A New Generation of Parton Distributions with Uncertainties from Global QCD Analysis

What's new in this Global QCD Analysis of PDF's?

- New Data sets (common to all recent analyses)
- New methods and techniques of analysis: enable
  ⇒ * quantitative treatments of systematic errors; D. Stump
  * reliable calculation of the Hessian matrix J. Pumplin

New Results and physical applications

- New generation of CTEQ PDF's: eigenvector sets that characterize the behavior of overall χ^2 in the ⇒ neighborhood of the global minimum, hence allow the calculation of uncertainties of any variable dependent on parton distributions. (Available in the traditional and in the Les Houches universal interface form)
- e.g. Precision W/Z physics at the Tevatron/LHC
- some general results: Parton Luminosities at the Tevatron, LHC, RHIC, VLHC ⇒ predictions on X-sections and their uncertainties for Higgs-, top-productions, high p_T jets, ... etc.

Outlook

J. Pumplin, D.R. Stump, J. Huston, H.L. Lai,
P. Nadolsky, and W.K. Tung

Michigan State Univ. hep-ph/0201195

Related Work

S. I. Alekhin, Eur. Phys. J. **C 10** (1999) 395 [hep-ph/9611213]; hep-ex/0005042.

M. Botje, Eur. Phys. J. **C 14** (2000) 285 [hep-ph/9912439].

V. Barone, C. Pascaud and F. Zomer, Eur. Phys. J. **C 12** (2000) 243 [hep-ph/9907512]; C. Pascaud and F. Zomer, Tech. Note LAL-95-05.

A. D. Martin, R. G. Roberts, W. J. Stirling, R. S. Thorne, Eur.Phys.J. **C14** (2000) 133.

W. T. Giele and S. Keller, Phys. Rev. **D 58** (1998) 094023 [hep-ph/9803393]; W. T. Giele, S. A. Keller and D. A. Kosower, "Parton Distribution Function Uncertainties," [hep-ph/0104052].

J. Pumplin, D. R. Stump and W. K. Tung, Phys. Rev. **D 65** (2002) 014011 [hep-ph/0008191].

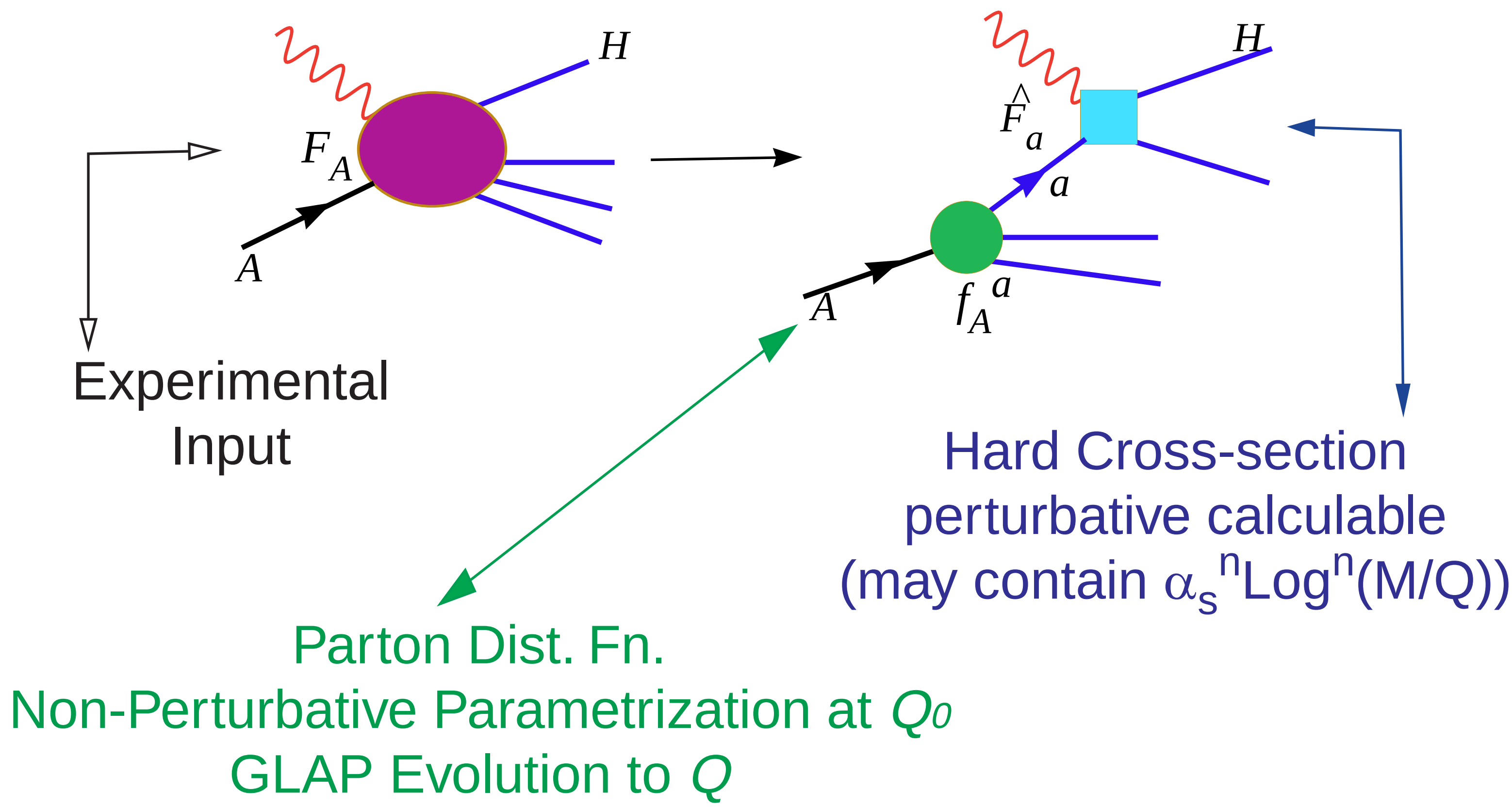
J. Pumplin *et al.*, Phys. Rev. **D 65** (2002) 014013 [hep-ph/0101032].

D. R. Stump *et al.*, Phys. Rev. **D 65** (2002) 014012 [hep-ph/0101051].

Global QCD Analysis in a Nutshell

Master Equation for QCD Parton Model
– the Factorization Theorem

$$F_A^\lambda\left(x,\frac{m}{Q},\frac{M}{Q}\right) = \sum_a f_A^a\left(x,\frac{m}{\mu}\right) \otimes \hat{F}_a^\lambda\left(x,\frac{Q}{\mu},\frac{M}{Q}\right) + \mathcal{O}\left(\left(\frac{\Lambda}{Q}\right)^2\right)$$



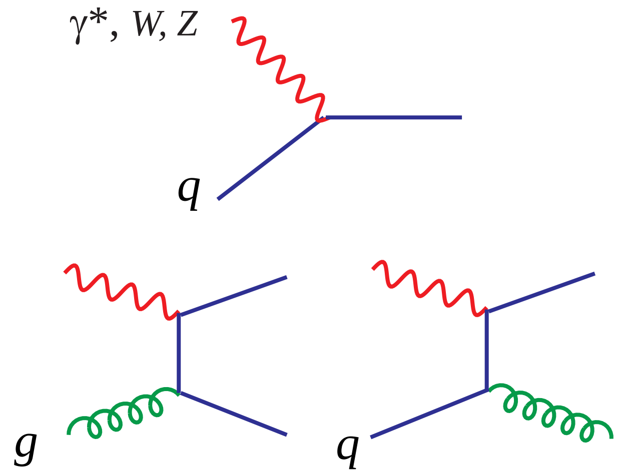
Sources of Uncertainties and Challenges:

- Experimental errors; (and uncertainties on errors!)
- Parametrization dependence;
- Higher-order corrections; Large Logarithms;
- Power-law (higher twist) corrections.

Experimental Input:
Phyical Processes & Experiments

DIS

e N
μ N
ν N
$\bar{\nu}$ N



SLAC
BCDMS
NMC, E665
H1, ZEUS

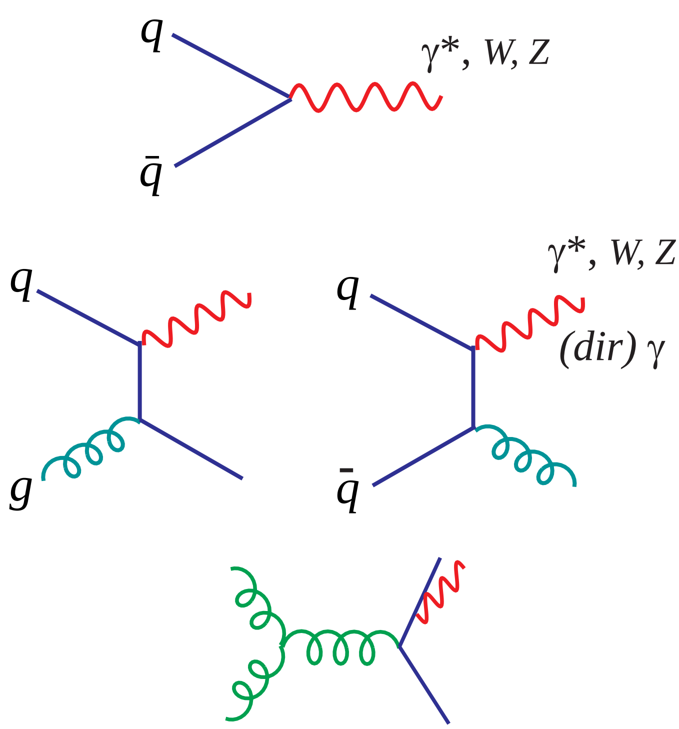
CDHS, CHARM
CCFR
CHORUS

DY

p N
π N
k N
$\bar{p}$ N

Dir.Ph.

p N
π N
k N
$\bar{p}$ N



E605, E772
NA51
E866

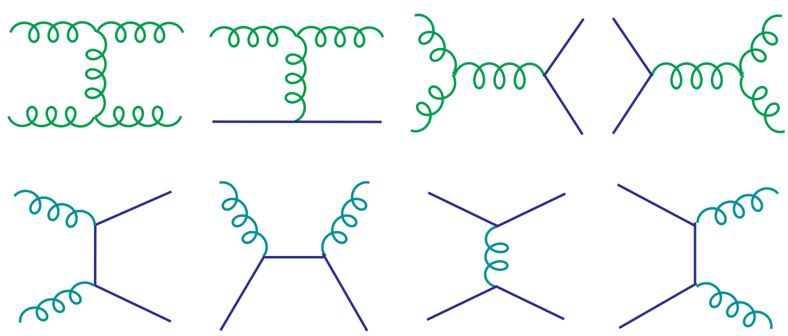
CDF, D0

WA70, UA6
E706

CDF, D0

Jet Inc.

$\bar{p}$ p



CDF, D0

Selection of Data

CTEQ5			CTEQ6		
	#	sys		#	sys
BCDMS μp	168	no	BCDMS μp	339	yes
BCDMS μd	156	no	BCDMS μd	251	yes
H1 ep	172	no	H1a ep •	104	yes
			H1b ep •	126	yes
ZEUS ep	186	no	ZEUS ep •	229	yes
NMC μp	104	no	NMC μp	201	yes
NMC $\mu p/\mu n$	123	no	NMC $\mu p/\mu n$	123	yes
CCFR F_2 νN	87	no	CCFR F_2 νN	159	yes
CCFR F_3 νN	87	no	CCFR F_3 νN	87	no
E605 pp DY	119	no	E605 pp	119	no
NA51 pd/pp DY	1	no	NA51 pd/pp	1	no
E866 pd/pp DY	15	no	E866 pd/pp	15	no
CDF W	11	no	CDF W	11	no
CDF jet	33	yes	CDF jet	33	yes
DØjet	24	yes	DØJet •	90	yes

• New Data

What's new in Fitting Procedure and Error Analysis?

- Comprehensive and efficient χ^2 minimization procedure, including correlated systematic errors;

- Deeper insight on the goodness-of-fit taking into account systematic errors;

What's new in Error Analysis?

The simplest χ^2 function is

$$\chi_0^2(a) = \sum_{\text{expt.}} \sum_{i=1}^{N_e} \frac{(D_i - T_i(\{a\}))^2}{\sigma_i^2}$$

where D_i is a data value, σ_i the experimental error, $T_i(\{a\})$ depend on the PDF model parameters $\{a\}$.

- Number of fitting parameters in the global analysis:
  - PDF parameters : ~ 20 (+ ~ 10 normalization factors)

When correlated systematic errors are included, must use a more general χ^2 function

$$\chi'^2(\{a\},\{r\}) = \sum_{\text{expt.}} \left[\sum_{i=1}^{N_e} \frac{1}{\alpha_i^2} \left(D_i - T_i(\{a\}) - \sum_{k=1}^{K} r_k \beta_{ki} \right)^2 + \sum_{k=1}^{K} r_k^2 \right]$$

where $\alpha_i^2 = \sigma_i^2 + u_i^2$ is the combined uncorrelated error; $\{\beta_{1i}, \ldots, \beta_{Ki}\}$ are K sources of correlated systematic errors; and r_k is the relative shift of the systematic error k.

Problem: (particularly for global analyses)

- Number of fitting parameters in the global analysis increases by # of sys. err. : (5 – 20) x 10 !!
- Fitting process, particularly uncertainty assessment of PDF parameters, become uncontrollable.

Solution:

- Minimization w.r.t. $\{r_k\}$ can be done *analytically*!

$$\hat{r}_k(\{a\}) = \sum_{k'=1}^{K} \left(A^{-1}\right)_{kk'} B_{k'}(\{a\})$$

where

$$B_k(\{a\}) = \sum_{i=1}^{N_e} \frac{\beta_{ki}\,(D_i - T_i(\{a\})}{\alpha_i^2} \quad \text{and} \quad A_{kk'} = \delta_{kk'} + \sum_{i=1}^{N_e} \frac{\beta_{ki}\beta_{k'i}}{\alpha_i^2}$$

- Now the numerical minimization *involves just the small # of (physical) fitting parameters* {a}, but w.r.t. the generalized χ^2 function

$$\chi^2(\{a\}) \equiv \chi'^2(\{a\}, \{\hat{r}(\{a\})\})$$

$$= \sum_{\text{expt.}} \left\{ \sum_{i=1}^{N_e} \frac{(D_i - T_i)^2}{\alpha_i^2} - \sum_{k,k'=1}^{K} B_{ki}(a) \left(A^{-1}\right)_{kk'} B_{k'i}(a) \right\}$$

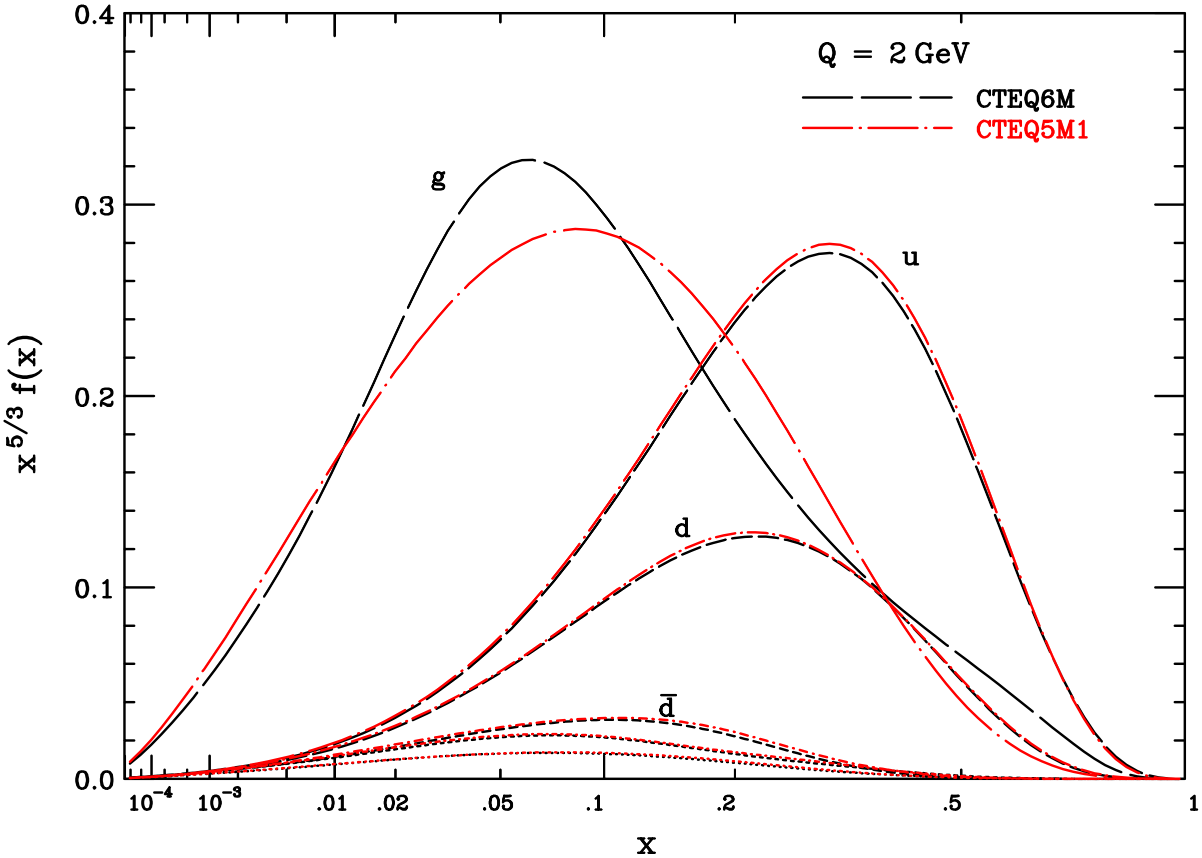
Bonuses:

- Gain more insight on the goodness-of-fit by examining the values of $\{r_k\}$ and their distribution.
- Get better feel on the goodness-of-fit in data/theory comparison plots, include the effects of systematic shifts $\{r_k\}$:

Compare $(\hat{D}_i \equiv D_i - \sum_{k=1}^{K} \hat{r}_k \beta_{ki})$ with $T_i(\hat{a})$, then

$$\chi^2 = \sum_{\text{expt.}} \left[\sum_{i=1}^{N_e} \frac{\left(\hat{D}_i - T_i(a)\right)^2}{\alpha_i^2} + \sum_{k=1}^{K} \hat{r}_k^2 \right]$$

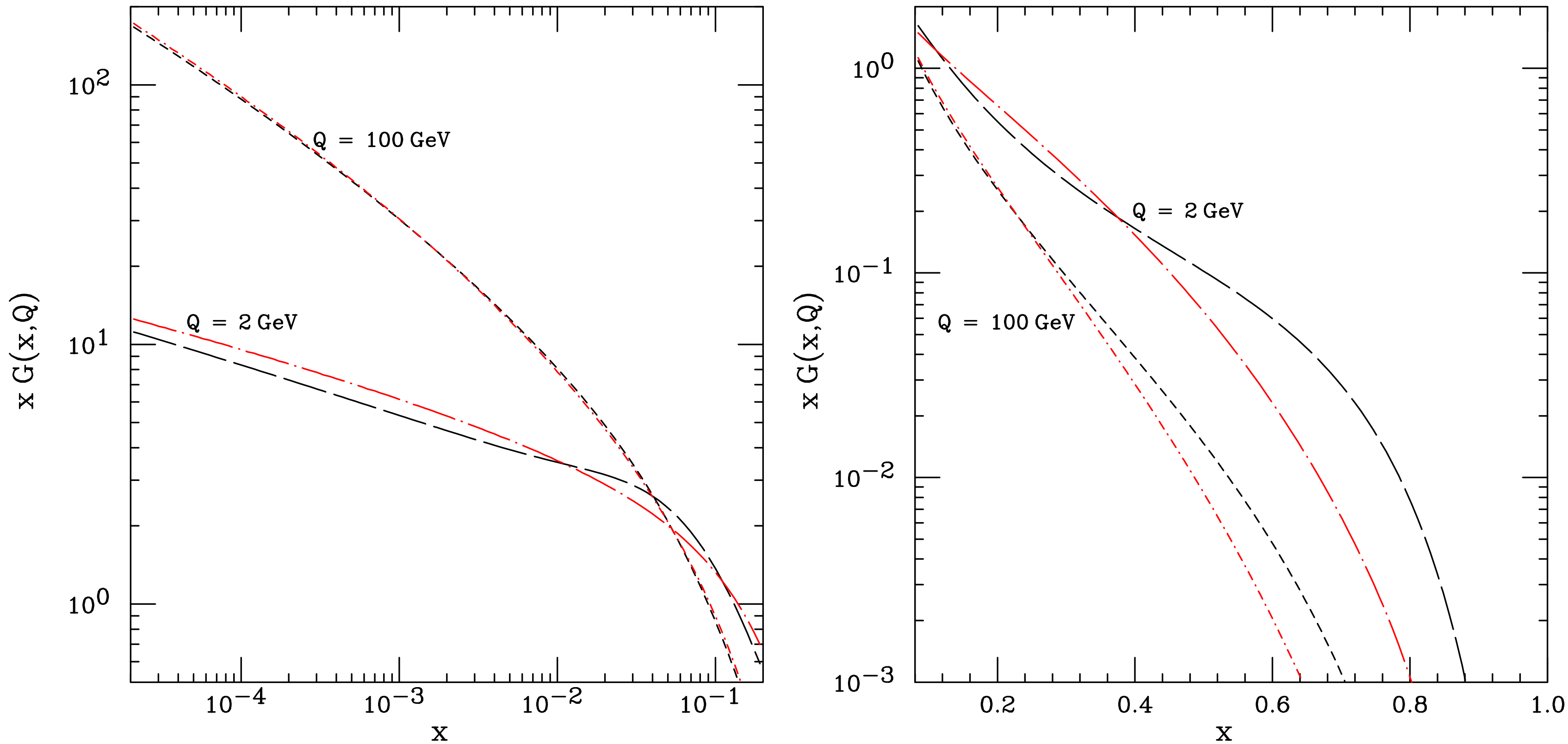
- Vis-à-vis the covariance matrix approach: avoid inverting *NxN* matrices that can be numerically unstable.



Comparison of CTEQ6M (dashed) to CTEQ5M1 (dot-dashed) PDF's at $Q = 2$ GeV. (The unlabeled curves are $\bar{u}$ and $s = \bar{s}$.)

★ Quarks have not changed much.

★ Gluon is noticeably different.



Comparison of CTEQ6M (dashed) to CTEQ5M1 (dot-dashed) gluon distributions at $Q = 2$ and 100 GeV. (a) The small-x region; (b) the large-x region.

THE GLUON DISTRIBUTION

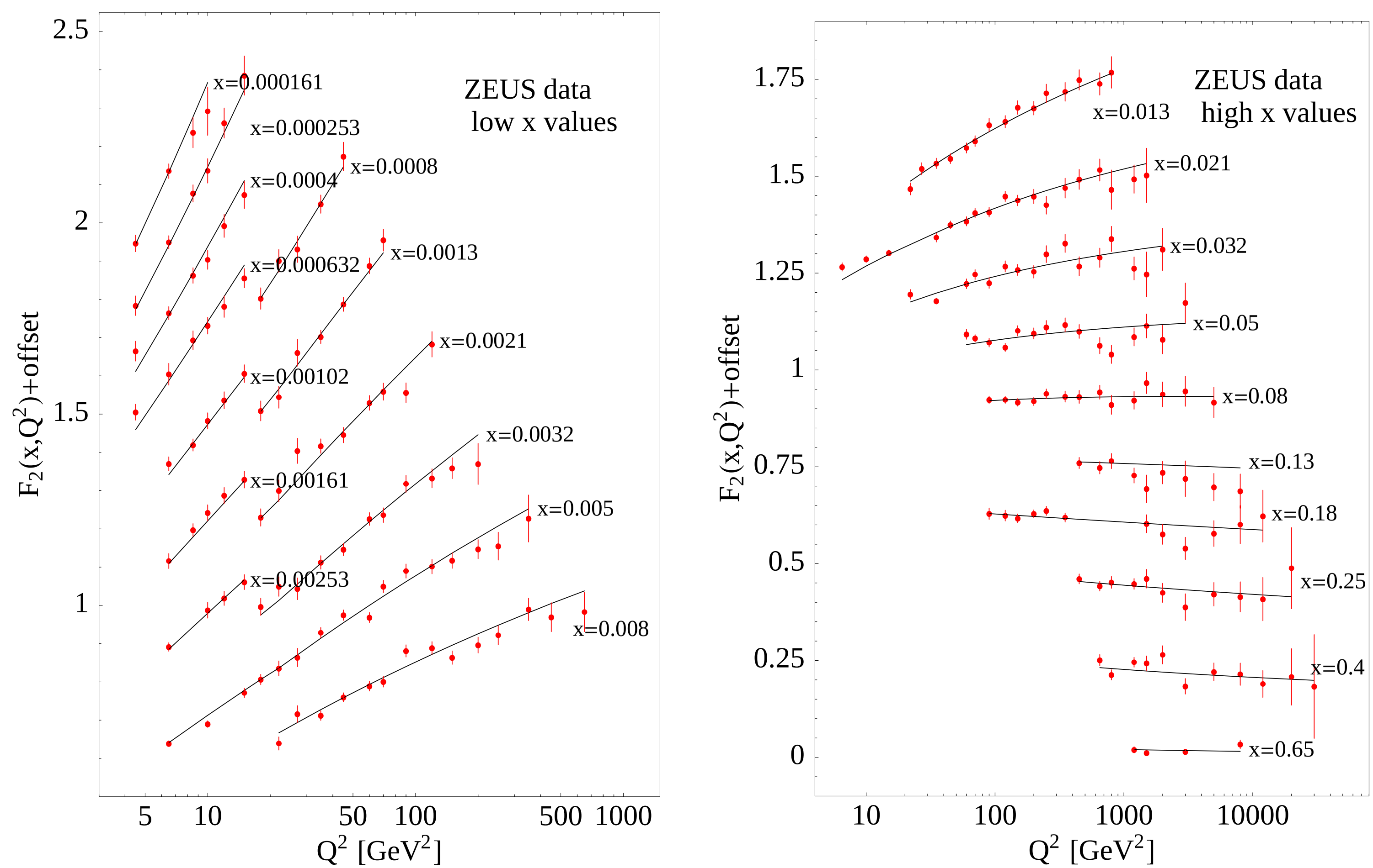
Comparison to Data

Comparison of the CTEQ6M fit to data with correlated systematic errors.

data set	N_e	χ_e^2	χ_e^2/N_e
BCDMS p	339	377.6	1.114
BCDMS d	251	279.7	1.114
H1a	104	98.59	0.948
H1b	126	129.1	1.024
ZEUS	229	262.6	1.147
NMC F2p	201	304.9	1.517
NMC F2d/p	123	111.8	0.909
DØ jet	90	64.86	0.721
CDF jet	33	48.57	1.472

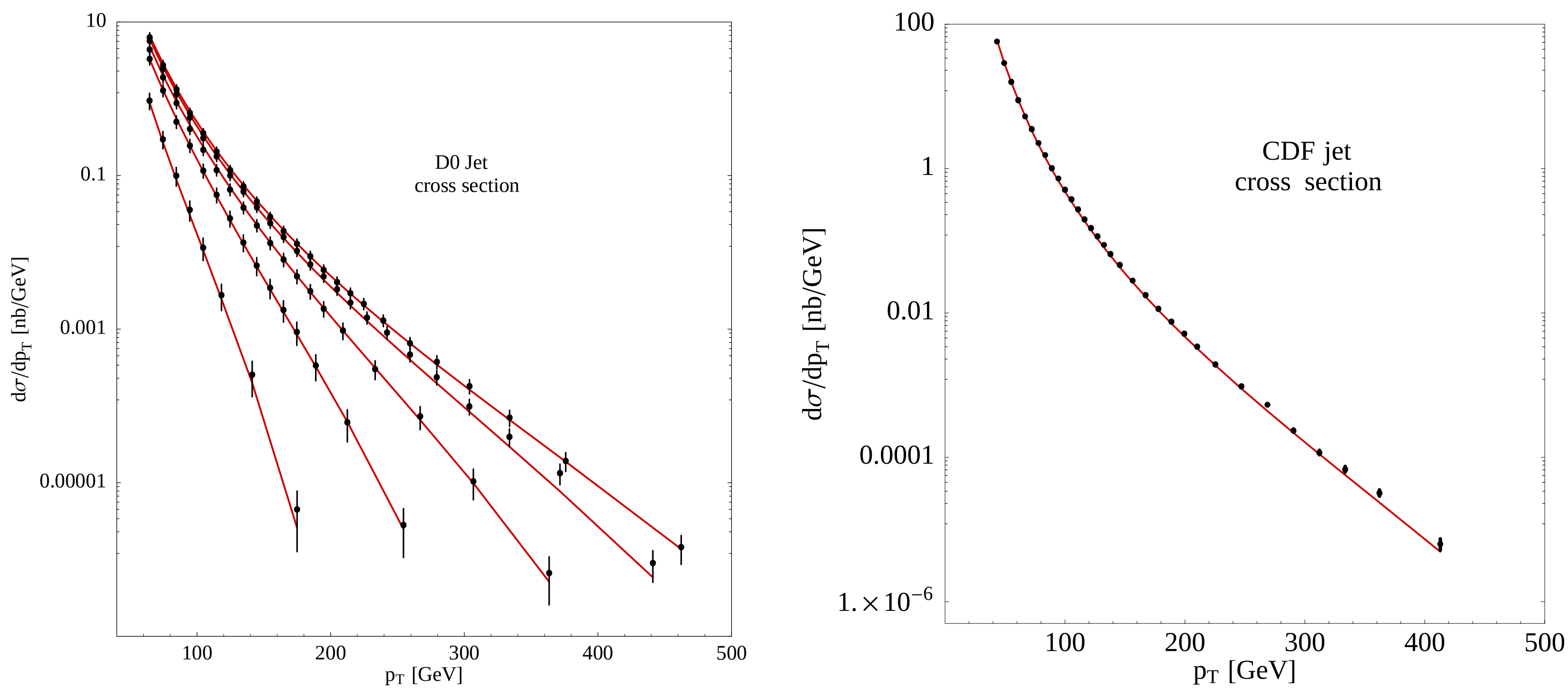
Other data sets:

CCFR	ν DIS	(150/156)
E605	Drell-Yan	(95/119)
E866	Drell-Yan	(6/15)
CDF	W-lepton asymmetry	(10/11)



Comparison of the CTEQ6M fit to the ZEUS data in separate x bins.* The data points include the estimated corrections for systematic errors. The error bars are statistical errors only.

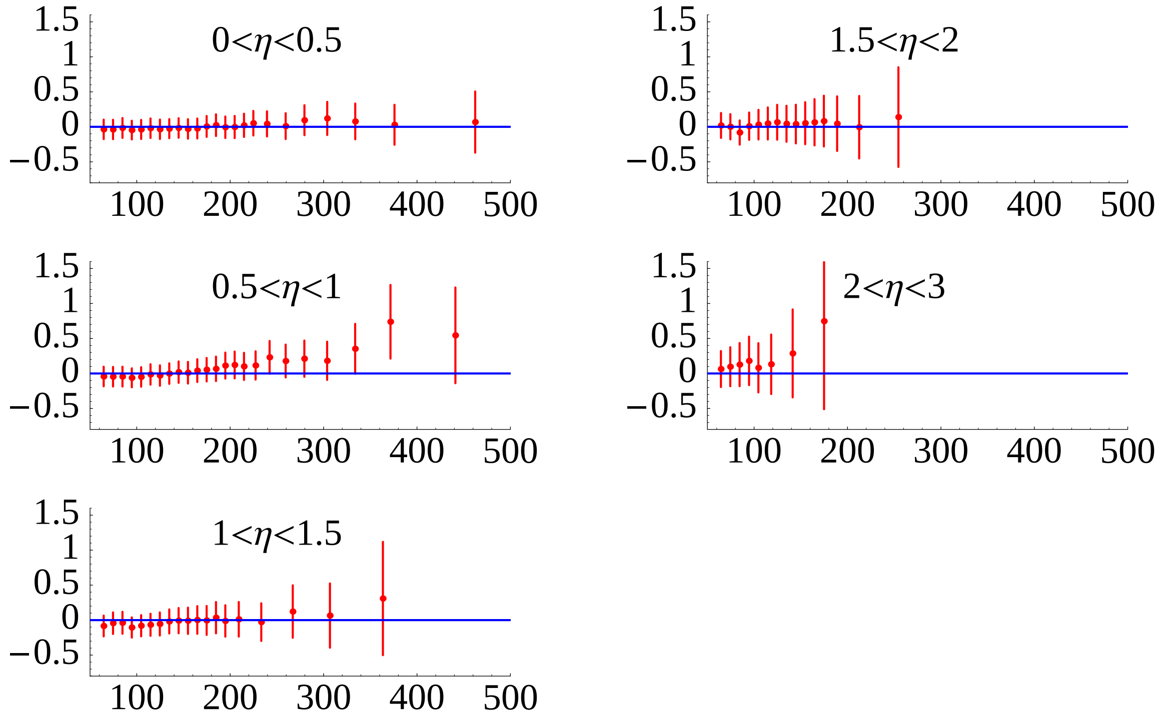
*ZEUS Collaboration: S. Chekanov *et al.*, Eur. Phys. J. **C 21** (2001) 443 [hep-ex/0105090].



Comparison of the CTEQ6M fit to the inclusive jet data.* (a) DØ (the boundary values of the 5 rapidity bins are $0, 0.5, 1.0, 1.5$, 2.0 and 3.0). (b) CDF (central rapidity, $0.1 < |\eta| < 0.7$).

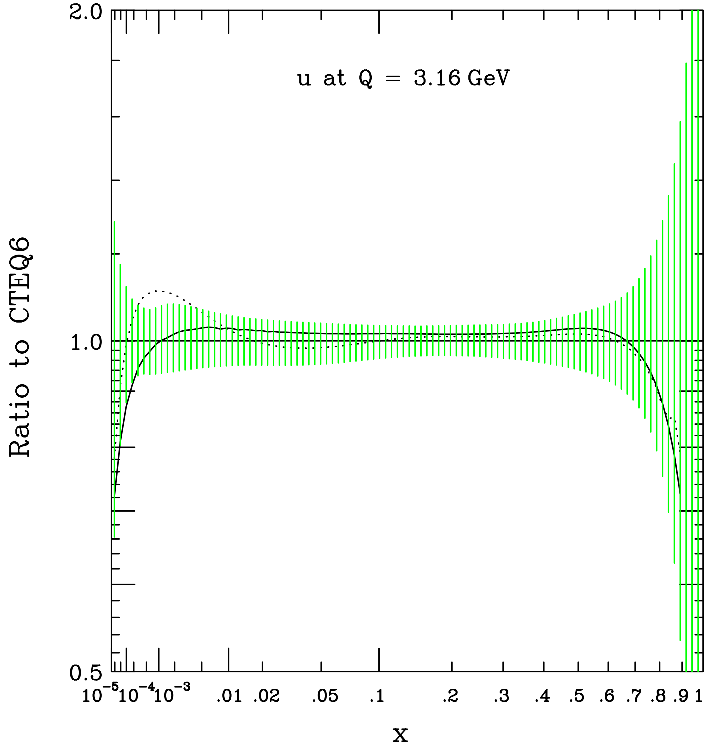
*DØ Collaboration: B. Abbott *et al.*; CDF Collaboration: T. Affolder *et al.*.

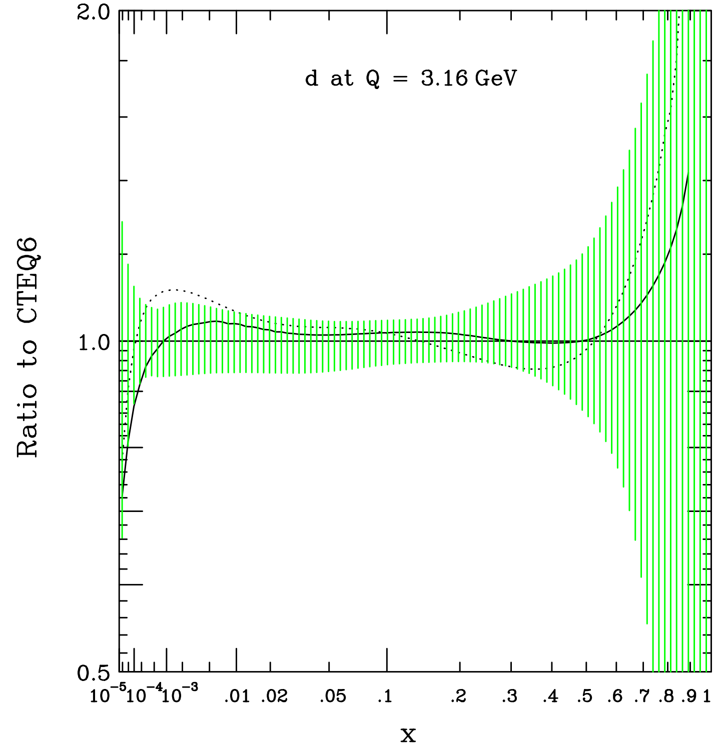
(data−theory)/theory versus p_T [GeV]



Closer comparison between CTEQ6M and the DØ jet data as fractional differences.

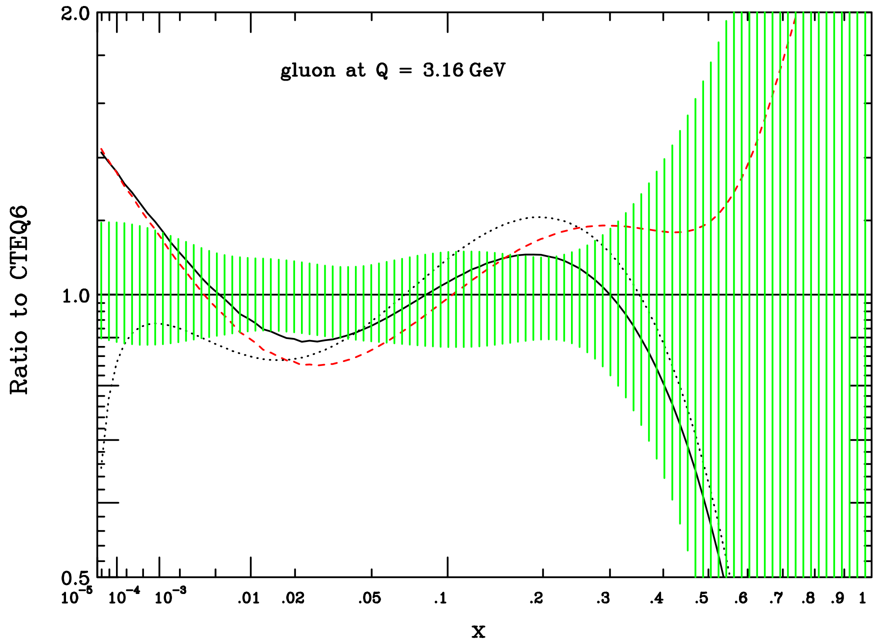
The Tevatron inclusive jet cross section implies a hard gluon: $g(x)$ is large at large x.
Recall CTEQ4HJ and CTEQ5HJ.





Uncertainty bands for the u- and d-quark distribution functions at $Q^2 = 10\,\mathrm{GeV}^2$. The solid line is CTEQ5M1 and the dotted line is MRST2001.

(Shaded: *envelopes* of extreme pdf's)



Uncertainty band for the gluon distribution function at $Q^2 = 10\,\text{GeV}^2$. The curves correspond to CTEQ5M1(solid), CTEQ5HJ (dashed), and MRST2001 (dotted).

The gluon is very uncertain for $x \gtrsim 0.4$.
[$g(x) \to 0$ as $x \to 1.$]

What are the *true* uncertainties of pdf parameters?

With Ideal Experiments and Perfect Theory Model:

- of course !! $\Delta\chi^2_{global} = 1$!! however,

Real life is never ideal and perfect.

High Road:

Hold your principles, stick to textbook recipes;
but then ...

Low Road:

Admit human imperfections
— make pragmatic, necessary compromises—
so do your best and see if sensible results emerge.

(My God! $\Delta\chi^2_{global} = 50, 100$?? !!)

(for N = 2000.)

Are the two approaches really different?

Not really—when all is said and done they actually
lead to the same conclusions!

... a little more details

What are the *true* uncertainties of pdf parameters?

With Ideal Experiments and Perfect Theory Model:

- of course !! $\Delta\chi^2_{global} = 1$!! however,

Reality #1: Real experiments are not ideal

- Many experimental results are *individually* "improbable" if errors are taken literally, i.e. $|\chi_e^2 / N_e - 1| \gg 1/\sqrt{2N_e}$ (e.g. NMC, CDF, ...)
- More than one experiment may be, strictly speaking (i.e., using the $\Delta\chi^2_e = 1$ criterion) statistically *incompatible*.

Reality #2: Theoretical Uncertainties differ widely between different processes, and are hard to quantify.

Idealistic approach: (Unique in principle, but ...)

Restrict to a few acceptable and compatible experiments, and apply textbook treatment. (Which expts to use?)

(In practice, the spread of predictions with different choices of experimental data sets becomes equivalent to below.)

Realistic approach: (not unique in principle, but ...)

Treat all experiments (with no known problems) on the same footing, and come up with more pragmatic treatment of error estimates.

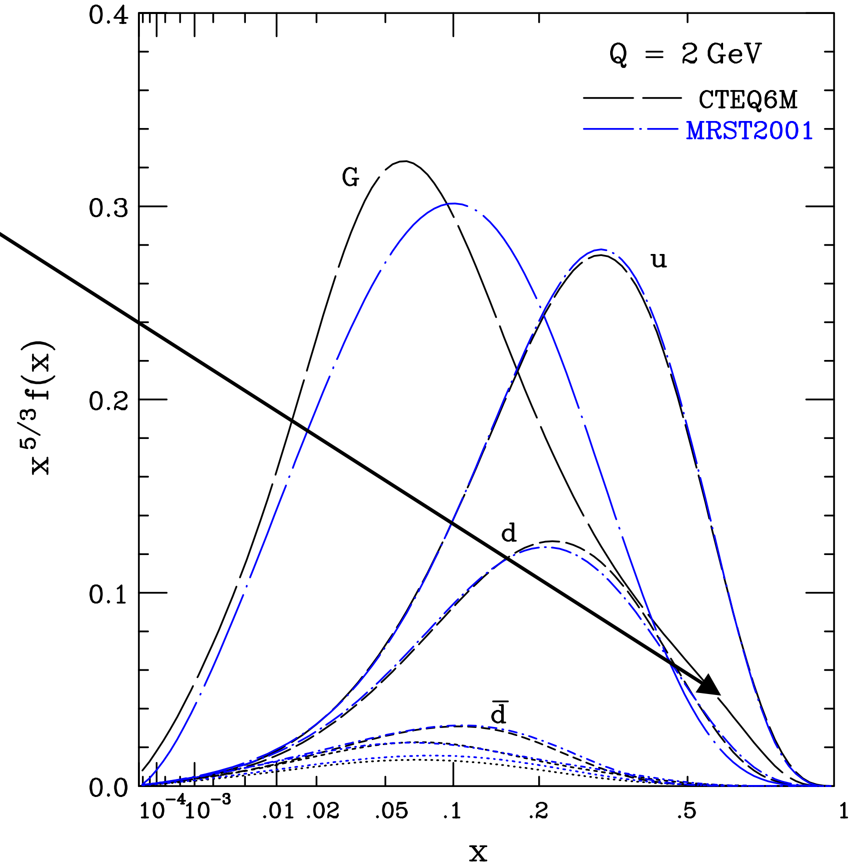
Adopt the ansatz that, *unless otherwise demonstrated*, all relevant experiments are acceptable and compatible.

$\Rightarrow$ Use relative (rather than absolute) probabilities; and assess uncertainties (tolerance) by examining the spread of certain reasonable measures. (*cf. the MC approach*)

Are the two approaches really different? No.

Compare CTEQ6 to MRST

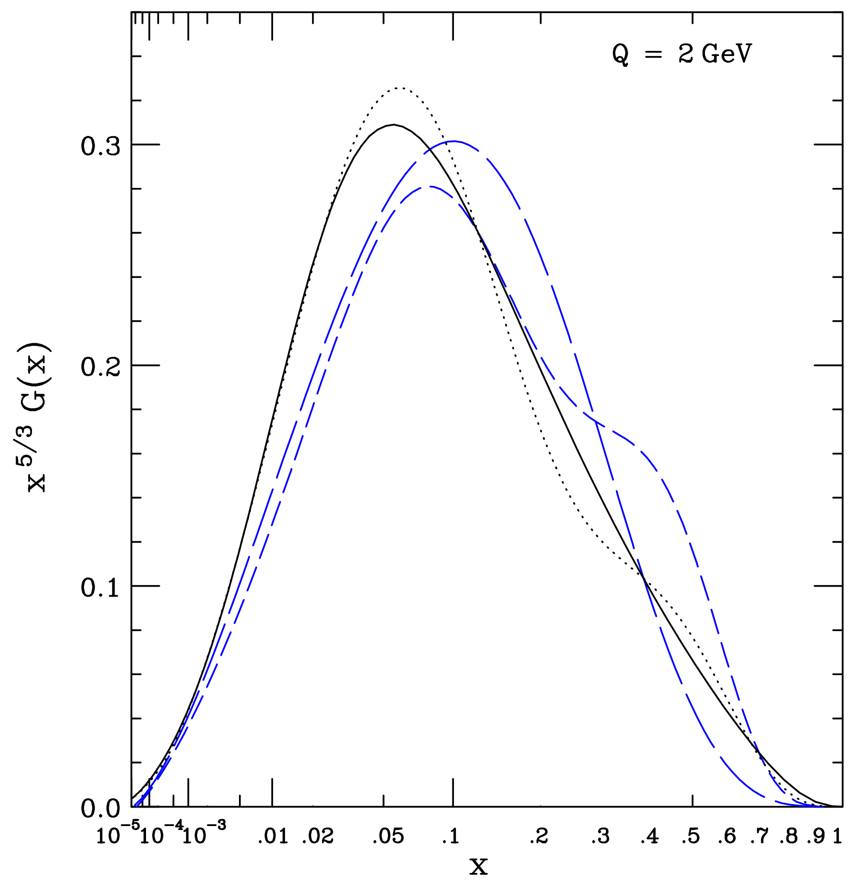
- Main difference is the gluon at high x



J. Huston

Compare CTEQ6 to MRST

- Solid: CTEQ6M
- Long-dashed: MRST2001
- Short-dashed: MRST2001J
- Dotted: MRST-like fit

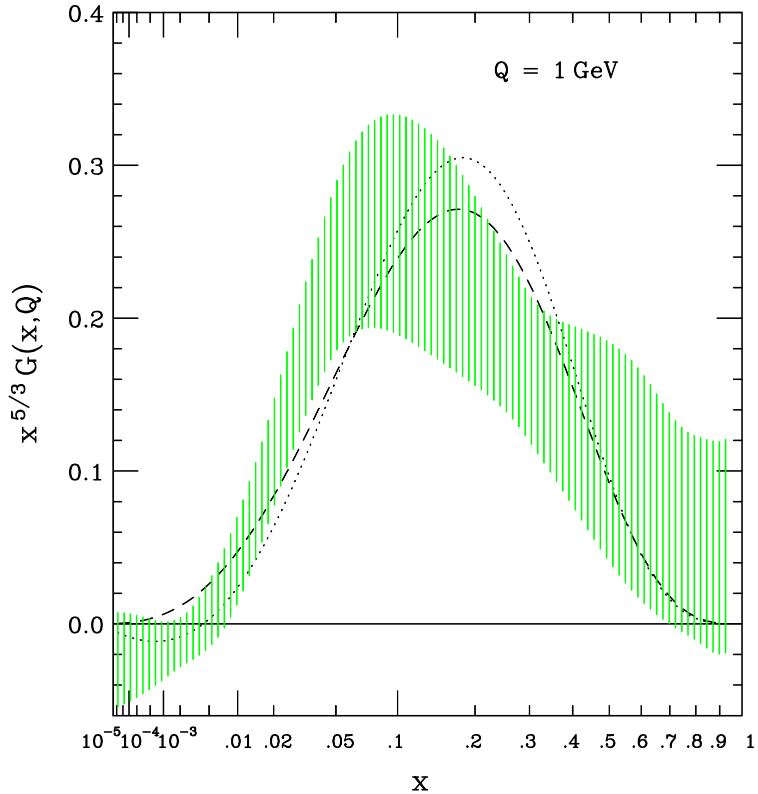


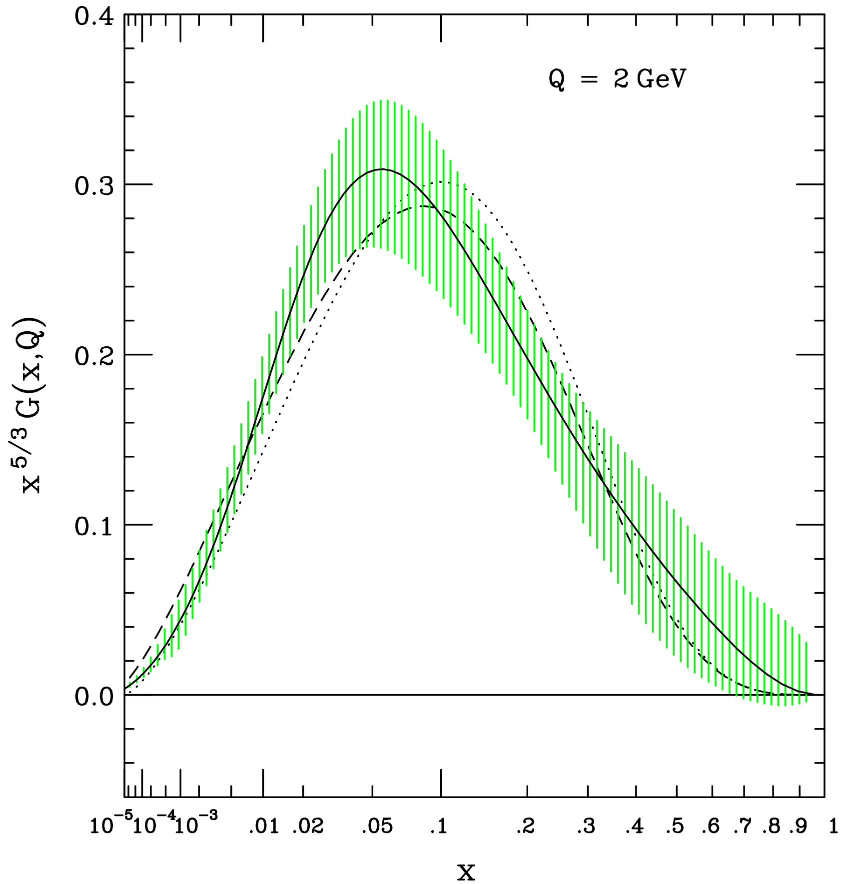
J. Huston

Uncertainty in gluon at small Q

- Note gluon can be negative at small x and **large x**

Evolves to positive by ~1.3 GeV



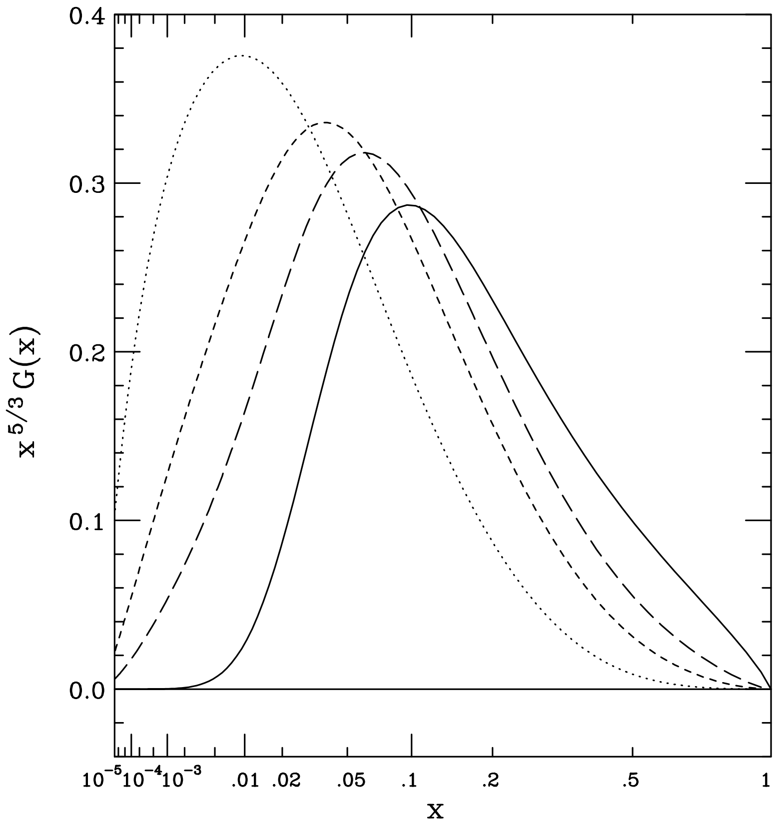


Dashed: CTEQ5M
Dotted: MRST2001
Solid: CTEQ6M

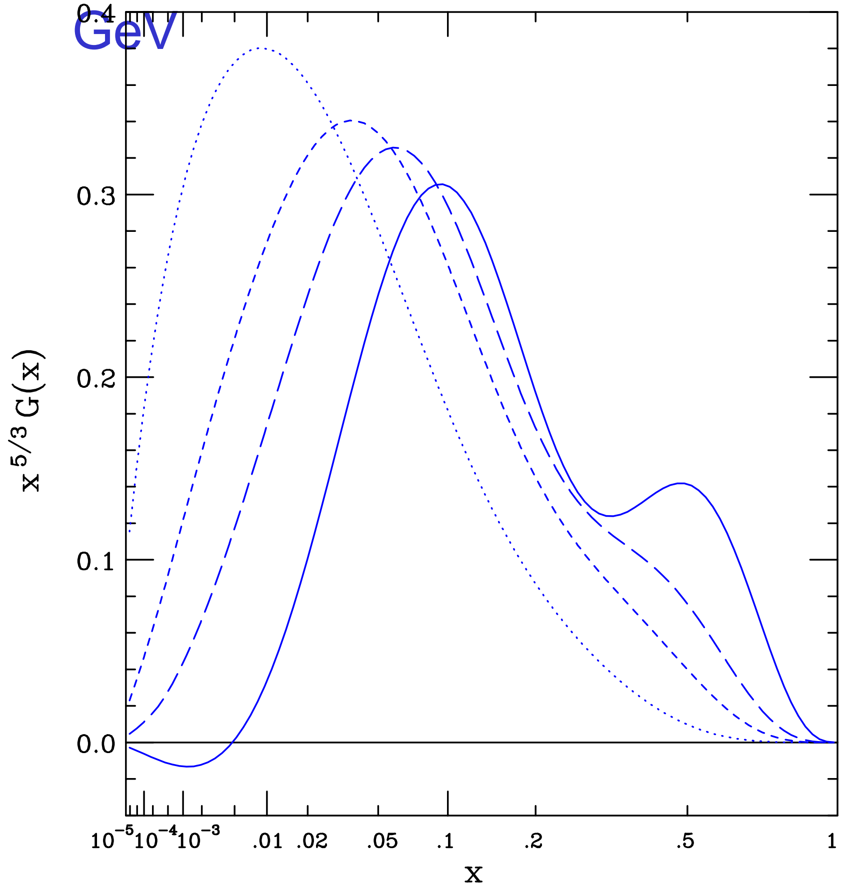
J. Huston

Gluon evolution

- CTEQ6M-like gluon at Q=1,2,5,100 GeV



MRST-like gluon at Q=1,2,5,100 GeV



J. Huston

Outlook

This is only the very beginning of studying uncertainties in global QCD analysis in a quantitative manner

- This work demonstrates that the new techniques for global analysis developed recently are viable and practical.
- The new results are very useful for the physics programs of the Tevatron, Hera, and LHC,
- There is a lot of room for collaboration among theorists and experimentalists

Many other sources of uncertainties in the overall global analysis have not yet been incorporated:

- Theoretical uncertainties due to higher-order PQCD corrections and resummation;
- Uncertainties introduced by the choice of parametrization have been explored extensively, but not yet quantitatively formulated.

Heavy quark effects and charm production data in NC and CC experiments will be systematically analyzed $\Longrightarrow$
More quantitative information on strange, charm, bottom distributions.

Continued progress in this venture is of vital importance for our understanding of *the parton structure of hadrons* (fundamental physics of its own right), for *precision SM physics studies* at future colliders, and for *New Physics searches*.