Updates of PDFs within the MSTW framework

Robert Thorne

May 16th, 2014

University College London

IPPP Research Associate

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

and thanks to Ben Watt, Graeme Watt and James Stirling

Updates in Fits with the MSTW Framework.

I will present results on continuing updates in PDFs within the MSTW framework due to some theory improvements and a variety of new data sets, including most of the up-to-date LHC data. A new set of PDFs is very close to being finalised, with no significant changes expected to the PDFs shown here.

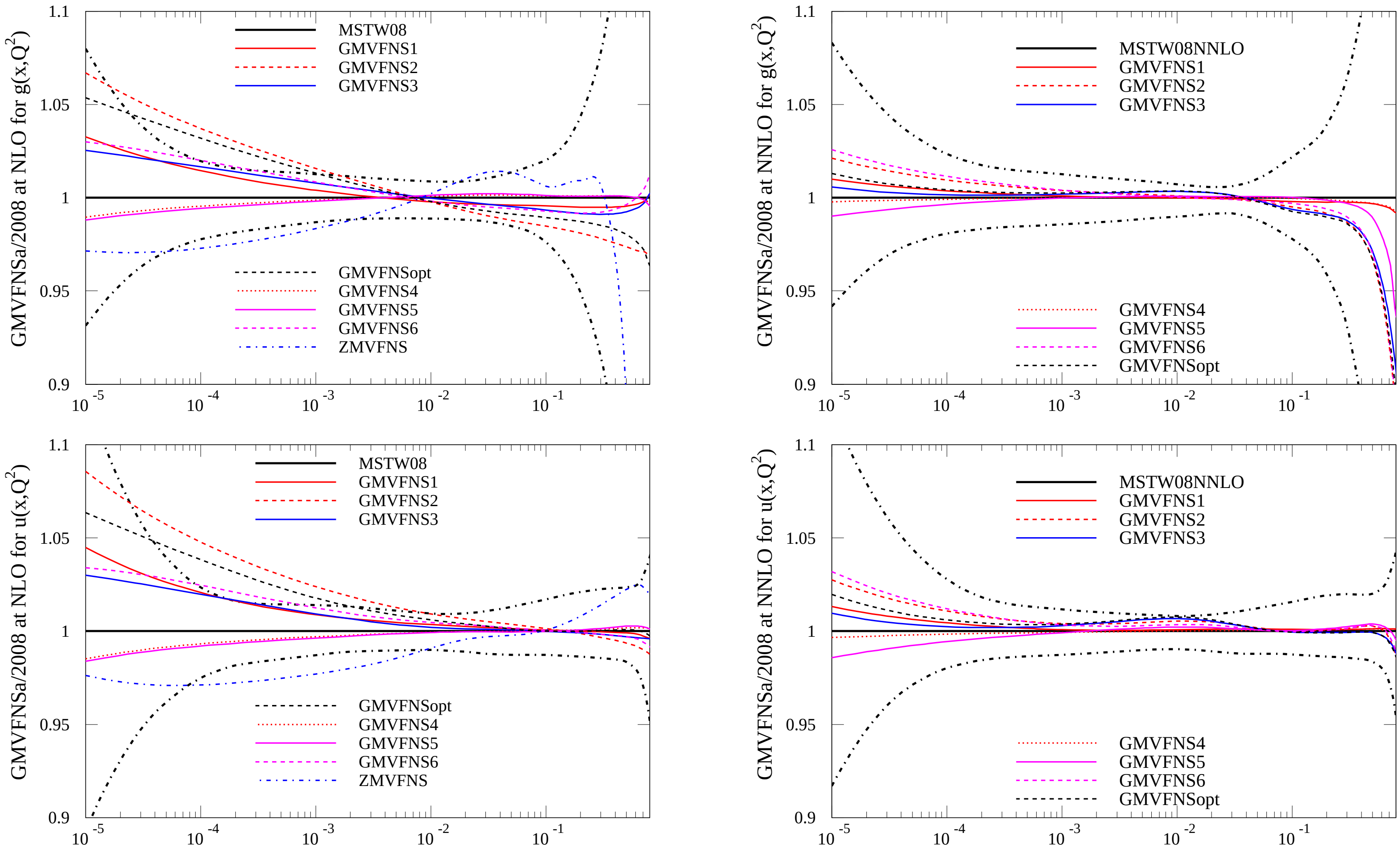
Changes in theoretical treatment or procedures.

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.



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 $\alpha_S(M_z^2)$.

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

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

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

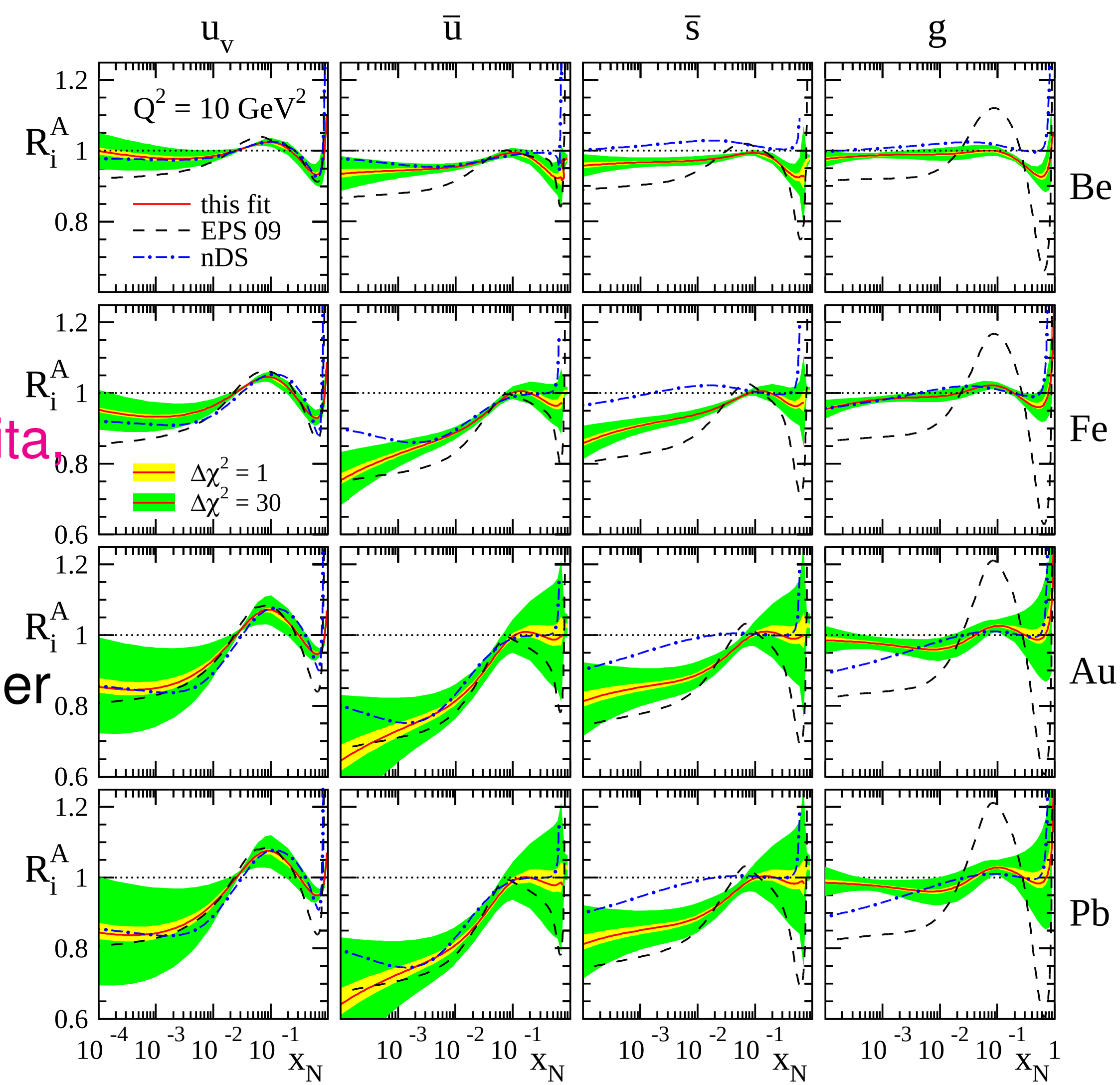
Strange branching ratio. Now avoid those determined by fits to dimuon data relying on PDF input. Also apply error which feeds into PDFs. Use $B_\mu = 0.092 \pm 10\%$ from hep-ex/9708014. Fits prefer $B_\mu = 0.087 - 0.091 \pm 15\%$, with NNLO at lower end.

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 smaller 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 $\sim$ 60-65 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.

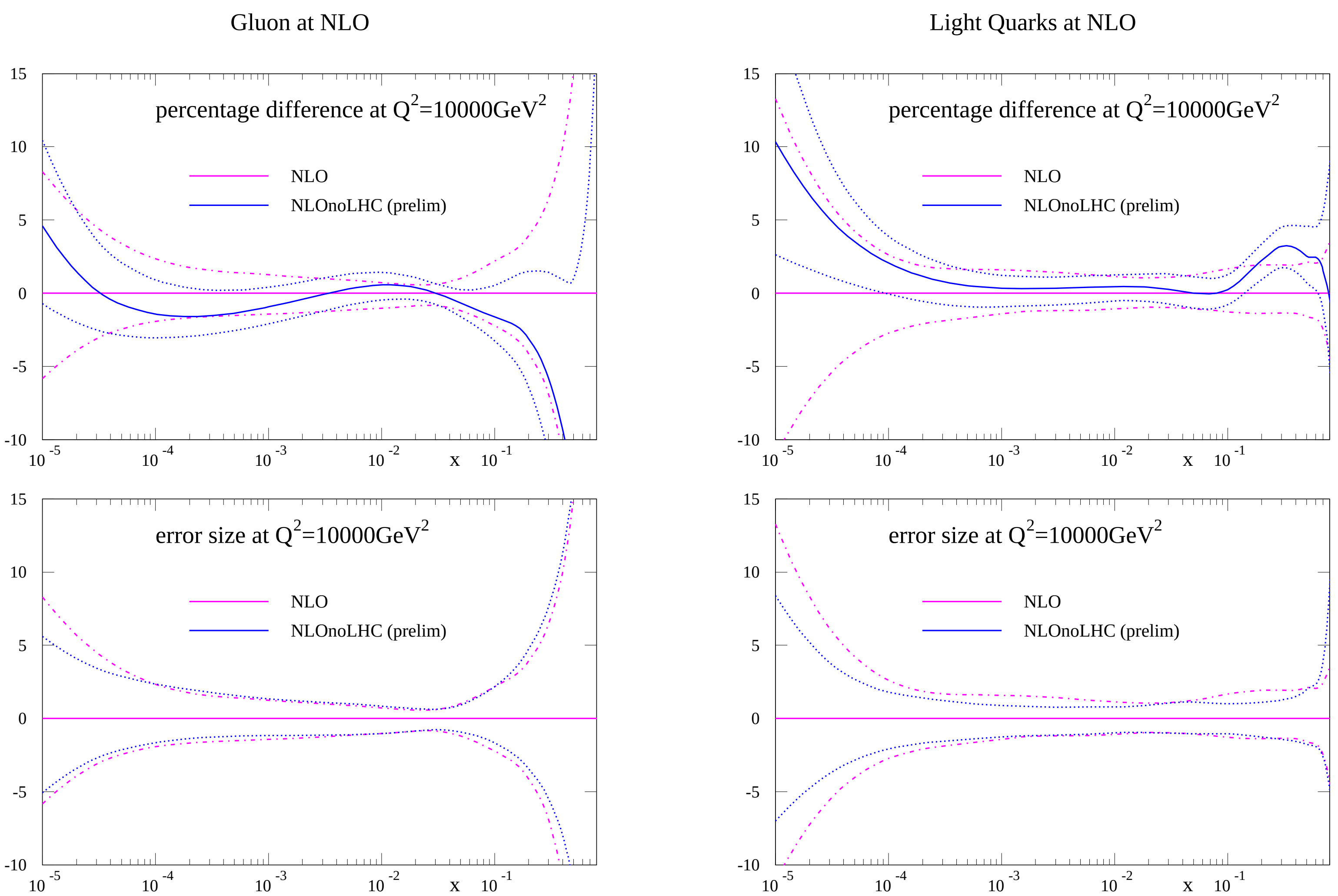
No inclusion of separate run II H1 and ZEUS data yet. Wait for Run II combination.

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}.

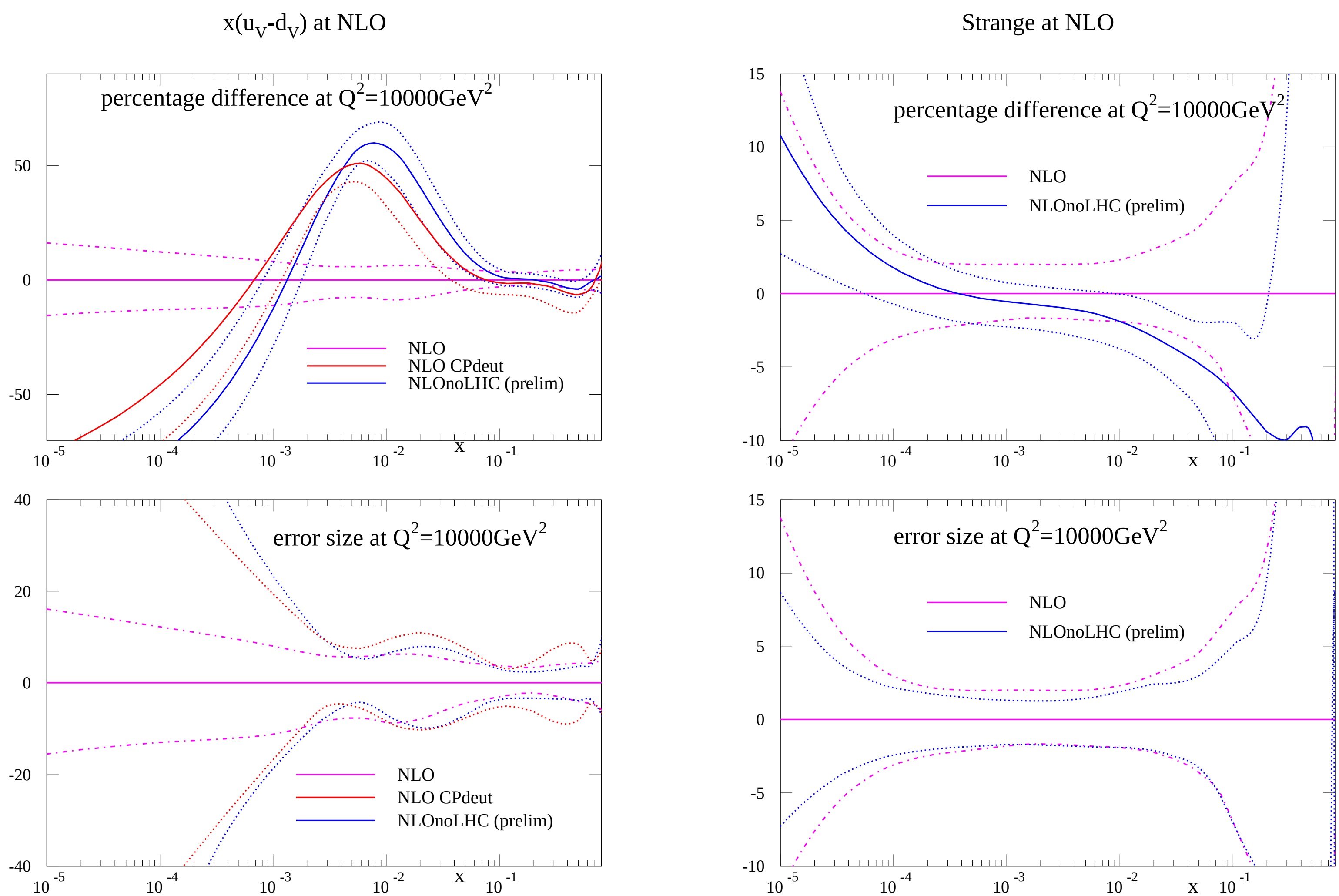
Include final numbers for CDF Z-rapidity data – final numbers changed after MSTW2008 fit. (Also include very small photon contribution in theory.) Very little change.

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.



Change in NLO PDFs from non-LHC data updates. Increase in d at high x.

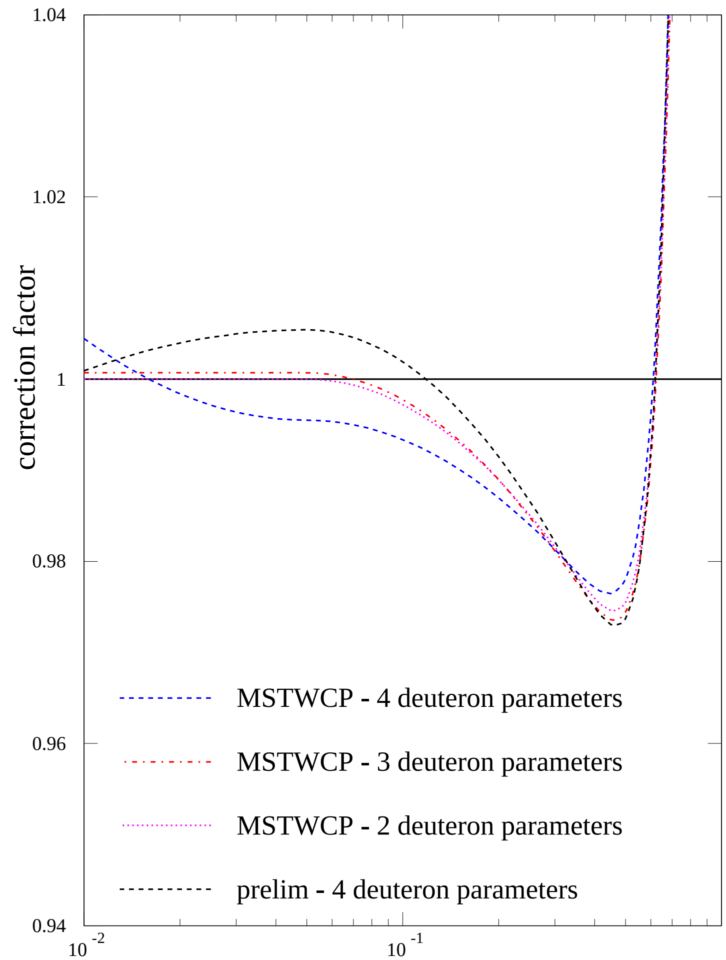


Change in NLO PDFs from non-LHC updates. Change in branching ratio for dimuon data not incorporated.

Result for fitted deuteron correction.

Previously big improvement in fit for MSTWCPdeut, 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.



LHC data on W,Z, $t\bar{t}$

Now with Harland-Lang and Motylinski using APPLGrid – MCFM and DYNNLO 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.

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 $7\mathrm{TeV}$ and one point at $8\mathrm{TeV}$. Use $m_t = 172.5\ \mathrm{GeV}$ (value used in Tevatron combination) with an error of $1\ \mathrm{GeV}$, with χ^2 penalty applied. Predictions and fit good, with NLO preferring masses slightly below $m_t = 172.5\ \mathrm{GeV}$ and NNLO masses slightly above.

Intend to include CMS double differential Drell Yan data extending to low mass. Problem with NNLO calculation (bug discovered), but seems clear NNLO fits better than NLO at lowest mass $\sim 20 - 40\mathrm{GeV}$.

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 $\chi^2 = 112/114$ and CMS $\chi^2 = 186/133$ before included directly – as good as any PDF.

Simultaneous fit of CMS data together with ATLAS 7 TeV + 2.76 TeV leads to bigger improvement for CMS, but a tiny amount for ATLAS.

The two experiments seem extremely compatible.

At NLO final extracted $\alpha_S(M_Z^2) = 0.1193$.

LHC jets not included at NNLO. At NNLO final extracted $\alpha_S(M_Z^2) = 0.1162$

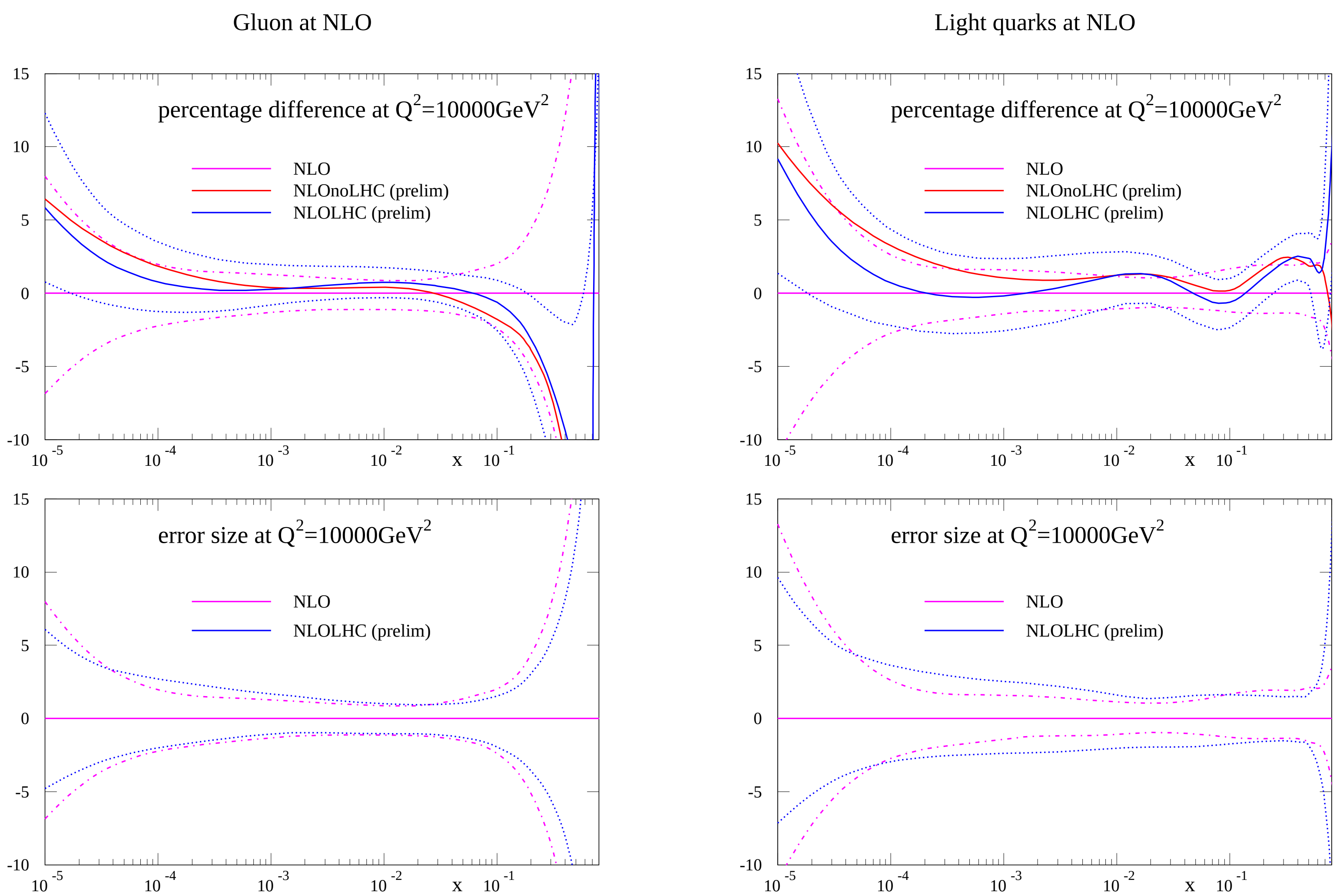
Fit quality for LHC data at NLO

data set	N_{pts}	CPdeut	no LHC	prelim
ATLAS jets (2.76TeV+7TeV)	114	111.9	111.7	110.6
CMS jets (7TeV)	133	179.6	186.3	173.0
ATLAS W^+, W^-, Z	30	46.9	44.2	40.5
CMS W asymm $p_T > 35\mathrm{GeV}$	11	8.5	12.9	6.9
CMS asymm $p_T > 25\mathrm{GeV}, 30\mathrm{GeV}$	24	8.6	16.7	7.1
LHCb $Z \to e^+e^-$	9	13.2	12.7	12.3
LHCb W asymm $p_T > 20\mathrm{GeV}$	10	12.5	13.8	12.2
CMS $Z \to e^+e^-$	35	20.8	19.9	22.7
ATLAS High mass DY	13	20.3	20.3	21.3
TeV, ATLAS, CMS $\sigma_{t\bar{t}}$	13	8.0	9.7	7.2

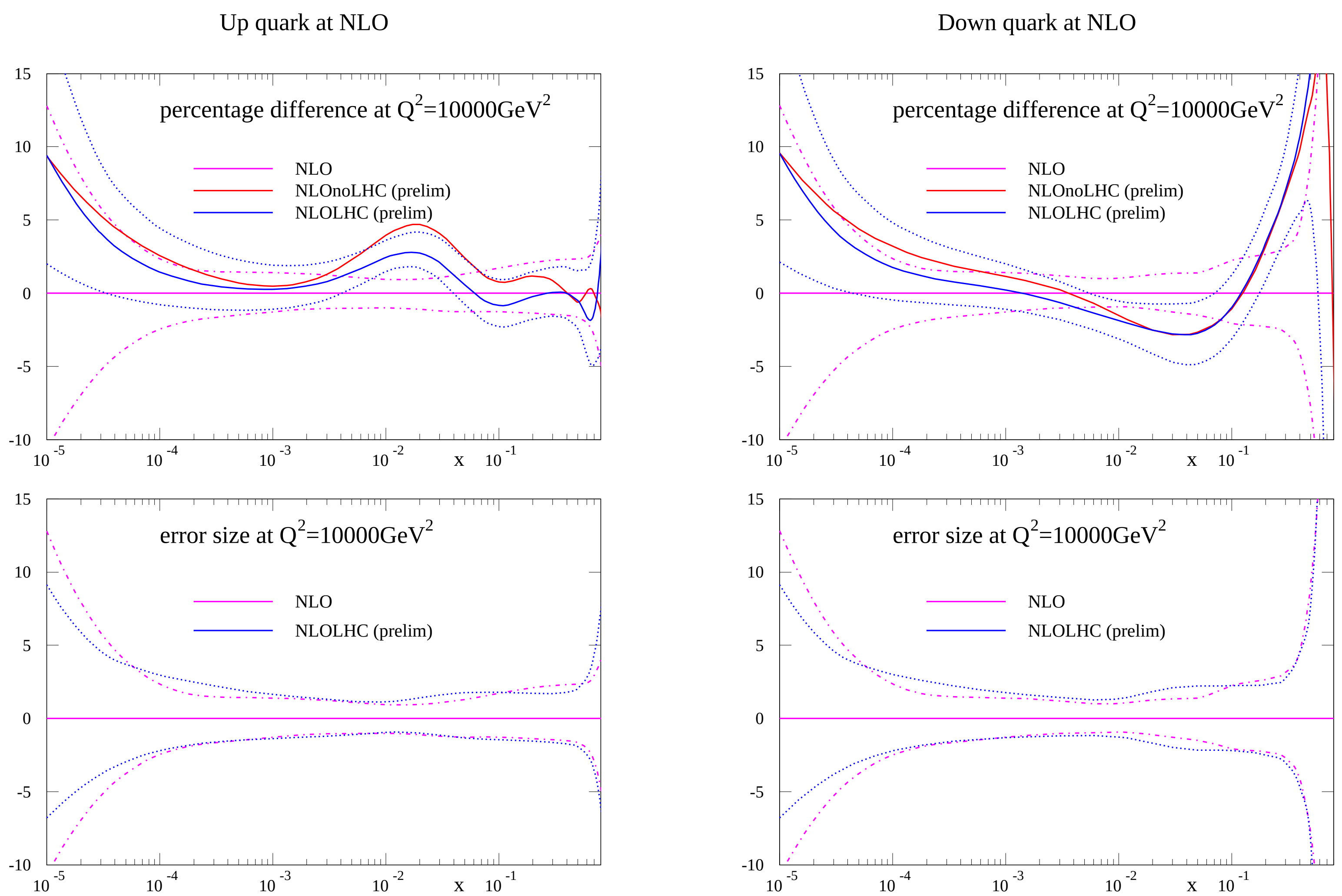
ATLAS W, Z data constrain a gluon eigenvector direction, as do $\sigma_{t\bar{t}}$ and CMS $Z \to e^+e^-$.

CMS W asymm. data constrains some flavour decomposition.

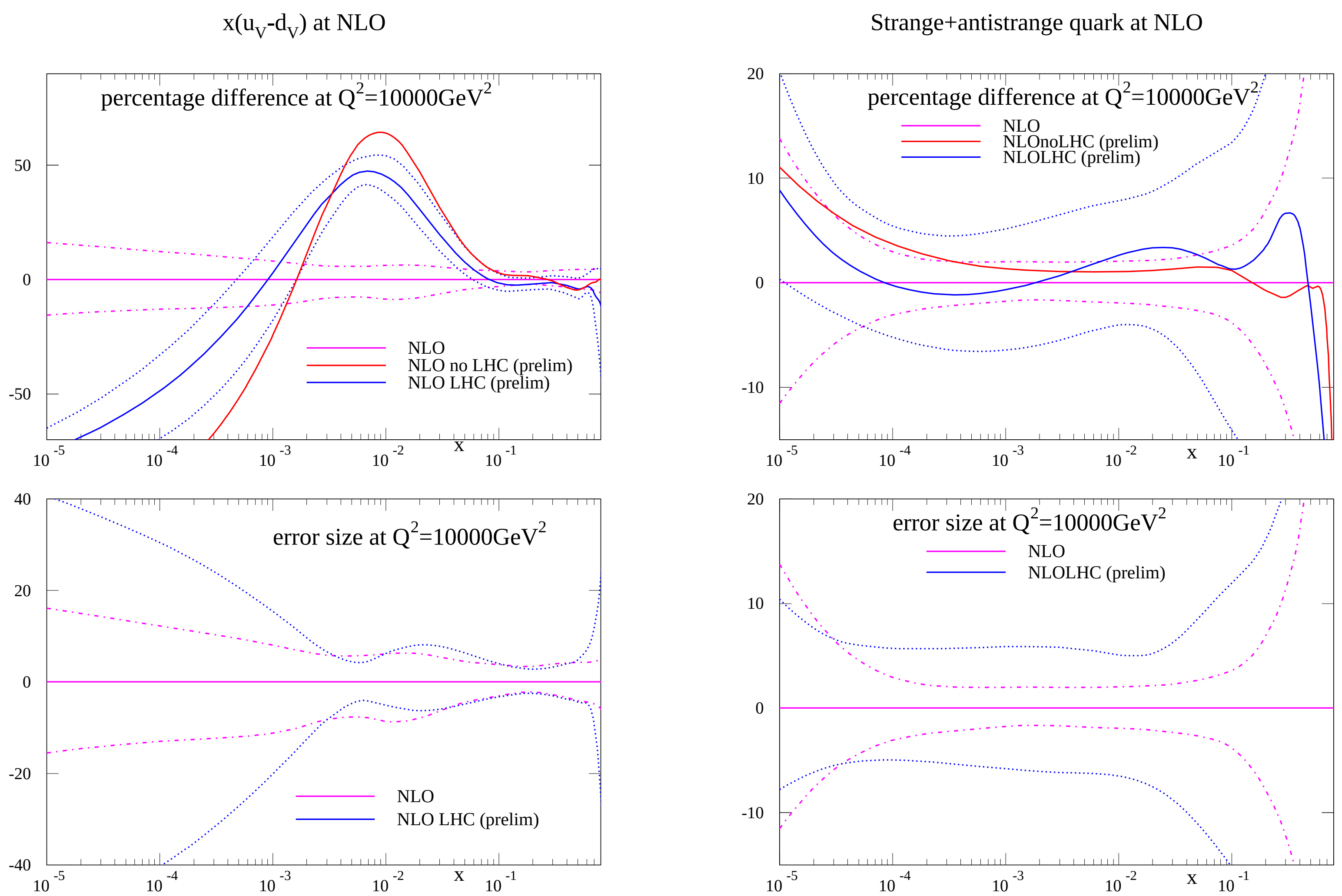
Also fit CMS double differential low and high mass Drell Yan data. No real change in PDFs. Fit very poor at NLO (effectively LO) in lowest mass bins, even when weighted.



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



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

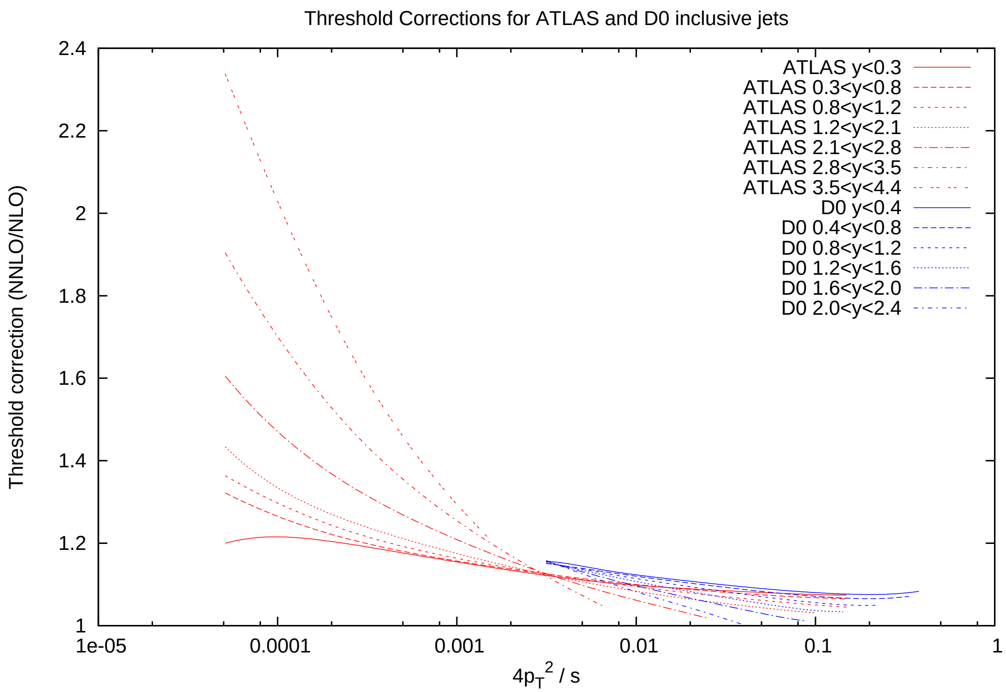


Change in NLO PDFs from all, including LHC data updates. Much expanded $s + \bar{s}$ uncertainty.

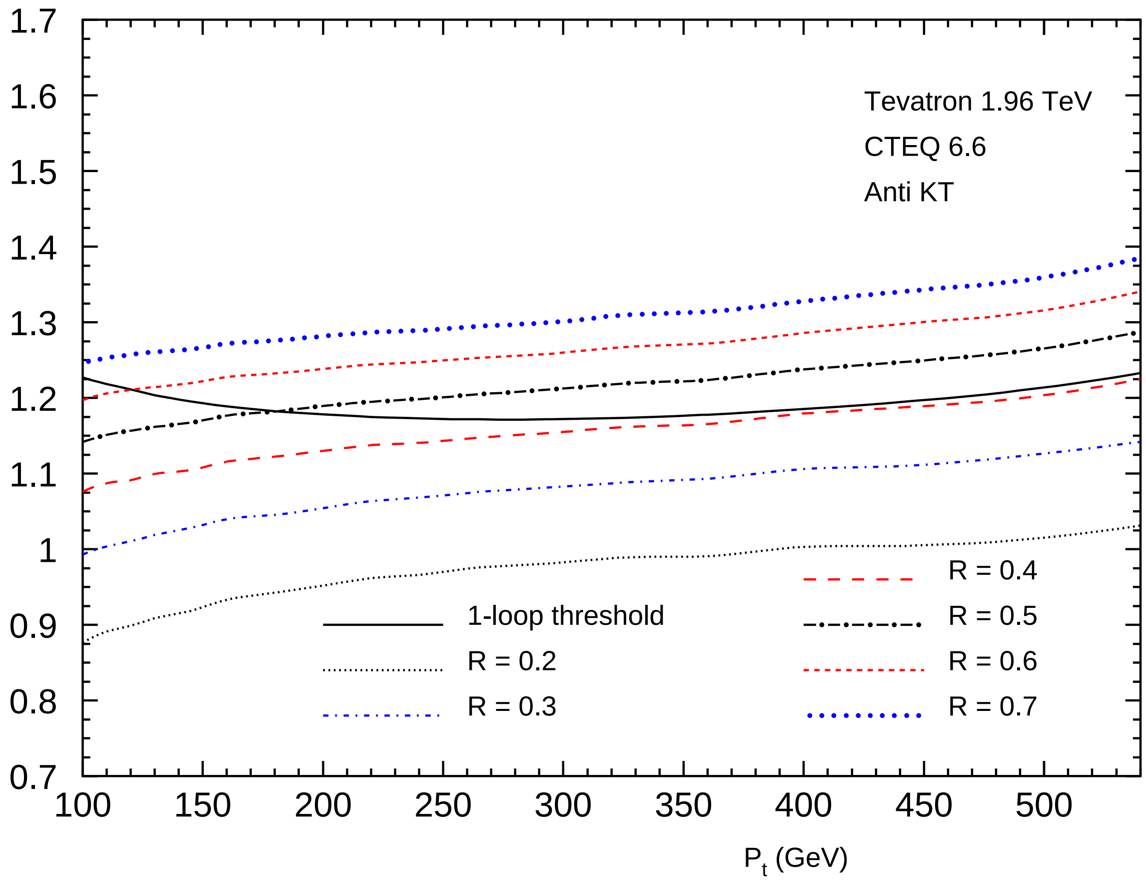
LHC jet data at NNLO?

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.

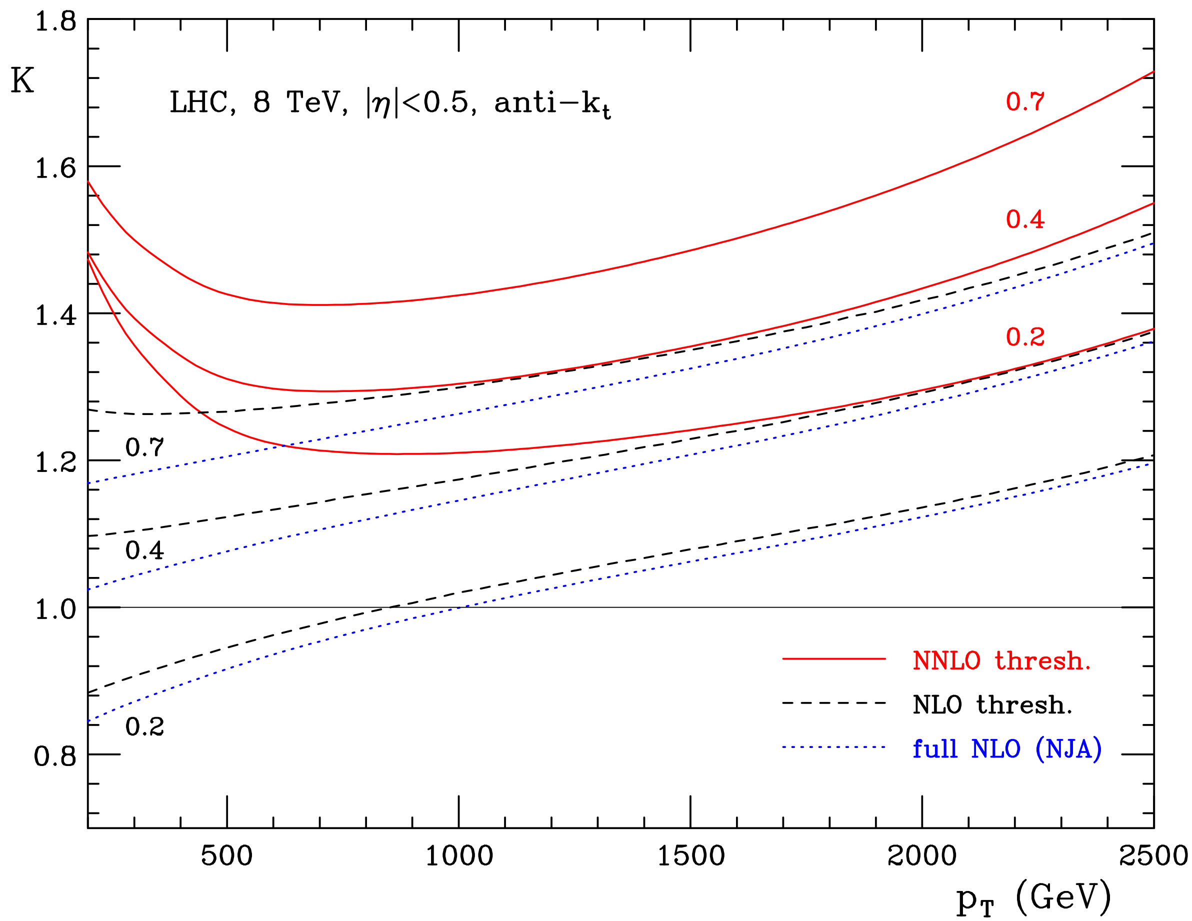


Enormous project of full NNLO calculation (Gehrmann-de-Ridder, Gehrmann, Glover and Pires) nearing completion. Some indications of full form of the correction.



Recent repeat of threshold calculations Kumar, Moch (arXiv:1309.5311) and comparison to exact NLO results for different jet radius R.

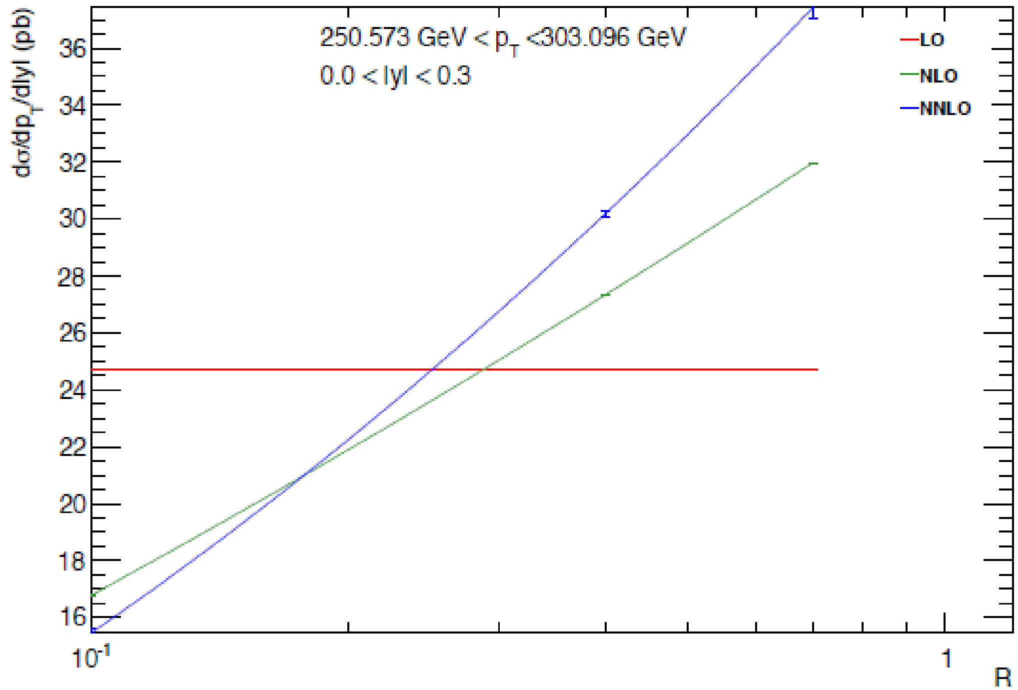
Big variation with R at NLO and threshold calculation which has no R dependence matches best with $R \sim 0.3 - 0.4$.



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

NNLO corrections to pp → 2jets

Single-jet inclusive: jet size dependence in anti-k_T algorithm

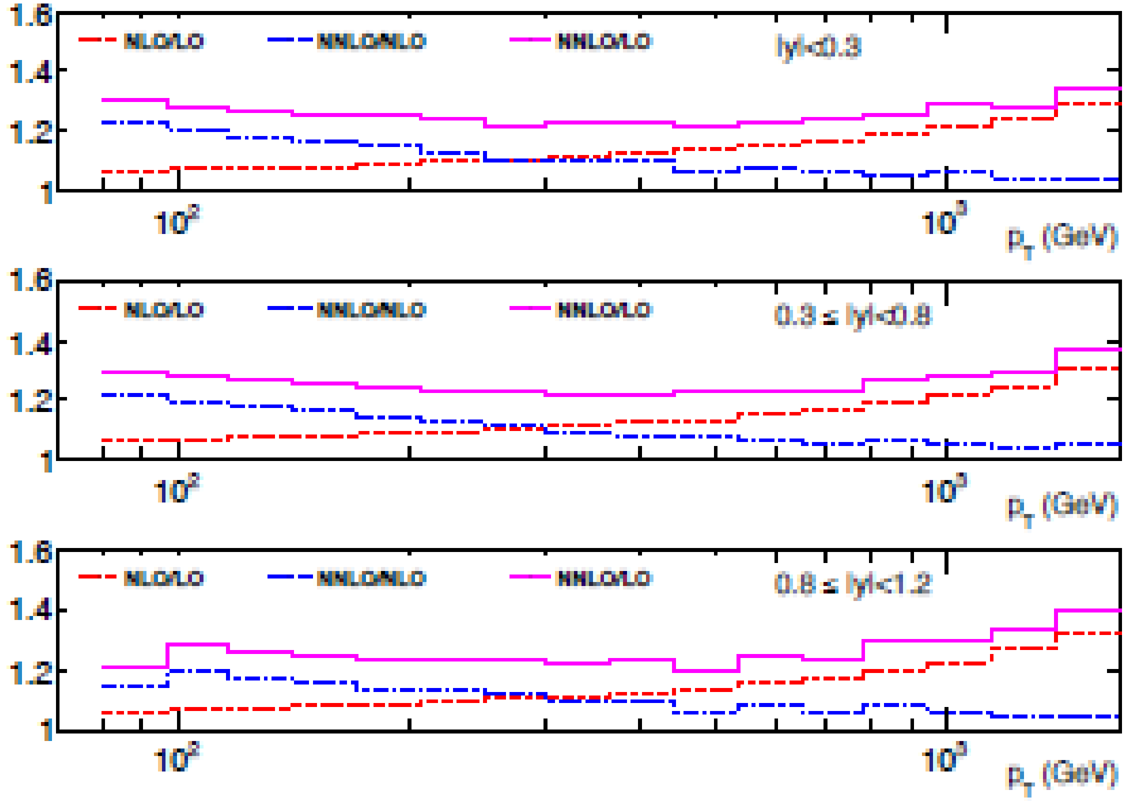


22 Aude Gehrmann-De Ridder SM@LHC 2013 Freiburg

Compatible with existing full results. Here at NLO ratio for $R = 0.7$ to $R = 0.4$ is 1.25 but shrinks to 1.06 at NNLO.

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.

NNLO PDF updates

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.

Have also tried repeating MSTW2008 fits with extreme modified K-factors for NNLO jets, i.e. multiply standard correction by 0 or 2 and use constant $K = 1.15$. Even at extremes changes almost entirely within one sigma. Similar to scale changes at NLO.

However, omit LHC data. Lowest p_T not stable in threshold corrections, and large uncertainty at highest rapidity.

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.

Fit quality for LHC data at NNLO. Jet data not fitted but quality checked using "smaller" K-factor.

data set	N_{pts}	CPdeut	no LHC	prelim
ATLAS jets (2.76TeV+7TeV)	114	(113.7)	(128.9)	(112.8)
CMS jets (7TeV)	133	(184.9)	(181.3)	(181.3)
ATLAS W^+, W^-, Z	30	76.8	57.1	40.1
CMS W asymm $p_T > 35\mathrm{GeV}$	11	21.4	18.1	9.0
CMS asymm $p_T > 25\mathrm{GeV}, 30\mathrm{GeV}$	24	18.5	16.6	10.8
LHCb $Z \rightarrow e^+e^-$	9	20.9	20.5	20.0
LHCb W asymm $p_T > 20\mathrm{GeV}$	10	24.1	21.5	13.5
CMS $Z \rightarrow e^+e^-$	35	31.0	28.8	19.2
ATLAS High mass DY	13	17.9	16.5	17.8
TeV, ATLAS, CMS $\sigma_{t\bar{t}}$	13	8.0	11.2	6.5

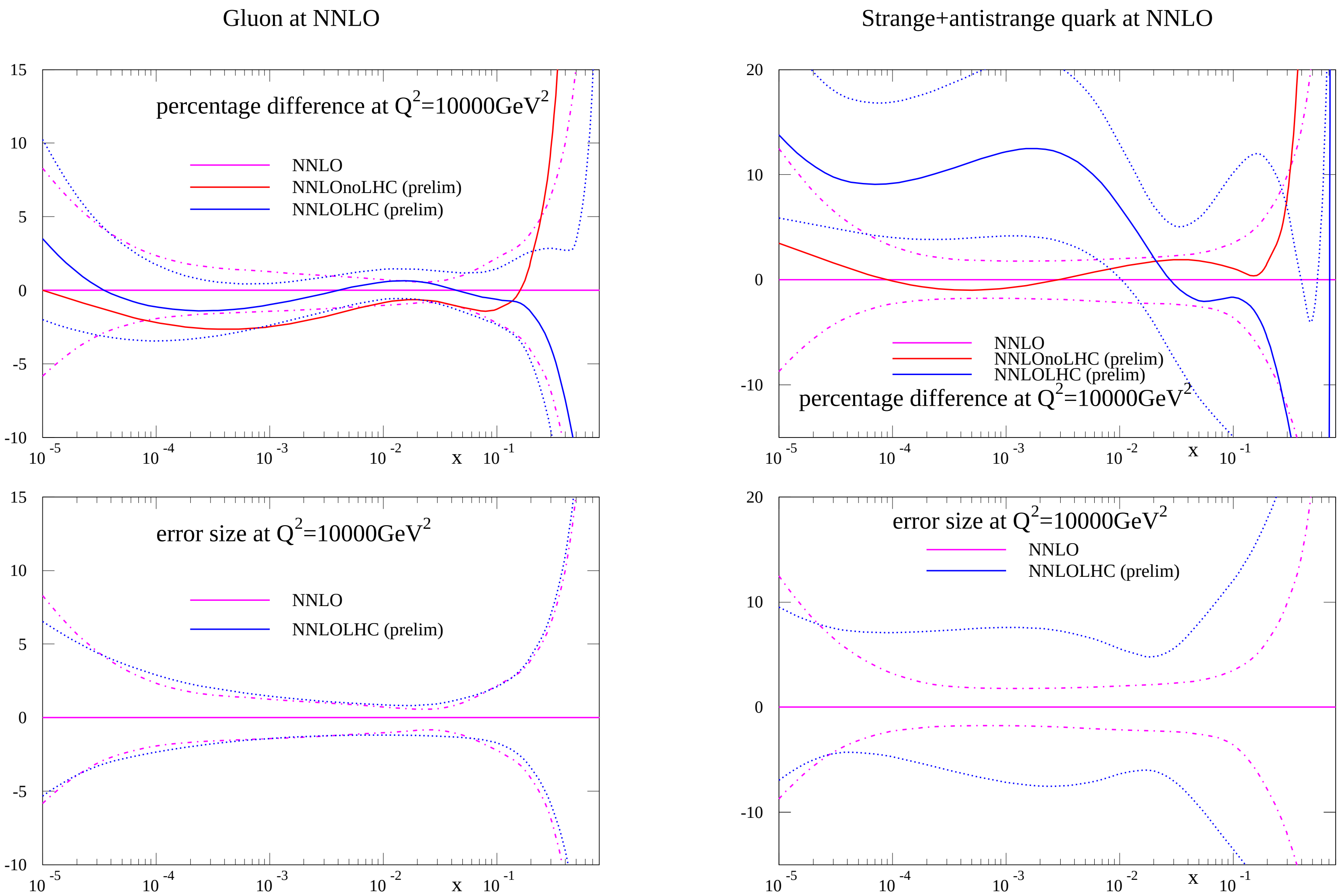
Large improvement in ATLAS W, Z data, mainly from strange quark, and in CMS $Z \rightarrow e^+e^-$ data and some extent to CMS W asymm. and LHCb W^+, W^- data.

CMS $Z \rightarrow e^+e^-$ data constrains a gluon eigenvector and CMS W asymm. data some flavour decomposition.

Fit quality for LHC data at NNLO. Jet data not fitted but quality checked using “larger” K-factor

data set	N_{pts}	CPdeut	no LHC	prelim
ATLAS jets (2.76TeV+7TeV)	114	(134.1)	(144.7)	(129.2)
CMS jets (7TeV)	133	(191.3)	(187.6)	(189.6)
ATLAS W^+, W^-, Z	30	76.8	57.1	40.1
CMS W asymm $p_T > 35\mathrm{GeV}$	11	21.4	18.1	9.0
CMS asymm $p_T > 25\mathrm{GeV}, 30\mathrm{GeV}$	24	18.5	16.6	10.8
LHCb $Z \to e^+e^-$	9	20.9	20.5	20.0
LHCb W asymm $p_T > 20\mathrm{GeV}$	10	24.1	21.5	13.5
CMS $Z \to e^+e^-$	35	31.0	28.8	19.2
ATLAS High mass DY	13	17.9	16.5	17.8
TeV, ATLAS, CMS $\sigma_{t\bar{t}}$	13	8.0	11.2	6.5

ATLAS jet data deteriorates more than CMS. Difficult to guess relative size of K-factor at two different energies.



Change in NNLO PDFs from all, including LHC data updates. Gluon uncertainty at high-x slightly greater than at NLO.

26 pairs of eigenvectors. 5 for sea and gluon, 4 for each valence quark, 3 for $\bar{u} - \bar{d}$ and $s + \bar{s}$ and 2 for $s - \bar{s}$. Really one too many.

Eigenvectors constrained by a wide variety of data sets.

data type	NLO eigenvectors	NNLO eigenvectors
HERA	11	13
Fixed target NC	4	6
Fixed target CC inclusive	3	1
Fixed target CC $\mu^+\mu^-$	8	8
Fixed target DY Drell Yan	8	8
Tevatron	10	10
LHC	8	6

HERA DIS data constrains many gross features, e.g. gluon, sea normalisation.

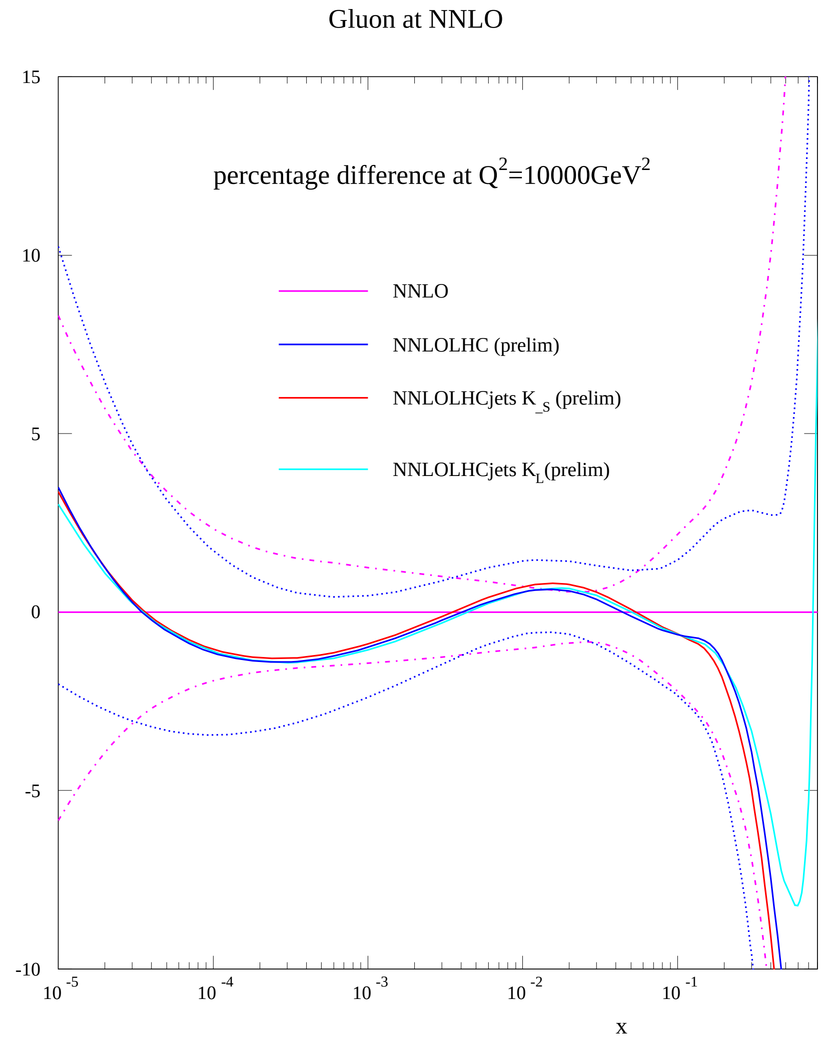
Dimuon data constrains most of the 10 strange eigenvectors.

Fixed target Drell Yan all $\bar{d} - \bar{u}$ differences, and high-x sea.

Tevatron and now LHC data constrain flavour separation and some constraint on gluon.

Ratio of $g(x,Q^2)$ for the default NNLO fit to that in MSTW2008, and also fits where jet data included with "smaller" and "larger" K-factor.

In both cases changes in gluon, $\alpha_S(M_Z^2)$ and fit to other data are extremely small.



For the "smaller" K-factor ATLAS $\chi^2 = 112.8/114 \rightarrow 112.2/114$ and CMS $\chi^2 = 181.3/133 \rightarrow 177.8/133$.

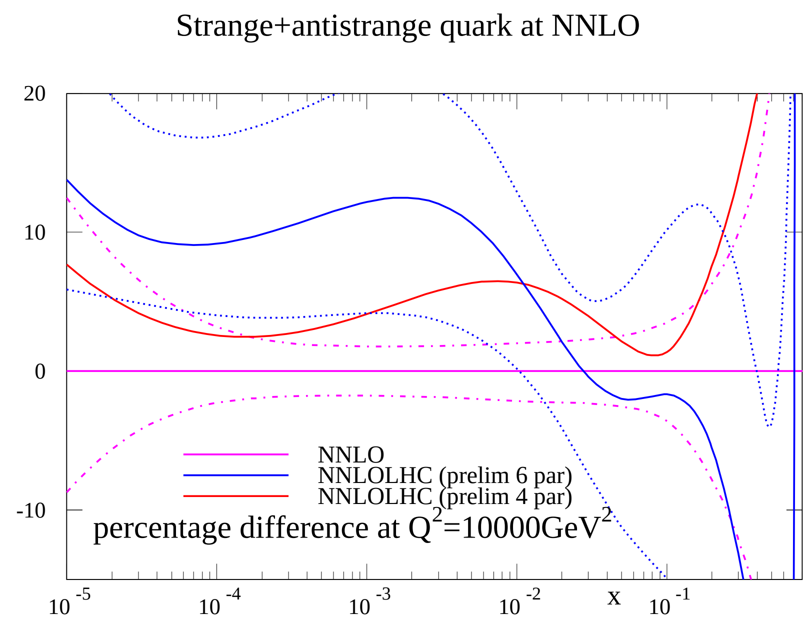
For the "larger" K-factor ATLAS $\chi^2 = 129.2/114 \rightarrow 129.4/114$ and CMS $\chi^2 = 189.6/133 \rightarrow 182.4/133$.

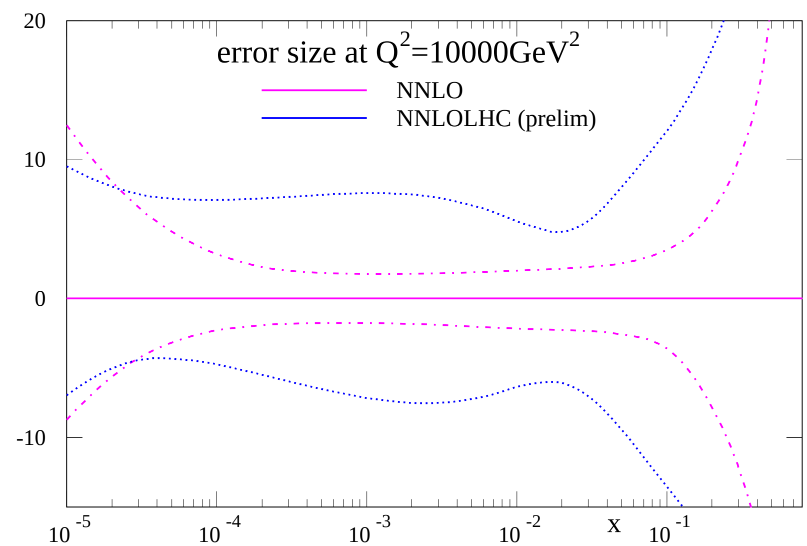
At NNLO strange quark with 4 polynomials becomes very large at very small x – larger than up and down.

Restrict fit at NLO and NNLO to 2 polynomials for strange (no change in eigenvectors).

No real change at NLO. At NNLO fit to LHC data deteriorates by a few units (dimuon data improves marginally), and $B_\mu = 0.083$.

$\alpha_S(m_Z^2) = 0.1170$. Overall little preference between 0.116–0.118 at NNLO.





Change in various cross section predictions compared to uncertainty for MSTW2008.

	no LHC NLO	no LHC NNLO	LHC NLO	LHC NNLO	unc.
W Tevatron (1.96 TeV)	$+1.0$	$+2.1$	-0.5	$+0.2$	1.8
Z Tevatron (1.96 TeV)	$+2.4$	$+2.6$	$+0.5$	$+0.1$	1.9
W^+ LHC (7 TeV)	$+2.5$	$+0.9$	$+0.3$	-1.1	2.2
W^- LHC (7 TeV)	-0.3	$+1.1$	-0.8	-1.9	2.2
Z LHC (7 TeV)	$+1.1$	$+1.1$	$+0.2$	-1.5	2.2
W^+ LHC (14 TeV)	$+3.0$	$+0.8$	$+0.7$	-0.9	2.4
W^- LHC (14 TeV)	$+0.6$	$+0.6$	-0.3	-1.6	2.4
Z LHC (14 TeV)	$+1.7$	$+0.6$	$+0.2$	-0.6	2.4
Higgs Tevatron	-3.5	$+2.8$	-3.1	-3.2	5.1
Higgs LHC (7 TeV)	-1.2	$+0.9$	-1.4	-2.1	3.3
Higgs LHC (14 TeV)	-2.0	$+0.1$	-1.2	-2.3	3.1
$t\bar{t}$ Tevatron	$+0.5$	$+4.9$	-1.6	-0.7	3.2
$t\bar{t}$ LHC (7 TeV)	-3.1	$+3.3$	-2.9	-2.5	3.9
$t\bar{t}$ LHC (14 TeV)	-2.0	$+1.7$	-2.0	-2.0	3.1

Some changes of order size of uncertainty - none dramatic.

$\alpha_S(m_Z^2)$ as a data point?

$\alpha_S(m_Z^2)$ coming out just slightly lower though NNLO varying a little around 0.116–0.118 depending on details. Still a NLO/NNLO difference. Compatible with global average. Perhaps input this as data point?

Try world average (minus DIS data) of $\alpha_S(m_Z^2) = 0.1187 \pm 0.0007$ (rather small uncertainty).

At NLO already within one sigma, essentially no change – $\alpha_S(m_Z^2) \to 0.1192$.

At NNLO best fit gives $\alpha_S(m_Z^2)$ very close to 0.118. Increase in global 4 units. All in BCDMS proton data.

Conclusions

Ongoing updates on PDFs – soon to release updated PDFs.

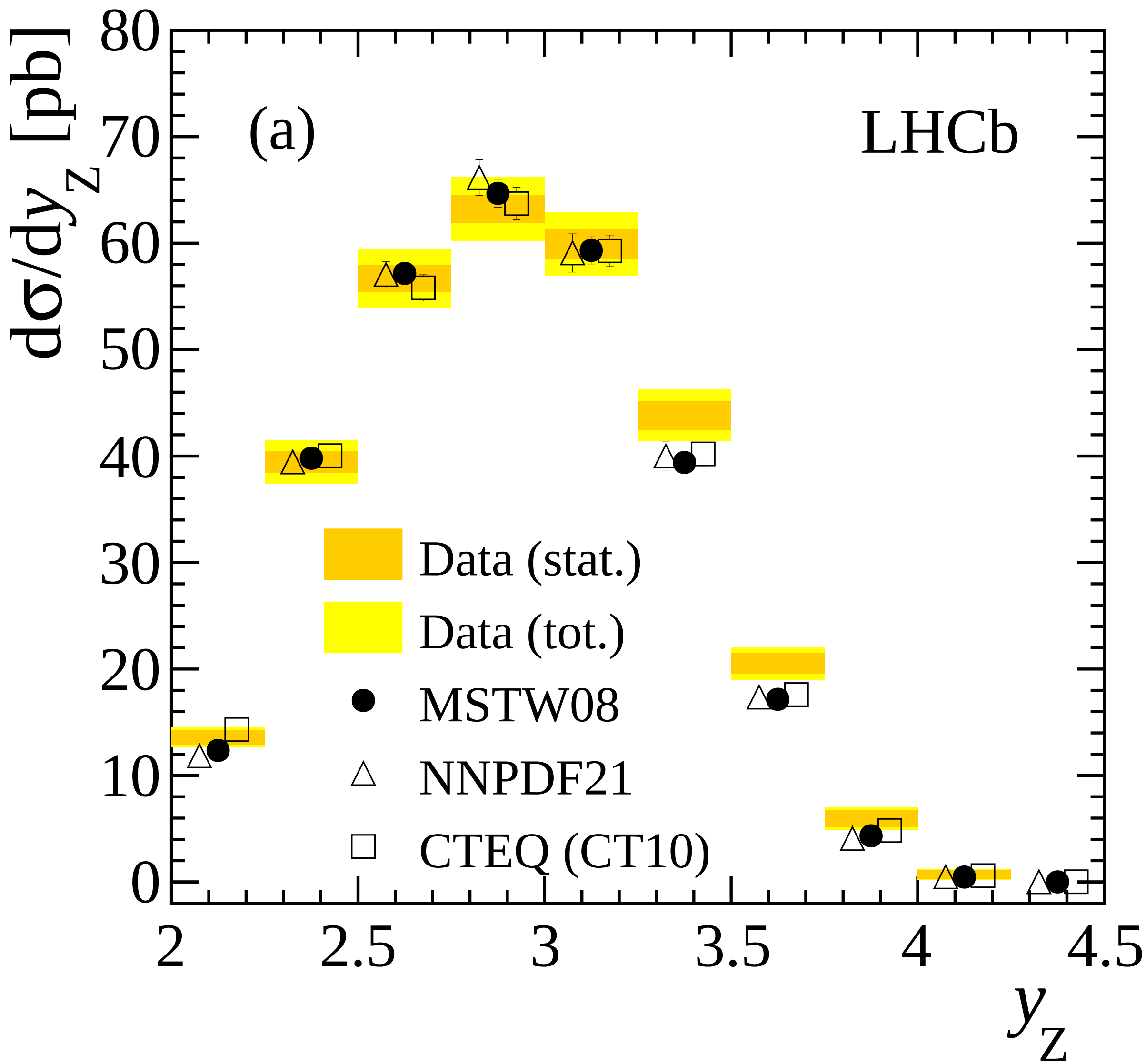
Improvement in parameterisation, heavy flavour treatments, nuclear corrections, and branching ratio for dimuon data. Inclusion of up-to-date HERA and Tevatron data.

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, though don't include these as default at NNLO. Fit good – no PDF conflicts.

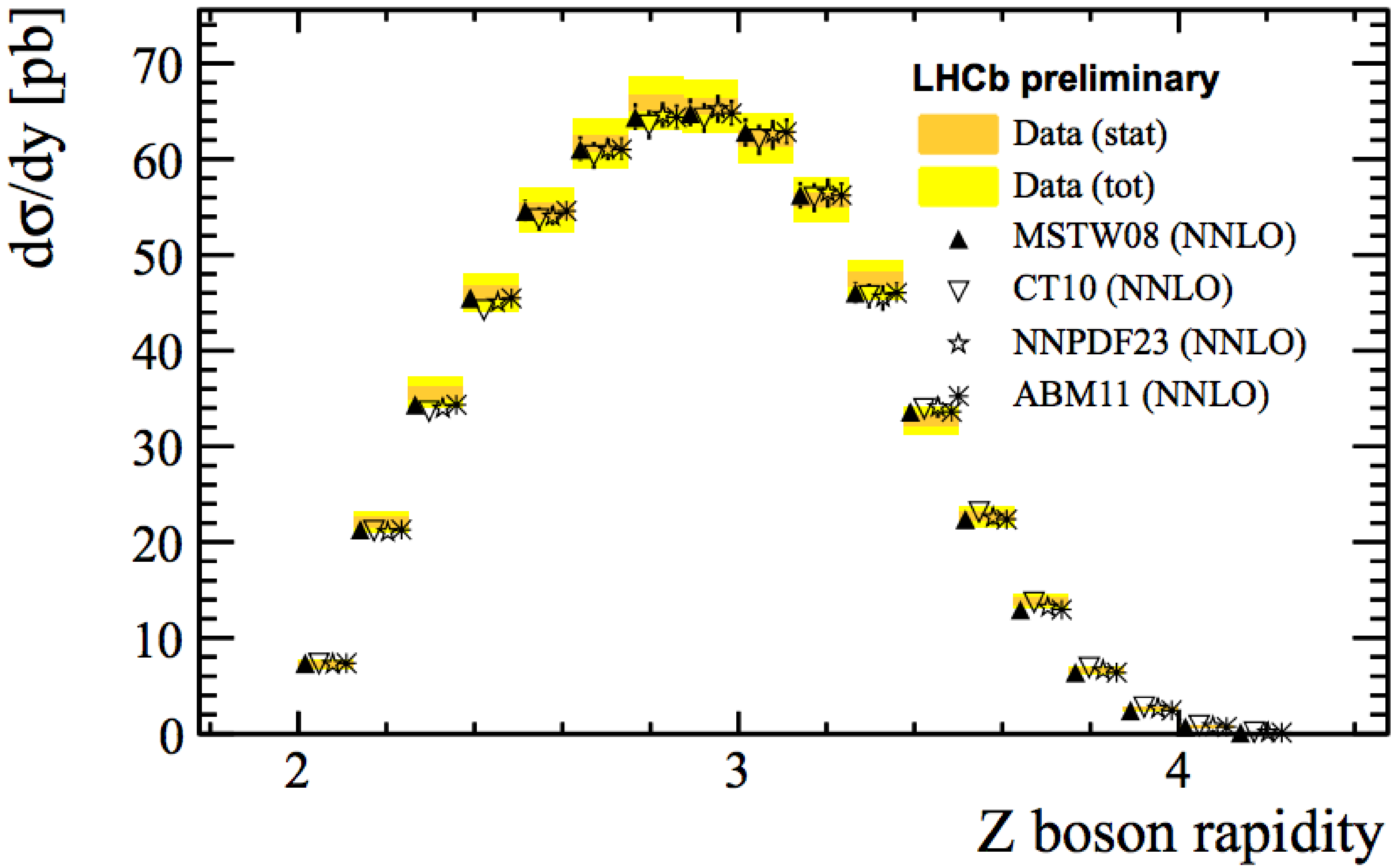
So far few dramatic effects on PDFs. Mainly strange quark (mainly at NNLO) and low-x valence quarks, largely due to change in methodology, but also to newer data. Larger strange uncertainty from branching ratio error.

Some uncertainty in the manner NNLO may affect jets. Away from vicinity of threshold decide at present to wait for full NNLO calculation.

Back-up



Points near to $y = 3.5$ overshoot predictions in general. Feature not present in prelim. higher luminosity $Z \to \mu^+\mu^-$ data.



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 $K = 1.15$.

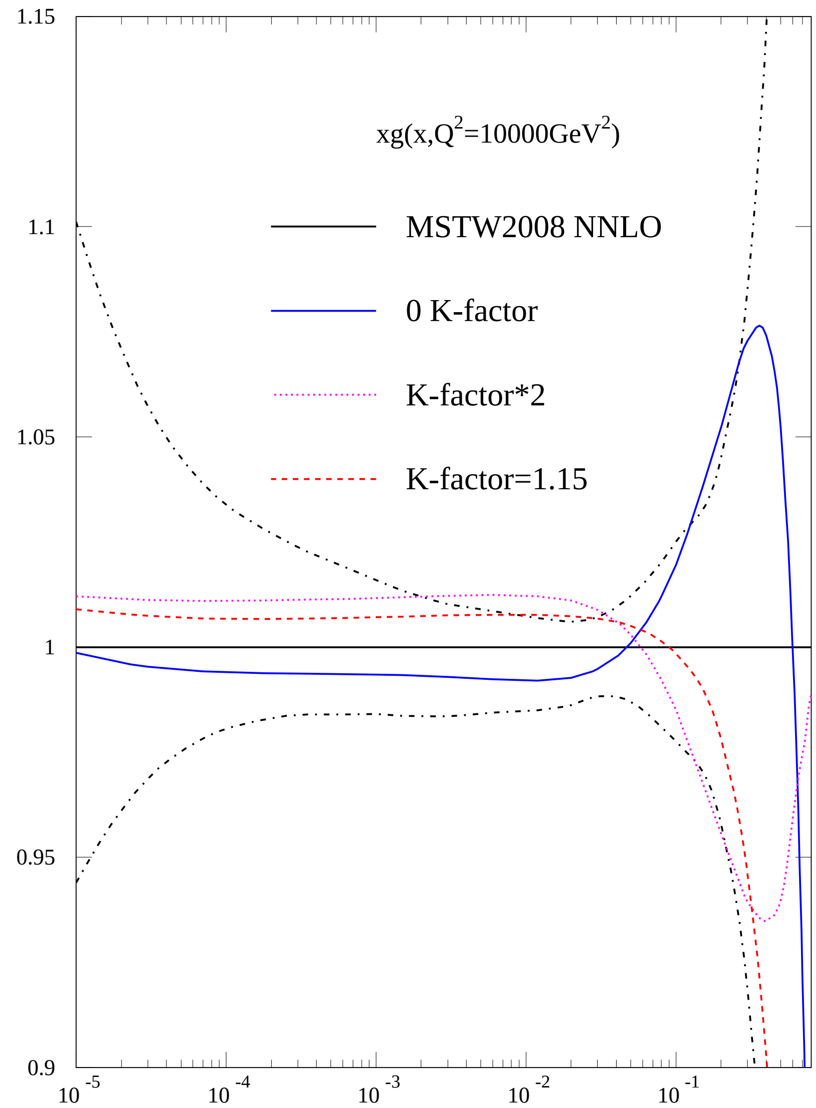
Extreme variations.

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.

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

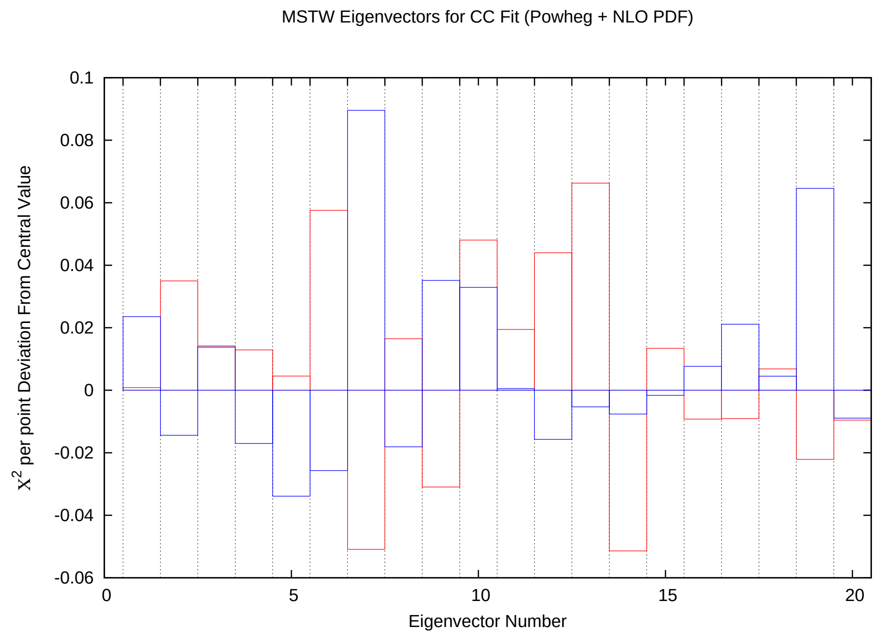
$K * 2 \ \alpha_S(M_Z^2) = 0.1159$

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

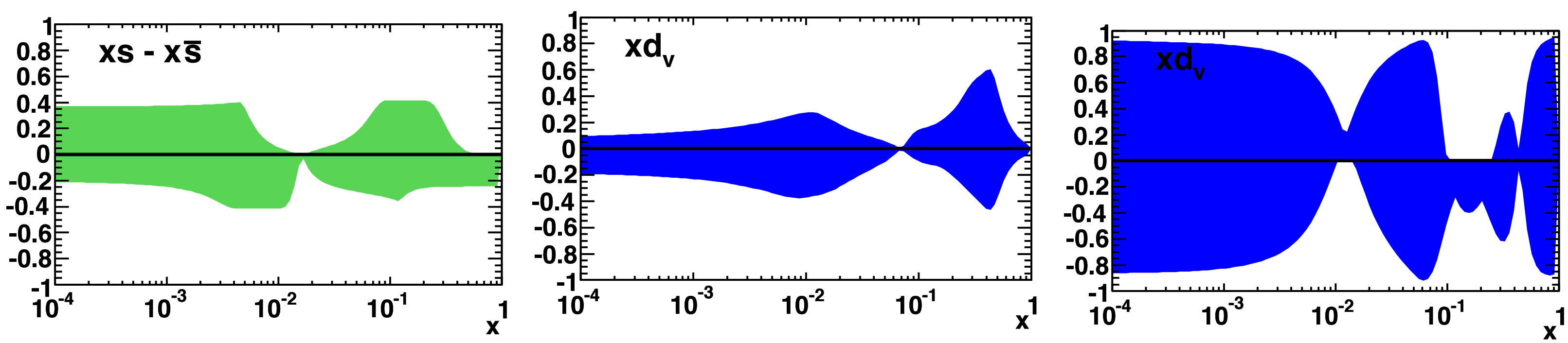


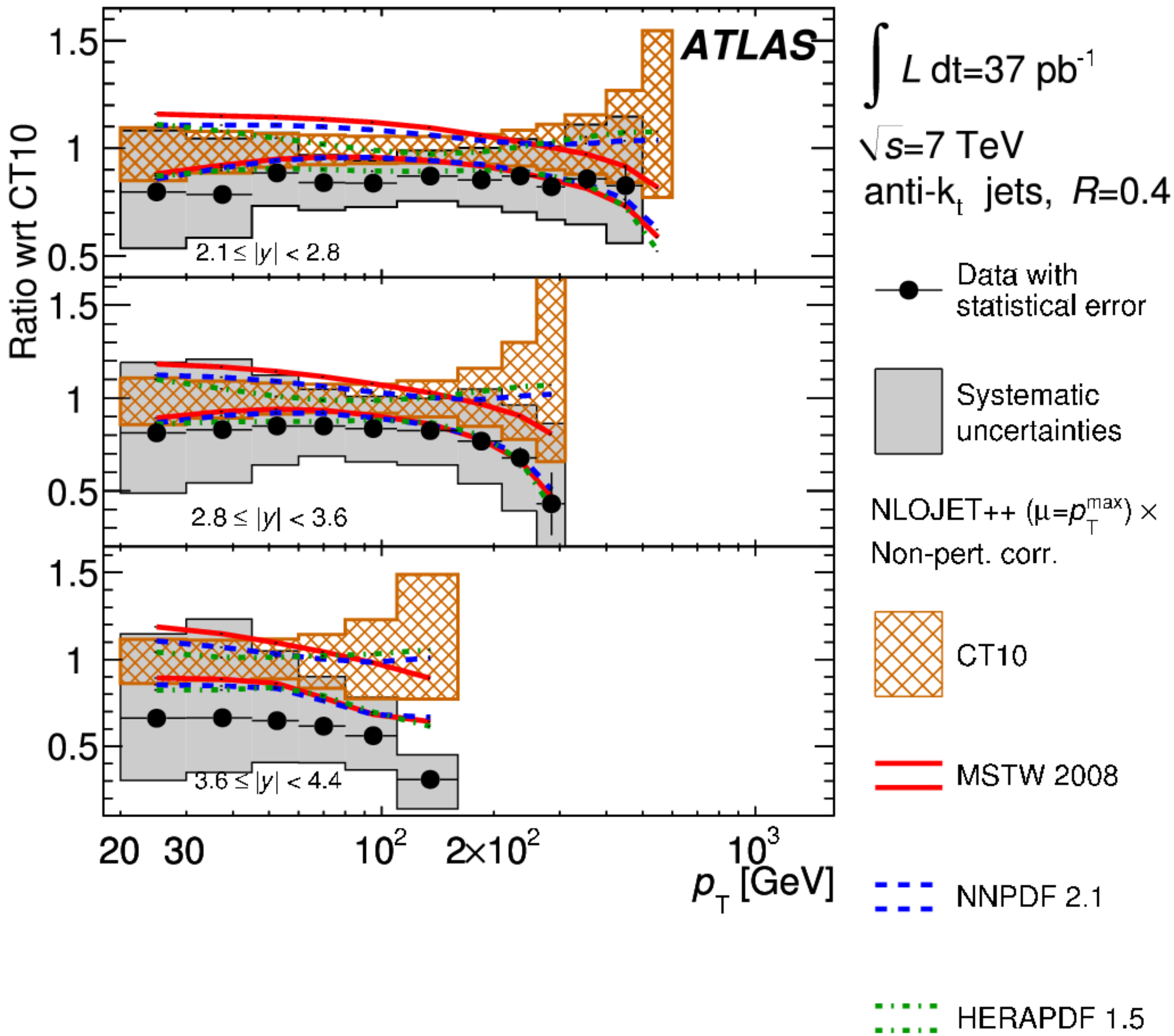
DIS jets - B. Watt

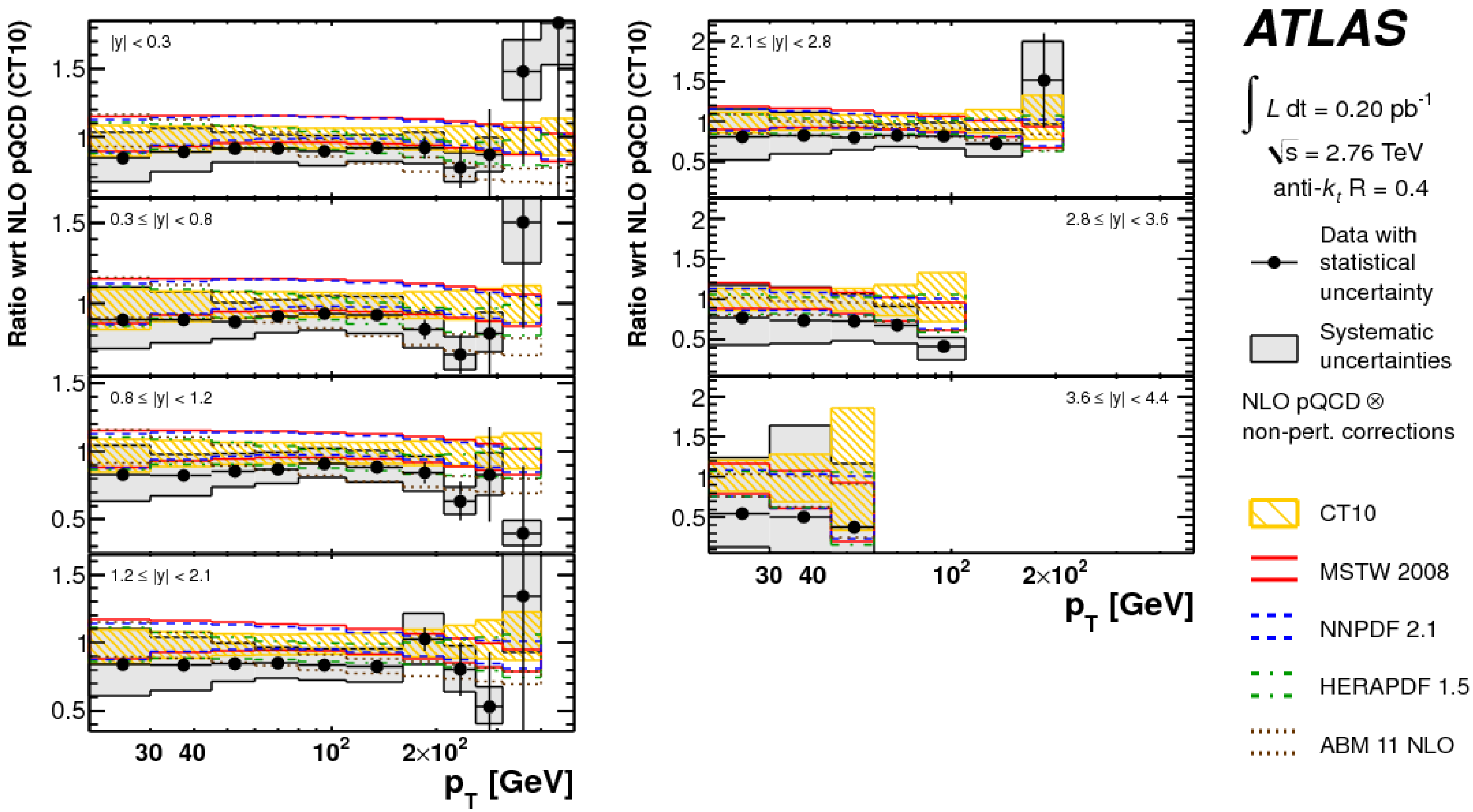
Interesting observation in eigenvector sensitivity to charged current ZEUS jet data usingPOWHEG.



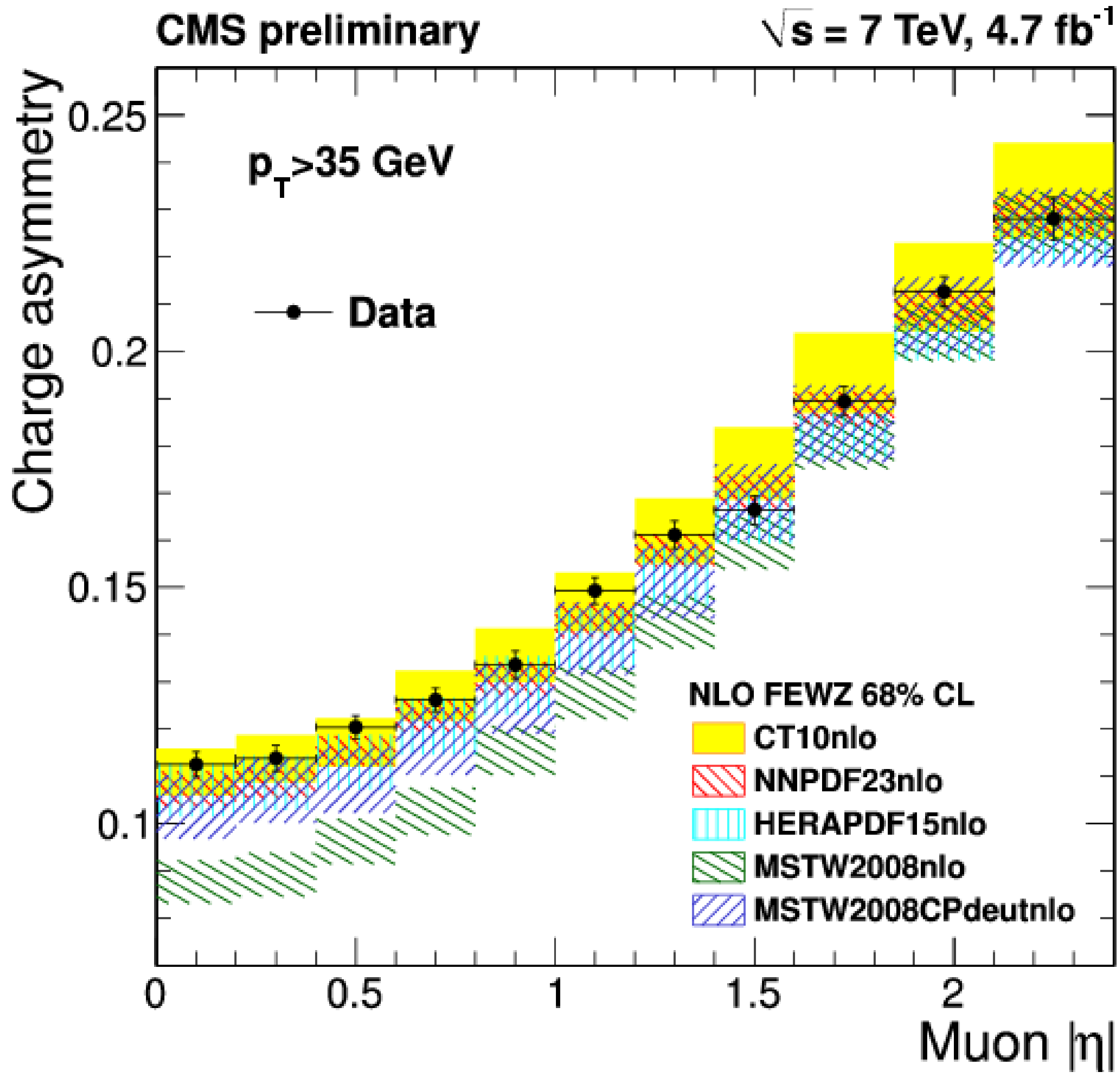
Some sensitivity to $s - \bar{s}$.



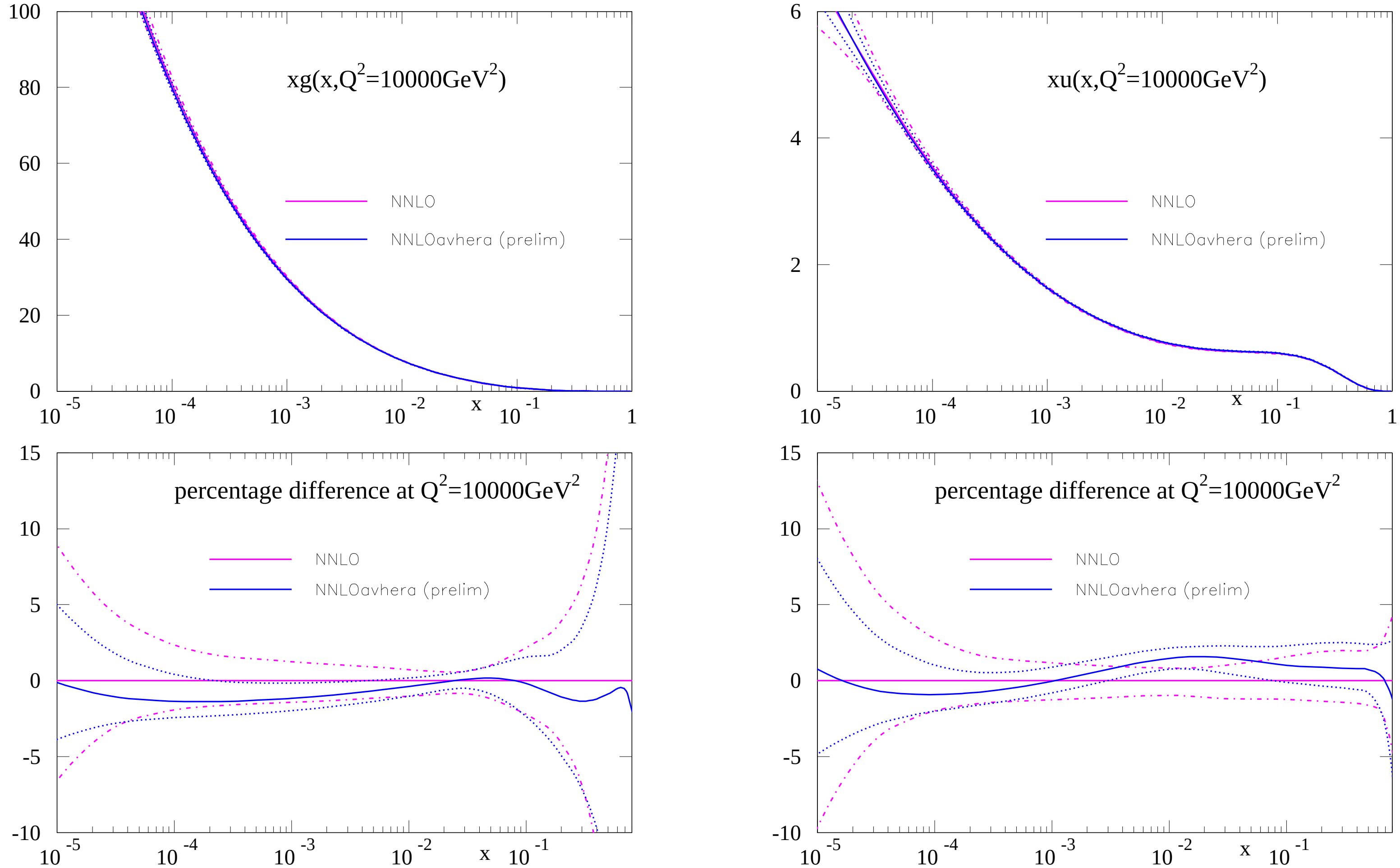

ATLAS
$\int L\,dt$=37 pb^{-1}
$\sqrt{s}$=7 TeV
anti-k$_t$ jets, R=0.4
Data with statistical error
Systematic uncertainties
NLOJET++ ($\mu=p_T^{max}$) × Non-pert. corr.
CT10
MSTW 2008
NNPDF 2.1
HERAPDF 1.5
Ratio wrt CT10
$2.1 \leq |y| < 2.8$
$2.8 \leq |y| < 3.6$
$3.6 \leq |y| < 4.4$
p_T [GeV]
1.5
1
0.5
20
30
10^2
2×10^2
10^3



ATLAS
∫ L dt = 0.20 pb^-1
√s = 2.76 TeV
anti-k_t 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
p_T [GeV]

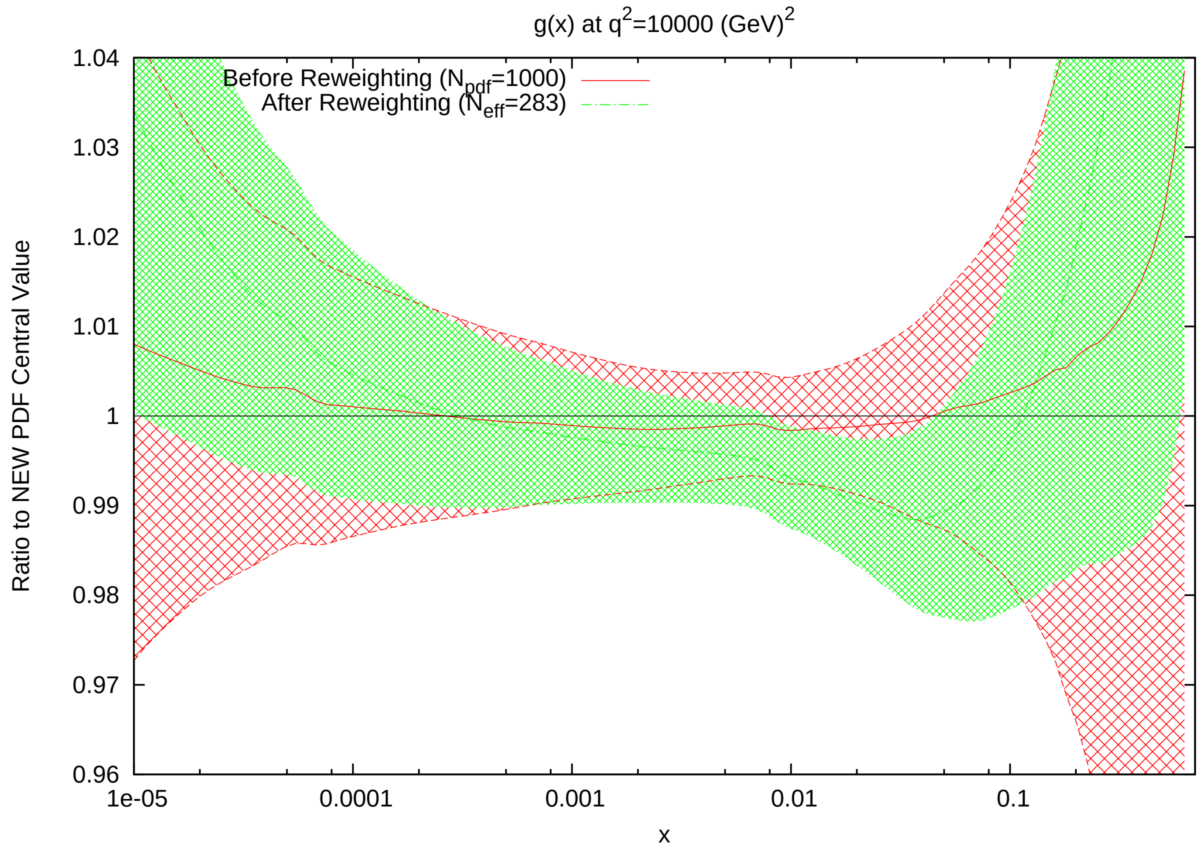


Comparison of various PDFs to CMS lepton asymmetry.



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.

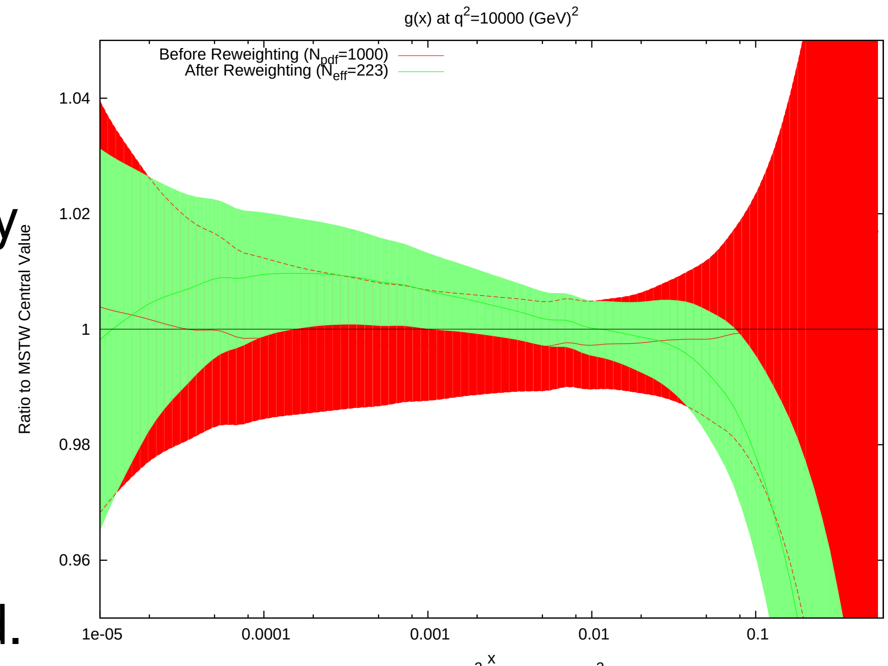
Reflection of different shape of higher order corrections?

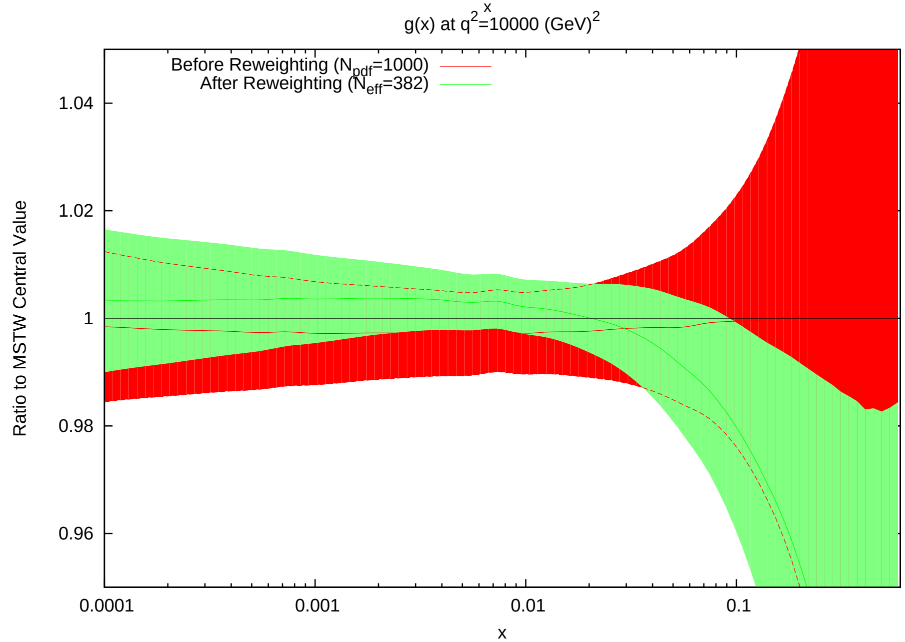
Different conclusions for fits to D0 and ATLAS dijet data, though they are not not necessarily incompatible.

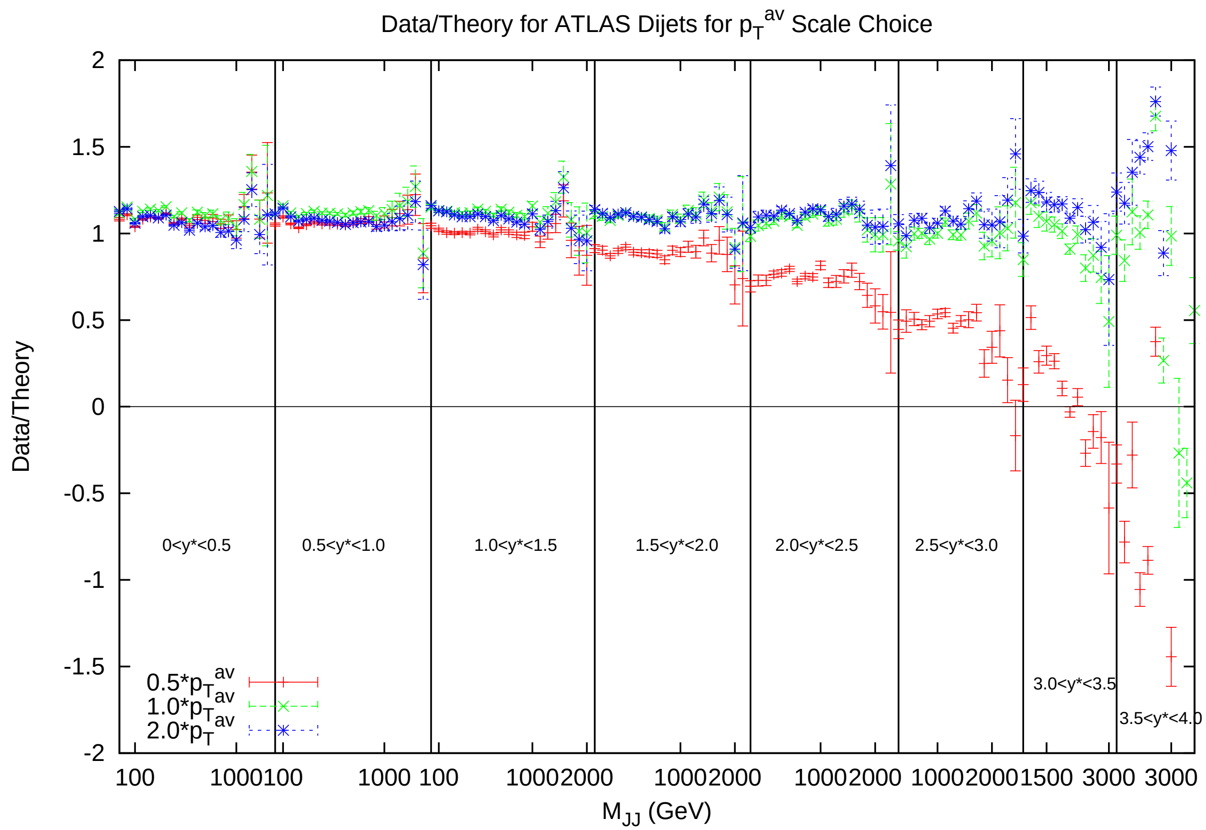
Similar to changes required by LHC inclusive jet data.

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.