Updates of PDFs and top quark production

Robert Thorne

September 27th, 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.

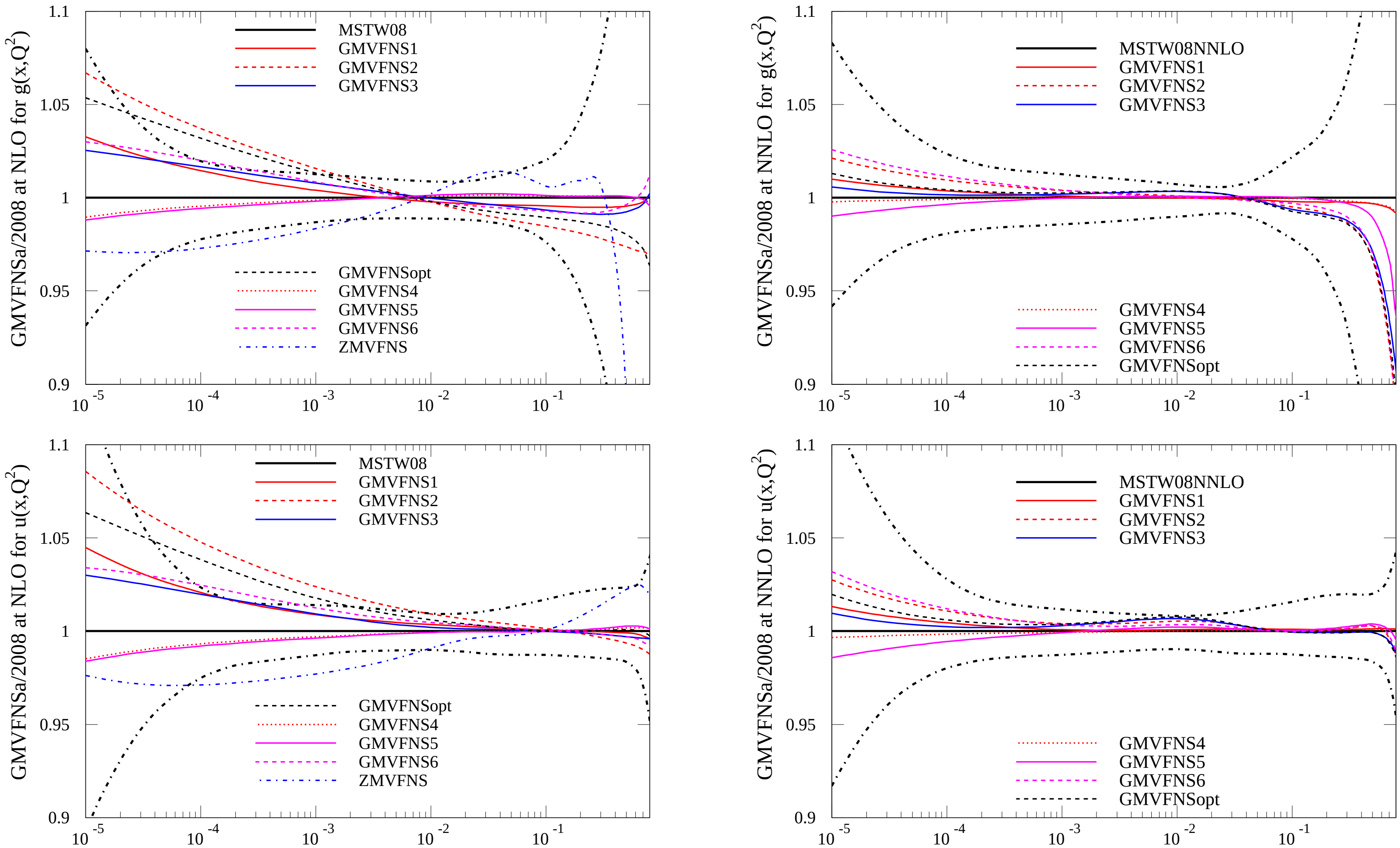
Changes in theoretical treatment or procedures.

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.

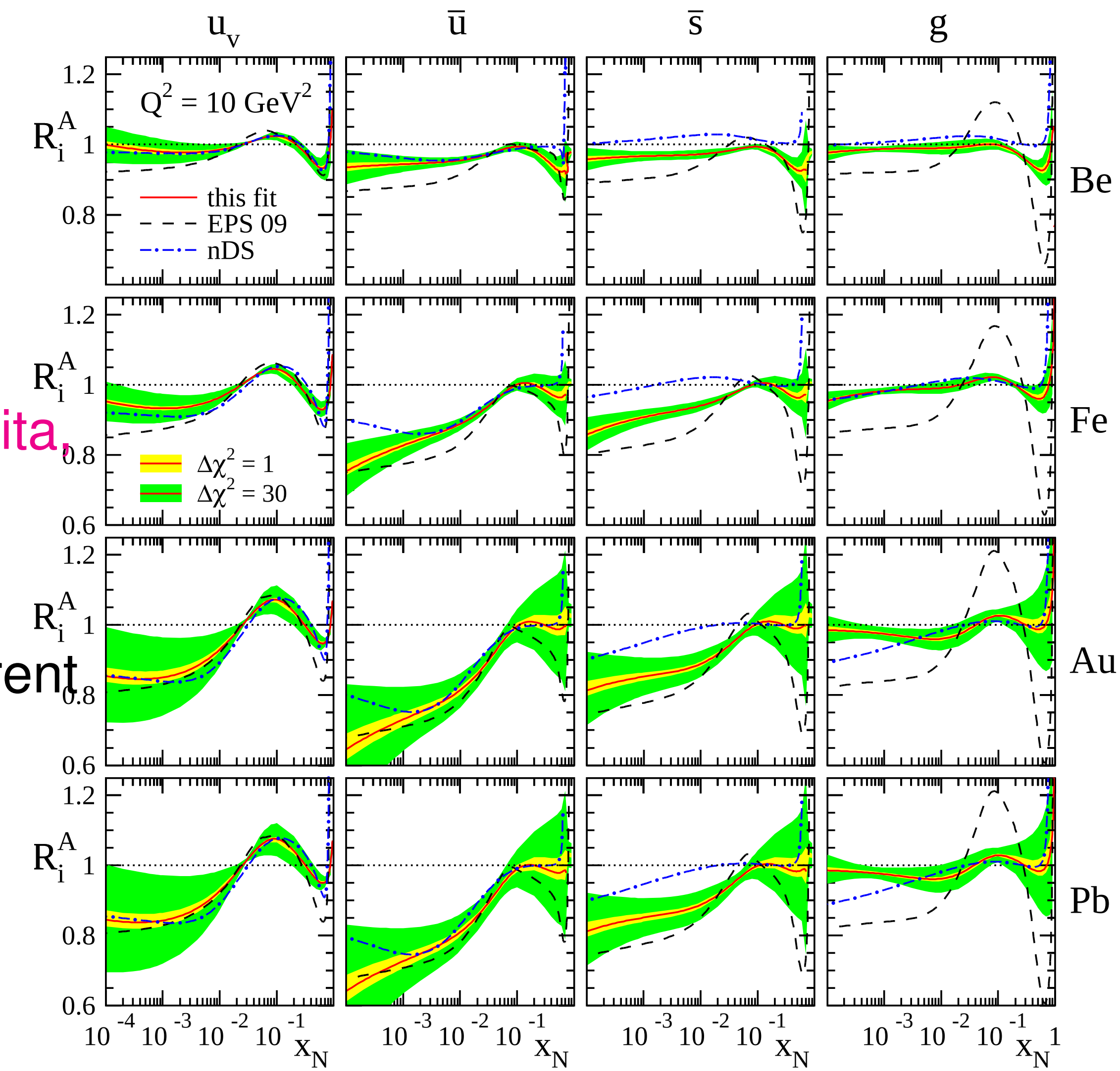
Strange branching ratio. Now avoid those determined by fits to dimuon data relying on PDF input. Also apply error which feeds into PDFs. Use $B_\mu = 0.092 \pm 10\%$ from hep-ex/9708014. Fits prefer $B_\mu = 0.082 - 0.090 \pm 15\%$.

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 $\sim$ 60-65 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.

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

Now with Harland-Lang and Motylinski 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 $Z \rightarrow e^+e^-$. Both predicted/fit well at NLO. For the latter theory a bit low at NNLO at $y \sim 3.5$. Not evident in preliminary $Z \rightarrow \mu^+\mu^-$ data with higher precision.

Include CMS data on $Z \rightarrow e^+e^-$, 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.

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

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 $\alpha_S(M_Z^2) = 0.1199$.

LHC jets not included at NNLO.

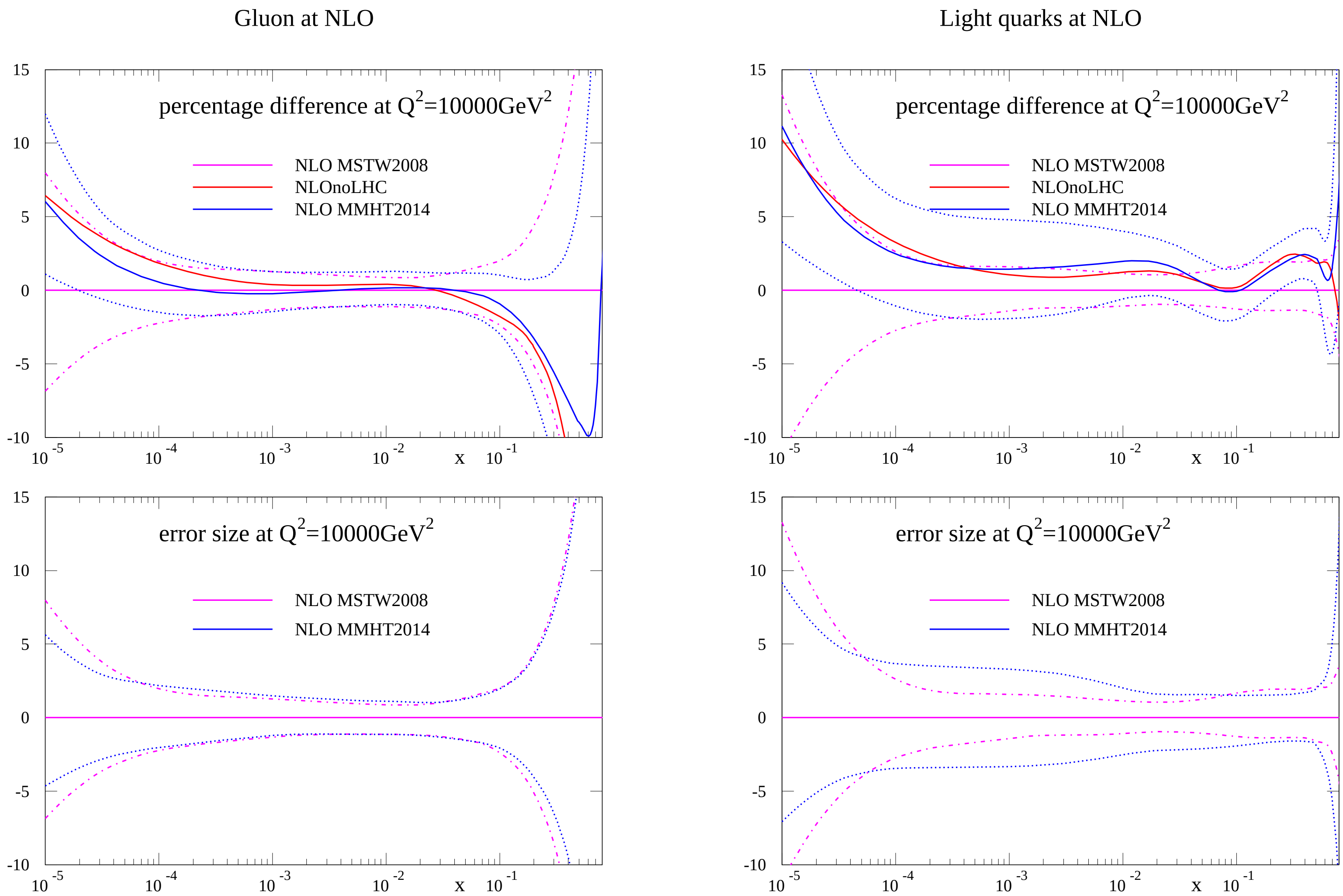
Fit quality for LHC data at NLO

data set	N_{pts}	CPdeut	no LHC	MMHT
ATLAS jets (2.76TeV+7TeV)	116	107	107	106
CMS jets (7TeV)	133	140	143	138
ATLAS W^+, W^-, Z	30	47	44	39
CMS W asymm $p_T > 35\mathrm{GeV}$	11	9	16	7
CMS asymm $p_T > 25\mathrm{GeV}, 30\mathrm{GeV}$	24	9	17	7
LHCb $Z \to e^+e^-$	9	13	13	13
LHCb W asymm $p_T > 20\mathrm{GeV}$	10	12	14	12
CMS $Z \to e^+e^-$	35	21	22	20
ATLAS High mass DY	13	20	20	21
TeV, ATLAS, CMS $\sigma_{t\bar{t}}$	13	8	10	7
CMS Low-high mass DY	132	385	396	373

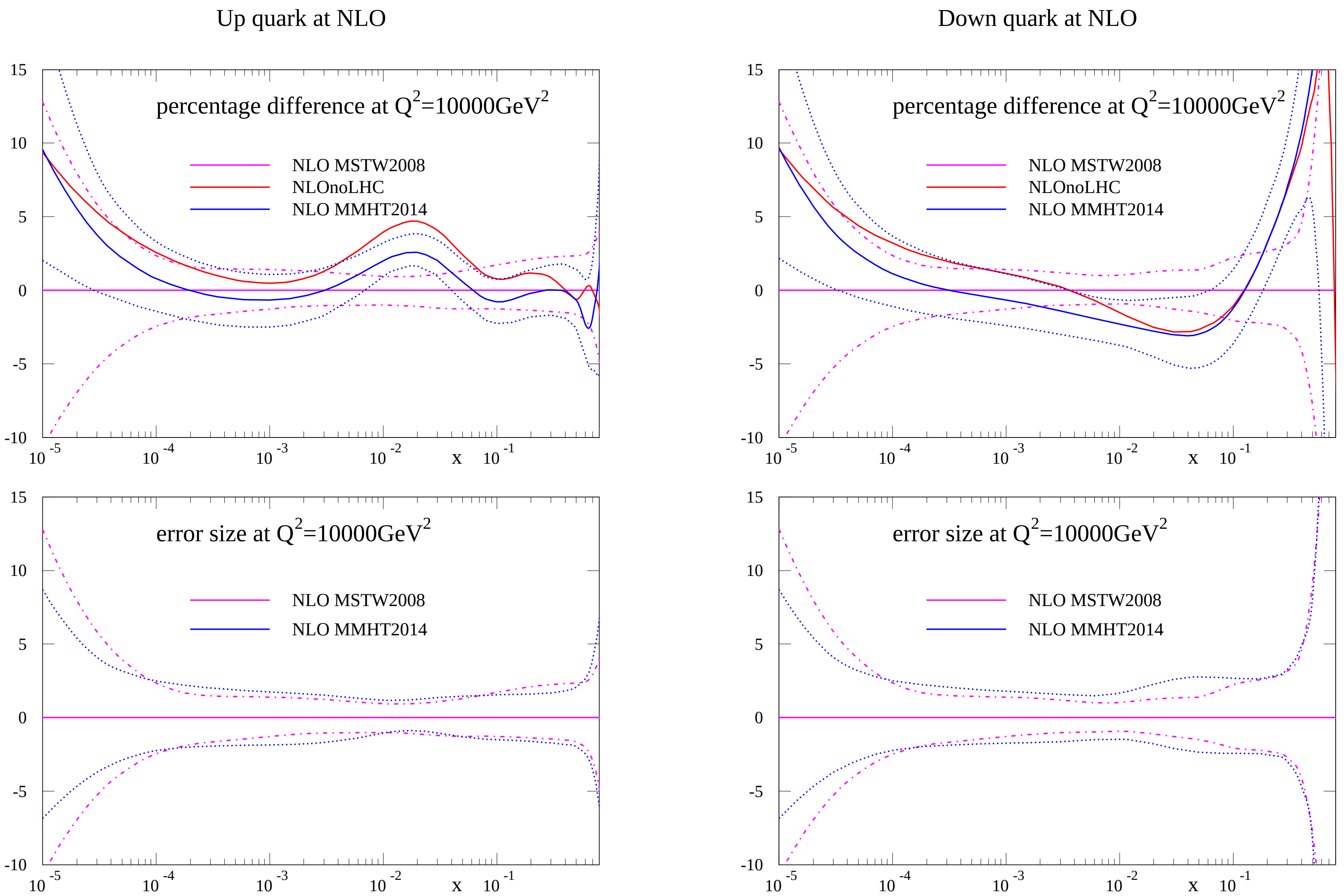
ATLAS W, Z data constrains the gluon as do $\sigma_{t\bar{t}}$ and CMS $Z \to e^+e^-$ data.

CMS W asymm. data constrains some flavour decomposition.

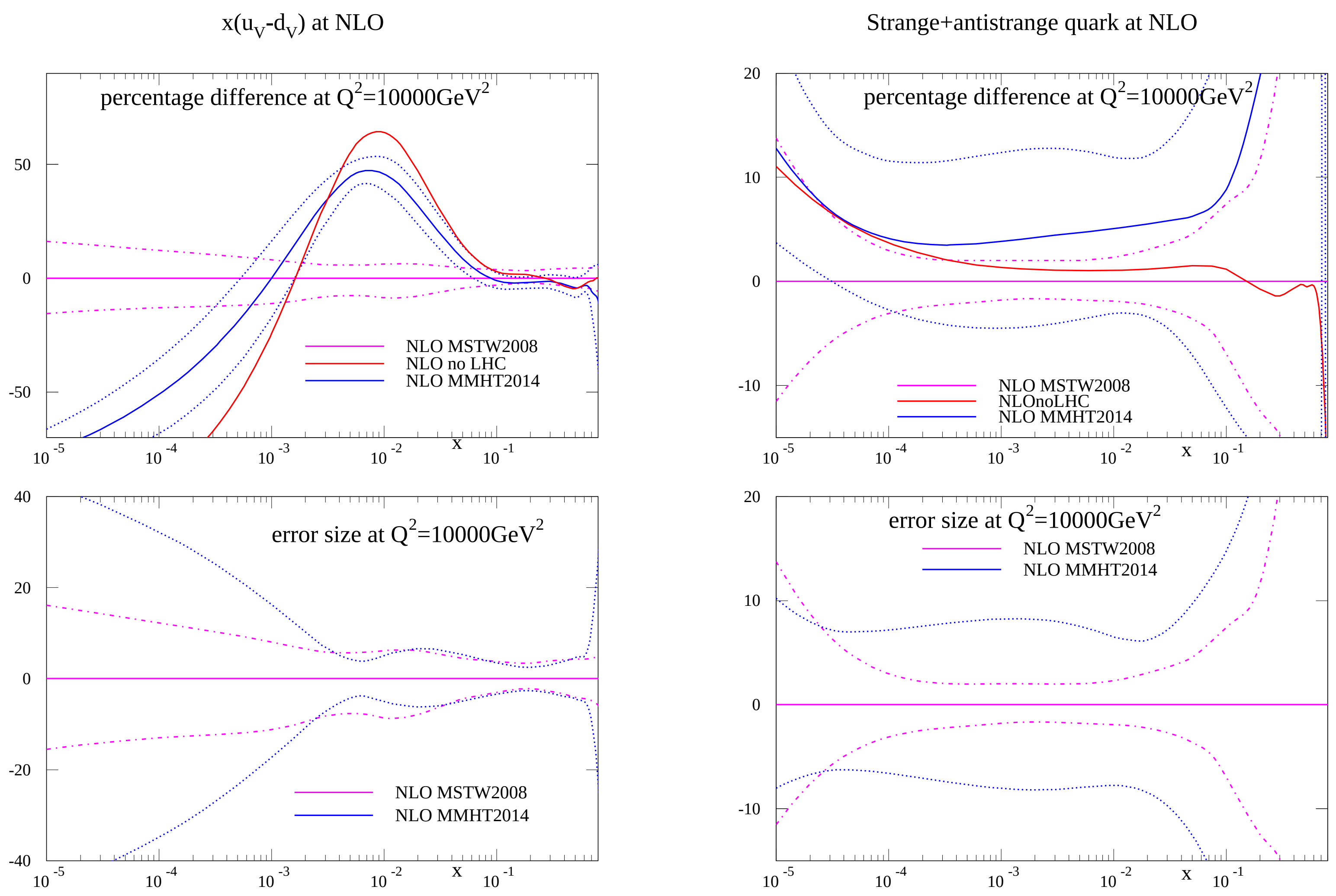
Fit CMS double differential low and high mass Drell Yan data. No real change in PDFs. Fit very poor.



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



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

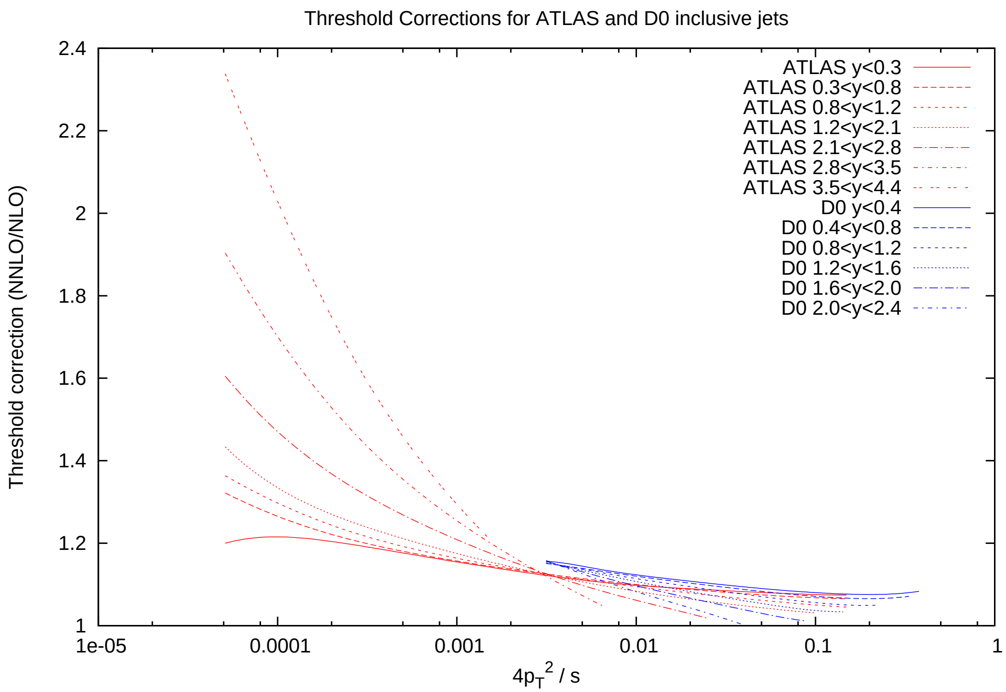


Change in NLO PDFs from all, including LHC data updates. Much expanded $s + \bar{s}$ uncertainty.

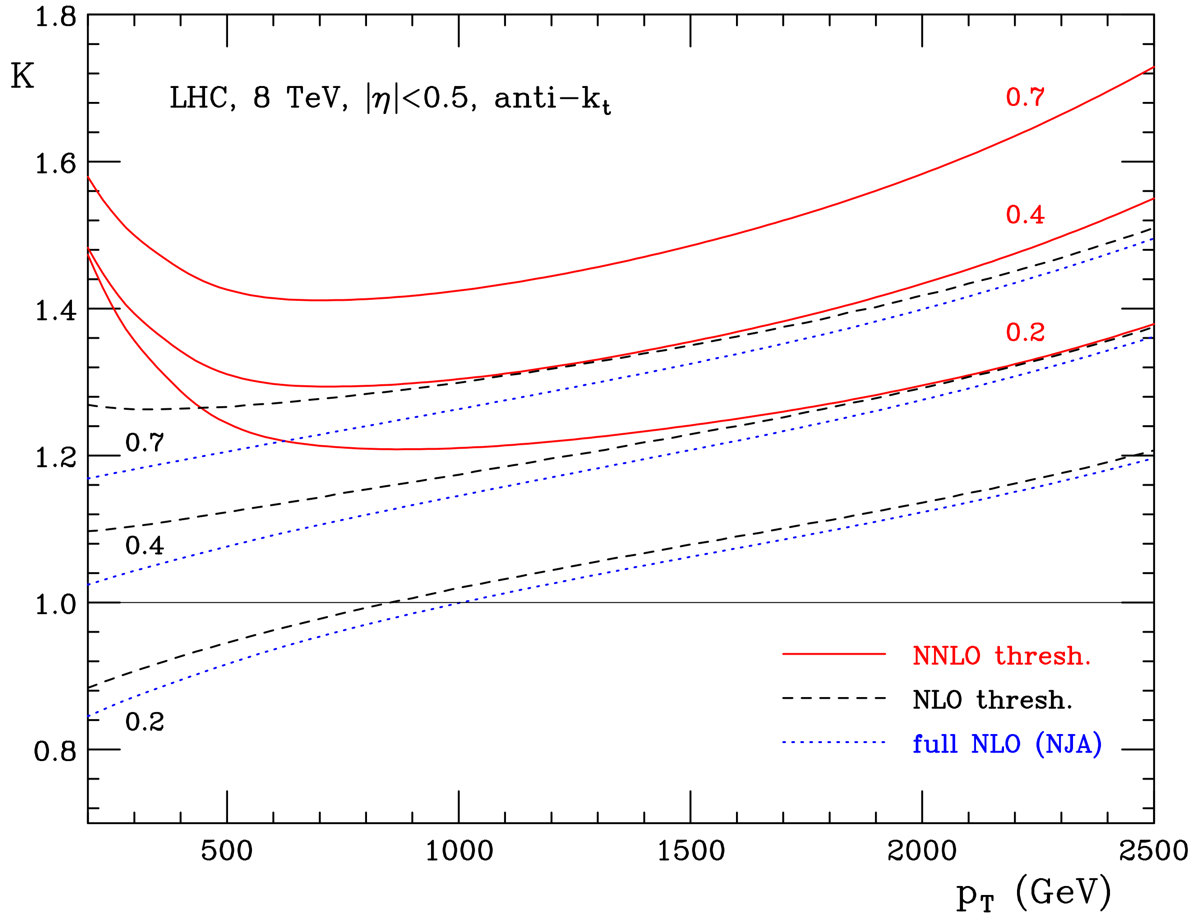
LHC jet data at NNLO?

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.



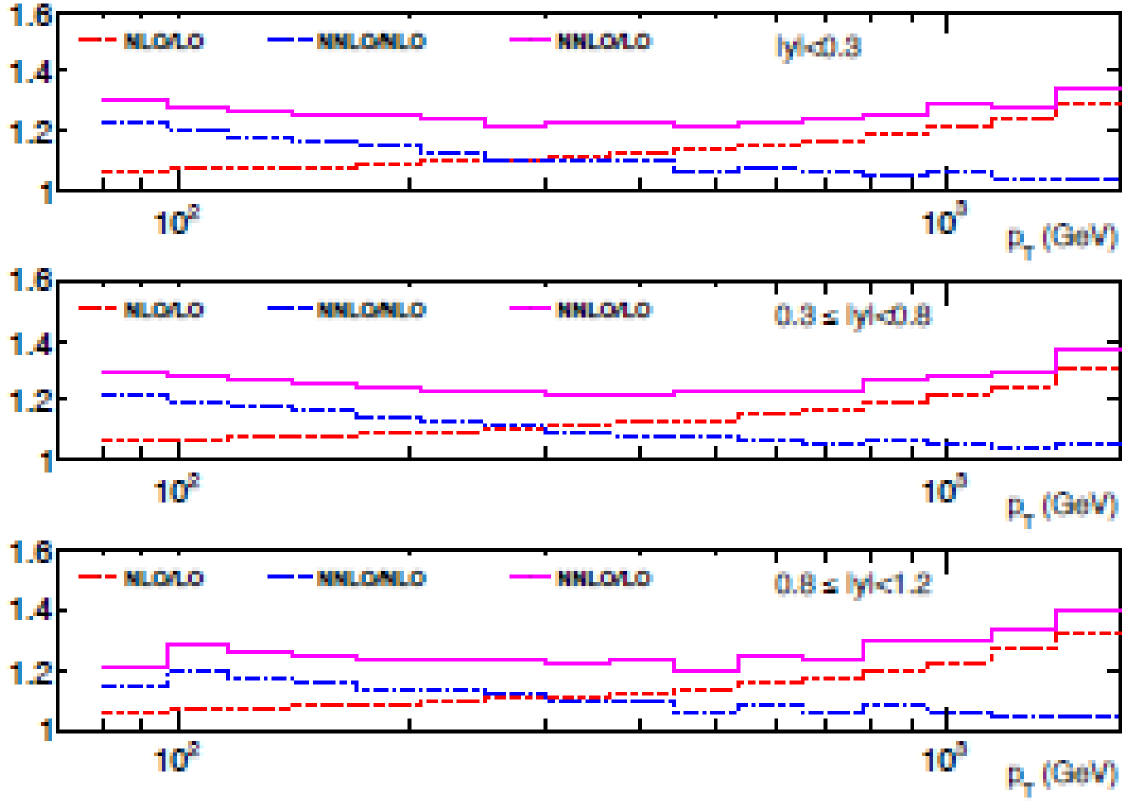
Enormous project of full NNLO calculation (Gehrmann-de-Ridder, Gehrmann, Glover and Pires) nearing completion. Some indications of full form of the correction.



Recent 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

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 $\sim 5\text{--}20\%$ 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.

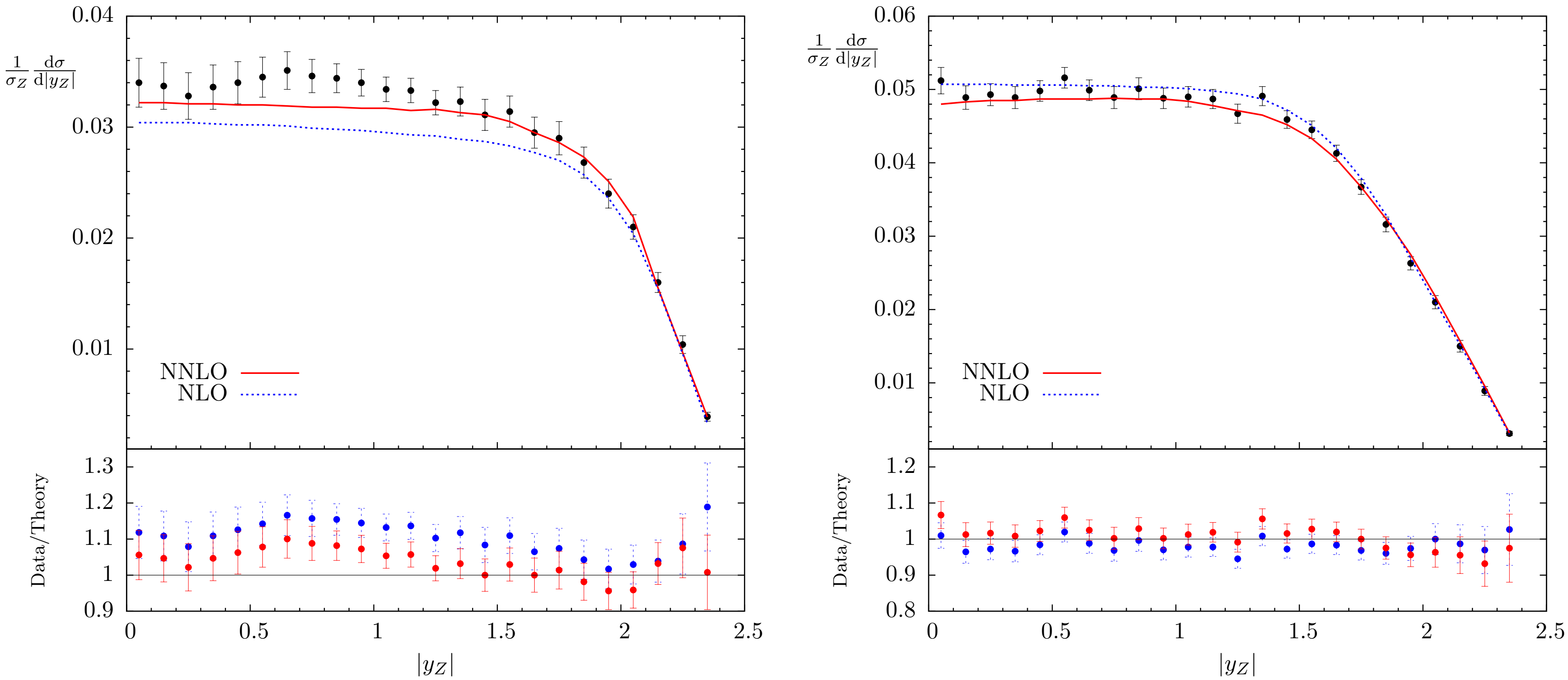
data set	N_{pts}	CPdeut	no LHC	MMHT
ATLAS jets (2.76TeV+7TeV)	116	(107)	(123)	(119)
CMS jets (7TeV)	133	(142)	(137)	(135)
ATLAS W^+, W^-, Z	30	72	53	39
CMS W asymm $p_T > 35\mathrm{GeV}$	11	18	15	9
CMS asymm $p_T > 25\mathrm{GeV}, 30\mathrm{GeV}$	24	18	17	10
LHCb $Z \to e^+e^-$	9	23	22	20
LHCb W asymm $p_T > 20\mathrm{GeV}$	10	24	21	13
CMS $Z \to e^+e^-$	35	30	24	22
ATLAS High mass DY	13	18	16	17
TeV, ATLAS, CMS $\sigma_{t\bar{t}}$	13	8	11	8
CMS Low-high mass DY	132	159	151	149

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

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

CMS Drell Yan data.

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



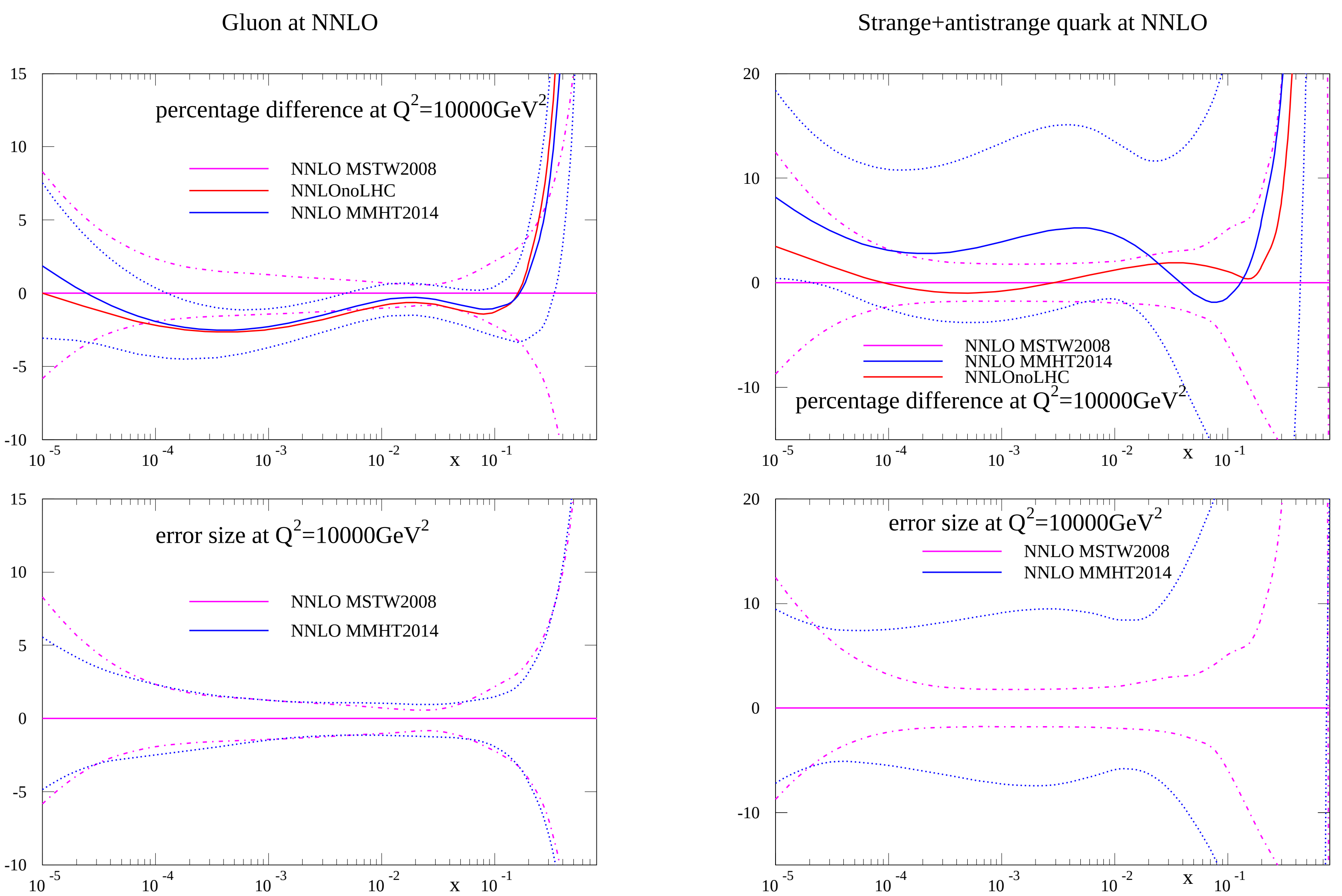
Enormously improved fit quality at NNLO due to improvement in cross-sections.

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

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	MMHT
ATLAS jets (2.76TeV+7TeV)	116	(117)	(132)	(128)
CMS jets (7TeV)	133	(145)	(137)	(139)
ATLAS W^+, W^-, Z	30	72	53	39
CMS W asymm $p_T > 35\mathrm{GeV}$	11	18	15	9
CMS asymm $p_T > 25\mathrm{GeV}, 30\mathrm{GeV}$	24	18	17	10
LHCb $Z \to e^+e^-$	9	23	22	20
LHCb W asymm $p_T > 20\mathrm{GeV}$	10	24	21	13
CMS $Z \to e^+e^-$	35	30	24	22
ATLAS High mass DY	13	18	16	17
TeV, ATLAS, CMS $\sigma_{t\bar{t}}$	13	8	11	8
CMS Low-high mass DY	132	159	151	149

ATLAS jet data deteriorates more than CMS, which with increase in systematics is largely insensitive to K-factor, though even prefers smaller one. 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). At NNLO final extracted $\alpha_S(M_Z^2) = 0.1172$

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

$\alpha_S(m_Z^2)$ coming out similar to 2008 fit. Still a NLO/NNLO difference. Both fairly compatible with global average. Try inputting 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 – $\alpha_S(m_Z^2) = 0.1120 \rightarrow 0.1195$ with $\Delta\chi^2 < 2$.

At NNLO best fit gives $\alpha_S(m_Z^2) = 0.1172 \rightarrow 0.1177$, i.e. very close to 0.118. $\Delta\chi^2 < 2$

Also force $\alpha_S(m_Z^2) = 0.118$. At NNLO basically no further change. At NLO $\Delta\chi^2 \sim 16$, but no single set deteriorates very significantly.

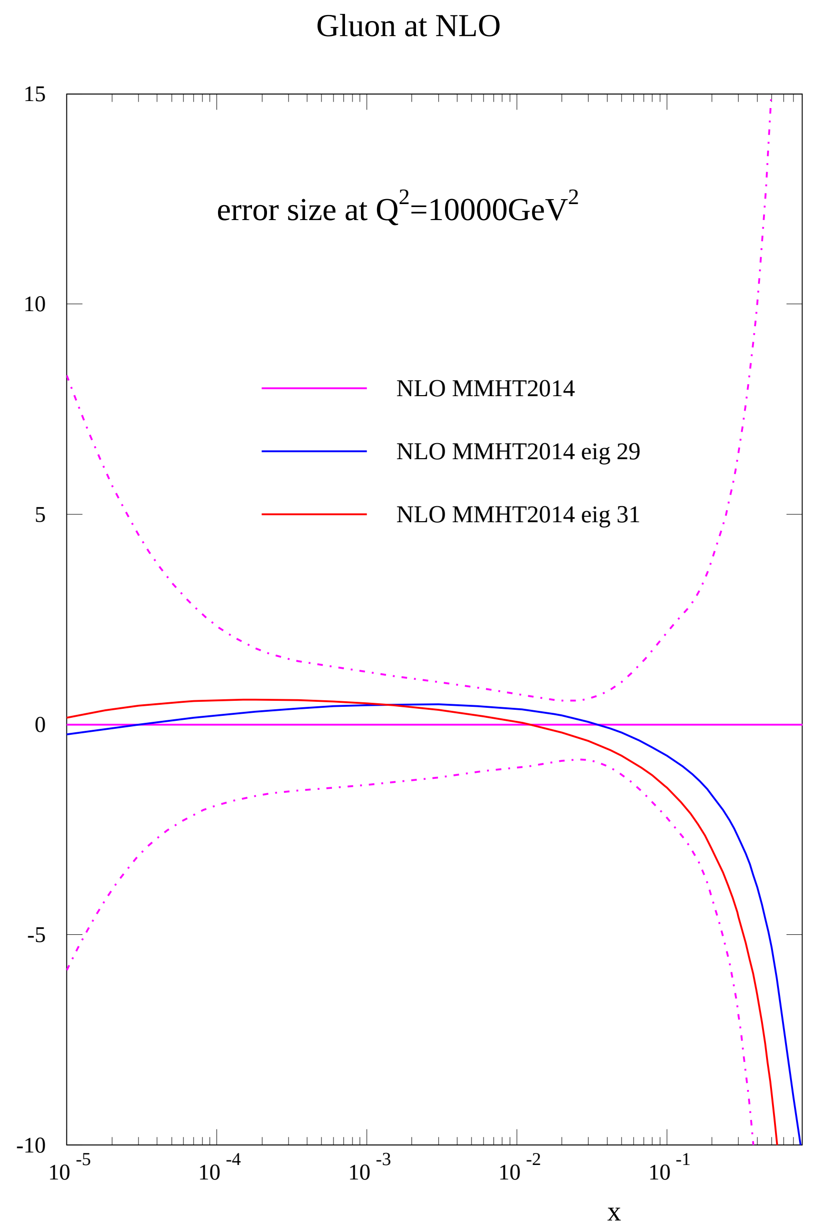
In the NLO fit the inclusive $t\bar{t}$ cross section does not constrain any eigenvectors. Best fit $m_t = 171.7\text{GeV}$ (lower if $\alpha_S(M_Z^2) = 0.118$).

Nearly constrains eigenvector number 29 and 31.

Both correspond to decreased gluon at high x only.

29 also corresponds to lower high-x sea and constrained mainly by NuTeV $F_3(x, Q^2)$ data.

31 primarily constrained by CDF jet data.



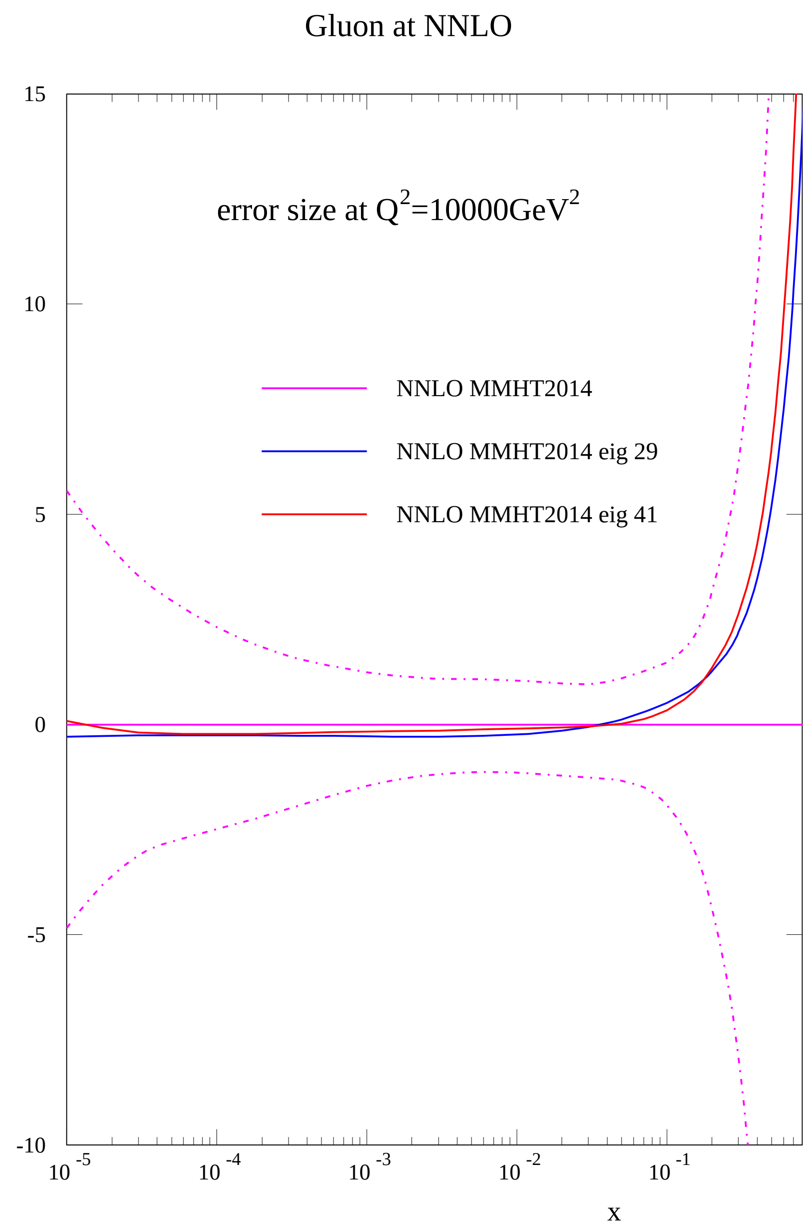
In the NNLO fit the inclusive $t\bar{t}$ cross section constrains one eigenvector.

At NNLO preferred $m_t = 174.1\mathrm{GeV}$.

Constrains eigenvector number 29 and (nearly) 41.

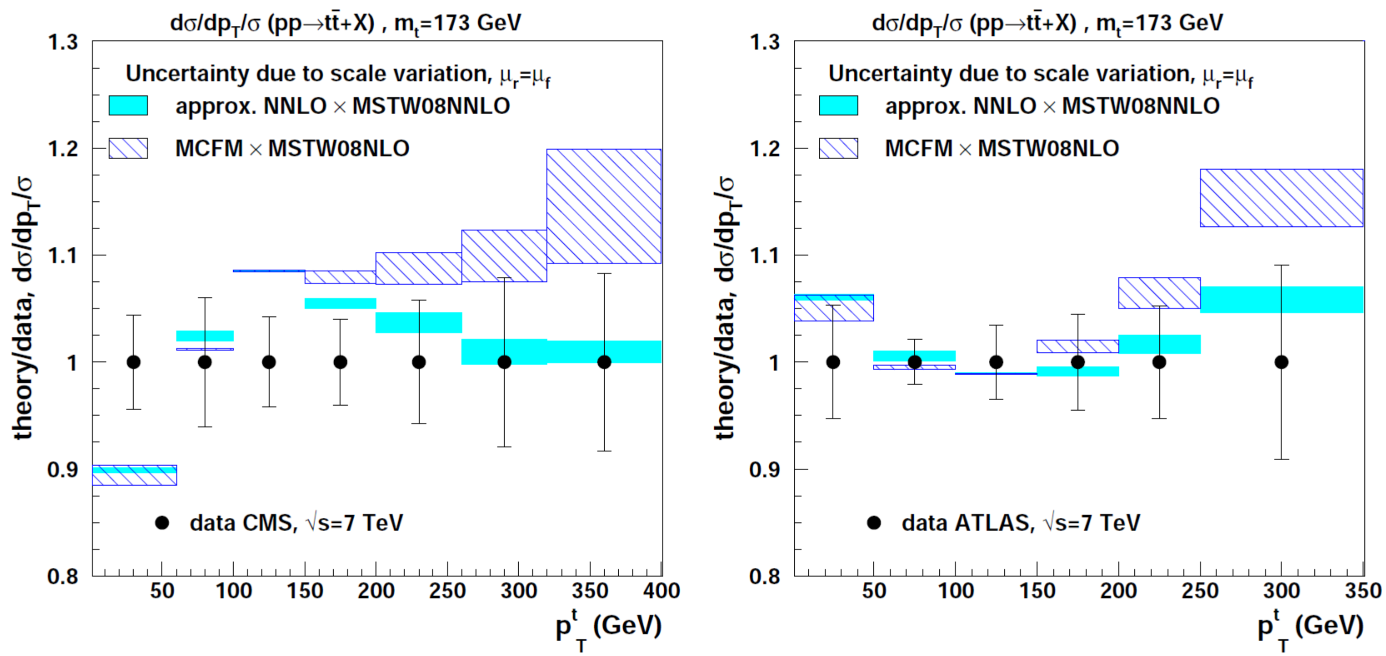
Both correspond to increased gluon at high x only.

41 also corresponds to strange normalisation and constrained also by ATLAS W, Z data.



Differential Data

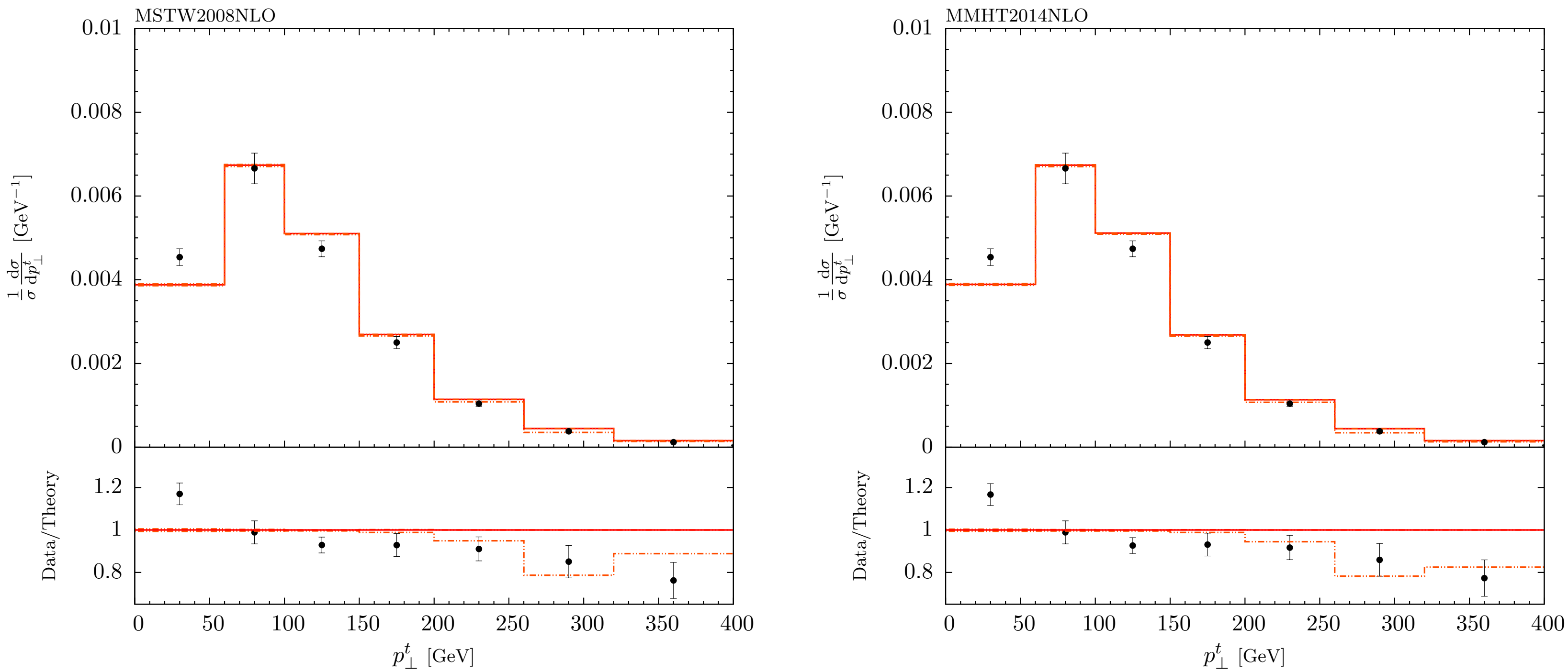
As it improves differential top production data will help constrain the gluon.



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.

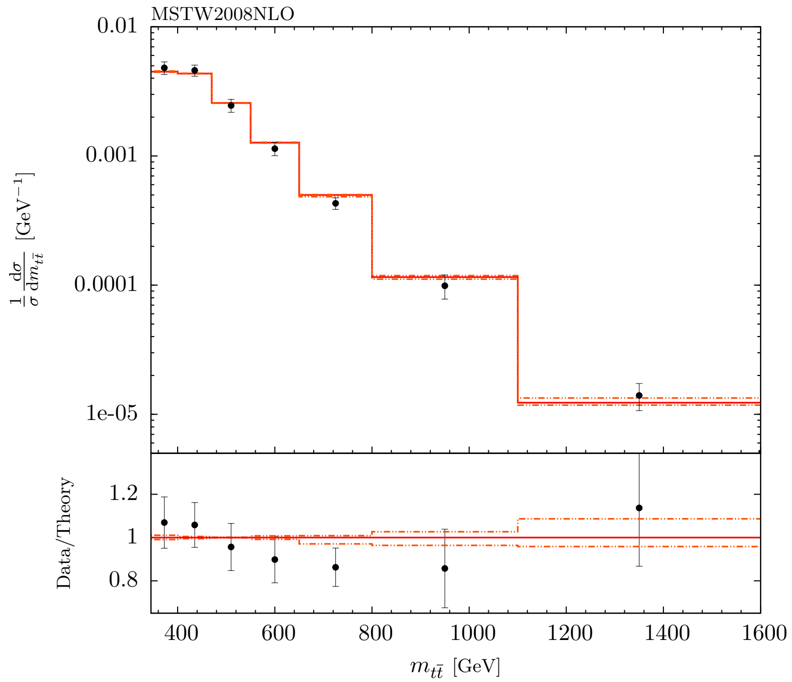
New results – plots by Harland-Lang

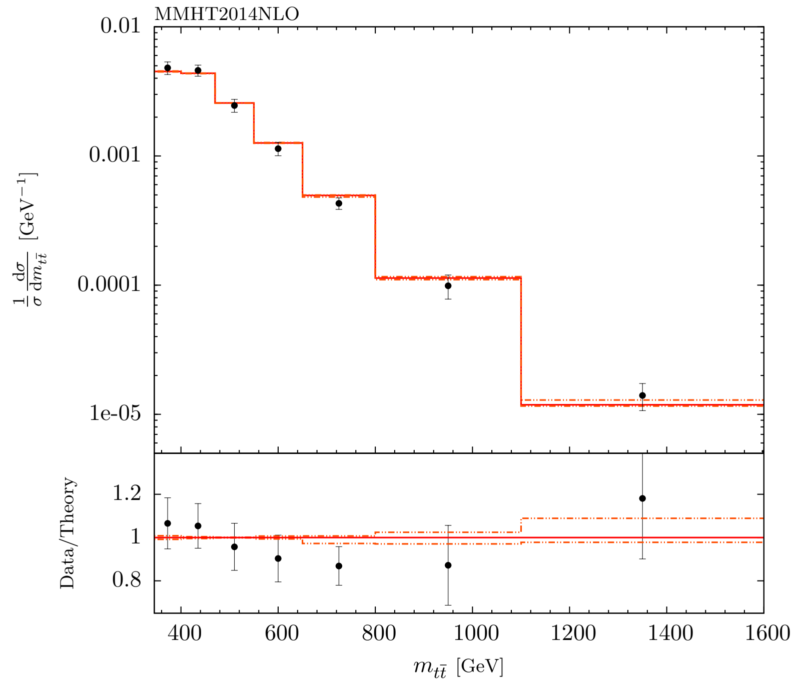
p_T distributions - CMS data



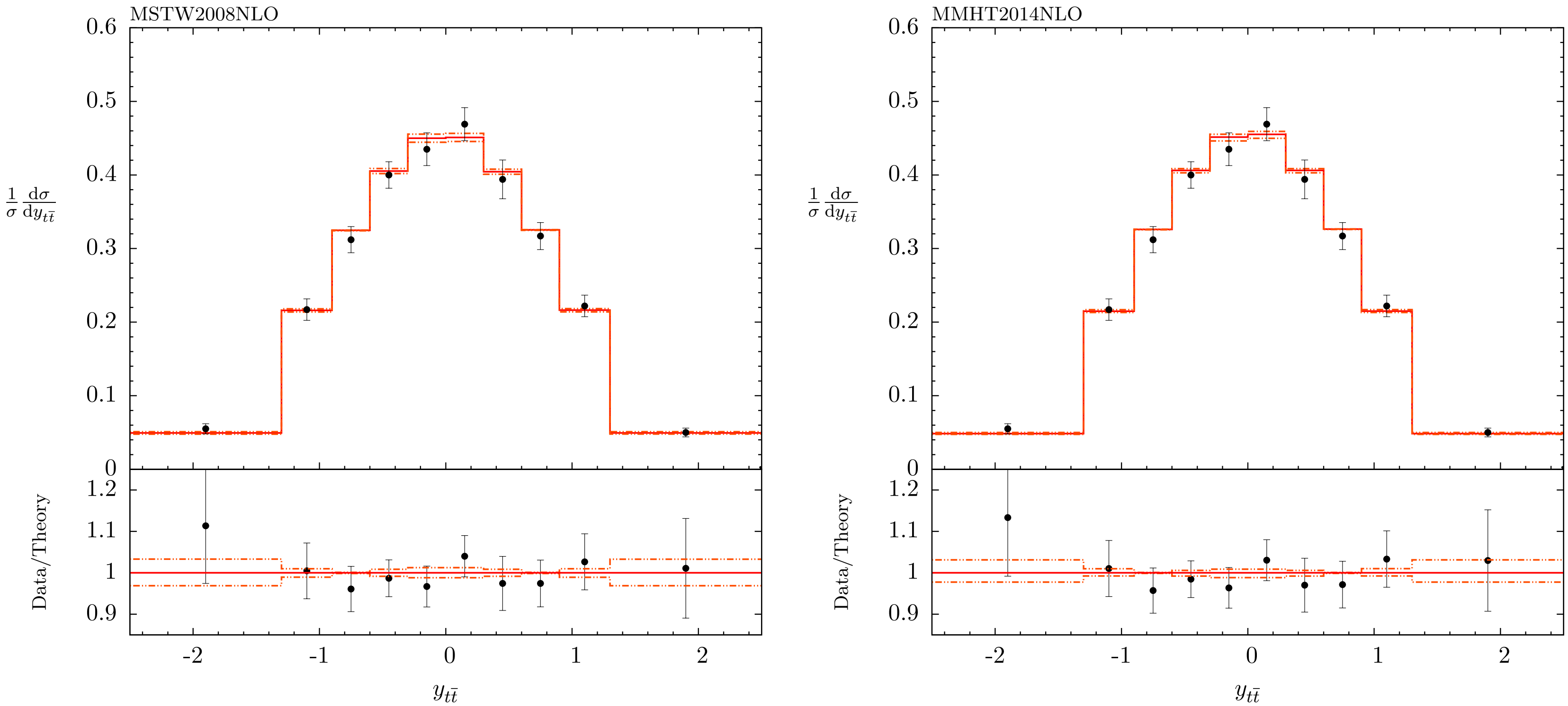
Very little difference between MSTW2008 and MMHT2014 predictions.

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





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



Conclusions

Soon to release updated PDFs. Assumed final.

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).

So far few dramatic effects on PDFs. Mainly slight increase in strange quark and increased strange uncertainty.

Inclusive $\sigma_{t\bar{t}}$ pulls pole top mass down a little at NLO and up a little at NNLO. Constrains (or nearly) a few eigenvectors. Essentially stops too small high-x gluon at NLO and stops too large high-x gluon at NNLO.

Some tensions between inclusive and differential conclusions at NLO. Hopefully reduced at NNLO?

MMHT2014 very similar to MSTW2008 for top disttributions.

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

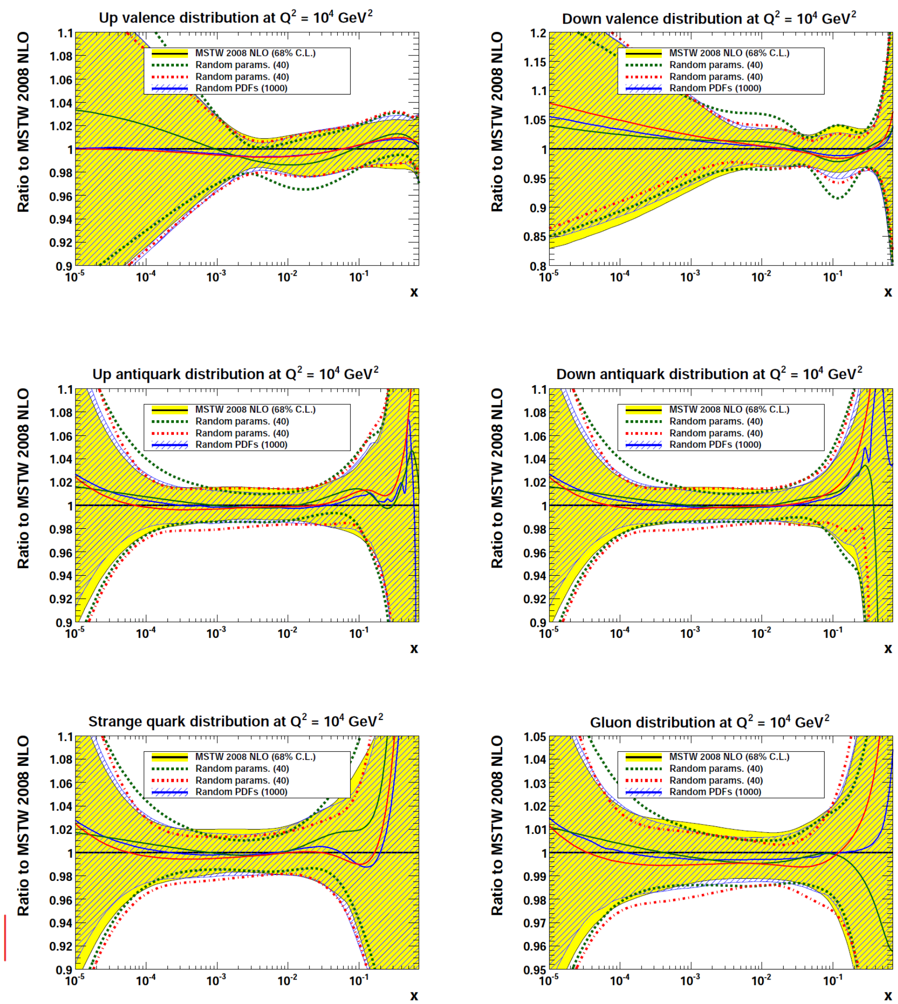
Study supported correctness of "dynamic tolerance" approach. Easiest in Hessian study with eigenvectors.

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

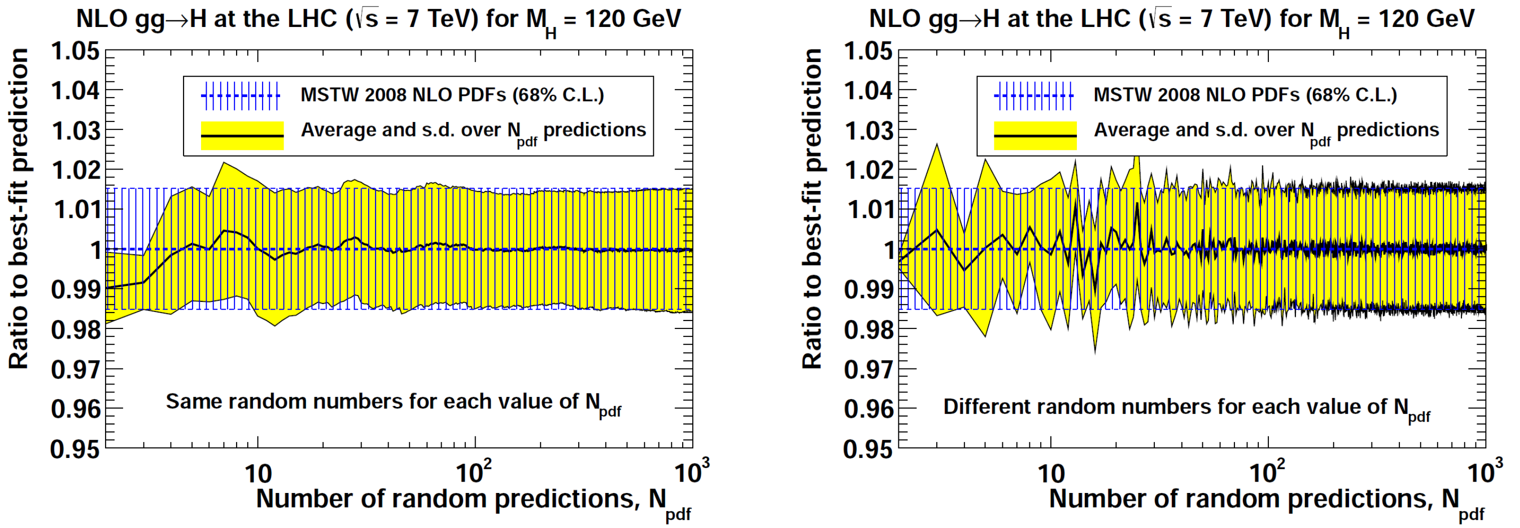
$$a_i(S_k) = a_i^0 + \sum_{j=1}^{n} e_{ij}(\pm t_j^{\pm})|R_{jk}|$$

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

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



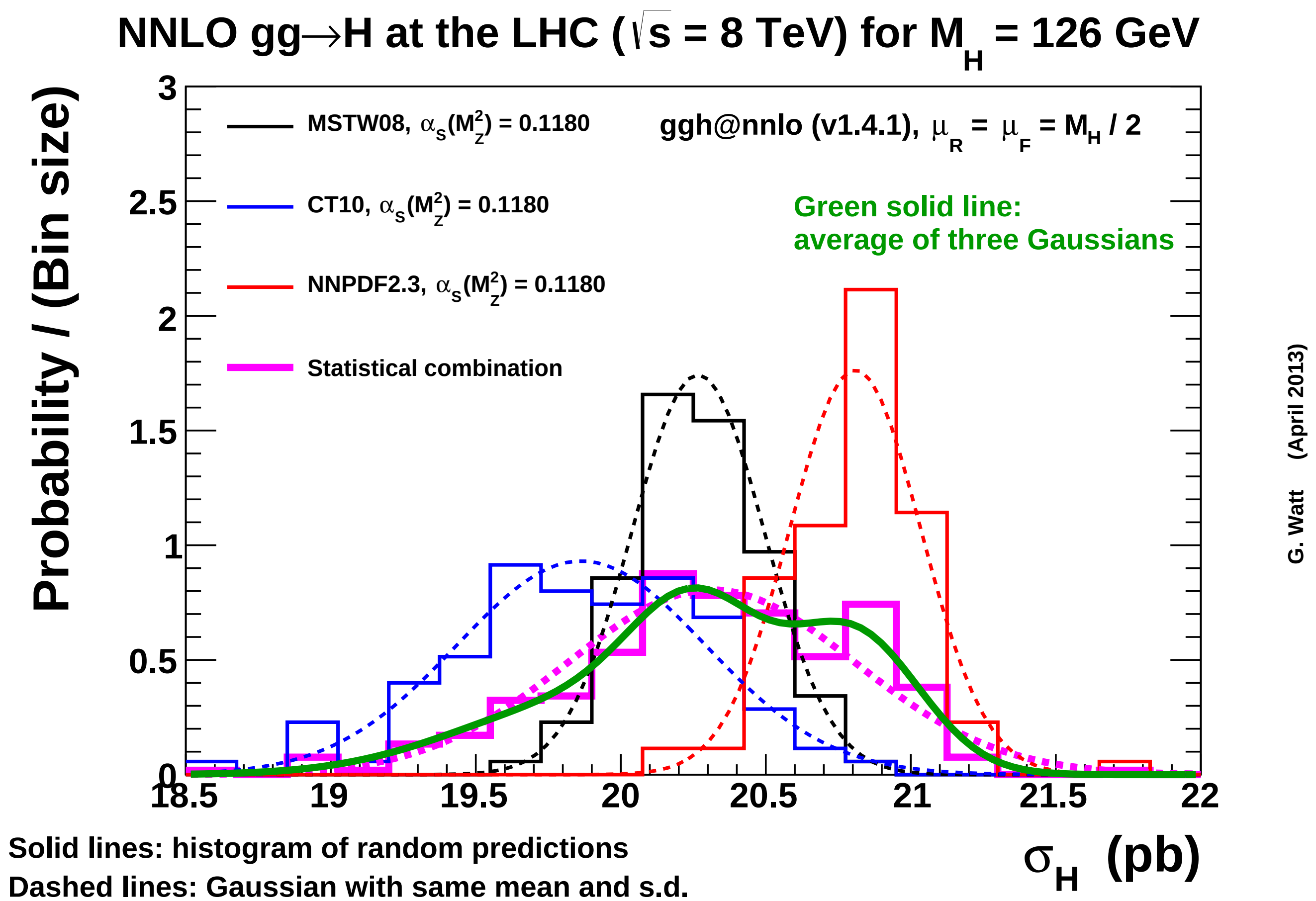
Speed of convergence of prediction for H cross section.



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.



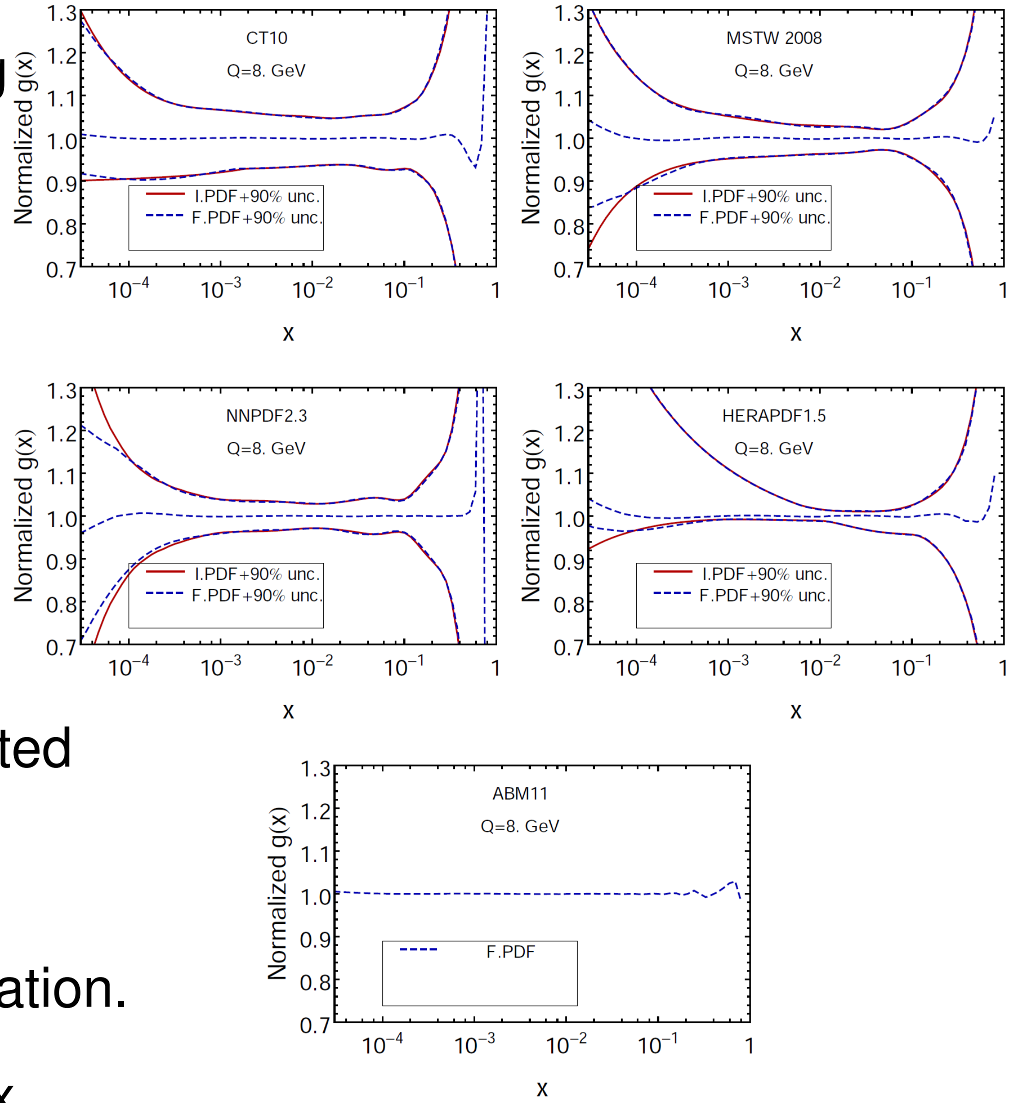
Slightly smaller uncertainty and shifted central value if disagreement between individual predictions. (Plot by G. Watt).

META PDFs (Gao, Nadolsky)

Fit each of the PDFs, including uncertainty sets, to a common functional form at a particular scale.

Check that both central values and uncertainties of original PDFs are reproduced.

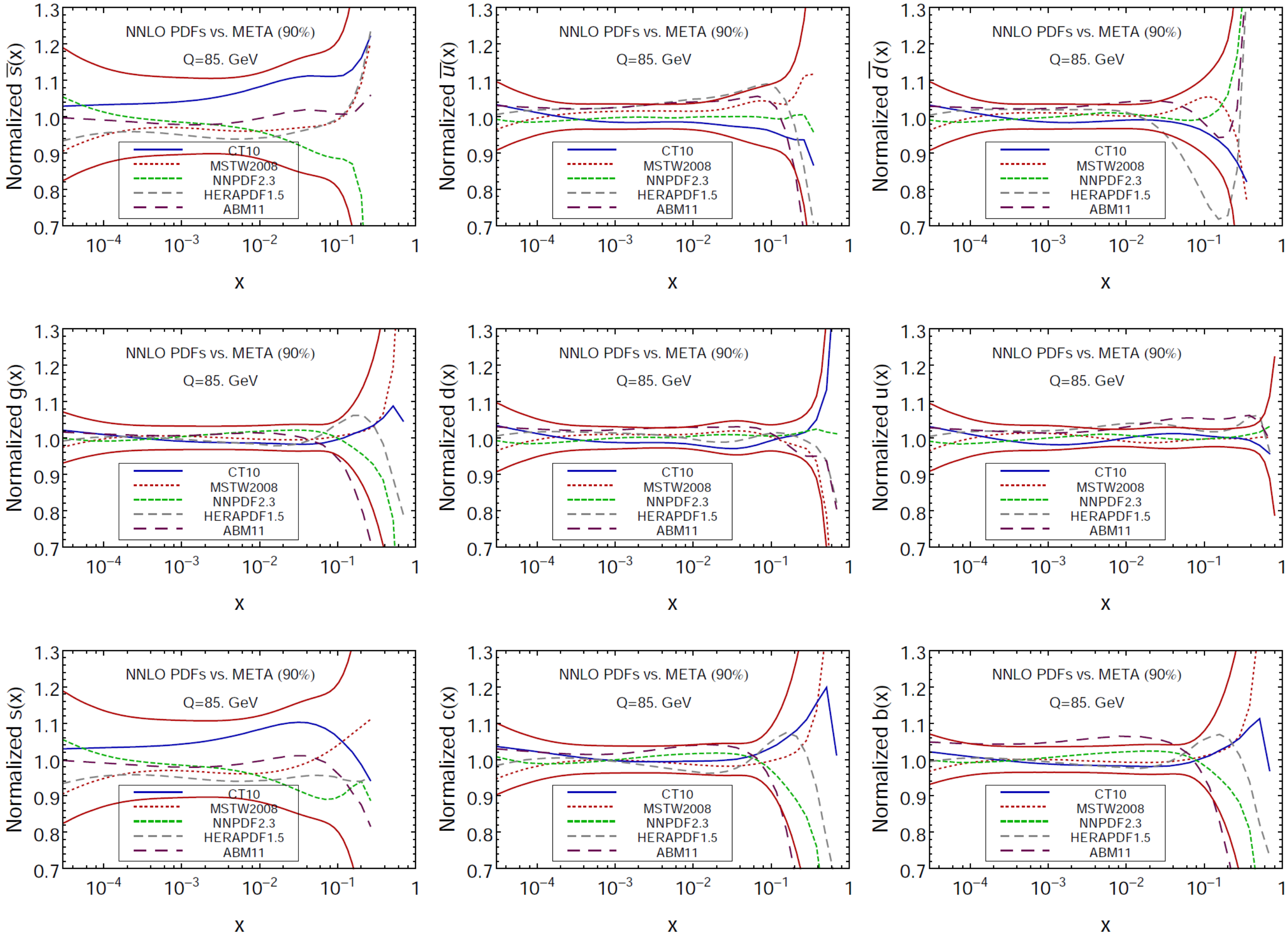
Take collection of equal numbers of CT10, MSTW, and NNPDF randomly distributed sets.



Fit average and uncertainty to same common parameterisation.

Diagonalise Covariance matrix and remove some redundant eigenvectors.

Produce META PDF set with 50 eigenvectors.



Comparison of META PDFS with 90% confidence level uncertainties to central values of a variety of PDF sets with $\alpha_S(M_Z^2) = 0.118$.

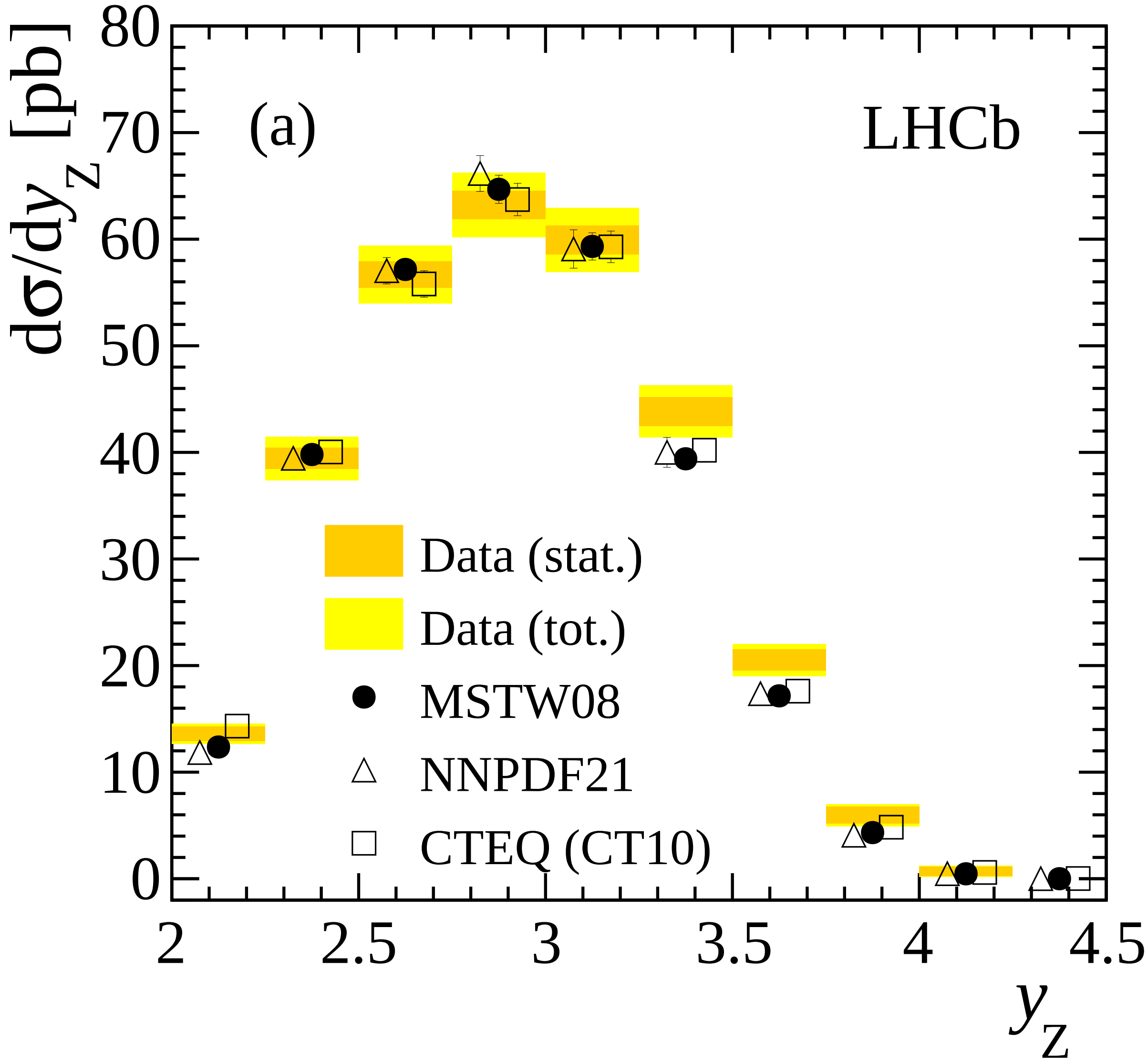
Could be used as more convenient basis for results with combination.

PDF4LHC recommendation - announcement made at meeting on 16th May, de Roeck.

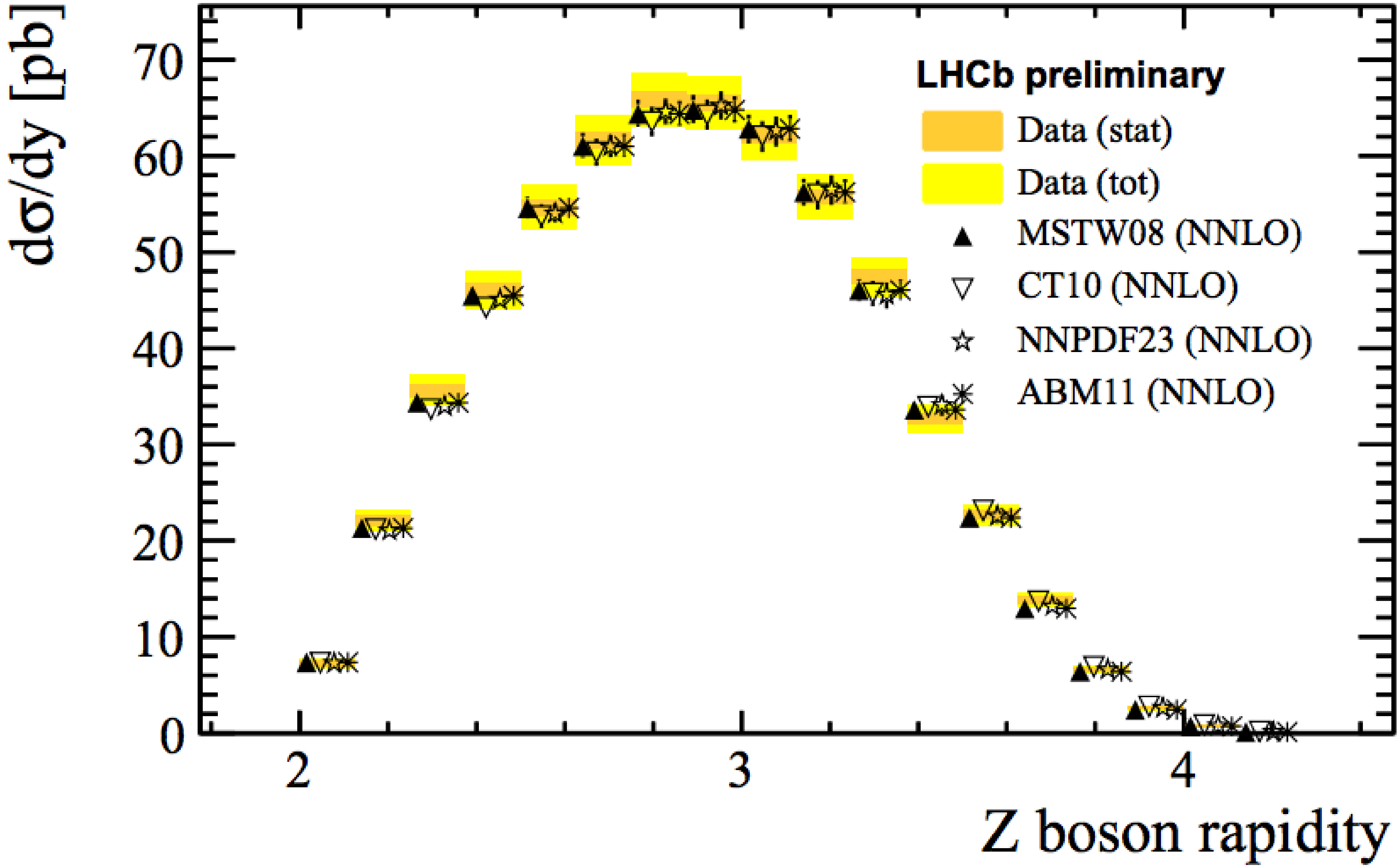
Pre-announcement

- Revisiting the PDF4LHC recommendations: we foresee two steps:
- De facto the published recommendation is used with the original PDFs replaced by the more recent incarnations. We have turned this now into a recommendation on the PDF4LHC web page that can be referred to, as of today. There will be no separate note on it. This should be useful for eg the Higgs cross section group...
- At the same time we plan to move towards using metaPDFs and have an envelope of a number of input PDF family uncertainty bands and prepare for a new recommendation & paper at the time scale of end of the year. Given that in a few months we expect updated PDFs by essentially all the groups, these would naturally by based on these new sets.
- This would become then the recommendation as of 2015 for the upcoming new data analyses. These envelopes can in principle have more sets than what we use now, where sensible.
- Discussion on this at the next PDF4LHC meeting (end of summer)

Back-up



Points near to $y = 3.5$ overshoot predictions in general. Feature not present in prelim. higher luminosity $Z \rightarrow \mu^+\mu^-$ 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 $K = 1.15$.

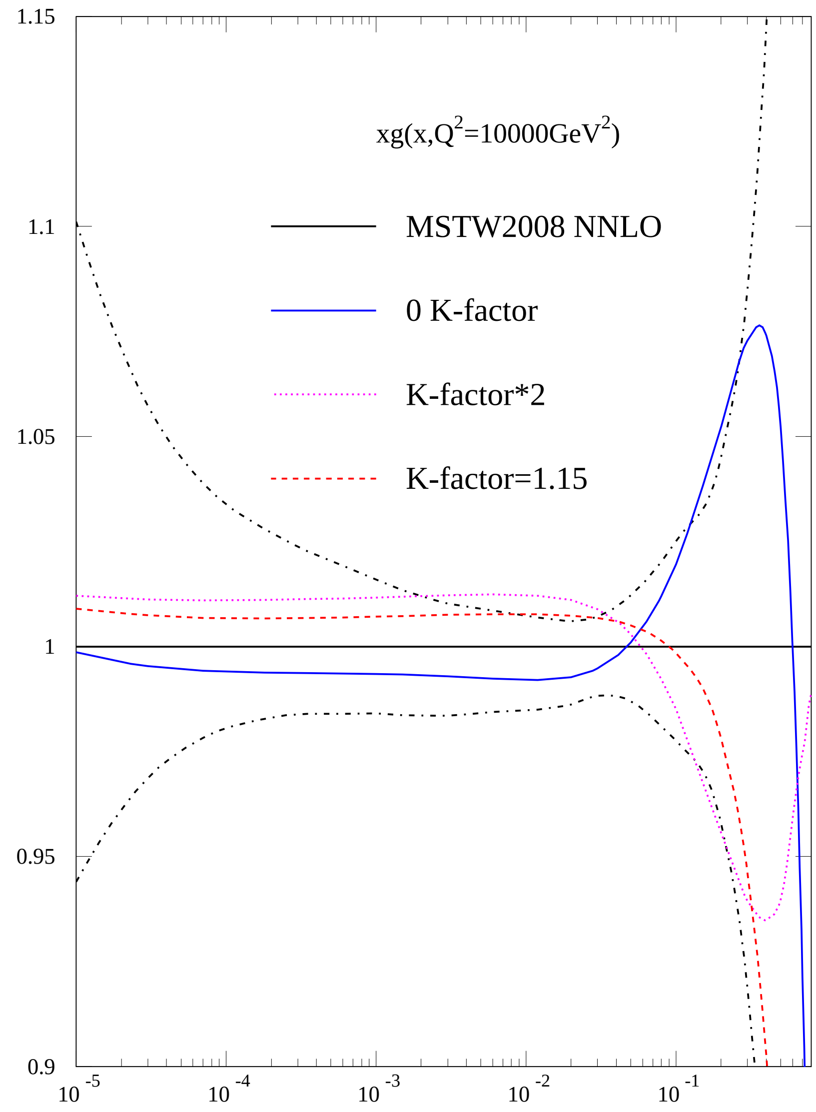
Extreme variations.

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.

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

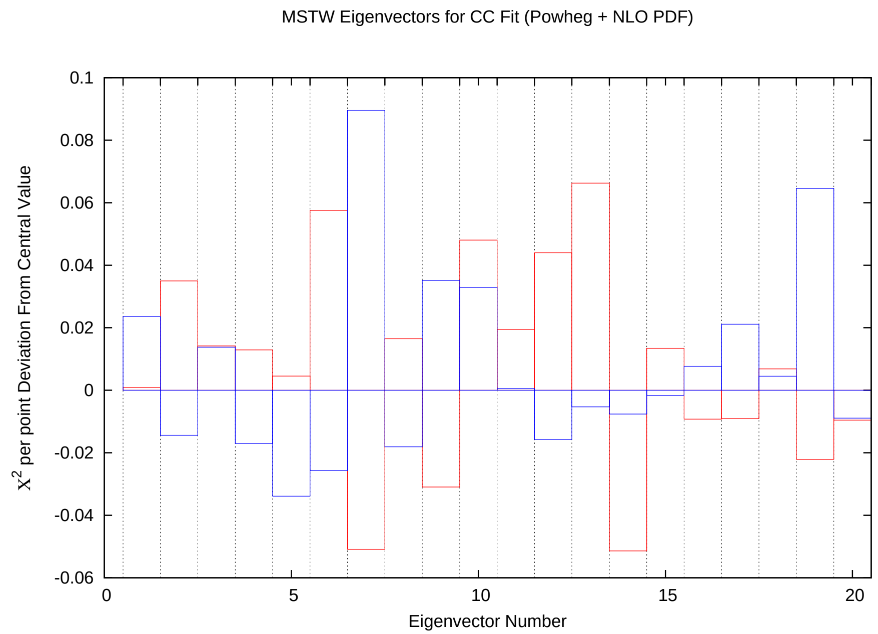
$K * 2 \;\; \alpha_S(M_Z^2) = 0.1159$

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

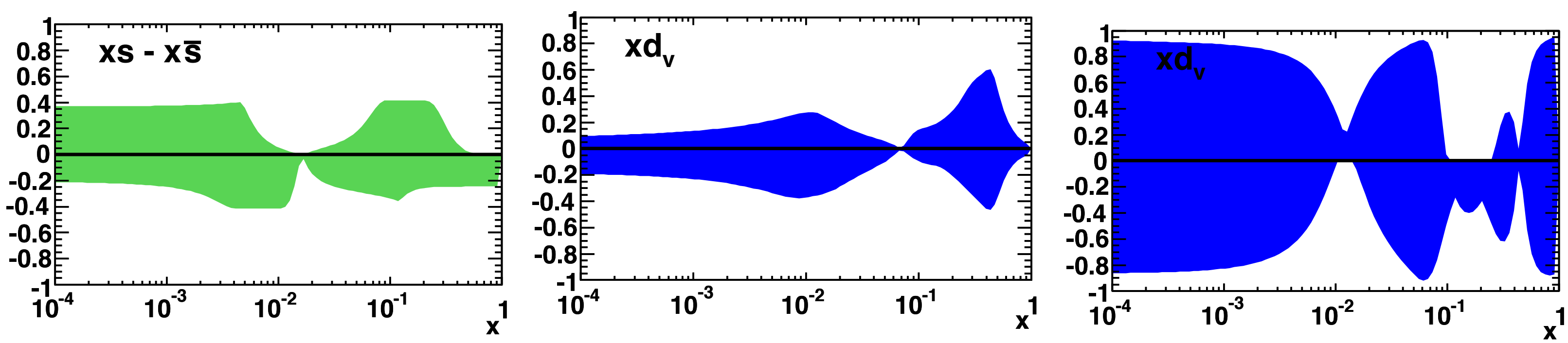


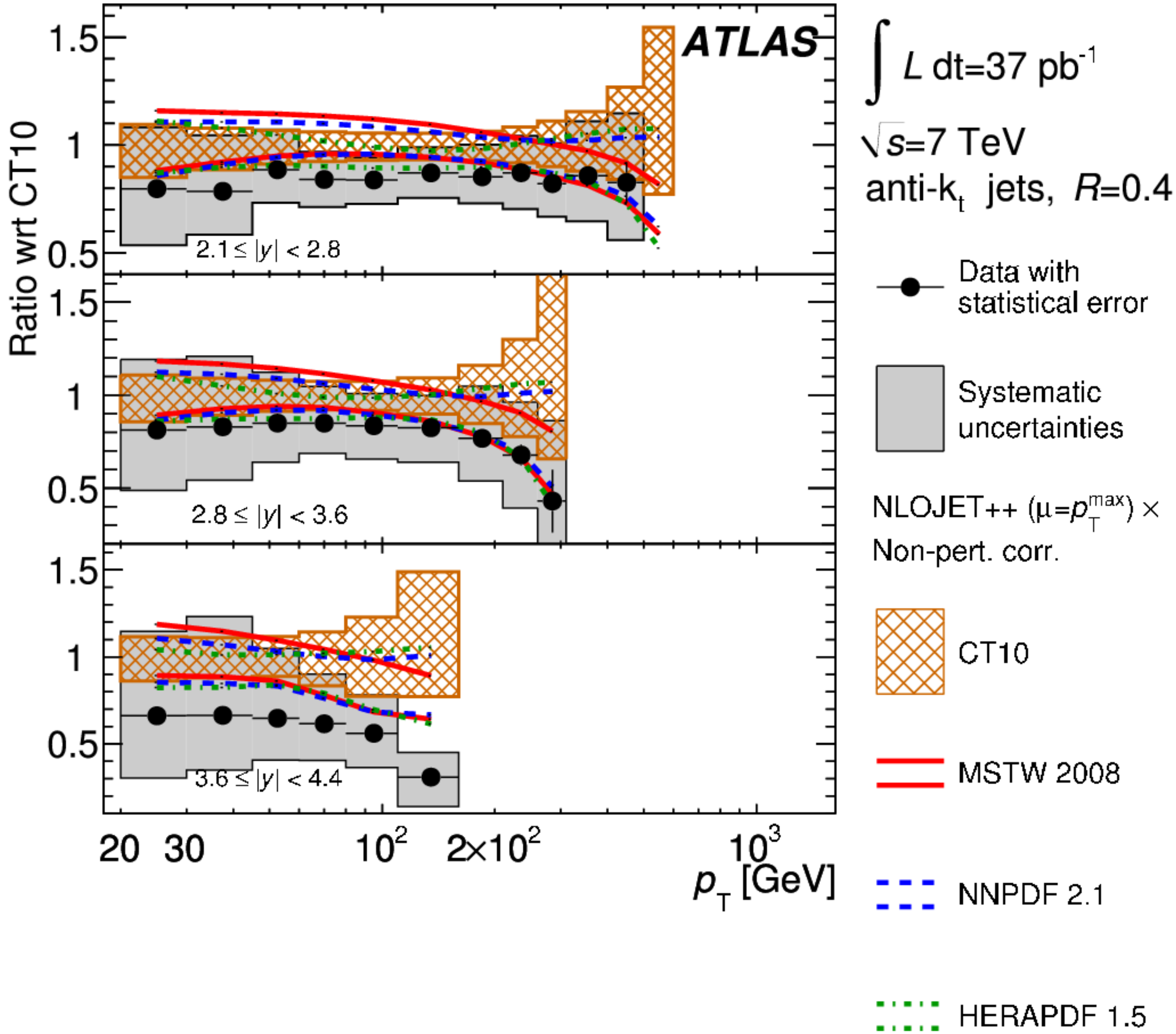
DIS jets - B. Watt

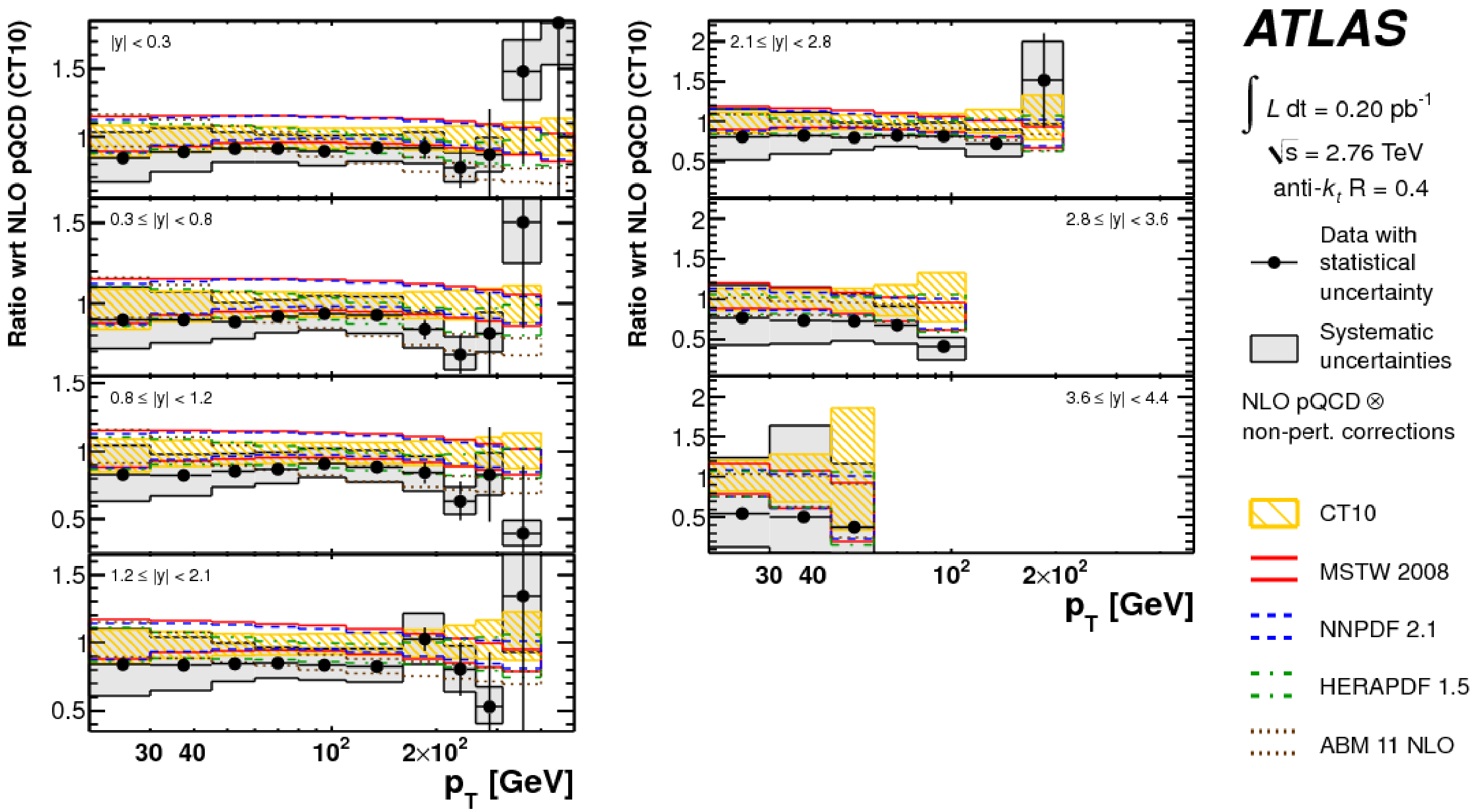
Interesting observation in eigenvector sensitivity to charged current ZEUS jet data usingPOWHEG.

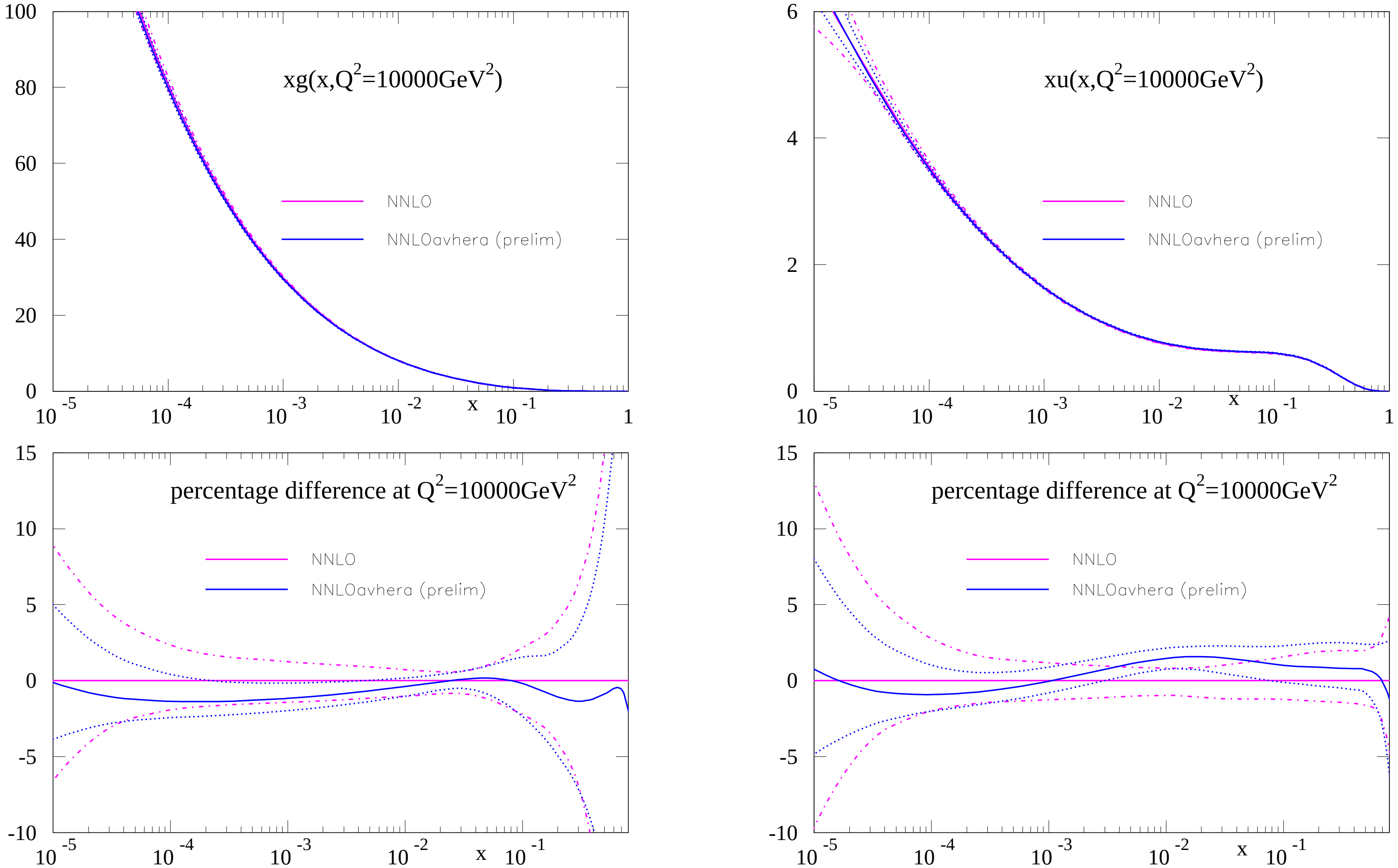


Some sensitivity to $s - \bar{s}$.



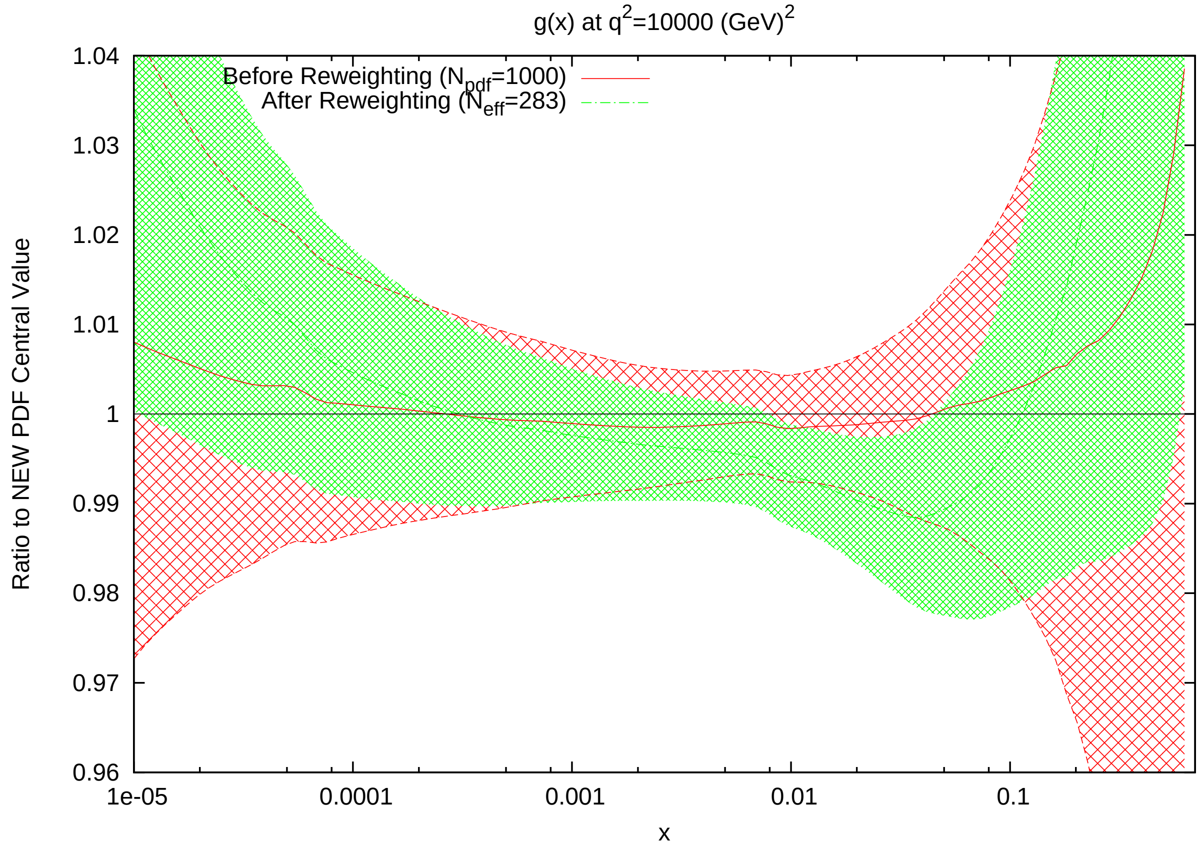

ATLAS
∫ L dt=37 pb-1
√s=7 TeV
anti-kt jets, R=0.4
Data with statistical error
Systematic uncertainties
NLOJET++ (μ=pTmax) × Non-pert. corr.
CT10
MSTW 2008
NNPDF 2.1
HERAPDF 1.5
Ratio wrt CT10
2.1 ≤ |y| < 2.8
2.8 ≤ |y| < 3.6
3.6 ≤ |y| < 4.4
pT [GeV]


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.

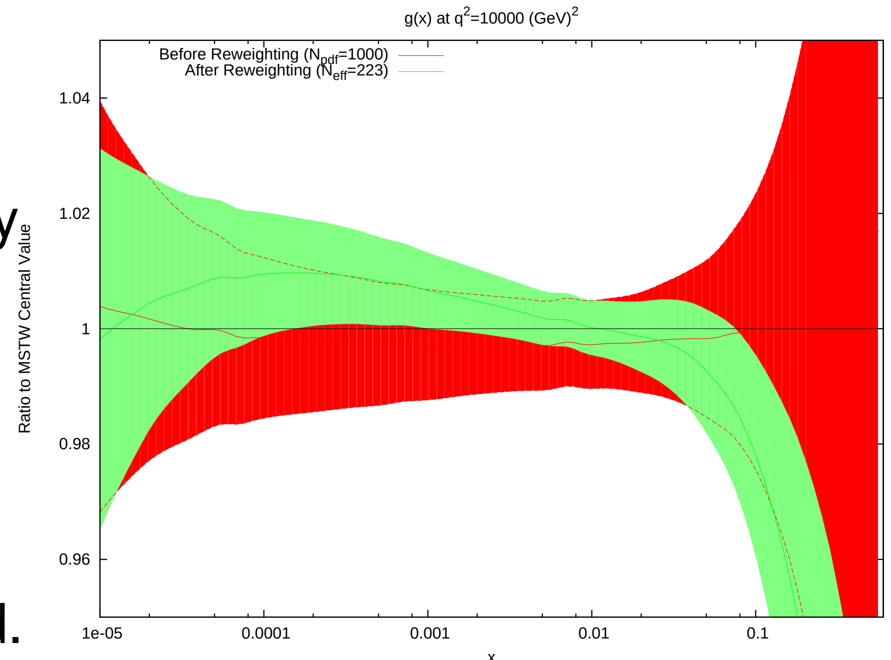
Reflection of different shape of higher order corrections?

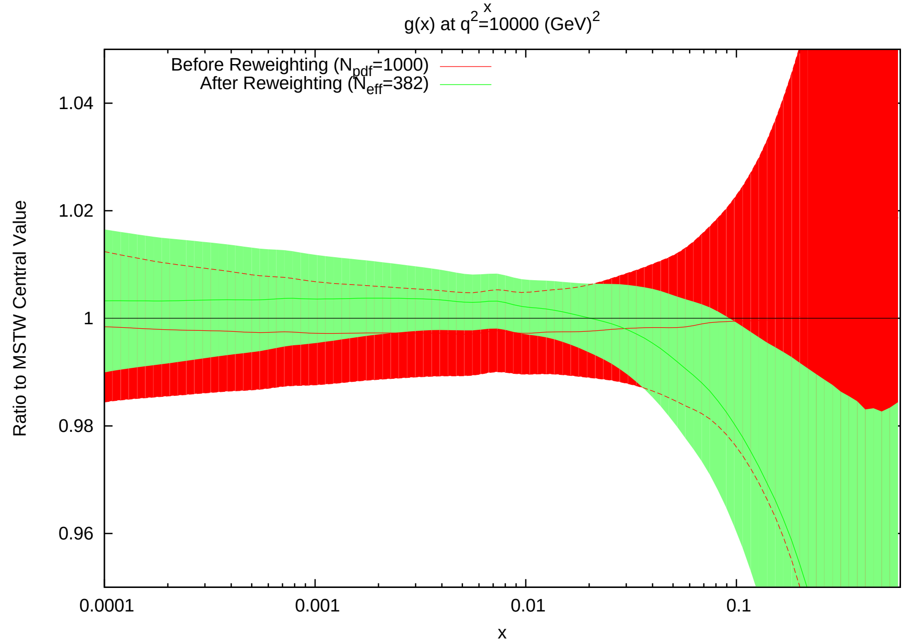
Different conclusions for fits to D0 and ATLAS dijet data, though they are not not necessarily incompatible.

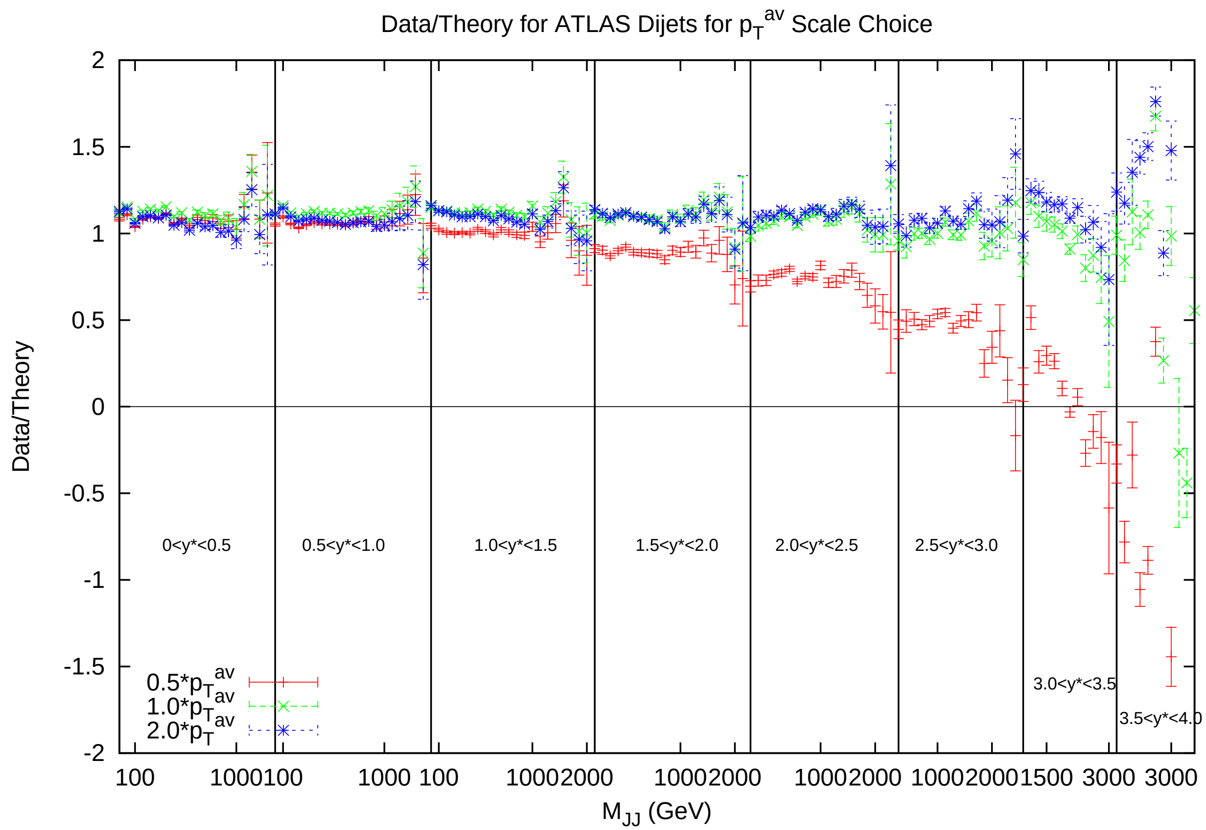
Similar to changes required by LHC inclusive jet data.

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}.



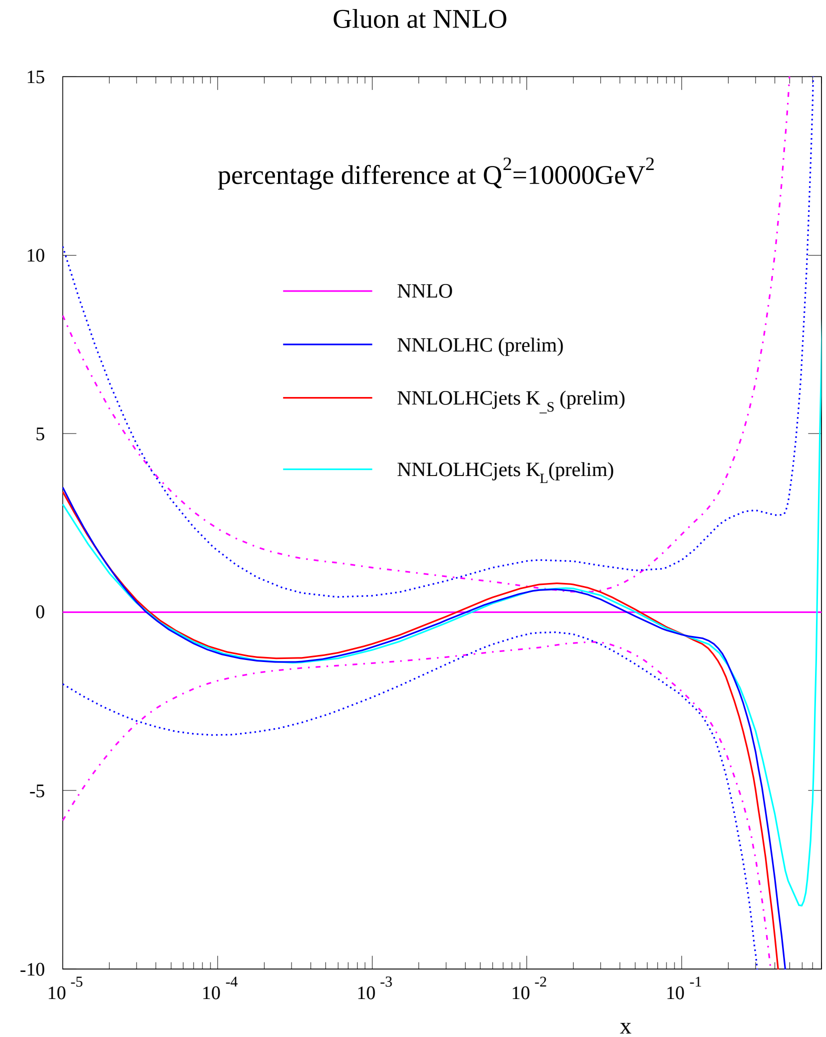




At high rapidity calculations unstable for scales equal to relativity low multiples of p_T.

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 $\chi^2 = 119/116 \rightarrow 106/116$ and CMS $\chi^2 = 138/133 \rightarrow 139/133$.

For the "larger" K-factor ATLAS $\chi^2 = 128/116 \rightarrow 118/116$ and CMS $\chi^2 = 139/133 \rightarrow 141/133$.