Parton Distribution Functions and LHC physics

A M Cooper-Sarkar
Corfu 5/9/14

Why the differences between PDFs?

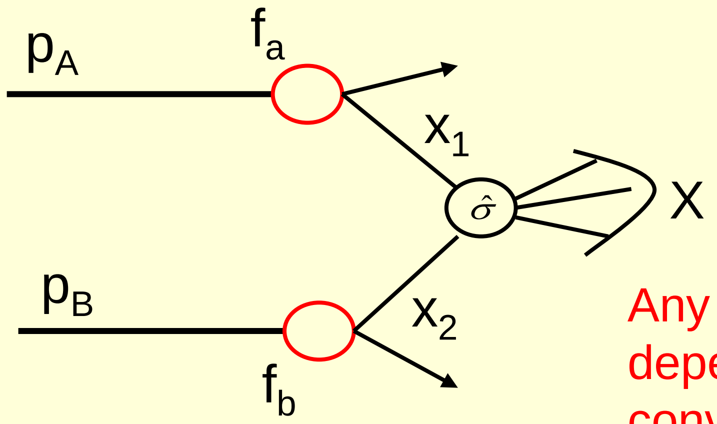
And what consequence do they have?

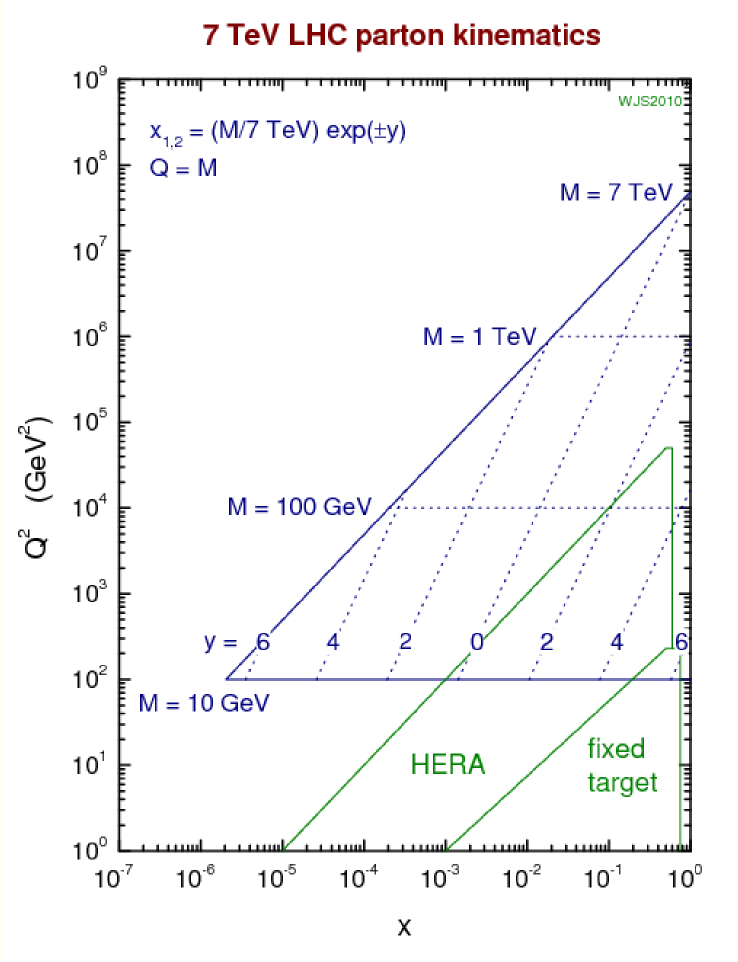
Confront PDFs with LHC data

Uncertainties on Parton Distribution Functions (PDFs) limit our knowledge of cross sections whether SM or BSM.

$$\sigma_X = \sum_{a,b} \int_0^1 dx_1 dx_2\, f_a(x_1, \mu_F^2)\, f_b(x_2, \mu_F^2) \times \hat{\sigma}_{ab\to X}\left(x_1, x_2, \{p_i^\mu\}; \alpha_S(\mu_R^2), \alpha(\mu_R^2), \frac{Q^2}{\mu_R^2}, \frac{Q^2}{\mu_F^2}\right)$$

where X=W, Z, D-Y, H, high-E_T jets, prompt-γ and σ is known to some fixed order in pQCD and EW or in some leading logarithm approximation (LL, NLL, …) to all orders via re-summation





Any claim for new physics at the highest masses is dependent on the PDF chosen to describe conventional physics.

The extent to which the Higgs that we are seeing agrees with the SM Higgs cross section predictions depends on the PDF used to make the prediction

We can use SM measurements to discriminate and improve current PDFs

Why might there be differences between PDF sets?

1. Different Data sets used for determination
2. Heavy quark scheme used and values of heavy quark mass chosen
3. The value of $\alpha_S(M_Z)$ used-- and there is a correlation between the value of alphas chosen/fitted and the gluon shape such that a larger value of $\alpha_S(M_Z)$ goes with a harder high-x gluon. Note many groups now provide PDFs for a series of $\alpha_S(M_Z)$ values.
4. Method of error estimation
5. Model /parmetrisation uncertainties considered

	DATASET	PERT. ORDER	HQ TREATMENT	α_s	PARAM.	UNCERT.
ABM11	DIS Drell-Yan	NLO NNLO	FFN (BMSN)	0.1135 ±0.0014	6 Indep. PDFs Polynomial (25 param.)	Hessian ($\Delta\chi^2$=1)
CT10	Global	LO NLO NNLO	GM-VFNS (S-ACOT)	0.118	6 Indep. PDFs Polynomial (26 param.)	Hessian ($\Delta\chi^2$=100)
JR09	DIS Drell-Yan	LO NLO NNLO	FFN VFN	0.1124 ±0.002	5 Indep. PDFs Polynomial (20 param)	Hessian $\Delta\chi2\sim20$
HERAPDF1.5	DIS (HERA)	LO NLO NNLO	GM-VFNS (TR)	0.1176	5 Indep. PDFs Polynomial (14 param.)	Hessian ($\Delta\chi^2$=1)
MSTW08	Global	LO NLO NNLO	GM-VFNS (TR)	0.1171 ±0.0012	7 Indep. PDFs Polynomial (20 param.)	Hessian ($\Delta\chi^2$~25)
NNPDF2.1/2.3	Global	LO NLO NNLO	GM-VFNS (FONLL)	0.119	7 Indep. PDFs Neural Nets (259 param.)	Monte Carlo

Higher twist terms are included

Plus model/param uncertainties

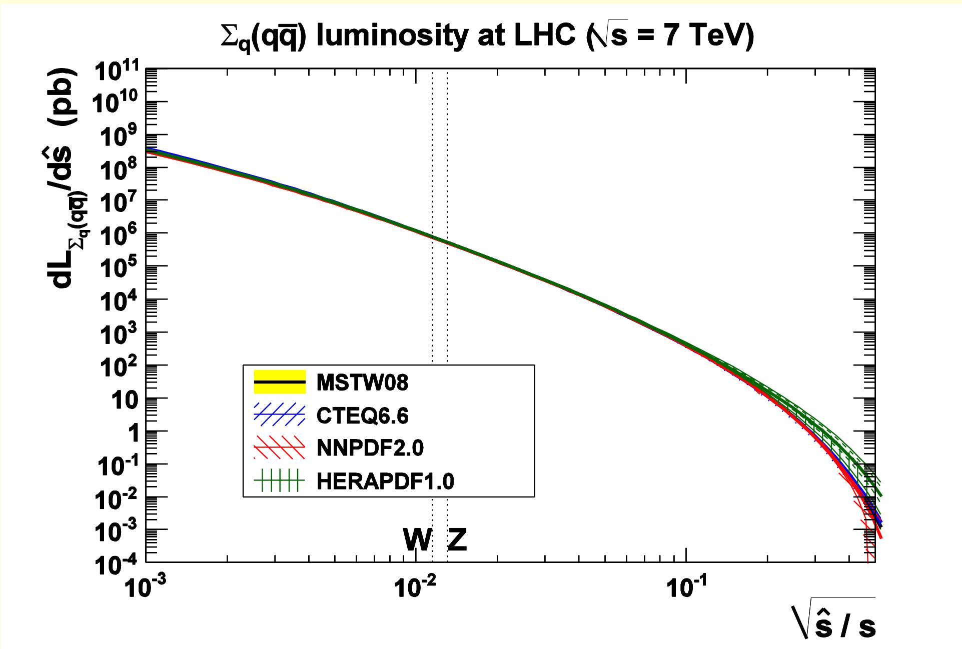
It is most useful for LHC physics if we compare PDFs in terms of parton-parton luminosities

q-qbar for W,Z production

$$\frac{\partial \mathcal{L}_{\Sigma_q(q\bar{q})}}{\partial \hat{s}} = \frac{1}{s}\int_\tau^1 \frac{dx}{x} \sum_{q=d,u,s,c,b} \left[f_q(x,\hat{s}) f_{\bar{q}}(\tau/x,\hat{s}) + f_{\bar{q}}(x,\hat{s}) f_q(\tau/x,\hat{s})\right],$$

And g-g for Top, Higgs

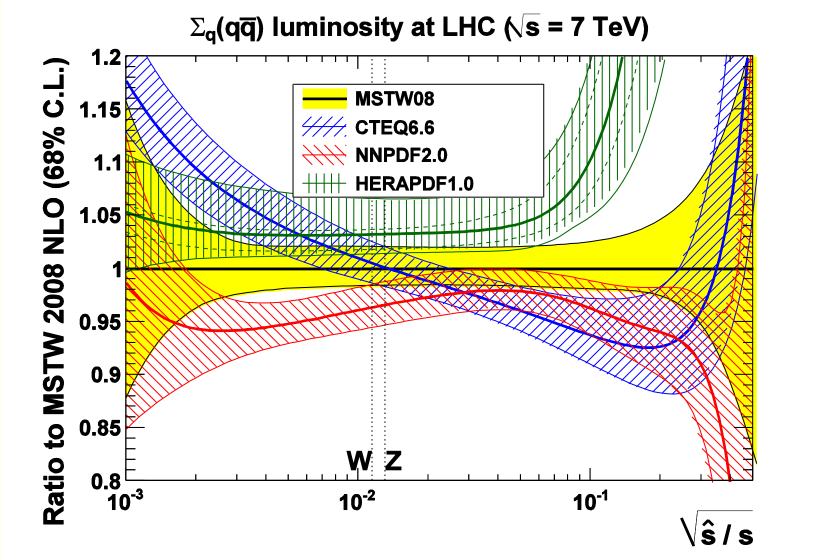
$$\frac{\partial \mathcal{L}_{gg}}{\partial \hat{s}} = \frac{1}{s}\int_\tau^1 \frac{dx}{x} f_g(x,\hat{s}) f_g(\tau/x,\hat{s}),$$

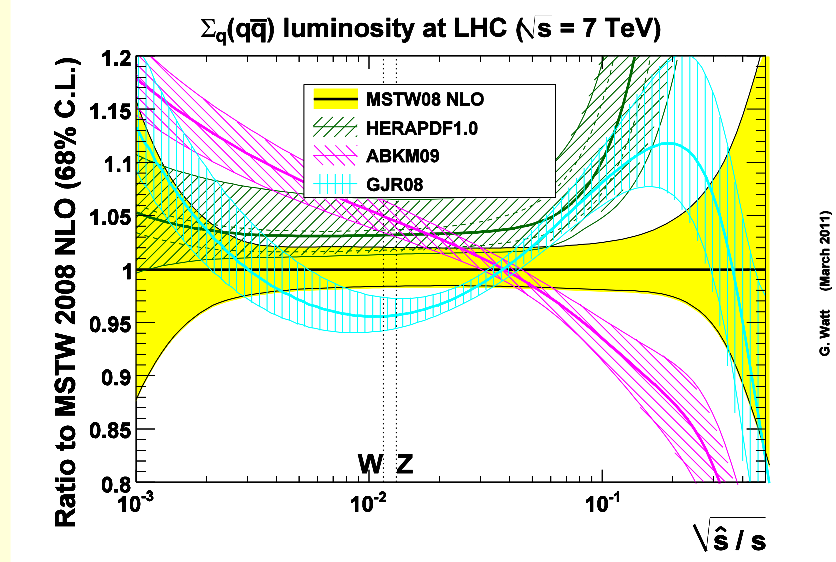


But since luminosities are very steeply falling functions of the invariant mass of the hard sub process

$$M^2 = x_1 x_2 s$$

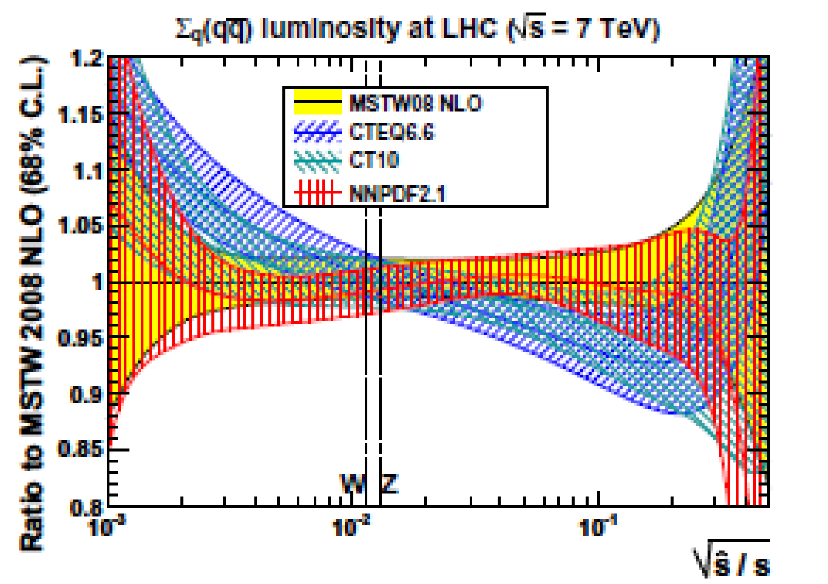
It is most instructive to compare them as ratios to a fixed PDF

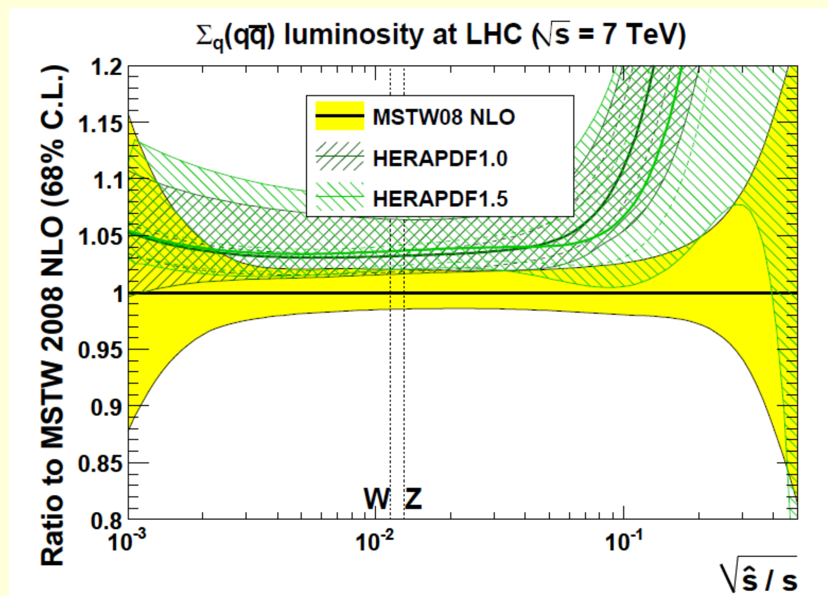


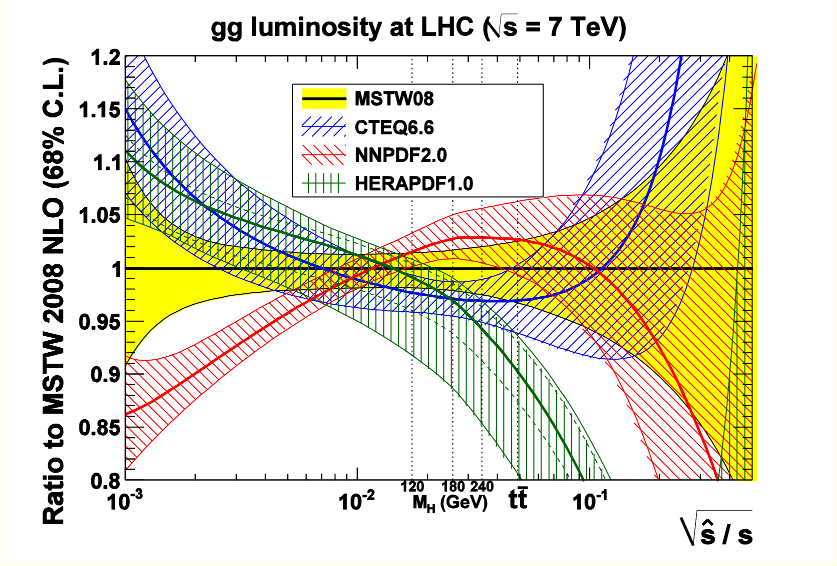


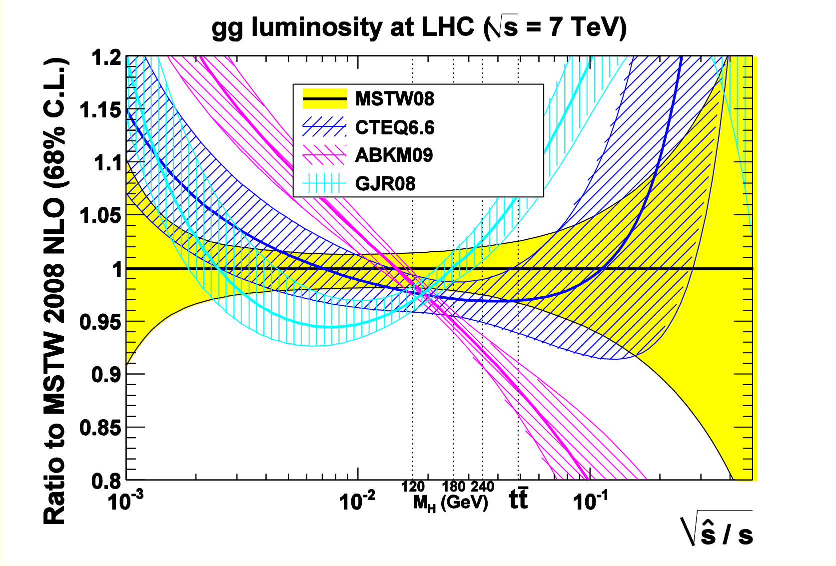
Here are the q-qbar luminosities first compared a few years ago and some of their recent updates. The NNPDF 2.0 →2.1 update involved a change in heavy quark scheme from Zero-Mass Variable Flavour Number to Massive Variable Flavour Number Scheme

The HERAPDF 1.0 →1.5 update involved new data









Here are the NLO g-g luminosities first compared a few years ago- updates are not dramatic

What is most noticeable?

The softer gluon at high scale (=high x) for HERAPDF and AB(K)M

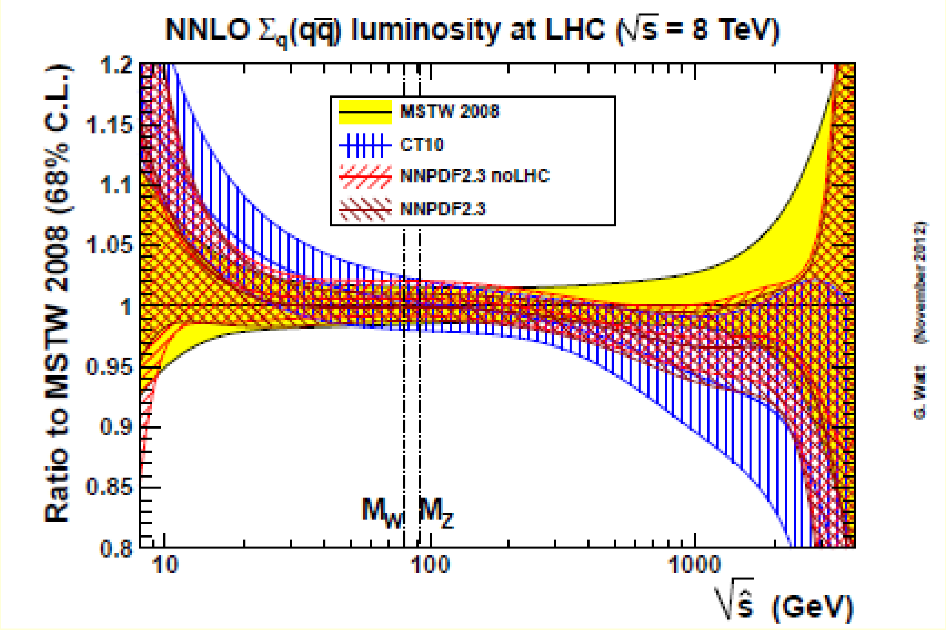
It is easy to say that not including Tevatron high-E_T jet data in these fits, leads to the softer high-x gluon BUT

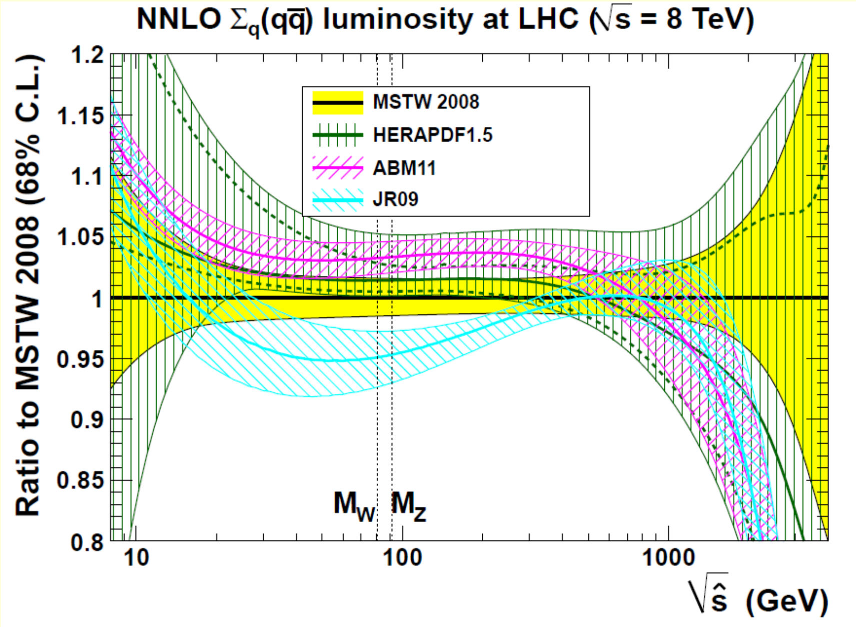
1. Both HERAPDF and AB(K)M can describe Tevatron jet data reasonably within their uncertainties. Arxiv1107.5170 (C-11-04-11-2)
2. HERAPDF at NNLO is not soft at high-x -and neither is JR09 which also does not have jet data......see next slides

Search for a deeper explanation

Look at NNLO for the most modern PDFs

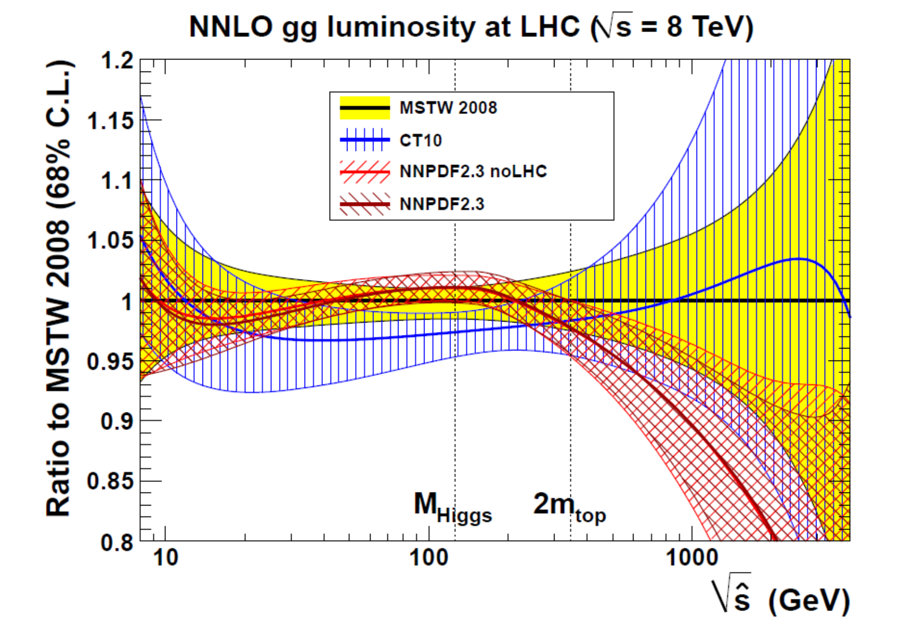
The q-qbar luminosities are somewhat closer at NNLO

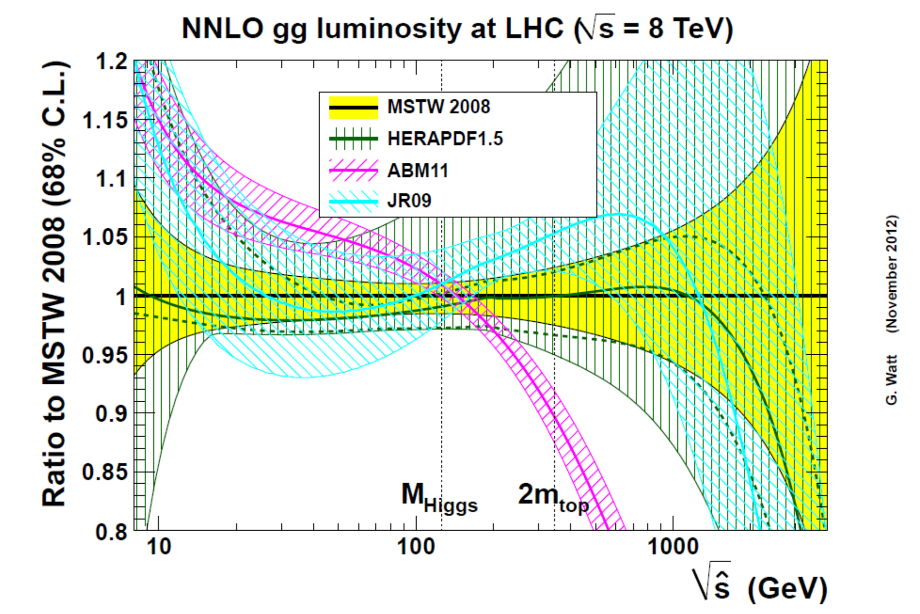




Notice how little difference leaving out LHC data makes--- but within this year thee will be many updates from the final 2011 and 2012 data which will change this– see later

But the g-g luminosity maintains a striking difference in shape for ABM





What are the **major** differences between ABM PDFs and others.

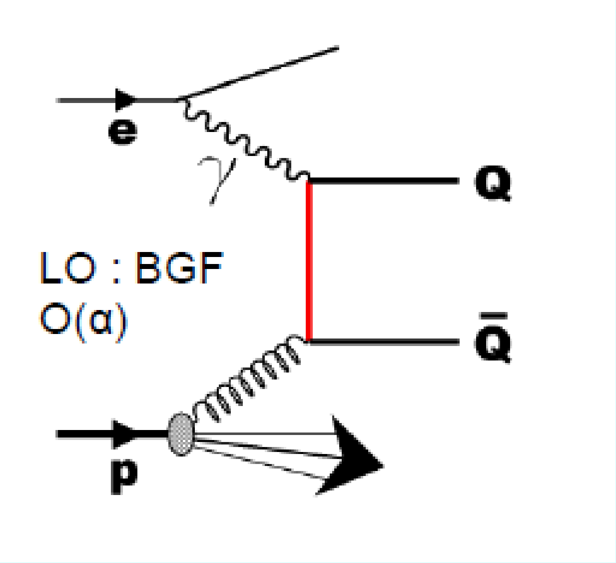
1. Higher twist terms at low W (low Q2, high x)
2. **Use of FFN not VFN for heavy quarks**

Investigation of the first by MSTW and NNPDF shows that they lead to only small differences – arXiv 1107.5170, Rojo, Thorne PDF4LHC Dec 12, Thorne PDF4LHC April 13

Heavy quark production in DIS

Fixed Flavour Number Scheme (FFNS)

- nf=3 active flavours in p
- heavy-quarks produced in hard scattering
- mass effects correctly included

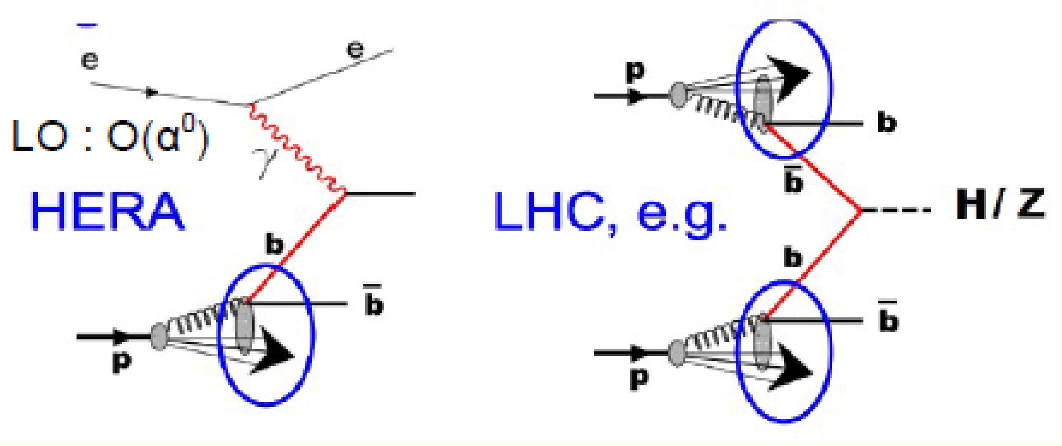


- spoiled by large logs of Q^2/m^2, p_T/m ...

Wrong at high Q2

Variable Flavour Number Scheme(s) (VFNS)

- c, b massless partons for $Q^2 > m_c^2$



- simplifies calculations at colliders (neglecting m_c)
- resums large $\log(Q^2/m^2)$

- Zero Mass (ZM) VFNS
  - neglects m_c at all Q^{2}s

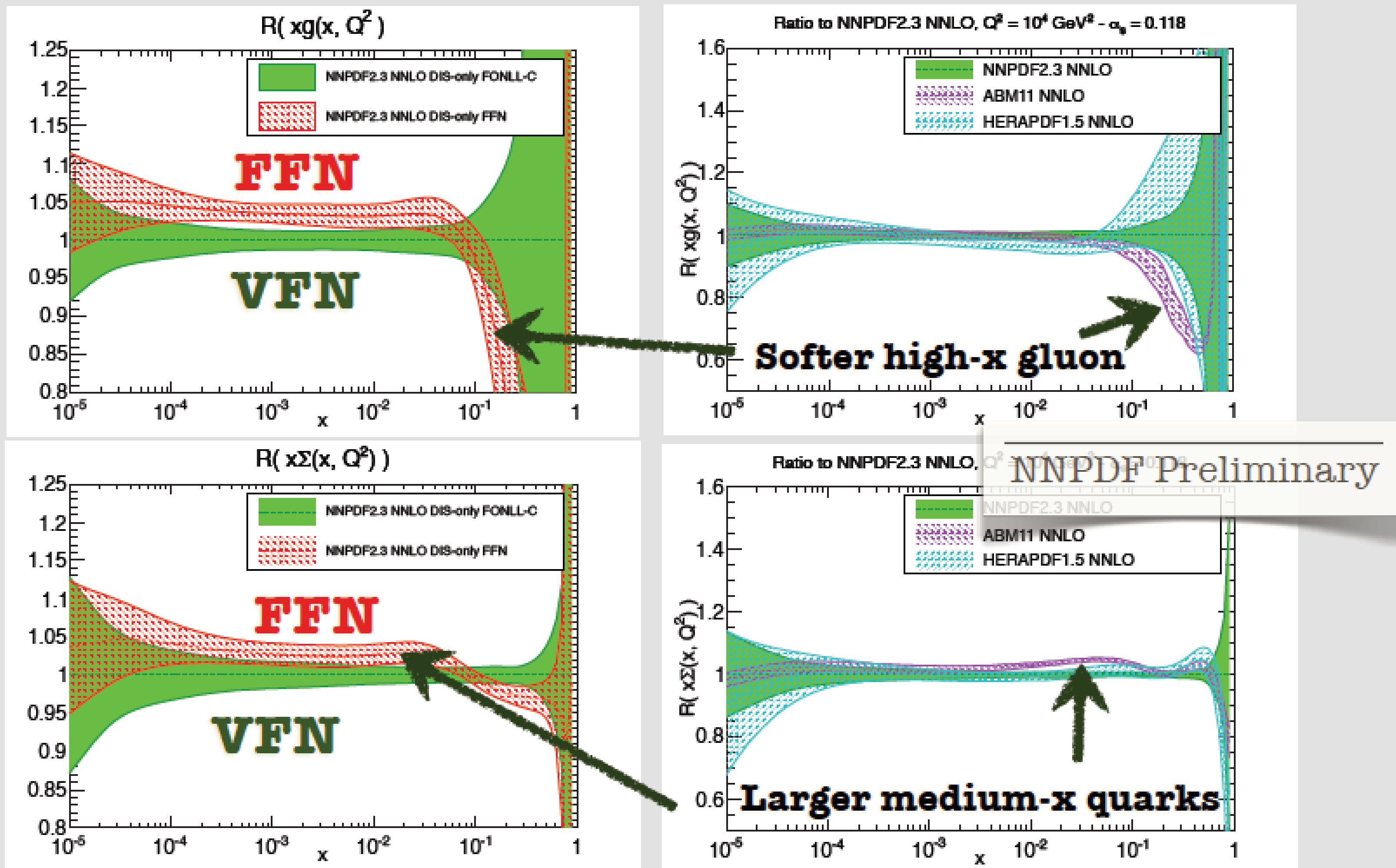
Wrong at threshold

- General Mass (GM) VFNS
  - FFNS at $Q^2 < m_c^2$, ZM-FNS at $Q^2 >> m^2$
  - Interpolating in between
  - different prescriptions available

Not unique

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**



These fits are done with the same value of alphas- so the PDF shape change does not come just from difference of $\alpha_S(M_Z)$

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 significant** and specially relevant for the **inclusive HERA-I**: due to **missing resummation of DGLAP logarithms**

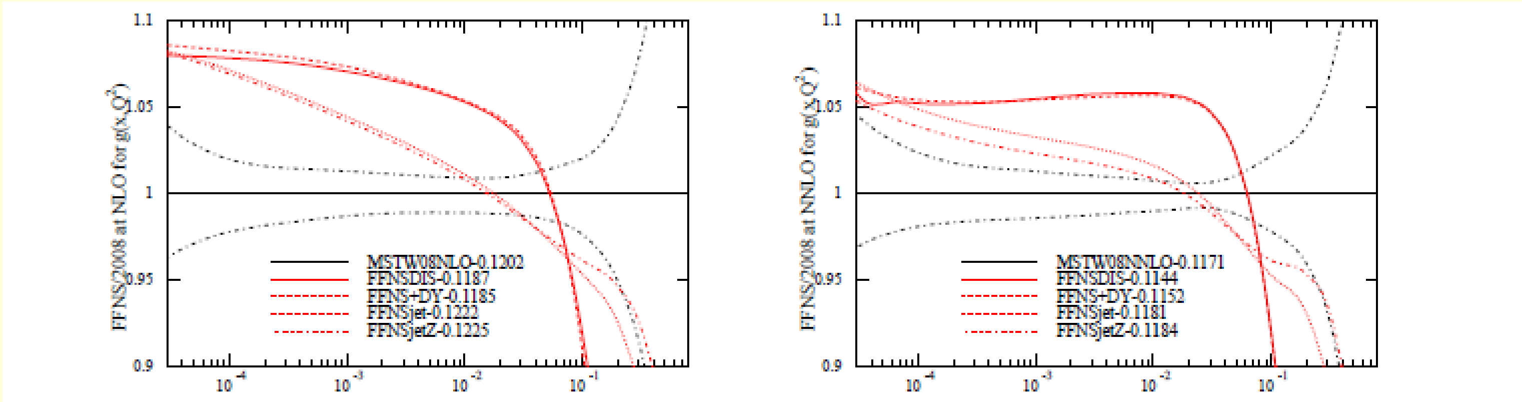
$x_{\rm min}$	$x_{\rm max}$	$Q^2_{\rm min}$	$Q^2_{\rm max}$	$\chi^2_{\rm tot}({\rm FFN}-{\rm VFN})$	$N_{\rm dat}^{\rm tot}$	$\chi^2_{\rm hera}({\rm FFN}-{\rm VFN})$	$N_{\rm dat}^{\rm hera}$
10^{-6}	1.0	3.0	10^6	28.26	2936	37.88	592
10^{-6}	1.0	3.0	10^6	68.88	1055	39.73	405
10^{-6}	1.0	3.0	10^6	28.54	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**

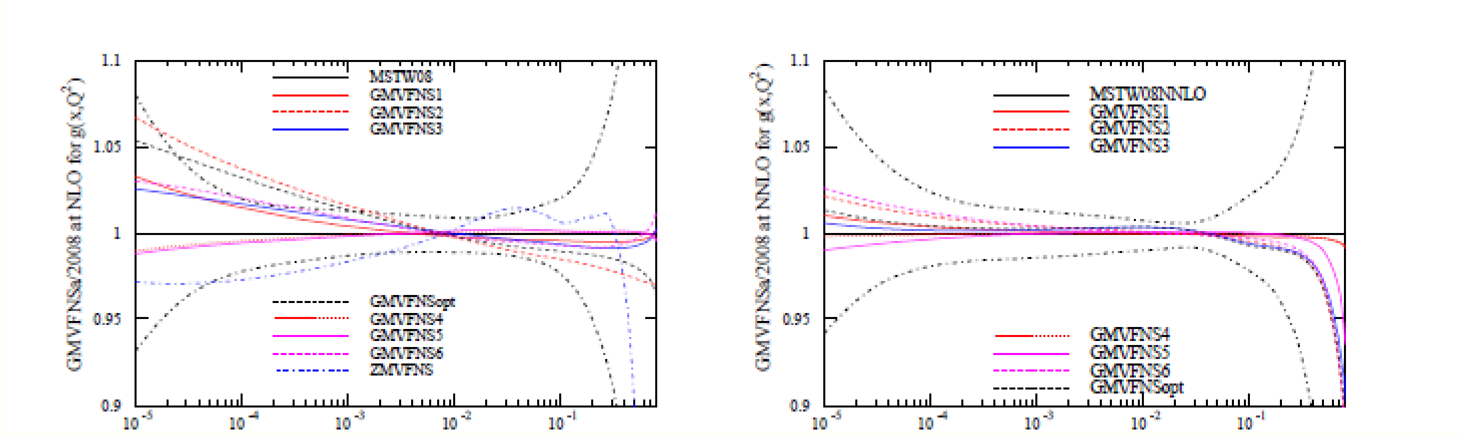
- Although FFN provides a reasonable description of HERA-I data, a **better fit quality** obtained by VFN thanks to **DGLAP resummation at moderate and large Q2**
- FFN and VFN similar χ^2 at **small Q2**

NNPDF Preliminary

MSTW find similar trends at NLO and NNLO, using various different data sets



And the changes using different VFN schemes are much smaller



Thorne also fits for $\alpha_S(M_Z)$ in both schemes, finding 0.1171 for GMVFN and 0.1136 for FFN- consistent with ABM.

In the FFN scheme the gluon needs to be bigger at x~0.001-0.1 and smaller at higher x to fit data. This results in a lower fitted value of $\alpha_S(M_Z)$.

Now let us see how the PDFs fare compared to LHC data

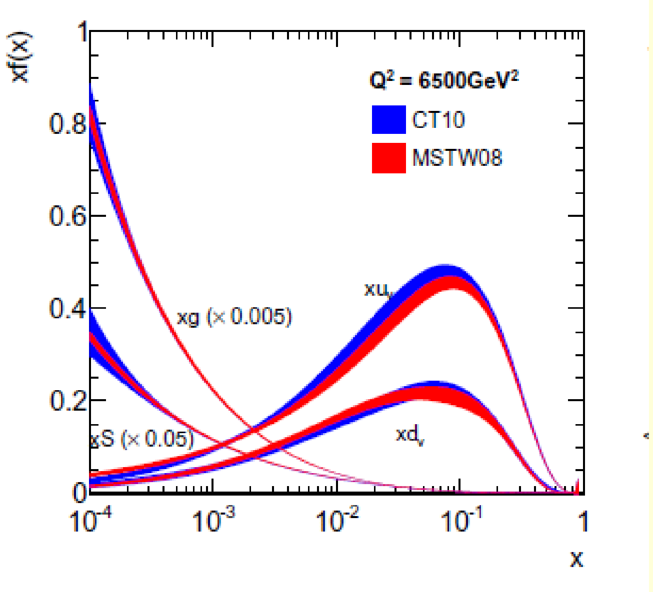
1. W-asymmetry data distinguishes between different PDFs in the valence quark sector- MSTW2008 is ruled out
2. W and Z data together suggest strangeness is not suppressed as traditionally thought- at least at low-x $< \sim 0.01$
3. W+c data from CMS and ATLAS marginally disagree on strangeness suppression
4. Jet production data suggest that the high-x gluon is less hard than determined from Tevatron data. CT10 does well, HERAPDF does surprisingly well, ABM is doing badly
5. Use of top production data to improve/discriminate PDFs
6. Predictions for Higgs

W and Z production are the best known sub-process cross-sections: known to NNLO, so how did current PDFs do in predicting what we have actually measured?

And at central rapidity $x_1 = x_2$ and assuming ubar = dbar (at small x)

$$\text{So } A_W \sim \frac{(u-d)}{(u+d)} = \frac{(u_v - d_v)}{(u_v + d_v + 2\,\text{qbar})}$$

And the PDF predictions for valence differ at small-x

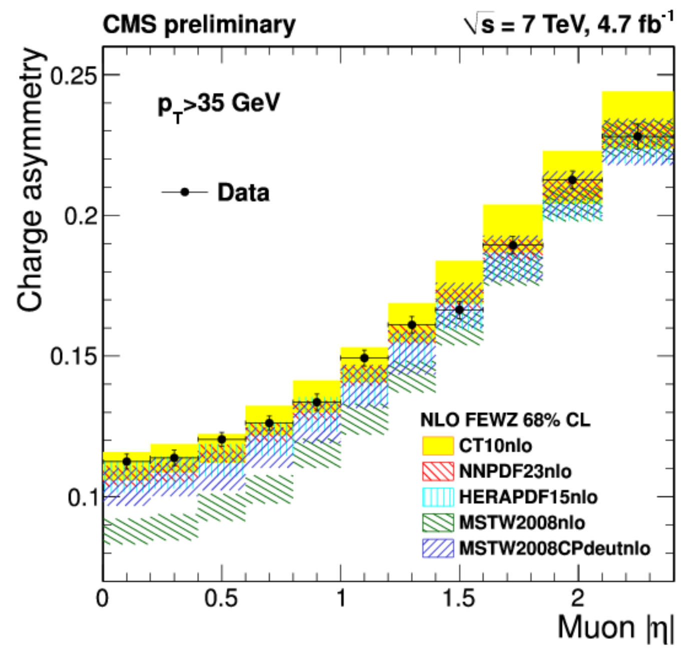


LHC data probe precisely the x range $10^{-3} < x < 10^{-1}$ where the difference is maximal

W-asymmetry

$$A_W = [\sigma(W^+) - \sigma(W^-)]/[\sigma(W^+) + \sigma(W^-)]$$

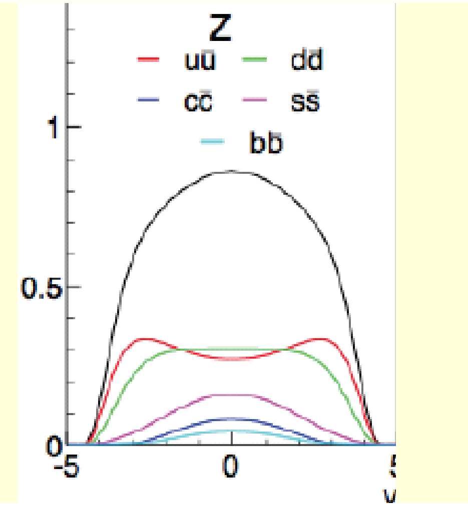
This translates into a difference in predictions for the W-lepton asymmetry pseudo-rapidity spectrum:



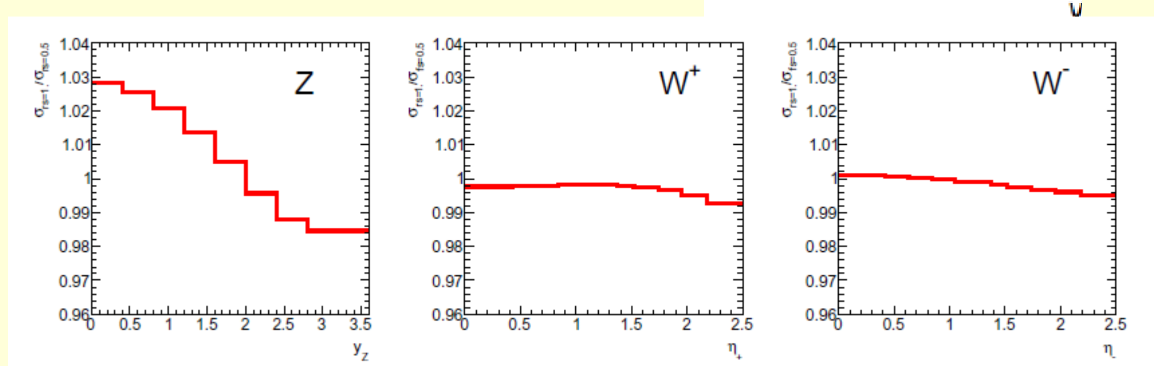
The CMS muon asymmetry data from 2011 (arXiv:1312.6283) clearly disfavour MSTW2008 (MSTW have addressed this in MSTWCPdeut)

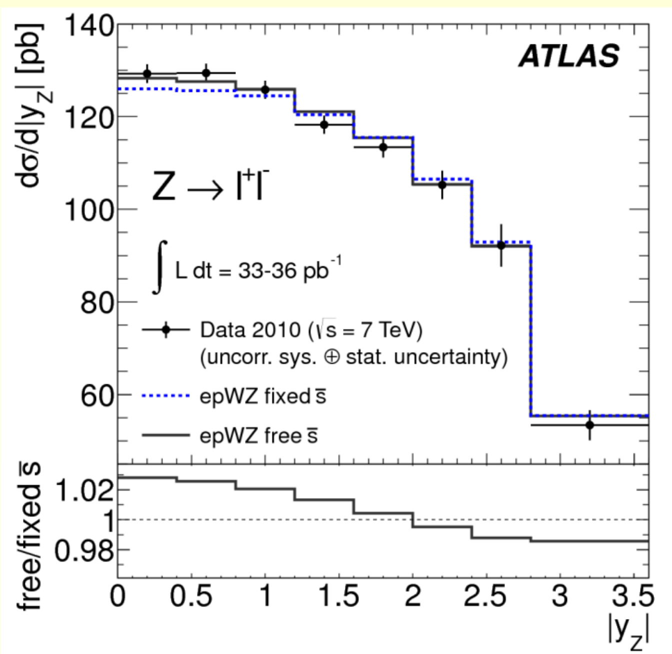
W and Z differential cross sections

Flavour contributions to W and Z show that s-sbar is prominent in Z production at central rapdidty.
This plots were made for the usual assumption that strange sea is suppressed ~0.5 of down sea.
This comes from di-muon production in neutrino induced deep inelastic scattering data.



How would Z and W rapidity spectra at the LHC change if strangeness were enhanced?Here is the ratio of Z and W cross-sections for strange = down sea in ratio to strange = 0.5 down sea
This is a small effect ~ 4%- can we see it?





YES WE CAN: ATLAS Phys Rev Lett 109(2012)012001

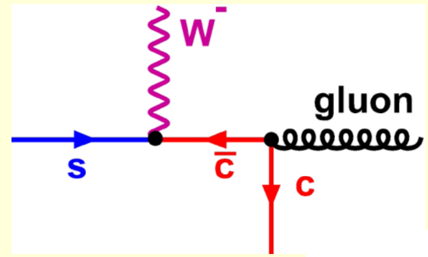
NNLO PDF fits to the ATLAS W,Z data plus HERA data are shown for two assumptions about strangeness: s/d = 0.5 fixed and $s/d = r_s\,(1-x)^{(Cs-Cd)}$ – fitted.

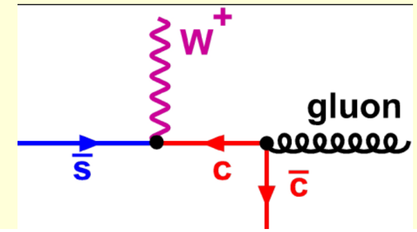
The fit gives $s/d = r_s = 1.0 \pm 0.25$

$$r_s = 1.00 \pm 0.20_{exp} \pm 0.07_{mod}\ {}^{+0.10}/_{-0.15\ par}\ {}^{+0.06}/_{-0.07\ \alpha s} \pm 0.08_{th}$$

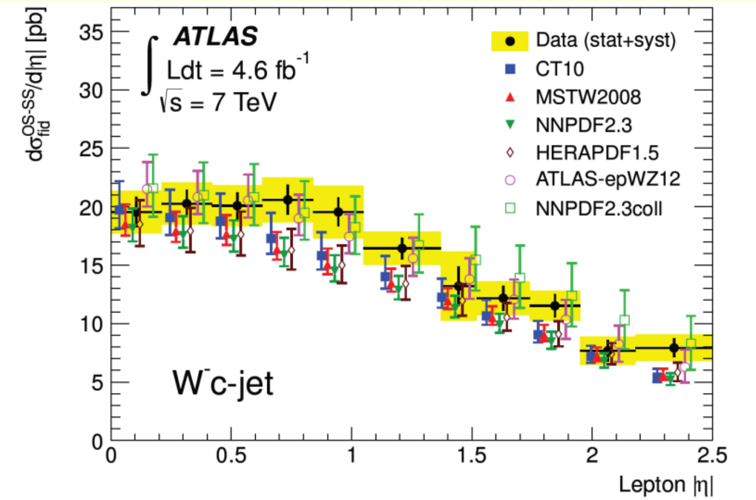
Another process which can yield information on strangeness is W+c production

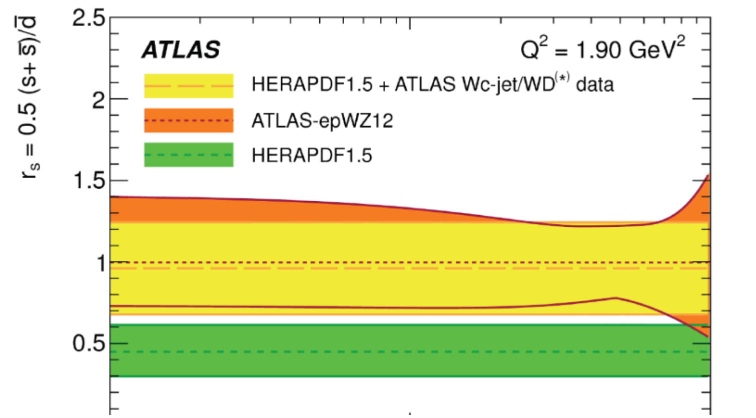
CMS SMP-12002
JHEP05(2014)068





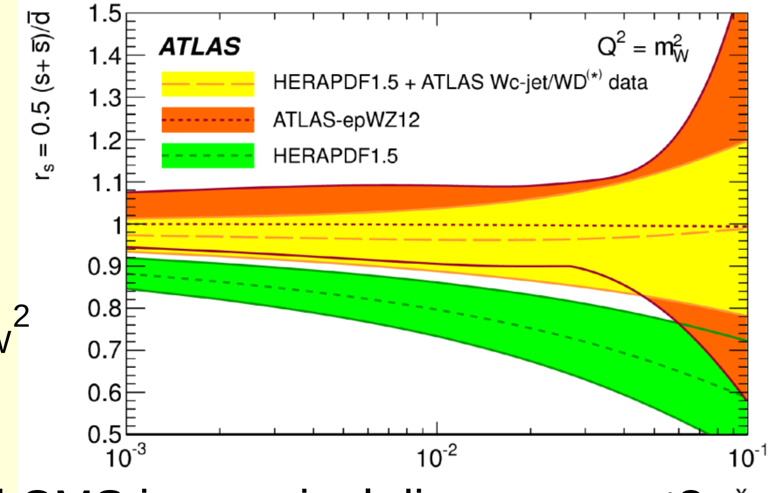
Both CMS and ATLAS have analysed such data



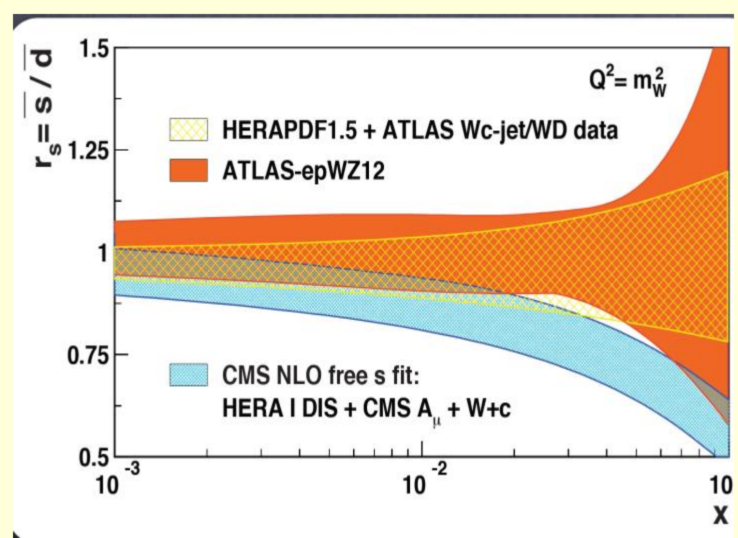


$$\rightarrow r_s \equiv 0.5(s+\bar{s})/\bar{d} = 0.96\,{}^{+0.16}_{-0.18}\,{}^{+0.21}_{-0.24}$$

And once you evolve to $Q^2 \sim M_W^2$



ATLAS and CMS in marginal diasgreement?



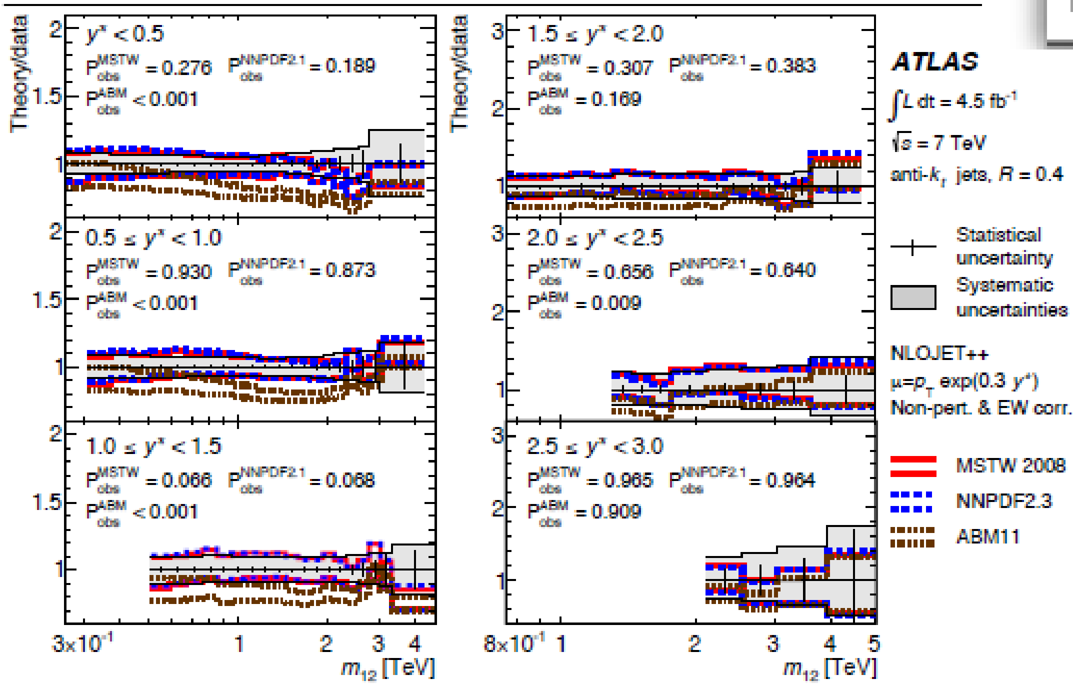
ATLAS di-jets JHEP05(2014)059

Comparison with PDFs

A frequentist method is employed to asses the probability that the measured cross sections are described by the SM predictions for each PDF considered. Different rapidity and mass ranges are considered.

PDF set	y^* ranges	mass range (full/high)	P_{obs} $R = 0.4$	P_{obs} $R = 0.6$
CT10	$y^* < 0.5$	high	0.742	0.785
	$y^* < 1.5$	high	0.080	0.066
	$y^* < 1.5$	full	0.324	0.168
HERAPDF1.5	$y^* < 0.5$	high	0.688	0.504
	$y^* < 1.5$	high	0.025	0.007
	$y^* < 1.5$	full	0.137	0.025
MSTW 2008	$y^* < 0.5$	high	0.328	0.533
	$y^* < 1.5$	high	0.167	0.183
	$y^* < 1.5$	full	0.470	0.352
NNPDF2.1	$y^* < 0.5$	high	0.405	0.568
	$y^* < 1.5$	high	0.151	0.125
	$y^* < 1.5$	full	0.431	0.242
ABM11	$y^* < 0.5$	high	0.024	$< 10^{-3}$
	$y^* < 1.5$	high	$< 10^{-3}$	$< 10^{-3}$
	$y^* < 1.5$	full	$< 10^{-3}$	$< 10^{-3}$

Dijet production from 2011 data



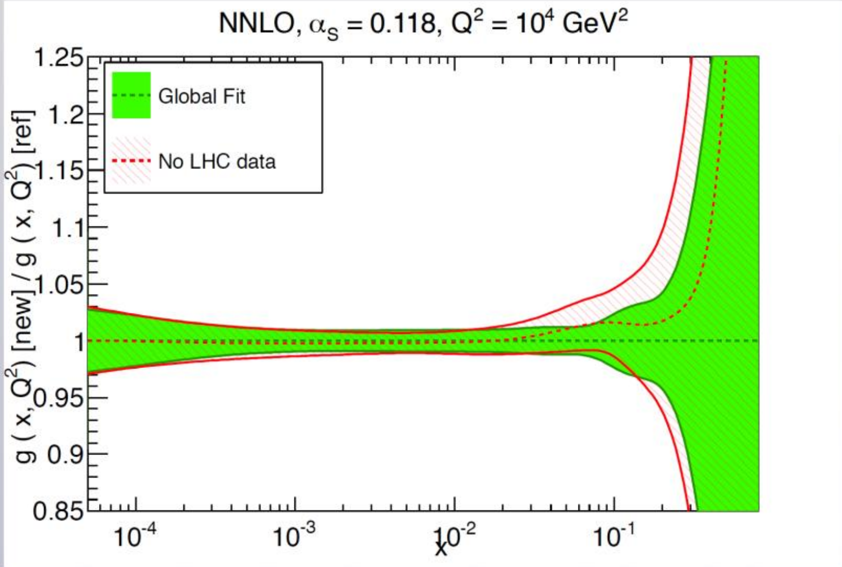
- Lower prediction from ABM, possibly due to the softer gluon

Comparison with PDFs

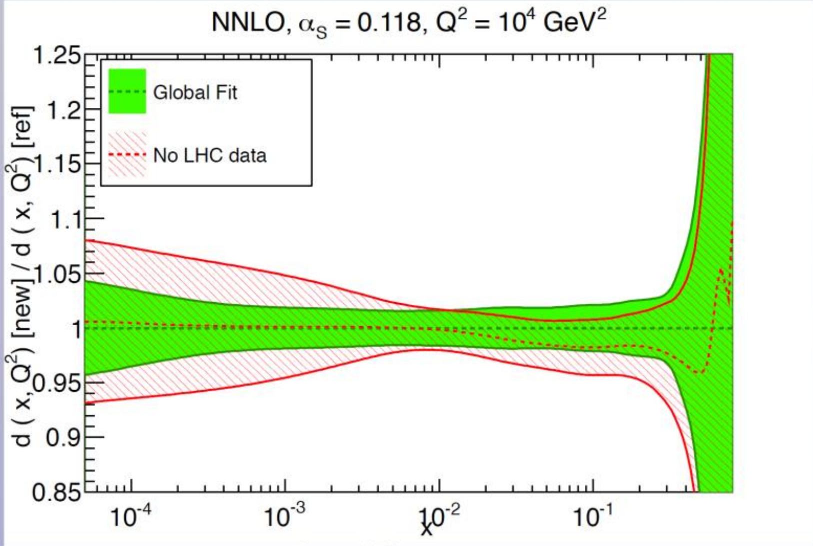
Looks OK for HERAPDF despite the fact that HERAPDF1.5 does not use Tevatron jet data and thus has a softer gluon than global fts.

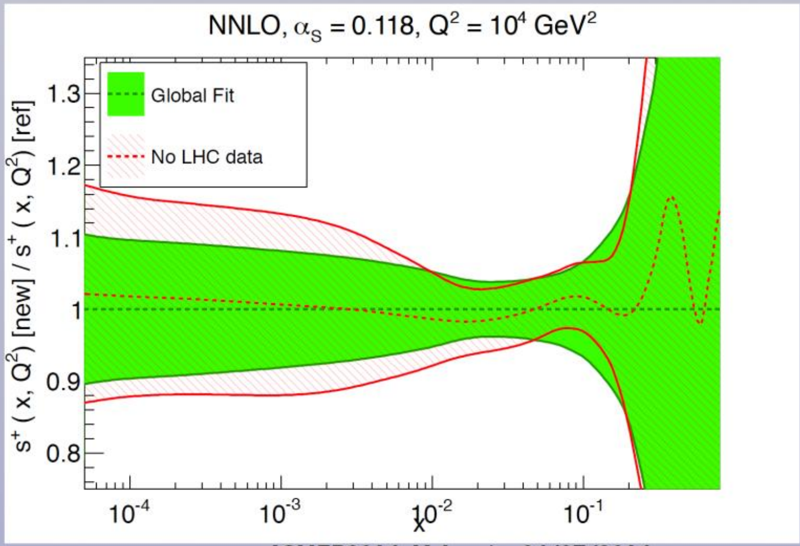
Looks worst for ABM11

Impact of LHC data



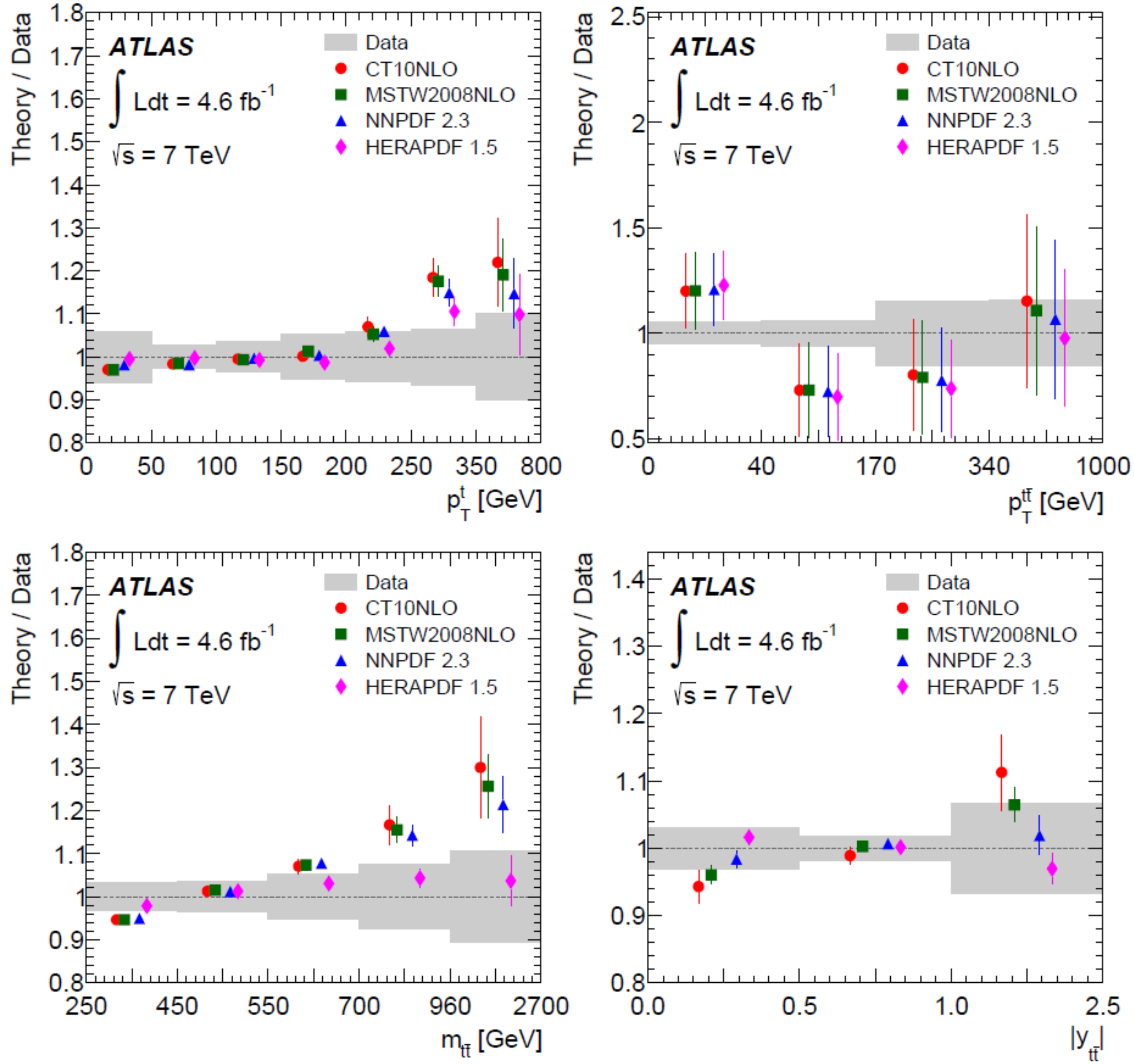
- Compare global NNPDF3.0 fit with a fit **without LHC data**
- PDF uncertainties on **large-x gluon** reduced due to **top quark and jet data**
- PDF uncertainties on **light quarks** reduced from the **Drell-Yan** and **W+charm data**
- The **description of all new LHC data**, already good in NNPDF2.3, is further improved in NNPDF3.0



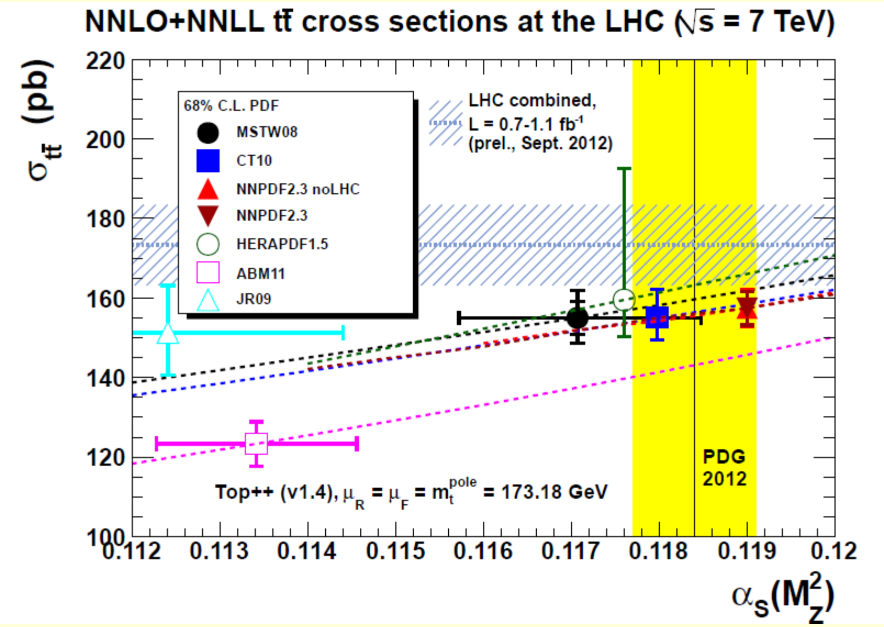


t-tbar production can improve the gluon PDF -NLO

ATLAS top differential distributions from 4.6fb^{-1} of 7 TeV arXiV:1407.0371
These look promising for PDF discrimination but calculations are so far only NLO, full NNLO is coming.

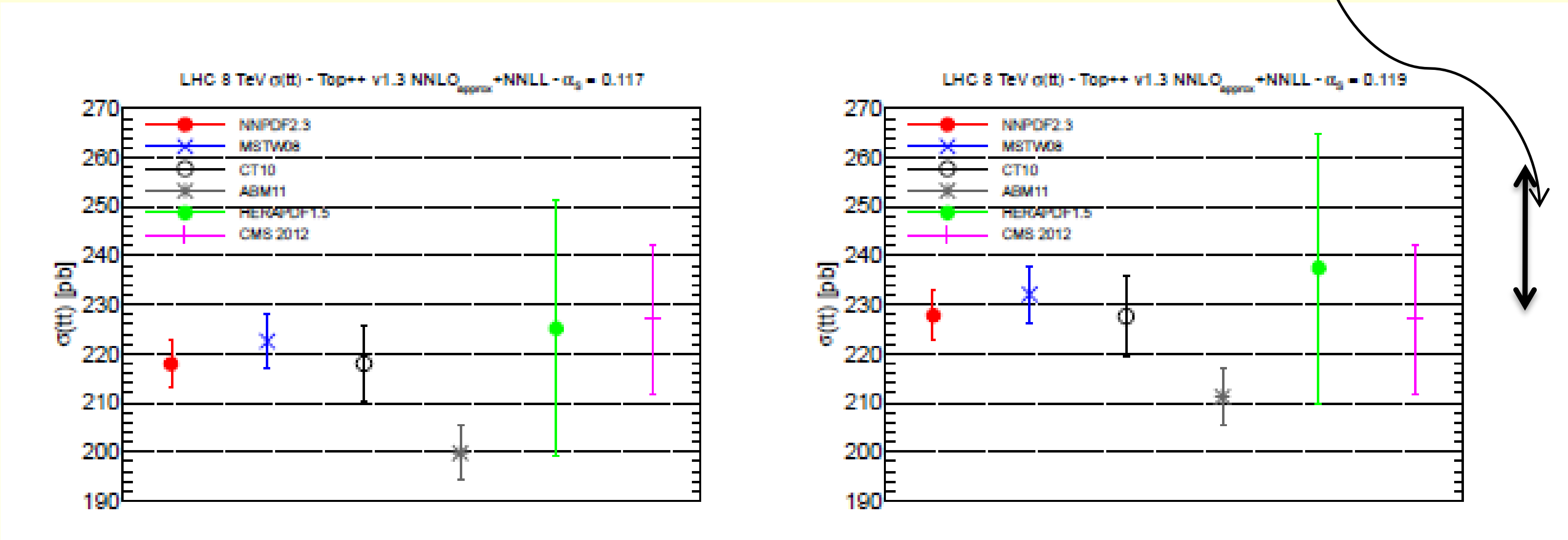


Do the top cross sections already provide PDF discrimination?



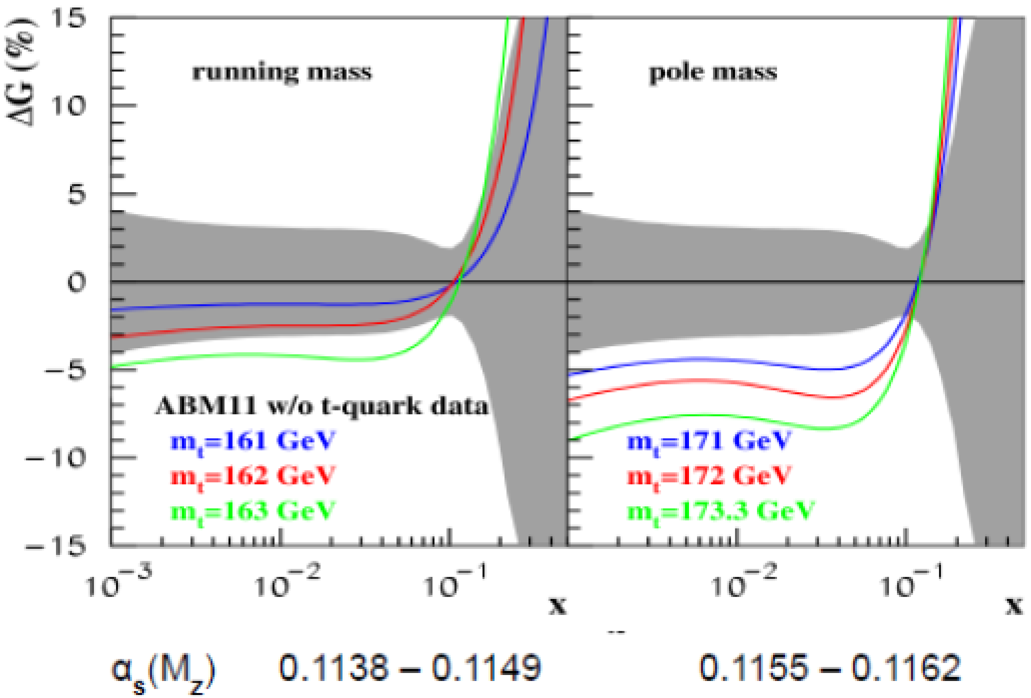
The ATLAS and CMS combined t-tbar cross section is 173 ± 2.3 ± 9.8pb at 7 TeV
ATLAS-CONF-2012-134/ CMS-TOP-12003
The predictions for this cross section have a strong $\alpha_S(M_Z)$ dependence.
But even if we use the same alphas values predictions differ

How about at 8 TeV? The range of the ATLAS and CMS top cross-sections presented at ICHEP14 is to the high side of the predictions



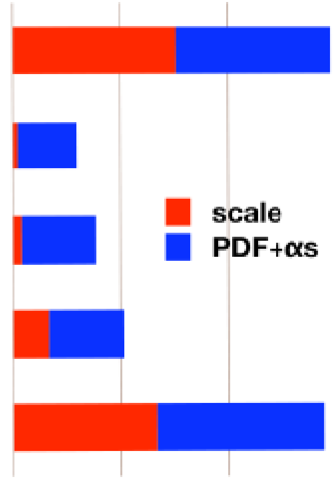
BUT the calculation of the t-tbar cross section Also depends on the top quark mass.
On the previous page the value 173.2 GeV was used.
The calculation also depends on whether running-mass or pole-mass is used

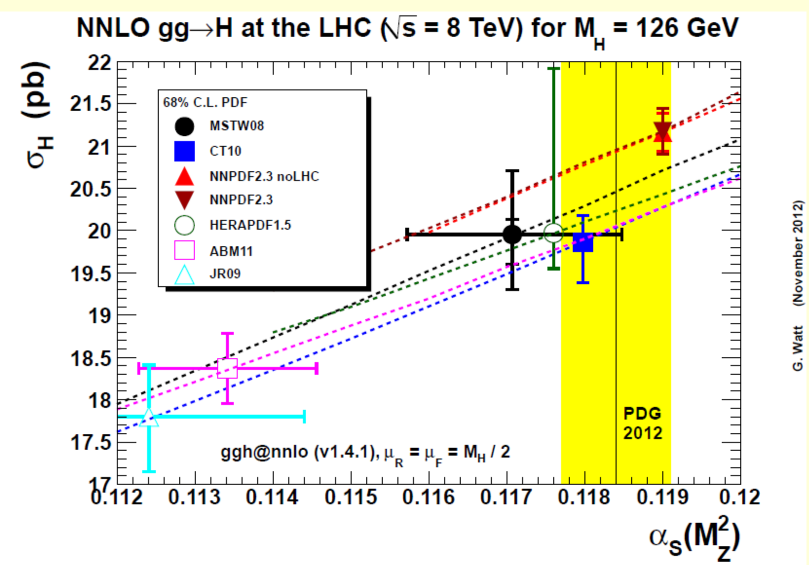
ABM have used the cross section data in their own fit and they find that a running mass calculation with Mt =161 GeV (and $\alpha_S(M_Z)$=0.1138) is compatible.
However it has a dramatic effect on the shape of the gluon, which is becoming MUCH harder at high-x.... Top measurements thus suggest that these PDFs need to be modified



PDFs and the Higgs

		σ (8 TeV)	uncertainty
NNLL QCD +NLO EW	gg→H	19.5 pb	14.7%
	VBF	1.56 pb	2.9%
NNLO QCD +NLO EW	WH	0.70 pb	3.9%
	ZH	0.39 pb	5.1%
NLO QCD	ttH	0.13 pb	14.4%





The gg →Higgs cross section is strongly $\alpha_S(M_Z)$ and gluon PDF dependent, rather like the t-tbar cross section

The extent to which the Higgs that we are seeing agrees with the SM Higgs cross section predictions depends on the PDF and $\alpha_S(M_Z)$ value used for these predictions.

Summary

•Uncertainties on Parton Distribution Functions (PDFs) limit our knowledge of cross sections whether SM or BSM.

•Any claim for new physics at the highest masses is dependent on the PDF chosen to describe conventional physics

•Standard Model LHC measurements can themselves contribute to PDF discrimination and PDF improvement

Extras

Sizes of uncertainties

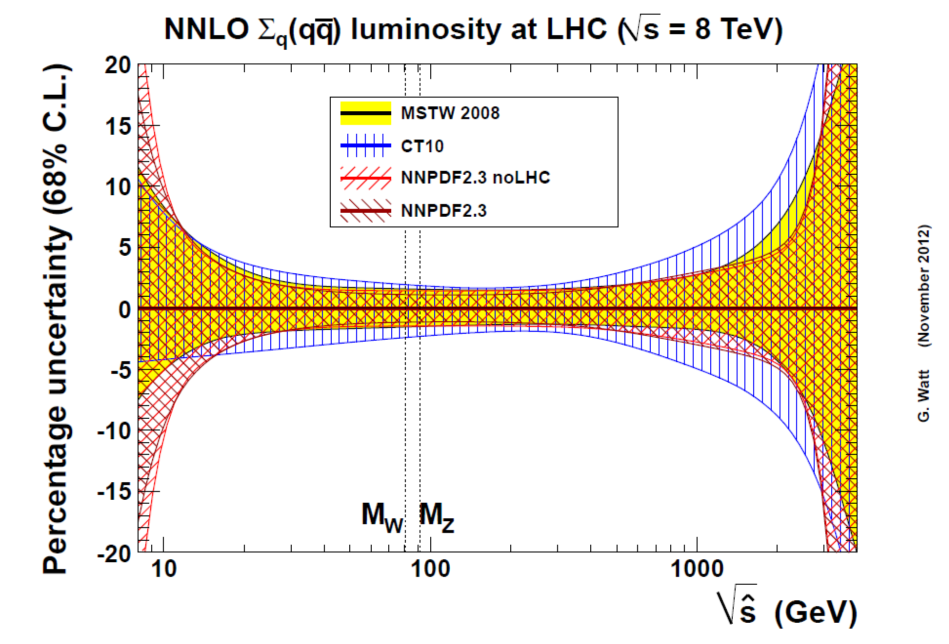
•Obviously more data sets –in the relevant kinematic region-should in general lead to lower uncertainties .But data sets are not all perfectly consistent. To account for this CT , MSTW, (G)JR do not apply simple Δχ2=1 criteria to set their errors. The have enhanced χ2 tolerances. Broadly speaking these come from the requirement that each indvidual data set be fit to within its 68% (or 90%) CL.

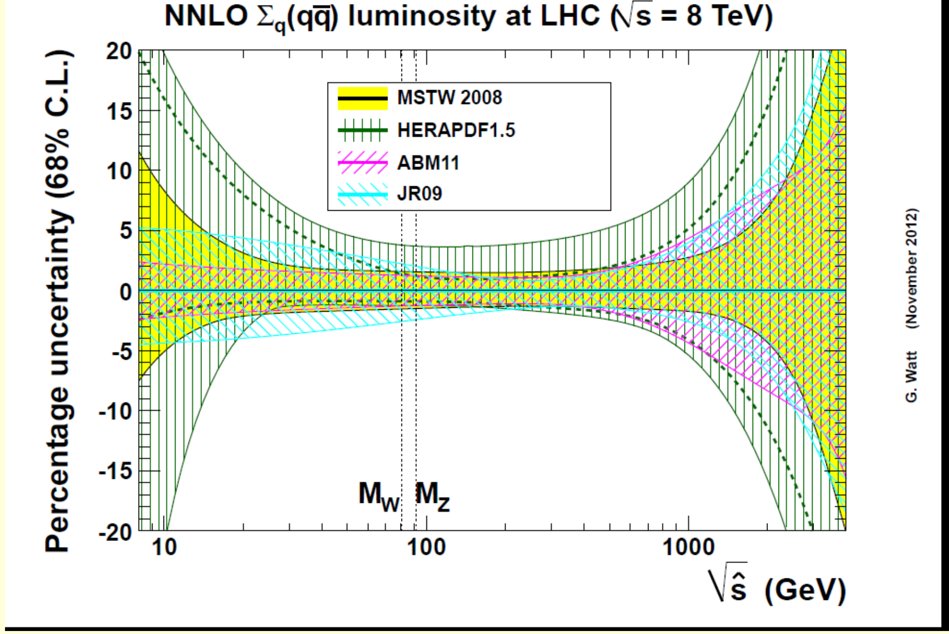
•NNPDF determine their uncertainties from fitting Mote-Carlo replicas which cover the full correlated uncertainties of each data set. The fact that the level of uncertainties of NNPDF, CT and MSTW are roughly comparable is a verification of the reasonableness of their methods (since they use almost the same data sets)

•ABM and HERAPDF use Δχ2=1 on more restricted data sets (for HERAPDF only the HERA data are used in order to have a consistent data set) but have extra uncertainties. For ABM these come from the higher twist. For HERAPDF these come from model and parametrisation variations.

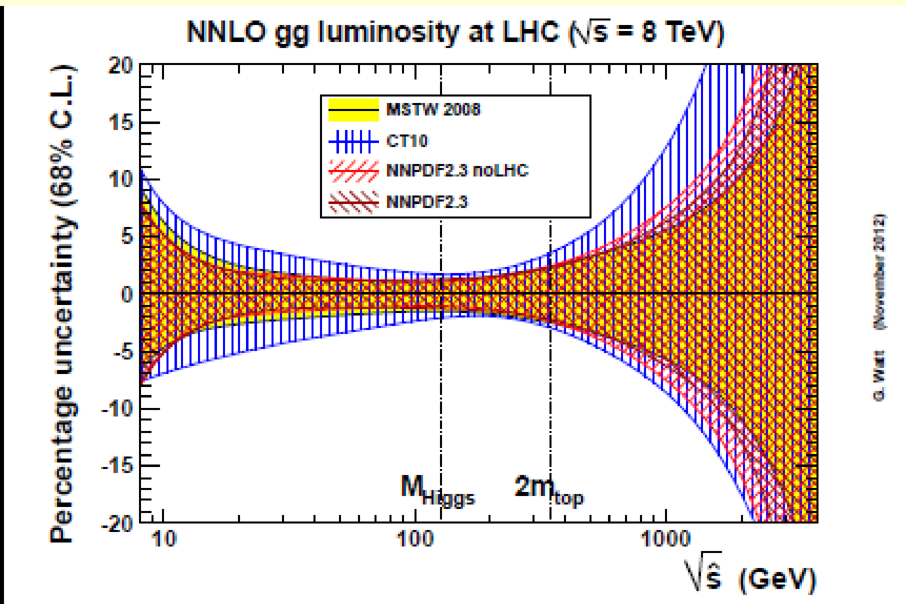
•One source of paramerisation variation which is not often considered is the choice of Q^2_0 the starting scale for Q^2 evolution (arXiv 1202.4262)

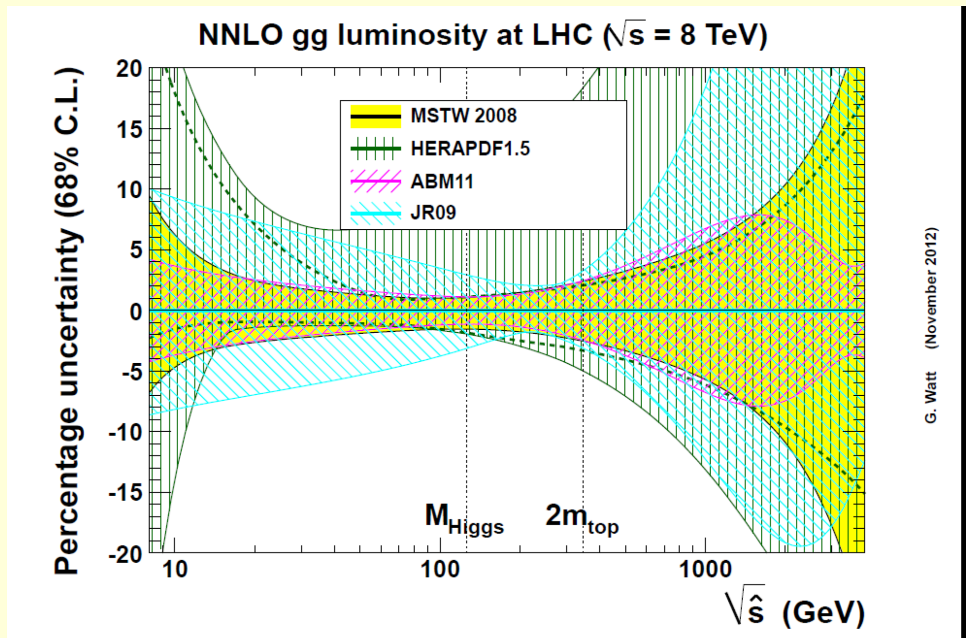
•One source of model uncertainty which is not often considered-in recent years- is the choice of the lower Q^2 cut for which data are allowed into the fit.

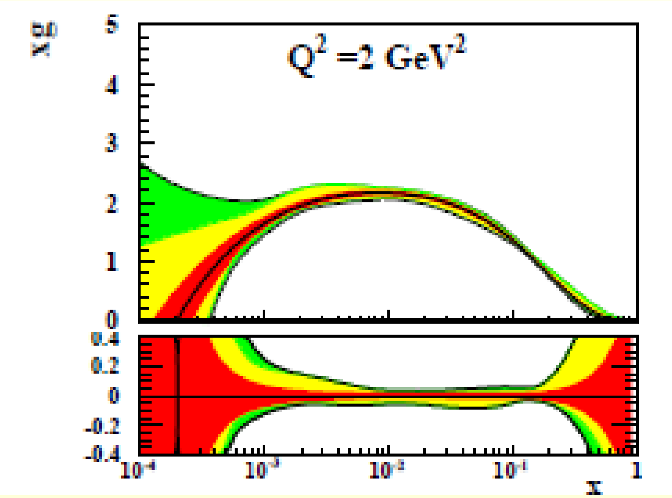


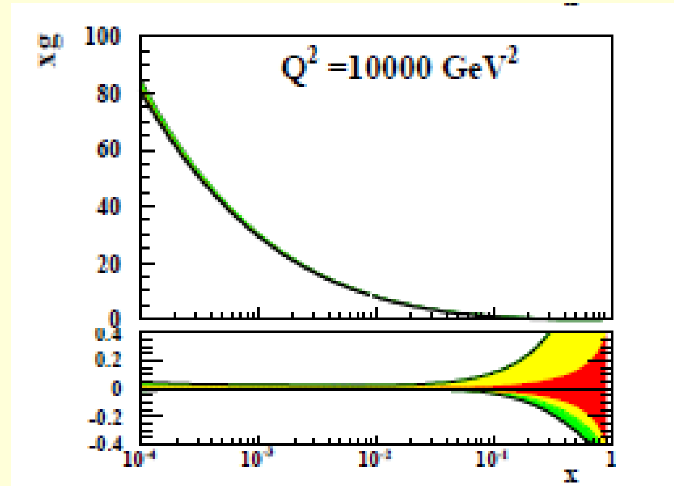


Look at the differences in uncertainties rather than in shape.
Are the larger errors of HERAPDF JUST because it has less input data sets?

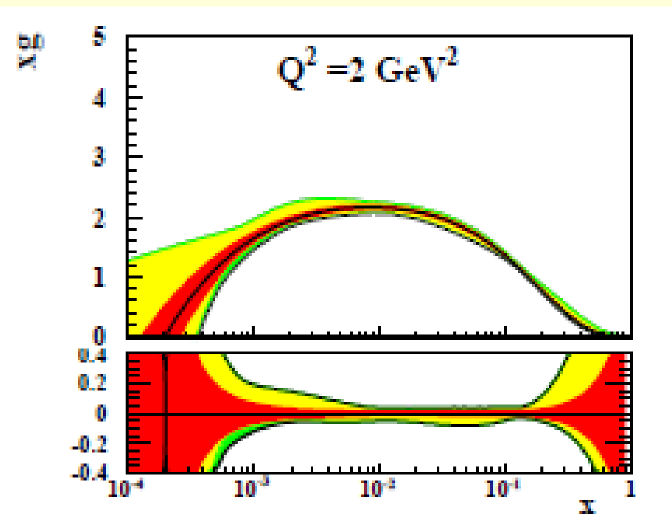


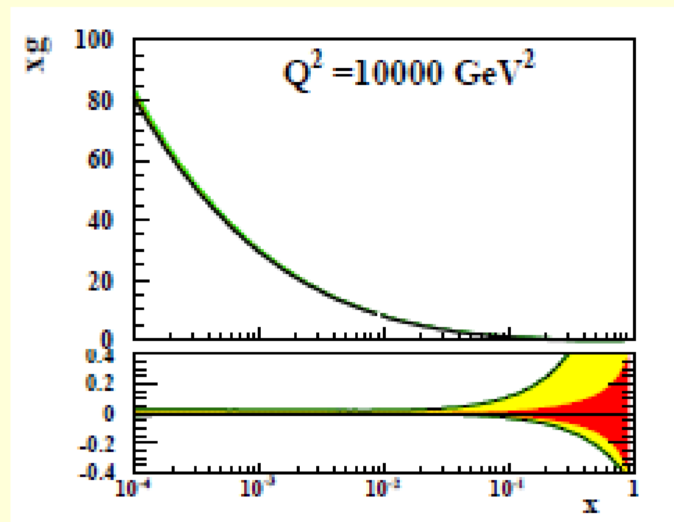






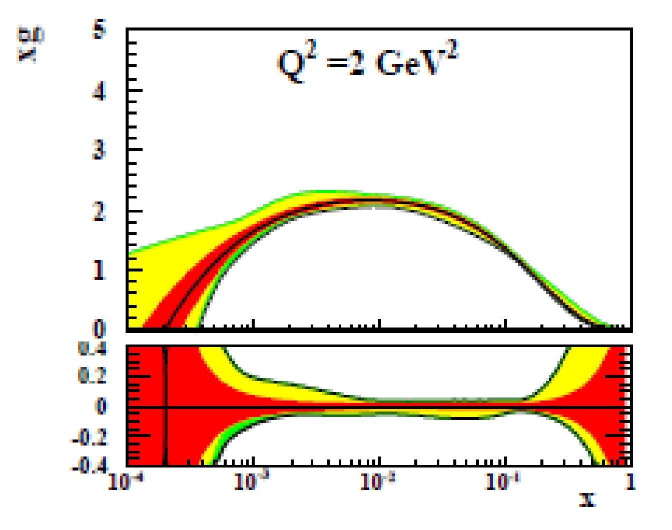
Part of the uncertainty is due to the choice of the starting scale- Q^2_0=1.9 is varied to Q^2_0=1.5 and 2.5 GeV^2. This is indicated in green

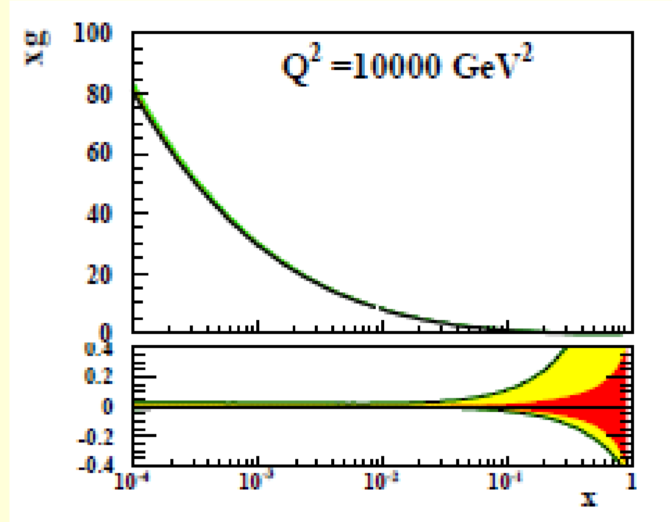




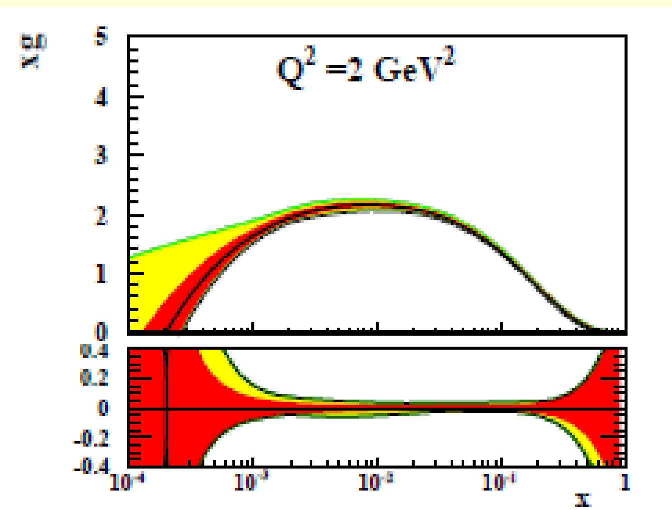
Remove the Q^2_0=2.5 variation

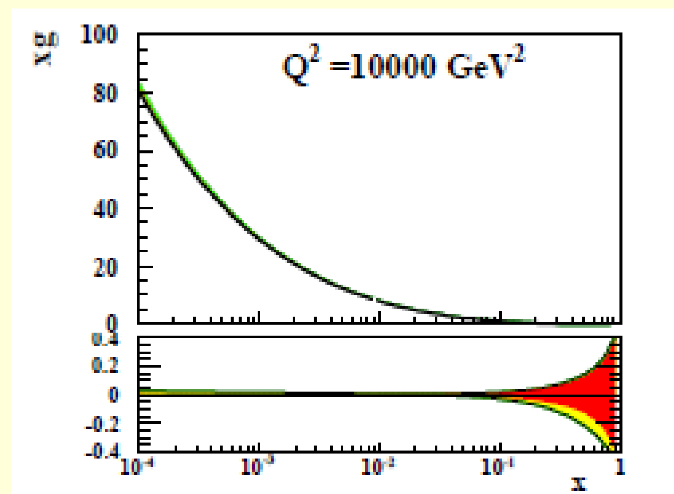
So a large part of the low scale error and a small part of the high-scale error is due to Q^2_0 variation



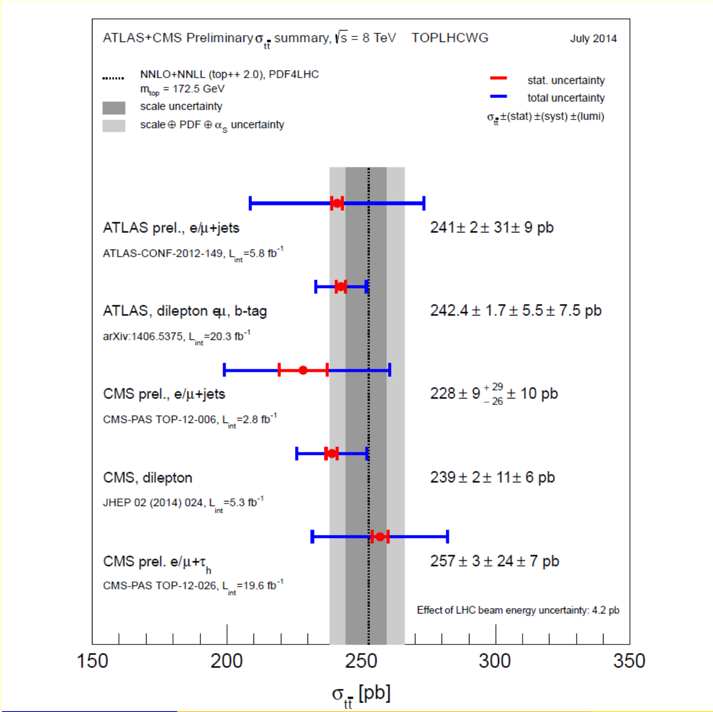


The greatest part of the contribution to the model uncertainties comes from the choice of the Q^2cut on the data. Q^2>3.5 is varied to.2.5 and 5.0 GeV^2. Model uncertainties are shown in yellow





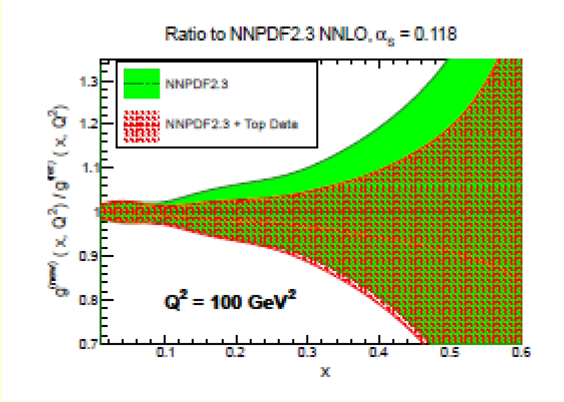
Without the Q^2>5 variation there is very little model error left on the high-x gluon at high scale- so **a LARGE part of the high-scale error is due to the Q^2cut**

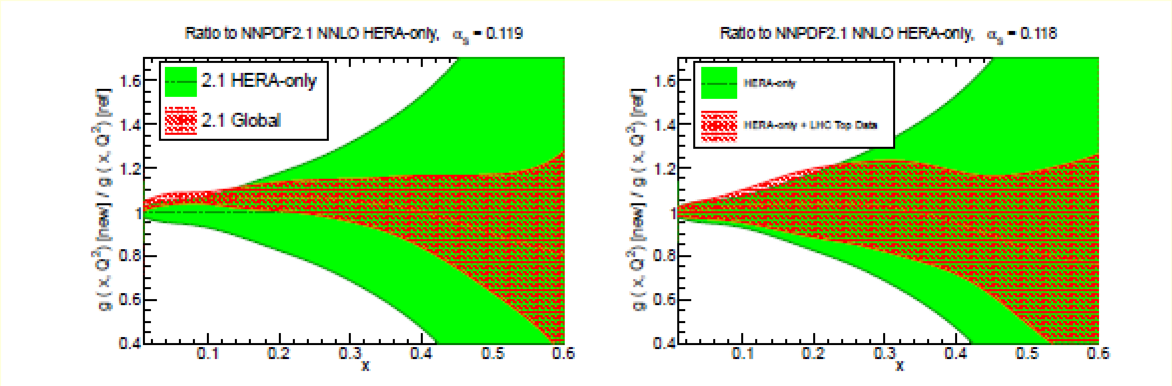

ATLAS+CMS Preliminary σ_tt̄ summary, √s = 8 TeV TOPLHCWG
July 2014
NNLO+NNLL (top++ 2.0), PDF4LHC
m_top = 172.5 GeV
scale uncertainty
scale ⊕ PDF ⊕ α_s uncertainty
stat. uncertainty
total uncertainty
σ_tt̄ ±(stat) ±(syst) ±(lumi)
ATLAS prel., e/μ+jets
ATLAS-CONF-2012-149, L_int=5.8 fb⁻¹
241 ± 2 ± 31 ± 9 pb
ATLAS, dilepton eμ, b-tag
arXiv:1406.5375, L_int=20.3 fb⁻¹
242.4 ± 1.7 ± 5.5 ± 7.5 pb
CMS prel., e/μ+jets
CMS-PAS TOP-12-006, L_int=2.8 fb⁻¹
228 ± 9 +29 −26 ± 10 pb
CMS, dilepton
JHEP 02 (2014) 024, L_int=5.3 fb⁻¹
239 ± 2 ± 11 ± 6 pb
CMS prel. e/μ+τ_h
CMS-PAS TOP-12-026, L_int=19.6 fb⁻¹
257 ± 3 ± 24 ± 7 pb
Effect of LHC beam energy uncertainty: 4.2 pb
150
200
250
300
350
σ_tt̄ [pb]

t-tbar production can already improve the gluon PDF?- NNLO total cross sections

Measurement	$\sigma_{t\bar{t}}$ (pb)	stat. (pb)	sys. (pb)	lumi. (pb)	total (pb)
Tevatron CDF+D0 (Ref. [47])	7.65	$\pm$ 0.20	$\pm$ 0.29	$\pm$ 0.22	7.65 $\pm$ 0.42 (5.5%)
Atlas 7 TeV (Ref. [48])	177	$\pm$ 3	$^{+8}_{-7}$	$\pm$ 7	177^{+10}_{-11} $\left(^{+5.6\%}_{-6.2\%}\right)$
CMS 7 TeV (Ref. [49])	160.9	$\pm$ 2.5	$^{+5.1}_{-5.0}$	$\pm$ 3.6	160.9 $\pm$ 6.6 (4.0%)
Atlas 8 TeV (Ref. [50])	241	$\pm$ 2	$\pm$ 31	$\pm$ 9	241 $\pm$ 32 (13.0%)
CMS 8 TeV (Refs. [51, 52])	227	$\pm$ 3	$\pm$ 11	$\pm$ 10	227 $\pm$ 15 (6.7%)

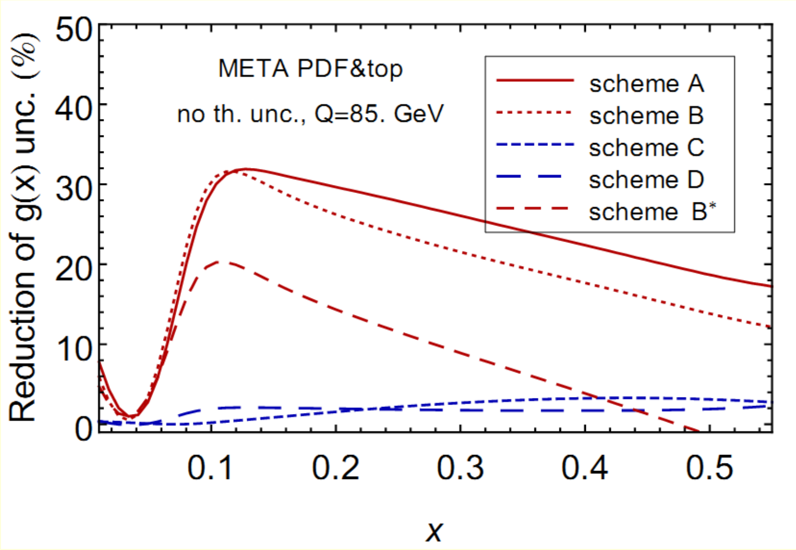
Czakon, Mangano, Mitov, Rojo arXiV:1303.7215
Uses just the top cross section data and an NNLO top calculation





Definitions of 68% c.l. error:
Red lines: $\Delta\chi^2_{\mathrm{global}} = 1$, various reweighting schemes

Short-dashed blue lines: $\Delta\chi^2_{\mathrm{global}} = 100$
Long-dashed blue line: MSTW-like



Not all PDF groups agree that such an improvement can be achieved. (Gao, Nadolsky Les Houches 2013)
It depends on exactly how the reweighting which is used to include the new data is done