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 2
July 2014

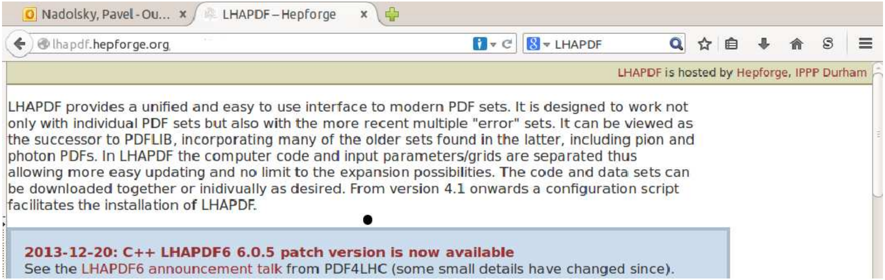
Recap, lecture 1

In the first lecture, we discussed the basic properties and behavior of parton distribution functions, motivation for precision studies of PDFs, and key experiments from which the PDFs are determined

Today we will review practical issues associated with the determination and usage of PDFs: differences between the PDF sets, computation of PDF uncertainties, implementation of correlated errors, combination of PDF sets, and dependence on the QCD coupling.

Les Houches Accord PDF library (LHAPDF)

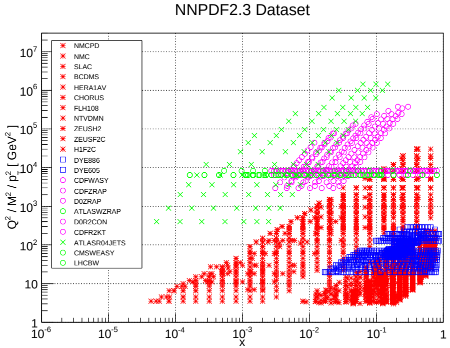
Almost all recent PDFs are included in the LHAPDF C++ library available at lhapdf.hepforge.org.



Hundreds of PDF sets are provided. Which one should you use?

Parton distribution functions in a nucleon

At NNLO QCD, general-purpose PDF parametrizations are available from ABM, CT, HERA, MSTW, NNPDF groups

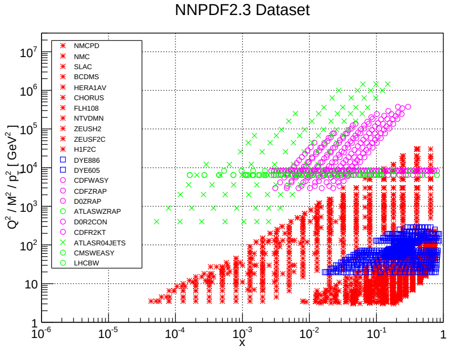


Typical PDFs are constrained by

- DIS at HERA
- Vector boson production at low Q, Tevatron
- Inclusive jet production
- early LHC data

Parton distribution functions in a nucleon

Fixed-target data sets are critical at $x > 0.01$, but may be replaced in the future by collider measurements in $t\bar{t}$, direct-γ,Wc, ... production



Any PDF set makes assumptions about poorly constrained PDF combinations, e.g., sea PDFs at $x < 0.01$ and $x > 0.3$.Photon PDF is largely unknown.

Our latest PDFs: CT10 and CT1X NNLO

- CT10 NNLO [arXiv:1302.6246] is an NNLO counterpart either to CT10 NLO or CT10W NNLO
  - In good agreement with early LHC data
- CT1X NNLO – a preliminary extension of CT10 NNLO that includes latest HERA data on $F_L(x,Q)$ and $F_2(x,Q)$, LHC 7 TeV data (ATLAS W & Z, ATLAS jets, CMS W asymmetry)
- The new data provide only minor improvements compared to the CT10 data set. We investigate its agreement with the CT10 data sets and await for more precise LHC data to be included in the CT1X public release

CT10 NNLO PDFs

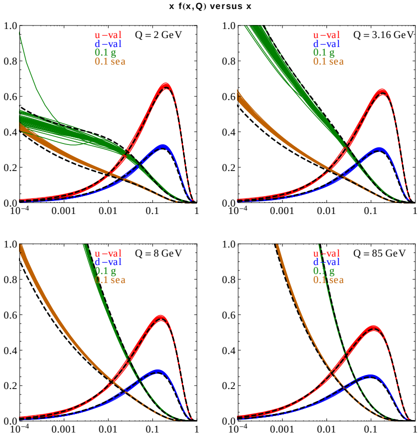
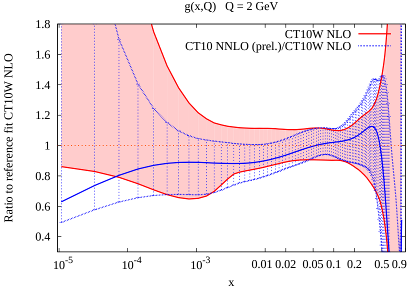
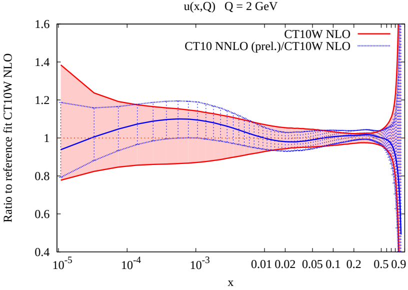
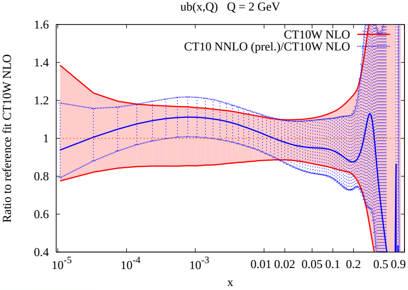


FIG. 2: CT10NNLO parton distribution functions. These figures show the Hessian error PDFs from the CT10NNLO analysis. Each graph shows $x\,u_{\rm valence} = x(u - \overline{u})$, $x\,d_{\rm valence} = x(d - \overline{d})$, $0.10\,x\,g$ and $0.10\,x\,q_{\rm sea}$ as functions of x for a fixed value of 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.

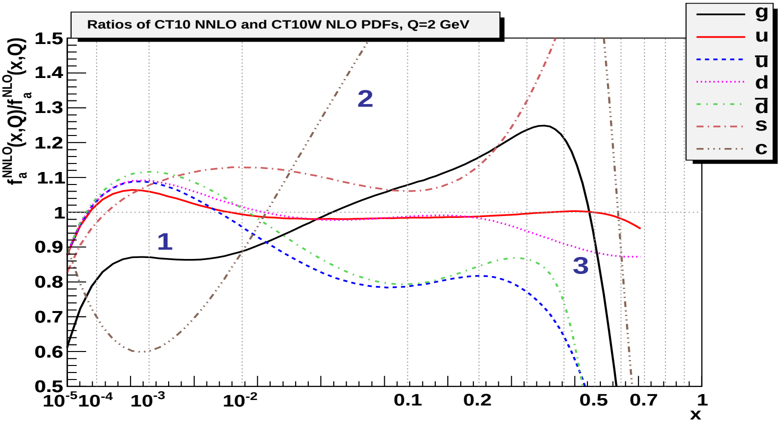
CT10 NNLO error PDFs (compared to CT10W NLO)





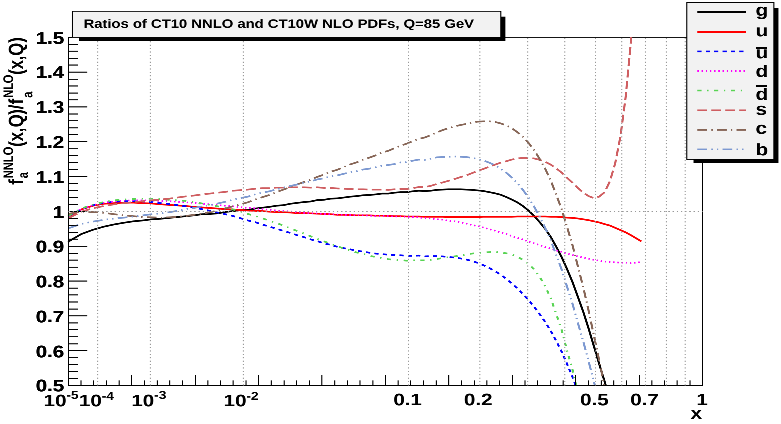


CT10 NNLO central PDFs, as ratios to NLO, Q=2 GeV

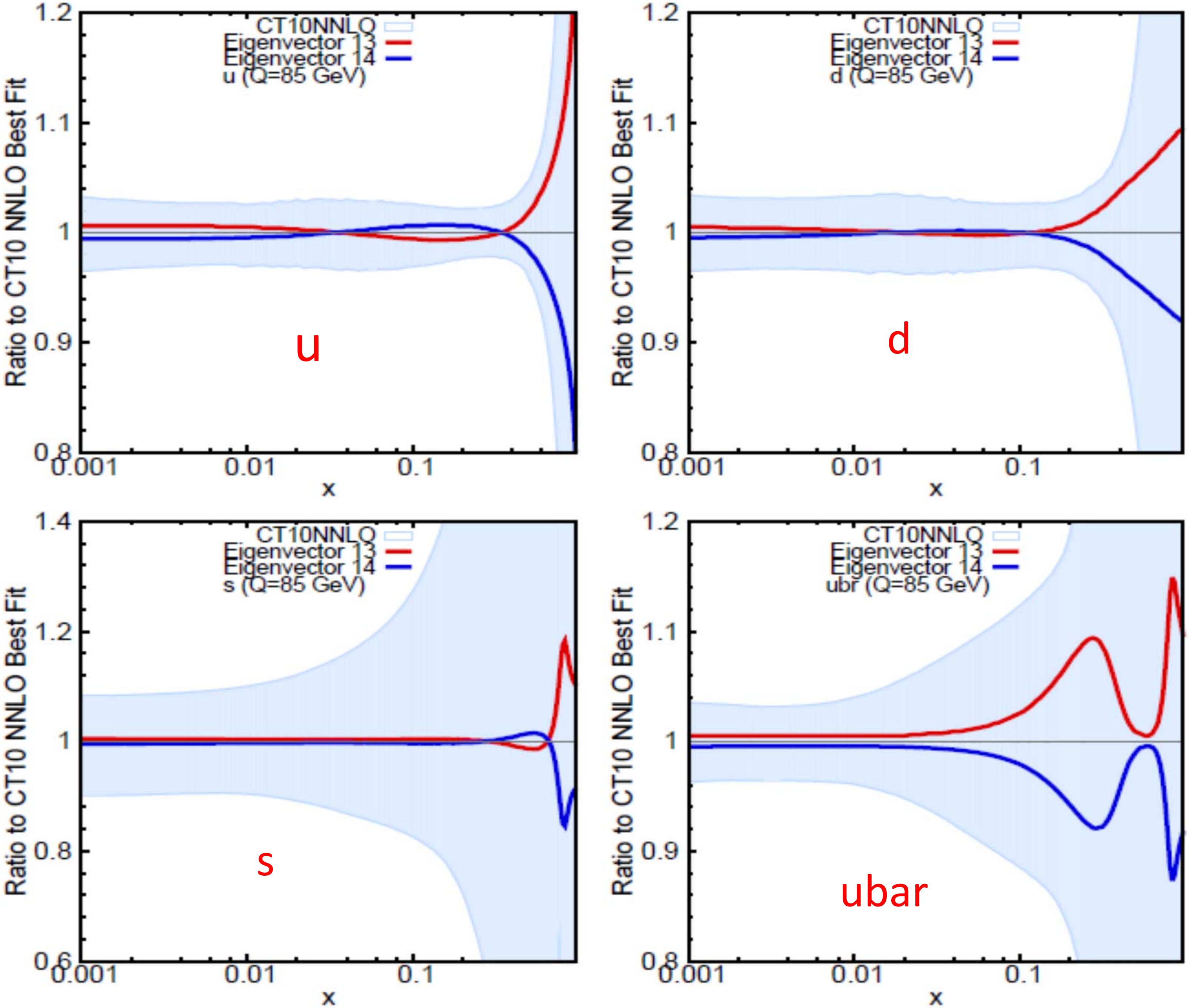


1. At $x < 10^{-2}$, $\mathcal{O}(\alpha_s^2)$ evolution suppresses $g(x,Q)$, increases $q(x,Q)$
2. $c(x,Q)$ and $b(x,Q)$ change as a result of the $\mathcal{O}(\alpha_s^2)$ GM VFN scheme
3. At $x > 0.1$, $g(x,Q)$ and $d(x,Q)$ are reduced by revised EW couplings, alternative treatment of correlated systematic errors, scale choices

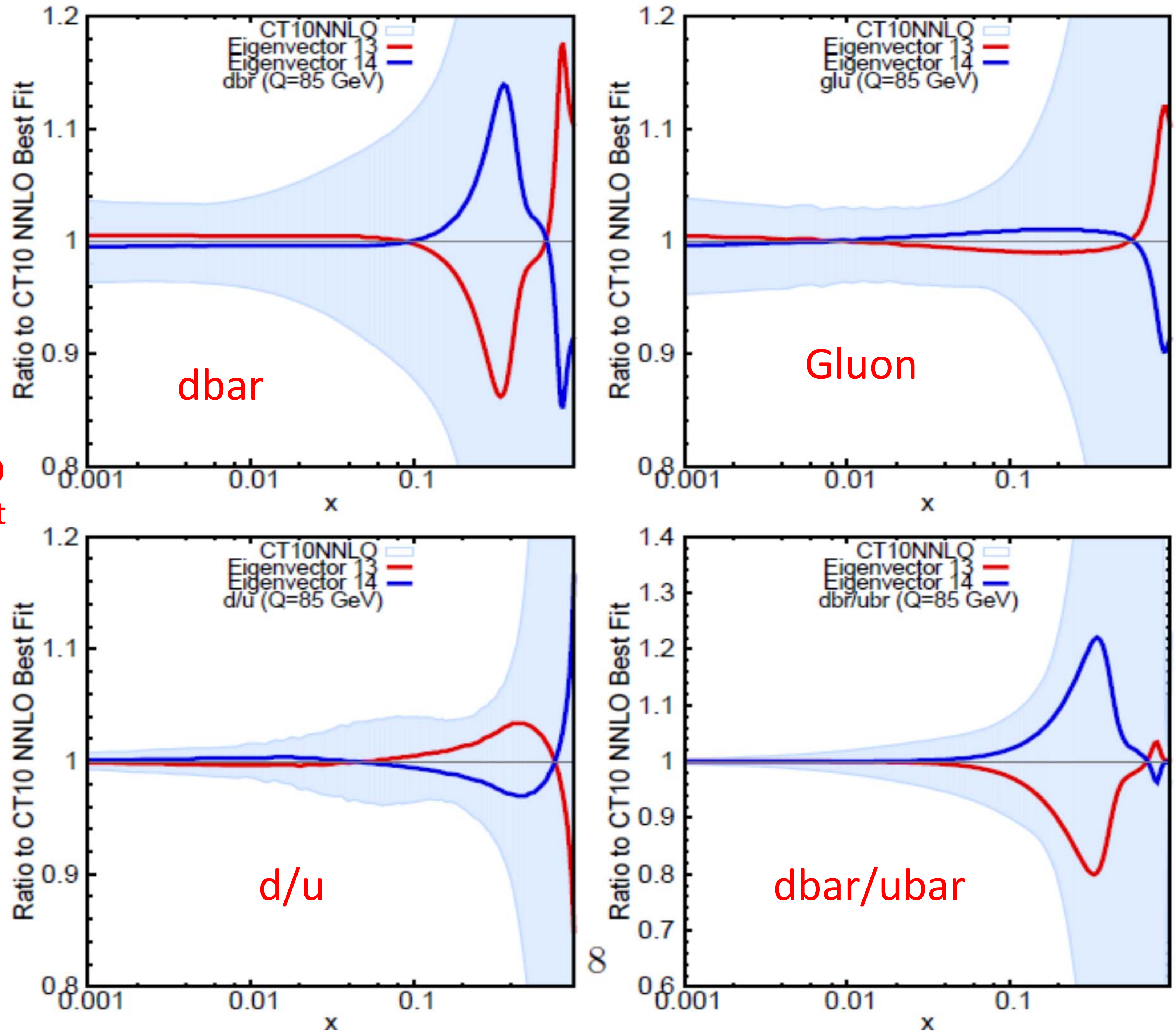
CT10 NNLO central PDFs, as ratios to NLO, Q=85 GeV



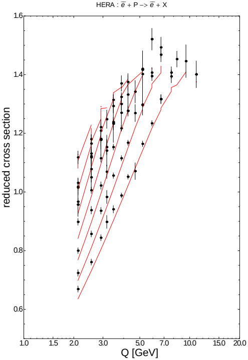
7th CT10 Eigenset

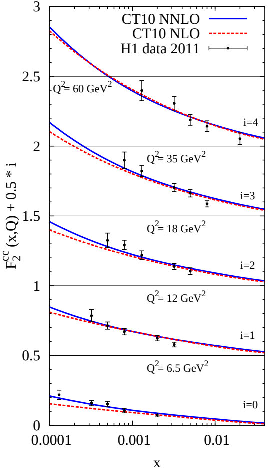


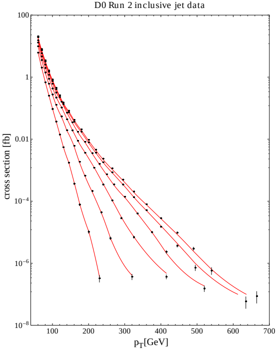
7th CT10 Eigenset



CT10 NNLO vs. fitted experiments

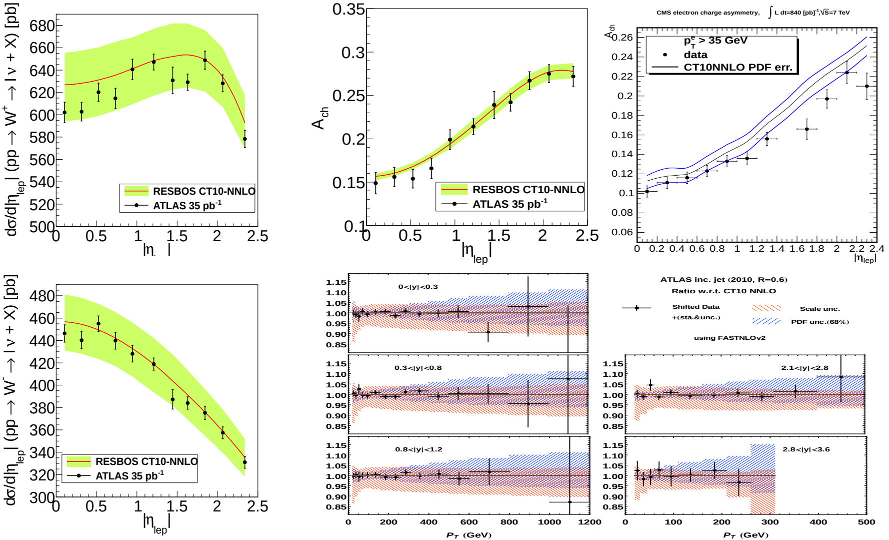






Good general agreement

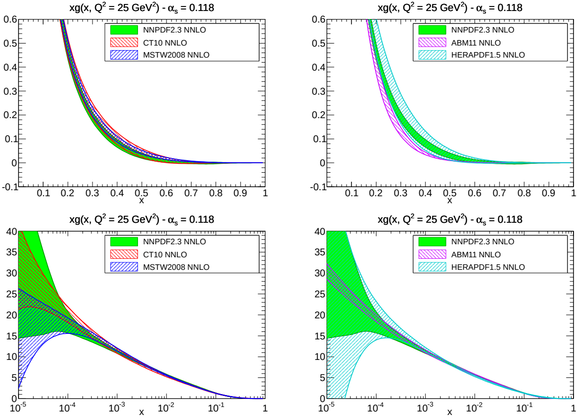
CT10 NNLO describes well LHC 7 TeV experiments



NNLO gluon PDF $xg(x,Q)$ from 5 groups

Linear x scale

Logarithmic x scale



R. Ball, et al., 1211.5142

CT10, MSTW'08, NNPDF2.3 PDFs are in good general agreement. We see some differences with HERAPDF and ABM sets. Where do these differences come from?

Origin of differences between PDF sets

1. Corrections of wrong or outdated assumptions

lead to significant differences between new (≈post-2010) and old (≈pre-2010) PDF sets

- inclusion of NNLO QCD, heavy-quark hard scattering contributions
  - the latest PDFs implement complete heavy-quark treatment; previous PDFs are obsolete without it
- relaxation of ad hoc constraints on PDF parametrizations
- improved numerical approximations

Origin of differences between PDF sets

2. PDF uncertainty

a range of allowed PDF shapes for plausible input assumptions, **partly** reflected by the PDF error band

is associated with

- the choice of fitted experiments
- experimental errors propagated into PDF's
- handling of inconsistencies between experiments
- choice of factorization scales, parametrizations for PDF's, higher-twist terms, nuclear effects,...

leads to non-negligible differences between the newest PDF sets

Stages of the PDF analysis

1. Select experimental data
2. Assemble all relevant theoretical cross sections and verify their mutual consistency
3. Choose the functional form for PDF parametrizations
4. Perform a fit
5. Make the new PDFs and their uncertainties available to end users

1. Selection of experimental data

- **Neutral-current ep DIS data from HERA are the most extensive and precise** among all data sets
  - In addition, their systematic errors were reduced recently by cross calibration of H1 and ZEUS detectors
- However, by their nature they constrain only a limited number of PDF parameters
- Thus, two complementary approaches to the selection of the data are possible

Two strategies for selection of experimental data

DIS-based analyses $\Rightarrow$ focus on the most precise (HERA DIS) data

- NC DIS, CC DIS, NC DIS jet, c and b production *(ABM; HERAPDF1.0,1.5, 2,0; JR)*

Global analyses *(**CT10**, MSTW'2008, NNPDF)* $\Rightarrow$ focus on completeness, reliable flavor decomposition

- all HERA data + fixed-target DIS data
  - notably, CCFR and NuTeV νN DIS constraining $s(x,Q)$
- low-Q Drell-Yan (E605, E866), W lepton asymmetry, Z rapidity *(CT10, MSTW'08, NNPDF2)*
- Tevatron Run-2 and LHC jet production, $t\bar{t}$ production

2. Available theoretical cross sections

Process	Number of QCD loops	Mass scheme*	
Neutral current	2	ZM	*Moch, Vermaseren, Vogt*
DIS	2	GM	*Riemersma, Harris, Smith, van Neerven*
			Buza, Matiounine, Smith, van Neerven
Charged current	2	ZM	*Moch, Vermaseren, Vogt*
DIS	2, in progress	GM	*Bluemlein et al.*
$pN \xrightarrow{\gamma^*, W, Z} \ell\bar{\ell}^{(\prime)} X$	2	ZM	*Anastasiou, Dixon, Melnikov, Petriello*
$p\bar{p} \to jX$	2, in progress	ZM	
$ep \to jjX$	2	ZM	

*ZM/GM: zero-mass/general-mass approximation for c, b contributions

3. Requirements for PDF parametrizations

A. A valid set of $f_{a/p}(x,Q)$ must satisfy QCD sum rules

Valence sum rule

$$\int_0^1 [u(x,Q) - \bar{u}(x,Q)]\, dx = 2 \qquad \int_0^1 \left[d(x,Q) - \bar{d}(x,Q)\right] dx = 1$$

$$\int_0^1 [s(x,Q) - \bar{s}(x,Q)]\, dx = 0$$

A proton has net quantum numbers of 2 u quarks + 1 d quark

Momentum sum rule

$$[\text{proton}] \equiv \sum_{a=g,q,\bar{q}} \int_0^1 x\, f_{a/p}(x,Q)\, dx = 1$$

momenta of all partons must add up to the proton's momentum

Through this rule, normalization of $g(x,Q)$ is tied to the first moments of quark PDFs

3. Requirements for PDF parametrizations

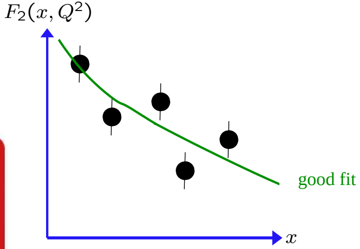
B. A valid PDF set must **not** produce unphysical predictions for observable quantities

Example

- Any concievable hadronic cross section σ must be non-negative: $\sigma \geq 0$
  - this is typically realized by requiring $f_{a/p}(x, Q) > 0$
- Any cross section asymmetry A must lie in the range $-1 \leq A \leq 1$
  - this constrains the range of allowed PDF pararametrizations

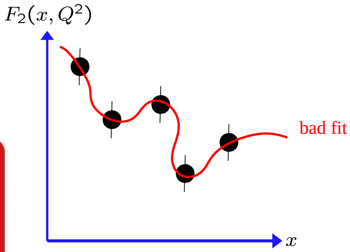
3. Requirements for PDF parametrizations

C. PDF parametrizations for $f_{a/p}(x,Q)$ must be “flexible just enough” to reach agreement with the data, without reproducing random fluctuations



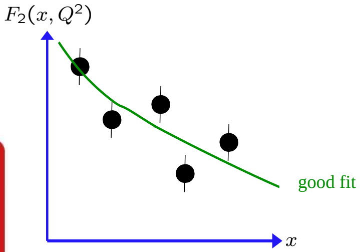
3. Requirements for PDF parametrizations

C. PDF parametrizations for $f_{a/p}(x, Q)$ must be "flexible just enough" to reach agreement with the data, without reproducing random fluctuations



3. Requirements for PDF parametrizations

C. PDF parametrizations for $f_{a/p}(x, Q)$ must be "flexible just enough" to reach agreement with the data, without reproducing random fluctuations



Traditional solution

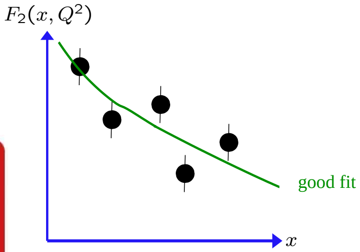
"Theoretically motivated" functions with 10-30 parameters

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

- $x \to 0$: $f \propto x^{a_1}$ – Regge-like behavior
- $x \to 1$: $f \propto (1-x)^{a_2}$ – quark counting rules
- $F(a_3, a_4, ...)$ affects intermediate x; just a convenient functional form

3. Requirements for PDF parametrizations

C. PDF parametrizations for $f_{a/p}(x,Q)$ must be "flexible just enough" to reach agreement with the data, without reproducing random fluctuations



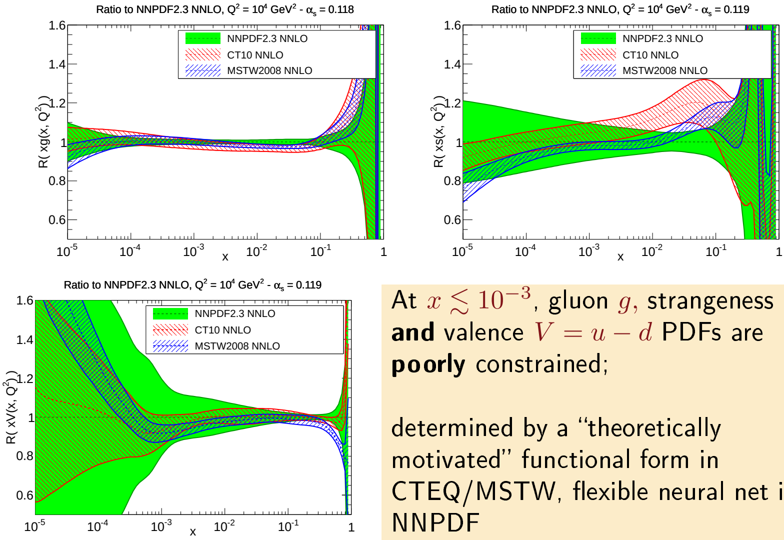
An alternative solution

Neural Network PDF collaboration

- Monte-Carlo sampling + parametrization of $f_{a/p}(x,Q)$ by ultra-flexible functions — neural networks, ~ 250 parameters
- Less bias from the choice of the functional form; PDF uncertainties are similar to the traditional methods in the $\{x,Q\}$ region covered by the data; much larger uncertainties in the unconstrained $\{x,Q\}$regions

Ratios of CT10 and MSTW NNLO PDFs to NNPDF2.3 at $Q = 100$ GeV

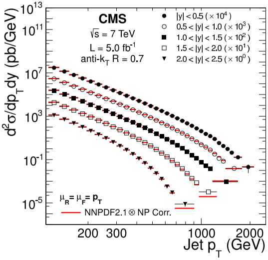
(arXiv:*1211.5142*)

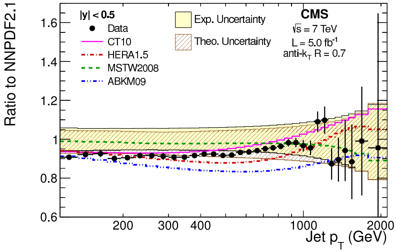


At $x \lesssim 10^{-3}$, gluon g, strangeness s, **and** valence $V = u - d$ PDFs are **poorly** constrained;

determined by a "theoretically motivated" functional form in CTEQ/MSTW, flexible neural net in NNPDF

Correlated systematic errors in the PDF analysis





At high luminosities, statistical errors in the collider data, such as jet production, are much smaller than correlated systematic errors.

The latest data sets are provided with many ($\sim 10 - 200$) sources of correlated errors.

Special methods are developed to include the effect of correlated systematic errors into the PDF uncertainties.

4. Statistical aspects

J. Pumplin et al., JHEP 0207, 012 (2002), and references therein; J. Collins & J. Pumplin, hep-ph/0105207

Suppose there are N PDF parameters $\{a_i\}$, N_{exp} experiments, M_k data points and N_k correlated systematic errors in each experiment

Each systematic error is associated with a random parameter r_n, assumed to be distributed as a Gaussian distribution with unit dispersion

The best external estimate of syst. errors corresponds to $\{r_n = 0\}$; but we must allow for $r_n \neq 0$

The most likely combination of $\{a\}$ and $\{r\}$ is found by minimizing

$$\chi^2 = \sum_{k=1}^{N_{exp}} w_k \chi_k^2$$

$w_k > 0$ are the weights applied to emphasize or de-emphasize contributions from individual experiments (default: $w_k = 1$)

4. Statistical aspects

J. Pumplin et al., JHEP 0207, 012 (2002), and references therein; J. Collins & J. Pumplin, hep-ph/0105207

χ^2 for one experiment is

$$\chi_k^2 = \sum_{i=1}^{M_k} \frac{1}{\sigma_i^2} \left(D_i - T_i(\{a\}) - \sum_{n=1}^{R_k} r_n \beta_{ni} \right)^2 + \sum_{n=1}^{R_k} r_n^2$$

D_i and T_i are **data** and **theory** values at each point

$\sigma_i = \sqrt{\sigma_{stat}^2 + \sigma_{syst,uncor}^2}$ is the total statistical + systematical **uncorrelated** error

$\sum_n \beta_{ni} r_k$ are **correlated** systematic shifts

β_{ni} is the **correlation** matrix; is provided with the data or theoretical cross sections before the fit

$\sum_n r_n^2$ is the penalty for deviations of r_n from their expected values, $r_n = 0$

4. Statistical aspects

J. Pumplin et al., JHEP 0207, 012 (2002), and references therein; J. Collins & J. Pumplin, hep-ph/0105207

Each χ_k can be **analytically** minimized with respect to **the Gaussian** r_n , with the result

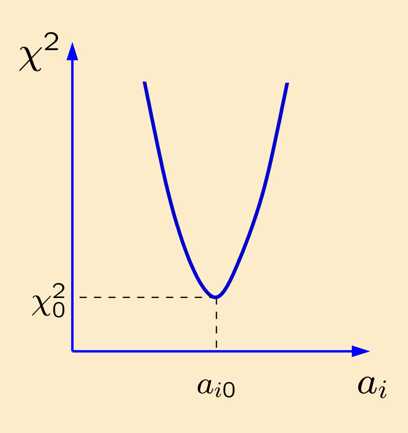
$$r_n\left(\{a\}\right) = \sum_{n'=1}^{R_k} (A^{-1})_{nn'} B_{n'}(\{a\})$$

$$A_{nn'} = \delta_{nn'} + \sum_{i=1}^{M_k} \frac{\beta_{ni}\beta_{n'i}}{\sigma_i^2}; \qquad B_n(\{a\}) = \sum_{i=1}^{M_k} \frac{\beta_{ni}(D_i - T_i)}{\sigma_i^2}$$

$$\chi_k^2 = \sum_{i=1}^{M_k} \frac{1}{\sigma_i^2} \left(D_{ki} - T_{ki}(\{a\})\right)^2 + \sum_{n,n'=1}^{R_k} B_n (A^{-1})_{nn'} B_{n'} \qquad (1)$$

Numerical minimization of $\sum_k w_k \chi_k^2(a, r(a))$ (with χ_k from Eq. (1)) then establishes the region of acceptable $\{a\}$, which includes the largest possible variations of $\{a\}$ that are allowed by the systematic effects

Finding the central PDF set and the PDF uncertainty

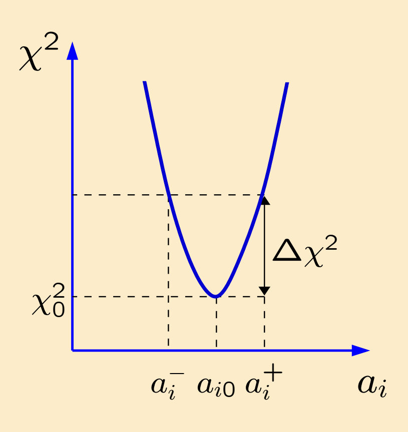


The central (most probable) PDF set corresponds to the minimum $\{a_{i0}, r_{k0}\}$ of the likelihood function (χ^2) in space of $N \sim 30$ theoretical (mostly PDF) parameters $\{a_i\}$ and $N_{exp} \cdot N_k > 100$ experimental systematical parameters

The χ^2 minimum, χ_0^2, is found

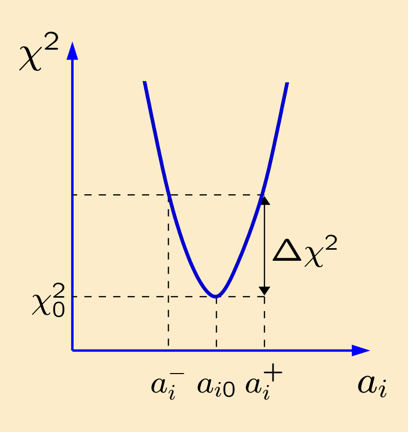
- by numerical minimization (CT, HERA, MSTW)
- by Monte-Carlo sampling of the $\{a_i\}$ space (NNPDF)

Finding the central PDF set and the PDF uncertainty



Then, a confidence region in $\{a_i\}$ space for a given tolerated increase in χ^2 is established. From this confidence region, the PDF uncertainty ΔX on a QCD observable X can be estimated by constructing additional PDF sets ("error sets"), besides the best-fit set.

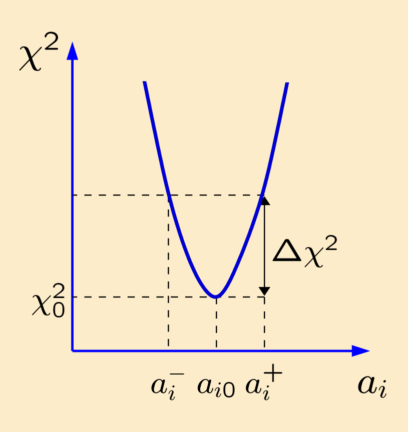
Finding the central PDF set and the PDF uncertainty



For arbitrary QCD observables, ΔX can be estimated using...

- $2N$ eigenvector PDF sets in the Hessian method (CTEQ)
- 100-1000 unweighted replicas in the Monte-Carlo method (NNPDF)

Finding the central PDF set and the PDF uncertainty

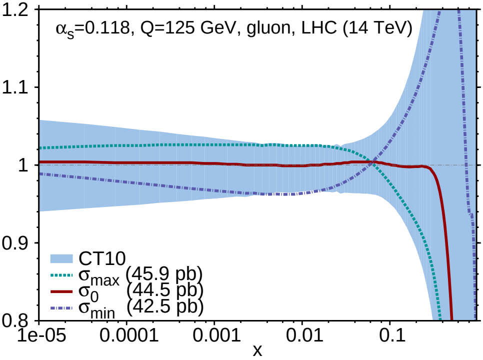


Alternatively, if some observable X is particularly prominent (e.g., the total cross section for $gg \to$ Higgs process), 2 special error sets can be constructed just for estimating ΔX of this observable in the Lagrange Multiplier (LM) method or the data set diagonalization method.

CT10(H) error PDFs for Higgs cross sections in the Hessian and LM methods

S. Dulat et al., arXiv:1310.7601 [hep-ph]

CT10H gluon PDFs at the momentum scale 125 GeV, compared to the CT10 error band, at the 90%c.l.



CT10 is our regular PDF ensemble consisting of 51 PDFs.

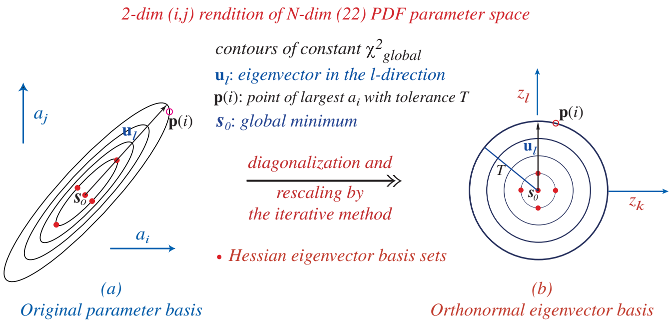
The CT10H fits give the central prediction (σ_0), the minimum (σ_{min}) and maximum (σ_{max}) predictions obtained using the Lagrange Multiplier method, for $\sigma(gg \to H)$ at the LHC with 14 TeV.

Key formulas of the Hessian method

The Hessian method is used by CTXX PDFs to propagate the PDF uncertainty into QCD predictions.

It provides several simple formulas for computing the PDF uncertainties and PDF-induced correlations.

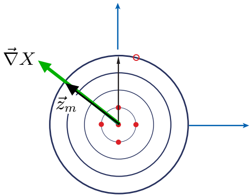
Tolerance hypersphere in the PDF space



A hyperellipse $\Delta\chi^2 \leq T^2$ in space of N physical PDF parameters $\{a_i\}$ is mapped onto a filled hypersphere of radius T in space of N orthonormal PDF parameters $\{z_i\}$

Tolerance hypersphere in the PDF space

2-dim (i,j) rendition of N-dim (22) PDF parameter space



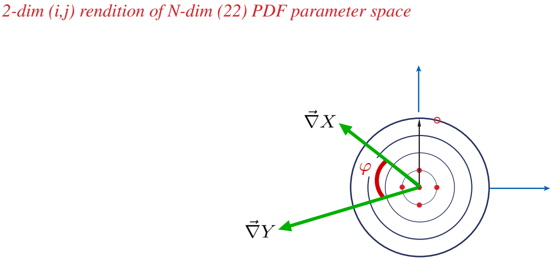
(b)
Orthonormal eigenvector basis

A symmetric PDF error for a physical observable X is given by

$$\Delta X = \vec{\nabla} X \cdot \vec{z}_m = \left|\vec{\nabla} X\right| = \frac{1}{2}\sqrt{\sum_{i=1}^{N}\left(X_i^{(+)} - X_i^{(-)}\right)^2}$$

Asymmetric PDF errors can be also computed.

Tolerance hypersphere in the PDF space



(b)
Orthonormal eigenvector basis

Correlation cosine for observables X and Y:

$$\cos\varphi = \frac{\vec{\nabla}X\cdot\vec{\nabla}Y}{\Delta X \Delta Y} = \frac{1}{4\Delta X\,\Delta Y}\sum_{i=1}^{N}\left(X_i^{(+)} - X_i^{(-)}\right)\left(Y_i^{(+)} - Y_i^{(-)}\right)$$

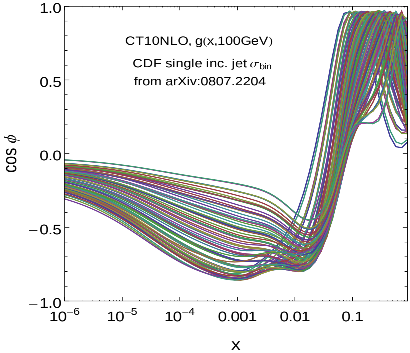
Sensitivity of PDFs to input cross sections

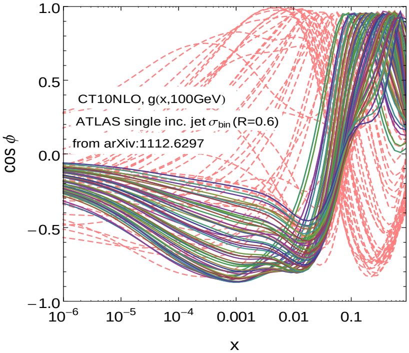
It is often interesting to know which specific data sets included in the PDF analysis constrain a given $f_{a/p}(x,Q)$; or which PDF flavors drive the PDF uncertainty in a given QCD observable X

This question can be addressed in several ways:

- by computing the correlation cosine $\cos\phi$ in the Hessian method *(Nadolsky et al., 0802.0007)*
- by a Lagrange multiplier scan of X *(Stump et al., hep-ph/0101051)*
- or by introducing effective Gaussian variables S_n dependent on X *(Dulat et al., 1309.0025, 1310.7601; Lai et al., 1007.2241)*

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

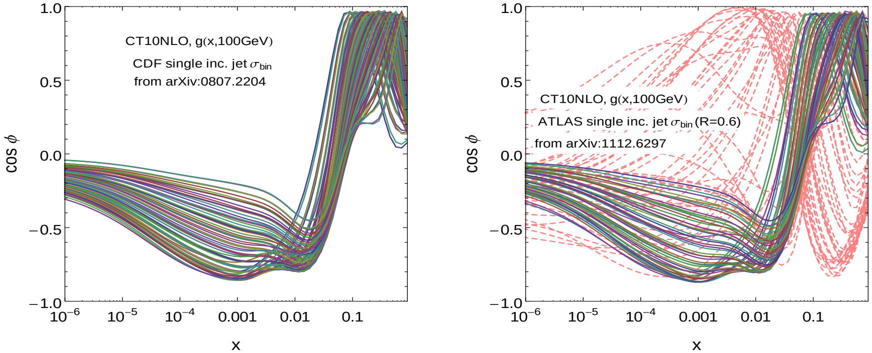




$\cos\phi(x)$ measures the correlation between the PDF uncertainty of $d^2\sigma/(dp_T dy)$ in inclusive jet production, and $g(x,Q)$ at a given x.

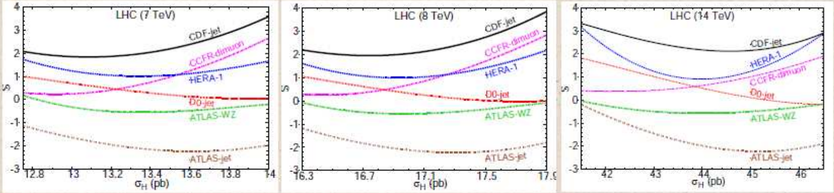
The curves show $\cos\varphi$ between the NLO theory cross sections in experimental $\{p_{Tj}, y_j\}$ bins and $g(x,Q)$, for various x values in $g(x,Q)$.

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



The CDF (left) and ATLAS (right) jet data are correlated with $g(x,Q)$ when $\cos\phi \gtrsim 0.7$ or $\cos\phi \lesssim -0.7$. The CDF (ATLAS) jet measurement are directly sensitive to $g(x,Q)$ with $x \gtrsim 0.1$ (0.01), and indirectly at $x \sim 0.001$ via the momentum sum rule. 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 at CDF.

Sensitivity to Data Sets



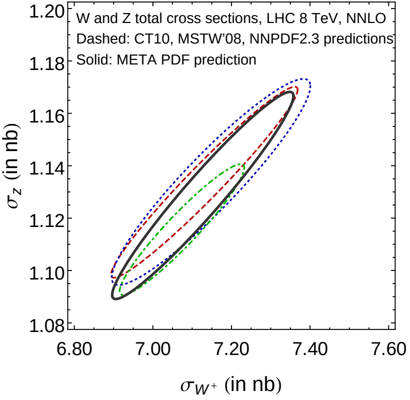
- How sensitive are included data sets to value of σ_H?
- "Effective Gaussian Variable" S maps cumulative χ^2 distribution for N_{pt} onto cumulative Gaussian distribution
  - +1,+2,+3,… equivalent to that many sigma deviations
  - Negative values correspond to anomalously well-fit data
- Most strongly correlated data: high p_T jet, inclusive HERA, CCFR-dimuon
  - HERA more strongly correlated with 14 TeV—smaller x
  - CCFR dimuon correlation due to gluon-strange interdependence

Combination of PDF uncertainties from several groups

As the final step, it may be necessary to combine QCD cross sections X_i from several PDF ensembles and find the ultimate PDF uncertainty. This combination can be done...

(a) ...according to the PDF4LHC convention, by combining X_i found for every individual error PDF *(M. Botje et al., (2011), arxiv:1101.0538; R. Ball, et al., 1211.5142)*

(b) ...by introducing meta-PDFs to first average the input PDFs in the PDF parameter space, before computing X_i *(J. Gao, P. Nadolsky, 1401.0013)*.



$\alpha_s(M_Z)$ and heavy-quark masses in the PDF fits

- The QCD coupling, $\alpha_s(M_Z)$, and $\overline{MS}$ heavy-quark masses, $m_c(m_c)$, $m_b(m_b)$, and $m_t(m_t)$, can be also varied in the PDF fits.

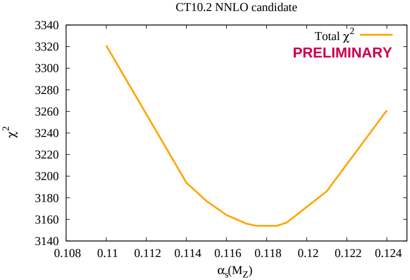
  However, the resulting constraints on these parameters are much weaker than from their dedicated measurements.

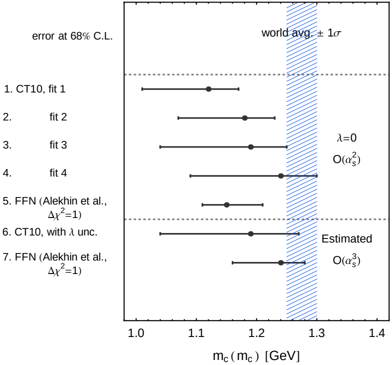
- CT10 NNLO assumes a fixed $\alpha_s(M_Z) = 0.118 \pm 0.002$ at 90%c.l., which is equal to the world average $\alpha_s(M_Z)$.

  For any X, the α_s uncertainty $\Delta_{\alpha_s} X$ is correlated with the PDF uncertainty $\Delta_{PDF} X$, the two must be combined in the full prediction.

  CT10 provides PDF error sets to compute the PDF+α_s uncertainty with full correlation.

Constraints on $\alpha_s(M_Z)$ and the $\overline{MS}$ charm mass $m_c(m_c)$ in the CT10 NNLO fit





- NLO: $\alpha_s(M_Z) = 0.11964 \pm 0.0064$ at 90% c.l.
- NNLO: $\alpha_s(M_Z) = 0.118 \pm 0.005$

Gao, Guzzi, Nadolsky, 1304.0494
$m_c(m_c) = 1.19^{+0.08}_{-0.15}$ GeV at 68% c.l.

$\alpha_s(M_Z)$ and heavy-quark masses in the PDF fits

- The QCD coupling, $\alpha_s(M_Z)$, and $\overline{MS}$ heavy-quark masses, $m_c(m_c)$, $m_b(m_b)$, and $m_t(m_t)$, can be also varied in the PDF fits.

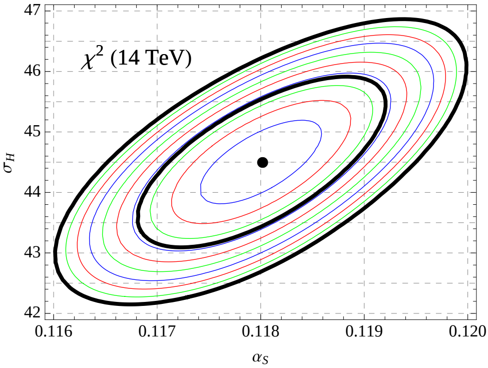
  However, the resulting constraints on these parameters are much weaker than from their dedicated measurements.

- CT10 NNLO assumes a fixed $\alpha_s(M_Z) = 0.118 \pm 0.002$ at 90%c.l., which is equal to the world average $\alpha_s(M_Z)$.

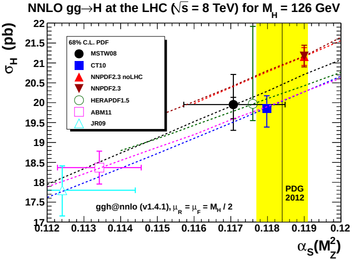
  For any X, the α_s uncertainty $\Delta_{\alpha_s} X$ is correlated with the PDF uncertainty $\Delta_{PDF} X$, the two must be combined in the full prediction.

  CT10 provides PDF error sets to compute the PDF+α_s uncertainty with full correlation.

PDF+α_s uncertainties of the total $gg \to H$ cross section



Dulat et al., 1310.7601



Forte, Watt, 1301.6754

CT10 estoimate of the PDF+α_s uncertainty

CT10 NNLO set provides

- $2N+1=51$ PDF error sets at $\alpha_s(M_Z)=0.118$, to compute

$$\Delta_{PDF}X=\frac{1}{2}\sqrt{\sum_{i=1}^{25}\left(X_i^{(+)}-X_i^{(-)}\right)^2}\quad \text{(the PDF uncertainty);}$$

- 2 best-fit PDFs for for $\alpha_s=0.116$ and 0.120, to compute

$$\Delta_{\alpha_s}X=\frac{X(\alpha_s=0.120)-X(\alpha_s=0.116)}{2}\quad \text{(the } \alpha_s \text{ uncertainty).}$$

The formula for the combined PDF+α_s uncertainty with **with full correlation** is *(Lai et al., 1004.4624)*

$$\Delta_{PDF+\alpha_s}X=\sqrt{(\Delta_{PDF}X)^2+(\Delta_{\alpha_s}X)^2}$$

What about $s(x,Q) \neq \bar{s}(x,Q)$?

This can be tested in subprocesses

$$W^+ s \to c \text{ and } W^- \bar{s} \to \bar{c}$$

In the experiment, charm quarks can be identified by their semileptonic decays,

$$c \to s\mu^+\nu \text{ and } \bar{c} \to \bar{s}\mu^-\bar{\nu}$$

So one sees

$$\nu N \to \mu^- c X \to \mu^- \mu^+ X$$
$$\bar{\nu} N \to \mu^+ c X \to \mu^+ \mu^- X$$

— SIDIS muon pair production (NuTeV)

Total strangeness and strangeness asymmetry

Denote

$$[q_i](Q) \equiv \int_0^1 x\, q_i(x,Q)\, dx \text{ —net moment fraction carried by } q_i$$

and introduce $s^{\pm}(x) = s(x) \pm \bar{s}(x)$ (total strangeness and its asymmetry). It is possible that

$$\int_0^1 s^-(x)dx = 0$$

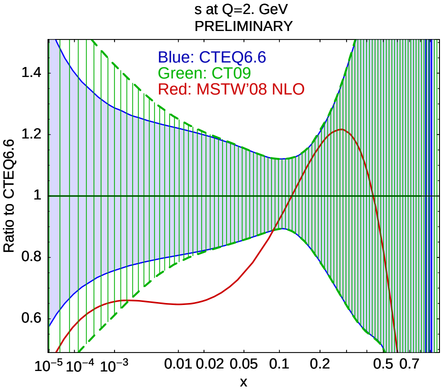
(a proton has no net strangeness), but

$$[S^-] \equiv \int_0^{-1} xs^-(x)dx \neq 0$$

(s and $\bar{s}$ have different x distributions)

A large non-vanishing $[S^-]$ might explain "the NuTeV weak angle anomaly"

CCFR (inclusive νN DIS) and NuTeV (SIDIS dimuon production): constraints on strangeness



$s^+(x,Q)$ is reasonably well constrained at $x > 0.01$; practically unknown at $x < 0.01$

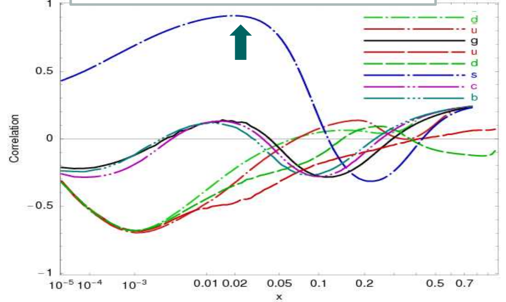
2009 NNPDF estimate (least biased by the parametrization of $s^-(x,Q)$):

$$[S^-](Q^2 = 20\ \mathrm{GeV}^2) = 0 \pm 0.009$$

No statistically significant $[S^-]$; but the PDF error is large enough to eliminate the NuTeV anomaly (!)

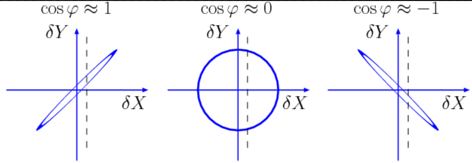
Constraining strangeness PDF by LHC W and Z cross sections

Correlation between σ_W/σ_Z and f(x,Q=85 GeV)



2008, CTEQ6.6 (arXiv:0802.0007): the ratio σ_W/σ_Z at LHC must be sensitive to the strange PDF $s(x,Q)$

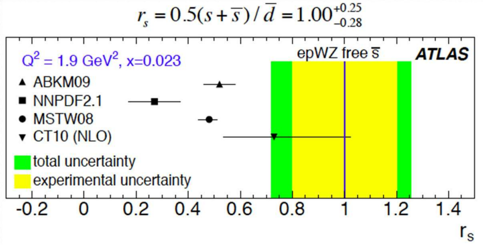
The uncertainty on $s(x,Q)$ limits the accuracy of the W boson mass measurement at the LHC



Correlation cosine $\cos\varphi \approx \pm 1$:

⇔ Measurement of X imposes tight constraints on Y

Constraining strangeness PDF by LHC W and Z cross sections

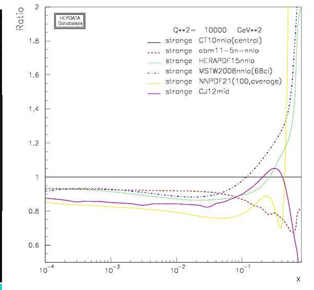


2012: the ATLAS analysis (arXiv:1203.4051) of W and Z production suggests

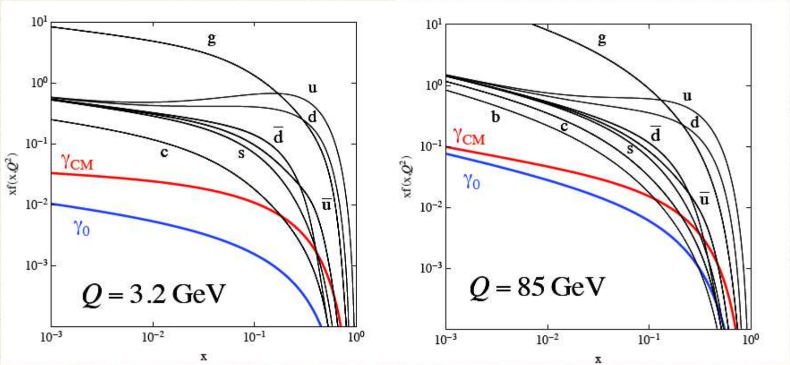
$\bar{s}(x,Q)/\bar{d}(x,Q) = 1.00^{+0.25}_{-0.28}$

at $x = 0.023$ and $Q^2 = 1.9$ GeV^2

Similarly, $\sigma(W^+c)/\sigma(W^-c)$ cross section ratios show preference for $\frac{\bar{s}(x,Q)}{\bar{d}(x,Q)} \sim 1$, larger than in most pre-LHC PDF sets



Photon PDFs (in proton)



γ momentum fraction:

$p^\gamma(Q)$	$\gamma(x,Q_0)=0$	$\gamma(x,Q_0)_{CM}$
$Q = 3.2$ GeV	0.05%	0.34%
$Q = 85$ GeV	0.22%	0.51%

Photon PDF can be larger than sea quarks at large x!

Initial Photon PDF still
← significant at large Q.

C. Schmidt et al.

Inclusion of Photon PDFs

LO QED + (NLO or NNLO) QCD evolution:

$$\frac{dq}{dt} = \frac{\alpha_s}{2\pi}\left(P_{qq}\circ q + P_{qg}\circ g\right) + \frac{\alpha}{2\pi}\left(e_q^2 \tilde{P}_{qq}\circ q + e_q^2 \tilde{P}_{q\gamma}\circ \gamma\right)$$

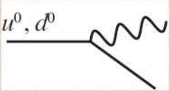
$$\frac{dg}{dt} = \frac{\alpha_s}{2\pi}\left(P_{gg}\circ g + P_{gq}\circ \sum (q+\bar{q})\right) \qquad t = \ln Q^2$$

$$\frac{d\gamma}{dt} = \frac{\alpha}{2\pi}\left(\tilde{P}_{\gamma\gamma}\circ \gamma + \tilde{P}_{\gamma q}\circ \sum e_q^2 (q+\bar{q})\right)$$

"Radiative ansatz" for initial Photon PDFs (generalization of MRST choice)

$$\gamma^p = \frac{\alpha}{2\pi}\left(A_u e_u^2 \tilde{P}_{\gamma q}\circ u^0 + A_d e_d^2 \tilde{P}_{\gamma q}\circ d^0\right)$$

$$\gamma^n = \frac{\alpha}{2\pi}\left(A_u e_u^2 \tilde{P}_{\gamma q}\circ d^0 + A_d e_d^2 \tilde{P}_{\gamma q}\circ u^0\right)$$



where u^0 and d^0 are "primordial" valence-type distributions of the proton.
Assumed approximate isospin symmetry for neutron.
Here, we take A_u and A_d as unknown fit parameters.

MRST choice: $A_q = \ln\left(Q_0^2/m_q^2\right)$ "Radiation from **C**urrent **M**ass" - **CM**

C. Schmidt et al.

Conclusion: the role of the PDF analysis now and in the future

It is a payoff decade for the PDF analysis efforts!

- A windfall of new data to compare with (LHC, HERA, Tevatron, JLab, RHIC,...)
- Tests of QCD factorization at new $\sqrt{s}$, targeting 1% precision
- New tools (HERA Fitter, APPLGRID, FastNLO,...) and methods (MC sampling, PDF reweighting) revolutionize the PDF analysis
- Understanding of the PDFs at the (N)NNLO level will be crucial for the success of the LHC physics program

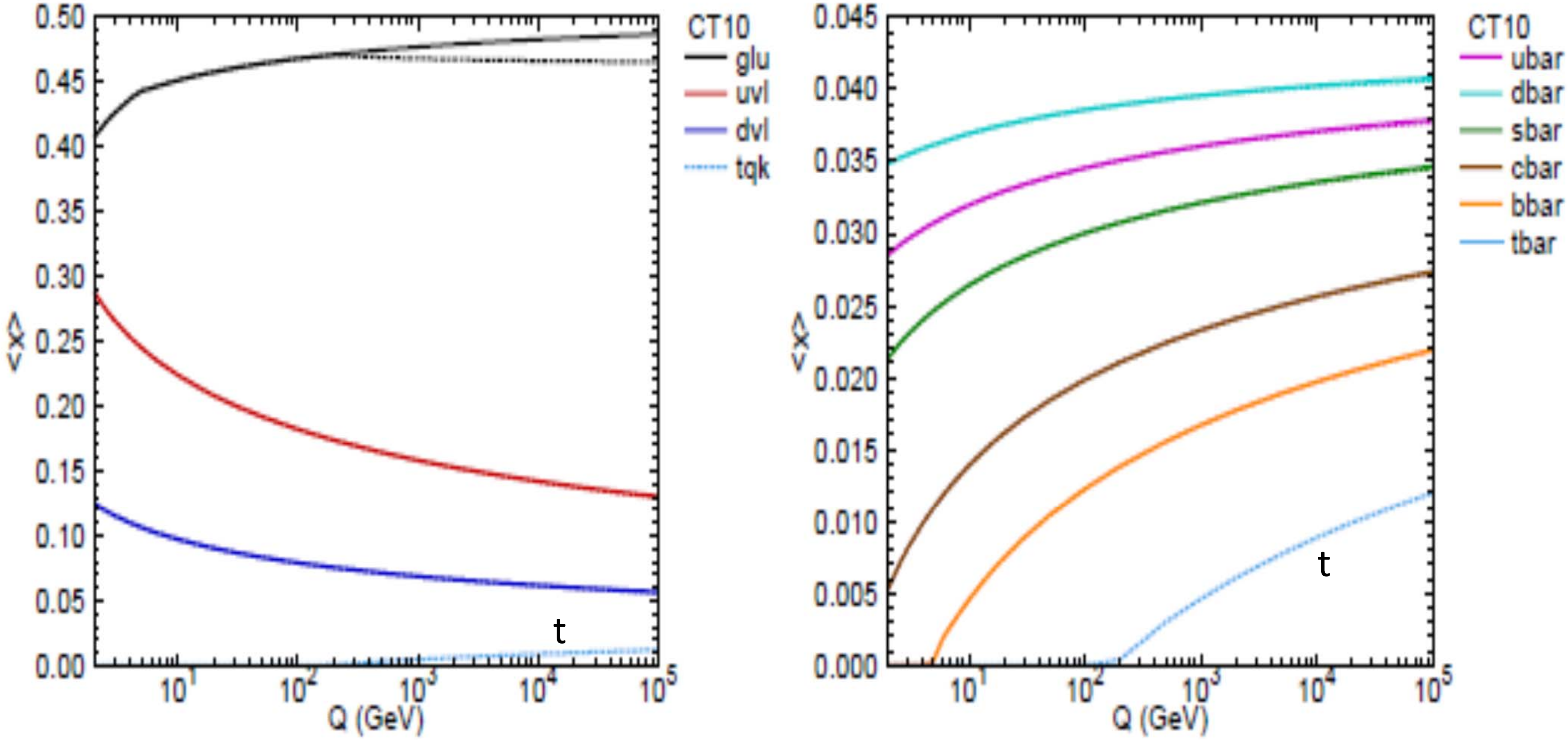
PDF analysis
continues to be
a fertile field
open to young researchers!

Backup slides

Top quark as a parton

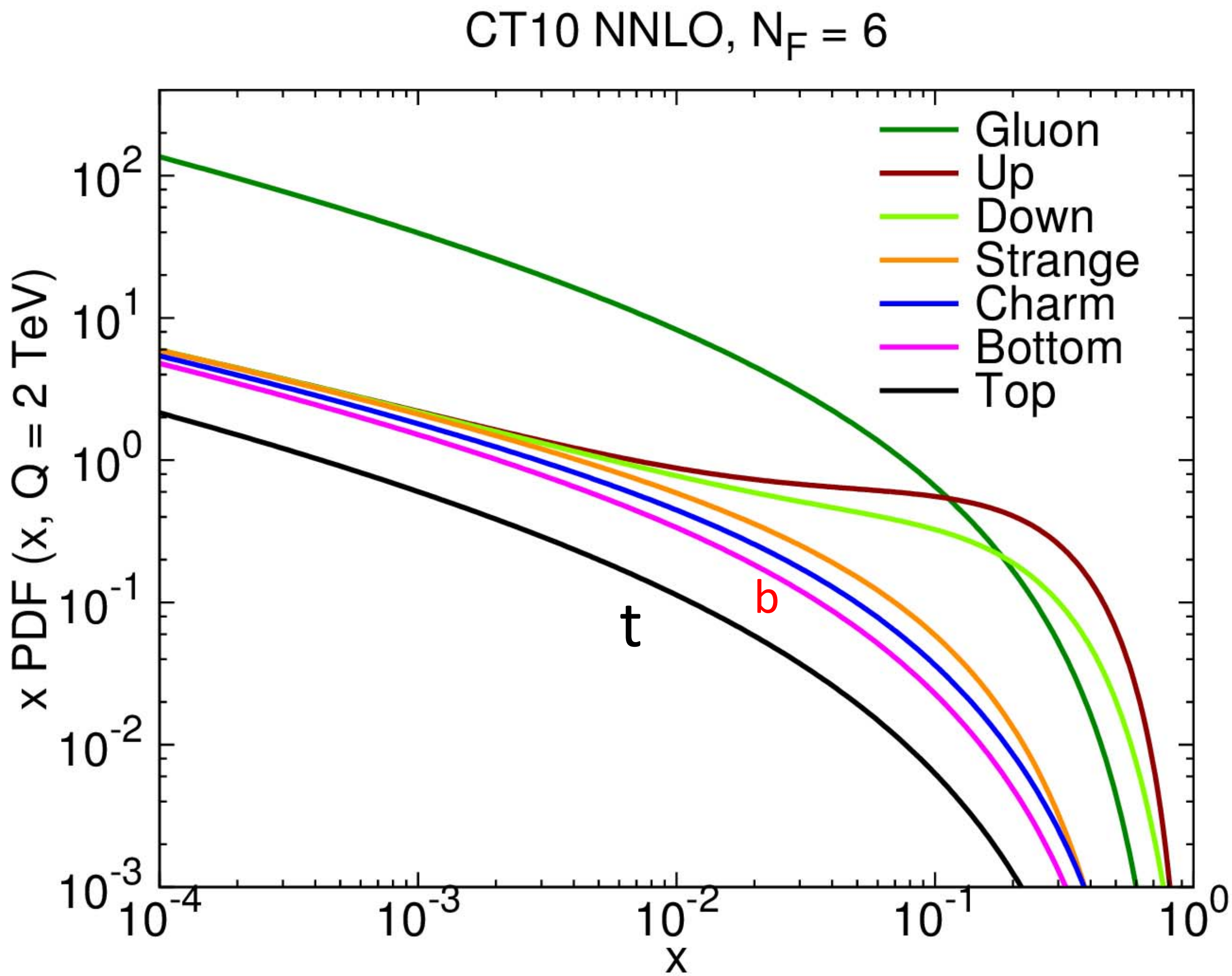
- For a 100 TeV SppC, top mass (172 GeV) can be ignored; top quark, just like bottom quark, can be a parton of proton.
- Top parton will take away some of the momentum of proton, mostly, from gluon (at NLO).
- Need to use s-ACOT scheme to calculate hard part matrix elements, to be consistent with CT10 PDFs.

Momentum fraction inside proton



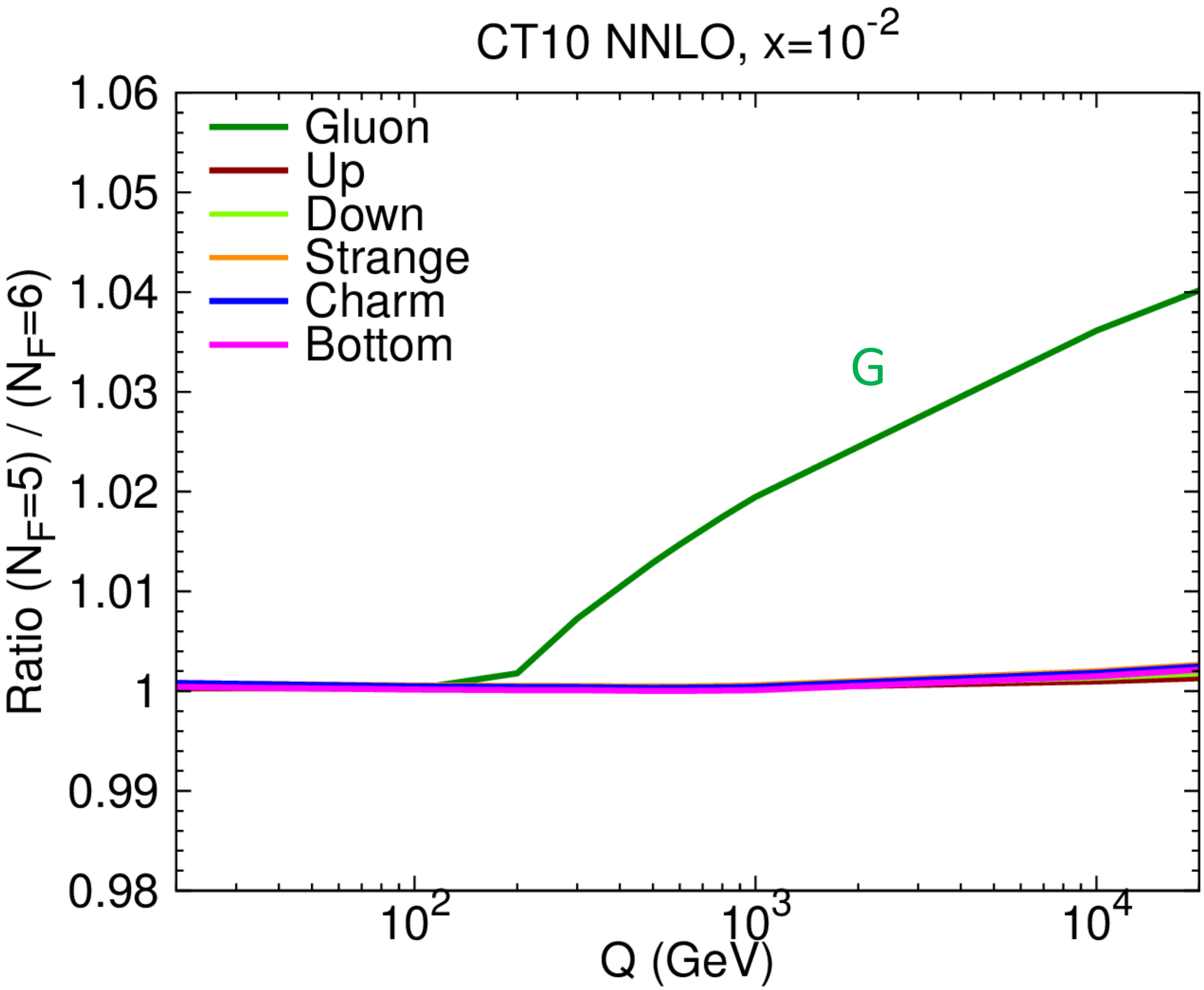
Solid curves: CT10 NNLO
Dashed curves: CT10Top NNLO

CT10Top PDFs (Q=2 TeV)

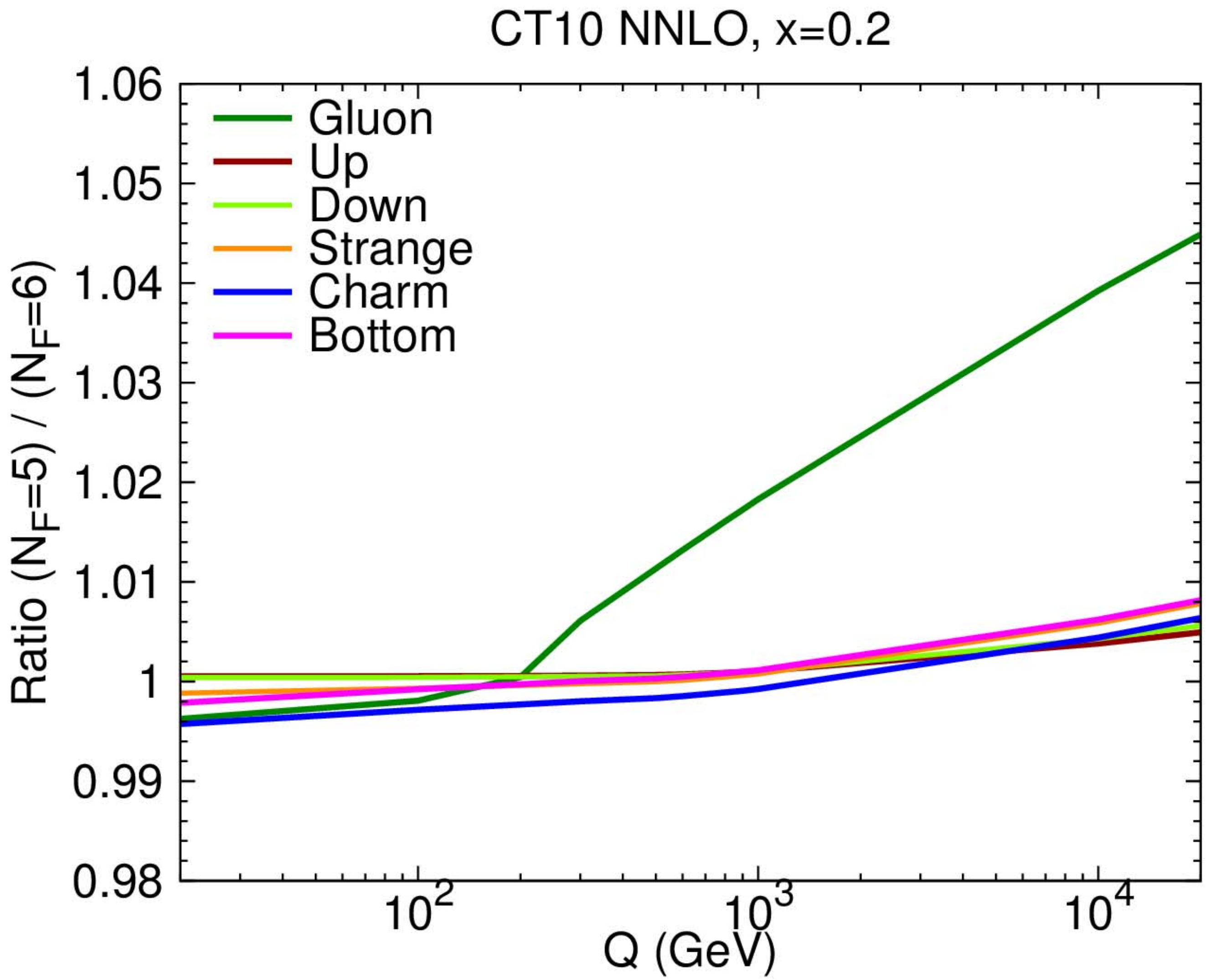


Top PDF is only a factor of 2 smaller than b PDF

CT10Top PDFs

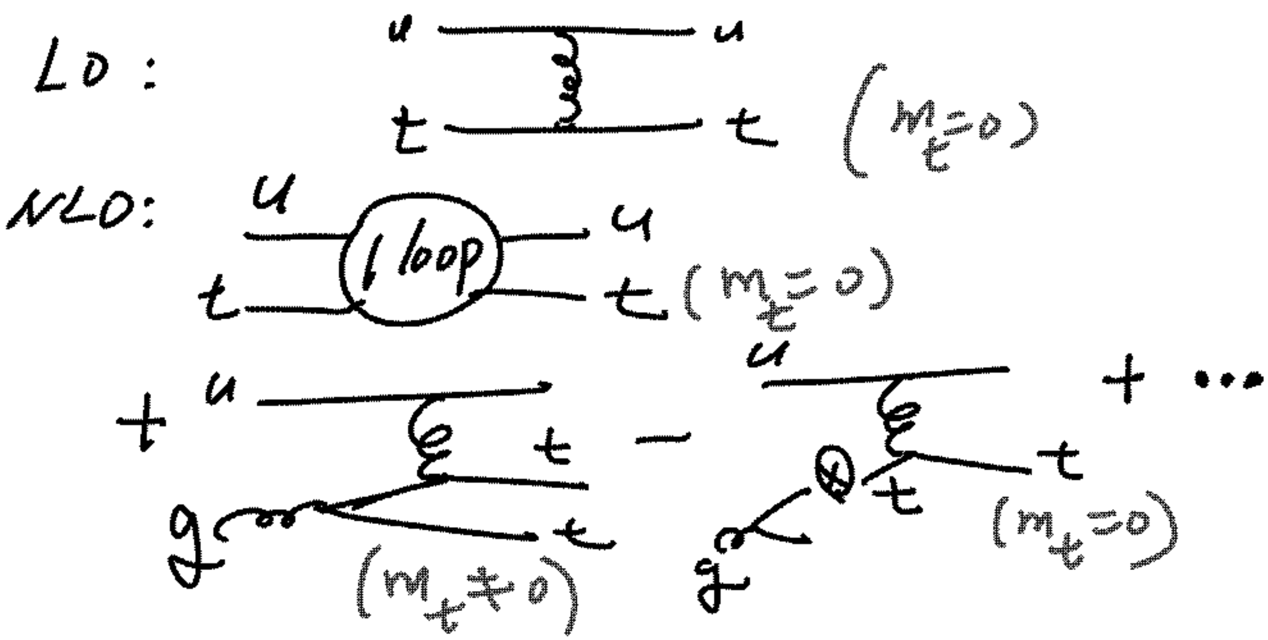


CT10Top PDFs

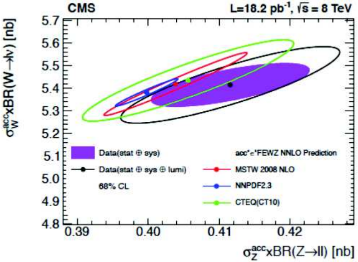


Hard part calculation

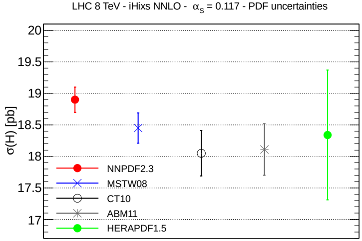
- S-ACOT scheme
- Example: single-top production



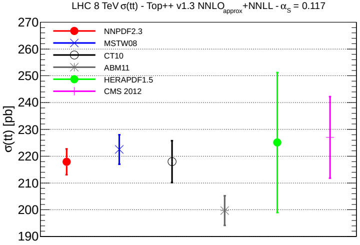
NNLO predictions for LHC total cross sections



W & Z production



SM Higgs production



$t\bar{t}$ production