

NNPDF4.0 aN³LO PDFs with QED corrections

Andrea Barontini,^a Niccolo Laurenti^a and Juan Rojo^{b,c,*}

^a*Tif Lab, Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano, Via Celoria 16, I-20133 Milano, Italy*

^b*Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands*

^c*Department of Physics and Astronomy, Vrije Universiteit, NL-1081 HV Amsterdam, The Netherlands*

E-mail: j.rojo@vu.nl

We review recent progress in the global determination of the collinear parton distributions (PDFs) of the proton within the NNPDF framework. This progress includes NNPDF4.0 variants with QED effects and with missing higher order uncertainties (MHOUs) via the theory covariance matrix formalism, as well as PDFs based on QCD calculations at approximate N³LO (aN³LO) accuracy. We present the combination of these theoretical developments resulting into NNPDF4.0 aN³LO variants accounting for QED corrections and with a photon PDF. We compare these aN³LO QED PDFs with analogous results obtained by the MSHT group, and briefly quantify their implications at the level of representative LHC cross-sections.

*31st International Workshop on Deep Inelastic Scattering (DIS2024)
8–12 April 2024
Grenoble, France*

*Speaker

Introduction. The precise determination of the quark and gluon substructure of the proton, as encoded by the parton distribution functions (PDFs) [1, 2], is a central ingredient of the LHC physics program. This sensitivity to the proton structure is highlighted by recent high-precision measurements of Standard Model (SM) parameters at the LHC [3], where the achieved precision is limited by PDF-related uncertainties. Relevant analyses include the measurements of the strong coupling $\alpha_s(m_Z)$ from the Z p_T spectrum from ATLAS [4], of the effective leptonic mixing angle $\sin^2 \theta_{\text{eff}}^\ell$ from CMS [5, 6], as well as the measurements of the W boson mass m_W from ATLAS [7] and from the CDF experiment at the TeVatron [8]. Similar considerations apply to searches of new physics in the high-energy tails of differential distribution, as demonstrated with the Drell-Yan forward-backward asymmetry as a case study [9] as well as by simultaneous PDF and SMEFT analyses of high- p_T data from the LHC [10, 11].

In this contribution we summarise recent progress within the NNPDF global analysis framework. This progress has focused on extending the NNPDF4.0 determination in three complementary directions: QED corrections, missing higher order uncertainties (MHOUs) by means of the theory covariance matrix approach, and PDFs determined at approximate N³LO (aN³LO) accuracy in the QCD coupling expansion. Here, for the first time, we present the combination of these theoretical developments into NNPDF4.0 aN³LO sets, with and without MHOUs, accounting for QED corrections and with a photon PDF. We also quantify the impact of this combination both at the level of PDFs and of benchmark LHC cross-sections, and compare them with their counterparts obtained by the MSHT group. These combined PDF sets are made available via the LHAPDF interface [12] as well as via the NNPDF Collaboration website.¹

NNPDF4.0 with QED corrections. The original NNPDF4.0 determination [13, 14] did not include neither QED corrections to the DGLAP evolution nor a photon PDF $\gamma(x, Q^2)$. Building up on the approach deployed for NNPDF3.1QED [15, 16] and based on the LuxQED formalism [17], we have recently presented the NNPDF4.0QED analysis [18]. This variant of NNPDF4.0 is based on the new theory pipeline [19] of the NNPDF framework and assembled around the EKO [20], YADISM [21], and PINEAPPL [22] programs. The NNPDF4.0QED variants are based on QCD \otimes QED DGLAP evolution accounting for $\mathcal{O}(\alpha)$, $\mathcal{O}(\alpha\alpha_s)$ and $\mathcal{O}(\alpha^2)$ corrections implemented and benchmarked in EKO. In [18] we investigate the impact of QED effects on NNPDF4.0, compare our results to other recent PDF sets that include $\gamma(x, Q^2)$, and quantify the impact of photon-initiated processes and electroweak corrections on a variety of LHC processes, finding that they can reach the several-percent level at the high-energy tail of differential distributions.

NNPDF4.0 with MHOUs. While most PDF determinations consider only experimental uncertainties when constructing the covariance matrix used for the χ^2 loss function, the precision of available and future measurements is such that theoretical uncertainties entering the cross-section predictions cannot be neglected anymore. One of the dominant sources of theory uncertainties are those associated to the missing higher orders (MHOs) in the QCD perturbative expansion. A first systematic study of MHOUs within a global NNPDF analysis was carried out in [23, 24], based on the theory covariance matrix formalism, where MHOUs and their correlations are parameterised in terms of nuisance parameters determined from scale variations. In this way, MHOUs are accounted

¹<https://nnpdf.mi.infn.it/for-users/unpolarized-pdf-sets/>

Fit ID	Perturbative accuracy	Theory cov. mat.
NNPDF40_nnlo_as_01180_qed_mhou	NNLO _{QCD} ⊗NLO _{QED}	MHOU _{7pt}
NNPDF40_an3lo_as_01180_qed	aN ³ LO _{QCD} ⊗NLO _{QED}	IHOU+MHOU _{3pt}
NNPDF40_an3lo_as_01180_qed_mhou	aN ³ LO _{QCD} ⊗NLO _{QED}	IHOU+MHOU _{7pt}

Table 1: The settings of the three combined PDF sets presented in this contribution.

for on a comparable footing as experimental errors. This initial study was restricted to NLO, and confirmed that MHOUs are potentially important given the accuracy of modern PDF fits. Thanks to the new NNPDF theory pipeline mentioned above, we have been able to produce variants of the NNPDF4.0 NNLO global fit which include MHOUs for all input processes [25]. This analysis finds that MHOU enable an improved overall consistency of the PDF determination, with an ensuing moderate reduction of PDF uncertainties at NNLO.

NNPDF4.0 at aN³LO accuracy. Calculations of hard-scattering cross-sections at $O(\alpha_s^3)$ (relative to the Born term) in the strong coupling (N³LO) have become available for a rapidly growing set of hadron collider processes such as inclusive and differential Higgs and Drell-Yan production, see [26] for an overview, in addition to the deep-inelastic scattering coefficient functions which have been known for a long time. To obtain predictions for LHC observables with this same accuracy, partonic cross-sections need to be combined with PDFs determined also at N³LO. These N³LO PDFs should be determined by comparing to experimental data theory predictions computed also at N³LO in the QCD expansion.

With this motivation, in [27] we presented a first NNPDF fit based on approximate N³LO (aN³LO) calculations. This result required constructing an approximation to the N³LO splitting functions and anomalous dimensions reproducing all available information from fixed-order computations and from small and large- x resummation; extending the FONLL general-mass scheme for DIS structure functions [28] to N³LO; and accounting both for the remaining MHOUs and for the incomplete higher order uncertainties (IHOU) associated to the derivation of aN³LO expressions for the splitting functions and the massive DIS coefficient functions. The NNPDF4.0 aN³LO analysis of [27] demonstrates that, within the NNPDF framework, aN³LO PDFs are consistent with their NNLO counterparts, that aN³LO corrections improve the description of the global dataset, and that the perturbative convergence of Higgs and Drell-Yan inclusive cross-sections is enhanced when consistently combining aN³LO PDFs with N³LO partonic matrix elements. We also find that, as expected, MHOU on the PDFs become smaller with the increase of perturbative order.

Combining aN³LO, MHOU, and QED corrections in NNPDF4.0. Given that the three theoretical developments outlined above (aN³LO corrections, MHOU, and QED effects) are separately beneficial for the precision and/or the accuracy of the fitted PDFs, it is only natural to unify them into a combined global PDF determination, along the lines of the recent study by the MSHT group [29, 30]. Therefore, benefiting from the flexibility of the new NNPDF theory pipeline, we

present here for the first time new variants of the NNPDF4.0 global analysis combining these three theory developments in various ways. Specifically, we have produced a variant of NNPDF4.0QED NNLO with MHOUs, and then two variants of NNPDF4.0 aN³LO with QED corrections, either with or without MHOUs, see Table 1. Recall that, as discussed in [27], in the aN³LO fits we always include 3-point MHOUs in the hard-scattering cross-sections of hadronic processes to compensate for the missing N³LO K -factors.

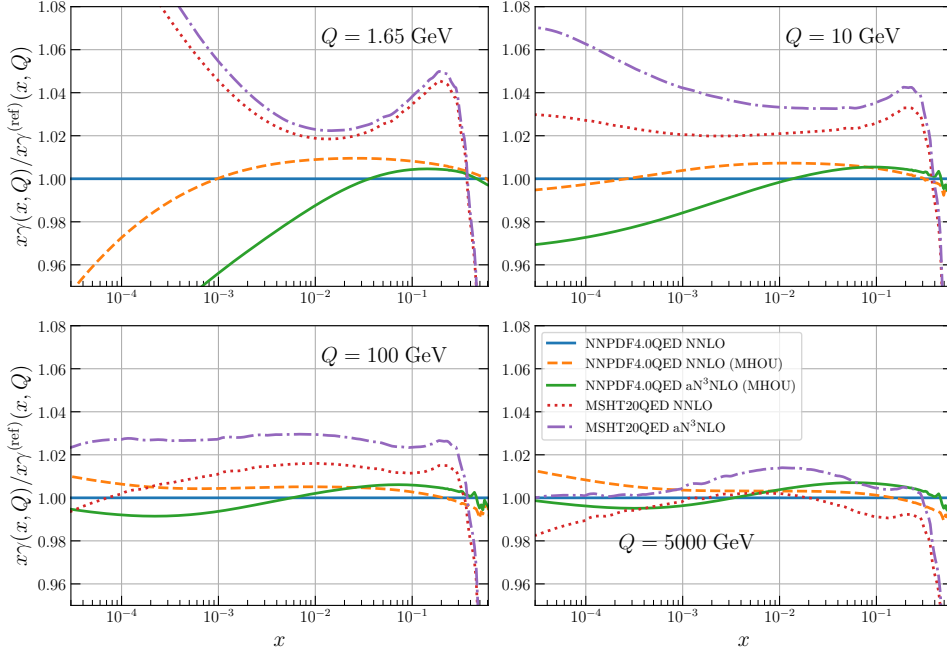


Figure 1: The central values of the photon PDF $\gamma(x, Q^2)$ in the NNPDF4.0 and MSHT20 QED sets, both for the NNLO and the aN³LO fits. For the NNPDF4.0QED fit we also include the variant with MHOUs in the covariance matrix. Results are shown normalised to the central value of NNPDF4.0QED NNLO.

We first assess the impact of MHOUs and of aN³LO corrections in the resulting photon PDF $\gamma(x, Q^2)$. Fig. 1 compares the central values of the photon PDF in the NNPDF4.0 and MSHT20 QED sets, both for the NNLO and the aN³LO fits. For the NNPDF4.0QED fit, we also include the variant with MHOUs in the covariance matrix: it is clear that MHOUs have a very small effect on the photon PDF, unless one goes to very low scales. Concerning the effect of aN³LO corrections, in NNPDF4.0 they lead to an increase of $\gamma(x, Q^2)$ as compared to the NNLO fit by up to 1% for $x \gtrsim 0.01$, and the same trend is found for their MSHT20 counterparts. For a smaller x values, aN³LO effects suppress the NNPDF4.0 photon while instead they enhance the MSHT20 photon.

We next compare in Fig. 2 the NNPDF4.0 NNLO and aN³LO fits, with and without QED corrections. We display the gluon, quark singlet, up valence, and antidown PDFs at $Q = 100$ GeV. The qualitative impact of the QED corrections on the quark and gluon PDFs is the same in the NNLO and aN³LO fits. The largest impact is associated to the gluon, where QED effects lead to an overall decrease of around 1% as a consequence of the photon PDF contributing to the momentum sum rule. For the quark PDFs, the impact of QED effects in the aN³LO fit is small, but nevertheless consistent with the results the corresponding NNLO fits.

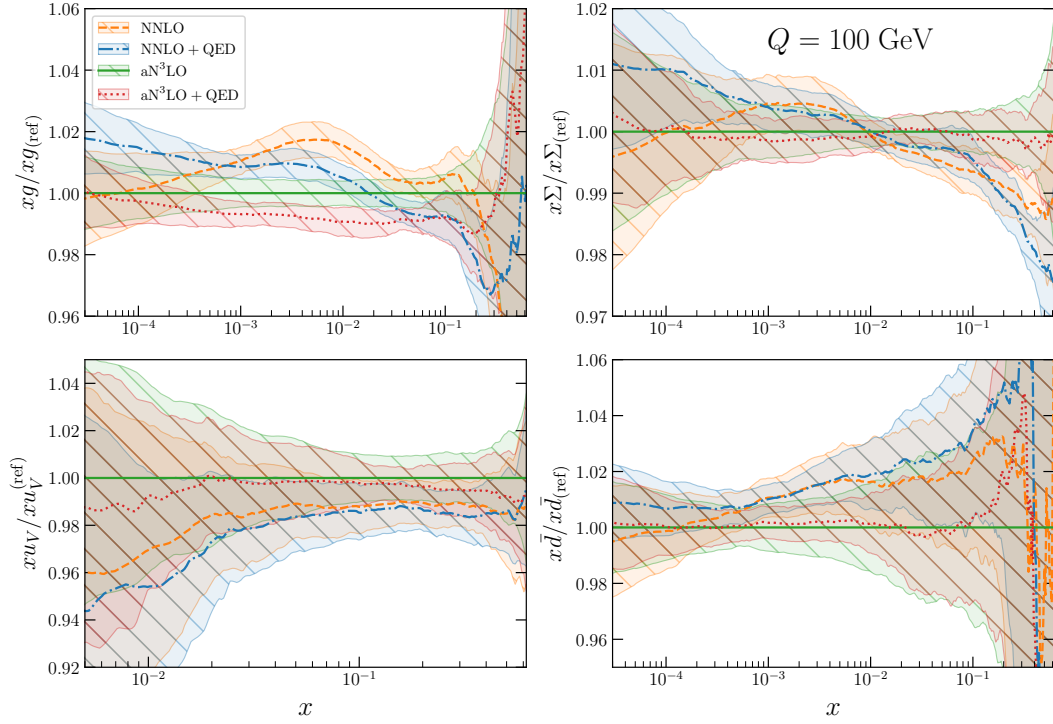


Figure 2: Comparison of the NNPDF4.0 NNLO and aN^3LO fits, with and without QED corrections, at $Q = 100$ GeV and normalised to the central value of the aN^3LO set. Bands indicate the 68% CL uncertainties.

One can also briefly assess the implications of the combined NNPDF4.0 sets listed in Table 1 for LHC phenomenology. To this end, in Fig. 3 we revisit the perturbative convergence analysis of representative inclusive cross-sections considered in [27] at the LHC with $\sqrt{s} = 13.6$ TeV. We display Higgs production in gluon fusion, in associated production with Z bosons, and W^+ production at NLO, NNLO, and N^3LO . For the N^3LO calculations, we also display predictions based on the QED variants of the NNPDF4.0 and MSHT20 aN^3LO sets. For this comparison, one observes that the qualitative impact of adding QED corrections to the aN^3LO sets is the same in the two determinations, namely a moderate reduction in the central values of the cross-sections by an amount well contained within the theory uncertainty band.

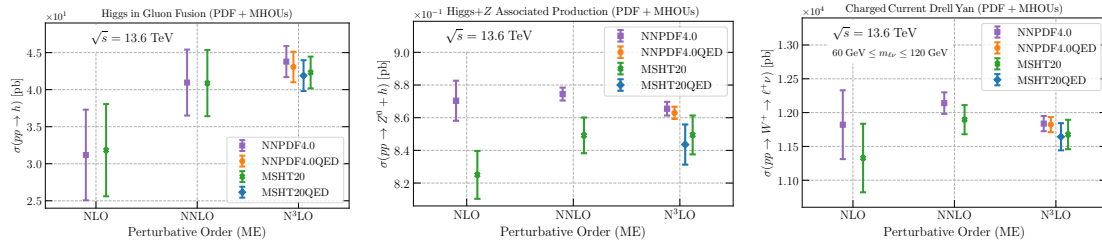


Figure 3: The perturbative convergence of Higgs production in gluon fusion and in associated production with Z bosons, and W^+ production at the LHC $\sqrt{s} = 13.6$ TeV. For the N^3LO calculations, we also display predictions based on the QED variants of the NNPDF4.0 and MSHT20 aN^3LO sets.

Additional NNPDF progress. Complementing the developments summarised above, other studies by the NNPDF Collaboration have either been presented in the past months or will be completed soon. To begin with, following our study of the total intrinsic charm (IC) content of the proton [31], we investigated the possible asymmetry between intrinsic charm and anticharm quarks. The crucial observation is that there is no perturbative mechanism in QCD that can generate such an asymmetry between charm and anticharm quarks, and thus its observation represents the ultimate smoking gun for IC in the proton. A global analysis of hard scattering data indicates a preference for a non-zero, positive intrinsic charm asymmetry [32], although its significance is smaller than for the total component. To either confirm or falsify this IC asymmetry suggested by current data, in [32] we propose two dedicated observables: Z +charm production at LHCb at the HL-LHC and F_2^c at the upcoming Electron Ion Collider (EIC) [33]. In both cases, tagging the charge of the final-state charm quark enables the calculation of asymmetries sensitive to the valence IC.

Second, Monte Carlo (MC) event generator programs modelling LHC processes often require PDFs sets with additional constraints as compared to the baseline fits, such as positive-definite PDFs down to $Q \sim 1$ GeV, smooth and physical extrapolation down to very small- and large- x as well as small Q^2 values, and the inclusion of a photon PDF enabling event generation in the presence of electroweak corrections. These so-called “MC PDFs” also enter tunes of non-perturbative QCD physics, such as the underlying event and multiple parton interactions, necessary for the description of LHC collisions. With this motivation we are developing LO, NLO, and NNLO variants of NNPDF4.0 with tailored settings for LHC event generators. These NNPDF4.0MC sets make possible a reliable description of high- p_T cross-sections as well as, subsequently to a dedicated tuning procedure, of processes sensitive to low- p_T QCD phenomena, while minimising the differences with respect to the baseline PDF sets.

Third, new methodologies are being studied with the aim to stress-test the NNPDF fitting methodology. These efforts include bypassing neural networks by parametrizing PDFs in terms of Gaussian Processes followed by Bayesian inference [34], the development of improved hyperparameter optimisation procedures based on the features of the complete PDF probability distribution (and not only of the first moments as in [13]), as well as the application of closure tests to the determination of PDF-related SM parameters such as the strong coupling constant.

Outlook. The progress along several directions in the NNPDF analysis framework outlined in this contribution contributes to the development of NNPDF4.1, an updated global fit to be released in 2025. In addition to accounting for QED, MHOU, and aN^3LO effects, the NNPDF4.1 determination will include significantly more data, in particular LHC high-precision measurements based on the complete Run II luminosity. Thanks to extending the PINEAPPL formalism to NNLO fixed-order calculations, NNPDF4.1 will achieve a global PDF determination without any reliance on the QCD K -factor approximation at NNLO. This upcoming release will also benefit from methodological improvements such as robust hyperparameter optimisation methods constructed from estimators based on the information contained on the full PDF probability distributions. Furthermore, the implementation of NNPDF4.1 in GPU architectures will result in a significant reduction of the overall computational overheads involved, enabling novel PDF studies to be carried out more efficiently.

References

- [1] J. Gao, L. Harland-Lang, and J. Rojo, “The Structure of the Proton in the LHC Precision Era,” [*Phys. Rept.* **742** \(2018\) 1–121](#), [arXiv:1709.04922 \[hep-ph\]](#).
- [2] K. Kovařík, P. M. Nadolsky, and D. E. Soper, “Hadronic structure in high-energy collisions,” [*Rev. Mod. Phys.* **92** no. 4, \(2020\) 045003](#), [arXiv:1905.06957 \[hep-ph\]](#).
- [3] S. Amoroso et al., “Snowmass 2021 Whitepaper: Proton Structure at the Precision Frontier,” [*Acta Phys. Polon. B* **53** no. 12, \(2022\) 12–A1](#), [arXiv:2203.13923 \[hep-ph\]](#).
- [4] **ATLAS** Collaboration, G. Aad et al., “A precise determination of the strong-coupling constant from the recoil of Z bosons with the ATLAS experiment at $\sqrt{s} = 8$ TeV,” [arXiv:2309.12986 \[hep-ex\]](#).
- [5] **CMS** Collaboration, A. M. Sirunyan et al., “Measurement of the weak mixing angle using the forward-backward asymmetry of Drell-Yan events in pp collisions at 8 TeV,” [*Eur. Phys. J. C* **78** no. 9, \(2018\) 701](#), [arXiv:1806.00863 \[hep-ex\]](#).
- [6] **CMS** Collaboration, “Measurement of the Drell-Yan forward-backward asymmetry and of the effective leptonic weak mixing angle using proton-proton collisions at 13 TeV,”
- [7] **ATLAS** Collaboration, G. Aad et al., “Measurement of the W-boson mass and width with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 7$ TeV,” [arXiv:2403.15085 \[hep-ex\]](#).
- [8] **CDF** Collaboration, T. Aaltonen et al., “High-precision measurement of the W boson mass with the CDF II detector,” [*Science* **376** no. 6589, \(2022\) 170–176](#).
- [9] R. D. Ball, A. Candido, S. Forte, F. Hekhorn, E. R. Nocera, J. Rojo, and C. Schwan, “Parton distributions and new physics searches: the Drell–Yan forward–backward asymmetry as a case study,” [*Eur. Phys. J. C* **82** no. 12, \(2022\) 1160](#), [arXiv:2209.08115 \[hep-ph\]](#).
- [10] A. Greljo, S. Iranipour, Z. Kassabov, M. Madigan, J. Moore, J. Rojo, M. Ubiali, and C. Voisey, “Parton distributions in the SMEFT from high-energy Drell-Yan tails,” [*JHEP* **07** \(2021\) 122](#), [arXiv:2104.02723 \[hep-ph\]](#).
- [11] Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. M. Alvarado, J. Rojo, and M. Ubiali, “The top quark legacy of the LHC Run II for PDF and SMEFT analyses,” [arXiv:2303.06159 \[hep-ph\]](#).
- [12] 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** \(2015\) 132](#), [arXiv:1412.7420 \[hep-ph\]](#).
- [13] **NNPDF** Collaboration, R. D. Ball et al., “The path to proton structure at 1% accuracy,” [*Eur. Phys. J. C* **82** no. 5, \(2022\) 428](#), [arXiv:2109.02653 \[hep-ph\]](#).

- [14] **NNPDF** Collaboration, R. D. Ball *et al.*, “An open-source machine learning framework for global analyses of parton distributions,” *Eur. Phys. J. C* **81** no. 10, (2021) 958, [arXiv:2109.02671 \[hep-ph\]](#).
- [15] **NNPDF** Collaboration, R. D. Ball *et al.*, “Parton distributions from high-precision collider data,” *Eur. Phys. J. C* **77** no. 10, (2017) 663, [arXiv:1706.00428 \[hep-ph\]](#).
- [16] **NNPDF** Collaboration, V. Bertone, S. Carrazza, N. P. Hartland, and J. Rojo, “Illuminating the photon content of the proton within a global PDF analysis,” *SciPost Phys.* **5** no. 1, (2018) 008, [arXiv:1712.07053 \[hep-ph\]](#).
- [17] A. V. Manohar, P. Nason, G. P. Salam, and G. Zanderighi, “The Photon Content of the Proton,” *JHEP* **12** (2017) 046, [arXiv:1708.01256 \[hep-ph\]](#).
- [18] **NNPDF** Collaboration, R. D. Ball *et al.*, “Photons in the proton: implications for the LHC,” *Eur. Phys. J. C* **84** no. 5, (2024) 540, [arXiv:2401.08749 \[hep-ph\]](#).
- [19] A. Barontini, A. Candido, J. M. Cruz-Martinez, F. Hekhorn, and C. Schwan, “Pineline: Industrialization of high-energy theory predictions,” *Comput. Phys. Commun.* **297** (2024) 109061, [arXiv:2302.12124 \[hep-ph\]](#).
- [20] A. Candido, F. Hekhorn, and G. Magni, “EKO: evolution kernel operators,” *Eur. Phys. J. C* **82** no. 10, (2022) 976, [arXiv:2202.02338 \[hep-ph\]](#).
- [21] A. Candido, F. Hekhorn, G. Magni, T. R. Rabemananjara, and R. Stegeman, “Yadism: Yet Another Deep-Inelastic Scattering Module,” [arXiv:2401.15187 \[hep-ph\]](#).
- [22] S. Carrazza, E. R. Nocera, C. Schwan, and M. Zaro, “PineAPPL: combining EW and QCD corrections for fast evaluation of LHC processes,” *JHEP* **12** (2020) 108, [arXiv:2008.12789 \[hep-ph\]](#).
- [23] **NNPDF** Collaboration, R. Abdul Khalek *et al.*, “Parton Distributions with Theory Uncertainties: General Formalism and First Phenomenological Studies,” *Eur. Phys. J. C* **79** no. 11, (2019) 931, [arXiv:1906.10698 \[hep-ph\]](#).
- [24] **NNPDF** Collaboration, R. Abdul Khalek *et al.*, “A first determination of parton distributions with theoretical uncertainties,” *Eur. Phys. J. C* (2019) 79:838, [arXiv:1905.04311 \[hep-ph\]](#).
- [25] **NNPDF** Collaboration, R. D. Ball *et al.*, “Determination of the theory uncertainties from missing higher orders on NNLO parton distributions with percent accuracy,” *Eur. Phys. J. C* **84** no. 5, (2024) 517, [arXiv:2401.10319 \[hep-ph\]](#).
- [26] F. Caola, W. Chen, C. Duhr, X. Liu, B. Mistlberger, F. Petriello, G. Vita, and S. Weinzierl, “The Path forward to N^3LO ,” in *Snowmass 2021*. 3, 2022. [arXiv:2203.06730 \[hep-ph\]](#).
- [27] **NNPDF** Collaboration, R. D. Ball *et al.*, “The Path to N^3LO Parton Distributions,” [arXiv:2402.18635 \[hep-ph\]](#).

- [28] S. Forte, E. Laenen, P. Nason, and J. Rojo, “Heavy quarks in deep-inelastic scattering,” [Nucl. Phys. B](#) **834** (2010) 116–162, [arXiv:1001.2312 \[hep-ph\]](#).
- [29] J. McGowan, T. Cridge, L. A. Harland-Lang, and R. S. Thorne, “Approximate N^3LO parton distribution functions with theoretical uncertainties: MSHT20a N^3LO PDFs,” [Eur. Phys. J. C](#) **83** no. 3, (2023) 185, [arXiv:2207.04739 \[hep-ph\]](#). [Erratum: [Eur.Phys.J.C](#) 83, 302 (2023)].
- [30] T. Cridge, L. A. Harland-Lang, and R. S. Thorne, “Combining QED and Approximate N^3LO QCD Corrections in a Global PDF Fit: MSHT20qed_an3lo PDFs,” [arXiv:2312.07665 \[hep-ph\]](#).
- [31] **NNPDF** Collaboration, R. D. Ball, A. Candido, J. Cruz-Martinez, S. Forte, T. Giani, F. Hekhorn, K. Kudashkin, G. Magni, and J. Rojo, “Evidence for intrinsic charm quarks in the proton,” [Nature](#) **608** no. 7923, (2022) 483–487, [arXiv:2208.08372 \[hep-ph\]](#).
- [32] **NNPDF** Collaboration, R. D. Ball, A. Candido, J. Cruz-Martinez, S. Forte, T. Giani, F. Hekhorn, G. Magni, E. R. Nocera, J. Rojo, and R. Stegeman, “Intrinsic charm quark valence distribution of the proton,” [Phys. Rev. D](#) **109** no. 9, (2024) L091501, [arXiv:2311.00743 \[hep-ph\]](#).
- [33] R. Abdul Khalek et al., “Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report,” [Nucl. Phys. A](#) **1026** (2022) 122447, [arXiv:2103.05419 \[physics.ins-det\]](#).
- [34] A. Candido, L. Del Debbio, T. Giani, and G. Petrillo, “Bayesian Inference with Gaussian Processes for the Determination of Parton Distribution Functions,” [arXiv:2404.07573 \[hep-ph\]](#).