# MSTW PDFs

Robert Thorne

August 22nd, 2011

University College London

Together with Alan Martin, James Stirling and Graeme Watt

# Variety of PDFs

MSTW make available PDFs in a very wide variety of forms.

- At , LO, NLO and NNLO, with some minor approximations at NNLO.
- Also a variety of extensions such as different $\alpha_S$ values, heavy quark masses, different flavour numbers. Some covered later in the week.
- Older MRST versions of modified LO* and LO** PDFs and of PDFs including QED evolution.

## Data fit

- Lepton-proton collider HERA – (DIS) $\rightarrow$ small-$x$ quarks, and gluons from evolution. Also, jets $\rightarrow$ moderate-$x$ gluon and $\alpha_S$ (not at NNLO).

- High-$p_T$ jets at colliders (Tevatron - Run II) – high-$x$ gluon distribution.

- $W$ and $Z$ production at colliders (Tevatron -Run II) (low luminosity Run II for $W$(lepton) asymmetry) – different quark contributions to DIS.

- Fixed target neutral current DIS – higher $x$ – leptons (BCDMS, NMC, E665, SLAC) $\rightarrow$ up quark (proton) or down quark (deuterium).

- Fixed target charged current DIS – neutrinos (CHORUS, NuTeV) (cut above $x = 0.5$ on latter) $\rightarrow$ valence or singlet combinations.

- Di-muon production in neutrino DIS – (CCFR, NuTeV) strange quarks and neutrino-antineutrino comparison $\rightarrow$ asymmetry .

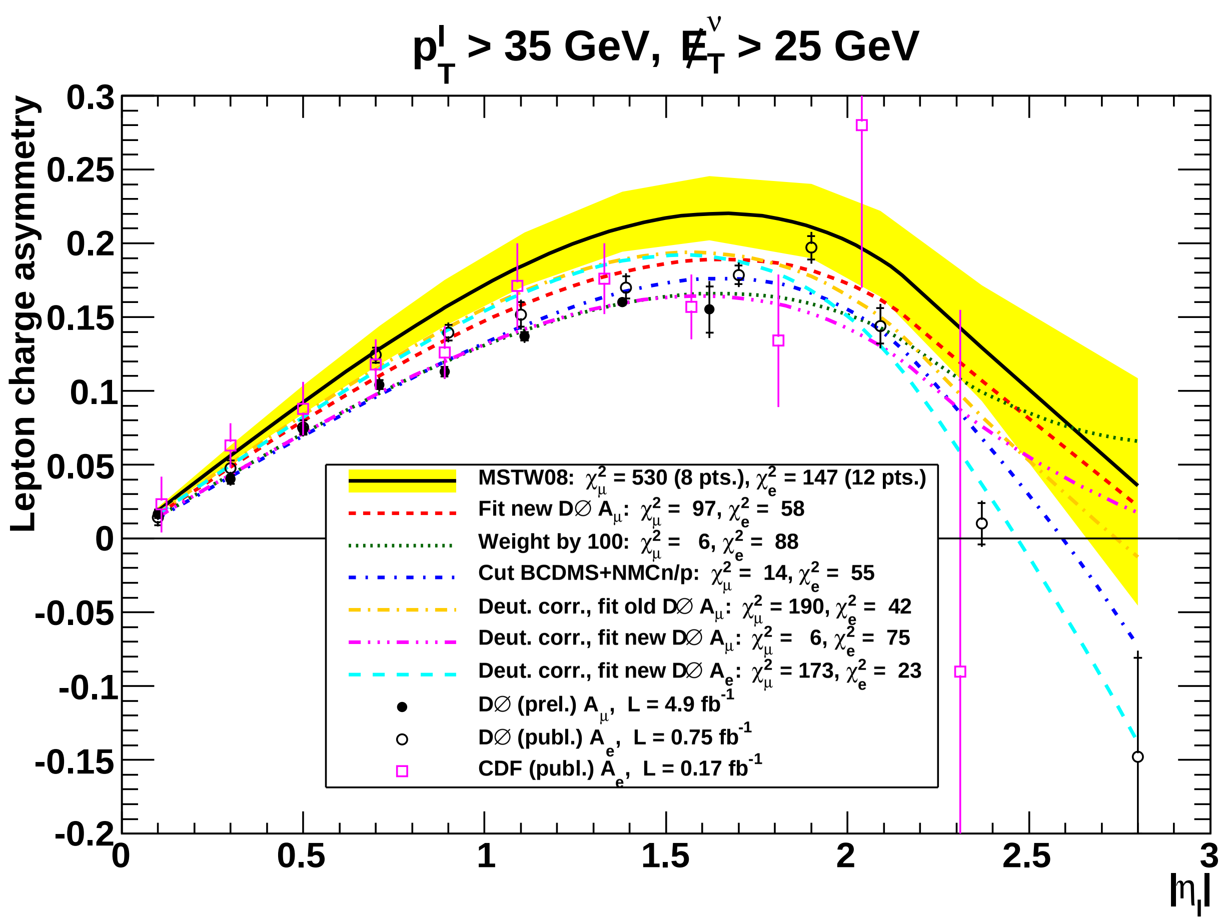
- Drell-Yan production of dileptons – quark-antiquark annihilation (E866 $pp$ experiment) – high-$x$ sea quarks. Deuterium target (E866) – $\bar{u}/\bar{d}$ asymmetry.

Keep first three in latter fits.

Fit data for scales above $2\mathrm{GeV}^2$. (most) DIS data for $W^2 > 15\mathrm{GeV}^2$. Will mention effect of cuts later.

Don't yet include combined HERA cross-section data. Have checked effects of this. In some cases predictions change by about $1\sigma$, in most cases less.

Major problems with high-luminosity D0 lepton asymmetry in some binnings. Same for other groups.

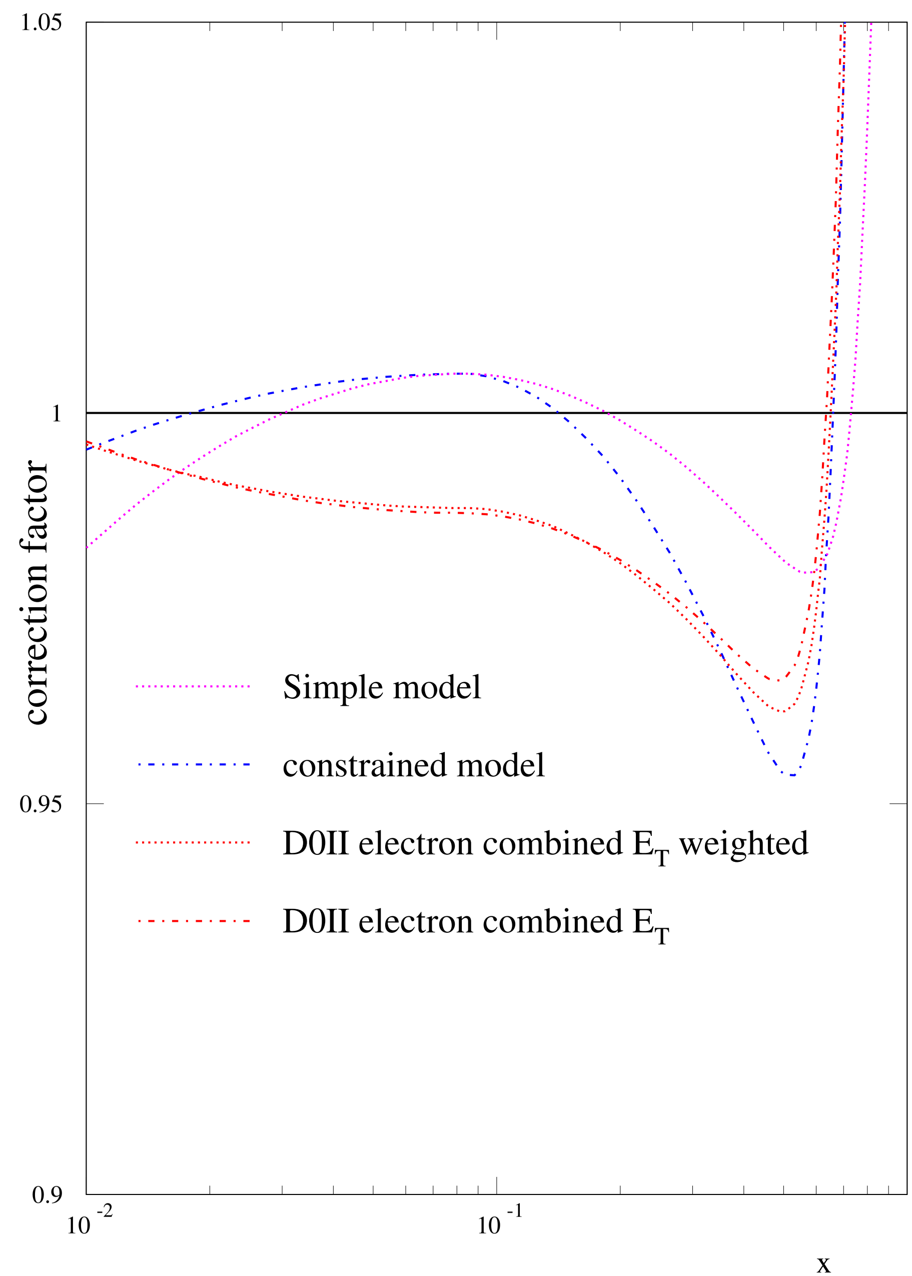



MSTW, NNPDF and CTEQ have difficulty fitting new D0 lepton asymmetry (particularly muon in different $E_T$ bins) along with other data.

NNPDF better for inclusive $p_T$ data. CTEQ produce CT10W. MSTW better when low number of data points sets given (slightly) more weight. Also improved using deuterium corrections.

Previously found improvement in fit to both global data set and lepton asymmetry with deuterium corrections, but $< 1$ for all but very high $x$.

Also find significant improvement with rather more plausible deuterium corrections.



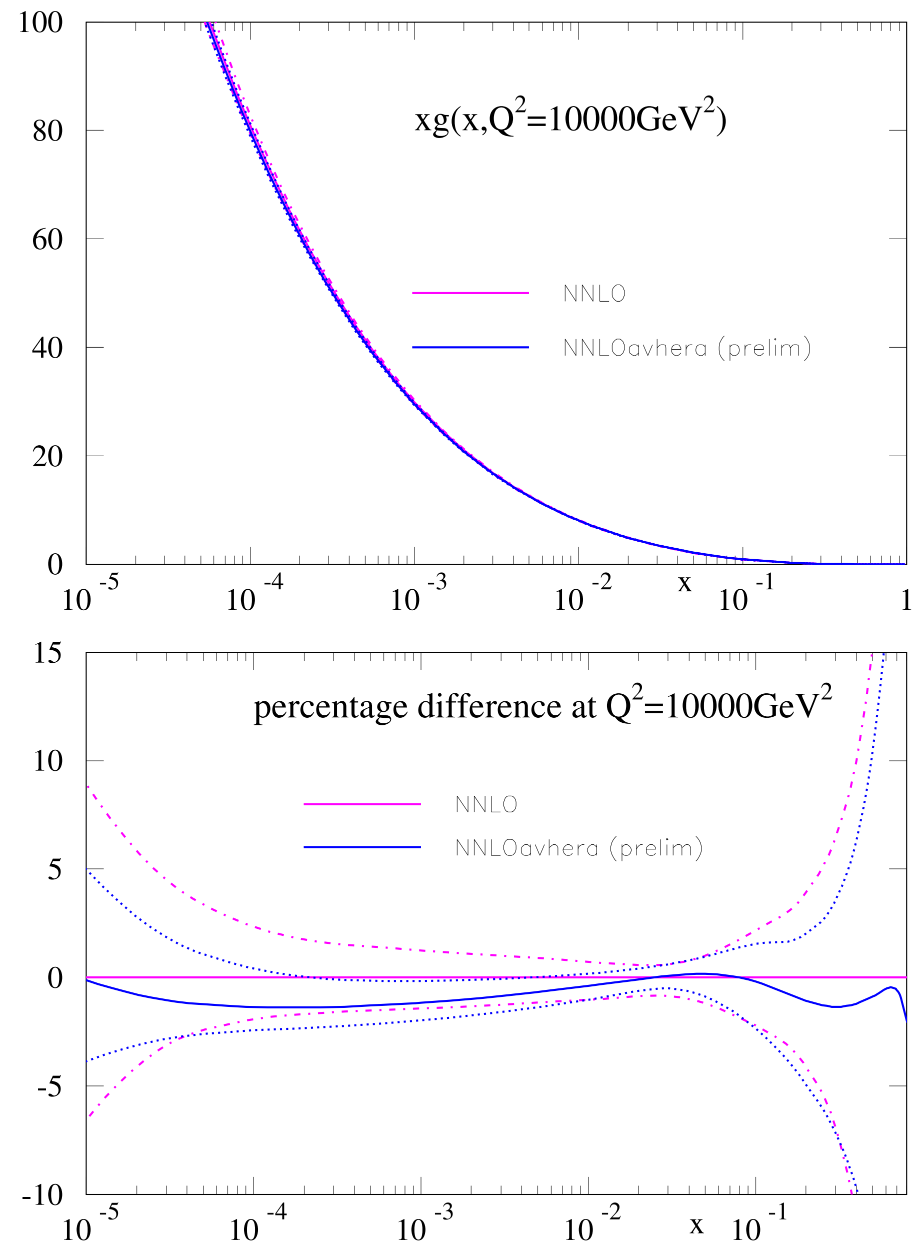
Comparison of gluon from fit using combined HERA data to MSTW2008 NNLO versions with $1-\sigma$, uncertainty shown.

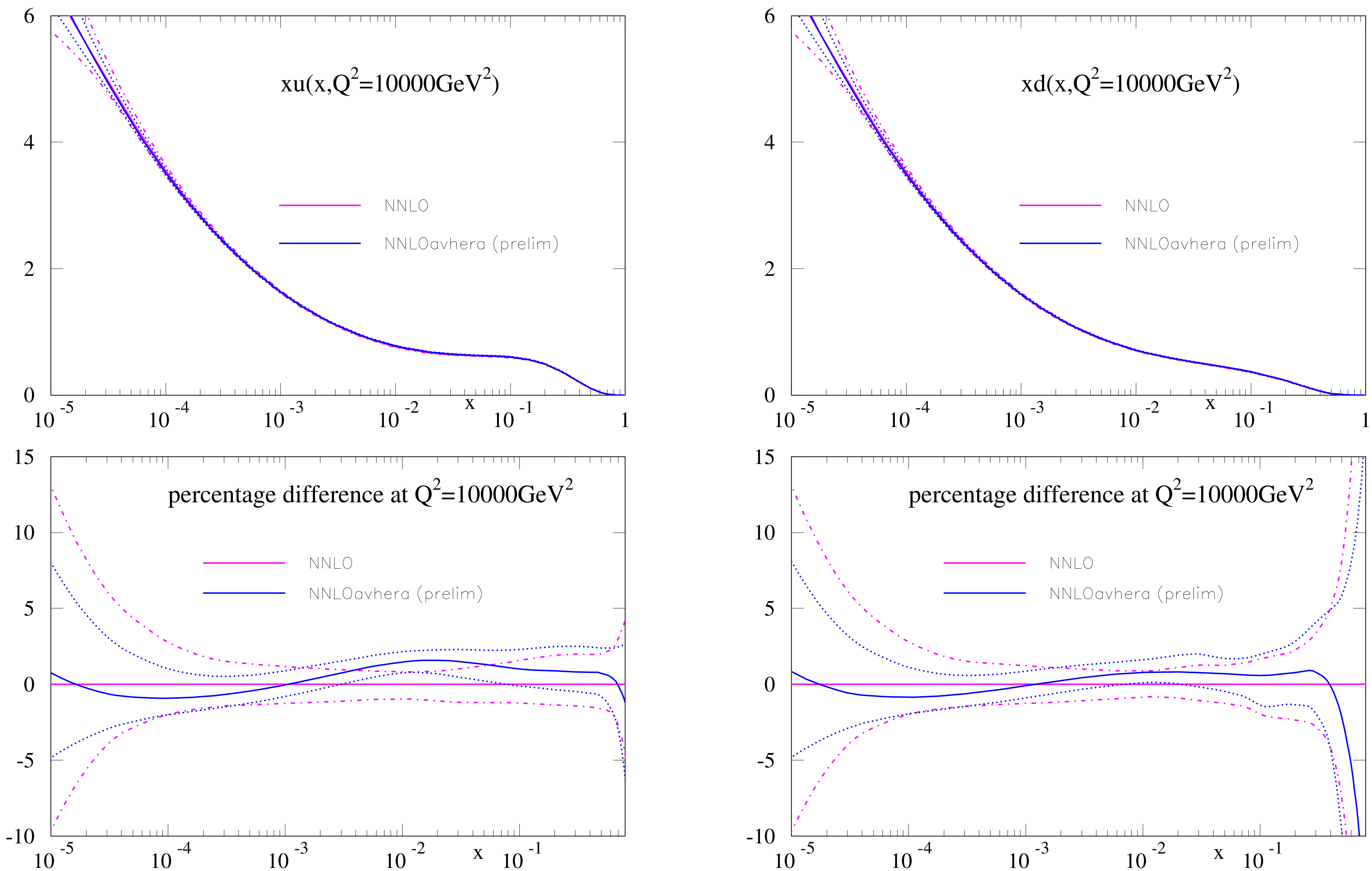
Slight difference in details of normalisation treatment compared to previous versions, still preliminary. First time have shown uncertainty.

Value of $\alpha_S(M_Z^2)$ moves slightly, $0.1171 \rightarrow 0.1178$.

Changes always within $1-\sigma$, and really less due to correlations with $\alpha_S$.

Uncertainty slightly smaller, especially at very small $x$.






Most dramatic change for up quark at about $x = 0.01$.

## Impact on Cross Sections - NNLO.

The values of the predicted cross-sections at NNLO for $Z$ and a 120 GeV Higgs boson at the Tevatron and the LHC (latter for 14 TeV centre of mass energy).

| PDF set | $B_{l^+l^-} \cdot \sigma_Z$(nb)TeV | $\sigma_H$(pb)TeV | $B_{l^+l^-} \cdot \sigma_Z$(nb)LHC | $\sigma_H$(pb)LHC |
|---|---|---|---|---|
| MSTW08 | 0.2507 | 0.9549 | 2.051 | 50.51 |
| Comb HERA | $+2.1\%$ | $+1.2\%$ | $+0.9\%$ | $+0.7\%$ |

For new global fits $2\%$ effect on $Z$ (and $W$) cross sections at Tevatron, but small change at LHC. Similar to, or less than $1-\sigma$ uncertainty in former case.

Maximum of $\sim 1\%$ for Higgs. Small effect.

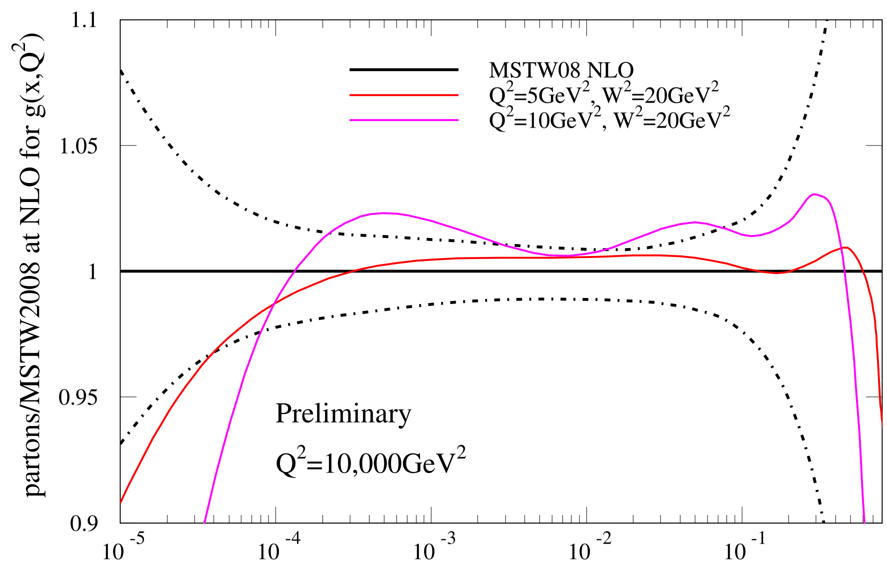
## Investigation to stability under changes in cuts.

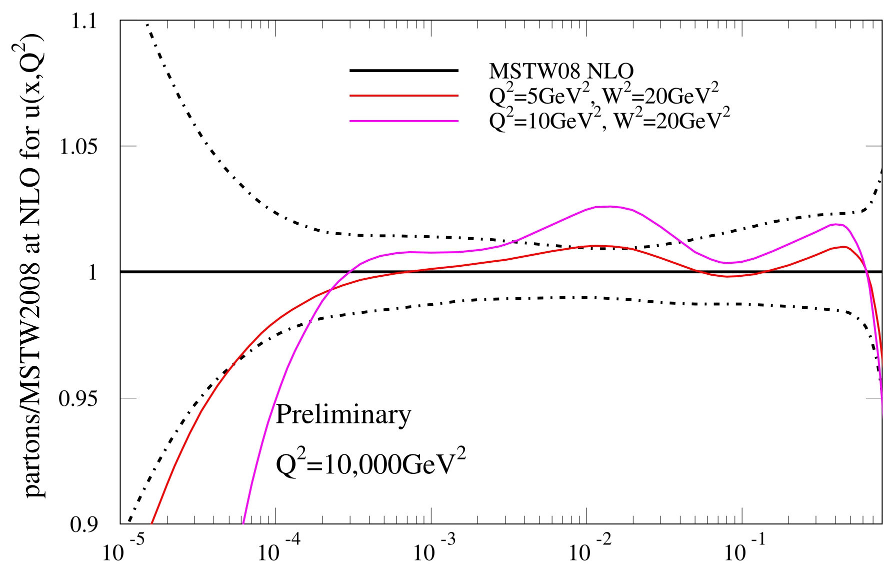
Raise $W^2_{\rm cut}$ to $20{\rm GeV}^2$, but no real changes.

Also raise $Q^2_{\rm cut}$ to $5{\rm GeV}^2$ and then $10{\rm GeV}^2$.

At NLO some movement just outside default error bands at general $x$.

Find $\alpha_S(M_Z^2) = 0.1202 \to 0.1193 \to 0.1175$, though for $Q^2 = 10{\rm GeV}^2$ cut error has roughly doubled to about $0.0025$.

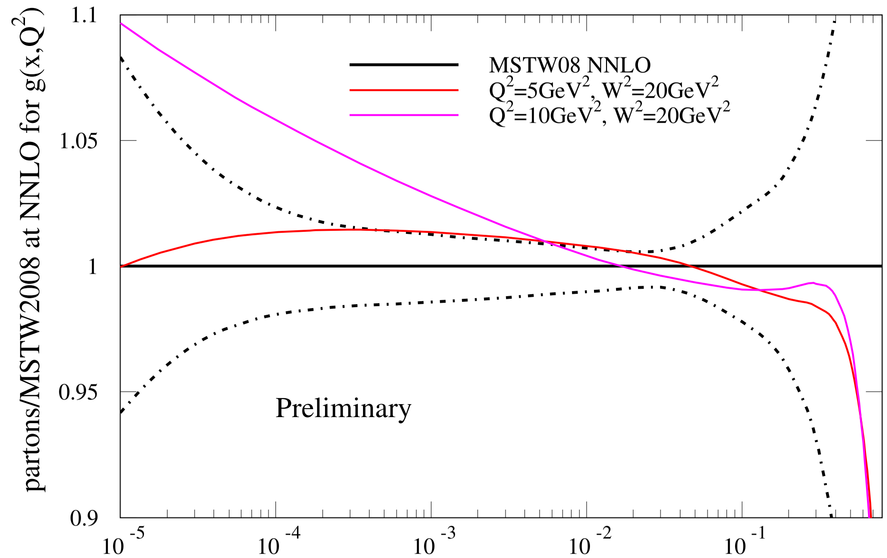


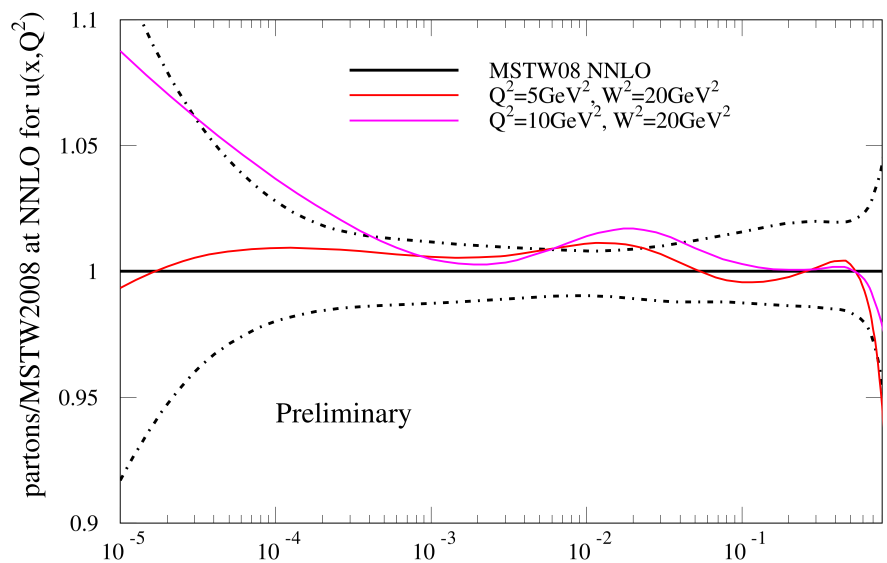



At NNLO most movement outside default error bands at low $x$, where constraint vanishes as $Q^2$ cut raises.

For $Q^2_{\rm cut} = 10{\rm GeV}^2$ no points below $x = 0.0001$, and little lever arm for evolution constraint for a bit higher.

Find $\alpha_S(M_Z^2) = 0.1171 \rightarrow 0.1171 \rightarrow 0.1164$, i.e. no change of significance.



The % change in the cross sections after cuts ($M_H = 165\mathrm{GeV}$).

| | NLO | | NNLO | |
|---|---|---|---|---|
| | $Q^2_{\rm cut}$=5GeV$^2$ | $Q_{\rm cut}2$=10GeV$^2$ | $Q^2_{\rm cut}$=5GeV$^2$ | $Q^2_{\rm cut}$=10GeV$^2$ |
| $W$ Tev | 0.0 | -2.4 | -0.7 | -0.4 |
| $Z$ Tev | 0.0 | -0.8 | -0.4 | 0.0 |
| $W$ LHC (7TeV) | -0.2 | -0.1 | -0.2 | -0.2 |
| $Z$ LHC (7TeV) | -0.2 | -0.3 | -0.4 | -0.5 |
| $W$ LHC (14TeV) | -0.6 | -1.1 | 0.3 | 0.8 |
| $Z$ LHC (14TeV) | -0.6 | -1.5 | 0.2 | 0.4 |
| Higgs TeV | -1.1 | -1.5 | -1.2 | -3.2 |
| Higgs LHC (7TeV) | -0.8 | -2.5 | 0.4 | -1.8 |
| Higgs LHC (14TeV) | -0.9 | -1.9 | 1.0 | -0.8 |

More variation at NLO than at NNLO, i.e. $7$ changes of $> 1\%$ compared to $4$.

However, both small, and changes with change in $Q^2_{\rm cut}$ slow. Does not suggest significant higher twist or problem with default cuts.

## Stability to changes in NMC data treatment.

In 1997 measurement of structure functions NMC obtained $R(x,Q^2) = F_L(x,Q^2)/(F_2(x,Q^2) - F_L(x,Q^2))$ for a few points directly by investigating cross-section measurement at common $x$ and $Q^2$ but different $y$ from different beam energy runs.

In previous measurements 1995 had not done this but assumed SLAC parameterisation $R_{1990}(x,Q^2)$.

Sensitivity to $R(x,Q^2)$ in relationship between $F_2(x,Q^2)$ and cross section.

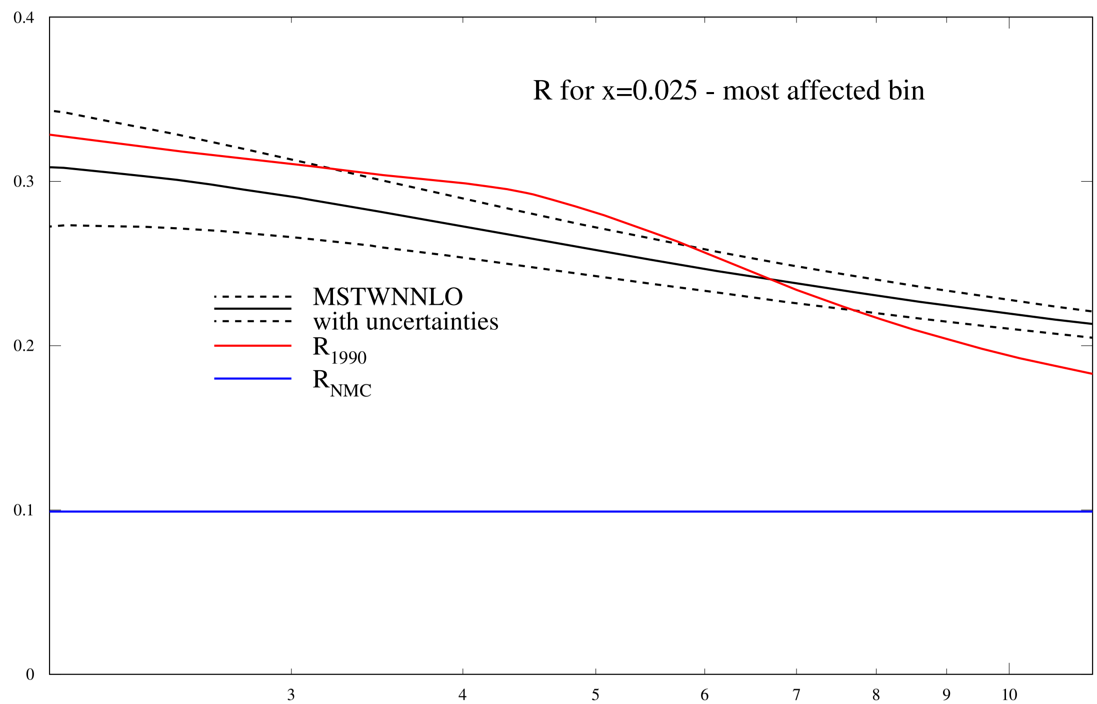
$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha_2}{xQ^4}\left[1 - y + \frac{y^2/2}{1+R(x,Q^2)}\right]F_2(x,Q^2)$$

In 1997 results used direct measurement of $R_{\rm NMC}(x)$ in $x$ bins for $x \leq 0.12$ (only one for each $x$ bin) to obtain $F_2(x,Q^2)$.

Using $R(x,Q^2)$ too small, as $R_{\rm NMC}$ often is leads to a smaller $F_2(x,Q^2)$, if $y$ is large.

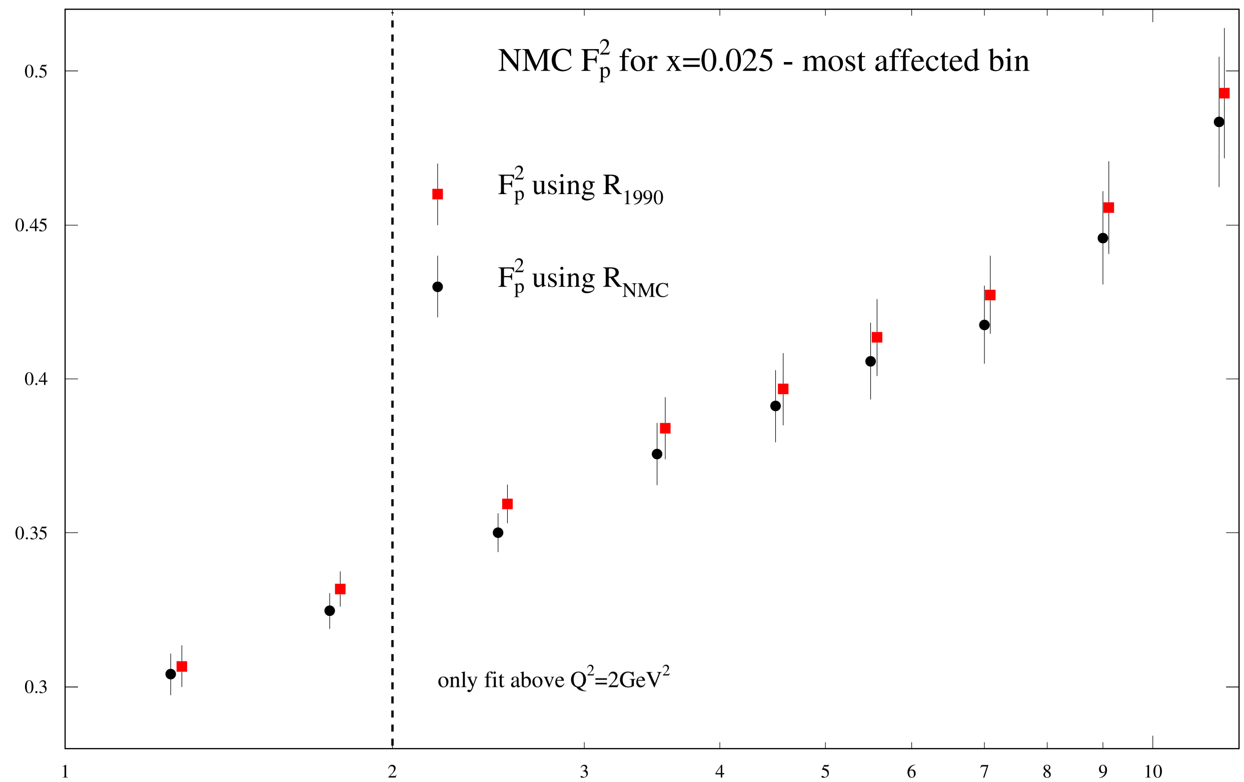
(Remember switch between 1995 and 1997 measurements when preparing MRST98 PDFs, and no real effect noticed).

Big different between $R_{\mathrm{NMC}}(x, Q^2)$ and $R_{1990}(x, Q^2)$ and $R_{\mathrm{QCD}}(x, Q^2)$ in some bins




Most consistent to fit to $\sigma$.

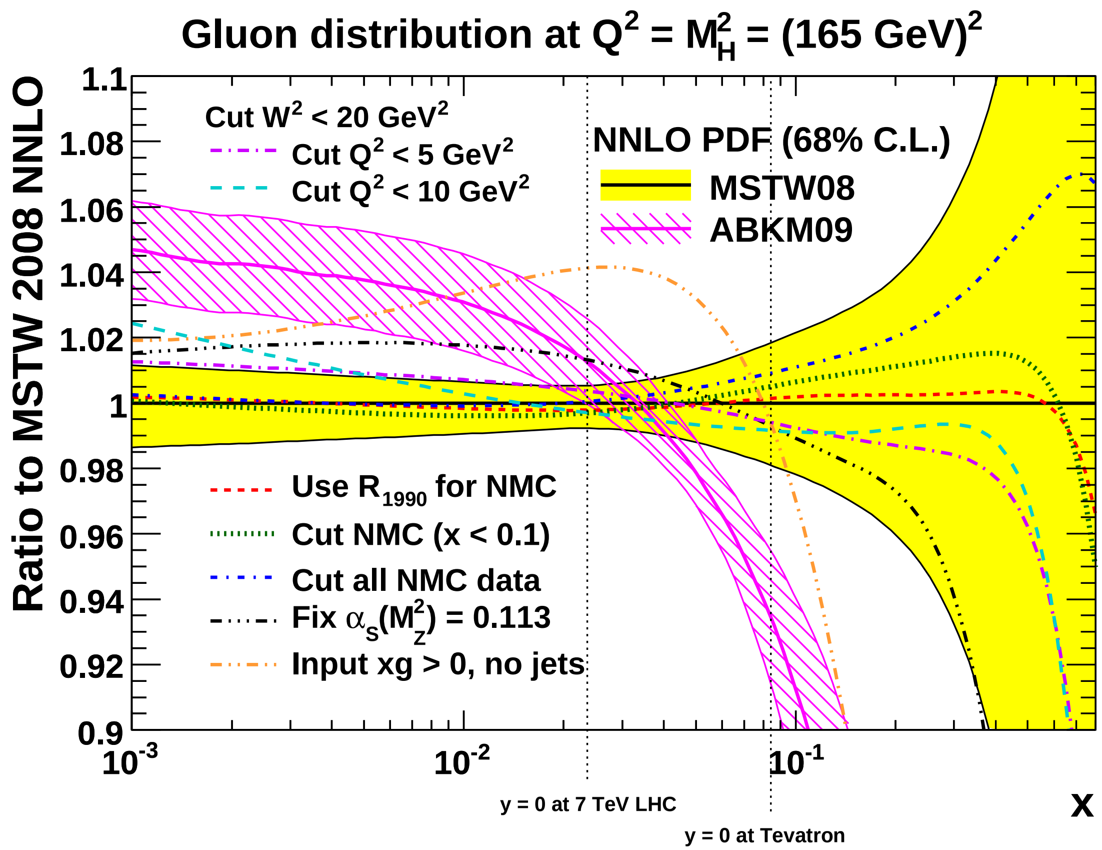
However, at NNLO $R_{\mathrm{MSTW}}(x, Q^2) \approx R_{1990}(x, Q^2)$ so using $F_2(x, Q^2)$ extracted using $R_{1990}(x, Q^2)$ very similar.




Because we use data averaged over energy bins effect not actually so large since $y$ is never large in all bins.

Show easily worst $x$ bin, i.e. $R_{\rm MSTW}(x,Q^2)$ and $R_{1990}(x,Q^2)$ very different, many points high $y$ and quite a lot of points survive $Q^2 \geq 2{\rm GeV}^2, W^2 \geq 15{\rm GeV}^2$.

Not much difference and we get $\Delta\alpha_S(M_Z^2) = 0.0012$ from a fit to $2500$ other data points.




Repeat global fit at NNLO changing the $F_2(x,Q^2)$ from NMC to that using $R_{1990}(x,Q^2)$, cutting NMC data sensitive to the change, cutting all NMC data changing $Q^2_{cut}$ up from $2\mathrm{GeV}^2$ (and losing much of the NMC data along with sensitivity to higher twist). None causes much change in the gluon.

Use the MSTW08 fit with $\alpha_S(M_Z^2) = 0.113$. More similar to ABKM09 gluon, but not all the way. Most similar, remove jet data from fit and use simpler gluon parameterisation ($4$ parameters) with one small-$x$ power.

| NNLO PDF | $\alpha_S(M_Z^2)$ | $\sigma_H$ Tevatron | $\sigma_H$ LHC (7 TeV) |
|---|---|---|---|
| **MSTW08** | **0.1171** | **0.342 pb** | **7.91 pb** |
| Use $R_{1990}$ for NMC | 0.1167 | $-0.7\%$ | $-0.9\%$ |
| Cut NMC ($x < 0.1$) | 0.1162 | $-1.2\%$ | $-2.1\%$ |
| Cut all NMC data | 0.1158 | $-0.7\%$ | $-2.1\%$ |
| Cut $Q^2 < 5$ GeV$^2$, $W^2 < 20$ GeV$^2$ | 0.1171 | $-1.2\%$ | $+0.4\%$ |
| Cut $Q^2 < 10$ GeV$^2$, $W^2 < 20$ GeV$^2$ | 0.1164 | $-3.0\%$ | $-1.7\%$ |
| Fix $\alpha_S(M_Z^2)$ | 0.1130 | $-11\%$ | $-7.6\%$ |
| Input $xg > 0$, no jets | 0.1139 | $-17\%$ | $-4.9\%$ |
| ABKM09 | 0.1135 | $-26\%$ | $-11\%$ |

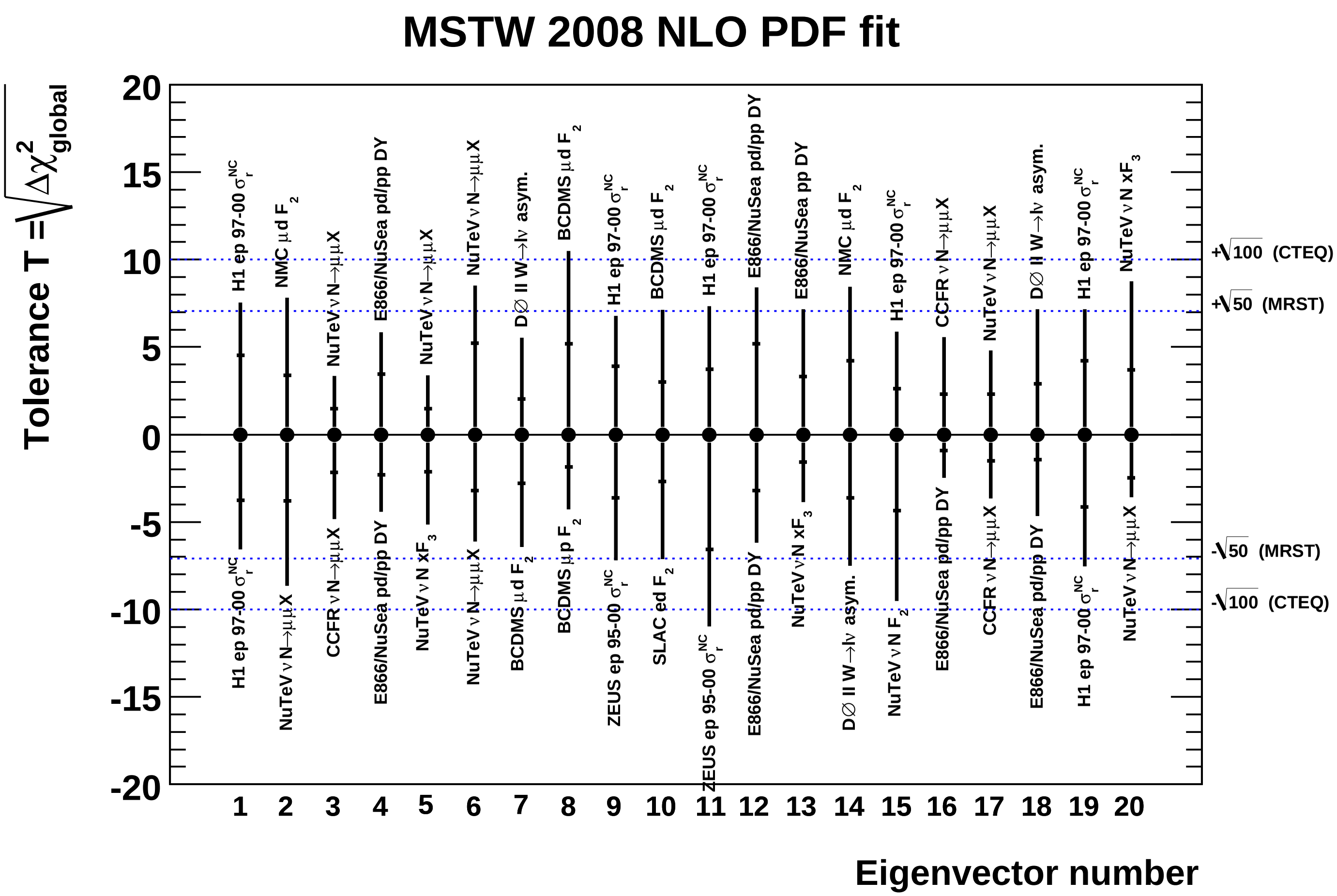
Change in $\alpha_S(M_Z^2)$ and Higgs production cross sections with fits outlined.

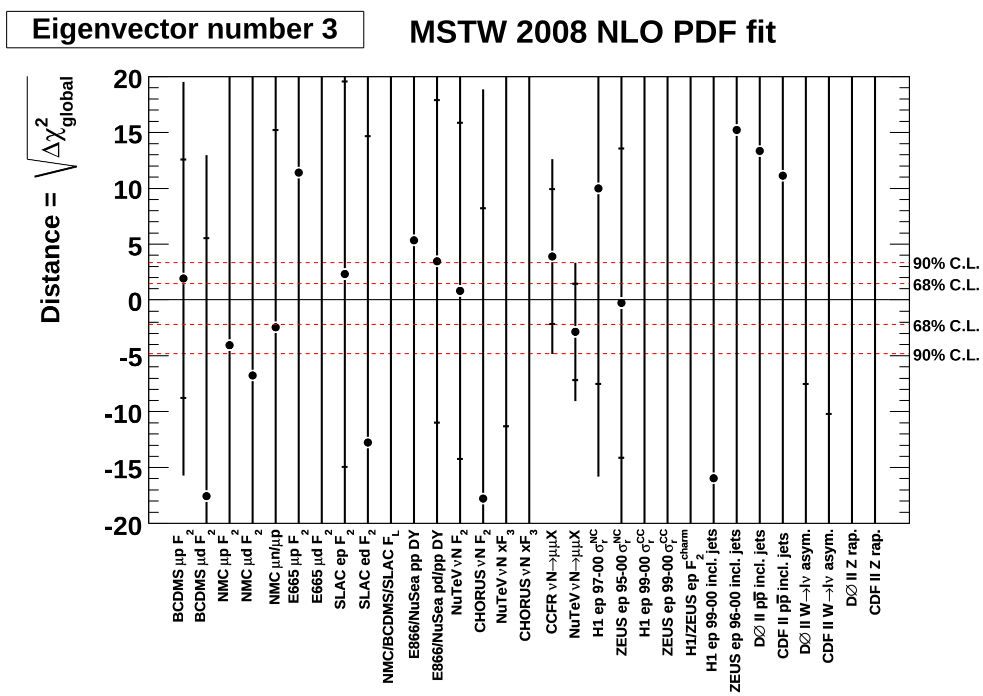
Only the imposition of $\alpha_S(M_Z^2) = 0.113$, and even more-so the fit with no jets and restricted parameterisation (which automatically gives $\alpha_S(M_Z^2) = 0.1139$) move much towards the ABKM09 values.

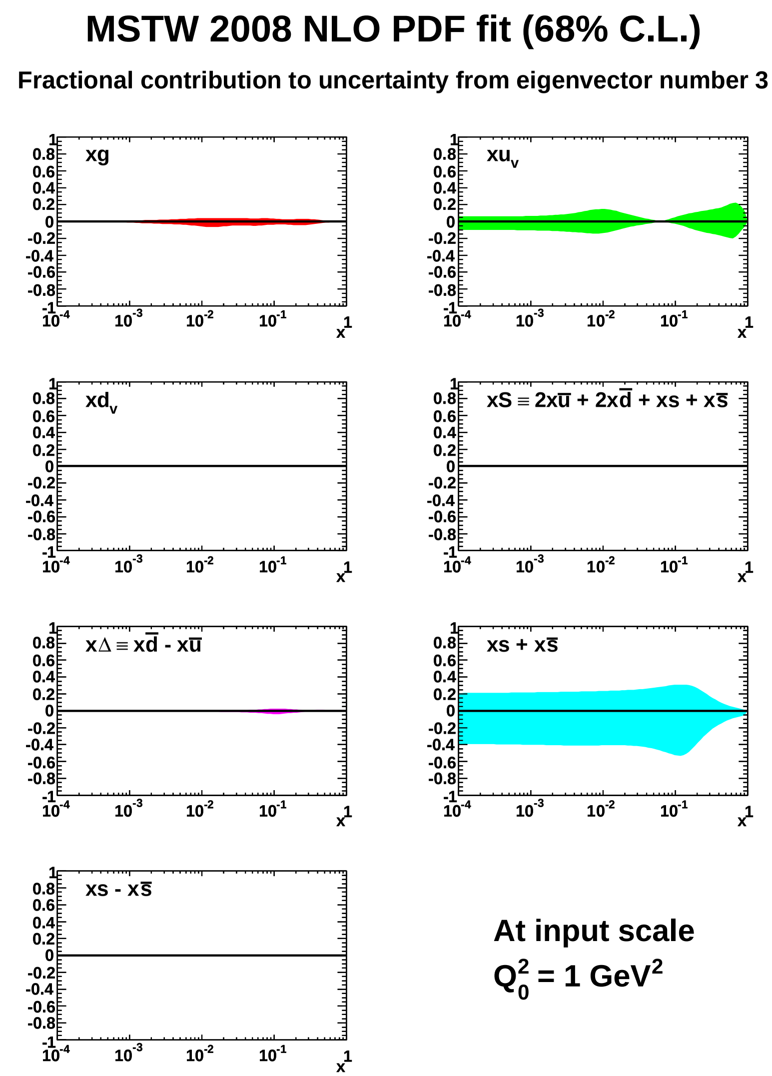
# Fits to only Collider data

Agreed for PDF4LHC to investigate the consequences of fitting to data sets not including any fixed-target data.

in MSTW08 PDFs have 20 eigenvectors and have to determine uncertainty on each. Constraining sets not always collider data.

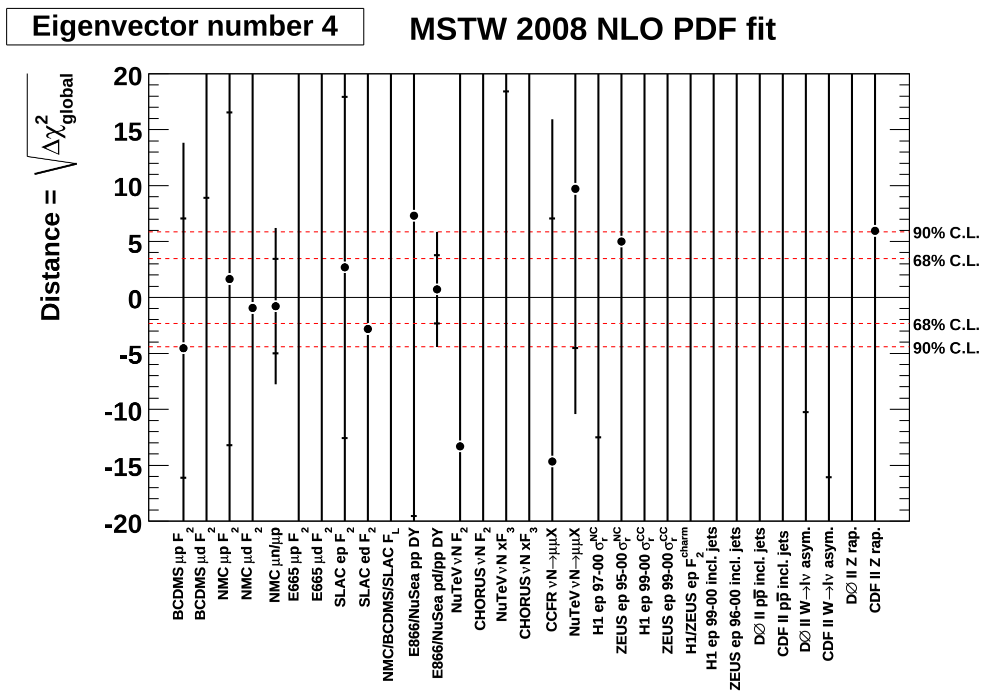


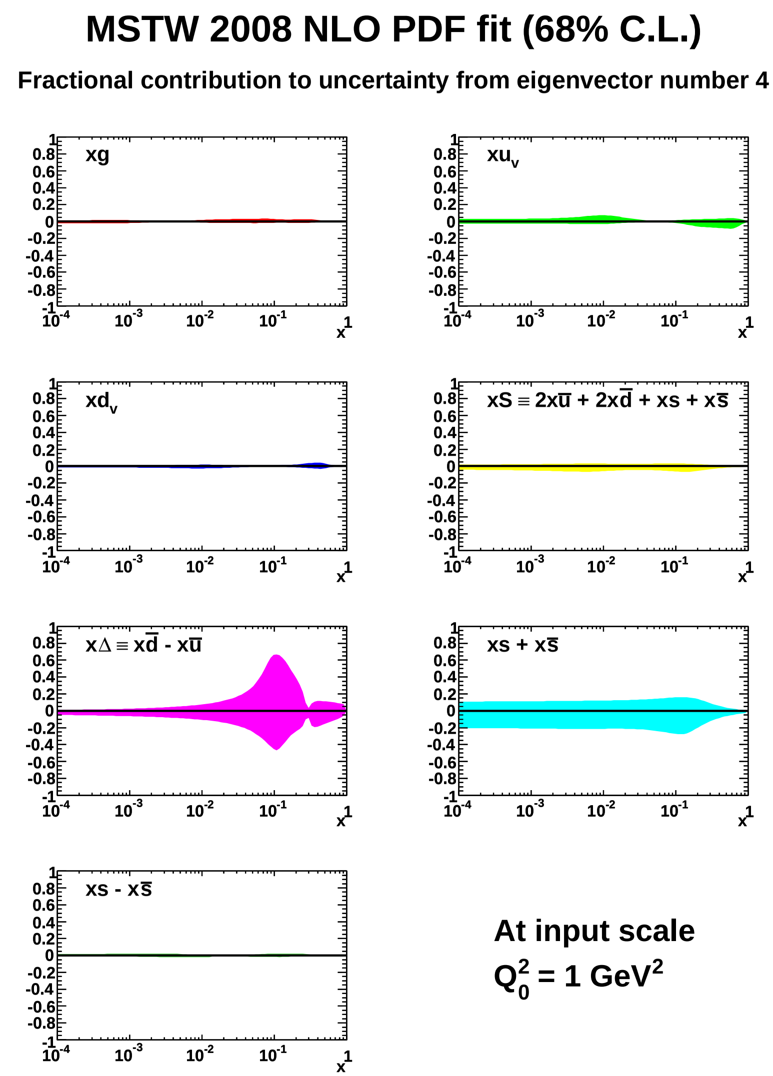


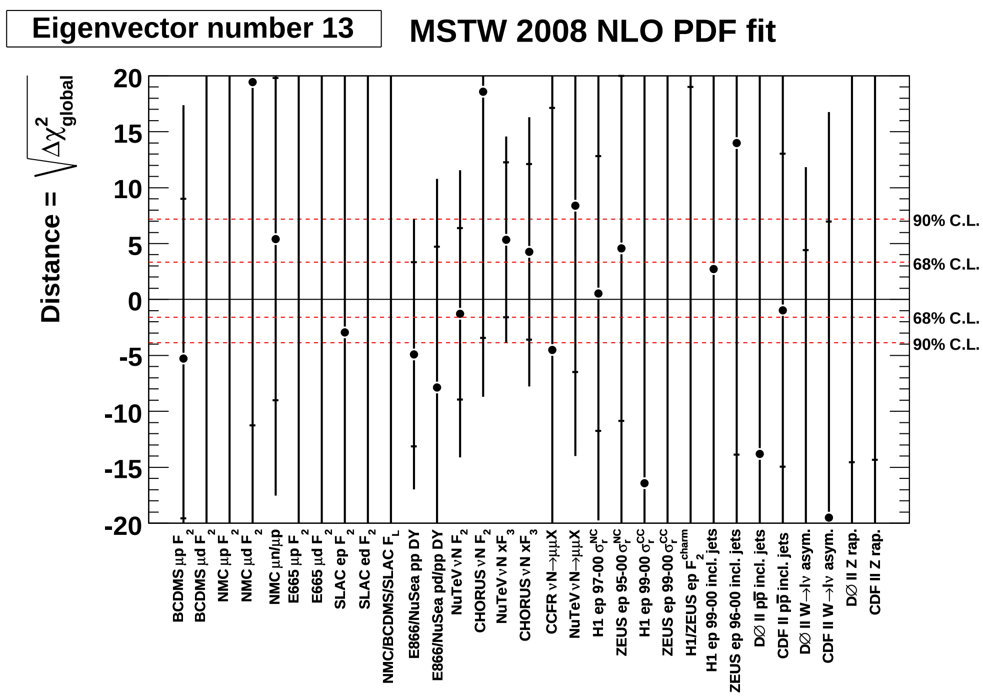

In MSTW08 fits see constraint on each eigenvector from different data sets.

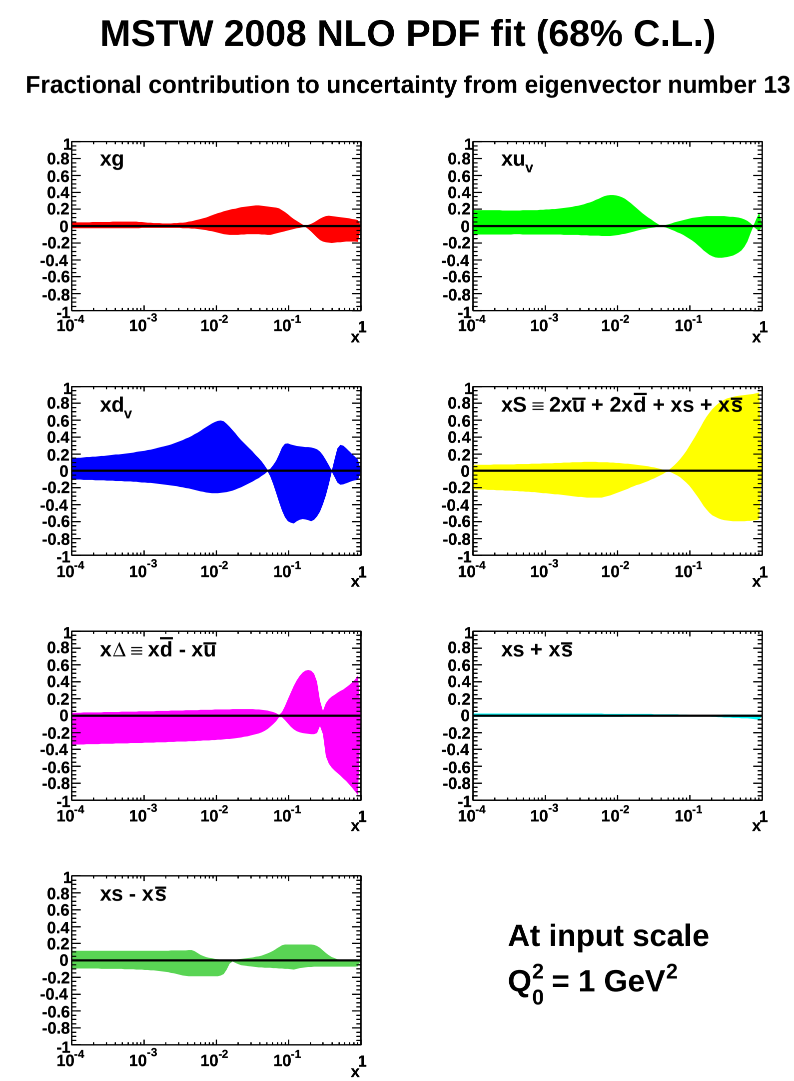
Eigenvector 3 has only a very weak, asymmetric constraint from collider data. Without dimuon data weak constraint on strange normalisation and push downwards.




Eigenvector 4 has almost no constraint from collider data. $\bar{d} - \bar{u}$ and again strange.




Eigenvector 14 has only a very weak, asymmetric constraint from collider data.

High-$x$ sea and flavours weakly pulled from default by Tevatron asymmetry data.

## Change in PDFs for fit to only collider data.

For default simply repeat MSTW2008 NLO fit with HERA structure function data updated to combined data.

Then try replacing default D0 low luminosity muon asymmetry data with higher luminosity electron data (in combined $p_T$ bin), which is the most constraining published asymmetry data (at present).

Fit quality to 1053 data points improves by $\Delta\chi^2 \sim -120$.

Improvement of $30$ in lepton asymmetry data and $47$ in HERA inclusive structure function data.

Small improvement in jet (Tevatron and HERA) data.

To avoid pathological behaviour have to fix some parameters though. $s - \bar{s}$ fixed otherwise negative quarks, and $(1 - x)$ power of strange also fixed to be same as averaged sea.

In eigenvectors also need one more fixed parameter in valence quarks to avoid redundancy of parameters. $\rightarrow 16$ rather than $20$ eigenvectors.

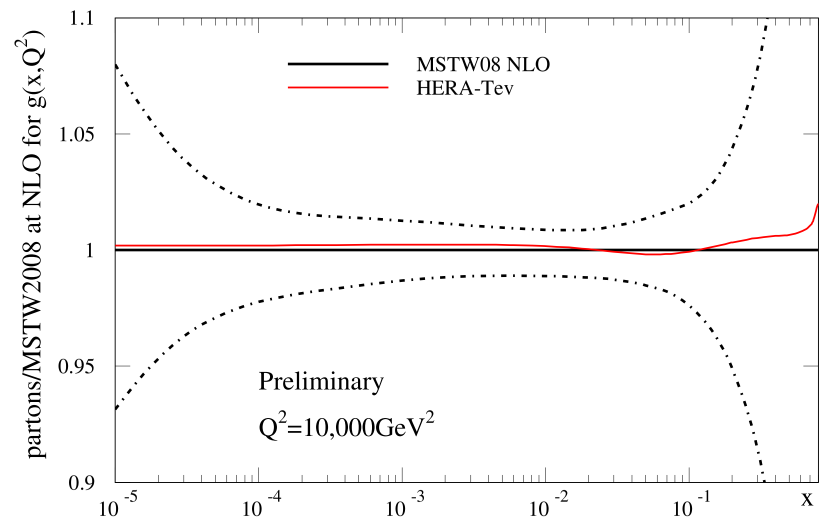
$\alpha_S(M_Z^2) = 0.1193$, but related to behaviour of strange sea. Uncertainty about $0.0025$.

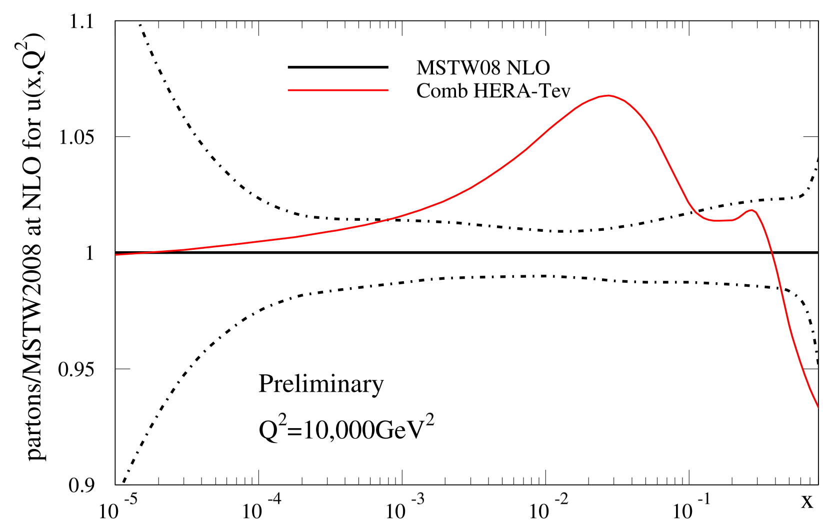
Almost no change in gluon distribution. All major constraints present.

Marginal improvement in jets and big improvement in HERA data due to quark changes.

Large increase in $u(x,Q^2)$ for $x \sim 0.02$ compensated by other quarks.

$(1-x)$ power of sea, usually constrained by Drell Yan and neutrino structure function data at least $5$ times as uncertain (with constraint from HERA charged current data and lepton asymmetry).

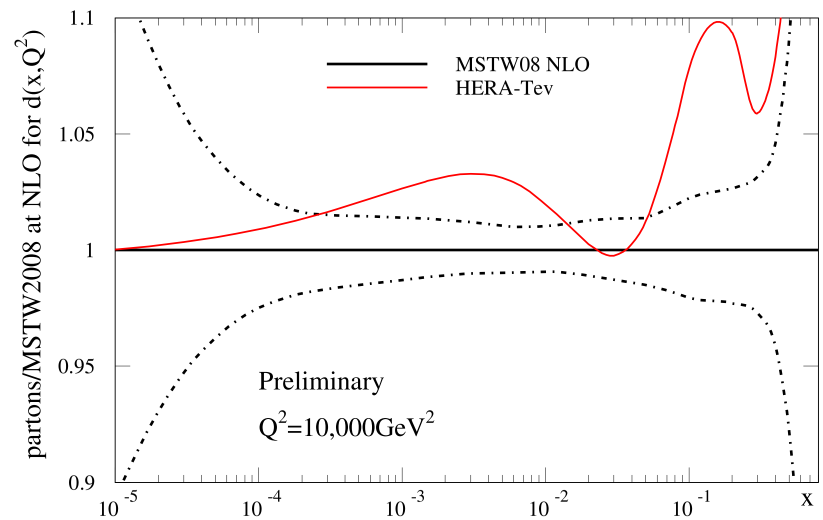


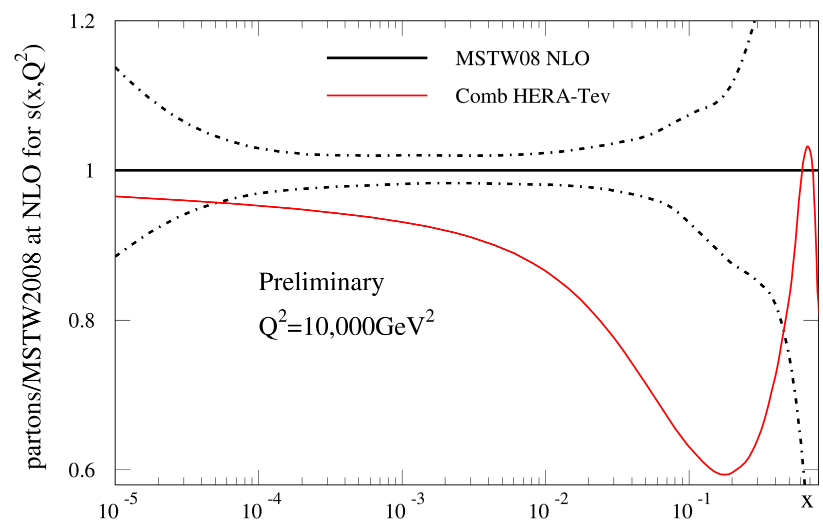



Not too much change in $d(x,Q^2)$ except near $x=0.2$.

However, $\int (\bar{d}(x,Q_0^2) - \bar{u}(x,Q_0^2))\, dx$ $1.5$ times bigger but with uncertainty similar to magnitude (normally $\sim 15\%$. Constrained by HERA $F_2(x,Q^2)$ (down) and CDF jets (up, rather that E866 DY ratio.

Input strange normalisation at $Q_0^2 = 1\mathrm{GeV}^2$ about zero. Can vary up to about $40\%$ of input sea (constrained by HERA $F_2(x,Q^2)$) and if allowed down to $-30\%$. Normally about $30\%$ of input sea with $\sim 15\%$ uncertainty, constrained by dimuon data from NuTeV and CCFR.

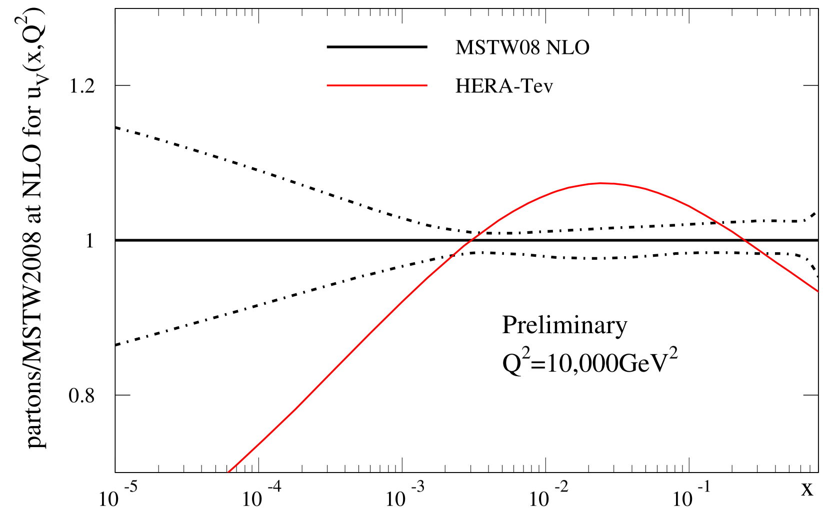


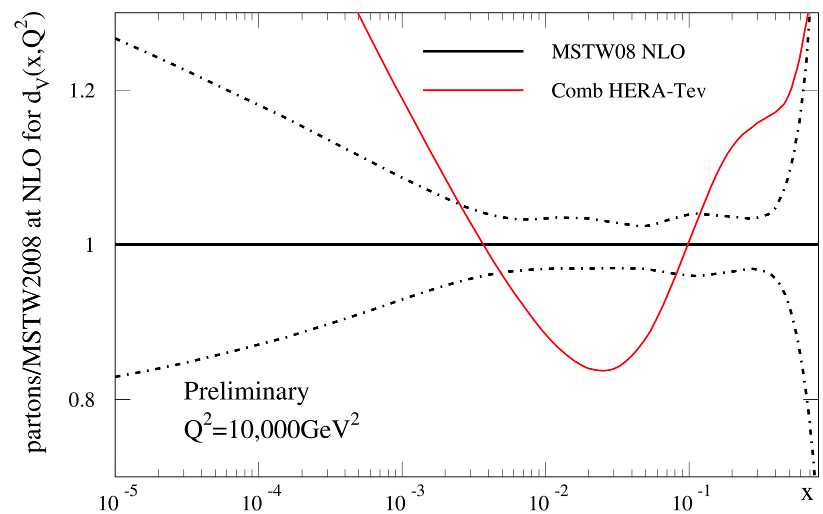



Both valence distributions much changed. Particularly at small $x$.

High-$x$ $u_V(x,Q^2)$ at least 3 times as uncertain. Constrained by HERA charged current data and lepton asymmetry – usually by most fixed target data. At low $x$ similar, with sensitivity to Tevatron rapidity data.

$d_V(x,Q^2)$ at least twice as uncertain. Constrained by HERA charged current data, lepton asymmetry and to some extent Tevatron jet data at high $x$. Usually by deuterium data and to some extent lepton asymmetry.



When the D0 asymmetry data is swapped the initial prediction is not very good.

Before refit over $40/12$ (good fit for $20/12$, i.e. lots of scatter). Actually worse than predictions made by variety of fits to all data where some deuterium corrections has been fit/modelled.

However, consistent within large uncertainties, i.e. very little change in quality of fit to other data in refit (mainly in D0 jets).

Few percent change in PDFs, almost all in high-$x$ ($x > 0.1$) $d_v(x, Q^2)$ (down) and sea quarks (up).

% change in cross sections for collider only fit ($M_H = 165\text{GeV}$).

| | MSTW comb HERA | D0$_\mu$ $0.4fb^{-1}$ | D0$_{\rm el}$ comb $p_T$ |
|---|---|---|---|
| $W$ Tev | +3.1 | +5.0 | +4.8 |
| $Z$ Tev | +3.0 | +5.9 | +5.5 |
| $W/Z$ Tev | +0.2 | -0.9 | -0.6 |
| $W^+/W^-$ Tev | +0.0 | +0.1 | +0.1 |
| $W$ LHC (7TeV) | +2.9 | +2.7 | +2.5 |
| $Z$ LHC (7TeV) | +2.7 | +2.3 | +1.9 |
| $W/Z$ LHC (7TeV) | +0.2 | +0.5 | +0.6 |
| $W^+/W^-$ LHC (7TeV) | +0.1 | 1.3 | 1.9 |
| $W$ LHC (14TeV) | +2.4 | +1.5 | +1.3 |
| $Z$ LHC (14TeV) | +2.5 | +0.9 | +0.9 |
| $W/Z$ LHC (14TeV) | -0.1 | +0.6 | +0.5 |
| $W^+/W^-$ LHC (14TeV) | -0.5 | +1.3 | +1.3 |
| Higgs TeV | -1.4 | -1.4 | +0.2 |
| Higgs LHC (7TeV) | +0.4 | -1.6 | -1.0 |
| Higgs LHC (14TeV) | +1.0 | -1.5 | -1.1 |

Changes not that large. In total cross sections largely due to inclusion of combined HERA data, (smaller at NNLO or if normalisation constraint relaxed). Changes greater than uncertainties in ratios. Mainly in $W/Z$ ratio due to change in strange quarks.

## Details from single charged-lepton cross sections and asymmetries – Stirling

Asymmetries can however, be very useful in relating features to PDFs.

$$A_W(y_W) = \frac{\mathrm{d}\sigma(W^+)/\mathrm{d}y_W - \mathrm{d}\sigma(W^-)/\mathrm{d}y_W}{\mathrm{d}\sigma(W^+)/\mathrm{d}y_W + \mathrm{d}\sigma(W^-)/\mathrm{d}y_W}$$

$$A_W(y_W) \approx \frac{u_V(x_1)S(x_2) + S(x_1)u_V(x_2) - d_V(x_1)S(x_2) - S(x_1)d_V(x_2)}{u_V(x_1)S(x_2) + S(x_1)u_V(x_2) + d_V(x_1)S(x_2) + S(x_1)d_V(x_2) + 4S(x_1)S(x_2)}$$

where $x_{1,2} = (M_W/\sqrt{s})\exp(\pm y_W)$ and $S(x) \approx \bar{u}(x) \approx \bar{d}(x)$. In particular at $y = 0$, where $x_0 = (M_W/\sqrt{s})$,

$$A_W(0) \approx \frac{u_V(x_0) - d_V(x_0)}{u_V(x_0) + d_V(x_0) + 2S(x_0)} \approx \frac{u_V(x_0) - d_V(x_0)}{u(x_0) + d(x_0)},$$

i.e. direct probe of valence quark difference (Cooper-Sarkar), the total quark distributions $u(x_0), d(x_0)$ being well-constrained by (mainly) HERA data.

However, really measure

$$A(y_\ell) = \frac{\mathrm{d}\sigma(\ell^+)/\mathrm{d}y_\ell - \mathrm{d}\sigma(\ell^-)/\mathrm{d}y_\ell}{\mathrm{d}\sigma(\ell^+)/\mathrm{d}y_\ell + \mathrm{d}\sigma(\ell^-)/\mathrm{d}y_\ell},$$

If $\theta^*$ is angle of lepton to proton beam $\cos^2\theta^* = 1 - 4p_T^2/M_W^2$ and

$$y_\ell = y_W + y^*, \quad y^* = \frac{1}{2}\ln\left(\frac{1+\cos\theta^*}{1-\cos\theta^*}\right)$$

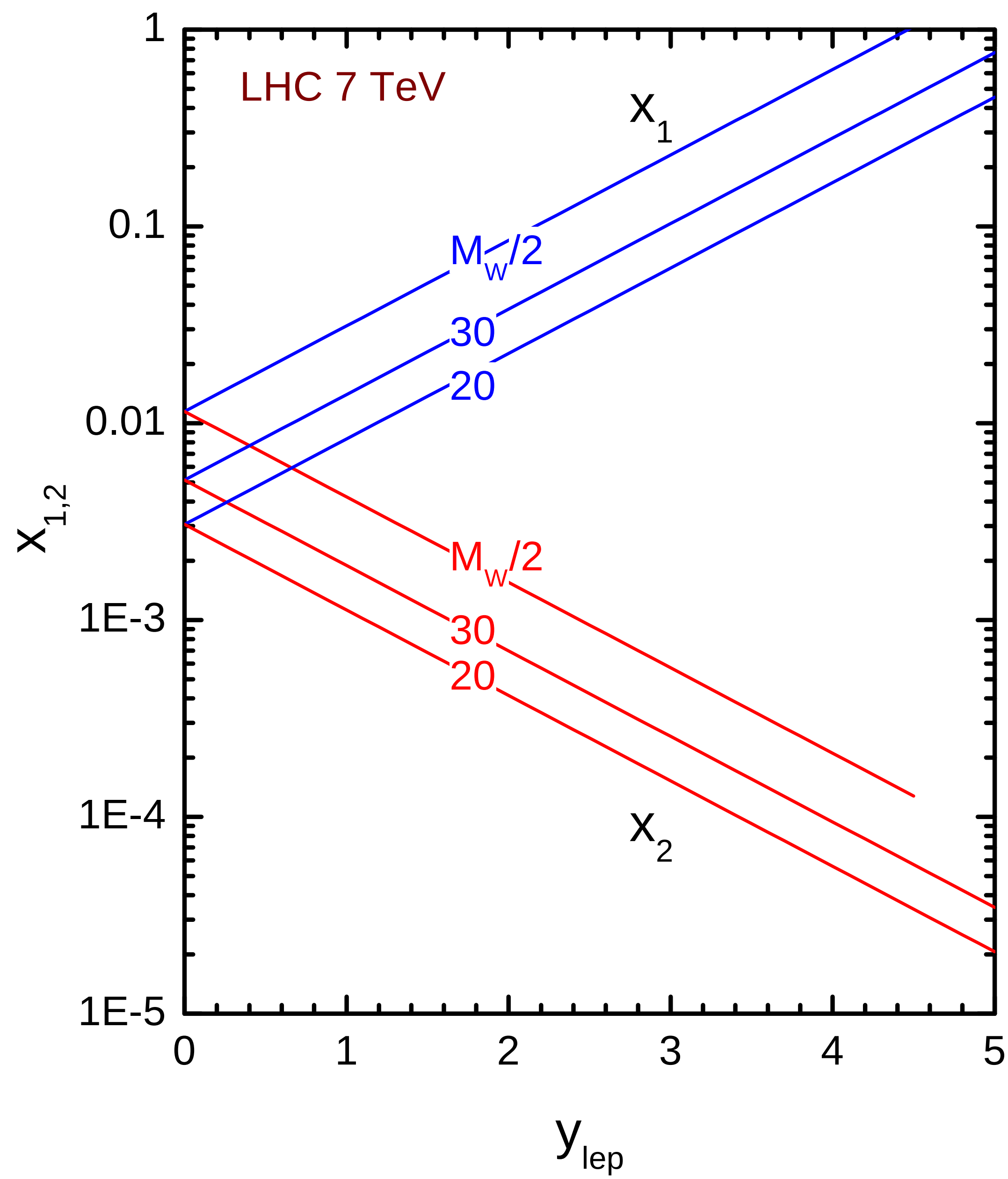
$$x_{1,2} = x_0 \exp(\pm y_W) = x_0 \exp(\pm y_\ell)\kappa^{\pm 1}, \quad \kappa = \left(\frac{1+|\cos\theta^*|}{1-|\cos\theta^*|}\right)^{1/2}$$

$$x_1^+ = x_0 \exp(+y_\ell)\kappa > x_1^- = x_0 \exp(+y_\ell)\kappa^{-1} \quad x_2^+ = x_0 \exp(-y_\ell)\kappa > x_2^- = x_0 \exp(-y_\ell)\kappa^{-1}.$$

Since all PDFs decrease with increasing $x$, the $x_{1,2}^-$ contributions are numerically dominant, particularly as we go away from central rapidity.

Values of generally dominant $x^-$ values probed shown opposite.

For $p_T = 20\mathrm{GeV}$ a factor of $3$ from naive estimate.



Ignoring sea-sea contributions

$$\frac{\mathrm{d}\sigma(\ell^+)}{\mathrm{d}y_\ell} \propto \left(u_V(x_1^+)S(x_2^+) + u_V(x_1^-)S(x_2^-)\right)(1-\cos\theta^*)^2 + \left(S(x_1^+)u_V(x_2^+) + S(x_1^-)u_V(x_2^-)\right)(1+\cos\theta^*)^2$$

$$\frac{\mathrm{d}\sigma(\ell^-)}{\mathrm{d}y_\ell} \propto \left(S(x_1^+)d_V(x_2^+) + S(x_1^-)d_V(x_2^-)\right)(1-\cos\theta^*)^2 + \left(d_V(x_1^+)S(x_2^+) + d_V(x_1^-)S(x_2^-)\right)(1+\cos\theta^*)^2$$

As $p_T$ deviates further from $M_W/2$, $(1+\cos\theta^*)^2$ dominates.

For large $y_\ell$, $x_1^- \gg x_2^-$ and the asymmetry becomes more and more dominated by the $d_V(x_1^-)S(x_2^-)$ contribution to $l^-$ and $A_\ell(y_\ell) \to -1$.

$d_V(x)$ decreases faster at large $x$ than $u_V(x)$, and so at some point at large $y_\ell$ the approximation

$$d_V(x_1^-)S(x_2^-)(1+\cos\theta^*)^2 \gg u_V(x_1^-)S(x_2^-)(1-\cos\theta^*)^2$$

breaks down, i.e. the $V \pm A$ unfavoured forward $u\bar{d} \to \ell^+\nu_\ell$ scattering process will eventually dominate, and $A_\ell(y_\ell) \to +1$

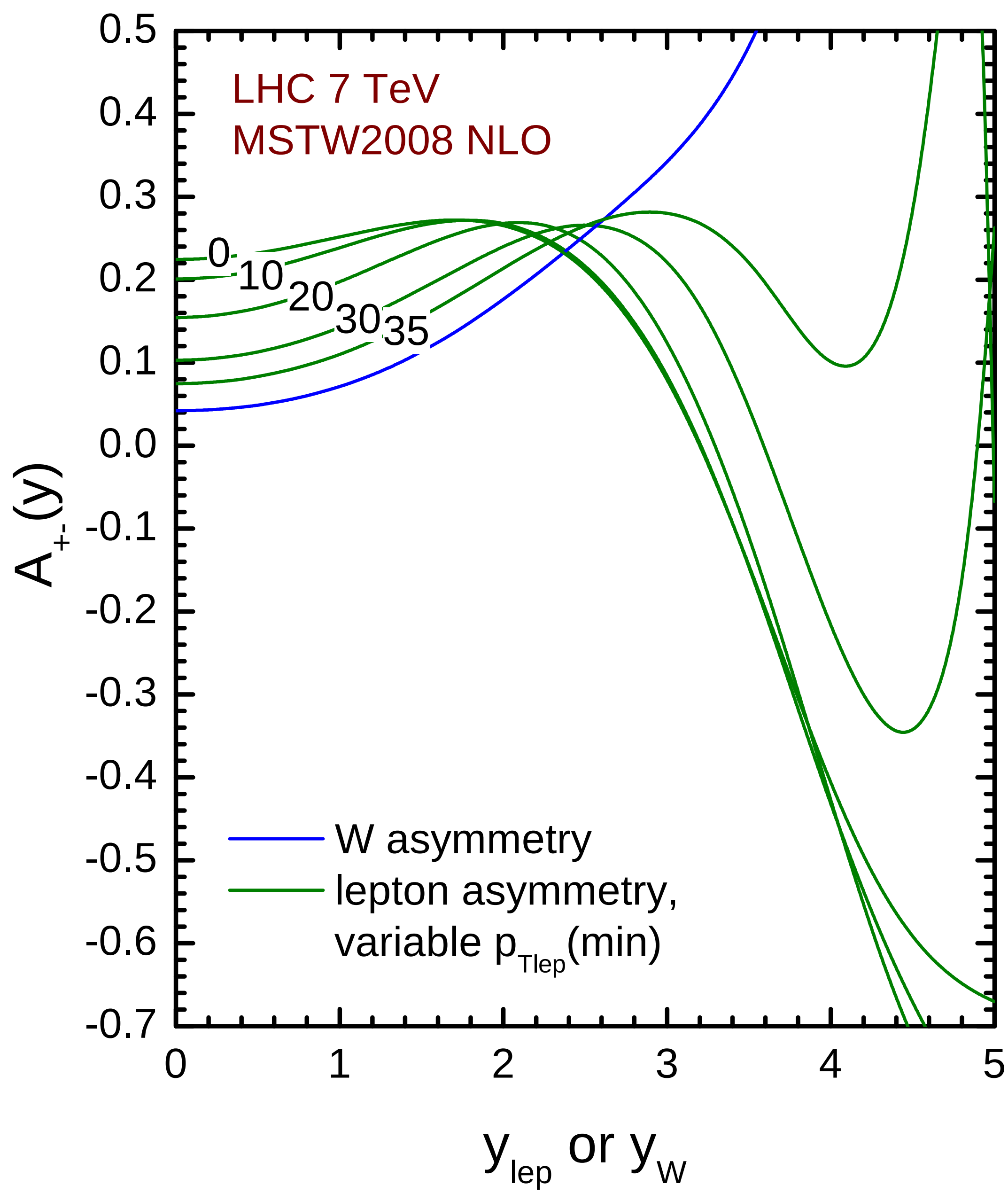
This will happen at the $y_\ell$ value for which

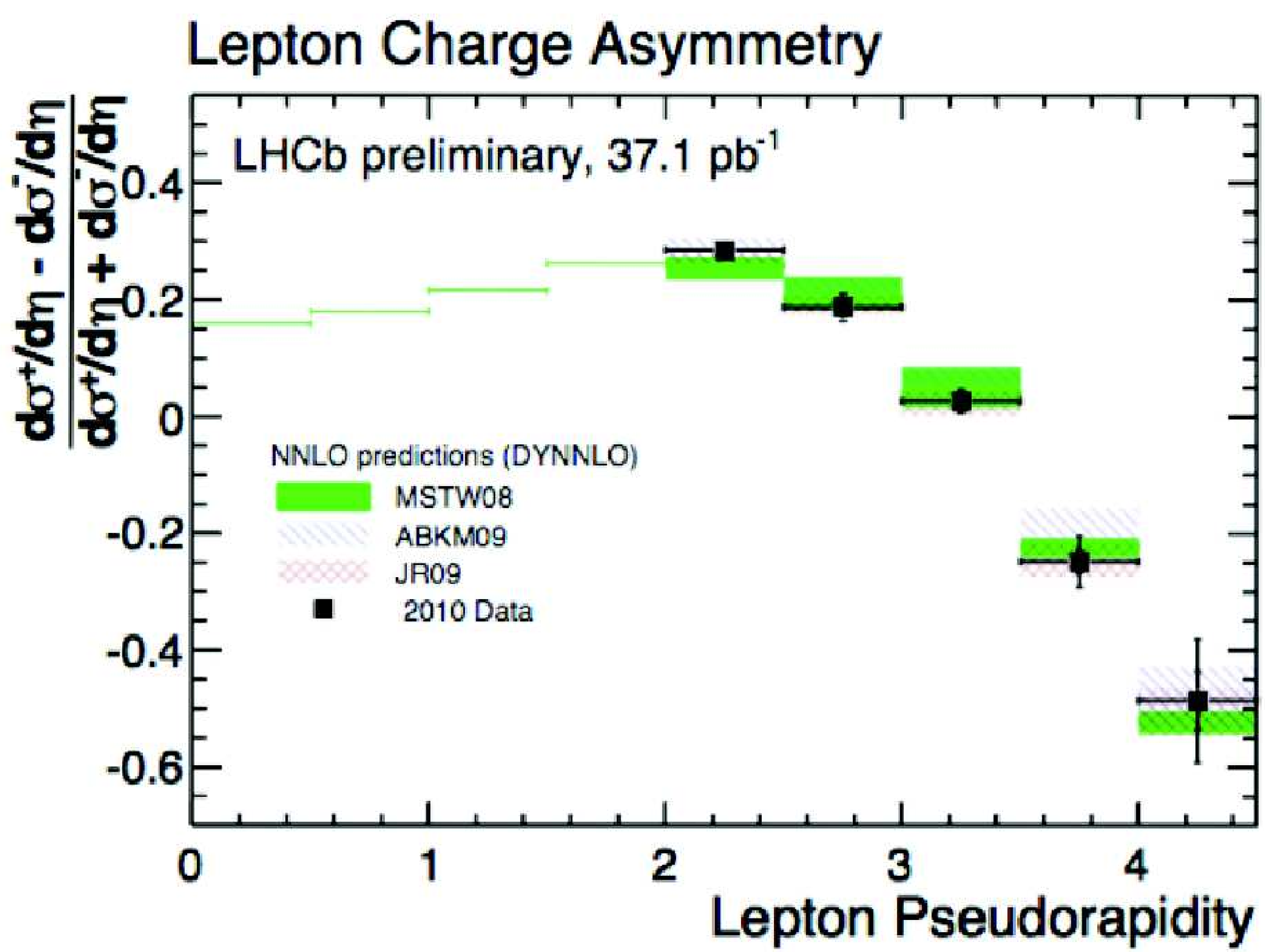
$$u_V(x_1^-)/d_V(x_1^-) \sim (1+\cos\theta^*)^2/(1-\cos\theta^*)^2 = \kappa^4.$$

The larger the lepton $p_T$ the earlier (in terms of increasing $y_\ell$) this will happen, and for $p_T \rightarrow m_W/2$ there is no $V \pm A$ dominance at all.

So asymmetry at large $y_\ell$ in terms of $p_T$ tells us about $d/u$ at large $x$.

It also confirms that at high $x$, $S(x) \ll u_V(x)(d_V(x))$, since dip is diminished by (approximately) equal sea-sea contribution to $\ell^+$ and $\ell^-$.






LHCb (with $p_T(\mathrm{min}) = 20\mathrm{GeV}$) already testing dip.

With higher $p_T(\mathrm{min})$ could potentially see upturn.

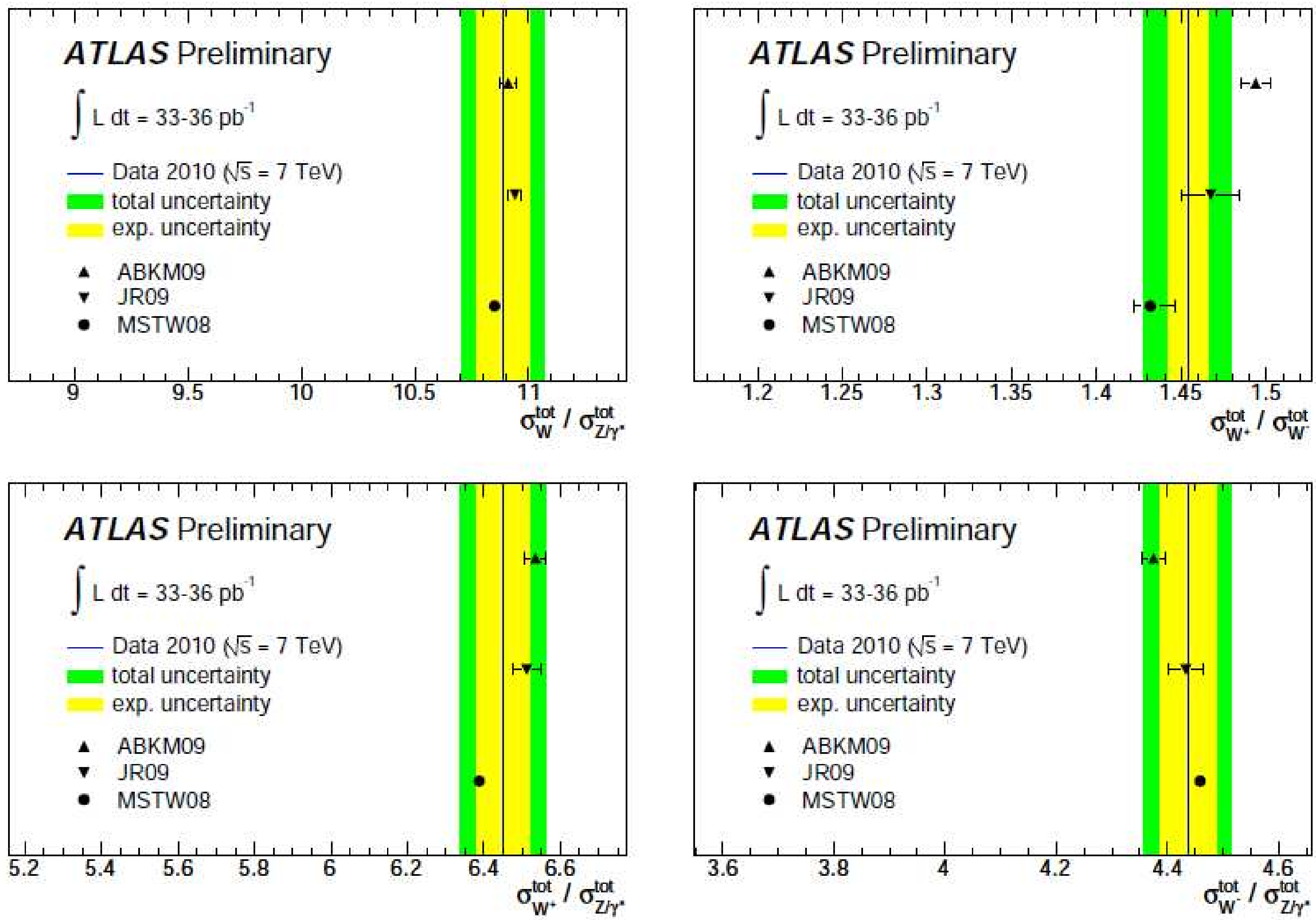
## Summary

MSTW/MRST have been providing NNLO PDFs for over 10 years. Have remained largely stable over this time. Stable under change in kinematic cuts. Stable to change in treatment of NMC data, though should be more consistent in future. Include all types of data other than HERA jets at NNLO. Only obvious significant reason for update is inclusion of new data. Prelim. sets with combined HERA data show only small changes.

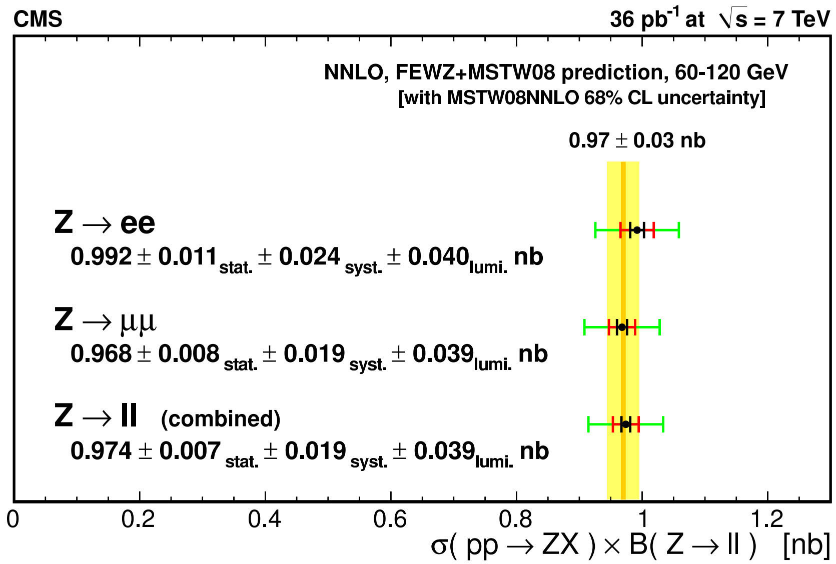
Fits to only collider data require some constraints to stop extreme central behaviour and variation in some PDFs. Uncertainties much bigger, though changes small for gluon. Total cross-sections fairly stable to change in fit, particularly ay the LHC, perhaps because dominated by evolution driven by gluon, but even ratios fairly similar to default. More change at Tevatron.

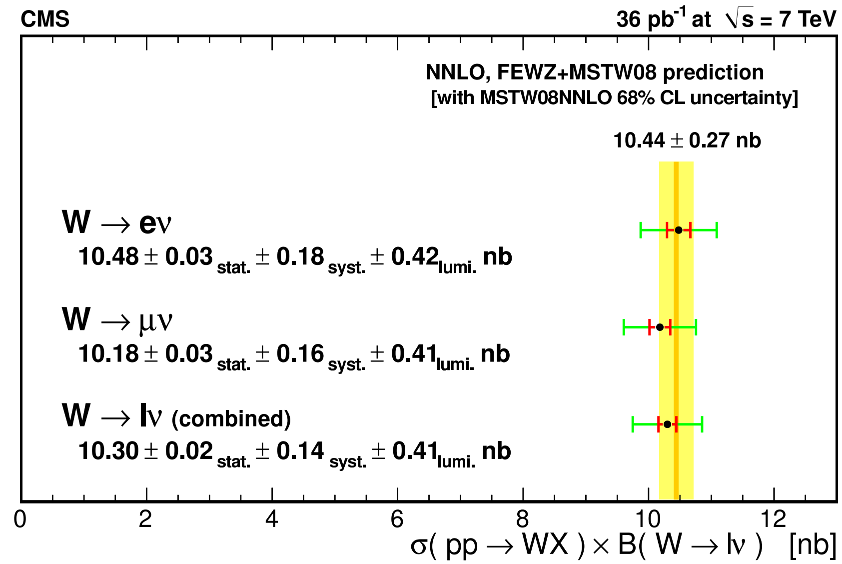
Personal opinion, better to include all data, e.g. if we want some constraint on strange normalisation it is from rather "unclean" nuclear target data - but take proper account of full uncertainties rather than make assumptions. Similar with deuterium data.

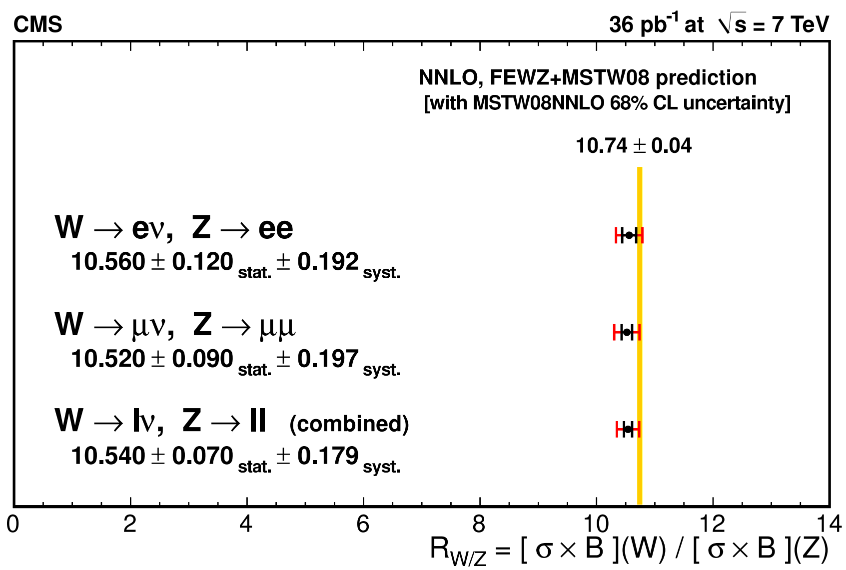
Variation with $p_T$ cuts of very high $y$ asymmetry data gives details on dominance of $u_V$ over $d_v$ at very high $x$.

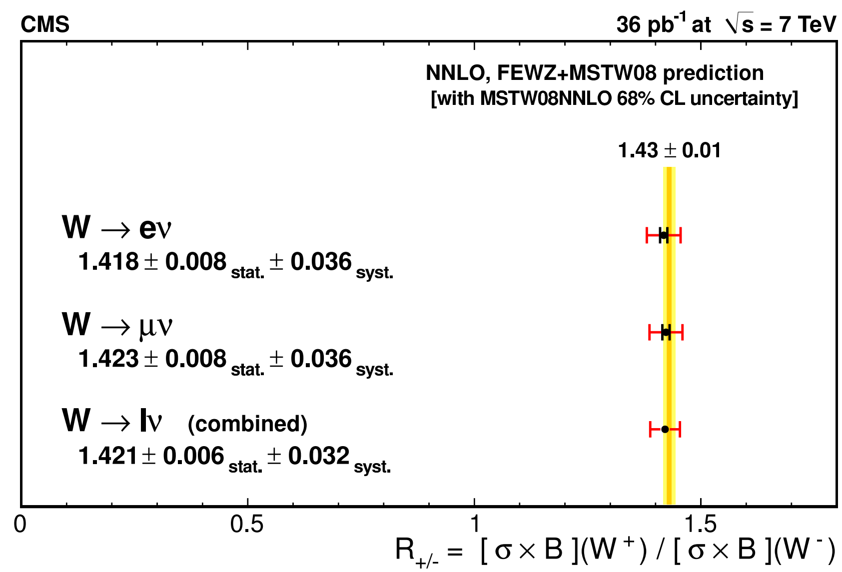



ATLAS results.

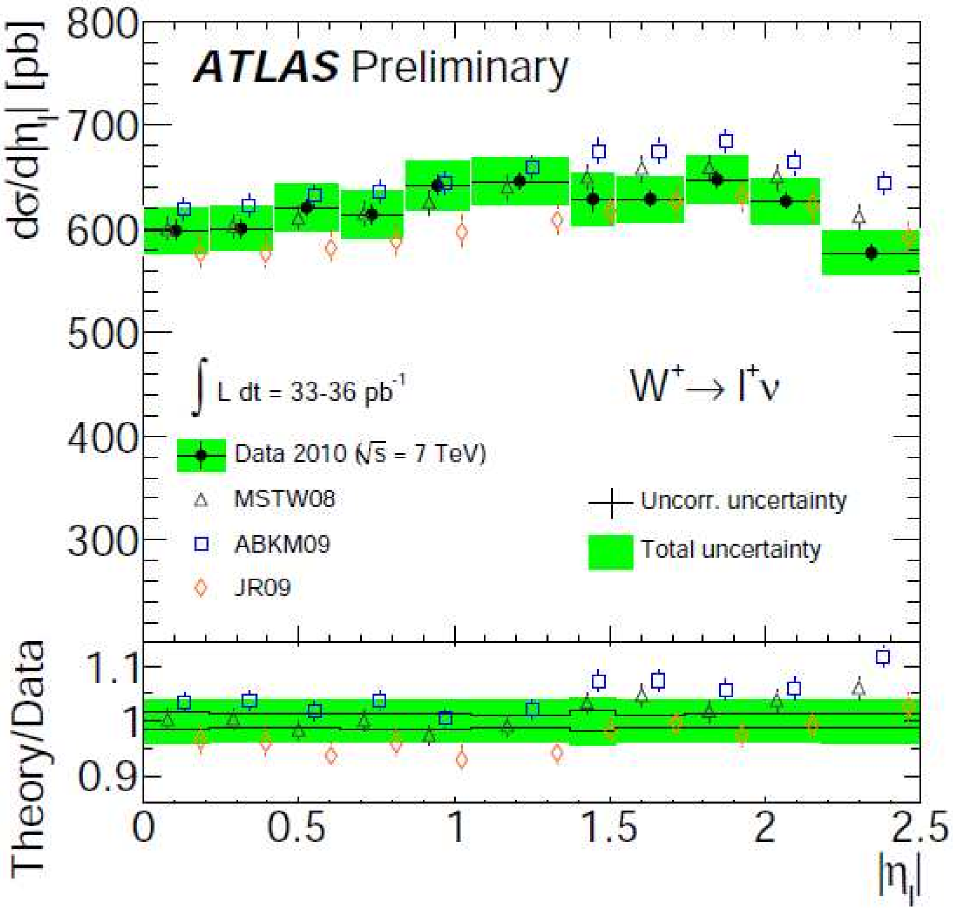


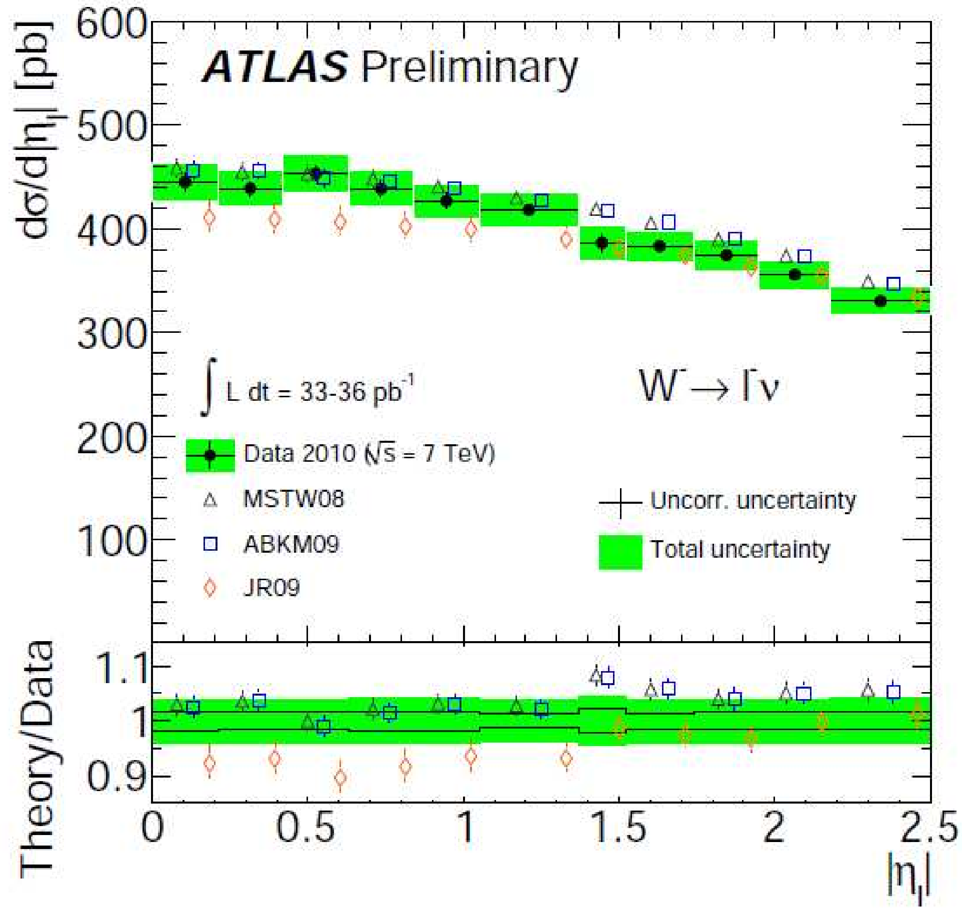



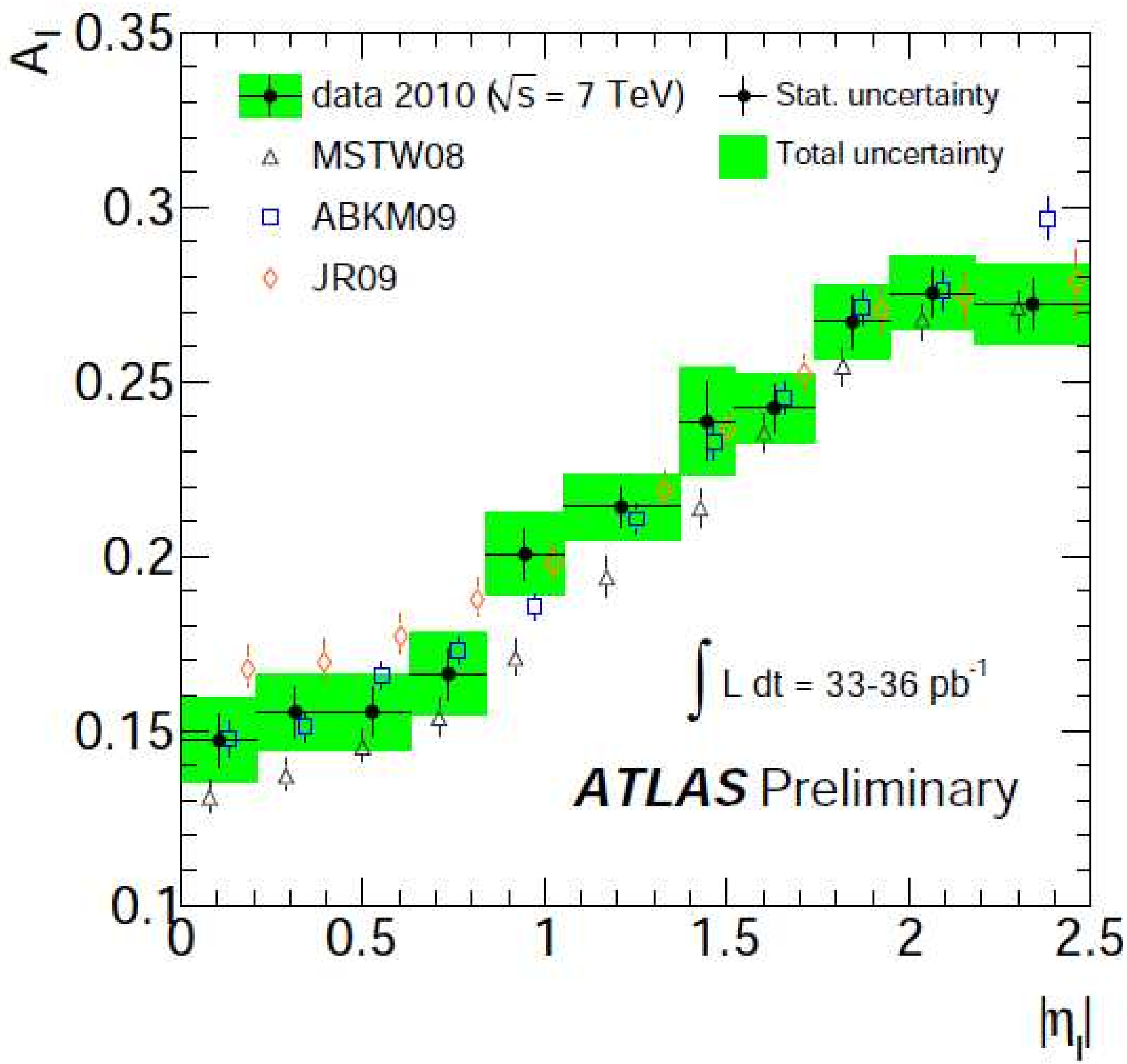




CMS results very similar.




Differential data on rapidity is becoming very constraining – on both shapes and on normalisations of predictions.




Clearly some of this information lost in ratios and asymmetries.

Ideally want individual distributions, with full correlations.