ETP-KA/2026-13

Karlsruher Institut für Technologie

Triple-differential measurement of Z+jet production with the CMS detector and comparison to Standard Model predictions

Zur Erlangung des akademischen Grades eines

DOKTORS DER NATURWISSENSCHAFTEN
(Dr. rer. nat.)

von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)
genehmigte

DISSERTATION

von

M.Sc. Cedric Verstege
aus Karlsruhe

Tag der mündlichen Prüfung: 3. Juli 2026

Referent: PD Dr. Klaus Rabbertz
Korreferent: Prof. Dr. Günter Quast

Abstract

In this thesis, a triple-differential cross section measurement for the production of a Z boson in association with at least one jet is presented. The analysed data were recorded by the CMS experiment at the CERN LHC from 2016 to 2018 in proton–proton collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\mathrm{TeV}$ and correspond to an integrated luminosity of $138\,\mathrm{fb}^{-1}$. The measurement is performed using the dimuon decay channel of the Z boson and is conducted triple-differentially in the transverse momentum of the Z boson, the boost in rapidity of the centre-of-mass system of the Z boson and the leading jet in the laboratory frame, and half the absolute difference in rapidity between the Z boson and the leading jet.

This choice of observables is directly sensitive to the parton momentum fractions and the scattering angle in the centre-of-mass frame of the hard interaction, which makes the measurement particularly useful for constraining the parton distribution functions of the proton. The measured event yields are corrected for detector effects and compared to predictions based on Monte Carlo event generators and calculations at next-to-next-to-leading order in perturbative QCD, which are corrected for electroweak and nonperturbative effects.

Acknowledgements

The past three and a half years working on this thesis have been an incredible journey. During this time, I had the opportunity to meet many scientists from around the world, especially during my research visits to CERN and through various conferences, workshops, and schools. The community that comes together at CERN, united by the shared goal of understanding the fundamental building blocks of the universe, has been a constant source of inspiration to me. It is a remarkable example of peaceful international collaboration in pursuit of knowledge.

First and foremost, I would like to thank my supervisor Klaus Rabbertz. I am especially grateful for his support in guiding this analysis through the CMS internal review process, which is quite a journey in itself. Throughout this work, he shared his deep understanding of particle physics and helped me to look at the analysis from many different perspectives.

I am also deeply grateful to my co-supervisor Günter Quast, who inspired me to pursue both my Master's degree and my PhD in particle physics. He has been a great mentor throughout my studies, and his enthusiasm for teaching and outreach has had a lasting influence on me. Teaching high-school students during the summer Science Camps and becoming a CERN guide, including for the CMS experiment, are experiences that I value greatly and that were strongly shaped by his encouragement.

I also want to thank Johannes for working together with Klaus on the theory predictions used in this analysis and for his ongoing work on using this measurement in determinations of the proton's structure. I am grateful to everyone who provided feedback on this thesis: Kylian, Nils, Max, Lars, Aritra, Nicolò, Benjamin, and Nikita. A special thank you goes to Robin, Max, and Alessandro, with whom I shared an office for most of my PhD. You made the everyday life of the PhD much more enjoyable, both inside and outside the office. I also want to thank Jost and Lars for the productive, memorable, and truly enjoyable time during our research stay at CERN.

A big thank you goes to everyone at the Institute of Experimental Particle Physics for the supportive and collegial working environment. In particular, I want to thank the admin team for keeping our computing infrastructure running. Without their work, this thesis would not have been possible.

Last but not least, I want to express my gratitude to my parents for their love and support throughout my studies.

Contents

CHAPTER 1

Introduction

With the discovery of the Higgs boson in 2012 [1, 2], all particles predicted by the Standard Model (SM) of particle physics had been experimentally observed. Nevertheless, the SM remains incomplete, leaving open questions such as the nature of dark matter, the origin of neutrino masses, and the matter–antimatter asymmetry of the Universe. In addition to direct searches for new phenomena, precision measurements of known SM processes provide a complementary approach that may reveal indirect signs of physics beyond the SM by deviating from precise theoretical predictions.

At the LHC, many precision measurements are limited by the knowledge of the proton structure. The momentum distributions of quarks and gluons inside the proton are described by parton distribution functions (PDFs), which are determined from global fits to experimental data. Their uncertainties propagate to precision observables, including measurements of the W boson mass [3, 4], the weak mixing angle [5, 6] and predictions for Higgs boson production cross sections [7, 8].

Precision measurements of vector-boson production provide important input to global PDF fits [9–12]. In particular, the production of a Z boson in association with at least one jet (Z + jet) is sensitive to the proton structure over a broad kinematic range. At leading order in perturbation theory, the process receives contributions from quark–antiquark-initiated and quark–gluon-initiated subprocesses, providing sensitivity to both quark and gluon PDFs.

Precision measurements of Z + jet production have been performed at the Fermilab Tevatron in proton–antiproton ($p\bar{p}$) collisions [13–16] and at the CERN LHC in proton–proton (pp) collisions [17–23]. Measurements differential in several observables are particularly valuable, since they separate different regions of parton momentum fractions and momentum-transfer scales and thereby reduce degeneracies in global PDF fits.

For $2 \to 2$ scattering processes, a particularly powerful strategy is to measure the cross section differentially in observables related to the hard scale, the longitudinal boost of the partonic system, and the scattering angle. These quantities are connected to the momentum fractions of the incoming partons and to the underlying partonic subprocesses. A triple-differential measurement can therefore separate the sensitivity to different PDF components more effectively than inclusive or single-differential measurements.

Recent advances in perturbative quantum chromodynamics (QCD) provide predictions for vector-boson production in association with one jet at next-to-next-to-leading order (NNLO) accuracy [24–26]. The large integrated luminosity collected by the CMS experiment during Run 2 of the LHC enables a precise triple-differential measurement of Z + jet production with uncertainties comparable to those of the theory predictions. Previous studies of this observable were presented in Refs. [27–30], paving the way for the measurement presented in this thesis and for the corresponding preliminary result by the CMS Collaboration [31].

In this thesis, a measurement of the triple-differential Z + jet production cross section is conducted using proton–proton collision data at a centre-of-mass energy of 13 TeV collected by the CMS experiment during Run 2 of the LHC. The data are corrected for detector effects and compared to state-of-the-art perturbative predictions. The measurement provides sensitivity to the proton structure over a broad kinematic range and constitutes an important input for future global PDF fits, while simultaneously serving as a stringent test of perturbative QCD. An example event display of a Z + jet candidate from the analysed dataset is shown in Figure 1.1.

This thesis is structured as follows: First, the theoretical foundations of the Standard Model of particle physics and the theory predictions for Z + jet production are presented in Chapter 2. This chapter focuses on the quantum chromodynamics that describes the strong force in Section 2.1, and gives an overview of factorization and PDFs in Section 2.2. It is followed by the definition of the triple-differential phase-space and binning in Section 2.3. The corresponding theoretical calculations for the production cross section of a Z boson in association with a jet are presented in Section 2.4. Additionally, these predictions are corrected for effects from higher-order calculations in electroweak theory, described in Section 2.5. Furthermore, nonperturbative (NP) corrections for Z + jet production are studied in Section 2.6 and are also already published in Ref. [32].

In Chapter 3, the experimental setup of the LHC and the CMS experiment is described. The measurement of the cross section using data collected by the CMS experiment is explained in Chapter 4, including the correction for detector effects in Section 4.4 to obtain the measured cross section at stable-particle level.

The measured data are then compared to predictions by Monte Carlo (MC) event generators and calculations at NNLO in perturbative QCD in Chapter 5. This includes a comparison to multiple PDF sets to assess the impact of these data on their determination. Finally, concluding remarks are given in Chapter 6.

Supplementary material, such as breakdowns across the full triple-differential phase-space for figures presented in the main matter, are given in Chapter A. Chapter B describes the software tools and computing resources used for the data analysis and derivation of NP corrections.

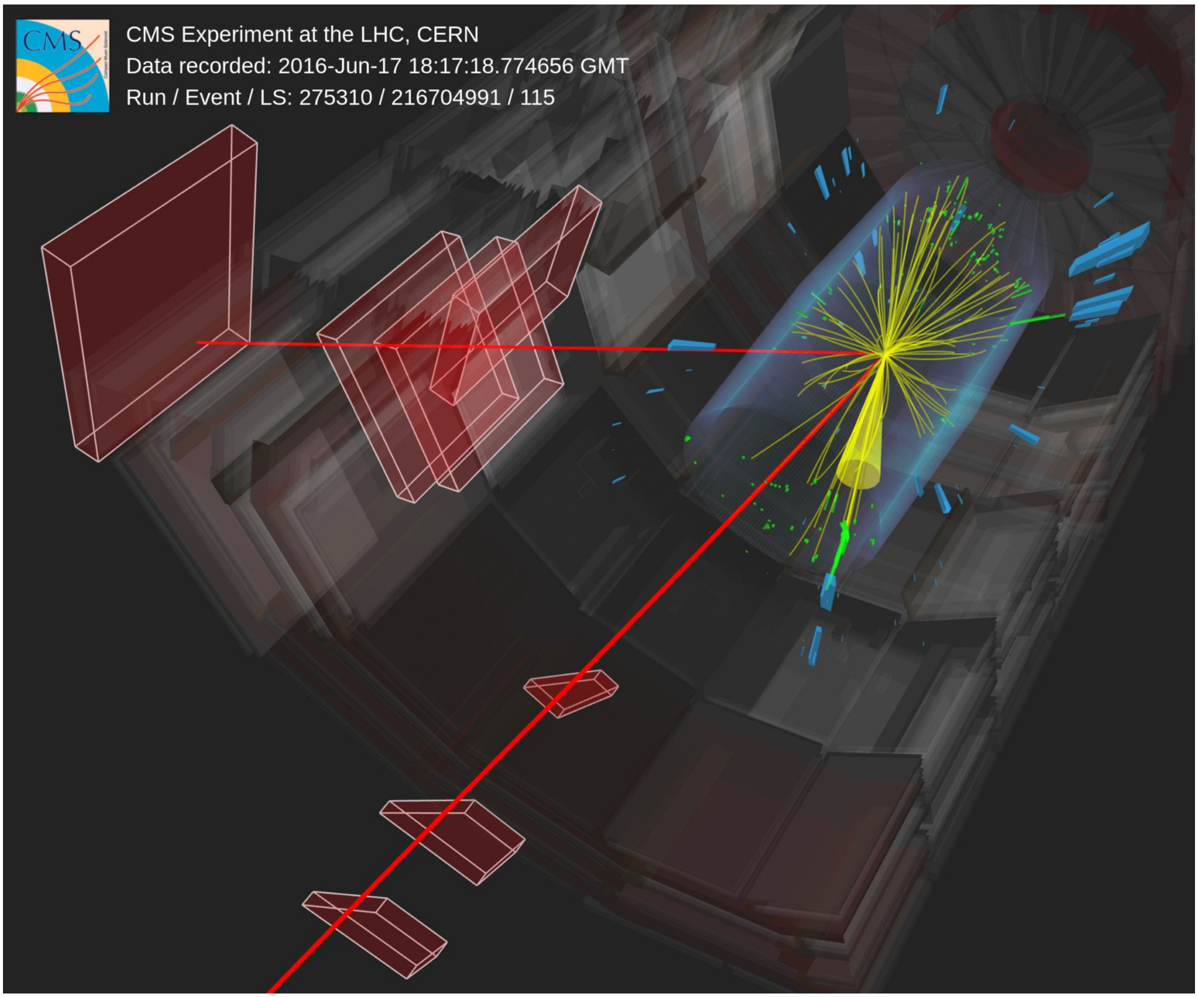


Figure 1.1: Display of a Z boson candidate decaying into two muons in association with a jet, recorded in proton–proton collisions with the CMS detector in 2016. The trajectories of the two muons forming the Z boson candidate are shown as red lines. Additional reconstructed particle trajectories are visible as curved yellow tracks, while energy deposits in the electromagnetic and hadronic calorimeters are shown as coloured bars. The jet, clustered from reconstructed particles, is indicated by the yellow cone. Parts of the muon system are shown semi-transparently in the background. The Z boson candidate and the jet are both reconstructed at large rapidities, illustrating a forward topology targeted by this measurement.

CHAPTER 2

Particle Physics

The Standard Model (SM) of particle physics has been tested to extraordinary precision and describes the fundamental particles and their interactions that govern our universe at the smallest scales. With the discovery of the Higgs boson in 2012 [1, 2] and the subsequent precision measurements of its properties [33, 34], the last missing piece of the SM was experimentally confirmed.

The fundamental particles and force carriers of the SM are summarised in Figure 2.1. All matter and antimatter is made up of spin-1/2 particles, the fermions, split into three generations. The quarks of the first generation are the building blocks of protons and neutrons, making up atomic nuclei, while the electron, the first-generation lepton, completes the structure of atoms. The interactions between particles are mediated by spin-1 particles, the gauge bosons, which are the force carriers of the SM. The photon is the massless mediator of the electromagnetic force, while the W and Z bosons are the massive mediators of the weak force. Gluons are the mediators of the strong force, which holds quarks together inside atomic nuclei. Particles obtain their mass through the Brout-Englert-Higgs mechanism [35, 36], which is mediated by the Higgs boson, a scalar particle with spin 0.

This chapter gives a brief overview of the theoretical framework of the SM, with a focus on the aspects relevant for the analysis presented in this thesis. It is largely based on the excellent descriptions in Refs. [37–39].

When colliding protons at the Large Hadron Collider (LHC), the strong interaction, described by quantum chromodynamics (QCD), is the dominant force, as discussed in Section 2.1. High-energy proton–proton collisions at the LHC are driven not by the protons themselves, but by their constituents, the partons. Each parton carries a fraction of the proton momentum, described by the parton distribution functions (PDFs). The factorization theorem allows this nonperturbative regime describing the proton structure to be separated from the hard scattering process, enabling precise theoretical predictions for proton–proton collisions. PDFs are discussed in Section 2.2.

Theoretical cross sections for scattering processes at the LHC are calculated perturbatively, order by order in the strong coupling constant α_S. For the production of a Z boson in association with jets, which is the process studied in this thesis, calculations are available up to next-to-next-to-leading order (NNLO) accuracy in perturbative QCD [24–26]. These predictions are presented for the triple-differential phase-space in Section 2.4. At the high

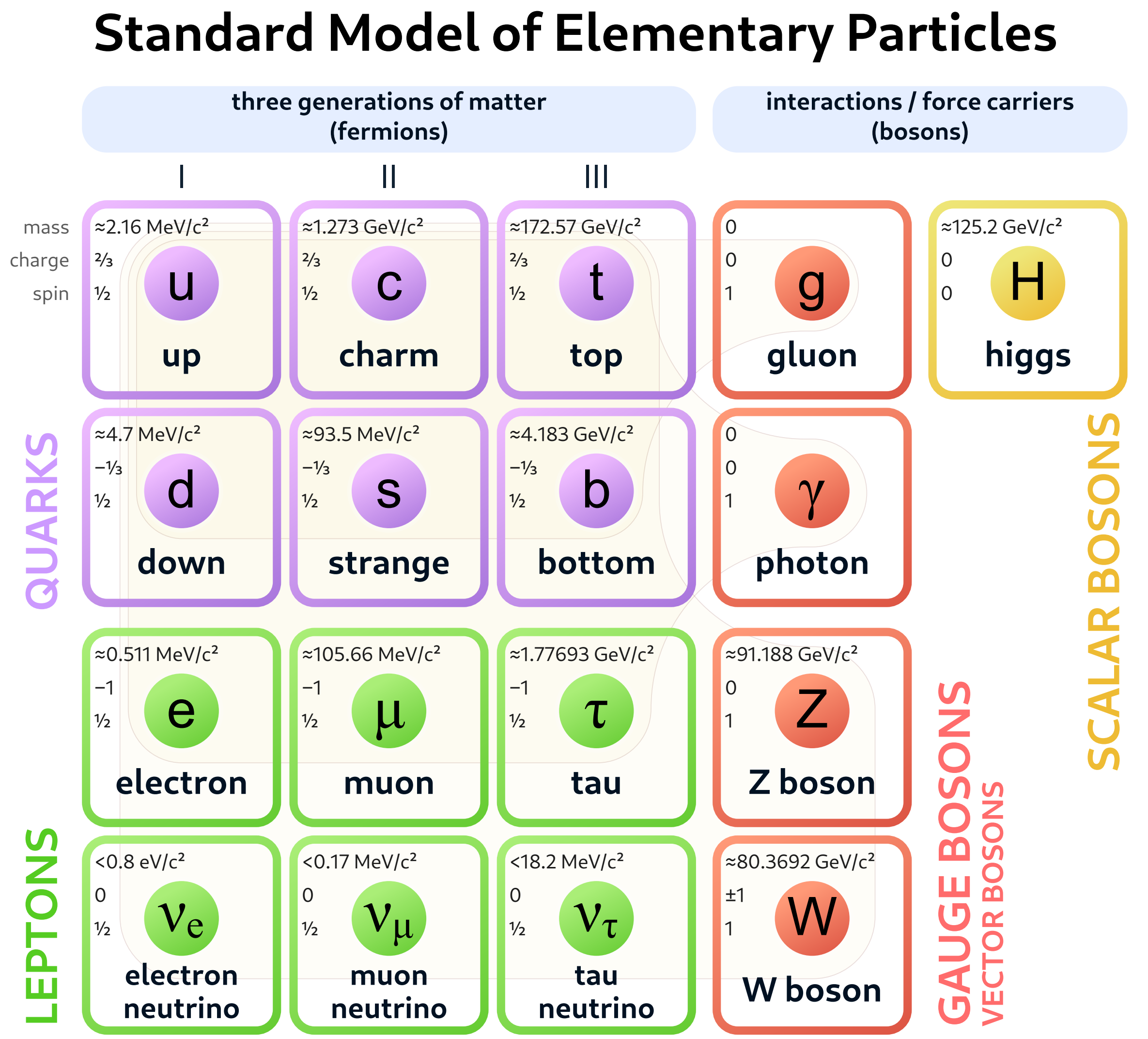


Figure 2.1: The Standard Model (SM) of particle physics. The fermions that make up all known matter in the universe are shown with their three generations on the left-hand side, while the bosons that mediate the forces between these particles are shown on the right-hand side. For each particle, the fundamental properties mass, charge, and spin are shown at the top left of each box. The colour-charged quarks interact via the strong force mediated by gluons. All particles with electric charge interact via the electromagnetic force mediated by the photon. The weak force is mediated by the W and Z bosons. Particles acquire their mass via the Higgs mechanism.

energies accessible at the LHC, higher-order electroweak (EW) corrections also become non-negligible and are reviewed in Section 2.5.

Finally, the hard scattering process alone is not sufficient to describe a full collision event. Nonperturbative effects, such as the hadronization of quarks and gluons into observable particles and the underlying event from additional softer parton interactions, need to be considered alongside the perturbative calculations. These effects, which are relevant at low energies, are discussed in Section 2.6.

2.1 Quantum chromodynamics

The strong interaction between quarks and gluons is described by quantum chromodynamics (QCD), the non-Abelian gauge theory of the strong force within the SM. QCD is based on a local $\mathrm{SU}(3)_\mathrm{C}$ gauge symmetry, where the index C denotes the colour charge associated with the strong interaction. The colour states of quarks form the fundamental, three-dimensional representation of $\mathrm{SU}(3)_\mathrm{C}$, referred to as a colour triplet. Gluons, in contrast, transform under the adjoint representation of $\mathrm{SU}(3)_\mathrm{C}$, which is eight-dimensional and corresponds to the eight generators of the gauge group.

The QCD Lagrangian density is given by

$$\mathscr{L}_\mathrm{QCD} = \sum_q \bar{\psi}_{q,a} \left(\mathrm{i}\gamma^\mu \partial_\mu \delta_{ab} - g_\mathrm{S} \gamma^\mu t^A_{ab} \mathscr{A}^A_\mu - m_q \delta_{ab} \right) \psi_{q,b} - \frac{1}{4} F^A_{\mu\nu} F^{A\,\mu\nu}, \tag{2.1}$$

where the Einstein summation convention over repeated indices is implied. The spinor fields $\psi_{q,a}$ describe quarks of flavour q and mass m_q, carrying the colour index $a = 1, \ldots, N_c$ with $N_c = 3$. The matrices t^A_{ab} are the generators of the SU(3) group in the fundamental representation and are related to the Gell-Mann matrices λ^A by

$$t^A = \frac{\lambda^A}{2}. \tag{2.2}$$

The gluon fields $\mathscr{A}^A_\mu$ with $A = 1, \ldots, 8$ are the gauge bosons associated with the eight generators of the symmetry group.

A fundamental parameter of QCD is the strong coupling constant g_S, or equivalently

$$\alpha_\mathrm{S} = \frac{g_\mathrm{S}^2}{4\pi}. \tag{2.3}$$

In contrast to quantum electrodynamics (QED), the non-Abelian structure of the SU(3) gauge group implies that gluons themselves carry colour charge and therefore interact directly with each other.

The gluon self-interactions originate from the non-linear term in the QCD field strength tensor

$$F^A_{\mu\nu} = \partial_\mu \mathscr{A}^A_\nu - \partial_\nu \mathscr{A}^A_\mu + g_\mathrm{S} f^{ABC} \mathscr{A}^B_\mu \mathscr{A}^C_\nu, \tag{2.4}$$

where f^{ABC} are the structure constants of the SU(3) group, defined through the commutation relation

$$[t^A, t^B] = \mathrm{i} f^{ABC} t^C. \tag{2.5}$$

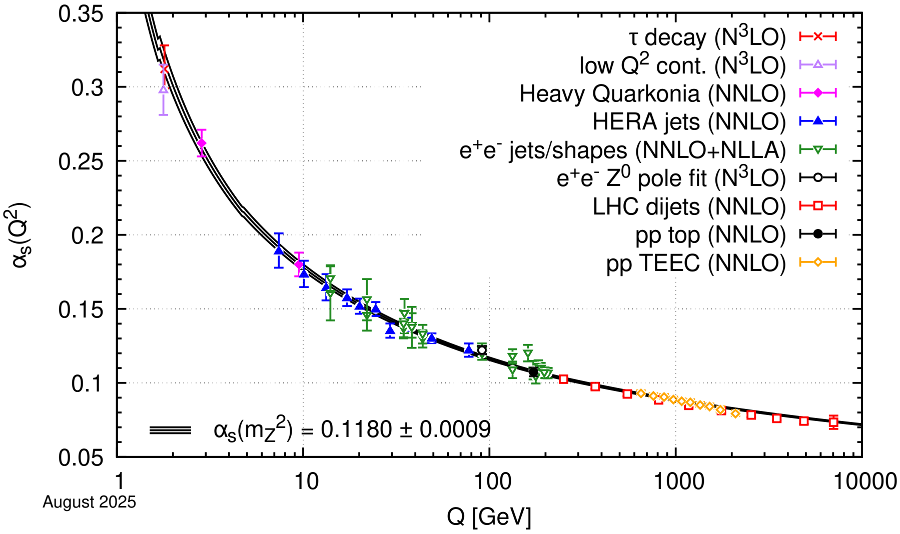


Figure 2.2: Running of the strong coupling constant α_S and summary of its determinations as a function of the energy scale Q. The running of α_S is computed at five-loop accuracy using the current PDG world average of $\alpha_S(m_Z) = 0.1180(9)$. Figure taken from the PDG review of 2025 [38].

The non-linear term proportional to g_S gives rise to three-gluon and four-gluon interaction vertices, which are absent in Abelian gauge theories such as QED.

Predictions in quantum chromodynamics can be obtained either through perturbative calculations or through nonperturbative approaches such as lattice gauge theory. Perturbative calculations are applicable at large momentum transfers or short distances, where the strong coupling constant becomes sufficiently small. This behaviour is known as asymptotic freedom [40, 41]. Conversely, at low energy scales the coupling strength increases, eventually causing perturbative calculations to break down and leading to the confinement of quarks and gluons into colour-neutral hadrons.

Beyond leading order, perturbative QCD calculations contain ultraviolet divergences arising from loop corrections. These divergences are absorbed into redefinitions of the theory parameters through the renormalization procedure, introducing the renormalization scale μ_R. As a consequence, the strong coupling constant becomes scale dependent. The scale dependence of α_S is described by the renormalization group equations and is referred to as the running of the strong coupling constant. At leading order, the running coupling is given by

$$\alpha_S(Q^2) = \frac{12\pi}{(33 - 2n_f)\ln(Q^2/\Lambda^2_{\mathrm{QCD}})}, \tag{2.6}$$

where n_f denotes the number of active quark flavours and Λ_{QCD} is the QCD scale parameter. In practical calculations, the characteristic momentum transfer Q of the process is commonly identified with the renormalization scale μ_R. The dependence of α_S on the energy scale Q is shown in Figure 2.2.

2.2 Factorization and PDFs

High-energy particle collisions probe progressively shorter distance scales. In proton–proton collisions at the LHC, it is therefore not the protons themselves that interact, but rather their constituent quarks and gluons, collectively referred to as partons. Since the proton is a composite object, the momentum participating in the hard interaction is only a fraction x of the incoming proton momentum. The description of the proton structure is encoded in the parton distribution functions (PDFs).

According to the QCD factorization theorem, the short-distance hard-scattering process can be separated from the long-distance nonperturbative proton structure. The separation between the perturbatively calculable hard interaction and the nonperturbative dynamics absorbed into the PDFs is introduced through the factorization scale μ_{F}. The PDFs $f_a(x_a, \mu_{\mathrm{F}})$ describe the probability density of finding a parton of flavour a carrying a momentum fraction x_a of the proton momentum when evaluated at the factorization scale μ_{F}.

Using the PDFs, the hadronic cross section for the production of a final state X can be expressed as

$$\sigma_{\mathrm{pp}\to X} = \sum_{a,b} \int_0^1 \mathrm{d}x_a \int_0^1 \mathrm{d}x_b \, f_a(x_a, \mu_{\mathrm{F}}) f_b(x_b, \mu_{\mathrm{F}}) \hat{\sigma}_{ab\to X}(x_a, x_b, \mu_{\mathrm{F}}, \mu_{\mathrm{R}}), \tag{2.7}$$

where a and b denote the two incoming partons, x_a and x_b their respective momentum fractions, and $\hat{\sigma}_{ab\to X}$ the perturbatively calculable partonic cross section. The renormalization scale μ_{R} enters through the running of the strong coupling constant.

The dependence of the PDFs on the factorization scale is described by the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi (DGLAP) equations [42–45]. These equations determine how the PDFs evolve with μ_{F} through collinear parton emissions that are perturbatively calculable in QCD.

As the scale at which the PDFs are evaluated increases, gluon radiation and parton splittings modify the momentum distribution of the proton constituents. In particular, the probability of finding low-momentum partons increases at higher scales due to successive parton splittings. As an example, the NNPDF 4.0 [11] set is shown in Figure 2.3 at two values of the PDF evolution scale, denoted by Q in the figure.

Different processes at hadron colliders probe different combinations of parton flavours and momentum fractions. Measurements of differential cross sections therefore provide important constraints on the PDFs and improve the understanding of the proton structure over a broad range of momentum fractions and energy scales.

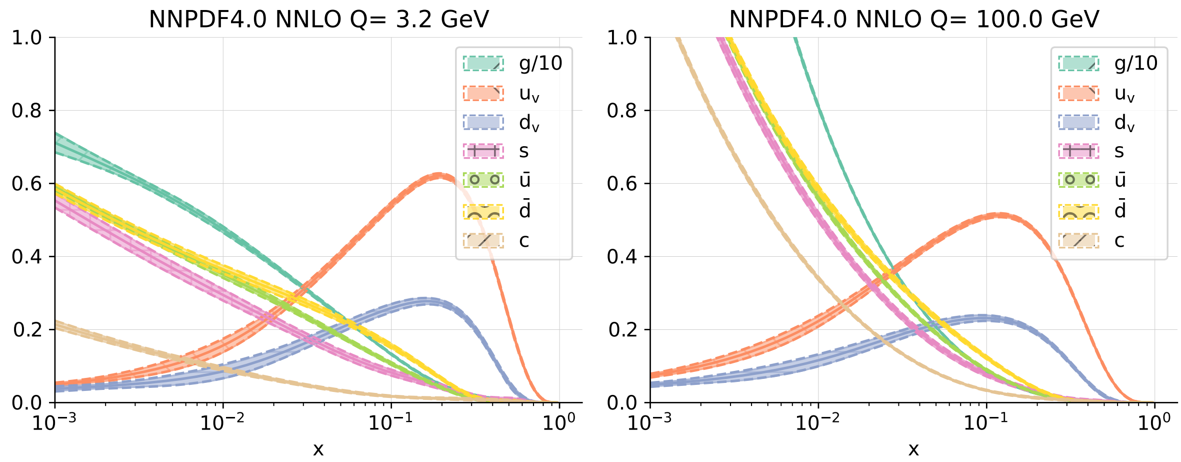


Figure 2.3: The NNPDF 4.0 parton distribution functions evaluated at the scales $Q = 3.2\,\mathrm{GeV}$ (left) and $Q = 100\,\mathrm{GeV}$ (right) [11]. The scale Q denotes the PDF evolution scale, corresponding to the factorization scale μ_F in the factorized cross section. The PDFs are shown for each parton flavour as a function of the momentum fraction x.

2.3 Triple-differential phase space for Z+jet production

The production of a Z boson in association with a jet provides direct sensitivity to the partonic structure of the proton. At leading order in perturbation theory, the kinematics of the hard scattering are fully determined by the transverse momentum and rapidities of the Z boson and the leading jet. A measurement differential in these observables therefore probes both the momentum fractions and the flavour composition of the incoming partons across a broad kinematic range.

This section introduces the rapidity observables y^* and y_b and explains why it is useful to measure the Z + jet cross section triple-differentially in these variables together with the transverse momentum of the Z boson, p_T^Z.

2.3.1 Rapidity observables

In high-energy collisions, particles are usually described using the transverse momentum p_T, which is perpendicular to the beam axis, and the rapidity for describing kinematics along the beam axis, since differences in rapidity are invariant under Lorentz boosts along the beam direction. The rapidity of a particle is defined as

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right), \tag{2.8}$$

where E is the energy and p_z is the momentum component along the beam axis.

For a $2 \to 2$ scattering process such as $\mathrm{pp} \to \mathrm{Z} + \mathrm{jet}$, the momentum fractions $x_{1,2}$ of the two incoming partons inside the protons can be expressed in terms of the final-state rapidities

y_Z and y_{jet} as

$$x_{1,2} = \frac{p_T^Z}{\sqrt{s}} \left(\exp(\pm y_Z) + \exp(\pm y_{jet}) \right) . \tag{2.9}$$

Here, $\sqrt{s}$ denotes the centre-of-mass energy of the proton–proton collision.

Rather than using the two rapidities directly, it is convenient to introduce the variables

$$y_b = \frac{1}{2} \left| y_Z + y_{jet} \right| , \tag{2.10}$$

$$y^* = \frac{1}{2} \left| y_Z - y_{jet} \right| . \tag{2.11}$$

The variable y_b corresponds to the average rapidity of the Z + jet system and therefore describes the longitudinal boost of the hard interaction. The variable y^* corresponds to half of the rapidity separation between the Z boson and the jet and is related to the scattering angle in the partonic centre-of-mass frame.

Expressing the momentum fractions in terms of y_b and y^* yields

$$x_{1,2} = \frac{2 p_T^Z}{\sqrt{s}} \cosh(y^*) \exp(\pm y_b) . \tag{2.12}$$

Since the absolute values are used in Equation 2.10 and 2.11, the sign information is lost and Equation 2.12 is invariant when swapping x_1 and x_2.

The dependence on y_b enters through the exponential factor. Large values of y_b correspond to a strongly boosted Z + jet system and therefore require a large imbalance between the two incoming parton momentum fractions. For large boosts, x_1 increases as $\exp(+y_b)$ while x_2 decreases as $\exp(-y_b)$, and vice versa. Consequently, events at large y_b probe configurations in which one parton carries a large fraction of the proton momentum, while the second parton carries only a small fraction. These regions are therefore particularly sensitive to the relative contributions of valence quarks, sea quarks, and gluons at different values of x.

The variable y^* affects both momentum fractions symmetrically through the factor $\cosh(y^*)$. As a consequence, increasing y^* requires both incoming partons to carry higher momentum fractions. High values of y^* correspond to a wide rapidity separation between the Z boson and the jet, or, in the partonic centre-of-mass frame, to a more forward scattering configuration with smaller polar angles relative to the beam axis. At fixed p_T^Z, producing such configurations requires an increased partonic centre-of-mass energy and therefore higher values of both x_1 and x_2.

The high-y^* region is therefore predominantly populated by initial states containing valence quarks in the protons, while contributions from gluons and sea quarks become less important. Therefore, the relative contribution from quark–quark initiated subprocesses increases in this region. Such channels first appear at next-to-leading order (NLO) in perturbative QCD through real-emission corrections to Z + jet production and are therefore particularly relevant at large y^*, as discussed later in Section 2.4.

The transverse momentum of the Z boson, p_T^Z, sets the overall energy scale of the hard interaction. Larger values of p_T^Z increase both momentum fractions proportionally and therefore probe the proton structure at higher scales and increased momentum fractions x. Together, the variables p_T^Z, y_b, and y^* provide a complete description of the leading-order kinematics of the hard scattering and form the basis of the triple-differential cross section measurement.

2.3.2 Binning scheme

The experimentally accessible phase space is limited by the acceptance of the Compact Muon Solenoid (CMS) detector. More information about the CMS detector is presented in Chapter 3. Muons can be reconstructed precisely only within the tracker and muon system acceptance of approximately $|\eta| < 2.4$. The definition of the pseudorapidity η in the CMS coordinate system is given in Equation 3.1. Since the Z boson is reconstructed from two muons, this also restricts the accessible rapidity range of the Z boson. Similarly, although jets can be reconstructed in the forward calorimeters up to larger pseudorapidities, precise and well-understood jet reconstruction with sufficient acceptance and resolution is only available in the central region.

Using

$$y_Z = y_b \pm y^* \,, \tag{2.13}$$

$$y_{jet} = y_b \mp y^* \,, \tag{2.14}$$

the maximum absolute rapidity reached by either the Z boson or the jet is

$$y_{max} \equiv \max(|y_Z|, |y_{jet}|) = y_b + y^* \,. \tag{2.15}$$

Consequently, simultaneously large values of y_b and y^* would require either the Z boson or the jet to be reconstructed close to the edge of the detector acceptance, where the reconstruction performance degrades rapidly. For this reason, the measurement is restricted to the phase space where $y_{max} < 3$. This requirement ensures that both the Z boson and the leading jet remain within the fiducial acceptance of the detector while still covering a broad range of partonic configurations.

The fiducial phase space is divided into five bins in y_b between 0 and 2.5 with equal width of 0.5. For each bin in y_b, up to five bins in y^* between 0 and 2.5 are defined with the same bin width satisfying the requirement on y_{max}. This results in a triangular binning scheme in the (y_b, y^*) plane, illustrated in Figure 2.4 on the left.

Since the production cross section falls steeply with increasing p_T^Z, progressively wider p_T^Z bins are required at high transverse momenta to retain sufficient statistical precision. In addition, the event yield decreases towards large values of y^*, while detector-acceptance effects become increasingly important near the edge of the accessible (y_b, y^*) phase space. To balance statistical precision and sensitivity to the underlying kinematics, three different p_T^Z binning schemes are employed across the (y_b, y^*) bins.

The central binning scheme is used in the densely populated central regions of the (y_b, y^*) plane, where sufficient event yields permit a fine granularity in p_T^Z, splitting the range from 25 GeV to 1000 GeV into 19 bins. For bins at larger y_{max}, bins in the high-p_T^Z region are merged to compensate for the reduced event yield, reducing the bin count to 17. In the

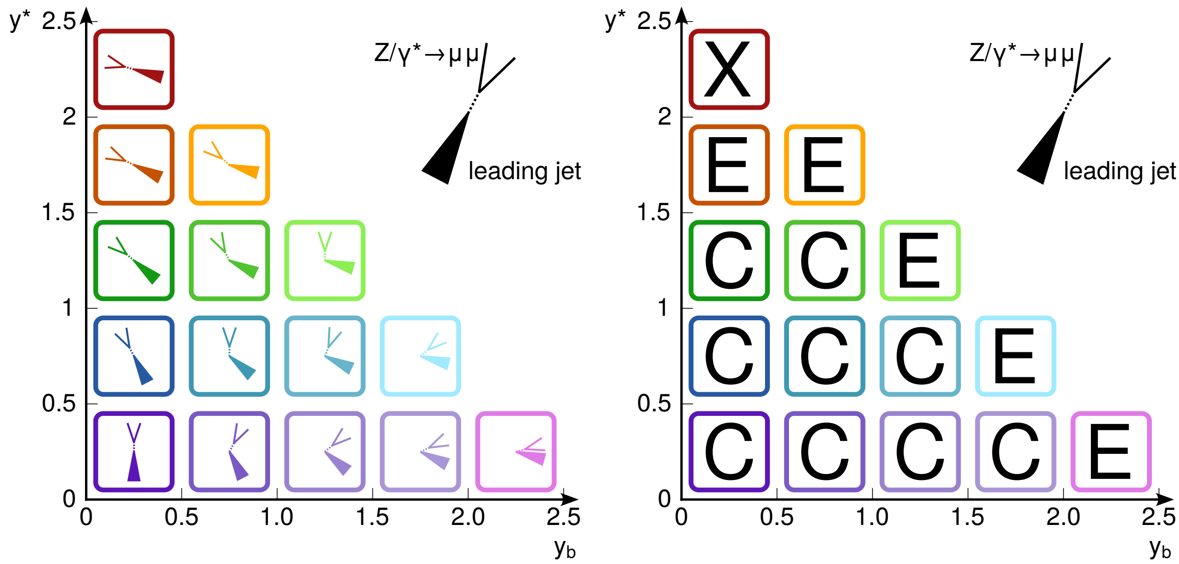


Figure 2.4: Binning scheme of the analysed phase space in the (y_b, y^*) plane. The horizontal axis corresponds to y_b and the vertical axis to y^*. Both variables are divided into bins of width 0.5, while only regions satisfying $y_b + y^* < 3$ are included in the measurement. The colours assigned to the individual bins are used consistently throughout this thesis to identify the corresponding phase-space regions. Illustrations of the final-state rapidities per (y_b, y^*)-bin are shown on the left with a horizontal beam axis. The binning schemes in p_T^Z per (y_b, y^*)-bin are given on the right, labelled "C" for central, "E" for edge, and "X" for extra, with the definitions given in Table 2.1.

Table 2.1: Bin boundaries in p_T^Z per binning scheme in GeV. The extra binning is truncated at 250 GeV due to the limited number of events. Figure 2.4 (right) shows which binning scheme is used in each (y_b, y^*)-bin.

Central (C)	25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 130, 150, 170, 190, 220, 250, 400, 1000
Edge (E)	25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 130, 150, 170, 190, 250, 1000
Extra (X)	25, 30, 40, 50, 70, 90, 110, 150, 250

high-y^* region, an additional coarse binning scheme is employed because the production cross section decreases towards high y^*, as discussed in Section 2.4. Here, only eight bins are used in the range from 25 GeV to 250 GeV. The assignment of the individual p_T^Z binning schemes to the different (y_b, y^*) bins is illustrated in Figure 2.4 on the right, while the corresponding bin boundaries are listed in Table 2.1.

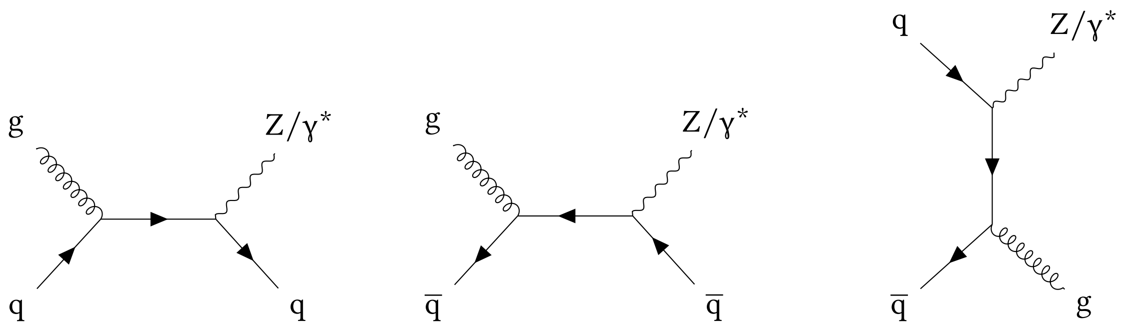
2.4 Fixed order QCD predictions for Z+jet production

The production of a Z boson in association with jets provides an important testing ground for perturbative QCD. In Subsection 2.4.1 the perturbative QCD calculations and higher-order corrections for Z + jet production are presented. A breakdown of the different initial states across the differential phase space is given in Subsection 2.4.2, illustrating the sensitivity of this measurement to different partonic channels. The comparison of these predictions to the measured triple-differential cross section is presented later in Chapter 5.

2.4.1 Perturbative QCD calculations

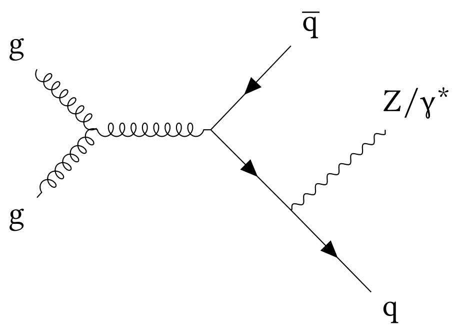
At leading order (LO), the production of a Z boson decaying into two muons and in association with one jet is described by Feynman diagrams of order $\mathcal{O}(\alpha^2, \alpha_S)$, containing two EW vertices and one QCD vertex. The Feynman diagrams for the three different initial-states available at LO, gluon–quark, gluon–antiquark and quark–antiquark, are shown in Figure 2.5. The decay of the Z/γ^* boson into a muon pair is omitted in the presented diagrams. The full leading-order cross section including all contributing subprocesses is denoted by σ_{LO}.

At NLO in perturbative QCD, additional contributions of order $\mathcal{O}(\alpha^2, \alpha_S^2)$ arise from virtual loop corrections and real-emission processes. The NLO calculations for Z + jet production were first presented in Ref. [46]. In addition to corrections to the leading-order subprocesses, some real-emission contributions introduce new partonic initial states that are absent at leading order. These subprocesses can be interpreted topologically as dijet production with EW radiation of a Z boson and contribute significantly to the large NLO QCD corrections. This includes gluon–gluon and quark–quark (antiquark–antiquark) initial-state subprocesses, as shown in Figure 2.6. In this figure the decay of the Z boson into a muon–antimuon pair is omitted, adding the missing second EW vertex when included. Especially the now available quark–quark initial state contributes significantly at high-y^* where large momentum fractions of the incoming partons are required, as described in Section 2.3.

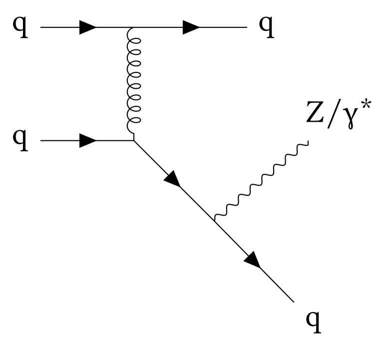


(a) gluon–quark **(b)** gluon–antiquark **(c)** quark–antiquark

Figure 2.5: Production of a Z boson in association with one jet at leading order. The decay of the Z/γ^* into a muon–antimuon pair is not shown. The Born diagrams are of order $\mathcal{O}(\alpha^2, \alpha_S)$ when the $Z/\gamma^* \to \mu^+\mu^-$ vertex is included.



(a) gluon–gluon



(b) quark–quark or antiquark–antiquark when changing the direction of the fermion propagators

Figure 2.6: Exemplary real corrections to the Z + jet production, allowing for gluon–gluon and quark–quark (antiquark–antiquark) subprocesses. The $Z/\gamma^* \to \mu^+\mu^-$ decay is not shown. When including this vertex, these diagrams are of order $\mathcal{O}(\alpha^2, \alpha_S^2)$. These subprocesses correspond topologically to dijet production with EW radiation of a Z boson and contribute significantly to the large NLO QCD corrections.

Including these higher-order contributions modifies both the normalization and the shape of the predicted cross sections. The total cross section including NLO QCD corrections for Z + jet production can be written as

$$\sigma_{\mathrm{NLO}}^{\mathrm{QCD}} = \sigma_{\mathrm{LO}} + \delta_{\mathrm{NLO}}^{\mathrm{QCD}} . \tag{2.16}$$

The NLO QCD K-factor is then defined as

$$K_{\mathrm{NLO}}^{\mathrm{QCD}} = \frac{\sigma_{\mathrm{NLO}}^{\mathrm{QCD}}}{\sigma_{\mathrm{LO}}} = 1 + \frac{\delta_{\mathrm{NLO}}^{\mathrm{QCD}}}{\sigma_{\mathrm{LO}}} . \tag{2.17}$$

Next-to-next-to-leading order (NNLO) corrections in perturbative QCD to Z+jet production have been calculated in Refs. [24–26]. At NNLO, the calculation includes double real-emission contributions, real-virtual corrections, and double-virtual loop corrections. The corresponding contributions are of order $\mathcal{O}(\alpha^2, \alpha_S^3)$. These contributions further improve the perturbative precision and reduce the residual scale dependence of the predictions. Compared to the sizeable NLO corrections, the additional NNLO contributions are significantly smaller and generally remain within the perturbative scale uncertainties of the NLO predictions. Similarly to the NLO corrections, the cross section including all QCD effects up to NNLO can be written as

$$\sigma_{\mathrm{NNLO}}^{\mathrm{QCD}} = \sigma_{\mathrm{NLO}}^{\mathrm{QCD}} + \delta_{\mathrm{NNLO}}^{\mathrm{QCD}} . \tag{2.18}$$

The NNLO K-factor as a correction to the LO cross section includes the NLO and NNLO corrections, defined as follows:

$$K_{\mathrm{NNLO}}^{\mathrm{QCD}} = \frac{\sigma_{\mathrm{NNLO}}^{\mathrm{QCD}}}{\sigma_{\mathrm{LO}}} = 1 + \frac{\delta_{\mathrm{NLO}}^{\mathrm{QCD}} + \delta_{\mathrm{NNLO}}^{\mathrm{QCD}}}{\sigma_{\mathrm{LO}}} . \tag{2.19}$$

Predictions for the triple-differential Z + jet cross section at NNLO accuracy in perturbative QCD are calculated using NNLOJet [47], interfaced via APPLfast [48] to FastNLO [49]. FastNLO uses interpolation grids for fast reevaluations of the NNLO cross section when varying the PDFs, α_S, or the factorization and renormalization scales μ_F and μ_R.

The ratio of the calculated cross sections to the LO predictions are shown in Figure 2.7 for selected bins in (y_b, y^*). They are derived using the selection requirements and binning used in the measurement (see Section 2.3 and Section 4.3). The band around the predictions includes scale variations and statistical uncertainties, as described in Section 2.7. A full breakdown for all bins in (y_b, y^*) is given in the appendix in Figure A.1.

At low y^*, the NLO and NNLO K-factors reach values of about 1.5 at large transverse momenta. Perturbative convergence between the NLO and NNLO predictions is observed only for p_T^Z above about 50 GeV in the low-y^* region, where the scale uncertainties cover the difference between the NLO and NNLO results. In contrast, the high-y^* bin exhibits very large K-factors, reaching values of up to about 12. This behaviour can be attributed to the PDF effects discussed in Section 2.3 that suppress the LO contribution in this region. As a result, important configurations receive their first substantial contribution only at NLO, so that the NLO prediction is effectively only of LO quality for these configurations.

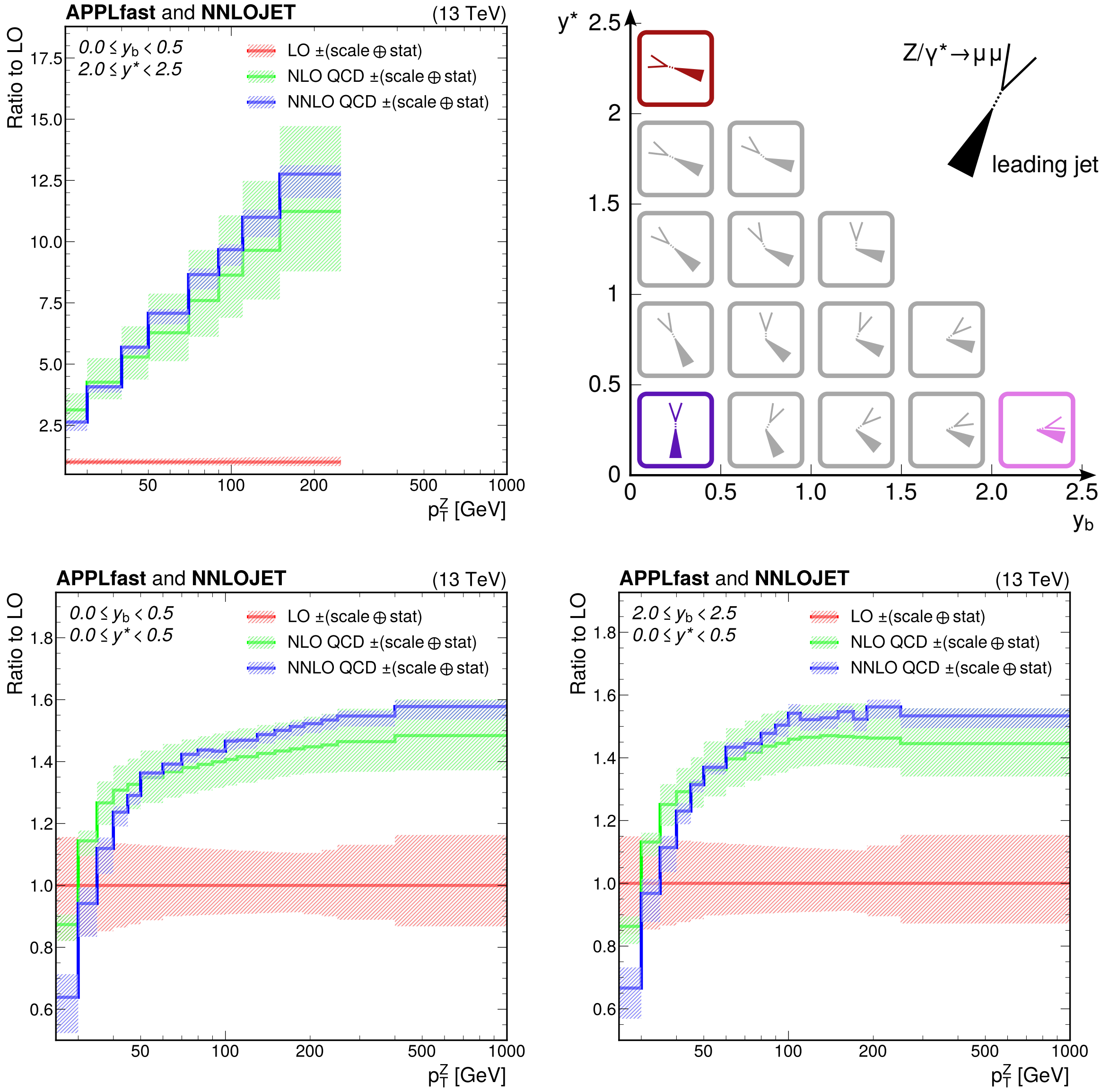


Figure 2.7: Higher order corrections for Z + jet production in perturbative QCD as a ratio to the LO predictions. The NLO and NNLO K-factors are shown as a function of p_T^Z for the central rapidity region (bottom left), high-y_b (bottom right) and high-y^* bin (top left), as illustrated in the top right. The uncertainty band includes statistical uncertainties in the numerical integration and scale variations.

2.4.2 Initial-state parton contributions at NNLO

The dependence of the cross section on the initial-state parton flavours is illustrated by looking at the subprocess decomposition in different regions of the phase space. Different partonic initial states contribute with varying strengths across the phase space, particularly as a function of the rapidity variable y^* as explained in Section 2.3 and the scale of the hard process, represented by $p_{\mathrm{T}}^{\mathrm{Z}}$.

A breakdown of the initial-state contributions to the full NNLO cross section is shown in Figure 2.8 for representative central, high-y^*, and high-y_{b} regions. The gluon–quark (shown in orange), gluon–antiquark (red), and quark–antiquark (grey) initial states are already present at leading order and dominate in most kinematic regions. The quark–quark initial state (purple), which first appears at NLO, becomes increasingly important in the high-y^* region. As discussed in Section 2.3, large values of y^* require both incoming partons to carry large momentum fractions $x_{1,2}$. In this region, the proton structure is increasingly dominated by valence-quark contributions, enhancing the relative contribution of quark–quark initiated subprocesses. The subprocess contributions are calculated using the PDF4LHC21 parton distribution functions [50]. A full breakdown for all bins in (y_{b}, y^*) is shown in the appendix in Figure A.2.

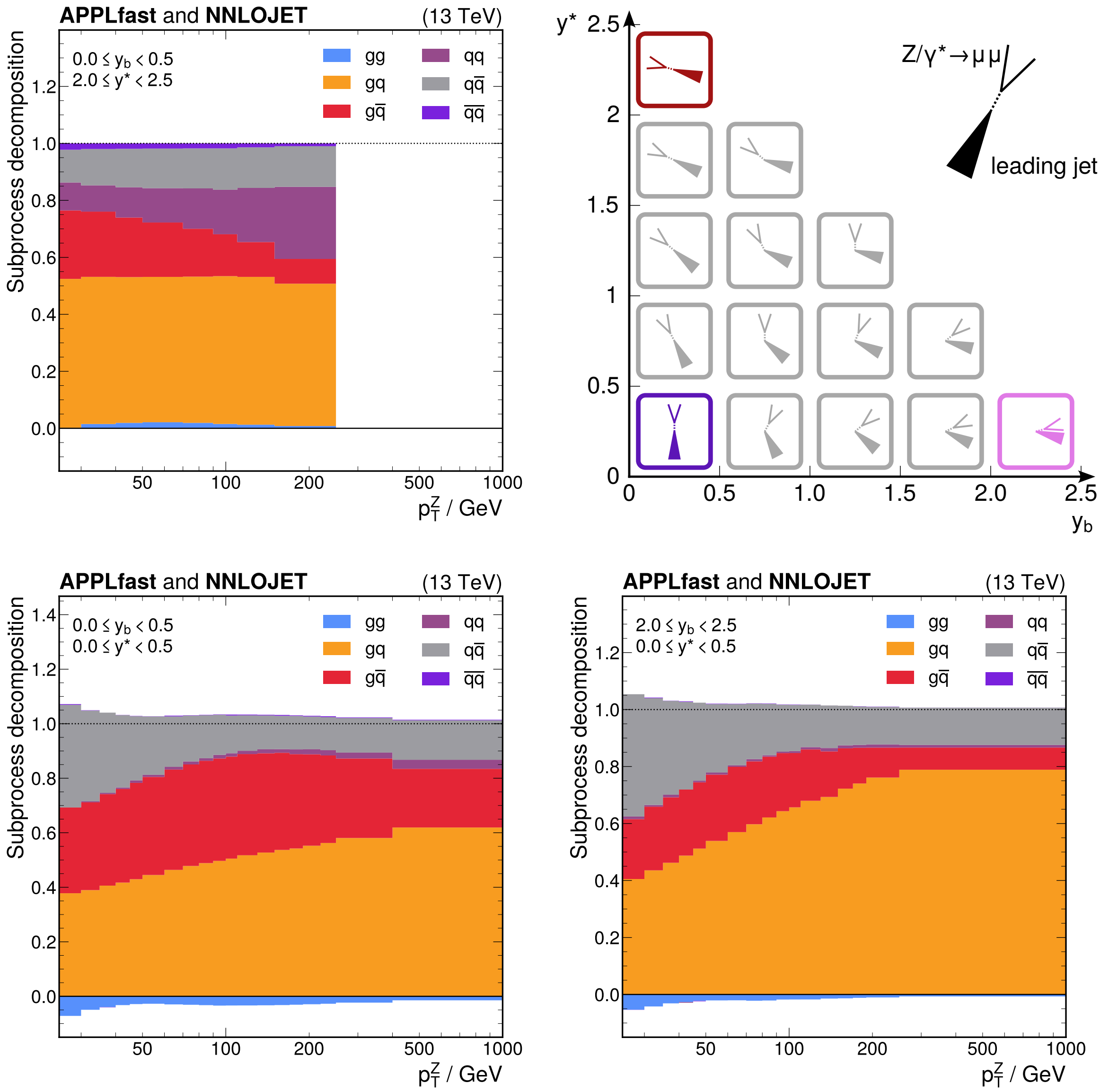


Figure 2.8: Subprocess contributions to the full NNLO perturbative QCD calculations using the PDF4LHC21 PDF set. The subprocess decomposition is shown in three different rapidity regions: a central back-to-back topology (lower left), a highly boosted region (lower right), and a large rapidity-separation region (upper left). An illustration of the three selected (y_b, y^*) bins with respect to the full analysed phase space is shown in the top right. The gluon–quark (orange), gluon–antiquark (red) and quark–antiquark (grey) initial states are already available at LO. At large values of y^* a sizeable contribution from the quark–quark initial state (purple), which first appears at NLO, is present. Individual subprocess contributions can become negative due to interference terms involving virtual loop corrections. Only the sum of all subprocesses corresponds to a physical observable.

2.5 Electroweak corrections

Higher-order electroweak (EW) corrections to vector-boson production in association with jets have been studied extensively in recent years [51–53]. They become increasingly important at the energy scales reached at the LHC, in particular in the tails of differential distributions. In these regions, EW corrections can reach the level of several percent, up to about 10% or more, and therefore become comparable in size to residual higher-order QCD effects.

For the production of a Z boson in association with a jet, where the Z boson decays into a pair of muons, the dominant Born contribution is of order $\mathcal{O}(\alpha^2, \alpha_S)$, while the corresponding NLO EW corrections are of order $\mathcal{O}(\alpha^3, \alpha_S)$. The cross section including NLO electroweak effects can be written as

$$\sigma_{\mathrm{EW}}^{\mathrm{NLO}} = \sigma_{\mathrm{LO}} + \delta_{\mathrm{NLO}}^{\mathrm{EW}}, \tag{2.20}$$

where σ_{LO} denotes the leading-order cross section and $\delta_{\mathrm{NLO}}^{\mathrm{EW}}$ the NLO EW correction.

The NLO electroweak contribution receives virtual corrections from one-loop diagrams and real corrections from processes with an additional emitted photon. At high energies, where the characteristic scale of the hard interaction is much larger than the masses of the weak gauge bosons, the virtual weak corrections are dominated by logarithmically enhanced terms of the form $\log^2(Q^2/m_{\mathrm{V}}^2)$, with $\mathrm{V} = \mathrm{W}, \mathrm{Z}$. These terms are referred to as electroweak Sudakov logarithms and lead to sizeable negative corrections in the high-$p_{\mathrm{T}}^{\mathrm{Z}}$ region. Real photon radiation can distort reconstructed observables through momentum loss from the charged leptons. The Feynman diagrams of three exemplary corrections are shown in Figure 2.9, including two loop diagrams and a real photon emission in the final state.

Using MadGraph5_aMC@NLO [53] the full fixed-order NLO electroweak corrections for Z+jet production can be calculated. The fixed-order NLO electroweak calculations include the virtual and real-emission contributions required for an infrared-finite prediction. Finite-width effects of intermediate massive gauge bosons are treated using the complex-mass scheme.

The EW K-factor is defined per measurement bin as

$$K_{\mathrm{NLO}}^{\mathrm{EW}} = \frac{\sigma_{\mathrm{NLO}}^{\mathrm{EW}}}{\sigma_{\mathrm{LO}}} = 1 + \frac{\delta\sigma_{\mathrm{NLO}}^{\mathrm{EW}}}{\sigma_{\mathrm{LO}}}. \tag{2.21}$$

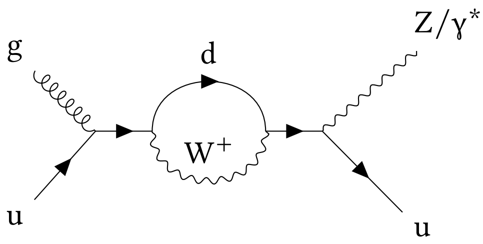
Assuming a factorisation of the EW corrections towards higher orders in QCD calculations, a multiplicative correction to the NNLO prediction in perturbative QCD is obtained as

$$\sigma^{\mathrm{QCD}\times\mathrm{EW}} = \sigma_{\mathrm{NNLO}}^{\mathrm{QCD}}\, K_{\mathrm{NLO}}^{\mathrm{EW}}, \tag{2.22}$$

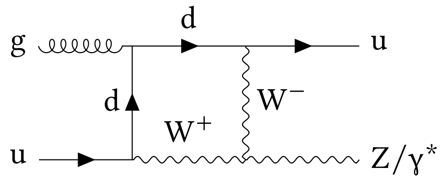
and used as the nominal correction. Alternatively, the additive combination of the NLO EW correction and the NNLO cross section prediction from perturbative QCD is defined as

$$\sigma^{\mathrm{QCD}+\mathrm{EW}} = \sigma_{\mathrm{NNLO}}^{\mathrm{QCD}} + \delta_{\mathrm{NLO}}^{\mathrm{EW}}. \tag{2.23}$$

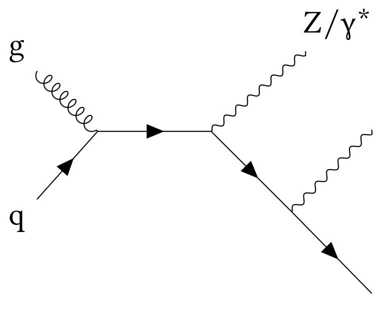
To account for the uncertainty associated with the factorisation approximation, half the difference between the multiplicative and additive prescriptions is taken as a systematic uncertainty in the EW correction.



(a) weak self-energy correction



(b) EW box diagram



(c) real photon emission

Figure 2.9: Representative EW corrections to Z + jet production at next-to-leading order. Shown are virtual weak corrections, such as self-energy and box diagrams, as well as real photon emission. The decay $Z/\gamma^* \to \mu^+\mu^-$ is not shown explicitly. Including the leptonic decay, these contributions are of order $\mathcal{O}(\alpha^3, \alpha_S)$.

The calculation is performed using the PDF4LHC21 parton distribution functions. Additional contributions from initial states including photons, e.g., photon–quark scattering, might also have an impact on the calculated cross section. A cross-check using the LUXqed17 [54, 55] PDF set for the photon content of the proton and the PDF4LHC15 [56] PDF at NNLO accuracy yield the same EW correction factors as calculated with PDF4LHC21.

In addition to the full fixed-order prediction, approximate EW corrections based on Sudakov logarithms are also considered. These approximate corrections are obtained using an implementation in the SHERPA event generator [57]. These corrections capture the dominant logarithmically enhanced virtual weak contributions at high energies.

In this thesis, the Sudakov corrections are not used as the nominal EW correction, but only for comparison with the full fixed-order calculation. At large transverse momenta, the full NLO EW correction is generally observed to be larger in magnitude than the Sudakov approximation. Differences are expected, since the Sudakov approximation includes only the dominant logarithmically enhanced virtual weak contributions, while the full fixed-order calculation additionally contains non-logarithmic terms, real-emission effects, and the full dependence on the fiducial phase-space definition.

Differences are also observed between the Sudakov approximations obtained from the LO and NLO QCD predictions, particularly in the high-y^* region. As discussed in Section 2.4, this region receives enhanced contributions at NLO in perturbative QCD. Since the EW Sudakov corrections depend on the flavour and kinematics of the underlying partonic subprocesses, the changing subprocess composition affects the resulting correction factors. This behaviour demonstrates that the assumption of a fully factorizing EW correction with respect to higher-order perturbative QCD corrections is only approximate. Due to the large NLO QCD corrections, noticeable differences are observed between the additive and multiplicative prescriptions for the full NLO EW correction, especially at high-y^*. The resulting uncertainty estimate also covers the observed difference between the Sudakov approximations obtained from the LO and NLO QCD predictions and the full NLO EW calculation.

The final electroweak correction factors used in this analysis are shown in Figure 2.10 for the central rapidity, the high-y^*, and the high-y_b bin. A full breakdown for all bins in (y_b, y^*) is shown in the appendix in Figure A.3. Displayed are the multiplicative NLO EW correction, the relative additive correction with respect to the NNLO QCD prediction, and the Sudakov approximations. Half the difference between the multiplicative and additive prescriptions is used as the systematic uncertainty is the EW correction.

The correction factors reach up to 15% at large transverse momenta in the central rapidity bin, while decreasing to 1–2% at low p_T^Z. At large y_b, the corrections are similar, decreasing at high-p_T^Z due to the merged binning scheme. Differences between the multiplicative and additive correction reach up to 7% at high p_T^Z or in the high-y^* region around a transverse momentum of 20 GeV. In some regions of the phase space the difference in the EW correction reaches more than 10%, as shown in the full (y_b, y^*) breakdown in Figure A.3.

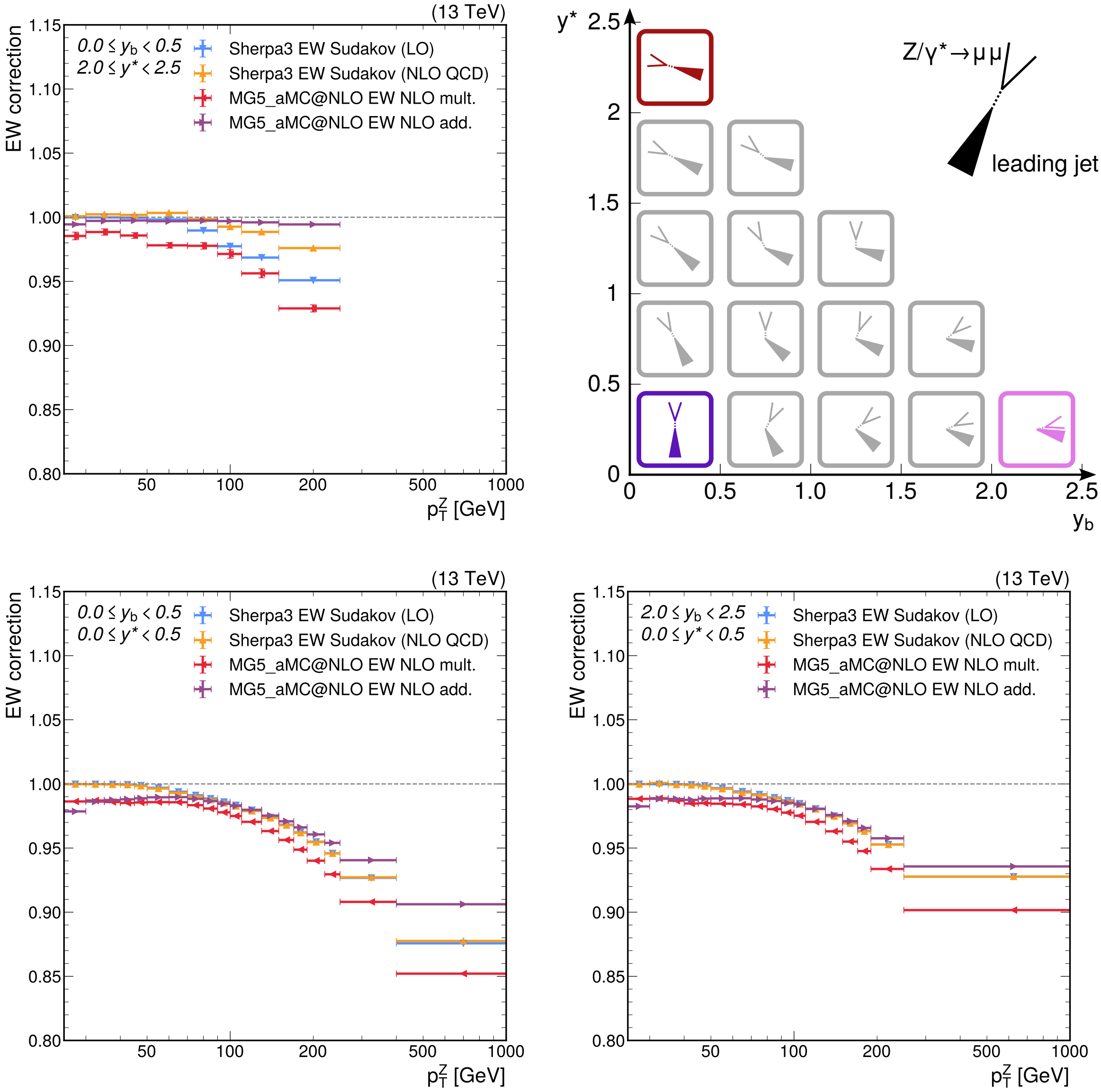


Figure 2.10: Electroweak correction factors as a function of p_T^Z in three representative bins in (y_b, y^*) as illustrated in the top right. Shown are the multiplicative NLO EW correction, the relative additive correction with respect to the NNLO QCD prediction, and the Sudakov approximations based on a calculation at LO and NLO in perturbative QCD. The nominal EW correction is obtained using the multiplicative NLO correction factor, while half the difference to the additive prescription is taken as the associated systematic uncertainty.

2.6 Nonperturbative effects

In order to fully describe a proton–proton collision, perturbative approximations of the hard interaction are not sufficient. An illustration of additional activity in a proton–proton interaction is shown in Figure 2.11 for the example of Drell–Yan lepton-pair production. The small red circle shows the hard interaction of the two incoming partons producing the lepton pair. Additionally, the initial-state partons radiate gluons, leading to parton showers (PS). If the final state of the hard scattering also contains coloured partons, as in Z + jet production, final-state radiation (FSR) occurs as well. The parton shower describes the successive emission of additional partons from coloured partons in the initial and final state, approximating higher-order QCD radiation in the soft and collinear limits [58, 59]. PS effects are simulated by Monte Carlo (MC) generators such as PYTHIA [60], HERWIG [61], and SHERPA [62].

The perturbative shower evolution terminates once the evolution scale approaches the hadronization scale, typically of the order of a few GeV. At this scale, the strong coupling becomes large and perturbation theory is no longer applicable. Phenomenological models are therefore required to describe the transition from coloured partons to colour-neutral hadrons, illustrated by the colourless clusters shown as white ellipses decaying into hadrons, shown as yellow circles in Figure 2.11. This process is referred to as hadronization. Since hadronization occurs at low momentum scales, it cannot be calculated from first principles in perturbative QCD and is instead described by phenomenological models.

In addition to radiation associated with the primary hard scattering, further activity can arise from the remnants of the two incoming protons. Since protons are composite objects, the partons that do not participate in the primary hard interaction remain as coloured beam remnants. Moreover, additional pairs of partons within the same proton–proton collision can undergo secondary semi-hard scatterings. These interactions are referred to as multiple parton interactions (MPI). Together with beam-remnant fragmentation and additional soft radiation, they form the underlying event (UE). The UE contributes extra particles and transverse momentum to the final state and therefore affects reconstructed jets and event-level observables. A sketch including MPI from UE contributions is shown in the appendix in Figure A.4. The modelling of MPI is likewise phenomenological, with model parameters tuned to measurements of underlying-event observables in data.

The perturbative shower evolution, hadronization, and modelling of the underlying event are implemented differently in different event generators. In PYTHIA 8, the default shower is ordered in transverse momentum and interleaves initial-state radiation, final-state radiation, and multiple-parton interactions in a common sequence of decreasing transverse-momentum scales [63]. Hadronization in PYTHIA 8 is described using the Lund string model [64–66]. In HERWIG 7, the default shower is angular ordered, such that subsequent emissions occur at decreasing opening angles [67, 68]. Hadronization in HERWIG 7 is based on the cluster model, in which colour-singlet clusters are formed from the partonic final state and subsequently decay into hadrons [69]. SHERPA uses a dipole shower based on Catani–Seymour factorisation, where emissions are generated from emitter–spectator dipoles rather than from independently evolving partons [70]. Its hadronization model is based on cluster fragmentation [71]. Both PYTHIA and SHERPA use the impact parameter model from Refs. [63, 72, 73] for modelling MPI, while HERWIG employs a multiple-scattering model based on the impact-parameter overlap of

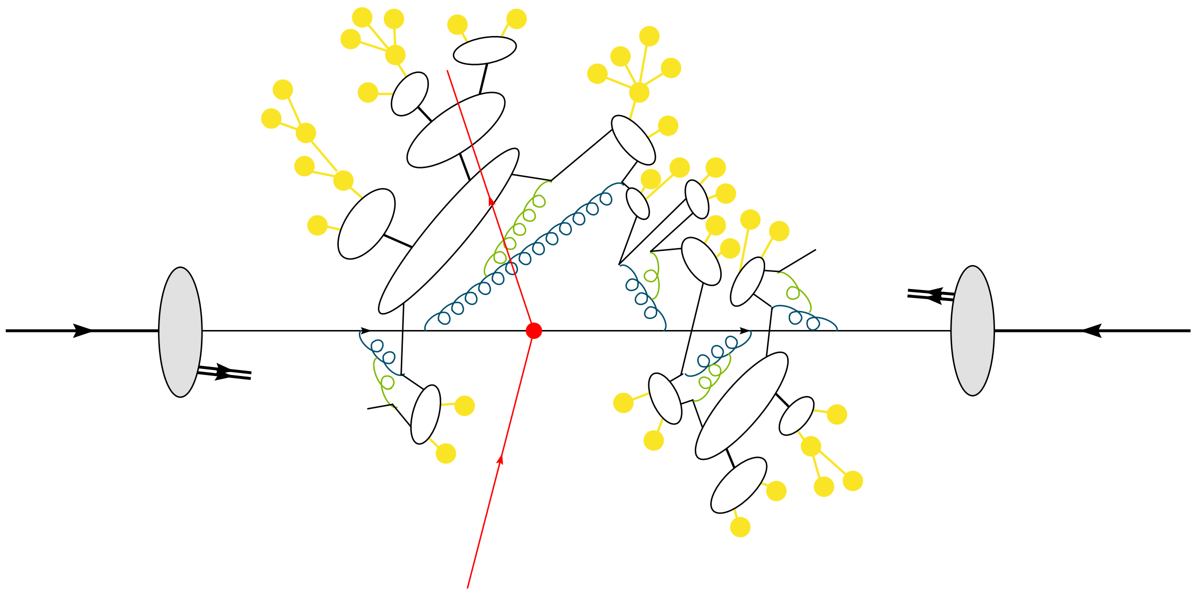


Figure 2.11: Sketch of nonperturbative effects in a proton–proton collision. The two protons are depicted as grey blobs on the left and right, and the hard interaction of the incoming partons is shown in red, producing two high-p_T leptons. Initial-state radiation is shown as gluons radiated from the incoming partons, leading to parton showers (PS). The subsequent hadronization is illustrated according to the cluster model used in HERWIG. The partons are combined into colourless clusters, shown as white ellipses, which decay into hadrons, shown as yellow circles, that can decay further. Additional contributions from the underlying event through the proton remnants are not shown. Sketch courtesy of S. Gieseke.

the colliding protons [74].

For perturbative calculations at LO, the merging of higher-multiplicity final states and subsequent matching to parton showers is possible, e.g. using the MLM [75] or CKKW [76] formalism. Matching perturbative calculations at NLO accuracy in QCD to parton showers is also possible, e.g. using the MC@NLO [77], MEPS@NLO [78], and POWHEG [79–81] methods. Even though recent progress has been made in matching calculations at NNLO accuracy to PS with the MiNNLO$_{\mathrm{PS}}$ [82, 83] formalism for some processes, this is not yet possible for Z + jet production. Therefore, the effect of nonperturbative (NP) corrections is estimated from simulations based on LO and NLO calculations in perturbative QCD. These corrections are assumed to factorize with respect to higher-order perturbative QCD calculations.

NP corrections are derived using HERWIG 7.2.3 [61] and SHERPA 2.2.15 [84]. They are derived by comparing simulations at the stable-particle level, including hadronization and MPI, to predictions at the parton level based on LO and NLO calculations matched to parton showers. Additionally, a multileg-merged prediction at NLO accuracy in QCD is used with SHERPA. The results have been published in Ref. [32], which also includes NP corrections for dijet production.

The nonperturbative correction factor is defined as the ratio of particle-level predictions including hadronization and MPI to predictions where these effects are disabled,

$$c_{\mathrm{NP}} = \frac{\sigma(\mathrm{ME} + \mathrm{PS} + \mathrm{Had} + \mathrm{MPI})}{\sigma(\mathrm{ME} + \mathrm{PS})}\,, \tag{2.24}$$

where ME denotes the matrix-element calculation, PS the perturbative parton shower evolution, Had the hadronization model, and MPI the model for multiparton interactions.

The calculated correction factors are shown in Figure 2.12 for the central rapidity, high-y_{b} and high-y^* region. The resulting nonperturbative corrections exhibit a significant dependence on y^*, while only a weak dependence on y_{b} is observed. Furthermore, differences are found between corrections derived from LO and NLO matrix-element calculations, indicating that the nonperturbative corrections do not perfectly factorize with respect to the perturbative order of the hard process. To account for this residual dependence, an envelope is constructed from the three NP correction factors obtained using different event generators and matching schemes based on HERWIG 7 and SHERPA, using baseline perturbative predictions with at least NLO accuracy. Statistical uncertainties of the individual correction factors are included in this construction. The centre of the envelope is used as the nominal NP correction, while the upward and downward extents define the corresponding uncertainty. For comparison, the LO-based correction factors from both MC generators are also shown, but they are not used in the construction of the final correction. A full breakdown of all bins in (y_{b}, y^*) is given in the appendix in Figure A.5.

The central NP correction is close to unity above 50 GeV across the whole (y_{b}, y^*)-plane. Negative corrections up to 4% at low transverse momenta are observed in the low-y^* region, increasing up to 12% in the high-y^* bin.

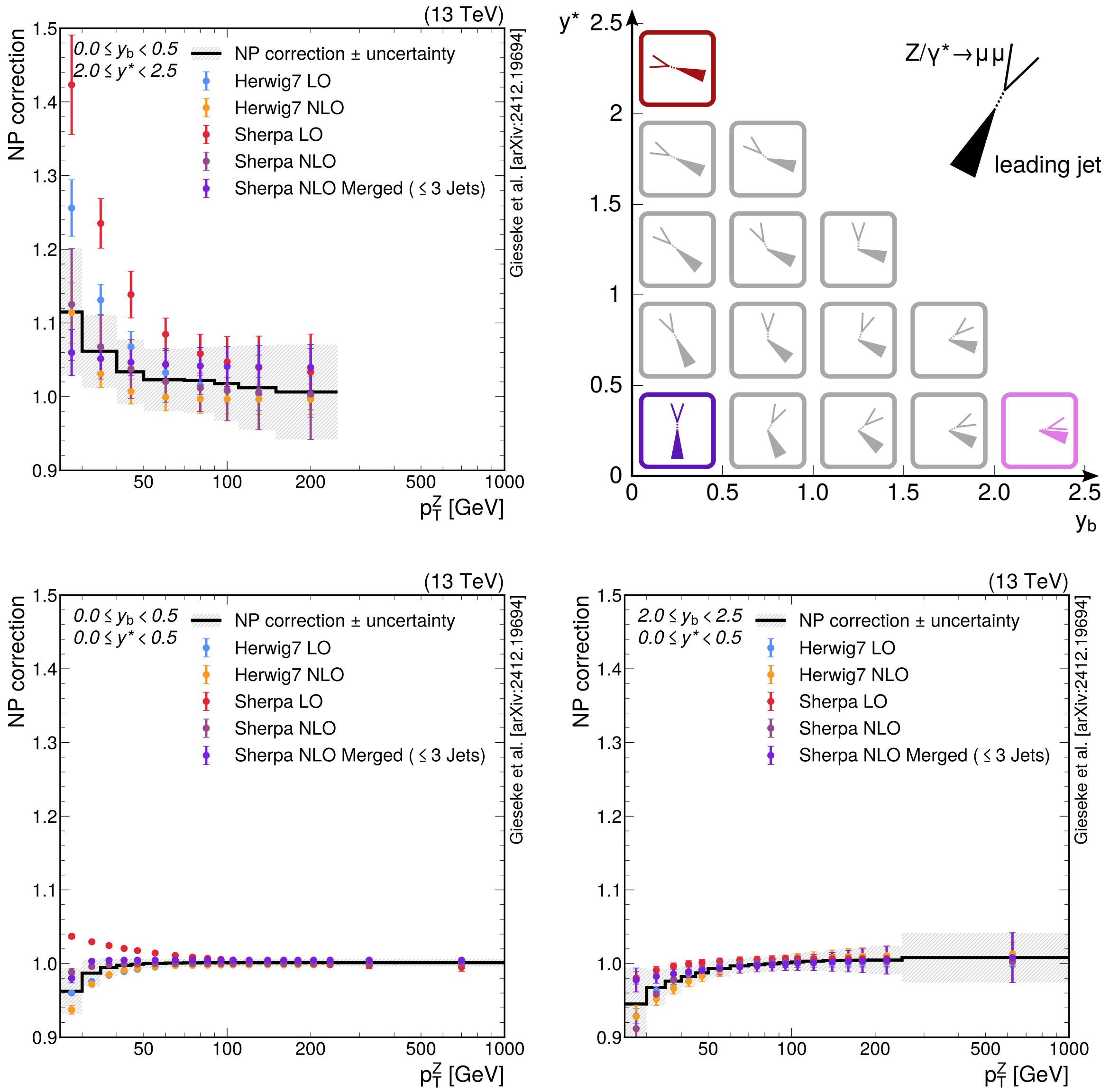


Figure 2.12: Nonperturbative corrections for Z + jet production as a function of $p_{\mathrm{T}}^{\mathrm{Z}}$ in three exemplary bins in (y_{b}, y^{*}), as indicated in the top right. The nominal correction and its associated uncertainty are obtained from the envelope of predictions with NLO accuracy in the matrix-element calculations. Predictions using matrix-element calculations at LO are shown for reference only.

2.7 Uncertainties in the NNLO predictions

The final theory predictions at NNLO accuracy in perturbative QCD are obtained by combining the fixed-order calculations described in Section 2.4, evaluated with the PDF4LHC21 PDF set, with the higher-order EW corrections described in Section 2.5 and the NP correction factors described in Section 2.6. The total uncertainty in these predictions is constructed from the individual uncertainty contributions associated with the fixed-order calculation, the PDFs, the EW corrections, and the NP correction factors.

Although the factorization and renormalization scales are introduced as part of the perturbative framework, physical observables are formally independent of their choice when computed to all orders in perturbation theory. However, fixed-order calculations are truncated at finite order and therefore retain a residual dependence on these scales. This remaining dependence is used to estimate uncertainties from missing higher-order contributions.

The corresponding uncertainties are typically evaluated by varying the factorization scale μ_{F} and renormalization scale μ_{R} independently by factors of 0.5 and 2 around their nominal value μ_0. For the calculations at NNLO accuracy presented in Section 2.4, both μ_{F} and μ_{R} are set to $p_{\mathrm{T}}^{\mathrm{Z}}$. The extreme combinations of the variations $(\mu_{\mathrm{R}}/\mu_0, \mu_{\mathrm{F}}/\mu_0) = (0.5, 2.0)$ and $(2.0, 0.5)$ are excluded. The envelope of the resulting predictions is then taken as an estimate of the perturbative uncertainty.

In addition to the perturbative uncertainties associated with the scale choices, PDF determinations provide uncertainty estimates related to the experimental and theoretical uncertainties entering the PDF determination. Two main approaches are commonly used to represent these uncertainties. In the Hessian approach [85, 86], the uncertainty is encoded through a set of orthogonal eigenvector variations corresponding to independent directions in the parameter space of the fit. Alternatively, the Monte Carlo approach [87, 88] represents the probability distribution of the fitted parameters by an ensemble of replicas generated from pseudo-data fluctuated within the experimental uncertainties. Statistical estimators evaluated over these replicas provide the corresponding central values and uncertainties of the PDFs. For the estimation of the uncertainty in the PDF4LHC21 PDF, a reduced set of 40 Hessian eigenvectors are used.

The uncertainty contributions for the cross section predictions at NNLO are shown in Figure 2.13 for representative central, large-y_{b}, and large-y^* bins. A breakdown of the uncertainty contributions across all (y_{b}, y^*) bins is shown in Figure A.6.

Numerical uncertainties from the integration of the fixed-order predictions are taken into account and remain below 1% in most bins of the analysed phase space. Scale variations provide the largest uncertainty contribution in parts of the phase space, ranging from about 1% to 18%, with the largest values observed at low transverse momenta. For the PDF4LHC21 PDF set, uncertainties are typically around 2% to 3% across the phase space and become dominant in the central rapidity region for transverse momenta between 50 GeV to 300 GeV. Nonperturbative corrections contribute uncertainties between 1% and 8%. These are particularly relevant at large y^*, where they provide a sizeable contribution to the total uncertainty in addition to the scale variations. At large transverse momenta, uncertainties in the higher-order EW corrections become noticeable and reach up to about 3%.

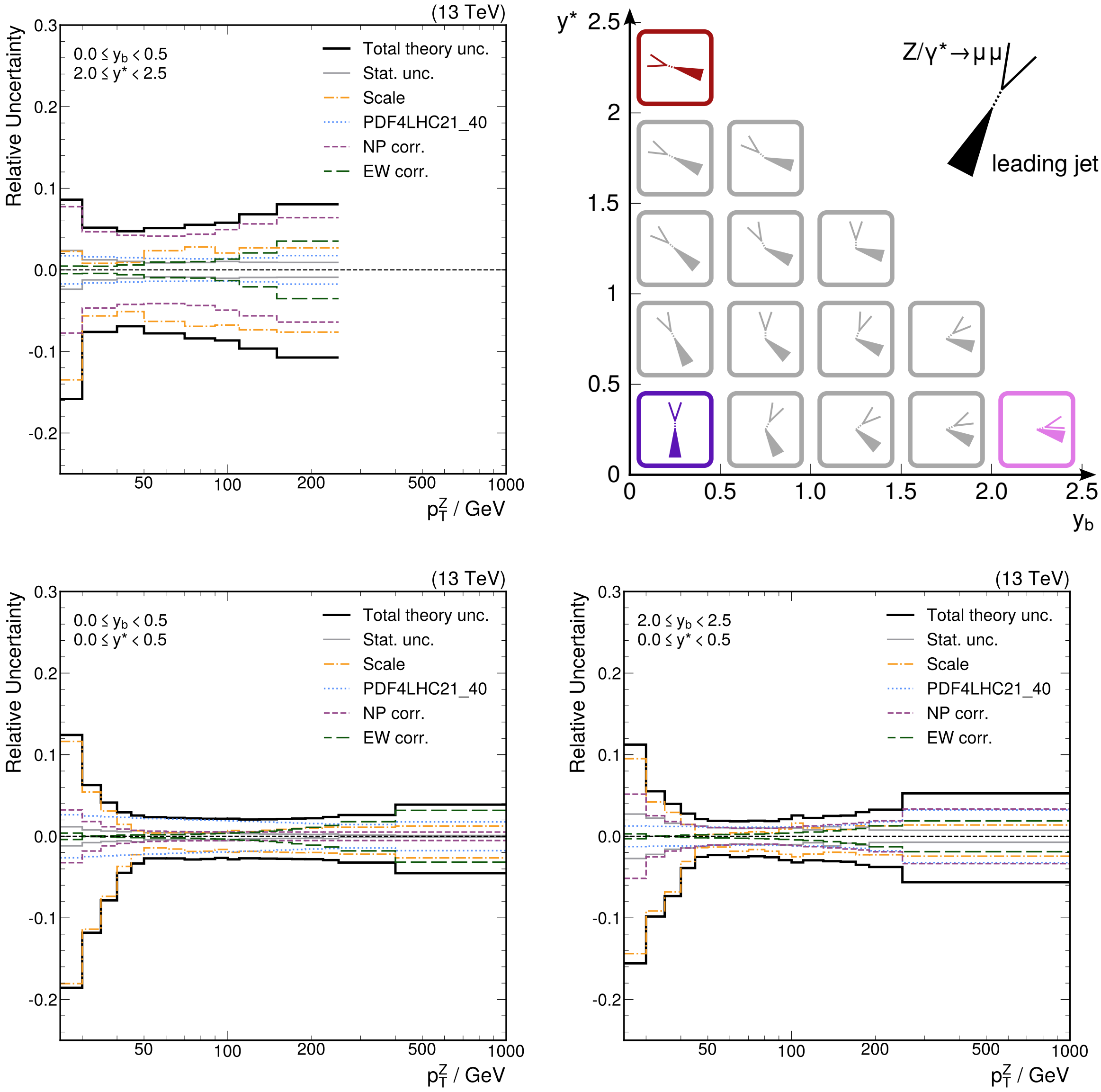


Figure 2.13: Uncertainties in the final NNLO theory predictions, including numerical uncertainties from the integration, scale variations, uncertainties in the PDF4LHC21 PDF set, NP effects, and NLO EW corrections. The uncertainties are shown as a function of p_T^Z in the central (bottom left), large-y_b (bottom right), and large-y^* (top left) bins, as indicated in the top right.

CHAPTER 3

The Large Hadron Collider and the CMS detector

One of the central institutions in high energy physics (HEP) is the European Organization for Nuclear Research, or in French, *Conseil Européen pour la Recherche Nucléaire (CERN)*. It is located west of Geneva, on the border between Switzerland and France. Founded in 1954, many important discoveries in particle physics have been made there, probing the fundamental building blocks of matter and their interactions with unprecedented precision.

Achieving a better understanding of nature requires investigating particles at tiny scales and therefore at very high energies. Therefore, larger and more powerful particle accelerators have been built at CERN over the last decades. Currently, the largest and most powerful particle accelerator in the world is the Large Hadron Collider (LHC), described in Section 3.1. It started operation in 2010 with the LHC Run 1, which lasted until 2013 at centre-of-mass energies of 7 TeV and 8 TeV [89]. After a long shutdown in which upgrades were installed, LHC Run 2 took place from 2015 to 2018, increasing the centre-of-mass energy to 13 TeV [90]. The second long shutdown took place after Run 2 and included upgrades enabling the LHC to reach a centre-of-mass energy of 13.6 TeV during Run 3, which lasts from 2022 to 2026. A third long shutdown will begin in 2026 for installing the upgrades for the High-Luminosity LHC [91].

The particles accelerated at the LHC collide at four separate collision points along the accelerator ring. At each of these collision points, one of the four main particle detectors of the LHC is located. One of the general-purpose detectors is the Compact Muon Solenoid (CMS) detector, introduced in Section 3.2. It is designed to measure the particles produced in proton–proton as well as heavy-ion collisions at the LHC, together with their properties. In this thesis, data recorded by the CMS detector at a centre-of-mass energy of 13 TeV during the LHC Run 2 from 2016 to 2018 are used.

The reconstruction of final-state particles produced in collisions at the CMS detector is discussed in Section 3.3. This section mainly focuses on the particle-flow (PF) algorithm and jet reconstruction.

3.1 The Large Hadron Collider

The LHC is a proton–proton collider designed to reach centre-of-mass energies of up to 14 TeV. In addition to proton–proton collisions, it is operated in heavy-ion mode, usually with lead nuclei, at centre-of-mass energies of up to 5.5 TeV per nucleon pair. The following overview is largely based on the LHC design report published in 2004 [92].

The LHC is located on the Franco-Swiss border near Geneva in a circular tunnel with a circumference of 27 km, about 100 m underground. It houses multiple experiments along four interaction points (IPs), where the particle beams collide. The two general-purpose detectors, A Toroidal LHC Apparatus (ATLAS) and CMS, are located at two of these IPs. The other two IPs host the LHC beauty (LHCb) and A Large Ion Collider Experiment (ALICE) detectors, which are mainly designed to study beauty quark and heavy-ion physics, respectively.

This section is structured as follows. The injection chain of the protons before they enter the LHC is described in Subsection 3.1.1. In Subsection 3.1.2, the characteristics of proton–proton collisions at the LHC are described as they take place e.g. at the IP where the CMS detector is located.

3.1.1 The LHC injection chain

Before injection into the LHC, protons undergo several stages of pre-acceleration to 450 GeV. The structure of the CERN accelerator complex during Run 2 is shown in Figure 3.1.

Until the end of Run 2, the proton injection chain started with the linear accelerator LINAC 2 [94]. Hydrogen gas was ionized by passing it through an electric field to strip the electrons from the protons. The resulting protons were then injected into LINAC 2 and accelerated to an energy of 50 MeV with the use of radio frequency cavities. The LINAC 2 was replaced by LINAC 4 during the long shutdown 2 of the LHC in 2020. LINAC 4 accelerates negatively charged Hydrogen ions compared to the positively charged Hydrogen ions used in LINAC 2. The negatively charged ions are stripped of their electrons before being injected into the next accelerator.

After the LINAC, the protons are injected into the Proton Synchrotron Booster, where they are accelerated to 1.4 GeV and subsequently delivered to the Proton Synchrotron. The Proton Synchrotron is CERN's first synchrotron and started operations in 1959 [95]. Here, the protons are accelerated to energies up to 26 GeV, before being delivered to the Super Proton Synchrotron (SPS) [96]. The SPS started operating in 1976 and enabled the discovery of the W and Z bosons at the UA1 and UA2 experiments [97–100] in a proton-antiproton collision configuration. In the SPS, the protons are accelerated to energies up to 450 GeV and then delivered to the LHC. The full injection into the LHC is described in detail in Ref. [101].

During the Run 2 operations, a nominal fill into the LHC consisted of a total of 2220 bunches per beam in 2016 and 2556 bunches per beam in 2017 and 2018, spaced 25 ns apart from each other [90]. In the LHC, the final acceleration stages take place up to 6.5 TeV per beam, resulting in collisions at a centre-of-mass energy of 13 TeV. Since the start of Run 3 in 2022, the LHC has reached beam energies of 6.8 TeV, resulting in centre-of-mass energies of 13.6 TeV.

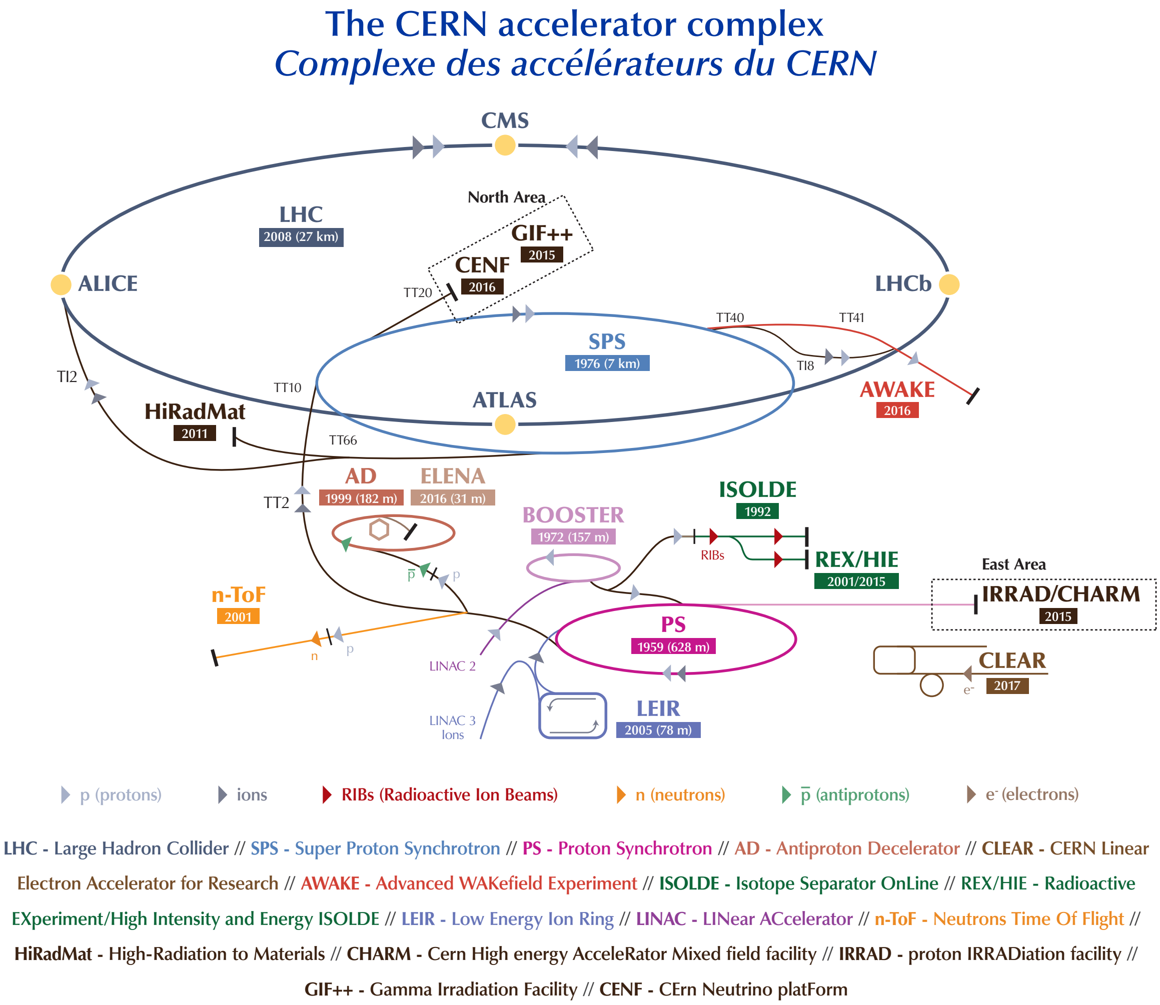


Figure 3.1: The CERN accelerator complex as of August 2018 at the end of the LHC Run 2. Protons are first accelerated in the linear accelerator LINAC 2. They are then injected into the Booster, followed by the Proton Synchrotron. In the Super Proton Synchrotron the protons are accelerated to an energy of 450 GeV, before finally being injected into the LHC, where they reach energies up to 6.5 TeV. Illustration from Ref. [93].

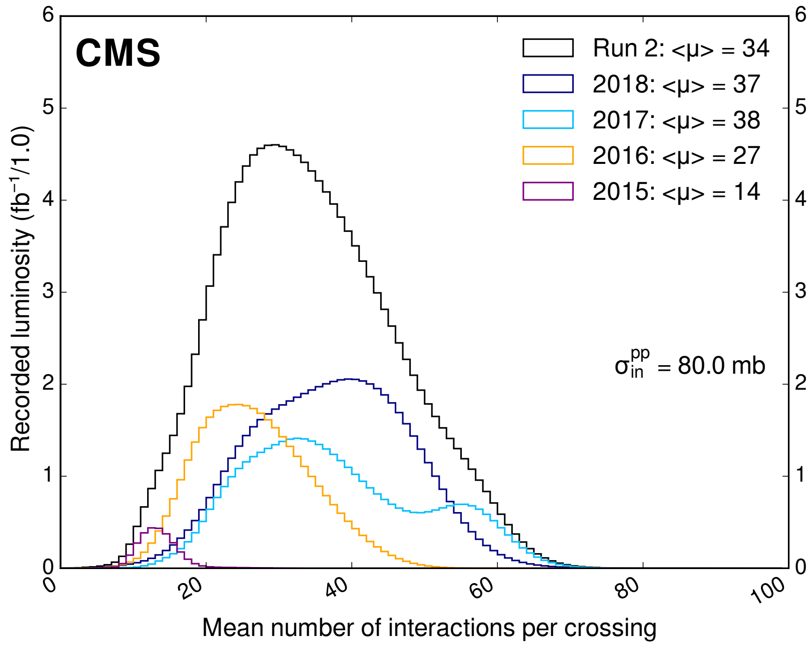


Figure 3.2: Recorded luminosity as a function of mean number of proton–proton interactions per bunch crossing at CMS during the data-taking period from 2015 to 2018. Here a total inelastic cross section for proton–proton collisions of 80 mb is assumed. This figure is taken from the public CMS luminosity information webpage [102].

3.1.2 Proton–proton collisions at the LHC

The proton bunches are spaced by 25 ns, corresponding to a bunch-crossing frequency of 40 MHz at one of the four IPs. At these four IPs, the ATLAS, CMS, LHCb and ALICE detectors are located. Up to 2×10^{11} protons per bunch were present during LHC Run 2, with the beams squeezed to β^* values between 40 cm to 25 cm at the IPs at the ATLAS and CMS detectors. This resulted in an average pileup (PU) of 34 proton–proton collisions per bunch crossing during the data-taking period from 2015 to 2018 at the CMS detector [102] as shown in Figure 3.2, assuming a total inelastic proton–proton scattering cross section of $\sigma_{\text{in}}^{\text{pp}} = 80\,\text{mb}$ at a centre-of-mass energy of 13 TeV.

The LHC delivered a total integrated luminosity of 164 fb^{-1} to the CMS experiment during Run 2 from 2015 to 2018. A total of 151 fb^{-1} was recorded by the CMS experiment as shown in Figure 3.3. In this thesis, data corresponding to a total integrated luminosity of 138 fb^{-1}, collected from 2016 to 2018, are used. These data passed certification by the CMS collaboration, ensuring that all subdetectors operated within nominal conditions during the data-taking.

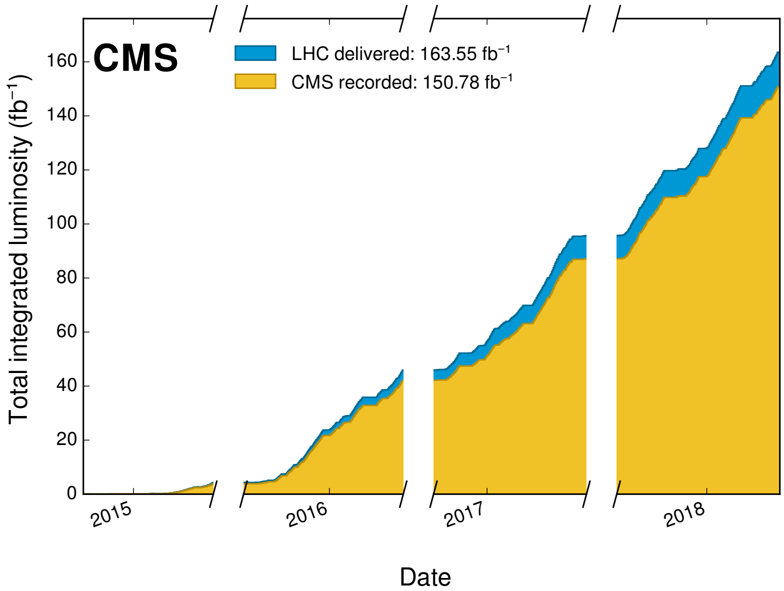


Figure 3.3: Total integrated luminosity for proton–proton collisions delivered by the LHC and recorded by the CMS experiment during Run 2 from 2015 to 2018 as a function of time. This figure is taken from the public CMS luminosity information webpage [102].

3.2 The CMS detector

The Compact Muon Solenoid (CMS) detector [103, 104] is, alongside ATLAS [105], one of the two general-purpose detectors at the LHC. It has an overall length of 22 m, a diameter of 15 m, and a mass of approximately 14 000 t. CMS is designed to measure a wide range of final-state particles produced in high-energy proton–proton and heavy-ion collisions, and to efficiently select interesting events using a dedicated trigger system. These particles include electrons, muons, photons, charged and neutral hadrons, as well as missing transverse momentum, which serves as a proxy for weakly or non-interacting particles such as neutrinos. The detector covers nearly the full solid angle around the interaction point and is composed of several layered and complementary subsystems. A transverse slice through the CMS detector is shown in Figure 3.4, illustrating both the detector layout and the interaction of different particle types with the individual components.

The following sections provide an overview of the different subsystems, starting from the interaction point (IP) at the centre of the detector and moving outwards. Located directly around the beam pipe at the collision point is the tracking system, discussed in Subsection 3.2.1. It consists of silicon pixel and strip detectors, which measure the trajectories of charged particles. Moving outward from the IP, the next layer is the electromagnetic calorimeter (ECAL), presented in Subsection 3.2.2, which measures the energy of electromagnetically interacting particles, mainly electrons and photons. In Subsection 3.2.3, the hadronic calorimeter (HCAL), located outside the ECAL, is introduced. It measures the energy of hadrons, such as protons,

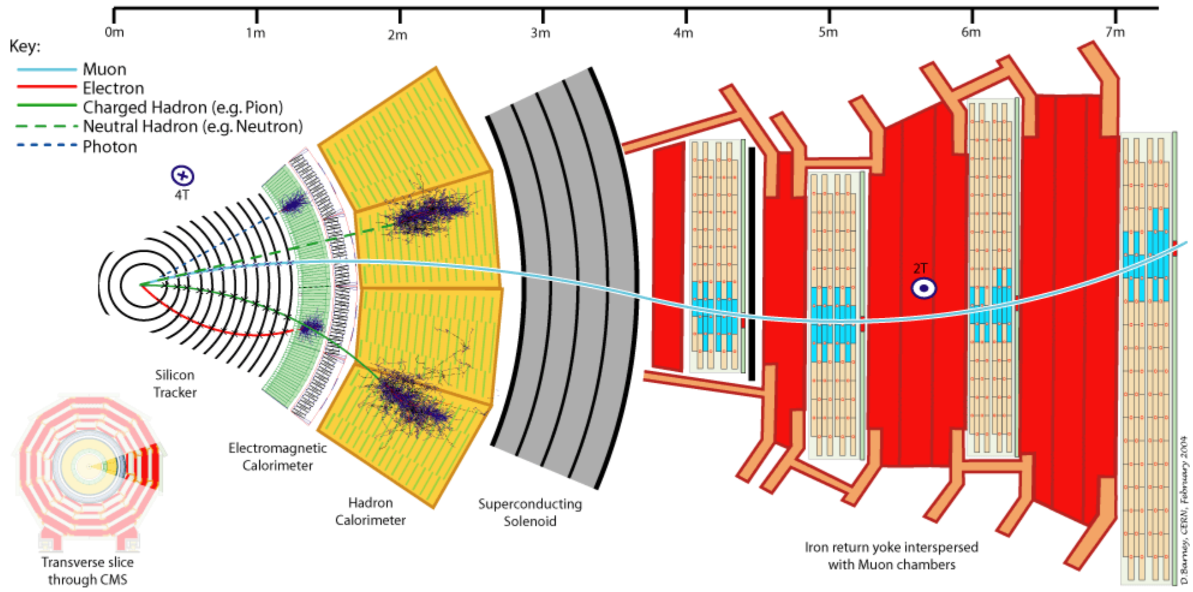


Figure 3.4: A transverse slice through the CMS detector. All subsystems from the IP outwards are shown. The innermost part around the beam pipe is the silicon tracker, which measures the trajectory of charged particles. Outside the tracker, the electromagnetic and hadronic calorimeters are located. They are designed to measure the energy of electrons and photons as well as hadrons. The superconducting magnet generates a magnetic field of 3.8 T, which bends the trajectories of charged particles, enabling the determination of their transverse momentum. The steel return yoke guides the magnetic flux and largely contains the field within the detector volume. Embedded in the return yoke are the muon chambers, designed to precisely measure the trajectories of muons. Combining information from all subsystems allows for the identification and precise measurement of different particle types. The interaction of a positron (red), negatively charged muon (blue), a charged hadron (green), a neutral hadron (green dashed), and a photon (blue dashed) with the detector is shown. Illustration taken from Ref. [106].

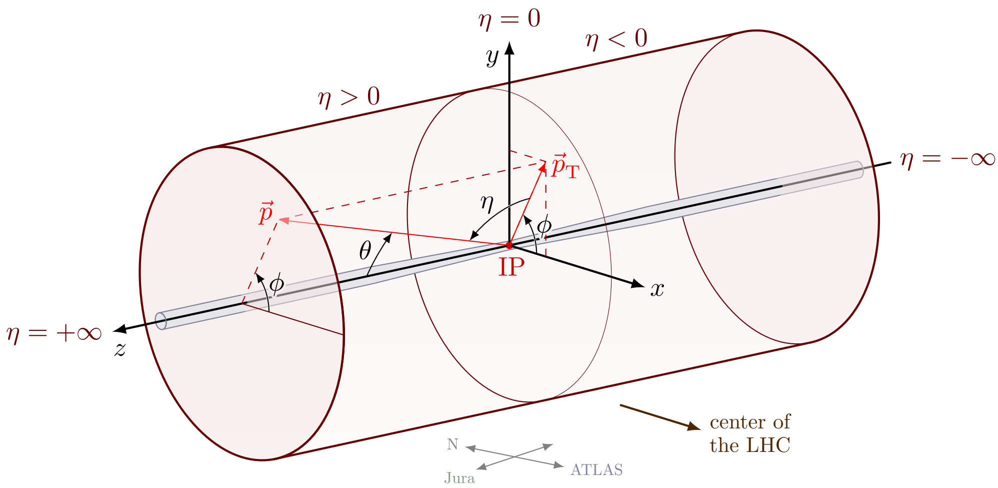


Figure 3.5: Coordinate system used in the CMS detector. The right-handed coordinate system is centred at the interaction point (IP), with the x-, y-, and z-axes indicated, as well as the direction towards the centre of the LHC ring. The beam axis is depicted along the z-axis, including indications of the regions where the pseudorapidity η is greater and smaller than 0. The momentum vector $\vec{p}$ of an example particle and its transverse component $\vec{p}_{\mathrm{T}}$ are illustrated. A compass indicating the orientation towards north, ATLAS, and the Jura mountains is also included below the detector sketch. Illustration from Ref. [107].

neutrons, and pions. Outside the calorimeters, a central feature of the CMS detector, the superconducting solenoid, is located. It generates a magnetic field of 3.8 T approximately parallel to the beam axis. This bends the paths of charged particles via the Lorentz force and enables the determination of a particle's transverse momentum. Finally, the outermost subsystem of the CMS detector is the muon system, described in Subsection 3.2.4. The muon system consists of several detector technologies, including drift tubes (DTs), cathode strip chambers (CSCs), and resistive plate chambers (RPCs), all embedded in the steel return yoke for the magnetic field. The detector description is followed by Subsection 3.2.5, which explains the trigger system used to reduce the data rate to manageable levels and make decisions on which collisions are of interest to be stored for further analysis.

The coordinate system used in the CMS detector, which is also employed to describe the momenta of reconstructed particles, is illustrated in Figure 3.5. The origin of the right-handed coordinate system is defined at the nominal IP, located at the centre of the detector. The x-axis points towards the centre of the LHC ring, the y-axis points vertically upwards, and the z-axis is oriented along the beam direction, with the positive direction defined by the counter-clockwise beam.

Kinematic quantities are commonly expressed using cylindrical coordinates (r, ϕ, z). The azimuthal angle ϕ is measured in the transverse $(x{-}y)$ plane with respect to the x-axis, while the polar angle θ is defined with respect to the z-axis. Instead of θ, the pseudorapidity η is frequently used, as it provides a convenient description of particle production in hadron collisions and coincides with rapidity for massless particles. Rapidity differences are invariant under Lorentz boosts along the beam axis, making η a useful angular variable in the ultrarelativistic limit. It is defined as

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]. \tag{3.1}$$

In addition, the transverse momentum p_{T} is defined as the component of a particle's momentum perpendicular to the beam axis, $p_{\mathrm{T}} = p\sin\theta$, and plays a central role in event reconstruction and selection.

3.2.1 Tracker system

The CMS tracking system is designed to measure the trajectories of charged particles with high precision using silicon detector modules, as described in detail in Ref. [108]. The precise reconstruction of particle trajectories allows charged particles to be associated with individual interaction vertices from the many proton–proton collisions occurring in a single bunch crossing. Each interaction point is reconstructed as a separate vertex, with the primary vertex (PV) defined as the vertex corresponding to the hardest scattering in the event, typically identified by the largest summed transverse momentum of its associated tracks, using only information from the tracking system, as described in Section 9.4.1 of Ref. [109]. In most analyses, only particles originating from the hard interaction associated with the PV are of primary interest. All other reconstructed vertices arise from additional, softer interactions and are classified as pileup (PU).

The tracking system consists of the pixel detector, providing very high spatial granularity close to the beam pipe, and the strip detector in the outer regions. Both subsystems are based on the same fundamental principle for detecting charged particles. A reverse-bias voltage is applied to the p–n junctions in the silicon sensors, such that the depleted region extends over most of the detector volume. When a charged particle traverses this region, it ionizes the material and creates electron–hole pairs. These charge carriers are collected at the electrodes, producing an electrical signal that is read out and used to reconstruct the passage of the particle.

The main difference between the pixel and strip detectors lies in the layout of the p–n junctions. In the pixel detector, the sensors are segmented into a two-dimensional grid, enabling precise measurements of hit positions in both transverse directions. The strip detector, on the other hand, uses elongated strips, providing position measurements in one dimension per layer. In combination with multiple layers, including stereo configurations, this allows for full track reconstruction while significantly reducing the number of readout channels. This design choice balances spatial resolution, channel count, and material budget. The highest spatial resolution is required close to the IP to resolve tracks from multiple nearby vertices.

During the extended technical stop of the LHC in 2016–2017, an upgraded pixel detector was installed, increasing the number of layers from three to four [110]. The strip detector surrounds the pixel system and consists of multiple layers of silicon strip modules arranged in barrel and endcap regions. The CMS tracker achieves a relative transverse momentum resolution of about 2.8% for isolated muons with $p_{\mathrm{T}} = 100\,\mathrm{GeV}$ in the central region ($|\eta| < 1.4$) [108].

During the early 2016 data-taking period, a decrease in the signal-to-noise ratio in the tracker was observed. This issue was more pronounced at increased instantaneous luminosity and caused an efficiency loss of up to 10% in the tracker [111]. It was found that this issue was caused by saturation effects in the preamplifier of the APV25 readout chip (analogue pipeline voltage (APV)) [112] under high occupancy in the tracker. The problem could be mitigated by adjusting the preamplifier feedback voltage bias, abbreviated as voltage feedback preamplifier (VFP), of the readout chips, thereby recovering the tracking efficiency to 99%. Unfortunately, this loss in efficiency cannot be recovered for the affected 20 fb^{-1} of recorded data and has to be included in the Monte Carlo (MC) simulation. The 2016 data-taking period of the CMS detector is therefore divided into two parts.

In centrally produced datasets by the CMS Collaboration, the affected early 2016 period is referred to as high ionizing particle multiplicity (HIPM), reflecting the high-occupancy conditions under which the effect becomes significant. In contrast, simulated samples use the label APV for the corresponding detector configuration, referring to the modified response of the APV25 readout chip. The same periods are also denoted as pre-VFP and post-VFP, referring to data before and after the mitigation of the issue.

3.2.2 Electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) of the CMS detector is designed to measure the energy of electromagnetically interacting particles, primarily electrons and photons [113]. High-energy electrons predominantly lose energy via Bremsstrahlung, while high-energy photons undergo electron–positron pair production in the field of nuclei. These processes lead to the development of electromagnetic showers, in which the number of particles increases until their energies fall below the pair production threshold. At low energies, the remaining charged particles deposit their energy in the material predominantly via ionization.

The characteristic scale governing these processes is the radiation length X_0, which determines both the mean distance over which electrons lose energy via Bremsstrahlung and the conversion probability of photons. In order to fully contain electromagnetic showers, a calorimeter must therefore span several radiation lengths. The ECAL is constructed as a homogeneous calorimeter using more than 82 000 scintillating lead tungstate crystals, providing a total depth of about $25X_0$. The short radiation length of lead tungstate, $X_0 = 0.89$ cm, allows for a compact and homogeneous detector design while containing nearly the entire electromagnetic shower.

The energy deposited via ionization leads to the emission of scintillation light proportional to the deposited energy. The optical transparency of lead tungstate allows this light to be efficiently collected and read out using avalanche photodiodes in the barrel region and vacuum phototriodes in the endcaps.

The ECAL configuration during Run 2 achieves an energy resolution for electrons ranging from 1.8% to 5% from the inner barrel to the endcaps [114]. While optimized for electromagnetic particles, the calorimeter also contributes to the measurement of hadronic activity. In combination with the hadron calorimeter, it plays an important role in the reconstruction and calibration of jets, where it provides a precise measurement of the electromagnetic component of hadronic showers and jets.

3.2.3 Hadronic calorimeter

The hadronic calorimeter (HCAL) is located directly outside the ECAL and is designed to measure the energy of hadrons [115]. Hadrons interact predominantly via the strong interaction, leading to hadronic showers that are more spatially extended and subject to larger fluctuations than electromagnetic showers, with a significant fraction of the energy deposited in nuclear interactions.

The characteristic scale governing hadronic interactions is the nuclear interaction length, which is considerably larger than the radiation length introduced in Subsection 3.2.2. To contain hadronic showers within the available detector volume, the HCAL is therefore implemented as a sampling calorimeter. It consists of alternating layers of brass absorber, which induces the shower development, and plastic scintillator serving as the active medium. Due to spatial constraints, the calorimeter covers about six nuclear interaction lengths in the central (barrel) region. The tracker and ECAL add approximately one additional interaction length. While this is sufficient to contain the majority of hadronic showers, a small fraction of the hadronic shower leaks out of the calorimeter at the per-mille level of the total shower energy. These residual particles can reach the muon system and produce signals, an effect referred to as hadronic punch-through.

Due to the sampling nature of the detector and the intrinsic fluctuations of hadronic showers, the achievable energy resolution is inferior to that of the ECAL. Together, the ECAL and HCAL form the calorimetric system of the CMS detector, where the HCAL measures the dominant fraction of hadronic energy. Both subsystems are essential for the reconstruction and calibration of jets, as well as for the determination of the missing transverse momentum.

During Run 2, the HCAL underwent a Phase-1 upgrade, in which the photodetectors were replaced by silicon photomultipliers and the readout electronics were modernized [116]. These improvements enhanced the signal-to-noise ratio and ensured stable operation under high luminosity conditions.

3.2.4 Muon system

Muons, which are about 200 times more massive than electrons, lose only a small amount of energy via bremsstrahlung and behave as minimum ionizing particles. Muons with a transverse momentum of $p_{\mathrm{T}} \gtrsim 10\,\mathrm{GeV}$ therefore traverse the entire detector and can be detected in the muon chambers.

The CMS muon system is located outside the HCAL and the superconducting solenoid. A detailed description can be found in Ref. [117]. It measures muons in the pseudorapidity range $|\eta| < 2.4$ using three types of gaseous detectors: drift tubes (DTs), cathode strip chambers (CSCs), and resistive plate chambers (RPCs). All detectors are embedded in the steel return yoke of the magnet, which also serves as the mechanical support structure. Additional gas electron multipliers (GEMs) were installed during LS2 and are used in Run 3, but are not relevant for the data collected during Run 2 considered in this analysis.

In the barrel region, DTs are used for precise muon track reconstruction, while CSCs are employed in the endcap regions, where particle rates and radiation levels are higher. The RPCs provide excellent timing resolution of about 1 ns and fast signal formation without drift time, thereby enabling robust bunch-crossing identification. They are therefore primarily used in the Level-1 (L1) trigger, in combination with information from the DTs and CSCs.

All muon detectors in CMS operate on a similar principle. A charged muon traversing the gas volume ionizes the gas, producing free electrons and ions. The electrons drift in an applied electric field towards the anode, where the strong electric field initiates an avalanche multiplication in the vicinity of the readout structures. The resulting charge signals are collected by the readout electronics, allowing for precise position and timing measurements. The spatial resolution depends on the detector technology, ranging from $\mathcal{O}(100\,\mu\text{m})$ for DTs and CSCs to $\mathcal{O}(1\,\text{cm})$ for RPCs.

The muon trigger efficiency during data-taking at 13 TeV is described in detail in Ref. [118] and the identification efficiency and reconstruction resolution in Ref. [119]. The efficiency of the single-muon trigger exceeds 90% over the full η range, with the muon identification and reconstruction efficiency exceeding 96%. In addition to standalone measurements in the muon system, muon tracks are combined with tracks reconstructed in the silicon tracker to form so-called global muons. This combined reconstruction significantly improves the momentum resolution, in particular at low and intermediate transverse momenta. The relative resolution of the transverse momentum for muons with $p_\text{T} \leq 100\,\text{GeV}$ is 1% in the barrel ($|\eta| < 0.9$) and degrades to about 3% in the endcaps ($1.2 < |\eta| < 2.4$).

3.2.5 Trigger system

Bunch crossings (BXs) at the LHC occur at a frequency of approximately 40 MHz, but the rate at which events can be permanently stored is only of the order of 1 kHz. The CMS detector therefore employs a two-tiered trigger system [120] to reduce the event rate to a manageable level while retaining events of potential physics interest.

The first stage is the L1 trigger [121], which is implemented in custom hardware and uses coarse-granularity information from the calorimeters and muon detectors. It identifies physics objects such as muons, electrons, photons, and jets, and reduces the event rate to approximately 100 kHz. The second stage is the high-level trigger (HLT) [122], a software-based system that performs a more detailed event reconstruction using full detector information, including the silicon tracker. The HLT further reduces the rate to about 1 kHz, which is then written to permanent storage and reconstructed offline using more precise algorithms.

Events used in this thesis are selected using single-muon triggers. At the L1 trigger, muon candidates are reconstructed from information in the muon chambers (DTs, CSCs, and RPCs) and are required to exceed a transverse momentum threshold. The HLT refines this selection by reconstructing muon tracks with improved resolution, combining muon-system and tracker information. Additional quality and isolation requirements are typically applied at the HLT level to suppress backgrounds from hadrons and non-prompt muons. The combination of L1 and HLT requirements results in a highly efficient and pure selection of prompt muons, with trigger efficiencies exceeding 90%.

During Run 2, timing effects in the L1 trigger system led to inefficiencies referred to as *prefiring*. Prefiring occurs when detector signals are incorrectly assigned to an earlier bunch crossing (BX−1) instead of the correct bunch crossing (BX0). If the L1 trigger accepts the event in BX−1, the subsequent event in BX0 is vetoed by the trigger rules, which are designed to prevent buffer overflows by limiting the number of accepted events in consecutive bunch crossings. As a consequence, the physically interesting event in BX0 is not recorded, leading to an inefficiency.

In the calorimeter system, prefiring was primarily caused by a gradual shift in the timing of the ECAL trigger primitives in the forward region ($|\eta| \gtrsim 2.5$), induced by radiation damage affecting the crystal transparency. This shift resulted in a fraction of energy deposits being assigned to BX−1, leading to a significant inefficiency for events containing high-p_T electrons or energetic forward jets. The effect depends strongly on the transverse momentum and pseudorapidity of the objects and reached values of up to $\mathcal{O}(10–20\%)$ in extreme regions of phase space.

Muon prefiring arises from the finite timing resolution of the muon detectors, which can lead to an incorrect bunch-crossing assignment of muon candidates at the L1 trigger. If a muon is assigned to BX−1 and triggers the event, the correct BX0 event is again vetoed by the trigger rules. In contrast to the calorimeter case, this effect is smaller, with a prefiring probability of about 1.5% per muon at the beginning of 2016, decreasing to about 0.5% after improvements in the timing synchronization and refinements of the combined muon trigger, in particular through a more effective use of RPCs information. To account for this effect, data-driven correction factors are applied to simulated events.

3.3 Event reconstruction

This section describes the event reconstruction and particle identification used in the CMS experiment. One of the central features of the particle reconstruction and identification is the PF algorithm described in Subsection 3.3.1. It combines information from all subdetectors to reconstruct and identify all visible particles traversing through the detector, also identifying particles from PU interactions. Using the information provided by the PF algorithm, jets are clustered from all final-state particles as described in Subsection 3.3.2 also using PU mitigation methods.

3.3.1 Particle-flow reconstruction

The particle-flow (PF) algorithm is described in detail in Ref. [123]. It uses a holistic approach to reconstruct and identify particles by combining information from all subdetectors of the CMS experiment.

Tracks of charged particles are first reconstructed in the tracker using an iterative procedure based on Kalman filtering. This approach allows for an efficient reconstruction of tracks with transverse momenta down to about 200 MeV. At high transverse momenta, the determination of the track curvature becomes increasingly limited by the spatial resolution of the tracker, requiring additional information from the other subdetectors.

In parallel to the track finding, clusters of energy deposits are reconstructed in the ECAL and HCAL. In the ECAL, energy deposits are grouped into superclusters to account for the spatial development of electromagnetic showers, in particular due to Bremsstrahlung and photon conversions. The PF algorithm links tracks and calorimeter clusters to form blocks of connected elements, which are then used to reconstruct individual particle candidates.

Electrons are reconstructed by matching tracks to ECAL superclusters, accounting for energy losses due to Bremsstrahlung. Photons are reconstructed from energy deposits in the ECAL that are not associated with charged particle tracks. Converted photons are identified by matching ECAL clusters to pairs of oppositely charged tracks originating from a common vertex in the tracker. Their energy is determined primarily from the ECAL measurement.

Clusters in the HCAL are combined with tracks to reconstruct charged hadrons, while clusters without associated tracks are identified as neutral hadrons. For charged hadrons, the momentum measurement from the tracker typically provides a more precise estimate of the particle energy than the calorimeter measurement. These reconstructed hadrons constitute the dominant input for jet clustering described in the following section.

Muons are reconstructed independently in the muon system using hits in the DTs, CSCs, and RPCs, forming so-called *standalone* muons. These are matched to tracks in the inner tracker. If the trajectories are compatible, a combined *global* fit is performed, which improves the momentum resolution, in particular at large transverse momenta. In addition, *tracker* muons are defined by matching inner tracker tracks to signals in the muon system. Muon candidates reconstructed with different algorithms are subsequently merged.

Energy deposits in the calorimeters are also associated with reconstructed muons. Charged hadrons can be misidentified as muons when remnants of hadronic showers reach the muon system, an effect referred to as hadronic punch-through. This background can be suppressed by applying dedicated identification criteria, as described in Section 4.3.

3.3.2 Jet reconstruction

In high-energy particle collisions, quarks and gluons are produced in the hard scattering process. Due to colour confinement, these partons cannot exist as free particles and instead undergo fragmentation and hadronization, producing collimated sprays of hadrons. These hadrons can decay further and interact with the detector material, resulting in a shower of stable particles that can be reconstructed.

In order to measure the properties, such as the energy and transverse momentum, of the parton initiating a shower, the reconstructed final-state particles are clustered into jets. Jets are an experimental proxy for these initial partons. In the CMS experiment, jet reconstruction is based on the PF candidates described in Subsection 3.3.1.

The clustering of PF candidates into jets is performed using the infrared- and collinear-safe anti-k_T algorithm with a distance parameter of $R = 0.4$, as implemented in the FastJet package [124, 125]. This algorithm defines a distance measure between particles in the η–ϕ plane and successively recombines them, resulting in jets with a regular, approximately conical shape. The anti-k_T algorithm is widely used due to its robustness against soft radiation and its stability under higher-order quantum chromodynamics (QCD) corrections. The four-momentum of a reconstructed jet is defined as the sum of the four-momenta of its constituent PF candidates.

Because part of the particle energy is deposited in passive detector material, in particular in the absorbers of the HCAL, the reconstructed jet momentum does not directly correspond to the momentum of the corresponding stable-particle jet. Simulation studies indicate that the raw reconstructed jet momentum differs from the true momentum by approximately 5% to 10% over the full kinematic range and detector acceptance. Additional particles originating from PU interactions can contribute extra tracks and calorimeter energy deposits, biasing the reconstructed jet momentum. To mitigate these effects, the CMS experiment employs the charged hadron subtraction (CHS) technique at the level of PF candidates.

The CHS algorithm exploits the precise tracking capabilities of the detector to associate charged particles with their production vertex. Charged PF candidates identified as originating from pileup vertices, rather than the PV, are removed from the list of particles used for jet clustering. In this way, the contribution of charged pileup particles to the jet energy is largely suppressed.

Since neutral particles cannot be directly associated with a specific vertex, their contribution from pileup remains and is corrected for at a later stage using dedicated jet energy corrections. The combination of CHS and subsequent corrections significantly reduces the impact of pileup on jet reconstruction.

Therefore, a sequence of jet energy corrections is applied to account for these effects and to bring the reconstructed jet energy on average to the level of stable-particle jets. The corrections are factorized into an offset correction removing the residual neutral pileup contribution, followed by relative corrections as a function of η and absolute corrections as a function of p_{T}. For data, an additional residual correction is applied to account for small remaining differences between data and simulation. The correction factors are derived from a combination of simulation and in situ measurements using well-understood reference processes, such as dijet, γ + jet, and Z + jet events.

In simulation, the jet energy resolution is typically better than in data. Therefore, an additional jet energy resolution correction is applied by smearing the reconstructed jet momentum in simulated events such that the resolution matches the one observed in data. The correction is derived as a function of the jet pseudorapidity and is applied either using the momentum of a matched generator-level jet or, if no such jet is found, using a stochastic smearing procedure.

After applying all corrections, the jet energy resolution is approximately 15% to 20% at 30 GeV, 10% at 100 GeV, and 5% at 1 TeV [126].

CHAPTER 4

Triple-differential Z+jet cross section measurement

As described in Chapter 2, measuring the Z + jet cross section triple-differentially provides valuable input for determining the proton's partonic structure. This measurement focuses on the muon decay channel of the Z boson, as muons are reconstructed and measured with excellent precision in the Compact Muon Solenoid (CMS) detector, as described in Subsection 3.2.4. Compared to electrons, muons typically exhibit higher reconstruction and identification efficiencies as well as smaller associated systematic uncertainties. Since this measurement is predominantly limited by systematic uncertainties, the inclusion of the electron decay channel would only provide a marginal improvement in the overall precision while substantially increasing the complexity of the analysis. The use of the muon decay channel therefore provides a robust experimental signature with high precision and well-controlled systematic uncertainties for this measurement.

In this chapter, the analysed data samples are described in Section 4.1, followed by the Monte Carlo (MC) simulation of signal and background samples in Section 4.2. Then, the object and event selection for this measurement are described in Section 4.3, as well as the correction for detector effects to obtain results at stable-particle level, called unfolding, in Section 4.4. Lastly, the description and evaluation of the systematic uncertainties of the measurement are given in Section 4.5. Together, these steps explain the methodology used to measure the triple-differential Z + jet cross section and to compare the results with theoretical predictions.

4.1 Data samples

The proton–proton collision data used for this measurement were collected from 2016 to 2018 with the CMS detector at the Large Hadron Collider (LHC). Overall, the analysed data correspond to an integrated luminosity of $138\,\mathrm{fb}^{-1}$, for which all relevant CMS detector subsystems were fully operational and the data quality was certified as good. Since the Z boson is reconstructed from two oppositely charged muons, events from the *SingleMuon* dataset are analysed. Although the signal final state contains two muons, single-muon triggers are chosen because they provide a simple and highly efficient trigger selection when at least one of the two muons satisfies the trigger requirement. The second muon is reconstructed only in the offline event selection. The offline muon transverse momentum requirement is chosen sufficiently above the trigger threshold such that the trigger efficiency is close to its plateau.

The data are split into four different data-taking periods: early and late 2016, 2017, and 2018. In 2016, an issue in the tracker readout boards, as described in Subsection 3.2.1, caused an efficiency loss in data. Therefore, the 2016 data-taking period is split into the early period affected by this issue, also referred to as pre-VFP, and the late period, where this issue was fixed, also called post-VFP. A full list of the used datasets and their corresponding number of events is given in the appendix in Table A.1.

4.2 Signal and background simulation

For the signal, MadGraph5_amc@nlo [77, 127] is used to simulate $\mathrm{pp} \to \mathrm{l}^+\mathrm{l}^-$ with $\mathrm{l}^\pm \in (\mathrm{e}^\pm, \mu^\pm, \tau^\pm)$ in association with up to two jets at next-to-leading order (NLO) in perturbative quantum chromodynamics (QCD) for an invariant mass $m_{\mathrm{ll}} > 50\,\mathrm{GeV}$ of the lepton pair. The simulated samples for 0, 1 and 2 jets in the final state are merged using the FxFx merging scheme [128] and interfaced to pythia 8 [60] with the CP5 tune [129]. This simulation is normalized to the fiducial cross section prediction for Drell–Yan production of 6077 pb ± 1 pb (integration) ± 15 pb (PDF) ± 2% (scale) at next-to-next-to-leading order (NNLO) accuracy in perturbative QCD and NLO accuracy in electroweak (EW) theory, calculated using fewz version 3.1.b2 [130–133] and the NNPDF3.1 parton distribution function (PDF) set at NNLO including the LUXqed determination of the photon PDF [54, 55, 134, 135].

Additional signal simulations for cross-checks of the model dependence of the unfolding procedure described in Section 4.4 are used. These include simulation of the Drell–Yan process in association with up to four jets at leading order (LO) in perturbative QCD merged with the MLM merging scheme [75, 136]. One sample is interfaced to pythia 8 with the CP5 tune, while another one is interfaced to herwig 7 [137] using the CH3 tune [138]. A third alternative signal simulation is performed using the powheg [79–81] MiNNLO$_{\mathrm{PS}}$ formalism [82, 83] interfaced to pythia 8 and photos for quantum electrodynamics (QED) final state radiation [139]. This sample has NNLO accuracy only for Drell–Yan production without additional jets and only NLO accuracy in perturbative QCD for Z + 1 jet production. A full list of all used simulated signal samples is given in the appendix in Table A.3.

The background processes for this measurement are heavily suppressed by the choice of the signal reconstruction. Muons are very precisely reconstructed by the CMS detector. Including tight identification and isolation criteria strongly suppresses nonprompt muons not originating from the hard interaction. Remaining contributions from other processes that pass the selection criteria are estimated from simulation. There are two main background contributions, diboson and nonresonant processes.

Firstly, resonant backgrounds from diboson production can yield similar event signatures. For example, two oppositely charged W bosons decaying into a muon and a muon neutrino can easily be misidentified as muons originating from a single Z boson, as neutrinos do not interact with the detector. Furthermore, the production of a Z boson in association with a W boson passes the selection criteria if the Z boson decays into two muons. Here, the W boson can decay in any mode, for example, into a quark-antiquark pair, which then form two jets. Additionally, Z boson pair production yields similar event signatures when one Z boson decays into muons while the other Z boson, for example, decays into a quark-antiquark pair, resulting in two additional jets. The diboson production is simulated at NLO accuracy in perturbative QCD with POWHEG [140, 141] or MADGRAPH5_aMC@NLO, interfaced to PYTHIA 8 using the CP5 tune. These samples are normalized to the cross section predicted by the respective MC generator. A full list of the used diboson simulation samples can be found in the appendix in Table A.4.

The second background category stems from nonresonant processes, including single top quark and top quark pair production. These samples are simulated using POWHEG [142–144] at NLO accuracy in perturbative QCD. The decays of the top quarks are performed in the narrow-width approximation with spin correlations using the MADSPIN formalism [145]. The top quark almost always decays into a bottom quark and a W boson, with the subsequent W boson decay into leptonic or hadronic final states. Leptonic decays of the W boson can produce prompt muons that enter the reconstruction of a Z boson candidate, thereby contributing to the background events of this measurement.

The simulation of single top quark production is split into two contributions, one in association with a W boson and another one in association with light quarks. The simulation of a single top quark in association with a W boson is normalized to a cross section of 71.7(39) pb, which is calculated at NNLO accuracy in Refs. [146, 147]. Meanwhile, the top quark and top antiquark production in association with a light quark is normalized to a cross section of $145.0^{+2.8}_{-1.9}$ pb and $87.2^{+1.8}_{-1.5}$ pb, respectively, calculated at NNLO accuracy in Ref. [148]. The cross section for top quark pair production of 831^{+40}_{-51} pb is obtained from TOP++ 2.0 [149] at NNLO perturbative QCD using a top quark mass of 172.5 GeV and includes the resummation of infrared gluons up to next-to-next-to-leading logarithmic accuracy. Table A.5 in the appendix contains a list of all used MC samples for background contributions from top quark production.

The interaction of the final-state particles, obtained from MC simulation described above, with the CMS detector is simulated using the GEANT4 package [150]. Here, the different detector configurations for each data-taking period are taken into account. Hence, there are a total of four MC samples for each simulated process, one for each of the 2016 pre-VFP, 2016 post-VFP, 2017, and 2018 data-taking periods, respectively. Each of these samples is scaled to match the measured integrated luminosity for the corresponding data-taking period described in Subsection 4.2.1. Furthermore, contributions from pileup (PU) are simulated using minimum-bias events and are overlaid on the hard-scattering process. The PU profile is matched to the one measured in the corresponding data-taking period, as explained in Subsection 4.2.2.

4.2.1 Integrated luminosity

The integrated luminosity of the data used in this analysis amounts to 19.5 fb^{-1} and 16.8 fb^{-1} for the 2016 pre-VFP and post-VFP periods, respectively, with an uncertainty in the com-

Table 4.1: Correlation matrix of the luminosity uncertainties between the different Run 2 data-taking periods (left) and decomposition of these correlations into three independent uncertainty sources (right). The uncertainties of the 2016 pre-VFP and post-VFP periods are treated as fully correlated and are therefore combined into a single 2016 period. The final uncorrelated uncertainties are chosen such that one component affects only the 2016 period, one is correlated between the 2016 and 2017 periods, and one is fully correlated across all Run 2 data-taking periods.

	2016	2017	2018
2016	100%	23%	28%
2017	23%	100%	74%
2018	28%	74%	100%

Nuisance parameter	2016	2017	2018
2016	1.18%	–	–
2016–2017	0.04%	0.55%	–
2016–2018	0.35%	0.61%	0.84%

bined luminosity of 1.2% [151]. For the 2017 and 2018 data-taking periods, the integrated luminosities are measured to be 42.1 fb^{-1} and 59.6 fb^{-1}, with uncertainties of 0.82% and 0.84%, respectively [152]. Taking the correlations between the different data-taking periods given in Table 4.1 into account, the total integrated luminosity of the full Run 2 dataset is determined to be 138 fb^{-1} with an uncertainty of 0.73%.

The simulated samples for each data-taking period are weighted by the corresponding integrated luminosity before being combined into a single full Run 2 simulation sample. This procedure ensures that the relative contributions of the different detector conditions and pileup profiles are modelled consistently with the recorded data.

The luminosity uncertainties are propagated to the final measurement while preserving the correlations between the different years. For this purpose, the full covariance matrix is decomposed into three independent nuisance parameters: one affecting only the 2016 data-taking period, one correlated between the 2016 and 2017 periods, and one fully correlated across all Run 2 data-taking periods.

4.2.2 Pileup reweighting

The average number of proton–proton interactions μ in a single bunch crossing can be estimated from the instantaneous luminosity L_{inst}, the revolution frequency of the LHC f_{rev}, the number of colliding bunch pairs n_{b} and the total inelastic proton–proton scattering cross section $\sigma_{\mathrm{in}}^{\mathrm{pp}}$:

$$\mu = \frac{L_{\mathrm{inst}}\,\sigma_{\mathrm{in}}^{\mathrm{pp}}}{f_{\mathrm{rev}}\,n_{\mathrm{b}}}\,. \tag{4.1}$$

Here, the revolution frequency of the LHC is $f_{\mathrm{rev}} = 11\,245\,\mathrm{Hz}$ [92] and the number of bunches reached a maximum of $n_{\mathrm{b}} = 2556$ during Run 2 [90]. The product $f_{\mathrm{rev}}\,n_{\mathrm{b}}$ corresponds to the bunch crossing frequency and converts the instantaneous luminosity into the luminosity delivered per bunch crossing. The instantaneous luminosity is measured by CMS in luminosity sections of 2×10^{18} bunch crossings, corresponding to about 23.3 s of data-taking [104]. Both f_{rev} and n_{b} are known exactly while the values for L_{inst} and $\sigma_{\mathrm{in}}^{\mathrm{pp}}$ are measured and therefore have uncertainties associated with them. According to Ref. [153], the uncertainty in the instantaneous luminosity is negligible compared to that in the total inelastic cross section.

Using the measured instantaneous luminosity and assuming a total inelastic proton–proton scattering cross section of

$$\sigma_{\text{in}}^{\text{pp}} = 69.2\,\text{mb} \pm 4.6\%\,, \tag{4.2}$$

the distribution of the average number of additional interactions per bunch crossing in data is estimated separately for each data-taking period. The value for the inelastic cross section was derived by the CMS Collaboration by optimizing the data to simulation agreement in PU sensitive variables. It is compatible with a measurement by the CMS Collaboration at 13 TeV of 67.5(32) mb [154].

In simulated samples, additional interactions from PU are overlaid during the event generation step using minimum bias events generated with PYTHIA 8. For each simulated hard-scattering event, a value of μ is first drawn from a predefined PU profile corresponding to the data-taking period for the MC production. These profiles were derived using preliminary reconstructed data for 2016, 2017, and 2018, respectively. The actual number of additional interactions in the bunch crossing in simulation is then sampled from a Poisson distribution with mean μ. The generated minimum bias events are overlaid both in-time and for neighbouring bunch crossings in order to model the detector response to out-of-time PU.

The PU profile used in the simulation campaigns is based on a preliminary reconstruction of the data and luminosity measurement, and therefore does not exactly reproduce the luminosity profile observed in data using the so-called ultra-legacy reconstruction, with improved knowledge of the detector and improved luminosity measurements in Ref. [152]. As a consequence, the distribution of the true number of interactions in simulation differs from the one inferred from data.

To correct for this difference, the simulated events are reweighted according to their true number of interactions n_{PU}. The weight applied to each event is given by

$$w(n_{\text{PU}}) = \frac{P_{\text{data}}(n_{\text{PU}})}{P_{\text{sim}}(n_{\text{PU}})}\,, \tag{4.3}$$

where $P_{\text{data}}(n_{\text{PU}})$ is the normalized distribution of the expected number of interactions in data and $P_{\text{sim}}(n_{\text{PU}})$ is the corresponding normalized distribution in simulation. The weights are derived separately for each year and, for 2016, separately for the pre-VFP and post-VFP periods.

An uncertainty in the PU reweighting is assigned by varying the assumed inelastic cross section by its uncertainty of 4.6% when deriving $P_{\text{data}}(n_{\text{PU}})$. This results in alternative PU weights corresponding to increased and decreased PU activity. The difference with respect to the nominal prediction is propagated through the full analysis and taken as the systematic uncertainty due to the modelling of PU.

The effect of the pileup reweighting on the agreement between data and simulation is illustrated in Figure 4.1 using the distribution of the number of reconstructed collision vertices per bunch crossing in the full Run 2 dataset. Since the PU profiles used in the MC production are estimated from data and preliminary luminosity measurements, the overall effect on the full Run 2 is relatively small. The effect of the reweighting is more pronounced in the separate data-taking periods, especially for the 2016 pre- and post-VFP periods, as they used a combined PU profile for the whole 2016 data-taking period. The separate periods are shown

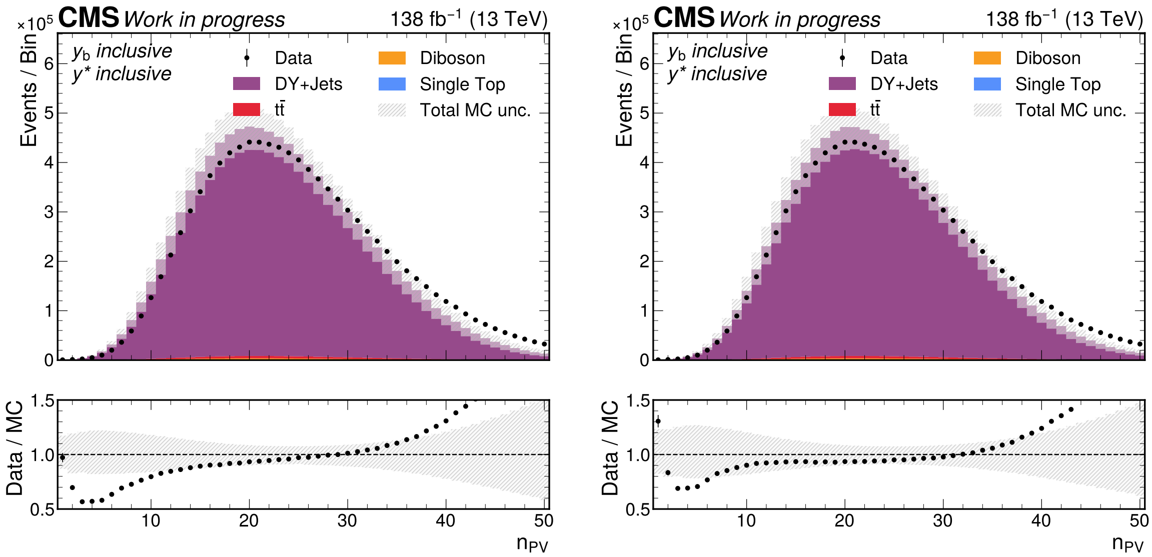


Figure 4.1: Distribution of the number of reconstructed collision vertices per bunch crossing in the Run 2 data used for this measurement compared to simulation before pileup reweighting (left) and after applying the pileup reweighting procedure (right).

in Figure A.7 to Figure A.10 in the appendix. Since the reweighting is derived at generator level using the true number of pileup interactions it therefore does not act directly on the reconstructed vertex multiplicity. Remaining differences between data and simulation are expected due to the imperfect correlation between the true number of interactions and the reconstructed vertex count, as well as differences in vertex reconstruction efficiency.

4.3 Event and object selection

This section describes the event and object selection used in this measurement. First, only events triggered by the presence of a muon candidate are analysed, and basic muon selection criteria are applied, as described in Subsection 4.3.1. Using these selected muons, Z boson candidates are reconstructed from the $Z \rightarrow \mu^+\mu^-$ decay, explained in Subsection 4.3.2. The Z boson candidates need to pass further selection criteria given in the same section. In Subsection 4.3.3 the additional jet selection criteria are explained. Combining all the reconstructed and selected objects, the final observables are calculated and additional requirements at event level outlined in Subsection 4.3.4 are applied.

4.3.1 Trigger and muon selection

In this analysis, the decay of the Z boson into two oppositely charged muons is considered. Events are selected using single-muon triggers. For the 2016 and 2018 data-taking periods, the trigger requires at least one muon with a transverse momentum of $p_T^\mu > 24\,\mathrm{GeV}$. During part of the 2017 data-taking period, this trigger path was prescaled such that only a fraction

Table 4.2: Requirements of the tight muon identification working point used in this analysis.

Muon reconstruction variable	Selection requirement
χ^2/ndf of the global muon track fit	< 10
Number of muon chamber hits	> 0
Number of matched muon stations	> 1
Number of tracker layers with hits	> 5
Number of pixel hits	> 0
Transverse impact parameter d_{xy} with respect to the PV	$< 2\,\mathrm{mm}$
Longitudinal impact parameter d_z with respect to the PV	$< 5\,\mathrm{mm}$

of the selected events was recorded. Therefore, the unprescaled trigger requiring at least one muon with $p_{\mathrm{T}}^{\mu} > 27\,\mathrm{GeV}$ is used for the full 2017 dataset [118].

Further quality selection criteria are applied to the muon candidates to ensure they originate from, e.g., prompt Z boson decays [119]. Muon candidates are required to fulfil the tight muon identification criteria recommended by CMS. These requirements ensure a high-quality reconstruction of the muon trajectory and are summarised in Table 4.2.

Furthermore, the muon candidates are required to be isolated from surrounding particles in order to suppress backgrounds from nonprompt muons inside jets. For this purpose, the relative particle-flow (PF) isolation of a muon is defined as

$$I_{\mathrm{PF}}^{\mu} = \frac{1}{p_{\mathrm{T}}^{\mu}} \left(\sum_{h^{\pm} \in \mathrm{PV}} p_{\mathrm{T}}^{h^{\pm}} + \sum_{\gamma} p_{\mathrm{T}}^{\gamma} + \sum_{h^{0}} p_{\mathrm{T}}^{h^{0}} - \frac{1}{2} \sum_{h^{\pm} \in \mathrm{PU}} p_{\mathrm{T}}^{h^{\pm}} \right), \tag{4.4}$$

where the sums run over charged hadrons associated with the PV, photons, and neutral hadrons within a cone around the muon of size

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4\,. \tag{4.5}$$

The final term in Equation 4.4 corrects for the contribution from neutral particles originating from PU. The factor of one-half is motivated by simulation studies of minimum bias events, where the average contribution from neutral particles is found to be approximately half of that from charged hadrons associated with PU vertices.

For the muon isolation, the tight PF isolation working point, corresponding to an efficiency of approximately 95%, is used. The resulting distribution of the relative muon isolation is shown in Figure 4.2.

Small differences between the reconstructed muon momentum in data and simulation arise from residual misalignments of the tracking system, imperfect modelling of the magnetic field, and inaccuracies in the detector material description. These effects lead to shifts in the muon momentum scale and to differences in the momentum resolution between data and simulation. To account for these differences, corrections derived by the CMS Collaboration for Run 2 are applied to all reconstructed muons [155]. The $\mathrm{Z} \to \mu^{+}\mu^{-}$ mass peak in data and simulation is used to determine corrections to the curvature of the reconstructed muon tracks.

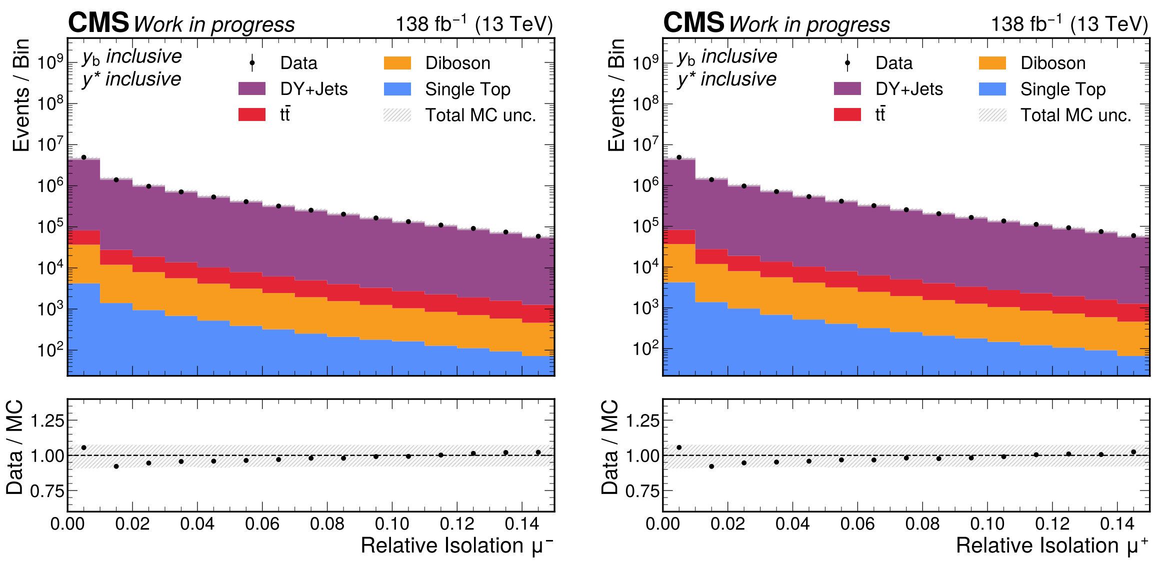


Figure 4.2: Event yields in data and simulation as a function of the relative PF isolation of the two selected muons. The isolation of the negatively charged muon μ^- is shown on the left and the isolation of the positively charged muon μ^+ on the right.

In data, the reconstructed muon momentum is corrected for residual scale biases. In simulated events, an additional smearing is applied to reproduce the momentum resolution observed in data. The correction and smearing factors depend on the muon charge, transverse momentum, pseudorapidity, and azimuthal angle. For simulated muons with a matched generator-level muon, the smearing is applied relative to the generator-level momentum. For unmatched muons, a stochastic smearing is used.

To account for energy losses via photons radiated from muons, a so-called muon dressing procedure is used. For this purpose, the four-momenta of all photons within a cone of

$$\Delta R(\mu, \gamma) < 0.1 \tag{4.6}$$

around the muon are added to the muon four-momentum. The dressed-muon definition reduces the sensitivity of the measurement to soft and collinear photon radiation and improves the correspondence between reconstructed and particle-level muons.

The kinematic requirements on the muons are applied after this dressing procedure. To ensure that the selected events are fully contained in the trigger efficiency plateau region for all data-taking periods, both dressed muons are required to satisfy $p_T^\mu > 29\,\text{GeV}$. In addition, both dressed muons are required to lie within the tracking acceptance of the CMS detector, defined by $|\eta| < 2.4$.

The efficiencies of the muon reconstruction, identification, isolation, and trigger requirements are determined with the tag-and-probe method in $Z \to \mu^+\mu^-$ events. Differences between data and simulation are corrected by applying efficiency scale factors to the simulated events [119, 156].

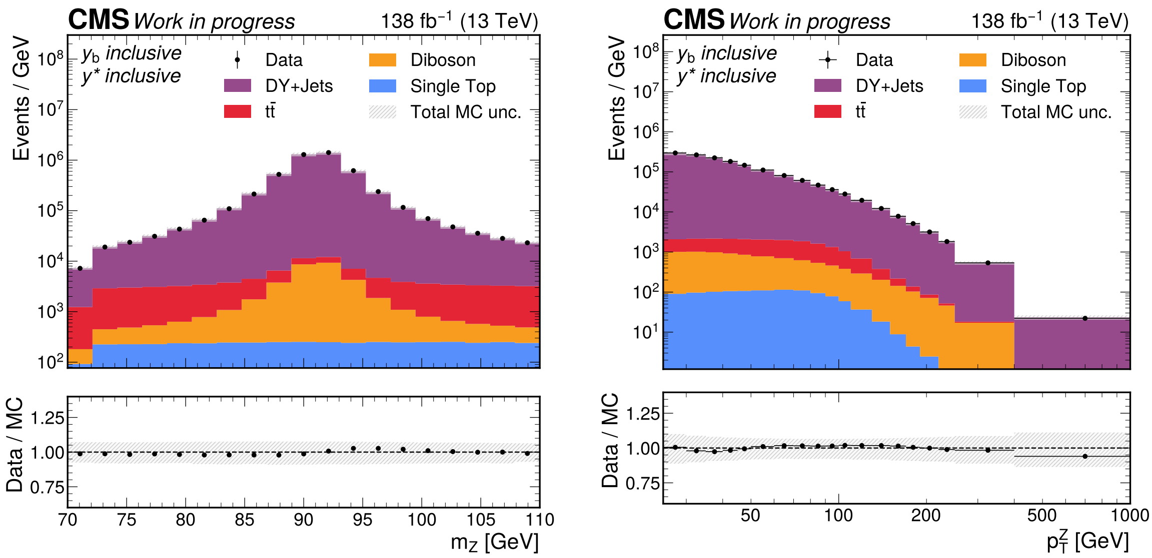


Figure 4.3: Number of events per GeV as a function of the reconstructed Z boson mass (left) and the transverse momentum of the Z boson (right). The region around the mass peak at m_Z is clearly dominated by signal contributions.

4.3.2 Z boson reconstruction and selection

Based on the previously selected muons, Z boson candidates are reconstructed. Two oppositely charged muons in an event form a candidate by adding their four-momentum vectors. If multiple candidates are found, the one closest to the mass of the Z boson $m_Z = 91.1876$ GeV [157] is used.

Additionally, the reconstructed Z boson is required to have a mass within $m_Z \pm 20$ GeV. This selection ensures that the analysed data are dominated by events from signal production while limiting the contribution from background processes. Furthermore, the reconstructed Z boson candidate is required to have a transverse momentum $p_T^Z > 25$ GeV. Only events where the Z boson candidate satisfies these requirements are used for further processing.

The reconstructed mass of the Z boson candidates m_Z and the transverse momentum p_T^Z in data are compared to simulation in Figure 4.3 for the full Run 2 data-taking period.

4.3.3 Jet selection

Jets are reconstructed from the PF candidates described in Subsection 3.3.1 using the anti-k_T algorithm with a distance parameter of $R = 0.4$. The contribution from PU is mitigated using the charged hadron subtraction (CHS) technique, and the reconstructed jet energy is corrected in data and simulation as described in Subsection 3.3.2.

Quality criteria are applied to suppress jets originating from detector noise, beam halo, or misreconstructed particles. These jet identification criteria are based on the composition of the reconstructed jet and require minimum fractions of charged and neutral hadrons, limits

Table 4.3: Jet identification criteria for CHS jets reconstructed with the anti-k_T algorithm with a distance parameter of $R = 0.4$. The requirements are shown separately for the detector regions and data-taking periods relevant for this analysis.

Criteria	2016 $\lvert\eta\rvert < 2.4$	2016 $2.4 < \lvert\eta\rvert < 2.7$	2017 and 2018 $\lvert\eta\rvert < 2.6$
Neutral hadron fraction	< 0.9	< 0.9	< 0.9
Neutral electromagnetic fraction	< 0.9	< 0.99	< 0.9
Number of constituents	> 1	–	> 1
Muon fraction	< 0.8	–	< 0.8
Charged hadron fraction	> 0	–	> 0
Charged multiplicity	> 0	–	> 0
Charged electromagnetic fraction	< 0.8	–	< 0.8

on electromagnetic and muon energy fractions, and a minimum number of jet constituents. The exact requirements for the different data-taking periods and detector regions relevant for this analysis are summarized in Table 4.3.

Jets at low transverse momentum are more susceptible to contamination from residual PU activity and may even originate entirely from PU vertices. Therefore, an additional pileup jet identification requirement, the so-called pileup jet identification (PUJetID), is applied to jets with $p_\mathrm{T}^\mathrm{jet} < 50\,\mathrm{GeV}$. The PUJetID is calculated using a boosted decision tree (BDT) trained on several jet shape and kinematic features [158]. In this measurement, the tight working point, corresponding to an efficiency of approximately 80% for genuine prompt jets, is employed. The efficiency is estimated in simulation using $\mathrm{Z} + \mathrm{jet}$ events and matching reconstructed jets with $p_\mathrm{T} > 20\,\mathrm{GeV}$ to jets clustered at generator level with $p_\mathrm{T} > 30\,\mathrm{GeV}$ within $\Delta R < 0.4$. In data, the performance is measured using events in which the leading jet is closely aligned with the Z boson in azimuth, $\Delta\phi(\mathrm{Z}, \mathrm{jet}) < 1.5$, as a pileup-enriched control region, while events with $\Delta\phi(\mathrm{Z}, \mathrm{jet}) > 2.5$ are classified as enriched in genuine hard-scatter jets.

Scale factors correcting the PUJetID efficiency in simulation to the one measured in data are applied to simulated events. For jets with $\lvert\eta\rvert < 2.5$, the scale factors are close to unity. An additional systematic uncertainty is assigned from the difference between the efficiencies obtained using PYTHIA and HERWIG++ simulated samples.

At low transverse momentum, one of the muons from the Z boson decay can geometrically overlap with the leading jet. To avoid ambiguities in the jet definition and double-counting of the muon momentum, any jet within $\Delta R(\mu, \mathrm{jet}) < 0.4$ of one of the two muons forming the Z boson candidate is removed from the jet collection.

During data-taking, localized problems in the calorimeter system and tracker can lead to anomalous rates of reconstructed PF jets. The CMS Collaboration therefore provides so-called jet-veto maps that identify the affected regions in the η–ϕ plane. Any reconstructed jet falling into one of these regions is removed from the analysis.

The distribution of the leading remaining jet, i.e., the jet with the highest transverse momentum after applying all identification and selection criteria, in the η–ϕ plane is shown in Figure 4.4 for data and simulation. The distributions are shown separately for each data-taking period. For 2016, the same jet identification criteria, jet-veto map and PU jet identification criteria are used for both the pre- and post-VFP eras, and the corresponding distributions are therefore combined. The vetoed regions from the jet-veto maps are clearly visible as localized deficits in the event yields.

All these requirements are applied both to jets reconstructed in data and to reconstructed jets in simulation, while no corresponding selection is applied at particle level. The corresponding inefficiencies arising from the selection criteria at reconstruction level are therefore corrected for in the unfolding procedure described in Section 4.4.

All reconstructed jets are required to satisfy the identification, pileup rejection, overlap removal, and jet-veto-map requirements described above. The event selection is then based on the remaining jet with the largest transverse momentum. The leading remaining jet is required to satisfy $p_{\mathrm{T}}^{\mathrm{jet1}} > 20\,\mathrm{GeV}$ and $|y^{\mathrm{jet1}}| < 2.4$. Events failing either requirement are discarded. The distribution of the transverse momentum of the leading jet $p_{\mathrm{T}}^{\mathrm{jet1}}$ and the distribution of the leading jet rapidity y^{jet1} in data compared to MC simulation are shown in Figure 4.5.

4.3.4 Event selection

Only events in which the reconstructed Z boson candidate and at least one jet pass the criteria described above are considered in this analysis. The jet with the highest transverse momentum is defined as the leading jet and used to construct the variables relevant for this measurement. In all events satisfying these requirements, the two variables

$$y_{\mathrm{b}} = \frac{1}{2}\left|y_{\mathrm{Z}} + y_{\mathrm{jet}}\right| \quad \text{and} \tag{4.7}$$

$$y^{*} = \frac{1}{2}\left|y_{\mathrm{Z}} - y_{\mathrm{jet}}\right| \tag{4.8}$$

are calculated, using the rapidity of the Z boson candidate y_{Z} and the rapidity of the leading jet y_{jet}.

The event yields measured in data differentially in y_{b} and y^{*} are compared to the expected event yields from simulation in Figure 4.6. Due to the PDF-suppressed LO production in simulation towards high y^{*}, as explained in Section 2.4, a worse description of the data by the simulation is expected and observed.

In each (y_{b}, y^{*}) bin, the event yields are measured differentially as a function of $p_{\mathrm{T}}^{\mathrm{Z}}$. The binning scheme in y^{*}, y_{b} and $p_{\mathrm{T}}^{\mathrm{Z}}$ is described in Section 2.3. Using the two muons that form the reconstructed Z boson candidate, as described in Subsection 4.3.2, the transverse momentum of the Z boson is calculated as

$$p_{\mathrm{T}}^{\mathrm{Z}} = \sqrt{\left(p_{\mathrm{T}}^{\mu^{+}}\right)^{2} + \left(p_{\mathrm{T}}^{\mu^{-}}\right)^{2} + 2p_{\mathrm{T}}^{\mu^{+}}\,p_{\mathrm{T}}^{\mu^{-}}\,\cos\left(\Delta\phi_{\mu^{+}\mu^{-}}\right)}. \tag{4.9}$$

A comparison of the event yields estimated from simulation and measured in data for the full unrolled phase space of this analysis is shown in Figure 4.7. To obtain a one-dimensional representation of the three-dimensional phase space, the bins are unrolled in increasing y^{*},

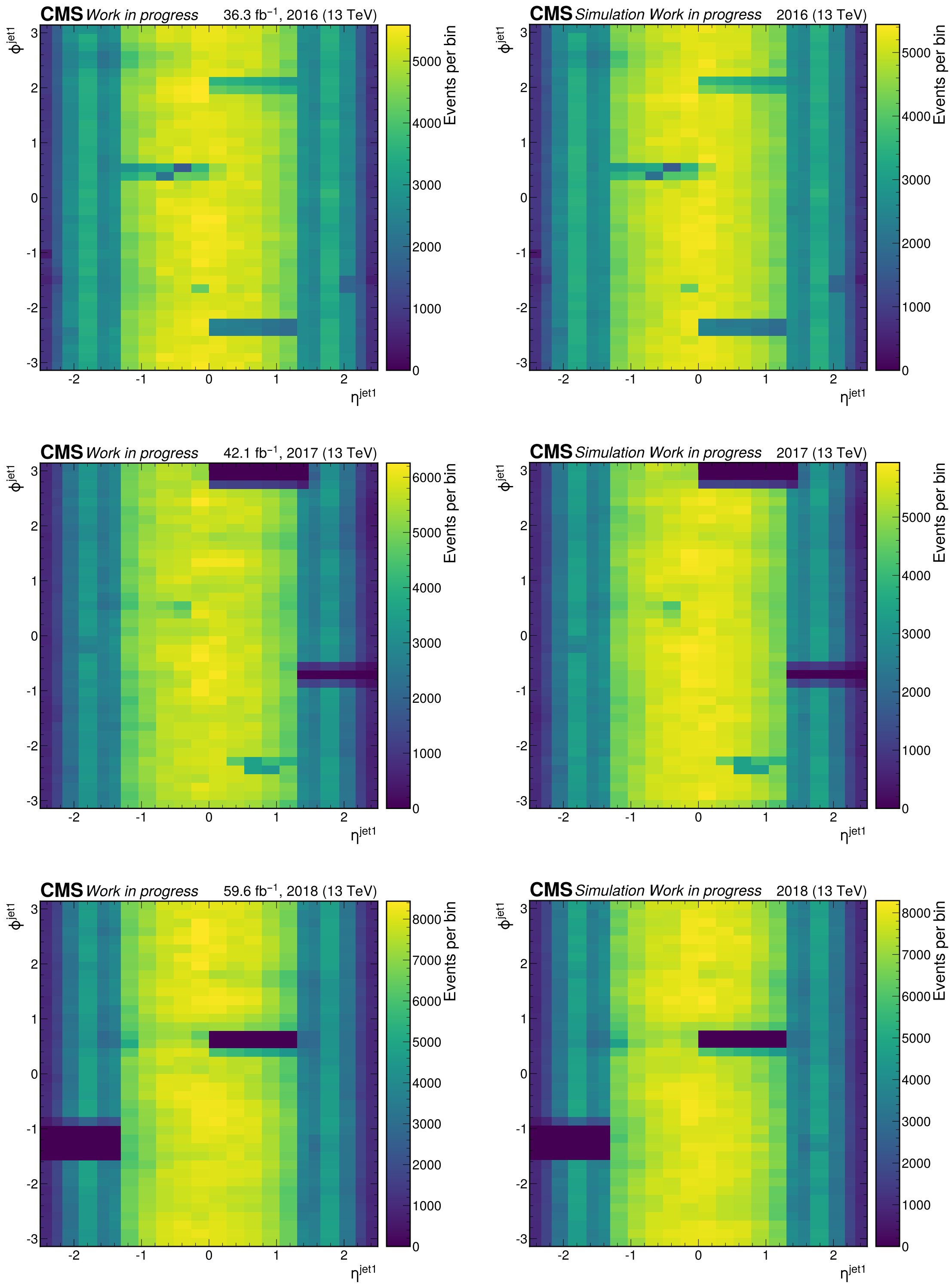


Figure 4.4: Number of events in data (left) and simulation (right) as a function of the η and ϕ coordinates of the leading jet. The data-taking periods are shown separately: 2016 in the upper row, 2017 in the middle row, and 2018 in the lower row.

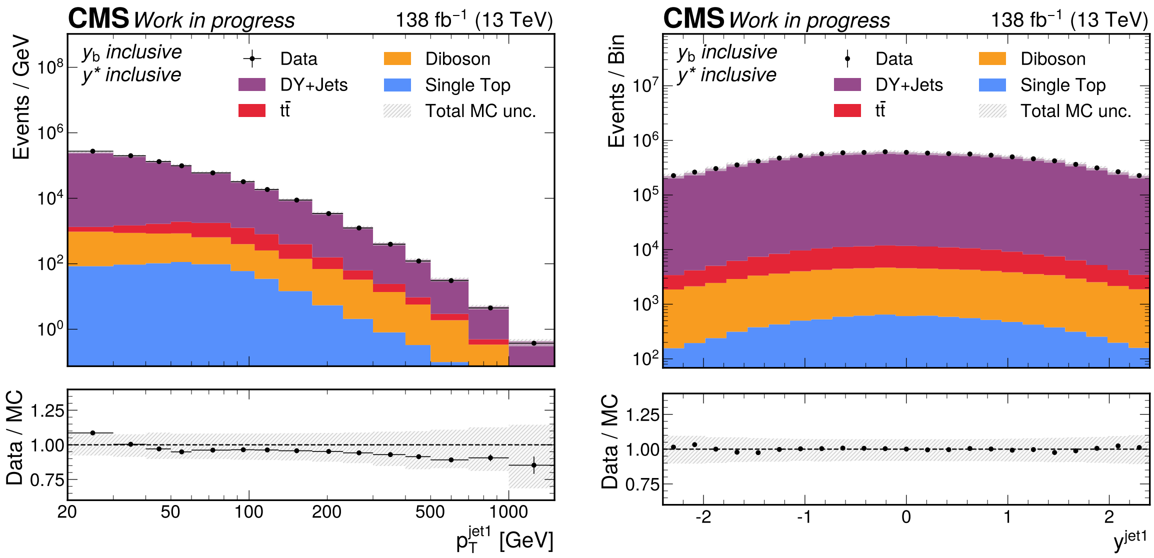


Figure 4.5: Transverse momentum of the leading jet p_T^{jet1} (left) and rapidity of the leading jet y^{jet1} (right) in data compared to MC simulation for the full Run 2 data-taking period.

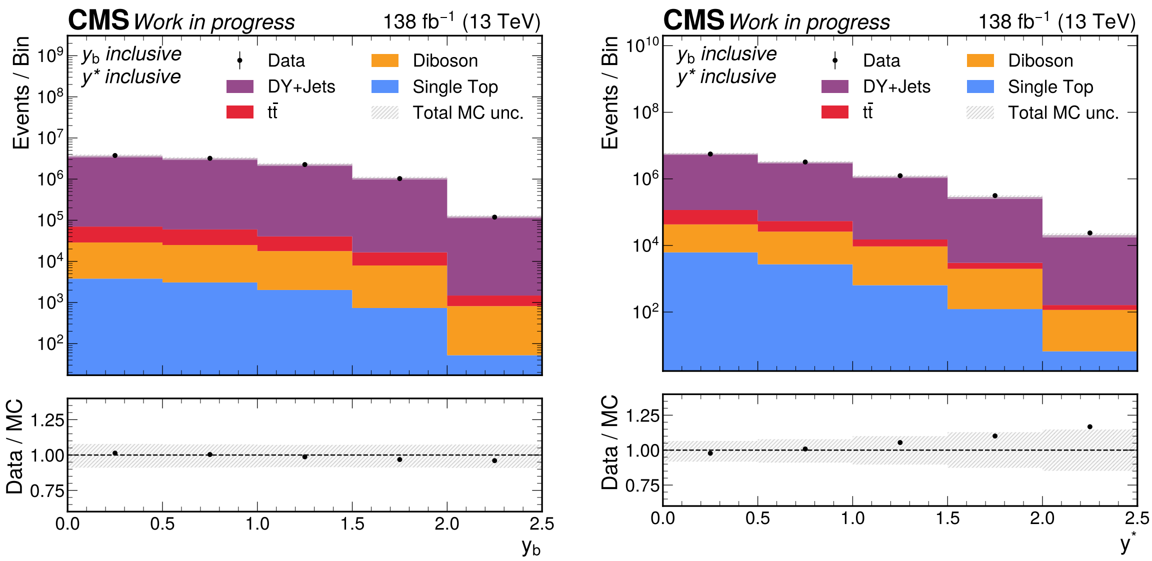


Figure 4.6: Measured event yields differentially in y_b (left) and y^* (right) compared to the expected event yields from simulation.

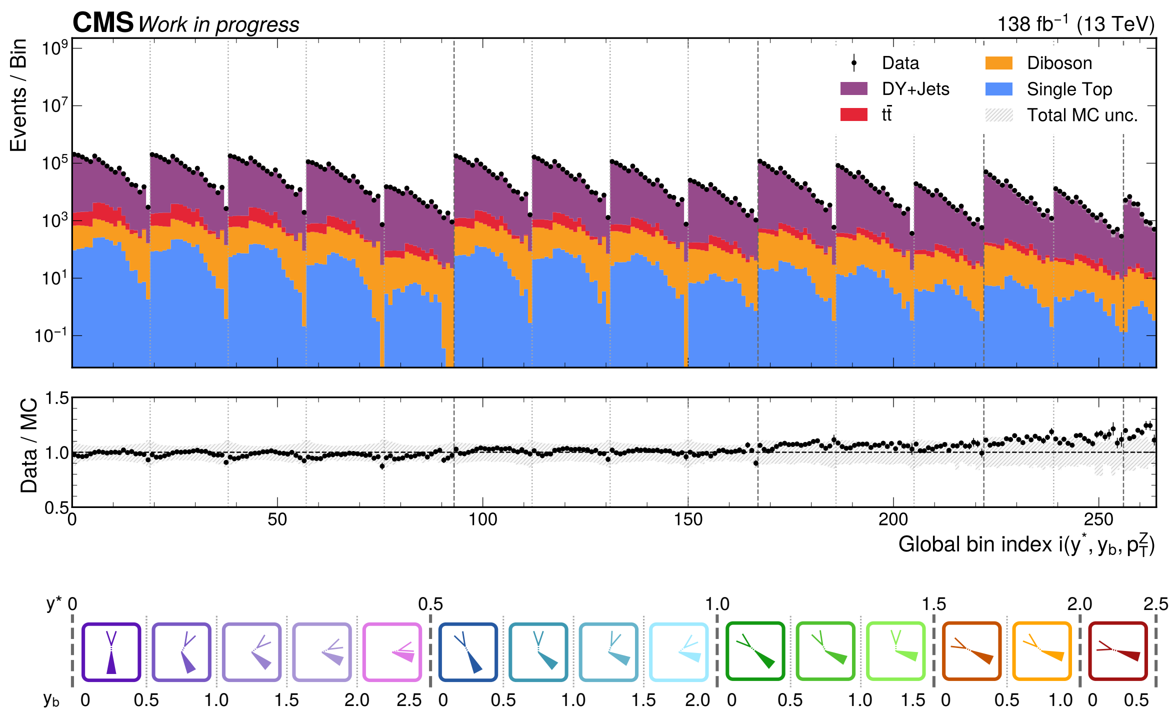


Figure 4.7: Comparison of event yields in simulation and data for the whole unrolled analysis phase space as a function of the global bin index. The event counts are not normalized, therefore changes in the bin width in p_T^Z induce visible structures. An illustration of the unrolling scheme in (y_b, y^*) is shown at the bottom.

y_b, and p_T^Z. For each bin in y^*, all the y_b bins are laid out, each containing their respective series of p_T^Z bins. Each of the bins is assigned a global bin index from 1 to 264 in the order they are unrolled. For illustration purposes Figure 4.7 (bottom) depicts the sequence of the unrolled (y_b, y^*) bins. The binning scheme in (y_b, y^*) and p_T^Z is described in Section 2.3.

4.4 Unfolding

The recorded data are unfolded in order to correct for detector effects, such as bin migrations caused by reconstruction, detector acceptance, and inefficiencies. Given a true distribution $\tilde{x}_j$ of m bins, the matrix A_{ij} of migration probabilities from bin j to any bin i out of a total of n bins at reconstruction level, and the events originating from background contributions at reconstruction level b_i, the expected observed event count $\tilde{y}_i$ can be expressed as:

$$\tilde{y}_i = \sum_{j=1}^{m} A_{ij}\tilde{x}_j + b_i \,. \tag{4.10}$$

The vector of observed event yields is then a random vector following a multivariate Gaussian distribution with mean $\tilde{y}$.

The process of estimating the stable-particle-level distribution $\tilde{\boldsymbol{x}}$ from the reconstructed event yields using the estimated background contribution $\boldsymbol{b}$ and the detector response $\boldsymbol{A}$ is called unfolding. Before the unfolding is performed, the estimated background contribution obtained from simulation is subtracted from the observed event yields. In the following, the vector $\boldsymbol{y}$ therefore denotes the background-subtracted event yields at reconstruction level. The vector $\boldsymbol{x}$ denotes the estimator of the stable-particle-level distribution. In this measurement, the TUnfold algorithm [159] is used following a least-squares minimization with:

$$\mathscr{L} = (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x})^{\mathrm{T}} \boldsymbol{V}_{yy}^{-1} (\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x})\,. \tag{4.11}$$

Here, $\boldsymbol{V}_{yy}$ is the covariance matrix of the background-subtracted vector $\boldsymbol{y}$. It is estimated from the statistical fluctuations of the observed event yields and includes the propagated statistical uncertainties introduced by the background subtraction. TUnfold also supports regularization and area constraints, which are not used in this measurement and therefore not described further.

The migrations between stable-particle level and reconstruction level are estimated from simulation. For this purpose, events passing the analysis selection at stable-particle level and reconstruction level are filled into a migration matrix, which estimates the detector response. A version of this matrix normalized in each stable-particle-level bin is shown in Figure 4.8. It represents the migration probabilities from stable-particle-level bins to reconstruction-level bins. On the x-axis, the global bin index of the unrolled distribution at stable-particle level is shown, while the corresponding reconstruction-level bins are shown on the y-axis.

The stability of the unfolding depends on the numerical properties of the response matrix. One measure of this stability is the condition number, which quantifies how strongly uncertainties or statistical fluctuations in the input distribution can be amplified when solving the inverse problem. Large condition numbers indicate an ill-conditioned response matrix and can lead to unstable unfolded results, making regularization necessary. The response matrix used in this measurement has a small condition number of 2.79. The unfolding problem is therefore well conditioned, and no regularization or area constraint is applied. This choice is validated with the closure checks in Subsection 4.4.2.

The unfolding procedure also corrects for losses due to identification requirements, reconstruction inefficiencies, or migrations out of the fiducial region. The corresponding detector efficiency accounts for events that are present at stable-particle level but are not reconstructed in the selected reconstruction-level phase space. In the response matrix, these events are filled into the reconstruction-level underflow bin for the corresponding stable-particle-level bin and are therefore taken into account during the unfolding.

For each stable-particle-level bin, the number of lost events is estimated from the difference between the total number of simulated events passing the stable-particle-level selection, N_{gen}, and the number of events passing both the stable-particle-level and reconstruction-level selections, N_{match}. All event counts in these definitions correspond to weighted simulated event yields. The corresponding efficiency is therefore defined as

$$\epsilon = \frac{N_{\mathrm{match}}}{N_{\mathrm{gen}}} = 1 - \frac{N_{\mathrm{loss}}}{N_{\mathrm{gen}}}\,, \tag{4.12}$$

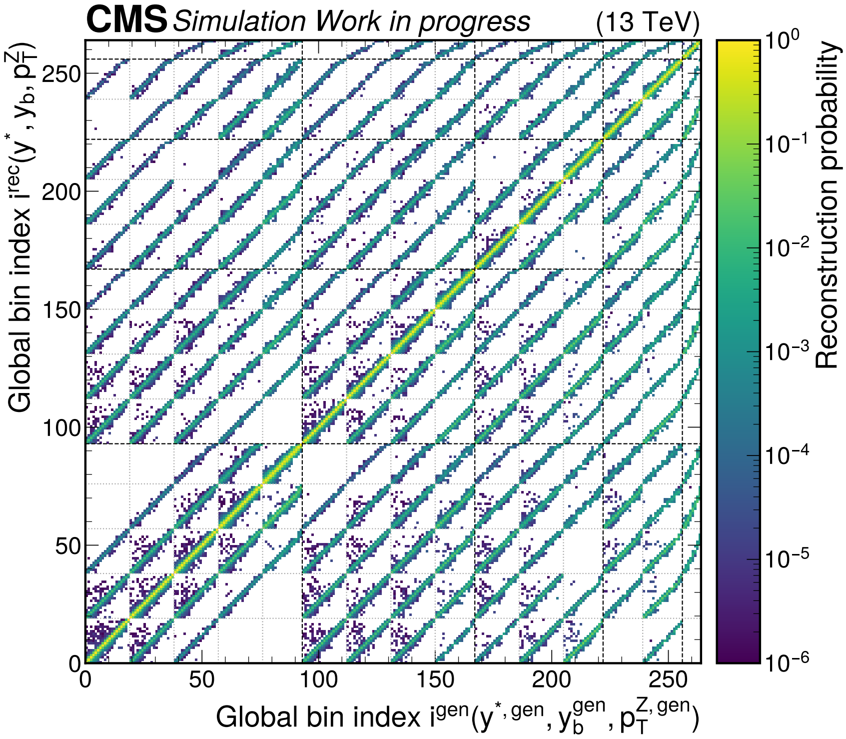


Figure 4.8: Migration probabilities from stable-particle level to reconstruction level estimated from simulation. The global bin index of the unrolled distribution at stable-particle level is shown on the x-axis, while the bin index at reconstruction level is shown on the y-axis. The condition number of the response matrix is 2.79.

with $N_{\mathrm{loss}} = N_{\mathrm{gen}} - N_{\mathrm{match}}$. The uncertainty in N_{loss} is estimated by error propagation assuming full correlation between N_{gen} and N_{match}, since the latter is a subset of the former.

Events that are present at reconstruction level but not at stable-particle level are referred to as out-of-acceptance contributions. They can arise, for example, from misidentified objects, pileup jets passing the jet selection, or resolution effects causing events outside the fiducial phase space to migrate into the selected reconstruction-level phase space. The out-of-acceptance contribution is estimated from simulated events passing the reconstruction-level selection but failing the stable-particle-level selection. It is calculated as a fraction of the full reconstruction-level event yield,

$$f_{\mathrm{out}} = \frac{N_{\mathrm{out}}}{N_{\mathrm{reco}}}\,, \tag{4.13}$$

where N_{out} is the number of out-of-acceptance events and N_{reco} is the total number of simulated events passing the reconstruction-level selection. This fraction is applied to the observed event yield in data and the resulting contribution is treated as an additional background contribution in the unfolding. The statistical uncertainty of the observed data yield is already included in the covariance matrix of the reconstructed input distribution. Therefore, only the additional uncertainty from the simulated estimate of the out-of-acceptance fraction is propagated separately.

The detector efficiency, $\epsilon = 1 - N_{\mathrm{loss}}/N_{\mathrm{gen}}$, and the fraction of true signal events at reconstruction level, $1 - f_{\mathrm{out}}$, are shown in Figure 4.9 for the central rapidity bin and two extreme bins in y_{b} and y^{*}.

The detector efficiency ranges from 40% at low transverse momenta to up to 80% at high transverse momenta in the central rapidity regions. At large y_{b} or y^{*}, the efficiency reaches values close to 70%. The out-of-acceptance contribution is largest at low transverse momenta, where pileup effects and migrations into the selected phase space play a sizeable role. In the central rapidity region, it starts at about 20% and decreases close to zero above 100 GeV. At large rapidities, the out-of-acceptance contribution reaches 30% to 40% in the lowest $p_{\mathrm{T}}^{\mathrm{Z}}$ bin and decreases to less than 5% above 100 GeV.

4.4.1 Statistical uncertainties and correlations

The statistical uncertainties of the background-subtracted reconstructed distribution are described by the covariance matrix $\boldsymbol{V}_{yy}$. This matrix contains the statistical uncertainty of the observed data and the propagated statistical uncertainties of the subtracted background contributions, including the out-of-acceptance contribution.

The covariance matrix of the unfolded result is obtained by propagating $\boldsymbol{V}_{yy}$ through the unfolding procedure. Since the unfolding is a global solution of the response equation, the resulting covariance matrix is in general nondiagonal, even if $\boldsymbol{V}_{yy}$ is diagonal. A statistical fluctuation in one reconstruction-level bin can affect several stable-particle-level bins, because events are redistributed according to the detector response. The off-diagonal entries of the unfolded covariance matrix therefore describe the statistical correlations between unfolded bins induced by bin migrations.

In the case without regularization and without area constraint, the unfolding result is a

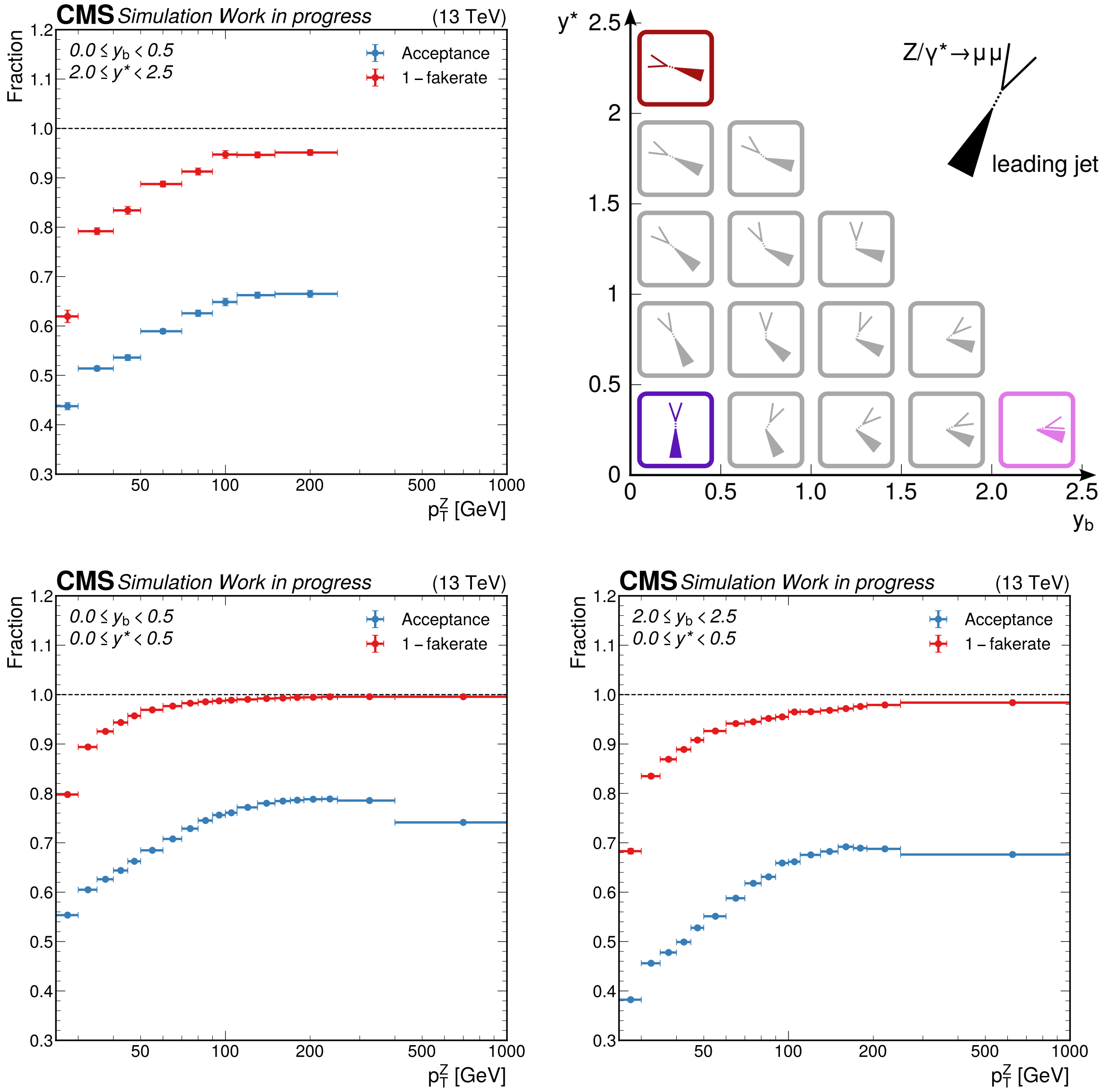


Figure 4.9: Detector efficiency and fraction of true signal events at reconstruction level shown as a function of p_T^Z. The distributions are shown for the central rapidity bin and for bins at large y_b and large y^*.

linear function of the input vector. The covariance matrix contribution originating from $\boldsymbol{V}_{\boldsymbol{yy}}$ can therefore be written as

$$\boldsymbol{V}_{\boldsymbol{xx}}^{\mathrm{stat}} = \boldsymbol{D}\,\boldsymbol{V}_{\boldsymbol{yy}}\,\boldsymbol{D}^{\mathrm{T}}\,, \tag{4.14}$$

where $\boldsymbol{D}$ is the matrix of derivatives of the unfolded result with respect to the reconstructed input yields,

$$D_{ij} = \frac{\partial x_i}{\partial y_j}\,. \tag{4.15}$$

This derivative matrix describes how a fluctuation in a reconstruction-level bin propagates to the unfolded stable-particle-level bins.

In addition to the covariance contribution from the reconstructed input distribution, the limited statistical precision of the simulated sample used to determine the detector response is taken into account. The unnormalized migration matrix is denoted by M_{ij}, where M_{ij} is the simulated event yield generated in stable-particle-level bin j and reconstructed in reconstruction-level bin i. Since this matrix is derived from a finite number of simulated events, its entries have statistical uncertainties. These uncertainties propagate first to the normalized response matrix A_{ij} used in the unfolding equation and then to the unfolded result.

Because the detector response is normalized in the unfolding procedure, a fluctuation in one entry of M_{ij} can affect several response probabilities. The resulting covariance contribution is therefore in general nondiagonal. Schematically, it can be written as

$$V_{kl}^{\mathrm{MC\ stat}} = \sum_{i,j} \frac{\partial x_k}{\partial M_{ij}} \frac{\partial x_l}{\partial M_{ij}} \sigma^2(M_{ij})\,, \tag{4.16}$$

where the derivatives include the normalization of the migration matrix to obtain A_{ij} and the subsequent unfolding step. For weighted simulated events, $\sigma^2(M_{ij})$ is estimated from the sum of squared event weights in the corresponding migration-matrix bin. This covariance matrix describes the statistical uncertainty due to the finite size of the simulated sample used to model the detector response.

The full statistical covariance matrix of the unfolded result is then obtained as the sum of the individual covariance contributions,

$$\boldsymbol{V}_{\boldsymbol{xx}}^{\mathrm{stat,total}} = \boldsymbol{V}_{\boldsymbol{xx}}^{\mathrm{data+bkg}} + \boldsymbol{V}_{\boldsymbol{xx}}^{\mathrm{MC\ stat}}\,. \tag{4.17}$$

Here, $\boldsymbol{V}_{\boldsymbol{xx}}^{\mathrm{data+bkg}}$ denotes the contribution from the statistical uncertainties of the reconstructed input distribution after background subtraction, while $\boldsymbol{V}_{\boldsymbol{xx}}^{\mathrm{MC\ stat}}$ denotes the contribution from the limited statistical precision of the response matrix. Both covariance matrices are symmetric and generally nondiagonal.

The corresponding correlation matrix is calculated from the covariance matrix according to

$$\rho_{ij} = \frac{V_{ij}}{\sqrt{V_{ii}V_{jj}}}\,, \tag{4.18}$$

where V_{ij} denotes an element of the covariance matrix of the unfolded result. The correlation matrix containing the total statistical uncertainty in the full Run 2 data, the detector response

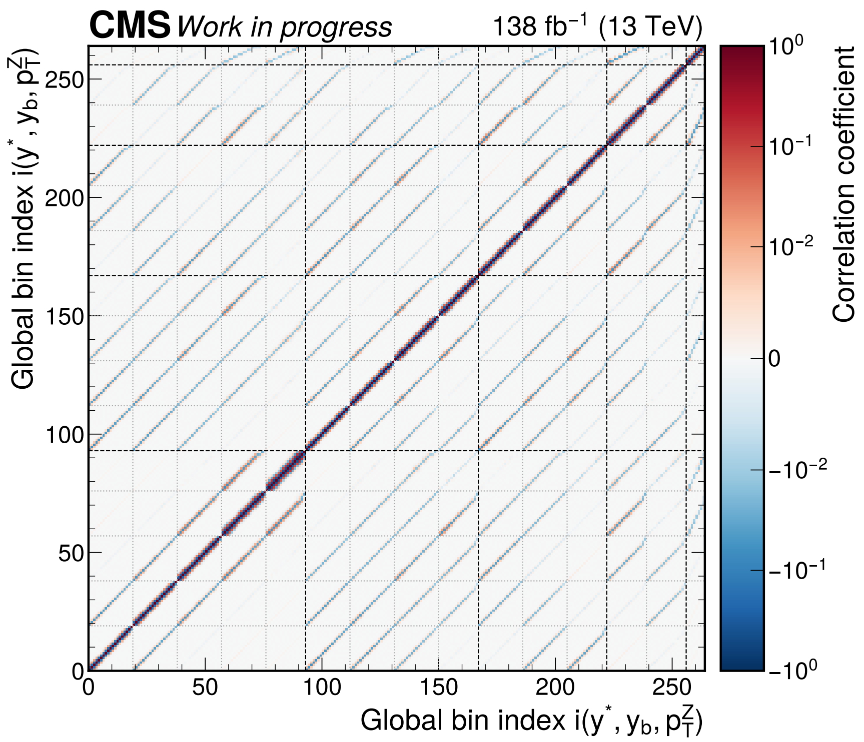


Figure 4.10: Correlation matrix including the statistical uncertainties on the full Run 2 data, the detector response and out-of-acceptance contribution obtained from the p_T^Z-binned MadGraph5 aMC@NLO FxFx + pythia 8 simulation, and simulated background contributions.

and out-of-acceptance contribution obtained from the p_T^Z-binned MadGraph5_aMC@NLO FxFx + pythia 8 simulation, and simulated background contributions is shown in Figure 4.10. Large anti-correlations can be seen in neighbouring p_T^Z bins, where the largest migrations occur, as well as smaller correlations of around 1% between (y_b, y^*) bins.

4.4.2 Closure check for the unfolding procedure

As a first technical closure check, simulated event yields at reconstruction level are unfolded using the migration matrix, detector efficiency, and out-of-acceptance fraction obtained from the same simulated sample. In this case, perfect agreement between the unfolded result and the corresponding stable-particle-level distribution is expected. This agreement is observed and shown in Figure 4.11.

Additionally, a cross-check of the numerical stability of the unfolding procedure is performed. The low condition number of the migration matrix already indicates that the unfolding problem is well conditioned. As an additional test, simulated event yields at reconstruction level are unfolded with a statistically independent response estimate obtained from the same MC generator. Here, the migration matrix, detector efficiency, and out-of-acceptance fraction are taken from an independent simulated sample. Differences between the unfolded result

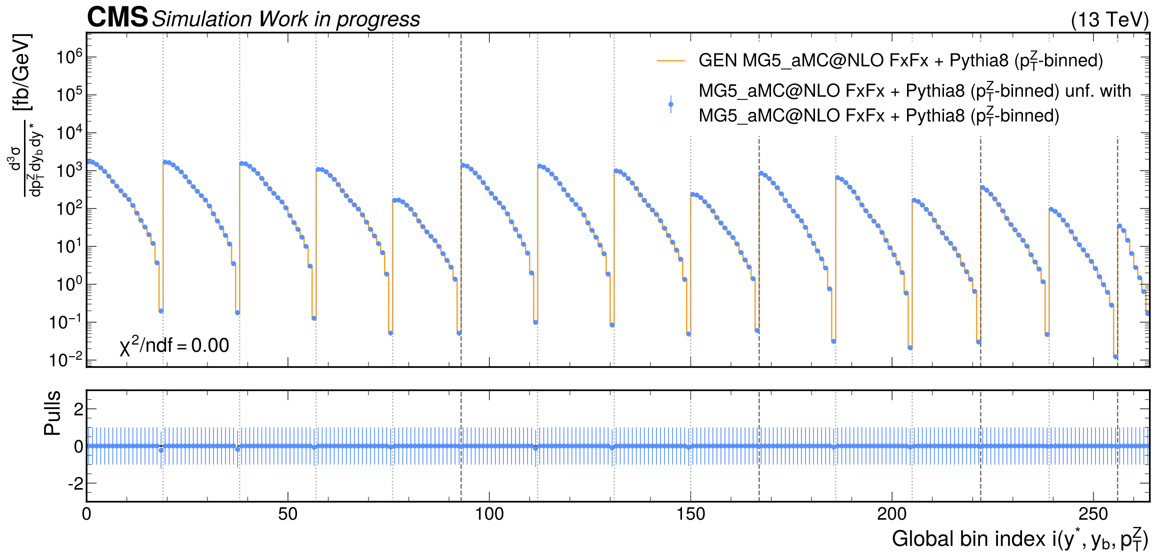


Figure 4.11: Closure check for the unfolding procedure using simulated event yields at reconstruction level and the detector response estimated from the same simulated sample. A near perfect closure with $\chi^2/\text{ndf} = 0$ is observed, with minimal numerical fluctuations in the high-p_T^Z bin.

and the corresponding stable-particle-level distribution are therefore expected to be covered by the statistical uncertainties of the response estimate.

For this test, the high-statistics p_T^Z-binned sample is used as pseudo-data at reconstruction level and unfolded with the response obtained from the lower-statistics inclusive signal sample. The unfolded result is then compared to the stable-particle-level distribution of the p_T^Z-binned sample. Since the two simulated samples are statistically independent and the pseudo-data sample has much larger statistical precision, the uncertainty of the comparison is dominated by the statistical precision of the response estimate. The non-trivial correlations between the pseudo-data at reconstruction level and the stable-particle-level reference distribution can therefore be neglected when quantifying the agreement. The χ^2 is calculated using the covariance matrix associated with the statistical uncertainty of the independently estimated response and the uncertainty in the estimated out-of-acceptance contribution. Figure 4.12 shows the result of this cross-check.

Since the samples used for the pseudo-data and for the response estimate are statistically independent, the closure check is expected to yield a value of χ^2/ndf close to unity. The observed value of $\chi^2/\text{ndf} = 1.08$ confirms the stability of the unfolding procedure.

4.4.3 Model dependence of the unfolding procedure

A cross-check is performed to test whether the unfolding procedure is sensitive to the choice of the event generator used to estimate the detector response from simulated events. First, the description of the reconstructed data is compared for the alternative samples. The effect of these alternative response models is then tested by unfolding the data with each response.

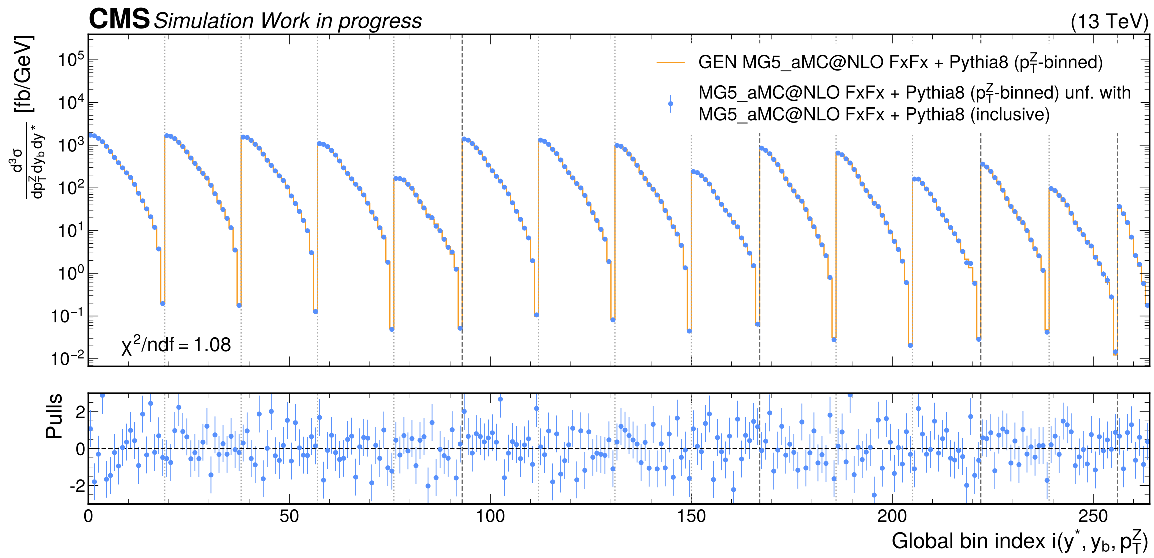


Figure 4.12: Closure check using simulated event yields at reconstruction level unfolded with a statistically independent detector response estimate. The unfolded result is compared to the corresponding stable-particle-level distribution. An overall compatibility of $\chi^2/\text{ndf} = 1.08$ is observed.

Alternative samples are simulated at LO for the production of a Z boson in association with up to four jets using MLM merging. One sample is interfaced to PYTHIA 8, while the other is interfaced to HERWIG 7 [137] using the CH3 tune [138]. The simulation of parton showering, hadronization, and the underlying event differs between the two samples. Additionally, the samples differ also in their perturbative precision being only at LO for the Z + 1 jet phase space, while also including real-emission contributions with additional jets.

The description of the data by these simulations at reconstruction level is shown in Figure 4.13. Overall, the agreement between data and simulation is worse than for the nominal simulation used in the unfolding. Especially the HERWIG 7 sample shows pronounced structures in the data-to-simulation ratio for the p_T^Z distribution in all (y_b, y^*) bins. This can affect the unfolding procedure, since the absolute number of migrations from bins with a badly modelled event yield to neighbouring p_T^Z bins is also affected. While the PYTHIA 8 sample also does not describe the data well, the ratio is smoother and the overall shape of the data is described better than by the HERWIG 7 sample. A smaller dependence on the response model is therefore expected for this sample.

The simulated migrations during reconstruction obtained with the two samples are shown in Figure 4.14. The condition number is 2.77 for the PYTHIA 8 sample and 2.93 for the HERWIG 7 sample. Both response matrices are therefore similarly well conditioned as the nominal response matrix and suitable for the same unfolding procedure without regularization.

For the 2018 data-taking period only, another alternative signal simulation using POWHEG MiNNLO$_{\text{PS}}$ + PYTHIA 8 + PHOTOS is available. This sample has NNLO accuracy in perturbative QCD for inclusive Drell–Yan production without requiring additional jets. For the Z + 1 jet phase space relevant for this analysis, it is therefore only of NLO accuracy. The nominal

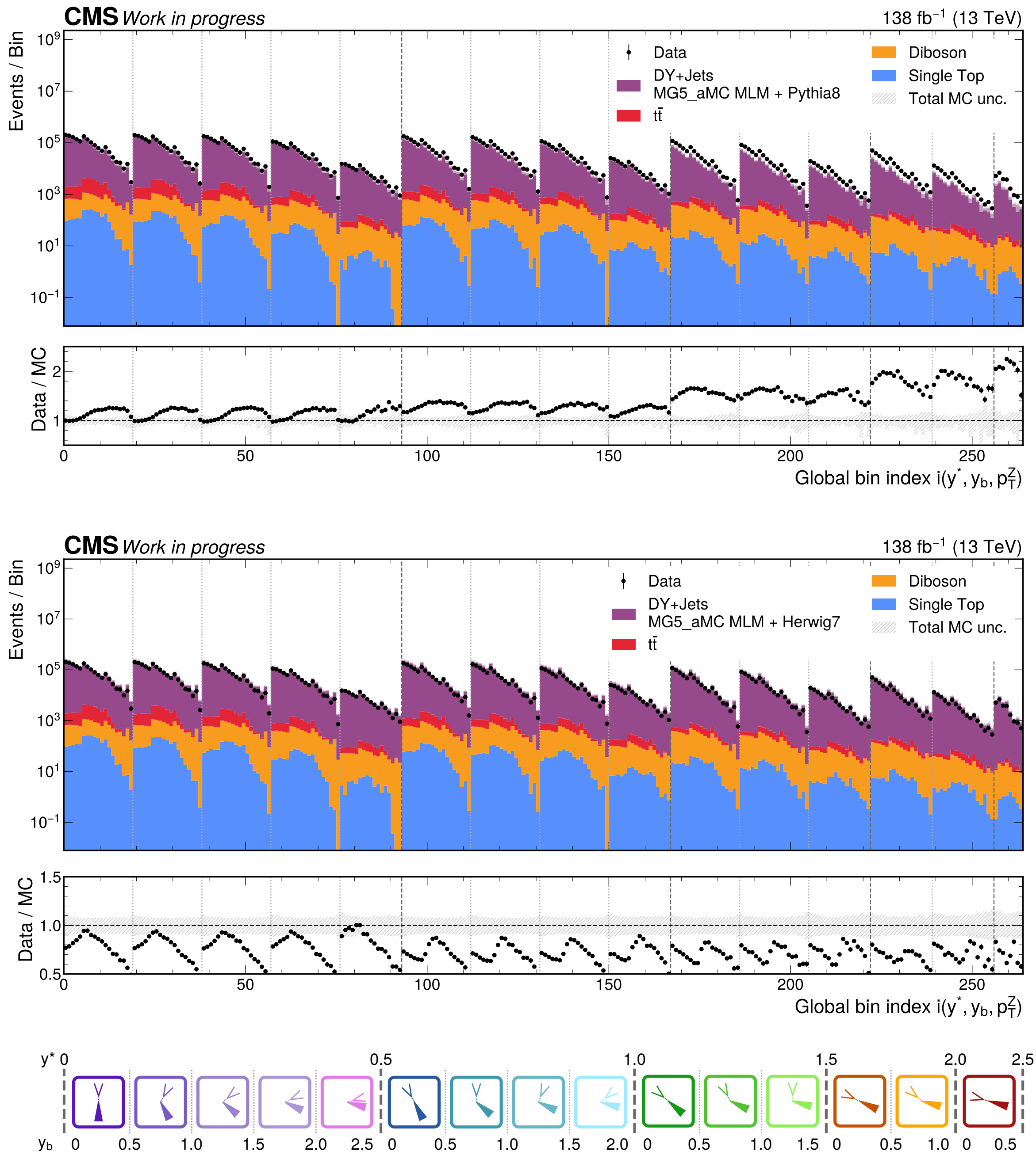


Figure 4.13: Unrolled triple-differential distribution comparing the event yields in data to LO simulation interfaced to PYTHIA 8 (top) and HERWIG 7 (bottom). The PYTHIA 8 sample gives a reasonable description of the shape of the p_T^Z distribution within individual (y_b, y^*) bins. The HERWIG 7 sample shows a worse description of the p_T^Z shape, visible as pronounced structures in the ratio. The overall normalization and rapidity dependence are not well described by either alternative sample compared to the nominal simulation.

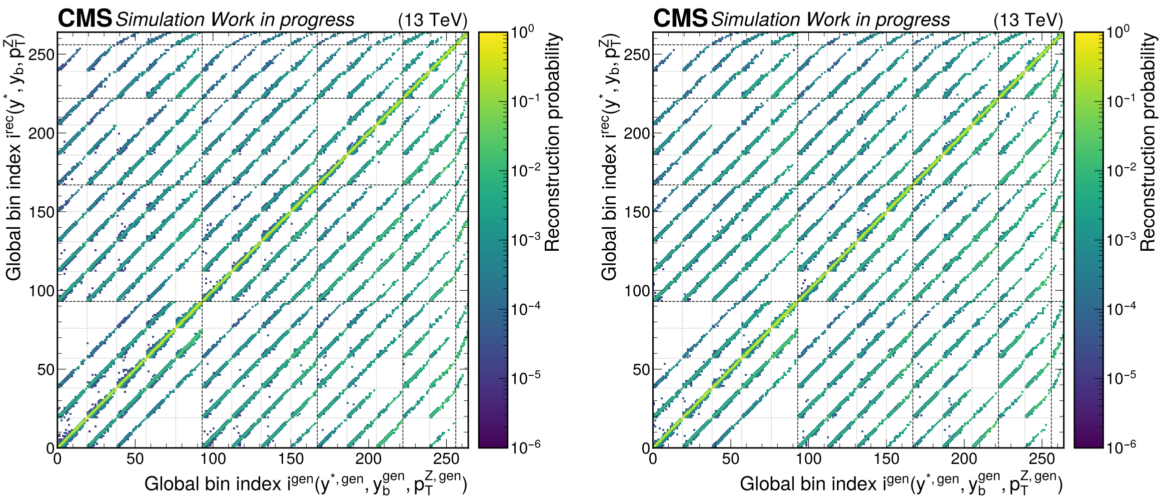


Figure 4.14: Migration matrices predicted by the alternative signal simulations. On the left, the matrix obtained from the LO sample interfaced to PYTHIA 8 is shown with a condition number of 2.77. On the right, the matrix obtained from the LO sample interfaced to HERWIG 7 is shown with a condition number of 2.93.

MADGRAPH5_aMC@NLO FxFx sample merges the Z + 2 jet contribution at NLO, while the MiNNLO$_{\text{PS}}$ sample includes this contribution only at LO. The comparison between data and simulated event yields at reconstruction level for the full unrolled phase space is shown in Figure 4.15, and the corresponding detector response is shown in Figure 4.16.

The data are unfolded using these three alternative simulations and compared to the nominal unfolded result across all 264 bins in Figure 4.17. The lower panels show the bin-wise pulls calculated from the diagonal entries of the combined covariance matrix from both results introduced later in Equation 4.25. These pulls indicate the bins in which the largest local differences occur. It does not, however, display the bin-to-bin correlations. The correlations are only taken into account in the global compatibility test using the full covariance matrix.

The agreement between the unfolded cross sections is quantified using a χ^2 calculation. For each alternative response model, the difference vector

$$\Delta \boldsymbol{x} = \boldsymbol{x}_{\text{alt}} - \boldsymbol{x}_{\text{nom}} \,. \tag{4.19}$$

is calculated, where $\boldsymbol{x}_{\text{nom}}$ is the nominal unfolded result and $\boldsymbol{x}_{\text{alt}}$ is the result obtained with the alternative detector response. Since both unfolded results are obtained from the same data, the statistical uncertainties of the observed event yields and of the subtracted background contributions are correlated between the two results. They are therefore propagated to the difference using the difference of the corresponding unfolding derivative matrices,

$$\boldsymbol{V}_{\Delta\boldsymbol{x}}^{\text{data}} = (\boldsymbol{D}_{\text{alt}} - \boldsymbol{D}_{\text{nom}})\, \boldsymbol{V}_{\boldsymbol{yy}}\, (\boldsymbol{D}_{\text{alt}} - \boldsymbol{D}_{\text{nom}})^{\text{T}} \,, \tag{4.20}$$

and

$$\boldsymbol{V}_{\Delta\boldsymbol{x}}^{\text{bkg}} = (\boldsymbol{D}_{\text{alt}} - \boldsymbol{D}_{\text{nom}})\, \boldsymbol{V}_{\boldsymbol{bb}}\, (\boldsymbol{D}_{\text{alt}} - \boldsymbol{D}_{\text{nom}})^{\text{T}} \,. \tag{4.21}$$

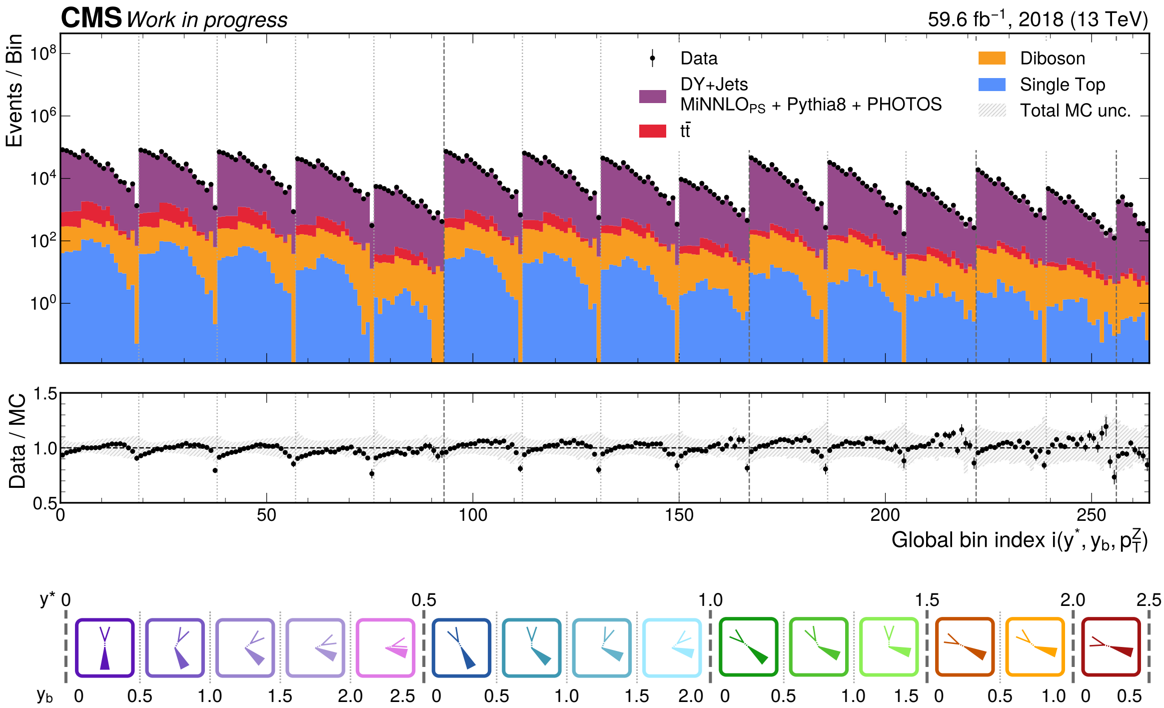


Figure 4.15: Data-to-MC comparison for the POWHEG MiNNLO$_{PS}$ + PYTHIA 8 + PHOTOS simulation for the 2018 data-taking period in the full unrolled phase space.

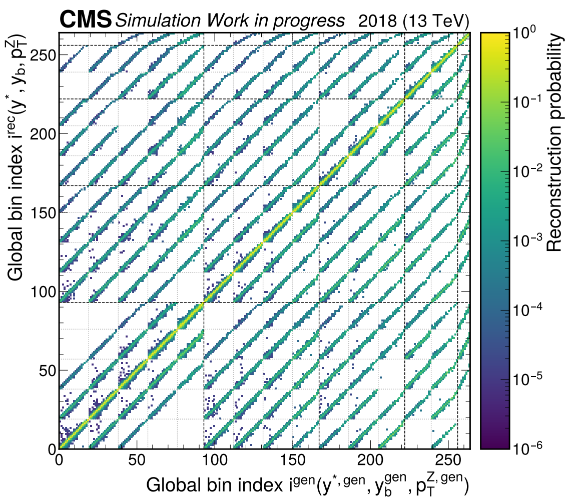
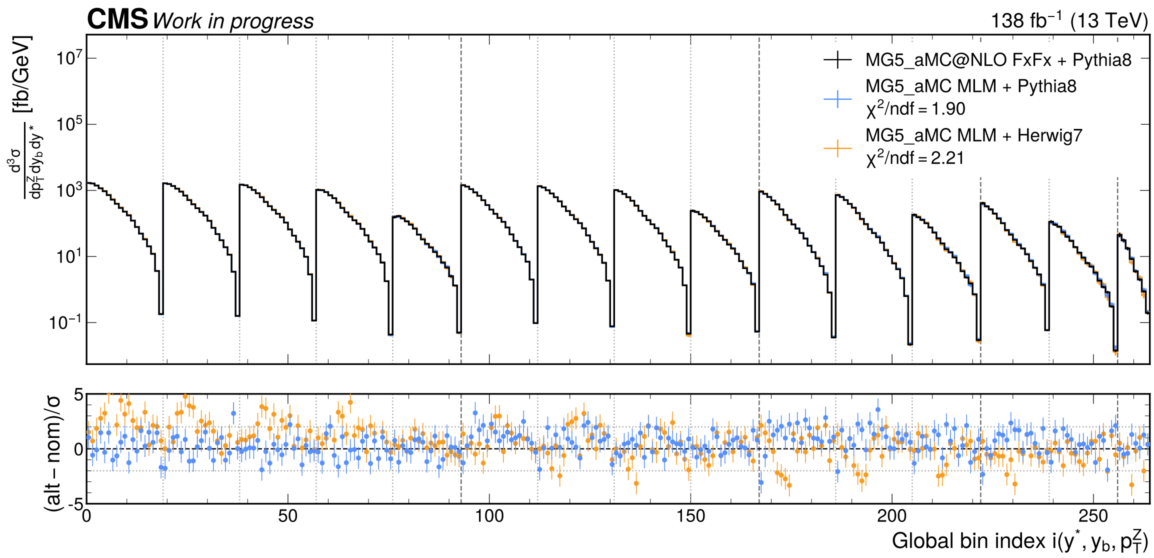


Figure 4.16: Migration matrix for the POWHEG MiNNLO$_{PS}$ + PYTHIA 8 + PHOTOS simulation. The condition number is 2.76.



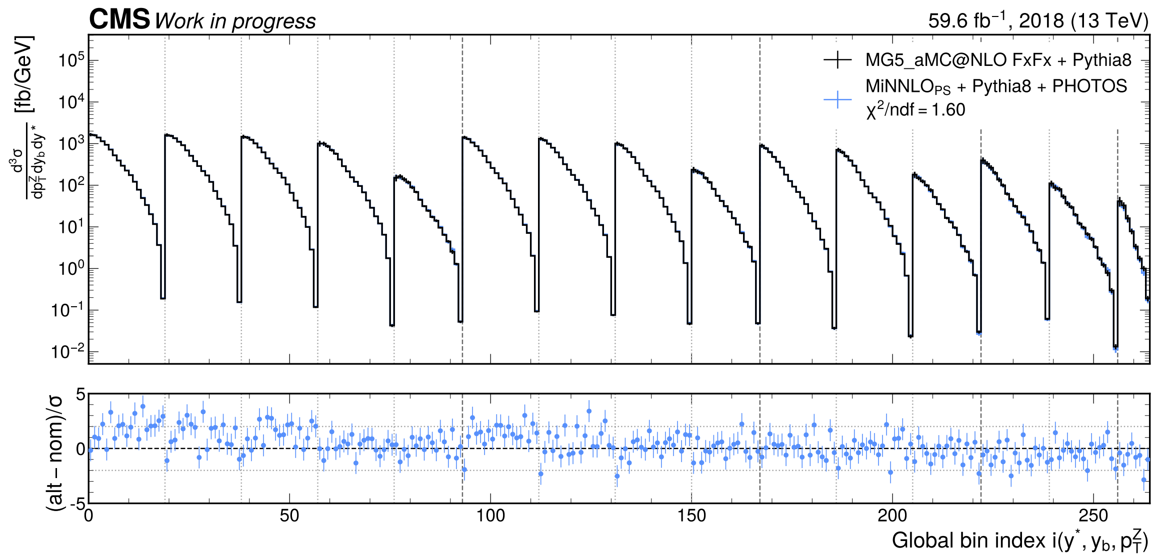


Figure 4.17: Comparison of the unfolded data using alternative simulated samples. The full Run 2 data unfolded with the nominal MadGraph5_amc@nlo FxFx + pythia 8 simulation is compared to the results obtained using samples simulated with MadGraph5_amc@nlo and MLM merging interfaced to pythia 8 or herwig 7 (top). Additionally, for the 2018 data-taking period only, the unfolded data obtained with the nominal MadGraph5_amc@nlo FxFx + pythia 8 simulation are compared to the result obtained using powheg MiNNLO$_{PS}$ interfaced to pythia 8 and photos (bottom). The lower panels show the bin-wise pull between the unfolded result obtained with the alternative response model and the nominal unfolded result. The pull is calculated using only the diagonal entries of the covariance matrix of the difference and therefore does not visualize the bin-to-bin correlations. These correlations are, however, included in the quoted χ^2/ndf values through the full covariance matrix.

Here, D_{nom} and D_{alt} are the derivative matrices of the unfolded result (given in Equation 4.14 and 4.15) with respect to the reconstructed input distribution for the nominal and alternative response models, respectively. The matrices V_{yy} and V_{bb} denote the statistical covariance matrices of the observed data and of the subtracted background estimates before unfolding. This treatment accounts for the fact that the same statistical fluctuations in data are propagated through two different response matrices.

The statistical uncertainties from the limited simulated samples used to estimate the detector response and the out-of-acceptance fraction are treated as uncorrelated between the nominal and alternative simulations. Their contribution to the covariance matrix of the difference is therefore given by

$$V_{\Delta x}^{\mathrm{MC\,stat}} = V_{\mathrm{nom}}^{\mathrm{MC\,stat}} + V_{\mathrm{alt}}^{\mathrm{MC\,stat}} \,. \tag{4.22}$$

For systematic uncertainties, the same nuisance parameter is assumed to be fully correlated between the two unfolded results. For each source, the signed one-standard-deviation shift is estimated from the up and down variations as

$$s = \frac{1}{2}\left(x^{\mathrm{up}} - x^{\mathrm{down}}\right). \tag{4.23}$$

The corresponding contribution to the covariance matrix of the difference is then calculated from the difference of the signed shifts,

$$V_{\Delta x}^{\mathrm{syst}} = \left(s_{\mathrm{alt}} - s_{\mathrm{nom}}\right)\left(s_{\mathrm{alt}} - s_{\mathrm{nom}}\right)^{\mathrm{T}} . \tag{4.24}$$

The total covariance matrix used for the compatibility test is obtained by adding all contributions,

$$V_{\Delta x} = V_{\Delta x}^{\mathrm{data}} + V_{\Delta x}^{\mathrm{bkg}} + V_{\Delta x}^{\mathrm{MC\,stat}} + \sum_{k} V_{\Delta x}^{\mathrm{syst},k} \,. \tag{4.25}$$

The compatibility is quantified as

$$\chi^2 = \Delta x^{\mathrm{T}} V_{\Delta x}^{-1} \Delta x \,. \tag{4.26}$$

Overall, the unfolded results obtained with the alternative response models are in reasonable agreement with the nominal unfolded result. Given the poor description of the data by the MLM-merged samples at reconstruction level, values of χ^2/ndf around two show that the unfolded result is stable even under sizeable changes of the simulated sample used to estimate the detector response. As shown in Figure 4.15, the MiNNLO$_{\mathrm{PS}}$ sample gives a better description of the data at reconstruction level towards large y^*, while shape differences in $p_{\mathrm{T}}^{\mathrm{Z}}$ within individual (y_{b}, y^*) are larger compared to the nominal MadGraph5_aMC@NLO FxFx simulation shown in Figure 4.7. The unfolding using the MiNNLO$_{\mathrm{PS}}$ simulation gives a result compatible with the nominal unfolding, with $\chi^2/\mathrm{ndf} = 1.6$. The observed differences between the two unfolded results are covered by the uncertainties and their correlations and do not indicate a coherent bias of the unfolded result.

4.5 Systematic uncertainties

In addition to the statistical uncertainties discussed in Subsection 4.4.1, which include nontrivial correlations introduced by the unfolding procedure, systematic uncertainties are considered in this measurement. Each systematic uncertainty source is treated as fully correlated across all bins and is propagated through the unfolding procedure by varying the response matrix, the acceptance correction, and the out-of-acceptance estimate according to the corresponding uncertainty. Upward and downward variations are evaluated separately to account for asymmetric uncertainties.

Uncertainties in the jet energy scale (JES) and jet energy resolution (JER) directly affect the reconstructed transverse momentum of jets and therefore play an important role in the low-p_T region of this measurement, where they can change whether jets pass or fail the minimum p_T requirement. In addition, these uncertainties affect the reconstructed jet rapidity and therefore also y_b and y^*. A smoothing procedure is applied to the individual sources of the full jet energy scale (JES) uncertainty to reduce unphysical fluctuations in their magnitude between neighbouring bins. A breakdown of the individual sources of the full JES uncertainty and the smoothing procedure is given in Subsection 4.5.1.

Uncertainties in the simulated background estimate are discussed in Subsection 4.5.2. Further sources of uncertainty in the final result arise from the integrated luminosity, the pileup jet identification, the muon identification and reconstruction efficiencies, and the estimation of the PU profile used in simulation. A full overview of all uncertainty sources and of the resulting total uncertainty is given in Subsection 4.5.3.

4.5.1 Jet energy uncertainties

The uncertainty in the jet energy calibration is split into several independent sources related to different aspects of the jet calibration. Each uncertainty is propagated separately through the full analysis chain. For each source, all jet four-momenta are shifted up and down by one standard deviation, the event selection and unfolding are repeated, and the corresponding change in the measured cross section is taken as the final uncertainty.

The individual JES uncertainty sources are assigned different correlations between the data-taking years 2016, 2017, and 2018, summarized in Table 4.4. The jet energy resolution (JER) uncertainties are treated as uncorrelated among the data-taking periods. To account for them consistently in the combined Run 2 result, each source is decomposed into a part that is correlated across all years and a part that is uncorrelated between the years. For a source s with correlation coefficient ρ_s, the variation in year y is split as

$$\Delta_s^y = \Delta_{s,\text{corr}}^y + \Delta_{s,\text{uncorr}}^y \tag{4.27}$$

with

$$\Delta_{s,\text{corr}}^y = \sqrt{\rho_s}\,\Delta_s^y \qquad \text{and} \qquad \Delta_{s,\text{uncorr}}^y = \sqrt{1-\rho_s}\,\Delta_s^y\,. \tag{4.28}$$

In this way, a source with $\rho_s = 1$ is treated as fully correlated, a source with $\rho_s = 0$ as fully uncorrelated. Uncertainties with 50% correlation are then given by a combination of both components. After this decomposition, all components are propagated independently and

Table 4.4: Correlations of the individual jet energy scale uncertainty sources between the data-taking years 2016, 2017, and 2018.

jet energy scale uncertainty	correlation
AbsoluteMPFBias	100%
AbsoluteScale	100%
AbsoluteStat	0%
FlavorQCD	100%
Fragmentation	100%
PileUpDataMC	50%
PileUpPtBB	50%
PileUpPtEC1	50%
PileUpPtEC2	50%
PileUpPtHF	50%
PileUpPtRef	50%
RelativeFSR	50%
RelativeJEREC1	0%
RelativeJEREC2	0%

jet energy scale uncertainty	correlation
RelativeJERHF	50%
RelativePtBB	50%
RelativePtEC1	0%
RelativePtEC2	0%
RelativePtHF	50%
RelativeBal	50%
RelativeSample	0%
RelativeStatEC	0%
RelativeStatFSR	0%
RelativeStatHF	0%
SinglePionECAL	100%
SinglePionHCAL	100%
TimePtEta	0%

treated in the same way as the other systematic uncertainties discussed in the beginning of this section.

For presentation and for later applications of the result, it is convenient to group the individual JES sources into subtotal categories. The composition of these subtotals is listed in Table 4.5.

The propagated JES uncertainties are susceptible to statistical fluctuations. In particular, neighbouring $p_{\mathrm{T}}^{\mathrm{Z}}$ bins can show sizeable bin-to-bin variations that do not reflect a physical change in the underlying uncertainty. Such fluctuations are undesirable when comparing the measurement to theory predictions or when using the result in PDF fits. Therefore, a smoothing procedure is applied to the final JES uncertainties similar to Ref. [160].

The smoothing is performed separately for each individual JES source and for each y_{b}-y^{*} bin as a function of $p_{\mathrm{T}}^{\mathrm{Z}}$. For a given source, both the nominal unfolded cross section and the corresponding shifted result are fitted with a smooth function. Since the cross section falls approximately exponentially with $p_{\mathrm{T}}^{\mathrm{Z}}$, the logarithm of the cross section is fitted instead of the cross section itself. In addition, the $p_{\mathrm{T}}^{\mathrm{Z}}$ range extends from 25 GeV to 1000 GeV, and the binning becomes wider towards large $p_{\mathrm{T}}^{\mathrm{Z}}$. Therefore, the fit is performed as a function of $\ln(p_{\mathrm{T}}^{\mathrm{Z}})$.

The cross section is fitted using a linear combination of Chebyshev polynomials, which are a set of orthogonal functions in the interval $[-1, 1]$. These polynomials are given by a recursive formula

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)\,, \qquad T_0(x) = 1\,, \qquad T_1(x) = x\,. \tag{4.29}$$

An arbitrary function in this interval can then be approximated with a linear combination of

Table 4.5: Composition of the jet energy scale uncertainty subtotals. The sources FlavorQCD and TimePtEta are not included in any subtotal.

subtotal	included jet energy scale uncertainty sources
SubTotalPileUp	PileUpDataMC, PileUpPtBB, PileUpPtEC1, PileUpPtEC2, PileUpPtHF, PileUpPtRef
SubTotalRelative	RelativeFSR, RelativeJEREC1, RelativeJEREC2, RelativeJERHF, RelativePtBB, RelativePtEC1, RelativePtEC2, RelativePtHF, RelativeBal, RelativeSample, RelativeStatEC, RelativeStatFSR, RelativeStatHF
SubTotalPt	Fragmentation, SinglePionECAL, SinglePionHCAL
SubTotalScale	AbsoluteMPFBias, AbsoluteScale, AbsoluteStat

the polynomials up to order k and the corresponding coefficients $\theta = (\theta_0, \dots, \theta_k)$:

$$f_k(\tilde{x}, \theta) = \sum_{n=0}^{k} \theta_n T_n(\tilde{x})\,. \tag{4.30}$$

To match the definition interval of the Chebyshev polynomials are defined, the logarithm of $p_{\mathrm{T}}^{\mathrm{Z}}$ is mapped to the interval $[-1, 1]$,

$$\tilde{x}(p_{\mathrm{T}}^{\mathrm{Z}}) = \frac{2\ln(p_{\mathrm{T}}^{\mathrm{Z}} / p_T^{\mathrm{Z,min}})}{\ln(p_T^{\mathrm{Z,max}} / p_T^{\mathrm{Z,min}})} - 1\,, \tag{4.31}$$

where $p_T^{\mathrm{Z,min}}$ and $p_T^{\mathrm{Z,max}}$ denote the lower and upper boundaries of the fitted $p_{\mathrm{T}}^{\mathrm{Z}}$ range in the respective y_{b}-y^* bin.

The corresponding fit function for the cross section is then given by

$$\sigma_{\mathrm{fit}}(p_{\mathrm{T}}^{\mathrm{Z}}; \theta) = \exp\Big(f_k(\tilde{x}(p_{\mathrm{T}}^{\mathrm{Z}}), \theta)\Big)\,. \tag{4.32}$$

The fit parameters θ are determined from a χ^2 minimization.

For each JES source, this procedure yields a smooth description of the nominal unfolded spectrum, $\sigma_{\mathrm{fit}}^{\mathrm{nom}}$, and of the shifted spectrum, $\sigma_{\mathrm{fit}}^{\mathrm{shift}}$. The smoothed relative uncertainty in bin i is then obtained from the relative difference of the two fitted spectra,

$$\delta_i^{\mathrm{smooth}} = \frac{\sigma_{\mathrm{fit},i}^{\mathrm{shift}} - \sigma_{\mathrm{fit},i}^{\mathrm{nom}}}{\sigma_{\mathrm{fit},i}^{\mathrm{nom}}}\,. \tag{4.33}$$

This replaces the raw bin-by-bin variation by a smooth uncertainty shape while preserving the overall size and trend of the corresponding systematic effect.

The impact of the smoothing on the total JES uncertainty is illustrated in Figure 4.18 for the FlavorQCD and TimePtEta JES uncertainty contributions as well as the combination of all

JES uncertainty sources before and after smoothing. A full breakdown for all (y_b, y^*)-bins is shown in Figure A.11 in the appendix. Overall a reduction in large unphysical fluctuations between neighbouring bins for the single uncertainty sources is observed while only having a small impact on the combined results.

The final jet energy uncertainties are shown in Figure 4.19 for the central, high-y_b and high-y^* bins including the PileUp, Relative, Pt, and Scale subtotal contributions. A full breakdown in all (y_b, y^*)-bins is shown in Figure A.12 in the appendix. In the high rapidity regions, the subtotal relative uncertainty is the main contributing factor to the total JES uncertainty, reaching up to 10% at low transverse momenta. In the central rapidity region, the pileup subtotal is the main source of JES uncertainty at low transverse momenta. Towards high-p_T^Z the relative subtotal dominates the total JES uncertainty across the full phase space. Also, the JER uncertainty is shown, and found to be one to two orders of magnitude smaller than the total JES uncertainty.

4.5.2 Background uncertainties

The background contribution that is subtracted from the observed event yields in data before unfolding is estimated from simulation. The simulated background samples are produced at NLO accuracy in perturbative QCD, as described in Section 4.2. The uncertainty due to missing higher-order QCD contributions is estimated by varying the factorization scale μ_F and the renormalization scale μ_R independently by factors of 0.5 and 2 relative to their nominal value μ_0. The uncertainty is obtained from the envelope of the corresponding yield variations in each analysis bin. The extreme combinations $(\mu_R/\mu_0, \mu_F/\mu_0) = (0.5, 2.0)$ and $(2.0, 0.5)$ are excluded.

In addition, the uncertainty associated with the choice of the PDF set is taken into account using the PDF4LHC15 set [56]. This set is constructed from the CT14 [161], MMHT2014 [162], and NNPDF3.0 [163] PDF sets. The corresponding uncertainty is evaluated using the prescription for PDF4LHC15 with the reduced set of 30 Hessian eigenvector variations [56, 164], including the associated α_S variations. The PDF and α_S contributions are added in quadrature to obtain the total PDF uncertainty.

Additional uncertainties arise from the modelling of initial-state radiation (ISR) and final-state radiation (FSR) radiation in the parton shower. These uncertainties are estimated using the PYTHIA 8 parton-shower variation weights. The weights vary the renormalization scale used for the evaluation of α_S in the shower evolution by factors of 0.5 and 2.0, separately for ISR and FSR. The nominal scale is given by the transverse-momentum-ordered shower evolution scale. Since rare events can receive very large variation weights, leading to large statistical fluctuations in sparsely populated bins, the shower-variation weights are truncated to the interval $[0, 100]$. The resulting uncertainty is obtained from the corresponding variations of the simulated background yields.

Furthermore, the overall normalization of the top quark background, consisting of top quark-antiquark pair production and single top quark production, is varied by ±10% to account for potential mismodelling effects. This variation is motivated by comparisons with eμ data performed in Ref. [165], which show small discrepancies of a few percent between data and

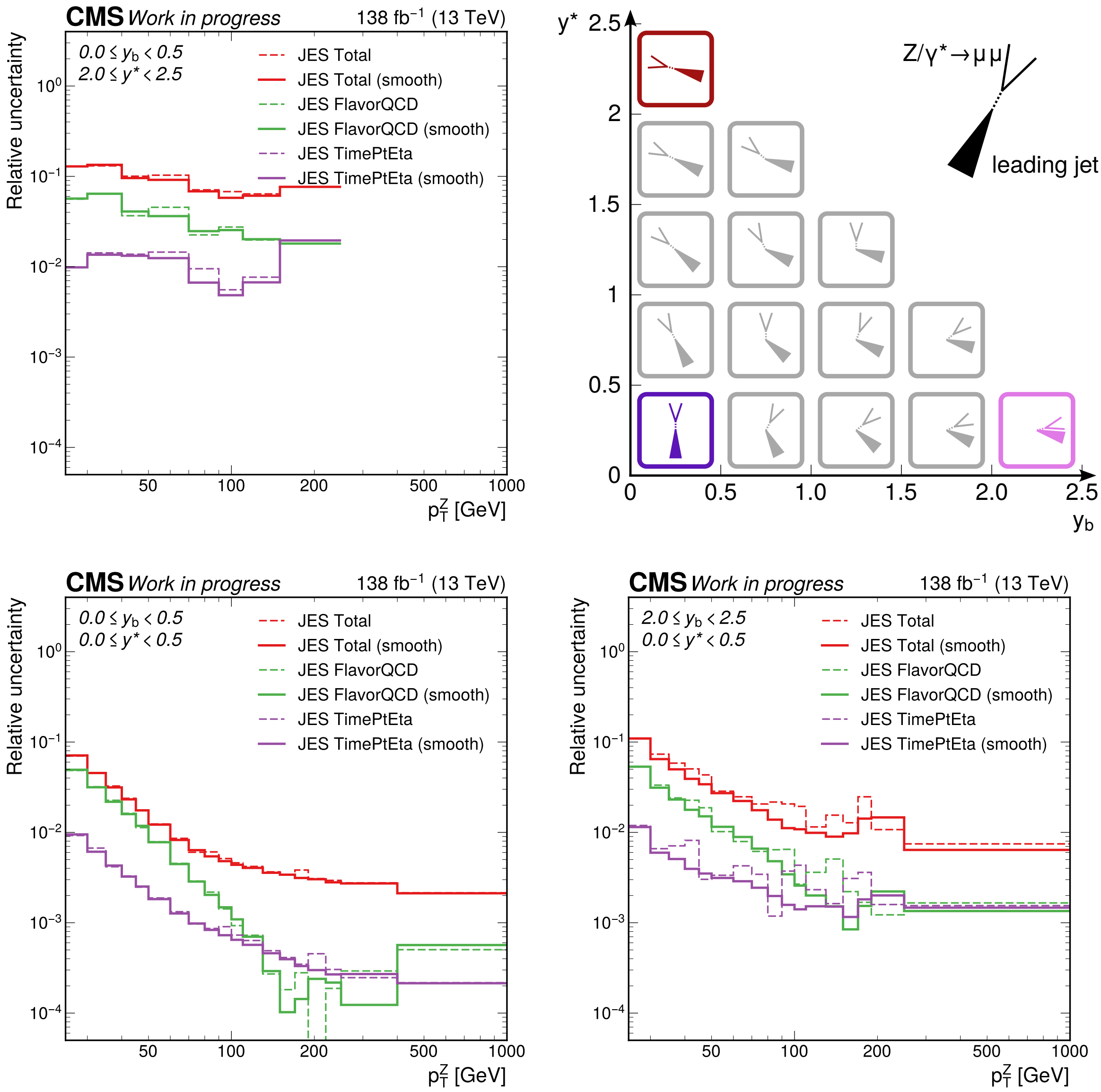


Figure 4.18: FlavorQCD and TimePtEta JES uncertainty contributions (green and purple) before smoothing (dashed lines) and after smoothing (solid lines). Also, the combination of all JES uncertainty sources (red) before and after smoothing is shown.

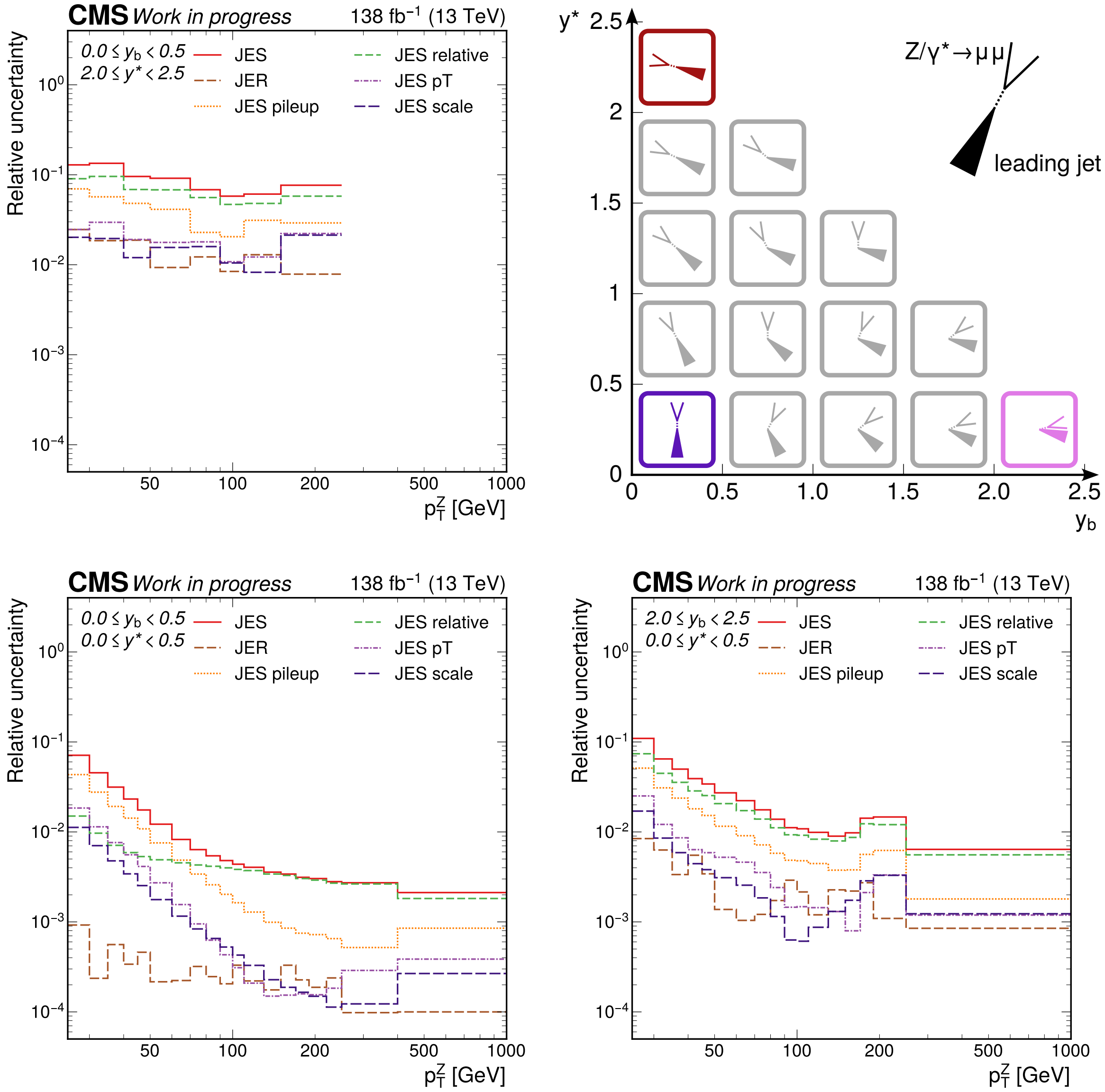


Figure 4.19: Final relative jet energy uncertainties as a function of p_T^Z for representative y_b-y^* bins. Shown are the total JES and JER uncertainties as well as the PileUp, Relative, Pt, and Scale contributions to the full JES uncertainty.

simulation in p_T^Z, y_b, and y^*. Since the overall contribution of this background is small, the resulting uncertainty in the final result is below 0.1%.

An additional contribution from events with at least one nonprompt or misidentified muon is evaluated using a W simulation merged with up to two additional jets at NLO. For this estimate, the jet-binned MC samples listed in Table A.6 in the appendix are used for the full Run 2 in order to obtain the highest possible statistical precision. The event yields obtained from this simulation are found to be below 0.01% in every bin of the measurement. This contribution is therefore neglected.

Since the overall background contribution from diboson and top quark production ranges only from 0.3% to 5.6% across all measurement bins, the total uncertainty in the final measurement from the background estimation is about 0.03% to 1.5% and plays only a minor role in the total uncertainty budget. Only in the central rapidity regions around $p_T^Z \approx 100\,\text{GeV}$, where the highest overall precision is achieved, the uncertainty in the background estimation contributes visibly to the total uncertainty. Even in these bins, it remains smaller than the uncertainties in the integrated luminosity and the PUJetID.

4.5.3 Total uncertainty

The per-bin total uncertainty is obtained by adding all uncertainty sources in quadrature. Figure 4.20 shows the total uncertainty and a breakdown of the individual uncertainty sources in the central rapidity region, the high-y_b bin, and the high-y^* bin. A full overview across all (y_b, y^*)-bins is given in the appendix in Figure A.13. For asymmetric uncertainties, the larger variation is shown.

The statistical uncertainty in the data, labelled as "Data stat." in Figure 4.20, is around 0.5% in the central rapidity bins and reaches up to 6.5% in the high-y^* and high-p_T^Z bins. The uncertainty labelled as "Unfold stat." includes the statistical uncertainty in the estimated detector response, efficiency, and out-of-acceptance contribution, as described in Subsection 4.4.1. It ranges from 0.25% in the central rapidity region and low-p_T^Z bins up to 5% in the highest-y^* and highest-p_T^Z bins.

The largest contribution at low transverse momenta across all (y_b, y^*)-bins arises from the uncertainty in the jet energy scale (JES). In the overview plot, the combined contribution from the jet energy scale and jet energy resolution (JER) uncertainties is shown. This uncertainty reaches up to 14% at low p_T^Z in the outer rapidity regions. In the central rapidity region, it is smaller than 8% and decreases to less than 0.5% at high p_T^Z. A breakdown of the jet-energy-related uncertainties, including the smoothing procedure applied to the individual JES uncertainty sources and their correlations across data-taking periods, is described in more detail in Subsection 4.5.1. Uncertainties in the suppression of jets originating from PU via the PUJetID are labelled accordingly in Figure 4.20 and reach up to 4% in the high-y^* and high-y_b bins at low p_T^Z.

Uncertainties in the integrated luminosity are propagated by varying the simulated event yields for each data-taking period according to the uncertainty breakdown in Subsection 4.2.1. This accounts for the different luminosity uncertainties and detector conditions in the indi-

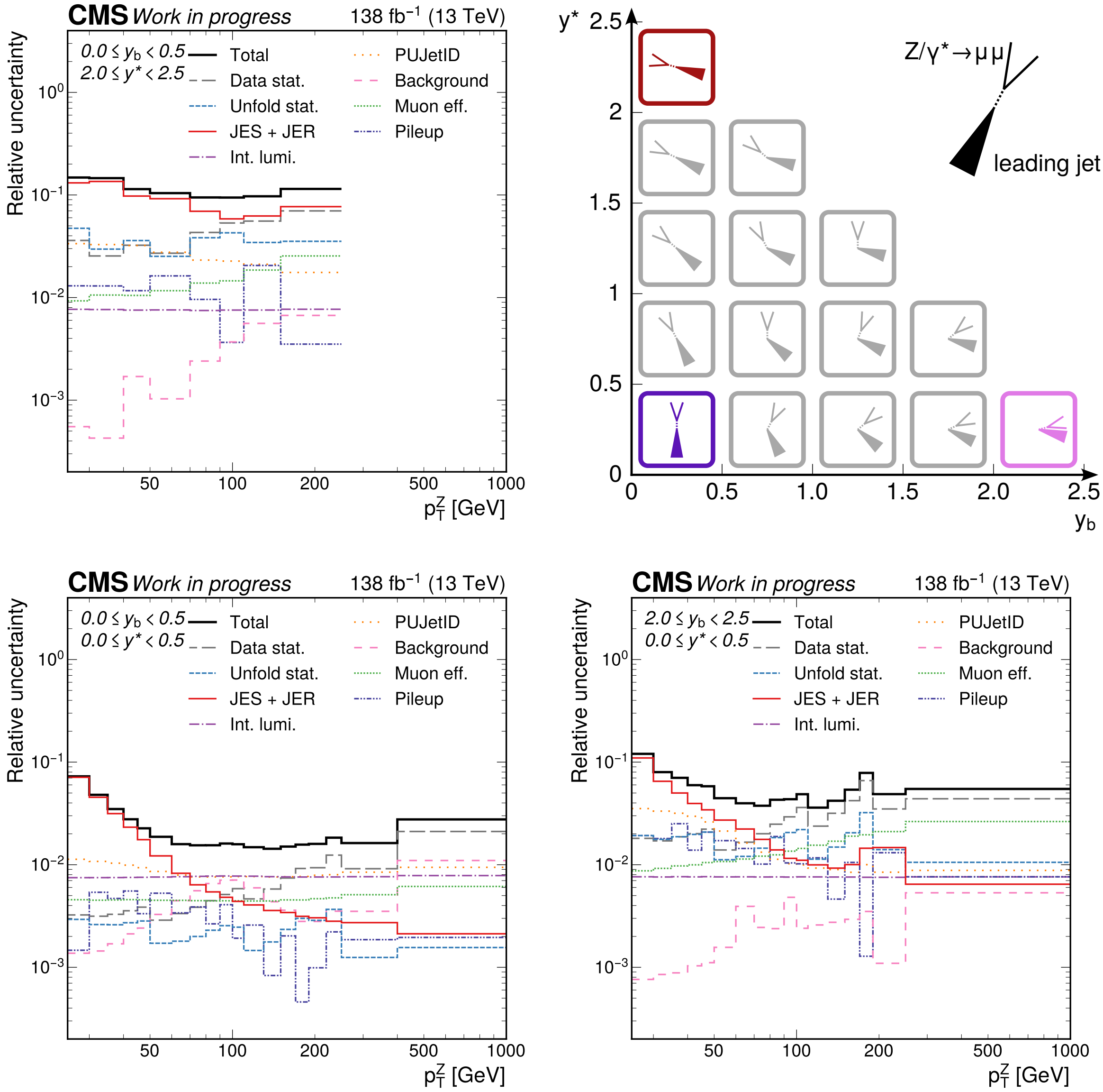


Figure 4.20: Uncertainty breakdown for three exemplary (y_b, y^*)-bins as a function of p_T^Z. The shown bins in (y_b, y^*) are indicated by the legend in the top right. The labels and contributions from each uncertainty source are explained in the text.

vidual data-taking periods. After propagation through the full analysis chain, the resulting uncertainty is flat across the measurement phase space and amounts to 0.73% in each bin. This contribution is labelled as “Int. lumi” in Figure 4.20.

Variations of the inelastic proton–proton cross section by ±4.6% for the pileup reweighting, as described in Subsection 4.2.2, are labelled as “Pileup” in Figure 4.20. They result in an uncertainty below 1% in the central rapidity region and reach up to 2% in the outer rapidity regions.

Uncertainties in the scale factors that correct the muon trigger, reconstruction, and identification efficiencies in simulation to match those observed in data are labelled as “Muon eff.” in Figure 4.20. This also includes the correction factors for the prefiring of the muon triggers, as described in Subsection 3.2.5. The combined uncertainty is below 0.5% in most bins and rises up to 2% in the outermost y^* and $p_{\mathrm{T}}^{\mathrm{Z}}$ bin.

Finally, uncertainties in the background estimation are taken into account, including the statistical uncertainties from simulation and their correlations after unfolding, as described in Subsection 4.4.1, as well as systematic uncertainties from the MC simulation. A detailed explanation of these uncertainties is given in Subsection 4.5.2. The total uncertainty in the background estimation is labelled as “Background” in Figure 4.20. This uncertainty only plays an important role in the central rapidity region around a $p_{\mathrm{T}}^{\mathrm{Z}}$ of 100 GeV, where it is slightly smaller than the uncertainty in the integrated luminosity.

The best precision is achieved in the central rapidity regions around a $p_{\mathrm{T}}^{\mathrm{Z}}$ of 70 GeV to 200 GeV, where the total uncertainty reaches values as low as 1.5%. In this region, several uncertainty sources contribute comparably to the total uncertainty. At low transverse momenta, the uncertainty in the jet energy scale corrections is the dominant source, while at large transverse momenta the measurement is limited by the size of the data sample.

In summary, a high-precision triple-differential measurement of the Z + jet cross section is performed as a function of y_{b}, y^*, and $p_{\mathrm{T}}^{\mathrm{Z}}$. The total uncertainty ranges from 1.5% in the central rapidity region for transverse momenta around 100 GeV to 10% or more in the outer rapidity regions and at the edges of the analysed $p_{\mathrm{T}}^{\mathrm{Z}}$ range. The unfolded data are compared to predictions obtained using MC generators at NLO accuracy in perturbative QCD and fixed-order calculations at NNLO accuracy in perturbative QCD in the following chapter.

CHAPTER 5

Measured Z+jet cross section and comparison to theory predictions

After the definition of the measurement strategy and the discussion of the relevant theory predictions, this chapter presents the unfolded triple-differential Z + jet cross section and its comparison to theoretical calculations. The measured cross section is based on 138 fb^{-1} of proton–proton collision data collected by the CMS experiment during 2016–2018, as described in Chapter 4. It is unfolded to stable-particle level correcting detector effects. The theory calculations used for the comparisons are introduced in Chapter 2.

The triple-differential measurement provides a detailed test of theory predictions for Z + jet production across different regions of the partonic phase space. First, the measured cross section is compared to Monte Carlo (MC) predictions obtained at leading order (LO) and next-to-leading order (NLO) in perturbative quantum chromodynamics (QCD) interfaced to PYTHIA 8 or HERWIG 7 for showering, hadronization, and underlying-event simulation. These comparisons are presented in Section 5.1.

The measured cross section is then compared to fixed-order theory predictions at next-to-next-to-leading order (NNLO) in perturbative QCD. These predictions are defined at parton level, as no matching formalism for Z + jet production at NNLO to parton showers currently exists. Therefore, correction factors for nonperturbative (NP) effects are derived in Section 2.6 and applied to the fixed-order calculations. In addition, NLO corrections from electroweak (EW) theory are calculated in Section 2.5 and applied to these predictions. The corrected predictions are then compared to data in Section 5.2. The potential sensitivity of these data to future parton distribution function (PDF) determinations is estimated from these comparisons using multiple different PDF sets.

5.1 Comparison to Monte Carlo simulation

The unfolded data are compared to predictions obtained from multiple MC generators in Figure 5.1 for the central-rapidity, high-y_b, and high-y^* bins. Figure A.14 in the appendix contains all 15 (y_b, y^*)-bins. The data and predictions from the different generators are shown as ratios to the stable-particle-level prediction obtained with the MADGRAPH5_aMC@NLO FxFx

+ PYTHIA 8 simulation that is also used for the unfolding procedure described in Section 4.4.

Here, the error bars on the data points correspond to the statistical uncertainty arising from the limited number of measured events, while the full uncertainty band corresponds to the total uncertainty in the measurement described in Section 4.5.

The nominal prediction is of NLO accuracy in perturbative QCD for the production of a Z boson in association with up to two jets. The three different jet multiplicities (Z + 0, 1, and 2 jets) are simulated with MadGraph5_aMC@NLO using the PDF4LHC15 PDF set and merged using the FxFx merging scheme. The simulated events are then matched to the parton shower using the MC@NLO formalism and showered with PYTHIA 8 using the CP5 tune.

A second prediction using MiNNLO$_{\mathrm{PS}}$ at NNLO accuracy for inclusive Drell–Yan production is shown. It is only of NLO accuracy for observables involving at least one jet. The PDF4LHC21 PDF set [50] is used, which combines the CT18 [9], MSHT20 [166], and NNPDF4.0 [11] PDFs. The simulated events are showered with PYTHIA 8 using the CP5 tune, while PHOTOS is used for quantum electrodynamics (QED) emissions. This prediction is shown as the blue line in Figure 5.1.

The uncertainty band around these predictions is composed of multiple sources. The renormalization and factorization scales μ_{R} and μ_{F} are varied around their nominal scale μ_0 by factors of 0.5 and 2, excluding the extreme combinations $(\mu_{\mathrm{R}}/\mu_0, \mu_{\mathrm{F}}/\mu_0) = (0.5, 2.0)$ and $(2.0, 0.5)$. Uncertainties in the PDF set and variations of the initial-state radiation (ISR) and final-state radiation (FSR) scales in the parton shower by factors of 0.5 and 2 are taken into account. Additionally, the statistical uncertainty arising from the limited number of simulated events is considered.

In addition, predictions at LO accuracy in perturbative QCD are shown without uncertainties, as the LO scale variations are very large compared to the uncertainties of the NLO predictions, as illustrated in Figure 2.7. These predictions are generated with MadGraph5 at LO accuracy for the production of a Z boson in association with up to four additional jets, merged using the MLM formalism, and utilizing the PDF4LHC15 PDF set. One sample is showered with PYTHIA 8 using the CP5 tune and is shown in orange in Figure 5.1, while the other sample is showered with HERWIG 7 using the CH3 tune and is shown in red.

The samples at LO accuracy fail to describe the unfolded data, especially at high y^*, where deviations from the data exceed 50%. In this phase-space region, large NLO QCD K-factors are present, as discussed in Section 2.4. At low y^*, the simulation showered with PYTHIA 8 describes the unfolded data well at low transverse momenta, with deviations of up to 30% at high $p_{\mathrm{T}}^{\mathrm{Z}}$. The simulation showered with HERWIG 7 deviates from the unfolded data by up to a factor of 2 at large $p_{\mathrm{T}}^{\mathrm{Z}}$.

Both simulations at NLO accuracy in perturbative QCD for Z + jet production describe the data within their respective uncertainties across the full phase space in y_{b} and y^*. The central values of the predictions obtained with POWHEG MiNNLO$_{\mathrm{PS}}$ show better agreement with the data than those obtained with MadGraph5_aMC@NLO FxFx only in the high-y^* region. Here differences between the two MC samples of up to 30% in the central values are observed. In the high-y^* region the uncertainty in the simulation is approximately 15% and is of similar size to the uncertainty in the data.

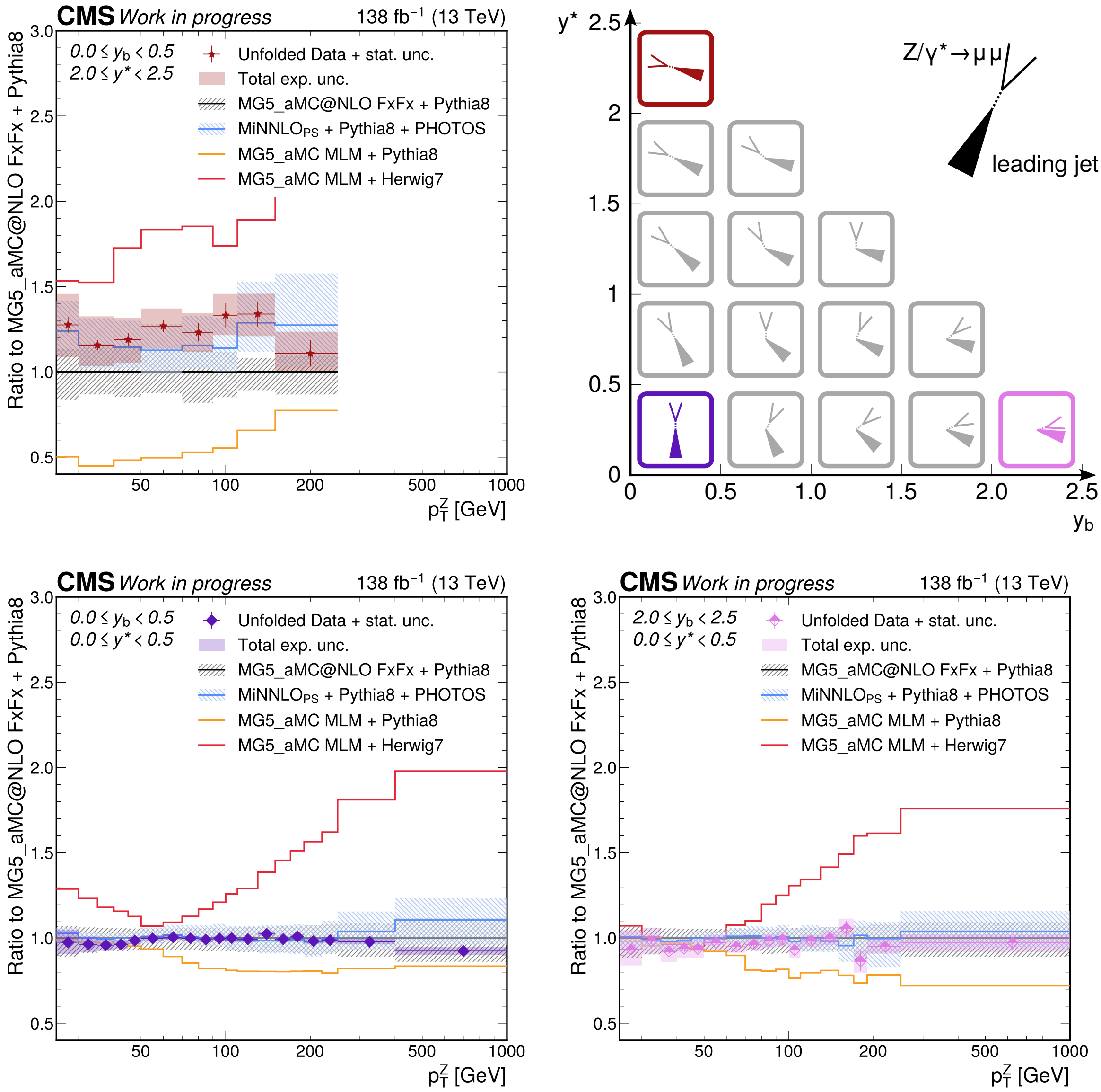


Figure 5.1: Comparison of cross section predictions from different MC generators to the unfolded data as a function of $p_{\mathrm{T}}^{\mathrm{Z}}$ in the high-y^{*} bin (top left), at central rapidity (bottom left), and in the high-y_{b} bin (bottom right). An illustration of the three selected (y_{b}, y^{*})-bins with respect to the full analysed phase space is shown in the top right. The data and MC predictions are shown as ratios to the prediction obtained with MadGraph5_amc@nlo FxFx + pythia 8. Both the MadGraph5_amc@nlo FxFx + pythia 8 and powheg MiNNLO$_{\mathrm{PS}}$ + pythia 8 predictions are shown with uncertainties including scale variations, initial- and final-state radiation, PDF uncertainties, and statistical uncertainties. The MadGraph5_amc@nlo MLM-merged samples interfaced with pythia 8 (orange) and herwig 7 (red) are shown without uncertainties, as they are simulated only at LO accuracy and therefore exhibit large scale uncertainties.

5.2 Comparison to NNLO fixed order calculations

Predictions for the triple-differential Z + jet cross section at NNLO accuracy in perturbative QCD are calculated using NNLOJet [47], interfaced via APPLfast [48] to FastNLO [49]. FastNLO uses interpolation grids for fast reevaluations of the NNLO cross section when varying the PDFs, α_S, or the factorization and renormalization scales μ_F and μ_R.

The fixed-order calculations at parton level are supplemented with correction factors for NLO EW contributions and NP effects, as described in Section 2.5 and 2.6. In Figure 5.2, the measured cross section is compared to theory predictions obtained by evaluating the interpolation grids with the PDF4LHC21 PDF set via the LHAPDF interface [167] for all (y_b, y^*) bins as a function of p_T^Z. Figure 5.3 shows the ratios of the unfolded data to the same NNLO QCD theory predictions. The error bars on the data points correspond to the statistical uncertainty due to the limited number of measured events, while the band around the data points corresponds to the total experimental uncertainty described in Section 4.5. In Figure 5.2, the total experimental uncertainty is included in the error bars on the data points. The hatched uncertainty band around the theory prediction in Figure 5.3 corresponds to the quadratic combination of the scale uncertainties, the uncertainties in the PDF4LHC21 PDF set, and the uncertainties associated with the NP and EW corrections as described in Section 2.7.

At low transverse momenta, $p_T^Z < 50\,\text{GeV}$, the data and theory predictions agree within uncertainties. This agreement is partly driven by the large uncertainties in both the measured cross section and the theory predictions in this region. Towards higher p_T^Z and y^*, the agreement deteriorates, with deviations exceeding the quoted uncertainties.

The ratios are shown in more detail in Figure 5.4 for the central-rapidity, high-y_b, and high-y^* bins. In addition, predictions obtained with the five-flavour ABMPtt NNLO [168], NNPDF4.0 NNLO [11], and approximate next-to-NNLO (aN^{3}LO) MSHT20 [169] PDF sets using the LHAPDF interface are shown. A complete breakdown of all 15 (y_b, y^*) bins is provided in the appendix in Figure A.15.

Here, the sensitivity of the differential variables to variations of the PDF set is clearly visible, with differences of up to 7% depending on the chosen PDF set. In the central rapidity region, where the highest experimental precision is achieved, differences between the measured data and the predictions reach up to 10% for $p_T^Z > 100\,\text{GeV}$. In the high-y^* and high-y_b bins, the data and predictions agree within the larger uncertainties.

The predicted decrease in the cross section towards larger y^* is confirmed by the data. The production of Z + jet events at high y^* requires both incoming partons to carry large momentum fractions $x_{1,2}$ of the proton, as given in Equation 2.12. For this kinematic configuration, the PDFs favour an initial state consisting of two quarks, but this initial state only becomes available at NLO accuracy in perturbative QCD. This leads to NLO K-factors of up to eleven in the high-y^* bin, as shown in Figure 2.7. As a result, the NLO QCD prediction is effectively only of leading-order quality in this region and exhibits large scale uncertainties. The NNLO prediction provides a substantially improved description of the data in this region and reduces the scale uncertainty to a level comparable to the experimental uncertainty.

Overall, good agreement between the data and theory predictions is observed at transverse momenta below 100 GeV or towards high y_b in the full p_T^Z range. For $p_T^Z > 100\,\text{GeV}$ in

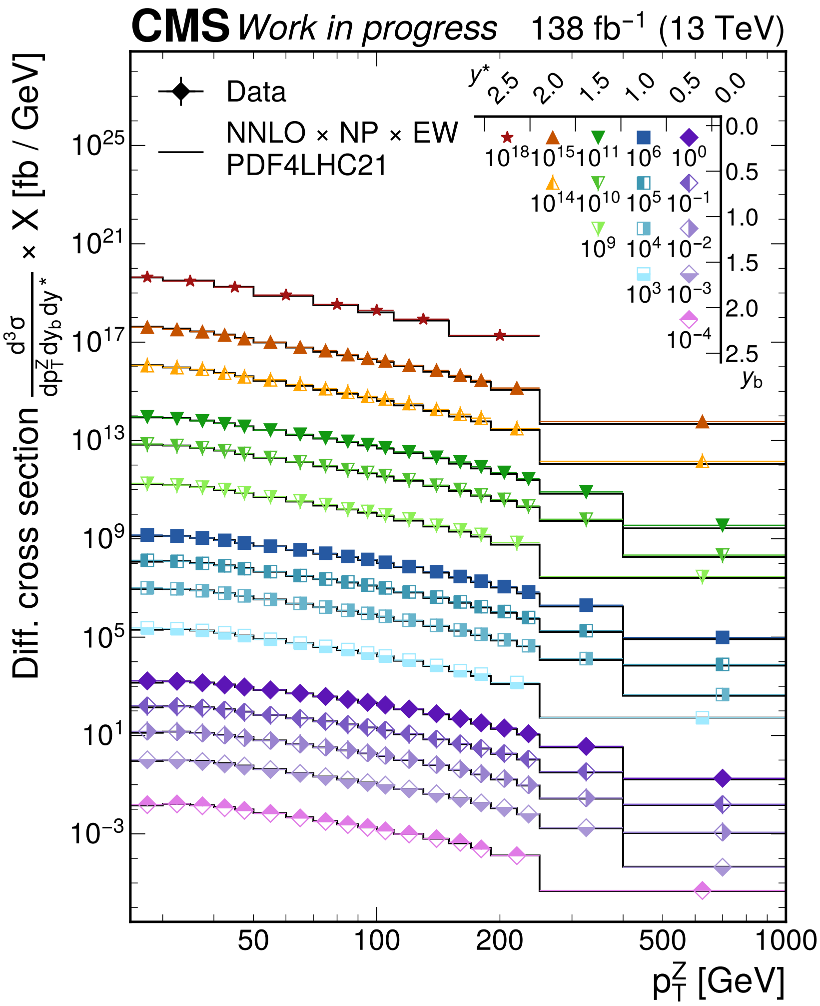


Figure 5.2: Unfolded triple-differential Z + jet cross section as a function of p_T^Z in bins of y_b and y^* compared to NNLO QCD predictions obtained with APPLFAST interfaced to NNLOJET using the PDF4LHC21 PDF set. The theory predictions are corrected for EW and NP effects. For visual clarity, the cross sections in different (y_b, y^*)-bins are scaled by multiplicative factors as indicated in the legend.

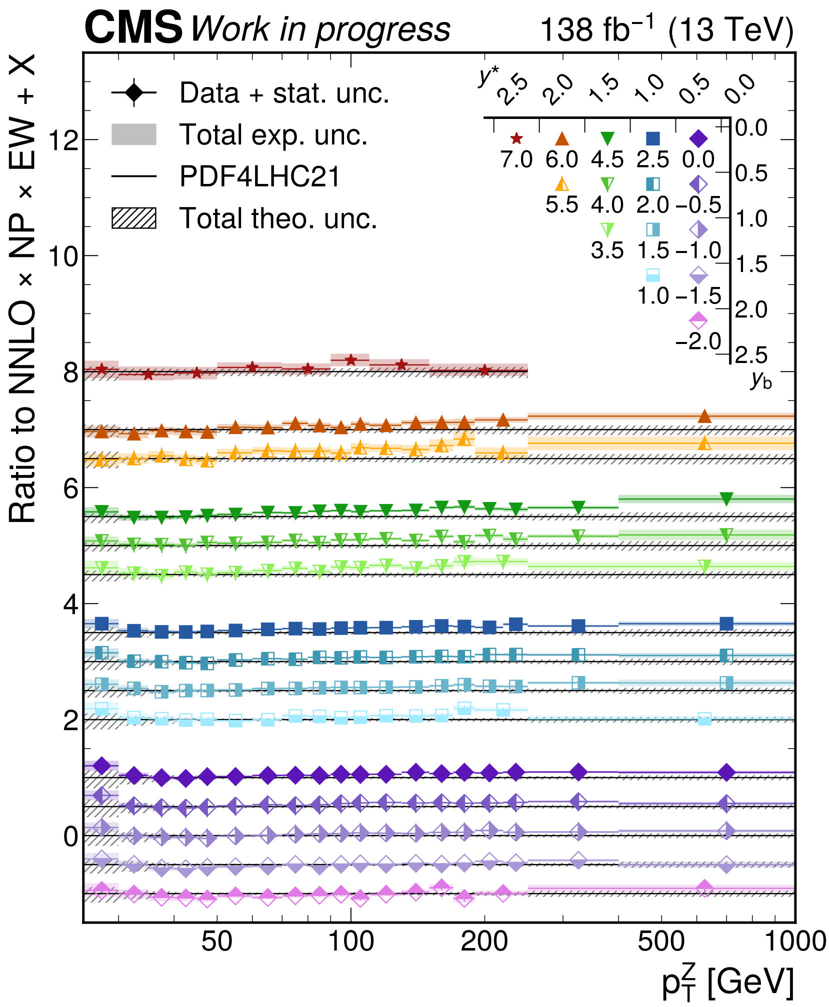


Figure 5.3: Ratio of the unfolded triple-differential Z + jet cross section to the NNLO QCD theory predictions obtained with APPLFAST interfaced to NNLOJET using the PDF4LHC21 PDF set as a function of p_T^Z in bins of y_b and y^*. The theory predictions are corrected for EW and NP effects. For visual clarity, the ratios are vertically offset by additive constants, as indicated in the legend.

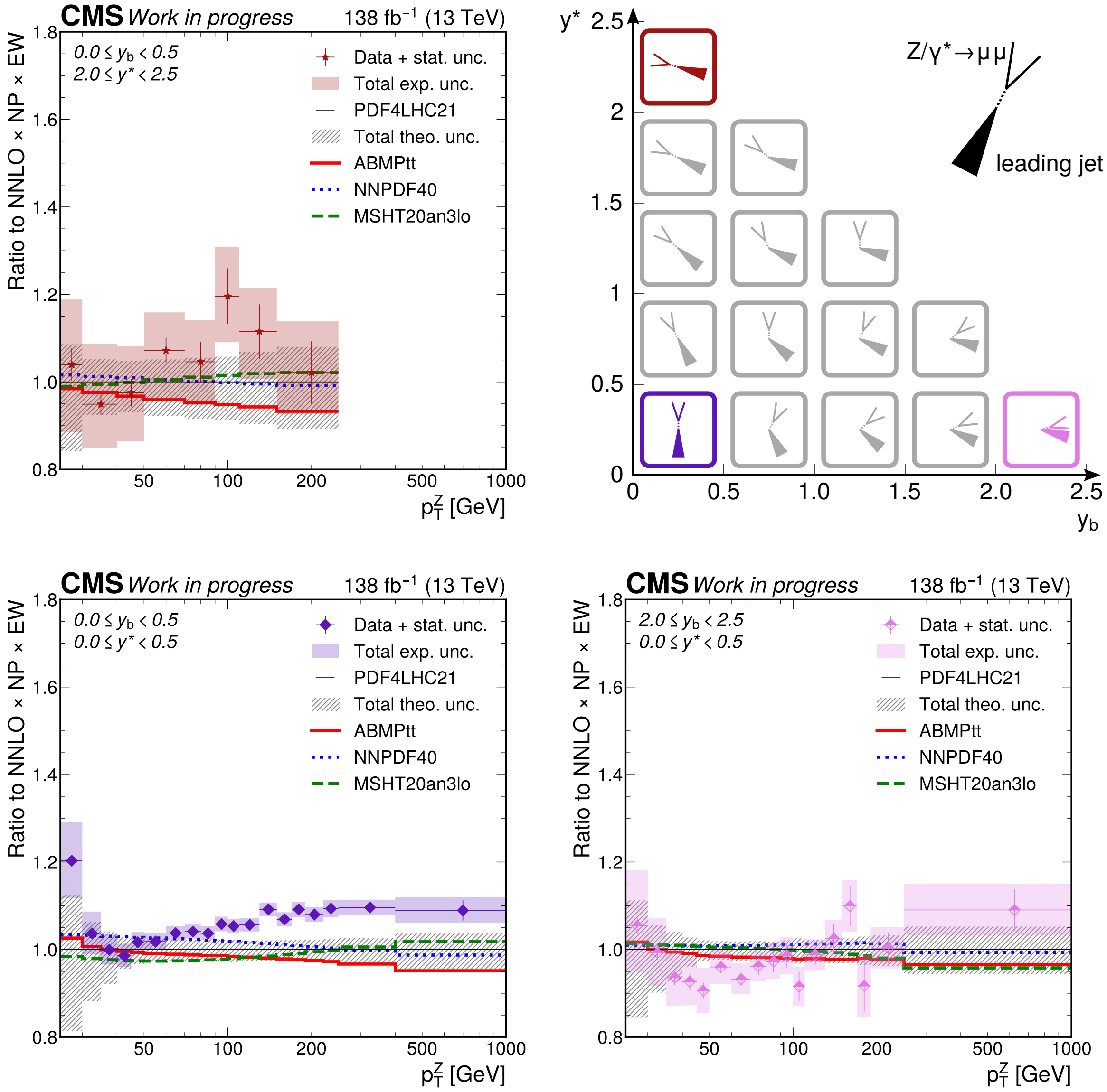


Figure 5.4: Unfolded cross section as a function of p_T^Z compared to the NNLO QCD predictions corrected for EW and NP effects in the central rapidity bin (bottom left), high-y_b bin (bottom right) and high-y^* bin (top left). An illustration of the three selected (y_b, y^*)-bins with respect to the full analysed phase space is shown in the top-right panel. The central theory prediction uses the PDF4LHC21 PDF set, alternative predictions using the ABMPtt (solid red), NNPDF 4.0 (dotted blue) and approximate N3LO MSHT20 (dashed green) PDF sets are shown.

the central rapidity region and towards higher y^*, differences between the predicted and measured cross sections are clearly visible. In the highest y^* bin, however, no significant differences between data and theory predictions are observed within the larger experimental and theoretical uncertainties. The observed differences indicate that the measured cross section is sensitive to further improvements in the theory predictions, including more precise perturbative calculations, resummation effects, and improved determinations of the PDFs. Conversely, the precision and differential structure of these data make them valuable input for future PDF determinations.

CHAPTER 6

Summary

The cross section for production of a Z boson in association with at least one jet has been measured triple-differentially as a function of the average boost in rapidity of the Z boson and the leading jet, y_b, half the absolute difference in rapidity between the Z boson and the leading jet, y^*, and the transverse momentum of the Z boson, p_T^Z. The measurement uses proton–proton collision data collected with the CMS detector in the years 2016–2018, corresponding to an integrated luminosity of $138\,\mathrm{fb}^{-1}$. The chosen observables provide sensitivity to complementary aspects of the hard scattering: p_T^Z probes the momentum scale of the interaction, while y^* and y_b are related to the scattering topology and the momentum imbalance of the colliding partons.

The measured distributions are unfolded to stable-particle level to correct for detector effects. The total experimental uncertainty ranges from about 1.5% in the central rapidity region to 10% or more in regions of large y_b or large y^*. Across all bins in (y_b, y^*), the uncertainty in the jet energy scale (JES) dominates at low transverse momenta, while the statistical precision of the data is the largest contribution to the total uncertainty at large p_T^Z.

The unfolded data are compared to predictions from Monte Carlo (MC) event generators at next-to-leading order (NLO) accuracy in perturbative quantum chromodynamics (QCD). Good agreement is observed within the experimental uncertainties and the uncertainties in the theory predictions. For these predictions, the uncertainties from hard-scattering scale variations and parton shower (PS) modelling are substantially larger than the experimental uncertainties in most regions of phase space.

In addition, the data are compared to fixed-order calculations at next-to-next-to-leading order (NNLO) accuracy in perturbative QCD, corrected for NLO electroweak (EW) contributions and nonperturbative (NP) effects. The uncertainties in these predictions range from about 3% around $p_T^Z \approx 100\,\mathrm{GeV}$ up to 18% at low p_T^Z, where they are dominated by the scale dependence. At high transverse momenta, the uncertainty in the EW corrections becomes dominant, while at large y^*, the uncertainties in the estimation of the NP effects play an important role. For transverse momenta around 50 GeV to 200 GeV, the uncertainties in the parton distribution functions (PDFs) also provide a sizeable contribution.

The theory predictions show a significant dependence on the chosen PDF set, with differences of up to 7% between PDF4LHC21, the five-flavour ABMPtt NNLO, NNPDF4.0 NNLO, and MSHT20 aN^{3}LO PDF sets. At the same time, differences of up to 10% or more are observed between the measured and predicted cross sections in some regions of phase space, which are not covered by the quoted uncertainties. In the large-y^* and large-y_b regions, where the experimental uncertainties are largest, the remaining differences between the central values are covered by the uncertainties in the data and theory predictions.

These results demonstrate that the measured cross sections can provide valuable constraints in future determinations of PDFs. The triple-differential binning separates regions with different sensitivity to the hard-scattering scale, the scattering topology, and the momentum fractions of the colliding partons. As a result, the measurement provides information that is complementary to more inclusive Z + jet measurements and, e.g., triple-differential dijet measurements in a similar phase space [170, 171]. A combined interpretation of triple-differential Z + jet and dijet cross section measurements is particularly interesting, since the dominant initial-state parton contributions differ between the two processes. A future combined fit of both measurements can therefore help disentangle the quark and gluon contributions to the proton structure more effectively than either measurement alone.

Overall, the measurement provides a precise test of perturbative QCD in a phase space that is directly sensitive to the partonic structure of the proton. Future theory predictions with reduced scale uncertainties, improved treatments of EW and NP effects, the inclusion of higher-order QCD or mixed QCD–EW corrections, resummation effects, and improved PDF sets may reduce the observed differences between the measured and predicted cross sections.

With the increase in recorded proton–proton collision data at the High-Luminosity LHC and improvements in experimental techniques, a precise understanding of Standard Model (SM) processes becomes ever more important for reducing theory uncertainties in precision measurements and searches for phenomena beyond the SM. This thesis contributes directly to this effort by providing valuable data for future PDF determinations and tests of theory predictions.

Preliminary results of the measurement presented in this thesis have been made available by the CMS Collaboration in Ref. [31] and were presented at the 2026 QCD session of the 60th Rencontres de Moriond [172]. The work on the nonperturbative correction factors is part of a paper published in Ref. [32].

APPENDIX A

Supplementary material

A.1 Datasets and simulation samples

In this section, the datasets and simulated samples used are listed. The data and simulation samples are each split into four data-taking periods: 2016 pre-VFP, 2016 post-VFP, 2017, and 2018. Table A.1 shows the datasets used in this thesis, split by data-taking period. The global tag used for the offline data reconstruction is `106X_dataRun2_v37`.

In Table A.2, the global tags for the Monte Carlo (MC) production are given. Table A.3 shows the simulated signal samples that were used in this thesis, Table A.4 shows the diboson background simulation samples and Table A.5 shows the background simulation of single top quark and top quark pair production. All tables also contain the number of simulated events for the combined full Run 2 data-taking period using the MiniAODv2 data format.

Table A.1: Datasets recorded by the CMS detector and used in this thesis. Horizontal lines separate the four data-taking periods: 2016 pre-VFP, 2016 post-VFP, 2017, and 2018. The right column shows the number of recorded events.

Dataset	# events
/SingleMuon/Run2016B-ver1_HIPM_UL2016_MiniAODv2_BParking-v1/MINIAOD	2 789 243
/SingleMuon/Run2016B-ver2_HIPM_UL2016_MiniAODv2_BParking-v1/MINIAOD	158 145 722
/SingleMuon/Run2016C-HIPM_UL2016_MiniAODv2_BParking-v1/MINIAOD	67 441 308
/SingleMuon/Run2016D-HIPM_UL2016_MiniAODv2_BParking-v1/MINIAOD	97 318 625
/SingleMuon/Run2016E-HIPM_UL2016_MiniAODv2_BParking-v1/MINIAOD	90 984 718
/SingleMuon/Run2016F-HIPM_UL2016_MiniAODv2_BParking-v1/MINIAOD	57 428 360
/SingleMuon/Run2016F-UL2016_MiniAODv2_BParking-v1/MINIAOD	8 024 195
/SingleMuon/Run2016G-UL2016_MiniAODv2_BParking-v1/MINIAOD	149 916 849
/SingleMuon/Run2016H-UL2016_MiniAODv2_BParking-v1/MINIAOD	174 035 164
/SingleMuon/Run2017B-UL2017_MiniAODv2-v1/MINIAOD	136 300 266
/SingleMuon/Run2017C-UL2017_MiniAODv2-v1/MINIAOD	165 652 756
/SingleMuon/Run2017D-UL2017_MiniAODv2-v1/MINIAOD	70 361 660
/SingleMuon/Run2017E-UL2017_MiniAODv2-v1/MINIAOD	154 618 774
/SingleMuon/Run2017F-UL2017_MiniAODv2-v1/MINIAOD	242 140 980
/SingleMuon/Run2018A-UL2018_MiniAODv2_BParking-v1/MINIAOD	241 181 837
/SingleMuon/Run2018B-UL2018_MiniAODv2_BParking-v1/MINIAOD	119 918 017
/SingleMuon/Run2018C-UL2018_MiniAODv2_BParking-v2/MINIAOD	109 921 698
/SingleMuon/Run2018D-UL2018_MiniAODv2_BParking-v1/MINIAOD	513 884 680
Total	2 561 023 995

Table A.2: MiniAOD campaign strings used for the simulated samples corresponding to each data-taking period.

Period	Campaign string
2016 preVFP	RunIISummer20UL16MiniAODAPVv2-106X_mcRun2_asymptotic_preVFP_v11
2016 postVFP	RunIISummer20UL16MiniAODv2-106X_mcRun2_asymptotic_v17
2017	RunIISummer20UL17MiniAODv2-106X_mc2017_realistic_v9
2018	RunIISummer20UL18MiniAODv2-106X_upgrade2018_realistic_v16_L1v1

Table A.3: Signal simulation samples. The right column shows the number of simulated events after combining the four different detector simulation campaigns. The POWHEG MiNNLO$_{PS}$ simulation sample is not available for the 2016 pre-VFP data-taking period.

Simulation sample	# events
DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8	557 761 821
DYJetsToLL_0J_TuneCP5_13TeV-amcatnloFXFX-pythia8	321 093 173
DYJetsToLL_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8	347 213 003
DYJetsToLL_2J_TuneCP5_13TeV-amcatnloFXFX-pythia8	176 598 366
DYJetsToLL_LHEFilterPtZ-0To50_MatchEWPDG20_TuneCP5_ 13TeV-amcatnloFXFX-pythia8	598 591 734
DYJetsToLL_LHEFilterPtZ-50To100_MatchEWPDG20_TuneCP5_ 13TeV-amcatnloFXFX-pythia8	366 795 450
DYJetsToLL_LHEFilterPtZ-100To250_MatchEWPDG20_TuneCP5_ 13TeV-amcatnloFXFX-pythia8	239 264 652
DYJetsToLL_LHEFilterPtZ-250To400_MatchEWPDG20_TuneCP5_ 13TeV-amcatnloFXFX-pythia8	72 882 453
DYJetsToLL_LHEFilterPtZ-400To650_MatchEWPDG20_TuneCP5_ 13TeV-amcatnloFXFX-pythia8	11 970 039
DYJetsToLL_LHEFilterPtZ-650ToInf_MatchEWPDG20_TuneCP5_ 13TeV-amcatnloFXFX-pythia8	12 047 863
DYJetsToMuMu_H2ErratumFix_PDFExt_TuneCP5_ 13TeV-powhegMiNNLO-pythia8-photos	740 184 575
DYJetsToTauTau_M-50_AtLeastOneEorMuDecay_H2ErratumFix_PDF_ TuneCP5_13TeV-powhegMiNNLO-pythia8-photos	1 019 459 113
DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8	377 197 381
DYJetsToLL_M-50_TuneCH3_13TeV-madgraphMLM-herwig7	118 569 423

Table A.4: Diboson background simulation samples. The right column shows the number of simulated events after combining the four different detector simulation campaigns.

Simulation sample	# events
WWTo1L1Nu2Q_4f_TuneCP5_13TeV-amcatnloFXFX-pythia8	120 251 433
WWTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	23 010 000
WZTo1L1Nu2Q_4f_TuneCP5_13TeV-amcatnloFXFX-pythia8	22 133 021
WZTo2Q2L_mllmin4p0_TuneCP5_13TeV-amcatnloFXFX-pythia8	86 874 762
WZTo3LNu_TuneCP5_13TeV-amcatnloFXFX-pythia8	40 248 649
ZZTo2Nu2Q_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8	14 830 716
ZZTo2L2Nu_TuneCP5_13TeV_powheg_pythia8	130 515 000
ZZTo2Q2L_mllmin4p0_TuneCP5_13TeV-amcatnloFXFX-pythia8	88 680 772
ZZTo4L_TuneCP5_13TeV_powheg_pythia8	290 201 000

Table A.5: Single-top-quark and top-quark-pair production background simulation samples. The right column shows the number of simulated events after combining the four different detector simulation campaigns.

Simulation sample	# events
ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	427 693 000
ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	227 387 000
ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	18 396 000
ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	18 277 000
TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	333 869 000

Table A.6: Simulation samples for the production of a W boson in association with up to two jets. The right column shows the number of simulated events after combining the four different detector simulation campaigns.

Simulation sample	# events
WJetsToLNu_TuneCP5_13TeV-amcatnloFXFX-pythia8	113 105 553
WJetsToLNu_0J_TuneCP5_13TeV-amcatnloFXFX-pythia8	653 828 448
WJetsToLNu_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8	705 763 669
WJetsToLNu_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8	364 713 316

A.2 Supplementary figures

In this section supplementary plots are presented. It is mostly a collection of plots across all 15 bins in (y_{b}, y^{*}), while in the main part of the thesis only the central. high-y_{b} and high-y^{*} bins were shown.

In Figure A.1 to A.6 extensions to the figures shown in Chapter 2 are shown. From Figure A.7 to A.10 the pileup (PU) reweighting is shown for each data-taking period, extending the full Run 2 Figure 4.1. This is followed by a detailed look into the uncertainties of this measurement in Figure A.11 to A.13. Finally, comparisons of MC samples and next-to-next-to-leading order (NNLO) theory predictions to the data are given in all (y_{b}, y^{*}) bins in Figure A.14 and A.15.

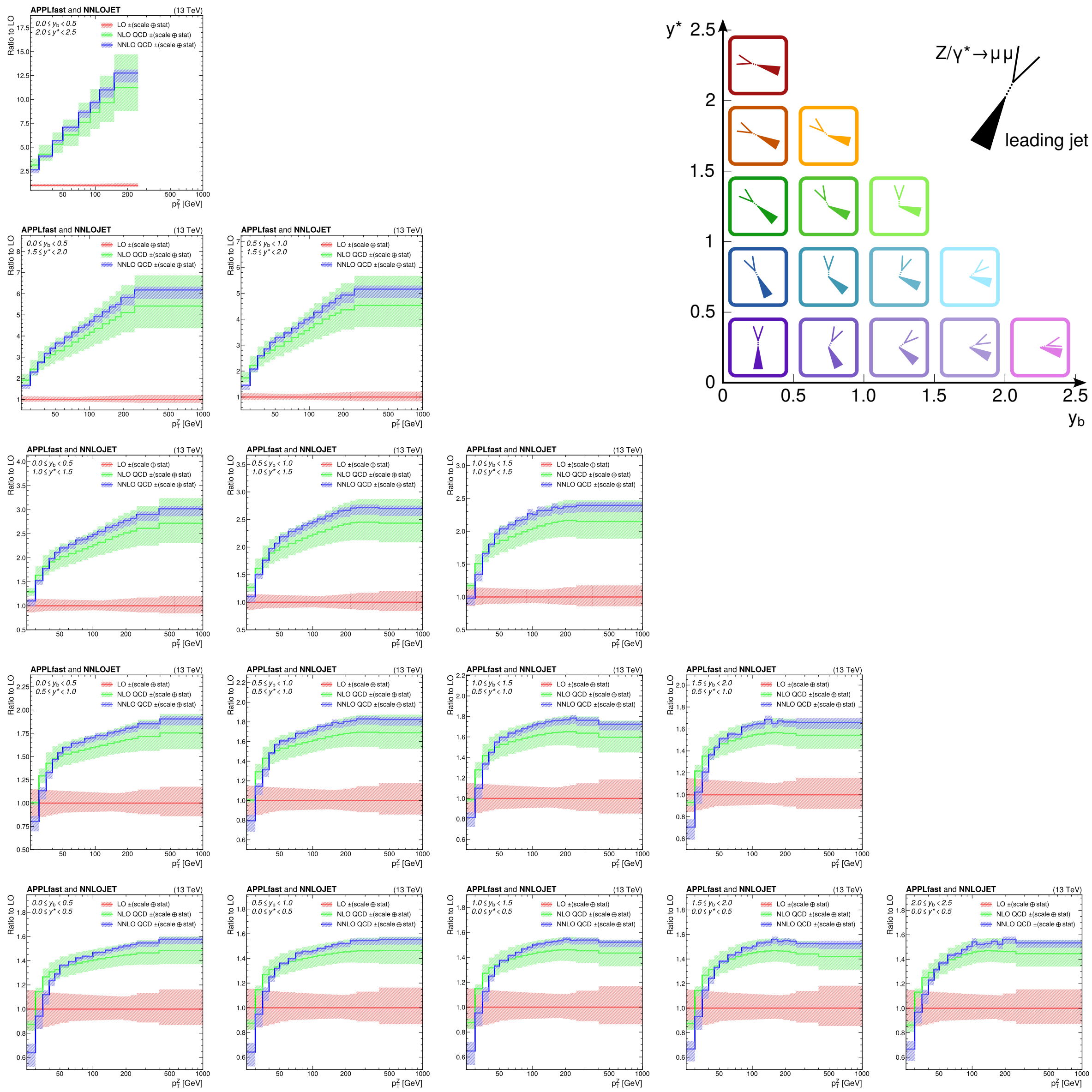


Figure A.1: Higher order corrections for Z + jet production in perturbative quantum chromodynamics (QCD) as a ratio to the leading order (LO) predictions for all 15 (y_b, y^*) bins as illustrated in the top right. The uncertainty band includes statistical uncertainties in the numerical integration and scale variations.

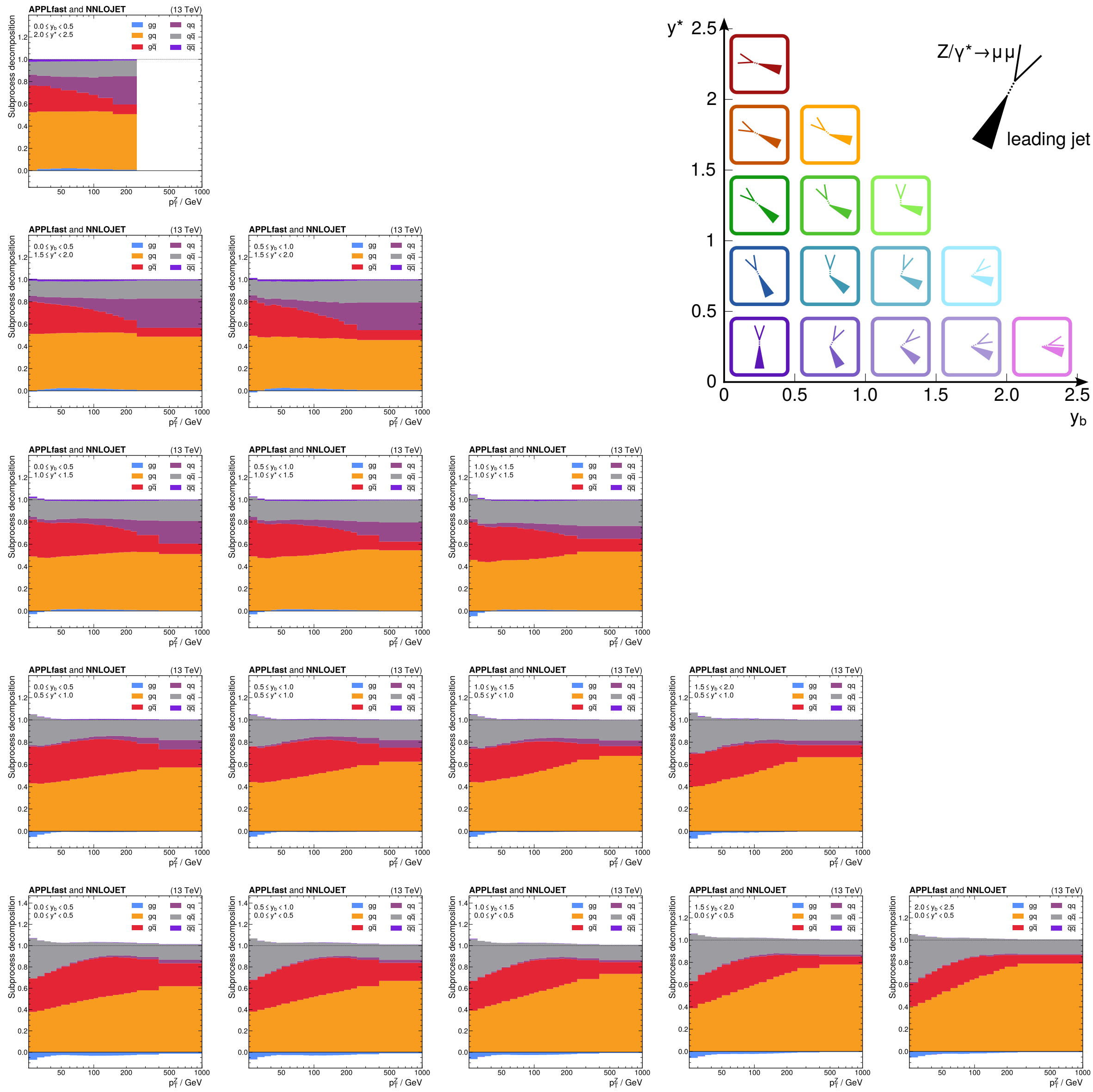


Figure A.2: Subprocess decomposition of the NNLO cross section calculated using the PDF4LHC21 parton distribution function (PDF) set in all 15 (y_b, y^*) bins. The corresponding breakdown of the (y_b, y^*)-bins is illustrated in the top right.

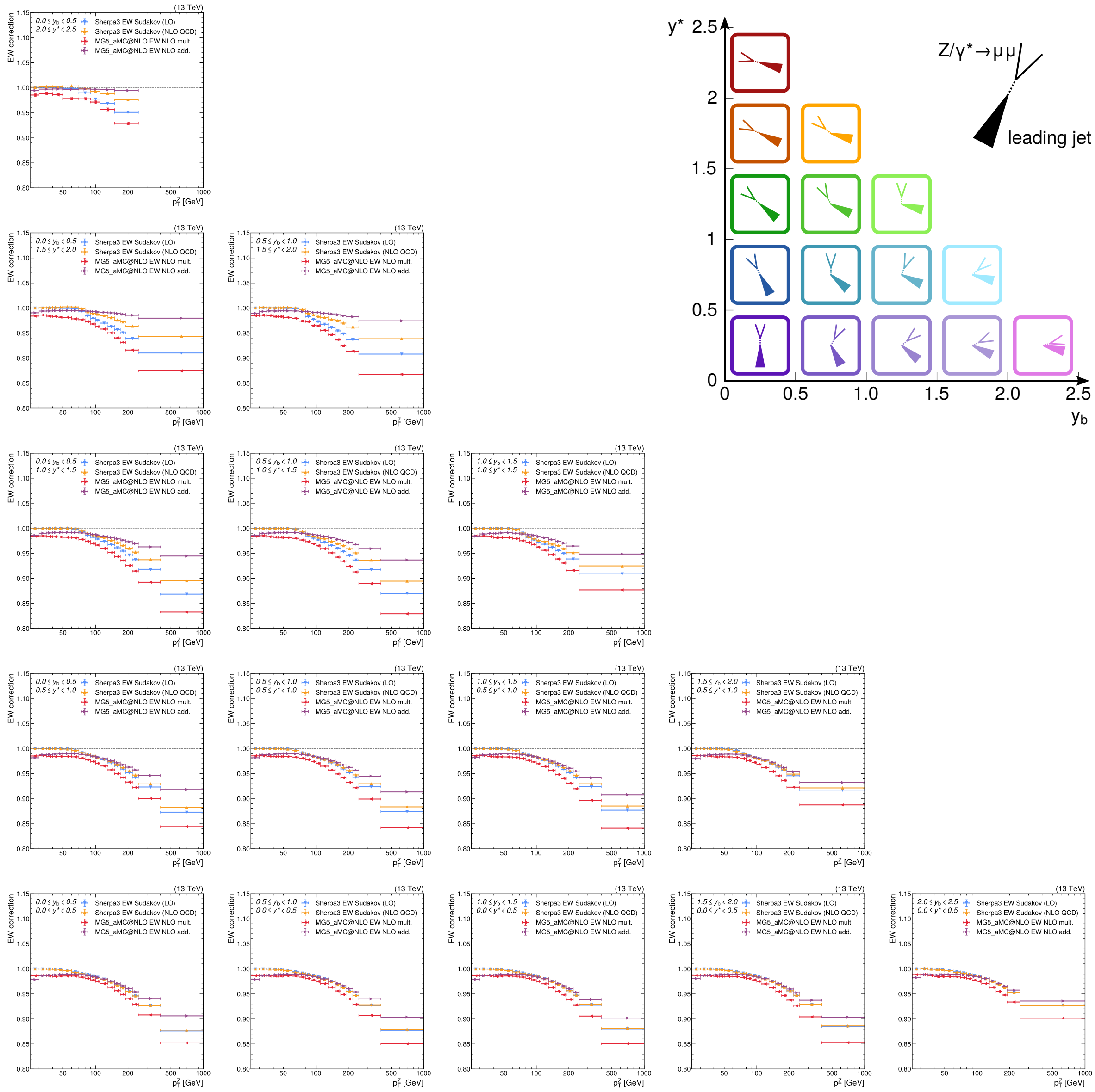


Figure A.3: Electroweak (EW) corrections as a function of p_T^Z in all 15 (y_b, y^*)-bins, as illustrated in the top right.

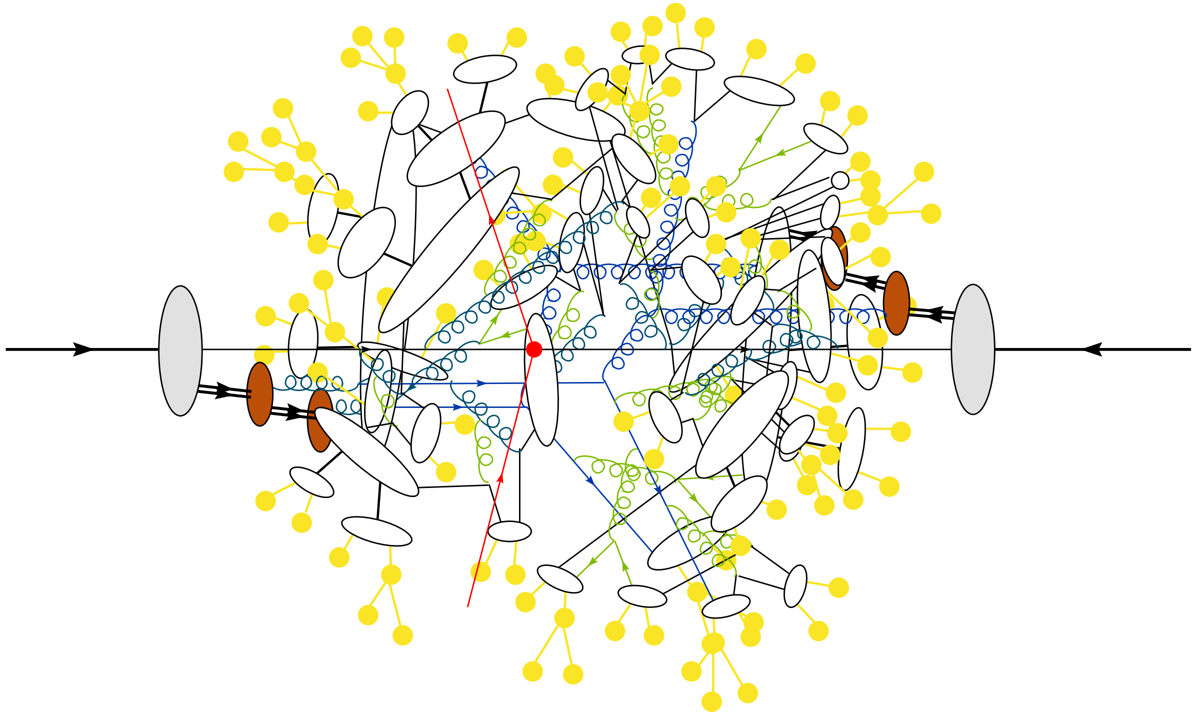


Figure A.4: Sketch of nonperturbative effects in a proton–proton collision. The two protons are depicted as grey blobs on the left and right, and the hard interaction of the incoming partons is shown in red, producing two high-p_T leptons. Initial-state radiation is shown as gluons radiated from the incoming partons, leading to parton showers. The subsequent hadronization is illustrated according to the cluster model used in HERWIG. The partons are combined into colourless clusters, shown as white ellipses, which decay into hadrons, shown as yellow circles, that can decay further. Additional contributions from the underlying event through the proton remnants and their interactions with the rest of the event simulation are included here. Sketch courtesy of S. Gieseke.

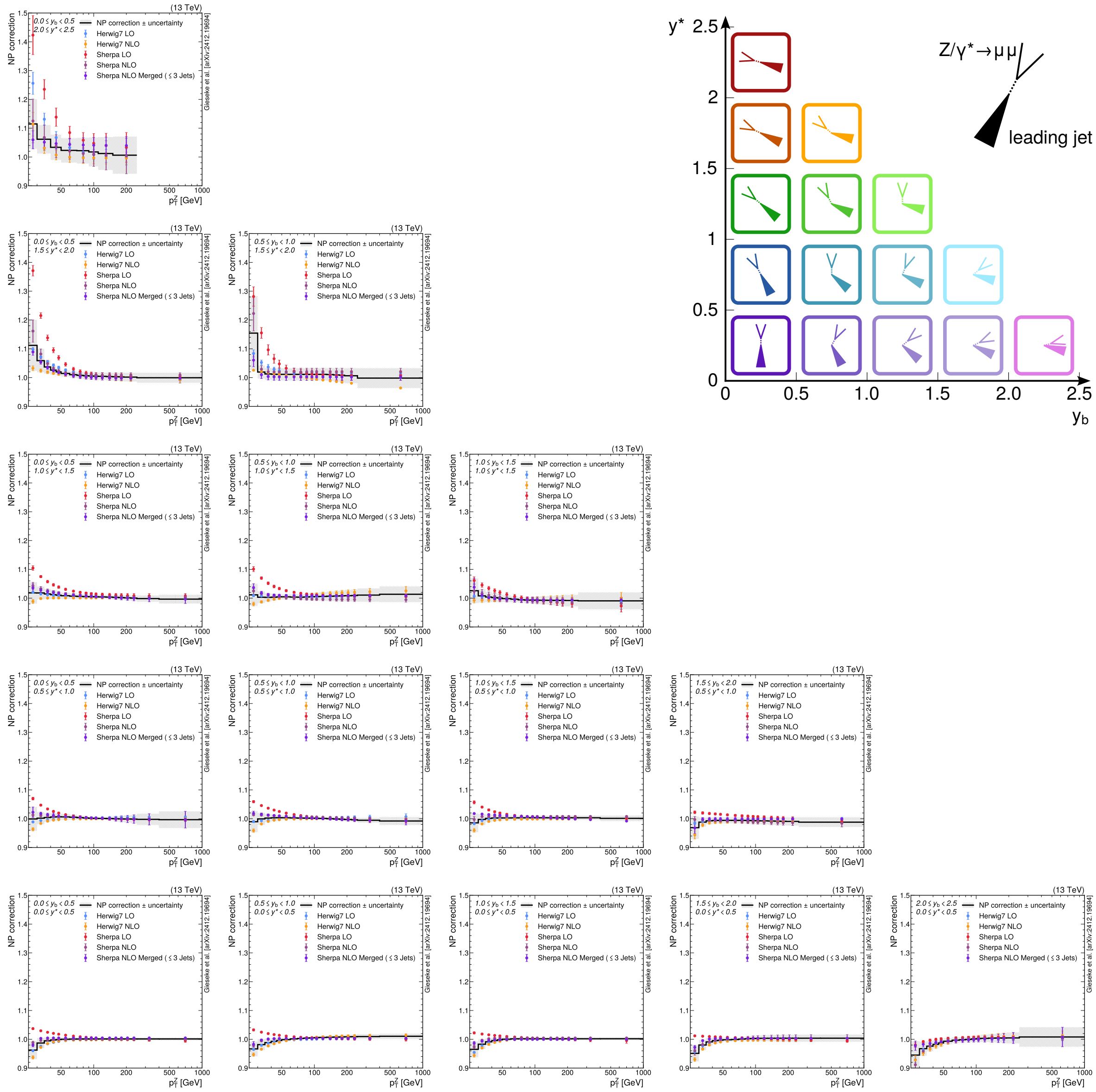


Figure A.5: Nonperturbative (NP) corrections as a function of p_T^Z in all 15 (y_b, y^*)-bins, as illustrated in the top right.

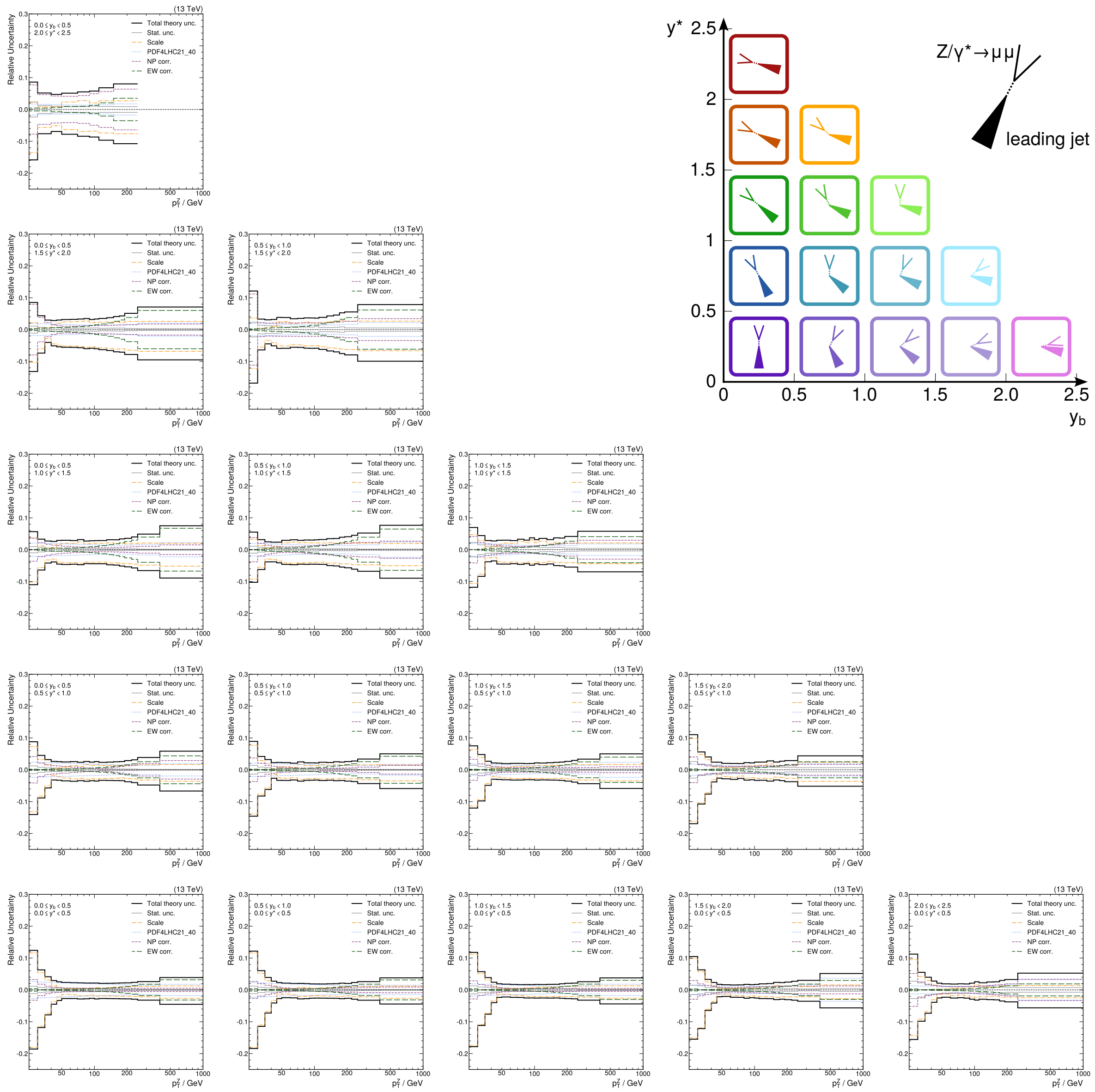


Figure A.6: Uncertainties in the final NNLO theory predictions, including statistical uncertainties in the integration, scale variations, uncertainties in the PDF4LHC21 PDF set, NP effects and next-to-leading order (NLO) EW corrections. The uncertainties are shown as a function of $p_{\mathrm{T}}^{\mathrm{Z}}$ in all 15 (y_{b}, y^{*})-bins as illustrated in the top right.

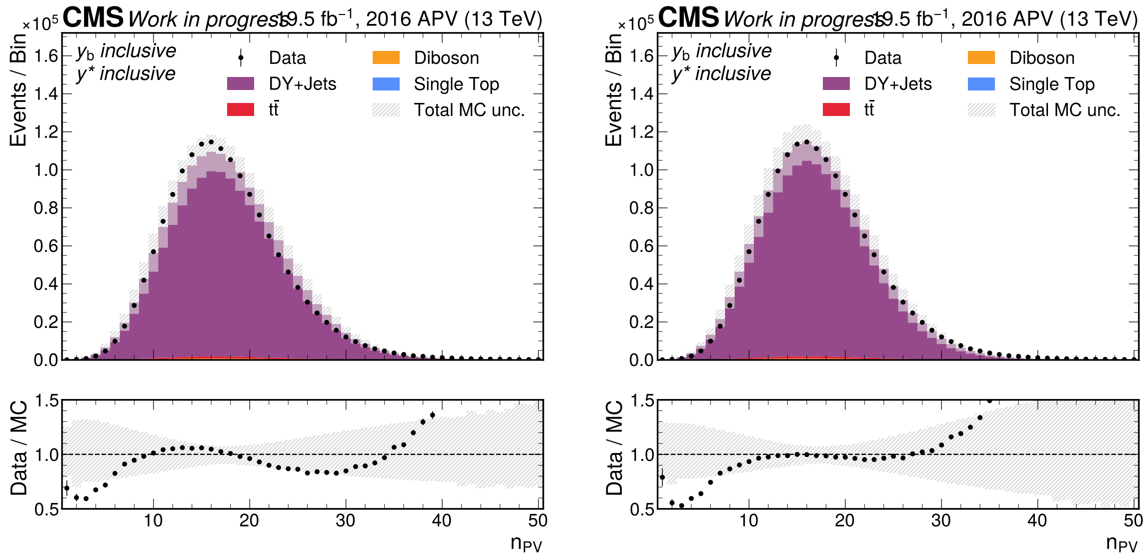


Figure A.7: Distribution of the number of reconstructed collision vertices in 2016 pre-VFP data compared to simulation before pileup reweighting (left) and after applying the pileup reweighting procedure (right).

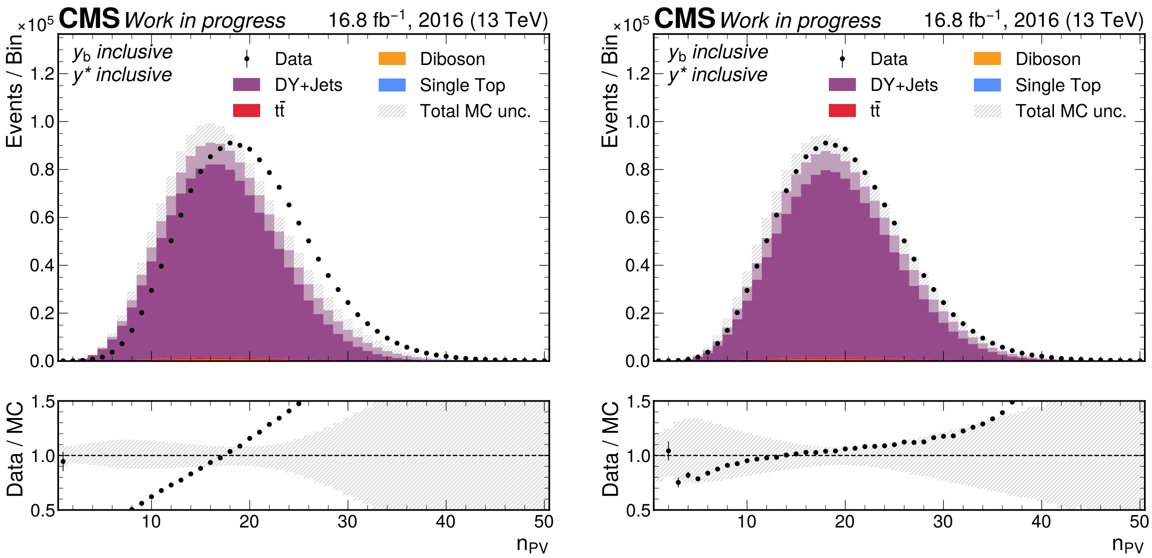


Figure A.8: Distribution of the number of reconstructed collision vertices in 2016 post-VFP data compared to simulation before pileup reweighting (left) and after applying the pileup reweighting procedure (right).

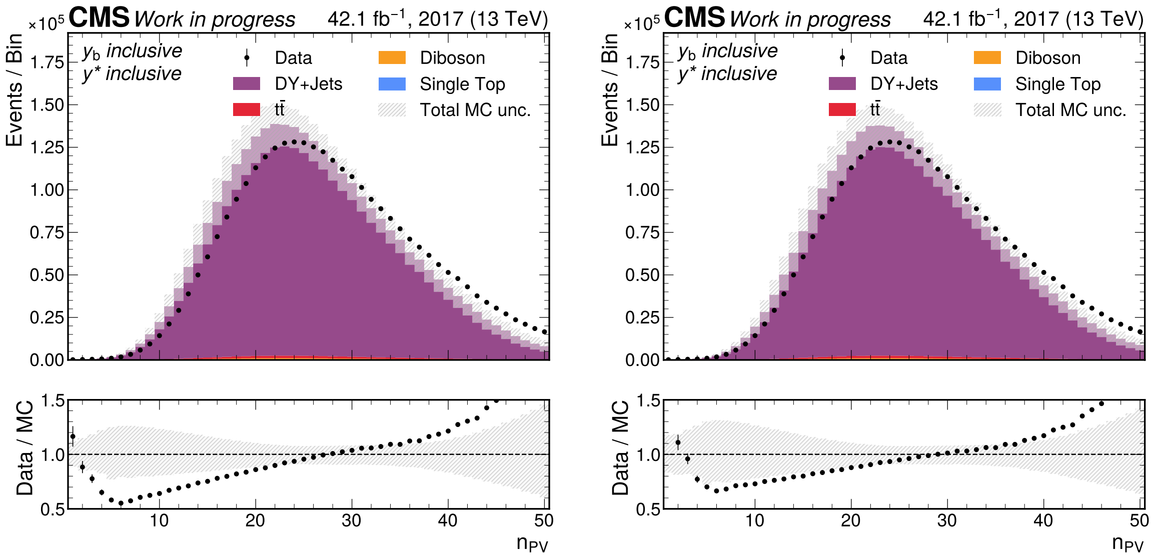


Figure A.9: Distribution of the number of reconstructed collision vertices in 2017 data compared to simulation before pileup reweighting (left) and after applying the pileup reweighting procedure (right).

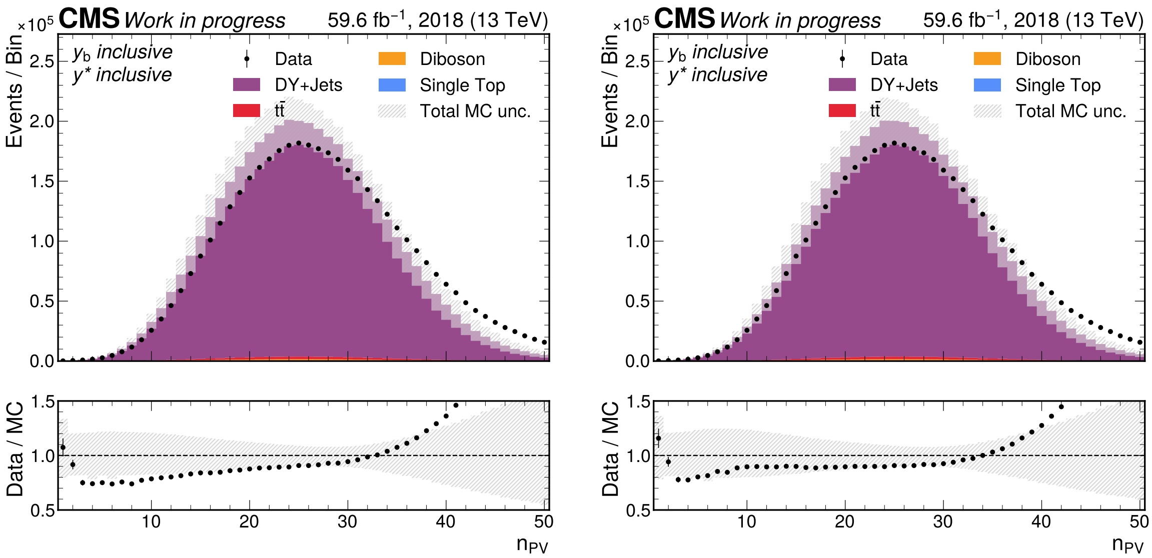


Figure A.10: Distribution of the number of reconstructed collision vertices in 2018 data compared to simulation before pileup reweighting (left) and after applying the pileup reweighting procedure (right).

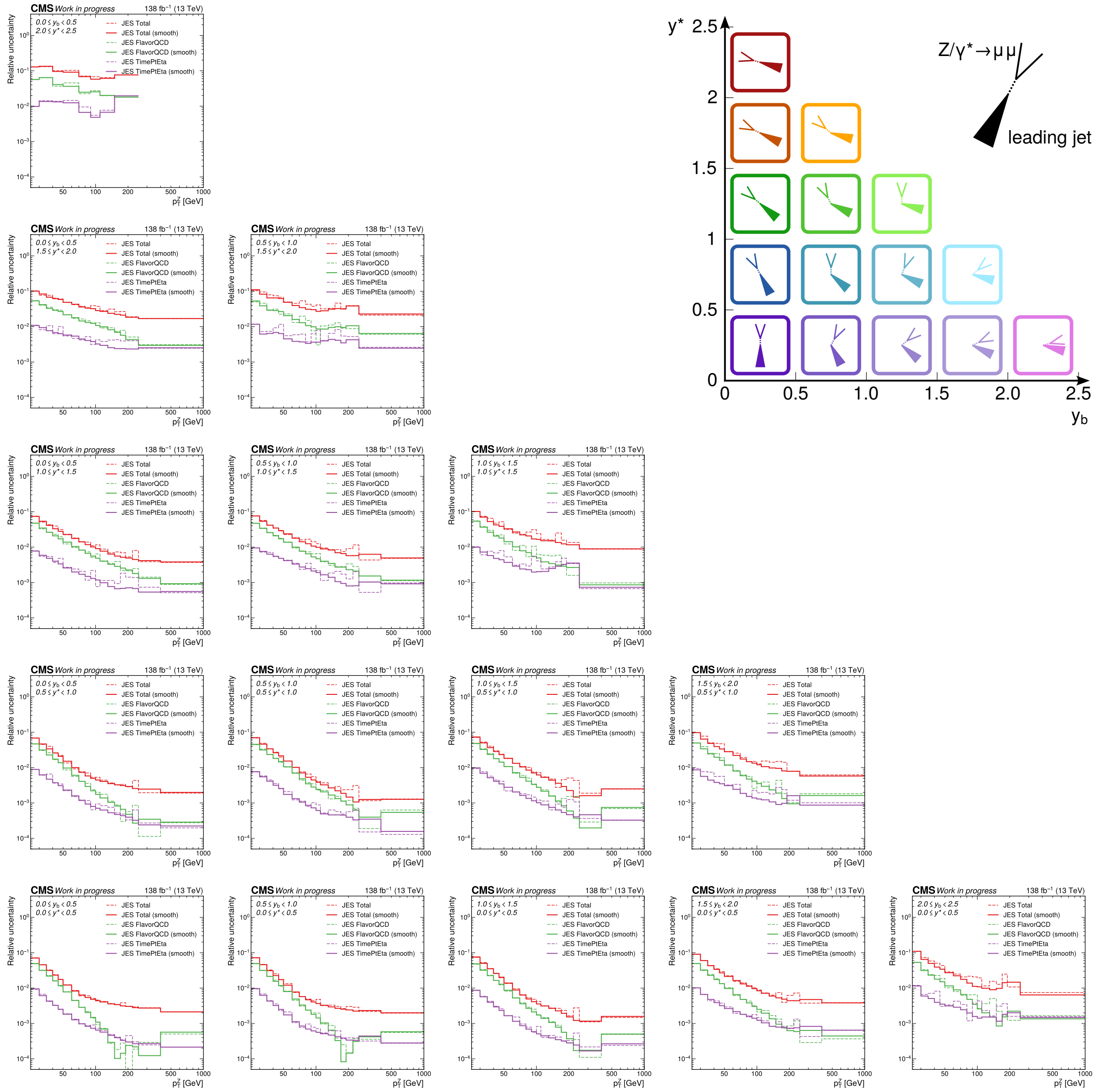


Figure A.11: FlavorQCD and TimePtEta jet energy scale (JES) uncertainty contributions (green and purple) before smoothing (dashed lines) and after smoothing (solid lines). Also, the combination of all JES uncertainty sources (red) before and after smoothing is shown.

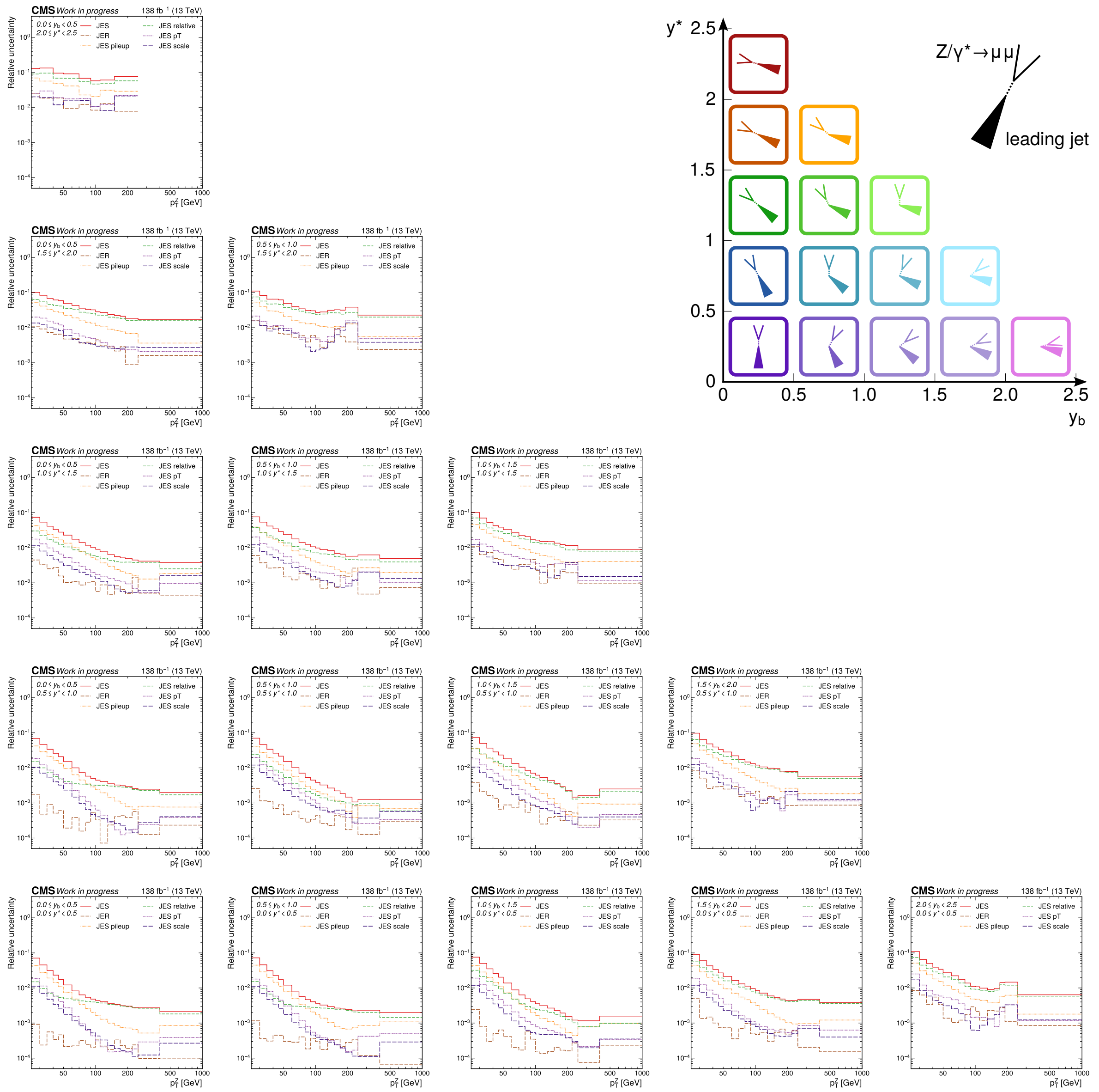


Figure A.12: Uncertainty breakdown for the JES and jet energy resolution (JER) uncertainties for all bins in (y_b, y^*). The plots also show the PileUp, Relative, Pt, and Scale subtotal contributions to the full JES uncertainty

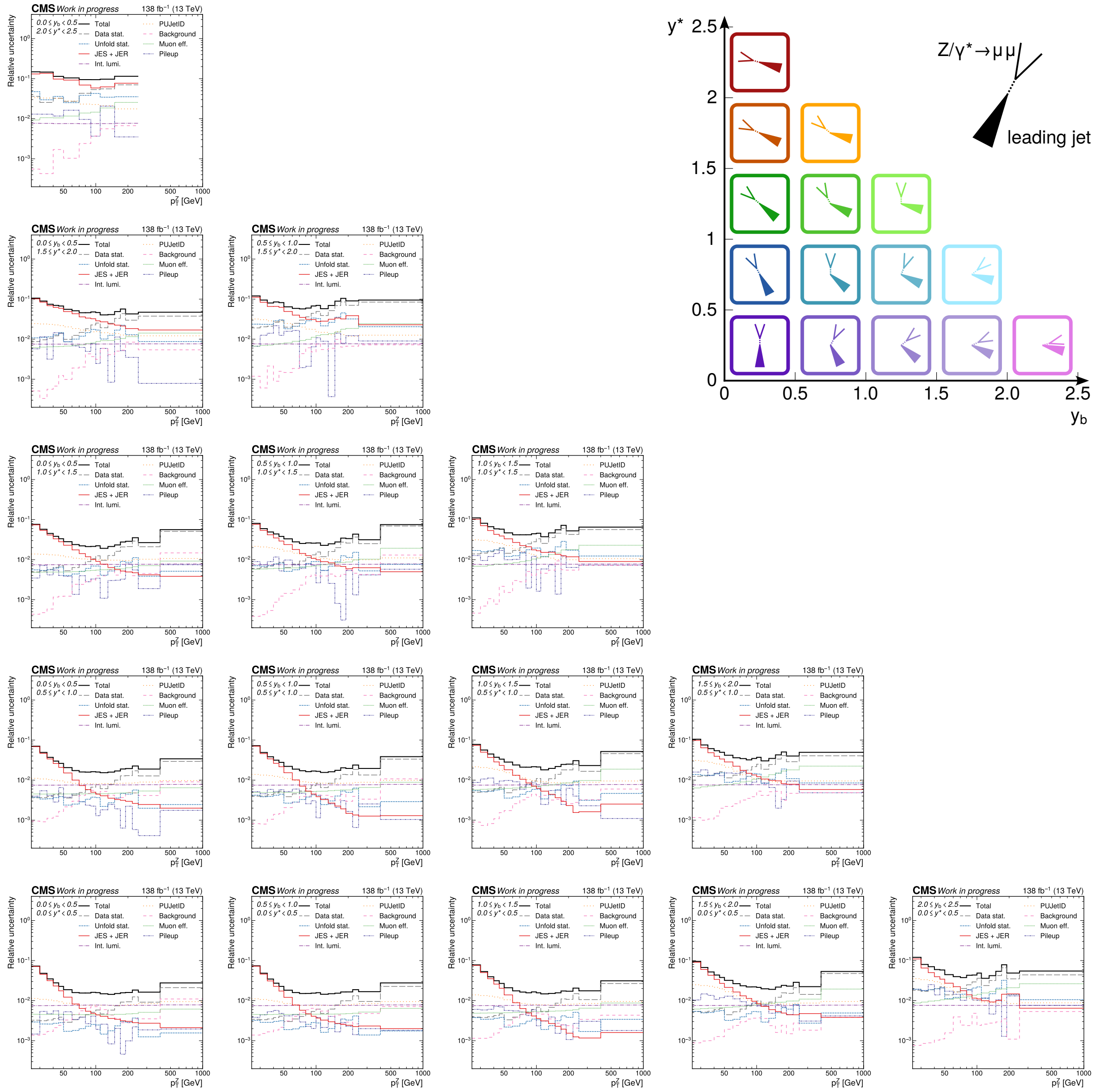


Figure A.13: Uncertainty breakdown of the unfolded cross section as a function of p_T^Z for all bins in (y_b, y^*).

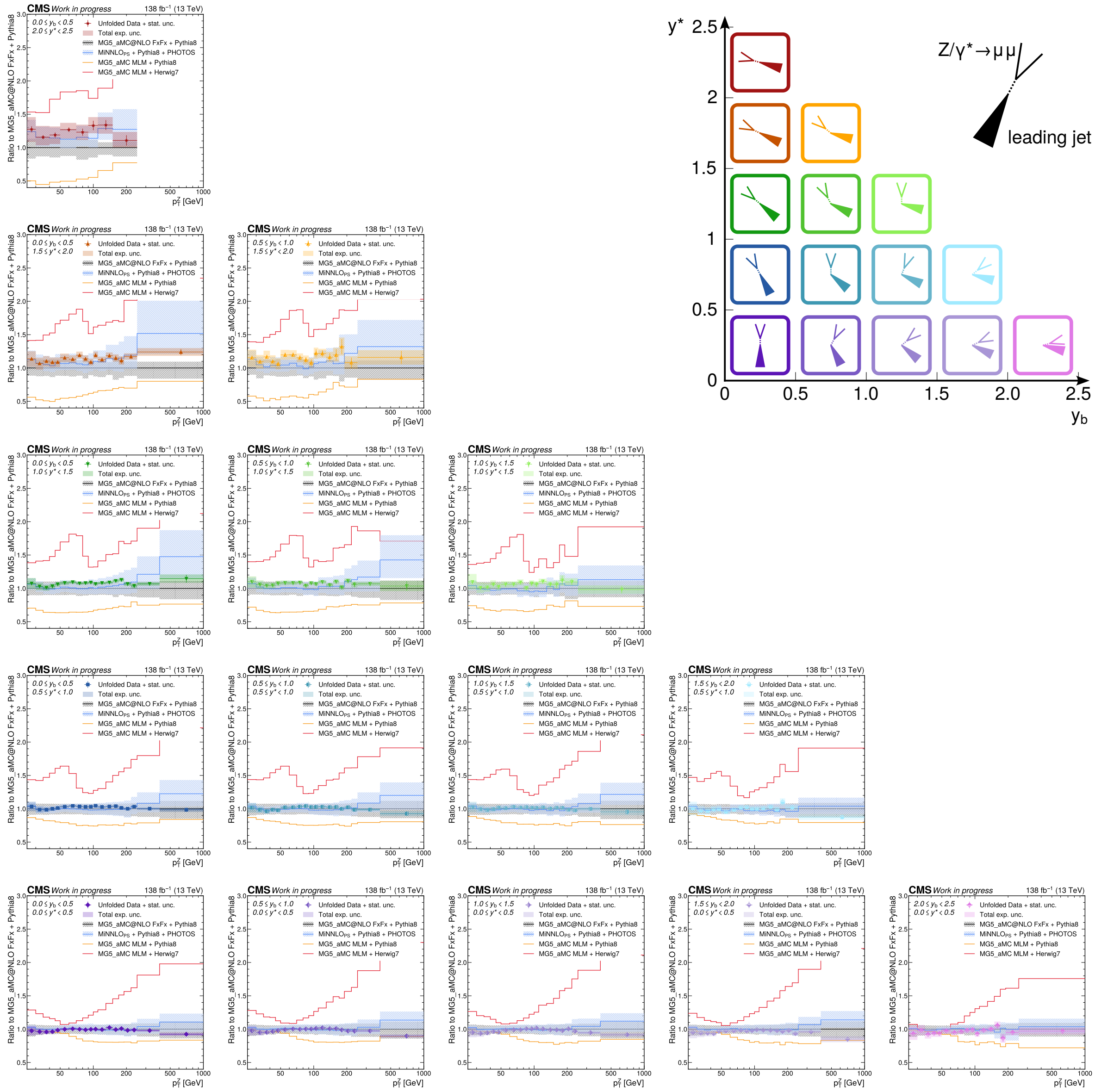


Figure A.14: Comparison of cross section predictions by different MC generators to the unfolded data in all 15 (y_{b}, y^{*}) bins. The corresponding breakdown of the (y_{b}, y^{*})-bins is illustrated in the top right. The data and MC predictions are shown as a ratio to the prediction by MadGraph5_aMC@NLO FxFx + PYTHIA 8. Both, the MadGraph5_aMC@NLO FxFx + PYTHIA 8 and POWHEG MiNNLO$_{\mathrm{PS}}$ + PYTHIA 8 predictions are shown with their uncertainties including scale variation, initial- and final-state radiation, PDF and their statistical uncertainty. The MadGraph5_aMC@NLO MLM merged samples interfaced to PYTHIA 8 (orange) and HERWIG 7 (red) are shown without uncertainties, as they are simulated at LO only and therefore have large scale uncertainty.

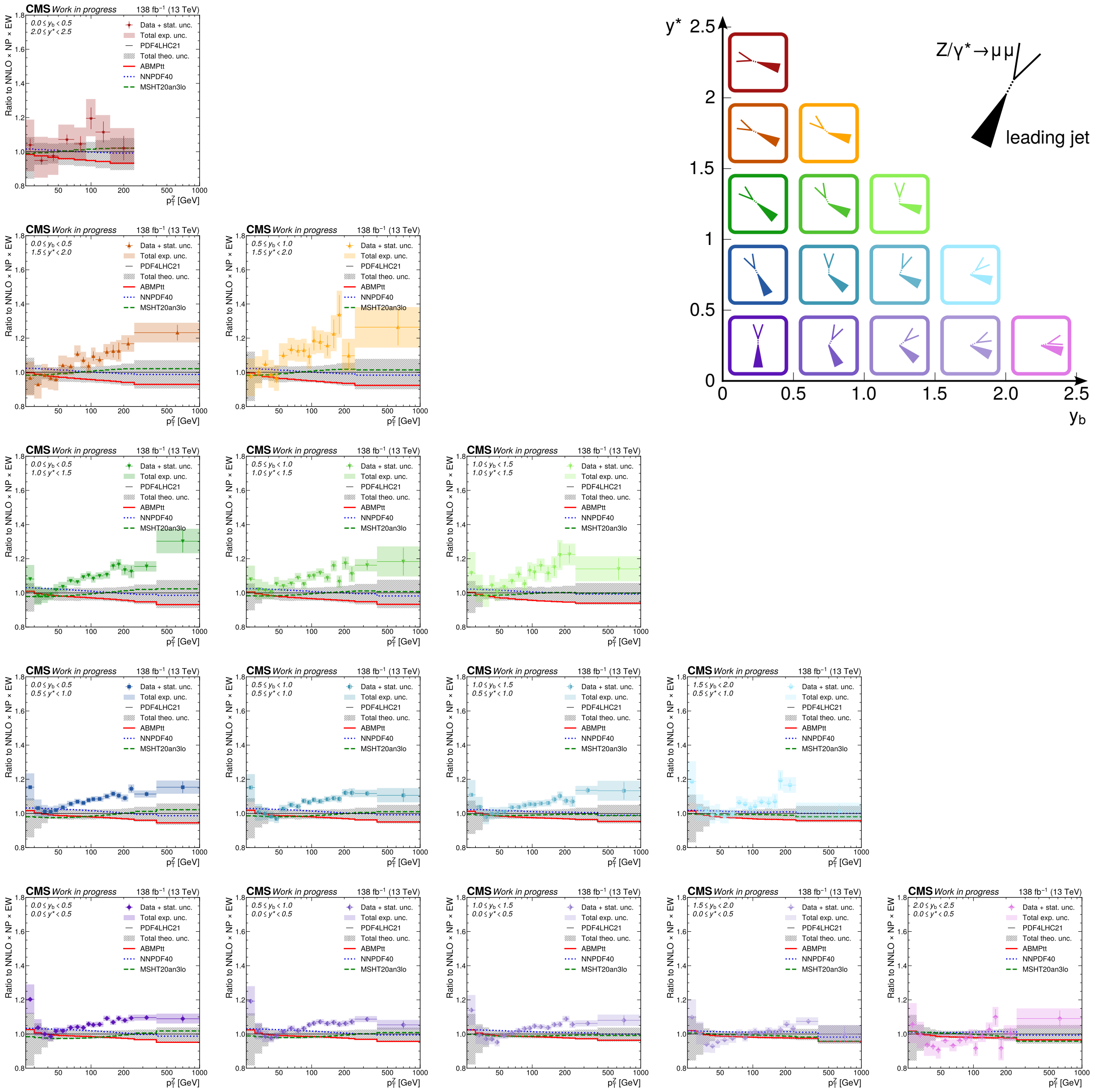


Figure A.15: Comparison of the unfolded data to theory predictions at NNLO in perturbative QCD corrected for NP and EW effects as a function of $p_{\mathrm{T}}^{\mathrm{Z}}$ in all 15 (y_{b}, y^{*})-bins. The corresponding breakdown of the (y_{b}, y^{*})-bins is illustrated in the top right. The error bars on the data points indicate the statistical uncertainty with the error band showing the full uncertainty on the measured cross sections. The uncertainty band on the theory prediction consists of scale variation, statistical uncertainties from the numerical integration in NNLOJET, the PDF4LHC21 uncertainties, the uncertainties in the determination of the NP correction factors and the uncertainty in the EW correction estimated from the multiplicative and additive correction approach. Additionally, the central values of the theory predictions using the five-flavour ABMPtt NNLO, NNPDF 4.0 and aN^{3}LO MSHT PDF sets are shown.

APPENDIX B

Software and computing

For the analysis of the data presented in this thesis, multiple software packages were used and contributed to in order to enable the analysis of the ultra-legacy Run 2 data and simulation. An overview of the software packages used for the analysis of CMS data and simulation is given in Section B.1. For the derivation of the nonperturbative (NP) correction factors, the workflow management tool MC-Run was developed. A brief overview is given in Section B.2. For the development of this software, as well as for writing this thesis, modern large language models were used, as outlined in Section B.3. A summary of the computing resources used over the past years to obtain the results presented in this thesis is given in Section B.4.

B.1 CMS data analysis

Central reconstruction of the data collected by the CMS experiment, as well as Monte Carlo (MC) sample production, is performed with the CMSSW software package.[1] The provided miniAODv2 data format is then skimmed using KAPPA.[2] This reduces the initial file size and allows storage on the Tier-3 disk storage at GridKA, enabling fast and reliable access to the data for further processing. Object and event selection, including the application of corrections and scale factors as well as the reconstruction of Z boson candidates, is performed with the EXCALIBUR[3] and ARTUS[4] frameworks. These produce flat n-tuples as output, which are then processed by KARMA[5] using a configuration specific to this analysis[6] and stored as histograms. The final plotting and overall orchestration of the workflow[7] are handled using the LAW [173] and LUIGI[8] workflow management tools.

[1]https://github.com/cms-sw/cmssw
[2]https://github.com/KIT-CMS/Kappa
[3]https://github.com/KIT-CMS/Excalibur
[4]https://github.com/KIT-CMS/Artus
[5]https://github.com/dsavoiu/Karma
[6]https://gitlab.etp.kit.edu/cverstege/zjet-postprocessing
[7]https://gitlab.etp.kit.edu/cverstege/zjet-analysis
[8]https://github.com/spotify/luigi

B.2 MC-Run

During the course of this analysis, one of the main challenges was the production and analysis of Monte Carlo simulation, in particular for deriving nonperturbative correction factors for a large number of MC generators and different settings. These calculations are computationally intensive and require substantial storage resources. To facilitate this, a dedicated MC-Run framework was developed to streamline the production, storage, and analysis of MC samples. The framework is designed to be modular and flexible, allowing for the easy integration of new MC generators and analysis modules.

Running and managing multiple dependent tasks, such as running MC generators and subsequently analysing the simulated events with RIVET [174] for further studies, is handled via the workflow manager Luigi and its extension LAW for use cases in high-energy physics.

MC-Run leverages the Grid infrastructure used in high-energy physics, such as distributed computing resources via HTCondor and data storage on Grid storage via XRootD or WebDAV. Support for this is implemented in the LAW framework.

MC-Run will be presented to a wider audience in the physics community at the 28th Conference on Computing in High Energy and Nuclear Physics.[9]

B.3 Use of artificial intelligence

In recent years, large language models (LLMs) have become widely available and have improved rapidly. They can support tasks such as code completion, debugging, proofreading, language refinement, and the restructuring of text. During the preparation of this thesis, LLM-based tools were used as auxiliary tools for selected tasks. In particular, ChatGPT[10], Gemini[11], and GitHub Copilot[12] were used.

The use of these tools was mainly limited to proofreading, improving clarity and wording, suggesting alternative formulations, and assisting with source-code development and debugging. All text, code, and analysis choices were reviewed, adapted where necessary, and checked.

The use of LLMs also introduces risks for research quality, including hallucinations, bias, and inaccurate or misleading output. Therefore, the tools were used in accordance with best practices for the responsible use of generative AI in research, as outlined in the report by the European Commission in Ref. [175]. All AI-assisted outputs relevant to this thesis were critically assessed and verified before inclusion.

B.4 Usage of computing resources

For the data analysis of the CMS data and the derivation of the NP corrections, a large amount of data storage and compute resources was required. At the end of writing this thesis, the

[9] https://indico.cern.ch/event/1471803/contributions/6968204/
[10] https://chatgpt.com
[11] https://gemini.google.com
[12] https://github.com/copilot

storage consumption for intermediate analysis products amounted to approximately 650 TB on the dCache storage system provided for the German CMS community at GridKA.

The computing tasks were distributed across several computing centres, including a total of 1 904 530 h of CPU time on the institute batch system, which integrates resources from Topas, NEMO, and NEMO2. The author would like to thank the state of Baden-Württemberg for its support of bwHPC and the German Research Foundation (DFG) through grant no INST 39/963-1 FUGG (bwForCluster NEMO) and for funding under Project number 455622343 (bwForCluster NEMO 2). Additional jobs were executed directly on NEMO2, accounting for 675 379 h of CPU time, while the CERN batch system (LxBatch) contributed 1 159 916 h. Furthermore, computing resources provided by the National Analysis Facility hosted at DESY were used, though no precise accounting of CPU usage is available. Here only a handful of analysis jobs were run, so the CPU usage is negligible compared to the other computing sites.

Overall, this results in a total CPU usage of approximately 3 739 825 h. A rough estimate of the associated energy consumption can be obtained using typical values for modern high-throughput scientific computing infrastructure. An average power consumption of 10 W per CPU core and a typical Power Usage Effectiveness (PUE) of 1.5 for WLCG sites are assumed [176, 177]. This yields an effective energy consumption of

$$E \approx 3\,739\,825\,\mathrm{h} \times 10\,\mathrm{W} \times 1.5 = 56\,097\,\mathrm{kW\,h}.$$

To convert energy consumption into greenhouse gas emissions, country-dependent grid carbon intensities are used. For 2025, typical values are approximately 31 g $\mathrm{CO_2e/kW\,h}$ for France[13] and 342 g $\mathrm{CO_2e/kW\,h}$ for Germany[14], based on Electricity Maps data. Assuming that computing at CERN is predominantly powered by low-carbon French electricity, while local computing resources at KIT are powered by the German grid, the total emissions are estimated as

$$\mathrm{CO_2^{CERN}} \approx 0.5\,\mathrm{t}, \qquad \mathrm{CO_2^{KIT}} \approx 10\,\mathrm{t}, \qquad \mathrm{CO_2^{NEMO2}} \approx 3.5\,\mathrm{t},$$

resulting in a total footprint of approximately $\mathcal{O}(14\,\mathrm{t\,CO_2e})$. To provide context, this corresponds to the emissions of approximately 6 round-trip flights for one person in economy class between Frankfurt and Sydney (2.3 t $\mathrm{CO_2e}$ per round trip via Abu Dhabi)[15]. This estimate relies on a number of simplifying assumptions and should therefore be interpreted as an order-of-magnitude illustration rather than a precise accounting of emissions.

[13] https://app.electricitymaps.com/map/zone/FR/all/yearly
[14] https://app.electricitymaps.com/map/zone/DE/all/yearly
[15] https://travelimpactmodel.org

List of Acronyms

ALICE A Large Ion Collider Experiment 32, 34

aN^{3}LO approximate next-to-NNLO 84, 90, 108

APV analogue pipeline voltage 39

ATLAS A Toroidal LHC Apparatus 32, 34, 35, 37

BDT boosted decision tree 54

BX bunch crossing 41, 42

CERN Conseil Européen pour la Recherche Nucléaire 1, 31, 32, 33, 111

CHS charged hadron subtraction 44, 53, 54

CMS Compact Muon Solenoid 2, 3, 12, 31, 32, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 51, 52, 54, 81, 89, 90, 92, 109, 110, 111

CSC cathode strip chamber 37, 40, 41, 43

DT drift tube 37, 40, 41, 43

ECAL electromagnetic calorimeter 35, 39, 40, 42, 43

EW electroweak 7, 14, 15, 20, 21, 22, 23, 28, 29, 46, 81, 84, 85, 86, 87, 89, 90, 98, 101, 108

FSR final-state radiation 24, 75, 82

GEM gas electron multiplier 40

HCAL hadronic calorimeter 35, 40, 42, 43

HEP high energy physics 31

HIPM high ionizing particle multiplicity 39

HLT high-level trigger 41

IP interaction point 32, 34, 35, 36, 37, 38

ISR initial-state radiation 75, 82

JER jet energy resolution 72, 75, 77, 78, 105

JES jet energy scale 72, 73, 74, 75, 76, 77, 78, 89, 104, 105

L1 Level-1 41, 42

LHC Large Hadron Collider 1, 2, 5, 7, 9, 20, 31, 32, 33, 34, 35, 37, 38, 41, 45, 48, 90

LHCb LHC beauty 32, 34

LO leading order 14, 16, 17, 19, 22, 23, 26, 27, 46, 55, 66, 67, 68, 81, 82, 83, 96, 107, 118

MC Monte Carlo 2, 24, 26, 39, 45, 47, 49, 55, 57, 64, 69, 78, 80, 81, 82, 83, 89, 91, 95, 107, 109, 110, 118

MPI multiple parton interaction 24, 26

NLO next-to-leading order 11, 14, 15, 16, 17, 18, 19, 20, 22, 23, 26, 27, 29, 46, 47, 66, 68, 75, 78, 80, 81, 82, 84, 89, 101

NNLO next-to-next-to-leading order 1, 2, 5, 16, 17, 18, 19, 20, 22, 23, 26, 28, 29, 46, 47, 66, 80, 81, 82, 84, 89, 95, 101, 108

NP nonperturbative 2, 26, 28, 29, 81, 84, 85, 86, 87, 89, 90, 100, 101, 108, 109, 110

PDF parton distribution function 1, 2, 5, 9, 10, 16, 19, 22, 28, 29, 46, 55, 73, 75, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 97, 101, 107, 108

PF particle-flow 31, 42, 43, 44, 51, 52, 53, 54

PS parton shower 24, 25, 26, 89

PU pileup 34, 38, 42, 44, 47, 49, 51, 53, 54, 55, 72, 78, 95

PUJetID pileup jet identification 54, 78

PV primary vertex 38, 44, 51

QCD quantum chromodynamics 1, 2, 5, 7, 8, 9, 11, 14, 15, 16, 17, 19, 20, 22, 23, 24, 26, 28, 43, 46, 47, 66, 75, 80, 81, 82, 84, 85, 86, 87, 89, 90, 96, 108

QED quantum electrodynamics 7, 8, 46, 82

RPC resistive plate chamber 37, 40, 41, 42, 43

SM Standard Model 1, 5, 6, 7, 90

SPS Super Proton Synchrotron 32

UE underlying event 24

VFP voltage feedback preamplifier 39, 46, 47, 48, 49, 55, 91, 92, 93, 102

List of Figures

List of Tables

Bibliography

[1] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," Phys. Lett. B **716**, 30 (2012) 10.1016/j.physletb.2012.08.021, arXiv:1207.7235 [hep-ex].

[2] ATLAS Collaboration, "Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC," Phys. Lett. B **716**, 1 (2012) 10.1016/j.physletb.2012.08.020, arXiv:1207.7214 [hep-ex].

[3] CMS Collaboration, "High-precision measurement of the W boson mass with the CMS experiment," Nature **652**, 321 (2026) 10.1038/s41586-026-10168-5, arXiv:2412.13872 [hep-ex].

[4] ATLAS Collaboration, "Measurement of the W-boson mass and width with the ATLAS detector using proton–proton collisions at $\sqrt{s} = 7$ TeV," Eur. Phys. J. C **84**, 1309 (2024) 10.1140/epjc/s10052-024-13190-x, arXiv:2403.15085 [hep-ex].

[5] 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," JHEP **09**, 049 (2015) 10.1007/JHEP09(2015)049, arXiv:1503.03709 [hep-ex].

[6] 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," PLB **866**, 139526 (2025) 10.1016/j.physletb.2025.139526, arXiv:2408.07622 [hep-ex].

[7] LHC Higgs Cross Section Working Group, S. Heinemeyer, et al., *Handbook of LHC Higgs cross sections: 3. Higgs properties*, CERN Report CERN-2013-004 (2013), 10.5170/CERN-2013-004, arXiv:1307.1347 [hep-ph].

[8] LHC Higgs Cross Section Working Group, D. de Florian, et al., *Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector*, CERN Report CERN-2017-002-M (2016), 10.23731/CYRM-2017-002, arXiv:1610.07922 [hep-ph].

[9] T.-J. Hou, J. Gao, T. J. Hobbs, K. Xie, S. Dulat, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, I. Sitiwaldi, D. Stump, and C.-P. Yuan, "New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC," Phys. Rev. D **103**, 014013 (2021) 10.1103/PhysRevD.103.014013, arXiv:1912.10053 [hep-ph].

[10] ATLAS Collaboration, "Determination of the parton distribution functions of the proton from ATLAS measurements of differential $W^{\pm}$ and Z boson production in association with jets," JHEP **07**, 223 (2021) 10.1007/JHEP07(2021)223, arXiv:2101.05095 [hep-ex].

[11] NNPDF Collaboration, "The path to proton structure at 1% accuracy," Eur. Phys. J. C **82**, 428 (2022) 10.1140/epjc/s10052-022-10328-7, arXiv:2109.02653 [hep-ph].

[12] T. Cridge, L. A. Harland-Lang, and R. S. Thorne, "The impact of LHC jet and Z p_T data at up to approximate N^3LO order in the MSHT global PDF fit," Eur. Phys. J. C **84**, 446 (2024) 10.1140/epjc/s10052-024-12771-0, arXiv:2312.12505 [hep-ph].

[13] D0 Collaboration, "Measurement of differential Z/γ^*+jet+X cross sections in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. B **669**, 278 (2008) 10.1016/j.physletb.2008.09.060, arXiv:0808.1296 [hep-ex].

[14] D0 Collaboration, "Measurements of differential cross sections of Z/γ^*+jet+X events in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. B **678**, 45 (2009) 10.1016/j.physletb.2009.05.058, arXiv:0903.1748 [hep-ex].

[15] D0 Collaboration, "Measurement of Z/γ^*+jet+X angular distributions in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Lett. B **682**, 370 (2010) 10.1016/j.physletb.2009.11.012, arXiv:0907.4286 [hep-ex].

[16] CDF Collaboration, "Measurement of differential production cross section for Z/γ^* bosons in association with jets in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV," Phys. Rev. D **91**, 012002 (2015) 10.1103/PhysRevD.91.012002, arXiv:1409.4359 [hep-ex].

[17] CMS Collaboration, "Measurements of jet multiplicity and differential production cross sections of Z+jets events in proton-proton collisions at $\sqrt{s} = 7$ TeV," Phys. Rev. D **91**, 052008 (2015) 10.1103/PhysRevD.91.052008, arXiv:1408.3104 [hep-ex].

[18] CMS Collaboration, "Measurements of differential production cross sections for a Z boson in association with jets in pp collisions at $\sqrt{s} = 8$ TeV," JHEP **04**, 022 (2017) 10.1007/JHEP04(2017)022, arXiv:1611.03844 [hep-ex].

[19] CMS Collaboration, "Measurement of differential cross sections for Z boson production in association with jets in proton-proton collisions at $\sqrt{s} = 13$ TeV," Eur. Phys. J. C **78**, 965 (2018) 10.1140/epjc/s10052-018-6373-0, arXiv:1804.05252 [hep-ex].

[20] ATLAS Collaboration, "Measurement of the inclusive cross-section for the production of jets in association with a Z boson in proton-proton collisions at 8 TeV using the ATLAS detector," Eur. Phys. J. C **79**, 847 (2019) 10.1140/epjc/s10052-019-7321-3, arXiv:1907.06728 [hep-ex].

[21] ATLAS Collaboration, "Cross-section measurements for the production of a Z boson in association with high-transverse-momentum jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," JHEP **06**, 080 (2023) 10.1007/JHEP06(2023)080, arXiv:2205.02597 [hep-ex].

[22] CMS Collaboration, "Measurement of differential cross sections for the production of a Z boson in association with jets in proton-proton collisions at $\sqrt{s} = 13$ TeV," Phys. Rev. D **108**, 052004 (2023) 10.1103/PhysRevD.108.052004, arXiv:2205.02872 [hep-ex].

[23] ATLAS Collaboration, "Simultaneous unbinned differential cross-section measurement of twenty-four Z+jets kinematic observables with the ATLAS detector," Phys. Rev. Lett. **133**, 261803 (2024) 10.1103/PhysRevLett.133.261803, arXiv:2405.20041 [hep-ex].

[24] R. Boughezal, J. M. Campbell, R. K. Ellis, C. Focke, W. T. Giele, X. Liu, and F. Petriello, "Z-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD," Phys. Rev. Lett. **116**, 152001 (2016) 10.1103/PhysRevLett.116.152001, arXiv:1512.01291 [hep-ph].

[25] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss, and T. A. Morgan, "Precise QCD predictions for the production of a Z boson in association with a hadronic jet," Phys. Rev. Lett. **117**, 022001 (2016) 10.1103/PhysRevLett.117.022001, arXiv:1507.02850 [hep-ph].

[26] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, A. Huss, and T. A. Morgan, "The NNLO QCD corrections to Z boson production at large transverse momentum," JHEP **07**, 133 (2016) 10.1007/JHEP07(2016)133, arXiv:1605.04295 [hep-ph].

[27] T. Berger, "Jet energy calibration and triple differential inclusive cross section measurements with $Z(\to \mu\mu)$ + jet events at 13 TeV recorded by the CMS detector," PhD thesis (Karlsruher Institut für Technologie, 2019), 10.5445/IR/1000104286.

[28] M. Schnepf, "Dynamic provision of heterogeneous computing resources for computation- and data-intensive particle physics analyses," PhD thesis (Karlsruher Institut für Technologie, 2022), 10.5445/IR/1000143165.

[29] C. Verstege, "Measurement of the triple-differential cross-section of Z+jet production with the CMS detector at 13 TeV," MA thesis (Karlsruher Institut für Technologie, 2022).

[30] M. M. Horzela, "Measurement of triple-differential Z+jet cross sections with the CMS detector at 13 TeV and modelling of large-scale distributed computing systems," PhD thesis (Karlsruher Institut für Technologie, 2023), 10.5445/IR/1000165566.

[31] CMS Collaboration, *Triple-differential measurement of Z+jet production in proton-proton collisions at $\sqrt{s} = 13$ TeV*, CMS Physics Analysis Summary CMS-PAS-SMP-24-010 (2026).

[32] S. Gieseke, M. Horzela, M. Kaur, D. Leonardi, K. Rabbertz, A. Singla, and C. Verstege, "Nonperturbative effects in triple-differential dijet and Z+jet production at the LHC," SciPost Phys. **21**, 033 (2026) 10.21468/SciPostPhys.21.2.033, arXiv:2412.19694 [hep-ph].

[33] CMS Collaboration, "A portrait of the Higgs boson by the CMS experiment ten years after the discovery," Nature **607**, [Author correction: doi: 10.1038/s41586-023-06164-8], 60 (2022) 10.1038/s41586-022-04892-x, arXiv:2207.00043 [hep-ex].

[34] ATLAS Collaboration, "A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery," Nature **607**, [Publisher correction: doi: 10.1038/s41586-022-05581-5, author correction: doi: 10.1038/s41586-023-06248-5], 52 (2022) 10.1038/s41586-022-04893-w, arXiv:2207.00092 [hep-ex].

[35] F. Englert and R. Brout, "Broken symmetry and the mass of gauge vector mesons," Phys. Rev. Lett. **13**, edited by J. C. Taylor, 321 (1964) 10.1103/PhysRevLett.13.321.

[36] P. W. Higgs, "Broken symmetries and the masses of gauge bosons," Phys. Rev. Lett. **13**, edited by J. C. Taylor, 508 (1964) 10.1103/PhysRevLett.13.508.

[37] J. Campbell, J. Huston, and F. Krauss, *The black book of quantum chromodynamics: a primer for the LHC era* (Oxford University Press, 2018), 10.1093/oso/9780199652747.001.0001.

[38] Particle Data Group Collaboration, "Review of particle physics," Phys. Rev. D **110**, 030001 (2024) 10.1103/PhysRevD.110.030001.

[39] K. Rabbertz, *Jet physics at the LHC: the strong force beyond the TeV scale*, Vol. 268, Springer Tracts in Modern Physics (Springer, Berlin, 2017), 10.1007/978-3-319-42115-5.

[40] D. J. Gross and F. Wilczek, "Ultraviolet behavior of nonabelian gauge theories," Phys. Rev. Lett. **30**, 1343 (1973) 10.1103/PhysRevLett.30.1343.

[41] H. D. Politzer, "Reliable perturbative results for strong interactions?" Phys. Rev. Lett. **30**, 1346 (1973) 10.1103/PhysRevLett.30.1346.

[42] V. N. Gribov and L. N. Lipatov, "Deep inelastic ep scattering in perturbation theory," Sov. J. Nucl. Phys. **15**, 438 (1972).

[43] L. N. Lipatov, "The parton model and perturbation theory," Sov. J. Nucl. Phys. **20**, 94 (1975).

[44] G. Altarelli and G. Parisi, "Asymptotic freedom in parton language," Nucl. Phys. B **126**, 298 (1977) 10.1016/0550-3213(77)90384-4.

[45] Y. L. Dokshitzer, "Calculation of the structure functions for deep inelastic scattering and $e^+ e^-$ annihilation by perturbation theory in quantum chromodynamics," Sov. Phys. JETP **46**, 641 (1977).

[46] W. T. Giele, E. W. N. Glover, and D. A. Kosower, "Higher order corrections to jet cross-sections in hadron colliders," Nucl. Phys. B **403**, 633 (1993) 10.1016/0550-3213(93)90365-V, arXiv:hep-ph/9302225.

[47] T. Gehrmann, X. Chen, J. Cruz-Martinez, J. Currie, R. Gauld, N. Glover, A. Gehrmann-De Ridder, A. Huss, T. Morgan, J. Niehues, J. Pires, and D. Walker, "Jet cross sections and transverse momentum distributions with NNLOJET," in Proc. 13th International Symposium on Radiative Corrections (RADCOR2017): St. Gilgen, Austria, September 24–27, 2017, [PoS (RADCOR2017) 074] (2018), 10.22323/1.290.0074, arXiv:1801.06415 [hep-ph].

[48] D. Britzger, A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, C. Gwenlan, A. Huss, J. Pires, K. Rabbertz, D. Savoiu, M. R. Sutton, and J. Stark, "NNLO interpolation grids for jet production at the LHC," Eur. Phys. J. C **82**, 930 (2022) 10.1140/epjc/s10052-022-10880-2, arXiv:2207.13735 [hep-ph].

[49] D. Britzger, K. Rabbertz, F. Stober, and M. Wobisch, "New features in version 2 of the fastNLO project," in Proc. 20th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS2012): Bonn, Germany, March 26–30, 2012 (2012), 10.3204/DESY-PROC-2012-02/165, arXiv:1208.3641 [hep-ph].

[50] PDF4LHC Working Group, R. D. Ball, et al., "The PDF4LHC21 combination of global PDF fits for the LHC Run 3," J. Phys. G **49**, 080501 (2022) 10.1088/1361-6471/ac7216, arXiv:2203.05506 [hep-ph].

[51] S. Kallweit, J. M. Lindert, P. Maierhofer, S. Pozzorini, and M. Schönherr, "NLO QCD+EW predictions for V+jets including off-shell vector-boson decays and multijet merging," JHEP **04**, 021 (2016) 10.1007/JHEP04(2016)021, arXiv:1511.08692 [hep-ph].

[52] J. M. Lindert et al., "Precise predictions for V+jets dark matter backgrounds," Eur. Phys. J. C **77**, 829 (2017) 10.1140/epjc/s10052-017-5389-1, arXiv:1705.04664 [hep-ph].

[53] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H. .-. Shao, and M. Zaro, "The automation of next-to-leading order electroweak calculations," JHEP **07**, [Erratum: JHEP 11, 085 (2021)], 185 (2018) 10.1007/JHEP11(2021)085, arXiv:1804.10017 [hep-ph].

[54] A. Manohar, P. Nason, G. P. Salam, and G. Zanderighi, "How bright is the proton? A precise determination of the photon parton distribution function," Phys. Rev. Lett. **117**, 242002 (2016) 10.1103/PhysRevLett.117.242002, arXiv:1607.04266 [hep-ph].

[55] A. V. Manohar, P. Nason, G. P. Salam, and G. Zanderighi, "The photon content of the proton," JHEP **12**, 046 (2017) 10.1007/JHEP12(2017)046, arXiv:1708.01256 [hep-ph].

[56] J. Butterworth et al., "PDF4LHC recommendations for LHC Run 2," J. Phys. G **43**, 023001 (2016) 10.1088/0954-3899/43/2/023001, arXiv:1510.03865 [hep-ph].

[57] E. Bothmann and D. Napoletano, "Automated evaluation of electroweak Sudakov logarithms in SHERPA," Eur. Phys. J. C **80**, 1024 (2020) 10.1140/epjc/s10052-020-08596-2, arXiv:2006.14635 [hep-ph].

[58] I. G. Knowles, "Spin correlations in parton - parton scattering," Nucl. Phys. B **310**, 571 (1988) 10.1016/0550-3213(88)90092-2.

[59] I. G. Knowles, "A linear algorithm for calculating spin correlations in hadronic collisions," Comput. Phys. Commun. **58**, 271 (1990) 10.1016/0010-4655(90)90063-7.

[60] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An introduction to PYTHIA 8.2," Comput. Phys. Commun. **191**, 159 (2015) 10.1016/j.cpc.2015.01.024, arXiv:1410.3012 [hep-ph].

[61] J. Bellm et al., "HERWIG 7.2 release note," Eur. Phys. J. C **80**, 452 (2020) 10.1140/epjc/s10052-020-8011-x, arXiv:1912.06509 [hep-ph].

[62] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, "Event generation with SHERPA 1.1," JHEP **02**, 007 (2009) 10.1088/1126-6708/2009/02/007, arXiv:0811.4622 [hep-ph].

[63] T. Sjostrand and P. Z. Skands, "Transverse-momentum-ordered showers and interleaved multiple interactions," Eur. Phys. J. C **39**, 129 (2005) 10.1140/epjc/s2004-02084-y, arXiv:hep-ph/0408302.

[64] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, "Parton fragmentation and string dynamics," Phys. Rept. **97**, 31 (1983) 10.1016/0370-1573(83)90080-7.

[65] B. Andersson, G. Gustafson, and B. Soderberg, "A general model for jet fragmentation," Z. Phys. C **20**, 317 (1983) 10.1007/BF01407824.

[66] T. Sjostrand, "The merging of jets," Phys. Lett. B **142**, 420 (1984) 10.1016/0370-2693(84)91354-6.

[67] G. Marchesini and B. R. Webber, "Monte carlo simulation of general hard processes with coherent QCD radiation," Nucl. Phys. B **310**, 461 (1988) 10.1016/0550-3213(88)90089-2.

[68] S. Gieseke, P. Stephens, and B. Webber, "New formalism for QCD parton showers," JHEP **12**, 045 (2003) 10.1088/1126-6708/2003/12/045, arXiv:hep-ph/0310083.

[69] B. R. Webber, "A QCD model for jet fragmentation including soft gluon interference," Nucl. Phys. B **238**, 492 (1984) 10.1016/0550-3213(84)90333-X.

[70] S. Schumann and F. Krauss, "A parton shower algorithm based on Catani–Seymour dipole factorisation," JHEP **03**, 038 (2008) 10.1088/1126-6708/2008/03/038, arXiv:0709.1027 [hep-ph].

[71] J.-C. Winter, F. Krauss, and G. Soff, "A modified cluster hadronization model," Eur. Phys. J. C **36**, 381 (2004) 10.1140/epjc/s2004-01960-8, arXiv:hep-ph/0311085.

[72] T. Sjostrand and M. van Zijl, "A multiple interaction model for the event structure in hadron collisions," Phys. Rev. D **36**, 2019 (1987) 10.1103/PhysRevD.36.2019.

[73] T. Sjostrand and P. Z. Skands, "Multiple interactions and the structure of beam remnants," JHEP **03**, 053 (2004) 10.1088/1126-6708/2004/03/053, arXiv:hep-ph/0402078.

[74] M. Bahr, S. Gieseke, and M. H. Seymour, "Simulation of multiple partonic interactions in HERWIG++," JHEP **07**, 076 (2008) 10.1088/1126-6708/2008/07/076, arXiv:0803.3633 [hep-ph].

[75] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, "Matching matrix elements and shower evolution for top-pair production in hadronic collisions," JHEP **01**, 013 (2007) 10.1088/1126-6708/2007/01/013, arXiv:hep-ph/0611129.

[76] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, "QCD matrix elements + parton showers," JHEP **11**, 063 (2001) 10.1088/1126-6708/2001/11/063, arXiv:hep-ph/0109231.

[77] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," JHEP **06**, 029 (2002) 10.1088/1126-6708/2002/06/029, arXiv:hep-ph/0204244.

[78] S. Hoeche, F. Krauss, M. Schönherr, and F. Siegert, "QCD matrix elements + parton showers: the NLO case," JHEP **04**, 027 (2013) 10.1007/JHEP04(2013)027, arXiv:1207.5030 [hep-ph].

[79] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms," JHEP **11**, 040 (2004) 10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[80] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method," JHEP **11**, 070 (2007) 10.1088/1126-6708/2007/11/070, arXiv:0709.2092 [hep-ph].

[81] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX," JHEP **06**, 043 (2010) 10.1007/JHEP06(2010)043, arXiv:1002.2581 [hep-ph].

[82] P. F. Monni, P. Nason, E. Re, M. Wiesemann, and G. Zanderighi, "MiNNLO$_{PS}$: a new method to match NNLO QCD to parton showers," JHEP **05**, [Erratum: doi: 10.1007/JHEP02(2022)031], 143 (2020) 10.1007/JHEP05(2020)143, arXiv:1908.06987 [hep-ph].

[83] P. F. Monni, E. Re, and M. Wiesemann, "MiNNLO$_{PS}$: optimizing $2 \to 1$ hadronic processes," Eur. Phys. J. C **80**, 1075 (2020) 10.1140/epjc/s10052-020-08658-5, arXiv:2006.04133 [hep-ph].

[84] Sherpa Collaboration, "Event generation with SHERPA 2.2," SciPost Phys. **7**, 034 (2019) 10.21468/SciPostPhys.7.3.034, arXiv:1905.09127 [hep-ph].

[85] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk, H. L. Lai, and W. K. Tung, "Uncertainties of predictions from parton distribution functions. II. The Hessian method," Phys. Rev. D **65**, 014013 (2001) 10.1103/PhysRevD.65.014013, arXiv:hep-ph/0101032.

[86] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, and W.-K. Tung, "New generation of parton distributions with uncertainties from global QCD analysis," JHEP **07**, 012 (2002) 10.1088/1126-6708/2002/07/012, arXiv:hep-ph/0201195 [hep-ph].

[87] NNPDF Collaboration, "A determination of parton distributions with faithful uncertainty estimation," Nucl. Phys. B **809**, [Erratum: doi: j.nuclphysb.2009.02.027], 1 (2009) 10.1016/j.nuclphysb.2008.09.037, arXiv:0808.1231 [hep-ph].

[88] NNPDF Collaboration, "Parton distributions with LHC data," Nucl. Phys. B **867**, 244 (2013) 10.1016/j.nuclphysb.2012.10.003, arXiv:1207.1303 [hep-ph].

[89] R. Alemany-Fernandez, E. Bravin, L. Drosdal, A. Gorzawski, V. Kain, M. Lamont, A. Macpherson, G. Papotti, M. Pojer, L. Ponce, S. Redaelli, G. Roy, M. Solfaroli Camillocci, W. Venturini, and J. Wenninger, *Operation and configuration of the LHC in Run 1*, CERN Report CERN-ACC-NOTE-2013-0041 (2013).

[90] J. Wenninger, *Operation and configuration of the LHC in Run 2*, CERN Report CERN-ACC-NOTE-2019-0007 (2019).

[91] G. Apollinari et al., *High-Luminosity Large Hadron Collider (HL-LHC): technical design report v.0.1*, CERN Technical Proposal CERN-2017-007-M (2017), 10.23731/CYRM-2017-004.

[92] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC design report. Volume I: the LHC main ring*, CERN report CERN-2004-003-V1 (2004), 10.5170/CERN-2004-003-V-1.

[93] E. Mobs, *The CERN accelerator complex - August 2018*, General Photo, (2018) https://cds.cern.ch/record/2636343.

[94] *Linear accelerator 2*, https://home.cern/science/accelerators/linear-accelerator-2 (visited on 2026-01-07).

[95] *The proton synchrotron*, https://home.cern/science/accelerators/proton-synchrotron (visited on 2026-01-07).

[96] *The super proton synchrotron*, https://home.cern/science/accelerators/super-proton-synchrotron (visited on 2026-01-07).

[97] UA1 Collaboration, "Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540\,\mathrm{GeV}$," Phys. Lett. B **122**, 103 (1983) 10.1016/0370-2693(83)91177-2.

[98] UA2 Collaboration, "Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the cern $\overline{\mathrm{p}}\mathrm{p}$ collider," Phys. Lett. B **122**, 476 (1983) 10.1016/0370-2693(83)91605-2.

[99] UA1 Collaboration, "Experimental observation of lepton pairs of invariant mass around $95\,\mathrm{GeV}/c^2$ at the CERN SPS collider," Phys. Lett. B **126**, 398 (1983) 10.1016/0370-2693(83)90188-0.

[100] UA2 Collaboration, "Evidence for $\mathrm{Z}^0 \to \mathrm{e}^+\mathrm{e}^-$ at the CERN $\overline{\mathrm{p}}\mathrm{p}$ collider," Phys. Lett. B **129**, 130 (1983) 10.1016/0370-2693(83)90744-X.

[101] M. Benedikt, P. Collier, V. Mertens, J. Poole, and K. Schindl, *LHC design report. Volume III: the LHC injector chain*, CERN report CERN-2004-003-V-3 (2004), 10.5170/CERN-2004-003-V-3.

[102] *Public CMS luminosity information*, https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults (visited on 2026-04-01).

[103] CMS Collaboration, "The CMS experiment at the CERN LHC," JINST **3**, S08004 (2008) 10.1088/1748-0221/3/08/S08004.

[104] CMS Collaboration, "Development of the CMS detector for the CERN LHC Run 3," JINST **19**, P05064 (2024) 10.1088/1748-0221/19/05/P05064, arXiv:2309.05466 [physics.ins-det].

[105] ATLAS Collaboration, "The ATLAS experiment at the CERN Large Hadron Collider," JINST **3**, S08003 (2008) 10.1088/1748-0221/3/08/S08003.

[106] CMS Collaboration, *Interactive slice of the CMS detector*, (2016) https://cds.cern.ch/record/2205172.

[107] I. Neutelings, *CMS coordinate system*, https://tikz.net/axis3d_cms/ (visited on 2026-01-08).

[108] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker," JINST **9**, P10009 (2014) 10.1088/1748-0221/9/10/P10009, arXiv:1405.6569 [physics.ins-det].

[109] CMS Collaboration, *Technical proposal for the Phase-II upgrade of the Compact Muon Solenoid*, CMS Technical Proposal CERN-LHCC-2015-010, CMS-TDR-15-02 (2015), 10.17181/CERN.VU8I.D59J.

[110] CMS Tracker Group Collaboration, "The CMS Phase-1 pixel detector upgrade," JINST **16**, P02027 (2021) 10.1088/1748-0221/16/02/P02027, arXiv:2012.14304 [physics.ins-det].

[111] CMS Collaboration, "Operation and performance of the CMS silicon strip tracker with proton-proton collisions at the CERN LHC," JINST **20**, P08027 (2025) 10.1088/1748-0221/20/08/P08027, arXiv:2506.17195 [physics.ins-det].

[112] M. French, L. Jones, Q. Morrissey, A. Neviani, R. Turchetta, J. Fulcher, G. Hall, E. Noah, M. Raymond, G. Cervelli, P. Moreira, and G. Marseguerra, "Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker," in Proceedings of the 4th int. symp. on development and application of semiconductor tracking detectors, [Nucl. Instrum. Meth. A 466 (2001) 359] (2001), 10.1016/S0168-9002(01)00589-7.

[113] CMS Collaboration, *The CMS electromagnetic calorimeter project*, CMS Technical Proposal CERN-LHCC-97-33, CMS-TDR-4 (1997).

[114] CMS Collaboration, "Performance of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 13\,\text{TeV}$," JINST **19**, P09004 (2024) 10.1088/1748-0221/19/09/P09004, arXiv:2403.15518 [physics.ins-det].

[115] CMS Collaboration, *The hadron calorimeter*, CMS Technical Proposal CERN-LHCC-97-31, CMS-TDR-2 (1997).

[116] CMS Collaboration, *CMS technical design report for the Phase 1 upgrade of the hadron calorimeter*, CMS Technical Proposal CERN-LHCC-2012-015, CMS-TDR-010 (2012).

[117] CMS Collaboration, *The CMS muon project*, CMS Technical Proposal CERN-LHCC-97-032, CMS-TDR-3 (1997).

[118] CMS Collaboration, "Performance of the CMS muon trigger system in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$," JINST **16**, P07001 (2021) 10.1088/1748-0221/16/07/P07001, arXiv:2102.04790 [hep-ex].

[119] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$," JINST **13**, P06015 (2018) 10.1088/1748-0221/13/06/P06015, arXiv:1804.04528 [physics.ins-det].

[120] CMS Collaboration, “The CMS trigger system,” JINST **12**, P01020 (2017) 10.1088/1748-0221/12/01/P01020, arXiv:1609.02366 [physics.ins-det].

[121] CMS Collaboration, “Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$,” JINST **15**, P10017 (2020) 10.1088/1748-0221/15/10/P10017, arXiv:2006.10165 [hep-ex].

[122] CMS Collaboration, “Performance of the CMS high-level trigger during LHC Run 2,” JINST **19**, P11021 (2024) 10.1088/1748-0221/19/11/P11021, arXiv:2410.17038 [physics.ins-det].

[123] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector,” JINST **12**, P10003 (2017) 10.1088/1748-0221/12/10/P10003, arXiv:1706.04965 [physics.ins-det].

[124] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-k_T jet clustering algorithm,” JHEP **04**, 063 (2008) 10.1088/1126-6708/2008/04/063, arXiv:0802.1189 [hep-ph].

[125] M. Cacciari, G. P. Salam, and G. Soyez, “FastJet user manual,” Eur. Phys. J. C **72**, 1896 (2012) 10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097 [hep-ph].

[126] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV,” JINST **12**, P02014 (2017) 10.1088/1748-0221/12/02/P02014, arXiv:1607.03663 [hep-ex].

[127] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” JHEP **07**, 079 (2014) 10.1007/JHEP07(2014)079, arXiv:1405.0301 [hep-ph].

[128] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO,” JHEP **12**, 061 (2012) 10.1007/JHEP12(2012)061, arXiv:1209.6215 [hep-ph].

[129] CMS Collaboration, “Extraction and validation of a new set of CMS PYTHIA 8 tunes from underlying-event measurements,” Eur. Phys. J. C **80**, 4 (2020) 10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179 [hep-ex].

[130] K. Melnikov and F. Petriello, “Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_S^2)$,” Phys. Rev. D **74**, 114017 (2006) 10.1103/PhysRevD.74.114017, arXiv:hep-ph/0609070.

[131] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, “FEWZ 2.0: a code for hadronic Z production at next-to-next-to-leading order,” Comput. Phys. Commun. **182**, 2388 (2011) 10.1016/j.cpc.2011.06.008, arXiv:1011.3540 [hep-ph].

[132] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, “W physics at the LHC with FEWZ 2.1,” Comput. Phys. Commun. **184**, 208 (2013) 10.1016/j.cpc.2012.09.005, arXiv:1201.5896 [hep-ph].

[133] Y. Li and F. Petriello, “Combining QCD and electroweak corrections to dilepton production in FEWZ,” Phys. Rev. D **86**, 094034 (2012) 10.1103/PhysRevD.86.094034, arXiv:1208.5967 [hep-ph].

[134] NNPDF Collaboration, "Parton distributions from high-precision collider data," Eur. Phys. J. C **77**, 663 (2017) 10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428 [hep-ph].

[135] NNPDF Collaboration, "Illuminating the photon content of the proton within a global PDF analysis," SciPost Phys. **5**, 008 (2018) 10.21468/SciPostPhys.5.1.008, arXiv:1712.07053 [hep-ph].

[136] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, "ALPGEN, a generator for hard multiparton processes in hadronic collisions," JHEP **07**, 001 (2003) 10.1088/1126-6708/2003/07/001, arXiv:hep-ph/0206293.

[137] J. Bellm et al., "HERWIG 7.0/HERWIG++ 3.0 release note," Eur. Phys. J. C **76**, 196 (2016) 10.1140/epjc/s10052-016-4018-8, arXiv:1512.01178 [hep-ph].

[138] CMS Collaboration, "Development and validation of HERWIG 7 tunes from CMS underlying-event measurements," Eur. Phys. J. C **81**, 312 (2021) 10.1140/epjc/s10052-021-08949-5, arXiv:2011.03422 [hep-ex].

[139] E. Barberio and Z. Wąs, "PHOTOS—a universal Monte Carlo for QED radiative corrections: version 2.0," Comput. Phys. Commun. **79**, 291 (1994) 10.1016/0010-4655(94)90074-4.

[140] T. Melia, P. Nason, R. Röntsch, and G. Zanderighi, "W^+W^-, WZ and ZZ production in the POWHEG BOX," JHEP **11**, 078 (2011) 10.1007/JHEP11(2011)078, arXiv:1107.5051 [hep-ph].

[141] P. Nason and G. Zanderighi, "W^+W^-, WZ and ZZ production in the POWHEG-BOX-V2," Eur. Phys. J. C **74**, 2702 (2014) 10.1140/epjc/s10052-013-2702-5, arXiv:1311.1365 [hep-ph].

[142] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method," Eur. Phys. J. C **71**, 1547 (2011) 10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450 [hep-ph].

[143] S. Alioli, P. Nason, C. Oleari, and E. Re, "NLO single-top production matched with shower in POWHEG: s- and t-channel contributions," JHEP **09**, [Erratum: doi: 10.1007/JHEP02(2010)011], 111 (2009) 10.1088/1126-6708/2009/09/111, arXiv:0907.4076 [hep-ph].

[144] S. Frixione, G. Ridolfi, and P. Nason, "A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction," JHEP **09**, 126 (2007) 10.1088/1126-6708/2007/09/126, arXiv:0707.3088 [hep-ph].

[145] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations," JHEP **03**, 015 (2013) 10.1007/JHEP03(2013)015, arXiv:1212.3460 [hep-ph].

[146] N. Kidonakis, "Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-," Phys. Rev. D **82**, 054018 (2010) 10.1103/PhysRevD.82.054018, arXiv:1005.4451 [hep-ph].

[147] N. Kidonakis, "Top quark production," in Proc. Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013): Dubna, Russia, July 15–28, 2013, [DESY-PROC-2013-03] (2013), 10.3204/DESY-PROC-2013-03/Kidonakis, arXiv:1311.0283 [hep-ph].

[148] J. Campbell, T. Neumann, and Z. Sullivan, "Single-top-quark production in the t-channel at NNLO," JHEP **02**, 040 (2021) 10.1007/JHEP02(2021)040, arXiv:2012.01574 [hep-ph].

[149] M. Czakon and A. Mitov, "Top++: a program for the calculation of the top-pair cross-section at hadron colliders," Comput. Phys. Commun. **185**, 2930 (2014) 10.1016/j.cpc.2014.06.021, arXiv:1112.5675 [hep-ph].

[150] GEANT4 Collaboration, "Geant4—a simulation toolkit," Nucl. Instrum. Meth. A **506**, 250 (2003) 10.1016/S0168-9002(03)01368-8.

[151] CMS Collaboration, "Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS," Eur. Phys. J. C **81**, 800 (2021) 10.1140/epjc/s10052-021-09538-2, arXiv:2104.01927 [hep-ex].

[152] CMS Collaboration, *Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV with the CMS detector*, CMS Physics Analysis Summary CMS-PAS-LUM-20-001 (2025).

[153] CMS Collaboration, *The Phase-2 upgrade of the CMS beam radiation, instrumentation, and luminosity detectors: conceptual design*, CMS Technical Proposal CMS-TDR-19-003 (2019).

[154] CMS Collaboration, "Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV," JHEP **07**, 161 (2018) 10.1007/JHEP07(2018)161, arXiv:1802.02613 [hep-ex].

[155] A. Bodek, A. van Dyne, J. Y. Han, W. Sakumoto, and A. Strelnikov, "Extracting muon momentum scale corrections for hadron collider experiments," Eur. Phys. J. C **72**, 2194 (2012) 10.1140/epjc/s10052-012-2194-8, arXiv:1208.3710 [hep-ex].

[156] CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV," JINST **7**, P10002 (2012) 10.1088/1748-0221/7/10/P10002, arXiv:1206.4071 [physics.ins-det].

[157] Particle Data Group, R. L. Workman, et al., "Review of particle physics," Prog. Theor. Exp. Phys. **2022**, 083C01 (2022) 10.1093/ptep/ptac097.

[158] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data," JINST **15**, P09018 (2020) 10.1088/1748-0221/15/09/P09018, arXiv:2003.00503 [hep-ex].

[159] S. Schmitt, "TUnfold, an algorithm for correcting migration effects in high energy physics," JINST **7**, T10003 (2012) 10.1088/1748-0221/7/10/T10003, arXiv:1205.6201 [physics.data-an].

[160] D. Savoiu, "Triple-differential measurement of the dijet cross section at $\sqrt{s} = 13$ TeV with the CMS detector," PhD thesis (Karlsruher Institut für Technologie, 2021), 10.5445/IR/1000129970.

[161] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, "New parton distribution functions from a global analysis of quantum chromodynamics," Phys. Rev. D **93**, 033006 (2016) 10.1103/PhysRevD.93.033006, arXiv:1506.07443 [hep-ph].

[162] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, "Parton distributions in the LHC era: MMHT 2014 PDFs," Eur. Phys. J. C **75**, 204 (2015) 10.1140/epjc/s10052-015-3397-6, arXiv:1412.3989 [hep-ph].

[163] NNPDF Collaboration, "Parton distributions for the LHC run II," JHEP **04**, 040 (2015) 10.1007/JHEP04(2015)040, arXiv:1410.8849 [hep-ph].

[164] S. Carrazza, S. Forte, Z. Kassabov, J. I. Latorre, and J. Rojo, "An unbiased Hessian representation for Monte Carlo PDFs," Eur. Phys. J. C **75**, 369 (2015) 10.1140/epjc/s10052-015-3590-7, arXiv:1505.06736 [hep-ph].

[165] CMS Collaboration, "Measurement of differential cross sections for the production of a Z boson in association with jets in proton-proton collisions at $\sqrt{s} = 13$ TeV," Phys. Rev. D **108**, 052004 (2023) 10.1103/PhysRevD.108.052004, arXiv:2205.02872 [hep-ex].

[166] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin, and R. S. Thorne, "Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs," Eur. Phys. J. C **81**, 341 (2021) 10.1140/epjc/s10052-021-09057-0, arXiv:2012.04684 [hep-ph].

[167] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, "LHAPDF6: parton density access in the LHC precision era," Eur. Phys. J. C **75**, 132 (2015) 10.1140/epjc/s10052-015-3318-8, arXiv:1412.7420 [hep-ph].

[168] S. Alekhin, M. V. Garzelli, S. .-. Moch, and O. Zenaiev, "NNLO PDFs driven by top-quark data," Eur. Phys. J. C **85**, 162 (2025) 10.1140/epjc/s10052-025-13832-8, arXiv:2407.00545 [hep-ph].

[169] J. McGowan, T. Cridge, L. A. Harland-Lang, and R. S. Thorne, "Approximate N^{3}LO parton distribution functions with theoretical uncertainties: MSHT20aN^{3}LO PDFs," Eur. Phys. J. C **83**, [Erratum: doi: 10.1140/epjc/s10052-023-11451-9], 185 (2023) 10.1140/epjc/s10052-023-11236-0, arXiv:2207.04739 [hep-ph].

[170] CMS Collaboration, "Measurement of the triple-differential dijet cross section in proton-proton collisions at $\sqrt{s} = 8$ TeV and constraints on parton distribution functions," Eur. Phys. J. C **77**, 746 (2017) 10.1140/epjc/s10052-017-5286-7, arXiv:1705.02628 [hep-ex].

[171] CMS Collaboration, "Measurement of multidifferential cross sections for dijet production in proton-proton collisions at $\sqrt{s} = 13$ TeV," Eur. Phys. J. C **85**, 72 (2025) 10.1140/epjc/s10052-024-13606-8, arXiv:2312.16669 [hep-ex].

[172] C. Verstege, "Standard model W, Z (+jet) at CMS and ATLAS," in 60th rencontres de Moriond on QCD and high energy interactions (Apr. 2026), arXiv:2604.15426 [hep-ex].

[173] M. Rieger, "End-to-end analysis automation over distributed resources with luigi analysis workflows," EPJ Web Conf. **295**, 05012 (2024) 10.1051/epjconf/202429505012, arXiv:2402.17949 [physics.data-an].

[174] C. Bierlich, A. Buckley, J. Butterworth, C. H. Christensen, L. Corpe, D. Grellscheid, J. F. Grosse-Oetringhaus, C. Gutschow, P. Karczmarczyk, J. Klein, L. Lönnblad, C. S. Pollard, P. Richardson, H. Schulz, and F. Siegert, "Robust independent validation of experiment and theory: RIVET version 3," SciPost Phys. **8**, 026 (2020) 10.21468/SciPostPhys.8.2.026, arXiv:1912.05451 [hep-ph].

[175] European Commission, *Living guidelines on the responsible use of generative AI in research* (Directorate-General for Research and Innovation, European Commission, 2024).

[176] ATLAS Collaboration, "The environmental impact, carbon emissions and sustainability of computing in the ATLAS experiment," Eur. Phys. J. C **85**, 1397 (2025) 10.1140/epjc/s10052-025-14976-3, arXiv:2505.08530 [hep-ex].

[177] L. Lannelongue, J. Grealey, and M. Inouye, "Green algorithms: quantifying the carbon emissions of computation," Advanced Science **8**, 2100707 (2021) 10.1002/advs.202100707, arXiv:2007.07610 [cs.CY].