Recent developments in QCD global analyses of proton's PDFs

Marco Guzzi
in collaboration with the CTEQ-TEA (CT) group

The University of Manchester

"XLV International Symposium on Multiparticle Dynamics"

Intro

Investigation of the structure of the nucleon crucial for a multitude of high-energy physics programs.

Interpretation of experimental measurements at hadron colliders relies on the precise knowledge of fundamental QCD parameters and Parton Distribution Functions (PDFs) of the proton.

- **Discrimination of New Physics signals at the LHC crucially depends on precise knowledge of PDFs**
- **Global QCD analysis of PDFs is a vast topic, I'll focus on:**
  - **basic facts,**
  - **brief overview of the current status of modern PDFs,**
  - **aspects of the CT14 analysis.**
  - **selected CT14 recent results.**

Making a long story short...

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, lepton-hadron beams, or hadron targets, is a convolution product of long-distance non-perturbative contributions (PDFs) and short-distance infrared-safe perturbatively calculable quantities (hard scatterings cross sections) . For Drell-Yan process in the collinear limit we have (Collins Soper Sterman (1984), (1985))

$$\sigma(h_1 h_2 \to l^+ l^- + X) =$$
$$\sum_{a,b} \int_{x_1}^{1} d\xi_1 \int_{x_2}^{1} d\xi_2 \, f_{h_1\to a}(\xi_1, \alpha_s(\mu_R), \mu_R, \mu_F) f_{h_2\to b}(\xi_2, \alpha_s(\mu_R), \mu_R, \mu_F)$$
$$\times \hat{\sigma}^{ab}\left(\frac{x_1}{\xi_1}, \frac{x_2}{\xi_2}; \alpha_s(\mu_R), Q, \mu_F, \mu_R\right) + \mathcal{O}\left(\frac{\Lambda^2}{Q^2}\right), \quad (1)$$

$\mathcal{O}\left(\frac{\Lambda^2}{Q^2}\right)$ are subleading terms: higher twists, target corrections,...

Complicated objects

The formal definition of PDFs in QCD, contains all the complications of "real life": UV regulator in DR, gauge invariance Collins (2011)

$$f_{(0)\,j/h}(\xi) = \int \frac{dw^-}{2\pi}\, e^{-i\xi P^+ w^-} \langle P|\overline{\psi}_j^{(0)}(0, w^-, \mathbf{0}_\mathrm{T}) W(w^-, 0) \frac{\gamma^+}{2} \psi_j^{(0)}(0)|P\rangle_\mathrm{c}, \tag{2}$$

that is for quarks, where the Wilson-line factor is

$$W(w^-, 0) = P\left[e^{-ig_0 \int_0^{w^-} dy^- A^+_{(0)\alpha}(0,y^-,\mathbf{0}_\mathrm{T}) t_\alpha}\right]. \tag{3}$$

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. $\overline{\mathrm{MS}}$ scheme.

Scale dependence

In the collinear picture, the use of RG invariance tells us how to predict scale dependence or "evolution" of PDFs by renormalization group equations (RGE's) once the "initial conditions" are given. Parton evolution is obtained in terms of integro-differential equations known as DGLAP (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi) equations

$$\frac{d\ f_i(x,\mu_R,\mu_F)}{d\log\mu_F} = \sum_{j=q\bar{q},g}\int_x^1 \frac{dy}{y} P_{ij}\left(\frac{x}{y};\alpha_s,\mu_R,\mu_F\right) f_j(y,\mu_R,\mu_F)\,, \quad (4)$$

The evolution kernels or "splitting functions" P_{ij} are known at 3-loop for the unpolarized case. Moch, Vermaseren, Vogt (2004)

Universal objects

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 hard-scattering cross sections, the analytic structure of the PDFs cannot be predicted by perturbative QCD, but
has to be determined by comparing standard sets of cross sections, such as Eq. 1, to experimental measurements by using a variety of analytical and statistical methods.

For this reason PDFs are "data-driven" quantities.

PDFs for the LHC in the NNLO QCD era

The increasing accuracy of the current data and LHC run I 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),
- Understanding of jets substructure
- theoretical progress in the combination of the fixed-order results with a parton shower codes,
- rapid developments of very sophisticated tools for phenomenology such as:
  HERAFITTER platform, (see H. Pirumov talk),
  META and MC-H PDFs (see A. Buckley talk),
- LHC run II: next challenge for precision

STATUS

Unpolarized collinear PDFs at NLO, NNLO in QCD:

Recent (2014-2015) determinations including LHC run I data:

- CTEQ TEA $\Longrightarrow$ CT10, CT14 (Hessian method)
- MMHT $\Longrightarrow$ MSTW, MMHT14 (Hessian method)
- NNPDF $\Longrightarrow$ NNPDF3.0 (MC sampling and neural networks)
- ABM $\Longrightarrow$ ABM12LHC (Hessian method)

Other recent determinations not including LHC data

- HERA2.0 (Hessian method)
- CTEQ-Jlab $\Longrightarrow$ CJ12 fit (Hessian method)
- JR (Hessian method)

In the past years, a lot of efforts have been put in organizing a sistematic library to access all PDFs with an organic C++ interface:
https://lhapdf.hepforge.org/

Extremely important tool for hadron collider phenomenology.

Different methodologies

Methodologies for PDF determination vary among recent PDF analyses:

- smaller/larger/different data sets considered,
- heavy-flavor treatment (GMVFN, FFN,),
- different values/treatment of $\alpha_s(M_Z)$,
- different parametrizations for input PDFs at Q_0,
- ....

$\Rightarrow$ differences in central predictions and error estimate

A couple of examples with pre LHC PDFs:

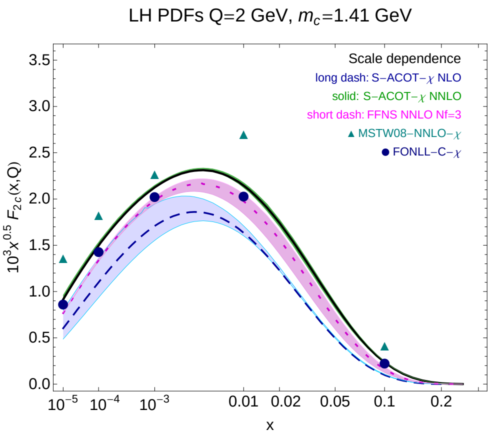
Results for $F_2^c(x, Q^2)$ in DIS at NLO/NNLO

At NNLO and $Q \approx m_c$:

■ S-ACOT-$\chi \approx$ FFN($N_f = 3$) without tuning

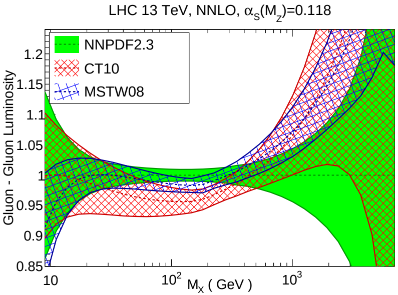
■ It is close to other NNLO schemes

■ S-ACOT-χ predictions are for a physically motivated rescaling variable $\zeta = x(1 + 4m_c^2/Q^2)$. Dependence on the form of ζ is also reduced

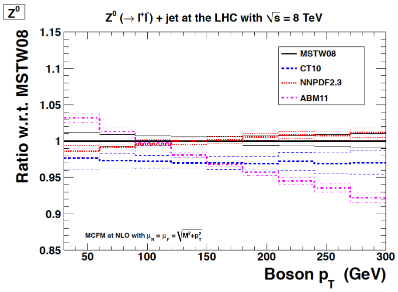


From: M.G., Nadolsky, Lai, Yuan, PRD (2012)

Other examples of pre LHC determinations



Gluon luminosity at the LHC 13 TeV.
From PDF4LHC 1507.00556, JPG (2015)



Z-boson production: p_T spectrum LHC 8 TeV.
From Malik and Watt JHEP (2014)

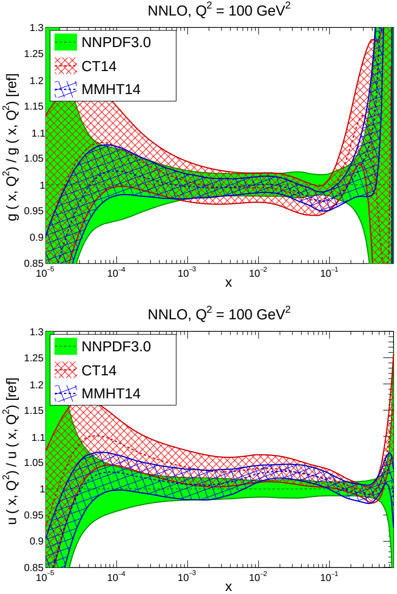
PDF4LHC: Comparisons and Benchmarking

The PDF4LHC Working Group:

- performing thorough benchmark studies of PDFs and of predictions at the LHC
- making recommendations for a standard method of estimating **PDFs** + $\alpha_s(M_Z^2)$ uncertainties at the LHC through a combination of the results from different individual groups.
- Forthcoming PDF4LHC LHC run II PDF recommendations arXiv:xxxx.xxxxx with more recent extensive comparisons between CT14, NNPDF3.0, MMHT14, ABM12LHC, HERA2.0 ...

http://www.hep.ucl.ac.uk/pdf4lhc/

Recent efforts in comparisons/benchmarking

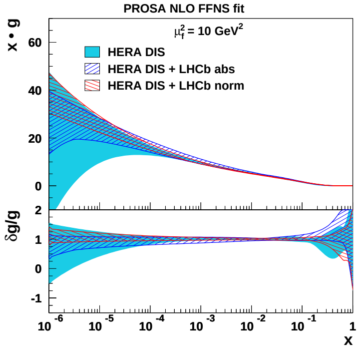


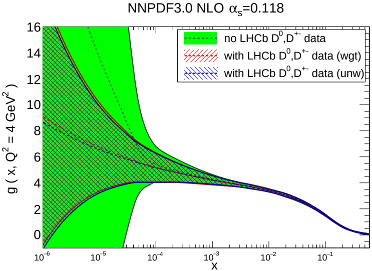
Comparison of PDFs at $Q2 = 10^2\ \text{GeV}^2$ between the NNPDF3.0, CT14 and MMHT14 sets at NNLO, with $\alpha_s(M_Z^2) = 0.118$.

From PDF4LHC 1507.00556 (July 2015)

SOME RECENT ANALYSES

Constraints on the gluon at low x from LHCb





"Impact of heavy-flavour production cross sections measured by the LHCb experiment on parton distribution functions at low x"

Zenaiev *et al.*, PROSA Collaboration EPJC 2015

(More in Achim Geiser's talk)

"Charm production in the forward region: constraints on the small-x gluon and backgrounds for neutrino astronomy"

Gauld, Rojo, Rottoli, Talbert, 1506.08025 (2015)

Procedures for the combination of the PDFs into the future PDF4LHC ensemble

Possible methods for the construction of the combined PDF4LHC ensemble:

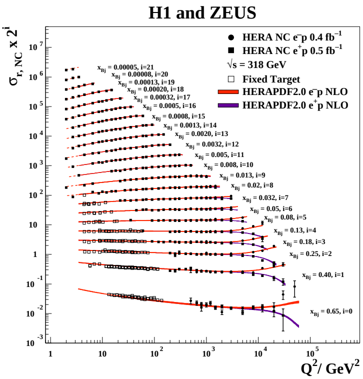
- Meta-parametrizations + MC replicas + Hessian data set diagonalization
  J. Gao, P. Nadolsky JHEP (2014)
  http://metapdf.hepforge.org (Gao, Huston, Nadolsky)

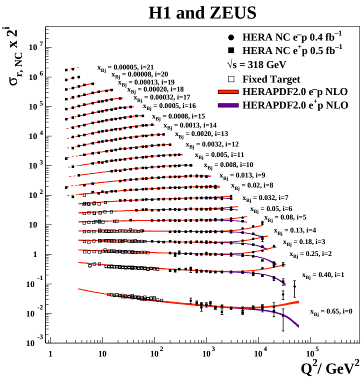
- Unweighting/compression of Monte-Carlo replicas
  G. Watt, R. Thorne,1205.4024; R. Ball et al., 1108.1758; S. Forte, G. Watt, 1301.6754

- A compression algorithm for the combination of PDF sets
  Carrazza, Latorre, Rojo, Watt, 1504.06469 (2015)
  (more in Andy Buckley's talk)

PDF's backbone: new high precision HERA data

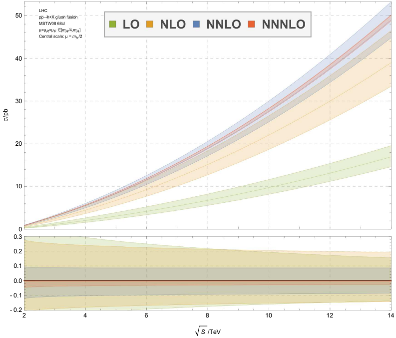
Measurements of lepton-proton deep-inelastic-scattering (DIS) reaction data from H1 and ZEUS coll. at HERA are the most important data sets in PDFs determination.





Recent: New combined data and HERA2.0 PDFs released arxiv:1506.06042 (More in Iris Abt's talk)

The need for precision



Inclusive Higgs production Xsec for gluon-gluon fusion at N^{3}LO, Anastasiou, Duhr, Dulat, Herzog, Mistlberger, PRL (2015).

- Significantly reduced the scale dependence of the Higgs cross section
- PDF and α_s uncertainties become the dominant remaining theoretical uncertainty.

The CT14 global QCD analysis

New results from the CTEQ-TEA group:

S. Dulat, T.J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky,
J. Pumplin, C. Schmidt, D. Stump, C.-P. Yuan

arXiv:1506.07443

What's new in CT14 NNLO PDFs

CT14 differs from CT10 PDFs in several respects:

new HERA data:

- Combined HERA charm production measurements ($F_2^{(c)}$)
- measurements of the longitudinal $F_L(x, Q^2)$ in DIS neutral currents

new Tevatron data:

- Tevatron Run 1 CDF and D0 inclusive jet data are dropped,
- old D0 data (0.75 fb^{-1}) superseded by the new D0 (9.7 fb^{-1}) W-electron rapidity asymmetry data.

LHC 7 TeV run I data included

- ATLAS and LHCb W and Z production,
- ATLAS, CMS and LHCb W-lepton charge asymmetry,
- ATLAS and CMS inclusive jet data.

CT14 has 2995 data points

CT14 Data sets ensemble I

ID#	Experimental data set	N_{pt}	χ^2_e	χ^2_e/N_{pt}	S_n
101	BCDMS F_2^p	337	384	1.14	1.74
102	BCDMS F_2^d	250	294	1.18	1.89
104	NMC F_2^d/F_2^p	123	133	1.08	0.68
106	NMC σ_{red}^p	201	372	1.85	6.89
108	CDHSW F_2^p	85	72	0.85	-0.99
109	CDHSW F_3^p	96	80	0.83	-1.18
110	CCFR F_2^p	69	70	1.02	0.15
111	CCFR xF_3^p	86	31	0.36	-5.73
124	NuTeV $\nu\mu\mu$ SIDIS	38	24	0.62	-1.83
125	NuTeV $\bar{\nu}\mu\mu$ SIDIS	33	39	1.18	0.78
126	CCFR $\nu\mu\mu$ SIDIS	40	29	0.72	-1.32
127	CCFR $\bar{\nu}\mu\mu$ SIDIS	38	20	0.53	-2.46
145	H1 σ_r^b	10	6.8	0.68	-0.67
147	Combined HERA charm production	47	59	1.26	1.22
159	HERA1 Combined NC and CC DIS	579	591	1.02	0.37
169	H1 F_L	9	17	1.92	1.7

CT14 Data sets ensemble II

ID#	**Experimental data set**	N_{pt}	χ^2_e	χ^2_e/N_{pt}	S_n
201	E605 Drell-Yan process	119	116	0.98	-0.15
203	E866 Drell-Yan process, $\sigma_{pd}/(2\sigma_{pp})$	15	13	0.87	-0.25
204	E866 Drell-Yan process, $Q^3 d^2\sigma_{pp}/(dQdx_F)$	184	252	1.37	3.19
225	CDF Run-1 electron A_{ch}, $p_{T\ell} > 25$ GeV	11	8.9	0.81	-0.32
227	CDF Run-2 electron A_{ch}, $p_{T\ell} > 25$ GeV	11	14	1.24	0.67
234	DØ Run-2 muon A_{ch}, $p_{T\ell} > 20$ GeV	9	8.3	0.92	-0.02
240	LHCb 7 TeV 35 pb^{-1} W/Z $d\sigma/dy_\ell$	14	9.9	0.71	-0.73
241	LHCb 7 TeV 35 pb^{-1} A_{ch}, $p_{T\ell} > 20$ GeV	5	5.3	1.06	0.30
260	DØ Run-2 Z rapidity	28	17	0.59	-1.71
261	CDF Run-2 Z rapidity	29	48	1.64	2.13
266	CMS 7 TeV 4.7 fb^{-1}, muon A_{ch}, $p_{T\ell} > 35$ GeV	11	12.1	1.10	0.37
267	CMS 7 TeV 840 pb^{-1}, elec. A_{ch}, $p_{T\ell} > 35$ GeV	11	10.1	0.92	-0.06
268	ATLAS 7 TeV 35 pb^{-1} W/Z cross sec., A_{ch}	41	51	1.25	1.11
281	DØ Run-2 9.7 fb^{-1} elec. A_{ch}, $p_{T\ell} > 25$ GeV	13	35	2.67	3.11
504	CDF Run-2 inclusive jet production	72	105	1.45	2.45
514	DØ Run-2 inclusive jet production	110	120	1.09	0.67
535	ATLAS 7 TeV 35 pb^{-1} incl. jet production	90	50	0.55	-3.59
538	CMS 7 TeV 5 fb^{-1} incl. jet production	133	177	1.33	2.51

Aspects of the CT14 analysis

- PDFs are parametrized at init scale $Q_0 = 1.3$ GeV.
- large-x data not included to avoid large non-perturbative contributions ($W > 3.5$ GeV)
- more flexible parametrizations for gluon, d/u at large x, both d/u and $\bar{d}/\bar{u}$ at small x, and strange ($\bar{s} = s$) PDFs.
- Non-perturbative parametrization employing Bernstein polynomials $P_a(x)$: $xf_a(x) = x^{a_1}(1-x)^{a_2}P_a(x)$

  This reduces the correlation among its coefficients.
- CT14: 28 shape parameters, while CT10 has 25.
- S-ACOT-χ NNLO for the heavy flavor treatment
- NNLO calculations for DIS, DY, W, Z cross sections, for the jet cross sections and DIS charged currents we only use the NLO calculation but with NNLO PDF.

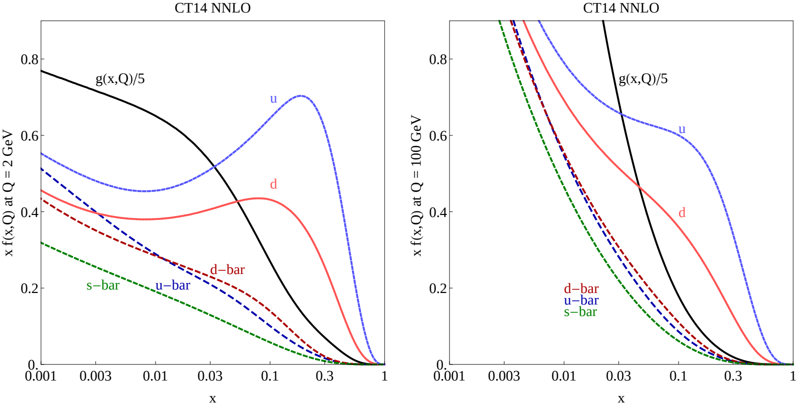
Aspects of the CT14 analysis: $\alpha_s(M_Z)$

- central value of $\alpha_s(M_Z) = 0.118$ has been assumed in the global fits at NLO and NNLO, but
- PDF sets at alternative values of $\alpha_s(M_Z)$ are provided.
- CT14 prefers $\alpha_s(M_Z) = 0.115^{+0.006}_{-0.004}$ at NNLO (0.117 ± 0.005 at NLO) at 90 % confidence level (C.L.).

Uncertainties from the global QCD fits are larger than those of the data from LEP and other experiments included into the world average *Chin.Phys.C* (2014).

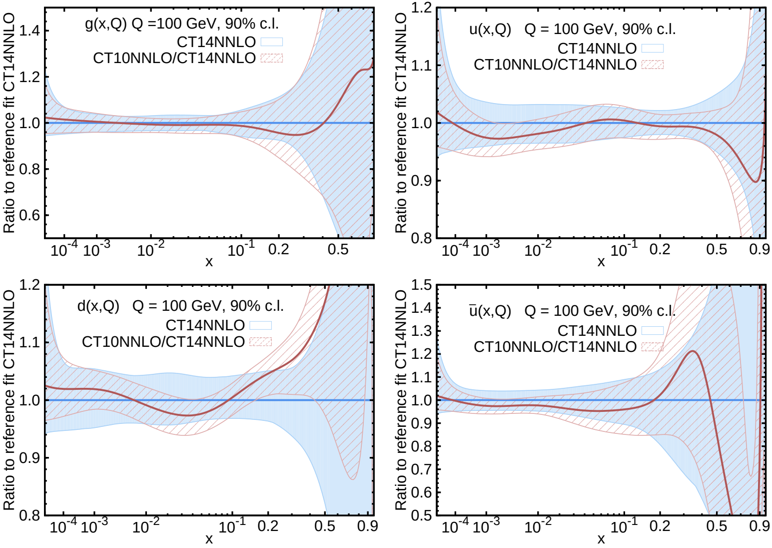
CT14 $\alpha_s(M_Z)$ central is consistent with the world average value.

CT14: DGLAP evolution



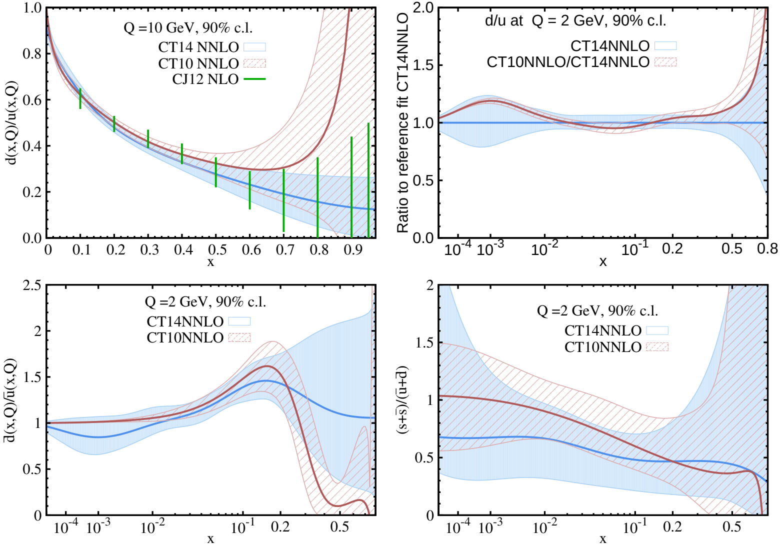
The CT14 PDFs $u, \overline{u}, d, \overline{d}, s = \overline{s}$, and g, evolved up to $Q = 2$ GeV and $Q = 100$ GeV.

CT14 vs CT10 at NNLO 90% C.L.



- ATLAS, CMS 7 TeV W/Z prod. $\Rightarrow$ d-quark increased by 5% at $x \approx 0.05$.
- D0 ele charge asy data $\Rightarrow$ d highly reduced at $x \geq 0.1$ and u moderately increased.

CT14 $d(x,Q)/u(x,Q)$ ratios

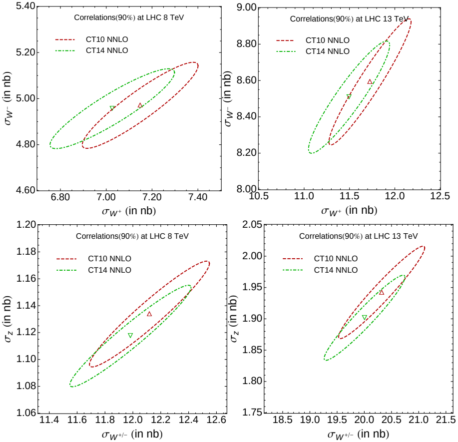


- 9.7 fb^{-1} D0 charge asy $\Rightarrow$ reduction of the central ratio at $x > 0.1$,
- new parametrization form $\Rightarrow$ increased uncertainty at $x < 0.05$
- s reduction at $x > 0.01 \Rightarrow$ smaller ratio $(s+\bar{s})/(\bar{u}+\bar{d})$. The $SU(3)$-symmetric asymptotic solution at $x \to 0$ is still allowed in CT14: bigger unc. $x \approx 10^{-5}$.

CT14 NNLO: agreement with data

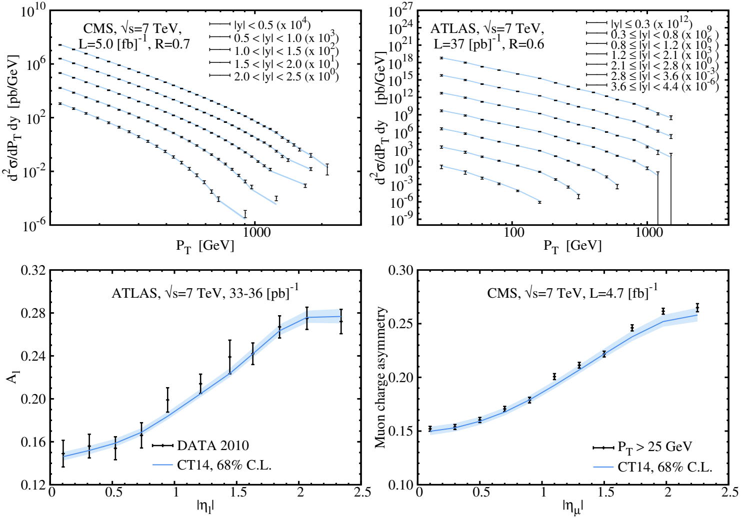
Total of 2947 data points from 33 experiments
$\chi^2 = 3252$ at the best fit CT14 NNLO,
$\chi^2/N_{pt} = 1.10$.

Data and theory are in reasonable good agreement for most experiments (next slides)



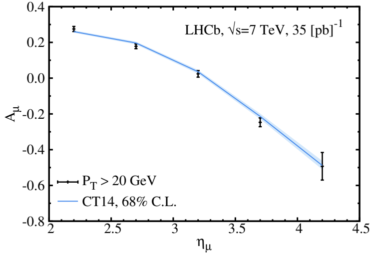
W/Z Correlations plots CT14 vs CT10 @ NNLO

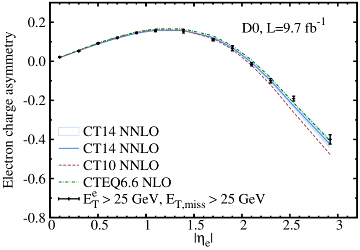
CT14 NNLO: agreement with data

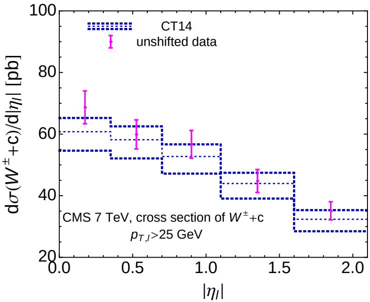


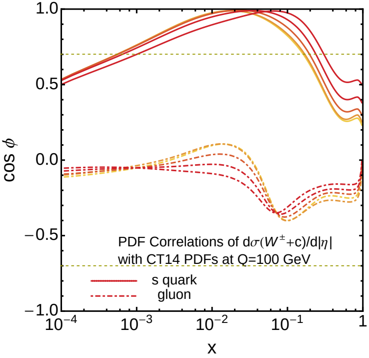
Inclusive jet production and *W* lepton charge asymmetry at the LHC 7 TeV

CT14 NNLO: agreement with data







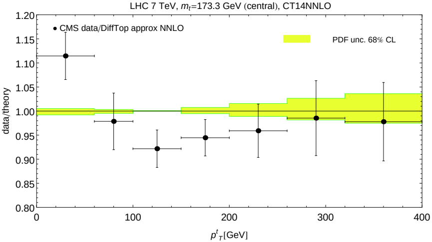


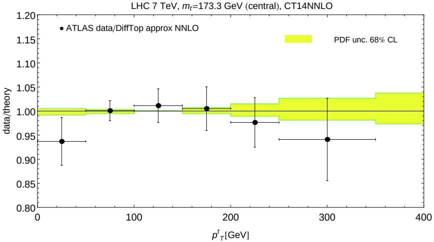
$W + c$ **is not included in the fit (theory available only at NLO).**

CT14 NNLO: agreement with data

Total inclusive $t\bar{t}$ cross section at NNLO in QCD with Top++, (Czakon, Mitov, CPC 2014)

$pp \to t\bar{t}$ (pb), PDF unc., $\alpha_s = 0.118$	7 TeV	8 TeV	13 TeV
68% C.L. (Hessian)	$177 + 4.8\% - 3.9\%$	$250 + 3.9\% - 3.5\%$	$820 + 2.6\% - 2.7\%$
68% C.L. (LM)		$+4.8\% - 4.6\%$	$+2.9\% - 2.9\%$
$pp \to t\bar{t}$ (pb), PDF+α_s	7 TeV	8 TeV	13 TeV
68% C.L. (Hessian)	$+5.5\% - 4.6\%$	$+5.2\% - 4.4\%$	$+3.6\% - 3.5\%$
68% C.L. (LM)		$+5.1\% - 4.7\%$	$+3.6\% - 3.5\%$





Approx NNLO p_T spectrum for the final state top-quark with DiffTop (M.G., Lipka, Moch, JHEP 2015)

Post CT14 analysis: ongoing/future work

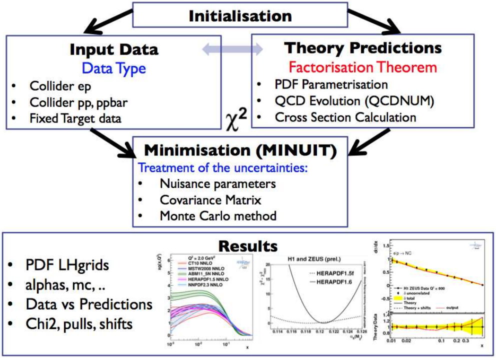
- impact of the new HERA II recent combined DIS measurements
- impact of $t\bar{t}$ inclusive and differential cross section on
- ....

Conclusions

- LHC unprecedented energies brought us in a new precision era
- A lot of efforts are ongoing to pin down PDFs uncertainties which still remain among the major sources of systematical theory uncertainties
- Things will be very interesting when many missing NNLO Xsecs will be consistently included in the next PDF iteration.
- Several future LHC programs for discovery of new physics interactions strongly depend on our knowledge of proton structure.

THANK YOU!

Backup



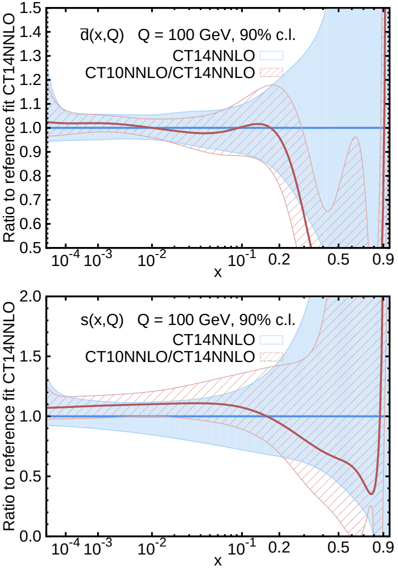
In global PDF fits a large number of iterations of the theory calculation programs (NLO, NNLO) is required to evaluate cross sections:

some of these computations are CPU time consuming!

Advanced tools have been developed to have theory calculations on grids: extremely fast!

$\Rightarrow$ **FastNLO** and **APPLgrid**

CT14 vs CT10 at NNLO 90% C.L.



Definitions of the covariance matrix

arXiv:1302.6246, appendix in R. Ball et al., arXiv:1211.5142

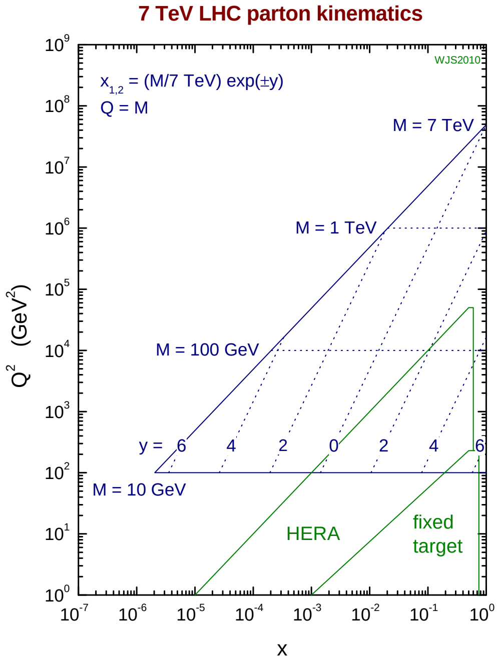
$$\chi^2 = \sum_{\{\text{exp.}\}} \left[\sum_{k=1}^{N_{pts}} \frac{1}{s_k^2} \left(D_k - T_k(\{a\}) - \sum_{\alpha=1}^{N_\lambda} \lambda_\alpha \beta_{k\alpha} \right)^2 + \sum_{\alpha=1}^{K_e} \lambda_\alpha^2 \right]$$

The experimental correlated systematic errors $\beta_{k\alpha}$ are often published as percentages. It can be taken to be a percentage of the theoretical prediction T_k ("truth") or the experimental datum D_k.

1. **Experimental (D) prescription:** normalize all $\beta_{k\alpha}$ to D_k
2. T **(T_0) prescription:** normalize luminosity & other multiplicative errors to (fixed) T_k, additive errors to D_k
3. **Extended** T **(T_0) prescription:** normalize all errors to (fixed) T_k

The methods are numerically equivalent if T_k is close to D_k. Additive (multiplicative) errors are to be normalized to T_k (D_k) to avoid/reduce biases. The available experimental data usually do not specify if the errors are additive or multiplicative.

Kinematics: LHC 7 TeV



Kinematics: LHC 14 TeV