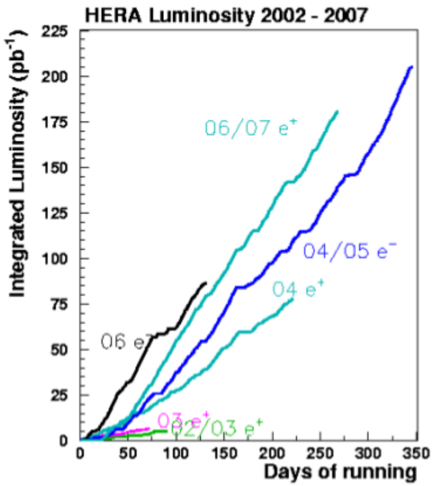
Study of proton parton distribution functions at high x

(Phys. Rev. D 101 (2020) 112009)

Ritu Aggarwal
SPPU, Pune - India

(on behalf of the ZEUS Collaboration)

Overview



HERA upgrade

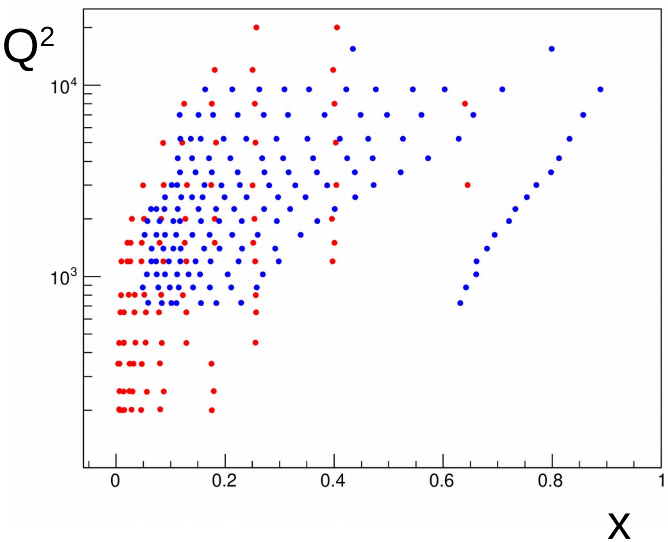
- Luminosity upgrade.
  - HERA I : 120 pb^{-1}
  - HERA II : 360 pb^{-1}

HERA-II

- *e* beam : 27.5 GeV
- *p* beam : 920 GeV
- Centre of mass E : 318 GeV
- H1 & ZEUS : General Purpose Detectors
- Data Taking : 1992-2007

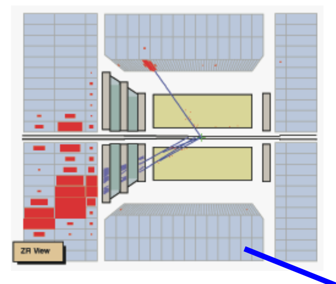
- HERA-II upgrade (2002-2007):
  Increased Luminosity
  Polarized Lepton Beam
  Improved Tracking system (MVD, STT)

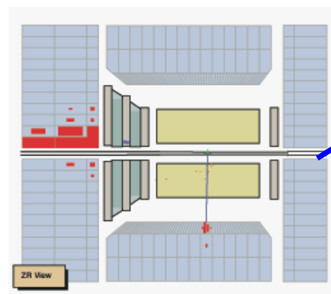
Motivation of studying published high-x data



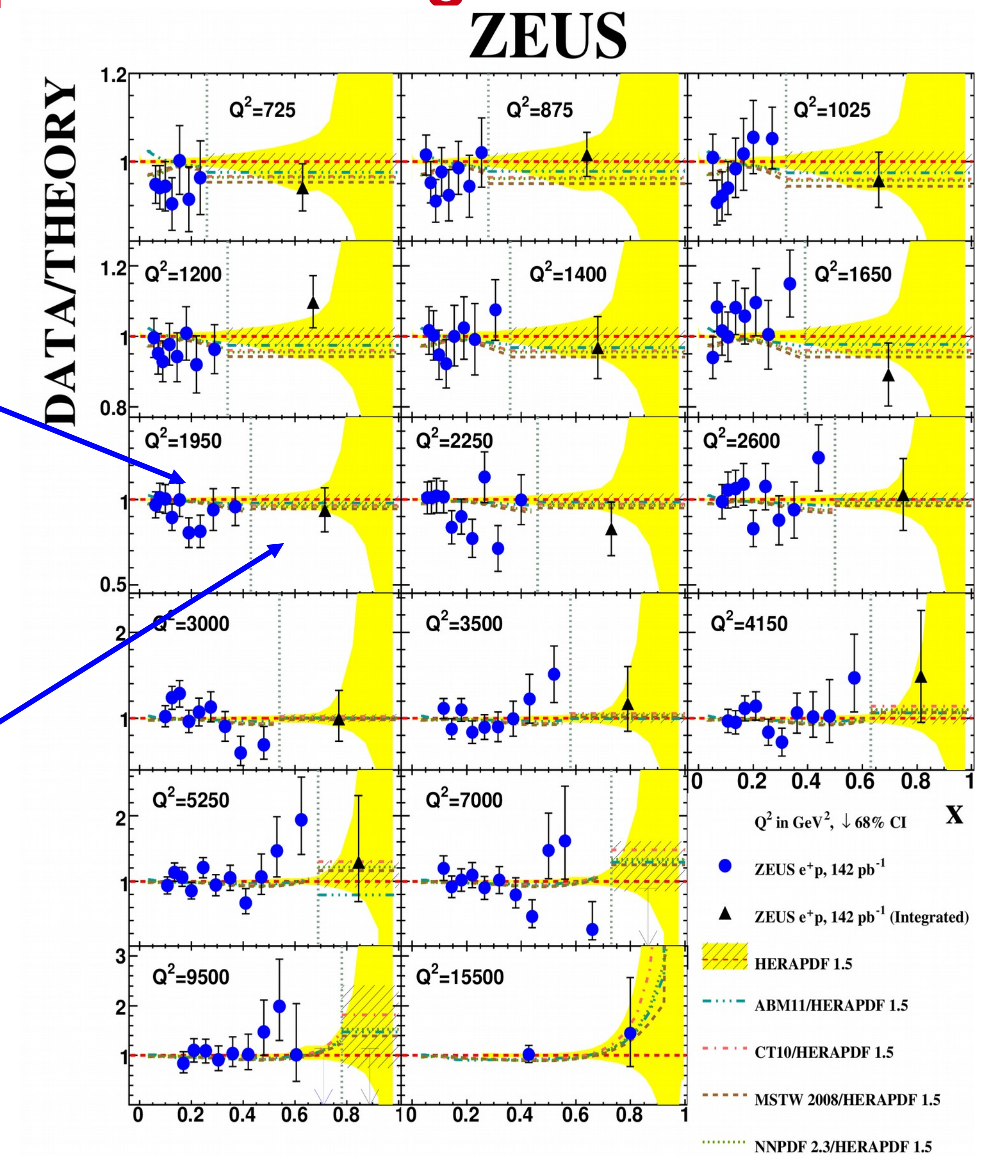
● high-Q^2 data

● high-x data





Integrated bins
(cross sections
Integrated in x)



At present x upto 0.65 ZEUS DIS data is included in PDF fits

Note the uncertainity bands above x ~ 0.65, can high-x data impact here?

ZEUS Collaboration; H. Abramowicz et al. Measurement of Neutral Current $e^{\pm}p$ Cross-Sections at High Bjorken x with the ZEUS Detector Phys. Rev. D 89 (2014) 072007

Technique to use high-x data

1) Some of the bins have low number of events / few have zero, poisson errors quoted.

2) Ofcourse it has a subset of data (high-Q^2 ZEUS data) already included in fits, but high-x data has more to say.

Transfer Matrix for the detector is developed using which number of events reconstructed in data can be predicted from any PDF ($\nu_{j,k}$) as below.

$$\nu_{j,k} = \mathcal{L} \int_{(\Delta x, \Delta Q^2)_j} \left[\int A(x_{\rm rec}, Q^2_{\rm rec} | x, Q^2) \times \frac{d^2\sigma(x, Q^2 | {\rm PDF}_k)}{dx dQ^2} dx dQ^2 \right] dx_{\rm rec} dQ^2_{\rm rec}.$$

$$\longrightarrow \quad \nu_{j,k} \approx \sum_i A_{ji} \lambda_{i,k}$$

$$T_{ji} = \frac{\sum_{m=1}^{M_i} \omega_m I(m \in j)}{\sum_{m=1}^{M_i} \omega_m^{\rm MC}}$$

$$\boldsymbol{\nu = T\, R\, \lambda}$$

$$R_{ii} = \frac{\sum_{m=1}^{M_i} \omega_m^{\rm MC}}{\mathcal{L}^{\rm MC} \sigma_{i,{\rm CTEQ5D}}}$$

L : data luminosity

R_{ii} : Radiative corrections (calculated using HERACLES)

$\lambda_{i,k}$: born level cross sections in i^{th} bin for k^{th} PDF

$\boldsymbol{T_{ji}}$: Transfer Matrix, has all detector and analysis effects
(probability of an event reconstructed in j^{th} bin to come from i^{th} true bin)

Transfer Matrix : Probability of an event reconstructed in j^{th} bin to come from i^{th} true bin

Tracing back the path of MC reconstructed events in the generated x-Q^2 phase space

$$T_{ji} = \frac{\sum_{n=1}^{M_i} \omega_n I(n \in j)}{\sum_{n=1}^{M_i} \omega_n^{MC}}$$

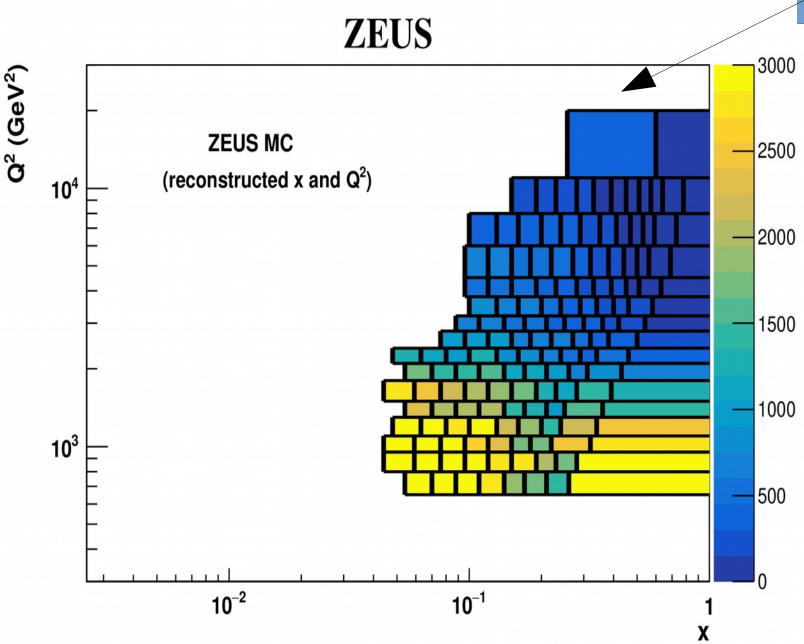
T_{ij} = probability of an event reconstructed in j^{th} bin to come from i^{th} bin

ω_m = MC weights given to m^{th} event in bin i

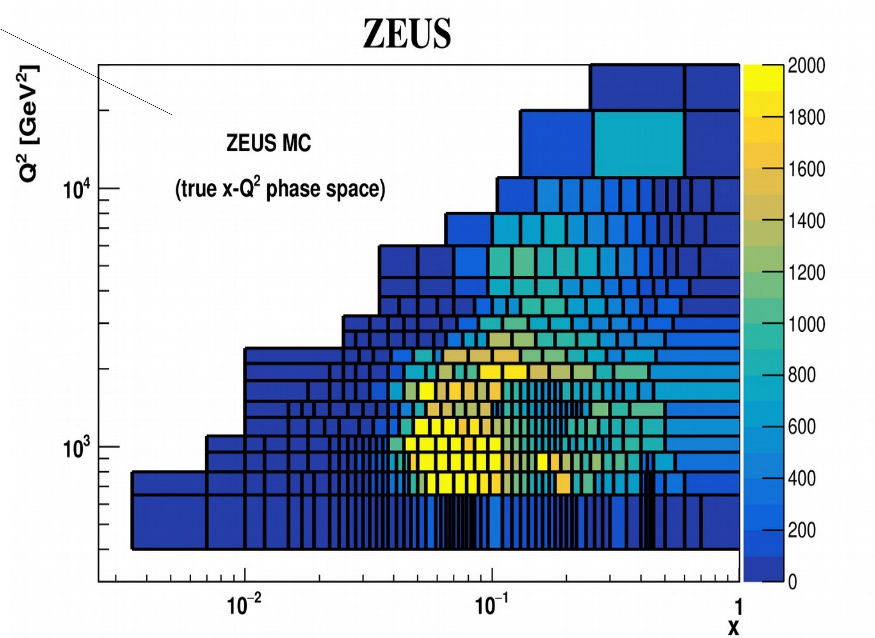
I = 1 if m^{th} event is reconstructed in bin j, else = 0

M_i = total events generated in i^{th} bin

$$\nu = T\, \underline{R}\, \lambda$$



Reconstructed MC events in Cross section bins



Generated distribution of these events

Comparison of Different PDFs

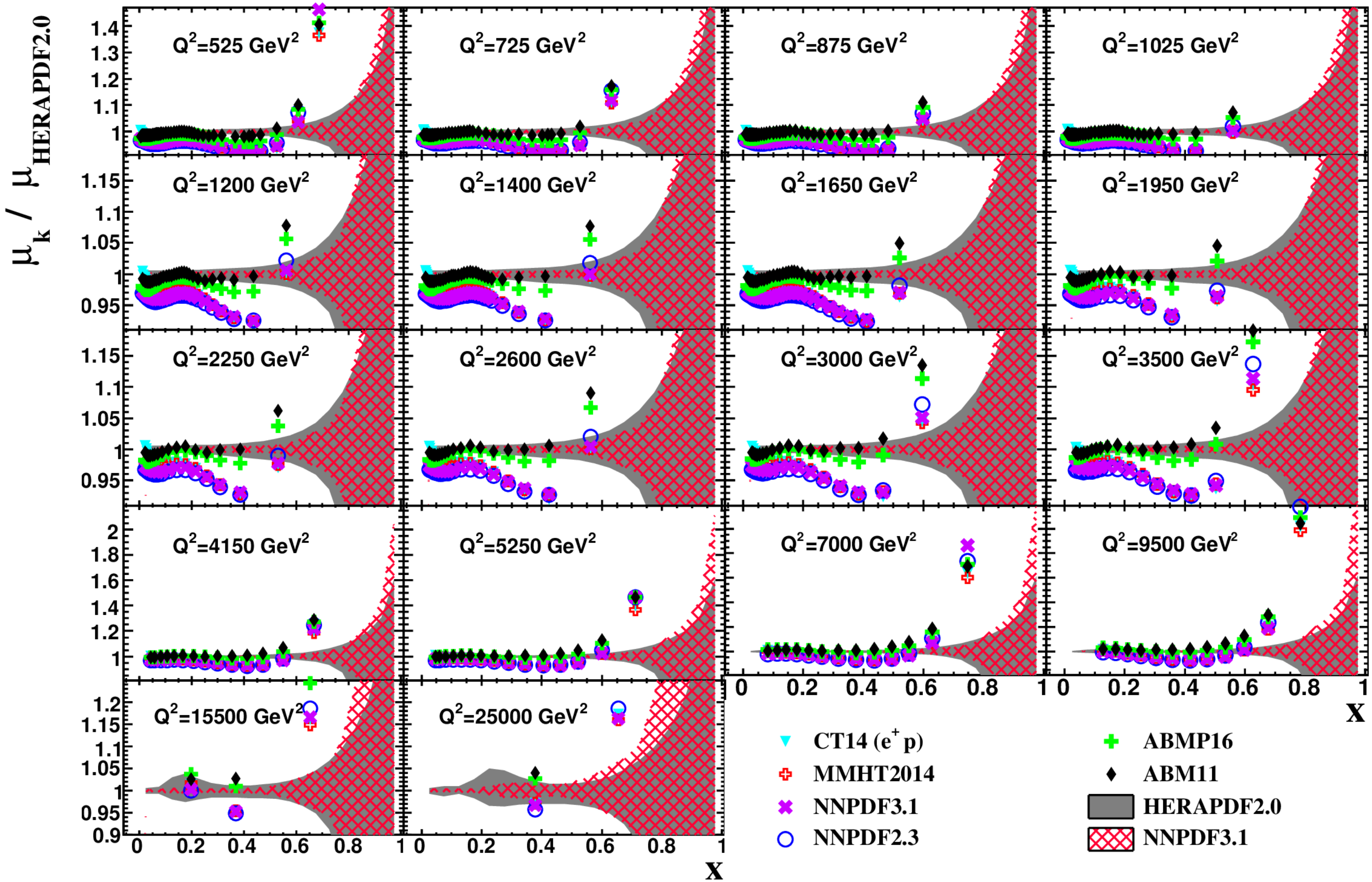
1) Comparison at generator level ($\mu = R\lambda$) from different PDFs.

2) Comparison at reconstructed level from different PDFs : Convolute μ with Transfer Matrix to get a prediction of number of events in the cross section b bins (ν) from different PDFs

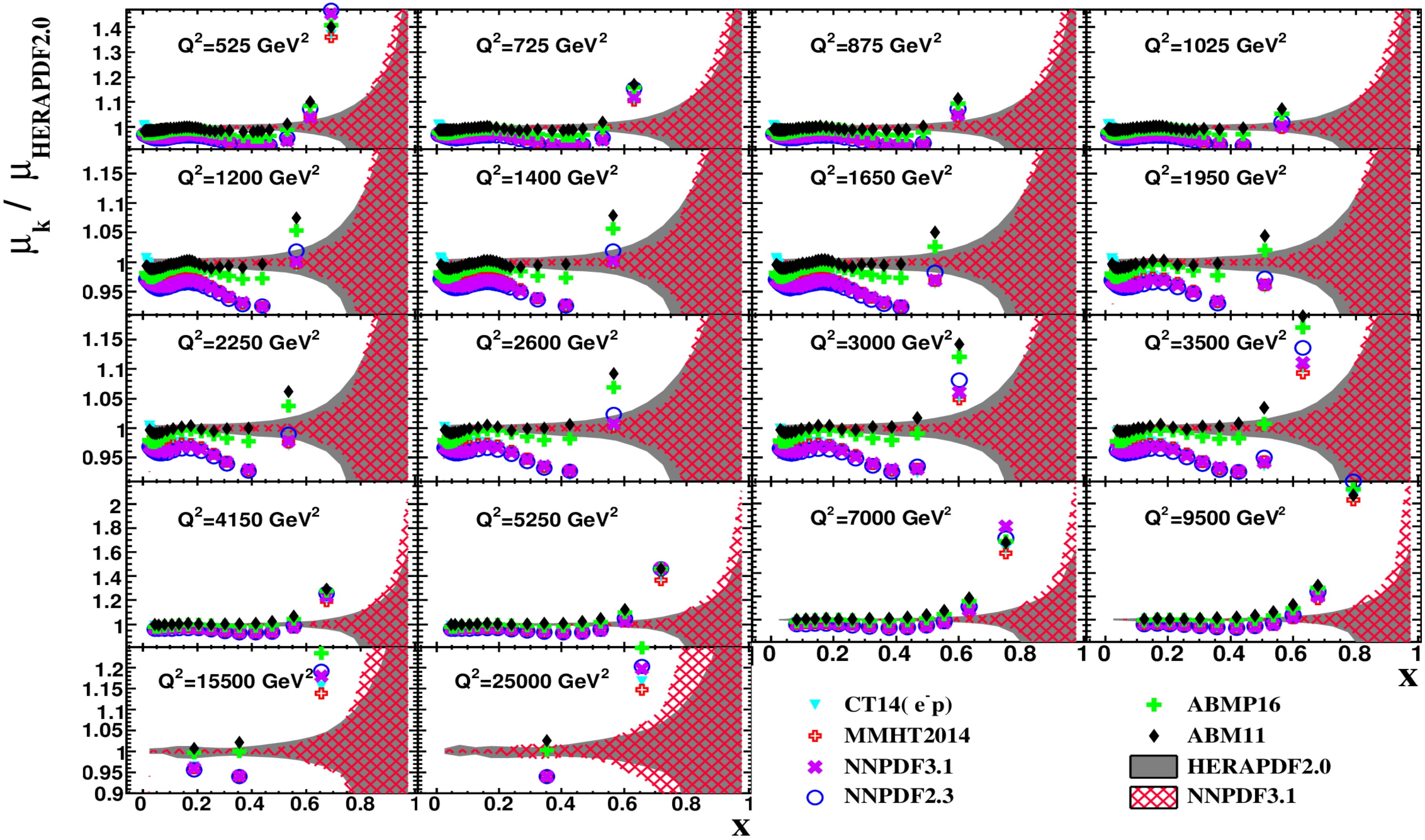
i.e. $\nu = T\mu$

-

Ratio of generator level cross sections in different PDFs to HERAPDF2.0NNLO (e+p)



Average ratio of Born level cross sections in different PDFs to HERAPDF2.0NNLO (e-p)



Comparison of Different PDFs

- ν (= Tμ) from different PDFs is compared to observed events from data (n)

Poisson statistics is used to probe how well given PDF is defining the data.

$$P(D|\mathrm{PDF_k}) = \prod_j \frac{e^{-\nu_{j,k}} \nu_{j,k}^{n_j}}{n_j!}$$

- Bayes factor, $\Delta\chi^2$ and p-value is determined for different PDFs

$$\Delta\chi^2_{k,l} = -2 \ln \frac{P(D|\mathrm{PDF}_k)}{P(D|\mathrm{PDF}_l)}$$

- Comparison of p-values in high-x and lower-x range is shown for different PDFs

Probability of explaining high-x data from different PDFs

PDF	e^-p p-value	P1/P2	$\Delta\chi^2$	e^+p p-value	P1/P2	$\Delta\chi^2$
HERAPDF2.0	2.8×10^{-2}	1.0	0.0	0.35	1.0	0.0
CT14	3.2×10^{-3}	7.6×10^{-3}	9.8	0.82	$5.9 \times 10^{+5}$	−27
MMHT2014	2.3×10^{-3}	2.1×10^{-3}	12	0.82	$4.7 \times 10^{+5}$	−26
NNPDF3.1	3.9×10^{-4}	3.2×10^{-6}	25	0.73	$9.0 \times 10^{+4}$	−23
NNPDF2.3	1.3×10^{-4}	2.3×10^{-7}	31	0.70	$4.2 \times 10^{+4}$	−21
ABMP16	2.6×10^{-2}	9.0×10^{-1}	0.21	0.64	$6.1 \times 10^{+2}$	−13
ABM11	3.3×10^{-2}	7.2×10^{-1}	0.67	0.45	2.8	−2.1

Conclusions :

- **p-values from MMHT2014, CT14nlo, NNPDF higher than HERAPDF2.0 for e^+p, much worse for e^-p**

	e^-p				e^+p			
	Lower x		Higher x		Lower x		Higher x	
PDF	P1/P2	$\Delta\chi^2$	P1/P2	$\Delta\chi^2$	P1/P2	$\Delta\chi^2$	P1/P2	$\Delta\chi^2$
HERAPDF2.0	1.0	0.0	1.0	0.0	1.0	0.0	1.0	0.0
CT14	6.0×10^{-1}	1.0	1.3×10^{-2}	8.7	$1.6 \times 10^{+3}$	−15	$3.6 \times 10^{+2}$	−12
MMHT2014	7.1×10^{-2}	5.3	2.9×10^{-2}	7.1	$1.3 \times 10^{+3}$	−14	$3.7 \times 10^{+2}$	−12
NNPDF3.1	9.1×10^{-5}	19	3.5×10^{-2}	6.7	$2.5 \times 10^{+2}$	−11	$3.6 \times 10^{+2}$	−12
NNPDF2.3	8.0×10^{-6}	23	2.9×10^{-2}	7.1	$1.2 \times 10^{+2}$	−9.5	$3.7 \times 10^{+2}$	−12
ABMP16	2.3×10^{-1}	3.0	4.0	−2.7	$4.8 \times 10^{+1}$	−7.8	$1.3 \times 10^{+1}$	−5.1
ABM11	2.3×10^{-1}	3.0	3.2	−2.3	4.2	−2.9	6.7×10^{-1}	0.8

F. Beaujean et al, p-values for model evaluation, Phys. Rev. D 83, 012004 (2011). arXiv:1011.1674

Statistical and systematic uncertainties

Type of Systematic Uncertainties :

1) Affecting the predictions at generator level (μ values)
2) Affecting the Transfer Matrix (T)

Type I :
1) Luminosity uncertainty scaling μ values
2) Knowledge of radiative effects
3) PDF uncertainities

Type II :
1) MC statistical fluctuations (uncorrelated uncertainty)
2) All correlated and uncorrelated systematic uncertainties as in high-x paper
3) Choice of PDF for building T

Conclusions – NC DIS high-x analysis at ZEUS

Technique of building Transfer Matrix for high-x ZEUS data Shown.

--Transfer Matrix can be used to predict number of events in the given cross section bins in MC.

-- Transfer Matrix can be used to compare number of events predicted by different PDFs.

Bayes factor, p-values from different PDFs calculated and shown on the basis of their explanation to the high-x data using Transfer Matrix.

-- Differences are seen between different PDFs
-- Differences are also there for e-p and e+p data sets and the higher and lower x ranges.

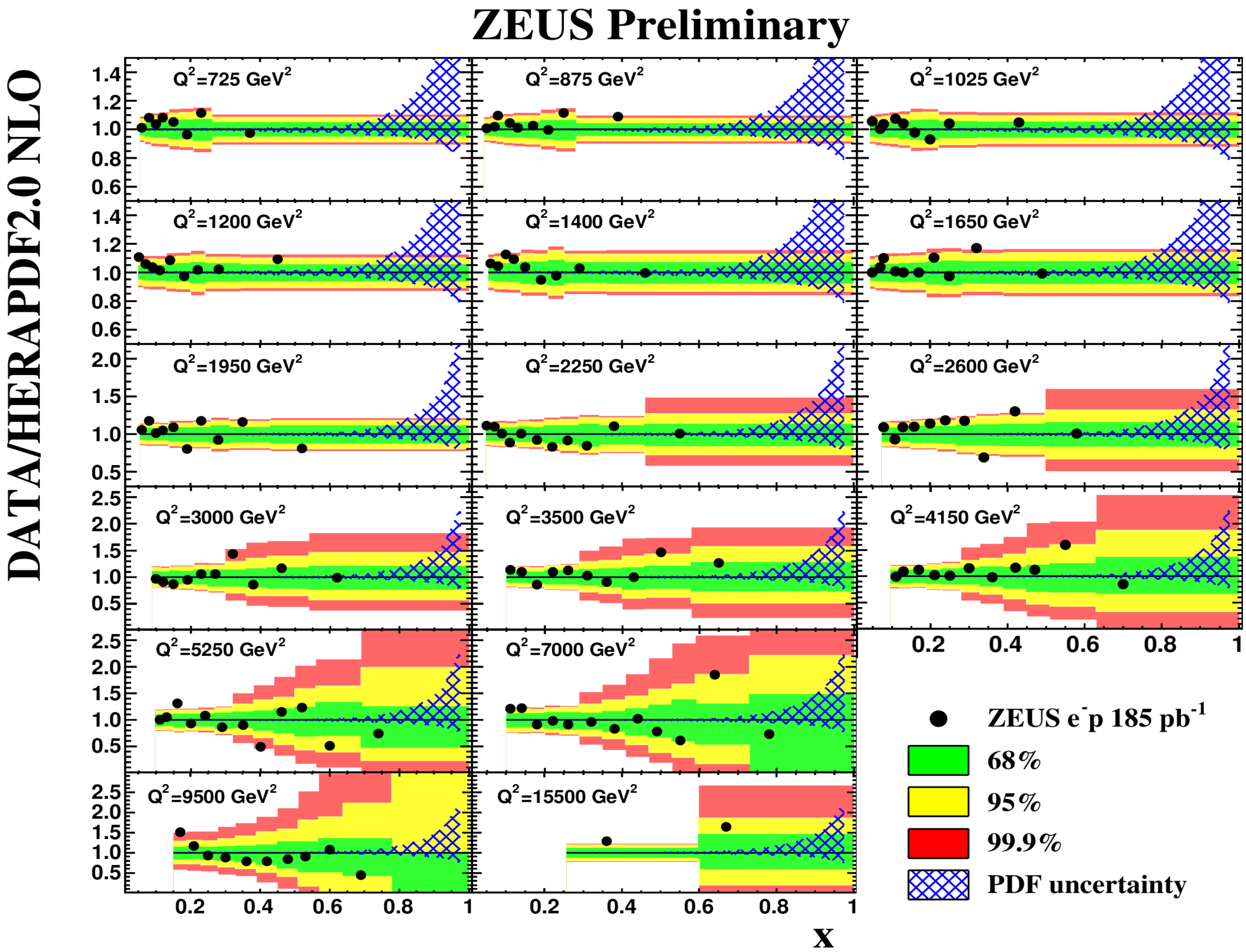
Dominant systematic uncertainty is due to uncertainty in Luminosity measurement.

Direction on how to include high-x data in PDF fits provided.

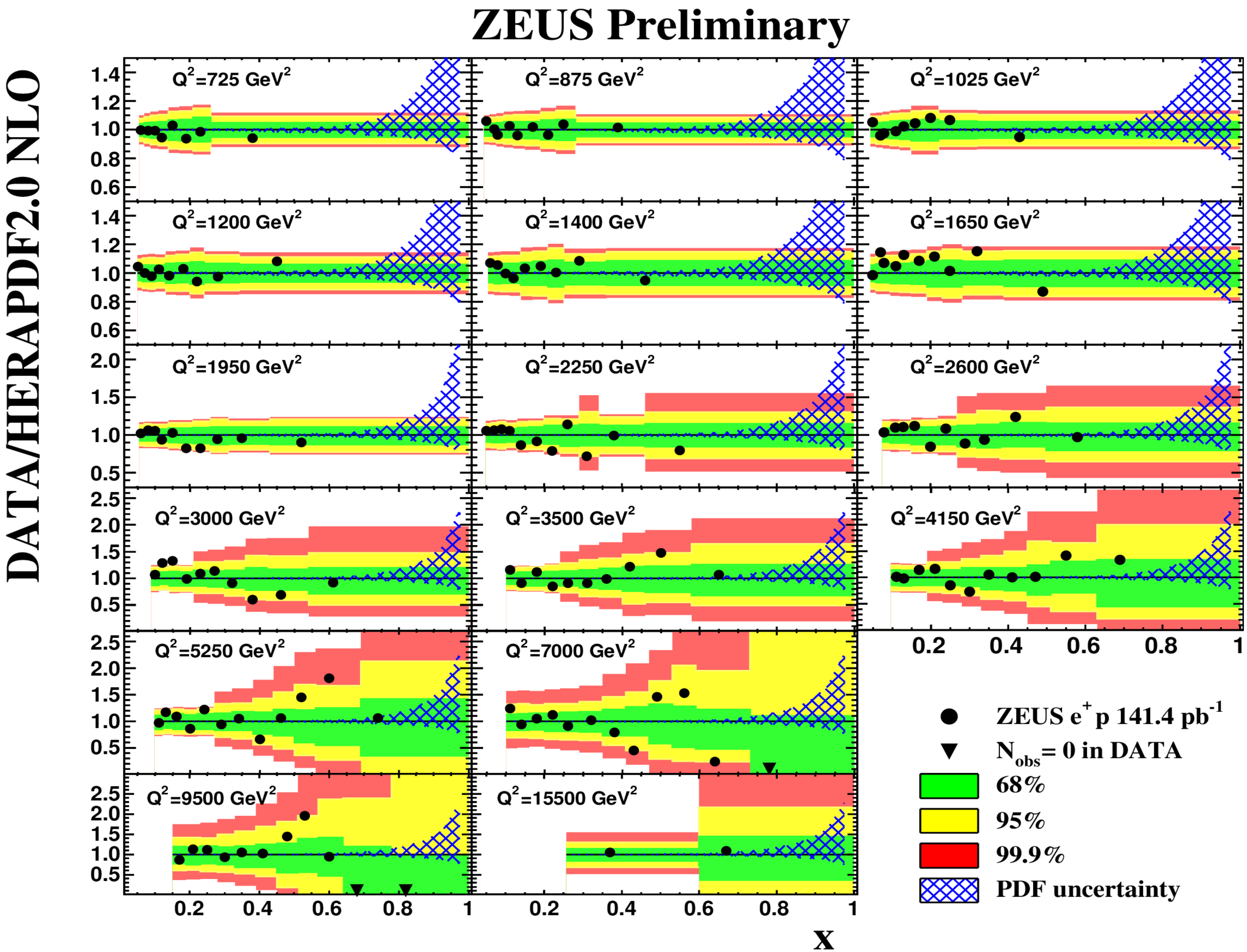
Thank you!

Back up

Ratio of No. of events in data to HERAPDF2.0 NLO and 1,2,3 sigma bands from Poisson Statistics



Ratio of No. of events in data to HERAPDF2.0 NLO and 1,2,3 sigma bands from Poisson Statistics



R. Aggarwal, A. Caldwell, Error bars for distributions of numbers of events. Eur. Phys. J. Plus 2012, 127, 1–8.

Nomalization Error : Vary generated events by 1.8 % up and down and calculate new p-value

	e^-p				e^+p			
	Lower x		Higher x		Lower x		Higher x	
PDF	P1/P2	$\Delta\chi^2$	P1/P2	$\Delta\chi^2$	P1/P2	$\Delta\chi^2$	P1/P2	$\Delta\chi^2$
			$+1.8\%$					
HERAPDF2.0	1.7×10^{-3}	13	0.08	5.0	8.6×10^{-5}	19	0.01	9.8
CT14	28	−6.7	16	−5.5	5.0×10^{-2}	6.0	0.37	2.0
MMHT2014	$1.1\times10^{+2}$	−9.5	13	−5.1	0.11	4.4	0.31	2.3
NNPDF3.1	$4.9\times10^{+3}$	−17	15	−5.4	2.2	−1.5	0.36	2.0
NNPDF2.3	$1.8\times10^{+4}$	−20	18	−5.8	5.9	−3.6	0.41	1.8
ABMP16	0.37	2.0	0.30	2.4	2.6×10^{-3}	12	0.02	7.8
ABM11	9.6×10^{-3}	9.3	0.07	5.3	2.9×10^{-4}	16	0.01	10
			-1.8%					
HERAPDF2.0	0.69	0.74	1.7	−1.1	93	−9.1	30	−6.8
CT14	4.9×10^{-5}	20	9.3×10^{-3}	9.3	0.18	3.4	0.66	0.83
MMHT2014	1.2×10^{-5}	23	1.1×10^{-2}	8.9	8.0×10^{-2}	5.0	0.77	0.51
NNPDF3.1	3.1×10^{-7}	30	9.8×10^{-3}	9.3	4.4×10^{-3}	11	0.67	0.79
NNPDF2.3	8.5×10^{-8}	33	8.2×10^{-3}	9.6	1.6×10^{-3}	13	0.59	1.1
ABMP16	3.5×10^{-3}	11	0.46	1.5	3.3	−2.4	12	−4.9
ABM11	0.12	4.2	1.9	−1.3	28	−6.6	33	−7.0

Conclusions : Dominant systematics : due to error in normalization of data quoted as 1.8 %

- **p-values from different PDFs change differently**
- **Similar behavior as when using only statistical fluctuations.**

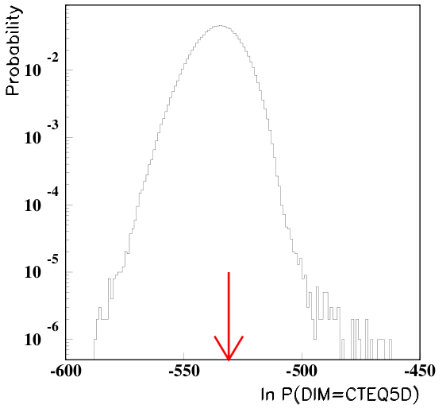
TABLE VI. List of $e^{\pm}p$ combined data points [9] to be removed and the ZEUS high-x data points [8] to be added. See Sec. VI for details.

HERA combined points to be removed	
x	Q^2 (GeV^2)
0.4	1200 1500 2000 3000 20000
0.6	3000 8000

ZEUS high-x data to be added	
x	Q^2 (GeV^2)
Integrated	725 875 1025 1200 1400 1650 1950 2250 2600 3000 3500 4150 5250 7000 9500
0.37	1950
0.4	2250
0.35, 0.44	2600
0.39, 0.48	3000
0.57	4150
0.53, 0.62	5250
0.56, 0.66	7000
0.54, 0.61, 0.71	9500
0.8	15500

P-value determination

Total probability for each PDF :
$$P(D|M_k) = \prod_j \frac{e^{-\nu_{j,k}} \nu_{j,k}^{n_j}}{n_j!}$$



P-value is calculated by integrating out the probability from the left edge till red for the given PDF