THE EUROPEAN
PHYSICAL JOURNAL C

Regular Article - Theoretical Physics

Photons in the proton: implications for the LHC

NNPDF Collaboration

Richard D. Ball[1], Andrea Barontini[2], Alessandro Candido[2,3], Stefano Carrazza[2], Juan Cruz-Martinez[3],
Luigi Del Debbio[1], Stefano Forte[2,a], Tommaso Giani[4,5], Felix Hekhorn[2,6,7], Zahari Kassabov[8], Niccolò Laurenti[2],
Giacomo Magni[4,5], Emanuele R. Nocera[9], Tanjona R. Rabemananjara[4,5], Juan Rojo[4,5], Christopher Schwan[10],
Roy Stegeman[1], Maria Ubiali[8]

[1] The Higgs Centre for Theoretical Physics, University of Edinburgh, JCMB, KB, Mayfield Rd, Edinburgh EH9 3JZ, Scotland
[2] Tif Lab, Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano, Via Celoria 16, 20133 Milan, Italy
[3] CERN, Theoretical Physics Department, 1211 Geneva 23, Switzerland
[4] Department of Physics and Astronomy, Vrije Universiteit, 1081 HV Amsterdam, Netherlands
[5] Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands
[6] Department of Physics, University of Jyvaskyla, P.O. Box 35, 40014 Jyvaskyla, Finland
[7] Helsinki Institute of Physics, University of Helsinki, P.O. Box 64, 00014 Helsinki, Finland
[8] DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK
[9] Dipartimento di Fisica, Università degli Studi di Torino and INFN, Sezione di Torino, Via Pietro Giuria 1, 10125 Turin, Italy
[10] Institut für Theoretische Physik und Astrophysik, Universität Würzburg, 97074 Würzburg, Germany

Abstract We construct a set of parton distribution functions (PDFs), based on the recent NNPDF4.0 PDF set, that also include a photon PDF. The photon PDF is constructed using the LuxQED formalism, while QED evolution accounting for $\mathcal{O}(\alpha)$, $\mathcal{O}(\alpha\alpha_s)$, and $\mathcal{O}(\alpha^2)$ corrections is implemented and benchmarked by means of the EKO code. We investigate the impact of QED effects on NNPDF4.0, and compare our results both to our previous NNPDF3.1QED PDF set and to other recent PDF sets that include the photon. We assess the impact of photon-initiated processes and electroweak corrections on a variety of representative LHC processes, and find that they can reach the 5% level in vector boson pair production at large invariant mass.

Contents

1 Introduction

High precision physics at the LHC, especially in its high-luminosity era [1–3], will demand the inclusion of electroweak corrections in the computation of theoretical predictions for hard processes [4]. This requires an extension of the set of proton parton distribution functions (PDFs). In particular a photon PDF has to be provided, and evolution equations need to be supplemented with QED splitting functions. The photon PDF enables the inclusion of photon-initiated processes, which typically are enhanced in the high-mass

[a] e-mail: stefano.forte@mi.infn.it (corresponding author)

Springer

and large transverse-momentum tails of the distributions. In principle, at high-enough scales, proton PDFs should also include PDFs for leptons [5], gauge bosons [6] and indeed for the full set of 52 standard model fields [7], or even new hypothetical particles such as dark photons [8]. Even at the LHC lepton PDFs are needed in searches for leptoquarks [9] or exotic resonances that couple to leptons [10]. However, in practice, the main requirement of current precision physics at the LHC is the availability of PDF sets that also include a photon PDF, that mixes upon combined QED×QCD evolution with the standard quark and gluon PDFs. We will henceforth refer to such PDF sets as "QED PDFs".

Initial attempts at the construction of QED PDF sets relied on models for the photon PDF at the initial evolution scale [11,12]. A first data-driven determination of QED PDFs, based on a fit using the NNPDF2.3 methodology [13,14], resulted in a photon PDF with large uncertainties. The fact that determining the photon PDF from the data yields a result affected by large uncertainties was more recently confirmed in determinations based on fitting to the data PDFs with a fixed functional form within the xFITTER methodology [15] applied to high-mass ATLAS Drell–Yan distributions [16].

A breakthrough in the determination of QED PDFs was achieved in 2016 in Refs. [17,18] (see also related results in Refs. [19,20]), where it was shown that the photon PDF can be computed perturbatively in QED, given as input the proton structure functions at all scales, from the elastic ($Q^2 \to 0$) to the deep-inelastic ($Q^2 \to \infty$) regimes – the so-called LuxQED method. Since then, this LuxQED framework has been the basis of all QED PDF sets [21–25].

While the effect of the inclusion of the photon PDF on other PDFs is small, it is not negligible within current uncertainties. For instance, the photon typically carries a fraction of the proton momentum which is about two orders of magnitude smaller than that of the gluon, so the corresponding depletion of the gluon momentum fraction is relevant at the percent level, which is comparable to the size of the current uncertainty on the gluon PDF. Precision calculations of LHC processes with electroweak effects thus require a consistent global QED PDF determination.

With this motivation, we construct a QED PDF set based on the recent NNPDF4.0 PDF determination [26,27], including QED corrections to parton evolution up to $\mathcal{O}(\alpha\alpha_s)$ and $\mathcal{O}(\alpha^2)$. We follow closely the methodology developed for the NNPDF3.1QED PDFs in Ref. [21], based on an iterative procedure. Namely, we evaluate the photon PDF at some fixed scale using the LuxQED formula and the structure function computed from an existing PDF set, now NNPDF4.0. We evolve it together with the other PDFs to the initial parametrization scale $Q_0 = 1.65$ GeV. We then re-determine the quark and gluon PDFs while also including this photon PDF as a boundary condition to the QCD×QED evolution.

We compute the photon PDF again using LuxQED, and we iterate until convergence.

All results presented in this paper are obtained using a new theory implementation, now adopted by default in the NNPDF public code [27]. In particular, the APFEL [28] and APFELGRID [29] codes have been replaced by a suite of newly developed tools including the evolution equations solver EKO [30], the DIS structure functions calculator YADISM [31], and the interpolator of hard-scattering cross-sections PINEAPPL [32,33]. Taken together, they provide a pipeline for the efficient automatization of the computation of theory predictions [34]. An integral component of this theory pipeline is the use of interpolation grids in the format provided by PINEAPPL [32,33], which can be interfaced to Monte Carlo generators. Among these is MG5_AMC@NLO [35], which automates NLO QCD+EW calculations for a wide range of LHC processes where knowledge of QED PDFs is crucial.

The outline of this paper is as follows. First, Sect. 2 reviews the theoretical framework underlying the NNPDF4.0QED determination and in particular the implementation of QED evolution in EKO. Then the NNPDF4.0QED PDFs are presented in Sect. 3, where they are compared to the previous NNPDF3.1QED PDF set, and to other recent QED PDF sets. Implications for LHC phenomenology are studied in Sect. 4 by means of the PINEAPPL interface to MG5_AMC@NLO. Conclusions and an outline of future developments are finally presented in Sect. 5. Because this paper is, as mentioned, the first to make use of a new theory pipeline, and specifically its implementation of combined QCD×QED evolution, a set of benchmarks is collected in two appendices. Specifically, in Appendix A we benchmark NNPDF4.0 PDFs determined using the old and new pipeline and show that they are indistinguishable; in Appendix B the new implementation of joint QCD and QED evolution of PDFs in the EKO code is discussed and benchmarked against the previous implementation in the APFEL code.

2 Evolution and determination of QED PDFs

In this section, we discuss the structure of combined QED×QCD evolution and we briefly review the methodology used to determine QED PDFs. This methodology is based on the LuxQED formalism [17,18], and was developed for the construction of the NNPDF3.1QED PDF set [21].

2.1 QCD×QED evolution equations and basis choice

The scale dependence of PDFs is determined by evolution equations of the form

$$\mu^2 \frac{df_i(N,\mu^2)}{d\mu^2}$$

$$= -\sum_j \gamma_{ij}\left(N, a_s(\mu^2), a_{em}(\mu^2)\right) f_j(N,\mu^2), \qquad (2.1)$$

where $f_i(N,\mu^2) = \int_0^1 dx\, x^{N-1} f_i(x,\mu^2)$ is the Mellin transform of the i-th PDF. The anomalous dimensions $\gamma_{ij}(N, a_s, a_{em})$ are determined as a simultaneous perturbative expansion in the strong coupling $a_s = \alpha_s/(4\pi)$ and in the electromagnetic coupling $a_{em} = \alpha/(4\pi)$:

$$\gamma_{ij}(N, a_s, a_{em}) = \sum_{\substack{n,m=0 \\ (n,m)\neq(0,0)}}^{\infty} a_s^n a_{em}^m \gamma_{ij}^{(n,m)}(N)$$

$$= a_s \gamma_{ij}^{(1,0)}(N) + a_s^2 \gamma_{ij}^{(2,0)}(N) + a_s^3 \gamma_{ij}^{(3,0)}(N)$$

$$+ a_{em}\gamma_{ij}^{(0,1)}(N) + a_{em}^2 \gamma_{ij}^{(0,2)}(N) + a_s a_{em}\gamma_{ij}^{(1,1)}(N) + \cdots . \qquad (2.2)$$

In this work we include pure QCD corrections up to NNLO, $\gamma^{(1,0)}$, $\gamma^{(2,0)}$, and $\gamma^{(3,0)}$ [36,37]; the pure NLO QED corrections $\gamma^{(0,1)}$ and $\gamma^{(0,2)}$ [38]; and the leading mixed correction $\gamma^{(1,1)}$ [39].

The scale dependence of the strong and electromagnetic couplings is in turn determined by coupled renormalization group equations of the form

$$\mu^2 \frac{da_s}{d\mu^2} = \beta_{\text{QCD}}(a_s, a_{em})$$

$$= -a_s^2\left(\beta_{\text{QCD}}^{(2,0)} + a_s\beta_{\text{QCD}}^{(3,0)} + a_{em}\beta_{\text{QCD}}^{(2,1)} + a_s^2\beta_{\text{QCD}}^{(4,0)}\cdots\right), \qquad (2.3)$$

$$\mu^2 \frac{da_{em}}{d\mu^2} = \beta_{\text{QED}}(a_s, a_{em})$$

$$= -a_{em}^2\left(\beta_{\text{QED}}^{(0,2)} + a_{em}\beta_{\text{QED}}^{(0,3)} + a_s\beta_{\text{QED}}^{(1,2)} + \cdots\right), \qquad (2.4)$$

in terms of the coefficients of the corresponding QCD and QED beta functions. Consistently with the treatment of evolution equations, we include pure QCD contributions up to NNLO, namely $\beta_{\text{QCD}}^{(2,0)}$, $\beta_{\text{QCD}}^{(3,0)}$, and $\beta_{\text{QCD}}^{(4,0)}$; pure QED up to NLO, namely $\beta_{\text{QED}}^{(0,2)}$ and $\beta_{\text{QED}}^{(0,3)}$; and the leading mixed terms $\beta_{\text{QCD}}^{(2,1)}$ and $\beta_{\text{QED}}^{(1,2)}$ [40]. We adopt the $\overline{\text{MS}}$ scheme. Schemes in which the electroweak coupling does not run, such as the G_μ scheme, are commonly used in the computation of electroweak corrections, but $\overline{\text{MS}}$ is more convenient when considering combined QCD and QED corrections [41,42]. Equations (2.1–2.4) must then be simultaneously solved with a common scale μ.

The solution is most efficiently obtained in a maximally decoupled basis in quark flavor space. This requires adopting a suitable combination of quark and antiquark flavors such that the sum over j in Eq. (2.1) contains the smallest possible number of entries. In the case of QCD-only evolution,

this is achieved in the so-called evolution basis, in which one separates off the singlet combination $\Sigma = \sum_i(q_i + \bar{q}_i)$, which mixes with the gluon, and then one constructs nonsinglet combinations of individual C-even (sea-like) and C-odd (valence-like) $q_i^\pm = q_i \pm \bar{q}_i$ quark and antiquark flavors, each of which evolves independently.

Because the photon couples differently to up-like $u_k = \{u, c, t\}$ and down-like $d_k = \{d, s, b\}$ quarks, a different basis choice is necessary when also including QED. To this end, given n_f active quark flavors, we split them into n_u up-like and n_d down-like flavors, such that $n_f = n_u + n_d$, and define the four combinations

$$\Sigma_u = \sum_{k=1}^{n_u} u_k^+, \quad \Sigma_d = \sum_{k=1}^{n_d} d_k^+,$$

$$V_u = \sum_{k=1}^{n_u} u_k^-, \quad V_d = \sum_{k=1}^{n_d} d_k^-. \qquad (2.5)$$

Also, we separate off the QCD contributions to the anomalous dimensions, by rewriting the perturbative expansion of the anomalous dimensions, Eq. (2.2), as

$$\gamma_{ij}(N, a_s, a_{em}) = \gamma_{ij}(N, a_s) + \tilde{\gamma}_{ij}(N, a_s, a_{em}), \qquad (2.6)$$

where $\gamma_{ij}(a_s)$ contains the pure QCD contributions and $\tilde{\gamma}_{ij}$ contains both the pure QED and the mixed QCD$\times$QED corrections.

The maximally decoupled evolution equations are then constructed as follows. The nonsinglet combinations

$$T_3^d = d^+ - s^+, V_3^d = d^- - s^-,$$
$$T_3^u = u^+ - c^+, V_3^u = u^- - c^-, \qquad (2.7)$$
$$T_8^d = d^+ + s^+ - 2b^+, V_8^d = d^- + s^- - 2b^-,$$
$$T_8^u = u^+ + c^+ - 2t^+, V_8^u = u^- + c^- - 2t^-, \qquad (2.8)$$

evolve independently according to nonsinglet evolution equations of the form

$$\mu^2 \frac{d}{d\mu^2} T_{3/8}^{u/d} = -(\gamma_{\text{ns},+} + \tilde{\gamma}_{u/d}^{\text{ns},+})T_{3/8}^{u/d}, \qquad (2.9)$$

$$\mu^2 \frac{d}{d\mu^2} V_{3/8}^{u/d} = -(\gamma_{\text{ns},-} + \tilde{\gamma}_{u/d}^{\text{ns},-})V_{3/8}^{u/d}. \qquad (2.10)$$

The valence sum and difference combinations, defined as

$$V = V_u + V_d, \quad V_\Delta = \frac{n_d}{n_u}V_u - V_d, \qquad (2.11)$$

satisfy coupled evolution equations

$$\mu^2 \frac{d}{d\mu^2}\begin{pmatrix} V \\ V_\Delta \end{pmatrix}$$

$$= -\begin{pmatrix} \gamma^{\text{ns},V} + \langle\tilde{\gamma}_q^{\text{ns},-}\rangle & \nu_u\tilde{\gamma}_{\Delta q}^{\text{ns},-} \\ \nu_d\tilde{\gamma}_{\Delta q}^{\text{ns},-} & \gamma^{\text{ns},-} + \{\tilde{\gamma}_q^{\text{ns},-}\} \end{pmatrix}\begin{pmatrix} V \\ V_\Delta \end{pmatrix}, \qquad (2.12)$$

in terms of the linear combinations of anomalous dimensions

$$
\nu_{u/d} = \frac{n_{u/d}}{n_f}, \quad \langle \tilde{\gamma}_q^{\mathrm{ns},\pm} \rangle = \nu_u \tilde{\gamma}_u^{\mathrm{ns},\pm} + \nu_d \tilde{\gamma}_d^{\mathrm{ns},\pm},
$$

$$
\{\tilde{\gamma}_q^{\mathrm{ns},\pm}\} = \nu_d \tilde{\gamma}_u^{\mathrm{ns},\pm} + \nu_u \tilde{\gamma}_d^{\mathrm{ns},\pm}, \tilde{\gamma}_{\Delta q}^{\mathrm{ns},\pm} = \tilde{\gamma}_u^{\mathrm{ns},\pm} - \tilde{\gamma}_d^{\mathrm{ns},\pm}. \tag{2.13}
$$

Finally, the gluon and photon satisfy coupled evolution equations together with the quark singlet sum and difference combinations, defined as

$$
\Sigma = \Sigma_u + \Sigma_d, \quad \Sigma_\Delta = \frac{n_d}{n_u} \Sigma_u - \Sigma_d. \tag{2.14}
$$

These coupled evolution equations read

$$
\mu^2 \frac{d}{d\mu^2} \begin{pmatrix} g \\ \gamma \\ \Sigma \\ \Sigma_\Delta \end{pmatrix} = -\Gamma \begin{pmatrix} g \\ \gamma \\ \Sigma \\ \Sigma_\Delta \end{pmatrix}, \tag{2.15}
$$

with Γ a 4×4 anomalous dimension matrix of the form

$$
\Gamma = \begin{pmatrix}
\gamma_{gg} + \tilde{\gamma}_{gg} & \tilde{\gamma}_{g\gamma} & \gamma_{gq} + \langle \tilde{\gamma}_{gq} \rangle & \nu_u \tilde{\gamma}_{g\Delta q} \\
\tilde{\gamma}_{\gamma g} & \tilde{\gamma}_{\gamma\gamma} & \langle \tilde{\gamma}_{\gamma q} \rangle & \nu_u \tilde{\gamma}_{\gamma \Delta q} \\
2n_f(\gamma_{qg} + \langle \tilde{\gamma}_{qg} \rangle) & 2n_f \langle \tilde{\gamma}_{q\gamma} \rangle & \gamma_{qq} + \langle \tilde{\gamma}_q^{\mathrm{ns},+} \rangle + \langle e_q^2 \rangle^2 \tilde{\gamma}_{\mathrm{ps}} & \nu_u \tilde{\gamma}_{\Delta q}^{\mathrm{ns},+} + \nu_u e_{\Delta q}^2 \langle e_q^2 \rangle \tilde{\gamma}_{\mathrm{ps}} \\
2n_f \nu_d \tilde{\gamma}_{\Delta qg} & 2n_f \nu_d \tilde{\gamma}_{\Delta q\gamma} & \nu_d \tilde{\gamma}_{\Delta q}^{\mathrm{ns},+} + \nu_d e_{\Delta q}^2 \langle e_q^2 \rangle \tilde{\gamma}_{\mathrm{ps}} & \gamma^{\mathrm{ns},+} + \{\tilde{\gamma}_q^{\mathrm{ns},+}\} + \nu_u \nu_d (e_{\Delta q}^2)^2 \tilde{\gamma}_{\mathrm{ps}}
\end{pmatrix},
$$

where $\tilde{\gamma}_{qq'}^{\mathrm{ps}} = e_q^2 e_{q'}^2 \tilde{\gamma}_{\mathrm{ps}}$ [38], and the combinations $\langle \tilde{\gamma}_{gq} \rangle$, $\langle \tilde{\gamma}_{qg} \rangle$, $\tilde{\gamma}_{g\Delta q}$, $\tilde{\gamma}_{\Delta qg}$, $\langle \tilde{\gamma}_{\gamma q} \rangle$, $\langle \tilde{\gamma}_{q\gamma} \rangle$, $\tilde{\gamma}_{\gamma \Delta q}$, and $\tilde{\gamma}_{\Delta q\gamma}$ are constructed analogously to those listed in Eq. (2.13).

The basis defined by Eqs. (2.7), (2.8), (2.11) and (2.14) is denoted as the unified evolution basis. Compared to the basis used in APFEL [28] to solve the QCD×QED evolution equations, our definitions of Σ_Δ and V_Δ differ due to the prefactors n_d/n_u that make the basis fully orthogonal. In the presence of scale-independent intrinsic heavy quarks (e.g. a charm-quark PDF in a three-flavor scheme [43]), a further decomposition needs to be applied [30].

The unified flavor basis has been implemented in the EKO code [30], which, as discussed in the introduction, is now used to solve evolution equations as part of the new theory pipeline. The basic ingredient of the EKO code is the construction of evolution kernel operators $E(Q^2 \leftarrow Q_0^2)$ (EKOs) such that

$$
f(Q^2) = E(Q^2 \leftarrow Q_0^2) f(Q_0^2), \tag{2.16}
$$

where $f(Q^2)$ is a vector whose components are all PDF flavors, including all active quark and antiquark flavors in a suitable basis, the photon and the gluon. Formally, the EKOs are given by

$$
E(Q^2 \leftarrow Q_0^2) = \mathcal{P} \exp
$$
$$
\times \left(-\int_{Q_0^2}^{Q^2} \frac{d\mu^2}{\mu^2} \gamma\left(a_s(\mu^2), a_{em}(\mu^2)\right) \right), \tag{2.17}
$$

where γ is the full matrix of anomalous dimensions, and $\mathcal{P}$ denotes path ordering. In the unified evolution basis the evolution kernel is a block diagonal matrix, with individual diagonal entries for the nonsinglet combinations Eqs. (2.7, 2.8) (without path-ordering), a 2×2 block for the valence combinations Eq. (2.11), and a 4×4 block in the singlet sector.

A variety of implementations of the solution of the evolution equations, each of those corresponding to a determination of the EKO, and which differ by higher-order terms, are available in the EKO code. These are discussed in Ref. [30]. At NNLO in QCD, the solutions are essentially indistinguishable in the data region. The solution adopted here is the iterated-exact (EXA) of Ref. [30], while the truncated (TRN) solution was adopted for the NNPDF3.1QED [21] and NNPDF4.0 [26] PDF sets, using the APFEL [28] implementation. These different implementations (at NNLO) are benchmarked and shown to be equivalent in Appendix B, where we also discuss the motivations for this choice.

2.2 Construction of the photon PDF

As mentioned in Sect. 1, the photon PDF is determined using the LuxQED [17,18] formalism, implemented in a PDF fit using the same methodology as the one used for the NNPDF3.1QED PDF set [21], but now with the new theory pipeline. The LuxQED result amounts to proving that the photon PDF is perturbatively determined in QED by knowledge of the proton inclusive structure functions F_2 and F_L:

$$
x\gamma(x, \mu^2) = \frac{2}{a_{em}(\mu^2)} \int_x^1 \frac{dz}{z} \left\{ \int_{\frac{m_p^2 x^2}{(1-z)}}^{\frac{\mu^2}{(1-z)}} \frac{dQ^2}{Q^2} a_{em}^2(Q^2) \right.
$$

$$
\times \left[-z^2 F_L(x/z, Q^2) \right.
$$

$$
\left. + \left(z P_{\gamma q}(z) + \frac{2x^2 m_p^2}{Q^2} \right) F_2(x/z, Q^2) \right]
$$

$$
\left. - a_{em}^2(\mu^2) z^2 F_2(x/z, \mu^2) \right\}, \tag{2.18}
$$

where m_p is the mass of the proton and $P_{\gamma q}$ is the photon-quark splitting function. Equation (2.18) holds in the $\overline{\mathrm{MS}}$ scheme, including terms of order a_{em} and $a_{em}^2 \ln \mu^2/m_p^2$, times the accuracy of the QCD determination of the structure functions.

The integration over the scale Q^2 and Bjorken-x dependence of the structure functions $F_i(x, Q^2)$ in Eq. (2.18) includes four different regions and corresponding contributions to the structure functions: an elastic contribution at $x = 1$, a resonance contribution when x is large and close to $x = 1$, an inelastic non-perturbative contribution at low Q^2 and intermediate x and an inelastic perturbative region for intermediate x and large Q^2. The first three contributions must be determined by fits to lepton-proton scattering data, but the latter contribution may be determined by expressing the structure functions through perturbative QCD factorization in terms of quark and gluon PDFs. In a global PDF determination, one may choose to determine the photon PDF using the LuxQED formula Eq. (2.18) at a single chosen scale $\mu^2 = Q_\gamma^2$, and then evolve jointly the photon and all other PDFs through the QCD×QED evolution equations discussed in Sect. 2.1. An alternative option is to determine the photon PDF using Eq. (2.18) at all scales. The two choices are equivalent because the LuxQED photon Eq. (2.18) satisfies the joint QED×QCD evolution equations to the accuracy of the LuxQED formula itself, though they differ by higher-order corrections and also due to the fact that the LuxQED photon is partly determined from a parametrization of data. The former choice was made for the construction of our previous NNPDF3.1QED set [21], which we follow here, also in the choice of parametrization of the structure functions in the non-perturbative region. The same choice was made in Refs. [21–25], though in Refs. [24,25] a PDF set in which the photon PDF is determined using Eq. (2.18) at all scales was also presented.

The relative size of the elastic, inelastic non-perturbative (including resonances) and elastic perturbative contributions to $\gamma(x, \mu^2)$ is displayed in Fig. 1 as a function of x for four different choices of the scale $\mu = Q_\gamma$ at which the LuxQED formula Eq. (2.18) is used to determine the photon (see also Fig. 18 in Ref. [18]). If the LuxQED formula is used at the scale $Q_\gamma = 1.65$ GeV at which the NNPDF4.0 PDFs are parametrized, the photon is entirely determined by the elastic and non-perturbative contributions, but as the scale is increased, an increasingly large contribution comes from the perturbative region. By the time a value $Q_\gamma \gtrsim 100$ GeV is reached, the photon at small $x \lesssim 10^{-3}$ is almost entirely determined by the perturbative contribution, while the elastic and non-perturbative contributions only remain dominant at large $x \gtrsim 0.1$ where however the photon PDF is tiny. Choosing a large value of Q_γ has the dual advantage that the LuxQED result is more accurate at high scale, because it includes contributions of order a_{em} and $a_{em}^2 \ln \mu^2/m_p^2$, and also that one is then mostly relying on the accurate perturbative determination of the structure functions, that exploits global information on quarks and gluon PDFs and not just the lepton-proton

scattering data. As in Ref. [21], we choose $\mu = Q_\gamma = 100$ GeV, as also advocated in Ref. [18].[1] We will discuss the dependence of our results on the choice of scale Q_γ in Sect. 3.

Equation (2.18) must be viewed as a constraint on the set of photon, quark and gluon PDFs that are simultaneously determined. Because of the small impact of the photon PDF on the other PDFs, it was suggested in Refs. [18,21] that the constraint can be implemented iteratively. Namely, we first determine the photon PDF using Eq. (2.18) by means of structure functions that are determined from an existing PDF set at a given scale $\mu = Q_\gamma$. This photon PDF is then evolved, by solving joint QCD×QED evolution equations with the unchanged given set, to a chosen PDF parametrization scale where it is taken as fixed. All other PDFs are then re-determined, with the constraint that the momentum sum rule now also includes a contribution from the given (fixed) photon, i.e.

$$\int_0^1 dx \left(x\Sigma(x, Q^2) + xg(x, Q^2) + x\gamma(x, Q^2) \right) = 1.$$
(2.19)

The photon PDF at a scale $\mu = Q_\gamma$ is then determined again from Eq. (2.18), in which structure functions are obtained from the new fit. The procedure is iterated until convergence, which was achieved in two iterations in the NNPDF3.1QED determination [21].

Here we follow the same procedure, starting with a re-determination of the NNPDF4.0 PDF set with the new pipeline, which is compared and shown to be equivalent to the published NNPDF4.0 in Appendix A.

3 The NNPDF4.0QED parton distributions

We present here the NNPDF4.0QED PDFs: we first summarize our procedure, then discuss the effect on fit quality of the inclusion of the photon PDF, examine the photon PDF itself, also in comparison to other determinations, and finally study the photon momentum fraction.

3.1 Construction of the NNPDF4.0QED parton set

As mentioned in Sect. 2 all the methodological aspects and settings of the PDF determination are the same as used for the underlying pure QCD NNPDF4.0 PDF [26], but now using a new theory pipeline. Even if in principle theory predictions should be independent of implementation details, in practice

---

[1] Note that in Ref. [44], Table 1, the value at which the photon is evaluated in Ref. [18] (called μ_{eval}, see Sect. 10.1 of that Ref.) is incorrectly reported as $Q = 10$ GeV, instead of the correct $Q = 100$ GeV.
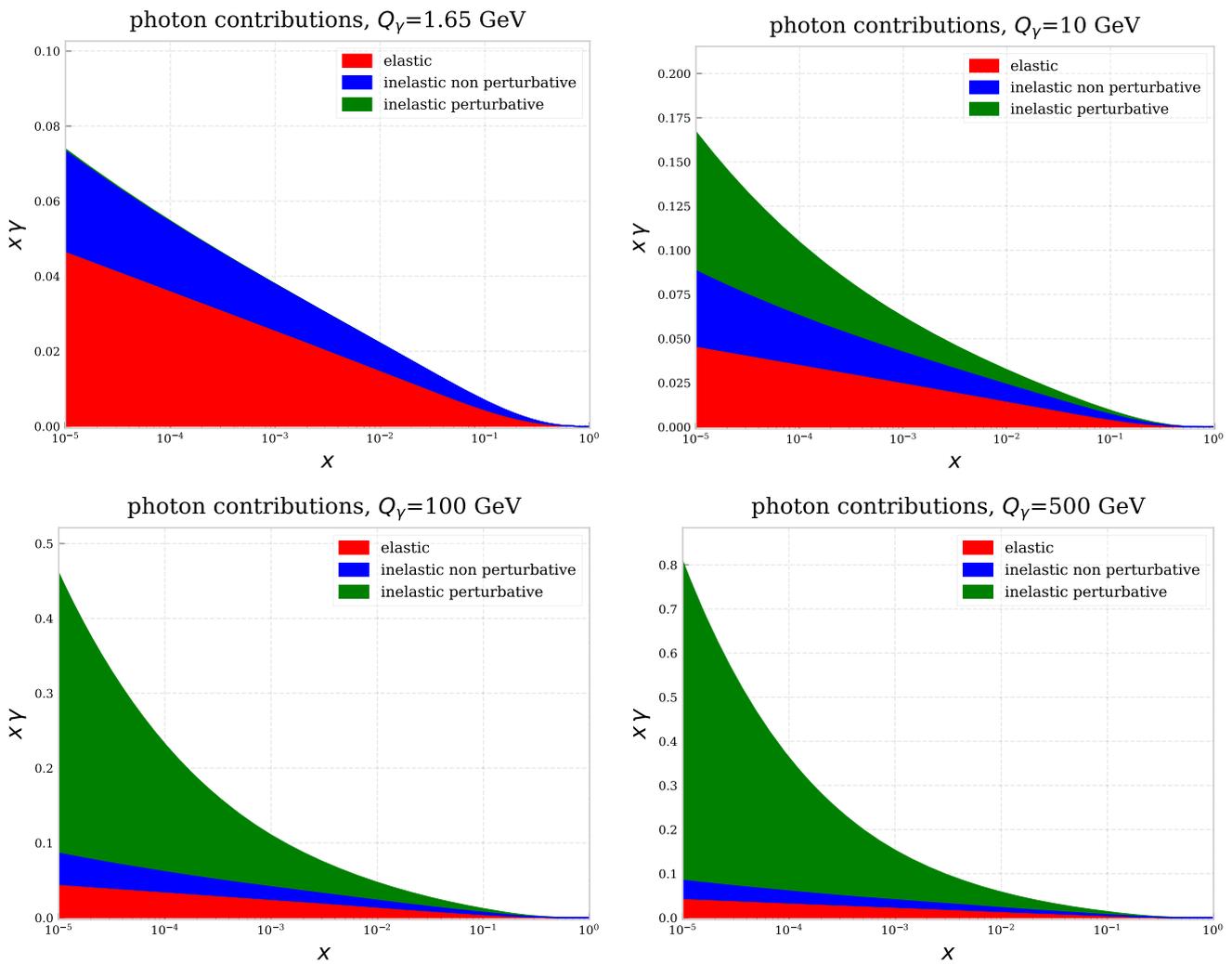
Fig. 1 Breakdown of the photon PDF $\gamma(x, \mu^2)$ into the contributions coming from different contributions to the proton structure functions $F_i(x, Q^2)$ that determine it according to the LuxQED formula Eq. (2.18). The result is shown as a function of x for four different choices of the scale $\mu = Q_\gamma$: $Q_\gamma = 1.65$ GeV (top left), 10 GeV (top right), 100 GeV (bottom left) and 500 GeV (bottom right)

differences may arise, e.g. due to issues of numerical accuracy. Also, in the process of transitioning to this new pipeline, a few minor bugs in data implementation were uncovered and fixed. Finally, the new pipeline includes a new implementation of heavy-quark mass effect in deep-inelastic structure functions that differs by subleading terms from the previous one. Benchmarks showing the equivalence of the old and new pipeline are briefly presented in Appendix A. Because of this equivalence, NNPDF4.0 PDF replicas produced with the new pipeline should be considered equivalent to the published ones, and indeed for phenomenological applications (specifically those presented in Sect. 4) we will compare the results obtained using NNPDF4.0QED PDFs to pure QCD results obtained using the published NNPDF4.0 replicas. Nevertheless, in this section only, for all comparisons between pure QCD and QCD×QED, we will use NNPDF4.0 pure QCD

replicas generated using the new pipeline, in order to avoid even small confounding effects. Note however that, as discussed in the end of Sect. 2.1, the QED PDFs are based on the EXA solution of the evolution equations, that differs by subleading terms from the TRN solution used in the pure QCD fit. The effect of this difference is assessed in Appendix B and is very small at NNLO. However, all comparisons between pure QCD and QED PDFs also include the effect of this change.

The NNPDF4.0QED PDFs are determined at NLO and NNLO by supplementing with a photon PDF the pure QCD PDF set, according to the methodology outlined in the previous section. However, all theory predictions are obtained as in the pure QCD determination: hence in particular no photon-induced contributions are included, and thus the only effect of the inclusion of a photon PDF is through its mixing with

other PDFs. This is justified because the NNPDF4.0 dataset was constructed including cuts that remove all datapoints for which the effect of electroweak corrections is larger than the experimental uncertainties (see Ref. [26], Sect. 4.1]). The inclusion of the electroweak corrections, which will allow relaxing these cuts, is left for future work. As in Ref. [21], the final PDF set is obtained after two iterations, with a third iteration providing a check of convergence.

3.2 Fit quality

Table 1 displays the statistical estimators obtained using a set of 100 NNPDF4.0QED NLO and NNLO PDF replicas, compared to their QCD-only counterparts, generated using the new theory pipeline. Specifically, we show the χ^2, for both the full dataset and for datasets grouped by process; for the full dataset we also show the average over replicas of the training and validation figures of merit $\langle E_{tr} \rangle_{rep}$ and $\langle E_{val} \rangle_{rep}$, and the average χ^2 over replicas $\langle \chi^2 \rangle_{rep}$, all as defined in Table 9 of Ref. [45]. Note that χ^2 and $\langle \chi^2 \rangle_{rep}$ are computed using the experimental covariance matrix, while, as in all NNPDF determinations, the figure of merit used for minimization is computed using the t_0 covariance matrix [46].

It is clear from Table 1 that all estimators are essentially unchanged by the inclusion of QED corrections, with slightly larger differences seen at NLO than at NNLO: the impact of QED effects on fit quality is negligible, both globally and for individual processes. Specifically, the training and validation figures of merit in the pure QCD and QCD×QED determinations differ by less than one sigma. As mentioned, these differences also include the effect of switching from the TRN solution of evolution equations in the pure QCD fit to the EXA solution. The impact of this change is yet smaller than that of the QED corrections, but it slightly reduces it.

3.3 The photon PDF

In Fig. 2, the NNLO NNPDF4.0QED photon PDF at $Q = 100$ GeV is compared to its counterpart in our previous NNPDF3.1QED [21] set, and in the recent QED PDF sets MSHT20QED [23], and CT18QED [25].[2] Here and elsewhere in this paper all uncertainties correspond to 1σ. Results agree at the percent level, despite the fact that quark and gluon PDFs in these sets can display much larger differences. This is a consequence of the fact that in all these PDF sets the photon PDF is determined with the LuxQED formalism in terms of the proton structure function, and that the latter, in turn, is well constrained by experimental data both at high and low scale. In fact, we have checked that the dominant con-

tribution to the difference between the NNPDF3.1QED and NNPDF4.0QED photon PDFs seen in Fig. 2 is due to the difference between the TRN and EXA solutions of evolution equations (see Sect. 3.1 and Appendix B), i.e. to higher-order corrections, with the residual difference being due to the change in the PDFs used in order to compute the structure function. The fact that the 3.1 and 4.0 PDFs are compatible within uncertainties is thus a consequence of the fact that the uncertainties due to higher-order corrections and to PDFs are correctly accounted for by the LuxQED construction, while the NNPDF4.0 PDFs are backward-compatible with the NNPDF3.1 PDFs [26].

The uncertainty on the photon PDF is completely dominated by theoretical uncertainties on the LuxQED procedure, which include [18] missing higher order corrections, uncertainties on the experimentally measured low-scale structure function and so on. In our uncertainty determination we follow Ref. [18]; MSHT also mostly follows this reference, with an extra higher-twist contribution due to the low choice of scale Q_γ, [22] and indeed finds a very similar uncertainty. A somewhat more conservative uncertainty estimate is provided by CT18QED [25], which also adopts a somewhat different determination of the elastic contribution.

In Fig. 3 the central value of the photon PDF in all these sets is compared at different scales, using the native scale dependence of each PDF set. Uncertainties are not shown in order not to clutter the plot. Note that differences in the scale dependence of the PDFs compared in the figure also arise due to somewhat different treatments of the QCD×QED evolution equations. Specifically, the NNPDF3.1QED PDF set adopts a numerical implementation of the TRN solution (see Appendix B). However, one would expect differences to be mostly driven by the mixing with the quark and gluon PDFs. The fact that differences grow at low scales suggests that this is indeed the case. Even so, all photon PDFs agree within about 3% for all $x \gtrsim 10^{-3}$, even at the lowest scale $Q = 1.65$ GeV.

In Fig. 3 we also show the central photon PDF which is found by repeating the NNPDF4.0QED determination with the photon PDF determined at a scale $Q_\gamma = 10$ GeV instead of the default $Q_\gamma = 100$ GeV. If a low value of Q_γ is adopted, the upper limit of integration in Q^2, Eq. (2.18), is accordingly lower. As discussed in Sect. 2.2 in such a case a sizable contribution to the LuxQED formula comes from the low-scale region in which the structure function is determined from a fit to the data, and $\mathcal{O}\left(\frac{m_p}{Q}\right)$ corrections to the LuxQED formula may then become relevant. We see that the shift in central photon PDF which is found by making this choice instead of the default one is at most of the order of the uncertainty on the photon PDF. Again, this shows that the uncertainty on the LuxQED procedure is correctly estimated.

---

[2] The MSHT group has recently also released the MSHT20qed_an3lo PDF set, in which a photon PDF set is added to PDFs treated with approximate N³LO QCD theory [47].

Table 1 Statistical estimators for NNPDF4.0QED NLO and NNLO, compared to NNPDF4.0 pure QCD. From top to bottom: total χ^2 per number of data points, average over replicas of the training and vali- dation figures of merit $\langle E_{tr}\rangle_{rep}$ and $\langle E_{val}\rangle_{rep}$, average χ^2 over replicas $\langle \chi^2 \rangle_{rep}$, χ^2 for datasets grouped by process. The total number of data points is 4424 (4616) at NLO (NNLO)

	Dataset	NNPDF4.0 NLO		NNPDF4.0 NNLO	
		QCD×QED	QCD	QCD×QED	QCD
χ^2	Global	1.31	1.26	1.17	1.17
$\langle E_{tr}\rangle_{rep}$		2.47±0.07	2.41±0.06	2.27±0.06	2.28±0.05
$\langle E_{val}\rangle_{rep}$		2.66±0.11	2.57±0.10	2.39±0.10	2.37±0.11
$\langle \chi^2 \rangle_{rep}$		1.337±0.016	1.286±0.017	1.192±0.014	1.195±0.015
χ^2	DIS neutral-current	1.38	1.31	1.22	1.23
	DIS charged-current	0.94	0.92	0.90	0.90
	Drell–Yan (inclusive and with one jet)	1.56	1.56	1.30	1.31
	Top-quark pair production	2.31	1.98	1.31	1.24
	Single-top production	0.38	0.36	0.39	0.36
	Inclusive jet production	0.83	0.85	0.93	0.96
	Dijet production	1.56	1.55	1.94	2.03
	Direct photon production	0.64	0.58	0.74	0.75

We finally assess the impact of the inclusion of a photon PDF on the other PDFs, by comparing the NNLO NNPDF4.0QED and NNPDF4.0 (pure QCD) PDFs. In Fig. 4 we present this comparison at $Q = 100$ GeV. It is clear that the impact of the inclusion of the photon is moderate, with the NNPDF4.0 (pure QCD) and NNPDF4.0QED PDFs generally differing by less than one sigma and always in agreement within uncertainties. The largest effects are seen in the gluon PDF, which is suppressed at the percent level due to the momentum fraction transferred from the gluon to the photon.

3.4 The photon momentum fraction

The main impact of the photon on other PDFs is through its contribution to the momentum sum rule, Eq. (2.19). We quantify this contribution by evaluating the photon momentum fraction

$$M\left[\gamma(Q)\right] \equiv \int_0^1 dx\, x\gamma(x, Q). \tag{3.1}$$

In Fig. 5 the momentum fractions carried by the photon and by the gluon PDFs in the NNPDF4.0QED set are shown (in percentage) as a function of scale. For the photon the result is compared to that of NNPDF3.1QED, and for the gluon to that of NNPDF4.0 (pure QCD). The photon momentum fraction in the NNPDF3.1QED and NNPDF4.0QED PDF sets is essentially the same: the photon carries around 0.2% of the proton momentum at a low (Q∼1 GeV) scale, growing logarithmically with Q up to around 0.6% at the multi-TeV scale. The momentum fraction carried by the gluon is reduced by a comparable amount upon inclusion of the photon.

4 Implications for LHC phenomenology

We now study the phenomenological implications of the NNPDF4.0QED PDF set. First we compare parton luminosities to those computed using other QED PDF sets. Then we assess the impact of QED corrections on selected processes, by comparing to the pure QCD case calculations that include photon-induced contributions, and also by directly comparing results obtained using NNPDF4.0 (pure QCD) and NNPDF4.0QED PDFs.

4.1 Luminosities

The phenomenological effect on the parton luminosity of the inclusion of a photon PDF is both direct, through the presence of photon-induced partonic channels, and indirect, through the effect of the photon on other PDFs, mostly through the depletion of the gluon that is necessary in order to preserve the momentum sum rule, as discussed in Sect. 3.3–3.4. The photon-induced contributions for NNPDF4.0QED NNLO are compared in Fig. 6 to their counterpart in NNPDF3.1QED (top) and in MSHT20QED and CT18QED (bottom). The level of agreement is high and directly follows from that seen between the respective photon PDFs in Fig. 2

The luminosities for all other parton channels are shown in Figs. 7, 8, where we compare NNPDF4.0QED both to the previous set NNPDF3.1QED and to the pure QCD NNPDF4.0 (Fig. 7) and to the MSHT20QED and CT18QED sets (Fig. 8). Because, as shown in Sect. 3, the effect of the inclusion of the photon on the other PDFs is moderate, the comparison between the two QED sets is very similar to the comparison between the pure QCD NNPDF4.0 and NNPDF3.1 shown in
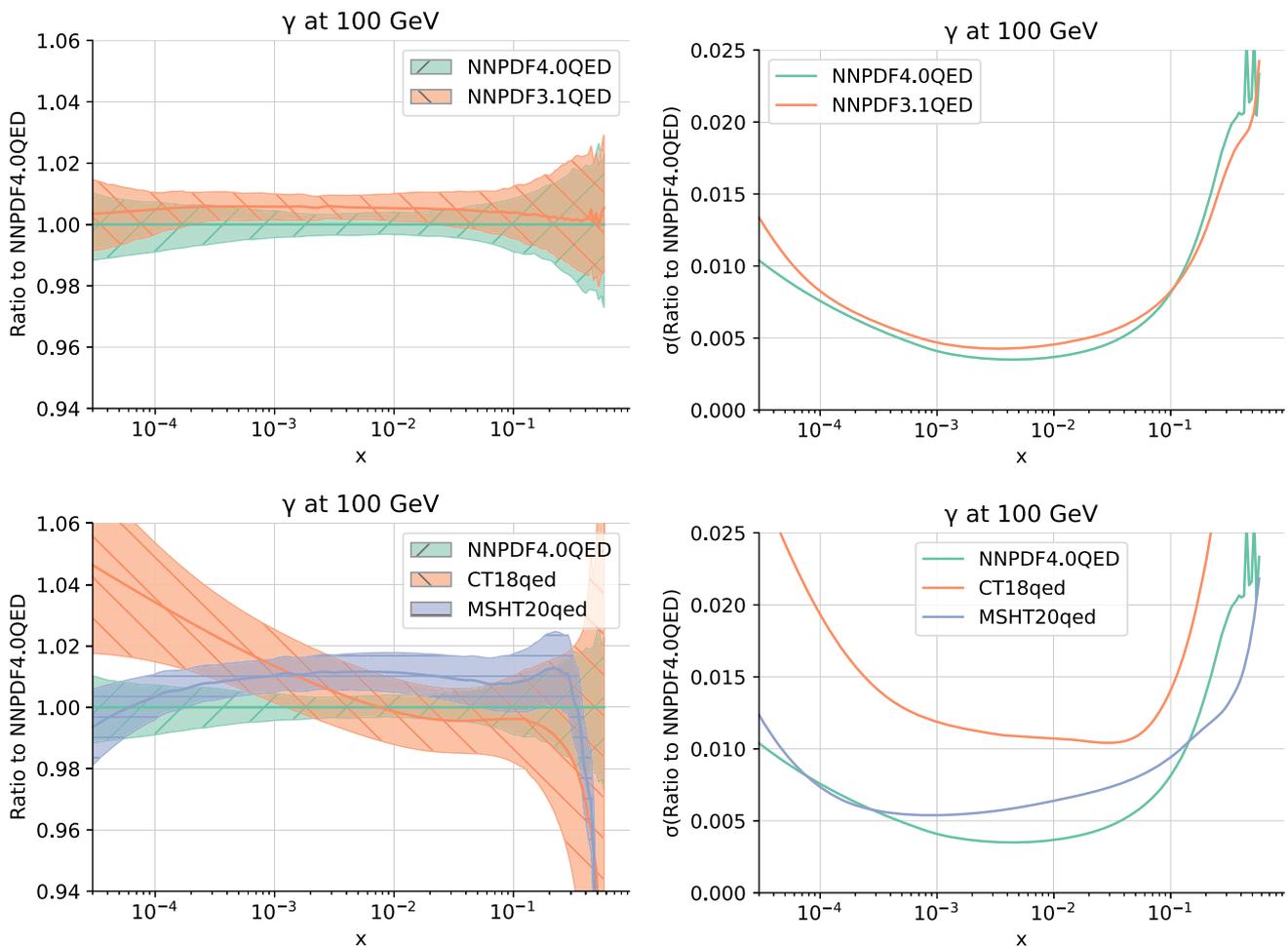
Fig. 2 Top left: the photon PDF at $Q = 100$ GeV in NNPDF4.0QED NNLO compared to its NNPDF3.1QED counterpart as a ratio to the central value of the former. Top right: the relative PDF uncertainties on the photon PDF in these two determinations. Bottom: same as top panels now comparing NNPDF4.0QED with MSHT20QED [23] and CT18QED [24] (all NNLO). Bands correspond to 1σ uncertainties

Fig. 9.1 of Ref. [26], where it was shown that NNPDF4.0 are backward compatible, i.e. generally agree with NNPDF3.1 within uncertainties. The effect of the inclusion of QED corrections is, as expected, mostly seen in the gluon–gluon channel, where it leads to a suppression of a few percent in the $m_X \sim 100$ GeV region in order to account for the transfer of a small amount of momentum to the photon. The comparison to other PDF sets is dominated by the differences in the quark and gluon PDFs, which are rather more significant than the difference in the photon PDF, and thus very similar to the corresponding comparison of pure QCD luminosities shown in Fig. 9.3 of Ref. [26].

4.2 Physics processes

We now study the impact of the photon PDF on a few representative processes: Drell–Yan production (neutral- and charged-current), Higgs production in gluon-gluon fusion, in

vector boson fusion, and in associated production with weak bosons, diboson production, and top-quark pair production, all at the LHC with center-of-mass energy $\sqrt{s} = 14$ TeV. We have computed theory predictions for these processes exploiting the PINEAPPL [32,33] interface to the automated QCD and EW calculations provided by MG5_AMC@NLO [35]. PINEAPPL produces interpolation grids, accurate to NLO in both the strong and electroweak couplings, that are independent of PDFs. They therefore make it easy to vary the input PDF set, since the same grid can be used for all PDF sets considered. As we are interested in assessing differences between PDF sets, rather than in doing precision phenomenology, we do not include NNLO QCD corrections, and we only show PDF uncertainties. For a detailed list of parameters and cuts used in the calculation of these grids, see Sect. 9 of Ref. [26].

For all processes, we display predictions in figures below with a standardized format, as follows. We show the abso-
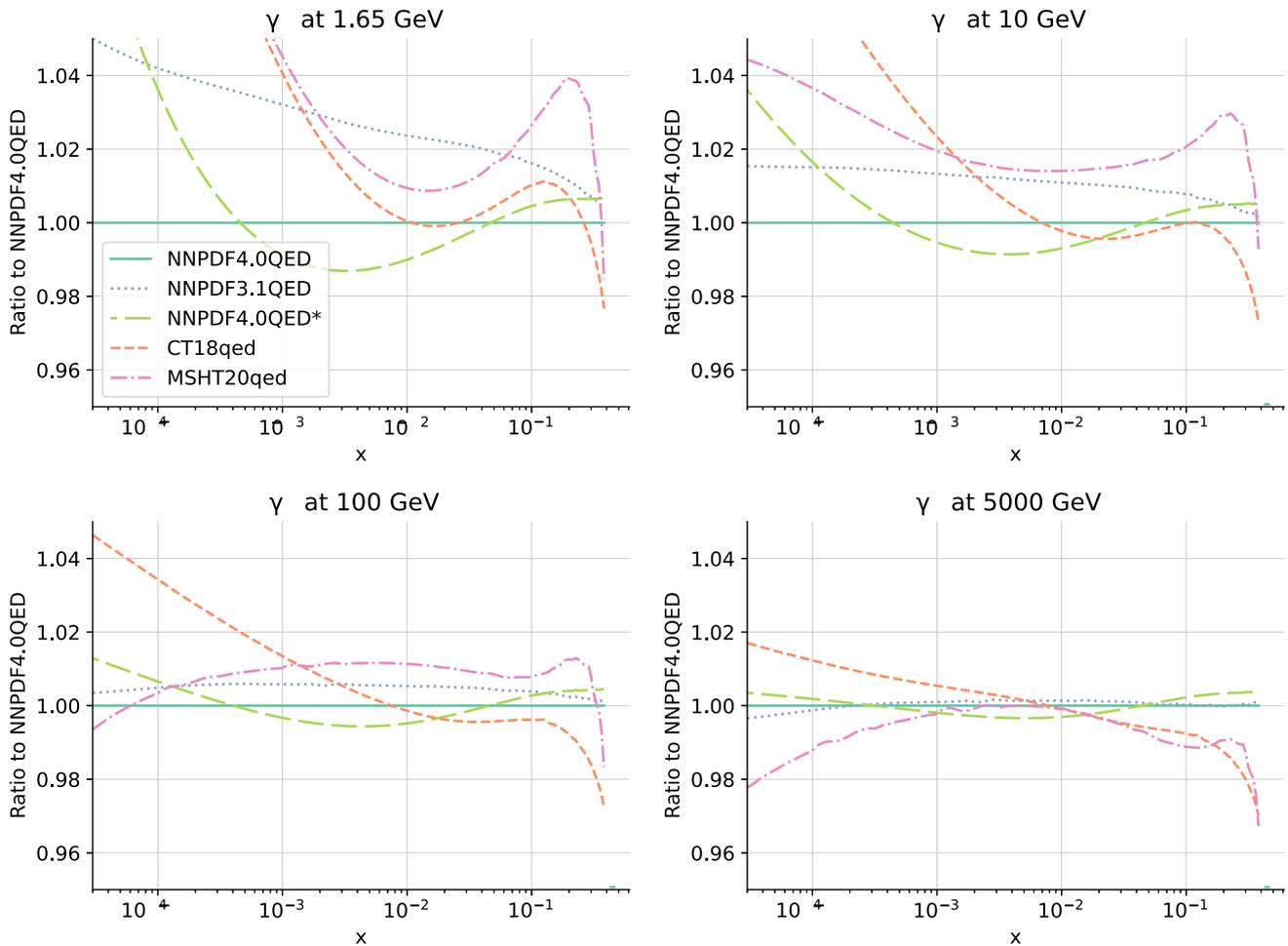
Fig. 3 Scale dependence of the central photon PDF in the NNPDF4.0QED, NNPDF3.1QED, MSHT20QED, and CT18QED PDF sets. The NNPDF4.0QED result in which the photon is determined at $Q = 10$ GeV (denoted as NNPDF4.0QED*) is also shown

lute distributions (top panels) and the ratio to the central value obtained using NNPDF4.0QED PDFs and including photon-induced channels, which we call NNPDF4.0QED (bottom panels). In the left plots we compare NNPDF4.0QED to: NNPDF4.0QED but with no photon-initiated channels; NNPDF4.0 pure QCD; NNPDF3.1QED. In the right plots we compare NNPDF4.0QED to MSHT20QED and CT18QED. The left plots allow for assessing the overall size of the QED corrections (by comparison of the QED and pure QCD results), and disentangling the size of the photon-initiated contributions (by comparison of predictions with the photon-initiated channels switched on and off) and the impact of the changes in the quark and gluon PDFs due to QED effects (by comparison of NNPDF4.0 with NNPDF4.0QED with the photon-initiated channels switched off).

In Fig. 9 we show results for inclusive Drell–Yan production both in the neutral-current and charged-current channels. For neutral current we show results for the invariant mass distribution of the dilepton pair, while for charged-current we display the rapidity distribution of the lepton. Predictions for

the Higgs rapidity distribution are shown in Fig. 10 for gluon fusion, associated production with a W^+ boson, and vector boson fusion. Finally, in Fig. 11 we show predictions for weak-boson pair production (W^+W^- and W^+Z), as a function of the dilepton transverse momentum, and for top-quark pair production as a function of the invariant mass of the top quark pair.

For charged-current Drell–Yan, the QED corrections have essentially no effect. In all the other cases that we consider, the effects of the QED correction fall in one of two categories. Either upon inclusion of QED effects we see an enhancement of the cross-section, which is only present when the photon-induced contribution is included, while the NNPDF4.0QED result without photon-induced contribution is very close to the pure QCD result. Or else we see a suppression of the cross-section, but with the NNPDF4.0QED result with and without the photon-induced contribution very close to each other. Of course, the former case can be explained with the presence of a sizable positive photon-induced contribution, while the latter case is explained by the suppression of the
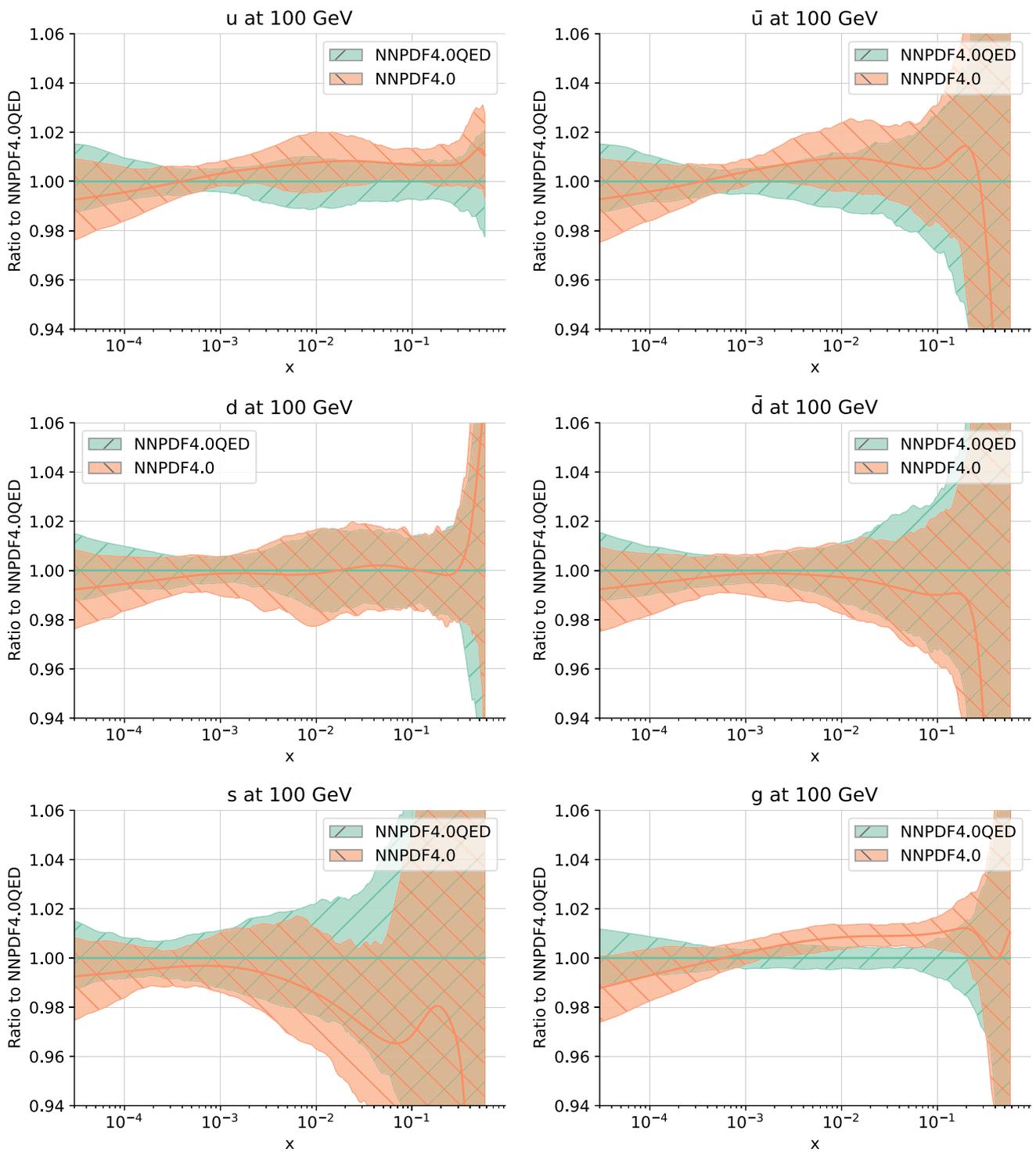
Fig. 4 Comparison of PDFs in the NNPDF4.0QED and the NNPDF4.0 (pure QCD) sets, shown as a ratio to the former at $Q = 100$ GeV. Bands correspond to 1σ uncertainties. From left to right and from top to bottom, the up, anti-up, down, anti-down, strange, and gluon PDFs are shown
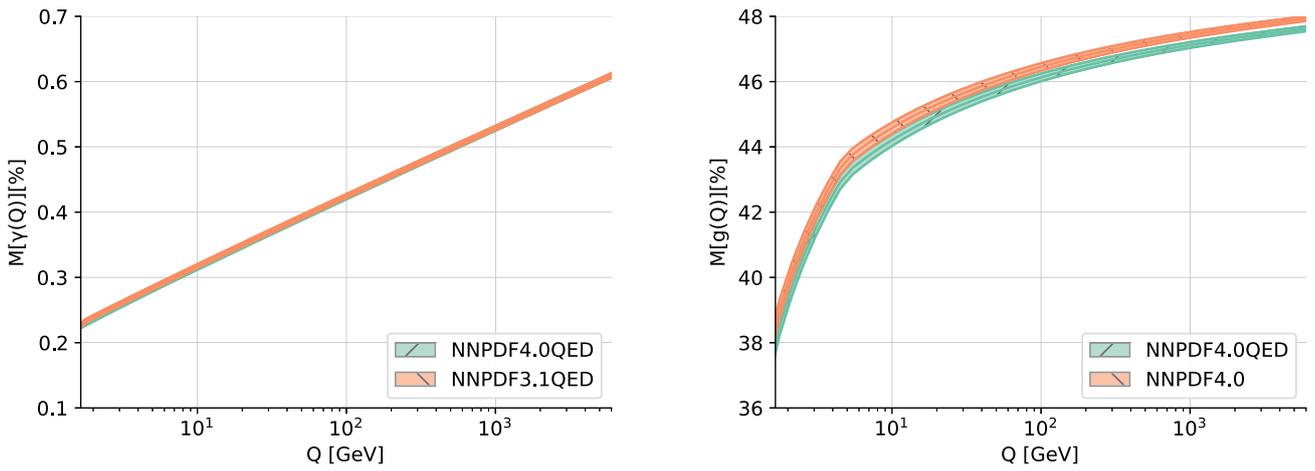
Fig. 5 Left: the percentage momentum fraction carried by the photon PDF $\gamma(x, Q^2)$ in NNPDF4.0QED and in NNPDF3.1QED as a function of the scale Q, where the bands indicate 1σ uncertainties. Right: same for the momentum fraction carried by the gluon PDF in NNPDF4.0QED and in NNPDF4.0 (pure QCD)
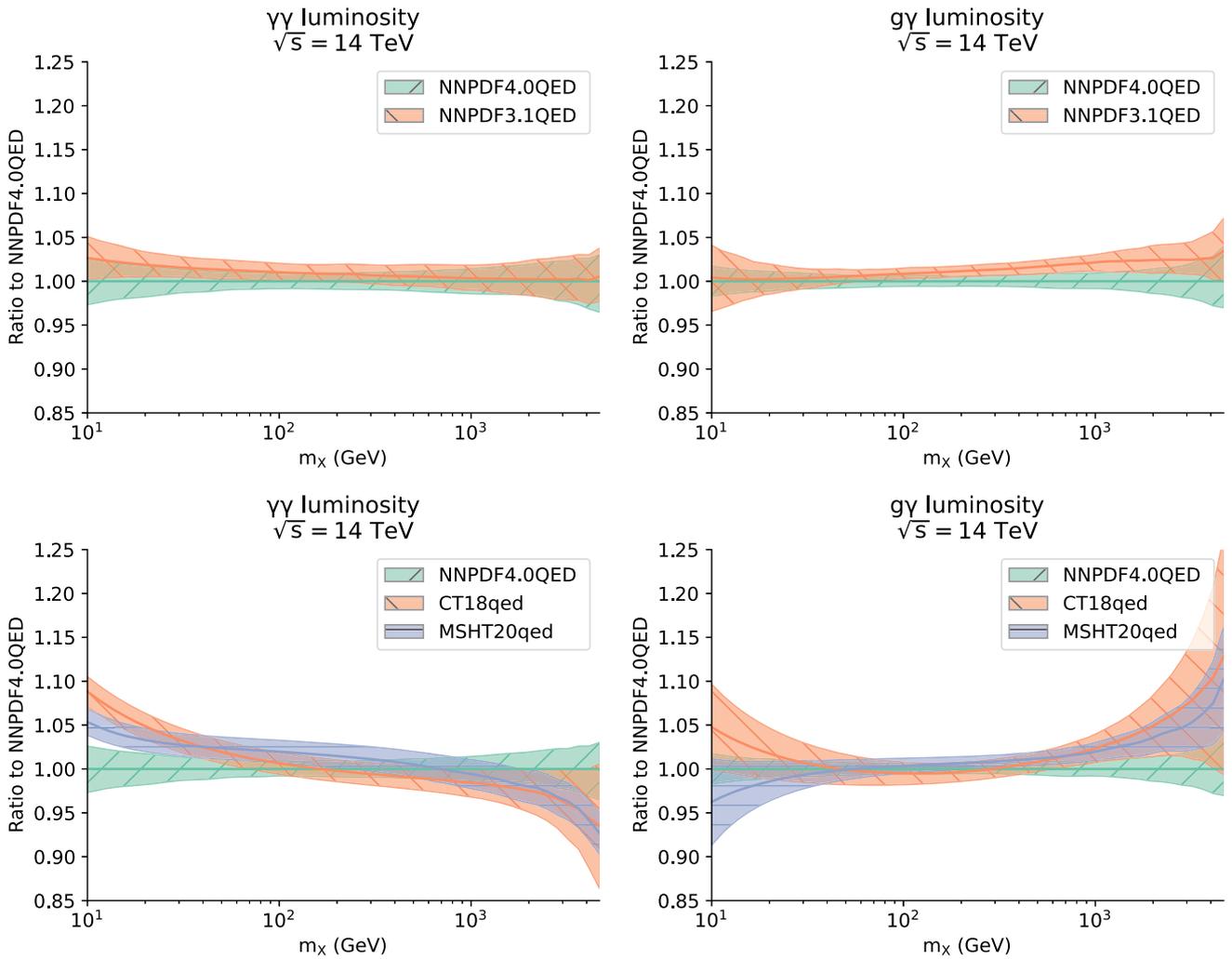
Fig. 6 The photon-induced contributions to the luminosity at the LHC with $\sqrt{s} = 14$ TeV as a function of the invariant mass m_X for NNPDF4.0QED NNLO compared to its NNPDF3.1QED counterpart (top) and compared to MSHT20QED and CT18QED, all shown as a ratio to NNPDF4.0QED
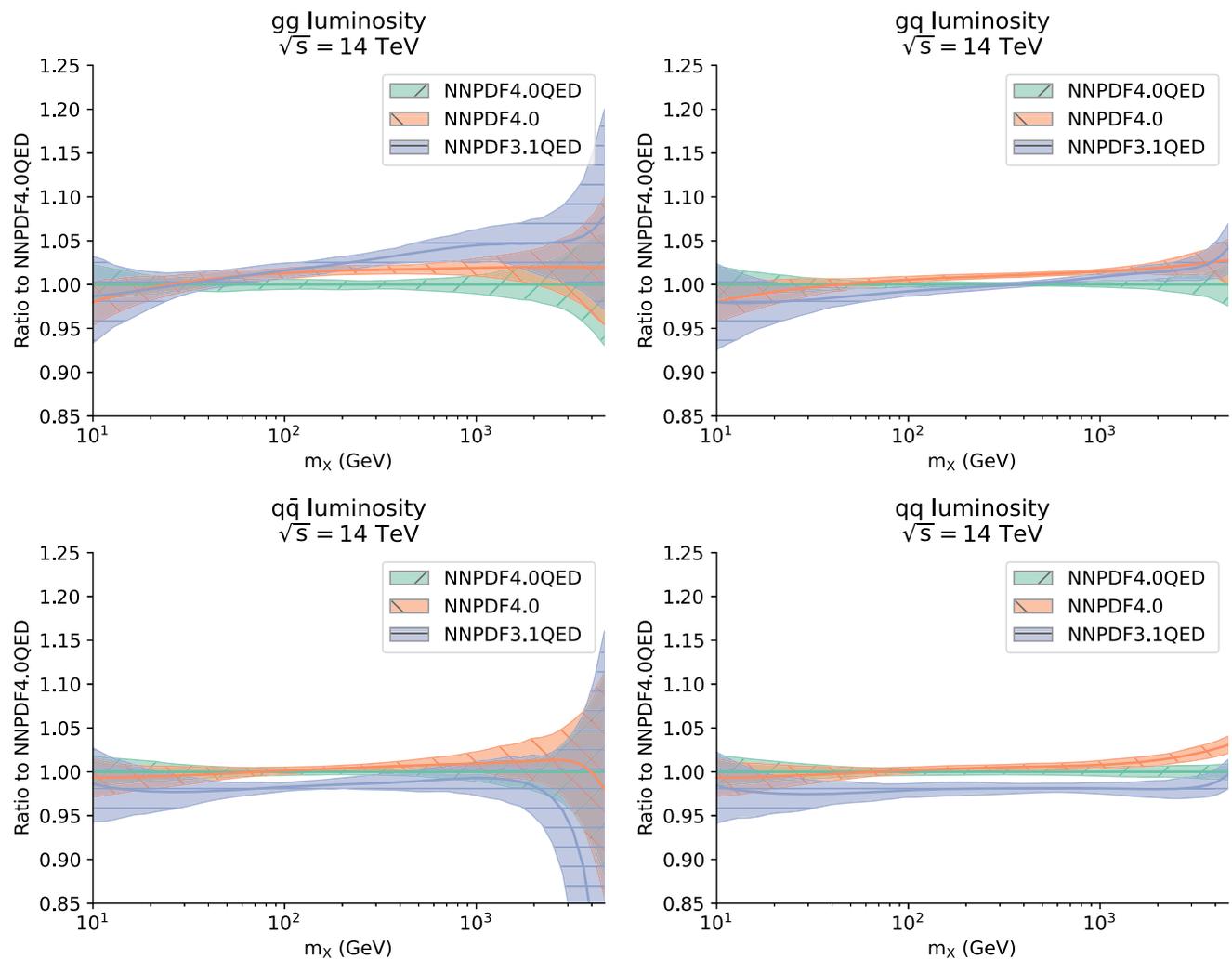
Fig. 7 Comparison of parton luminosities in the NNPDF4.0QED and the NNPDF4.0 (pure QCD) sets, shown as a ratio to the former for the LHC with $\sqrt{s} = 14$ TeV. From left to right and from top to bottom the gluon-gluon, gluon-quark, quark-antiquark and quark-quark luminosities are shown

gluon luminosity due to the transfer of momentum fraction from the gluon to the photon seen in Fig. 7.

The first situation – enhancement due to the photon–induced contribution – is observed in neutral-current dilepton production, where the enhancement increases with invariant mass and can reach up to 5% at the TeV scale. Note that the photon-induced contribution to the dilepton final state can proceed also through t- and u-channel leading-order diagrams. The fact that the QED enhancement is absent at the Z peak suggests that this non-resonant contribution provides the dominant part of the photon-induced contribution. A similar situation occurs in W^+W^- and ZW^+, where the enhancement increases with p_T and reaches 5% in the former case and 2% in the latter case for transverse momenta in the TeV range. Finally, the enhancement is also observed in associate Higgs production with W^+ and in vector boson fusion. In the former case the enhancement is largest at for-

ward rapidity, where it reaches 4%, and it decreases to 2% for the largest rapidity $y_H = 2.5$. In the latter case the enhancement is very moderate, around 1%, and almost independent of the rapidity (though slightly decreasing as the rapidity increases).

The second situation – suppression, independent of the photon-induced contribution – is found in processes that proceed through gluon fusion. The effect is clearly seen in Higgs production in gluon fusion, where the suppression is weakly dependent on rapidity, varying between 2% at central rapidity and about 1% at the largest rapidity. A similar, but more moderate effect, is also seen in top pair production, where the suppression is of order 1%, essentially independent of the invariant mass of the top pair.

When comparing results obtained using different QED PDF sets, be they NNPDF3.1QED, MSHT20QED or CT18QED, we observe that differences are essentially driven by the dif-
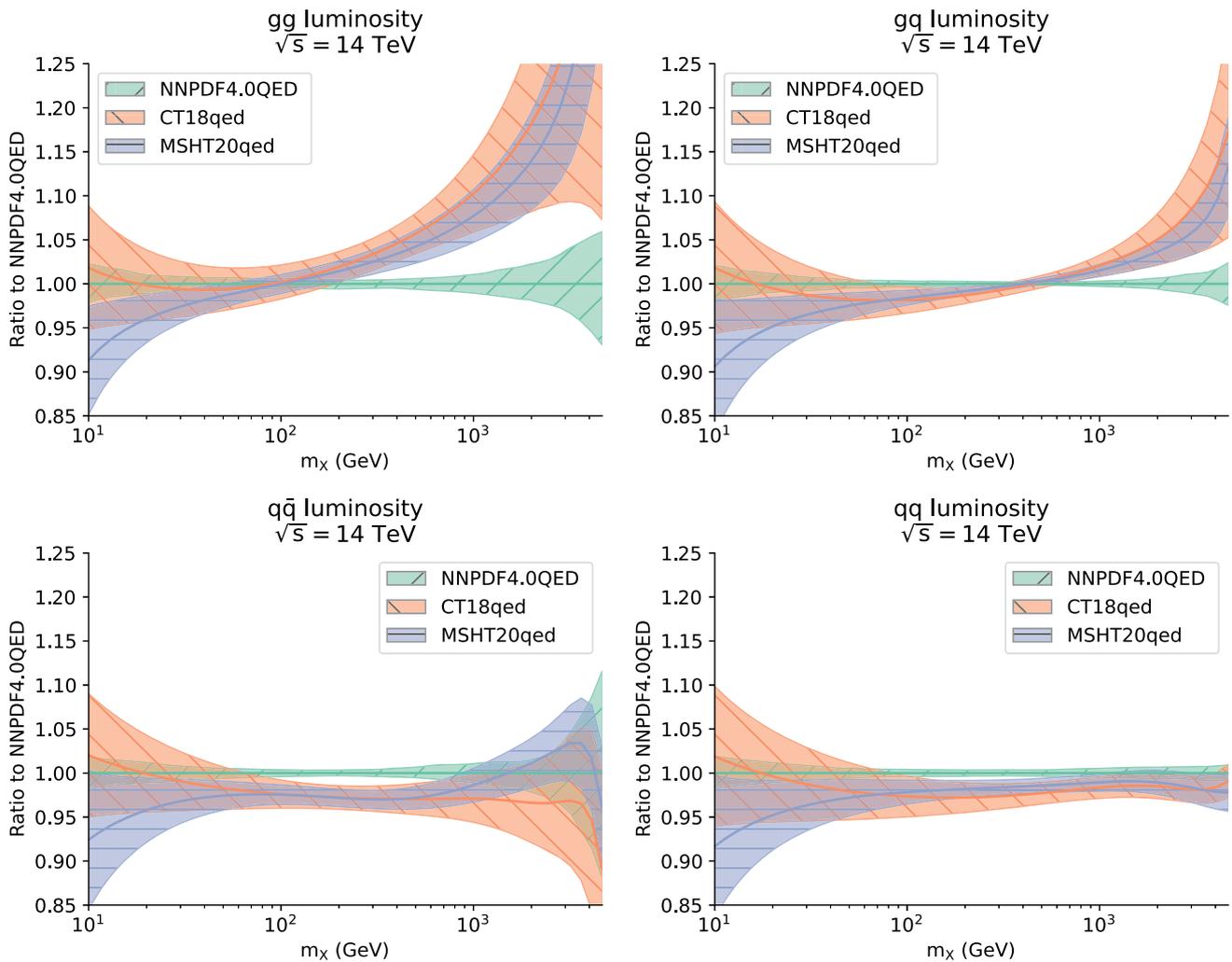
Fig. 8 Same as Fig. 7, now comparing NNPDF4.0QED to MSHT20 [23] and CT18qed [25]

ference in quark and gluon PDFs. This is a direct consequence of the similarity of the photon PDF in all sets. Indeed, all comparisons are quite similar to those shown in Sect. 9.3 of Ref. [26], where the same processes were studied in pure QCD.

All in all, we conclude that the inclusion of QED corrections is important for precision phenomenology at the percent level, even in cases in which the photon-induced contribution is negligible, such as Higgs production in gluon fusion. Here neglecting the indirect effect of including the photon PDF results in an overestimation of the peak cross-section (and thus the total cross-section) by about 2%, thus biasing the prediction by an amount that is of the same order as the PDF uncertainty, but not included in it.

5 Summary and outlook

We have presented a new determination of QED PDFs based on the NNPDF4.0 set of parton distributions, using the methodology previously adopted for the construction of the NNPDF3.1QED PDFs [21]. This methodology implements the LuxQED procedure [17,18] to determine the photon PDF. Results are consistent with previous studies: specifically, we find that QED effects have a small but non-negligible impact mostly on the gluon PDF, and that photon-initiated contributions are most important for high-mass process such as neutral-current Drell–Yan production. This PDF determination is based on a new NNPDF pipeline for producing theory predictions [34]. Thanks to the integration of QED evolution in this pipeline, and in particular thanks to the use of the EKO evolution code [30], the production of QED variants of NNPDF determinations is now essentially automated, and will become the default in future releases. Indeed, there is in general no reason to switch off the photon PDF, even if it is a small correction, nor to neglect its effect on the momentum fraction carried by the gluon.

The NNPDF4.0QED PDF sets are made publicly available via the LHAPDF6 interface,
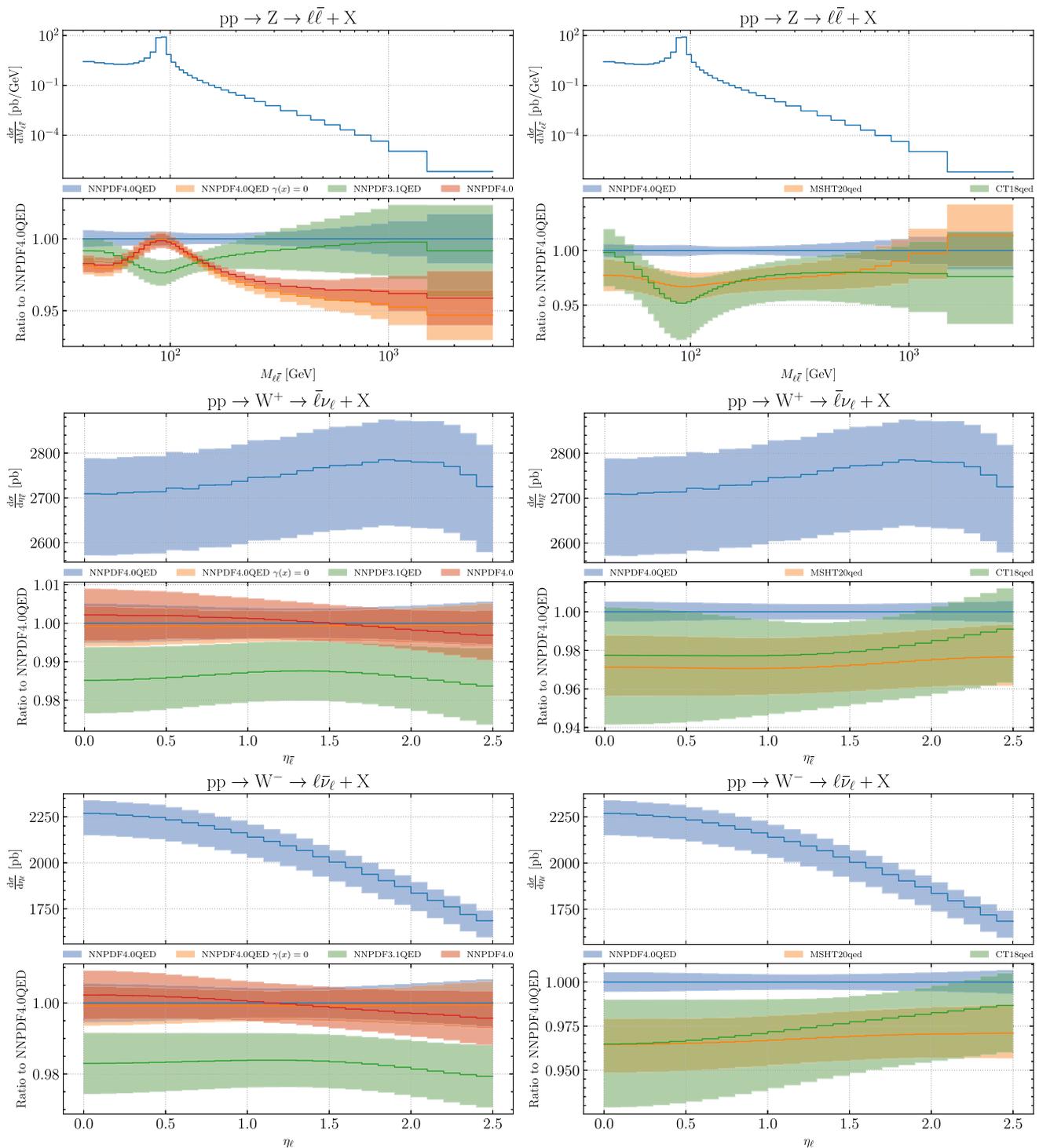
Fig. 9 Predictions for inclusive Drell–Yan production at the LHC with center-of-mass energy $\sqrt{s} = 14$ TeV, computed at NLO accuracy in the QCD and electroweak couplings. All uncertainties shown are PDF uncertainties only. From top to bottom: neutral-current dilepton production as a function of the dilepton invariant mass $m_{\ell\bar{\ell}}$; W^+ production as a function of the antilepton pseudo-rapidity $\eta_{\bar{\ell}}$; W^- production as a function of the antilepton pseudo-rapidity η_l. For each process, we display the absolute distributions (top panels) and the ratio to the central value obtained with NNPDF4.0QED PDFs, including photon-initiated channels (bottom panels). In the left panels the full NNPDF4.0 result is compared to NNPDF4.0 (QCD only), NNPDF3.1QED, NNPDF4.0QED (no photon-initiated); in the right panels it is compared to MSHT20QED [23] and CT18QED [25]. Note that the experimentally measurable quantity is shown, so for dilepton production the t- and u-channel photon-induced contribution is also included
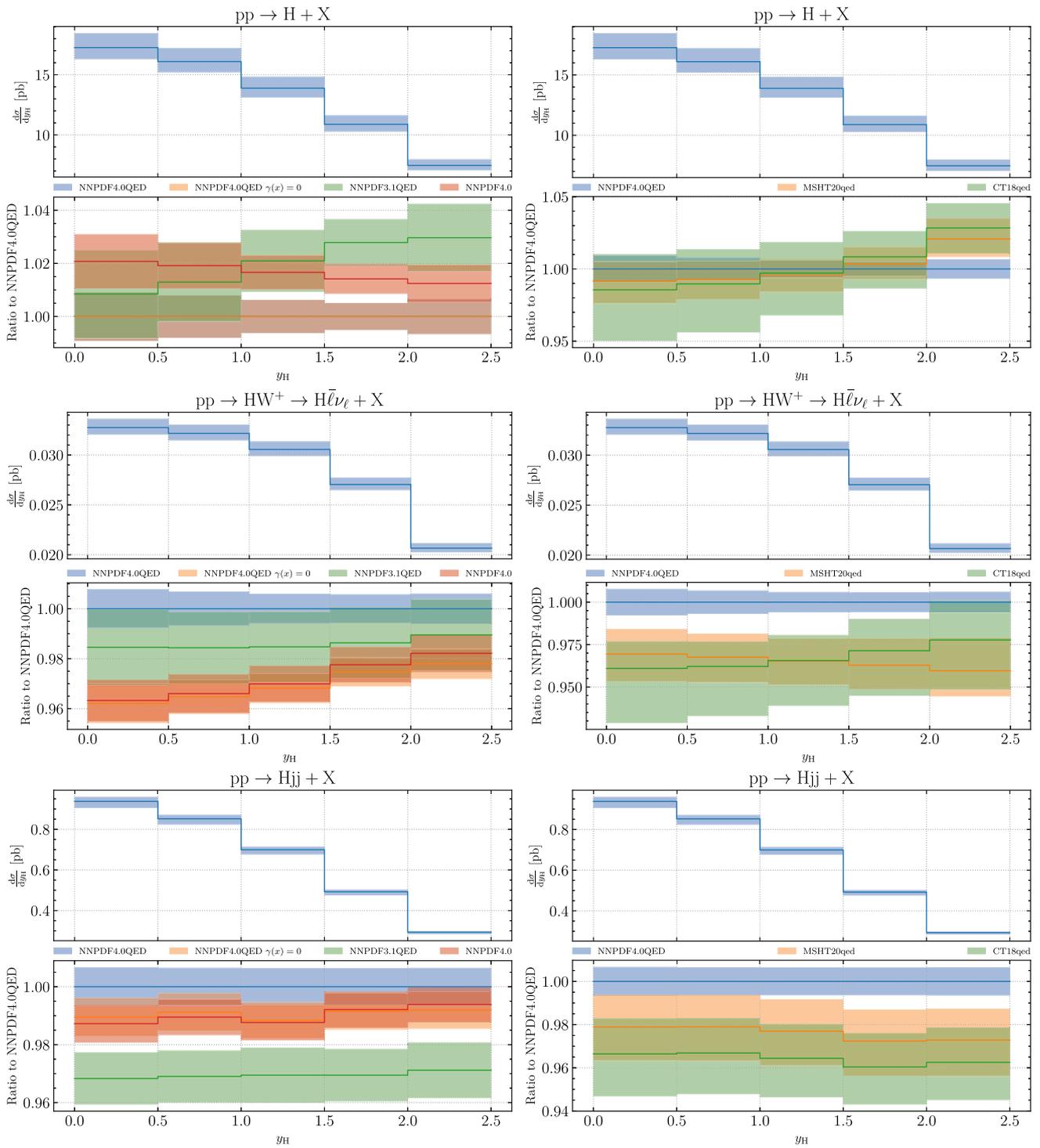
Fig. 10 Same as Fig. 9 but for the rapidity distribution of the Higgs for production in gluon-gluon fusion (top panel), in association with a W^+ boson (middle panel) and in vector-boson fusion (bottom panel)
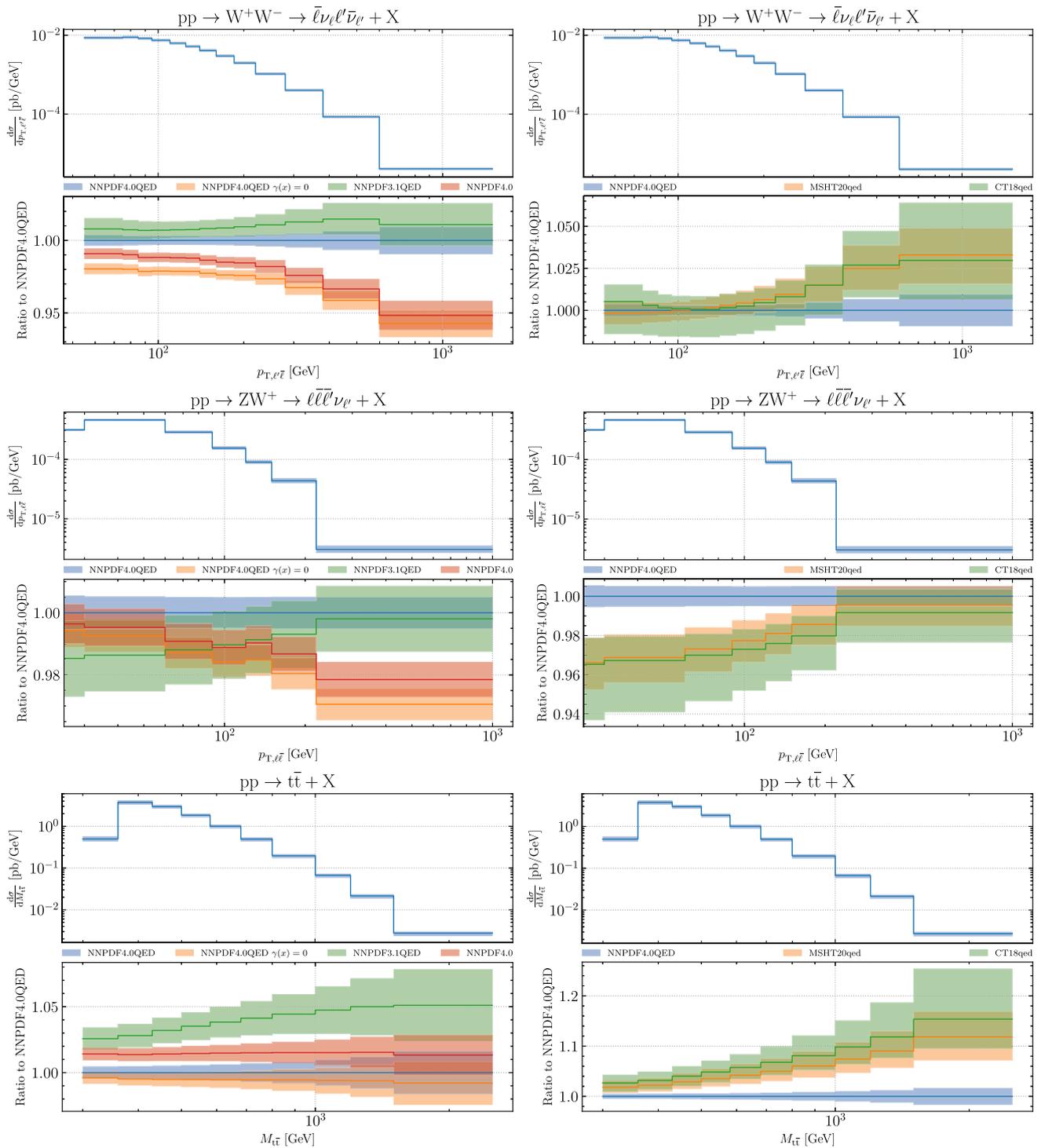
Fig. 11 Same as Fig. 9 but for the dilepton transverse momentum distribution for weak boson pair production (W^+W^- and W^+Z) and for the invariant mass distribution for top-quark pair production. No acceptance cuts on the decay productions of the W, Z bosons and top quark have been imposed

http://lhapdf.hepforge.org/ .

All sets are delivered as sets of $N_{rep} = 100$ Monte Carlo replicas. Specifically, we provide the NLO and NNLO global fits constructed with the settings defined in Sect. 2 and denoted as

```
NNPDF40_nlo_as_01180_qed
NNPDF40_nnlo_as_01180_qed
```

These sets are also made available via the NNPDF collaboration website

https://nnpdf.mi.infn.it/nnpdf4-0-qed/ .

They should be considered the QED PDF counterparts of the published NNPDF4.0 QCD-only PDF sets [26].

We also make available the set of 100 NNPDF4.0 (QCD only) replicas used to produce the comparisons shown in Sect. 3. These differ from the published NNPDF4.0 PDFs because they have been produced using the new theory pipeline, and include some minor bug corrections in the implementation of the NNPDF4.0 dataset. These are called:

```
NNPDF40_nlo_as_01180_qcd
NNPDF40_nnlo_as_01180_qcd.
```

The equivalence of these replicas to the published NNPDF4.0 replicas is demonstrated in Appendix A. They are made available for completeness, on the NNPDF website only.

The NNPDF4.0QED determination is part of a family of developments based on the NNPDF4.0 PDF set, and aimed at increasing its accuracy. These will also include a determination of the theory uncertainty on NNPDF4.0 PDFs based on the methodology of Refs. [48,49], and a first PDF determination based on NNPDF methodology at approximate N^3LO [50]. All of these, as well as their combination, will become part of the default NNPDF methodology in future releases.

A natural development of QED PDFs is the full inclusion of electroweak corrections in theory predictions used for PDF determination, which enables a widening of both the set of processes and the kinematic range that may be used for PDF determination, which are currently constrained by the need of ensuring that electroweak corrections are small. This is especially important as electroweak effects become more relevant in regions of phase space sensitive to the large-x PDFs, which are in turn relevant for new physics searches [51,52]. This development will be greatly facilitated by the availability, within the new pipeline, of the PINEAPPL interface to automated Monte Carlo generators, such as MG5_AMC@NLO. All these developments will help achieve PDF determination with percent or sub-percent accuracy.

Appendix A: The new NNPDF theory pipeline

As mentioned in Sect. 3, this work is based on a new pipeline for the calculation of theoretical predictions. This new theory pipeline is described in Ref. [34]; it supersedes and replaces the one used for the NNPDF4.0 determination, which was based on a combination of different pieces of code of various origin, specifically APFEL [28] for PDF evolution and DIS structure function computation, and APFELGRID [29] for the generation of interpolation grids of NLO partonic matrix elements. This last piece of code made use of APPLGRID [53] and FASTNLO [54] interpolators. The main benefit of the new pipeline is its unified, yet modular and flexible, structure.

This theory pipeline has been extensively benchmarked for numerical accuracy, including QCD evolution, the computation of structure functions, the interpolation of grids, and the interfacing to MG5_AMC@NLO. Using this new pipeline, interpolation grids for a number of processes have been recomputed, as discussed below. Therefore Tables 2.1–2.5 in Ref. [26] should be updated accordingly.

- All fully inclusive DIS processes (Table 2.1 in Ref. [26]) have been recomputed using YADISM [31,55]. Specifically, in this code heavy quark mass effects are accounted for using a new implementation of the FONLL prescription [56].
- All fixed-target Drell–Yan processes (see Table 2.3 in Ref. [26]) have been recomputed using a modified version [34] of VRAP [57].
- The following datasets of Table 2.4 and Table 2.5 in Ref. [26] have been recomputed using MG5_AMC@NLO (all grids, in the PINEAPPL format, are available at https://github.com/NNPDF/pineapplgrids):

  – ATLAS W, Z 7 TeV ($\mathcal{L} = 4.6$ fb^{-1})
  – ATLAS low-mass DY 7 TeV
  – ATLAS high-mass DY 7 TeV
  – ATLAS $\sigma_{tt}^{\mathrm{tot}}$ 7, 8
  – ATLAS $\sigma_{tt}^{\mathrm{tot}}$ 13 ($\mathcal{L} = 139$ fb^{-1})
  – ATLAS $t\bar{t}$ lepton+jets 8 TeV
  – ATLAS $t\bar{t}$ dilepton 8 TeV
  – CMS Drell–Yan 2D 7 TeV
  – CMS $t\bar{t}$ 2D dilepton 8 TeV
  – CMS $t\bar{t}$ lepton+jets 13 TeV
  – CMS $t\bar{t}$ dilepton 13 TeV
  – CMS $\sigma_{tt}^{\mathrm{tot}}$ 5.02 TeV
  – CMS $\sigma_{tt}^{\mathrm{tot}}$ 7, 8 TeV
  – CMS $\sigma_{tt}^{\mathrm{tot}}$ 13 TeV
  – LHCb Z 7 TeV ($\mathcal{L} = 940$ pb^{-1})
  – LHCb $Z \rightarrow ee$ 8 TeV ($\mathcal{L} = 2$ fb^{-1})
  – LHCb W, $Z \rightarrow \mu$ 7 TeV
  – LHCb W, $Z \rightarrow \mu$ 8 TeV

Grids for all remaining datasets have been converted to the PINEAPPL format without recomputing them. Of course, none of these changes except the new FONLL implementation should make any difference to the extent that the previous and new theory implementations are both numerically accurate. The new FONLL implementation differs from the previous one by subleading corrections and thus it does introduce NNLO differences in the NLO fit and N³LO differences in the NNLO fit; these differences are confined to the charm and bottom mass corrections to deep-inelastic structure functions and thus only affect a small number of datapoints, at NNLO at the sub-percent level.

Finally, in the process of transitioning to the new pipeline, a few bugs were discovered in the implementation of a few datapoints, such as incorrect normalization, incorrect scale assignment or incorrect bin size. These corrections in practice have a negligible impact, as they involve a handful of points out of more than 4500.

We have generated variants of the published NNPDF4.0 NLO and NNLO fits using the new theory pipeline. In order to illustrate the equivalence between these fits and the NNPDF4.0 fits, obtained with the previous theory pipeline, we compare each pair of fits. All fits are made of 100 replicas. The corresponding statistical estimators are reported in Table 2. In addition to the fit quality estimators, we also show the ϕ estimator, defined in Eq. (4.6) of Ref. [58], which measures the ratio of the average (correlated) PDF uncertainty on datapoints over the experimental uncertainty, and is thus a fairly sensitive measure of fit quality. All estimators are seen to coincide at NNLO (note that the statistical uncertainty on the χ^2 with the given number of datapoints is 0.02), demonstrating explicitly the equivalence between the two replica sets. All estimators are also very close at NLO: differences are compatible with the statistical uncertainty on the χ^2.

Parton distributions are compared in Fig. 12, which displays the distances (defined in Ref. [45]) between central values and uncertainties computed using the two replica sets at $Q = 100$ GeV. For a set of 100 replicas, $d \sim 1$ indicates statistical equivalence, while $d \sim 10$ indicates one-sigma differences. At NNLO, distances fluctuate about 1, indicating that the two replica sets are drawn by the same underlying distribution. At NLO, distances are somewhat larger, around the half-sigma level for most PDFs and reaching 1σ in some cases; this is mostly due to the different FONLL implementation [56]. The PDFs themselves are shown in Figs. 13 and 14, and the parton luminosities for the LHC at $\sqrt{s} = 14$ TeV are shown in Figs. 15 and 16, respectively at NLO and NNLO. The complete agreement between the two pairs of NNLO replica sets, and the small differences at NLO, are manifest.

Appendix B: Solution of evolution equations

We summarize results and benchmarks concerning the solution to evolution equations. First, we comparatively review solutions to evolution equations, for which a different choice is made in this work in comparison to the previous NNPDF3.1QED and NNPDF4.0 PDF determinations: namely the use of an exact (EXA) solution instead of the truncated (TRN) solution [30]. We then study the impact of this different choice on PDFs. Having established the equivalence of (pure QCD) PDFs obtained using either solution, we perform an inversion test of the exact QED×QCD solution implemented in EKO, which we adopt, in order to verify its accuracy. We finally benchmark this EKO QED×QCD solution against the implementation in the APFEL code, which was used in the previous NNPDF3.1QED PDF determination.

B.1 Exact vs. expanded solutions: formal aspects

Perturbative solutions of QCD evolution equations of the form

Table 2 Statistical indicators (defined as in Table 1 for a set of 100 NNPDF4.0 NNLO PDF replicas produced using the new theory pipeline compared to published NNPDF4.0 NNLO 100 replica set. The ϕ estimator, defined in Ref. [58], Eq. (4.6), is also shown (see text)

	NLO QCD		NNLO QCD	
	New	Published	New	Published
χ^2	1.26	1.24	1.17	1.16
$\langle E_{\rm tr} \rangle_{\rm rep}$	2.41 ± 0.06	2.43 ± 0.08	2.28 ± 0.05	2.27 ± 0.07
$\langle E_{\rm val} \rangle_{\rm rep}$	2.57 ± 0.10	2.62 ± 0.13	2.37 ± 0.11	2.35 ± 0.11
$\langle \chi^2 \rangle_{\rm rep}$	1.29 ± 0.02	1.27 ± 0.02	1.20 ± 0.02	1.18 ± 0.02
$\langle {\rm TL} \rangle_{\rm rep}$	12900 ± 2000	13200 ± 2100	12400 ± 2600	13400 ± 2400
ϕ	0.156 ± 0.006	0.178 ± 0.007	0.153 ± 0.005	0.162 ± 0.005
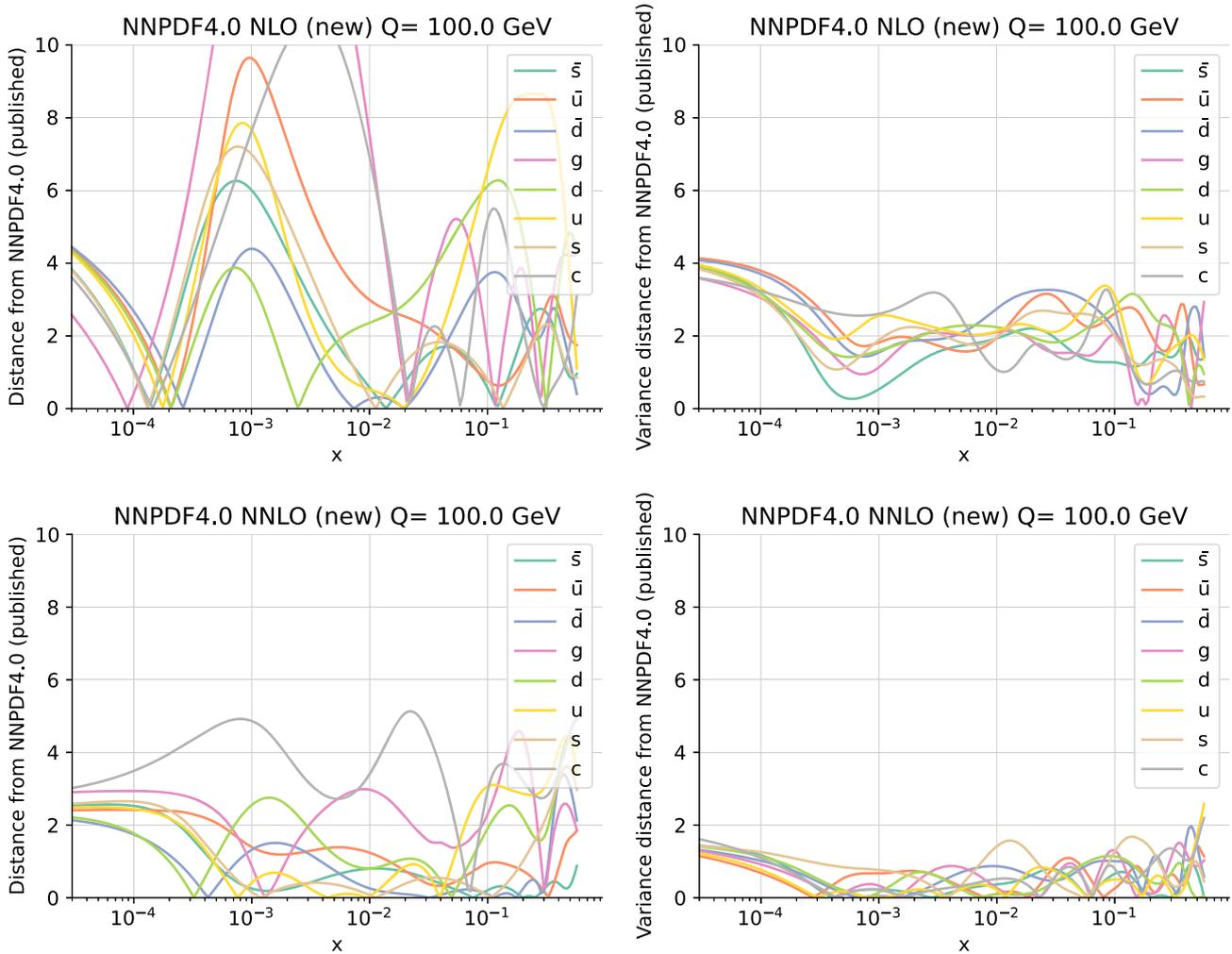
Fig. 12 Distances at $Q = 100$ GeV between the central values (left) and uncertainties (right) of the 100 NNPDF4.0 PDF replicas at NLO (top) and NNLO (bottom) whose statistical indicators are compared in Table 2

$$\mu^2 \frac{df_i}{d\mu^2} = -\left(a_s \gamma_{ij}^{(1,0)} + a_s^2 \gamma_{ij}^{(2,0)} + \cdots \right) f_j, \qquad \text{(B.1)}$$

are based on the observation that if the equation is solved exactly at leading order, with the leading-order term in the beta function Eq. (2.3), then the solution $f_{\rm LO}(Q^2)$ in

terms of a boundary condition $f_{\rm LO}(Q_0^2)$ is a pure leading log (LL) function, i.e. it is a function of $a_s L$ only, where $L = \ln(Q^2/Q_0^2)$, rather than depending on a_s and L separately. However, if the next-to leading order contributions $\gamma_{ij}^{(2,0)}$ and β_1 are also included, then the solution is next-to-leading log (NLL) accurate, but the exact solution Eqs. (2.16)
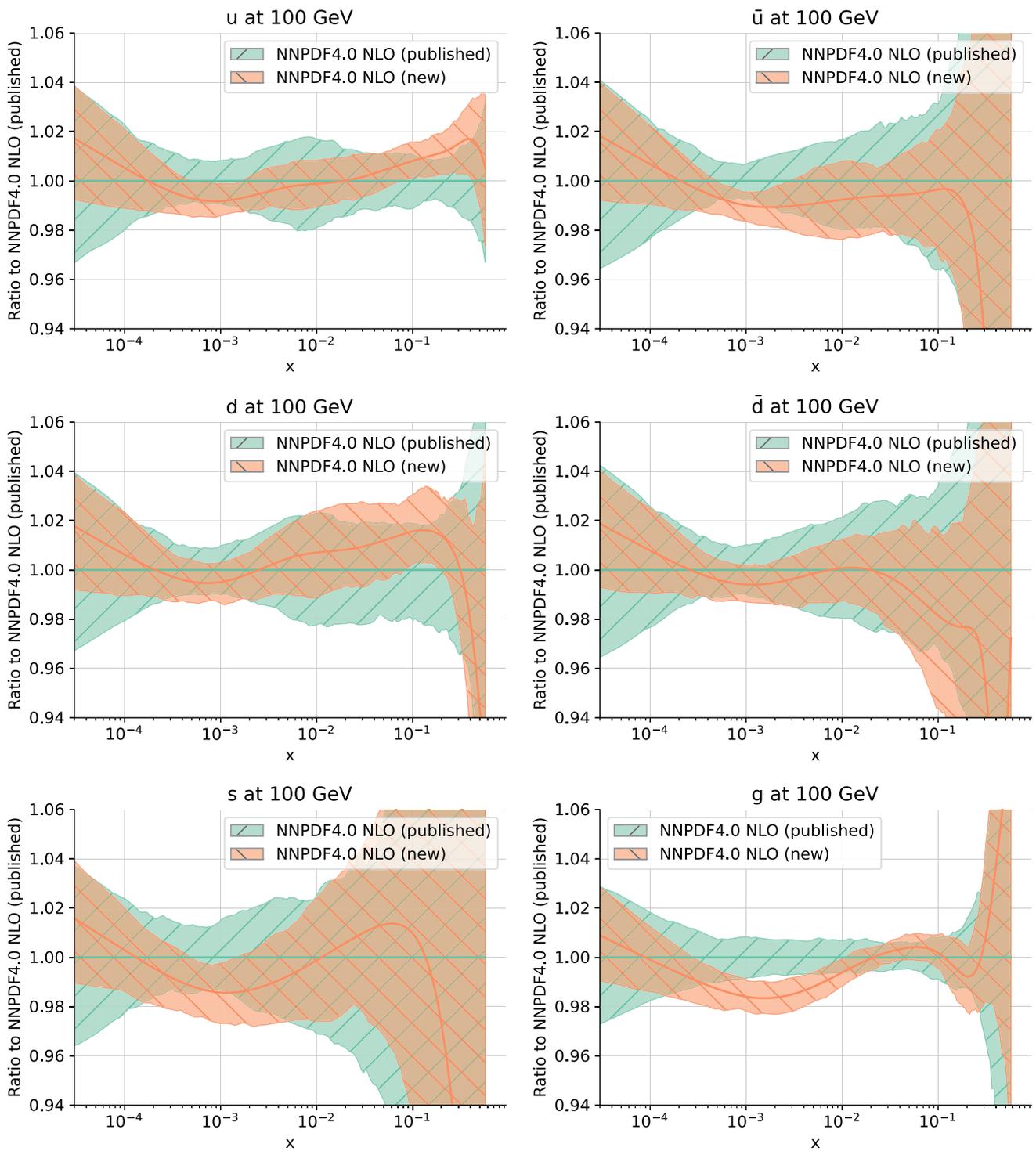
Fig. 13 The NLO PDFs at $Q = 100$ GeV from the 100 replica sets compared in Table 2 and Fig. 12. Results are shown normalized to the central value of the published set. Bands correspond to 1σ uncertainties. From left to right and from top to bottom, we show the up, anti-up, down, anti-down, strange and gluon PDFs

and (2.17) includes terms to all logarithmic orders. A solution that only includes NLL terms, namely with the structure

$$f = f_{\text{LO}} + a_s f_{\text{NLO}}, \qquad (\text{B.2})$$

where both f_{LO} and f_{NLO} are pure LL functions, can be constructed by linearization of the exact solution, namely by expanding the exact solution and by truncating the expansion. Various intermediate options are also possible. The argument
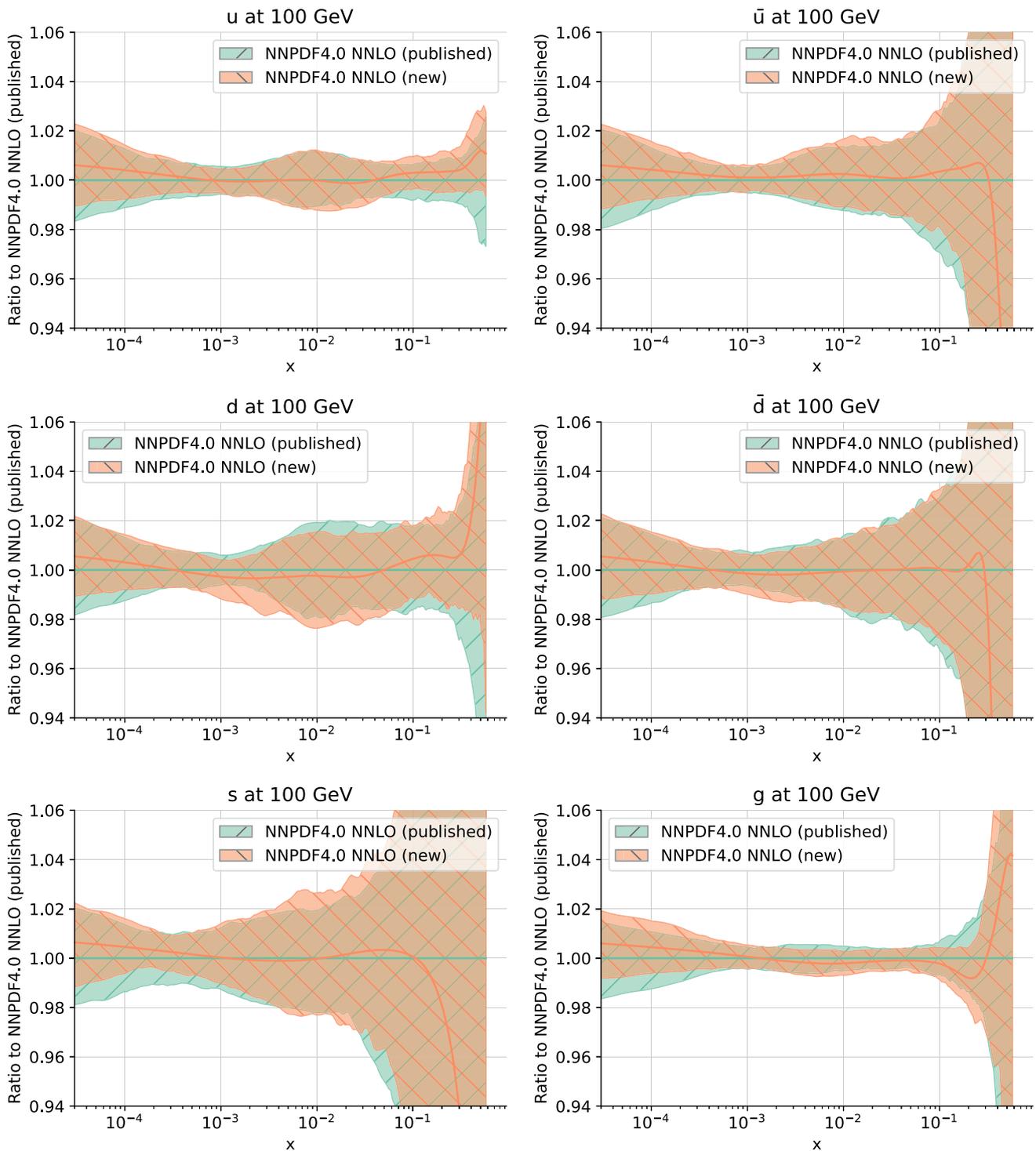
Fig. 14 Same as Fig. 13 at NNLO

can be repeated at any logarithmic accuracy. Of course all these solutions are equivalent up to subleading logarithmic terms.

A truncated solution to the pure QCD evolution equation of the form Eq. (B.2) can be determined in closed form to all

orders in Mellin space [59], by diagonalizing order by order the anomalous dimension matrix. This corresponds to the solution referred to as truncated in Ref. [30]. However, this strategy fails for combined QED×QCD evolution, because the QED and QCD anomalous dimension matrices do not
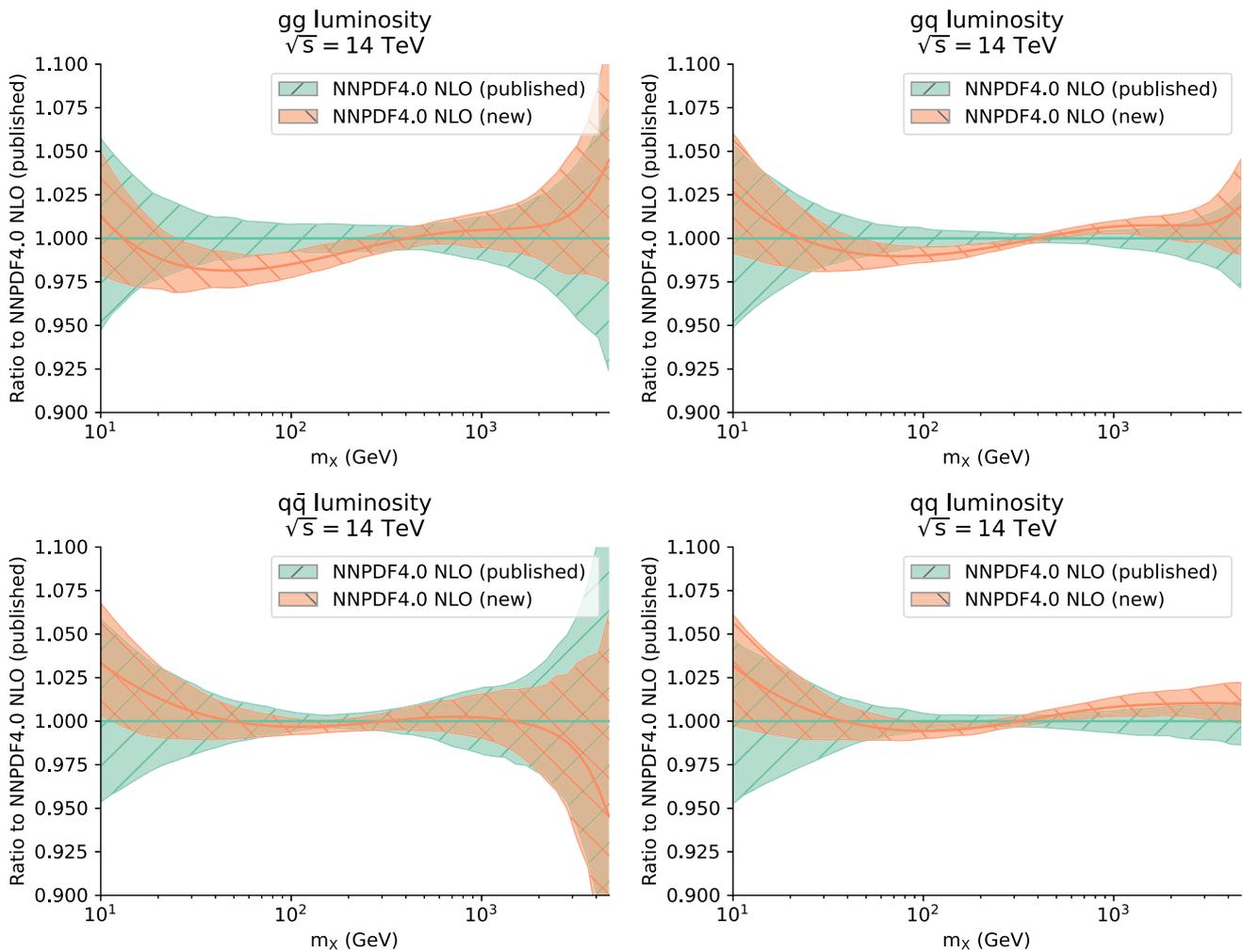
Fig. 15 Same as in Fig. 13 for the partonic luminosities at the LHC with $\sqrt{s} = 14$ TeV as a function of the invariant mass m_X. The gg (left) and qg (right) luminosities are shown in the top row, the $q\bar{q}$ (left), and qq (right) luminosities in the bottom row

commute, and thus cannot be diagonalized simultaneously. Because a_s and a_{em} depend on scale in different ways, this implies that the anomalous dimension matrices γ_{ij} evaluated at different scales do not commute:

$$
\left[\gamma\left(N, a_s(\mu^2), a_{em}(\mu^2)\right), \gamma\left(N, a_s(\mu'^2), a_{em}(\mu'^2)\right)\right] \neq 0
$$
$$
\text{if} \quad \mu \neq \mu'. \tag{B.3}
$$

This means that the LO solution, constructed only including the LO contributions to the anomalous dimensions, takes the form of a path-ordered exponential

$$
f_{\text{LO}}(Q^2) = \mathcal{P}
$$
$$
\times \exp\left(-\int_{Q_0^2}^{Q^2} \frac{d\mu^2}{\mu^2}\left(a_s(\mu^2)\gamma^{(1,0)} + a_{em}(\mu^2)\gamma^{(0,1)}\right)\right)\bigg|_{\text{LO}} (Q_0^2)
$$
$$
\tag{B.4}
$$

that cannot be written in closed form. Note that the commutator terms are not subleading: for example, the quadratic term is proportional to $a_s a_{em} L^2$, so it is a LL contribution.

The problem persists to all orders, and requires a truncated solution to be also given as a path-ordered exponential. For instance, including NLO QCD and LO QED contributions, the evolution equation is

$$
\mu^2 \frac{df_i}{d\mu^2} = -\left(a_s \gamma_{ij}^{(1,0)} + a_{em}\gamma_{ij}^{(0,1)} + a_s^2 \gamma_{ij}^{(2,0)}\right) f_j. \tag{B.5}
$$

The perturbative NLO QCD solution then has the form Eq. (B.2), where in the general case f_{LO} and f_{NLO} are LO accurate in QCD and QED, but may also include subleading contributions, i.e. $f_{\text{N}^k\text{LO}} = f_{\text{N}^k\text{LO}}(a_s L, a_{em}L)[1 + \mathcal{O}(a_s, a_{em})]$. Substituting the perturbative expansion Eq. (B.2) in the NLO Eq. (B.5), and using the LO solution, Eq. (B.4), leads to

$$
\mu^2 \frac{df_{\text{NLO},i}}{d\mu^2} = -a_s\left(\gamma_{ij}^{(1,0)} - \beta_{\text{QCD}}^{(2,0)}\delta_{ij}\right) f_{\text{NLO},j}
$$
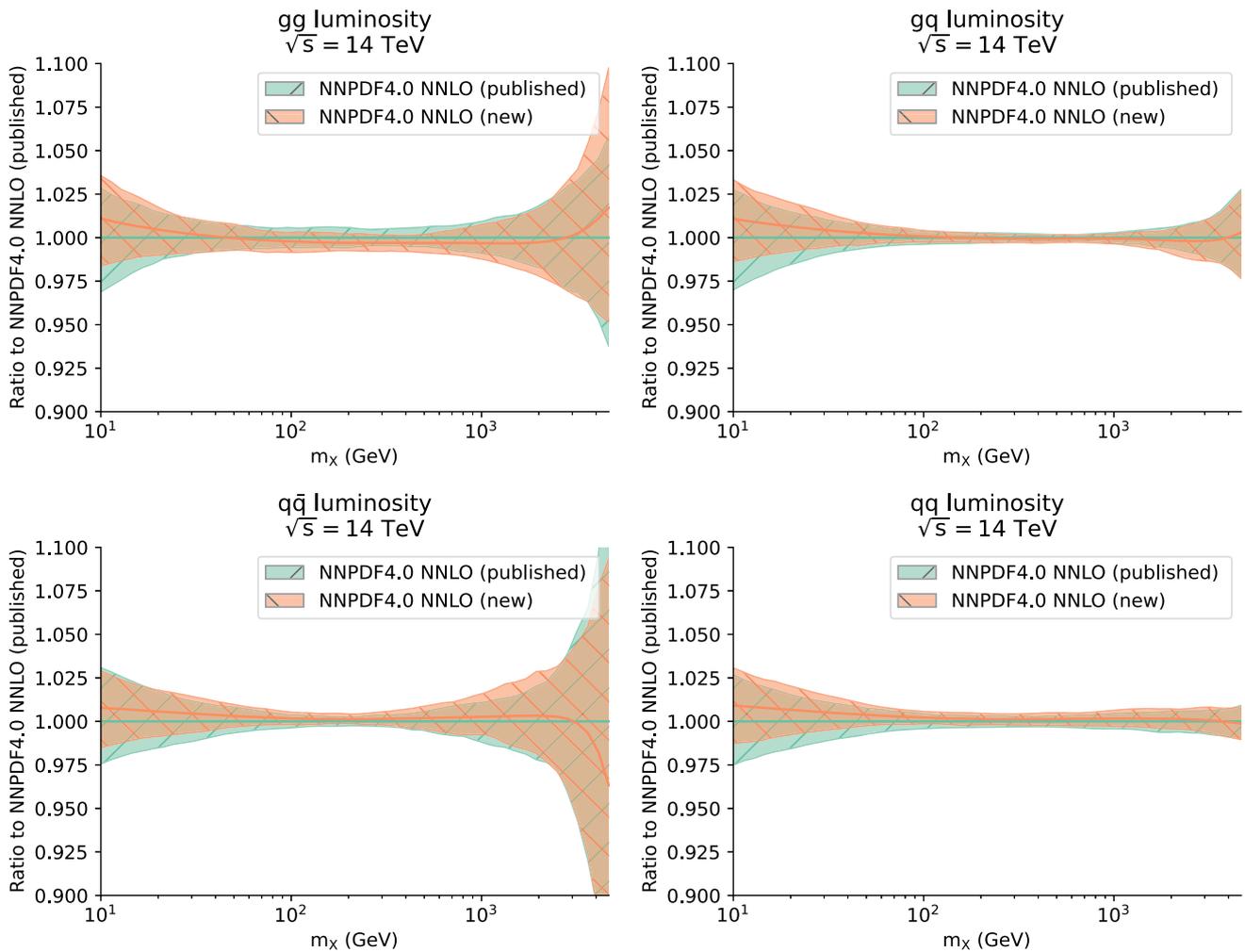
Fig. 16 Same as in Fig. 15 at NNLO

$$-a_s\gamma_{ij}^{(2,0)}f_{\mathrm{LO},j} - a_{em}\gamma_{ij}^{(0,1)}f_{\mathrm{NLO},j}. \qquad (B.6)$$

Again, due to the non-commutativity of $\gamma_{ij}^{(1,0)}$ and $\gamma_{ij}^{(0,1)}$, the solution for f_{NLO} is a path-ordered exponential and cannot be given in closed analytic form. This continues to be the case at higher orders.

A truncated solution can be constructed numerically, by starting with the exact path-ordered solution Eqs. (2.16–2.17), and then expanding numerically. Such an approach, however, involves approximating higher-order derivatives by finite differences [28], which may lead to numerical instabilities. This method was adopted in APFEL: APFEL is an x-space evolution code, so the analytic truncated solution of Ref. [59] cannot be used, and the expansion is performed numerically anyway, even in the case of pure QCD. On the other hand, in the EKO evolution code the pure QCD truncated solution is implemented through the analytic solution of Ref. [59], while the numerical path-ordered solution only needs to be used for the exact solution Eqs. (2.16) and (2.17). Therefore, in order to ensure greater accuracy, we adopt the

exact solution for the QED×QCD case. The impact of this choice is benchmarked in the next subsection.

B.2 Exact vs. expanded solution: impact on PDFs

We wish to quantify the impact of the choice of solution: exact, Eqs. (2.16–2.17), vs. truncated, based on an expansion of the form of Eq. (B.2). To this goal, we have determined a set of NNPDF4.0 NNLO (pure QCD) PDF replicas, but now using the exact solution, instead of the truncated solution used for the published NNPDF4.0 PDF sets. We have used the new theory pipeline.

The statistical indicators for these two sets of replicas are compared in Table 3, and are seen to be indistinguishable. The corresponding PDFs are compared in Fig. 17, where we display the distances (defined in Ref. [45]) between the central values and uncertainties of all PDFs in either set, and in Fig. 18, where we specifically compare the gluon and antiup PDFs at $Q = 100$ GeV. It is clear that for $x \gtrsim 10^{-3}$ the PDFs are statistically indistinguishable. At small x they
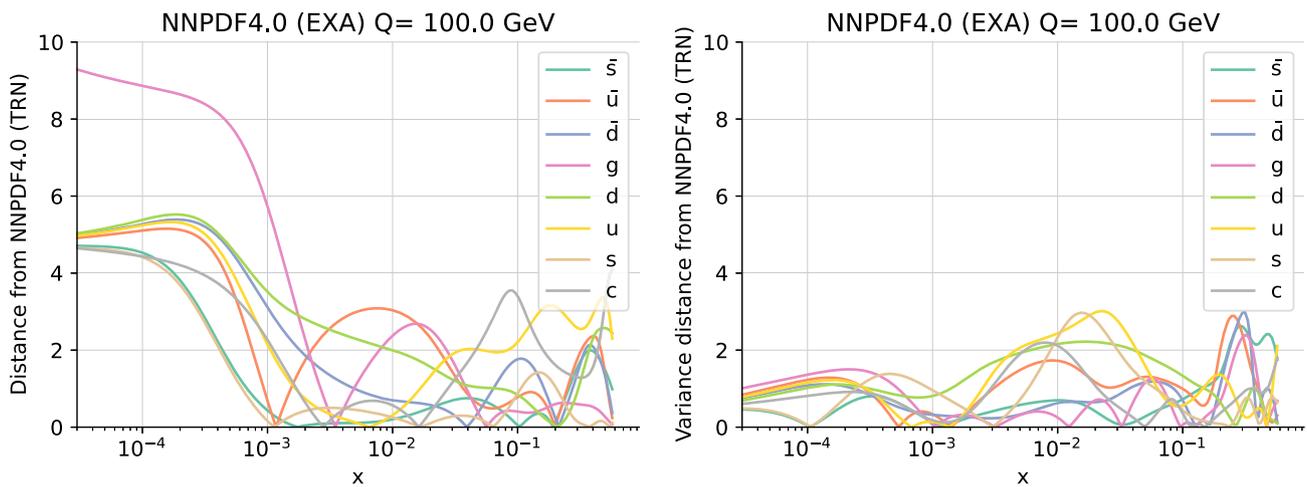
Fig. 17 Statistical distances between the central values (left) and uncertainties (right) of the 100 NNPDF4.0 PDF replicas whose statistical indicators are compared in Table 3
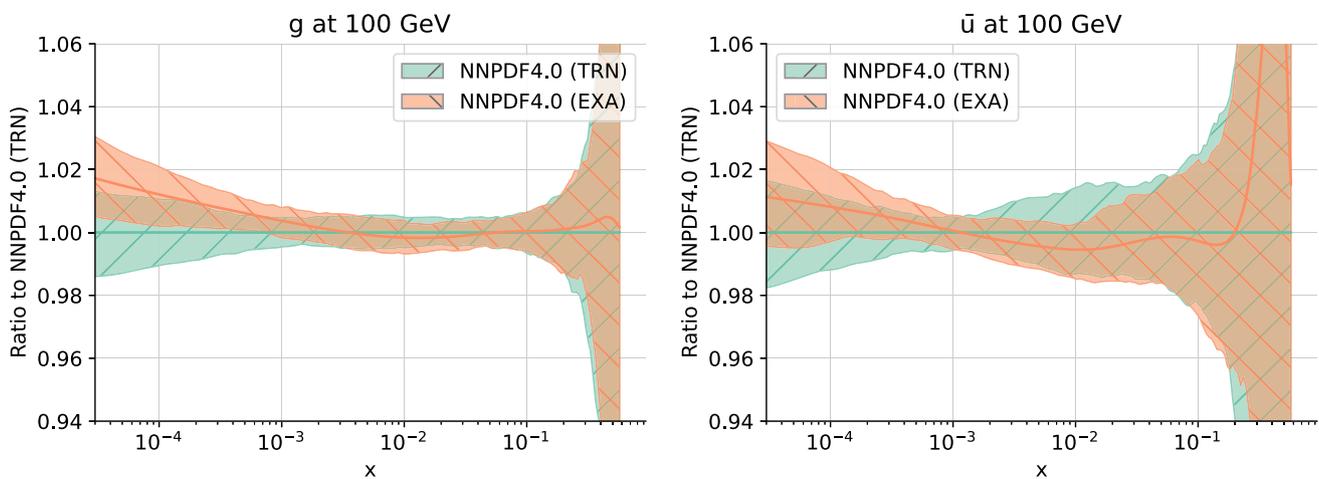
Fig. 18 The gluon (left) and anti-up quark (right) PDFs at $Q = 100$ GeV from the PDF sets compared in Table 3 and Fig. 17. Results are shown normalized to the central value of the PDFs obtained with truncated evolution (published NNPDF4.0)

Table 3 Same as Table 2, now comparing a set of 100 NNPDF4.0 replicas obtained using the default truncated solution of evolution equations with a set obtained using the exact solution (pure NNLO QCD theory). Both PDF sets are produced using the new theory pipeline

	NNPDF4.0 NNLO QCD	
	Exact solution	Truncated solution
χ^2	1.17	1.17
$\langle E_{tr}\rangle_{rep}$	2.26 ± 0.06	2.28 ± 0.05
$\langle E_{val}\rangle_{rep}$	2.34 ± 0.10	2.37 ± 0.11
$\langle \chi^2\rangle_{rep}$	1.19 ± 0.01	1.20 ± 0.02
$\langle TL\rangle_{rep}$	12100 ± 2500	12400 ± 2600
ϕ	0.147 ± 0.005	0.153 ± 0.005

start differing at the half-σ level, with differences at most reaching the one-σ level for the gluon.

This is consistent with the fact that the exact and truncated solutions differ by higher-order perturbative corrections that

go beyond the NNLO accuracy of the computation. Indeed, it is well known that at small x the perturbative convergence starts deteriorating because of high-energy logarithms that need resummation in order for PDF determination to be accurate [60]. The qualitative behavior of the PDFs shown in Fig. 18 agrees with this explanation: the exact solution exponentiates a set of subleading small-x logarithms which are linearized in the truncated solution. It follows that in the small-x region, where there is no data, these contributions lead to a stronger rise of the gluon, which then feeds back onto the quark-antiquark sea.

B.3 Exact solution: invertibility

The unexpanded EKO Eq. 2.17 manifestly satisfies the exact inversion property
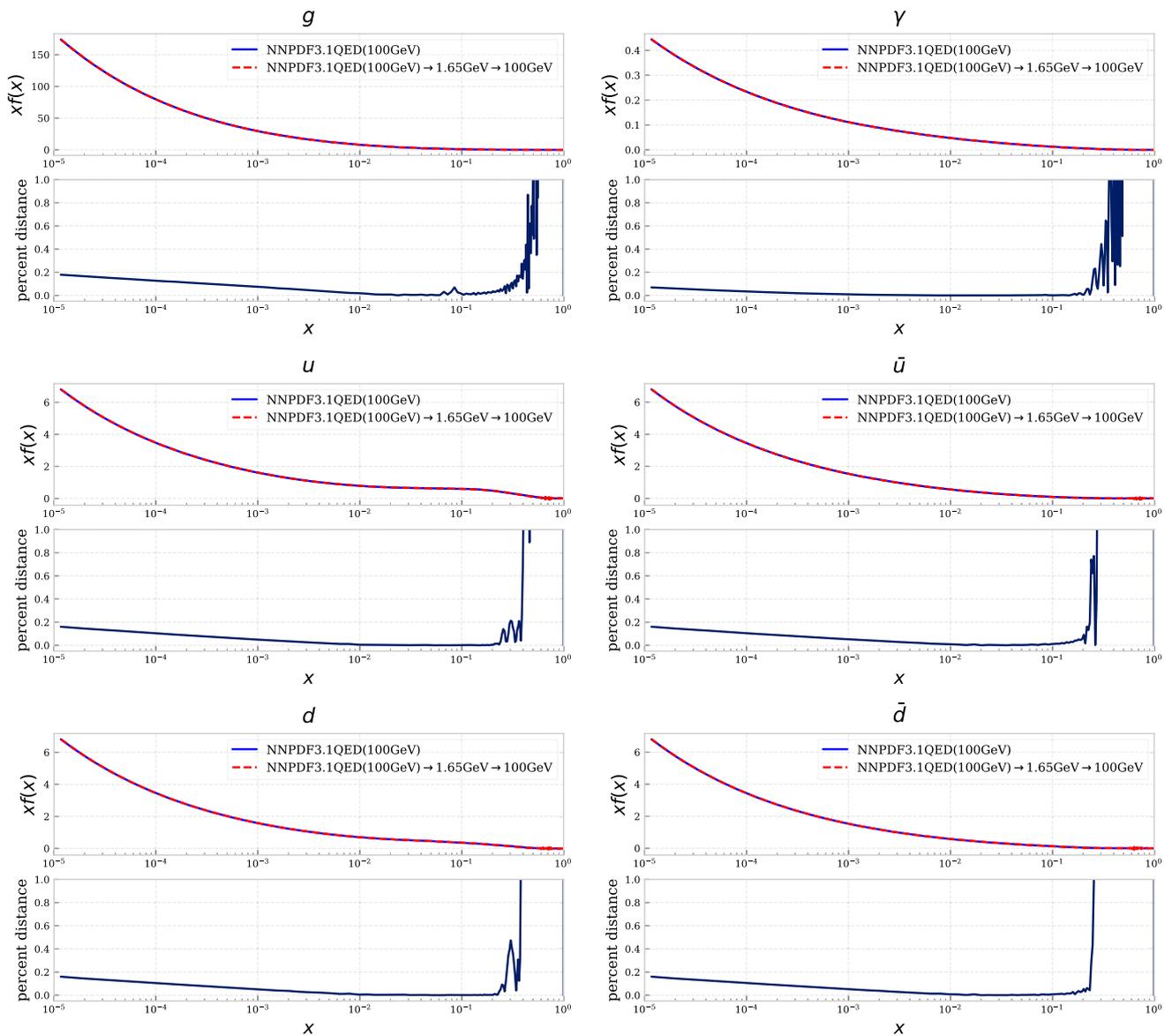
Fig. 19 Comparison between the NNPDF3.1QED PDFs at $Q_a = 100$ GeV, with the results of evolution of the same PDFs to $Q_b = 1.65$ GeV followed by evolution back to $Q_a = 100$ GeV. We show, from top to bottom, the gluon, up and down (left) and the photon, anti-up and anti-down (right). In each case we show both the pair of PDFs, and their percentage relative difference

$$E(Q^2 \leftarrow Q_0^2)E(Q_0^2 \leftarrow Q^2) = \mathbb{1}. \tag{B.7}$$

Hence, checking that evolving a PDF back and forth between two scales gives back the starting PDF provides a stringent test of the implementation of evolution equations and its accuracy.

We have performed this check, by starting with the NNPDF3.1QED NNLO PDFs at $Q_a = 100$ GeV, then evolving down to $Q_b = 1.65$ GeV and back to $Q_a = 100$ GeV. Note that this evolution crosses back and forth the bottom quark threshold, so this also checks the accuracy of the inversion of matching conditions between the $N_f = 4$ and $N_f = 5$

flavor schemes. Results are displayed in Fig. 19, where we compare the initial and final PDFs. Differences are at most at the permille level, except at very large x where they can reach the percentage level, but PDFs are becoming rapidly very small. We conclude that exact evolution as implemented in the EKO code has (at least) this permille accuracy.

B.4 Benchmarking of QCD×QED evolution: EKO vs. APFEL

We finally compare the implementation of QCD×QED evolution in the EKO code used in this paper, with the
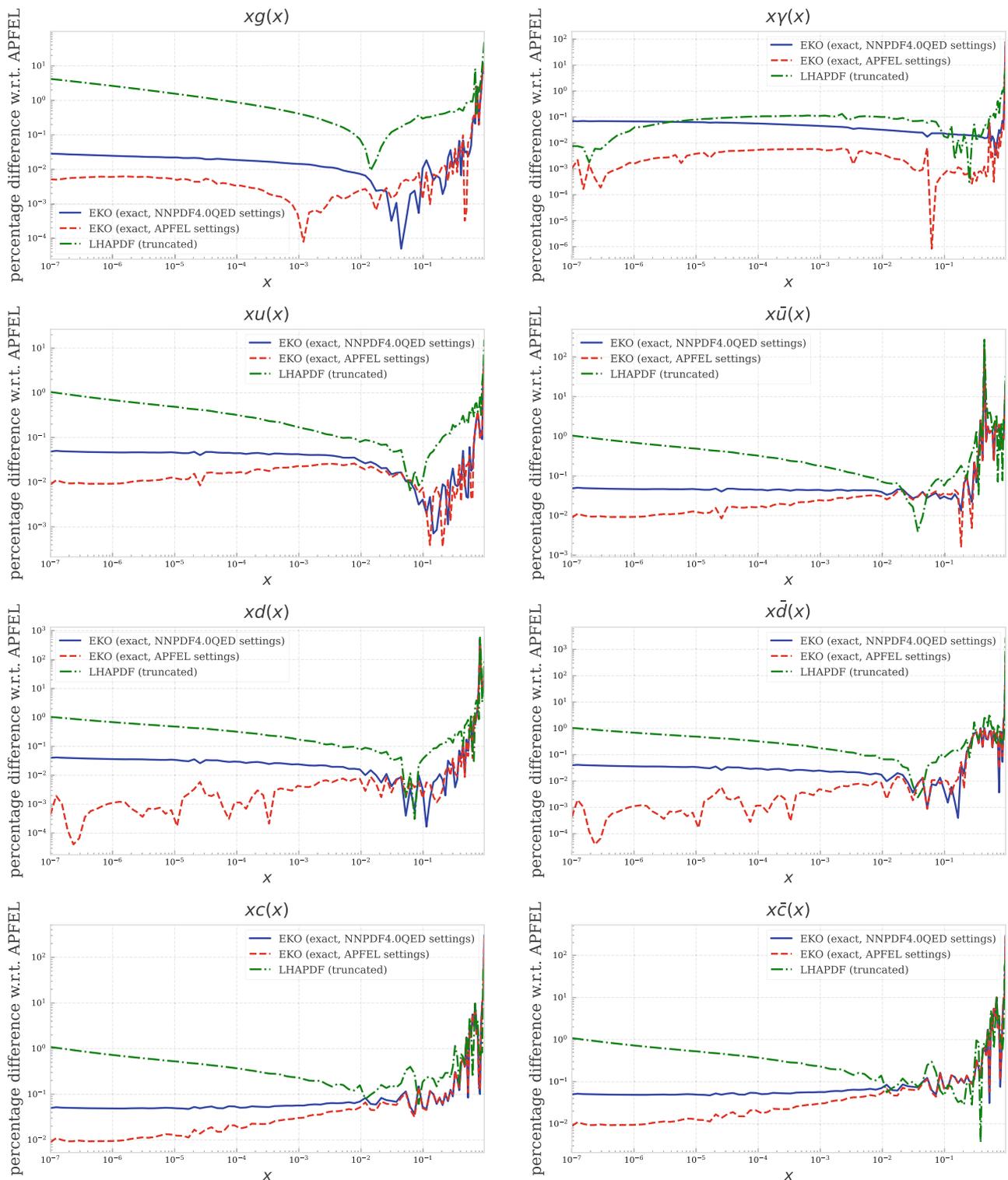
Fig. 20 Percentage difference between pairs of PDFs at $Q = 100$ GeV obtained evolving NNPDF3.1QED PDFs from $Q_0 = 1.65$ GeV with different implementation of the QED$\times$QCD evolution. From top to bottom the gluon, up, down and charm (left), photon, antiup, antidown and anticharm (right) are shown. The three curves compare: APFEL exact vs. truncated evolution (green, dot-dashed); APFEL vs. EKO exact with in each case default settings for the running of the couplings (see text) (blue, solid); APFEL vs. EKO exact both with APFEL settings for the running of the coupling (red, dashed). Note the logarithmic scale on the y axis; note also that the range on the y axis for gluon plot differs from that of all other PDFs

APFEL code used for the NNPDF3.1QED PDF determination. The comparison is performed by taking as input the NNPDF3.1QED NNLO PDF set at the initial parametrization scale of $Q_0 = 1.65$ GeV, evolving to $Q = 100$ GeV and determining the percentage difference for all PDFs.

Results are shown in Fig. 20 for three pairwise comparison. First, we compare APFEL exact vs. truncated evolution (green curve); the APFEL truncated result is the published NNPDF3.1 PDF set, as given by public LHAPDF grids. Then we compare EKO vs. APFEL evolution using the exact solution in both cases, and the default settings of either code as respectively used in this work and for the NNPDF3.1QED PDF set (blue curve). These settings differ in the running of the couplings: in the APFEL settings the coefficients $\beta_{\mathrm{QCD}}^{(2,1)}$, $\beta_{\mathrm{QED}}^{(0,3)}$ and $\beta_{\mathrm{QED}}^{(1,2)}$ in Eqs. (2.3) and (2.4) are neglected, i.e., the two equations are decoupled and a_{em} runs at leading order. Finally, we compare EKO vs. APFEL evolution using the same solution and the same settings (red curve), namely exact solution and running of the couplings as in APFEL. Note that the scale on the y axis is logarithmic, and that it shows percentage difference (so 10^{-3} denotes a relative difference of 10^{-5}).

The percentage differences between EKO and APFEL with common settings are always below 10^{-2}, for all PDFs except at very large x and for charm, where they are about a factor 10 larger. This sets the accuracy of the evolution codes that are being compared. The impact of the different running of the couplings is moderate: it increases the percentage difference by about a factor 3, and then only for the photon for all x values, for other PDFs only at small $x \lesssim 10^{-2}$. Even so, the difference between EKO and APFEL with their respective settings is at the sub-permille level. The difference between exact and truncated evolution is at the permille level for $x \gtrsim 0.003$, and it can grow up to a few percent for the gluon and a few permille for all other PDFs at very small $x \sim 10^{-6}$, in agreement with what already discussed when comparing exact and truncated evolution at the level of PDFs in Fig. 18.

References

1. X. Cid Vidal et al., Report from Working Group 3: beyond the standard model physics at the HL-LHC and HE-LHC. CERN Yellow Rep. Monogr. **7**, 585–865 (2019). arXiv:1812.07831
2. HL-LHC, HE-LHC Working Group Collaboration, P. Azzi et al., Standard Model Physics at the HL-LHC and HE-LHC. arXiv:1902.04070
3. Physics of the HL-LHC Working Group Collaboration, M. Cepeda et al., Higgs Physics at the HL-LHC and HE-LHC. arXiv:1902.00134
4. G. Heinrich, Collider physics at the precision frontier. Phys. Rep. **922**, 1–69 (2021). arXiv:2009.00516
5. V. Bertone, S. Carrazza, D. Pagani, M. Zaro, On the impact of lepton PDFs. JHEP **11**, 194 (2015). arXiv:1508.07002
6. B. Fornal, A.V. Manohar, W.J. Waalewijn, Electroweak Gauge boson parton distribution functions. JHEP **05**, 106 (2018). arXiv:1803.06347
7. C.W. Bauer, N. Ferland, B.R. Webber, Standard model parton distributions at very high energies. JHEP **08**, 036 (2017). arXiv:1703.08562
8. M. McCullough, J. Moore, M. Ubiali, The dark side of the proton. JHEP **08**, 019 (2022). arXiv:2203.12628
9. L. Buonocore, U. Haisch, P. Nason, F. Tramontano, G. Zanderighi, Lepton-quark collisions at the large hadron collider. Phys. Rev. Lett. **125**, 23 231804 (2020). arXiv:2005.06475
10. L. Buonocore, P. Nason, F. Tramontano, G. Zanderighi, Photon and leptons induced processes at the LHC. JHEP **12**, 073 (2021). arXiv:2109.10924
11. A.D. Martin, R.G. Roberts, W.J. Stirling, R.S. Thorne, Parton distributions incorporating QED contributions. Eur. Phys. J. C **39**, 155 (2005). arXiv:hep-ph/0411040
12. C. Schmidt, J. Pumplin, D. Stump, C.P. Yuan, CT14QED parton distribution functions from isolated photon production in deep inelastic scattering. Phys. Rev. D **93**(11), 114015 (2016). arXiv:1509.02905
13. R.D. Ball et al., Parton distributions with LHC data. Nucl. Phys. B **867**, 244 (2013). arXiv:1207.1303
14. NNPDF Collaboration, R.D. Ball et al., Parton distributions with QED corrections. Nucl. Phys. B **877**, 290–320 (2013). arXiv:1308.0598
15. xFitter Developers' Team Collaboration, F. Giuli et al., The photon PDF from high-mass Drell-Yan data at the LHC. Eur. Phys. J. C **77**(6), 400 (2017). arXiv:1701.08553
16. ATLAS Collaboration, G. Aad et al., Measurement of the double-differential high-mass Drell-Yan cross section in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. JHEP **08**, 009 (2016). arXiv:1606.01736
17. A. Manohar, P. Nason, G.P. Salam, G. Zanderighi, How bright is the proton? A precise determination of the photon parton distribution function. Phys. Rev. Lett. **117**(24), 242002 (2016). arXiv:1607.04266
18. A.V. Manohar, P. Nason, G.P. Salam, G. Zanderighi, The photon content of the proton. JHEP **12**, 046 (2017). arXiv:1708.01256
19. L.A. Harland-Lang, V.A. Khoze, M.G. Ryskin, Photon-initiated processes at high mass. Phys. Rev. D **94**(7), 074008 (2016). arXiv:1607.04635
20. L.A. Harland-Lang, V.A. Khoze, M.G. Ryskin, Sudakov effects in photon-initiated processes. Phys. Lett. B **761**, 20–24 (2016). arXiv:1605.04935
21. **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** (2018), no. 1 008, [arXiv:1712.07053]
22. L.A. Harland-Lang, A.D. Martin, R. Nathvani, R.S. Thorne, Ad Lucem: QED parton distribution functions in the MMHT framework. Eur. Phys. J. C **79**(10), 811 (2019). arXiv:1907.02750
23. T. Cridge, L.A. Harland-Lang, A.D. Martin, R.S. Thorne, QED parton distribution functions in the MSHT20 fit. Eur. Phys. J. C **82**(1), 90 (2022). arXiv:2111.05357
24. K. Xie, T. Hobbs, T.-J. Hou, C. Schmidt, M. Yan, C.-P. Yuan, The photon content of the proton in the CT18 global analysis. SciPost Phys. Proc. **8**, 074 (2022). arXiv:2107.13580
25. K. Xie, B. Zhou, T.J. Hobbs, The photon content of the neutron. arXiv:2305.10497
26. NNPDF Collaboration, R.D. Ball et al., The path to proton structure at 1% accuracy. Eur. Phys. J. C **82**(5), 428 (2022). arXiv:2109.02653
27. NNPDF Collaboration, R.D. Ball et al., An open-source machine learning framework for global analyses of parton distributions. Eur. Phys. J. C **81**(10), 958 (2021). arXiv:2109.02671

28. V. Bertone, S. Carrazza, J. Rojo, APFEL: a PDF evolution library with QED corrections. Comput. Phys. Commun. **185**, 1647 (2014). arXiv:1310.1394

29. V. Bertone, S. Carrazza, N.P. Hartland, APFELgrid: a high performance tool for parton density determinations. Comput. Phys. Commun. **212**, 205–209 (2017). arXiv:1605.02070

30. A. Candido, F. Hekhorn, G. Magni, EKO: evolution kernel operators. Eur. Phys. J. C **82**(10), 976 (2022). arXiv:2202.02338

31. A. Candido, F. Hekhorn, G. Magni, T. R. Rabemananjara, R. Stegeman, Yadism: yet another deep-inelastic scattering module. arXiv:2401.15187

32. C. Schwan, A. Candido, F. Hekhorn, S. Carrazza, A. Barontini, Nnpdf/pineappl: v0.6.0 (2023)

33. S. Carrazza, E.R. Nocera, C. Schwan, M. Zaro, PineAPPL: combining EW and QCD corrections for fast evaluation of LHC processes. JHEP **12**, 108 (2020). arXiv:2008.12789

34. A. Barontini, A. Candido, J.M. Cruz-Martinez, F. Hekhorn, C. Schwan, Pineline: industrialization of high-energy theory predictions. Comput. Phys. Commun. **297**, 109061 (2024). arXiv:2302.12124

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

36. A. Vogt, S. Moch, J.A.M. Vermaseren, The three-loop splitting functions in QCD: the singlet case. Nucl. Phys. B **691**, 129–181 (2004). arXiv:hep-ph/0404111

37. S. Moch, J.A.M. Vermaseren, A. Vogt, The Three loop splitting functions in QCD: the nonsinglet case. Nucl. Phys. B **688**, 101–134 (2004). arXiv:hep-ph/0403192

38. D. de Florian, G.F.R. Sborlini, G. Rodrigo, Two-loop QED corrections to the Altarelli–Parisi splitting functions. JHEP **10**, 056 (2016). arXiv:1606.02887

39. D. de Florian, G.F.R. Sborlini, G. Rodrigo, QED corrections to the Altarelli–Parisi splitting functions. Eur. Phys. J. C **76**(5), 282 (2016). arXiv:1512.00612

40. L.R. Surguladze, O(alpha**n alpha-s**m) corrections in e+ e- annihilation and tau decay. arXiv:hep-ph/9803211

41. S. Fanchiotti, B.A. Kniehl, A. Sirlin, Incorporation of QCD effects in basic corrections of the electroweak theory. Phys. Rev. D **48**, 307–331 (1993). arXiv:hep-ph/9212285

42. A. Denner, S. Dittmaier, Electroweak radiative corrections for collider physics. Phys. Rep. **864**, 1–163 (2020). arXiv:1912.06823

43. R.D. Ball, M. Bonvini, L. Rottoli, Charm in deep-inelastic scattering. JHEP **11**, 122 (2015). arXiv:1510.02491

44. CTEQ-TEA Collaboration, K. Xie, T.J. Hobbs, T.-J. Hou, C. Schmidt, M. Yan, C.P. Yuan, Photon PDF within the CT18 global analysis. Phys. Rev. D **105**(5), 054006 (2022). arXiv:2106.10299

45. The NNPDF Collaboration, R.D. Ball et al., A first unbiased global NLO determination of parton distributions and their uncertainties. Nucl. Phys. B **838**, 136 (2010). arXiv:1002.4407

46. The NNPDF Collaboration, R.D. Ball et al., Fitting parton distribution data with multiplicative normalization uncertainties. JHEP **05**, 075 (2010). arXiv:0912.2276

47. T. Cridge, L.A. Harland-Lang, R.S. Thorne, Combining QED and approximate N^{3}LO QCD corrections in a global PDF fit: MSHT20qed_an3lo PDFs. arXiv:2312.07665

48. NNPDF Collaboration, R. Abdul Khalek et al., A first determination of parton distributions with theoretical uncertainties. Eur. Phys. J. C **79**, 838 (2019). arXiv:1905.04311

49. NNPDF Collaboration, R. Abdul Khalek et al., Parton Distributions with Theory Uncertainties: General Formalism and First Phenomenological Studies. Eur. Phys. J. C **79**(11), 931 (2019). arXiv:1906.10698

50. F. Hekhorn, G. Magni, DGLAP evolution of parton distributions at approximate N^{3}LO. arXiv:2306.15294

51. R.D. Ball, A. Candido, S. Forte, F. Hekhorn, E.R. Nocera, J. Rojo, C. Schwan, Parton distributions and new physics searches: the Drell–Yan forward–backward asymmetry as a case study. Eur. Phys. J. C **82**(12), 1160 (2022). arXiv:2209.08115

52. E. Hammou, Z. Kassabov, M. Madigan, M.L. Mangano, L. Mantani, J. Moore, M.M. Alvarado, M. Ubiali, Hide and seek: how PDFs can conceal new physics. JHEP **11**, 090 (2023). arXiv:2307.10370

53. T. Carli et al., 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). arXiv:0911.2985

54. fastNLO Collaboration, M. Wobisch, D. Britzger, T. Kluge, K. Rabbertz, F. Stober, Theory-data comparisons for jet measurements in hadron-induced processes. arXiv:1109.1310

55. A. Candido, A. Garcia, G. Magni, T. Rabemananjara, J. Rojo, R. Stegeman, Neutrino structure functions from GeV to EeV energies. JHEP **05**, 149 (2023). arXiv:2302.08527

56. A. Barontini, A. Candido, F. Hekhorn, G. Magni, R. Stegeman, An FONLL prescription with coexisting flavor number PDFs **(in preparation)**

57. C. Anastasiou, L.J. Dixon, K. Melnikov, F. Petriello, High precision QCD at hadron colliders: electroweak gauge boson rapidity distributions at NNLO. Phys. Rev. D **69**, 094008 (2004). arXiv:hep-ph/0312266

58. NNPDF Collaboration, R.D. Ball et al., Parton distributions for the LHC Run II. JHEP **04**, 040 (2015). arXiv:1410.8849

59. A. Vogt, Efficient evolution of unpolarized and polarized parton distributions with qcd-pegasus. Comput. Phys. Commun. **170**, 65–92 (2005). arXiv:hep-ph/0408244

60. R.D. Ball, V. Bertone, M. Bonvini, S. Marzani, J. Rojo, L. Rottoli, Parton distributions with small-x resummation: evidence for BFKL dynamics in HERA data, Eur. Phys. J. C **78**(4), 321 (2018). arXiv:1710.05935