Updates of PDFs for the 2nd LHC run

Lucian Harland-Lang, IPPP Durham

Diffraction 2014
Primošten, Croatia, September 11 2014

In collaboration with Alan Martin, Patrick Motylinski and Robert Thorne

and thanks to Ben Watt, Graeme Watt and James Stirling

Outline

In this talk I will present results on continuing updates in PDFs within the MSTW framework. A new set is very close to being finalized, with no significant changes expected to the PDFs shown here. Updates:

- Changes in theoretical procedures (updated parameterisation, error treatment, nuclear corrections...).

- Inclusion of a variety of new data sets, including the most up-to-date LHC data:
    - $W^{\pm}$, Z, $t\bar{t}$
    - HERA updates
    - Jet data
    - DY@CMS double differential
    - Jet data @NNLO (not included in fit)

Theory changes - extended parameterization

- Continue to use extended parameterisation with Chebyshev polynomials as in recent MSTWCPdeut study (Eur.Phys.J. C73 (2013) 2318).

- For valence and sea quarks, instead of taking

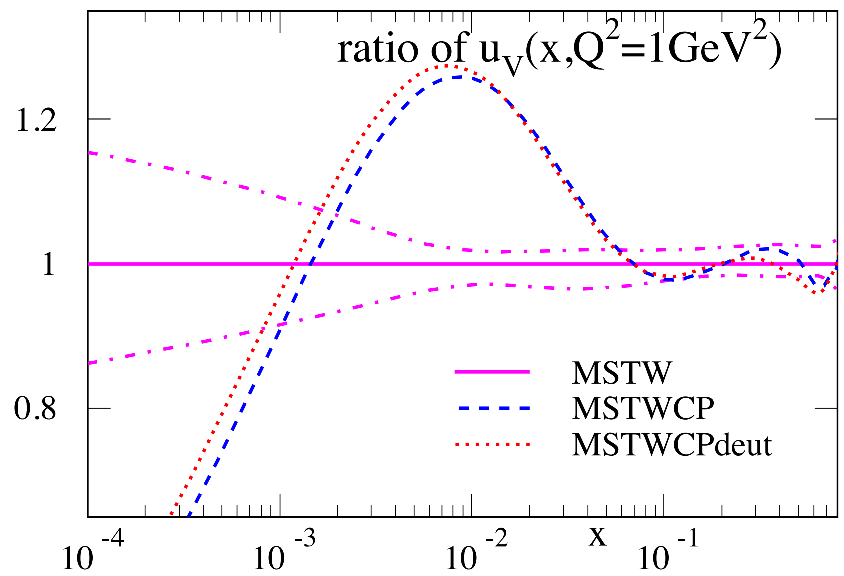
$$xf(x,Q_0^2) = A(1-x)^\eta x^\delta (1+\epsilon x^{0.5}+\gamma x)$$

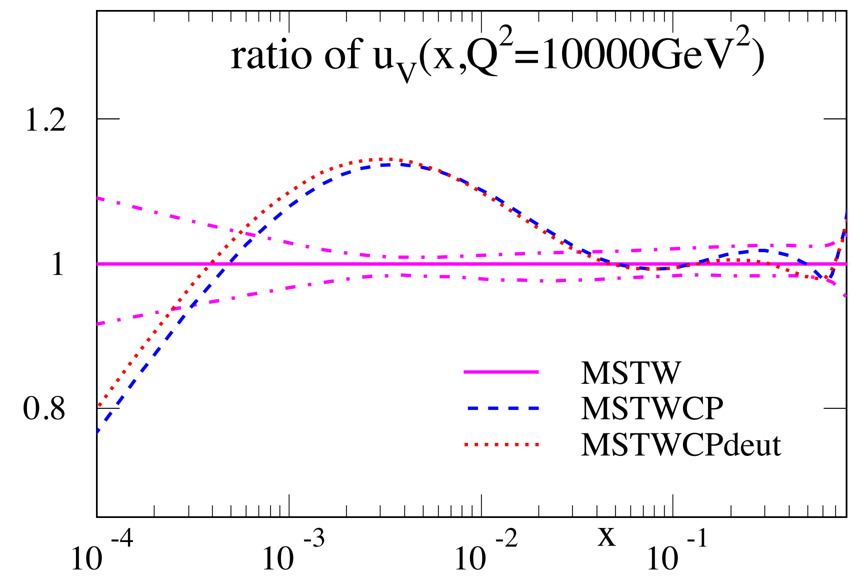
take

$$xf(x,Q_0^2) = A(1-x)^\eta x^\delta \left(1+\sum_{i=1}^{n} a_i T_i(y(x))\right) \qquad y(x) = 1-2\sqrt{x}$$

where T_i are Chebyshev polynomials - convenient choice of basis for interpolating polynomial. By considering different n can perform systematic study.

- Taking $n=4$ find significant improvement in global $\Delta\chi^2$ and change in u_V for $x \lesssim 0.03$.





Deuteron correction - fit

- In order to separate u and d at moderate to large x use DIS data on deuteron targets.

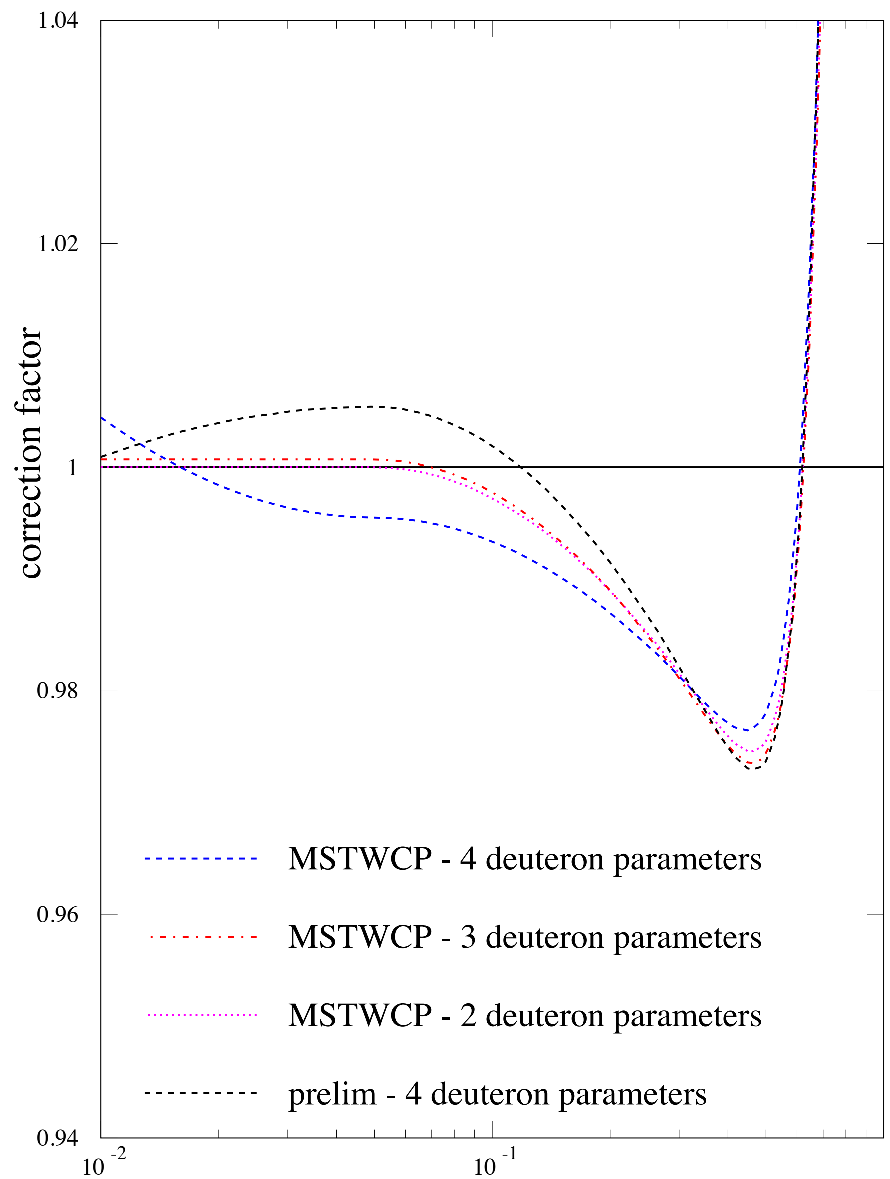
  $\longrightarrow$ Need to include nuclear corrections.

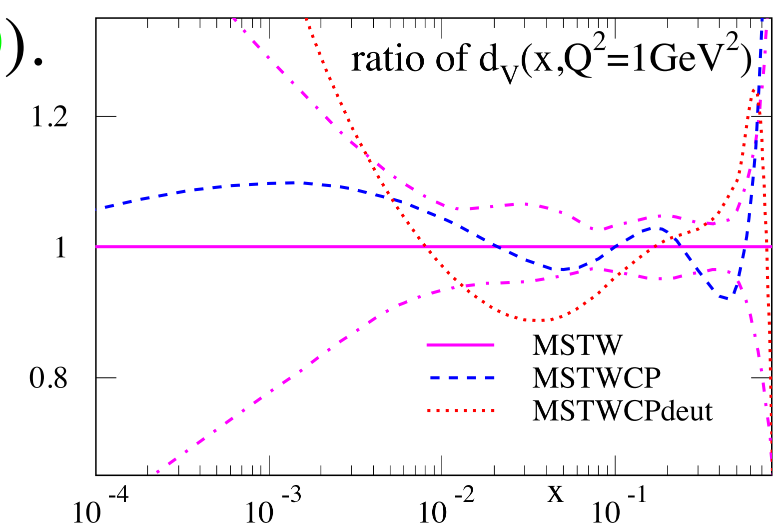
$$F^d(x,Q^2) = \mathrm{c}(x)(F^p(x,Q^2) + F^n(x,Q^2))/2$$

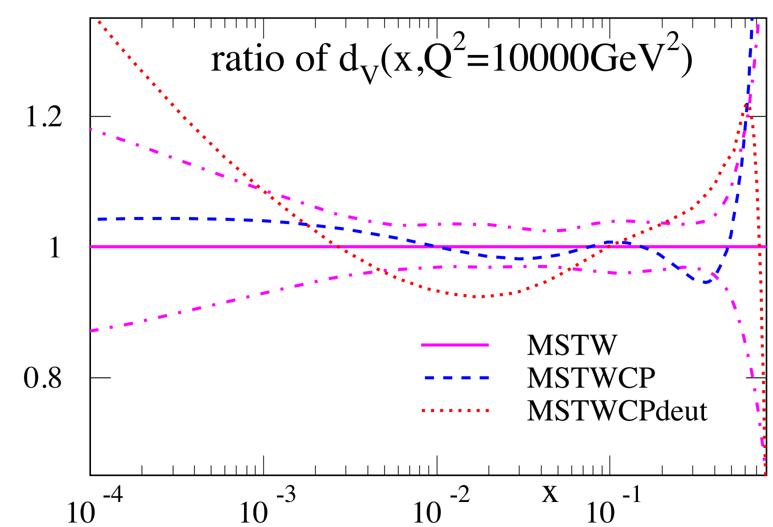
- Parameterise $c(x)$ and allow to vary with no penalty in fit.

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

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







Treatment of errors

- Systematic errors generally multiplicative (uncertainty $\propto$ measured value, as in e.g. overall normalization uncertainty $\longrightarrow$ percentage error).

- 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 normalization uncertainty.

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

- And so:

both data and error scaled

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

incorrectly treating multiplicative errors as additive will bias results.

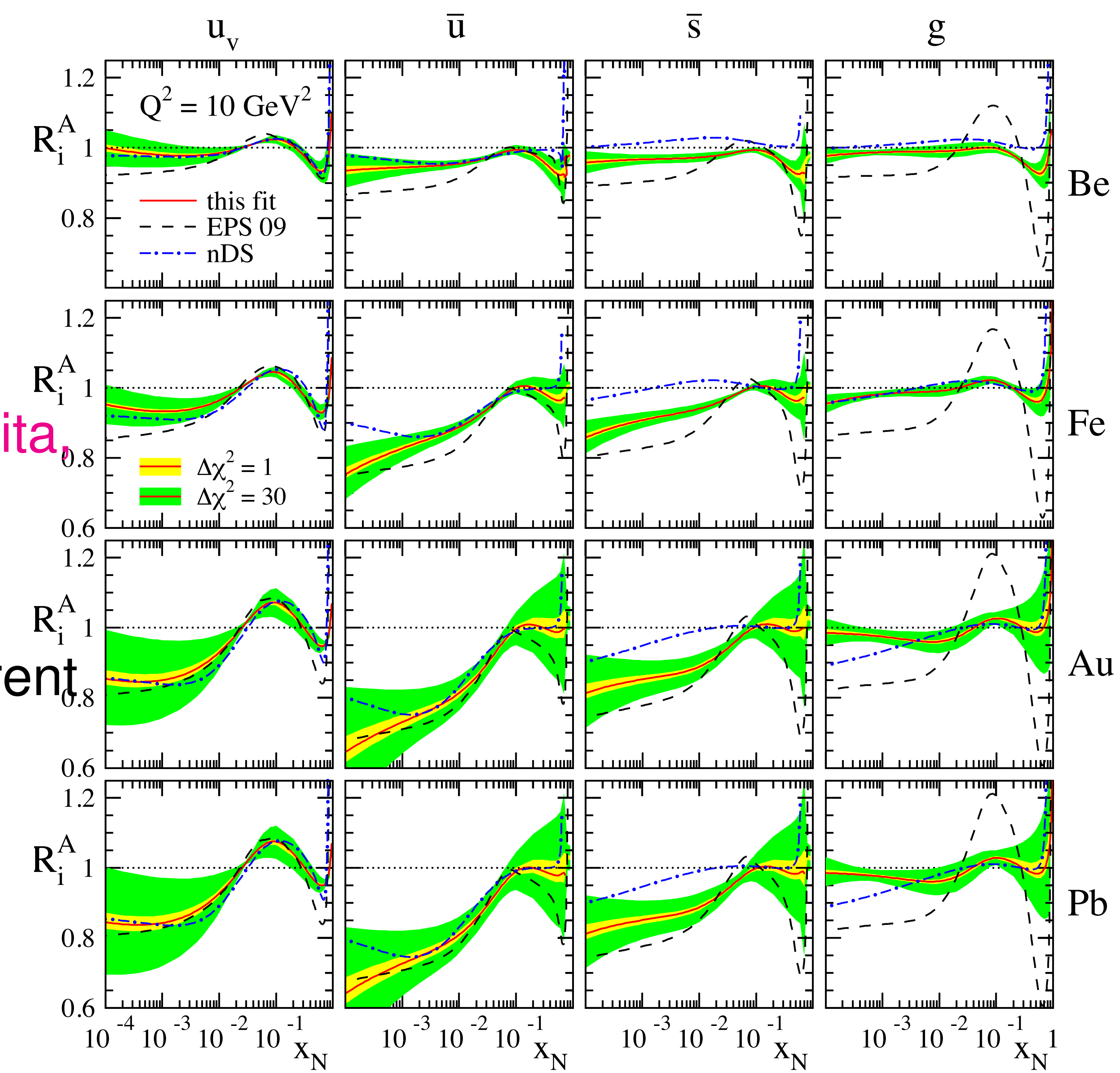
Nuclear corrections

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).

Other changes in theoretical procedures

• 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.

• Branching ratio $B_\mu \equiv B(D \to \mu)$ (for strangeness): 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 $B_\mu = 0.082 - 0.090 \pm 15\%$ with NNLO at lower end.

Changes in data sets - HERA

- HERA Run-I neutral and charged current data from H1 and ZEUS replaced with combined data set with full treatment of correlated errors. Fit to data very good. Slightly better fit at NNLO.
- HERA combined data on $F_2^c(x,Q^2)$ included. Fit quality $\sim 60-65$ for 52 points.
- All direct published HERA $F_L(x,Q^2)$ measurements included. Undershoot data a little at lower Q^2, but χ^2 not much more than one per point.
- Separate Run-II H1 and ZEUS data not included yet. Will wait for Run-II combination.

Changes in data sets - Tevatron

- New(er) Tevatron data sets included:
  - CDF W-asymmetry data
  - D0 electron asymmetry data ($p_\perp > 25\,\mathrm{GeV}$ based on $0.75\,\mathrm{fb}^{-1}$)
  - New D0 muon asymmetry data ($p_\perp > 25\,\mathrm{GeV}$ based on $7.3\,\mathrm{fb}^{-1}$)

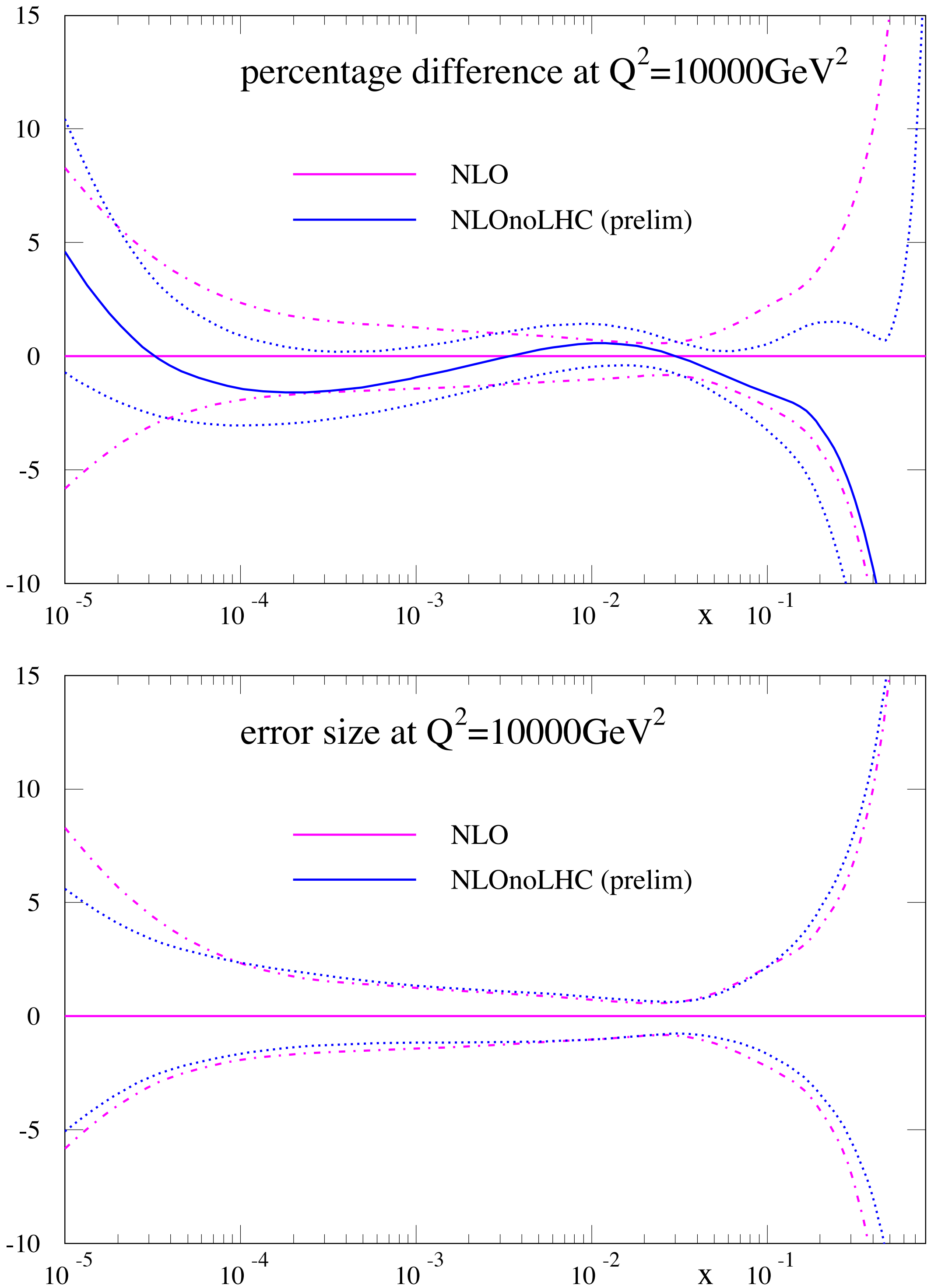
- Include final numbers for CDF Z-rapidity data - final numbers changed after MSTW2008 fit. (Also include very small photon contribution in theory.) Very little change.

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

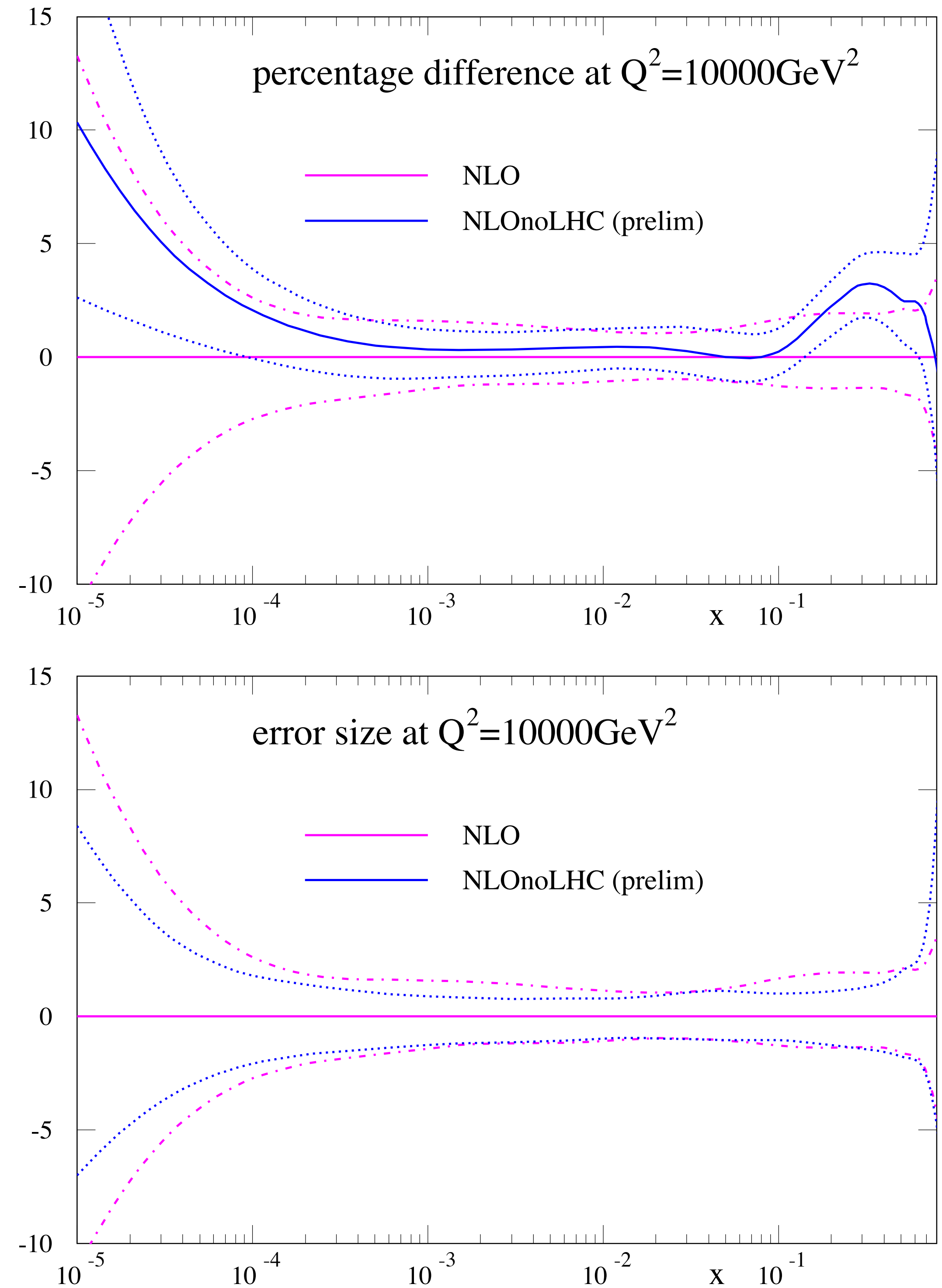
- At NLO $\alpha_S(M_Z^2) = 0.1199$ from 0.1202 and at NNLO $\alpha_S(M_Z^2) = 0.1180$ from 0.1171.

Changes in NLO PDFs - pre-LHC

Gluon at NLO



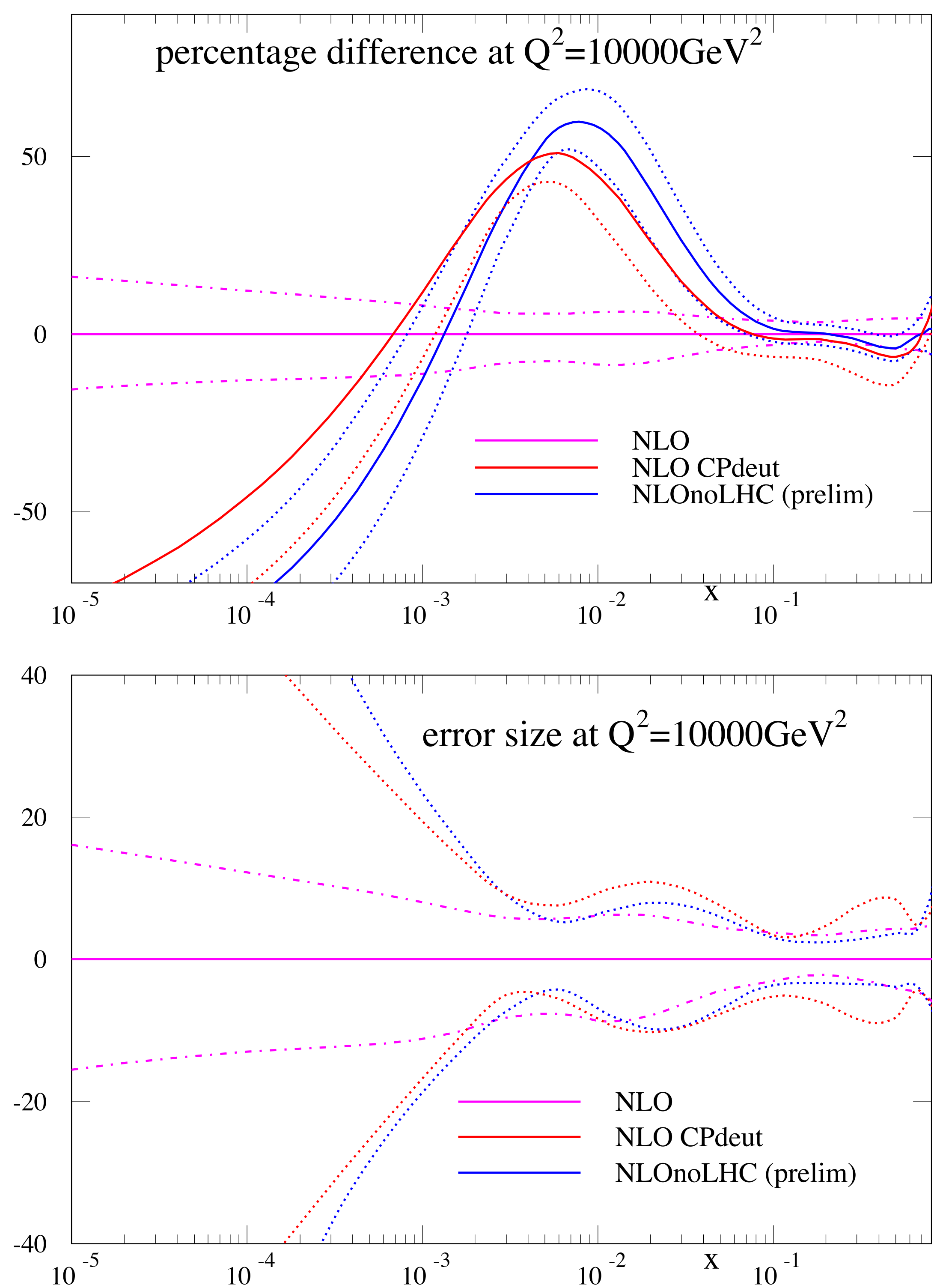
Light Quarks at NLO



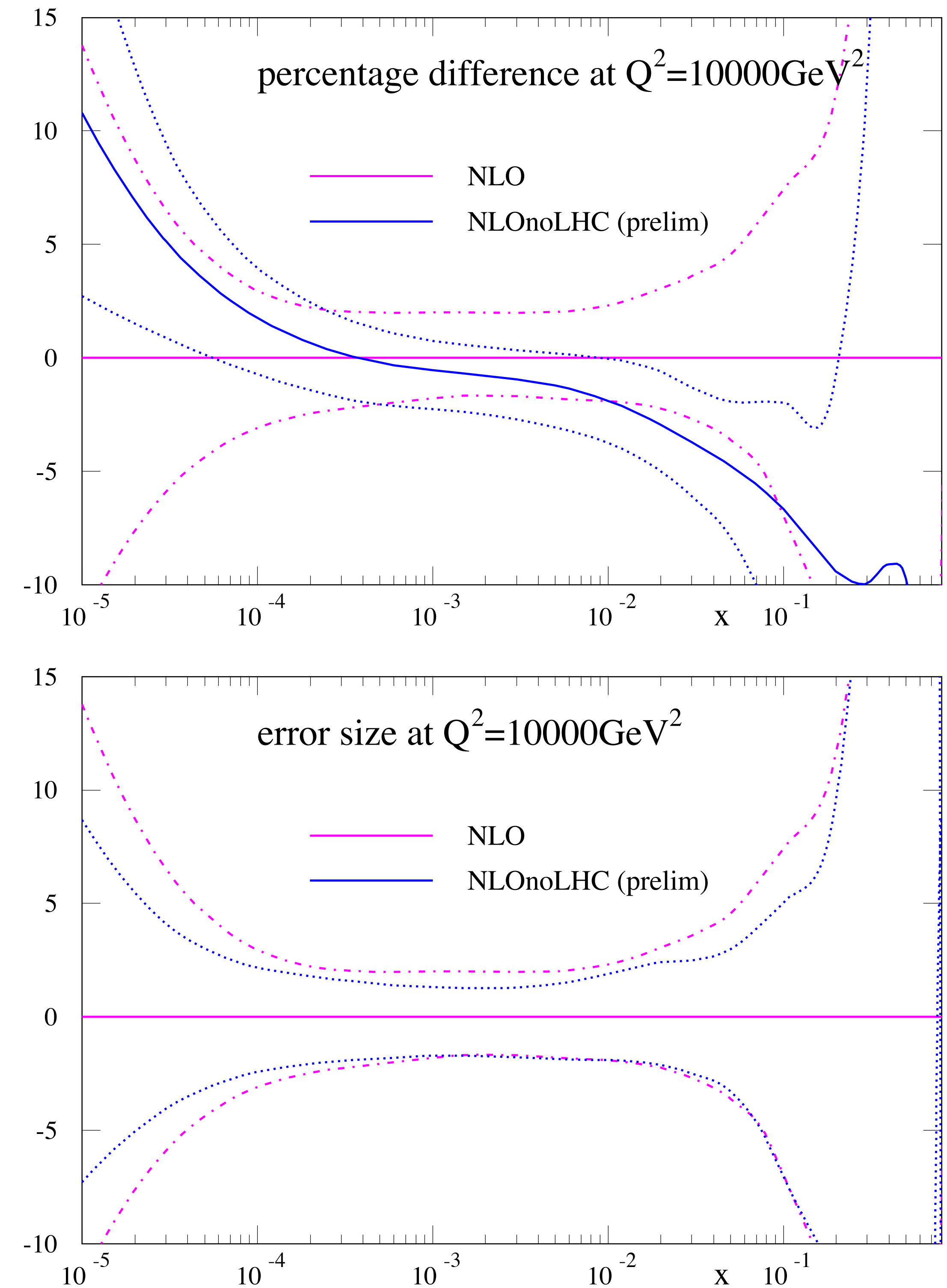
Increase in d at high x (mainly due to deuteron corr.). Overall small to moderate changes.

Changes in NLO PDFs - pre-LHC

$x(u_V - d_V)$ at NLO



Strange at NLO



Note large changes in $u_V - d_V$ and strange distributions.

$\longrightarrow$ Due to extended parameterisation/deuteron corr.

Change in branching ratio for dimuon data not included yet

Inclusion of LHC data

- Work done with R.S. Thorne and P. Motylinksi using `FastNLO`, `APPLGrid`, `MCFM` and `DYNNLO/FEWZ`. Allows direct inclusion of data at NLO into fit. At NNLO still rely on K-factor approximation:

$$K^{\rm NNLO}(M,y) = 1 + \left(\frac{\alpha_s(M)}{\pi}\right) D(M,y) + \left(\frac{\alpha_s(M)}{\pi}\right)^2 E(M,y)$$

- ATLAS $W^\pm$, Z rapidity data now included:
  - Before inclusion $\chi^2 \sim 1.6$ per point at NLO, $\chi^2 \sim 2$ per point at NNLO.
  - Inclusion results in some extra improvement at NLO, $\chi^2 \sim 1.3$ with strongest pull on gluon PDF.
  - Also goes to $\chi^2 \sim 1.3$ at NNLO. The most obvious change is in the strange quark (balance of W and Z production depends on strange).

Inclusion of LHC data - W, Z

- ATLAS and CMS $W^+ - W^-$ asymmetry data both included, and no longer an issue at all (c.f. extended parameterisation and deuteron corrections). Fit slightly better at NLO.

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

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

- CMS Z double differential (rapidity distributions for $20 < M_{ll} < 1500\,\mathrm{GeV}$ divided into six bins) measurement included. Extends down to low mass: NNLO much better fit than NLO at lowest mass $\sim 20 - 45\,\mathrm{GeV}$ (more later).

Inclusion of LHC data - $t\bar{t}$

- Include $t\bar{t}$ data into fit:
  - $\sigma_{t\bar{t}}$ from the Tevatron (combined CDF and D0 cross section measurement).
  - 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.

Inclusion of LHC data - jets

- At NLO, LHC jet data is included in fit:
  - CMS together with ATLAS 7 TeV + 2.76 TeV data.
  - Before fitting, the ATLAS 7 TeV + 2.76 TeV $\chi^2 = 112/114$ and CMS $\chi^2 = 186/133$ - comparable to the best PDFs of other groups.
  - Simultaneous fit of CMS data together with ATLAS 7 TeV + 2.76 TeV gives some improvement for CMS, and a small amount for ATLAS. The experiments seem extremely compatible.
- At NLO, final extracted $\alpha_S(M_Z^2) = 0.1193$
- LHC jets not included at NNLO - more later.

Fit quality for LHC data - NLO

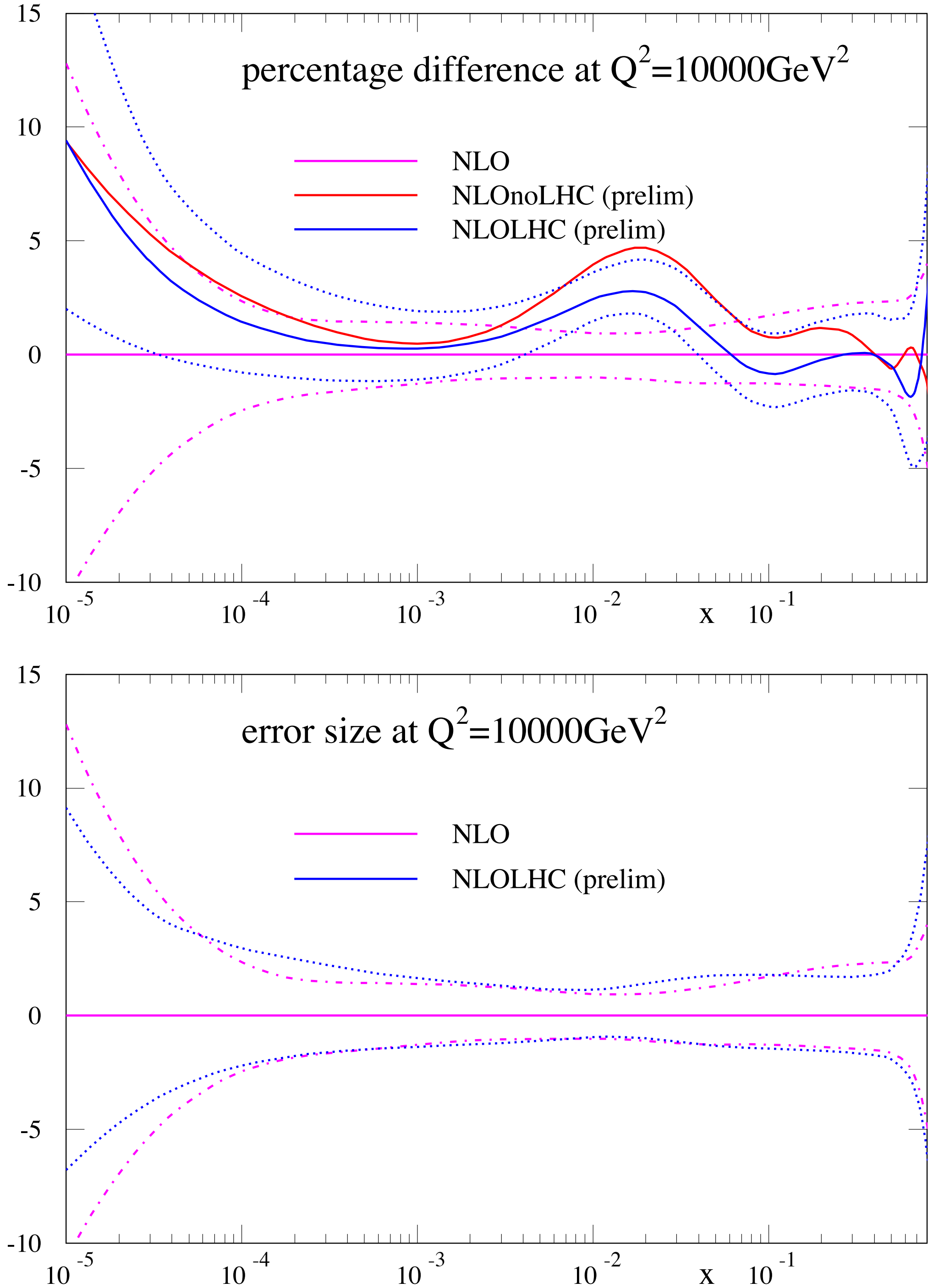
data set	N_{pts}	CPdeut	no LHC	prelim
ATLAS jets (2.76TeV+7TeV)	116	107	107	106
CMS jets (7TeV)	133	140	143	138
ATLAS W^+, W^-, Z	30	47	44	39
CMS W asymm $p_T > 35\mathrm{GeV}$	11	9	16	7
CMS asymm $p_T > 25\mathrm{GeV}, 30\mathrm{GeV}$	24	9	17	7
LHCb $Z \to e^+e^-$	9	13	13	13
LHCb W asymm $p_T > 20\mathrm{GeV}$	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

- W, Z data constrain gluon, as does $\sigma_{t\bar{t}}$.
- CMS W asymmetry data constrains some flavour decomposition.
- CMS double differential and ATLAS high mass DY have little impact on PDFs.

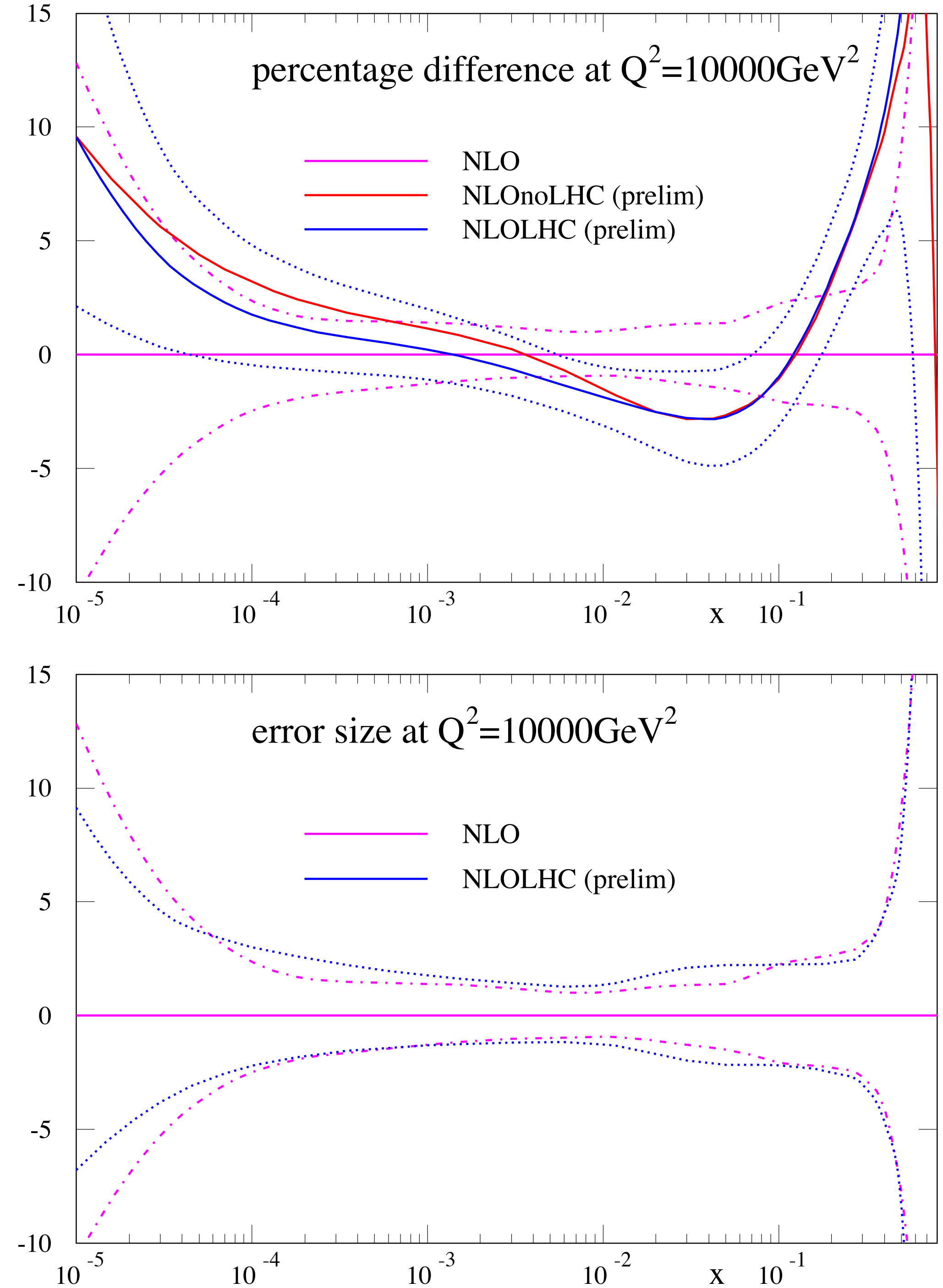
However fit very poor at NLO in lowest mass bins (more later).

Change in NLO PDFs

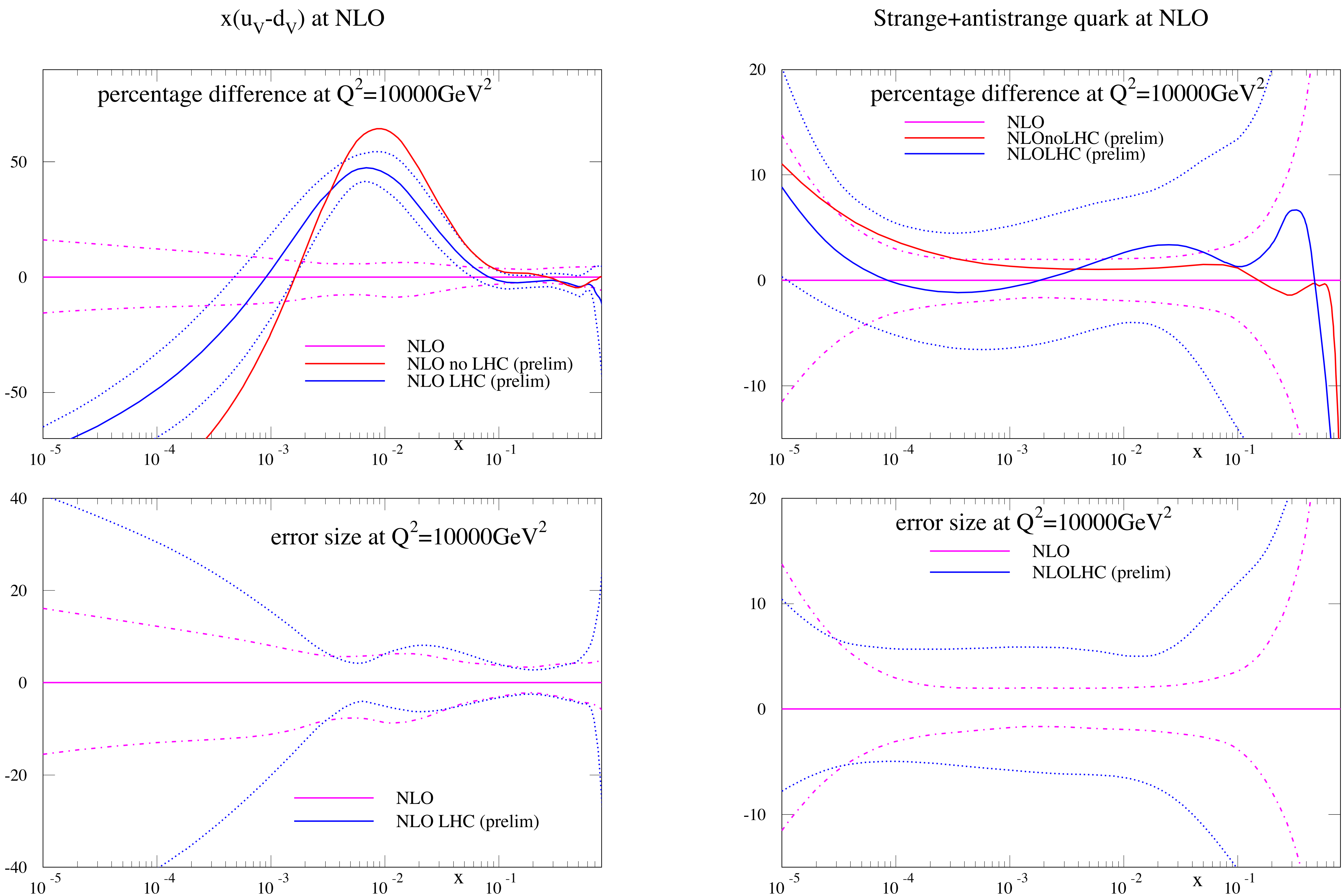
Up quark at NLO



Down quark at NLO



Includes theoretical updates and LHC data (not completely final version)

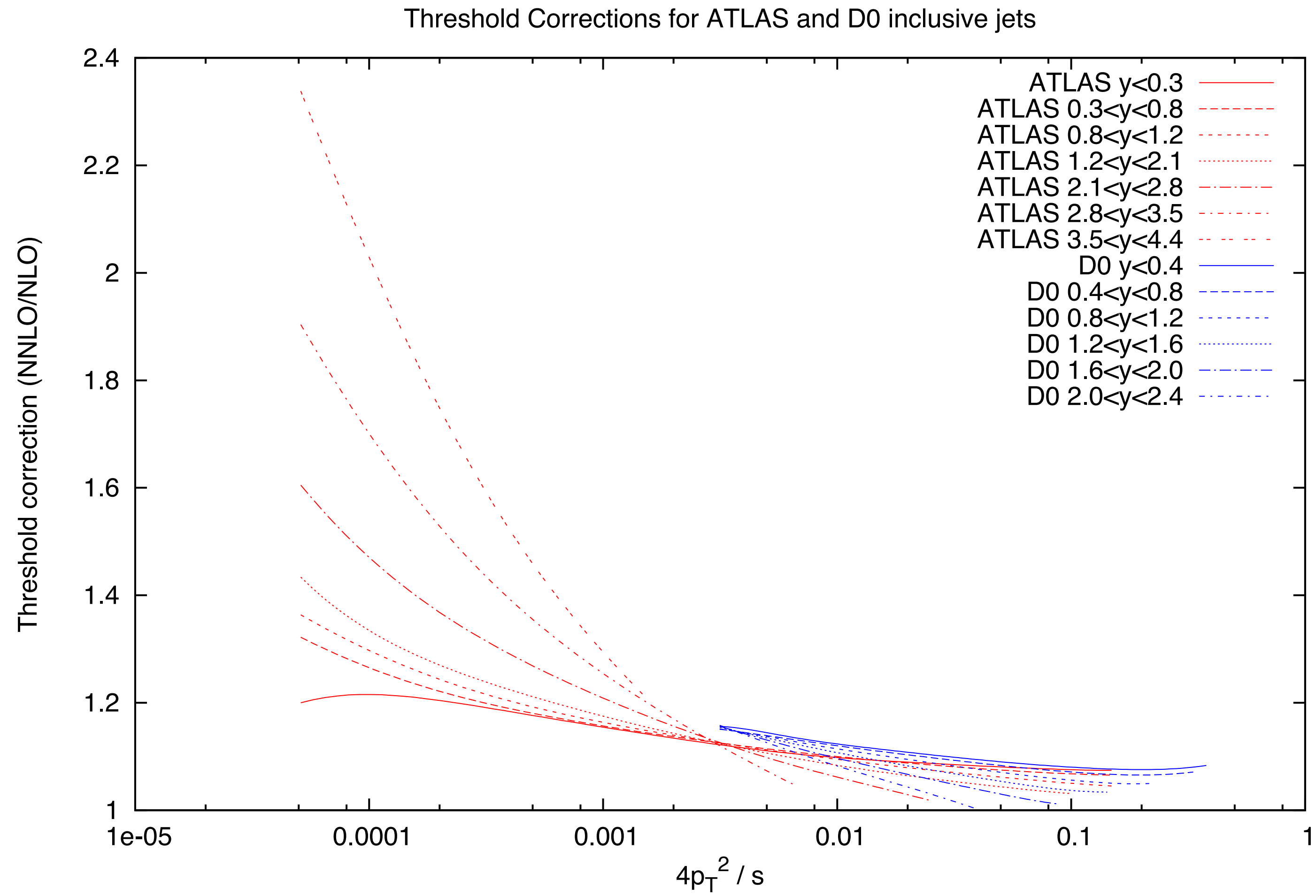


Much expanded $s + \overline{s}$ uncertainty is clear (error on B_μ).

NNLO: LHC jet data?

$$\frac{2p_\perp}{\sqrt{s}} \sim 1$$

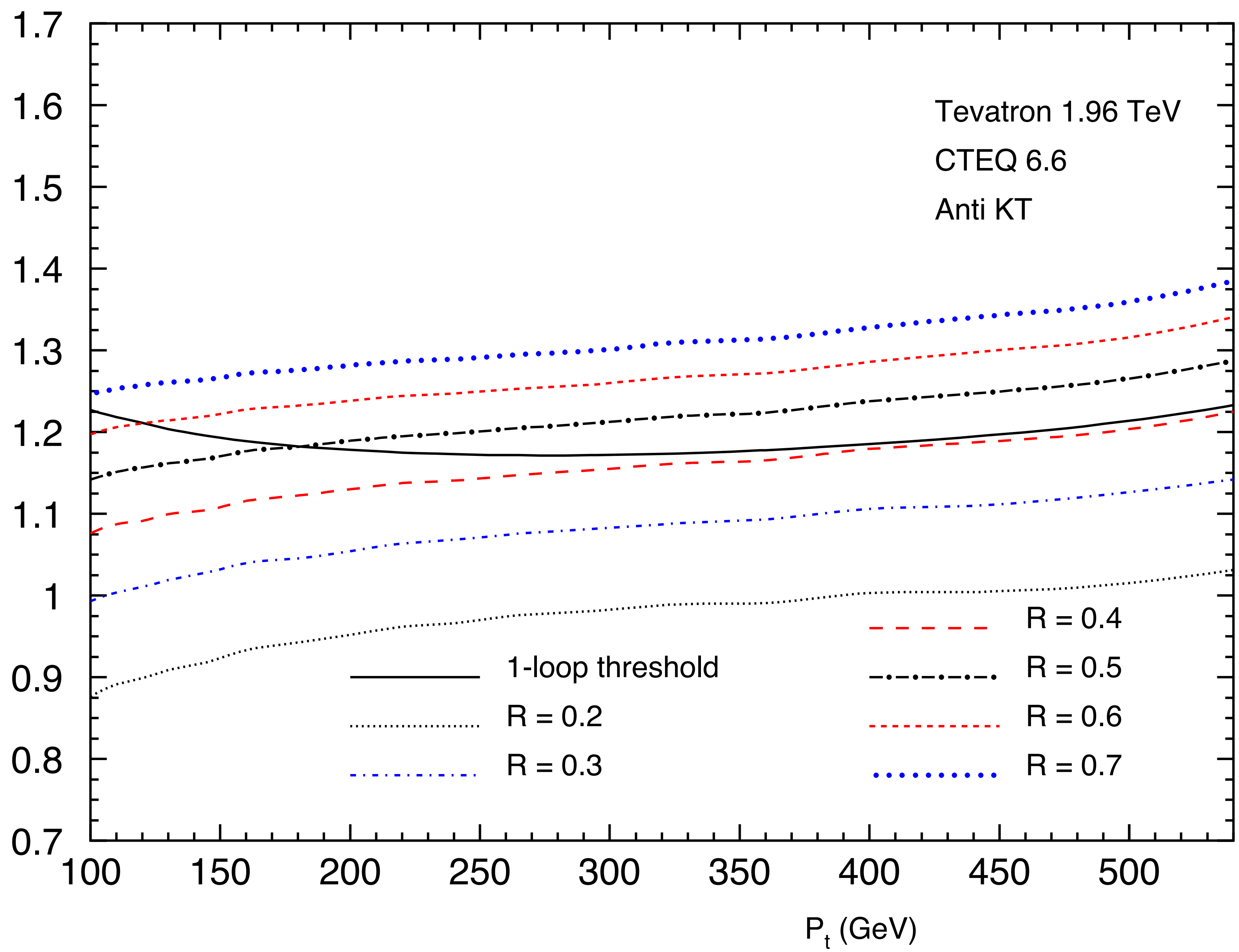
- For Tevatron data use approximate "threshold" corrections (Kidonakis and Owens), $\sim 10\%$ positive correction.
- LHC corrections very similar in the highish x region (as probed at the Tevatron), however these blow up for low x, i.e. far from threshold, which is probed at the LHC.



B.J.A. Watt, P. Motylinski and R.S. Thorne, arXiv: 1311.5703

$\longrightarrow$ need full NNLO calculation. Enormous project for full NNLO calculation (Gehrmann-de-Ridder, Gehrmann, Glover and Pires) nearing completion. Some channels calculated, and some indications of the full form of the correction.

NNLO threshold corrections- recent results



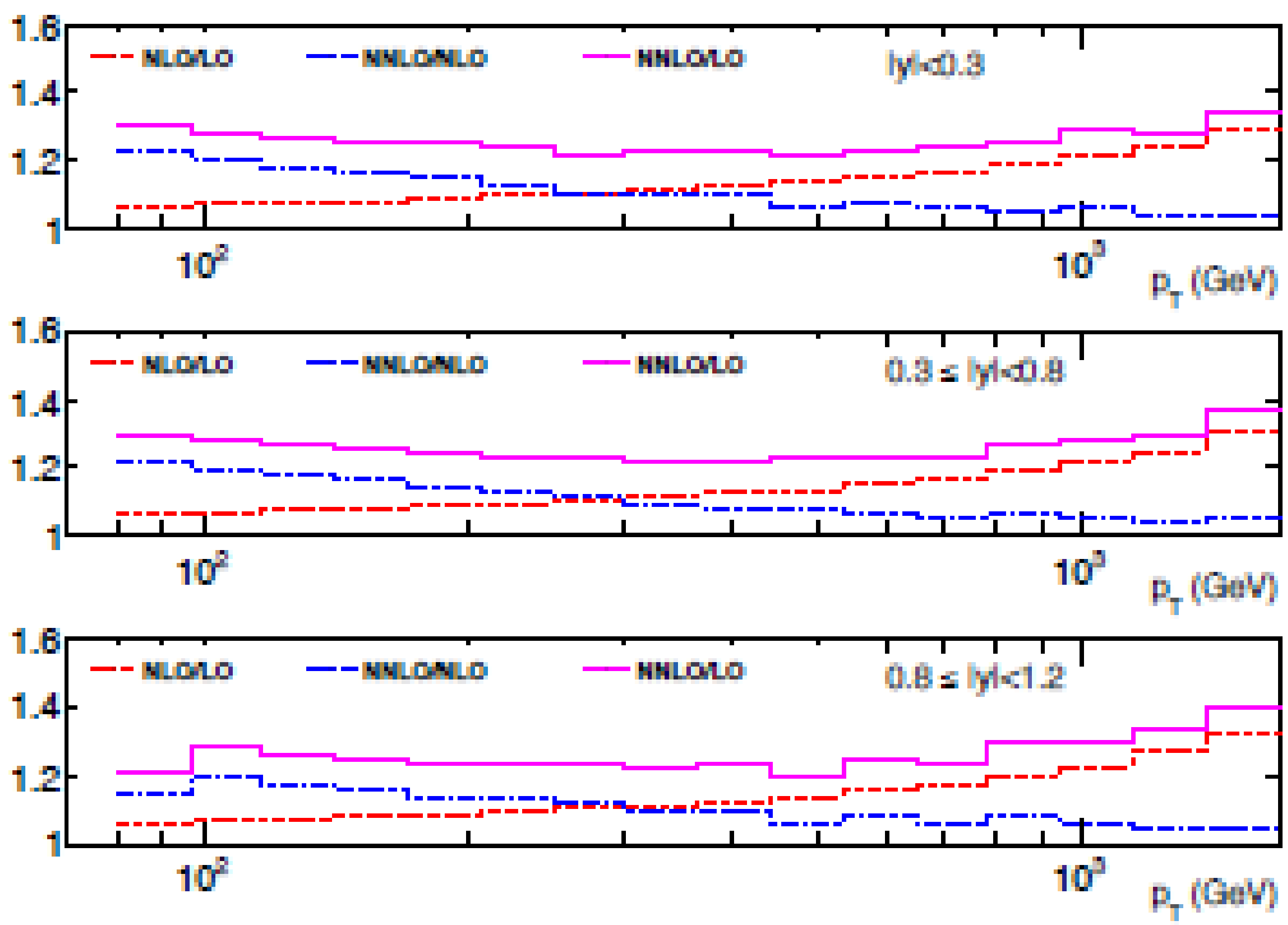
- Recent repeat of threshold calculation by Kumar, Moch (arXiv:1309.5311) and comparison to exact NLO results for different jet radii R.
- Big variation with R at NLO and threshold calculation which has no R dependence matches best with $R \sim 0.3 - 0.4$. Bit lower than typical Tevatron value, however adjusting leads to little change.

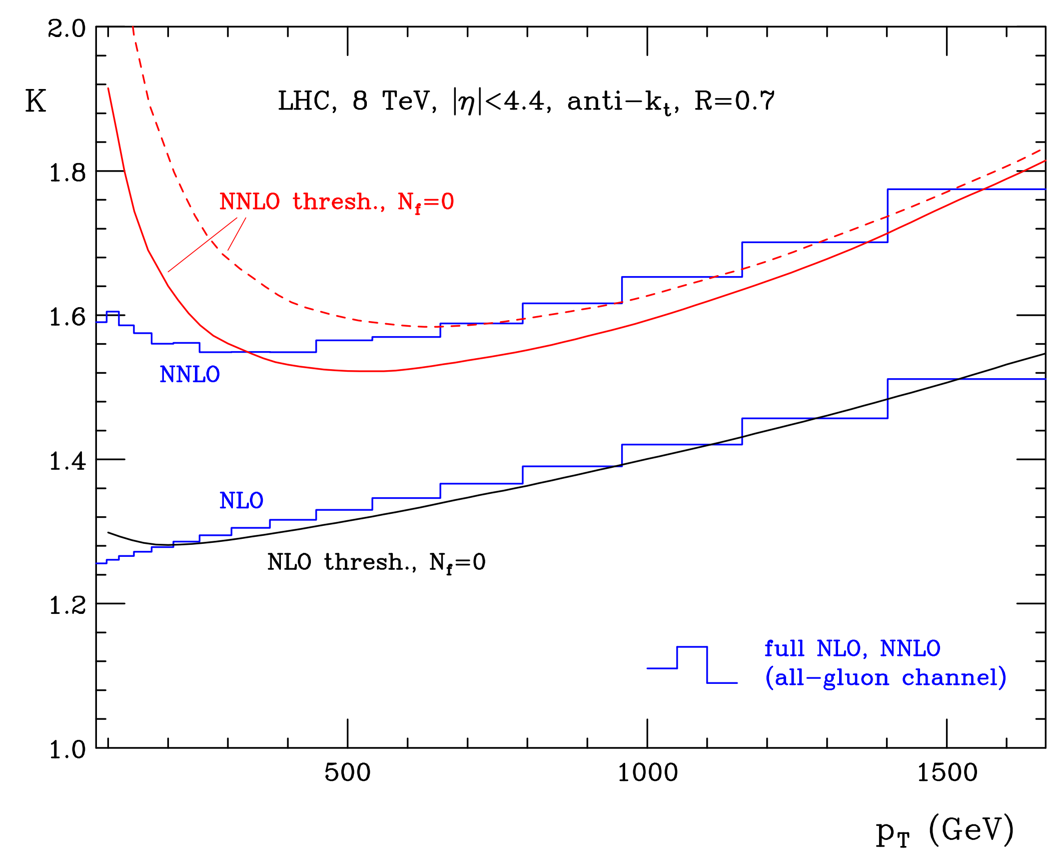
NNLO correction- existing results

Inclusive jet production:
double differential distributions

R = 0.4 Gehrmann-de-Ridder et al.

de Florian et al.





- Calculation of all-gluon contribution to jet production performed by Gehrmann-de-Ridder et al.
- Result appears to be fairly similar to threshold correction near threshold by de Florian et al (arXiv:1310.7192). Overall $\sim 5 - 10\%$ positive correction which increases at lower $p_\perp$.

NNLO jets - PDF updates

• As default NNLO set still fits Tevatron data. Seems safe as data are always relatively near to threshold, and corrections do not obviously break down at lowest probed $p_\perp$.

• In order to test robustness: have repeated MSTW2008 fits with extreme modified K-factors for NNLO jets, i.e. multiply standard corrections by 0 or 2 and use constant $K = 1.15$. All within one sigma, even for extreme changes.

• Different story for LHC data. In general much farther away from threshold, lowest not stable in threshold corrections, and large uncertainty at highest rapidity. Therefore do not include in fit.

• Test: try putting in very approximate NNLO correction of $\sim 5-20\%$ which grows at lower $p_\perp$. "Smaller" and "larger" K-factor with corrections of $\sim 10\%$ and $\sim 20\%$ at $p_\perp = 100\,\mathrm{GeV}$ - rapidity independent. Prediction is good: fit quality a small amount worse than at NLO, though deteriorates slowly with larger K-factor.

Fit quality for LHC data - NNLO

- Jet data not fitted but quality checked using "smaller" K-factor, $\sim 10\%$.

data set	N_{pts}	CPdeut	no LHC	prelim
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 $p_T > 35\text{GeV}$	11	18	15	9
CMS asymm $p_T > 25\text{GeV}, 30\text{GeV}$	24	18	17	10
LHCb $Z \rightarrow e^+e^-$	9	23	22	20
LHCb W asymm $p_T > 20\text{GeV}$	10	24	21	13
CMS $Z \rightarrow e^+e^-$	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 after fit in ATLAS W, Z data, mainly from strange quark, and in CMS $Z \rightarrow e^+e^-$ data, CMS W asymmetry and LHCb W^+, W^- data.
- CMS W asymmetry data constrains some flavour decomposition.

Fit quality for LHC data - NNLO

- Jet data not fitted but quality checked using "larger" K-factor, $\sim 20\%$.

data set	N_{pts}	CPdeut	no LHC	prelim
ATLAS jets (2.76TeV+7TeV)	116	(117)	(132)	(128)
CMS jets (7TeV)	133	(145)	(137)	(139)

and "smaller" K-factor, $\sim 10\%$

data set	N_{pts}	CPdeut	no LHC	prelim
ATLAS jets (2.76TeV+7TeV)	116	(107)	(123)	(119)
CMS jets (7TeV)	133	(142)	(137)	(135)

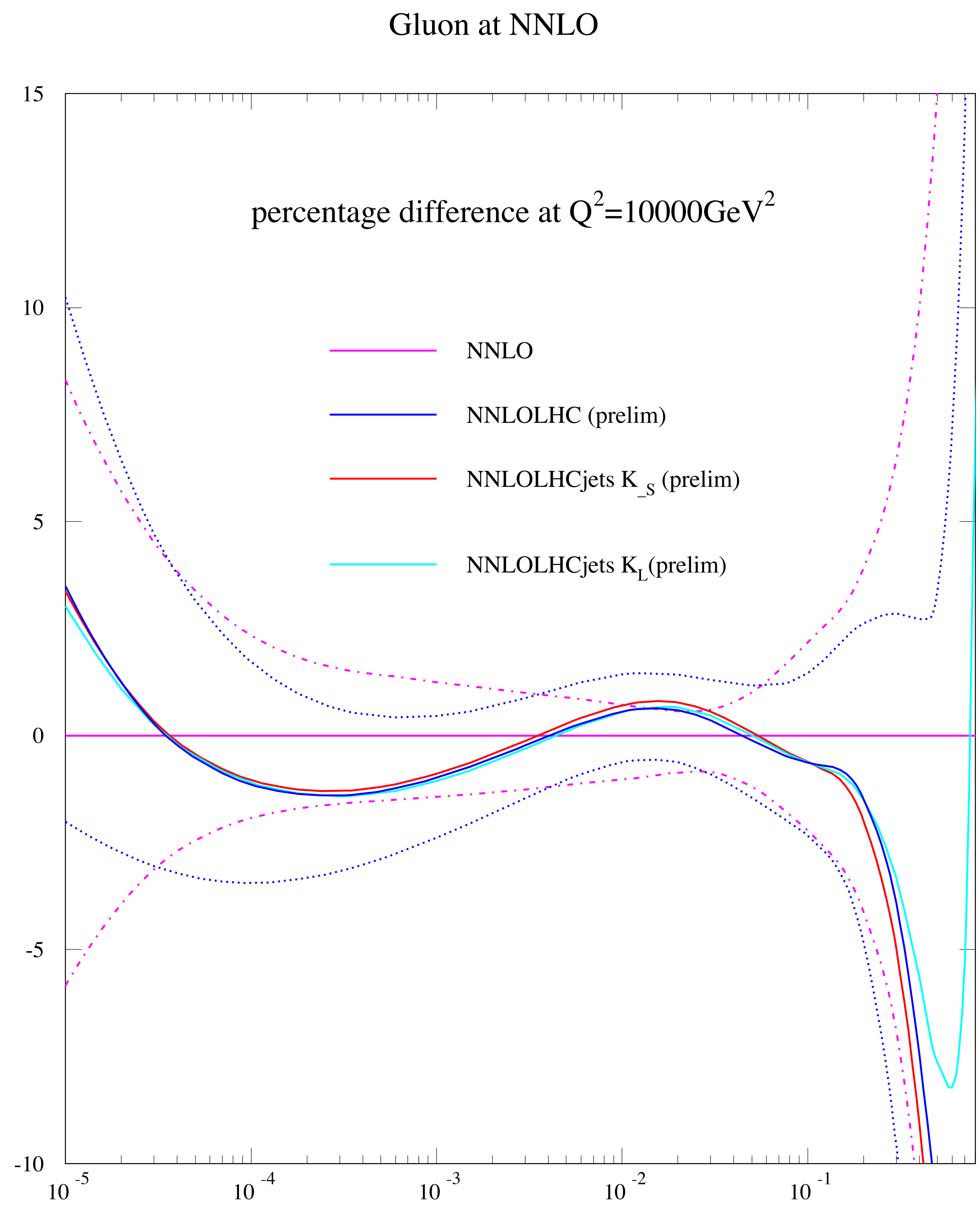
- ATLAS jet data deteriorates more than CMS, which with increase in systematics is largely insensitive to K-factor, though even prefers smaller one. Difficult to guess relative size of K-factor at two different energies.

NNLO fit with jets - change in gluon

- Ratio of $g(x, Q^2)$ for the default NNLO fit to that in MSTW2008, and also fits where jet data are 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.

"Smaller" K-factor ($\sim 10\%$)
ATLAS $\chi^2 = 119/116 \rightarrow 106/116$
and CMS $\chi^2 = 138/133 \rightarrow 139/133$

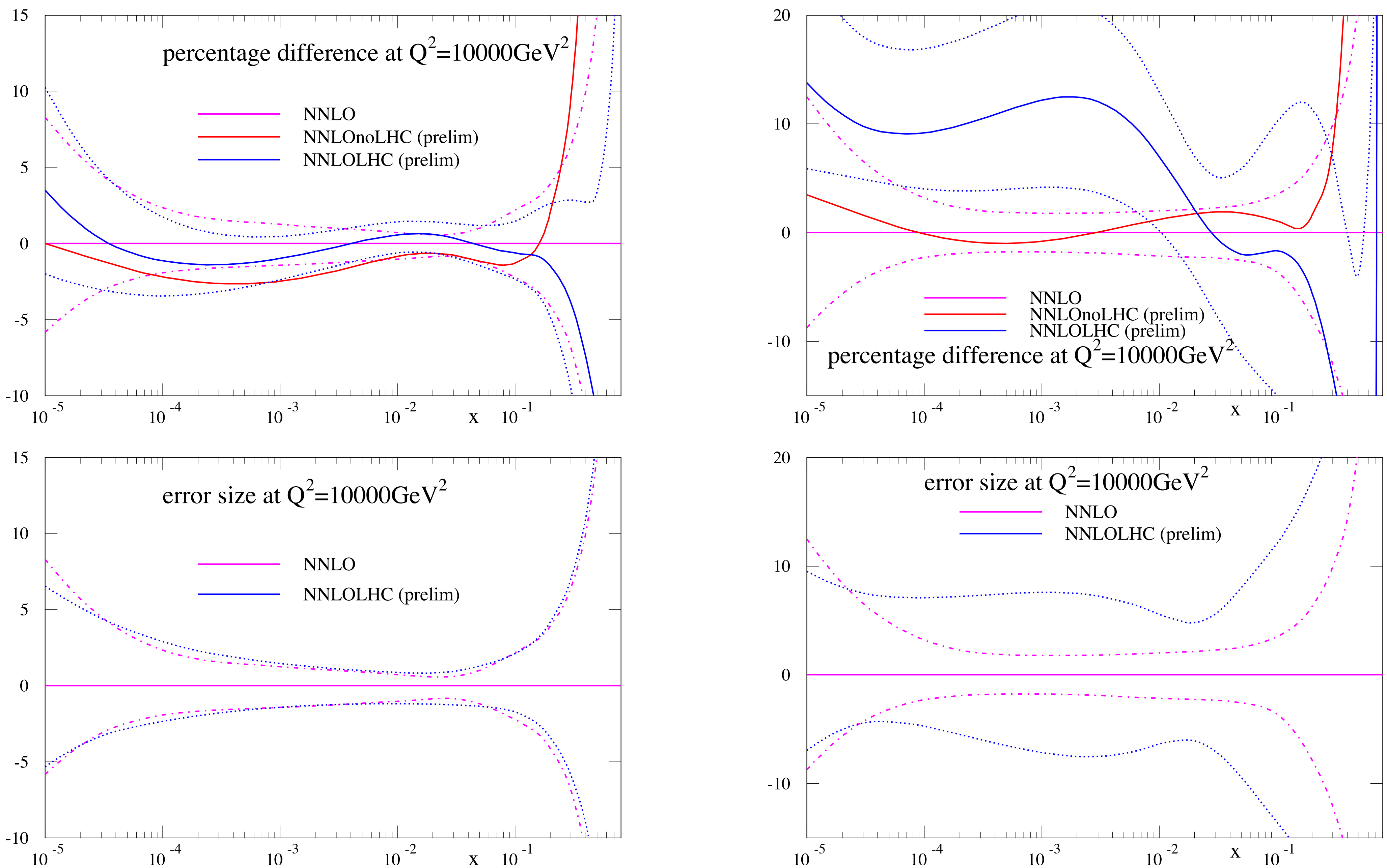
"Larger" K-factor ($\sim 20\%$)
ATLAS $\chi^2 = 128/116 \rightarrow 118/116$
and CMS $\chi^2 = 139/133 \rightarrow 141/133$



Change in NNLO PDFs

Gluon at NNLO

Strange+antistrange quark at NNLO



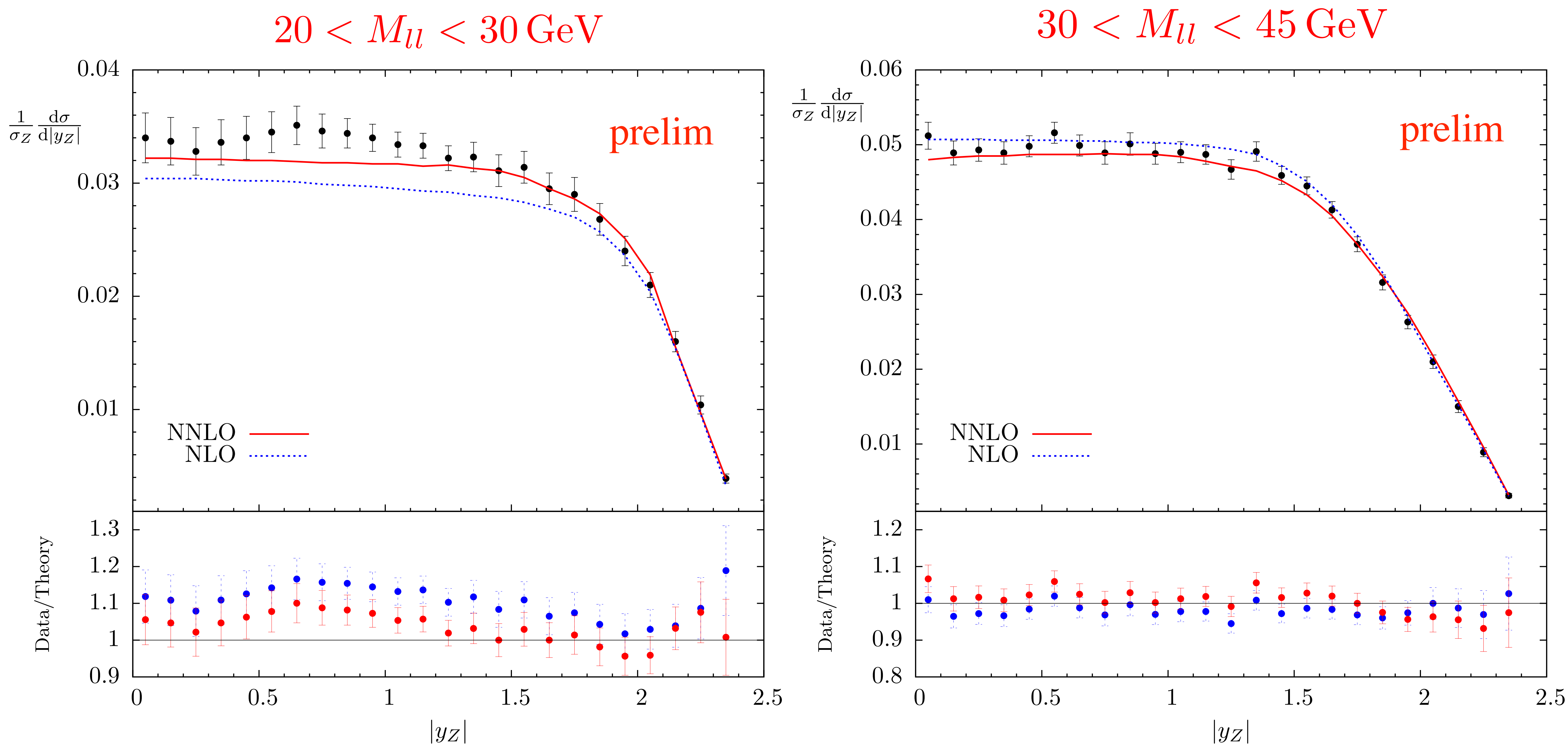
Includes theoretical updates and LHC data (not absolutely final). Gluon uncertainty slightly larger at high x - no jet data in fit.

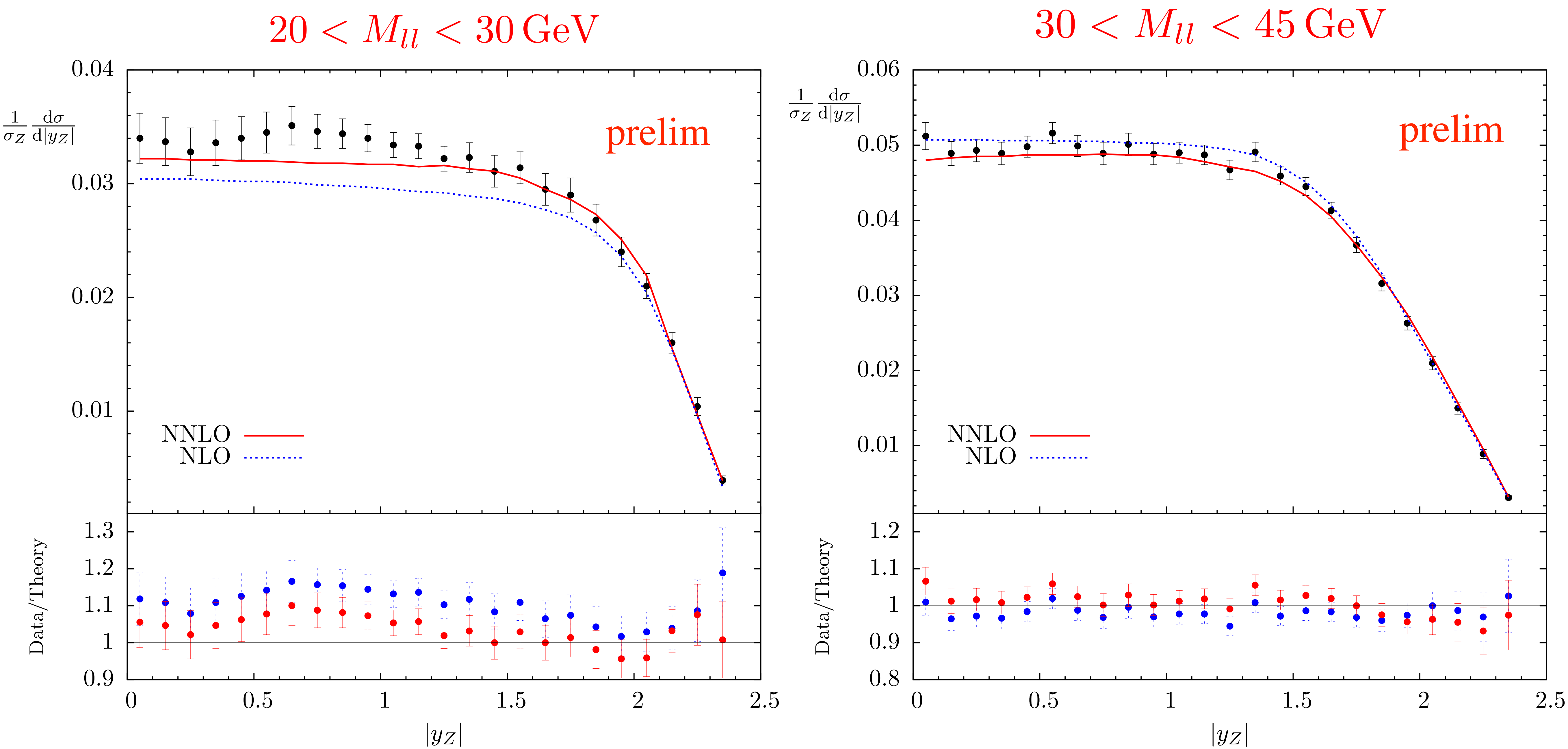
CMS Drell-Yan data

- Fit very poor at NLO:

data set		N_{pts}	CPdeut	no LHC	prelim
CMS Low-high mass DY	NNLO	132	159	151	149
CMS Low-high mass DY	NLO	132	385	396	373

- In lowest mass bins, $20 < M_{ll} < 45\,\mathrm{GeV}$, cuts on leptons ($p_\perp > 9, 14\,\mathrm{GeV}$) mean NLO is effectively LO, and fit is very poor:

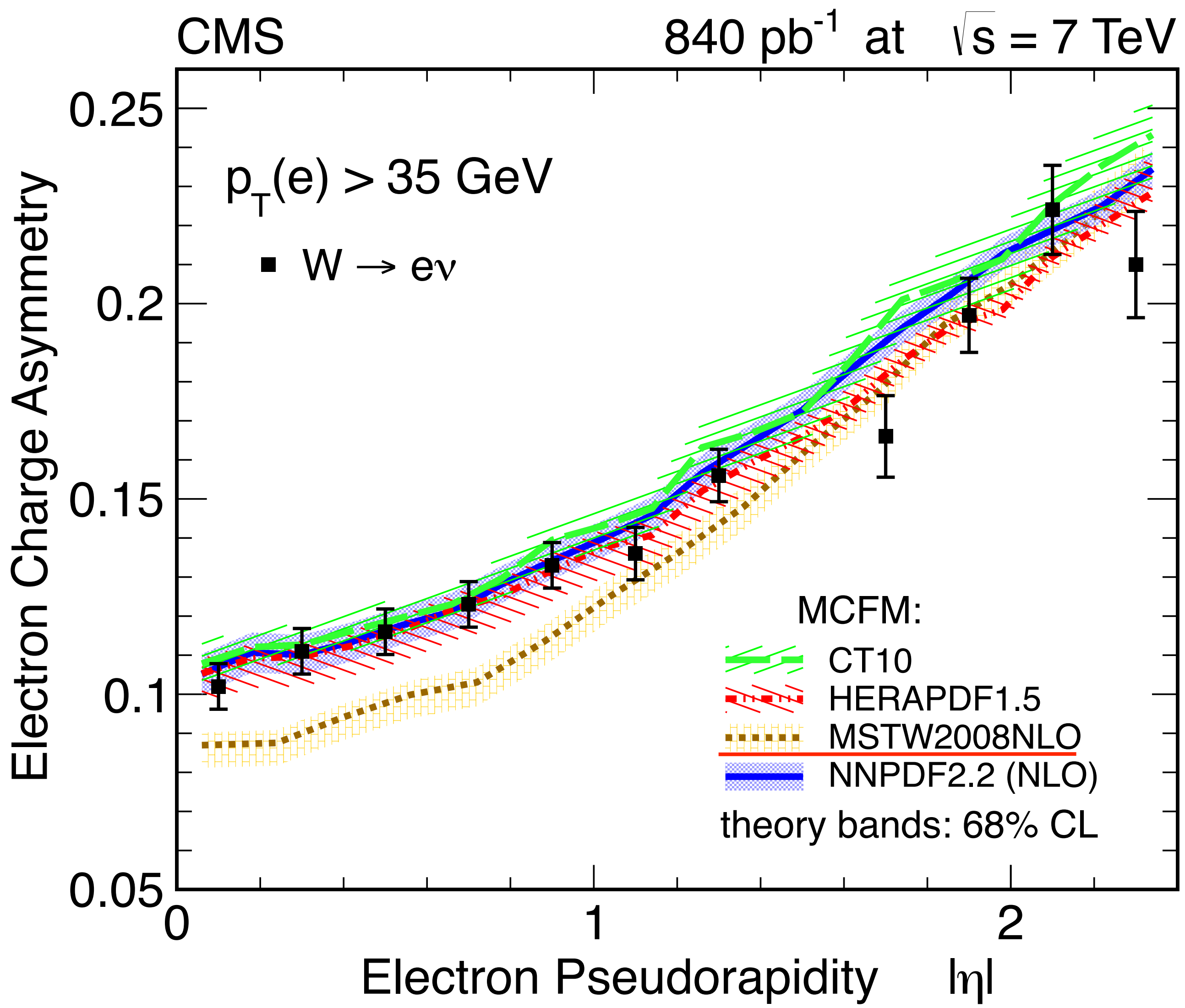


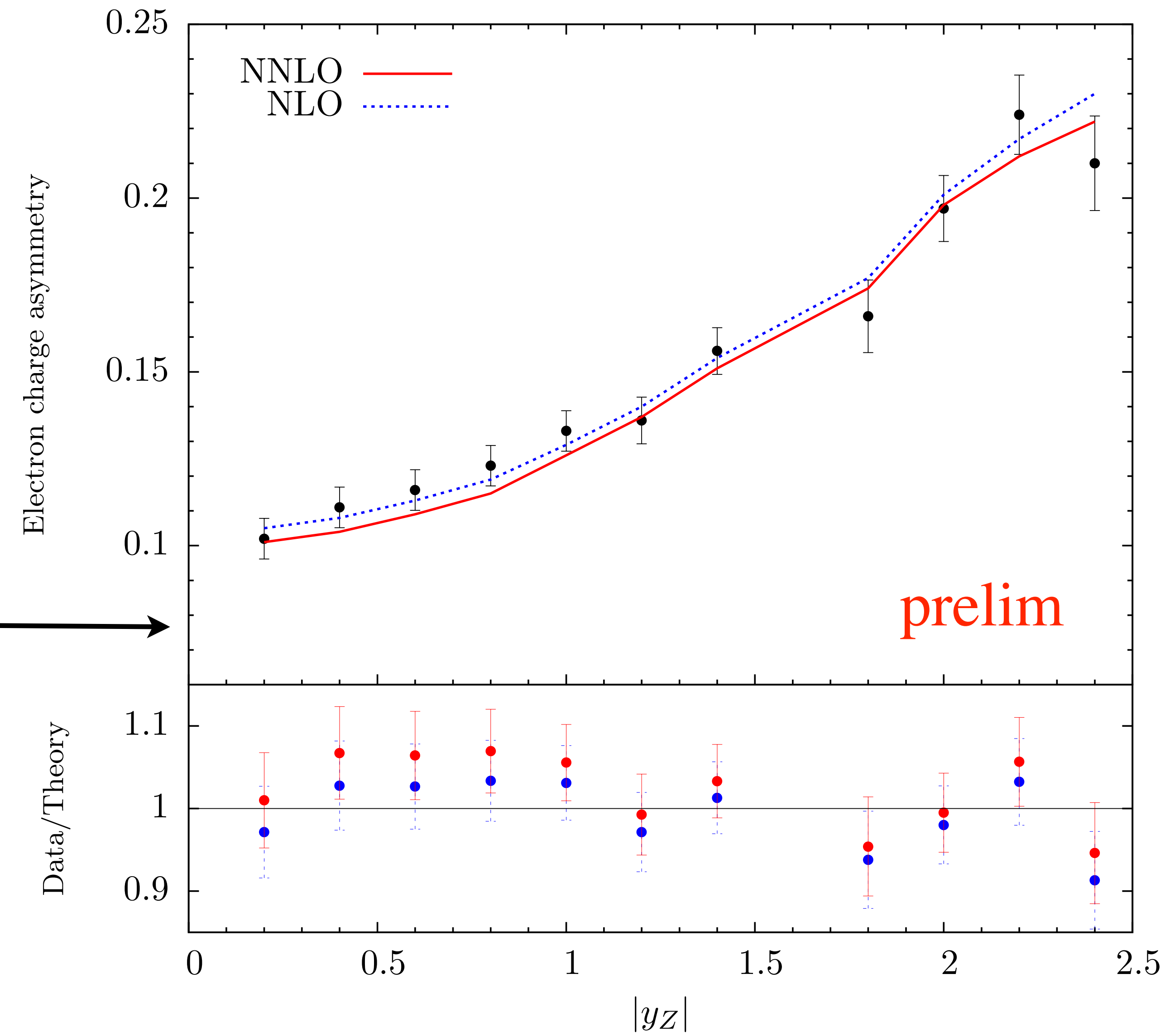


- Enormously improved fit quality at NNLO due to improvement in theory.
- Sensitivity to strange fraction in quarks, but differs at NLO and NNLO and weak compared to direct constraint from dimuon data.

CMS W asymmetry - fit

- CMS W asymmetry (arXiv:1206.2598) - no longer an issue (extended parametrization and deuteron corrections).





data set	N_{pts}	CPdeut	no LHC	prelim
CMS W asymm $p_T > 35$GeV	11	18	15	9

NNLO

$\alpha_S(M_Z^2)$ 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 $\longrightarrow$ try inputing this as a data point.

- Try world average (minus DIS data) of $\alpha_S(M_Z^2) = 0.1187 \pm 0.0007$ (rather small uncertainty, mainly from lattice):

NLO : already within one sigma, essentially no change - $\alpha_S(M_Z^2) = 0.1199 \rightarrow 0.1195$ with $\Delta\chi^2 < 2$.

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 $\alpha_S(M_Z^2) = 0.118$:

NLO : $\Delta\chi^2 \sim 16$, but not single set deteriorates significantly.

NNLO : basically no further change.

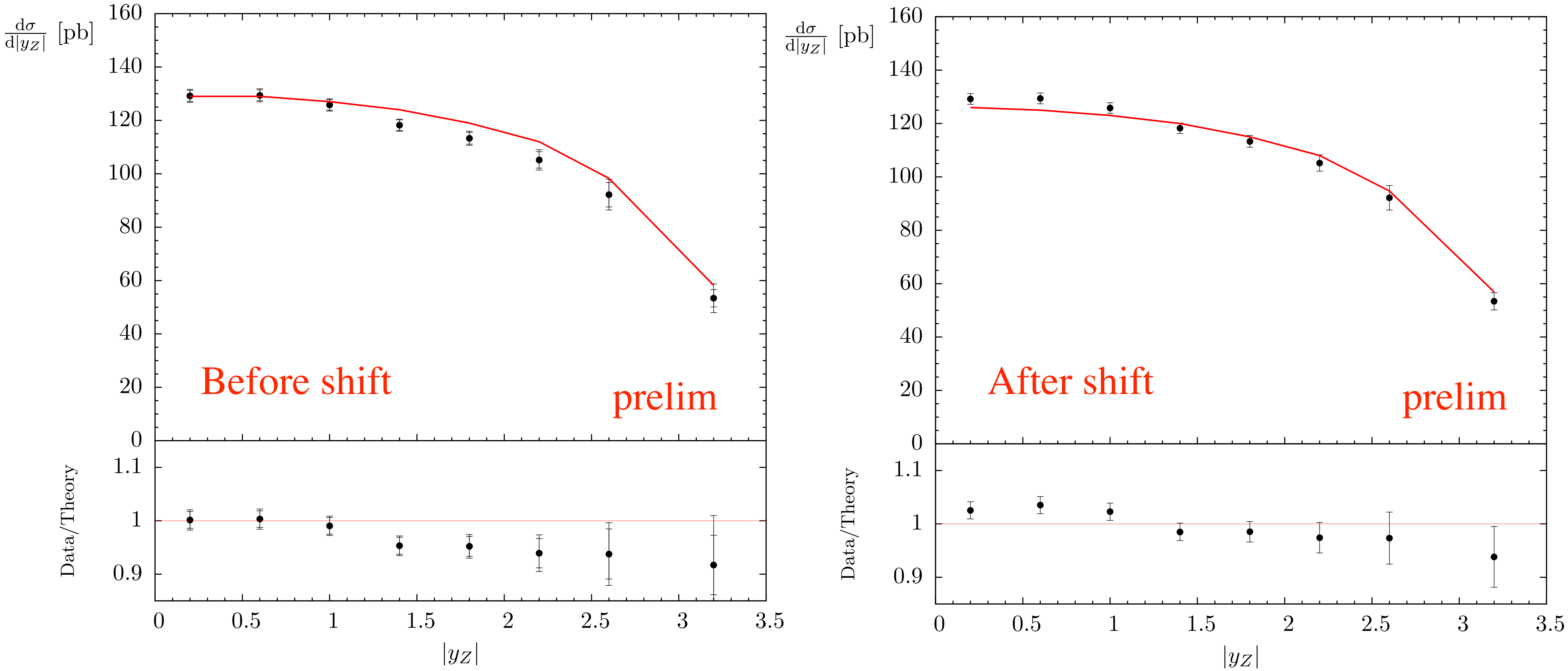
Conclusions

- Ongoing, but very near final, updates on PDFs - soon to be released.
- Various theoretical improvements described: parameterisation, deuteron corrections, heavy flavour treatments, nuclear corrections, branching ratio for dimuon data.
- Inclusion of up-to-date HERA and Tevatron data
- Directly included most relevant published LHC data: ATLAS, CMS, LHC W,Z rapidity data, top cross sections and all published ATLAS and CMS inclusive jet data (but not at NNLO).
- Fit good (except for CMS double differential at NLO - but clear reason for this). No PDF conflicts.
- So far few dramatic effects on PDFs. Mainly strange quark and low x valence quarks, largely due to change in methodology, but also newer data. Larger strange uncertainty from branching ratio error.
- Some uncertainty in NNLO effect on jets. Have decided at present to wait for full NNLO calculation. However, comparison now suggests NNLO fit is happy with moderate guesses for K-factors, with no real change in PDFs or coupling.

Backup

Correlated errors - data shift

ATLAS Z rapidity (arXiv:1109.5141)



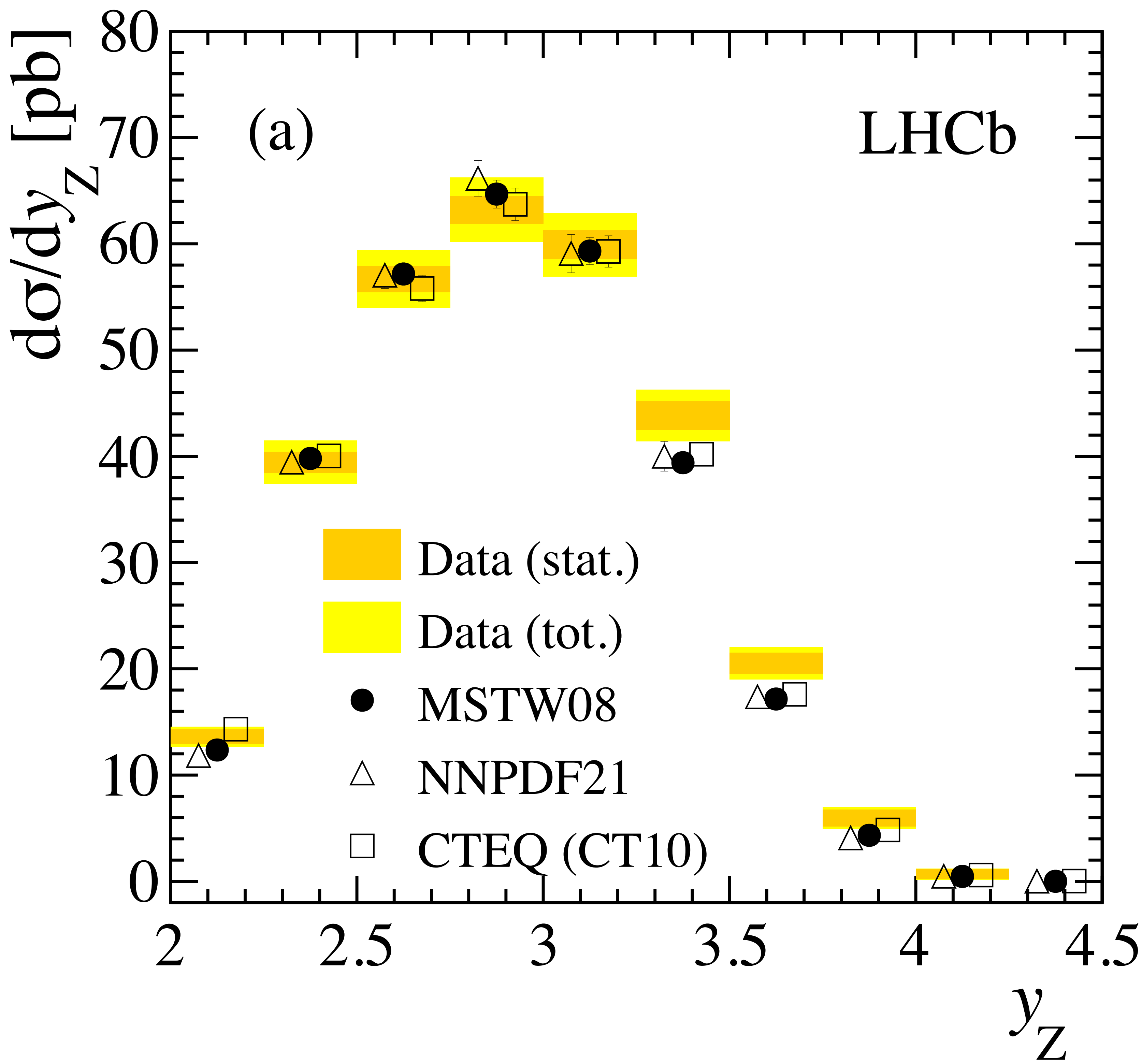
Recall treatment of correlated errors: $\chi^2 \sim \left(\frac{D_i - T_i/f}{\sigma_i^{uncorr}}\right)^2 = \left(\frac{f * D_i - T_i}{f * \sigma_i^{uncorr}}\right)^2$

$\longrightarrow$ Data (and error) allowed to shift by fraction f to give best fit.

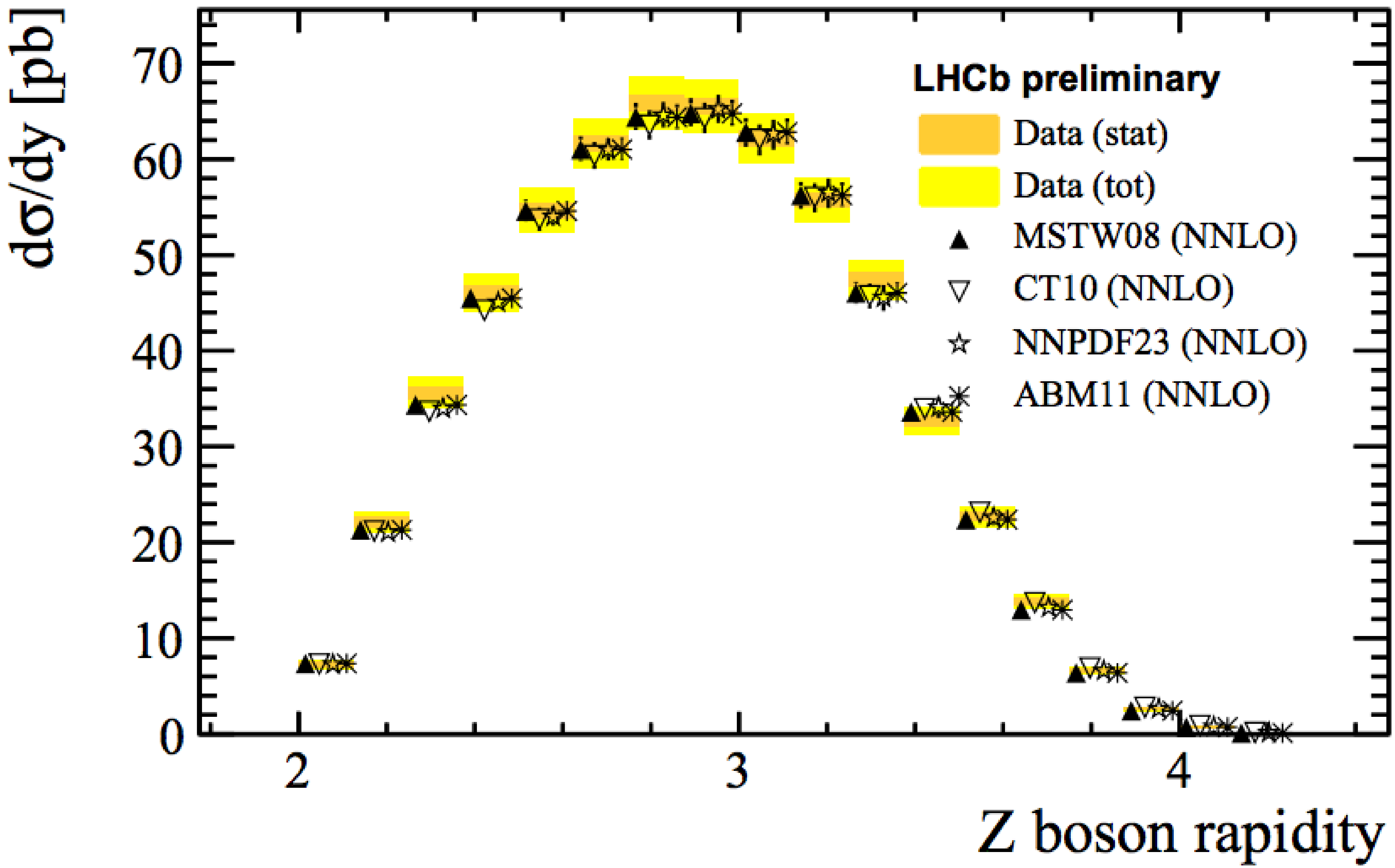
Change in various cross section predictions compared to uncertainty for MSTW2008.

	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.



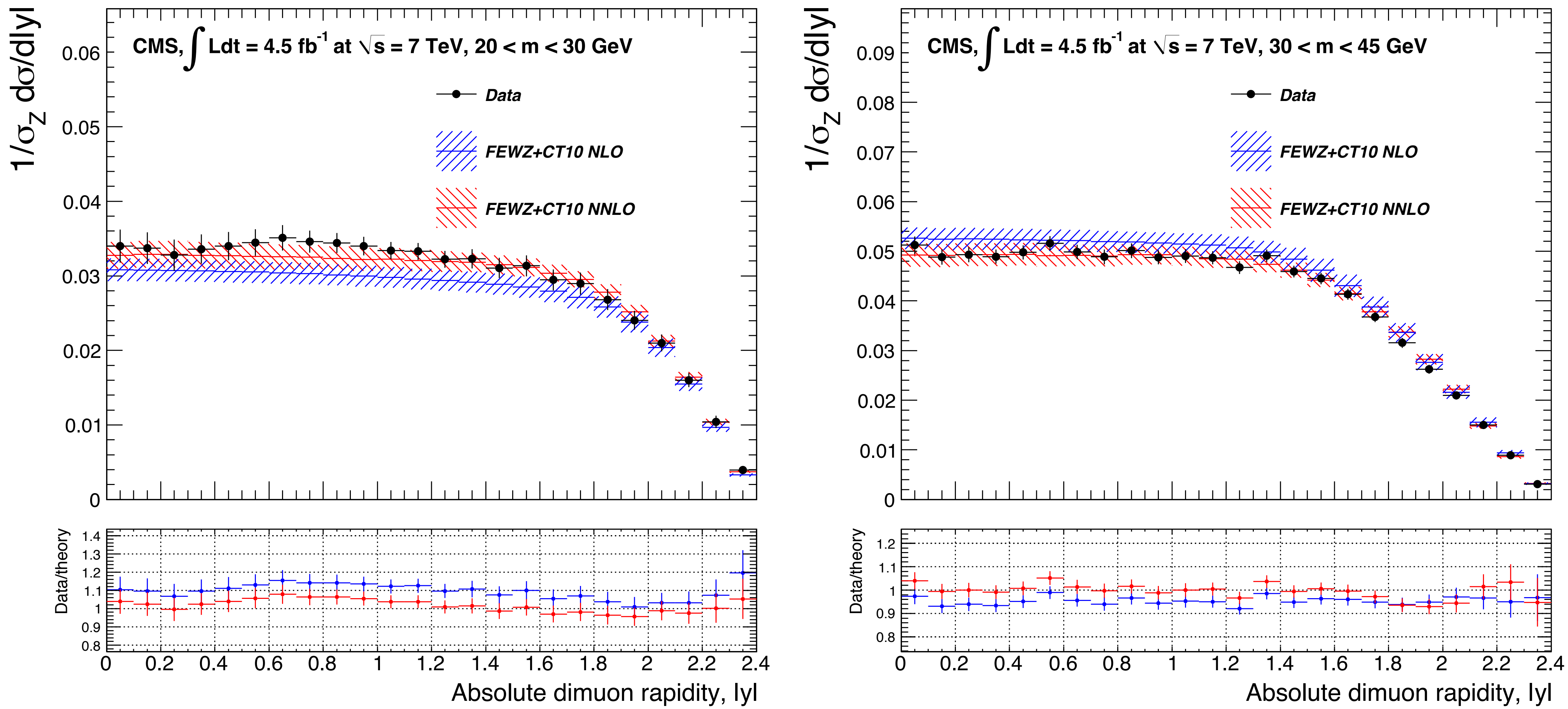
Points near to $y = 3.5$ overshoot predictions in general. Feature not present in prelim. higher luminosity $Z \to \mu^+\mu^-$ data.



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

CMS Drell Yan data.

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



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

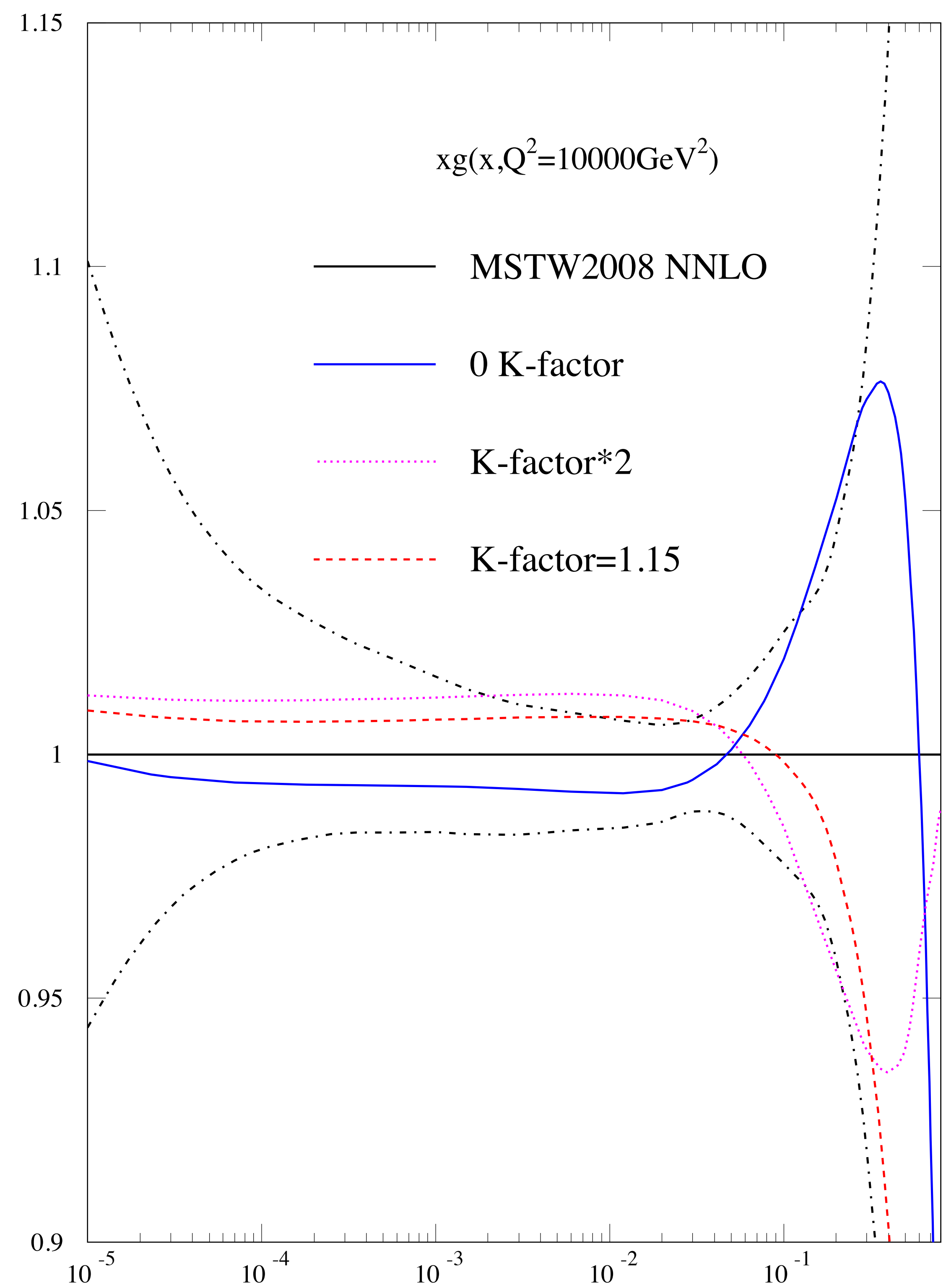
Extreme variations.

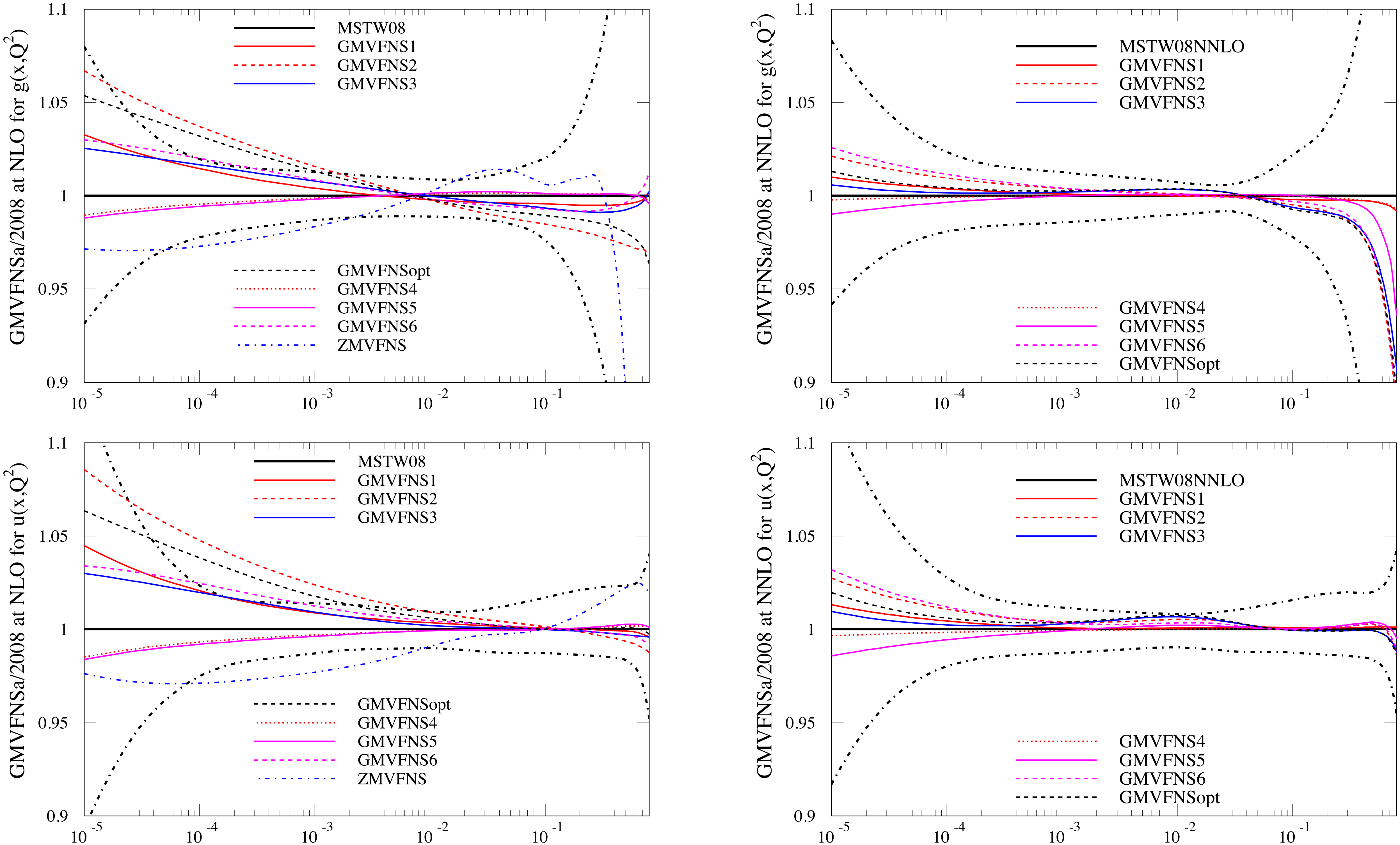
Changes in gluon relatively small. Larger K-factor slightly worse χ^2. Zero K-factor slightly better χ^2, $K = 1.15$ almost no change.

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

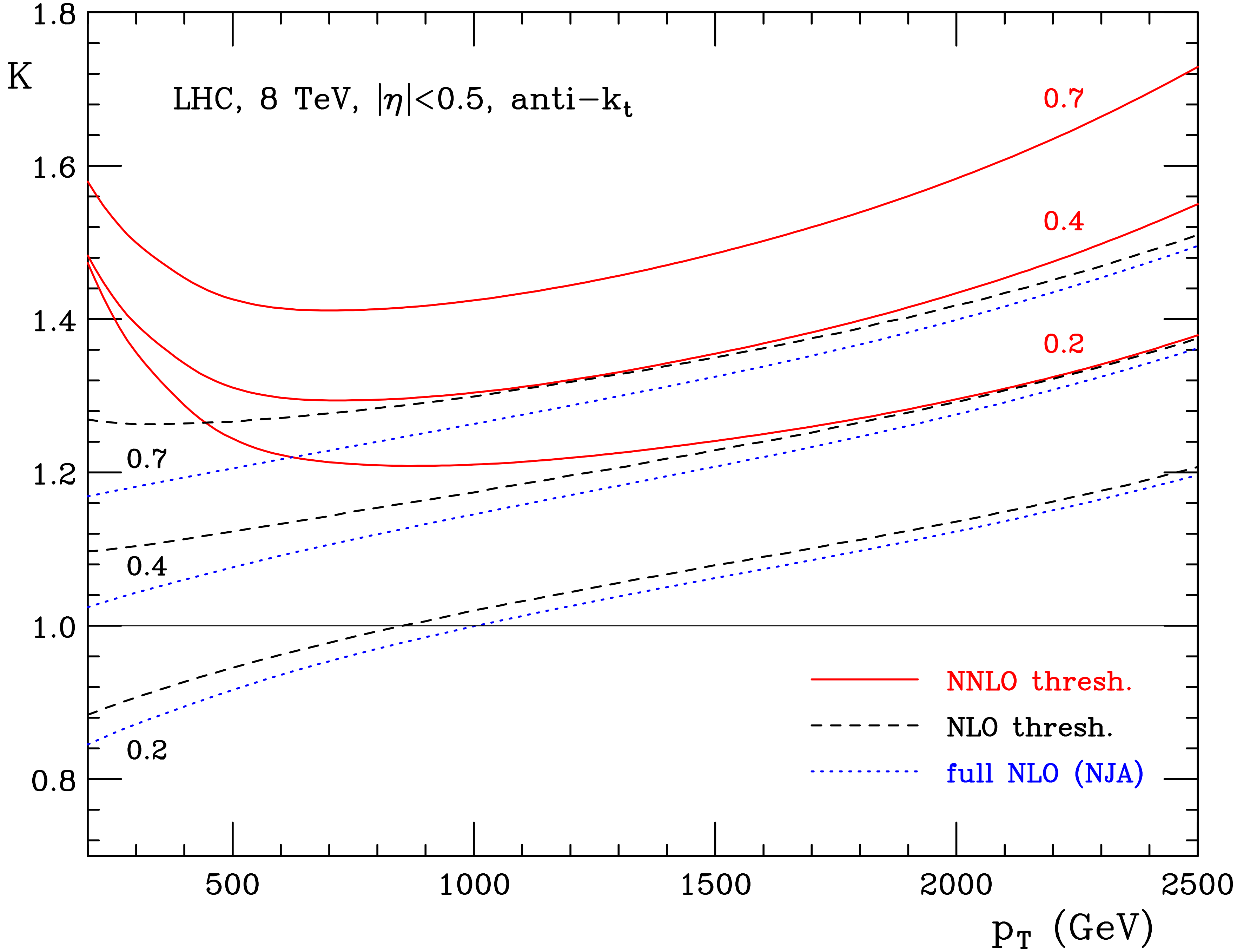
$K * 2 \ \alpha_S(M_Z^2) = 0.1159$

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





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 $\alpha_S(M_z^2)$.



• Very recent improved calculation by 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 $p_\perp$.