

PDF School 2008

DESY, Zeuthen, 12-14 November 2008

Introduction to Parton Distributions

and

MSTW* Analysis

James Stirling

Cambridge University



* Alan Martin, JS, Robert Thorne, Graeme Watt

references

“QCD and Collider Physics”

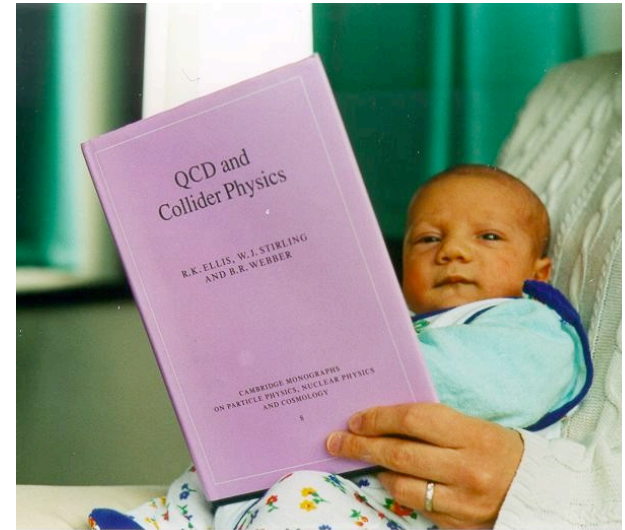
RK Ellis, WJ Stirling, BR Webber
Cambridge University Press (1996)

also

“Hard Interactions of Quarks and Gluons: a Primer for LHC Physics ”

JM Campbell, JW Huston, WJ Stirling (CSH)

www.pa.msu.edu/~huston/seminars/Main.pdf
Rep. Prog. Phys. **70**, 89 (2007)



past, present and future proton/antiproton colliders...

**Tevatron (1987 →)
Fermilab
proton-antiproton collisions
 $\sqrt{S} = 1.8, 1.96 \text{ TeV}$**

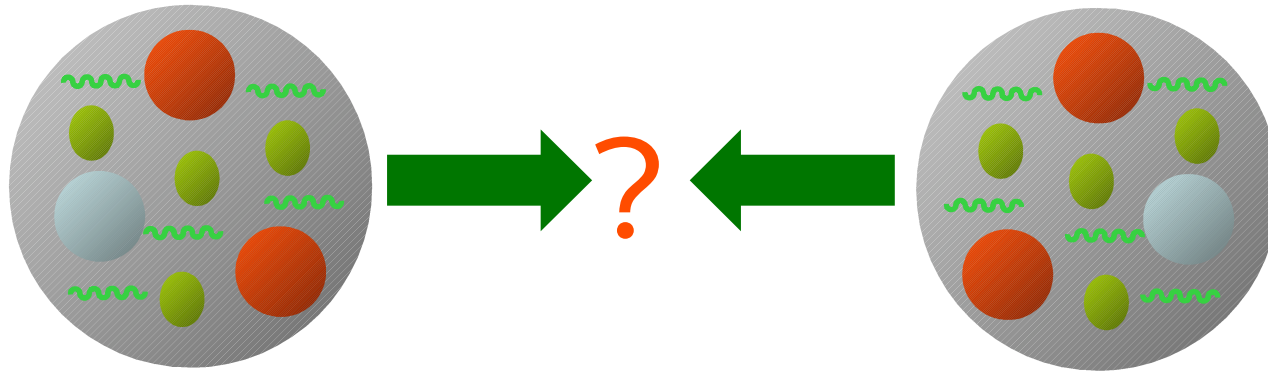


**Spp̄S (1981 → 1990)
CERN
proton-antiproton
collisions
 $\sqrt{S} = 540, 630 \text{ GeV}$**



**LHC (2009 →)
CERN
proton-proton and
heavy ion collisions
 $\sqrt{S} = 10 \rightarrow 14 \text{ TeV}$**

protons are not fundamental – what happens when they collide?



Most of the time – nothing of much interest, the protons break up and the final state consists of many low energy particles (pions, kaons, photons, neutrons,).

But, very occasionally, something dramatic happens ...violent collision between two 'parton' (hard, fundamental) constituents in the proton, which can produce a wide-angle scattering, or annihilation into new heavy objects.

We aim to quantify this.

What can we calculate?

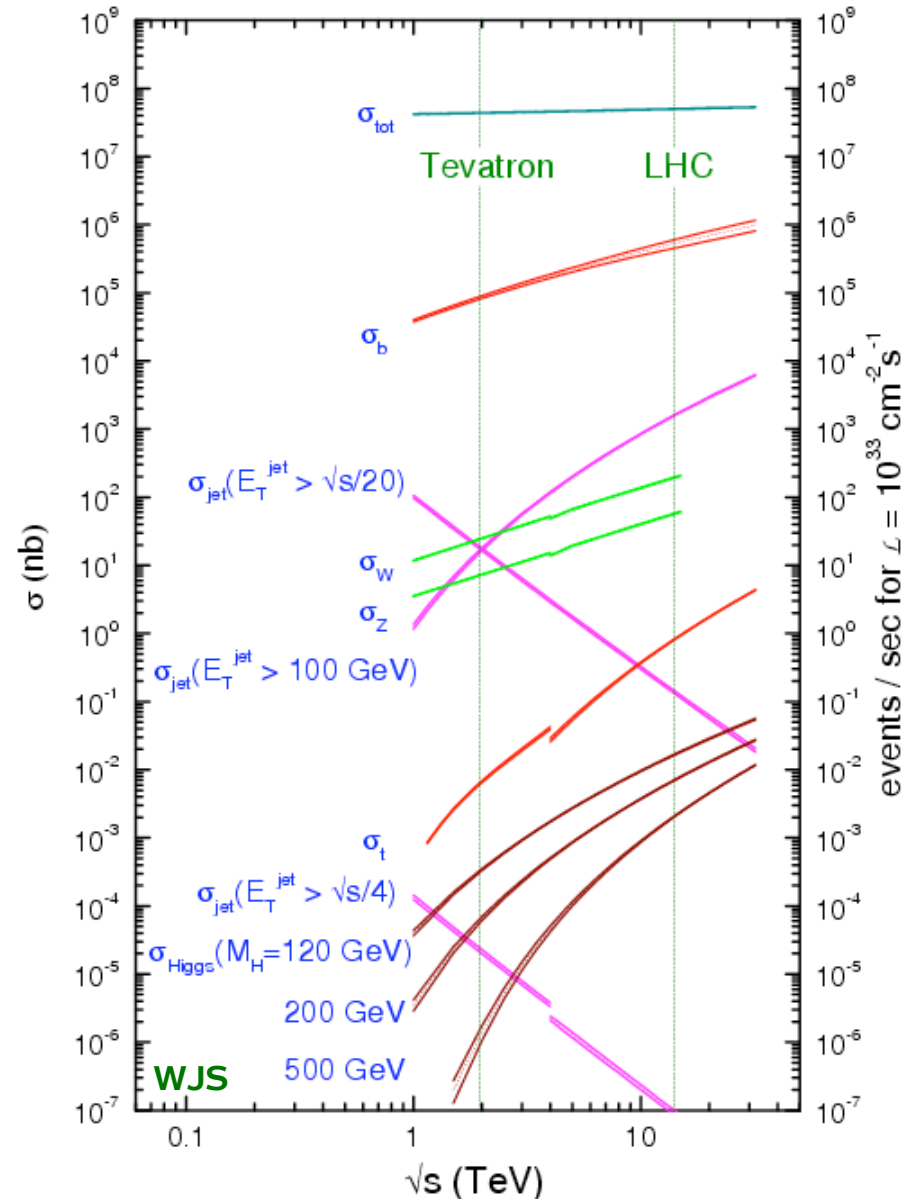
Scattering processes at high energy hadron colliders can be classified as either **HARD** or **SOFT**

Quantum Chromodynamics (QCD) is the underlying theory for **all** such processes, but the approach (and the level of understanding) is very different for the two cases

For **HARD** processes, e.g. **W** or **high- E_T jet** production, the rates and event properties can be predicted with some precision using **perturbation theory**

For **SOFT** processes, e.g. the **total cross section** or **diffractive** processes, the rates and properties are dominated by **non-perturbative** QCD effects, which are much less well understood

proton - (anti)proton cross sections



→ the QCD **factorization theorem** for hard-scattering (short-distance) inclusive processes

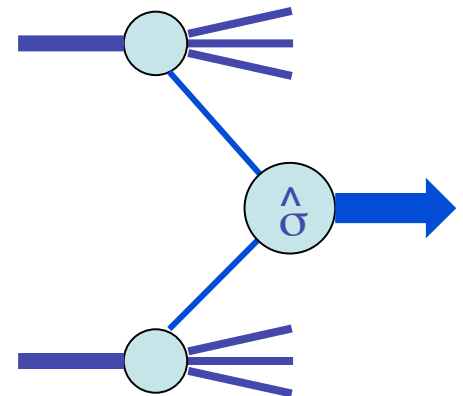
$$\sigma_X = \sum_{a,b} \int_0^1 dx_1 dx_2 f_a(x_1, \mu_F^2) f_b(x_2, \mu_F^2) \times \hat{\sigma}_{ab \rightarrow X} \left(x_1, x_2, \{P_i^\mu\}; \alpha_S(\mu_R^2), \alpha(\mu_R^2), \frac{Q^2}{\mu_R^2}, \frac{Q^2}{\mu_F^2} \right)$$

where $X=W, Z, H, \text{ high-}E_T \text{ jets, SUSY sparticles, black hole, ...}$, and Q is the 'hard scale' (e.g. $= M_X$), usually $\mu_F = \mu_R = Q$, and $\hat{\sigma}$ is known ...

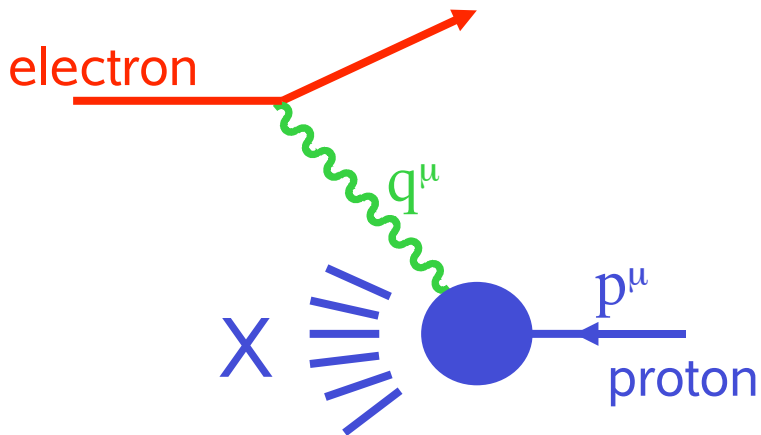
- to some fixed order in pQCD. e.a. high- E_T jets

$$\hat{\sigma} = A\alpha_S^2 + B\alpha_S^3$$

- or in some leading logarithm approximation (LL, NLL, ...) to all orders via resummation



deep inelastic scattering



- variables

$$Q^2 = -q^2$$

$$x = Q^2 / 2p \cdot q \quad (\text{Bjorken } x)$$

$$(y = Q^2 / x s)$$

- resolution

$$\lambda = \frac{h}{Q} = \frac{2 \times 10^{-16} \text{ m GeV}}{Q}$$

at HERA, $Q^2 < 10^5 \text{ GeV}^2$

$$\Rightarrow \lambda > 10^{-18} \text{ m} = r_p / 1000$$

- inelasticity

$$x = \frac{Q^2}{Q^2 + M_X^2 - M_p^2}$$

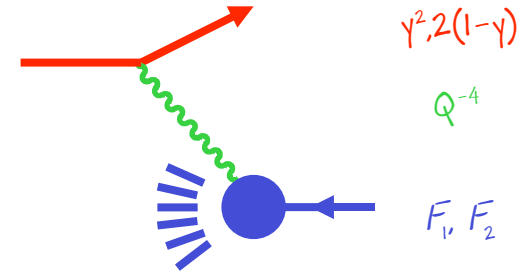
$$\Rightarrow 0 < x \leq 1$$

structure functions

- in general, we can write

$$\frac{d\sigma}{dx dQ^2} = \frac{4\pi\alpha^2}{Q^4} [y^2 F_1 + 2(1-y)x^{-1} F_2]$$

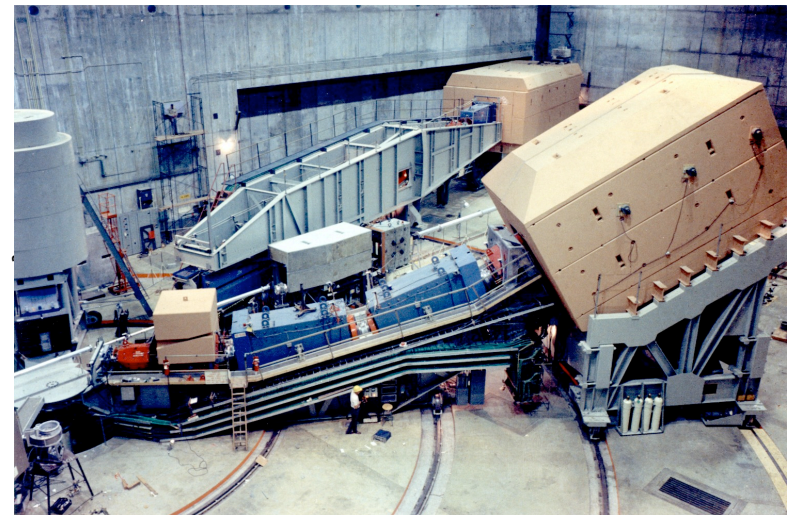
where the $F_i(x, Q^2)$ are called structure functions



SLAC, ~1970

- experimentally,
for $Q^2 > 1 \text{ GeV}^2$
 - $F_i(x, Q^2) \rightarrow F_i(x)$
“scaling”
 - $F_2(x) \approx 2 x F_1(x)$

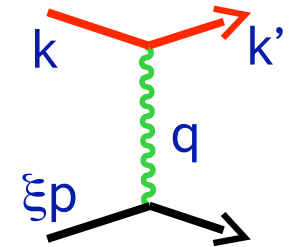
Bjorken 1968



toy model

- suppose that the electron scatters off a pointlike, \sim massless, spin $\frac{1}{2}$ particle a of charge e_a moving collinear with the parent proton with four-momentum $p_a^\mu = \xi p^\mu$
- calculate the scattering cross section $ea \rightarrow ea$

$$\begin{aligned} \frac{1}{4} \sum_{\text{spins}} |\mathcal{M}|^2 &= 2 e^4 e_a^2 \frac{s^2 + u^2}{t^2} \\ \frac{d\sigma^{ea \rightarrow ea}}{dt} &= \frac{e^4 e_a^2}{8\pi s^2} \frac{s^2 + u^2}{t^2} \\ \frac{d\sigma^{ea \rightarrow ea}}{dQ^2} &= \frac{2\pi\alpha^2 e_a^2}{Q^4} [1 + (1 - y)^2] \\ \frac{d\sigma}{dx dQ^2} &= \frac{2\pi\alpha^2}{Q^4} [y^2 + 2(1 - y)] e_a^2 \delta(x - \xi) \\ &\Rightarrow F_2 = x e_a^2 \delta(x - \xi) = 2x F_1 \end{aligned}$$

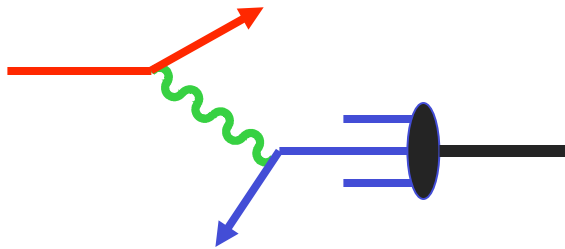


- **Exercise:** show that if a has spin-zero, then $F_1 = 0$

the parton model (Feynman 1969)

- photon scatters incoherently off massless, pointlike, spin-1/2 **quarks**
- probability that a quark carries fraction ξ of parent proton's momentum is $q(\xi)$, ($0 < \xi < 1$)

infinite
momentum
frame



$$\begin{aligned} F_2(x) &= \sum_{q,q} \int_0^1 d\xi e_q^2 \xi q(\xi) \delta(x-\xi) = \sum_{q,q} e_q^2 x q(x) \\ &= \frac{4}{9} x u(x) + \frac{1}{9} x d(x) + \frac{1}{9} x s(x) + \dots \end{aligned}$$

- the functions $u(x)$, $d(x)$, $s(x)$, ... are called **parton distribution functions** (pdfs) - they encode information about the proton's deep structure

extracting pdfs from experiment

- different beams (e, μ , ν , ...) & targets (H, D, Fe, ...) measure different combinations of quark pdfs
- thus the individual $q(x)$ can be extracted from a set of structure function measurements
- gluon not measured directly, but carries about 1/2 the proton's momentum

$$F_2^{ep} = \frac{4}{9}(u + \bar{u}) + \frac{1}{9}(d + \bar{d}) + \frac{1}{9}(s + \bar{s}) + \dots$$

$$F_2^{en} = \frac{1}{9}(u + \bar{u}) + \frac{4}{9}(d + \bar{d}) + \frac{1}{9}(s + \bar{s}) + \dots$$

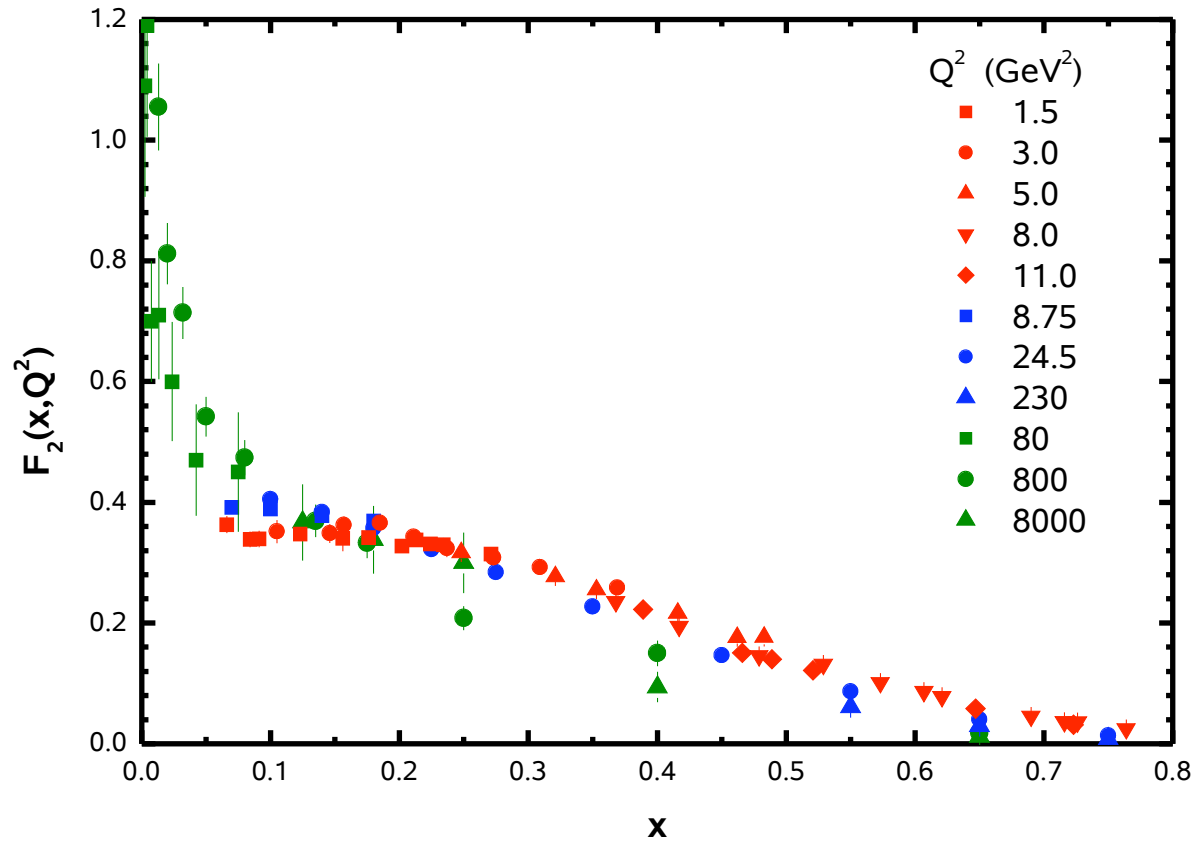
$$F_2^{\nu p} = 2 \left[d + s + \bar{u} + \dots \right]$$

$$F_2^{\nu n} = 2 \left[u + \bar{d} + \bar{s} + \dots \right]$$

$$s = \bar{s} = \frac{5}{6} F_2^{\nu N} - 3 F_2^{eN}$$

$$\sum_q \int_0^1 dx x (q(x) + \bar{q}(x)) = 0.55$$

40 years of Deep Inelastic Scattering



HERA

HERA

e^+, e^- (28 GeV)

p (920 GeV)

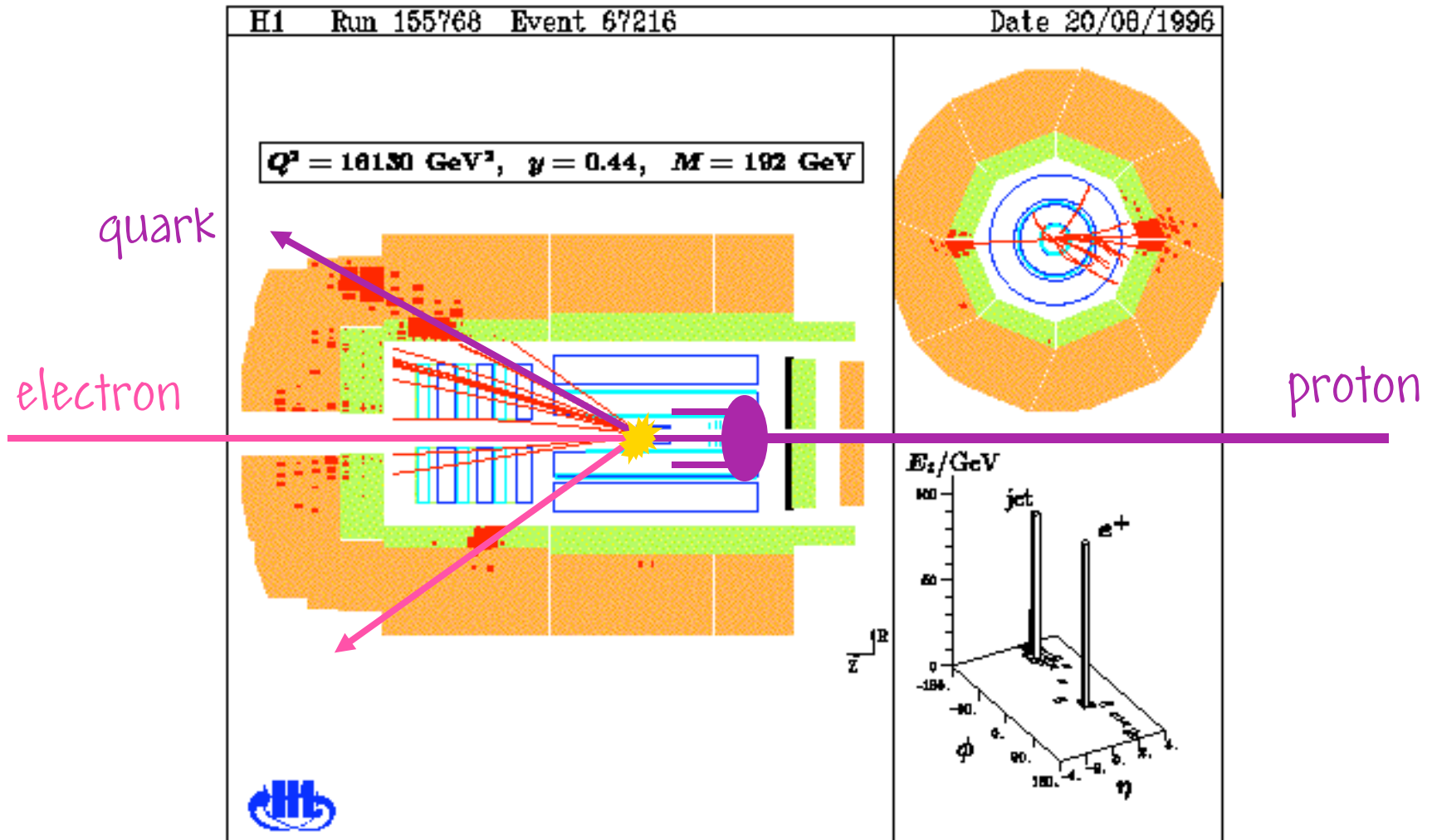


EXPERIMENT
AT HERA

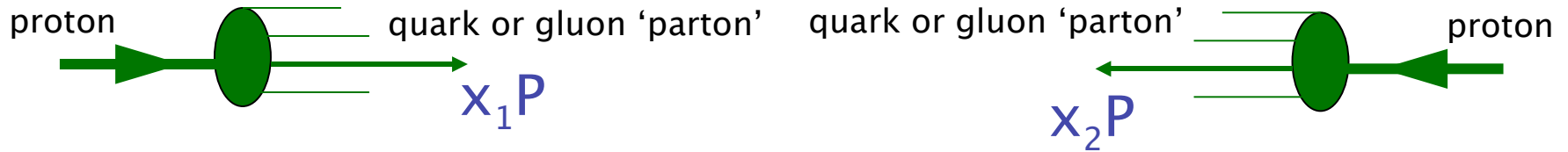


STERN

a deep inelastic scattering event at HERA



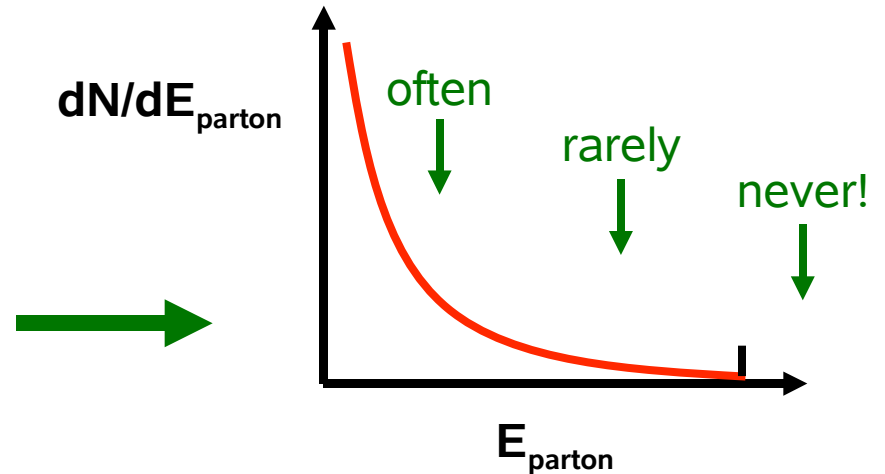
...and so in proton-proton collisions



$$\Rightarrow E_{\text{parton}} = \sqrt{x_1 x_2} E_{\text{collider}} \leq E_{\text{collider}}$$

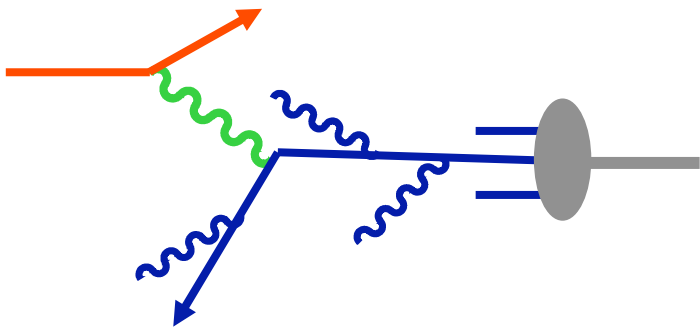
↑
relativistic kinematics

this collision energy distribution is just a convolution of the two parton probability distribution functions $f(x_1) * f(x_2)$

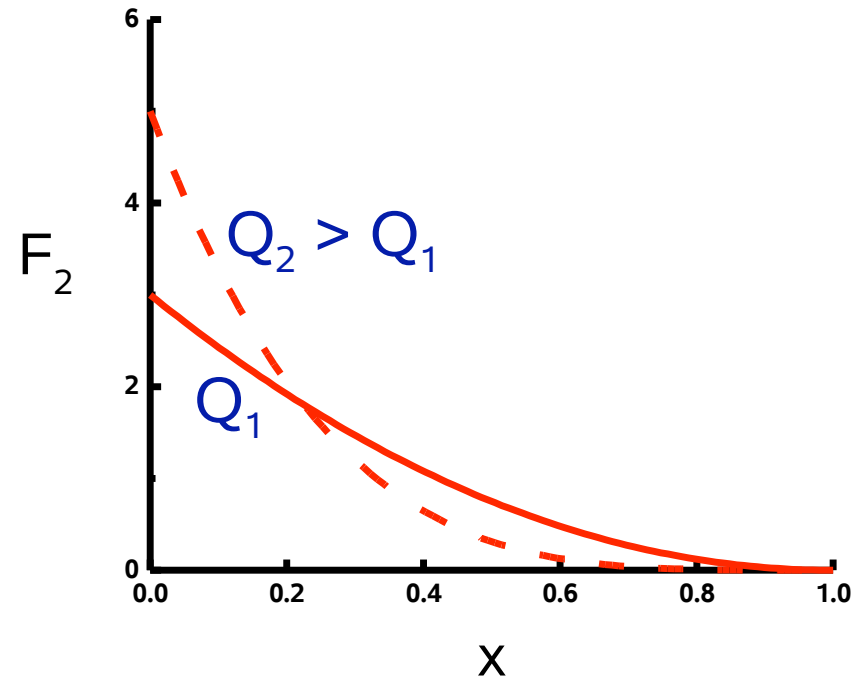


scaling violations and QCD

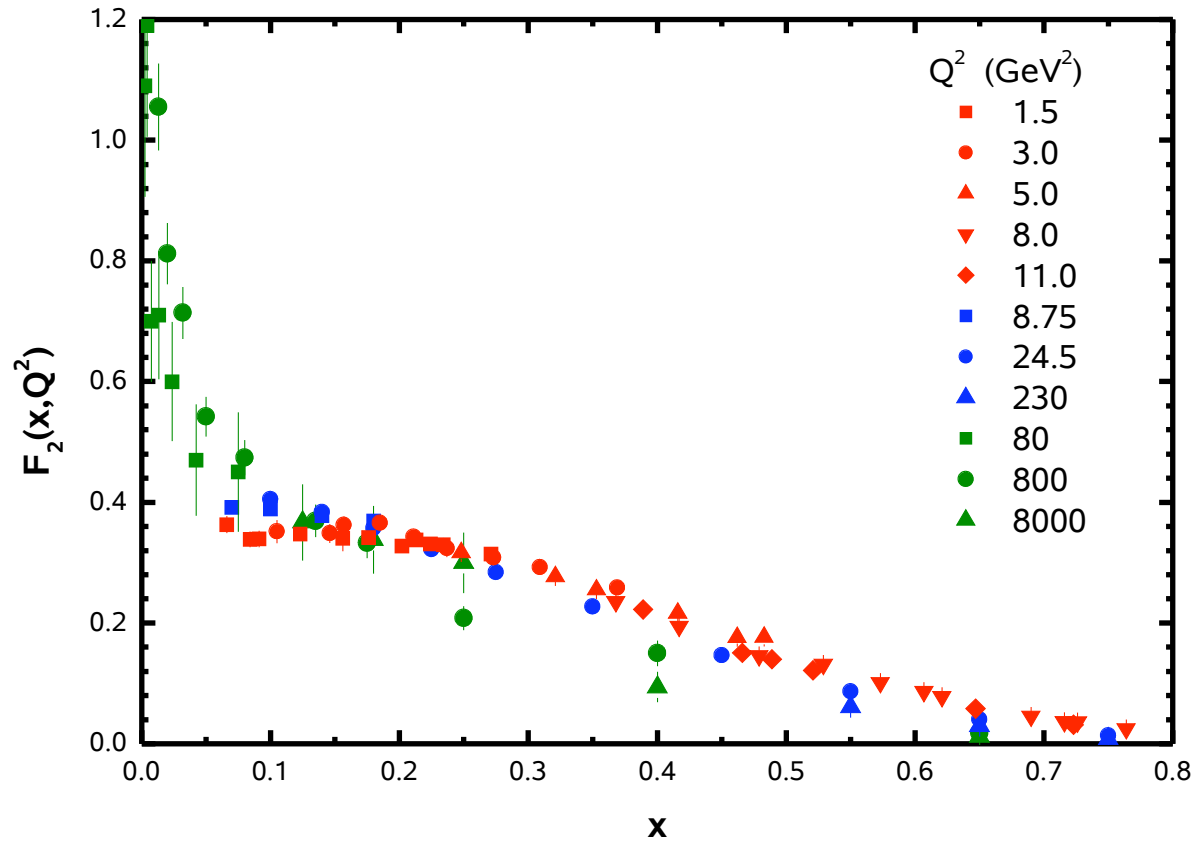
The structure function data exhibit systematic violations of Bjorken scaling:

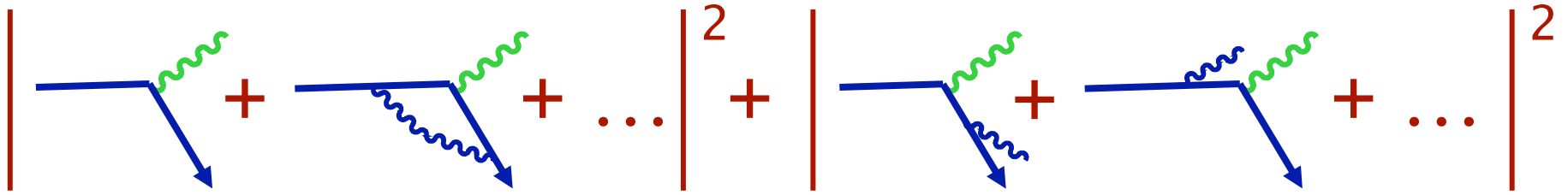


quarks emit gluons!



40 years of Deep Inelastic Scattering



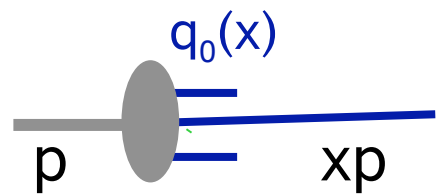


$$\longrightarrow \hat{F}_2 = e_q^2 \delta(1-x) + e_q^2 \frac{\alpha_S}{2\pi} x \left[P(x) \ln(Q^2/\kappa^2) + C(x) \right]$$

where the logarithm comes from $\int_0^{\sim Q^2} \frac{dk_T^2}{k_T^2} \rightarrow \int_{\kappa^2}^{\sim Q^2} \frac{dk_T^2}{k_T^2} \rightarrow \ln(Q^2/\kappa^2)$ ('collinear singularity') and

$$P(x) = C_F \left[\frac{1+x^2}{(1-x)_+} + \frac{3}{2} \delta(1-x) \right] \quad \boxed{\int_0^1 dx = \frac{f(x)}{(1-x)_+} = \int_0^1 \frac{f(x) - f(1)}{1-x}}$$

then convolute with a 'bare' quark distribution in the proton:



$$F_2(x, Q^2) = x \sum_q e_q^2 \left[q_0(x) + \frac{\alpha_S}{2\pi} \int_x^1 \frac{dy}{y} q_0(y) \left\{ P(x/y) \ln(Q^2/\kappa^2) + C(x/y) \right\} \right]$$

next, factorise the collinear divergence into a ‘renormalised’ quark distribution, by introducing the factorisation scale μ^2 :

$$q(x, \mu^2) = q_0(x) + \frac{\alpha_S}{2\pi} \int_x^1 \frac{dy}{y} q_0(y) \left\{ P(x/y) \ln(\mu^2/\kappa^2) + \bar{C}(x/y) \right\}$$

then $\frac{1}{x} F_2(x, Q^2) = x \sum_q e_q^2 \int_x^1 \frac{dy}{y} q(y, \mu^2) \left\{ \delta\left(1 - \frac{x}{y}\right) + \frac{\alpha_S}{2\pi} \left(P(x/y) \ln(Q^2/\mu^2) + C_q(x/y) \right) \right\}$

finite, by construction

note arbitrariness of $C_q = C - \bar{C} \rightarrow$ ‘factorisation scheme dependence’

we can choose \bar{C} such that $C_q = 0$, the DIS scheme, or use dimensional regularisation and remove the poles at $N=4$, the $\overline{\text{MS}}$ scheme, with $C_q \neq 0$

$q(x, \mu^2)$ is not calculable in perturbation theory,* but its scale (μ^2) dependence is:

$$\mu^2 \frac{\partial}{\partial \mu^2} q(x, \mu^2) = \frac{\alpha_S(\mu^2)}{2\pi} \int_x^1 \frac{dy}{y} q(y, \mu^2) P(x/y)$$

Dokshitzer
Gribov
Lipatov
Altarelli
Parisi

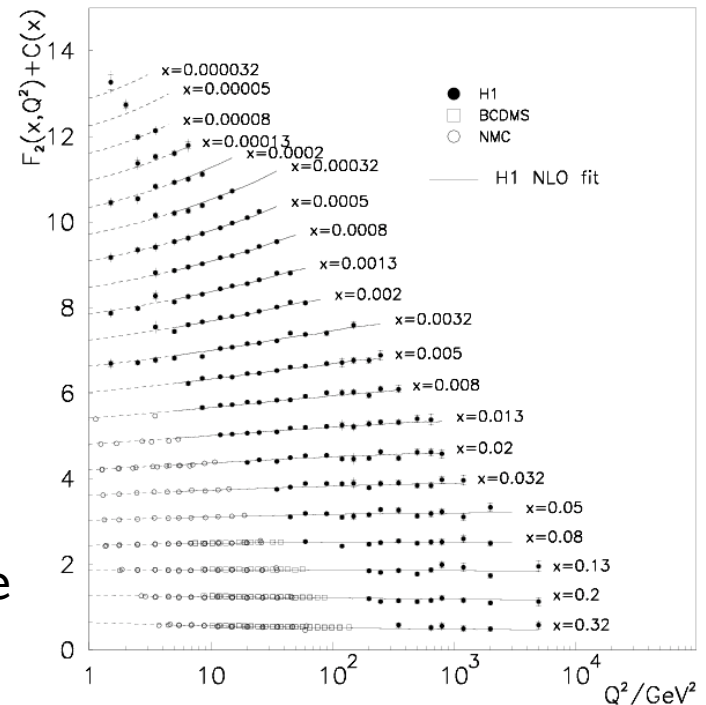
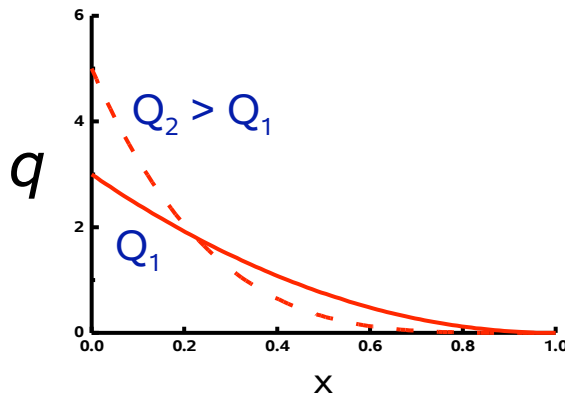
*lattice QCD?

note that we are free to choose $\mu^2 = Q^2$ in which case

$$\frac{1}{x}F_2(x, Q^2) = x \sum_q e_q^2 \int_x^1 \frac{dy}{y} q(y, Q^2) \left\{ \delta\left(1 - \frac{x}{y}\right) + \frac{\alpha_s}{2\pi} C_q(x/y) \right\}$$

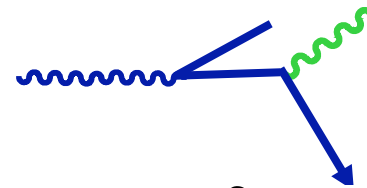
coefficient function,
see QCD book

... and thus the scaling violations of the structure function follow those of $q(x, Q^2)$ predicted by the DGLAP equation:



the rate of change of F_2 is proportional to α_s (DGLAP), hence structure function data can be used to measure the strong coupling!

however, we must also include the gluon contribution



$$\frac{1}{x}F_2(x, Q^2) = x \sum_q e_q^2 \int_x^1 \frac{dy}{y} q(y, Q^2) \left\{ \delta\left(1 - \frac{x}{y}\right) + \frac{\alpha_s(Q^2)}{2\pi} C_q(x/y) \right\}$$

$$+ x \sum_q e_q^2 \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{dy}{y} g(y, Q^2) C_g(x/y)$$

coefficient functions
- see QCD book

... and with additional terms in the DGLAP equations

$$\mu^2 \frac{\partial q_i(x, \mu^2)}{\partial \mu^2} = \frac{\alpha_S(\mu^2)}{2\pi} (P^{qq} * q_i + 2n_f P^{qg} * g)$$

$$\mu^2 \frac{\partial g(x, \mu^2)}{\partial \mu^2} = \frac{\alpha_S(\mu^2)}{2\pi} (P^{gq} * \sum_i q_i + P^{gg} * g)$$

$q_i = u, \bar{u}, d, \bar{d}, \dots$

* = convolution integral

note that at small (large) x , the gluon (quark) contribution dominates the evolution of the quark distributions, and therefore of F_2

$$P^{qq} = \frac{4}{3} \left(\frac{1+x^2}{1-x} \right)_+$$

$$P^{qg} = \frac{1}{2} (x^2 + (1-x)^2)$$

$$P^{gq} = \frac{4}{3} \left(\frac{1+(1-x)^2}{x} \right)$$

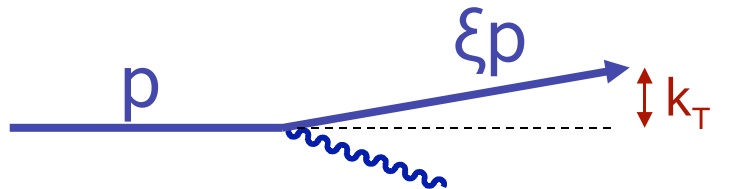
$$P^{gg} = 6 \left(\frac{1-x}{x} + x(1-x) + \left(\frac{x}{1-x} \right)_+ \right) - \left(\frac{1}{2} + \frac{n_f}{3} \right) \delta(1-x)$$

splitting functions

DGLAP evolution: physical picture

Altarelli, Parisi (1977)

- a fast-moving quark loses momentum by emitting a gluon:



$$d\mathcal{P} \simeq \frac{\alpha_S(k_T^2)}{2\pi} \frac{dk_T^2}{k_T^2} P^{qq}(\xi) d\xi$$

- ... with phase space $k_T^2 < O(Q^2)$, hence

$$d\mathcal{P} \simeq \frac{\alpha_S}{2\pi} \ln Q^2 P^{qq}(\xi) d\xi$$

- similarly for other splittings



$$\sim P^{gq}$$

- the combination of all such probabilistic splittings correctly generates the leading-logarithm approximation to the all-orders in pQCD solution of the DGLAP equations



basis of parton
shower Monte Carlos!

beyond lowest order in pQCD

going to higher orders in pQCD is straightforward in principle, since the above structure for F_2 and for DGLAP generalises in a straightforward way:

$$\frac{\partial q_i(x, Q^2)}{\partial \log Q^2} = \frac{\alpha_S}{2\pi} \int_x^1 \frac{dy}{y} \left\{ P_{q_i q_j}(y, \alpha_S) q_j\left(\frac{x}{y}, Q^2\right) + P_{q_i g}(y, \alpha_S) g\left(\frac{x}{y}, Q^2\right) \right\}$$

$$\frac{\partial g(x, Q^2)}{\partial \log Q^2} = \frac{\alpha_S}{2\pi} \int_x^1 \frac{dy}{y} \left\{ P_{g q_j}(y, \alpha_S) q_j\left(\frac{x}{y}, Q^2\right) + P_{g g}(y, \alpha_S) g\left(\frac{x}{y}, Q^2\right) \right\}$$

1972-77

1977-80

2004

DGLAP:
$$P(x, \alpha_S) = P^{(0)} + \alpha_S P^{(1)}(x) + \alpha_S^2 P^{(2)}(x) + \dots$$

see above

see book

see next slide!

The 2004 calculation of the complete set of $P^{(2)}$ splitting functions by Moch, Vermaseren and Vogt ([hep-ph/0403192](https://arxiv.org/abs/hep-ph/0403192), [0404111](https://arxiv.org/abs/hep-ph/0404111)) completes the calculational tools for a consistent NNLO pQCD treatment of Tevatron & LHC hard-scattering cross sections

- and for the structure functions...

$$\frac{1}{x}F_2(x, Q^2) = x \sum_q e_q^2 \int_x^1 \frac{dy}{y} q(y, Q^2) \left\{ \delta\left(1 - \frac{x}{y}\right) + \frac{\alpha_s(Q^2)}{2\pi} C_q^{(1)}(x/y) \right\}$$

$$x \sum_q e_q^2 \frac{\alpha_s(Q^2)}{2\pi} \int_x^1 \frac{dy}{y} g(y, Q^2) C_g^{(1)}(x/y) + \mathcal{O}(\alpha_s^2(Q^2))$$

... where up to and including the $\mathcal{O}(\alpha_s^3)$ coefficient functions are known

- terminology:

- LO: $P^{(0)}$
- NLO: $P^{(0,1)}$ and $C^{(1)}$
- NNLO: $P^{(0,1,2)}$ and $C^{(1,2)}$

- the more pQCD orders are included, the weaker the dependence on the (unphysical) factorisation scale, μ_F^2

– and also the (unphysical) renormalisation scale, μ_R^2 ; note above has $\mu_R^2 = Q^2$

how pdfs are obtained

- choose a factorisation scheme (e.g. MSbar), an order in perturbation theory (see below, e.g. LO, NLO, NNLO) and a ‘starting scale’ Q_0 where pQCD applies (e.g. 1-2 GeV)
- parametrise the quark and gluon distributions at Q_0 , e.g.

$$f_i(x, Q_0^2) = A_i x^{a_i} [1 + b_i \sqrt{x} + c_i x] (1 - x)^{d_i}$$

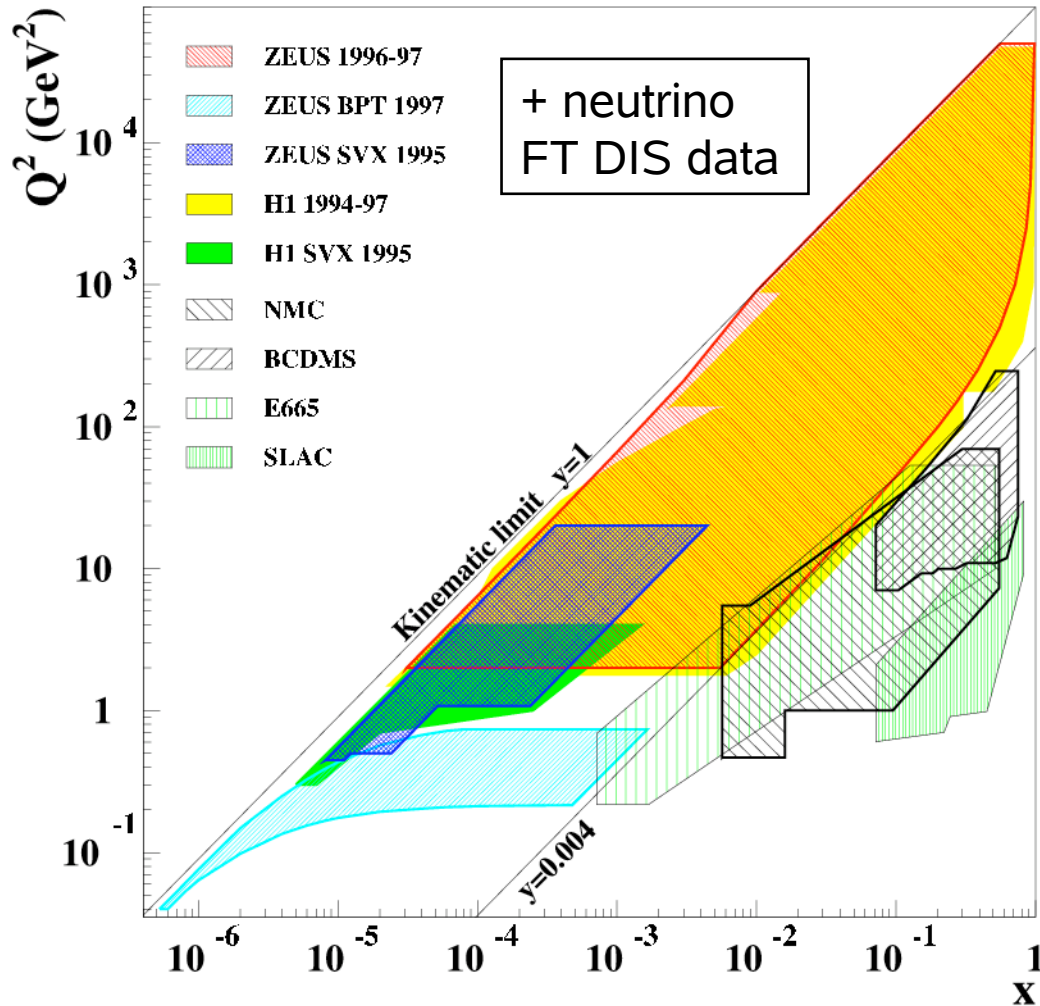
- solve DGLAP equations to obtain the pdfs at any x and scale $Q > Q_0$; fit data for parameters $\{A_i, a_i, \dots, \alpha_S\}$
- approximate the exact solutions (e.g. interpolation grids, expansions in polynomials etc) for ease of use; thus the output ‘global fits’ are available ‘off the shelf’, e.g.

`SUBROUTINE PDF (X, Q, U, UBAR, D, DBAR, ..., BBAR, GLU)`

input |

output

summary of DIS data



Note: must impose cuts on DIS data to ensure validity of leading-twist DGLAP formalism in the global analysis, typically:

$$Q^2 > 2 - 4 \text{ GeV}^2$$

$$W^2 = (1-x)/x Q^2 > 10 - 15 \text{ GeV}^2$$

pdfs from global fits – summary

Formalism

LO, NLO, NNLO DGLAP
MSbar factorisation
 Q_0^2
functional form @ Q_0^2
sea quark (a)symmetry
etc.

$$f_i(x, Q^2) \pm \delta f_i(x, Q^2)$$

$$\alpha_s(M_Z)$$

Data

DIS (SLAC, BCDMS, NMC, E665, CCFR, H1, ZEUS, ...)
Drell-Yan (E605, E772, E866, ...)
High E_T jets (CDF, D0)
W rapidity asymmetry (CDF, D0)
Z rapidity distribution (CDF, D0)
 νN dimuon (CCFR, NuTeV)
etc.

Who?

CTEQ, MSTW, Alekhin, NNPDF,
H1, ZEUS, Dortmund, Zeuthen, ...

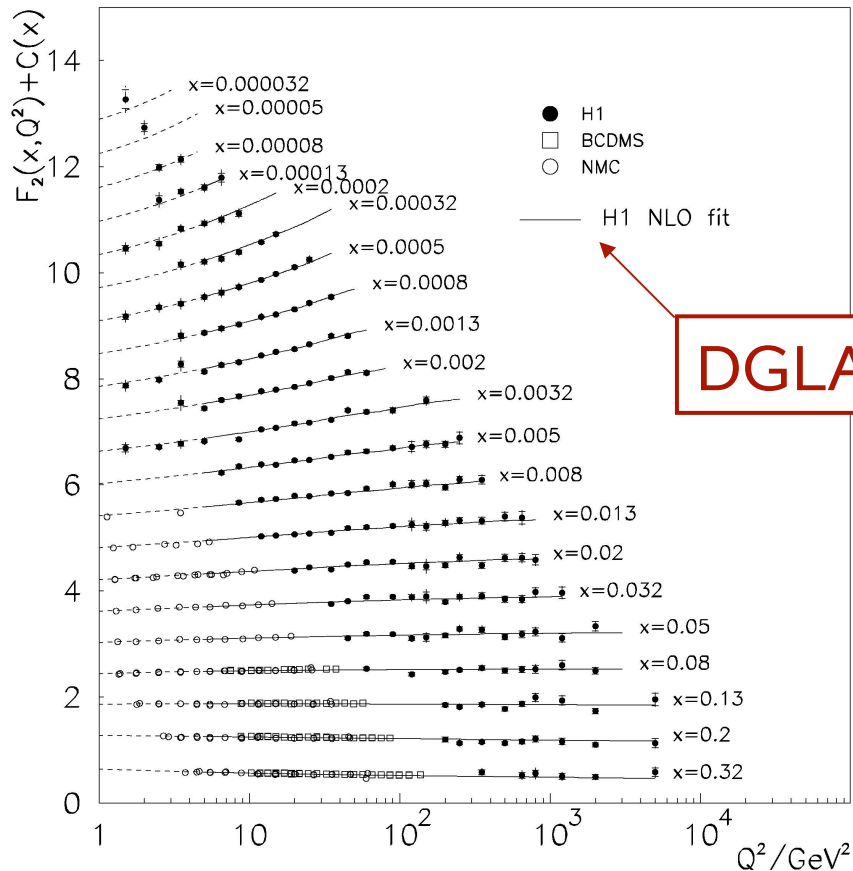
testing QCD

structure function data
from H1, BCDMS, NMC

- precision test of QCD
- measurement of the strong coupling:

$$\alpha_S^{\text{NNLO}}(M_Z) = 0.117 \pm 0.003$$


(MSTW 2008, from global fit)



where to find parton distributions

HEPDATA pdf website
<http://durpdg.dur.ac.uk/hepdata>

- access to code for MRST/MSTW, CTEQ etc
- online pdf plotting
- FORTRAN and C++ code

**Parton Distribution Functions**

Unpolarized Parton Distributions

Access the parton distribution code, on-line calculation and graphical display of the distributions, from CTEQ, GRV, MRS and Alekhin.

CTEQ distributions, [fortran code and grids](#)
GRV distributions, [fortran code and grids](#)
MRST distributions, [fortran code and grids](#), [C++ code](#)
ALEKHIN distributions, [fortran, C++ and Mathematica code, and grids](#)

On-line Parton Distribution Calculator with Graphical Display.
- now includes PDF error calculations from MRST2001E and CTEQ6.

Public access to the [ZEUS 2002 PDFs](#), [ZEUS 2005 jet fit PDFs](#) and [H1 PDF 2000 sets](#).

Polarized Parton Distributions

Currently available parametrizations:

E.Leader, A.V.Sidorov and D.B.Stamenov, Eur.Phys.J.C23 (2002) 479: [LSS2001](#)
E.Leader, A.V.Sidorov and D.B.Stamenov, Phys.Rev.D73 (2006) 034023: [LSS2005](#)
M. Glueck, E. Reya, M. Stratmann and W. Vogelsang, Phys. Rev. D53 (1996) 4775: [GRSV](#)
M. Glueck, E. Reya, M. Stratmann and W. Vogelsang, Phys. Rev. D63 (2001) 094005: [GRSV2000](#)
T. Gehrmann and W.J. Stirling, Phys. Rev. D53 (1996) 6100: [GS](#)
J. Bluemlein and H. Boettcher - hep-ph/0203155 [BB](#)
Asymmetry Analysis Collaboration - M. Hirai et al - Phys. Rev. D69 (2004) 054021 [AAC](#)
D. de Florian and R. Sassot, Phys. Rev. D62 (2000) 094025: [DS2000](#)
D. de Florian, G.A. Navarro and R. Sassot, Phys. Rev. D71 (2005) 094018: [DNS2005](#)

Polarized Parton Distributions

Currently available parametrizations:

E.Leader, A.V.Sidorov and D.B.Stamenov, Eur.Phys.J.C23 (2002) 479: [LSS2001](#)
E.Leader, A.V.Sidorov and D.B.Stamenov, Phys.Rev.D73 (2006) 034023: [LSS2005](#)
M. Glueck, E. Reya, M. Stratmann and W. Vogelsang, Phys. Rev. D53 (1996) 4775: [GRSV](#)
M. Glueck, E. Reya, M. Stratmann and W. Vogelsang, Phys. Rev. D63 (2001) 094005: [GRSV2000](#)
T. Gehrmann and W.J. Stirling, Phys. Rev. D53 (1996) 6100: [GS](#)
J. Bluemlein and H. Boettcher - hep-ph/0203155 [BB](#)
Asymmetry Analysis Collaboration - M. Hirai et al - Phys. Rev. D69 (2004) 054021 [AAC](#)
D. de Florian and R. Sassot, Phys. Rev. D62 (2000) 094025: [DS2000](#)
D. de Florian, G.A. Navarro and R. Sassot, Phys. Rev. D71 (2005) 094018: [DNS2005](#)

Pion Parton Distributions

Access the parton distribution code for pions

MRS pion distributions, [fortran code and grids](#)

PDFs from nuclei

M. Hirai, S. Kumano and M. Miyama - Phys. Rev. D64 (2001) 034003 [PDFs from nuclei](#)
K.J.Eskola, V.J.Kolhinen and C.A. Salgado - Eur.Phys.J C9(1999)61
and K.J.Eskola, V.J.Kolhinen and P.V.Ruuskanen - Nucl.Phys.B535(1998)351 [EKS98 parametrization](#)


==> **Other related topics**

Deeply Virtual Compton Scattering at NLO in pQCD

Access to the [DVCS code of Freund and McDermott](#)

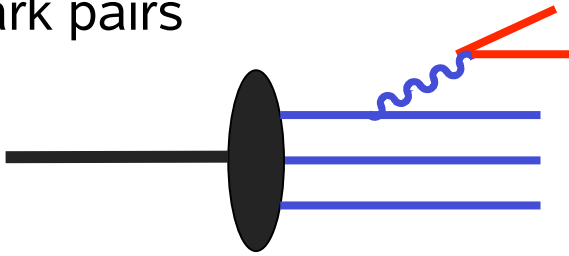
Fragmentation Functions

Access to the [Fragmentation Distribution database](#) site compiled by [Marco Radice](#) and Rainer Jakob.

 Questions and Comments to m.r.whalley@durham.ac.uk
Updated: Dec 11, 2002

the asymmetric sea

- the sea presumably arises when 'primordial' valence quarks emit gluons which in turn split into quark-antiquark pairs, with suppressed splitting into heavier quark pairs

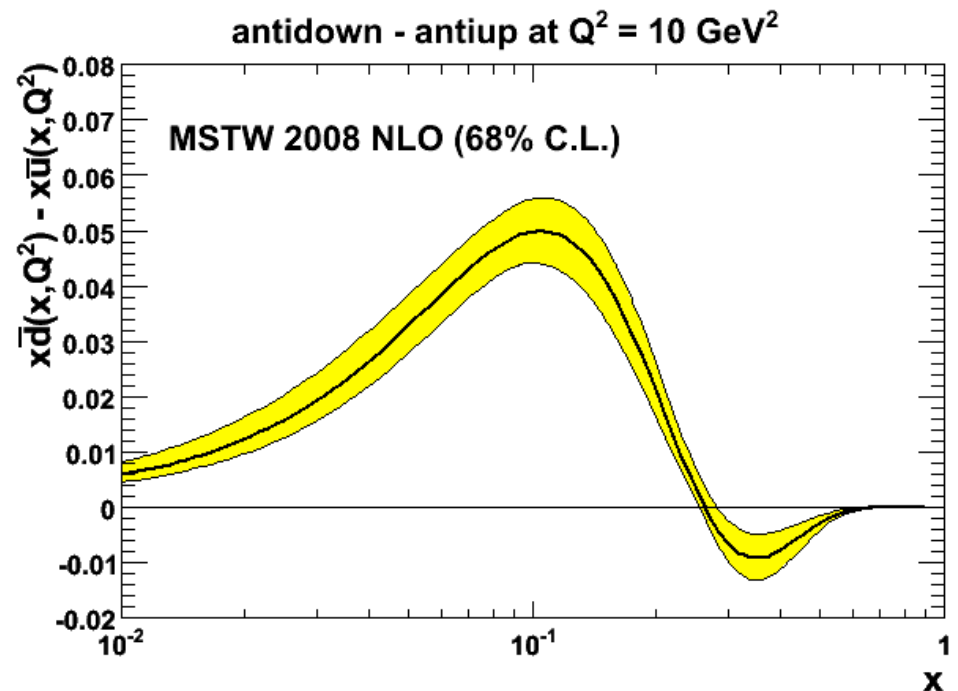


- so we naively expect

$$\bar{u} \approx \bar{d} > \bar{s} > \bar{c} > \dots$$

- but why such a big d-u asymmetry? Meson cloud, Pauli exclusion, ...?

The ratio of Drell-Yan cross sections for $pp, pn \rightarrow \mu^+\mu^- + X$ provides a measure of the difference between the u and d sea quark distributions



strange

earliest pdf fits had SU(3) symmetry: $s(x, Q_0^2) = \bar{s}(x, Q_0^2) = \bar{u}(x, Q_0^2) = \bar{d}(x, Q_0^2)$

later relaxed to include (constant) strange suppression (cf. fragmentation):

$$s(x, Q_0^2) = \bar{s}(x, Q_0^2) = \frac{\kappa}{2} [\bar{u}(x, Q_0^2) + \bar{d}(x, Q_0^2)]$$

with $\kappa = 0.4 - 0.5$

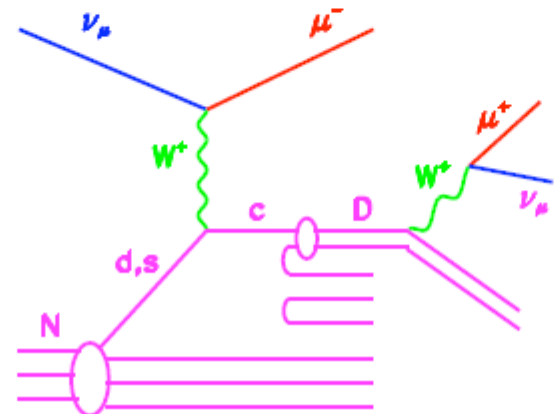
nowadays, dimuon production in νN DIS (CCFR, NuTeV) allows 'direct' determination:

$$\frac{d\sigma}{dx dy} (\nu_\mu (\bar{\nu}_\mu) N \rightarrow \mu^+ \mu^- X) = B_c \mathcal{N} \mathcal{A} \frac{d\sigma}{dx dy} (\nu_\mu s (\bar{\nu}_\mu \bar{s}) \rightarrow c \mu^- (\bar{c} \mu^+) X)$$

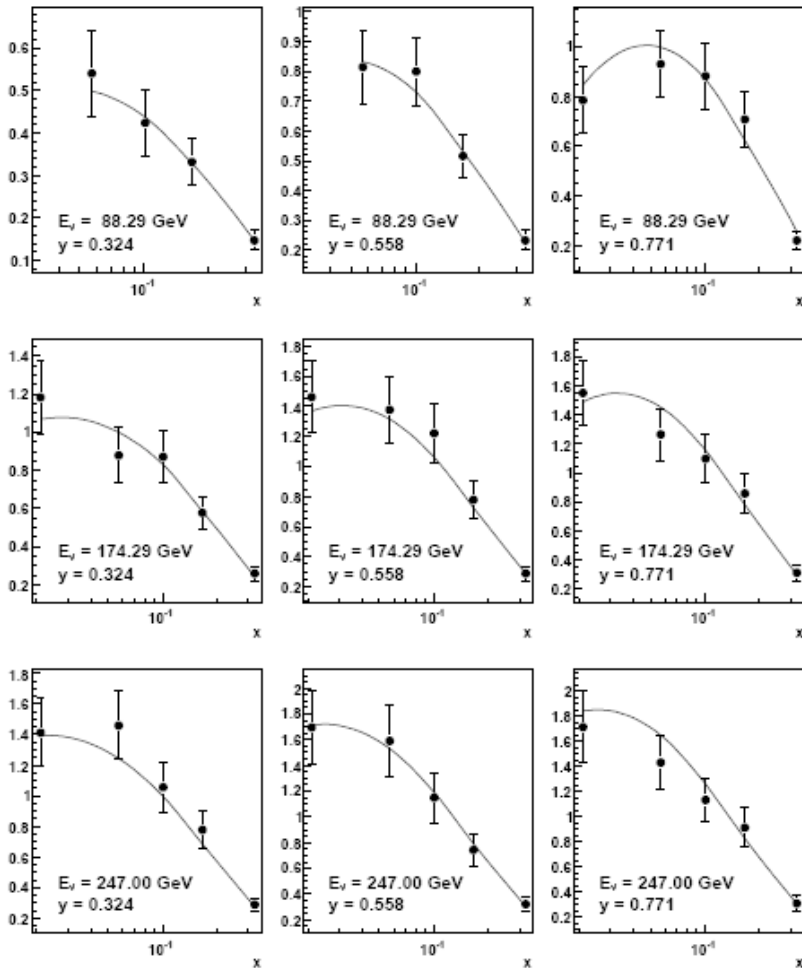
in the range $0.01 < x < 0.4$

data seem to prefer $s(x, Q_0^2) - \bar{s}(x, Q_0^2) \neq 0$

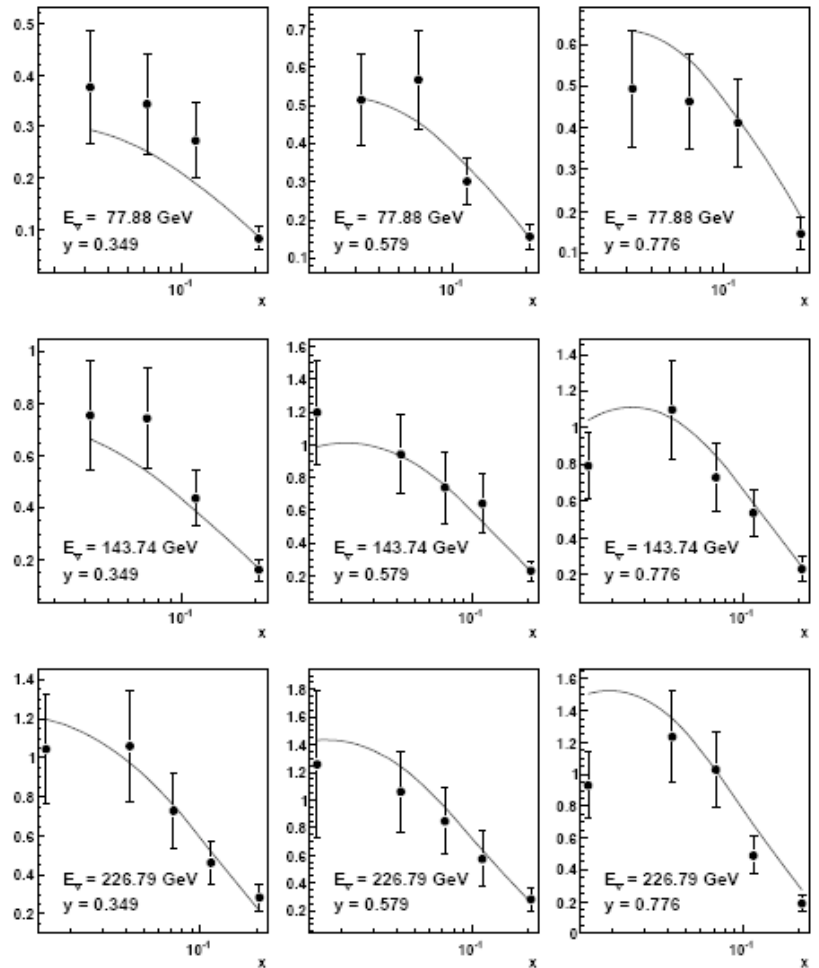
theoretical explanation?!



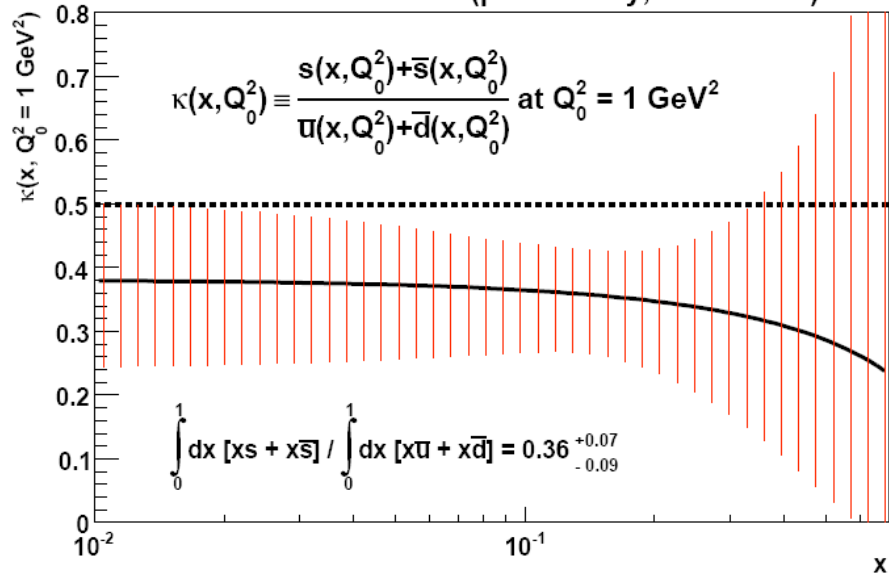
$\text{NuTeV } \frac{100\pi}{G_F^2 M_N E_\nu} \frac{d\sigma}{dx dy} (\nu_\mu N \rightarrow \mu^+ \mu^- X) \text{ in } \text{GeV}^{-2}, \chi^2 = 11/21 \text{ DOF}$
 MSTW NLO PDF fit (preliminary, 17/10/2007)



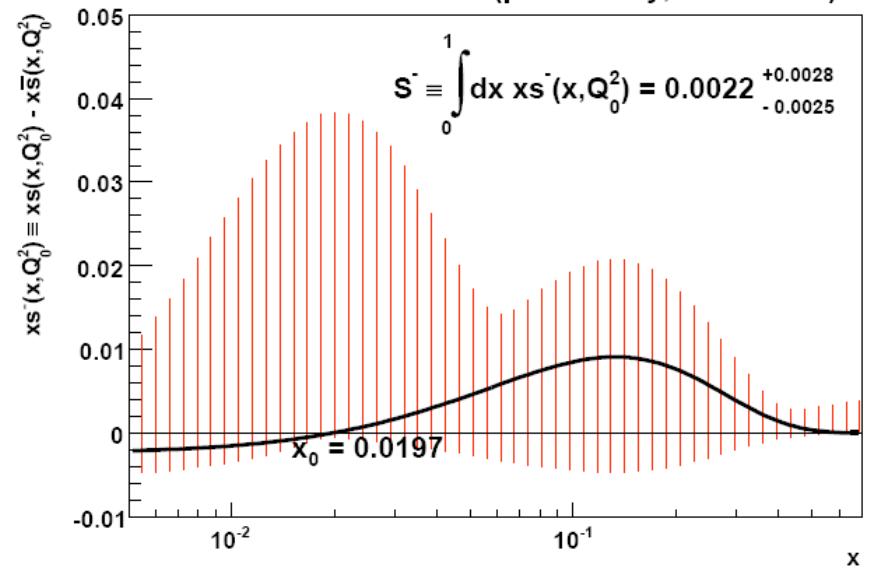
$\text{NuTeV } \frac{100\pi}{G_F^2 M_N E_\nu} \frac{d\sigma}{dx dy} (\bar{\nu}_\mu N \rightarrow \mu^+ \mu^- X) \text{ in } \text{GeV}^{-2}, \chi^2 = 28/19 \text{ DOF}$
 MSTW NLO PDF fit (preliminary, 17/10/2007)



MSTW NNLO PDF fit (preliminary, 30/11/2007)



MSTW NNLO PDF fit (preliminary, 30/11/2007)



MSTW

charm, bottom

considered sufficiently massive to allow pQCD treatment: $g \rightarrow Q\bar{Q}$

distinguish two regimes:

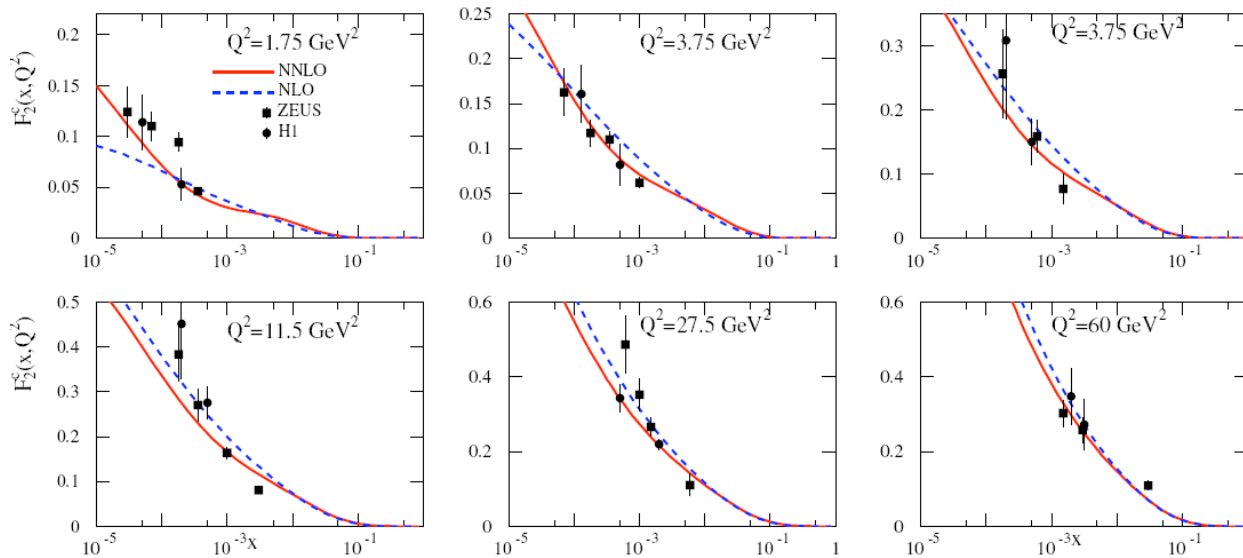
- (i) $Q^2 \sim m_H^2$ include full m_H dependence to get correct threshold behaviour
- (ii) $Q^2 \gg m_H^2$ treat as \sim massless partons to resum $\alpha_s^n \log^n(Q^2/m_H^2)$ via DGLAP

FFNS: OK for (i) only **ZM-VFNS:** OK for (ii) only

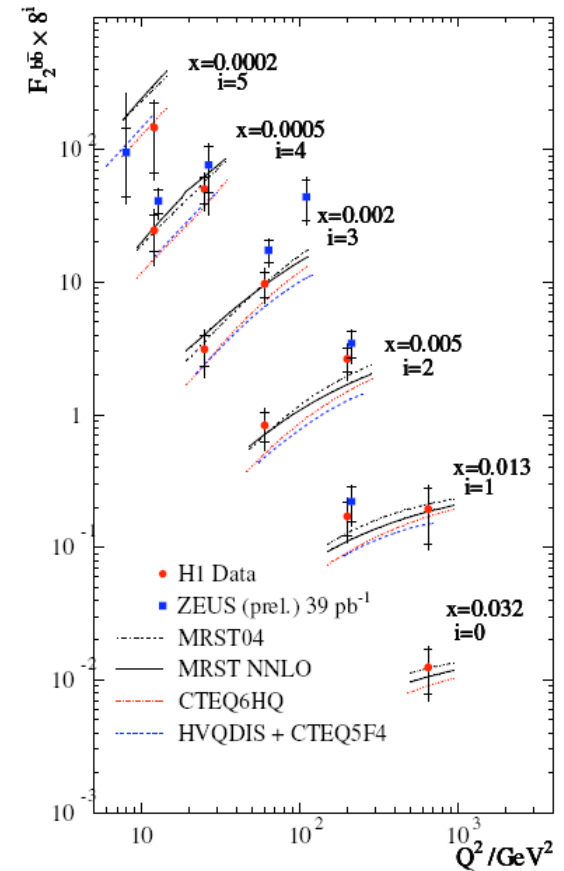
consistent **GM(=general mass)-VFNS** now available (e.g. ACOT(χ), Roberts-Thorne) which interpolates smoothly between the two regimes

Note: definition of these is tricky and non-unique (ambiguity in assignment of $O(m_H^2/Q^2)$ contributions), and the implementation of improved treatment (e.g. in going from MRST2006 to MSTW 2008) can have a big effect on light partons

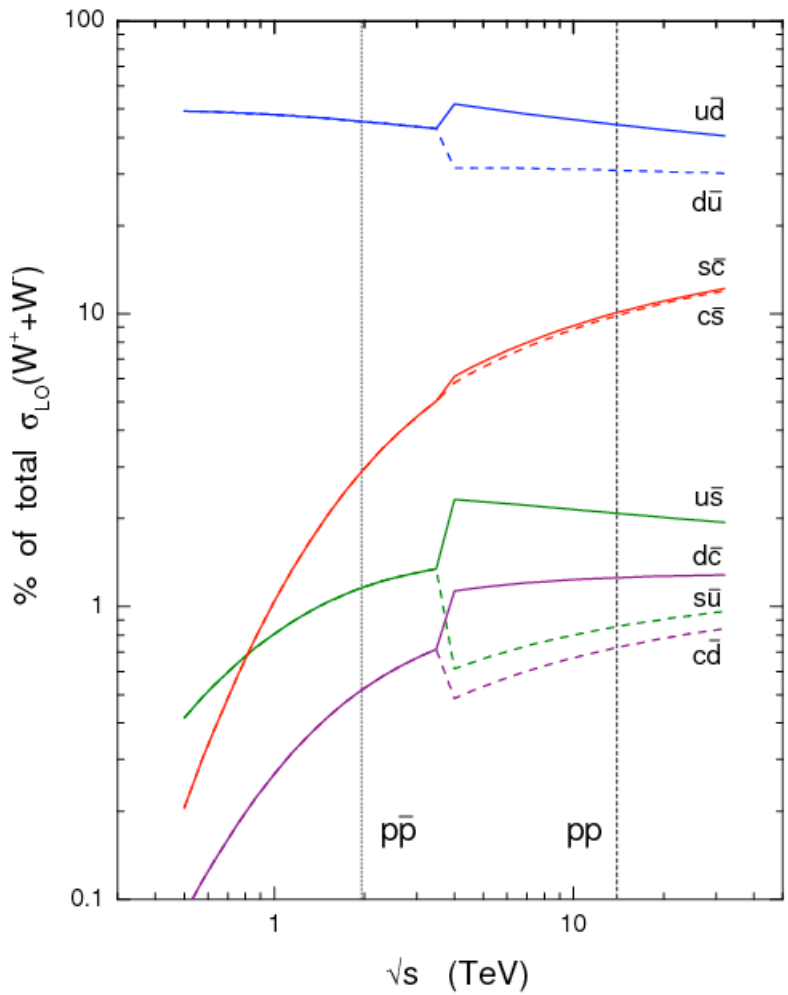
charm and bottom structure functions



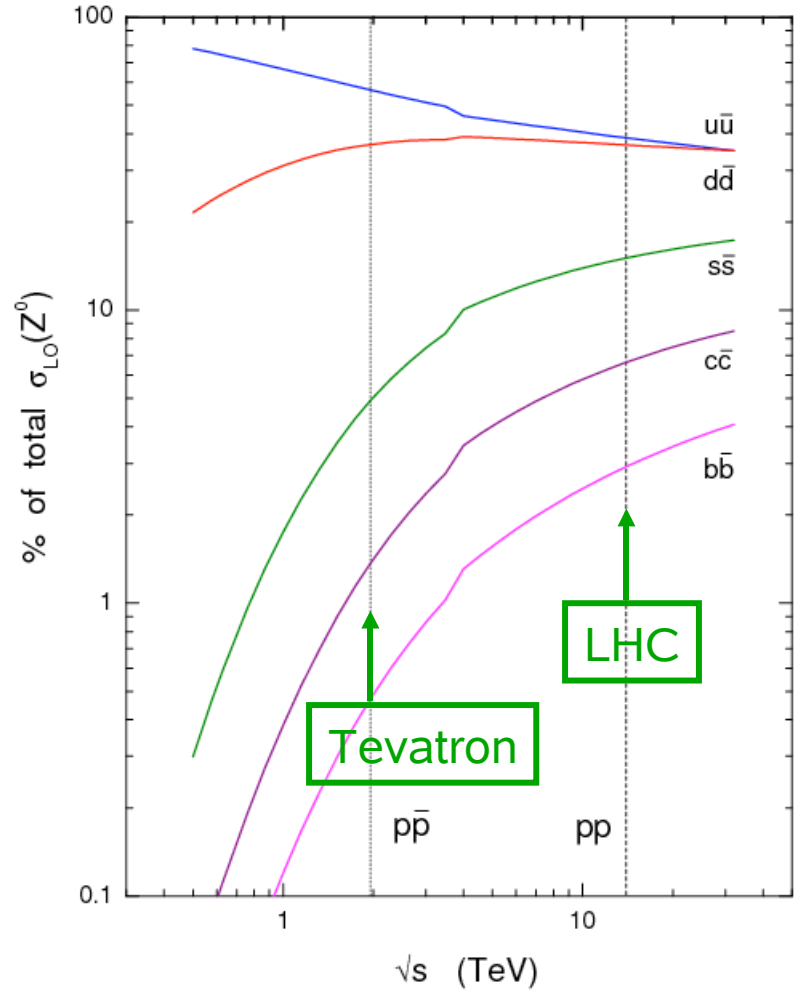
MSTW 2008



flavour decomposition of W cross sections



flavour decomposition of Z⁰ cross sections



at LHC, ~30% of W and Z total cross sections involves s,c,b quarks

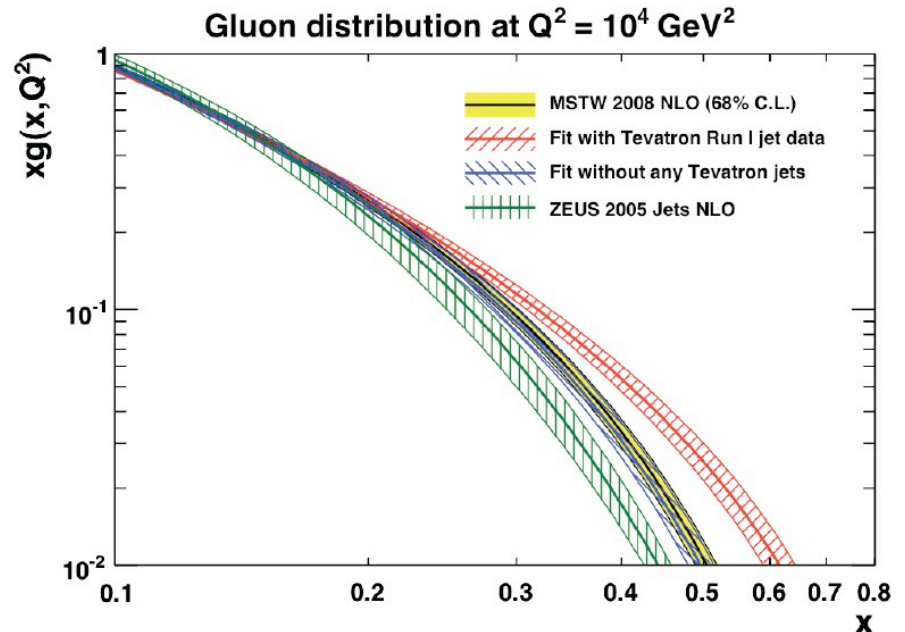
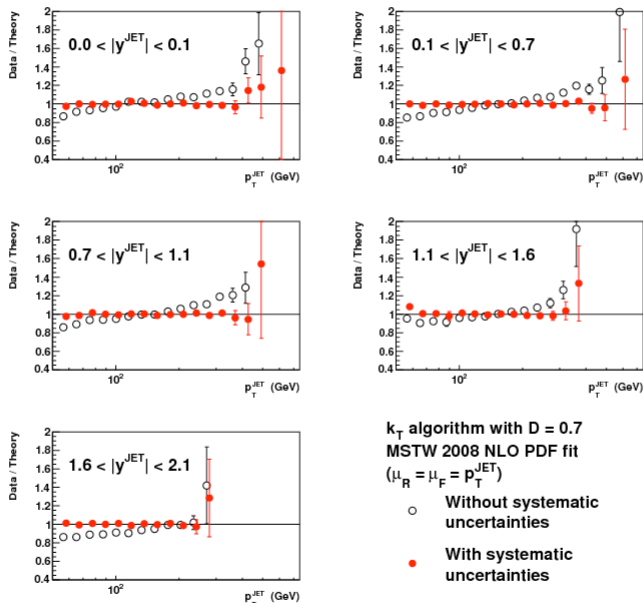
why do 'best fit' pdfs and errors differ?

- different data sets in fit
 - different subselection of data (cuts etc)
 - different treatment of exp. sys. errors
- different choice of
 - tolerance to define $\pm \delta f_i$ ($\Delta\chi^2=1$ or ??)
 - factorisation/renormalisation scheme/scale
 - Q_0^2
 - parametric form at Q_0^2 : $Ax^a(1-x)^b[.]$ etc
 - α_s
 - treatment of heavy flavours
 - theoretical assumptions about $x \rightarrow 0, 1$ behaviour
 - theoretical assumptions about sea flavour symmetry
 - evolution and cross section codes (removable differences!)
generally not straightforward to disentangle!

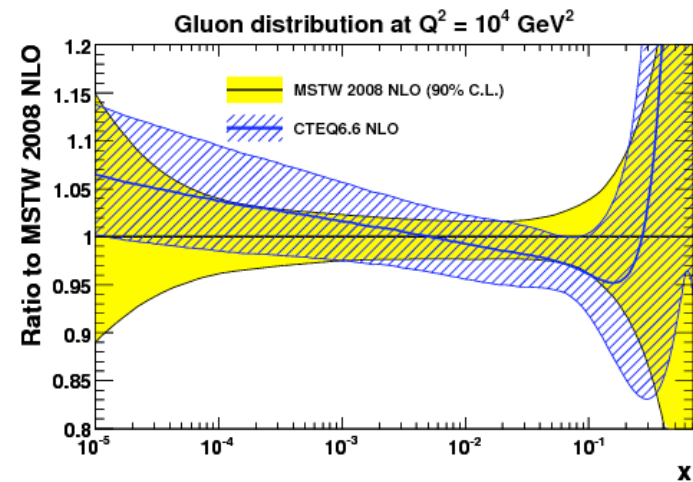
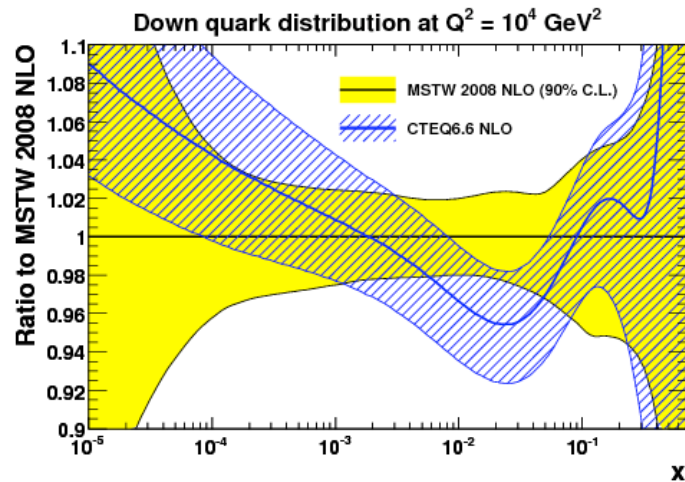
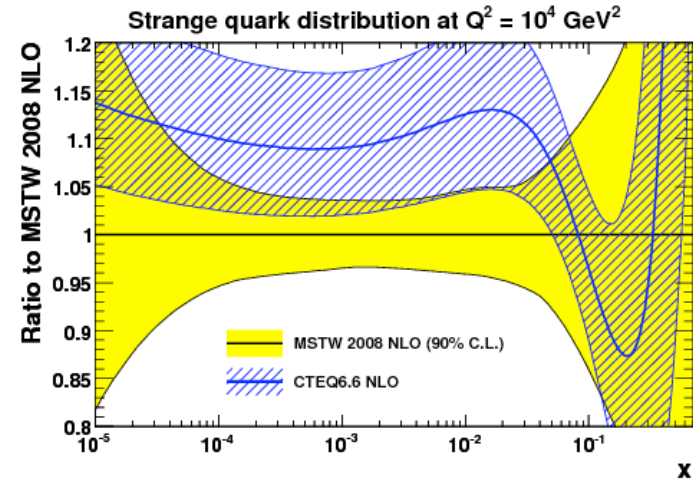
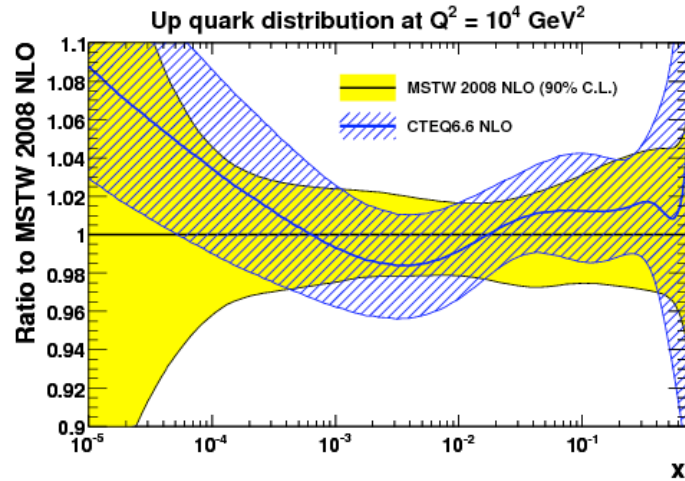
impact of high E_T jet data on fits

- a distinguishing feature of pdf sets is whether they use (MRST/MSTW, CTEQ,...) or do not use (H1, ZEUS, Alekhin, NNPDF,...) Tevatron jet data in the fit: the impact is on the *high- x gluon*
(Note: Run II data requires slightly softer gluon than Run I data)
- the (still) missing ingredient is the full NNLO pQCD correction to the cross section, but not expected to have much impact in practice
- note that large-mass pN Drell-Yan also probes the gluon indirectly via $g \rightarrow q\bar{q}$ generation of *sea antiquarks at high x*

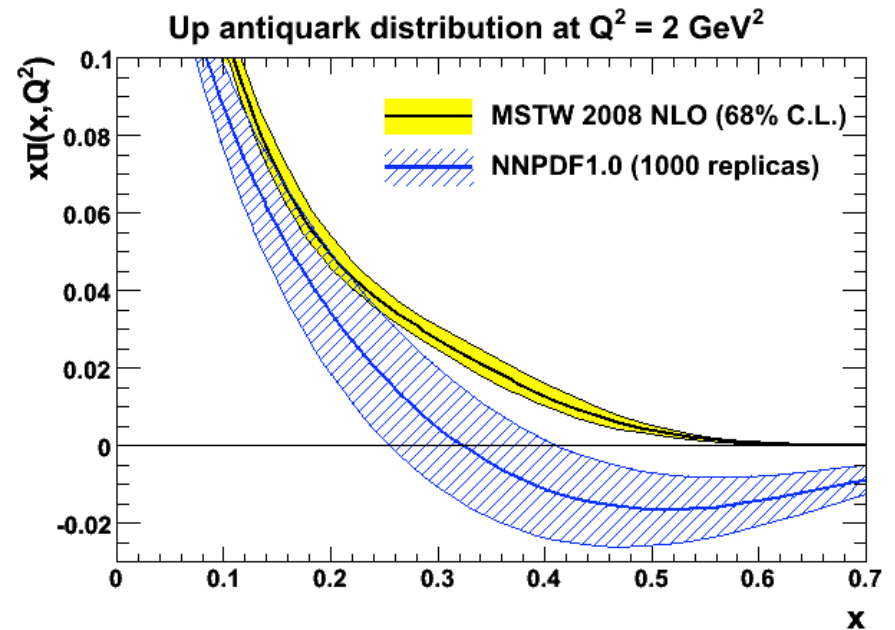
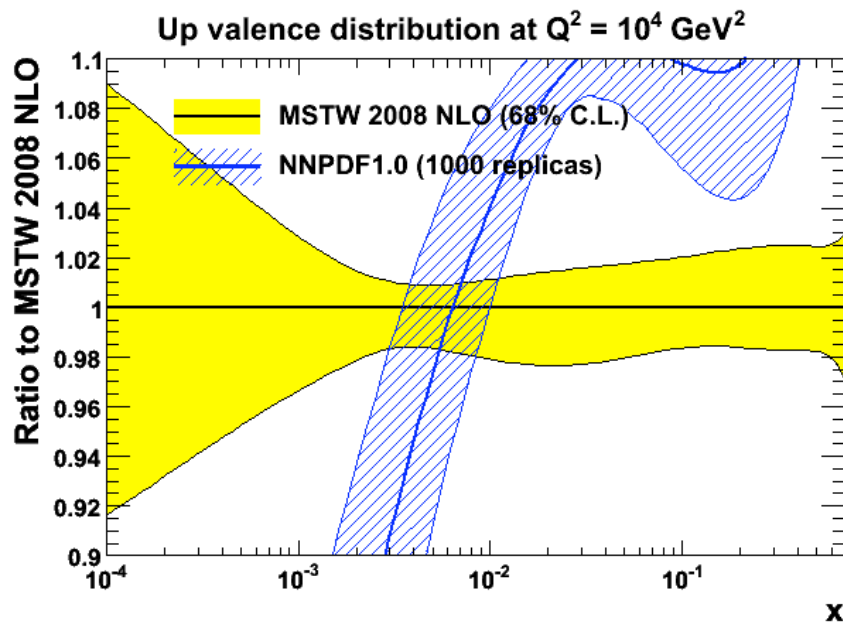
CDF Run II inclusive jet data, $\chi^2 = 56$ for 76 pts.



MSTW2008(NLO) vs. CTEQ6.6

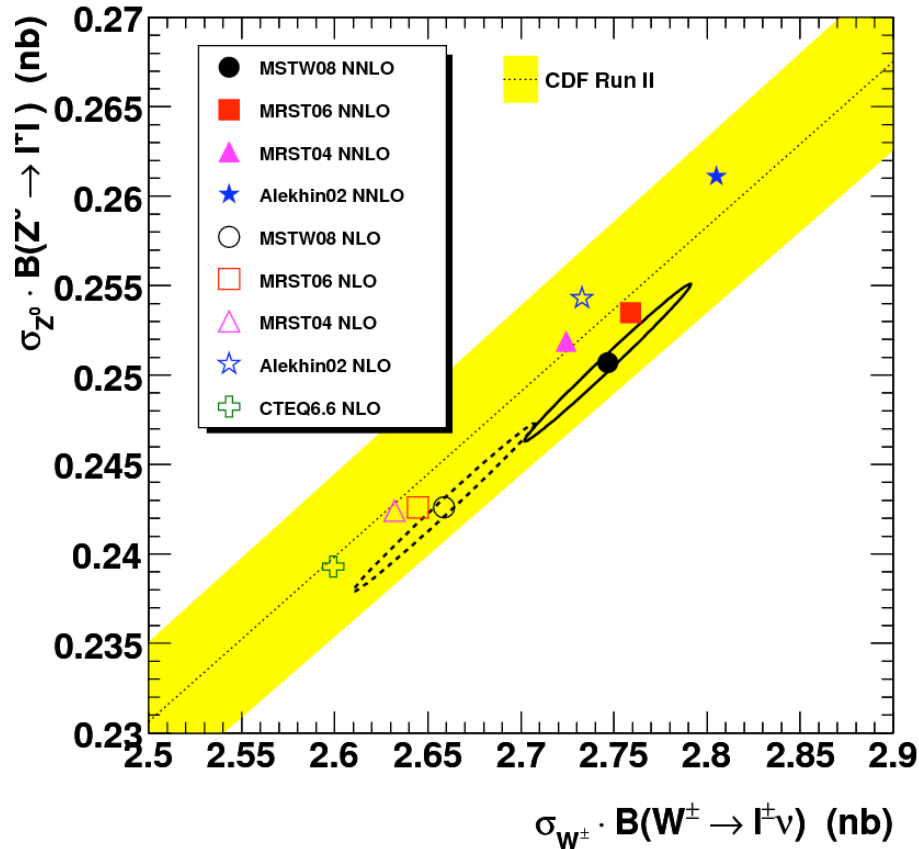


MSTW2008(NLO) vs. NNPDF1.0

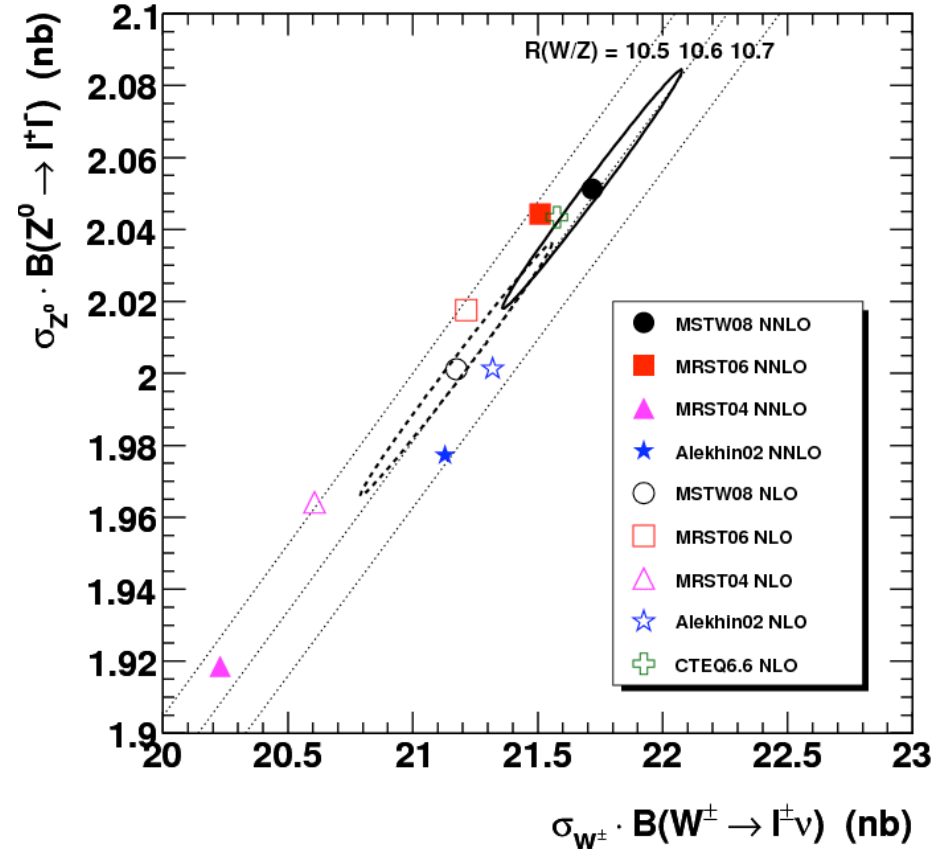


$\sigma(W), \sigma(Z)$ @ Tevatron & LHC

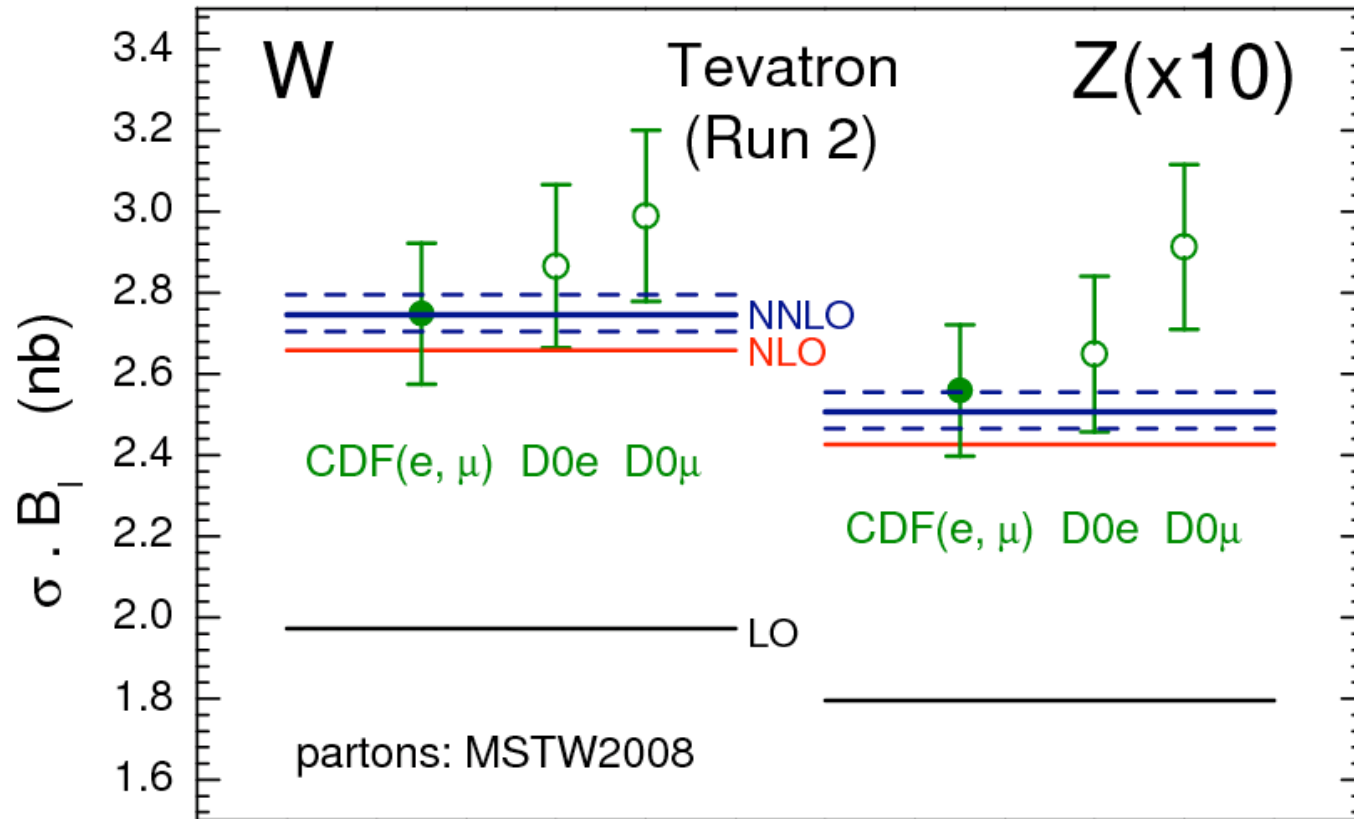
W and Z total cross sections at the Tevatron



W and Z total cross sections at the LHC



CDF 2007: $R = 10.84 \pm 0.15$ (stat) ± 0.14 (sys)



Tevatron, $\sqrt{s} = 1.96$ TeV	$B_{l\nu} \cdot \sigma_W$ (nb)	$B_{l+l-} \cdot \sigma_Z$ (nb)	R_{WZ}
MSTW 2008 LO	$1.963^{+0.025}_{-0.028}$ (+1.2%) (-1.4%)	$0.1788^{+0.0023}_{-0.0025}$ (+1.3%) (-1.4%)	$10.98^{+0.02}_{-0.03}$ (+0.2%) (-0.3%)
MSTW 2008 NLO	$2.659^{+0.057}_{-0.045}$ (+2.1%) (-1.7%)	$0.2426^{+0.0054}_{-0.0043}$ (+2.2%) (-1.8%)	$10.96^{+0.03}_{-0.02}$ (+0.3%) (-0.2%)
MSTW 2008 NNLO	$2.747^{+0.049}_{-0.042}$ (+1.8%) (-1.5%)	$0.2507^{+0.0048}_{-0.0041}$ (+1.9%) (-1.6%)	$10.96^{+0.03}_{-0.03}$ (+0.2%) (-0.2%)

LHC, $\sqrt{s} = 10$ TeV	$B_{l\nu} \cdot \sigma_W$ (nb)	$B_{l+l-} \cdot \sigma_Z$ (nb)	R_{WZ}
MSTW 2008 LO	$12.57^{+0.13}_{-0.19}$ (+1.1%) (-1.5%)	$1.163^{+0.011}_{-0.017}$ (+1.0%) (-1.5%)	$10.81^{+0.02}_{-0.02}$ (+0.2%) (-0.2%)
MSTW 2008 NLO	$14.92^{+0.31}_{-0.24}$ (+2.1%) (-1.6%)	$1.390^{+0.029}_{-0.022}$ (+2.1%) (-1.5%)	$10.73^{+0.02}_{-0.02}$ (+0.2%) (-0.2%)
MSTW 2008 NNLO	$15.35^{+0.26}_{-0.25}$ (+1.7%) (-1.6%)	$1.429^{+0.024}_{-0.022}$ (+1.7%) (-1.6%)	$10.74^{+0.02}_{-0.02}$ (+0.2%) (-0.2%)

LHC, $\sqrt{s} = 14$ TeV	$B_{l\nu} \cdot \sigma_W$ (nb)	$B_{l+l-} \cdot \sigma_Z$ (nb)	R_{WZ}
MSTW 2008 LO	$18.51^{+0.22}_{-0.32}$ (+1.2%) (-1.7%)	$1.736^{+0.019}_{-0.028}$ (+1.1%) (-1.6%)	$10.66^{+0.02}_{-0.02}$ (+0.2%) (-0.2%)
MSTW 2008 NLO	$21.17^{+0.42}_{-0.36}$ (+2.0%) (-1.7%)	$2.001^{+0.040}_{-0.032}$ (+2.0%) (-1.6%)	$10.58^{+0.02}_{-0.02}$ (+0.2%) (-0.2%)
MSTW 2008 NNLO	$21.72^{+0.36}_{-0.36}$ (+1.7%) (-1.7%)	$2.051^{+0.035}_{-0.033}$ (+1.7%) (-1.6%)	$10.59^{+0.02}_{-0.03}$ (+0.2%) (-0.3%)

Note: at NNLO, factorisation and renormalisation scale variation M/2 \rightarrow 2M gives an additional $\pm 2\%$ change in the LHC cross sections

MSTW 2008 update

- new data (see next slide)
- new theory/infrastructure
 - δf_i from new dynamic tolerance method
 - new definition of α_s (no more Λ_{QCD})
 - new GM-VFNS for c, b (see Martin et al., arXiv:0706.0459)
 - new fitting codes: FEWZ, fastNLO
 - new grids: denser, broader coverage
 - slightly extended parametrisation at Q_0^2 : 34-4=30 free parameters including α_s

data sets used in fit

Data set	$N_{\text{pts.}}$
H1 MB 99 e^+p NC	8
H1 MB 97 e^+p NC	64
H1 low Q^2 96–97 e^+p NC	80
H1 high Q^2 98–99 e^-p NC	126
H1 high Q^2 99–00 e^+p NC	147
ZEUS SVX 95 e^+p NC	30
ZEUS 96–97 e^+p NC	144
ZEUS 98–99 e^-p NC	92
ZEUS 99–00 e^+p NC	90
H1 99–00 e^+p CC	28
ZEUS 99–00 e^+p CC	30
H1/ZEUS $e^\pm p F_2^{\text{charm}}$	83
H1 99–00 e^+p incl. jets	24
ZEUS 96–97 e^+p incl. jets	30
ZEUS 98–00 $e^\pm p$ incl. jets	30
DØ II $p\bar{p}$ incl. jets	110
CDF II $p\bar{p}$ incl. jets	76
CDF II $W \rightarrow l\nu$ asym.	22
DØ II $W \rightarrow l\nu$ asym.	10
DØ II Z rap.	28
CDF II Z rap.	29

Data set	$N_{\text{pts.}}$
BCDMS $\mu p F_2$	163
BCDMS $\mu d F_2$	151
NMC $\mu p F_2$	123
NMC $\mu d F_2$	123
NMC $\mu n/\mu p$	148
E665 $\mu p F_2$	53
E665 $\mu d F_2$	53
SLAC $ep F_2$	37
SLAC $ed F_2$	38
NMC/BCDMS/SLAC F_L	31
E866/NuSea pp DY	184
E866/NuSea pd/pp DY	15
NuTeV $\nu N F_2$	53
CHORUS $\nu N F_2$	42
NuTeV $\nu N xF_3$	45
CHORUS $\nu N xF_3$	33
CCFR $\nu N \rightarrow \mu\mu X$	86
NuTeV $\nu N \rightarrow \mu\mu X$	84
All data sets	2743

- Red = New w.r.t. MRST 2006 fit.

MSTW input parametrisation

At input scale $Q_0^2 = 1 \text{ GeV}^2$:

$$xu_v = A_u x^{\eta_1} (1-x)^{\eta_2} (1 + \epsilon_u \sqrt{x} + \gamma_u x)$$

$$xd_v = A_d x^{\eta_3} (1-x)^{\eta_4} (1 + \epsilon_d \sqrt{x} + \gamma_d x)$$

$$xS = A_S x^{\delta_S} (1-x)^{\eta_S} (1 + \epsilon_S \sqrt{x} + \gamma_S x)$$

$$x\bar{d} - x\bar{u} = A_\Delta x^{\eta_\Delta} (1-x)^{\eta_S+2} (1 + \gamma_\Delta x + \delta_\Delta x^2)$$

$$xg = A_g x^{\delta_g} (1-x)^{\eta_g} (1 + \epsilon_g \sqrt{x} + \gamma_g x) + A_{g'} x^{\delta_{g'}} (1-x)^{\eta_{g'}}$$

$$xs + x\bar{s} = A_+ x^{\delta_+} (1-x)^{\eta_+} (1 + \epsilon_S \sqrt{x} + \gamma_S x)$$

$$xs - x\bar{s} = A_- x^{\delta_-} (1-x)^{\eta_-} (1 - x/x_0)$$

Note: 20 parameters allowed to go free for eigenvector PDF sets, *cf.* 15 for MRST sets

PDF eigenvector sets

- the Hessian matrix $\chi^2 - \chi_{min}^2 \equiv \Delta\chi^2 = \sum_{i,j} H_{ij} (a_i - a_i^{(0)}) (a_j - a_j^{(0)})$
- diagonalise the **covariance matrix** $C \equiv H^{-1}$

$$\sum_j C_{ij} v_{jk} = \lambda_k v_{ik}$$

- produce **eigenvector pdf sets** S_k^\pm with parameters a_i shifted from the global minimum

$$a_i(S_k^\pm) = a_i^0 \pm t \sqrt{\lambda_k} v_{ik}$$

with t adjusted to give the desired **tolerance** $T = \sqrt{\Delta\chi_{global}^2}$

- then calculate

$$\Delta F = \frac{1}{2} \sqrt{\sum_k [F(S_k^+) - F(S_k^-)]^2},$$

criteria for choice of tolerance T

Parameter-fitting criterion

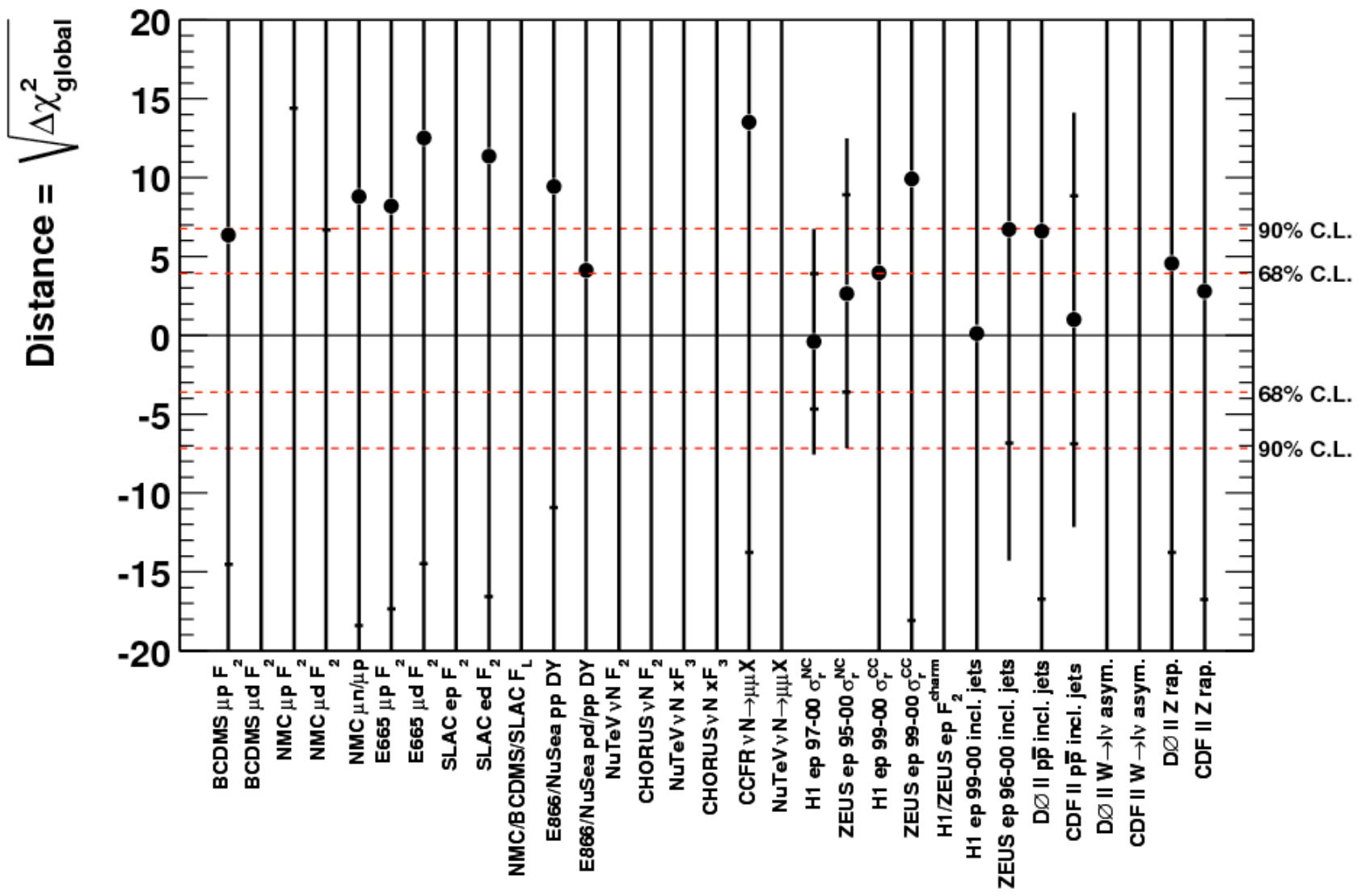
- $T^2 = 1$ for 68% (1σ) c.l., $T^2 = 2.71$ for 90% c.l., etc
- appropriate if fitting consistent data sets with ideal Gaussian errors to a well-defined theory
- in practice: minor inconsistencies between fitted data sets, and unknown experimental and theoretical uncertainties, so
- therefore not appropriate for *global* PDF analysis

Hypothesis-testing criterion (CTEQ)

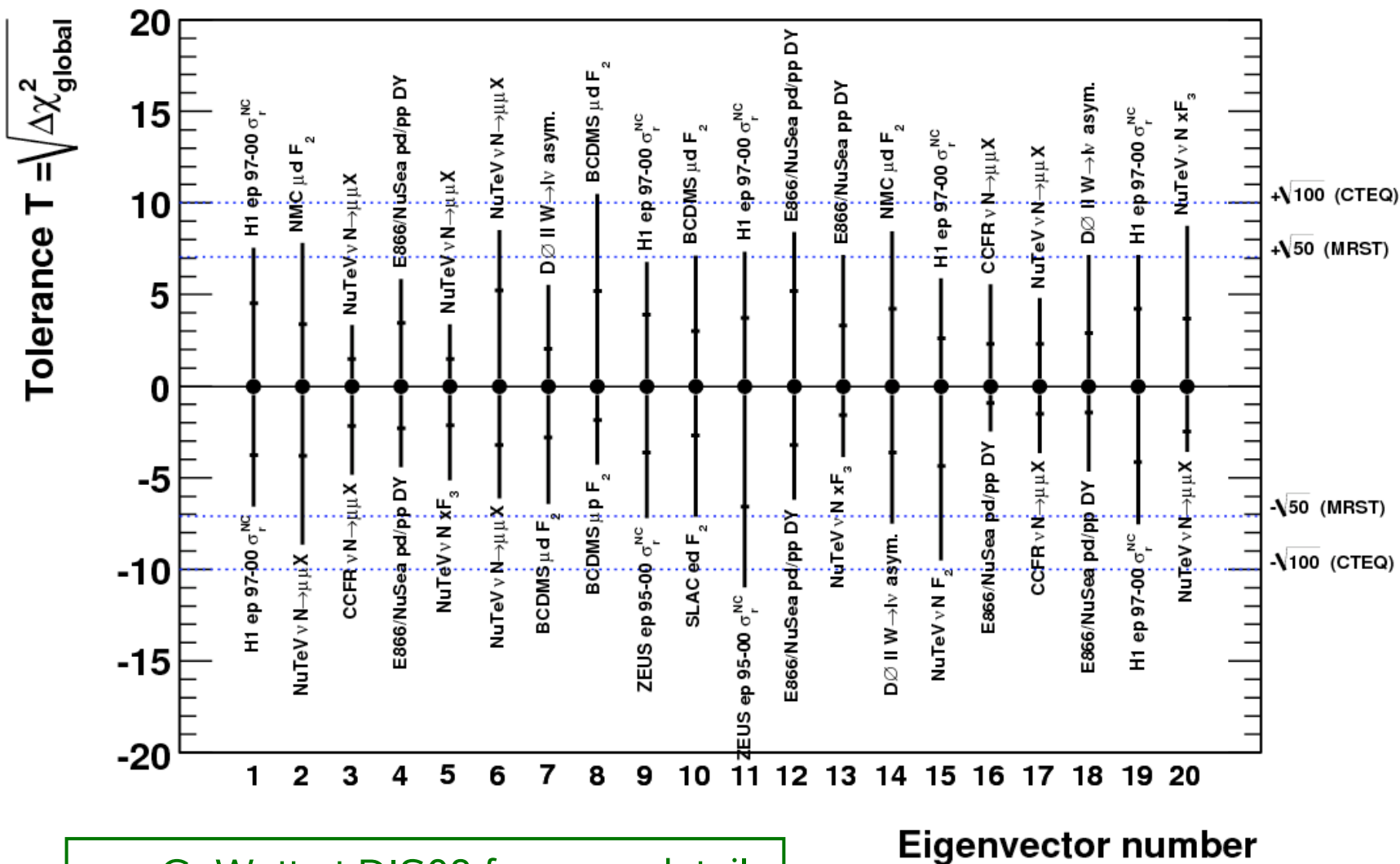
- much weaker than the parameter-fitting criterion: treat eigenvector pdf sets as **alternative hypotheses**
- determine T^2 from the criterion that **each data set should be described within its 90% c.l. limit**

Eigenvector number 9

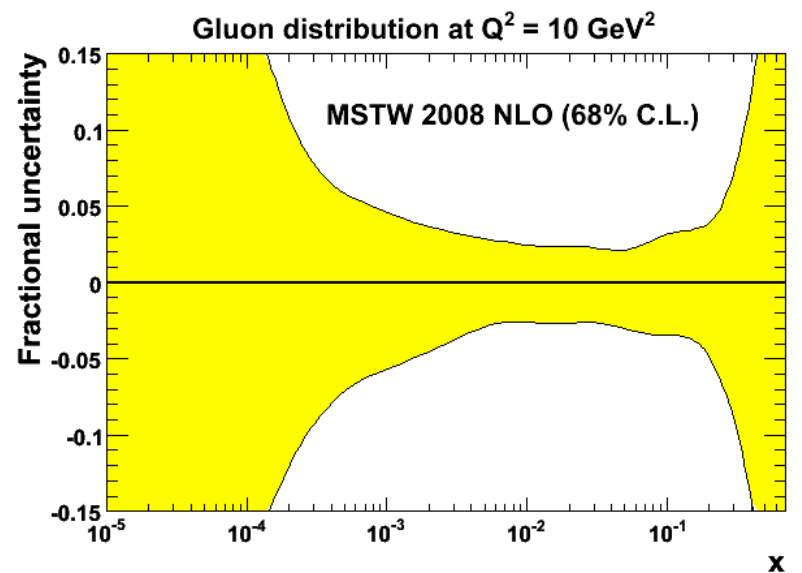
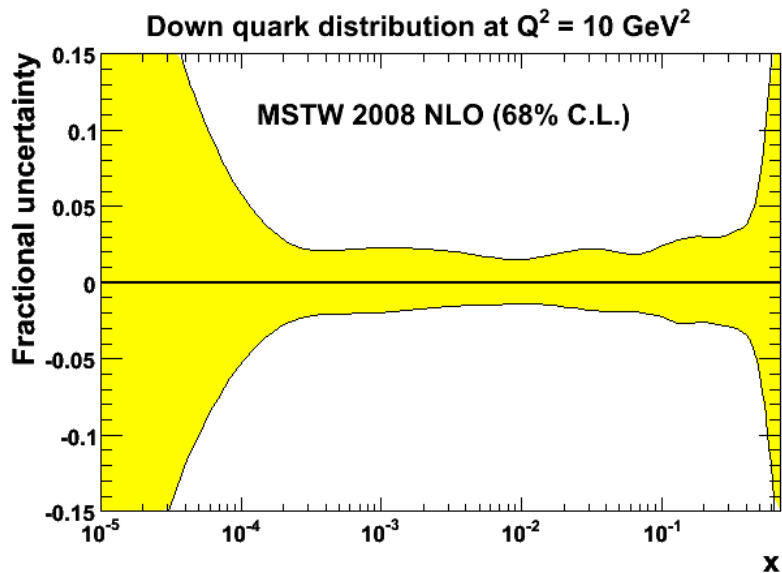
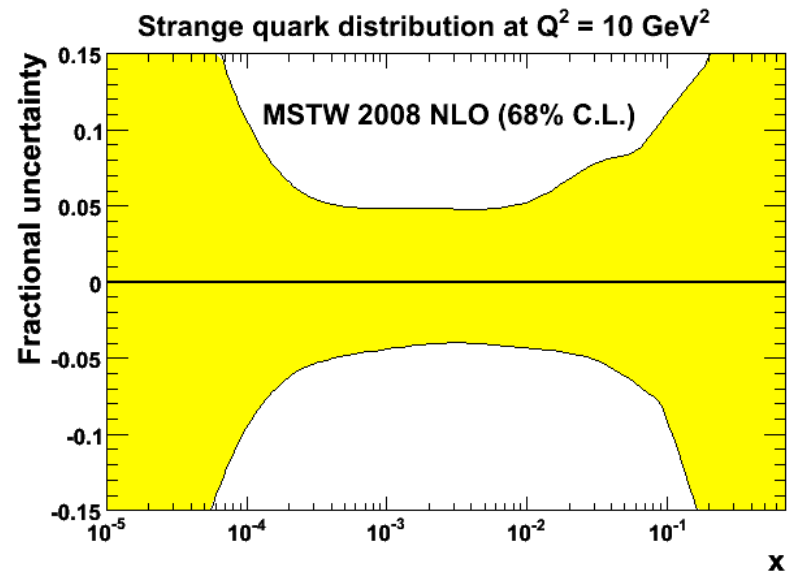
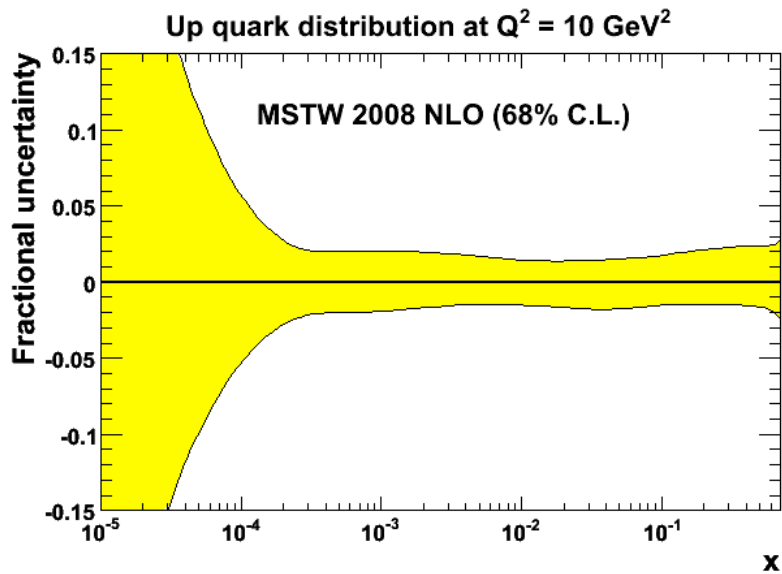
MSTW 2008 NLO PDF fit



MSTW 2008 NLO PDF fit



see G. Watt at DIS08 for more details

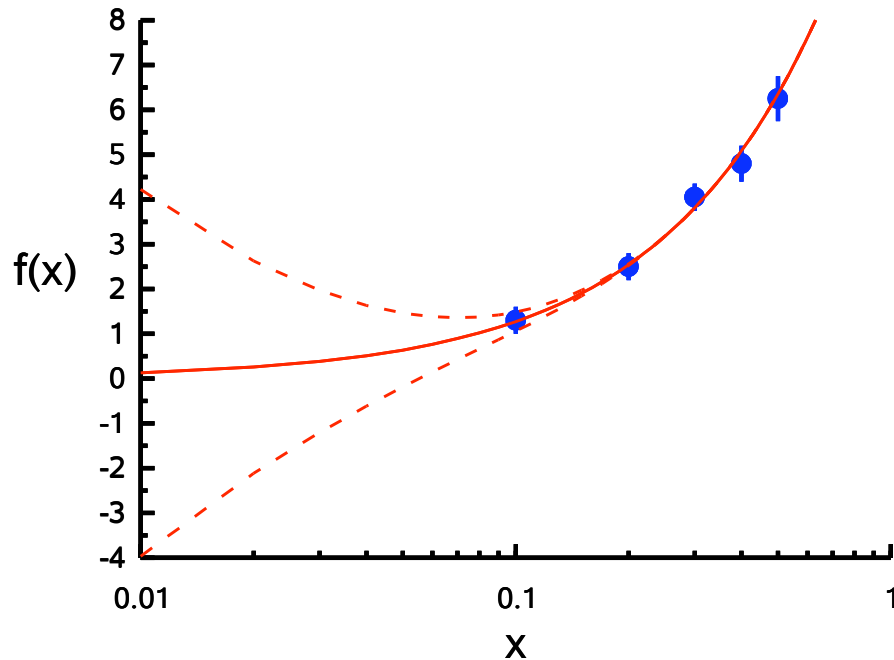


summary

- precision phenomenology at high-energy colliders such as the **LHC** requires an accurate knowledge of the distribution functions of partons in hadrons
- determining pdfs from global fits to data is now a major industry... the MSTW collaboration is about to release its latest (2008) LO, NLO, NNLO sets
- watch out for differences between pdf sets > quoted uncertainties!
- at a proton-proton collider such as the **LHC**, the quark **sea** plays an important role in new-physics processes; parton analyses reveal interesting quark flavour asymmetries, which are not well understood theoretically

extra slides

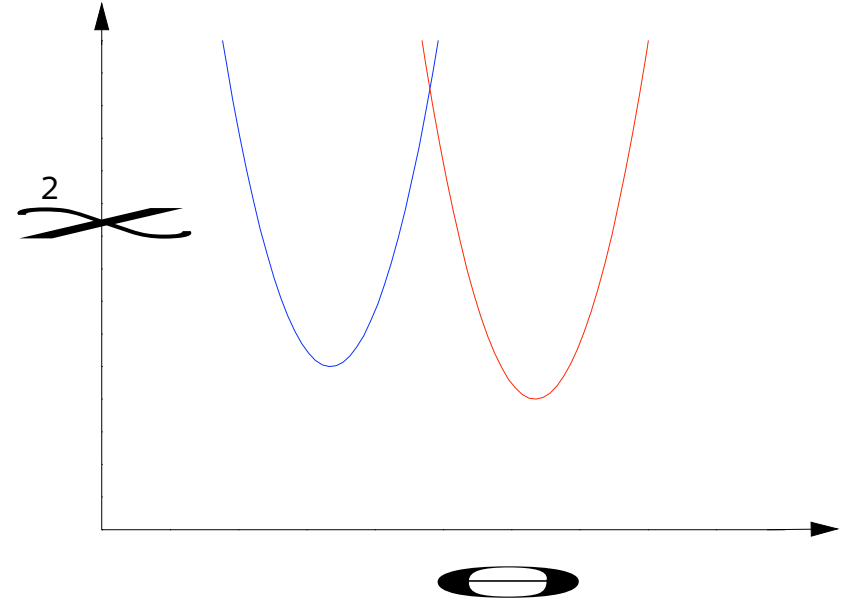
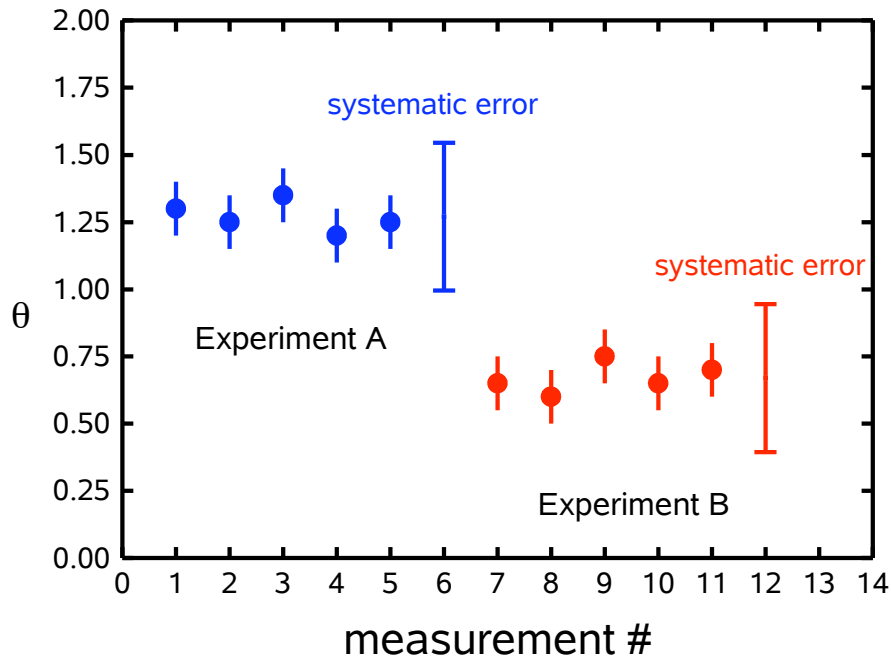
extrapolation uncertainties



theoretical insight/guess: $f \sim Ax$ as $x \rightarrow 0$

theoretical insight/guess: $f \sim \pm Ax^{-0.5}$ as $x \rightarrow 0$

tensions within the global fit



- with dataset A in fit, $\Delta\chi^2=1$; with A and B in fit, $\Delta\chi^2=?$
- in practice modest 'tensions' between data sets do arise, for example,
 - between DIS data sets (e.g. $\mu\mathbf{H}$ and $\nu\mathbf{N}$ data, α_s , ...)
 - when jet and Drell-Yan data are combined with DIS data

F_L

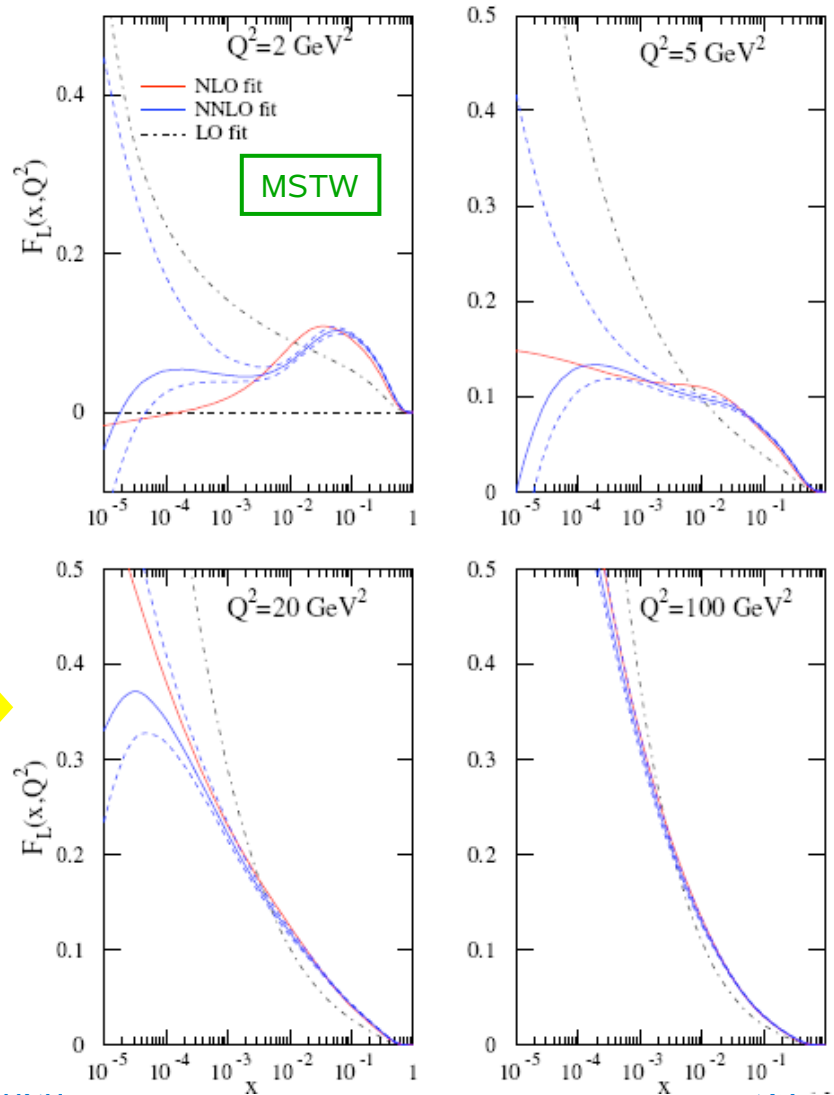
$$\frac{\partial F_2}{\partial \ln Q^2} \simeq \alpha_S P^{qg} \otimes g + \dots$$

$$F_L \simeq \alpha_S C_{Lg} \otimes g + \dots$$

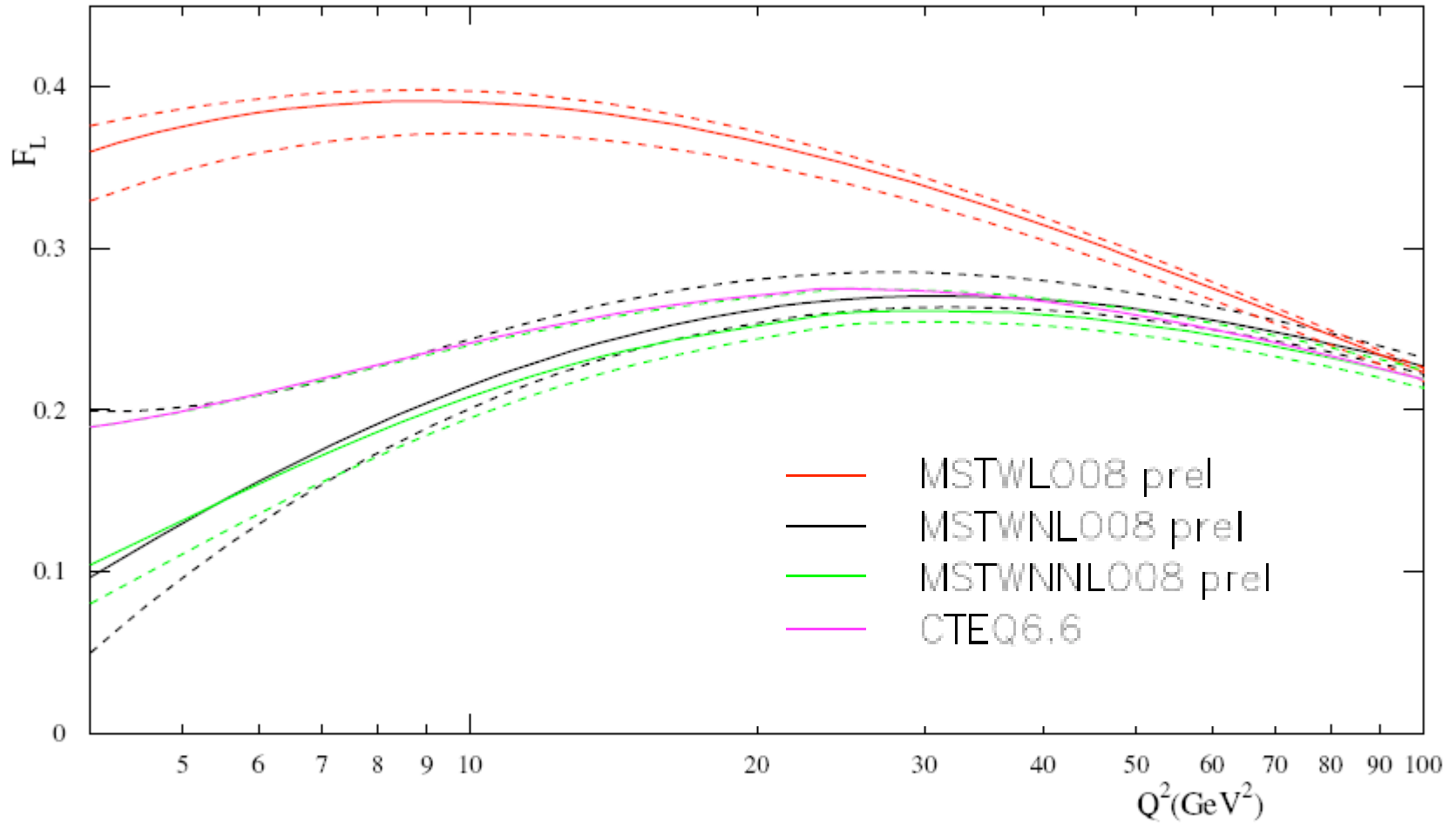
- an independent measurement of the small-x gluon
- a test of the assumptions in the DGLAP LT pQCD analysis of small-x F_2
- visible instability in MSTW analysis at small x and Q^2 (impact of negative gluon and large NNLO coefficient function)
- higher-order $\ln(1/x)$ and higher-twist contributions could be important



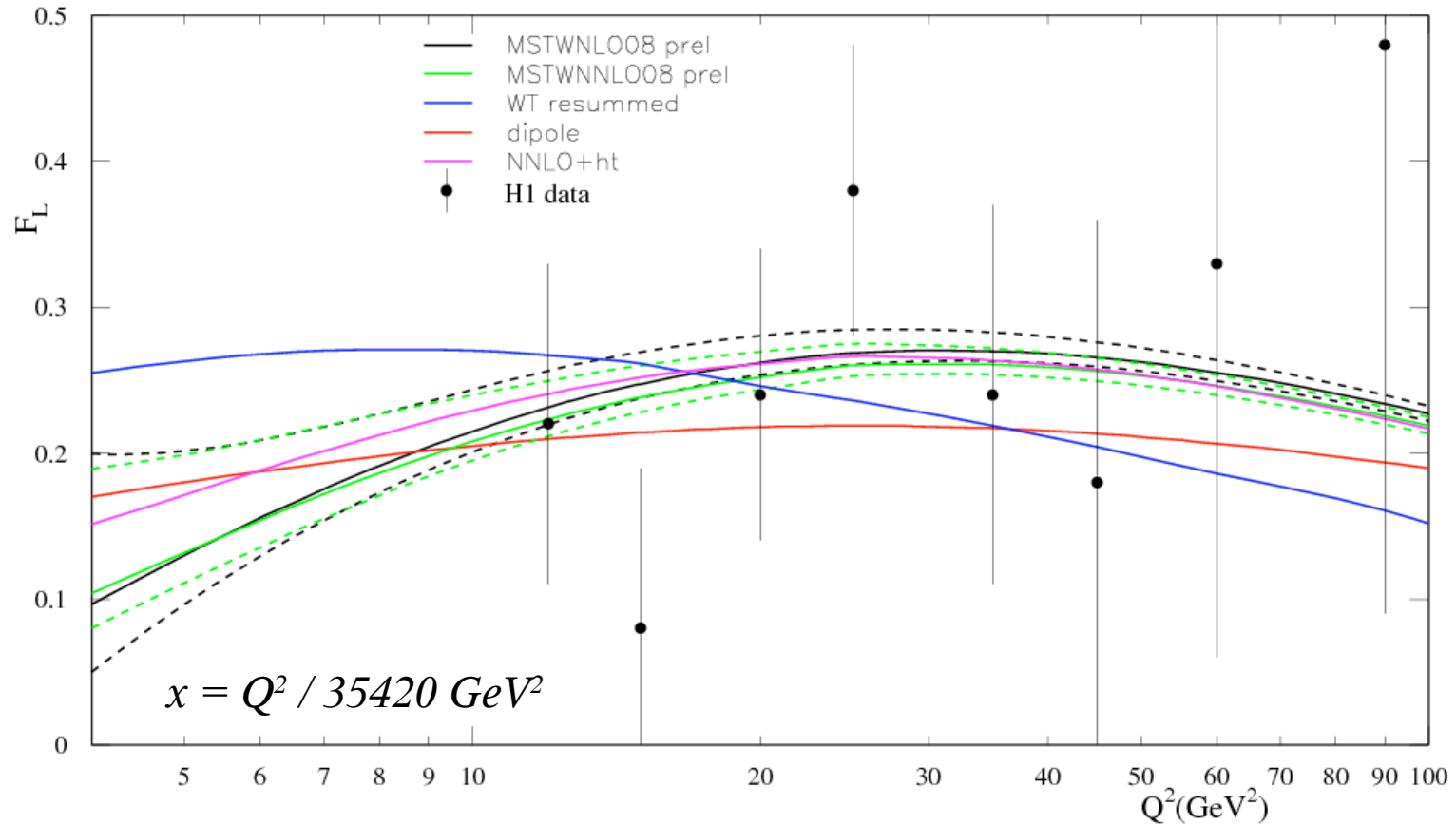
F_L LO, NLO and NNLO



pQCD F_L predictions



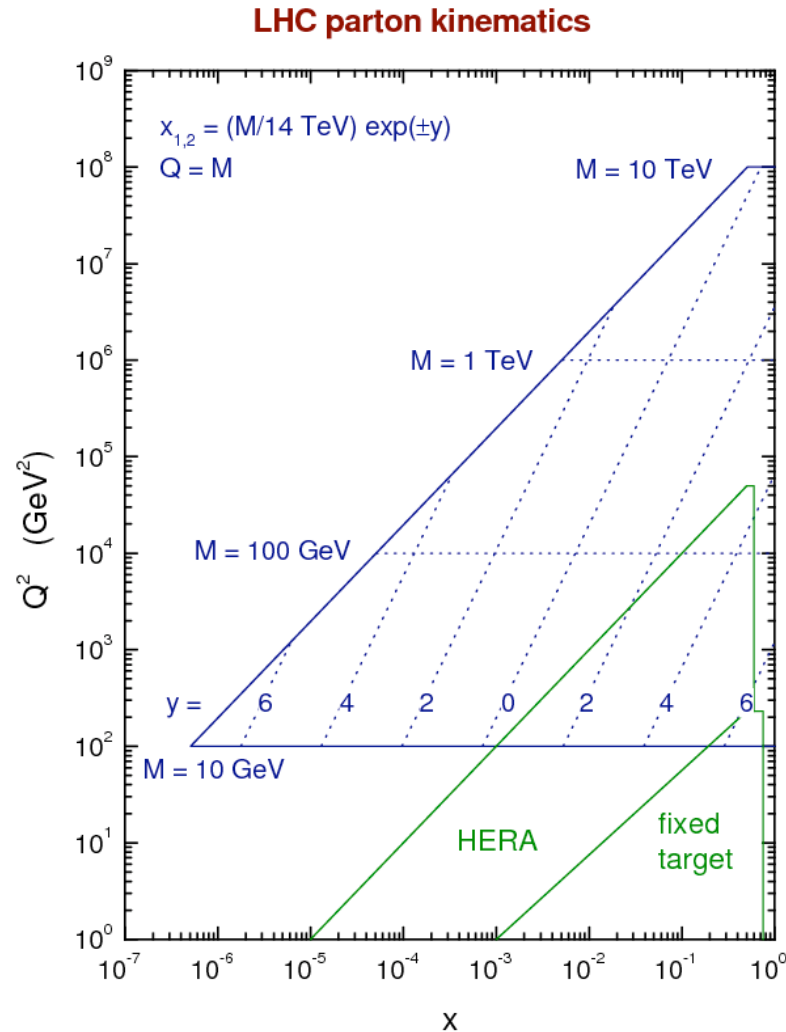
F_L predictions and H1 data



Thorne arXiv:0808.1845

impact of LHC measurements on pdfs

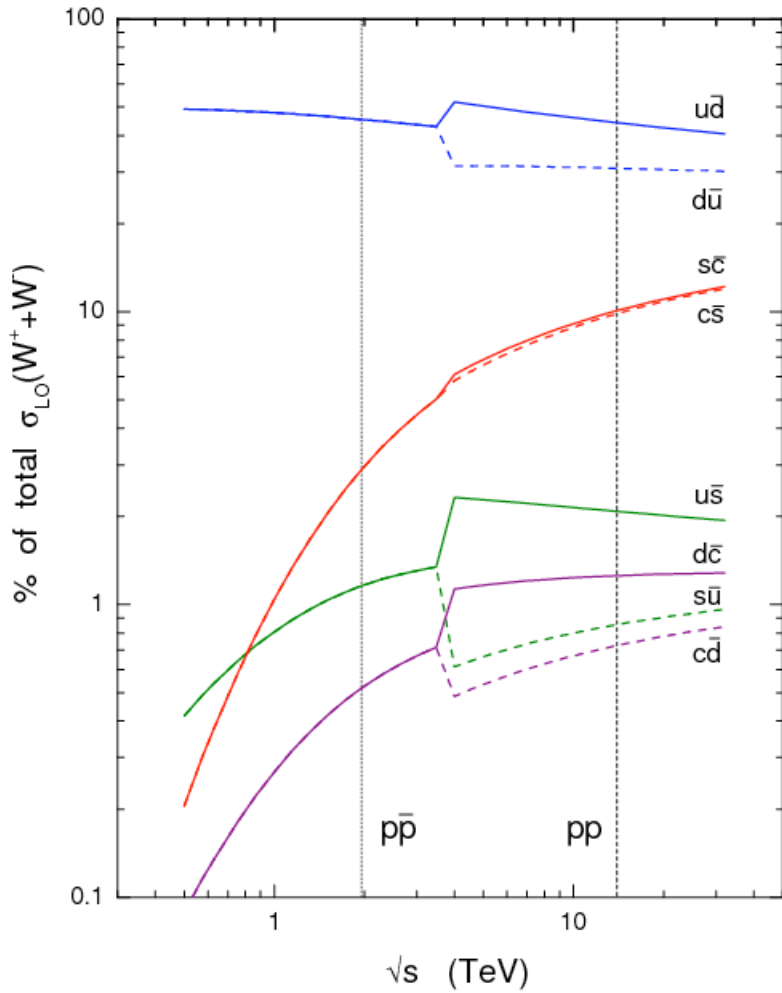
- the standard candles:
 central $\sigma(W,Z,tt,jets)$ as a probe and test of pdfs in the $x \sim 10^{-2\pm 1}$, $Q^2 \sim 10^{4-6} \text{ GeV}^2$ range where most New Physics is expected (H, SUSY,)
 - ongoing studies of uncertainties and correlations
- forward production of (relatively) low-mass states (e.g. $\gamma^*, W, Z, \text{dijets}$) to access partons at $x \ll 1$ (and $x \sim 1$)



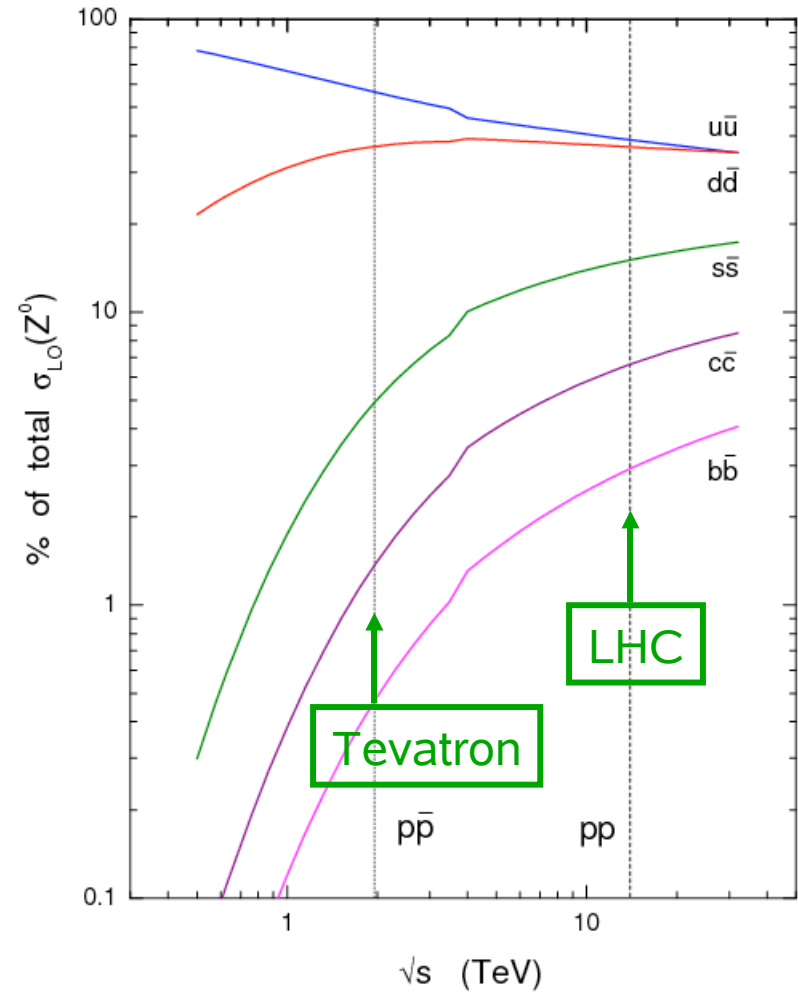
$\sigma(W,Z)$ @ LHC

- in understanding differences between $\sigma(W,Z)$ predictions from different pdf sets (due to the pdfs, *not* choice of pQCD order, e/w parameters, etc) a number of factors are important, particularly
 - the rate of evolution from the Q^2 of the fitted DIS data, to $Q^2 \sim 10^4 \text{ GeV}^2$ (driven by α_s , gluon)
 - the mix of quark flavours: F_2 and $\sigma(W,Z)$ probe *different* combinations of u,d,s,c,b
- precise measurement of cross section *ratios* at LHC (e.g. $\sigma(W^+)/\sigma(W^-)$, $\sigma(W^\pm)/\sigma(Z)$) will allow these subtle effects to be explored further

flavour decomposition of W cross sections

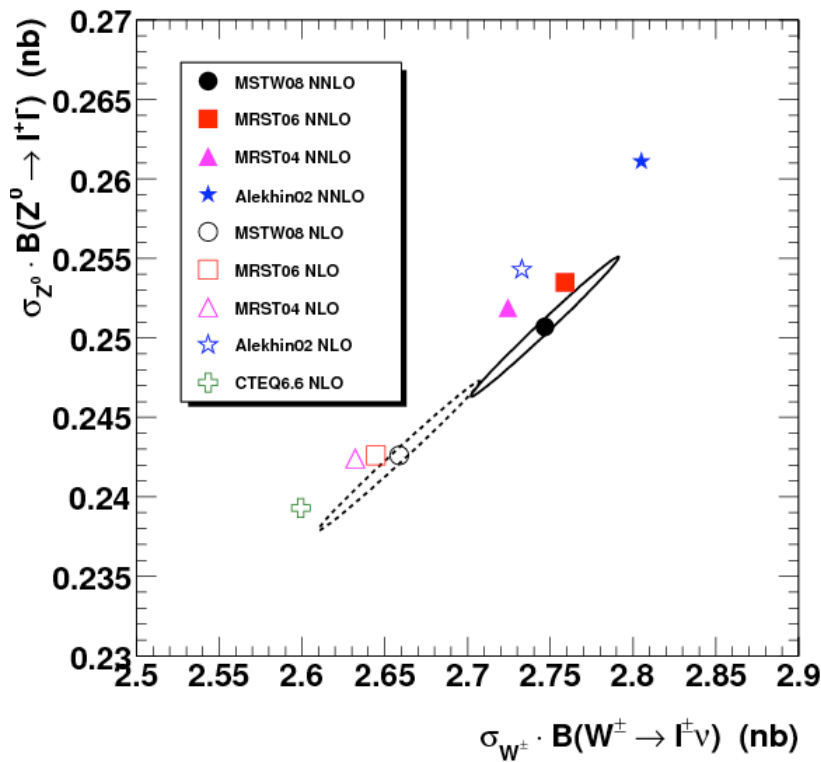


flavour decomposition of Z^0 cross sections

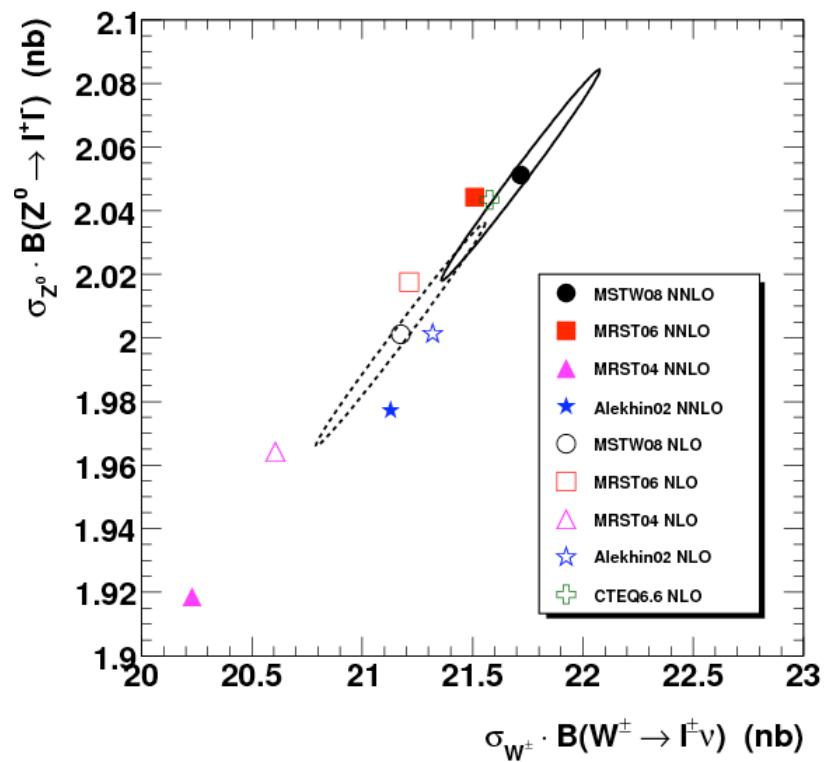


impact on $\sigma(W,Z)$ @ LHC

W and Z total cross sections at the Tevatron



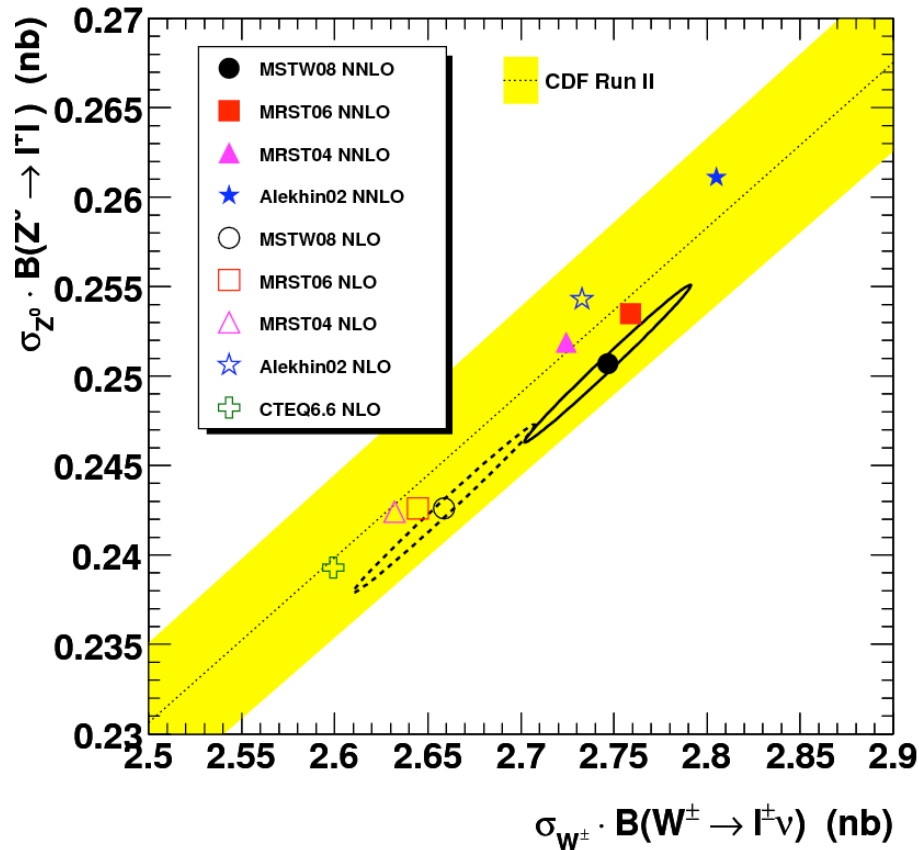
W and Z total cross sections at the LHC



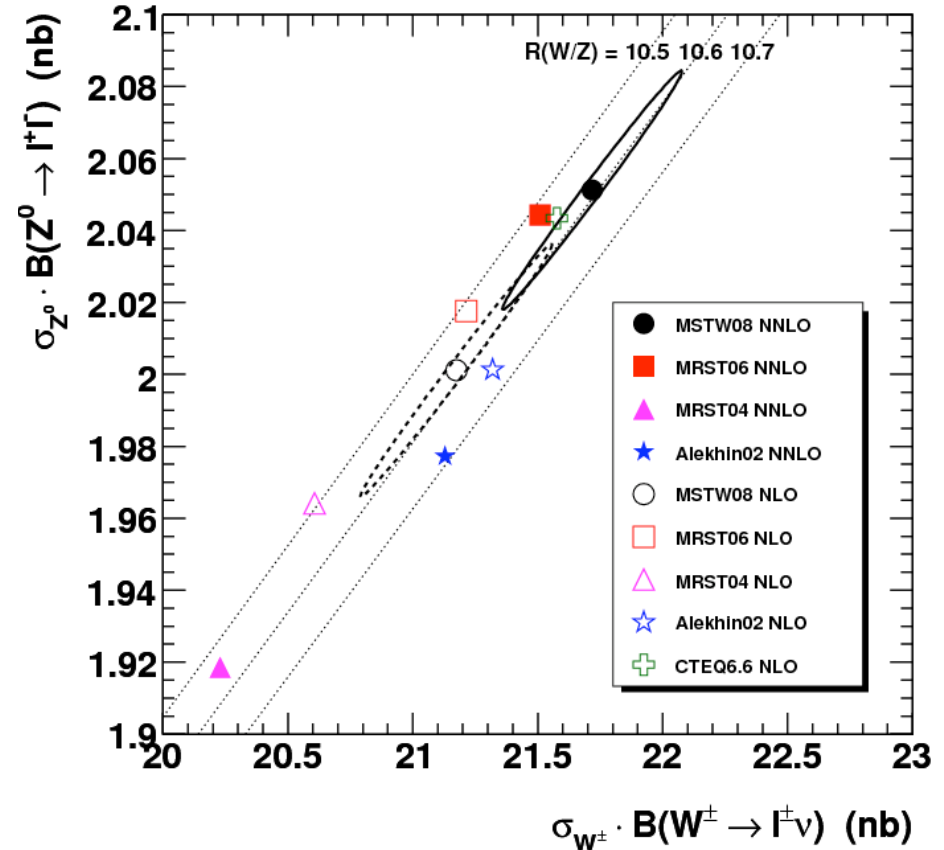
- MRST/MSTW NNLO: 2008 ~ 2006 > 2004 mainly due to changes in treatment of *charm*
- CTEQ: 6.6 ~ 6.5 > 6.1 due to changes in treatment of *s,c,b*
- NLO: CTEQ6.6 2% higher than MSTW 2008 at LHC, because of slight differences in quark (u,d,s,c) pdfs, difference within quoted uncertainty

R(W/Z)= $\sigma(W)/\sigma(Z)$ @ Tevatron & LHC

W and Z total cross sections at the Tevatron



W and Z total cross sections at the LHC



CDF 2007: $R = 10.84 \pm 0.15$ (stat) ± 0.14 (sys)

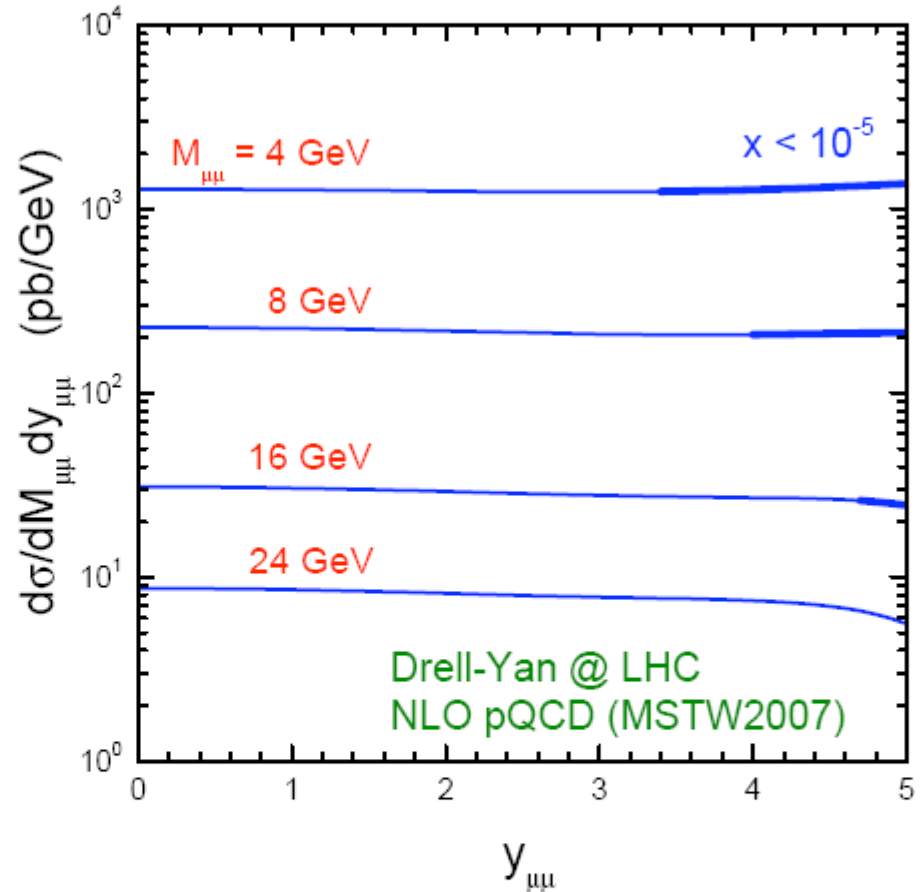
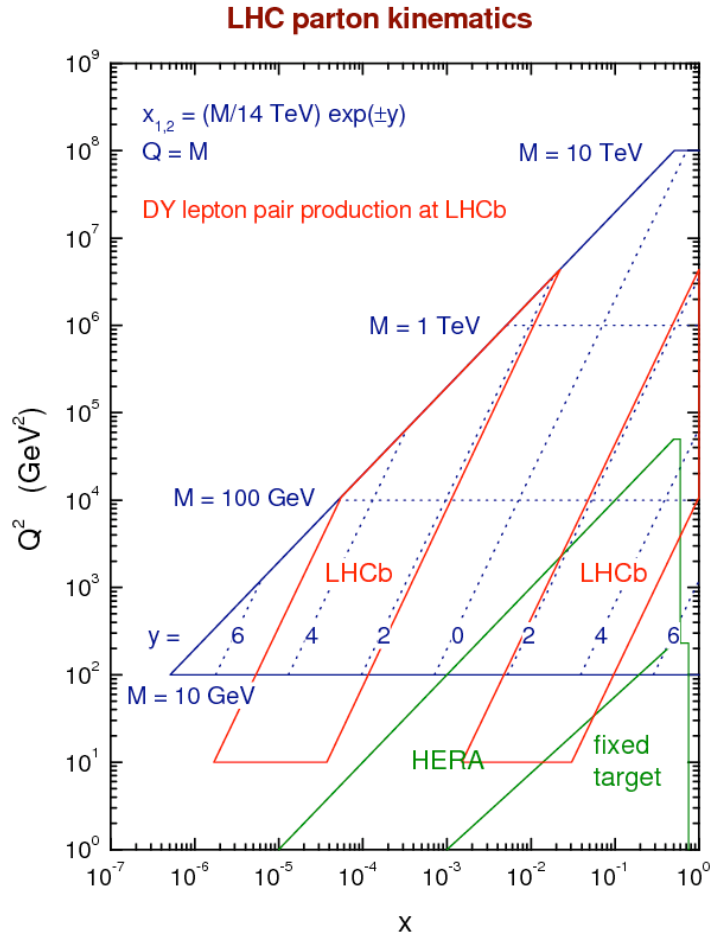
predictions for $\sigma(W,Z)$ @ LHC (Tevatron)

	$B_{lv} \cdot \sigma_W$ (nb)	$B_{ll} \cdot \sigma_Z$ (nb)
MSTW 2008 NLO	21.17 (2.659)	2.001 (0.2426)
MSTW 2008 NNLO	21.72 (2.747)	2.051 (0.2507)

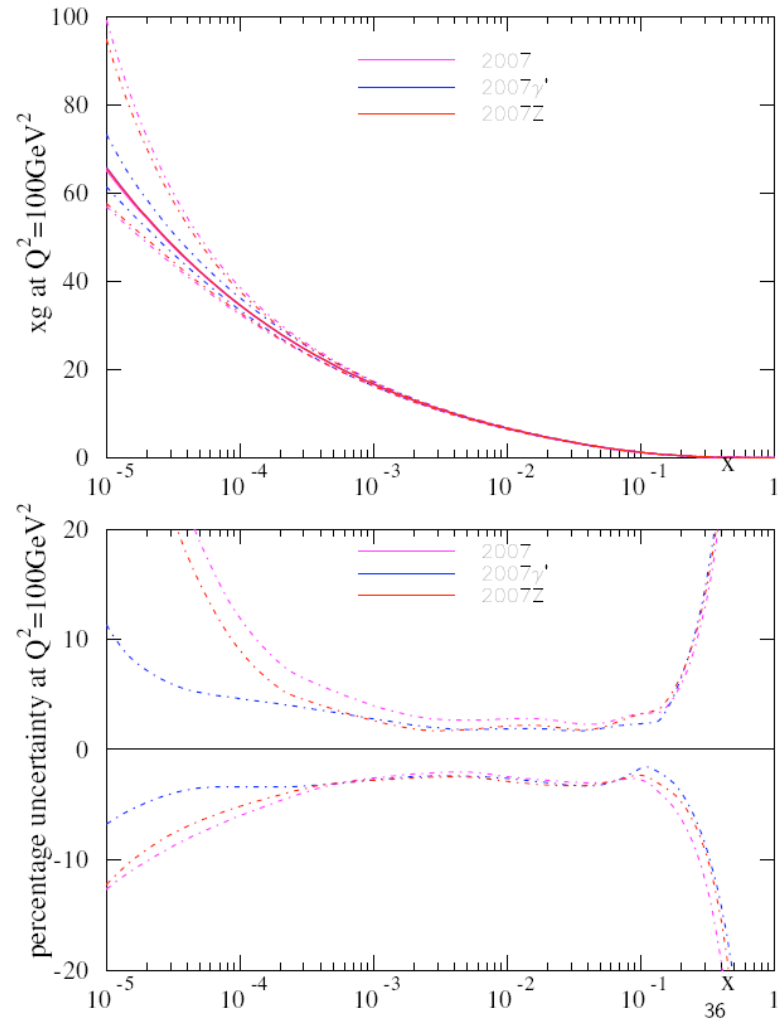
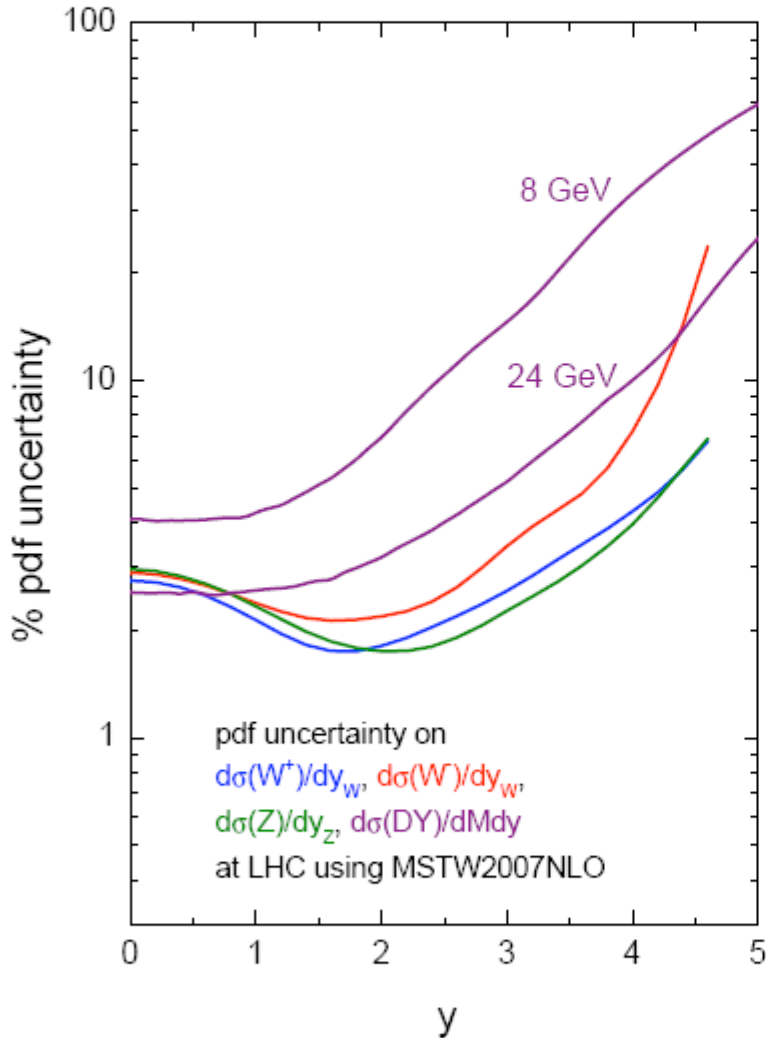
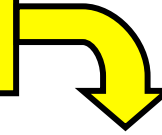
MRST 2006 NLO	21.21 (2.645)	2.018 (0.2426)
MRST 2006 NNLO	21.51 (2.759)	2.044 (0.2535)
MRST 2004 NLO	20.61 (2.632)	1.964 (0.2424)
MRST 2004 NNLO	20.23 (2.724)	1.917 (0.2519)
CTEQ6.6 NLO	21.58 (2.599)	2.043 (0.2393)
Alekhin 2002 NLO	21.32 (2.733)	2.001 (0.2543)
Alekhin 2002 NNLO	21.13 (2.805)	1.977 (0.2611)

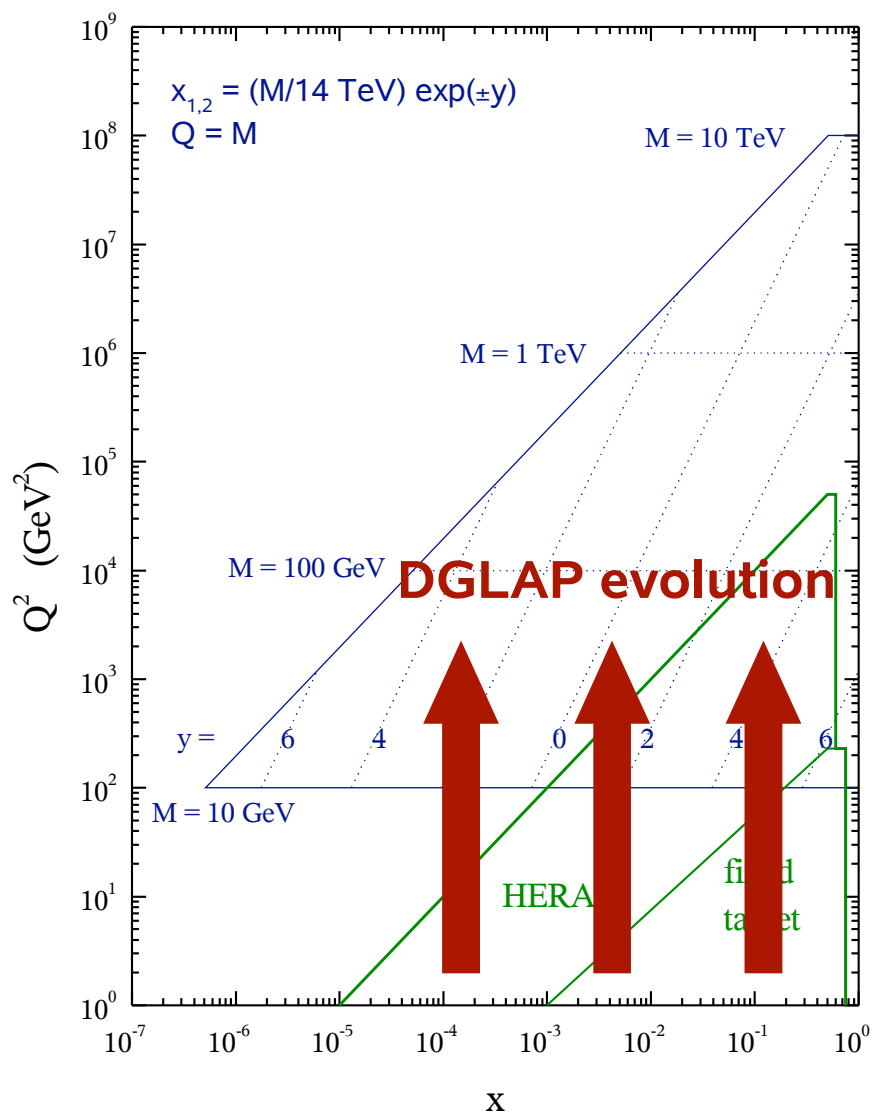
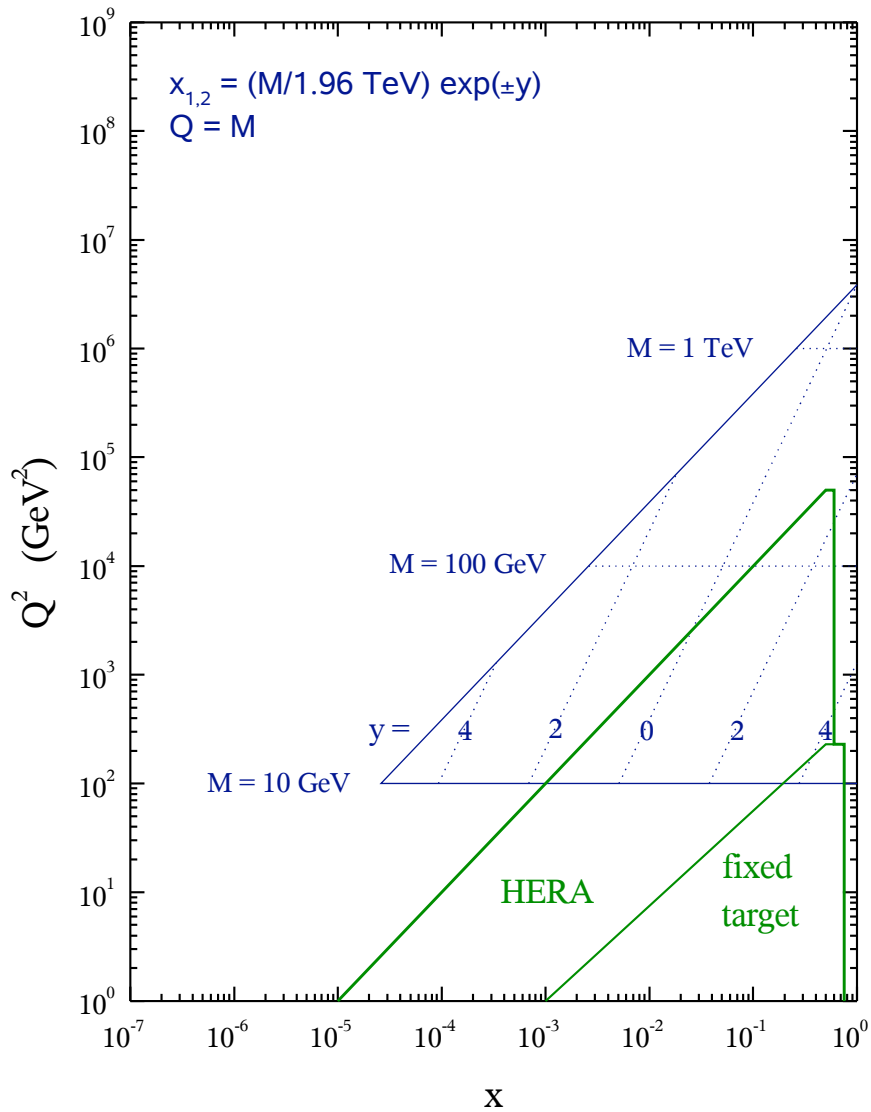
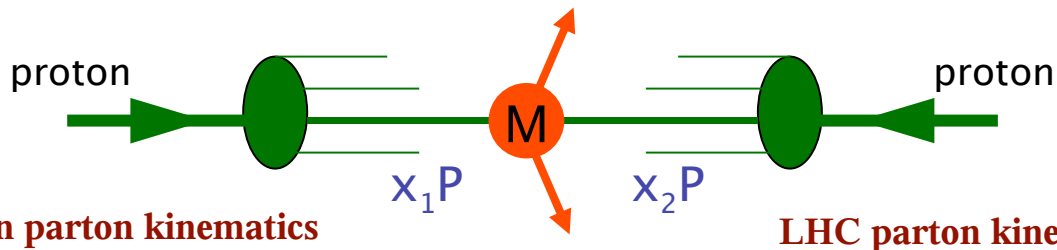
MSTW

→ detect forward, low p_T muons from $q\bar{q} \rightarrow \mu^+ \mu^-$



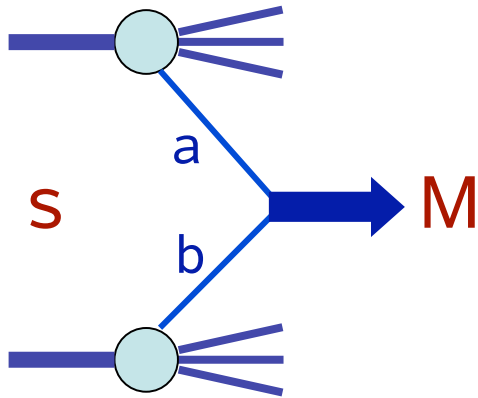
Impact of 1 fb^{-1} LHCb data for forward Z and γ^* ($M = 14 \text{ GeV}$) production on the gluon distribution uncertainty





parton luminosity functions

- a quick and easy way to assess the mass and collider energy dependence of production cross sections



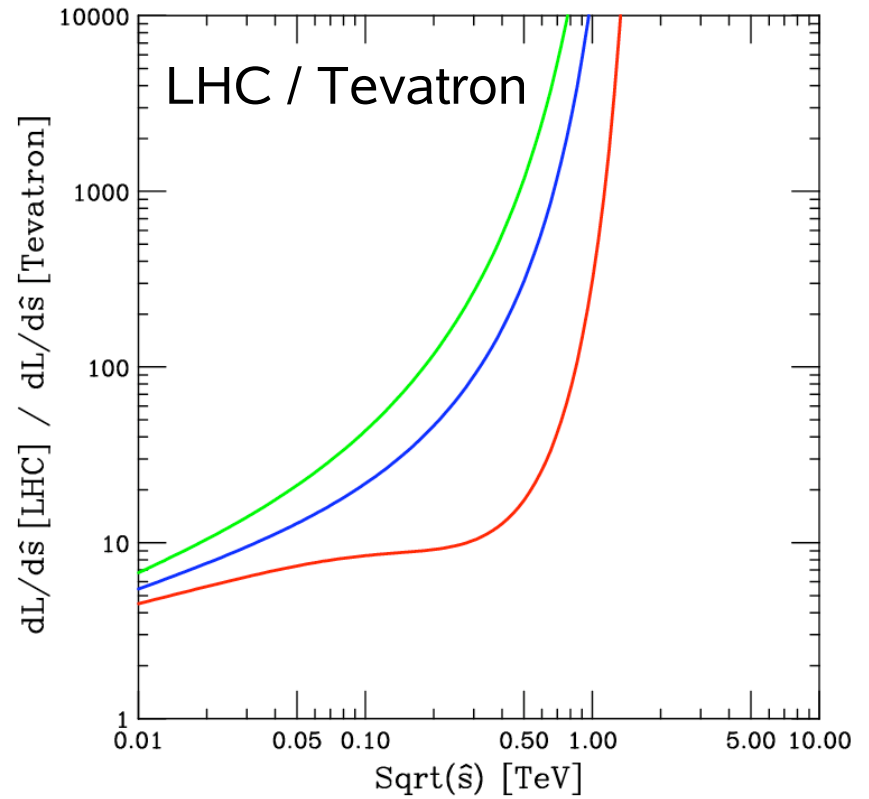
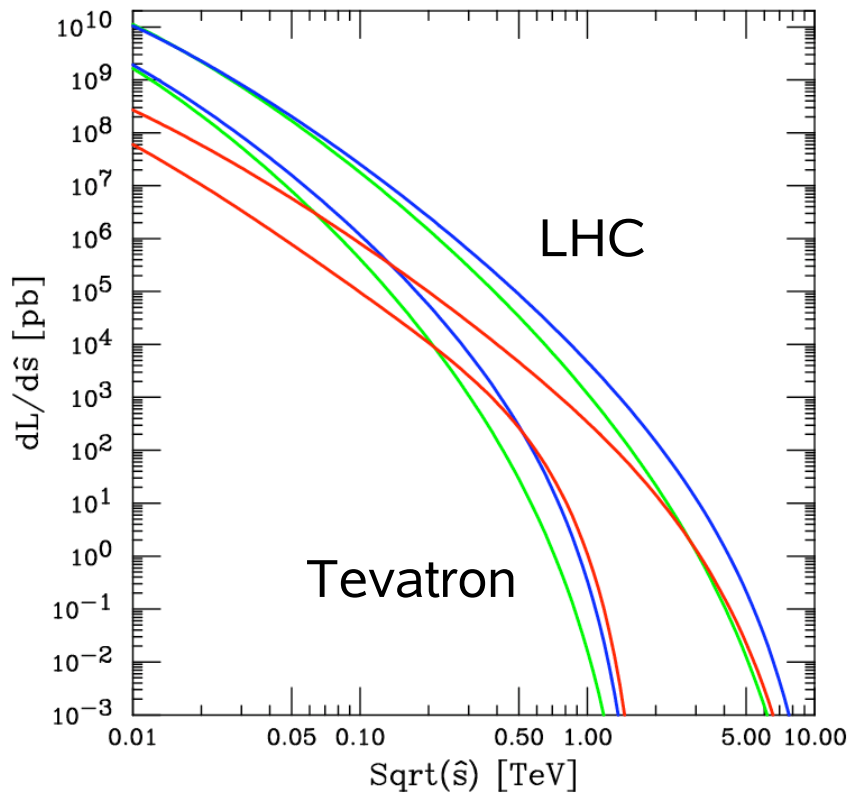
$$\hat{\sigma}_{ab \rightarrow X} = C_X \delta(\hat{s} - M^2)$$

$$\sigma_X = \int_0^1 dx_a dx_b f_a(x_a, M^2) f_b(x_b, M^2) C_X \delta(x_a x_b - \tau)$$

$$\equiv C_X \left[\frac{1}{s} \frac{\partial \mathcal{L}_{ab}}{\partial \tau} \right] \quad (\tau = M^2/s)$$

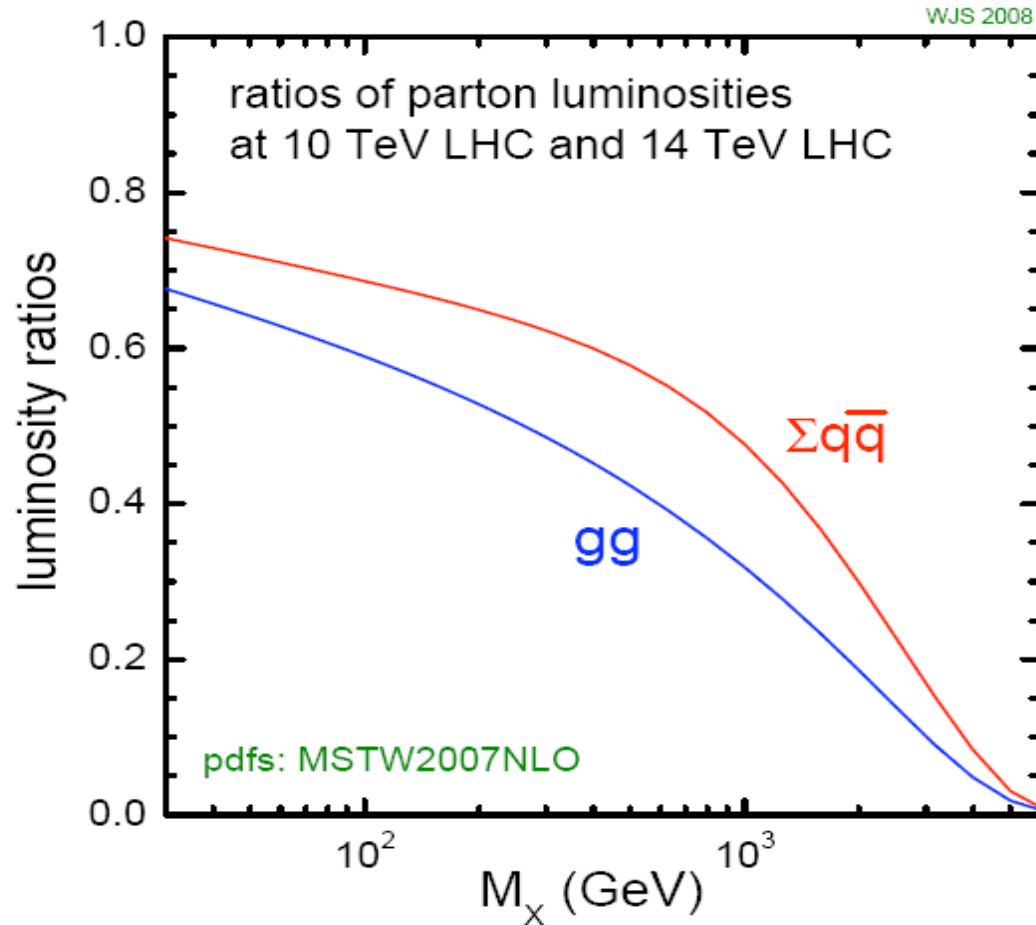
$$\frac{\partial \mathcal{L}_{ab}}{\partial \tau} = \int_0^1 dx_a dx_b f_a(x_a, M^2) f_b(x_b, M^2) \delta(x_a x_b - \tau)$$

- i.e. all the mass and energy dependence is contained in the X-independent parton luminosity function in []
- useful combinations are $ab = gg, \sum_q q\bar{q}, \dots$
- and also useful for assessing the uncertainty on cross sections due to uncertainties in the pdfs (see later)



- = gg
- = $\sum_i (gq_i + g\bar{q}_i + q_i g + \bar{q}_i g)$
- = $\sum_i (q_i \bar{q}_i + \bar{q}_i q_i)$

LHC at 10 TeV




future hadron colliders: energy vs luminosity?

recall parton-parton luminosity:

$$\frac{\partial \mathcal{L}_{ab}}{\partial \tau} = \int_{\tau}^1 \frac{dx}{x} f_a(x, Q^2) f_b(\tau/x, Q^2)$$

so that

$$\sigma_X \propto \frac{1}{s} \frac{\partial \mathcal{L}_{ab}}{\partial \tau}$$


with $\tau = M_X^2/s$

for $M_X > O(1 \text{ TeV})$, energy $\times 3$ is better than luminosity $\times 10$ (everything else assumed equal!)

