Experimental error propagation: Hessian versus Monte Carlo

Graeme Watt

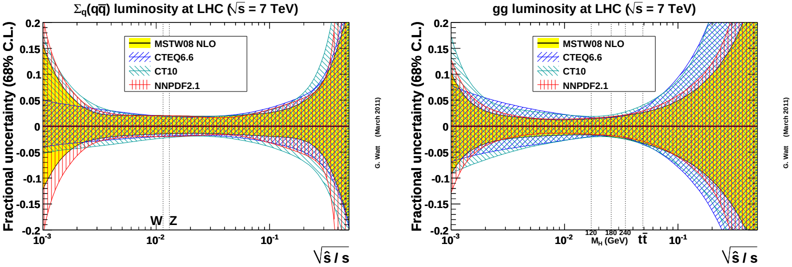
CERN PH-TH

PDF4LHC Meeting
CERN, Geneva, 28th November 2011

(With thanks to Robert Thorne for discussions.)

Introduction

- Motivation of adding LHC data to global PDF fits is to reduce (experimental) PDF uncertainties. Are these well-defined?
- Compare $q\bar{q}$ and gg uncertainties from NLO global fits:



- Different groups fitting similar data ⇒ similar uncertainties.
- Everything looks consistent on the surface. Is it really so?

Hessian method (used by MSTW and CTEQ groups)

- Assume χ^2 is quadratic about the global minimum $\{a_i^0\}$:

$$\Delta\chi^2 \equiv \chi^2 - \chi^2_{\rm min} = \sum_{i,j=1}^{n} H_{ij}\,(a_i - a_i^0)\,(a_j - a_j^0).$$

- Convenient to diagonalise covariance matrix $C \equiv H^{-1}$:

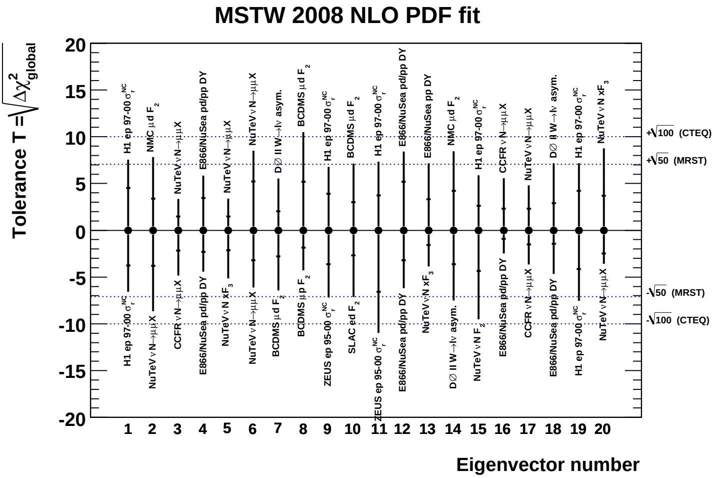
$$\sum_{j=1}^{n} C_{ij}\, v_{jk} = \lambda_k v_{ik}.$$

- Generate eigenvector PDF sets $S_k^\pm$ with parameters given by

$$a_i(S_k^\pm) = a_i^0 \pm t\,\sqrt{\lambda_k}\, v_{ik},$$

with t adjusted to give the desired tolerance $T = \sqrt{\Delta\chi^2}$.
- Non-standard values of $T > 1$ used to accommodate minor data inconsistencies (average $T \approx 3$ in **MSTW08** fit).

Dynamic tolerance: different for each eigenvector



- Outer (inner) error bars give tolerance for 90% (68%) C.L.

Monte Carlo sampling using data replicas (used by NNPDF)

- Generate replica data sets with shifted central values:

$$D_{m,i} \to \left(D_{m,i} + R^{\text{uncorr.}}_{m,i}\, \sigma^{\text{uncorr.}}_{m,i} + \sum_{k=1}^{N_{\text{corr.}}} R^{\text{corr.}}_{m,k}\, \sigma^{\text{corr.}}_{m,k,i} \right) \cdot \left(1 + R^{\mathcal{N}}_m\, \sigma^{\mathcal{N}}_m\right)$$

- Perform a separate PDF fit to each replica data set.
- Calculate average and s.d. over $N_{\text{rep}} \sim \mathcal{O}(100)$ PDF sets.
- Equivalent to Hessian method with $\Delta\chi^2 = 1$ [e.g. J. Feltesse, A. Glazov, V. Radescu, H1 PDF fit, arXiv:0903.3861].
- But more useful when fitting weakly-constrained parameters.

Puzzle: NNPDF *w/o* tolerance ≈ MSTW/CTEQ *with* tolerance.

1. Tolerance compensates for MSTW/CTEQ parameterisation bias?
2. Tolerance mimics premature stopping criterion for NN training?

Implementing MC method in **MSTW** fit allows a controlled study.

Monte Carlo sampling using data replicas (used by NNPDF)

- Generate replica data sets with shifted central values:

$$D_{m,i} \to \left(D_{m,i} + R_{m,i}^{\text{uncorr.}} \, \sigma_{m,i}^{\text{uncorr.}} + \sum_{k=1}^{N_{\text{corr.}}} R_{m,k}^{\text{corr.}} \, \sigma_{m,k,i}^{\text{corr.}} \right) \cdot \left(1 + R_m^{\mathcal{N}} \, \sigma_m^{\mathcal{N}}\right)$$

- Perform a separate PDF fit to each replica data set.
- Calculate average and s.d. over $N_{\text{rep}} \sim \mathcal{O}(100)$ PDF sets.
- Equivalent to Hessian method with $\Delta\chi^2 = 1$ [e.g. J. Feltesse, A. Glazov, V. Radescu, H1 PDF fit, arXiv:0903.3861].
- But more useful when fitting weakly-constrained parameters.

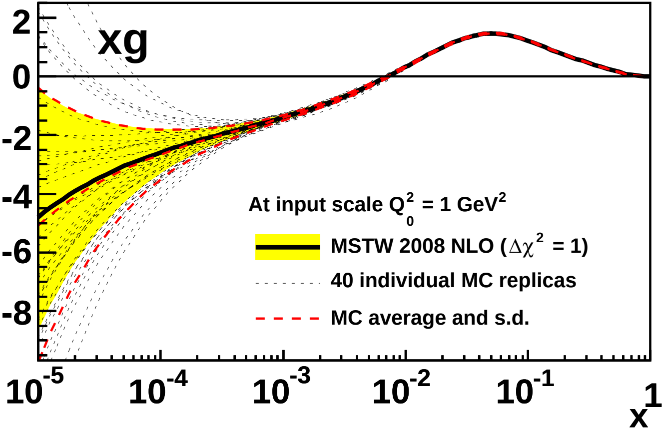
Puzzle: NNPDF *w/o* tolerance ≈ MSTW/CTEQ *with* tolerance.

1. Tolerance compensates for MSTW/CTEQ parameterisation bias?
2. Tolerance mimics premature stopping criterion for NN training?

Implementing MC method in **MSTW** fit allows a controlled study.

Input gluon distribution, $xg(x, Q_0^2 = 1\ \mathrm{GeV}^2)$

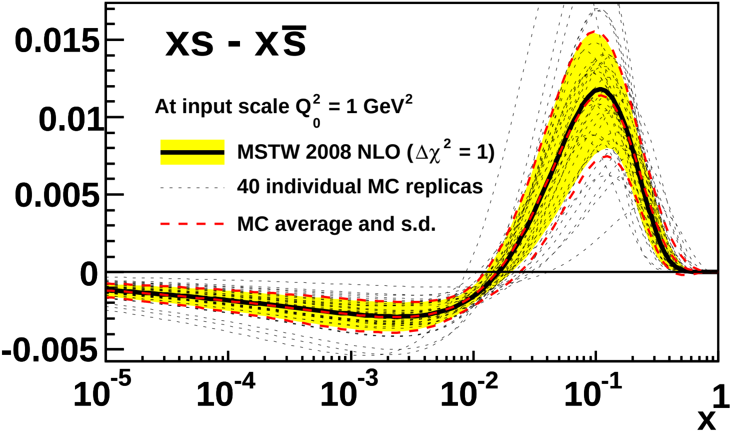
- Compare **Hessian** (w/o tolerance) to Monte Carlo method:



- Good agreement of MC average and s.d. with **Hessian** uncertainty.

Input strange asymmetry, $x(s-\bar{s})(x, Q_0^2 = 1\ \mathrm{GeV}^2)$

- Compare **Hessian** (w/o tolerance) to Monte Carlo method:



- Good agreement of MC average and s.d. with **Hessian** uncertainty.

Input parameterisation in MSTW 2008 NLO fit

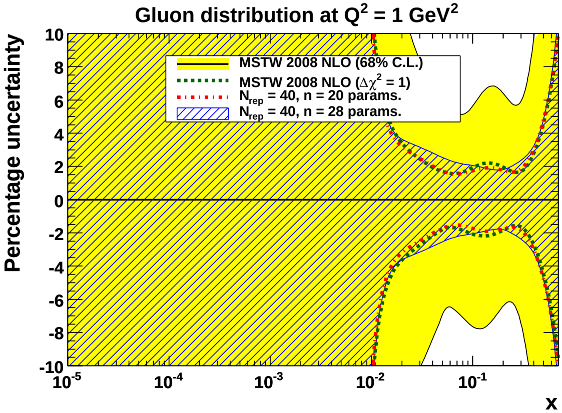
Input parameterisation ($Q_0^2 = 1$ GeV^2) in MSTW 2008 fit

$$xu_v = A_u\, x^{\eta_1}(1-x)^{\eta_2}(1+\epsilon_u\sqrt{x}+\gamma_u\, x)$$
$$xd_v = A_d\, x^{\eta_3}(1-x)^{\eta_4}(1+\epsilon_d\sqrt{x}+\gamma_d\, x)$$
$$xS = A_S\, x^{\delta_S}(1-x)^{\eta_S}(1+\epsilon_S\sqrt{x}+\gamma_S\, x)$$
$$x(\bar{d}-\bar{u}) = A_\Delta\, x^{\eta_\Delta}(1-x)^{\eta_S+2}(1+\gamma_\Delta\, x+\delta_\Delta\, x^2)$$
$$xg = A_g\, x^{\delta_g}(1-x)^{\eta_g}(1+\epsilon_g\sqrt{x}+\gamma_g\, x)+A_{g'}\, x^{\delta_{g'}}(1-x)^{\eta_{g'}}$$
$$x(s+\bar{s}) = A_+\, x^{\delta_S}\,(1-x)^{\eta_+}(1+\epsilon_S\sqrt{x}+\gamma_S\, x)$$
$$x(s-\bar{s}) = A_-\, x^{0.2}(1-x)^{\eta_-}(1-x/x_0)$$

- A_u, A_d, A_g and x_0 are determined from sum rules.
- 28 parameters allowed to go free to find overall best fit.
  But restrict to 20 parameters for Hessian error propagation.
- MC sampling allows all 28 parameters free $\Rightarrow$ check potential bias!

Parameterisation bias for xg at $Q_0 = 1$ GeV

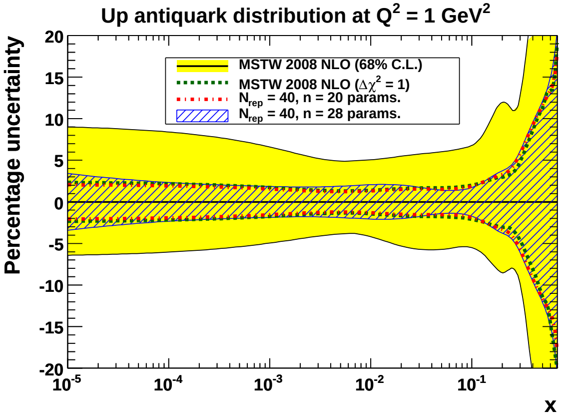
- Compare Hessian **with** or without tolerance (4 parameters for xg) to Monte Carlo sampling (either 4 or 7 parameters for xg).



$$xg = A_g\, x^{\delta_g}(1-x)^{\eta_g}(1+\epsilon_g\sqrt{x}+\gamma_g\, x) + A_{g'}\, x^{\delta_{g'}}(1-x)^{\eta_{g'}}$$

Parameterisation bias for $x\bar{u}$ at $Q_0 = 1$ GeV

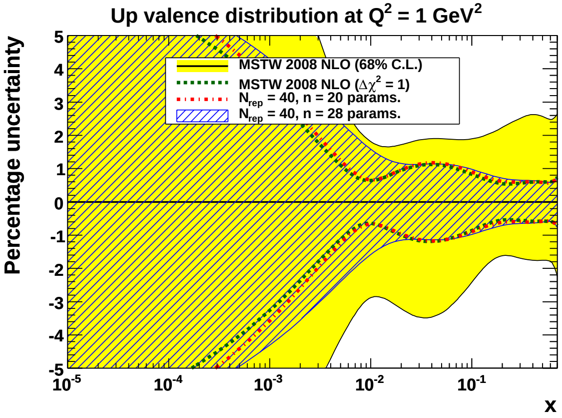
- Compare Hessian **with** or without tolerance (3 parameters for xS) to Monte Carlo sampling (either 3 or 5 parameters for xS).



$$xS \equiv 2x\bar{u} + 2x\bar{d} + xs + x\bar{s} = A_S\, x^{\delta_S}(1-x)^{\eta_S}(1 + \epsilon_S\sqrt{x} + \gamma_S\, x)$$

Parameterisation bias for xu_v at $Q_0 = 1$ GeV

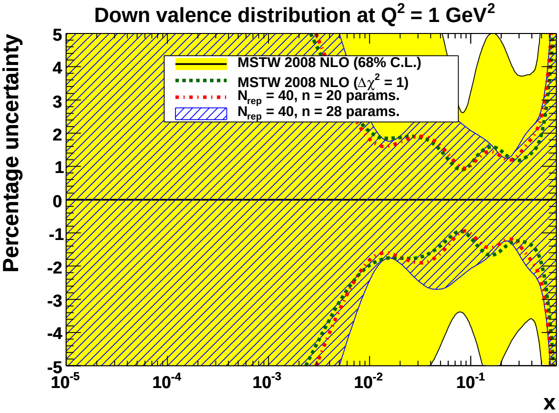
- Compare Hessian **with** or without tolerance (3 parameters for xu_v) to Monte Carlo sampling (either 3 or 4 parameters for xu_v).



$$xu_v = A_u\, x^{\eta_1}(1-x)^{\eta_2}(1 + \epsilon_u\,\sqrt{x} + \gamma_u\, x)$$

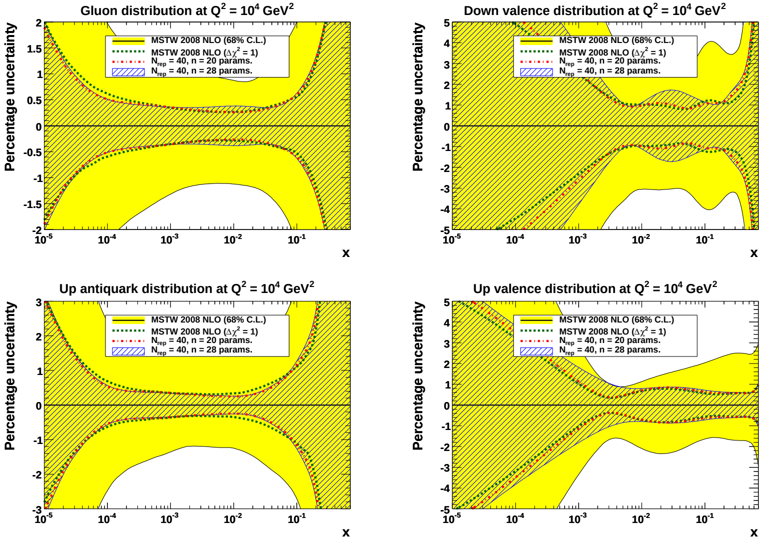
Parameterisation bias for xd_v at $Q_0 = 1$ GeV

- Compare Hessian **with** or without tolerance (3 parameters for xd_v) to Monte Carlo sampling (either 3 or 4 parameters for xd_v).



$$xd_v = A_d\, x^{\eta_3}(1-x)^{\eta_4}(1+\epsilon_d\sqrt{x}+\gamma_d\, x)$$

Parameterisation bias for PDFs evolved to $Q = 100$ GeV



- **Conclusion:** param. bias is likely to be small (except $s, \bar{s}$).

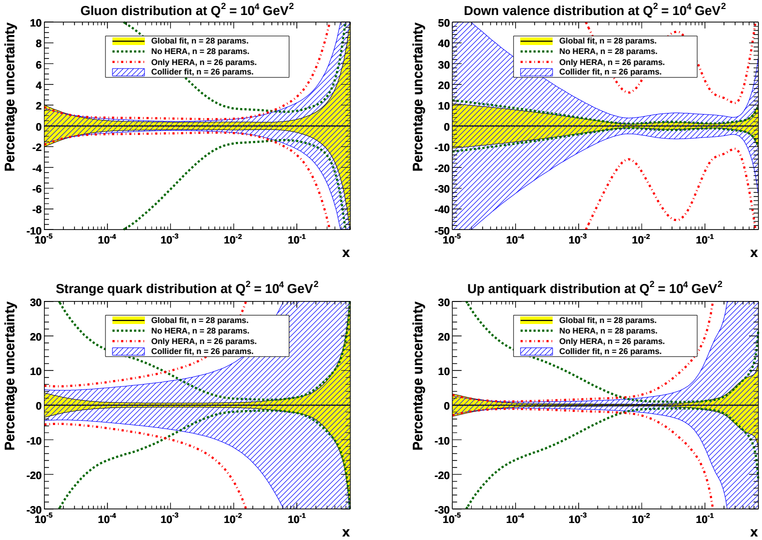
Input parameters ($Q_0^2 = 1.9\ \mathrm{GeV}^2$) in HERAPDF1.0/1.5

$$
\begin{aligned}
xu_v &= A_{u_v}\, x^{B_{qv}} (1-x)^{C_{u_v}} (1 + E_{u_v}\, x^2) \\
xd_v &= A_{d_v}\, x^{B_{qv}} (1-x)^{C_{d_v}} \\
x\bar{u} &= A_{\bar{q}}\, x^{B_{\bar{q}}} (1-x)^{C_{\bar{u}}} \\
x\bar{d} &= A_{\bar{q}}\, x^{B_{\bar{q}}} (1-x)^{C_{\bar{d}}} \\
x\bar{s} &= 0.45\, x\bar{d} \\
xs &= x\bar{s} \\
xg &= A_g\, x^{B_g} (1-x)^{C_g}
\end{aligned}
$$

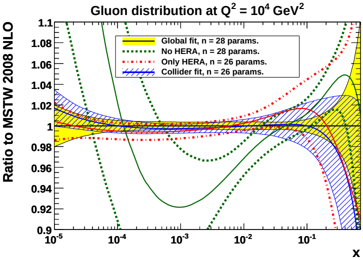
- 10 parameters for central fit and "experimental" uncertainties, additional "model" and "parameterisation" uncertainties.
- 4 more params. for HERAPDF1.5 NNLO (2 for g, 1 each for u_v, d_v).
- Does HERAPDF get the "right" answer for the "wrong" reason? (i.e. constrained by *parameterisation* in **absence** of relevant data.)
- What is *genuine* HERA constraint with flexible parameterisation?

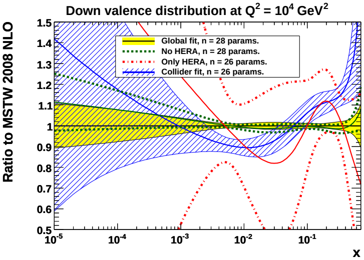
Constraints provided by subsets of data using MC sampling

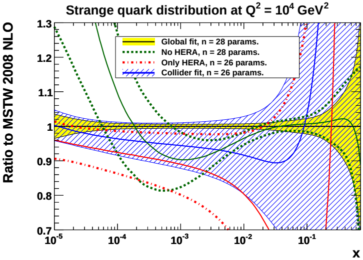
- Fit subsets of global data in MSTW 2008 NLO analysis:

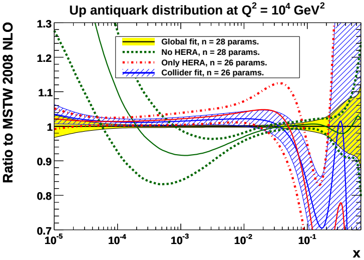


Relative changes in PDFs when fitting subsets of data





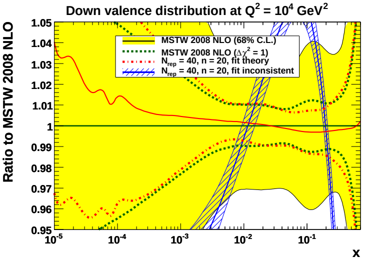


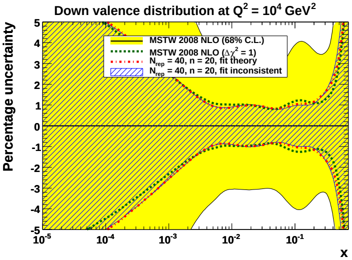


- Uncertainty bands not always consistent $\Rightarrow$ need tolerance.

Study of consistent/inconsistent idealised pseudodata

- Generate perfectly **consistent** data from global best-fit theory.
- $\Rightarrow$ MC fit to idealised data gives $\chi^2/N_{\rm pts.} = 0.98 \pm 0.03$.
- Then introduce deliberate inconsistencies to idealised pseudodata.
- $\Rightarrow$ MC fit to inconsistent data gives $\chi^2/N_{\rm pts.} = 1.07 \pm 0.04$.

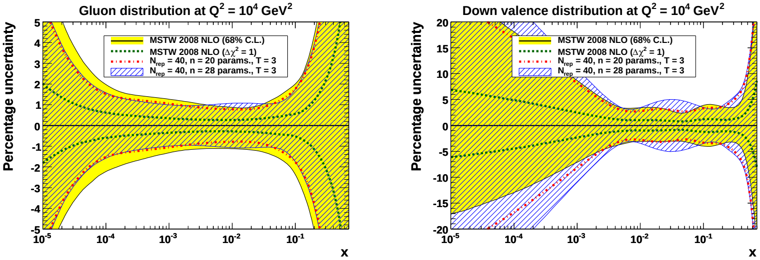




- Data set inconsistencies not obvious to spot $\Rightarrow$ need tolerance.

How to include tolerance in Monte Carlo method?

- **Simplest approach:** scale all experimental uncertainties in generation of data replicas by "average" tolerance ($T \approx 3$).



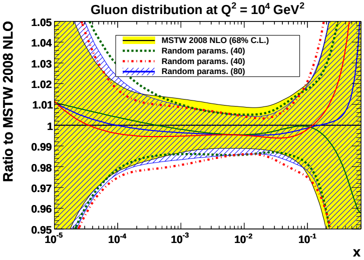
- MC uncertainties (20 parameters, $T = 3$) close to normal Hessian uncertainties. But not clear how to implement "dynamic" tolerance (different for each eigenvector direction).

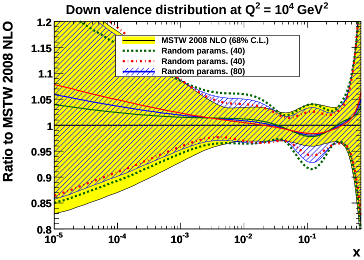
Random PDFs generated in space of fit parameters

- Generate random PDFs from covariance matrix [H. Prosper].
- Done by Giele and Keller [hep-ph/9803393] (using Alekhin's fit).
- Expand parameter displacements in basis of eigenvectors:

$$a_i - a_i^0 = \sum_{j=1}^{20} \sqrt{\lambda_j}\, v_{ij}\, z_j.$$

- Usually, $z_j = \pm t_j^{\pm}\,\delta_{jk}$ for $\pm k$th eigenvector PDF set.
- Instead generate random PDF sets with $z_j = \pm t_j^{\pm}\,R_j$:





Bayesian reweighting to include new data sets

[Giele, Keller, hep-ph/9803393; NNPDF, arXiv:1012.0836, arXiv:1108.1758]

- Now have a set of random PDFs $\{f_k\}$ with equal probability. (Irrelevant whether generated in space of data or parameters.)
- Can apply **reweighting** technique exactly as for NNPDF.
- Compute χ_k^2 for new data set using each f_k, then calculate the mean value of a PDF-dependent quantity $\mathcal{O}[f]$ as:

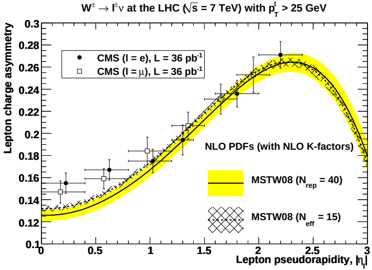
$$\langle\mathcal{O}\rangle_{\rm old} = \frac{1}{N_{\rm rep}}\sum_{k=1}^{N_{\rm rep}}\mathcal{O}[f_k], \quad \langle\mathcal{O}\rangle_{\rm new} = \frac{1}{N_{\rm rep}}\sum_{k=1}^{N_{\rm rep}} w_k(\chi_k^2)\,\mathcal{O}[f_k],$$

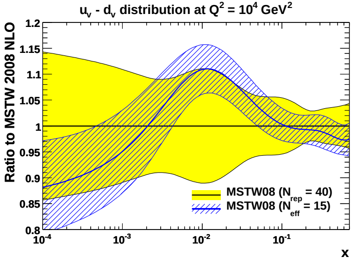
where the weights are given by

$$w_k(\chi_k^2) \propto \left(\chi_k^2\right)^{\frac{1}{2}(N_{\rm pts.}-1)} \exp\left(-\frac{1}{2}\chi_k^2\right).$$

$W \to \ell\nu$ charge asymmetry from CMS [arXiv:1103.3470]

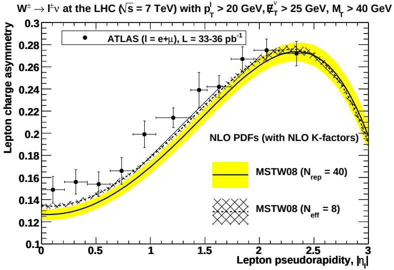
- **Disclaimer:** aim to simply demonstrate reweighting technique, *not* to show true impact of LHC data on **MSTW08** PDF fit.

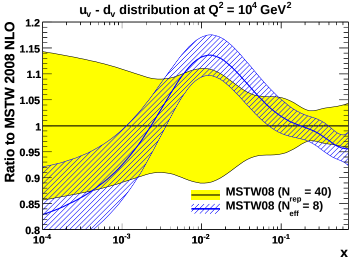




- Include CMS data by reweighting: $\chi^2/N_{\text{pts.}} = 2.43 \to 1.32$.
- Reweighting shifts $u_v - d_v$ at $x \sim M_W/\sqrt{s} \sim 0.01$.

$W \to \ell\nu$ charge asymmetry from ATLAS [arXiv:1109.5141]





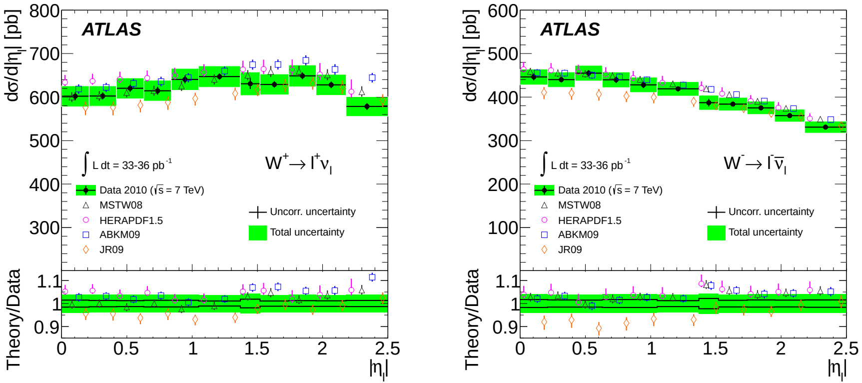
- Include ATLAS data by reweighting: $\chi^2/N_{\rm pts.} = 3.39 \to 1.48$.
- Large shift in central value of $u_v - d_v$ outside original uncertainty.
- Clear **tension** of LHC $W \to \ell\nu$ asymmetry with existing data in **MSTW08** fit (e.g. Tevatron $W \to \ell\nu$ asymmetry, NMC F_2^d/F_2^p ratio, E866/NuSea DY σ^{pd}/σ^{pp} ratio). Further work needed (e.g. reexamination of nuclear corrections for F_2^d).
- Example where reweighting is useful, but does not indicate the true impact of including the new data in a global PDF fit.

Summary

- First implementation of *Monte Carlo sampling* in **MSTW** fit.
- Parameterisation bias is likely to be small (except $s, \bar{s}$).
- Studies of fitting restricted data sets and consistent or inconsistent pseudodata suggest need tolerance ($\Delta\chi^2 > 1$).
- New method of generating random PDFs in parameter space using basis of eigenvectors including dynamic tolerance.
- Random PDFs can be used for Bayesian reweighting *à la* NNPDF: can make public to allow independent studies.

Differential cross sections: $\mathrm{d}\sigma(\ell^+)/\mathrm{d}\eta_\ell$ and $\mathrm{d}\sigma(\ell^-)/\mathrm{d}\eta_\ell$

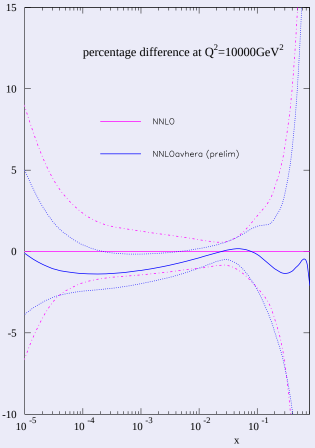
[ATLAS Collaboration, `arXiv:1109.5141`]



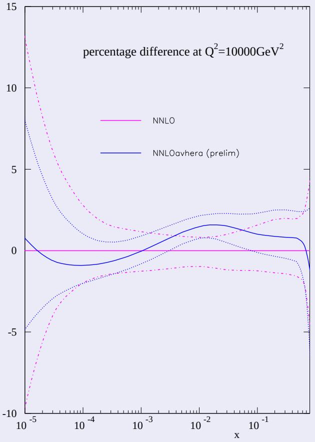
- **MSTW08** describes ATLAS differential cross sections $\mathrm{d}\sigma/\mathrm{d}\eta_\ell$ for ℓ^+ and ℓ^- better than the charge asymmetry $A_\ell(\eta_\ell)$.

Impact of combined HERA I data [arXiv:0911.0884] (R. Thorne)

Gluon distribution



Up quark distribution



- Changes not large enough to warrant an immediate update.