7.14}$$

where $i, j\eta^\pm$ is the index of the positive or negative muon from the decay of the J/Ψ. The calibration coefficients are derived via a minimum χ^2 fit using formula 7.14 to the muon scale correction factors derived from the J/Ψ invariant mass fits as a function of the p_T and η of the two muons.

In the following, it is explained briefly how the formula in equation 7.13 is derived: the bias in curvature due to differences in the magnetic field is:

$$\delta k = \frac{1}{qR}\left(\frac{1}{B_{\text{corr.}}} - \frac{1}{B}\right) = k\left(\frac{B}{B_{\text{corr.}}} - 1\right) = kA, \tag{7.15}$$

where the relation between the reconstructed radius in the transverse plane R, the charge q, p_T and magnetic field B ($B_{\text{corr.}}$ indicates the actual magnitude of the magnetic field)

$$p_T = qBR \tag{7.16}$$

has been used.

The correction in energy loss results in a difference in the energy of the muon:

$$E_{\text{corr.}} = E + \epsilon, \tag{7.17}$$

and, in the ultrarelavistic limit where the energy and the momentum can be considered the same, this is related to the transverse momentum through $p_T = E\sin\theta$, with θ the polar angle with respect to the beam line. Therefore, the correction in energy loss is related to the curvature through:

$$\delta k = \frac{1}{(E+\epsilon)\sin\theta} - \frac{1}{E\sin\theta} = \frac{k}{1+k\epsilon\sin\theta} - k = -\epsilon\sin\theta k^2 = -\epsilon' k^2, \tag{7.18}$$

where the approximation

$$\frac{k}{1+k\epsilon\sin\theta} = k(1 - k\epsilon\sin\theta), \tag{7.19}$$

valid for $|\epsilon| \ll E$, has been used and $\epsilon' = \epsilon\sin\theta$ is defined (the corrections are derived as a function of η regardless, so that reabsorbing $\sin\theta$ into ϵ does not lead to biases in the corrections).

Residual misalignment due to weak modes and localized detector effects such as the differences in simulation of the Lorentz angle, i.e. the angle with respect to the orientation of tracker modules of the drift of the electrons, due to the relative directions of the electric field in the silicon modules and the external magnetic field with respect to one another, can still be present after the corrections in the first step. These biases are charge dependent; in particular they have the opposite effect on positive and negative charge, and are independent of the curvature, as they only affect the relative position of the measured points used in the determination of the sagitta. The effect of the misalignment of the modules on track reconstruction is shown in figure 7.60. In the assumption that the shifts due to the misalignment are small, this bias is parameterized as an additive term in the corrected curvature, with a charge dependence:

$$\delta k = qM. \tag{7.20}$$

In Monte Carlo simulations, the geometry, magnetic field and material for both the generation of events and the reconstruction are the design ones, and it was verified that the p_T^μ reconstructed by the CVH algorithm (after the application of the modular corrections in the first step), agrees on average with the true one.

It was tested in Monte Carlo simulations that the formula in equation 7.13 is able to model the biases in the muon reconstruction: this was done by generating events using the design configurations of the geometry, magnetic field, and material, and then reconstructing p_T^μ assuming that the geometry, magnetic field and material are those determined in data, as, for example, determined in [254, 255]. This can in principle result in differences in the tracker configuration of magnitude similar to that of the residual differences corrected by the muon scale calibration procedure. The relative difference between the true and the reconstructed p_T^μ is evaluated in each η bin as a function of qp_T^μ, where q is the charge of the muon, and this is then fitted using the model in equation 7.13. This is shown in figure 7.61 for two representative η bins, for the

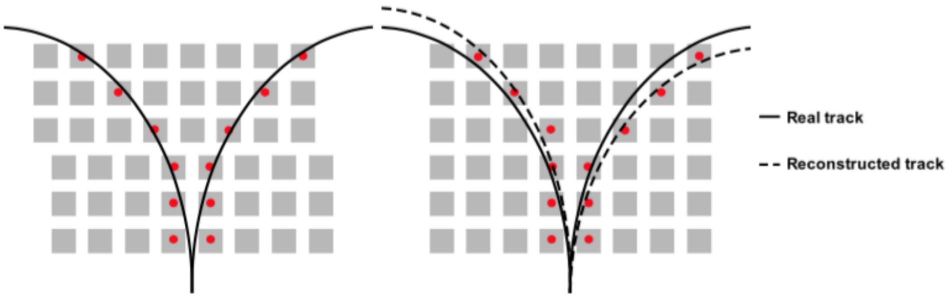


Figure 7.60: Effect of the misalignment of the modules. Left: actual position of the modules and real track, for positive and negative muon. Right: positions of the modules as implemented in the track reconstruction and reconstructed track, compared to the real tracks. From: `https://indico.cern.ch/event/1441575/`

standard CMS reconstruction, after the CVH refit is run and after the modular corrections are applied. It can be seen that the model does not reproduce the bias on p_T if the standard CMS reconstruction is used. The model works better if the CVH refit is used, and agreement is reached after the application of the modular corrections described above.

The final corrections for the magnetic field, energy loss, and alignment extracted in data from the procedure described above are shown, and used to correct the reconstructed muon p_T in data, in figure 7.62. The magnitude of the A term, due to biases in the magnetic field, is approximately 5×10^{-3} (which would result in a 40 MeV shift in m_W, if uncorrected), mostly flat in the barrel. The energy loss term ϵ, relevant mainly for muons at low p_T, is in most cases around 4 MeV. The alignment term is between ± 50 GeV^{-1} with a trend in η. As previously mentioned, for Monte Carlo simulations only the modular corrections are applied, as the detector geometry used in the simulation and CVH refit are consistent, so that no other scale correction needs to be applied.

The calibration procedure, in particular the model used to parameterize the bias as a function of p_T, was validated using $\Upsilon(1S) \to \mu\mu$ and $Z \to \mu\mu$ events. After all the corrections are applied, fits to the invariant mass distributions for $\Upsilon(1S)$ and Z are run to extract the mass scale as a function of the p_T and η of the muons from the decay, and additional correction terms for the magnetic field and the alignment are extracted using the relationship in formula 7.14. These terms are expected to be compatible with 0 within statistical uncertainties if the calibration procedure corrects for the biases in the reconstructed p_T.

The results of this comparison are shown in figure 7.63 for the charge-independent term A', which parameterizes the same dependence on the magnetic field as in equation 7.13 and the charge-dependent term M', which parameterizes the same dependence on the alignment as in equation 7.13, for J/Ψ (in green), $\Upsilon(1S)$ (in red) and Z (in blue). The values for this closure test are extracted in 24 bins in η, compared to the 48 for the standard corrections, so that small differences, compatible with 0, are present also for the J/Ψ. The yellow band in the figure shows the statistical uncertainty on the A and M terms from the nominal corrections, properly correlated across different η bins.

For $\Upsilon(1S)$, the additional corrections are compatible with 0 with the combined uncertainties from the $\Upsilon(1S)$ sample and the J/Ψ, but are not available at high $|\eta|$ values, due to the trigger selection applied in CMS to select $\Upsilon(1S)$ events and resolution effects that cause the reconstructed $\Upsilon(1S)$ mass distributions to merge with those from $\Upsilon(2S)$ and $\Upsilon(3S)$. The charge-independent corrections extracted in Z data are compatible with 0 within statistical uncertainties, and this can also be seen from the χ^2 value reported in the top right, which properly incorporates the correlations across bins for Z and J/Ψ events and the uncertainty in the world average value of the Z mass. Instead, the charge-dependent corrections extracted in Z events seem to exhibit a trend in η, and this is reflected in the χ^2 in the top right of the figure. This must be due to effects that might not have been taken into full account by the calibration procedure, for example, weak modes with different sensitivities to the Z and the J/Ψ, or the trigger selection in J/Ψ events. To

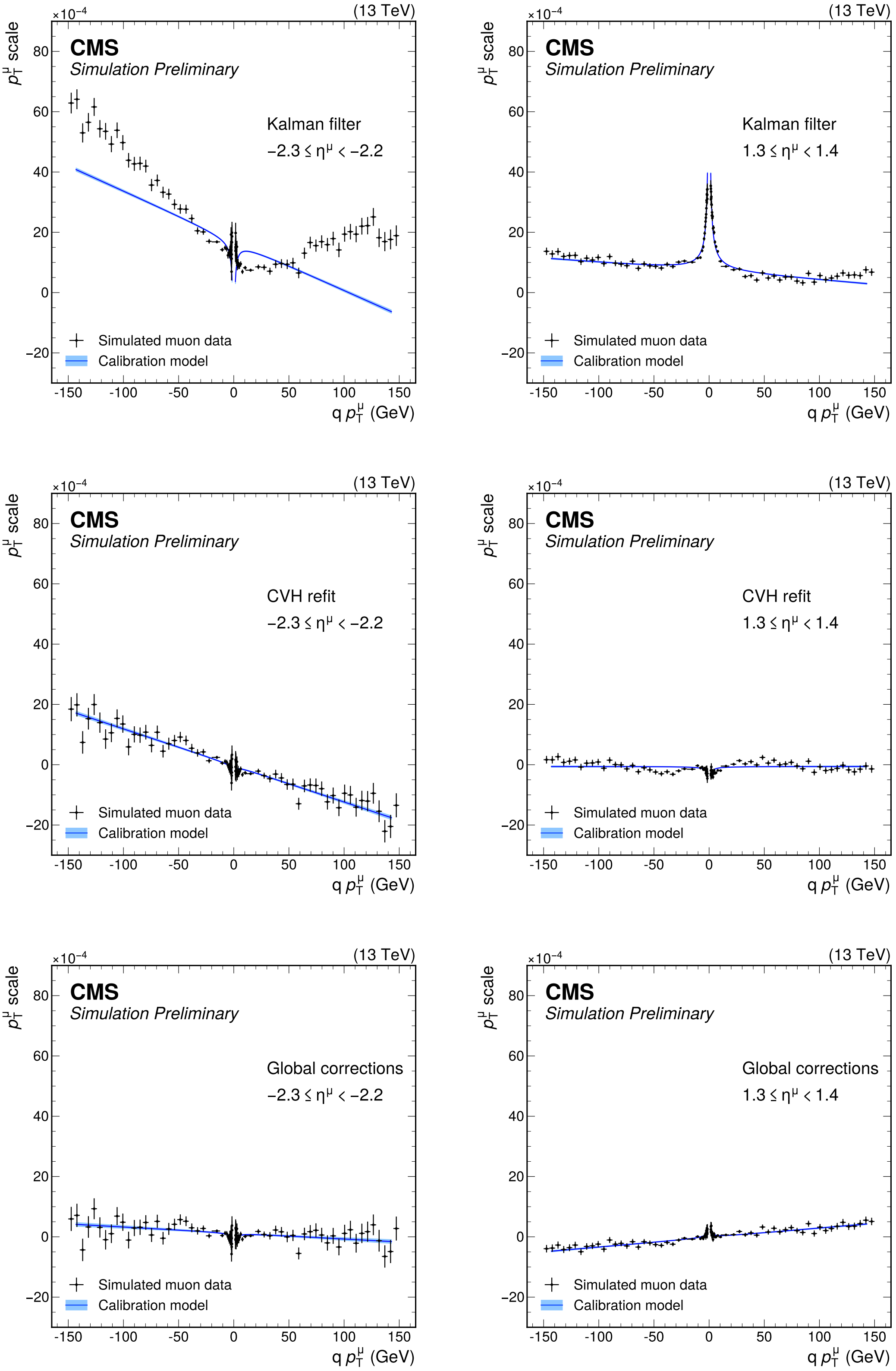


Figure 7.61: Validation of the formula in equation 7.13. First row: p_T^μ scale evaluated using the standard CMS reconstruction. Second row: p_T^μ scale evaluated using p_T^μ from the CVH refit. Third row: p_T^μ scale evaluated using the p_T^μ from the CVH refit after the modular corrections described in the text are applied. Each plot is shown for the two η bins indicated in the figure, and a fit using the model in equation 7.13 is superimposed. From [1].

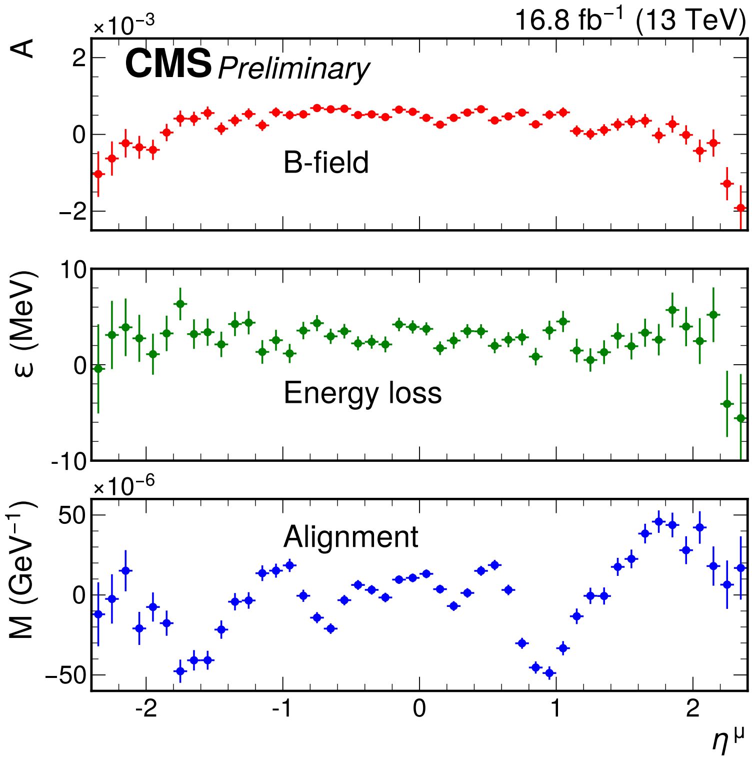


Figure 7.62: Correction terms in equation 7.13 derived from data. From [1].

account for these differences, the statistical uncertainty in the J/Ψ calibration parameters is scaled by a factor 2.1, determined so that the χ^2 in the right of figure 7.63 is equal to the number of degrees of freedom. Biases in the alignment should cancel out when both W^+ and W^- are considered together, as they have opposite contributions to the reconstructed p_T, as per 7.13, so that a negligible bias in the measured value of m_W should be induced. This is studied in more detail in section 7.10.

The muon momentum resolution is calibrated using both $J/\Psi \to \mu\mu$ and $Z \to \mu\mu$ decays, by fitting their invariant mass distributions as a function of the p_T and η of the positive and negative muon from the decay, after the corrections to the muon momentum scale are applied. The resolution is parameterized as:

$$\frac{\sigma_{i\eta}^2}{k^2} = a_{i\eta}^2 + \frac{c_{i\eta}^2}{k^2} + \frac{b_{i\eta}^2}{1 + d_{i\eta}^2 k^2}, \tag{7.21}$$

for 24 uniform η bins from -2.4 to 2.4, where the parameters $a_{i\eta}$, $c_{i\eta}$, $b_{i\eta}$, and $d_{i\eta}$ parameterize the contributions to the curvature resolution from multiple scattering, hit resolution, and the correlations between them induced by the track fit. These resolution corrections are applied to simulation by smearing the reconstructed curvature of the simulated reconstructed muons with a Gaussian distribution of width given by the difference in quadrature between the resolution measured in data and simulation. The relative agreement between the measured and simulated samples is mostly affected by the different number of pixel hits in the data and simulation after the quality requirements imposed by the CVH fit. A systematic uncertainty is evaluated by reweighting the pixel hit multiplicity distribution in Monte Carlo to match the one observed in data in each η bin, taking the full difference as an uncertainty. This also affects the nominal resolution corrections that are derived: their statistical uncertainty is inflated by a factor of 10 to cover the observed differences.

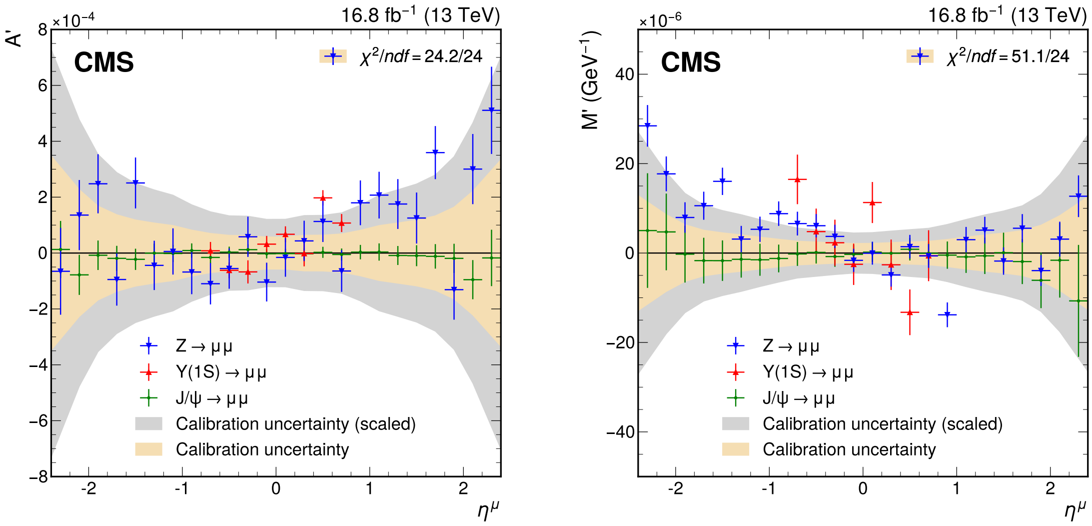


Figure 7.63: Charge-independent (A', left) and charge-dependent (M', right) residual scale differences using $J/\Psi \to \mu\mu$, $\Upsilon(1S) \to \mu\mu$, and $Z \to \mu\mu$ events. The points with error bars represent the scale parameters and statistical uncertainties associated with the closure test performed with J/Ψ (green), $Y(1S)$ (red), and Z (blue) events. The yellow band represents the corresponding statistical uncertainty in the calibration parameters derived from the J/Ψ calibration sample. The gray band shows the statistical uncertainty scaled by a factor of 2.1. The χ^2 values correspond to the compatibility of the additional scale parameters with zero for the closure test performed with $Z \to$ events and take into account the statistical uncertainties for these parameters, as well as for the calibration parameters derived from the J/Ψ sample, without additional scaling. From [1].

Table 7.10 summarizes the muon scale and resolution calibration model and impact of the various groups on m_W. The largest source of uncertainty comes from the statistical uncertainty on the coefficients in 7.13 due to the finite statistics of the J/Ψ sample, scaled as explained above. This is implemented as a total of 144 nuisance parameters, which corresponds to three nuisance parameters per η bin, by diagonalizing the covariance matrix, so that each nuisance parameter corresponds to a statistically independent variation of the $p_T - \eta$ templates. A similar strategy is implemented for the statistical uncertainty on the closure test on $Z \to \mu\mu$ events: since the additional term for energy loss is not determined, and since the validation was done on 24 η bins instead of 48, a total of 48 nuisance parameters is assigned, each corresponding to an independent variation as in the J/Ψ case. One fully correlated nuisance parameter to take into account the finite uncertainty on the measured Z boson mass is also assigned. When the Z information is used to define the uncertainties, the central value of the Z boson mass is not used. A similar procedure is followed for the statistical uncertainty in the resolution. 49 nuisance parameters (48 uncorrelated and fully correlated across η) are assigned to the impact on the resolution of the pixel multiplicity. The total impact on m_W is 4.8 MeV, if the nominal breakdown of uncertainties is used.

7.6 Recoil calibration

The recoil is only used in the analysis to define the event selection and sidebands to estimate the nonprompt background (described in section 7.7) through a m_T cut. As already mentioned, p_T^{miss} is reconstructed using the `DeepMET` algorithm [216, 217] that uses machine-learning techniques to improve the resolution on the reconstructed p_T^{miss}, particularly in determining the contribution from the neutral PF can-

Source of uncertainty	Nuisance parameters	Impact on m_W [MeV]
J/Ψ calibration stat. (scaled $\times$2.1)	144	3.7
Z closure stat.	48	1.0
Z closure (LEP measurement)	1	1.7
Resolution stat.	72	1.4
Pixel multiplicity	49	0.7
Total	314	4.8

Table 7.10: Breakdown of muon scale and resolution calibration uncertainties. From [1].

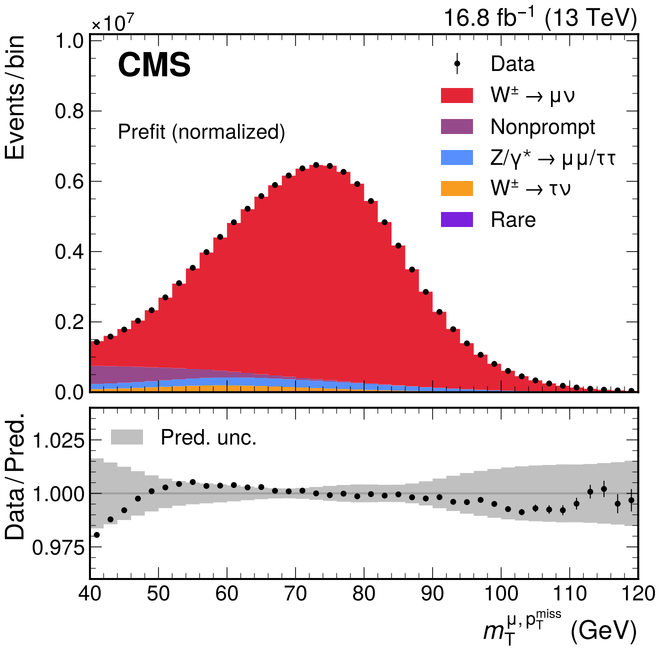


Figure 7.64: Measured and predicted m_T distributions in $W \to \mu\nu$ events, after calibrating the hadronic recoil. The total uncertainties (statistical and systematic) are represented by the gray band and the simulation is scaled to the measured distribution in data. From [1].

didates.

The recoil has been calibrated with a procedure similar to that used by the ATLAS experiment from $Z \to \mu\mu$ events: the templates for the parallel and perpendicular components of the hadronic recoil are parameterized as a function of the reconstructed dimuon transverse momentum ($p_T^{\mu\mu}$), and a new missing transverse momentum value (p_T^{miss}) is calculated by applying an inverse cumulative distribution function transformation that maps the simulated templates onto the data. Corrections obtained from $Z \to \mu\mu$ events are then applied to simulated $W \to \mu\nu$ events by using the generator-level W boson transverse momentum (p_T^W). Any differences in recoil modelling between W and Z bosons are incorporated as systematic uncertainties. Statistical uncertainties in the template parameterization are included, but they are found to have a negligible effect on the m_W measurement, since is based on p_T^μ, and m_T only affects the amount of events passing or failing the selection, so that uncertainties on it have a second-order effect on p_T^μ.

Figure 7.64 shows the m_T distribution for data and Monte Carlo events that pass the $W \to \mu\nu$ selection, where the Monte Carlo is normalized to the data. A good agreement within the uncertainties is found, indicating that the recoil scale and resolution are corrected with subpercent precision.

The uncertainty in the corrections is determined from the statistical uncertainties of the fits that parameterize the correction. Their effect on m_W is assessed in terms of the eigenvectors of the covariance matrix of the fit to the recoil distributions. It was verified that these variations contribute less than 0.3 MeV to the overall uncertainty in m_W. Since assessing these variations requires significant computational resources and their impact on m_W is negligible, they were not included in the nominal fit configuration.

7.7 Background estimate

Several electroweak backgrounds give small contributions to the m_W analysis. The amount of events that pass the final analysis selection are determined from Monte Carlo simulations:

- W or Z decays to τ leptons: generated using POWHEG MINNLO$_{\text{PS}}$ interfaced to PYTHIA8, with the same theory corrections on the boson production kinematic distributions as those applied in the muon decays.
- EW production of lepton pairs or of a W boson in association with a quark through photon-photon or photon-quark scattering. Simulated at leading order using PYTHIA 8.240
- top quark production: simulated at next-to-leading order QCD accuracy using MADGRAPH5_aMC@NLO v2.6.5 [260], interfaced to PYTHIA 8.240
- diboson production: simulated at next-to-leading order QCD accuracy using POWHEG-BOX-V2, interfaced with PYTHIA 8.240
- quarkonia production: simulated using PYTHIA 8 interfaced with PHOTOS++ v3.61 for final-state photon radiation (only used for the muon scale and resolution calibration).

For each of these processes, all the experimental and theoretical corrections, described in the previous sections, are applied, and the relevant uncertainties are propagated accordingly.

Another important source of background, the contribution of which cannot be easily modelled via Monte Carlo simulations, is the nonprompt background where muons originate from decays of heavy-flavor hadrons. Although suppressed by the requirements on the impact parameter with respect to the beamspot and isolation variable (and m_T), there is still a significant component that passes the final analysis selection. This contribution is evaluated by defining sidebands enriched in events from nonprompt background. This is done by inverting the requirement on the transverse mass m_T, the relative isolation defined in 5.1, or both.

The QCD background is estimated using a refined version of the $ABCD$ method: defining D the signal region, i.e. the one in which events pass the standard analysis selection, and B and C sidebands obtained by reversing the requirements on one of the two variables considered to define regions enriched in the background of interest, and A the sideband obtained by reversing the requirements on both variables, the amount of background in the signal region D is obtained as:

$$N^D_{\text{bkg}} = N^C_{\text{bkg}} \frac{N^B_{\text{bkg}}}{N^A_{\text{bkg}}}, \tag{7.22}$$

where N^i_{bkg} is the number of events due to the background of interest in the region i. For the m_W measurement by CMS, the two variables are the relative isolation computed from PF candidates used in the analysis and the transverse mass m_T. In estimating these numbers, the contamination of processes different from the background of interest need to be subtracted from data (which is the reason why it was necessary to derive the anti-Isolation efficiency corrections as described in 7.4).

This implementation of the $ABCD$ method gives a reliable estimate of the background contamination in the signal region D only if the two variables used to define the sideband are independent. The refined approach used by CMS is called "extended $ABCD$ method" [261]: in this case, the region of events with

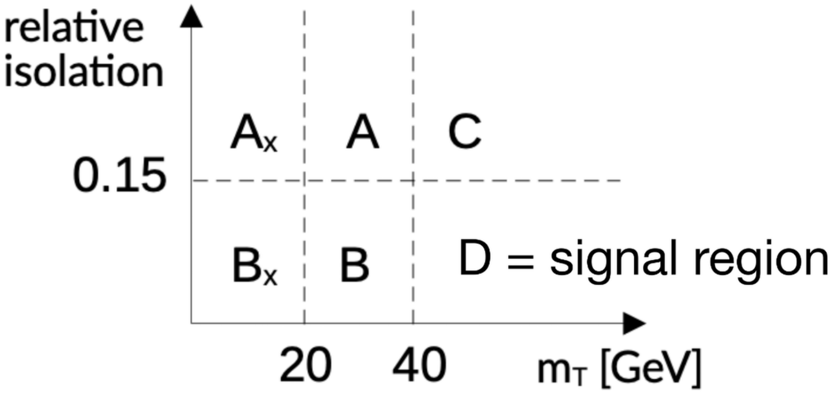


Figure 7.65: Sketch that shows the different sidebands and the signal region used in the "extended $ABCD$ method". `https://indico.cern.ch/event/1441575/`.

$m_T < 40$ GeV is divided in two bins with events with $m_T < 20$ GeV and 20 GeV$\leq m_T < 40$ GeV, while the same splitting regarding the relative isolation is used as in the standard $ABCD$ method. The number of background events in the signal region D isthen estimated through:

$$N^D_{\mathrm{bkg}} = N^C_{\mathrm{bkg}} \left(\frac{N^B_{\mathrm{bkg}}}{N^A_{\mathrm{bkg}}}\right)^2 \frac{N^{Ax}_{\mathrm{bkg}}}{N^{Bx}_{\mathrm{bkg}}}, \tag{7.23}$$

where the regions are defined as in figure 7.65. Equation 7.23 differs from equation 7.22 for the extra

$$\left(\frac{N^B_{\mathrm{bkg}}}{N^A_{\mathrm{bkg}}}\right) / \left(\frac{N^{Bx}_{\mathrm{bkg}}}{N^{Ax}_{\mathrm{bkg}}}\right) \tag{7.24}$$

term, which is used to correct for the correlation between the two variables used in the background estimation. This is done independently in bins of $p_T - \eta - q$.

To suppress statistical fluctuations in the background estimation, the QCD nonprompt background contribution in the signal region is fitted with an analytic function of p_T^μ, chosen to be an exponential of a third-order Chebyshev polynomial, independently for each bin of η and charge. The statistical uncertainties of the data are accounted for by propagating the uncertainties in the function parameters in the signal region. This procedure results in 384 variations, reflecting the four coefficients of the polynomials, the two charges, and the 48 η^μ bins. Since the contributions from the prompt processes to the sideband regions are evaluated from simulation with all the experimental and theoretical corrections properly applied, their uncertainties are propagated to the sideband regions by repeating the subtraction of the prompt component and the determination of the smoothing parameters in the sideband regions for each variation.

This procedure has been validated on a QCD Monte Carlo simulation (simulated muons from b or c hadron decays) and in an enriched region in data obtained by requiring that the muons originate from a secondary vertex. The background estimation obtained using the extended $ABCD$ method is compared with the actual number of events in the signal region: in both the QCD Monte Carlo simulation and the data control region, an additional scaling factor of 0.85 was necessary to reproduce the number of observed events in the signal region. Residual shape differences are accounted for by varying the coefficients of the smoothing polynomials. These two additional variations are fully correlated across η^μ and charge.

As a validation of the procedure to evaluate the QCD nonprompt background, figure 7.66 shows the pre-fit p_T^μ and η^μ distribution for events that pass the full m_W analysis selection and the additional requirement that the muon originates from a secondary vertex. It can be seen that the data and the background prediction (together with a small contamination from prompt processes) agree within uncertainties, within

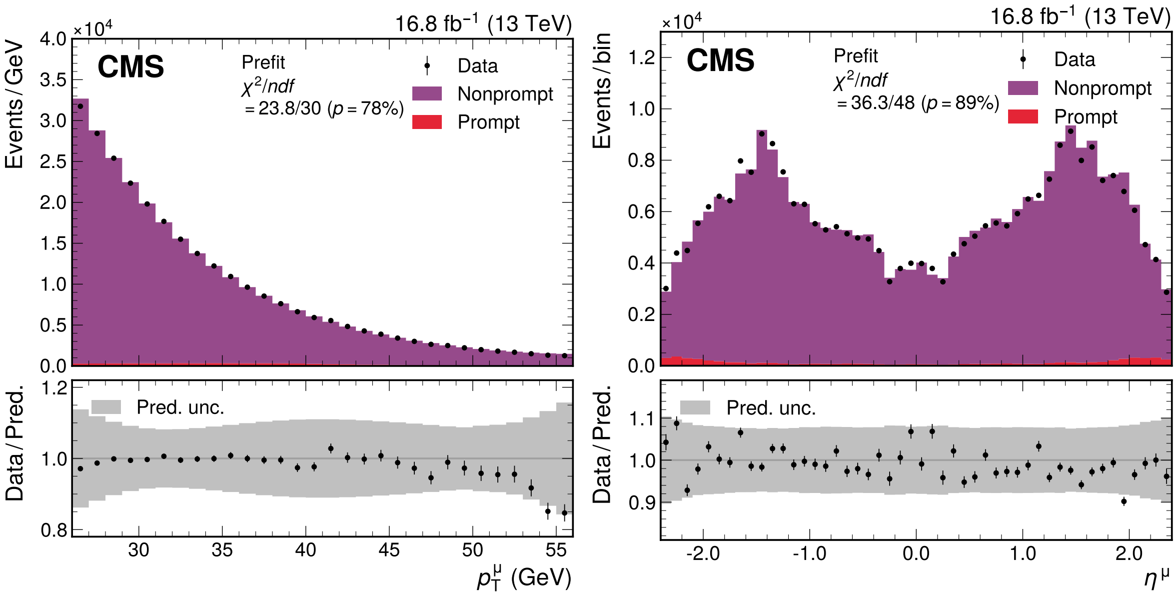


Figure 7.66: The observed data and the pre-fit prediction of the extended $ABCD$ method, for the p_T^μ (left) and η^μ (right) distributions, in a region obtained by selecting muons compatible with being produced in a secondary vertex. Small contributions from events with a prompt muon are shown by the red histogram. The total uncertainties (statistical and systematic) are represented by the gray bands. From [1].

an agreement at 2% between the number of events in data and background prediction: a normalization uncertainty of 5% is assigned to cover for this difference. A likelihood fit with the uncertainty model used in the m_W analysis was performed, leading to a good agreement between data and prediction and a $p-$value of 98%, serving as a validation of the QCD nonprompt background model.

The total impact on m_W from the QCD nonprompt background estimate is 3.2 MeV, in which 2.5 MeV are due to the normalization uncertainty and systematic variations of the coefficients to cover for differences in shape, and 1.9 MeV are due to the finite amount of statistics. This uncertainty does not include the contribution on the QCD nonprompt estimate from the theoretical and experimental correction uncertainties, that are included in the corresponding uncertainties in m_W from the respective sources.

7.8 Ancillary measurements

The muon scale and resolution calibration procedure and the uncertainty model described in 7.5 are validated by performing a likelihood fit to the $m_{\mu\mu}$ distribution on $Z \to \mu\mu$ events to extract m_Z, where all the systematic uncertainties that enter in the W-like fit of m_Z are considered. The pre-fit and post-fit distributions are shown in figure 7.67. The post-fit distribution in simulation is in agreement with data. In the pre-fit distribution, the effect of the different resolution in data and simulation can be seen, in particular the effect of the pixel multiplicity in the tails.

The m_Z value extracted from this fit is:

$$m_{\mu\mu}^{Z,\mathrm{CMS}} - m_Z^{\mathrm{PDG}} = -2.2 \pm 4.8\ \text{MeV}, \tag{7.25}$$

where the total uncertainty is dominated by scale and resolution calibration (4.6 MeV). The agreement with m_Z value from PDG serves as a validation of the calibration procedure. However, this result cannot be interpreted as an independent measurement of the mass of the Z boson, as the $Z \to \mu\mu$ data have been used to tune the scale calibration uncertainty model. In addition to this, the standard CMS alignment proce-

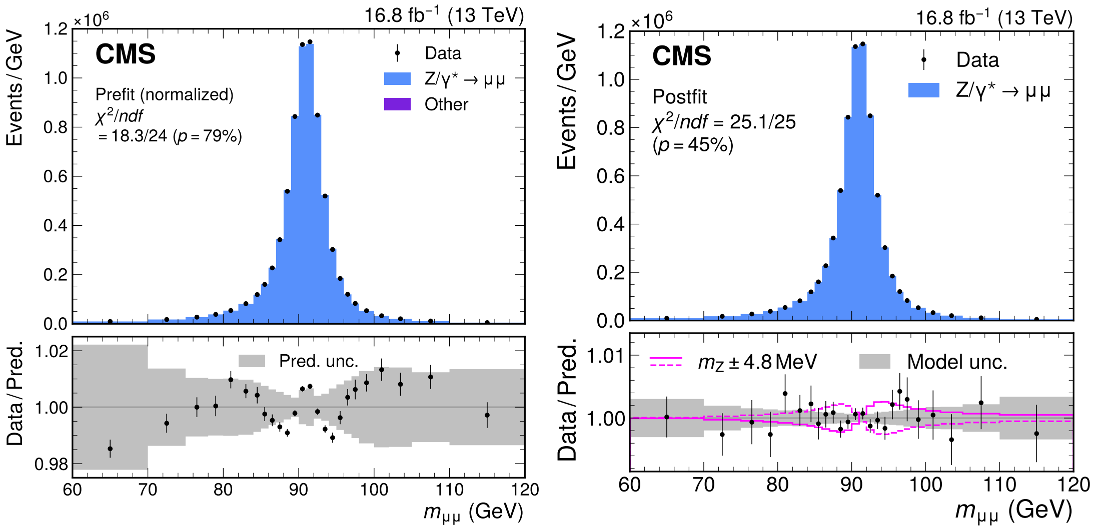


Figure 7.67: Measured and simulated $Z \to \mu\mu$ dimuon mass distributions, after applying the muon momentum scale and resolution corrections, before (left) and after (right) the fit to extract m_Z to validate the calibration procedure. For the pre-fit distribution, the simulated predictions and uncertainties are scaled to match the number of observed data events. From [1].

dure already uses $Z \to \mu\mu$ data and the measurement of m_Z, so that also our calibration procedure might depend, at least in part, on m_Z.

As a validation of the theory model described in section 7.3 and the procedure to tune it from data, m_Z is extracted in the W-like fit described in section 7.2, where m_Z is extracted together with all other nuisance parameters from the $p_T^\mu - \eta^\mu - q^\mu$ distribution.

The post-fit nuisance parameters relative to the theory model and their uncertainties are used to determine the postFSR generator level p_T^Z distribution. This is compared to the measured $p_T^{\mu\mu}$ spectrum unfolded at the postFSR generator level, and to the postFSR generator level p_T^Z distribution as determined from post-fit values of the nuisance parameters as determined from the fit to the $p_T^{\mu\mu} - y^{\mu\mu}$ distribution. The results of this comparison are shown in figure 7.68, where the pre-fit predictions are also shown. The post-fit predictions are in agreement with the unfolded data in both fits. The uncertainties at low p_T, the region most impacting on the p_T^μ spectrum, are significantly reduced in both cases, particularly for the $p_T^{\mu\mu} - y^{\mu\mu}$ spectrum. This gives us confidence that the model is accurate enough to reproduce the p_T^V distribution and that p_T^W can be accurately determined likewise.The total impact on m_W from the p_T^V modelling is 2.0 MeV.

As a cross-check, m_W was also extracted from the simultaneous $\eta^\mu - p_T^\mu$ spectrum on $W \to \mu\nu$ events and the $p_T^{\mu\mu} - y^{\mu\mu}$ spectrum on $Z \to \mu\mu$ events: the shift with respect to the nominal m_W value is $+0.6$ MeV, with a decrease in the total uncertainty of 0.3 MeV. The nominal result, however, does not use the information from $Z \to \mu\mu$ events, to avoid the choice of a correlation scheme between the Z and W bosons.

The difference between the value of m_Z extracted from the W-like fit and the experimental world average determined by PDG is:

$$\begin{aligned} m_Z^{W\text{-like}} - m_Z^{\text{PDG}} &= -6 \pm 7_{\text{stat}} \pm 12_{\text{syst}}\ \text{MeV} \\ &= -6 \pm 14\ \text{MeV}, \end{aligned} \tag{7.26}$$

providing further validation of the theoretical model and experimental calibrations used in the analysis. Since events are split into according to their parity, the fit was also repeated by reversing the splitting logic, obtaining a value of $\Delta m_Z = 8 \pm 14$ MeV. Figure 7.69 shows the post-fit p_T^μ distribution from the W-like fit, with all the uncertainties and a 14 MeV variation superimposed in the ratio plot, showing that the agreement

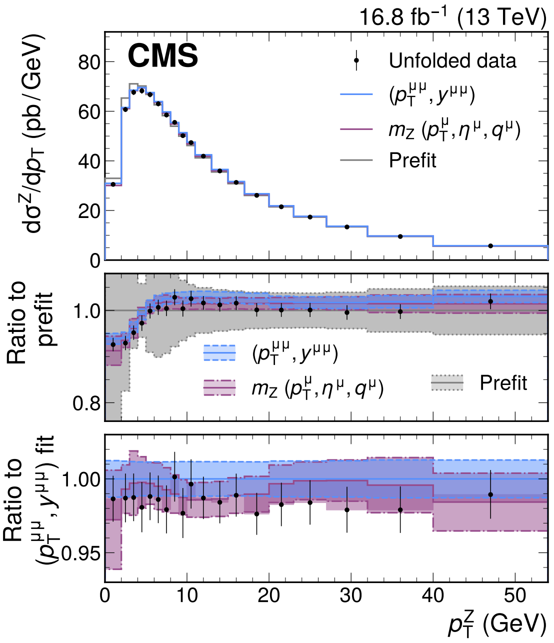


Figure 7.68: Validation of the theory model. Unfolded measured p_T^Z distribution (points) compared with the generator-level SCETlib+MiNNLO$_{\text{PS}}$ predictions before (pre-fit, gray) and after adjusting the nuisance parameters to the best fit values obtained from the W-like m_Z fit (magenta) or from the fit to the $p_T^{\mu\mu} - y^{\mu\mu}$ distribution (blue). The center panel shows the ratio of the post-fit predictions and the unfolded data to the pre-fit prediction, while the bottom panel shows the ratio of the predictions and unfolded data to the prediction adjusted to the best fit values obtained from the fit to the $p_T^{\mu\mu} - y^{\mu\mu}$ distribution. The uncertainty in the pre-fit prediction is shown by the shaded gray area. The uncertainties in the predictions after the maximum likelihood fits are shown in the shaded magenta (W-like fit) and blue ($p_T^{\mu\mu} - y^{\mu\mu}$) bands. From [1].

between data and simulation is good, as also indicated by the χ^2 and $p-$value shown.

7.9 Fit to the $p_T - \eta$ spectrum

The $p_T - \eta$ spectrum is fitted minimizing the negative log-likelihood given the data spectrum. Both m_W and the nuisance parameters relative to the systematic uncertainties are extracted simultaneously. The function being minimized is:

$$\begin{aligned} -\ln\mathcal{L}\left(m_W, \vec{\theta}|\vec{n}\right) = & \sum_{i=0}^{\text{Nbins - 1}} \left(\mu_i\left(m_W, \vec{\theta}\right) - n_i \ln\left(\mu_i\left(m_W, \vec{\theta}\right)\right)\right) \\ & + \frac{1}{2}\sum_{j=0}^{\text{Nnuisances - 1}} \left(\frac{\hat{\theta}_j - \theta_j}{\sigma_j}\right)^2 \\ & + \sum_{i=0}^{\text{Nbins - 1}} \left(n_i^{\text{MCstat}}\beta_i - n_i^{\text{MCstat}}\beta_i^0 \ln\beta_i\right). \end{aligned} \tag{7.27}$$

The first term comes from the product of the Poisson distributions of mean μ_i for each of the $p_T - \eta$ bins in the analysis (the n_i are the number of events in each bin). The second term contains all the nuisance parameters to be determined from the fit θ_j and their best estimation $\hat{\theta}_j$ and the uncertainty on their estima-

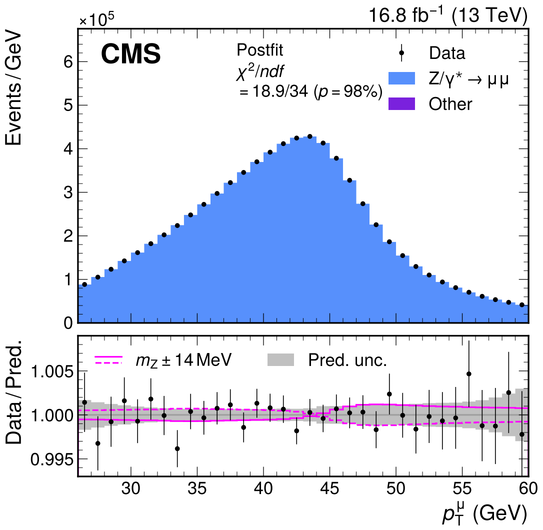


Figure 7.69: Post-fit measured and simulated p_T^μ distributions from the maximum likelihood fit of the W-like m_Z analysis. The solid and dashed purple lines represent, respectively, the relative impact of a change in m_Z by 14 MeV. The uncertainties in the predictions, after the systematic uncertainty profiling in the maximum likelihood fit, are shown by the shaded band. From [1].

tion σ_j. The $\hat{\theta}_j$ estimates are obtained from dedicated external measurements in the case of experimental systematic uncertainties[12], or from variations in the theoretical calculations in the case of the theoretical uncertainties. The $\vec{\theta}$ vector, which contains all the nuisance parameters in the second term, has been introduced to shorten the notation. The third term implements the statistical uncertainty due to the finite size of the Monte Carlo sample used, in the Barlow-Beeston method [224], as simplified by Conway [225], with $n_i^{\text{MCstat}} = \frac{(\sum w)^2}{\sum w^2}$ where w is the event weight, and β_i is the parameter fit to model the statistical fluctuations on Monte Carlo[13].

In the second term, the nuisance parameters are redefined so that $\theta_j = 0$ corresponds to the nominal template (with the nominal configuration for the source of systematic uncertainty considered) and $\theta_j = \pm 1$ corresponds to a $\pm 1\sigma$ variation, so that in practice the second term can be simplified to $\frac{1}{2}\sum_{j=0}^{\text{Nnuisances - 1}} \theta_j^2$. With this simplification, the relationship between the expected number of events μ_i and the systematic variations expressed in terms of the nuisance parameters θ_k is the following:

$$\mu_i\left(m_W, \vec{\theta}\right) = \sum_{j=0}^{\text{Nprocess - 1}} \beta_i n_{i,j}^0 \prod_{k=0}^{\text{Nsyst - 1}} \kappa_{i,j,k}^{\theta_k}, \qquad \text{where } \kappa_{i,j,k} = \frac{n_{i,j}^0 + v_{i,j,k}}{n_{i,j}^0} \tag{7.28}$$

with $n_{i,j}^0$ corresponding to the number of events in the $i-$th bin for the $j-$th process, $v_{i,j,k}$ corresponding to a 1σ the variation in the number of events of the $j-$th process in the $i-$th bin due to the k-th systematic uncertainty[14].

---

[12] In the case of experimental uncertainties, $\hat{\theta}_j$ is the maximum likelihood estimator from θ_j from a dedicated set of measurements, and σ_j is the standard deviation obtained in the estimation. In principle, the second term in 7.27 could be replaced with the equivalent of the first one for the measurements to determine θ_j, and this would result in a direct simultaneous estimation of the nuisance parameter θ_j.

[13] β_i^0 is the parameter corresponding to the nominal template. β_i are normalized so that $\beta_i = 1$ corresponds to having the nominal event yield, and $\beta_i^0 = 1$ in most cases, it can be different when Monte Carlo toys are generated

[14] A $+1\sigma$ variation in a single systematic uncertainty k corresponds to $\mu_i = \sum_{j=0}^{\text{Nprocess - 1}} \beta_i n_{i,j}^0 \kappa_{i,j,k}$, while a -1σ variation corresponds to $\mu_i = \sum_{j=0}^{\text{Nprocess - 1}} \beta_i n_{i,j}^0 \frac{1}{\kappa_{i,j,k}}$

The $v_{i,j,k}$ values are derived by producing alternate $p_T^\mu - \eta^\mu$ histograms for each variation corresponding to a 1σ variation and then taking the difference with respect to the nominal histogram, while equation 7.28 is used to interpolate and extrapolate for $\theta_k \neq \pm 1$.

All variations are symmetrized assuming that the logarithm of the yield for the original asymmetric variation varies quadratically as a function of the nuisance parameter for the up and down variations at values of ± 1, and then two symmetric variations are computed: one is the average of the up and down variations, and the second is a residual contribution needed to produce the equivalent variance for the yields as the original asymmetric variation. This is done to avoid spurious contributions to the Hessian matrix that would otherwise degrade the accuracy of the Hessian approximation for determining statistical uncertainties. In cases in which the variations are approximately (or exactly) symmetrical, the second term is not evaluated.

The results, particularly the extracted value of m_W, are also cross-checked by modifying the relation in equation 7.28 to be purely linear as a function of θ_k and by running a minimum χ^2 fit (using this linear parameterization).

m_W is fitted together with the other parameters, and the impact of changes in m_W on the spectrum are implemented using the same formalism as in equation 7.28, but no quadratic term as those for the nuisance parameters is present in the total likelihood in equation 7.27. In this case, the $\pm 1\sigma$ variations corresponding to varying m_W by ± 100 MeV: the stability and accuracy of this procedure was checked by running fits on pseudodata obtained by varying m_W by specified amounts, obtaining post-fit values of m_W in agreement with the input values within 0.1 MeV.

The construction and minimization of the likelihood is implemented using the COMBINETF framework (now RABBIT [262]) based on the TENSORFLOW software package [248], which allows the m_W and W-like m_Z likelihood fits to be computationally feasible and numerically stable with the number of bins and parameters in this analysis. Table 7.11 summarizes all the systematic uncertainties, with the total number of nuisance parameters associated to them, for the W-like m_Z and m_W fits, the majority of which have already been discussed in the previous sections. Other uncertainties come from the integrated luminosity (one nuisance parameter fully correlated across all processes, properly propagated to the nonprompt background), uncertainties on the cross-sections for prompt background processes[15] described in 7.7, the widths of the W and Z boson, the mass of the Z boson (only for the m_W analysis).

7.10 m_W extraction

Figure 7.70 shows the p_T^μ distributions in data and simulation after the fit on the $p_T^\mu - \eta^\mu - q^\mu$ spectrum described in 7.9 has been run (the parameters of the simulation are tuned to the post-fit values, and the post-fit covariance matrix is used to evaluated the uncertainties). The value of m_W extracted by the CMS experiment is:

$$m_W^{\text{CMS}} = 80360.2 \pm 9.9 \text{ MeV}, \tag{7.29}$$

in agreement with the Standard Model prediction and the other experimental measurements, except for the one from CDF. This is summarized in figure 7.71.

Table 7.12 summarizes the impacts on m_W and m_Z (in the W-like presented in 7.8) in both the nominal and global definitions discussed when presenting the ATLAS measurement in 6.3.2. The measurements are dominated by statistics, as can be seen in the global impacts, particularly for the W-like m_Z measurement.

Concerning theoretical uncertainties, it can seen that the PDFs have the largest impact on m_W, with an impact of 4.4 MeV in the nominal definition, while this decreases to 2.8 MeV when the global definition

---

[15] Some conservative assumptions are made, and different groups of backgrounds are considered together, for the top background here actually includes $t\bar{t}$ or single top production, but since the contaminations from these backgrounds are small this choice of uncertainties should have only a small impact on m_W or m_Z.

<table>
<tr><th>Systematic uncertainties</th><th>W-like m_Z</th><th>m_W</th></tr>
<tr><td>Muon efficiency</td><td>3127</td><td>3658</td></tr>
<tr><td>Muon eff. veto</td><td>-</td><td>531</td></tr>
<tr><td>Muon eff. syst.</td><td colspan="2">343</td></tr>
<tr><td>Muon eff. stat.</td><td colspan="2">2784</td></tr>
<tr><td>Nonprompt background</td><td>-</td><td>387</td></tr>
<tr><td>Prompt background</td><td>2</td><td>3</td></tr>
<tr><td>Muon momentum scale</td><td colspan="2">314</td></tr>
<tr><td>L1 prefiring</td><td colspan="2">14</td></tr>
<tr><td>Integrated luminosity</td><td colspan="2">1</td></tr>
<tr><td>PDF (CT18Z)</td><td colspan="2">60</td></tr>
<tr><td>Angular coefficients</td><td>177</td><td>353</td></tr>
<tr><td>W MiNNLO μ_F, μ_R</td><td>-</td><td>176</td></tr>
<tr><td>Z MiNNLO μ_F, μ_R</td><td colspan="2">176</td></tr>
<tr><td>PYTHIA shower k_T</td><td colspan="2">1</td></tr>
<tr><td>p_T^V modeling</td><td>22</td><td>32</td></tr>
<tr><td>Non-perturbative</td><td>4</td><td>10</td></tr>
<tr><td>Perturbative</td><td>4</td><td>8</td></tr>
<tr><td>Theory nuisance parameters</td><td colspan="2">10</td></tr>
<tr><td>q_c, q_b quark mass</td><td colspan="2">4</td></tr>
<tr><td>Higher-order EW</td><td>6</td><td>7</td></tr>
<tr><td>Z boson width</td><td colspan="2">1</td></tr>
<tr><td>Z boson mass</td><td>-</td><td>1</td></tr>
<tr><td>W boson width</td><td>-</td><td>1</td></tr>
<tr><td>$\sin^2\theta_W$</td><td colspan="2">1</td></tr>
<tr><td>Total</td><td>3725</td><td>4833</td></tr>
</table>

Table 7.11: Number of nuisance parameters for the main groups of systematic uncertainties, for the W-like m_Z and m_W fits. Subgroups of parameters are also reported as indented labels for a few groups.

is used. This is a consequence of the fact that the nuisance parameters related to the PDFs are being determined in data with an uncertainty that is smaller than the pre-fit one, in agreement with the conclusions made in the CMS measurement of the helicity and rapidity of the W boson [212].

The p_T^V modelling has an impact of 2.0 MeV at most. Regarding the experimental uncertainties, the main source is the determination of the muon momentum scale: the impact is 4.8 MeV in the nominal definition and 4.4 MeV in the global definition. The similarity of the impact between the two definitions suggests that this source of uncertainty is not being significantly constrained in data, and this is to be expected given that this source of uncertainty modifies the $p_T^\mu - \eta^\mu - q^\mu$ distribution similarly to a m_W variation. The uncertainties in the estimate of the nonprompt background have an impact of 3.2 MeV on m_W.

Multiple tests were run to check the consistency of the extracted value of m_W under different variations. m_W was also extracted for each of the PDF sets mentioned in 7.3.2, both with and without the inflation of their uncertainties. The results are shown in table 7.13 and figure 7.72. When the uncertainties are inflated, all of the extracted values of m_W are compatible with the PDF uncertainties indicated in table 7.4. With the standard uncertainties, the extracted value when NNPDF4.0 is incompatible within PDF uncertainties with the other values (although the shift is still within the total uncertainty on m_W). This does not necessarily indicate that the PDF uncertainties on NNPDF4.0 are underestimated, but simply that they do not cover the differences in the extracted m_W value with the other PDF sets. It is also interesting to note that when the NNPDF4.0 set is used both the contribution from the PDFs and the total uncertainty are the smallest when the uncertainties are not inflated and are the largest when they are.

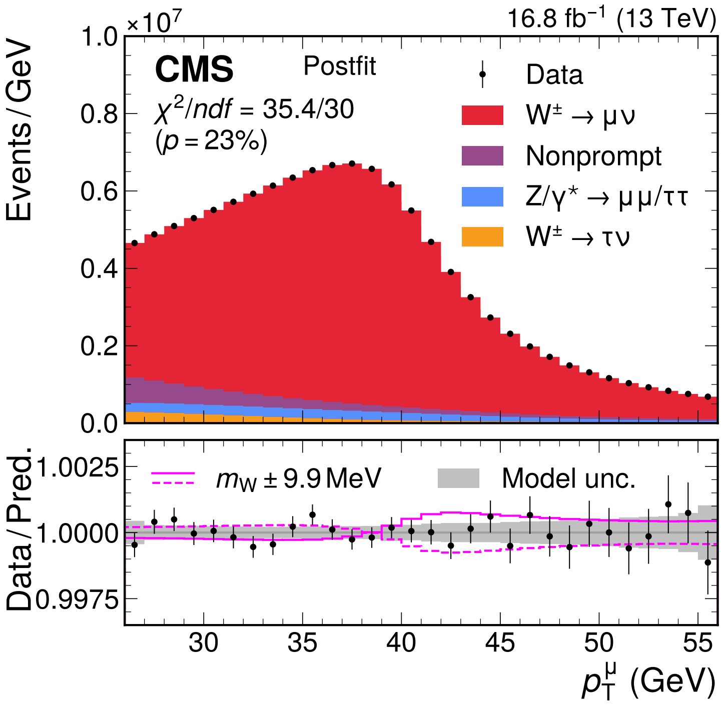


Figure 7.70: The W boson mass measurement. Measured and post-fit predicted p_T^μ distributions. The predicted $W \to \mu\nu$ contribution is shown in red, while the background contributions are shown in blue, purple, and yellow (with additional backgrounds giving much smaller contributions, not visible in the figure). The lines in magenta in the ratio plot shows the impact on the p_T^μ distribution of a variation of m_W of 9.9 MeV. The total contribution of all theoretical and experimental uncertainties in the predictions, after the systematic uncertainty profiling in the maximum likelihood fit, is shown by the gray band. From [1].

The W-like m_Z and m_W measurements were repeated using predictions from SCETLIB matched to DYTURBO with the different approaches to incorporate the TNPs described in 7.3.1. When using N^{3+1}LL and N^{4+0}LL predictions and uncertainties, the measured value of m_W is shifted by less than 0.5 MeV, within the 2.0 MeV p_T^W −modeling uncertainty of the nominal result.

$Z \to \mu\mu$ data has also been incorporated in the m_W for cross-checks. As a first test, the p_T^W predictions from SCETLIB+DYTURBO are corrected using the ratio between the unfolded $p_T^{\mu\mu}$ and fitted p_T^Z spectra, and no change is made on the uncertainty model. This results in a shift in m_W of less than 0.5 MeV. As a second test, m_W is extracted from the simultaneous fit to the $p_T^\mu - \eta^\mu - q^\mu$ spectrum from $W \to \mu\nu$ events and the $p_T^{\mu\mu} - y^{\mu\mu}$ spectrum in $Z \to \mu\mu$ events. In this fit, the TNPs and the CS anomalous rapidity are correlated, while the other p_T^V uncertainties are treated as uncorrelated. This results in a shift of $+0.6$ MeV in m_W, with a moderate reduction in the total uncertainty. However, this is not used as the nominal result as assumptions in the correlation between W and Z processes are made, and the nominal result is expected to be more robust with respect to them (this is true also for the first test, in which it is assumed that the $p_T^{\mu\mu}/p_T^Z$ ratio can be used directly to correct p_T^W, assuming an even larger degree of correlation, but since these tests are only used as cross-checks, their results should still hold).

Figure 7.73 shows the comparison between the pre-fit and post-fit generator level p_T^W spectra, extracted from the standard fit to extract m_W from the $p_T^\mu - \eta^\mu - q^\mu$ spectrum and from the simultaneous fit to $p_T^\mu - \eta^\mu - q^\mu$ in $W \to \mu\nu$ events and $p_T^{\mu\mu} - y^{\mu\mu}$ in $Z \to \mu\mu$; agreement between the two post-fit spectra is found within uncertainties. The results of these comparisons are summarized in figure 7.74.

To further cross-check the theoretical modelling, an alternative fit has been developed and performed, with the aim of reducing the dependence on external theoretical models (that heavily rely on QCD, and in

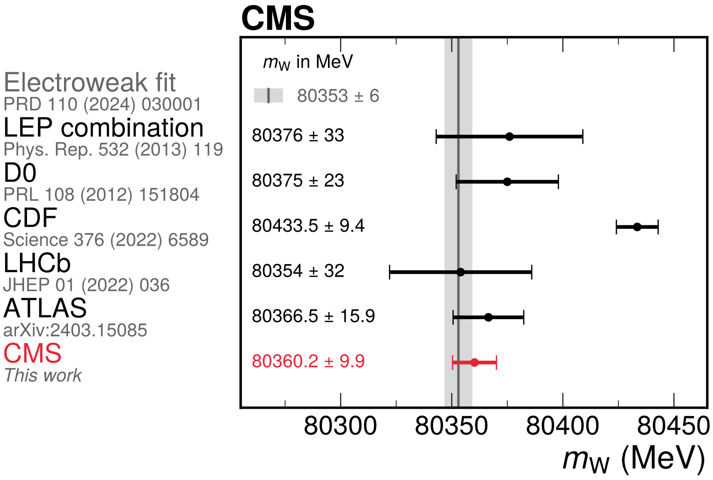


Figure 7.71: Comparison with other experiments and the EW fit prediction [9] (in gray). The m_W measurement from CMS (in red) is compared with the combined measurement of experiments at LEP [60], and with the measurements performed by the D0 [74, 166], CDF [3], LHCb [38], and ATLAS [37] experiments. From [1].

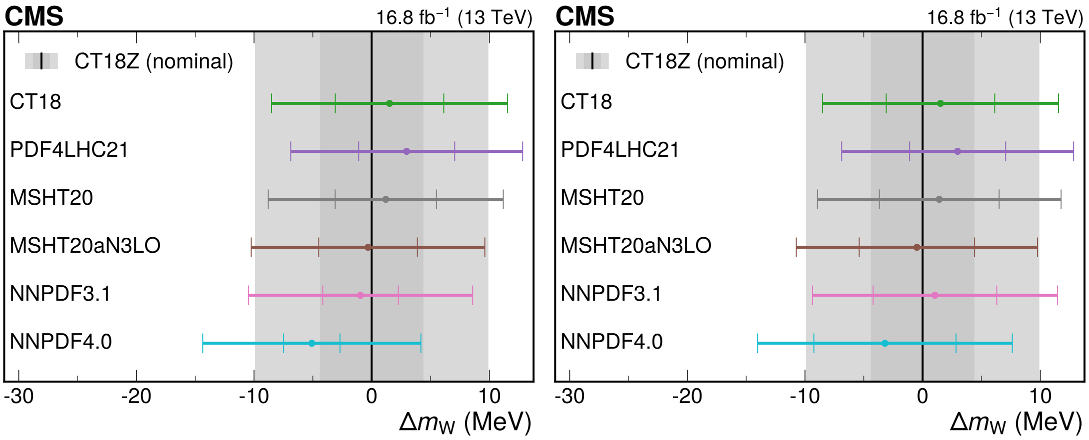


Figure 7.72: Difference in m_W values for the PDF sets mentioned in 7.3.2, when using the original uncertainty for the given set (left) and when the uncertainties are scaled to accommodate the central prediction of the other sets (right). The inner bar shows the uncertainty from the PDF, and the outer bar shows the total uncertainty. The nominal result, using CT18Z, is shown by the black line, with the CT18Z PDF and total uncertainty shown in dark and light gray, respectively. From [1].

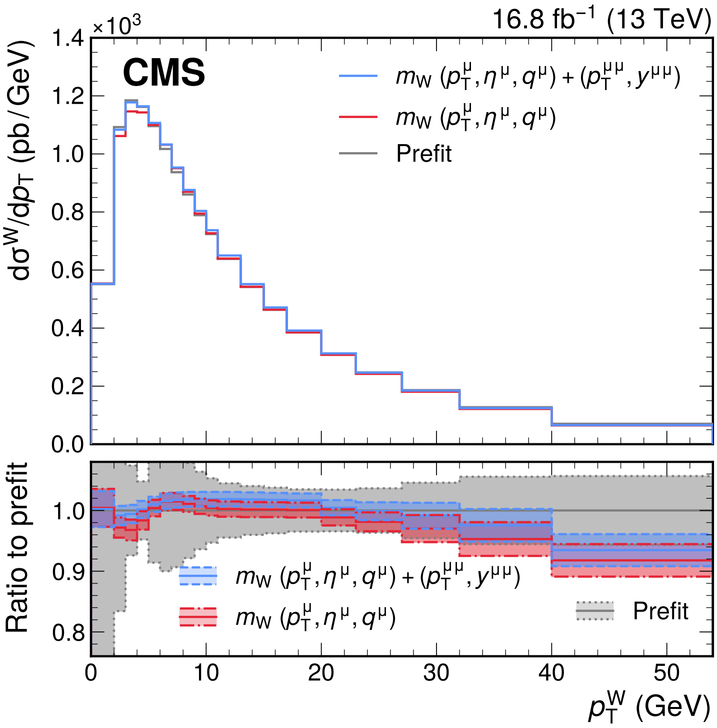


Figure 7.73: Pre-fit and post-fit generator-level p_T^W distribution. The distribution and uncertainties obtained from the simultaneous fit to the $p_T^\mu - \eta^\mu - q^\mu$ and $p_T^{\mu\mu} - y^{\mu\mu}$ distributions is shown in red, while the purple band shows the distribution obtained from the nominal fit to extract m_W that uses only the $p_T^\mu - \eta^\mu - q^\mu$. The ratio of the post-fit predictions to the pre-fit prediction (in gray), as well as their uncertainties, are shown by the shaded bands in the lower panel. From [1].

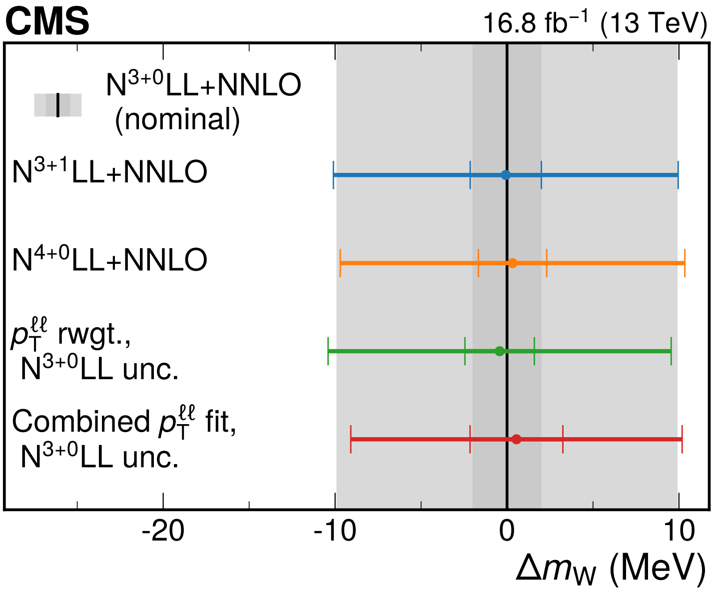


Figure 7.74: Difference between the m_W value extracted when the SCETLIB+DYTURBO predictions at N^{3}LL+NNLO (used in the nominal result) and alternative approaches to the p_T^W modeling and uncertainty are used. The p_T^W modeling uncertainties are shown as the inner bars while the outer bars denote the total uncertainty. From [1].

Source of uncertainty	Impact (MeV) Nominal in m_Z	Nominal in m_W	Global in m_Z	Global in m_W
Muon momentum scale	5.6	4.8	5.3	4.4
Muon reco. efficiency	3.8	3.0	3.0	2.3
W and Z angular coeffs.	4.9	3.3	4.5	3.0
Higher-order EW	2.2	2.0	2.2	1.9
p_T^V modeling	1.7	2.0	1.0	0.8
PDF	2.4	4.4	1.9	2.8
Nonprompt background	-	3.2	-	1.7
Integrated luminosity	0.3	0.1	0.2	0.1
MC sample size	2.5	1.5	3.6	3.8
Data sample size	6.9	2.4	10.1	6.0
Total uncertainty	13.5	9.9	13.5	9.9

Table 7.12: Uncertainties in the W-like m_Z and m_W measurements, with contributions to the total uncertainty from individual sources separated according to the "nominal" [212] and "global" [200] definitions of the impacts.

PDF set	Extracted m_W (MeV) Original σ_{PDF}	Inflated σ_{PDF}
CT18Z	$80\,360.2 \pm 9.9$	
CT18	$80\,361.8 \pm 10.0$	
PDF4LHC21	$80\,363.2 \pm 9.9$	
MSHT20	$80\,361.4 \pm 10.0$	$80\,361.7 \pm 10.4$
MSHT20aN^{3}LO	$80\,359.9 \pm 9.9$	$80\,359.8 \pm 10.3$
NNPDF3.1	$80\,359.3 \pm 9.5$	$80\,361.3 \pm 10.4$
NNPDF4.0	$80\,355.1 \pm 9.3$	$80\,357.0 \pm 10.8$

Table 7.13: The m_W values measured for different PDF sets indicated in 7.3.2, with both inflated and unscaled uncertainties.

principle might not capture effects from physics beyond the Standard Model in the production mechanism). This is referred to as "helicity cross-section fit": in this fit, the predictions on p_T^V, y^V and angular coefficients in equation 6.13 are replaced with parameters of interest that are extracted together with m_W and the other sources of uncertainty. In this way, the theoretical uncertainties are traded for statistical uncertainties. As was discussed in section 6.3 and shown in figure 6.10, variations in m_W and angular coefficients modify the $p_T^\mu - \eta^\mu - q^\mu$ spectrum very differently, in a way that can be disentangled by the fit, and it was discussed how this approach can reduce the PDF uncertainties. The "helicity cross-section fit" can be considered a less theory dependent approach compared to the one used in the nominal result. The helicity cross-sections σ_i, corresponding to the product of the A_i coefficients in equation 6.13 and the unpolarized cross-sections are fitted to have a set of parameters that is linear (so that the total cross-section is given by the sum of all the cross-section, while instead using the A_i coefficients directly leads to having to fit the product of multiple parameters). Terms with $i > 4$ are neglected, as no selection in ϕ^* in equation 6.13 is applied, so that those terms go to zero when integrated over ϕ^*. The parameters are fitted in the phase space of $p_T^V < 64$ GeV and $|y^V| < 3$, separately for the positive and negative charges, and this phase space is divided into 7 bins in $|y^V|$ and 8 bins in p_T^V, corresponding to a total of 864 parameters (the 5 helicity cross-sections for $i < 4$ and the unpolarized cross-section are fitted for each $p_T^V - |y^V|$ bin and charge). As the amount of statistics is not enough to sustain a fit in which the helicity cross-section parameters are

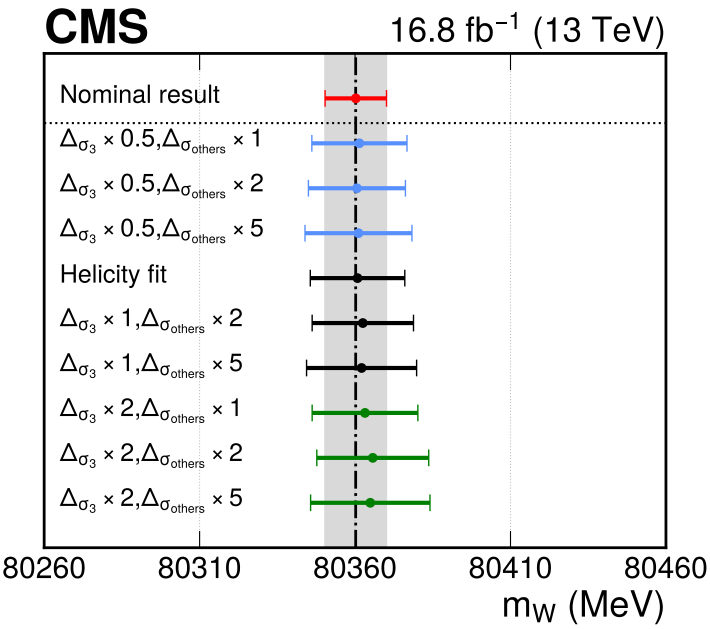


Figure 7.75: The W boson mass measured with the helicity cross-section fit for different pre-fit helicity cross section uncertainty configurations, separated for the σ_3 and the other components, respectively. From [1].

allowed to float freely and to reduce the impact of issues due to the finite amount of Monte Carlo statistics, for example discussed in [226], the nuisance parameters are constrained around the theoretical predictions with uncertainties that are relaxed with respect to the values used for the nominal result. The σ^{U+L} and σ^4 parameters have pre-fit uncertainties of $\pm 50\%$ and $\pm 100\%$ of the predicted cross sections, respectively, but these are also varied to check how the extracted value of m_W changes. To cover for all the possible correlated variations allowed by the theory model used for the nominal result, the uncertainties on the PDFs and missing higher-order terms, as well as the primordial$-k_T$ smearing and non-perturbative uncertainties in the angular coefficients, are included in the fit, with the latter also retained for the unpolarized cross-section, due to their impact at low p_T^V. Uncertainties in the unpolarized cross section from resummation, matching, and missing higher orders are not included as the variations induced by them should already be captured by the explicit σ^{U+L} variations. The results of the helicity cross-section fits, for several set of initial constraints on the helicity cross-sections, compared to the nominal result, are shown in figure 7.75. Using this approach, it was noticed that the uncertainty in m_W is only mildly dependent on the uncertainties on the helicity cross-sections, except for σ_3: this is reflected by the uncertainty configurations shown in figure 7.75, where the constraints on σ_3 are typically tighter than those of the other 5 cross-sections (grouped as σ_{others}). The results from the helicity cross-section fit are in agreement with the nominal result, with the larger deviations corresponding to the results with the looser constraints (that also have larger uncertainties due to the helicity cross-sections themselves).

In order to test the experimental inputs, m_W was also fitted separately in each η bin, and no particular trend was seen within uncertainties (mostly dominated by the finite amount of statistics). For the standard fit, the W-like m_Z and m_W analyses were also performed splitting events by the sign of η^μ, for each half of the CMS detector separately, and also by splitting events to separate between muons in the central region, $|\eta^\mu| < 0.9$, or in the forward region, $|\eta^\mu| \geq 0.9$ (for the W-like this was done for both muons from the $Z \to \mu\mu$ decay). The m_W analysis was also performed reducing the p_T^μ range considered. The alternative ranges 26 GeV$< p_T^\mu < 52$ GeV, 30 GeV$< p_T^\mu < 56$ GeV and 30 GeV$< p_T^\mu < 52$ GeV were considered. m_W was also fitted together with a scaling factor that multiplies the normalization for the W signal process, obtaining a scaling factor of 0.979 ± 0.026. In all cases only a negligible shift in the extracted value of m_W

was obtained.

The mass difference between W^+ and W^- was also fitted to validate the experimental and theoretical model, to validate the sources of uncertainty that behave differently for the two charges, the impact of which might be reduced when both charges are considered. This was done by replacing the parameter corresponding to variations in m_W with two separate parameters, one corresponding to the average value of m_W for the positive and negative W boson:

$$m_W^{\text{avg}} = \frac{m_{W^+} + m_{W^-}}{2} \tag{7.30}$$

and the other one corresponding to the difference between the masses of the positive and negative W boson:

$$\Delta m_W = m_{W^+} - m_{W^-}. \tag{7.31}$$

In this way, both the shift in the average value and the difference can be determined properly taking into account the correlation of uncertainties across channels and parameters. The value obtained is:

$$m_{W^+} - m_{W^-} = 57.0 \pm 30.3 \text{ MeV}, \tag{7.32}$$

compatible with 0 within two standard deviations, with a $p-$value for the compatibility with 0 of 6%, with a negligible difference between $m_W^{(1)}$ and the nominal value of m_W in 7.29. The correlation between $m_W^{(1)}$ and Δm_W is 0.02. As a cross-check, the same test was repeated for the W-like analysis:

$$m_{Z^+} - m_{Z^-} = 30.9 \pm 32.5 \text{ MeV} \tag{7.33}$$

for the nominal splitting of events, and:

$$m_{Z^+} - m_{Z^-} = 6.4 \pm 32.3 \text{ MeV} \tag{7.34}$$

for the alternate splitting, compatible in both cases with 0.

For both W and Z, a significant increase in the uncertainty in the mass differences compared to the mass average can be observed, related directly to the negative correlation of some uncertainties between the positive and negative charge. Table 7.14 summarizes the impacts, in both the nominal and global definitions, for the mass difference and average for both the m_W and W-like m_Z analyses: the increase in uncertainties can be linked to the increase in the statistical uncertainty, the uncertainty in the muon momentum scale, particularly in the corrections in the alignment, that are expected to have an opposite contribution to positive and negative charges, the angular coefficients, particularly A_3 for m_W, and the PDFs (although they are expected to be less relevant for these measurements). These anticorrelations can also explain the relatively small correlation between mass average and difference. The increase in the impact of statistical uncertainty is particularly relevant for the W-like m_Z extraction, as it can be seen in the difference of values extracted by changing the splitting of events, and this can in principle dilute the impact on the mass difference from systematic differences between the positive and negative muon. Because of this, several additional checks were performed to verify that the differences from zero in Δm_W does not reflect a bias or an underestimation of uncertainties that would impact result of this analysis.

Mass difference measurements were repeated after varying the central value of these parameters by one standard deviation to reduce the mass difference, without changing their relative uncertainty. When both are varied in such a way as to maximally reduce the mass difference, this moves by 1.2σ. An extreme configuration in which only the alignment term is varied by 3σ, to exclude that potential biases in the muon scale calibration[16] not covered by the uncertainty model of this analysis might induce biases in the extracted

---

[16] Some of these potential biases were already discussed in section 7.5, and might be related to the non-closure in the charge-dependent term (that should behave like the alignment term in equation 7.13) shown in figure 7.63

value of m_W. This variation of the alignment term results in $m_{W^+} - m_{W^-} \approx 0$, and in a shift of 0.6 MeV in the extracted value of m_W (shifts of similar or smaller order were obtained when other variations of A_3 and alignment were tested).

Source of uncertainty	Nominal impact (MeV)			
	$m_{Z^+} - m_{Z^-}$	m_Z	$m_{W^+} - m_{W^-}$	m_W
Muon momentum scale	23.1	5.6	21.6	4.8
Muon reco. efficiency	7.1	3.8	7.2	3.0
W and Z angular coeffs.	14.5	4.9	18.7	3.3
Higher-order EW	0.2	2.2	1.5	2.0
p_T^V modeling	0.6	1.7	7.4	2.0
PDF	0.9	2.4	11.8	4.4
Nonprompt background	-	-	7.5	3.2
Integrated luminosity	< 0.1	0.3	0.1	0.1
MC sample size	4.9	2.5	3.0	1.5
Data sample size	13.9	6.9	4.7	2.4
Total uncertainty	32.5	13.5	30.3	9.9

Source of uncertainty	Global impact (MeV)			
	$m_{Z^+} - m_{Z^-}$	m_Z	$m_{W^+} - m_{W^-}$	m_W
Muon momentum scale	21.2	5.3	20.0	4.4
Muon reco. efficiency	6.5	3.0	5.8	2.3
W and Z angular coeffs.	13.9	4.5	13.7	3.0
Higher-order EW	0.2	2.2	1.5	1.9
p_T^V modeling	0.4	1.0	2.7	0.8
PDF	0.7	1.9	4.2	2.8
Nonprompt background	-	-	4.8	1.7
Integrated luminosity	< 0.1	0.2	0.1	0.1
MC sample size	6.4	3.6	8.4	3.8
Data sample size	18.1	10.1	13.4	6.0
Total uncertainty	32.5	13.5	30.3	9.9

Table 7.14: Uncertainties in the W-like m_Z and m_W measurements, comparing the mass difference between charges and the nominal charge combination, using nominal (upper) and global (lower) impacts. From [1].

Further ongoing tests include splitting the dataset used in this analysis in different data-taking periods or in bins of instantaneous luminosity.

The m_W result presented here is in agreement with the Standard Model and has an uncertainty that is comparable with the most precise measurement by CDF in 2022. This result could help disentangle the W boson mass puzzle that arose after the publication of the latest CDF measurement. The methodology employed in this measurement relies on the most precise theoretical predictions and uncertainty model, particularly regarding the uncertainties for the non-perturbative and resummation uncertainties in the p_T^V modelling, out of all the measurements at hadron colliders, as well as a novel approach in the extraction of m_W and the theoretical model itself, as this is constrained by data and does not require inputs from $Z \to \mu\mu$ events, which was only used for validation purposes as described in section 7.8, so that no assumption about the degree of similarity and correlation between p_T^W and p_T^Z needs to be made. The new CVH refitting algorithm, specifically developed for this analysis, has been deployed and it was verified to improve the accuracy with which the muon momenta can be reconstructed, and the residual differences between true and reconstructed momenta are traced back to differences in the magnetic field, alignment and material budget within the tracker volume, account for by deriving the corrections as described in 7.5. The corrections are parameterized in terms of a physics-driven model, the robustness of which has been thoroughly tested. The multiple tests that have been presented seem to confirm that the result is robust under changes in the theoretical model (e.g. the PDF set chosen, the order up to which TNPs are included, other modifications of the p_T^V spectrum) or in possible missing sources of experimental uncertainties (as for example those that might account for the 2σ tension in $m_{W^+} - m_{W^-}$).

Chapter 8

Future prospects

The analysis presented in this thesis can be relatively easily extended to use the full Run 2 dataset collected by the CMS experiment. As seen in the previous chapter, the precision in the new CMS measurement of m_W is limited by statistics (6.0 MeV), and the dataset used in the analysis corresponds to roughly 10% of the total Run 2, so that a lot of information can be extracted from the remaining data. One of the possible challenges comes from the fact that the tracker geometry is different for 2017 and 2018, because of the installation of the new Phase-1 pixel detector shown in figure 3.6. Preliminary tests have been run, and the new CVH refit algorithm seems to converge successfully for the Pixel-1 geometry, and a preliminary version of the comparisons shown in figure 7.61 seems to indicate that the model in equation 7.13 holds also for the new geometry.

I have been involved in porting the CVH algorithm to the central CMSSW framework, so that it can be employed for other analysis or as part of the tracking effort, for example, it could be employed in the standard CMS alignment procedure. The CVH algorithm is being restudied to understand if there are possible improvements: it was found that the dip in event counts seen in data with respect to Monte Carlo observed around $\eta = 0.2$ in figure 7.39 is due to a non-negligible fraction of muon for which the CVH refit does not convergence successfully. Figure 8.1 shows the probability (measured using the same tag-and-probe procedure described in 5.2.1) that the CVH refit is successful for global muons (with all the analysis requirements) for 2016 (a similar situation was observed for 2018). A clear dip in efficiency around $\eta = 0.2$ can be observed, mostly flat in η. The implementation of this requirement in the $\mathrm{Tracking}$ efficiency derivation is being developed: it was left out because the efficiencies computed using the standard CMS reconstruction and the CVH refit are expected to be very similar, as the difference in p_T between the two reconstruction should be $O(10^{-3})$, and the inclusion of the momenta reconstructed with the CVH algorithm in the nTuples to derive the efficiency corrections (see 5.2.4 for more details) is not trivial and would result in an increase in the time required to produce the nTuples. However, since this inefficiency is restricted to a few η bins, and it does not induce a variation in p_T, no bias on m_W is expected by neglecting this correction.

Assuming there are no changes in the muon definition and in the event selection in the analysis, muon reconstruction and identification efficiencies have been derived for both 2017 and 2018. As of this thesis, the additional muon scale and resolution and recoil corrections have not been derived yet, and no changes have been made to the estimation of the backgrounds. With this in mind, figure8.2 shows a preliminary p_T^μ spectrum for 2018. To define a reference point for comparisons, figure 8.3 shows the p_T^μ spectrum without the corrections on the muon scale and resolution for 2016.

As already discussed, the resolution on p_T^{miss} does not allow CMS to measure m_T with enough precision to provide a competitive measurement of m_W, at least for the luminosity conditions of Run 2. Preliminary estimates have been made on the integrated luminosity needed to have a competitive measurement of m_W on so-called "low pileup data", i.e. a dataset in which the instantaneous luminosity is reduced so that the average number of interactions per bunch-crossing $\langle\mu\rangle \approx 1 - 3$: it was estimated that about 1 fb^{-1} could be sufficient for a measurement with a 10 MeV uncertainty. However, possible improvements in the

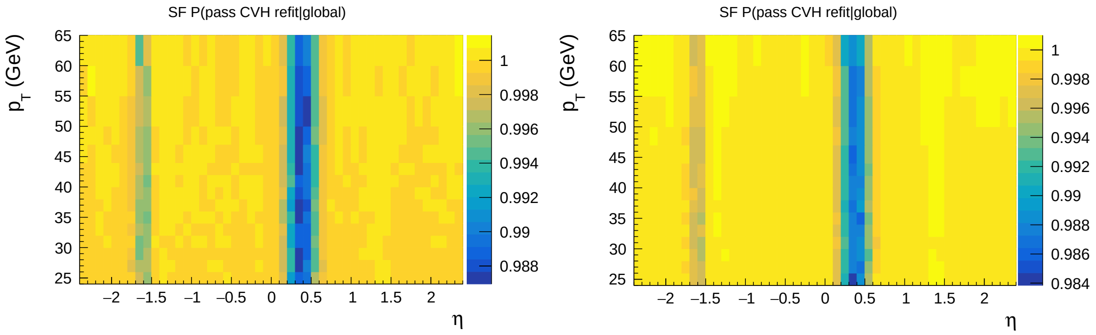


Figure 8.1: Probability that the CVH refit is successful for global muons (with all the analysis requirements), separately for positive (left) and negative (right) muons.

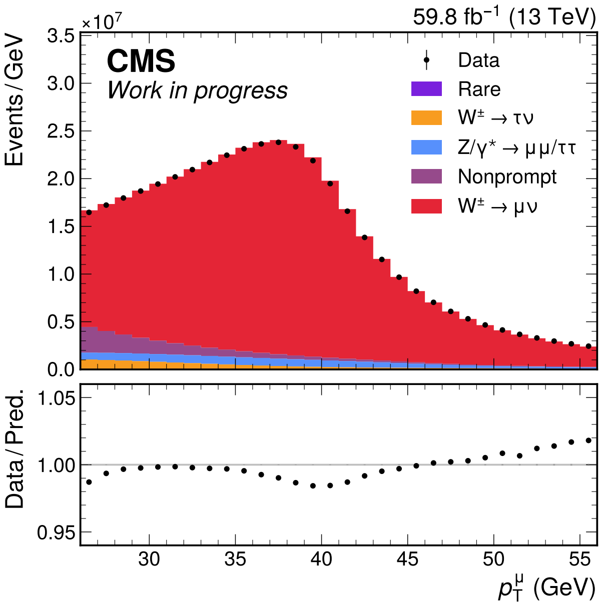


Figure 8.2: Preliminary p_T^μ on 2018 data. No recoil, momentum scale and resolution corrections are applied (p_T^μ is reconstructed using the CVH algorithm).

`DeepMET` algorithm used to reconstruct m_T could also lead to a competitive measurement of m_W also under standard Run 2 (or even Run 3) pileup conditions. For example, figure 8.4 shows a comparison in m_T distribution when m_T is reconstructed using the Particle Flow MET algorithm, the standard algorithm used to reconstruct p_T^{miss} in Run 2, and `DeepMET` (which has actually become the standard algorithm for Run 3). A significant improvement in the resolution can be observed when `DeepMET` is used, so that even larger improvements could be expected in future updates. The m_T variable is also sensitive to the width of the W boson Γ_W, another electroweak precision observable (see section 2.2), as shown for example by the latest m_W measurement by the ATLAS experiment [37]. This is another a potential target for CMS. New m_W measurements might also benefit from improvements in the theoretical uncertainties, for example related to the choice of future PDF sets that include new data on the theoretical accuracy, but also from increases in geometrical acceptance: the acceptance of both the inner tracker and the muon spectrometer will be extended when LHC moves to its High-Luminosity phase.

As of this thesis, work is ongoing to perform a new measurement of m_Z, such as the one reported in section 7.8, fully independent from the measurement at LEP and with improvements in the muon calibration with respect to this analysis. Another line of research that could be pursued using the data and analysis

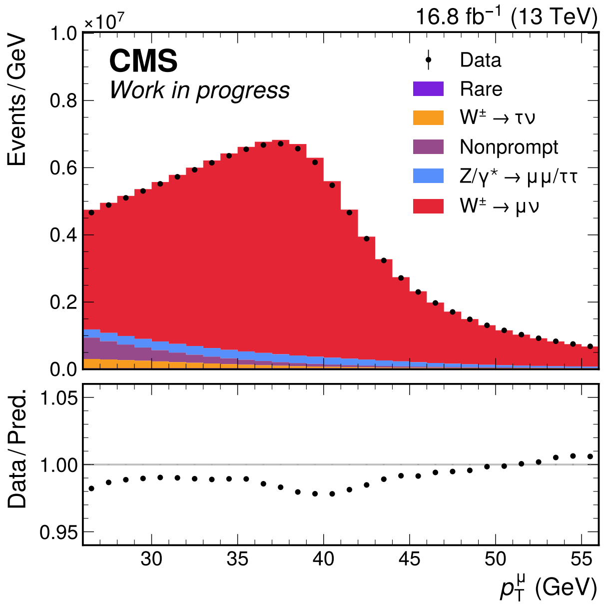


Figure 8.3: Same spectrum as in figure 8.2, but for 2016.

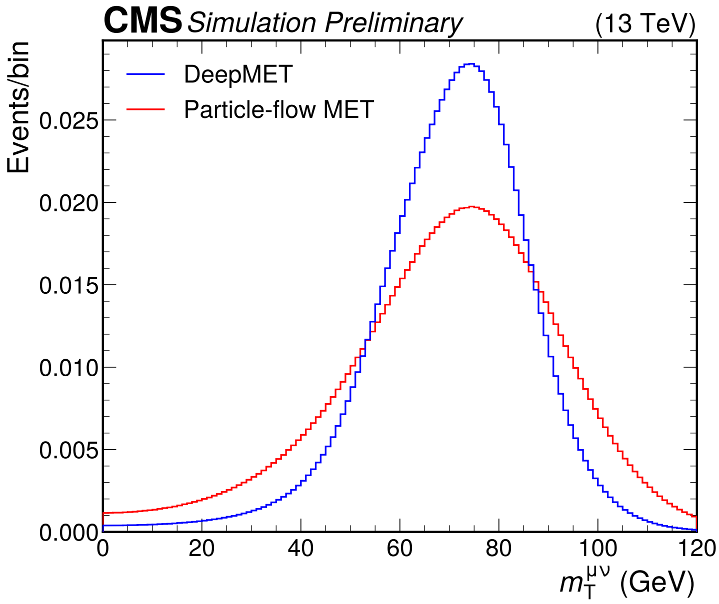


Figure 8.4: Comparison between the m_T distribution reconstructed using the Particle Flow MET algorithm used for standard Run 2 analyses or the DeepMET algorithm used in the m_W measurement by CMS.

framework developed for m_W is the measurement of the strong coupling constant α_S from the $p_T^{\mu\mu} - y^{\mu\mu}$ spectrum in $Z \to \mu\mu$ data, similar to the one presented by the ATLAS experiment [263], that could benefit from all the improvements in the theoretical modelling and uncertainties described in this thesis.

Chapter 9

Acknowledgements

Davide Bruschini acknowledges financial support from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement N. 101001205) [4].

Bibliography

[1] CMS Collaboration. *High-precision measurement of the W boson mass with the CMS experiment at the LHC*. 2024. URL: `https://arxiv.org/abs/2412.13872`.

[2] S. Amoroso et al. (LHC-TeV MW Working Group). *Compatibility and combination of world W-boson mass measurements*. 2023. URL: `https://arxiv.org/abs/2308.09417`.

[3] CDF Collaboration. "High-precision measurement of the W boson mass with the CDF II detector". In: *Science 376, 170-176* (2022). DOI: `10.1126/science.abk1781`.

[4] European Research Council. *Power to the LHC data: an ASYmptotically MOdel-independent measurement of the W boson mass*. 2025. URL: `https://cordis.europa.eu/project/id/101001205/results`.

[5] H. Greaves and T. Thomas. *The CPT Theorem*. URL: `https://users.ox.ac.uk/~mert2255/papers/cpt.pdf`.

[6] Super-Kamiokande Collaboration. "Evidence for Oscillation of Atmospheric Neutrinos". In: *Phys. Rev. Lett. 81, 1562* (1998). DOI: `10.1103/PhysRevLett.81.1562`.

[7] KATRIN collaboration. *Direct neutrino-mass measurement based on 259 days of KATRIN data*. URL: `https://arxiv.org/abs/2406.13516`.

[8] The ATLAS and CMS Collaborations. "Combination of measurements of the top quark mass from data collected by the ATLAS and CMS experiments at $\sqrt{s}$ = 7 and 8 TeV". In: *ATLAS-CONF-2023-066, CMS-PAS-TOP-22-001* (2023). URL: `https://cds.cern.ch/record/2872484`.

[9] S. Navaset al. (Particle Data Group). "Review of Particle Physics". In: *Phys. Rev. D110, 030001* (2024). DOI: `10.1103/PhysRevD.110.030001`.

[10] F.J. Ynduráin. "Limits on the mass of the gluon". In: *Phys.Lett.B 345 524-526* (1995). DOI: `10.1016/0370-2693(94)01677-5`.

[11] S. L. Glashow. "The renormalizability of vector meson interactions". In: *Nuclear Physics 10 107* (1959). DOI: `10.1016/0029-5582(59)90196-8`.

[12] G.'t Hooft. "Renormalizable Lagrangians for massive Yang-Mills fields". In: *Nuclear Physics B 35 167* (1971). DOI: `10.1016/0550-3213(71)90139-8`.

[13] N. Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett. 10, 531* (1963). DOI: `10.1103/PhysRevLett.10.531`.

[14] M. Kobayashi and T. Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics, Volume 49, Issue 2, Pages 652–657* (1973). DOI: `10.1143/PTP.49.652`.

[15] L.-L. Chau and W.-Y. Keung. "Comments on the Parametrization of the Kobayashi-Maskawa Matrix". In: *Phys. Rev. Lett. 53, 1802* (1984). DOI: `10.1103/PhysRevLett.53.1802`.

[16] L.-L. Chau and W.-Y. Keung. "Comments on the Parametrization of the Kobayashi-Maskawa Matrix". In: *Phys. Rev. Lett. 53, 1802* (1984). DOI: `https://doi.org/10.1103/PhysRevLett.53.1802`.

[17] S. Glashow. "The renormalizability of vector meson interactions". In: *Nucl. Phys. 10, 107* (1959).

[18] A. Salam. "Weak and electromagnetic interactions". In: *Nuovo Cimento. 11 (4): 568–577* (1959). DOI: `10.1007%2FBF02726525`.

[19] S. Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett. 19 (21): 1264–66* (1967). DOI: `10.1103%2FPhysRevLett.19.1264`.

[20] M. Gell-Mann. "Symmetries of Baryons and Mesons". In: *Phys. Rev. 125, 1067* (1962). DOI: `10.1103/PhysRev.125.1067`.

[21] CMS Collaboration. "Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at $\sqrt{s} = 8$ TeV and cross section ratios to 2.76 and 7 TeV". In: *JHEP 03 2017 156* (2017). DOI: `10.1007/JHEP03%282017%29156`.

[22] R. Gupta. "Introduction to Lattice QCD". In: *LAUR-98-3174* (1998). URL: `https://arxiv.org/abs/hep-lat/9807028`.

[23] V. Trimble. "Existence and Nature of Dark Matter in the Universe". In: *Ann.Rev.Astron.Astrophys. 25 425-472* (1987). DOI: `10.1146/annurev.aa.25.090187.002233`.

[24] A. G. Riess et al. "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant". In: *AJ 116 1009* (1998). DOI: `10.1086/300499`.

[25] B. A. Robson. "The Matter-Antimatter Asymmetry Problem". In: *Journal of High Energy Physics, Gravitation and Cosmology, 4, 166-178* (2018). DOI: `10.4236/jhepgc.2018.41015`.

[26] G. F. Giudice. *Naturalness after LHC8*. 2013. URL: `https://arxiv.org/abs/1307.7879`.

[27] D. Buttazzo et al. "Investigating the near-criticality of the Higgs boson". In: *J. High Energy Phys. 2013, 089* (2013). DOI: `10.1007/jhep12(2013)089`.

[28] GFitter group. *Gfitter. A Generic Fitter Project for HEP Model Testing*. URL: `http://project-gfitter.web.cern.ch/project-gfitter/`.

[29] J. de Bias et al. "Electroweak precision observables and Higgs-boson signal strengths in the Standard Model and beyond: present and future". In: *JHEP 12 (2016) 135* (2016). DOI: `10.1007/JHEP12(2016)135`.

[30] A. D. Linde. "Vacuum structure in gauge theories. The problem of strong CP violation and cosmology". In: *Phys.Lett.B 93 327-330* (1980). DOI: `10.1016/0370-2693(80)90524-9`.

[31] X. Fan et al. "Measurement of the Electron Magnetic Moment". In: *Phys. Rev. Lett. 130, 071801* (2023). DOI: `10.1103/PhysRevLett.130.071801`.

[32] R. Carey et al. *Mulan: a part-per-million measurement of the muon lifetime and determination of the Fermi constant*. 2021. URL: `https://arxiv.org/abs/2108.09182`.

[33] The ALEPH Collaboration, the DELPHI Collaboration, the L3 Collaboration, the OPAL Collaboration, the SLD Collaboration, the LEP Electroweak Working Group, the SLD electroweak, heavy flavour groups. "Precision Electroweak Measurements on the Z Resonance". In: *Phys.Rept.427:257-454* (2006). DOI: `10.1016/j.physrep.2005.12.006`.

[34] Y. Fischer et al. "The global electroweak fit in the SM and SMEFT". In: *PoS* EPS-HEP2023 (2024), p. 304. DOI: `10.22323/1.449.0304`.

[35] I. Dubovyk et al. "The two-loop electroweak bosonic corrections to $\sin^2 \theta b_{eff}$". In: *Phys. Lett. B762 184* (2016). DOI: `10.1016/j.physletb.2016.09.012`.

[36] I. Dubovyk et al. "Complete electroweak two-loop corrections to Z boson production and decay". In: *Phys. Lett. B 783 86* (2018). DOI: `10.1016/j.physletb.2018.06.037`.

[37] ATLAS Collaboration. "Measurement of the W-boson mass and width with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 7 TeV". In: *Eur. Phys. J. C 84 1309* (2024). DOI: `10.1140/epjc/s10052-024-13190-x`.

[38] LHCb Collaboration. "Measurement of the W boson mass". In: *JHEP 01 2022 036* (2022). DOI: `10.1007/JHEP01(2022)036`.

[39] G. Voutsinas et al. "Beam-beam effects on the luminosity measurement at LEP and the number of light neutrino species". In: *Physics Letters B 800 135068* (2020). DOI: `10.1016/j.physletb.2019.135068`.

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

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

[42] ATLAS Collaboration. "Measurement of the effective leptonic weak mixing angle using electron and muon pairs from Z-boson decay in the ATLAS experiment at $\sqrt{s}$ = 8 TeV". In: *ATLAS-CONF-2018-037* (2018). URL: `https://cds.cern.ch/record/2630340`.

[43] CMS Collaboration. "Measurement of the weak mixing angle using the forward-backward asymmetry of Drell-Yan events in pp collisions at 8 TeV". In: *Eur. Phys. J. C 78 701* (2018). DOI: `10.1140/epjc/s10052-018-6148-7`.

[44] LHCb Collaboration. "Measurement of the forward-backward asymmetry in $Z/\gamma^* \rightarrow \mu^+\mu^+$-decays and determination of the effective weak mixing angle". In: *JHEP 11 2015 190* (2015). DOI: `10.1007/JHEP11(2015)190`.

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

[46] M. Davier et al. "Reevaluation of the hadronic vacuum polarisation contributions to the Standard Model predictions of the muon g - 2 and $\alpha(M_Z)$ using newest hadronic cross-section data". In: *Eur. Phys. J. C77 827* (2017). DOI: `10.1140/epjc/s10052-017-5161-6`.

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

[48] Particle Data Group. "Review of Particle Physics". In: *PTEP 2022 083C01* (2022). DOI: `10.1093/ptep/ptac097`.

[49] CDF and D0 Collaborations. *Combination of CDF and D0 results on the mass of the top quark using up 9.7 fb^1 at the Tevatron*. 2016. URL: `https://arxiv.org/abs/1608.01881`.

[50] CMS Collaboration. "Measurement of the top quark mass using a profile likelihood approach with the lepton + jets final states in proton–proton collisions at $\sqrt{s} = 13\,TeV$". In: *Eur. Phys. J. C 83, 963* (2023). DOI: `10.1140/epjc/s10052-023-12050-4`.

[51] CMS Collaboration. "Measurement of the $t\bar{t}$ production cross section, the top quark mass, and the strong coupling constant using dilepton events in pp collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C 79, 368* (2019). DOI: `10.1140/epjc/s10052-019-6863-8`.

[52] CMS Collaboration. "Measurement of the top quark mass in the all-jets final state at $\sqrt{s} = 13$ TeV and combination with the lepton+jets channel". In: *Eur. Phys. J. C 79, 313* (2019). DOI: `10.1140/epjc/s10052-019-6788-2`.

[53] CMS Collaboration. “Measurement of the top quark mass using events with a single reconstructed top quark in pp collisions at $\sqrt{s} = 13$ TeV”. In: *J. High Energ. Phys. 2021, 161* (2021). DOI: `10.1007/JHEP12(2021)161`.

[54] ATLAS Collaboration. “Measurement of the top-quark mass using a leptonic invariant mass in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector”. In: *J. High Energ. Phys. 2023, 19* (2023). DOI: `10.1007/JHEP06(2023)019`.

[55] ATLAS Collaboration. “Measurement of the top-quark mass in $t\bar{t} \rightarrow$ dilepton events with the ATLAS experiment using the template method in 13 TeV pp collision data”. In: *ATLAS-CONF-2022-058* (2022). URL: `https://cds.cern.ch/record/2826701`.

[56] ATLAS Collaboration. “Measurement of the Higgs boson mass in the $H \rightarrow ZZ^*$,$H \rightarrow \updownarrow\updownarrow$ and $H \rightarrow \gamma\gamma$ channels with TeV pp collisions using the ATLAS detector”. In: *Physics Letters B* 784 (2018), pp. 345–366. DOI: `10.1016/j.physletb.2018.07.050`.

[57] CMS Collaboration. “A measurement of the Higgs boson mass in the diphoton decay channel”. In: *Physics Letters B* 805 (2020), p. 135425. DOI: `10.1016/j.physletb.2020.135425`.

[58] ATLAS Collaboration. “Evidence of off-shell Higgs boson production from ZZ leptonic decay channels and constraints on its total width with the ATLAS detector”. In: *Physics Letters B* 846 (2023), p. 138223. DOI: `10.1016/j.physletb.2023.138223`.

[59] CMS Collaboration. “Measurement of the Higgs boson width and evidence of its off-shell contributions to ZZ production”. In: *Nat. Phys. 18 1329* (2022). DOI: `10.1038/s41567-022-01682-0`.

[60] The ALEPH Collaboration, The DELPHI Collaboration, The L3 Collaboration, The OPAL Collaboration, The LEP Electroweak Working Group. “Electroweak Measurements in Electron-Positron Collisions at W-Boson-Pair Energies at LEP”. In: *Physics Reports* 532.4 (2013), pp. 119–244. DOI: `10.1016/j.physrep.2013.07.004`.

[61] Tevatron Electroweak Working Group. *Combination of CDF and D0 Results on the Width of the W boson*. 2010. URL: `https://arxiv.org/abs/1003.2826`.

[62] CMS Collaboration. “Precision measurement of the W boson decay branching fractions in proton-proton collisions at $\sqrt{s} = 13\,TeV$”. In: *Phys. Rev. D 105, 072008* (2022). DOI: `10.1103/PhysRevD.105.072008`.

[63] SLAC E158 Collaboration. “Precision Measurement of the Weak Mixing Angle in Møller Scattering”. In: *Phys. Rev. Lett. 95, 081601* (2005). DOI: `10.1103/PhysRevLett.95.081601`.

[64] The Jefferson Lab Qweak Collaboration. “Precision measurement of the weak charge of the proton”. In: *Nature 557, 207–211* (2018). DOI: `10.1038/s41586-018-0096-0`.

[65] C. S. Wood et al. “Measurement of Parity Nonconservation and an Anapole Moment in Cesium”. In: *Science 275, 1759-1763* (1997). DOI: `10.1126/science.275.5307.1759`.

[66] J. Guéna et al. “Measurement of the parity violating $6S - 7S$ transition amplitude in cesium achieved within 2×10^{-13} atomic-unit accuracy by stimulated-emission detection”. In: *Phys. Rev. A 71, 042108* (2005). DOI: `10.1103/PhysRevA.71.042108`.

[67] N. H. Edwards et al. “Precise Measurement of Parity Nonconserving Optical Rotation in Atomic Thallium”. In: *Phys. Rev. Lett. 74, 2654* (1995). DOI: `10.1103/PhysRevLett.74.2654`.

[68] P. A. Vetter et al. “Precise Test of Electroweak Theory from a New Measurement of Parity Nonconservation in Atomic Thallium”. In: *Phys. Rev. Lett. 74, 2658* (1995). DOI: `10.1103/PhysRevLett.74.2658`.

[69] The Jefferson Lab PVDIS Collaboration. “Measurement of parity violation in electron–quark scattering”. In: *Nature 506, 67–70* (2014). DOI: `10.1038/nature12964`.

[70] CHARM Collaboration, J. Dorenbosch, F. Udo, et al. "Experimental results on neutrino-electron scattering". In: *Particles and Fields 41, 567–589* (1989 (Erratum: Z.Phys.C 51, 142, 1991)). DOI: `10.1007/BF01564701`.

[71] L. Ahrens et al. "Determination of electroweak parameters from the elastic scattering of muon neutrinos and antineutrinos on electrons". In: *Phys. Rev. D 41, 3297* (1990). DOI: `10.1103/PhysRevD.41.3297`.

[72] CHARM II Collaboration. "Precision measurement of electroweak parameters from the scattering of muon-neutrinos on electrons". In: *Physics Letters B* 335.2 (1994), pp. 246–252. DOI: `10.1016/0370-2693(94)91421-4`.

[73] Muon $g-2$ Collaboration. "Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm". In: *Phys. Rev. Lett. 126, 141801* (2021). DOI: `{10.1103/PhysRevLett.126.141801}`.

[74] D0 Collaboration. "Measurement of the W Boson Mass with the D0 Detector". In: *Phys. Rev. Lett. 108, 151804* (2012). DOI: `10.1103/PhysRevLett.108.151804`.

[75] J. Erler, M. Luo. "Precision determination of heavy quark masses and the strong coupling constant". In: *Physics Letters B* 558.3 (2003), pp. 125–131. DOI: `10.1016/S0370-2693(03)00276-4`.

[76] The SLD Collaboration. "High-Precision Measurement of the Left-Right Z Boson Cross-Section Asymmetry". In: *Phys. Rev. Lett. 84, 5945* (2000). DOI: `10.1103/PhysRevLett.84.5945`.

[77] The SLD Collaboration. "Improved Direct Measurement of Leptonic Coupling Asymmetries with Polarized Z Bosons". In: *Phys. Rev. Lett. 86, 1162* (2001). DOI: `10.1103/PhysRevLett.86.1162`.

[78] CMS Collaboration. "Measurement of the τ lepton polarization in Z boson decays in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *J. High Energ. Phys. 2024, 101* (2024). DOI: `10.1007/JHEP01(2024)101`.

[79] ATLAS collaboration. "Measurement of the forward-backward asymmetry of electron and muon pair-production in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *J. High Energ. Phys. 2015, 49* (2015). DOI: `10.1007/JHEP09(2015)049`.

[80] CMS Collaboration. *Measurement of the Drell-Yan forward-backward asymmetry and of the effective leptonic weak mixing angle in proton-proton collisions at $\sqrt{s} = 13$ TeV*. 2024. URL: `https://arxiv.org/abs/2408.07622`.

[81] M. E. Peskin and T. Takeuchi. "Estimation of oblique electroweak corrections". In: *Phys. Rev. D 46, 381–409* (1992). DOI: `10.1103/PhysRevD.46.381`.

[82] J. Fan, L. Li, T. Liu, K. Lyu. "W-boson mass, electroweak precision tests, and SMEFT". In: *Phys. Rev. D 106, 073010* (2022). DOI: `10.1103/PhysRevD.106.073010`.

[83] O. S. Brüning et al. "LHC Design Report". In: *CERN Yellow Reports: Monographs* (CERN, Geneva, 2004). URL: `http://cds.cern.ch/record/782076`.

[84] I. Béjar Alonso et al. "High-Luminosity Large Hadron Collider (HL-LHC): Technical design report". In: *CERN Yellow Reports: Monographs* (CERN, Geneva, 2020). DOI: `10.23731/CYRM-2020-0010`.

[85] CERN Bulletin. "The particle suppliers. Les fournisseurs de particules". In: *BUL-NA-2010-077* (2010). URL: `https://cds.cern.ch/record/1255151`.

[86] ATLAS Collaboration. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST 3, S08003* (2008). DOI: `10.1088/1748-0221/3/08/S08003`.

[87] CMS Collaboration. "The CMS Experiment at the CERN LHC". In: *JINST 3, S08004* (2008). DOI: `10.1088/1748-0221/3/08/S08004`.

[88] ALICE Collaboration. "The ALICE experiment at the CERN LHC". In: *JINST 3, S08002* (2008). DOI: `10.1088/1748-0221/3/08/S08002`.

[89] LHCb Collaboration. “The LHCb Detector at the LHC”. In: *JINST 3, S08005* (2008). DOI: `10.1088/1748-0221/3/08/S08005`.

[90] CMS Collaboration. *CMS Luminosity - Public Results*. 2024. URL: `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`.

[91] CMS Collaboration. “Measurement of the weak mixing angle with the Drell-Yan process in proton-proton collisions at the LHC”. In: *Phys. Rev. D 84, 112002* (2011). DOI: `10.1103/PhysRevD.84.112002?`.

[92] CMS Collaboration. “Description and performance of track and primary-vertex reconstruction with the CMS tracker”. In: *JINST, 9 P10009* (2014). DOI: `10.1088/1748-0221/9/10/P10009`.

[93] The Tracker Group of the CMS Collaboration. “The CMS Phase-1 Pixel Detector Upgrade”. In: *CERN-CMS-NOTE-2020-005* (2020). URL: `https://cds.cern.ch/record/2745805`.

[94] CMS Collaboration. “The Electromagnetic Calorimeter Project Technical Design Report”. In: *CERN/LHCC 97-33, CMS TDR4* (1997).

[95] CMS Collaboration. “The CMS hadron calorimeter project : Technical Design Report”. In: *CERN-LHCC-97-031, CMS TDR2* (1997). URL: `https://cds.cern.ch/record/357153/`.

[96] CMS Collaboration. “The CMS magnet project: Technical Design Report”. In: *Technical Design Report CMS* (CERN, Geneva, 1997). URL: `https://cds.cern.ch/record/331056`.

[97] CMS Collaboration. “Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays”. In: *JINST 5, T03021. 35 p* (2009). DOI: `10.1088/1748-0221/5/03/T03021`.

[98] CMS Collaboration. “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$ = 13 TeV”. In: *JINST 13, P06015. 53 p* (2018). DOI: `10.1088/1748-0221/13/06/P06015`.

[99] D. Abbaneo et al. “Status report of the upgrade of the CMS muon system with Triple-GEM detectors”. In: *Nucl. Instrum. Methods Phys. Res. A 824, Frontier Detectors for Frontier Physics: Proceedings of the 13th Pisa Meeting on Advanced Detectors, 521–525* (2016). DOI: `10.1016/j.nima.2015.11.125`.

[100] CMS Collaboration. “Particle-flow reconstruction and global event description with the CMS detector”. In: *JINST 12 P10003* (2017). DOI: `10.1088/1748-0221/12/10/P10003`.

[101] CMS Collaboration. “CMS TriDAS project : Technical Design Report, Volume 1: The Trigger Systems”. In: *CERN-LHCC-2000-038 ; CMS-TDR-6-1* (2000). URL: `https://cds.cern.ch/record/706847`.

[102] The CMS Trigger, Data Acquisition Group. “The CMS High Level Trigger”. In: *Eur.Phys.J.C46:605-667* (2006). URL: `https://arxiv.org/abs/hep-ex/0512077`.

[103] CMS Collaboration. *CMS and CMSSW*. 2023. URL: `https://cms-sw.github.io`.

[104] CMS Collaboration. “Development of the CMS detector for the CERN LHC Run 3”. In: *JINST 19 05, P05064* (2024). DOI: `10.1088/1748-0221/19/05/P05064`.

[105] CMS Collaboration. “Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *JINST 15 (2020) P10017* (2020). DOI: `10.1088/1748-0221/15/10/P10017`.

[106] CMS Collaboration. “Early 2018 High-Level Trigger rates”. In: *CMS-DP-2018-057* (2018). URL: `https://cds.cern.ch/record/2644379`.

[107] R. Frühwirth. “Application of Kalman filtering to track and vertex fitting”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2 (1987), pp. 444–450. ISSN: 0168-9002. DOI: `10.1016/0168-9002(87)90887-4`.

[108] W. Adam, B. Mangano, T. Speer et al. “Track Reconstruction in the CMS tracker”. In: *CMS-NOTE-2006-041* (2006). URL: `%7Bhttps://cds.cern.ch/record/934067%7D`.

[109] F. Pantaleo et al. *New Track Seeding Techniques for the CMS Experiment*. 2017. URL: `http://cds.cern.ch/record/2293435`.

[110] J. R. Cash and A. H. Karp. "A Variable Order Runge-Kutta Method for Initial Value Problems with Rapidly Varying Right-hand Sides". In: *ACM Trans. Math. Softw. 16, 201* (1990). DOI: `10.1145/79505.79507`.

[111] S. Lantz et al. "Speeding up particle track reconstruction using a parallel Kalman filter algorithm". In: *Journal of Instrumentation* 15.09 (Sept. 2020), P09030. DOI: `10.1088/1748-0221/15/09/P09030`.

[112] CMS Collaboration. "Performance of Run 3 track reconstruction with the mkFit algorithm". In: *CMS-DP-2022-018* (2022). URL: `https://cds.cern.ch/record/2814000`.

[113] CMS Collaboration. "2017 tracking performance plots". In: *CMS-DP-2017-015* (2017). URL: `https://cds.cern.ch/record/2290524`.

[114] CMS Collaboration. "Tracking Performance in the CMS High Level Trigger - June 2018". In: *CMS-DP-2018-038* (2018). URL: `https://cds.cern.ch/record/2629370`.

[115] CMS Collaboration. "Performance of Run-3 HLT Track Reconstruction". In: *CMS-DP-2022-014* (2022). URL: `https://cds.cern.ch/record/2814111/`.

[116] A. Bocci et al. "Heterogeneous Reconstruction of Tracks and Primary Vertices With the CMS Pixel Tracker". In: *Front.Big Data 3 601728* (2020). DOI: `10.3389/fdata.2020.601728`.

[117] CMS Collaboration. "Performance of Track Reconstruction at the CMS High-Level Trigger in 2023 data". In: *CMS-DP-2024-013* (2024). URL: `https://cds.cern.ch/record/2890676`.

[118] K. Rose. "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems". In: *Proceedings of the IEEE* 86.11 (1998), pp. 2210–2239. DOI: `10.1109/5.726788`.

[119] W. Waltenberger et al. "Adaptive vertex fitting". In: *J. Phys. G: Nucl. Part. Phys. 34 N343* (2007). DOI: `10.1088/0954-3899/34/12/N01`.

[120] M. Cacciari et al. "The anti-k_T jet clustering algorithm". In: *J. High Energy Phys. 2008, 063–063* (2008). DOI: `10.1088/1126-6708/2008/04/063`.

[121] CMS Collaboration. "Early Run 3 tracking performance". In: *CMS-DP-2022-064* (2022). URL: `https://cds.cern.ch/record/2843180`.

[122] CMS Collaboration. "CMS tracking performance in 2023". In: *CMS-DP-2023-090* (2023). URL: `https://cds.cern.ch/record/2882249`.

[123] CMS Collaboration. "CMS tracking performance in 2024". In: *CMS-DP-2025-026* (2025). URL: `https://cds.cern.ch/record/2936309`.

[124] D. Bruschini. "Comparison of track properties between Data and MC at CMS using early Run-3 data". In: 2022. URL: `https://2022.congresso.sif.it/talk/579`.

[125] D. Barney. *CMS Detector Slice*. 2016. URL: `https://cds.cern.ch/record/2120661`.

[126] CMS Collaboration. "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV". In: *JINST 7 P10002* (2012). DOI: `10.1088/1748-0221/7/10/P10002`.

[127] UA1 Collaboration. "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s}$ = 540 GeV". In: *Physics Letters B 122, 103–116* (1983). DOI: `10.1016/0370-2693(83)91177-2`.

[128] UA2 Collaboration. "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider". In: *Physics Letters B 122, 476–485* (1983). DOI: `10.1016/0370-2693(83)91605-2`.

[129] CERN. "The 300 GeV programme". In: *Geneva : CERN, 1972 - 238.* (1972). URL: `https://cds.cern.ch/record/104068`.

[130] L. Di Lella and C. Rubbia. "The Discovery of the W and Z Particles". In: *Adv. Ser. Dir. High Energy Phys. 23, 137–163* (2015). DOI: `10.1142/9789814644150_0006`.

[131] ALEPH Collaboration. "ALEPH: technical report 1983". In: *CERN-LEPC-83-2* (1983). URL: `https://cds.cern.ch/record/300680`.

[132] DELPHI Collaboration. "DELPHI: technical proposal". In: *CERN-LEPC-83-3* (1983). URL: `https://cds.cern.ch/record/300668`.

[133] L3 Collaboration. "Technical proposal: L3". In: *CERN-LEPC-83-5* (1983). URL: `https://cds.cern.ch/record/297266`.

[134] OPAL Collaboration. "The OPAL detector : technical proposal". In: *CERN-LEPC-83-4* (1983). URL: `https://cds.cern.ch/record/300669`.

[135] ALEPH Collaboration. "Measurement of W-pair production in e+e- collisions at 183 GeV". In: *Phys. Lett. B 453, 107–120* (1999). DOI: `{10.1016/S0370-2693(99)00304-4}`.

[136] OPAL Collaboration. "WW Production Cross Section and W Branching Fractions in e^+e^- Collisions at 189 GeV". In: *Phys. Lett. B 493, 249-265* (2000). DOI: `{10.1016/S0370-2693%2800%2901085-6}`.

[137] ALEPH Collaboration. "Measurement of the W boson mass and width in e^+e^- collisions at LEP". In: *Eur. Phys. J. C 47, 309–335* (2006). DOI: `{10.1140/epjc/s2006-02576-8}`.

[138] P. Charitos. "FCC-ee Research Programme". In: *Zenodo* (2024). DOI: `10.5281/zenodo.15052382`.

[139] M. Benedikt. "Future Circular Collider Feasibility Study Report Volume 1: Physics and Experiments". In: *CERN Document Server* (2025). DOI: `10.17181/CERN.9DKX.TDH9`.

[140] M. Benedikt. "Future Circular Collider Feasibility Study Report Volume 2: Accelerators, technical infrastructure and safety". In: *CERN Document Server* (2025). DOI: `10.17181/CERN.EBAY.7W4X`.

[141] M. Benedikt. "Future Circular Collider Feasibility Study Report Volume 3: Civil Engineering, Implementation and Sustainability". In: *CERN Document Server* (2025). DOI: `10.17181/CERN.I26X.V4VF`.

[142] A. Blondel, E. Gianfelice. "The challenges of beam polarization and keV-scale centre-of-mass energy calibration at the FCC-ee". In: *Eur. Phys. J. Plus 136, 1103* (2021). DOI: `10.1140/epjp/s13360-021-02038-y`.

[143] P. Charitos. "Correction to: Articles from Article Collection "Focus Point on A Future Higgs & Electroweak Factory (FCC)"". In: *Eur. Phys. J. Plus 137, 792* (2022). DOI: `10.1140/epjp/s13360-022-02959-2`.

[144] T. Behnke et al. "The International Linear Collider Technical Design Report - Volume 1: Executive Summary". In: *ILC-REPORT-2013-040* (2013). URL: `https://arxiv.org/abs/1306.6327`.

[145] H. Baer et al. "The International Linear Collider Technical Design Report - Volume 2: Physics". In: *ILC-REPORT-2013-040* (2013). URL: `https://arxiv.org/abs/1306.6352`.

[146] C. Adolphsen et al. "The International Linear Collider Technical Design Report - Volume 3.I: Accelerator R&D in the Technical Design Phase". In: *ILC-REPORT-2013-040* (2013). URL: `https://arxiv.org/abs/1306.6353`.

[147] C. Adolphsen et al. "The International Linear Collider Technical Design Report - Volume 3.II: Accelerator Baseline Design". In: *ILC-REPORT-2013-040* (2013). URL: `https://arxiv.org/abs/1306.6328`.

[148] T. Behnke et al. “The International Linear Collider Technical Design Report - Volume 4: Detectors”. In: *ILC-REPORT-2013-040* (2013). URL: `https://arxiv.org/abs/1306.6329`.

[149] O. Brunner et al. *The CLIC project*. 2022. URL: `https://arxiv.org/abs/2203.09186`.

[150] P. Azzurri. “The W mass and width measurement challenge at FCC-ee”. In: *Eur. Phys. J. Plus 136, 1203* (2021). DOI: `10.1140/epjp/s13360-021-02211-3`.

[151] M. Freytsis et al. “Prospects for a Measurement of the W Boson Mass in the All-Jets Final State at Hadron Colliders”. In: *J. High Energy Phys. 02, 2019, 003* (2019). DOI: `{10.1007/JHEP02(2019)003}`.

[152] CMS Collaboration. “Search for the rare decay of the W boson into a pion and a photon in proton-proton collisions at $\sqrt{s}$ = 13 TeV”. In: *Phys. Lett. B 819 136409* (2021). DOI: `10.1016/j.physletb.2021.136409`.

[153] E. Jones, W. J. Murray. “Mass biases in exclusive radiative hadronic decays of W bosons at the LHC”. In: *New J. Phys. 23 113035* (2021). DOI: `10.1088/1367-2630/ac3572`.

[154] J. Collins et al. “Transverse momentum distribution in Drell-Yan pair and W and Z boson production”. In: *Nucl. Phys. B 250, 199–224* (1985). DOI: `10.1016/0550-3213(85)90479-1`.

[155] S. Catani et al. “Universality of non-leading logarithmic contributions in transverse-momentum distributions”. In: *Nucl. Phys. B 596, 299–312* (2001). DOI: `10.1016/S0550-3213(00)00617-9`.

[156] S. Camarda et al. “DYTurbo: Fast predictions for Drell-Yan processes”. In: *Eur. Phys. J. C 80, 251* (2020). DOI: `10.1140/epjc/s10052-020-7757-5`.

[157] S. Camarda et al. “Drell–Yan lepton-pair production: q_T resummation at N^{4}LL accuracy”. In: *Phys. Lett. B 845 (2023) 138125* (2023). DOI: `10.1016/j.physletb.2023.138125`.

[158] V. Bertacchi. “Towards a simultaneous measurement of W boson mass and production properties with the CMS detector”. In: *CERN-THESIS-2021-100* (2021). URL: `https://cds.cern.ch/record/2776894`.

[159] M. Cipriani. “Measurement of the helicity of the W boson with the CMS experiment”. In: *Ph.D. thesis at Sapienza Università di Roma* (2019). URL: `https://iris.uniroma1.it/handle/11573/1241411`.

[160] E. Mirkes. “Angular decay distribution of leptons from W bosons at NLO in hadronic collisions”. In: *Nucl. Phys. B 387 3* (1992). DOI: `10.1016/0550-3213(92)90046-E`.

[161] R. Wilson. “The Tevatron”. In: *Phys.Today 30N10 (1977) 23-30* (1977). DOI: `10.1063/1.3037746`.

[162] CDF Collaboration. “The CDF detector: an overview”. In: *Nucl.Instrum.Meth.A271:387-403* (1988). DOI: `10.1016/0168-9002(88)90298-7`.

[163] CDF Collaboration. “The CDF II Detector Technical Design Report”. In: *FERMILAB-DESIGN-1996-01* (1996). URL: `https://inspirehep.net/literature/428832`.

[164] D0 Collaboration. “Design Report: The D0 Experiment at the Fermilab Antiproton - Proton Collider”. In: *FERMILAB-DESIGN-1984-02* (1984). DOI: `10.2172/1879497`.

[165] D0 Collaboration. “The upgraded D0 detector”. In: *Nuclear Instruments and Methods in Physics Research A 565 (2006) 463–537* (2006). DOI: `10.1016/j.nima.2006.05.248`.

[166] D0 collaboration. “Measurement of the W boson mass with the D0 detector”. In: *Phys. Rev. D 89, 012005* (2014). DOI: `10.1103/PhysRevD.89.012005`.

[167] D0 Collaboration. “Measurement of the W Boson Mass”. In: *Phys. Rev. Lett. 103, 141801* (2009). DOI: `10.1103/PhysRevLett.103.141801`.

[168] C. Balázs and C.-P. Yuan. “Soft gluon effects on lepton pairs at hadron colliders”. In: *Phys. Rev. D 56, 5558* (1997). DOI: `10.1103/PhysRevD.56.5558`.

[169] P. Golonka and Z. Wąs. “PHOTOS Monte Carlo: a precision tool for QED corrections in Z and W decays”. In: *Eur. Phys. J. C 45, 97–107* (2006). DOI: `10.1140/epjc/s2005-02396-4`.

[170] F. Landry et al. “Fermilab Tevatron run-1 Z boson data and the Collins-Soper-Sterman resummation formalism”. In: *Phys. Rev. D 67, 073016* (2003). DOI: `10.1103/PhysRevD.67.073016`.

[171] D0 Collaboration. “Measurement of the Shape of the Boson-Transverse Momentum Distribution in $p\bar{p} \to Z/\gamma^* \to ee + X$ Events Produced at $\sqrt{s} = 1.96$ TeV”. In: *Phys. Rev. Lett. 100, 102002* (2008). DOI: `10.1103/PhysRevLett.100.102002`.

[172] P.M. Nadolsky et al. “Implications of CTEQ global analysis for collider observables”. In: *Phys. Rev. D 78, 013004* (2008). DOI: `10.1103/PhysRevD.78.013004`.

[173] U. Baur, S. Keller, and D. Wackeroth. “Electroweak radiative corrections to W boson production in hadronic collisions”. In: *Phys. Rev. D 59, 013002* (1998). DOI: `10.1103/PhysRevD.59.013002`.

[174] U. Baur, S. Keller, and W. K. Sakumoto. “QED radiative corrections to Z boson production and the forward-backward asymmetry at hadron colliders”. In: *Phys. Rev. D 57, 199* (1998). DOI: `10.1103/PhysRevD.57.199`.

[175] U. Baur, O. Brein, W. Hollik, C. Schappacher, and D. Wackeroth. “Electroweak radiative corrections to neutral-current Drell-Yan processes at hadron colliders”. In: *Phys. Rev. D 65, 033007* (2002). DOI: `10.1103/PhysRevD.65.033007`.

[176] G. A. Ladinsky, C. Yuan. “Nonperturbative regime in QCD resummation for gauge boson production at hadron colliders”. In: *Phys.Rev.D50:4239* (1994). DOI: `10.1103/PhysRevD.50.R4239`.

[177] G. Bozzi, S. Catani, G. Ferrera, D. de Florian, M. Grazzini. “Transverse-momentum resummation: A perturbative study of Z production at the Tevatron”. In: *Nucl. Phys. B 815, 174–197* (2009). DOI: `10.1016/j.nuclphysb.2009.02.014`.

[178] G. Bozzi, S. Catani, G. Ferrera, D. de Florian, M. Grazzini. “Production of Drell–Yan lepton pairs in hadron collisions: Transverse-momentum resummation at next-to-next-to-leading logarithmic accuracy”. In: *Phys. Lett. B 696, 207–213* (2011). DOI: `10.1016/j.physletb.2010.12.024`.

[179] C. M. Carloni Calame, G. Montagna, O. Nicrosini, A. Vicini. “Precision electroweak calculation of the production of a high transverse-momentum lepton pair at hadron colliders”. In: *JHEP 0710:109,2007* (2007). DOI: `10.1088/1126-6708/2007/10/109`.

[180] A. V. Kotwal, B. Jayatilaka. “Comparison of `HORACE` and `PHOTOS` Algorithms for Multiphoton Emission in the Context of W Boson Mass Measurement”. In: *Adv.High Energy Phys. 2016 1615081* (2016). DOI: `10.1155/2016/1615081`.

[181] NNPDF Collaboration. “Parton distributions from high-precision collider data”. In: *Eur. Phys. J. C 77, 663* (2017). DOI: `10.1140/epjc/s10052-017-5199-5`.

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

[183] L. A. Harland-Lang, A. D. Martin, P. Motylinski, R. S. Thorne. “Parton distributions in the LHC era: MMHT 2014 PDFs”. In: *Eur. Phys. J. C 75, 204* (2015). DOI: `10.1140/epjc/s10052-015-3397-6`.

[184] ATLAS Collaboration. “Measurement of the $W-$boson mass in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector”. In: *Eur. Phys. J. C 78 110* (2018). DOI: `10.1140/epjc/s10052-017-5475-4`.

[185] P. Nason. “A new method for combining NLO QCD with shower Monte Carlo algorithms”. In: *JHEP11 2004 040* (2004). DOI: `10.1088/1126-6708/2004/11/040`.

[186] S. Frixione, P. Nason and C. Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method". In: *JHEP11 2007 070* (2007). DOI: `10.1088/1126-6708/2007/11/070`.

[187] S. Alioli, P. Nason, C. Oleari and E. Re. "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX". In: *J. High Energ. Phys. 2010, 43* (2010). DOI: `10.1007/JHEP06(2010)043`.

[188] T. Sjöstrand, S. Mrenna and P. Skands. "PYTHIA 6.4 physics and manual". In: *JHEP05 2006 026* (2006). DOI: `10.1088/1126-6708/2006/05/026`.

[189] T. Sjöstrand, S. Mrenna and P. Skands. "A brief introduction to PYTHIA 8.1". In: *Comput. Phys. Commun. 178 (2008) 852* (2008). DOI: `10.1016/j.cpc.2008.01.036`.

[190] ATLAS Collaboration. "Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *J. High Energ. Phys. 2014, 145* (2014). DOI: `10.1007/JHEP09(2014)145`.

[191] ATLAS Collaboration. "Precise measurements of W and Z transverse momentum spectra with the ATLAS detector at $\sqrt{s} = 5.02$ TeV and 13 TeV". In: *CERN-EP-2024-080* (2024). URL: `https://cds.cern.ch/record/2894925`.

[192] S. Catani, L. Cieri, G. Ferrera, D. de Florian and M. Grazzini. "Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO". In: *Phys. Rev. Lett. 103, 082001* (2009). DOI: `10.1103/PhysRevLett.103.082001`.

[193] ATLAS Collaboration. "Measurement of the angular coefficients in $Z-$boson events using electron and muon pairs from data taken at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *JHEP 08 (2016) 159* (2016). DOI: `10.1007/JHEP08(2016)159`.

[194] H.-L. Lai et al. "New parton distributions for collider physics". In: *Phys. Rev. D 82, 074024* (2010). DOI: `10.1103/PhysRevD.82.074024`.

[195] J. Pumplin et al. "New Generation of Parton Distributions with Uncertainties from Global QCD Analysis". In: *JHEP07(2002)012* (2002). DOI: `10.1088/1126-6708/2002/07/012`.

[196] S. Jadach W. P laczek and M. W. Krasny. "Drell-Yan processes with WINHAC". In: *Acta Phys. Polon. B 44 2171–2178* (2013). DOI: `10.5506/APhysPolB.44.2171`.

[197] ATLAS Collaboration. "The ATLAS Simulation Infrastructure". In: *Eur. Phys. J. C 70 823* (2010). DOI: `10.1140/epjc/s10052-010-1429-9`.

[198] S. Agostinelli et al. "Geant4–a simulation toolkit". In: *Nucl. Instrum. Meth. A 506 250* (2003). DOI: `10.1016/S0168-9002(03)01368-8`.

[199] NNPDF Collaboration. "The path to proton structure at 1% accuracy". In: *Eur. Phys. J. C 82, 428* (2022). DOI: `10.1140/epjc/s10052-022-10328-7`.

[200] A. Pinto et al. "Uncertainty components in profile likelihood fits". In: *Eur. Phys. J. C 84 593* (2024). DOI: `10.1140/epjc/s10052-024-12877-5`.

[201] Bozzi, G., Citelli, L., Vesterinen, M. et al. "Prospects for improving the LHC W boson mass measurement with forward muons". In: *Eur. Phys. J. C 75, 601* (2015). DOI: `10.1140/epjc/s10052-015-3810-1`.

[202] I. Belyaev et al. "Handling of the generation of primary events in Gauss, the LHCb simulation framework". In: *J. Phys. Conf. Ser. 331 032047* (2011). DOI: `10.1088/1742-6596/331/3/032047`.

[203] T. Sjöstrand et al. "An introduction to PYTHIA 8.2". In: *Comput. Phys. Commun. 191 159* (2015). DOI: `10.1016/j.cpc.2015.01.024`.

[204] M. Clemencic et al. "The LHCb simulation application, Gauss: Design, evolution and experience". In: *J. Phys. Conf. Ser. 331 032023* (2011). DOI: `10.1088/1742-6596/331/3/032023`.

[205] R. Aaij et al. "Performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC". In: *JINST 14 P04013* (2019). DOI: `10.1088/1748-0221/14/04/P04013`.

[206] W. D. Hulsbergen. "The global covariance matrix of tracks fitted with a Kalman filter and an application in detector alignment". In: *Nucl. Instr. Meth. 600 471-477* (2009). DOI: `10.1016/j.nima.2008.11.094`.

[207] S. Bailey et al. "Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs". In: *Eur. Phys. J. C81 341* (2021). DOI: `10.1140/epjc/s10052-021-09057-0`.

[208] J. Bellm et al. "Herwig 7.0/Herwig++ 3.0 release note". In: *Eur. Phys. J. C76 196* (2016). DOI: `10.1140/epjc/s10052-016-4018-8`.

[209] L. Barze et al. "Implementation of electroweak corrections in the POWHEG BOX: single W production". In: *JHEP 04 037* (2012). DOI: `10.1007/JHEP04(2012)037`.

[210] L. Barze et al. "Neutral current Drell-Yan with combined QCD and electroweak corrections in the POWHEG BOX". In: *Eur. Phys. J. C73 2474* (2013). DOI: `10.1140/epjc/s10052-013-2474-y`.

[211] C. M. Carloni Calame et al. "Precision Measurement of the W-Boson mass: Theoretical Contributions and Uncertainties". In: *Phys. Rev. D96 093005* (2017). DOI: `10.1103/PhysRevD.96.093005`.

[212] CMS Collaboration. "Measurements of the W boson rapidity, helicity, double-differential cross sections, and charge asymmetry in pp collisions at $\sqrt{s} = 13$ TeV". In: *Phys. Rev. D 102 092012* (2020). DOI: `10.1103/PhysRevD.102.092012`.

[213] T. Cridge, L. A. Harland-Lang, and R. S. Thorne. "Combining QED and approximate N3LO QCD corrections in a global PDF fit: MSHT20qed_an3lo PDFs". In: *SciPost Phys. 17 026* (2024). DOI: `10.21468/SciPostPhys.17.1.026`.

[214] CMS Collaboration. *CMS Standard Model summary plots*. 2025. URL: `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined`.

[215] CMS EGamma POG. *EGamma ID recipes for Run 2*. Accessed on 04-06-2025. URL: `https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaIDRecipesRun2`.

[216] Y. Feng. "A New Deep-Neural-Network–Based Missing Transverse Momentum Estimator, and its Application to W Recoil". In: *CERN-THESIS-2020-194* (2020). URL: `https://repository.cern/records/xy5wh-y7z82`.

[217] A. Belloni, Y. Feng, M. Seidel, J. Steggemann. "DeepMET: Improving missing transverse momentum estimation with a deep neural network". In: *CMS Analysis Note 2024/056* (2024).

[218] P. F. Monni et al. "MiNNLO$_textPS$: a new method to match NNLO QCD to parton showers". In: *JHEP 05 143* (2020). DOI: `10.1007/JHEP05(2020)143`.

[219] P. F. Monni et al. "Erratum to: MiNNLO$_{\text{PS}}$: a new method to match NNLO QCD to parton showers". In: *J. High Energ. Phys. 2022, 31* (2022). DOI: `10.1007/JHEP02(2022)031`.

[220] P. F. Monni, E. Re, and M. Wiesemann. "MiNNLO$_{\text{PS}}$: optimizing $2 \to 1$ hadronic processes". In: *Eur. Phys. J. C 80 1075* (2020). DOI: `10.1140/epjc/s10052-020-08658-5`.

[221] N. Davidson, T. Przedzinski, and Z. Wąs. "PHOTOS interface in C++: Technical and physics documentation". In: *Comput. Phys. Commun. 199 86* (2016). DOI: `10.1140/epjc/s2005-02396-4`.

[222] CMS Collaboration. "Extraction and validation of a new set of CMS PYTHIA 8 tunes from underlying-event measurements". In: *Eur. Phys. J. C 80 4* (2020). DOI: `10.1140/epjc/s10052-019-7499-4`.

[223] CMS Collaboration. "Measurements of differential Z boson production cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *JHEP 12 061* (2019). DOI: `10.1007/JHEP12(2019)061`.

[224] R. J. Barlow and C. Beeston. “Fitting using finite Monte Carlo samples”. In: *Comput. Phys. Commun. 77 219* (1993). DOI: `10.1016/0010-4655(93)90005-W`.

[225] J. S. Conway. “Incorporating nuisance parameters in likelihoods for multisource spectra”. In: *PHYSTAT 2011: Workshop on statistical issues related to discovery claims in search experiments and unfolding* (H. B. Prosper and L. Lyons, eds., CERN Yellow Reports: Conf. Proc., p. 115. CERN, Geneva, 2011). DOI: `10.5170/CERN-2011-006.115`.

[226] C.-A. Alexe, J. L. Bendavid, L. Bianchini, and D. Bruschini. “Undercoverage in high-statistics counting experiments with finite MC samples”. In: (2024). URL: `http://www.arxiv.org/abs/2401.10542`.

[227] B. Efron. “Bootstrap methods: another look at the jackknife”. In: *Breakthroughs in statistics: methodology and distribution* (S. Kotz and N. L. Johnson, eds., p. 569. Springer New York, New York, NY, 1992). DOI: `http://dx.doi.org/10.1007/978-1-4612-4380-9_41`.

[228] M. A. Ebert, J. K. L. Michel, I. W. Stewart, and F. J. Tackmann. “Drell–Yan q_T resummation of fiducial power corrections at N^{3}LL”. In: *JHEP 04 102* (2021). DOI: `10.1007/JHEP04(2021)102`.

[229] G. Billis, J. K. L. Michel, and F. J. Tackmann. “Drell–Yan transverse-momentum spectra at N^{3}LL′ and approximate N^{4}LL with SCETlib”. In: *J. High Energ. Phys. 2025, 170* (2025). DOI: `10.1007/JHEP02(2025)170`.

[230] G. Billis, M. A. Ebert, J. K. L. Michel, and F. J. Tackmann. “A toolbox for q_T and $0-$jettiness subtractions at N^{3}LO”. In: *Eur. Phys. J. Plus 136 214* (2021). DOI: `10.1140/epjp/s13360-021-01155-y`.

[231] S. Camarda et al. “Erratum to: DYTurbo: fast predictions for Drell–Yan processes”. In: *Eur. Phys. J. C 80, 440* (2020). DOI: `10.1140/epjc/s10052-020-7972-0`.

[232] S. Camarda, L. Cieri, and G. Ferrera. “Drell–Yan lepton-pair production: q_T resummation at N^{3}LL accuracy and fiducial cross sections at N^{3}LO”. In: *Phys. Rev. D 104 L111503* (2021). DOI: `10.1103/PhysRevD.104.L111503`.

[233] W. Bizoń et al. “The transverse momentum spectrum of weak gauge bosons at N3 LL+NNLO”. In: *Eur. Phys. J. C 79 868* (2019). DOI: `{10.1140/epjc/s10052-019-7324-0}`.

[234] M. Grazzini, S. Kallweit, and M. Wiesemann. “Fully differential NNLO computations with MATRIX”. In: *Eur. Phys. J. C 78 537* (2018). DOI: `10.1140/epjc/s10052-018-5771-7`.

[235] T. Becher and T. Neumann. “Fiducial q_T resummation of color-singlet processes at N^{3}LL+NNLO”. In: *JHEP 03 199* (2021). DOI: `10.1007/JHEP03(2021)199`.

[236] J. Campbell and T. Neumann. “Precision phenomenology with MCFM”. In: *JHEP 12 034* (2019). DOI: `10.1007/JHEP12(2019)034`.

[237] F. J. Tackmann. “Beyond scale variations: perturbative theory uncertainties from nuisance parameters”. In: (2024). URL: `https://www.arxiv.org/abs/2411.18606`.

[238] J. C. Collins and D. E. Soper. “Back-to-back jets in QCD”. In: *Nucl. Phys. B 193 381* (1981). DOI: `10.1016/0550-3213(81)90339-4`.

[239] J. C. Collins and D. E. Soper. “Errata: Back-to-back jets in QCD”. In: *Nucl. Phys. B 213 545-546* (1983). DOI: `10.1016/0550-3213(83)90235-3`.

[240] PDF4LHC Working Group Collaboration. “The PDF4LHC21 combination of global PDF fits for the LHC Run III”. In: *J. Phys. G 49 080501* (2022). DOI: `10.1088/1361-6471/ac7216`.

[241] C. M. Carloni Calame, G. Montagna, O. Nicrosini, and M. Treccani. “Higher order QED corrections to W boson mass determination at hadron colliders”. In: *Phys. Rev. D 69 037301* (2004). DOI: `10.1103/PhysRevD.69.037301`.

[242] C. M. Carloni Calame, G. Montagna, O. Nicrosini, and M. Treccani. "Multiple photon corrections to the neutral-current Drell–Yan process". In: *JHEP 05 019* (2005). DOI: `10.1088/1126-6708/2005/05/019`.

[243] L. Barze et al. "Neutral current Drell–Yan with combined QCD and electroweak corrections in the POWHEG BOX". In: *Eur. Phys. J. C 73 2474* (2013). DOI: `10.1140/epjc/s10052-013-2474-y`.

[244] M. Chiesa, C. L. Del Pio, and F. Piccinini. "On electroweak corrections to neutral current Drell–Yan with the POWHEG BOX". In: *Eur. Phys. J. C 84 539* (2024). DOI: `10.1140/epjc/s10052-024-12908-1`.

[245] S. Bondarenko et al. "Hadron-hadron collision mode in ReneSANCe-v1.3.0". In: *Comput. Phys. Commun. 285 108646* (2023). DOI: `10.1016/j.cpc.2022.108646`.

[246] CERN. *RooFit - ROOT*. URL: `https://root.cern/manual/roofit/`.

[247] CERN. *MINUIT - Function Minimization and Error Analysis*. URL: `https://root.cern.ch/download/minuit.pdf`.

[248] Martìn Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Sys- tems*. URL: `www.tensorflow.org`.

[249] CMS Collaboration. "The CMS trigger system". In: *JINST 12 (2017) P01020* (2017). DOI: `10.1088/1748-0221/12/01/P01020`.

[250] CMS Collaboration. "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s}$ = 13 TeV". In: *JINST 15 (2020) P10017* (2020). DOI: `10.1088/1748-0221/15/10/P10017`.

[251] A. Bodek et al. "Extracting muon momentum scale corrections for hadron collider experiments". In: *EPJC 72 2194* (2012). DOI: `10.1140/epjc/s10052-012-2194-8`.

[252] V. Blobel, C. Kleinwort, and F. Meier. "Fast alignment of a complex tracking detector using advanced track models". In: *Comput. Phys. Commun. 182 1760* (2011). DOI: `10.1016/j.cpc.2011.03.017`.

[253] V. Blobel. "A new fast track-fit algorithm based on broken lines". In: *Nucl. Instrum. Meth. A 566 14* (2006). DOI: `http://dx.doi.org/10.1016/j.nima.2006.05.156`.

[254] CMS Collaboration. "Precision measurement of the structure of the CMS inner tracking system using nuclear interactions". In: *JINST 13 P10034* (2018). DOI: `10.1088/1748-0221/13/10/P10034`.

[255] CMS Collaboration. "Strategies and performance of the CMS silicon tracker alignment during LHC Run 2". In: *Nucl. Instrum. Meth. A 1037 166795* (2022). DOI: `10.1016/j.nima.2022.166795`.

[256] CMS Collaboration. "Alignment of the CMS tracker with LHC and cosmic ray data". In: *JINST 9 P06009* (2014). DOI: `10.1088/1748-0221/9/06/P06009`.

[257] J. Allison et al. "GEANT4 developments and applications". In: *IEEE Trans. Nucl. Sci. 53 270* (2006). DOI: `10.1109/TNS.2006.869826`.

[258] J. Allison et al. "Recent developments in GEANT4". In: *Nucl. Instrum. Meth. A 835 186* (2016). DOI: `10.1016/j.nima.2016.06.125`.

[259] V. Klyukhin et al. "The CMS magnetic field measuring and monitoring systems". In: *Symmetry 14 169* (2022). DOI: `10.3390/sym14010169`.

[260] 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". In: *JHEP 07 079* (2014). DOI: `10.1007/JHEP07(2014)079`.

[261] S. Choi and H. Oh. "Improved extrapolation methods of data-driven background estimations in high energy physics". In: *Eur. Phys. J. C 81 643* (2021). DOI: `10.1140/epjc/s10052-021-09404-1`.

[262] J. Bendavid and D. Walter. *GitHub - WMass/rabbit*. 2025. URL: `https://github.com/WMass/rabbit`.

[263] ATLAS Collaboration. *A precise determination of the strong-coupling constant from the recoil of Z bosons with the ATLAS experiment at $\sqrt{s} = 8$ TeV*. 2023. URL: `https://arxiv.org/abs/2309.12986`.