Parton Distribution Functions and their applications

Pavel Nadolsky

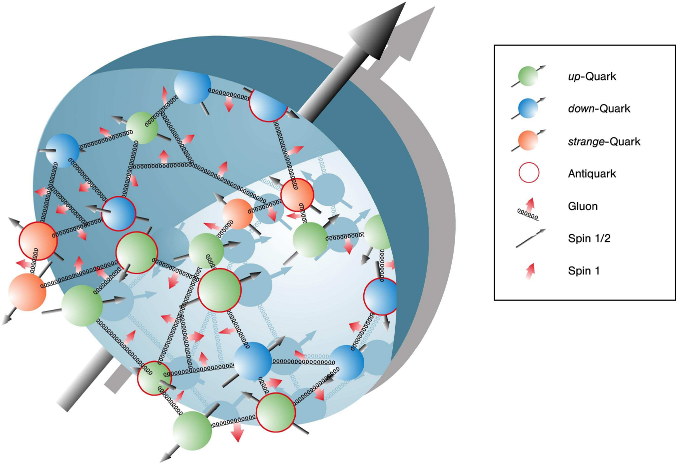
Department of Physics
Southern Methodist University (Dallas, TX)

Lecture 1
July 21, 2017

1. Key ideas

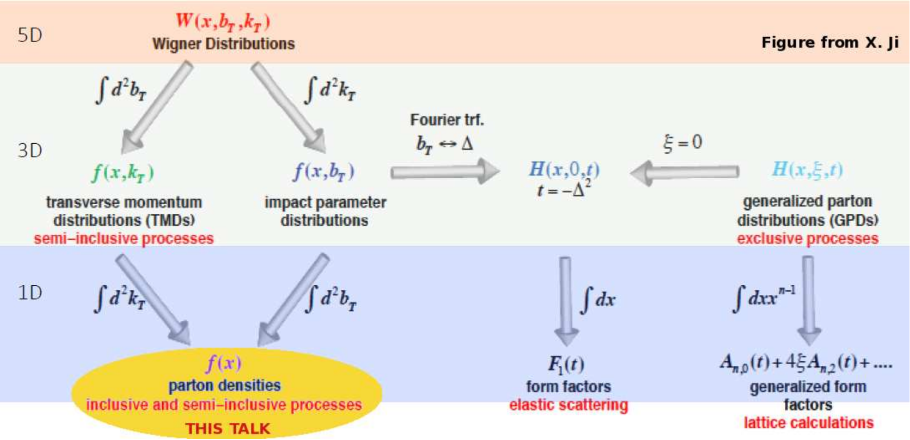
Parton distribution functions $f_{a/p}(x, Q)$...

... are the key functions describing the nucleon structure in QCD



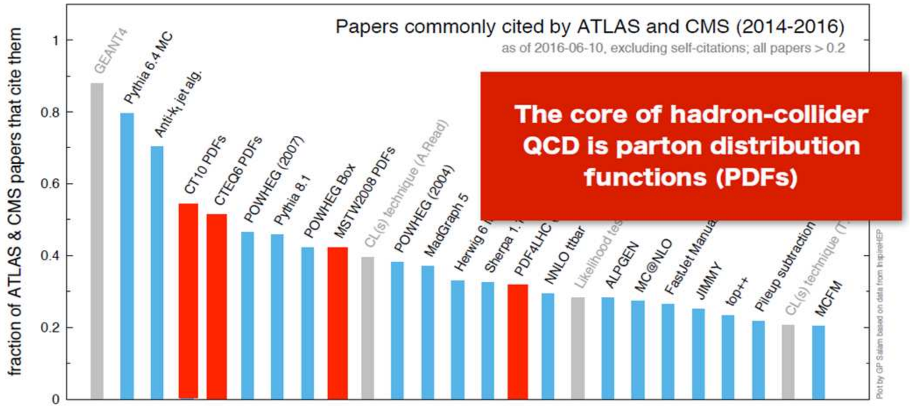
Parton distribution functions $f_{a/p}(x,Q)$...

... the best-known nonperturbative functions introduced in QCD



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

...are indispensable in computations of inclusive hadronic reactions at CERN and other laboratories

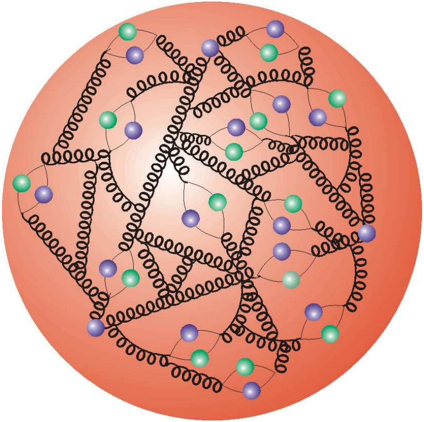


Basic definitions

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

  (Feynman; Bjorken, Paschos - 1969)

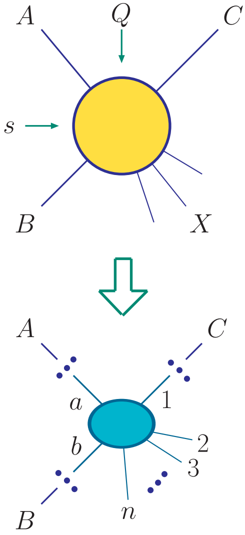
  - assumed to be pointlike at present



Basic definitions

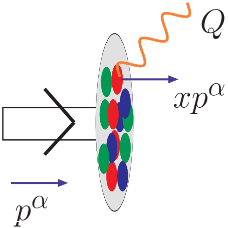
- Partons are most easily detected in **inclusive** hadronic scattering $A + B \to C + X$ at large collision energy $\sqrt{s} \gg 1$ GeV, with typical energy transfer Q of order $\sqrt{s}$

- 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

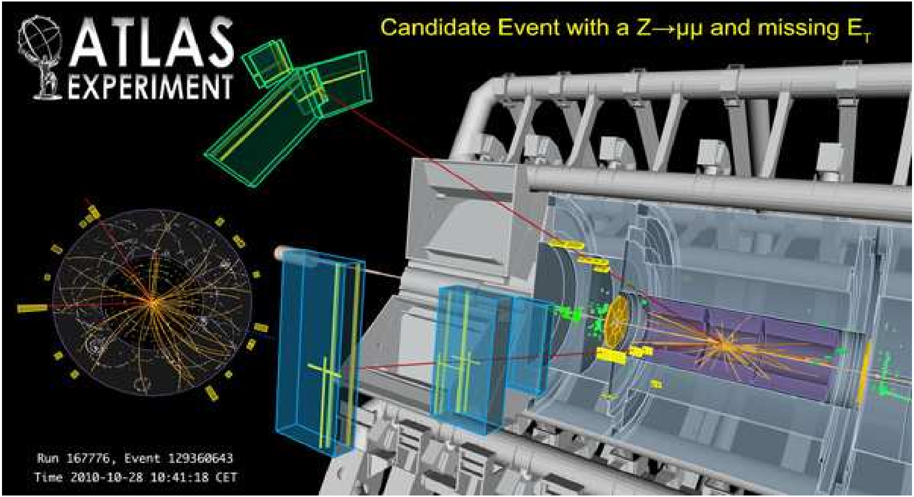


Basic definitions

- 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

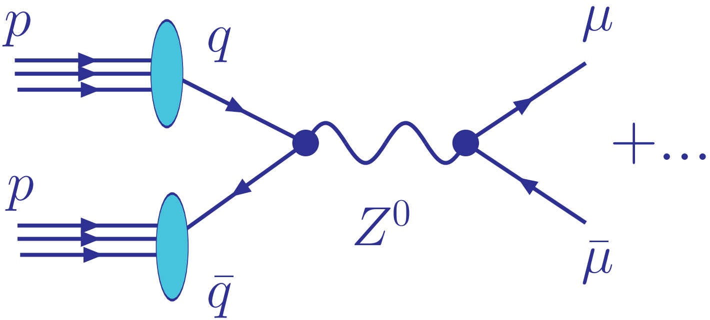


PDFs and QCD factorization



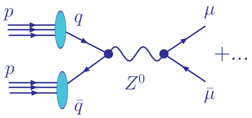
Drell-Yan process $pp \rightarrow (Z^0 \rightarrow \ell\bar{\ell})X$ at the LHC ($\ell\bar{\ell} = e\bar{e}$ or $\mu\bar{\mu}$)

PDFs and QCD factorization



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

PDFs and QCD factorization

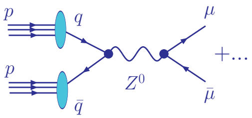


According to QCD factorization theorems, typical cross sections (e.g., for $p(k_1)p(k_2) \to \left[Z(q) \to \ell(k_3)\bar{\ell}(k_4)\right] X$) take the form

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

- $\widehat{\sigma}_{ab\to Z\to\ell\bar{\ell}}$ 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

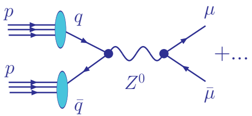


According to QCD factorization theorems, typical cross sections (e.g., for $p(k_1)p(k_2) \to \left[Z(q) \to \ell(k_3)\bar{\ell}(k_4)\right] X$) take the form

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

- 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:

$$f_{q/p}(x,\mu) = \frac{1}{4\pi}\int_{-\infty}^{\infty} dy^- e^{iy^- p^+} \langle p \left| \overline{\psi}_q(0,y^-,\vec{0}_T)\gamma^+\psi_q(0,0,\vec{0}_T) \right| p\rangle, \text{ etc.}$$

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:

$$f_{q/p}(x,\mu) = \frac{1}{4\pi}\int_{-\infty}^{\infty} dy^- e^{iy^-p^+} \langle p \left| \overline{\psi}_q(0,y^-,\vec{0}_T)\gamma^+\psi_q(0,0,\vec{0}_T) \right| p\rangle, \text{ etc.}$$

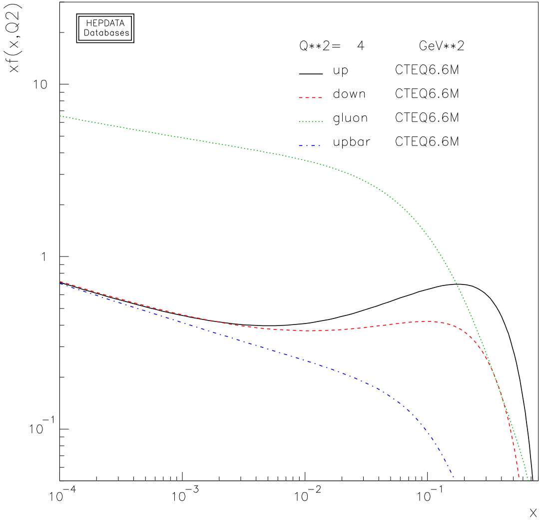
The exact form of $f_{a/p}$ is not known; but its μ dependence is described by **D**okshitzer-**G**ribov-**L**ipatov-**A**ltarelli-**P**arisi (**DGLAP**) equations

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

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

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

Example of DGLAP evolution



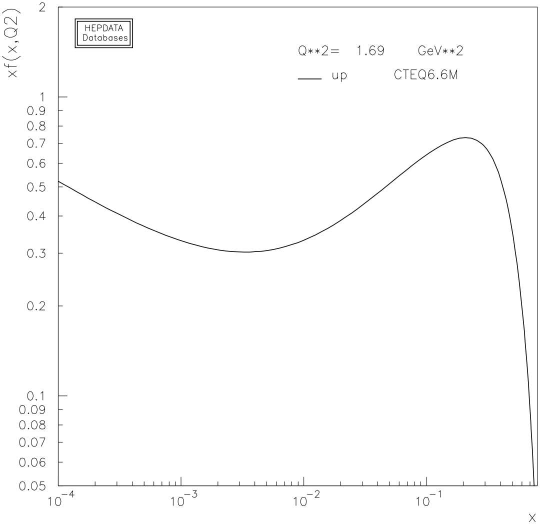
Durham PDF plotter, http://durpdg.dur.ac.uk/hepdata/pdf3.html

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

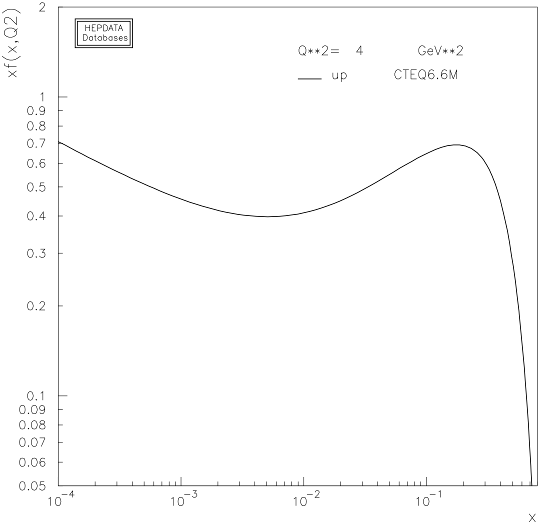


Durham PDF plotter, http://durpdg.dur.ac.uk/hepdata/pdf3.html

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 $x \lesssim 0.1$

Example of DGLAP evolution

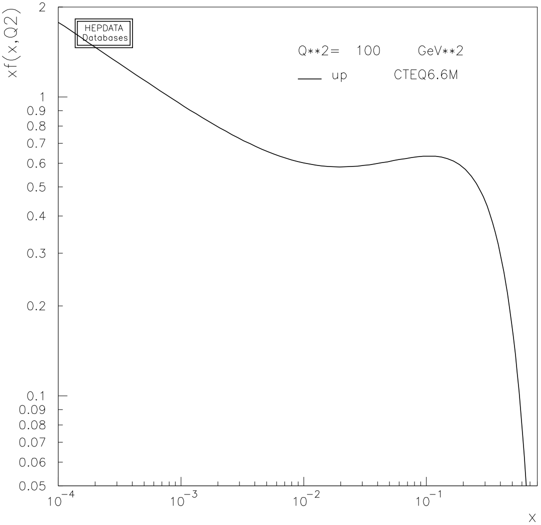


Durham PDF plotter, http://durpdg.dur.ac.uk/hepdata/pdf3.html

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 $x \lesssim 0.1$

Example of DGLAP evolution

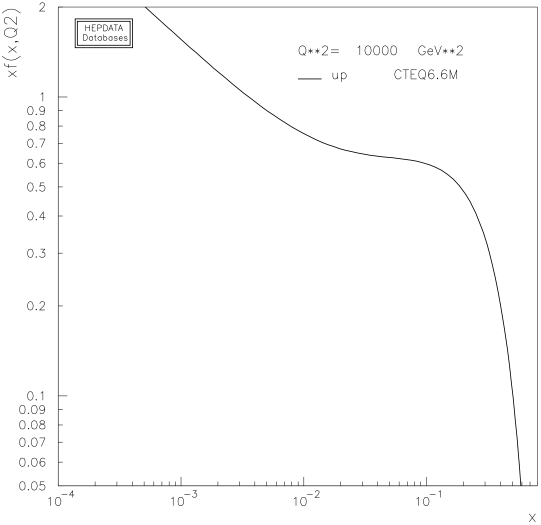


Durham PDF plotter, http://durpdg.dur.ac.uk/hepdata/pdf3.html

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 $x \lesssim 0.1$

Example of DGLAP evolution

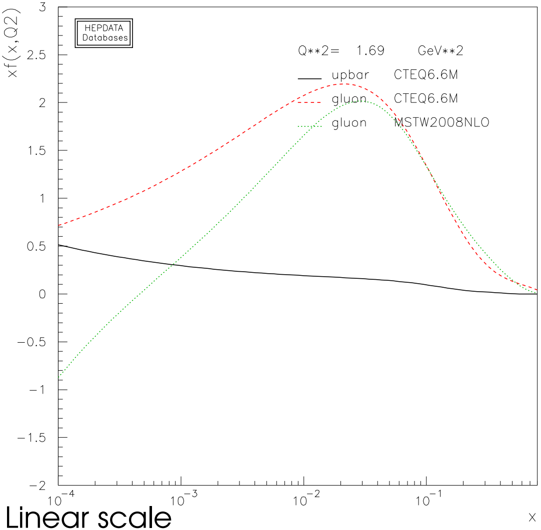


Durham PDF plotter, http://durpdg.dur.ac.uk/hepdata/pdf3.html

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 $x \lesssim 0.1$

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



Linear scale

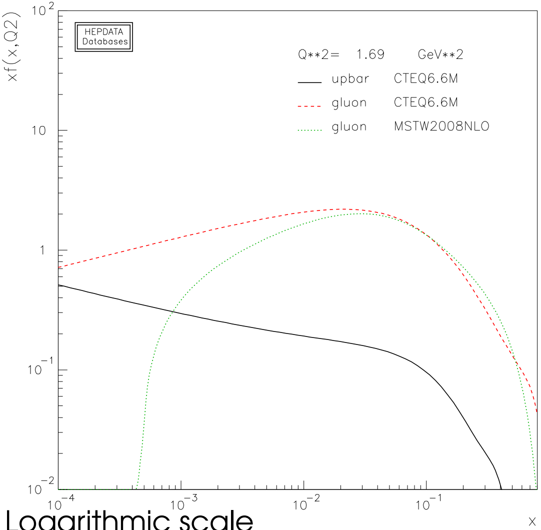
$g(x,Q)$ can become negative at $x < 10^{-2}$, $Q < 2$ GeV

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

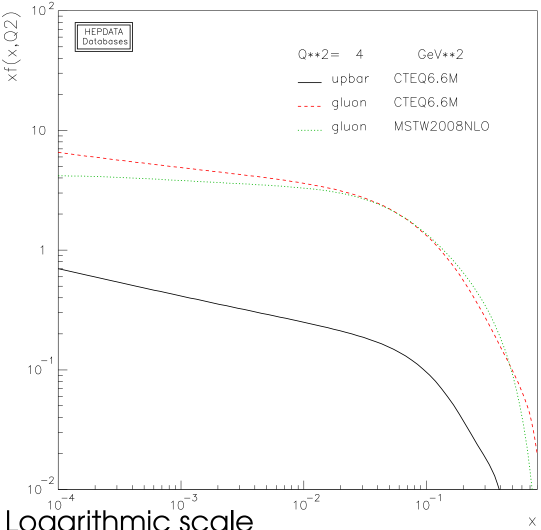
$g(x,Q)$ can become negative at $x < 10^{-2}$, $Q < 2$ GeV

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



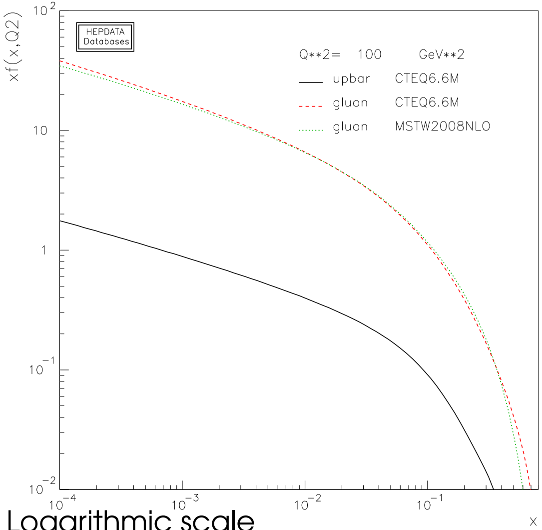
Logarithmic scale

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



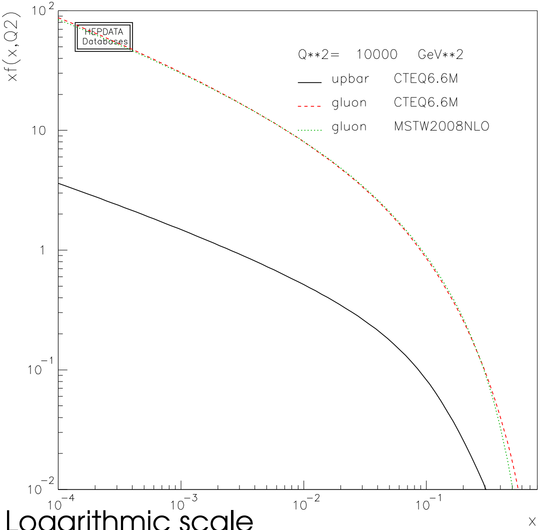
Logarithmic scale

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:

$$f_{q/p}(x,\mu) = \frac{1}{4\pi}\int_{-\infty}^{\infty} dy^- e^{iy^- p^+} \langle p \left| \overline{\psi}_q(0, y^-, \vec{0}_T)\gamma^+ \psi_q(0,0,\vec{0}_T) \right| p\rangle, \text{ etc.}$$

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 $\mu > Q_0$

Factorized QCD predictions

Lepton-hadron scattering

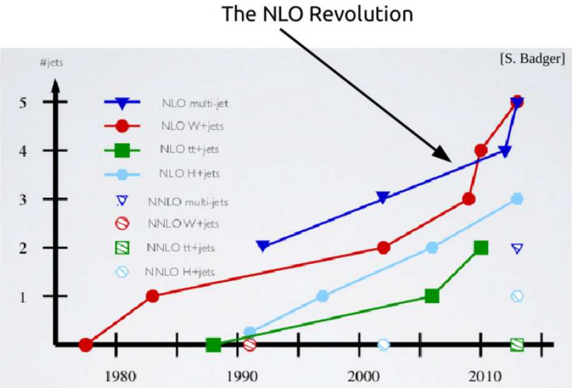
$$\sigma = \sum_a \widehat{\sigma}_a \otimes f_{a/p}$$

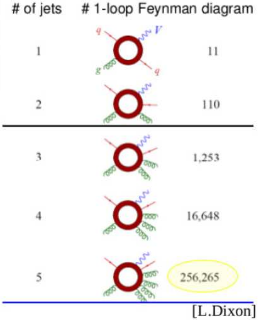
Hadron-hadron scattering

$$\sigma = \sum_{a_1,a_2} \widehat{\sigma}_{a_1 a_2} \otimes f_{a_1/p_1} \otimes f_{a_2/p_2}$$

The accuracy in determination of PDFs $f_{a/p}$ must match the accuracy of hard cross sections $\widehat{\sigma}$

Perturbative QCD loop revolution



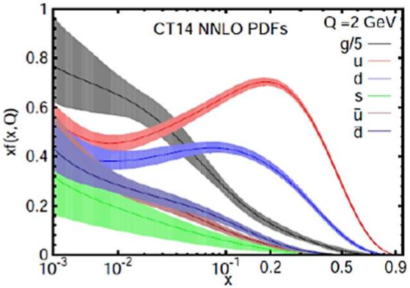


Since 2005, generalized unitarity and related methods dramatically advanced the computations of **perturbative** NLO/NNLO/N3LO hard cross sections $\hat{\sigma}$.

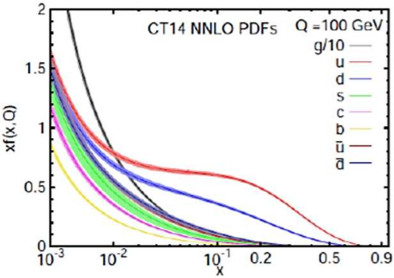
To make use of it, accuracy of PDFs $f_{a/p}(x, \mu)$ must keep up

General-purpose CT14 PDFs

(S. Dulat et al., arXiv:1506.07443)



Q= 2 GeV



Q= 100 GeV

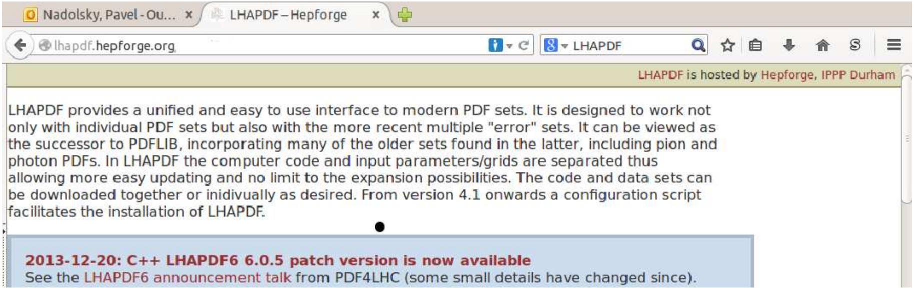
Phenomenological parametrizations of PDFs are provided with estimated uncertainties of multiple origins (**uncertainties of measurement, theoretical model, parametrization form, statistical analysis**, …)

The shape of PDFs is optimized w.r.t. hundreds of **nuisance parameters**

Where do the PDFs come from?

Practical answer: from the Les Houches Accord PDF library (LHAPDF)

Almost all recent PDFs are included in the LHAPDF C++ library available at lhapdf.hepforge.org.


Nadolsky, Pavel - Ou... LHAPDF – Hepforge
lhapdf.hepforge.org LHAPDF

LHAPDF is hosted by Hepforge, IPPP Durham

LHAPDF provides a unified and easy to use interface to modern PDF sets. It is designed to work not only with individual PDF sets but also with the more recent multiple "error" sets. It can be viewed as the successor to PDFLIB, incorporating many of the older sets found in the latter, including pion and photon PDFs. In LHAPDF the computer code and input parameters/grids are separated thus allowing more easy updating and no limit to the expansion possibilities. The code and data sets can be downloaded together or inidivually as desired. From version 4.1 onwards a configuration script facilitates the installation of LHAPDF.

2013-12-20: C++ LHAPDF6 6.0.5 patch version is now available
See the LHAPDF6 announcement talk from PDF4LHC (some small details have changed since).

Thousands of PDF sets are provided and can be linked to your computer code. Which one should you use?

Where do the PDFs come from?

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

LHC
Tevatron

HERA
RHIC
EIC

Fixed-target experiments

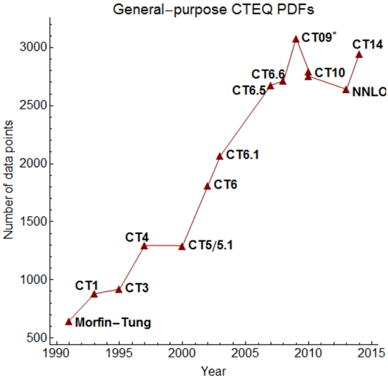
+ lattice QCD

Coordinated Theoretical-Experimental Analysis of QCD (CTEQ)

Several groups in CTEQ work on determination of PDFs:
CTEQ-Tung Et Al. (CT), CTEQ-JLab (CJ),...

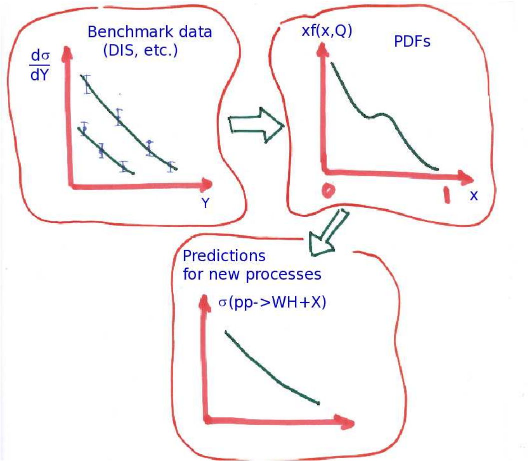
Global analysis *(term promoted by J. Morfin & W.-K. Tung in 1990):*

constrains PDFs or other nonperturbative functions with data from diverse hadronic experiments



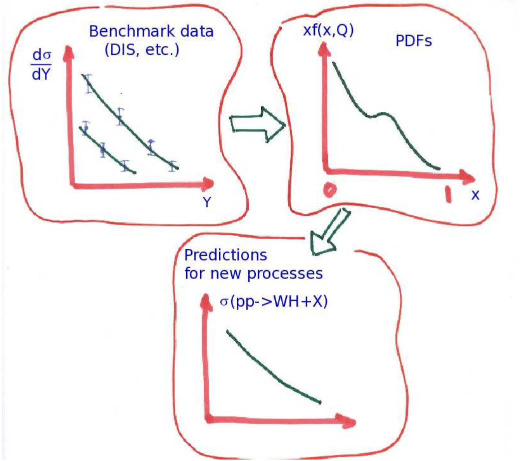
Morfin–Tung:	DIS, low–Q Drell–Yan process
CT1:	DIS, DY, direct γ
CT3:	low–x DIS, W charge asymmetry
CT4:	– direct γ; + CDF high–p_T jets
CT5/5.1:	$+\sigma_{pp}/\sigma_{pd}$ in DY process; D0 jets
CT6:	Error PDFs; correlated syst. errors
CT6.1:	Tevatron Run–1b jets
CT6.5:	GM–VFN scheme; free s(x)
CT6.6:	PDF correlations
CT09*:	+Tevatron Run–2 jets
CT10:	+NNLO, combined HERA, Run–2 W asymmetry
CT14:	+LHC Run–1 W, Z, jet production

The flow of the global 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 global analysis

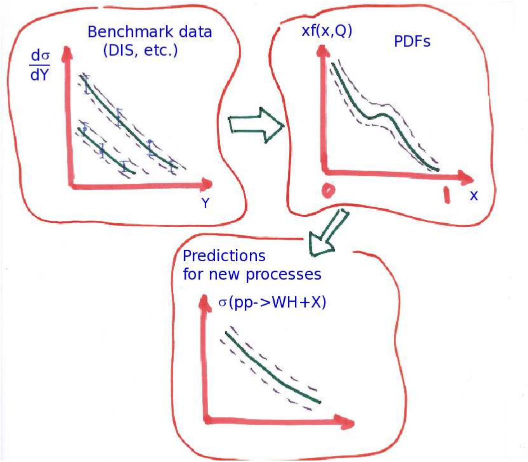


Data sets in the CT10 analysis

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**

Modern fits involve up to 40 experiments, $3000+$ data points, and 100+ free parameters

The flow of the global 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

A question to you (think for 1 minute)

Among Standard Model particles, which particles can have a non-zero PDF?

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 QCD+QED PDFs by CT, LUX, MRST, NNPDF... groups)
  - a QCD+QED fit is more complicated than one might think: it must account for violation of charge symmetry by EM effects,

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

PDFs for heavy flavors

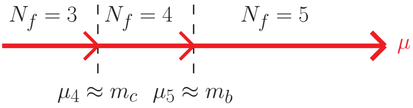
PDFs for heavy partons h can be generated via DGLAP evolution at $Q \geq m$, using a boundary condition $f_{h/p}(x, Q) = 0$ at $Q \leq m$

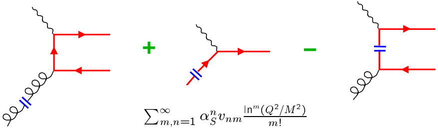
In practice:

- PDFs are usually introduced for c and b quarks
  - starting from $\mathcal{O}(\alpha_s^2)$, an initial condition $f_{c/p}(x, Q_0) \neq 0$ can be generated at $Q_0 = m_c$ from twist-4 intrinsic charm DIS terms *(arXiv:1707.00657)*
- 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

General-mass variable-flavor number scheme

- A series of factorization schemes with N_f active quark flavors in $\alpha_s(Q)$ and $f_{a/p}(x, Q)$
  - N_f is incremented sequentially at momentum scales $\mu_{N_f} \approx m_{N_f}$
- incorporates essential $m_{c,b}$ dependence near, and away from, heavy-flavor thresholds
- implemented in all latest PDF fits except ABM



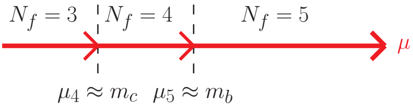


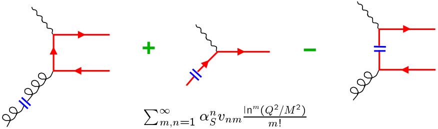
General-mass variable-flavor number scheme

Proved for *inclusive DIS* by J. Collins (1998)

$$F_2(x, Q, m_c) = \sum_a \int_\chi^1 \frac{d\xi}{\xi} C_a(\frac{\chi}{\xi}, \frac{Q}{\mu}, \frac{m_c}{Q}) f_a(\xi, \frac{\mu}{m_c}) + \mathcal{O}\left(\frac{\Lambda_{QCD}}{Q}\right)$$

- $\lim_{Q\to\infty} C$ exists and is infrared safe
- collinear logarithms $\sum_{k,n=1}^{\infty} \alpha_s^k v_{kn} \ln^n(\mu/m_c)$ are resummed in $f_c(x, \mu/m_c)$
- no terms $\mathcal{O}(m_c/Q)$ in the remainder





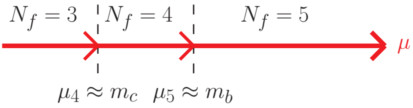
$\sum_{m,n=1}^{\infty} \alpha_S^n v_{nm} \frac{\ln^m(Q^2/M^2)}{m!}$

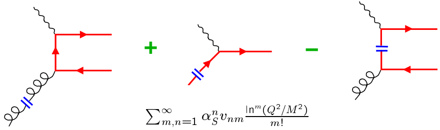
General-mass variable-flavor number scheme

Proved for *inclusive DIS* by J. Collins (1998)

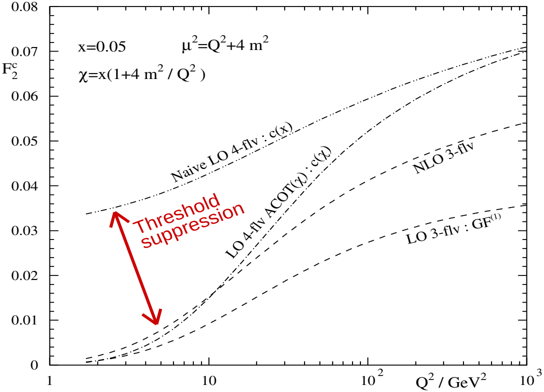
$$F_2(x, Q, m_c) = \sum_a \int_\chi^1 \frac{d\xi}{\xi} C_a\left(\frac{\chi}{\xi}, \frac{Q}{\mu}, \frac{m_c}{Q}\right) f_a\left(\xi, \frac{\mu}{m_c}\right) + \mathcal{O}\left(\frac{\Lambda_{QCD}}{Q}\right)$$

- Works most effectively in DIS and Drell-Yan-like processes; practical implementation requires
  1. efficient treatment of mass dependence, rescaling of momentum fractions χ in processes with incoming c, b
  2. physically motivated factorization scale to ensure fast PQCD convergence (e.g., $\mu = Q$ in DIS)





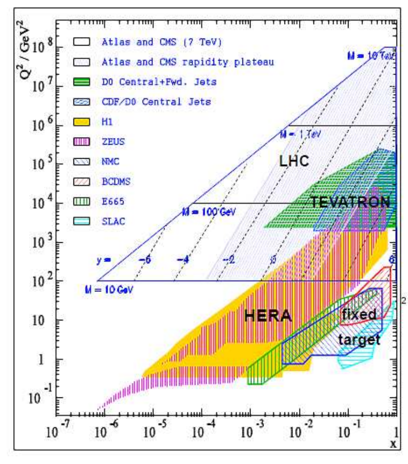
An example of GM-VFN factorization scheme



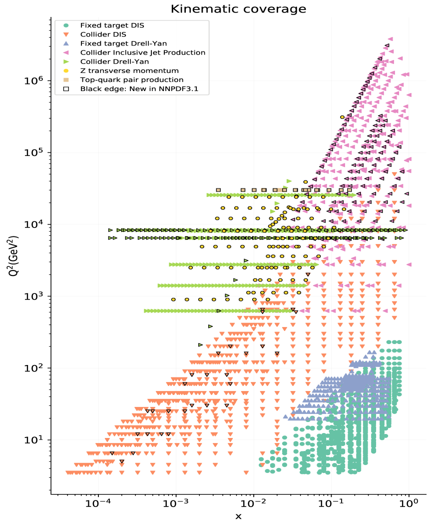
- Charm Wilson coefficient function is suppressed at $Q \to m_c$
- To keep agreement with F_2 data, u, d, $\bar{u}$, $\bar{d}$ PDF's are enhanced at small x, as compared to the zero-mass (ZM-VFN) scheme

2. Experimental observables constraining the PDFs in global fits

x, Q coverage of various experiments

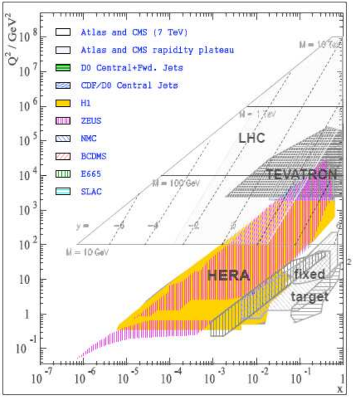


Experiments included in the NNPDF3.1 PDF analysis



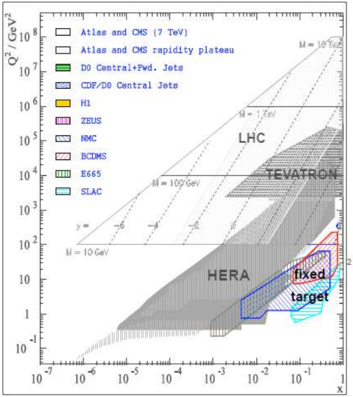
Selection of experimental data for PDF fits (2017)

- **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



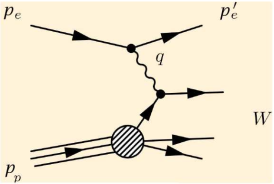
Selection of experimental data for PDF fits (2017)

- **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
    - $eN, \mu N, \nu N$ 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



$$\frac{d^2\sigma(e^\pm p)}{dQ^2 dx} = \frac{2\pi\alpha^2}{Q^4 x} Y_+ \left(F_2 - \frac{y^2}{Y_+} F_L \pm \frac{Y_-}{Y_+} x F_3 \right),$$

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

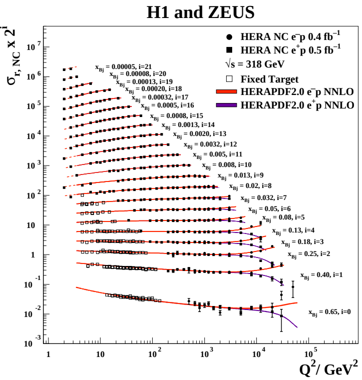
The data is fitted either in the form of $F_2(x, Q^2)$ or $d^2\sigma/(dQ^2 dx)$

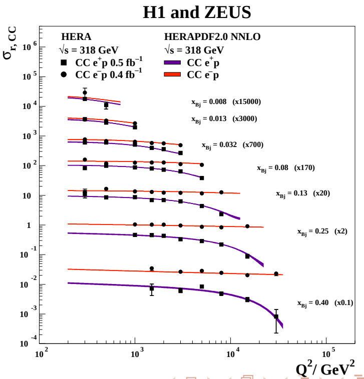
Final combined DIS cross sections at HERA

(arXiv:1506.06042)

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 :

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

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

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,

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

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 $x < 0.1$

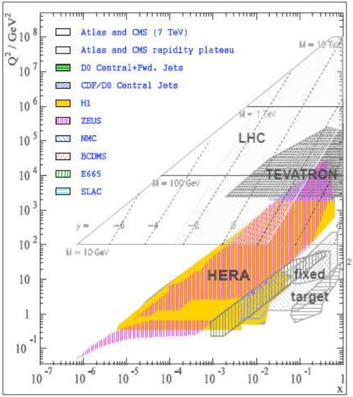
DIS cannot fully separate quarks from antiquarks, or s, c, b contributions from u and d contributions; more so because of systematic effects in fixed-target DIS experiments (higher-order terms, nuclear corrections,...)

Selection of experimental data for PDF fits (2017)

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, NOMAD, ...)



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

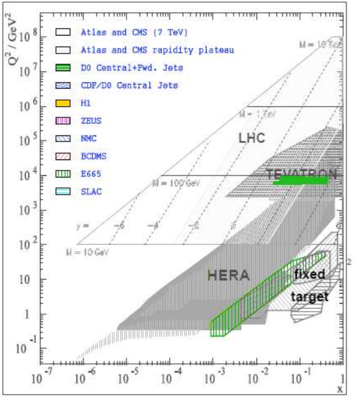
Selection of experimental data for PDF fits (2017)

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, NOMAD, ...)

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



Selection of experimental data for PDF fits (2017)

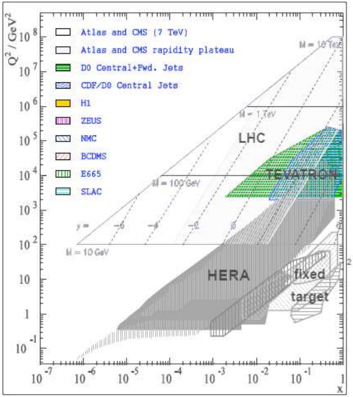
The modern PDF fits include
Inclusive deep-inelastic scattering...

+ Semi-inclusive DIS:

- charm production $ep \to ecX$ (HERA)
- $\mu\mu$ production $\nu N \to \mu(c \to \mu)X$ (NuTeV, NOMAD, ...)

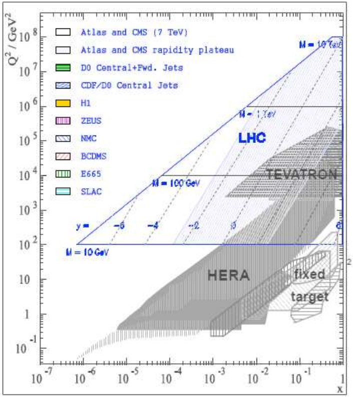
+ Lepton pair production $pN \xrightarrow{\gamma^*, W, Z} \ell\bar{\ell}' X$ (Tevatron, fixed-target experiments)

+ 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, jet, $t\bar{t}$ production

Selection of experimental data for PDF fits (2017)



Dozens of data sets from LHC!

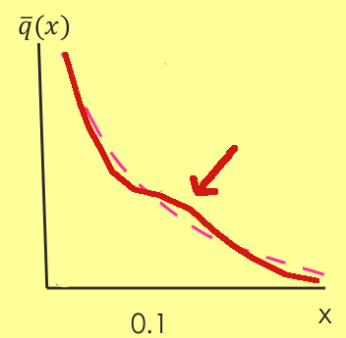
- CT14, MMHT14, NNPDF3.0 include early 7 TeV W/Z, jet production data sets
- NNPDF3.1 *(arXiv:1706.00428)* includes high-luminosity data on W/Z, Z p_T, $t\bar{t}$ production
  - moderate reduction in PDF uncertainty, especially for $g(x,Q)$

Zooming in on quark sea PDFs

Various QCD effects produce non-trivial sea PDFs

- breaking of SU(2) symmetry and charge symmetry
- 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

SU(2) and charge symmetry breaking

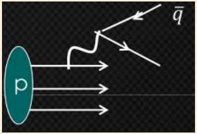
$$\bar{d}(x) \neq \bar{u}(x), \quad \bar{q}(x) \neq q(x)$$

May be cause by

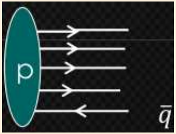
- DGLAP evolution
- Fermi motion
- Electromagnetic effects
- Nonperturbative meson fluctuations
- Chiral symmetry breaking
- Instantons
- ...

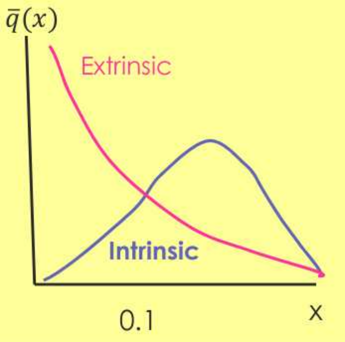
Extrinsic and intrinsic sea PDFs

"Extrinsic" sea



"Intrinsic" sea





Extrinsic and intrinsic sea PDFs

Smooth $\bar{u}+\bar{d}$ parametrizations can hide existence of two components

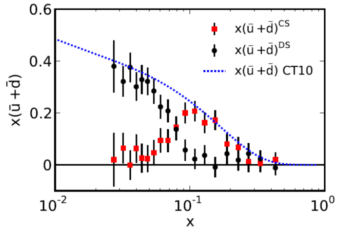
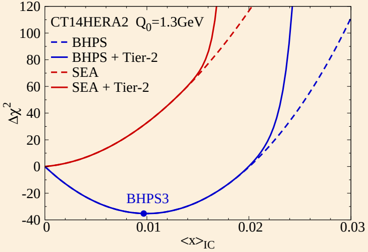


FIG. 5: $x(\bar{u}^{cs}(x)+\bar{d}^{cs}(x))$ obtained from Eq. (1) is plotted together with $x(\bar{u}(x)+\bar{d}(x))$ from CT10 and $\frac{1}{R}x(s(x)+\bar{s}(x))$ which is taken to be $x(\bar{u}^{ds}(x)+\bar{d}^{ds}(x))$.

Liu, Chang, Cheng, Peng, 1206.4339

Intrinsic charm (IC) can carry up to 1% of the proton momentum



CT14 IC NNLO PDFs

T.-J. Hou et al., 1707.00657

Constraints on quark sea from $pN \to \ell^+\ell^- X$ ($N = p,\, d,\, Fe,\, Cu, ..$)

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

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

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

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

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 $pN \to \ell^+\ell^- X$ ($N = p,\, d,\, Fe,\, Cu,\, ..$)

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

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

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

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

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 $pN \to \ell^+\ell^- X$ ($N = p,\, d,\, Fe,\, Cu, ..$)

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

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

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

$\frac{d\sigma_{pn}}{dQ^2 dy} \sim (\frac{2}{3})^2 \left[u_A \bar{d}_B + \bar{u}_A d_B\right] + (-\frac{1}{3})^2 \left[d_A \bar{u}_B + \bar{d}_A u_B\right] +$ 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 $pN \to \ell^+\ell^- X$ ($N = p,\, d,\, Fe,\, Cu, ..$)

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

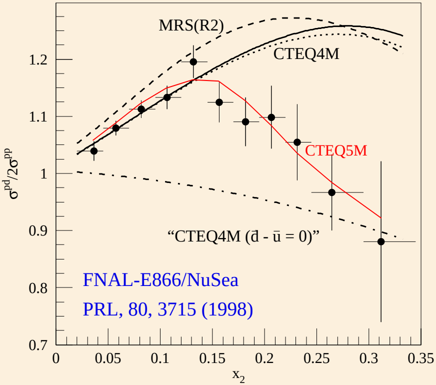
$\therefore \sigma_{pd}/(2\sigma_{pp})$ constrains $\bar{d}(x,Q)/\bar{u}(x,Q)$ at moderate x

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)

$\sigma_{pd}/(2\sigma_{pp})$ 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

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

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

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

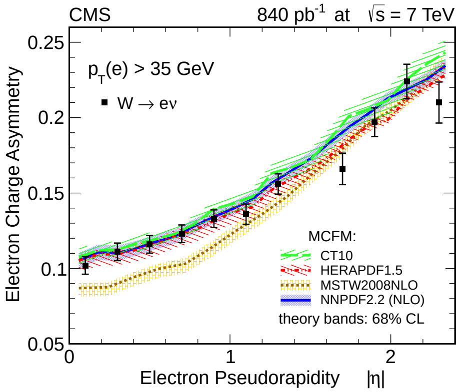
y_e and $\eta \approx y_e$ are rapidity and pseudorapidity of an electron from W decay

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

$\therefore A_{ch}(y_e)$ 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)

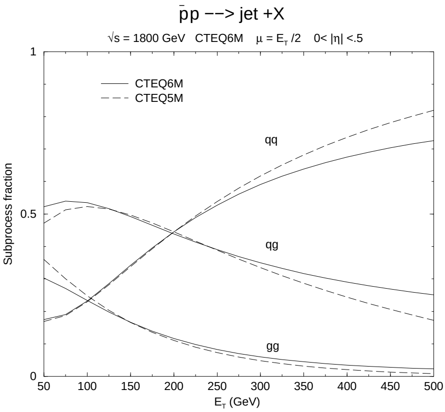
Charge asymmetry at the Tevatron and LHC



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

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

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

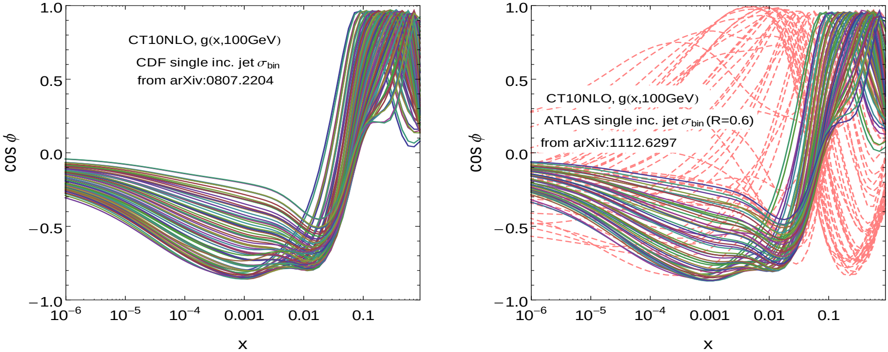


High-E_T jets are mostly produced in qq scattering; yet most of the PDF uncertainty arises from qg and gg contributions

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 $g(x,Q)$

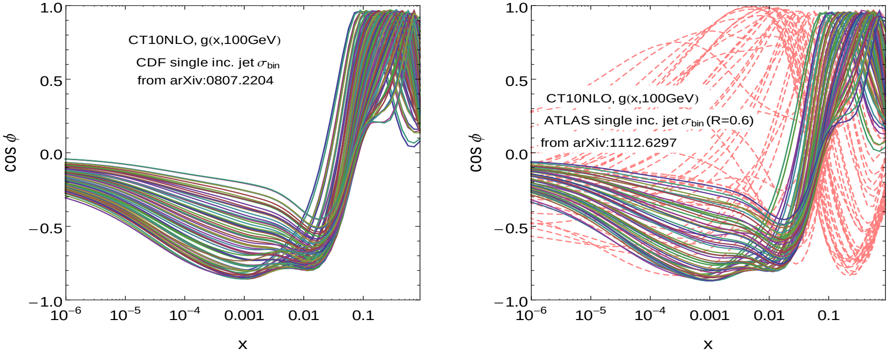
Sensitivity of $p\overset{(-)}{p} \to$ jet $+ X$ to gluon PDF



$\cos\phi(x)$ measures sensitivity of the data to $g(x,Q)$ at various x (see Lecture 2).

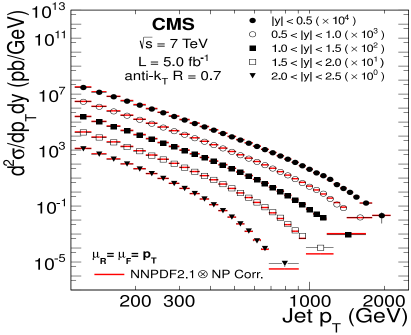
The CDF (left) and ATLAS (right) jet production are sensitive to $g(x,Q)$ in the x regions with $\cos\phi \gtrsim 0.7$ and $\cos\phi \lesssim -0.7$

Sensitivity of $p\overset{(-)}{p} \to$ jet $+ X$ to gluon PDF



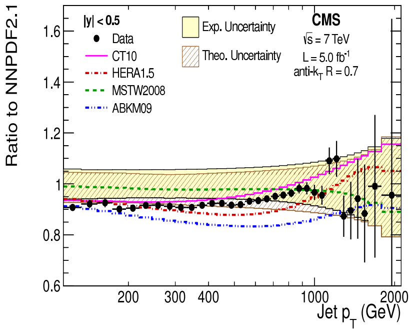
The curves show $\cos\varphi$ between the NLO theory cross sections in experimental p_T^j bins and $g(x,Q)$, for various x values in $g(x,Q)$. The CDF (ATLAS) jet measurement are mostly sensitive to the gluon PDF with $x \gtrsim 0.1\ (0.01)$. In the ATLAS jet data, the bins with $p_T^j < 200$ GeV (pink dashed curves) probe $g(x,Q)$ in a wider range than CDF.

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



- The cross sections span 12 orders of magnitude
- (Almost) negligible statistical error

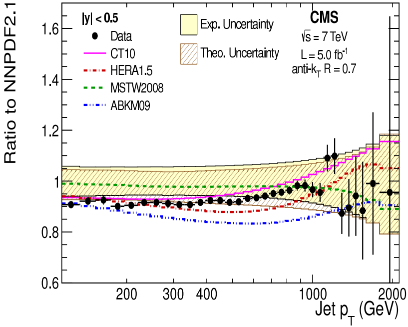
Inclusive jet production in $pp \to$ jet $+X$ (7 TeV)



- The cross sections span 12 orders of magnitude
- (Almost) negligible statistical error

- 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

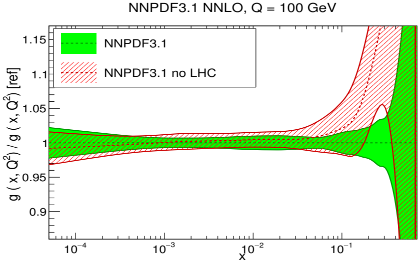
Inclusive jet production in $pp \to$ jet $+X$ (7 TeV)

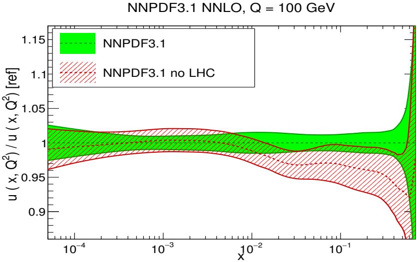


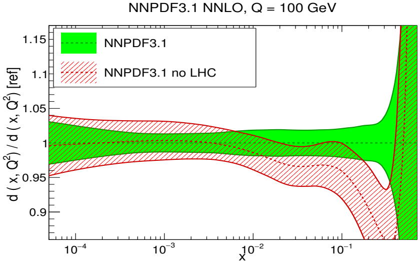
- The cross sections span 12 orders of magnitude
- (Almost) negligible statistical error

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

Effect of the LHC data on NNPDF3.1 PDFs



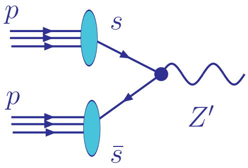




Note: NNPDF3.1 (no LHC) does not include D0 Run-2 jet data, has softer gluon

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?

3. Choice of PDF parametrization

4. Statistical aspects

5. Practical applications