MMHT2014 PDFs

Robert Thorne

November 3rd, 2014

University College London

In collaboration with Lucian Harland-Lang, Patrick Motylinski and Alan Martin

and thanks to Ben Watt, Graeme Watt and James Stirling

Updates in PDF Fits – the MMHT 2014 PDFs.

I will present results on an update in PDFs within the general MSTW framework due to some theory improvements and a variety of new data sets, including most of the up-to-date LHC data. The release of a new set of MMHT PDFs is finalised.

Changes in theoretical treatment or procedures.

Continue to use extended parameterisation with Chebyshev polynomials, and freedom in deuteron nuclear corrections (and heavy nuclear corrections), as in recent MSTWCPdeut study (Eur.Phys.J. C73 (2013) 2318) – change in $u_V - d_V$ distribution.

Now use "optimal" GM-VFNS choice (Phys.Rev. D86 (2012) 074017) which is smoother near to heavy flavour transition points (more so at NLO).

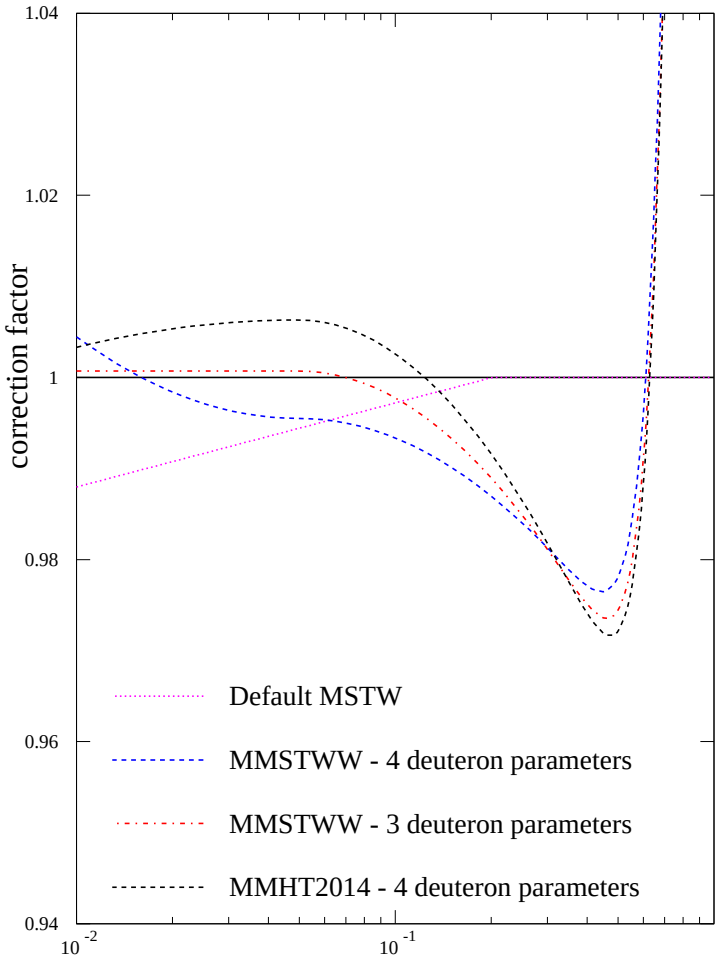
Correct dimuon cross-sections for missing small contribution, i.e. where charm is produced away from the interaction point. Previously assumed this was accounted for by acceptance corrections. Previous checks showed correction is a small effect on strange distribution.

Use NMC structure function data with $F_L(x, Q^2)$ correction very close to theoretical $F_L(x, Q^2)$ value. Very little effect.

Result for fitted deuteron correction.

Previously big improvement in fit for MMSTWW, but not exactly as expected at lower x.

Now more like that expected, and 4 parameters left free at NLO (and now NNLO). Uncertainty of about $0.5 - 1\%$. Feeds into PDF uncertainty.



Errors multiplicative not additive. Using χ^2 definition

$$\chi^2 = \sum_{i=1}^{N_{pts}} \left(\frac{D_i + \sum_{k=1}^{N_{corr}} r_k \sigma_{k,i}^{corr} - T_i}{\sigma_i^{uncorr}} \right)^2 + \sum_{k=1}^{N_{corr}} r_k^2,$$

where $\sigma_{k,i}^{corr} = \beta_{k,i}^{corr} T_i$ and $\beta_{k,i}^{corr}$ are the percentage error. Additive would use $\sigma_{k,i}^{corr} = \beta_{k,i}^{corr} D_i$. Previously did this for all but normalisation uncertainty.

Effectively if

$$D_i + \sum_{k=1}^{N_{corr}} \beta_{k,i}^{corr} D_i \sim f * D_i \qquad \text{or} \qquad T_i - \sum_{k=1}^{N_{corr}} \beta_{k,i}^{corr} T_i \sim T_i / f,$$

then

$$\chi^2 \sim \left(\frac{f * D_i - T_i}{\sigma_i^{uncorr}} \right)^2 \qquad \text{rather than} \qquad \chi^2 \sim \left(\frac{D_i - T_i/f}{\sigma_i^{uncorr}} \right)^2 = \left(\frac{f * D_i - T_i}{f * \sigma_i^{uncorr}} \right)^2.$$

Use standard penalty for normalisation shifts, rather than previous quartic penalty.Extremely little difference.

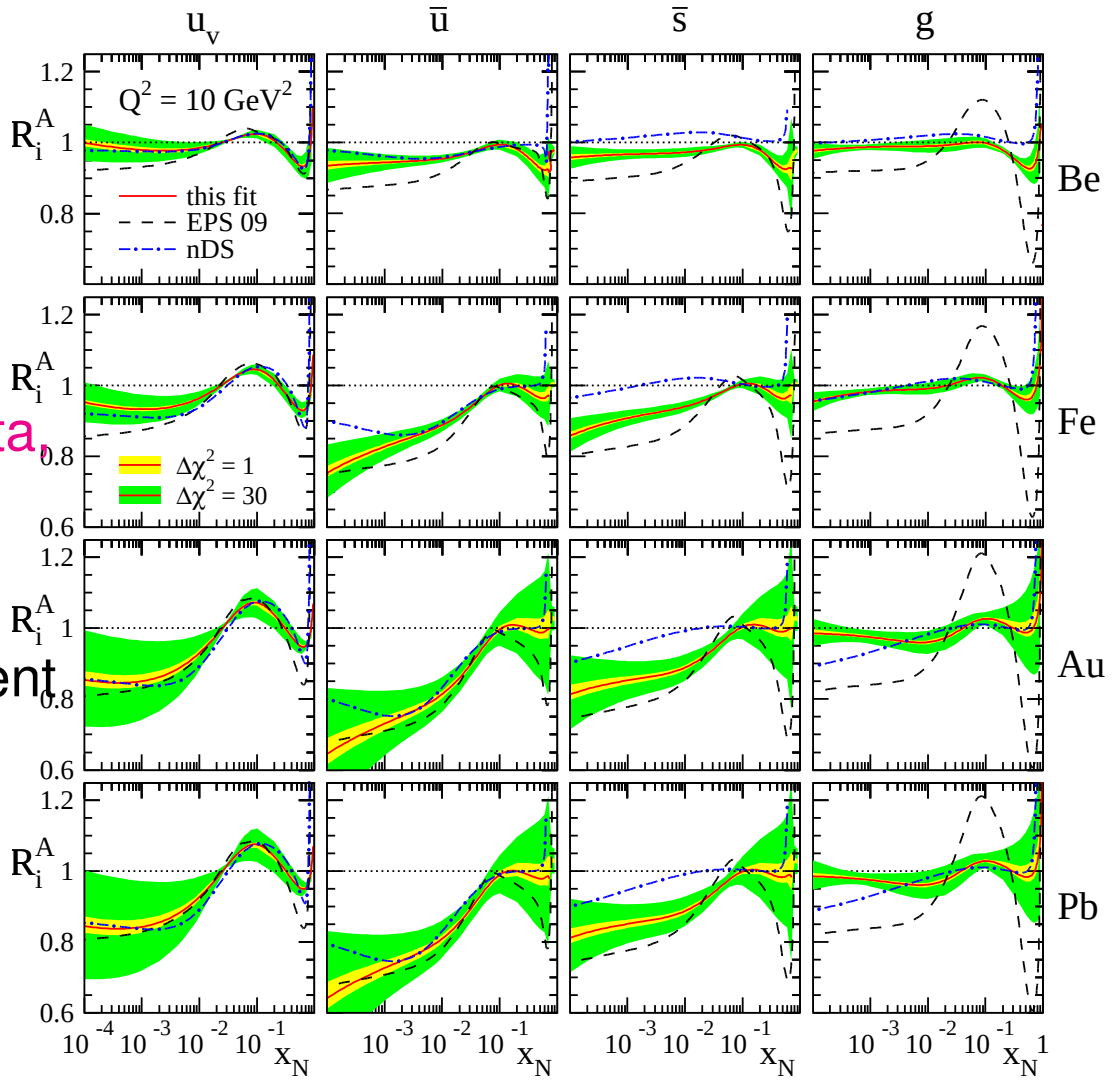
Strange branching ratio. Now avoid those determined by fits to dimuon data relying on PDF input. Also apply error which feeds into PDFs. Use $B_\mu = 0.092 \pm 10\%$ from hep-ex/9708014. Fits prefer $B_\mu = 0.085 - 0.091 \pm 15\%$.

Have been using de Florian, Sassot nuclear corrections.

Update to more recent version, de Florian, Sassot, Stratmann, Zurita, Phys.Rev. D85 (2012) 074028.

Mainly similar, but different correction for small-x strange.

Improves global fit by ~ 25 units - NuTeV F_2, HERA F_2, CMS jet.



Only small change in strange quark, (no effect on ATLAS, W,Z fit).

Changes in data sets.

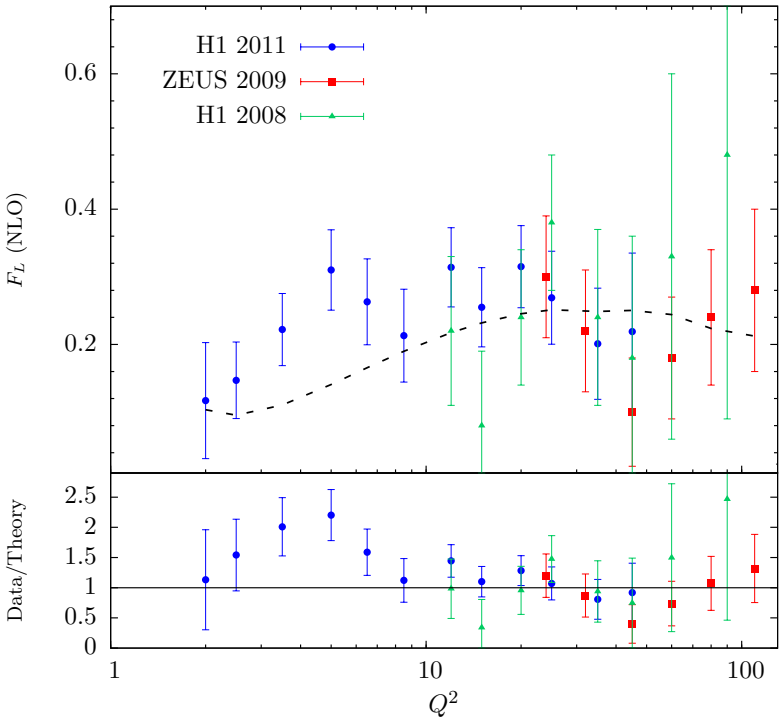
Replacement of HERA run I neutral and charged current data from HERA and ZEUS with combined data set with full treatment of correlated errors. Fit to data very good. Slightly better fit at NNLO.

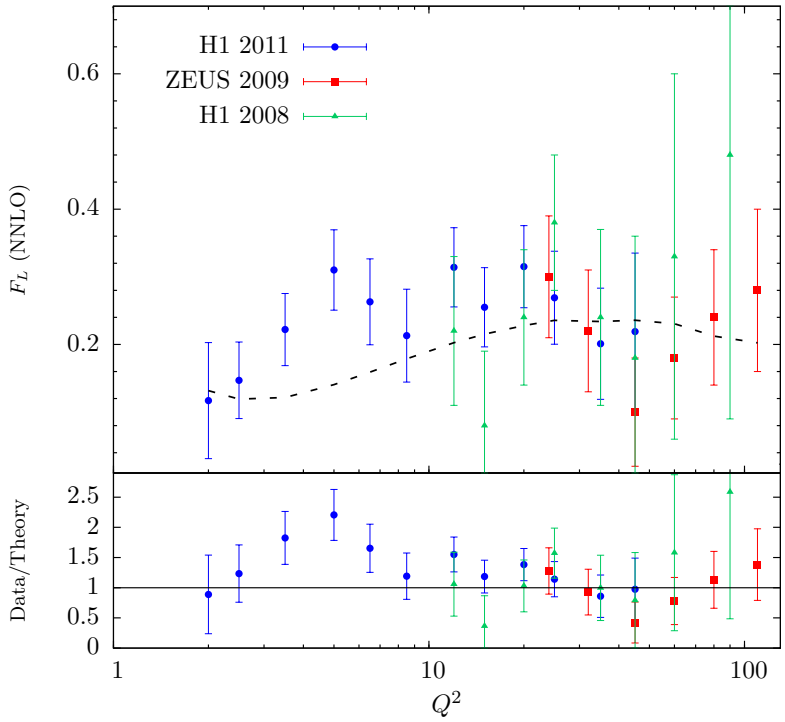
Inclusion of HERA combined data on $F_2^c(x, Q^2)$. Fit quality $\sim$ 60-80 for 52 points.

Inclusion of all direct published HERA $F_L(x, Q^2)$ measurements. Undershoot data a little at lower Q^2, but χ^2 not much more than one per point.

No inclusion of separate run II H1 and ZEUS data yet. Wait for Run II combination.

Fit quality at NLO and NNLO for $F_L(x, Q^2)$ data





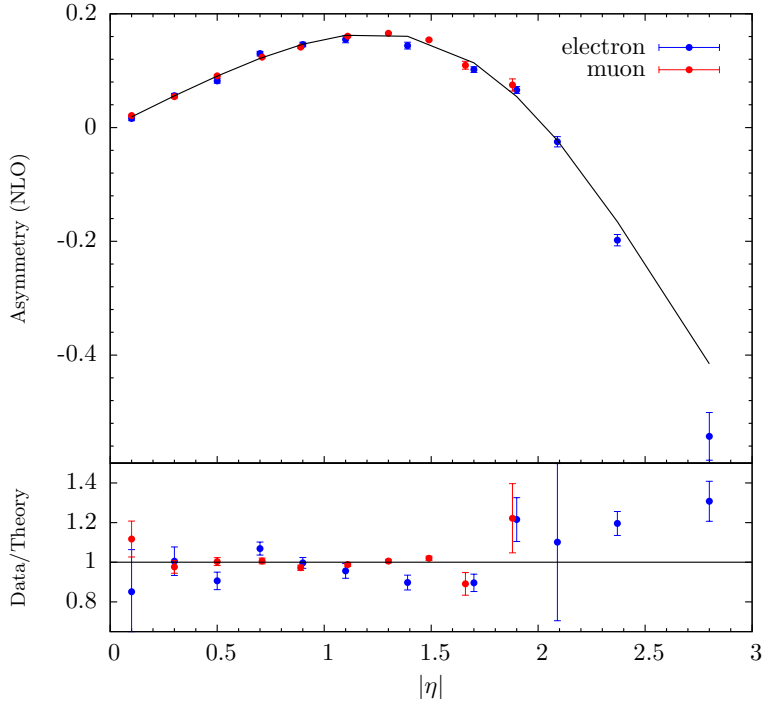
Inclusion of the CDF W-asymmetry data, the D0 electron asymmetry data $p_T > 25\mathrm{GeV}$ based on 0.75 fb^{-1} and new D0 muon asymmetry data for $p_T > 25\mathrm{GeV}$ based on 7.3 fb^{-1}.

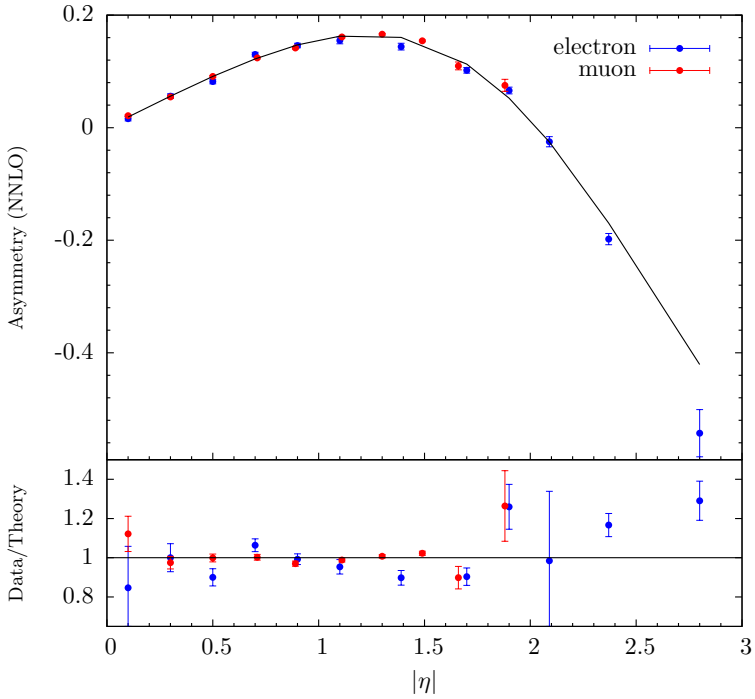
Include final numbers for CDF Z-rapidity data – final numbers changed after MSTW2008 fit. (Also include very small photon contribution in theory.) Very little change.

Not much change in PDFs (other than already seen in $u_V - d_V$).

At NLO $\alpha_S(M_Z^2) = 0.1199$ from 0.1202 and at NNLO $\alpha_S(M_Z^2) = 0.1180$ from 0.1171.

Fit quality at NLO and NNLO for D0 asymmetry data





LHC data on W,Z

Now using APPLGrid – MCFM and DYNNLO/FEWZ include the ATLAS W,Z rapidity data directly in the fit.

Before inclusion $\chi^2 \sim 1.6$ per point at NLO and actually $\chi^2 \sim 2$ per point at NNLO.

Inclusion leads some extra improvement at NLO, $\chi^2 \sim 1.3$, strongest pull on gluon PDF. Also goes to $\chi^2 \sim 1.3$ at NNLO. The most obvious change is in the strange quark.

$W^+ - W^-$ asymmetry no longer an issue at all both for ATLAS and CMS asymmetry data. Slightly better at NLO.

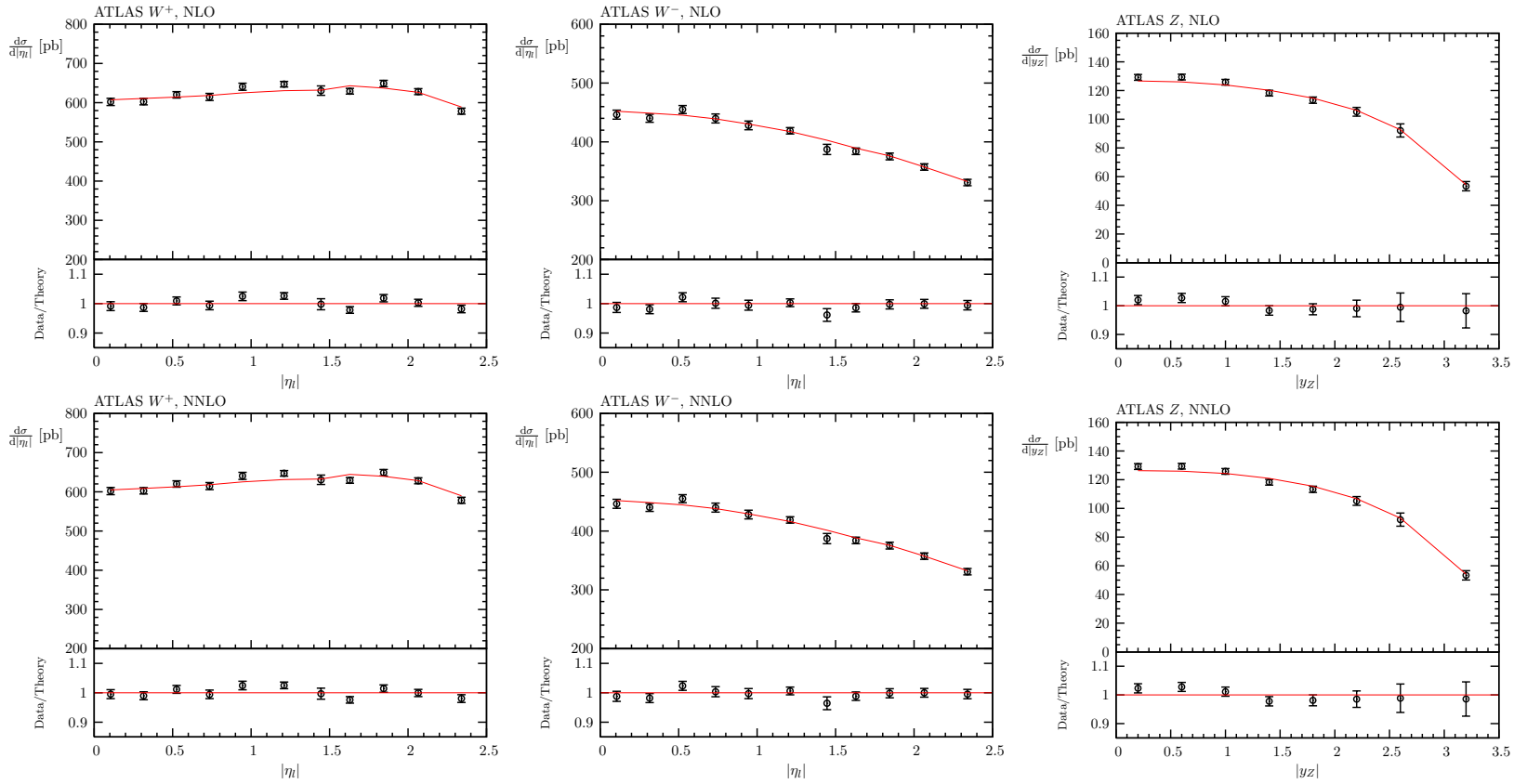
Include LHCb data on W^+, W^-, and $Z \to e^+e^-$. Both predicted/fit well at NLO. For the latter theory a bit low at NNLO at $y \sim 3.5$. Not evident in preliminary $Z \to \mu^+\mu^-$ data with higher precision.

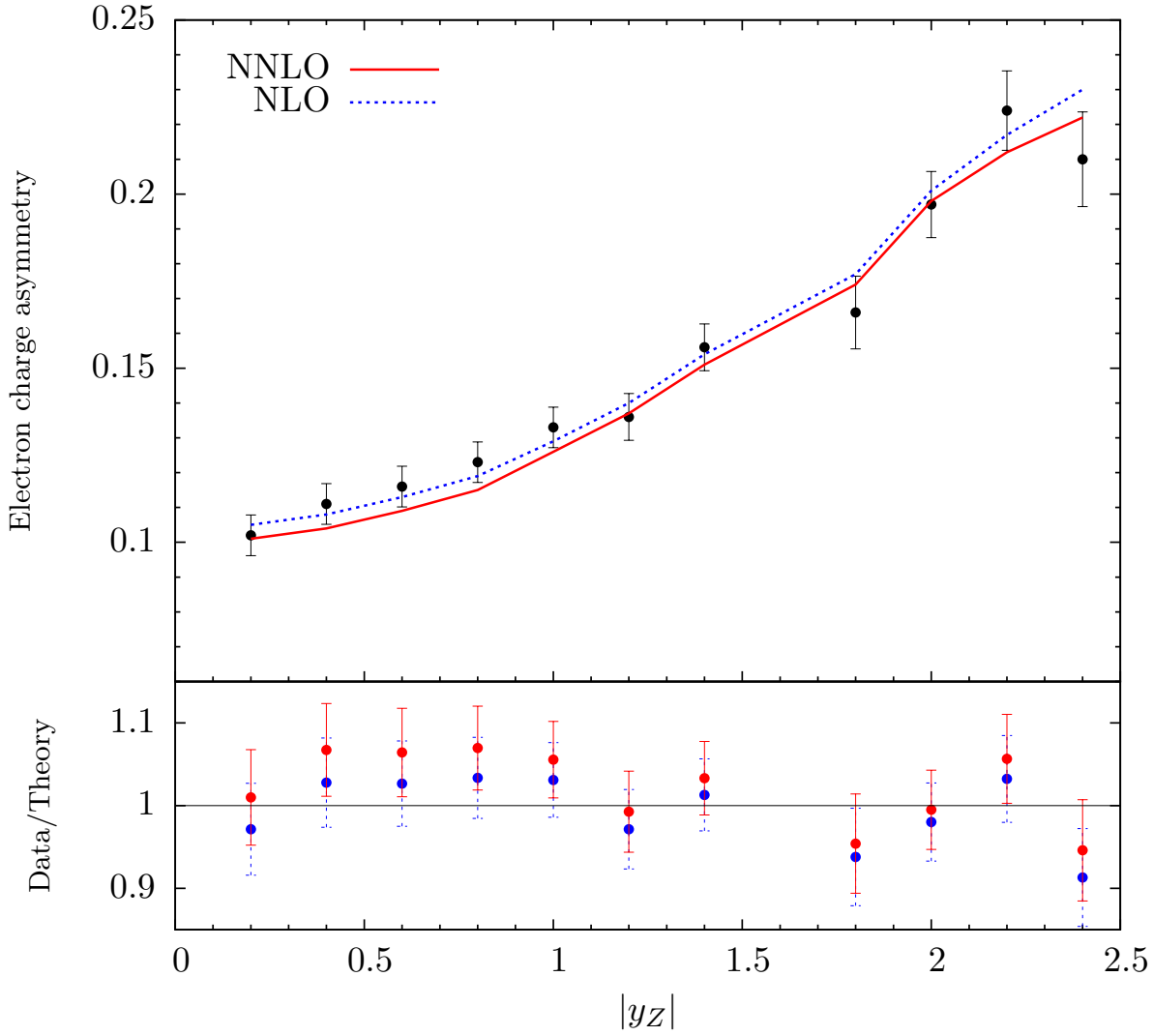
Include CMS data on $Z \to e^+e^-$, and ATLAS high mass Drell-Yan data. Again both predicted/fit well.

Fit CMS double differential Drell Yan data extending to low mass. NNLO fits enormously better than NLO at lowest mass $\sim 20 - 45\mathrm{GeV}$.

data set	$N_{\rm pts}$	MMSTWW	MMHT14 (no LHC)	MMHT14 (LHC)
NLO				
ATLAS W^+, W^-, Z	30	47	44	38
CMS W asym $p_T > 35$ GeV	11	9	16	7
CMS asym $p_T > 25, 30$ GeV	24	9	17	8
LHCb $Z \to e^+e^-$	9	13	13	13
LHCb W asym $p_T > 20$ GeV	10	12	14	12
CMS $Z \to e^+e^-$	35	21	22	19
ATLAS high-mass Drell-Yan	13	20	20	21
CMS double diff. Drell-Yan	132	385	396	372
NNLO				
ATLAS W^+, W^-, Z	30	72	53	39
CMS W asym $p_T > 35$ GeV	11	18	15	8
CMS asym $p_T > 25, 30$ GeV	24	18	17	9
LHCb $Z \to e^+e^-$	9	23	22	21
LHCb W asym $p_T > 20$ GeV	10	24	21	18
CMS $Z \to e^+e^-$	35	30	24	22
ATLAS high-mass Drell-Yan	13	18	16	17
CMS double diff. Drell Yan	132	159	151	150

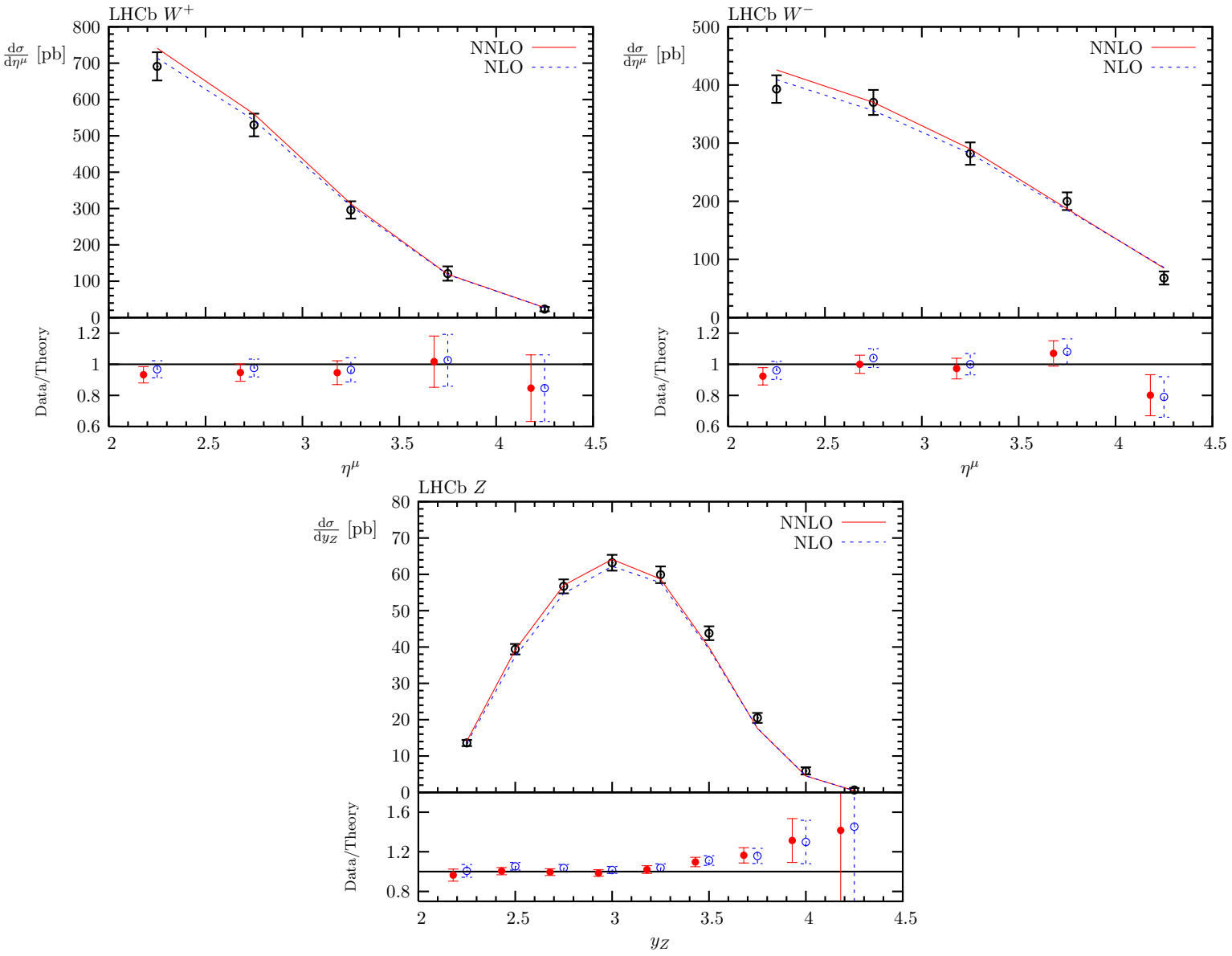
Fit quality at NLO and NNLO for ATLAS W, Z rapidity data





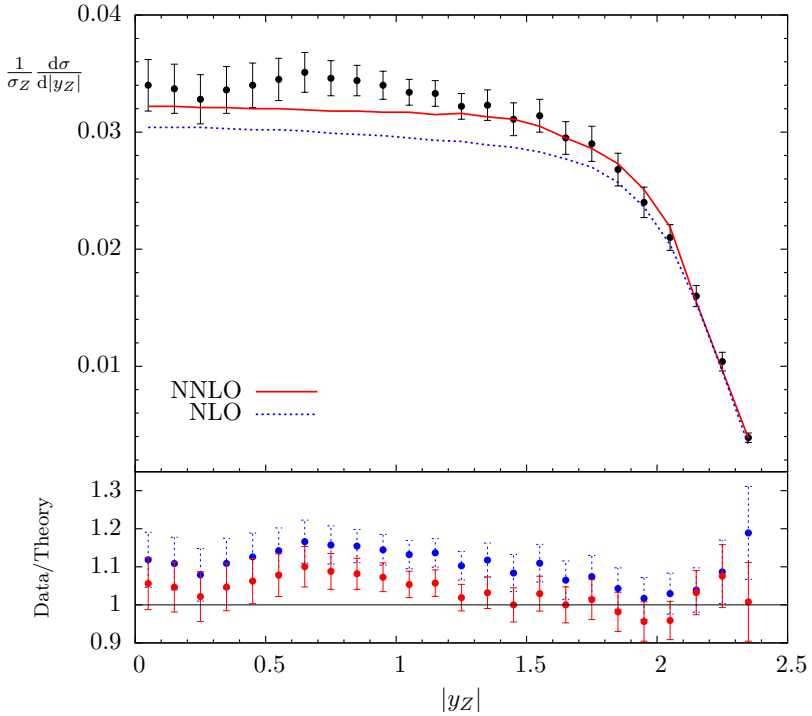
W asymmetry data now fit very well, though a little better at NLO than at NNLO.

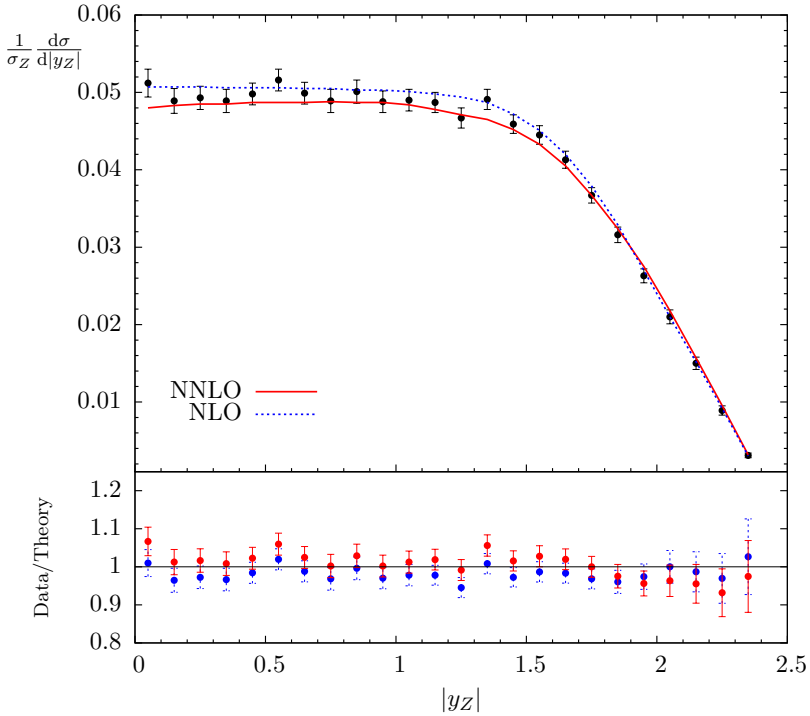
Fit quality at NLO and NNLO for LHCb W, Z rapidity data



CMS Drell Yan data.

Fit very poor at NLO in lowest mass bins (where it is effectively LO), even when data highly weighted.





Enormously improved fit quality at NNLO due to improvement in cross-sections.

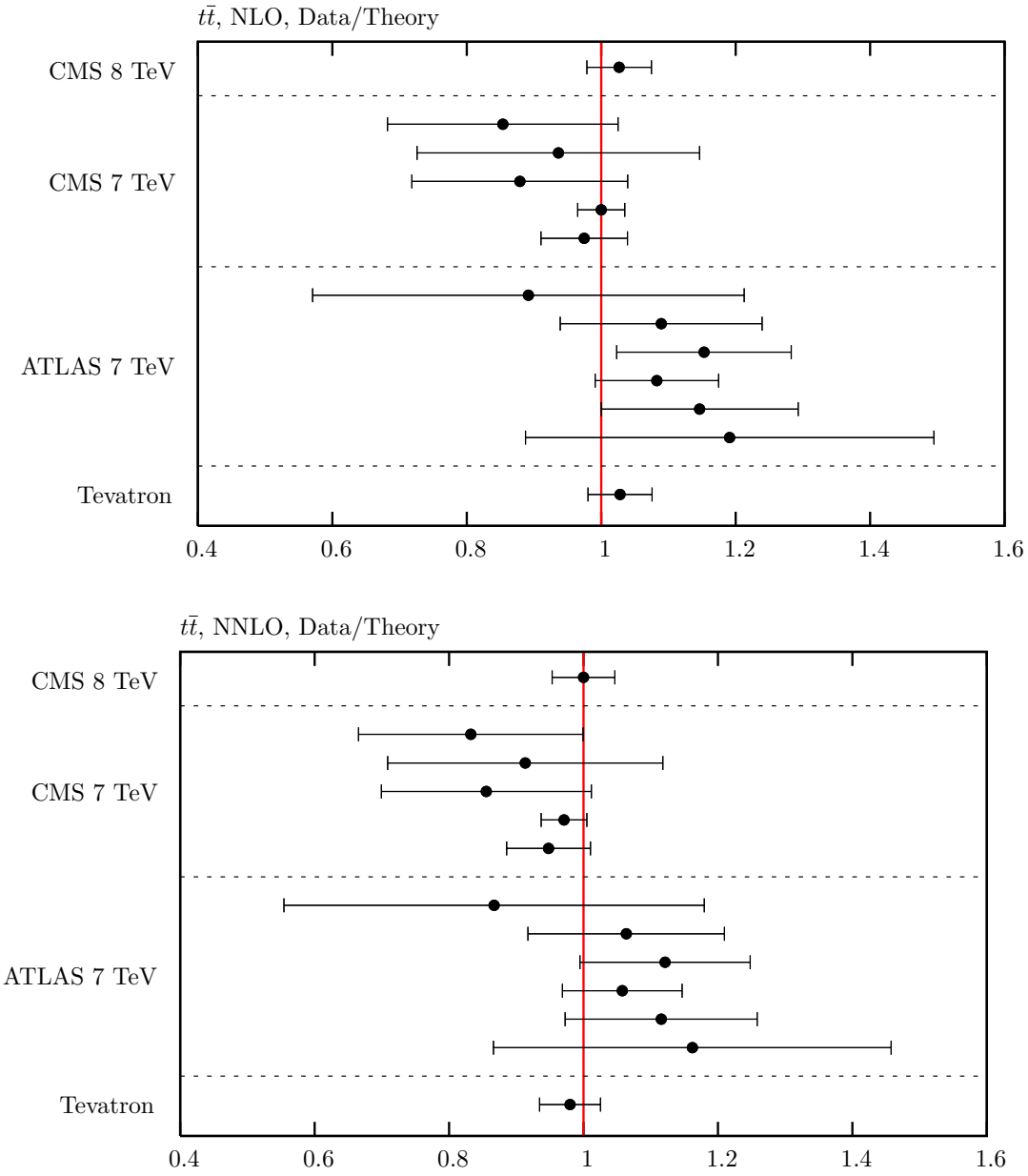
Sensitivity to strange fraction in quarks, but differs at NLO and NNLO and weak compared to direct constraint from di-muon data.

LHC data on $t\bar{t}$

Include data on $\sigma_{t\bar{t}}$ from Tevatron (combined cross section measurement from D0 and CDF), and all published data from ATLAS and CMS for 7TeV and one point at 8TeV. Use $m_t = 172.5\ \text{GeV}$ (value used in Tevatron combination) with an error of $1\ \text{GeV}$, with χ^2 penalty applied. Predictions and fit good, with NLO preferring masses slightly below $m_t = 172.5\ \text{GeV}$ and NNLO masses slightly above.

data set	N_{pts}	MMSTWW	MMHT14 (no LHC)	MMHT14 (LHC)
NLO				
Tevatron/ATLAS/CMS $\sigma_{t\bar{t}}$	13	8	10	7
NNLO				
Tevatron/ATLAS/CMS $\sigma_{t\bar{t}}$	13	8	11	8

Comparison of data and theory for $\sigma_{t\bar{t}}$ at NLO and NNLO.



LHC data on jets

At NLO also include CMS data together with ATLAS 7 TeV + 2.76 TeV data. Use ATLAS/HERAPDF study cuts, which eliminate lowest two p_T points in each bin and some highest p_T points.

The ATLAS $\chi^2 = 107/116$ and CMS $\chi^2 = 143/133$ before included directly – comparable to the best of the PDFs of other groups.

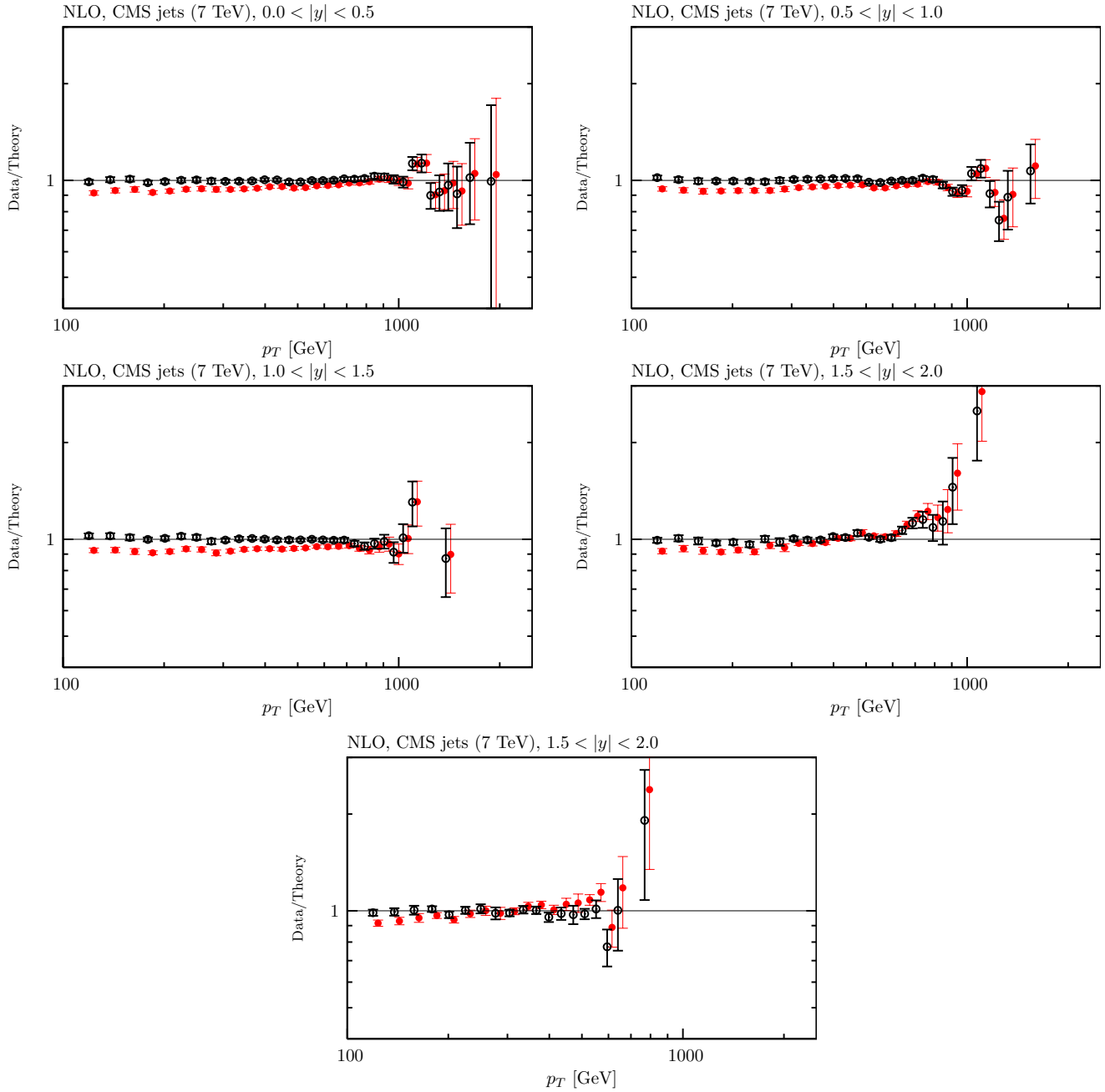
Simultaneous fit of CMS data together with ATLAS 7 TeV + 2.76 TeV leads to some improvement for CMS, and a small amount for ATLAS. The two experiments seem extremely compatible.

CMS inclusive jet data - updated. Previously the single pion uncertainties all correlated. Decision within collaboration made to decorrelate single pion systematics, i.e. to split the single pion source into 5 parts. Lowers χ^2 significantly, but no real change in PDFs. Allows slightly higher α_S.

At NLO final extracted $\alpha_S(M_Z^2) = 0.12001$.

LHC jets at NNLO?

Fit to CMS inclusive jet data.



LHC jet data at NNLO?

Enormous project of full NNLO calculation (Gehrmann-de-Ridder, Gehrmann, Glover and Pires) nearing completion. Some indications of full form of the correction.

As default at NNLO still fit Tevatron data which seems safe, since are always relatively near to threshold, and corrections do not obviously break down at lowest p_T.

However, omit LHC data. Lowest p_T not stable in threshold corrections, and large uncertainty at highest rapidity. (See also Carrazza and Pires.)

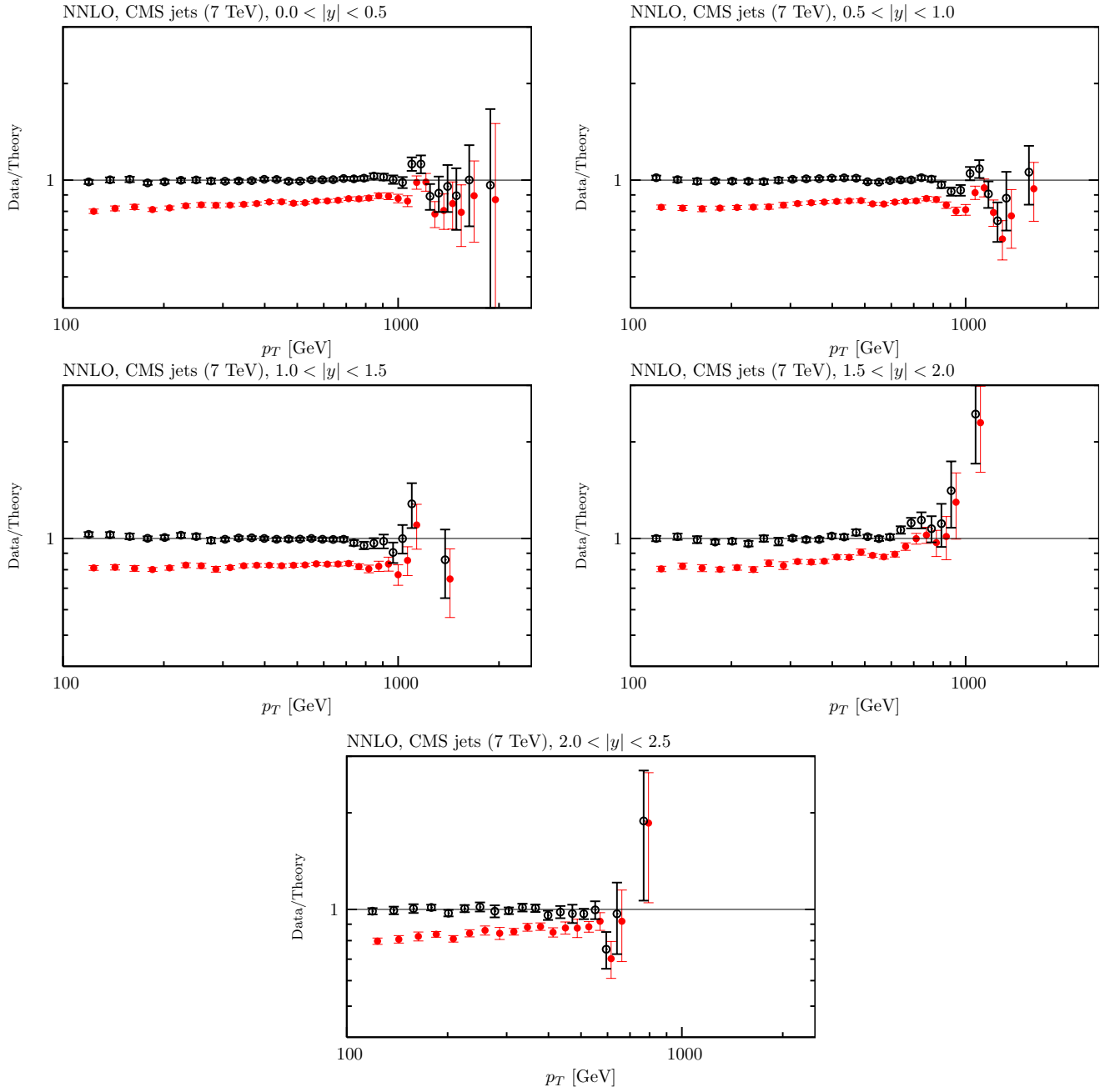
Try putting in very approx NNLO correction of $\sim 5 - 20\%$ positive correction growing at lower p_T. "Smaller" and "larger" K-factor with corrections of about $\sim 10\%$ and $\sim 20\%$ at $p_T = 100$ GeV - rapidity independent.

Prediction good. Fit quality a small amount worse than at NLO, though deteriorates slowly with larger K-factor.

data set	N_{pts}	MMSTWW	MMHT14 (no LHC)	MMHT14 (with LHC)
NLO				
ATLAS (2.76+7 TeV)	116	107	107	106
CMS (7 TeV)	133	140	143	138
NNLO small K-factor				
ATLAS (2.76+7 TeV)	116	(107)	(123)	(122) 115
CMS (7 TeV)	133	(142)	(137)	(138) 137
NNLO large K-factor				
ATLAS (2.76+7 TeV)	116	(117)	(132)	(132) 126
CMS (7 TeV)	133	(145)	(137)	(139) 139

ATLAS jet data deteriorates more than CMS, which with increase in systematics is largely insensitive to K-factor, though each prefers smaller one. Difficult to guess relative size of K-factor at two different energies.

Fit to CMS inclusive jet data at NNLO with large K-factor.



MMHT2014 PDFs compared to MSTW2008 PDFs.

Use same "dynamic tolerance" prescription to determine eigenvectors.

Typical tolerance $T = \Delta\chi^2 \sim 10$.

We now have 25 eigenvector pairs, rather than the 20 in MSTW or even the 23 in MMSTWW.

4 for each of $u_V(x, Q^2), d_v(x, Q^2)$.

5 for $S(x, Q^2)$, i.e. total light sea .

3 for $\bar{d}(x, Q^2) - \bar{u}(x, Q^2)$.

4 for $g(x, Q^2)$.

3 for $s^+(x, Q^2)$.

2 for $s^-(x, Q^2)$.

NLO

HERA structure – 6 eigenvector directions.

fixed target data DIS data – 13 eigenvector directions

LHC data – 4 eigenvector directions

Tevatron data– 9 eigenvector directions

Dimuon data – 8 eigenvector directions

E866 Drell Yan data – 10 eigenvector directions.

NNLO

HERA structure – 11 eigenvector directions.

fixed target data DIS data – 10 eigenvector directions

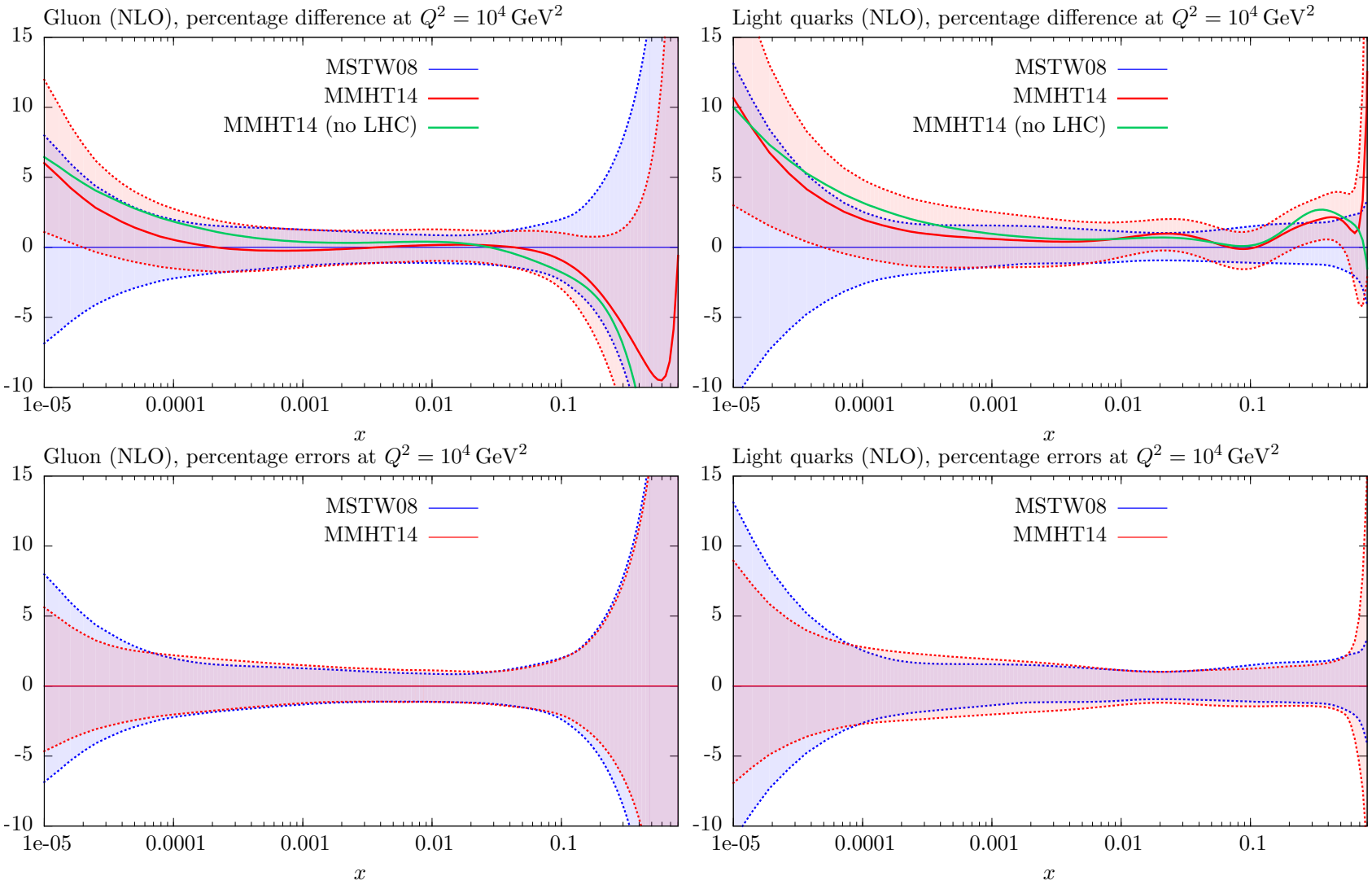
LHC data – 8 eigenvector directions

Tevatron data– 6 eigenvector directions

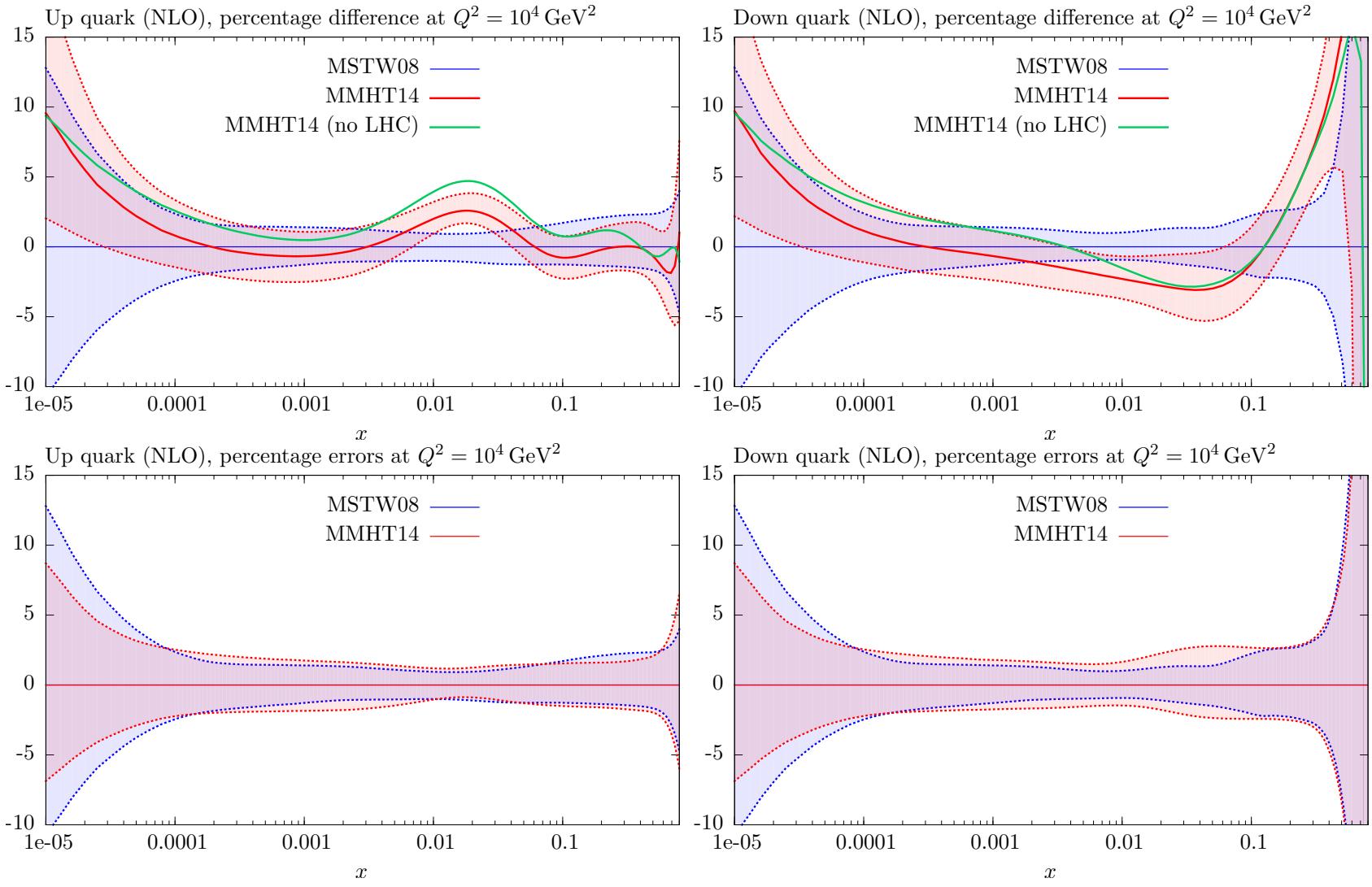
Dimuon data – 9 eigenvector directions

E866 Drell Yan data – 6 eigenvector directions.

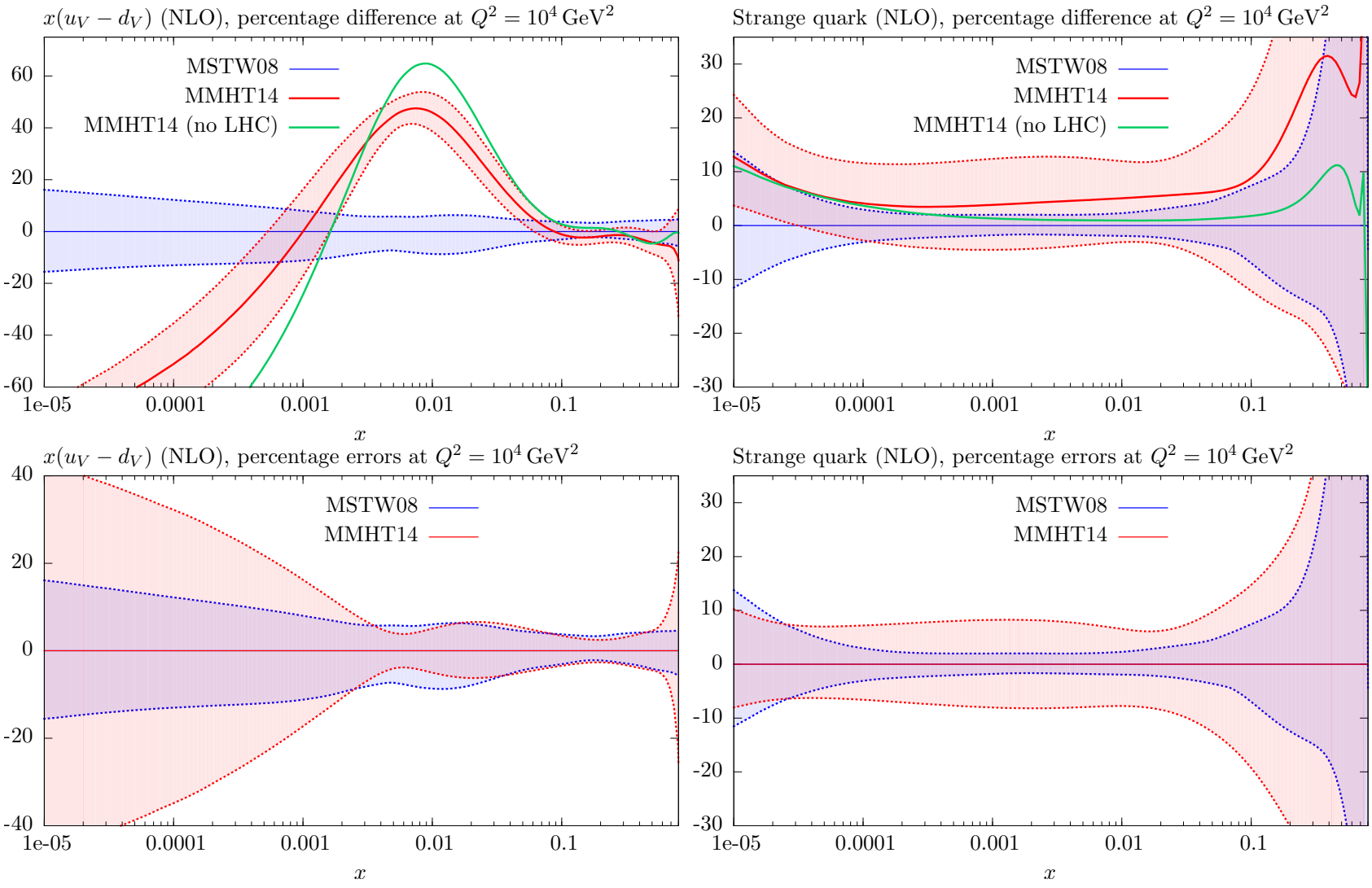
Comparison of PDFs at NLO



Change in NLO PDFs from all, including LHC data updates.

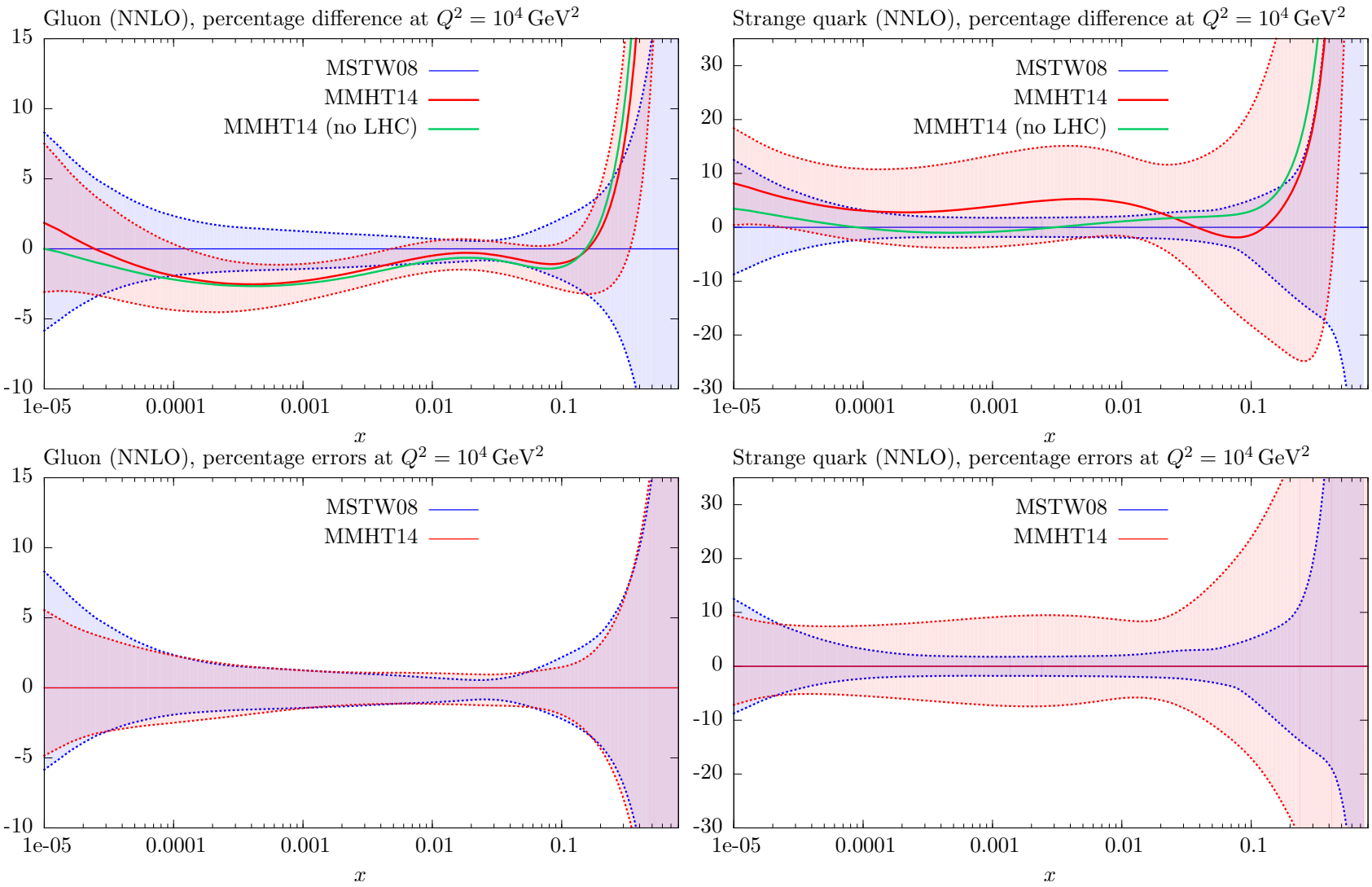


Change in NLO PDFs from all, including LHC data updates.



Change in NLO PDFs from all, including LHC data updates. Much expanded $s + \bar{s}$ uncertainty.

Comparison of PDFs at NNLO



Change in NNLO PDFs from all updates. Gluon uncertainty at high-x slightly greater than at NLO). At NNLO final extracted $\alpha_S(M_Z^2) = 0.11722$, but PDFs presented for $\alpha_S(M_Z^2) = 0.1180$.

$\alpha_S(m_Z^2)$ and PDF sets.

$\alpha_S(m_Z^2)$ coming out similar to 2008 fit. Still a NLO/NNLO difference. Both fairly compatible with global average, i.e.

NLO – $\alpha_S(m_Z^2) = 0.12001$.

NNLO – $\alpha_S(m_Z^2) = 0.11722$.

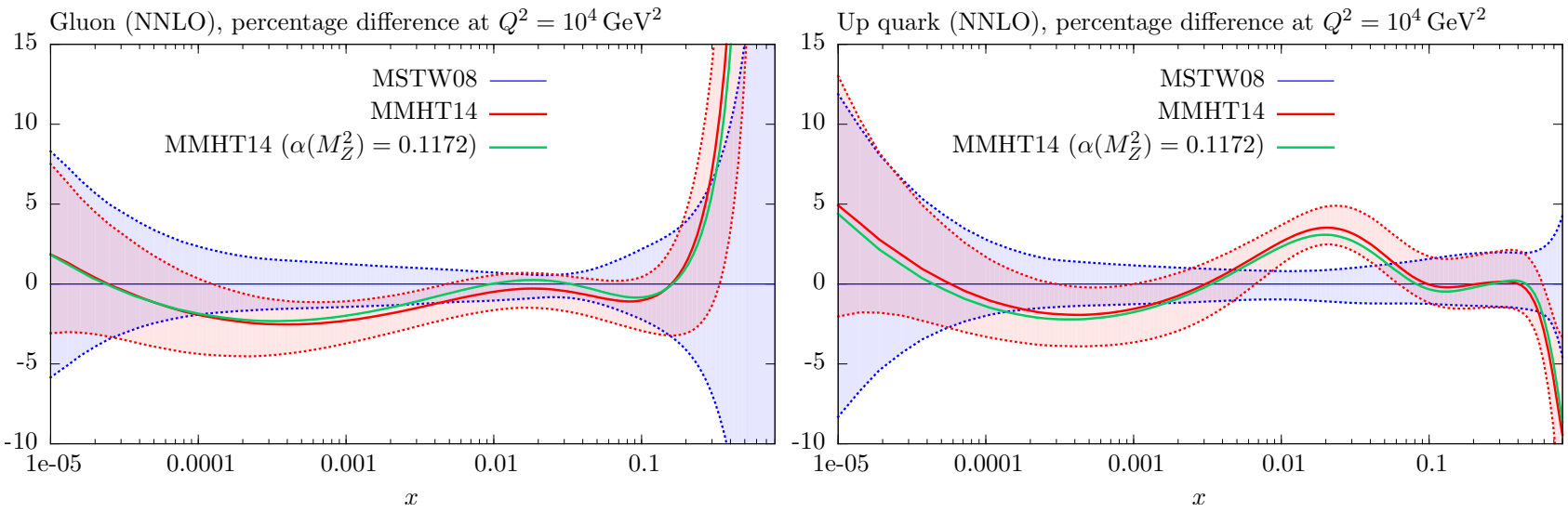
$\alpha_S(m_Z^2)^{\rm world} = 0.1186 \pm 0.0006$.

Decide to present MMHT2014 PDFs with eigenvectors at round value of $\alpha_S(m_Z^2) = 0.118$ at NNLO and at NLO also at $\alpha_S(m_Z^2) = 0.120$.

At NNLO forcing 0.118 leads to $\Delta\chi^2 < 2$

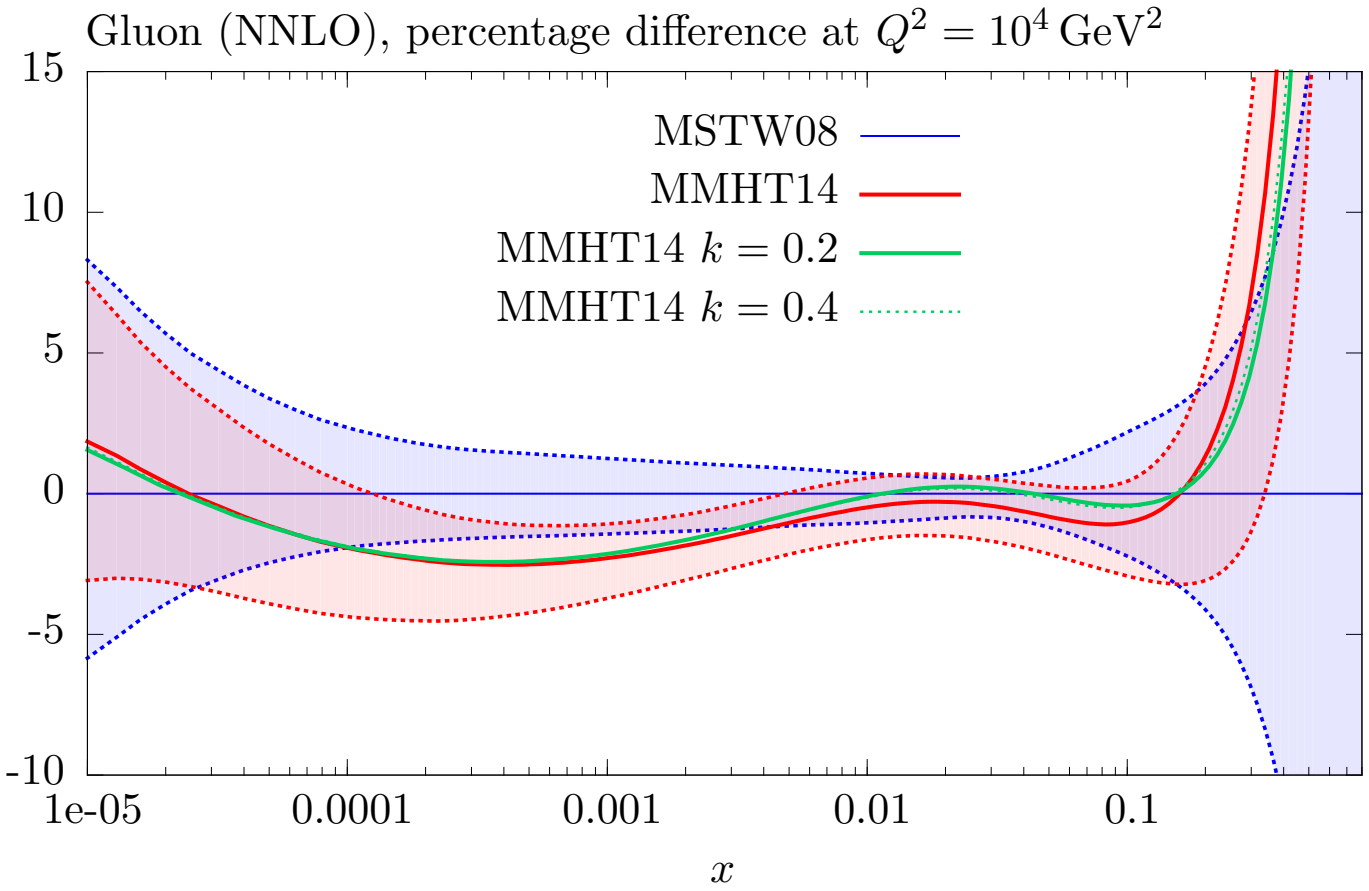
Forcing $\alpha_S(m_Z^2) = 0.118$ at NLO leads to $\Delta\chi^2 \sim 16$, but no single set deteriorates very significantly.

Comparison of PDFs at NNLO



NNLO PDFs for best fit value $\alpha_S(M_Z^2) = 0.11722$ compared to PDFs for $\alpha_S(M_Z^2) = 0.1180$.

Comparison of PDFs at NNLO when LHC jet data included



NNLO PDFs with LHC jet data included in fit with both K-factor choices compared to standard NNLO PDFs.

NLO total cross sections

	MSTW08	MMHT14 no LHC	MMHT14
W Tevatron (1.96 TeV)	$2.659^{+0.057}_{-0.045}$	2.685	$2.645^{+0.058}_{-0.049}$
Z Tevatron (1.96 TeV)	$0.2426^{+0.0054}_{-0.0043}$	0.2486	$0.2442^{+0.0049}_{-0.0043}$
W^+ LHC (7 TeV)	$5.960^{+0.129}_{-0.097}$	6.107	$5.974^{+0.092}_{-0.086}$
W^- LHC (7 TeV)	$4.192^{+0.092}_{-0.071}$	4.181	$4.163^{+0.069}_{-0.061}$
Z LHC (7 TeV)	$0.931^{+0.020}_{-0.014}$	0.941	$0.932^{+0.013}_{-0.013}$
W^+ LHC (14 TeV)	$12.07^{+0.24}_{-0.21}$	12.43	$12.17^{+0.20}_{-0.18}$
W^- LHC (14 TeV)	$9.107^{+0.19}_{-0.16}$	9.16	$9.10^{+0.15}_{-0.14}$
Z LHC (14 TeV)	$2.001^{+0.040}_{-0.032}$	2.035	$2.016^{+0.031}_{-0.033}$
Higgs Tevatron	$0.658^{+0.021}_{-0.027}$	0.636	$0.644^{+0.021}_{-0.022}$
Higgs LHC (7 TeV)	$11.39^{+0.16}_{-0.19}$	11.26	$11.28^{+0.21}_{-0.20}$
Higgs LHC (14 TeV)	$37.93^{+0.42}_{-0.60}$	37.67	$37.63^{+0.67}_{-0.59}$
$t\bar{t}$ Tevatron	$6.85^{+0.19}_{-0.13}$	6.89	$6.82^{+0.18}_{-0.17}$
$t\bar{t}$ LHC (7 TeV)	$162.0^{+4.3}_{-5.4}$	157.0	$158.6^{+4.5}_{-4.5}$
$t\bar{t}$ LHC (14 TeV)	903.8^{+16}_{-17}	886.7	891.9^{+18}_{-18}

Few changes greater than one sigma (PDF uncertainty only).

NNLO total cross sections

	MSTW08	MMHT14 no LHC	MMHT14
W Tevatron (1.96 TeV)	$2.746^{+0.049}_{-0.042}$	2.803	$2.782^{+0.056}_{-0.056}$
Z Tevatron (1.96 TeV)	$0.2507^{+0.0048}_{-0.0041}$	0.2574	$0.2559^{+0.0052}_{-0.0046}$
W^+ LHC (7 TeV)	$6.159^{+0.111}_{-0.099}$	6.214	$6.197^{+0.103}_{-0.092}$
W^- LHC (7 TeV)	$4.310^{+0.078}_{-0.069}$	4.355	$4.306^{+0.067}_{-0.076}$
Z LHC (7 TeV)	$0.9586^{+0.020}_{-0.014}$	0.9695	$0.9638^{+0.014}_{-0.013}$
W^+ LHC (14 TeV)	$12.39^{+0.22}_{-0.21}$	12.49	$12.48^{+0.22}_{-0.18}$
W^- LHC (14 TeV)	$9.33^{+0.16}_{-0.16}$	9.39	$9.32^{+0.15}_{-0.14}$
Z LHC (14 TeV)	$2.051^{+0.035}_{-0.033}$	2.069	$2.065^{+0.035}_{-0.030}$
Higgs Tevatron	$0.853^{+0.028}_{-0.029}$	0.877	$0.874^{+0.024}_{-0.030}$
Higgs LHC (7 TeV)	$14.40^{+0.17}_{-0.23}$	14.54	$14.56^{+0.21}_{-0.29}$
Higgs LHC (14 TeV)	$47.50^{+0.47}_{-0.74}$	47.61	$47.69^{+0.63}_{-0.88}$
$t\bar{t}$ Tevatron	$7.19^{+0.17}_{-0.12}$	7.54	$7.51^{+0.21}_{-0.20}$
$t\bar{t}$ LHC (7 TeV)	$171.1^{+4.7}_{-4.8}$	176.5	$175.9^{+3.9}_{-5.5}$
$t\bar{t}$ LHC (14 TeV)	953.3^{+16}_{-18}	969.0	969.9^{+16}_{-20}

Again few changes greater than one sigma (PDF uncertainty only).

Conclusions

Release of updated MMHT2014 PDFs imminent.

Improvement in parameterisation, heavy flavour treatments, nuclear corrections, and branching ratio for dimuon data. Inclusion of up-to-date HERA and Tevatron data.

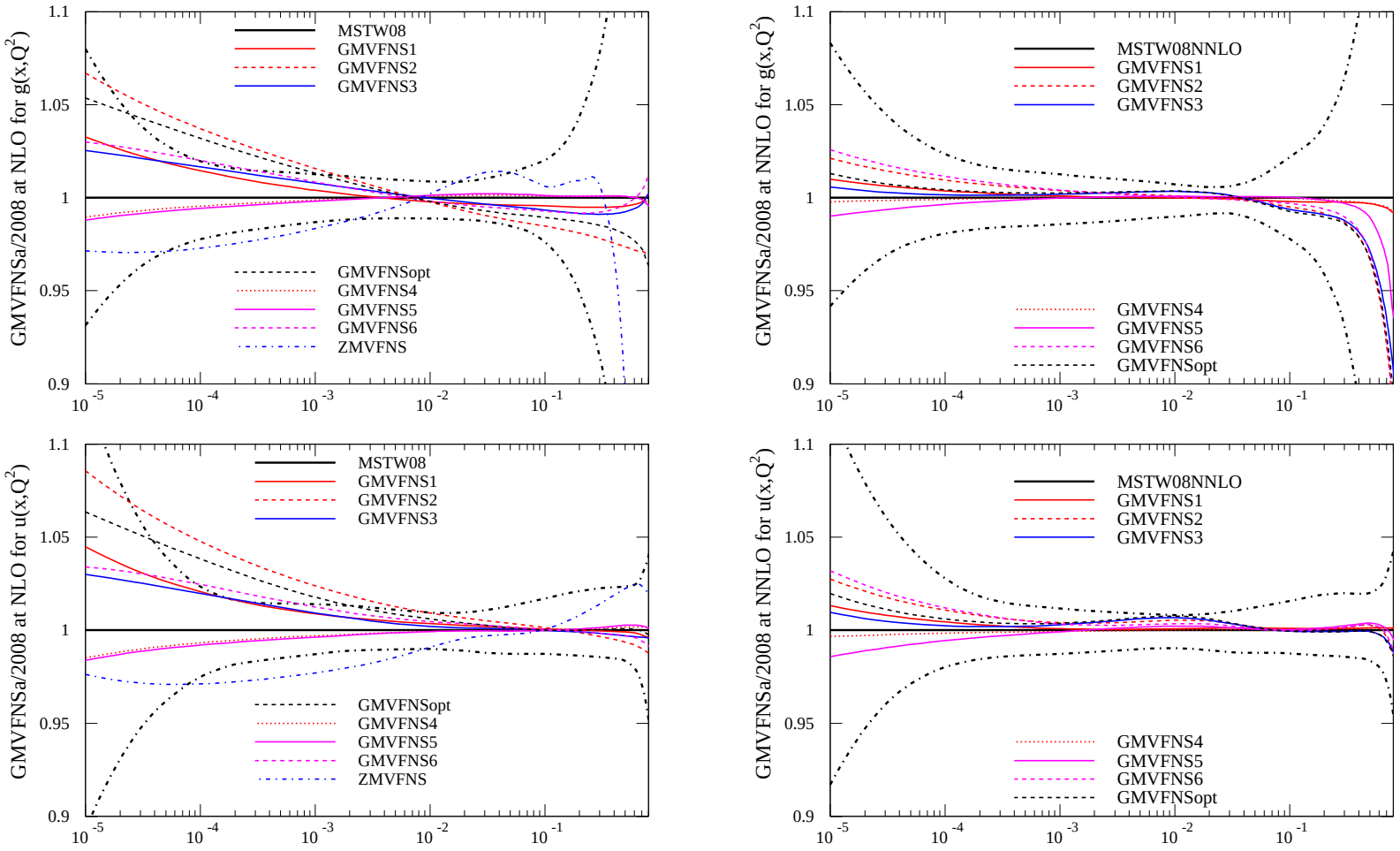
Also directly included most relevant published LHC data, i.e. ATLAS, CMS, LHCb W, Z rapidity data, top cross section data and all published ATLAS and CMS inclusive jet data (not at NNLO). Predictions for most of these data good. The fit is always good (except low mass Drell-Yan at NLO). Overall fit better at NNLO than at NLO.

Few dramatic effects on PDFs. Mainly slight increase in strange quark and increased strange uncertainty.

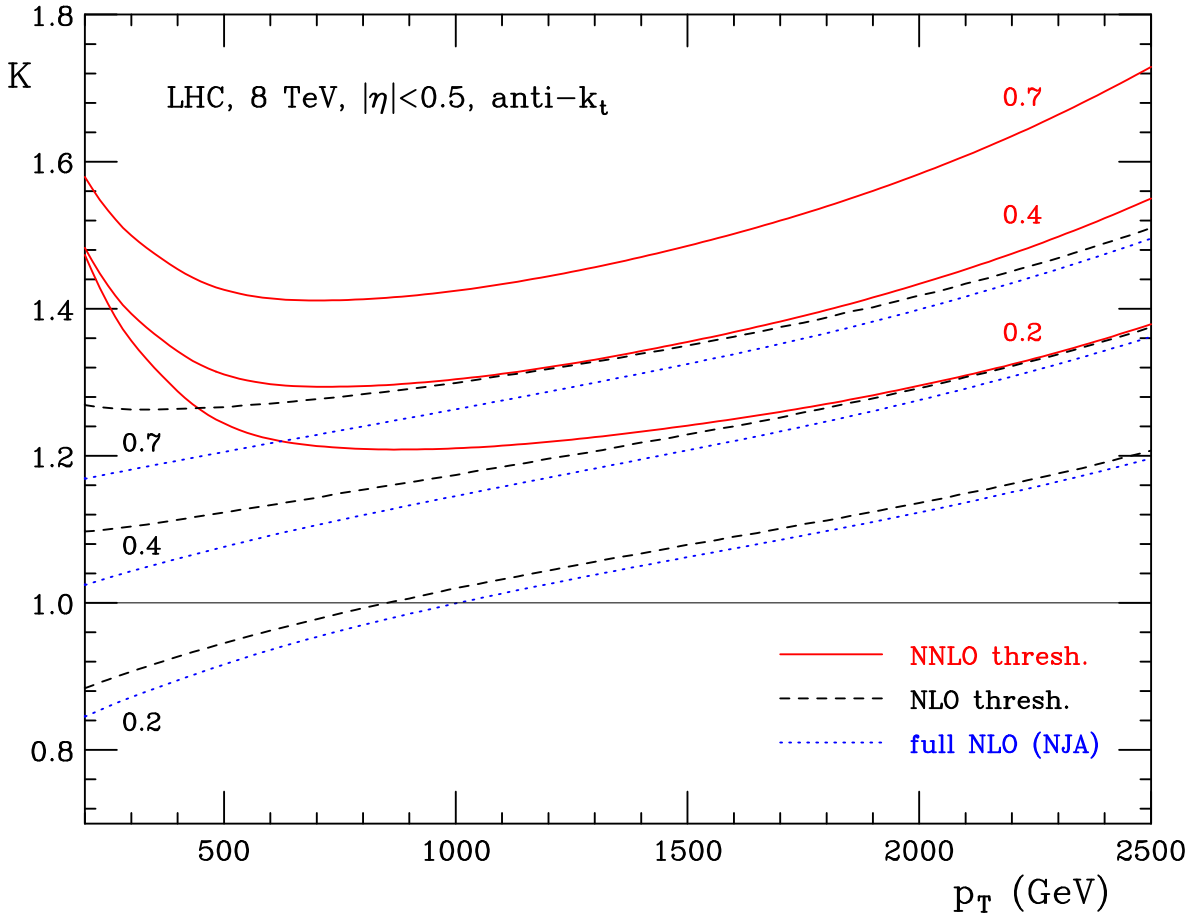
Set with uncertainties available with $\alpha_S(M_Z^2) = 0.135$ at LO, with $\alpha_S(M_Z^2) = 0.118$ and 120 at NLO and with $\alpha_S(M_Z^2) = 0.118$ at NNLO. Sets contain 25 eigenvector pairs.

In general predictions remain very close to those with MSTW2008 PDFs. The details of quark and anti-quark breakdown slightly changed.

Back -up



Using smoother schemes leads to some change in PDFs, with tendency for slight increase at small x and slight decrease at high x for gluon. Much smaller at NNLO than NLO. No real change in $\alpha_S(M_z^2)$.

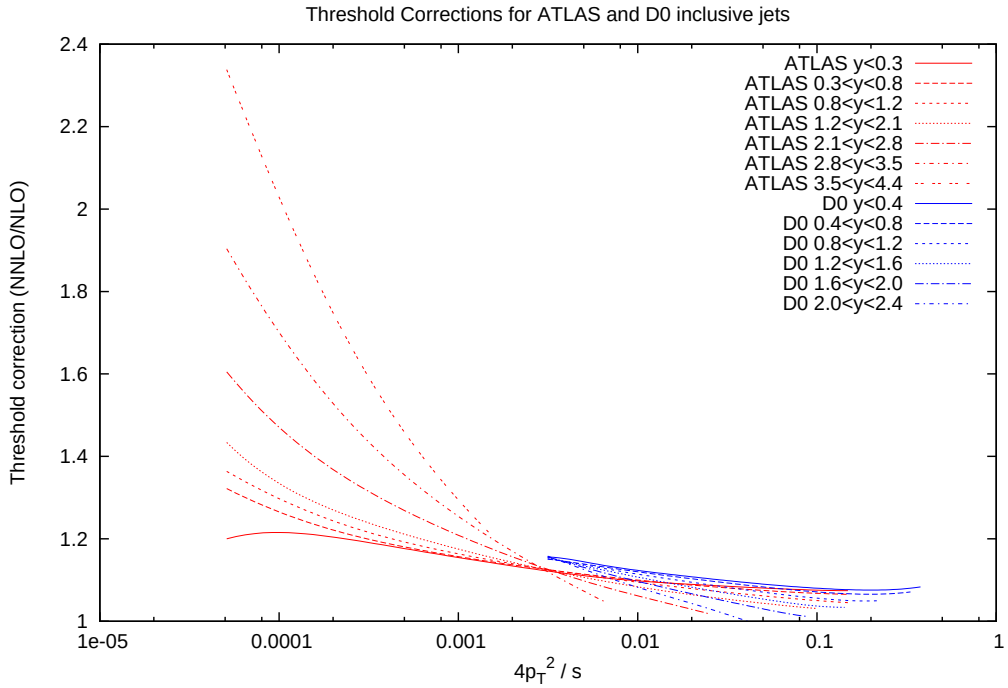


Recent improved threshold calculation from de Florian et al. (arXiv:1310.7192) has built in R dependence. Shows variation at NLO but little extra R dependence at NNLO. Still has problems at low p_T

LHC jet data at NNLO?

For Tevatron data use approximate "threshold" corrections (Kidonakis and Owens), $\sim 10\%$ positive correction.

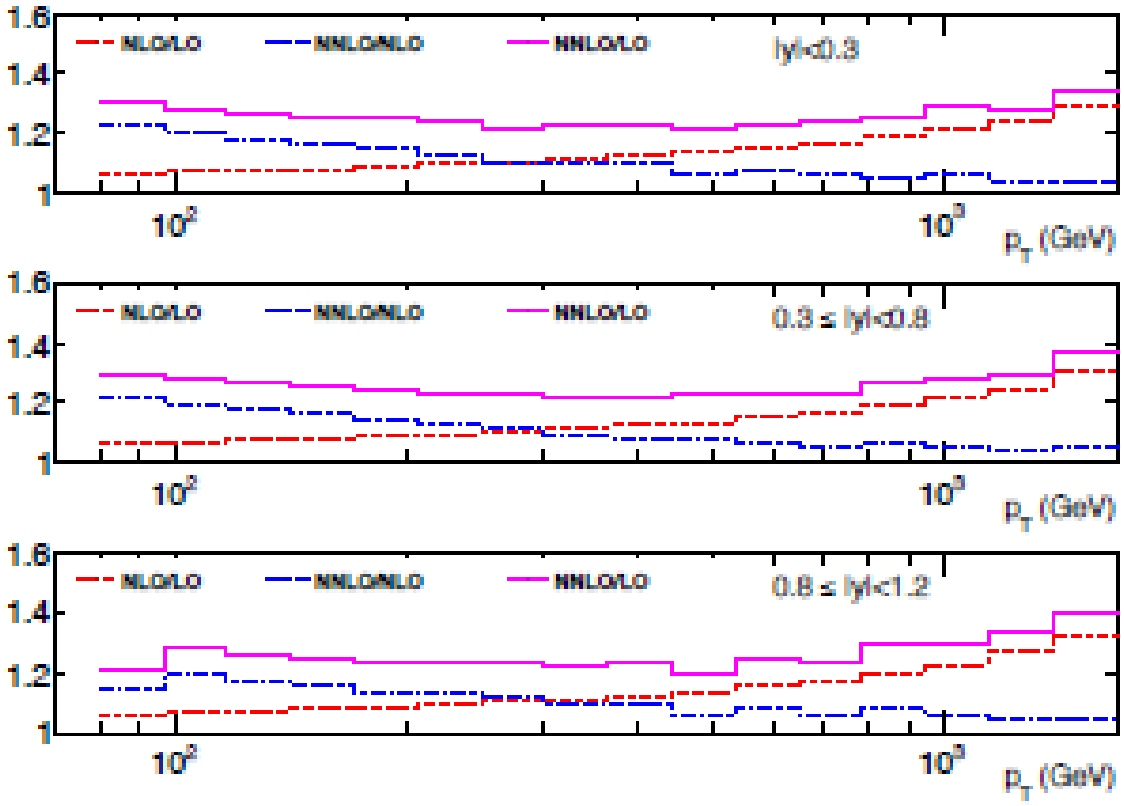
LHC corrections very similar for highish x probed at the Tevatron, but blow up when low x probed at the LHC, i.e. far from threshold.



Enormous project of full NNLO calculation (Gehrmann-de-Ridder, Gehrmann, Glover and Pires) nearing completion. Some indications of full form of the correction.

Inclusive jet production: double differential distributions

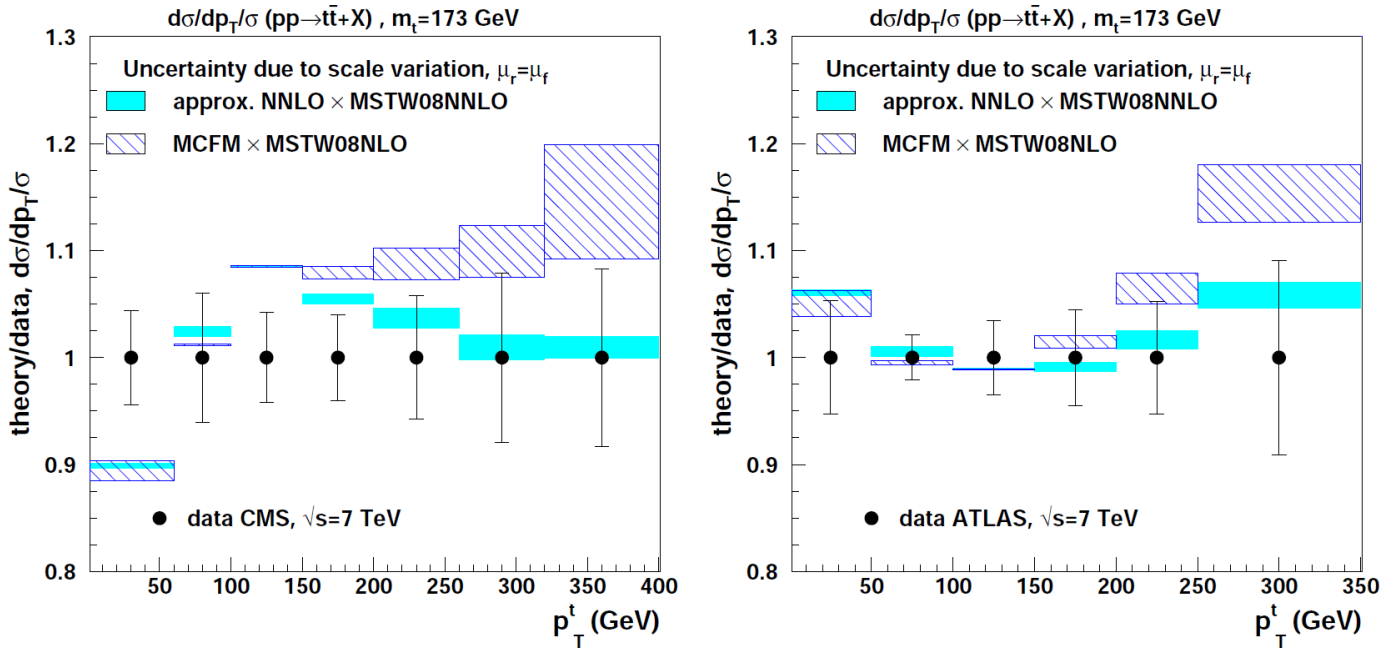
R = 0.4



Appears to be fairly similar to "threshold" correction near threshold, now verified by de Florian et al.. Overall $\sim 5-20\%$ positive correction growing at lower p_T.

Differential Data

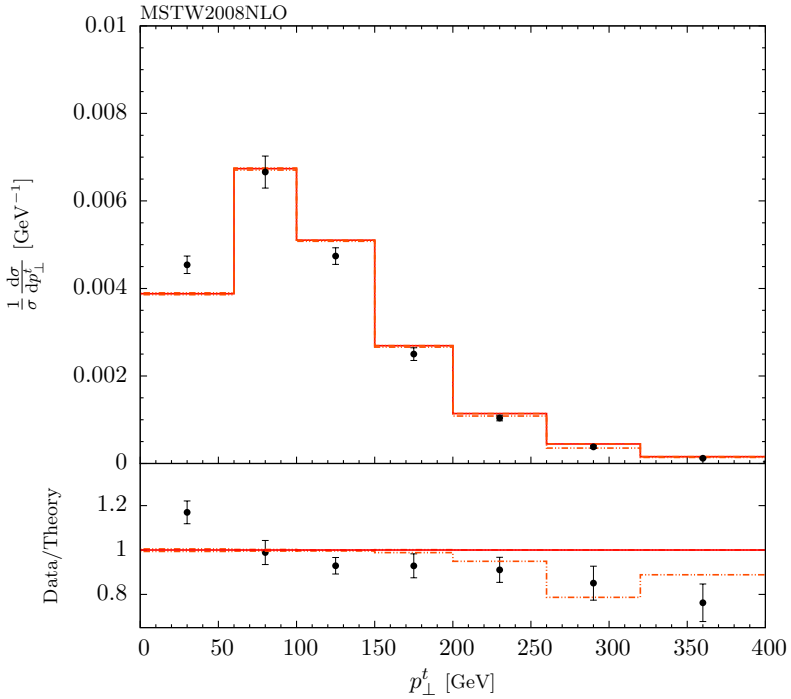
As it improves differential top production data will help constrain the gluon.

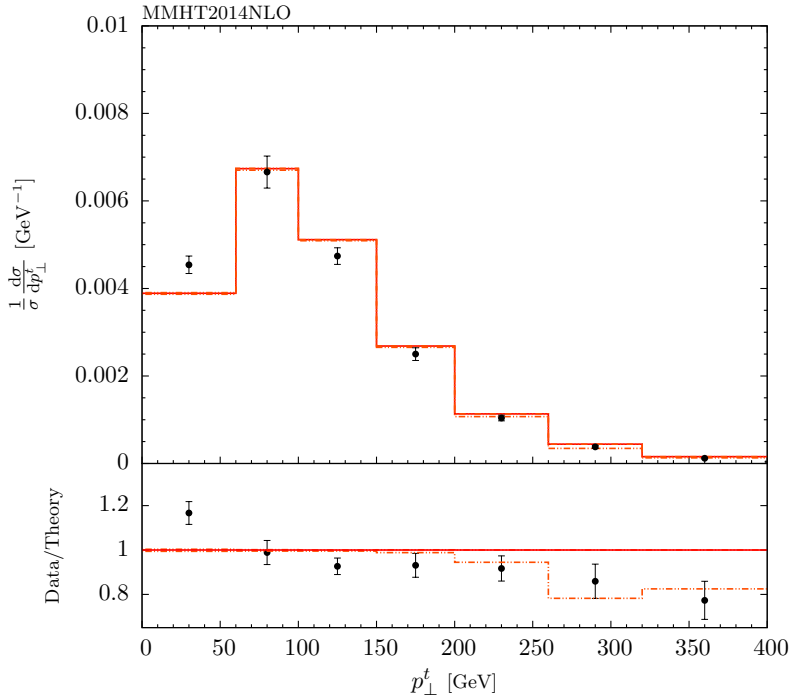


However, here potentially inclusion of NNLO is very important as available approximation using threshold resummation (Guzzi, Lipka, Moch) implies. Softer PDF currently preferred at NLO, contrary to requirement of inclusive cross-section, may be misleading.

New results – plots by Harland-Lang

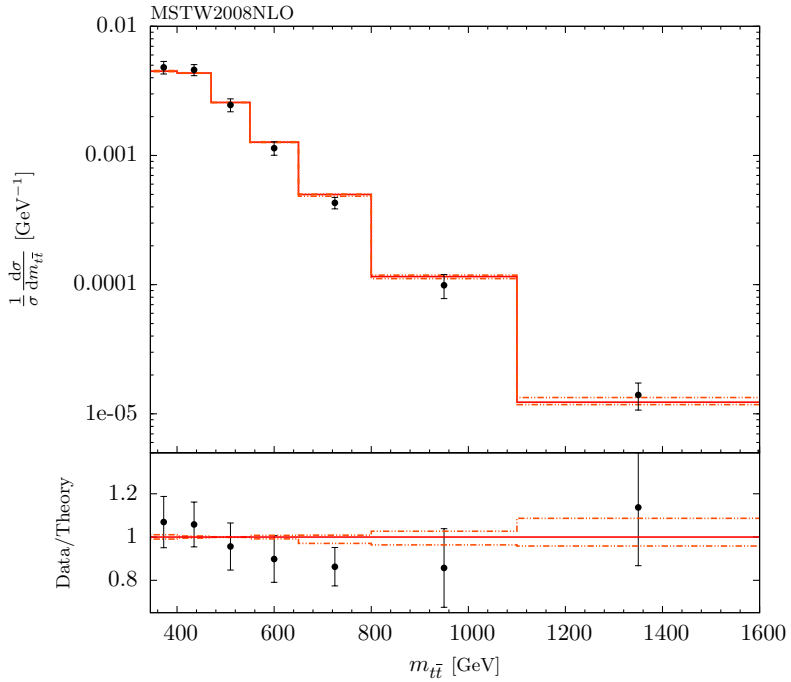
p_T distributions - CMS data

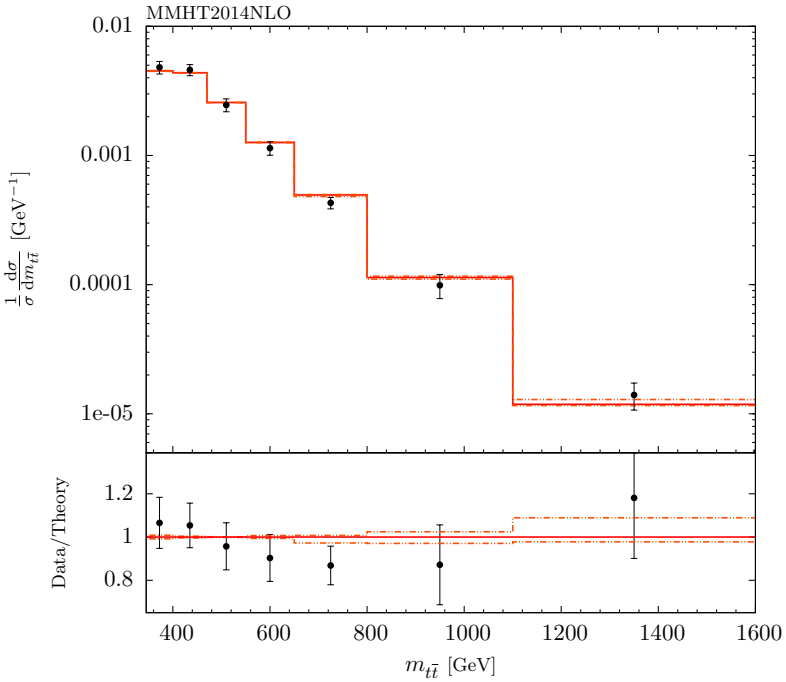




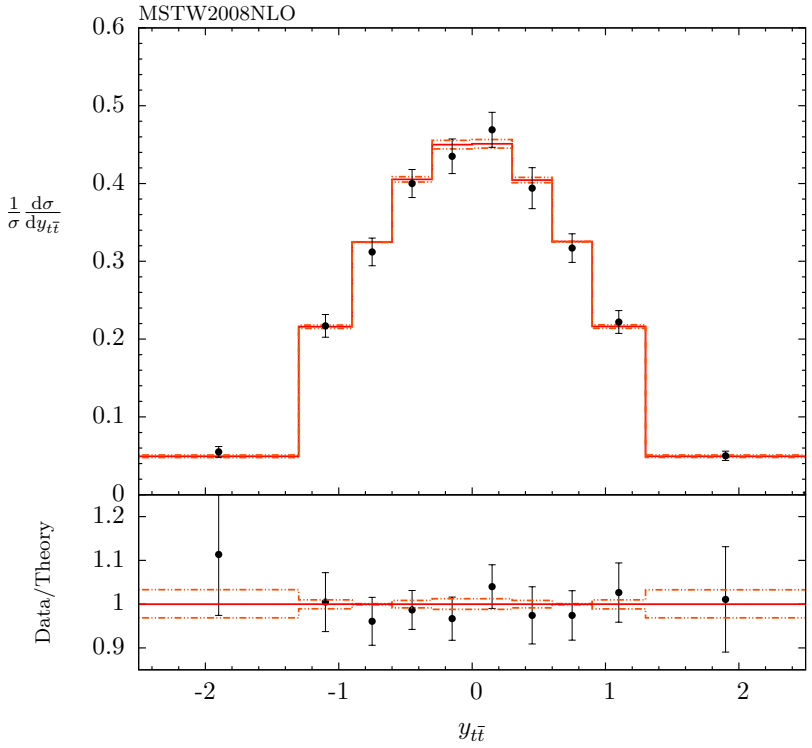
Very little difference between MSTW2008 and MMHT2014 predictions.

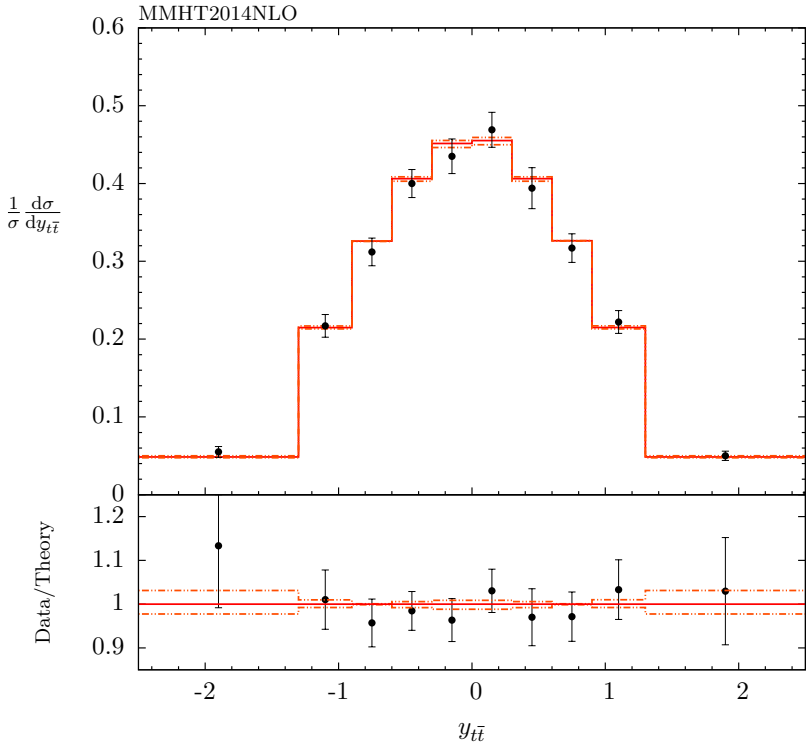
$m_{t\bar{t}}$ distributions - CMS data





$y_{t\bar{t}}$ distributions - CMS data





Randomly distributed "Hessian" PDF sets. G. Watt, RST

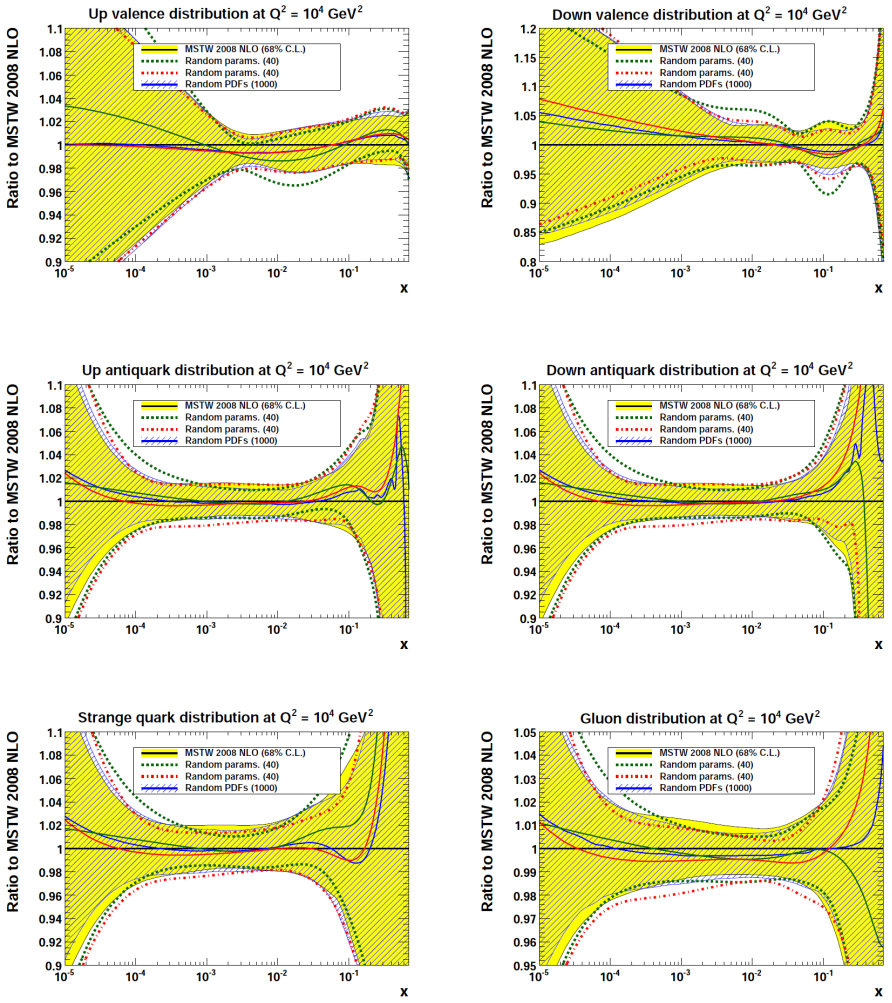
Study supported correctness of "dynamic tolerance" approach. Easiest in Hessian study with eigenvectors.

However, can generate "random" PDF sets directly from parameters and variation from eigenvectors.

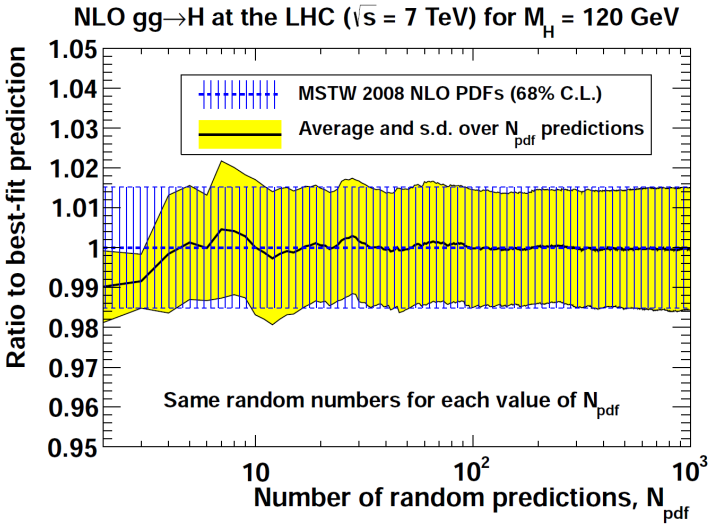
$$a_i(S_k) = a_i^0 + \sum_{j=1}^{n} e_{ij}\left(\pm t_j^{\pm}\right)|R_{jk}|$$

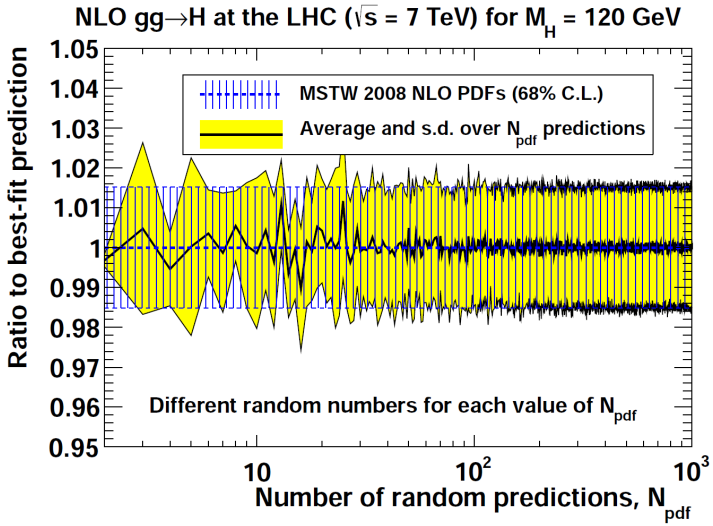
$(k = 1, \ldots, N_{\mathrm{pdf}})$. Or from eigenvectors directly (see LHCb study and De Lorenzi thesis). Far quicker.

$$F(S_k) = F(S_0) + \sum_j \left[F(S_j^{\pm}) - F(S_0)\right]|R_{jk}|$$



Speed of convergence of prediction for H cross section.

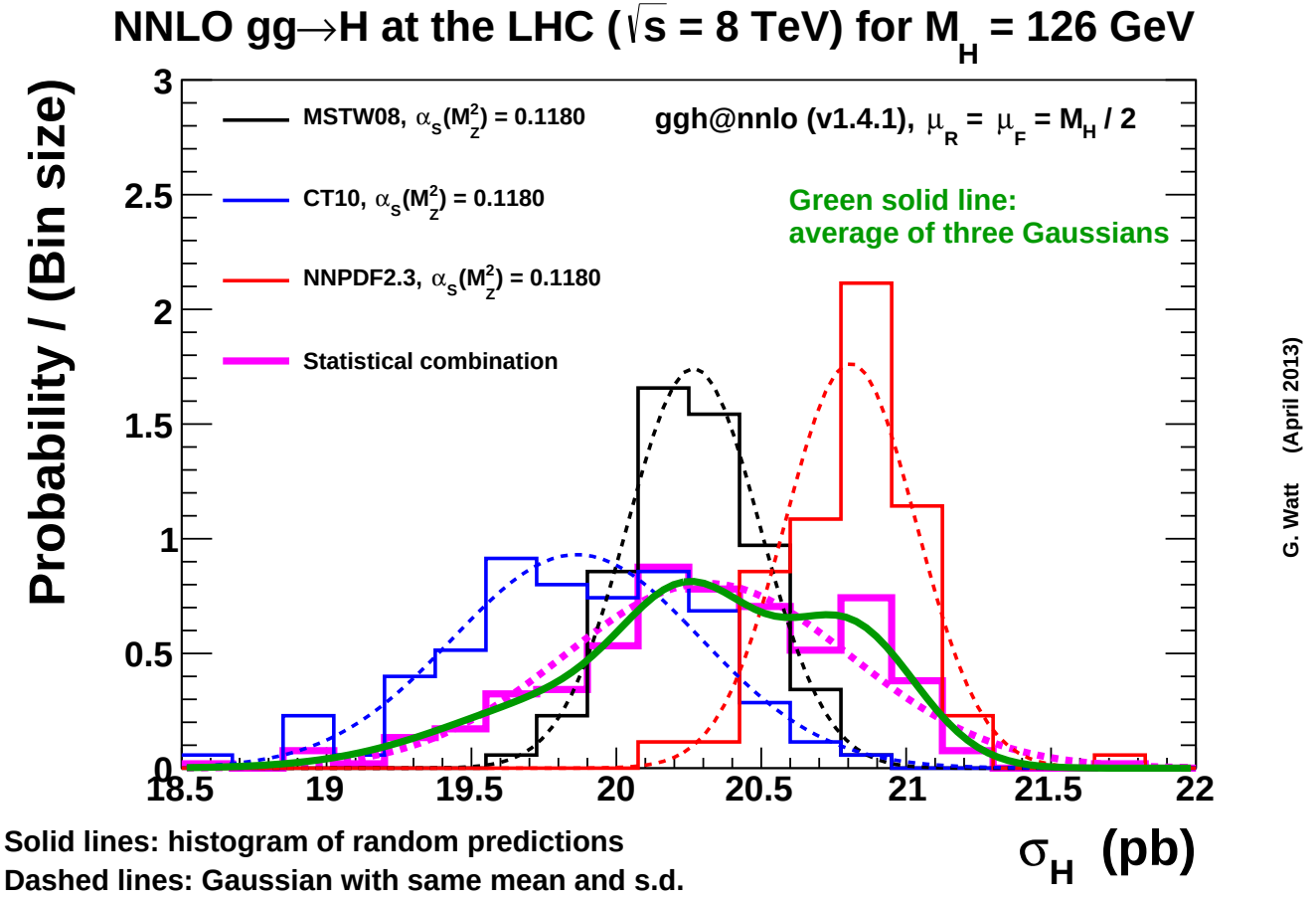




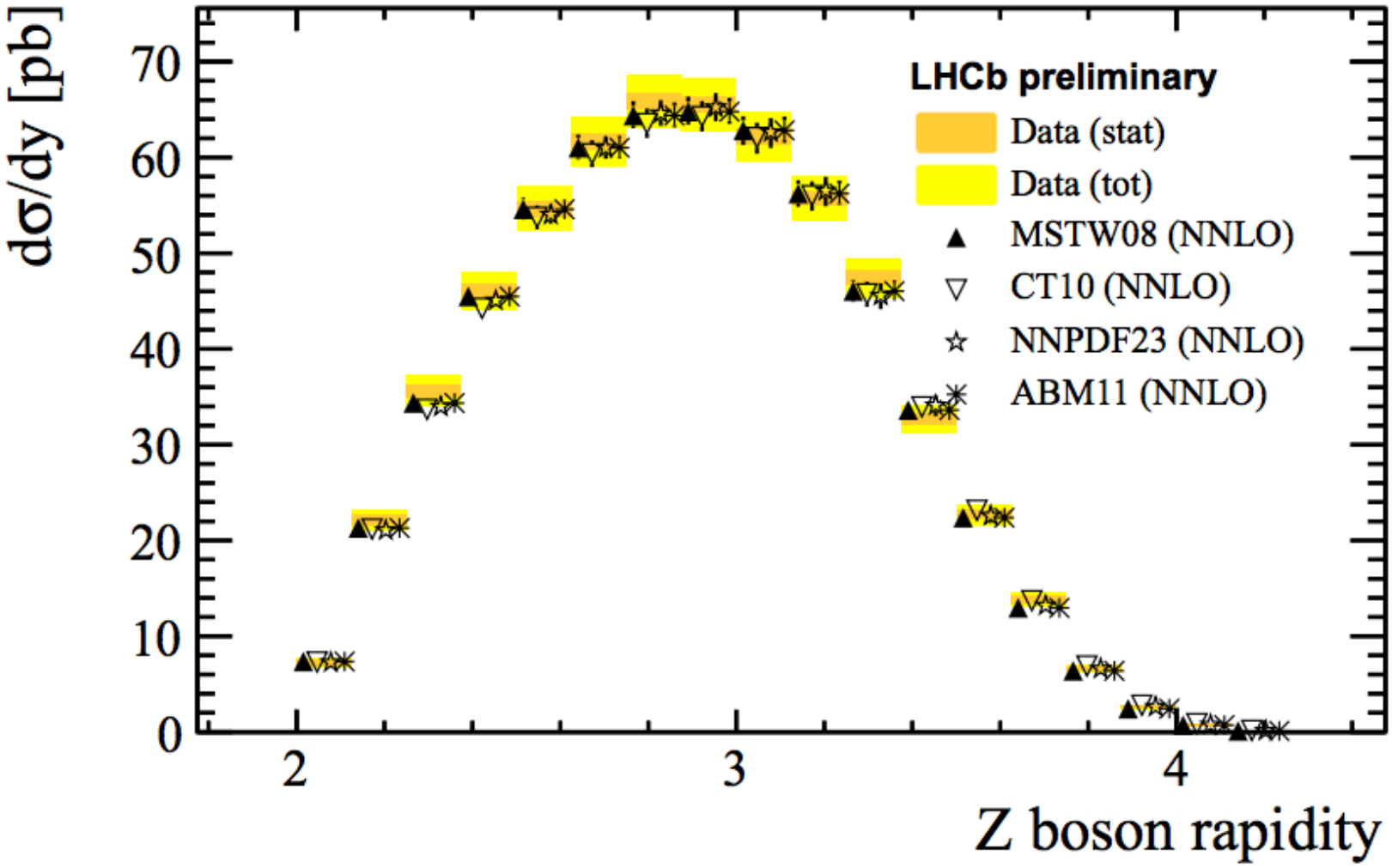
Left, add a new random set to existing ones sequentially. Right, increasing numbers of independent random sets.

Very good with 40 sets. Excellent with 100 sets.

Can combine different PDF sets, e.g. comparison to PDF4LHC prescription.



Slightly smaller uncertainty and shifted central value if disagreement between individual predictions. (Plot by G. Watt).



Higher luminosity LHCb $Z \to \mu^+\mu^-$ data.

Repeat MSTW2008 fits with modified K-factors for NNLO jets, i.e. multiply standard correction by 0 or 2 and use constant $K = 1.15$.

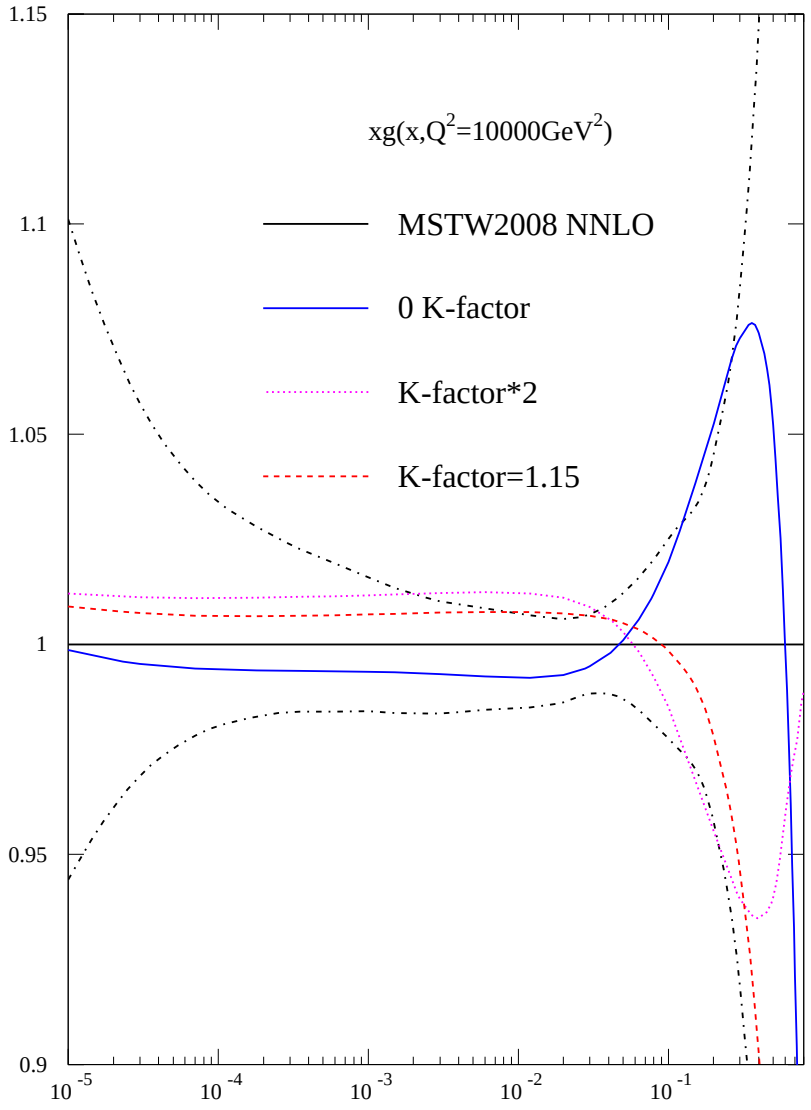
Extreme variations.

Changes in gluon relatively small. Larger K-factor slightly worse χ^2. Zero K-factor slightly better χ^2, $K = 1.15$ almost no change.

$K = 0 \ \alpha_S(M_Z^2) = 0.1181$

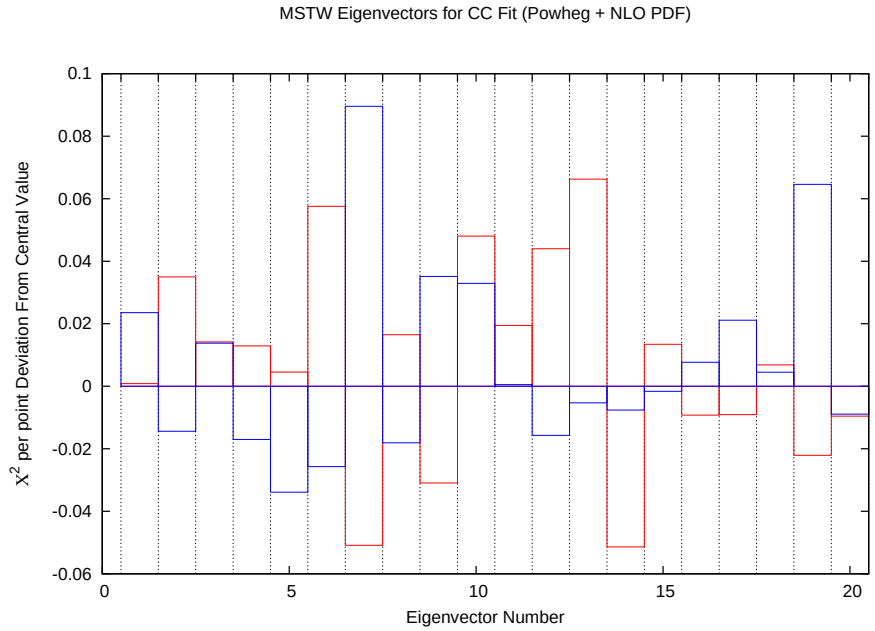
$K * 2 \ \alpha_S(M_Z^2) = 0.1159$

$K = 1.15 \ \alpha_S(M_Z^2) = 0.1167$

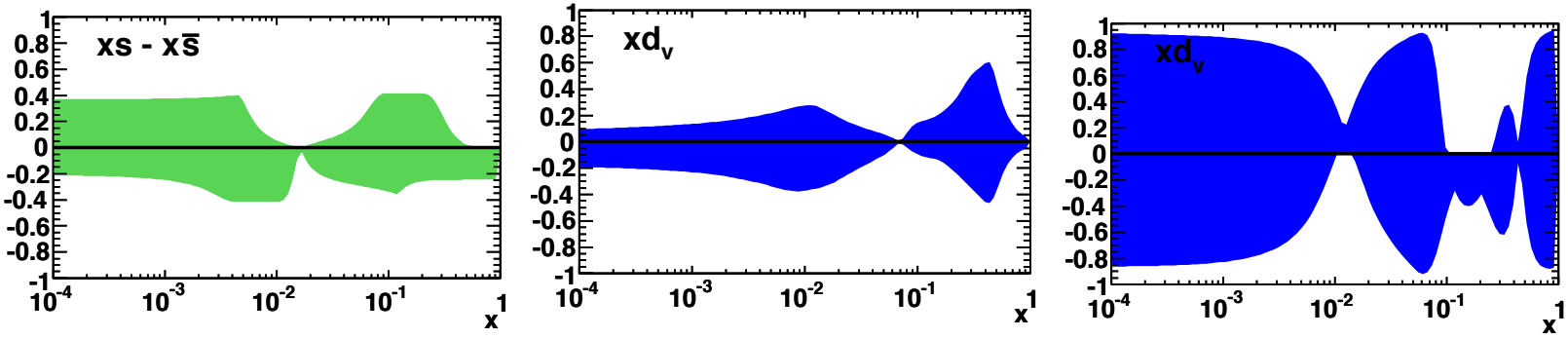


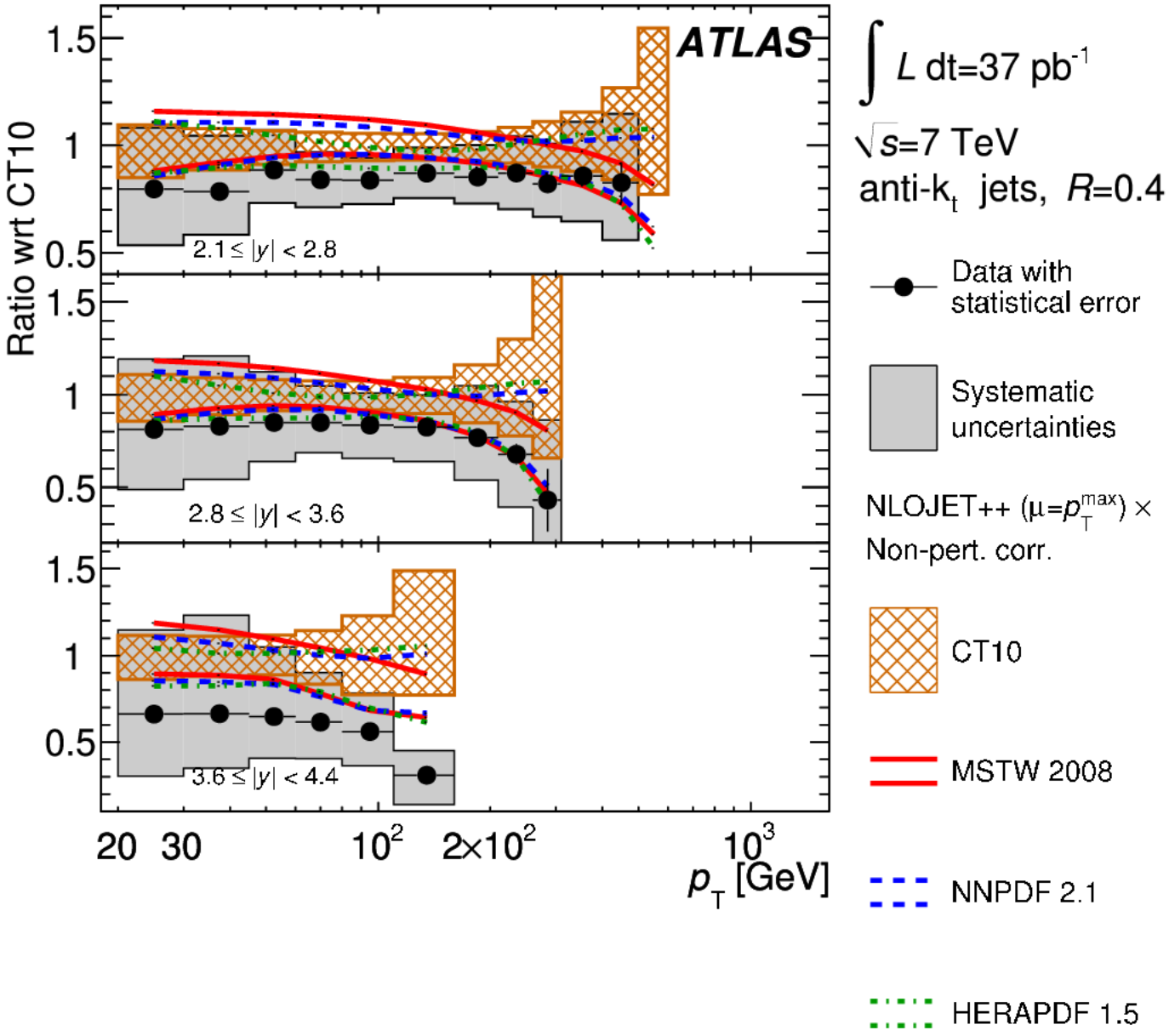
DIS jets - B. Watt

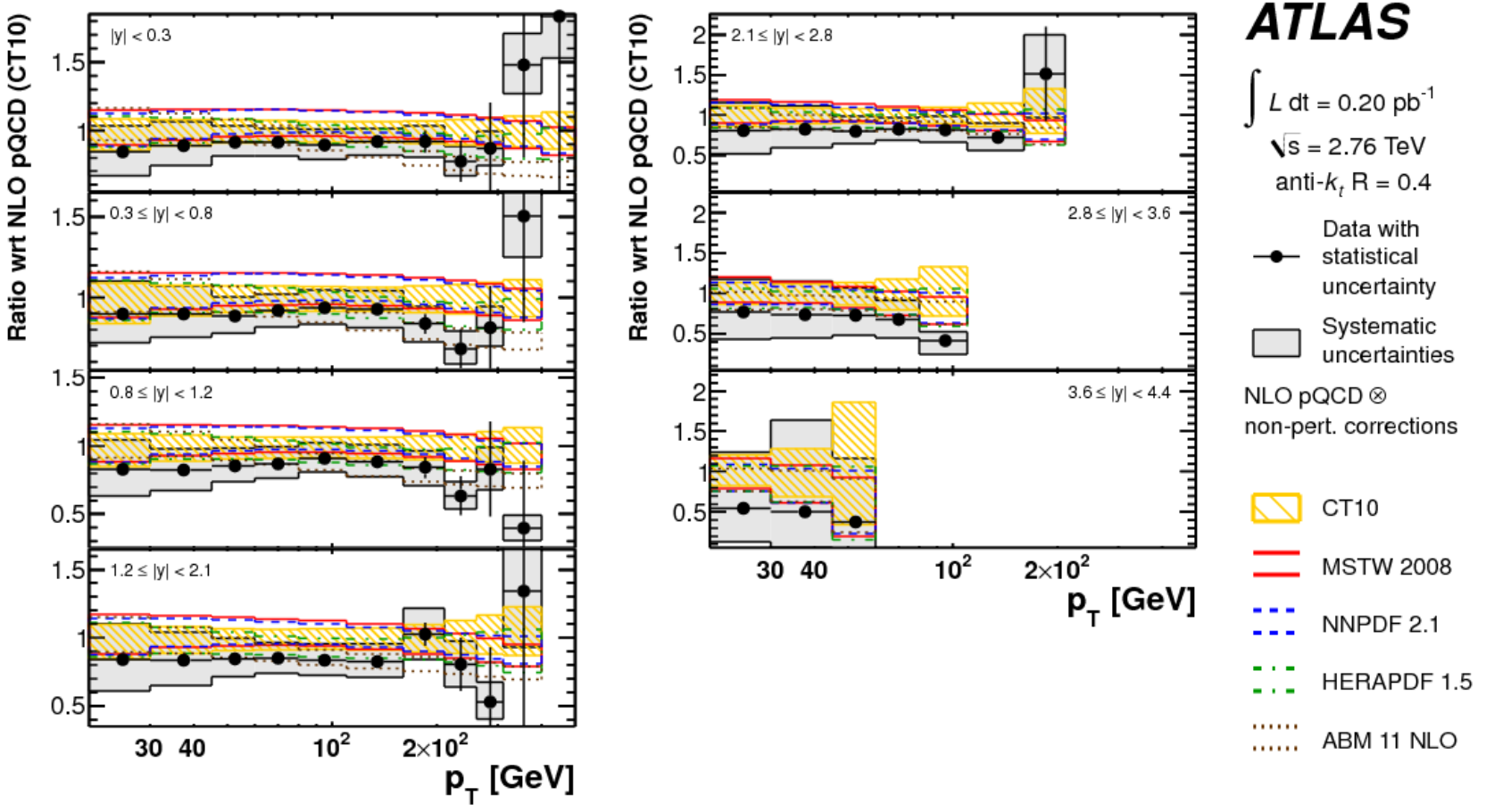
Interesting observation in eigenvector sensitivity to charged current ZEUS jet data usingPOWHEG.

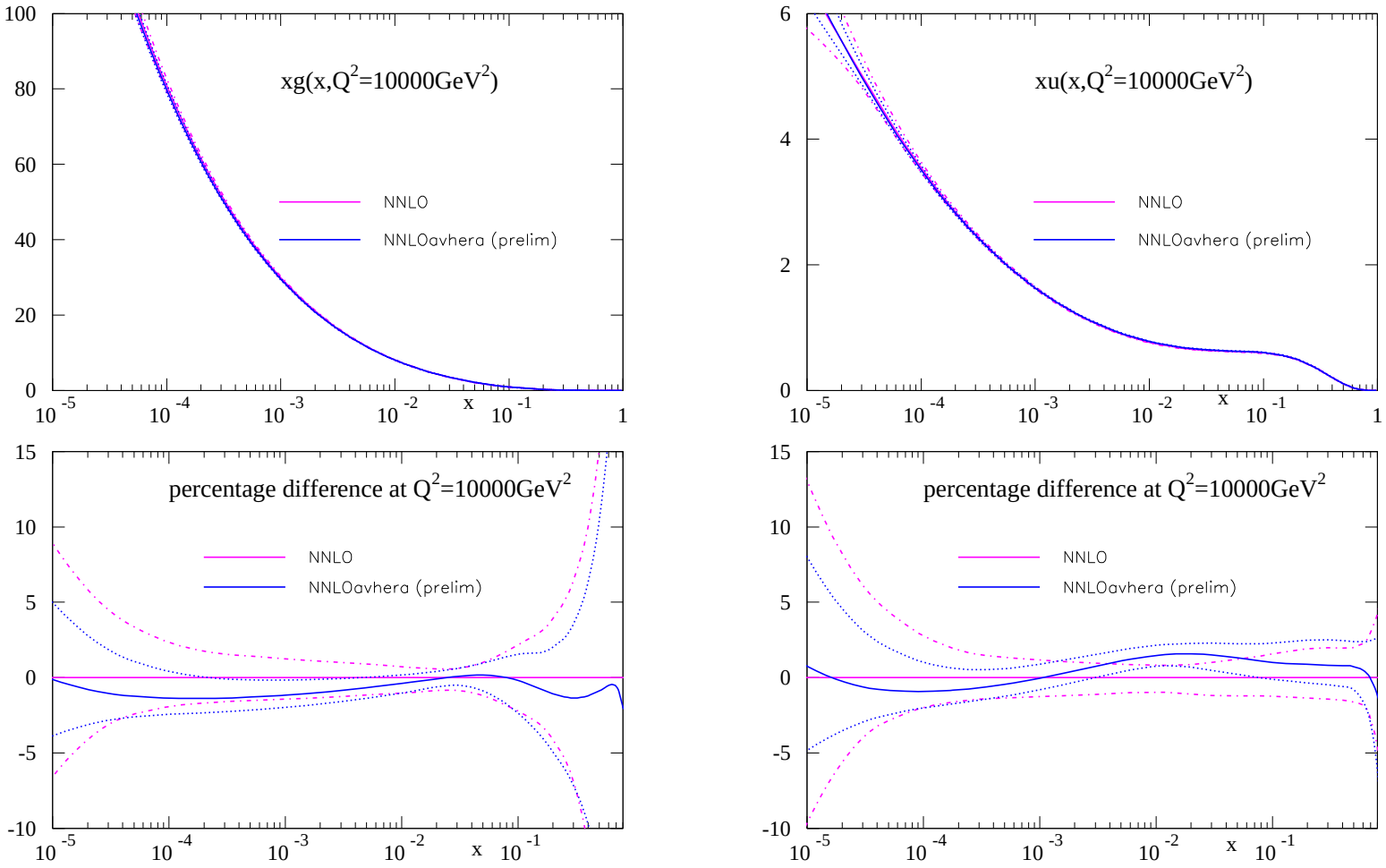


Some sensitivity to $s - \bar{s}$.



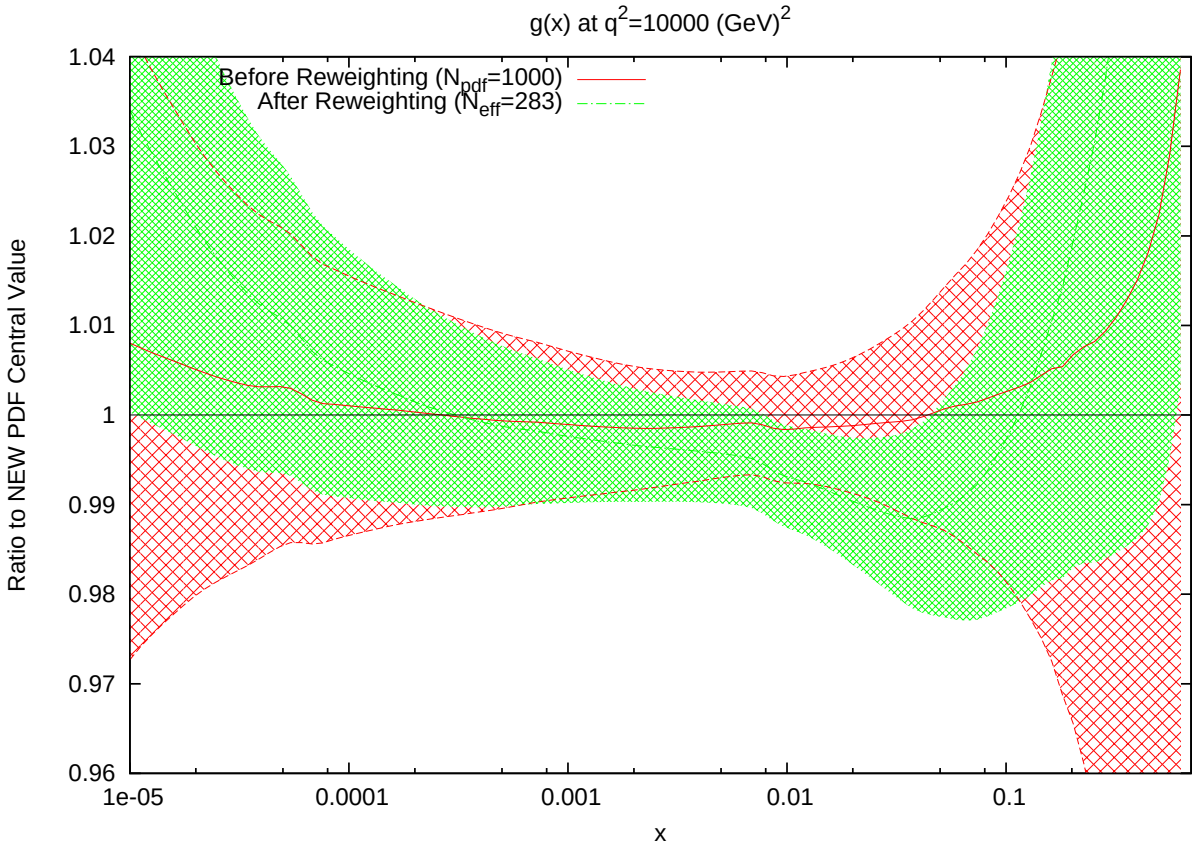

ATLAS
∫ L dt=37 pb-1
√s=7 TeV
anti-kt jets, R=0.4
Data with statistical error
Systematic uncertainties
NLOJET++ (μ=pTmax) × Non-pert. corr.
CT10
MSTW 2008
NNPDF 2.1
HERAPDF 1.5
Ratio wrt CT10
2.1 ≤ |y| < 2.8
2.8 ≤ |y| < 3.6
3.6 ≤ |y| < 4.4
pT [GeV]


ATLAS
∫ L dt = 0.20 pb-1
√s = 2.76 TeV
anti-kt R = 0.4
Data with statistical uncertainty
Systematic uncertainties
NLO pQCD ⊗ non-pert. corrections
CT10
MSTW 2008
NNPDF 2.1
HERAPDF 1.5
ABM 11 NLO
Ratio wrt NLO pQCD (CT10)
|y| < 0.3
0.3 ≤ |y| < 0.8
0.8 ≤ |y| < 1.2
1.2 ≤ |y| < 2.1
2.1 ≤ |y| < 2.8
2.8 ≤ |y| < 3.6
3.6 ≤ |y| < 4.4
pT [GeV]



Change in MSTW2008 NNLO PDFs when fitting HERA combined data.

Dijets



Using reweighting exercise for CMS dijets results in a rather modified shape of gluon.

Not as high rapidity as other sets – dependence on renormalisation/factorisation scales not so severe.

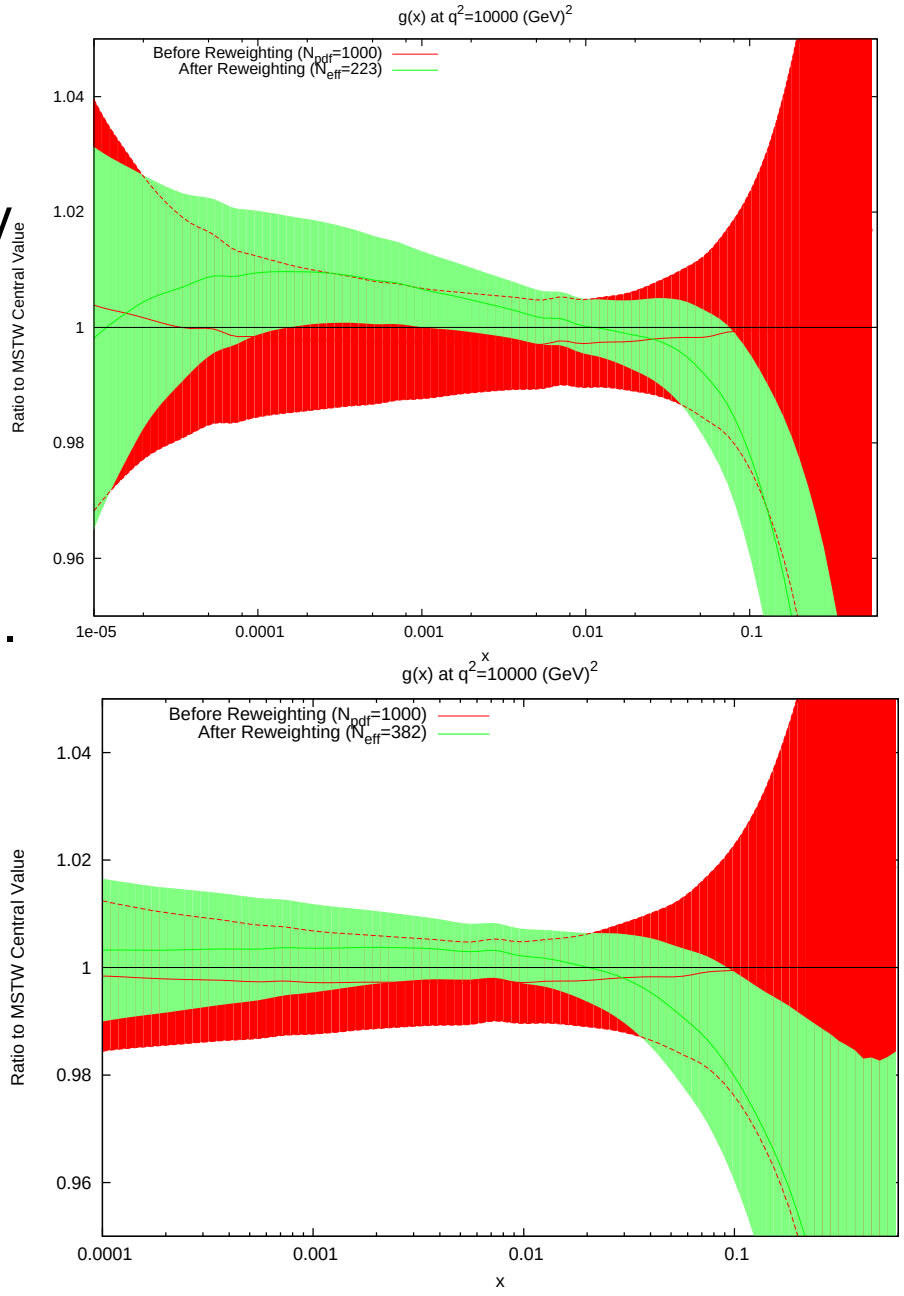
Reflection of different shape of higher order corrections?

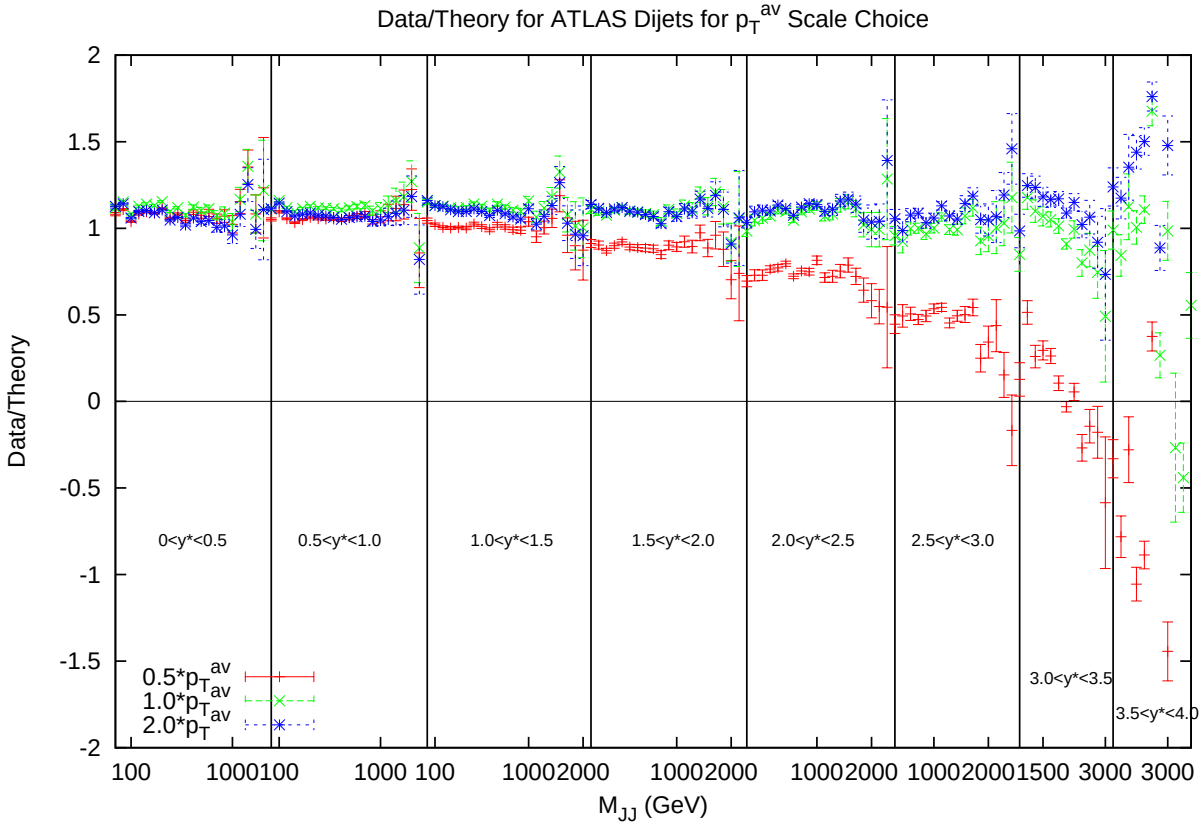
Different conclusions for fits to D0 and ATLAS dijet data, though they are not not necessarily incompatible.

Similar to changes required by LHC inclusive jet data.

Different range of rapidity spanned. Need to use scale other than p_T to get good fits. $\mu = 2p_T$ best for ATLAS and $\mu = M_{JJ}$ best for D0.

For ATLAS rapidity dependent scale choices give results more like that for CMS, but with a worse fit and lower value of N_{eff}.





At high rapidity calculations unstable for scales equal to relativity low multiples of p_T.