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 15, 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:

$$k_F = \frac{NNLO\ QCD + NLO\ EW\sigma}{NLO\ QCD + LO\ EW\sigma} \tag{1}$$

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. *and maybe refine the quark/antiquark content, refer sensitivity study.* 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 NLO fit a $\chi^2/ndf = 1225.3/1084 = 1.13$, with a partial $\chi^2/ndp = 46.9/48$ for the high-mass Drell-yan data [*please separate contribution of log term and correlated term in the partial chisq table to allow this calculation to be accurate- I estimated it*], is achieved for the following parametrisation, which has 15 parameters for the quarks and gluons and 5 parameters for the photon:

$$xu_v(x) = A_{u_v}x^{B_{uv}}(1-x)^{C_{uv}}(1 + D_{u_v}x + E_{u_v}x^2), \tag{3}$$

$$xd_v(x) = A_{d_v}x^{B_{d_v}}(1-x)^{C_{d_v}}, \tag{4}$$

$$x\bar{U}(x) = A_{\bar{U}}x^{B_{\bar{U}}}(1-x)^{C_{\bar{U}}}(1 + D_{\bar{U}}x + E_{\bar{U}}x^2), \tag{5}$$

$$x\bar{D}(x) = A_{\bar{D}}x^{B_{\bar{D}}}(1-x)^{C_{\bar{D}}}, \tag{6}$$

$$xg(x) = A_gx^{B_g}(1-x)^{C_g}(1 + E_gx^2), \tag{7}$$

$$x\gamma(x) = A_\gamma x^{B_\gamma}(1-x)^{C_\gamma}(1 + D_\gamma x + E_\gamma x^2) \tag{8}$$

$$\tag{9}$$

Figures **show the PDF distributions $x_{u_v}, xd_{d_v}, x\bar{u}, x\bar{d}, xg$ at $Q^2 = 10^2$ GeV^2, while Figures **show them at $Q^2 = 10^4$ GeV^2. *Redo figures for the just Dubar final parametrisation, conisder adding +Dg and Eubar as parametrisation variations. Also consider model variations like change of Q2cut, Q20, fs, mc,mb.* [*When showing PDF distributions for the NLO fit at two scales. show ubar and dbar as well, sbar,cbar,bbar are not interesting in the current context. Then repeat with this parametrisation at NNLO and quote NNLO partial chisq for DY data. Make plots of NNLo vs NLo only for the photon PDF. Here you will almost certainly find a worse overall chisq, but it will be worse for HERA which really likes the negative gluon term. It will probably not be worse for the DY. We do not need to discuss the HERA NNLO features here*]. In these figures comparisons are made to the NNPDF3.0PDF set [*use NNPDF3.0 not the rwg version*] and the HERAPDF2.0 set. One can see that 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. *More comments?*

Fig. shows the comparison between the hmDY two-dimensional distribution and the predictions. *Just make it for the 15 parameter Dubar fit.* The χ^2 values for each separate fitted dataset and the output parameters from the various fits can be found in Tables * and in Table *respectively. *Do this ONLy for the central +Dubar fit*

Focusing now on the photon PDF distribution, there is some impact on the photon PDF from adding parameters to the quark and gluon PDFS, particularly from adding a D_g and an $E_{\bar{U}}$ parameter. This impact has been included in the parametrisation variations and is shown both at the starting scale (7.5 GeV^2) and at 10^4 GeV^2 in Fig *

Fig shows the photon distribution in the range $0.045 < x < 0.35$, region where high mass DY data are most sensitive to this quantity. The new fit results have an uncertainty between 20% and 30%, which is considerably reduced compared to the NNPDF30qed NLO photon PDF which is also shown for comparison. *Not the rwg version.* The predictions for the LUXqed [**?**] photon PDF and the HKR photon PDF [**?**] are shown compared to the NNLO PDF from the present analysis in Fig.** *Compare our NNLO PDF when you have it.* In this kinematic region, the fit predictions agree with LUXqed and with the HKR photon PDF at the 1-σ level.

IV. CONCLUSIONS