Parton Distribution Functions Introductory Lecture: PDF2

CTEQ School - 20 July 2019

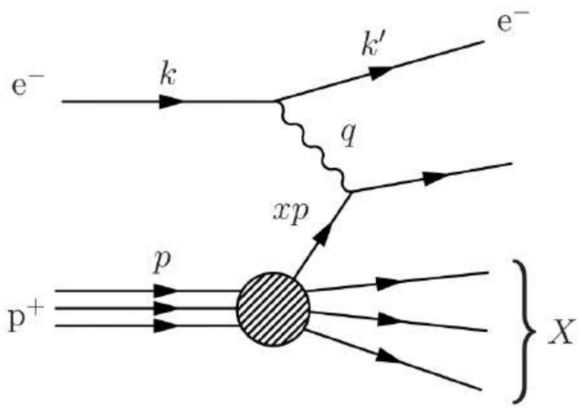
University of Pittsburgh

Outline

- PDFs and how to determine them
- The Hessian method
- Constraints from hadronic data
- QCD global analyses

Factorization

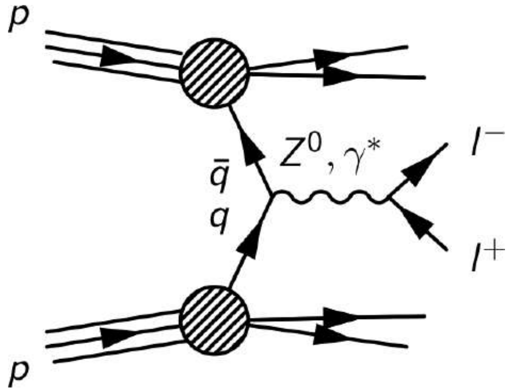
In previous lectures, we learned that in DIS the inclusive cross section can be written as



$$\sigma = f \otimes H + \mathcal{O}\left(\frac{\Lambda_{\text{QCD}}}{Q}\right)$$

where H is a calculable hard function and f a universal PDF. In the DIS case, we introduced f as a non-perturbative object that accounts for the nucleon's probability of emitting a parton carrying a fraction of the incoming nucleon's momentum.

A similar factorized form form for the cross section can also be used for proton-proton collisions



$$\sigma = f \otimes f \otimes H + \mathcal{O}\left(\frac{\Lambda_{\text{QCD}}}{Q}\right)$$

Factorization

Parton distribution functions (PDFs) of the proton are essential ingredients of factorization theorems in QCD:

The general structure of the inclusive cross section for high-energy collisions involving hadron-hadron beams, lepton-hadrons, or hadron targets, is a convolution product where long-distance non-perturbative contributions (PDFs) and short-distance infrared-safe perturbatively calculable quantities (hard scattering) are separated.

For Drell-Yan we have (Collins, Soper, Sterman (PLB 1984), (NPB 1985))

$$\sigma(H_1 H_2 \to l^+ l^- + X) = \sum_{a,b} \int_{x_1}^{1} d\xi_1 \int_{x_2}^{1} d\xi_2 f_{H_1\to a}(\xi_1, \alpha_s(\mu_R), \mu_R, \mu_F) f_{H_2\to b}(\xi_2, \alpha_s(\mu_R), \mu_R, \mu_F)$$

$$\times \hat{\sigma}^{ab}\left(\frac{x_1}{\xi_1}, \frac{x_2}{\xi_2}; \alpha_s(\mu_R), Q, \mu_F, \mu_R\right) + \mathcal{O}\left(\frac{\Lambda^2}{Q^2}\right)$$

Can PDFs be defined in a more rigorous way?

- Operator Product Expansion:

perform a Taylor expansion at large Q^2 in which we could drop sub-leading terms.

$$T\{J^\mu(x)J^\nu(y)\} = \sum_n C_n(x-y)\mathcal{O}_n^{\mu\nu}(x)$$

- Light-cone variables:

use light-cone projections of proton matrix elements. This approach is more friendly to more general factorization arguments (see Dave Soper's talk).

To all orders in α_s, PDFs can be **defined** as correlator functions:

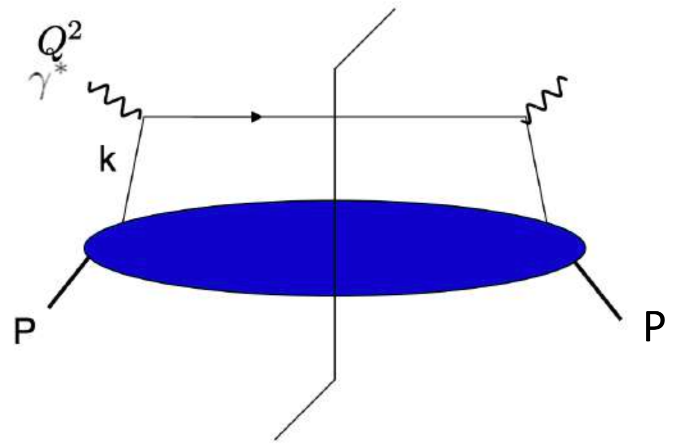
$$n^\mu = (1, \vec{n})$$

$$f_q(\xi) = \int_{-\infty}^{\infty} \frac{dt}{2\pi} e^{-it\xi(n\cdot P)} \langle P|\bar{\psi}_q(tn^\mu)\,\frac{\not{n}}{2} W_n \psi_q(0)|P\rangle$$

$$W_n = P\exp\left\{ig_s n_\nu \int_0^t ds\, A^\nu(sn^\mu)\right\}$$ Gauge link for having a gauge invariant correlator

The formal definition of PDFs in QCD, contains all the complications of "real life":
UV regulator in DR, gauge invariance,…

$$f_q(\xi) = \int_{-\infty}^{\infty} \frac{dt}{2\pi} e^{-it\xi(n\cdot P)} \langle P|\bar{\psi}_q(tn^\mu)\, \frac{\not{n}}{2} W_n \psi_q(0)|P\rangle$$



Similarly to the case of renormalization scheme, a set of rules has to be provided
in order to define the PDFs when a cross section calculation is performed, e.g. MSbar scheme.

PDFs universality

- Gluons, quarks and antiquarks are the known constituents of the proton.
- Their distributions as a function of x and generic scale μ, at which partons are probed, are universal quantities that do not depend on the specific hard process under consideration.

Differently from the hard-scattering cross section, the analytic structure of the PDFs cannot be fully predicted by perturbative QCD, but has to be determined by comparing standard sets of cross sections, to experimental measurements by using a variety of analytical/statistical methods.

For this reason PDFs are “data-driven” quantities.

In previous lectures we also learned that PDF's energy behavior can be predicted by RGE DGLAP equations

$$\mu \frac{d}{d\mu} f_i(x,\mu) = \frac{\alpha_s}{\pi} \int_x^1 \frac{d\xi}{\xi} f_i(\xi,\mu)\, P_{qq}\left(\frac{x}{\xi}\right)$$

where to find a solution we need to impose initial conditions: we need PDFs at a certain initial scale Q_0.

- We choose a factorization scheme (e.g. MSbar), an order in perturbation theory (e.g. LO, NLO, NNLO) and a starting scale Q_0 where pQCD applies (typically > 1 or 2 GeV)

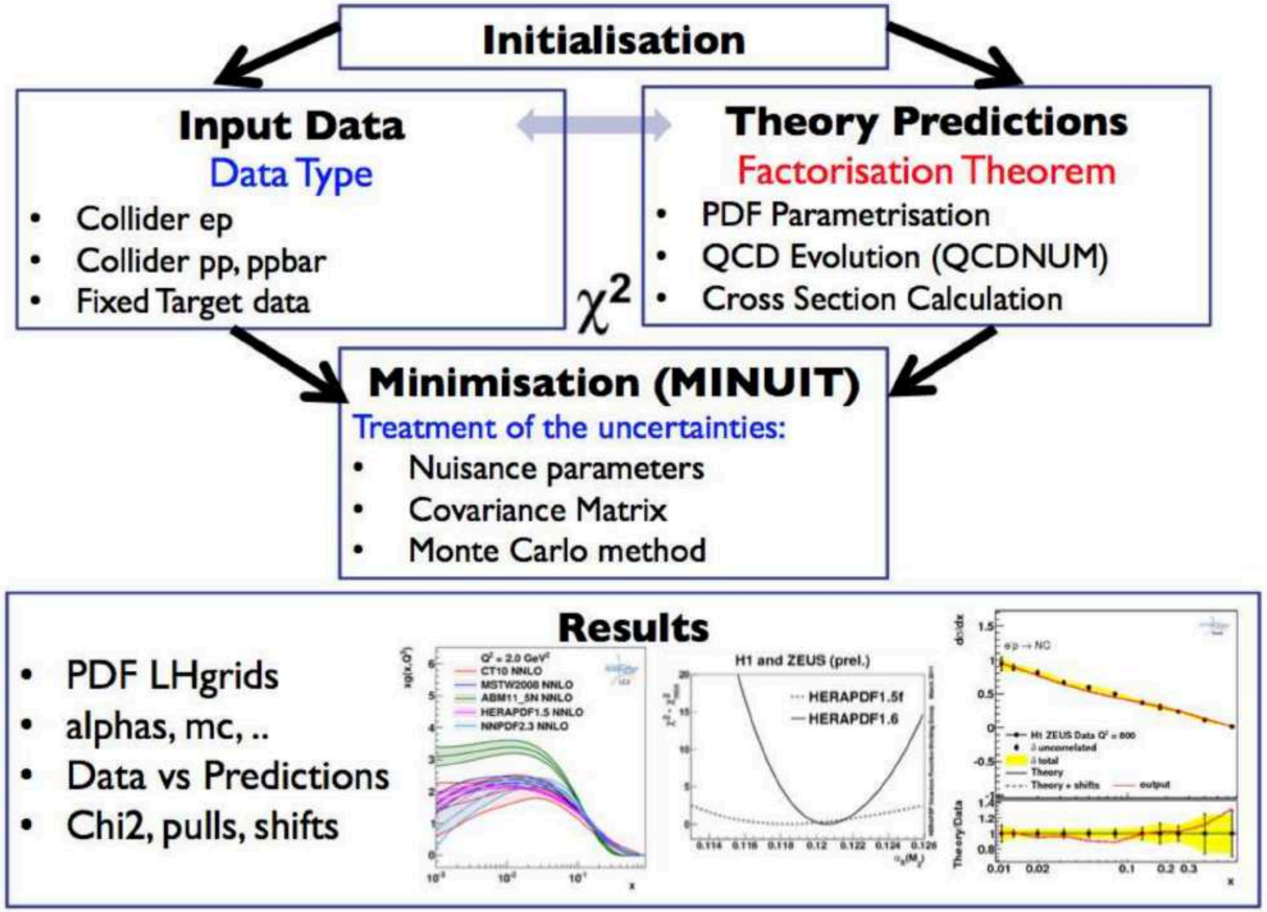
- parametrize the quark and gluon distributions at Q_0. For example:

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

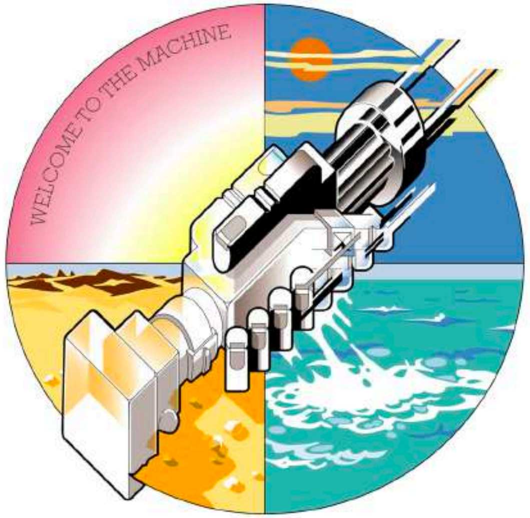
The parameters are determined from a fit to world data of hadronic collisions

Global QCD analyses combine:

- Advanced QCD theory
- Precise measurements from HEP hadronic world data
- Sophisticated statistical methods



Welcome to the fitting machine!

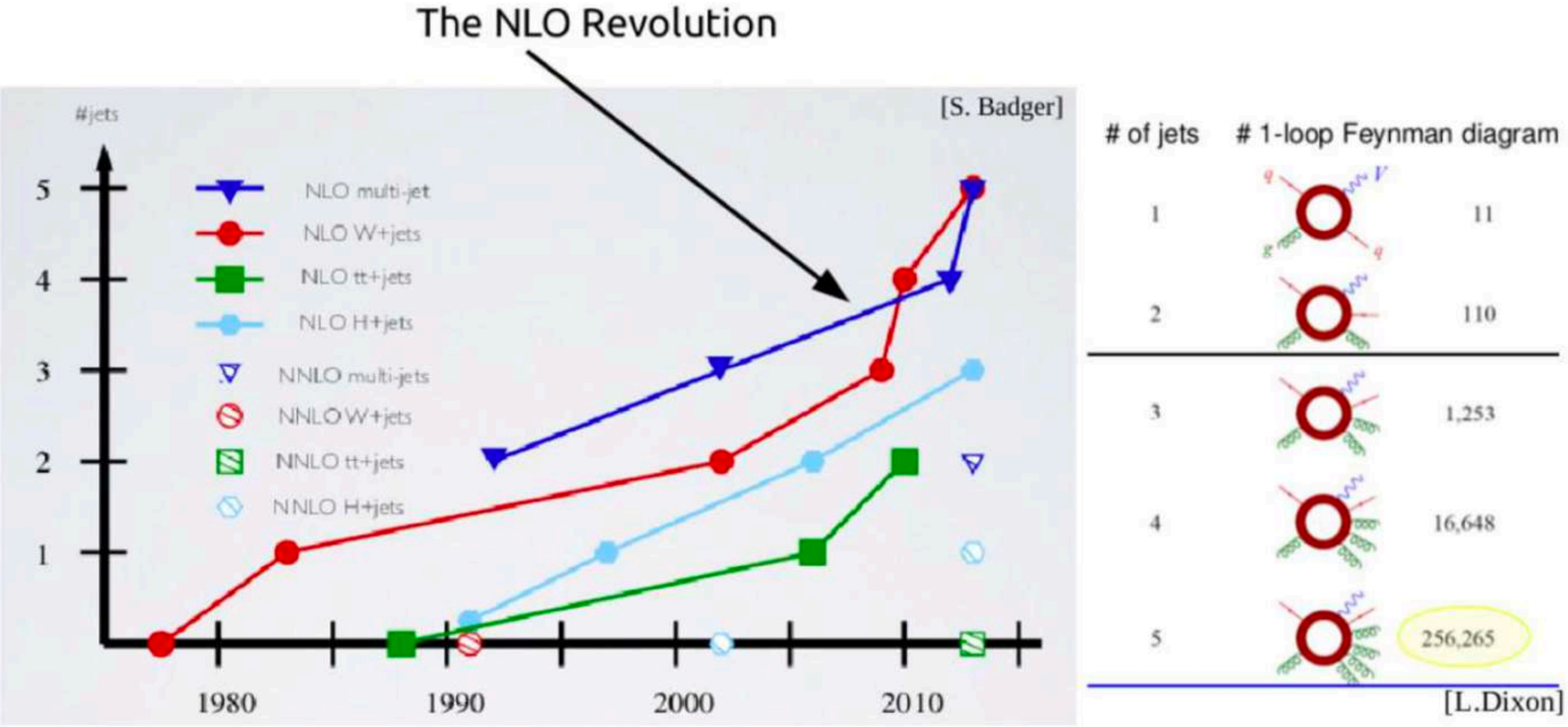


LHC and PDFs in the NNLO QCD era

The increasing accuracy of the current data and the LHC unprecedented energies pushed the high-energy physics community towards a new realm of precision calculations:

- Enormous progress in perturbative NNLO QCD calculations (e.g. unitarity based methods),
- semi-automated calculations of multi-leg NLO processes,
- NLO calculation of complex multi-leg processes such as for the production of vector boson plus 5 jets (e.g. $W + 5$ jets; $H+3$ jets),
- theoretical progress in the combination of the fixed-order results with a parton shower codes,
- rapid developments of very sophisticated tools for phenomenology

Perturbative QCD loop revolution



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

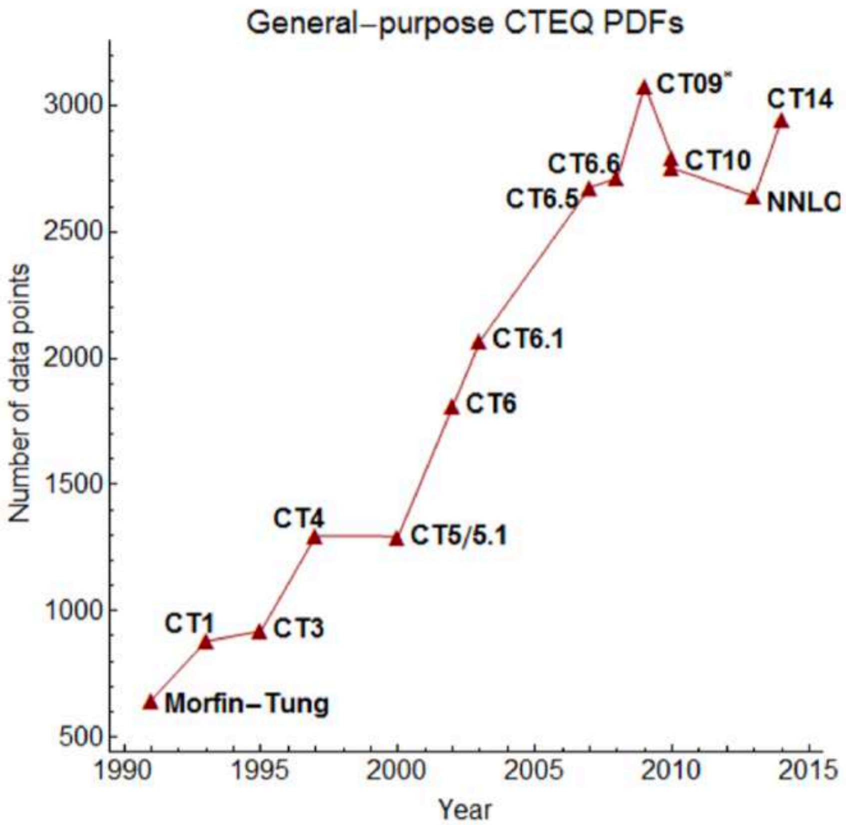
To make use of it, accuracy of PDFs $f_{a/p}(x, \mu)$ must keep up

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
CT18:	+new high precision LHC Run-I measurements

How does a PDF global analysis work?

- What are the most relevant data?
- What are the constraints on PDFs?
- How do we choose the initial scale parametrization?
- A bit of statistics

Data sets included in the CT18 global analysis

ID#	Experimental data set		$N_{pt,n}$
160	HERAI+II 1 fb^{-1}, H1 and ZEUS NC and CC $e^{\pm}p$ reduced cross sec. comb.	[11]	1120
101	BCDMS F_2^p	[12]	337
102	BCDMS F_2^d	[13]	250
104	NMC F_2^d/F_2^p	[14]	123
108	CDHSW† F_2^p	[15]	85
109	CDHSW† F_3^p	[15]	96
110	CCFR F_2^p	[16]	69
111	CCFR xF_3^p	[17]	86
124	NuTeV $\nu\mu\mu$ SIDIS	[18]	38
125	NuTeV $\bar{\nu}\mu\mu$ SIDIS	[18]	33
126	CCFR $\nu\mu\mu$ SIDIS	[19]	40
127	CCFR $\bar{\nu}\mu\mu$ SIDIS	[19]	38
145	H1 σ_r^b	[20]	10
147	Combined HERA charm production	[21]	47
169	H1 F_L	[22]	9
201	E605 Drell-Yan process	[23]	119
203	E866 Drell-Yan process $\sigma_{pd}/(2\sigma_{pp})$	[24]	15
204	E866 Drell-Yan process $Q^3 d^2\sigma_{pp}/(dQdx_F)$	[25]	184
225	CDF Run-1 electron A_{ch}, $p_{T\ell} > 25$ GeV	[26]	11
227	CDF Run-2 electron A_{ch}, $p_{T\ell} > 25$ GeV	[27]	11
234	DØ Run-2 muon A_{ch}, $p_{T\ell} > 20$ GeV	[28]	9
260	DØ Run-2 Z rapidity	[29]	28
261	CDF Run-2 Z rapidity	[30]	29
266	CMS 7 TeV 4.7 fb^{-1}, muon A_{ch}, $p_{T\ell} > 35$ GeV	[31]	11
267	CMS 7 TeV 840 pb^{-1}, electron A_{ch}, $p_{T\ell} > 35$ GeV	[32]	11
268	ATLAS 7 TeV 35 pb^{-1} W/Z cross sec., A_{ch}	[33]	41
281	DØ Run-2 9.7 fb^{-1} electron A_{ch}, $p_{T\ell} > 25$ GeV	[34]	13
504	CDF Run-2 inclusive jet production	[35]	72
514	DØ Run-2 inclusive jet production	[36]	110

Neutral current DIS: the backbone of PDFs

Neutrino DIS & di-muon

Drell-Yan data and charge asymmetries

New high-precision data from LHC Run-I

ID#	Experimental data set		$N_{pt,n}$
245	LHCb 7 TeV 1.0 fb^{-1} W/Z forward rapidity cross sec.	[37]	33
246	LHCb 8 TeV 2.0 fb^{-1} $Z \to e^-e^+$ forward rapidity cross. sec.	[38]	17
248	ATLAS‡ 7 TeV 4.6 fb^{-1}, W/Z combined cross sec.	[39]	34
249	CMS 8 TeV 18.8 fb^{-1} W cross sec. and A_{ch}	[40]	11
250	LHCb 8 TeV 2.0fb^{-1} W/Z cross sec.	[41]	34
251	ATLAS 8 TeV 20.3 fb^{-1} single diff. high-mass cross sec.	[42]	12
253	ATLAS 8 TeV 20.3 fb^{-1}, Z p_T cross sec.	[43]	27
542	CMS 7 TeV 5 fb^{-1}, single incl. jet cross sec., $R = 0.7$ (extended in y)	[44]	158
544	ATLAS 7 TeV 4.5 fb^{-1}, single incl. jet cross sec., $R = 0.6$	[45]	140
545	CMS 8 TeV 19.7 fb^{-1}, single incl. jet cross sec., $R = 0.7$, (extended in y)	[46]	185
573	CMS 8 TeV 19.7 fb^{-1}, $t\bar{t}$ norm. double-diff. top p_T & y cross sec.	[47]	16
580	ATLAS 8 TeV 20.3^{-1}, $t\bar{t}$ p_T^t and $m_{t\bar{t}}$ abs. spectrum	[48]	15

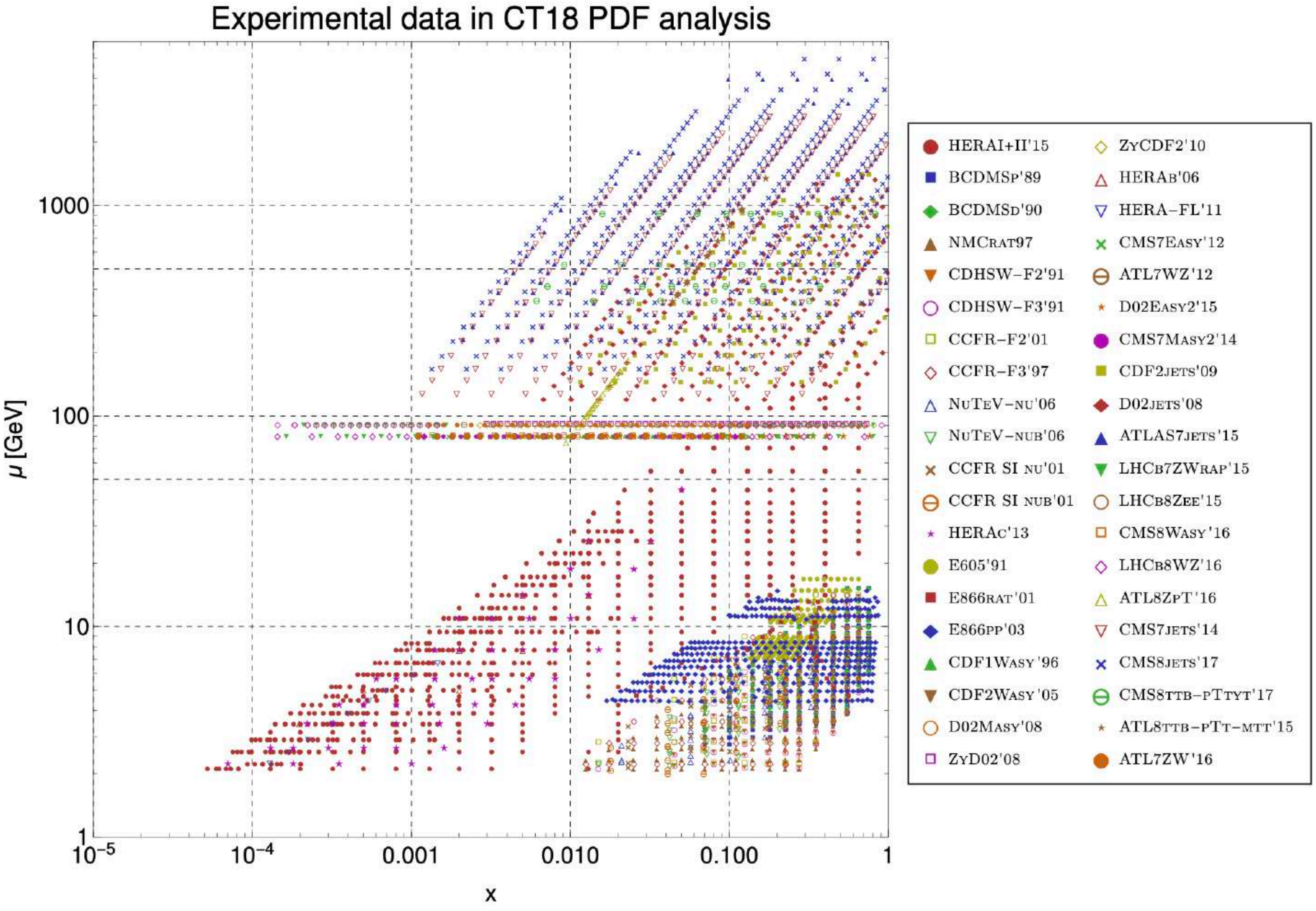


FIG. 1. The CT18 data set, in x and Q plane.

Constraints on PDFs: a few examples

Process	Sensitivity
Drell-Yan	Flavour decomposition of the sea, u_v, d_v, γ PDF
W+charm	Strange PDF
Jets	High-x gluon PDF
Photon	Medium-x gluon PDF
Top pair	Medium- and high-x gluon PDF

Of course, there are many other measurements which are not yet fully exploited

DIS high precision data from HERA

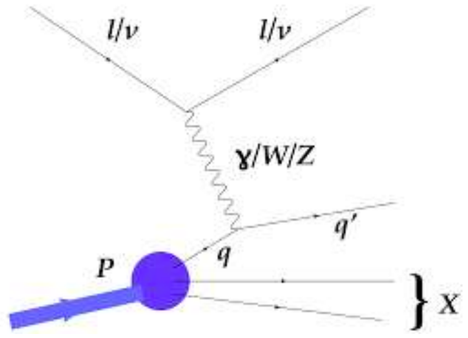
$$\sigma^{\pm}_{r,\mathrm{NC}} = \frac{\mathrm{d}^2\sigma^{e^{\pm}p}_{\mathrm{NC}}}{\mathrm{d}x_{\mathrm{Bj}}\mathrm{d}Q^2} \cdot \frac{Q^4 x_{\mathrm{Bj}}}{2\pi\alpha^2 Y_+} = \tilde{F}_2 \mp \frac{Y_-}{Y_+} x\tilde{F}_3 - \frac{y^2}{Y_+}\tilde{F}_{\mathrm{L}}$$

Reduced cross section

$$Y_{\pm} = 1 \pm (1-y)^2$$

$$\begin{aligned} \tilde{F}_2 &= F_2 - \kappa_Z v_e \cdot F_2^{\gamma Z} + \kappa_Z^2 (v_e^2 + a_e^2) \cdot F_2^Z , \\ \tilde{F}_{\mathrm{L}} &= F_{\mathrm{L}} - \kappa_Z v_e \cdot F_{\mathrm{L}}^{\gamma Z} + \kappa_Z^2 (v_e^2 + a_e^2) \cdot F_{\mathrm{L}}^Z , \\ x\tilde{F}_3 &= -\kappa_Z a_e \cdot xF_3^{\gamma Z} + \kappa_Z^2 \cdot 2 v_e a_e \cdot xF_3^Z , \end{aligned}$$

$$\begin{aligned} (F_2, F_2^{\gamma Z}, F_2^Z) &\approx [(e_u^2, 2e_u v_u, v_u^2 + a_u^2)(xU + x\bar{U}) + (e_d^2, 2e_d v_d, v_d^2 + a_d^2)(xD + x\bar{D})] , \\ (xF_3^{\gamma Z}, xF_3^Z) &\approx 2[(e_u a_u, v_u a_u)(xU - x\bar{U}) + (e_d a_d, v_d a_d)(xD - x\bar{D})] , \end{aligned}$$



The quark flavors are probed with different weights

$$xU = xu + xc , \qquad x\bar{U} = x\bar{u} + x\bar{c} , \qquad xD = xd + xs , \qquad x\bar{D} = x\bar{d} + x\bar{s} ,$$

$$xu_v = xU - x\bar{U} , \qquad xd_v = xD - x\bar{D}$$

<u>Precise determination of valence PDFs and total sea</u>

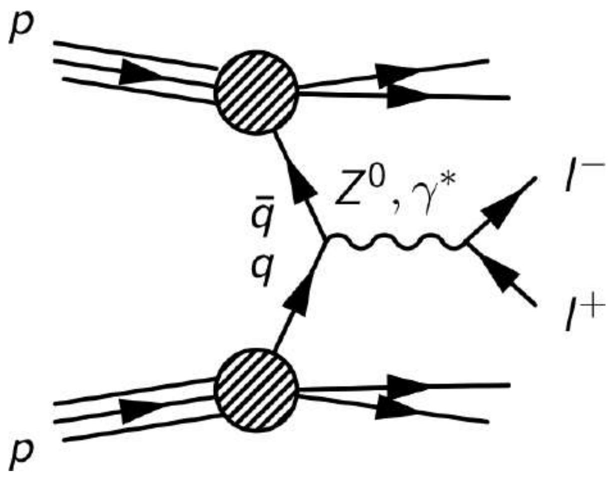
$$xF_3^{\gamma Z} \approx \frac{x}{3}(2u_v + d_v)$$

$$\sigma^+_{r,\mathrm{CC}} \approx (x\bar{U} + (1-y)^2 xD) , \qquad \sigma^-_{r,\mathrm{CC}} \approx (xU + (1-y)^2 x\bar{D})$$

Drell-Yan Measurements

$$\frac{d\sigma}{dQ^2 dy} = \frac{4\pi\alpha^2}{9Q^2 s} \sum_i e_i^2 \left[q_i(x_a, Q^2)\bar{q}_i(x_b, Q^2) + a \leftrightarrow b \right]$$

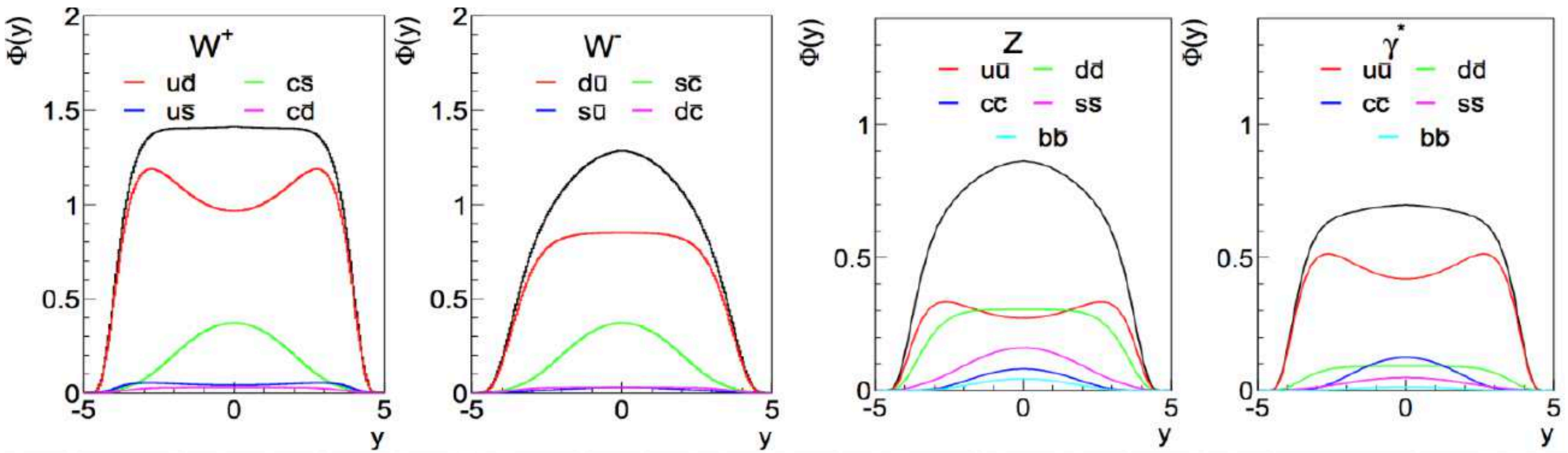
LO Xsec: photon dominated when away from Z and W resonances



Important to constrain PDF combinations and in particular anti-quarks

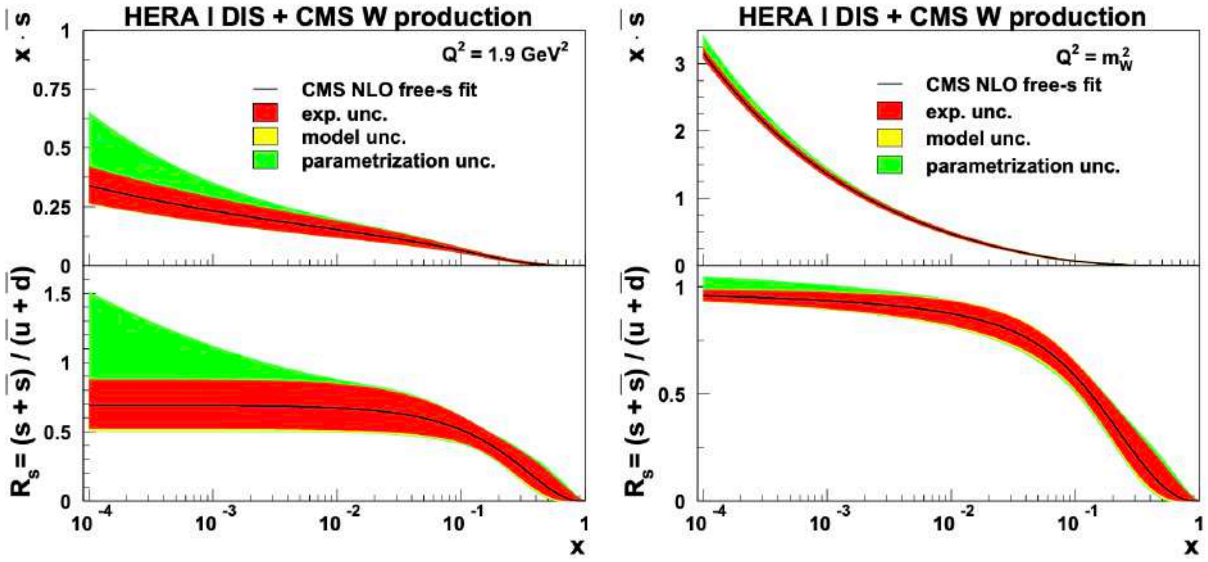
$$\frac{d\sigma^W}{dy} = \frac{\sqrt{2}\pi G_F m_W^2}{3s} \sum_{i,j} |V_{ij}^{\mathrm{CKM}}| \left[q_i(x_a, Q^2)\bar{q}_j(x_b, Q^2) + a \leftrightarrow b \right]$$

$$\frac{d\sigma^Z}{dy} = \frac{\sqrt{2}\pi G_F m_Z^2}{3s} \sum_i (V_i^2 + A_i^2) \left[q_i(x_a, Q^2)\bar{q}_i(x_b, Q^2) + a \leftrightarrow b \right]$$

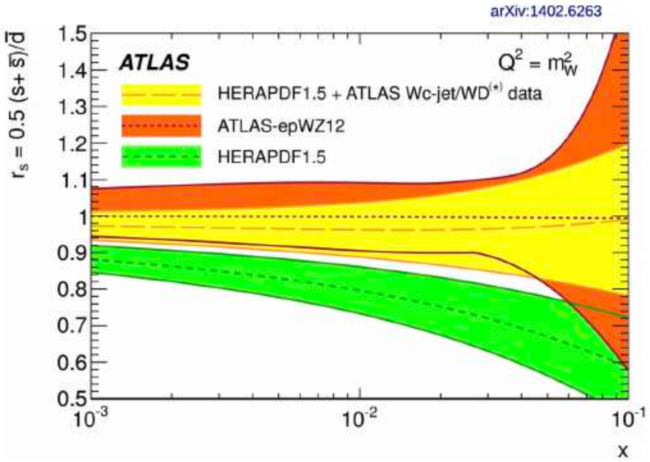


W + Charm production at the LHC

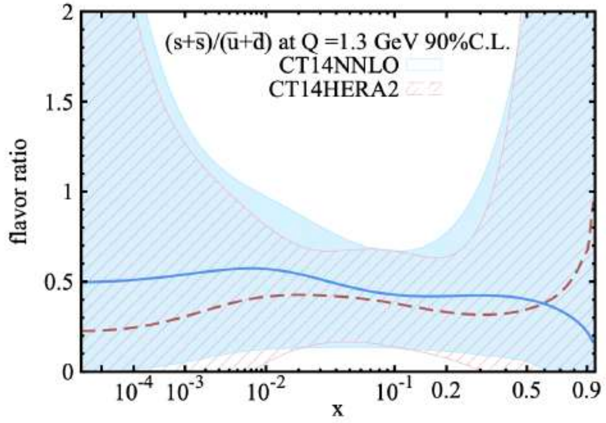
- W + single charm at LO is mainly produced by *gs* → *W+c*
- Sensitive to the strange PDF
- Exploit charge correlation to distinguish between *W + c* and *W + (g → c$\bar{c}$)*



The CMS coll. (PRD90 2014)



ATLAS Coll (2013)

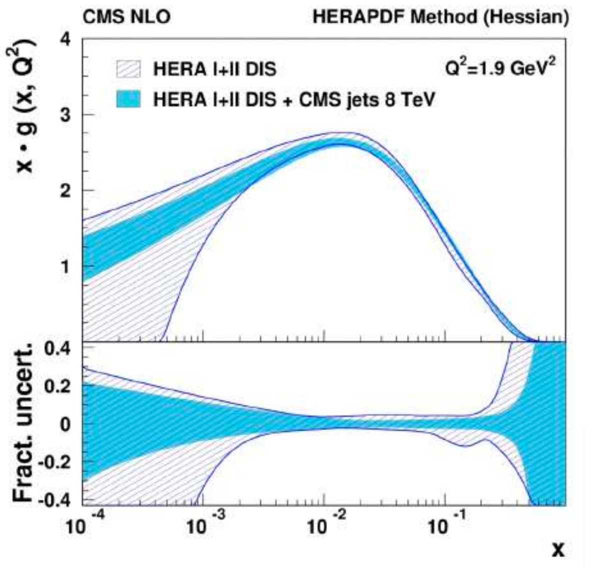


CT14HERA2
PRD(2017)

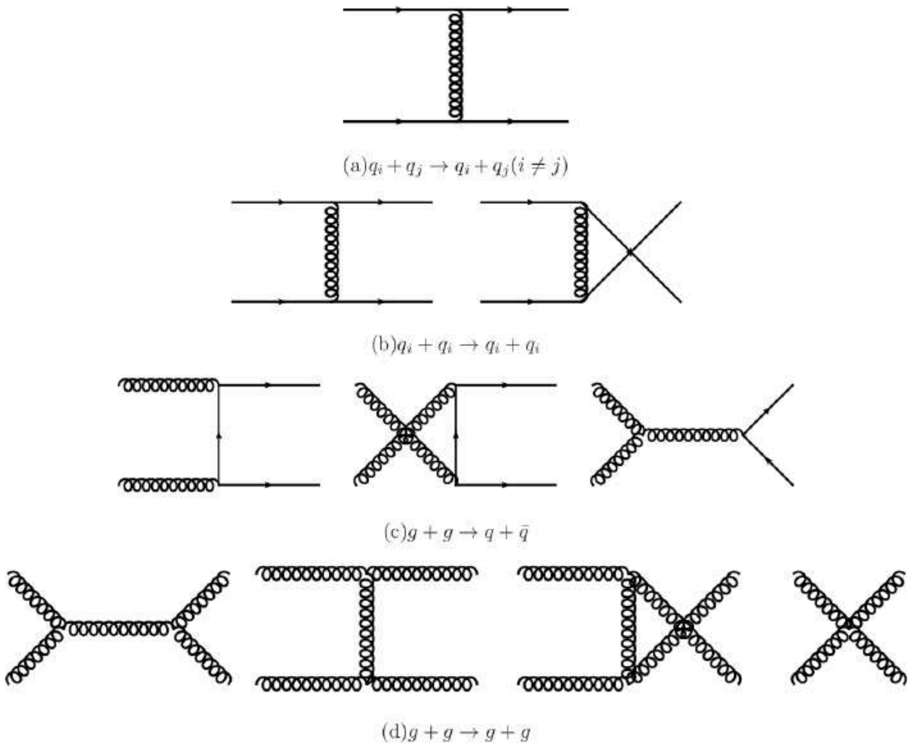
Jet production at hadron colliders

Important to constrain the gluon PDF at large-x

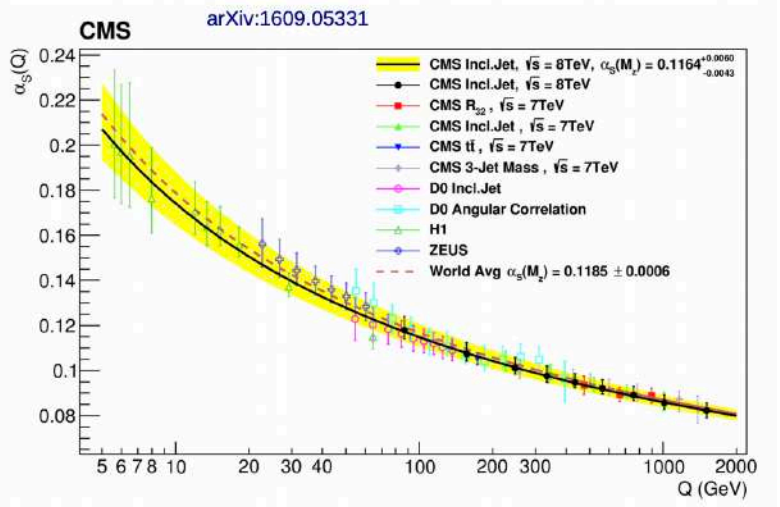
- combined with low-x constraints on gluon PDF from DIS and with sum rules one has strong constraints on the gluon PDF



JHEP 1703 (2017) 156



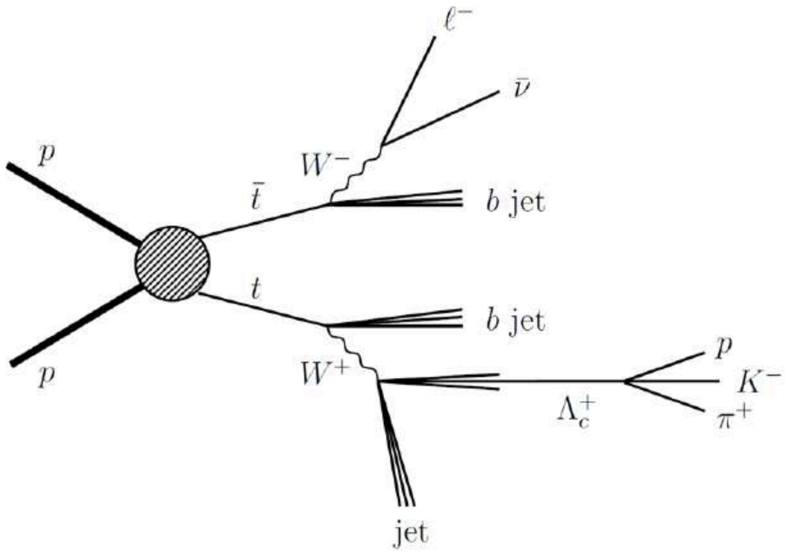
Also crucial to probe the running of the qcd coupling at high energy

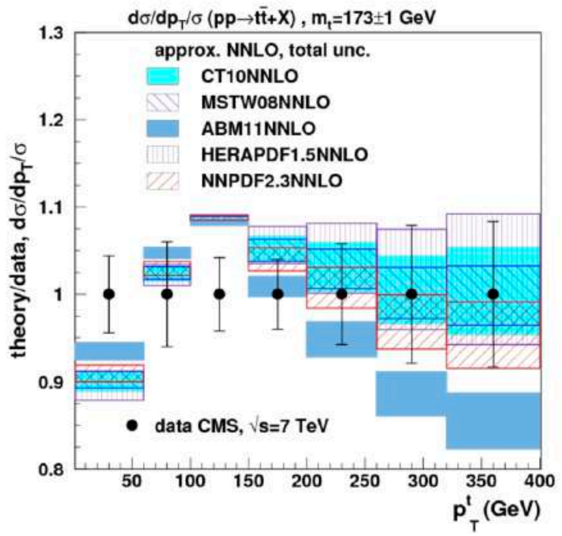


arXiv: 1609.05331

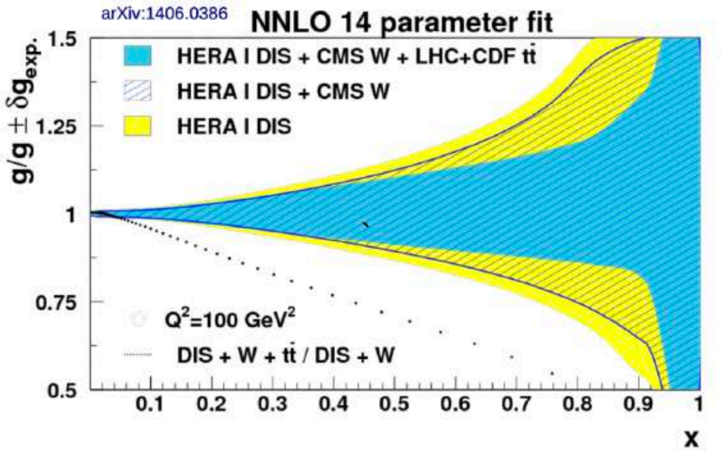
Top quark production at hadron colliders

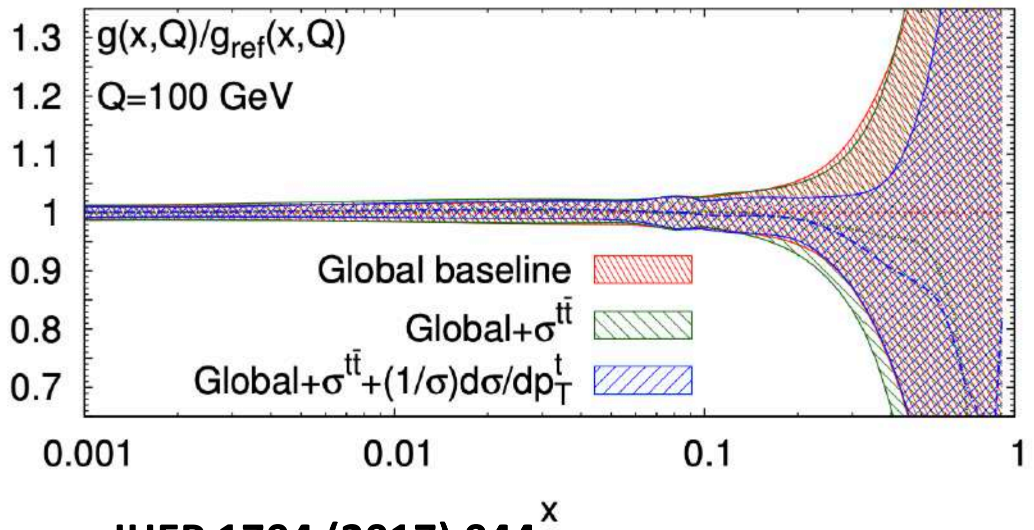
Top-quark pair production differential Xsec measurements constrain the gluon at large x (weaker than jet data due to statistics)





JHEP 1501 (2015) 082





JHEP 1704 (2017) 044

Initial scale parametrization

PDF parametrizations for $f_{a/p}(x, Q_0)$ must be “flexible just enough” to reach agreement with the data, without violating QCD constraints (sum rules, positivity, ...) or reproducing random fluctuations.

They are constructed by using an ansatz based on models for the two asymptotic behaviors x → 0 and x → 1

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

- Regge-like behavior $x \to 0 \Longrightarrow f \propto x^{a_1}$
- Quark counting rules $x \to 1 \Longrightarrow f \propto (1-x)^{a_2}$
- F(x, a3,…,an) interpolates between the small and large x regions

Initial scale parametrization: CT14/CT18 fit

$$x\, f_a(x, Q_0) \;=\; x^{a_1}\,(1-x)^{a_2}\,P_a(x)$$

$$P_{u_v} \;=\; c_0 \,+\, c_1\, y \,+\, c_2\, y^2 \,+\, c_3\, y^3 \,+\, c_4\, y^4$$

Instead of using coefficients c1,…, c4, we use a linear combination of Bernstein polynomials

$$P_{u_v} \;=\; d_0\, p_0(y) \,+\, d_1\, p_1(y) \,+\, d_2\, p_2(y) \,+\, d_3\, p_3(y) \,+\, d_4\, p_4(y)$$

$$\begin{aligned} p_0(y) &= (1-y)^4, \\ p_1(y) &= 4\,y\,(1-y)^3, \\ p_2(y) &= 6\,y^2\,(1-y)^2, \\ p_3(y) &= 4\,y^3\,(1-y), \\ p_4(y) &= y^4. \end{aligned}$$

Features:

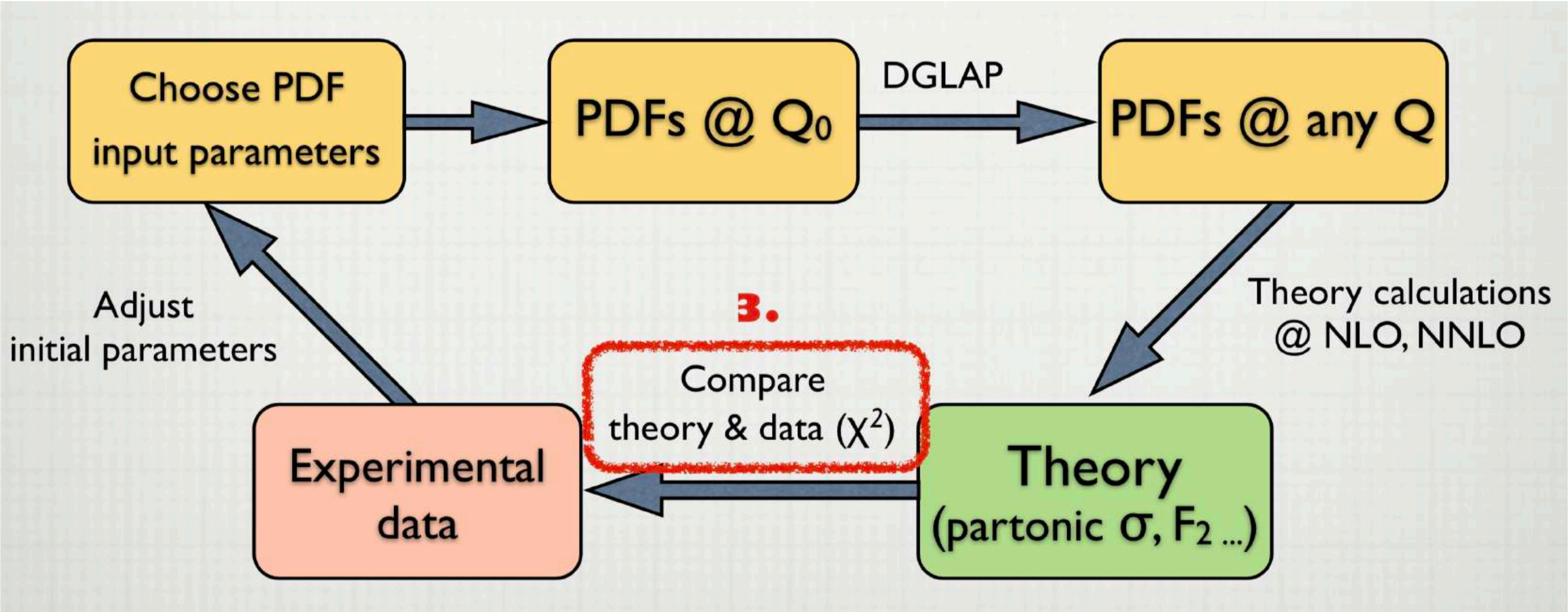
- Reduces rapid variations at the boundaries x = 0,1
- Reduces correlations between coefficients d1,…,d4

The CT14 global analysis uses 28 parameters to parametrize the PDFs at Q_0

Flow of a PDF analysis

1. Select valid experimental data
2. Assemble most precise theoretical cross sections and verify their mutual consistency
3. Choose the functional form for PDF parametrizations
4. Implement a procedure to handle nuisance parameters
5. (>200 sources of correlated experimental errors)
6. Perform a fit to determine the new central value PDFs and their uncertainties

Flow of a PDF analysis



Modern NNLO PDF analyses challenges

- full NNLO calculation are lengthy and CPU time consuming
- PDF fits go through a large (~10^3) num. of iterations before converging
- Need a tool to speed up Xsec calculation

fastNLO

fast pQCD calculations for hadron-induced processes

the APPLgrid project

Fast interpolation for the cross section evaluation in PDF fits: create and evaluate fast interpolation tables of pre-computed coefficients in perturbation theory for observables in hadron-induced processes.

Statistics: Hessian method

- find N parameters {a_i} of the PDF (N=28 in CT14/CT18) ensemble from N_{exp} experiments with N_λ correlated systematic errors in each experiment
- Each systematic error is associated with a random parameter λ_α, assumed to be distributed as a Gaussian distribution with unit dispersion
- The most likely combination of a_i and λ_α is found by minimizing the χ^2

$$\chi^2_E(\{a\},\{\lambda\}) = \chi^2_D + \chi^2_\lambda$$

$$\chi^2_D \equiv \sum_{k=1}^{N_{pt}} \frac{1}{s_k^2}\left(D_k - T_k - \sum_{\alpha=1}^{N_\lambda} \beta_{k,\alpha}\lambda_\alpha\right)^2$$

Correlated systematic shifts

D_k and $T_k(ai)$ are the data and theory values at each point.

$$s_k = \sqrt{\sigma^2_{stat} + \sigma^2_{sys,uncor}}$$

is the total statistical + systematical **uncorrelated** error

$$\chi^2_\lambda \equiv \sum_{\alpha=1}^{N_\lambda} \lambda^2_\alpha$$

Penalty term for deviations of λ_α from their expected λ_α = 0 values

Statistics: Hessian method

Hessian matrix

After diagonalization

$$\chi^2(a) = \chi_0^2 + \frac{1}{2}\frac{\partial\chi^2}{\partial a_i \partial a_j}(a-a_0)_i(a-a_0)_j + \cdots \to \chi_0^2 + \sum_i z_i^2$$

$$y_i = a_i - a_i^0$$

$$\chi^2 = \chi_0^2 + \sum_{i,j} H_{ij}\, y_i\, y_j\,,$$

$$H_{ij} = \frac{1}{2}\left(\frac{\partial^2\chi^2}{\partial y_i\, \partial y_j}\right)_0\,,$$

$$\sum_j H_{ij}\, v_{jk} = \epsilon_k\, v_{ik}$$ Eigenvalue equation

$$\sum_i v_{ij}\, v_{ik} = \delta_{jk}\,.$$ Orthonormality

$$z_i = \sqrt{\epsilon_i}\sum_j y_j\, v_{ji}$$ Change of basis in terms of the eigenvalues

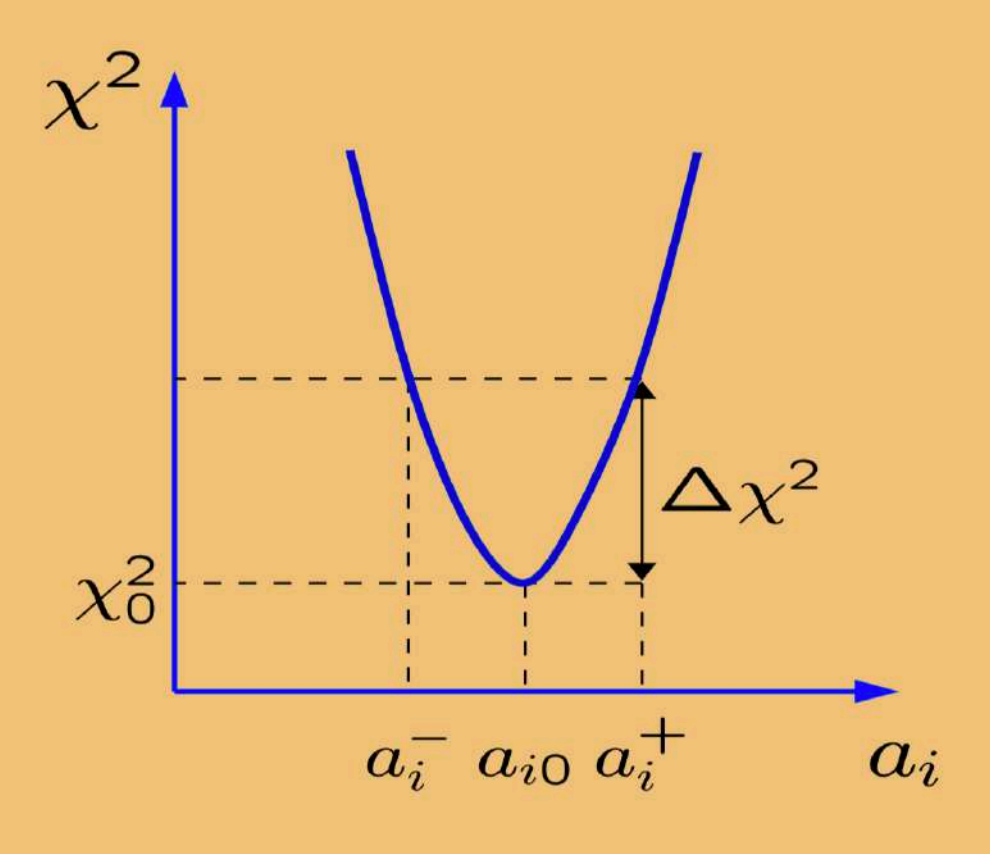
$$\Delta\chi^2 = \chi^2 - \chi_0^2 = \sum_i z_i^2$$

the surfaces of constant χ2 are spheres in z_i space, with Δχ2 the squared distance from the minimum.

Tolerance Criterion

- We need to establish a confidence region for {ai} for a given tolerated increase in χ2

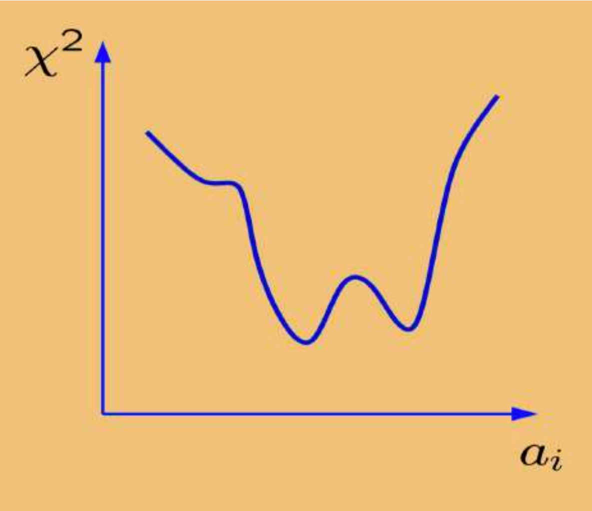
- In the ideal case of perfectly compatible Gaussian errors, 68% c.l. on a physical observable X corresponds to Δχ2 = 1 independently of the number N of PDF parameters



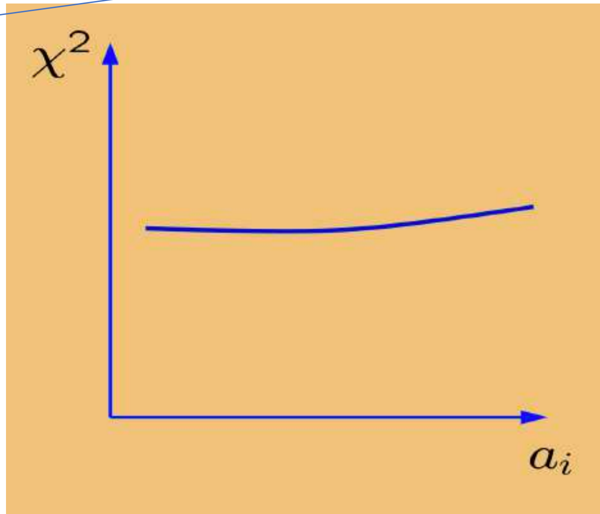
$$\Delta\chi^2 = 1$$ Ideal choice -> 68% CL -> 1 sigma

Real-life case

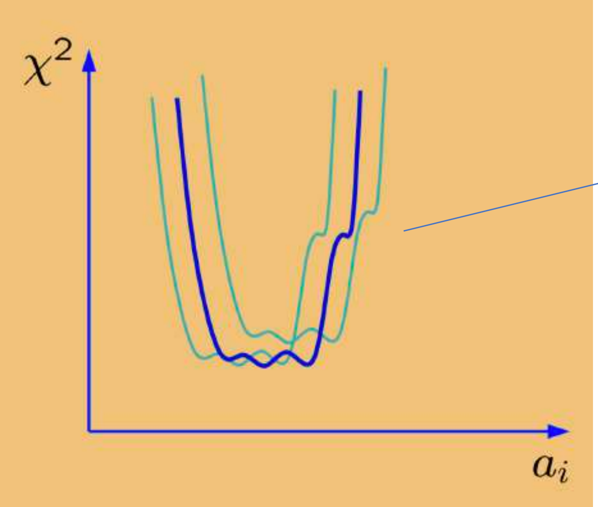
Pitfalls



Disagreements between
the experiments

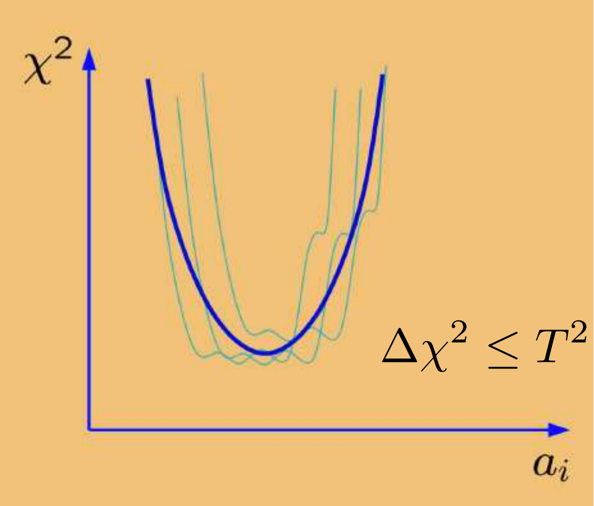


Flat directions: unconstrained combinations
of PDF parameters. PDF error sets unreliable.



Actual function

- a well pronounced global minimum
- weak tensions between data sets in the vicinity of χ_0^2 (mini-landscape)
- some dependence on assumptions about flat directions



PDF induced errors on a physical observable

The Hessian method is used in CT PDFs to propagate the PDF uncertainty into QCD predictions.

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

$$\Delta X = X - X_0 \cong \sum_i \frac{\partial X}{\partial y_i} y_i = \sum_i X_i z_i$$

$$X_i \equiv \frac{\partial X}{\partial z_i}$$

$$(\Delta X)^2 = (X \cdot Z)^2 = \Delta\chi^2 \sum_i X_i^2$$

Since χ2 increases uniformly in all directions in z-space, the gradient vector X_i gives the direction in which the physical observable X varies fastest with increasing χ2 .
The maximum deviation in X for a given increase in χ2 is therefore obtained by the dot product of the gradient vector X_i and a displacement vector Z_i in the same direction with length $\sqrt{\Delta\chi^2}$ for example:

$$Z_i = X_i \sqrt{\Delta\chi^2 / \Sigma_j X_j^2}$$

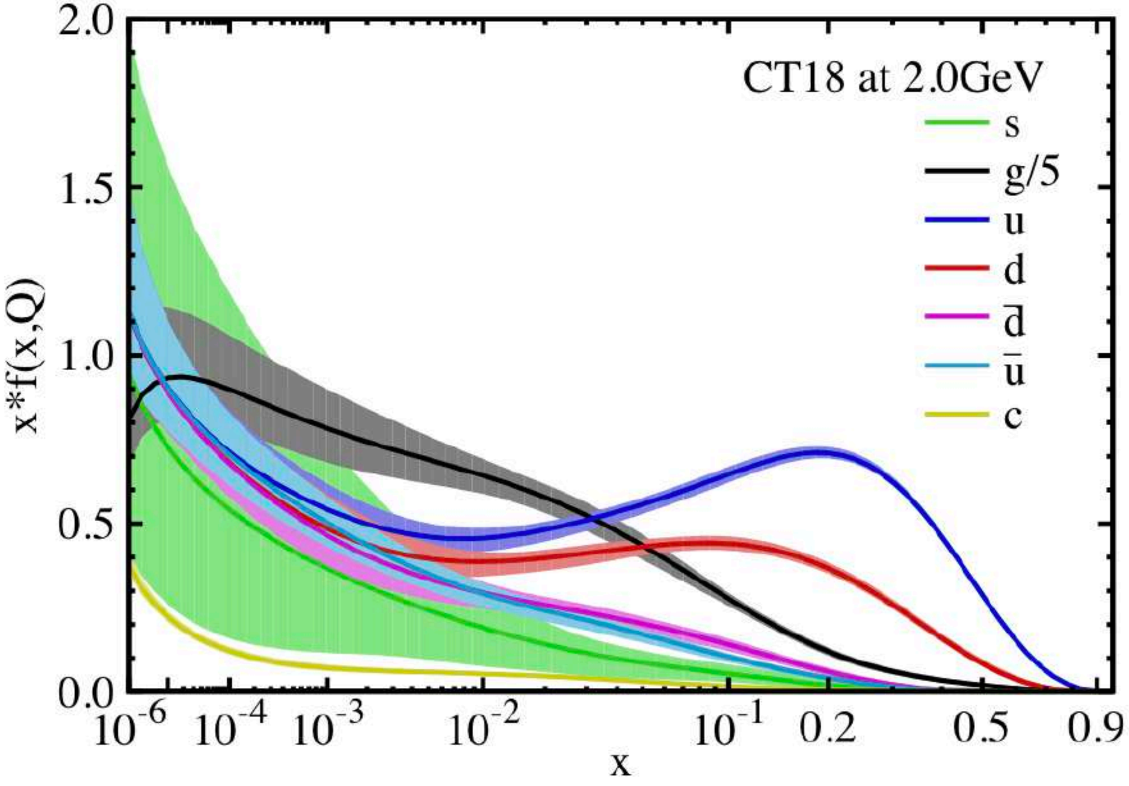
$$X_i^{\pm}(z) = X_i^{\pm}(0, 0, \ldots, \pm\sqrt{\Delta\chi^2}, \ldots, 0, 0)$$

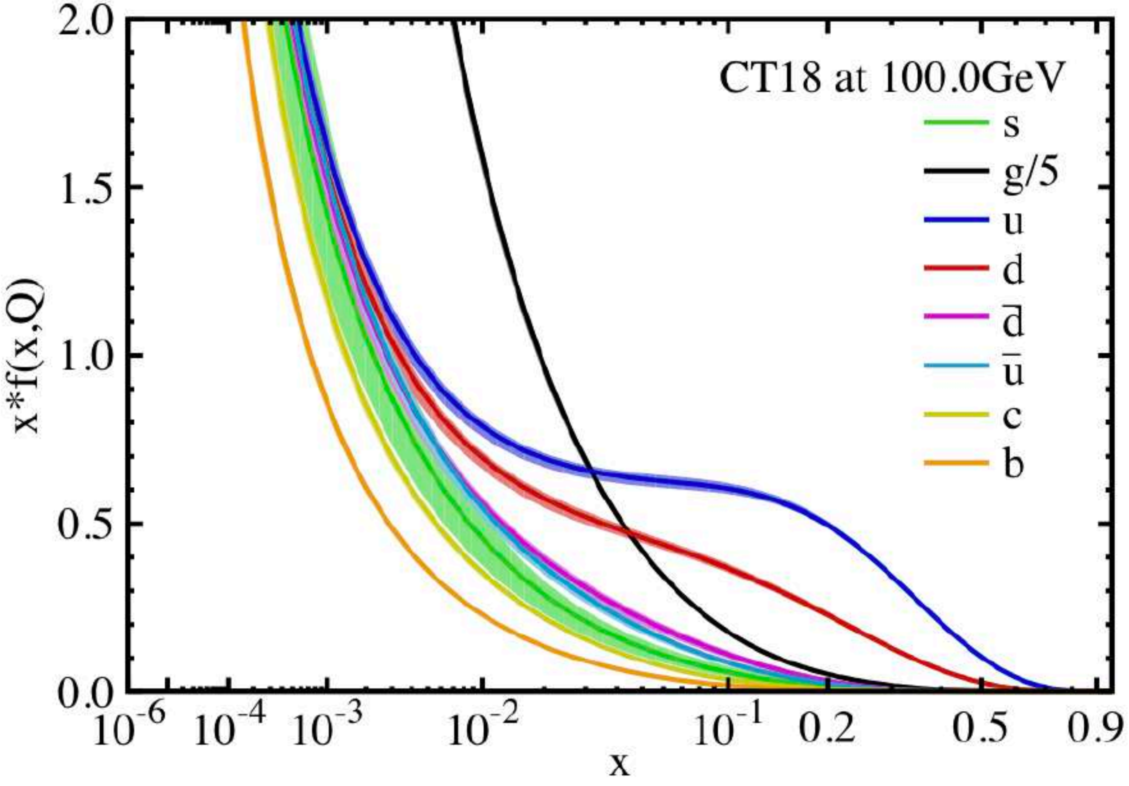
Error PDF

$$(\Delta\sigma)^2 \approx \frac{1}{4} \sum_i^{N_p} \left(\sigma(X_i^+) - \sigma(X_i^-)\right)^2$$

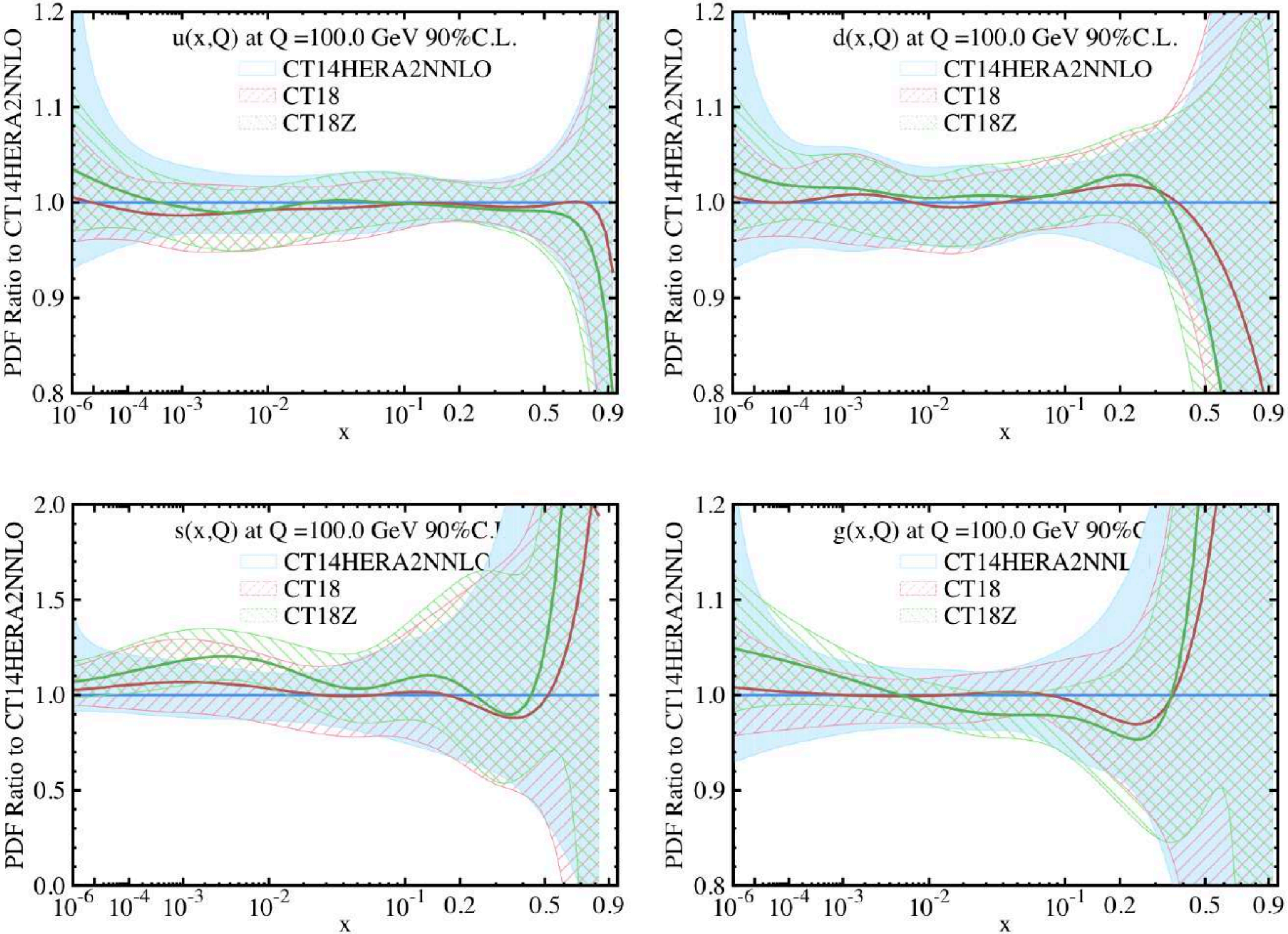
Symmetric formula

Most recent results of the CT18 global PDF analysis

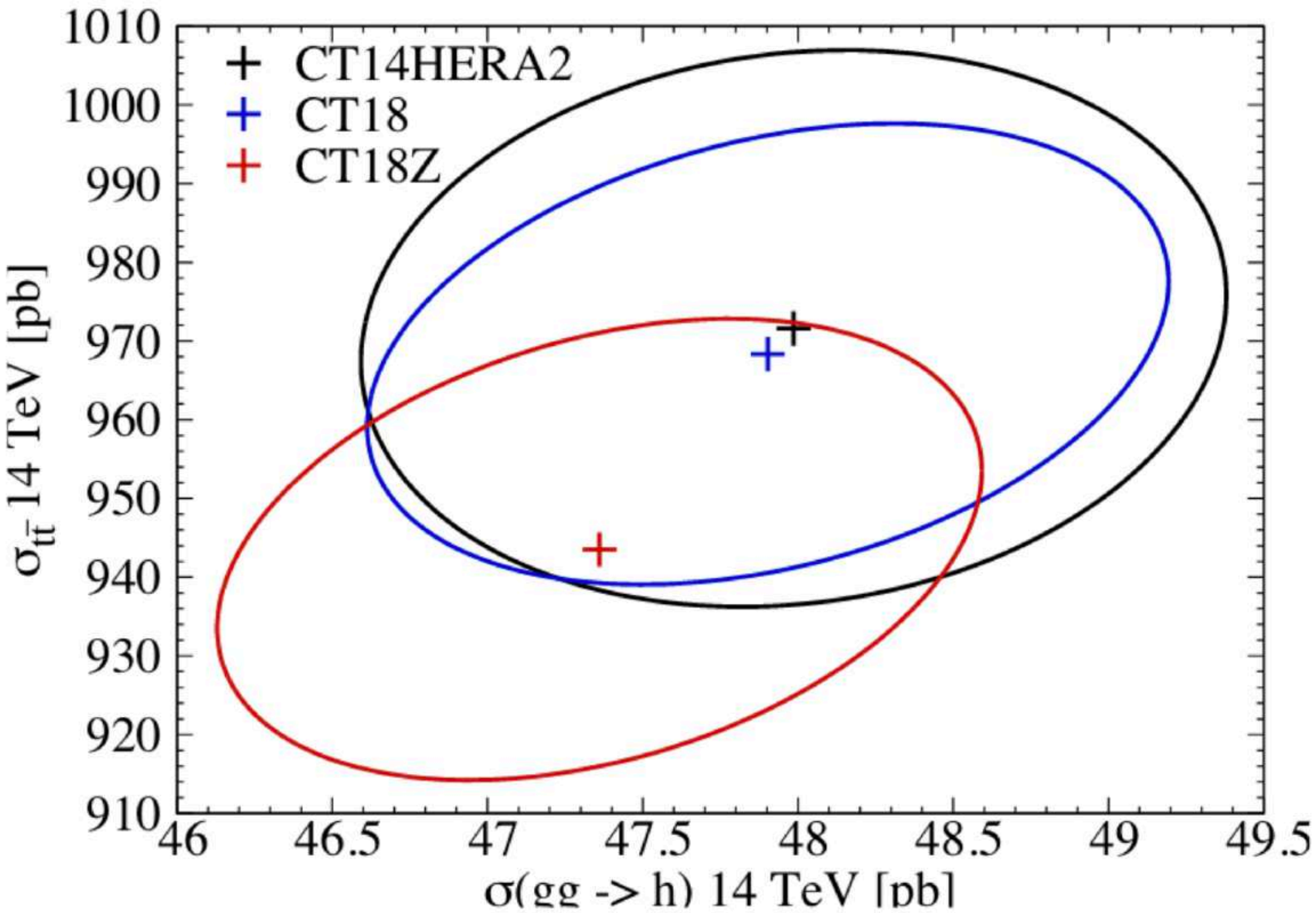




Some CT18 individual flavors



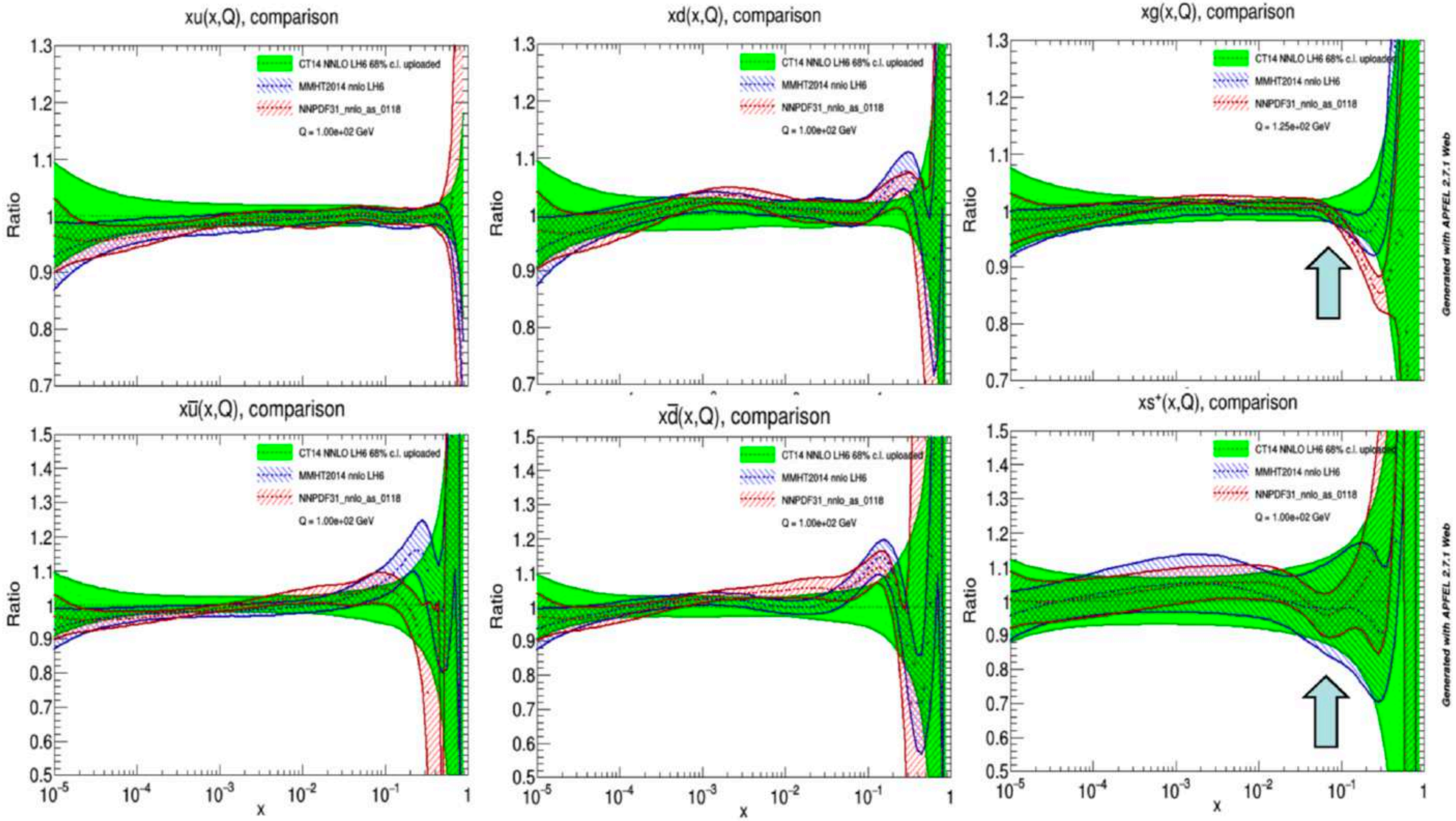
Induced PDF uncertainties on physical cross sections



Unpolarized proton PDFs: where we are now

- CT (CTEQ-TEA)
- MMHT
- NNPDF
- ABM
- CTEQ-Jlab
- HeraFitter/xFitter

CT14, MMHT14, NNPDF3.1 PDFs



Central PDFs of 3 sets are compatible. The NNPDF3.1 uncertainty is moderately reduced on $g(x, Q)$ at $x > 0.05$, $s + \bar{s}$ at all x. The effect of the LHC data on the error bands does not exceed dependence on the definition of the PDF uncertainty (CT vs. MMHT vs. NNPDF).

Origin of differences

1. Corrections of wrong or outdated assumptions

lead to significant differences between new (≈post-2014) and old (≈pre-2014) PDF sets

- inclusion of NNLO QCD, heavy-quark hard scattering contributions
- relaxation of ad hoc constraints on PDF parametrizations
- improved numerical approximations

Origin of differences

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 $\alpha_s(M_Z)$,m_c,m_b, , factorization scales, parametrizations for PDF's, higher-twist terms, nuclear effects,...

leads to non-negligible differences between the newest PDF sets

Where do we get PDFs from?

LHAPDF 6.2.3

Main Page | Related Pages | Namespaces ▾ | Classes ▾ | Files ▾ | Examples | Search

LHAPDF Documentation

Introduction

LHAPDF is a general purpose C++ interpolator, used for evaluating PDFs from discretised data files. Previous versions of LHAPDF were written in Fortran 77/90 and are documented at http://lhapdf.hepforge.org/lhapdf5/.

LHAPDF6 vastly reduces the memory overhead of the Fortran LHAPDF (from gigabytes to megabytes!), entirely removes restrictions on numbers of concurrent PDFs, allows access to single PDF members without needing to load whole sets, and separates a new standardised PDF data format from the code library so that new PDF sets may be created and released easier and faster. The C++ LHAPDF6 also permits arbitrary parton contents via the standard PDG ID code scheme, is computationally more efficient (particularly if only one or two flavours are required at each phase space point, as in PDF reweighting), and uses a flexible metadata system which fixes many fundamental metadata and concurrency bugs in LHAPDF5.

Compatibility routines are provided as standard for existing C++ and Fortran codes using the LHAPDF5 and PDFLIB legacy interfaces, so you can keep using your existing codes. But the new interface is much more powerful and pleasant to work with, so we think you'll want to switch once you've used it!

LHAPDF6 is documented in more detail in http://arxiv.org/abs/1412.7420

Table of Contents

https://lhapdf.hepforge.org/

Conclusions

- Why all these efforts in PDFs determination?
- PDFs still represent one of the major sources of uncertainty in cross section determinations at hadron colliders
- LHC measurements reached a new realm of precision.
- This has to match with theory predictions in order to maximize the outcome of the LHC run II and after.
- It's critical to explore Higgs and top quark properties. They are doors to a better understanding of the EW sector and to search for new physics.

THANK YOU!