Determination of photon PDF from High Mass Drell Yan data at LHC

F. Giulli and The xFitter Collaboration: V. Berone, A. Cooper-Sarkar,
A. Glazov, R. Placakyte, V. Radescu, J. Rojo, A. Sapronov, etc.
(Dated: November 24, 2016)

abstract goes here: ...

CONTENTS

I. INTRODUCTION

II. THEORY

Two processes contribute to opposite sign, same family, dilepton production at the LHC: the Drell-Yan quark-antiquark process and the photon-induced process. Both the contributions can be simulated with MadGraph5_aMC@NLO (version 2.4.3) and interfaced to APPLgrid (version 01-04-70) and aMCfast (version 01-03-00). A special release of APPLgrid is used to account for the photon PDF within the proton *need references for the programmes.* Both contributions are generated in the 5-flavour scheme, where all the quarks, except for the *top* quark, are treated as massless quarks; all the calculations are performed at fixed-order (FO) without parton showers.

Theoretical predictions for both the one-dimensional $\frac{d\sigma}{dm_{ll}}$ distribution (where m_{ll} is the invariant mass of the dilepton pair in the final state) and the double-differential distributions $\frac{d^2\sigma}{dm_{ll}d|y_{ll}|}$ (where $|y_{ll}|$ is the rapidity of the dilepton pair) and $\frac{d^2\sigma}{dm_{ll}\Delta\eta_{ll}}$ (where $\Delta\eta_{ll}$ represents the difference in pseudorapidity between the two leptons) are generated for both the electron and the muon channels.

These predictions are generated using the same selections as in reference [?] as follows:

- the invariant mass of the lepton pair is required to be greater than 116 GeV;
- the absolute value of the pseudorapidity of each lepton is required to be less than 2.5;
- the transverse momentum (p_T) of the leading lepton has to be greater than 40 GeV;
- the p_T of the sub-leading lepton has to be greater than 30 GeV.

The binning used is the same as used in reference [?]. For the invariant mass distribution, there are 12 bins between 116 GeV and 1.5 TeV with variable bin widths; and for both of the the two-dimensional distributions, there are five different histograms, each one for a different invariant mass range: (a) 116 GeV $< m_{ll} <$ 150 GeV; (b) 150 GeV $< m_{ll} <$ 200 GeV; (c) 200 GeV $< m_{ll} <$ 300 GeV; (d) 300 GeV $< m_{ll} <$ 500 GeV; (e) 500 GeV $< m_{ll} <$ 1500 GeV. The APPLgrids for the first three m_{ll} intervals are divided into 12 bins with fixed bin width between $|y_{ll}^{mim}|$ ($|\Delta\eta_{ll}|$) = 0.0 (0.0) and $|y_{ll}^{max}|$ ($|\Delta\eta_{ll}|$) = 2.4 (3.0), while the final two m_{ll} intervals are divided into 6 bins with fixed bin width scanning the same $|y_{ll}|$ and $|\Delta\eta_{ll}|$ ranges.

Dynamical renormalization (μ_R) and factorization (μ_R) scales are used in the calculations and both are set to m_{ll}. The theoretical calculations were validated by comparing both the NLO QCD + LO EW predictions and the LO PI predictions to those computed using the FEWZ 3.1 framework. These calculations are evaluated in the G_F electroweak scheme, with the following values for the couplings: $\alpha_S = 0.118$; $1/\alpha_{EW} = 1/127$. The difference between the two predictions is at most 1%, for both the 1-dimensional and the 2-dimensional distributions.

In order to make a next-to-next-to-leading order (NNLO) fit k-factors (k_F) are computed matching the NLO QCD + LO EW cross sections to higher order (HO) calculations. These are computed using FEWZ, with the same input parameters as for the NLO computations. The k_F are defined as:

$$k_F = \frac{NNLO\ QCD + NLO\ EW\sigma}{NLO\ QCD + LO\ EW\sigma} \quad (1)$$

The MMHT2014NNLO PDF set is used to compute both numerator and denominator. The k_F are close to the unity and their variation is $\sim 2\%$. *provide Table of Final k-factors?*

Discuss theory improvements: addition of the NLO QED+QCD piece

III. RESULTS

A. Sensitivity

show impact of HM DY on PDFs using sensitivity studies based on pseudo-data, for which we only use the data uncertainties, while central value are fixed: HERA I+II vs HERA I+II + HMDY –> see the sensitivity plots from the previous email

conclusion: HMDY data has a large impact on photonPDF

B. PDF Fits

In order to make a full PDF fit the ATLAS Drell-Yan data data are fitted together with the final combined inclusive cross section data from HERA [?]. The HERA data provide information on the quark/antiquark and gluon content of the proton and the Drell-Yan data add information on the photon content of the proton. The NLO and NNLO pQCD predictions are fitted to the data using the xFitter open source pQCD fitting platform [?]. The DGLAP equations [?] are solved using the programme QCDNUM which has been modified to include the photon PDF in the proton [?]. The DGLAP equations yield the PDFs at all scales if they are input as finctions of x at a starting scale Q_0^2, which should be large enough that perturbative QCD can be assumed to be valid. For the present analysis this value is chosen to be $Q_0^2 = 7.5$ GeV^2. This is also the value chosen for the minimum value of Q^2 for data entering the fit. The charm and beauty masses are chosen to be $m_c = 1.47$ GeV and $m_b = 4.5$ GeV following the HERA analysis. The value of $\alpha_s(M_Z)$ is chosen to be $\alpha_s(M_Z) = 0.118$ [?]. The value of Q_0^2 is above the charm mass squared, however a version of the programme is used which displaces the charm threshold from the charm mass [?] such that the threshold is at Q_0^2. The form of the χ^2 used for the fit is that defined in the H1 paper [?]. Alternative forms have also been tried with no significant difference to our results.

The PDF parametrisation input at Q_0^2 is determined by the technique of saturation of the χ^2 [?]. The parametrised PDFs are the valence distributions xu_v and xd_v, the gluon distribution xg, and the u-type and d-type sea, $x\bar{U}$, $x\bar{D}$, where $x\bar{U} = x\bar{u}$ and $x\bar{D} = x\bar{d} + x\bar{s}$, and finally the photon distribution $x\gamma$. The following standard functional form is used to parametrise them:

$$xf(x) = Ax^B(1-x)^C(1 + Dx + Ex^2) \tag{2}$$

where the normalisation parameters A_{u_v}, A_{d_v} and A_g are constrained by the number sum-rules and the momentum sum-rule, respectively. The B parameters $B_{\bar{U}}$ and $B_{\bar{D}}$ are set equal, such that there is a single B parameter for the sea distribution. The data are not sensitive to the strangeness content of the proton which is thus set such that $x\bar{s} = 0.5\bar{D}$, following the ATLAS analysis [?]. The further constraint $A_{\bar{U}} = 0.5A_{\bar{D}}$ is imposed such that $\bar{u} = x\bar{d}$ as $x \to 0$. The D and E parameters are introduced one by one until no significant improvement in χ^2 is found.

For the NNLO fit a $\chi^2/ndf = 1.18$, with a partial $\chi^2/ndp = 1.15$ for the high-mass Drell-yan data [*update with final numbers*], is achieved for the following parametrisation, which has 11 parameters for the quarks and gluons and 5 parameters for the photon:

$$\begin{aligned}
xu_v(x) &= A_{u_v}x^{B_{u_v}}(1-x)^{C_{u_v}}(1 + E_{u_v}x^2), && (3)\\
xd_v(x) &= A_{d_v}x^{B_{d_v}}(1-x)^{C_{d_v}}, && (4)\\
x\bar{U}(x) &= A_{\bar{U}}x^{B_{\bar{U}}}(1-x)^{C_{\bar{U}}}, && (5)\\
x\bar{D}(x) &= A_{\bar{D}}x^{B_{\bar{D}}}(1-x)^{C_{\bar{D}}}, && (6)\\
xg(x) &= A_g x^{B_g}(1-x)^{C_g}(1 + E_g x^2), && (7)\\
x\gamma(x) &= A_\gamma x^{B_\gamma}(1-x)^{C_\gamma}(1 + D_\gamma x + E_\gamma x^2) && (8)\\
& && (9)
\end{aligned}$$

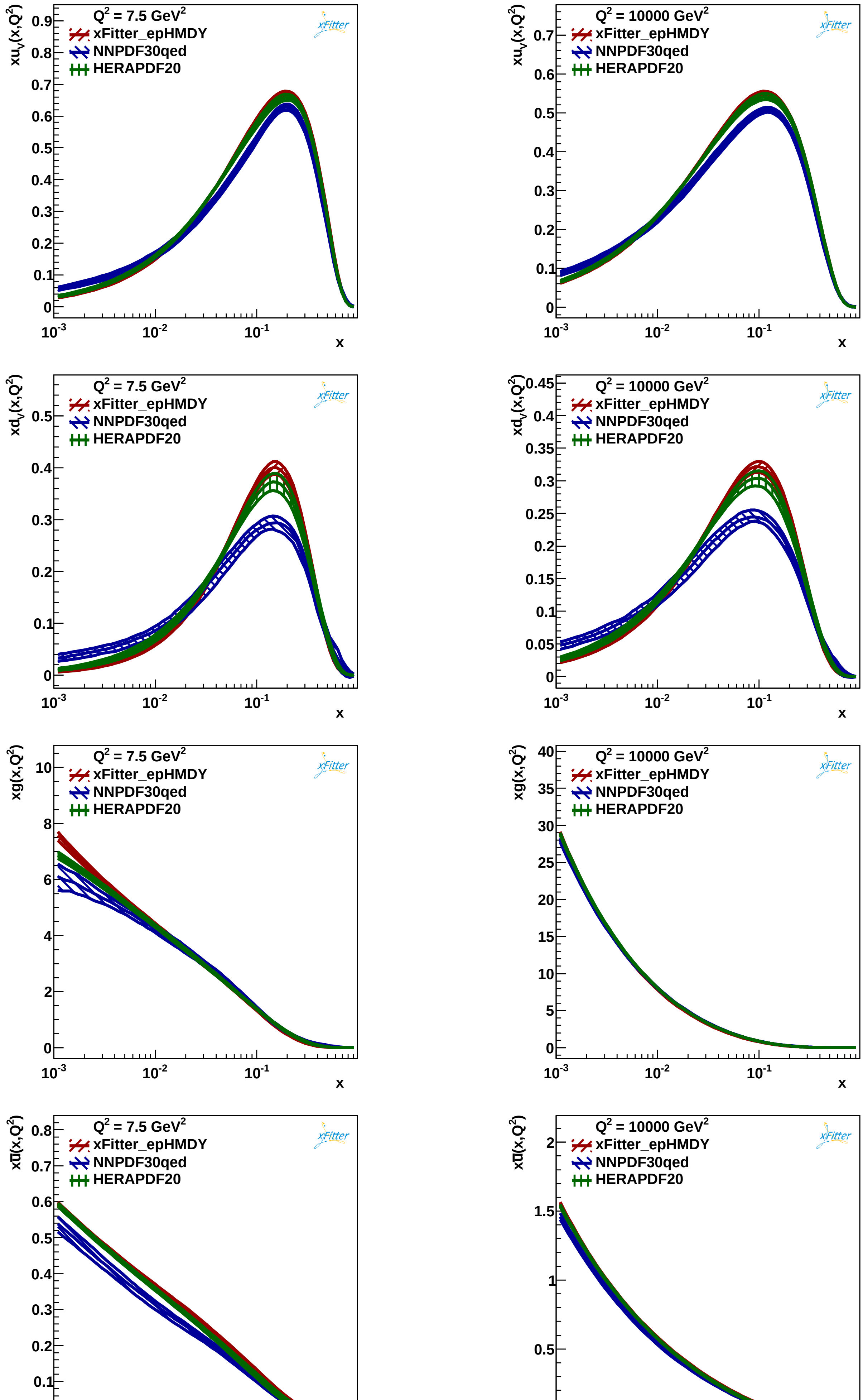
The parametrisation for HERA data differs from that of the HERAPDF2.0 PDF since the starting scale Q_0^2 is higher and the additional negative term in the gluon parametrisation is not necessary. Parametrisation and model uncertainties are considered according to the HERAPDF procedure [?] by adding extra terms which make little difference to the χ^2 of the fit, but which can change the shape of the PDFs. Additonal parameters considered are: the extra negative term for the gluon; D_{u_v}, $D_{\bar{u}}$ and $E_{\bar{d}}$. Model variations considered are the variation of: m_b from 4.25 to 4.75 GeV; m_c from 1.41 to 1.53 GeV; Q_0^2 up to 10 GeV^2; Q^2_{cut} up to 10 GeV^2; the strangeness fraction down to $f_s = 0.4$; the value of $\alpha_s(M_Z)$ from 0.116 to 0.120.

Fig. 1 shows the PDF distributions $x_{u_v}, xd_{d_v}, x\bar{u}, x\bar{d}, xg$ at $Q^2 = 7.5^2$ GeV^2, including model and parametrisation uncertainties, while Fig. 2 shows them at $Q^2 = 10^4$ GeV^2. *Add model and parametrisation variations, Use NNLO MC central.* In these figures comparisons are made to the NNPDF3.0PDF set and the HERAPDF2.0 set. The shape of the xd_{u_v} distribution is close to that of HERAPDF2.0 because of the dominance of HERA data in the fit.

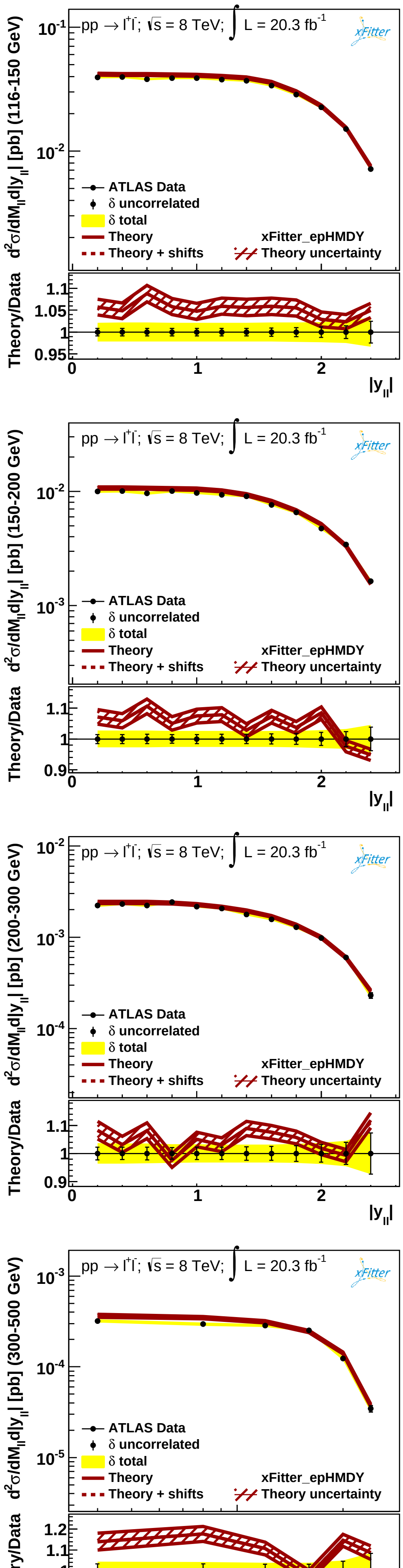
Fig. 3 shows the comparison between the high-mass Drell-Yan double differential distribution and the predictions. The χ^2 values for the high-mass Drell Yan data and the output parameters from NNLO fit can be found in Table. 4 and Table. 5 respectively.

The NNLO photon PDF distribution is shown both at the starting scale (7.5 GeV^2) and at 10^4 GeV^2 in Fig. 6, where it is also compared to an NLO extraction of the photon distribution. The x-range of the figure is restricted to the range of sensitivity of the high-mass drell-Yan data; $0.045 < x < 0.35$. The NLO and NNLO photon PDFs are compatible over this range.

Fig. 7 shows the photon distribution in the restricted range compared to the NNPDF3.0qed NNLO photon PDF. The uncertainties are considerably reduced. The comparison is shown at scale 100 GeV^2 and at 10^4 GeV^2, where the value of 100 GeV^2 is chosen such that comparisons can also be made to the LUXqed [?] photon PDF, which is only defined above this scale. The HKR photon PDF [?] is also shown in this figure. The fit predictions from the present analysis agree with the LUXqed and the HKR photon PDFs at the 1-σ level.



Q² = 7.5 GeV²
xFitter_epHMDY
NNPDF30qed
HERAPDF20
xFitter
xu_v(x,Q²)
x
Q² = 10000 GeV²
xFitter_epHMDY
NNPDF30qed
HERAPDF20
xFitter
xu_v(x,Q²)
x
Q² = 7.5 GeV²
xFitter_epHMDY
NNPDF30qed
HERAPDF20
xFitter
xd_v(x,Q²)
x
Q² = 10000 GeV²
xFitter_epHMDY
NNPDF30qed
HERAPDF20
xFitter
xd_v(x,Q²)
x
Q² = 7.5 GeV²
xFitter_epHMDY
NNPDF30qed
HERAPDF20
xFitter
xg(x,Q²)
x
Q² = 10000 GeV²
xFitter_epHMDY
NNPDF30qed
HERAPDF20
xFitter
xg(x,Q²)
x
Q² = 7.5 GeV²
xFitter_epHMDY
NNPDF30qed
HERAPDF20
xFitter
xū(x,Q²)
Q² = 10000 GeV²
xFitter_epHMDY
NNPDF30qed
HERAPDF20
xFitter
xū(x,Q²)



Dataset	xFit epH
HMDY rap 116-150	9.3 /
HMDY rap 150 200	17 /
HMDY rap 200 300	15 /
HMDY rap 300 500	3.8 /
HMDY rap 500 1500	4.2 /
Correlated χ^2	4.98
Log penalty χ^2	-3.64
Total χ^2 / dof	55 /
χ^2 p-value	0.01

Figure 4. χ^2 for high-mass Drell Yan data, for the NNLO fit

IV. CONCLUSIONS

Parameter	xFitter epHMDY
'Bg'	$-0.220^{+0.014}_{-0.013}$
'Cg'	$6.92^{+0.62}_{-0.61}$
'Buv'	$0.761^{+0.017}_{-0.015}$
'Cuv'	$5.060^{+0.064}_{-0.092}$
'Euv'	$8.07^{+0.80}_{-0.82}$
'Bdv'	$1.009^{+0.050}_{-0.056}$
'Cdv'	$5.61^{+0.24}_{-0.30}$
'Cubar'	$6.37^{+0.56}_{-0.38}$
'Adbar'	$0.3226^{+0.0078}_{-0.0083}$
'Bdbar'	$-0.1921^{+0.0033}_{-0.0033}$
'Cdbar'	$14.0^{+2.0}_{-1.7}$
'alphas'	0.1180
'rs'	1.0000
'Aph'	$0.00120^{+0.031}_{-0.00089}$
'Bph'	$-0.62^{+0.63}_{-0.36}$
'Cph'	$10.0^{+13}_{-5.9}$
'Dph'	-4^{+210}_{-15}
'Eph'	87^{+257}_{-140}
Fit status	MC-replica
Uncertainties	median±68cl

Figure 5. PDF parameters for the NNLO fit.

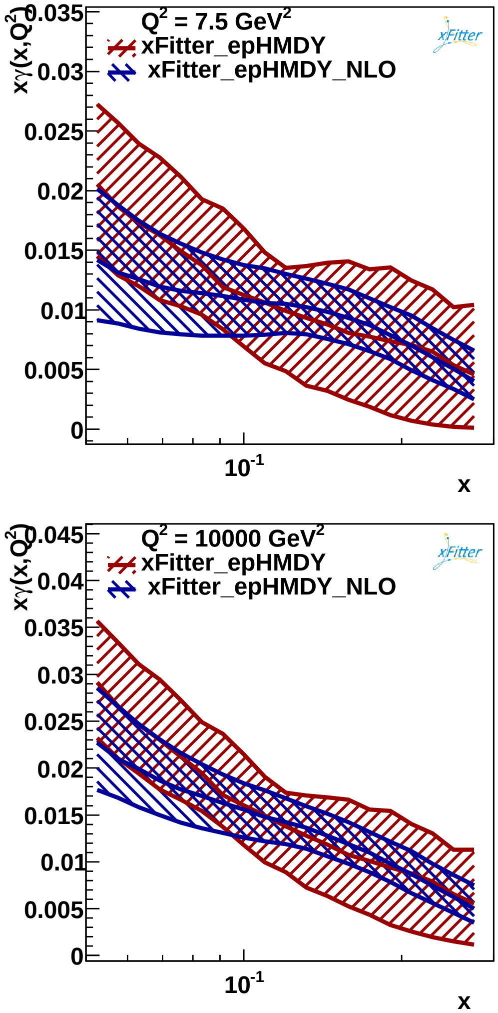


Figure 6. Comparison between the photon PDF distributions at NNLO and NLO: (a) at the starting scale; (b) at the evolved scale.

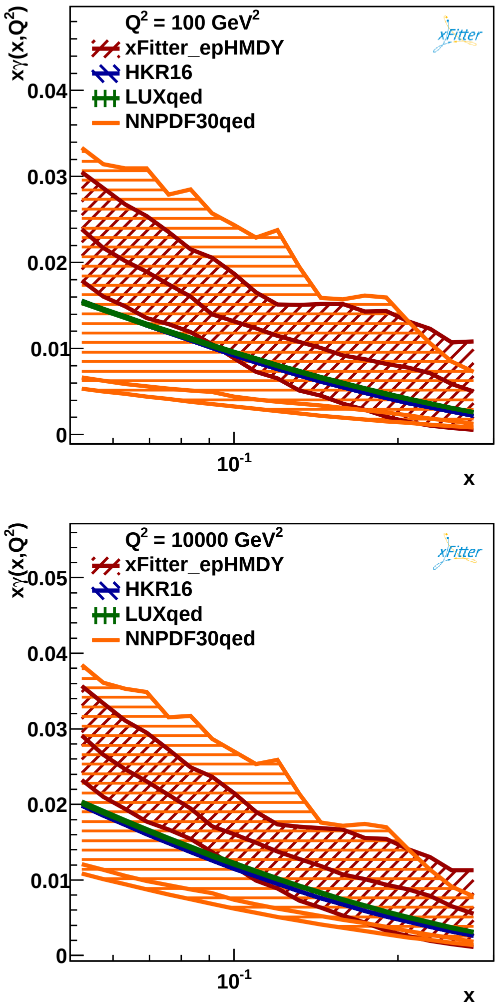


Figure 7. Comparison between the NNLO photon PDF distributions for the present analysis, NNPDF3.0QED, LUXqed, HKR: (a) at scale 100 GeV^2; (b) at the evolved scale 10000 GeV^2.