Parton Distribution Functions & the LHC

Alberto Guffanti

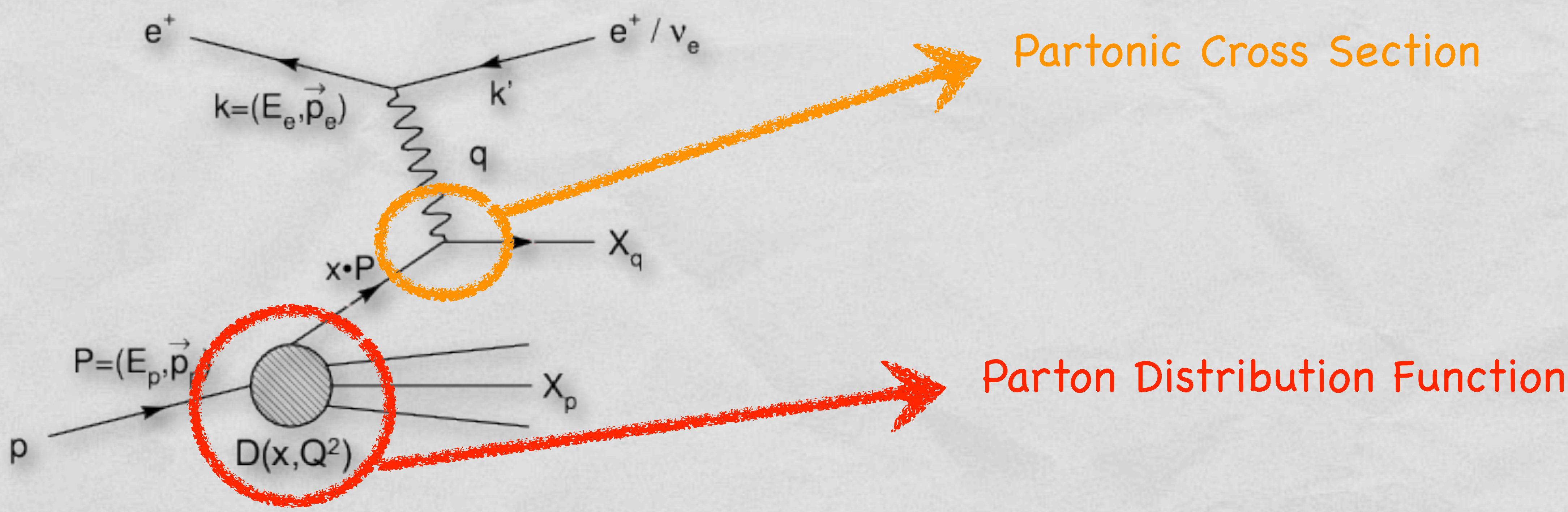
Niels Bohr International Academy & Discovery center
Niels Bohr Institute - Copenhagen

Zürich Phenomenology Workshop 2013
Zürich, 7-9.1.2013

Parton Distribution Functions

What are they? - Definition

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? - DGLAP evolution

- 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)]

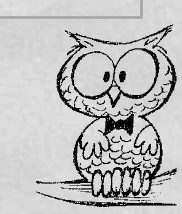
[**NLO** - Floratos, Ross, Sachrajda; Gonzalez-Arroyo, Lopez, Yndurain; Curci, Furmanski, Petronzio (1981)]

[**NNLO** - Moch, Vermaseren, Vogt (2004)]

The PDF fitting game

The players

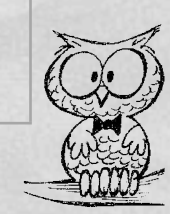
Collaboration	Authors	arXiv
ABM	S. Alekhin, J. Blümlein, S. Moch	1202.2281, 1105.5349, 1101.5261, 1107.3657, 0908.3128, 0908.2766, ...
CTEQ/TEA	M. Guzzi J. Huston, H.-L. Lai, P. Nadolsky, J. Pumplin, D. Stump, C.-P. Yuan	1108.5112, 1101.0561, 1007.2241, 1004.4624, 0910.4183, 0904.2424, 0802.0007, ...
GJR/JR	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	1211.1215, 1205.4024, 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	1207.1300, 1110.2483, 1108.2758, 1107.2652, 1103.2369, 1102.3182, 1101.1300, 1005.0397, 1002.4407, ...



The PDF fitting game

Present 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 ($\Delta\chi^2$~25)
NNPDF2.3	**Global** (includes LHC data)	LO (2.1) NLO NNLO	GM-VFNS (FONLL)	External (multiple values available)	7 indep. PDFs Neural Nets (259 param.)	Monte Carlo

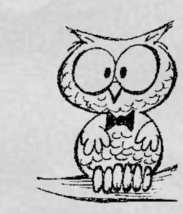


Benchmarking PDFs with LHC data

The rules of the game

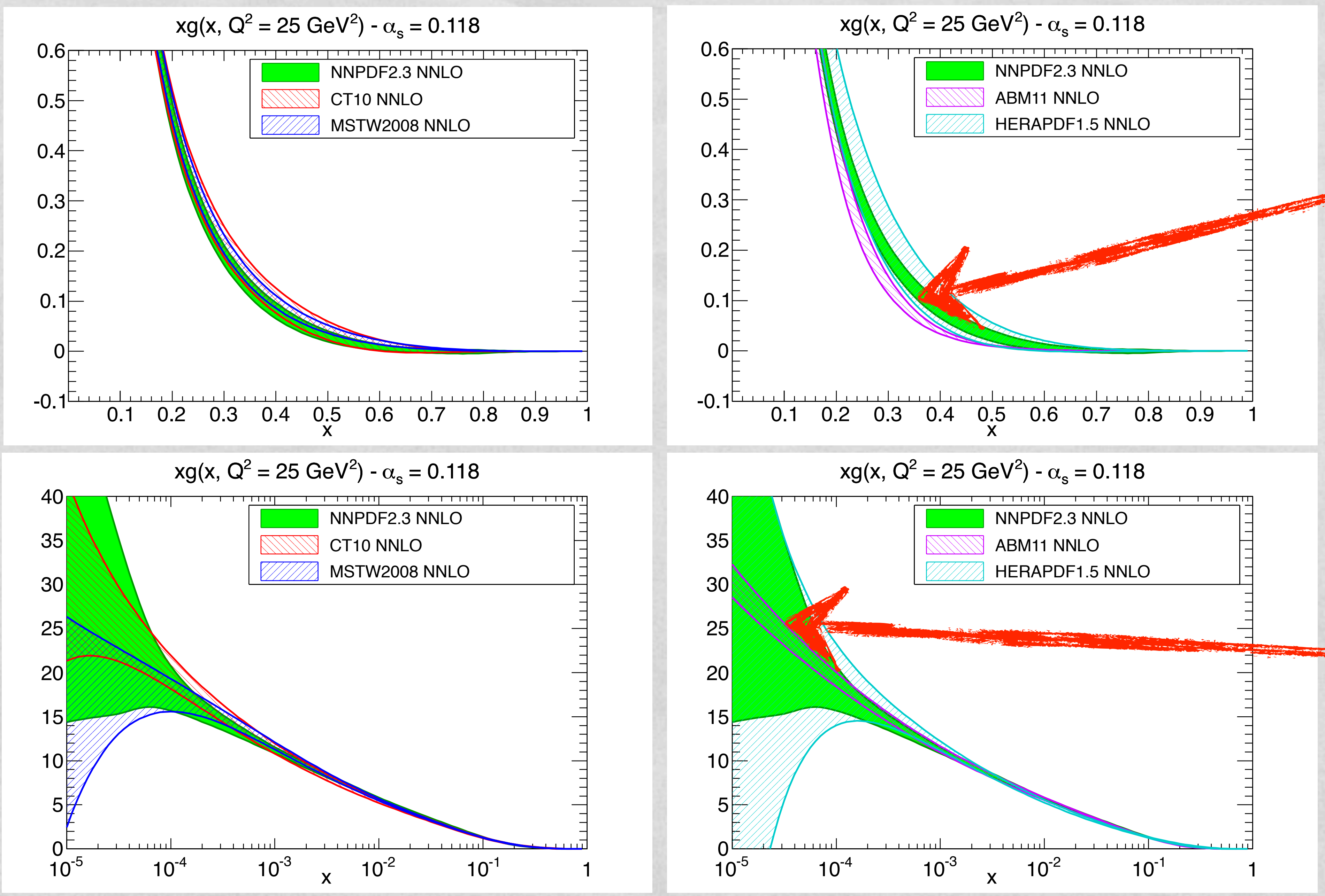
[J. Rojo et al., arXiv:1211.5142]

- **Systematically compare** PDFs, parton luminosities and predictions for observables at the LHC using the most up-to-date releases from different PDF fitters groups
    - ABM11
    - CT10
    - HERAPDF1.5
    - MSTW08
    - NNPDF2.3
- Compare consistently using a **common value for** α_s in order to disentangle PDF and strong coupling uncertainties (JR not available for multiple α_s values)
- When comparing to **differential distributions**, **quantitatively** describe the agreement giving χ^2 values, **DO NOT** rely only on plots, which could be misleading due to large systematic uncertainties



Benchmarking PDFs

Parton Distributions - Gluon

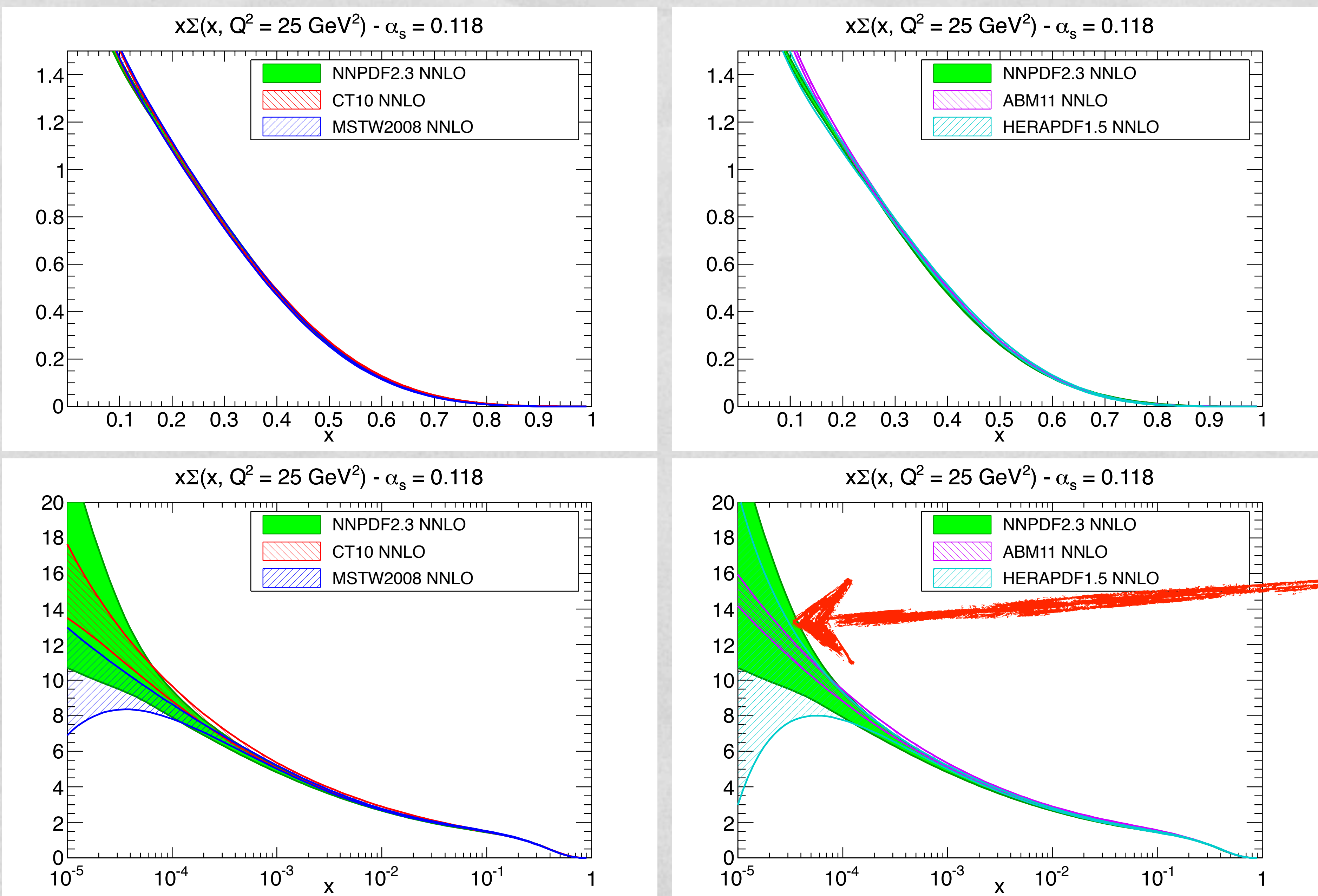


ABM gluon softer at medium-/large-x (Inclusive Jet data not used in the fit)

Unnaturally small uncertainties in extrapolation regions (where there are no data) might signal a parametriation bias

Benchmarking PDFs

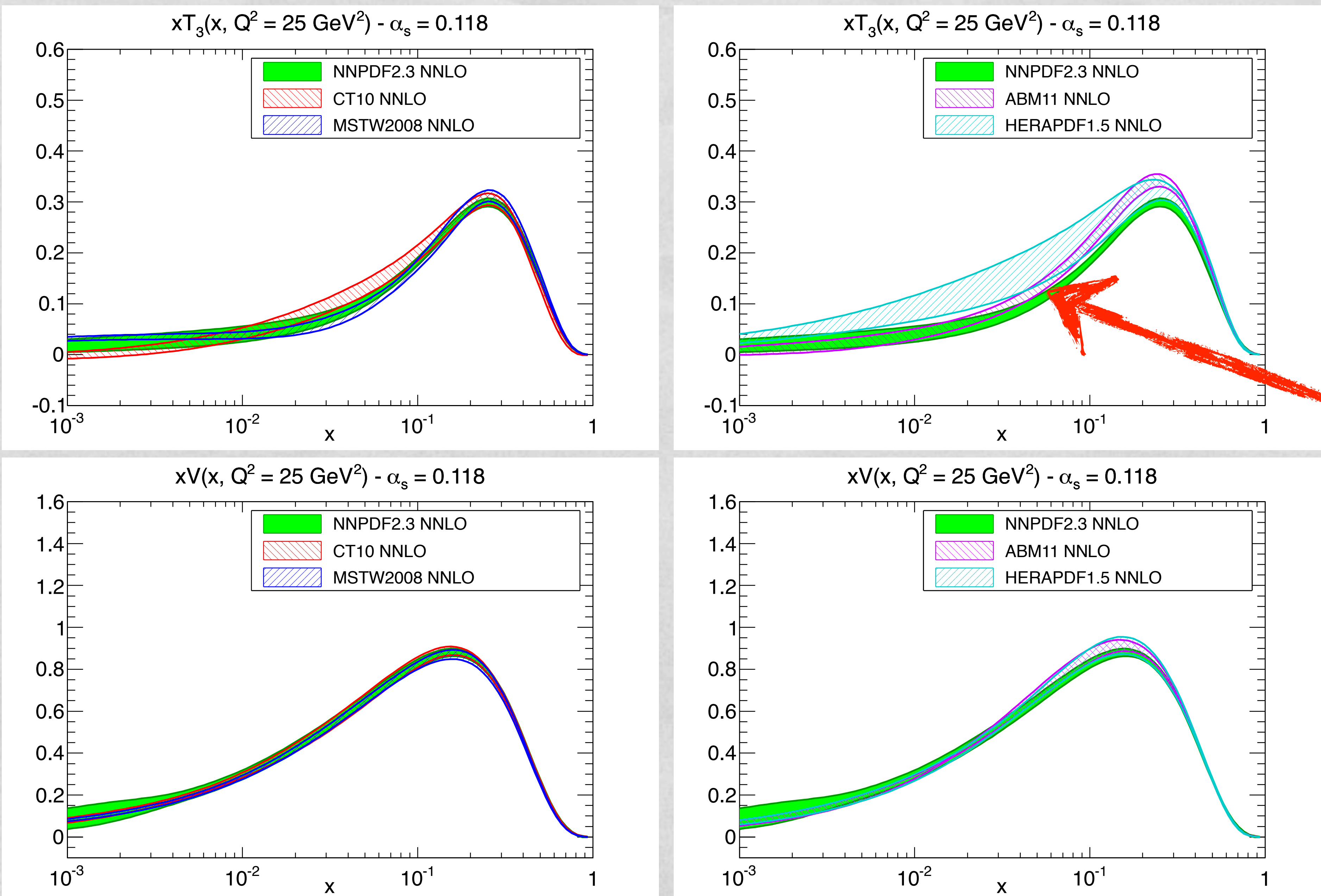
Parton Distributions - Sea Quarks



Unnaturally small uncertainties in extrapolation regions (where there are no data) might signal a parametriation bias

Benchmarking PDFs

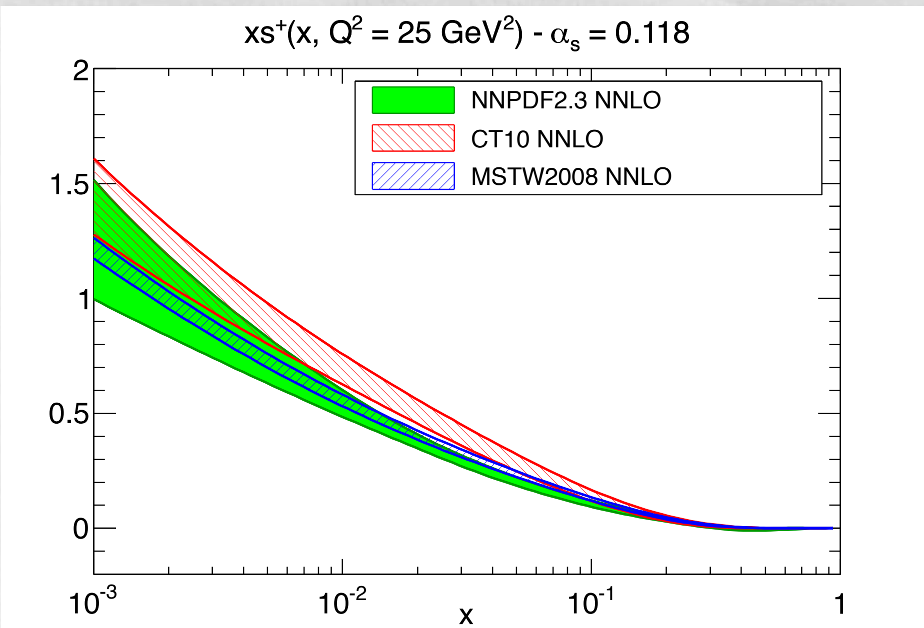
Parton Distributions - Valence Quarks

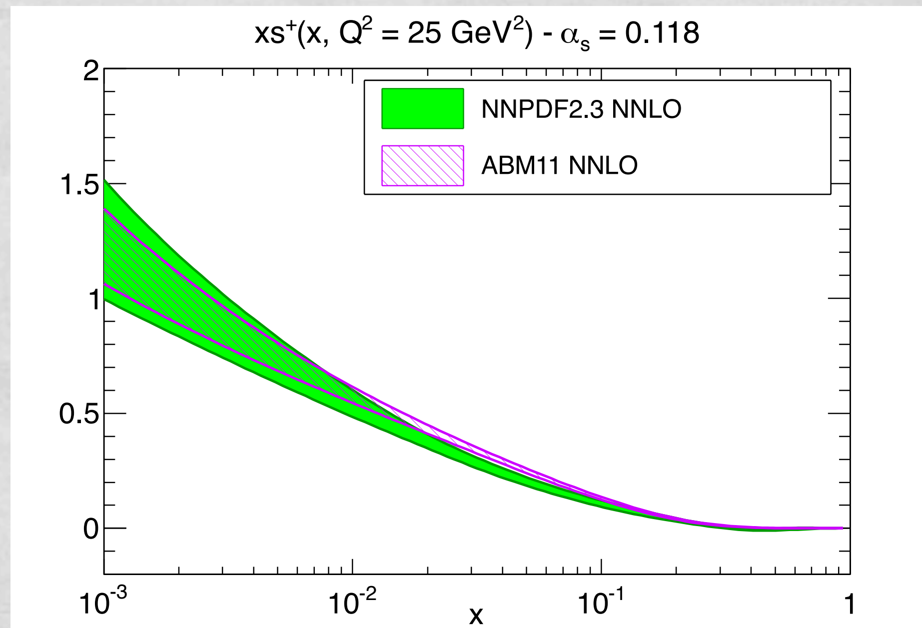


HERA data provide little information on flavour separation at medium-/large-x

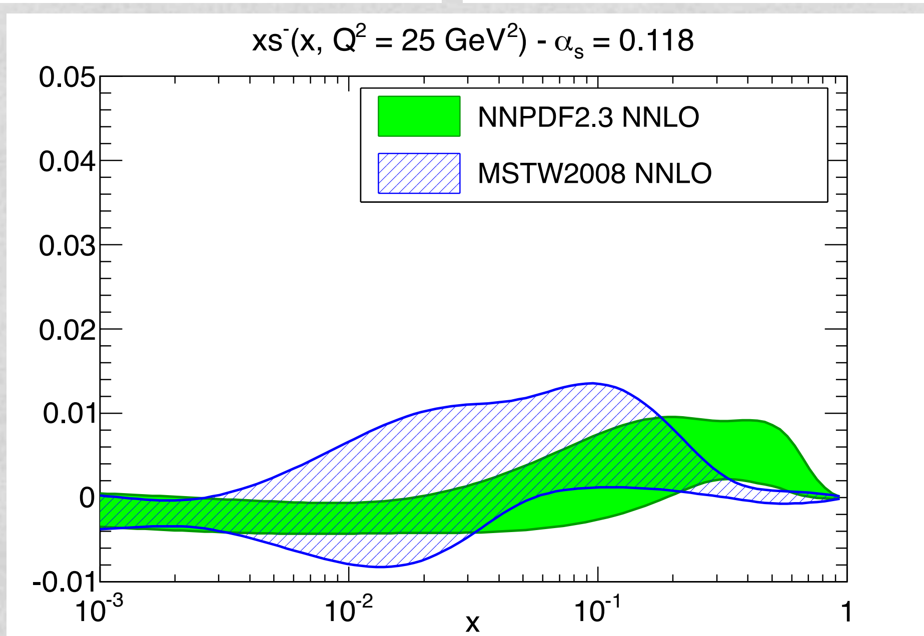
Benchmarking PDFs

Parton Distributions - Strangeness





Strangeness is where global fits differ the most



Only MSTW and NNPDF provide an independent parametrization for $(s-\bar{s})$

Benchmarking PDFs

Parton Luminosities

- Factorized cross-sections at hadron colliders can be written in terms of **parton lumnosities**

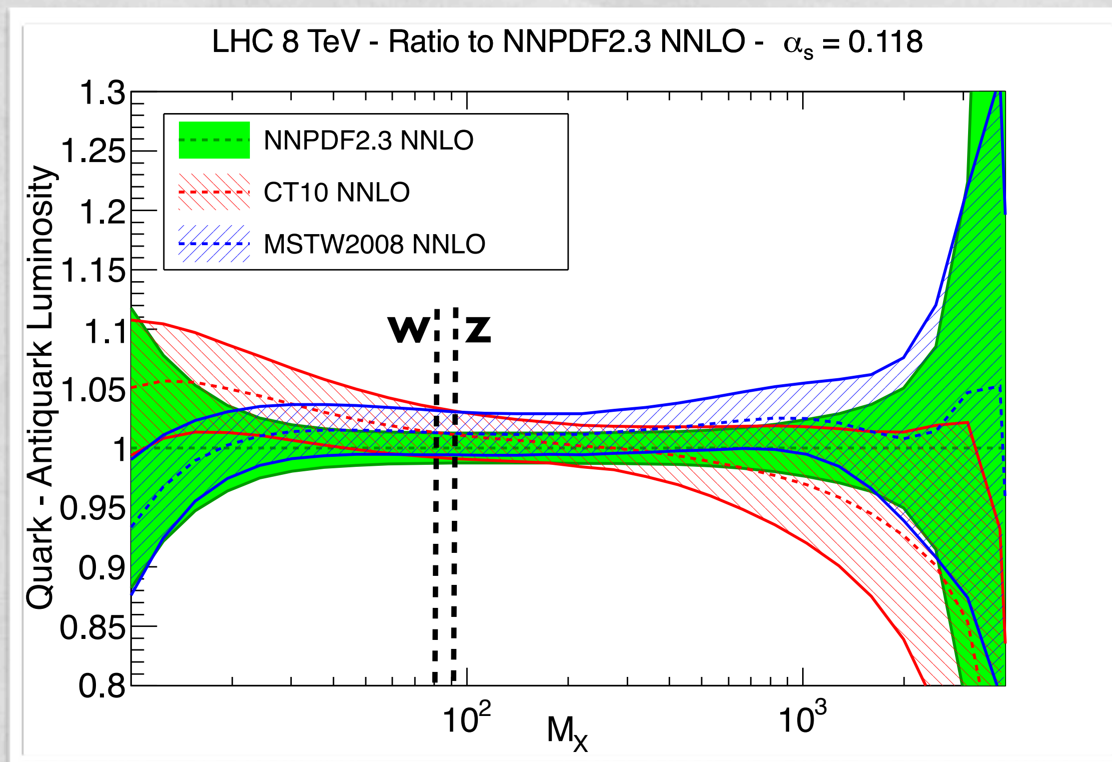
$$\Phi_{ij} = \frac{1}{s}\int_{\tau}^{1} \frac{dx'}{x'} f_i(x', M_X^2) f_j(\tau / x', M_X^2)$$

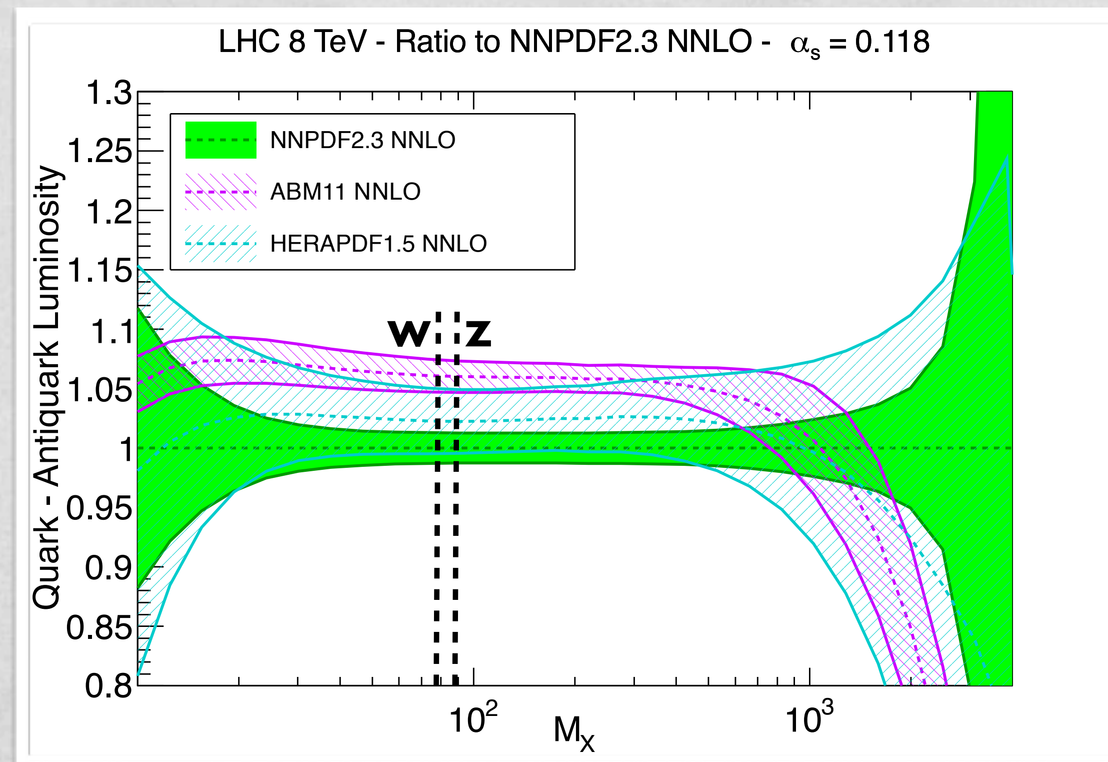
with f_i the PDFs and M_X the invariant mass of the produced final state

- Parton luminosities encode the whole dependence of hadronic cross-sections on PDFs

Benchmarking PDFs

Parton Luminosities

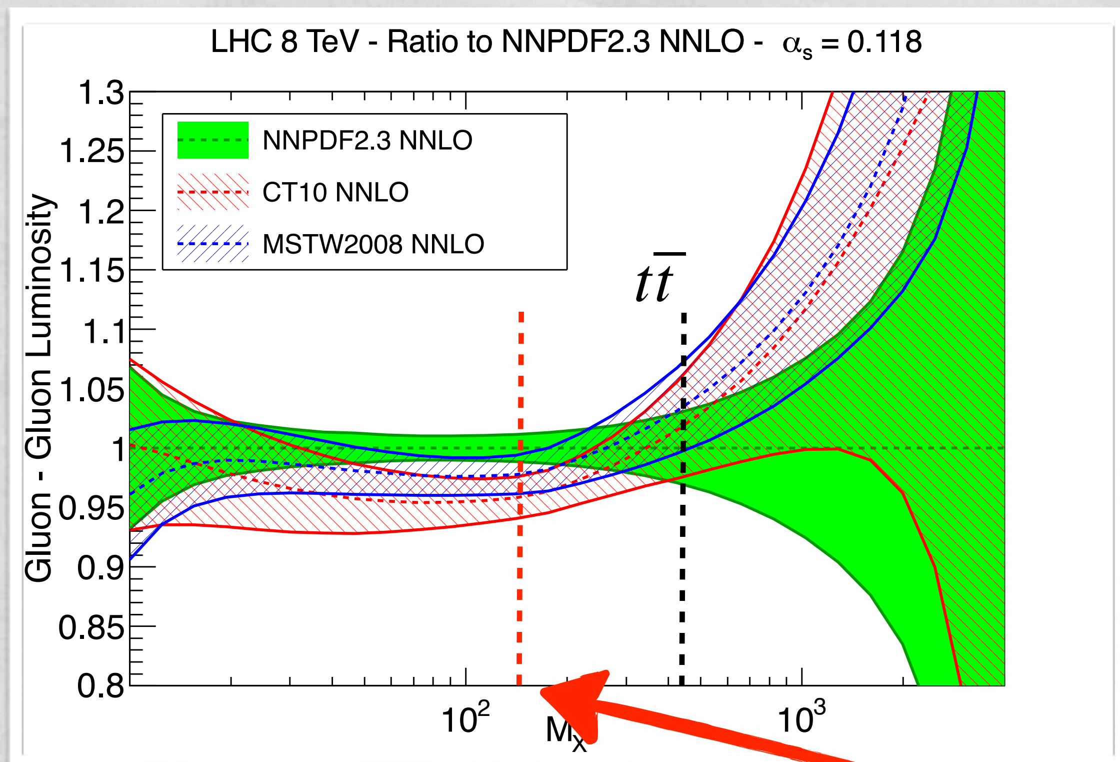


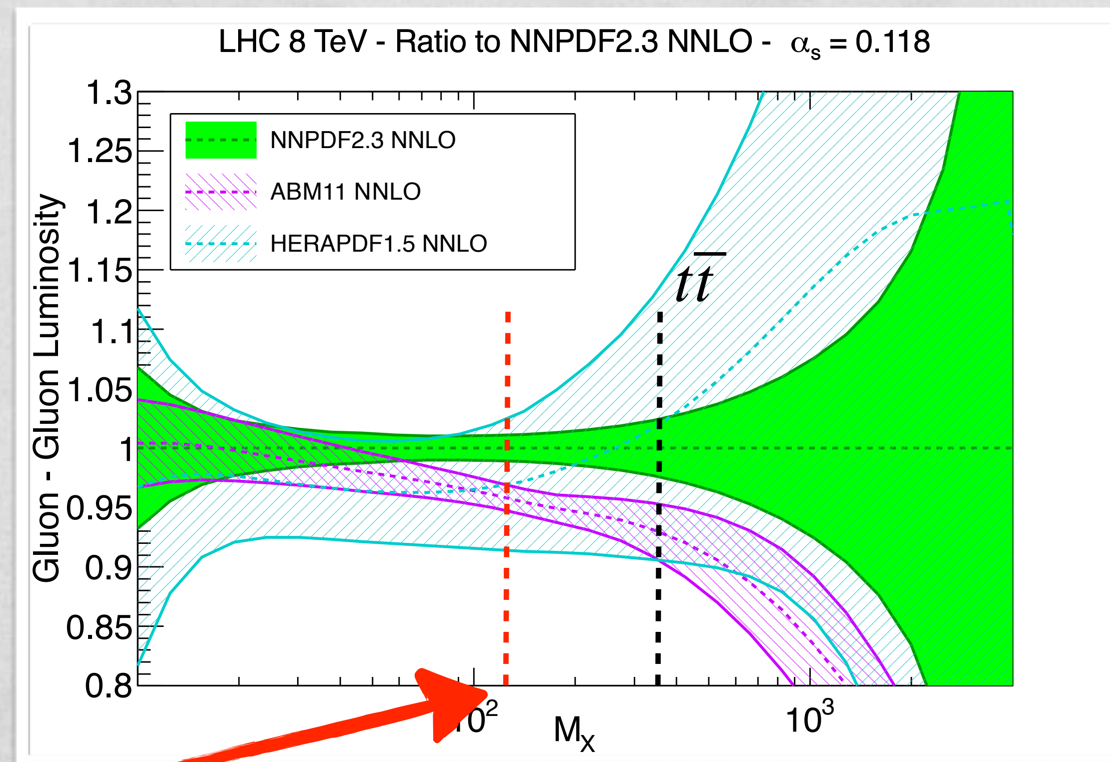


- Directly related to Elctroweak Vector Boson production
- Good agreement between CT10, MSTW, NNPDF and HERAPDF over the whole kinematical range (larger uncertainties on the latter due to smaller dataset)
- ABM11 somewhat larger qqbar luminosity (~8%), partly due to the use of FFN scheme for heavy flavour treatment

Benchmarking PDFs

Parton Luminosities



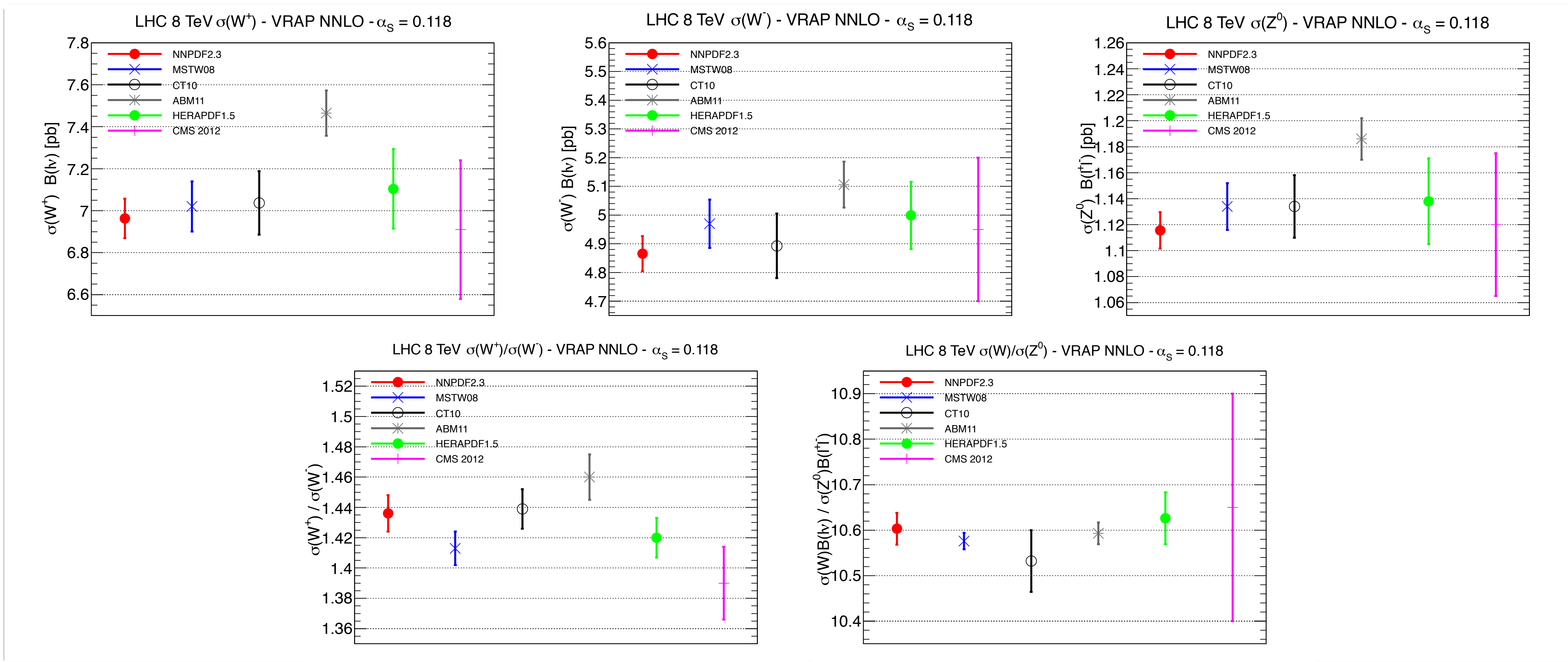


m_h=125GeV

- Probed by **top pair** and **Higgs** production **in gluon-gluon fusion**
- Good agreement between CT10, MSTW, NNPDF and HERAPDF over the whole kinematical range (larger uncertainties on the latter due to smaller dataset)
- ABM11 luminosity generally smaller, except for small invariant masses of the final state

Benchmarking PDFs

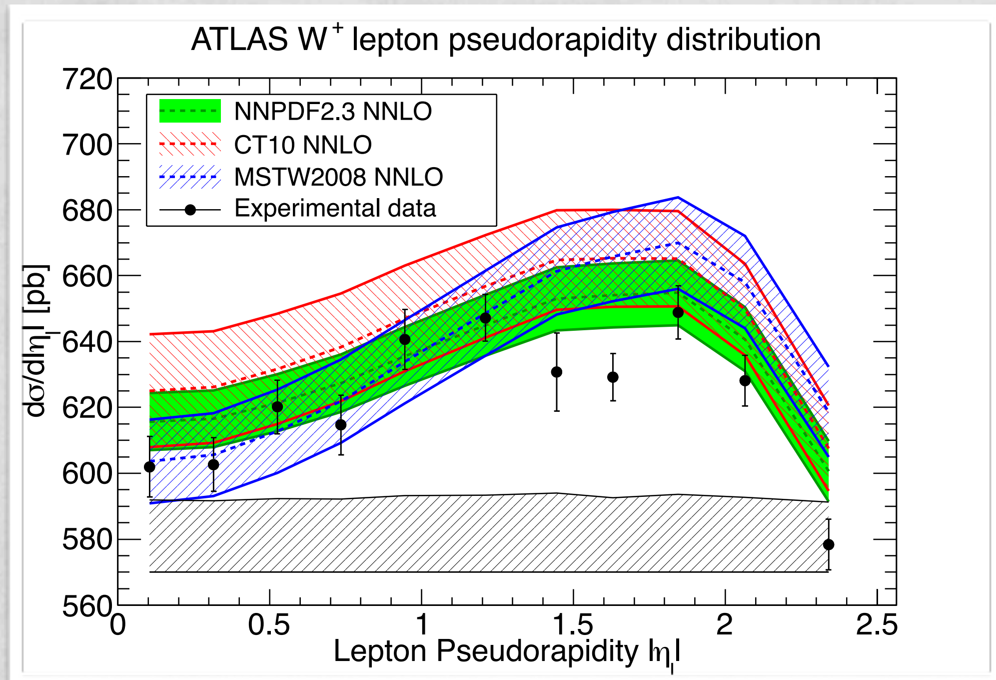
LHC data - W/Z production

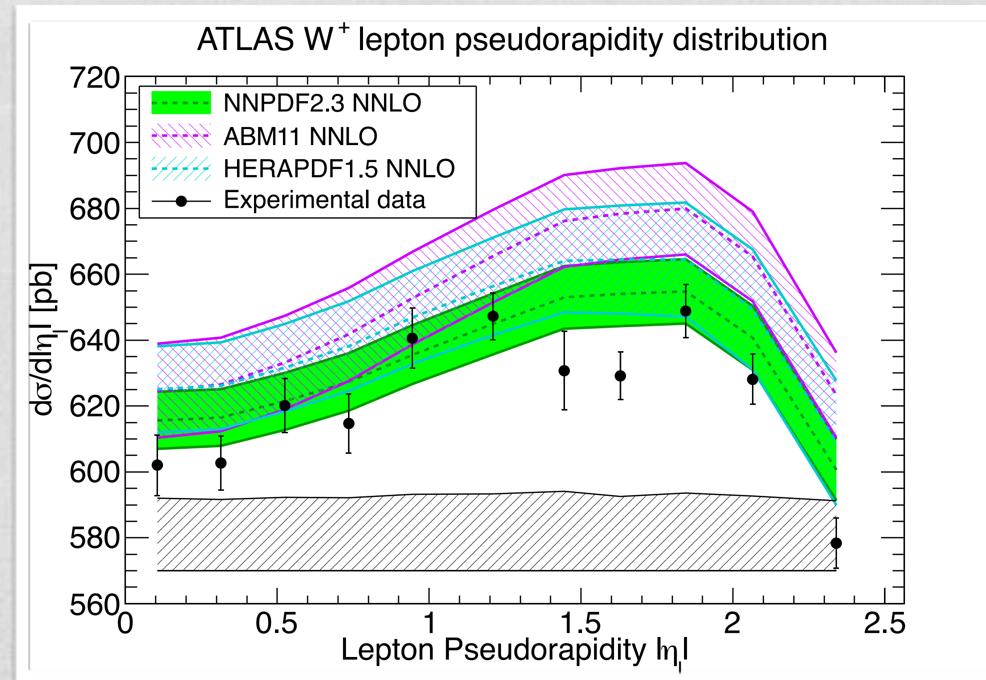


- General agreement among predictions from different PDF sets (albeit with some discrepancies, in particular with ABM) and data will soon be precise to discriminate
- Look at ratios (and double ratios) at different center-of-mass energies

Benchmarking PDFs

LHC data - W/Z differential distributions





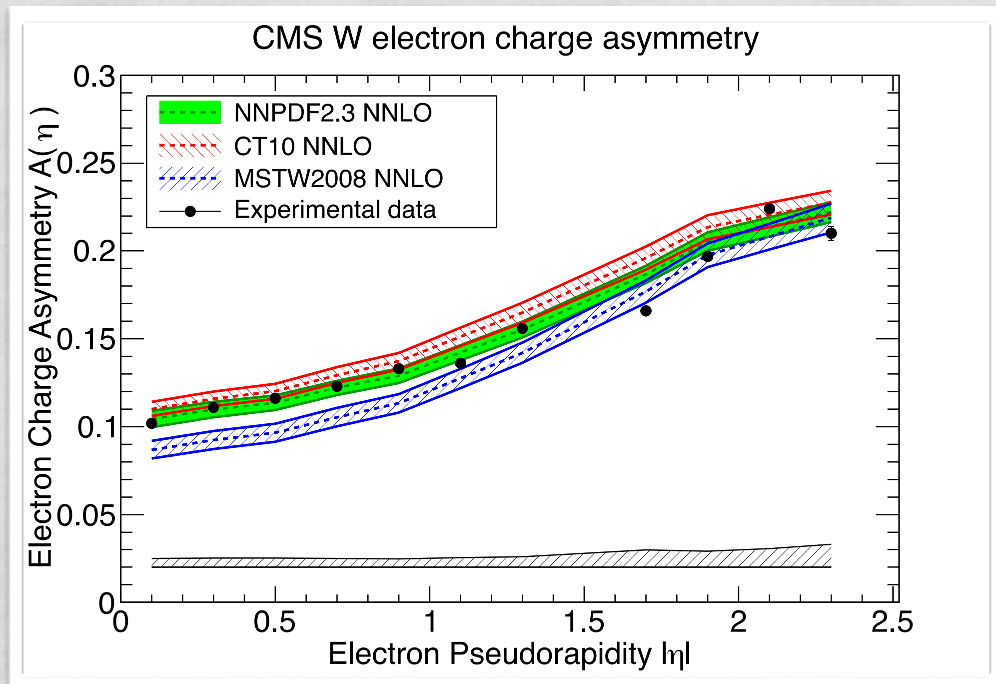
	ATLAS W/Z (2010)
ABM11	1.923
CT10	1.125
HERAPDF1.5	1.845
MSTW08	3.194
NNPDF2.3	1.382

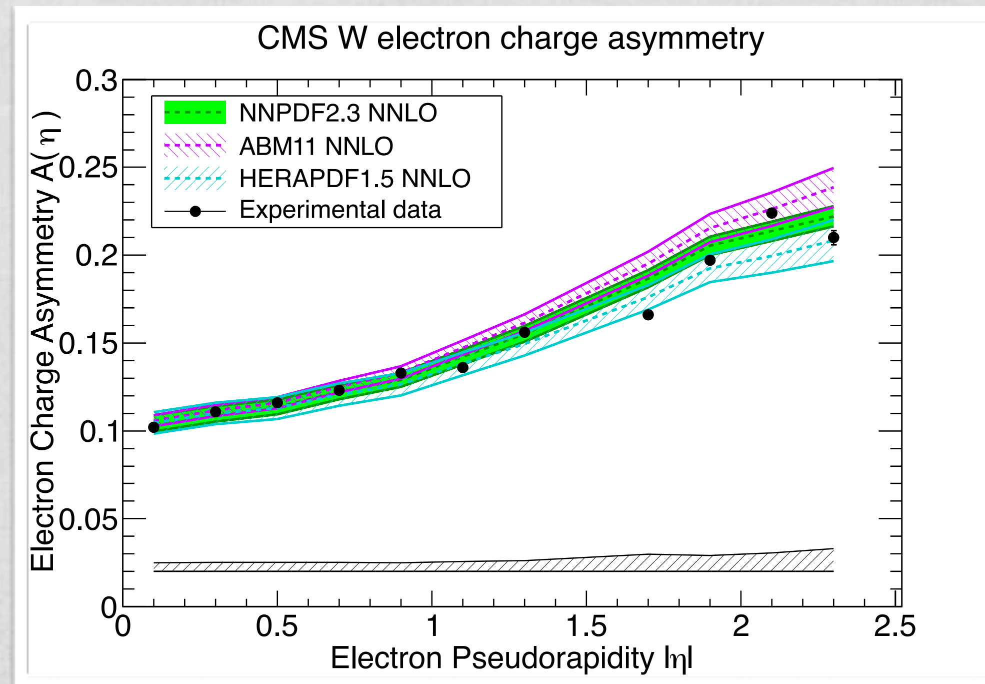
- Constrain **valence and sea quarks** (absolute normalization and flavour separation)
- Limited sensitivity to strangeness
- Poor description from MSTW due to constrained valence quark (**u_v** & **d_v**) parametrization at small-x

[A.D. Martin et al., arXiv:1211.1215]

Benchmarking PDFs

LHC data - W lepton asymmetry

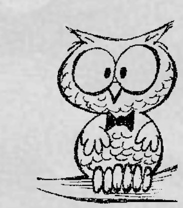




	CMS W Asy (2011)
ABM11	1.602
CT10	1.778
HERAPDF1.5	0.817
MSTW08	4.140
NNPDF2.3	0.828

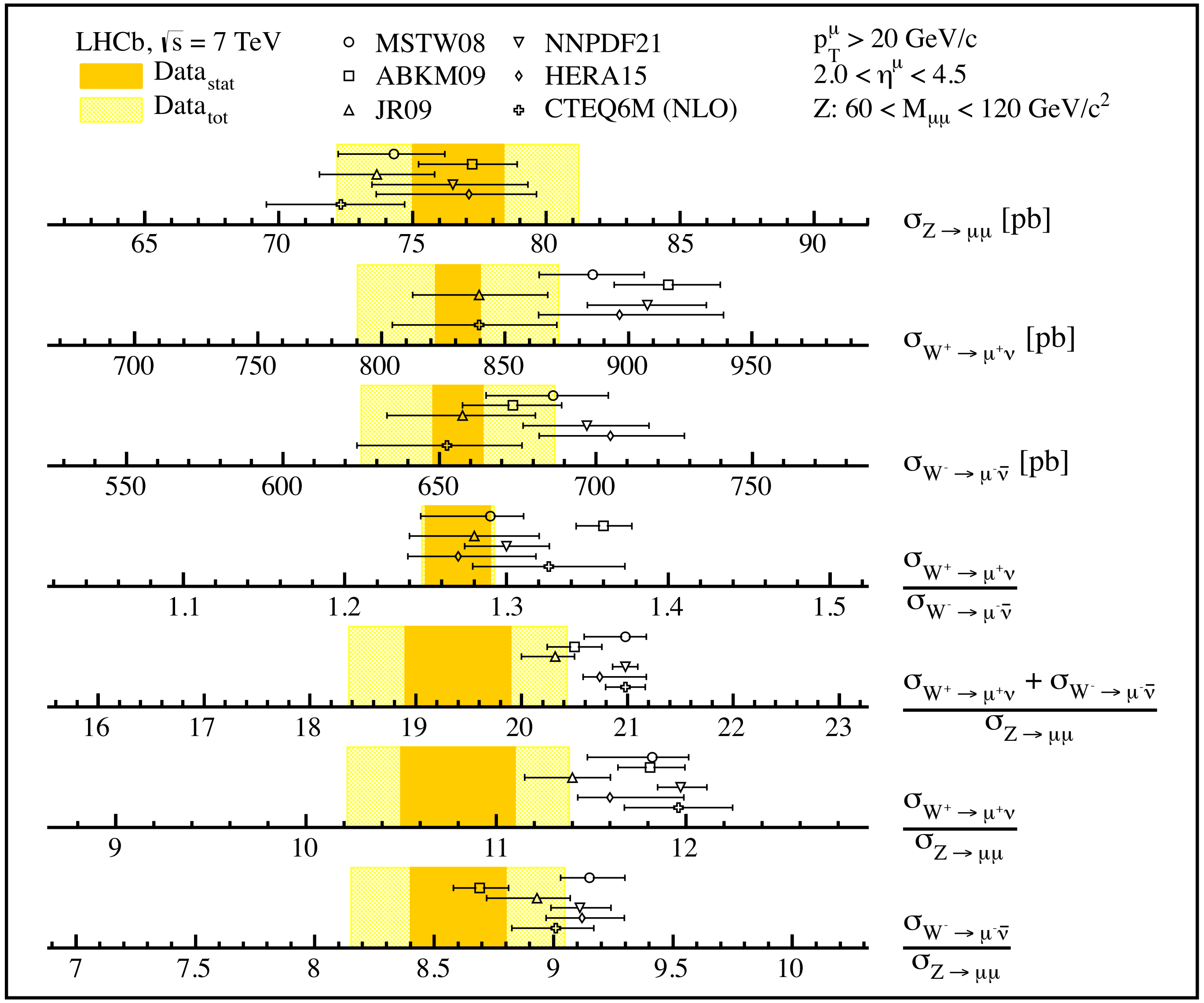
- Direct probe of **light quark flavour separation**
- Limited sensitivity to strangeness
- Poor description from MSTW due to constrained valence quark (**u_v** & **d_v**) parametrization at small-x

[A.D. Martin et al., arXiv:1211.1215]



Benchmarking PDFs

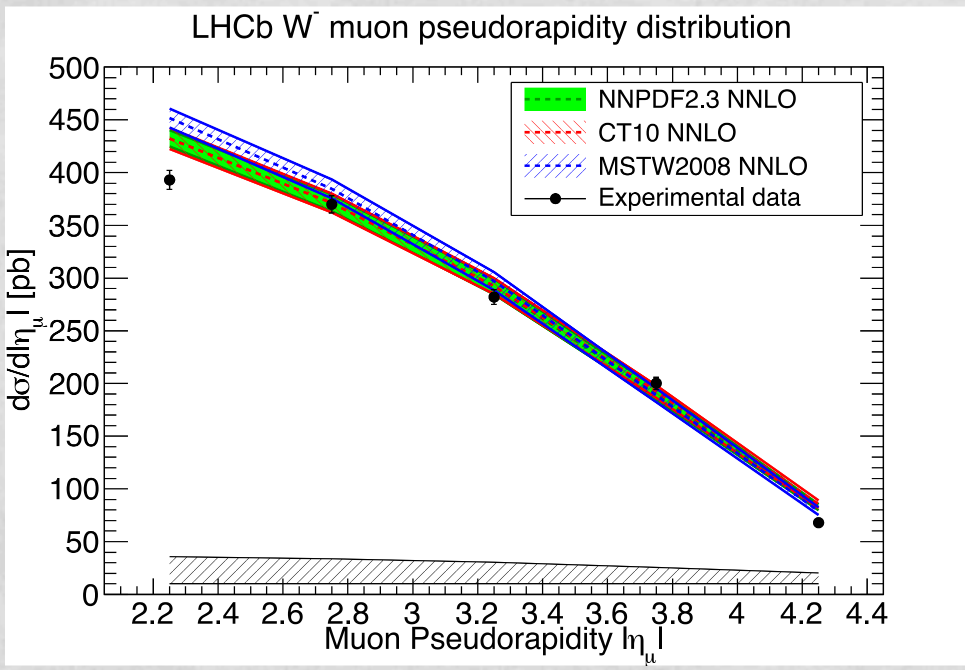
LHC data - W/Z in the forward region

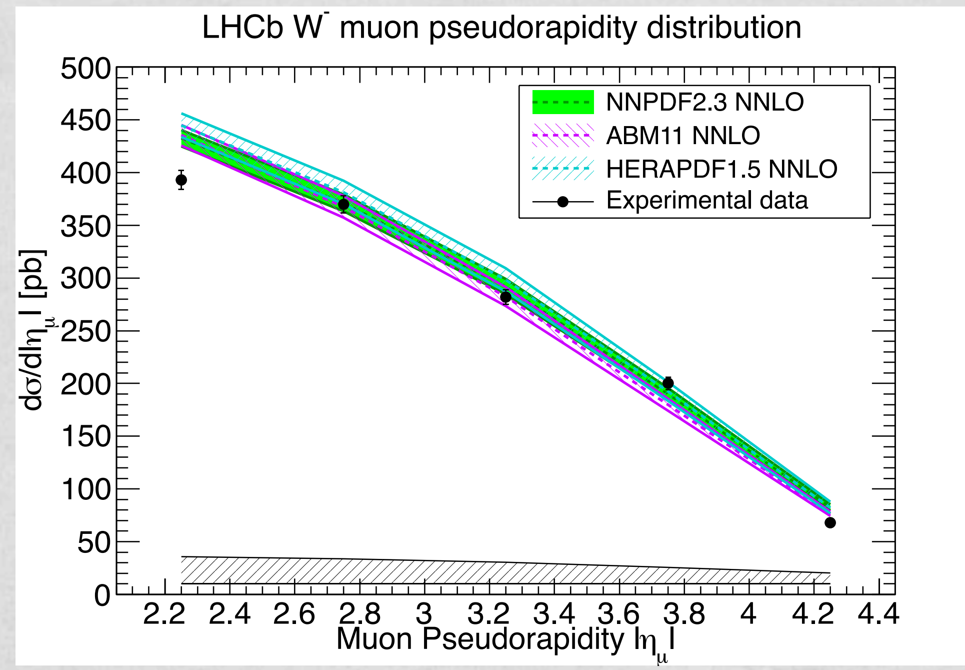


- **LHCb** provides a unique probe of the **forward region**
- **Test** of our knowledge of **PDFs** at **small-x**
- Z measurement recently updated

Benchmarking PDFs

LHC data - W/Z in the forward region



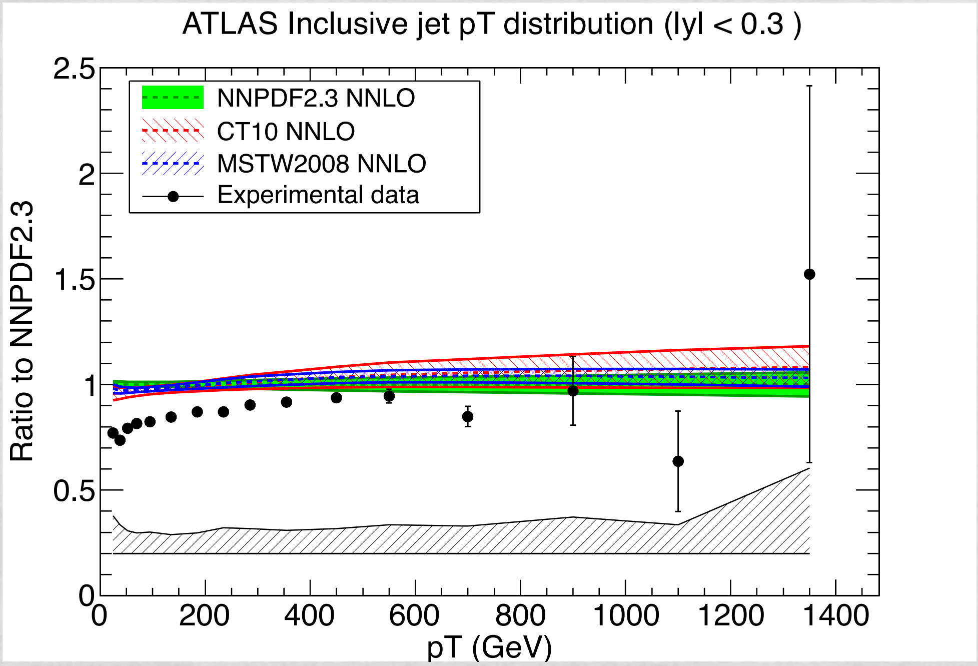


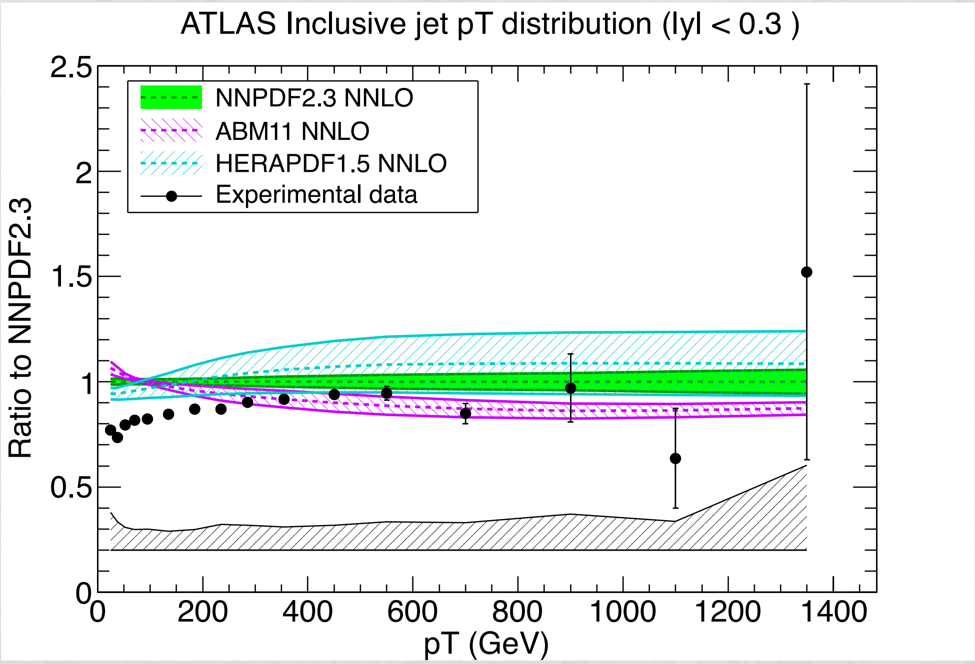
	LHCb W (2010)
ABM11	1.873
CT10	0.892
HERAPDF1.5	0.744
MSTW08	0.956
NNPDF2.3	0.741

- LHCb provides a unique probe of small-x of **light quark distributions**
- Moderate discriminating power among predictions from different sets

Benchmarking PDFs

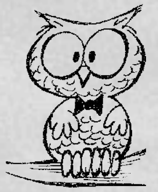
LHC data - Inclusive Jet production





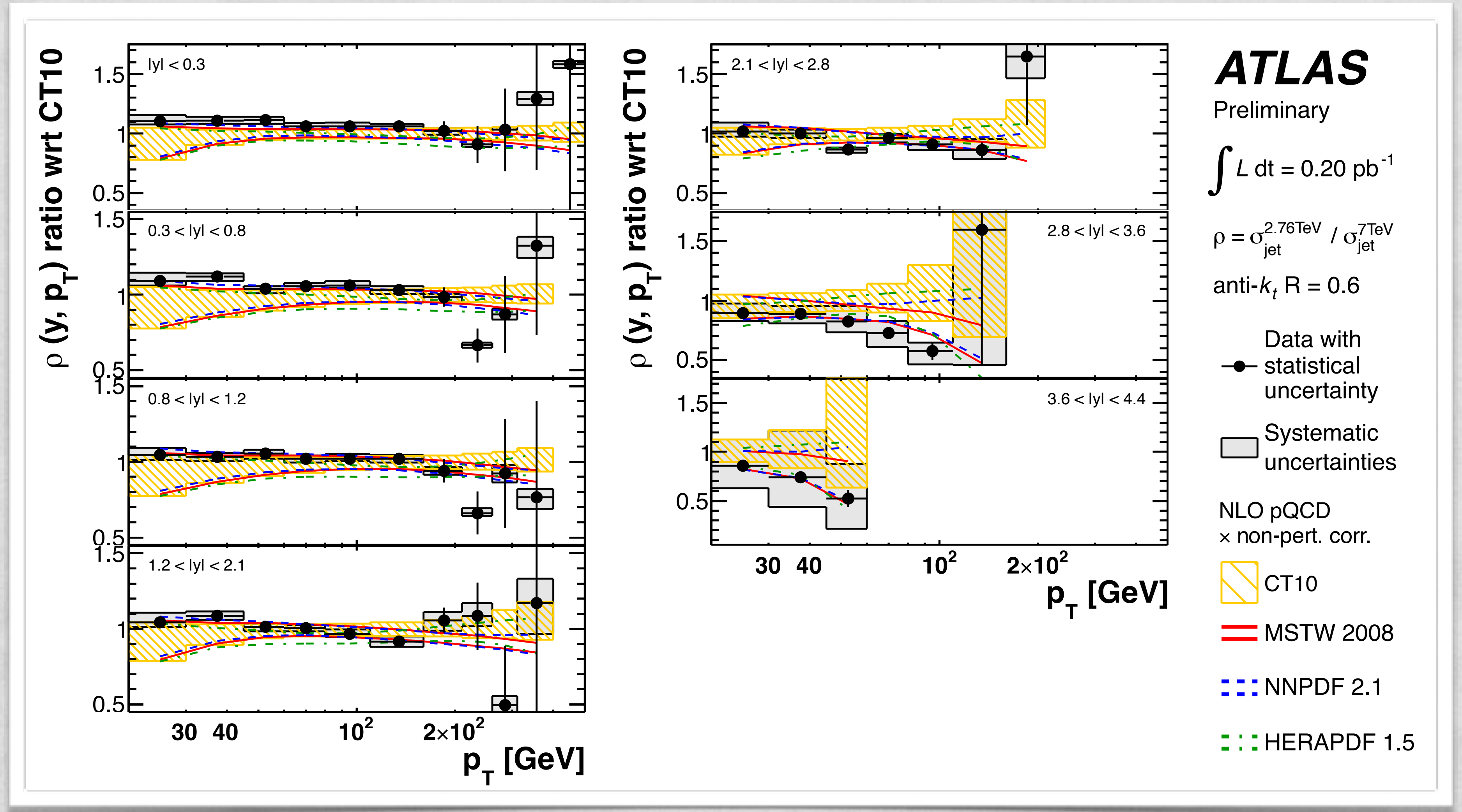
	ATLAS jets (2010)
ABM11	0.963
CT10	0.940
HERAPDF1.5	0.848
MSTW08	0.828
NNPDF2.3	0.862

- Inclusive jet production provides the best constraint on large-x gluon
- All sets give a good description of available data (ATLAS, 2010)
- Beware of visual data-theory comparisons in presence of large systematic uncertainties



Benchmarking PDFs

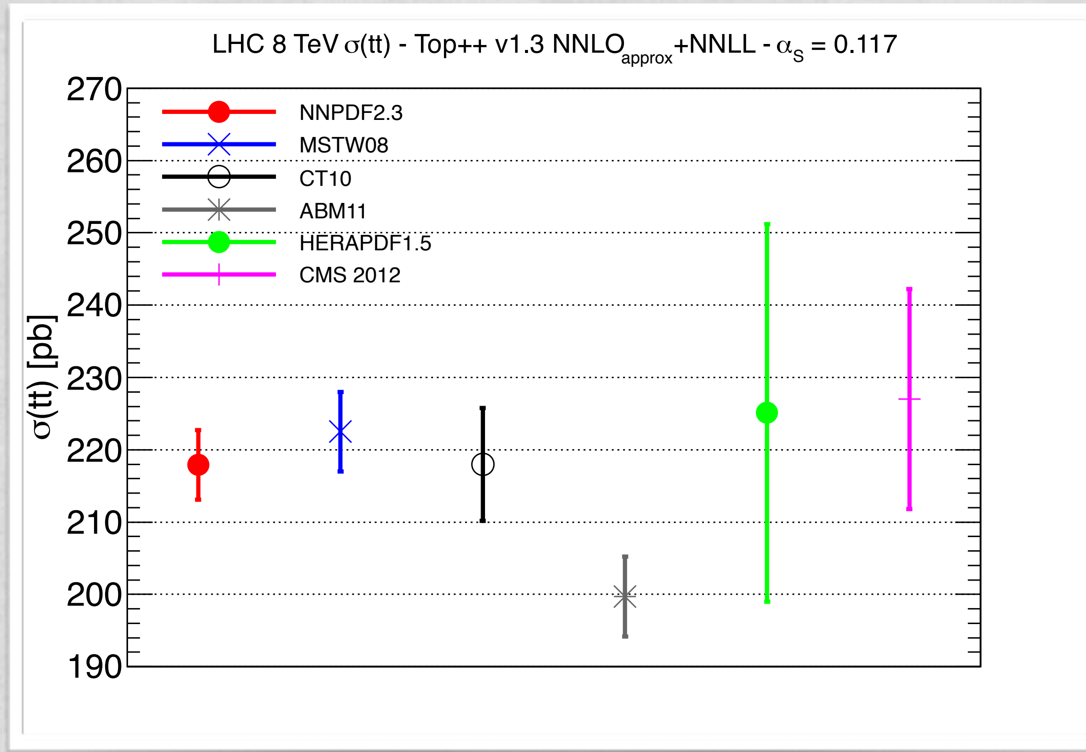
Ratios of Inclusive Jet prod. at different CM energy

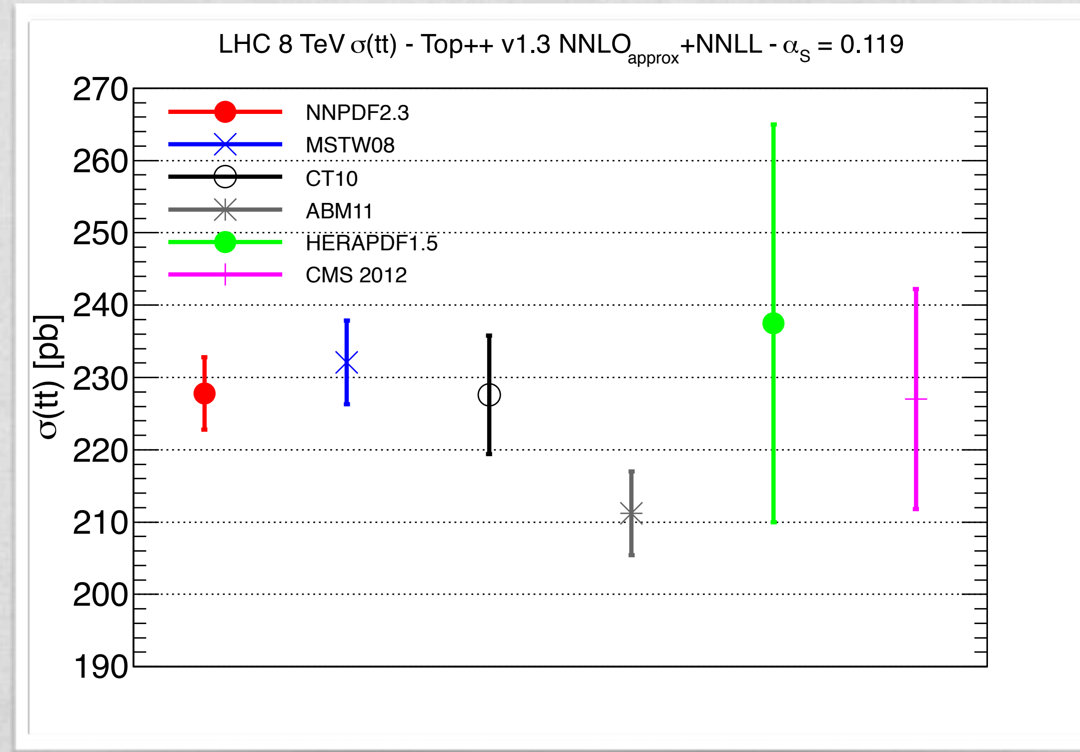


- Substantial **reduction of the experimental uncertainties** in the ratio, especially at low p_T
- Interesting exercise to be repeated with the higher statistics data from the 8 TeV run (expect more substantial reduction in theory uncertainties)

Benchmarking PDFs

LHC data - ttbar production



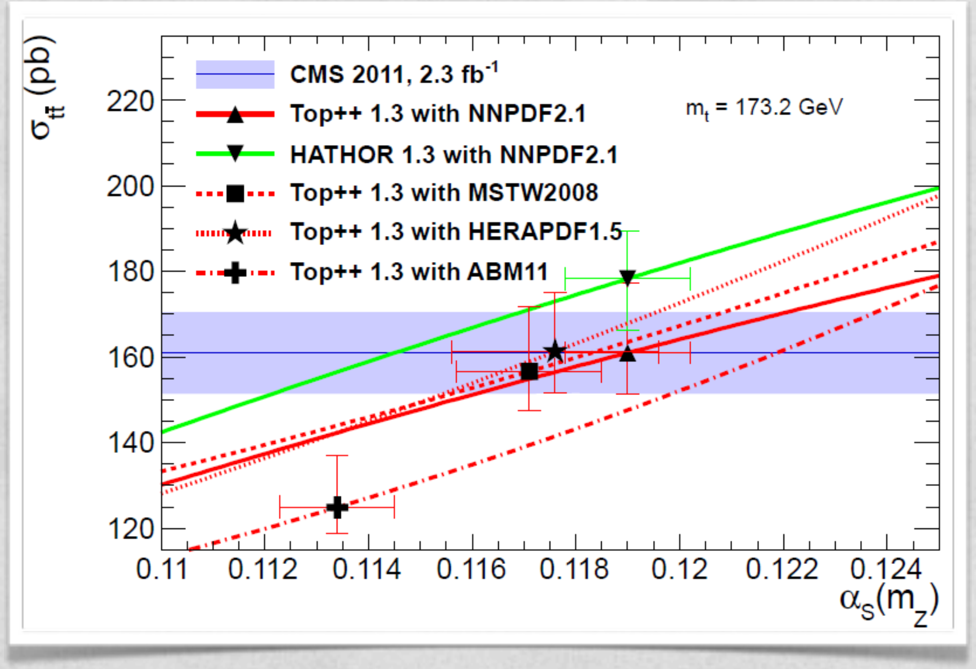


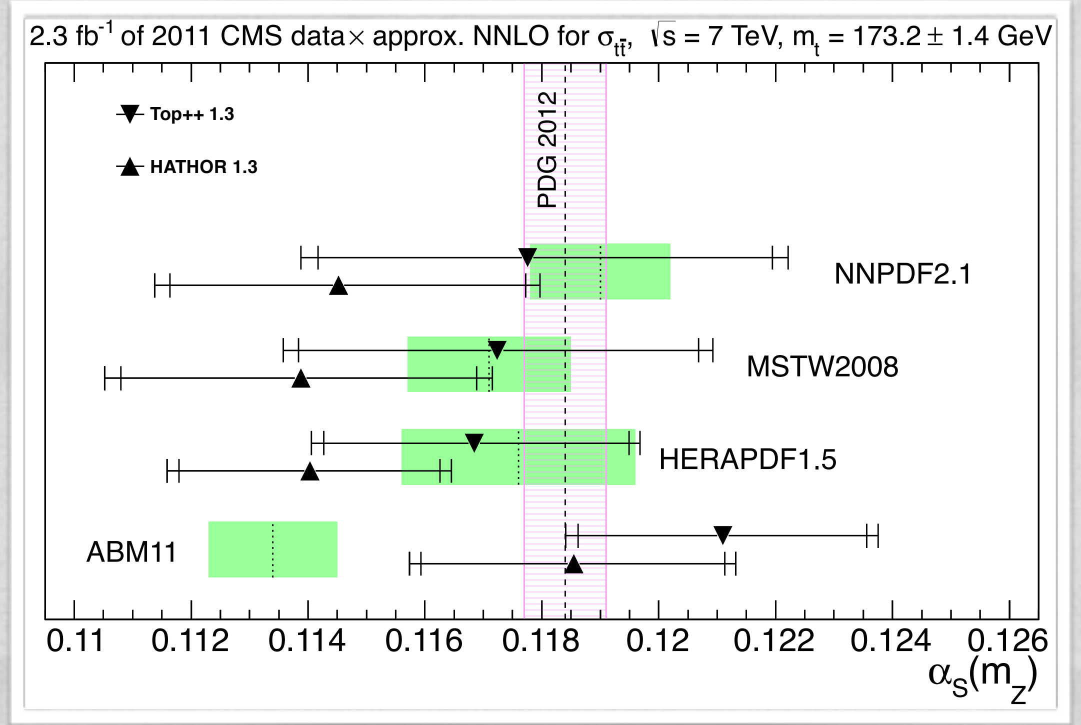
- The ttbar cross-section is **directly sensitive to the gluon** PDF at medium-/large-x
- Dependence of the cross-section on the value of the strong coupling constant suggests that it is possible to **use the ttbar cross-section to extract** α_S

Benchmarking PDFs

LHC data - α_s determination from ttbar production

[CMS PAS TOP-12-022]

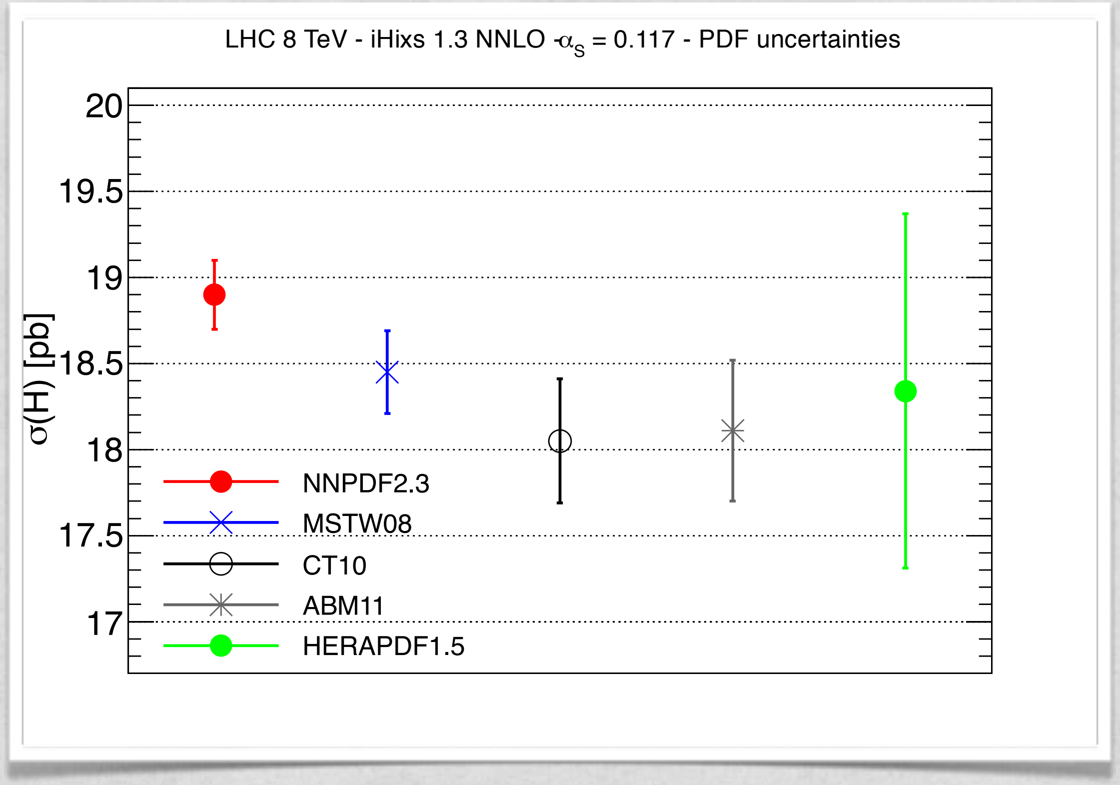


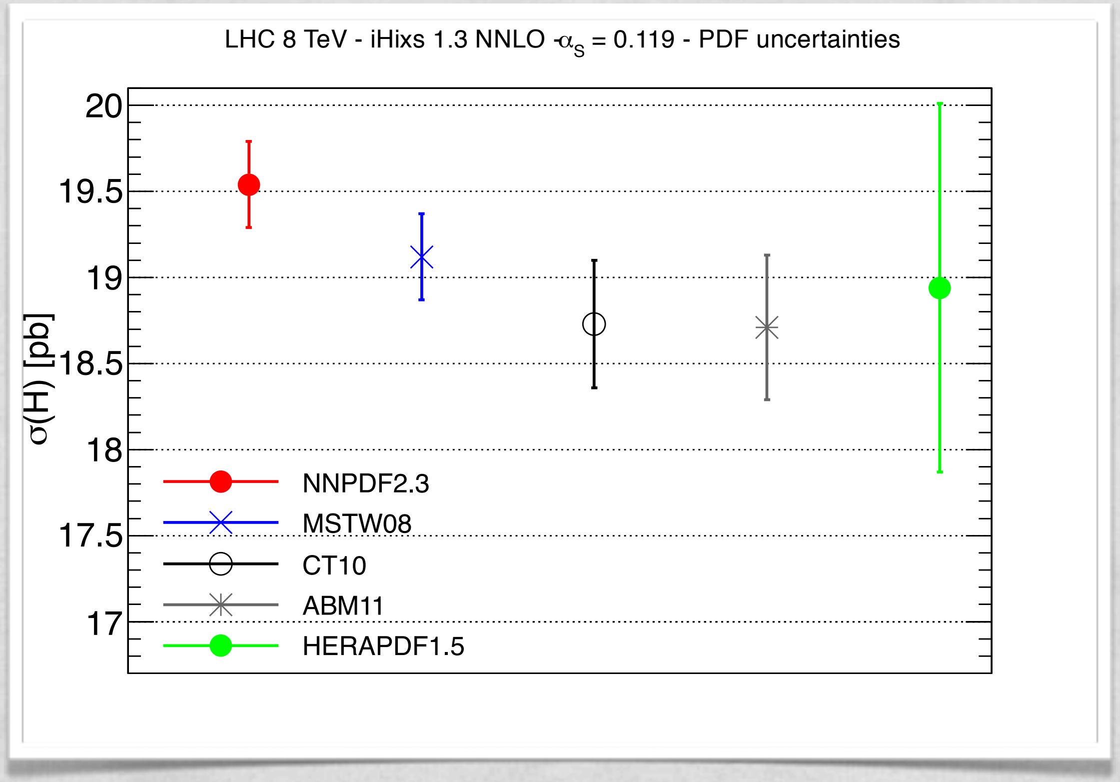


- Once **m_t is fixed**, it is possible to **determine** α_s through comparison of the predicted and measured cross-section
- Results obtained with **HERAPDF, MSTW and NNPDF similar** and **compatible** with preferred values for each set
- **ABM yields** a **large** α_s due to smaller gluon, in contrast with preferred value
- New **high-energy approximation** in **HATHOR** results in 1.3% **lower** value for **extracted** α_s

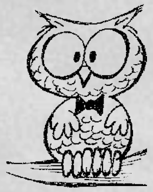
Benchmarking PDFs

Higgs prduction - gluon fusion



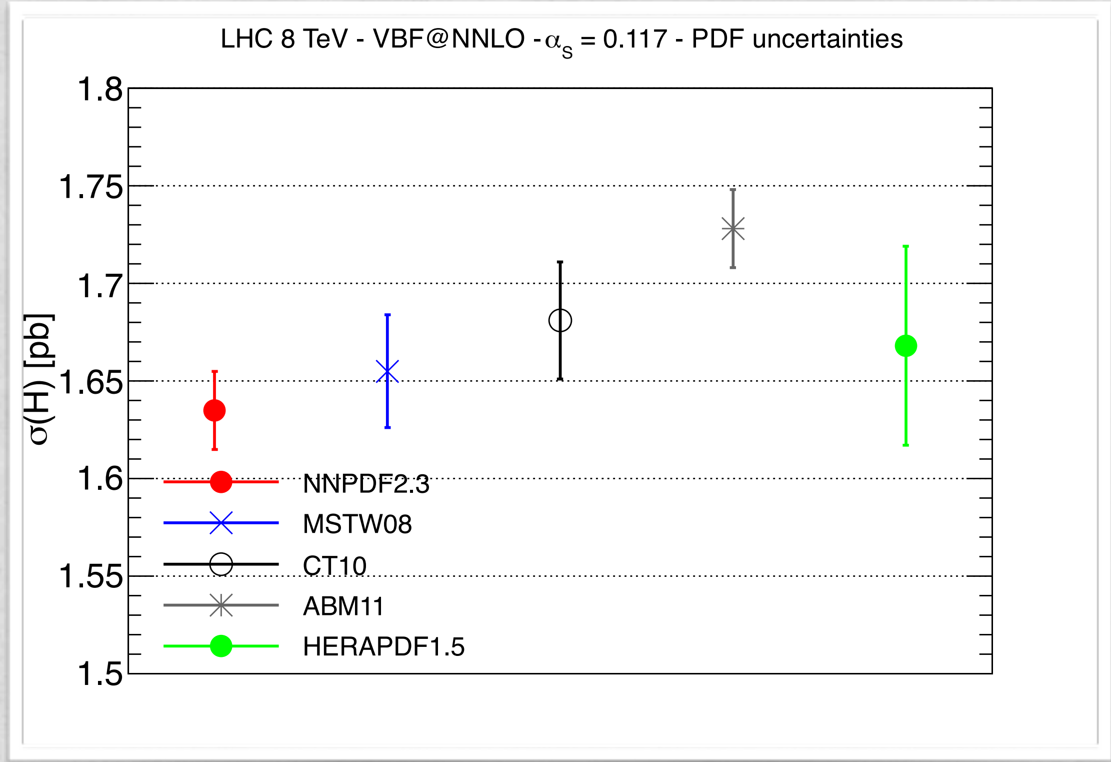


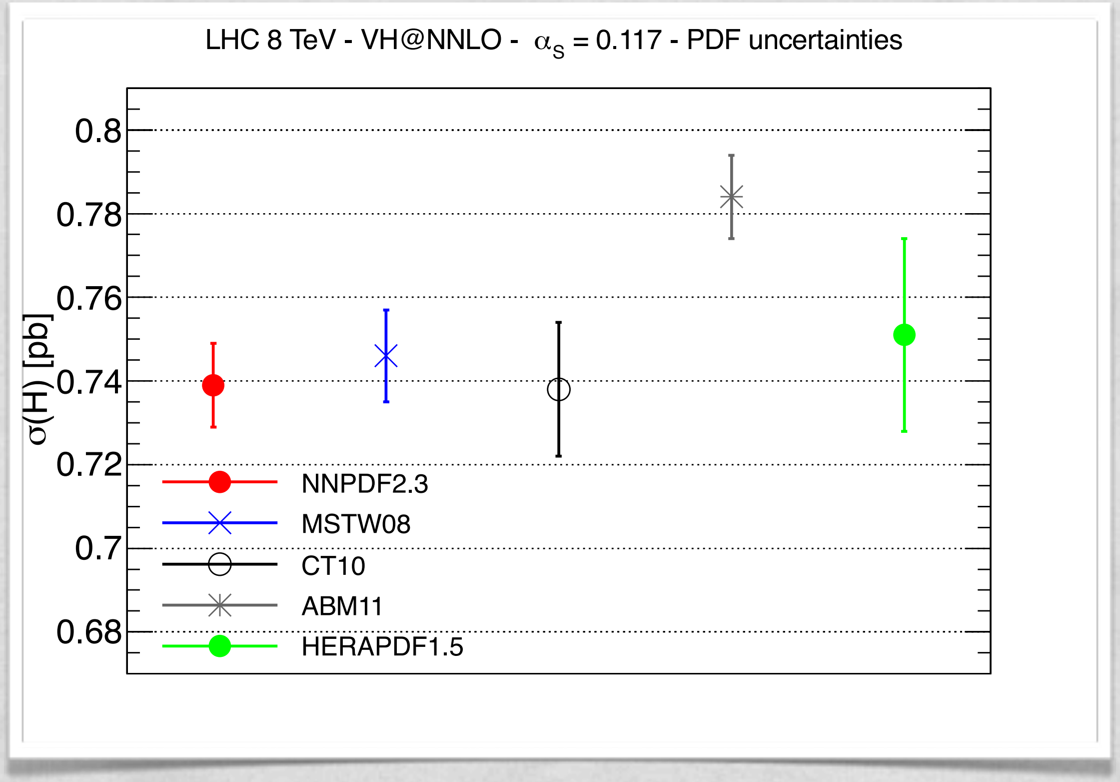
- **Higgs** production cross-section in **gluon-gluon fusion** computed at **NNLO** with iHix (with Q = $\mathbf{M_H}$ **= 125 GeV**)
- **Predictions** from different sets are **compatible** when **common value for** α_s is used (ABM prediction substantially lower if default value for $\alpha_s = 0.1134$ is used)



Benchmarking PDFs

Higgs prduction - VBF & WH





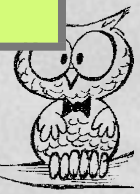
- **Higgs** production cross-section in **Vector Boson fusion** and **associated WH production** computed at **NNLO** (with Q = M_H **= 125 GeV**)
- **Predictions** from global sets and HERAPDF are **compatible** within errors, ABM prediction substantially larger due to differences in quark-quark luminosities

LHC for PDFs

Impact of LHC data

- Compare the quality of the fit to all datasets before and after inclusion of LHC data 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

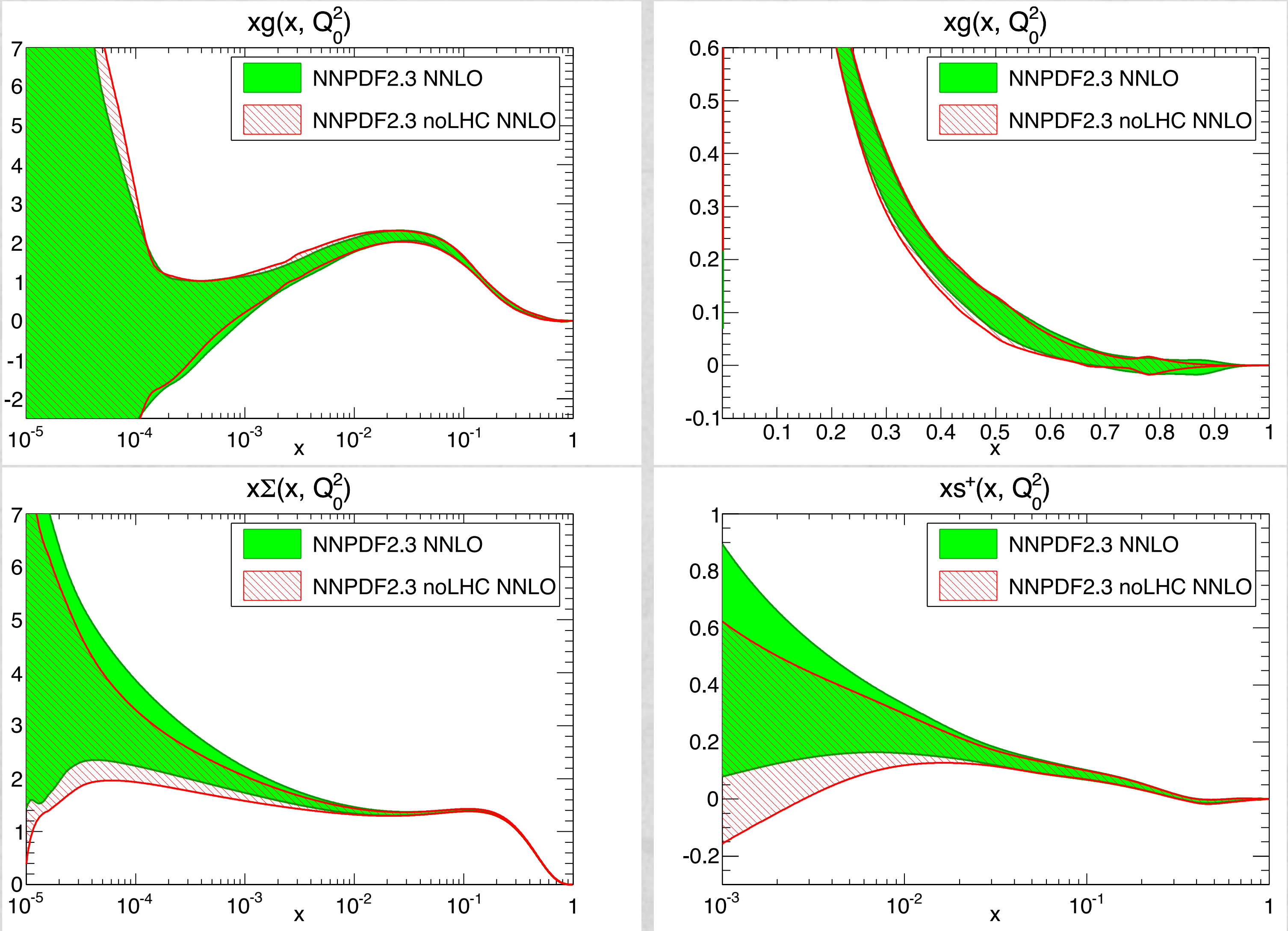
NNLO	NNPDF2.3 noLHC	NNPDF2.3
NMCpd	0.94	0.95
NMC	1.56	1.59
SLAC	1.04	1.00
BCDMS	1.28	1.28
HERA-I	1.03	1.01
CHORUS	1.07	1.07
NuTeV	0.48	0.56
DYE605	1.06	1.08
DYE866	1.55	1.69
CDFWASY	1.67	1.64
CDFZRAP	2.13	2.03
D0ZRAP	0.63	0.61
ATLAS-WZ	1.94	1.43
CMS-WEASY	1.20	0.81
LHCb-W	1.03	0.83
CDFR2KT	0.67	0.68
D0R2CON	0.94	0.94
ATLAS-JETS-2010	0.94	0.94



LHC for PDFs

Impact of LHC data on PDFs

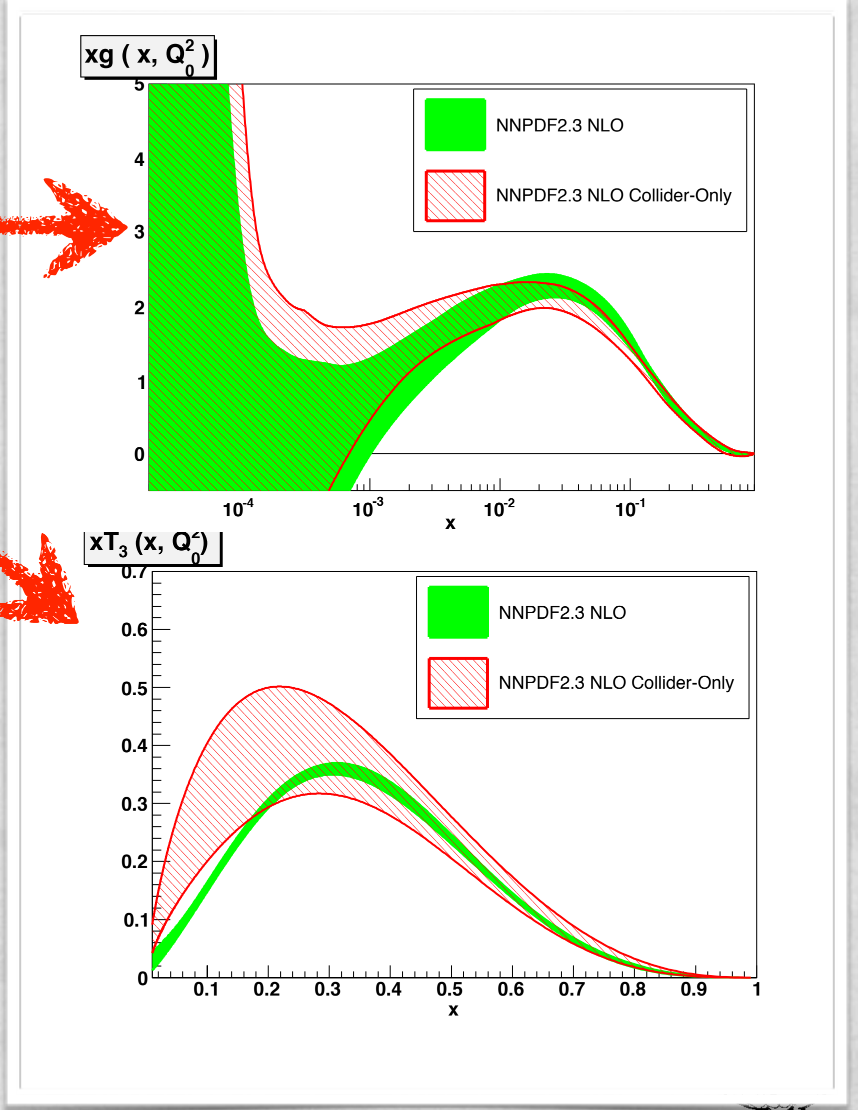
- **Moderate impact of LHC data** on extracted PDFs
- Largest **impact** on **singlet quark** and **strange** distribution
- Effect is **at most half a sigma** shift in central values



Collider-only fit

Are we there yet?

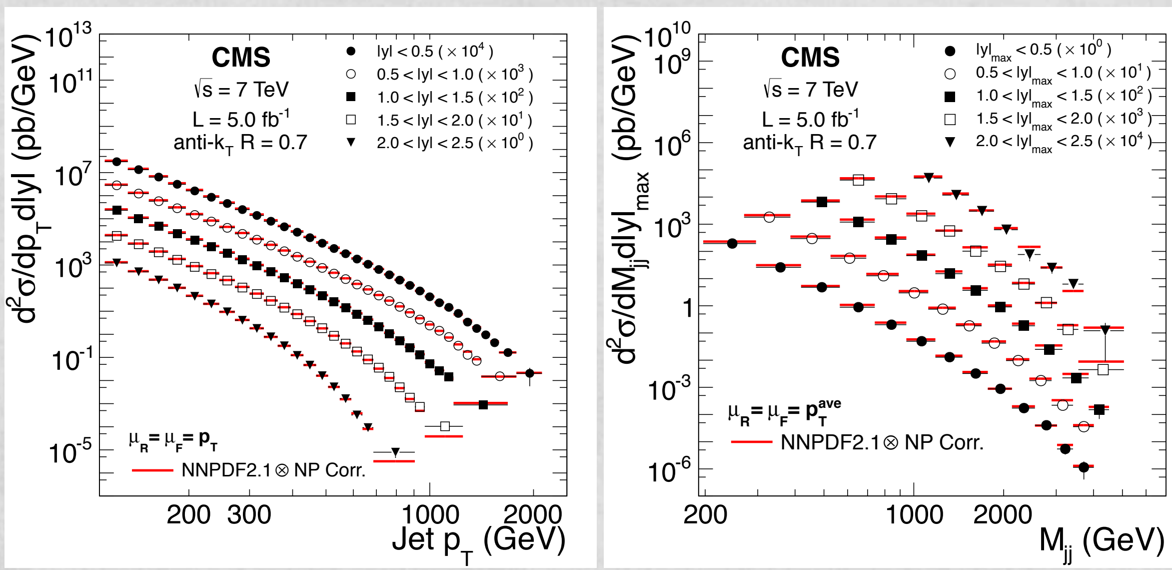
- It is the fit we would love to have
  - Only **high energy data**: minimize the effects of higher-twist contributions
  - Only **proton data**: no assumptions based on models for nuclear corrections
- **Gluon distribution** is very **well constrained** both at small-x (HERA) and large-x (Tevatron/LHC 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**
- **Improvement** with respect to NNPDF2.1 collider fit thanks to **inclusion of LHC data**
- … **things can only get better** with more data coming from HERA (combined F2c and HERA-II data) and the LHC (W+c, low mass DY, photons, high pt Z/W …)



LHC for PDFs

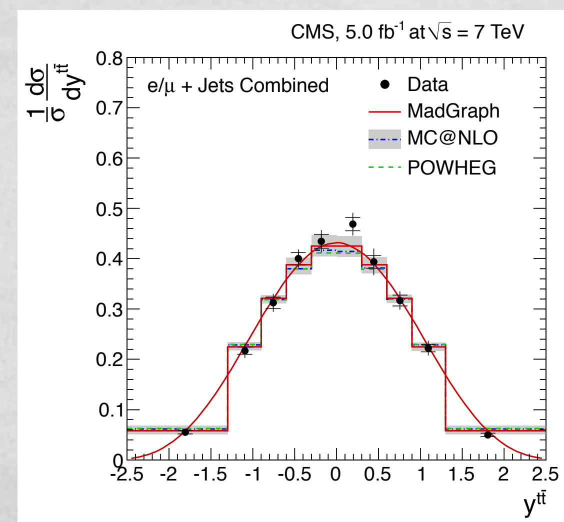
More data are coming!

More inclusive jet (and dijet) data



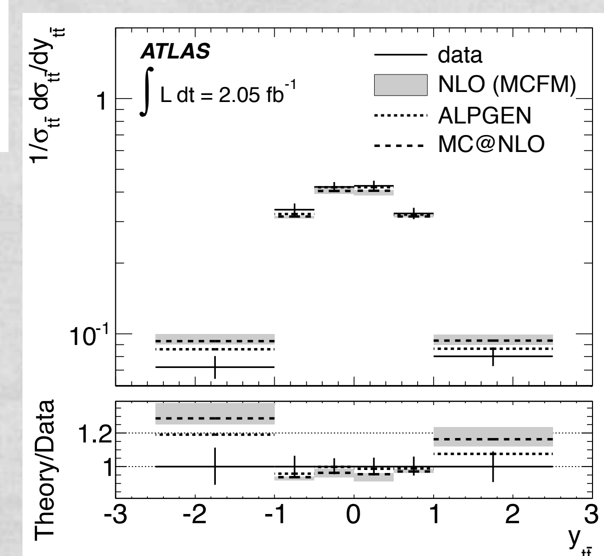
[CMS, arXiv:1212.6660]

top-pair differential distributions

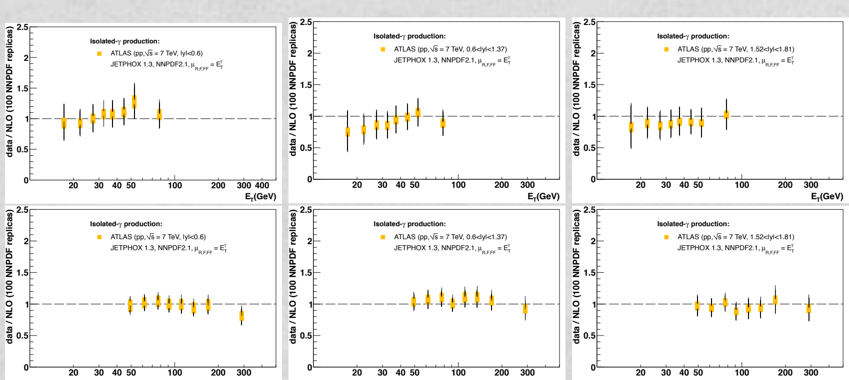


[CMS, arXiv:1211.2220]

[ATLAS, arXiv:1207.5644]



Prompt photons



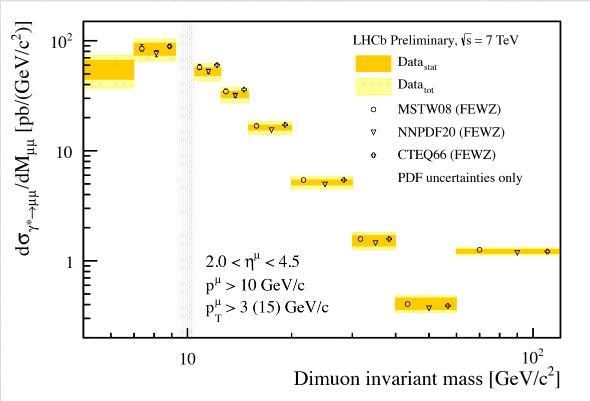
D. d'Enterria & J. Rojo, arXiv:1202.1762

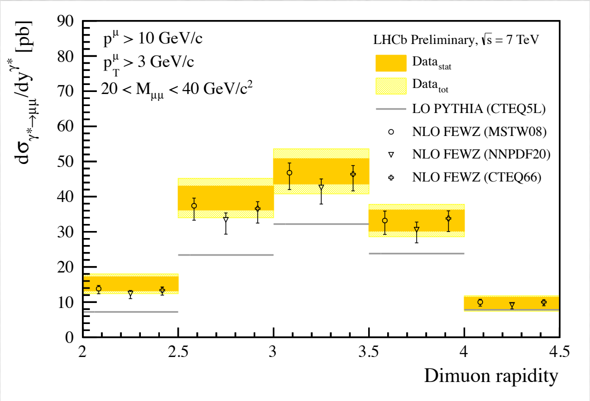
Constraining the **medium/large-x gluon**

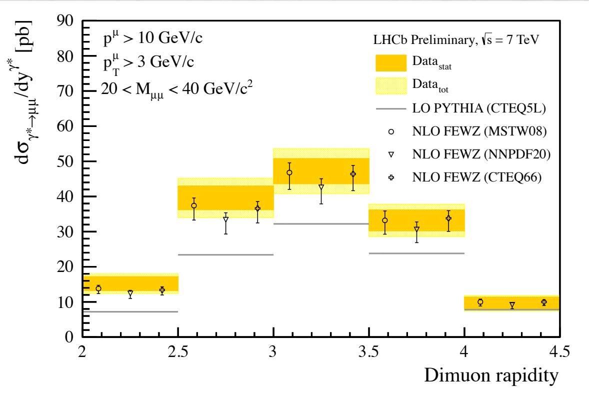
LHC for PDFs

More data are coming!

Low-mass Drell-Yan

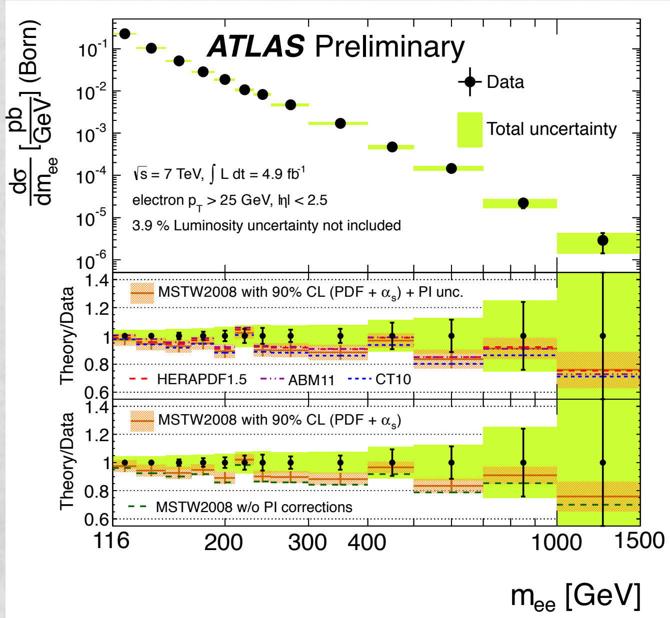


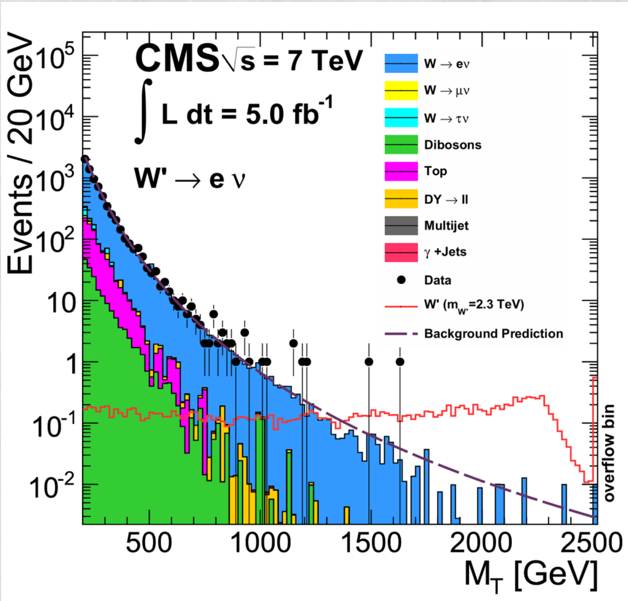




[LHCb-CONF-2012-013]

High mass Drell-Yan/W production

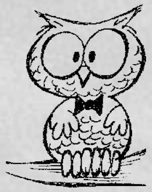




[ATLAS-CONF-2012-159]

(**CAVEAT:** needs to be turned from a search into a measurement)

Constraining **light quark flavour separation**



LHC for PDFs

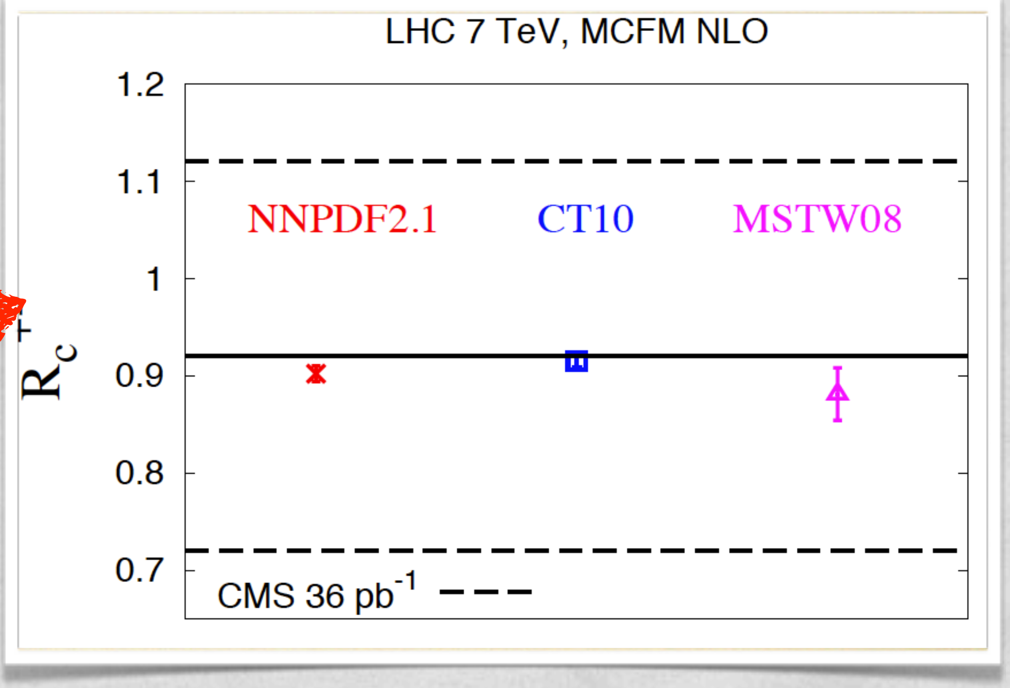
More data are coming!

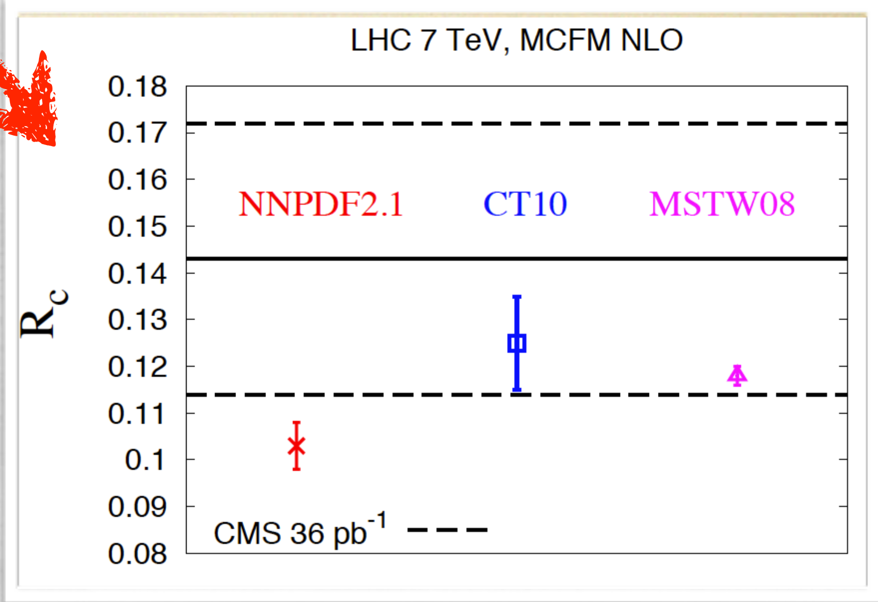
W+charm production

[CMS-PAS-EWK-11-013]

$R_c^{\pm} = \sigma(W^+\bar{c})/\sigma(W^-c)$ $R_c = \sigma(W+c)/\sigma(W+jets)$

$$R_c^{\pm} = 0.92 \pm 0.19\,(stat.) \pm 0.04\,(syst.)$$
$$R_c = 0.143 \pm 0.015\,(stat.) \pm 0.024\,(syst.)$$





Z+bottom

Z+charm

γ+charm

single-top

....

Constraining **strangeness** & **heavy quark PDFs**

Conclusions & Outlook

The way ahead

- **Parton Distribution Functions** are a **fundamental ingredient** of theoretical predictions for observables at **hadron colliders**
- A **systematic benchmarking** of available Parton Distribution Functions against LHC data is a mandatory exercise to **understand** the origin of **differences** in predictions obtained using **different PDF sets**
- **LHC data** will soon allow to **distinguish** among predictions from **different sets**
- **LHC data** are making their way into PDF fits and having a **moderate impact on PDF** determinations
- **More data** are **coming** that will provide **further constraints** and eventually allow for **fits based on collider data** only

Backup Slides

PDFs & the LHC

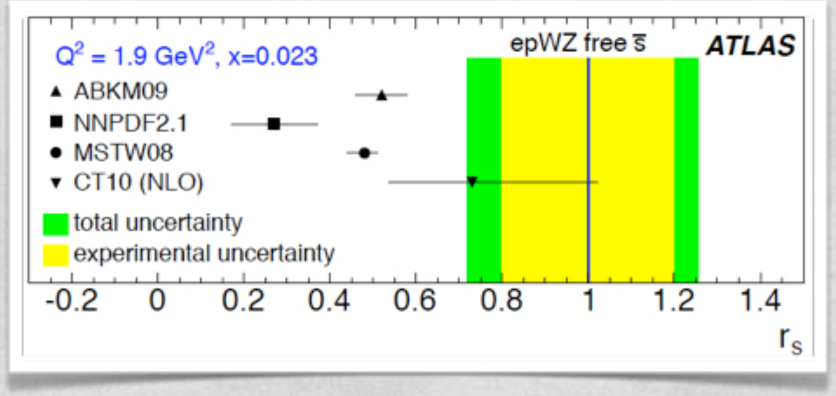
A "strange" story

- ATLAS recently presented evidence for a larger than thought strange distribution at low Q2 and x, leading to a value for

$$r_s(x,Q^2) = \frac{s(x,Q^2)+\bar{s}(x,Q^2)}{2\bar{d}(x,Q^2)} \approx 1 \quad \text{(for } Q^2\text{=1.9 GeV}^2 \text{ and x=0.023)}$$

which systematically larger than prediction from global PDF fits

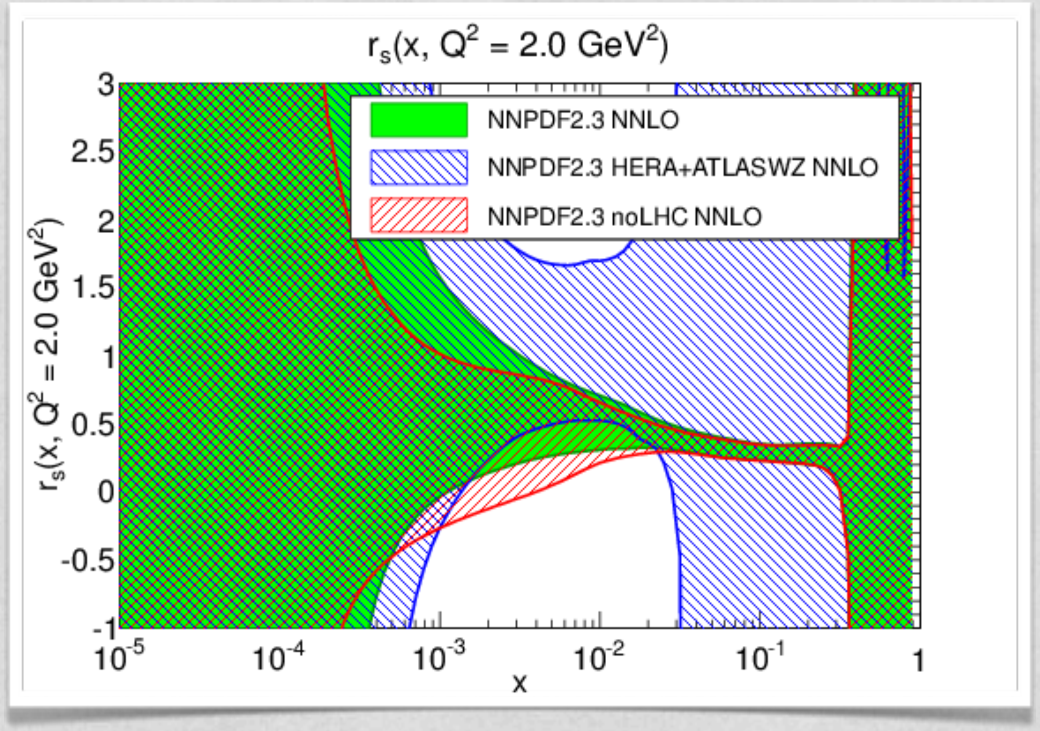
- The ATLAS analysis is based on a fit combining the **HERA-I data** with the **ATLAS W/Z rapidity distribution** data

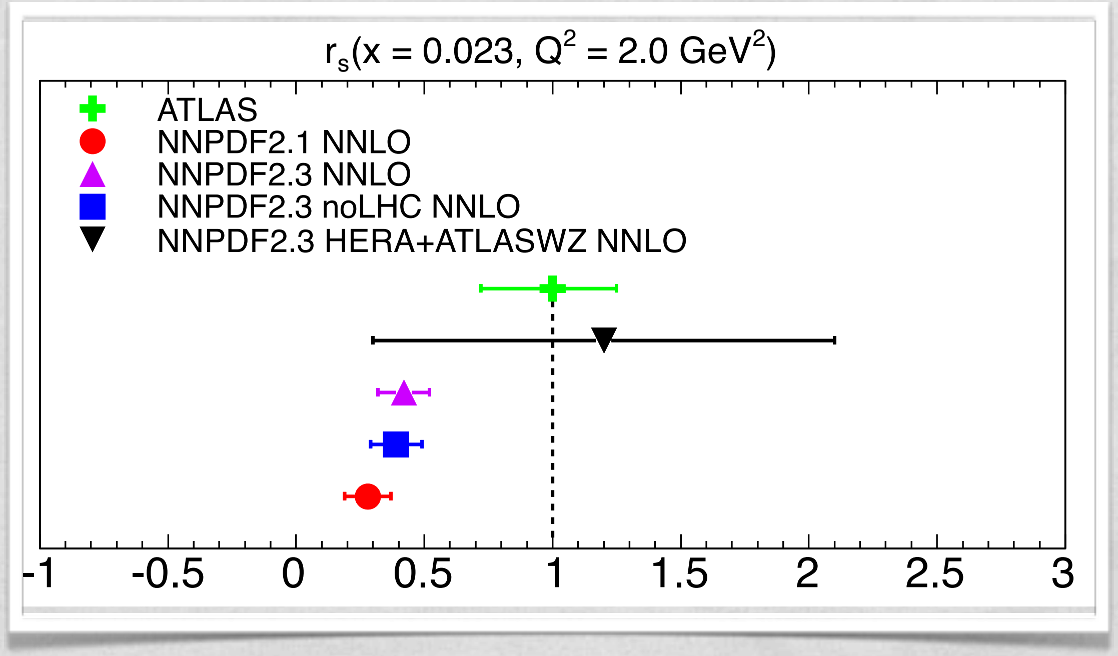


- In order to check the ATLAS claim we produced a version of the NNPDF2.3 fit based on the same dataset used in the ATLAS study

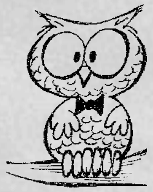
PDFs & the LHC

A "strange" story



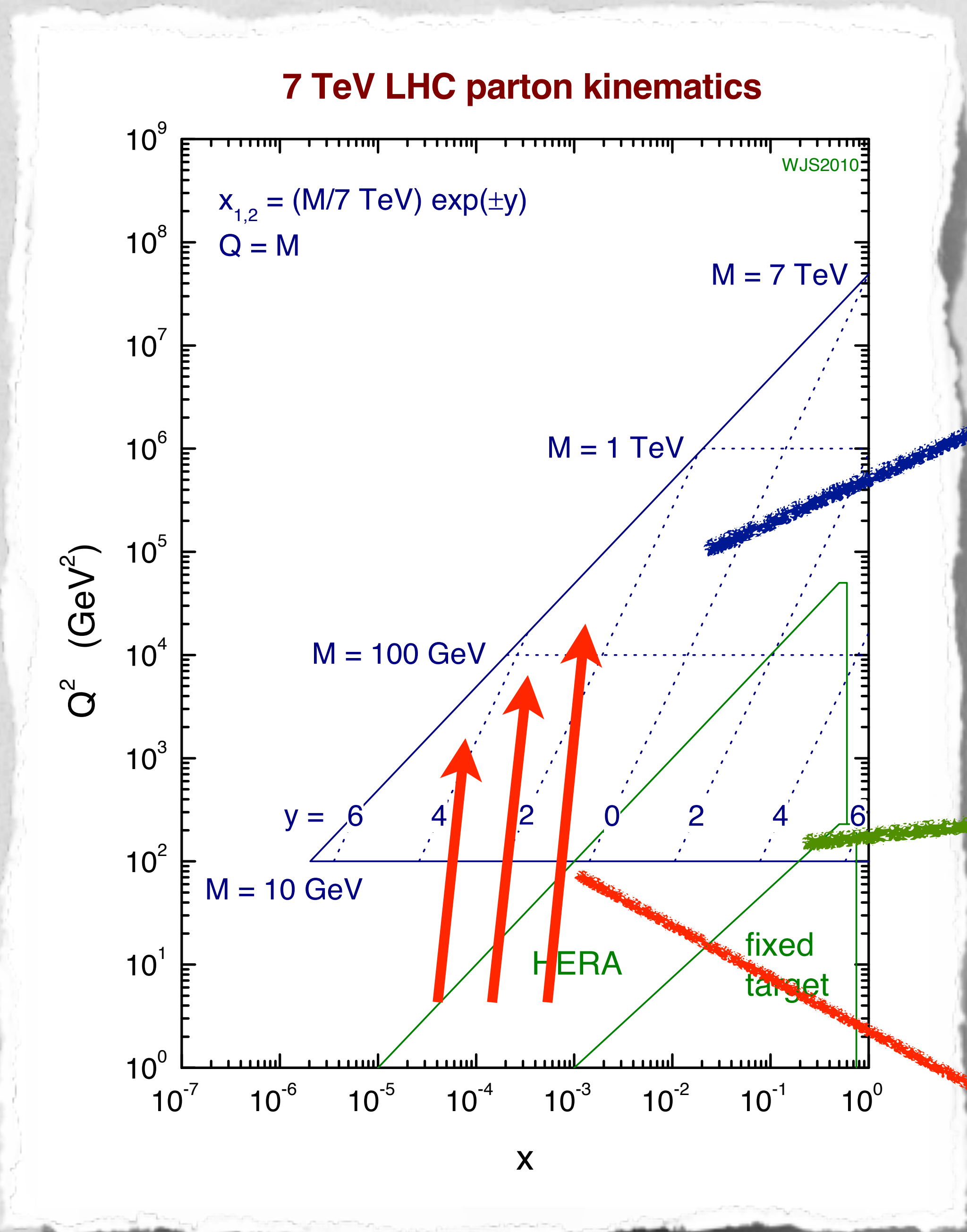


- Strangeness in NNPDF2.3 somewhat larger than NNPDF2.3noLHC in the range $10^{-3}<x<10^{-1}$, though still compatible within errors shows minor small impact of LHC data
- Determination from HERA+ATLAS dataset yields $r_s \sim 1$ but has much larger uncertainties (substantially larger than ATLAS determination)



Parton Distribution Functions

Why do we care?



This is the region where we want to use PDFs for predictions at the LHC

This is the region where we measure PDFs, from DIS and Tevatron experiments

DGLAP evolution

LHC for PDFs

Impact of LHC data

- Compare the quality of the fit to all datasets before and after inclusion of LHC data 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	0.95
NMC	1.59	1.61
SLAC	1.28	1.24
BCDMS	1.20	1.20
HERA-I	1.01	1.00
CHORUS	1.09	1.10
NuTeV	0.42	0.43
DYE605	0.85	0.86
DYE866	1.24	1.27
CDFWASY	1.45	1.57
CDFZRAP	1.77	1.80
D0ZRAP	0.57	0.56
ATLAS-WZ	1.37	1.26
CMS-WEASY	1.32	0.72
LHCb-W	1.03	0.67
CDFR2KT	0.60	0.60
D0R2CON	0.84	0.84
ATLAS-JETS-2010	1.01	1.00

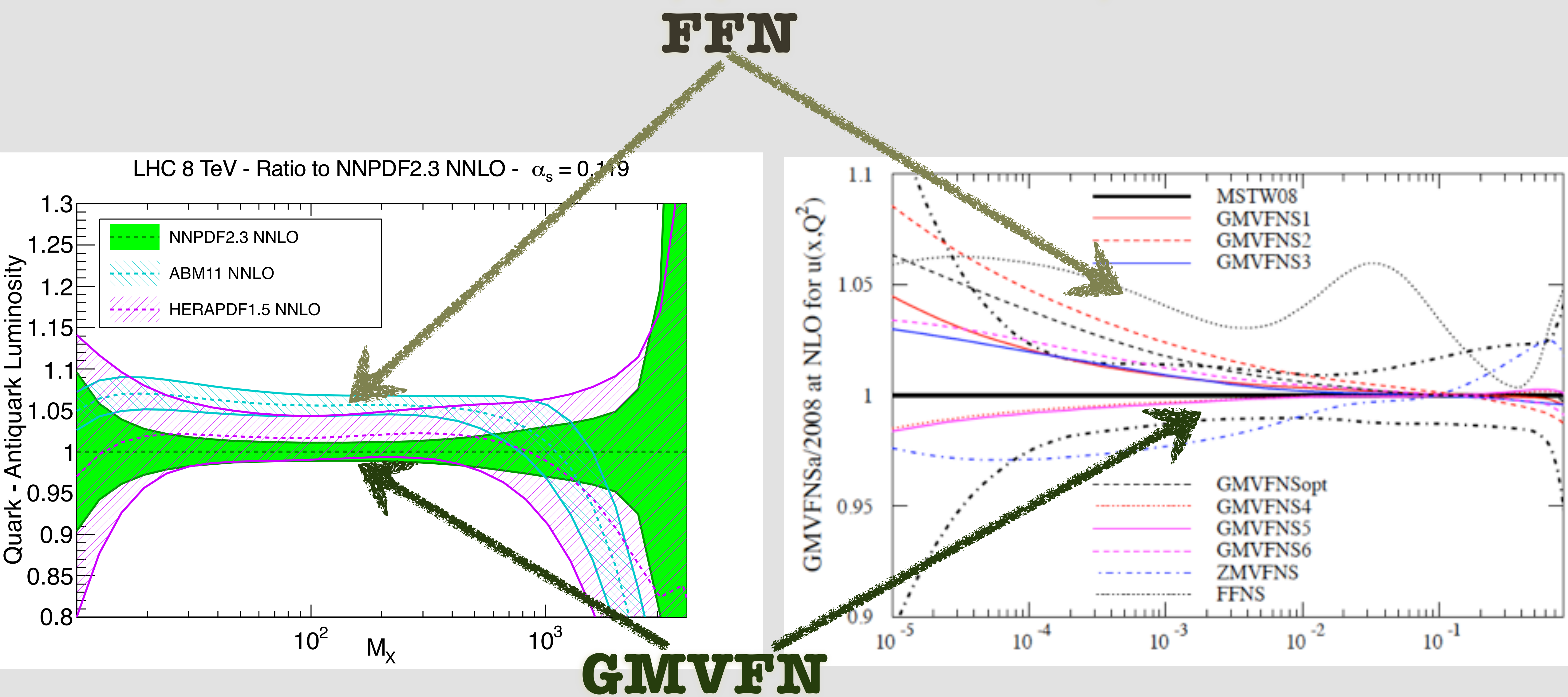
Beyond Benchmarking

(next step:
trying to understand the differences)

NNPDF Preliminary

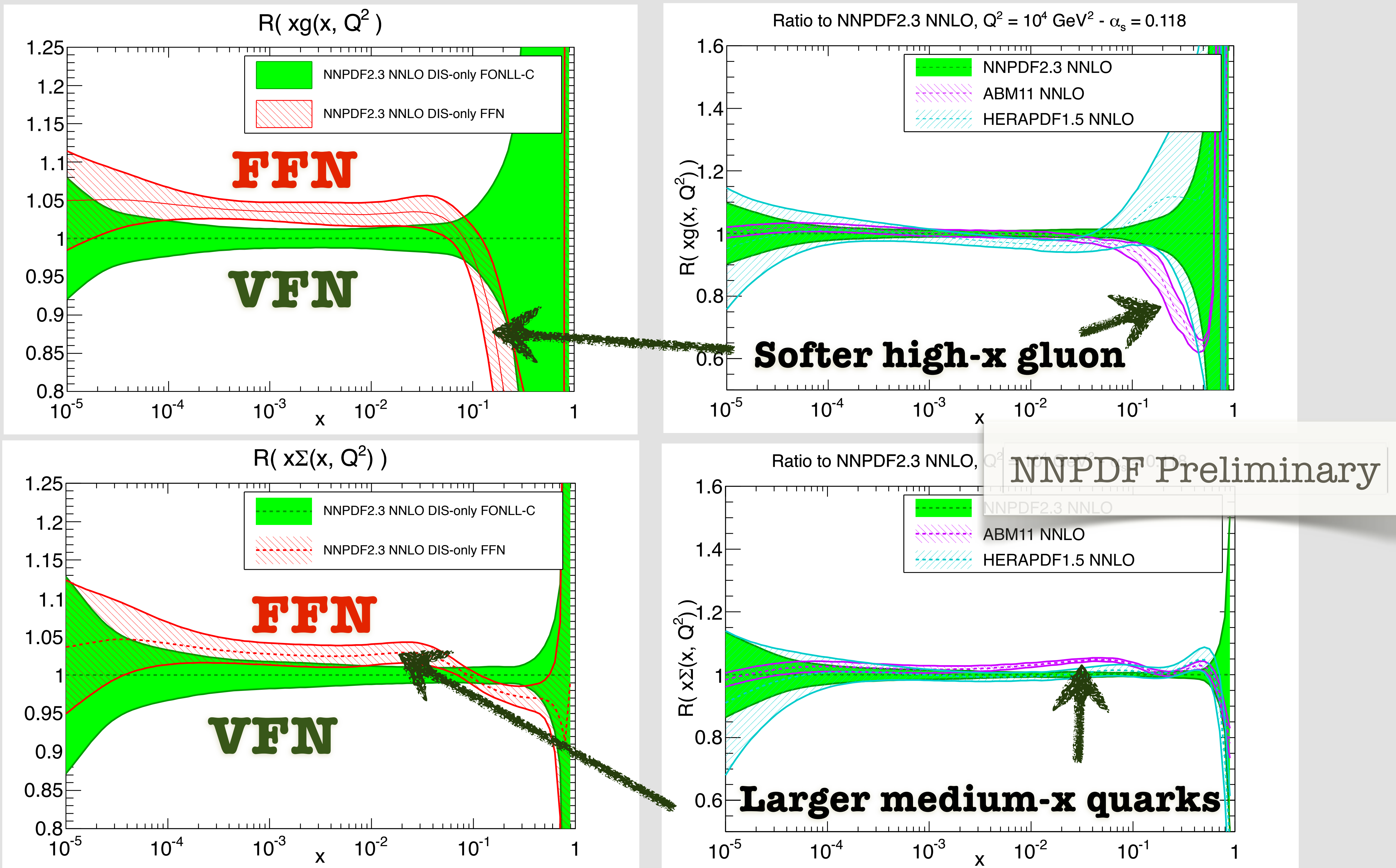
Variable vs Fixed Flavor Number Schemes

- Differences between ABM11 and other PDF sets partly arise from **different HQ treatment**: Fixed Flavor Number vs General Mass VFN (Thorne, arXiv:1201.6180)
- The FFN fit leads to a **harder small-x gluon**, and thus (via the momentum sum rule) a **softer large-x gluon**, and to a **harder quarks at small-x** at LHC scales through evolution
- Differences between various **GM-VFN** schemes much smaller that between **FFN** and **GM-VFN.** See also Les Houches heavy quark benchmark study.



Variable vs Fixed Flavor Number Schemes

- The impact of **FFN vs GM-VFN** also studied by **NNPDF:** consistent results with those of Thorne
- **Similar trend** observed as between NNPDF2.3 and ABM11: **softer large-x gluon, harder medium-x quarks**



Variable vs Fixed Flavor Number Schemes

- Are all heavy quark schemes equally valid? Or some of them **describe better** exp data?
- Compute the difference in χ^2 between the VFN and FFN fits with **various kinematical cuts**
- The **FFN fit quality is poorer than the VFN**, the difference is **statistically significan**t and specially relevant for the **inclusive HERA-I**: due to **missing resummation of DGLAP logarithms**

$x_{\rm min}$	$x_{\rm max}$	$Q^2_{\rm min}$	$Q^2_{\rm max}$	$\chi^2_{\rm tot}({\rm FFN}-{\rm VFN})$	$N_{\rm dat}^{\rm tot}$	$\chi^2_{\rm hera}({\rm FFN}-{\rm VFN})$	$N_{\rm dat}^{\rm hera}$
10^{-6}	1.0	3.0	10^6	28.26	2936	37.88	592
10^{-6}	1.0	3.0	10^6	68.88	1055	39.73	405
10^{-6}	1.0	3.0	10^6	28.54	422	10.65	202
10^{-6}	1.0	10^2	10^6	38.80	620	46.67	412
10^{-6}	0.1	10	10^6	49.67	583	32.43	350
10^{-6}	0.1	10^2	10^6	45.92	321	47.26	227
10^{-6}	0.1	10	10^3	31.17	510	13.52	298
10^{-6}	0.1	10^2	10^3	27.21	248	28.11	175

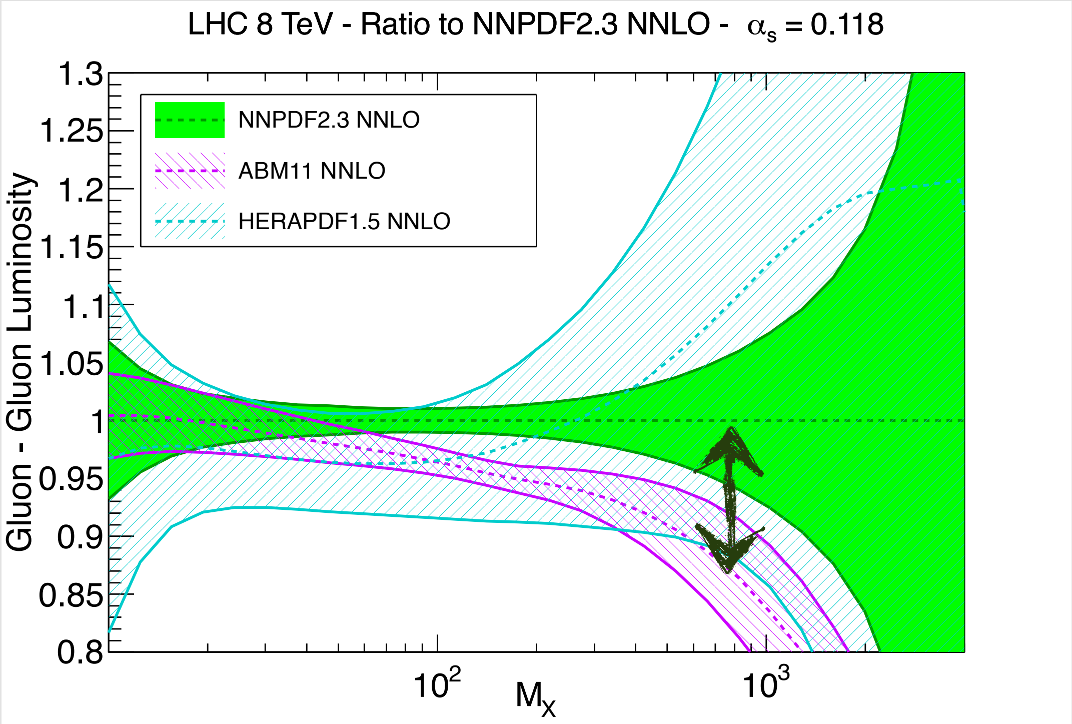
kin cuts **all DIS data** **HERA-I data**

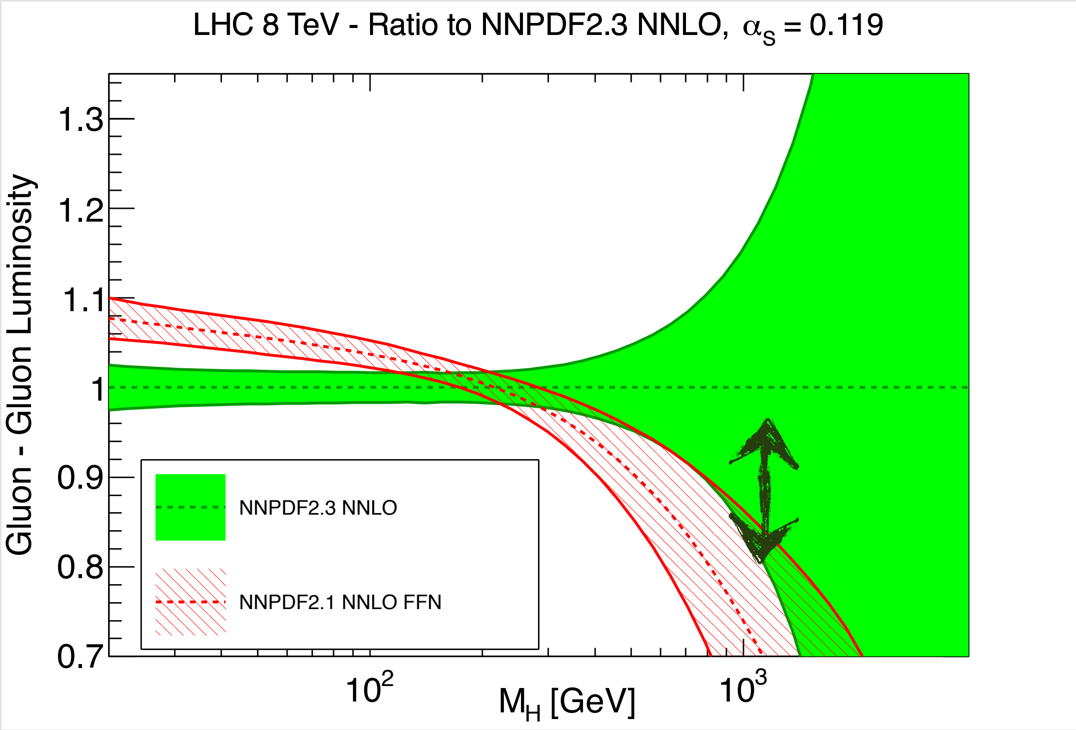
- Although FFN provides a reasonable description of HERA-I data, a **better fit quality** obtained by VFN thanks to **DGLAP resummation at moderate and large Q2**
- FFN and VFN similar χ^2 at **small Q2**

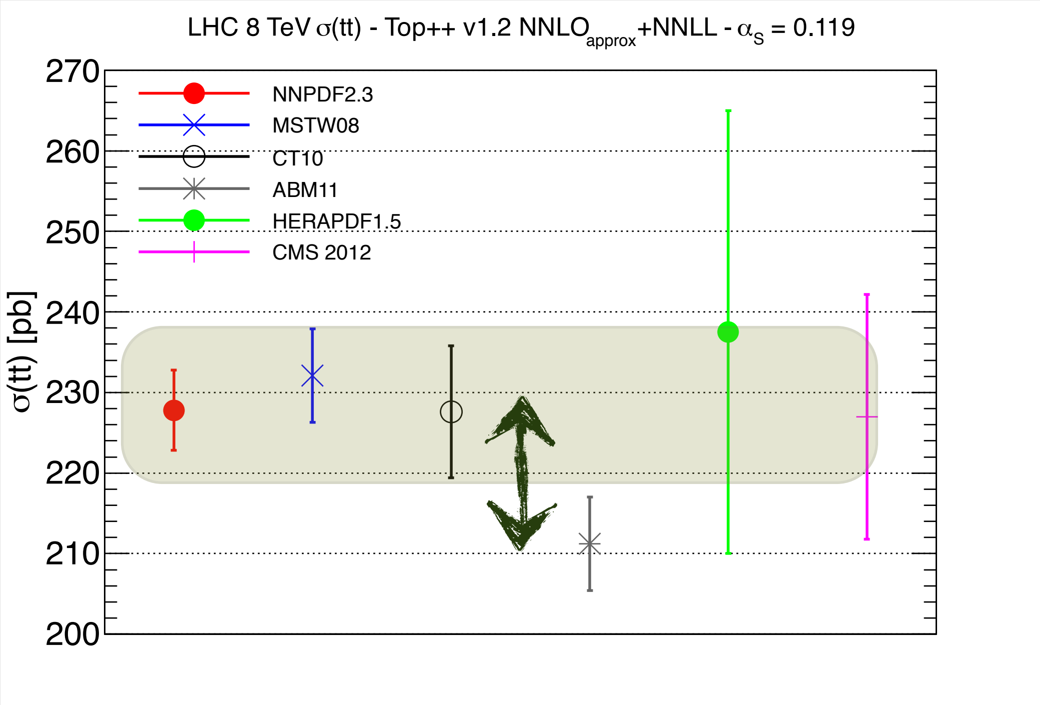
NNPDF Preliminary

Variable vs Fixed Flavor Number Schemes

- The different treatment of **DGLAP logarithms** (fixed order vs resummation) might explain most differences in LHC cross sections between ABM11 and NNPDF2.3 (even for common α_S)



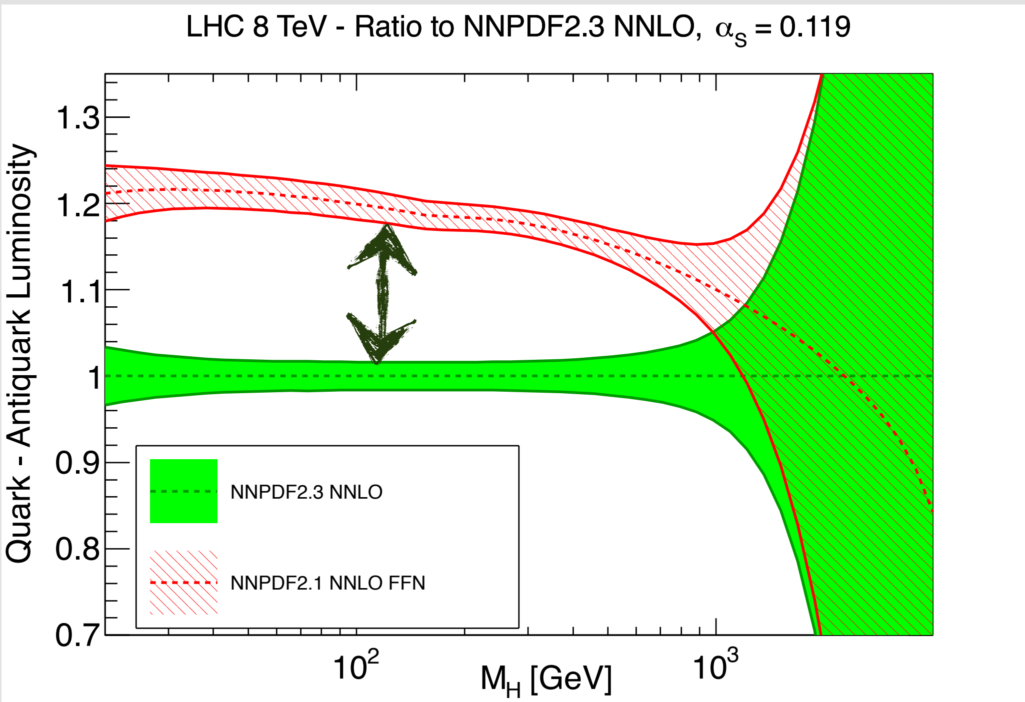


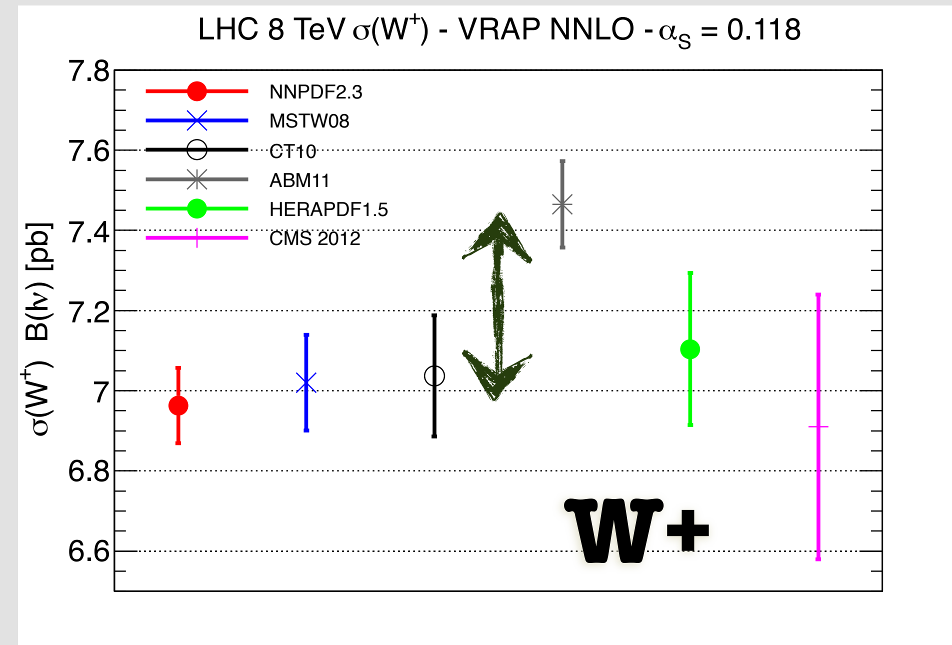


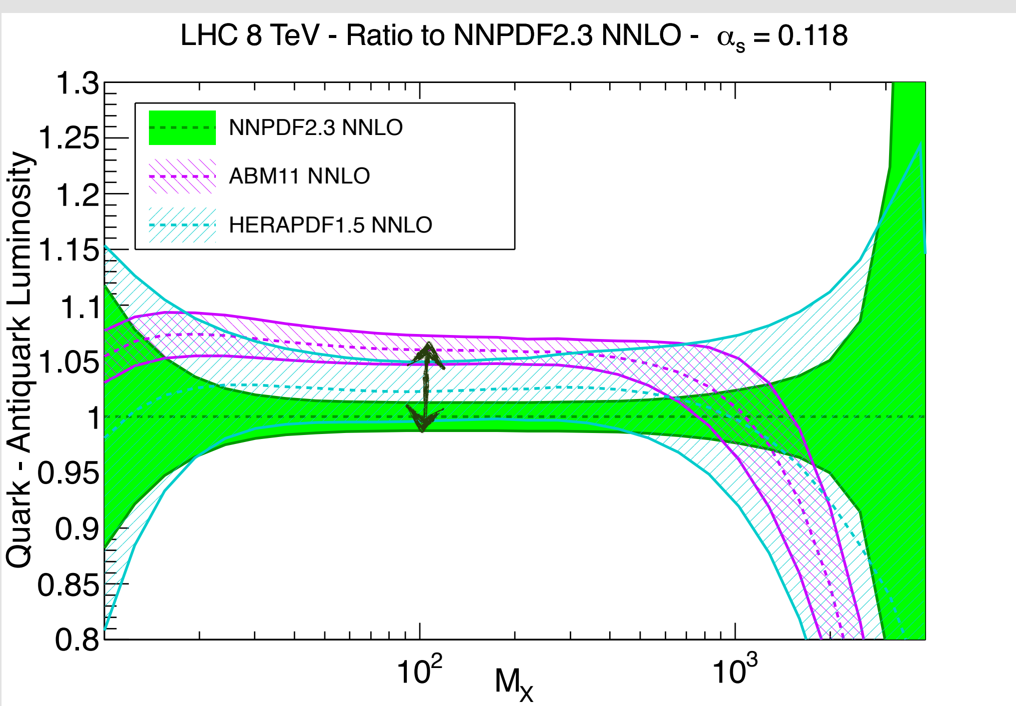
Softer **large-x gluon** -> Smaller **tT, ttH** cross sections

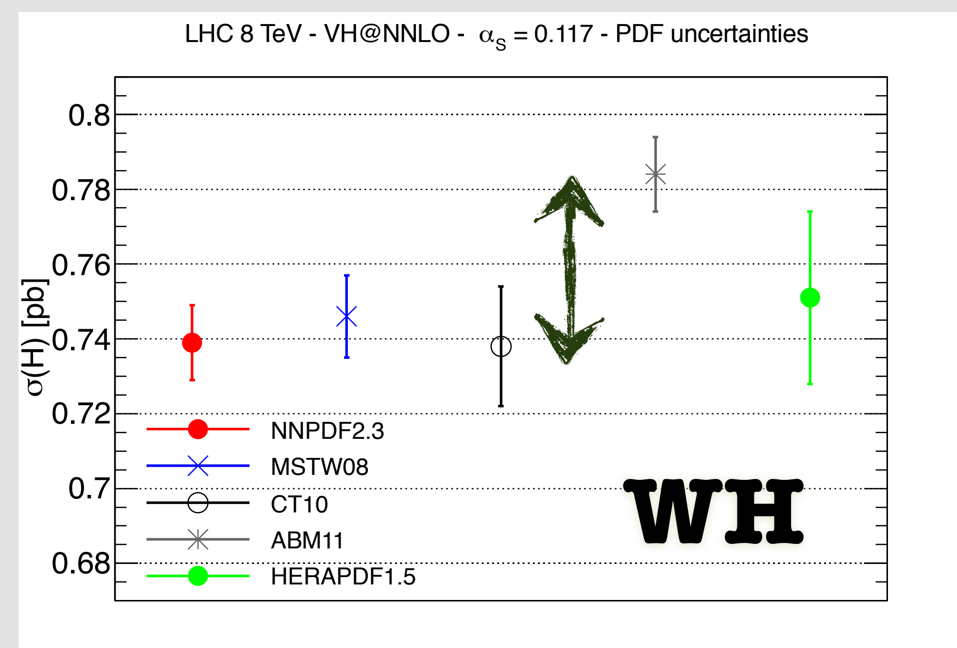
Variable vs Fixed Flavor Number Schemes

- The different treatment of **DGLAP logarithms** (fixed order vs resummation) might explain most differences in LHC cross sections between ABM11 and NNPDF2.3 (even for common α_S)









Harder **smaller-x quarks** -> Larger **W,Z, WH, VBF** cross sections

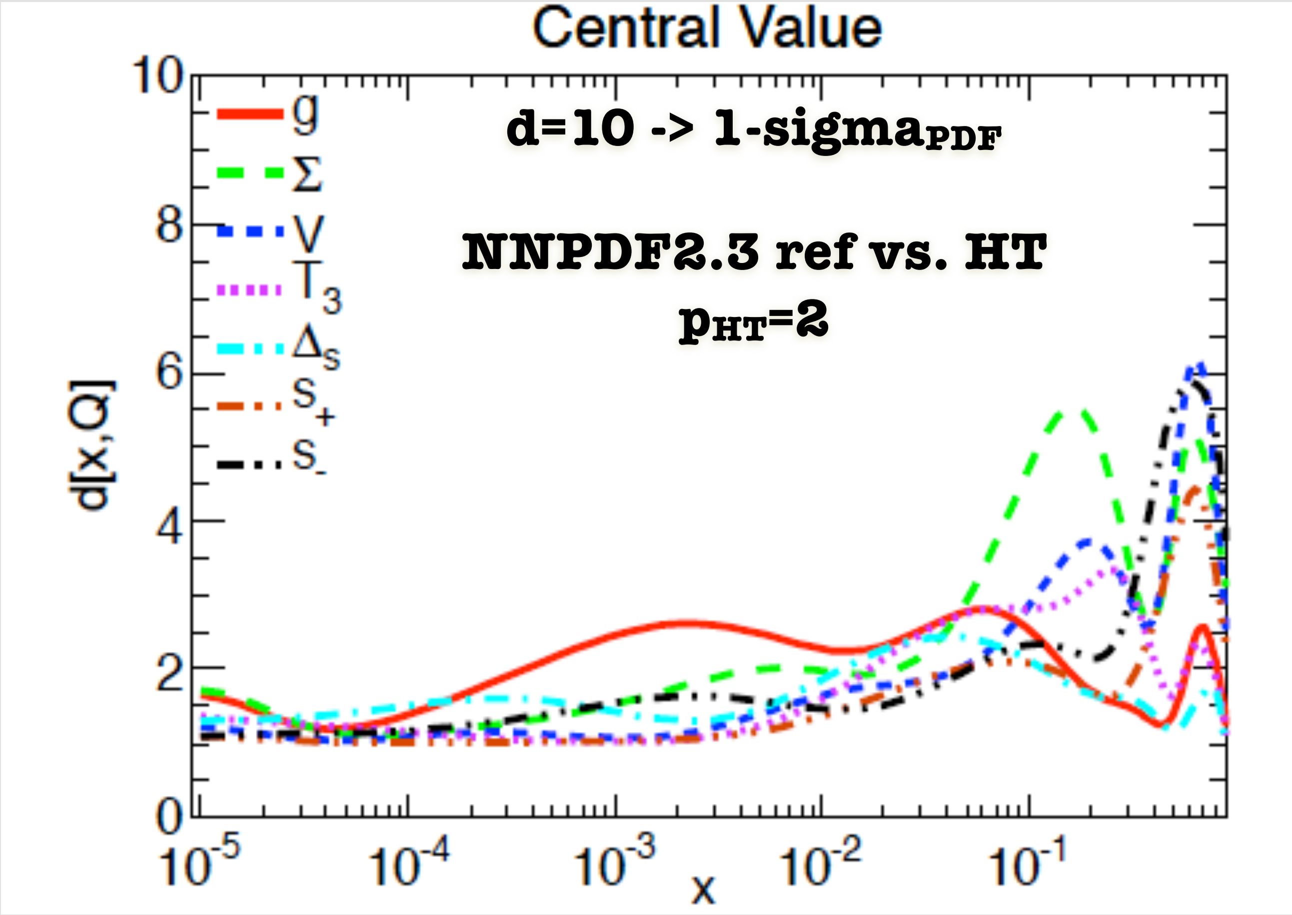
Impact of Higher Twists

- **NNPDF2.3** adopts a kin cut of **$W^2 > 12.5$ GeV^2** and includes exactly **kinematical higher twists** (target mass corrections)

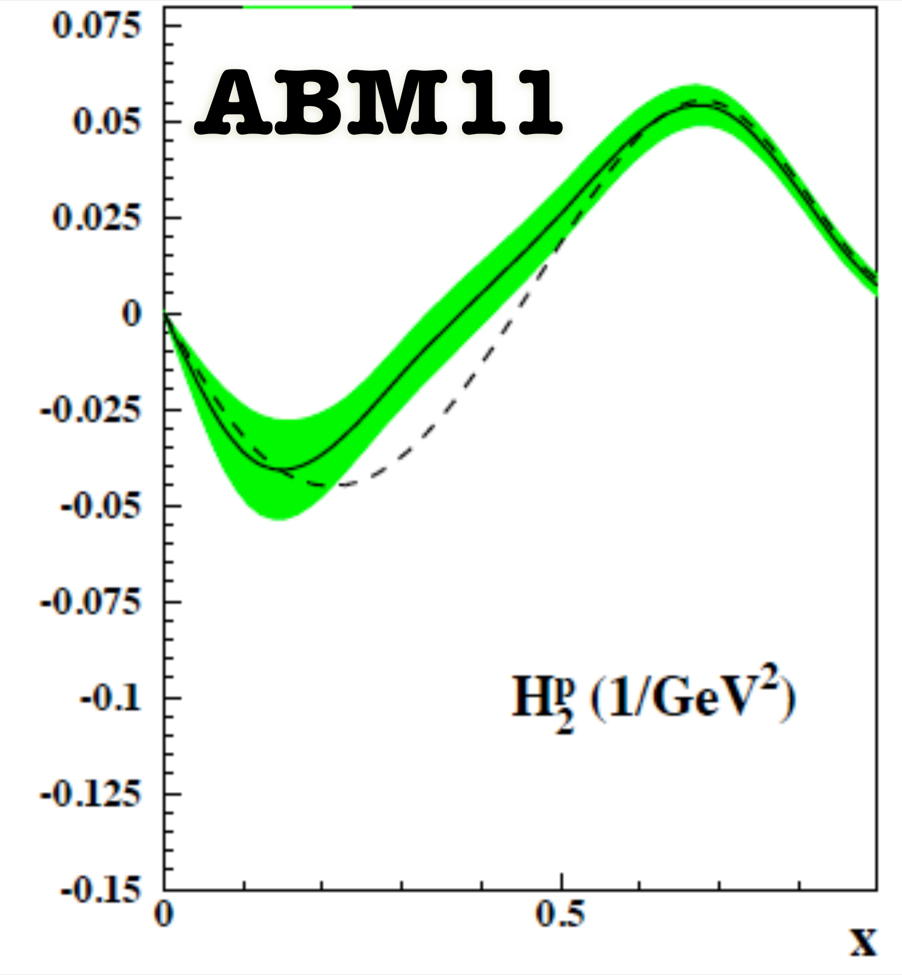
- To explore possible impact of residual dynamical higher twists, redo NNPDF2.3 using the **ABM HT parametrization** varying the overall normalization

$$F_i^{\rm HT}(x,Q^2) = F_i^{\rm TMC}(x,Q^2) + p_{\rm HT}\frac{H_i^{\tau=4}}{Q^2} \ ,$$

- Even for a HT correction **twice** the size of ABM11, differences in PDFs **much smaller than PDF uncertainties.** Similar conclusions from MSTW study (arXiv:1106.5789)

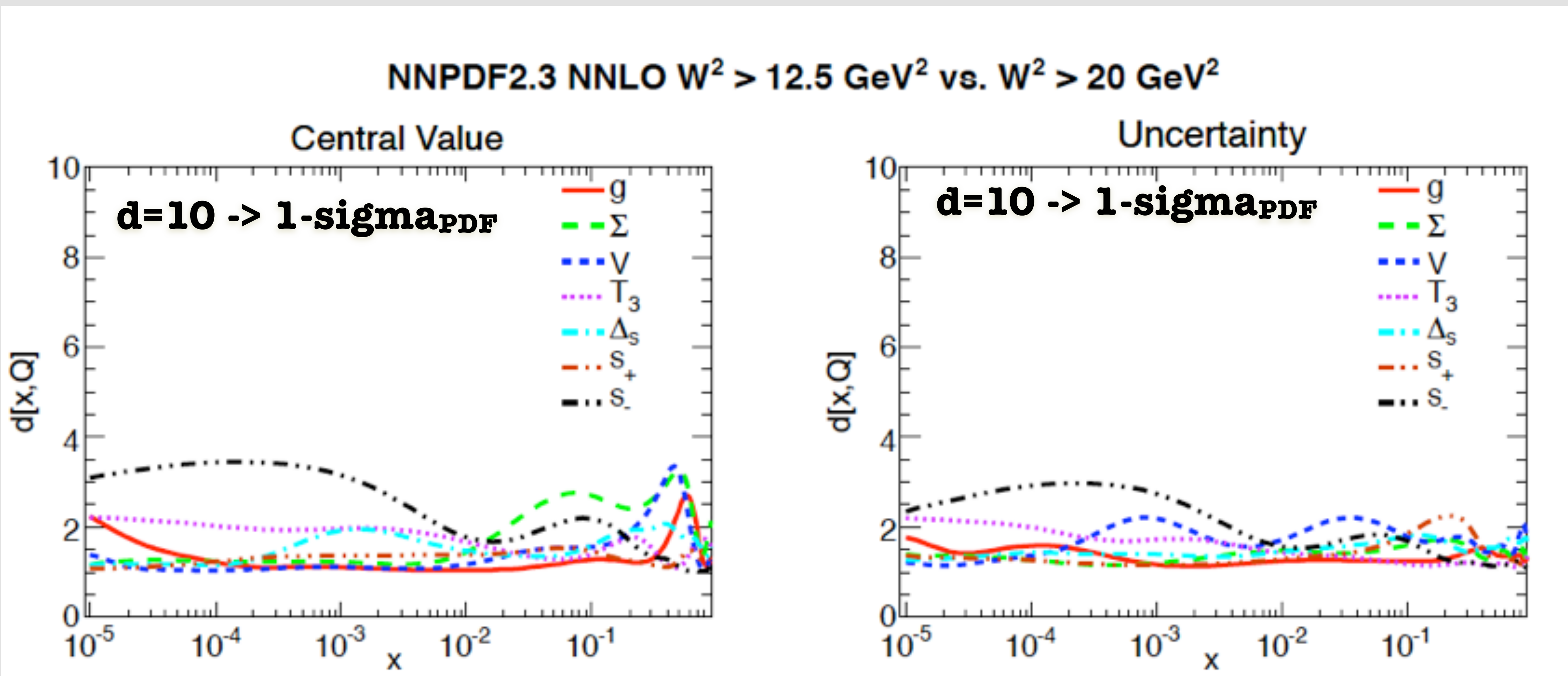


NNPDF Preliminary



Impact of Higher Twists

- **NNPDF2.3** adopts a kin cut of **W^2 > 12.5 GeV^2** and includes exactly **kinematical higher twists** (target mass corrections)
- If **higher twists contaminate** the default NNPDF2.3, one should expect a systematic difference when the cut is raised say to **W^2 > 20 GeV^2**
- We find however no statistically significant differences varying the **W^2** cut: **higher twists are irrelevant** for the NNPDF fits



NNPDF Preliminary