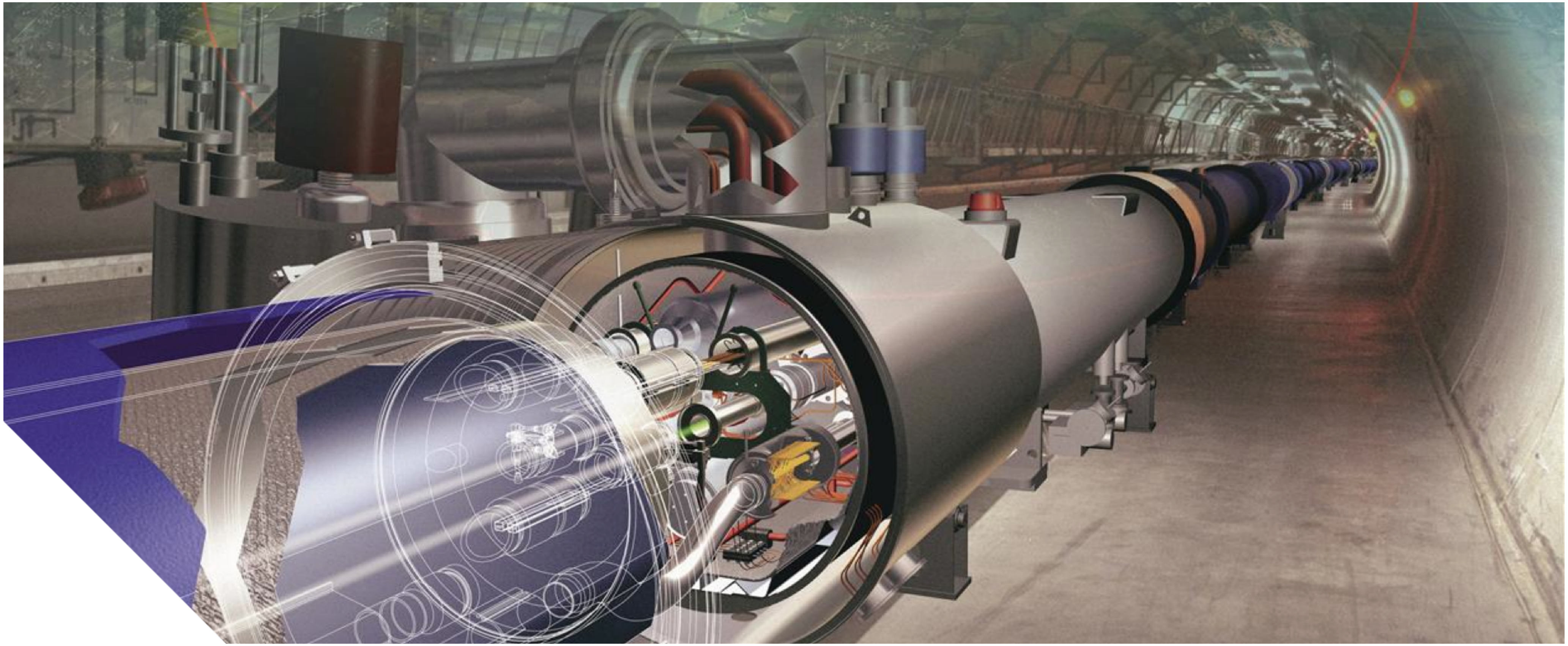


QCD – Lecture 4

PDFs and their uncertainties

Claire Gwenlan, Oxford, HT

PDF determination

How do we determine Parton Distributions? They are not perturbatively calculable …
(Lattice Gauge Theory not yet good enough…) Must extract from experimental measurements!

- Parametrise PDFs at some low starting scale **$Q0^2$** (typically a few GeV^2), so $\alpha s(Q0^2)$ small
- Evolve with DGLAP equations to **$Q^2 > Q0^2$** and confront with data via a **X^2** fit

General Parametrisation:

$$x f_i(x, Q_0^2) = A_{f_i} x^{a_{f_i}} (1-x)^{b_{f_i}} \mathscr{F}(x, \{c_{f_i}\})$$

$\mathscr{F}(x, \{c_{f_i}\})$ is a smooth function which remains finite both when $x \to 0$ and $x \to 1$

One specific example:

$$x u_v(x) = A_u x^{a_u}(1-x)^{b_u} \left(1 + \epsilon_u \sqrt{x} + \gamma_u x\right)$$
$$x d_v(x) = A_d x^{a_d}(1-x)^{b_d} \left(1 + \epsilon_d \sqrt{x} + \gamma_d x\right)$$
$$x S(x) = A_s x^{-a_s}(1-x)^{b_s} \left(1 + \epsilon_s \sqrt{x} + \gamma_s x\right)$$
$$x g(x) = A_g x^{-a_g}(1-x)^{b_g} \left(1 + \epsilon_g \sqrt{x} + \gamma_g x\right)$$

$\mathscr{F}(x, \{c_{f_i}\})$

EG. simple polynomials, Chebyshev polynomials, Bernstein polynomials; multi-layer NNs; …

- **Not all parameters are independent:**
- Ag is determined from the momentum sum rule

$$\int_0^1 \sum_{q,\bar{q}} (x q(x) + x \bar{q}(x)) + x g(x)\, dx = 1$$

- Au, Ad from the number sum rules

$$\int_0^1 u_v(x)\, dx = 2 \quad \int_0^1 d_v(x)\, dx = 1$$

- Various other choices restrict parameters

Then measurable quantities like, $F_2,\ xF_3$ for $\nu,\ \bar{\nu},\ e^{\pm},\ \mu^{\pm} \rightarrow p, D$ NC CC

depend on a finite number of parameters (~15-20)

These structure functions are measured over a very wide, (x,Q^2) range
~2500 data points

So you evolve the partons – using the DGLAP equations – to a Q^2 value at which you have data and then you predict the measured structure functions from them:

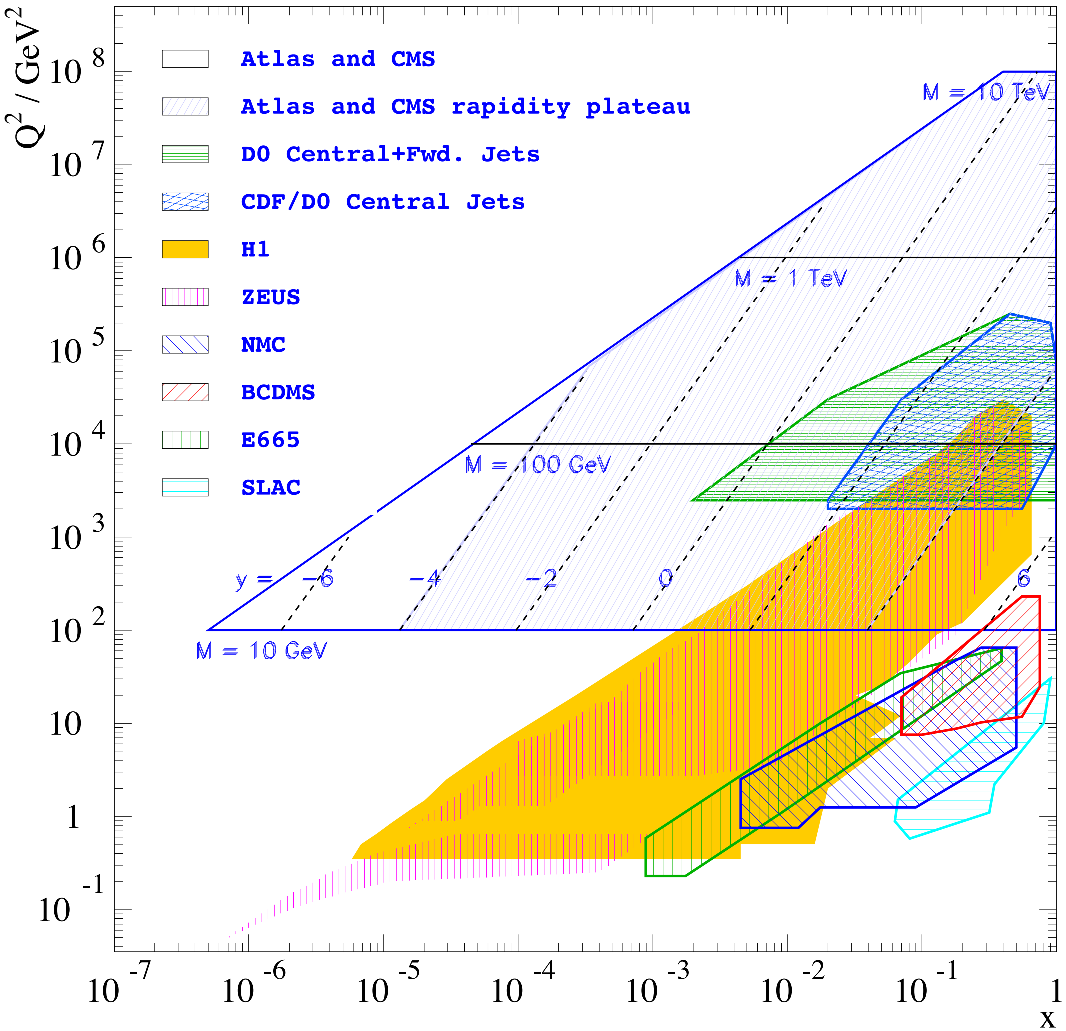
Simply at LO

And by convolution with QCD calculable coefficient functions at NLO and NNLO

Then you fit the data to determine the parameters of the PDFs

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)

(x,Q^2) kinematic coverage



Note that, for now, we concentrate on constraints from **DIS**; we will return to talk about additional constraints provided by **pp** collision measurements in a later lecture

Traditionally

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

$$F_2(e/\mu_p) \sim \frac{4}{9}x(u+\bar{u}) + \frac{1}{9}x\left(d+\bar{d}\right) + \frac{4}{9}x(c+\bar{c}) + \frac{1}{9}x(s+\bar{s})$$

$$F_2(e/\mu_D) \sim \frac{5}{18}x\left(u+\bar{u}+d+\bar{d}\right) + \frac{4}{9}x(c+\bar{c}) + \frac{1}{9}x(s+\bar{s})$$

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

Also use ν, $\bar{\nu}$ fixed target data from **CCFR, NUTEV, CHORUS , (NOMAD)**
(Beware, Fe target needs corrections; even deuterium is not safe)

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

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

We have 4 equations so we can get ~4 distributions from this: E.G. $u,\ d,\ \bar{u},\ \bar{d}$

BUT note we have assumed:

- u in proton = d in neutron and q=qbar in the sea (in practice violations are very small)
- And need further assumptions like sbar = ¼ (ubar+dbar) and a heavy quark treatment

→ The assumption on sbar is questionable

→ But the heavy quark contributions can be calculated from pQCD

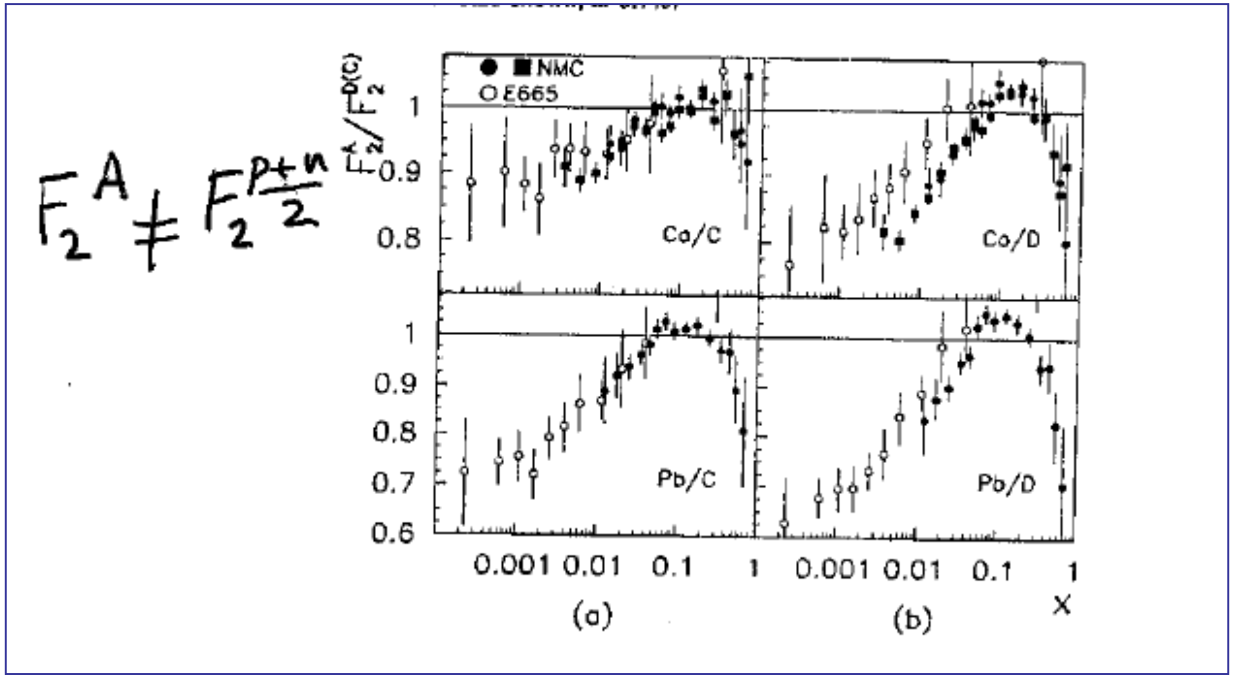
Note, the **gluon** enters **indirectly** via the DGLAP equations for Q^2 evolution, and **directly** in the longitudinal structure function F_L at $\mathcal{O}(\alpha_s)$

The 4 equations could also come from e+/e- NC/CC scattering on pure proton target – HERA

And why might we want to do that?

Because of this – the EMC effect

Heavy targets – and even deuterium – require uncertain nuclear corrections.



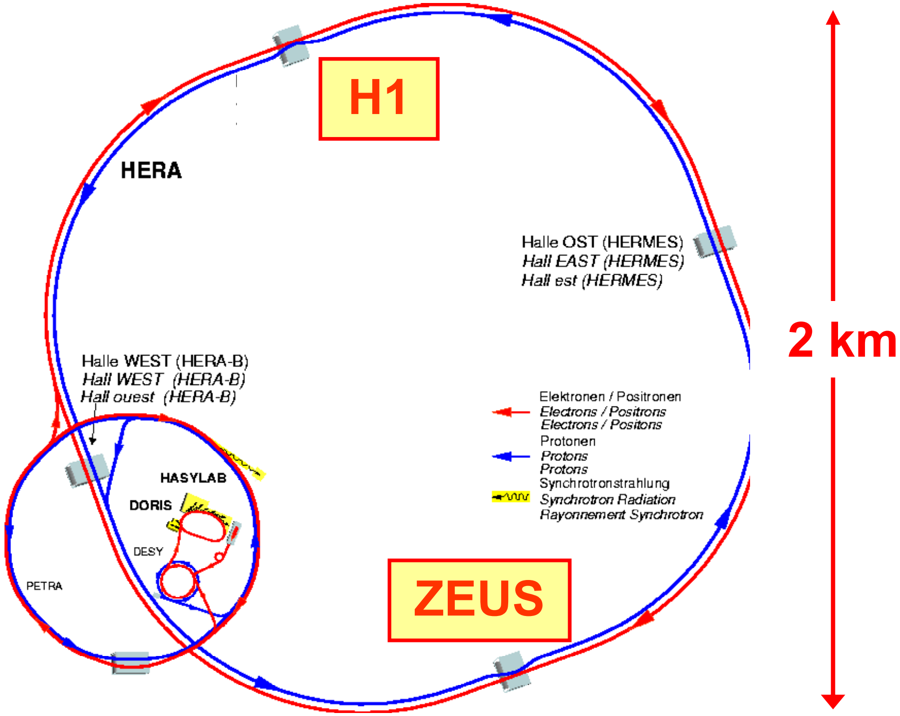
But with *e p* scattering you can get 4 equations at high energy because you need *W, Z* as well as γ exchange.

HERA (1992 – 2007)

- DESY (Deutsches Elektronen-Synchrotron) Laboratory, Hamburg, Germany

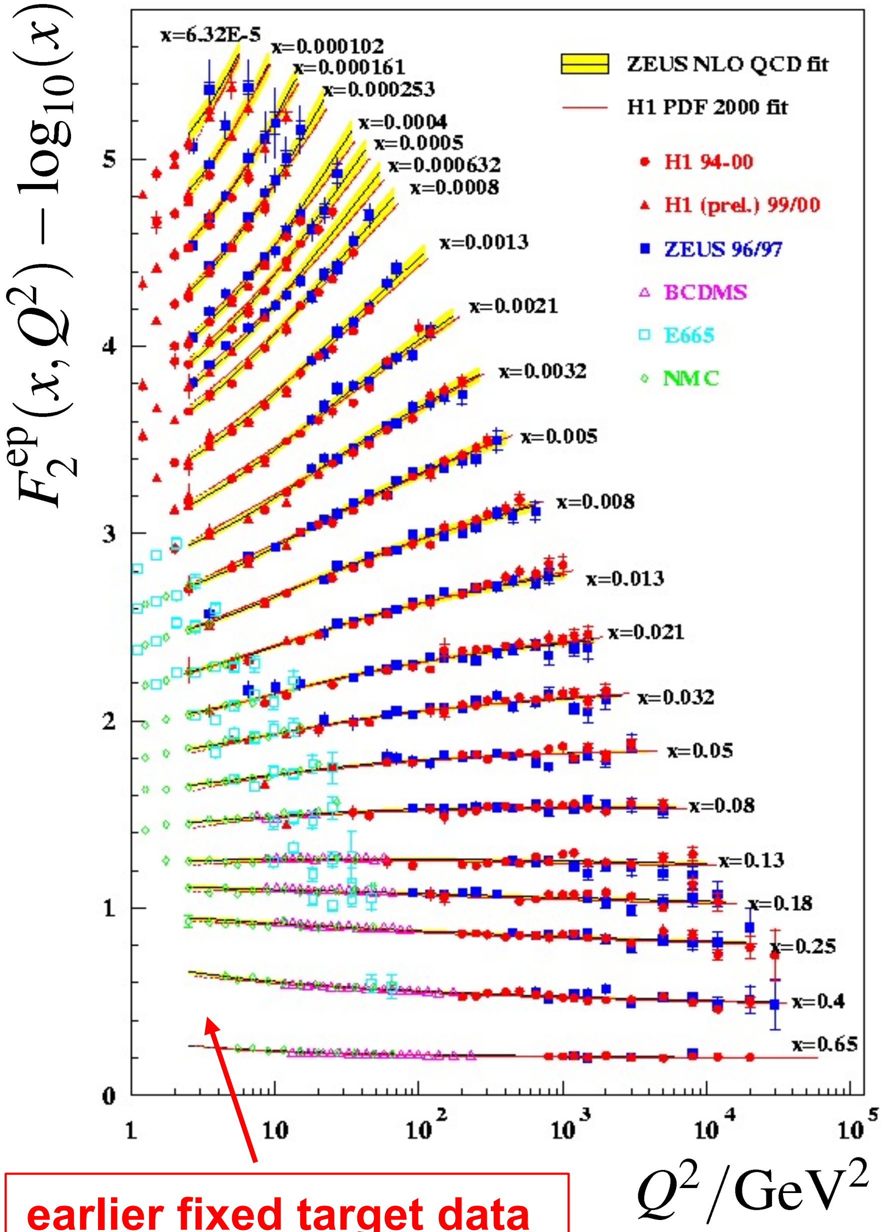
$e^{\pm}$ 27.5 GeV → ← 820 – 920 GeV p

$\sqrt{s}$ = 300 GeV

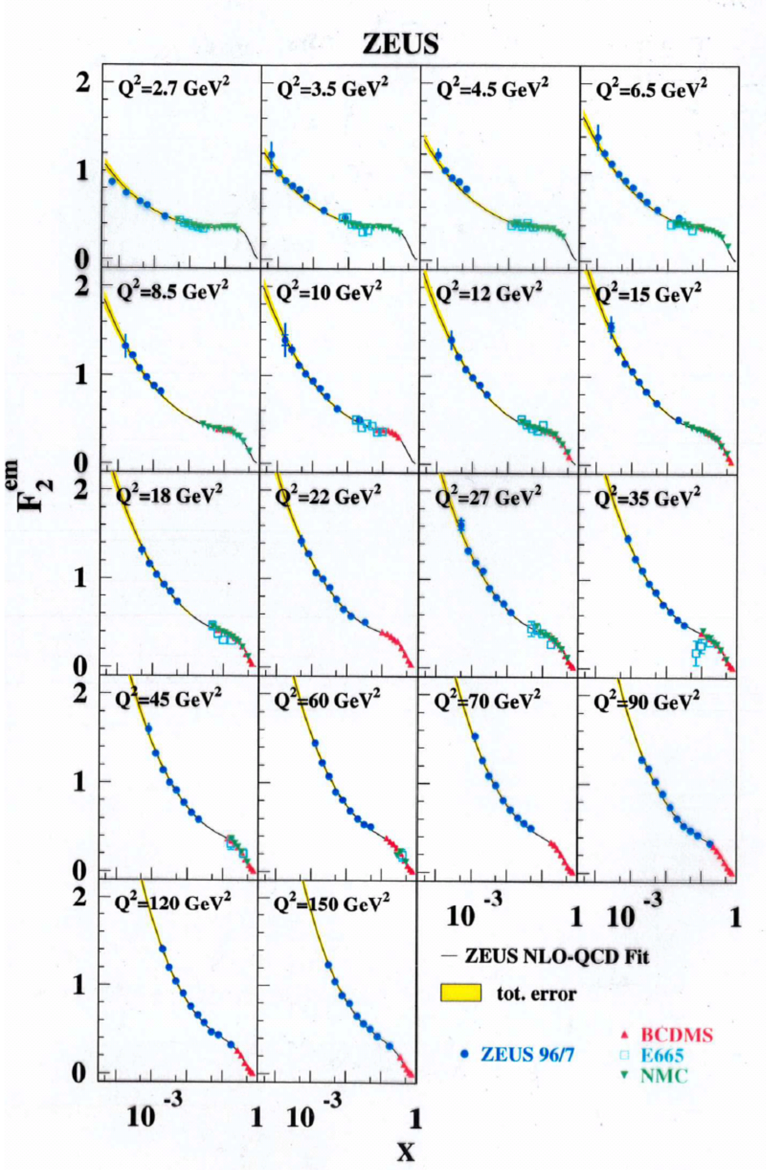


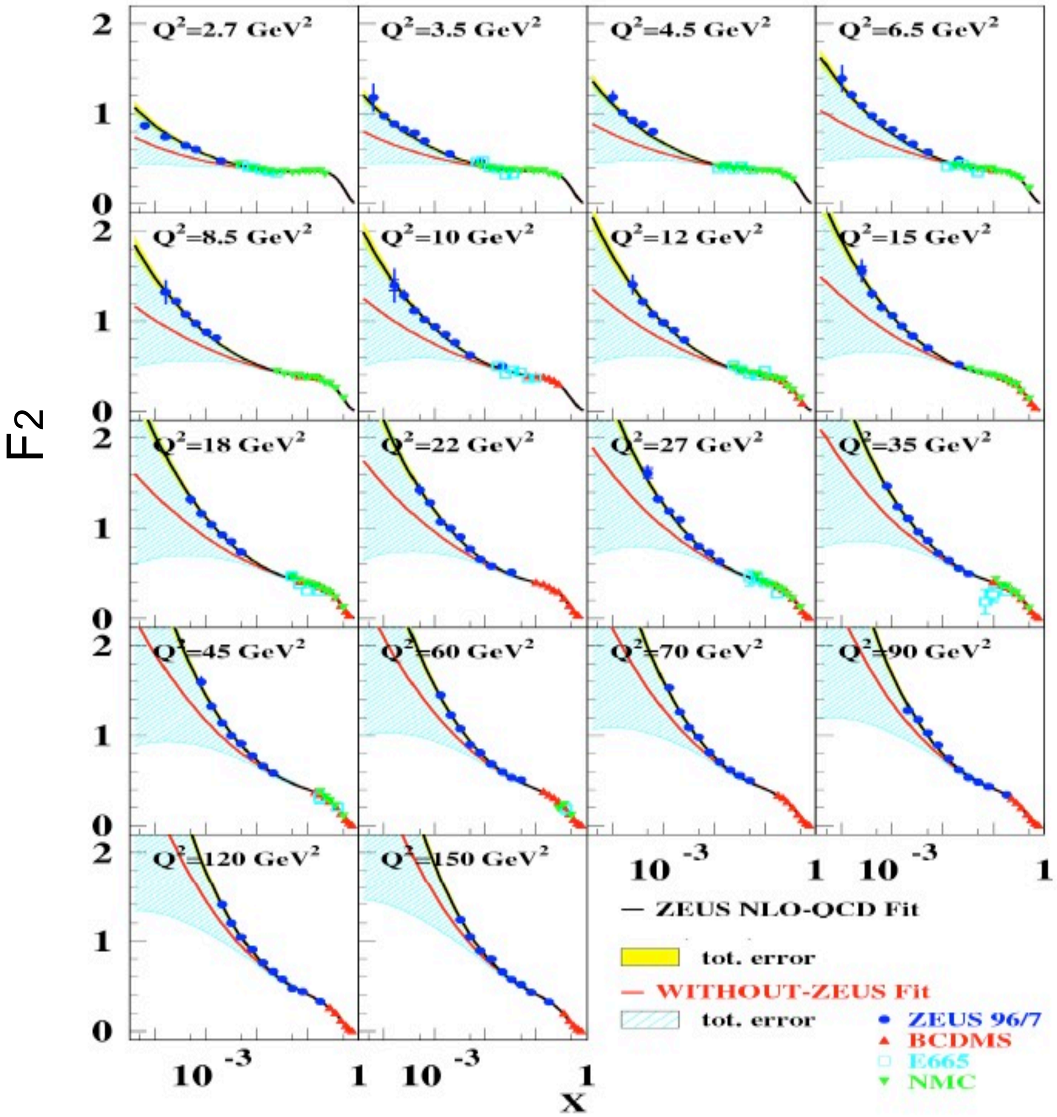
- two large experiments: H1 and ZEUS
- probe proton at very high Q^2 and very low x

F2(x,Q²) from HERA



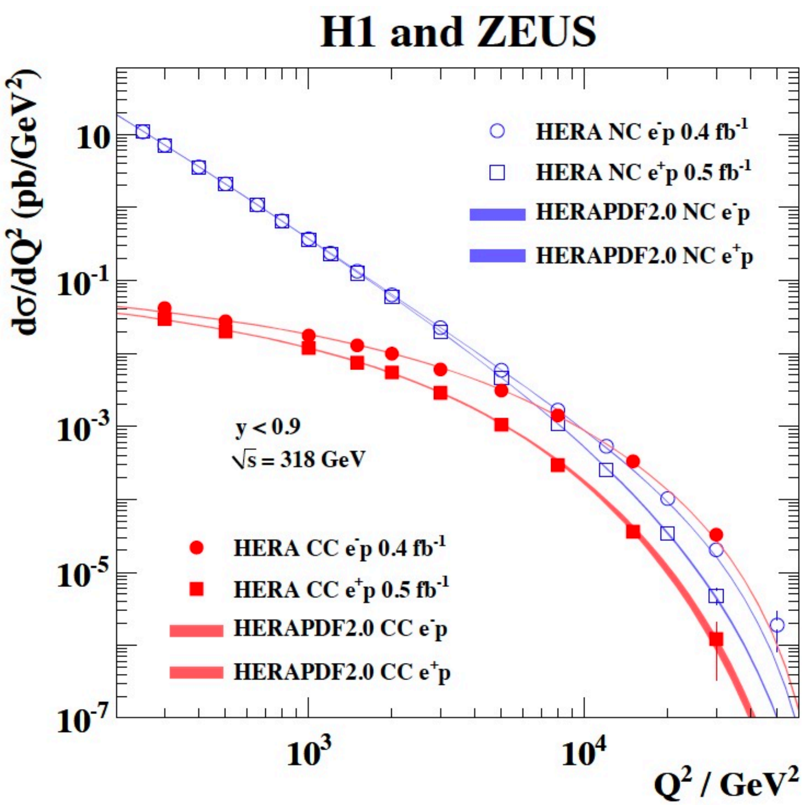
earlier fixed target data

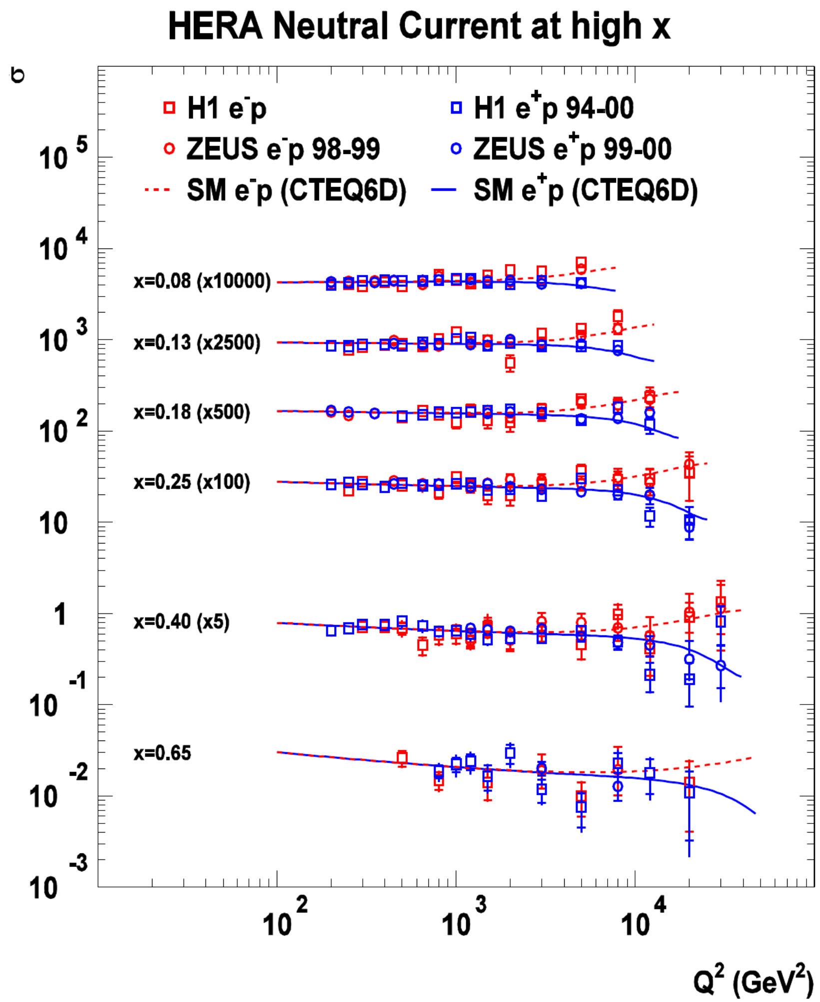




- HERA constitutes the single-most important dataset in any current PDF determination

HERA measurements at large Q^2





- HERA has also provided information at high Q^2
- **Z** and **W** become as important as **γ** exchange
- **NC** and **CC** cross sections become comparable
- electroweak effects in NC visible at highest Q^2
  → direct evidence for Z exchange

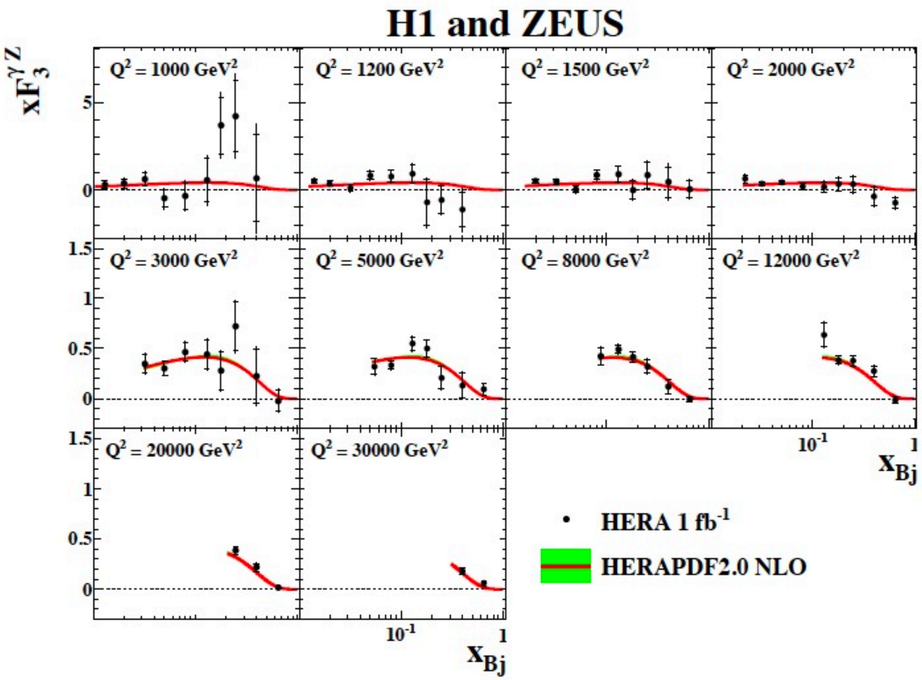
$$\sigma_{r,\mathrm{NC}}^{e^{\pm}p}(x,Q^2) = \left[F_2(x,Q^2) - \frac{y^2}{Y_+}F_\mathrm{L}(x,Q^2) \mp \frac{Y_-}{Y_+}xF_3(x,Q^2)\right]$$

HERA measurements at large Q^2

- F2 gives the usual sea information
- BUT also new valence structure function xF3 due to Z exchange (parity violating structure function)
- extraction of xF3 needs both e^+p and e^-p measurements

$$xF_3(x,Q^2) = \frac{Y_+}{2Y_-}\left(\sigma_{r,\mathrm{NC}}^{e^-p} - \sigma_{r,\mathrm{NC}}^{e^+p}\right)$$

- measurable from low to high x on pure proton target; **no heavy target corrections** and **no strong isospin assumptions**



$F2 = \Sigma_i\, A_i(Q^2)\, [xq_i(x,Q^2) + x\bar{q}_i(x,Q^2)]$

$xF3 = \Sigma_i\, B_i(Q^2)\, [xq_i(x,Q^2) - x\bar{q}_i(x,Q^2)]$

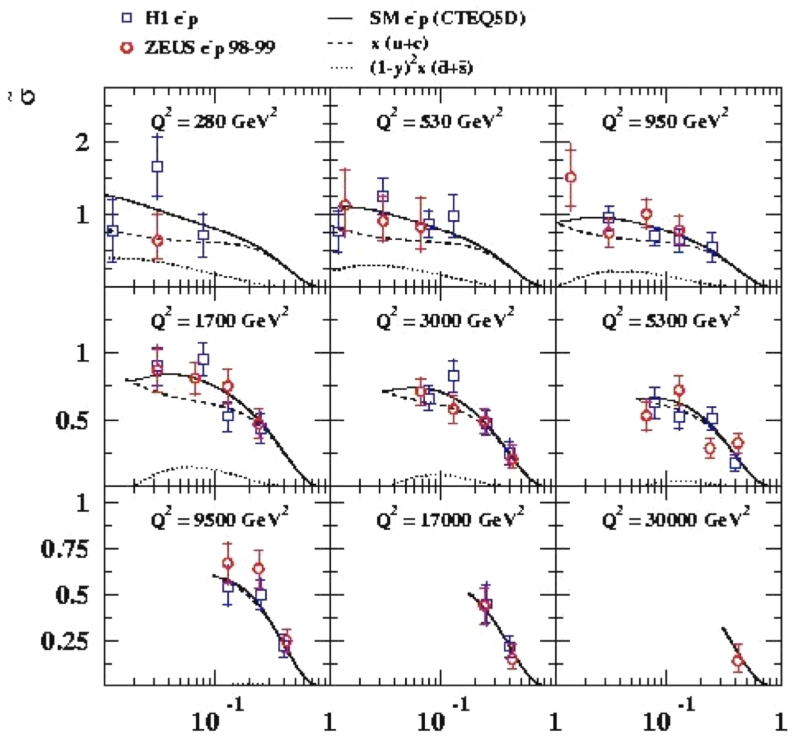
$A_i(Q^2) = e_i^2 - 2\, e_i v_i v_e P_Z + (v_e^2+a_e^2)(v_i^2+a_i^2) P_Z^2$

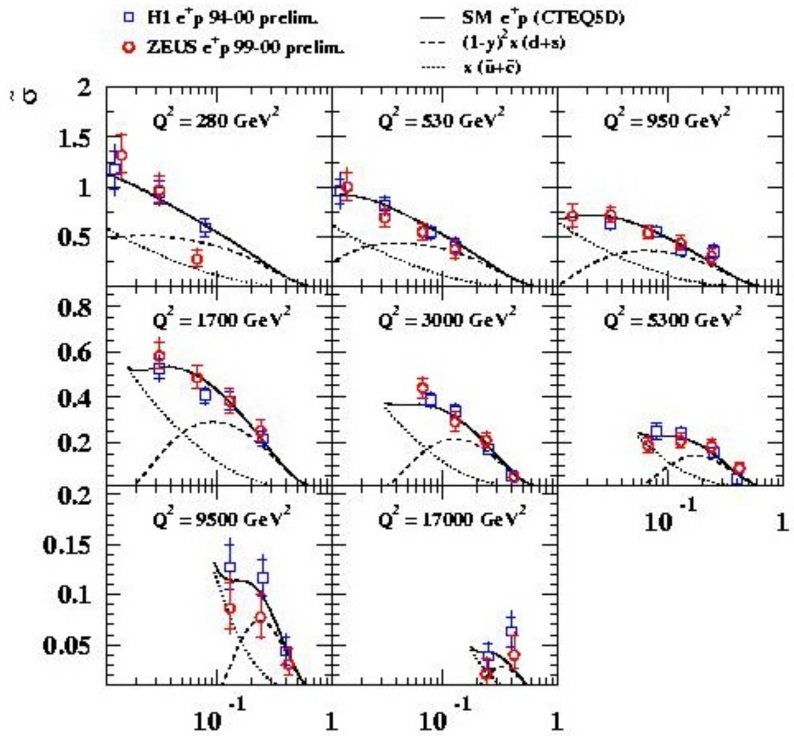
$B_i(Q^2) = \quad - 2\, e_i a_i a_e P_Z + 4\, a_i a_e v_i v_e P_Z^2$

$P_Z^2 = Q^2/[(Q^2+M_Z^2)(4 \sin^2\theta_W \cos^2\theta_W)]$

CC at HERA

gives flavour information





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

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

M_W information

u_v at high x

d_v at high x

Measurement of high-x d_v on a pure proton target (one caveat: data only up to x~0.65)

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. And you have to assume d in proton = u in neutron

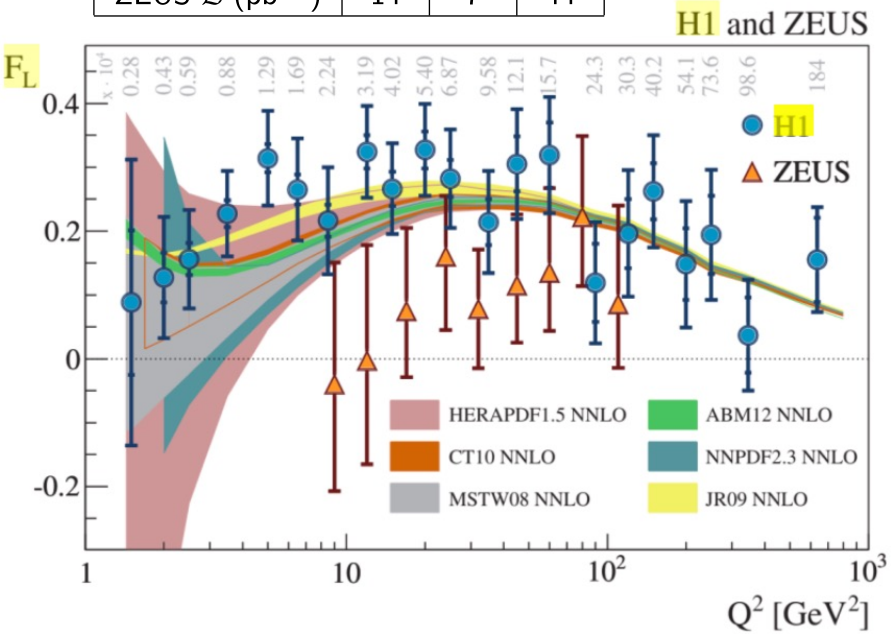
FL

- DGLAP QCD: to lowest order, FL: $F_L(x,Q^2) = \frac{\alpha_S}{4\pi} x^2 \int_x^1 \frac{dz}{z^3} \cdot \left[\frac{16}{3} F_2(z,Q^2) + 8 \sum e_q^2 \left(1 - \frac{x}{z}\right) zg(z,Q^2)\right]$
- at sufficiently low Q^2 can neglect xF3 and write reduced cross section:

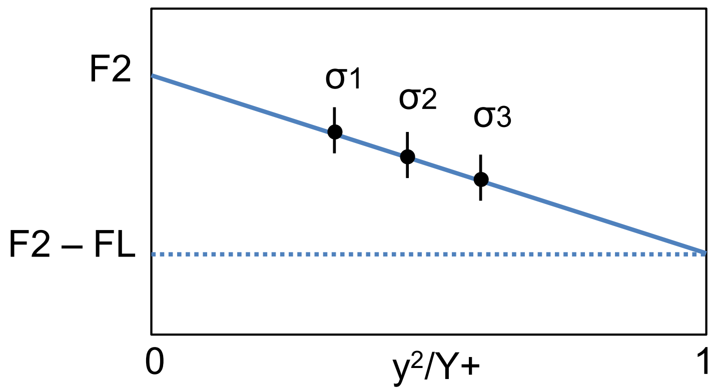
$$\sigma_r(x,Q^2;y) = \left[F_2(x,Q^2) - \frac{y^2}{Y_+} F_L(x,Q^2)\right] \quad \text{where} \quad \frac{y^2}{Y_+} = \frac{y^2}{1+(1-y)^2}$$

- need to measure at the same **x,Q²**, different **y** – use different beam energies (**Q^2 = s.x.y**)

E_p (GeV)	460	575	920
H1 $\mathcal{L}$ (pb^{-1})	12	6	22
ZEUS $\mathcal{L}$ (pb^{-1})	14	7	44



schematically:



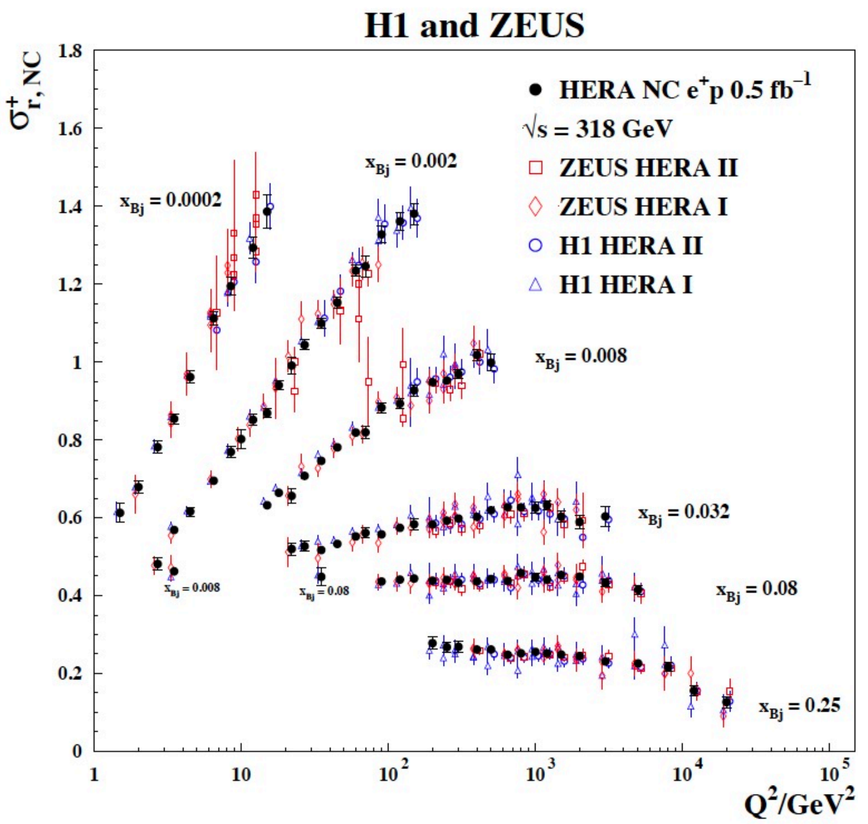
at a given x and Q^2:

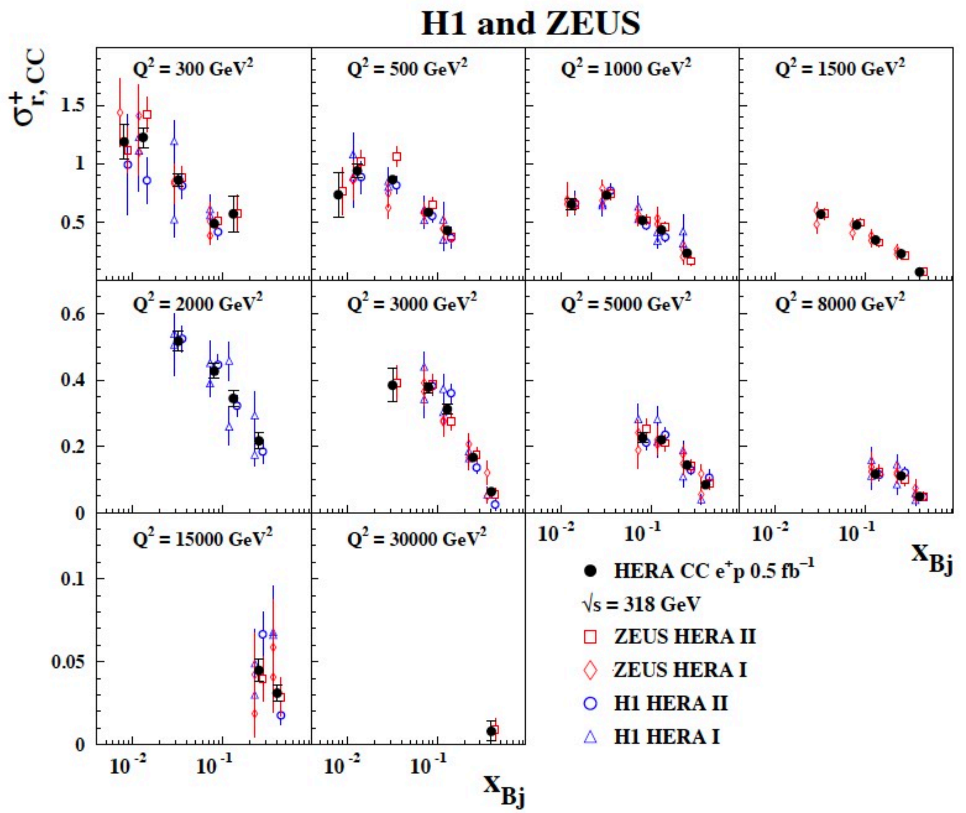
F2 is intercept at y-axis

FL is negative slope

HERA combination

arXiv:[1506.06042]

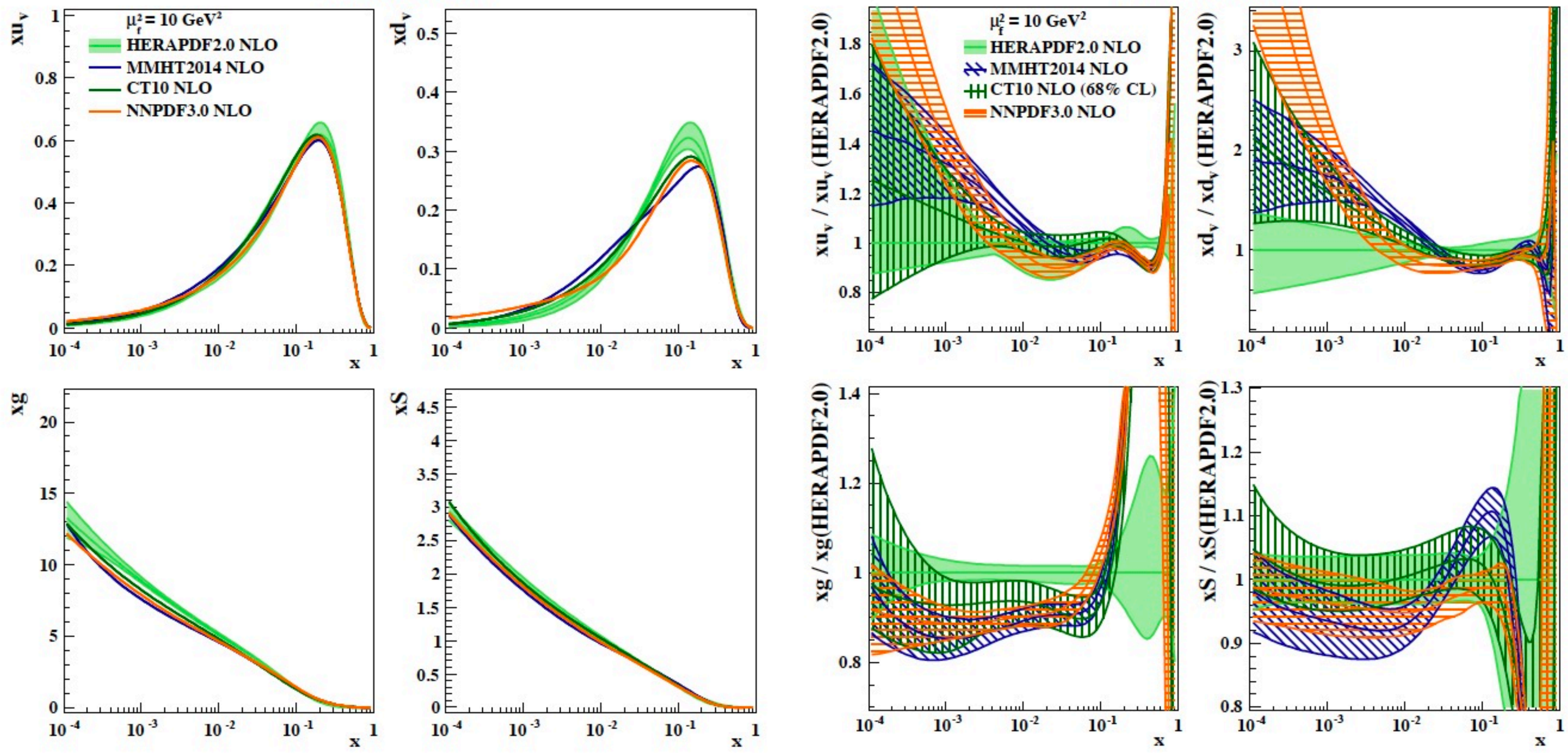




- H1 and ZEUS measurements combined in **generalised averaging procedure**, taking account of correlated systematics within and between experiments
- experiments cross-calibrate each other – reduced systematics in combined dataset
- total uncertainties improved by more than √2 in systematics dominated regions

- HERA constitutes the single-most important dataset in **any** modern PDF determination

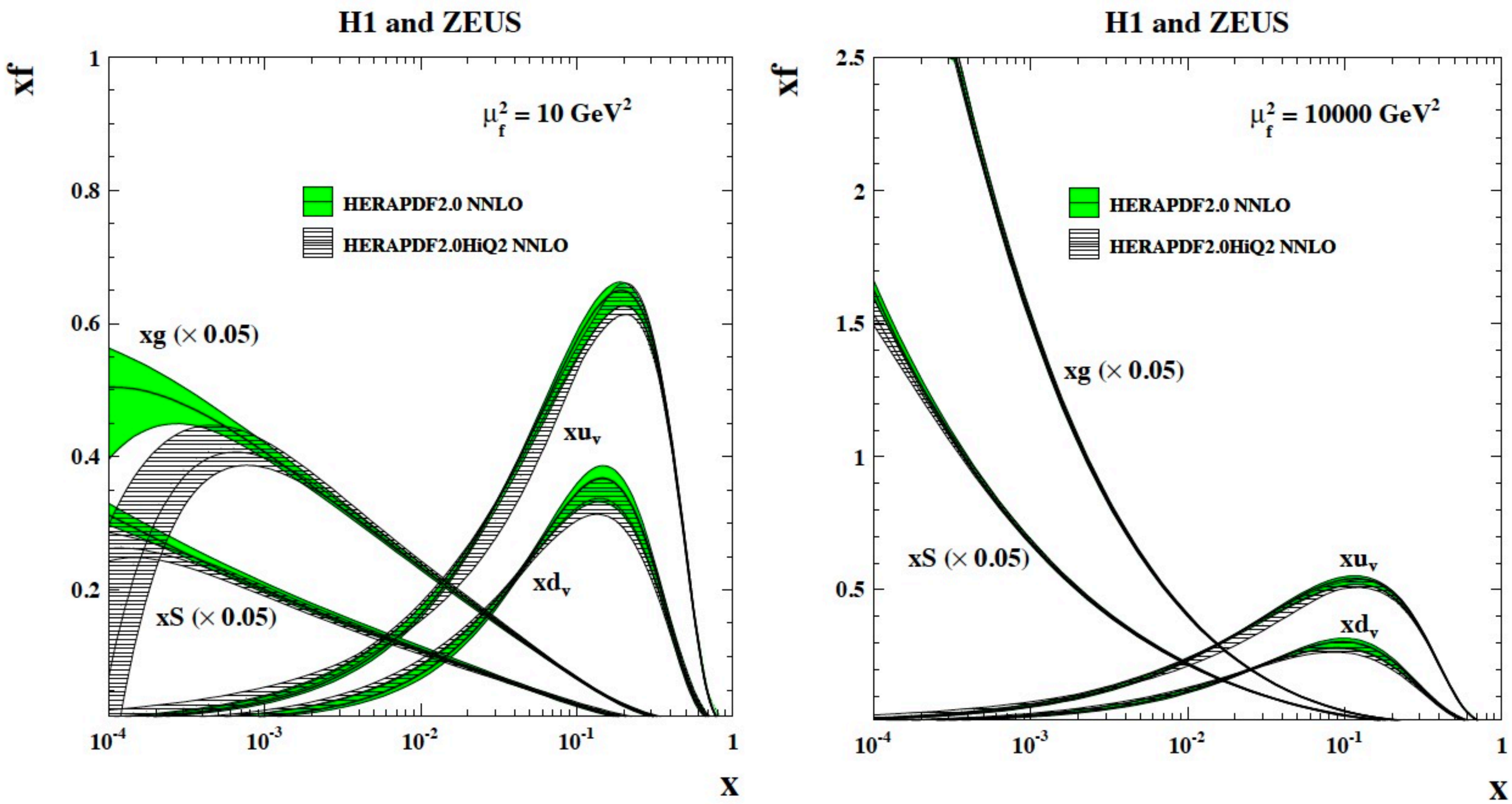
PDFs from a variety of groups



- this is what the PDFs look like – these are measured, not theoretical!
- PDFs are extracted by various groups:
- 3 main global fitters: **CT**, **MSHT** (previously known as MMHT, MSTW and MRS(T)), **NNPDF**
- others, usually using subsets of data: ABM(P); (J)GR(V); HERAPDF (HERA only); ATLAS, CMS and LHCb Colls., …

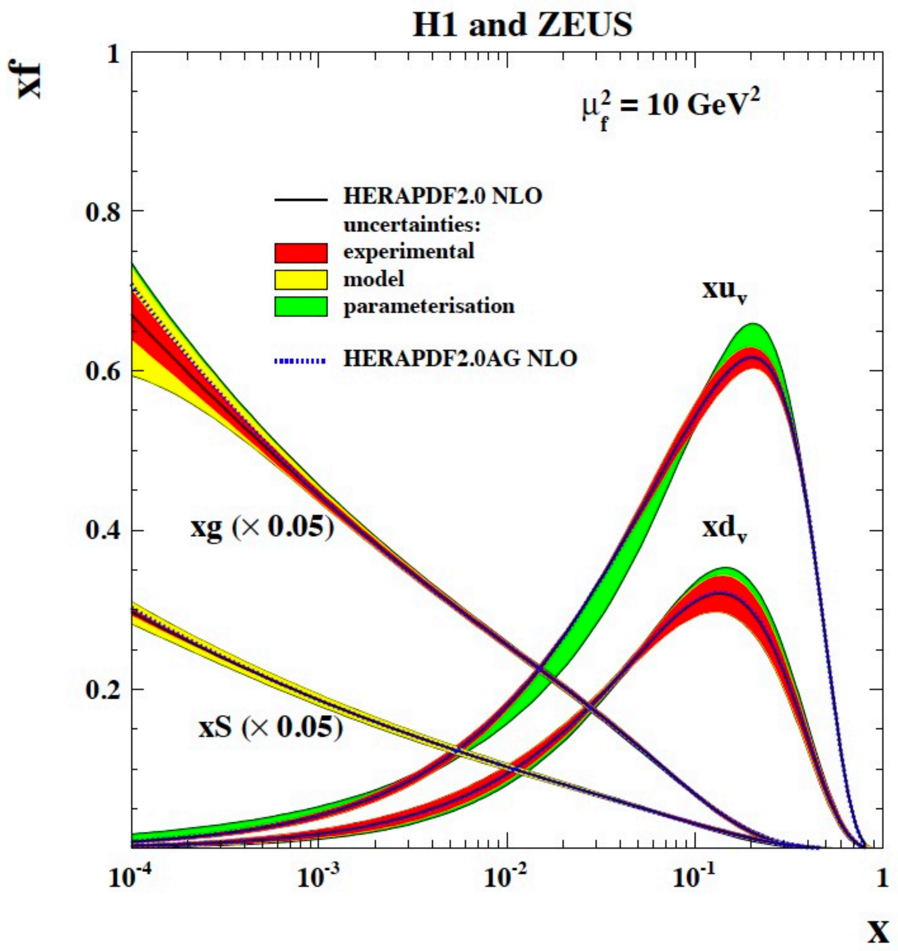
various illustrative PDF plots will be shown from now; many historical, not necessarily with the latest PDF versions!

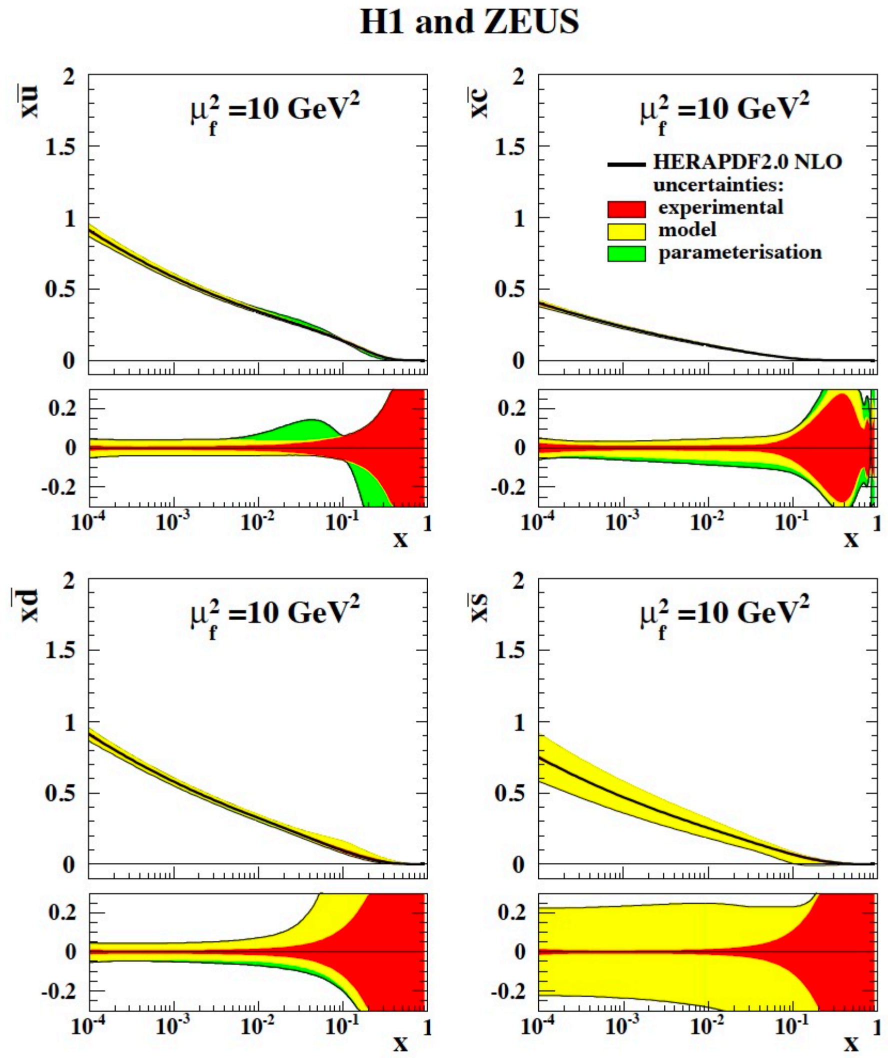
PDF evolution with Q^2



- as seen previously, valence evolve slowly, while sea and gluon evolve very rapidly!

PDF uncertainties



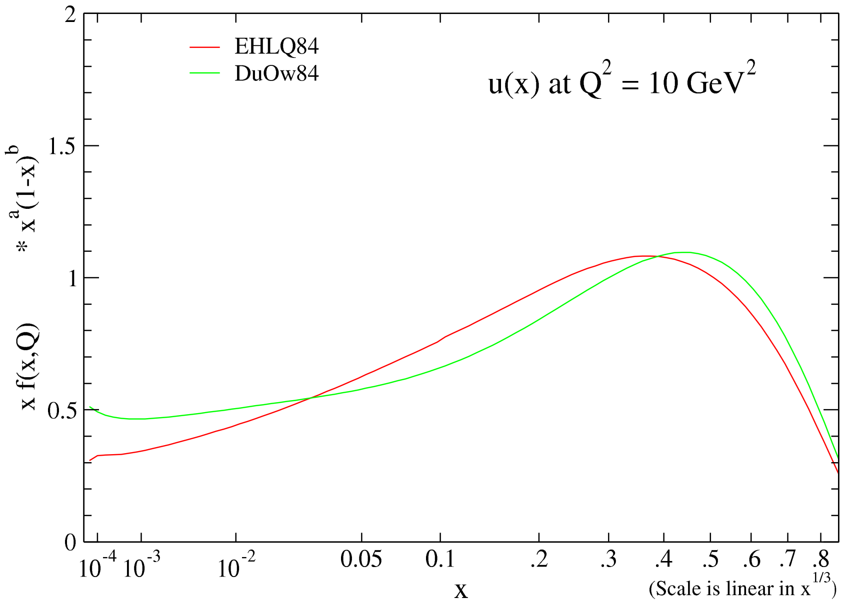


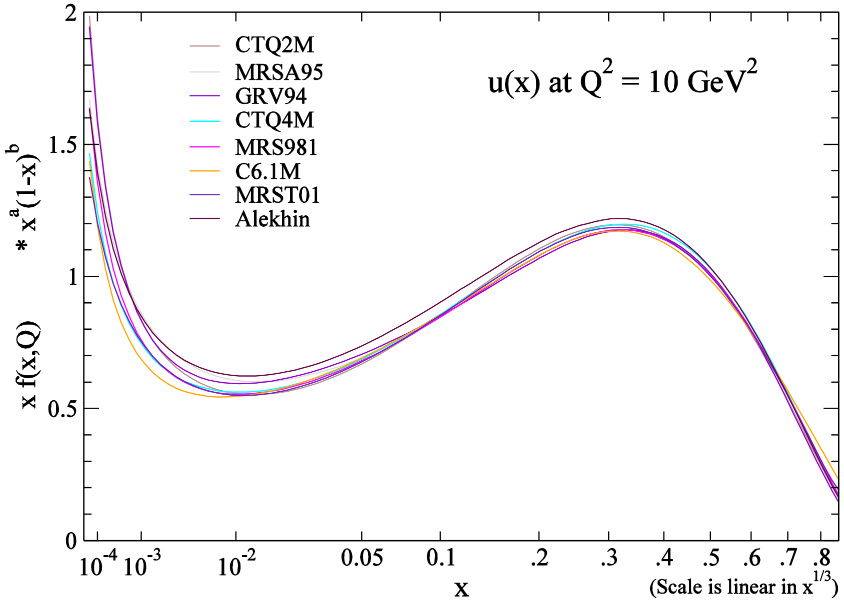
what do the error bands mean?

part is directly from experimental uncertainties on the measurements

part is due to assumptions – let's first consider assumptions …

progress over 20 years of PDF fitting





- the u-quark from 1984 looks rather different than the u-quark of 2004!

WHY?

obviously experiment has contributed

HERA data has shown that at low-x the gluon rises very steeply and generates a steep behavior in the quarks (see later)

BUT also development in relaxing model assumptions …

model assumptions I

- The mathematical form of the parameterisation

(The NNPDF use a neural net to learn the shape of the data rather than imposing a specific form of parameterisation)

- For $Q^2 >> Q_0^2$ this gets "washed out" provided it's reasonable…
- Value of Q_0^2

- No longer assume:
  - $\bar{u} = \bar{d}$, $\bar{s} = 0$ as in early work
  - $\frac{d_v}{u_v} = \frac{1}{2}$ independent of *x*, $b_u = b_d$
  - Or $a_s = a_g$, $a_u = a_d$
  - Or impose values on these parameters like

$$a_s = a_g = 0 \qquad a_u = a_d = 0.5 \qquad b_u = b_d = 3 \qquad b_g = 5,\ b_s = 7$$

(NB, parameters a, b are as defined at the start of this Lecture; control low and high-x behaviour: $x^a (1-x)^b$

Where did these prejudices from?

- Regge theory and counting rules

model assumptions II

- We now know that,

$$\bar{d} - \bar{u} \neq 0 \qquad \int_0^1 \frac{dx}{x}\left(F_2^p - F_2^n\right) \neq 0.33$$

$$= \Delta \qquad = \frac{1}{3}\int dv(u_v - d_v) + \frac{2}{3}\int dx(\bar{u} - \bar{d})$$

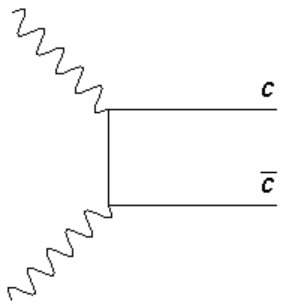
we will return to this later!

$$\bar{s} = \frac{\bar{u} + \bar{d}}{4}$$

rather than ½(ubar+dbar)

(from neutrino dimuons – maybe!)

- Charm sea generated by Boson Gluon Fusion (BGF)



- We still assume:
  - $d_{proton} = u_{neutron}$ — MRST QED 2004 challenges this
  - $u_{proton} = d_{neutron}$
  - $q_{sea} = \bar{q}$ — Maybe not for strange sector

valence flavour structure

Historically an SU(3) symmetric sea was assumed

$u = u_v + u_{sea}$, $d = d_v + d_{sea}$

$u_{sea} = \text{ubar} = d_{sea} = \text{dbar} = s = \text{sbar} = K$ and $c = \text{cbar} = 0$

Measurements of $\dfrac{F_2^{\mu n}}{F_2^{\mu p}} = \dfrac{u_v + 4d_v + 4/3K}{4u_v + d_v + 4/3K}$

Establish no valence quarks at small-x because

$F_2^{\mu n}/F_2^{\mu p} \rightarrow 1$

But $F_2^{\mu n}/F_2^{\mu p} \rightarrow 1/4$ as $x \rightarrow 1$

Not to 2/3 as it would for $d_v/u_v = 1/2$,

Hence it looks as if $d/u \rightarrow 0$ as $x \rightarrow 1$

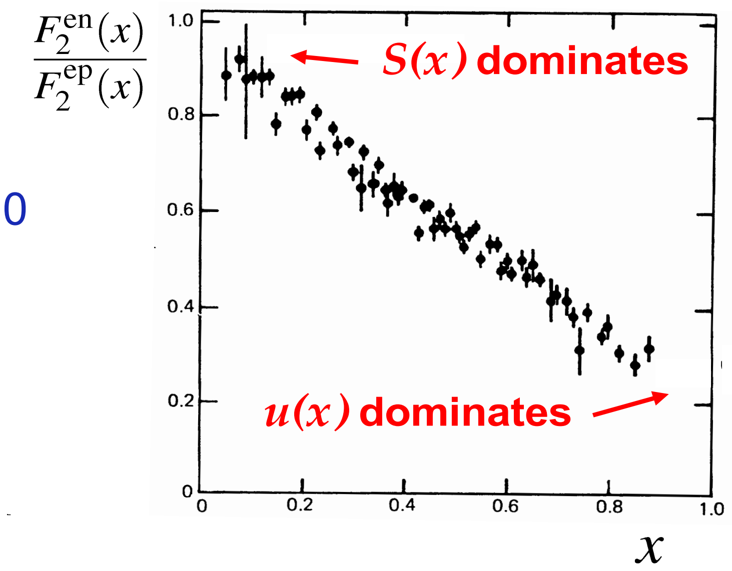
i.e. the dv momentum is softer than that of uv

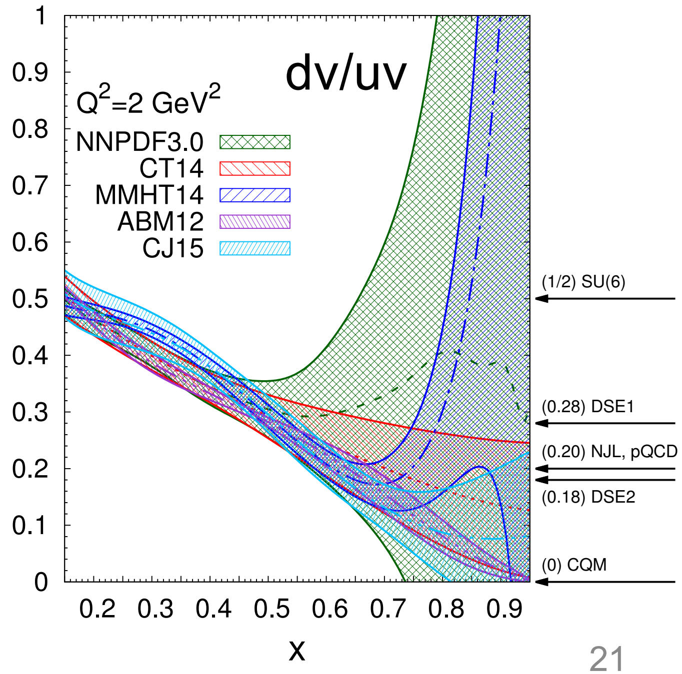
WHY? Non-perturbative physics...

… BUT precise behavior not understood:

data inconclusive, with large nuclear uncertainties; no predictive power from current pdfs; conflicting theory pictures;

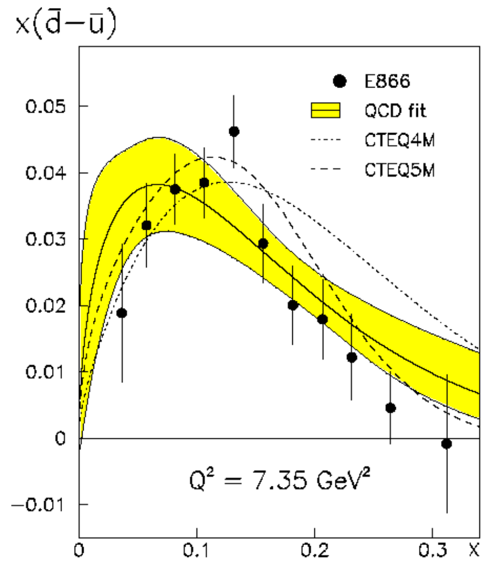
dv/uv essentially unknown at large x!





flavour structure in the sea

$\bar{d} \neq \bar{u}$ in the sea



Consider the Gottfried sum-rule (at LO)

$$\int d\,(F_{2\,p} - F_{2\,n}) = \frac{1}{3}\int d\,x\,(u_v - d_v) + \frac{2}{3}\int d\,x\left(\bar{u} - \bar{d}\right)$$

If $\bar{u} - \bar{d}$ then the sum should be 0.33

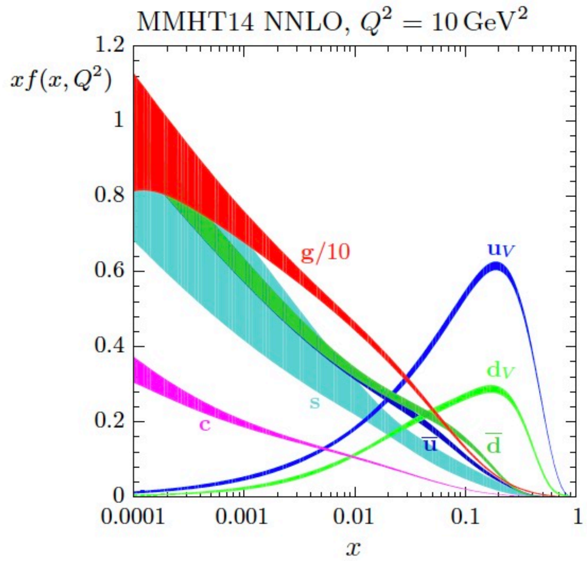
the measured value from NMC = 0.235 ± 0.026

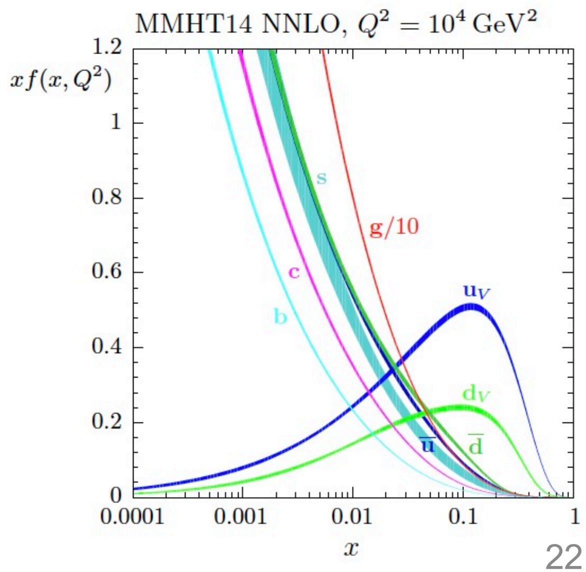
Clearly $\bar{d} > \bar{u}$…why? low Q² non-perturbative effects,

Pauli blocking from valence, suppression of g→uubar relative to g→ddbar;
meson cloud models: $p \to n\,\pi^{+},\ p\,\pi^{0},\ \Delta^{++}\,\pi^{-}$; …

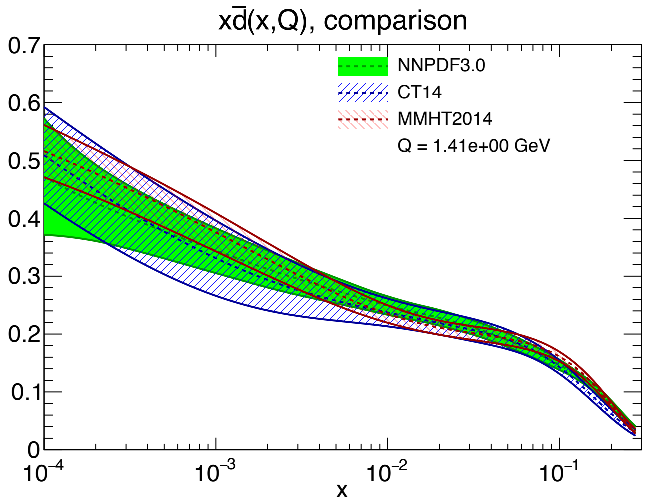
we now have more detailed shape information from Drell-Yan qqbar→μμ data using pp and pD scattering

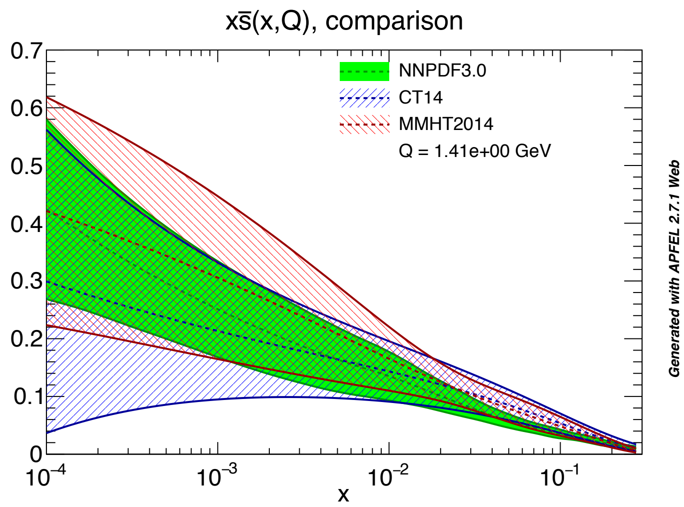
Note, if we compare this difference to the overall size of the **ubar** and **dbar** PDFs it is quite a small effect, and even less important at higher scales

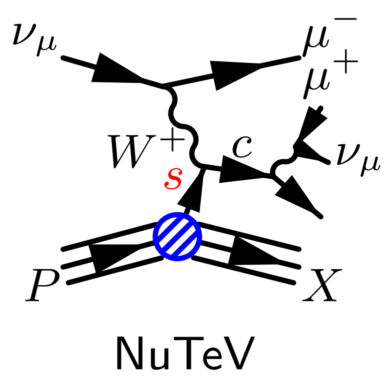




exactly how strange is the sea?







A first guess was: sbar = ½(ubar+dbar)

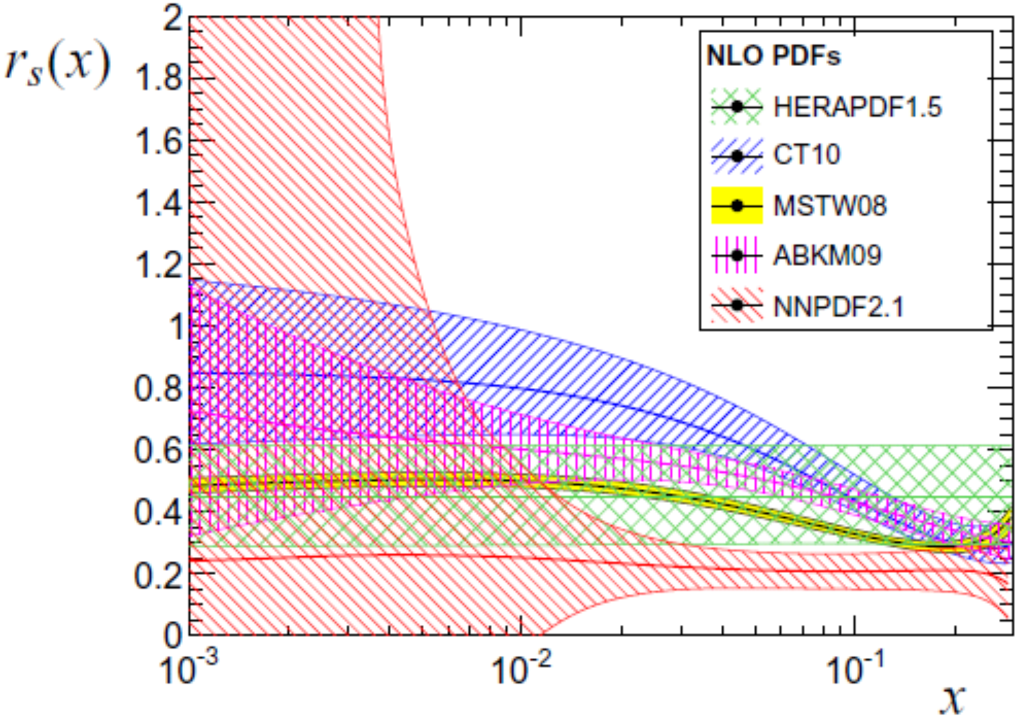
BUT quickly became **sbar = ¼(ubar+dbar)**

when neutrino dimuon data seemed to indicate suppression (NuTeV, now CHORUS, NOMAD)

- BUT what is the level of suppression as a function of x? **Nobody really knows!**
- appears suppressed at high x BUT not necessarily at low x?
- neutrino measurements only provide information for x ~ 0.1
- PLUS it needs nuclear target corrections, understanding of s → c threshold transition, and of progress of the charm quark through the nuclear medium
- and modern LHC data on W and Z production from ATLAS suggest it is not suppressed at low x (see later)

strange ratio

ratio: **xs(x)/xd(x)**



(slightly older PDFs shown here, plot just for illustration purposes)

s = sbar?

BUT Is the strangeness sector even charge symmetric?

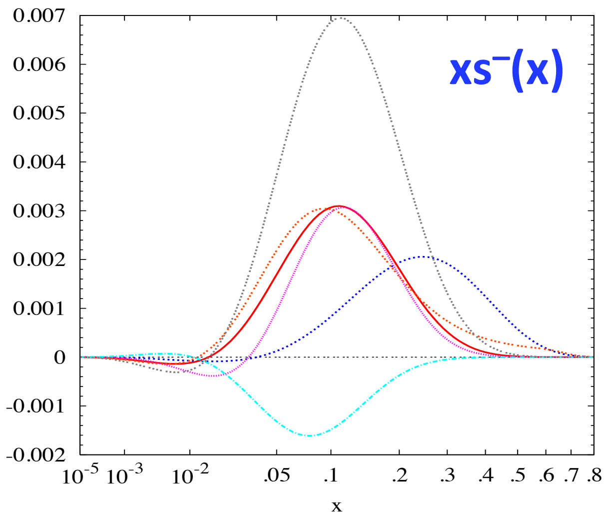
– is this the cause of the NuTeV $\sin^2\theta_W$ anomaly?

- CTEQ say that current global analysis does not **require** a non-zero $xs_-(x)= x(s-\bar{s})$ Its value is in the range

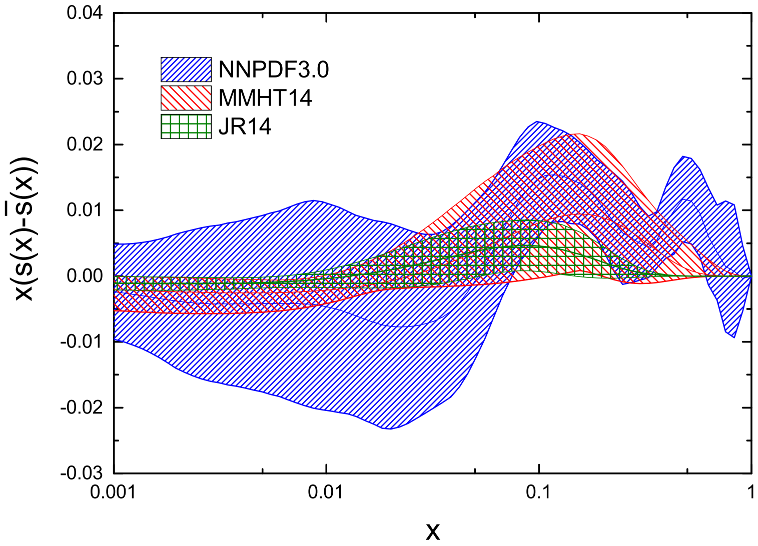
$$-0.001 < \langle x \rangle_{s_-} < 0.005$$

At 90%CL

They also give a range of possible shapes



- Other groups say there is an x(s-sbar) asymmetry



BUT this is a very **small effect**

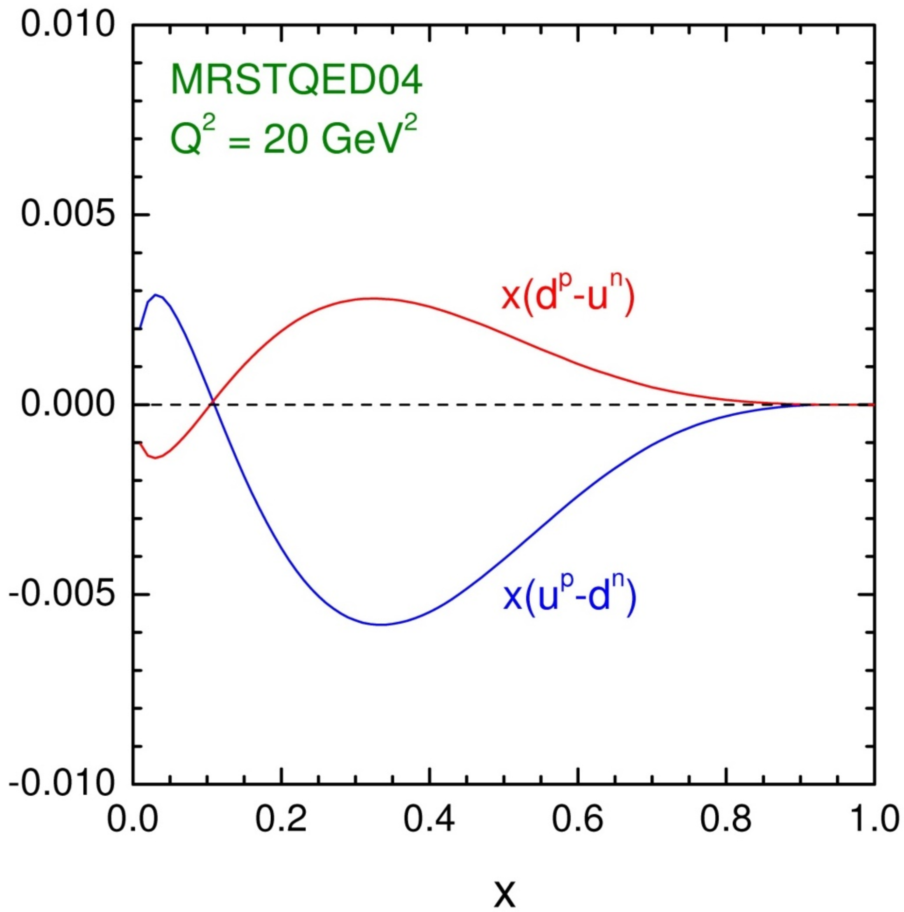
(NB, $\langle x \rangle_{s_-} = \int_0^1 x\, s_-(x, Q_0)\, dx$)

isospin symmetry assumption?

Is it true that u in proton = d in neutron ?

NOT if QED corrections are incorporated in the analysis

– is this the cause of the NuTeV $\sin^2\theta_W$ anomaly?



And this is an even smaller effect

extras

progress 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
Alekhin'03	1576	572	892	270	3309

(even the most recent fits shown here are now rather old, but illustrate a point!)