PineAPPL: NLO EW corrections for PDF processes

Christopher Schwan⋆

TIF Lab and Dipartimento di Fisica, Università degli Studi di Milano and INFN, Milan, Italy

⋆ christopher.schwan@mi.infn.it

Abstract

Modern parton distribution function (PDF) determinations either neglect higher-order corrections in the electroweak (EW) coupling or implement them approximately for a subset of PDF processes. We present a new tool, PINEAPPL, which supports perturbative correction in arbitrary powers of the EW and strong coupling constant, and stores theoretical predictions independently of the used PDFs in so-called interpolation grids. These are the foundation of any PDF determination, and will allow us to perform a PDF fit taking into account EW corrections for all processes consistently. Apart from PDF determinations PINEAPPL can also be used in precision phenomenology to study the impact of different choices of PDF sets, without rerunning the time-consuming computation. As an application of this tool we show Drell–Yan (DY) lepton-pair production were larger effects of NLO EW corrections can be seen that will potentially influence the PDF fit and comment about which observables are particularly suited and NLO EW PDF fit and which are not.

1 Introduction

PDF determinations (for a recent review see e.g. Ref. [1]) typically neglect the effects of higher-order effects in the EW coupling constant, α, and only include corrections in the strong coupling constant α_s, up to next-to-next-to-leading order (NNLO) in recent PDF fits; a notable exception is the MSHT20 [2] PDF set, which however uses K-factors that approximately implement the effect of EW corrections for a subset of PDF processes. With uncertainties of about 5 % — of the PDFs themselves, and by convolution also the cross-sections — neglecting/approximating NLO EW effects is certainly justified, given that the size of EW corrections are often of the size of 1 % or even smaller. However, it is foreseeable that uncertainties of PDFs will be much smaller in the future, so that the task of including EW corrections in a PDF fit becomes increasingly important. Apart from this more general statement we note the

following specific observations: 1) in phase-space regions where many events are recorded from the experiments the PDFs are typically much more precise, so that even small EW correction can become important; 2) for more extreme phase-space regions, for example large transverse momentum of the Z boson, EW corrections can be of the order of -20% data are still measured. At the LHC for run I and II the datapoints typically have large experimental uncertainties to accommodate potentially large theoretical corrections; however, this is expected to change in run III. Finally we note that including EW corrections in a PDF fit requires a more consistent treatment of data, which in practice means not subtracting collinear photons, which is typically done for a large number of gauge-boson production datasets at the LHC, for which this procedure can have a sizeable impact.

2 PineAPPL: EW corrections and interpolation grids

In this article we use the term "EW corrections" loosely and let it refer to any perturbative order that is not purely of a strong origin. For PDF processes this includes, for example, the $\mathcal{O}(\alpha^3)$ for DY (with decays), which is a purely EW correction, but for jet production we also include the additional leading orders at $\mathcal{O}(\alpha_s\alpha)$ and $\mathcal{O}(\alpha^2)$, and the NLOs at $\mathcal{O}(\alpha_s^2\alpha)$, $\mathcal{O}(\alpha_s\alpha^2)$, which we sometimes call combined QCD–EW corrections, and finally the purely EW correction of $\mathcal{O}(\alpha^3)$.

Including these orders into a PDF fit requires the calculation of interpolation grids, which are essentially cross sections of processes differential with respect to the parton momentum fractions x_1 and x_2, their initial-state flavours a and b and finally the choice of the scale Q^2, in separate powers of $\alpha_s(Q^2)$ and α:

$$\frac{\mathrm{d}\sigma_{ab}}{\mathrm{d}\mathcal{O}}(x_1,x_2,Q^2) = \sum_{k,l,m,n} \alpha_s^k(Q^2)\,\alpha^l \log^m(\xi_R^2)\log^n(\xi_F^2) W_{ab}^{(k,l,m,n)}(x_1,x_2,Q^2,\mathcal{O}). \quad (1)$$

The scale-variation grids, which have $m > 0$ or $n > 0$, are important when the user wants to vary the renormalisation, μ_R, and factorisations scale, μ_F, around the central scale choice Q^2 using the ratios $\xi_R^2 = \mu_R^2/Q^2$ and $\xi_F^2 = \mu_F^2/Q^2$.

These grids are the preferred representation of theory predictions for PDF fits, but outside of PDF fits they are also useful to quickly study the impact of different PDF sets to observables. They can be generated once, and after their generation convolutions with arbitrary PDFs are performed in a matter of seconds.

An alternative to interpolation grids are K-factors, in our case EW K-factors,

$$K_{EW} = \frac{\mathrm{d}\sigma_{EW}/\mathrm{d}\mathcal{O}}{\mathrm{d}\sigma_{QCD}/\mathrm{d}\mathcal{O}}, \quad (2)$$

which are the ratios of cross sections with EW (and NNLO QCD) corrections to the cross section without EW corrections, only including higher orders in QCD. Eq. (1) is then approximated using

$$\frac{\mathrm{d}\sigma_{ab}}{\mathrm{d}\mathcal{O}}(x_1,x_2,Q^2) \approx K_{EW} \sum_k \alpha_s^k(Q^2)\,\alpha^{l_{min}} W_{ab}^{(k,l_{min})}(x_1,x_2,Q^2,\mathcal{O}), \quad (3)$$

where l_{min} is the smallest power of α for the process at LO. The advantage of the method is that an NNLO QCD PDF fit can be easily upgraded to approximately include NLO EW effects (see for example Refs. [2–4]), however, there are several pitfalls stemming from the fact that the K-factor is applied after the PDF convolution (i.e. is a hadronic K-factor and not a partonic one): it is blind to the initial-state flavours, the partonic momentum fractions and the scale. For example, it is possible that at NLO EW a new partonic channel opens up, typically $q\gamma$, which is

positive and cancels negative corrections of a virtual correction. This results in small K-factors, which leave the PDF fit unaffected, and therefore underestimating the EW corrections. To be on the safe side, especially for the accuracy mentioned above, we therefore take EW corrections into account using interpolation grids instead of K-factors.

To this end we developed PINEAPPL [5], which is an interpolation library [6], similar to AP-PLGRID [7] and FASTNLO [8–10], but it supports arbitrary perturbative predictions in powers of α_s and α as shown in Eq. (1). While extending the two previously mentioned interpolation libraries would have been possible in principle, we found that the increased memory requirements for supporting the additional orders in α made a rewrite from scratch inevitable in practice.

PINEAPPL is the technical solution that will allow us to generate the interpolation grids needed for PDF fits including EW corrections. It consists of a library, written in Rust [11], which we interfaced to MADGRAPH5_AMC@NLO [12,13], similarly to AMCFAST [14], which it replaces for MADGRAPH5_AMC@NLO version 3 and higher. In addition to a library a command-line program `pineappl` is provided, which allows for quick and easy convolutions of the interpolation grids with arbitrary PDF sets through LHAPDF [15], evaluations of the PDF uncertainties, scale variations and plotting. An example output of the plotting command is shown in Fig. 1. See also Sec. 5 for instructions on how to produce a similar plot using the toolchain described above.

3 CMS Drell–Yan lepton-pair production at 13 TeV

As an example of the application of PINEAPPL together with MADGRAPH5_AMC@NLO v3.1 we show the theoretical prediction for the CMS measurement [16] of DY lepton-pair production at 13 TeV in Fig. 1. This measurement shows the invariant mass $M_{\ell\bar{\ell}}$ of the lepton-pair very finely resolved around the Z-boson peak, with bins as small as 4 GeV. As a result of this, QED effects produce very large higher-order corrections, which are part of the NLO EW corrections and can be seen in the second panel of Fig. 1. This effect is typically not so pronounced in the ATLAS and CMS measurements at 7 TeV and 8 TeV, where the bins are larger and are chosen to be symmetric around the peak, so that the corrections mostly cancel themselves out.

The second panel also shows the uncertainty band for the NLO QCD+EW correction (blue) and for the NLO QCD correction (red) only. Note that the large corrections below the Z-boson peak also enhance the uncertainties significantly (the largest difference is found for the range from 76 GeV to 81 GeV where the uncertainty in the positive direction grows from 3.5 % to 6.4 %, which is an increase of 83 %). As a consequence including this dataset into a PDF fit with missing-higher order uncertainties (see e.g. Ref. [17, 18], but without this dataset) without NLO EW corrections and Born-level observables would seriously underestimate the theory uncertainties.

The third and fourth panel show the relative PDF uncertainties for NNPDF4.0 (preliminary fit, see Ref. [19]) against NNPDF3.1 [20], MSHT20 [2], CT18 [21] and ABMP16 [22], for which we chose the PDF sets with NNLO QCD accuracy and $\alpha_s(M_Z^2) = 0.118$. The pull is calculated as

$$\text{Pull} = \frac{X_2 - X_1}{\sqrt{(\delta X_1)^2 + (\delta X_2)^2}}, \tag{4}$$

where $X_1 = \mathrm{d}\sigma/\mathrm{d}M_{\ell\bar{\ell}}$ is the differential cross section calculated with the NNPDF4.0 PDF set and X_2 the same for any of the other sets, and δX_1 and δX_2 are the corresponding PDF uncertainties. The pull of each PDF with respect to NNPDF4.0 is below 2σ for most of the PDFs and for most of the invariant mass, however, around the Z-boson peak the pull is as large as 3σ, also due to the very small uncertainties of the NNPDF4.0 PDF set. In this region we expect the

inclusion of theoretical uncertainties to produce larger PDF uncertainties in NNPDF4.0, and in turn smaller pulls. However, given that the pull between NNPDF4.0 and 3.1 are always below 2σ shows that the two PDF sets are consistent with each other.

4 Lepton observables for different theoretical predictions

As can be seen in the second panel of Fig. 1 the NLO EW corrections are very large and distort the shape of the $M_{\ell\bar{\ell}}$ distribution around the Z-boson peak. This effect is of course well known to come from soft photons, and therefore measurements of leptonic observables are typically published at *Born-level*, as shown in Fig. 2. Born-level leptons are a measurement combined with a theoretical prediction that removes the effect from soft photons (typically obtained with PHOTOS [23–25]). The advantage of using leptons "before they radiate" is that the large corrections from soft photons are taken care of and that these observables can be compared to pure QCD predictions with good accuracy, which is what is done in most global PDF fits. However, if we want to take into account weak corrections, which are important e.g. in DY in the tail of the $M_{\ell\bar{\ell}}$ distribution, Born-level leptons pose the problem that some QED effects have been subtracted from the data, but the rest of the EW corrections have not been. A comparison of data against theory with full EW corrections is therefore not possible without having some sort of double-counting, and therefore we have to use different lepton-observables, the most convenient which are *dressed-lepton observables* (see Fig. 2). Dressed-lepton observable combine the momenta of leptons with the momenta of photon around a certain radius, where typically $\Delta R = 0.1$ is chosen. For the observables both measurements and predictions can be easily done.

5 Conclusion

We discussed the aim of performing a PDF fit with the consistent inclusion of EW corrections, for which we wrote the interpolation library PINEAPPL. Together with MADGRAPH5_AMC@NLO, to which it is interfaced, we are able to produce predictions for all PDF processes including the EW corrections, and we discussed an application for the CMS DY lepton-pair production, showcasing features of the EW corrections: 1) large and shape-changing QED corrections, 2) enlarged theory uncertainties and 3) large and negative weak corrections in the tails of the invariant mass distribution. These predictions, when compared to measurements require an appropriate definition of lepton observables — dressed leptons — which we discussed is different from the ones used for purely NNLO QCD predictions.

Finally, we note that we published a script [26] can be used to produce a plot similar to the one shown in Fig. 1 and which gives an in-depth showcase of how the toolchain is used and what it can accomplish. Before running the script MADGRAPH5_AMC@NLO v3.1 or higher (see https://launchpad.net/mg5amcnlo) and of PineAPPL v0.4.1 or higher (see https://github.com/N3PDF/pineappl) must be installed.

Funding information C.S. is supported by the European Research Council under the European Union's Horizon 2020 research and innovation Programme (grant agreement no. 740006).
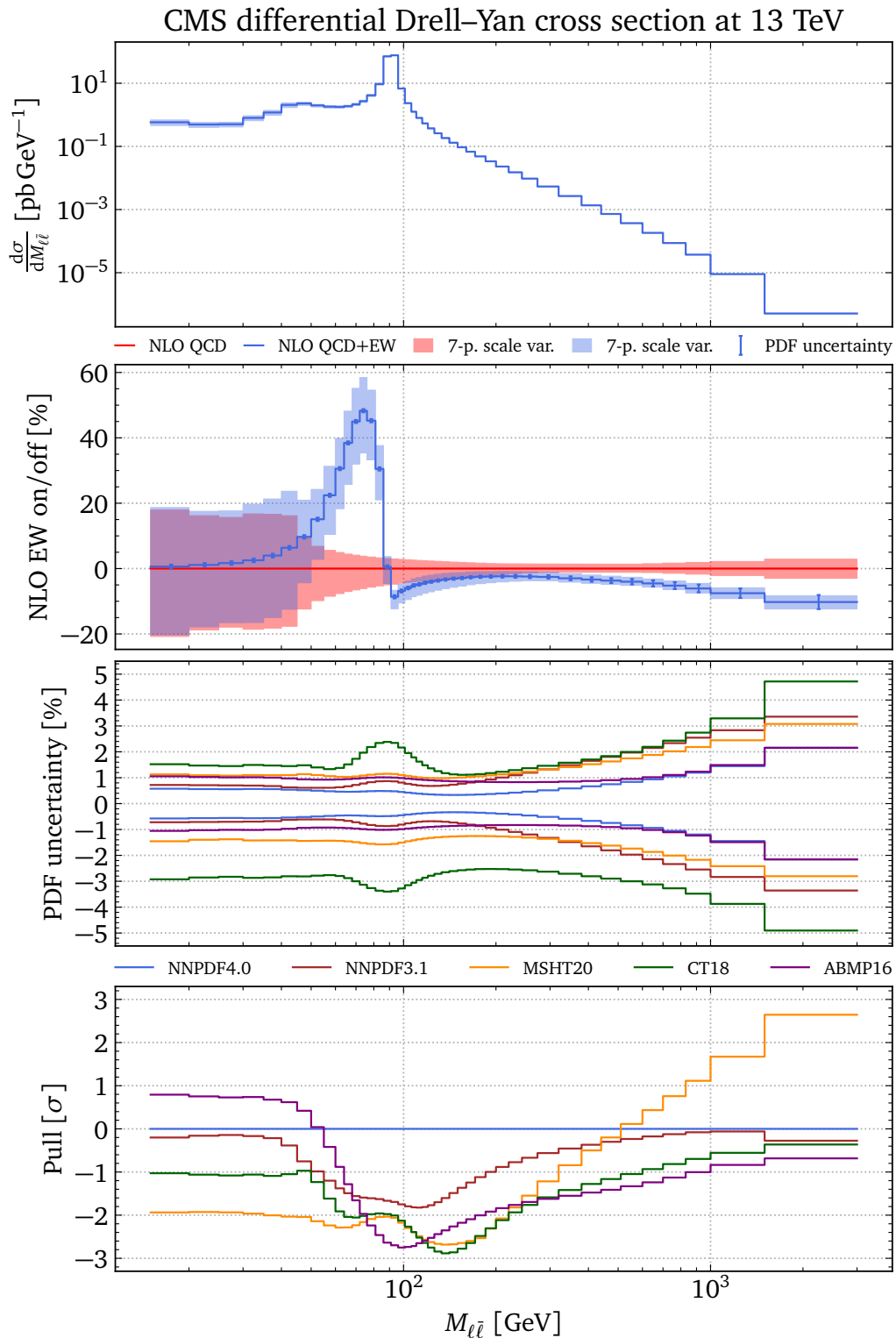
Figure 1: Theoretical predictions for the CMS DY lepton-pair production at 13 TeV. Top panel: absolute prediction for the cross section differentially in the invariant mass of the lepton pair. Second panel: NLO EW+QCD predictions relative to NLO QCD only. Third panel: PDF uncertainties for the four global PDF sets NNPDF4.0 (preliminary fit), MSHT20, CT18 and ABMP16. Bottom panel: Pull between NNPDF4.0 and the remaining three PDF sets.
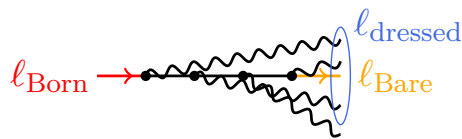
Figure 2: Different lepton observables for which data are published: 1) Born-level leptons, and 2) bare leptons, which are leptons before and after radiation, respectively. Finally, 3) dressed-leptons, whose momenta include the surrounding photons.

References

[1] J. J. Ethier and E. R. Nocera, *Parton Distributions in Nucleons and Nuclei*, Annu. Rev. Nucl. Part. Sci. **70**, 43 (2020), doi:10.1146/annurev-nucl-011720-042725.

[2] 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), doi:10.1140/epjc/s10052-021-09057-0.

[3] R. Boughezal, A. Guffanti, F. Petriello and M. Ubiali, *The impact of the LHC Z-boson transverse momentum data on PDF determinations*, J. High Energy Phys. **07**, 130 (2017), doi:10.1007/JHEP07(2017)130.

[4] R. Abdul Khalek *et al.*, *Phenomenology of NNLO jet production at the LHC and its impact on parton distributions*, Eur. Phys. J. C **80**(8), 797 (2020), doi:10.1140/epjc/s10052-020-8328-5, 2005.11327.

[5] S. Carrazza, E. R. Nocera, C. Schwan and M. Zaro, *PineAPPL: combining EW and QCD corrections for fast evaluation of LHC processes*, J. High Energy Phys. **12**, 108 (2020), doi:10.1007/JHEP12(2020)108.

[6] C. Schwan and S. Carrazza, *N3pdf/pineappl: v0.4.1*, Zenodo (2021), doi:10.5281/zenodo.4636076 .

[7] T. Carli, D. Clements, A. Cooper-Sarkar, C. Gwenlan, G. P. Salam, F. Siegert, P. Starovoitov and M. Sutton, *A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: the APPLGRID project*, Eur. Phys. J. C **66**, 503 (2010), doi:10.1140/epjc/s10052-010-1255-0.

[8] T. Kluge, K. Rabbertz and M. Wobisch, *FastNLO: Fast pQCD calculations for PDF fits*, In 14th International Workshop on Deep Inelastic Scattering (2006), doi:10.1142/9789812706706_0110.

[9] M. Wobisch, D. Britzger, T. Kluge, K. Rabbertz and F. Stober, *Theory-Data Comparisons for Jet Measurements in Hadron-Induced Processes*, arXiv:1109.1310.

[10] D. Britzger, K. Rabbertz, F. Stober and M. Wobisch, *New features in version 2 of the fastNLO project*, In 20th International Workshop on Deep-Inelastic Scattering and Related Subjects (2012), doi:10.3204/DESY-PROC-2012-02/165.

[11] N. D. Matsakis and F. S. Klock, *The rust language*, In Proceedings of the 2014 ACM SIGAda Annual Conference on High Integrity Language Technology, p. 103, Association for Computing Machinery, New York, NY, USA, ISBN 9781450332170 (2014), doi:10.1145/2663171.2663188.

[12] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, J. High Energy Phys. **07**, 079 (2014), doi:10.1007/JHEP07(2014)079.

[13] R. Frederix, S. Frixione, V. Hirschi, D. Pagani, H.-S. Shao and M. Zaro, *The automation of next-to-leading order electroweak calculations*, J. High Energy Phys. **07**, 185 (2018), doi:10.1007/JHEP07(2018)185.

[14] V. Bertone, R. Frederix, S. Frixione, J. Rojo and M. Sutton, *aMCfast: automation of fast NLO computations for PDF fits*, J. High Energy Phys. **08**, 166 (2014), doi:10.1007/JHEP08(2014)166.

[15] 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), doi:10.1140/epjc/s10052-015-3318-8.

[16] A. M. Sirunyan et al., *Measurement of the differential Drell-Yan cross section in proton-proton collisions at $\sqrt{s}$ = 13 TeV*, J. High Energy Phys. **12**, 059 (2019), doi:10.1007/JHEP12(2019)059.

[17] R. Abdul Khalek et al., *A first determination of parton distributions with theoretical uncertainties*, Eur. Phys. J. C **79**, 838 (2019), doi:10.1140/epjc/s10052-019-7364-5.

[18] R. Abdul Khalek et al., *Parton distributions with theory uncertainties: general formalism and first phenomenological studies*, Eur. Phys. J. C **79**, 931 (2019), doi:10.1140/epjc/s10052-019-7401-4.

[19] J. Rojo, *Progress in the NNPDF global analyses of proton structure*, arXiv:2104.09174.

[20] R. D. Ball et al., *Parton distributions from high-precision collider data*, Eur. Phys. J. C **77**, 663 (2017), doi:10.1140/epjc/s10052-017-5199-5.

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

[22] S. Alekhin, J. Blümlein, S. Moch and R. Plačakytė, *Parton distribution functions, α_s, and heavy-quark masses for LHC Run II*, Phys. Rev. D **96**, 014011 (2017), doi:10.1103/PhysRevD.96.014011.

[23] P. Golonka and Z. Was, *PHOTOS Monte Carlo: a precision tool for QED correctionsin Z and W decays*, Eur. Phys. J. C **45**, 97 (2006), doi:10.1140/epjc/s2005-02396-4.

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

[25] E. Barberio, B. van Eijk and Z. Was, *Photos — a universal Monte Carlo for QED radiative corrections in decays*, Comput. Phys. Commun. **66**, 115 (1991), doi:10.1016/0010-4655(91)90012-A.

[26] *Example run script for PineAPPL and Madgraph5_aMC@NLO v3*, https://github.com/N3PDF/pineappl/blob/master/examples/mg5amcnlo/DY_14TEV.sh.