Parton distribution functions PDFs

YETI06

A.M.Cooper-Sarkar

Oxford

1. What are they?
2. How do we determine them?
3. What are the uncertainties -experimental

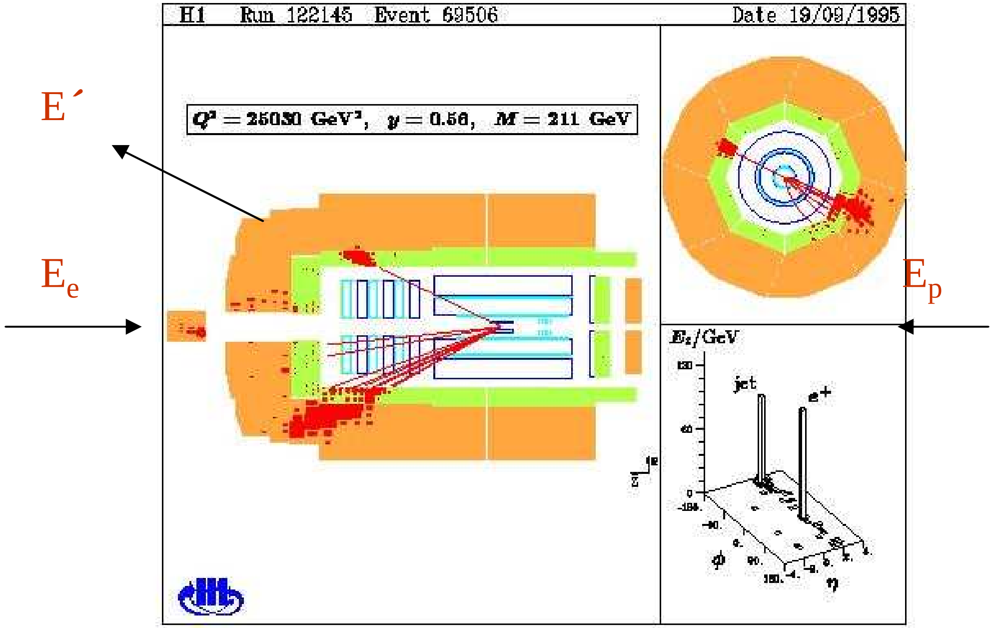
   -model

   -theoretical (see also Thorne's talk)

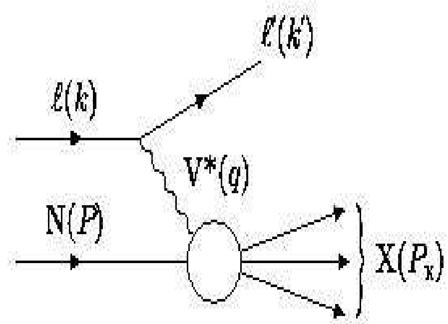
4. Why are they important?

PDFs were first investigated in deep inelastic lepton-hadron scattering -DIS

Candidate from NC sample



dσ ~



2 $L^{\mu\nu} W_{\mu\nu}$

Leptonic tensor - calculable

Hadronic tensor- constrained by Lorentz invariance

$q = k - k'$, $Q^2 = -q^2$

$P_x = p + q$, $W^2 = (p + q)^2$

$s = (p + k)^2$

$x = Q^2 / (2p.q)$

$y = (p.q)/(p.k)$

$W^2 = Q^2 (1/x - 1)$

$Q^2 = s\,x\,y$

$s = 4 E_e E_p$
$Q^2 = 4 E_e E' \sin^2\theta_e/2$
$y = (1 - E'/E_e \cos^2\theta_e/2)$
$x = Q^2/sy$

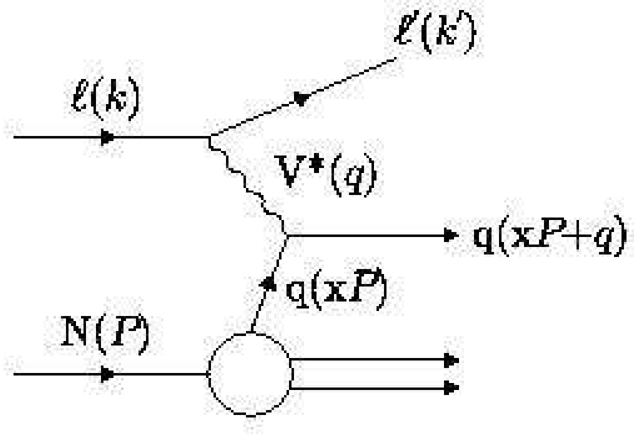
The kinematic variables are measurable

Completely generally the double differential cross-section for e-N scattering

$$\frac{d^2\sigma(e\pm N)}{dxdy} = \frac{2\pi\alpha^2 s}{Q^4}\left[\, Y_+ \mathbf{F_2}(x,Q^2) - y^2\, \mathbf{F_L}(x,Q^2) \pm Y\_x\mathbf{F_3}(x,Q^2)\right], \quad Y\pm = 1 \pm (1-y)^2$$

Leptonic part

hadronic part



F_2, F_L and xF_3 are structure functions

which express the dependence of the cross-section on the structure of the nucleon–

The Quark-Parton model interprets these structure functions as related to the momentum distributions of quarks or partons within the nucleon AND the measurable kinematic variable $x = Q2/(2p.q)$ is interpreted as the FRACTIONAL momentum of the incoming nucleon taken by the struck quark

$$(xP+q)^2=x^2p^2+q^2+2xp.q \sim 0$$

for massless quarks and $p^2 \sim 0$

so

$$x = Q^2/(2p.q)$$

The FRACTIONAL momentum of the incoming nucleon taken by the struck quark is the MEASURABLE quantity x

e.g. for charged lepton beams

F2$(x,Q^2) = \Sigma_i e_i^2(xq(x) + xq(x))$ – Bjorken scaling

FL(x,Q2) = 0 - spin ½ quarks

xF3(x,Q2) = 0 - only γ exchange

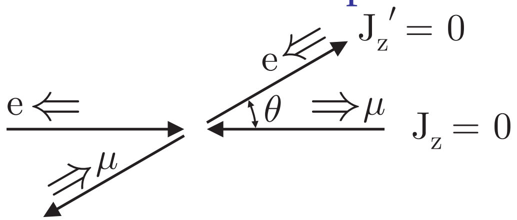
However for **neutrino beams**

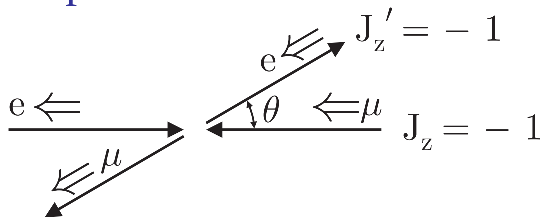
xF3(x,Q2)= Σ_i (xq(x) - xq(x)) ~ valence quark distributions of various flavours

Consider electron muon scattering

$$\frac{d\sigma}{dy} = \frac{2\pi\alpha^2}{Q^4} s\,[\,1 + (1-y)^2]\,, \text{ for elastic } e\mu$$

isotropic (1) non-isotropic ($(1-y)^2$)





$$\frac{d\sigma}{dy} = \frac{2\pi\alpha^2}{Q^4} e_i^2\, s\,[\,1 + (1-y)^2]\,, \text{ for elastic eq, quark charge } e_i\, e$$

$$\frac{d^2\sigma}{dxdy} = \frac{2\pi\alpha^2}{Q^4} s\,[\,1 + (1-y)^2]\,\Sigma_i\, e_i^2(xq(x) + x\bar{q}(x))$$

for eN, where eq has c. of m. energy^2 equal to xs, and q(x) gives probability that such a quark is in the Nucleon

Now compare the general equation to the QPM prediction to obtain the results

F_2 $(x,Q2) = \Sigma_i\, e_i^2(xq(x) + x\bar{q}(x))$ – Bjorken scaling

F_L $(x,Q2) = 0$ - spin ½ quarks

xF_3 $(x,Q2) = 0$ - only γ exchange

Consider $\nu,\bar{\nu}$ scattering: neutrinos are handed

$$\frac{d\sigma(\nu)}{dy} = \frac{G_F^2 \, x \, s}{\pi} \qquad \frac{d\sigma(\bar{\nu})}{dy} = \frac{G_F^2 \, x \, s}{\pi}(1-y)^2$$

For ν q (left-left) | For $\bar{\nu}$ q (left-right)

$$\frac{d^2\sigma(\nu)}{dxdy} = \frac{G_F^2 \, s}{\pi} \Sigma_i \, [xq_i(x) + (1-y)^2 x\bar{q}_i(x)]$$

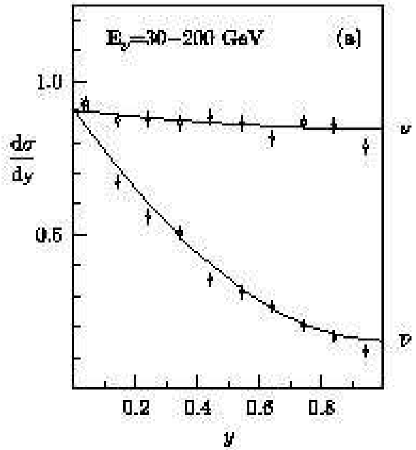
For νN

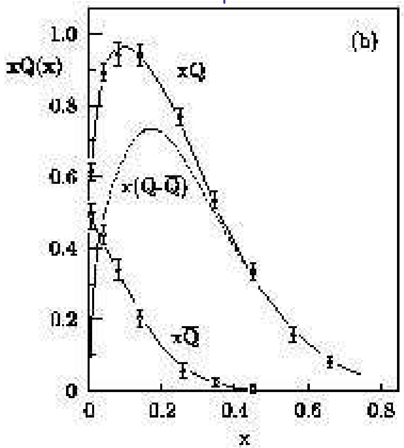
$$\frac{d^2\sigma(\bar{\nu})}{dxdy} = \frac{G_F^2 \, s}{\pi} \Sigma_i \, [x\bar{q}_i(x) + (1-y)^2 xq_i(x)]$$

For $\bar{\nu}$N

Clearly there are antiquarks in the nucleon

3 Valence quarks plus a flavourless $q\bar{q}$ Sea





$q = q_{valence} + q_{sea}$ $\quad \bar{q} = \bar{q}_{sea}$ $\quad q_{sea} = \bar{q}_{sea}$

Compare to the general form of the cross-section for $\nu/\bar{\nu}$ scattering via $W_{+/-}$

$$F_L(x,Q2) = 0$$

$$xF_3(x,Q2) = 2\Sigma_i x(q_i(x) - \bar{q}_i(x))$$

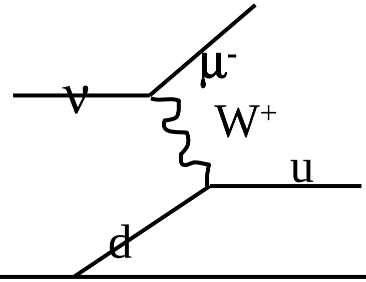
Valence

$$F_2(x,Q2) = 2\Sigma_i x(q_i(x) + \bar{q}_i(x))$$

Valence and Sea

And there will be a relationship between F_2^{eN} and $F_2^{\nu N}$

Also NOTE $\nu,\bar{\nu}$ scattering is FLAVOUR sensitive



W+ can only hit quarks of charge -e/3 or antiquarks -2e/3

$\sigma(\nu p) \sim (d + s) + (1-y)^2 (\bar{u} + \bar{c})$

$\sigma(\bar{\nu} p) \sim (u + c)(1-y)^2 + (\bar{d} + \bar{s})$

So in ν νbar scattering the sums over q, qbar ONLY contain the appropriate flavours BUT- high statistics ν,νbar data are taken on isoscalar targets e.g. Fe=(p+n)/2=N

d in proton = u in neutron

u in proton = d in neutron

$$xF_3(\nu,\bar{\nu}N) = x(u-\bar{u}+d-\bar{d}) = x(u_v+d_v)$$

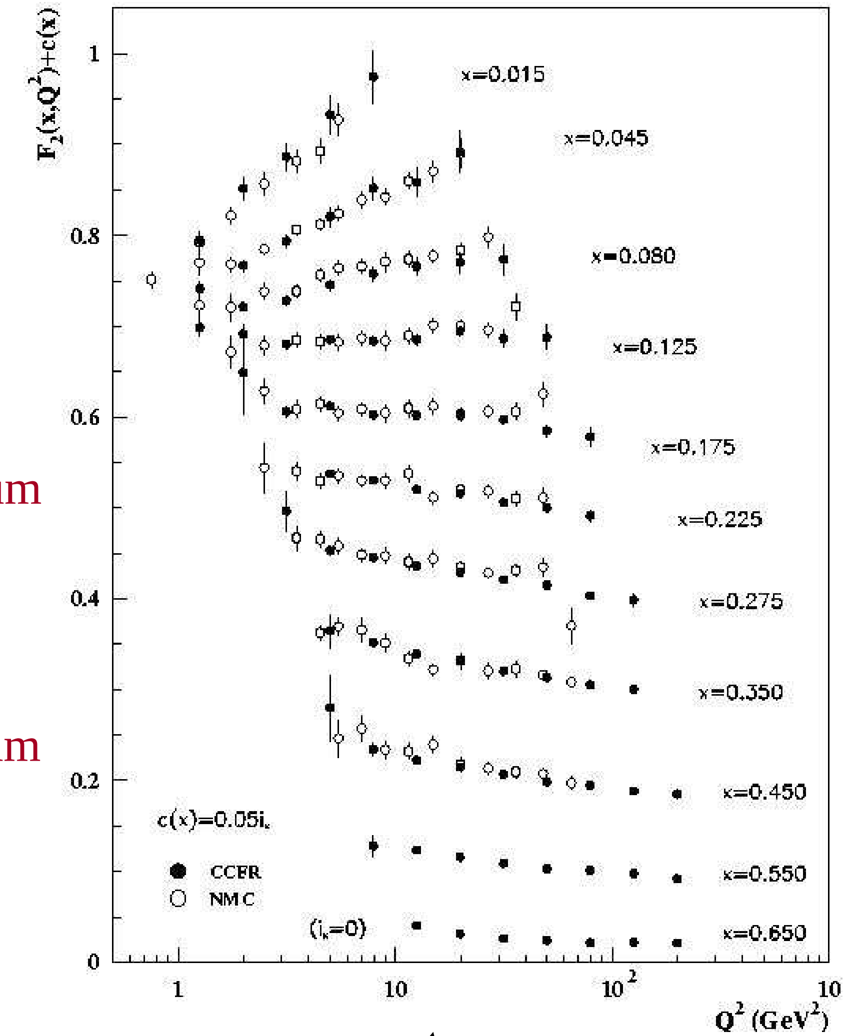
$$\int \frac{xF_3}{x}dx = \int_0^1 (u_v + d_v)\,dx = 3$$

GLS sum rule

$$F_2(\nu,\bar{\nu}N) = x(u+\bar{u}+d+\bar{d}+s+\bar{s}+c+\bar{c})$$

$$\int_0^1 F_2(\nu,\bar{\nu}N)dx = 1?$$

Total momentum of quarks



$$F_2(lp) = x\left(\frac{4}{9}(u+\bar{u})+\frac{1}{9}(d+\bar{d})+\frac{1}{9}(s+\bar{s})+\frac{4}{9}(c+\bar{c})\right)$$

$$F_2(lN) = \frac{5}{18}\,x\left[u+\bar{u}+d+\bar{d}+\frac{2}{5}(s+\bar{s})+\frac{8}{5}(c+\bar{c})\right]$$

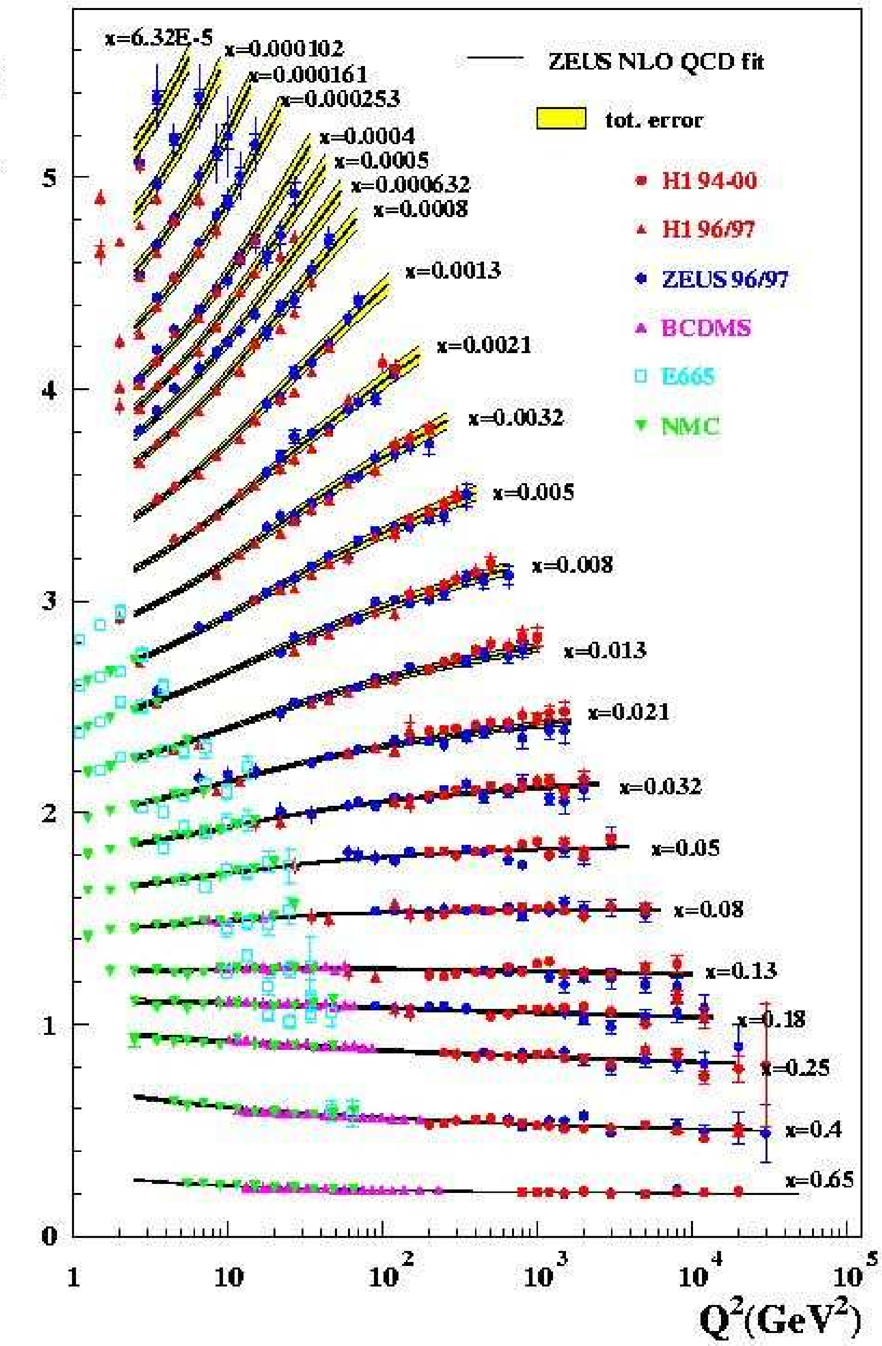
$$F_2(lN) \approx \frac{5}{18}\,F_2(\nu N)$$

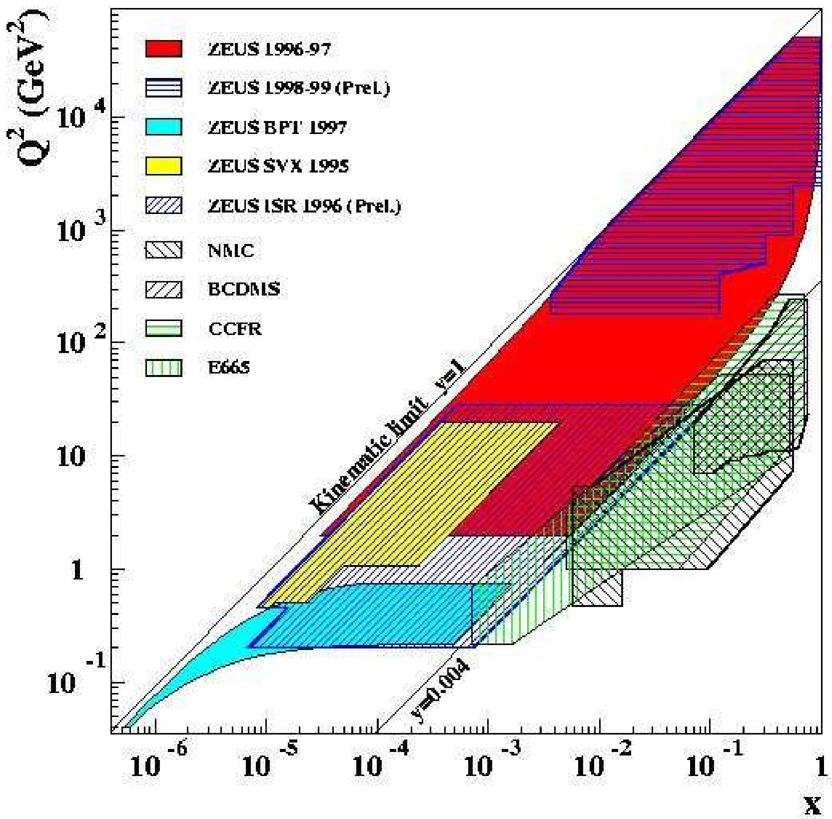
A TRIUMPH

(and 20 years of understanding the c $\bar{c}$ contribution)

Bjorken scaling is broken – $\ln(Q^2)$

Note strong rise at small x





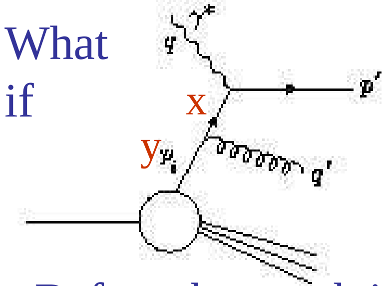
Terrific expansion in measured range across the x, Q^2 plane throughout the 90's

HERA data

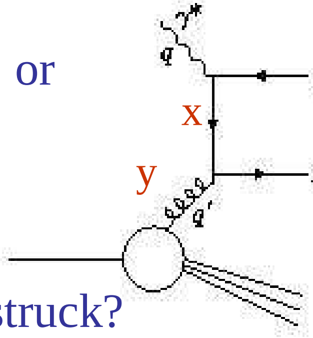
Pre HERA fixed target μp,μD NMC, BDCMS, E665 and ν,$\bar{\nu}$ Fe CCFR

QCD improves the Quark Parton Model

What if



or



Before the quark is struck?

y > x, z = x/y

$$\frac{F_2(x)}{x} = \int dy\,dz\,\delta(x - zy)\,q(y)\left[\sigma^{POINT}(z) + \sigma(\gamma^* q \to qg)\right]$$

$$\sigma(\gamma^* q \to qg) = \frac{e_i^2 \alpha_s}{2\pi} P_{qq}(z) \ln\frac{Q^2}{Q_0^2}$$

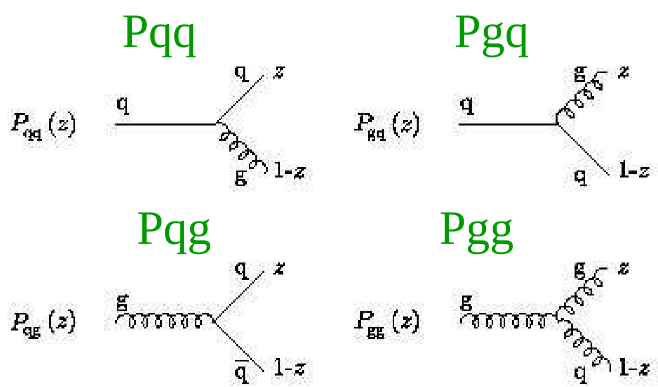
$$\frac{F_2(x,Q^2)}{x} = e_i^2\left[q(x) + \Delta q(x,Q^2)\right] = e_i^2 q(x,Q^2)$$

$$\Delta q(x,Q^2) = \frac{\alpha_s}{2\pi} \ln\frac{Q^2}{Q_0^2} \int_0^1 \frac{dy}{y} q(y) P_{qq}\left(\frac{x}{y}\right)$$

$$\frac{dq(x,Q^2)}{d\ln Q^2} = \frac{\alpha_s(Q^2)}{2\pi} \int_0^1 \frac{dy}{y}\left[P_{qq}(z) q(y,Q^2) + P_{qg}(z) g(y,Q^2)\right]$$

$$\frac{dg(x,Q^2)}{d\ln Q^2} = \frac{\alpha_s(Q^2)}{2\pi} \int_0^1 \frac{dy}{y}\left[\Sigma_q P_{gq}(z) q(y,Q^2) + P_{gg}(z) g(y,Q^2)\right]$$

The DGLAP parton evolution equations



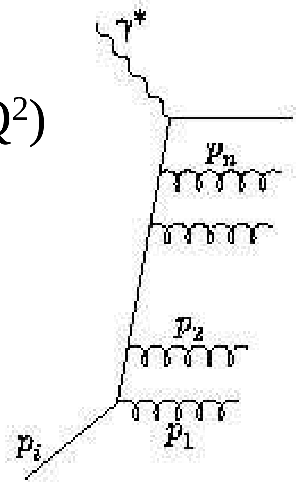
So $F_2(x,Q^2) = \Sigma_i e_i^2 (xq(x,Q^2) + x\bar{q}(x,Q^2))$

in LO QCD

The theory predicts the rate at which the parton distributions (both quarks and gluons) evolve with Q^2- (the energy scale of the probe) -BUT it does not predict their shape

Note $q(x,Q^2) \sim \alpha_s \ln Q^2$, but $\alpha_s(Q^2) \sim 1/\ln Q^2$, so $\alpha_s \ln Q^2$ is O(1), so we must sum all terms

$\alpha_s \to \alpha_s(Q^2)$



$\alpha_s^n \ln Q^{2n}$

Leading Log

Approximation

x decreases from

target to probe

$x_{i-1} > x_i > x_{i+1} \ldots$

p_t^2 of quark relative to proton increases from target to probe

$p_t^2{}_{i-1} < p_t^2{}_i < p_t^2{}_{i+1}$

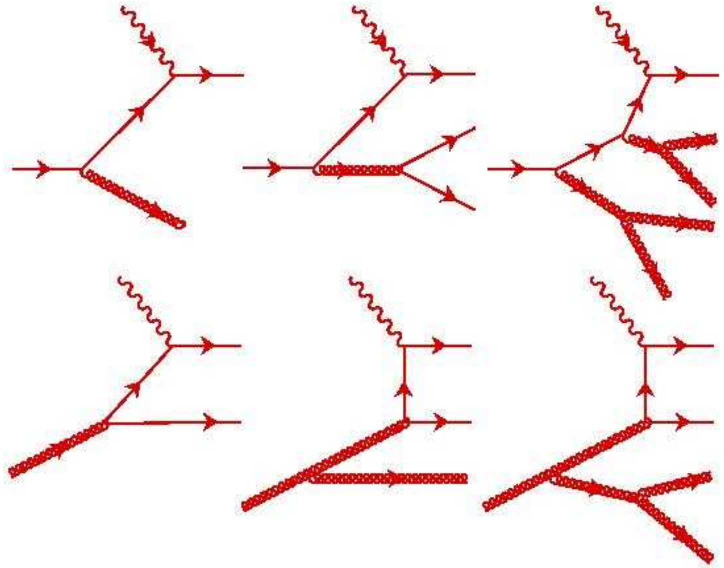
Dominant diagrams have STRONG p_t ordering

F2 is no longer so simply expressed in terms of partons -

convolution with coefficient functions is needed –

but these are calculable in QCD

What if higher orders are needed?



$Pqq(z) = P^0qq(z) + \alpha_s P^1qq(z) + \alpha_s^2 P^2qq(z)$

LO NLO NNLO

$$\frac{F_2(x,Q^2)}{x} = \int_0^1 \frac{dy}{y}\left[\Sigma_i C_2(z,\alpha_s) q_i(x,Q^2) + C_g(z,\alpha_s) g(y,Q^2)\right]$$

$$C_2(z,\alpha_s) = e_i^2\left[\delta(1-z) + \alpha_s f_2(z)\right]$$

$$C_g(z,\alpha_s) = \alpha_s f_g(z)$$

$$F_L(x,Q^2) = \frac{\alpha_s}{x}\left[\frac{4}{3}\int_0^1 \frac{dy}{y} z^2 F_2(y,Q^2) + 2\Sigma_i e_i^2 \int_0^1 \frac{dy}{y} z^2(1-z) yg(y,Q^2)\right]$$

How do we determine Parton Distribution Functions ?

Parametrise the parton distribution functions (PDFs) at Q^2_0 (~1-7 GeV^2)

$$xu_v(x) = A_u x^{au} (1-x)^{bu} (1+ \varepsilon_u \sqrt{x} + \gamma_u x)$$

$$xd_v(x) = A_d x^{ad} (1-x)^{bd} (1+ \varepsilon_d \sqrt{x} + \gamma_d x)$$

$$xS(x) = A_s x^{-\lambda s} (1-x)^{bs} (1+ \varepsilon_s \sqrt{x} + \gamma_s x)$$

$$xg(x) = A_g x^{-\lambda g} (1-x)^{bg} (1+ \varepsilon_g \sqrt{x} + \gamma_g x)$$

These parameters control the low-x shape

These parameters control the middling-x shape

These parameters control the high-x shape

Parameters Ag, Au, Ad are fixed through momentum and number sum rules – explain other parameters may be fixed by model choices-

Model choices ⇒Form of parametrization at Q^2_0, value of Q^2_0,, flavour structure of sea, cuts applied, heavy flavour scheme → typically ~15 parameters

Use QCD to evolve these PDFs to $Q^2 > Q^2_0$

Construct the measurable structure functions by convoluting PDFs with coefficient functions: make predictions for ~1500 data points across the x,Q2 plane

Perform χ2 fit to the data

The fact that so few parameters allows us to fit so many data points established QCD as the THEORY OF THE STRONG INTERACTION and provided the first measurements of α_s (as one of the fit parameters)

These days we assume the validity of the picture to measure parton distribution functions

PDFs are extracted by MRST, CTEQ, ZEUS, H1 ... http://durpdg.dur.ac.uk/hepdata/pdf.html

But where is the information coming from?

LHAPDF v5

Fixed target e/μ p/D data from **NMC, BCDMS, E665, SLAC**

$F_2(e/\mu p) \sim 4/9\, x(u+\bar{u}) + 1/9 x(d+\bar{d}) + 4/9\, x(c+\bar{c}) + 1/9 x(s+\bar{s})$

$F_2(e/\mu D) \sim 5/18\, x(u+\bar{u}+d+\bar{d}) + 4/9\, x(c+\bar{c}) + 1/9 x(s+\bar{s})$

Assuming u in proton = d in neutron – strong-isospin

Also use ν, $\bar{\nu}$ data from **CCFR** (Beware Fe target needs corrections)

$F2(\nu,\bar{\nu}N) = x(u+\bar{u}+d+\bar{d}+s+\bar{s}+c+\bar{c})$

$xF_3(\nu,\bar{\nu}N) = x(u_v + d_v)$ (provided $s = \bar{s}$)

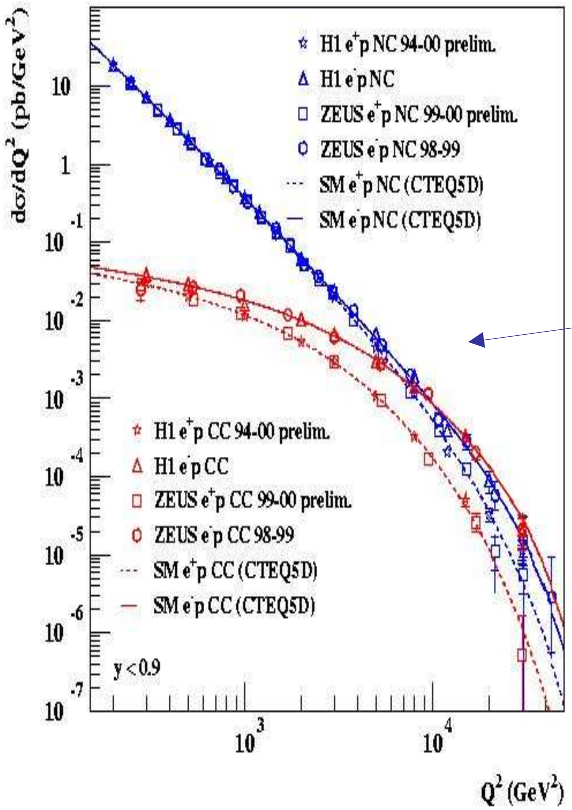
Valence information for 0< x < 1

Can get ~4 distributions from this: e.g. u, d, ubar, dbar – but need assumptions like q=qbar for all flavours, sbar=1/4 (ubar+dbar), dbar=ubar (wrong!) and need heavy quark treatment…(.*not part of this talk..see Devenish & Cooper-Sarkar 'Deep Inelastic Scattering', OUP 2004)*

Note gluon enters only indirectly via DGLAP equations for evolution

HERA ep neutral current (γ-exchange) data give much more information on the sea and gluon at small x….. xSea directly from F_2

xGluon from scaling violations $dF_2/d\ln Q^2$ – the relationship to the gluon is much more direct at small-x



HERA data have also provided information at high $Q^2 \rightarrow Z^0$ and $W^{+/-}$ become as important as γ exchange $\rightarrow$ NC and CC cross-sections comparable

For NC processes

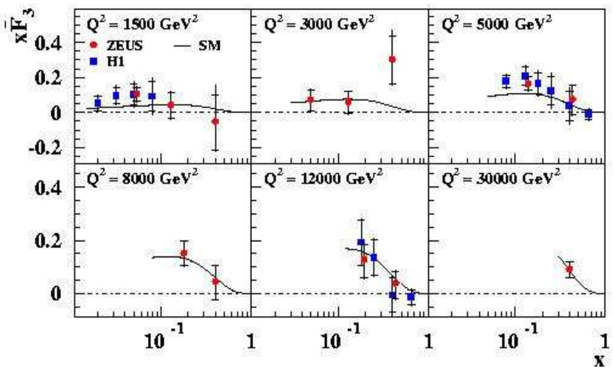
$$F_2 = \Sigma_i\, A_i(Q^2)\, [xq_i(x,Q^2) + x\bar{q}_i(x,Q^2)]$$

$$xF_3 = \Sigma_i\, B_i(Q^2)\, [xq_i(x,Q^2) - x\bar{q}_i(x,Q^2)]$$

$$A_i(Q^2) = e_i^2 - 2\, e_i\, v_i\, v_e\, P_Z + (v_e^2+a_e^2)(v_i^2+a_i^2)\, P_Z^2$$

$$B_i(Q^2) = \quad - 2\, e_i\, a_i\, a_e\, P_Z + \quad 4a_i\, a_e\, v_i\, v_e \quad P_Z^2$$

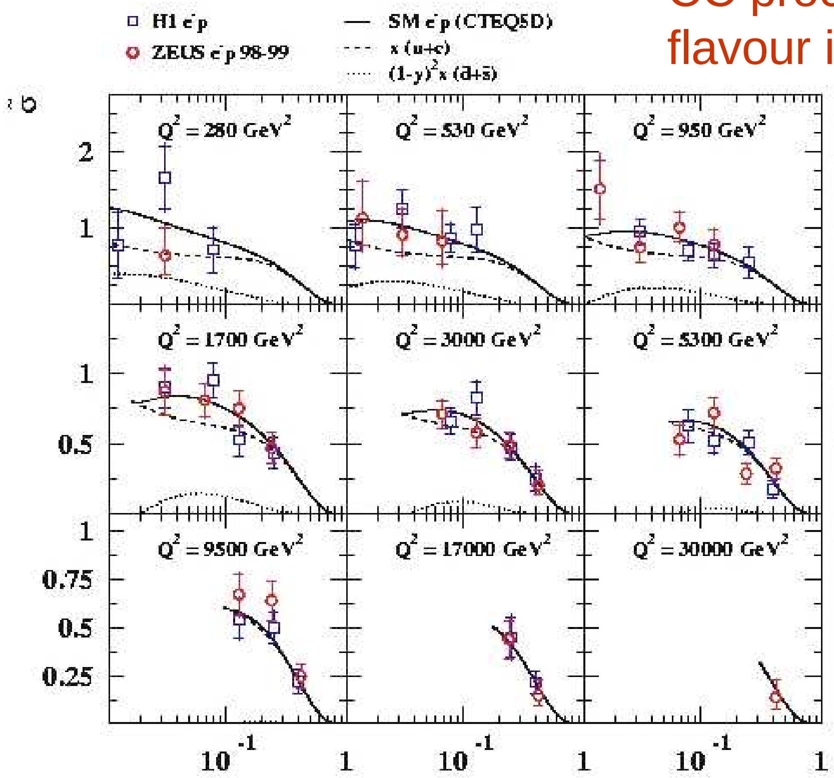
$$P_Z^2 = Q^2/(Q^2 + M^2_Z)\; 1/\sin^2\theta_W$$

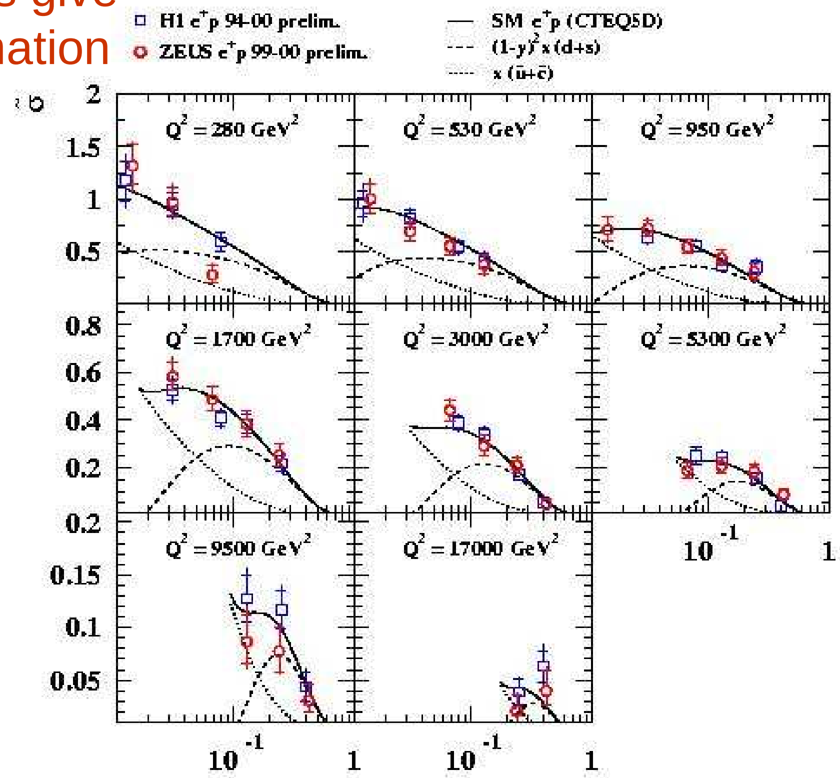


$\rightarrow$a new valence structure function xF_3 due to Z exchange is measurable from low to high x- on a pure proton target $\rightarrow$ no heavy target corrections- no assumptions about strong isospin

$\rightarrow$ e- running at HERA-II is already improving this measurement (to be released April'06

CC processes give flavour information





$$\frac{d^2\sigma(e^-p)}{dxdy} = \frac{G_F^2\, M^4_W}{2\pi x(Q^2+M^2_W)^2}\left[x\,(u+c) + (1-y)^2 x\,(\bar{d}+\bar{s})\right]$$

$$\frac{d^2\sigma(e^+p)}{dxdy} = \frac{G_F^2\, M^4_W}{2\pi x(Q^2+M^2_W)^2}\left[x\,(\bar{u}+\bar{c}) + (1-y)^2 x\,(d+s)\right]$$

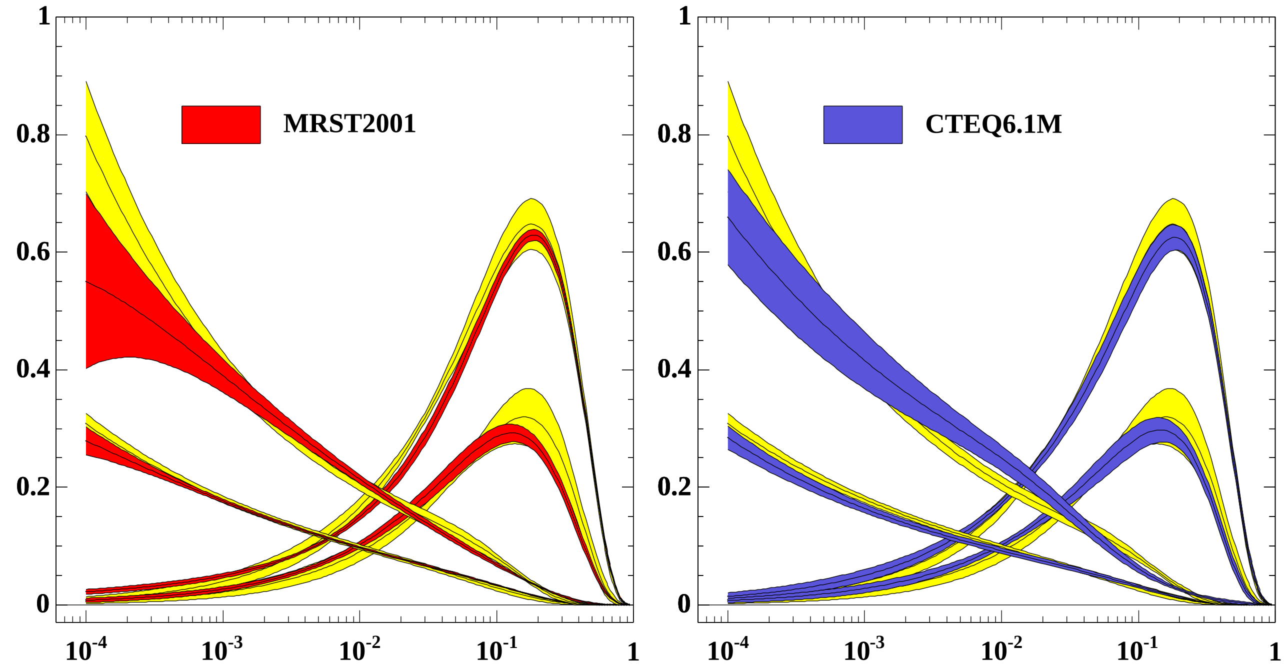
M_W information

u_v at high x

d_v at high x

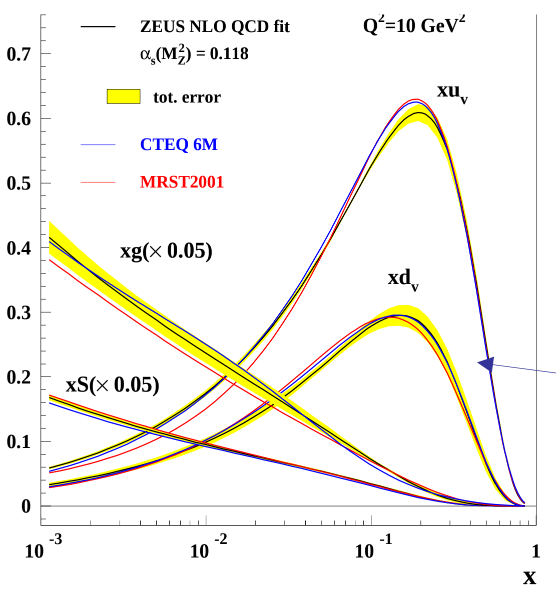
Measurement of high-x d_v on a pure proton target

d is not well known because u couples more strongly to the photon. Historically information has come from deuterium targets –but even Deuterium needs binding corrections. Open questions: does u in proton = d in neutron?, does dv/uv ⇒ 0, as x ⇒ 1?



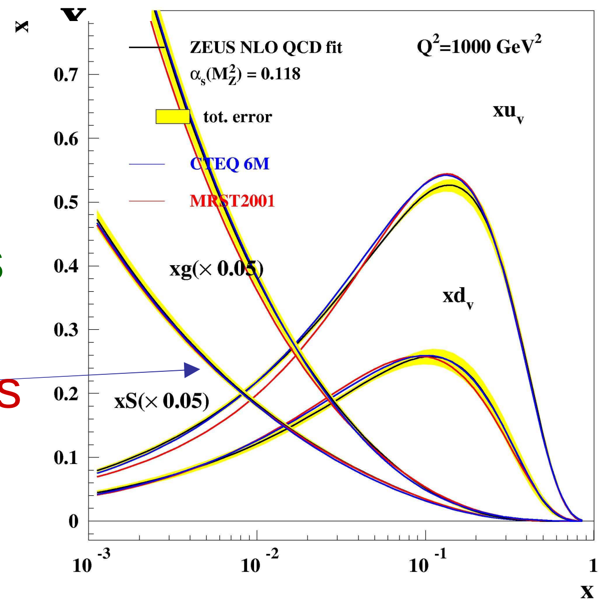
New millenium PDFs

With uncertainty estimates



–Valence distributions evolve slowly

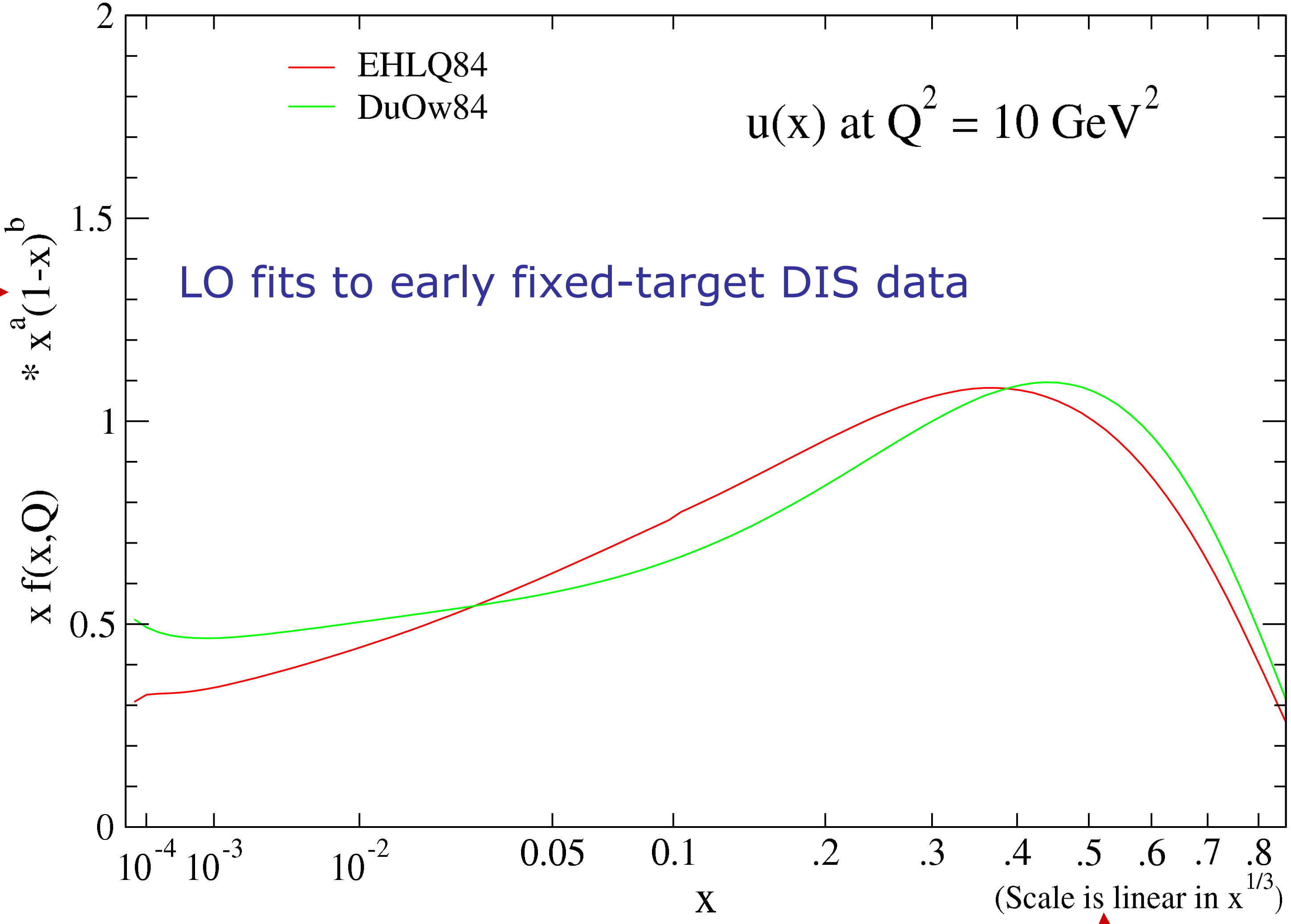
Sea/Gluon distributions evolve fast



So how certain are we? First, some quantitative measure of the progress made over 20 years of PDF fitting (thanks to Wu-ki Tung)

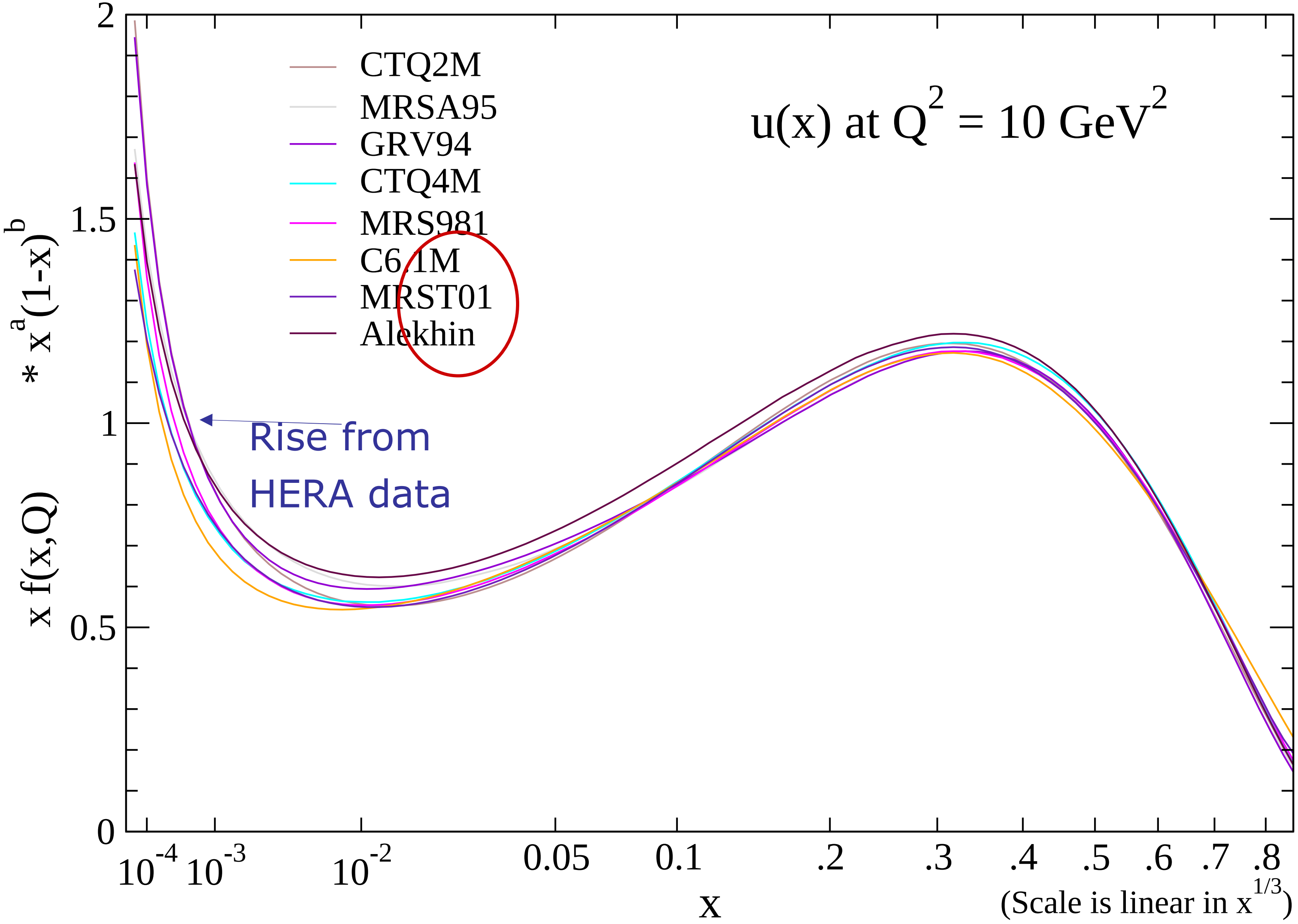
	Fixed-tgt	HERA	DY-W	Jets	Total
# Expt pts.	1070	484	145	123	1822
EHLQ ‘84	11475	7750	2373	331	21929
DuOw ‘84	8308	5005	1599	275	15187
MoTu ~‘90	3551	3707	857	218	8333
KMRS ~‘90	1815	7709	577	280	10381
CTQ2M ~’94	1531	1241	646	224	3642
MRSA ~’94	1590	983	249	231	3054
GRV94 ~’94	1497	3779	302	213	5791
CTQ4M ~’98	1414	666	227	206	2513
MRS98 ~’98	1398	659	111	227	2396
CTQ6M 02	1239	508	159	123	2029
MRST01/2	1378	530	120	236	2264

The u quark



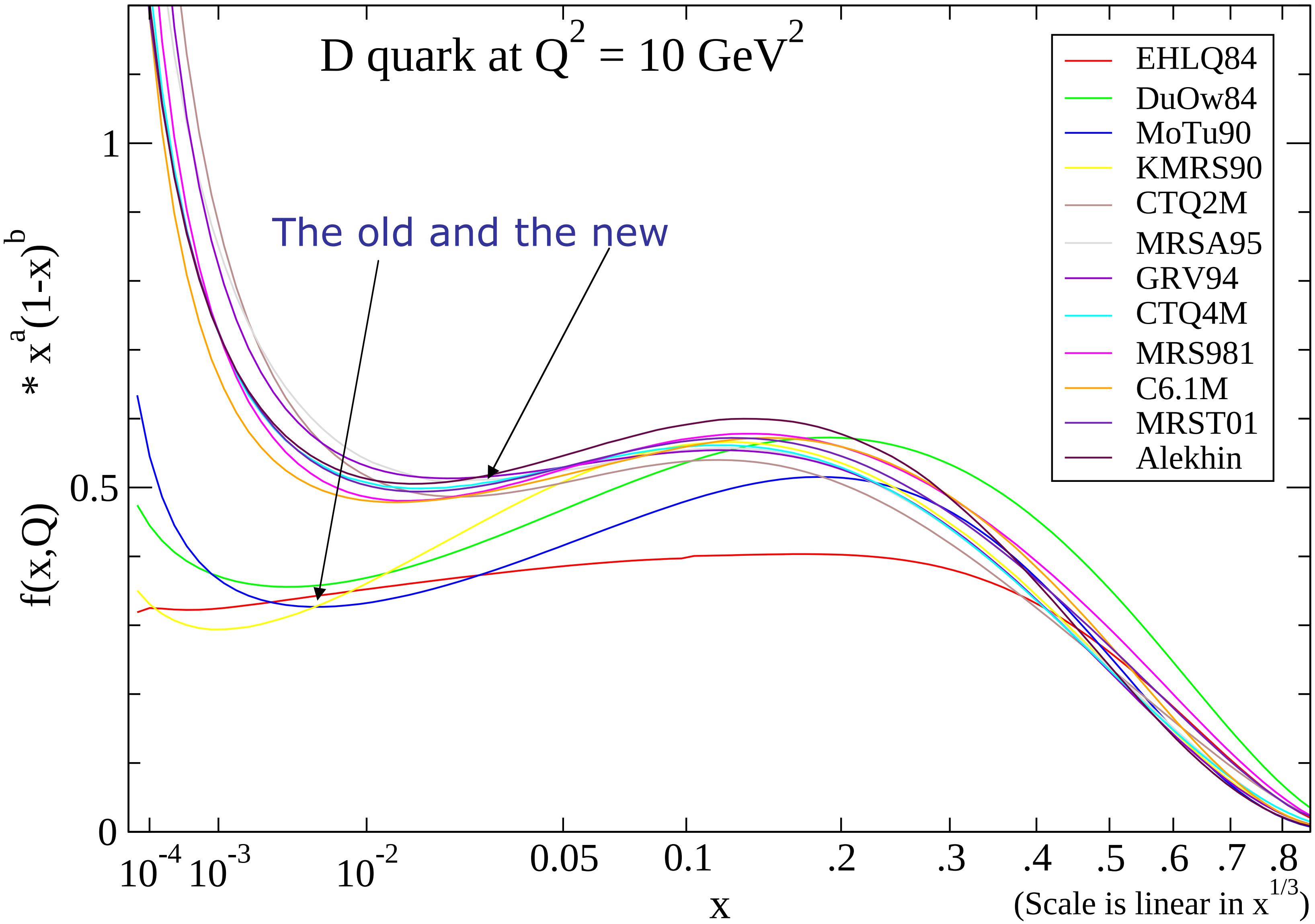
To reveal the difference in both large and small x regions

To view small and large x in one plot



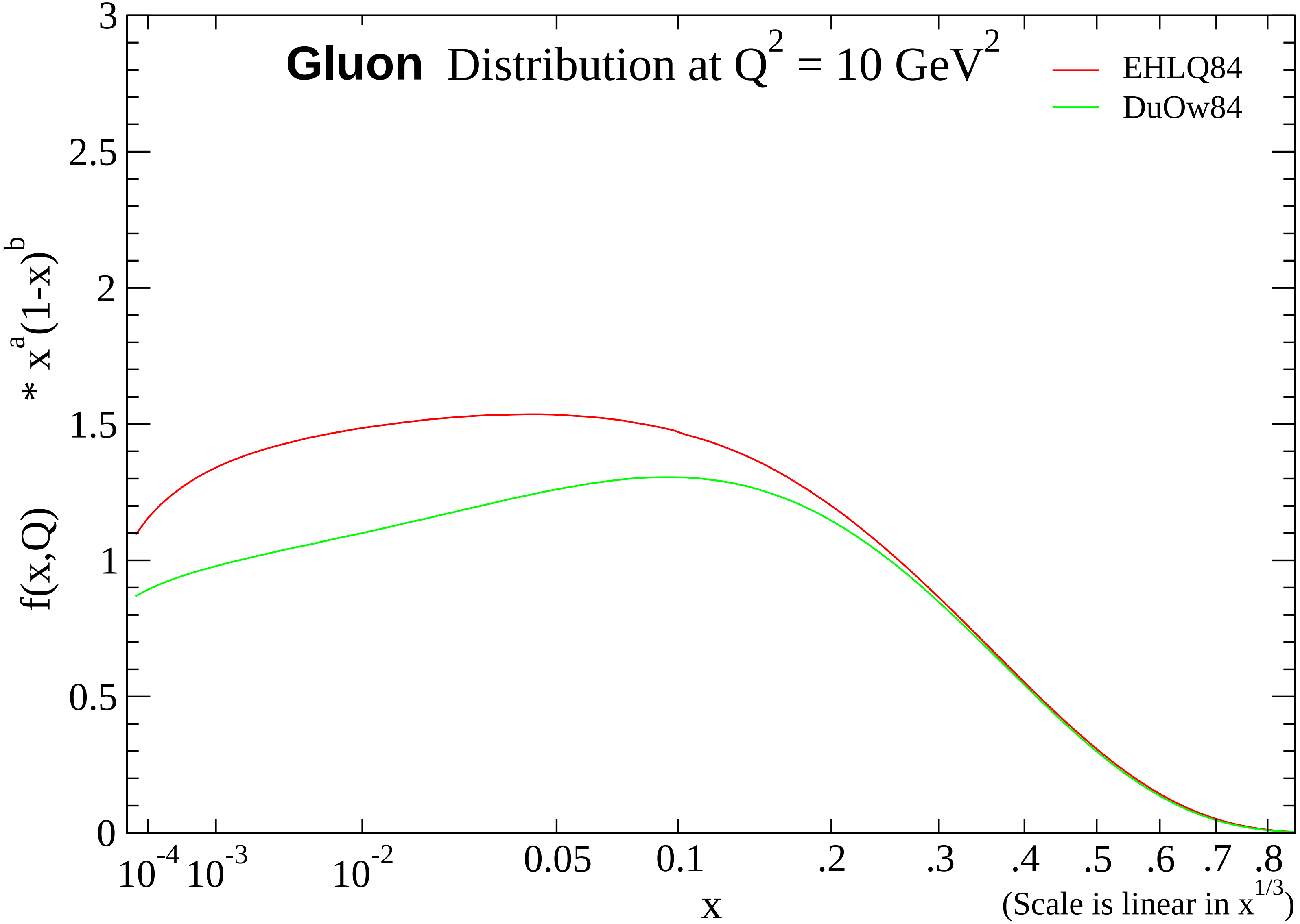
CTQ2M
MRSA95
GRV94
CTQ4M
MRS981
C6.1M
MRST01
Alekhin
u(x) at $Q^2 = 10$ GeV^2
Rise from
HERA data
x f(x,Q) * $x^a(1-x)^b$
0
0.5
1
1.5
2
10^{-4} 10^{-3} 10^{-2} 0.05 0.1 .2 .3 .4 .5 .6 .7 .8
x
(Scale is linear in $x^{1/3}$)

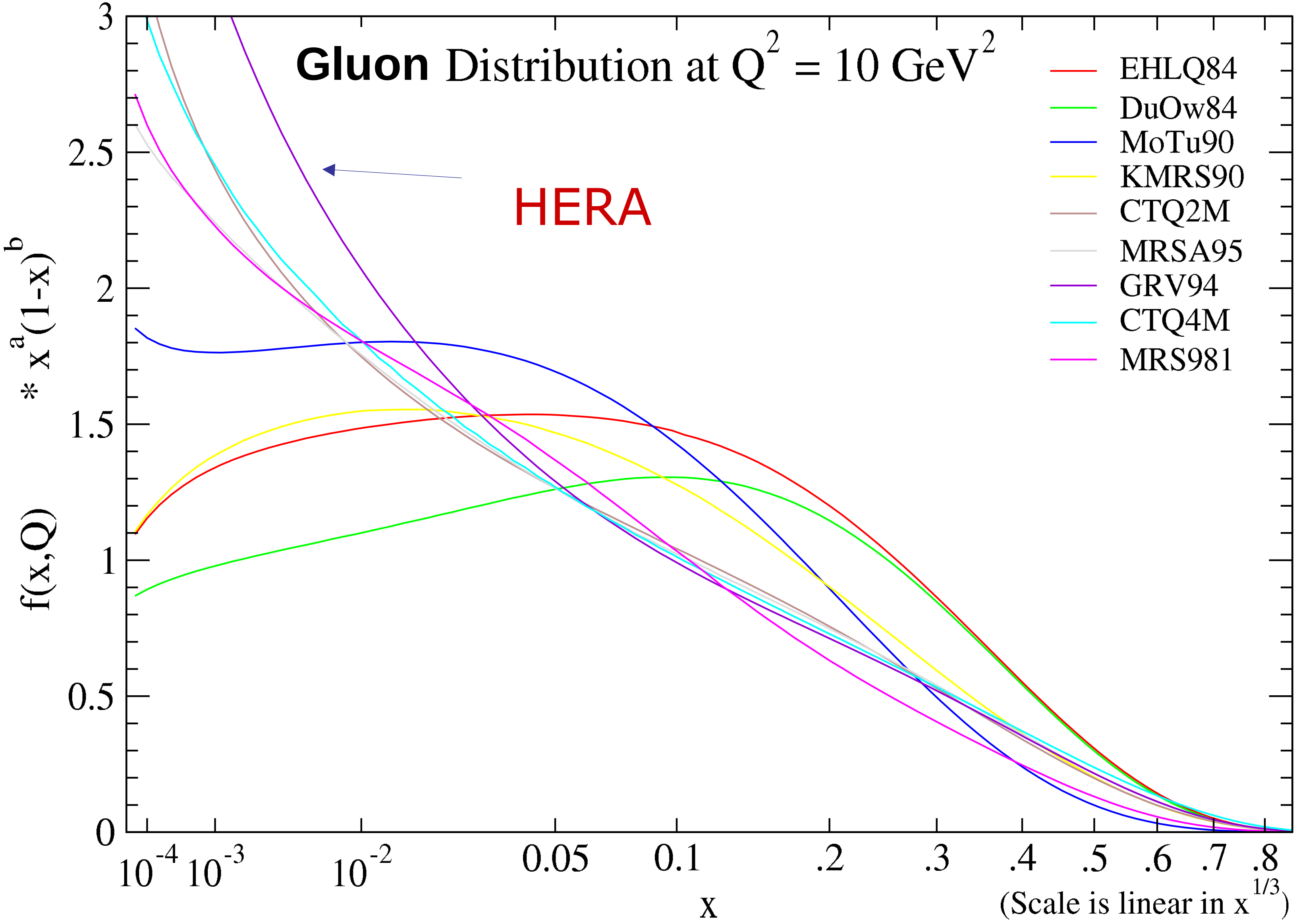
d quark

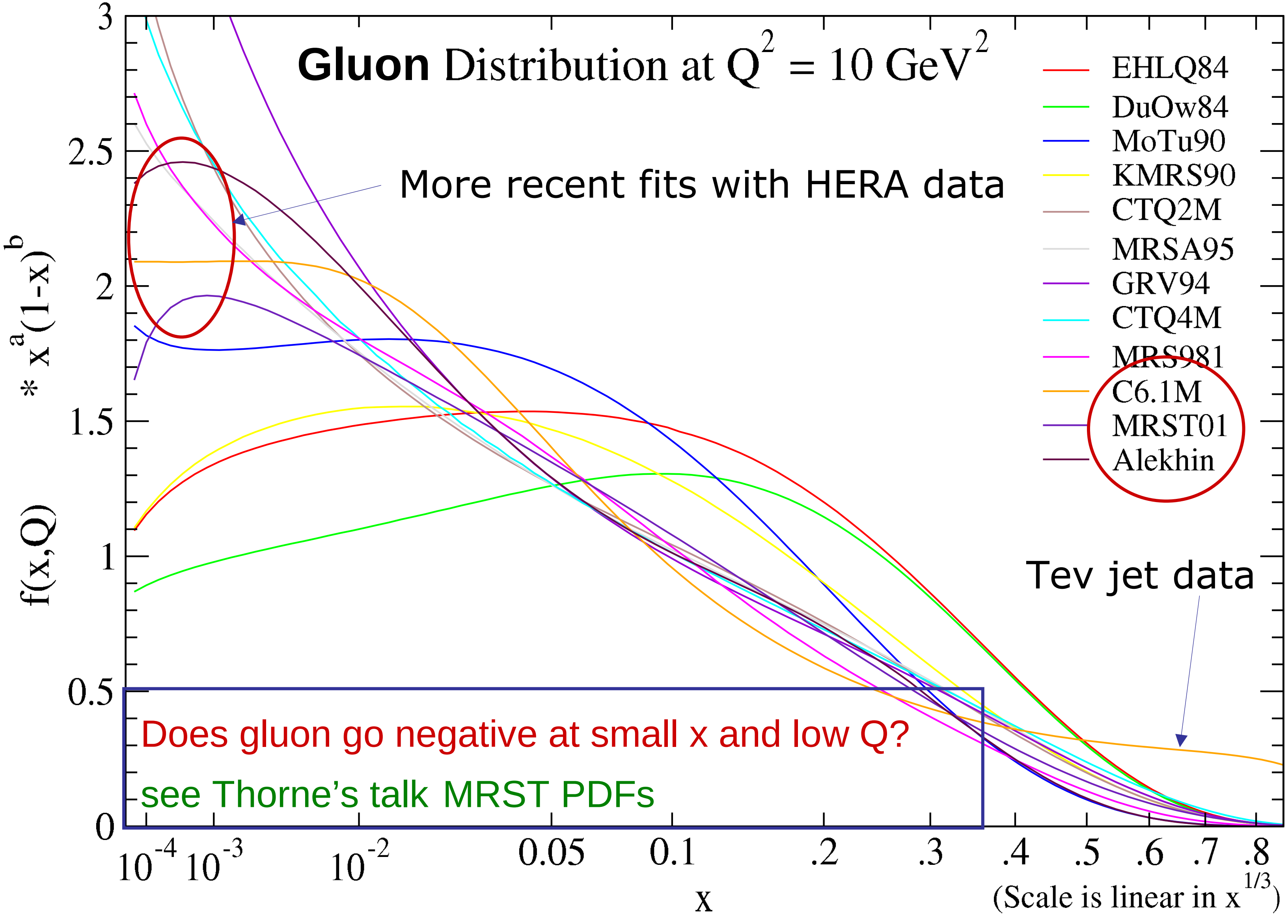


D quark at Q^2 = 10 GeV^2
The old and the new
f(x,Q) * x^a(1-x)^b
x
(Scale is linear in x^1/3)
EHLQ84
DuOw84
MoTu90
KMRS90
CTQ2M
MRSA95
GRV94
CTQ4M
MRS981
C6.1M
MRST01
Alekhin
0
0.5
1
10^-4
10^-3
10^-2
0.05
0.1
.2
.3
.4
.5
.6
.7
.8

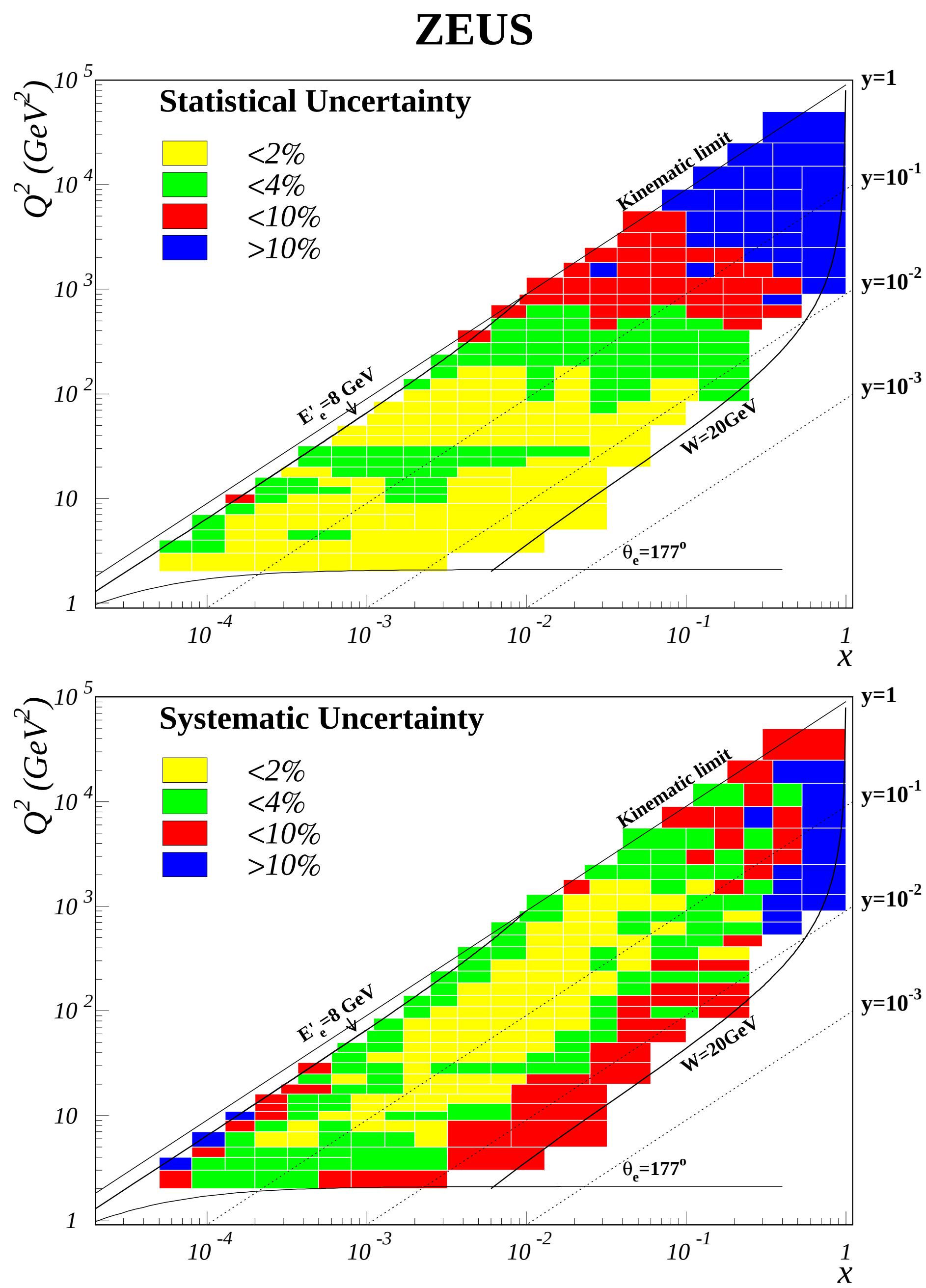
The story about the gluon is more complex




Gluon Distribution at $Q^2 = 10\ GeV^2$
HERA
EHLQ84
DuOw84
MoTu90
KMRS90
CTQ2M
MRSA95
GRV94
CTQ4M
MRS981
$f(x,Q) * x^a(1-x)^b$
3
2.5
2
1.5
1
0.5
0
10^{-4} 10^{-3} 10^{-2} 0.05 0.1 .2 .3 .4 .5 .6 .7 .8
x
(Scale is linear in $x^{1/3}$)



Gluon Distribution at Q2 = 10 GeV2
EHLQ84
DuOw84
MoTu90
KMRS90
CTQ2M
MRSA95
GRV94
CTQ4M
MRS981
C6.1M
MRST01
Alekhin
More recent fits with HERA data
Tev jet data
Does gluon go negative at small x and low Q?
see Thorne's talk MRST PDFs
f(x,Q) * x^a(1-x)^b
x
(Scale is linear in x^1/3)



Modern analyses assess PDF uncertainties within the fit

Clearly errors assigned to the data points translate into errors assigned to the fit parameters --

and these can be propagated to any quantity which depends on these parameters— **the parton distributions or the structure functions and cross-sections which are calculated from them**

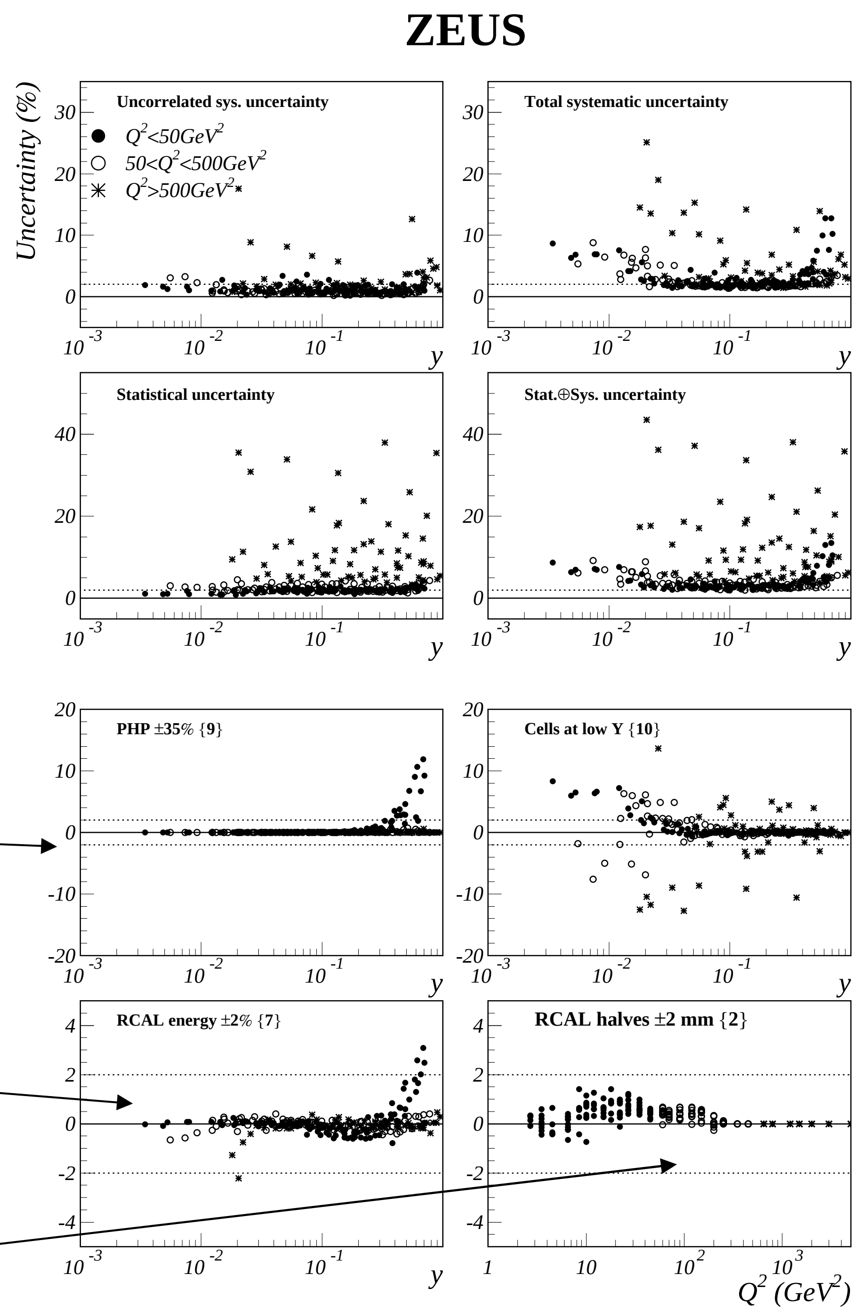
$$< \sigma^2 F > = \Sigma_j \Sigma_k \frac{\partial F}{\partial p_j} V_{jk} \frac{\partial F}{\partial p_k}$$

The errors assigned to the data are both statistical and systematic **and for much of the kinematic plane the size of the point-to-point correlated systematic errors is ~3 times the statistical errors.**

What are the sources of correlated systematic errors?

Normalisations are an obvious example

BUT there are more subtle cases- e.g. **Calorimeter energy scale/angular resolutions** can move events between x,Q2 bins and thus **change the shape** of experimental distributions



Vary the estimate of the photo-production background

Vary energy scales in different regions of the calorimeter

Vary position of the RCAL halves

Why does it matter?

Treatment of correlated systematic errors

$$\chi 2 = \Sigma_i \left[\frac{F_i^{QCD}(p) - F_i^{MEAS}}{(\sigma_i^{STAT})^2 + (\Delta_i^{SYS})^2} \right]^2$$

Errors on the fit parameters, p, evaluated from $\Delta\chi 2 = 1$,

THIS IS NOT GOOD ENOUGH if experimental systematic errors are correlated between data points-

$$\chi^2 = \Sigma_i \Sigma_j \left[F_i^{QCD}(p) - F_i^{MEAS} \right] V_{ij}^{-1} \left[F_j^{QCD}(p) - F_j^{MEAS} \right]$$

$$V_{ij} = \delta_{ij}(\sigma_i^{STAT})^2 + \Sigma_\lambda \Delta_{i\lambda}^{SYS} \Delta_{j\lambda}^{SYS}$$

Where $\Delta_{i\lambda}^{SYS}$ is the correlated error on point **i** due to systematic error source $\boldsymbol{\lambda}$

It can be established that this is equivalent to

$$\chi^2 = \Sigma_i \frac{\left[F_i^{QCD}(p) - \Sigma_\lambda s_\lambda \Delta_{i\lambda}^{SYS} - F_i^{MEAS} \right]^2}{(\sigma_i^{STAT})^2} + \Sigma s_\lambda^2$$

Where s_λ are systematic uncertainty fit parameters of zero mean and unit variance

This has modified the fit prediction by each source of systematic uncertainty

CTEQ, ZEUS, H1, MRST have all adopted this form of $\chi 2$ – but use it differently in the OFFSET and HESSIAN methods …hep-ph/0205153

How do experimentalists usually proceed: OFFSET method

1. Perform fit without correlated errors ($s_\lambda = 0$) for central fit
2. Shift measurement to upper limit of one of its systematic uncertainties ($s_\lambda = +1$)
3. Redo fit, record differences of parameters from those of step 1
4. Go back to 2, shift measurement to lower limit ($s_\lambda = -1$)
5. Go back to 2, repeat 2-4 for next source of systematic uncertainty
6. Add all deviations from central fit in quadrature (positive and negative deviations added in quadrature separately)
7. This method does not assume that correlated systematic uncertainties are Gaussian distributed

A1

Fortunately, there are smart ways to do this (Pascaud and Zomer LAL-95-05, Botje hep-ph-0110123)

Slide 25

A1 Cooper-Sarkar, 15/03/2004

There are other ways to treat correlated systematic errors- HESSIAN method (covariance method)

Allow s_λ parameters to vary for the central fit.

If we believe the theory why not let it calibrate the detector(s)? Effectively the theoretical prediction is not fitted to the central values of published experimental data, but allows these data points to move collectively according to their correlated systematic uncertainties

The fit determines the optimal settings for correlated systematic shifts such that the most consistent fit to all data sets is obtained. In a global fit the systematic uncertainties of one experiment will correlate to those of another through the fit

The resulting estimate of PDF errors is much smaller than for the Offset method for $\Delta\chi 2 = 1$

We must be very confident of the theory to trust it for calibration– but more dubiously **we must be very confident of the model choices** we made in setting boundary conditions

We must check that superficial changes of model choice (values of Q^2_0, form of parametrization…) do not result in large changes of s_λ

We must also check that $|s_\lambda|$ values are not >>1, so that data points are not shifted far outside their one standard deviation errors - Data inconsistencies!

In practice there are problems. Some data sets incompatible/only marginally compatible?

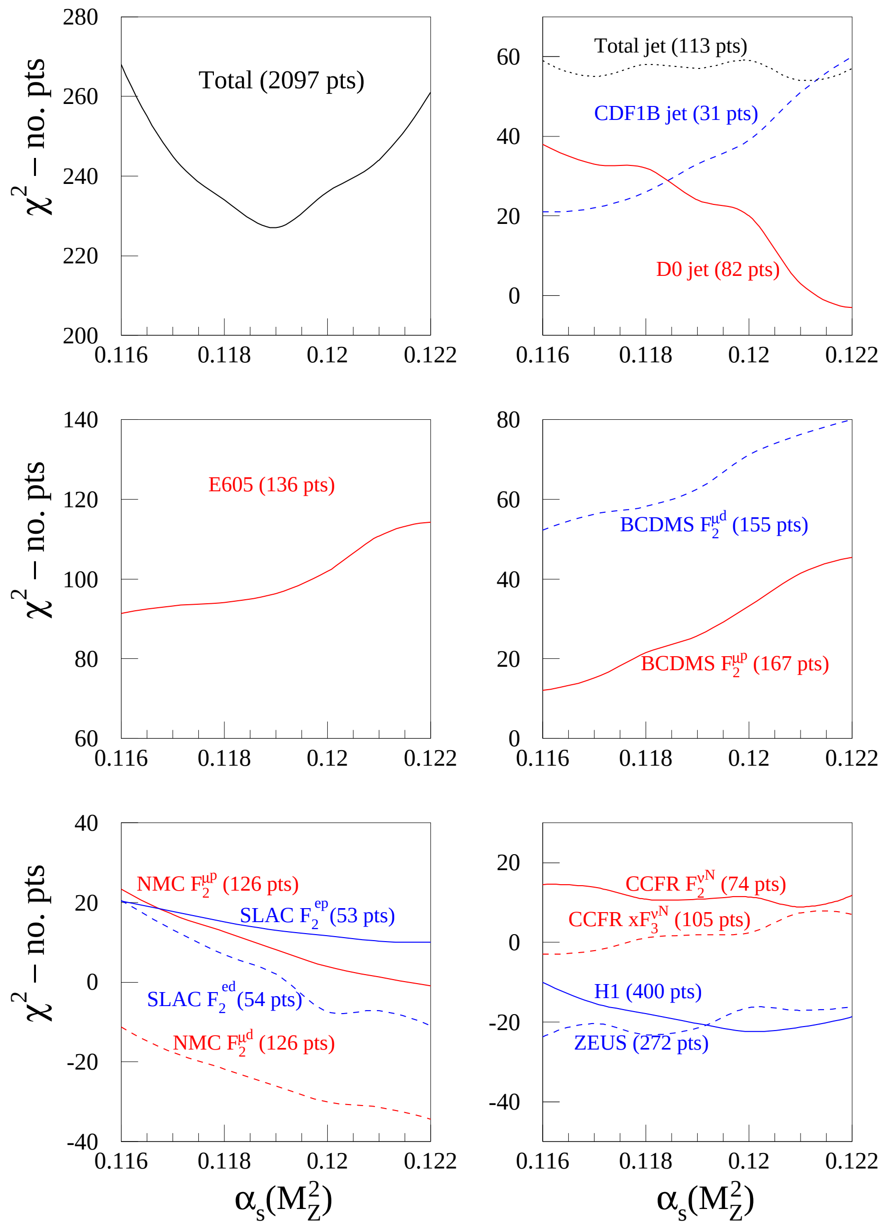
One could restrict the data sets to those which are sufficiently consistent that these problems do not arise – (H1, GKK, Alekhin)

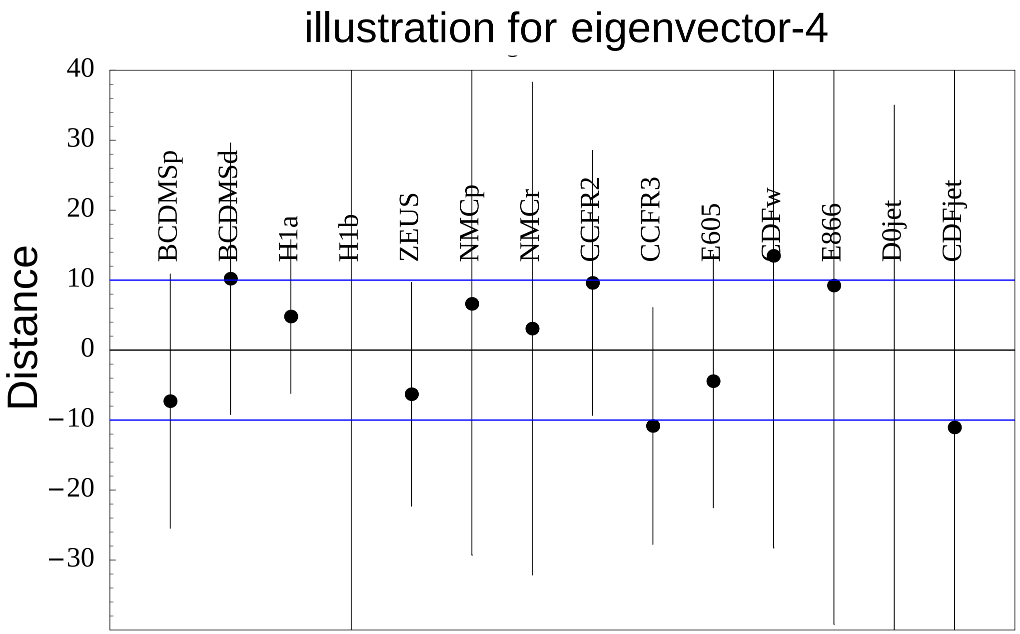
But one loses information since partons need constraints from many different data sets – no-one experiment has sufficient kinematic range / flavour info.

To illustrate: the $\chi 2$ for the MRST global fit is plotted versus the variation of a particular parameter (α_s).

The individual $\chi 2_e$ for each experiment is also plotted versus this parameter in the neighbourhood of the global minimum. Each experiment favours a different value of. α_s

PDF fitting is a compromise. Can one evaluate acceptable ranges of the parameter value with respect to the individual experiments?





CTEQ look at eigenvector combinations of their parameters rather than the parameters themselves. They determine the 90% C.L. bounds on the distance from the global minimum from

P(χe2, Ne) dχe2 =0.9

for each experiment

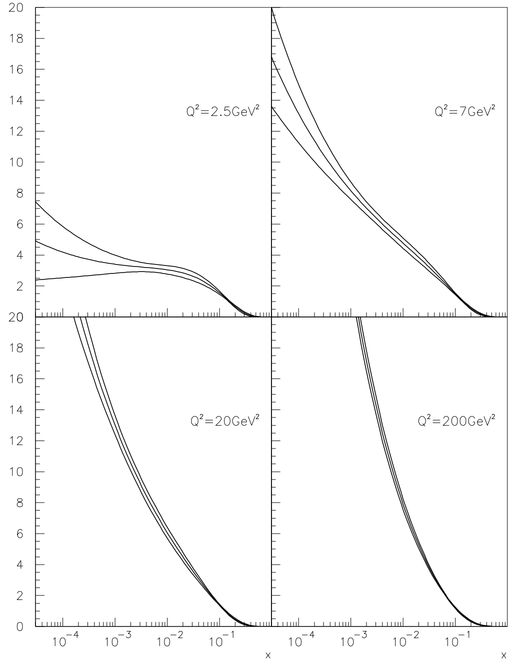
This leads them to suggest a modification of the χ2 tolerance, Δχ2 = 1, with which errors are evaluated such that Δχ2 = T^2, T = 10.

Why? Pragmatism. The size of the tolerance T is set by considering the distances from the χ2 minima of individual data sets from the global minimum for all the eigenvector combinations of the parameters of the fit.

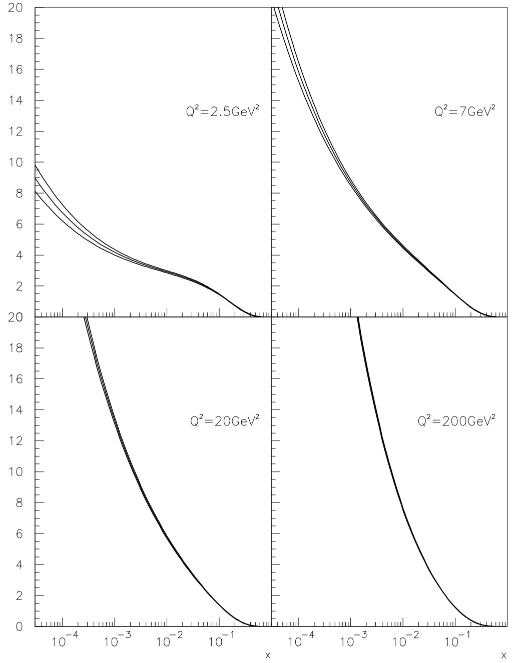
All of the world's data sets must be considered acceptable and compatible at some level, even if strict statistical criteria are not met, since the conditions for the application of strict statistical criteria, namely Gaussian error distributions are also not met.

One does not wish to lose constraints on the PDFs by dropping data sets, but the level of inconsistency between data sets must be reflected in the uncertainties on the PDFs.

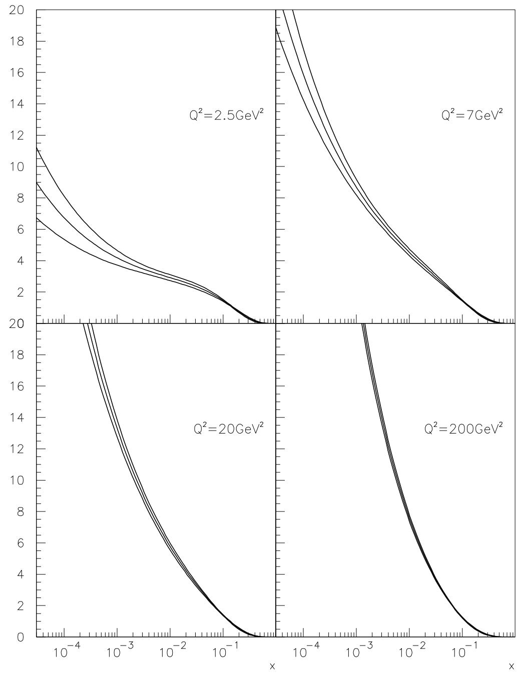
Compare gluon PDFs for Hessian and Offset methods for the ZEUS fit analysis



Offset method



Hessian method T=1



Hessian method T=7

The Hessian method gives comparable size of error band as the Offset method, when the tolerance is raised to T ~ 7 – (similar ball park to CTEQ, T=10)

Note this makes the error band large enough to encompass reasonable variations of model choice. (For the ZEUS global fit $\sqrt{2N}$=50, where N is the number of degrees of freedom)

Aside on model choices

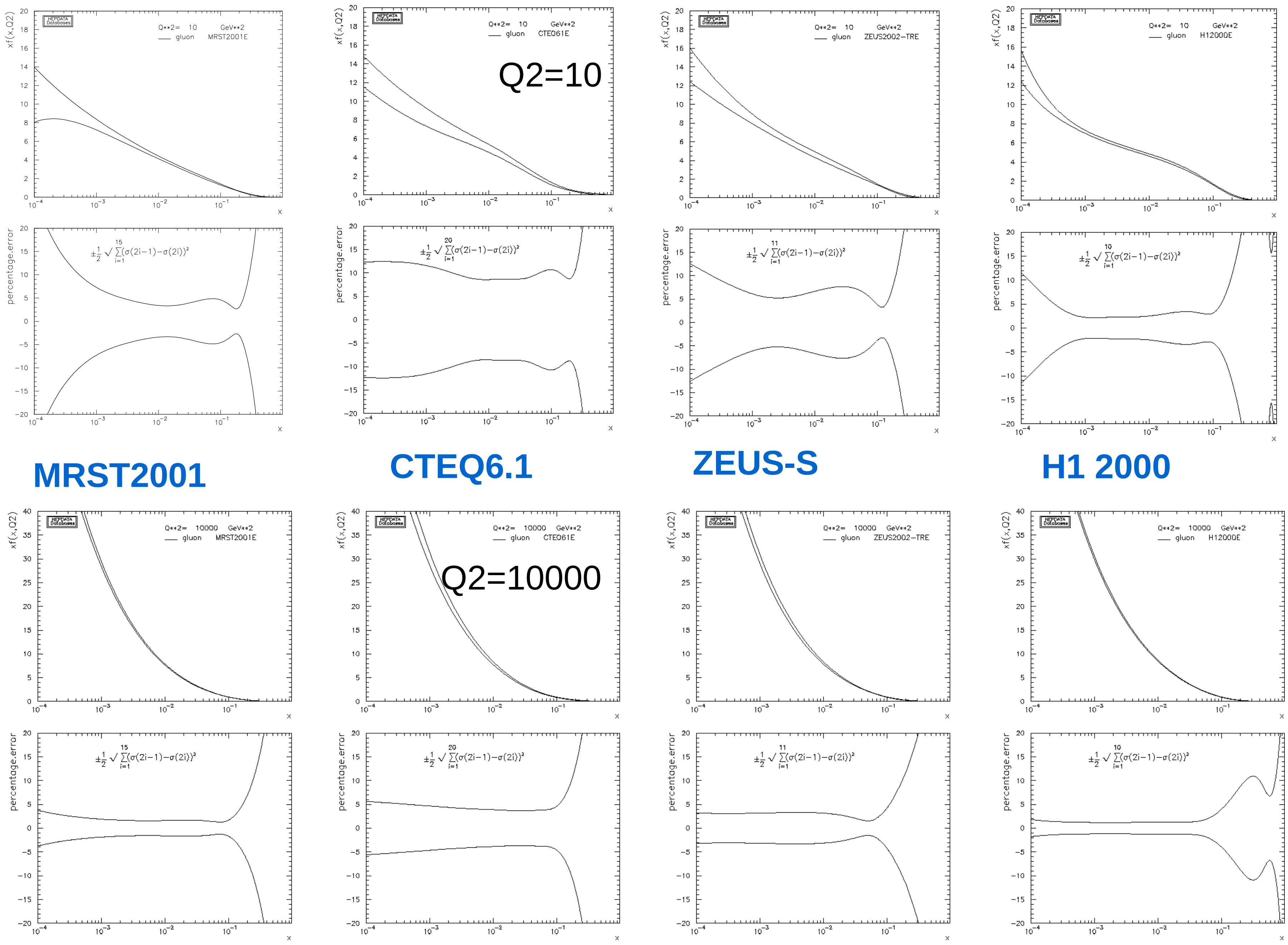
We trust NLO QCD– but are we sure about every choice which goes into setting up the boundary conditions for QCD evolution? – form of parametrization etc.

The statistical criterion for parameter error estimation within a particular hypothesis is $\Delta\chi2 = T^2 = 1$. But for judging the acceptability of an hypothesis the criterion is that $\chi2$ lie in the range $N \pm\sqrt{2N}$, where N is the number of degrees of freedom

There are many choices, such as the form of the parametrization at Q^2_0, the value of Q_0^2 itself, the flavour structure of the sea, etc., which might be considered as superficial changes of hypothesis, but the $\chi2$ change for these different hypotheses often exceeds $\Delta\chi2=1$, while remaining acceptably within the range $N \pm\sqrt{2N}$.

In this case the model error on the PDF parameters usually exceeds the experimental error on the PDF, if this has been evaluated using T=1, with the Hessian method.

This leads to somewhat different uncertainty estimates e.g gluon comparison



The general trend of PDF uncertainties is that

The u quark is much better known than the d quark

The valence quarks are much better known than the gluon at high-x

The valence quarks are poorly known at small-x but they are not important for physics in this region

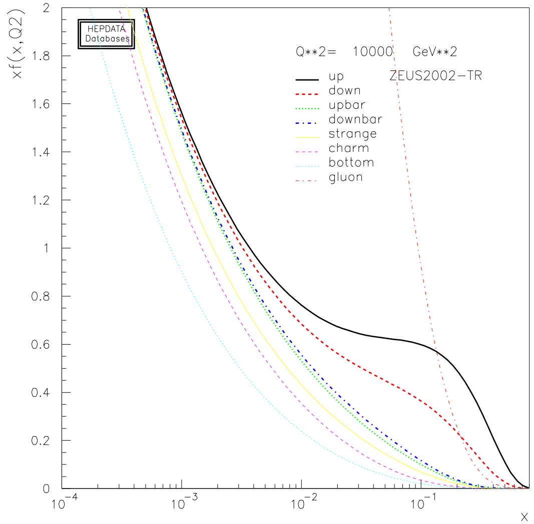
The sea and the gluon are well known at low-x

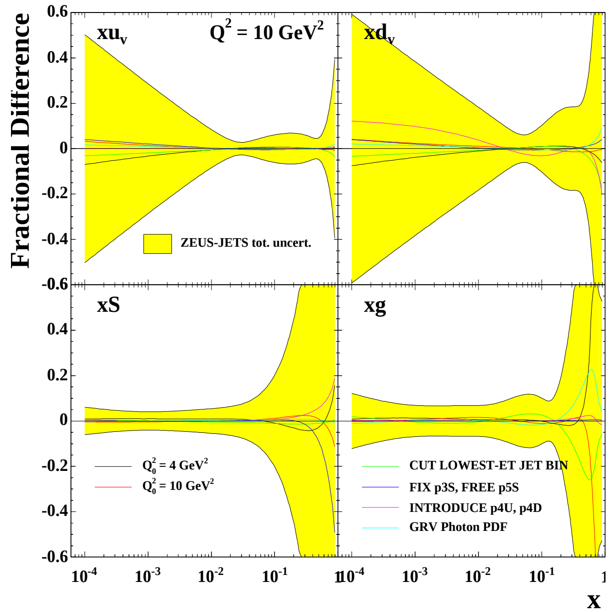
The sea is poorly known at high-x, but the valence quarks are more important in this region

The gluon is poorly known at high-x

And it can still be very important for physics e.g.– high ET jet xsecn

need to tie down the high-x gluon





Why are PDF's important

At the LHC high precision (SM and BSM) cross section predictions require precision Parton Distribution Functions (PDFs)

How do PDF uncertainties affect discovery physics?

- Higgs cross-sections
- high ET jets..contact interactions/extra dimensions

Investigate 'standard candle' processes which are insensitive to PDF uncertainties to

- calibrate experiment
- measure machine luminosity?

HERA and the LHC- transporting PDFs to hadron-hadron cross-sections

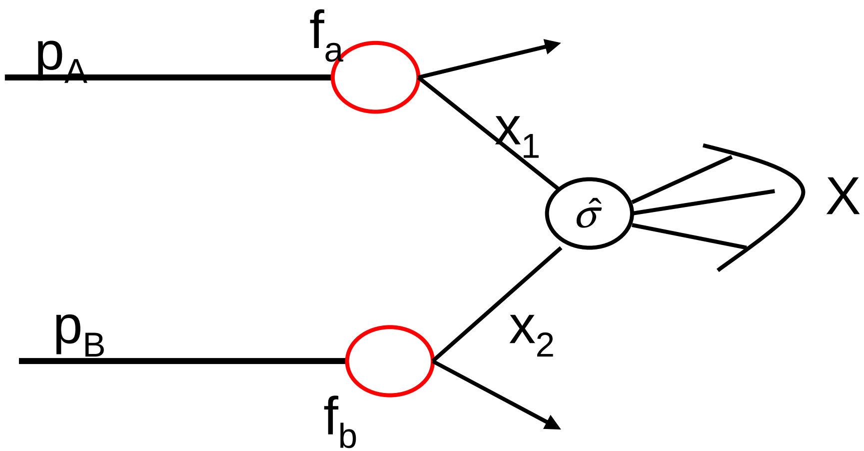
QCD factorization theorem for short-distance inclusive processes

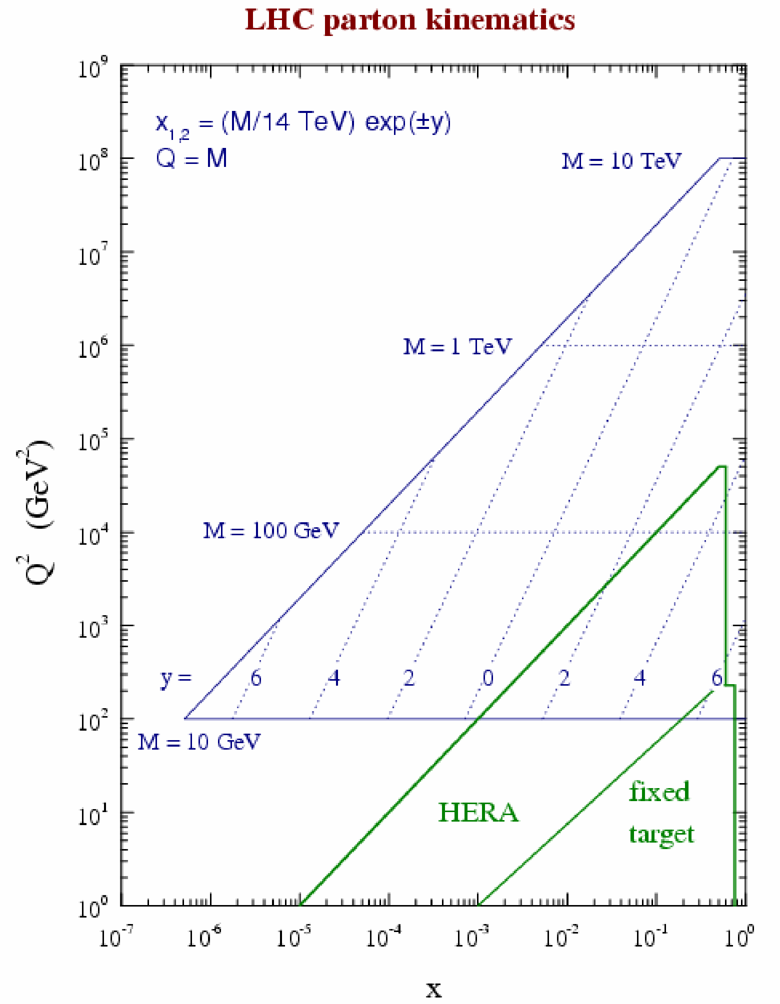
$$\sigma_X = \sum_{a,b} \int_0^1 dx_1 dx_2\, f_a(x_1, \mu_F^2)\, f_b(x_2, \mu_F^2) \times \hat{\sigma}_{ab\to X}\left(x_1, x_2, \{p_i^\mu\}; \alpha_S(\mu_R^2), \alpha(\mu_R^2), \frac{Q^2}{\mu_R^2}, \frac{Q^2}{\mu_F^2}\right)$$

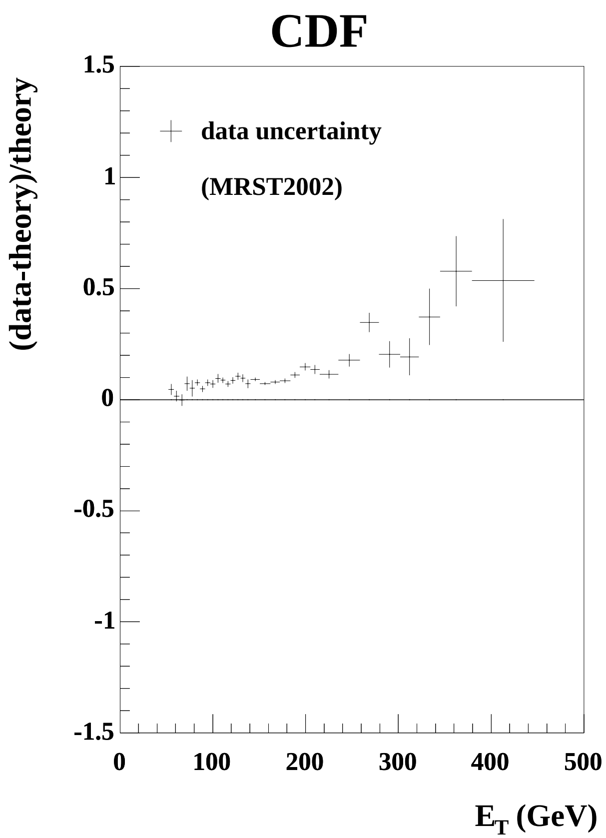
where X=W, Z, D-Y, H, high-E_T jets, prompt-γ

and $\hat{\sigma}$ is known

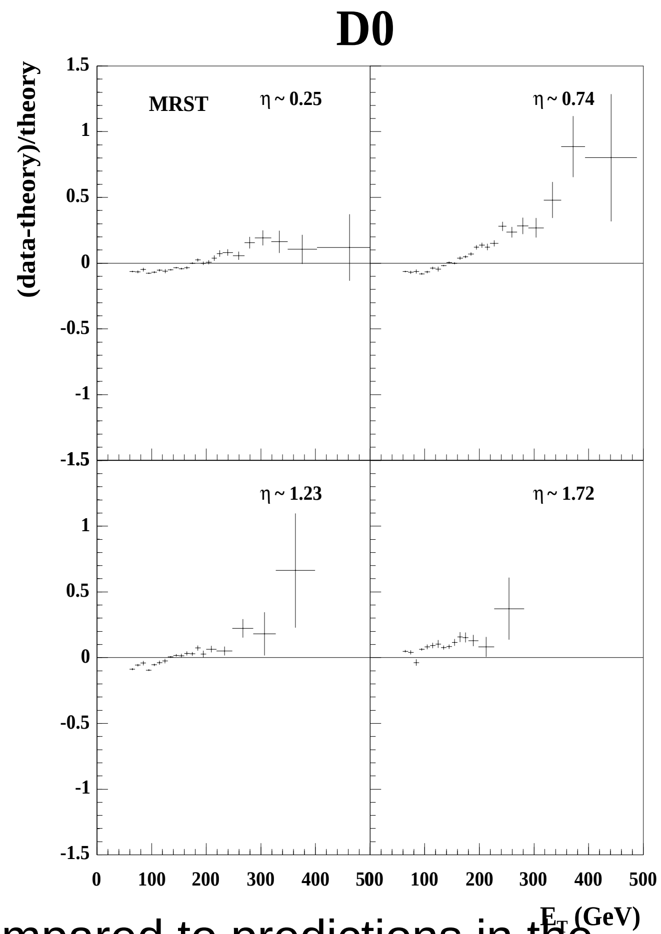
- to some fixed order in pQCD and EW
- in some leading logarithm approximation (LL, NLL, …) to all orders via resummation







Example of how PDF uncertainties matter for BSM physics– Tevatron jet data were originally taken as evidence for new physics--

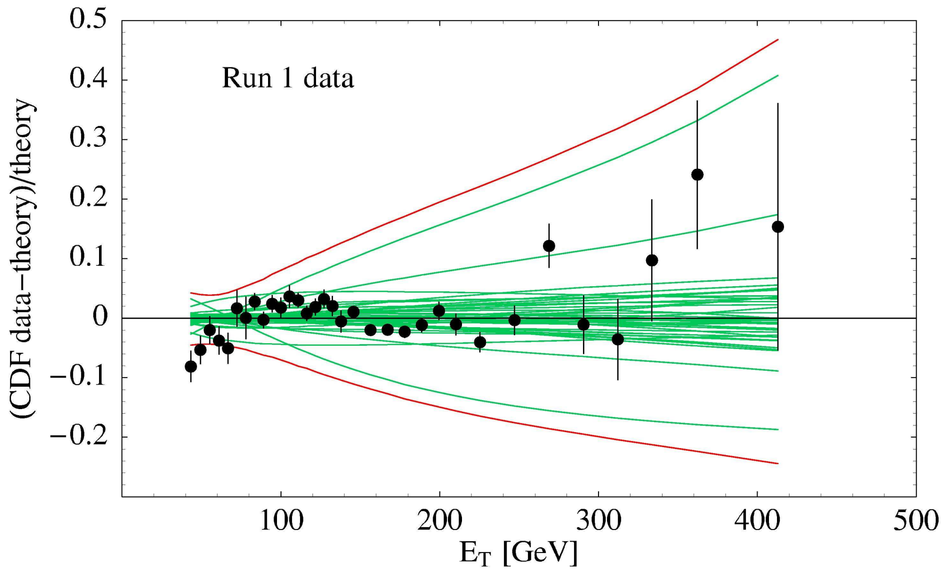


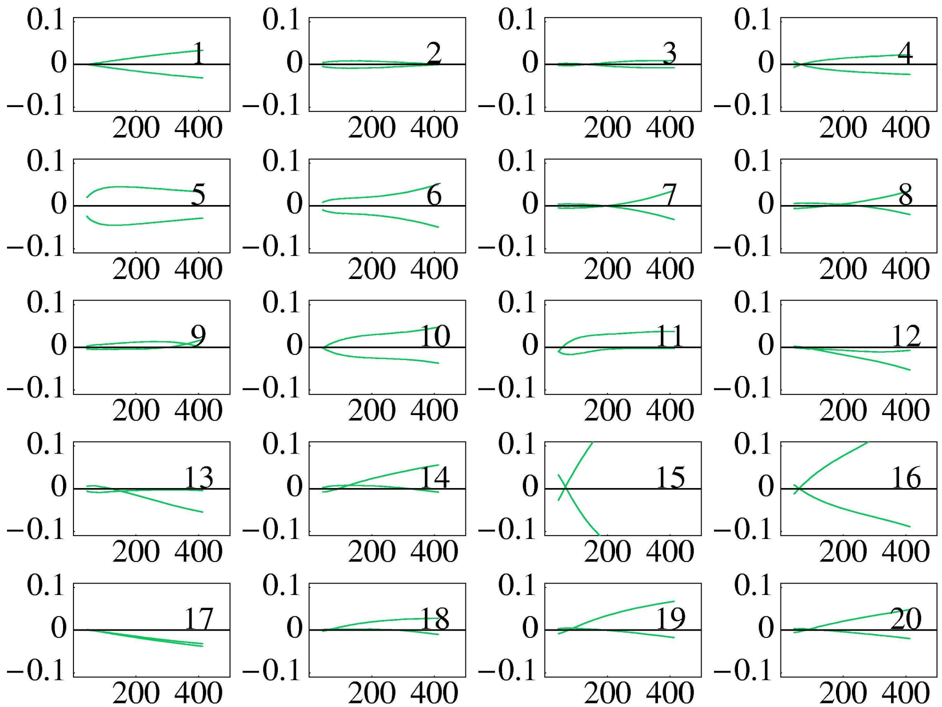
These figures show inclusive jet cross-sections compared to predictions in the form (data - theory)/ theory

Something seemed to be going on at the highest E_T

And special PDFs like CTEQ4/5HJ were tuned to describe it better- note the quality of the fits to the rest of the data deteriorated.

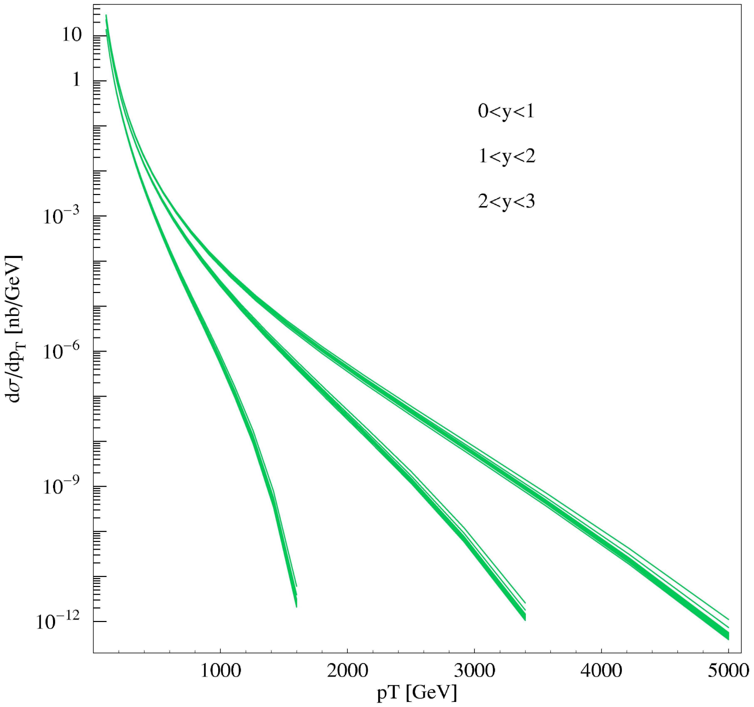
But this was before uncertainties on the PDFs were seriously considered

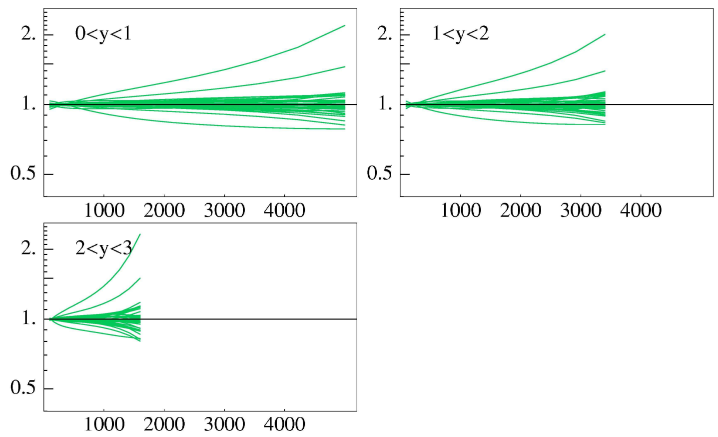




Today Tevatron jet data are considered to lie within PDF uncertainties. (Example from CTEQ hep-ph/0303013)

We can decompose the uncertainties into eigenvector combinations of the fit parameters-the largest uncertainty is along eigenvector 15 –which is dominated by the high x gluon uncertainty

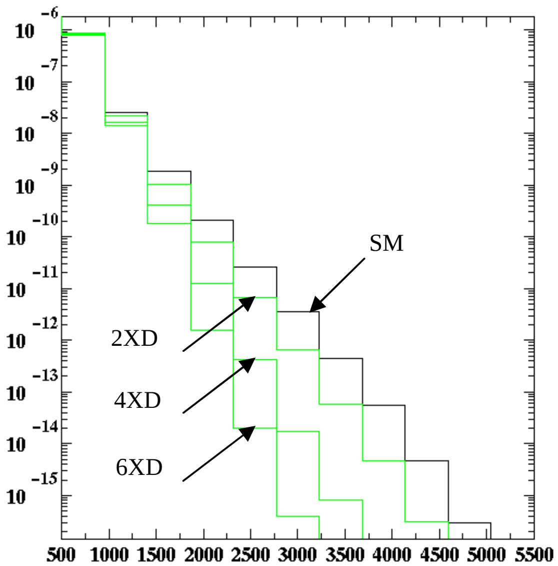


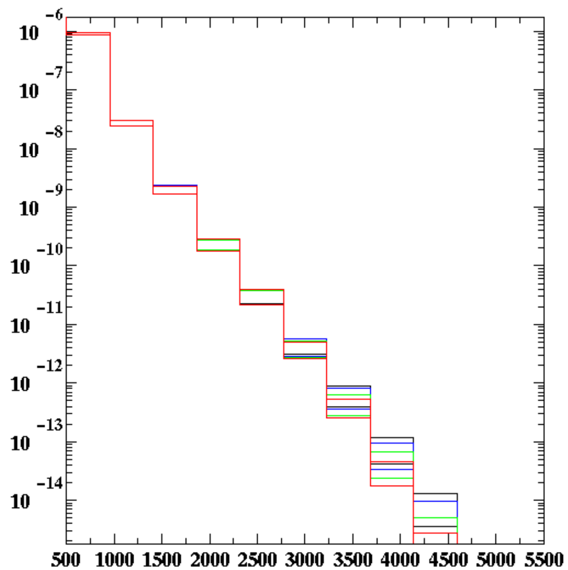


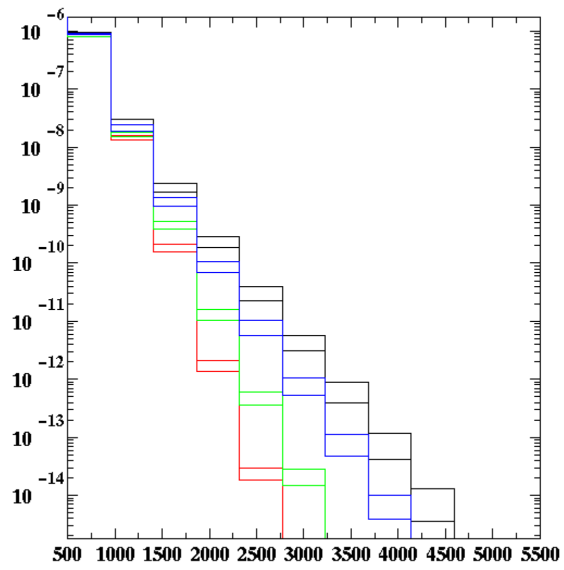
And we can translate the current level of PDF uncertainty into the uncertainty on LHC jet cross-sections. This has consequences for any new BSM physics which can be described by a contact interaction-consider the case of extra dimensions

Such PDF uncertainties on the jet cross sections compromise the potential for discovery.
E.G. Dijet cross section potential sensitivity to compactification scale of extra dimensions (M_c) reduced from ~6 TeV to 2 TeV. (Ferrag et al)

M_c = 2 TeV, no PDF error	M_c = 6 TeV, no PDF error	M_c = 2 TeV, with PDF error

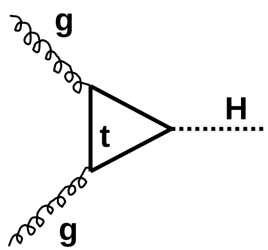




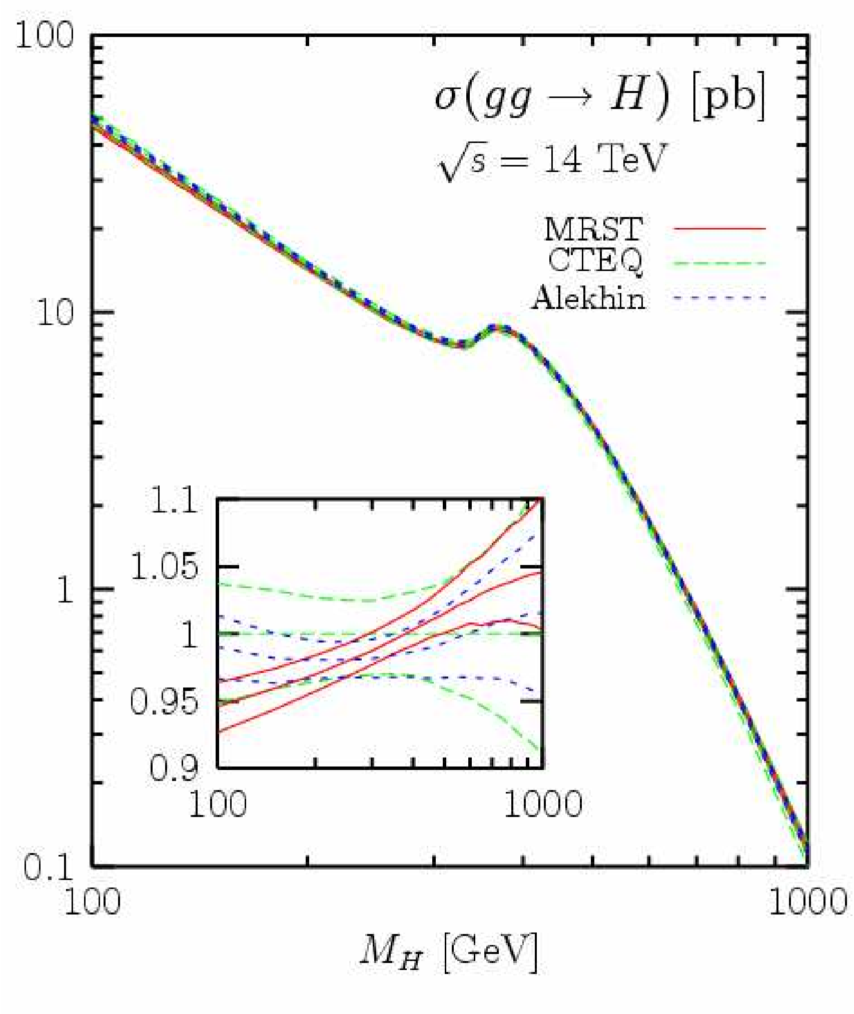


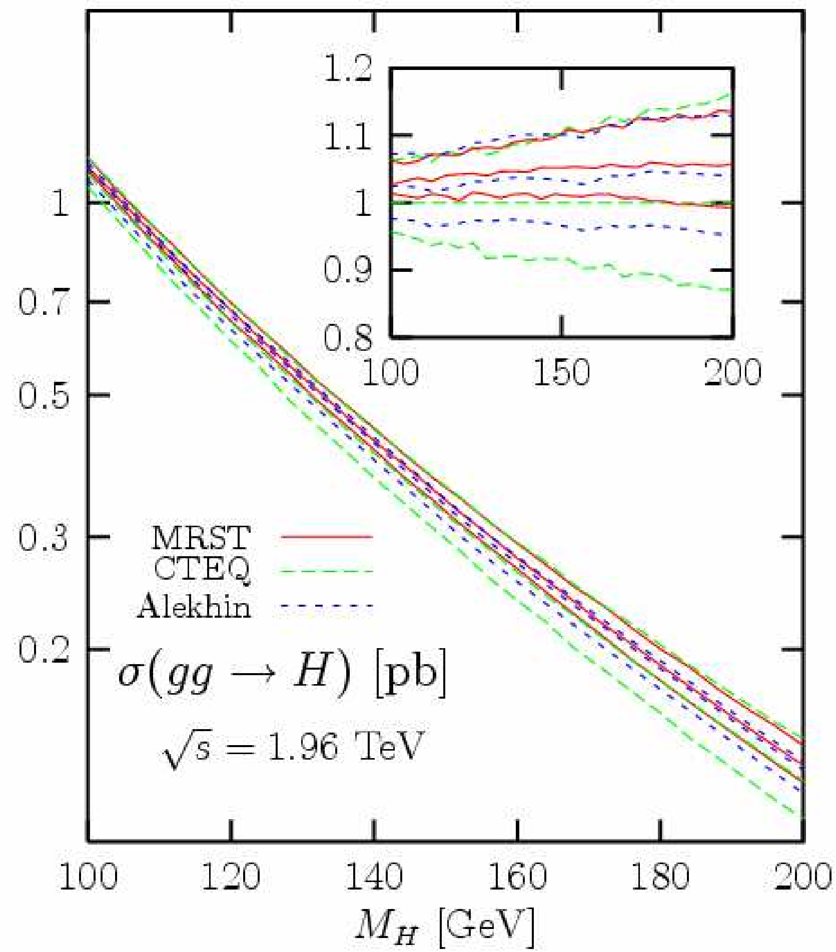
And how do PDF uncertainties affect the Higgs discovery potential?

Higgs at LHC

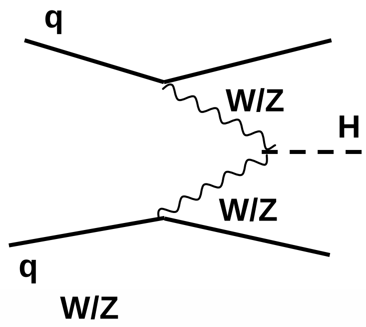


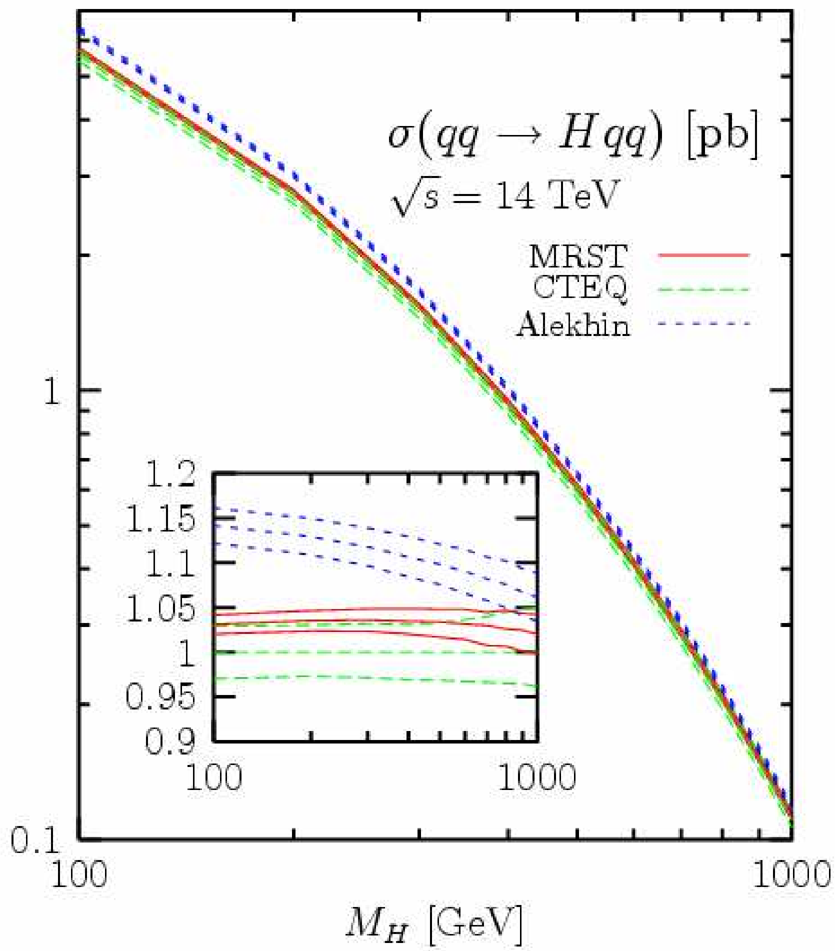
Higgs at Tevatron

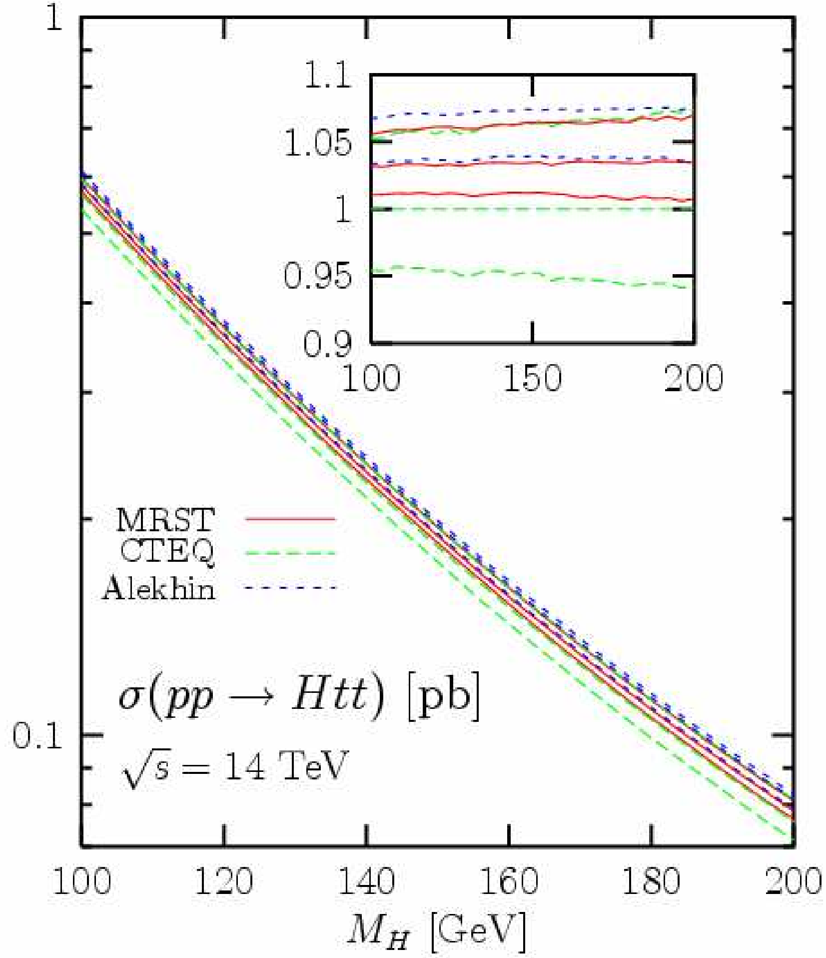




Higgs from qq at LHC







Good news: PDF uncertainties will decrease before LHC comes on line

HERA-II and Tevatron Run-II will improve our knowledge

•HERA now in second stage of operation (HERA-II)

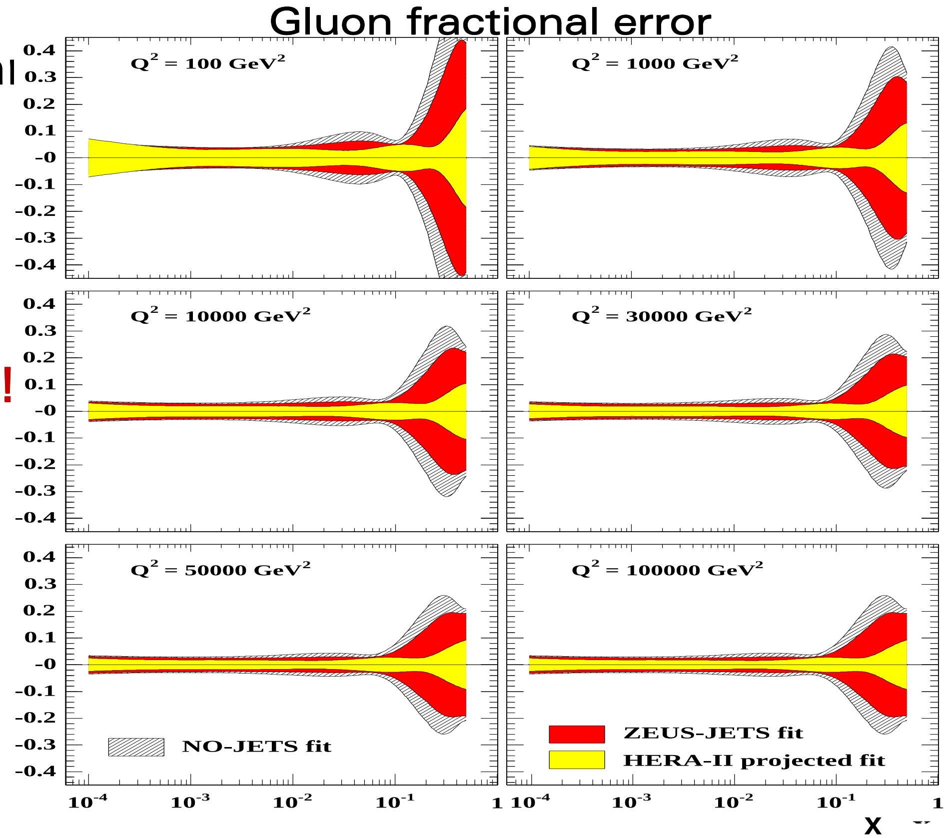
substantial increase in luminosity

possibilities for new measurements

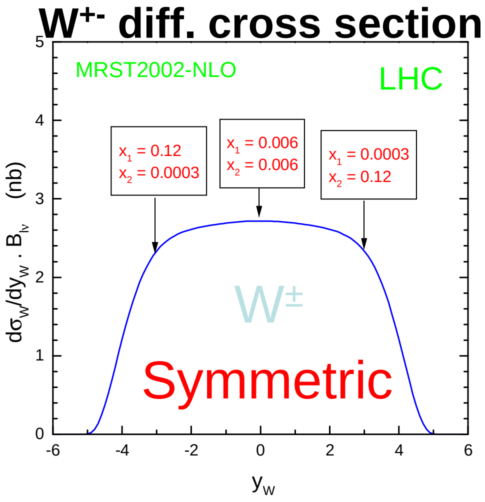
HERA-II projection shows significant improvement to high-x PDF uncertainties

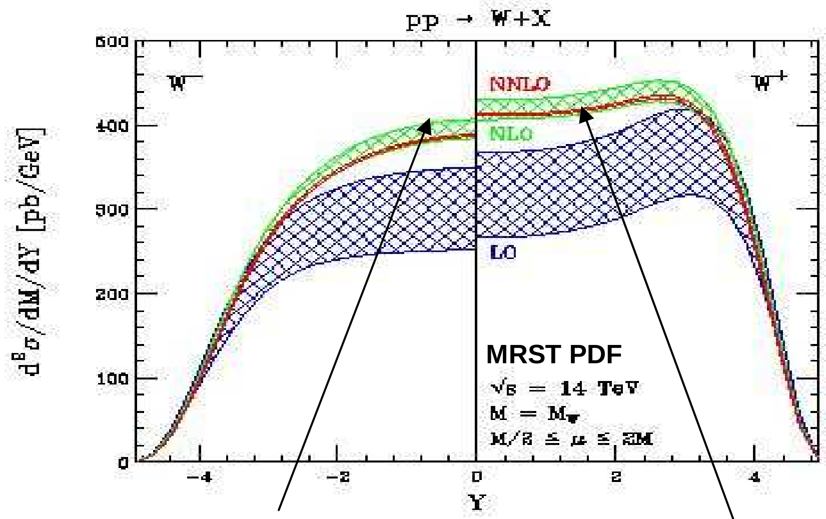
⇒ relevant for high-scale physics at the LHC

→ where we expect new physics !



Example of how PDF uncertainties matter for SM physics: W/Z production have been considered as good standard candle processes possibly even monitors of the luminosity **But are they really well known cross-sections?**





NNLO corrections small ~ few%

NNLO residual scale dependence < 1%

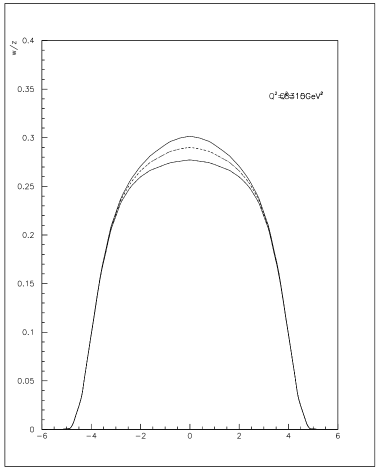
PDF Set	$\sigma_{W^+} \cdot B_{W\to l\nu}$ (nb)	$\sigma_{W^-} \cdot B_{W\to l\nu}$ (nb)	$\sigma_Z \cdot B_{Z\to ll}$ (nb)
ZEUS-S	12.07 ± 0.41	8.76 ± 0.30	1.89 ± 0.06
CTEQ6.1	11.66 ± 0.56	8.58 ± 0.43	1.92 ± 0.08
MRST01	11.72 ± 0.23	8.72 ± 0.16	1.96 ± 0.03

Theoretical uncertainties dominated by PDFs **note that** **central values differ by more than the MRST estimate of the error**

To improve the situation we NEED to be more accurate than this:**~3%**

Statistics are no problem we are dominated by systematic uncertainty

The **uncertainty on the W/Z rapidity distributions** is dominated by –- low-x **gluon PDF** dominated eigenvectors

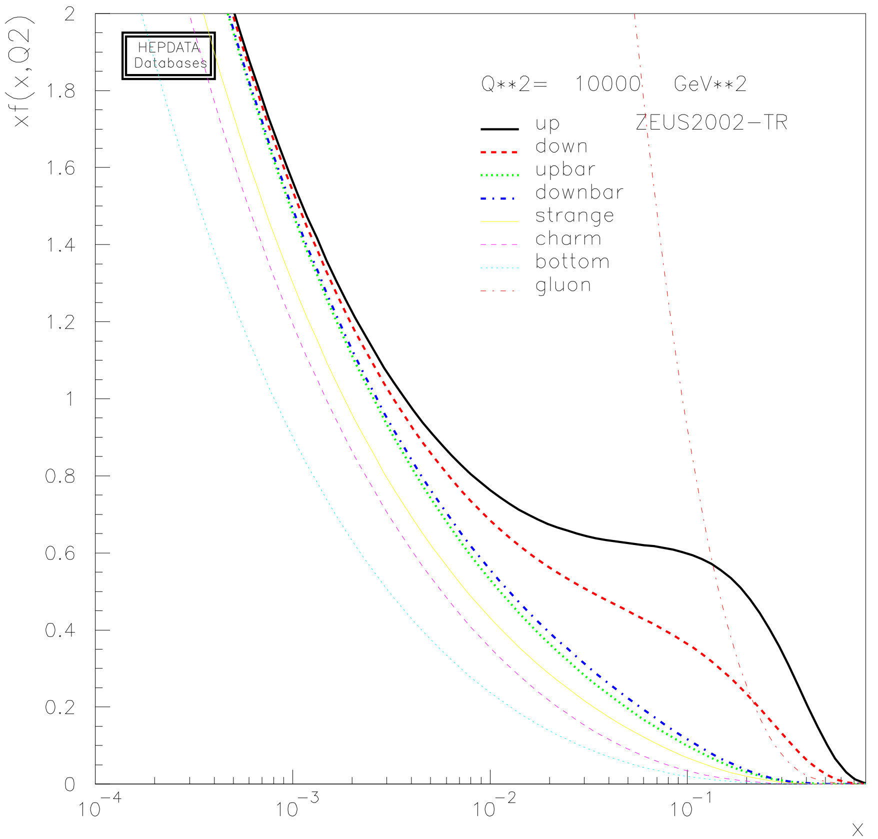


Differences are visible within the measurable range of rapidity

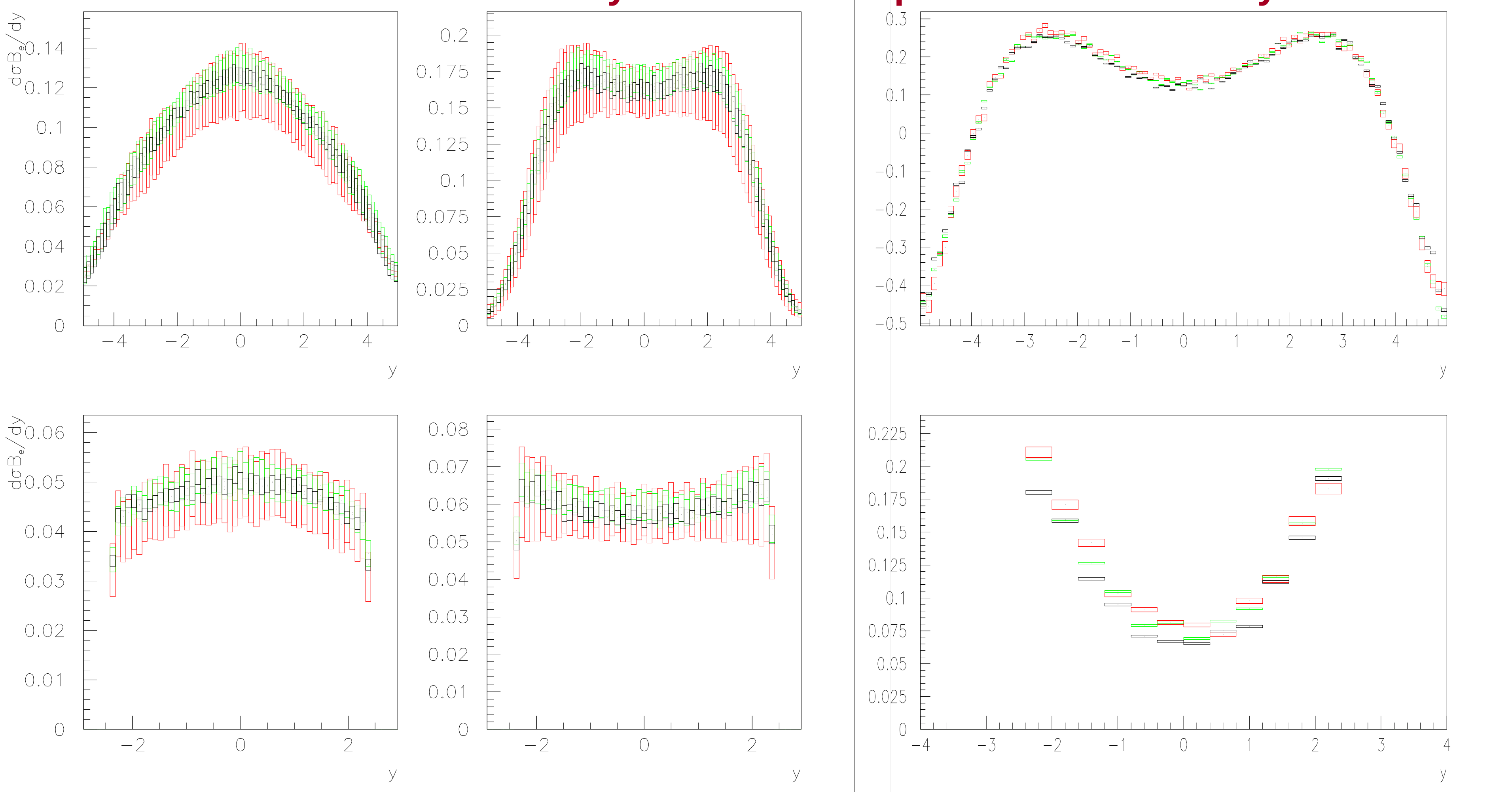
It may at first sight be surprising that **W/Z distns are sensitive to gluon parameters** BUT our experience is based on the Tevatron where Drell-Yan processes can involve valence-valence parton interactions.

At the LHC we will have dominantly sea-sea parton interactions at low-x

And at Q2~M_Z2 **the sea is driven by the gluon-** which is far less precisely determined for all x values



For the W+/W- we will actually observe the leptons from the decays



Look at the lepton rapidity spectra and asymmetry at generator level -TOP

and after passing through ATLFAST –BOTTOM

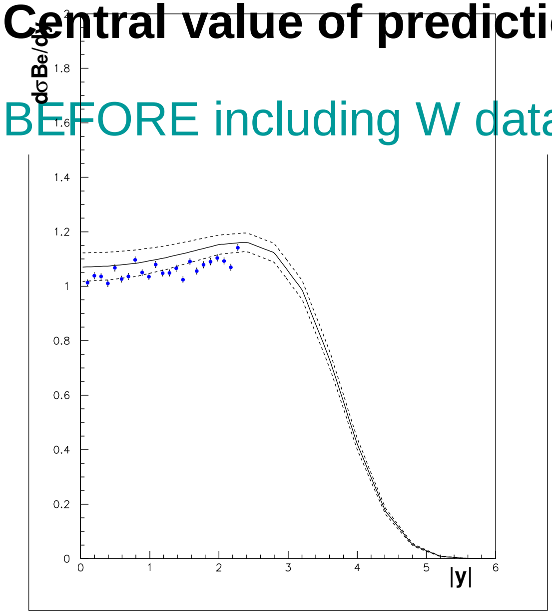
Generation with HERWIG+k-factors using CTEQ6.1M ZEUS_S MRST2001 PDFs with full uncertainties

Study of the effect of including the LHC W Rapidity distributions in global PDF fits **by how much can we reduce the PDF errors with early LHC data?**

Generate data with CTEQ6.1 PDF, pass through ATLFAST detector simulation and then include this pseudo-data in the global ZEUS PDF fit.

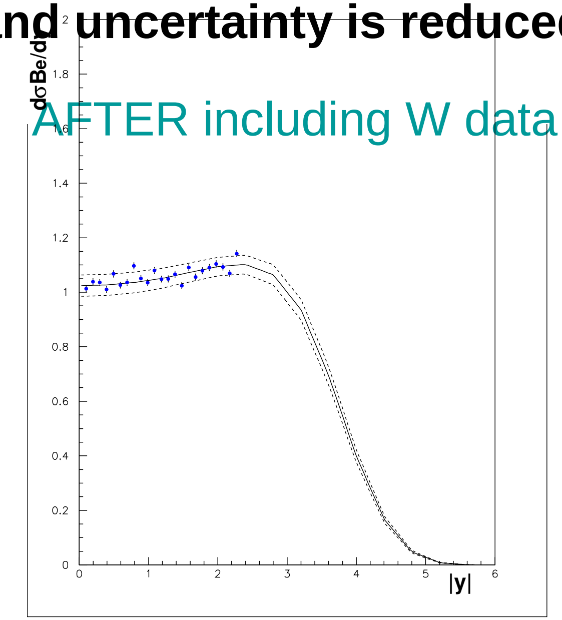
Central value of prediction shifts and uncertainty is reduced

BEFORE including W data



W+ to lepton rapidity spectrum data generated with CTEQ6.1 PDF compared to predictions from ZEUS PDF

AFTER including W data



W+ to lepton rapidity spectrum data generated with CTEQ6.1 PDF compared to predictions from ZEUS PDF AFTER these data are included in the fit

Specifically the low-x gluon shape parameter λ, $xg(x) = x^{-\lambda}$, was
λ = -.199 ± .046 for the ZEUS PDF before including this pseudo-data
It becomes λ = -.181 ± .030 after including the pseudodata

LHC is a low-x machine (at least for the early years of running)

Low-x information comes from evolving the HERA data

Is NLO (or even NNLO) DGLAP good enough?

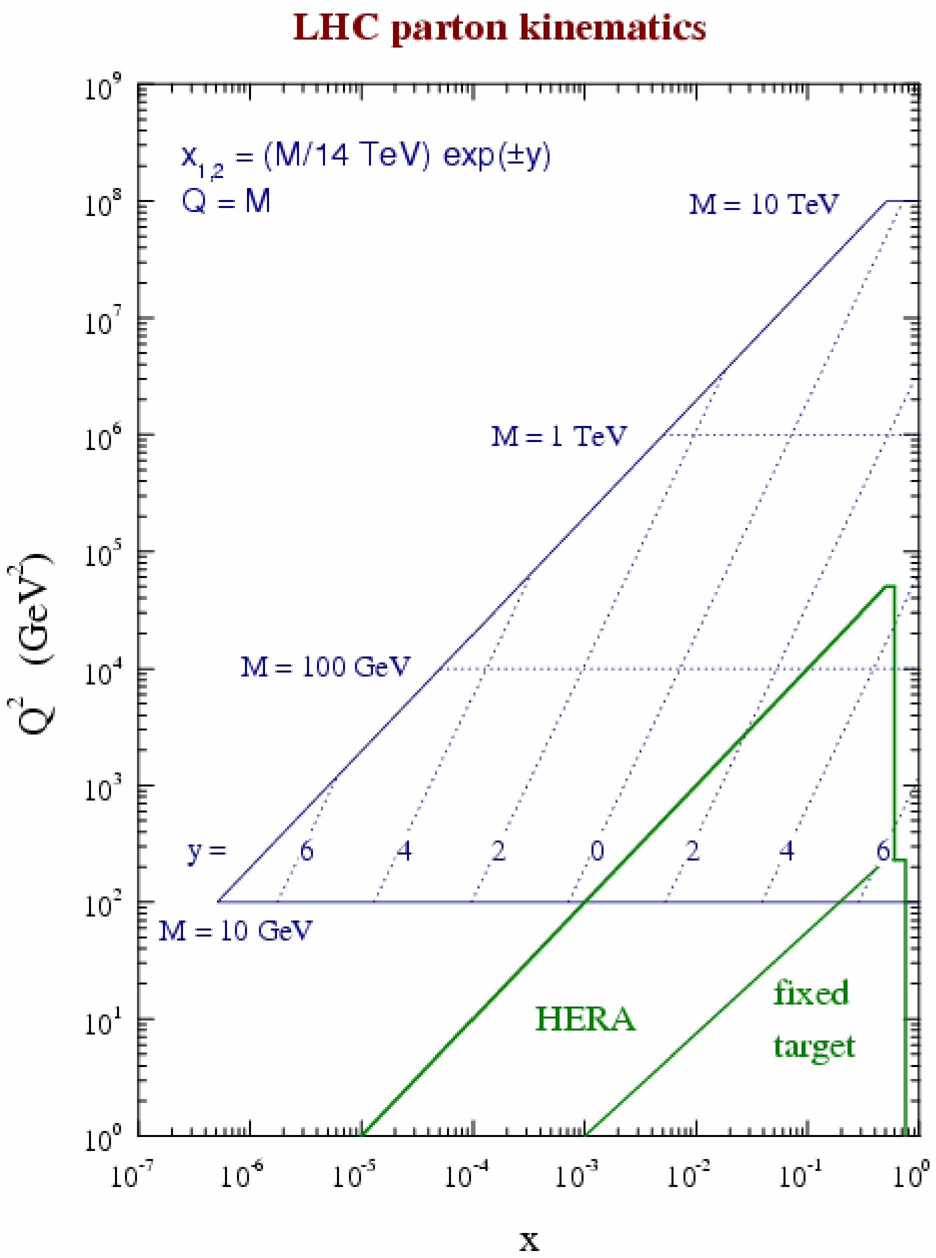
The QCD formalism may need extending at small-x

BFKL ln(1/x) resummation

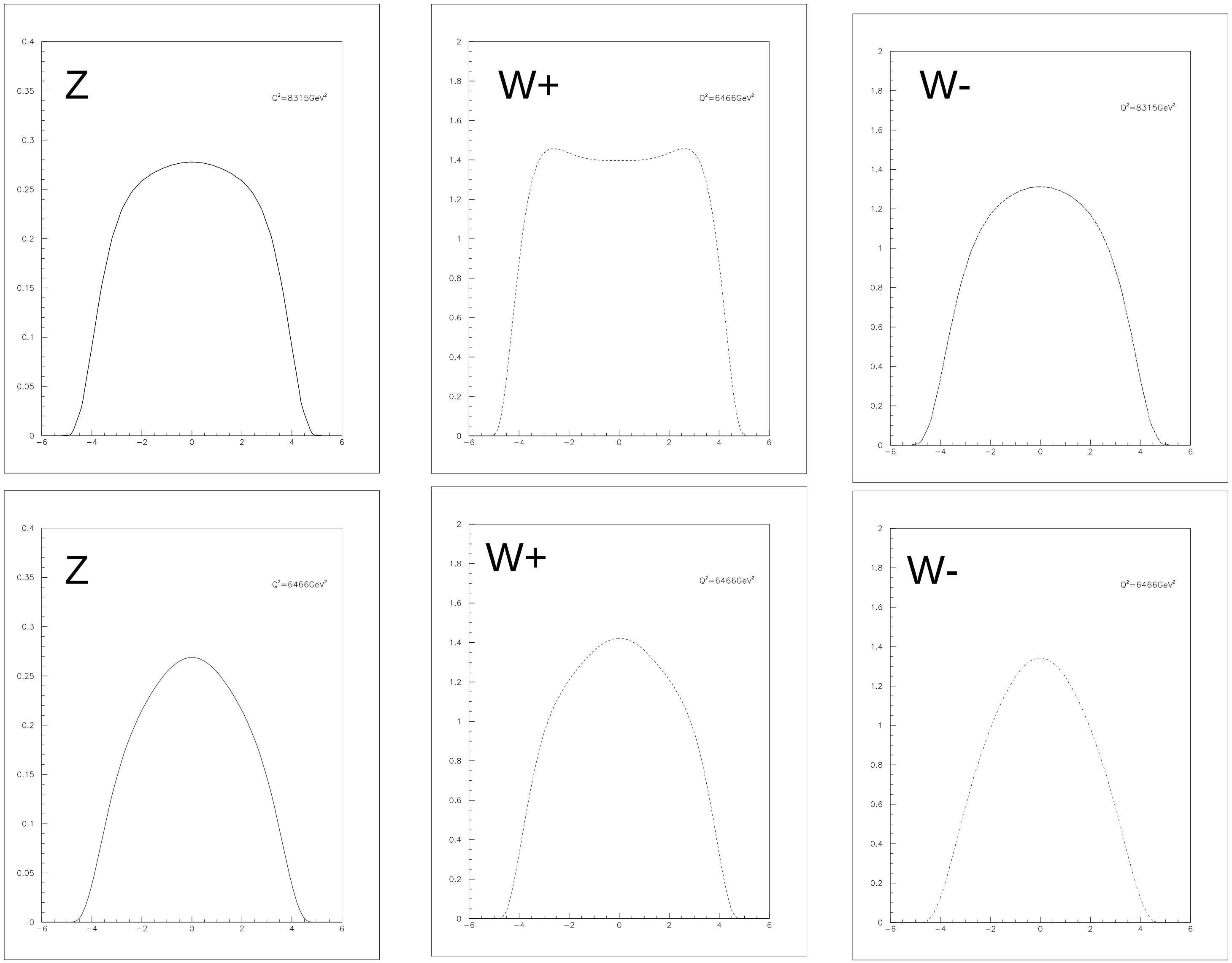
High density non-linear effects etc.

(Devenish and Cooper-Sarkar, ‘Deep Inelastic Scattering’, OUP 2004, Section 6.6.6 and Chapter 9 for details!)

Thorne will talk about this



MRST have produced a set of PDFs derived from a fit without low-x data –ie do not use the DGLAP formalism at low-x- called MRST03 ‘conservative partons’. These give VERY different predictions for W/Z production to those of the ‘standard’ PDFs.



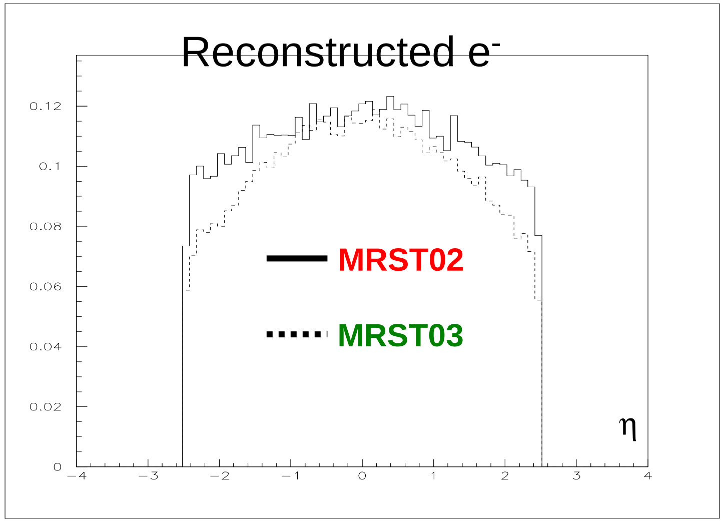
MRST02

MRST03

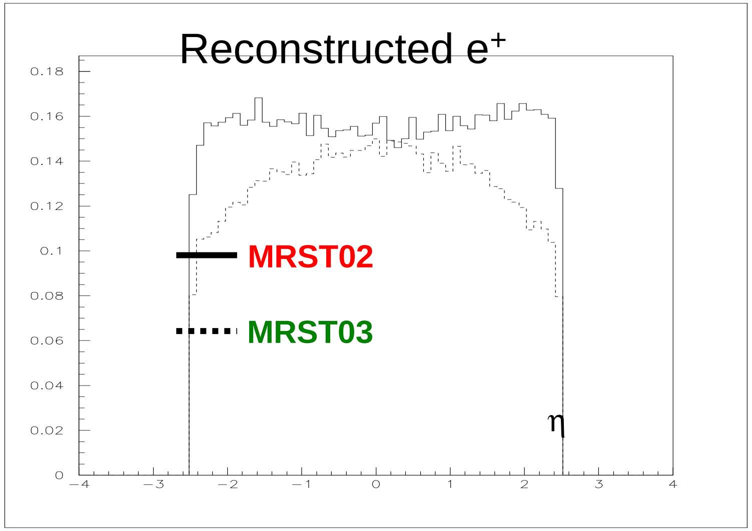
Differences persist in the decay lepton spectra and even in their ratio and asymmetry distributions

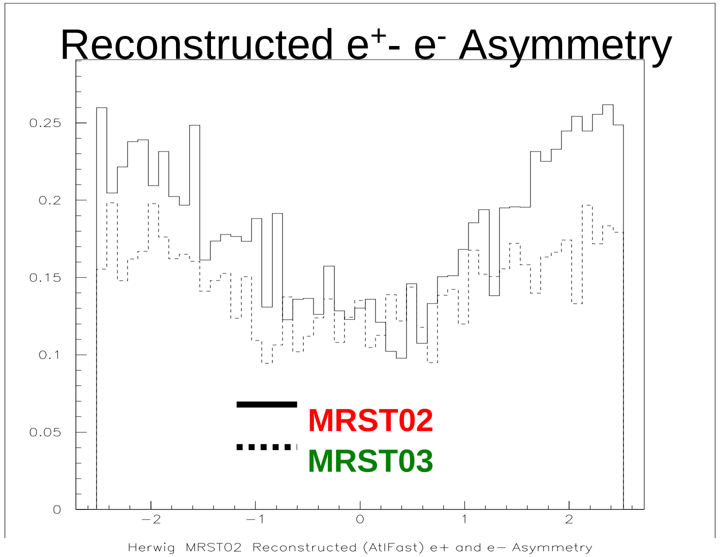
Reconstructed Electron Pseudo-Rapidity Distributions (ATLAS fast simulation)

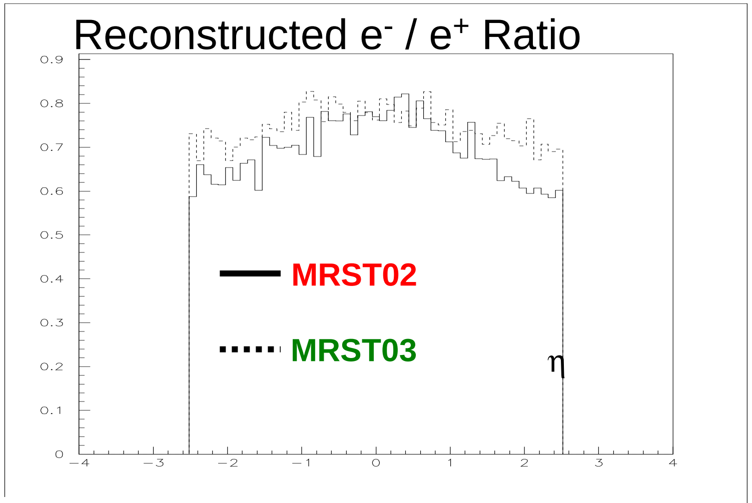
200k events of $\mathbf{W^{+-} \to e^{+-}}$ generated with HERWIG 6.505 + NLO K factors



6 hours running







Note of caution. MRST03 conservative partons DO NOT describe the HERA data for x< 5 10^{-3} which is not included in the fit which produces them. So there is no reason why they should correctly predict LHC data at non-central y, which probe such low x regions.

What is really required is an alternative theoretical treatment of low-x evolution which would describe HERA data at low-x, and could then predict LHC W/Z rapidity distributions reliably – also has consequences for pt distributions.

The point of the MRST03 partons is to illustrate that this prediction COULD be very different from the current ‘standard’ PDF predictions. When older standard predictions for HERA data were made in the early 90’s they did not predict the striking rise of HERA data at low-x. This is a warning against believing that a current theoretical paradigm for the behaviour of QCD at low-x can be extrapolated across decades in Q2 with full confidence.

→ **The LHC measurements may also tell us something new about QCD**

Summary

Parton distributions are extracted from NLOQCD fits to DIS data- But they are needed for predictions of all cross-sections involving hadrons.

I have introduced you to the history of this in order to illustrate that it's not all cut and dried- our knowledge evolves continually as new data come in to confirm or confound our input assumptions

You need to appreciate the sources of uncertainties on PDFs – experimental, model and theoretical- in order to appreciate how reliable predictions for interesting collider cross-sections are.

At the LHC high precision (SM and BSM) cross section predictions require precision Parton Distribution Functions

We will improve our current knowledge from the HERA data, and the Tevatron data, before the LHC turns on

We can begin LHC physics by measuring 'standard candle' processes which are insensitive to PDF uncertainties

We can even use early LHC measurements, at low scales where BSM physics is not expected, to increase precision on PDFs and thus improve limits for discovery physics

But there is some possibility that the Standard Model is wrong not due to exciting exotic physics, but because the standard QCD framework is not fully developed at small-x, hence we may first learn more about QCD!