Parton Distribution Functions and their applications

Pavel Nadolsky (slides)
Southern Methodist University (Dallas, TX, USA)
C.-P. Yuan (presentation)
Michigan State University (E. Lansing, MI, USA)

Lecture 1
July 2014

Objectives of these lectures

- introduce the basic methods for determination of PDFs from hadronic scattering data
- convey the richness of ideas encountered in the PDF analysis – contributed by diverse branches of theory, experiment, and mathematics
- discuss how our knowledge of PDFs affects practical applications

Selection of the topics is far from complete – complementary material can be found in excellent lectures on PDFs by A. Cooper-Sarkar, W. Giele, J. Owens, A. Martin, W. Melnitchouk, D. Stump, A. Accardi at recent CTEQ schools

Key ideas

Parton distribution functions $f_{a/p}(x, Q)$...

... arise as nonperturbative functions in QCD factorization

... describe probabilities for finding partons inside parent hadrons

... are universal – independent of the hard-scattering process

... cannot be computed systematically

... obey perturbative evolution (DGLAP) equations

... are determined from hadronic experiments

Parton distribution functions $f_{a/p}(x,Q)$...

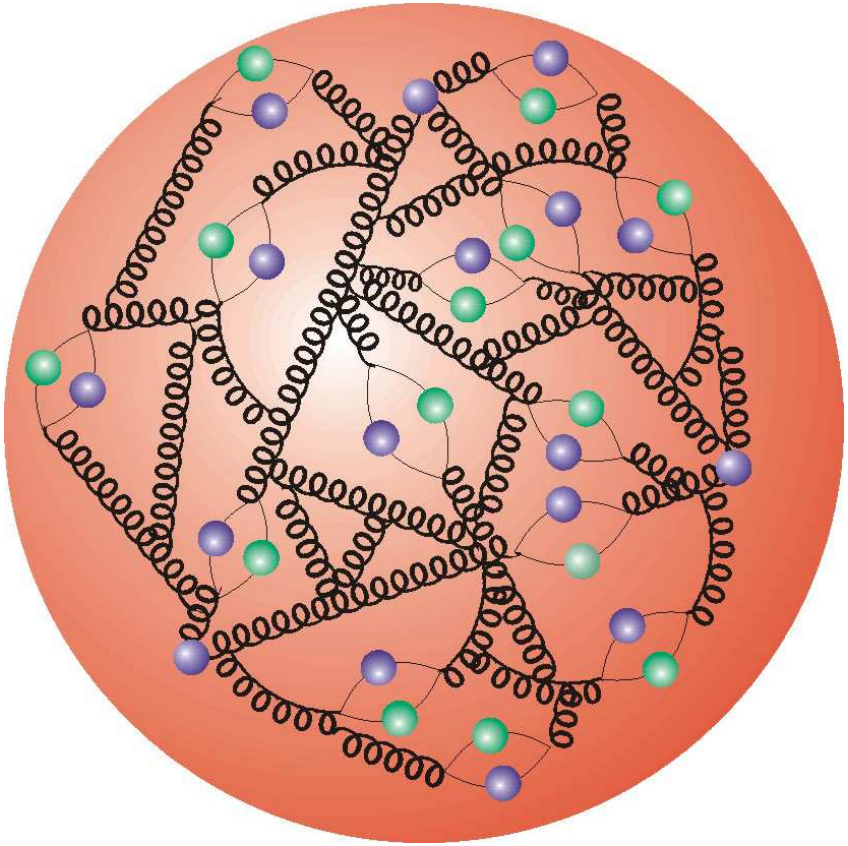
We will focus on $f_{a/p}(x,Q)$, unpolarized PDFs in the proton (for unpolarized processes at the LHC, etc.) Much of our discussion will apply to polarized PDFs and PDFs in the nucleus.

Basic definitions

- **Partons** are weakly bound constituents of hadrons with small typical size
  $(r \ll r_{nucleon} \approx 1 \text{ fm})$

  (Feynman; Bjorken, Paschos - 1969)

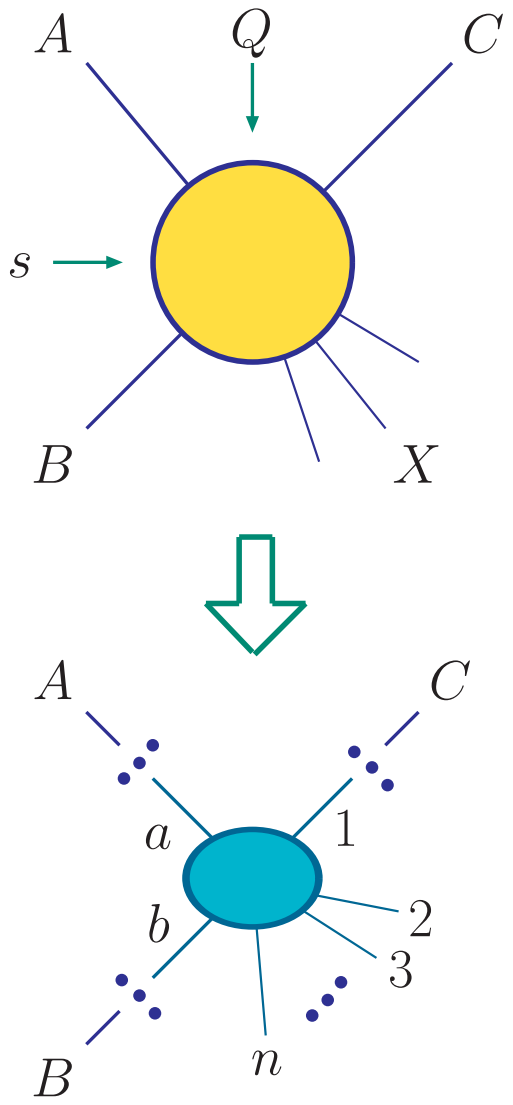
  - assumed to be pointlike at present



Basic definitions

- Partons are most easily detected in **inclusive** hadronic scattering $A + B \to C + X$ at large collision energy $\sqrt{s} \gg 1$ GeV, with typical energy transfer Q of order $\sqrt{s}$

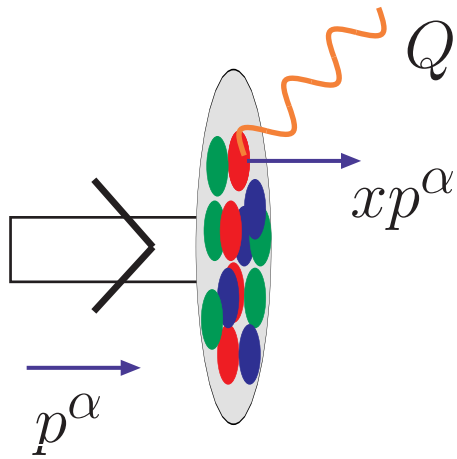
- Such scattering is dominated by **rare independent collisions** $a + b \to 1 + 2 + ... + n$ of a parton a from A on a parton b from B, proceeding through **perturbative** QCD and electroweak interactions



Basic definitions

- In the simplest (leading-order) interpretation, the PDF $f_{a/p}(x,Q)$ is a probability for finding a parton a with 4-momentum xp^α in a proton with 4-momentum p^α

- $f_{a/p}(x,Q)$ depends on **nonperturbative** QCD interactions

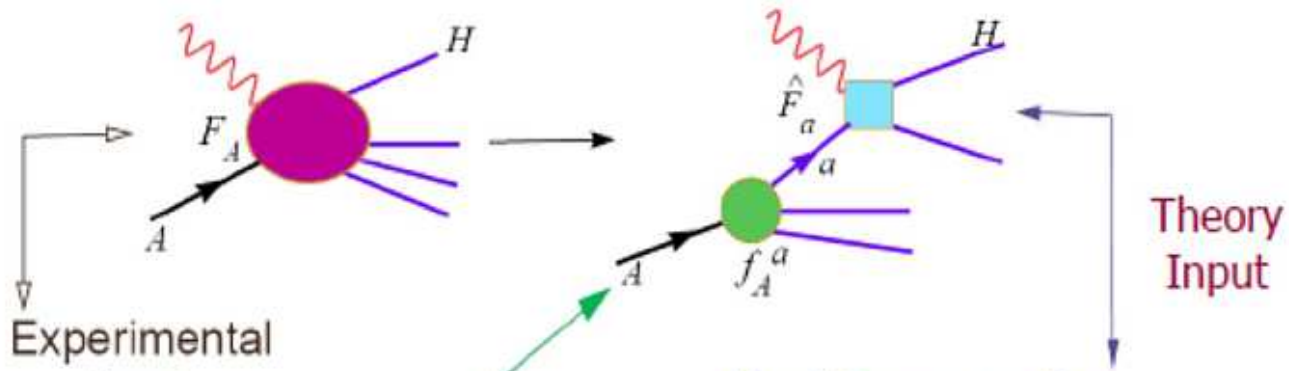


QCD factorization in DIS

Deep Inelastic Scattering process

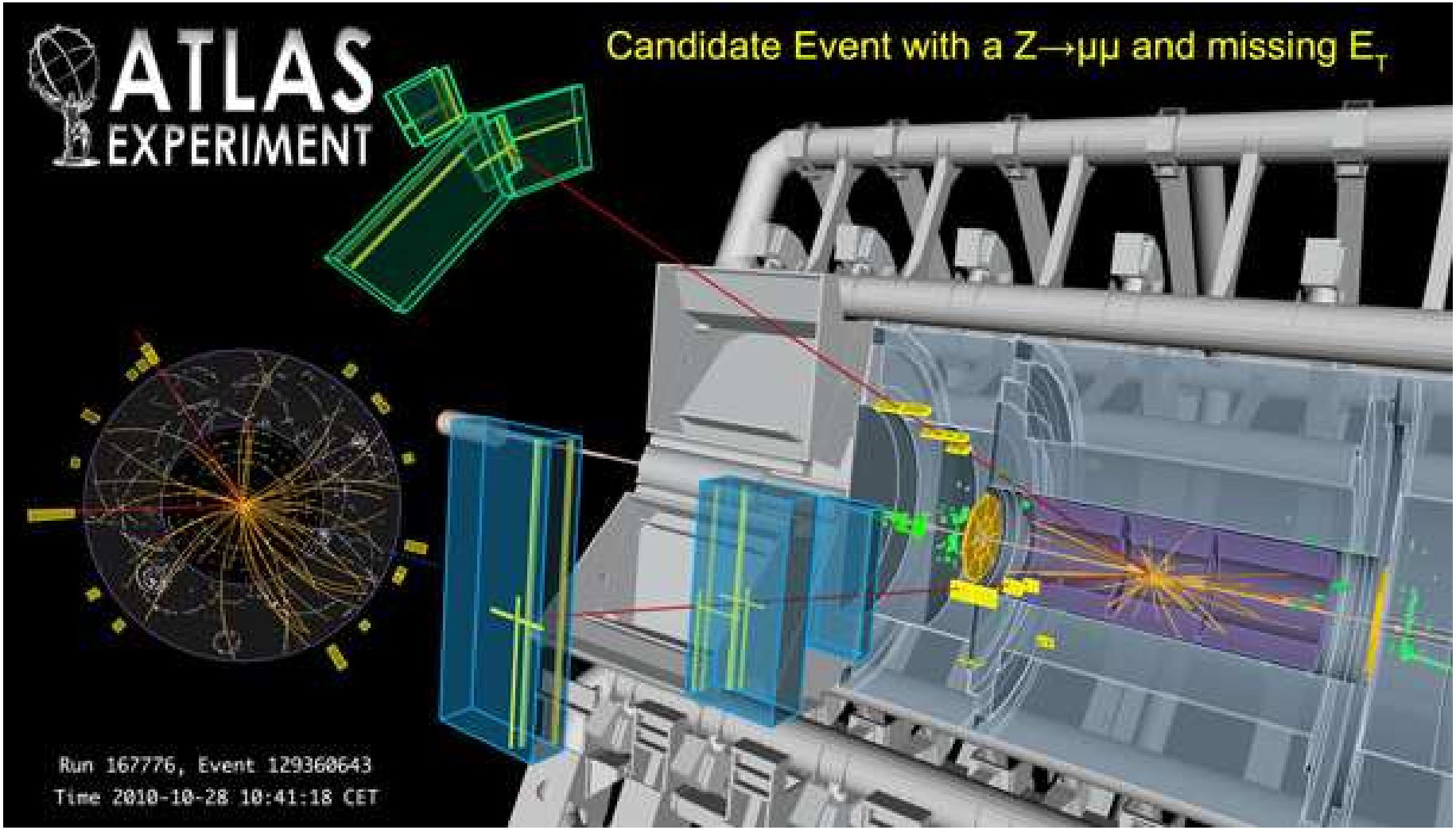
Master Equation for QCD Parton Model
– the Factorization Theorem

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



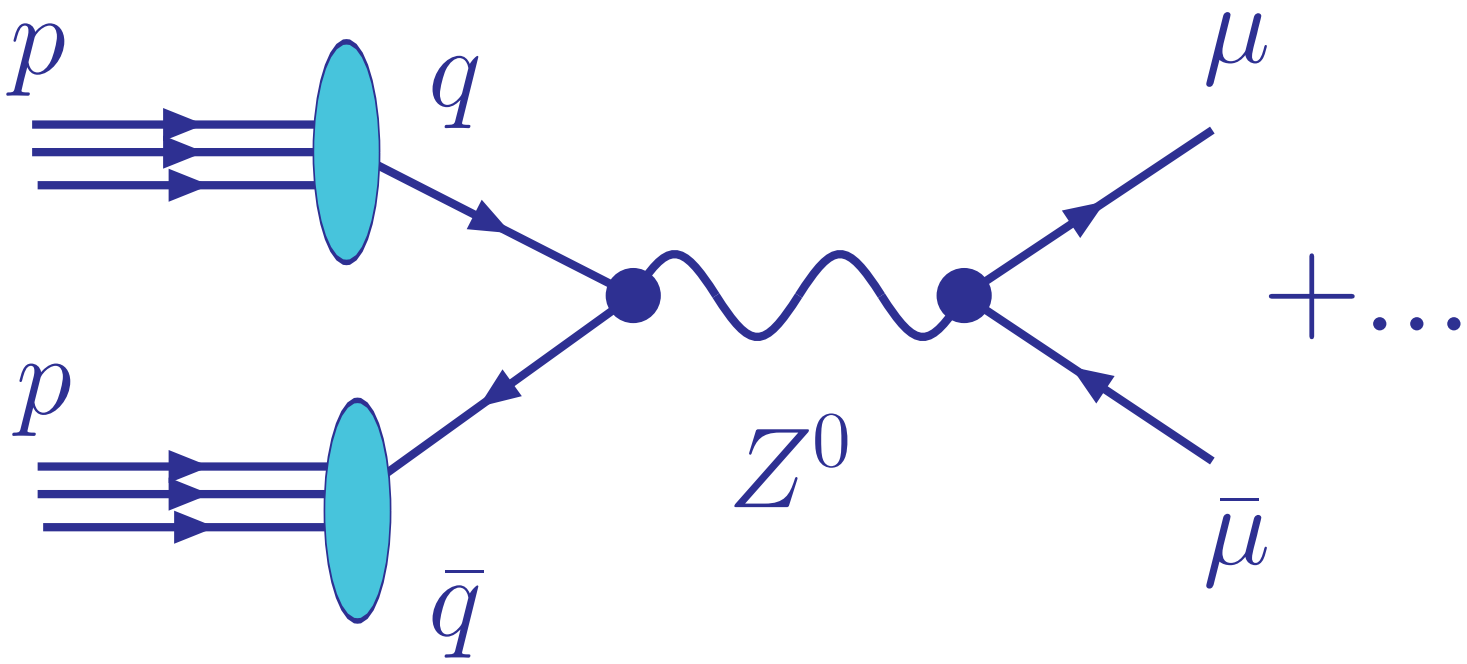
universal Parton Dist. Fn.
Non-Perturbative Parametrization at Q_0
DGLAP Evolution to Q
extracted by global analysis

QCD factorization in the Drell-Yan process



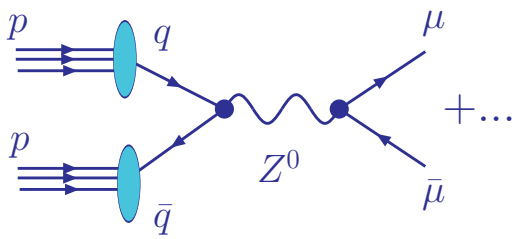
Drell-Yan process $pp \to (Z^0 \to \ell\bar{\ell})X$ at the LHC ($\ell\bar{\ell} = e\bar{e}$ or $\mu\bar{\mu}$)

QCD factorization in the Drell-Yan process



$pp \to (Z^0 \to \mu\bar{\mu})X$: Feynman diagram at the leading order in QCD

PDFs and QCD factorization

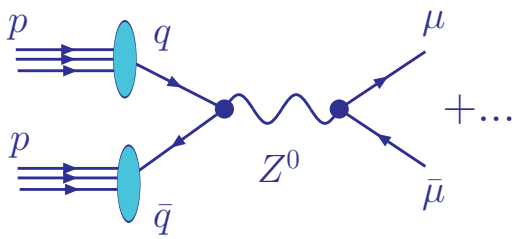


According to QCD factorization theorems, typical cross sections (e.g., for $p(k_1)p(k_2) \to \left[Z(q) \to \ell(k_3)\bar{\ell}(k_4)\right] X$) take the form

$$\sigma_{pp\to\ell\bar{\ell}X} = \sum_{a,b=q,\bar{q},g} \int_0^1 d\xi_1 \int_0^1 d\xi_2\, \widehat{\sigma}_{ab\to Z\to\ell\bar{\ell}}\left(\frac{x_1}{\xi_1}, \frac{x_2}{\xi_2}; \frac{Q}{\mu}\right) f_{a/p}(\xi_1,\mu) f_{b/p}(\xi_2,\mu) + \mathcal{O}\left(\Lambda^2_{QCD}/Q^2\right)$$

- $\widehat{\sigma}_{ab\to Z\to\ell\bar{\ell}}$ is the **hard-scattering cross section**
- $f_{a/p}(\xi,\mu)$ are the **PDFs**
- $Q^2 = (k_3+k_4)^2$, $x_{1,2} = (Q/\sqrt{s})\, e^{\pm y_V}$ — measurable quantities
- ξ_1, ξ_2 are partonic momentum fractions (integrated over)
- μ is a factorization scale (=renormalization scale from now on)

PDFs and QCD factorization

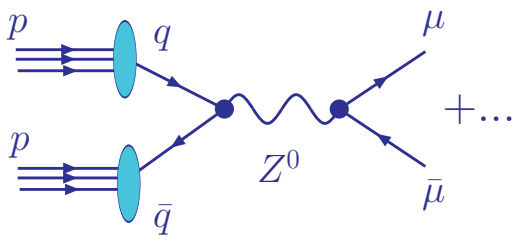


According to QCD factorization theorems, typical cross sections (e.g., for $p(k_1)p(k_2) \to \left[Z(q) \to \ell(k_3)\bar{\ell}(k_4)\right] X$) take the form

$$\sigma_{pp\to\ell\bar{\ell}X} = \sum_{a,b=q,\bar{q},g} \int_0^1 d\xi_1 \int_0^1 d\xi_2\, \widehat{\sigma}_{ab\to Z\to\ell\bar{\ell}}\left(\frac{x_1}{\xi_1}, \frac{x_2}{\xi_2}; \frac{Q}{\mu}\right) f_{a/p}(\xi_1,\mu) f_{b/p}(\xi_2,\mu) + \mathcal{O}\left(\Lambda^2_{QCD}/Q^2\right)$$

- μ is naturally set to be of order Q
- Factorization holds up to terms of order Λ^2_{QCD}/Q^2

PDFs and QCD factorization



According to QCD factorization theorems, typical cross sections (e.g., for $p(k_1)p(k_2) \to \left[Z(q) \to \ell(k_3)\bar{\ell}(k_4)\right] X$) take the form

$$\sigma_{pp\to\ell\bar{\ell}X} = \sum_{a,b=q,\bar{q},g} \int_0^1 d\xi_1 \int_0^1 d\xi_2\, \widehat{\sigma}_{ab\to Z\to\ell\bar{\ell}}\left(\frac{x_1}{\xi_1}, \frac{x_2}{\xi_2}; \frac{Q}{\mu}\right) f_{a/p}(\xi_1,\mu) f_{b/p}(\xi_2,\mu) + \mathcal{O}\left(\Lambda^2_{QCD}/Q^2\right)$$

Purpose of this arrangement:

- Subtract large collinear logarithms $\alpha_s^n \ln^k(Q^2/m_q^2)$ from $\widehat{\sigma}$
- Resum them in $f_{a/p}(\xi,\mu)$ to all orders of α_s

Operator definitions for PDFs

To all orders in α_s, PDFs are **defined** as matrix elements of certain correlator functions:

$$f_{q/p}(x,\mu) = \int_{-\infty}^{\infty} \frac{dy^-}{4\pi} e^{iy^- p^+} \langle p \left| \overline{\psi}_q(0,y^-,\vec{0}_T) F(y^-,0) \gamma^+ \psi_q(0,0,\vec{0}_T) \right| p \rangle, \text{ etc.}$$

Several types of definitions, or **factorization schemes** ($\overline{MS}$, DIS, etc.), exist

They all correspond to the probability density for finding a in p at LO; they differ at NLO and beyond

To prove factorization, one must show that $f_{a/p}(x,\mu)$ correctly captures higher-order contributions for the considered observable

This condition can be violated for multi-scale observables (e.g., DIS or Drell-Yan process at $x \sim Q/\sqrt{s} \ll 1$)

Operator definitions for PDFs

To all orders in α_s, PDFs are **defined** as matrix elements of certain correlator functions:

$$f_{q/p}(x,\mu) = \int_{-\infty}^{\infty} \frac{dy^-}{4\pi} e^{iy^- p^+} \langle p \left| \overline{\psi}_q(0,y^-,\vec{0}_T) F(y^-,0) \gamma^+ \psi_q(0,0,\vec{0}_T) \right| p \rangle, \text{ etc.}$$

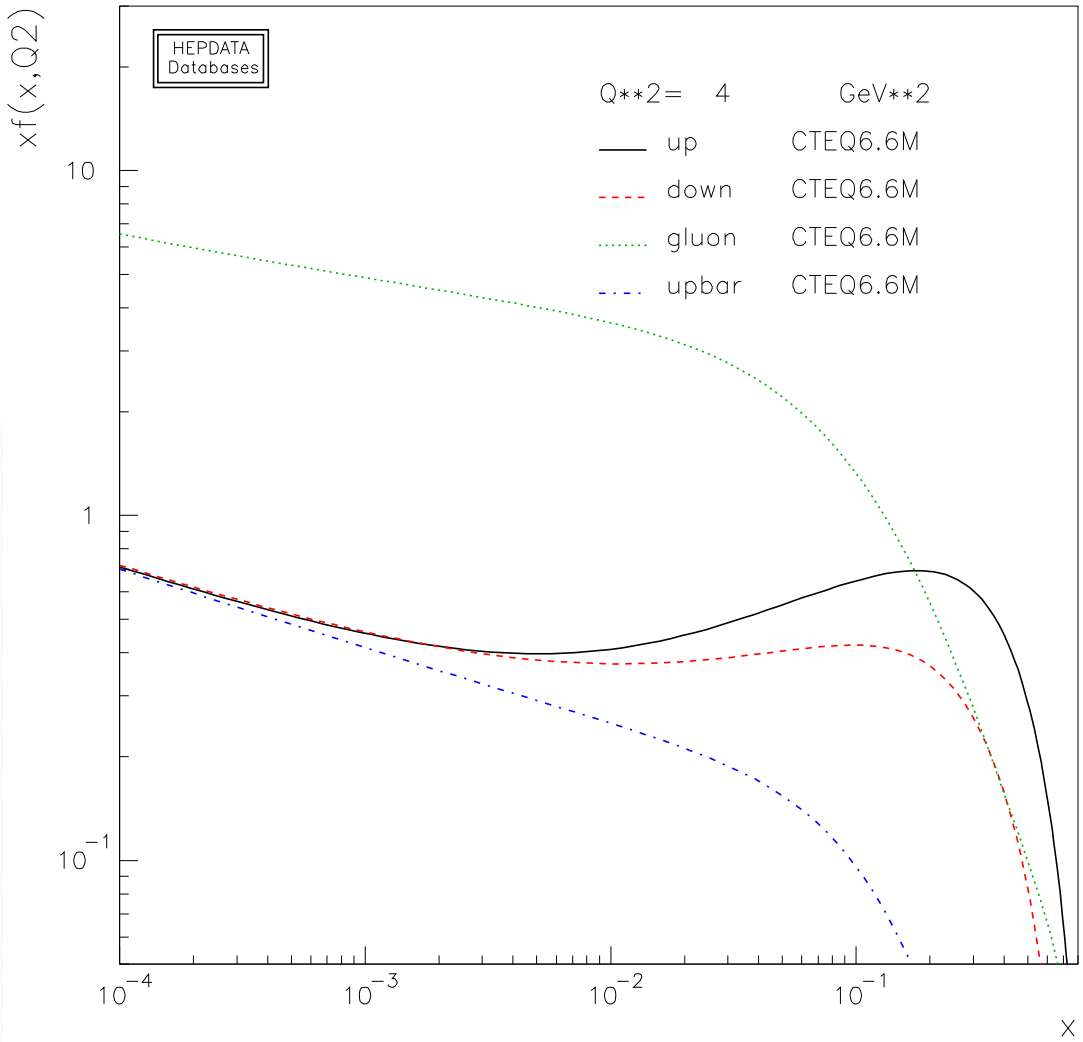
The exact form of $f_{a/p}$ is not known; but its μ dependence is described by **D**okshitzer-**G**ribov-**L**ipatov-**A**ltarelli-**P**arisi (**DGLAP**) equations:

$$\mu \frac{df_{i/p}(x,\mu)}{d\mu} = \sum_{j=g,u,\bar{u},d,\bar{d},....} \int_x^1 \frac{dy}{y} P_{i/j}\left(\frac{x}{y}, \alpha_s(\mu)\right) f_{j/p}(y,\mu)$$

$P_{i/j}$ are probabilities for $j \to ik$ collinear splittings;
are known to order α_s^3 (NNLO):

$$P_{i/j}(x,\alpha_s) = \alpha_s P_{i/j}^{(1)}(x) + \alpha_s^2 P_{i/j}^{(2)}(x) + \alpha_s^3 P_{i/j}^{(3)}(x) + ...$$

Example of DGLAP evolution



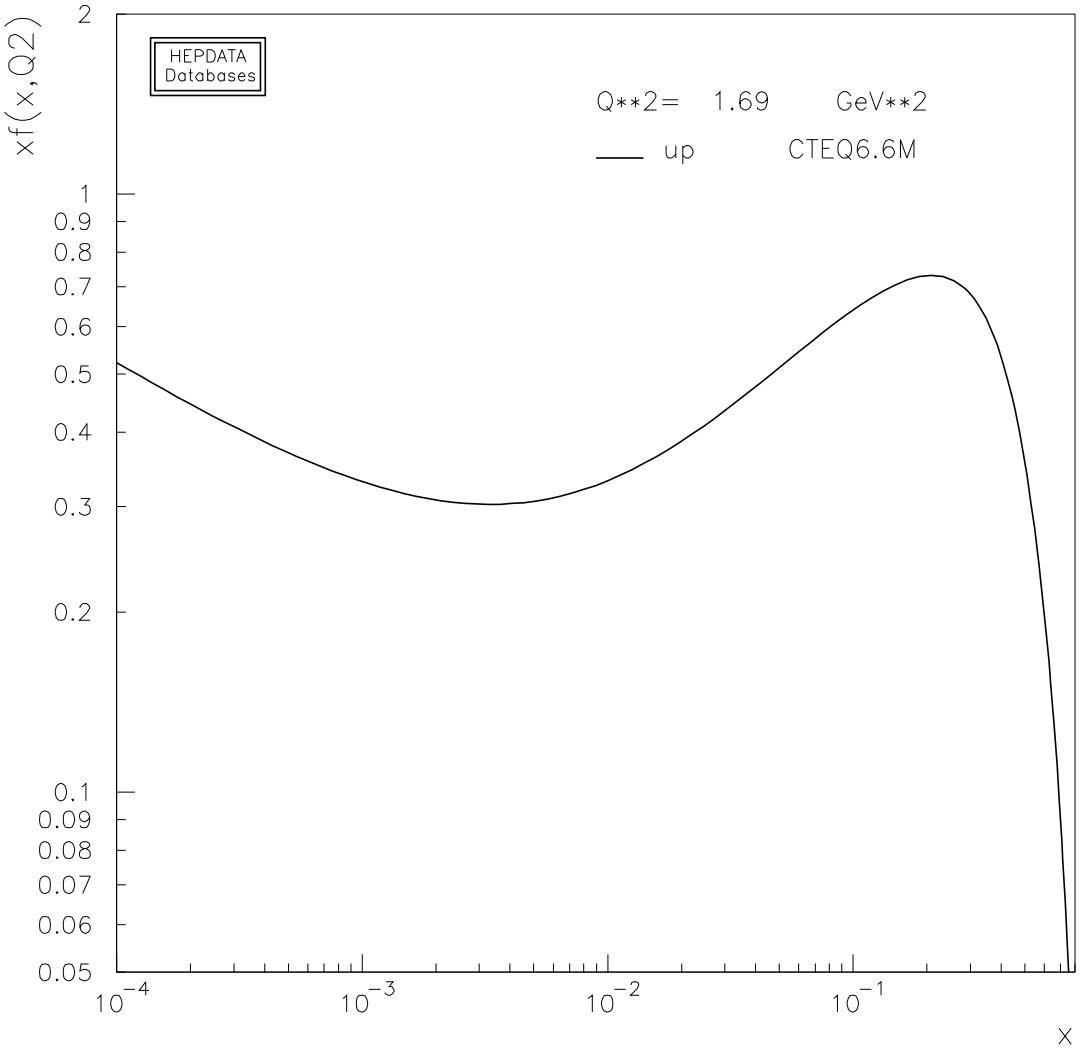
Durham PDF plotter, http://durpdg.dur.ac.uk/hepdata/pdf3.html

Compare μ dependence of u quark PDF and the gluon PDF

The u, d PDFs have a characteristic bump at $x \sim 1/3$ – reminiscent of early valence quark models of the proton structure

The PDFs rise rapidly at $x < 0.1$ as a consequence of perturbative evolution

Example of DGLAP evolution

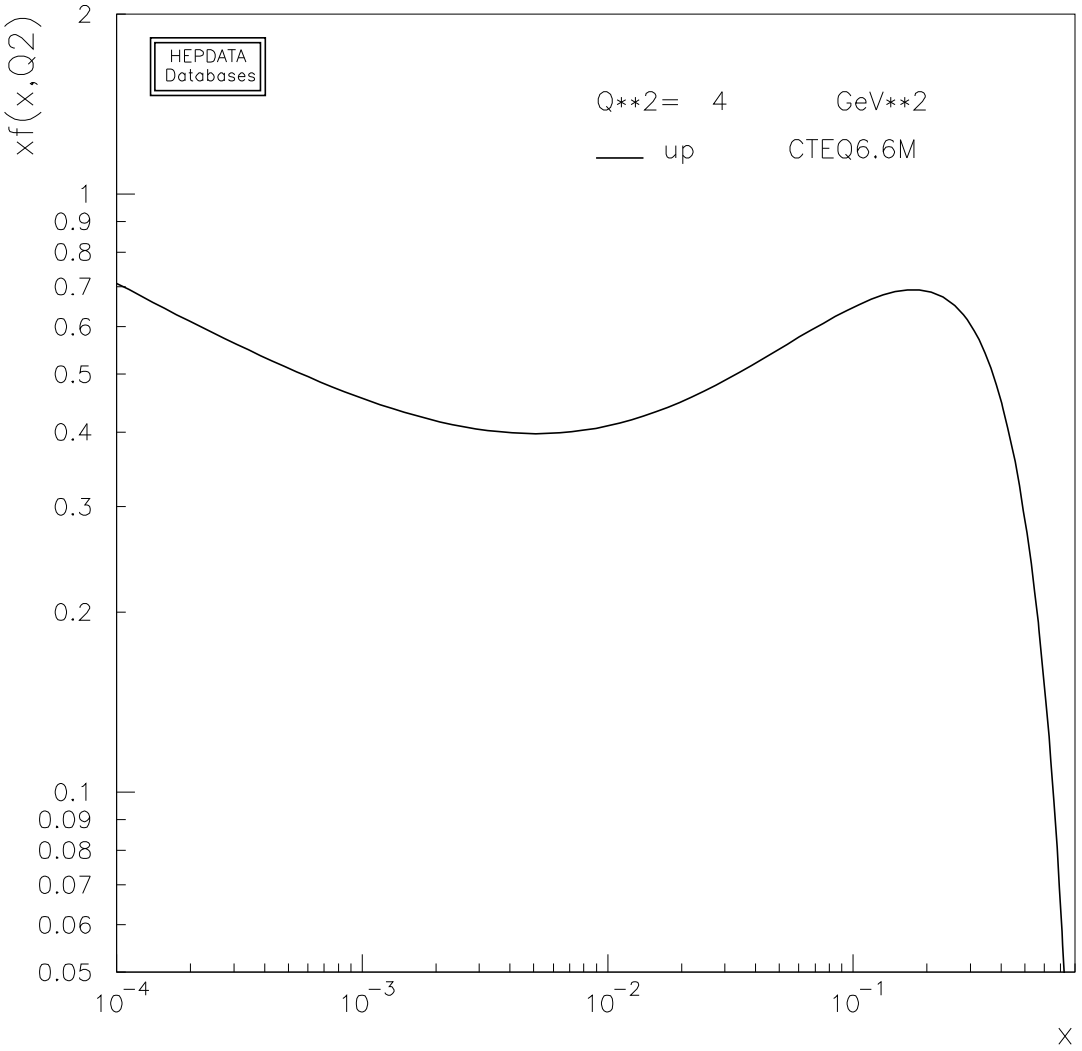


As Q increases, it becomes more likely that a high-x parton loses some momentum through QCD radiation

$\Rightarrow u(x,Q)$ reduces at $x \gtrsim 0.1$, increases at $x \lesssim 0.1$

Durham PDF plotter, http://durpdg.dur.ac.uk/hepdata/pdf3.html

Example of DGLAP evolution

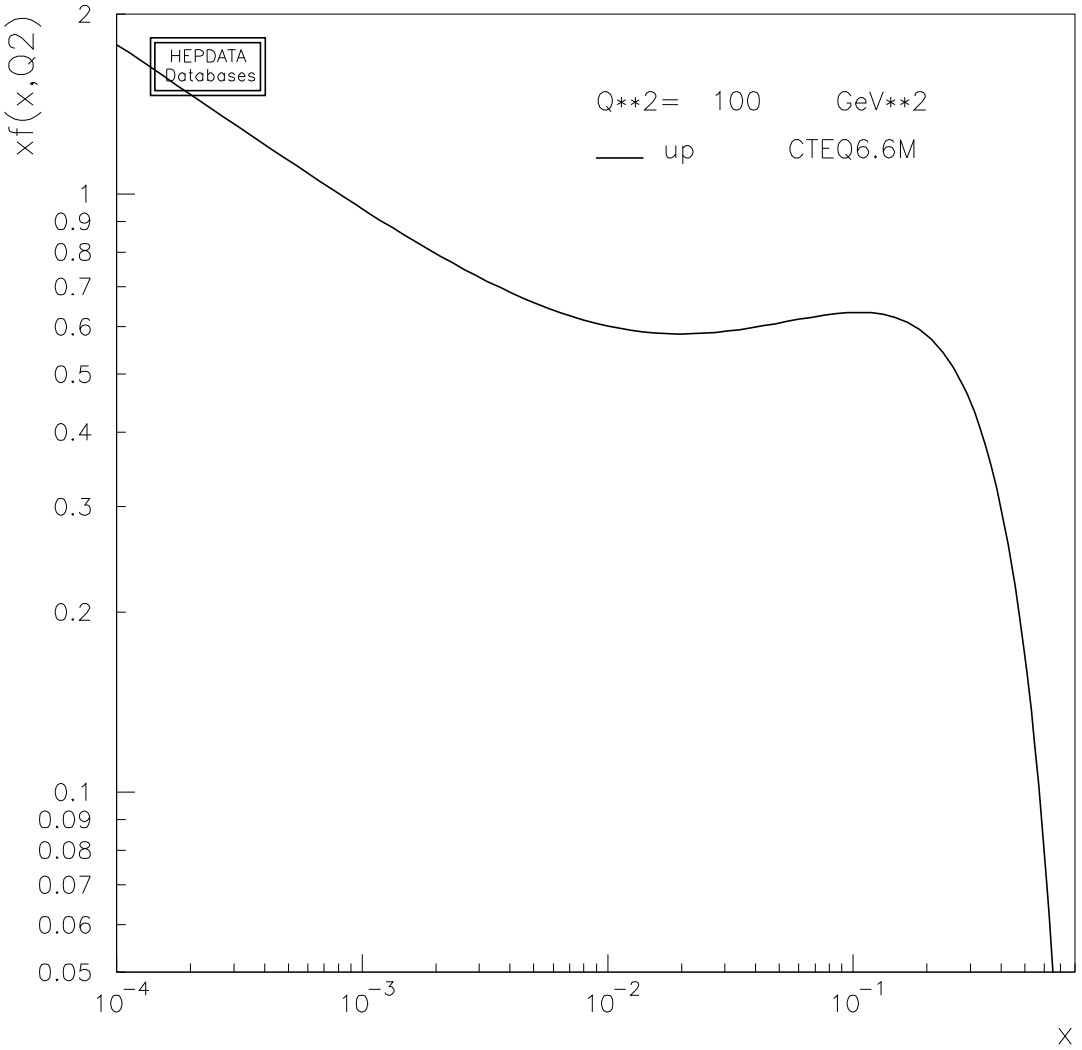


As Q increases, it becomes more likely that a high-x parton loses some momentum through QCD radiation

$\Rightarrow u(x,Q)$ reduces at $x \gtrsim 0.1$, increases at $x \lesssim 0.1$

Durham PDF plotter, http://durpdg.dur.ac.uk/hepdata/pdf3.html

Example of DGLAP evolution

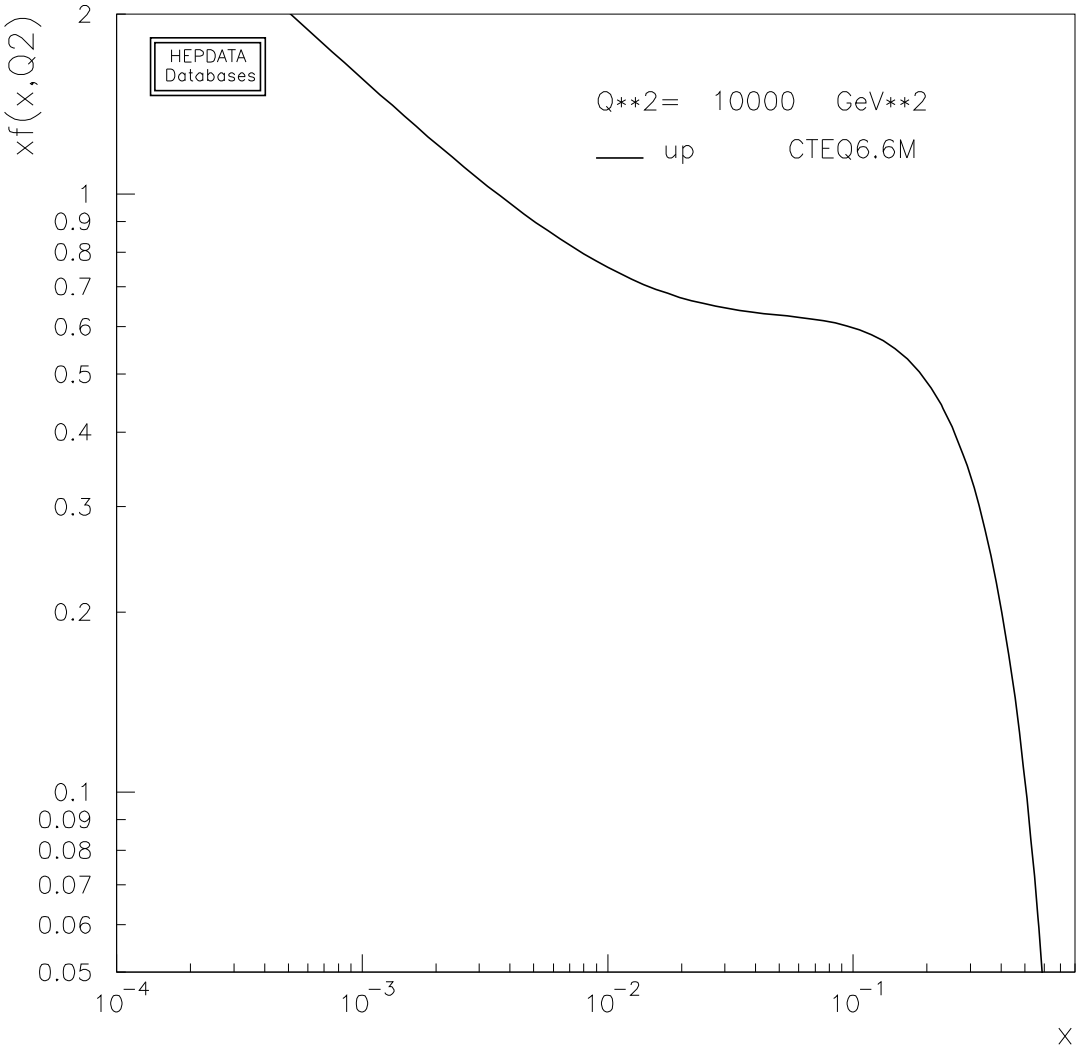


As Q increases, it becomes more likely that a high-x parton loses some momentum through QCD radiation

$\Rightarrow u(x,Q)$ reduces at $x \gtrsim 0.1$, increases at $x \lesssim 0.1$

Durham PDF plotter, http://durpdg.dur.ac.uk/hepdata/pdf3.html

Example of DGLAP evolution

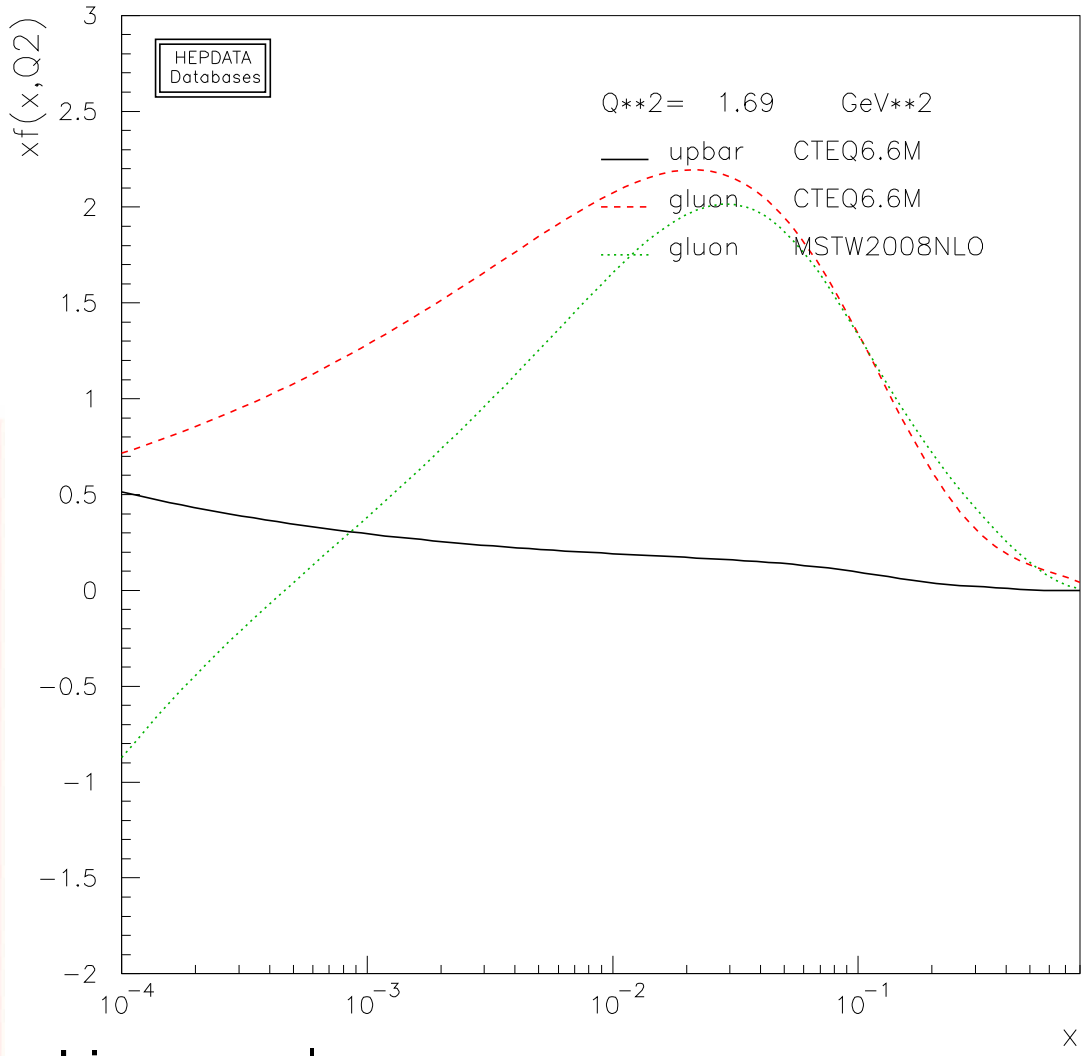


As Q increases, it becomes more likely that a high-x parton loses some momentum through QCD radiation

$\Rightarrow u(x,Q)$ reduces at $x \gtrsim 0.1$, increases at $x \lesssim 0.1$

Durham PDF plotter, http://durpdg.dur.ac.uk/hepdata/pdf3.html

Example of DGLAP evolution: $\bar{u}$ and gluon PDF



Linear scale

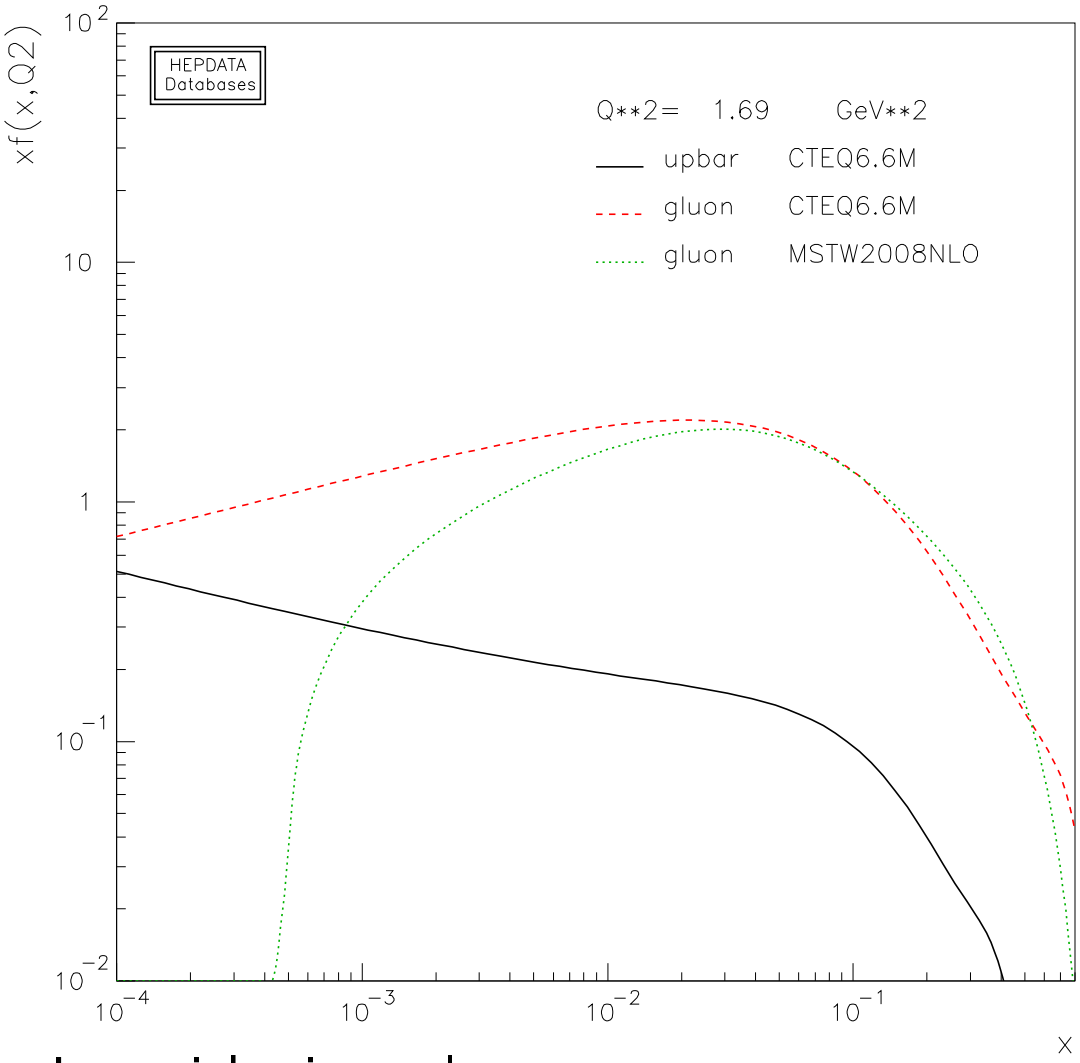
$g(x,Q)$ can become negative at $x < 10^{-2}$, $Q < 2$ GeV

may lead to unphysical predictions

This is an indication that DGLAP factorization experiences difficulties at such small x and Q

Large $\ln^k(1/x)$ in $P_{i/j}(x)$ break PQCD expansion at $x \sim Q/\sqrt{s} \ll 1$

Example of DGLAP evolution: $\bar{u}$ and gluon PDF



Logarithmic scale

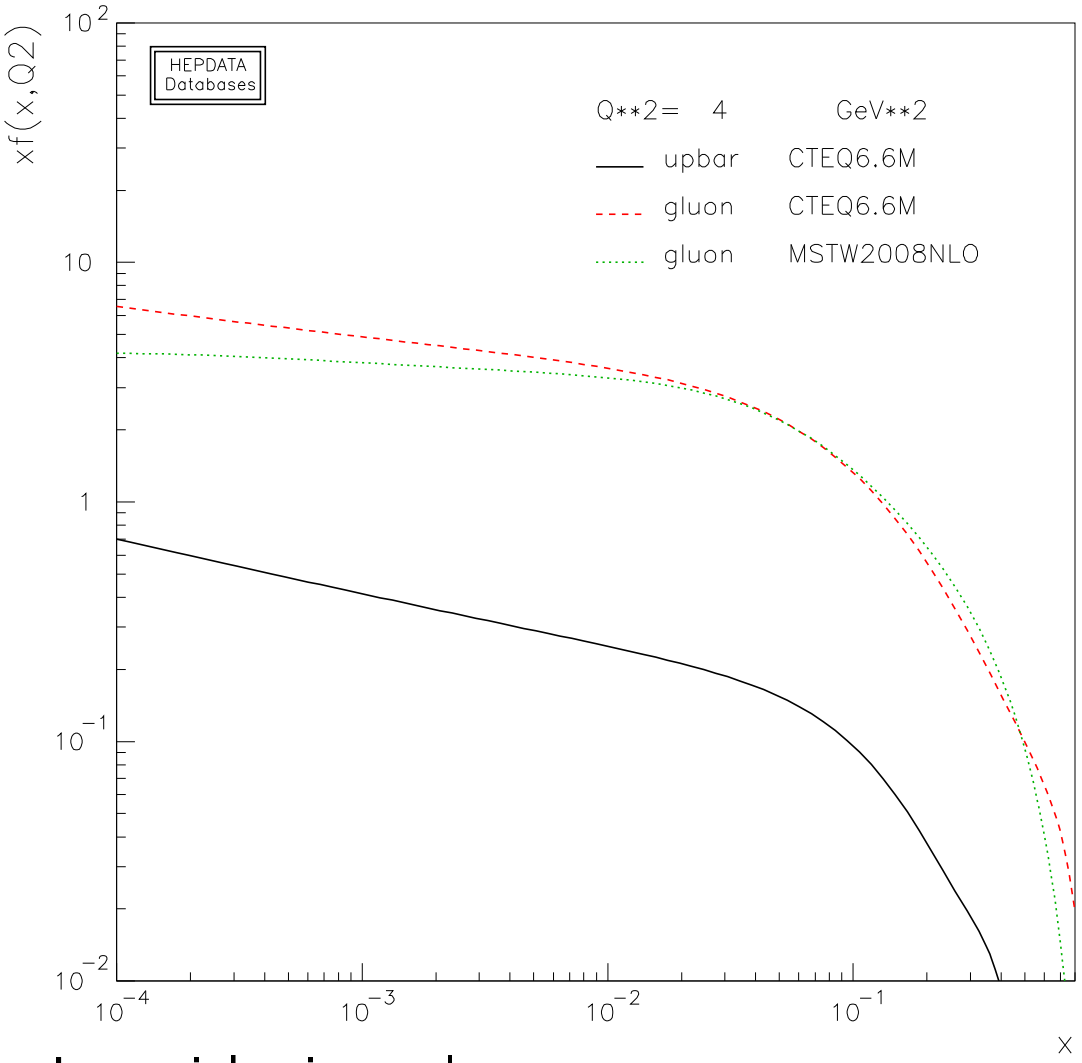
$g(x, Q)$ can become negative at $x < 10^{-2}$, $Q < 2$ GeV

may lead to unphysical predictions

This is an indication that DGLAP factorization experiences difficulties at such small x and Q

Large $\ln^k(1/x)$ in $P_{i/j}(x)$ break PQCD expansion at $x \sim Q/\sqrt{s} \ll 1$

Example of DGLAP evolution: $\bar{u}$ and gluon PDF



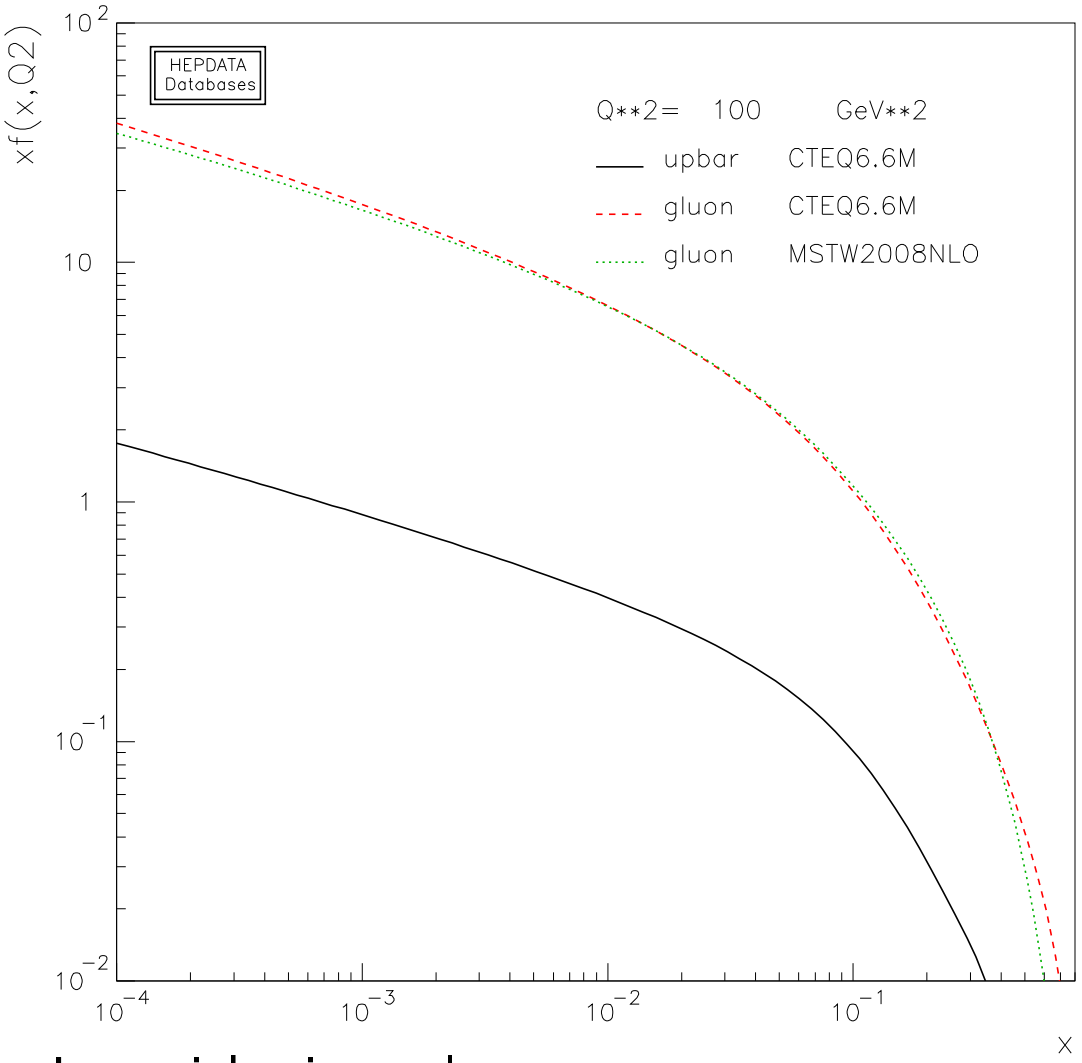
Logarithmic scale

As Q increases, $g(x, Q)$ grows rapidly at small x

$\alpha_s(Q)$ becomes small enough to suppress $\ln^k(1/x)$ terms

small-x behavior stabilizes

Example of DGLAP evolution: $\bar{u}$ and gluon PDF



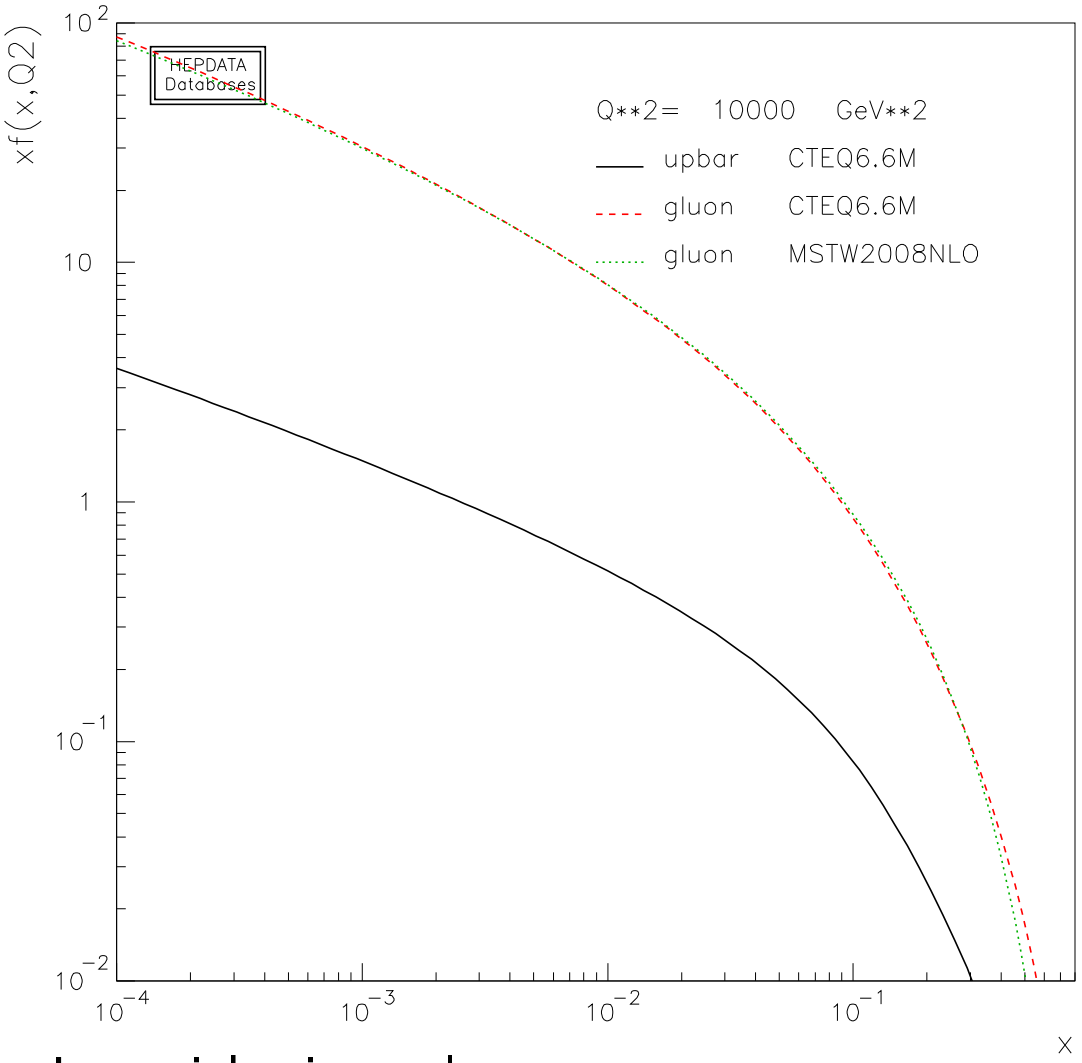
Logarithmic scale

As Q increases, $g(x, Q)$ grows rapidly at small x

$\alpha_s(Q)$ becomes small enough to suppress $\ln^k(1/x)$ terms

small-x behavior stabilizes

Example of DGLAP evolution: $\bar{u}$ and gluon PDF



Logarithmic scale

As Q increases, $g(x, Q)$ grows rapidly at small x

$\alpha_s(Q)$ becomes small enough to suppress $\ln^k(1/x)$ terms

small-x behavior stabilizes

Universality of PDFs

To all orders in α_s, PDFs are **defined** as matrix elements of certain correlator functions:

$$f_{q/p}(x,\mu) = \int_{-\infty}^{\infty} \frac{dy^-}{4\pi} e^{iy^- p^+} \langle p \left| \overline{\psi}_q(0, y^-, \vec{0}_T) F(y^-, 0) \gamma^+ \psi_q(0, 0, \vec{0}_T) \right| p \rangle, \text{ etc.}$$

PDFs are **universal** – depend only on the type of the hadron (p) and parton (q, $\bar{q}$, g)

... can be **parametrized** as

$$f_{i/p}(x, Q_0) = a_0 x^{a_1} (1-x)^{a_2} F(a_3, a_4, ...) \text{ at } Q_0 \sim 1 \text{ GeV}$$

... predicted by solving DGLAP equations at $\mu > Q_0$

Where do the PDFs come from?

CT10 NNLO PDFs

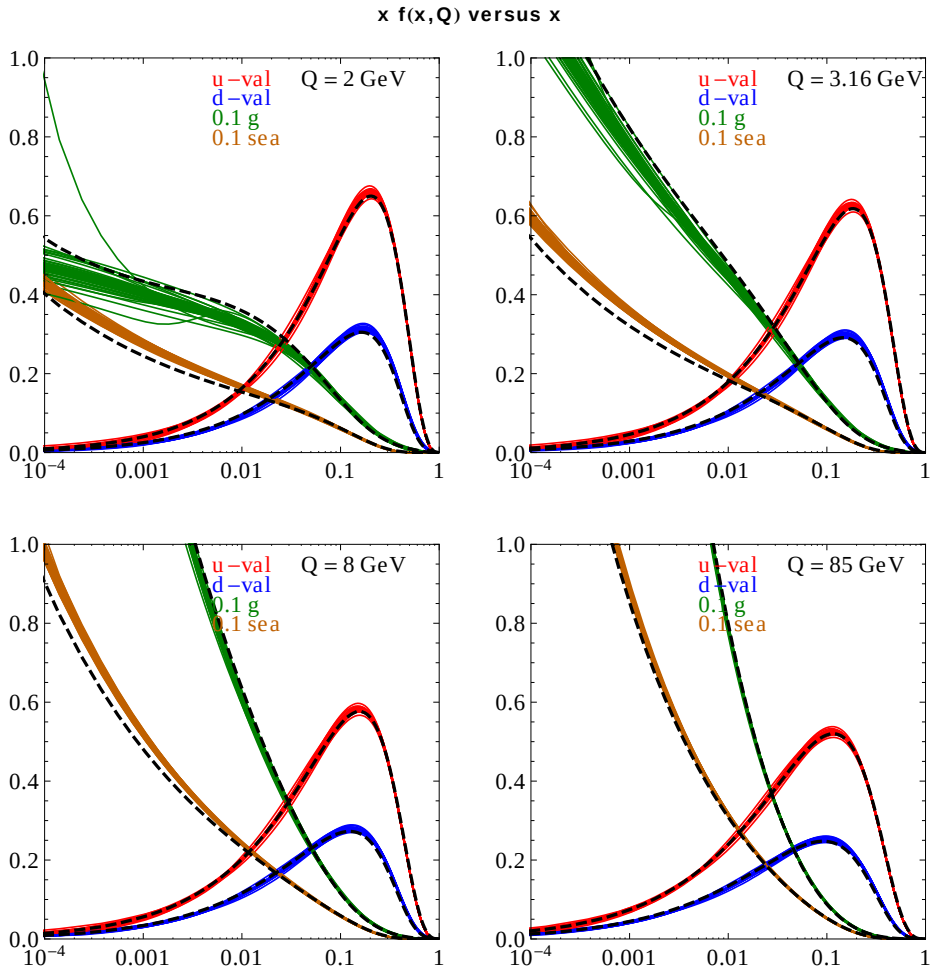


FIG. 2: CT10NNLO parton distribution functions. These figures show the Hessian error PDFs from the CT10NNLO analysis. Each graph shows $x\,u_{\rm valence} = x(u-\overline{u}), x\,d_{\rm valence} = x(d-\overline{d}), 0.10\,x\,g$ and $0.10\,x\,q_{\rm sea}$ as functions of x for a fixed value of Q. The values of Q are 2, 3.16, 8, 85 GeV. The quark sea contribution is $q_{\rm sea} = 2(\overline{d}+\overline{u}+\overline{s})$. The dashed curves are the central CT10 NLO fit.

Where do the PDFs come from?

- From a combination of BIG, medium, and **small** experiments
- Complementarity in
  - kinematical ranges
  - systematics

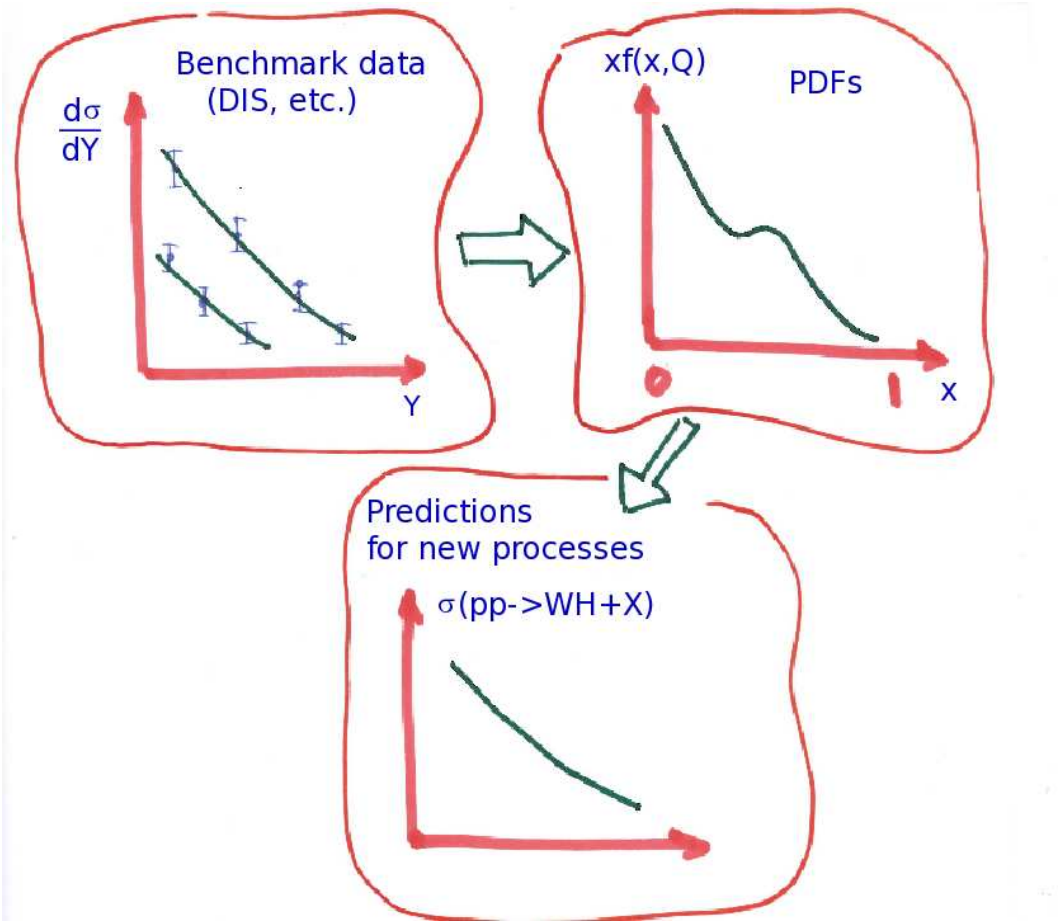
+ lattice QCD

LHC
Tevatron

HERA
RHIC
EIC

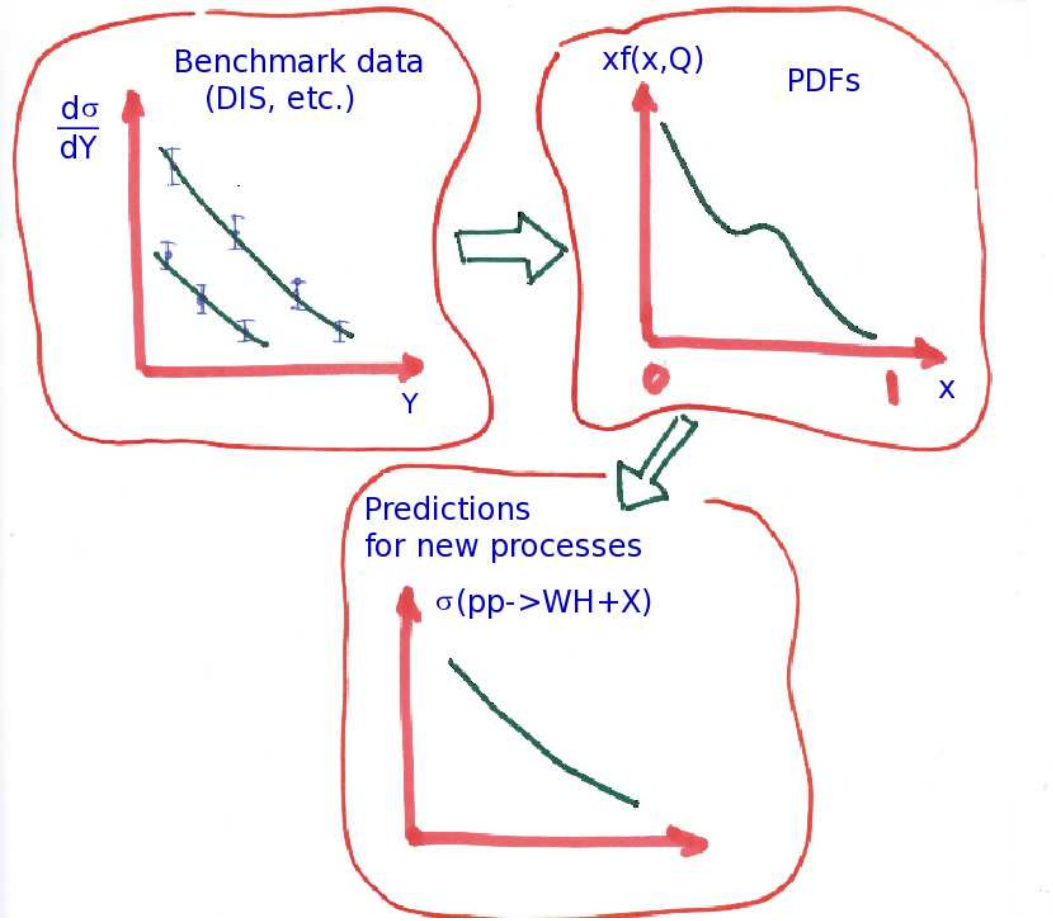
Fixed-target experiments

The flow of the PDF analysis



PDFs are not measured directly, but some data sets are sensitive to specific combinations of PDFs. By constraining these combinations, the PDFs can be disentangled in a combined (global) fit.

The flow of the PDF analysis

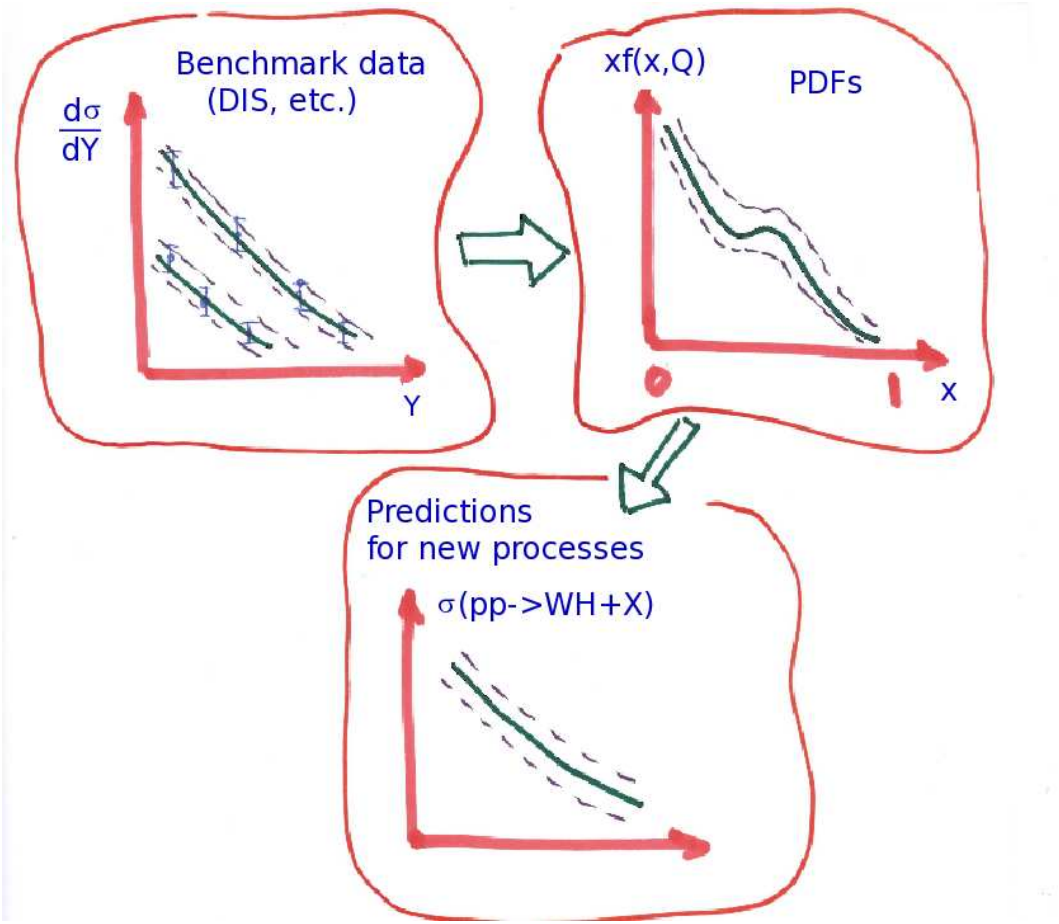


Data sets and $\chi^2/d.o.f.$ in CT10 NNLO and CT10W NLO analyses

Experimental data set	N_{pt}	CT10NNLO	CT10W
Combined HERA1 NC and CC DIS [74]	579	1.07	1.17
BCDMS F_2^p [75]	339	1.16	1.14
BCDMS F_2^d [76]	251	1.16	1.12
NMC F_2^p [77]	201	1.66	1.71
NMC F_2^d/F_2^p [77]	123	1.23	1.28
CDHSW F_2^p [78]	85	0.83	0.66
CDHSW F_3^p [78]	96	0.81	0.75
CCFR F_2^p [79]	69	0.98	1.02
CCFR xF_3^p [80]	86	0.40	0.59
NuTeV neutrino dimuon SIDIS [81]	38	0.78	0.94
NuTeV antineutrino dimuon SIDIS [81]	33	0.86	0.91
CCFR neutrino dimuon SIDIS [82]	40	1.20	1.25
CCFR antineutrino dimuon SIDIS [82]	38	0.70	0.78
H1 F_2^c [83]	8	1.17	1.26
H1 σ_r^c for $c\bar{c}$ [59, 84]	10	1.63	1.54
ZEUS F_2^c [57]	18	0.74	0.90
ZEUS F_2^c [58]	27	0.62	0.76
E605 Drell-Yan process, $\sigma(pA)$ [85]	119	0.80	0.81
E866 Drell Yan process, $\sigma(pd)/(2\sigma(pp))$ [86]	15	0.65	0.64
E866 Drell-Yan process, $\sigma(pp)$ [87]	184	1.27	1.21
CDF Run-1 W charge asymmetry [88]	11	1.22	1.24
CDF Run-2 W charge asymmetry [89]	11	1.04	1.02
DØ Run-2 $W \to e\nu_e$ charge asymmetry [90]	12	2.17	2.11
DØ Run-2 $W \to \mu\nu_\mu$ charge asymmetry [91]	9	1.65	1.49
DØ Run-2 Z rapidity distribution [92]	28	0.56	0.54
CDF Run-2 Z rapidity distribution [93]	29	1.60	1.44
CDF Run-2 inclusive jet production [94]	72	1.42	1.55
DØ Run-2 inclusive jet production [95]	110	1.04	1.13
Total:	**2641**	**1.11**	**1.13**

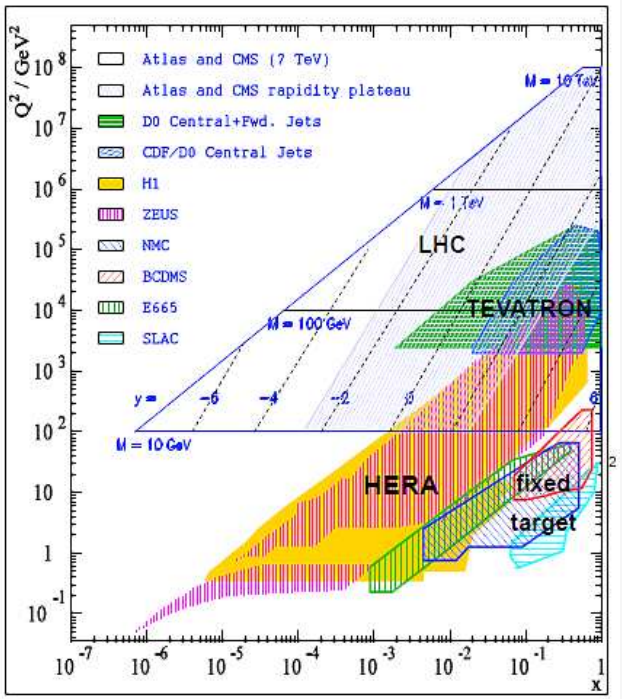
CT10 NNLO fit involves 28 experiments, 2640 data points,
25 PDF parameters, 100+ correlated systematic parameters

The flow of the PDF analysis

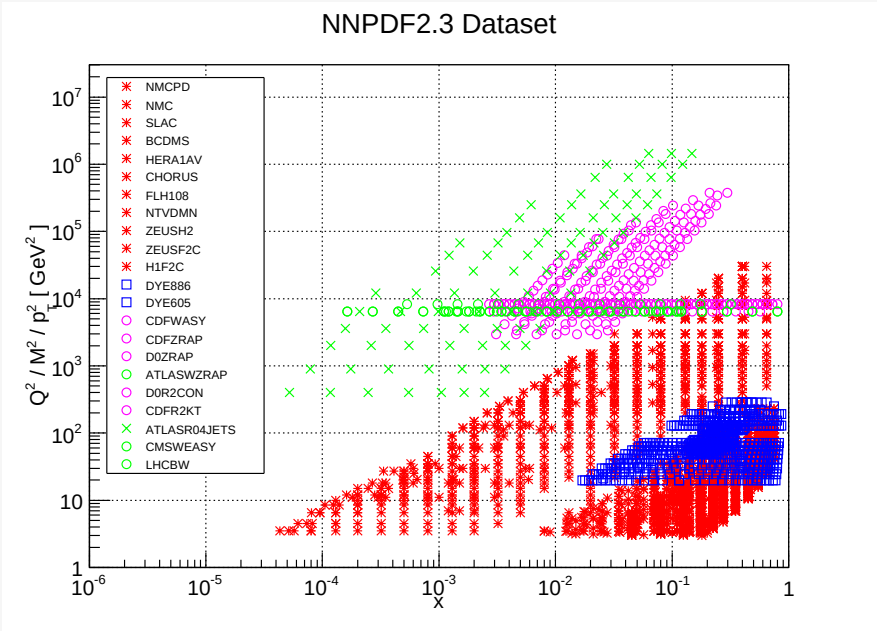


We are interested not just in one best fit, but also in the uncertainty of the resulting PDF parametrizations and theoretical predictions based on them. This will be covered in Lecture 2

x, Q coverage of various experiments



Experiments included in the NNPDF2.3 PDF analysis



Which SM particles can have a non-zero PDF in the proton?

Should we introduce $f_{a/p}(x, Q)$ for any of the following particles?

1. light partons u, d, s, g
2. heavy quarks c, b, t
3. photon γ; leptons e, μ, τ, ν
4. massive electroweak bosons W and Z

Which SM particles can have a non-zero PDF in the proton?

Should we introduce $f_{a/p}(x, Q)$ for any of the following particles?

1. light partons u, d, s, g
2. heavy quarks c, b, t
3. photon γ; leptons e, μ, τ, ν
4. massive electroweak bosons W and Z

Answer

All of them – the PDF can be defined for any particle

However, only partons with mass $\lesssim 1$ GeV are expected to have a non-negligible $f_{a/p}(x, Q_0)$ at the initial scale $Q_0 \approx 1$ GeV

Boundary conditions at Q_0

In practice, independent parametrizations $f_{a/p}(x, Q_0)$ are introduced for

- $g, u, d, s, \bar{u}, \bar{d}, \bar{s}$ (always)
  contribute $> 97\%$ of the proton's energy E_p at Q_0
  - even in this case, the data are usually insufficient for constraining all PDF parameters; some of them can be fixed by hand
  - e.g., $\bar{u} = \bar{d} = \bar{s}$ in outdated fits
- c and or b (occasionally; in a model allowing nonperturbative "intrinsic heavy-quark production")
- photons γ (in MRST'03 QCD+QED PDFs, in the upcoming sets by all groups)
  - a QCD+QED fit is more complicated than one might think: it must account for violation of charge symmetry by EM effects,

$$u_p(x, Q) \neq d_n(x, Q); \;\; d_p(x, Q) \neq u_n(x, Q)$$

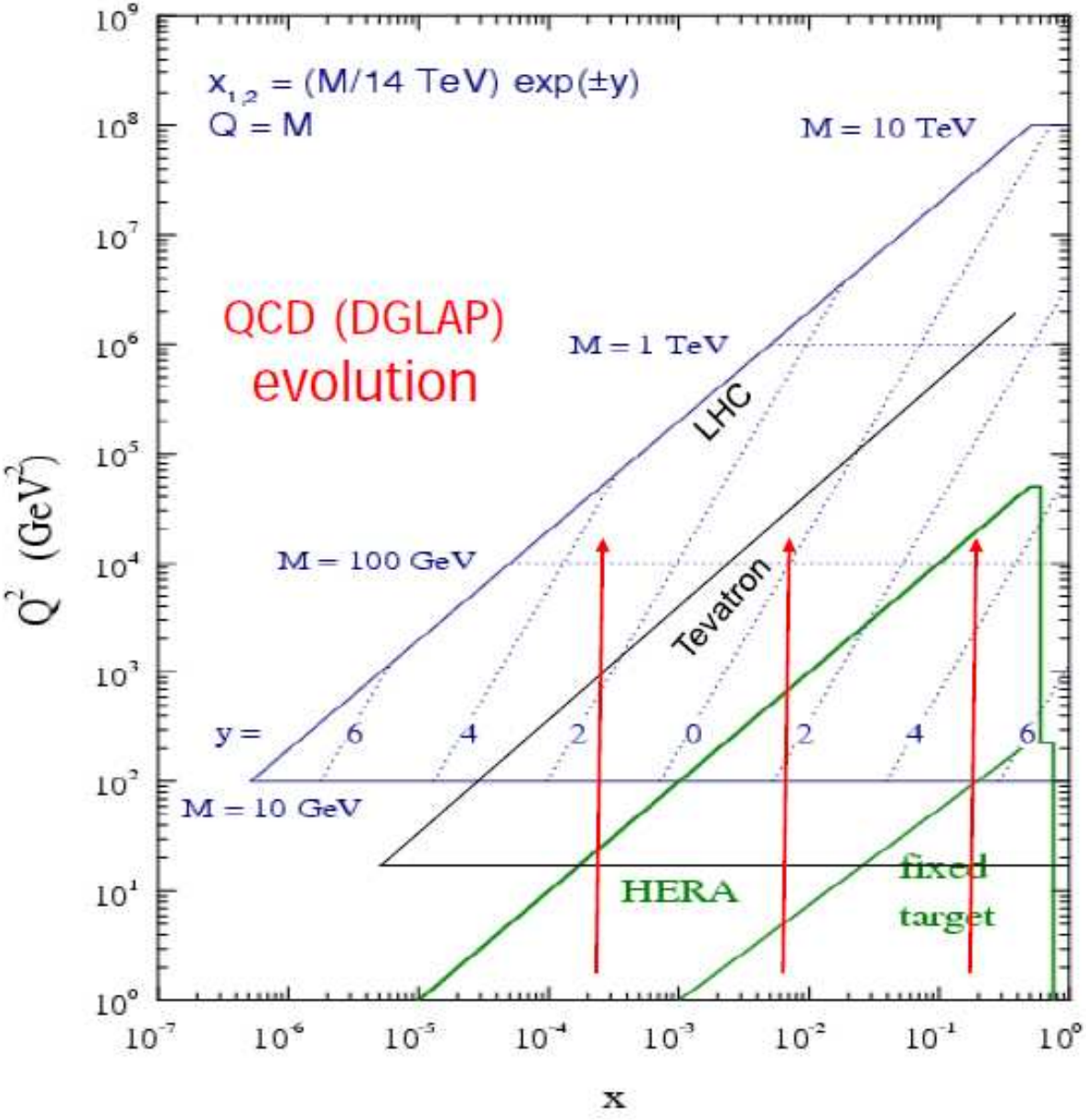
PDFs for heavy flavors

- PDFs for heavy partons h can be generated via DGLAP evolution at $Q \geq m$, using a boundary condition $f_{h/p}(x, Q) = 0$ at $Q \leq m$
- In practice:
  - perturbative PDFs are usually introduced for c and b quarks
  - QCD coupling $\alpha_s(Q)$ and PDFs are evaluated with 5 active flavors at all $Q \geq m_b$
  - Logarithmic enhancements may exist in collinear t, W, Z production at $Q \gtrsim 1$ TeV; PDFs for t, W, Z "partons" may be introduced at such Q

Experimental observables constraining the PDFs in global fits

Kinematics of Parton variables

Predictive power of global analysis of PDFs is based on the renormalization group properties of the universal Parton Distributions f(x,Q).



Some basics about PDFs

- Parton Distribution Function $f(x, Q)$
- Given a heavy resonance with mass Q produced at hadron collider with c.m. energy $\sqrt{S}$
- What's the typical x value?

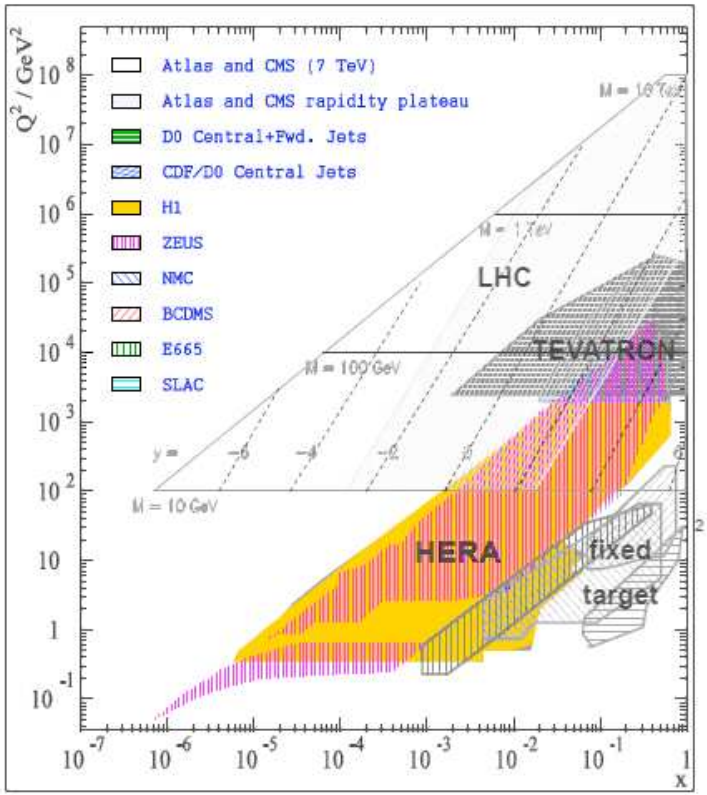
$$< x >= \frac{Q}{\sqrt{S}}$$ at central rapidity (y=0)

- Generally, $x_1 = \frac{Q}{\sqrt{S}} e^{y}$ *and* $x_2 = \frac{Q}{\sqrt{S}} e^{-y}$

$$x_1 + x_2 = 2\frac{Q}{\sqrt{S}}\cosh(y) \quad \Rightarrow \quad y_{\max}: \; x_1 + x_2 = 1$$

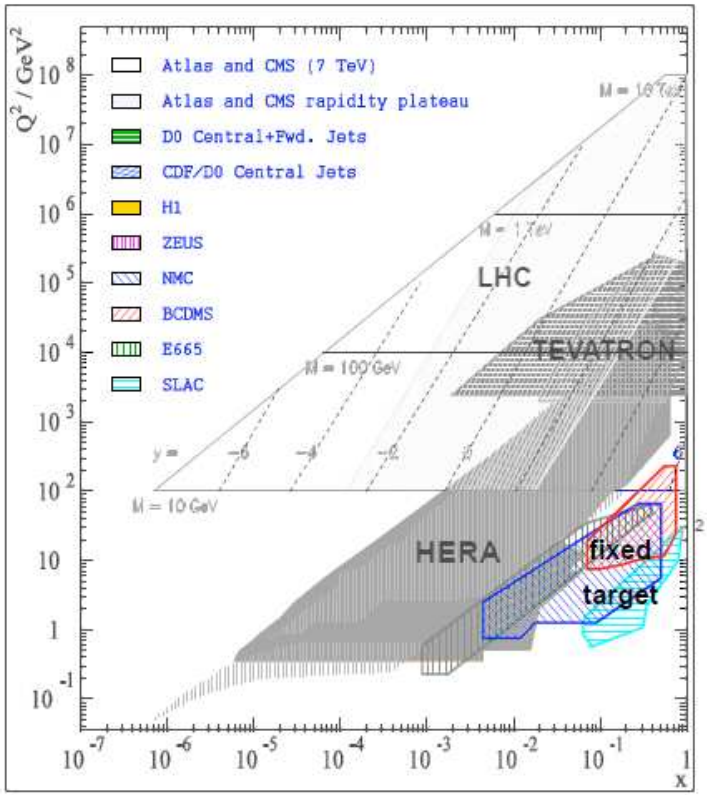
Selection of experimental data for PDF fits (2014)

- **Inclusive deep-inelastic scattering**
  - At HERA:
    neutral-current $e^{\pm}p \to e^{\pm}X$;
    charged-current $ep \to \nu X$
    - ♢ the largest data set in the fit
  - Fixed-target experiments
    - ♢ eN, μN, νN scattering



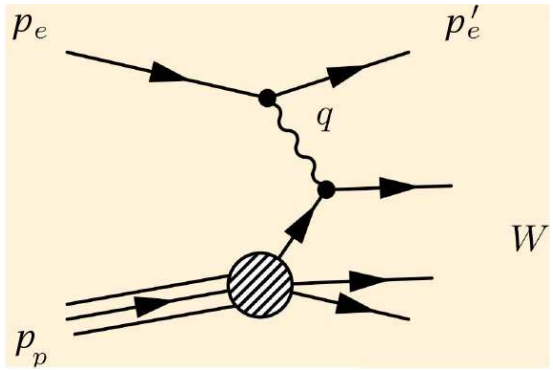
Selection of experimental data for PDF fits (2014)

- **Inclusive deep-inelastic scattering**
  - At HERA:
    neutral-current $e^{\pm}p \to e^{\pm}X$;
    charged-current $ep \to \nu X$
    - ♢ the largest data set in the fit
  - Fixed-target experiments
    - ♢ eN, μN, νN scattering



Neutral-current ep DIS: kinematics

- $s = (p_e + p_p)^2$ –total energy
- $Q^2 = -q^2 = -(p_e - p'_e)^2$– momentum transfer
- $x = Q^2/(2p_p \cdot q)$ – Bjorken scaling variable
- $y = Q^2/(xs)$ – inelasticity
- $W^2 = Q^2(1-x)/x$ – energy of the hadronic final state



$$\frac{d^2\sigma(e^{\pm}p)}{dQ^2dx} = \frac{2\pi\alpha^2}{Q^4x}Y_+\left(F_2 - \frac{y^2}{Y_+}F_L \pm \frac{Y_-}{Y_+}xF_3\right),$$

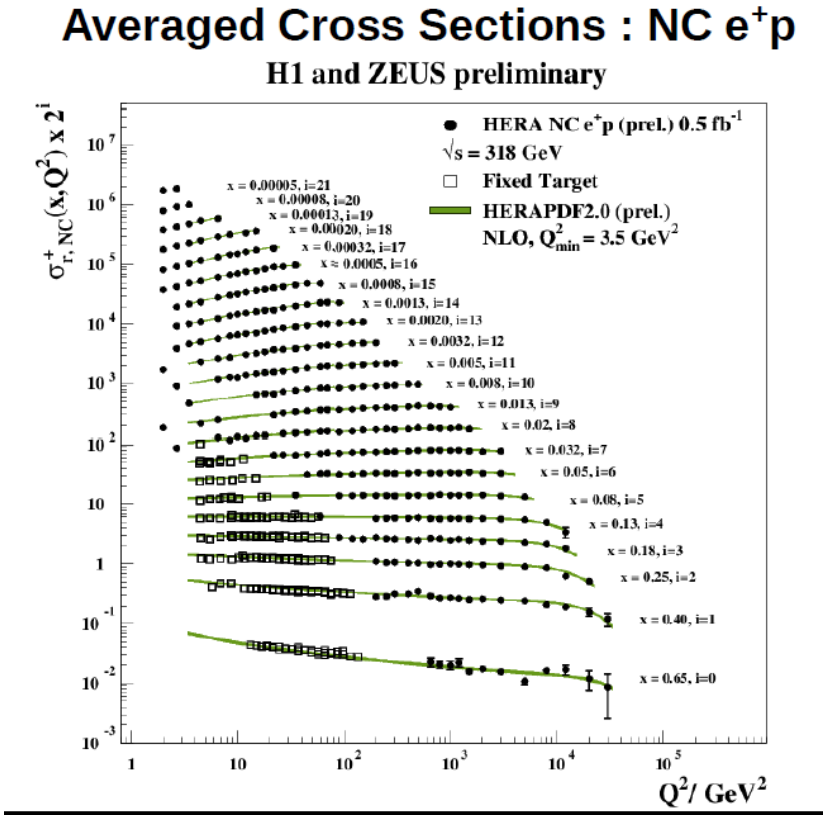
with $Y_\pm \equiv 1 \pm (1-y)^2$

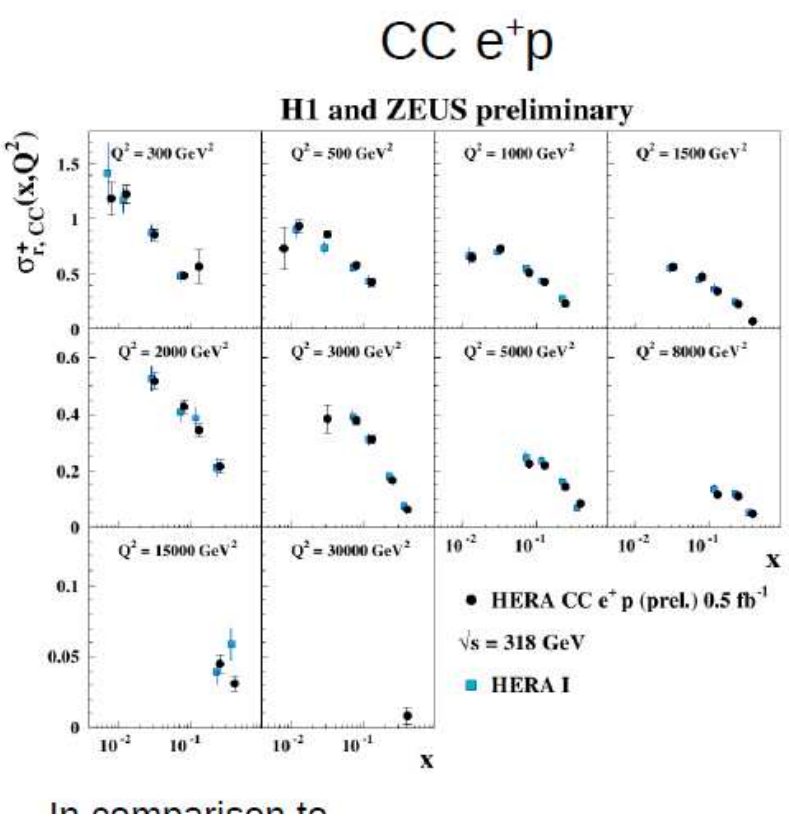
The data is fitted either in the form of $F_2(x, Q^2)$ or $d^2\sigma/(dQ^2dx)$

Combined data from HERA DIS experiments

41 data sets on NC and CC DIS from H1 and ZEUS are combined into 1 set.

2927 data points are combined into 1307 data points. 165 correlated systematic errors are reanalyzed and calibrated.





PDF combinations in DIS at the lowest order

- Neutral current $\ell^{\pm}p$:

$$F_2^{\ell^{\pm}p}(x,Q^2) = \frac{4}{9}\left(u+\bar{u}+c+\bar{c}\right) + \frac{1}{9}\left(d+\bar{d}+s+\bar{s}+b+\bar{b}\right)$$

  - PDFs are weighted by the fractional EM quark coupling $e_i^2 = 4/9$ or $1/9$
  - 4 times more sensitivity to u and c than to d, s, and b
  - No sensitivity to the gluon at this order

- Neutral current $(\ell^{\pm}N)$ DIS on isoscalar nuclei $(N = (p+n)/2)$:

$$F_2^{\ell^{\pm}N}(x,Q^2) = \frac{5}{9}\left(u+\bar{u}+d+\bar{d}+\text{smaller } s,\ c,\ b \text{ contributions}\right)$$

- Charged current (νN) DIS :

$$F_2^{\nu N}(x,Q^2) = x\sum_{i=u,d,s\ldots}\left(q_i+\bar{q}_i\right)$$

$$xF_3^{\nu N}(x,Q^2) = x\sum_{i=u,d,s}\left(q_i-\bar{q}_i\right)$$

DIS at next-to-leading order (NLO) and beyond

Logarithmic corrections to Bjorken scaling (Q dependence of $F_2(x,Q^2)$) are sensitive to the gluon PDF through DGLAP equations,

$$\mu\frac{df_{i/p}(x,\mu)}{d\mu} = \sum_{j=g,u,\bar{u},d,\bar{d},....} \int_x^1 \frac{dy}{y} P_{i/j}\left(\frac{x}{y},\alpha_s(\mu)\right) f_{j/p}(y,\mu)$$

Thus, when examined at NLO, the DIS data constrains

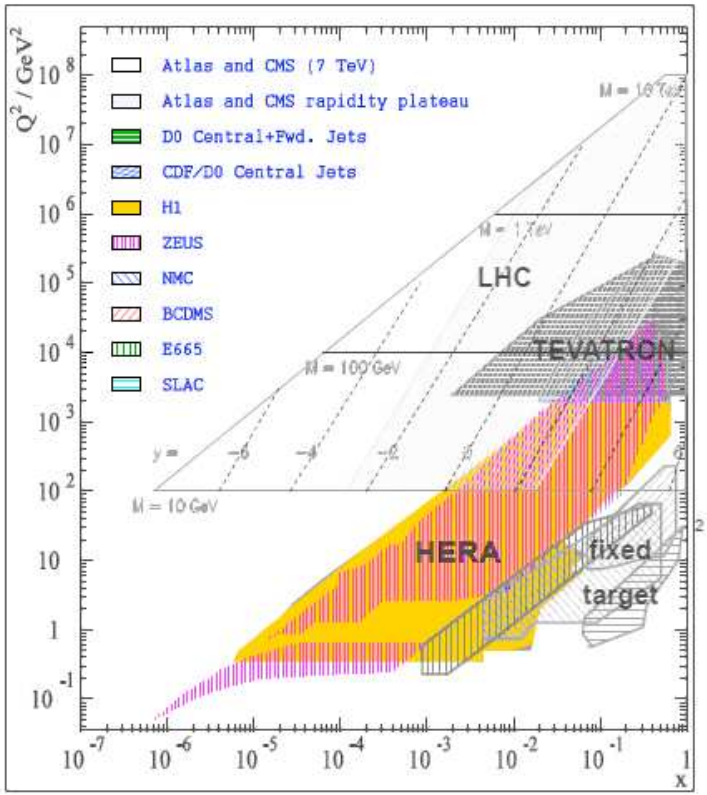
- $\sum_i e_i^2(q_i+\bar{q}_i)$ in an amazingly large range $10^{-5} < x < 0.5$
- u and d at $10^{-2} < x < 0.3$
- $g(x,Q)$ at $x < 0.1$

DIS cannot fully separate quarks from antiquarks, or s, c, b contributions from u and d contributions; more so because of systematic effects in fixed-target DIS experiments (higher-order terms, nuclear corrections,...)

Selection of experimental data for PDF fits (2014)

The modern PDF fits include

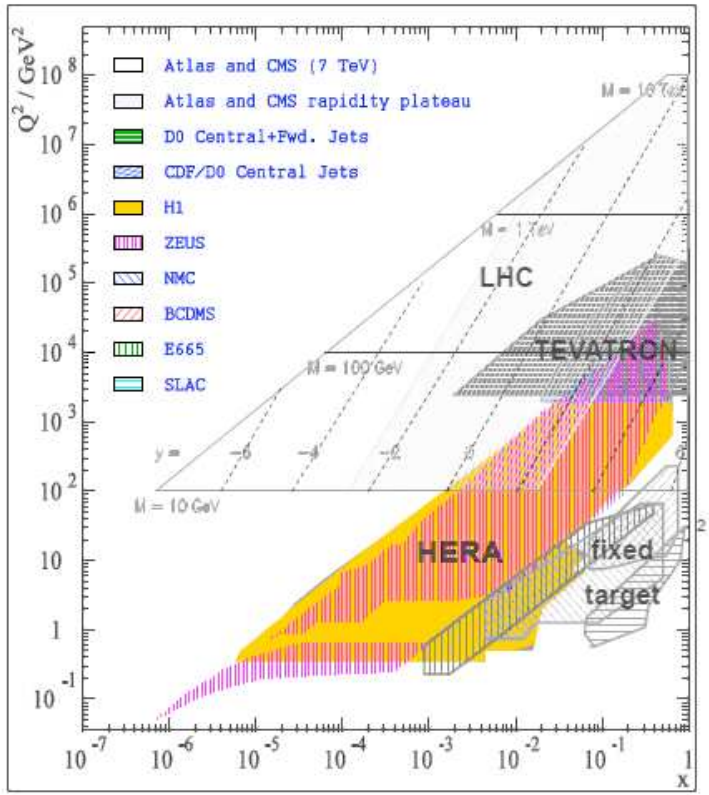
Inclusive deep-inelastic scattering...



Selection of experimental data for PDF fits (2014)

The modern PDF fits include
Inclusive deep-inelastic scattering...

+ **Semi-inclusive DIS:**
  - charm production $ep \to ecX$ (HERA)
  - $\mu\mu$ production $\nu N \to \mu(c \to \mu)X$ (NuTeV, not shown)



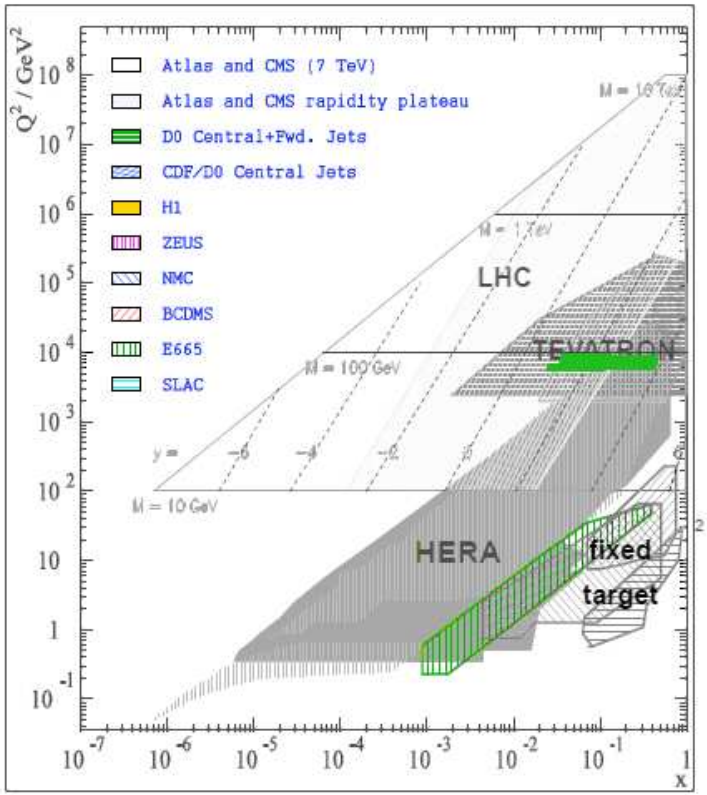
Hard cross sections are known at NNLO (two QCD loops) for inclusive DIS and $ep \to ecX$; partly at NNLO for $\nu N \to \mu\mu X$

Selection of experimental data for PDF fits (2014)

The modern PDF fits include
Inclusive deep-inelastic scattering...

+ **Semi-inclusive DIS:**
  - charm production $ep \to ecX$ (HERA)
  - $\mu\mu$ production $\nu N \to \mu(c \to \mu)X$ (NuTeV, not shown)

+ **Lepton pair production** $pN \xrightarrow{\gamma^*,W,Z} \ell\bar{\ell}'X$ (Tevatron, fixed-target experiments)



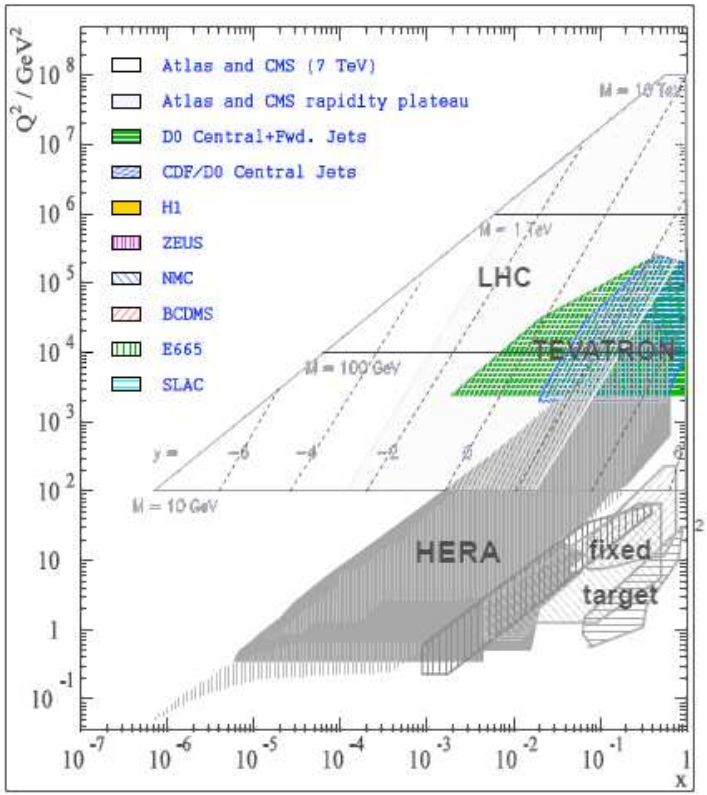
Selection of experimental data for PDF fits (2014)

The modern PDF fits include
Inclusive deep-inelastic scattering...

+ **Semi-inclusive DIS:**
  - charm production $ep \to ecX$ (HERA)
  - $\mu\mu$ production $\nu N \to \mu(c \to \mu)X$ (NuTeV, not shown)

+ **Lepton pair production** $pN \xrightarrow{\gamma^*,W,Z} \ell\bar{\ell}'X$ (Tevatron, fixed-target experiments)

+ **Inclusive jet production:** $p\bar{p} \to jX$ (Tevatron), $ep \to j(j)X$ (HERA)

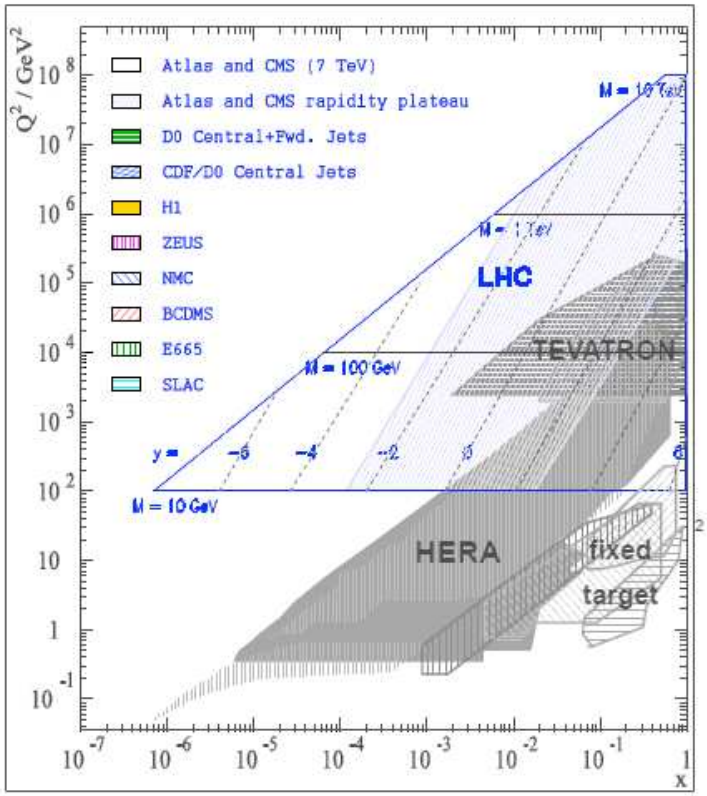


Hard cross sections are known at NNLO (two loops) for lepton pair production, partial NNLO for jet production

Selection of experimental data for PDF fits (2014)

LHC revolutionizes the PDF analysis

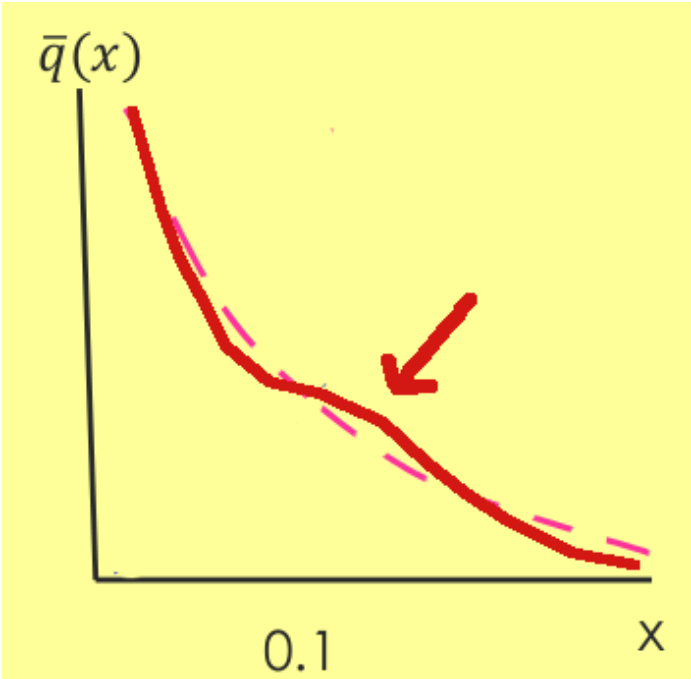
- Multiple data sets from γ^*, W, Z, jet, Wc, $t\bar{t}$,photon,... production at high luminosity
- to be included in the PDF fits very soon
- fits to "collider-only" data will become competitive



Zooming in on quark sea PDFs

Various QCD effects produce non-trivial sea PDFs

- breaking of SU(2) symmetry ($\bar{d}(x) \neq \bar{u}(x)$) and charge symmetry ($\bar{q}(x) \neq q(x)$)
- non-trivial shape of sea PDFs, cf. the figure



1% accuracy can distinguish between these effects.

Unpolarized integrated PDFs must be known to $\sim$1% to determine polarized and TMD PDFs, fragmentation functions, and for LHC physics

Constraints on quark sea from $pN \to \ell^+\ell^- X$

$(N = p,\, d,\, Fe,\, Cu, ..)$

$$\frac{d\sigma_{pp}}{dQ^2dy} \sim (\tfrac{2}{3})^2 \left[u_A\bar{u}_B + \bar{u}_A u_B\right] + (-\tfrac{1}{3})^2 \left[d_A\bar{d}_B + \bar{d}_A d_B\right] + \text{ smaller terms}$$

$\Rightarrow$ sensitivity to $\bar{q}(x, Q)$

Assuming charge symmetry between protons and neutrons $(u_p = d_n,\, u_n = d_p)$:

$$\frac{d\sigma_{pn}}{dQ^2dy} \sim (\tfrac{2}{3})^2 \left[u_A\bar{d}_B + \bar{u}_A d_B\right] + (-\tfrac{1}{3})^2 \left[d_A\bar{u}_B + \bar{d}_A u_B\right] + \text{ smaller terms}$$

If deuterium binding corrections are neglected: $q_d(x) \approx q_p(x) + q_n(x)$

At $x_A \gg x_B$ (large y): $\bar{q}(x_A) \sim 0$ and $4u(x_A) \gg d(x_A)$

Constraints on quark sea from $pN \to \ell^+\ell^- X$

($N = p,\ d,\ Fe,\ Cu, ..$)

$$\frac{d\sigma_{pp}}{dQ^2dy} \sim (\tfrac{2}{3})^2 \left[u_A\bar{u}_B + \bar{u}_A u_B\right] + (-\tfrac{1}{3})^2 \left[d_A\bar{d}_B + \bar{d}_A d_B\right] + \text{ smaller terms}$$

$\Rightarrow$ sensitivity to $\bar{q}(x, Q)$

Assuming charge symmetry between protons and neutrons ($u_p = d_n,\ u_n = d_p$):

$$\frac{d\sigma_{pn}}{dQ^2dy} \sim (\tfrac{2}{3})^2 \left[u_A\bar{d}_B + \bar{u}_A d_B\right] + (-\tfrac{1}{3})^2 \left[d_A\bar{u}_B + \bar{d}_A u_B\right] + \text{ smaller terms}$$

If deuterium binding corrections are neglected: $q_d(x) \approx q_p(x) + q_n(x)$

At $x_A \gg x_B$ (large y): $\bar{q}(x_A) \sim 0$ and $4u(x_A) \gg d(x_A)$

Constraints on quark sea from $pN \to \ell^+\ell^- X$

($N = p,\, d,\, Fe,\, Cu, ..$)

$$\frac{d\sigma_{pp}}{dQ^2dy} \sim (\tfrac{2}{3})^2 \left[u_A\bar{u}_B + \bar{u}_A u_B\right] + (-\tfrac{1}{3})^2 \left[d_A\bar{d}_B + \bar{d}_A d_B\right] + \text{ smaller terms}$$

$\Rightarrow$ sensitivity to $\bar{q}(x, Q)$

Assuming charge symmetry between protons and neutrons ($u_p = d_n,\, u_n = d_p$):

$$\frac{d\sigma_{pn}}{dQ^2dy} \sim (\tfrac{2}{3})^2 \left[u_A\bar{d}_B + \bar{u}_A d_B\right] + (-\tfrac{1}{3})^2 \left[d_A\bar{u}_B + \bar{d}_A u_B\right] + \text{ smaller terms}$$

If deuterium binding corrections are neglected: $q_d(x) \approx q_p(x) + q_n(x)$

At $x_A \gg x_B$ (large y): $\bar{q}(x_A) \sim 0$ and $4u(x_A) \gg d(x_A)$

Constraints on quark sea from $pN \to \ell^+\ell^- X$

$(N = p,\, d,\, Fe,\, Cu, ..)$

At large $x_A \gg x_B$ (large y):

$$\frac{\sigma_{pd}}{2\sigma_{pp}} \approx \frac{1}{2}\frac{(1+\frac{d_A}{4u_A})[1+r]}{(1+\frac{d_A}{4u_A}r)} \approx \frac{1}{2}(1+r), \text{ where } r \equiv \overline{d}(x_B)/\overline{u}(x_B)$$

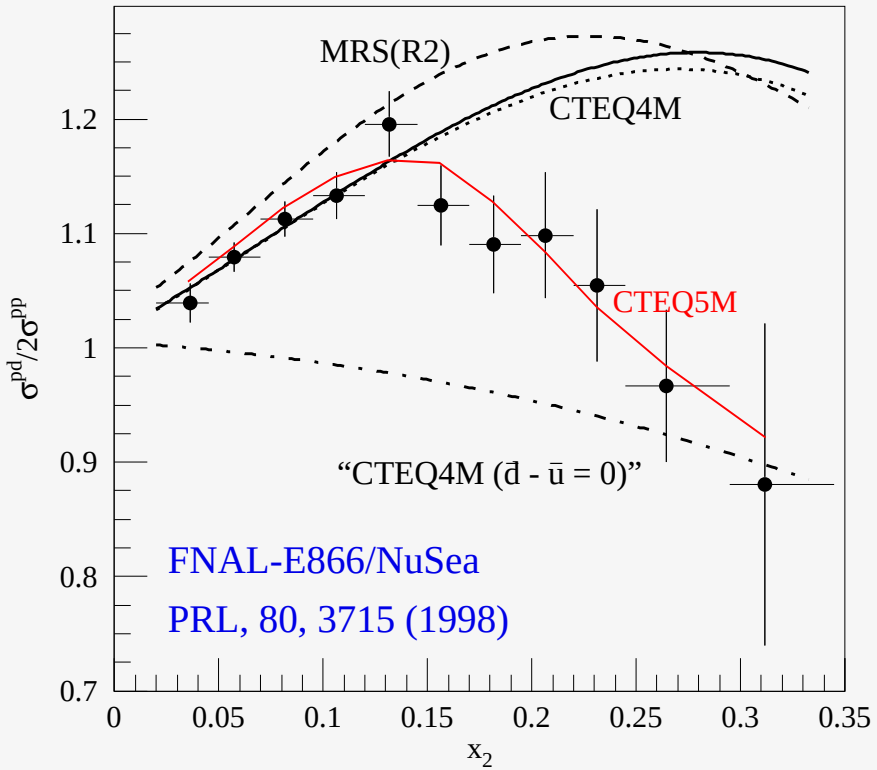
$\therefore \sigma_{pd}/(2\sigma_{pp})$ constrains $\bar{d}(x_B, Q)/\bar{u}(x_B, Q)$ at moderate x_B

Experimental evidence for SU(2) symmetry breaking

E866 Drell-Yan pair production:
$\bar{d}(x) - \bar{u}(x) \neq 0$ at $x > 0.1$
(large difference)

LHC W/Z production:
$\bar{d}(x) - \bar{u}(x) \neq 0$ at $x < 0.1$
(a few percent, see next slides)

$\sigma_{pd}/(2\sigma_{pp})$ at large $x_F = x_A - x_B$



Theory curves reflect different assumptions about $\bar{d}/\bar{u}$

PDF fits (e.g., CTEQ5M) quantitatively account for the violation of $SU(2)$ symmetry in the quark sea

Charged lepton asymmetry in $AB \to (W \to e\nu_e)X$ ($A, B = p$ or $\bar{p}$)

y_e and $\eta \approx y_e$ are rapidity and pseudorapidity of an electron from W decay

$$A_{ch}(y_e) \equiv \frac{\frac{d\sigma^{W^+}}{dy_e} - \frac{d\sigma^{W^-}}{dy_e}}{\frac{d\sigma^{W^+}}{dy_e} + \frac{d\sigma^{W^-}}{dy_e}}$$

$A_{ch}(y_e)$ relates to the boson asymmetry $A_{ch}(y) = \frac{(d\sigma^{W^+}/dy) - (d\sigma^{W^-}/dy)}{(d\sigma^{W^+}/dy) + (d\sigma^{W^-}/dy)}$, where

$$\left(d\sigma^{W^+}/dy\right) \propto u_A(x_A, M_W)\bar{d}_B(x_B, M_W) + \bar{d}_A(x_A, M_W)u_B(x_B, M_W) + ...$$

$$\left(d\sigma^{W^-}/dy\right) \propto \bar{u}_A(x_A, M_W)d_B(x_B, M_W) + d_A(x_A, M_W)\bar{u}_B(x_B, M_W) + ...$$

Charged lepton asymmetry in $AB \to (W \to e\nu_e)X$ ($A, B = p$ or $\bar{p}$)

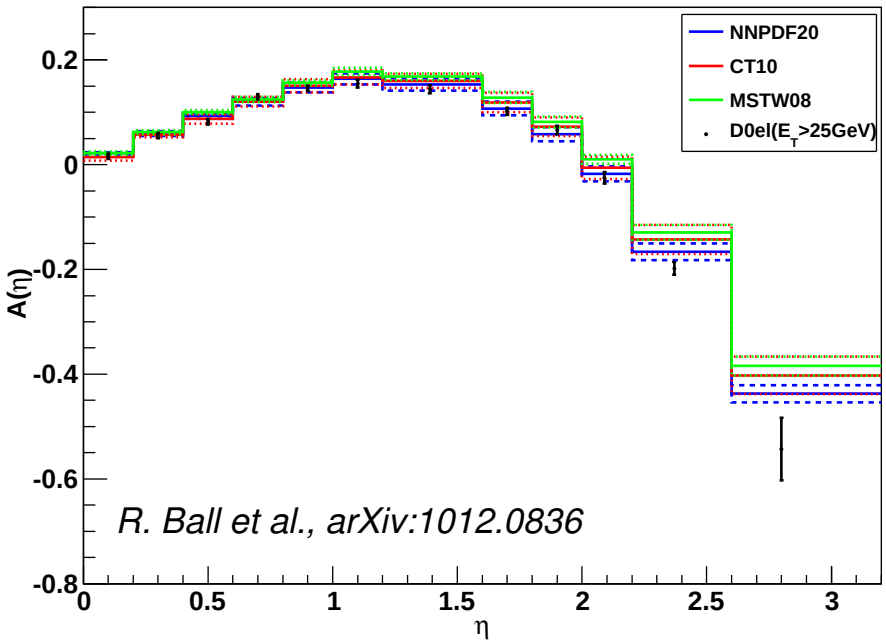
y_e and $\eta \approx y_e$ are rapidity and pseudorapidity of an electron from W decay

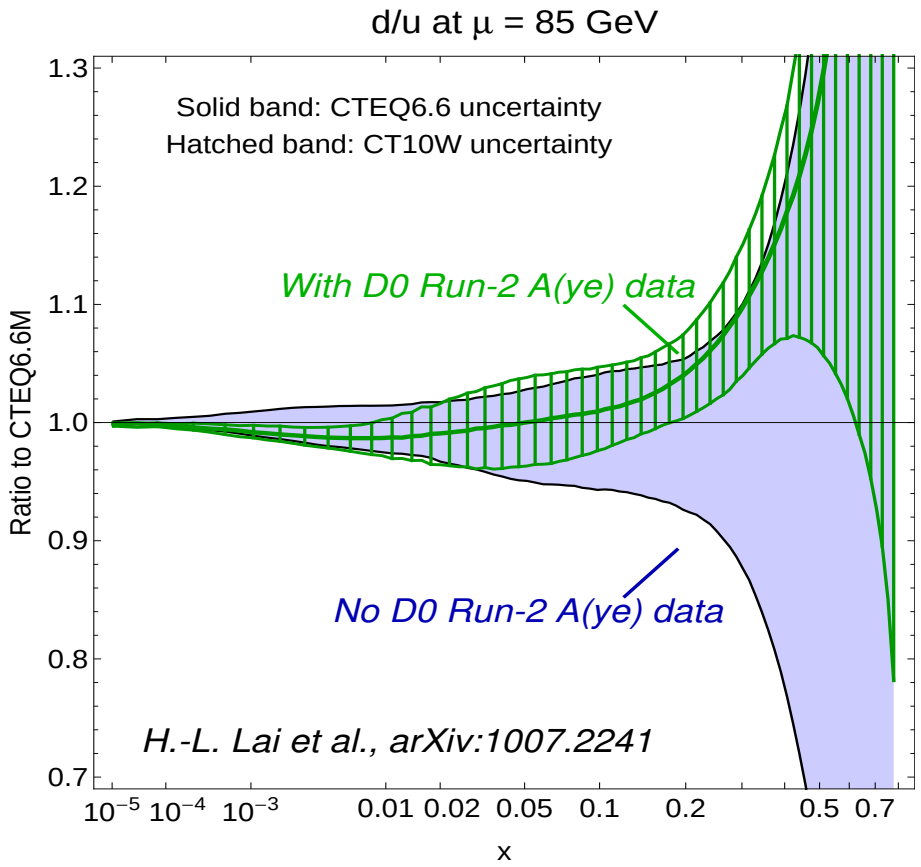
$$A_{ch}(y_e) \equiv \frac{\frac{d\sigma^{W^+}}{dy_e} - \frac{d\sigma^{W^-}}{dy_e}}{\frac{d\sigma^{W^+}}{dy_e} + \frac{d\sigma^{W^-}}{dy_e}}$$

$\therefore A_{ch}(y_e)$ constrains PDF ratios at $Q \approx M_W$:

- d/u at $x \to 1$ at the Tevatron 1.96 TeV ($p\bar{p}$);
- d/u at $x > 0.1$ **and** $\bar{u}/\bar{d}$ at $x \sim 0.01$ at the LHC 7 TeV (pp)

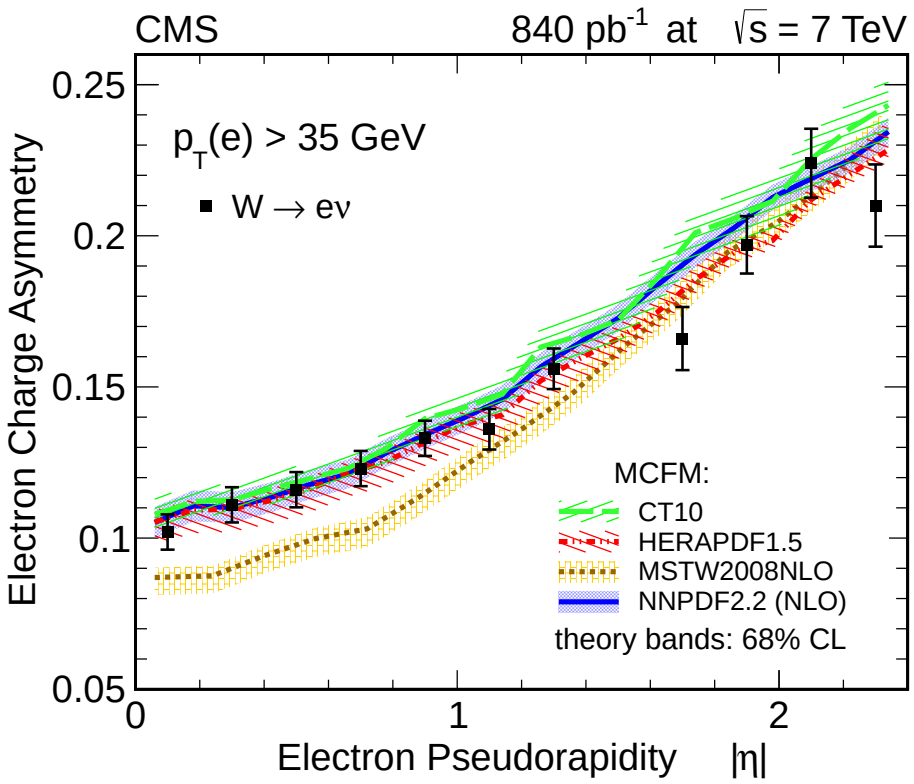
Impact of the D0 Run-2 A_{ch} data on PDFs





- The A_{ch} data from the Tevatron D0 collaboration distinguish between the PDF models, reduce the PDF uncertainty

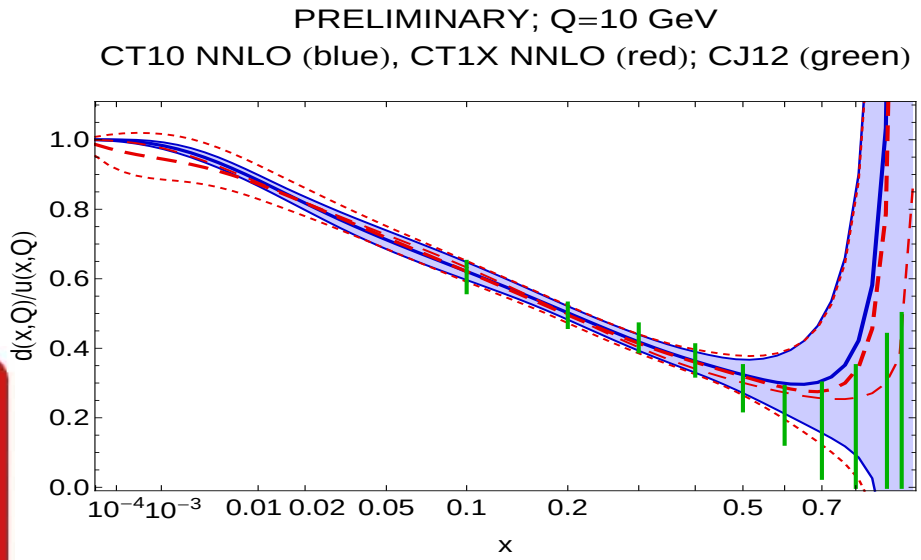
Charge asymmetry at the Tevatron and LHC

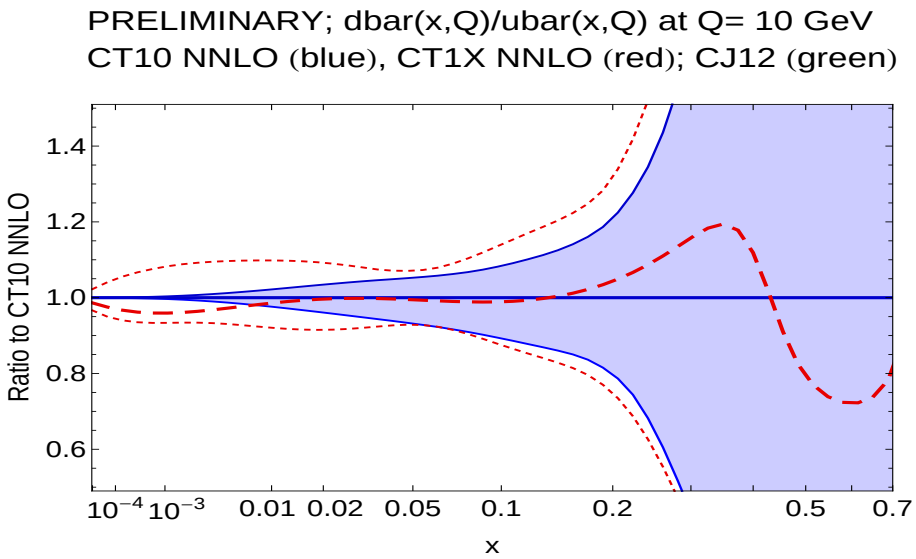


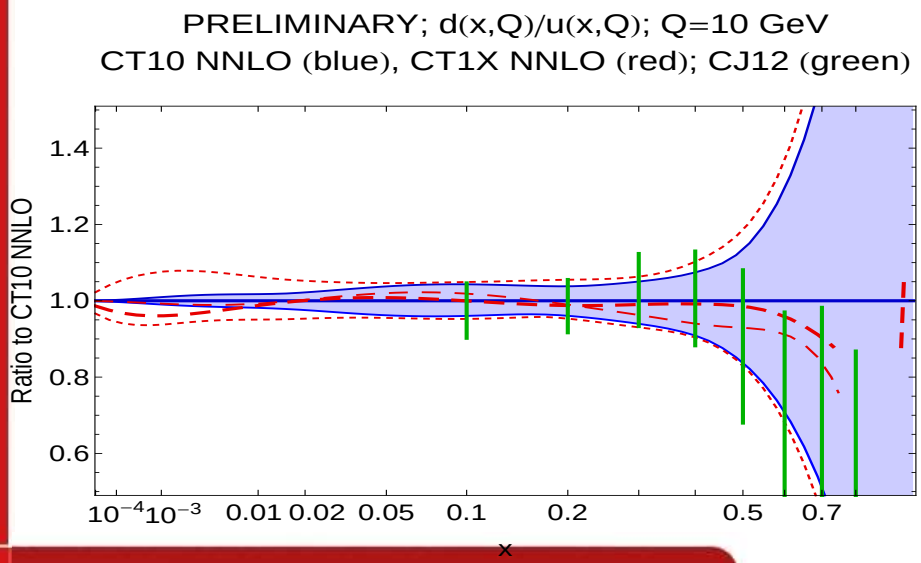
$$A_{ch}(\eta) \equiv \frac{\frac{d\sigma^{W^+}}{d\eta} - \frac{d\sigma^{W^-}}{d\eta}}{\frac{d\sigma^{W^+}}{d\eta} + \frac{d\sigma^{W^-}}{d\eta}}$$

CMS $A_{ch}(\eta)$ data disfavor some d/u parametrizations, motivated an update in MSTW'2008 PDFs

d/u and $\bar{d}/\bar{u}$: CT1X NNLO vs. CT10 NNLO and CJ 12 analysis of large-x DIS

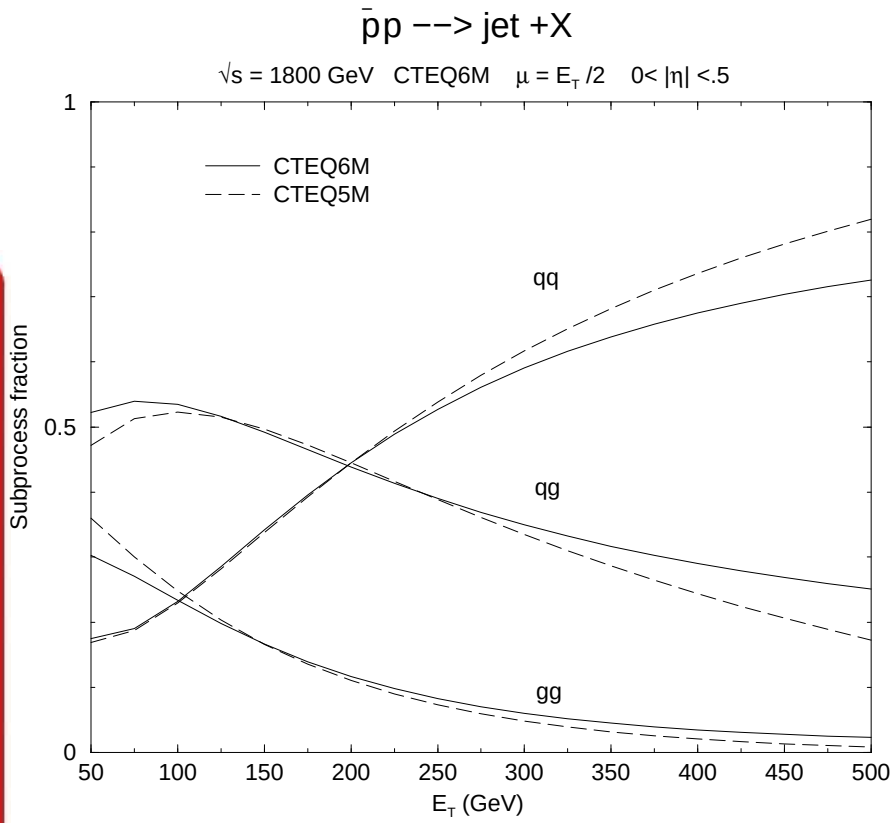






CT1X PDF uncertainty is larger at $x \to 0$ and 1, is compatible with the d/u band from the CJ12 analysis *(Owens et al., 1212.1702)* of large-x DIS for PDFs+nuclear+higher-twist corrections

Inclusive jet production, $p\overset{(-)}{p} \to \text{jet} + X$

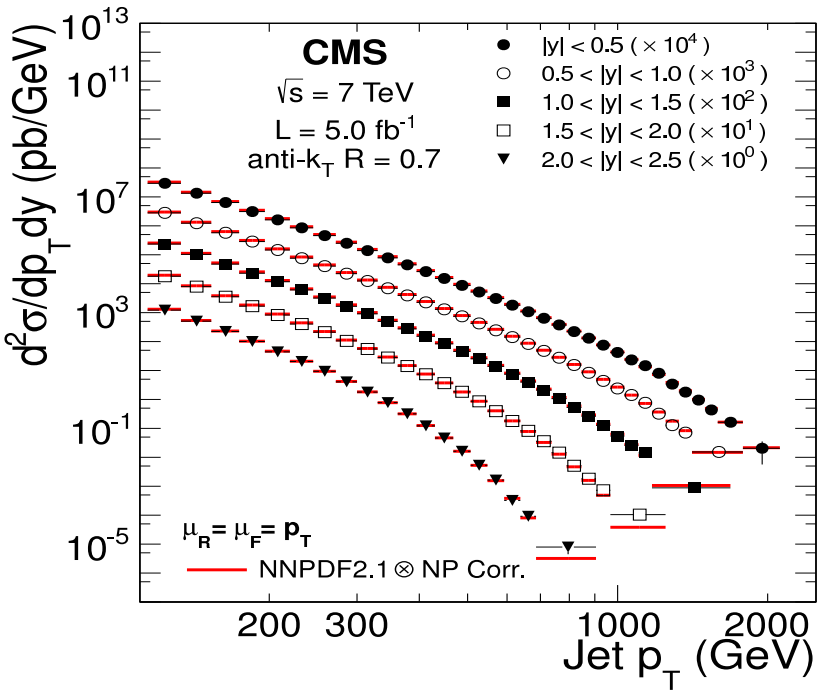


High-E_T jets are mostly produced in qq scattering; yet most of the PDF uncertainty arises from qg and gg contributions

Here typical x is of order $2E_T/\sqrt{s} \gtrsim 0.1$;
e.g., $x \approx 0.2$ for $E_T = 200$ GeV, $\sqrt{s} = 1.8$ TeV

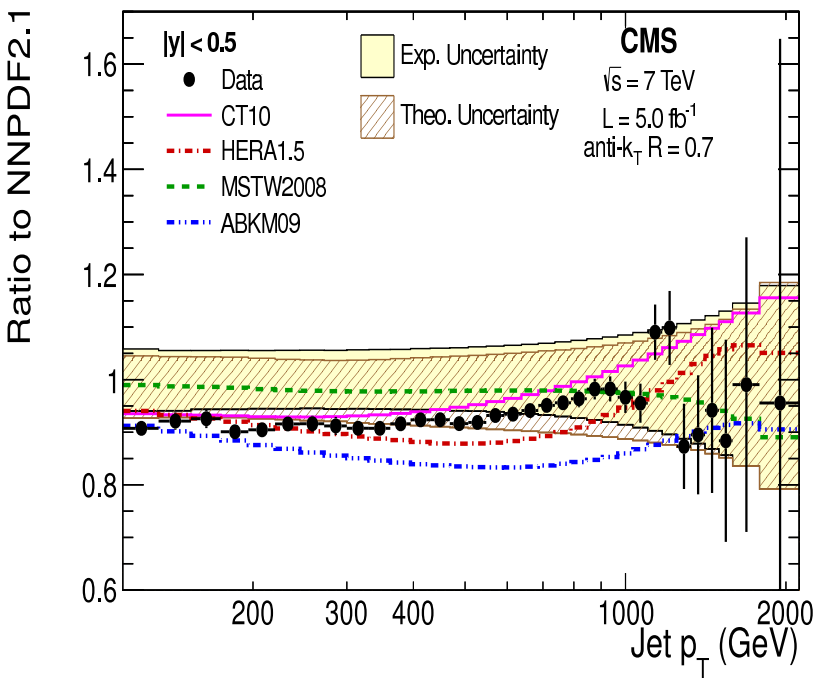
At such x, $u(x,Q)$ and $d(x,Q)$ are known very well; uncertainty arises mostly from $g(x,Q)$

Inclusive jet production in $pp \to$ jet $+X$ (7 TeV)



- The cross sections span 12 orders of magnitude
- (Almost) negligible statistical error
- Sensitive to $g(x, Q)$ at $x \gtrsim 0.001$, cf. Lecture 2

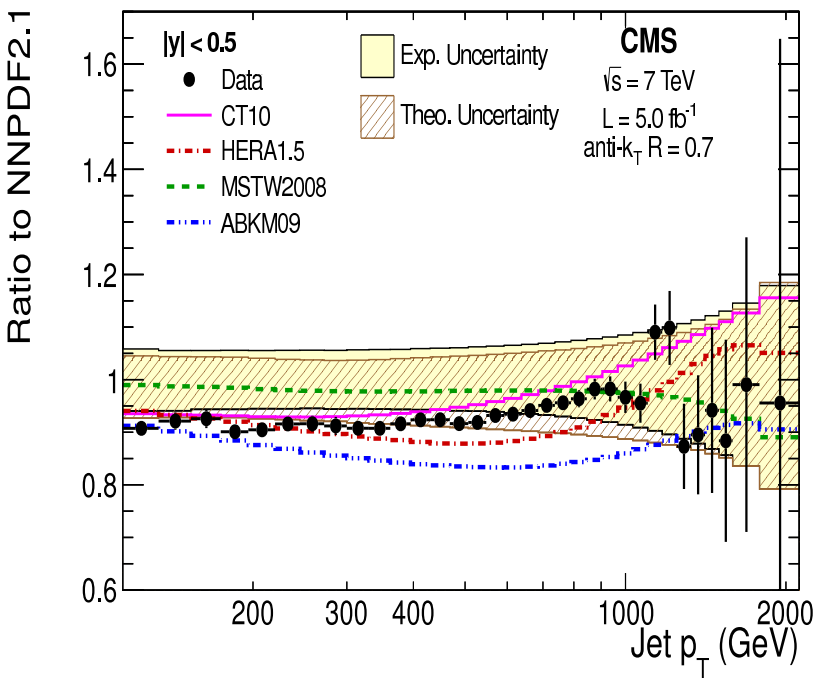
Inclusive jet production in $pp \to$ jet $+X$ (7 TeV)



- The cross sections span 12 orders of magnitude
- (Almost) negligible statistical error
- Sensitive to $g(x, Q)$ at $x \gtrsim 0.001$, cf. Lecture 2

- Systematic uncertainties dominate, both from the experiment (up to 90 correlated sources of uncertainty) and NLO theoretical cross section (QCD scale dependence)
- The PDF uncertainty would be strongly underestimated if these systematic errors are not included

Inclusive jet production in $pp \to$ jet $+X$ (7 TeV)

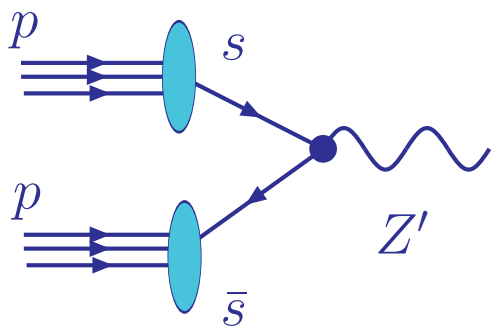


- The cross sections span 12 orders of magnitude
- (Almost) negligible statistical error
- Sensitive to $g(x, Q)$ at $x \gtrsim 0.001$, cf. Lecture 2

- Lecture 2 will discuss how to include the correlated systematic errors into the PDF analysis

Homework assignment

An exotic boson Z' with mass $Q = 2$ TeV is produced similarly to SM Z bosons, but only via the $s\bar{s} \to Z'$ vertex [Z' does not interact with non-strange (anti-)quarks].



Z'couples only to s,$\bar{s}$

You need to compute $\sigma(pp \to Z'X)$ at the LHC $\sqrt{s} = 13000$ GeV, but for that you need to precisely know the strange (anti-)quark PDFs, $s(x, Q)$ and $\bar{s}(x, Q)$. Propose one or two scattering processes to constrain $s(x, Q)$ and $\bar{s}(x, Q)$ at the relevant $\{x, Q\}$. Specify $\sqrt{s}$ and other kinematic parameters of these processes. Can you use non-LHC measurements to constrain $s(x, Q)$ at the LHC? Why or why not?

Choice of PDF parametrization

Statistical aspects

Practical applications