Using ZEUS high-x, high-Q^2 data to constrain proton parton densities

Allen Caldwell

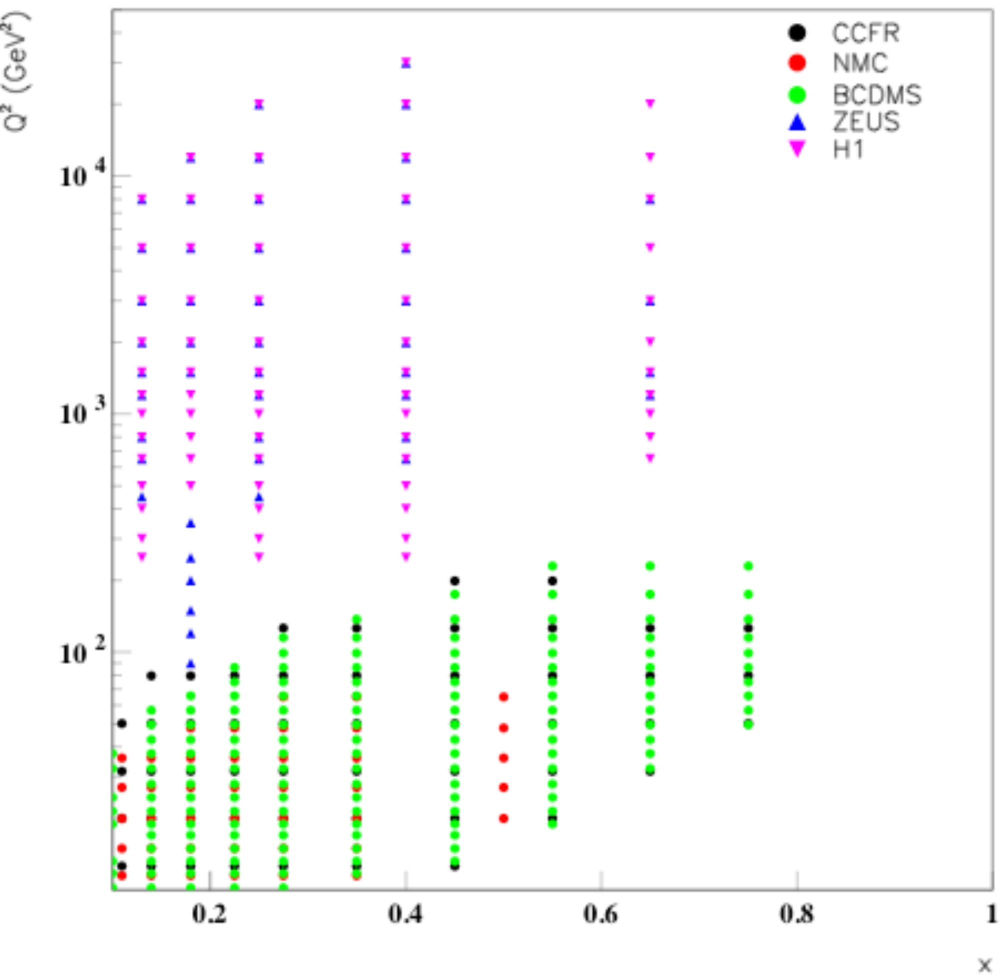
Max Planck Institute for Physics

MAX-PLANCK-GESELLSCHAFT

Max-Planck-Institut für Physik
(Werner-Heisenberg-Institut)

Motivation

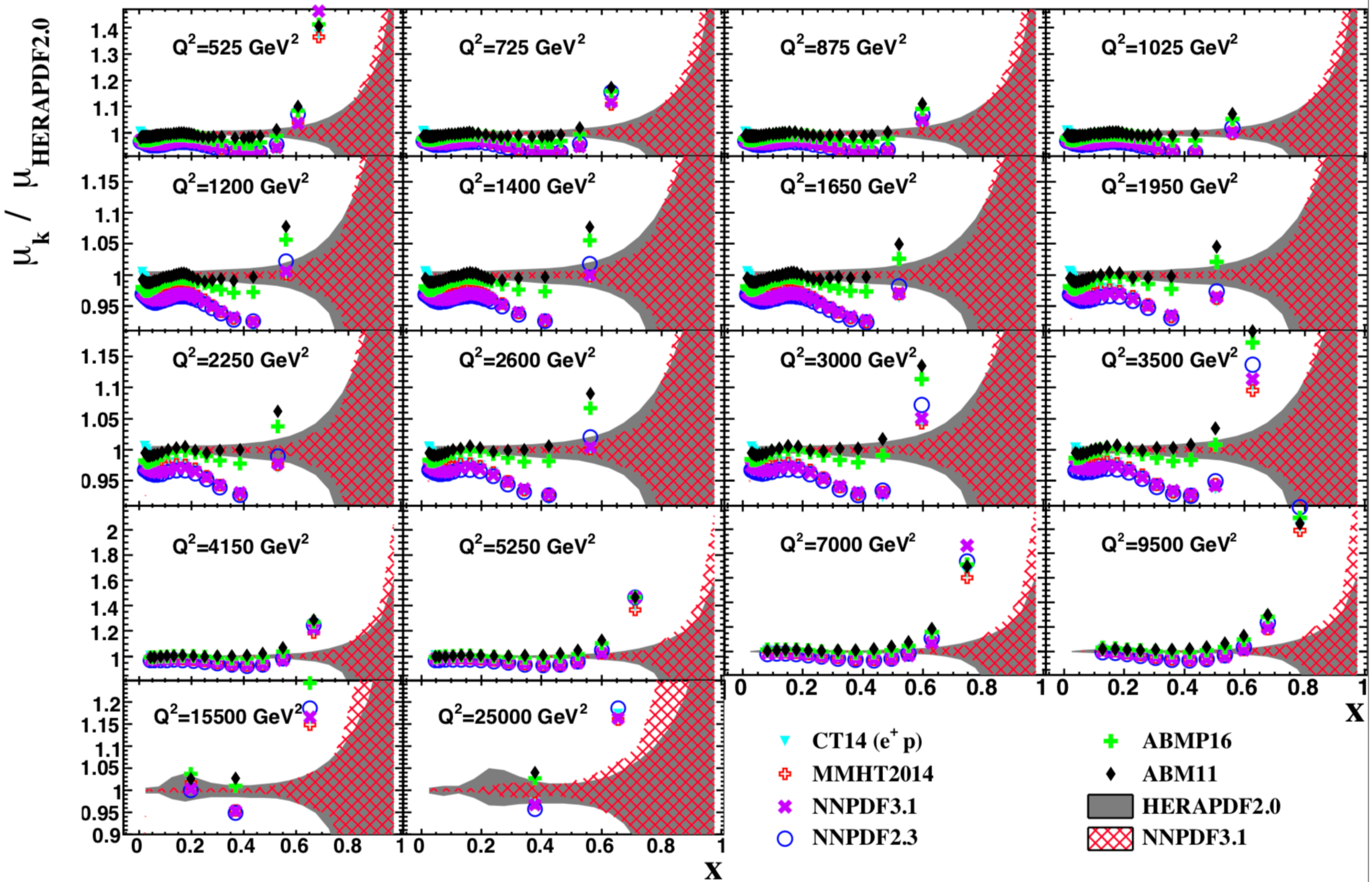
Our information on the very high x behavior of the parton densities in the DGLAP validity regime is primarily theoretical and assumption-based.



BCDMS has measured F_2 up to x=0.75

H1, ZEUS have previously measured F_2 up to x=0.65

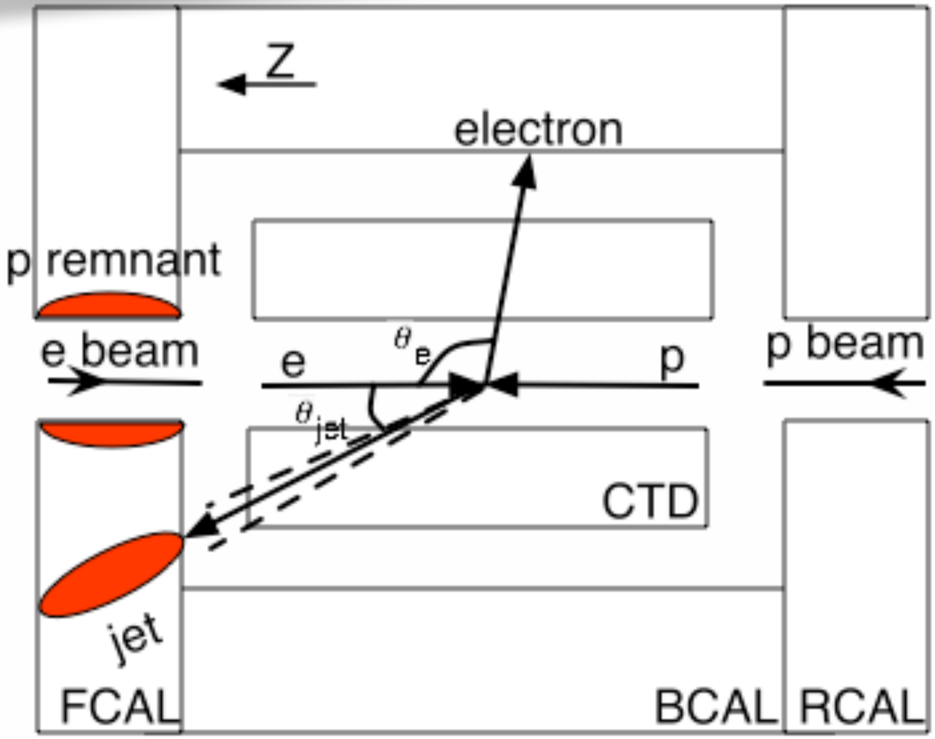
ZEUS has measured up to x=1, but these data are not (yet) included in PDF fits.

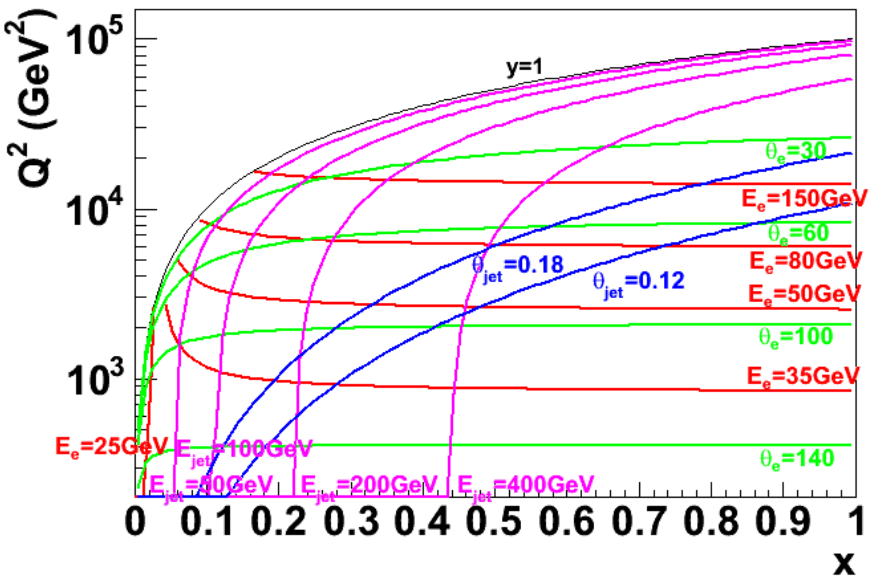


The PDF's are poorly determined at high-x. Sizable differences in expectations (much bigger than quoted uncertainties) despite the fact that fits typically use similar parametrization $xq \propto (1-x)^{\eta}$. Is it possible to improve this situation ?

Measurement of neutral current $e^{\pm}p$ cross sections at high Bjorken x with the ZEUS detector

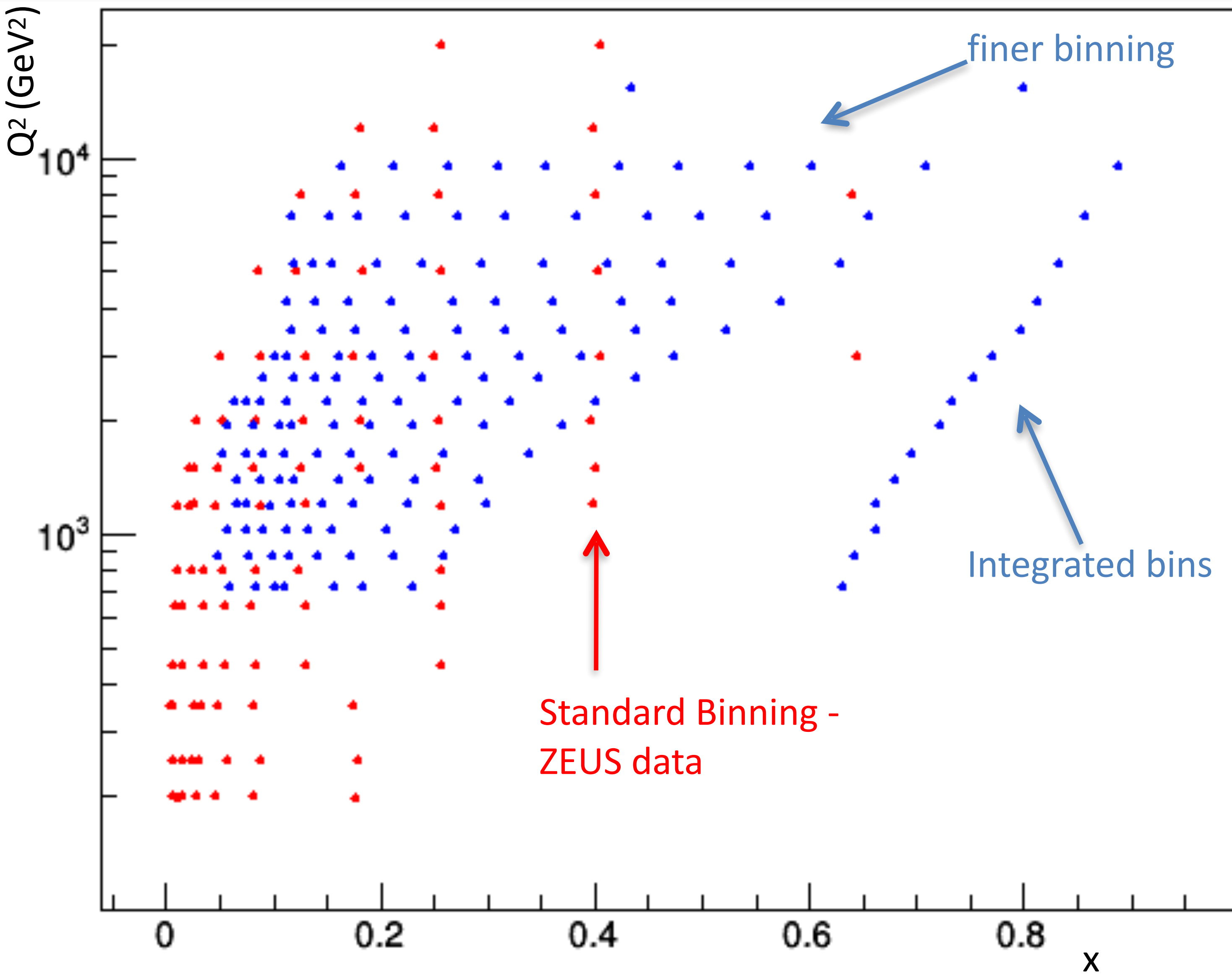
H. Abramowicz *et al.* (ZEUS Collaboration)
Phys. Rev. D **89**, 072007 – Published 8 April 2014

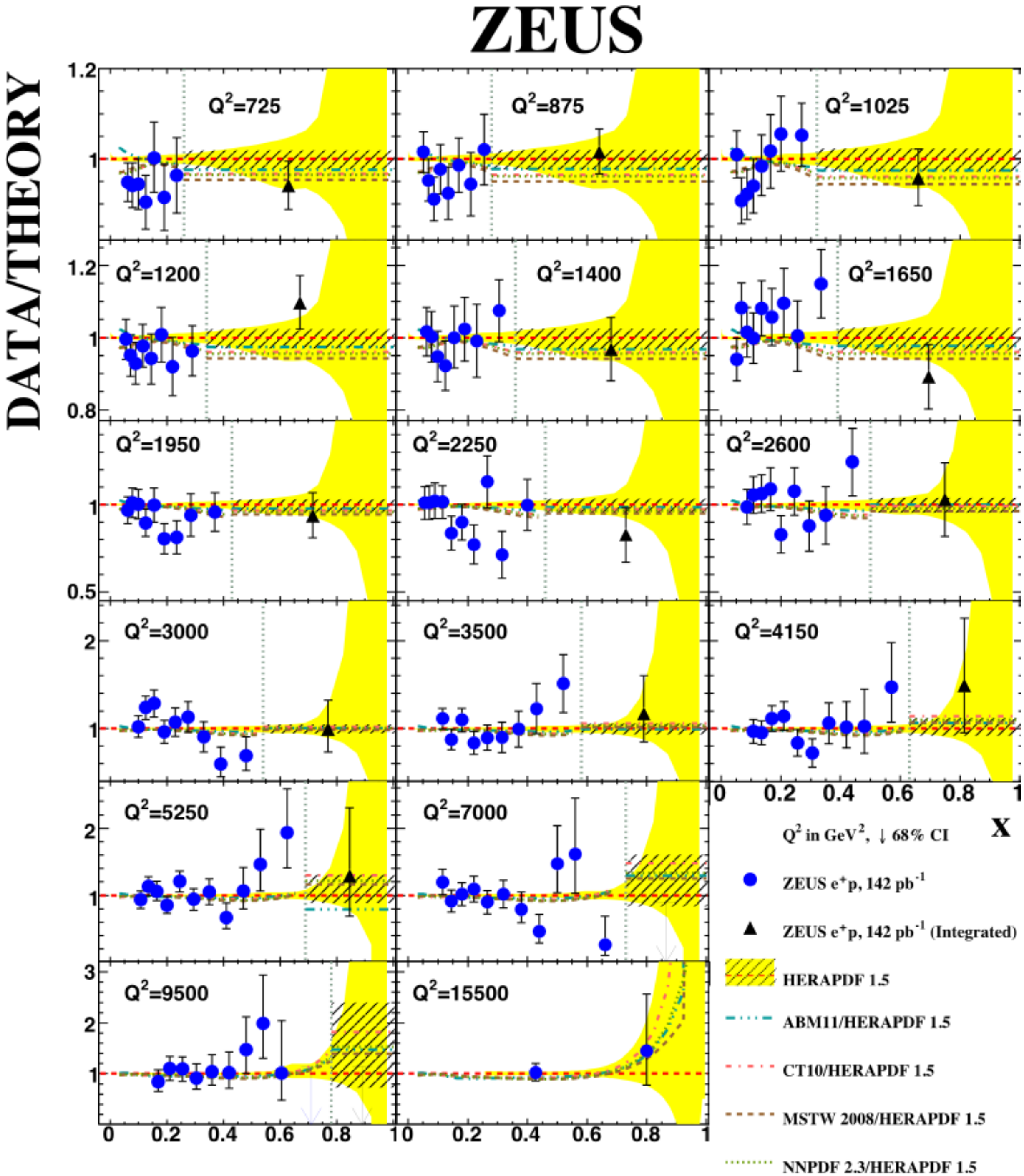




At high Q^2, scattered electron seen with ≈100% acceptance
For not too high x, measure x from jet and measure in small bins
For $x > x_{Edge}$, measure integrated cross section to x=1

Cross Sections





Error bars indicate the range of probable values for the underlying cross section given the measured data. How to use this information in a fit ?

Use the observed number of events & calculate the probability to see this number given the model expectation.

5250	0.62	5	$1.76e-04$	$^{+55.2}_{-35.2}$	$^{+11.1}_{-10.5}$	$^{+9.2}_{-9.1}$	$^{-3.2}_{+4.6}$	$^{+0.1}_{+0.1}$	$^{+3.7}_{-3.7}$	$^{+0.6}_{-0.6}$
7000	0.12	93	$1.61e-02$	$^{+10.4}_{-10.4}$	$^{+4.0}_{-5.1}$	$^{+3.3}_{-3.6}$	$^{-1.1}_{+0.8}$	$^{+0.8}_{-0.5}$	$^{-0.7}_{+0.7}$	$^{+0.0}_{-0.0}$
7000	0.14	89	$1.25e-02$	$^{+10.6}_{-10.6}$	$^{+3.7}_{-5.2}$	$^{+3.4}_{-3.5}$	$^{-1.3}_{+1.2}$	$^{-0.5}_{+0.2}$	$^{-0.9}_{+0.9}$	$^{+0.0}_{-0.0}$
7000	0.18	68	$7.02e-03$	$^{+12.1}_{-12.1}$	$^{+3.9}_{-3.6}$	$^{+3.4}_{-3.4}$	$^{-0.6}_{+0.6}$	$^{-0.6}_{+0.4}$	$^{-0.4}_{+0.4}$	$^{+0.0}_{-0.0}$
7000	0.22	56	$5.60e-03$	$^{+13.4}_{-13.4}$	$^{+4.2}_{-4.2}$	$^{+3.9}_{-3.9}$	$^{-1.4}_{+1.1}$	$^{-0.4}_{+1.0}$	$^{-0.4}_{+0.4}$	$^{+0.0}_{-0.0}$
7000	0.26	49	$3.79e-03$	$^{+14.3}_{-14.3}$	$^{+4.6}_{-4.8}$	$^{+3.9}_{-4.0}$	$^{-0.2}_{+2.1}$	$^{+0.2}_{-0.2}$	$^{-0.5}_{+0.5}$	$^{+0.0}_{-0.0}$
7000	0.32	41	$2.70e-03$	$^{+15.6}_{-15.6}$	$^{+5.1}_{-4.7}$	$^{+5.3}_{-4.5}$	$^{-1.4}_{+0.8}$	$^{-0.4}_{+0.4}$	$^{-0.2}_{+0.2}$	$^{+0.0}_{-0.0}$
7000	0.38	23	$1.52e-03$	$^{+20.9}_{-20.9}$	$^{+6.4}_{-6.2}$	$^{+5.5}_{-5.5}$	$^{-1.8}_{+1.7}$	$^{-0.7}_{+0.4}$	$^{+2.0}_{-2.0}$	$^{+0.0}_{-0.0}$
7000	0.44	17	$1.15e-03$	$^{+27.2}_{-21.3}$	$^{+8.4}_{-7.9}$	$^{+7.1}_{-7.1}$	$^{-2.7}_{+2.7}$	$^{-0.0}_{+0.2}$	$^{-2.4}_{+2.4}$	$^{+0.0}_{-0.0}$
7000	0.50	8	$5.38e-04$	$^{+41.8}_{-29.4}$	$^{+9.7}_{-10.3}$	$^{+9.5}_{-9.5}$	$^{-1.6}_{+1.4}$	$^{-0.3}_{+0.4}$	$^{+2.4}_{-2.4}$	$^{+0.1}_{-0.1}$
7000	0.56	4	$2.37e-04$	$^{+63.2}_{-38.2}$	$^{+12.3}_{-11.8}$	$^{+11.3}_{-11.3}$	$^{-3.4}_{+3.4}$	$^{+0.1}_{-0.0}$	$^{+1.2}_{-1.2}$	$^{+0.2}_{-0.2}$
7000	0.66	10	$2.30e-04$	$^{+36.7}_{-26.8}$	$^{+12.6}_{-13.6}$	$^{+12.1}_{-12.3}$	$^{-4.3}_{+2.5}$	$^{-0.3}_{-0.0}$	$^{+2.0}_{-2.0}$	$^{+0.9}_{-0.9}$
9500	0.17	76	$6.77e-03$	$^{+11.5}_{-11.5}$	$^{+5.6}_{-7.7}$	$^{+4.9}_{-4.9}$	$^{-2.0}_{+2.3}$	$^{+0.2}_{-0.2}$	$^{-0.6}_{+0.6}$	$^{+0.0}_{-0.0}$
9500	0.21	53	$3.87e-03$	$^{+13.7}_{-13.7}$	$^{+5.8}_{-5.1}$	$^{+4.3}_{-4.5}$	$^{-1.1}_{+1.8}$	$^{-0.7}_{+0.4}$	$^{-1.1}_{+1.1}$	$^{+0.0}_{-0.0}$
9500	0.25	40	$2.27e-03$	$^{+15.8}_{-15.8}$	$^{+4.8}_{-4.9}$	$^{+4.5}_{-4.5}$	$^{-2.0}_{+1.5}$	$^{+0.2}_{+0.4}$	$^{+0.1}_{-0.1}$	$^{+0.0}_{-0.0}$
9500	0.31	27	$1.50e-03$	$^{+19.2}_{-19.2}$	$^{+5.7}_{-8.1}$	$^{+5.2}_{-5.3}$	$^{-2.6}_{+1.5}$	$^{-0.5}_{+0.2}$	$^{+1.5}_{-1.5}$	$^{+0.0}_{-0.0}$
9500	0.36	19	$8.89e-04$	$^{+25.5}_{-20.3}$	$^{+6.6}_{-6.1}$	$^{+5.9}_{-5.9}$	$^{-1.0}_{+1.9}$	$^{-0.4}_{+0.2}$	$^{-0.3}_{+0.3}$	$^{+0.0}_{-0.0}$
9500	0.42	12	$5.64e-04$	$^{+33.1}_{-24.8}$	$^{+11.3}_{-7.5}$	$^{+13.4}_{-7.3}$	$^{-1.0}_{+2.4}$	$^{-0.8}_{+0.5}$	$^{-0.8}_{+0.8}$	$^{+0.0}_{-0.0}$
9500	0.48	8	$3.63e-04$	$^{+41.7}_{-29.4}$	$^{+10.5}_{-10.4}$	$^{+9.2}_{-9.2}$	$^{-2.6}_{+2.4}$	$^{-0.4}_{+0.6}$	$^{-3.4}_{+3.4}$	$^{+0.0}_{-0.0}$
9500	0.54	5	$2.31e-04$	$^{+55.2}_{-35.2}$	$^{+14.3}_{-13.7}$	$^{+12.5}_{-12.2}$	$^{-1.7}_{+3.6}$	$^{-0.2}_{+0.6}$	$^{+5.7}_{-5.7}$	$^{+0.1}_{-0.1}$
9500	0.61	4	$1.39e-04$	$^{+63.3}_{-38.3}$	$^{+15.5}_{-15.4}$	$^{+14.6}_{-14.8}$	$^{-4.2}_{+4.2}$	$^{+0.0}_{-0.1}$	$^{+0.4}_{-0.4}$	$^{+0.4}_{-0.4}$
9500	0.71	1	$1.50e-05$	$^{+158.0}_{-58.0}$	$^{+21.1}_{-19.8}$	$^{+18.9}_{-18.9}$	$^{-3.3}_{+4.5}$	$^{-0.4}_{+0.3}$	$^{+4.8}_{-4.8}$	$^{+1.3}_{-1.3}$

Not many events at high x

This uncertainty refers to how well we know the underlying cross section assuming that our only knowledge is the observed number of events.

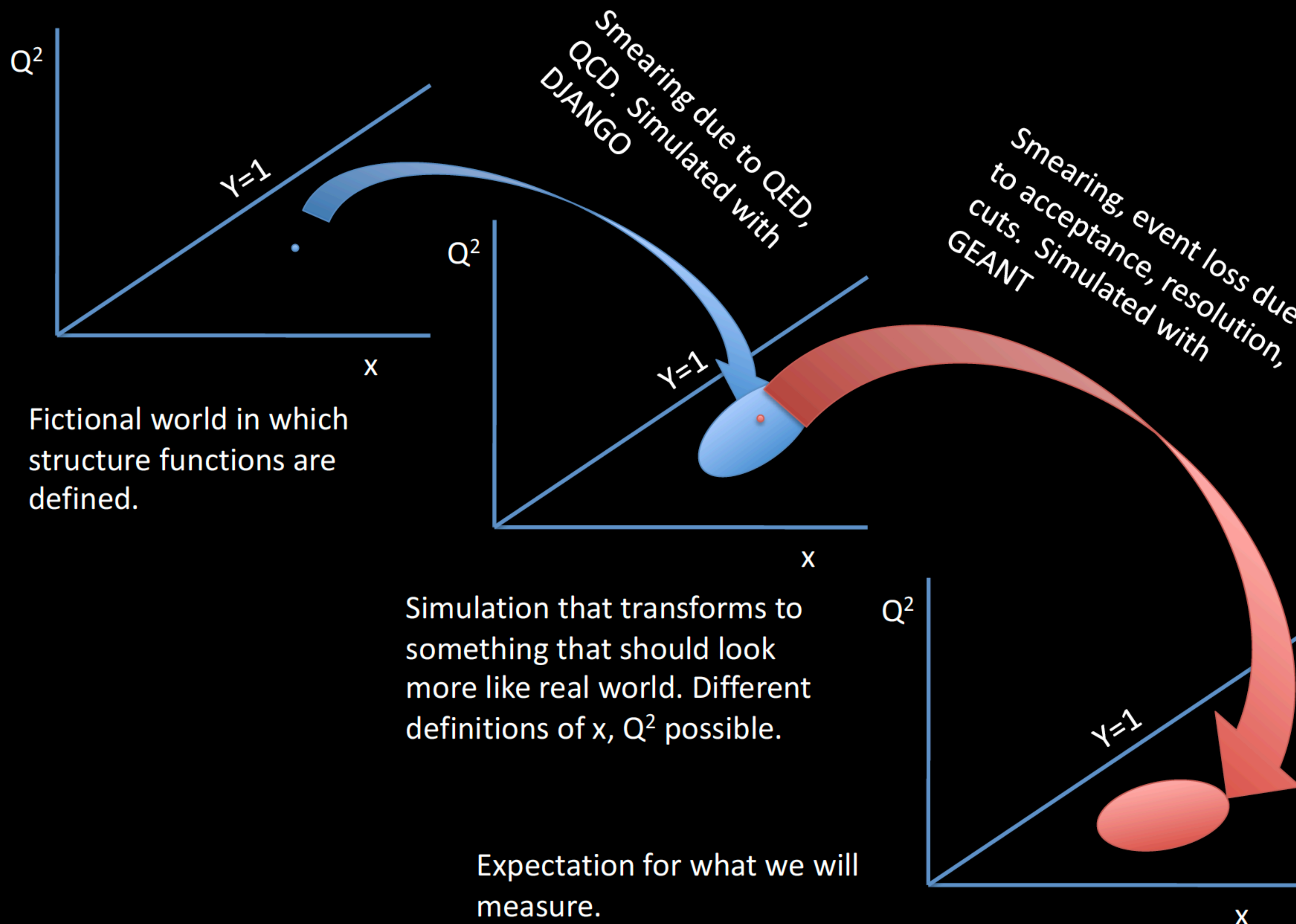
Not the uncertainty that belongs in a fit.

Study of proton parton distribution functions at high x using ZEUS data

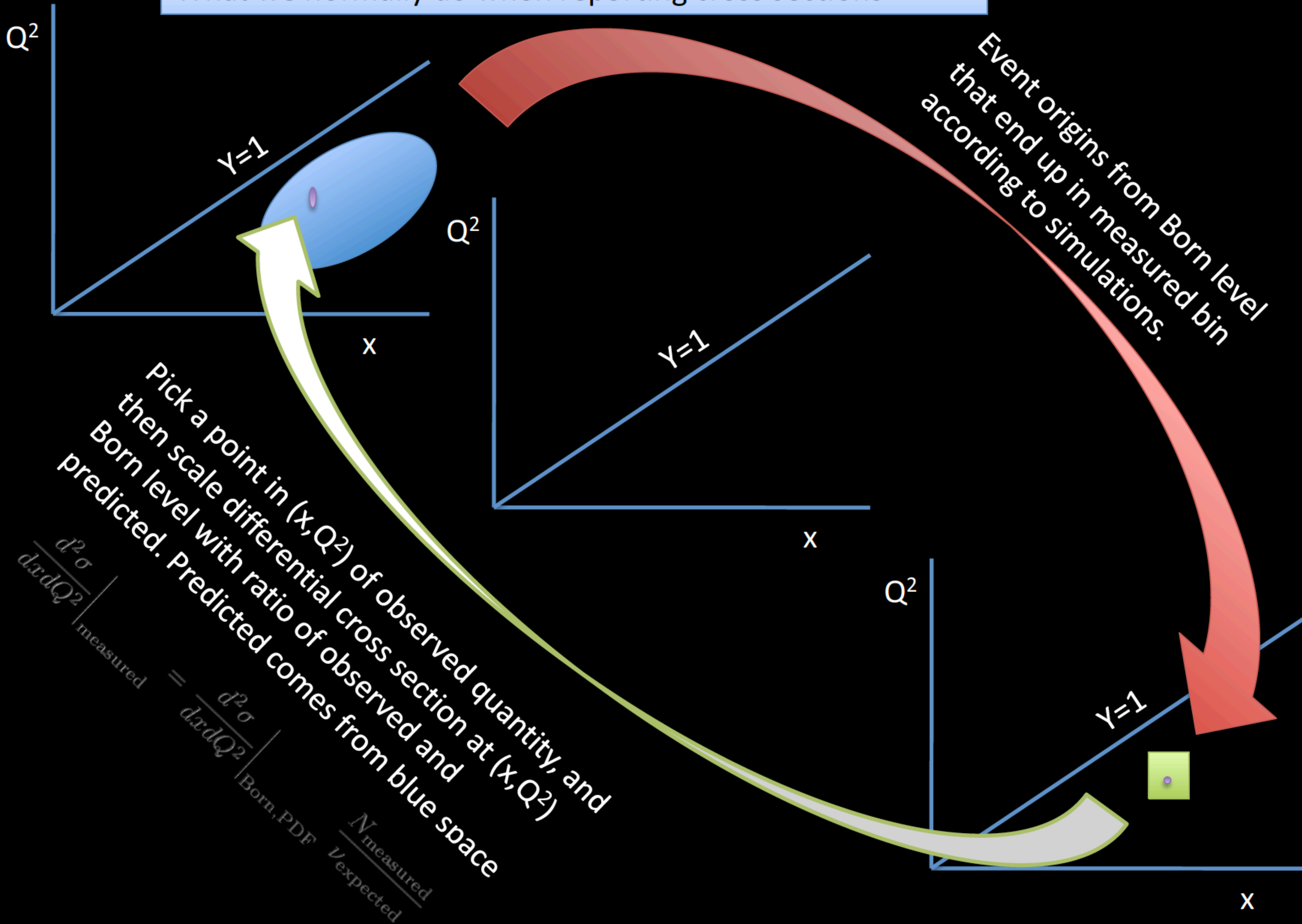
I. Abt *et al.* (ZEUS Collaboration)
Phys. Rev. D **101**, 112009 – Published 26 June 2020

Described how to use a forward modeling for analysis of the data:

Define pdfs -> apply radiative effects
-> predict cross sections
-> apply detector/analysis effects
-> calculate expected number of events
-> calculate a Poisson probability

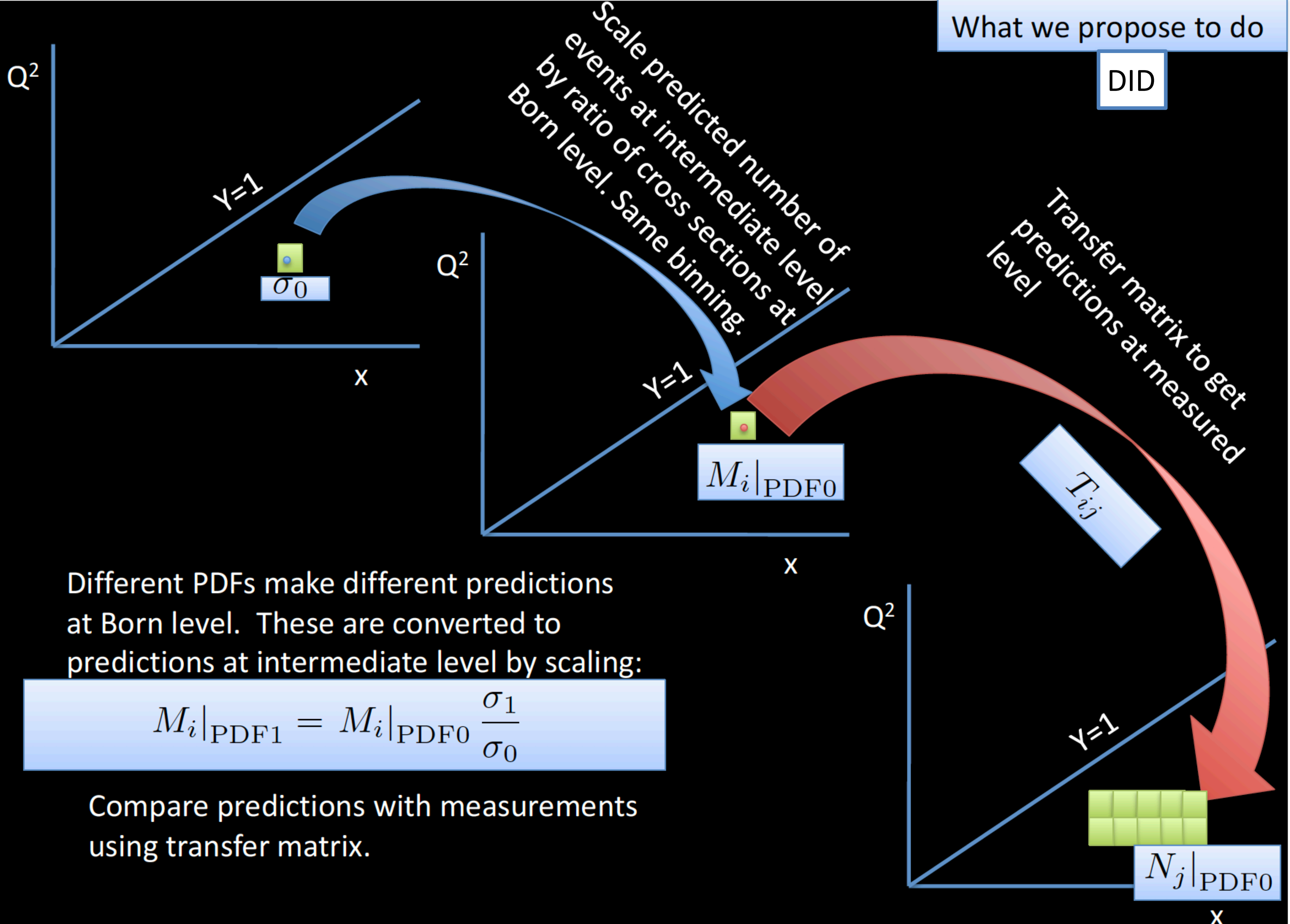

Q2
Y=1
x
Fictional world in which structure functions are defined.
Smearing due to QED, QCD. Simulated with DJANGO
Q2
Y=1
x
Simulation that transforms to something that should look more like real world. Different definitions of x, Q2 possible.
Smearing, event loss due to acceptance, resolution, cuts. Simulated with GEANT
Q2
Y=1
x
Expectation for what we will measure.

What we normally do when reporting cross sections



What we propose to do

DID

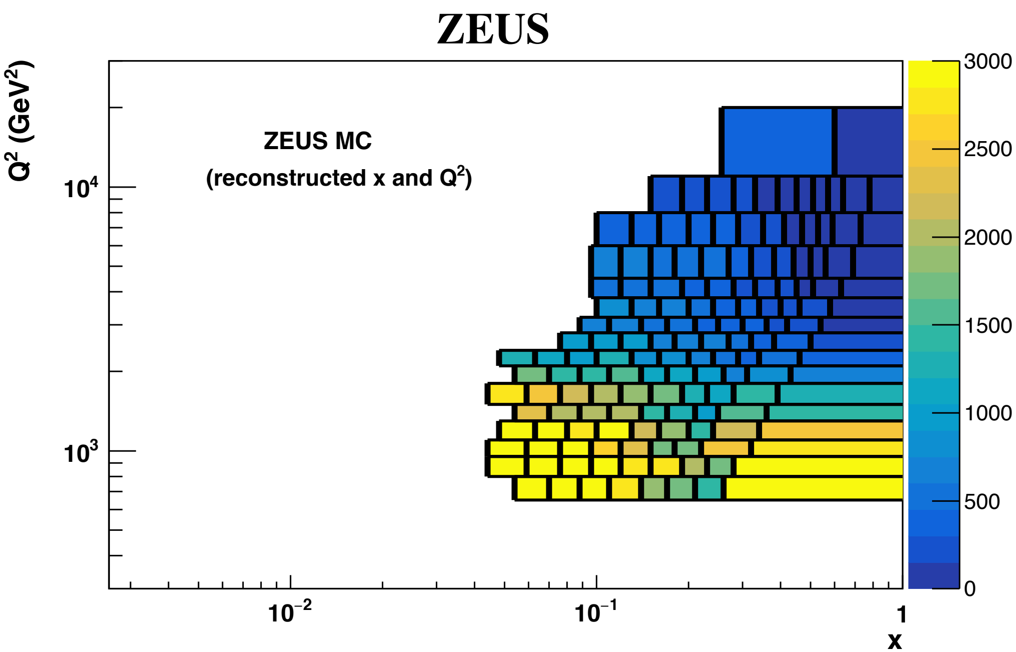


Different PDFs make different predictions at Born level. These are converted to predictions at intermediate level by scaling:

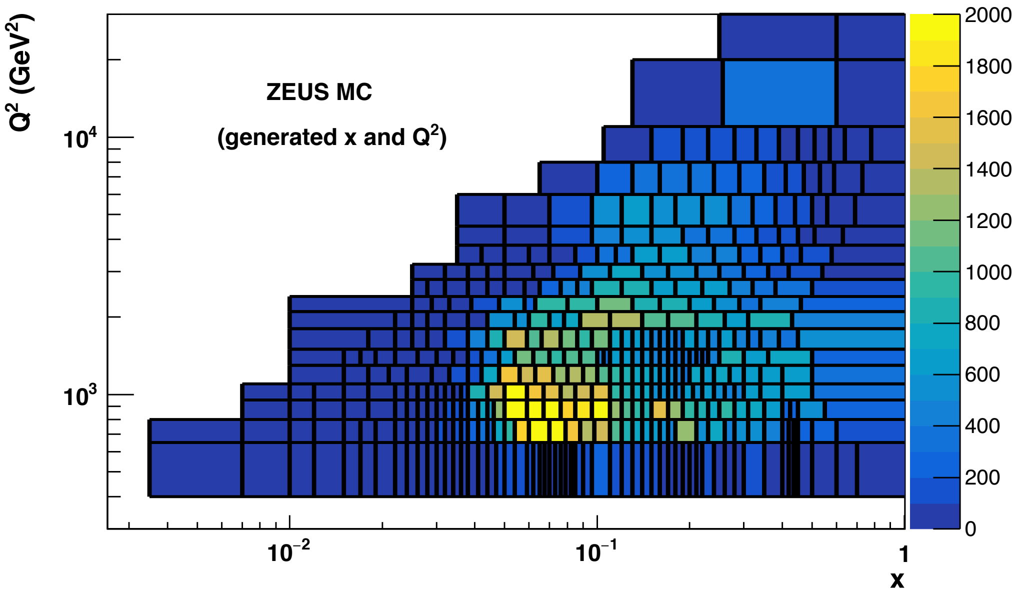
$$M_i|_{\mathrm{PDF1}} = M_i|_{\mathrm{PDF0}} \frac{\sigma_1}{\sigma_0}$$

Compare predictions with measurements using transfer matrix.

Transfer Matrices



Reconstructed variables/bins



generated variables/bins. Much finer binning used.

Separate transfer matrices exist for producing radiative cross sections and detector/analysis effects.

Studies of systematic effects

Categories investigated:

A. Uncertainties on expectation at generator level
   1. **normalization due to luminosity uncertainty**
   2. knowledge of radiative effects
   3. pdf set uncertainties

B. Uncertainties on transfer matrix
   4. binning effects (too crude ?)
   5. finite Monte Carlo statistics
   6. imperfect events & detector simulation

By far the dominant effect in terms of the probability of the data given a pdf set is the normalization.

Project: Use this approach/data in pdf extractions

Bayesian fit using the BAT.jl package

Team:

Ritu Aggarwal, SPPU, Pune

Performed analysis of ZEUS high-x data and extracted transfer matrices

Michiel Botje, **Nikhef (retired but very active!)**

Developer of QCDNUM, providing fast cross section integration routines, expert knowledge

Francesca Capel, Technical University, Munich

ORIGINS Data Science Lab postdoc, developing fitting code, providing expert advice on statistical methods

Oliver Schulz, Max Planck Institute for Physics

Developer of the BAT.jl package, statistics and coding expert, consulting

A.C. - coordination

BAT.jl

Bayesian Analysis Toolkit

docs stable | docs dev | license MIT | CI passing | codecov 59% | DOI 10.5281/zenodo.2587213

Welcome to BAT, a Bayesian analysis toolkit in Julia.

BAT.jl currently includes:

- Metropolis-Hastings MCMC sampling — and Hamiltonian MC, Nested Sampling
- Adaptive Harmonic Mean Integration (AHMI)
- Plotting recipes for MCMC samples and statistics

Additional sampling algorithms and other features are in preparation.

BAT.jl originated as a rewrite/redesign of BAT, the Bayesian Analysis Toolkit in C++. BAT.jl now offer a different set of functionality and a wider variety of algorithms than it's C++ predecessor.

Status/Goals

The project has started - approximately 6 meetings so far

QCDNUM Julia wrapper developed, can be called from BAT

Start with 5 flavors, FFN scheme, simple beta function parametrization

Dirichlet sampling distributions to satisfy sum rule constraints

Test comparison with xFitter of differential cross section calculations done

Integration routines nearing completion

Expect to start first fit attempts on order of weeks …

Status/Goals

We plan to develop the fitting code and provide an example of how the fitting of pdfs to the ZEUS data can be performed using forward modeling.

This can hopefully be integrated into existing fitting packages so the ZEUS high-x, high-Q2 data can be used in pdf extractions

Not planning to compete with existing pdf extraction teams

Summary

- There is very little data as x-> 1 and high virtuality
- ZEUS high-x data unique, but not used in PDF fits
- uncertainties in the PDFs appear to be underestimated, judged by comparing their predictions to each other (my personal conclusion)
- a transfer matrix formulation makes it possible to compare PDF set predictions to the ZEUS high-x data directly and calculate probabilities
- we are working to develop an example pdf fit using this approach