

ATLAS data which feed into PDFs:

- W^\pm, Z production- strange and valence
- Drell-Yan (mostly the photon PDF)
- Jet production- gluon and high-x d
- Top-antitop production- gluon
- $Z_{pt}, Z+jets$ -gluon
- $W^\pm+jets$ — valence
- $W+c$ —strange
- $\gamma +c$ — intrinsic charm

An ATLAS global fit to various types of data-

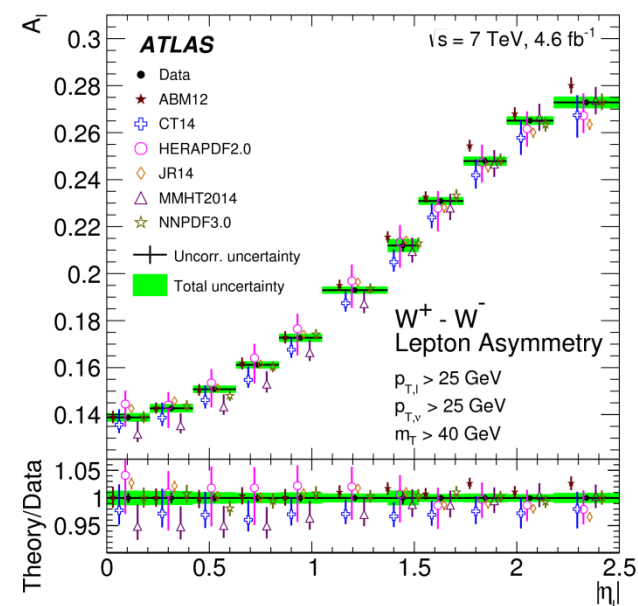
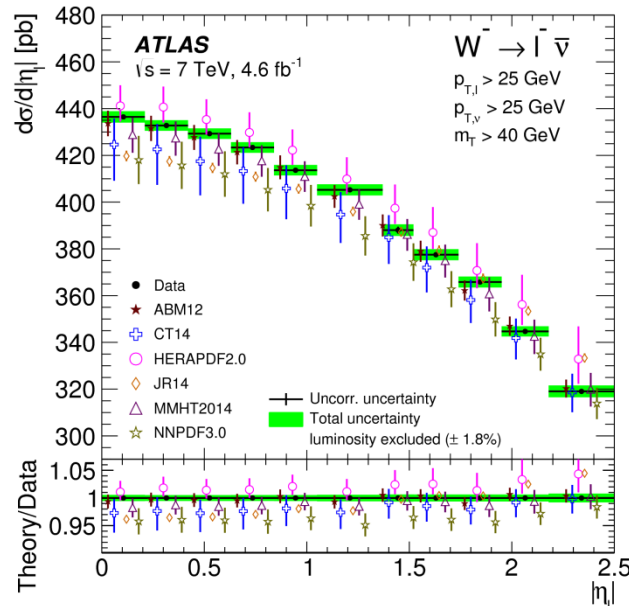
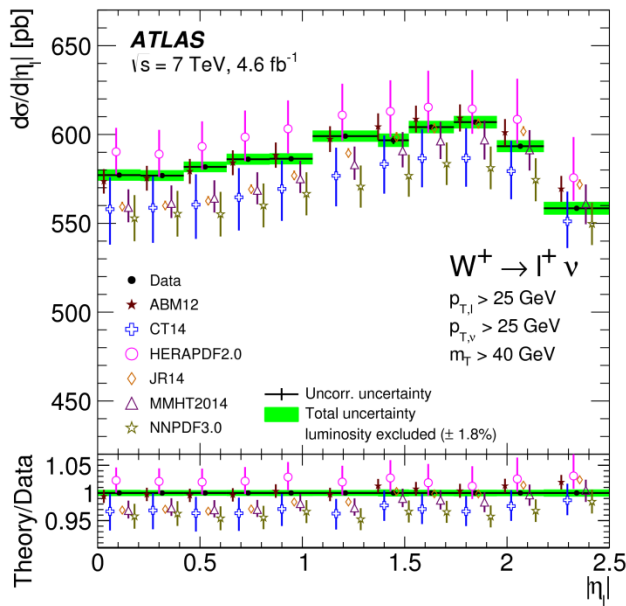
ATLASepWZ16 PDF set is the start

Recommendations for PDF usage---In particular for MC generation

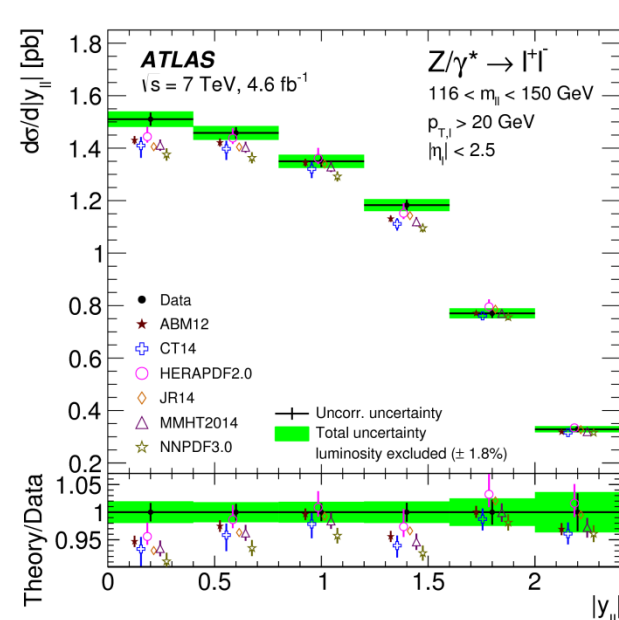
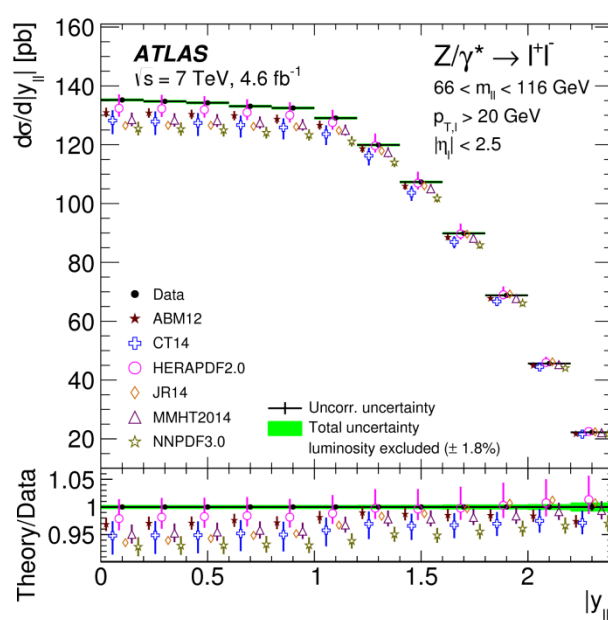
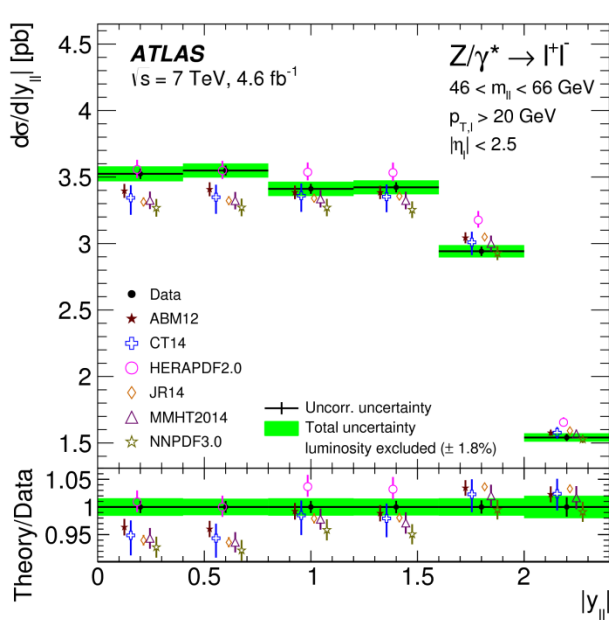
ATLAS/CMS cooperation?—see next talk from CMS

ATLAS high precision W and Z production arXiv:1612.03016

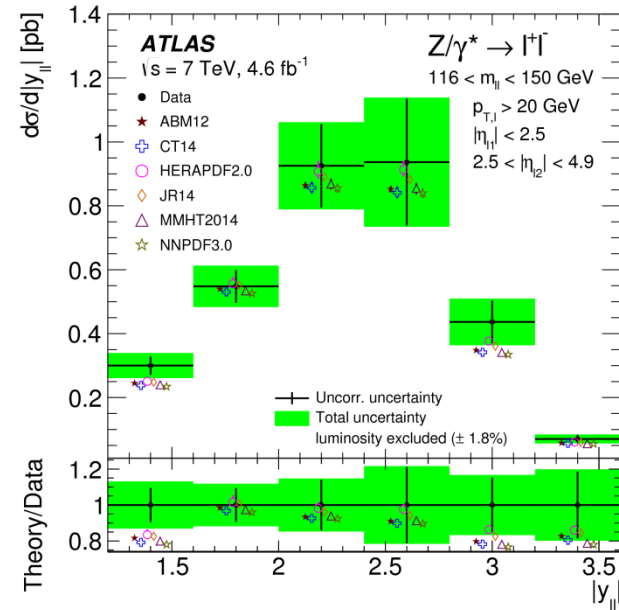
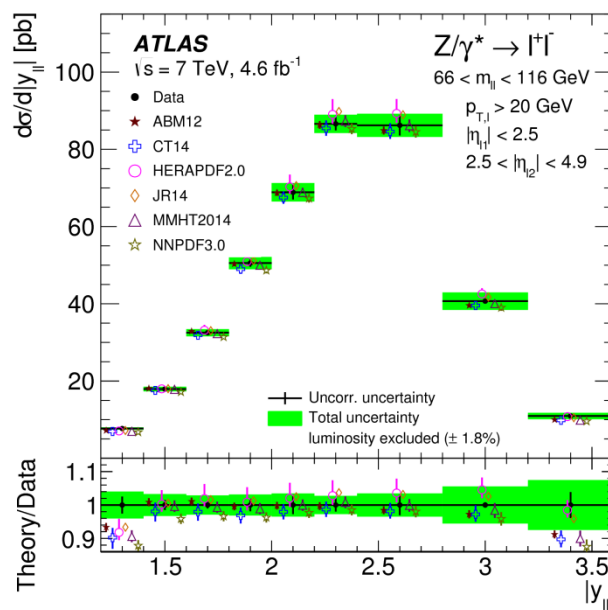
- W: Total (0.6–1.0%), multijet background (0.3–0.7%)
- Z Central: Total (0.4%), reconstruction efficiency (0.2–0.3%)
- Z Forward: Total (2.3%), identification efficiency (1.5%)
- 1.8% luminosity uncertainty

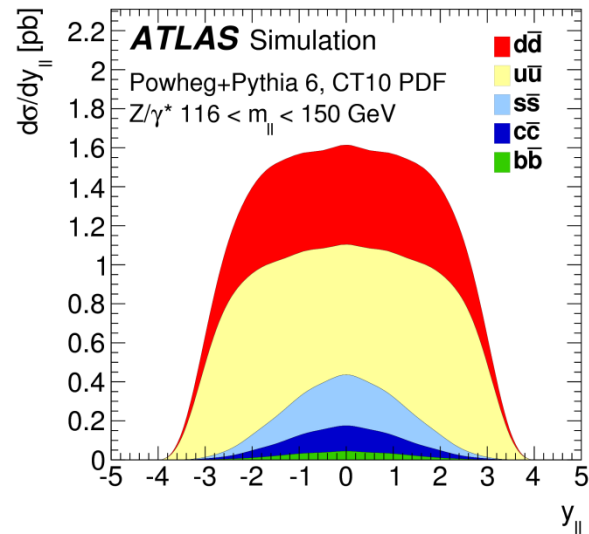
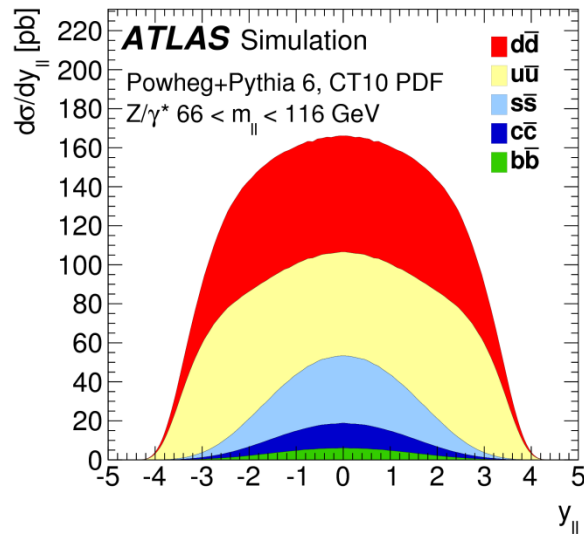
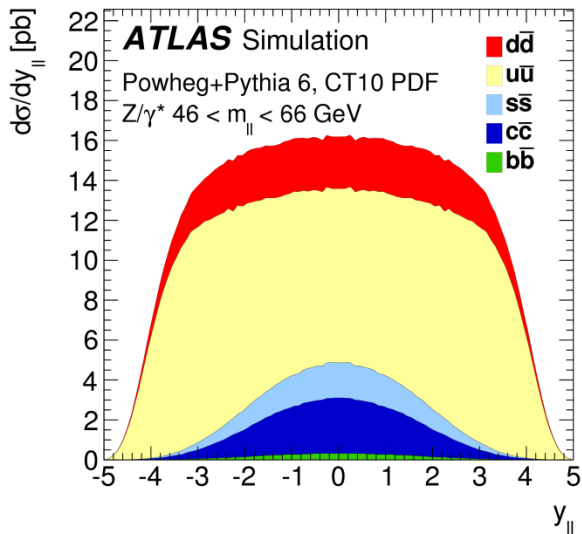


W+ and W- rapidity distributions separately, have more information than W-asymmetry

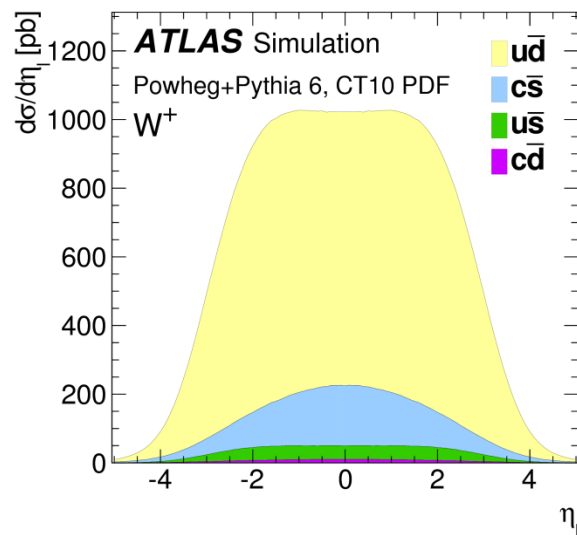
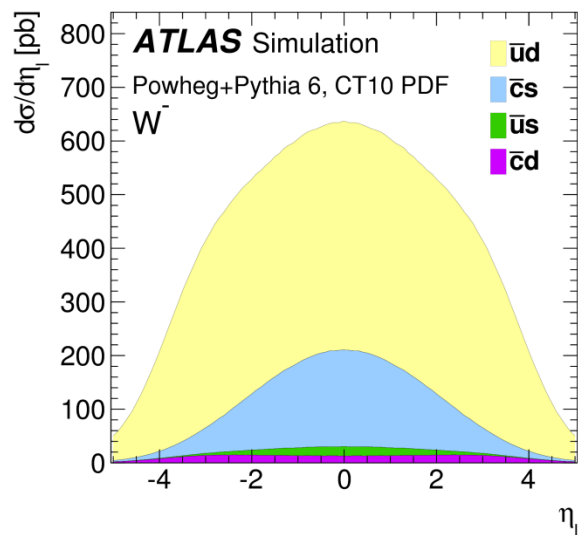


Z(*) rapidity distributions
 For low mass 46-66 GeV
 And high mass 116-150 GeV
 As well as central 66-116 GeV
 And for forward as well as
 central rapidity
 'Off- the-shelf' PDFs have
 difficulty describing these data
 --Particularly the central Z
 rapidity bin





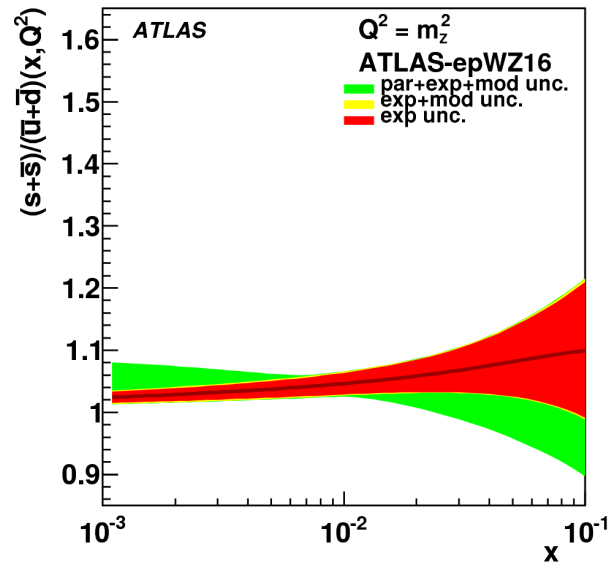
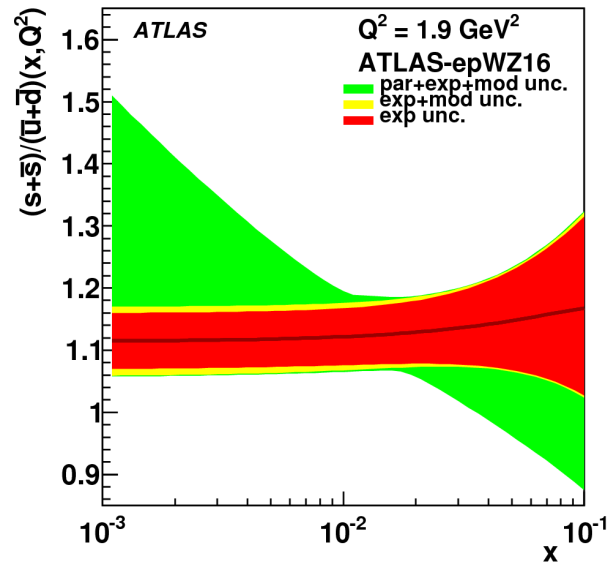
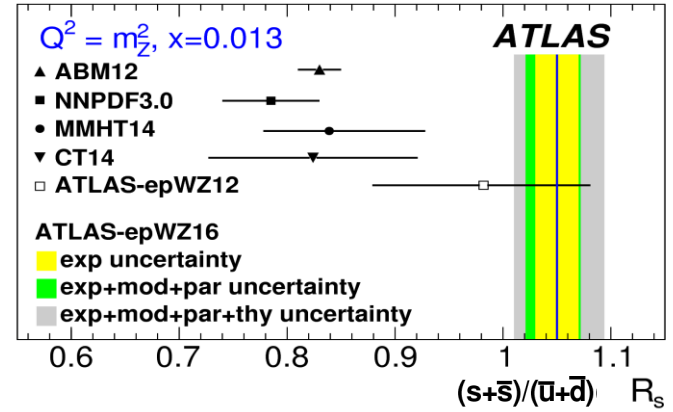
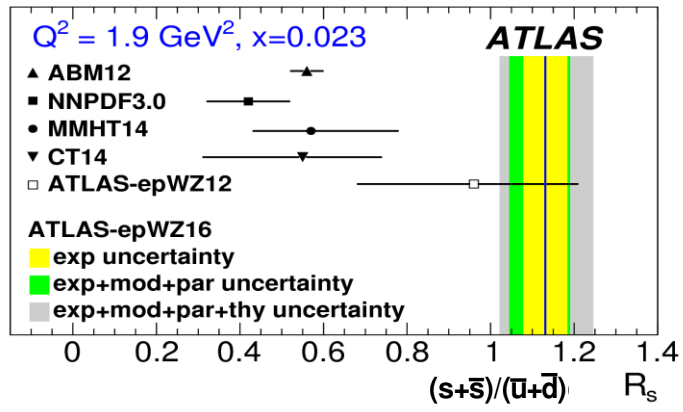
The central rapidity bin has a particularly large strange/anti-strange contribution

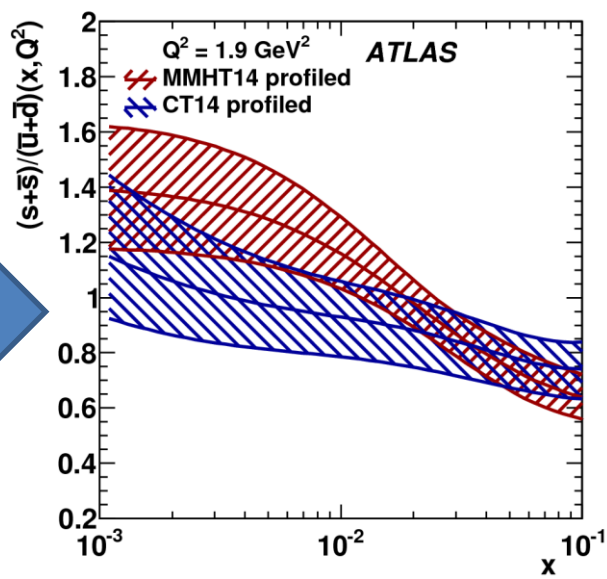
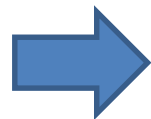
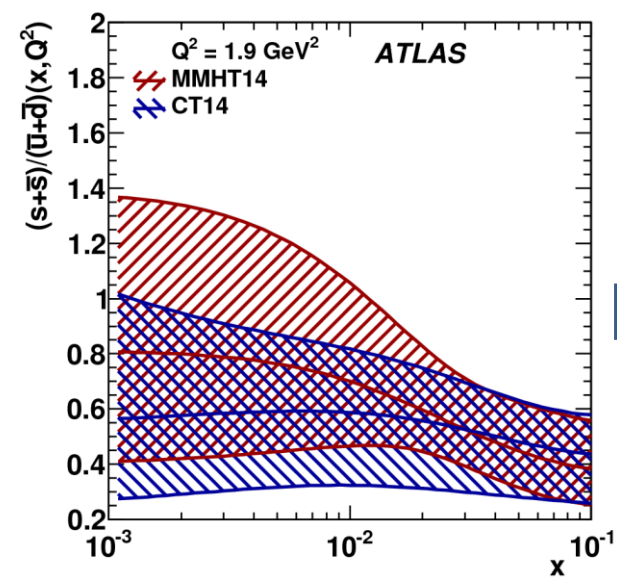
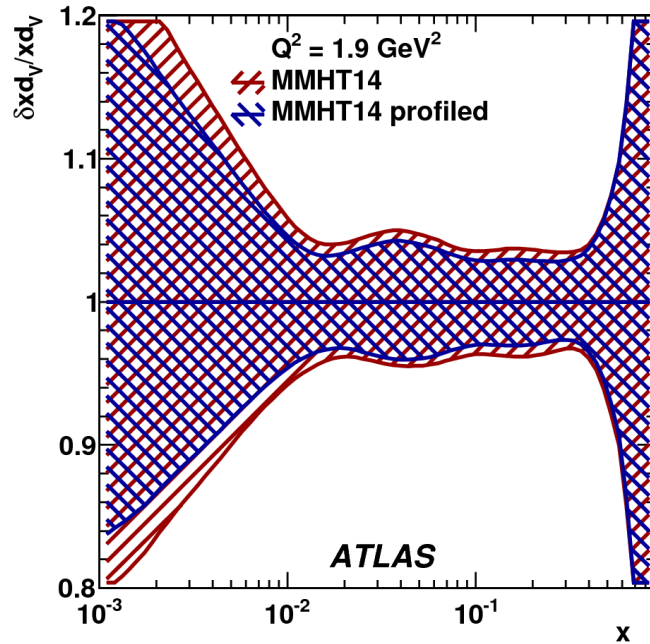
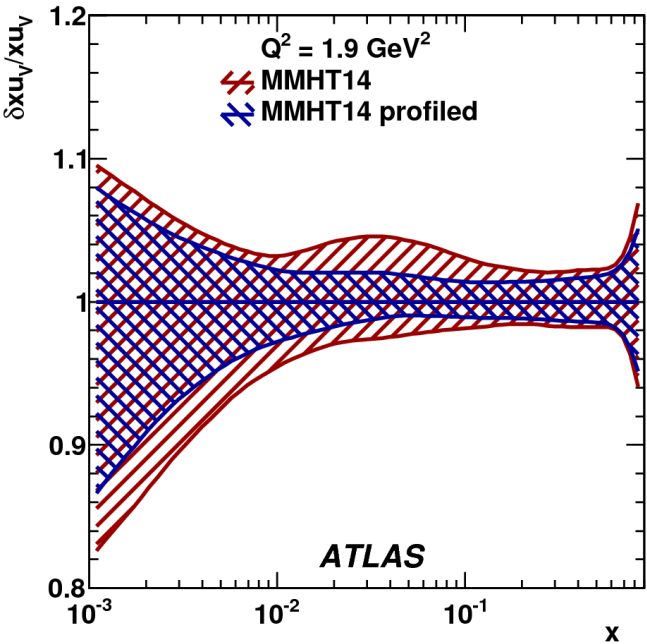


ATLAS has made a **complete NNLO QCD PDF fit** to these data, plus the HERA Deep Inelastic Scattering data. **The PDFs called ATLAS-epWZ16 are available on LHAPDF.**

This is much more than a single number at a single, x , Q^2 point.

This simple plot just summarises the conclusions: namely that there is more strangeness at low- x than in most of the off-the-shelf PDFs





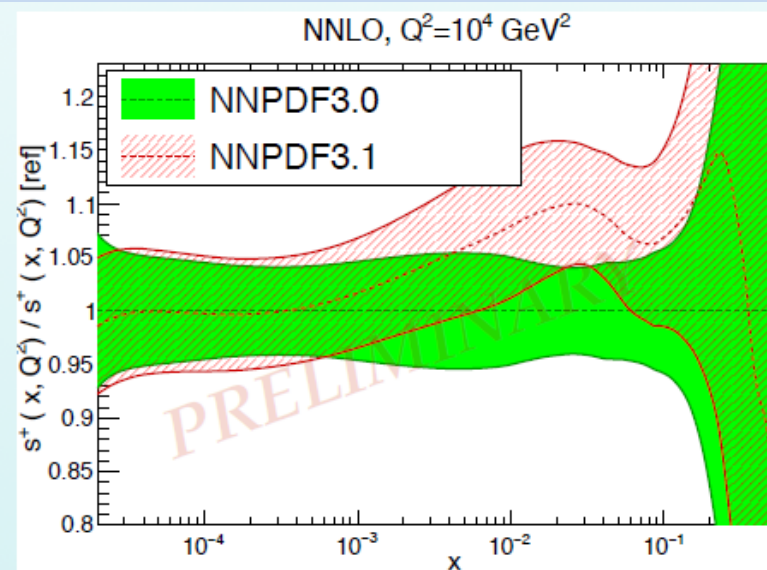
Impact of the ATLAS W,Z
 data on MMHT and CT
 global PDFs

The global PDF groups have started to look at these data

They are now in NNPDF3.1

Observe in particular how the strangeness has increase between NNPDF3.0 and 3.1

There is 'mild tension' between older heavy target neutrino data- which gave rise to the usual assumption of suppressed strangeness- and the ATLAS W,Z 2011 data



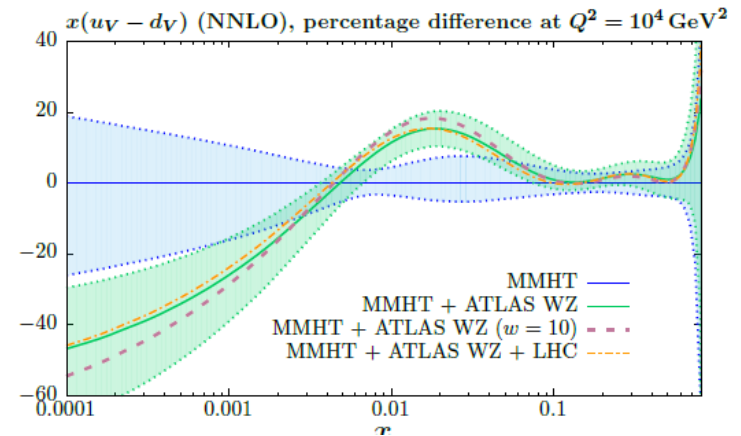
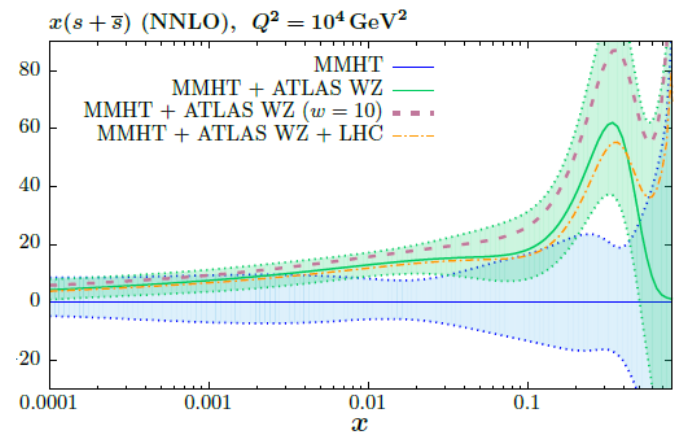
PDF set	$(s+\bar{s})/(\bar{u}+\bar{d})$ $R_s(0.023, 2 \text{ GeV}^2)$	$(s+\bar{s})/(\bar{u}+\bar{d})$ $R_s(0.013, M_Z^2)$
NNPDF3.0	0.47 ± 0.09	0.79 ± 0.04
NNPDF3.1	0.61 ± 0.14	0.83 ± 0.06
NNPDF3.1 collider-only	0.85 ± 0.16	0.93 ± 0.06
NNPDF3.1 HERA + ATLAS W, Z	0.96 ± 0.20	0.98 ± 0.09
ATLAS W, Z 2010 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

MMHT have also included these data and found increased strangeness— these PDFs are not public yet.

They also observe mild tension with older neutrino data AND with CMS 7 TeV differential Z/γ data and W+c data BUT they state that newer LHC data (mostly LHCb) are compatible and ‘pull in the same direction’

All groups find the CMS 8 TeV differential Z/γ data – which would be relevant to the strangeness question— ----difficult to fit ($\chi^2/ndp \sim 3.3$).

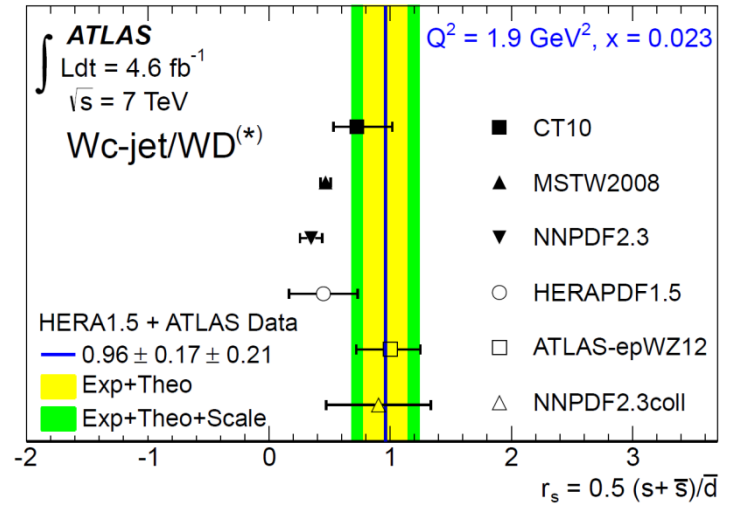
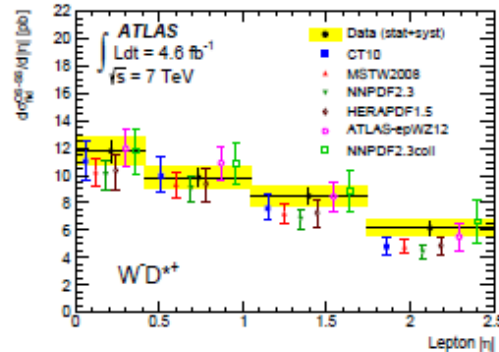
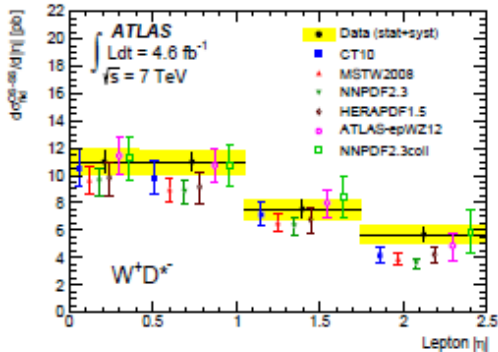
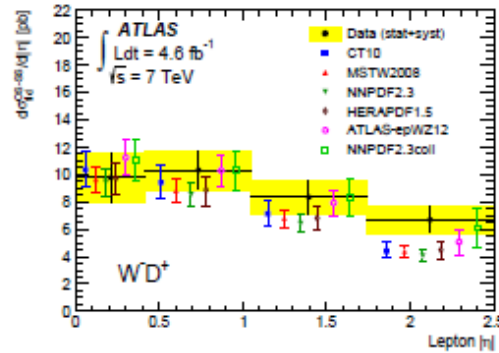
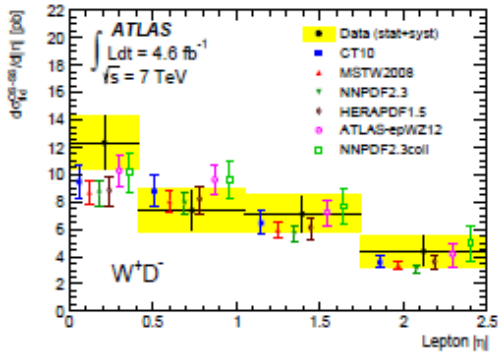
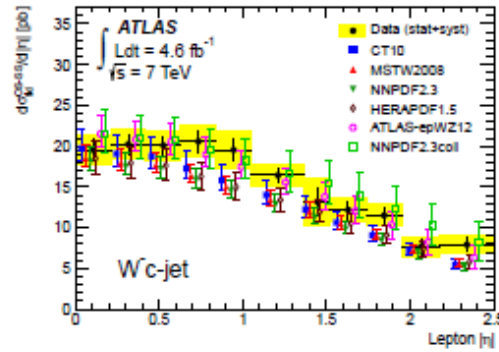
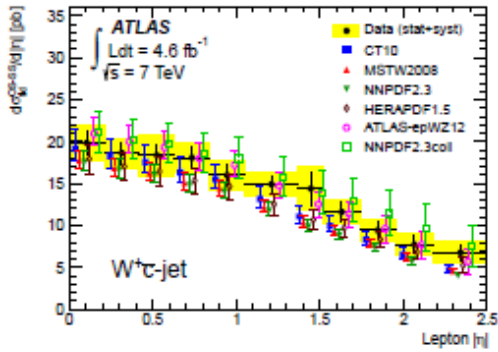
The ATLAS W,Z data also impact the valence shapes and decrease valence uncertainty.



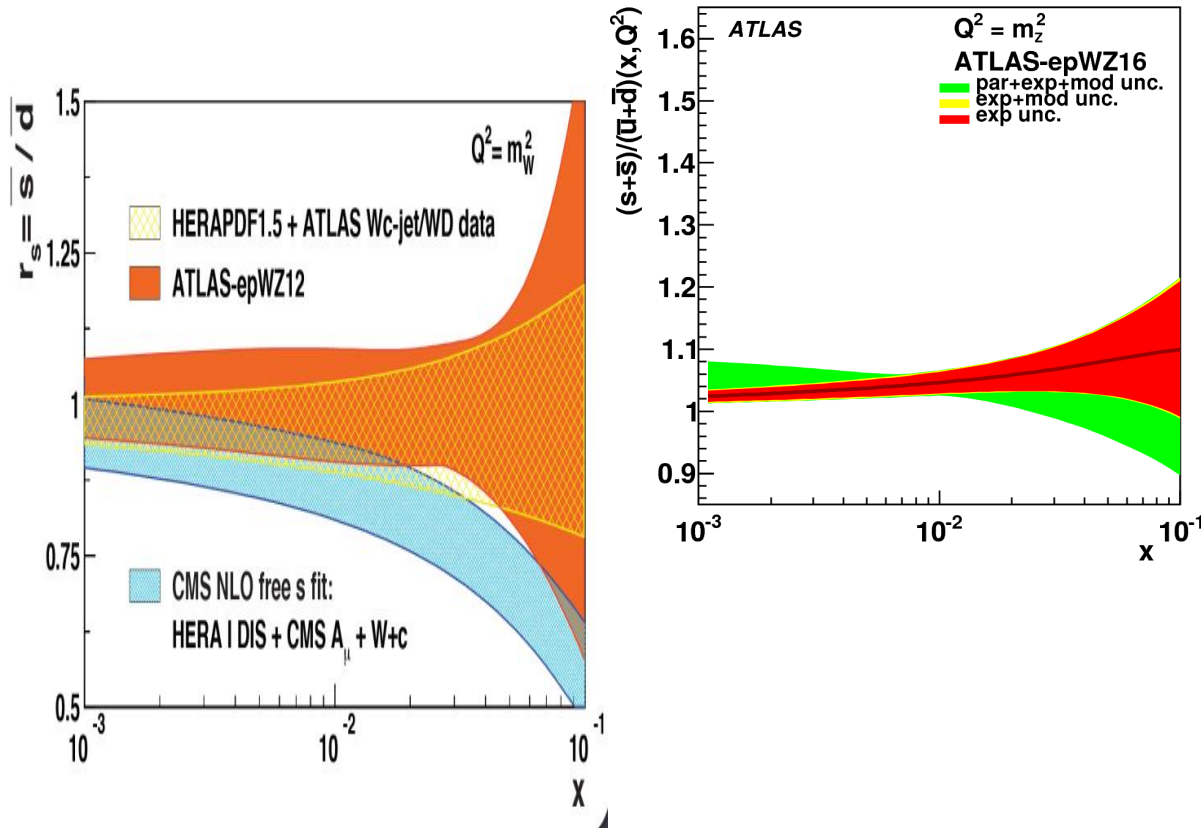
So before leaving the strangeness question let us look at the results on W+charm data from ATLAS W+c-jet W+D, W+D* channels JHEP05(2014)068

In all cases the data favour the predictions based on unsuppressed strangeness

We can quantify this..



Let us look at the results on W+charm data from BOTH ATLAS and CMS
 Compare the older results and the new ATLAS W,Z result on the same scale
 ATLASepWZ12 was based on ATLAS W,Z 2010 data
 ATLASepWZ16 is based on the ATLAS W,Z 2011 precision data
 ATLAS W+c data is in agreement



CMS W+c data is somewhat lower

But the analysis techniques of CMS and ATLAS are not sufficiently similar
 ATLAS use W+D(*) as well as W+c-jet and the hadronisation corrections are not agreed
 between the collaborations--- see next talk from CMS

ATLAS jet data:2011 7 TeV AND 2012 8 TeV

The MMHT group report the following

Attempted fit to high luminosity **ATLAS 7 TeV inclusive jet data**
(**JHEP 02 (2015) 153**)

Take as default $R = 0.4$ and $\mu = p_{T,1}$ and work at **NLO**.

Prediction at **NLO** gives $\chi^2/N_{pts} = 413.1/140$.

Refit gives improvement only to $\chi^2/N_{pts} = 400.4/140$.

Deterioration in other data only $\Delta\chi^2 \sim 3$, so failure not due to strong tensions.

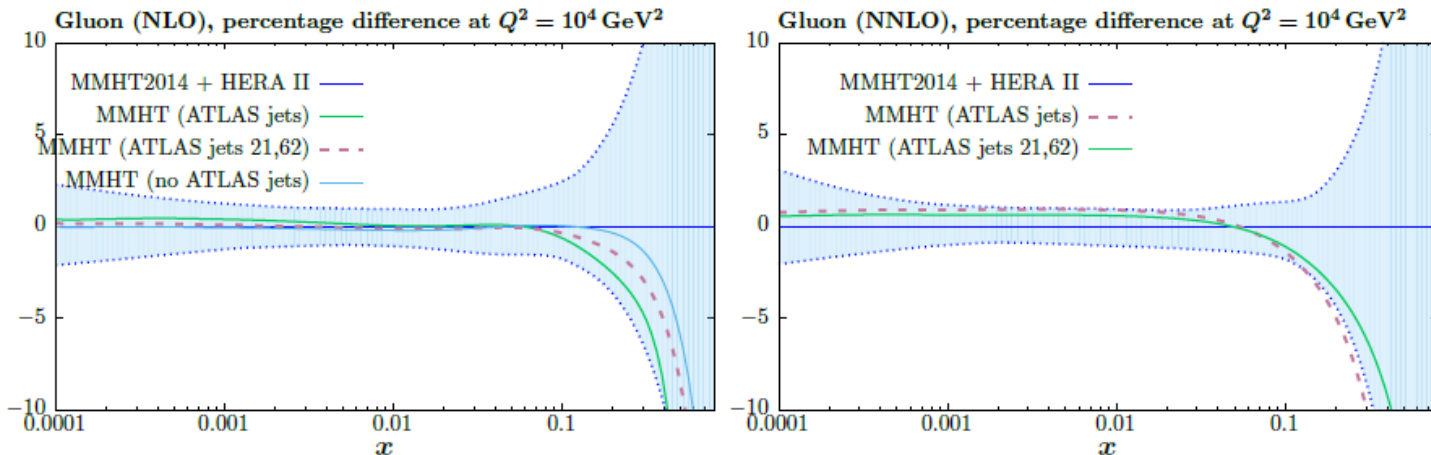
Cannot simultaneously fit data in all bins. Mismatch in one rapidity bin different in form to neighbouring bins probing PDFs of similar flavour, x and Q^2 .

However they perform an exercise decorrelating some of the Jet Energy Scale systematics

With correlations between rapidity bins relaxed for just two sources of systematics $\chi^2/N_{pts} = 178/140 = 1.27$.

NNLO corrections are now available for the ATLAS 7 TeV jet data and MMHT also used them

Gluon including ATLAS jet data at NLO and NNLO



Gluon similar whether decorrelation is used or NOT

Gluon is similar whether NLO or NNLO is used

Values of χ^2 without (with) the ATLAS jets data in the fit

	Full Corr.	21,62 decorr.
χ^2 , NLO	(413)400	(180)178
χ^2 , NNLO	(443)427	(211)204

Find significant, if not dramatic, deterioration in fit quality in all cases from NLO to NNLO
 Not an issue of tension with other data.

CT group agree that one cannot get a good fit to all ATLAS jet rapidity distributions at once

CT see tension between ATLAS and CMS jet data sets

[CT see tension between ATLAS and CMS top data sets in the opposite direction to the tension in the jet data sets]

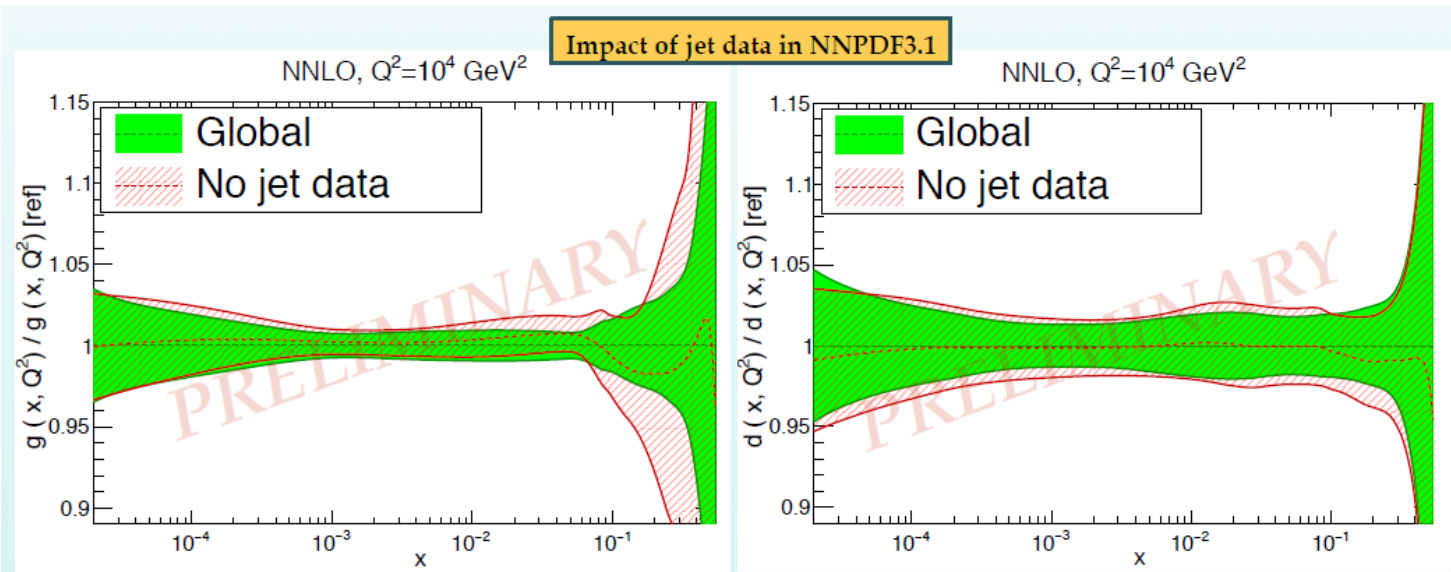
Currently under investigation; what is the envelope of PDFs resulting from fits to individual rapidity bins, i.e. how unanimous are the rapidity bin data

NNPDF also agree that one cannot get a good fit to all rapidity bins so for NNPDF3.1

ONLY the central rapidity bin of the ATLAS 7 TeV jet data is being used

By contrast all CMS 7 TeV jet data is used.

NNLO results are only available for the ATLAS 7 TeV jets SO NLO matrix elements are used, with P_t as central scale and scale uncertainties added

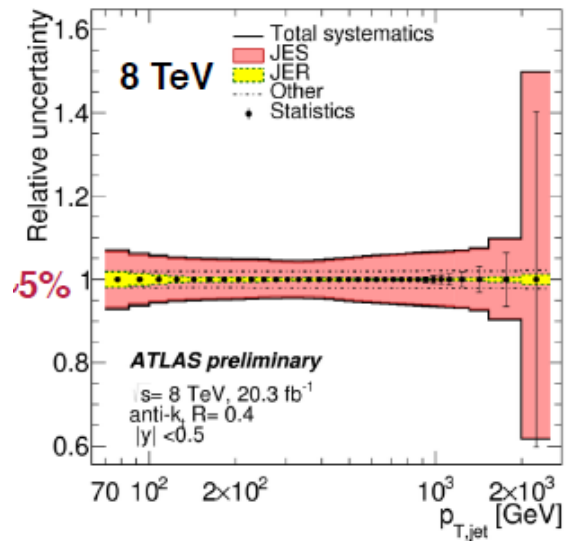
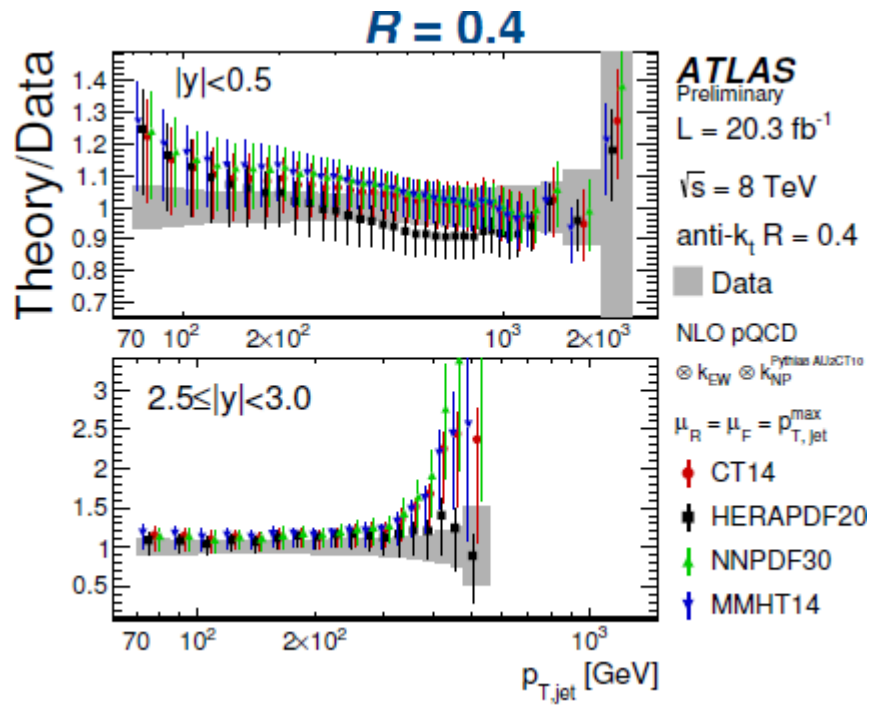
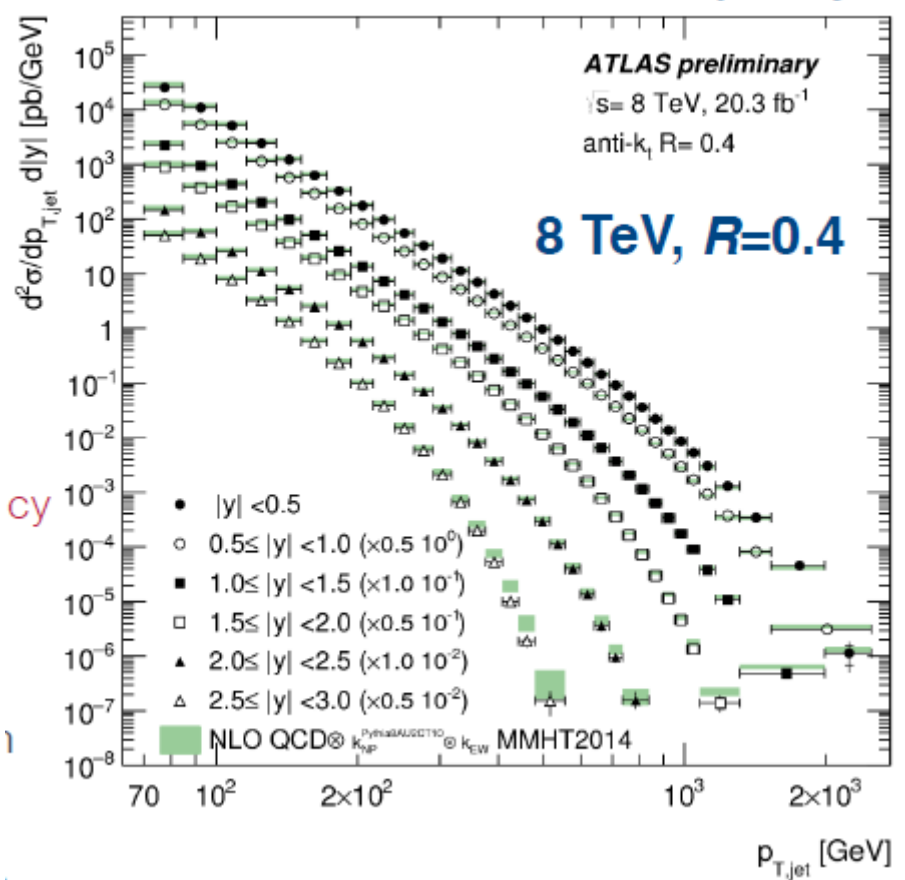


• If the jet p_T is used as central scale, NNLO/NLO K-factors only a few percent

• NNLO/NLO shift within NLO scale uncertainties

HOW about the NEW 8 TeV ATLAS jet data ATL-COM-PHYS-2016-371 preliminary

Double-differential distribution ($p_{T,jet}$ & $|y|$)



- **JES & JER** are dominant experimental uncertainties in inclusive measurement
- Significant reductions over 7 TeV result

However there is still a problem fitting all rapidity bins together for the 8TeV ATLAS jet data

- Not localized in $|y|$, and no central-forward tension
- Little sensitivity to scale / non-perturbation correction
- Similar to 7 TeV data

ATLAS are considering decorrelation scenarios--- but not completely decorrelating rapidity bins ---instead decorrelations that are smooth functions of y and/or p_T are considered for '2-point systematics'

χ^2 in combined $|y|$ bins

	χ^2/ndf	$p_T^{\text{jet,max}}$ $R = 0.4$
low-p_T	$p_T > 70$ GeV	
	CT14	345/171
	HERAPDF20	415/171
	NNPDF30	261/171
	MMHT2014	355/171
high-p_T	$100 < p_T < 400$ GeV	
	CT14	127/72
	HERAPDF20	148/72
	NNPDF30	124/72
	MMHT2014	132/72

- **Jet energy scale:** flavor response / MJB fragmentation / pileup ρ topology
- **Alternative correlation scenarios** split smoothly by $p_T / |y|$ into sub-components

Split option	Sub-component(s) definition(s), completed by complementary
1	$L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot \text{uncertainty}$
7	$L(y , 0, 3) \cdot \text{uncertainty}$
9	$L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot L(y , 0, 3) \cdot \text{uncertainty}$
18	$\sqrt{1 - L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5))^2} \cdot \sqrt{1 - L(y , 0, 2)^2} \cdot \text{uncertainty}$ $\sqrt{1 - L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5))^2} \cdot L(y , 2, 3) \cdot \text{uncertainty}$

Example 2-point correlation definitions

- Can improve χ^2 by **up to 96** points for $p_T > 70\text{GeV}$

Top-antitop differential distributions in the lepton +jets channel from 8 TeV data: arXiv:1511.04716

Top differential distributions can be included in PDF fits now using NNLO calculations,
This gives information on the high-x gluon.

ATLAS $d\sigma/dp_T^t$	a	8	$0 < p_T^t < 500$ GeV
ATLAS $d\sigma/d y_t $	a	5	$0 < y_t < 2.5$
ATLAS $d\sigma/d y_{t\bar{t}} $	a	5	$0 < y_{t\bar{t}} < 2.5$
ATLAS $d\sigma/dm_{t\bar{t}}$	a	7	$345 < m_{t\bar{t}} < 1600$ GeV
ATLAS $(1/\sigma)d\sigma/dp_T^t$		8	$0 < p_T^t < 500$ GeV
ATLAS $(1/\sigma)d\sigma/d y_t $		5	$0 < y_t < 2.5$
ATLAS $(1/\sigma)d\sigma/d y_{t\bar{t}} $		5	$0 < y_{t\bar{t}} < 2.5$
ATLAS $(1/\sigma)d\sigma/dm_{t\bar{t}}$		7	$345 < m_{t\bar{t}} < 1600$ GeV

Fit either absolute or normalized distributions (in the latter case adding total cross-sections) and compare results

But only one spectrum can be fitted at once because there is no information on statistical correlations between spectra

For $y_t, y_{t\bar{t}}, m_{t\bar{t}}$ $\mu_R = \mu_F = \mu = H_T/4$, $H_T \equiv \sqrt{m_t^2 + (p_T^t)^2} + \sqrt{m_{\bar{t}}^2 + (p_T^{\bar{t}})^2}$

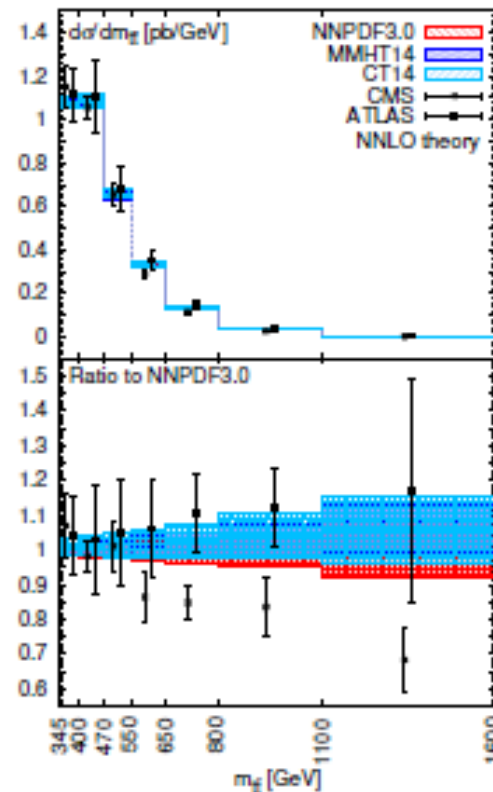
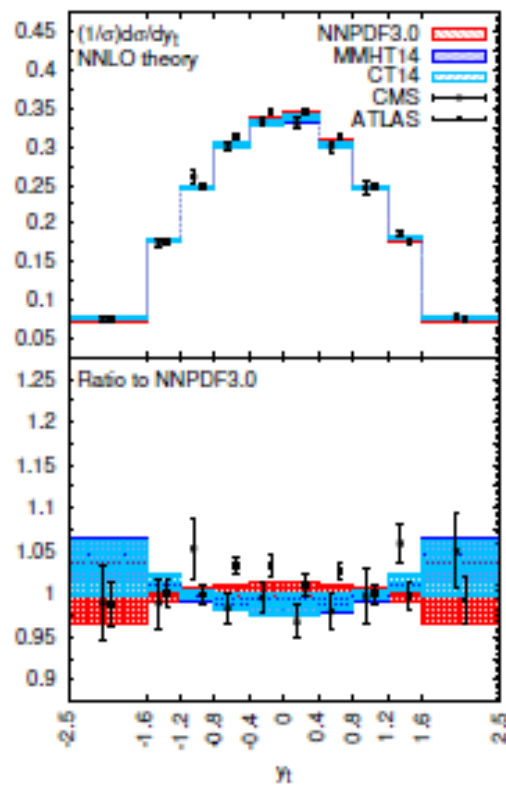
For p_T^t $\mu'_R = \mu'_F = \mu' = \sqrt{m_t^2 + (p_T^t)^2}/2$

There is similar data from CMS and they are not always fully consistent between CMS and ATLAS

The CT PDF fitting group say

- ATLAS and CMS have different trends; in this case, ATLAS favors harder gluon at high x , CMS weaker gluon
- In general, the ATLAS and CMS top results are in tension internally, and with each other

Thus adding in both ATLAS and CMS together could even increase uncertainties



By contrast the NNPDF group make particular choices of distributions to fit

- the normalized top-quark rapidity distribution $(1/\sigma)d\sigma/dy_t$ from ATLAS;
- the normalized top-quark pair rapidity distribution $(1/\sigma)d\sigma/dy_{t\bar{t}}$ from CMS;
- and the total inclusive cross-section $\sigma_{t\bar{t}}$ from ATLAS and CMS at $\sqrt{s} = 8$ TeV.

By contrast the NNPDF group make particular choices of distributions to fit for NNPDF3.1

- the normalized top-quark rapidity distribution $(1/\sigma)d\sigma/dy_t$ from ATLAS;
- the normalized top-quark pair rapidity distribution $(1/\sigma)d\sigma/dy_{t\bar{t}}$ from CMS;
- and the total inclusive cross-section $\sigma_{t\bar{t}}$ from ATLAS and CMS at $\sqrt{s} = 8$ TeV.

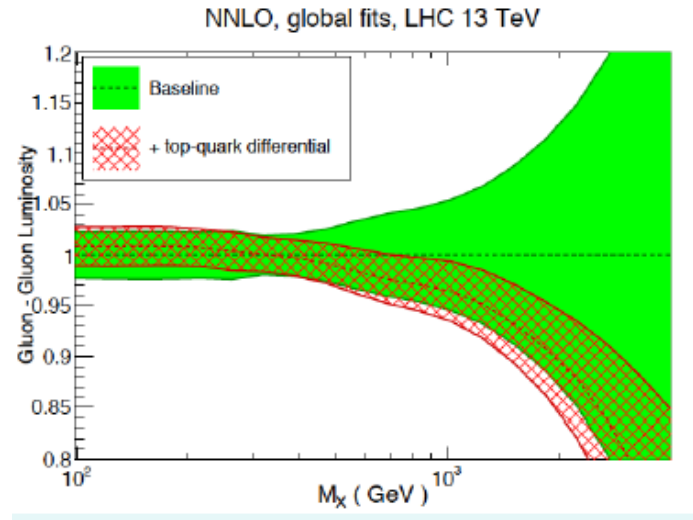
	ATLAS $d\sigma/dp_T^t$	ATLAS $d\sigma/dy_t$	ATLAS $d\sigma/dy_{t\bar{t}}$	ATLAS $d\sigma/dm_{t\bar{t}}$	ATLAS $(1/\sigma)d\sigma/dp_T^t$	ATLAS $(1/\sigma)d\sigma/dy_t$	ATLAS $(1/\sigma)d\sigma/dy_{t\bar{t}}$	ATLAS $(1/\sigma)d\sigma/dm_{t\bar{t}}$	ATLAS $\sigma_{t\bar{t}}$	CMS $d\sigma/dp_T^t$	CMS $d\sigma/dy_t$	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

Overall good description of the $t\bar{t}$ differential distributions included in the fit
no evident signs of tension between ATLAS/CMS and with the rest of the dataset

Distributions not included in the fit not well described, except companion absolute distr.

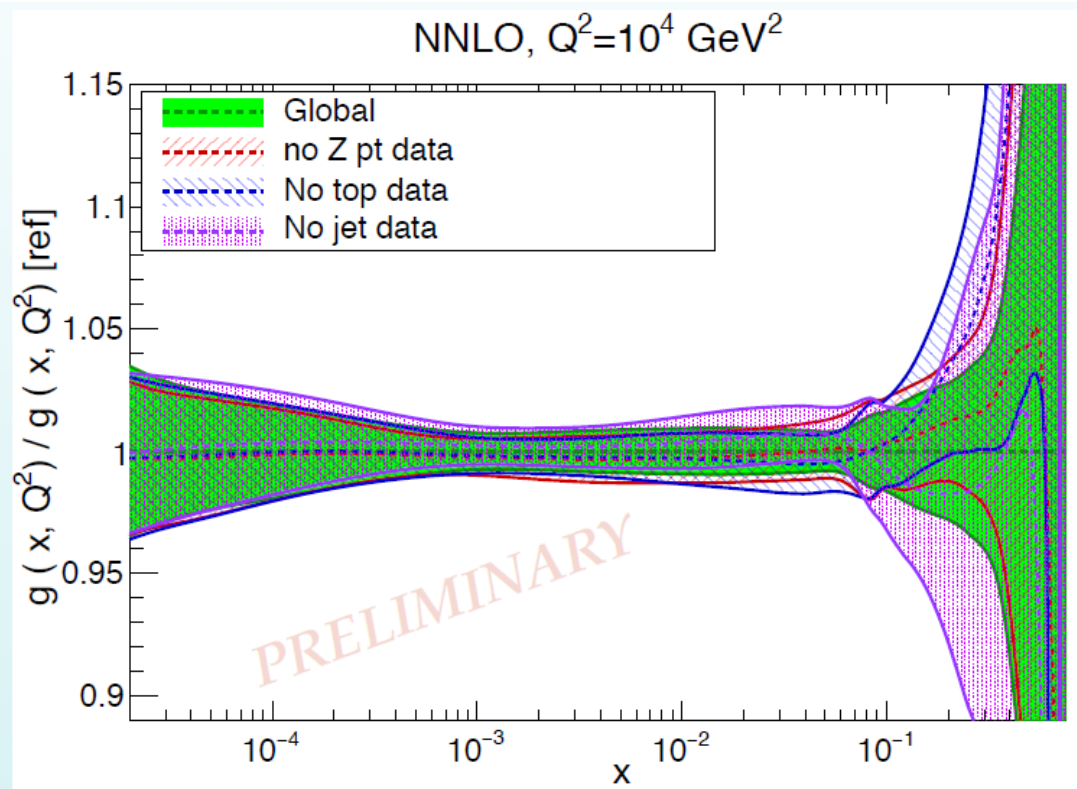
And the resulting gluon is somewhat softer than the NNPDF3.0 gluon

NNPDF also note that the top differential data give comparable constraints to the jets despite far fewer data points



Further data sets considered by the global fitters CT, MMHT, NNPDF include:
the ATLAS 7 and 8 TeV Z-pt data

But are top, jets and Zpt all consistent in how they affect the gluon?
More or less, IF the data you use are picked as NNPDF pick them



However CT see tension between jets and Zpt- and tension between jets and top

The data sets considered so far

- Top production
- Jet production
- W,Z production

are not the only data sets we consider as relevant for PDFs

We also have:

- Drell-Yan 8TeV [Published High Mass 8 TeV impacts the photon PDF arXiv: 1701.08553](#)
- Z pt data- 7 and 8 TeV used by NNPDF
- Z+jets 8 TeV- forthcoming
- W+jets 8TeV- forthcoming
- W+c --new analysis at 13TeV
- γ +HF ---New analysis coming on 8 TeV data

We plan to make a ‘global fit’ using many sorts of ATLAS data, obviously such a PDF fit would ideally describe our data for MC generation

Describing the way our data differs from some of the global fits has been important for

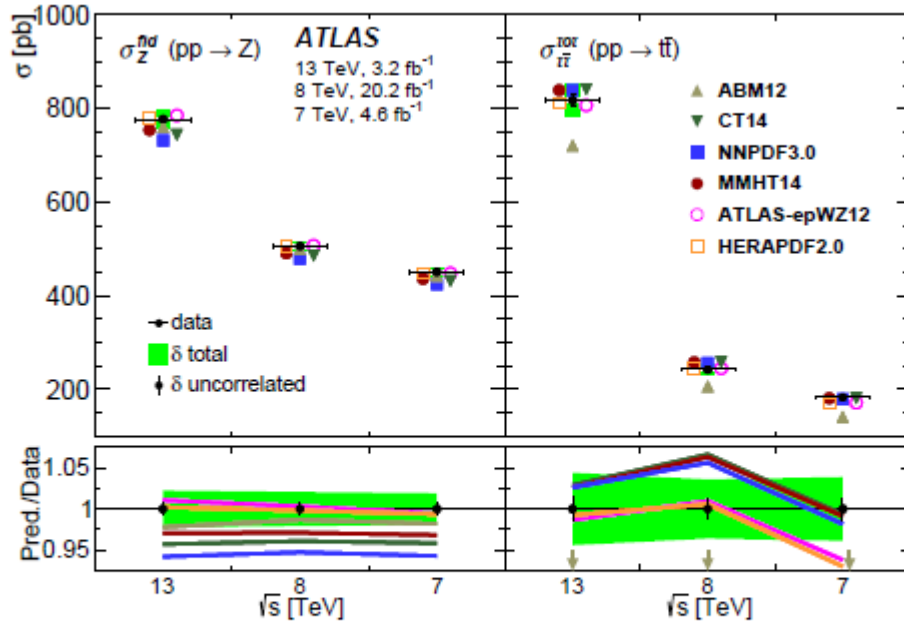
- $\sin^2\theta_w$ from Z forward/backward asymmetry
- W-mass measurement
- W' and Z' searches

To do this we need to consider correlations between data sets. A start on such a study has been made by the evaluation of top-antitop/Z ratios.

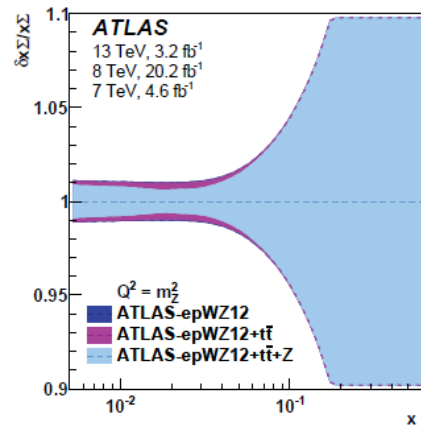
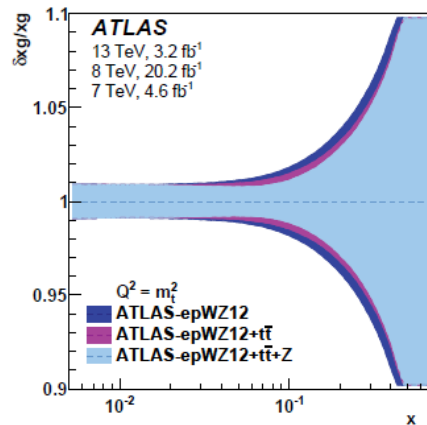
Top-anti-top/Z ratios at 7,8 and 13GeV : JHEP02(2017)117

It is necessary to evaluate the correlations between different classes of data

Source / \sqrt{s} [TeV]	$\delta \sigma_Z^{\text{fid}}$			$\delta \sigma_{t\bar{t}}^{\text{tot}}$		
	13	8	7	13	8	7
Luminosity	A	B	C	A	B	C
Beam energy	A	A	A	A	A	A
Muon (lepton) trigger	A	A*	A	A	B	B
Muon reconstruction/ID	A	B	C	A	D	D
Muon isolation	A	A	A	B	C	D
Muon momentum scale	A	A	A	A	A	A
Electron trigger	A	A	A	A	—	—
Electron reconstruction/ID	A	B	C	A	D	D
Electron isolation	A	A	—	B	C	D
Electron energy scale	A	A	A	A	A	A
Jet energy scale	—	—	—	A	B	B
b -tagging	—	—	—	A	B	B
Background	A	A	A	B	B	B
Signal modelling (incl. PDF)	A	A	A	B*	B	B



	ATLAS-epWZ12	CT14	MMHT14	NNPDF3.0	HERAPDF2.0	ABM12
χ^2/NDF	8.3 / 6	15 / 6	13 / 6	17 / 6	10 / 6	25 / 6
p-value	0.22	0.02	0.05	0.01	0.11	< 0.001



Recommendation for PDFs for MC generation

It is recommended

A) to use a particular PDF set for simulations, which is

- Successfully tested against ATLAS data
- Used in crucial MC tunings
- Includes where relevant QED (photon) PDFs
- Exists in LO, NLO and NNLO
- In use already by the Collaboration

That selects NNPDF3 as an obvious baseline choice



Right now this points at NNPDF3.1 which includes our latest data and which is available, **but it is not perfect**

B) To store weights as is needed and technically possible (wrt space and computing time) for allowing to reweight efficiently.

C) To not deny requests for special simulations with different sets where reasonable, by either full simulations or generator level MCs.

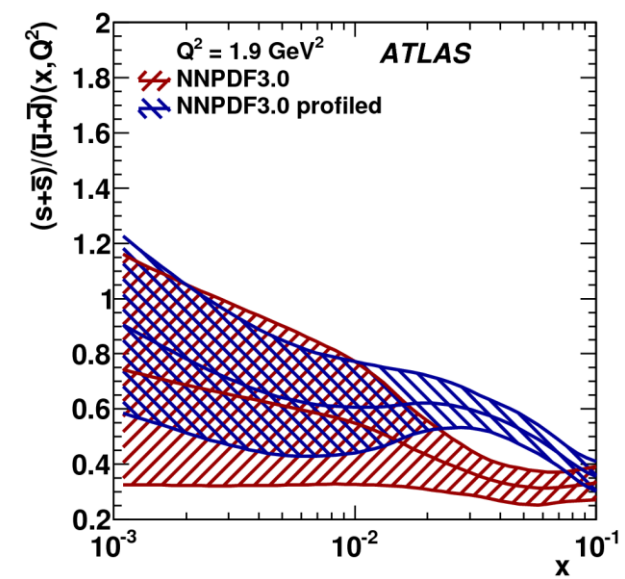
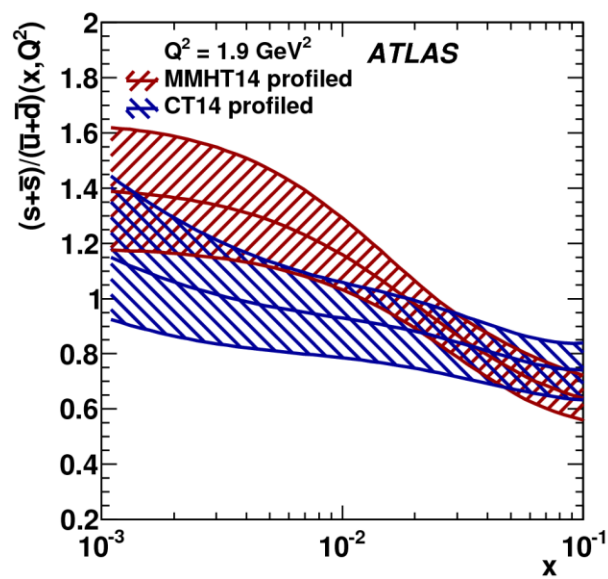
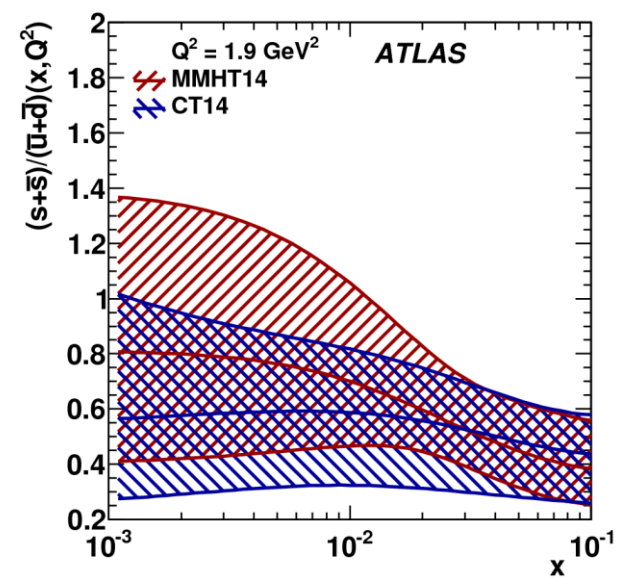
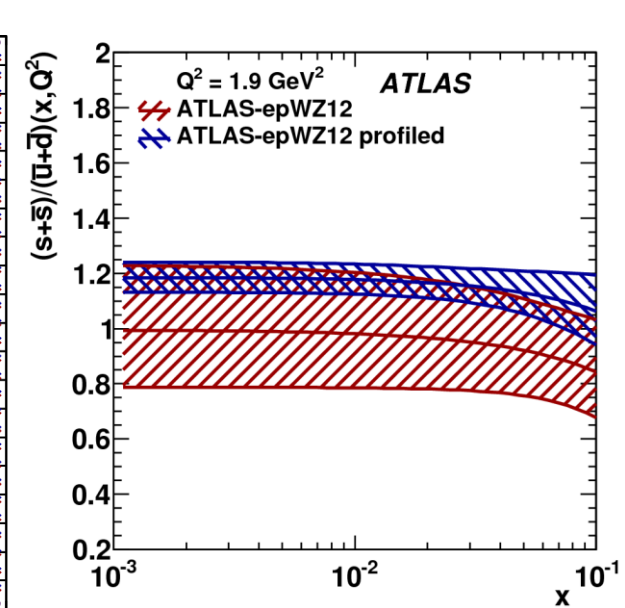
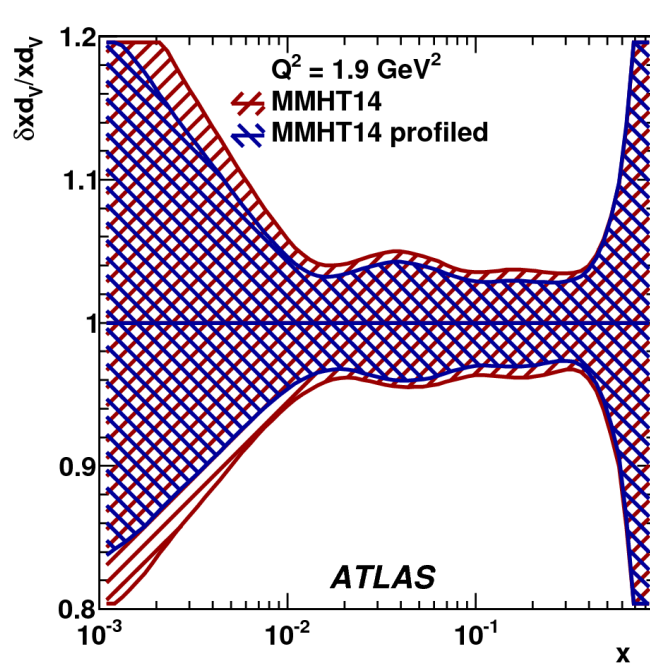
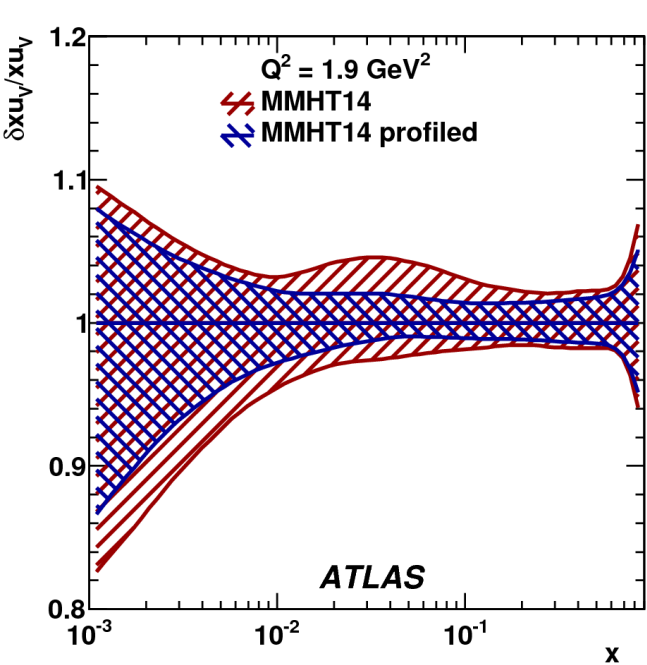
(An example may be the Higgs signal simulation with pdf4lhc15 for which an interest had been raised).

For NLO generators it is better to store weights for some alternative PDF sets. Recommended to add (NNPDF3.0 case given in parentheses):

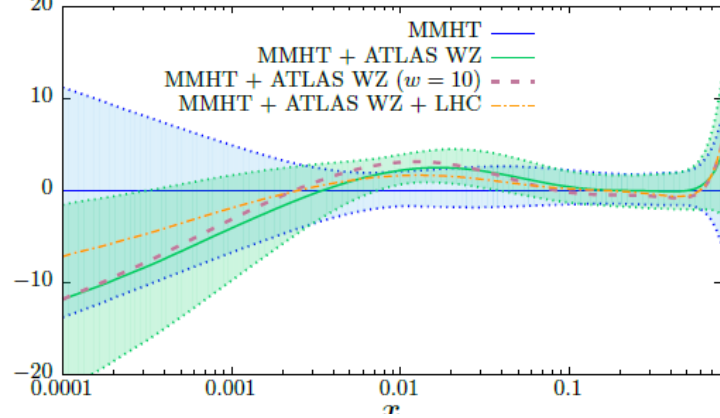
- PDF internal error sets (all error sets for NNPDF3.0 at $\alpha_s = 0.118$)
- α_s variations (NNPDF3.0 central at $\alpha_s = 0.117, 0.119$)
- Central values of the other main families (CT14, MMHT 2014)

Alternatively one could use the PDF4LHC15 PDF sets for generation, (preferably the 100 error set versions unless it has been shown that using the 30 set versions makes no difference), and then add the central value of all of the contributing families.

extras



Up valence (NNLO), percentage difference at $Q^2 = 10^4 \text{ GeV}^2$



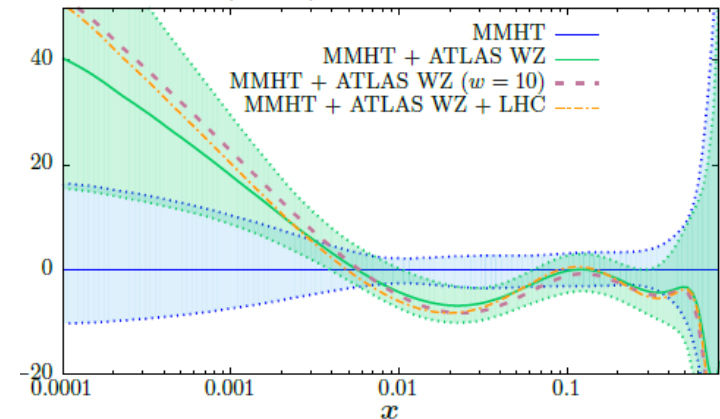
So look at the change from MMHT14 when our W,Z data is included

See the increased strange

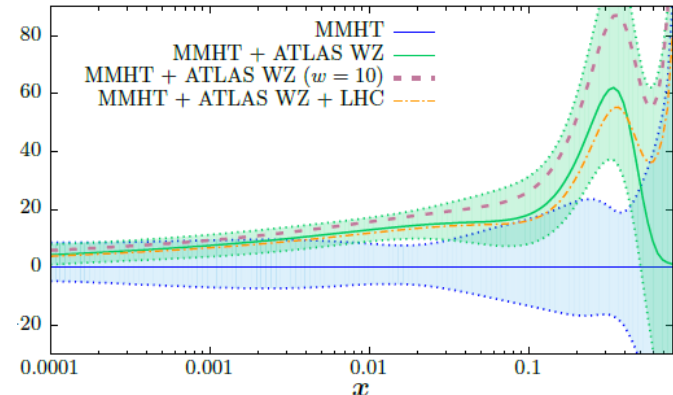
Plus changes in u_v, d_v and Sea

Little impact on the gluon so this is not shown

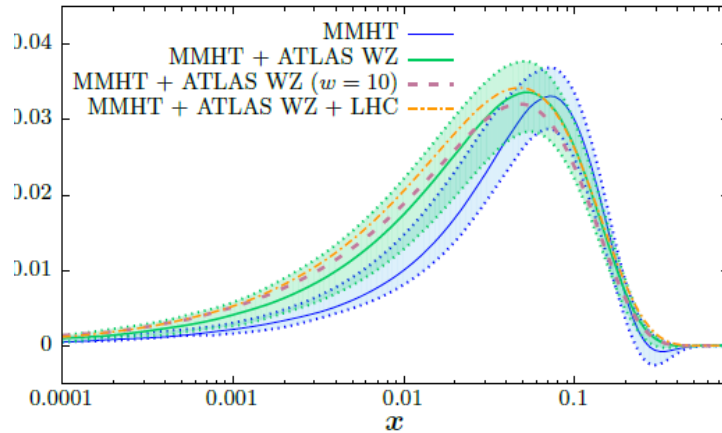
Down valence (NNLO), percentage difference at $Q^2 = 10^4 \text{ GeV}^2$



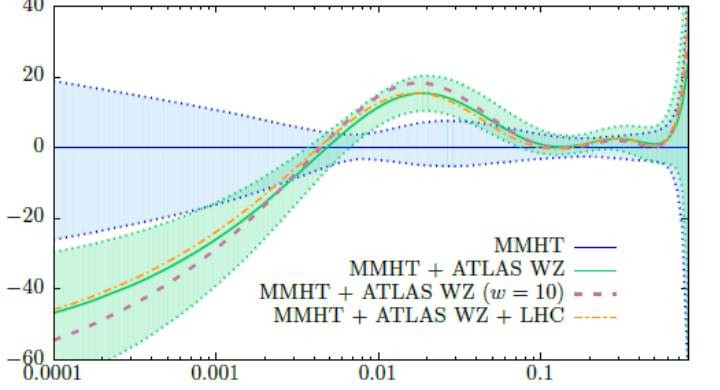
$x(s + \bar{s})$ (NNLO), $Q^2 = 10^4 \text{ GeV}^2$



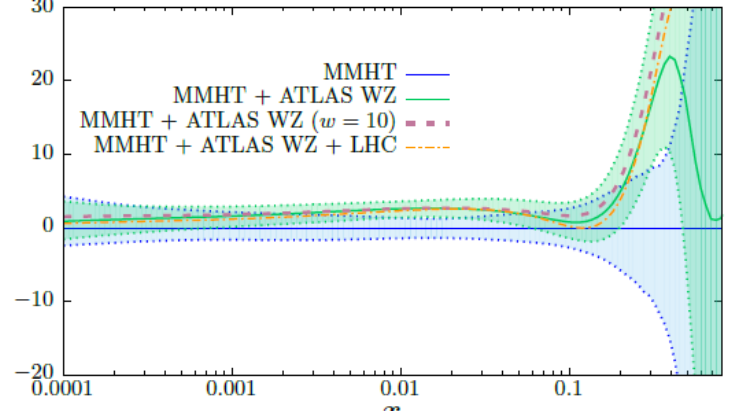
$x(\bar{d} - \bar{u})$ (NNLO), $Q^2 = 10^4 \text{ GeV}^2$



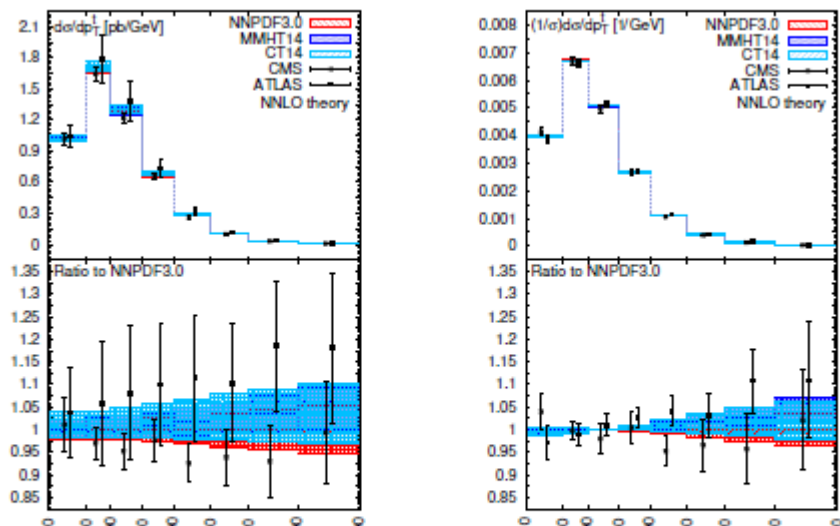
$x(u_v - d_v)$ (NNLO), percentage difference at $Q^2 = 10^4 \text{ GeV}^2$



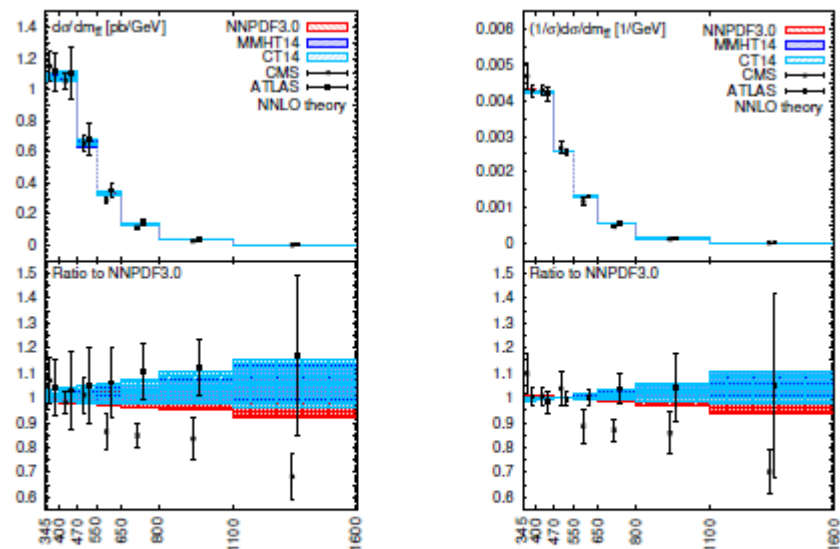
Light sea (NNLO), percentage difference at $Q^2 = 10^4 \text{ GeV}^2$



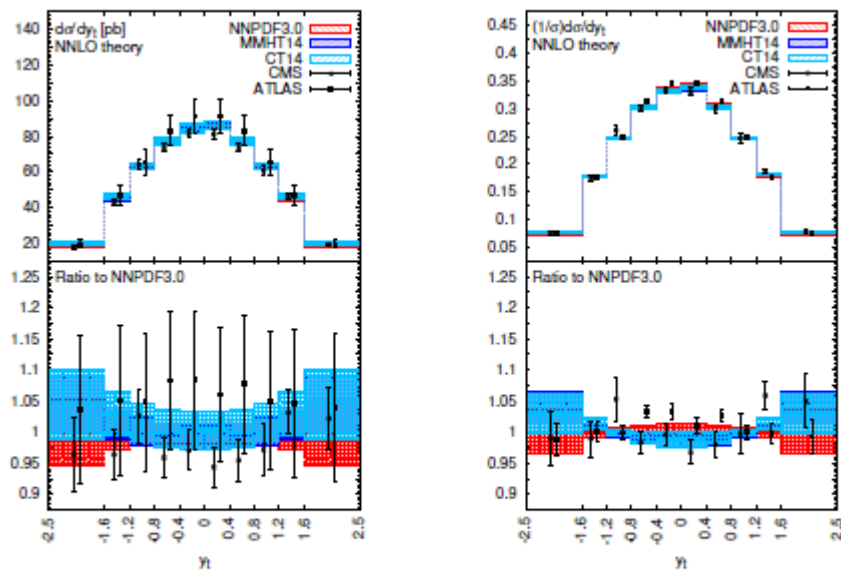
Data/Theory comparison: p_T^t



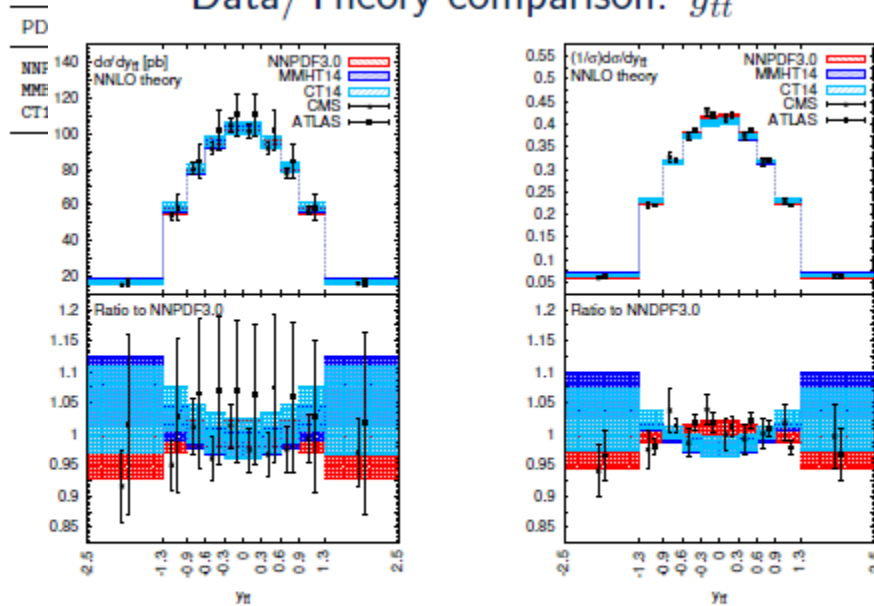
Data/Theory comparison: $m_{t\bar{t}}$



Data/Theory comparison: y_t



Data/Theory comparison: $y_{t\bar{t}}$

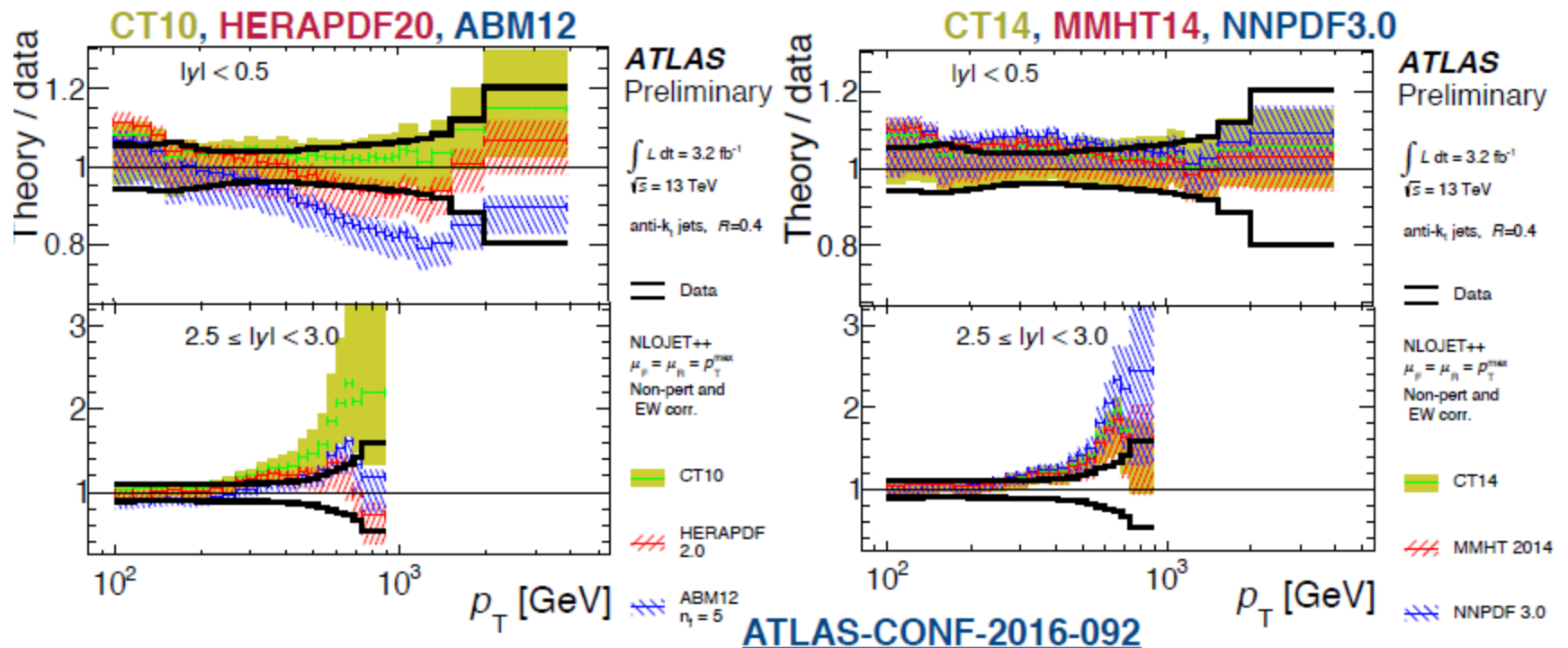


PDF set	ATLAS $d\sigma/dy_t$	CMS $d\sigma/dy_t$	ATLAS $(1/\sigma)d\sigma/dy_t$	CMS $(1/\sigma)d\sigma/dy_t$
NNPDF3.0	0.73 (0.28)	3.04 (1.05)	4.06 (2.85)	3.29 (1.49)
MMHT14	1.36 (0.29)	2.12 (0.98)	12.1 (6.82)	2.40 (1.09)
CT14	1.28 (0.20)	2.23 (1.47)	10.3 (5.71)	2.33 (0.96)

PDF set	ATLAS $d\sigma/dy_{t\bar{t}}$	CMS $d\sigma/dy_{t\bar{t}}$	ATLAS $(1/\sigma)d\sigma/dy_{t\bar{t}}$	CMS $(1/\sigma)d\sigma/dy_{t\bar{t}}$
NNPDF3.0	0.84 (0.21)	0.99 (0.74)	3.59 (1.48)	1.17 (0.75)
MMHT14	2.36 (0.29)	2.27 (1.52)	15.6 (5.49)	3.33 (2.10)
CT14	2.69 (0.19)	1.88 (1.67)	12.7 (5.26)	2.53 (1.51)

- Similar comparisons at 13 TeV using various PDFs
 - Better inter-PDF agreement between CT14, NNPDF3.0, and MMHT14 (right)
- Reduced low- p_T non-perturbative corrections (partially driven by AU2→AU14)

Theory-data ratios



Recommendations for the use of PDFs

<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/StandardModelPDF>

Comment on the [PDF4LHC](#) recommendation. The original recommendation 2011 was to use an envelope of the 'big three' global PDF sets: CTEQ, MSTW, NNPDF. This has been replaced in 2015 by the [PDF4LHC15](#) PDF sets which are a statistical combination of the most recent versions of these three. If an analysis used the recommendation before it is appropriate to update to the new recommendations. Have an eye on your intended accuracy. For processes like **Higgs** which access the middle of the LHC kinematic plane the [PDF4LHC_30](#) set is enough, for processes which are on the limits of x_{bjorken} the [PDF4LHC_100/MC](#) sets are better. (A hessian set and an MC set is available) Note the [PDF4LHC](#) recommendations were never meant for ALL cases, see our recommendations below.

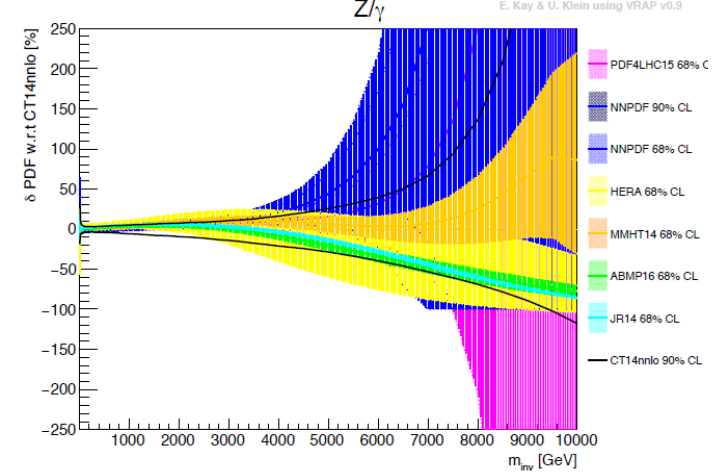
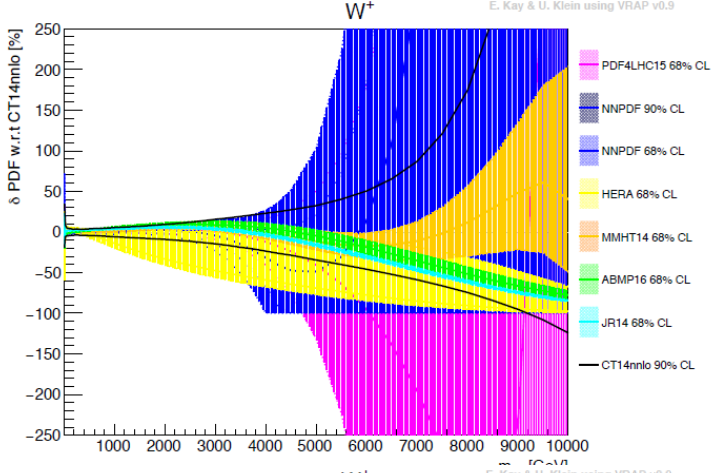
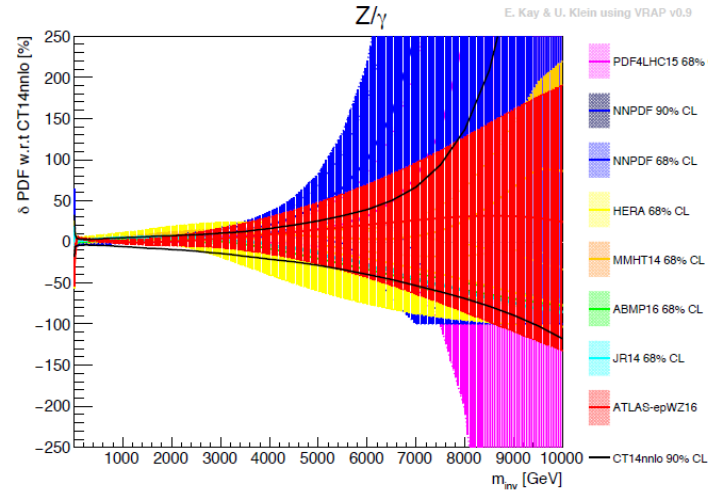
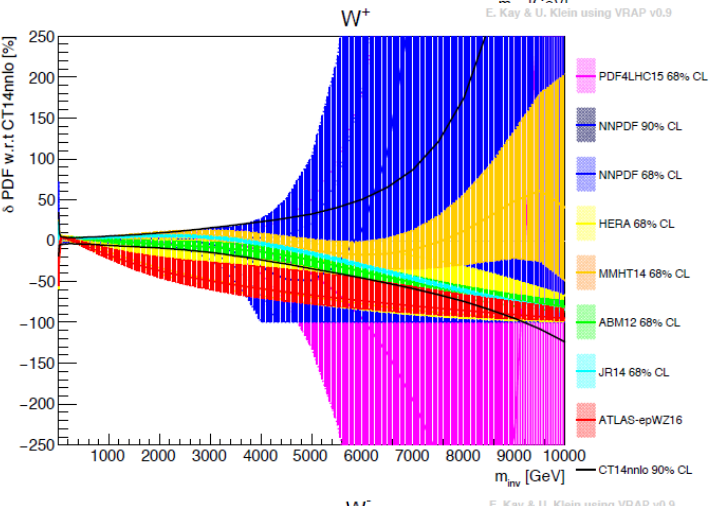
0. Comparison of precision data and theory Compare to all modern PDFs: ABM12, CT14, HERAPDF2.0, MMHT14, NNPDF3.0

1. Acceptance and background calculations. For these PDF uncertainties are usually sub-dominant. Use the PDF which was used in the MC generation. If this was an LO set without uncertainties, then perform PDF reweighting to a more modern LO set which has uncertainties. If several MC samples with different PDFs were used consider reweighting to a common PDF.

2. Limit calculations. For these PDF uncertainties can be dominant. Careful consideration on a case by case basis is needed. Consider more unusual PDFs such as HERAPDF and ABM as well as the 'big three' NNPDF, MMHT and CT. You can save time by using the appropriate [PDF4LHC15](#) set instead of these three different PDFs, see above

3. PDFs for MC generation We need to use a set which is successfully tested against ATLAS data/ used in MC tunings/ exists in LO, NLO and NNLO/ includes the photon PDF where needed. This selected NNPDF30 last year, but things move on. Requests have been made for generation with the [PDF4LHC](#) PDFs and MC tunes for them are under consideration.

Weights are included with the NNPDF3.0 central set to allow reweighting for i) the 100 replicas for uncertainties; ii) scale variation, iii) alphas variation iv) variation to MMHT14 or CT14 central

Z/γ  W^+  Z/γ  W^+  W^- 