Experimental inputs to PDF fits

Amanda Cooper-Sarkar, Oxford University
on behalf of ATLAS and CMS

Oct 31st 2017 – HE and HL-LHC meeting

Inputs to current global PDF fits

- W,Z Drell-Yan
- Jets
- Top-antitop

-------and combinations -- like ratios t-tbar/Z, or ratios at different c.m energies

Further ideas

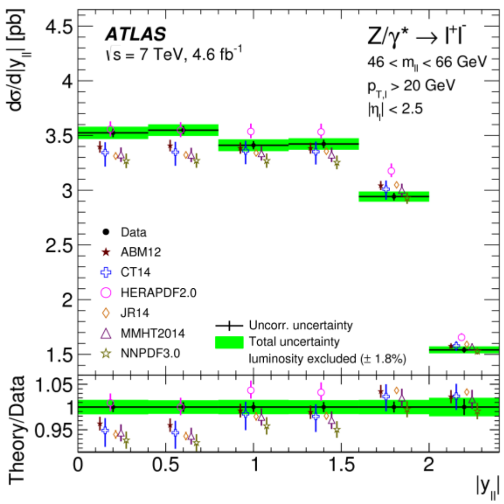
- Boson+jet, boson pt
- Boson +heavy flavour
- Direct photon
- Open charm and beauty

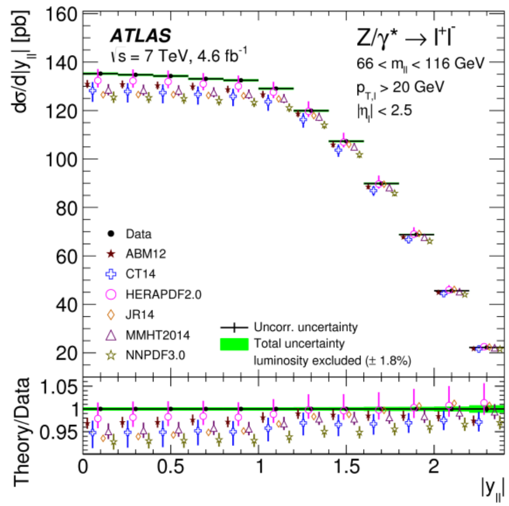
Exp.	Obs.	Ref.	$N_{\rm dat}$	Kin$_1$	Kin$_2$ (GeV)	Theory
ATLAS	W, Z 2010	[49]	30 (30/30)	$0 \le \lvert\eta_l\rvert \le 3.2$	$Q = M_W, M_Z$	MCFM+FEWZ
	W, Z 2011 (*)	[72]	34 (34/34)	$0 \le \lvert\eta_l\rvert \le 2.3$	$Q = M_W, M_Z$	MCFM+FEWZ
	high-mass DY 2011	[50]	11 (5/5)	$0 \le \lvert\eta_l\rvert \le 2.1$	$116 \le M_{ll} \le 1500$	MCFM+FEWZ
	low-mass DY 2011 (*)	[77]	6 (4/6)	$0 \le \lvert\eta_l\rvert \le 2.1$	$14 \le M_{ll} \le 56$	MCFM+FEWZ
	[Z p_T 7 TeV (p_T^Z, y_Z)] (*)	[78]	64 (39/39)	$0 \le \lvert y_Z\rvert \le 2.5$	$30 \le p_T^Z \le 300$	MCFM+NNLO
	Z p_T 8 TeV (p_T^Z, M_{ll}) (*)	[71]	64 (44/44)	$12 \le M_{ll} \le 150$ GeV	$30 \le p_T^Z \le 900$	MCFM+NNLO
	Z p_T 8 TeV (p_T^Z, y_Z) (*)	[71]	120 (48/48)	$0.0 \le \lvert y_Z\rvert \le 2.4$	$30 \le p_T^Z \le 150$	MCFM+NNLO
	7 TeV jets 2010	[57]	90 (90/90)	$0 \le \lvert y^{\rm jet}\rvert \le 4.4$	$25 \le p_T^{\rm jet} \le 1350$	NLOjet++
	2.76 TeV jets	[58]	59 (59/59)	$0 \le \lvert y^{\rm jet}\rvert \le 4.4$	$20 \le p_T^{\rm jet} \le 200$	NLOjet++
	7 TeV jets 2011 (*)	[76]	140 (31/31)	$0 \le \lvert y^{\rm jet}\rvert \le 0.5$	$108 \le p_T^{\rm jet} \le 1760$	NLOjet++
	$\sigma_{\rm tot}(t\bar{t})$	[74, 75]	3 (3/3)	-	$Q = m_t$	top++
	$(1/\sigma_{t\bar{t}})d\sigma(t\bar{t})/y_t$ (*)	[73]	10 (10/10)	$0 < \lvert y_t\rvert < 2.5$	$Q = m_t$	Sherpa+NNLO
CMS	W electron asy	[52]	11 (11/11)	$0 \le \lvert\eta_e\rvert \le 2.4$	$Q = M_W$	MCFM+FEWZ
	W muon asy	[53]	11 (11/11)	$0 \le \lvert\eta_\mu\rvert \le 2.4$	$Q = M_W$	MCFM+FEWZ
	$W + c$ total	[60]	5 (5/0)	$0 \le \lvert\eta_l\rvert \le 2.1$	$Q = M_W$	MCFM
	$W + c$ ratio	[60]	5 (5/0)	$0 \le \lvert\eta_l\rvert \le 2.1$	$Q = M_W$	MCFM
	2D DY 2011 7 TeV	[54]	124 (88/110)	$0 \le \lvert\eta_{ll}\rvert \le 2.2$	$20 \le M_{ll} \le 200$	MCFM+FEWZ
	[2D DY 2012 8 TeV]	[84]	124 (108/108)	$0 \le \lvert\eta_{ll}\rvert \le 2.4$	$20 \le M_{ll} \le 1200$	MCFM+FEWZ
	$W^\pm$ rap 8 TeV (*)	[79]	22 (22/22)	$0 \le \lvert\eta_l\rvert \le 2.3$	$Q = M_W$	MCFM+FEWZ
	Z p_T 8 TeV (*)	[83]	50 (28/28)	$0.0 \le \lvert y_Z\rvert \le 1.6$	$30 \le p_T^Z \le 170$	MCFM+NNLO
	7 TeV jets 2011	[59]	133 (133/133)	$0 \le \lvert y^{\rm jet}\rvert \le 2.5$	$114 \le p_T^{\rm jet} \le 2116$	NLOjet++
	2.76 TeV jets (*)	[80]	81 (81/81)	$0 \le \lvert y_{\rm jet}\rvert \le 2.8$	$80 \le p_T^{\rm jet} \le 570$	NLOjet++
	$\sigma_{\rm tot}(t\bar{t})$	[82, 88]	3 (3/3)	-	$Q = m_t$	top++
	$(1/\sigma_{t\bar{t}})d\sigma(t\bar{t})/y_{t\bar{t}}$ (*)	[81]	10 (10/10)	$-2.1 < y_{t\bar{t}} < 2.1$	$Q = m_t$	Sherpa+NNLO
LHCb	Z rapidity 940 pb	[55]	9 (9/9)	$2.0 \le \eta_l \le 4.5$	$Q = M_Z$	MCFM+FEWZ
	$Z \to ee$ rapidity 2 fb	[56]	17 (17/17)	$2.0 \le \eta_l \le 4.5$	$Q = M_Z$	MCFM+FEWZ
	$W, Z \to \mu$ 7 TeV (*)	[85]	33 (33/29)	$2.0 \le \eta_l \le 4.5$	$Q = M_W, M_Z$	MCFM+FEWZ
	$W, Z \to \mu$ 8 TeV (*)	[86]	34 (34/30)	$2.0 \le \eta_l \le 4.5$	$Q = M_W, M_Z$	MCFM+FEWZ

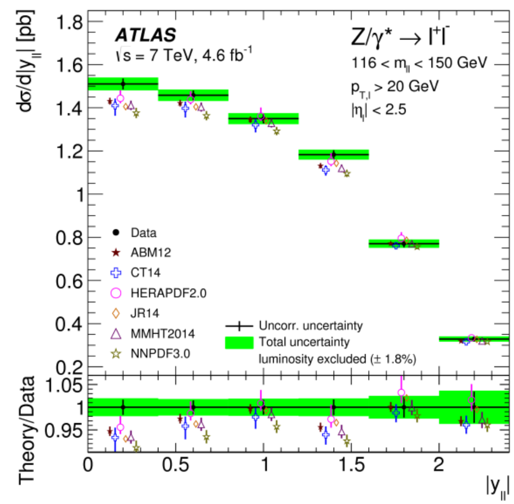
W,Z and Drell-Yan distributions

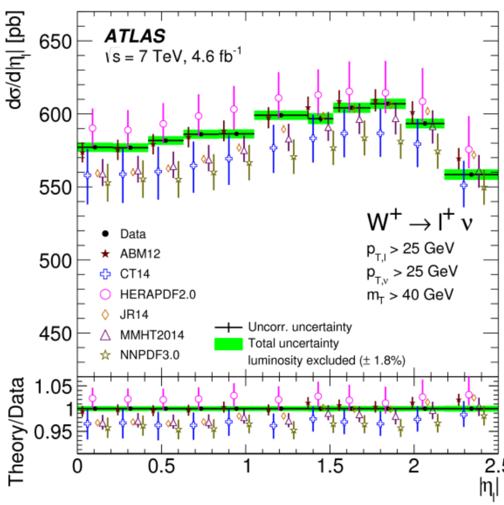
ATLAS inclusive W and Z differential distributions
arXiv:1612.03016
Very high precision

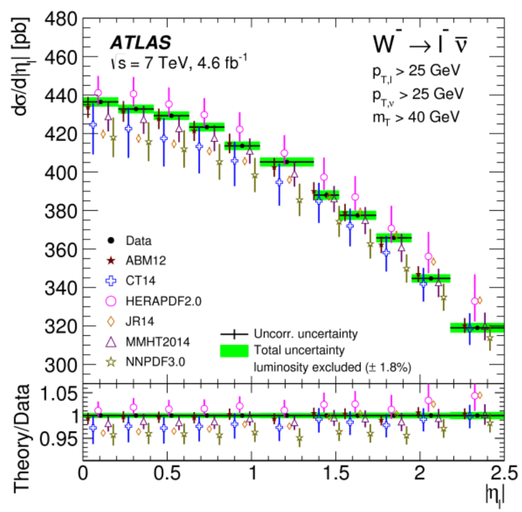
- W: Total (0.6–1.0%), multijet background (0.3–0.7%)
- Z Central: Total (0.4%), reconstruction efficiency (0.2–0.3%)
- Z Forward: Total (2.3%), identification efficiency (1.5%)
- 1.8% luminosity uncertainty



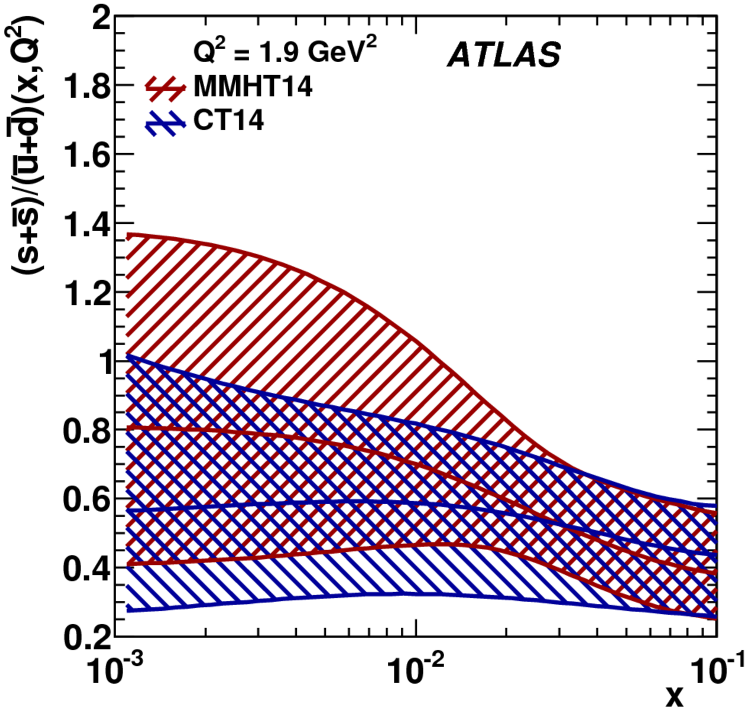


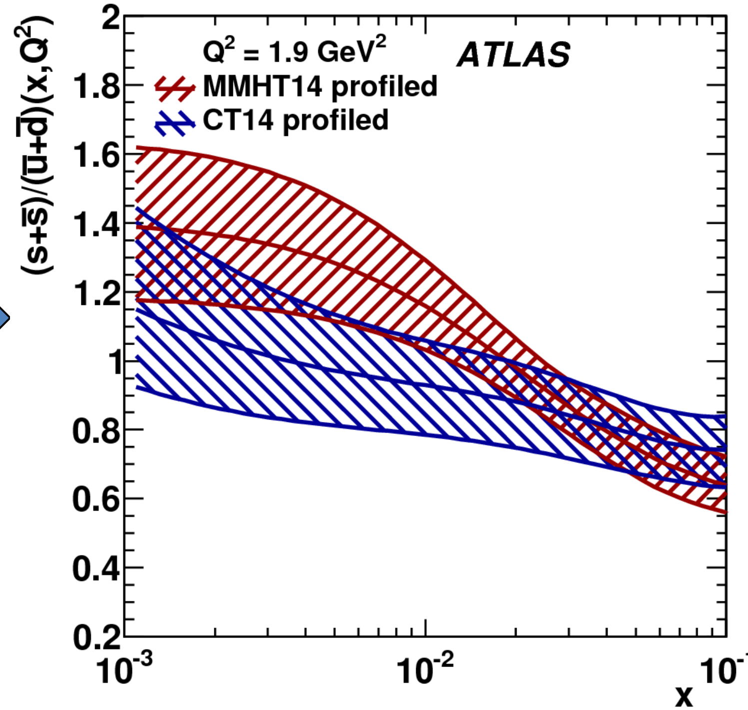


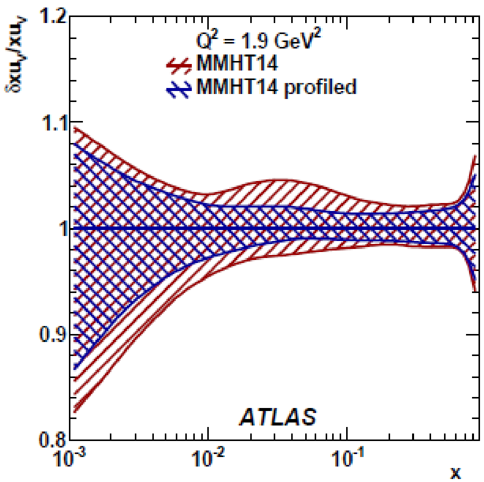


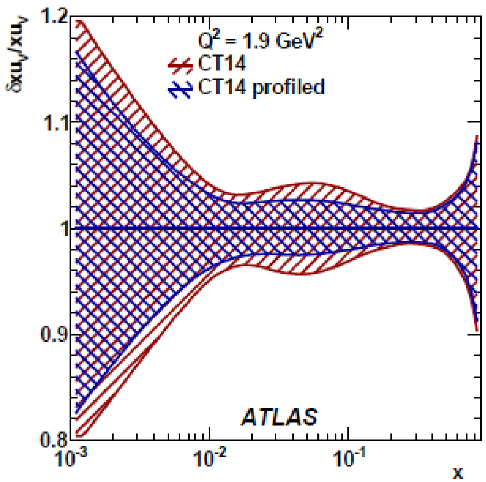


State of the art predictions at NNLO

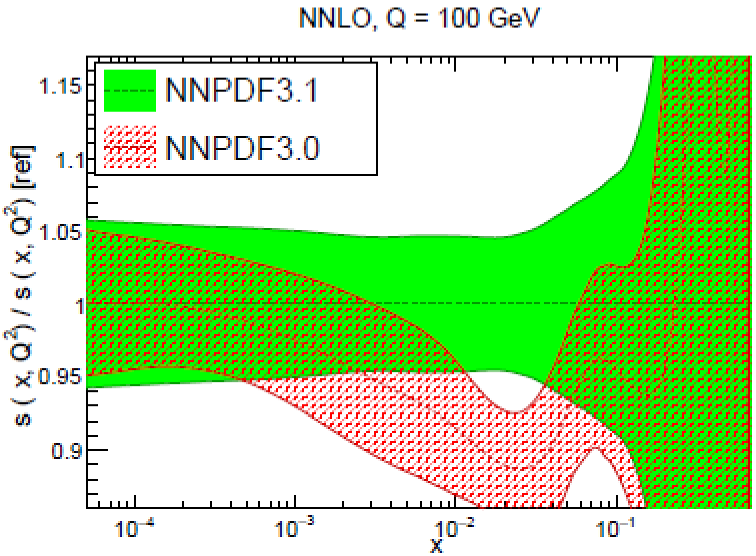


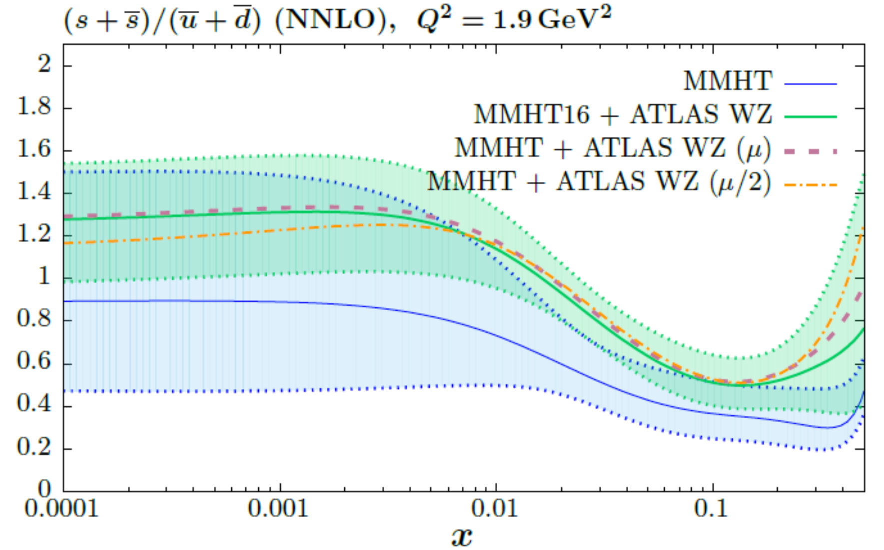






It reduces the uncertainties on the strange sea- as well as pulling up its absolute value at low-x.
It also reduces valence uncertainties





NNPDF and MMHT both see larger strangeness when using ATLAS W,Z data

ATLAS precision W,Z data are compatible with earlier CMS 7 and 8 TeV W data

arXIV: 1312.6283,1603.01803

There is mild tension with CMS 7 TeV double differential Drell-Yan Z/γ* (arXiV:1310.7291)–

BUT these data are also compatible with a higher than conventional strangeness fraction-

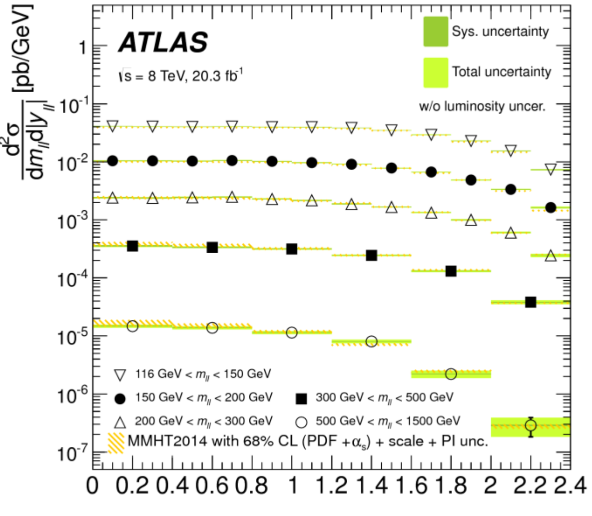
NNPDF 3.1 collider PDFs use HERA+ CMS, ATLAS LHCb and Tevatron data to obtain strange to light quark ratio, Rs=0.82±0. 18 at x=0.023, Q^2=2 GeV^2 , where conventional values have been Rs~0.5

CAN one improve in future?– maybe at high rapidity

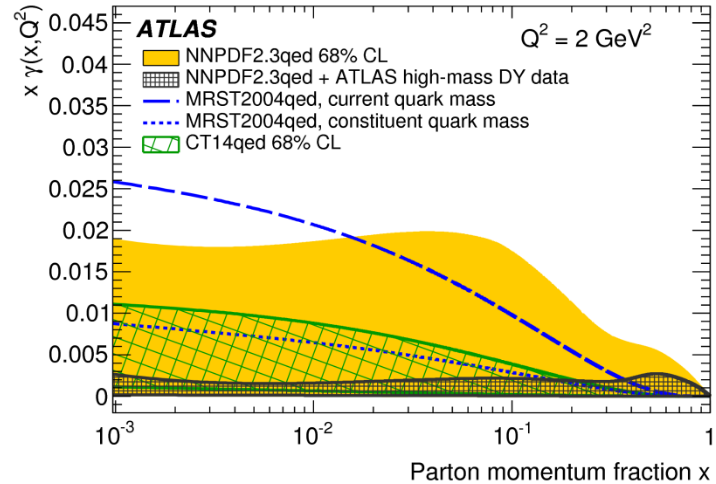
ATLAS peak W,Z data has already reached systematic uncertainties of ~0.5%, experimental improvement unlikely and this is already challenging NNLO calculations-need beyond fixed order when data are cut in pt?

Furthermore, the reach to lower x at 13,14,27 TeV brings more theoretical challenges- need for ln(1/x) resummation- see arXIV:1710.05935

ATLAS 8 TeV high-mass Drell-Yan and the photon PDF arXiv:1606.01736



At high-mass di-lepton pairs may be photon induced rather than true Drell-Yan processes. These data have been used to constrain the photon-PDF



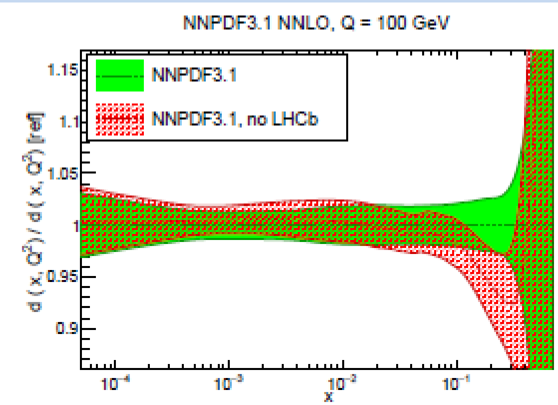
LHCb W,Z data probe a different kinematic region to both lower and higher-x values
Their impact is mostly seen on high-x quarks. Low-x can present theoretical challenges

Off-peak Drell-Yan can still improve both statistically and systematically- and there is greater reach to low and high-x from HE running

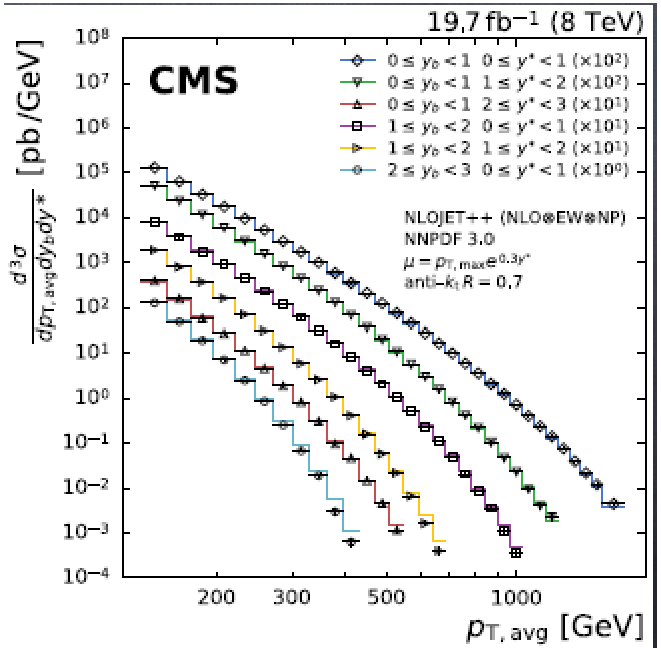
BUT low-mass brings the **low-x theory challenges-ln(1/x)** resummation etc

This also affects the LHCb data, NOTE the low- and high-x regions are of course coupled- both come from high rapidity

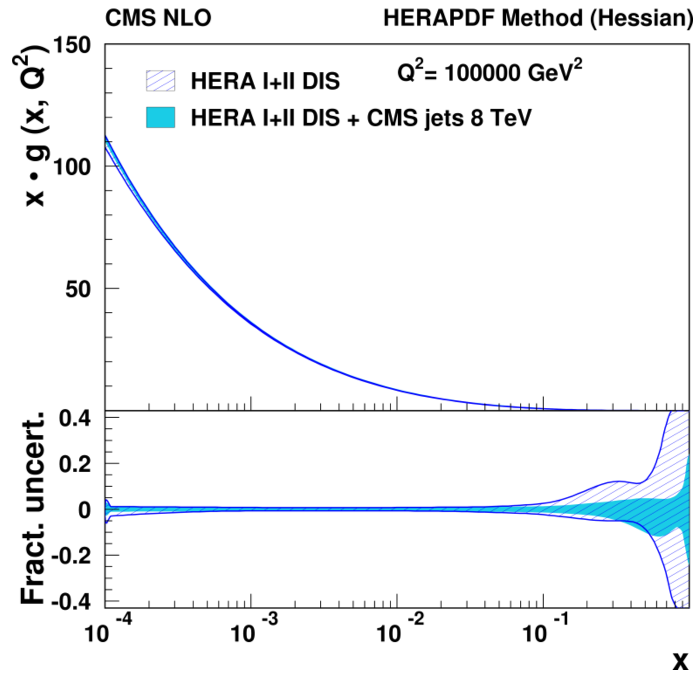
High-mass requires good understanding of the **NLO-EW corrections and photon PDF** (considerable recent progress)



Jet distributions



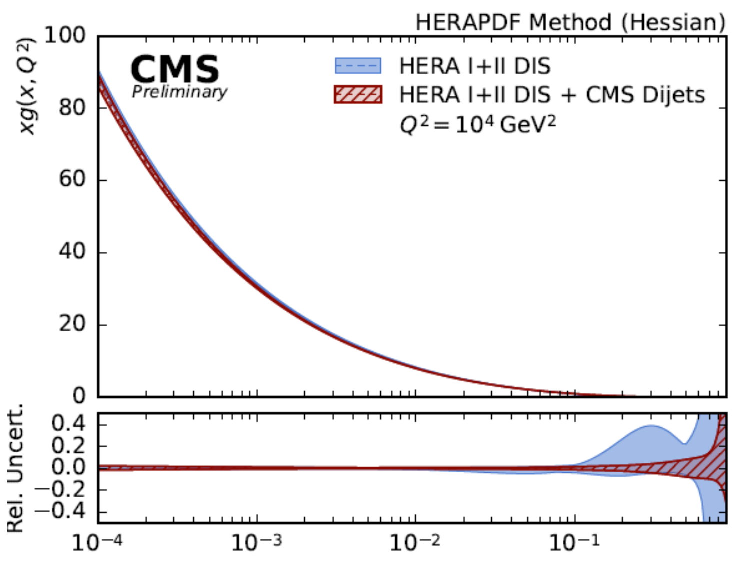
CMS 8 TeV inclusive jet data: arXIV:1609.05331



- **Strong improvement of gluon distribution, precise α_s extraction:**

$$(M_Z) = 0.1185^{+0.0019}_{-0.0021}\,(\mathrm{exp})^{+0.0002}_{-0.0015}\,(\mathrm{model})^{+0.0000}_{-0.0004}(\mathrm{param})^{+0.0022}_{-0.0018}\,(\mathrm{scale})$$

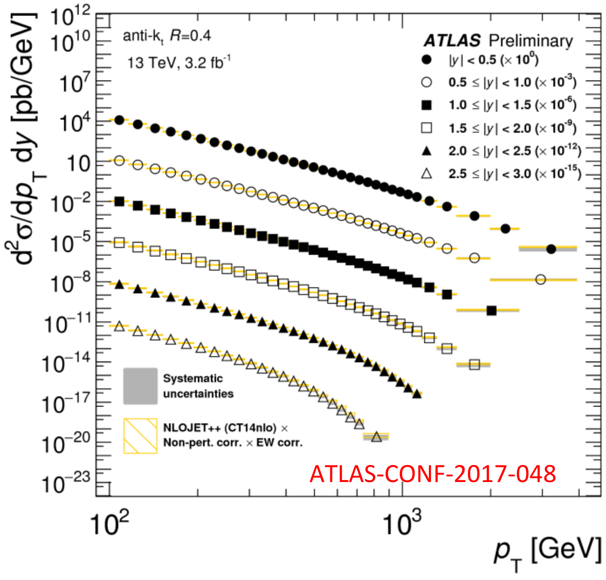
- **Ratios $2.76/8$, $7/8$ available: partial reduction of uncertainties**



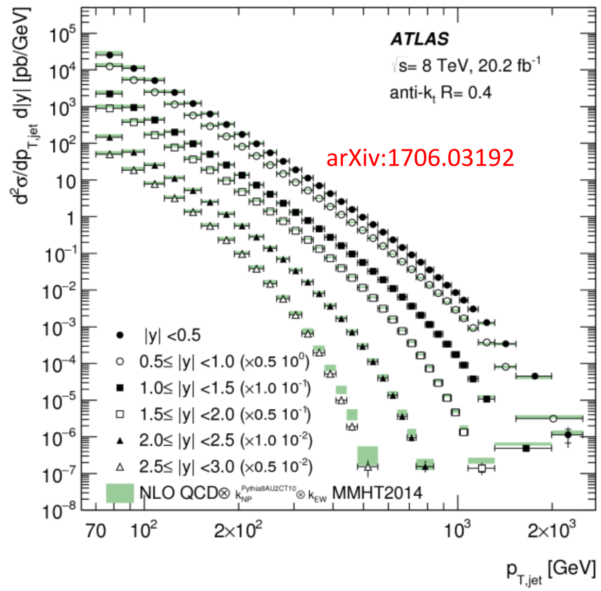
$$\alpha_S(M_Z) = 0.1199 \pm 0.0015\,(\mathrm{exp}) \pm 0.0002\,(\mathrm{mod})\,^{+0.0002}_{-0.0004}\,(\mathrm{par})\,^{+0.0031}_{-0.0019}\,(\mathrm{scale})$$

Ratios of 2.76 and 7 TeV data were already used by ATLAS: arXiv:1304.4739

New ATLAS jet production data at 8 and 13 TeV



(Older at 7 TeV, inclusive dijet,trijet)



State of the art prediction only just becomes NNLO- many studies still at NLO

Large χ2 when fitting different rapidity bins simultaneously for all inclusive jet samples at NLO. This has been found both by ATLAS and by global fitters
Much work on considering realistic de-correlations for 2-point systematics and on alternative scale variations choices and one still obtains χ2/ndp ~ 260/159.

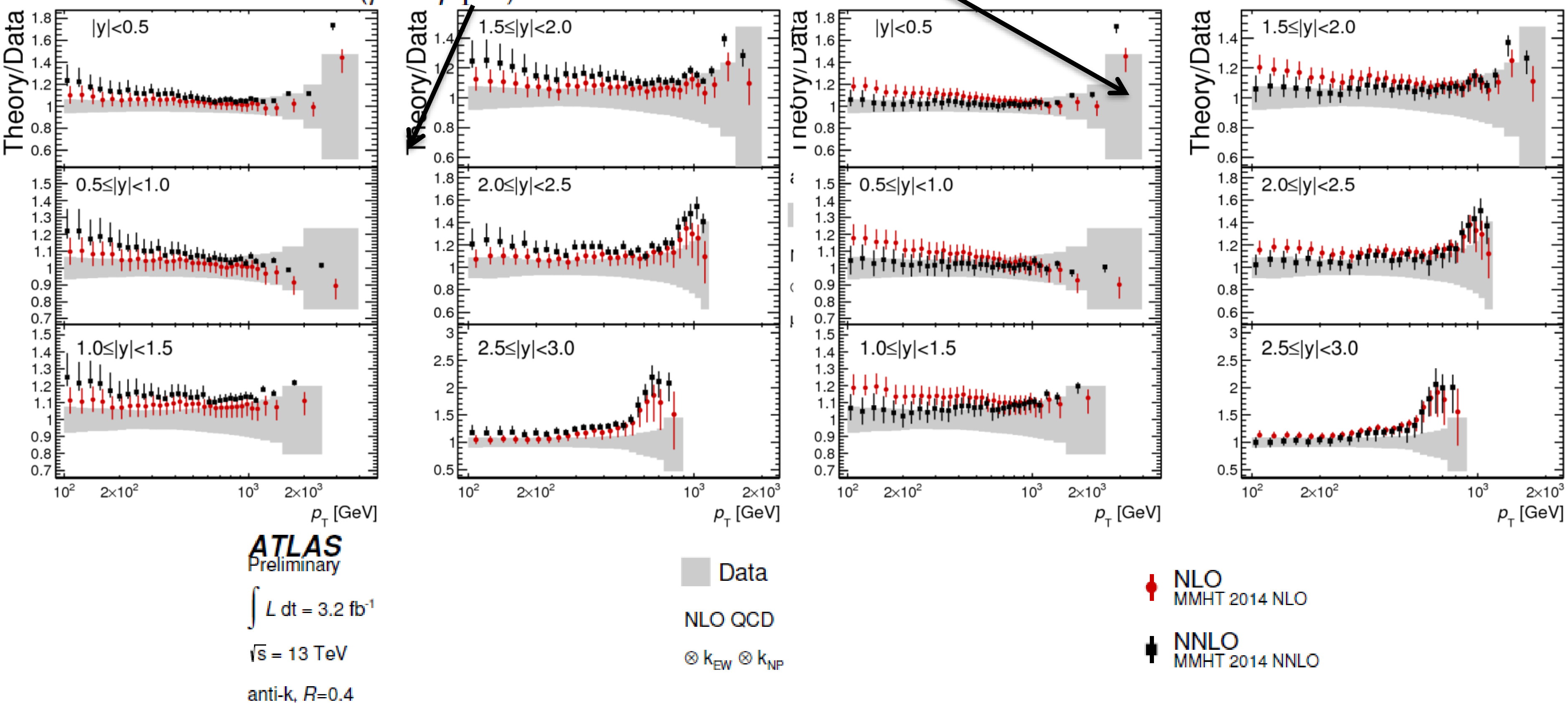
BUT NNLO can describe the data better?....

There is progress on the NNLO corrections- scale choice matters.

P_T^{jet} as the scale choice and larger cone size R=0.6, gives the most compatible results

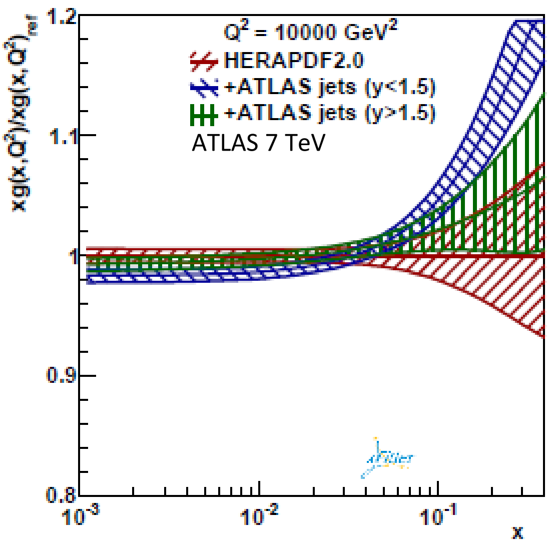
NLO/Data vs NNLO/Data

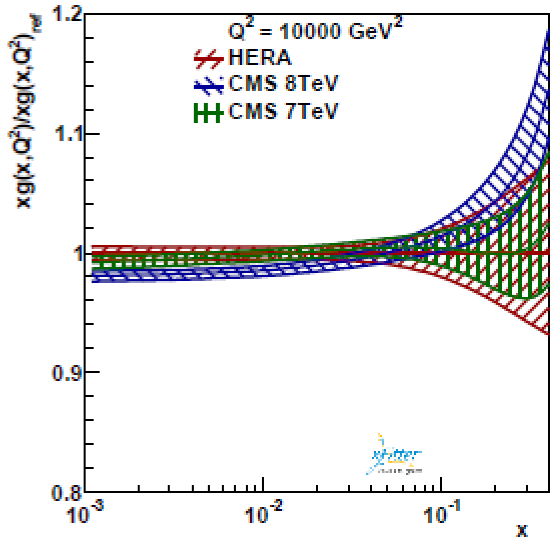
- NNLO ($\mu = p_T^{jet}$) describes the data within uncertainties
  - except at $|y| > 2$ and high-p_T
- NNLO ($\mu = p_T^{max}$) overestimates the measured cross-section

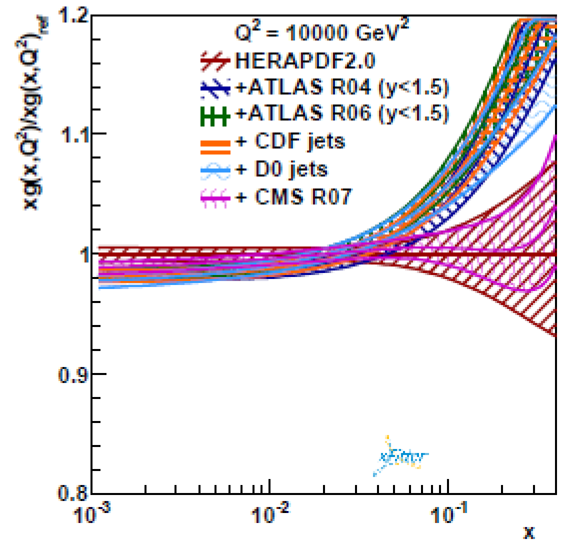


Can one improve in future ?

There is concern about data consistency in jets, both between rapidity bins within ATLAS and between ATLAS and CMS, for which more data is always helpful.





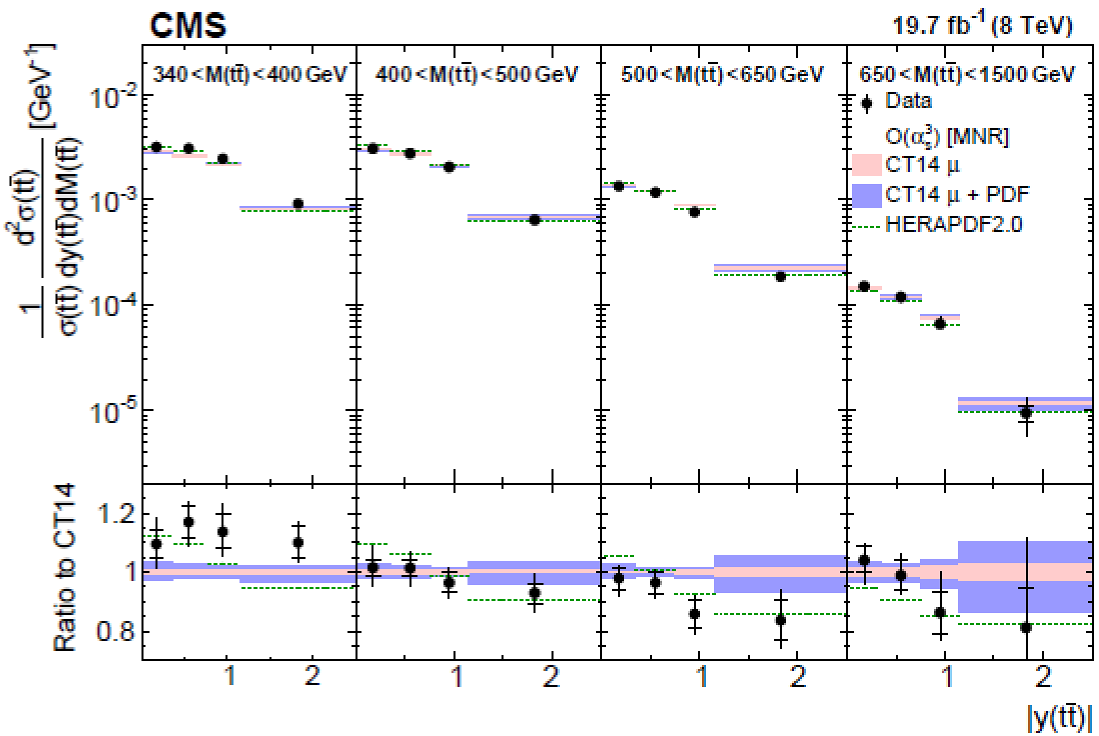


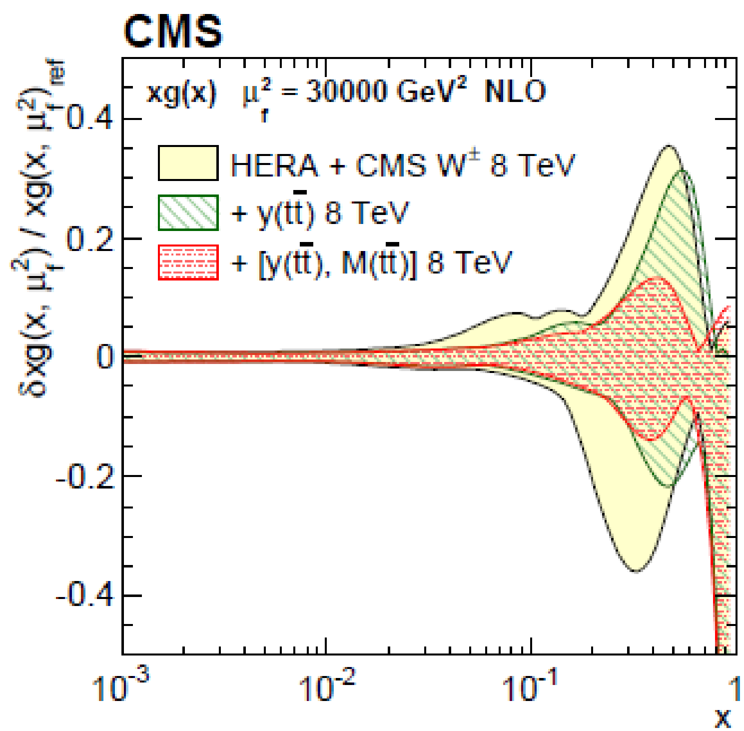
Since jet data do not suffer from lack of statistics this points up the fact that it is data systematic uncertainties which really matter. More data always helps us to improve systematics but it is not easy to quantify this.

NNLO calculations have already improved the description of data, experimentalists would like clarity on scale choice

Top distributions

CMS have recently (arXiv:1703.01630) presented double differential top distributions in mass and rapidity of the t-tbar pair

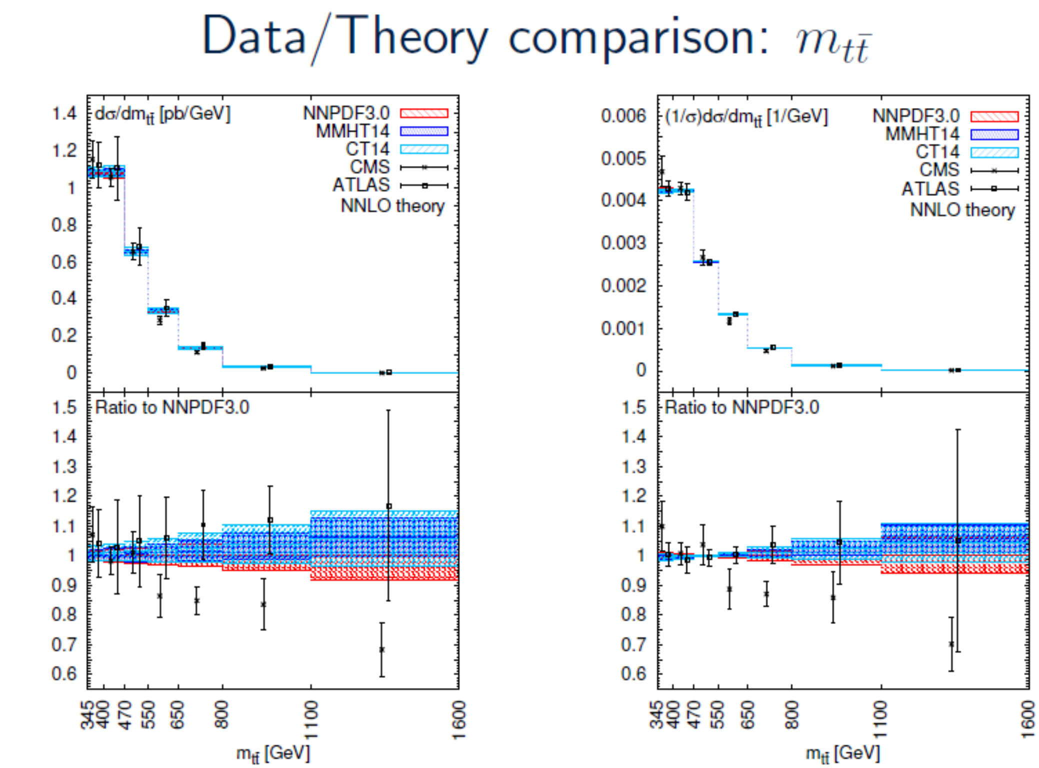




When input to a PDF fit these double differential is much more constraining than the single BUT analysis can only be done at NLO presently since there are no predictions at NNLO for the double differential distributions

CMS top data at 5 TeV (27pb-1) are also coming CMS –PAS TOP-16-023

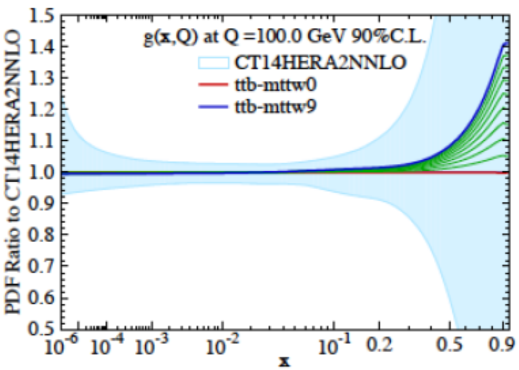
NNLO predictions are now available for ATLAS (1511.04716) and CMS (1505.04480) 8 TeV lepton +jets single differential distributions (arXiv: 1611.08609 and 1704.08551). EW corrections arXiv:1705.04105.

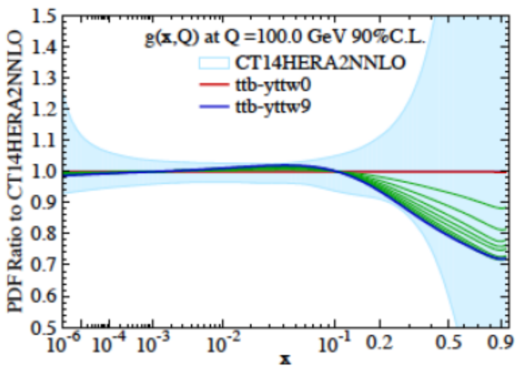


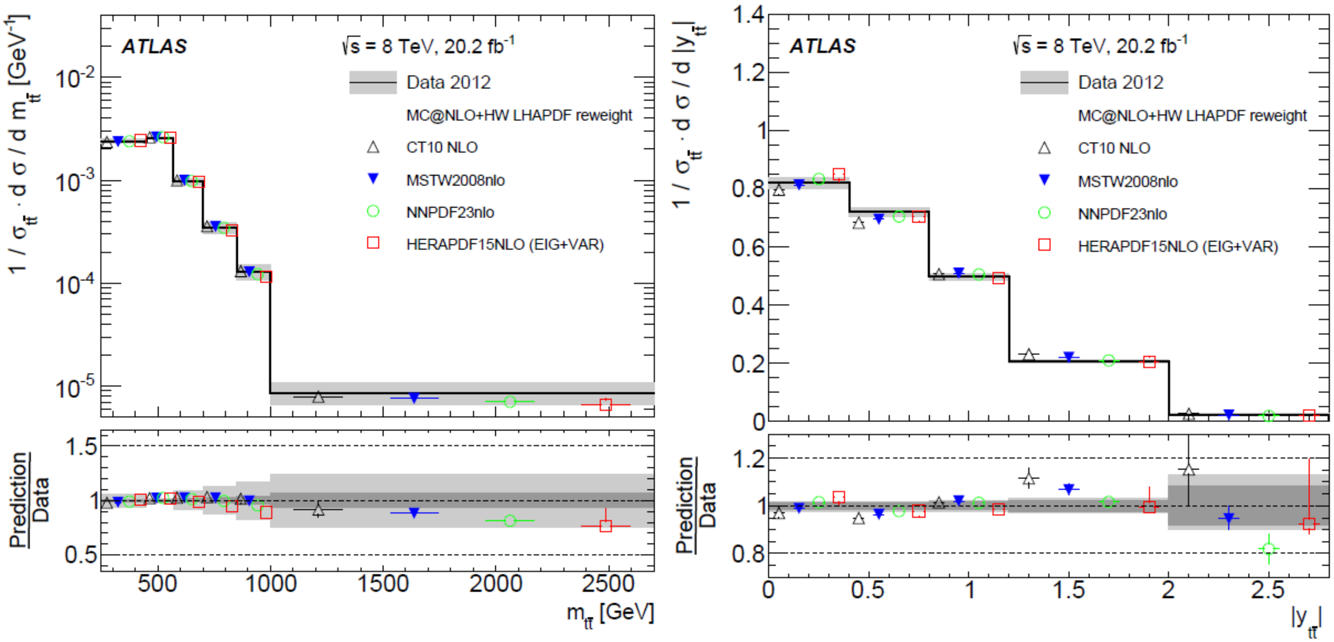
There are several distributions that constrain the high-x gluon:
mass t-tbar, rapidity t-tbar, rapidity-top and pt-top
Both normalised and absolute spectra have been compared to various PDFs

There are some issues:

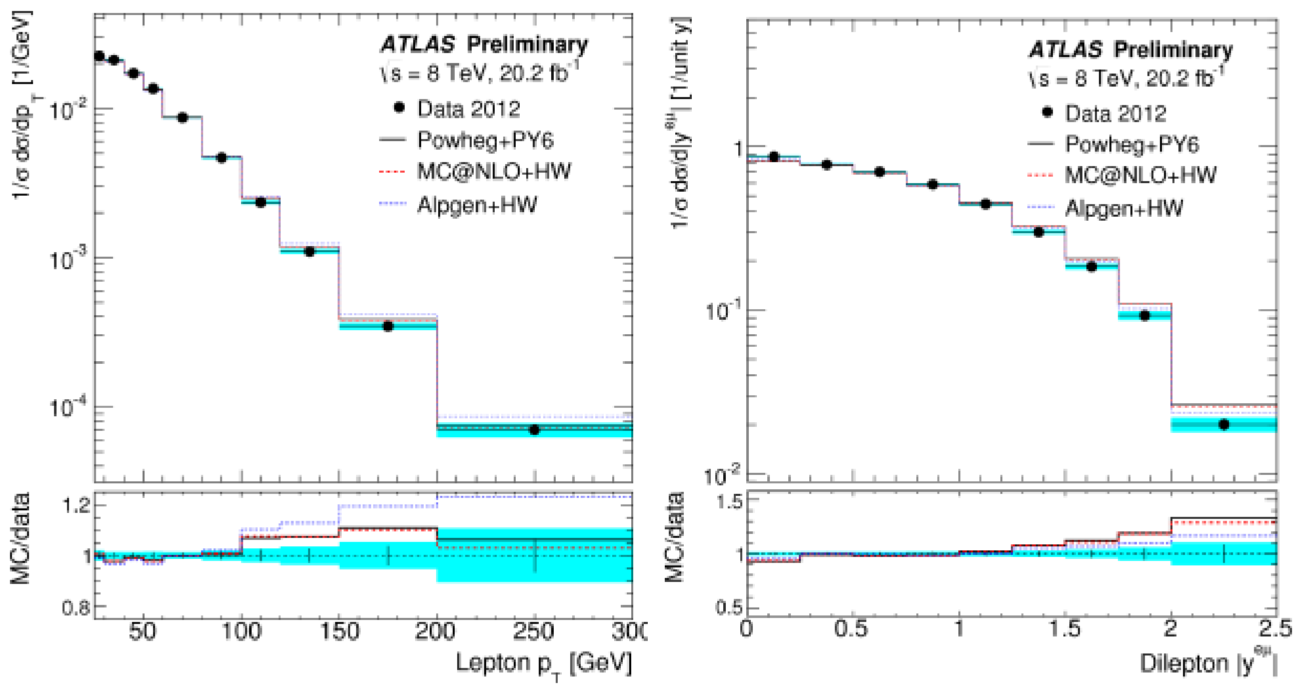
1. The CMS and ATLAS data are not always consistent with each other for the same spectra- and nor are their uncertainty estimates
2. Within the experiments the different spectra are not consistent with each other E.g--- for ATLAS M-tt wants a harder gluon, Y-tt wants a softer gluon CMS data gives similar inconsistencies
3. To fit more than one spectrum at a time one needs statistical correlation matrices







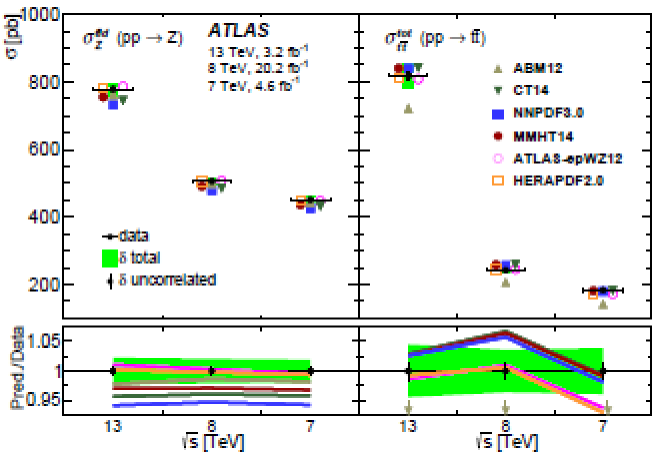
ATLAS has also presented data for the normalised M-tt and Y-tt spectra for the dilepton mode ArXiv:1607.07281 --these can be analysed at NNLO

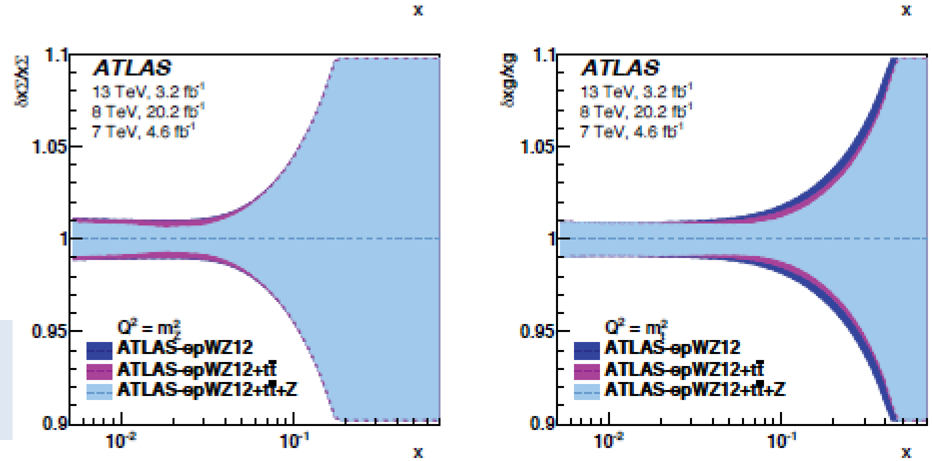


The data in the dilepton channel can also be analysed in terms on the lepton decay variables ATLAS-CONF-2017-044 But so far this can only be analysed at NLO

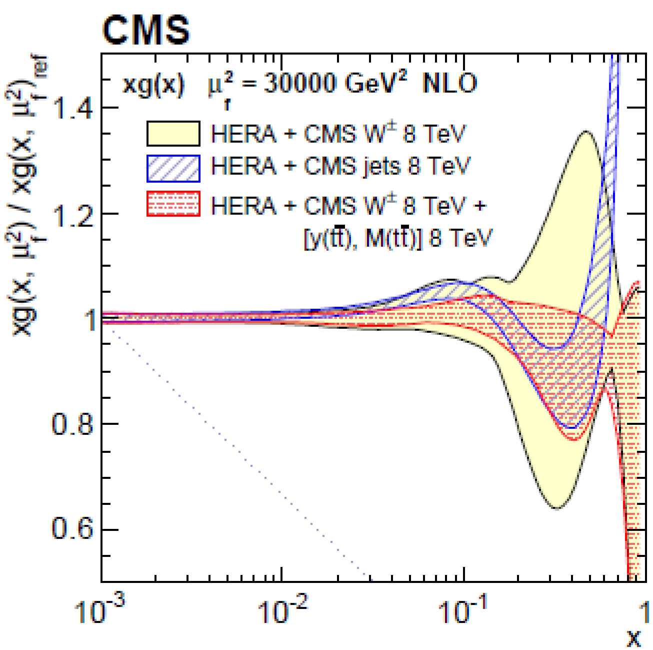
Simultaneous analyses of different data sets:

ATLAS measurement **of inclusive t-tbar to Z cross-sections** at 7, 8 and 13 TeV (arXiv:1612.03636) With accounting for correlations between them





T-tbar data mostly affects the gluon
Z data mostly affects the quarks



CMS analysis of **W, jets and top** arXiv:1703.01630

Can one improve in future?

Top distributions have not yet hit their potential systematic uncertainty limit

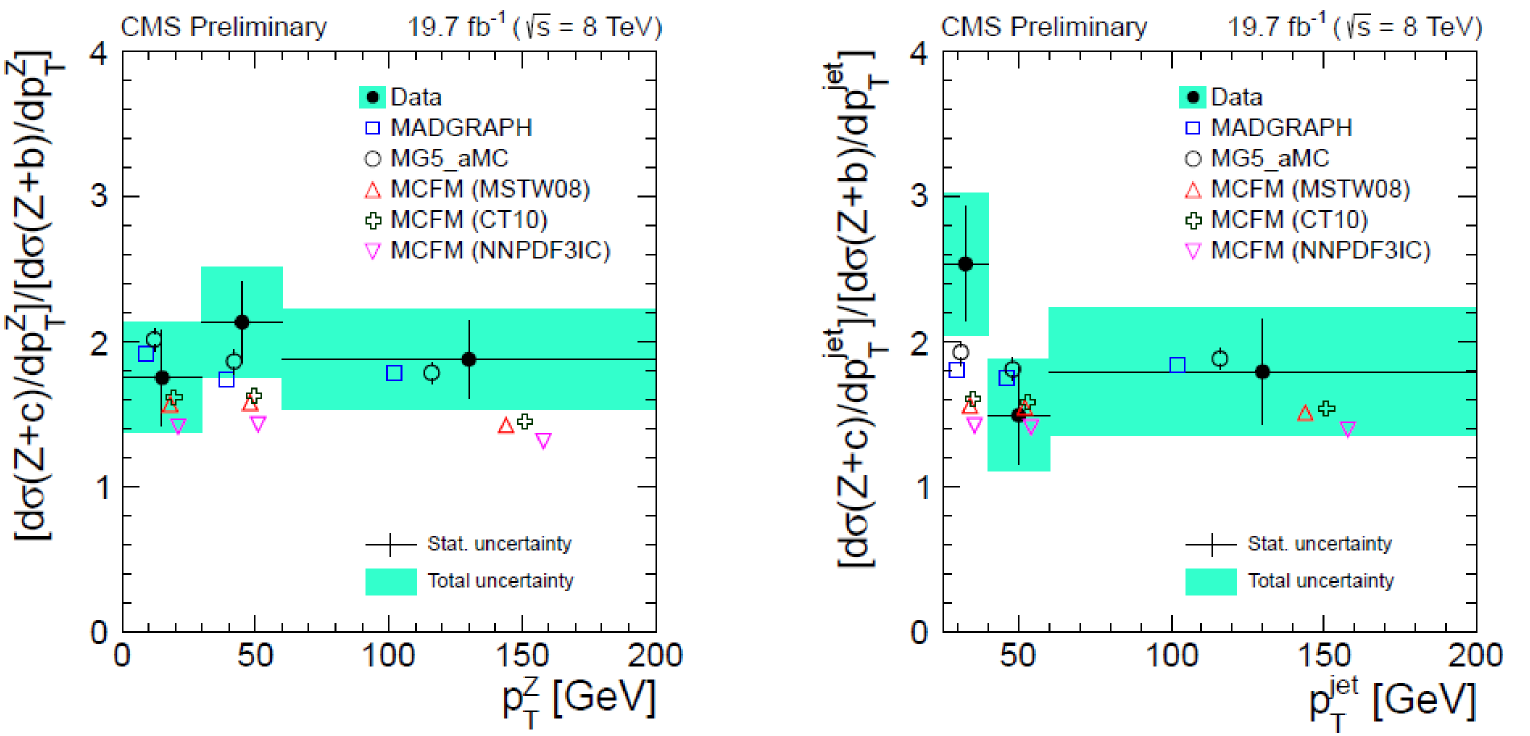
We can have more clever ideas like taking ratios of different quantities

And ratios of different CM energies

Boson (W,Z or γ) +heavy flavour distributions

Measurement of associated Z + charm production [CMS-PAS-SMP-15-009]

- Measurement at 8 TeV, $L = 19.7\ \text{fb}^{-1}$
- Cross section of $Z + c$ and ratio $Z + c/Z + b$ as function of p_T
- **Important for searches beyond SM, sensitive to possible intrinsic charm**

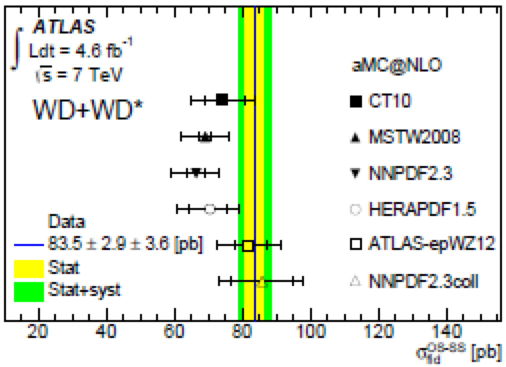


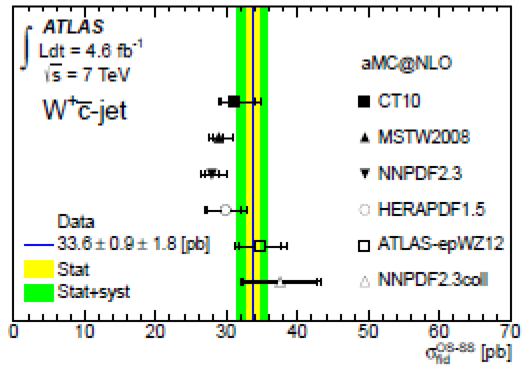
Z +c data is not yet very discriminating

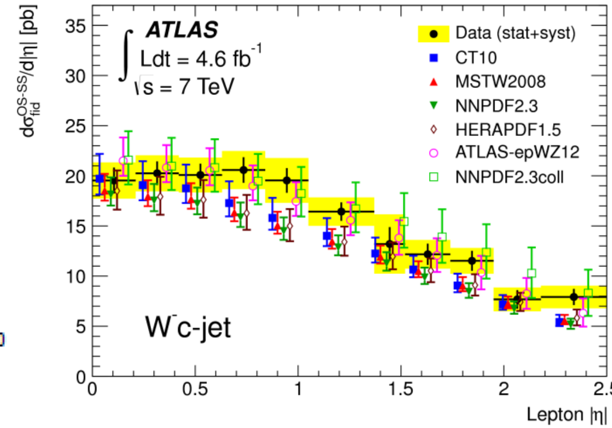
There is also VERY RECENTLY γ+c/b – ATLAS arXIV:1710.0 0560 which favours a 5-flavour scheme vas 4-flavour. However it is not discriminating against different intrinsic charm models, see back-up

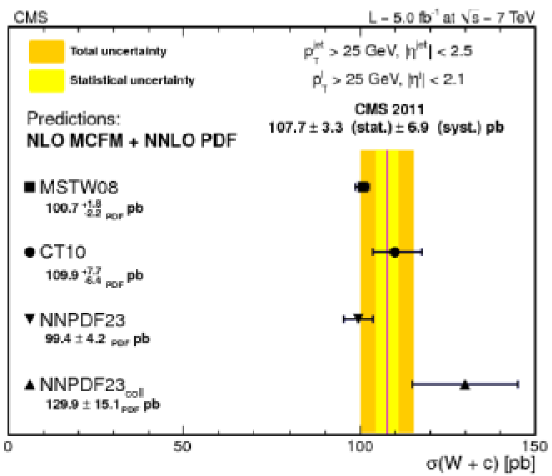
Measurements of W+c from ATLAS and CMS

arXiv:1402.6263

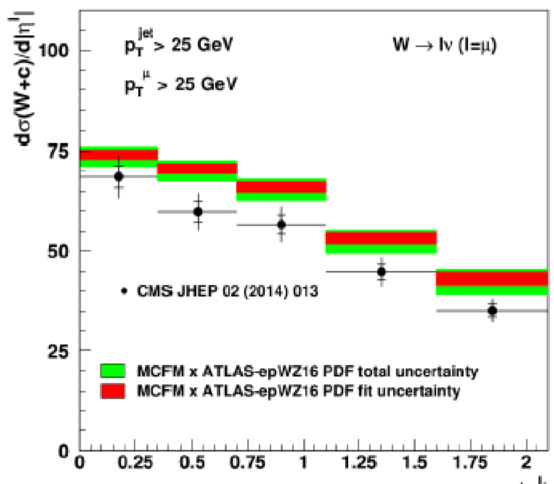








arXiv:1310.1138

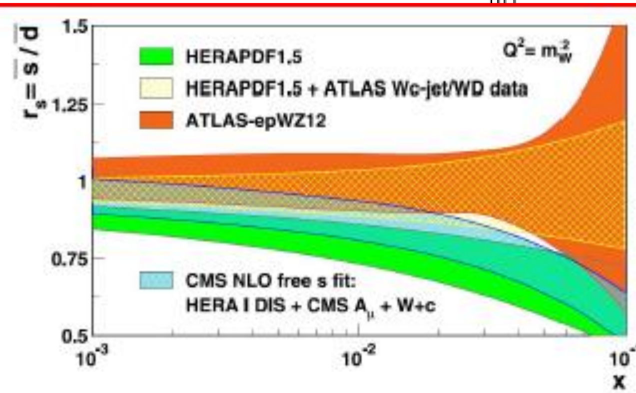


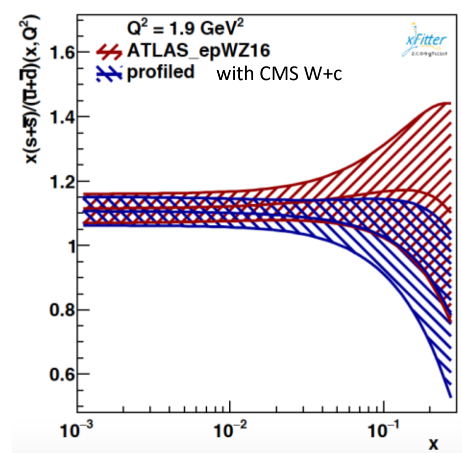
ATLAS data agrees with PDFs which have unsuppressed strangeness

CMS data has a smaller cross section and less strangeness

BUT CMS data still implies larger strangeness than the conventional suppression

NEW DATA is coming from both collaborations



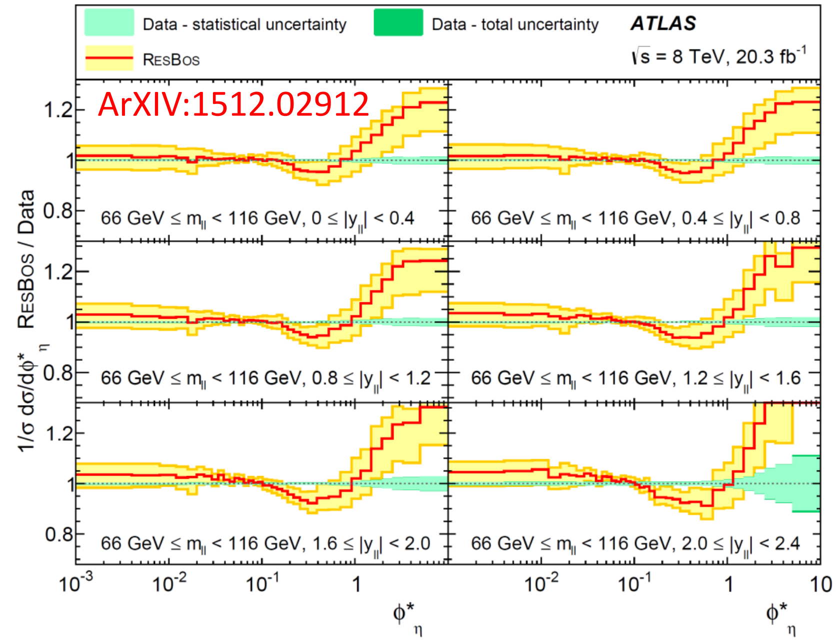


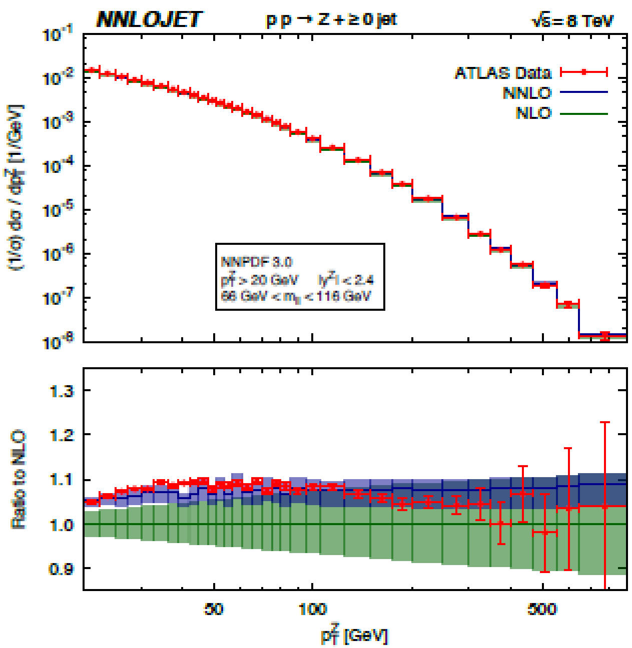
Boson (W,Z) +jets distributions

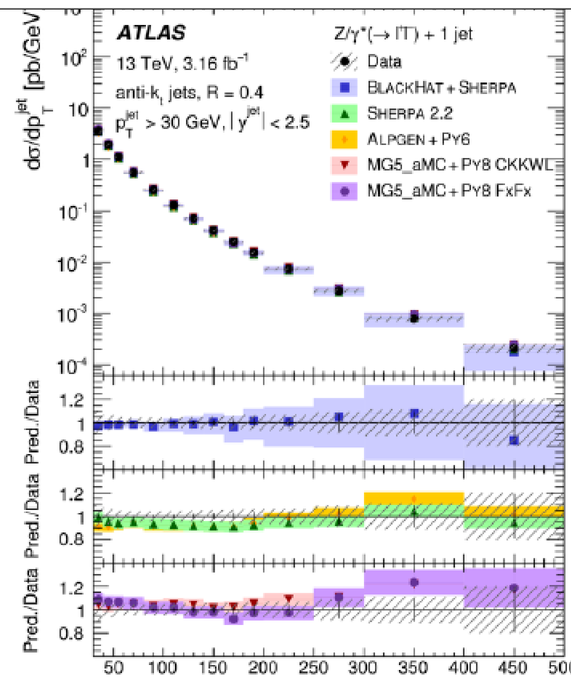
There are now NNLO predictions for Z +jets, Zpt and W+jets arXIV: 1607.01749, 1605.04295

There is new data– and more in the pipeline.

The data on Zpt or ZΦ* is very accurate –and have stimulated these developments, which even aim to cover quite low pt – impact on fits is not large so far



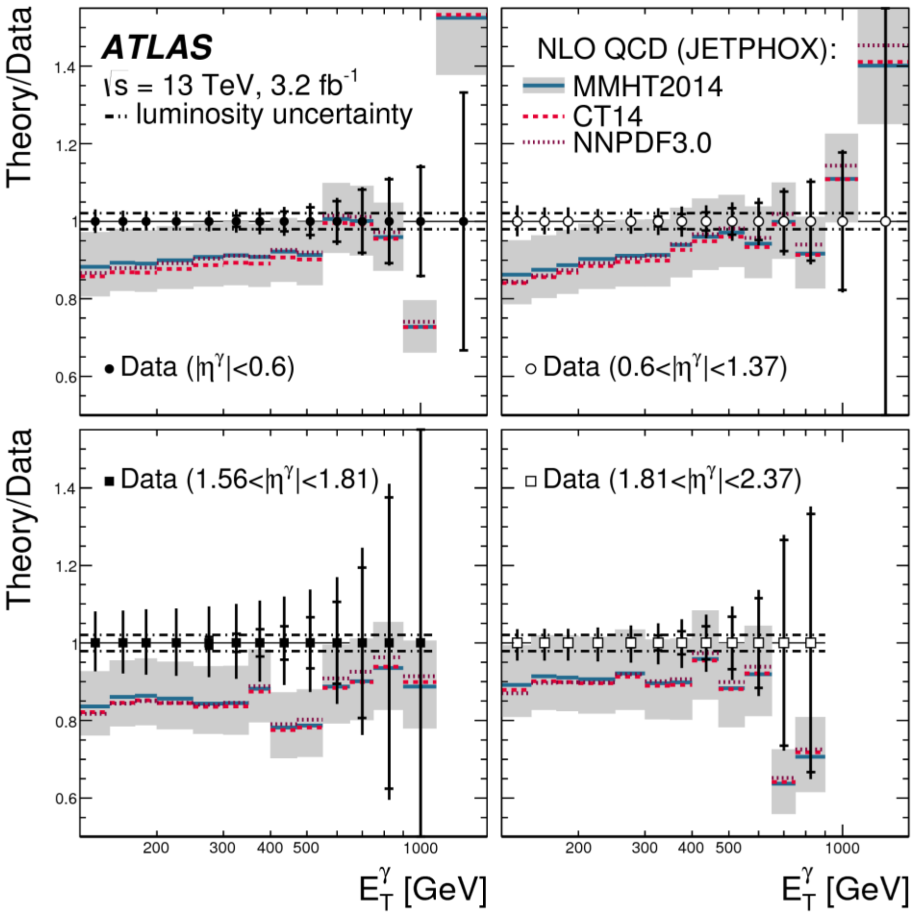




Eur. Phys. J. C77 (2017) 361

The data on Z+jets and W+jets is much less accurate and can improve in future

Direct photon production



There is new data at 8 TeV arXiV:1704.03839 and 13 TeV arXiv: 1701.06882

Direct photon data were abandoned in PDF fits more than 15 years ago due to lack of theoretical understanding. It has now been established that at collider energies these data can give useful information on the gluon

Studies at NLO have been done, but there are now NNLO predictions arXiv: 1701.06882

Summary: where can we improve in future?

- **W,Z and Drell-Yan distributions** – sensitivity to valence quarks, strangeness, photon PDF

ATLAS peak W,Z data has already reached systematic uncertainties of ~0.5%, experimental improvement unlikely and this is already challenging NNLO calculations

The reach to lower x at 13,14,27TeV brings more theoretical challenges- need for ln(1/x) resummation- see arXIV:1710.05935

Off-peak Drell-Yan can still improve BUT low-mass brings the same low-x challenges.

This also affects the LHCb data

And high-mass requires good understanding of the NLO-EW corrections and photon PDF

- **Inclusive, di-jet and tri-jet distributions**------sensitivity to gluon

Already challenging theoretical understanding -NNLO is needed but scale choice is still an issue

- **Top-antitop distributions** –sensitivity to gluon

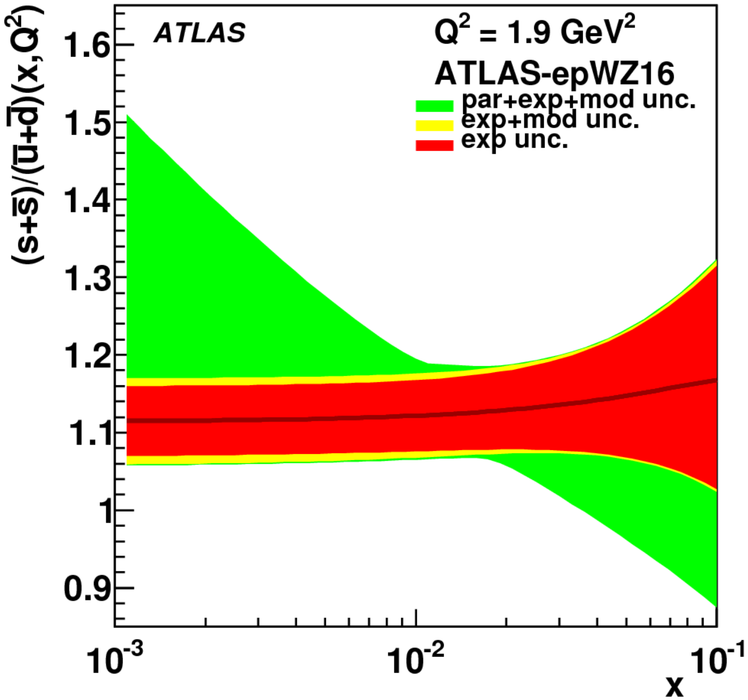
NNLO calculations already required, data can also improve (data consistency?)

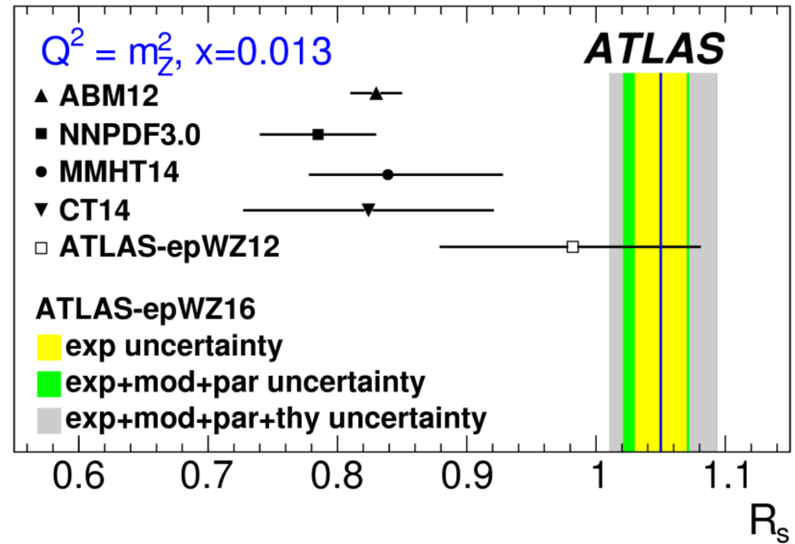
Combinations of types of data and different beam energies –accounting for their correlations- can help

For all of these below: precision of the data can improve

- **W,Z +jets** --------sensitivity to gluon- so far limited, can improve
- **W,Z/γ +heavy flavour** -sensitivity to strangeness and intrinsic charm- can improve
- **Direct photon**-------sensitivity to gluon—studies needed

Back-up





ATLAS new W,Z data are compatible with earlier CMS 7 and 8 TeV W data
There is mild tension with CMS 7 TeV double differential Drell-Yan– these data also favour a higher than conventional strangeness fraction.

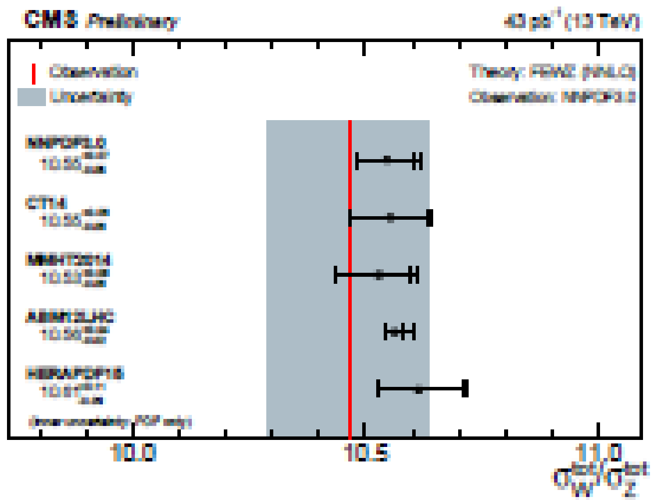
PDF set	$R_s(0.023, 2\ \text{GeV}^2)$	$R_s(0.013, M_Z^2)$
NNPDF3.0	0.47±0.09	0.79±0.04
NNPDF3.1	0.61±0.14	0.83±0.06
NNPDF3.1 collider-only	0.85±0.16	0.93±0.06
NNPDF3.1 HERA + ATLAS W, Z	0.96±0.20	0.98±0.09
ATLAS W, Z 2010 HERAfitter (Ref. [100])	$1.00^{+0.25}_{-0.28}$ (*)	$1.00^{+0.09}_{-0.10}$ (*)
ATLAS W, Z 2011 xFitter (Ref. [72])	$1.13^{+0.11}_{-0.11}$	1.05 ± 0.04

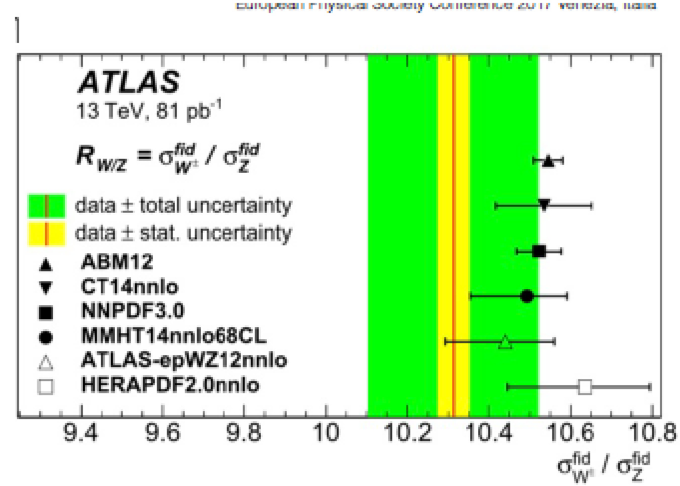
PRELIMINARY

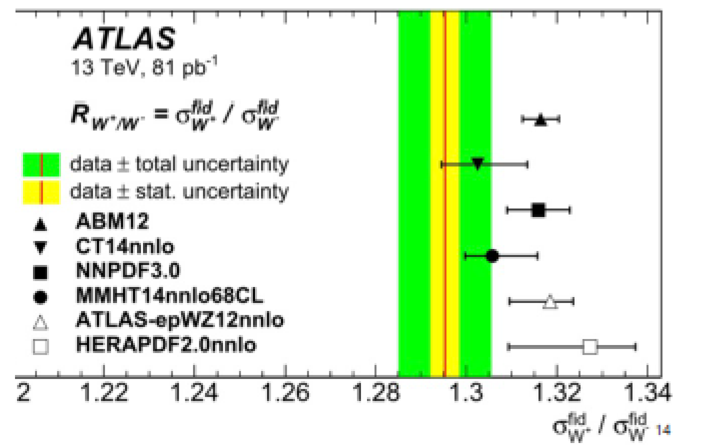
Collider data means CMS ATLAS and Tevatron data-

There is also 13 TeV data from 2015/2016

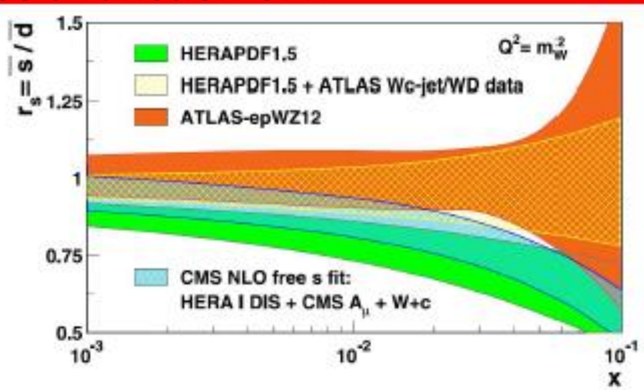
W/Z ratios are lower than most predictions – as you would expect if more strangeness is needed



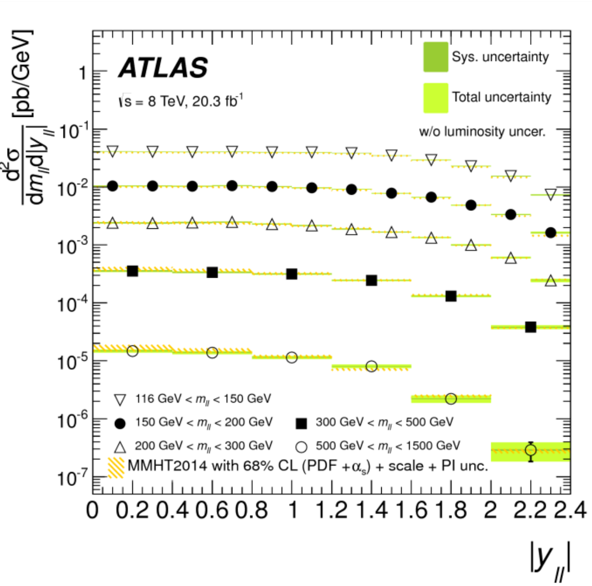




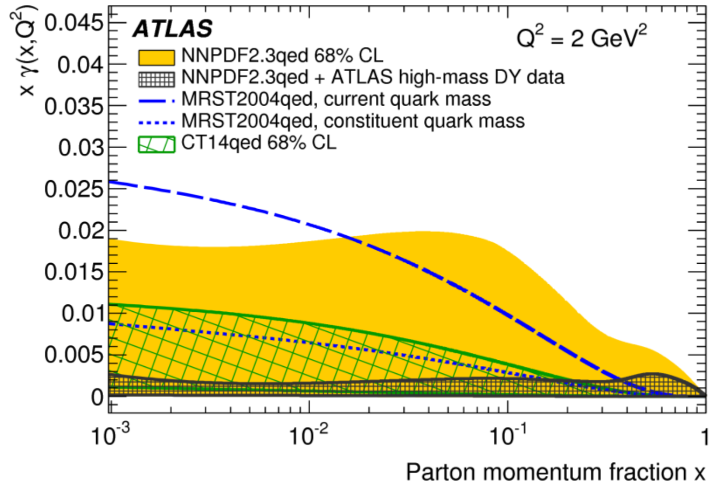
ATLAS and CMS strangeness ratio as a function of x



ATLAS 8 TeV high-mass Drell-Yan and the photon PDF arXiv:1606.01736

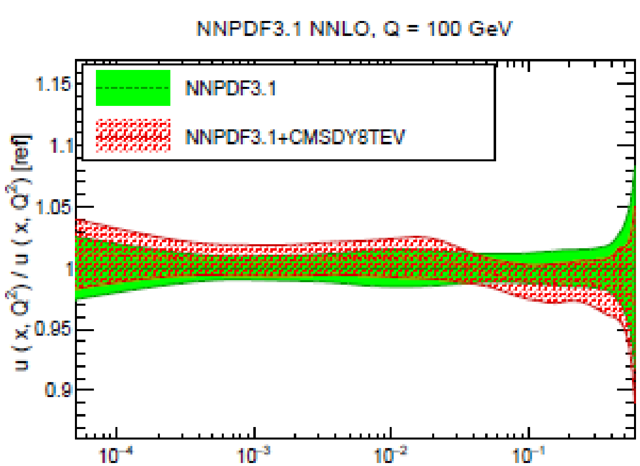


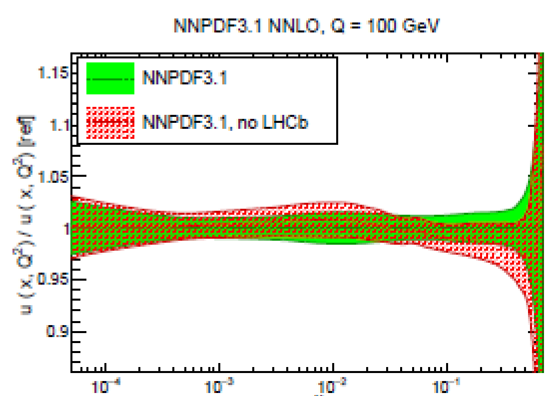
At high-mass di-lepton pairs may be photon induced rather than true Drell-Yan processes. These data have been used to constrain the photon-PDF

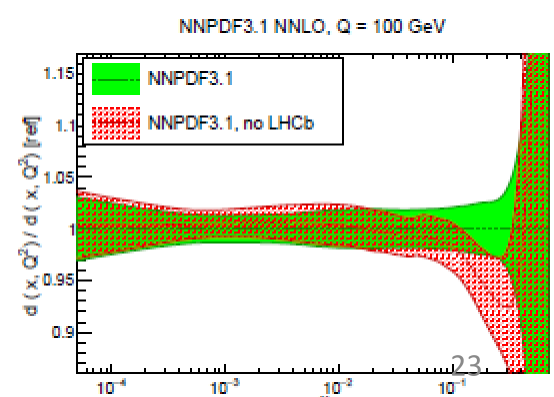


There is also CMS 8 TeV Z/γ* double differential Drell-Yan data (arXiv:1412.1115). However these data have very poor χ2/ndp~3.3 These data do not have a big pull on PDF fits

LHCb W,Z data probe a different kinematic region to both lower and higher-x values Their impact is mostly seen on high-x quarks. Low-x can present theoretical challenges

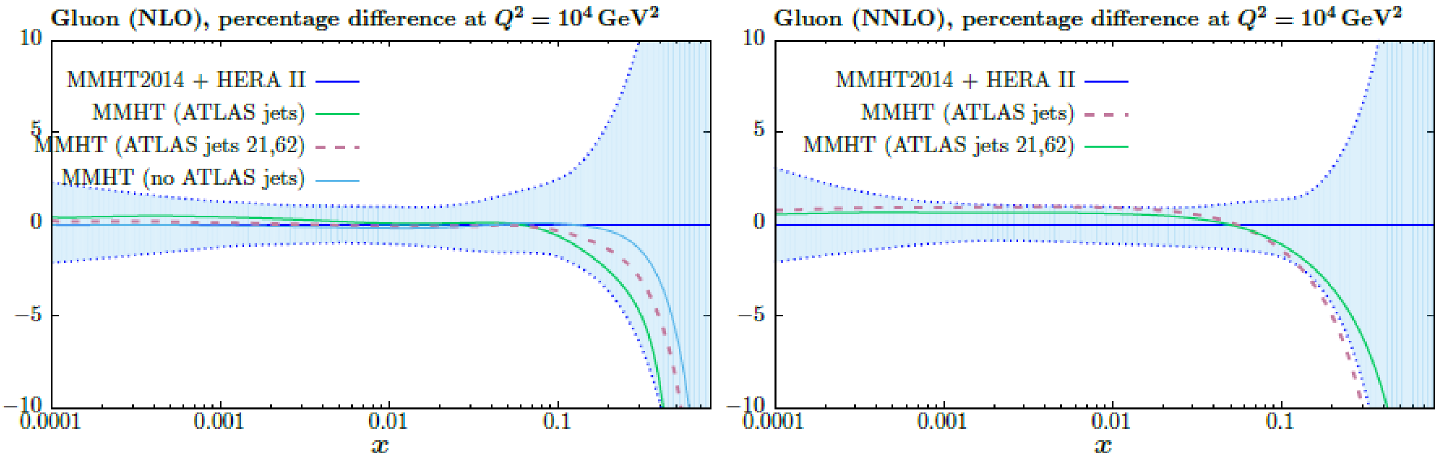






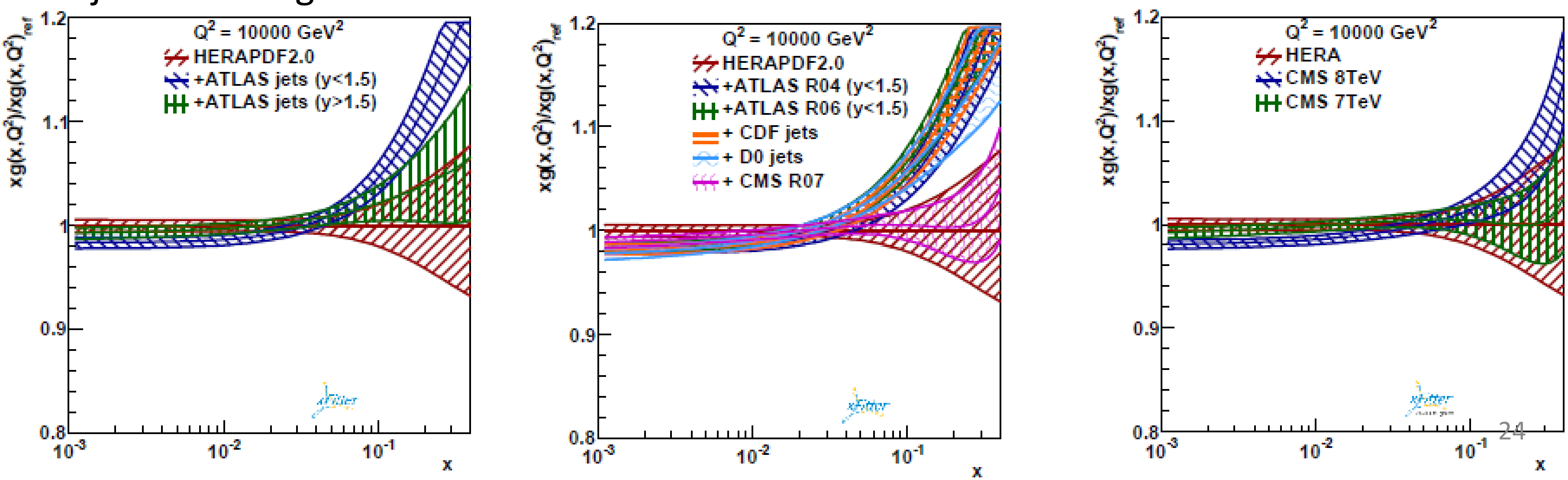
Gluon including ATLAS jet data at NLO and NNLO

MMHT 7 TeV study-gives slightly softer high-x gluon



The effect on the best fit gluon is noticeable, but within (or at boundary) of uncertainties. Softer at very high x.

Adding jet data to HERAPDF2.0 shows great consistency between the jets (apart from CMS 7 TeV)
ATLAS jets in this figure is 7 TeV inclusive



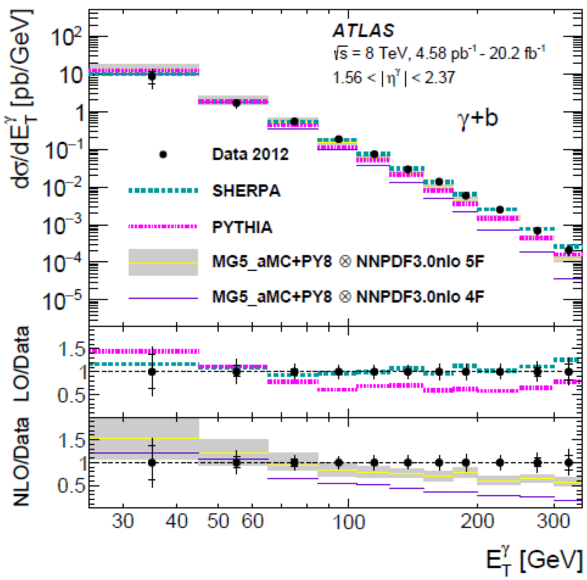
These spectra cannot be fitted at the same time because a statistical covariance matrix does not exist– although the systematic shift information IS provided

NNPDF have made fits and concluded that not only do CMs and ATLAS not agree so well but that also WITHIN an experiment the different spectra do not agree so well.

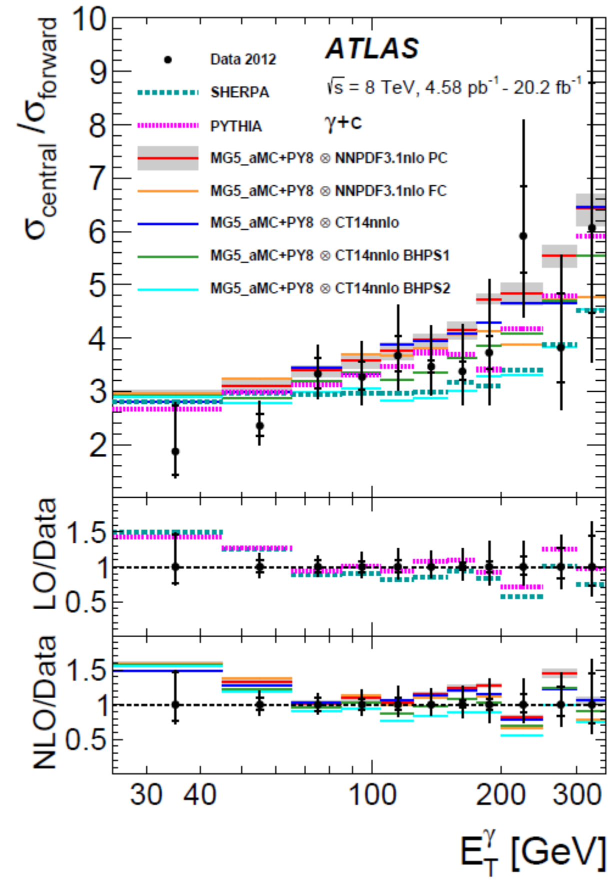
The chose to fit y_t from ATLAS and y_tt from CMS
When they do this they do NOT describe the other spectra very well

	ATLAS $d\sigma/dp_T^t$	ATLAS $d\sigma/dy_t$	ATLAS $d\sigma/dy_{t\bar{t}}$	ATLAS $d\sigma/dm_{t\bar{t}}$	ATLAS $(1/\sigma)d\sigma/dp_T^t$	ATLAS $(1/\sigma)d\sigma/dy_t$	ATLAS $(1/\sigma)d\sigma/dy_{t\bar{t}}$	ATLAS $(1/\sigma)d\sigma/dm_{t\bar{t}}$	ATLAS $\sigma_{t\bar{t}}$	CMS $d\sigma/dp_T^t$	CMS $d\sigma/dy_t$	CMS $d\sigma/dy_{t\bar{t}}$	CMS $d\sigma/dm_{t\bar{t}}$	CMS $(1/\sigma)d\sigma/dp_T^t$	CMS $(1/\sigma)d\sigma/dy_t$	CMS $(1/\sigma)d\sigma/dy_{t\bar{t}}$	CMS $(1/\sigma)d\sigma/dm_{t\bar{t}}$	CMS $\sigma_{t\bar{t}}$
Fit opt	2.19	0.64	1.84	5.01	2.49	**1.16**	3.81	4.55	**0.78**	2.91	4.98	1.07	4.77	3.33	5.78	**1.05**	8.05	**0.50**

arXiV:1710.09560 ATLAS γ+c and γ+b at 8 TeV



γ+b at high rapidity:
5-flavour scheme favoured over 4-flavout scheme

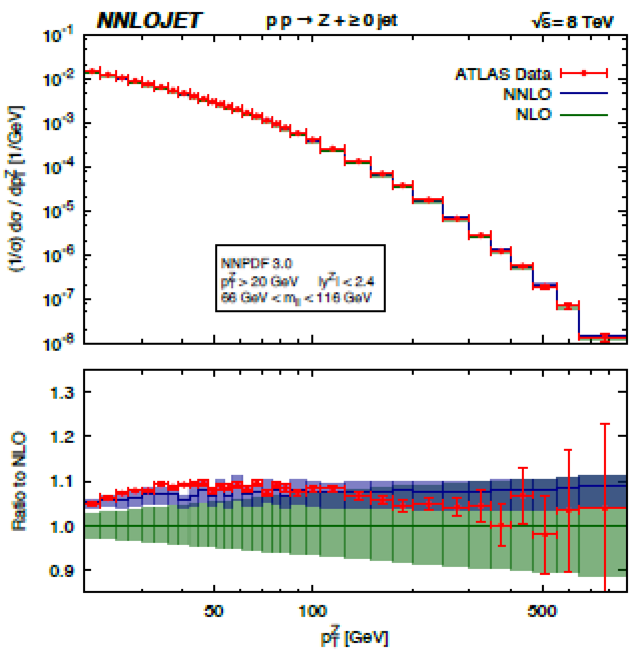


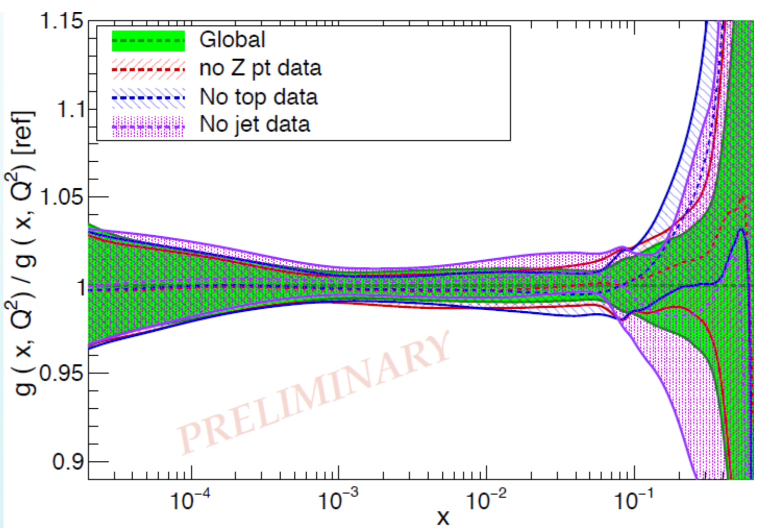
Central rapidity/forward rapidity for γ+c
No discrimination of intrinsic charm

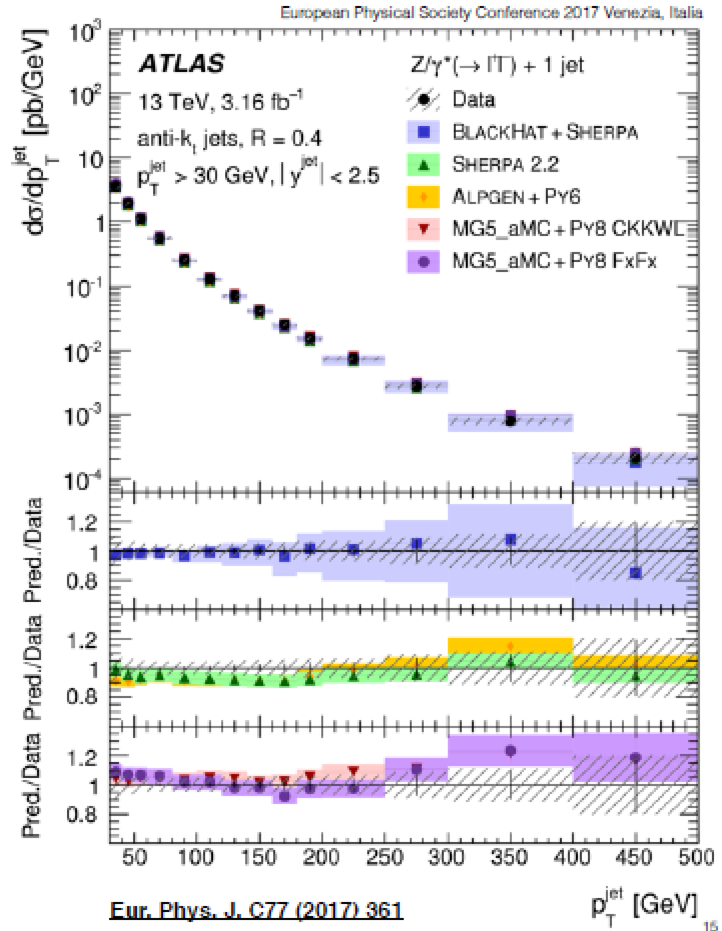
Boson (W,Z) +jets distributions

There are now NNLO predictions for Z +jets, Zpt and W+jets arXIV: 1607.01749, 1605.04295

There is new data– and more in the pipeline







NNPDF have compared the impact of Zpt to that of jets and top

Open charm and beauty production from LHCb added to HERA data
arXiv:1503.04581
Improves uncertainty on low-x gluon.
But this is a theoretically contentious region