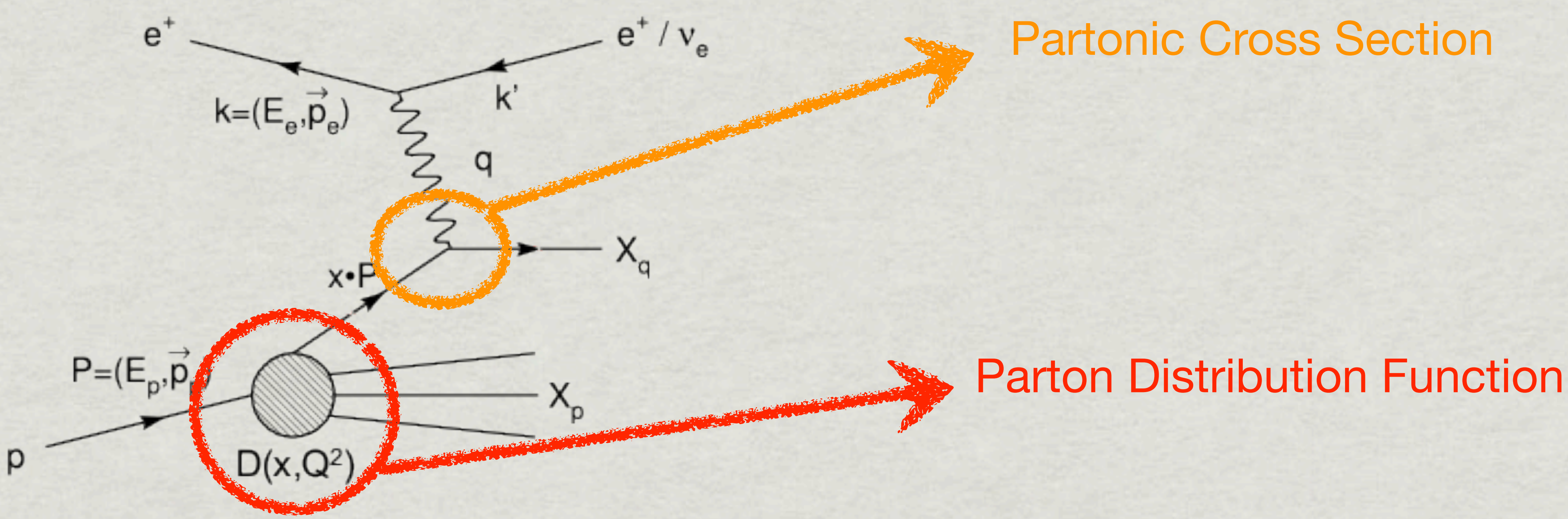
(NN)PDFs, the LHC & the LHeC

Alberto Guffanti
Niels Bohr International Academy & Discovery Center
Niels Bohr Intitute - Copenhagen

Parton Distribution Functions
What are they?

Consider a process with a single hadron in the initial state



The cross section for such a process can be written (**Factorization Theorem**) as

$$d\sigma = \sum_a \int_0^1 \frac{d\xi}{\xi} D_a(x,\mu^2)\, d\hat{\sigma}_a\left(\frac{x}{\xi}, \frac{\hat{s}}{\mu^2}, \alpha_s(\mu^2)\right) + O\left(\frac{1}{Q^p}\right)$$

Parton Distribution Functions
What are they?

- Parton Distribution Functions are non-perturbative objects and their value at given x and Q^2 cannot be computed in QCD Perturbation Theory (Lattice?)
- ... but the **scale dependence** of PDFs is governed by the **DGLAP** evolution equations

$$\frac{\partial q_i(x,\mu^2)}{\partial \ln \mu^2} = \frac{\alpha_s(\mu^2)}{2\pi}\left[P_{qq}(x)\otimes q_i(x,\mu^2)\right] + \frac{\alpha_s(\mu^2)}{2\pi}\left[P_{qg}\otimes g(x,\mu^2)\right]$$

$$\frac{\partial g(x,\mu^2)}{\partial \ln \mu^2} = \frac{\alpha_s(\mu^2)}{2\pi}\left[P_{gq}(x)\otimes \sum_i \left(q_i(x,\mu^2)+\bar{q}_i(x,\mu^2)\right)\right] + \frac{\alpha_s(\mu^2)}{2\pi}\left[P_{gg}\otimes g(x,\mu^2)\right]$$

- ... where the **splitting functions** (P_{ij}) can be computed in Perturbation Theory and are known up to **NNLO**

[**LO** - Dokshitzer; Gribov, Lipatov; Altarelli, Parisi (1977)]
[**NLO** - Floratos, Ross, Sachrajda; Gonzalez-Arroyo, Lopez, Ynduirain; Curci, Furmanski, Petronzio (1981)]
[**NNLO** - Moch, Vermaseren, Vogt (2004)]

The PDF fitting game

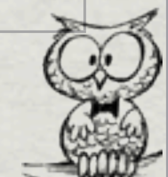
The players

Collaboration	Authors	arXiv
ABM	S. Alekhin, J. Blümlein, S. Moch	1105.5349, 1101.5261, 1107.3657, 0908.3128, 0908.2766, ...
CTEQ/TEA	M. Guzzi J. Huston, H.-L. Lai, P. Nadolsky, J. Pumplin, D. Stump, C.-P. Yuan	1108.5112, 1101.0561, 1007.2241, 1004.4624, 0910.4183, 0904.2424, 0802.0007, ...
GJR	M. Glück, P. Jimenez-Delgado, E. Reya	1003.3168, 0909.1711, 0810.4274, ...
HERAPDF	H1 and ZEUS Collaborations	1107.4193, 1006.4471, 0906.1108, ...
MSTW	A. Martin, J. Stirling, R. Thorne, G. Watt	1107.2624, 1006.2753, 0905.3531, 0901.0002, ...
NNPDF	R. D. Ball, V. Bertone, S. Carrazza, F. Cerutti, C. S. Deans, L. Del Debbio, S. Forte, AG, N. P. Hartland, J. I. Latorre, J. Rojo, M. Ubiali	1110.2483, 1108.2758, 1107.2652, 1103.2369, 1102.3182, 1101.1300, 1005.0397, 1002.4407, 0912.2276, 0906.1958, ...

The PDF fitting game

Status of PDF fits

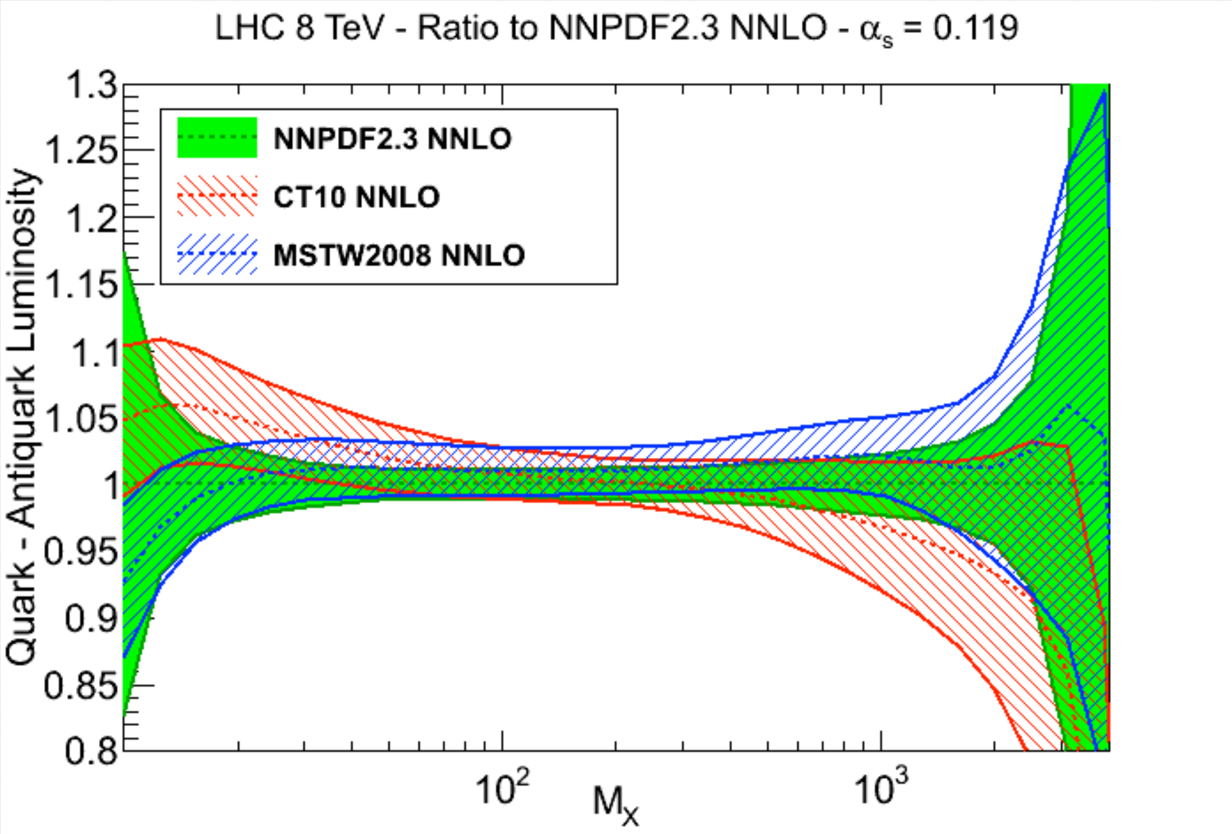
	DATASET	PERT. ORDER	HQ TREATMENT	α_s	PARAM.	UNCERT.
ABM11	DIS Drell-Yan	NLO NNLO	FFN (BMSN)	Fit (multiple values available)	6 indep. PDFs Polynomial (25 param.)	Hessian ($\Delta\chi^2$=1)
CT10	**Global**	LO NLO NNLO	GM-VFNS (S-ACOT)	External (multiple values available)	6 indep. PDFs Polynomial (26 param.)	Hessian ($\Delta\chi^2$=100)
JR09	DIS Drell-Yan Jets	NLO NNLO	FFN VFN	Fit	5 indep. PDFs Polynomial (15 param.)	Hessian ($\Delta\chi^2$=1)
HERAPDF1.5	DIS (HERA)	NLO NNLO	GM-VFNS (TR)	External (multiple values available)	5 indep. PDFs Polynomial (14 param.)	Hessian ($\Delta\chi^2$=1)
MSTW08	**Global**	LO NLO NNLO	GM-VFNS (TR)	Fit (multiple values available)	7 indep. PDFs Polynomial (20 param.)	Hessian ($\Delta\chi^2 \simeq 25$)
NNPDF2.3	**Global**	LO NLO NNLO	GM-VFNS (FONLL)	External (multiple values available)	7 indep. PDFs Neural Nets (259 param.)	Monte Carlo

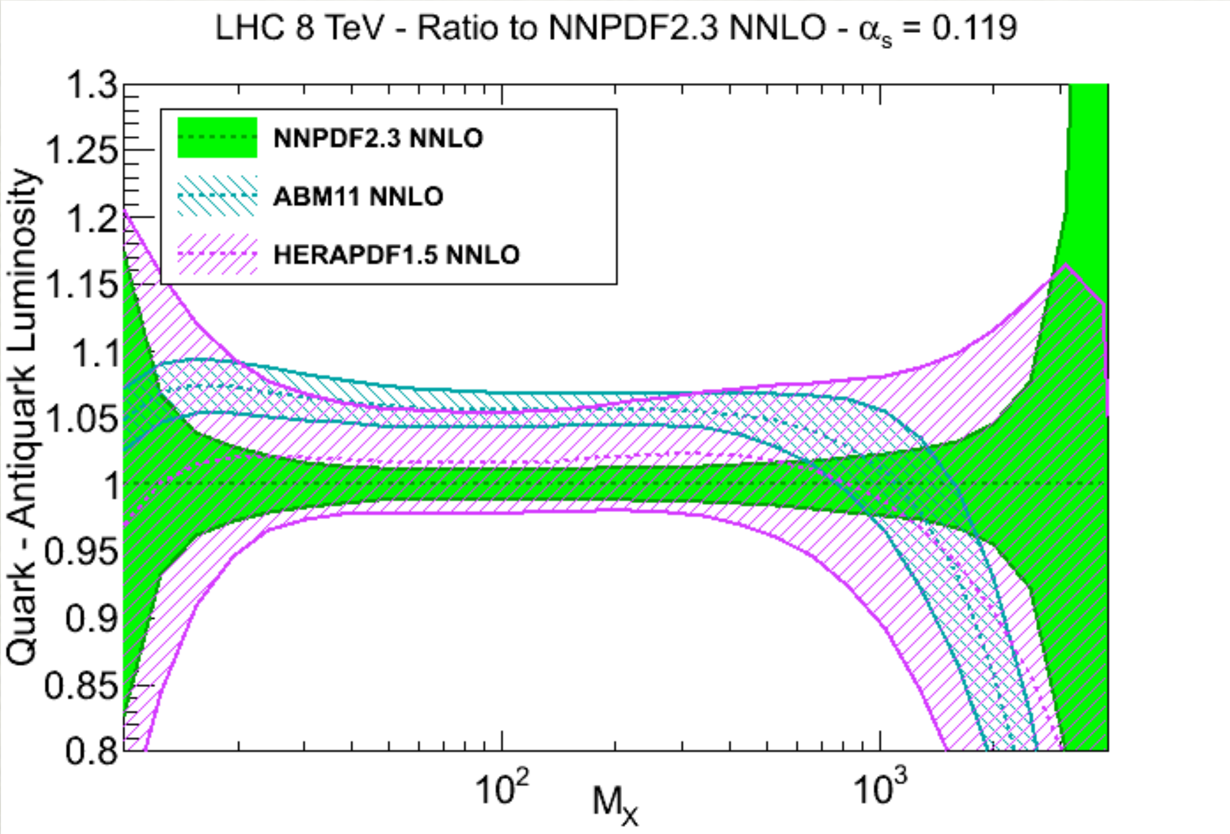


The PDF fitting game

Status of PDF fits - parton luminosities

$$\Phi_{ij}\left(M_X^2\right) = \frac{1}{s}\int_\tau^1 \frac{dx_1}{x_1} f_i\left(x_1, M_X^2\right) f_j\left(\tau/x_1, M_X^2\right)$$



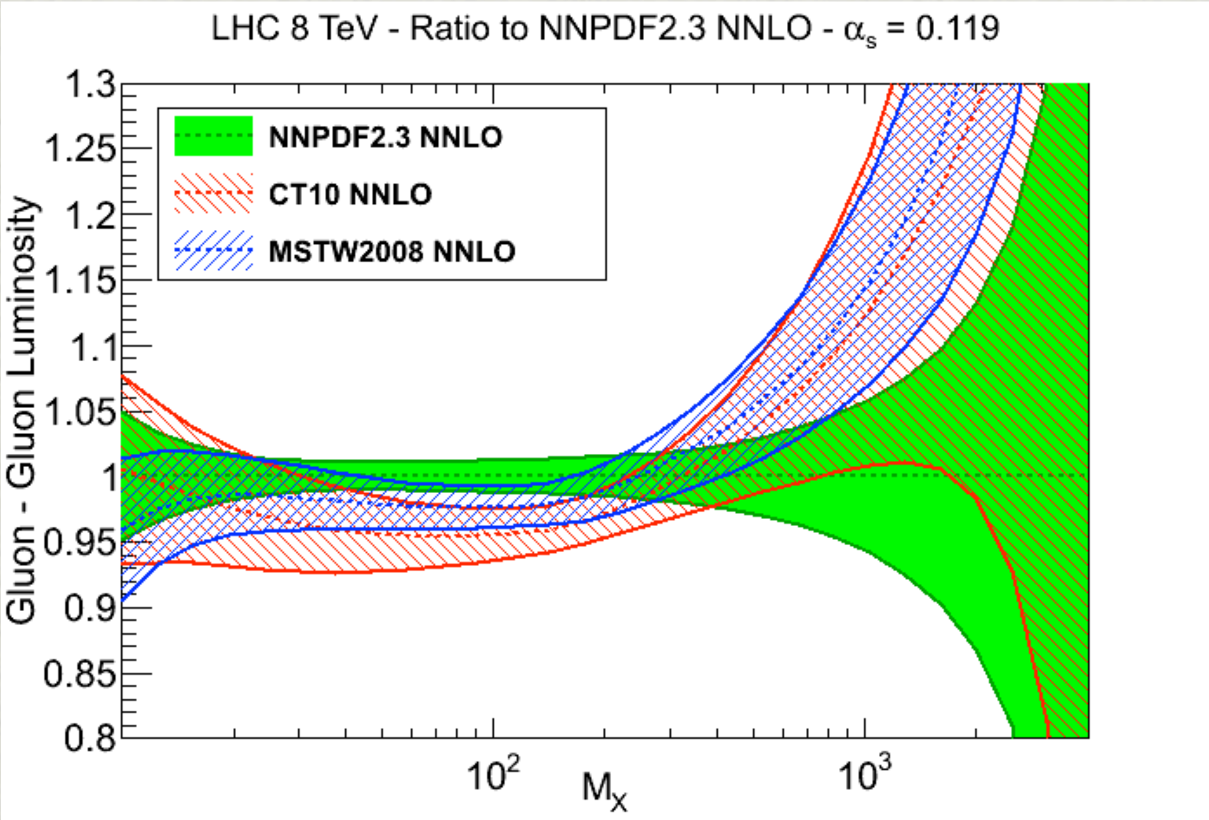


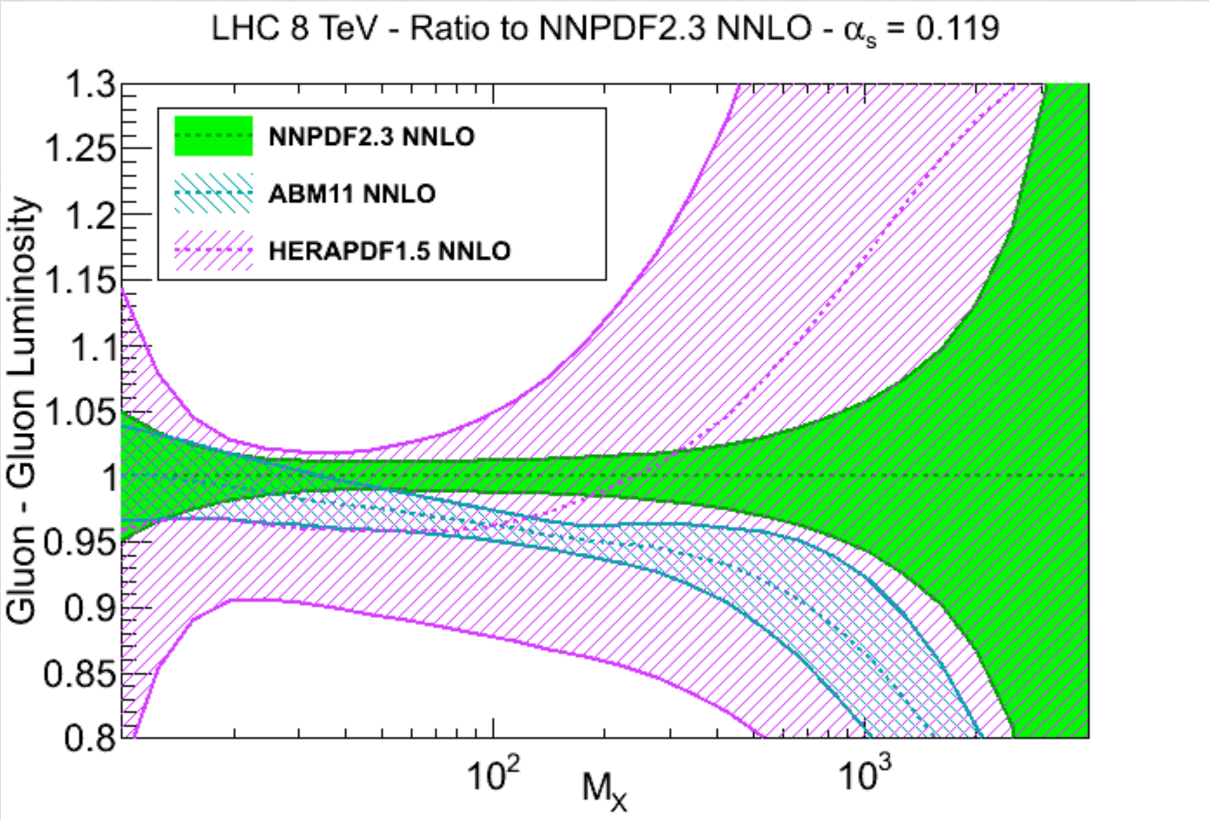
- **Good compatibility** among global fits
- Differences are more marked when comparing to restricted dataset fits

The PDF fitting game

Status of PDF fits - parton luminosities

$$\Phi_{ij}\left(M_X^2\right) = \frac{1}{s}\int_{\tau}^{1}\frac{dx_1}{x_1} f_i\left(x_1, M_X^2\right) f_j\left(\tau/x_1, M_X^2\right)$$

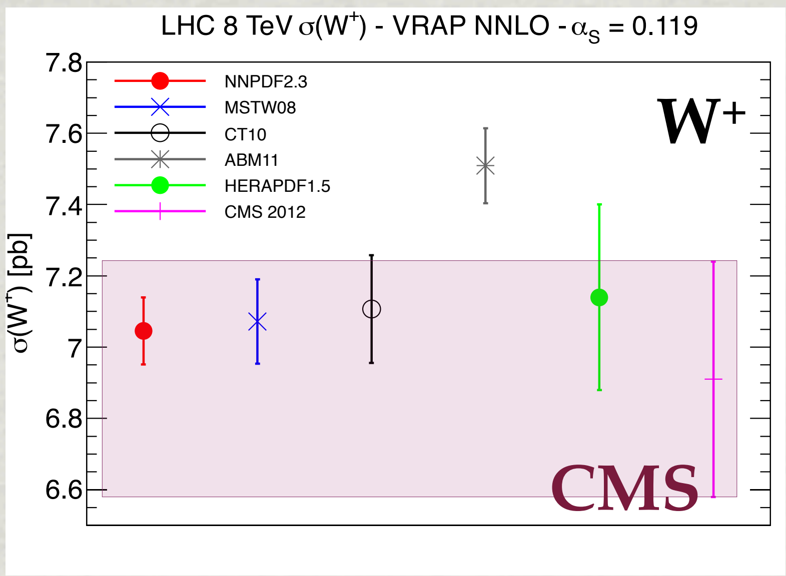


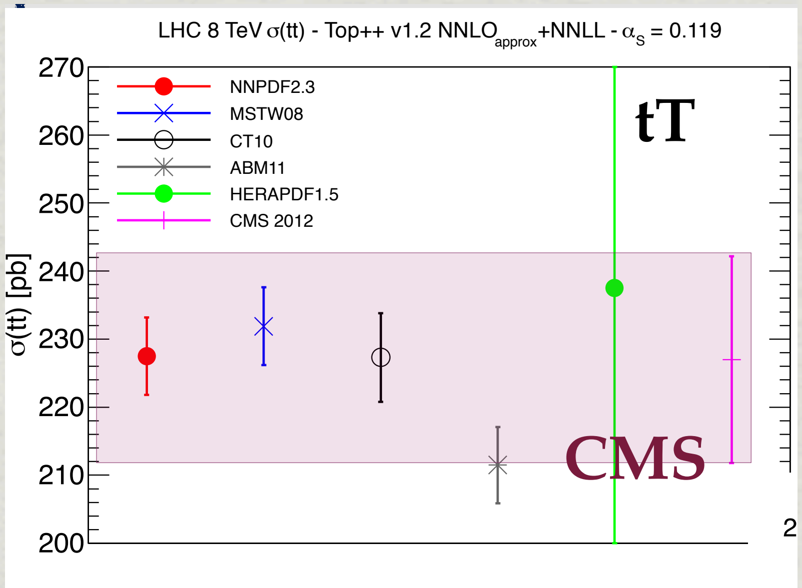


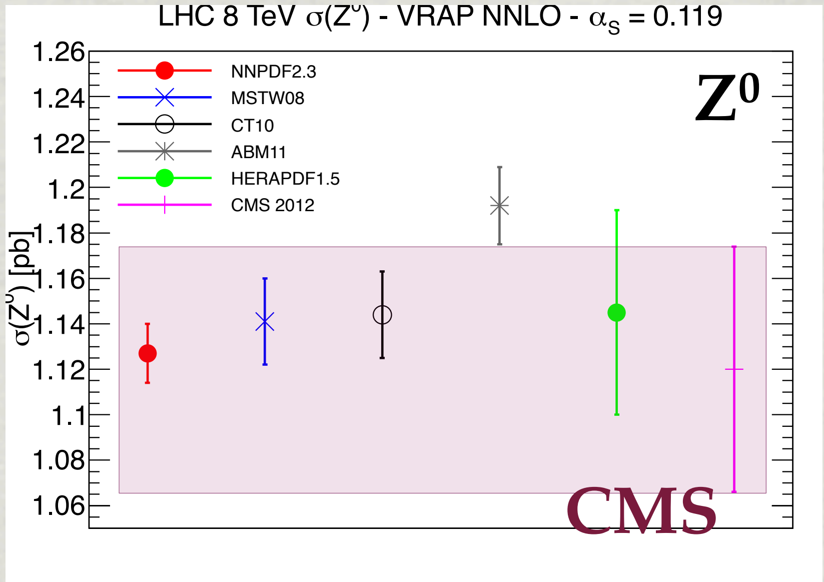
- **Good compatibility** among global fits
- Differences are more marked when comparing to restricted dataset fits

The PDF fitting game

Status of PDF fits - LHC cross-section



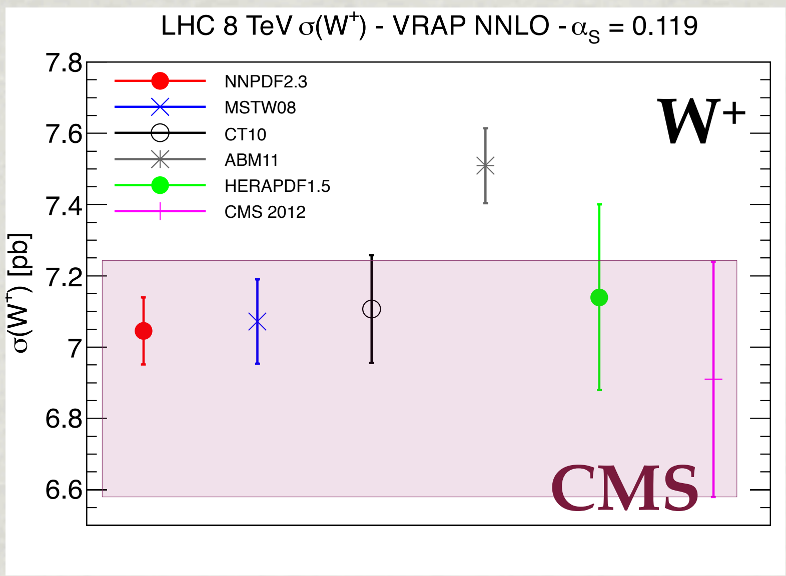


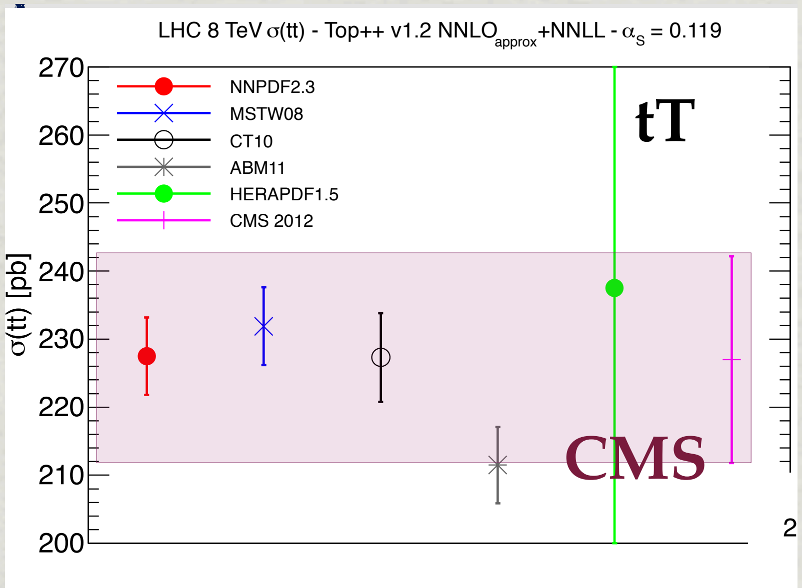


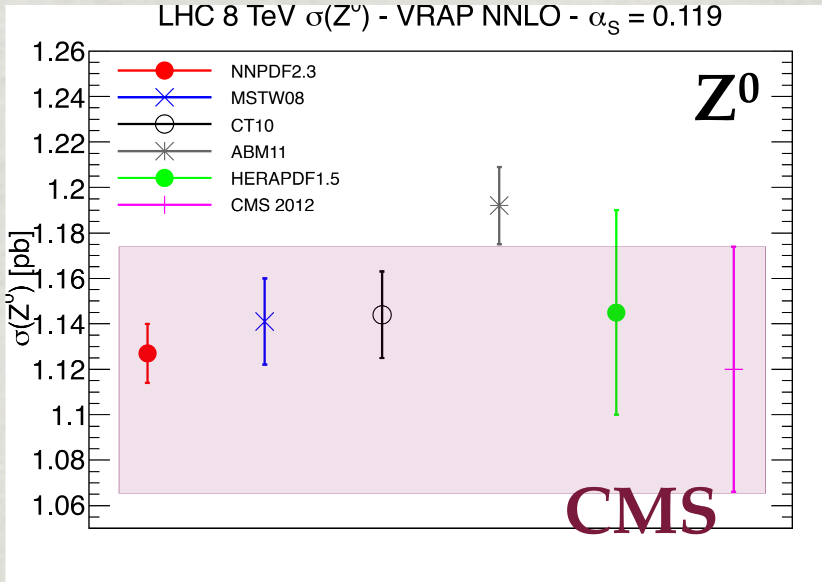
LHC data are already starting to **discriminate among predictions**

The PDF fitting game

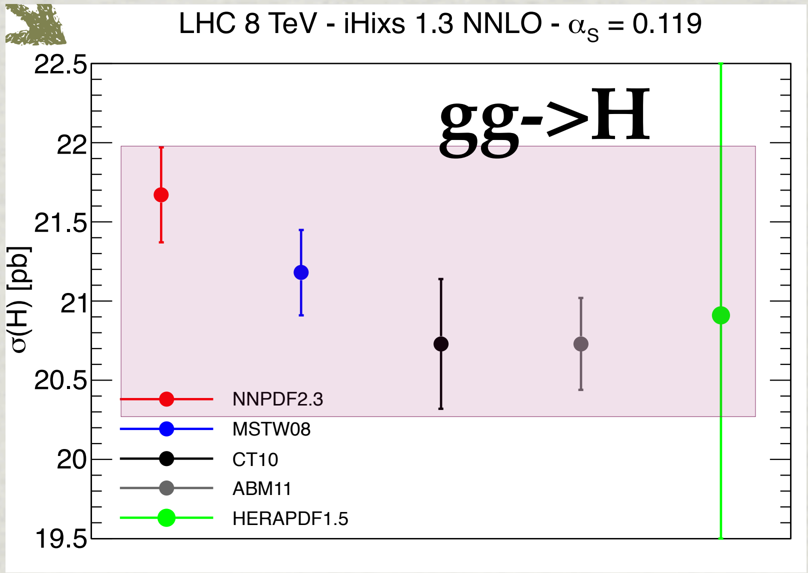
Status of PDF fits - LHC cross-section







LHC data are already starting to **discriminate among predictions**



NNPDF Methodology
Main ingredients

- **Monte Carlo** determination of **uncertainties**
  - **No** need to rely on **linear propagation** of errors
  - Possibility to test the impact of **non-gaussianly** distributed uncertainties
  - Possibility to test for **non-gaussian behaviour** of uncertainties of fitted PDFs
- Parametrization of PDFs using **Neural Networks**
  - Provide an **unbiased parametrization**
- Determine the **best fit** PDFs using **Cross-Validation**
  - Ensures **proper fitting**, avoiding overlearning

NNPDF Methodology
... in a Nutshell

- **Generate** N_{rep} **Monte Carlo replicas** of the experimental data, taking into account all experimental correlations

- **Fit** a set of Parton Distribution Functions, parametrized at the initial scale using Neural Networks, to each replica

- **Expectation values** for observables are then given by

$$\langle \mathcal{F}[f_i(x,Q^2)] \rangle = \frac{1}{N_{rep}} \sum_{k=1}^{N_{rep}} \mathcal{F}\left(f_i^{(net)(k)}(x,Q^2) \right)$$

.... and corresponding formulae are used to compute uncertainties, correlations, etc.

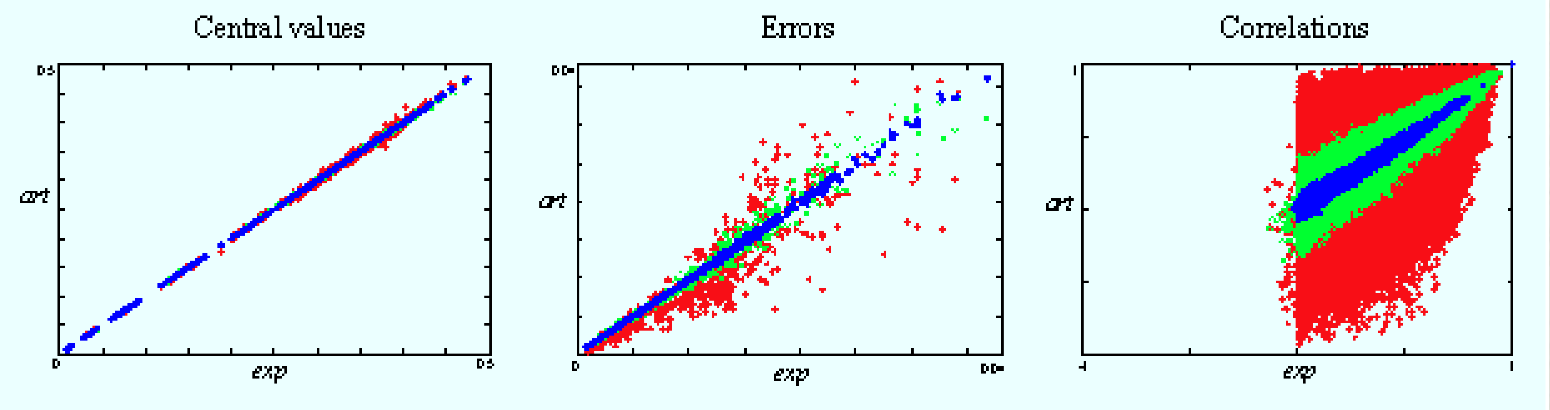
NNPDF Methodology
Monte Carlo replicas generation

- **Monte Carlo replicas** are generated according to the distribution

$$O_i^{(art)\,(k)} = \left(1 + r_N^{(k)}\,\sigma_N\right)\left[O_i^{(exp)} + \sum_{p=1}^{N_{sys}} r_p^{(k)}\,\sigma_{i,p} + r_{i,s}^{(k)}\,\sigma_s^i\right]$$

where r_i are (gaussianly distributed) random numbers

- **Validate** Monte Carlo replicas against experimental data



- O(1000) replicas needed to **reproduce correlations** in experimental data to percent accuracy

Reweighting (NN)PDFs

The reweighting idea

[R. D. Ball et al, arXiv:1012.0836]
[R. D. Ball et al, arXiv:1108.1758]

- The N replicas of an NNPDF fit give the probability density in the space of PDFs

- Expectation values for observables are computed as

$$\langle \mathcal{F}[f_i(x, Q^2)]\rangle = \frac{1}{N_{rep}} \sum_{k=1}^{N_{rep}} \mathcal{F}\left(f_i^{(net)(k)}(x, Q^2)\right)$$

- We can then assess the impact of including new data in the fit updating the probability density without performing a complete refit

Reweighting (NN)PDFs
The reweighting formula

[R. D. Ball et al, arXiv:1012.0836]
[R. D. Ball et al, arXiv:1108.1758]

- We can apply **Bayes Theorem** to determine the conditional probability of the PDF upon inclusion of the new data

$$\mathcal{P}_{\rm new}(\{f\}) = \mathcal{N}_\chi \mathcal{P}(\chi^2|\{f\})\mathcal{P}_{\rm init}(\{f\})\,, \quad \mathcal{P}(\chi^2|\{f\}) = [\chi^2(y,\{f\})]^{\frac{n_{dat}-1}{2}}\, e^{-\frac{\chi^2(y,\{f\})}{2}}$$

- Averages over the sample are no **weighted sums**

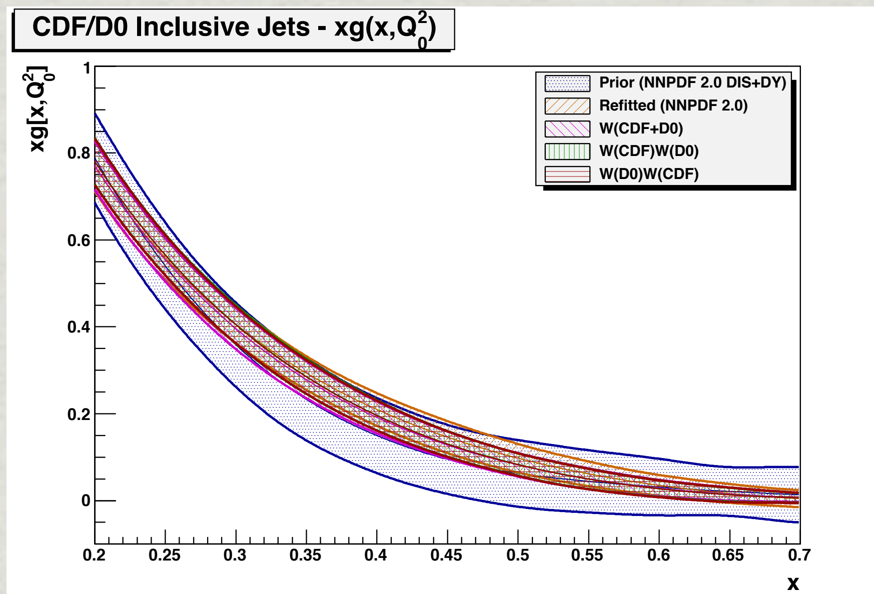
$$\langle \mathcal{F}[f_i(x,Q^2)]\rangle = \sum_{k=1}^{N_{rep}} w_k \mathcal{F}\left(f_i^{(net)(k)}(x,Q^2)\right)$$

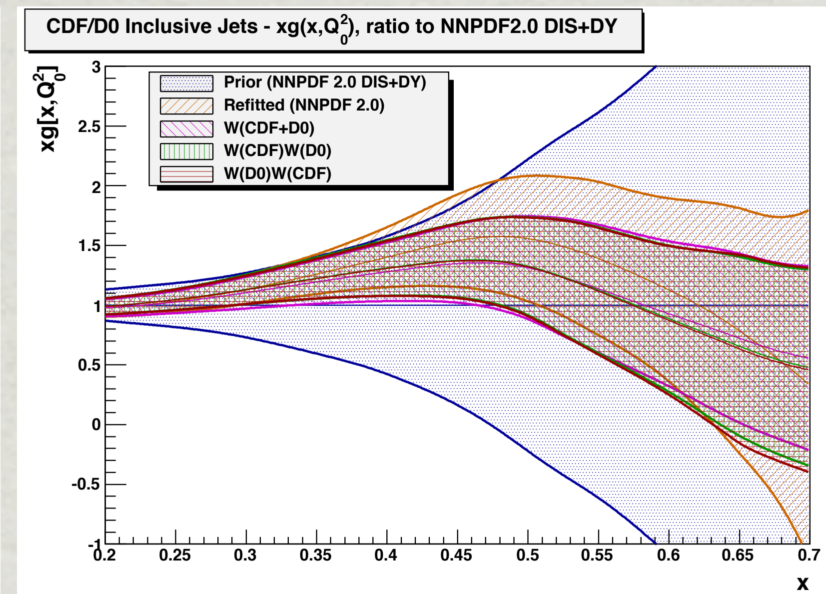
and the **weights** are given by

$$w_k = \frac{[\chi^2(y,f_k)]^{\frac{n_{dat}-1}{2}}\, e^{-\frac{\chi^2(y,f_k)}{2}}}{\sum_{i=1}^{N_{\rm rep}} [\chi^2(y,f_i)]^{\frac{n_{dat}-1}{2}}\, e^{-\frac{\chi^2(y,f_i)}{2}}}$$

Reweighting (NN)PDFs

Validating the reweighting procedure



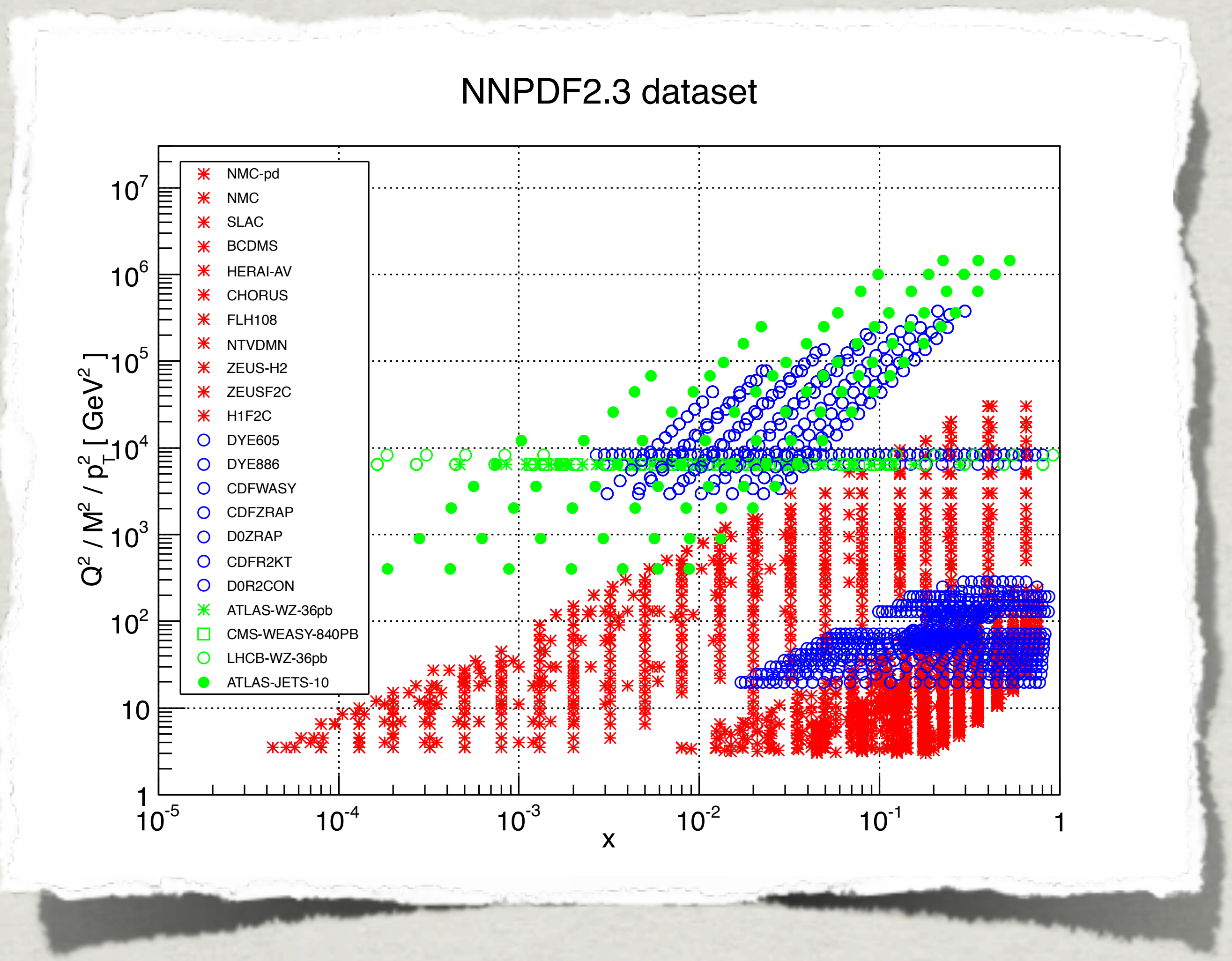


- We started from a fit to DIS and Drell-Yan data and included Tevatron inclusive jet data (CDF & D0) via refitting and reweighting

- Reweighting with the two dataset at the same time or separately (in either order) yields identical results

- Reweighting and refitting yield statistically equivalent results in the region constrained by the new data

NNPDF2.3
Dataset



Experiment	Data
Fixed Target DIS	1952
HERA DIS	834
Fixed Target DY	318
Tevatron W/Z	70
Tevatron Jets	186
LHC W/Z	56
LHC Jets	90

3506 data points
(in the NNLO global fit)

NNPDF2.3
The LHC data - Fit quality

- Compare the quality of the fit to LHC data before and after inclusion in the global fit
- Including LHC data in the fit improves the quality of their description, w/o deteriorating quality of the fit to other datasets
- Moderate impact of the LHC data, supporting consistency of the global fit framework
- Fit quality is comparable at NLO and NNLO, thought the former marginally better

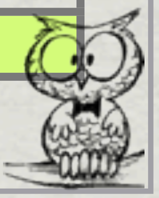
NLO	NNPDF2.3 noLHC	NNPDF2.3
NMCpd	0.93	0.93
NMC	1.59	1.61
SLAC	1.28	1.26
BCDMS	1.20	1.19
HERA-I	1.01	1.00
CHORUS	1.09	1.10
NuTeV	0.42	0.45
DYE605	0.85	0.88
DYE866	1.24	1.28
CDFWASY	1.45	1.54
CDFZRAP	1.77	1.79
D0ZRAP	0.57	0.57
ATLAS-WZ	1.37	1.27
CMS-WEASY	1.50	1.04
LHCb-WZ	1.24	1.21
CDFR2KT	0.60	0.61
D0R2CON	0.84	0.84
ATLAS-JETS-2010	1.57	1.55

NNPDF2.3

The LHC data - Fit quality

- Compare the quality of the fit to LHC data before and after inclusion in the global fit
- Including LHC data in the fit improves the quality of their description, w/o deteriorating quality of the fit to other datasets
- Moderate impact of the LHC data, supporting consistency of the global fit framework
- Fit quality is comparable at NLO and NNLO, thought the former marginally better

NNLO	NNPDF2.3 noLHC	NNPDF2.3
NMCpd	0.94	0.94
NMC	1.56	1.57
SLAC	1.04	1.02
BCDMS	1.28	1.29
HERA-I	1.03	1.01
CHORUS	1.07	1.06
NuTeV	0.48	0.55
DYE605	1.07	1.02
DYE866	1.61	1.62
CDFWASY	1.66	1.70
CDFZRAP	2.15	2.12
D0ZRAP	0.64	0.63
ATLAS-WZ	1.94	1.46
CMS-WEASY	1.37	0.96
LHCb-WZ	1.33	1.22
CDFR2KT	0.67	0.67
D0R2CON	0.94	0.93
ATLAS-JETS-2010	1.45	1.42

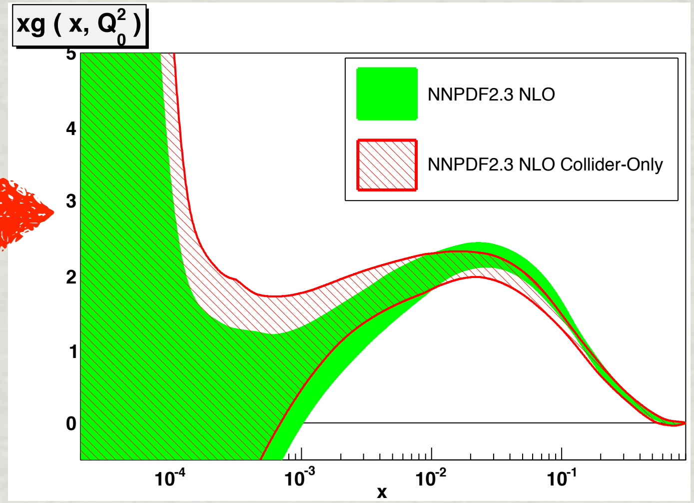


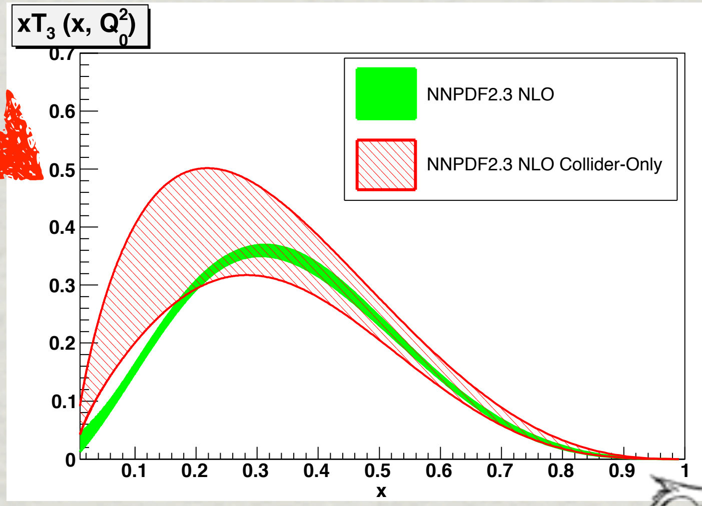
NNPDF2.3
Collider fit - are we there yet?

It is the fit we would love to have

- Only **high energy data**: minimize the effects of higher-twist contributions
- Only **proton data**: no assumptions based on models for nuclear corrections

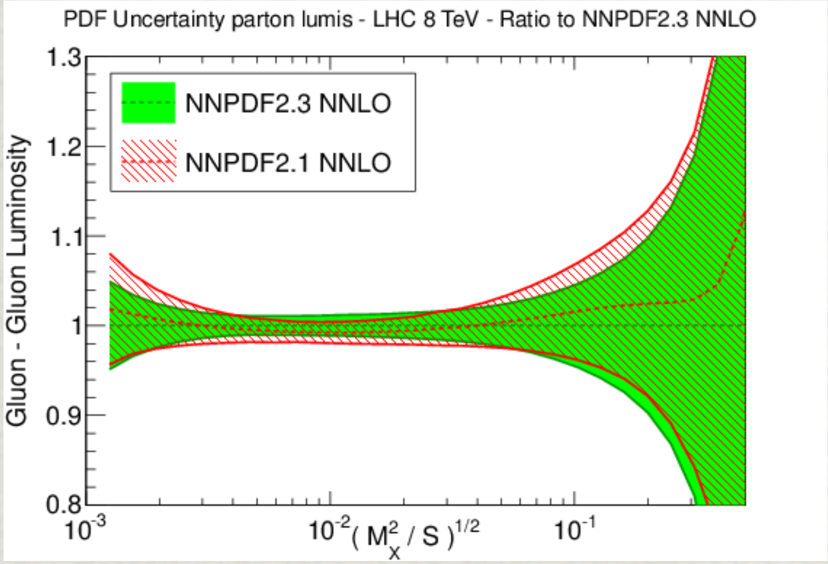
- **Gluon distribution** is very **well constrained** both at small-x (HERA) and large-x (Tevatron/LHC jets)
- PDF combinations sensitive to **light flavour separation** have **substantially larger uncertainties** (missing constraints from fixed target DIS/DY data)
- **Uncertainties** on **“fixed target” observables** are still unacceptably **large**
- ... things can only get better with more LHC data coming (W+c, low mass DY, photons, high pt Z/W ...)

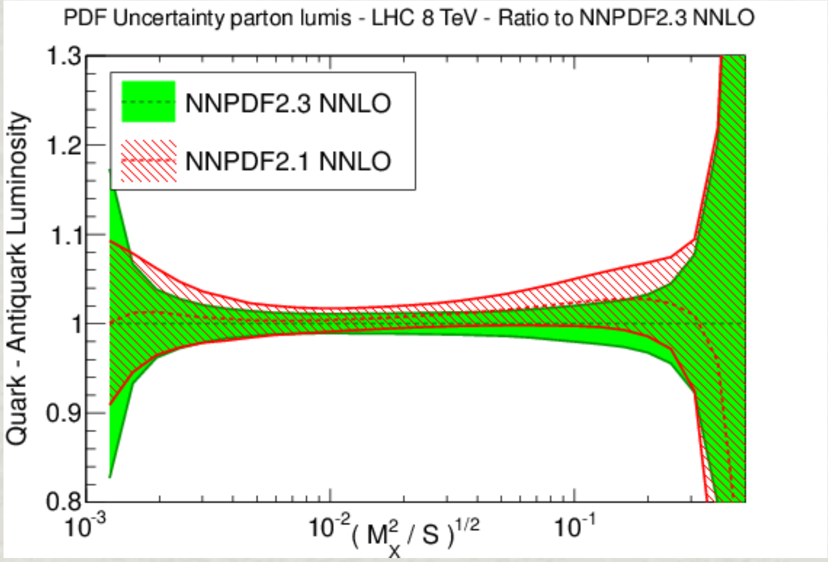


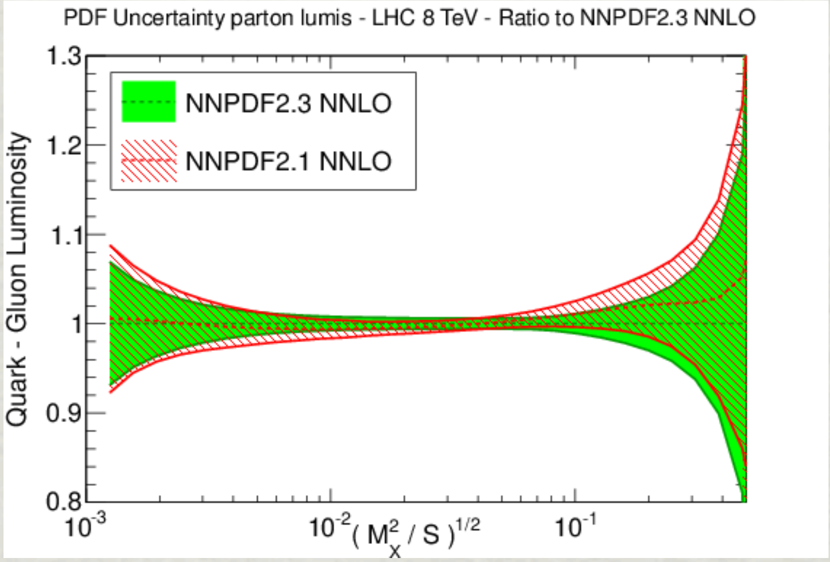


NNPDF2.3
Phenomenology - parton luminosities

$$\Phi_{ij}\left(M_X^2\right) = \frac{1}{s}\int_\tau^1 \frac{dx_1}{x_1} f_i\left(x_1, M_X^2\right) f_j\left(\tau/x_1, M_X^2\right)$$



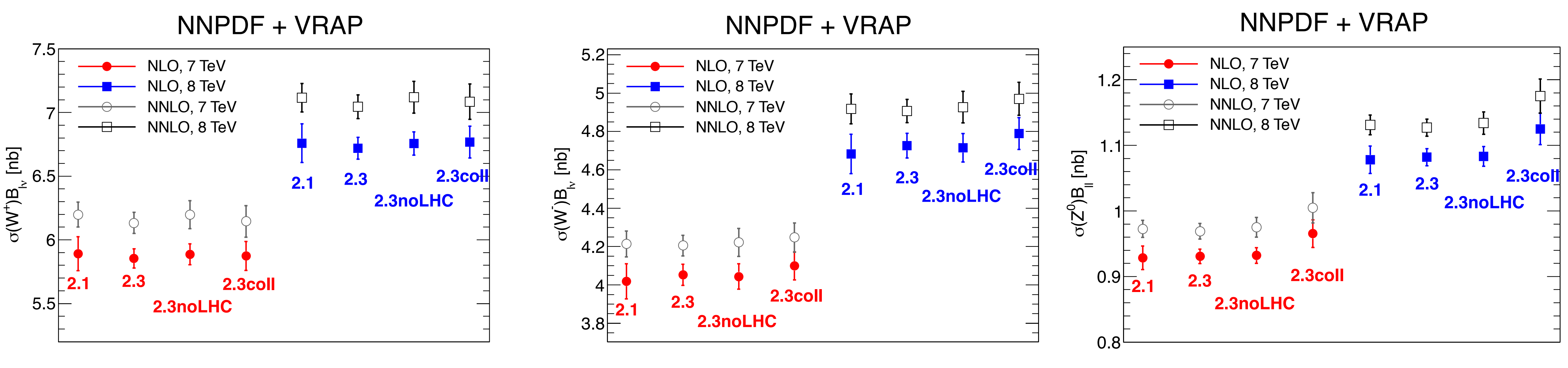




- Reduction in uncertainty on gluon-gluon luminosity for larger final state invariant masses when going from NNPDF2.1 to NNPDF2.3
- NNPDF2.3 quark-antiquark luminosity at large invariant masses somewhat smaller than NNPDF2.1

NNPDF2.3

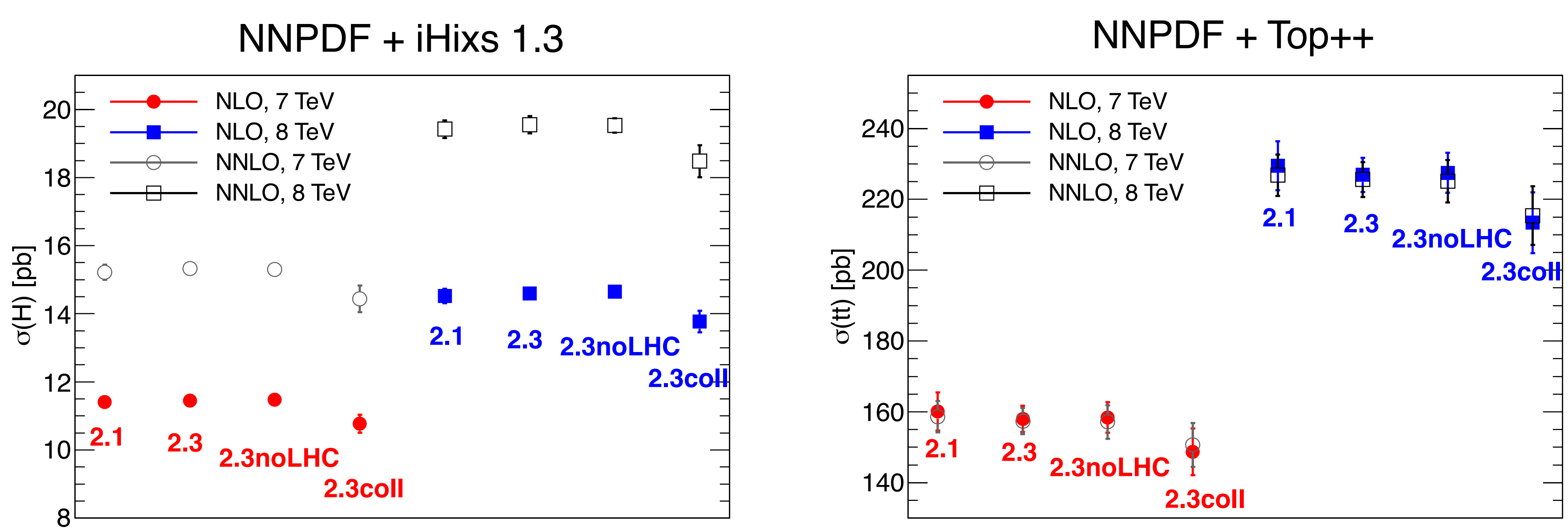
Phenomenology - W/Z production



- Mostly sensitive to quark parton luminosities
- Predictions from NNPDF2.3 sets are compatible with each other and with predictions obtained using the NNPDF2.1 global set
- Largest differences with collider only fit, although the latter has larger uncertainties

NNPDF2.3

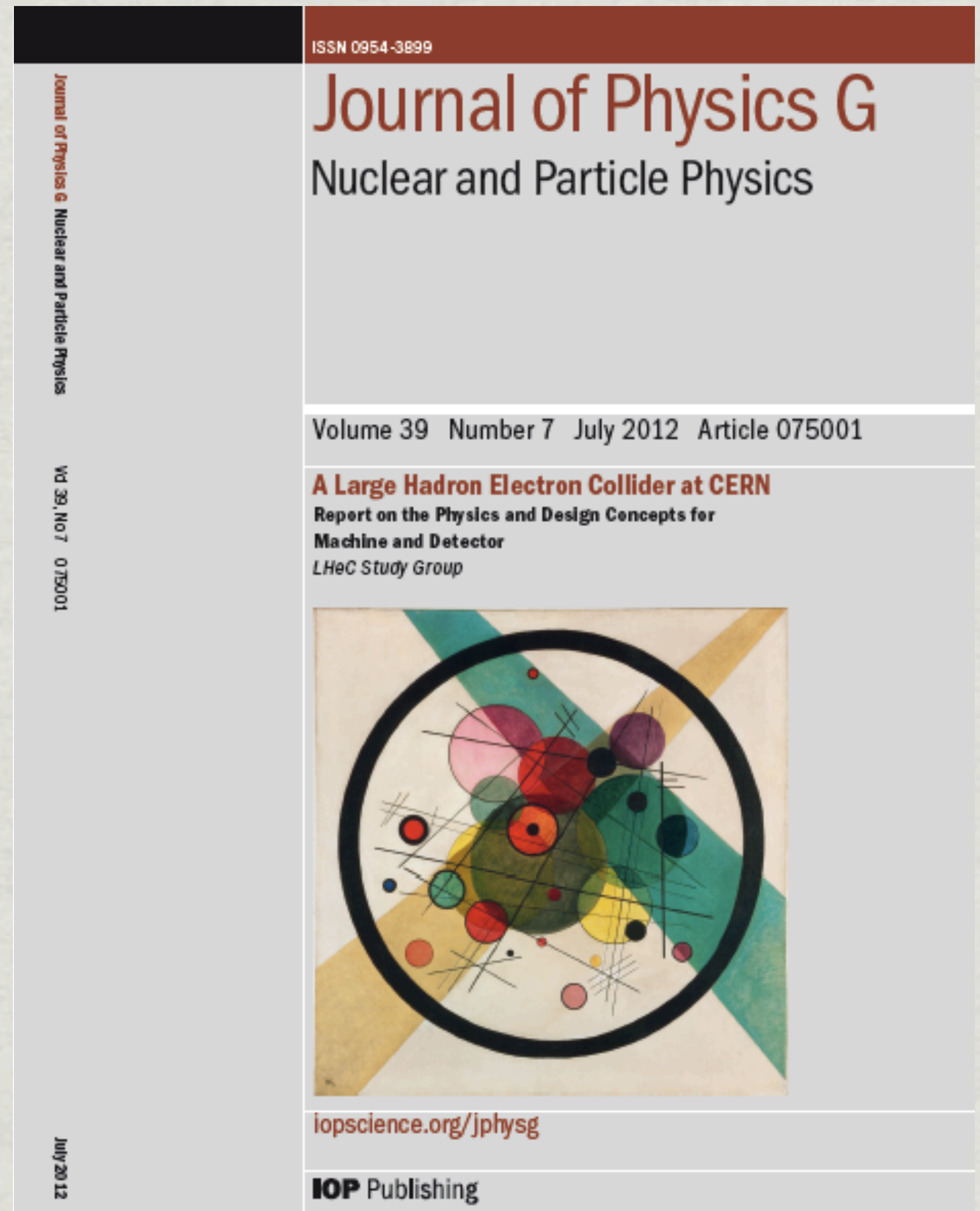
Phenomenology - top/Higgs production



- Mostly sensitive to quark parton luminosities
- Predictions from NNPDF2.3 sets are compatible with each other and with predictions obtained using the NNPDF2.1 global set
- Largest differences with collider only fit, although the latter has larger uncertainties

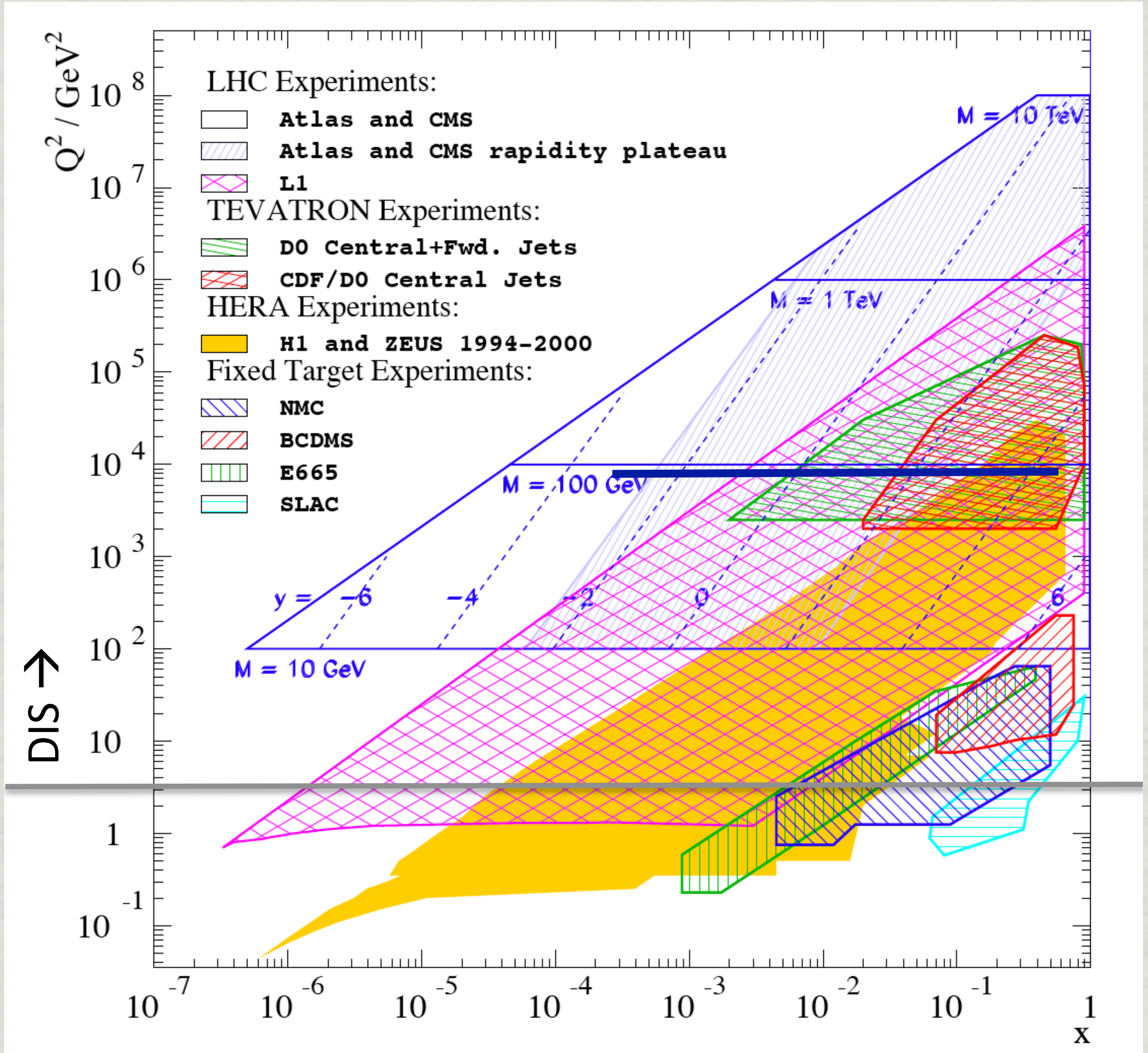
The LHC
Lepton-hadron collision at the LHC

- The LHeC is a proposed facility at CERN, where a (7 TeV) LHC proton beam is brought into collision with a lepton beam
- The LHeC is designed to operate simultaneously with the LHC
- Two options for the lepton beam are considered
  - Linac-Ring option
  - Ring-Ring option
- Many more details can be found in the recently published Conceptual Design Report (arXiv:1206.2913)



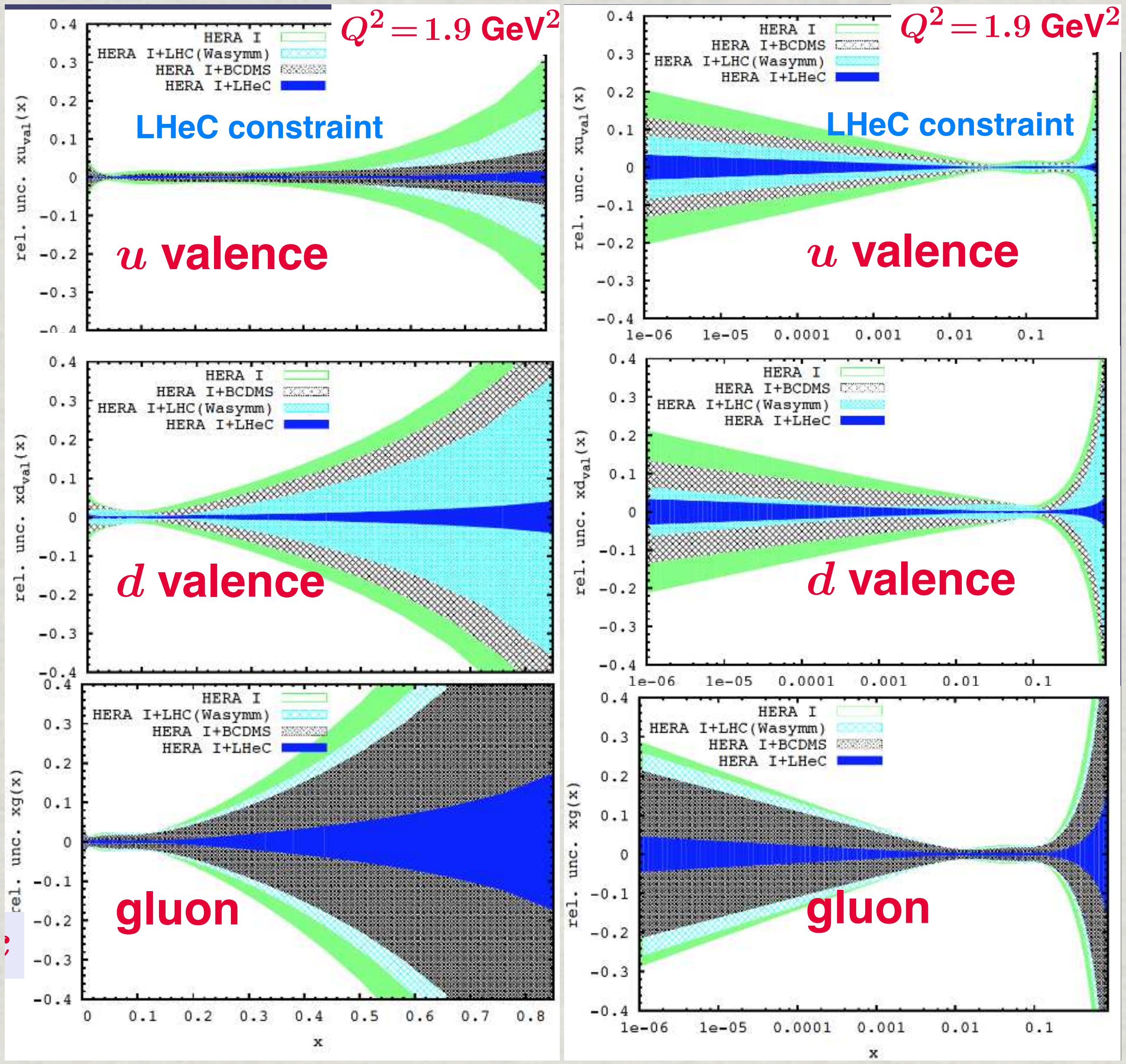
The LHC
Potential to constrain PDFs

- The measurements at the LHeC will nicely complement the pp and pA measurements from the LHC experiments
- Unique possibility to explore the small-x region in DIS
- Precise measurements of the Neutral and Charrged Current DIS cross-section at large-x for accurate flavour separation
- Precise determination of heavy flavour parton distributions



The LHC
Potential to constrain PDFs

- The measurements at the LHeC will nicely complement the pp and pA measurements from the LHC experiments
- Unique possibility to explore the small-x region in DIS
- Precise measurements of the Neutral and Charrged Current DIS cross-section at large-x for accurate flavour separation
- Precise determination of heavy flavour parton distributions



Conclusions & Outlook

- Parton Distribution Function are an essential ingredient of theoretical predictions for observables at hadron-hadron and lepton-hadron colliders

- The NNPDF methodology is ideally suited to tackle the shortcomings of the standard PDF fitting methodology

- The reweighting technique can be used to quickly and reliably assess the impact of new data in parton distribution functions fits without the need for refitting

- NNPDF2.3 is the first global PDF fit to include LHC data

- The proposed LHeC experiment will be a unique tool to understand the structure on the proton and accurately determine parton densities