MMHT2014 PDFs and connection to Xfitter

Robert Thorne

March 20th, 2016

University College London

In collaboration with Lucian Harland-Lang, Ricky Nathvani and Alan Martin

and thanks to Patrick Motylinski, Ben Watt, Graeme Watt and James Stirling

I will cover a number of topics.

– The impact of fitting new LHC and Tevatron data. $\rightarrow$ clear improvements in some PDF uncertainties.

Specific issues with ATLAS 7 TeV jet data. Some discussion of NNLO effects.

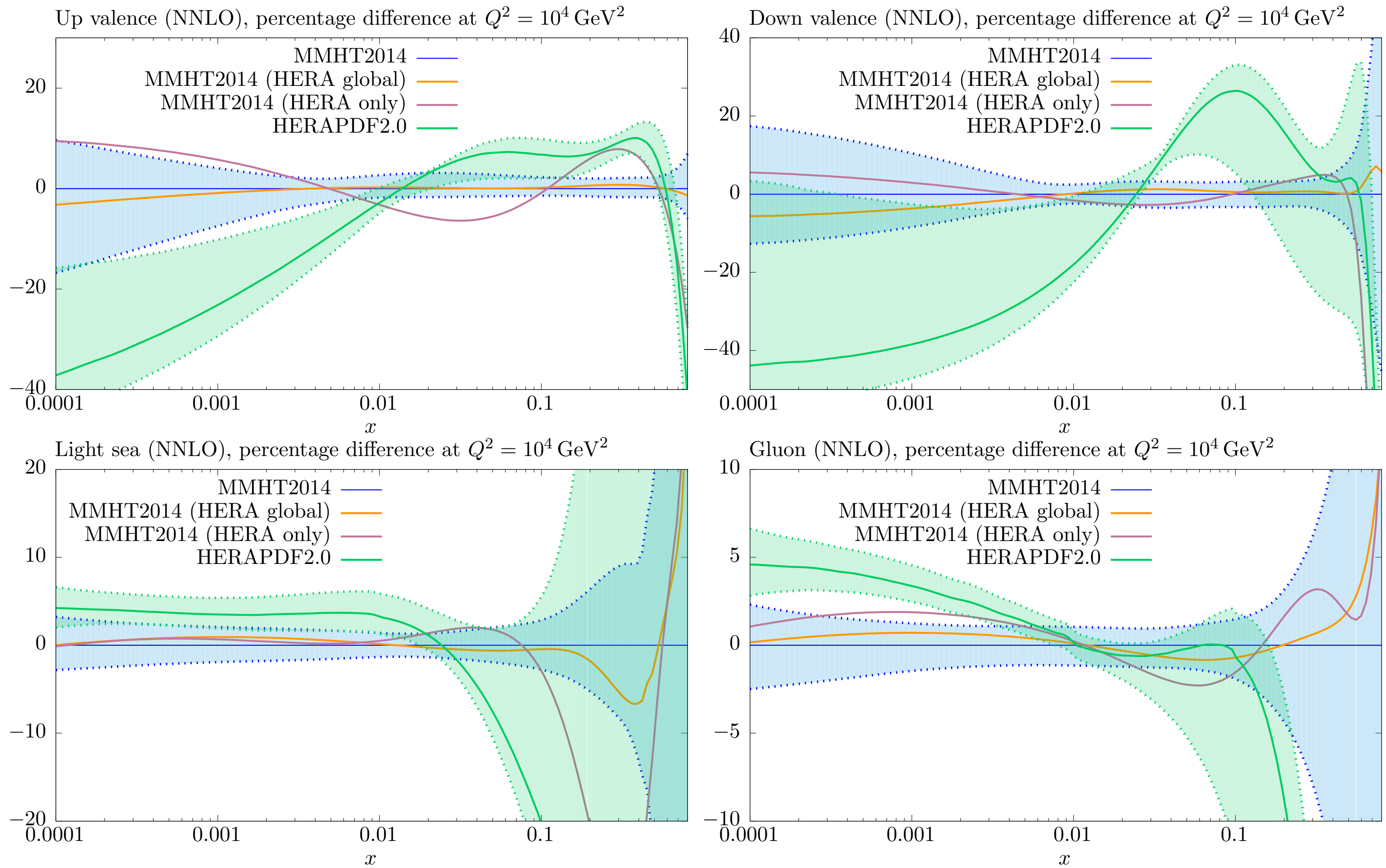
Preliminary results with inclusion of final ATLAS 7 TeV W, Z rapidity dependent data.

A brief intro to topics just starting investigation. Extension of parameterisation and QED corrections.

Connection of MSTW/MMHT to xFitter

HERA II Combined data

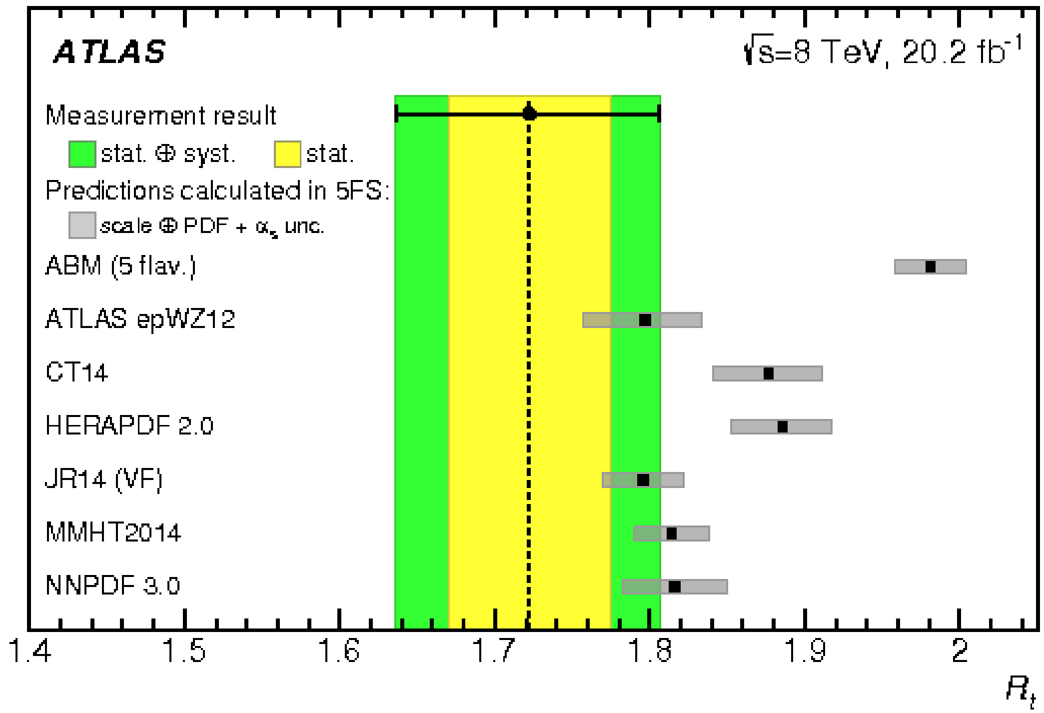
Fit quality in global fit at NNLO $Q^2 > 2.5 \sim 1435/1168$



HERA II modified PDFs very well within MMHT2014 uncertainties. PDFs from HERA II data only fit in some ways similar to HERAPDF2.0.

Up to 10% reduction in uncertainties. Small changes in central values.

Note that highish-x enhancement in up quark preferred by HERA e^- charged-current data in tension with recent most accurate measurement of single top ratio.



Also disfavours any other reason for enhanced $u(x)/d(x)$ for $x \sim 0.1$.

Breakdown of fit quality to new hadron collider data

We now also fit to high rapidity W, Z data from LHCb at 7 and $8\ \mathrm{TeV}$, $W + c$ jets from CMS, which constrains strange quarks, high precision CMS data on $W^{+,-}$ rapidity distributions which can also be interpreted as an asymmetry measurement, and also the final e asymmetry data from D0 (lepton, not W asymmetry).

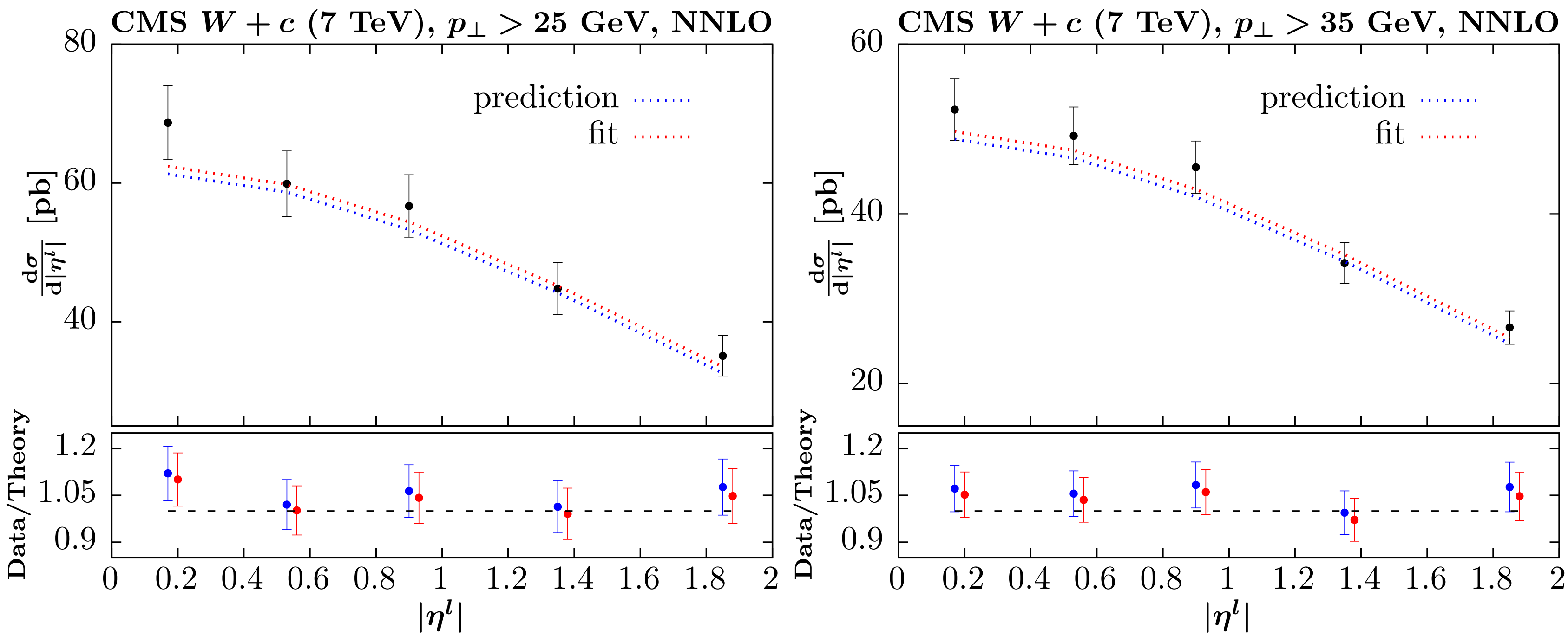
	no. points	NLO χ^2_{pred}	NLO χ^2_{new}	NNLO χ^2_{pred}	NNLO χ^2_{new}
$\sigma_{t\bar{t}}$ Tevatron +CMS+ATLAS	18	19.6	20.5	14.7	15.5
LHCb 7 TeV $W + Z$	33	50.1	45.4	46.5	42.9
LHCb 8 TeV $W + Z$	34	77.0	58.9	62.6	59.0
LHCb 8TeV e	17	37.4	33.4	30.3	28.9
CMS 8 TeV W	22	32.6	18.6	34.9	20.5
CMS 7 TeV $W + c$	10	8.5	10.0	8.7	8.0
D0 e asymmetry	13	22.2	21.5	27.3	25.8
total	3738/3405	4375.9	4336.1	3741.5	3723.7

Predictions good, and no real tension with other data when refitting, i.e. changes in PDFs relatively small. Slightly (~ 10 units) better than previous report due to improvements (and one correction) in K-factors.

At NLO $\Delta\chi^2 = 9$ for the remainder of the data and at NNLO $\Delta\chi^2 = 8$.

When couplings left free at NLO $\alpha_S(M_Z^2)$ stays very close to 0.120 but at NNLO $\alpha_S(M_Z^2)$ marginally above 0.118, higher than MMHT2014.

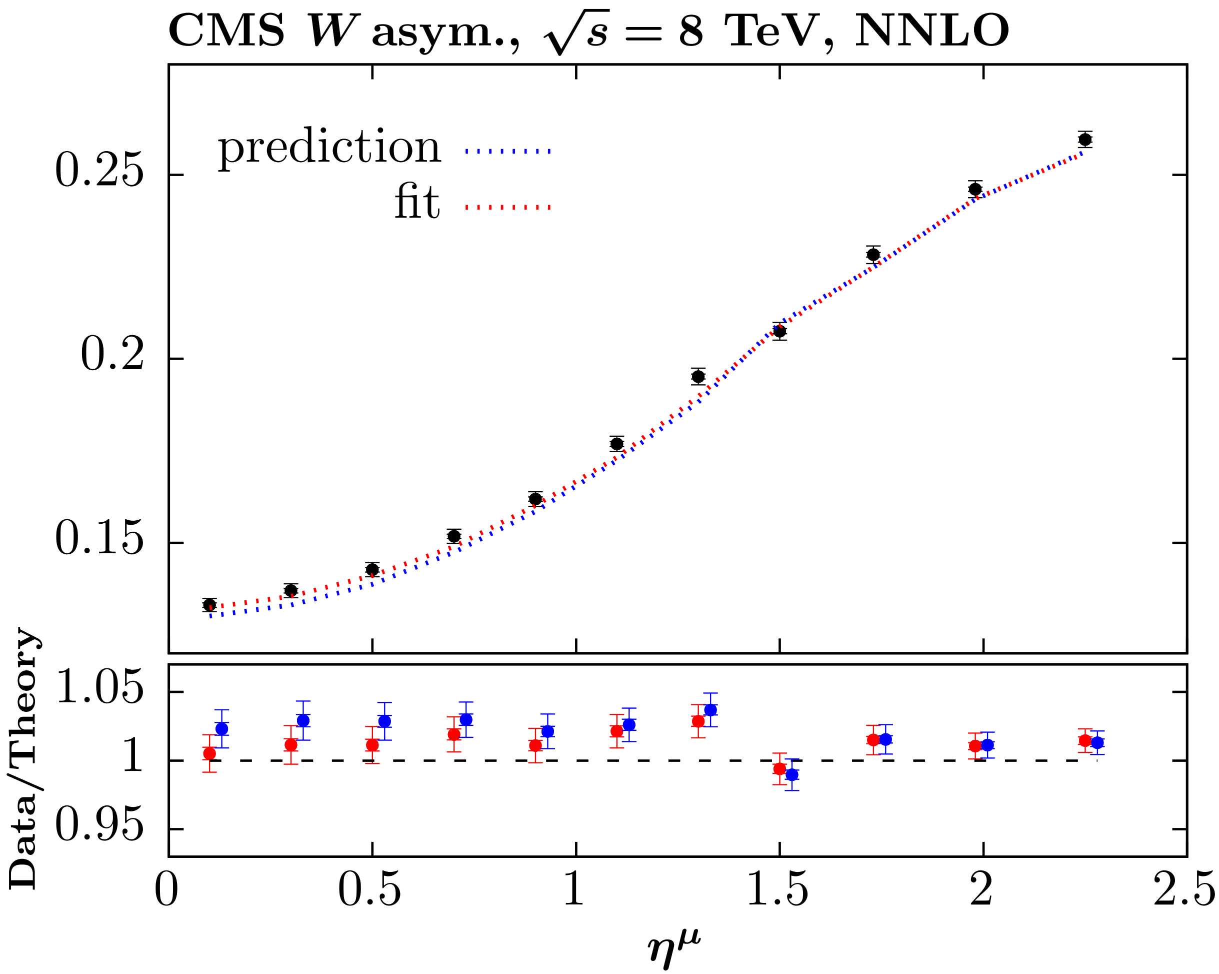
New data sets for fit – $W + c$ differential distributions.



Data on plot use uncertainties added in quadrature.

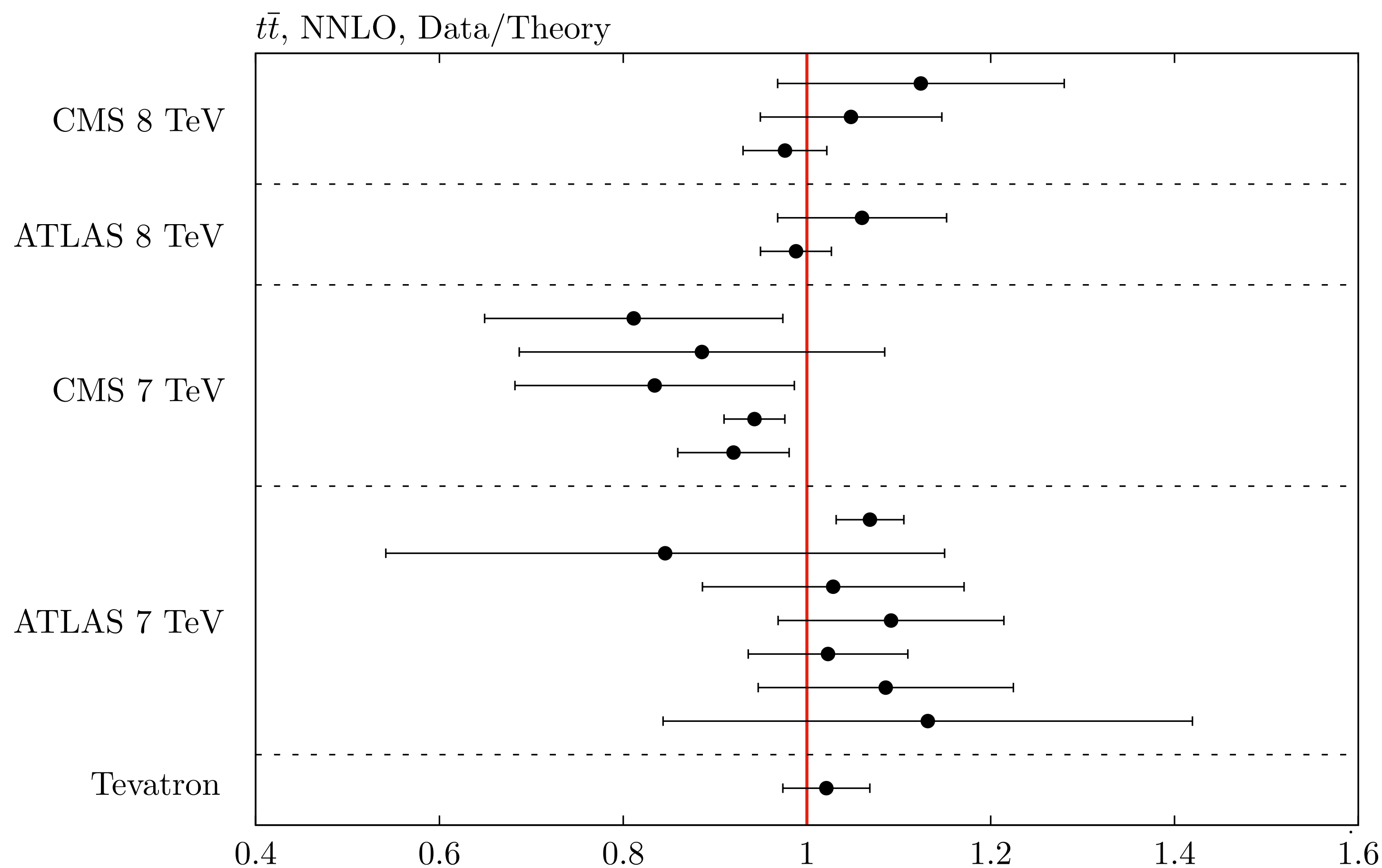
Very little change after fit. By eye comparison looks worse, but slightly better when covariance matrix used (as in fit).

Good agreement with new 8 TeV CMS $W^{\pm}$ rapidity and asymmetry data (shown). (Fit to individual distributions not asymmetry.)



Small-x valence quarks require some modification of order the size of uncertainty. Scope for reduced uncertainty with new data inclusion.

Included some more up-to-date results on $\sigma_{\bar{t}t}$.



Fit very good and with $\alpha_S(M_Z^2) = 0.118$ the fitted $m_t^{pole} = 173.4$ GeV. At NLO $m_t^{pole} = 170.2$ GeV. MMHT values $m_t^{pole} = 174.2$ GeV and $m_t^{pole} = 171.7$ GeV

Helps drive slight increase in $\alpha_S(M_Z^2)$

PDF sets generated

We generate a preliminary (not for distribution) central set at NLO and NNLO for fit to new data – labelled MMHT (2016 fit).

Also generate PDF eigenvector sets for uncertainties at NNLO.

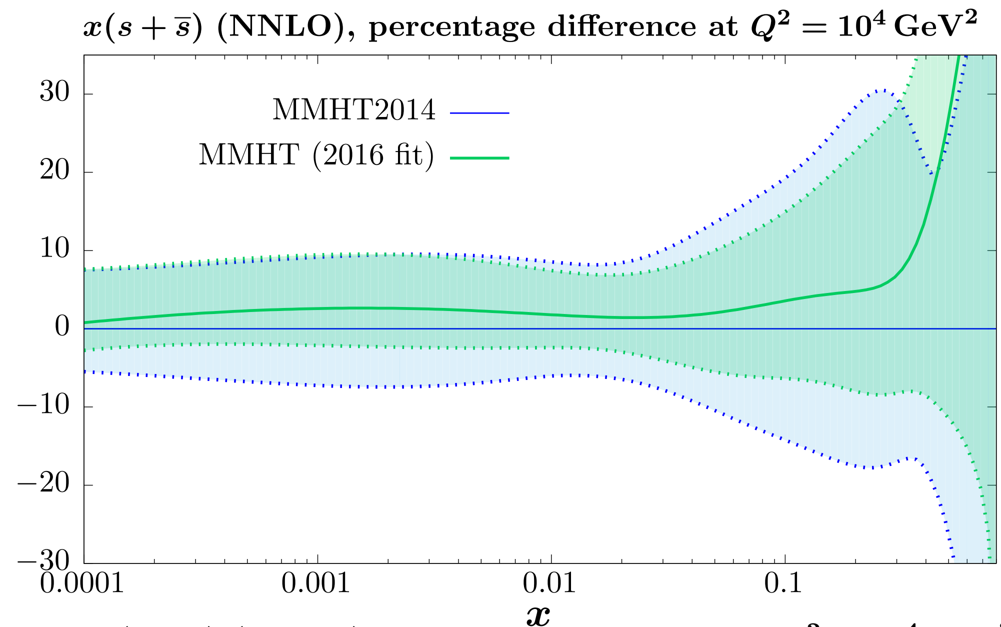
Use same basis of 25 free PDF parameters as in MMHT2014 (this is subject to possible change in the future).

Hence, 50 eigenvector directions.

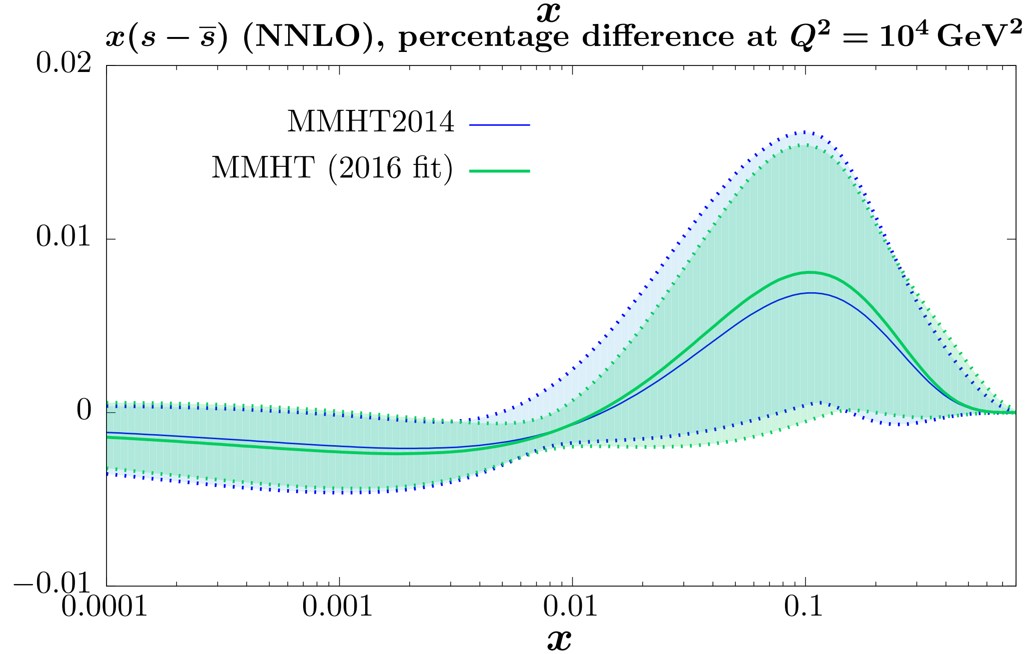
14 of these are best constrained by one of the new (LHC) data sets, CMS $8\ \mathrm{TeV}\ W$ data and $W + c$ jets and the new LHCb data.

Effect on PDFs

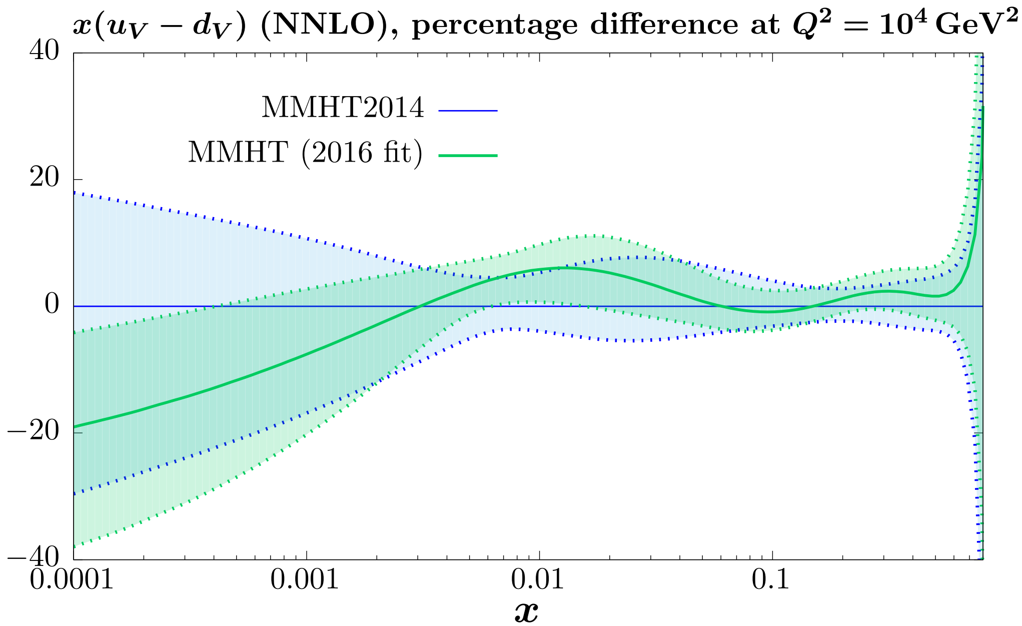
Large reduction in the $s + \bar{s}$ uncertainty, but little change in central value. Due to $W + c$ jets data.

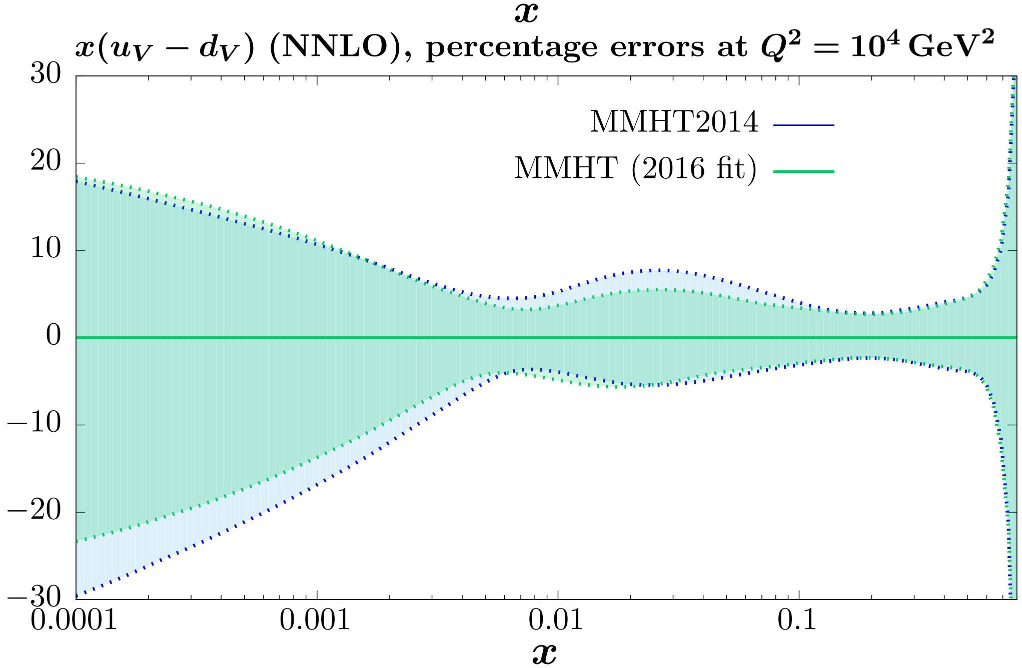


There is some impact on the $s - \bar{s}$ uncertainty, from (effective) asymmetry data.



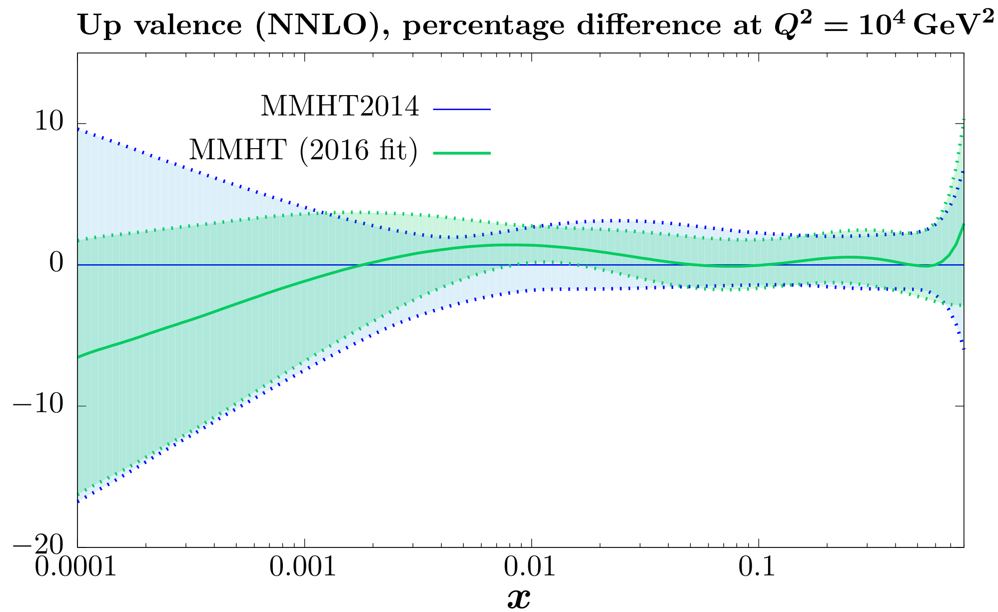
A significant change in $u_v - d_v$, and reduction in the uncertainty, from (effective) CMS asymmetry data.

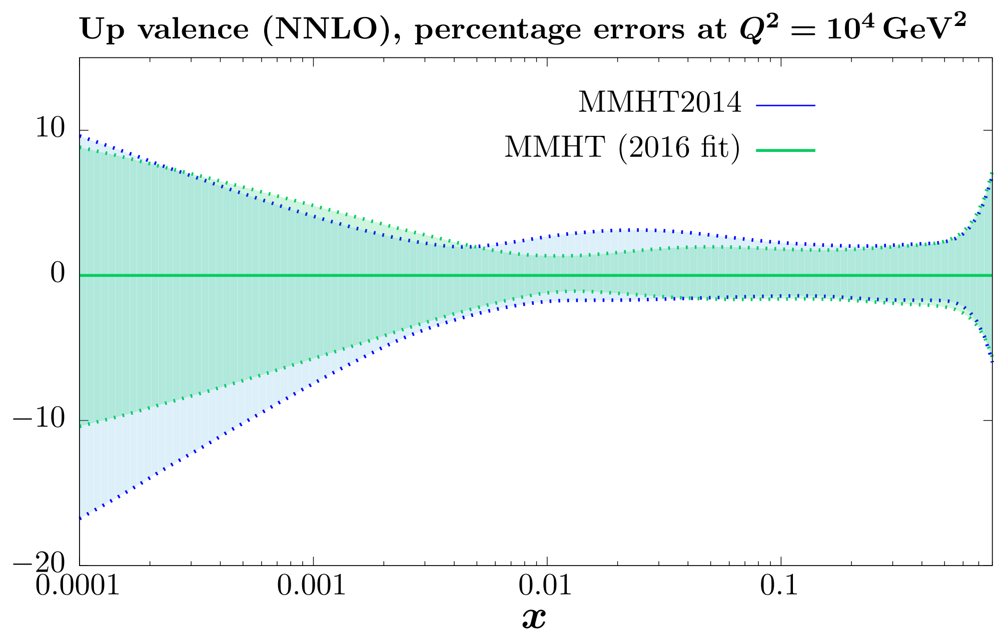




Main change and uncertainty reduction in u_v rather than d_v.

Mainly CMS data, but some impact of new HERA data.





Attempted fit to high luminosity ATLAS $7\ \mathrm{TeV}$ 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 $\chi^2/N_{pts} = 413.1/140$.

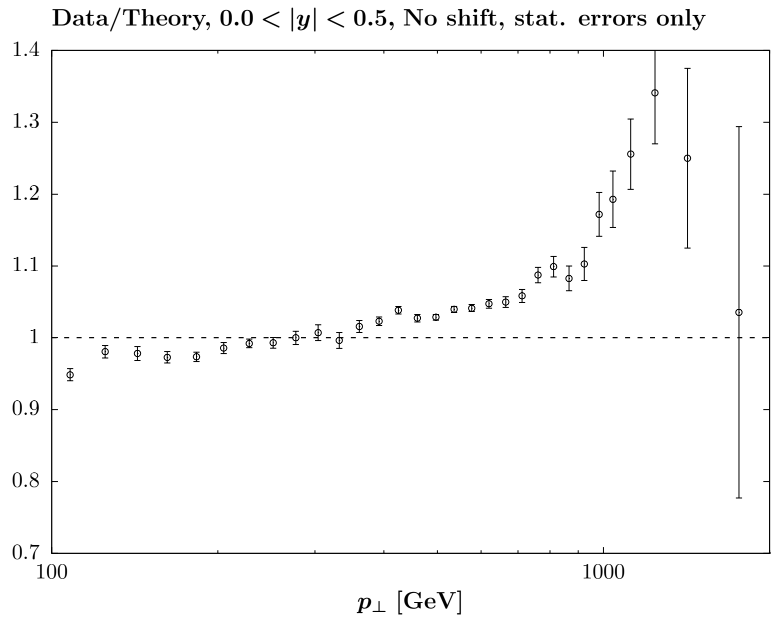
Refit gives improvement only to $\chi^2/N_{pts} = 400.4/140$.

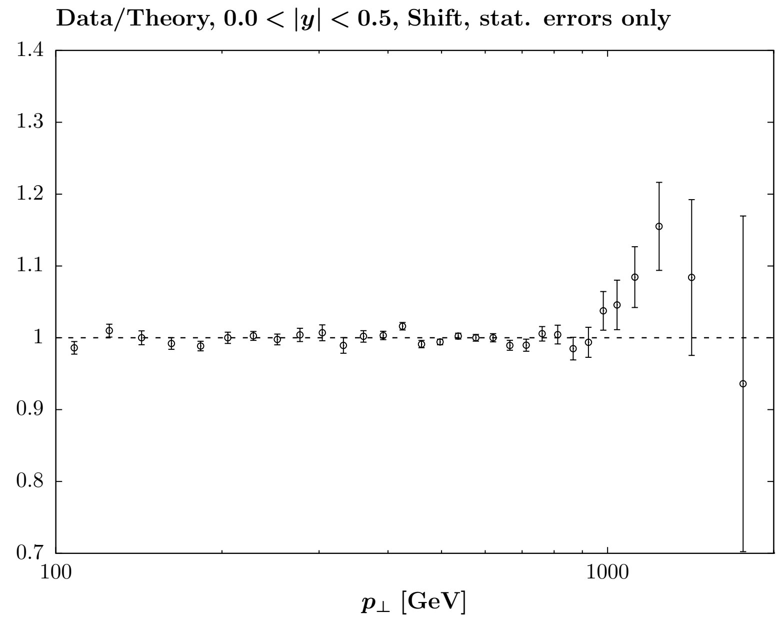
Deterioration in other data only $\Delta\chi^2 \sim 3$, so failure not due to strong tensions.

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.

Similar results also seen by other groups.

Large correlated systematics. Dominate over uncorrelated uncertainties.



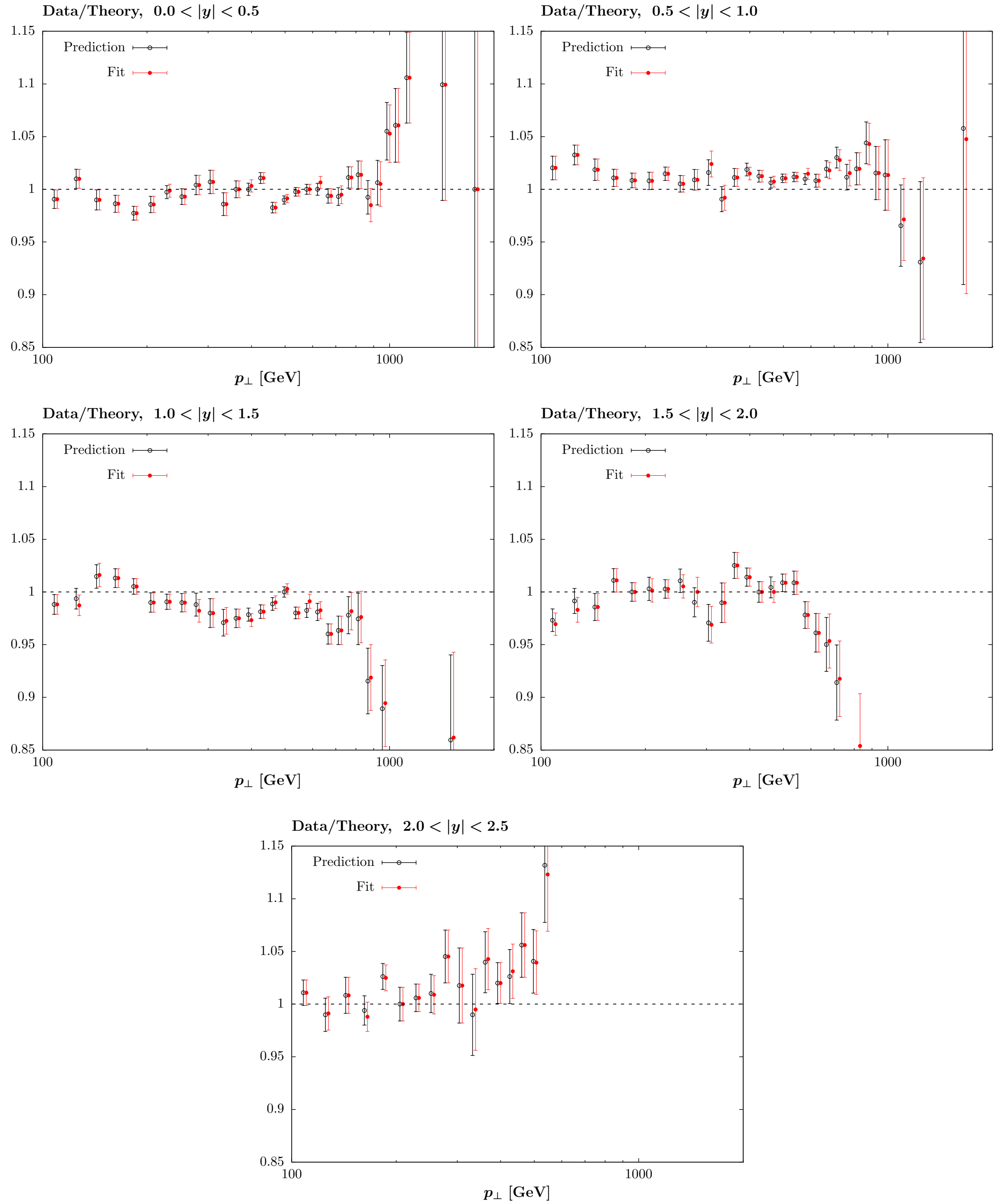


For “weak” assumption about correlations there are 71 correlated systematic uncertainties, 65 related to jet energy scale.

Best fit requires shift in data against theory. Highly correlated from one rapidity bin to another.

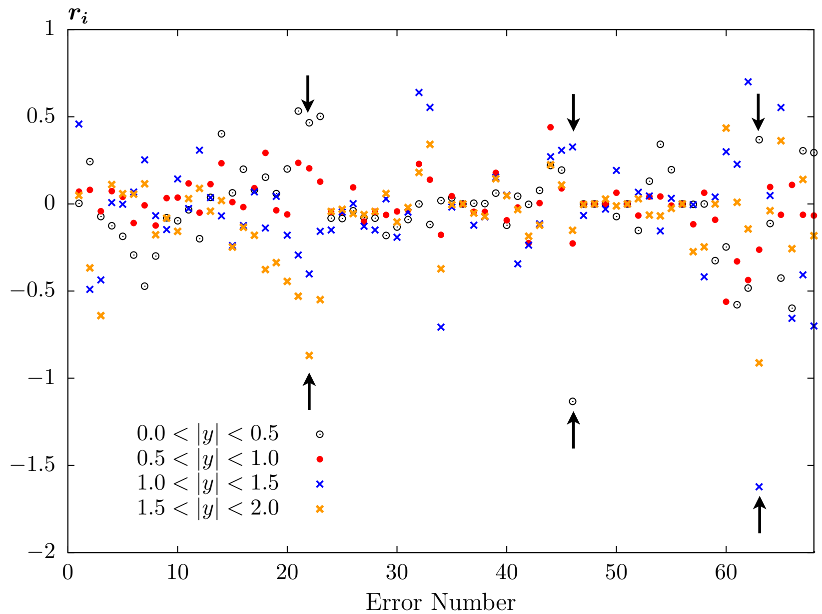
Cannot simultaneously fit data in all bins. Mismatch in one rapidity bin different to others probing PDFs of similar flavour, x and Q^2.

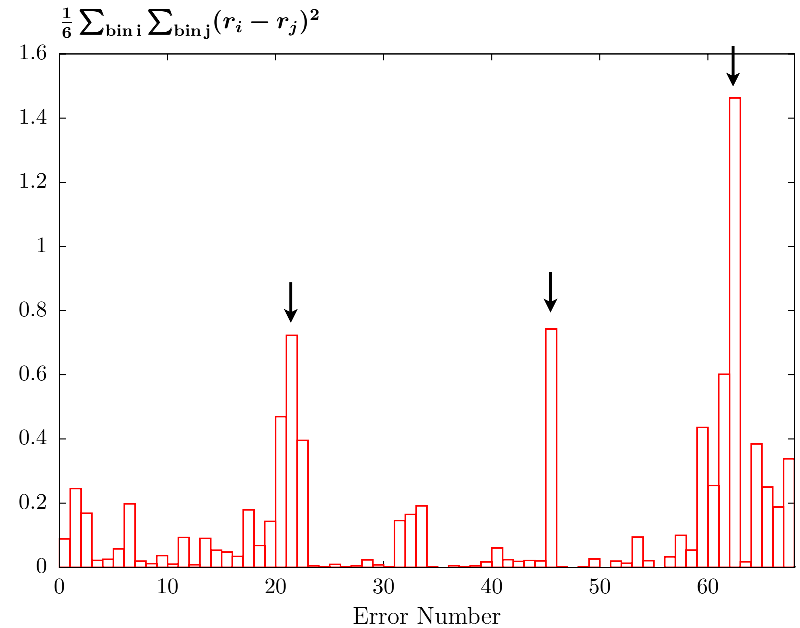
Similar results also seen by other groups.



Good fit ($\chi^2/N_{pts} \sim 1$) possible when fitting all individual rapidity bins.

Look at shifts due to each source of correlated uncertainty, i.e. preferred r_k when fitting each separate rapidity bin.





A small number of sources prefer very different values when fits to different bins performed.

These are jes21, jes45 and jes62 (Multi-jet balance asym., JES pile-up and JES close by jets 1406.0076).

Exercise on decorrelating uncertainties

We consider the effect on the χ^2 of the simultaneous fit to all data of decorrelating these three error sources, i.e. making them independent between the 6 rapidity bins.

Compared to the original $\chi^2/N_{pts} = 400/140$ we get instead

	21	45	62
χ^2	221	316	330

	21,45	21,62	45,62
χ^2	213	178	230

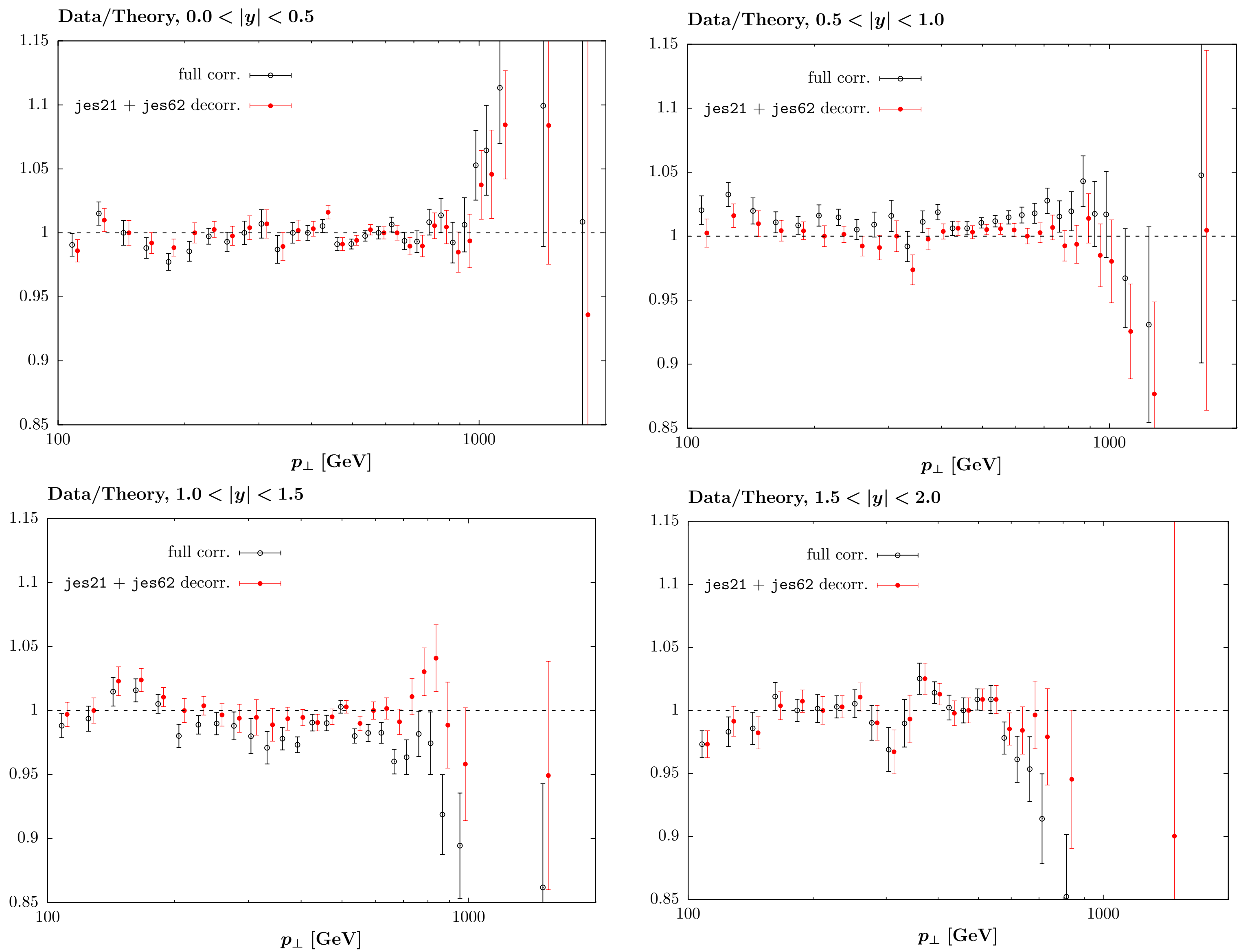
	21,45,62
χ^2	172

Very significant improvement, particularly from decorrelating jes21.

Little improvement if jes45 decorrelated on top of jes21 and jes62.

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

Anything other than just an interesting observation?



Fit quality evidently reasonable in all rapidity bins when jes21 and jes 62 decorrelated between $|y|$ bins.

The effect of NNLO corrections.

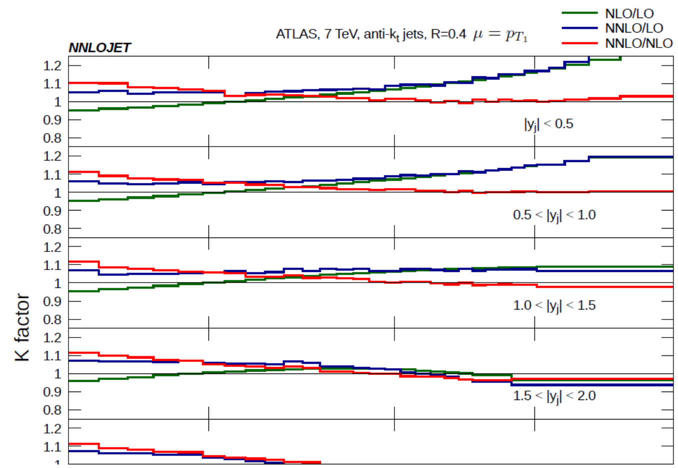
The NNLO corrections are now completed (arXiv:1611.01460) Currie *et al.*. Explicit K-factors available for ATLAS 7 TeV data.

Does this make a significant change to the conclusions just presented for the NLO fit?

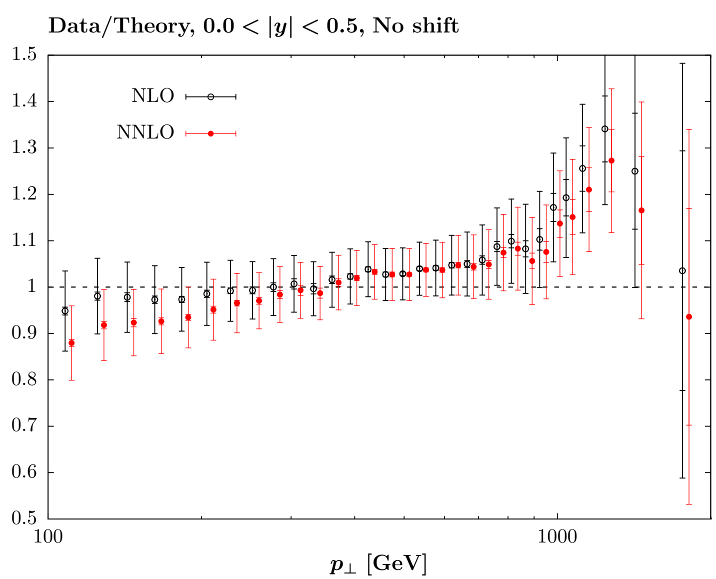
Fit to MMHT data set with HERA data updated to final combined set and Tevatron jet data excluded (usually fit using NNLO threshold approximation).

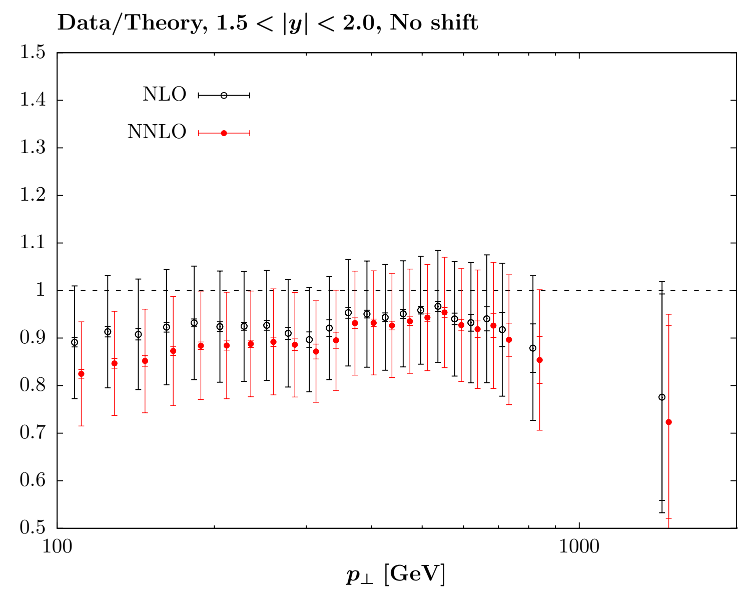
As at NLO take as default $R = 0.4$ and $\mu = p_{T,1}$.

At low p_T NNLO correction significant ($\mathcal{O}(10\%)$) and positive.



Moves unshifted data and theory further apart. Expect worse fit quality?





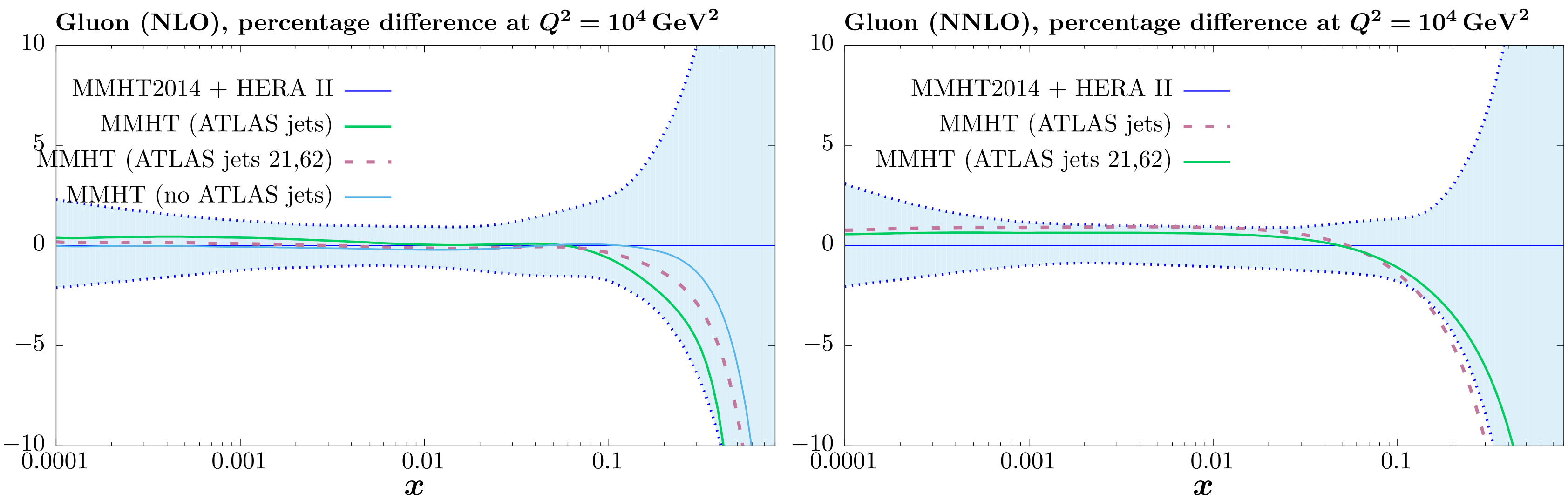
Values of χ^2 without (with) the ATLAS jets data in the fit

	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. Not an issue of tension with other data.

By eye fit quality looks very similar to NLO but slightly higher penalty from shifts.

Gluon including ATLAS jet data at NLO and NNLO



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

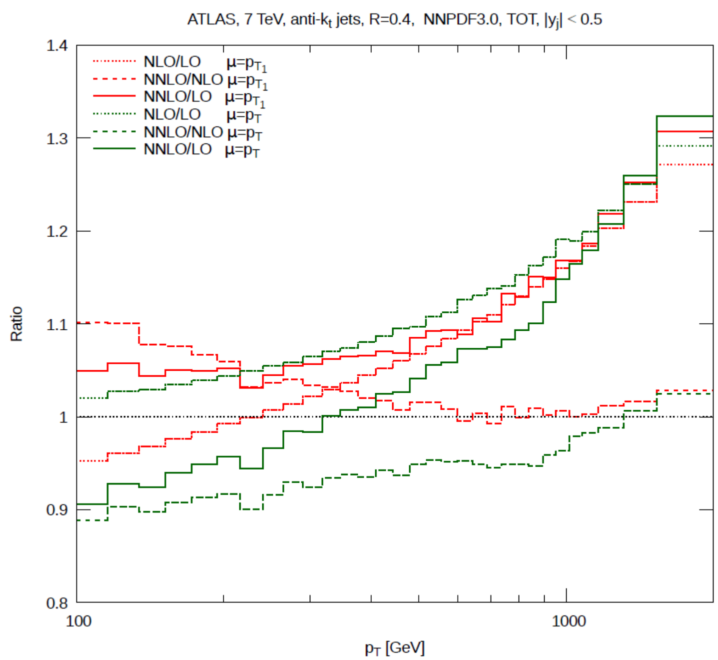
Slightly smaller effect at NLO than at NNLO.

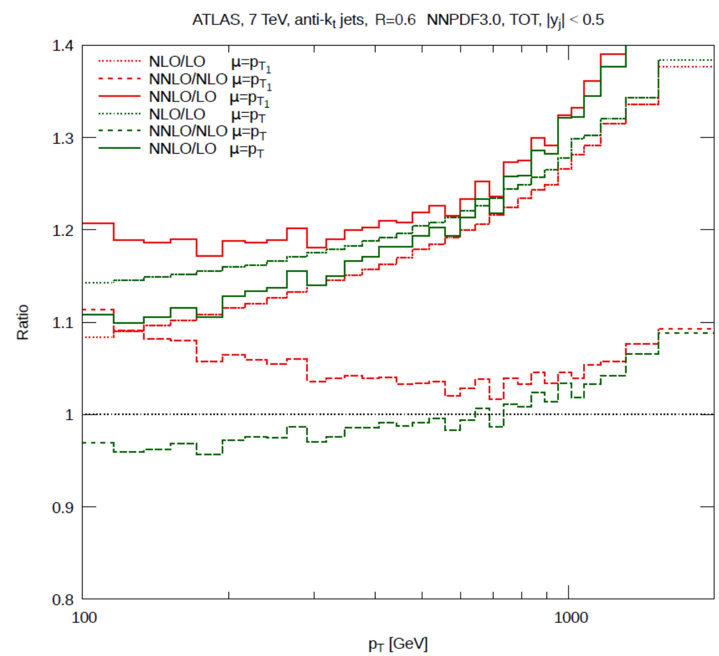
Not very dependent on whether jes21 and jes 62 are decorrelated or not.

CMS 7 TeV jets.

Here take as default $R = 0.7$ and $\mu = p_T$.

Larger R, and $\mu = p_T$ rather than $\mu = p_{T1}$, lead to more stable NNLO corrections.





Therefore good NLO fit ($\chi^2/N_{pts} = 138/133$) very likely to be maintained at NNLO and little change in gluon expected.

More stability from NLO to NNLO expected for ATLAS jets if larger R and different scale chosen for fit?

Inclusion of high precision ATLAS W, Z data arXiv:1612.03016

Confirm we obtain $\chi^2/N_{pts} \sim 400/61$ from MMHT14 PDFs at NNLO.

For slightly modified PDFs with final HERA combined data (and some more $\sigma_{t\bar{t}}$ points) obtain $\chi^2/N_{pts} \sim 387/61$. Use this as our "baseline". Essentially the same PDFS as from MMHT + final HERA data.

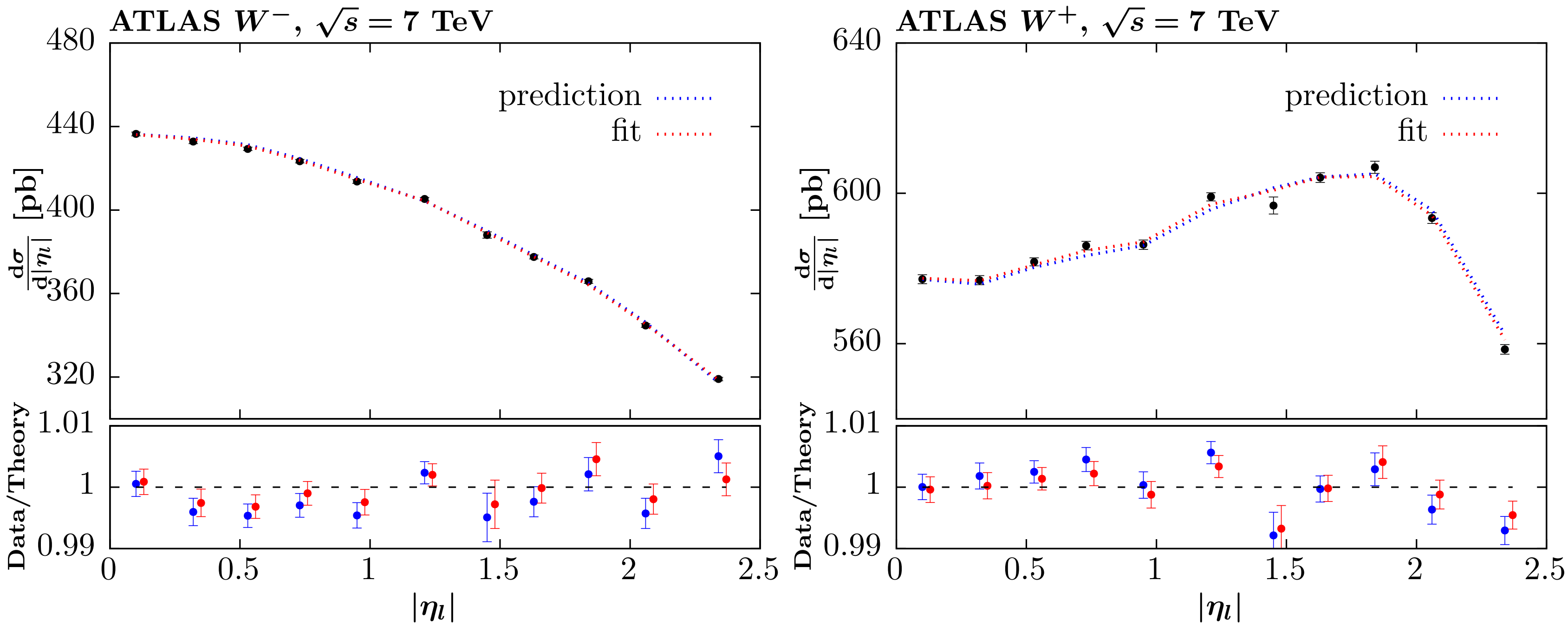
Including ATLAS W, Z data in fit goes to $\chi^2/N_{pts} \sim 130/61$, similar to ATLAS profiling. Use this as basis for study of effects on PDFs.

Deterioration in fit to other data $\Delta\chi^2 = 54$. Worst for CMS double differential Z/γ data ($\Delta\chi^2 = 17$) and CCFR/NuTeV dimuon data ($\Delta\chi^2 = 16$). For latter branching ratio requires 25% shift, but has uncertainty of 15%.

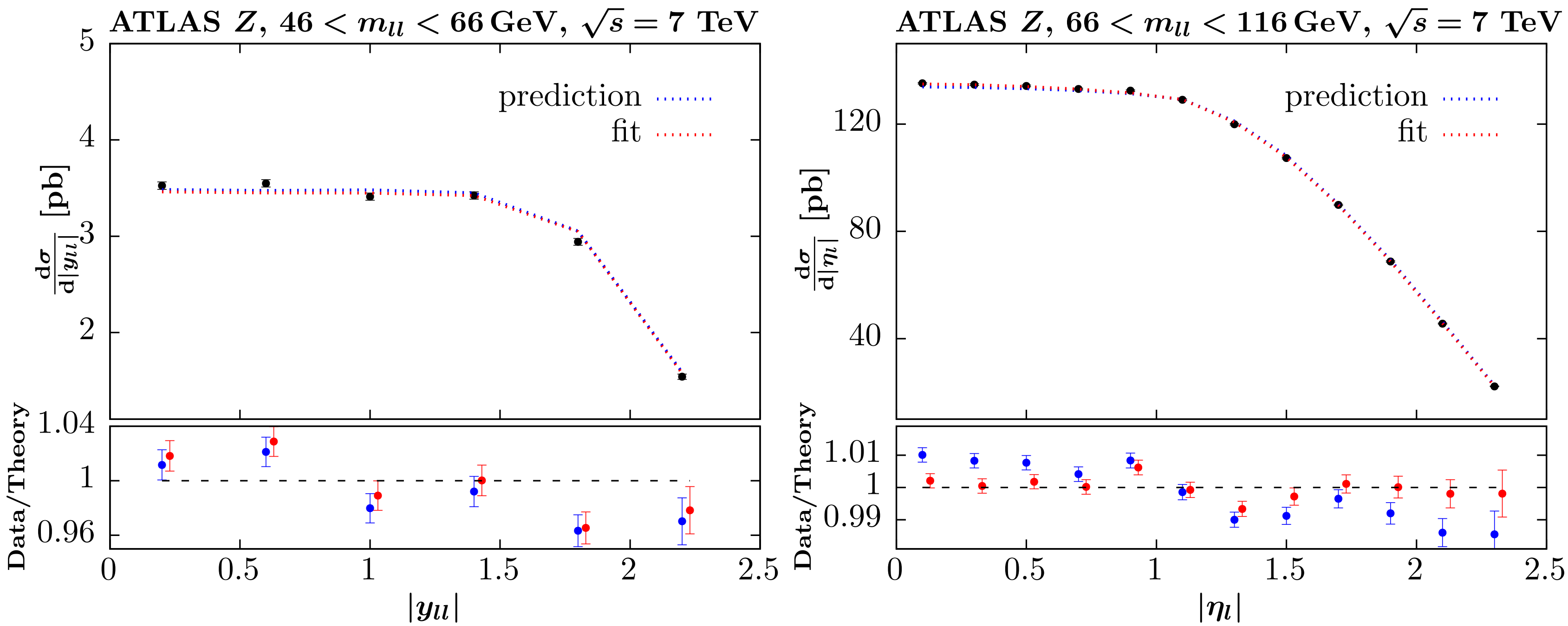
Other deterioration in fixed target DIS data, E866 Drell-Yan asymmetry and CDF W-asymmetry.

Generate PDFs with uncertainties including new ATLAS W, Z data, using same basis for 25 eigenvectors as for MMHT2014. ATLAS W, Z data constrains 5 eigenvector directions.

Prediction and Fit to data



Slight reduction in lower $|\eta|$ W^- required and opposite for W^+.



Significant change in shape required for Z production, Higher at low $|\eta|$ and lower at high $|\eta|$

Even with fit difficulty in shape for lower mass data.

Additional fits with high precision ATLAS W, Z data.

Increase weight of new ATLAS data by factor 10.

χ^2 improves to $\chi^2/N_{pts} \sim 121/61$. Deterioration in fit to other data $\Delta\chi^2 = 92$.

Further increase in CMS double differential Z/γ data ($\Delta\chi^2 = 24$) and E866 Drell-Yan asymmetry. Dimuon data not any worse.

Now also deterioration in HERA combined data, both NC and CC and CDF differential Z/γ data.

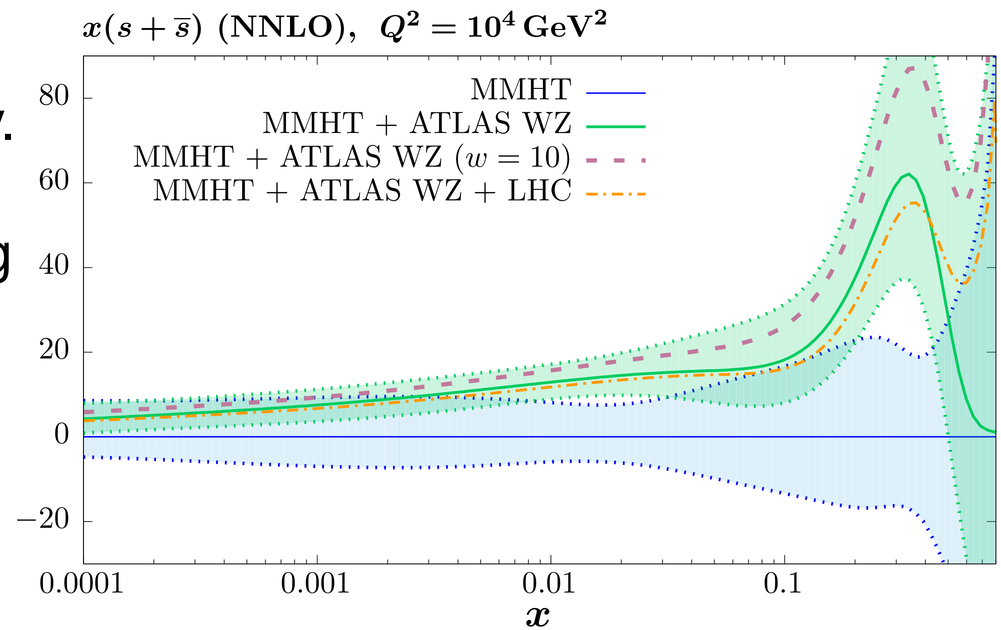
Also try fit where all other new LHC data from LHCb and CMS included. Compared to baseline plus ATLAS W, Z data very little change $\Delta\chi^2 = 3$ in total, and essentially no change in ATLAS W, Z data.

However, inclusion of ATLAS W, Z data lowers χ^2 for new LHC (plus final D0) data by $\Delta\chi^2 = -10$.

Hence ATLAS W, Z data and other new LHC data compatible and pull in same direction. Only CMS $W + c$ deteriorates slightly.

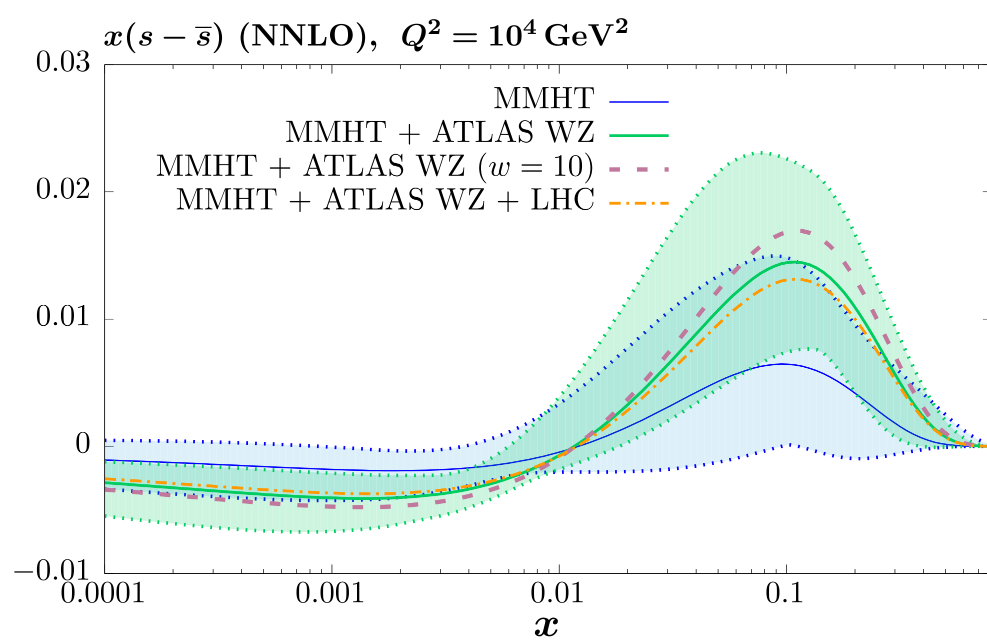
Effect on PDFs

Large increase in $s + \bar{s}$ and decrease in uncertainty. Correlation with fit to dimuon data (lower branching ratio) leads to increase at high x. (Note negative NNLO correction Phys. Rev. Lett. 116 (2016), Berger *et al.*.)

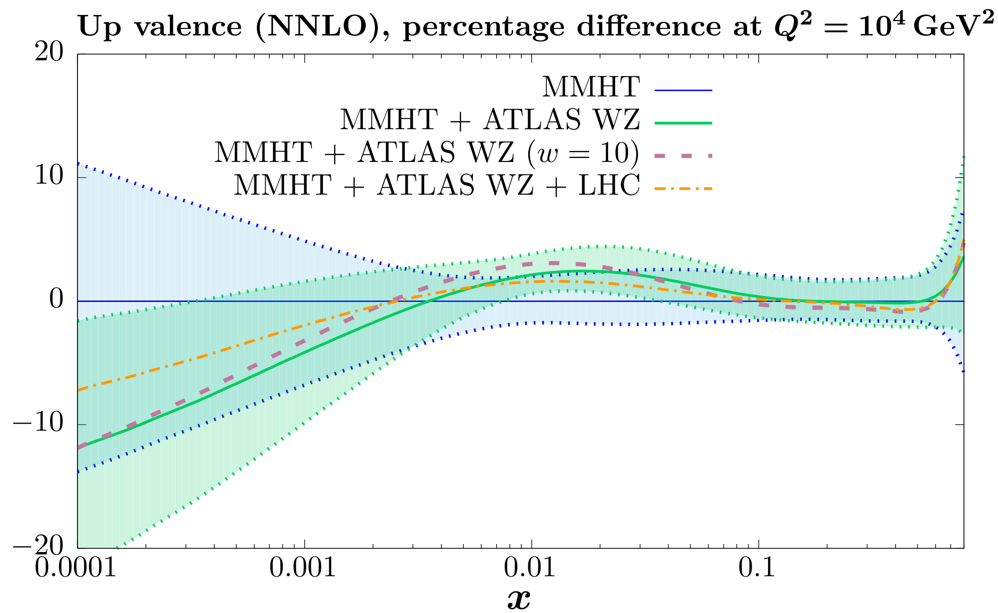


Larger for $x > 0.1$ due to significant down quark contribution in this region despite Cabibbo suppression.

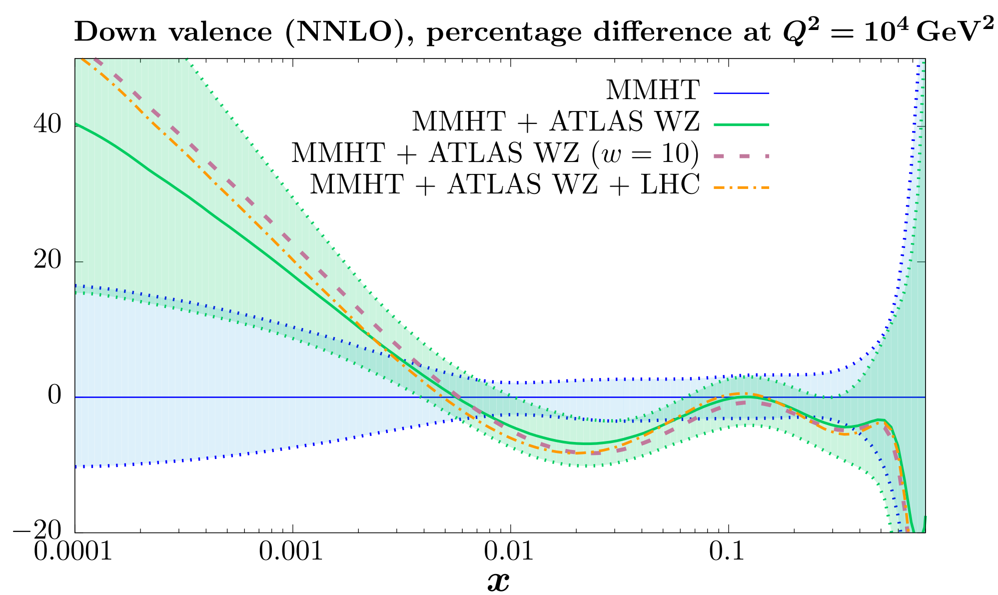
There is impact on $s - \bar{s}$ uncertainty, from the change in branching ratio.



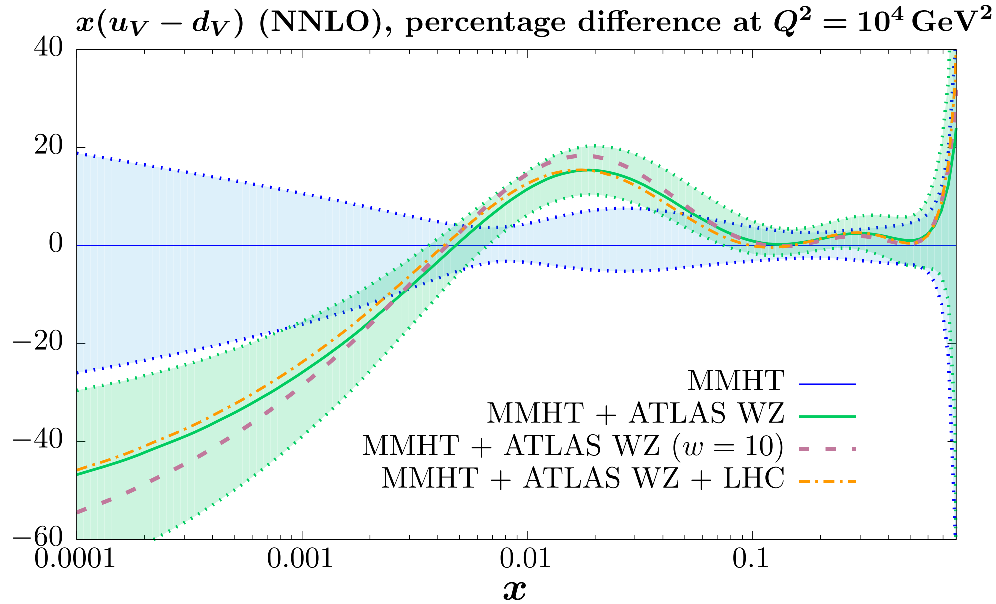
Significant impact on shape of valence quarks.



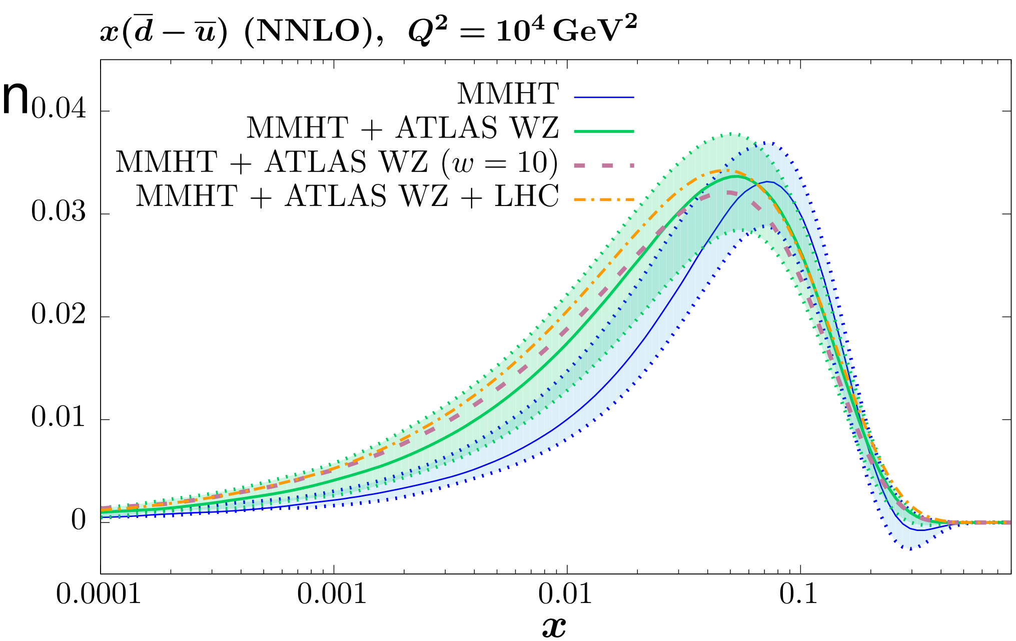
Same direction as impact of other new LHC data.



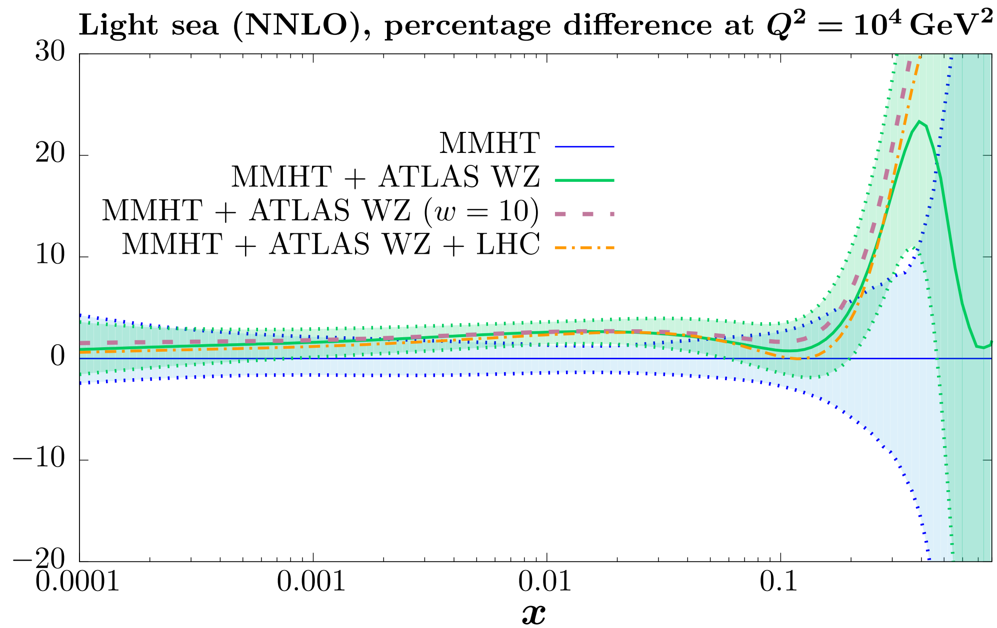
As implied by individual distributions, significant change in $u_V - d_V$.



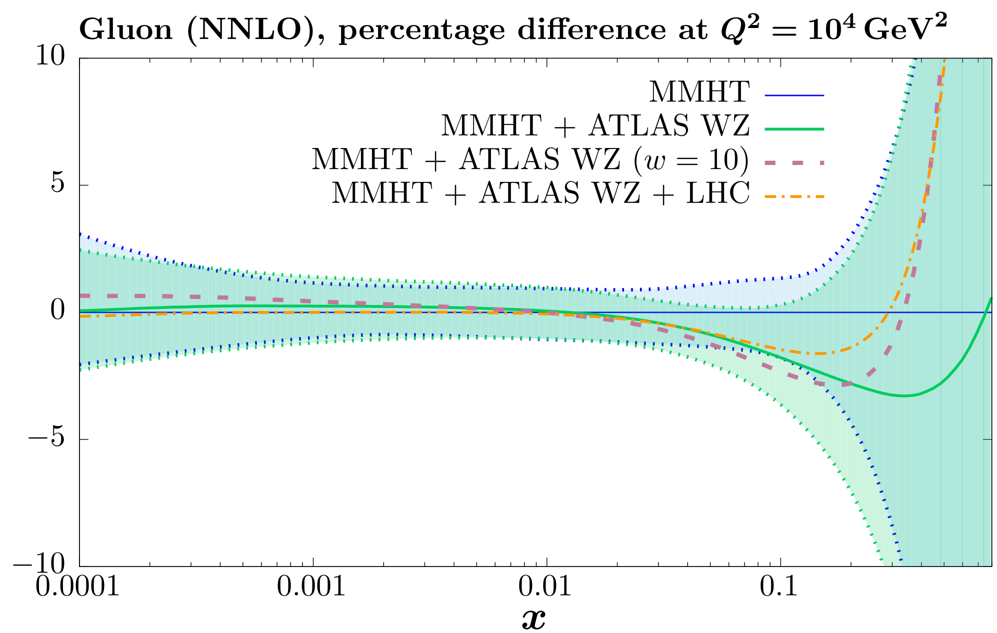
Shift in best fit $\bar{d} - \bar{u}$ accompanying deterioration in fit to E866 Drell-Yan asymmetry.



Change in strange quark affects sea, making it generally larger.



Little impact on gluon distribution. .



Extension of $\bar{d} - \bar{u}$ parameterisation.

Currently use 3 free parameters, i.e.

$$(\bar{d} - \bar{u})(x, Q_0^2) = A(1-x)^{\eta_{sea}+2} x^{\delta}(1 + \gamma x + \Delta x^2),$$

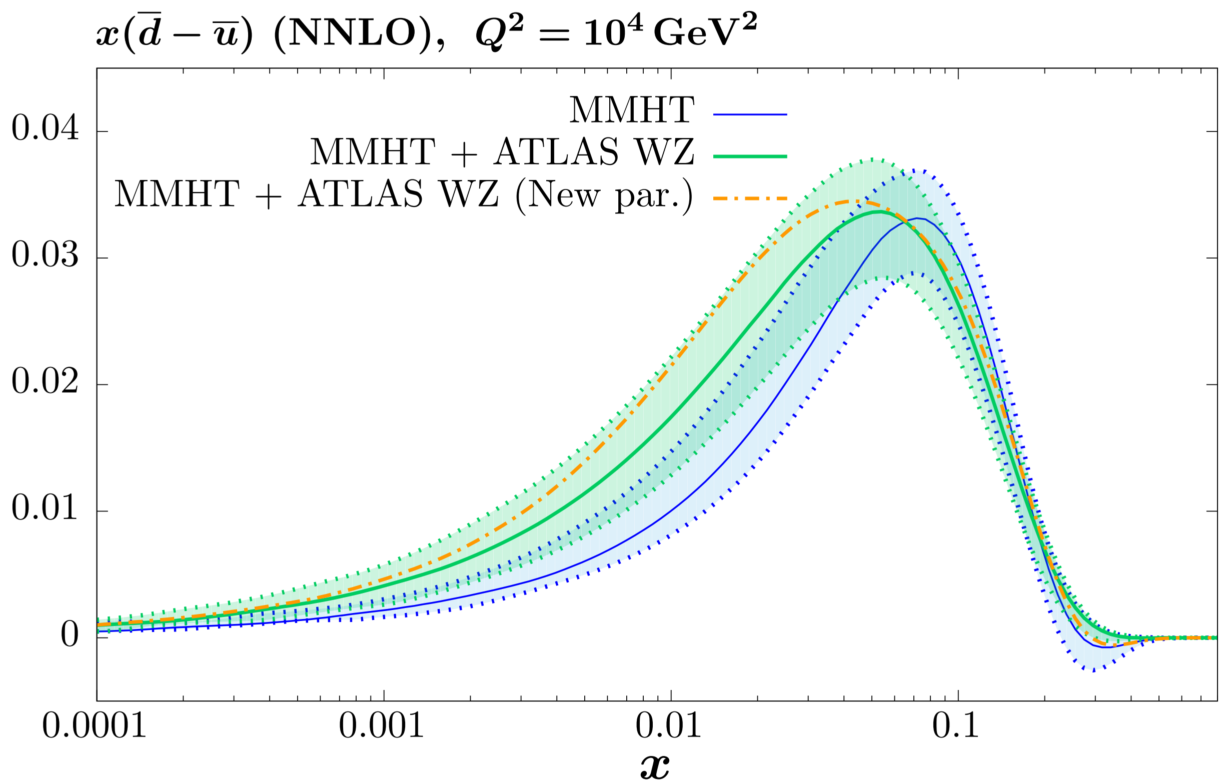
Extend to

$$(\bar{d} - \bar{u})(x, Q_0^2) = A(1-x)^{\eta_{sea}+2} x^{\delta}(1 + \sum_{i=1}^{4} a_i T_i(1 - 2x^{\frac{1}{2}})),$$

where $T_i(1 - 2x^{\frac{1}{2}})$ are Chebyshev polynomials. So 5 free parameters. Easily allows multiple turning points (seen in first fit iteration).

Global fit including new LHC data and new ATLAS W, Z data improves by 10 units, but over 5 of this in E866 Drell Yan asymmetry.

Parameterisation alleviates some tension between ATLAS data and Drell Yan asymmetry.



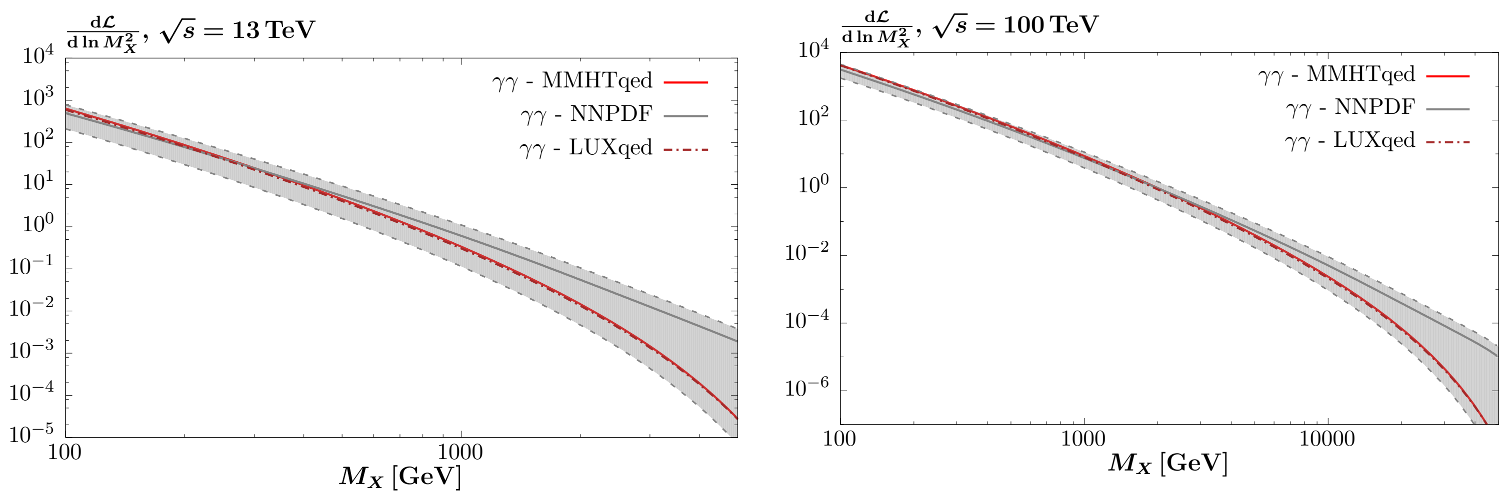
New $(\bar{d} - \bar{u})$ distribution similar at high x to previous one. (Dips to negative values at low-x allowed by, and seen using new parameterisation.)

Now a smaller decrease towards zero at low x beyond edge of previous uncertainty band.

PDFs with QED corrections

We will base photon input for PDFs at low Q^2 on the LUXqed prescription which demonstrated determination in terms of structure function, and hence precision of at worst a few percent.

Effect of photon evolution fully incorporated to couple with that of quarks and gluon for both proton and neutron.



Final details of transition from low Q^2 to be finalised (e.g. elastic (coherent) contribution still important above $Q_0^2 = 1\mathrm{GeV}^2$, but MMHT photon (Nathvani) very similar to LUXqed.

Connection to xFitter

I will briefly discuss a number of connections working in both directions.

Main historical connection, the general mass variable flavour number scheme GM-VFNS used by MSTW/MMHT has been used for a long time by HERAFitter/xFitter.

Also discuss various possibilities for the future - very much a case of “to be discussed”.

The GM-VFNS can be defined by demanding equivalence of the n_f light flavour and n_f+1 light flavour descriptions at all orders – above transition point $n_f \to n_f+1$

$$F(x,Q^2) = C_k^{FF,n_f}(Q^2/m_H^2) \otimes f_k^{n_f}(Q^2) = C_j^{VF,n_f+1}(Q^2/m_H^2) \otimes f_j^{n_f+1}(Q^2)$$
$$\equiv C_j^{VF,n_f+1}(Q^2/m_H^2) \otimes A_{jk}(Q^2/m_H^2) \otimes f_k^{n_f}(Q^2).$$

Hence, the VFNS coefficient functions satisfy

$$C_k^{FF,n_f}(Q^2/m_H^2) = C_j^{VF,n_f+1}(Q^2/m_H^2) \otimes A_{jk}(Q^2/m_H^2),$$

which at $\mathcal{O}(\alpha_S)$ gives (in $\overline{MS}$ scheme)

$$C_{2,Hg}^{FF,n_f,(1)}\left(\frac{Q^2}{m_H^2}\right) = C_{2,HH}^{VF,n_f+1,(0)}\left(\frac{Q^2}{m_H^2}\right) \otimes P_{qg}^0 \ln(Q^2/m_H^2) + C_{2,Hg}^{VF,n_f+1,(1)}\left(\frac{Q^2}{m_H^2}\right),$$

The VFNS coefficient functions tend to the $m=0$ limits as $Q^2/m_H^2 \to \infty$.

However, $C_j^{VF}(Q^2/m_H^2)$ only uniquely defined in this limit.

Can swap $\mathcal{O}(m_H^2/Q^2)$ terms between $C_{2,HH}^{VF,0}(Q^2/m_H^2)$ and $C_{2,g}^{VF,1}(Q^2/m_H^2)$.

Have the freedom to modify the heavy quark coefficient function, by default

$$C^{VF,0}_{2,HH}(Q^2/m_H^2, z) = \delta(z - x_{\max}).$$

Appears in convolutions for higher order subtraction terms, so do not want complicated x dependence. Simple choice.

$$C^{VF,0}_{2,HH}(Q^2/m_H^2, z) \to (1 + b(m_H^2/Q^2)^c)\delta(z - x_{\max})),$$

where c really encompasses (m_H^2/Q^2) with logarithmic corrections.

Can also modify argument of δ-function, as in Intermediate Mass (IM) scheme of Nadolsky, Tung. Let argument of heavy quark contribution change like

$$\xi = x/x_{\max} \to x\big(1 + (x(1 + 4m_H^2/Q^2))^d 4m_H^2/Q^2\big),$$

so kinematic limit stays the same, but if $d > 0$ small x less suppressed, or if $d < 0$ (must be > -1) small x more suppressed.

Default a, b, c, d all zero. Limit either by fit quality or *sensible* choices.

6 extreme variations tried.

GMVFNS1 – $b = -1, c = 1$.

GMVFNS2 – $b = -1, c = 0.5$.

GMVFNS1 – $a = 1$.

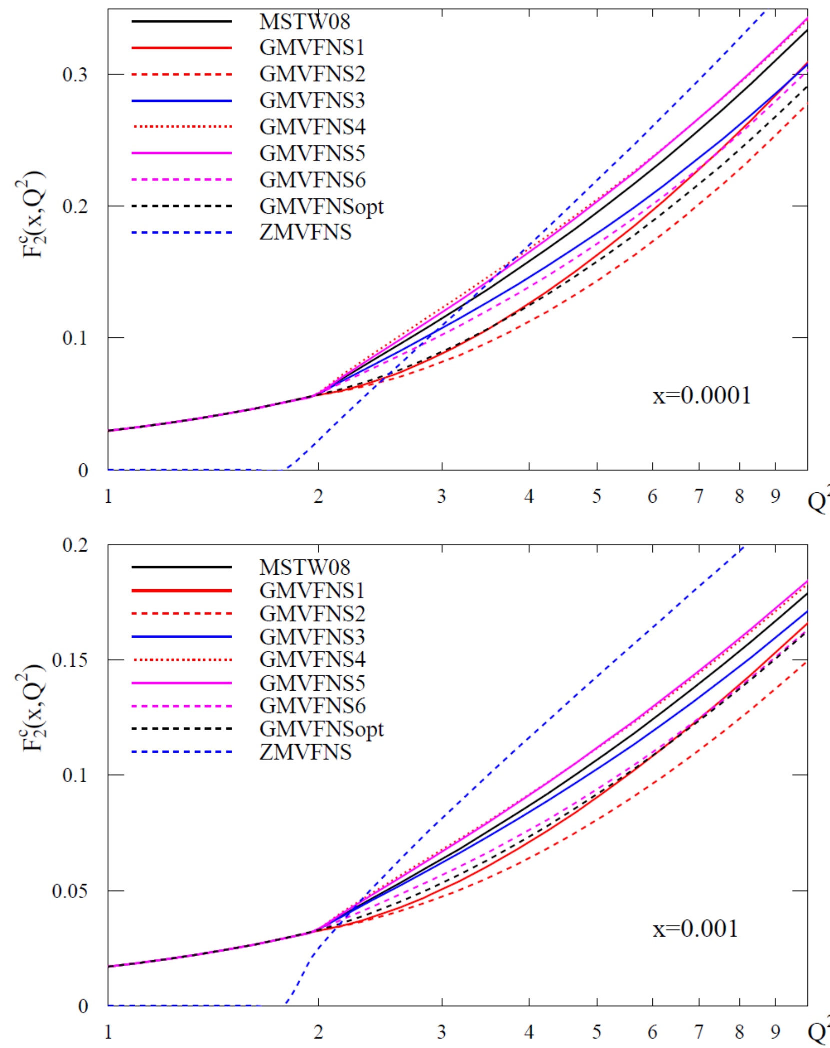
GMVFNS1 – $b = +0.3, c = 1$ – fit.

GMVFNS1 – $d = 0.1$ – fit.

GMVFNS1 – $d = -0.2$ – fit.

Variations in $F_2^c(x, Q^2)$ near the transition point at NLO due to different choices of GM-VFNS.

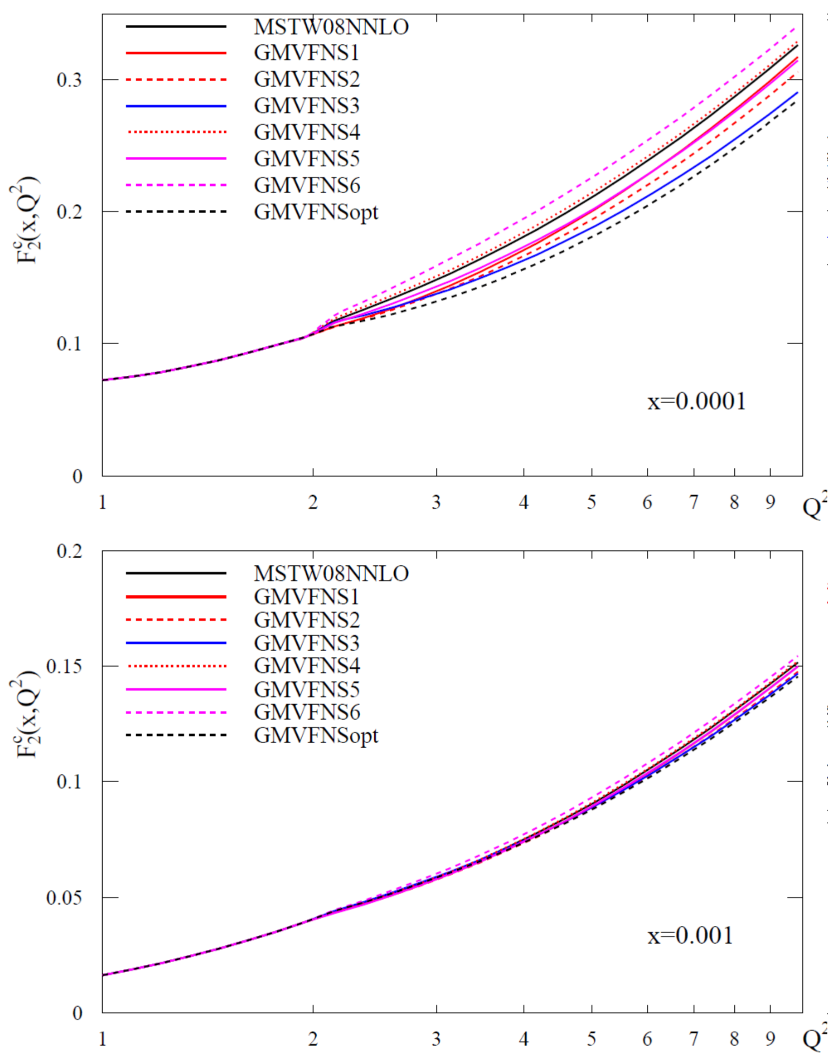
Optimal, $a = 1, b = -2/3, c = 1$, smooth behaviour.



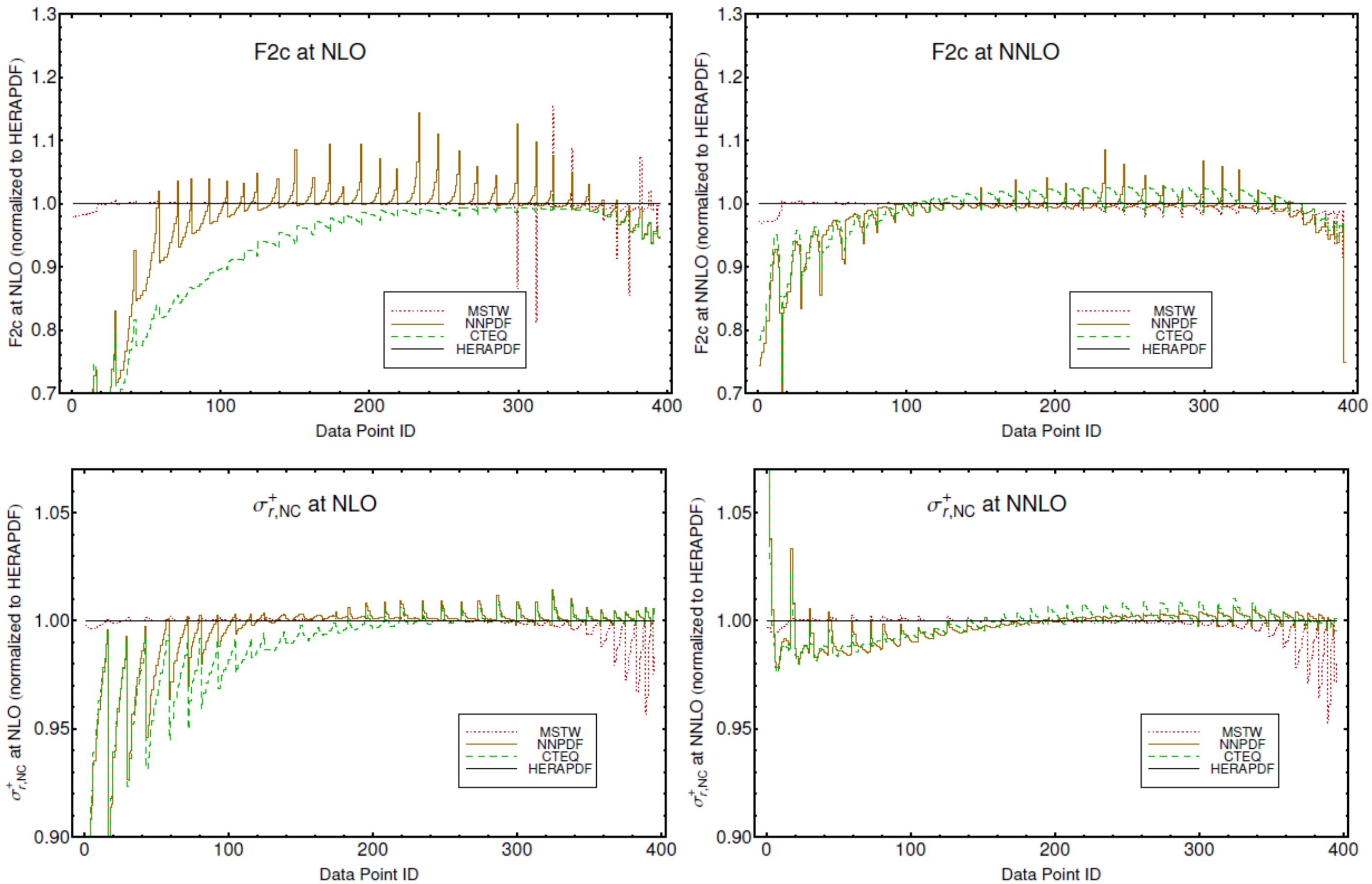
Variations in $F_2^c(x, Q^2)$ near the transition point due to different choices of GM-VFNS at NNLO.

Very much reduced, almost zero variation until very small x.

Shows that NNLO evolution effects most important in this regime.



Also see convergence between groups in Les Houches benchmark study.



NNLO TR scheme larger at lowest Q^2 due to use of $\mathcal{O}(\alpha_S^3)$ coefficient function.

Not currently implemented for charged current cross sections – not too important for HERA data where zero mass for charm is a good approximation.

In principle available for NLO and for an old version of approximate NNLO.

Better to wait for the inclusion of the real NNLO corrections of Phys. Rev. Lett. 116 (2016), Berger *et al.*?

MMHT uses of xFitter

Used for aid in fitting newly appearing LHC data.

Both data points, uncertainties and applgrids and K-factors.

For example, most recent ATLAS W, Z data information obtained from `http://www.hepforge.org/archive/xfitter/1612.03016.tar.gz`.

(Aspire to check theory calculations independently).

Will definitely find it useful to use similar information provided for processes where initial state photon is important, e.g. recent ATLAS high mass Drell Yan.

Other possible exchange of ideas?

Best fit procedure - based on Numerical Recipes.

Uses Levenberg-Marquardt algorithm based on a combination of steepest decent and Gauss-Newton iteration. For parameters a_i

$$a_{i+1} = a_i - (H + \lambda \mathrm{diag[H]})^{-1}_{\mathrm{ik}} \boldsymbol{\nabla}_{\mathrm{k}} \chi^2(\mathrm{a})$$

where H is the Hessian matrix and λ a constant. λ is varied to maximize convergence.

Chebyshev polynomial parameterisation. Not sure this requires any code.

Dynamical tolerance procedure (as alternative to $\Delta\chi^2 = 1$ if constraints from a variety of data sets start to seem too constraining?

Conclusions

New HERA II combined data studied with context of MMHT2014 PDFs. Fit quality good – better at NNLO. No very significant changes in PDFs or predictions. Slight reduction in uncertainties.

Predictions turn out to be good for many LHC data previously not included in the fit. Few changes in central values, but some data reduce uncertainties, mainly in strange and low-x valence quarks.

Failure to fit ATLAS 7 TeV jet data at NLO - common with other groups. NNLO leads to little change. Poor fit driven by a small number of correlated systematic uncertainties. Much improved with model of reduced rapidity correlation.

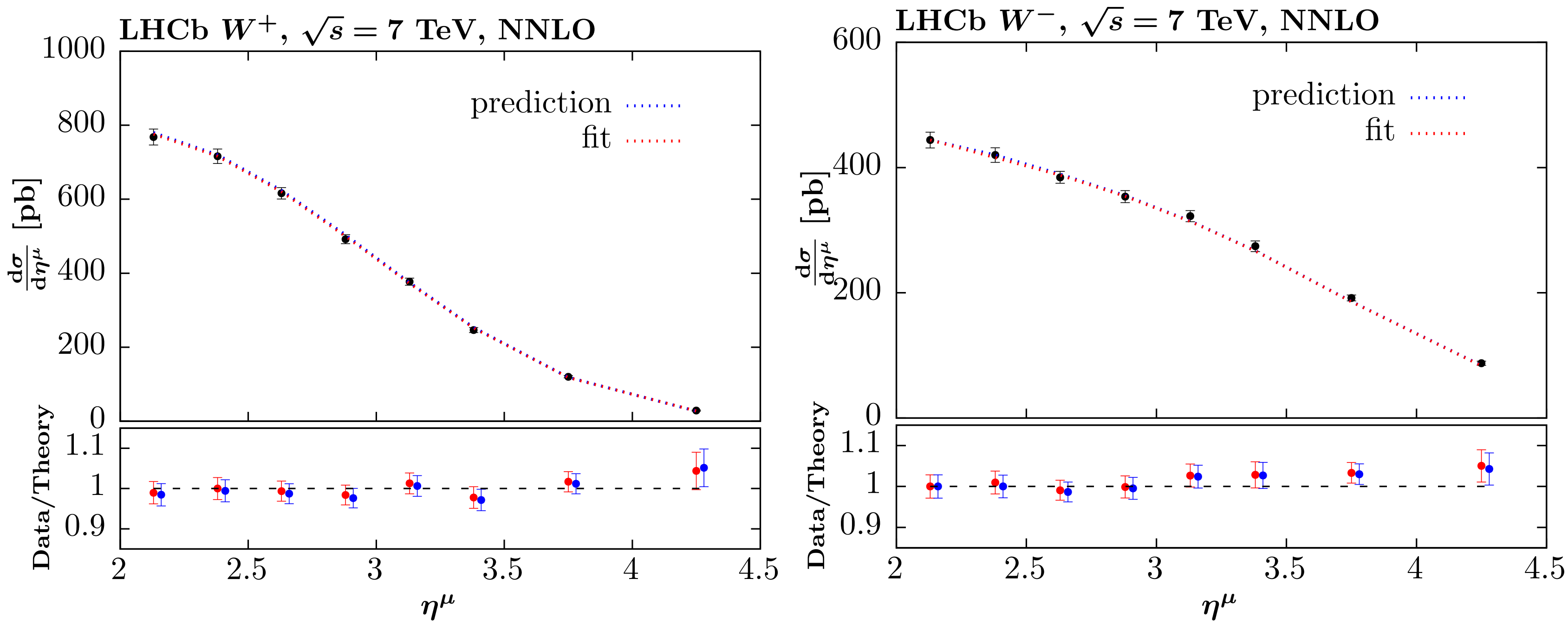
Prediction for new ATLAS W, Z data poor but much improved with new fit. Tensions with older data, particularly dimuon, and more-so CMS 7 TeV Drell Yan data. Increase in strange quark. Very compatible with most recent CMS, LHCb data in the fit. Improved $(\bar{d} - \bar{u}$ helps.

Work on updated PDFs with QED corrections (Nathvani).

Variety of possible areas of reciprocity with xFitter.

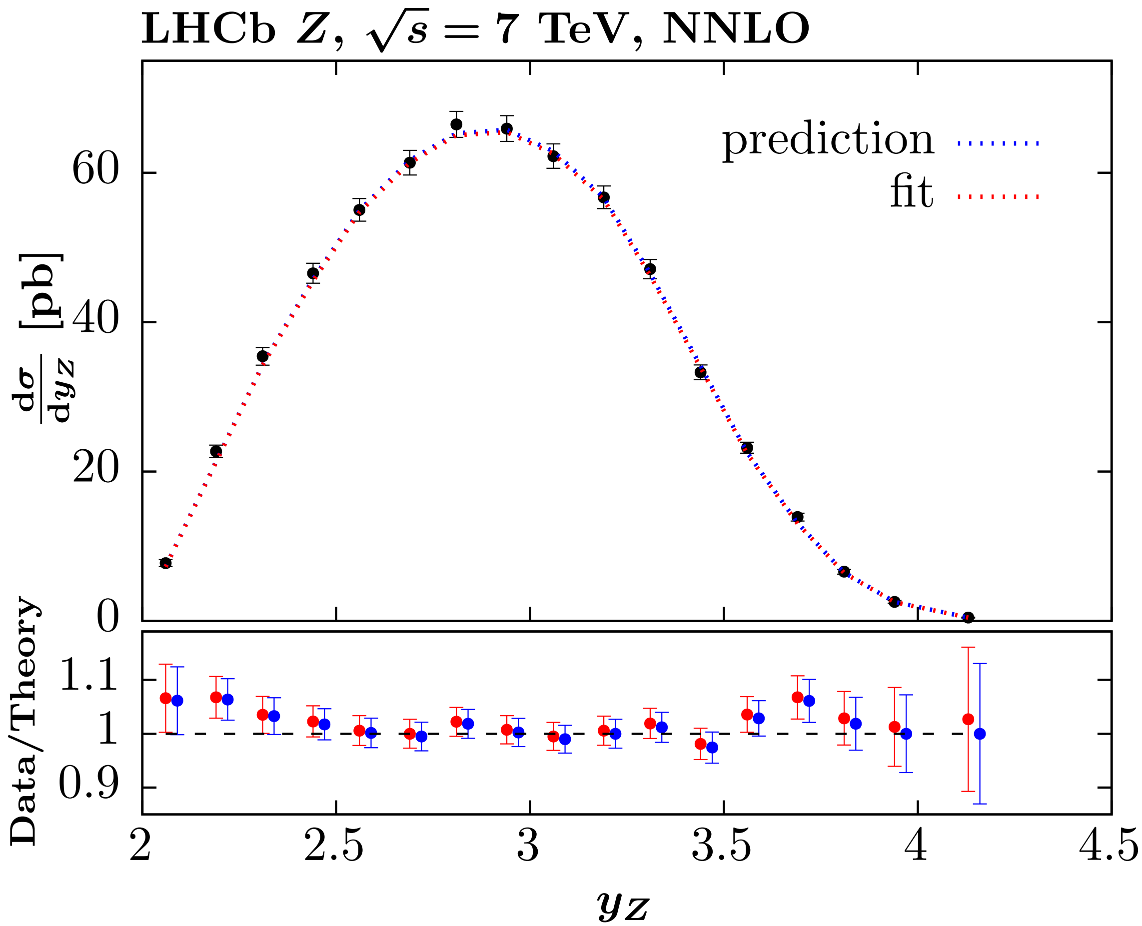
Back -up

New data on high rapidity W production at LHCb at 7 TeV.



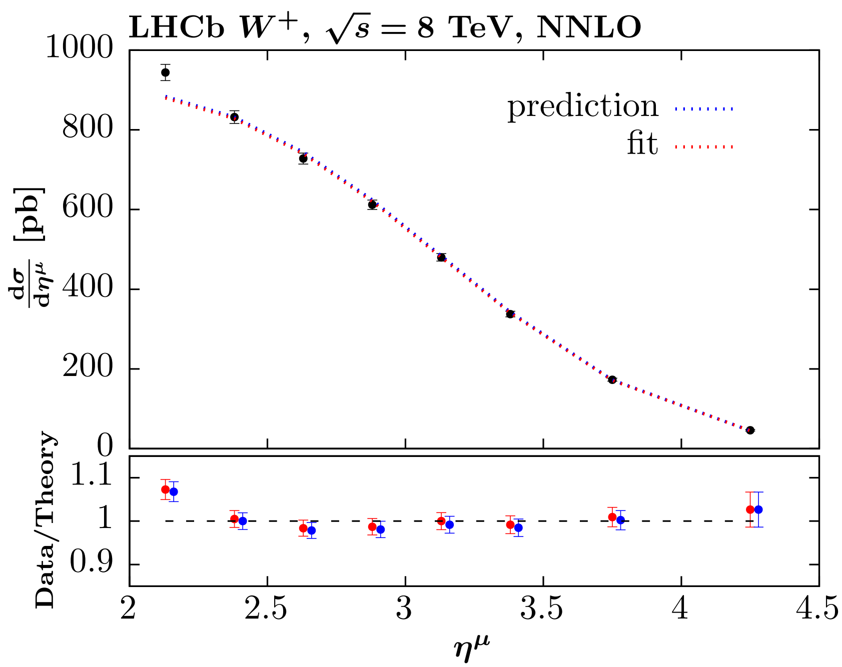
Generally perfectly good agreement using NNLO. Uncertainties added in quadrature on plot, but covariance matrix used in fit.

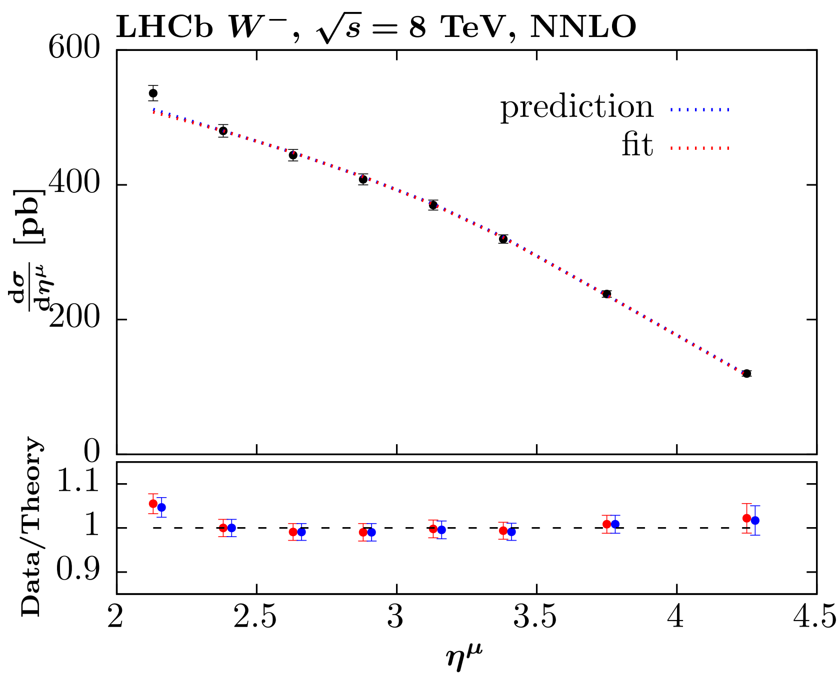
New data on high rapidity Z production at LHCb at $7\ \mathrm{TeV}$.



Generally perfectly good agreement using NNLO. A little low at low y_Z.

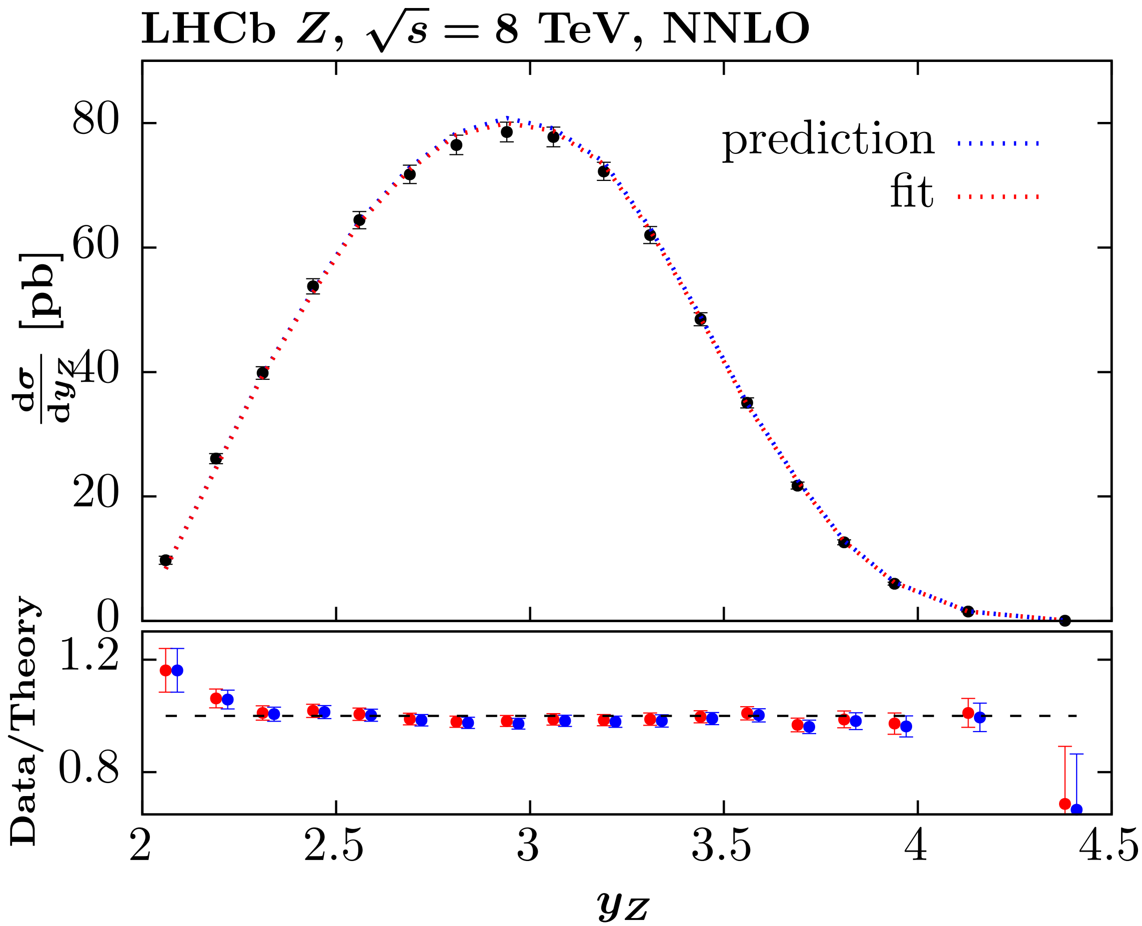
New data on high rapidity W production at LHCb at $8\ \mathrm{TeV}$.





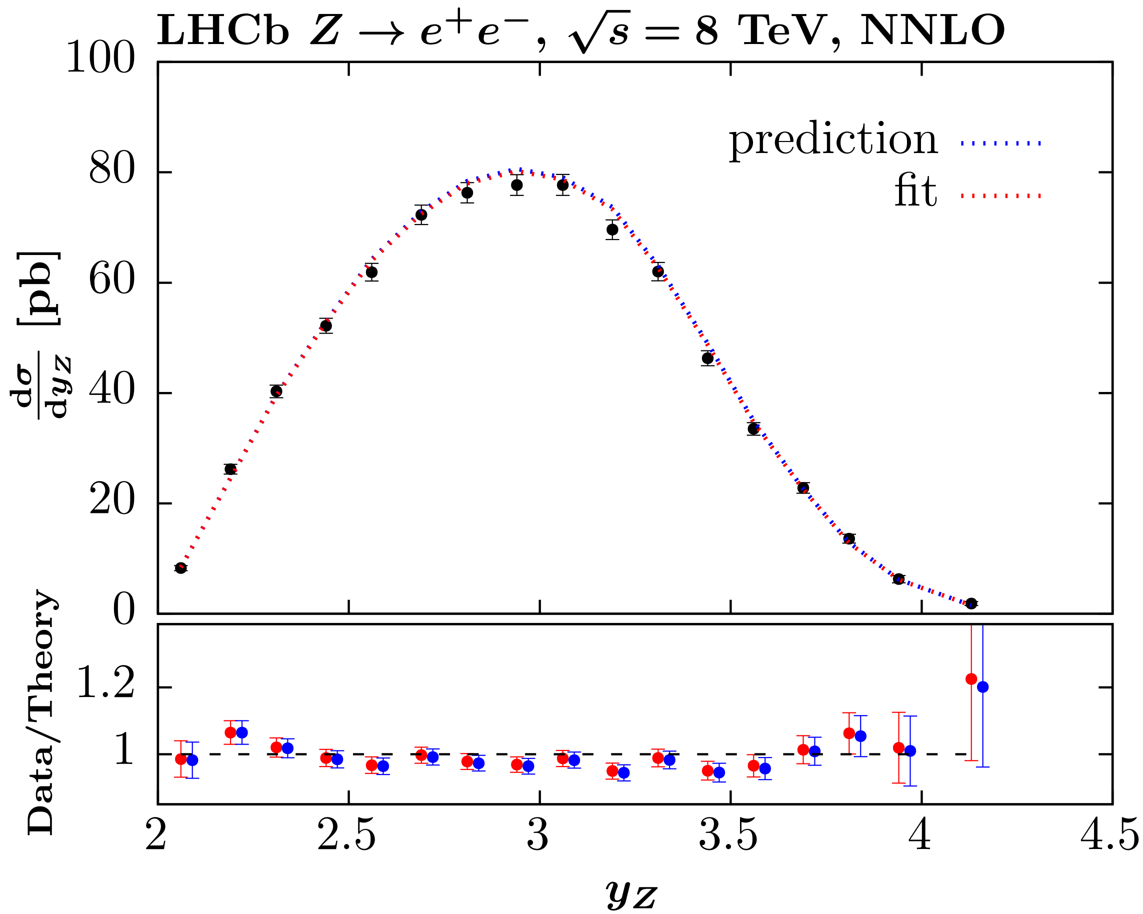
Good fit except at lowest η_μ point in each case.

New data on high rapidity Z production at LHCb at 8 TeV.



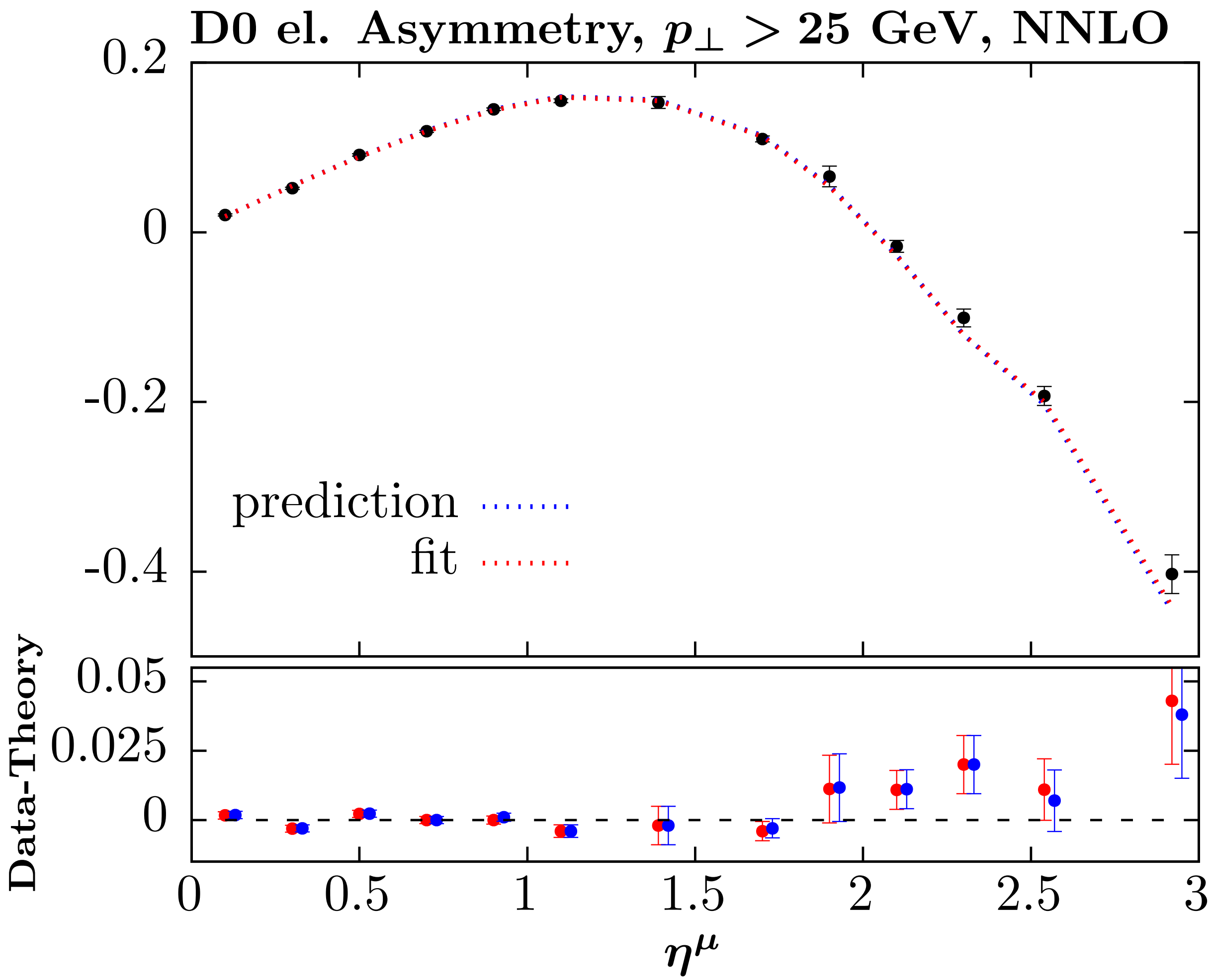
Same issue with lowest y_z point. PDFs at moderate x for these points and well constrained by DIS data.

New data on high rapidity Z production at LHCb at 8 TeV with electrons.



No issue at lowest y_z with these data. Relatively large χ^2 only down to fluctuations.

Good agreement with new D0 e asymmetry data



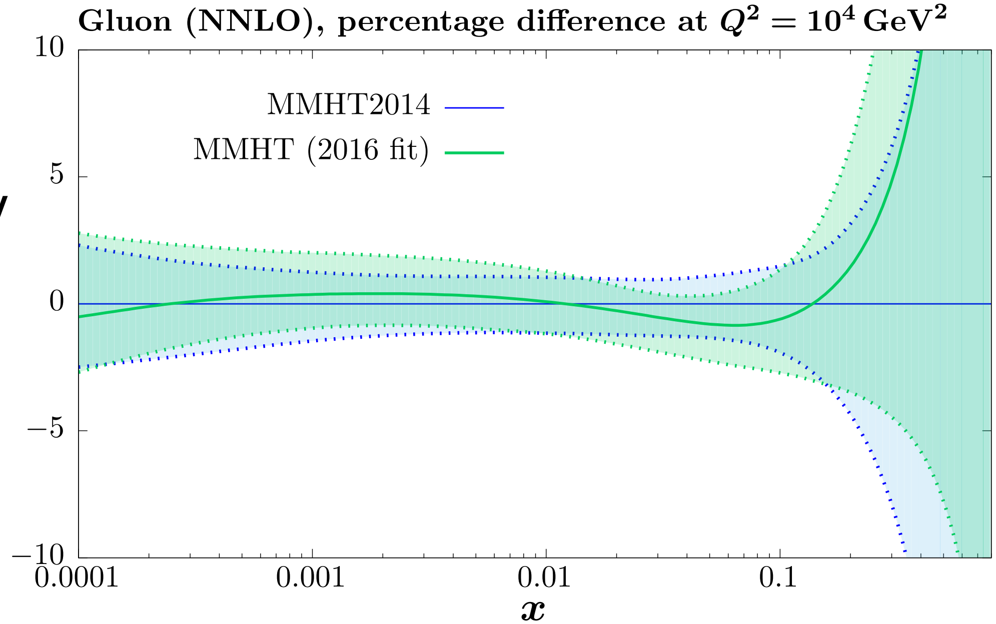
Slight undershooting at highest η_e. Implies slightly smaller down quark, but other data does not prefer this.

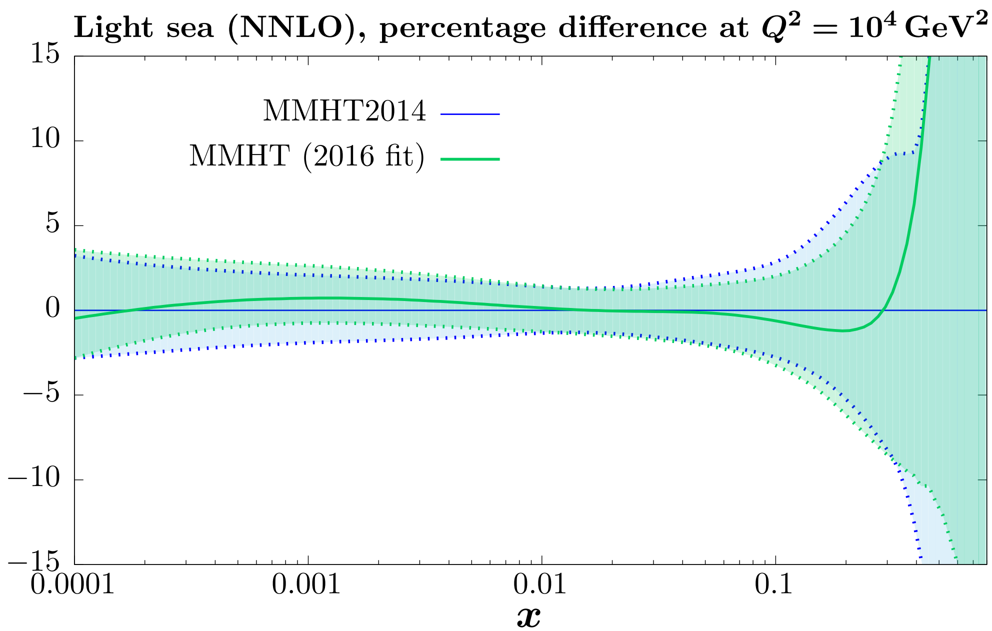
(Use the prescription for systematic uncertainties advocated in Eur.Phys.J. C75 (2015) no.9, 458 for these and other Tevatron data.)

Effect on PDFs

No significant change in gluon or light sea.

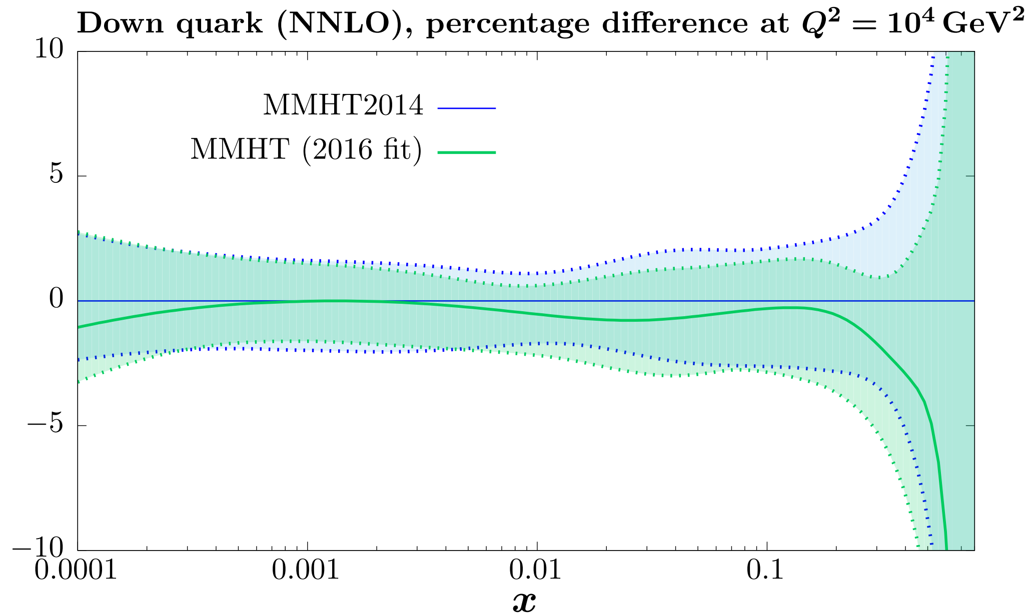
Small decrease in uncertainty in some small-x regions due to new HERA data.

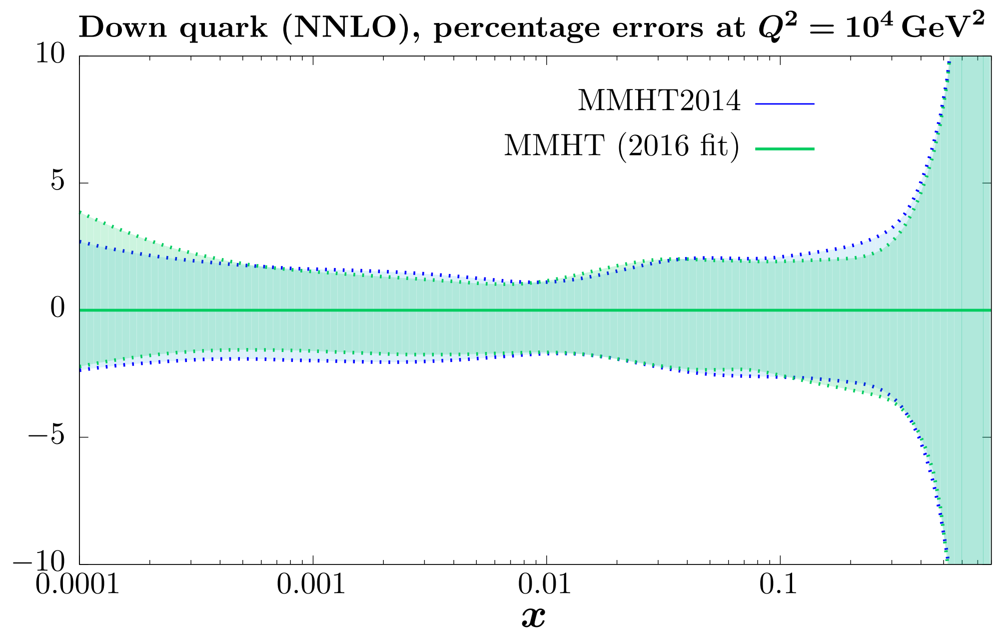


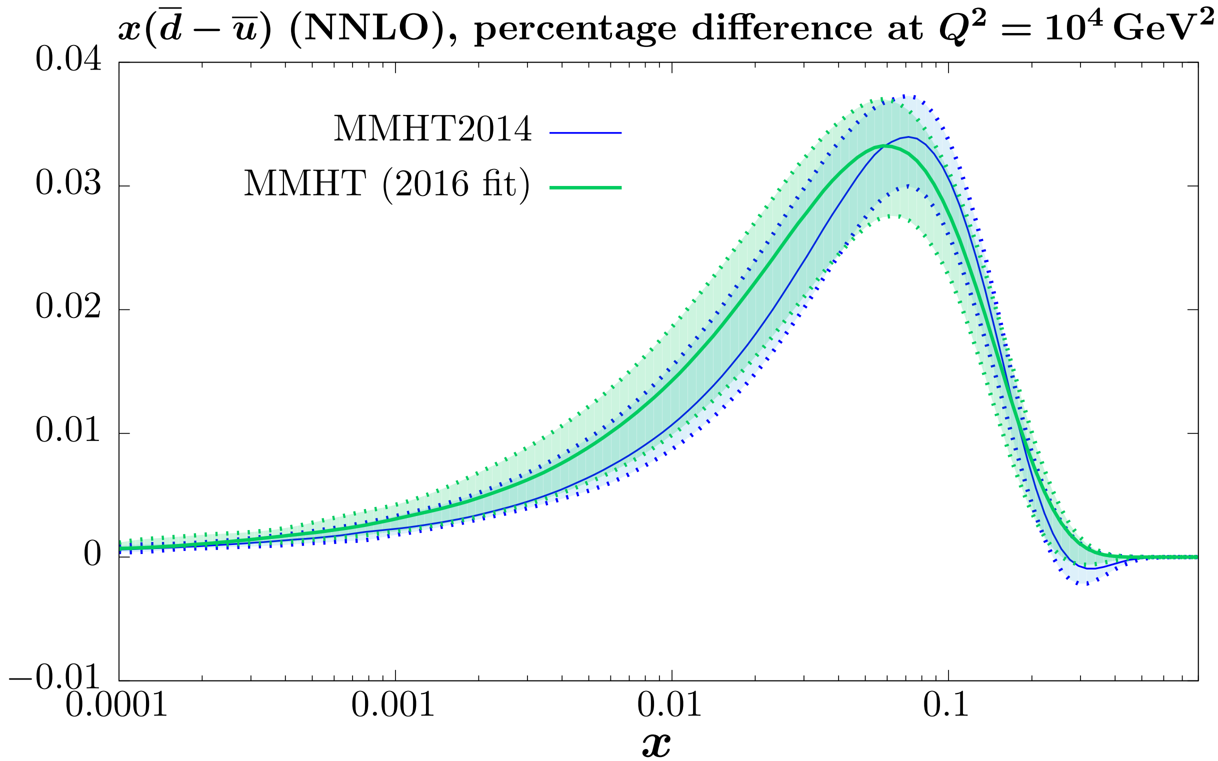


Small change in d at $x \sim 0.01$ and some reduction in uncertainty.

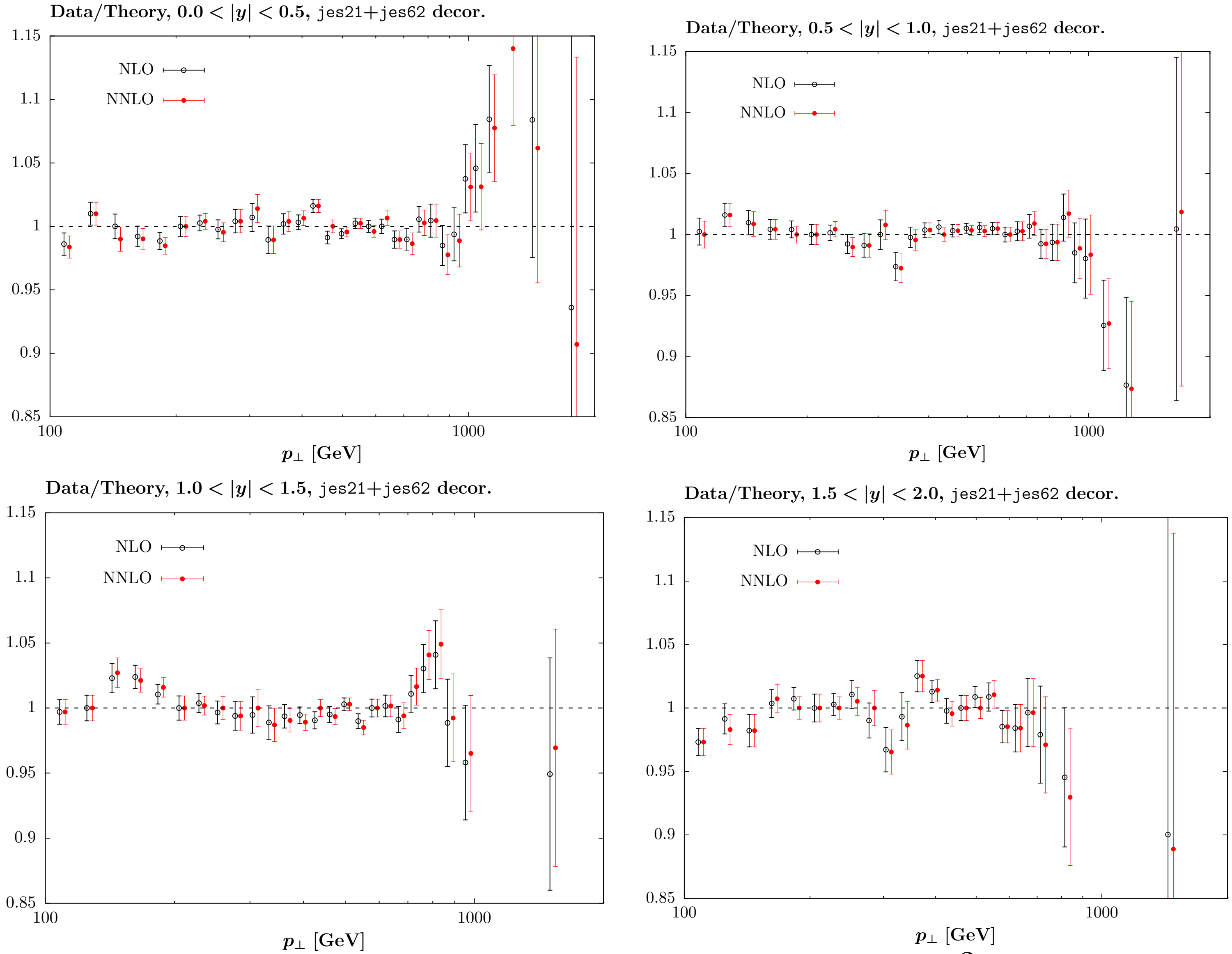
Significant change in d at high x and some reduction in uncertainty for $x \sim 0.2$.



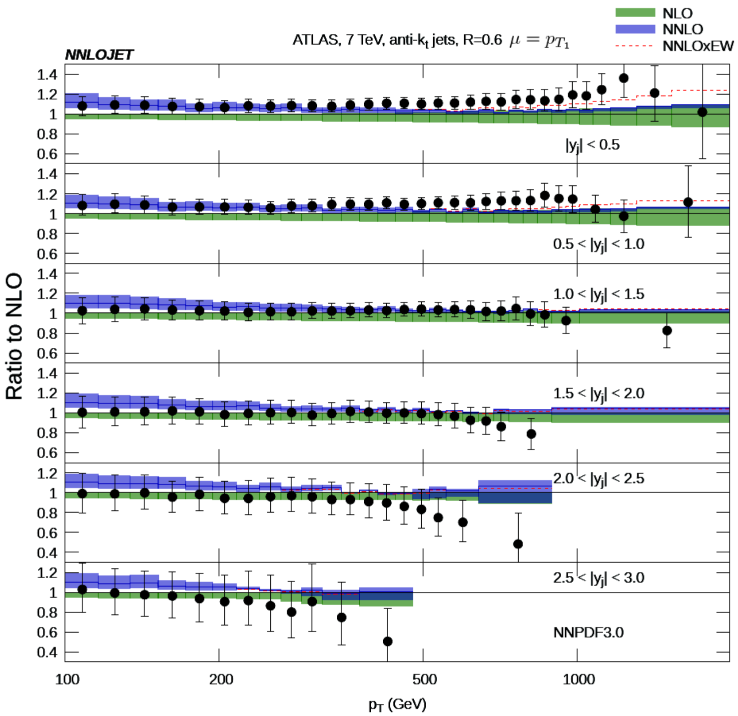


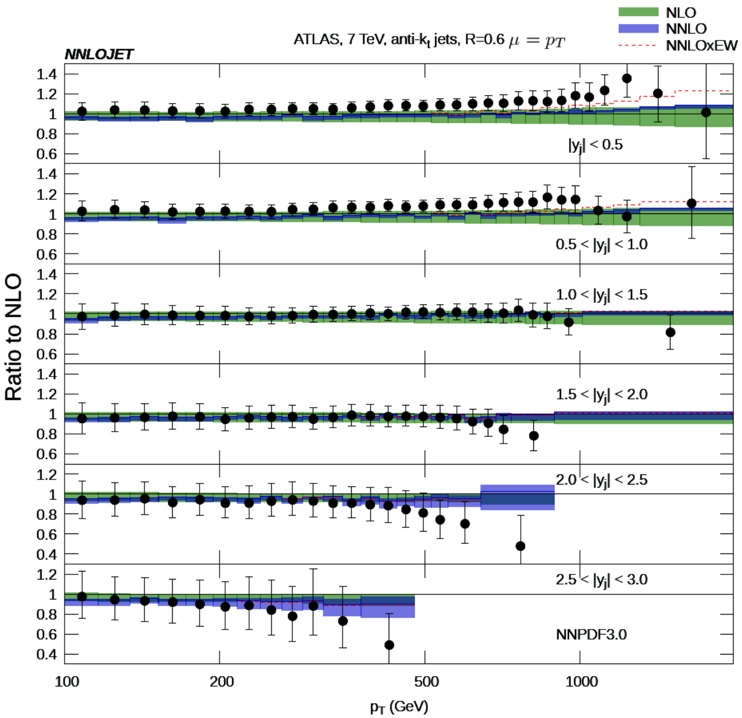


No major change in $\bar{d} - \bar{u}$, but even less inclination towards a change in sign at high x which was a feature of earlier sets.

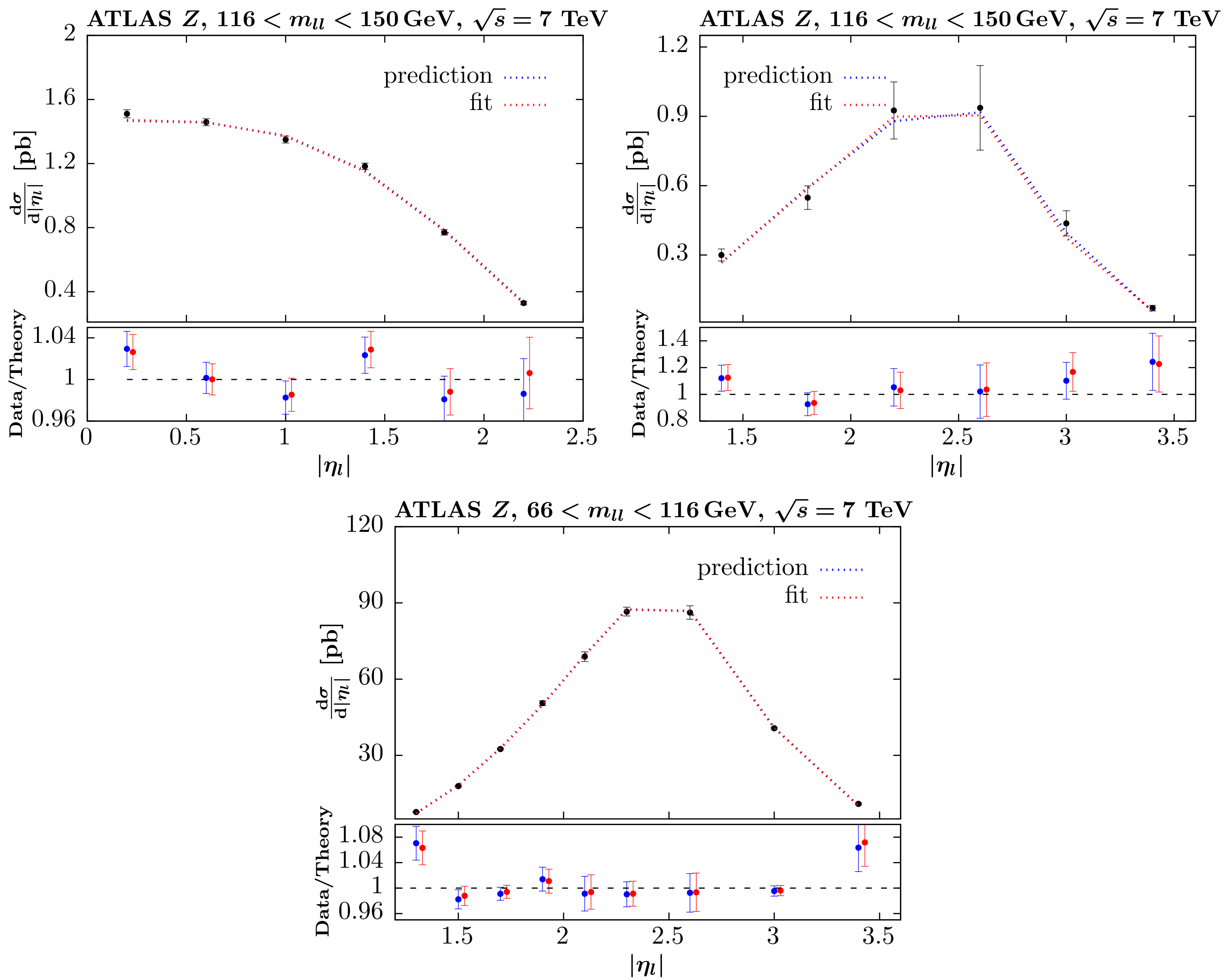


Fit quality at NLO and NNLO when jes21 and jes 62 decorrelated between $|y|$ bin.

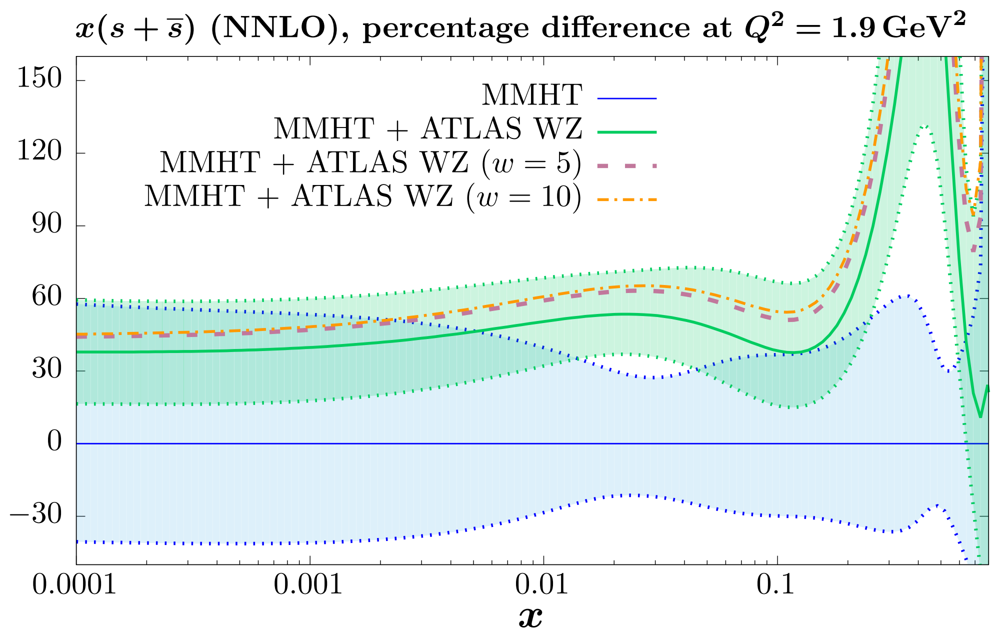

NNLOJET
ATLAS, 7 TeV, anti-kt jets, R=0.6 $\mu = p_{T_1}$
NLO
NNLO
NNLOxEW
Ratio to NLO
|yj| < 0.5
0.5 < |yj| < 1.0
1.0 < |yj| < 1.5
1.5 < |yj| < 2.0
2.0 < |yj| < 2.5
2.5 < |yj| < 3.0
NNPDF3.0
pT (GeV)


NNLOJET
ATLAS, 7 TeV, anti-kt jets, R=0.6 $\mu = p_T$
NLO
NNLO
NNLOxEW
Ratio to NLO
|yj| < 0.5
0.5 < |yj| < 1.0
1.0 < |yj| < 1.5
1.5 < |yj| < 2.0
2.0 < |yj| < 2.5
2.5 < |yj| < 3.0
NNPDF3.0
pT (GeV)

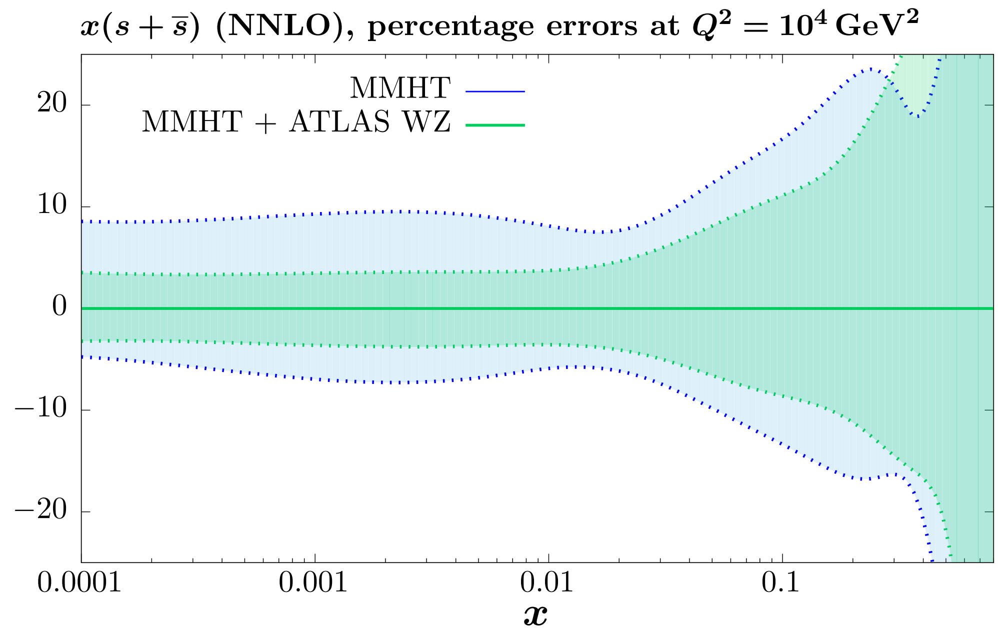
Fit to new ATLAS W, Z data



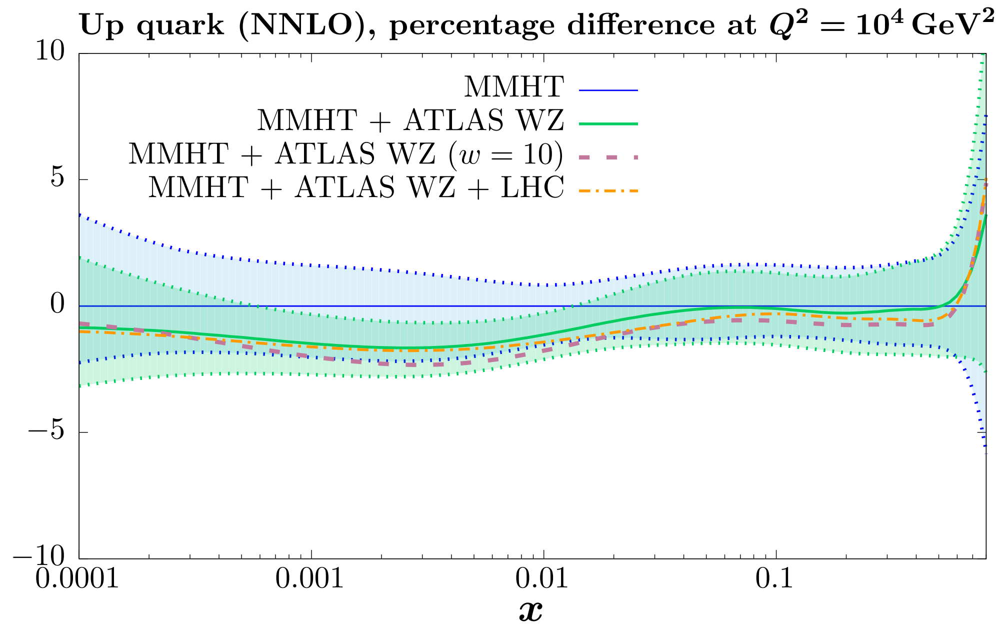
Strange quark at lower Q^2 with addition of new ATLAS W, Z data.



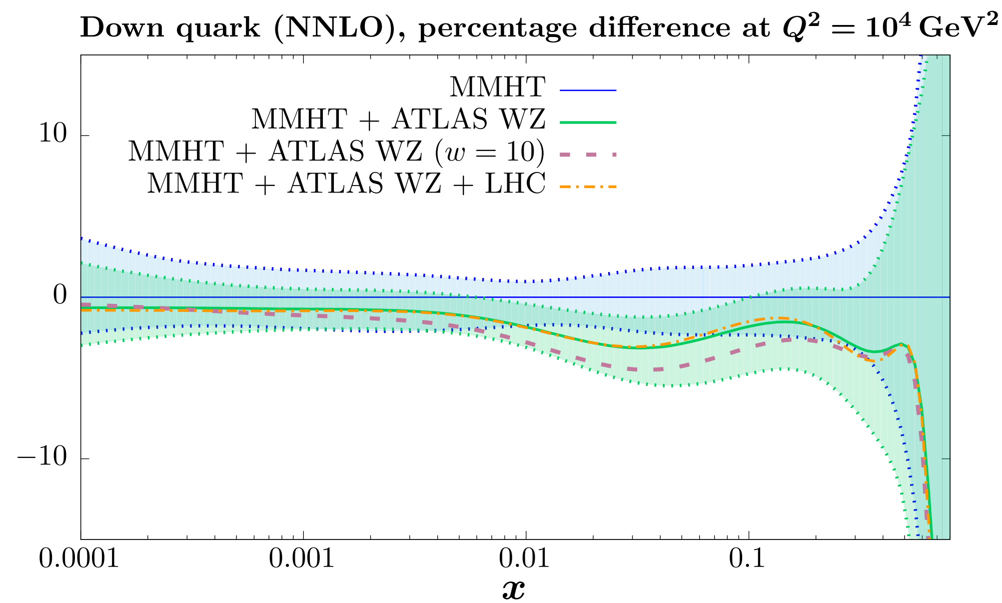
Uncertainty reduction in strange.



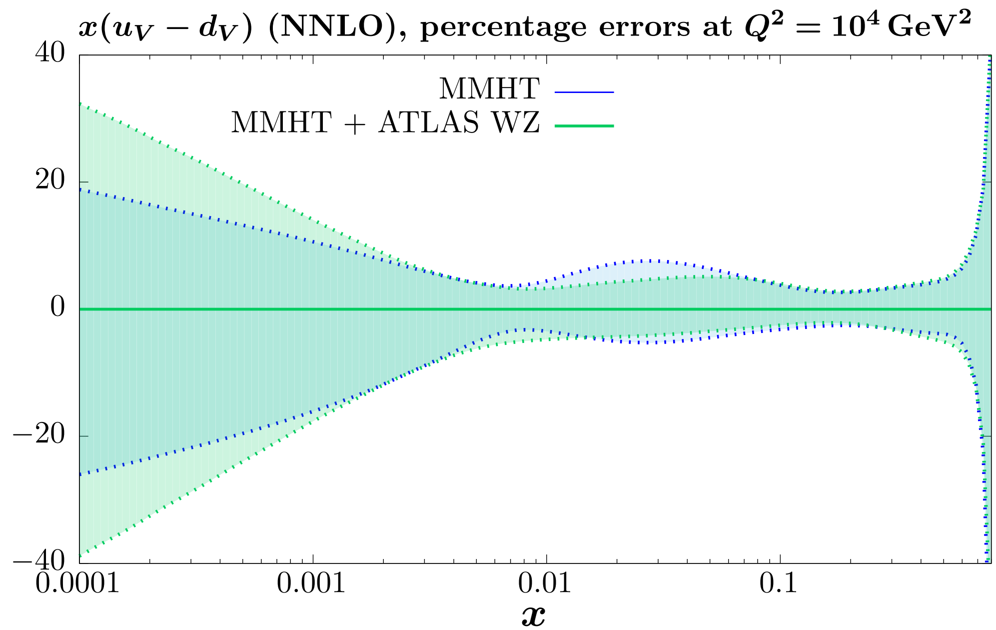
Slight decrease in low x up quark from increase in strange.



More pronounced effect in down quark.



Decrease in $u_V - d_V$ uncertainty for $0.01 < x < 0.1$.



Improved constraint on strange feeds into total sea uncertainty.