

Study of Transfer Matrix for high-x ZEUS data

R. Aggarwal^{*a}, **C. Gwenlan**^{b †}

On behalf of the ZEUS Collaboration

^a*Savitribai Phule Pune University*

^b*University of Oxford*

E-mail: ritu.aggarwall@gmail.com, claire.gwenlan@physics.ox.ac.uk

The knowledge of the proton parton densities for large x is very important in the search for new physics signals at the LHC. For Bjorken- x larger than 0.6 they are however poorly constrained by the data used in extracting the proton parton density functions (pdfs) and different pdf sets have large uncertainties, and differ considerably, in this regime. We compare the pdf sets most widely used by the LHC community to the ZEUS high- x data. This data has not been previously used in pdf set determinations. Due to the small expected and observed numbers of events in this kinematic regime, Poisson statistics is used in the evaluation of the probabilities assigned to the different pdf sets. A wide variation is found in the ability of the pdf sets to predict the observed results.

*The 39th International conference on High Energy Physics (ICHEP2018)
4-11 July, 2018
Seoul, Korea*

*Supported by Dept. of Science and Tech. (DST), India.

†Speaker.

1. Introduction

Deep Inelastic Scattering is the cleanest source of information to the proton structure. The data provided by HERA covers six orders in magnitude in Q^2 and ranges in Bjorken- x from as low as $6 \cdot 10^{-7}$ for BPC data to as high as 1. The Neutral current data included in PDF sets has an upper limit on x of 0.65 from the DIS data [1]. Apart from that, our knowledge on the parton distribution functions primarily rely on the fixed target experiment data [2, 3]. The DIS data has an advantage over fixed target data in a way that it needs no corrections as needed by the latter. The ZEUS NC high- x ep cross sections [4] are the only high Q^2 DIS data available that spans x up to the value of 1. But this data has not been included in any of the PDF fits yet due to the following reason. At very high- x values where the true cross section is small, some of the bins have few number of events measured. If the number of events expected in different cross sections bins are provided by different theory PDF groups, these can then be compared to the number of events measured in data through poisson statistics and the problem of low statistics is no longer relevant. This paper presents the technique of building Transfer Matrix for the high- x ZEUS data which includes all the detector and simulation effects and would be helpful to predict the number of events in the cross sections bins.

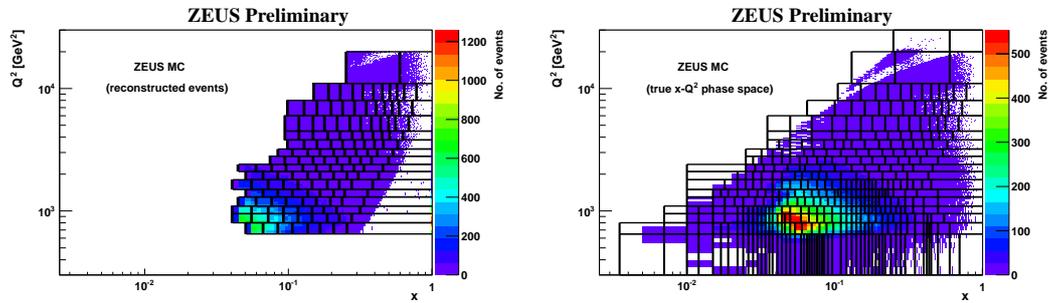


Figure 1: Distribution of event weights in the MC simulation after all analysis selections are applied, shown as functions of the (Left:) reconstructed kinematic quantities in cross section bins used in [4], (Right:) generated kinematic quantities with the binning scheme used in calculating the elements $V_{i,k}$.

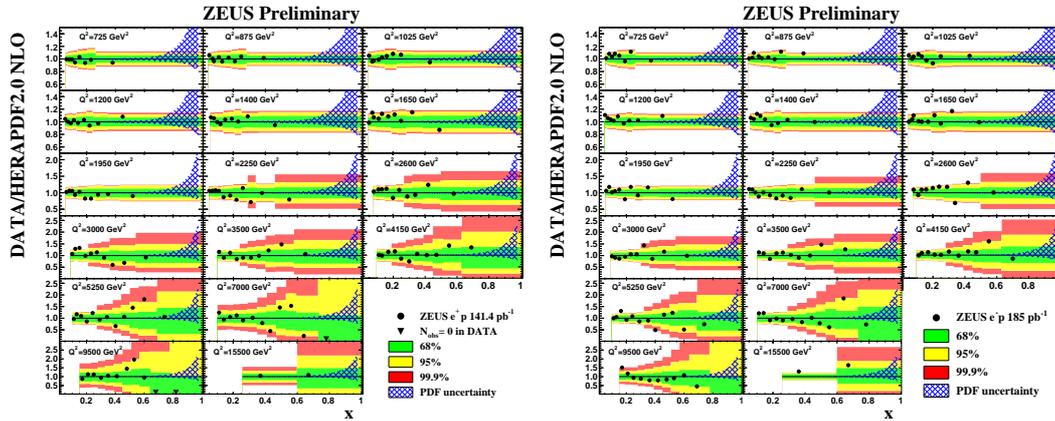


Figure 2: The ratio of the number of observed events to the expectations from the HERA2.0 PDF set for e^+p (left) and e^-p (right) data respectively. The green, yellow and red bands give the smallest intervals containing 68, 95, 99.9 % probability calculated using poisson statistics.

2. Transfer Matrix

Transfer Matrix gives the probability, an event generated with true x - Q^2 co-ordinates will have to appear in the cross section bins. Figure 1(left) shows the distribution of events reconstructed in the cross section bins. When we plot the true x - Q^2 distributions of these events we get what is

shown in Figure 1(right). To contain each of those events, a new binning is defined in the true x - Q^2 phase space (i.e. at generator level, using exchanged photon information) as shown which is finer than the cross section bins.

The number of events reconstructed in j^{th} cross section bin can be calculated from the born level cross sections $\frac{d^2\sigma(x, Q^2|M_k)}{dx dQ^2}$ at (x, Q^2) for PDF set k , using the equation given below.

$$v_{j,k} = \mathcal{L} \int_{(\Delta x, \Delta Q^2)_j} \left[\int T(x_{\text{rec}}, Q_{\text{rec}}^2 | x, Q^2) \frac{d^2\sigma(x, Q^2|M_k)}{dx dQ^2} dx dQ^2 \right] dx_{\text{rec}} dQ_{\text{rec}}^2 \quad (2.1)$$

where \mathcal{L} is the luminosity and $T(x_{\text{rec}}, Q_{\text{rec}}^2 | x, Q^2)$ is a transformation of the Born-level cross sections to an observed cross section that includes all relevant effects (radiative corrections, detector resolution and acceptance, selection criteria, etc.). Above integral can be approximated by the equation given below

$$v_{j,k} \approx \mathcal{L} \sum_i a_{ij} K_{ii} v_{i,k} \quad (2.2)$$

where $v_{i,k}$ are the number of events generated in the given Born level x - Q^2 bin (no radiative corrections included), K_{ii} are the radiative corrections for the given x - Q^2 bin which are calculated using HERACLES [5] and are applied as a scaling factor. Here a_{ij} are the the Transfer Matrix elements and it takes care of all detector effects including limited detector acceptance and various corrections applied on the simulated data. It is calculated in the MC as

$$a_{ij} = \frac{\sum_{m=1}^{M_i} \omega_m I(m \in j)}{\sum_{m=1}^{M_i} \omega_m^{MC}} \quad (2.3)$$

with M_i the number of Monte Carlo events generated in bin i , ω_m the total weight given to the m^{th} event, ω_m^{MC} the generated event weight and $I(m \in j) = 1$ if event m is reconstructed in bin j , else $I(m \in j) = 0$. The number of events in the cross section bins can be predicted using Equation 2.2.

Each element $v_{i,k}$ in Equation 2.2 for k^{th} PDF set was evaluated by reweighting the generated events using the cross sections calculated with the xFitter package [6].

$$v_{i,k} = \sum_m \frac{d^2\sigma(x, Q^2|M_k)/dx dQ^2}{d^2\sigma(x, Q^2|M_0)/dx dQ^2} \omega_m^{MC} \quad (2.4)$$

with $d^2\sigma(x, Q^2|M_0)/dx dQ^2$ the differential cross section evaluated using the HERAPDF2.0 PDF set.

PDF	e ⁻ p	e ⁺ p
HERAPDF2.0	0.05	0.5
CT14	0.002	0.8
MMHT2014	0.002	0.8
NNPDF2.3	0.00007	0.6
NNPDF3.0	0.0002	0.7
ABMP16	0.01	0.8
ABM11	0.001	0.6

Table 1: p-values for full Bjorken-x range for different PDFs (at NLO)

PDF	e ⁻ p		e ⁺ p	
	$x < 0.6$	$x \geq 0.6$	$x < 0.6$	$x \geq 0.6$
HERAPDF2.0	0.06	0.2	0.6	0.1
CT14	0.0008	0.2	0.7	0.6
MMHT2014	0.00003	0.1	0.6	0.6
NNPDF2.3	0.00007	0.2	0.6	0.6
NNPDF3.0	0.00003	0.2	0.6	0.6
ABMP16	0.01	0.2	0.8	0.5
ABM11	0.03	0.3	0.7	0.4

Table 2: p-values for two different x ranges

3. Results

Figure 2 gives a comparison of the observed number of events relative to the HERAPDF2.0 expectations for both the e⁺p (left) and e⁻p (right) data sets respectively. Also shown are the 68 %, 95 % and 99.9 % probabilities for the expected value calculated using the poisson statistics [7]. It is observed that the data generally are in the 68 % range of expected results.

The p -values [8] for the e⁺p and e⁻p data *wrt* HERAPDF2.0 were found to be 0.5 and 0.05 respectively. The summary of p -values calculated from different PDF sets as compared to the observed data is given in Table 1. For the e⁺p data, all PDF sets give good p -values indicating good

PDF	+1.8 %				-1.8 %			
	e ⁻ p		e ⁺ p		e ⁻ p		e ⁺ p	
	$x < 0.6$	$x \geq 0.6$						
HERAPDF2.0	0.02	0.1	0.2	0.3	0.03	0.3	0.8	0.2
CT14	0.02	0.3	0.8	0.5	0.0	0.08	0.4	0.6
MMHT2014	0.008	0.2	0.8	0.5	0.0	0.04	0.2	0.6
NNPDF2.3	0.009	0.3	0.8	0.4	0.0	0.08	0.2	0.6
NNPDF3.0	0.008	0.3	0.8	0.4	0.0	0.08	0.2	0.6
ABMP16	0.04	0.3	0.6	0.4	0.0003	0.1	0.7	0.6
ABM11	0.03	0.3	0.4	0.2	0.004	0.2	0.7	0.5

Table 3: The p -values for two different x ranges for the e⁻p and e⁺p data sets with generated events varied systematically by 1.8 % owing to the error in the luminosity measurement.

overall agreement with the observed data. For the e⁻p data set large differences are seen, with the HERAPDF2.0 [1] PDF set yielding the best results. The p -values from CT14 [10], MMHT2014 [9] and NNPDF2.3 [12] are quite small in this case, indicating that the observed data is in the tail of the expected distribution. The p -values from ABM [11] PDF sets are also shown and are found to be higher than other PDFs for e⁻p data set. The probability evaluation was performed separately for the data with $x < 0.6$ and $x \geq 0.6$. The results for the different x -ranges for both e⁺p and e⁻p data are given in Table 2. As can be seen in this table, there are significant differences observed in the two x ranges, particularly in the e⁻p data. Here, all data sets yield approximately the same probability for the larger x data whereas MMHT2014, CT14 and NNPDF2.3 give much worse probabilities for the smaller x data. For the e⁺p data, in contrast, MMHT2015, CT14 and NNPDF2.3 PDF sets yield the best probabilities.

It has been observed that the primary source of systematic uncertainty in the analysis is due to the uncertainty in the luminosity determination which is quoted as 1.8% [4]. The effect of this uncertainty is that it scales the number of generated events ($v_{i,k}$ values in Equation 2.2) systematically. Therefore, the events at generator level are scaled up and down by 1.8% and new p -values are obtained for the given expectations. The results from the study are summarized in Table 3. The effect of all other sources of statistical and systematic uncertainties are found to be negligible. We summarize and conclude here that large differences are observed in p -values from different PDFs. Despite the fact that the event numbers in ZEUS high- x data are small, this data set contains significant information on the behavior of the parton densities at the highest values of x .

References

- [1] ZEUS Coll., H. Abramowicz et al., Eur. Phys. J. C **75**, 580 (2015).
- [2] BCDMS Coll., A.C. Benvenuti et al., Phys. Lett. B **223**, 485 (1989).
- [3] L.W. Whitlow et al., Phys. Lett. B **282**, 475 (1992).
- [4] ZEUS Coll., H. Abramowicz et al., Phys. Rev. D **89**, 072007 (2014).
- [5] A. Kwiatkowski, H. Spiesberger and H.-J. Möhring, Comp. Phys. Comm. **69**, 155 (1992). Also in *Proc. Workshop Physics at HERA*, eds. W. Buchmüller and G. Ingelman, (DESY, Hamburg, 1991).
- [6] HERA-Fitter developersâ team: S. Alekhin et al. HERA-Fitter. Eur. Phys. J C **75**, 304 (2015).
- [7] R. Aggarwal, A. Caldwell, Eur. Phys. J. Plus **127** 1-8 (2012).
- [8] F. Beaujean et al., Phys. Rev. D **83**, 012004 (2011).
- [9] L. Harland-Lang, A. D. Martin, P. Motylinski and R. Thorne (2014), [arXiv:1412.3989].
- [10] M. Guzzi et al., Phys. Rev. D **82**, 074024 (2010), [arXiv:1007.2241].
- [11] S. Alekhin et al., Phys. Rev. D **86**, 054009 (2012), [arXiv:1202.2281].
[NNPDF Collaboration], Nucl. Phys. B **809**, 1 (2009), [arXiv:0808.1231]
- [12] R. D. Ball et al. [NNPDF Collaboration], JHEP **1004**, 040 (2015), [arXiv:1410.8849v2]