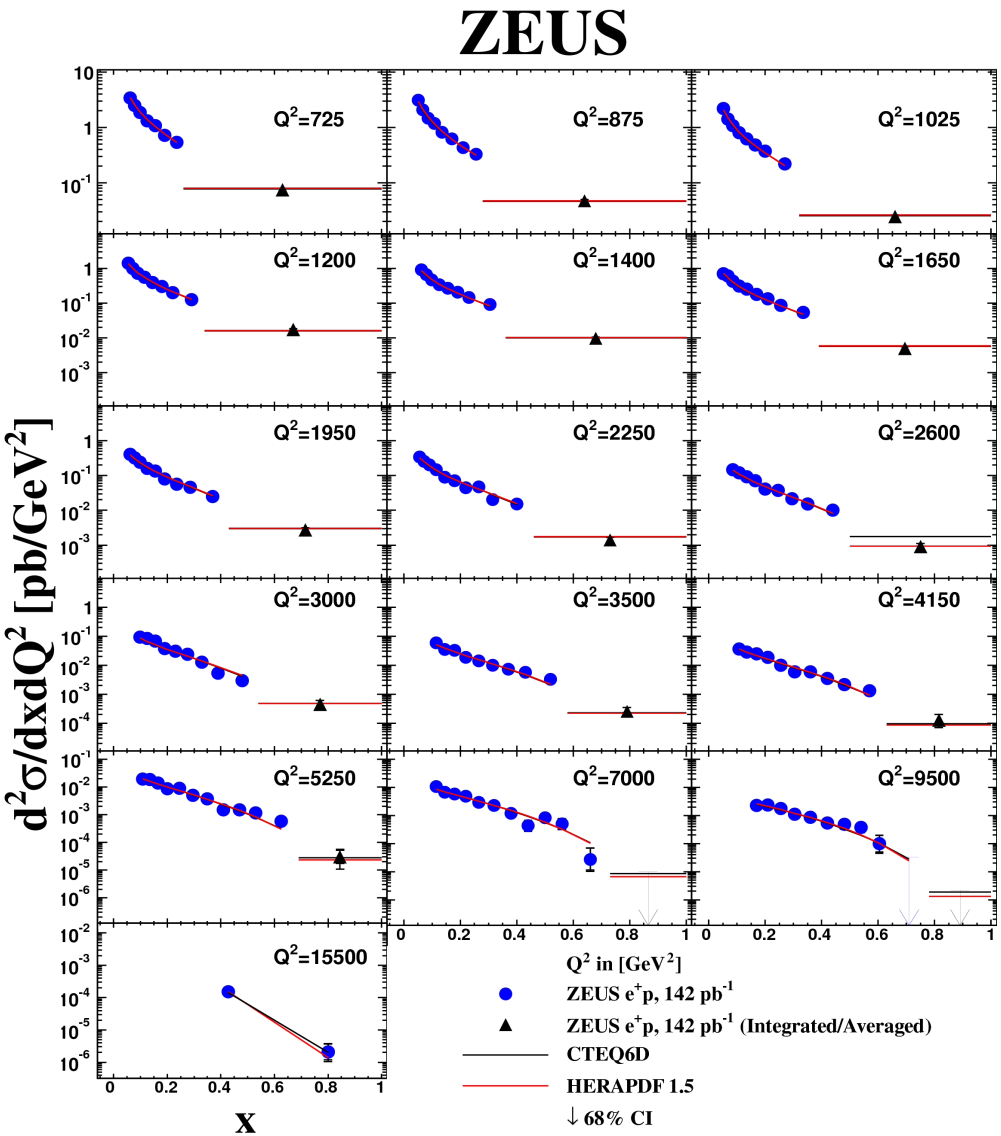
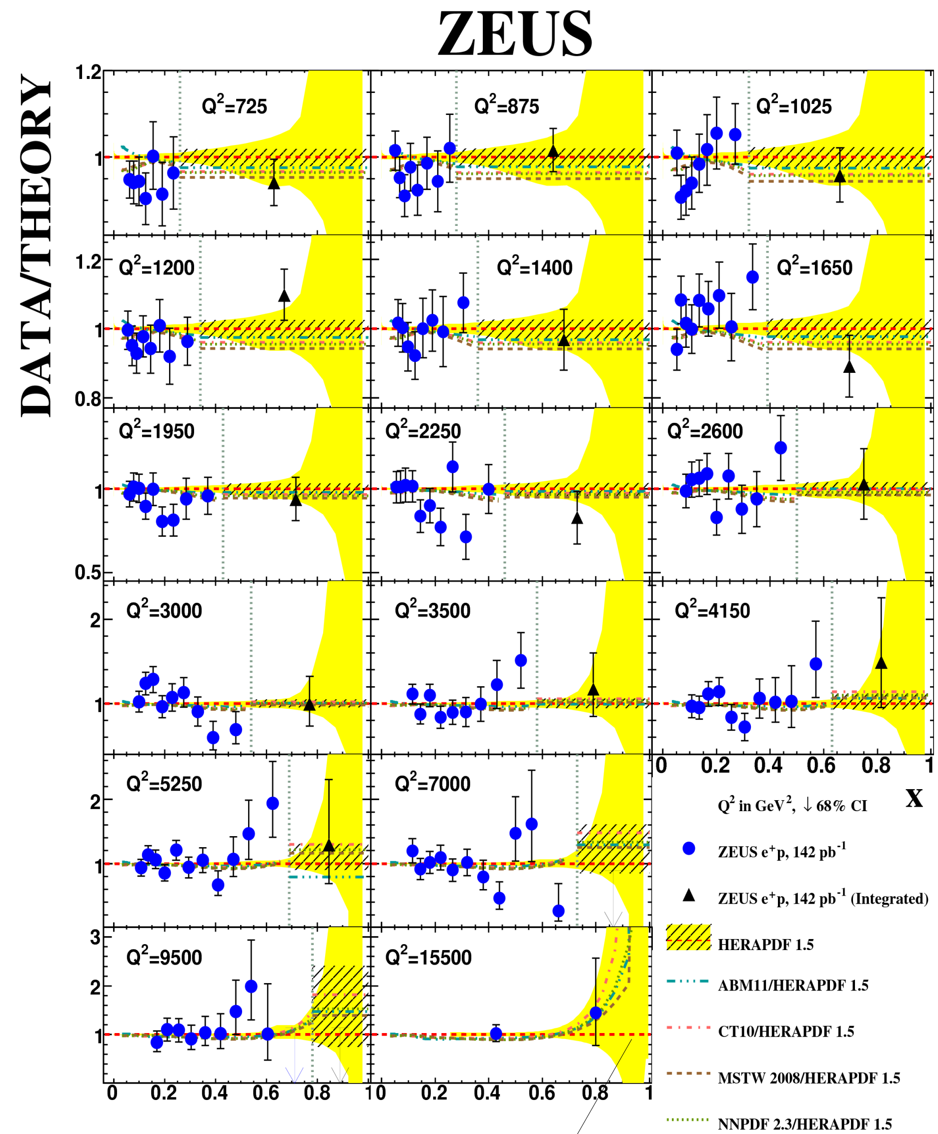
Transfer matrix for high-x MC
and
Study of impact of high-x data on parton distribution functions

Ritu Aggarwal, Allen Caldwell

Motivation of studying published high-x data



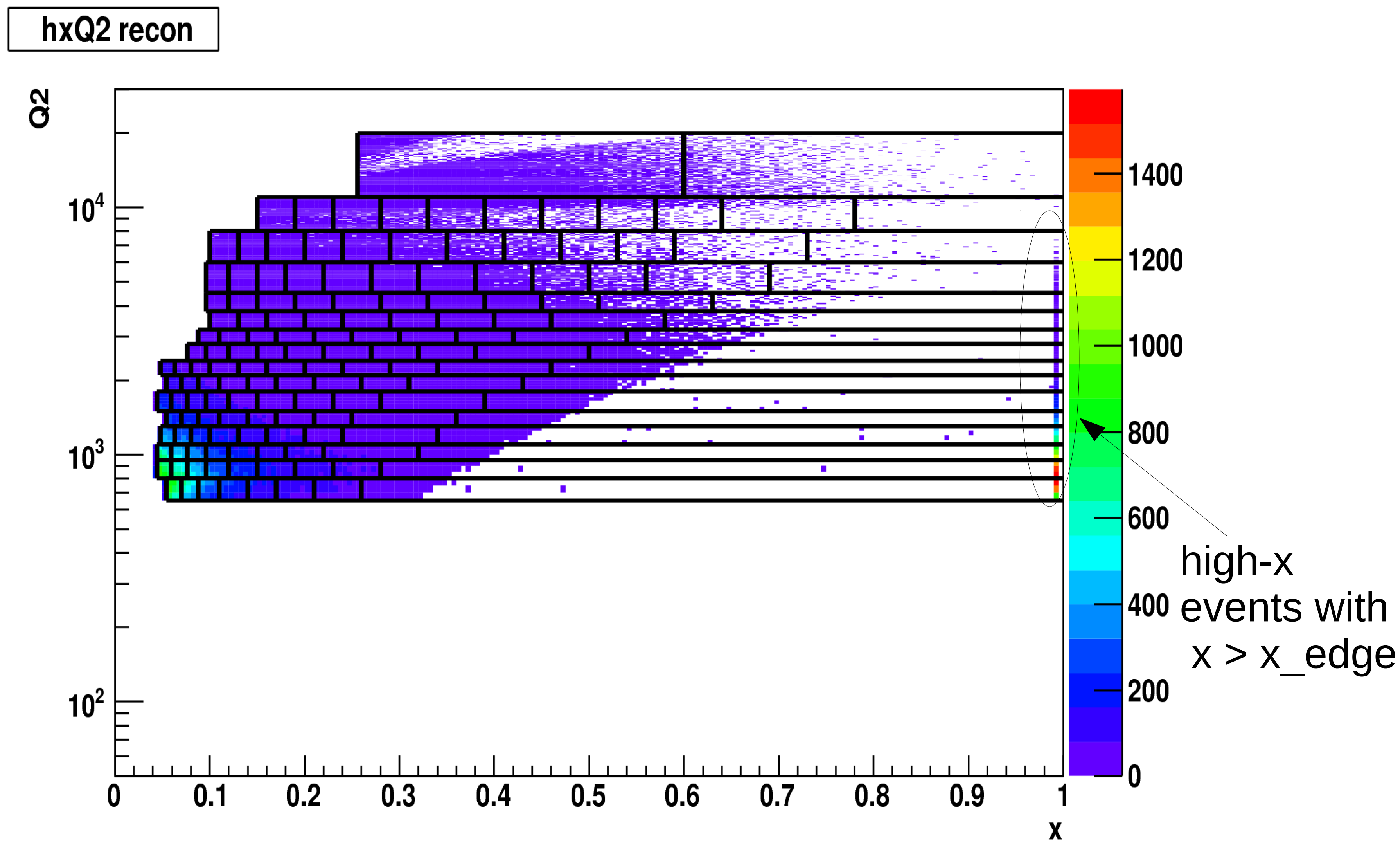


At present x upto 0.65 ZEUS data is included in PDF fits

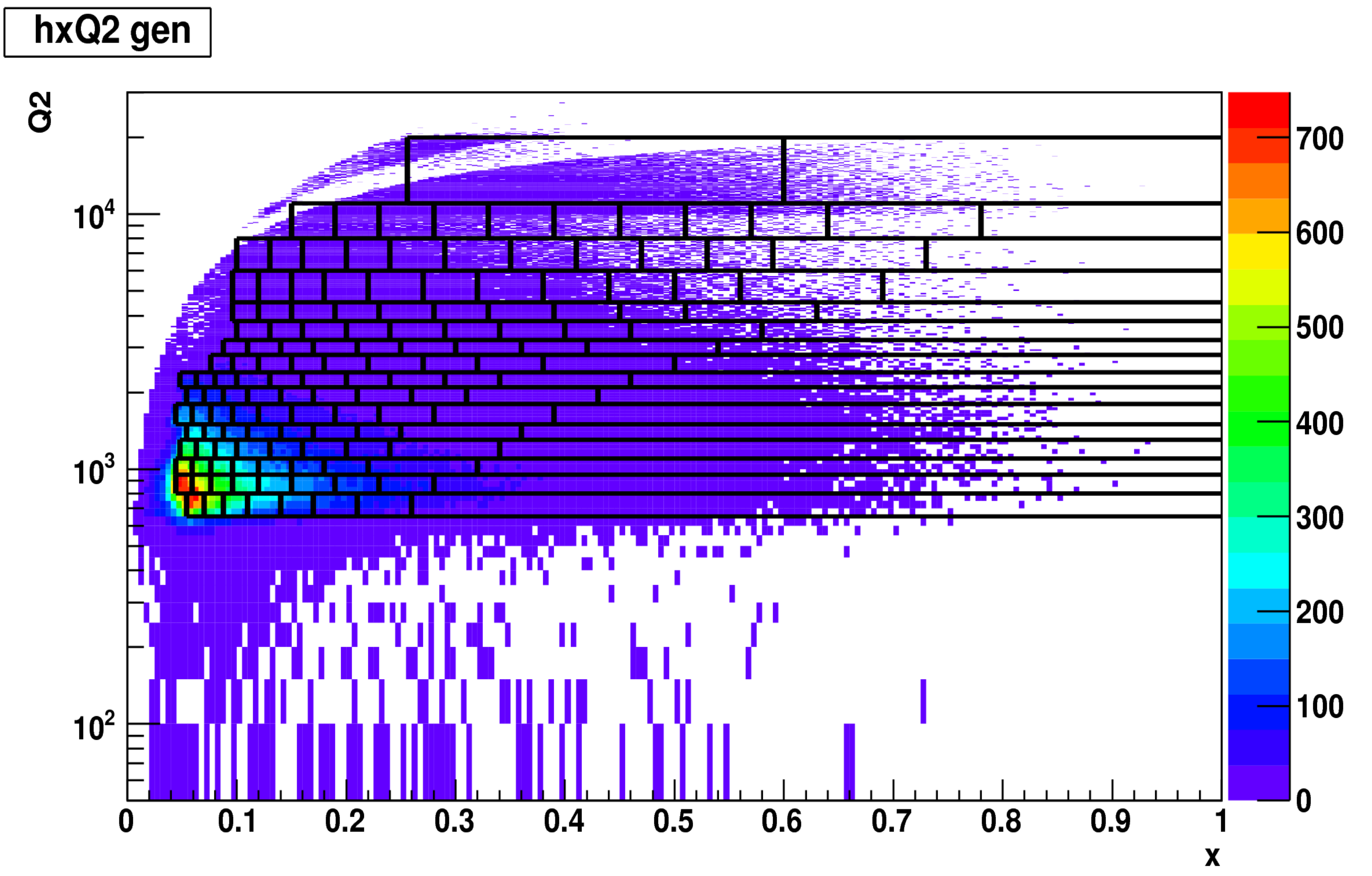
Note the uncertainity bands above x ~ 0.65, can high-x data impact here

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

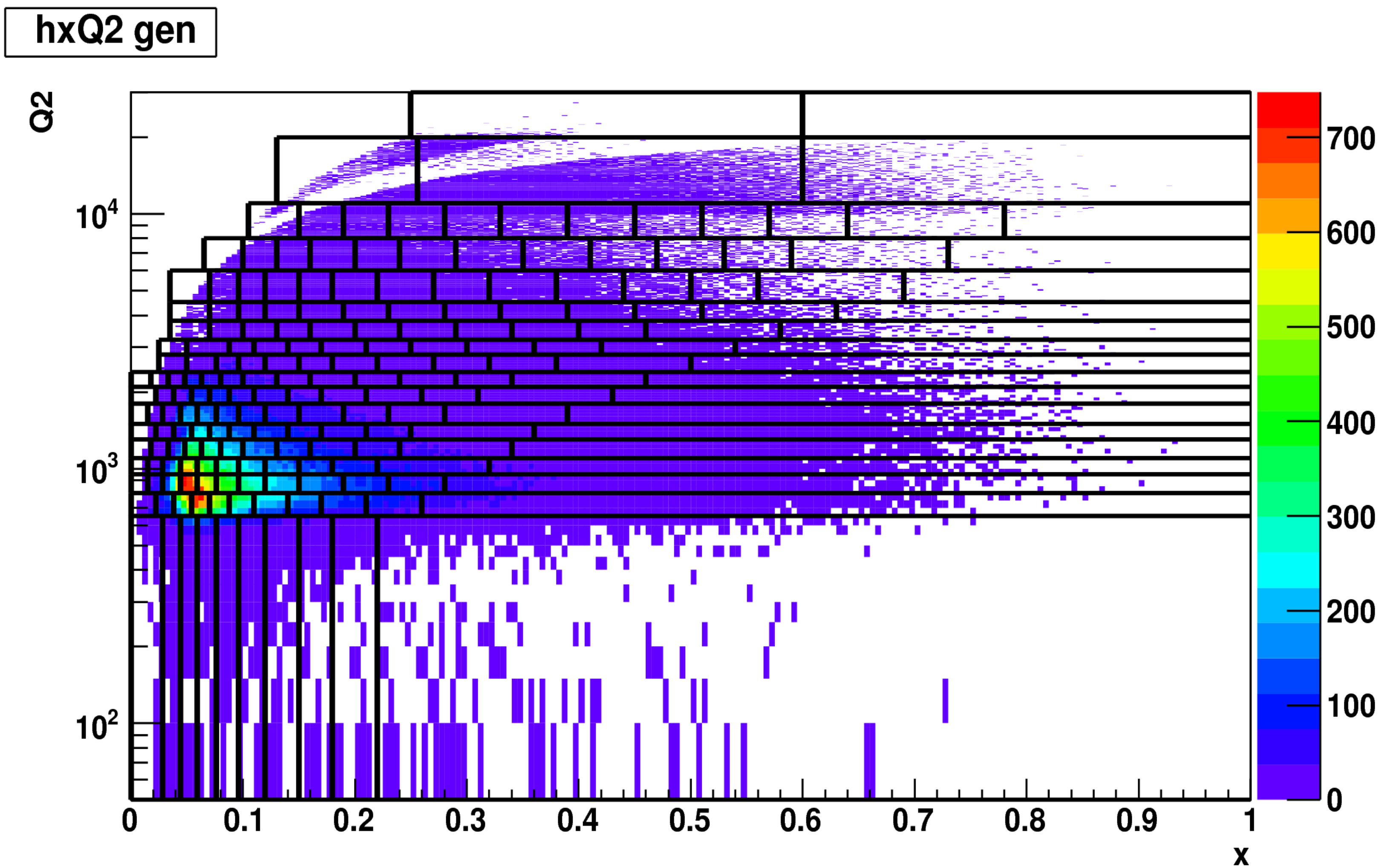
Reconstructed MC events in Xsection Binning (total 153 bins)



True x-Q2 distribution of events in Xsection Binning



Generated x-Q2 distribution in extended binning (total 203 bins)



Definition of Transfer matrix

Each element in Transfer Matrix is represented as

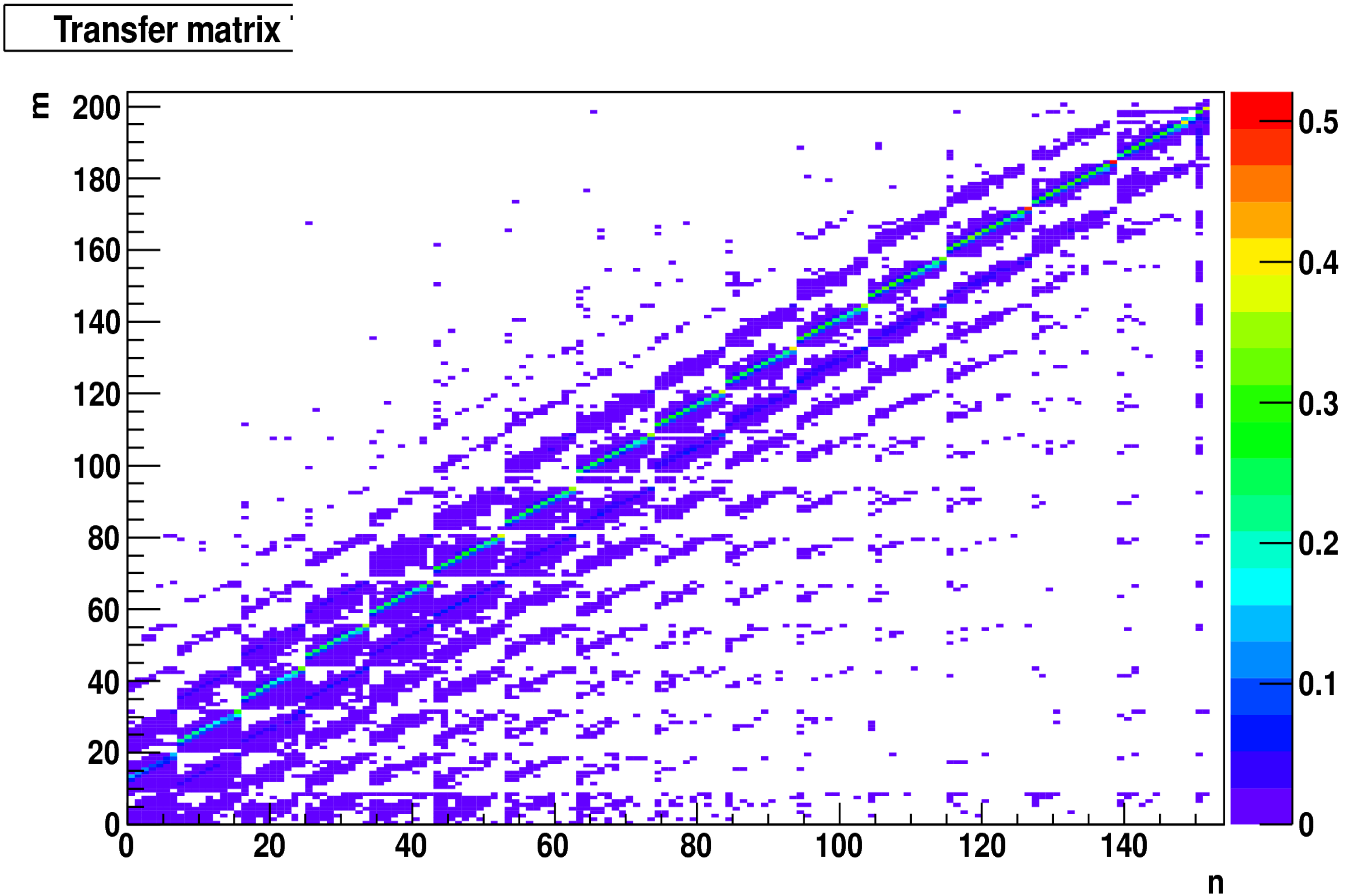
$$a_{ij} = \frac{\sum_{k=1}^{M_i} \omega_k I(k \in j)}{\sum_{k=1}^{M_i} \omega_k}$$

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

ω_k = weight given to k^{th} event in bin i

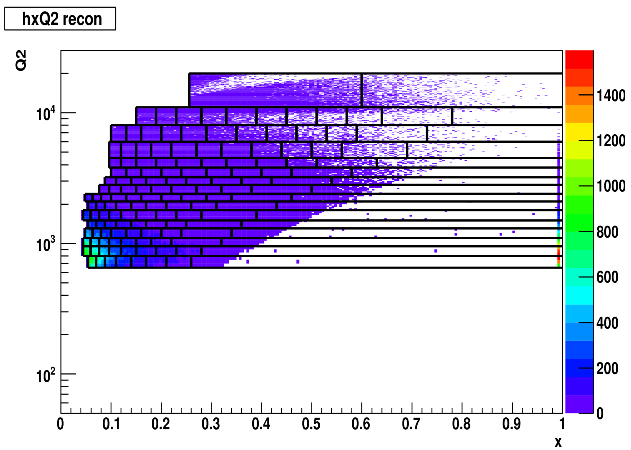
I = 1 if k^{th} event is reconstructed in bin j, else = 0

Transfer Matrix

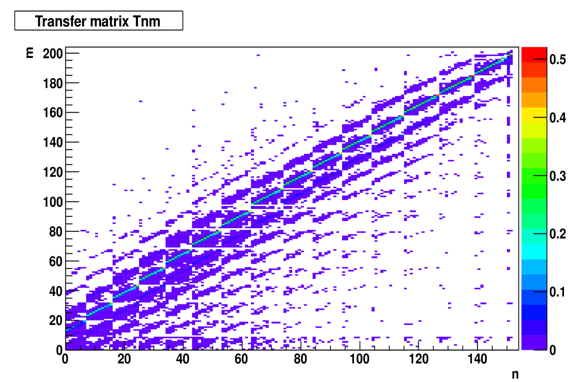


Using Transfer matrix to predict no. of events reconstructed in a given cross section bin

N = A M



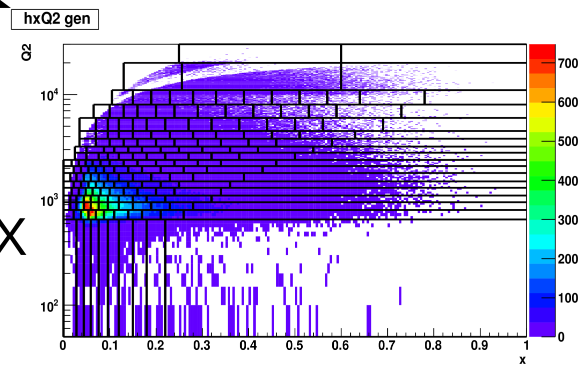
Predicted x-Q2 events in Cross section binning
(153 elements in N Vector = number of cross section bins)



Transfer Matrix

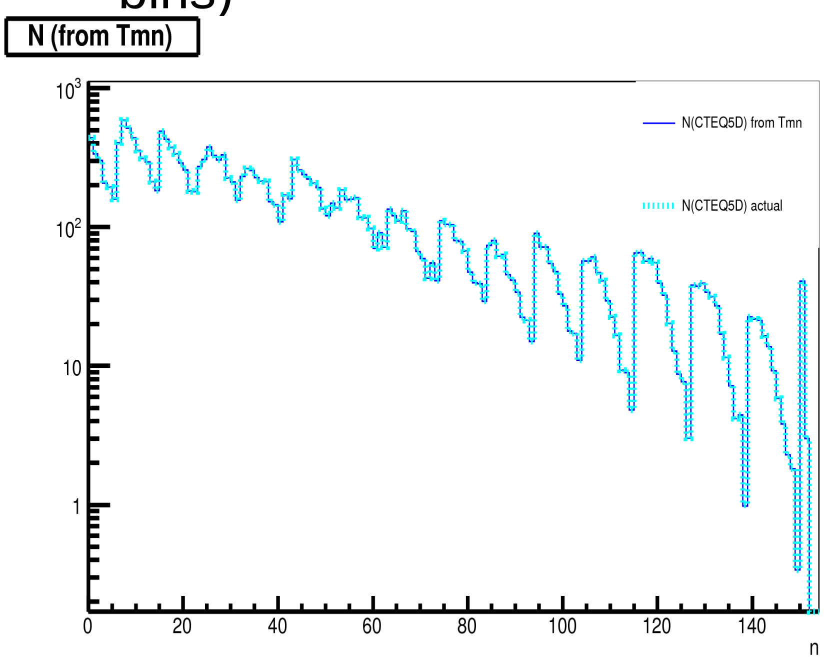
(153 X 203 elements)

X



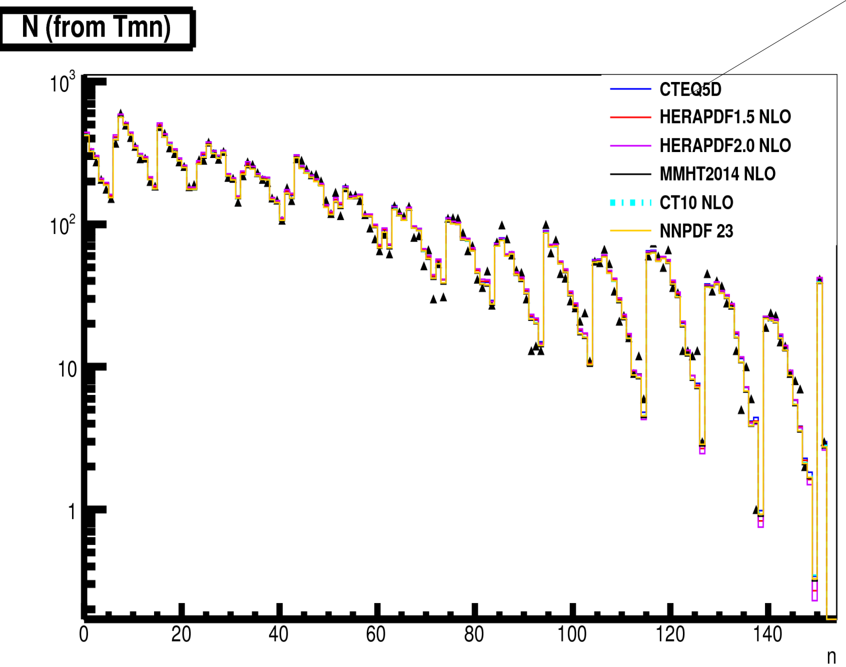
Generated x-Q2 events in Extended binning

(203 elements in M Vector = number of generated bins)



Using Transfer matrix to predict no. of events reconstructed in a given cross section bin

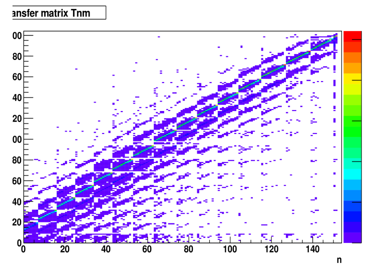
N = AM



Predicted x-Q2 events in Cross section binning

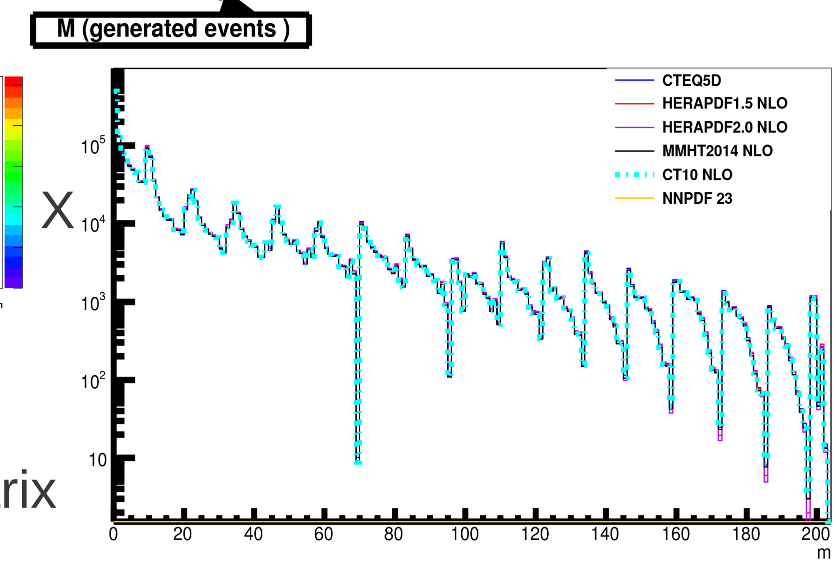
(for different PDFs)

=



Transfer Matrix

X



Generated x-Q2 events in Extended binning

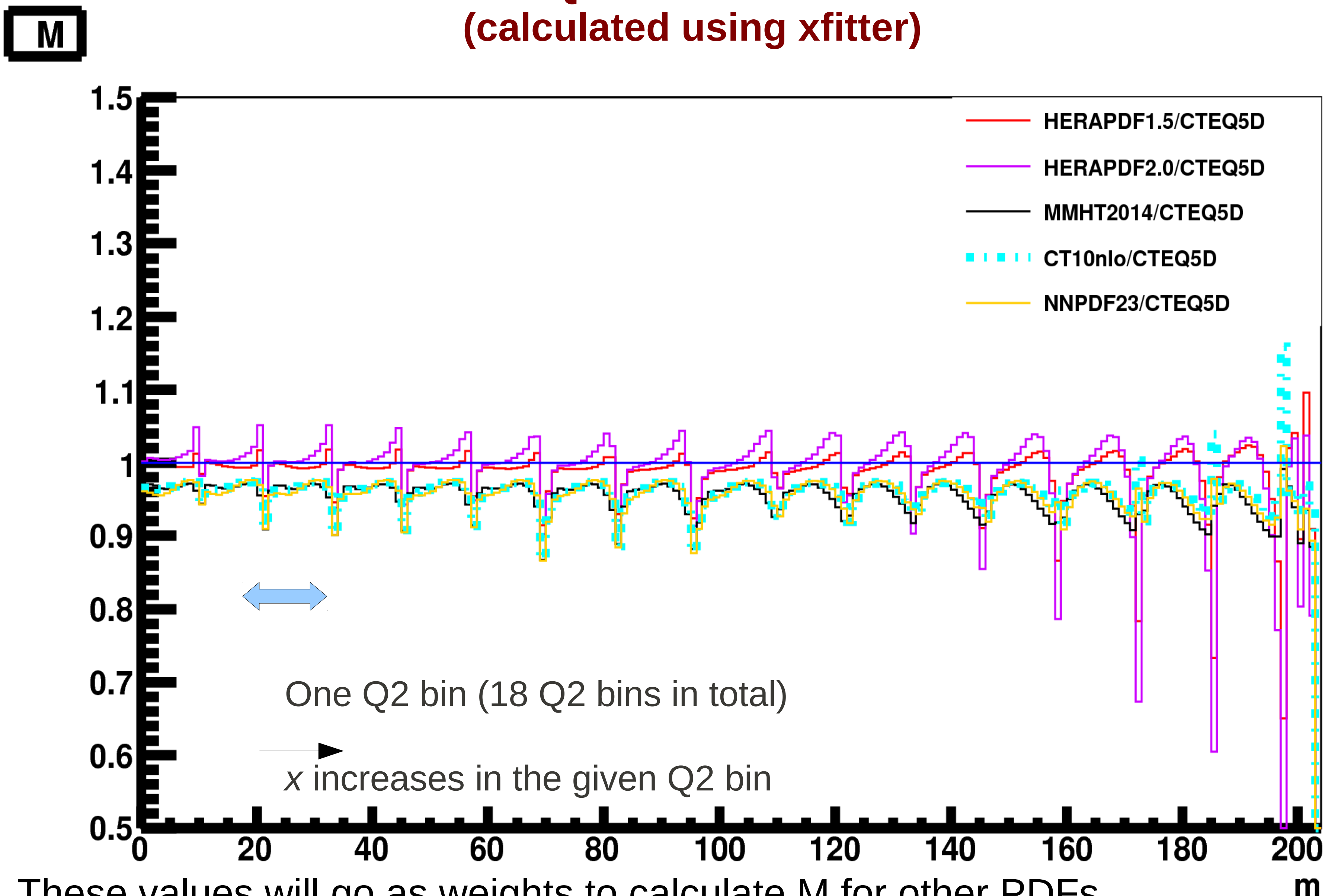
(for different PDFs)

Where for each (Q2,x) bin **M_PDF = M_CTEQ5d *(<u>RedXsec_PDF / RedXsec_CTEQ5D)</u>**

Shown on next slide

Note : the reduced cross sections are calculated at the bin average values of Q2 and x from CTEQ5d

Ratio of Reduced Xsec for the given PDF in extended binning (e+p) to CTEQ5D Reduced Xsec (calculated using xfitter)

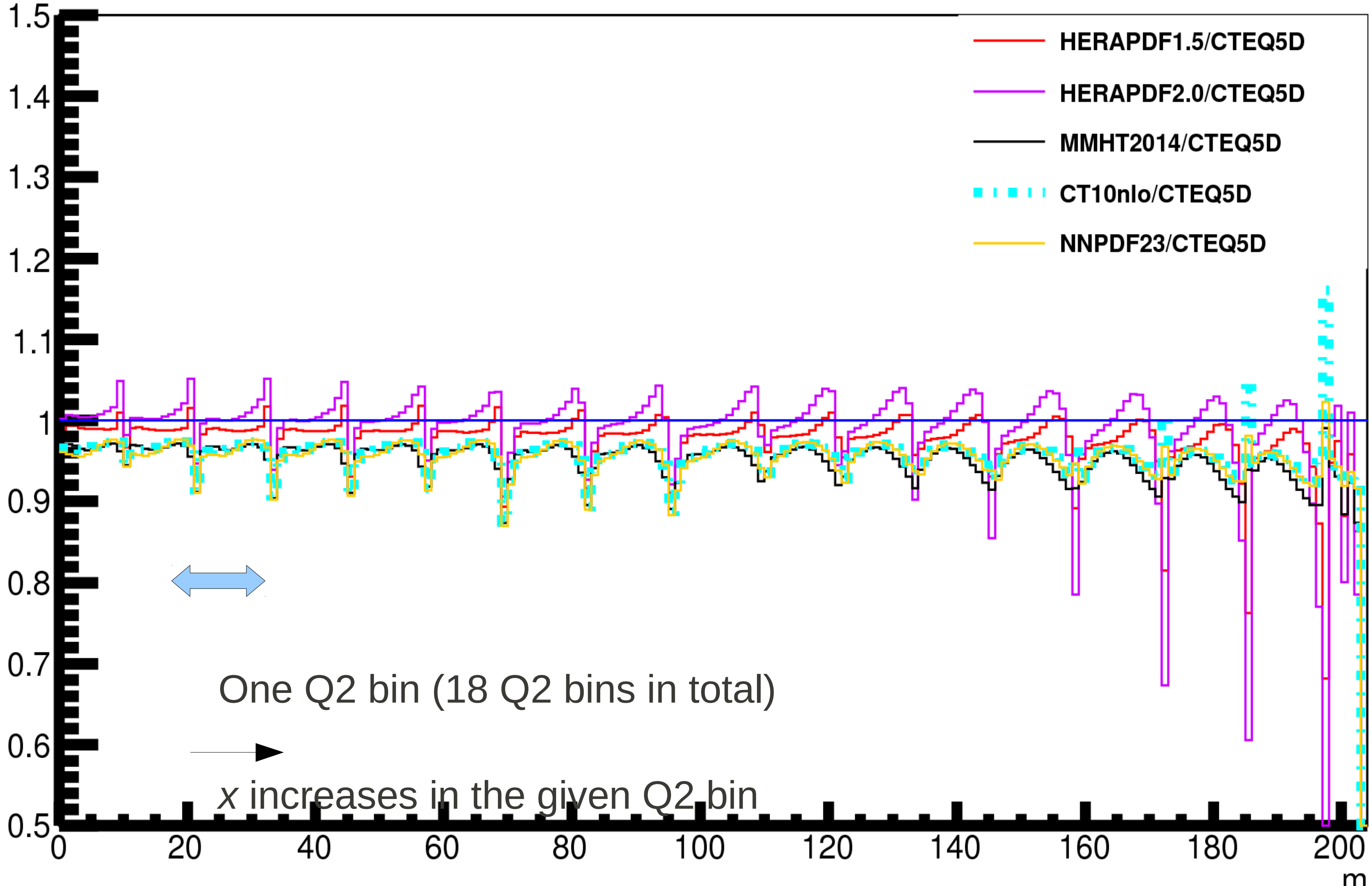


These values will go as weights to calculate M for other PDFs

There is a normalization difference between HERAPDF & other PDFs!

Ratio of Reduced Xsec for the given PDF in extended binning (e-p) to CTEQ5D Reduced Xsec (calculated using xfitter)

M

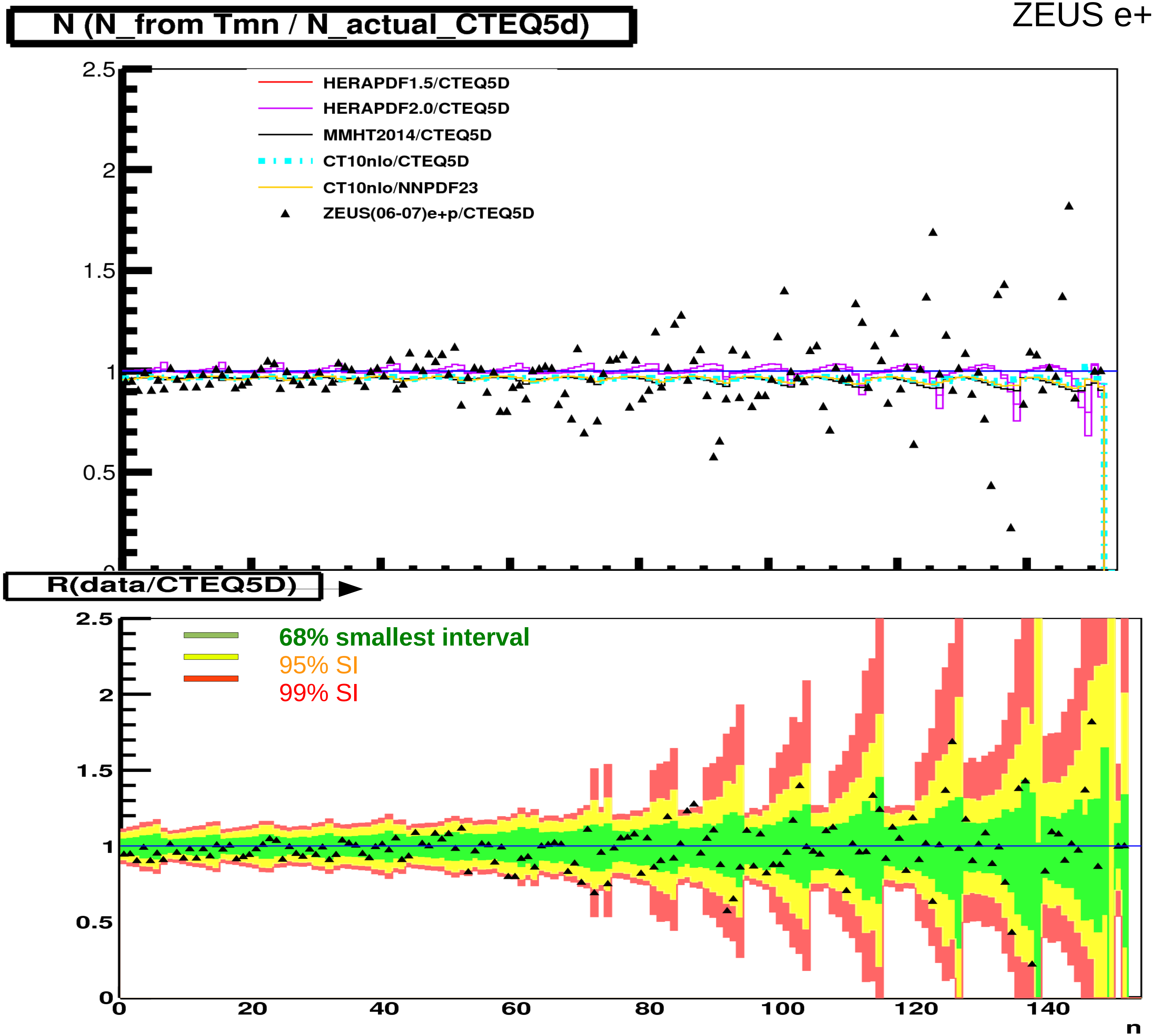


These values will go as weights to calculate M for other PDFs

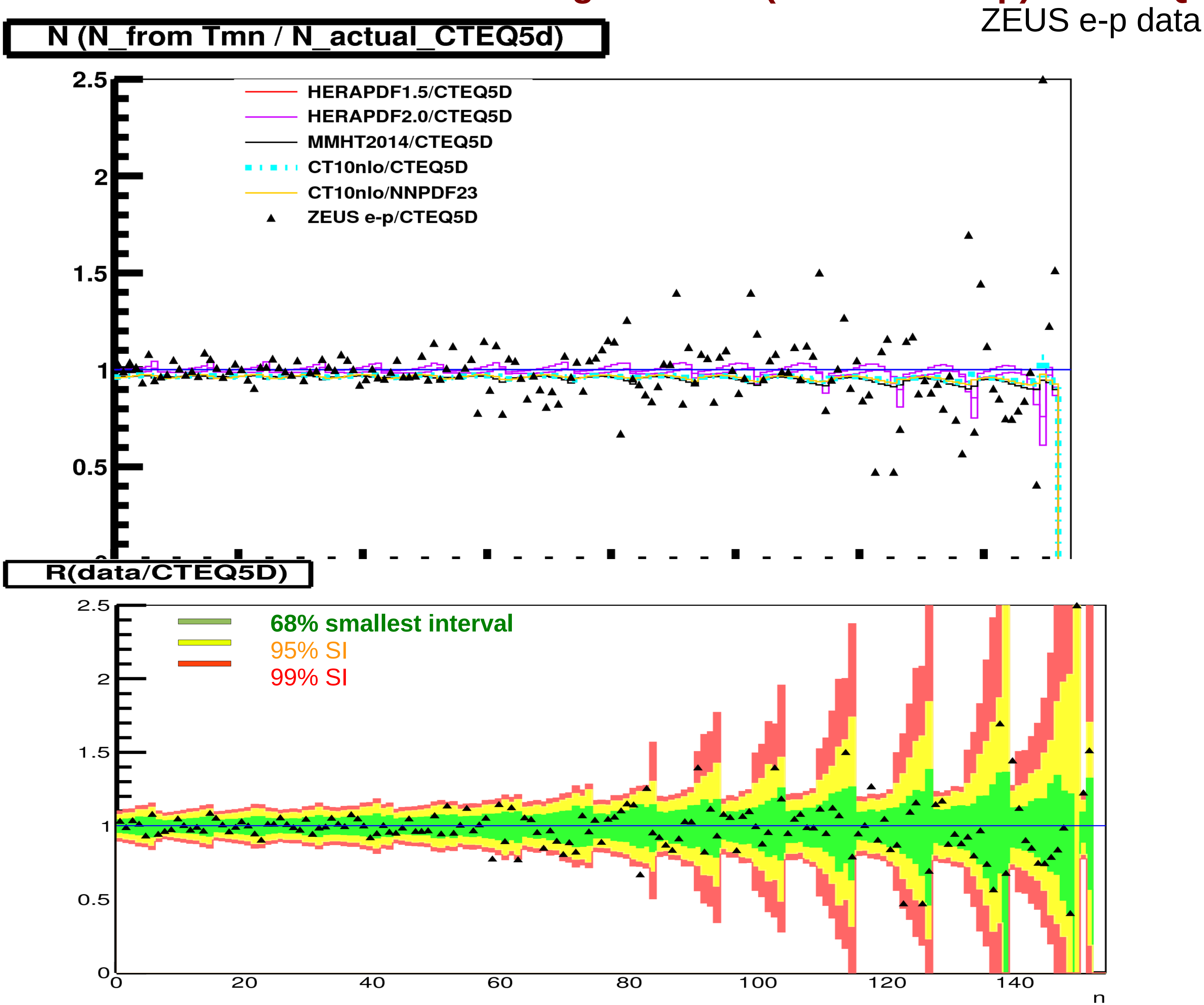
There is a normalization difference between HERAPDF & other PDFs!

Ratio of No. of events for the given PDF (shown for e+p) to CTEQ5D

ZEUS e+p data



Ratio of No. of events for the given PDF (shown for e-p) to CTEQ5D



Relative probability for data from different PDFs

Calculate Poisson probability P(n|ν) for each bin of [N] from different PDFs

$P(n|\nu) = (\nu^n \; e^{-\nu}) / n!$ Where $\nu_i = \Sigma_j A_{ij} m_j$ and reweighted to luminosity of data

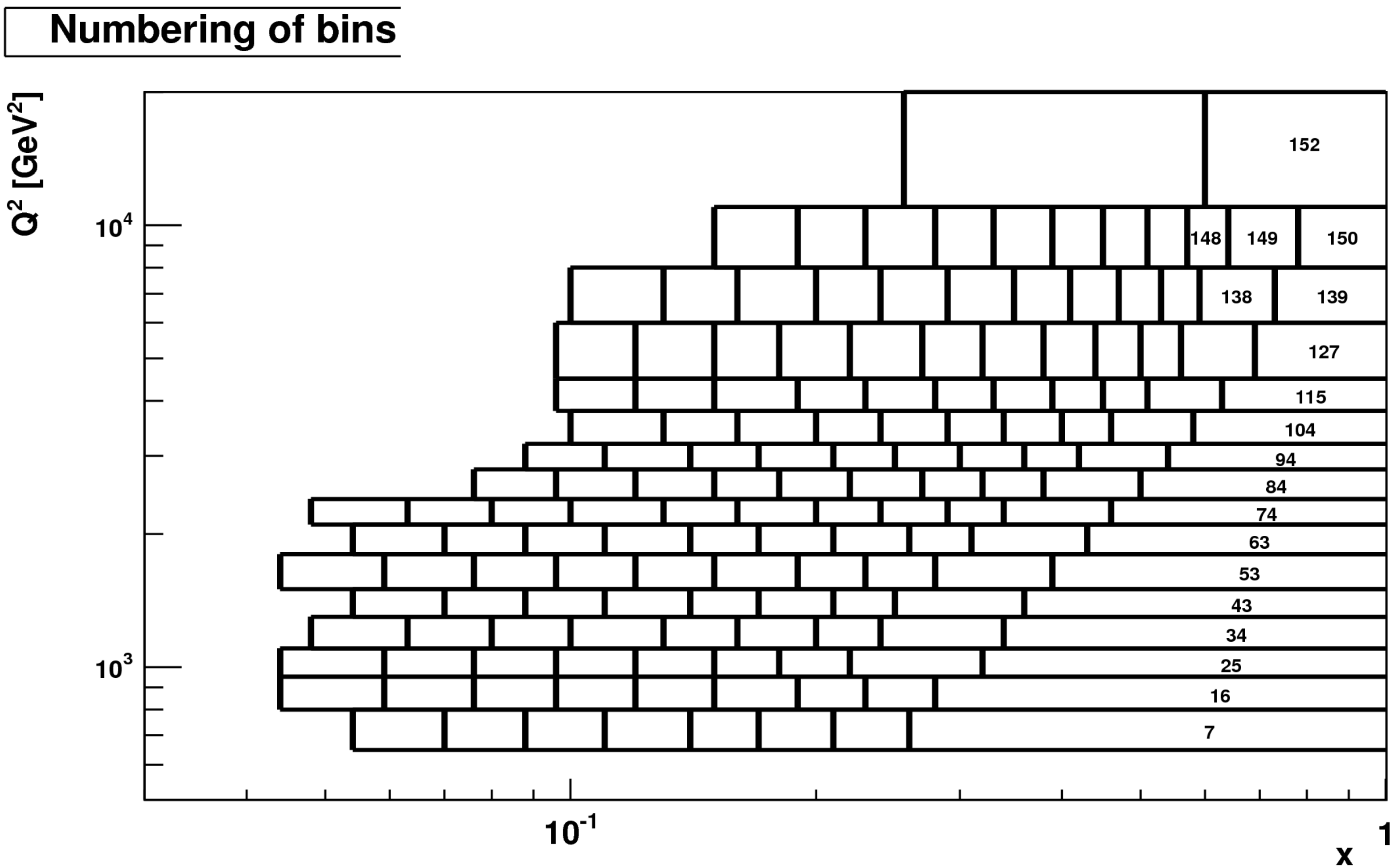
Total probability for each PDF = $\prod_i P_i(n|\nu)$

Calculating the **relative Probablity wrt. CTEQ5D**

	HERAPDF1.5	HERAPDF2.0	MMHT2014	CT10nlo	NNPDF2.3
e+p	0.755281	6.92953×10^{-05}	74.2666	35.3451	21.9212
e-p	0.0252737	0.0178363	2.11919×10^{-10}	1.55081×10^{-08}	1.011×10^{-09}

MMHT2014, CT10nlo, NNPDF2.3 much better than CTEQ5D for e^+P, much worse for e^-P. HERAPDF2.0 lower probability for both

Position of 20 high-x bins used for further analysis



Check 1 : Comparing N (calculated from Transfer Matrix) for different Pdfs
For high-x bins only (~20 bins)

Calculate Poisson probability P(n|ν) for each bin of [N] from different PDFs

$$\prod_i P(n|\nu) = \prod_i (\nu^n \; e^{-\nu}) / n!$$

Calculating the **relative Probablity wrt. CTEQ5D**

	HERAPDF1.5	HERAPDF2.0	MMHT2014	CT10nlo	NNPDF2.3
e+p	0.5781	0.01344	64.8686	31.0703	21.7235
e-p	0.67311	0.10109	0.0555	0.2793	35.8303

Excluding these 20 bins high-x bins the probability is as follows :

	HERAPDF1.5	HERAPDF2.0	MMHT2014	CT10nlo	NNPDF2.3
e+p	1.3065	0.00515603	1.14488	1.13759	0.611075
e-p	0.03755	0.17644	3.8159x10^-09	5.5512x10^-08	4.035x10^-09

at high x, MMHT, CT, NNPDF better for e+P data. disagreement comes primarily from lower Q2 and e-p

Why do we study in Probability numbers

What types of probabilities do we expect ?

E.g., imagine you expect 1 event, and measure 1, then the probability is

$$P(n|\nu) = e^{-\nu}\frac{\nu^n}{n!} = e^{-1} \approx 0.37$$

E.g., imagine you expect 10 events, and measure 8, then the probability is

$$P(n|\nu) = e^{-\nu}\frac{\nu^n}{n!} = e^{-10}\frac{10^8}{8!} \approx 0.11$$

E.g., imagine you expect 100 events, and measure 90, then the probability is

$$P(n|\nu) = e^{-\nu}\frac{\nu^n}{n!} = e^{-100}\frac{100^{90}}{90!} \approx 0.02$$

If we have 150 bins with probabilities ranging from a few % to few 10 %, then

$$P(\{n\}|\{\nu\}) = \prod_{i=1}^{150} e^{-\nu_i}\frac{\nu_i^{n_i}}{n_i!} \text{ maybe } 10^{-200} \quad \ln P \approx -500$$

Slide from Allen

Why do we study in Probability numbers

If the likelihood (product of the data probabilities) is a product of Gaussian distributions, then we have

$$\mathcal{L} \propto e^{-\chi^2/2}$$

and

$$\ln \mathcal{L}_1 - \ln \mathcal{L}_2 = \frac{1}{2}(\chi_2^2 - \chi_1^2)$$

So we can translate differences in the ln of the probabilities (multiplied by -2) to equivalent chi squared differences

If we look at ratios of probabilities, and again assuming Gaussian distributions, then

$$\frac{P_1}{P_2} = e^{-(\chi_1^2 - \chi_2^2)/2}$$

so taking -2* the natural logarithm of a probability ratio is again equivalent to a chi squared difference

Slide from Allen

Study 1 : Applying Normalization correction (e+p) in ν (BAT)

ν = ν * (1 + parameters[0])
Where penality is added = LogGaus (parameters[0],0.,1.8,0)

Parameters [0] is allowed to vary in [-0.9,0.9]

Luminosity error = 1.8%

Centered at zero

Global Mode for Parameters[0] from BAT is as follows :

	HERAPDF1.5	**HERAPDF2.0**	**MMHT2014**	**CT10nlo**
Parameter[0]	-0.027	-0.027	0.009	0.009
CorFactor	0.973	0.973	1.009	1.009
LogProb (w/o Cor)	-531.448	-540.745	-526.86	-527.603
LogProb (after Cor)	-528.092	-530.845	-525.445	-526.578
Delta_chi2	6.712	19.8	2.83	2.05

normalization of HERA-pdf shows big change, not for others!

Study 1 : Applying Normalization correction (e-p) in ν (BAT)

ν = ν * (1 + parameters[0])
Where penality in added = LogGaus (parameters[0],0.,1.8,0)

Luminosity error = 1.8%

Centered at zero

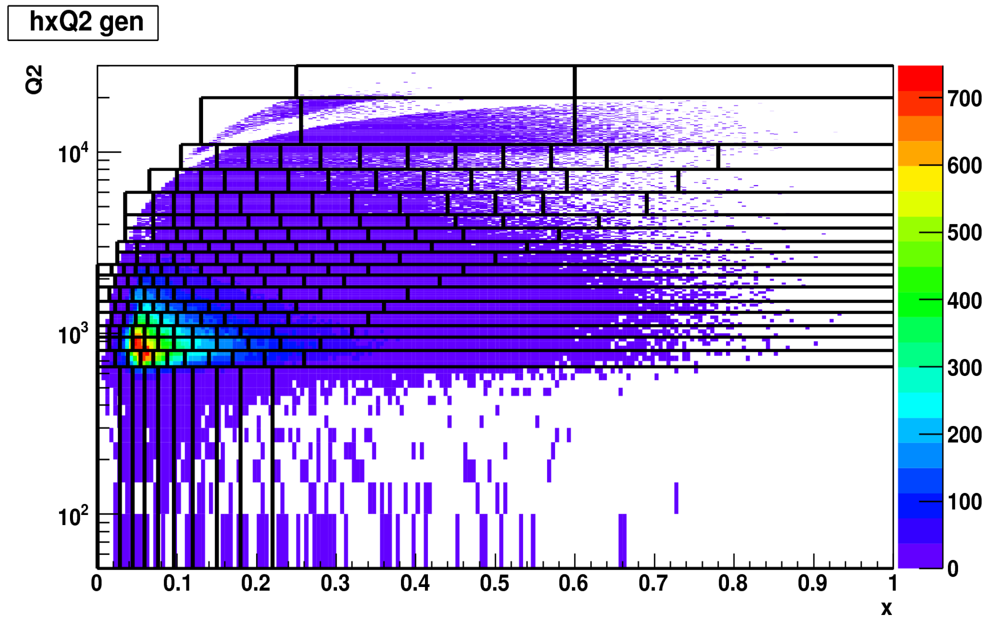
Parameters [0] is allowed to vary in [-0.9,0.9]

Global Mode for Parameters[0] from BAT is as follows :

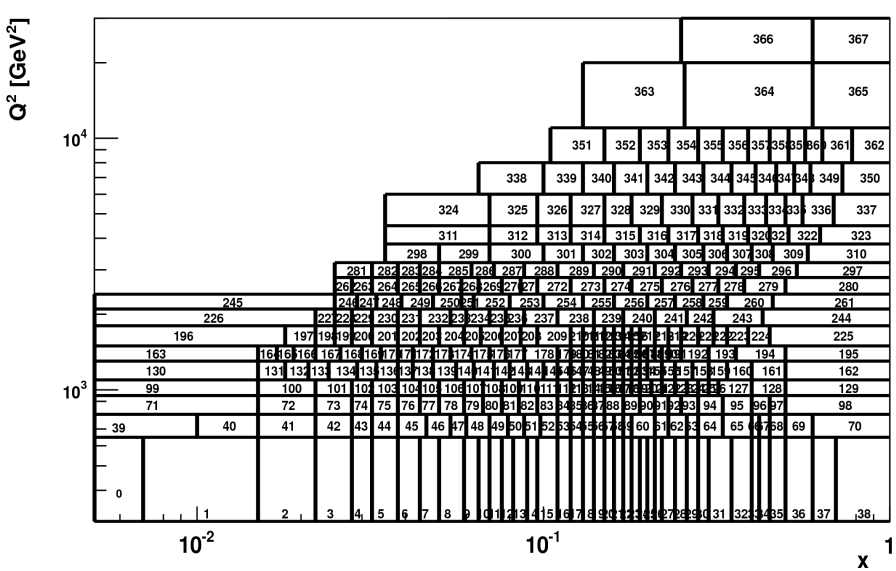
	HERAPDF1.5	HERAPDF2.0	MMHT2014	CT10nlo
Parameter[0]	0.009	-0.009	0.045	0.027
CorFactor	1.009	0.991	1.045	1.027
LogProb (w/o Cor)	-580.685	-581.034	-599.282	-594.989
LogProb (after Cor)	-578.89	-580.453	-579.458	-579.25
Delta_chi2	3.59	1.162	39.648	31.478

normalization of CT, MMHT shows big change. After shifting normalization, all agree

Check II- Increase the M bins : 203 ---> increased to 368 bins

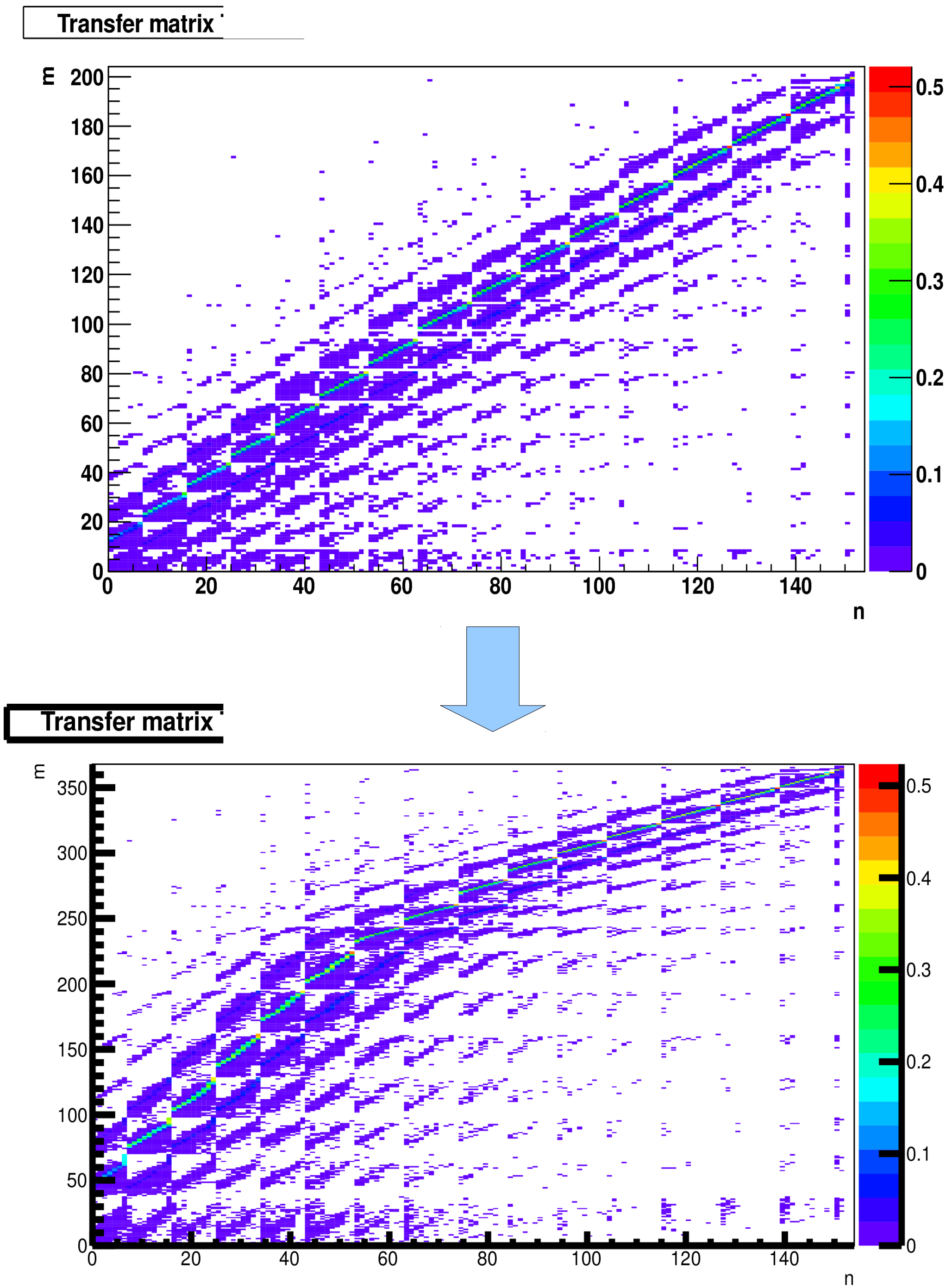


203 bins



368 bins

Check II- Increase the M bins, build a new Transfer Matrix

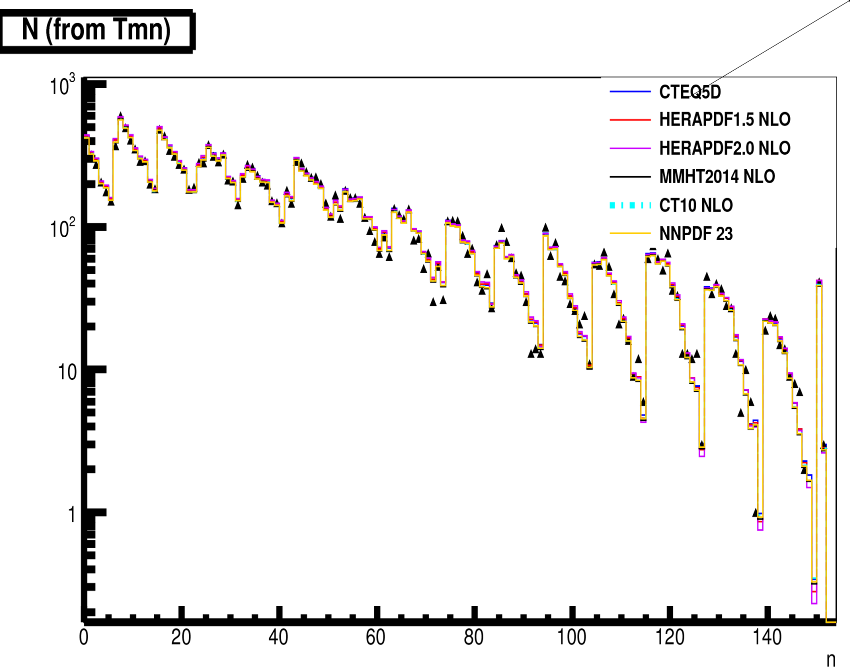


153 X 203 bins

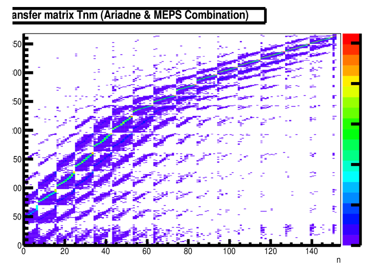
153 X 368 bins

Check II- Repeat the study using new Transfer Matrix

N = AM

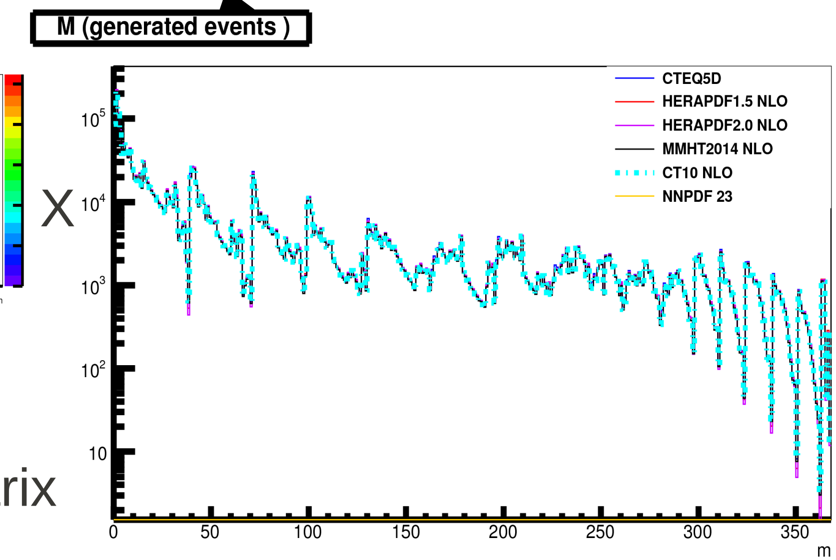


=



Transfer Matrix

X



Predicted x-Q2 events in
Cross section binning

(for different PDFs)

Generated x-Q2
events in
Extended binning

(for different PDFs)

Where for each (Q2,x) bin **M_PDF = M_CTEQ5d *<u>(RedXsec_PDF / RedXsec_CTEQ5D)</u>**

Check II: Comparing N for different PDFs for new Transfer Matrix

Calculate Poisson probability P(n|ν) for each bin of [N] from different PDFs

$$P(n|\nu) = (\nu^n \ e^{-\nu}) / n!$$

Total probability for each PDF =

$$\prod_i P_i(n|\nu)$$

Calculating the **relative Probablity wrt. CTEQ5D**

	HERAPDF1.5	HERAPDF2.0	MMHT2014	CT10nlo	NNPDF2.3
e+p	**0.755281**	**6.92953×10^{-05}**	**74.2666**	**35.3451**	**21.9212**
	4.2284	**4.15343×10^{-4}**	**76.5756**	**35.8755**	**21.7235** New values

Only for 20 high-x bins

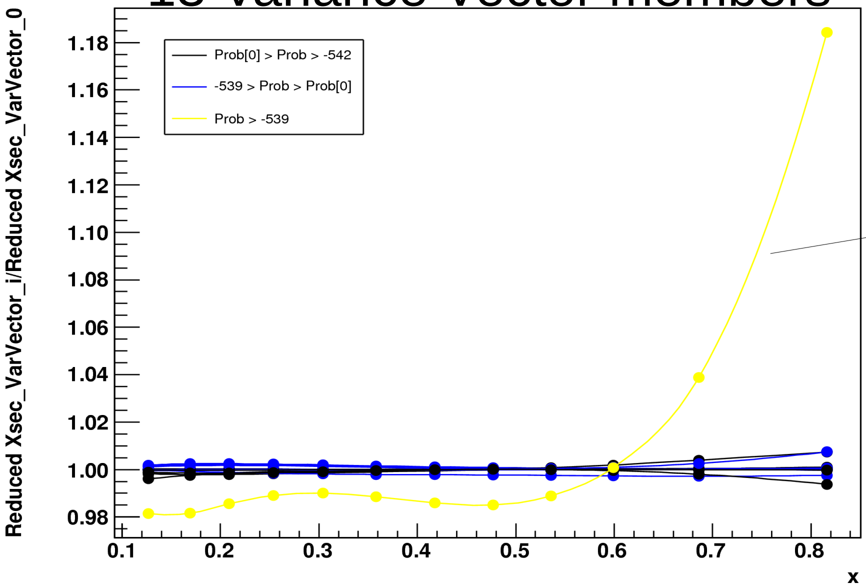
	HERAPDF1.5	HERAPDF2.0	MMHT2014	CT10nlo	NNPDF2.3
e+p	**0.5781**	**0.01344**	**64.8686**	**31.0703**	**21.7235**
	0.8403	**0.0788**	**64.9413**	**31.5864**	**35.8303**

HERAPDF shows some sensitivity to the finer binning, other pdfs do not!

For high-x bins, conclusion stays the same!

Study II - Check the Probability for each member in HERAPDF2.0 Error band For Q^2 ~ 9200 GeV2

13 Variance Vector members

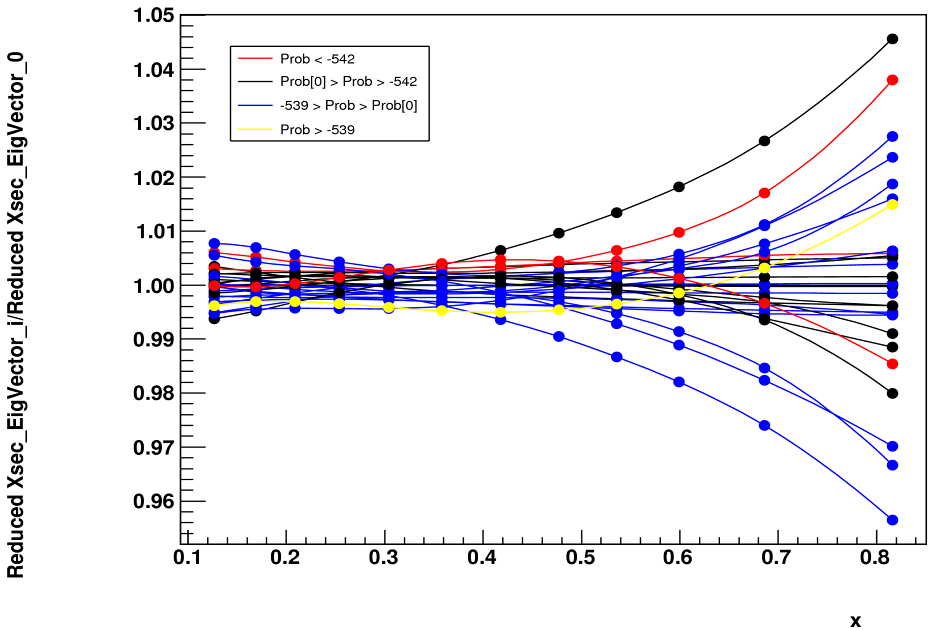


Prob[0] : Central value

13th (last) member in the Vector
Corresponds to u-valence parameter

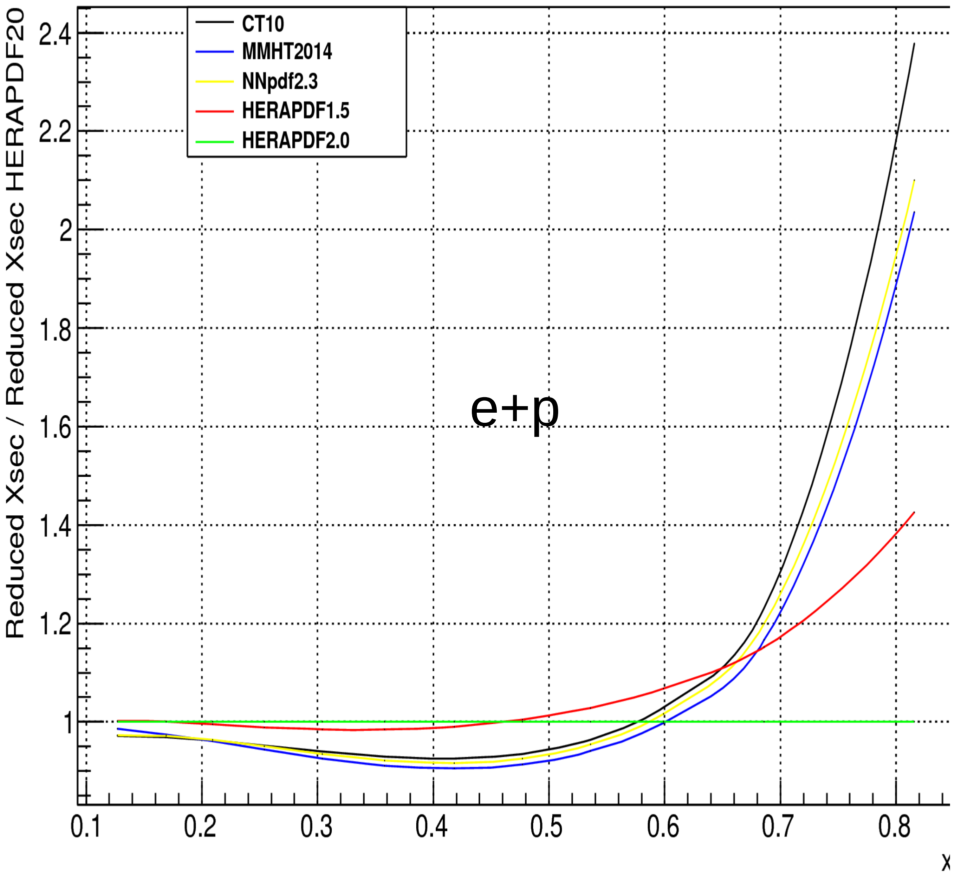
Variation in the PDFs within HERAPDF2.0 much smaller than the inter PDF variation. Big change comes from #13 which has different normalization!

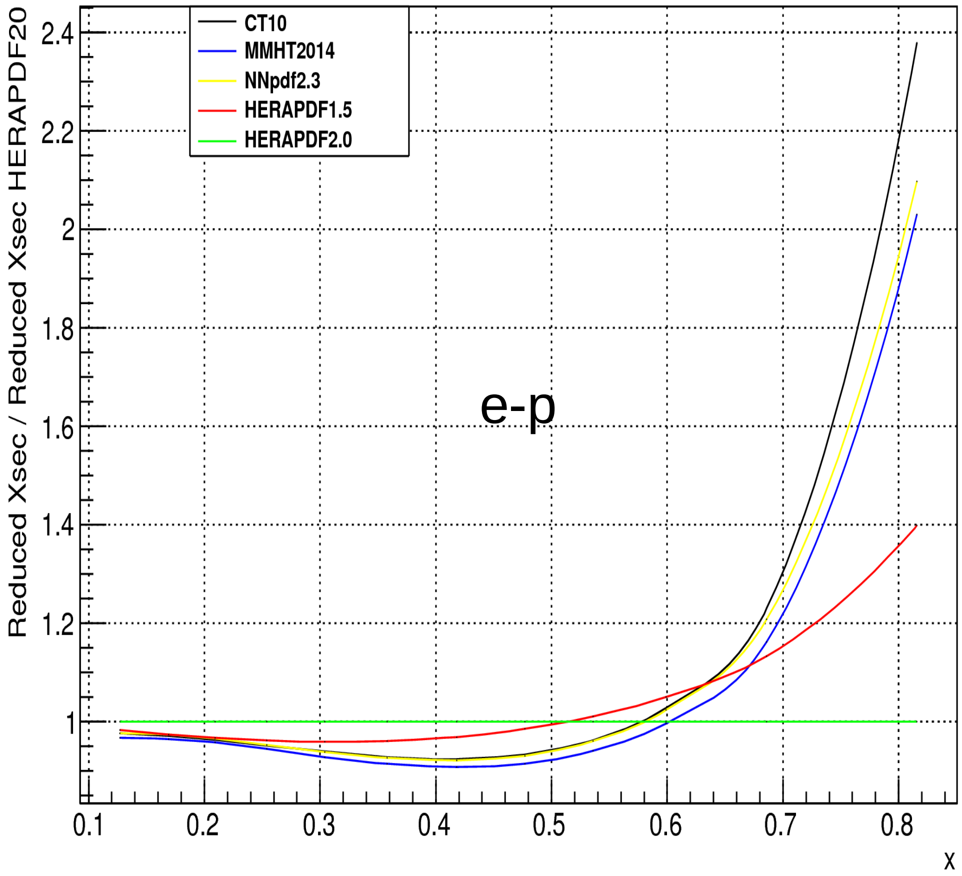
28 Eigen Vector members



How other PDFs behave in the same x-Q^2 region

For Q^2 ~ 9200 GeV2





Difference between PDFs much bigger than difference within HERAPDF set!

Results

1)Technique of building Transfer Matrix Shown.

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

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

4) A comparison of different PDFs can be done on the basis of best explanation to the high-x data using Transfer Matrix.

5)PDFs have been checked for significant systematic trends in normalization.

Back Up
(some Old slides)

Check 1 : Comparing N (calculated from Transfer Matrix) for different Pdfs

For high-x bins only (~20 bins) e+p

Q2bin	x bin	N_data	CTEQ5D	CT10nlo	HERAPDF1.5	HERAPDF2.0	MMHT2014	NNPDF2.3
650 - 800	0.26 - 1.00	371	405.72	394.06	410.30	423.25	389.24	393.99
800 - 950	0.28 - 1.00	482	489.51	474.57	495.66	511.02	468.23	474.34
950 - 1100	0.32 - 1.00	281	306.00	295.75	309.94	319.02	291.59	295.59
1100 - 1300	0.34 - 1.00	275	263.15	253.57	266.32	273.48	249.89	253.43
1300 - 1500	0.36 - 1.00	146	159.65	153.43	161.38	165.35	151.13	153.32
1500 - 1800	0.39 - 1.00	115	137.63	131.77	138.76	141.57	129.71	131.65
1800 - 2100	0.43 - 1.00	62	71.67	68.33	71.88	72.89	67.21	68.25
2100 - 2400	0.46 - 1.00	31	40.99	38.95	40.84	41.15	38.27	38.89
2400 - 2800	0.50 - 1.00	27	29.22	27.68	28.83	28.79	27.16	27.62
2800 - 3200	0.54 - 1.00	13	15.03	14.20	14.62	14.43	13.90	14.15
3200 - 3800	0.58 - 1.00	11	11.01	10.41	10.53	10.26	10.15	10.35
3800 - 4500	0.63 - 1.00	6	4.82	4.57	4.47	4.26	4.43	4.53
4500 - 6000	0.69 - 1.00	3	3.03	2.92	2.67	2.47	2.79	2.86
6000 - 8000	0.59 - 0.73	1	4.44	4.16	4.18	3.98	4.03	4.11
6000 - 8000	0.73 - 1.00	0	0.98	0.96	0.83	0.75	0.90	0.93
8000 - 11000	0.57 - 0.64	2	2.29	2.13	2.21	2.13	2.07	2.11
8000 - 11000	0.64 - 0.78	0	1.82	1.72	1.63	1.49	1.64	1.68
8000 - 11000	0.78 - 1.00	0	0.34	0.35	0.27	0.23	0.32	0.33
11000 - 20000	0.60 - 1.00	3	2.99	2.82	2.80	2.60	2.70	2.75

Check 1 : Comparing N (calculated from Transfer Matrix) for different Pdfs

For high-x bins only (~20 bins) e-p

Q2bin	x bin	N_data	CTEQ5D	CT10nlo	HERAPDF1.5	HERAPDF2.0	MMHT2014	NNPDF2.3
650 - 800,	0.26 - 1.00,	504,	532.79,	517.39,	537.46,	555.91,	511.00,	517.30
800 - 950,	0.28 - 1.00,	671,	635.27,	615.70,	642.07,	663.36,	607.38,	615.41
950 - 1100,	0.32 - 1.00,	414,	407.28,	393.53,	412.05,	424.52,	388.02,	393.37
1100 - 1300,	0.34 - 1.00,	368,	348.28,	335.45,	352.39,	361.86,	330.60,	335.31
1300 - 1500,	0.36 - 1.00,	202,	210.08,	201.77,	212.44,	217.51,	198.75,	201.66
1500 - 1800,	0.39 - 1.00,	173,	181.26,	173.43,	182.95,	186.35,	170.75,	173.33
1800 - 2100,	0.43 - 1.00,	74,	95.75,	91.18,	96.29,	97.25,	89.70,	91.12
2100 - 2400,	0.46 - 1.00,	51,	53.00,	50.29,	53.01,	53.10,	49.43,	50.25
2400 - 2800,	0.50 - 1.00,	36,	37.61,	35.57,	37.30,	36.94,	34.90,	35.52
2800 - 3200,	0.54 - 1.00,	19,	20.34,	19.21,	19.95,	19.49,	18.80,	19.16
3200 - 3800,	0.58 - 1.00,	17,	14.32,	13.52,	13.81,	13.28,	13.18,	13.47
3800 - 4500,	0.63 - 1.00,	5,	6.32,	6.00,	5.93,	5.55,	5.80,	5.95
4500 - 6000,	0.69 - 1.00,	3,	4.34,	4.18,	3.88,	3.50,	3.98,	4.11
6000 - 8000,	0.59 - 0.73,	10,	5.88,	5.49,	5.53,	5.22,	5.32,	5.46
6000 - 8000,	0.73 - 1.00,	1,	1.47,	1.43,	1.26,	1.11,	1.34,	1.39
8000 - 11000,	0.57 - 0.64,	4,	4.05,	3.75,	3.86,	3.73,	3.64,	3.73
8000 - 11000,	0.64 - 0.78,	1,	2.46,	2.32,	2.21,	2.02,	2.21,	2.28
8000 - 11000,	0.78 - 1.00,	1,	0.32,	0.34,	0.24,	0.19,	0.30,	0.31
11000 - 20000,	0.60 - 1.00,	8,	5.28,	4.94,	4.82,	4.58,	4.75,	4.90