

The photon PDF from high-mass Drell Yan data at the LHC

F. Giuliani¹, and the xFitter Developers' team: V. Bertone^{2,3}, D. Britzger⁴, S. Carrazza⁵, A. Cooper-Sarkar¹, A. Glazov⁴, K. Lohwasser⁶, A. Luszczak⁷, F. Olness⁸, R. Plačákytė⁹, V. Radescu^{1,5}, J. Rojo^{2,3}, R. Sadykov¹⁰, P. Shvydkin¹⁰, O. Zenaiev⁴, and M. Lisovyi¹¹

¹University of Oxford, 1 Keble Road, Oxford OX1 3NP, United Kingdom

²Department of Physics and Astronomy, VU University, NL-1081 HV Amsterdam, The Netherlands

³Nikhef Theory Group Science Park 105, 1098 XG Amsterdam, The Netherlands

⁴DESY Hamburg, Notkestrasse 85 D-22609, Hamburg, Germany

⁵CERN, CH-1211 Geneva 23, Switzerland

⁶DESY Zeuthen, Platanenallee 6 D-15738, Zeuthen, Germany

⁷T.Kosciuszko Cracow University of Technology, 30-084 Cracow, Poland

⁸SMU Physics, Box 0175 Dallas, TX 75275-0175, United States of America

⁹Institut für Theoretische Physik, Universität Hamburg, Luruper Chaussee 149, D-22761 Hamburg, Germany

¹⁰Joint Institute for Nuclear Research (JINR), Joliot-Curie 6, 141980, Dubna, Moscow Region, Russia

¹¹Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany

Received: date / Accepted: date

Abstract Achieving the highest precision for theoretical predictions at the LHC requires the calculation of hard-scattering cross-sections that include perturbative QCD corrections up to (N)NNLO and electroweak (EW) corrections up to NLO. Parton distribution functions (PDFs) need to be provided with matching accuracy, which in the case of QED effects involves introducing the photon parton distribution of the proton, $x\gamma(x, Q^2)$. In this work a determination of the photon PDF from fits to recent ATLAS measurements of high-mass Drell-Yan dilepton production at $\sqrt{s} = 8$ TeV is presented. This analysis is based on the xFitter framework, and has required improvements both in the APFEL program, to account for NLO QED effects, and in the aMCfast interface to account for the photon-initiated contributions in the EW calculations within MadGraph5_aMC@NLO. The results are compared with other recent QED fits and determinations of the photon PDF, consistent results are found.

1 Introduction

Precision phenomenology at the LHC requires theoretical calculations which include not only QCD corrections, where NNLO is rapidly becoming the standard, but also electroweak (EW) corrections, which are particularly significant for observables directly sensitive to the TeV region, where EW Sudakov logarithms are enhanced. An important ingredient of these electroweak corrections is the photon parton distribution function (PDF) of the proton, $x\gamma(x, Q^2)$, which must be introduced to absorb the collinear divergences arising in initial-state QED emissions.

The first PDF fit to include both QED corrections and a photon PDF was MRST2004QED [1], where the photon PDF was taken from a model and tested on HERA data for direct photon production. Almost 10 years later, the NNPDF2.3QED analysis [2, 3] provided a first model-independent determination of the photon PDF based on Drell-Yan (DY) data from the LHC. The resulting photon PDF was however affected by large uncertainties due to the limited sensitivity of the data used as input to that fit. The determination of $x\gamma(x, Q^2)$ from NNPDF2.3QED was later combined with the state-of-the-art quark and gluon PDFs from NNPDF3.0, together with an improved QED evolution, to construct the NNPDF3.0QED set [4, 5]. The CT group has also released a QED fit using a similar strategy as the MRST2004QED one, named the CT14QED set [6].

A recent breakthrough concerning the determination of the photon content of the proton has been the realization that $x\gamma(x, Q^2)$ can be calculated in terms of inclusive lepton-proton deep-inelastic scattering (DIS) structure functions. The photon PDF resulting from this strategy is called LUXqed [7] and its residual uncertainties are now at the few percent level, not too different from those of the quark and gluon PDFs. A related approach by the HKR [8] group, denoted by HKR16 in the following, also leads to a similar photon PDF as compared to the LUXqed calculation, although in this case no estimate for the associated uncertainties is provided.

The aim of this work is to perform a direct determination of the photon PDF from recent high-mass Drell-Yan measurements from ATLAS at $\sqrt{s} = 8$ TeV [9], and to compare it with some of the existing determinations of $x\gamma(x, Q^2)$ mentioned above. Note that earlier measurements of high-mass

DY from ATLAS and CMS were presented in Refs. [10–12]. The ATLAS 8 TeV DY data are provided in terms of both single-differential cross-section distributions in the dilepton invariant mass, m_{ll} , and of double-differential cross-section distributions in m_{ll} and $|y_{ll}|$, the absolute value of rapidity of the lepton pair, and in m_{ll} and $\Delta\eta_{ll}$, the difference in pseudo-rapidity between the two leptons. Using the Bayesian reweighting method [13, 14] applied to NNPDF2.3QED, it was shown in the same publication [9] that these measurements provided significant information on $x\gamma(x, Q^2)$.

The goal of this study is therefore to investigate further these constraints from the ATLAS high-mass DY measurements on the photon PDF, this time by means of a direct PDF fit performed within the open-source xFitter framework [15]. State-of-the-art theoretical calculations are employed, in particular the inclusion of NNLO QCD and NLO QED corrections to the PDF evolution and the computation of the DIS structure functions as implemented in the APFEL program [16]. The implementation of NLO QED effects in APFEL is presented here for the first time. The inclusion of NLO QED evolution effects is cross checked using the independent QEDevol code [17] based on the QCDNUM evolution program [18].

The resulting determination of $x\gamma(x, Q^2)$ represents an important validation test of recent developments in theory and data concerning our understanding of the nature and implications of the photon PDF.

The outline of this paper is as follows. Sect. 2 reviews the ATLAS 8 TeV high-mass DY data together with the theoretical formalism of the DIS and Drell-Yan cross-sections used in the analysis. Sect. 3 presents the settings of the PDF fit within the xFitter framework. The fit results are then discussed in Sect. 4, where they are compared to determinations by other groups. Finally, Sect. 5 summarises and discusses the results and future lines of investigation. Appendix A contains a detailed description of the implementation and validation of NLO QED corrections to the DGLAP PDF evolution equations and DIS structure functions, which are available now in APFEL.

2 Data and theory

In this work, the photon content of the proton $x\gamma(x, Q^2)$ is extracted from a PDF analysis based on the combined inclusive DIS cross-section data from HERA [19] supplemented by the ATLAS measurements of high-mass Drell-Yan differential cross sections at $\sqrt{s} = 8$ TeV [9]. The HERA inclusive data are the backbone of modern PDF fits, providing information on the quark and gluon content of the proton, while the high-mass Drell-Yan data provide a direct sensitivity to the photon PDF. As illustrated in Fig. 1, dilepton production at hadron colliders can arise from either quark-

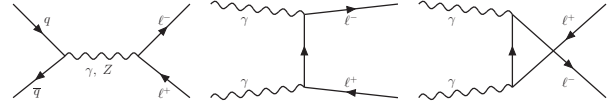


Figure 1 Diagrams that contribute to lepton-pair production at hadron colliders at the Born level.

antiquark s -channel scattering, or from photon-photon t - and u -channel scattering mediated by a lepton.

The ATLAS high-mass Drell-Yan 8 TeV measurements are presented in terms of both the single-differential (1D) invariant-mass distribution, $d\sigma/dm_{ll}$, as well as double-differential (2D) distributions in m_{ll} and y_{ll} , namely- $d^2\sigma/dm_{ll}d|y_{ll}|$, and in m_{ll} and $\Delta\eta_{ll}$, $d^2\sigma/dm_{ll}\Delta\eta_{ll}$. For the invariant-mass 1D distribution, there are 12 bins between $m_{ll} = 116$ GeV and 1.5 TeV; and for both double-differential distributions, there are five different bins in invariant mass, from the lowest bin with $116 \text{ GeV} < m_{ll} < 150 \text{ GeV}$ to the highest bin with $500 \text{ GeV} < m_{ll} < 1500 \text{ GeV}$. The first three (last two) m_{ll} bins of the 2D distributions are divided into 12 (6) bins with fixed width, extending up to 2.4 and 3.0 for the $|y_{ll}|$ and $|\Delta\eta_{ll}|$ distributions, respectively.

The photons which undergo hard scattering in the $\gamma\gamma \rightarrow ee$ process from Fig. 1 can be produced by either emission from the proton as a whole (the “elastic” component) or radiated by the constituent quarks (the “inelastic” component). From the theory point of view, the photon PDF extracted from the fit is by construction the sum of the elastic and inelastic contributions, though this analysis is mostly sensitive to the latter.

For the calculation of NLO high-mass Drell-Yan cross sections, the MadGraph5_aMC@NLO [20] program is used, which includes the contribution from photon-initiated diagrams, interfaced to APPLgrid [21] through aMCfast [22]. A tailored version of APPLgrid is used, accounting for the contribution of the photon-initiated processes¹. The calculation is performed in the $n_f = 5$ scheme neglecting mass effects of charm and bottom quarks in the matrix elements, as appropriate for a high-scale process. These NLO theoretical predictions match the analysis cuts of the data, with $m_{ll} \geq 116$ GeV, $\eta_l \leq 2.5$, and $p_T^l \geq 40$ GeV (30) GeV for the leading (sub-leading) lepton being the most important ones. As discussed below, the NLO calculations are then supplemented by NNLO/NLO K -factors obtained from FEWZ [23]. The NLO EW corrections to the DY processes are also estimated using FEWZ. The photon-initiated process is taken at LO since this corresponds to the APPLgrid implementation and the NLO corrections are very small compared to the data accuracy.

¹Modified version of APPLgrid available at: <https://github.com/scarrazza/applgridphoton>

The DIS structure functions and PDF evolution are computed with the APFEL program [16], which is currently accurate up to NNLO in QCD and NLO in QED, including the relevant mixed QCD+QED corrections. This means that, on top of the pure QCD contributions, the DGLAP evolution equations [24–26] are solved including the $\mathcal{O}(\alpha_s\alpha)$ and $\mathcal{O}(\alpha^2)$ corrections to the splitting functions. Corrections of $\mathcal{O}(\alpha)$ are also included leading to a (weak) explicit dependence of the predictions on the photon PDF. Details of the implementation of these corrections and of their numerical impact are given in Appendix A. Heavy-quark (charm and bottom) mass effects to DIS structure functions are taken into account using the FONLL-B (C) general-mass scheme [27] for the NLO (NNLO) fits. The numerical values of the heavy-quark masses in the mass parameter scheme are taken to be $m_c = 1.47$ GeV and $m_b = 4.5$ GeV as determined in [19], consistent with the latest PDG averages [28]. The reference values of the QCD and QED coupling constants are chosen to be $\alpha_s(m_Z) = 0.118$ and $\alpha(m_\tau = 1.777 \text{ GeV}) = 1/133.4$, again consistent with the PDG recommended values.

In the calculation of the Drell-Yan cross section, the dynamical renormalisation μ_R and factorisation μ_F scales are used, which are set equal to the scale of invariant mass m_{ll} , both for the quark- and gluon-induced and for the photon-induced contributions. The choice of other values for these scales in the QED diagrams, such as a fixed scale $\mu_R = \mu_F = M_Z$, leads to variations of the photon-initiated cross-sections of at most a few percent. The choice of the scale for the photon PDF is further discussed in [29, 30]. For the kinematics of the ATLAS DY data, the ratio between the photon-initiated contributions and quark- and gluon-induced dilepton production is largest for central rapidities and large invariant masses. For the most central (forward) rapidity bin, $0 < |y_{ll}| < 0.2$ ($2.0 < |y_{ll}| < 2.4$), the ratio between the QED and QCD contributions varies between 2.5% (2%) at low invariant masses and 12% (2.5%) for the highest m_{ll} bin. Therefore, data from the central region will exhibit the highest sensitivity to $x\gamma(x, Q^2)$.

The MadGraph5_aMC@NLO NLO QCD and LO QED calculations used in this work have been benchmarked against the corresponding predictions obtained with the FEWZ code [23], finding agreement within statistical uncertainties of the predictions for both the 1D and the 2D distributions.

In order to achieve NNLO QCD and NLO EW accuracy in our theoretical calculations, the NLO QCD and LO QED cross-sections computed with MadGraph5_aMC@NLO have been supplemented by bin-by-bin K -factors defined as:

$$K(m_{ll}, |y_{ll}|) \equiv \frac{\text{NNLO QCD} + \text{NLO EW}}{\text{NLO QCD} + \text{LO EW}}, \quad (1)$$

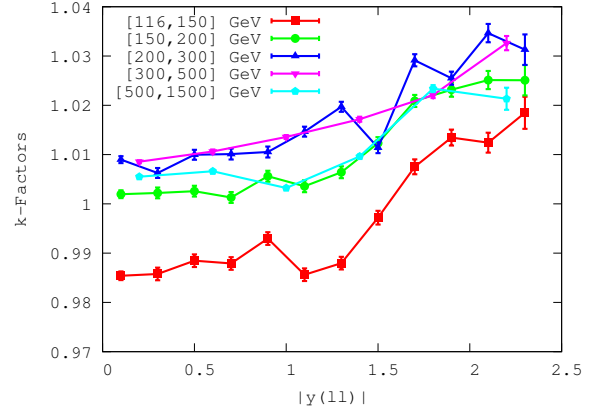


Figure 2 The NNLO/NLO K -factors, defined in Eq. (1), that account for higher order QCD and EW effects to the high-mass Drell-Yan cross sections with the photon induced contribution subtracted, as a function of the dilepton rapidity $|y_{ll}|$. Each set of points corresponds to a different bin in the dilepton invariant mass m_{ll} .

using the MMHT2014 NNLO [31] PDF set both in the numerator and in the denominator. The K -factors have been computed using FEWZ with the same settings and analysis cuts as the corresponding NLO calculations of MadGraph5_aMC@NLO. This approximation is justified since NNLO K -factors depend very mildly on the input PDF set, see for example [32]. The photon induced contribution, as provided in [9], has been explicitly subtracted from the FEWZ predictions. Fig. 2 shows the K -factors of Eq. (1) corresponding to the double differential $(m_{ll}, |y_{ll}|)$ cross sections as a function of the dilepton rapidity $|y_{ll}|$, where each set of points corresponds to a different dilepton invariant mass m_{ll} bin. The K -factors vary between 0.98 and 1.04, highlighting the fact that higher-order corrections to the Drell-Yan process are moderate, in particular at low values of m_{ll} and in the central region. Even at forward rapidities, the K -factors modify the NLO result by at most 4%.

3 Settings

This section presents the settings of the PDF fits, including the details about the parametrisation of the photon PDF $x\gamma(x, Q^2)$, which have been carried out using the open-source xFitter framework [15]. First of all, the scale Q_0^2 at which PDFs are parametrised is taken to be $Q_0^2 = 7.5 \text{ GeV}^2$, which coincides with the value Q_{\min}^2 that defines the kinematic cut $Q^2 \geq Q_{\min}^2$ for the data points that are used as input to the fits. The charm PDF is then generated perturbatively from quarks and gluons by means of DGLAP evolution, exploiting recent developments in APFEL which allow the setting of heavy-quark thresholds μ_h differently from the heavy quark masses m_h , such that $\mu_c = Q_0 > m_c$. Hence a high threshold can be used without having to parametrise the charm PDF [33].

The expression for the χ^2 function used for the fits is that of Ref. [34], which includes corrections for possible biases from statistical fluctuations and treats the systematic uncertainties multiplicatively. Alternative forms that do not include these corrections, such as those defined in [19, 35], have also been studied but no significant differences in the results have been observed.

In this analysis, the parametrised PDFs are the valence distributions $xu_v(x, Q_0^2)$ and $xd_v(x, Q_0^2)$, the gluon distribution $xg(x, Q_0^2)$, and the u -type and d -type sea-quark distributions, $x\bar{U}(x, Q_0^2)$, $x\bar{D}(x, Q_0^2)$, where $x\bar{U}(x, Q_0^2) = x\bar{u}(x, Q_0^2)$ and $x\bar{D}(x, Q_0^2) = x\bar{d}(x, Q_0^2) + x\bar{s}(x, Q_0^2)$. In addition, the photon distribution $x\gamma(x, Q_0^2)$ is also parametrised at the starting scale. The following general functional form is adopted:

$$xf(x, Q_0^2) = Ax^B(1-x)^C(1+Dx+Ex^2), \quad (2)$$

where some of the normalisation parameters, in particular A_{u_v} , A_{d_v} and A_g , are constrained by the valence and momentum sum rules (note that the photon PDF also enters the momentum sum rule). The parameters $B_{\bar{U}}$ and $B_{\bar{D}}$ are set equal to each other, so that the two quark sea distributions share a common small- x behaviour. Since the measurements used here are not sensitive to the strangeness content of the proton, strangeness is fixed such that $x\bar{s}(x, Q_0^2) = r_s x\bar{d}(x, Q_0^2)$, where $r_s = 1.0$ is consistent with the ATLAS analysis of inclusive W and Z production [36, 37]. The further constraint $A_{\bar{U}} = 0.5A_{\bar{D}}$ is imposed, such that $x\bar{u}(x, Q_0^2) \rightarrow x\bar{d}(x, Q_0^2)$ as $x \rightarrow 0$.

The explicit form of PDF parametrisation Eq. (2) at the scale Q_0^2 is determined by the technique of saturation of the χ^2 , namely the number of parameters is increased one by one until the χ^2 does not improve further, employing Wilks' theorem [38]. Following this method, the optimal parametrisation for the quark and gluon PDFs found for this analysis is:

$$\begin{aligned} xu_v(x) &= A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} (1 + E_{u_v} x^2), \\ xd_v(x) &= A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}}, \\ x\bar{U}(x) &= A_{\bar{U}} x^{B_{\bar{U}}} (1-x)^{C_{\bar{U}}}, \\ x\bar{D}(x) &= A_{\bar{D}} x^{B_{\bar{D}}} (1-x)^{C_{\bar{D}}}, \\ xg(x) &= A_g x^{B_g} (1-x)^{C_g} (1 + E_g x^2), \end{aligned} \quad (3)$$

while for the photon PDF it is used:

$$x\gamma(x) = A_\gamma x^{B_\gamma} (1-x)^{C_\gamma} (1 + D_\gamma x + E_\gamma x^2). \quad (4)$$

The parametrisation of the quark and gluon PDFs in Eq. (4) differs from the one used in the HERAPDF2.0 analysis in various ways. First of all, a higher value of the input evolution scale Q_0^2 is used, which is helpful to stabilise the fit of the photon PDF. Second, an additional negative term in the parametrisation of the gluon is not required here, because of the increased value of Q_0^2 which assures the positiveness of the gluon distribution. Third, the results of the

parametrisation scan are different because of the inclusion of the ATLAS high-mass Drell-Yan cross-section data.

PDF uncertainties are estimated using the Monte Carlo replica method [39–41], cross-checked with the Hessian method [42] using $\Delta\chi^2 = 1$. The former is expected to be more robust than the latter, due to the potential non-Gaussian nature of the photon PDF uncertainties [3]. In Section 4 it is shown that these two methods to estimate the PDF uncertainties on the photon PDF lead to similar results.

In addition, a number of cross-checks have been performed to assess the impact of various model and parametrisation uncertainties. For the model uncertainties, variations of the charm mass between $m_c = 1.41$ GeV to 1.53 GeV, of the bottom mass between $m_c = 4.25$ GeV to 4.75 GeV, of the strong coupling constant $\alpha_s(m_Z)$ between 0.116 to 0.120 are considered, and additionally the strangeness fraction is decreased down to $r_s = 0.75$. For the parametrisation uncertainties, the impact of increasing the input parametrisation scale up to $Q_0^2 = 10$ GeV² is considered as well as the impact of including additional parameters in Eq. (4). These extra parameters make little difference to the χ^2 of the fit, but they can change the shape of the PDFs in a non-negligible way. Such additional parameters are D_{u_v} , $D_{\bar{u}}$, $E_{\bar{d}}$, as well as the extra negative term in the gluon PDF used in HERAPDF2.0. The impact of these model and parametrisation uncertainties on the baseline results is quantified in Sect. 4.3.

4 Results

In this section the determination of the PDFs from a fit to HERA inclusive structure functions and ATLAS high-mass Drell-Yan cross sections, with an emphasis on the photon PDF is presented. First the fit quality is assessed and the fit results are compared with the experimental data. Then, the resulting photon PDF is shown and compared with other recent determinations. The impact of the high-mass DY data on the quark PDFs is also quantified. Following this, the robustness of the fits of $x\gamma(x, Q^2)$ with respect to varying the model, parametrisation, and procedural inputs is assessed. Finally, perturbative stability is addressed by comparing NLO and NNLO results.

4.1 Fit quality and comparison between data and fit results

In the following, the results that will be shown correspond to those obtained from fitting the double-differential (m_{ll}, y_{ll}) cross-section distributions. It has been verified that comparable results are obtained if the $(m_{ll}, \Delta\eta_{ll})$ cross-section distributions are fitted instead.

For the baseline NNLO fit, the value $\chi_{min}^2/N_{dof} = 1284./1083$ is obtained where N_{dof} is the number of degrees of freedom in the fit which is equal to total number of data

Dataset	χ^2/N_{dat}
HERA I+II	1236/1056
high-mass DY $116 \text{ GeV} \leq m_{ll} \leq 150 \text{ GeV}$	9/12
high-mass DY $150 \text{ GeV} \leq m_{ll} \leq 200 \text{ GeV}$	15/12
high-mass DY $200 \text{ GeV} \leq m_{ll} \leq 300 \text{ GeV}$	14/12
high-mass DY $300 \text{ GeV} \leq m_{ll} \leq 500 \text{ GeV}$	5/6
high-mass DY $500 \text{ GeV} \leq m_{ll} \leq 1500 \text{ GeV}$	4/6
Correlated (high-mass DY) χ^2	1.17
Log penalty (high-mass DY) χ^2	-0.12
Total (high-mass DY) χ^2/N_{dat}	48/48
Combined HERA I+II and high-mass DY χ^2/N_{dof}	1284/1083

Table 1 The χ^2/N_{dat} in the NNLO fits for the HERA inclusive structure functions and for the various invariant mass m_{ll} bins of the ATLAS high-mass DY data. In the latter case, the contribution to the χ^2 arising from the correlated and log-penalty terms are indicated, as well as the overall χ^2/N_{dof} is provided, where N_{dof} is the number of degree of freedom in the fit.

points minus number of free parameters. The contribution from the HERA inclusive data is $\chi^2/N_{\text{dat}} = 1236/1056$ and from the ATLAS high-mass DY data is $\chi^2/N_{\text{dat}} = 48/48$, where N_{dat} the number of the data points for the corresponding data sample. These values for χ^2/N_{dat} , together with the corresponding values for the various invariant mass m_{ll} bins of the ATLAS data, are summarised in Table 1. The quality of the agreement with the HERA cross sections is of comparable quality to that found in the HERAPDF2.0 analysis. Note that in the calculation of the total χ^2 for the ATLAS data, the correlations between the different m_{ll} bins have been taken into account.

Figs. 3–5 then show the comparison between the results of the NNLO fit, denoted by xFitter_epHMDY, and the ATLAS data for the $(m_{ll}, |y_{ll}|)$ double-differential Drell-Yan cross-sections as function of $|y_{ll}|$, for the five bins in m_{ll} separately. The comparisons are shown both on an absolute scale and as ratios to the central value of the experimental data. The error bars on the data points correspond to the bin-to-bin uncorrelated uncertainties, while the bands indicate the size of the correlated systematic uncertainties. The solid lines indicate the theory calculations obtained using the results of the fit.

Figs. 3–5 demonstrate a good agreement between ATLAS data and the NNLO theory predictions obtained from the xFitter_epHMDY fit. This agreement is also quantitatively expressed by the values of the χ^2 reported in Table 1, where for the ATLAS data a $\chi^2/N_{\text{dat}} = 1$ is found. This is particularly remarkable given the high precision of the data, with total experimental uncertainties at the few percent level in most of the kinematic range.

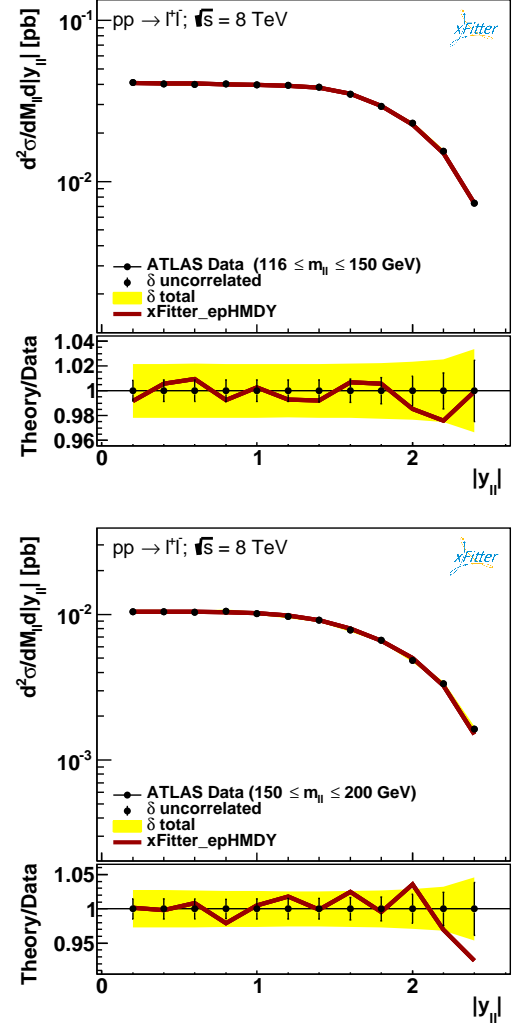


Figure 3 Comparison between the results of the fit and the ATLAS data for the $(m_{ll}, |y_{ll}|)$ double-differential Drell-Yan cross-sections as function of $|y_{ll}|$, for the first two m_{ll} bins. The comparisons are shown both in an absolute scale (upper plots) and as ratios to the central value of the experimental data in each y_{ll} bin (lower plots). The error bars on the data points correspond to the bin-to-bin uncorrelated uncertainties, while the yellow bands indicate the size of the correlated uncertainties. The solid lines indicate the theory calculations obtained using the results of the fit xFitter_epHMDY.

4.2 The photon PDF from LHC high-mass DY data

In Fig. 6, the photon PDF, $x\gamma(x, Q^2)$, is shown at $Q^2 = 10^4 \text{ GeV}^2$, and it is compared to the corresponding LUXqed, HKR16 and NNPDF3.0QED results. In the left plot the comparison is presented in an absolute scale, while in the right plot the ratio of different results normalized to the central value of the fit is shown. For the present fit, xFitter_epHMDY, the experimental PDF uncertainties at the 68% confidence level (CL) are obtained from the Monte Carlo method, while model and parametrisation uncertainties are discussed below. Likewise, the NNPDF3.0QED

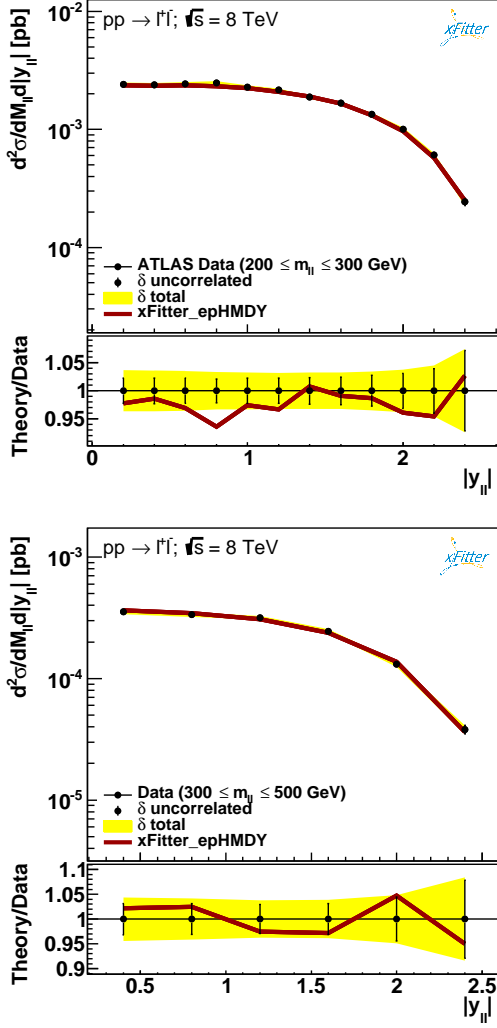


Figure 4 Same as Fig. 3 for the third and fourth m_{Tl} bins.

PDF set is shown the 68% CL uncertainty band, while for LUXqed the associated PDF uncertainty band is computed according to the prescription of Ref. [7]. For HKR16, only the central value is available. The x -range in Fig. 6 has been restricted to the region $0.02 \leq x \leq 0.9$, since beyond that region there is only limited sensitivity to $x\gamma(x, Q^2)$.

Fig. 6 shows that for $x \geq 0.1$ the four determinations of the photon PDF are consistent within PDF uncertainties. For smaller values of x , the photon PDF from LUXqed and HKR16 is somewhat smaller than xFitter_epHMDY, but still in agreement at the $2\text{-}\sigma$ level. This agreement is further improved if the PDF uncertainties in xFitter_epHMDY arising from variations of the input parametrisation are added to experimental uncertainties, as discussed in Sec. 4.3. Moreover, the results of this work and NNPDF3.0QED agree at the 68% CL for $x \geq 0.03$, and the agreement extends to smaller values of x once the parametrisation uncertainties in xFitter_epHMDY are accounted for. The LUXqed and

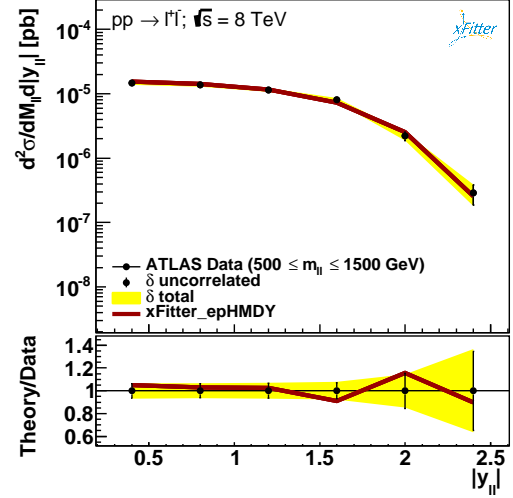


Figure 5 Same as Fig. 3 for the highest m_{Tl} bin.

the HKR16 calculations of $x\gamma(x, Q^2)$ are very close to each other across the entire range of x .

Fig. 6 shows that for $0.04 \leq x \leq 0.2$ the present analysis exhibits smaller PDF uncertainties as compared to those from NNPDF3.0QED. Indeed, the experimental uncertainty on the xFitter_epHMDY turns out to be at the $\sim 30\%$ level for $x \leq 0.1$. At larger x it increases rapidly specially in the positive direction. The reason for this behaviour at large x can be understood by recalling that variations of $x\gamma(x, Q^2)$ in the negative direction are constrained by positiveness. The limited sensitivity of the ATLAS data does not allow a determination of $x\gamma(x, Q^2)$ with uncertainties competitive with those of LUXqed, which are at the few percent level.

It is also interesting to assess the impact of the high-mass Drell-Yan 8 TeV measurements on the light quark and gluon PDFs. For this purpose, the fits have been repeated freezing the photon PDF to the xFitter_epHMDY shape. This is necessary because HERA inclusive data alone, which are the benchmark for this comparison, have no sensitivity to the photon PDF. This way, a meaningful comparison between the quark and gluon PDFs from a HERA-only baseline and the HERA+HMDY fit can be performed.

This comparison is shown in Fig. 7 for the up and down antiquarks $x\bar{u}(x, Q^2)$ and $x\bar{d}(x, Q^2)$, for which the effect of the high-mass Drell-Yan data is expected to be most pronounced, since HERA inclusive cross sections provide little information on quark flavour separation. In Fig. 7, the $x\bar{u}(x, Q^2)$ and $x\bar{d}(x, Q^2)$ together with the associated MC uncertainties have been computed at the initial parametrisation scale of $Q^2 = 7.5 \text{ GeV}^2$ and are shown as ratios to the central value of the xFitter_epHMDY fit. The modifications in the medium and large- x antiquark distributions from the high mass DY data are rather moderate. It has been verified that the same conclusions can be derived from fits obtained by

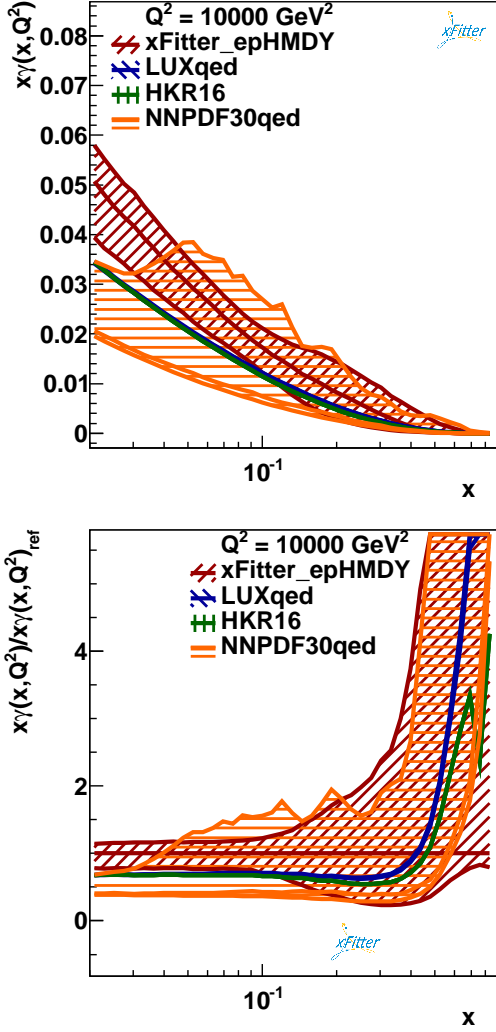


Figure 6 Left plot: comparison between the photon $x\gamma(x, Q^2)$ at $Q^2 = 10^4 \text{ GeV}^2$ from the present NNLO analysis (xFitter_epHMDY) with the corresponding results from NNPDF3.0QED, LUXqed and HKR16. Right plot: the same comparison, now with the results normalized to the central value of xFitter_epHMDY. For the present fit, the PDF uncertainties are shown at the 68% CL obtained from the MC method, while model and parametrisation uncertainties are discussed below. For HKR16 only the central value is shown, while for LUXqed the associated PDF uncertainty band [7] is included.

switching off the QED effects for both the HERA only fits and the HERA+HMDY fits. Therefore, while the ATLAS high-mass Drell-Yan measurements have a significant constraint on the photon PDF, their impact on the quark and gluon PDFs is moderate.

4.3 Robustness and perturbative stability checks

Following the presentation of the main result of this work, the xFitter_epHMDY determination of the photon PDF $x\gamma(x, Q^2)$, the robustness of this determination with respect to a number of variations is assessed. Firstly, variations in

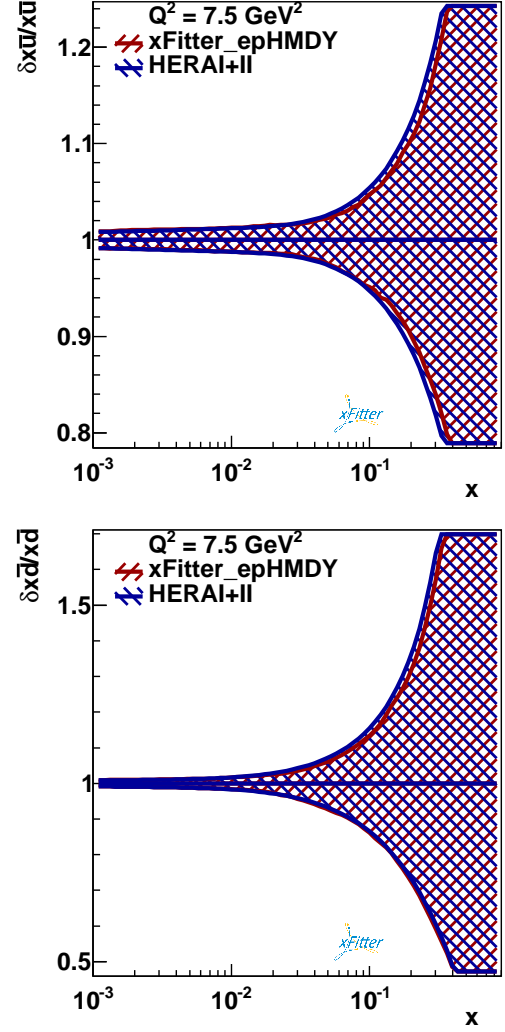


Figure 7 The impact of the ATLAS high-mass 8 TeV Drell-Yan measurements on the $x\bar{u}$ and $x\bar{d}$ sea quark PDFs at the input parametrisation scale $Q^2 = 7.5 \text{ GeV}^2$. The results are shown normalized to the central value of xFitter_epHMDY.

the values of the input physical parameters, such as α_s or the charm mass are explored. Secondly, variations of the choices made for the PDF input parametrisation are considered. Finally, variations associated to different methodological choices in the fitting procedure are quantified. In each case, one variation at a time is performed and compared with the central value of $x\gamma(x, Q^2)$ and its experimental PDF uncertainties computed using the Monte Carlo method.

First the impact of uncertainties associated to either the choice of input physical parameters or of specific settings adopted in the fit is considered. Fig. 8 shows the comparison between the xFitter_epHMDY determination of $x\gamma(x, Q^2)$ at $Q^2 = 10^4 \text{ GeV}^2$, including the experimental MC uncertainties, with the central value of those fits for which a number of variations have been performed. Specifically:

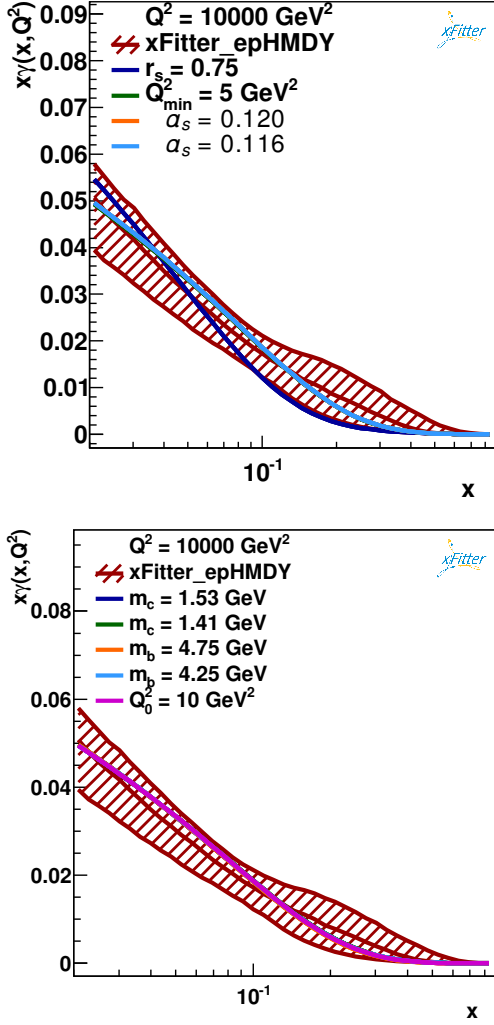


Figure 8 Comparison between the baseline determination of $x\gamma(x, Q^2)$ at $Q^2 = 10^4 \text{ GeV}^2$ in the present analysis, xFitter_epHMDY, with the central value of a number of fits for which one input parameter has been varied. The following variations have been considered: $r_s = 0.75$, $Q_{\min}^2 = 5 \text{ GeV}^2$, $\alpha_s = 0.116$ and 0.118 (left plot); and $m_c = 1.41$ and 1.53 GeV , $m_b = 4.25$ and 4.75 GeV , and $Q_0^2 = 10 \text{ GeV}^2$ (right plot). See text for more details about these variations.

- The strong coupling constant is varied by $\delta\alpha_s = \pm 0.002$ around the central value.
- The ratio of strange to non-strange light quark PDFs is decreased to $r_s = 0.75$ instead of $r_s = 1$.
- The value of the charm mass is varied between $m_c = 1.41 \text{ GeV}$ and $m_c = 1.53 \text{ GeV}$, and that of the bottom mass between $m_b = 4.25 \text{ GeV}$ and $m_b = 4.75 \text{ GeV}$.
- The minimum value Q_{\min}^2 of the fitted data is decreased down to 5 GeV^2 .
- The input parametrisation scale Q_0^2 is raised to 10 GeV^2 as compared to the baseline value of $Q_0^2 = 7.5 \text{ GeV}^2$.

The results of Fig. 8 highlight that in all cases effect of the variations considered here is contained within (and typically much smaller than) the experimental PDF uncertainty bands

of the reference fit. The largest variation comes from the strangeness ratio r_s , where the resulting central value turns out to be at the bottom end of the PDF uncertainty band for $x \gtrsim 0.1$.

Another important check of the robustness of the present determination of $x\gamma(x, Q^2)$ can be obtained by comparing the baseline fit with further fits where a number of new free parameters are allowed in the PDF parametrisation, in addition to those listed in Eq. (4). Fig. 9 shows the impact of three representative variations (others have been explored, leading to smaller differences): more flexibility to the gluon distribution, allowing it to become negative at the initial scale (labeled by “neg”), in addition to D_{u_v} , and then $D_{\bar{u}} + D_{\bar{d}}$. As before, all variations are contained within the experimental PDF uncertainty bands, though the impact of the parametrisation variations is typically larger than that of the model variations: in the case of the neg + $D_{\bar{u}} + D_{\bar{d}}$ variations, the central value is at the lower edge of the PDF uncertainty band in the entire range of x shown.

A cross-check of the robustness of the estimated experimental uncertainty of the photon PDF in this analysis is provided by the comparison of the Monte Carlo and Hessian methods. Fig. 9 shows this comparison indicating a reasonable agreement between the two methods. In particular, the central values of the photon obtained with the two fitting techniques are quite similar to each other. As expected, the MC uncertainties tend to be larger than the Hessian ones, specially in the region $x \gtrsim 0.2$, indicating deviations with respect to the Gaussian behaviour of the photon PDF.

To complete these studies, an interesting exercise is to quantify the perturbative stability of the xFitter_epHMDY determination of the photon PDF $x\gamma(x, Q^2)$ with respect to the inclusion of NNLO QCD corrections in the analysis. To study this, Fig. 10 shows a comparison between the baseline fit of $x\gamma(x, Q^2)$, based on NNLO QCD and NLO QED theoretical calculations, with the central value resulting from a corresponding fit based instead on NLO QCD and QED theory. In other words, the QED part of the calculations is identical in both cases. For the NNLO fit, only the experimental PDF uncertainties, estimated using the Monte Carlo method, are shown. From the comparison of Fig. 10, it is clear that the fit of $x\gamma(x, Q^2)$ exhibits a reasonable perturbative stability, since the central value of the NLO fit is always contained in the one-sigma PDF uncertainty band of the baseline xFitter_epHMDY fit. The agreement between the two fits is particularly good for $x \gtrsim 0.1$, where the two central values are very close to each other. This comparison is shown at low scale, $Q^2 = 7.5 \text{ GeV}^2$ and high scales $Q^2 = 10^4 \text{ GeV}^2$, indicating that perturbative stability is not scale dependent.

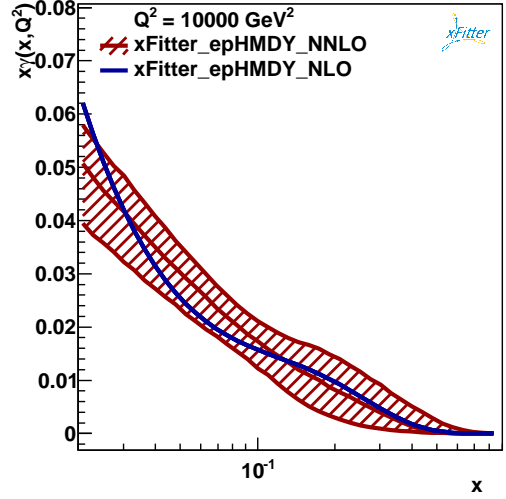
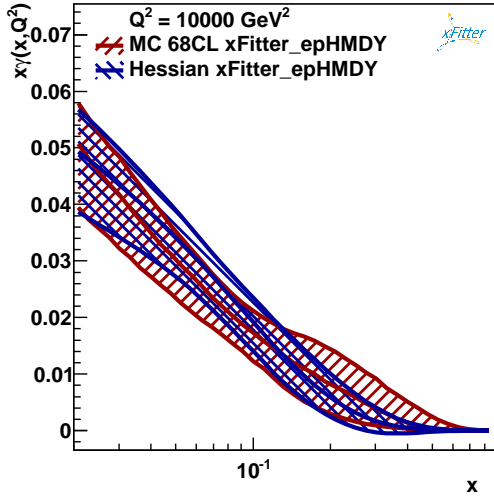
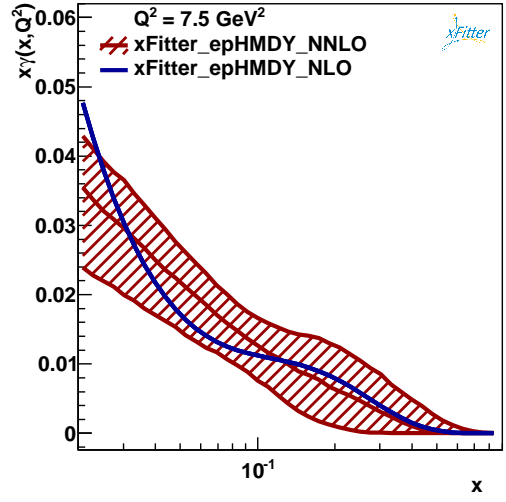
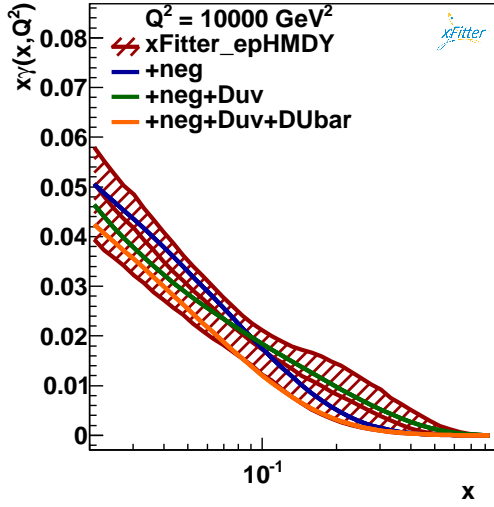


Figure 9 Left: the impact on the photon PDF $x\gamma(x, Q^2)$ from `xFitter_epHMDY` in fits where a number of additional free parameters are allowed in the PDF parametrisation Eq. (4). The parametrisation variations that have been explored are: more flexibility to the gluon distribution, allowing it to become negative (labeled by “neg”), adding on top D_{u_v} , and then adding $D_{\bar{u}} + D_{\bar{d}}$. Right: comparison between the `xFitter_epHMDY` determinations obtained with the Monte Carlo (baseline) and with the Hessian methods, where in both cases the PDF error band shown corresponds to the 68% CL uncertainties.

Figure 10 Left plot: comparison between the reference `xFitter_epHMDY` fit of $x\gamma(x, Q^2)$, based on NNLO QCD and NLO QED theoretical calculations, with the central value of the corresponding fit based on NLO QCD and QED theory, at $Q^2 = 7.5 \text{ GeV}^2$. In the former case, only the experimental Monte Carlo PDF uncertainties are shown. Right plot: same comparison, now presented at the higher scale of $Q^2 = 10^4 \text{ GeV}^2$.

5 Summary

In this work, a new determination of the photon PDF from a fit of HERA inclusive DIS structure functions supplemented by ATLAS data on high-mass Drell-Yan cross sections has been presented, based on the `xFitter` framework. As suggested by a previous reweighting analysis [9], this high-mass DY data provides significant constraints on the photon PDF, allowing a determination of $x\gamma(x, Q^2)$ with uncertainties at the 30% level for $0.02 \leq x \leq 0.1$. The results of the present study, dubbed `xFitter_epHMDY`, are in agreement and exhibit smaller PDF uncertainties than the only other ex-

isting photon PDF fit from LHC data, the NNPDF3.0QED analysis, based on previous LHC Drell-Yan measurements.

The results are in agreement within uncertainties with two recent calculations of the photon PDF, LUXqed and HKR16. For $x \geq 0.1$, the agreement is at the $1\text{-}\sigma$ level already including only the experimental MC uncertainties, while for $0.02 \leq x \leq 0.1$ it is important to account for parametrisation uncertainties. The findings indicate that a direct determination of the photon PDF from hadron collider data is still far from being competitive with the LUXqed and HKR calculations, which are based instead on precise measurements of the inclusive DIS structure functions of the proton.

The results of this study, which are available upon request in the LHAPDF6 format [43], have been made possible by a number of technical developments that should be of direct application for future PDF fits accounting for QED corrections. First of all, the full NLO QED corrections to the DGLAP evolution equations and the DIS structure functions have been implemented in the APFEL program. Moreover, our results illustrate the flexibility of the xFitter framework to extend its capabilities beyond the traditional quark and gluon PDF fits. Finally, the extension of aMCfast and APPLgrid to allow for the presence of photon-initiated channels in the calculations provided by MadGraph5_aMC@NLO, significantly streamlines the inclusion of future LHC measurements in PDF fits with QED corrections by consistently including diagrams with initial-state photons. All these technical improvements will certainly be helpful for future studies of the photon content of the proton.

Acknowledgements We thank L. Harland-Lang for providing us the LHAPDF6 grid of the HKR16 photon determination. We thank A. Sapronov for providing support to the xFitter platform. We thank P. Starovoitov for discussions related to the APPLgrid files. We thank M. Zinser for discussions related to the ATLAS data. We thank M. Dyndal for discussions on the scale choices for the photon induced contributions. The work of V. B., F. G., J. R. has been supported by the European Research Council Starting Grant “PDF4BSM”. The work of S. C. is supported by the HICCUP ERC Consolidator grant (614577). A. L. thanks for the support from the Mobility Plus grant no. 1320/MOB/IV/2015/0. The work of F. O. has been supported by the US DoE Grant DE-SC0010129. The work of P. S and R. S. has been supported by the BMBF-JINR cooperation. We are grateful to the DESY IT department for their support of the xFitter developers.

Appendix A: Implementation of NLO QED corrections in APFEL

In this appendix, the details of the implementation of the combined NLO QCD+QED corrections in the APFEL program are presented. As discussed in Ref. [16], the implementation of the LO QED corrections to the DGLAP evolution equations presents many simplifications, in particular the fact that QED and QCD corrections do not mix and therefore the DGLAP equations, as well as the evolution equations for the running of α_s and α , are decoupled. When considering NLO corrections, this property does not hold anymore, and QED and QCD contributions mix both in the DGLAP and in the coupling evolution equations. On top of this, NLO QED corrections induce the presence of diagrams with a real photon in the initial state, and these have to be consistently included in the computation of the DIS structure functions resulting in an explicit dependence on the photon PDF of these observables.

In the following, the discussion starts by discussing the generalisation of the equations for the running of the QCD and QED couplings, finding that the mixed QCD+QED terms have a negligible impact. Then the extension of the DGLAP evolution equations to account for the complete NLO QCD+QED effects is discussed. Finally, the modifications introduced by the NLO QED corrections in both the neutral-current and the charged-current DIS structure functions are discussed.

Appendix A.1: Evolution of the couplings

As mentioned above, the NLO QCD+QED corrections induce the presence of mixed terms in the evolution equations of α_s and α . In practice, the QCD β -function receives corrections proportional to α and, vice-versa, the QED β -function receives corrections proportional to α_s , in such a way that the coupling evolution equations read:

$$\begin{aligned}\mu^2 \frac{\partial \alpha_s}{\partial \mu^2} &= \beta^{\text{QCD}}(\alpha_s, \alpha), \\ \mu^2 \frac{\partial \alpha}{\partial \mu^2} &= \beta^{\text{QED}}(\alpha_s, \alpha).\end{aligned}\tag{A.1}$$

As a consequence, these evolution equations form a set of coupled differential equations. Up to three loops (*i.e.* NLO), the β -functions can be expanded as:

$$\begin{aligned}\beta^{\text{QCD}}(\alpha_s, \alpha) &= -\alpha_s \left[\beta_0^{(\alpha_s)} \left(\frac{\alpha_s}{4\pi} \right) + \beta_1^{(\alpha_s, \alpha)} \left(\frac{\alpha_s}{4\pi} \right) \left(\frac{\alpha}{4\pi} \right) \right. \\ &\quad \left. + \beta_1^{(\alpha_s^2)} \left(\frac{\alpha_s}{4\pi} \right)^2 + \dots \right],\end{aligned}\tag{A.2}$$

and:

$$\begin{aligned}\beta^{\text{QED}}(\alpha_s, \alpha) &= -\alpha \left[\beta_0^{(\alpha)} \left(\frac{\alpha}{4\pi} \right) + \beta_1^{(\alpha, \alpha_s)} \left(\frac{\alpha}{4\pi} \right) \left(\frac{\alpha_s}{4\pi} \right) \right. \\ &\quad \left. + \beta_1^{(\alpha^2)} \left(\frac{\alpha}{4\pi} \right)^2 + \dots \right].\end{aligned}\tag{A.3}$$

where the mixed terms, $\beta_1^{(\alpha_s, \alpha)}$ and $\beta_1^{(\alpha, \alpha_s)}$, and the pure NLO QED term, $\beta_1^{(\alpha^2)}$, can be found in Ref. [44]. Taking into account a factor four due to the different definitions of the expansion parameters, one finds:

$$\begin{aligned}\beta_1^{(\alpha_s, \alpha)} &= -2 \sum_{i=1}^{n_f} e_q^2, \\ \beta_1^{(\alpha, \alpha_s)} &= -\frac{16}{3} N_c \sum_{i=1}^{n_f} e_q^2, \\ \beta_1^{(\alpha^2)} &= -4 \left(n_l + N_c \sum_{i=1}^{n_f} e_q^2 \right),\end{aligned}\tag{A.4}$$

where $N_c = 3$ is the number of colours, e_q is the electric charge of the quark flavour q , and n_f and n_l are the number of active quark and lepton flavours, respectively.

Eq. (A.1) can be written in the vectorial form:

$$\mu^2 \frac{\partial \alpha}{\partial \mu^2} = \beta(\alpha(\mu)), \quad (\text{A.5})$$

with:

$$\alpha = \begin{pmatrix} \alpha_s \\ \alpha \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} \beta^{\text{QCD}} \\ \beta^{\text{QED}} \end{pmatrix}. \quad (\text{A.6})$$

Eq. (A.5) is an ordinary differential equation that can be numerically solved using, for example, Runge-Kutta methods.

The first two terms in eq. (A.4) are responsible for the coupling of the evolution of α_s and α , and thus they introduce a complication that affects both the implementation and the performance of the code. One can then ask what is the effect of their presence and whether their removal makes a substantial difference.

Fig. 11 shows the comparison between the evolution at NLO of both couplings α_s and α including and excluding the mixed terms in the respective β -functions. The evolution is performed between the Z mass scale M_Z and 10 TeV with 5 active quark flavours and 3 active lepton flavours and uses as boundary conditions $\alpha_s(M_Z) = 0.118$ and $\alpha(M_Z) = 1/128$. The two curves in Fig. 11 are normalised to the respective curves without mixed terms. It is clear that the mixed terms lead to tiny relative differences that are at most of $\mathcal{O}(10^{-4})$ at 10 TeV for α_s and $\mathcal{O}(10^{-3})$ at the same scale for α . Thus it is safe to conclude that the mixed terms in the β -functions have a negligible effect on the evolution of the couplings and thus they are excluded to simplify the code and to improve the performance without introducing any significant inaccuracy.

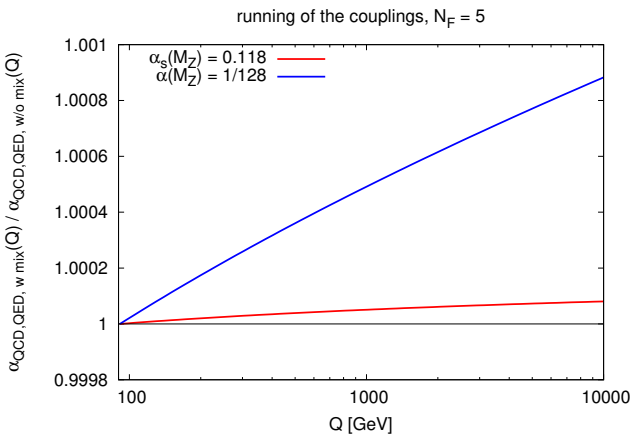


Figure 11 Comparison between the running with the scale Q of the QCD and QED couplings, α_s and α , including or not the mixed terms in the corresponding β -functions. The curves are normalised to the result of the respective coupling running without the mixed terms included in the β functions.

Appendix A.2: PDF evolution with NLO QED corrections

Next the implementation of the full NLO QCD+QED corrections to the DGLAP evolution equations is considered. The discussion is limited to consider of the photon PDF only, while the possible presence of the lepton PDFs is not addressed here. The first step towards an efficient implementation of the solution of the DGLAP equations in the presence of QED corrections is the adoption of a suitable PDF basis that diagonalises the splitting function matrix, decoupling as many equations as possible. Such a basis was already introduced in the Appendix A of Ref. [45] and will be used also here. Excluding the lepton PDFs, this basis contains 14 independent PDF combinations and reads:

$$\begin{aligned} 1 : g & & 9 : V = V_u + V_d \\ 2 : \gamma & & 10 : \Delta_V = V_u - V_d \\ 3 : \Sigma = \Sigma_u + \Sigma_d & & 11 : V_1^u = u^- - c^- \\ 4 : \Delta_\Sigma = \Sigma_u - \Sigma_d & & 12 : V_2^u = u^- + c^- - 2t^- \\ 5 : T_1^u = u^+ - c^+ & & 13 : V_1^d = d^- - s^- \\ 6 : T_2^u = u^+ + c^+ - 2t^+ & & 14 : V_2^d = d^- + s^- - 2b^- \\ 7 : T_1^d = d^+ - s^+ & & \\ 8 : T_2^d = d^+ + s^+ - 2b^+ & & \end{aligned} \quad (\text{A.7})$$

where $q^\pm = q \pm \bar{q}$ with $q = u, d, s, c, b, t$. In addition:

$$\begin{aligned} \Sigma_u &= u^+ + c^+ + t^+, & V_u &= u^- + c^- + t^-, \\ \Sigma_d &= d^+ + s^+ + b^+, & V_d &= d^- + s^- + b^-. \end{aligned} \quad (\text{A.8})$$

The second step is the construction of the splitting function matrix that determines the evolution of each of the combinations listed in Eq. (A.7). To this end, the splitting function matrix P is split into a pure QCD term \tilde{P} , which only depends on α_s , and a mixed QCD+QED correction term \bar{P} , which instead contains contributions proportional to at least one power of the QED coupling α . In practice, this means that:

$$P = \tilde{P} + \bar{P}, \quad (\text{A.9})$$

where the pure QCD term reads:

$$\tilde{P} = \alpha_s \mathcal{P}^{(1,0)} + \alpha_s^2 \mathcal{P}^{(2,0)} + \dots, \quad (\text{A.10})$$

while the term containing the QED coupling is given by:

$$\bar{P} = \alpha \mathcal{P}^{(0,1)} + \alpha_s \alpha \mathcal{P}^{(1,1)} + \alpha^2 \mathcal{P}^{(0,2)} + \dots \quad (\text{A.11})$$

Note that in the r.h.s. of Eqs. (A.10) and (A.11) the convention of Refs. [46, 47] is followed to indicate the power of α_s and α that each splitting function multiplies.

The structure of the pure QCD splitting function matrix \tilde{P} as well as the first term in \bar{P} , which represents the pure LO QED correction, were already discussed in Ref. [45]. It is now necessary to analyse the structure of the two additional terms, namely $\mathcal{P}^{(1,1)}$ and $\mathcal{P}^{(0,2)}$. Starting with the $\mathcal{O}(\alpha_s \alpha)$ term, the resulting evolution equations read:

$$\mu^2 \frac{\partial}{\partial \mu^2} \begin{pmatrix} g \\ \gamma \\ \Sigma \\ \Delta_\Sigma \end{pmatrix} \bigg|_{\mathcal{O}(\alpha_s \alpha)} = \begin{pmatrix} e_\Sigma^2 \mathcal{P}_{gg}^{(1,1)} & e_\Sigma^2 \mathcal{P}_{g\gamma}^{(1,1)} & \eta^+ \mathcal{P}_{gq}^{(1,1)} & \eta^- \mathcal{P}_{gq}^{(1,1)} \\ e_\Sigma^2 \mathcal{P}_{\gamma g}^{(1,1)} & e_\Sigma^2 \mathcal{P}_{\gamma\gamma}^{(1,1)} & \eta^+ \mathcal{P}_{\gamma q}^{(1,1)} & \eta^- \mathcal{P}_{\gamma q}^{(1,1)} \\ 2e_\Sigma^2 \mathcal{P}_{qg}^{(1,1)} & 2e_\Sigma^2 \mathcal{P}_{q\gamma}^{(1,1)} & \eta^+ \mathcal{P}^{+(1,1)} & \eta^- \mathcal{P}^{+(1,1)} \\ 2\delta_e^2 \mathcal{P}_{qg}^{(1,1)} & 2\delta_e^2 \mathcal{P}_{q\gamma}^{(1,1)} & \eta^- \mathcal{P}^{+(1,1)} & \eta^+ \mathcal{P}^{+(1,1)} \end{pmatrix} \otimes \begin{pmatrix} g \\ \gamma \\ \Sigma \\ \Delta_\Sigma \end{pmatrix}, \quad (\text{A.12})$$

$$\mu^2 \frac{\partial}{\partial \mu^2} \begin{pmatrix} V \\ \Delta_V \end{pmatrix} \bigg|_{\mathcal{O}(\alpha_s \alpha)} = \begin{pmatrix} \eta^+ \mathcal{P}^{-(1,1)} & \eta^- \mathcal{P}^{-(1,1)} \\ \eta^- \mathcal{P}^{-(1,1)} & \eta^+ \mathcal{P}^{-(1,1)} \end{pmatrix} \otimes \begin{pmatrix} V \\ \Delta_V \end{pmatrix}, \quad (\text{A.13})$$

$$\begin{aligned} \mu^2 \frac{\partial T_{1,2}^u}{\partial \mu^2} \bigg|_{\mathcal{O}(\alpha_s \alpha)} &= e_u^2 \mathcal{P}^{+(1,1)} \otimes T_{1,2}^u, \quad \mu^2 \frac{\partial T_{1,2}^d}{\partial \mu^2} \bigg|_{\mathcal{O}(\alpha_s \alpha)} = e_d^2 \mathcal{P}^{+(1,1)} \otimes T_{1,2}^d, \\ \mu^2 \frac{\partial V_{1,2}^u}{\partial \mu^2} \bigg|_{\mathcal{O}(\alpha_s \alpha)} &= e_u^2 \mathcal{P}^{-(1,1)} \otimes V_{1,2}^u, \quad \mu^2 \frac{\partial V_{1,2}^d}{\partial \mu^2} \bigg|_{\mathcal{O}(\alpha_s \alpha)} = e_d^2 \mathcal{P}^{-(1,1)} \otimes V_{1,2}^d. \end{aligned} \quad (\text{A.14})$$

where \otimes indicates the Mellin convolution and where:

$$\begin{aligned} e_\Sigma^2 &\equiv N_c(n_u e_u^2 + n_d e_d^2), \\ \delta_e^2 &\equiv N_c(n_u e_u^2 - n_d e_d^2), \\ \eta^\pm &\equiv \frac{1}{2}(e_u^2 \pm e_d^2), \end{aligned} \quad (\text{A.15})$$

with e_u and e_d the electric charges of the up- and down-type quarks, and n_u and n_d the number of up- and down-type active quark flavours (such that $n_u + n_d = n_f$).

Next the $\mathcal{O}(\alpha^2)$ corrections as considered. The expressions of the splitting functions at this order have been presented in Ref. [47]. There are two relevant new features that distinguish these corrections from the $\mathcal{O}(\alpha)$ and the

$\mathcal{O}(\alpha_s \alpha)$ ones. The first one is that, contrary to the other cases in which the electric charges appears to the second power at most, here they appear up to the fourth power. As a consequence, new couplings must be introduced:

$$\begin{aligned} e_\Sigma^4 &= N_c(n_u e_u^4 + n_d e_d^4), \\ \delta_e^4 &= N_c(n_u e_u^4 - n_d e_d^4). \end{aligned} \quad (\text{A.16})$$

The second feature is that the dependence on the electric charges of some of the $\mathcal{O}(\alpha^2)$ splitting functions is not factorisable as was the case for all the $\mathcal{O}(\alpha)$ and $\mathcal{O}(\alpha_s \alpha)$ ones and therefore a distinction must be made between up- and down-type splitting functions. Taking into account these features, it is possible to show that the $\mathcal{O}(\alpha^2)$ contributions to the DGLAP equations take the following form:

$$\mu^2 \frac{\partial}{\partial \mu^2} \begin{pmatrix} g \\ \gamma \\ \Sigma \\ \Delta_\Sigma \end{pmatrix} \bigg|_{\mathcal{O}(\alpha^2)} = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2e_\Sigma^4 \mathcal{P}_{\gamma\gamma}^{(0,2)} & e_u^4 \mathcal{P}_{\gamma u}^{(0,2)} + e_d^4 \mathcal{P}_{\gamma d} & e_u^4 \mathcal{P}_{\gamma u}^{(0,2)} - e_d^4 \mathcal{P}_{\gamma d}^{(0,2)} \\ 0 & 4e_\Sigma^4 \mathcal{P}_{q\gamma}^{(0,2)} & e_u^4 \mathcal{P}_{uu}^{+(0,2)} + e_d^4 \mathcal{P}_{dd}^{+(0,2)} + 2\eta^+ e_\Sigma^2 \mathcal{P}_{qq}^{S(0,2)} & e_u^4 \mathcal{P}_{uu}^{+(0,2)} - e_d^4 \mathcal{P}_{dd}^{+(0,2)} + 2\eta^- e_\Sigma^2 \mathcal{P}_{qq}^{S(0,2)} \\ 0 & 4\delta_e^4 \mathcal{P}_{q\gamma}^{(0,2)} & e_u^4 \mathcal{P}_{uu}^{+(0,2)} - e_d^4 \mathcal{P}_{dd}^{+(0,2)} + 2\eta^- \delta_e^2 \mathcal{P}_{qq}^{S(0,2)} & e_u^4 \mathcal{P}_{uu}^{+(0,2)} + e_d^4 \mathcal{P}_{dd}^{+(0,2)} + 2\eta^+ \delta_e^2 \mathcal{P}_{qq}^{S(0,2)} \end{pmatrix} \otimes \begin{pmatrix} g \\ \gamma \\ \Sigma \\ \Delta_\Sigma \end{pmatrix}, \quad (\text{A.17})$$

$$\mu^2 \frac{\partial}{\partial \mu^2} \begin{pmatrix} V \\ \Delta_V \end{pmatrix} \bigg|_{\mathcal{O}(\alpha^2)} = \frac{1}{2} \begin{pmatrix} e_u^4 \mathcal{P}_{uu}^{-(0,2)} + e_d^4 \mathcal{P}_{dd}^{-(0,2)} & e_u^4 \mathcal{P}_{uu}^{-(0,2)} - e_d^4 \mathcal{P}_{dd}^{-(0,2)} \\ e_u^4 \mathcal{P}_{uu}^{-(0,2)} - e_d^4 \mathcal{P}_{dd}^{-(0,2)} & e_u^4 \mathcal{P}_{uu}^{-(0,2)} + e_d^4 \mathcal{P}_{dd}^{-(0,2)} \end{pmatrix} \otimes \begin{pmatrix} V \\ \Delta_V \end{pmatrix}, \quad (\text{A.18})$$

$$\begin{aligned} \mu^2 \frac{\partial T_{1,2}^u}{\partial \mu^2} \bigg|_{\mathcal{O}(\alpha^2)} &= e_u^4 \mathcal{P}_{uu}^{+(0,2)} \otimes T_{1,2}^u, \quad \mu^2 \frac{\partial T_{1,2}^d}{\partial \mu^2} \bigg|_{\mathcal{O}(\alpha^2)} = e_d^4 \mathcal{P}_{dd}^{+(0,2)} \otimes T_{1,2}^d, \\ \mu^2 \frac{\partial V_{1,2}^u}{\partial \mu^2} \bigg|_{\mathcal{O}(\alpha^2)} &= e_u^4 \mathcal{P}_{uu}^{-(0,2)} \otimes V_{1,2}^u, \quad \mu^2 \frac{\partial V_{1,2}^d}{\partial \mu^2} \bigg|_{\mathcal{O}(\alpha^2)} = e_d^4 \mathcal{P}_{dd}^{-(0,2)} \otimes V_{1,2}^d. \end{aligned} \quad (\text{A.19})$$

It should be noted that, as compared to the expressions for $\mathcal{P}^{(0,2)}$ presented in Ref. [47], the electric charges haven been factored out in such a way that the expressions of the splitting functions are either independent from the electric charges themselves or depend on them only through the ratio e_Σ^2/e_q^2 .

As an illustration, the effects of the $\mathcal{O}(\alpha_s\alpha)$ and $\mathcal{O}(\alpha^2)$ corrections to the DGLAP evolution equations on the $\gamma\gamma$ luminosity at $\sqrt{s} = 13$ TeV are quantified. This luminosity is defined as:

$$\mathcal{L}_{\gamma\gamma}(M_X) = \frac{1}{s} \int_{M_X^2/s}^1 \frac{dx}{x} \gamma(x, M_X^2) \gamma\left(\frac{M_X^2}{xs}, M_X^2\right), \quad (\text{A.20})$$

as a function of the final state invariant mass M_X . Fig. 12 illustrates the behaviour of $\mathcal{L}_{\gamma\gamma}$ computed using the photon PDF from the NNPDF3.0QED NLO set as an input at $Q_0 = 1$ GeV and evolved to $Q = M_X$ including, on top of the pure QCD NLO evolution, the following corrections:

- the $\mathcal{O}(\alpha)$ corrections only,
- same as above, adding also the mixed $\mathcal{O}(\alpha_s\alpha)$ corrections, and
- the complete NLO QCD+QED corrections accounting for the $\mathcal{O}(\alpha + \alpha_s\alpha + \alpha^2)$ effects.

The results are shown normalised to the predictions obtained with LO QED corrections only. It is clear that the $\mathcal{O}(\alpha_s\alpha)$ and $\mathcal{O}(\alpha^2)$ corrections have a small but non-negligible impact on the $\gamma\gamma$ -luminosity. In particular, these corrections suppress $\mathcal{L}_{\gamma\gamma}$ by around 10% at relatively small values of M_X , while the suppression gradually shrinks to 1-2% as M_X increases. As expected, most of this effect comes from the $\mathcal{O}(\alpha_s\alpha)$ corrections, while the impact of the $\mathcal{O}(\alpha^2)$ ones is substantially smaller. These corrections to the DGLAP evolution have more recently been implemented also in the QEDDEVOL package [17] based on the QCDNUM evolution code [18]. APFEL and QEDDEVOL have been found to be in excellent agreement.

Appendix A.3: DIS structure functions

When considering NLO QCD+QED corrections to the DIS structure functions, it becomes necessary to include into the hard cross sections all the $\mathcal{O}(\alpha)$ diagrams where one photon is either in the initial state or emitted from an incoming quark (or possibly an incoming lepton). Such diagrams, being of purely QED origin, have associated coefficient functions that can be easily derived from the QCD expressions by properly adjusting the colour factors. This correspondence holds irrespective of whether mass effects are included.

The main complication of the inclusion of these corrections arises from their flavour structure. In fact, in the

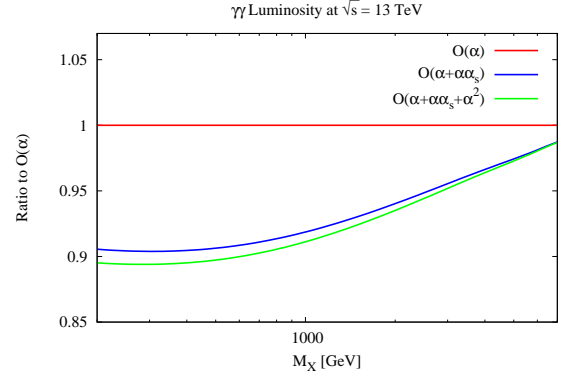


Figure 12 The photon-photon PDF luminosity $\mathcal{L}_{\gamma\gamma}$ at $\sqrt{s} = 13$ TeV as a function of the final state invariant mass M_X . The results with the photon evolved with only the $\mathcal{O}(\alpha)$ corrections are compared with the corresponding results taking into account the $\mathcal{O}(\alpha + \alpha_s\alpha)$ corrections and the complete $\mathcal{O}(\alpha + \alpha_s\alpha + \alpha^2)$ effects, normalised in all three cases to the $\mathcal{O}(\alpha)$ result. The calculation has been performed using the central value of the NNPDF3.0QED NLO fit.

case of quarks the isospin symmetry is broken due to the fact that the coupling of the photon is proportional to the squared charge of the parton to which it couples (a quark or a lepton). In the following, the neutral-current (NC) case, where lepton and proton exchange a neutral boson γ^*/Z , and the charged-current (CC) case, where instead lepton and proton exchange a charged W boson, are addressed separately.

First the $\mathcal{O}(\alpha)$ contributions to a generic NC structure function F are considered. Due to the fact that to this order there is no mixing between QCD and QED, such corrections can easily be derived from the $\mathcal{O}(\alpha_s)$ coefficient functions just by adjusting the colour factors by setting $C_F = T_R = 1$ and $C_A = 0$. Referring, *e.g.*, to the expressions reported in Ref. [48], the coefficient functions become:

$$C_{i;q}^{(\alpha)} = \frac{C_{i;q}^{(\alpha_s)}}{C_F} \quad i = 2, L, 3, \quad (\text{A.21})$$

$$C_{i;\gamma}^{(\alpha)} = \frac{C_{i;\gamma}^{(\alpha_s)}}{T_R}$$

In order to construct the corresponding structure functions, considering that the coupling between a photon and a quark of flavour q is proportional to e_q^2 , the electroweak couplings should be adjusted as follows:

$$\begin{aligned} \tilde{B}_q &= B_q e_q^2 \quad \text{for } F_2, F_L, \\ \tilde{D}_q &= D_q e_q^2 \quad \text{for } F_3, \end{aligned} \quad (\text{A.22})$$

where B_q and D_q are defined, *e.g.*, in Ref. [49]. Following this prescription, it is possible to write the $\mathcal{O}(\alpha)$ contributions to the NC structure functions as:

$$\begin{aligned} F_{2,L}^{\text{NC},(\alpha)} &= x \sum_q \tilde{B}_q \left[C_{2,L;q}^{(\alpha)} \otimes (q + \bar{q}) + C_{2,L;\gamma}^{(\alpha)} \otimes \gamma \right], \\ xF_3^{\text{NC},(\alpha)} &= x \sum_q \tilde{D}_q \left[C_{3;q}^{(\alpha)} \otimes (q - \bar{q}) + C_{3;\gamma}^{(\alpha)} \otimes \gamma \right]. \end{aligned} \quad (\text{A.23})$$

This structure holds for both massless and massive structure functions. This aspect is relevant to the construction of the FONLL general-mass structure functions.

For the CC case the procedure to obtain the expressions of the $\mathcal{O}(\alpha)$ coefficient functions is exactly the same as in the NC case (see Eq. (A.21)). However, this case is more complicated because the flavour structure of CC structure functions is more complex. Taking into account the presence of a factor e_q^2 every time that a quark of flavour q couples to a photon, the $\mathcal{O}(\alpha)$ corrections to the CC structure functions F_2 and F_L for the production of a neutrino or an anti-neutrino take the form:

$$\begin{aligned} F_{2,L}^{\text{CC},v,(\alpha)} &= x \sum_{\substack{U=u,c,t \\ D=d,s,b}} |V_{UD}|^2 \left[C_{2,L;q}^{(\alpha)} \otimes (e_D^2 D + e_U^2 \bar{U}) + 2C_{2,L;\gamma}^{(\alpha)} \otimes \gamma \right], \\ F_{2,L}^{\text{CC},\bar{v},(\alpha)} &= x \sum_{\substack{U=u,c,t \\ D=d,s,b}} |V_{UD}|^2 \left[C_{2,L;q}^{(\alpha)} \otimes (e_D^2 \bar{D} + e_U^2 U) + 2C_{2,L;\gamma}^{(\alpha)} \otimes \gamma \right], \end{aligned} \quad (\text{A.24})$$

where V_{UD} are the elements of the CKM matrix. The flavour structure of F_3 is instead slightly different:

$$\begin{aligned} xF_3^{\text{CC},v,(\alpha)} &= x \sum_{\substack{U=u,c,t \\ D=d,s,b}} |V_{UD}|^2 \left[C_{3;q}^{(\alpha)} \otimes (e_D^2 D - e_U^2 \bar{U}) + 2C_{3;\gamma}^{(\alpha)} \otimes \gamma \right], \\ xF_3^{\text{CC},\bar{v},(\alpha)} &= x \sum_{\substack{U=u,c,t \\ D=d,s,b}} |V_{UD}|^2 \left[C_{3;q}^{(\alpha)} \otimes (-e_D^2 \bar{D} + e_U^2 U) + 2C_{3;\gamma}^{(\alpha)} \otimes \gamma \right]. \end{aligned} \quad (\text{A.25})$$

In order to simplify the implementation, it is advantageous to assume that, in these particular corrections, the CKM matrix is a 3×3 unitary matrix. Note however that the exact CKM matrix is still used in the QCD part of the structure functions. This approximation introduces an inaccuracy of the order of the QED coupling α times the value of the off-diagonal elements of the CKM matrix and therefore it is numerically negligible.

As an illustration of the impact of the $\mathcal{O}(\alpha)$ correction on the DIS structure functions, Fig. 13 shows the effect of introducing these contributions on top of the pure QCD computation at NLO. The plots are produced by evolving the NNPDF3.0QED NLO set from $Q_0 = 1$ GeV to $Q = 100$ GeV including the full NLO QCD+QED corrections discussed in the previous section and using the resulting evolved PDFs to compute the NC (left panel) and the CC (right panel) DIS structure functions in the FONLL-B scheme, including

the $\mathcal{O}(\alpha)$ corrections to the coefficient functions discussed above. The predictions are shown normalised to the pure QCD computation where the QED corrections are absent both in the evolution and in the computation of the structure functions.

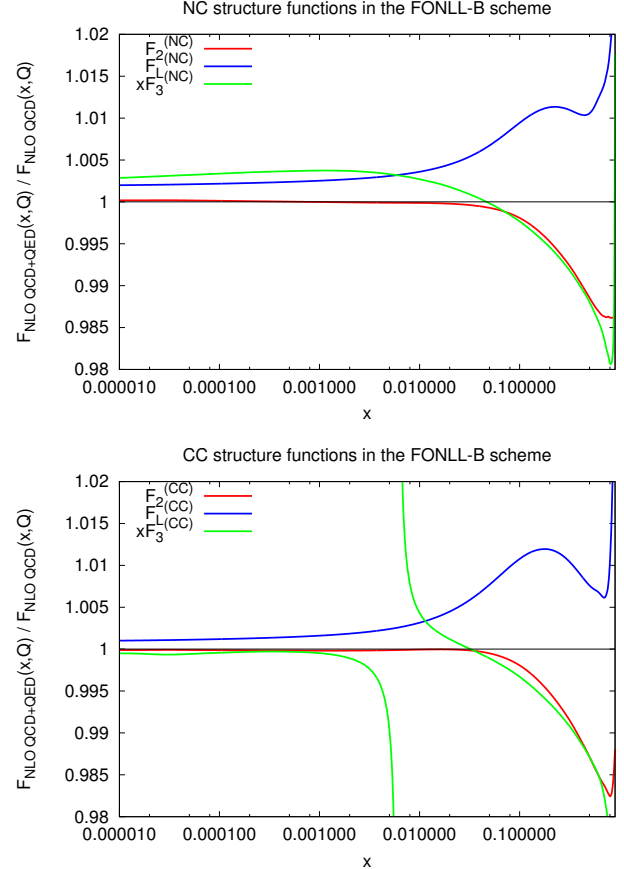


Figure 13 The effects of the NLO QED corrections on the neutral-current (left) and charged-current (right) DIS structure functions F_2, F_L and xF_3 , normalised to the pure QCD results. The calculation has been performed in the FONLL-B general-mass scheme using the central NNPDF3.0QED NLO set as input. Note that QED effects enter both via DGLAP evolution and the $\mathcal{O}(\alpha)$ DIS coefficient functions. The behaviour of xF_3 in the right plot for $x \sim 0.007$ is explained by the fact that this structure function exhibits a node in that region.

It is clear that the impact of the full NLO QCD+QED corrections is pretty small especially in the low- x region where it is well below 1%. In the large- x region, instead, the presence of a photon-initiated contribution has a more significant effect because of the suppression of the QCD distributions (quarks and gluon) relative to the photon PDF and the impact of the QED corrections reaches the 2% level. It should be stressed that the behaviour around $x = 10^{-2}$ of the CC xF_3 (green curve in the right panel) is driven by a change of sign of the predictions so that the ratio diverges.

References

1. A.D. Martin, R.G. Roberts, W.J. Stirling, R.S. Thorne, *Eur. Phys. J.* **C39**, 155 (2005). DOI 10.1140/epjc/s2004-02088-7
2. R.D. Ball, et al., *Nucl.Phys.* **B867**, 244 (2013). DOI 10.1016/j.nuclphysb.2012.10.003
3. R.D. Ball, et al., *Nucl.Phys.* **B877**, 290 (2013). DOI 10.1016/j.nuclphysb.2013.10.010
4. V. Bertone, S. Carrazza, (2016)
5. R.D. Ball, et al., *JHEP* **04**, 040 (2015). DOI 10.1007/JHEP04(2015)040
6. C. Schmidt, J. Pumplin, D. Stump, C.P. Yuan, *Phys. Rev.* **D93**(11), 114015 (2016). DOI 10.1103/PhysRevD.93.114015
7. A. Manohar, P. Nason, G.P. Salam, G. Zanderighi, (2016)
8. L.A. Harland-Lang, V.A. Khoze, M.G. Ryskin, *Phys. Rev.* **D94**(7), 074008 (2016). DOI 10.1103/PhysRevD.94.074008
9. G. Aad, et al., *JHEP* **08**, 009 (2016). DOI 10.1007/JHEP08(2016)009
10. V. Khachatryan, et al., *Eur. Phys. J.* **C75**(4), 147 (2015). DOI 10.1140/epjc/s10052-015-3364-2
11. S. Chatrchyan, et al., *JHEP* **12**, 030 (2013). DOI 10.1007/JHEP12(2013)030
12. G. Aad, et al., *Phys.Lett.* **B725**, 223 (2013). DOI 10.1016/j.physletb.2013.07.049
13. R.D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, et al., *Nucl.Phys.* **B855**, 608 (2012). DOI 10.1016/j.nuclphysb.2011.10.018
14. R.D. Ball, et al., *Nucl. Phys.* **B849**, 112 (2011). DOI 10.1016/j.nuclphysb.2011.03.017
15. S. Alekhin, et al., *Eur. Phys. J.* **C75**(7), 304 (2015). DOI 10.1140/epjc/s10052-015-3480-z
16. V. Bertone, S. Carrazza, J. Rojo, *Comput.Phys.Comm.* **185**, 1647 (2014). DOI 10.1016/j.cpc.2014.03.007
17. R. Sadykov, (2014)
18. M. Botje, *Comput.Phys.Comm.* **182**, 490 (2011). DOI 10.1016/j.cpc.2010.10.020
19. H. Abramowicz, et al., *Eur. Phys. J.* **C75**(12), 580 (2015). DOI 10.1140/epjc/s10052-015-3710-4
20. J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al., *JHEP* **1407**, 079 (2014). DOI 10.1007/JHEP07(2014)079
21. T. Carli, D. Clements, A. Cooper-Sarkar, C. Gwenlan, G.P. Salam, F. Siegert, P. Starovoitov, M. Sutton, *Eur. Phys. J.* **C66**, 503 (2010). DOI 10.1140/epjc/s10052-010-1255-0
22. V. Bertone, R. Frederix, S. Frixione, J. Rojo, M. Sutton, *JHEP* **1408**, 166 (2014). DOI 10.1007/JHEP08(2014)166
23. R. Gavin, Y. Li, F. Petriello, S. Quackenbush, *Comput.Phys.Comm.* **184**, 208 (2013). DOI 10.1016/j.cpc.2012.09.005
24. V.N. Gribov, L.N. Lipatov, *Sov. J. Nucl. Phys.* **15**, 438 (1972). [*Yad. Fiz.*15,781(1972)]
25. Y.L. Dokshitzer, *Sov. Phys. JETP* **46** (1977)
26. G. Altarelli, G. Parisi, *Nucl. Phys.* **B126**, 298 (1977). DOI 10.1016/0550-3213(77)90384-4
27. S. Forte, E. Laenen, P. Nason, J. Rojo, *Nucl. Phys.* **B834**, 116 (2010). DOI 10.1016/j.nuclphysb.2010.03.014
28. K. Olive, et al., *Chin.Phys.* **C38**, 090001 (2014). DOI 10.1088/1674-1137/38/9/090001
29. L.A. Harland-Lang, V.A. Khoze, M.G. Ryskin, *Phys. Lett.* **B761**, 20 (2016). DOI 10.1016/j.physletb.2016.08.004
30. S. Dittmaier, M. Huber, *JHEP* **01**, 060 (2010). DOI 10.1007/JHEP01(2010)060
31. L.A. Harland-Lang, A.D. Martin, P. Motylinski, R.S. Thorne, *Eur. Phys. J.* **C75**(5), 204 (2015). DOI 10.1140/epjc/s10052-015-3397-6
32. M. Czakon, N.P. Hartland, A. Mitov, E.R. Nocera, J. Rojo, (2016)
33. F. Olness, V. Bertone, work in progress
34. F. Aaron, et al., *JHEP* **1209**, 061 (2012). DOI 10.1007/JHEP09(2012)061
35. F. Aaron, et al., *JHEP* **1001**, 109 (2010). DOI 10.1007/JHEP01(2010)109
36. G. Aad, et al., *Phys.Rev.Lett.* (2012)
37. M. Aaboud, et al., (2016)
38. S.S. Wilks, *Annals Math. Statist.* **9**(1), 60 (1938). DOI 10.1214/aoms/1177732360
39. L. Del Debbio, S. Forte, J.I. Latorre, A. Piccione, J. Rojo, *JHEP* **03**, 080 (2005). DOI 10.1088/1126-6708/2005/03/080
40. L. Del Debbio, S. Forte, J.I. Latorre, A. Piccione, J. Rojo, *JHEP* **03**, 039 (2007)
41. R.D. Ball, et al., *Nucl. Phys.* **B809**, 1 (2009). DOI 10.1016/j.nuclphysb.2008.09.037
42. J. Pumplin, et al., *Phys. Rev.* **D65**, 014013 (2001). DOI 10.1103/PhysRevD.65.014013
43. A. Buckley, J. Ferrando, S. Lloyd, K. Nordstrom, B. Page, et al., *Eur.Phys.J.* **C75**, 132 (2015). DOI 10.1140/epjc/s10052-015-3318-8
44. L.R. Surguladze, (1996)
45. V. Bertone, S. Carrazza, D. Pagani, M. Zaro, *JHEP* **11**, 194 (2015). DOI 10.1007/JHEP11(2015)194
46. D. de Florian, G.F.R. Sborlini, G. Rodrigo, *Eur. Phys. J.* **C76**(5), 282 (2016). DOI 10.1140/epjc/s10052-016-4131-8
47. D. de Florian, G.F.R. Sborlini, G. Rodrigo, *JHEP* **10**, 056 (2016). DOI 10.1007/JHEP10(2016)056

- 48. R.K. Ellis, W.J. Stirling, B.R. Webber, *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **8**, 1 (1996)
- 49. C. Adloff, et al., *Eur. Phys. J.* **C30**, 1 (2003). DOI 10.1140/epjc/s2003-01257-6