# Updates on PDFs in the MMHT framework – "MMHT19"

Thomas Cridge (University College London)

with Shaun Bailey, Lucian Harland-Lang, Alan Martin,

Ricky Nathvani and Robert Thorne

t.cridge@ucl. ac.uk

**UCL**

# Talk Outline

1. **Introduction – MMHT14, MMHT16, "MMHT19".**

2. **New Data**

   - Hera Combined updated data

   - LHC data – LHCb W,Z at 7,8 TeV e,μ; CMS W asymmetry 8 TeV; CMS 7 TeV W+c; high precision ATLAS W,Z 7 TeV; Atlas and CMS 7 TeV jets; ATLAS 8TeV differential $t\bar{t}$ and others.

   - Tevatron D0 electron asymmetry data.

3. **Further Data to be added for "MMHT19"**

4. **New Extended Parametrisation and eigenvector sets**

   - Extension to 6 Chebyshevs throughout.

   - $\bar{d}/\bar{u}$ now parametrised rather than $\bar{d} - \bar{u}$

5. **MMHT QED**

   - new MMHT2015QED set with DGLAP QED evolution and photon PDF.

# Introduction

# MMHT14 datasets

| Data set | LO | NLO | NNLO |
|---|---|---|---|
| BCDMS $\mu p$ $F_2$ [125] | 162 / 153 | 176 / 163 | 173 / 163 |
| BCDMS $\mu d$ $F_2$ [19] | 140 / 142 | 143 / 151 | 143 / 151 |
| NMC $\mu p$ $F_2$ [20] | 141 / 115 | 132 / 123 | 123 / 123 |
| NMC $\mu d$ $F_2$ [20] | 134 / 115 | 115 / 123 | 108 / 123 |
| NMC $\mu n/\mu p$ [21] | 122 / 137 | 131 / 148 | 127 / 148 |
| E665 $\mu p$ $F_2$ [22] | 59 / 53 | 60 / 53 | 65 / 53 |
| E665 $\mu d$ $F_2$ [22] | 52 / 53 | 52 / 53 | 60 / 53 |
| SLAC $ep$ $F_2$ [23, 24] | 21 / 18 | 31 / 37 | 31 / 37 |
| SLAC $ed$ $F_2$ [23, 24] | 13 / 18 | 30 / 38 | 26 / 38 |
| NMC/BCDMS/SLAC/HERA $F_L$ [20, 125, 24, 63, 64, 65] | 113 / 53 | 68 / 57 | 63 / 57 |
| E866/NuSea $pp$ DY [88] | 229 / 184 | 221 / 184 | 227 / 184 |
| E866/NuSea $pd/pp$ DY [89] | 29 / 15 | 11 / 15 | 11 / 15 |
| NuTeV $\nu N$ $F_2$ [29] | 35 / 49 | 39 / 53 | 38 / 53 |
| CHORUS $\nu N$ $F_2$ [30] | 25 / 37 | 26 / 42 | 28 / 42 |
| NuTeV $\nu N$ $xF_3$ [29] | 49 / 42 | 37 / 42 | 31 / 42 |
| CHORUS $\nu N$ $xF_3$ [30] | 35 / 28 | 22 / 28 | 19 / 28 |
| CCFR $\nu N \to \mu\mu X$ [31] | 65 / 86 | 71 / 86 | 76 / 86 |
| NuTeV $\nu N \to \mu\mu X$ [31] | 53 / 40 | 38 / 40 | 43 / 40 |

| | | | |
|---|---|---|---|
| HERA $e^+p$ NC 820 GeV [61] | 125 / 78 | 93 / 78 | 89 / 78 |
| HERA $e^+p$ NC 920 GeV [61] | 479 / 330 | 402 / 330 | 373 / 330 |
| HERA $e^-p$ NC 920 GeV [61] | 158 / 145 | 129 / 145 | 125 / 145 |
| HERA $e^+p$ CC [61] | 41 / 34 | 34 / 34 | 32 / 34 |
| HERA $e^-p$ CC [61] | 29 / 34 | 23 / 34 | 21 / 34 |
| HERA $ep$ $F_2^{\text{charm}}$ [62] | 105 / 52 | 72 / 52 | 82 / 52 |
| H1 99–00 $e^+p$ incl. jets [126] | 77 / 24 | 14 / 24 | — |
| ZEUS incl. jets [127, 128] | 140 / 60 | 45 / 60 | — |
| DØ II $p\bar{p}$ incl. jets [119] | 125 / 110 | 116 / 110 | 119 / 110 |
| CDF II $p\bar{p}$ incl. jets [118] | 78 / 76 | 63 / 76 | 59 / 76 |
| CDF II $W$ asym. [66] | 55 / 13 | 32 / 13 | 30 / 13 |
| DØ II $W \to \nu e$ asym. [67] | 47 / 12 | 28 / 12 | 27 / 12 |
| DØ II $W \to \nu\mu$ asym. [68] | 16 / 10 | 19 / 10 | 21 / 10 |
| DØ II $Z$ rap. [90] | 34 / 28 | 16 / 28 | 16 / 28 |
| CDF II $Z$ rap. [70] | 95 / 28 | 36 / 28 | 40 / 28 |
| ATLAS $W^+, W^-, Z$ [10] | 94 / 30 | 38 / 30 | 39 / 30 |
| CMS $W$ asymm $p_T > 35$ GeV [9] | 10 / 11 | 7 / 11 | 9 / 11 |
| CMS asymm $p_T > 25$ GeV, 30 GeV [77] | 7 / 24 | 8 / 24 | 10 / 24 |
| LHCb $Z \to e^+e^-$ [79] | 76 / 9 | 13 / 9 | 20 / 9 |
| LHCb $W$ asymm $p_T > 20$ GeV [78] | 27 / 10 | 12 / 10 | 16 / 10 |
| CMS $Z \to e^+e^-$ [84] | 46 / 35 | 19 / 35 | 22 / 35 |
| ATLAS high-mass Drell-Yan [83] | 42 / 13 | 21 / 13 | 17 / 13 |
| CMS double diff. Drell-Yan [86] | — | 372 / 132 | 149 / 132 |
| Tevatron, ATLAS, CMS $\sigma_{t\bar{t}}$ [91]–[97] | 53 / 13 | 7 / 13 | 8 / 13 |
| ATLAS jets (2.76 TeV+7 TeV) [108, 107] | 162 / 116 | 106 / 116 | — |
| CMS jets (7 TeV) [106] | 150 / 133 | 138 / 133 | — |
| All data sets | 3706 / 2763 | 3267 / 2996 | 2717 / 2663 |

- **MMHT14 was the last public release** of the MRST, MSTW, MMHT PDF collaboration sets.

- Now is a time for a new public PDF set – "MMHT19" with a significant amount of LHC Run 1 data at 7 and 8 TeV included - due to come out shortly (order a couple/few months).

# New Data

# HERA combined dataset  *[H. Abramowicz et al., (2015), 1506.06042]

- New since MMHT14 but in MMHT16.

- Little change in fit quality → little change in central value of PDFs.

- New PDFs well within the MMHT14 uncertainty bands.

- Differences and tensions relative to HERA only fit are evident:

  - HERA only favours smaller/larger $u_v$ in intermediate/large $x$ due to CC e⁻p data.

  - No softer gluon at large $x$.

- HERA data does reduce pdf uncertainties:

  - g uncertainty reduced for $x < 0.01$.

  - Small improvements in $u_v$ and $d_v$ for small $x$.
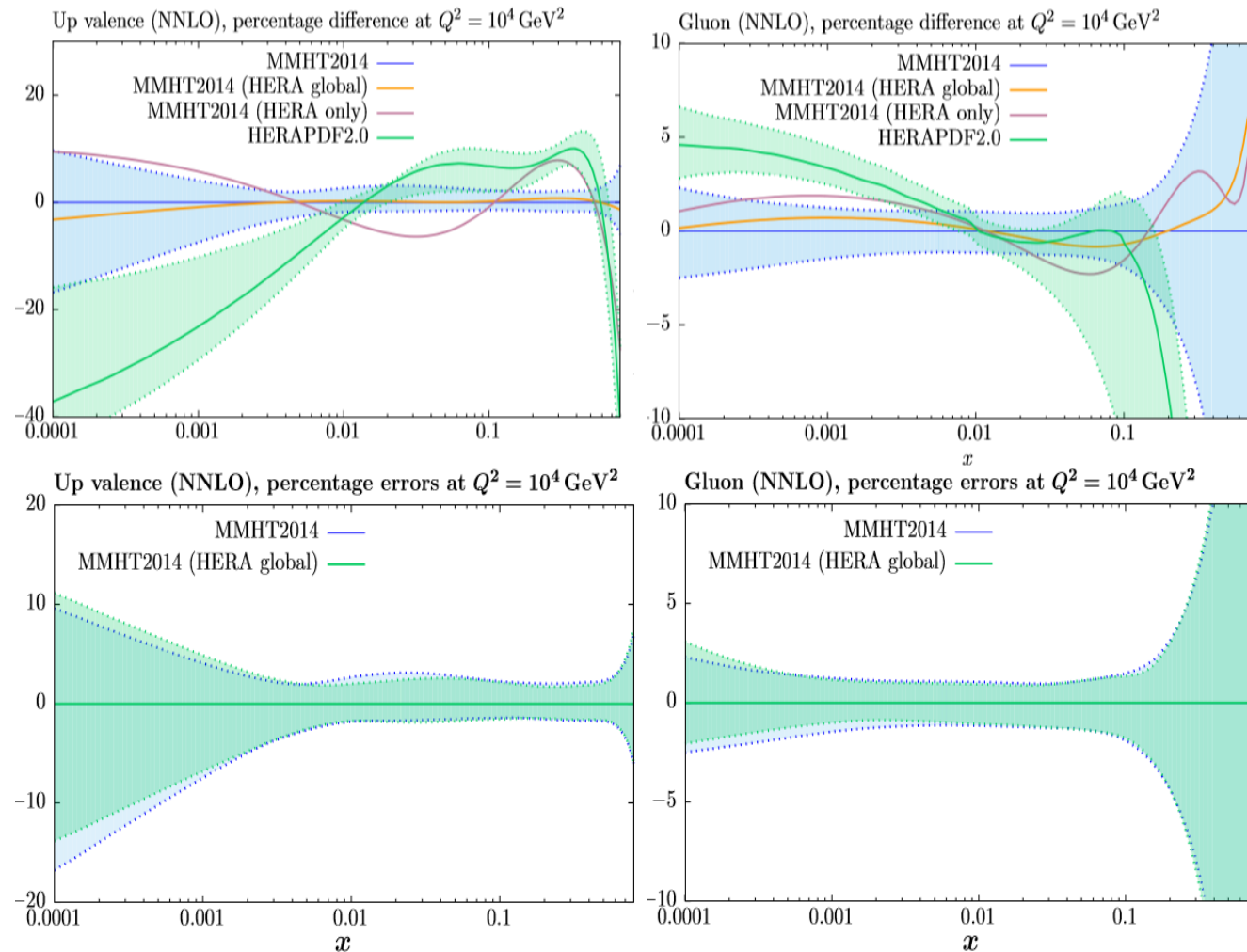
[See Harland-Lang, Martin, Motylinski, Thorne 1601.03413 for more details.]

# LHC(+Tevatron) Electroweak and other Data

- Since MMHT14 a great deal of further electroweak data and jet has been added:

  - High rapidity W, Z data from LHCb at 7 and 8TeV.

  - High precision CMS data on $W^{+-}$ rapidity distributions (also interpreted as an asymmetry measurement).

  - W + c jets from CMS at 7TeV.

  - Final D0 electron asymmetry data from Tevatron.

  - Some additional $t\bar{t}$ data.

  - High precision Atlas W, Z 7 TeV data.

- Predictions good (with a couple of exceptions – see later slides), no real tension with other data when refitting:

| | Points | NLO $\chi^2$ | NNLO $\chi^2$ |
|---|---|---|---|
| $\sigma_{t\bar{t}}$ | 18 | 19.6 (20.5) | 14.7 (15.3) |
| LHCb 7 TeV $W + Z$ | 33 | 50.1 (45.4) | 46.5 (42.9) |
| LHCb 8 TeV $W + Z$ | 34 | 77.0 (58.9) | 62.6 (59.0) |
| LHCb 8 TeV $Z \rightarrow ee$ | 17 | 37.4 (33.4) | 30.3 (28.9) |
| CMS 8 TeV $W$ | 22 | 32.6 (18.6) | 34.9 (20.5) |
| CMS 7 TeV $W + c$ | 10 | 8.5 (10.0) | 8.7 (7.8) |
| D0 $e$ asymmetry | 13 | 22.2 (21.5) | 27.3 (25.8) |
| Total | 3405 (3738) | 4375.9 (4336.1) | 3741.5 (3723.7) |

For the rest of the data:
$\Delta\chi^2 = 9$ at NLO
$\Delta\chi^2 = 15$ at NNLO
-> Little changes to PDFs

Shown are $\chi^2$ values without(with) refitting. Therefore unbracketed are MMHT14 predictions.

[See Thorne, Harland-Lang, Martin 1708.00047 for more details.]

# CMS 7 TeV $W$+c differential in lepton rapidity

- Constrains strange and antistrange quark distributions.

$$R_c^{\pm} = \frac{\sigma(pp \to W^+ + \bar{c} + X)}{\sigma(pp \to W^- + c + X)}$$

| | GeV | data | MSTW2008 | MMHT2014 |
|---|---|---|---|---|
| $\sigma(W+c)$ | $p_T^{lep} > 25$ | $107.7 \pm 3.3(\text{stat.}) \pm 6.9(\text{sys.})$ | $102.8 \pm 1.7$ | $110.2 \pm 8.1$ |
| $\sigma(W+c)$ | $p_T^{lep} > 35$ | $84.1 \pm 2.0(\text{stat.}) \pm 4.9(\text{sys.})$ | $80.4 \pm 1.4$ | $86.5 \pm 6.5$ |
| $R_c^{\pm}$ | $p_T^{lep} > 25$ | $0.954 \pm 0.025(\text{stat.}) \pm 0.004(\text{sys.})$ | $0.937 \pm 0.029$ | $0.924 \pm 0.026$ |
| $R_c^{\pm}$ | $p_T^{lep} > 35$ | $0.938 \pm 0.019(\text{stat.}) \pm 0.006(\text{sys.})$ | $0.932 \pm 0.030$ | $0.904 \pm 0.027$ |

- Good agreement between prediction (MMHT14) and new fit, little change in pdf central values (marginal increase) but significant reduction in uncertainty.



Note for later-> $s + \bar{s}$ in agreement with MMHT14.
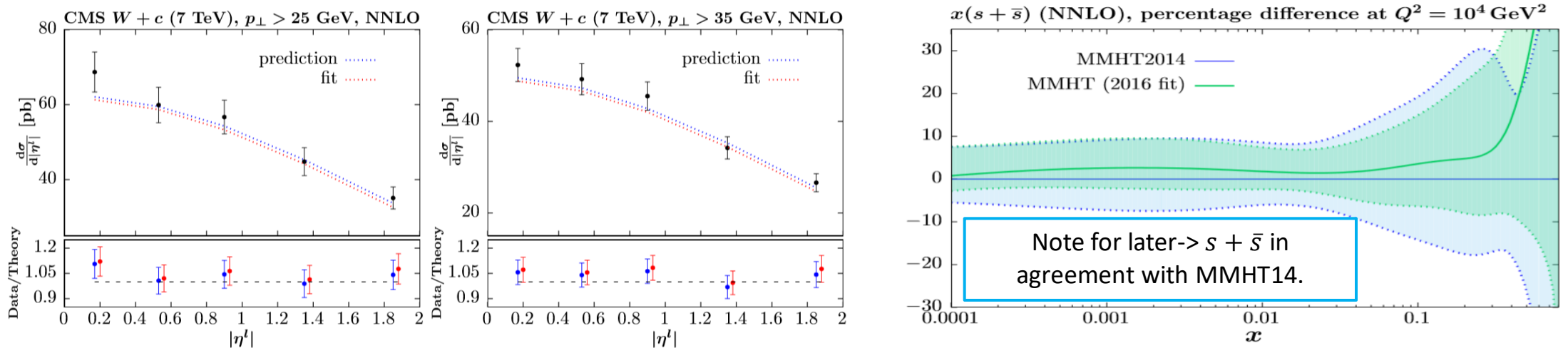
[See Harland-Lang, Nathvani, Thorne 1704.00162 for more details.]

# High precision Atlas data

- Interpreted as causing an increase in strange quark PDFs in profiled MMHT14 PDFs and also in an ATLAS PDF fit (ATLAS-epWZ16) → strangeness "unsuppression" at low $x$.

- ATLAS observe $R_s = \frac{s+\bar{s}}{\bar{u}+\bar{d}} = 1.13^{+0.08}_{-0.13}$.

- Any increase in $\bar{s}(x)$ will cause a reduction in $(\bar{u} + \bar{d})$ as a result of HERA structure function ($F_2$) charge weighted data. But overall singlet grows due to charge weighting – reduction in $\bar{u}$ will cause 4x increase in $\bar{s}$.

- In global fits further constraints on strange arise from:
  - CMS W+c data
  - Neutrino DIS (dimuon) data from NuTeV and CCFR

  these are more constraining and lead to the small $R_s \sim 0.4 - 0.6$ observed in previous global fits.

- As noted earlier, MMHT14 agrees with CMS 7TeV W+c on strange quark distribution.
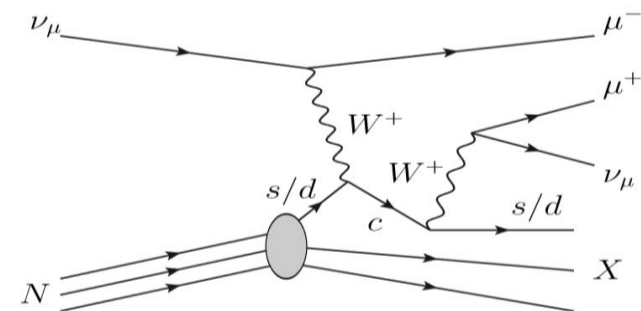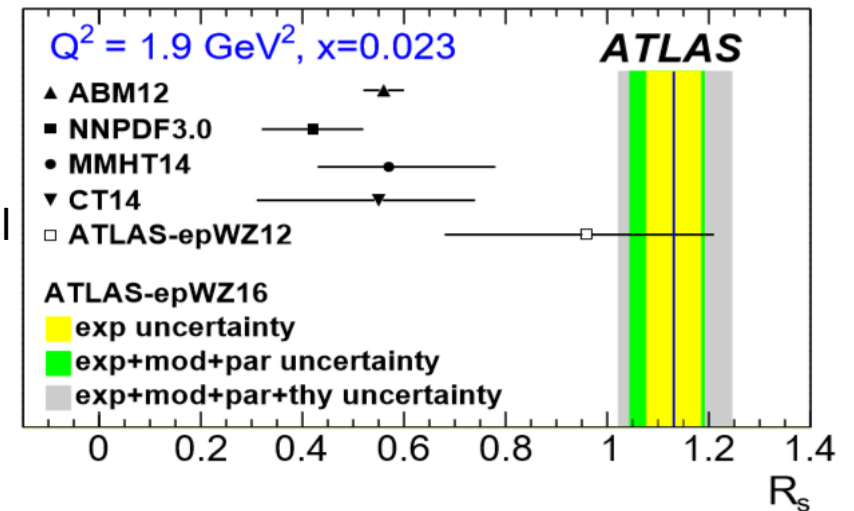
# High precision Atlas data

- Adding this data to MMHT14 PDFs: $\chi^2 \sim 115/61$.

- Refitting results in a slight reduction in $W^-$ prediction, a slight increase in $W^+$ and a shape change in $Z$ production. Further reduction to $\chi^2 \sim 106/61$ with scales $\mu_{R,F} = M_{W/Z}/2$, noticeable improvement in $W^+$

- Clear tension with existing global fit data $\Delta\chi^2 = 53$; significant increase in $\chi^2$ for CMS $Z/\gamma^*$, dimuon data and others.

- If fit with new LHC data improvement of $\Delta\chi^2 = -10$, consistent with new LHC data (except CMS 7 TeV W+c).

- Effect on PDFs: Large increase in $s + \bar{s}$, and a reduction in its uncertainty.

- We find $R_s \sim 0.83 \pm 0.15$, at $x = 0.023$, $Q^2 = 1.9$ GeV$^2$.

- Imposes a 25% reduction on the Branching ratio of the charm meson to muon to maintain the CCFR/NuTeV cross-section.



$(s+\bar{s})/(\bar{u}+\bar{d})$ (NNLO), $Q^2 = 1.9\,\mathrm{GeV}^2$

MMHT
MMHT16 + ATLAS WZ
MMHT + ATLAS WZ $(\mu)$
MMHT + ATLAS WZ $(\mu/2)$
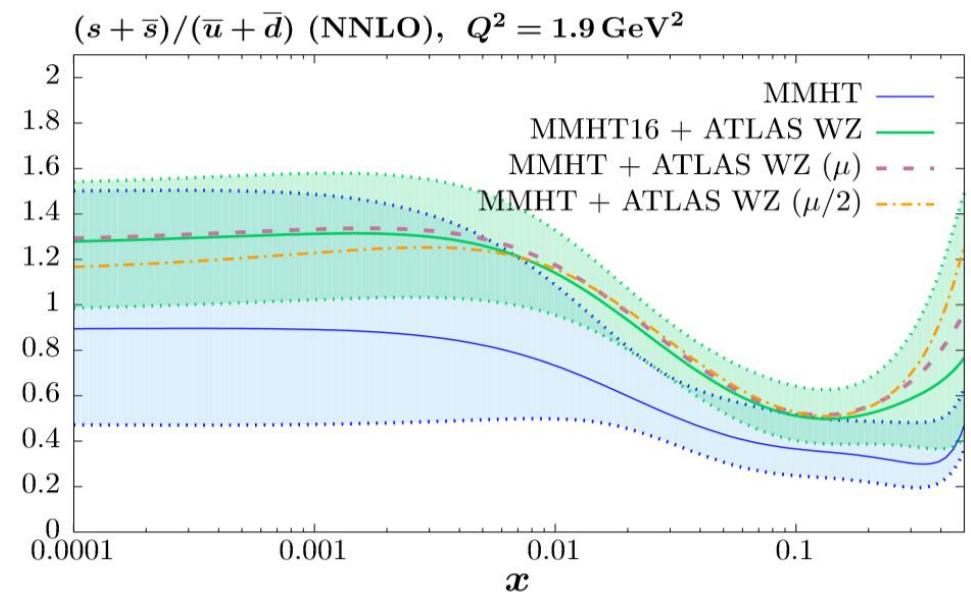
[See Thorne, Harland-Lang, Martin 1708.00047 for more details.]

# High precision Atlas data

- NNLO corrections to dimuon production calculated lowers the cross-section, potentially avoiding the need to reduce the charm meson to muon branching ratio as much.

- Included in "MMHT19" (Bailey), this required some improvement in threshold treatment for charged current VFNS.

- Little effect on global $\chi^2$ but some small reduction in that of Atlas W,Z data, whilst relieves issue of charm meson to muon branching ratio, raising it to a more normal value.
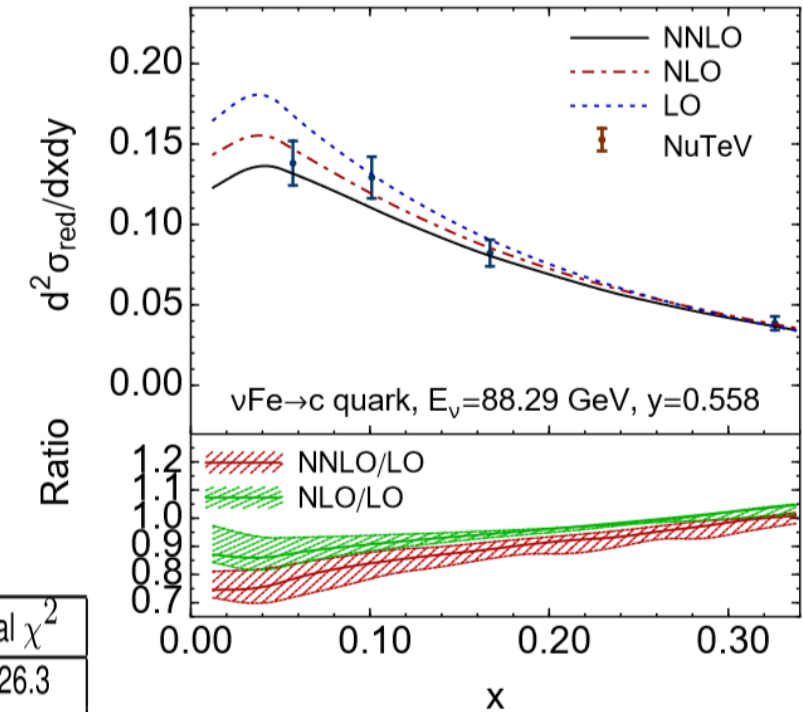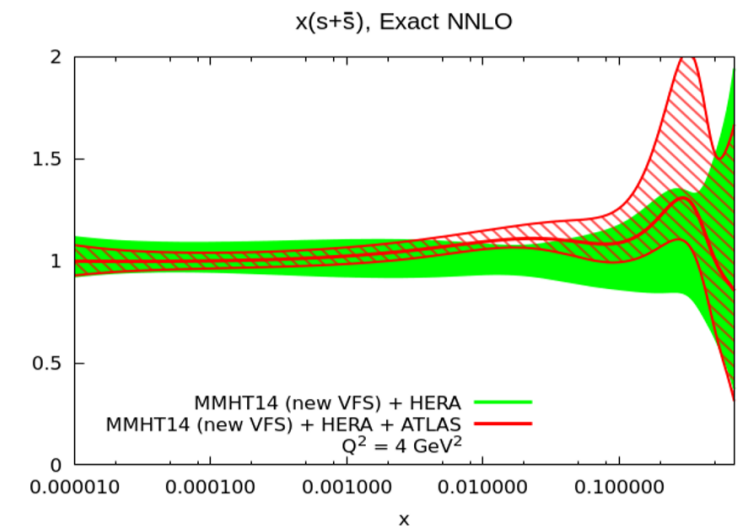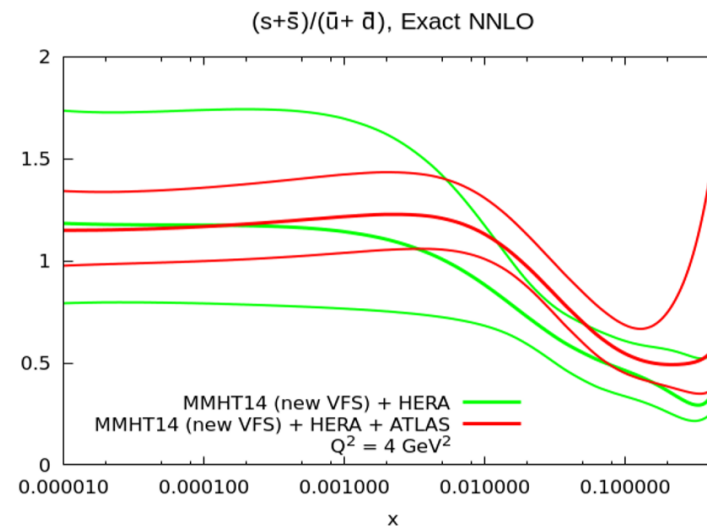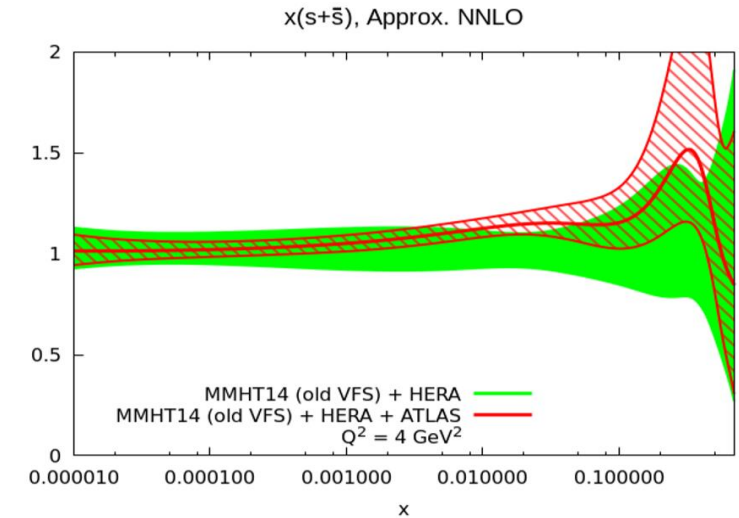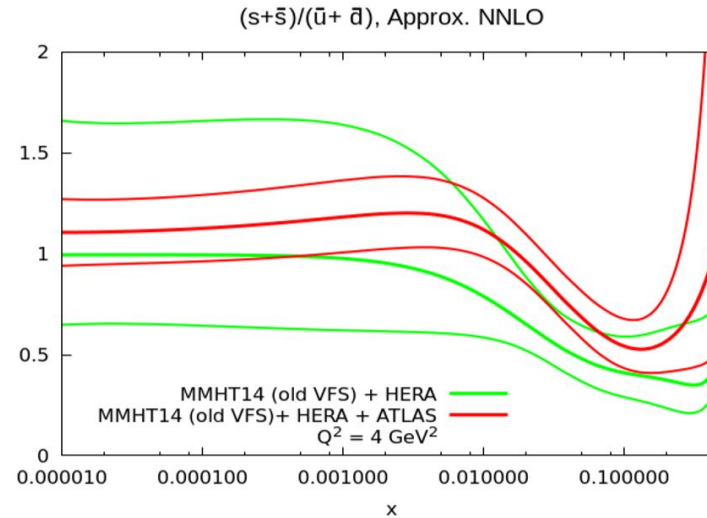


Default $BR(c \to \mu) = 0.092 \pm 10\%$

| | BR($c \to \mu$) | CCFR/NuTeV $\chi^2$ | ATLAS $W, Z\chi^2$ | Total $\chi^2$ |
|---|---|---|---|---|
| MMHT+HERAII | 0.090 | 120.5 | | 3526.3 |
| MMHT+HERAII (NNLO dimuon ) | 0.102 | 122.7 | | 3527.3 |
| MMHT+HERAII (NNLO VFNS dimuon) | 0.101 | 123.9 | | 3531.3 |
| MMHT+HERAII+ATLAS($W, Z$) | 0.073 | 127.3 | 108.6 | 3684.7 |
| MMHT+HERAII+ATLAS($W, Z$) (NNLO dimuon ) | 0.084 | 137.8 | 106.8 | 3688.4 |
| MMHT+HERAII+ATLAS($W, Z$) (NNLO VFNS dimuon) | 0.086 | 137.0 | 106.8 | 3688.5 |
| $N_{pts}$ | | 126.25 | 61 | 3337 |

# High precision Atlas data

- Only small effect on strange PDF of adding the full NNLO dimuon corrections with the increase in the strange reduced.

- Without NNLO corrections at top, with NNLO corrections bottom. Green is MMHT14 + HERA, red is MMHT14+HERA+ATLASWZ.
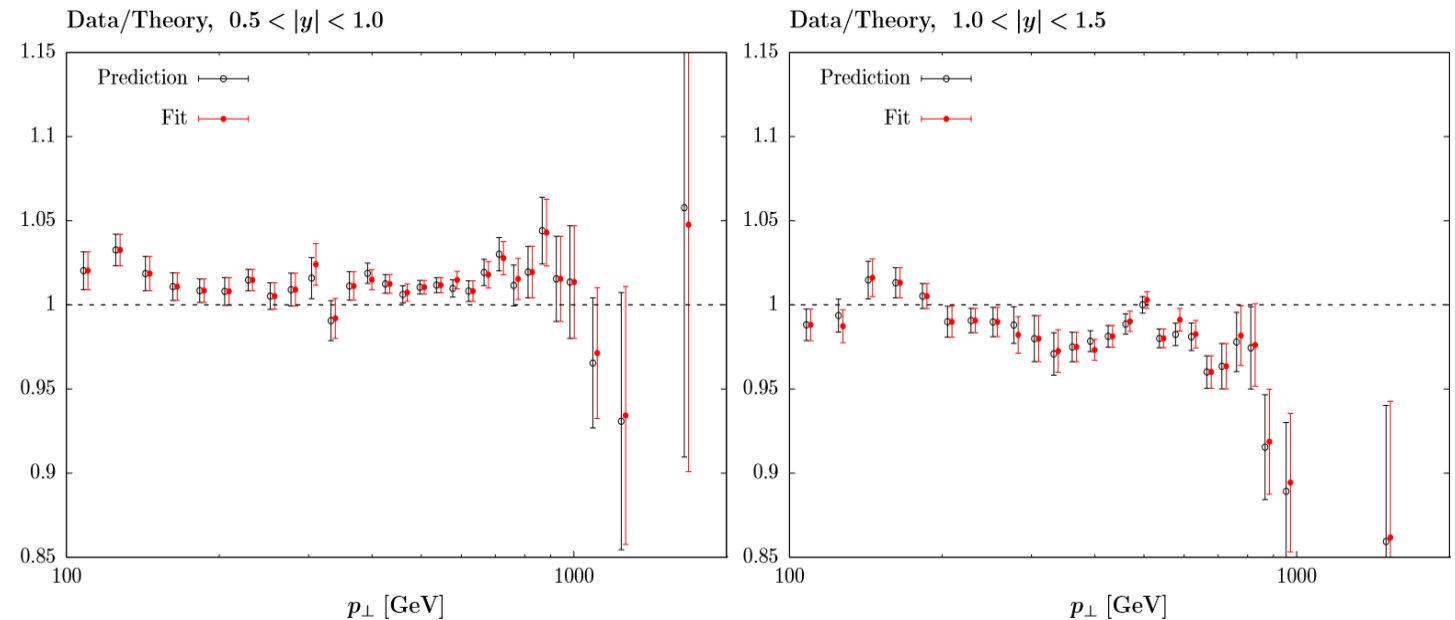
# Atlas 7TeV jets data

- LHC jet data was not included in MMHT14 due to lack of full NNLO result (some studies with toy k-factors).

- Including Atlas 7TeV jet data fully differential in $p_T$ in different jet rapidity bins, comparison of prediction and fit is poor (seen also by several other groups and ATLAS).

  $\chi^2 = 413$ before refitting and $\chi^2 = 400$ after refitting for 140 data points -> very poor fit.

- Comparing different rapidity bins we see a systematic offset in data/theory between neighbouring $y_j$ bins.

- These probe the same pdfs at similar $x$ and $Q^2$ -> refitting cannot improve this tension.
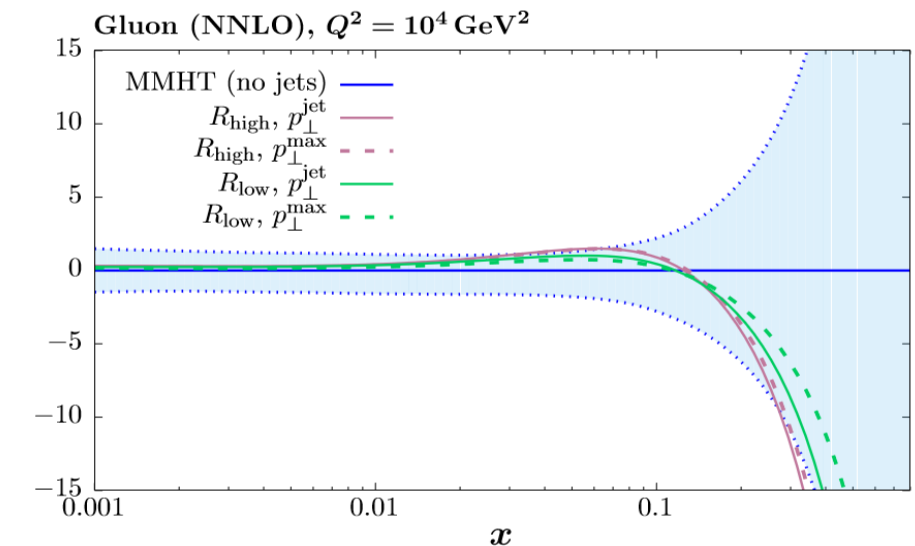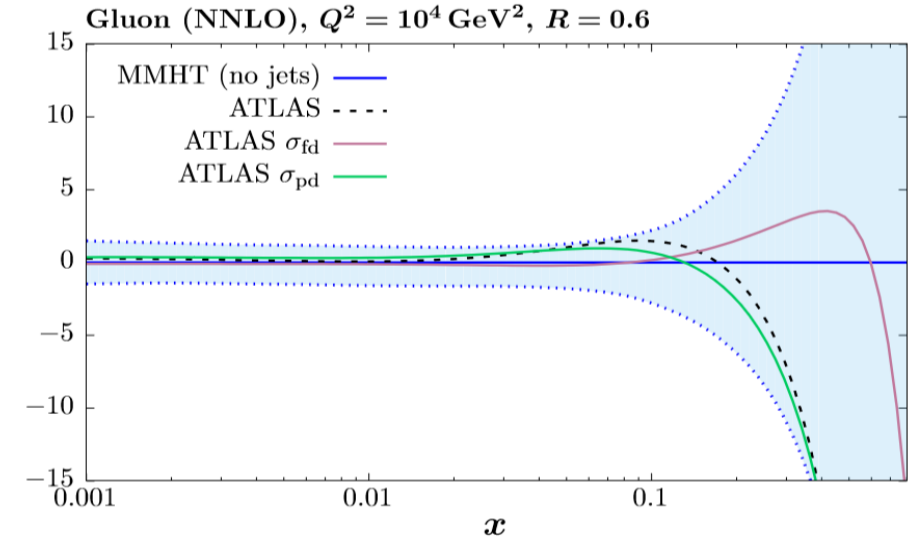


[See Harland-Lang, Nathvani, Thorne 1704.00162 and Harland-Lang, Martin, Thorne 1711.05757 for more details.]

# Atlas 7TeV jets data

- Atlas data dominated by a large number of correlated systematic errors – investigate by decorrelating these systematic errors between rapidity bins, whilst leaving them correlated within each bin, then a good fit for all rapidity bins can be found.

- In fact, the shift required between data and theory is very large for two main systematics (21, 62), indicating a clear tension.

- These correspond to a multi-jet balance asymmetry and the jet energy scale close by jets.

- Decorrelating just these between rapidity bins we can see the improvement of the $\chi^2$ per data point is significant and the data/theory is also then visibly better.

- Overall PDF effect is to soften the gluon at high $x$, and reduce the uncertainty bands there.

- Effect on gluon pdf stable with respect to R and scale choices.

|  | Full | 21 | 62 | 21,62 |
|---|---|---|---|---|
| $\chi^2/N_{\text{pts.}}$ | 2.85 | 1.58 | 2.36 | 1.27 |



Gluon (NNLO), $Q^2 = 10^4\,\text{GeV}^2$, $R = 0.6$

MMHT (no jets)
ATLAS
ATLAS $\sigma_{\text{fd}}$
ATLAS $\sigma_{\text{pd}}$



Gluon (NNLO), $Q^2 = 10^4\,\text{GeV}^2$

MMHT (no jets)
$R_{\text{high}}, p_\perp^{\text{jet}}$
$R_{\text{high}}, p_\perp^{\text{max}}$
$R_{\text{low}}, p_\perp^{\text{jet}}$
$R_{\text{low}}, p_\perp^{\text{max}}$

# Atlas 8 TeV $t\bar{t}$ differential data issues (ATL-PHYSPUB-2018-017)

- Problem – distributions differential in $y_t$, $y_{t\bar{t}}$, $p_T^t$, $M_{t\bar{t}}$, cannot be fit together, indeed problems even when fitting $y_t$, $y_{t\bar{t}}$ alone (seen by MMHT, CT, ATLAS but not NNPDF).



|  | Fitted data set(s) | | | | |
|---|---|---|---|---|---|
| Contribution | $p_T$ | $y_t$ | $y_{t\bar{t}}$ | $M_{t\bar{t}}$ | All |
| $p_T$ | 0.08 | | | | 2.38 |
| $y_t$ | | 1.23 | | | 1.84 |
| $y_{t\bar{t}}$ | | | 1.09 | | 2.22 |
| $M_{t\bar{t}}$ | | | | 0.29 | 1.81 |
| Penalty | 0.24 | 1.83 | 2.35 | 0.17 | 0.88 |
| Total | 0.32 | 3.06 | 3.44 | 0.47 | 2.96 |

Tension between data sets for MMHT (fitted)
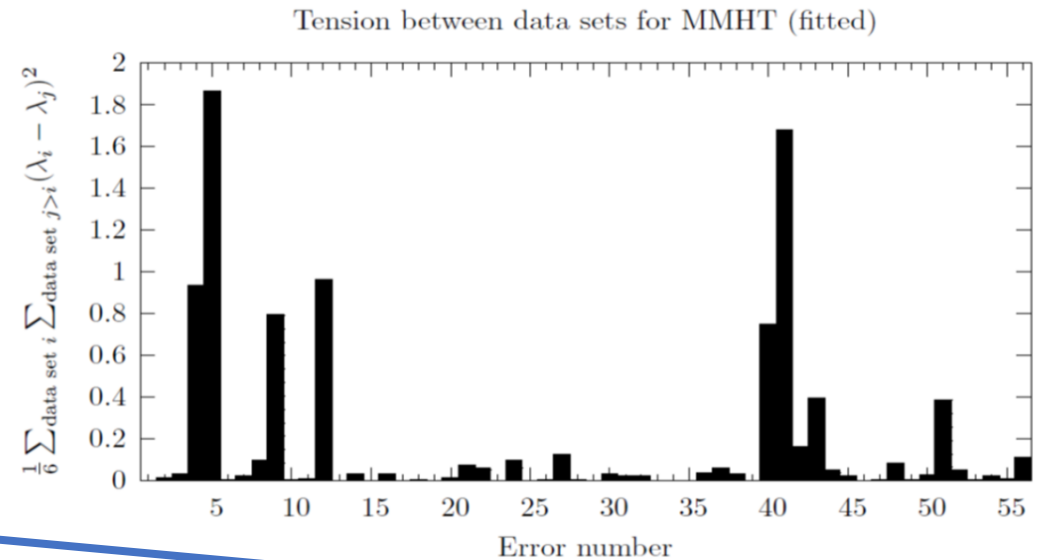


- $\chi^2$ per data point high in simultaneous fit of all the distributions, and in the individual fits to $y_t$, $y_{t\bar{t}}$. (Bailey)

- Dominated by 3 systematics – hard-scattering model, ISR/FSR and parton shower, which are all Monte Carlo related. Let's attempt to decorrelate these systematics.

- CMS do not provide correlations between distributions so only one can be fit at once.

[See Harland-Lang, Martin, Thorne 1711.05757 for more details, contains a study of ATLAS 7 TeV differential jet data with similar issues observed.]

# Atlas 8 TeV $t\bar{t}$ data decorrelation

- First decorrelate between distributions – observe that still $y_t$, $y_{t\bar{t}}$ fits are poor, reflective of fact that fitting alone saw issues.

- Reason is the desired shift in the data-theory is very different point-to-point within these distributions, e.g. for $y_{t\bar{t}}$ :

- Suggests investigating decorrelating also within individual sets.

- Given structure in shifts seen, we model the decorrelation required rather than allowing it to be completely free:

  – sin-cosine decorrelation is best and significantly improves fit
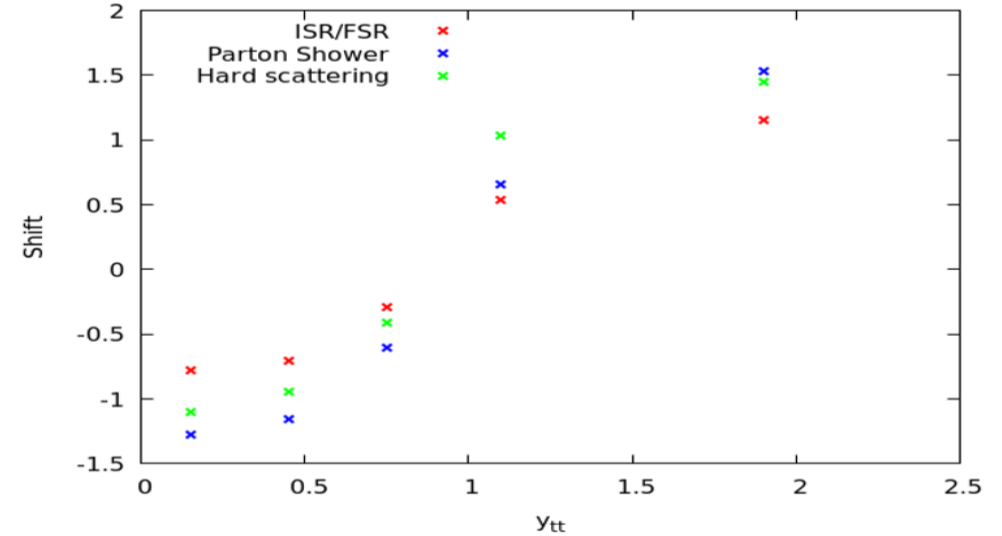
$$\beta_i^1 = \cos\left[\pi\left(\frac{y_{tt,i} - y_{tt,min}}{y_{tt,max} - y_{tt,min}}\right)\right]\beta_i^{tot}$$

$$\beta_i^2 = \sin\left[\pi\left(\frac{y_{tt,i} - y_{tt,min}}{y_{tt,max} - y_{tt,min}}\right)\right]\beta_i^{tot}$$

- Data makes gluon softer at high $x$, exact details more sensitive to the decorrelation.



|  | Before decorrelating | After decorrelating |
|---|---|---|
| pT | 2.38 | 0.52 |
| yt | 1.84 | 2.11 |
| ytt | 2.21 | 0.79 |
| mtt | 1.81 | 0.72 |
| pen | 0.88 | 0.72 |
| tot | 2.96 | 1.67 |

# Further Data

# D0 W asymmetry ("new" data)

- Include D0 electron asymmetry data instead as a $W$ asymmetry.

- Requires correcting back from leptons to W bosons, however this prevents the W asymmetry from being washed out in the lepton asymmetry and the impact on the PDF uncertainty of this correction is small.

- In contrast statistical errors are larger, particularly at high rapidity.

- Overall reduced uncertainty bands seen when using the D0 data as a $W$ asymmetry rather than electron asymmetry for both MMHT14+Hera and MMHT16. PDFs nonetheless consistent within errorbands.

- Smaller reduction seen for MMHT16 as various additional LHC datasets reduce its effect.
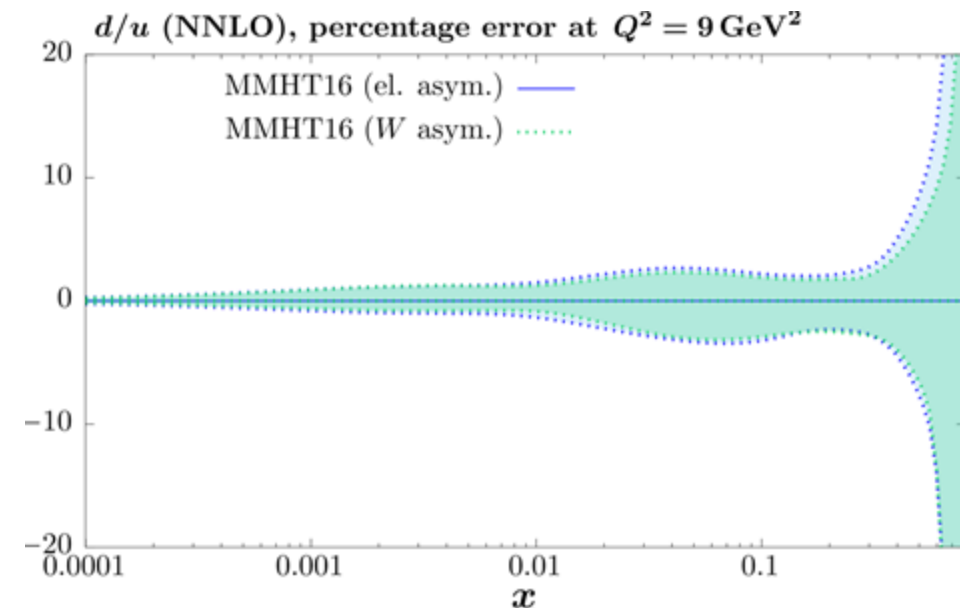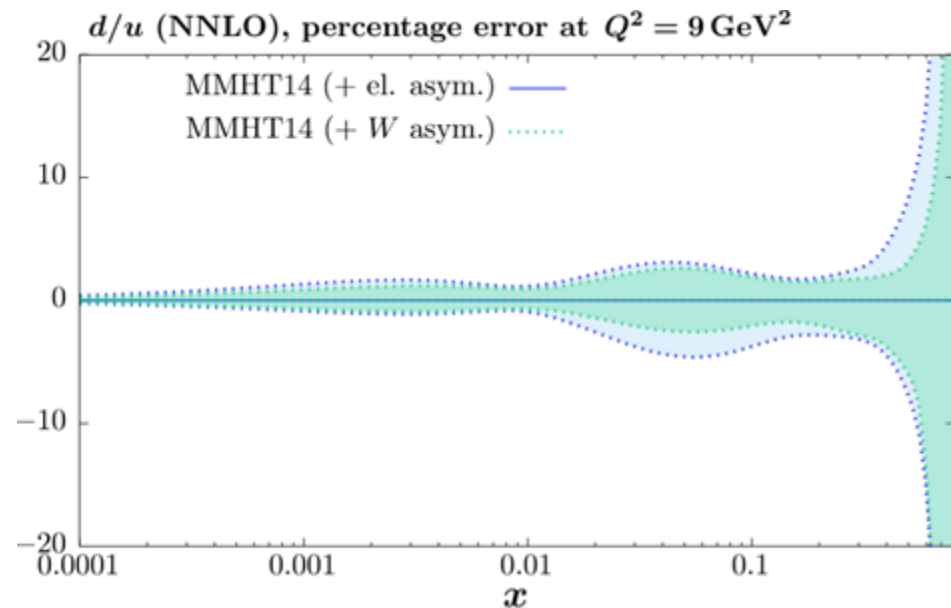
# D0 W asymmetry ("new" data)
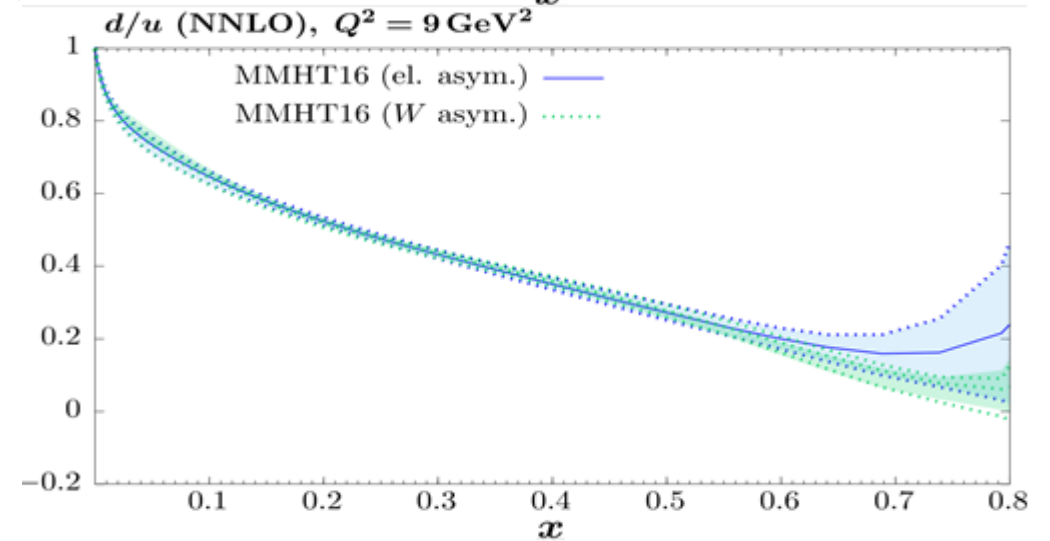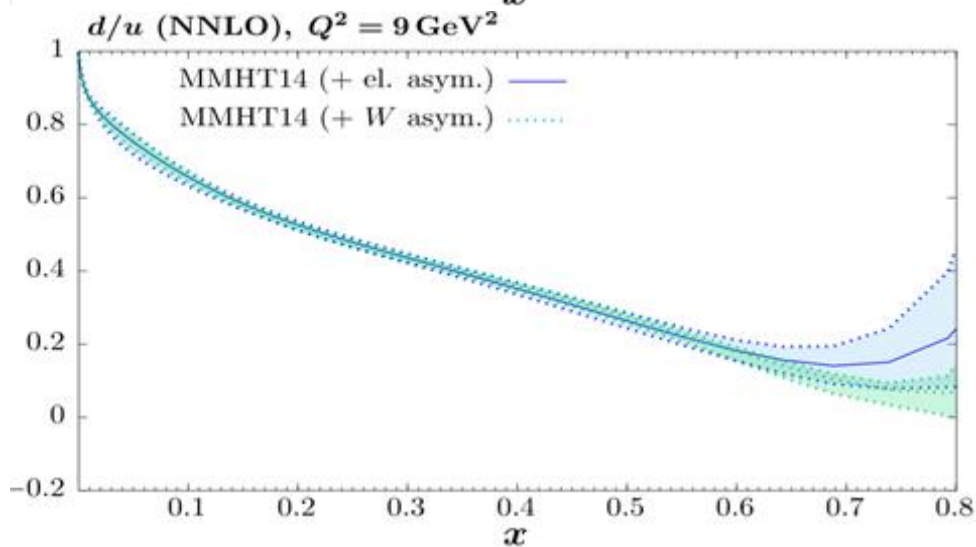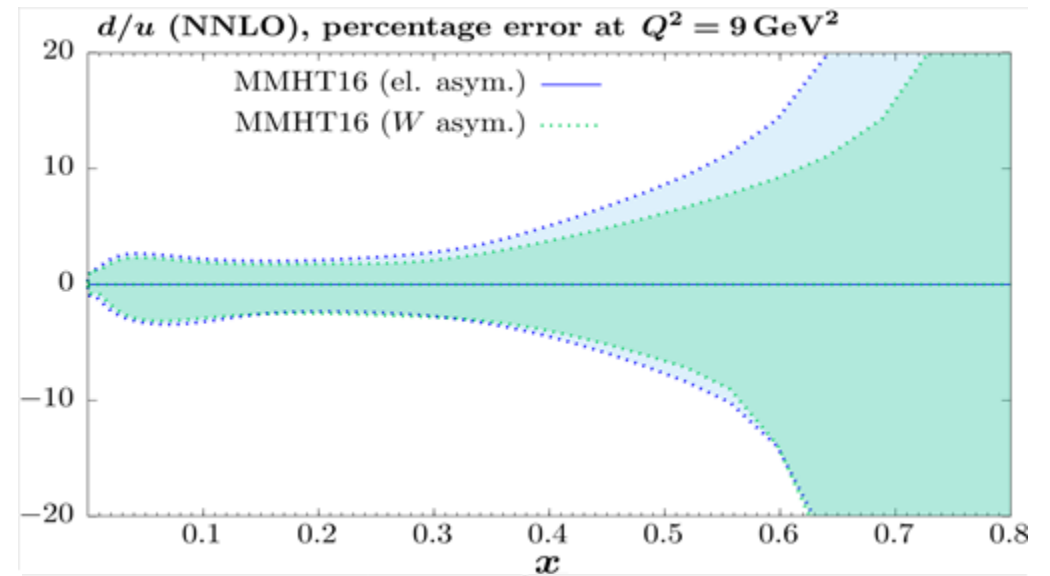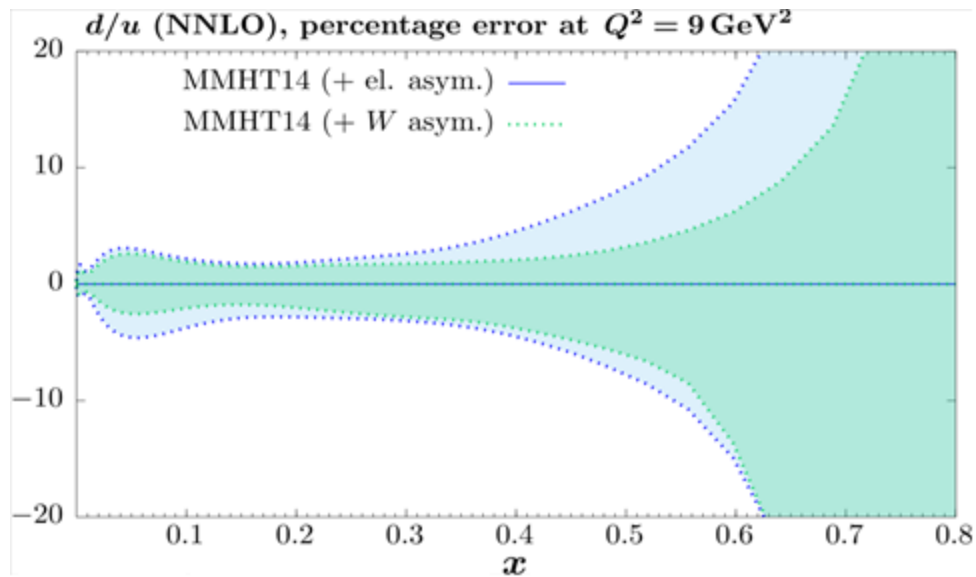
# Future Data to be added for "MMHT19"

- D0 W asymmetry already described, under investigation.

- ATLAS and CMS 8 TeV Z pT data [1512.02192, 1504.03511], issues (NNPDF) fitting Atlas 7TeV data.

- CMS 8 TeV jets [1609.005331], (currently not ATLAS 8TeV jets as no NNLO K-factors available yet).

- HERA combined heavy flavour [1804.01019] (under preliminary investigation).

- ATLAS, CMS 8 TeV differential top data (lepton + jet, dilepton) [1511.04716, 1505.04480].

All currently being added, should be all included in "MMHT19" for release within a few months.

– Further data (e.g. at 13TeV) will not be added for "MMHT19" but will be compared against, for example CMS 13TeV W+c [1811.10021], favours reduced $R_s$ (i.e. reduced $s + \bar{s}$).

– ATLAS W + jet data may be included if we can get the theory calculation.

– Note CMS 8 TeV double differential DY (1412.1115) is not fit by us as we cannot get a good fit (nor can other groups).
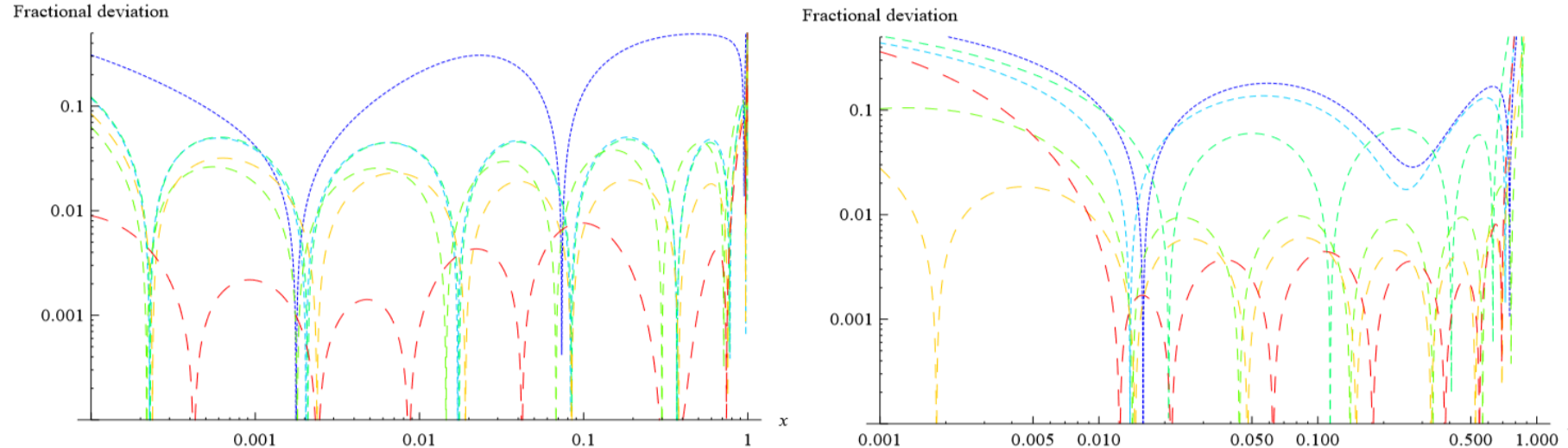
# New Extended Parametrisation

# Reminder MMHT14 parametrisation

- General MMHT14 parameterisation used is:

$$f(x) = A(1-x)^{\eta} x^{\delta}(1 + \sum_{i=1}^{n} a_i T_i(1 - 2\sqrt{x})),$$ where $T_i(1 - 2\sqrt{x})$ are Chebyshev polynomials.

- Higher powers $n$ will allow greater precision for fitting data, however also need sufficient data to constrain any free parameters -> balance to be made.



[Plot from Mathiejssen 1211.1215]

- Illustration of precision possible with increasing $n$ for sea-like (left) and valence-like (right) for pseudodata above.

- MMHT14 used $n = 4$ in most places (leads to few % precision), with some additional variation due to constraints. $n = 6$ (6 Chebyshevs) would lead to better than 1% precision over whole range.

# Reminder MMHT14 parametrisation

- General MMHT14 parameterisation used is:

$$u_V(x, Q_0^2) = A^{u_v}(1-x)^{\eta_{u_v}} x^{\delta_{u_v}}(1 + \sum_{i=1}^{4} a_{i,u_V} T_i(1 - 2\sqrt{x}))$$

6 free parameters as $A^{u_v}$ constrained by sum rule
$$\int_0^1 dx\, u_v = 2$$

$$d_V(x, Q_0^2) = A^{d_v}(1-x)^{\eta_{d_v}} x^{\delta_{d_v}}(1 + \sum_{i=1}^{4} a_{i,d_V} T_i(1 - 2\sqrt{x}))$$

6 free parameters as $A^{d_v}$ constrained by sum rule
$$\int_0^1 dx\, u_v = 1$$

$$s_{+,-} = s \pm \bar{s}$$

$$S(x, Q_0^2) = A^S(1-x)^{\eta_S} x^{\delta_S}(1 + \sum_{i=1}^{4} a_{i,S} T_i(1 - 2\sqrt{x}))$$

7 free parameters in sea

$\delta_{s_+} = \delta_S, a_{3,4,s_+} = a_{3,4,S}$, so 4 free parameters in $s_+$

$$s_+(x, Q_0^2) = A^{s+}(1-x)^{\eta_{s+}} x^{\delta_{s+}}(1 + \sum_{i=1}^{4} a_{i,s+} T_i(1 - 2\sqrt{x}))$$

$$s_-(x, Q_0^2) = A^{s-}(1-x)^{\eta_{s-}} x^{\delta_{s-}}(1 - x/x_0)$$

2 free parameters in $s_-$, $\delta_{s_-} = 0.2$, $x_0$ set by strange sum rule

$$\Delta = \bar{d} - \bar{u} \qquad \Delta(x, Q_0^2) = A^\Delta(1-x)^{\eta_\Delta} x^{\delta_\Delta}(1 + \gamma_\Delta x + \epsilon_\Delta x^2)$$
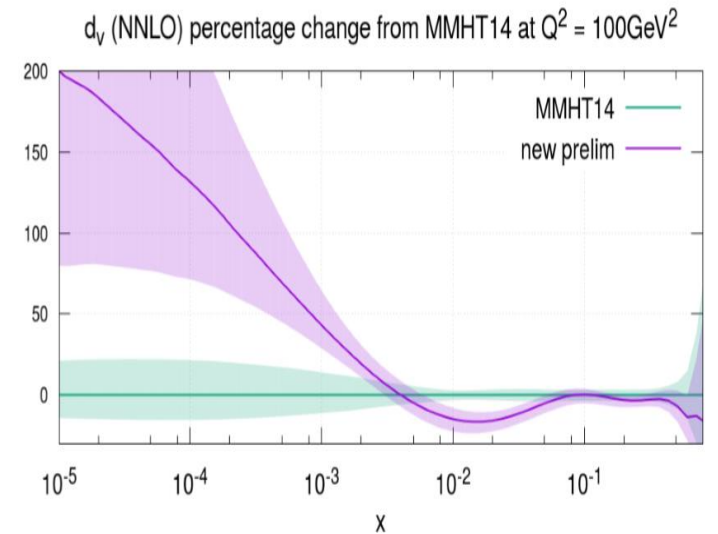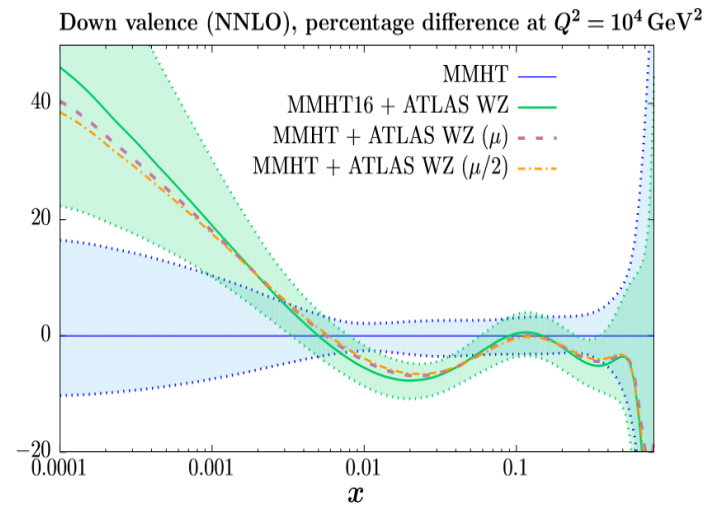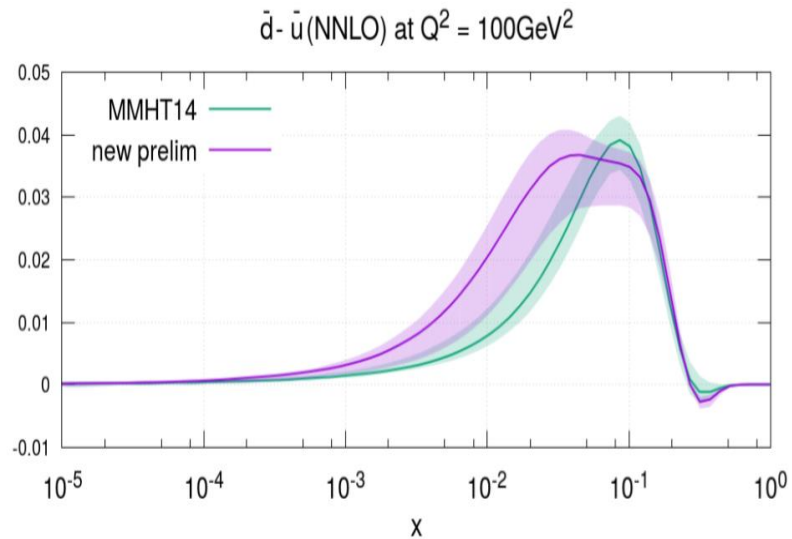
$\eta_\Delta = \eta_s + 2$ so 4 free parameters in $\Delta$

$$g(x, Q_0^2) = A^g(1-x)^{\eta_g} x^{\delta_g}(1 + \sum_{i=1}^{2} a_{i,g} T_i(1 - 2\sqrt{x})) + A^{g'}(1-x)^{\eta_{g'}} x^{\delta_{g'}}$$

7 free parameters as $A^g$ constrained by momentum sum rule $\int_0^1 dx\, x[u_v + d_v + S + g] = 1$

-> 36 free parameters in total + $\alpha_S$ free -> 37, formed into 25 eigenvector pairs for determination for uncertainties.

# Extending PDF parametrisation

- Aim: With more data now available we can aim for 6 Chebyshevs to enable a better fit to data -> also gives more eigenvectors and a better estimation of pdf errorbands.

- Start by writing $\Delta(x, Q_0^2) = (\bar{d} - \bar{u})(x, Q_0^2)$ in terms of 6 Chebyshevs.

- Allows multiple turning points (including going negative around $x = 0.3$) with clear peak shape change and eases tension between Atlas W,Z data and E866 DY ratio data -> fit improves by $\Delta\chi^2 = 17.6$.

- Then also extending $d_v$ to 6 Chebyshevs allows a further improvement in the fit of LHC data.

- Biggest change occurs in $d_v(x, Q^2)$ due to combined effect of ATLAS W,Z 7 TeV data and extended parametrisation.

- Get increased uncertainty at very small and very large $x$ due to extended parametrisation.



$\bar{d} - \bar{u}$(NNLO) at $Q^2 = 100 \text{GeV}^2$

Down valence (NNLO), percentage difference at $Q^2 = 10^4 \text{GeV}^2$

$d_v$ (NNLO) percentage change from MMHT14 at $Q^2 = 100\text{GeV}^2$

# Extending PDF parametrisation

- Extending $d_v$ allows better fit of LHC data, E866 data fit was already close to optimal with extended $\bar{d} - \bar{u}$.

- LHC lepton asymmetry improves, but D0 worsens.

- Extending the parametrisations of the other PDFs ($u_v$, $d_v$, $S$, $s_+$, g but not $s_-$) shows main improvements from $d_v$ (Atlas W,Z data) and gluon.

  – Gluon improvement down to both HERA and LHC data.

  – Increase number of eigenvectors from 25 (MMHT14) to 30 pairs, one additional parameter for each PDF other than the light sea ($S$) and $s_-$.

  – Additional eigenvectors allow better determination of uncertainties.

| Data set | $-\Delta\chi^2$ $(\bar{d}-\bar{u})$ | $-\Delta\chi^2$ $(\bar{d}-\bar{u}),d_v$ | $-\Delta\chi^2$ All |
|---|---|---|---|
| Total | 17.6 | 34.0 | 48.9 |
| BCDMS $F_2^p$ | -4.6 | -3.3 | -2.7 |
| BCDMS $F_2^d$ | -2.7 | 4.9 | 8.5 |
| NMC $F_2^n/F_2^p$ | 6.5 | 6.1 | 6.0 |
| NuTeV $F_3^N$ | -0.3 | 1.7 | 3.2 |
| E866 $\sigma(pd)/\sigma(pp)$ | 8.2 | 10.1 | 11.0 |
| NuTeV dimuon | 0.7 | 1.0 | 3.0 |
| HERA I+II $\sigma(e^+p)$ 920 GeV | 1.1 | 1.7 | 4.6 |
| CMS $pp \to l^+l^-$ | 0.7 | 1.8 | 3.1 |
| D0 $\sigma(e^+) - \sigma(e^-)$ | -1.2 | -3.4 | -1.4 |
| CMS $8\,\mathrm{TeV}\,\sigma(l^+) - \sigma(l^-)$ | 4.4 | 5.0 | 4.6 |
| ATLAS $7\,\mathrm{TeV}\,W,Z$ | -0.5 | 2.2 | 4.3 |
| CMS $7\,\mathrm{TeV}$ jets | -0.5 | 0.2 | 3.2 |

# Extending PDF parametrisation

- However, errorbands for $\Delta(x, Q^2) = (\bar{d} - \bar{u})(x, Q^2)$ still go to zero as $x \to 0$ due to parametrisation as the small x power in the $(\bar{d} - \bar{u})$ parametrisation is positive and so suppresses the central value and its errorbands, rather than due to data.

- We can avoid this by instead parametrising $\rho(x, Q^2) = \frac{\bar{d}}{\bar{u}}(x, Q^2)$, again we use 6 Chebyshevs:

$$\rho(x, Q^2) = A^\rho (1 - x)^{\eta_\rho} (1 + \sum_{i=1}^{6} a_{i,\rho} T_i(1 - 2\sqrt{x}))$$

- No small x power to ensure $\frac{\bar{d}}{\bar{u}} \to$ constant as $x \to 0$, also ensures errorbands are not suppressed to 0 as $x \to 0$.

- Similar to CT14, who ensure $\bar{d}$ and $\bar{u}$ are parametrised (although with 4 Bernstein polynomials) so $\frac{\bar{d}}{\bar{u}} \to 1$ as $x \to 0$, whilst ABM format their $\bar{d} - \bar{u}$ parametrisation to allow large non-zero values at small $x$.

- Whilst Chebyshev coefficients are free, the fit actually finds $1 + \sum_{i=1}^{6} a_{i,\rho} \to 1$ as $x \to 0$, thereby finding $\frac{\bar{d}}{\bar{u}} \to 1$.

- Refitting with our new $\frac{\bar{d}}{\bar{u}}$ parametrisation we find a further improvement of $\Delta\chi^2 = 6$ in the global fit.

# Extending PDF parametrisation

- Enhanced PDF parametrisation now fitting $\rho = \bar{d}/\bar{u}$ (with 6 Chebyshevs), shown is $-\Delta\chi^2$ relative to MMHT14.:

| Dataset | $-\Delta\chi^2(\bar{d} - \bar{u}$ 6 Chebyshevs) | $-\Delta\chi^2(\bar{d}/\bar{u}$ 6 Chebyshevs) |
|---|---|---|
| Total | 48.9 | 55.2 |
| BCDMS | 5.8 | 3.0 |
| NMC $F_2^n/F_2^p$ | 5.1 | 6.1 |
| NuTeV $F_3^N$ | 3.2 | 2.5 |
| E866 DY ratio $\sigma(pd)/\sigma(pp)$ | 11.0 | 10.1 |
| NuTeV dimuon | 3.0 | -0.3 |
| Hera CC $\sigma(e^+p)$ | 0.6 | 5.7 |
| Hera CC $\sigma(e^-p)$ | -0.2 | -3.7 |
| Hera I+II $\sigma(e^+p)$ NC | 4.2 | 6.4 |
| Hera I+II $\sigma(e^-p)$ NC | -2.1 | 1.9 |
| D0 $\sigma(e^+) - \sigma(e^-)$ old+new | 0.1 | -2.4 |
| D0 $\sigma(\mu^+) - \sigma(\mu^-)$ | -1.4 | -2.9 |
| CDF W asymmetry | 1.7 | 4.0 |
| LHCb 7 TeV W+$Z$ | -0.3 | 1.6 |
| LHCb 7+8TeV W,Z muon | -2.5 | -5.1 |
| CMS Neutral current($Z/\gamma$) DY | 3.1 | 10.0 |
| CMS 8TeV W $\sigma(l^+) - \sigma(l^-)$ | 4.6 | 0.5 |
| ATLAS 7TeV W,Z | 4.3 | 4.8 |
| CMS 7 TeV jets | 3.2 | 4.9 |

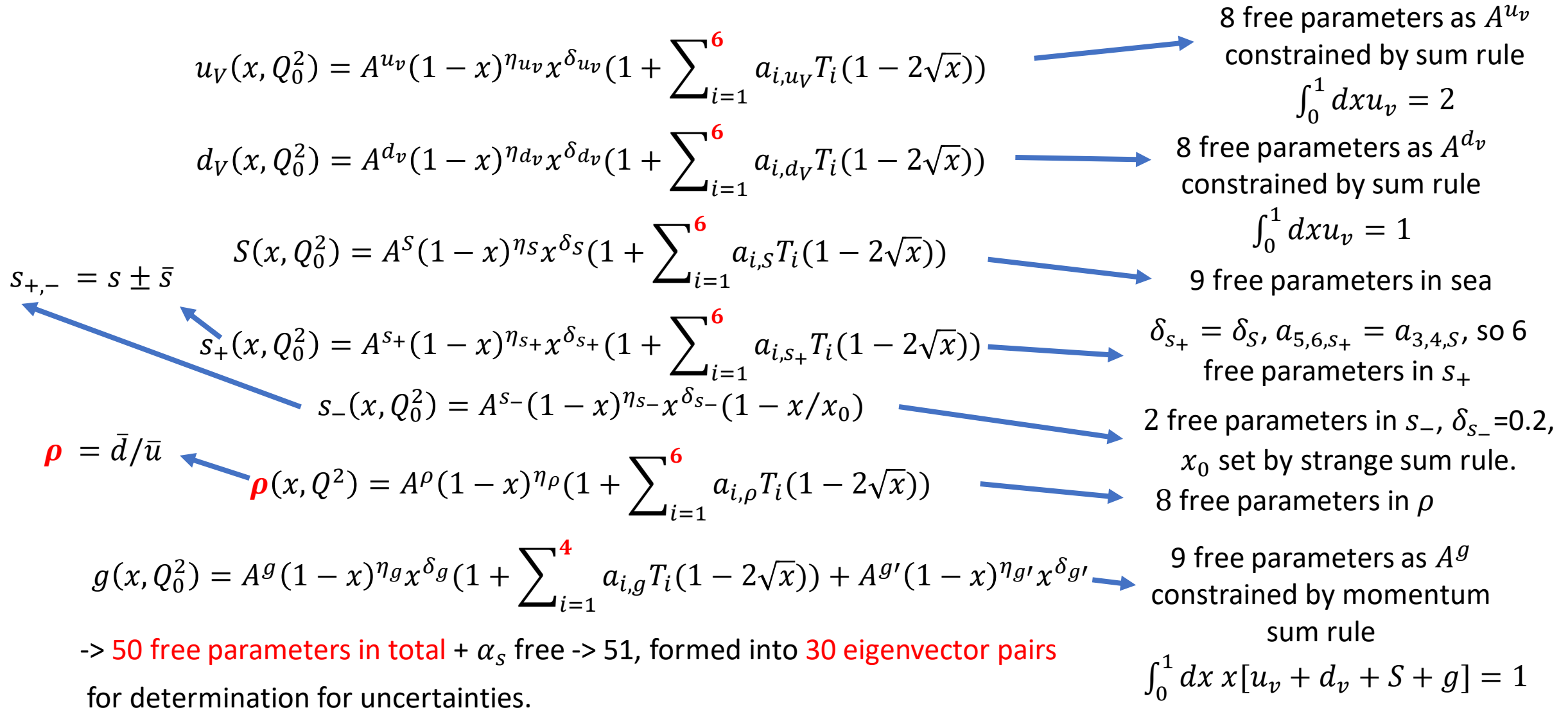Tension between E866 DY ratio data and LHC data (Atlas W,Z mainly) relieved still

Tension between improvement in Hera CC $\sigma(e^+p)$ and worsening in Hera CC $\sigma(e^-p)$, affected by $d_v$, $\bar{d}$, $\bar{u}$ changes around $x \sim 10^{-2}$

Hera improves now, although marginally.

CMS NC DY improves markedly due to change in $d_v$, $\bar{d}$, $\bar{u}$; but CMS CC lepton asymmetry worsens.

CMS 7 TeV jet fit improves.
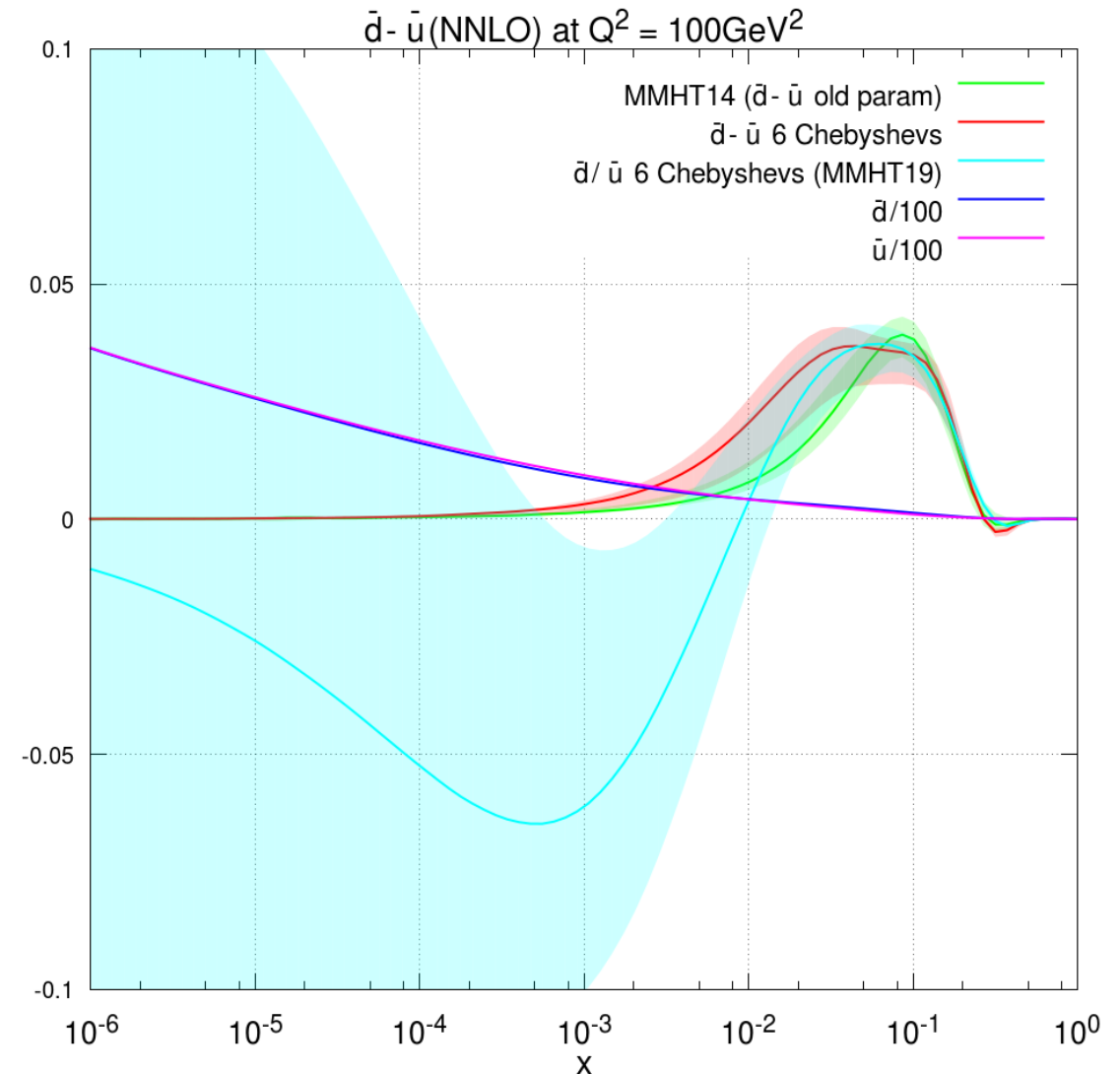
# New "MMHT19" parametrisation

$$u_V(x, Q_0^2) = A^{u_v}(1-x)^{\eta_{uv}} x^{\delta_{uv}}\left(1 + \sum_{i=1}^{6} a_{i,u_V} T_i(1-2\sqrt{x})\right)$$

8 free parameters as $A^{u_v}$ constrained by sum rule
$$\int_0^1 dx\, u_v = 2$$

$$d_V(x, Q_0^2) = A^{d_v}(1-x)^{\eta_{dv}} x^{\delta_{dv}}\left(1 + \sum_{i=1}^{6} a_{i,d_V} T_i(1-2\sqrt{x})\right)$$

8 free parameters as $A^{d_v}$ constrained by sum rule
$$\int_0^1 dx\, u_v = 1$$

$$S(x, Q_0^2) = A^{S}(1-x)^{\eta_S} x^{\delta_S}\left(1 + \sum_{i=1}^{6} a_{i,S} T_i(1-2\sqrt{x})\right)$$

9 free parameters in sea

$$s_{+,-} = s \pm \bar{s}$$

$$s_+(x, Q_0^2) = A^{s+}(1-x)^{\eta_{s+}} x^{\delta_{s+}}\left(1 + \sum_{i=1}^{6} a_{i,s+} T_i(1-2\sqrt{x})\right)$$

$\delta_{s_+} = \delta_S$, $a_{5,6,s_+} = a_{3,4,S}$, so 6 free parameters in $s_+$

$$s_-(x, Q_0^2) = A^{s-}(1-x)^{\eta_{s-}} x^{\delta_{s-}}(1 - x/x_0)$$

2 free parameters in $s_-$, $\delta_{s_-} = 0.2$, $x_0$ set by strange sum rule.

$$\boldsymbol{\rho} = \bar{d}/\bar{u}$$

$$\boldsymbol{\rho}(x, Q^2) = A^{\rho}(1-x)^{\eta_\rho}\left(1 + \sum_{i=1}^{6} a_{i,\rho} T_i(1-2\sqrt{x})\right)$$

8 free parameters in $\rho$

$$g(x, Q_0^2) = A^{g}(1-x)^{\eta_g} x^{\delta_g}\left(1 + \sum_{i=1}^{4} a_{i,g} T_i(1-2\sqrt{x})\right) + A^{g'}(1-x)^{\eta_{g'}} x^{\delta_{g'}}$$

9 free parameters as $A^{g}$ constrained by momentum sum rule

-> 50 free parameters in total + $\alpha_s$ free -> 51, formed into 30 eigenvector pairs for determination for uncertainties.

$$\int_0^1 dx\, x[u_v + d_v + S + g] = 1$$

# Extending PDF parametrisation

- 30 eigenvector pairs found out of the 50 free parameters.

- Mean Tolerance T=3.47, similar to MMHT14.

- 29 of the eigenvector directions are constrained most by LHC datasets, predominantly CMS W (and W+c) data and ATLAS W,Z 7 TeV data, with also some constrained by LHCb data.

- 8 directions are constrained by Tevatron data (lepton, W asymmetries and jet data).

- 8 directions are constrained by E866 Drell-Yan ratio data as it is key for constraining $\bar{d} - \bar{u}$ or equally $\bar{d}/\bar{u}$.

- 10 directions are constrained by Fixed target DIS data (BCDMS, SLAC, NMC, some NuTeV), constraining mainly high $x$.

- 2 directions are constrained by each of Fixed target Dimuon data (CCFR, NuTeV) and HERA data.

- Hence a fully global fit is still necessary for full constraint of the PDFs.


- This is all an overly-simplistic view -> lists only most constraining data for each direction, this is influenced by number of data points (which is why LHC does well) and also many directions are constrained by several different datasets to a similar degree.

- HERA data is close to being dominant on a variety of datasets.

# Extending PDF parametrisation – "MMHT19"

- Now plotting $(\bar{d} - \bar{u})(x)$ at $Q^2 = 100$ GeV$^2$, new enhanced parametrisation allows it have the distinct peak shape previously seen at moderate $x$.

- It also allows is to go negative at small $x$, where there are few constraints, before rising at very small $x$.

- Central value tends to 0 as $x \rightarrow 0$ as $\frac{\bar{d}}{\bar{u}} \rightarrow 1$ as one would expect, even though this is not imposed.

- The errorbands are improved, growing as $x \rightarrow 0$ and encompassing the full range of behaviours possible as $x \rightarrow 0$ in different parametrisations.

- Near agreement with previous MMHT14 parametrisation and extended $(\bar{d} - \bar{u})$ parametrisation throughout $x$ range, with biggest departure around $1.2\sigma$.

- Nonetheless, the errors on $(\bar{d} - \bar{u})$ at very small $x$ are still much smaller than the individual $\bar{d}$ and $\bar{u}$.



$\bar{d} - \bar{u}$ (NNLO) at Q$^2$ = 100GeV$^2$

MMHT14 ($\bar{d}$- $\bar{u}$ old param)
$\bar{d}$- $\bar{u}$ 6 Chebyshevs
$\bar{d}$/ $\bar{u}$ 6 Chebyshevs (MMHT19)
$\bar{d}$/100
$\bar{u}$/100

# Extending PDF parametrisation – "MMHT19"

- Now plotting $(\bar{d} - \bar{u})(x)$ at $Q^2 = 100$ GeV², new enhanced parametrisation allows it have the distinct peak shape previously seen at moderate $x$.

- It also allows is to go negative at small $x$, where there are few constraints, before rising at very small $x$.

- Central value tends to 0 as $x \to 0$ as $\frac{\bar{d}}{\bar{u}} \to 1$ as one would expect, even though this is not imposed.

- The errorbands are improved, growing as $x \to 0$ and encompassing the full range of behaviours possible as $x \to 0$ in different parametrisations.

- Near agreement with previous MMHT14 parametrisation and extended $(\bar{d} - \bar{u})$ parametrisation throughout $x$ range, with biggest departure around $1.2\sigma$.

- Nonetheless, the errors on $(\bar{d} - \bar{u})$ at very small $x$ are still much smaller than the individual $\bar{d}$ and $\bar{u}$.



$\bar{d}/\bar{u}$ (NNLO) at $Q^2 = 100$ GeV²

MMHT14 ($\bar{d} - \bar{u}$ old parametrisation)
MMHT14 extended $\bar{d} - \bar{u}$ 6 Chebyshevs
$\bar{d}/\bar{u}$ 6 Chebyshevs

# Extending PDF parametrisation – "MMHT19"

- $s + \bar{s}$, noticeable change in shape, higher over all x, note no NNLO dimuon corrections were included here. Increased uncertainty for $x > 0.6$.

- $s - \bar{s}$, same shape as MMHT14 (as parametrisation not altered) but enhanced amplitude driven by data – overwhelming the Atlas W,Z 7TeV data.

- New data drives $d_v$ higher at low $x$, as allowed by new parametrisation. Also increased uncertainties at very small and very high $x$.



s + s̄ (NNLO) percentage change from MMHT14 at $Q^2 = 100\text{GeV}^2$
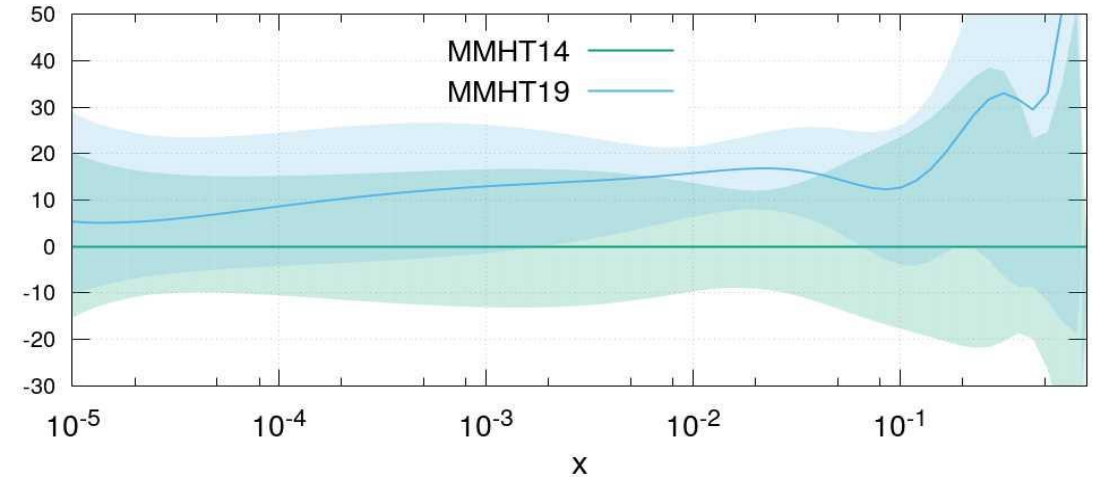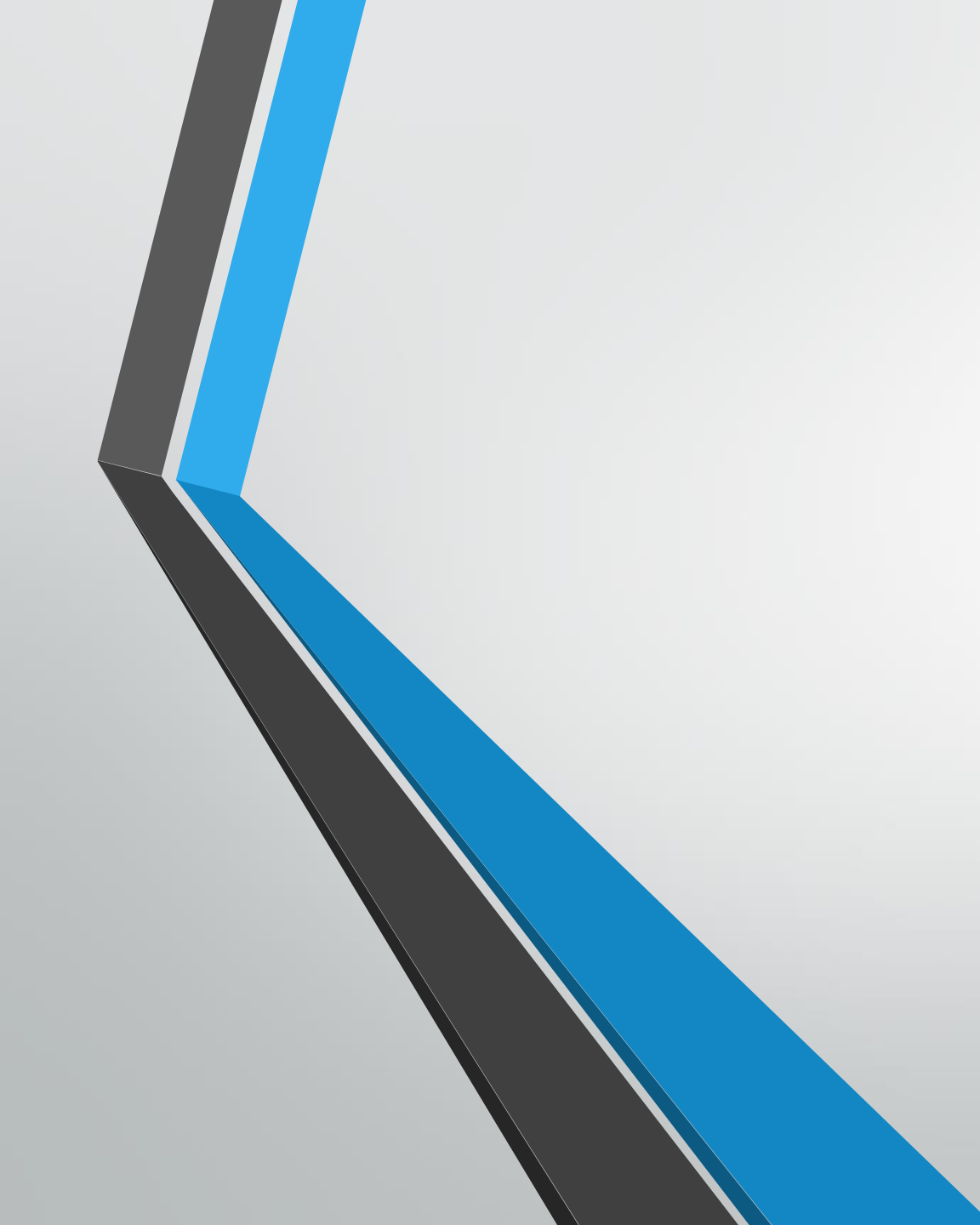


s - s̄ (NNLO) at $Q^2 = 100\text{GeV}^2$
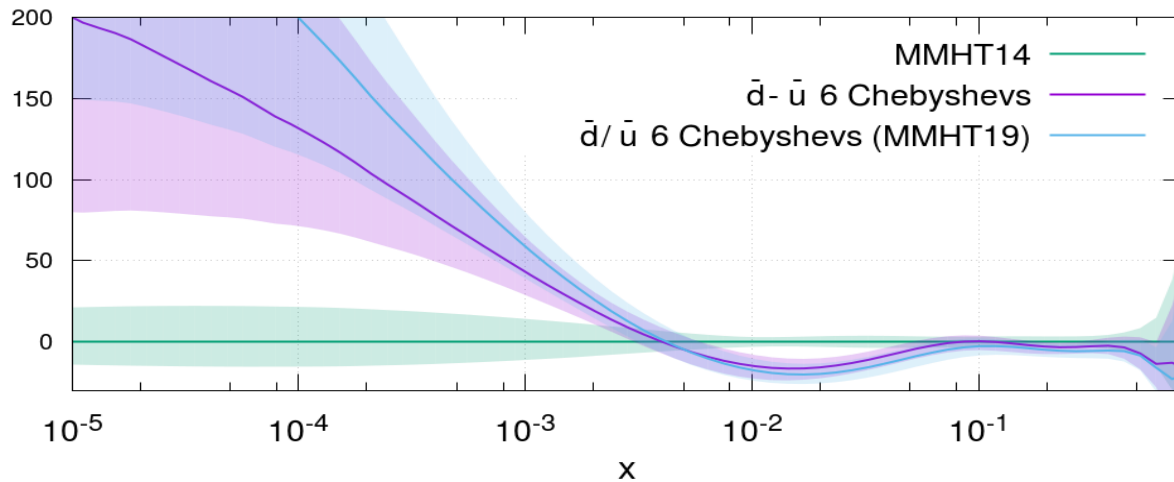


$d_v$ (NNLO) percentage change from MMHT14 at $Q^2 = 100\text{GeV}^2$
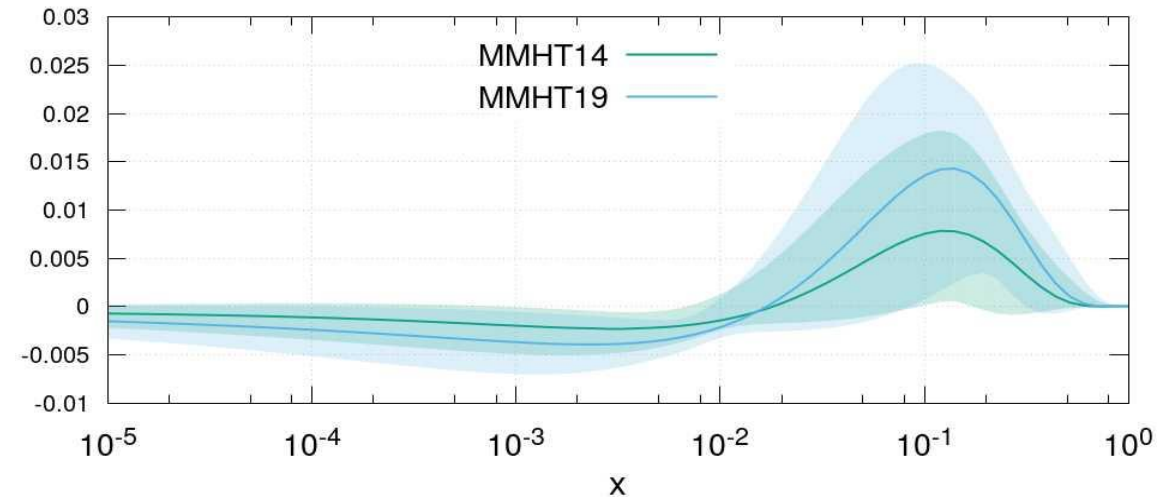
# QED + Photon PDF

# MMHT2015QED – Photon PDF + QED DGLAP

- Precise constraints from elastic and inelastic structure functions used, as in LUXqed*, to give a precise photon PDF at input. In MMHT A1 data -> $F_2^{(el)}$; HERMES and CLAS data -> $F_2^{(inel)}$ .

- Photon PDF and QED DGLAP evolution with $\mathcal{O}(\alpha), \mathcal{O}(\alpha\alpha_s), \mathcal{O}(\alpha^2)$ evolution now added to MMHT (Nathvani).

- Input photon obtained by integrating LUX expression to input scale $Q_0^2 = 1 GeV^2$, therefore higher twist effects, target mass corrections modelled for low $Q^2$ evolution.

- Input photon momentum fraction found to be 0.00195.

- Momentum conservation with a photon PDF and DGLAP QED evolution requires the momentum sum rule be adapted to include the photon PDF. This is imposed at input in MMHT2015QED.

$$\int_0^1 x \left( \Sigma(x, Q_0^2) + g(x, Q_0^2) + \gamma(x, Q_0^2) \right) = 1$$

> Elastic contribution evolution and higher twist corrections -> very small amount ($\mathcal{O}(10^{-4})$) of momentum non-conservation, smaller than PDF uncertainties.

- QED evolution of partons is unidirectional in MMHT, unlike NNPDF3.1luxQED.

- Photon contribution splits into elastic and inelastic contribution, both determined at input by fitting to data:

$$\gamma(x, Q^2) = \gamma^{(el)}(x, Q^2) + \gamma^{(inel)}(x, Q^2)$$

[See Harland-Lang, Martin, Nathvani, Thorne 1907.02750 for more details.]     *[Manohar et al, 1607.04266]

# MMHT2015QED – Photon PDF + QED DGLAP

- Elastic and inelastic contributions represent distinct physical processes -> MMHT2015QED includes their separate contributions in the grids.

- Inelastic contribution evolves as for other partons via DGLAP.

- Elastic contribution evolves independently of parton splittings as it receives contributions above $Q_0^2$ from $F_2^{(el)}$.

- Elastic contribution dominates at low $Q^2$ where you coherently scatter off entire proton.

- Elastic contribution dominates at high $x$, however here uncertainties become increasingly large too.
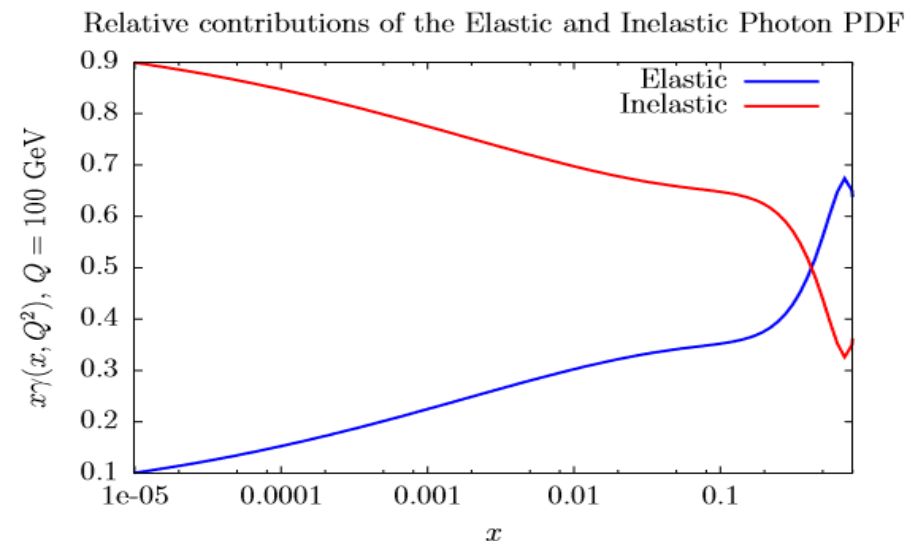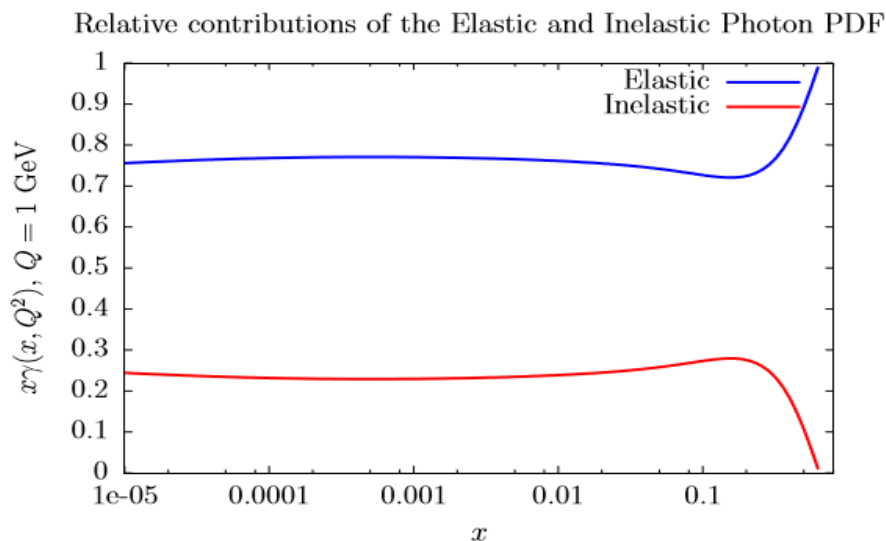


Relative contributions of the Elastic and Inelastic Photon PDF

[See Harland-Lang, Martin, Nathvani, Thorne 1907.02750 for more details.]

# MMHT2015QED – Photon PDF + QED DGLAP

- The effect of the photon and QED evolution on the PDFs is modest.

- $q \rightarrow q + \gamma$ emission, reduces the momenta in quark singlet distributions at high $x$ and correspondingly increases $x\gamma(x, Q^2)$.

- Must be a net loss in quark and gluon momentum to account for photon, given strange is less constrained it reduces most

    -> flavour rearrangement.

- Gluon is reduced over most of $x$ due to need for $\gamma$ momentum.

- Changes in quark singlet distributions and gluon are within PDF uncertainties.

- The up and down valence quarks are most sensitive to QED effects, with a $\mathcal{O}(2 - 5\%)$ increase at low $x$ in their central values.

- Also there is the expected reduction in $u_v$ and $d_v$ at large $x$.

- Essentially no impact on the global fit $\chi^2$ (very small increase), whilst $\alpha_s$ is barely affected, reducing from $0.1181$ to $0.1180$ at NNLO.

[See Harland-Lang, Martin, Nathvani, Thorne 1907.02750 for more details.]
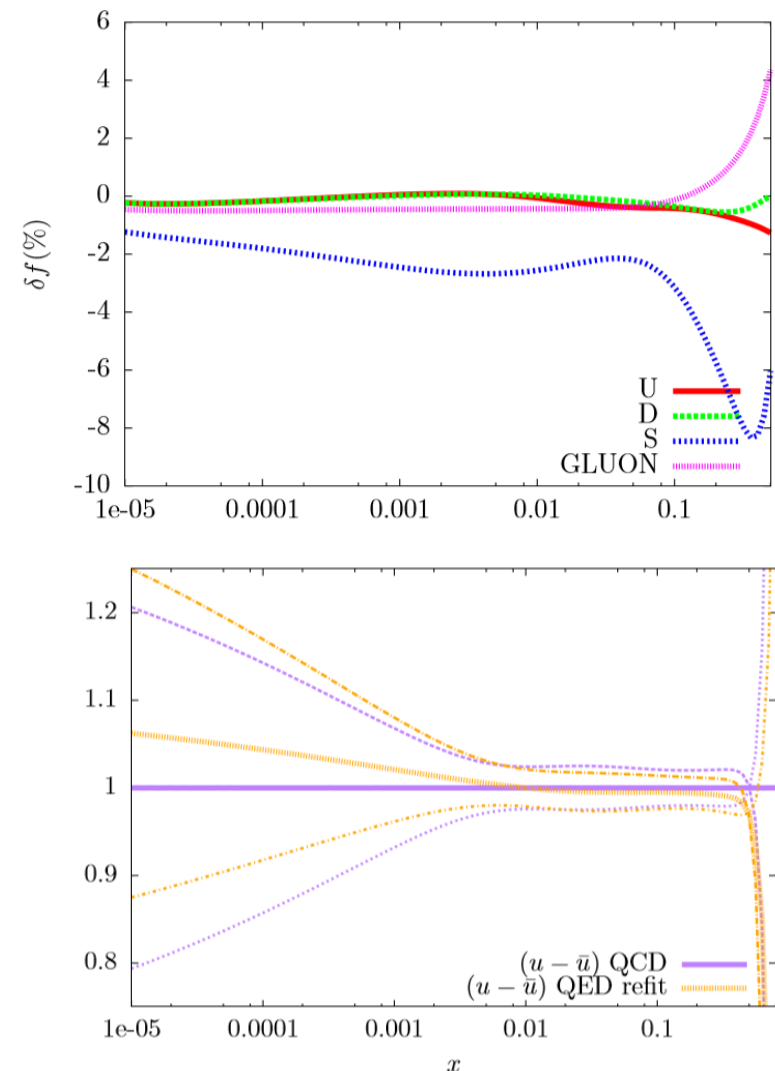
# MMHT2015QED – Photon PDF + QED DGLAP

- Excellent agreement of MMHT2015QED photon PDF with LUXqed and similar modern LUX-based $\gamma(x, Q^2)$.

- All modern sets show remarkable $\mathcal{O}(1\%)$ errors on $\gamma(x, Q^2)$.

- Agreement within 2% over broad range of $x$, diverge at high $x$ where uncertainties are large.

- MMHT2015QED photon slightly larger at intermediate $x$ and slightly smaller at low $x$.

- Reasons for intermediate $x$ difference are twofold:

  1. Difference in quarks – charge weighted singlet $\sum_i e_i^2 (q + \bar{q})$ larger in MMHT2015 than PDF4LHC2015_nnlo_100 due to updated HERA data, this leads to less photon from quark $\gamma$ emission. (Difference at large $x \sim 0.5$ also from $\sum_i e_i^2 (q + \bar{q})$ ).

  2. We exclude lepton splitting in our DGLAP (as no lepton PDFs), this is shown to cause $\mathcal{O}(1-2\%)$ reduction in $\gamma(x, Q^2)$.

[See Harland-Lang, Martin, Nathvani, Thorne 1907.02750 for more details.]



Comparisons of the Photon PDFs
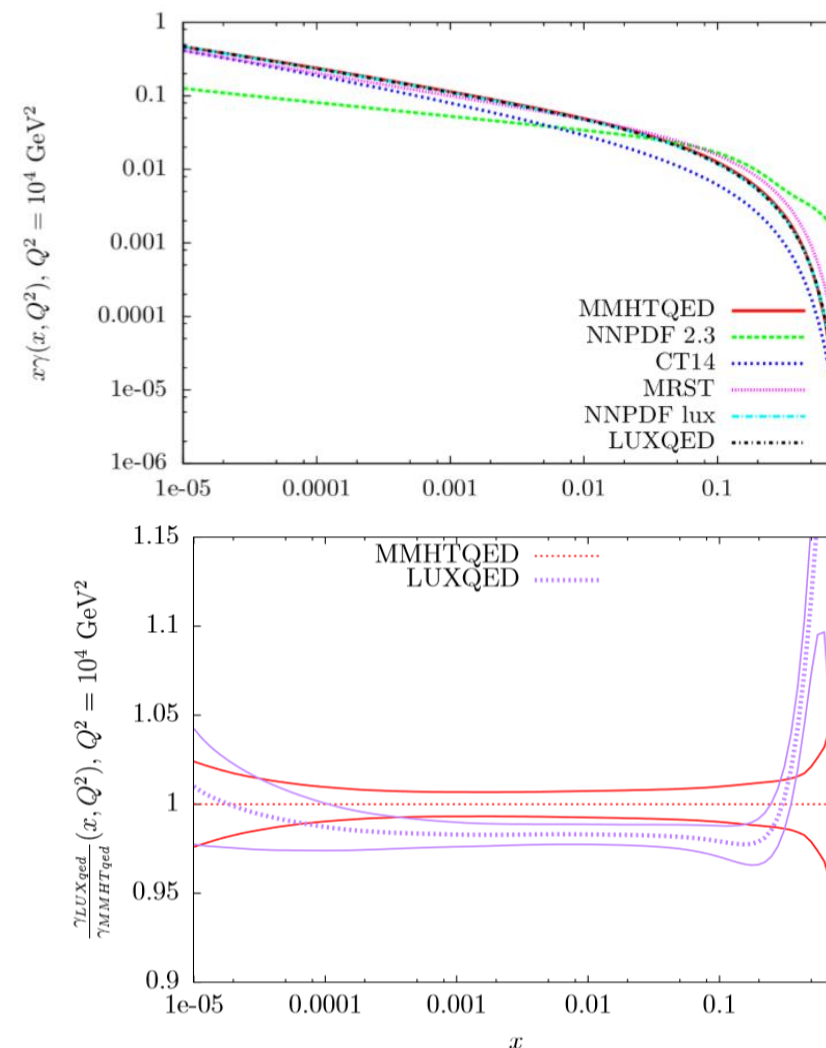
# MMHT2015QED – Effect on High mass DY

- Expect effects of QED on high mass DY (to avoid Z contribution) to be significant, in particular the photon-initiated (PI) contributions.



Theory Prediction/Data (ATLAS 8 TeV 2016), 116 GeV < $M_{ll}$ < 150 GeV
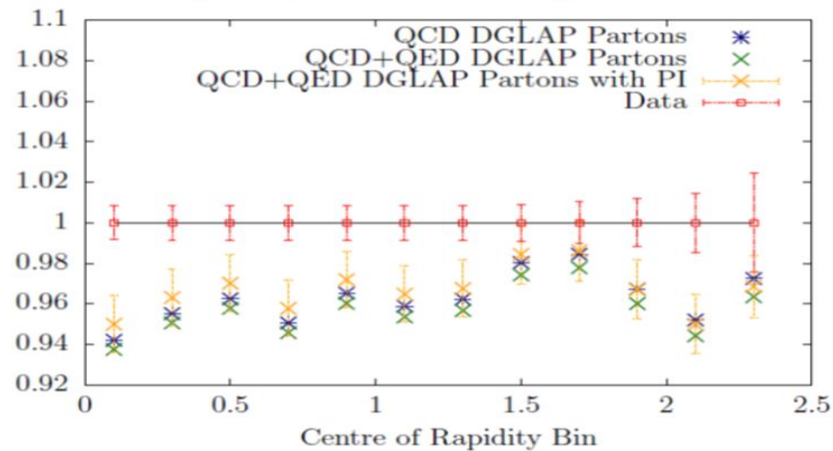
Theory Prediction/Data (ATLAS 8 TeV 2016), 500 GeV < $M_{ll}$ < 1500 GeV
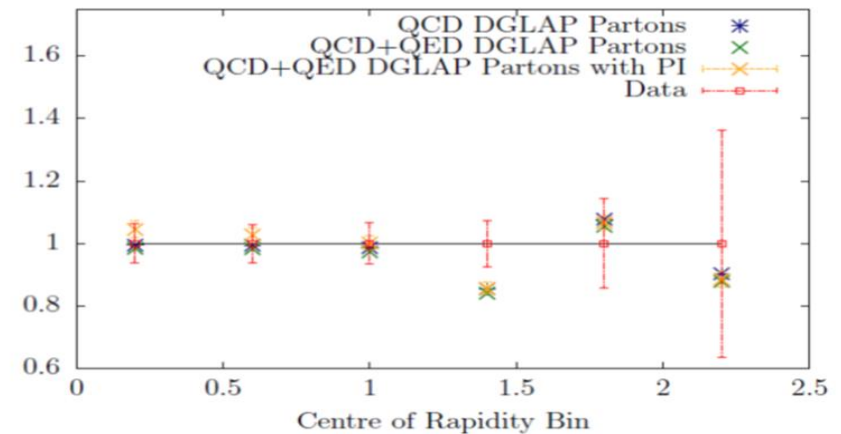
- QED DGLAP causes decrease in dominant $q\bar{q}$ contribution due to reduced quark and antiquark PDFs at high $x$.

- QED PI contributions increase cross-section across all bins as new channel for lepton pair production opens up.

- Refitting causes no further improvement to the fit $\chi^2/N_{points} \sim 65/48$, -> this data offers no meaningful constraint on the photon.

- Nonetheless obtain very mildly reduced PDF uncertainties upon refitting, shown to suggest mild preference for QED DGLAP evolution -> may see stronger hints with more accurate data.

[See Harland-Lang, Martin, Nathvani, Thorne 1907.02750 for more details.]

Conclusions

# Summary and Conclusions

- "MMHT19" PDF set is under development and will be publicly available soon (order a couple/few months) .

- Variety of new datasets included – HERA combined data, final D0 data, LHCb data, ATLAS and CMS EW and jet data.

- Some potential issues observed in some datasets, notably ATLAS 7 TeV jet data and 8 TeV fully differential $t\bar{t}$ data and need for decorrelation.

- Greater number of datasets allow the PDFs to be more precisely constrained and determined -> extended, enhanced parametrisation now in "MMHT19": 6 Chebyshevs for nearly all PDFs and now parametrise $\bar{d}/\bar{u}$, also allows better PDF uncertainty determination.

- MMHT2015QED now available including full QED DGLAP and photon PDF. LHAPDF grids, available including splitting up elastic and inelastic contributions.

- "MMHT19" will include this QED DGLAP and corrections as standard.

# Backup

# MMHT14,16,19

- MMHT14 was the last public release of the MRST, MSTW, MMHT PDF collaboration sets.

- It included for the first time early LHC data, as well as further Tevatron and HERA data and developing the MMHT framework for PDF fitting and uncertainty determination.

- Since then much more (largely LHC) data is available and has been steadily added to the PDF set; MMHT16 was an unofficial set including the final HERA combined data (and some new LHC data), and there has been various LHC EW and jet data also added since.

- Now is a time for a new public PDF set – "MMHT19" with a significant amount of LHC Run 1 data at 7 and 8 TeV included.

- Further significant improvements in the methodology and inclusion of higher order QCD predictions and also QED corrections for the first time.

- "MMHT19" due to come out shortly (order a couple/few months).

# HERA combined dataset   *[H. Abramowicz et al., (2015), 1506.06042]

- New since MMHT14 but not MMHT16.

$$\chi^2{}_{NNLO}{}^{no\ refit} = 1503/1185 = 1.27 \text{ per point}$$

$$\chi^2{}_{NNLO}{}^{refit} = 1457/1185 = 1.23 \text{ per point}$$

with $\Delta\chi^2 = 12$ in non-HERA data

- Fit just HERA run I+II data

$$\chi^2{}_{NNLO}{}^{refit} = 1385/1185 = 1.17 \text{ per point}$$

- Indicates some potential tension between HERA combined dataset and other data in MMHT14 global fit.

- Tensions in CC e⁻p data and NC 920GeV $e^+p$ with other global fit data.

- HERA data prefers $u_v$ smaller in intermediate $x$ and larger at large $x$ than predicted from the fixed target proton data.

| | no. points | NNLO $\chi^2_{\mathrm{HERA}}$ | NNLO $\chi^2_{\mathrm{global}}$ |
|---|---|---|---|
| correlated penalty | | 73.0 | 92.1 |
| CC $e^+p$ | 39 | 42.2 | 48.4 |
| CC $e^-p$ | 42 | 47.0 | 59.3 |
| NC $e^-p$ $E_p = 920$ GeV | 159 | 213.5 | 226.7 |
| NC $e^+p$ $E_p = 920$ GeV | 377 | 422.8 | 450.1 |
| NC $e^+p$ $E_p = 820$ GeV | 70 | 71.2 | 69.5 |
| NC $e^-p$ $E_p = 575$ GeV | 254 | 229.1 | 231.8 |
| NC $e^-p$ $E_p = 460$ GeV | 204 | 220.2 | 225.6 |
| total | 1145 | 1319.0 | 1403.5 |

# HERA combined dataset *[H. Abramowicz et al., (2015), 1506.06042]

- Little change in fit quality → little change in central value of PDFs.

- New PDFs well within the MMHT14 uncertainty bands.

- Differences and tensions relative to HERA only fit are evident:

  – HERA only favours smaller/larger $u_v$ in intermediate/large $x$ due to CC e⁻p data. Note bump at $x{\sim}0.3$

  – HERA only has slight preference for softer sea at high $x$ but within errorbands.

  – No softer gluon at large $x$, different to seen in HERAPDF-2.0.



Up valence (NNLO), percentage difference at $Q^2 = 10^4\,\mathrm{GeV}^2$

Down valence (NNLO), percentage difference at $Q^2 = 10^4\,\mathrm{GeV}^2$

Light sea (NNLO), percentage difference at $Q^2 = 10^4\,\mathrm{GeV}^2$

Gluon (NNLO), percentage difference at $Q^2 = 10^4\,\mathrm{GeV}^2$

MMHT2014
MMHT2014 (HERA global)
MMHT2014 (HERA only)
HERAPDF2.0
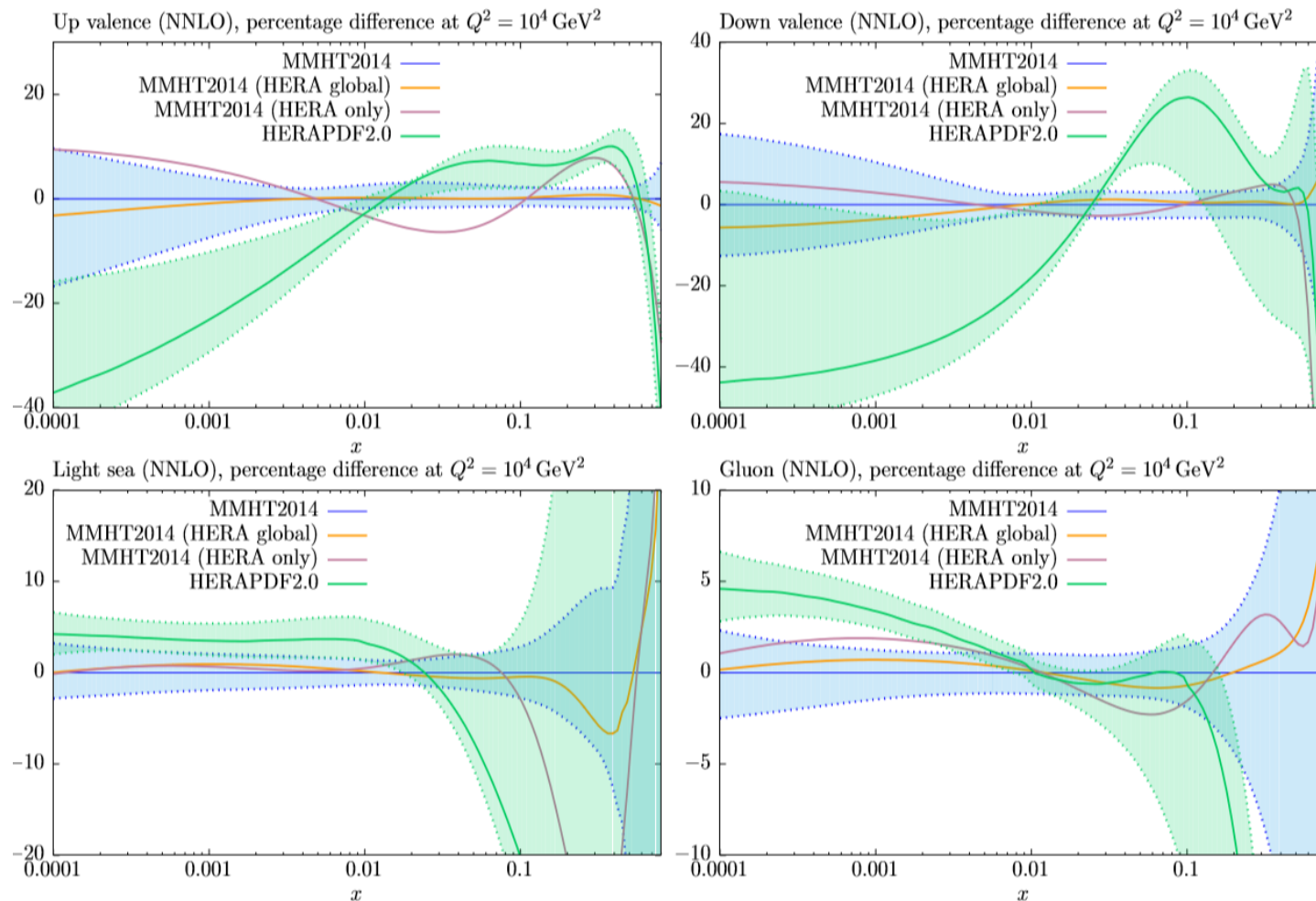
[See Harland-Lang, Martin, Motylinski, Thorne 1601.03413 for more details.]

# HERA combined dataset  *[H. Abramowicz et al., (2015), 1506.06042]

- Whilst little effect on pdf central values HERA data does affect pdf uncertainties:
  - g uncertainty reduced for $x < 0.01$.
  - Small improvements in $u_v$ and $d_v$ for small $x$.
- Reduced uncertainties can effect benchmark cross-section errors:
  - No change in central values for W, Z, Higgs or ttbar production at Tevatron or LHC.
  - But reduction in uncertainties of a few % in W,Z and of 10% in Higgs and ttbar.
- HERA combined dataset provides useful extra constraints, mainly reducing the uncertainty on the gluon pdf.
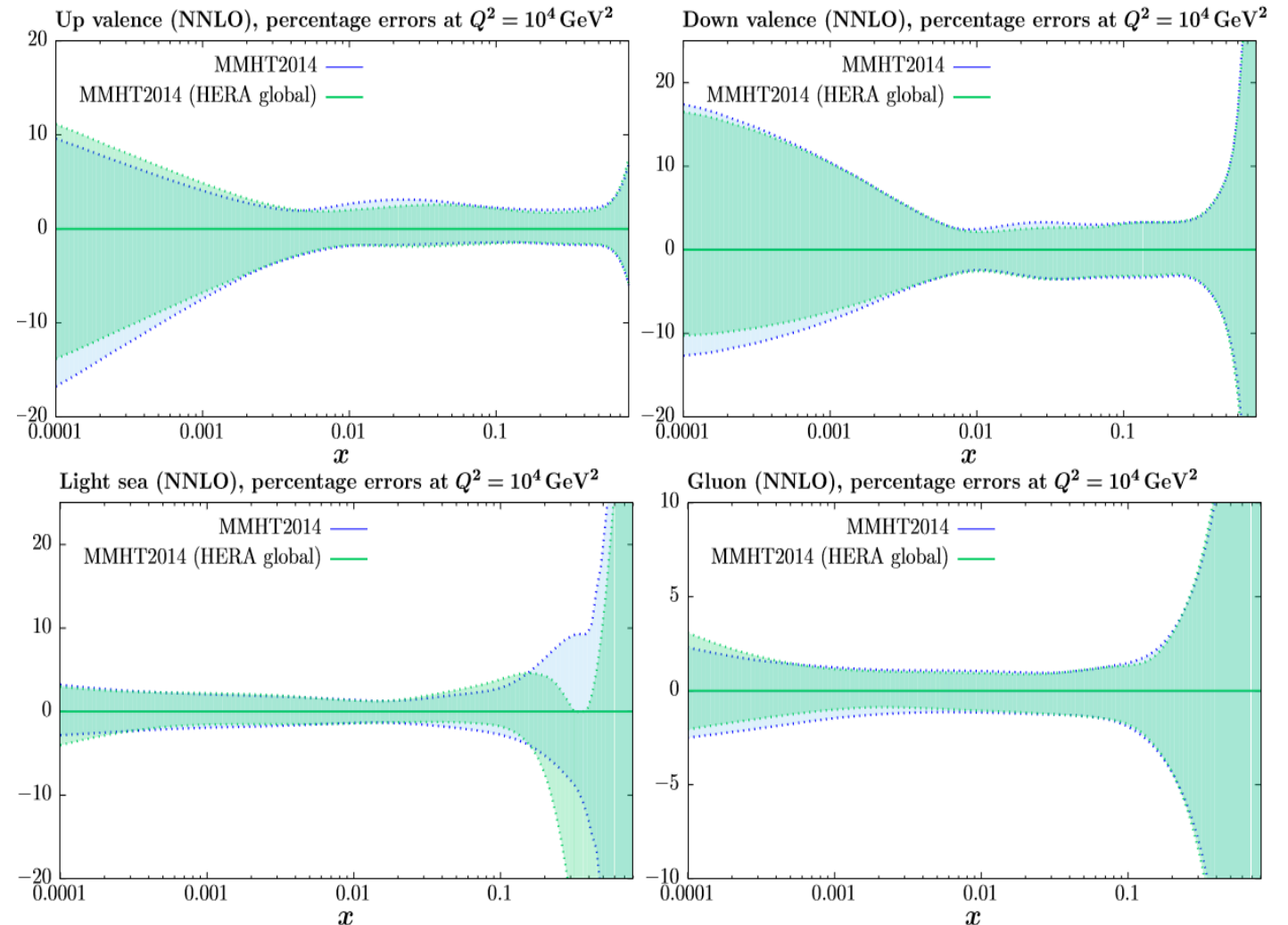
# LHCb W,Z data 7,8 TeV e,μ at high rapidity

- Good agreement between prediction and refitting.

- Z production in muon channel shows theory predictions are a little low at low Z rapidity, issue not seen with electrons, similar issue of low theory at small muon rapidity for W production at 8TeV.

- Involves PDFs at intermediate $x$, here they are already well constrained by DIS data -> therefore little effect on PDF fit, nonetheless in agreement.
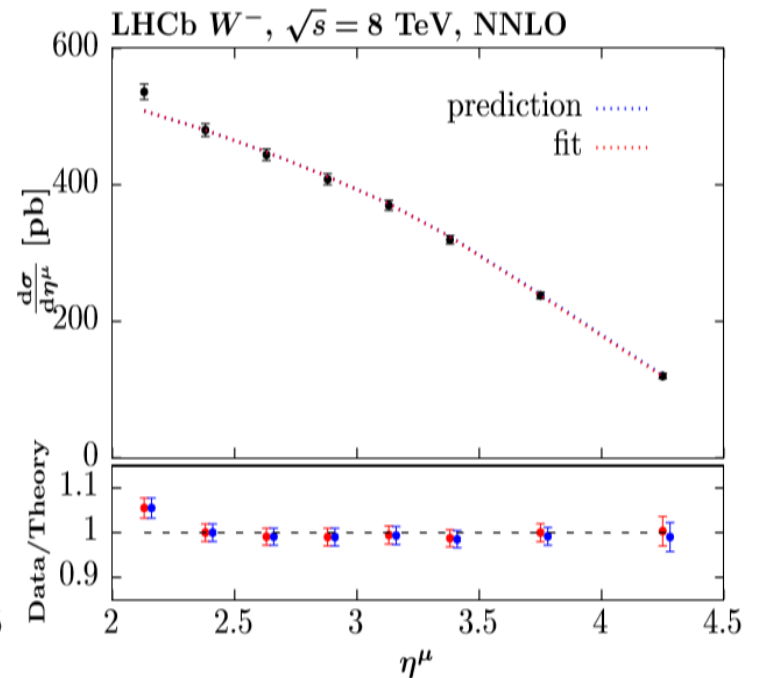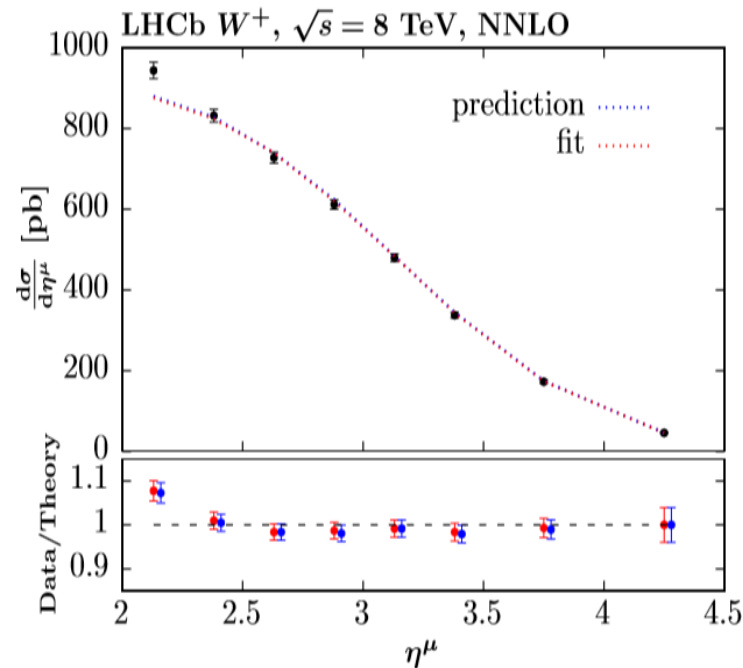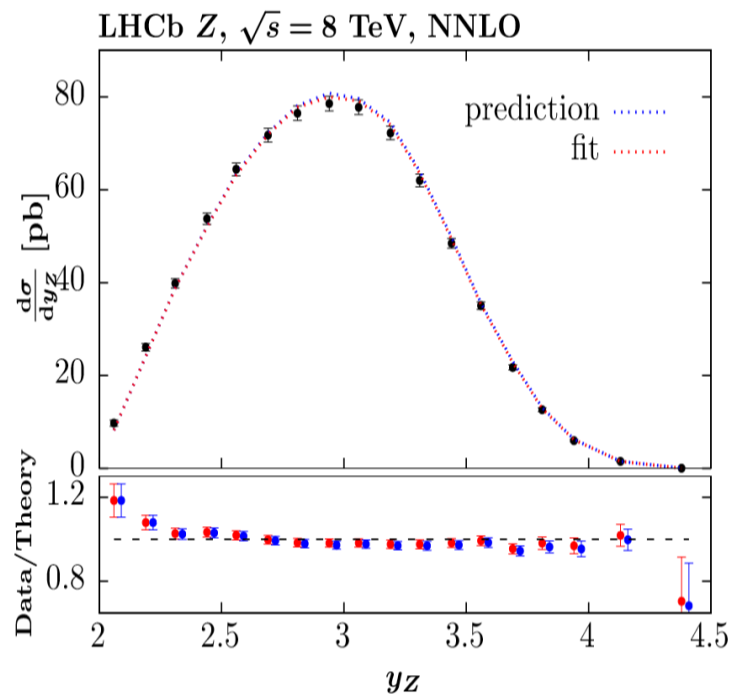
# CMS 8 TeV $W^{\pm}$ rapidity and asymmetry

- Good agreement, individual $W^+$ and $W^-$ are fit as more constraining, can be interpreted as asymmetry.

- After refitting there is a significant improvement in $\chi^2$ -> indicates effect on PDF shape.

- Modifies the small $x$ valence quarks by order of their MMHT14 uncertainty.

- As new LHC data is added small $x$ valence quarks will become more constrained and errorbands will reduce.

-  Most of uncertainty reduction seen is below right from this CMS asymmetry data but some is also from HERA.
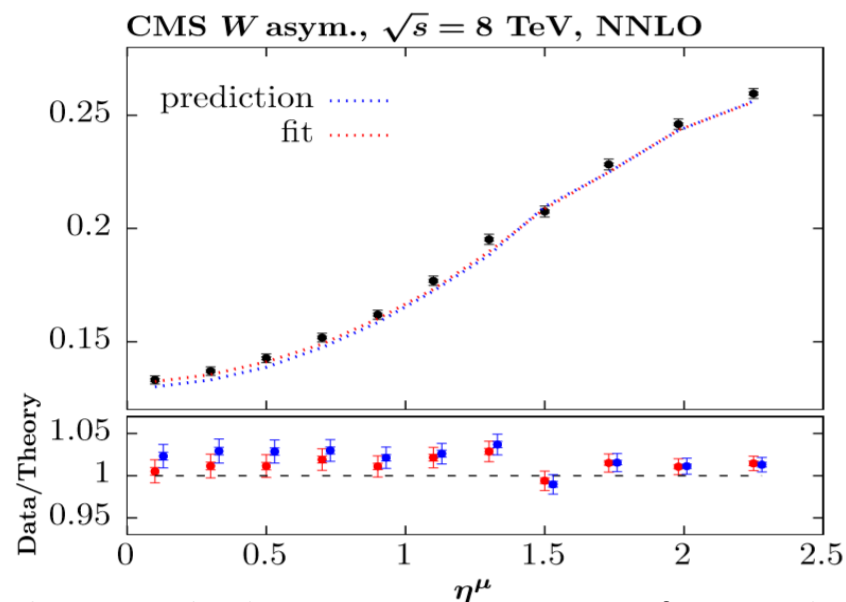


[See Thorne, Harland-Lang, Martin 1708.00047 for more details.]
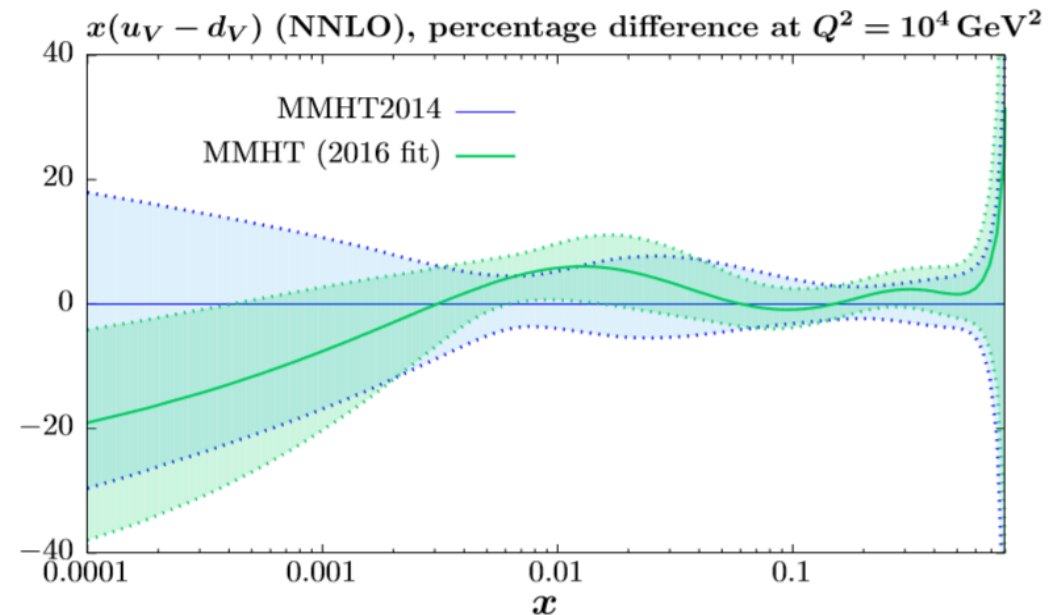
# CMS 8 TeV $W^\pm$ rapidity and asymmetry backup

- Again shows good agreement, individual $W^+$ and $W^-$ are fit as more constraining, can be interpreted together as asymmetry.

- However after refitting we saw significant improvement in $\chi^2$ -> indicates effect on PDF shape.

- Modifies the small x valence quarks by order of their MMHT14 uncertainty.

- Up and down valence distribution shown separately below, can see change in up valence at low x in central value and uncertainty and change in down valence central value at high x with slight reduction in uncertainties around $x \approx 0.1 - 0.2$:
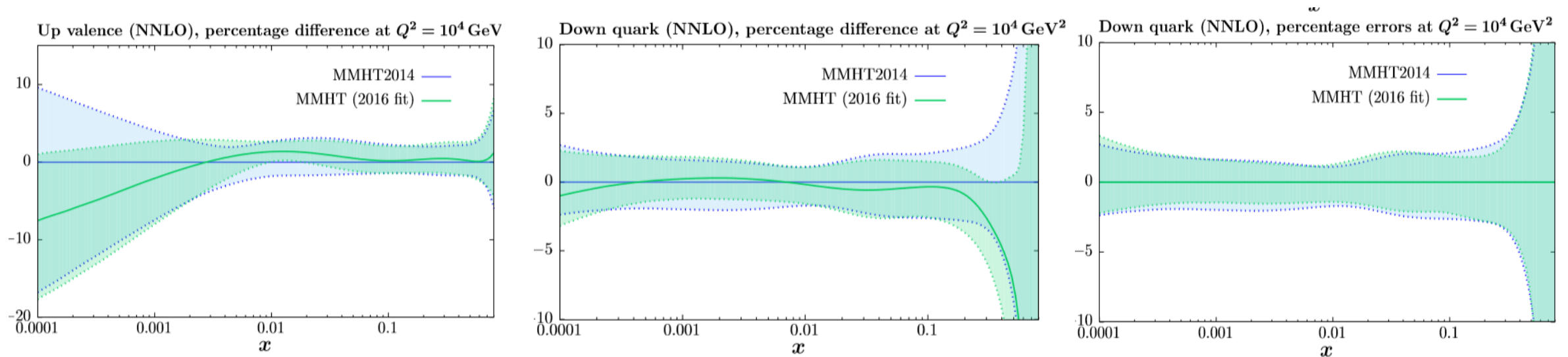
# High precision Atlas data

- Atlas 7 TeV data on high precision $W^+, W^-, Z/\gamma^*$ differential in lepton rapidity for 3 dilepton invariant mass ranges.

- Observe PDF predictions consistently undershoot data for absolute differential cross-sections in rapidity for $W^+, W^-$ (but asymmetry ok) and for $Z/\gamma^*$ at low rapidity.
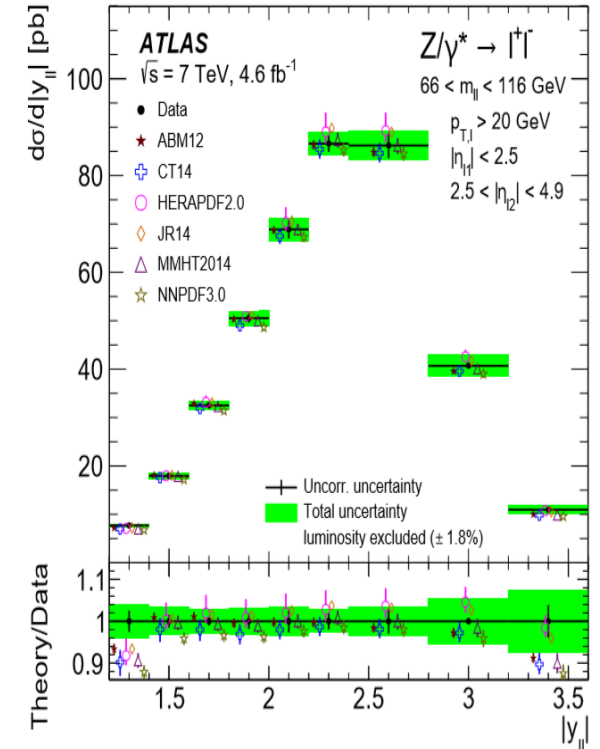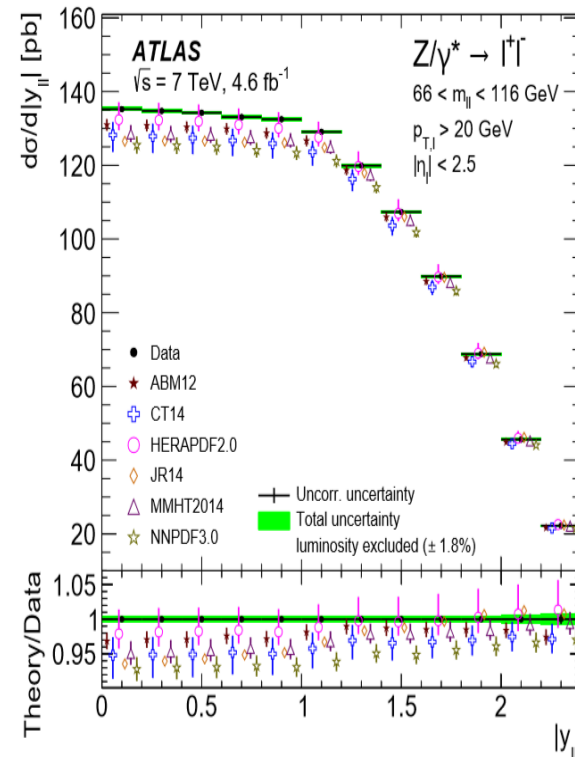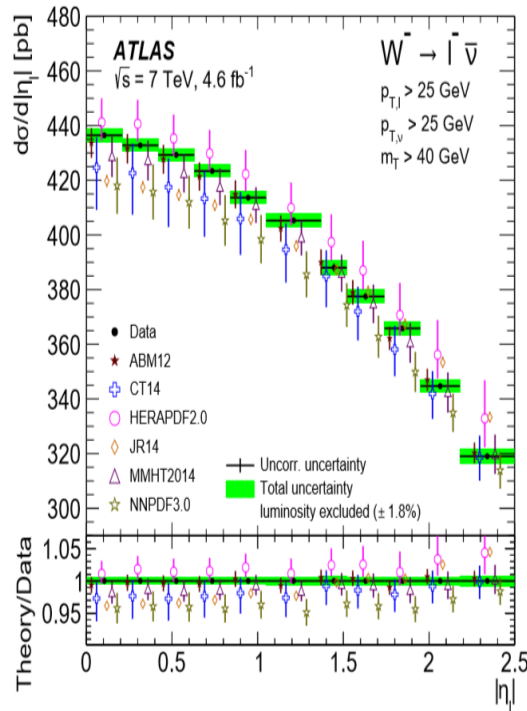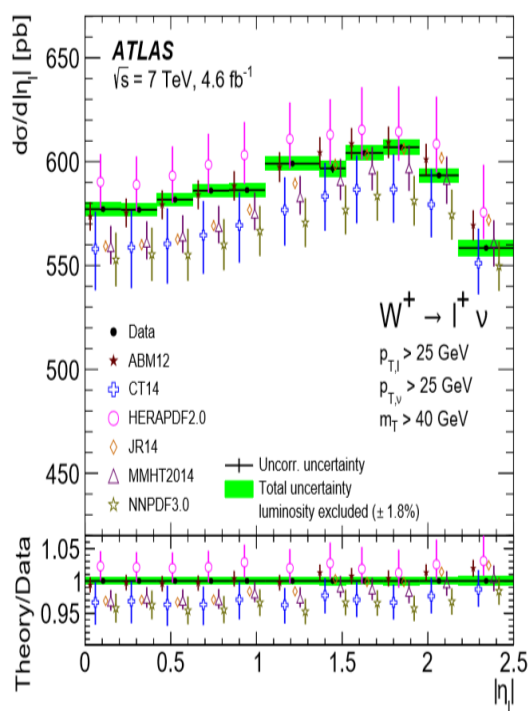
# High precision Atlas data

- Adding this data to MMHT14 PDFs: $\chi^2 \sim 387/61$ (note no PDF errors included in evaluation here).

- Refitting with this ATLAS W,Z data added reduces this to $\chi^2 \sim 130/61$.

- This results in a slight reduction in $W^-$ prediction, a slight increase in $W^+$ and a shape change in $Z$ production.*

(if PDF errors are included: $\chi^2 \sim 115/61$, so possibly compatible)

- Further reduction to $\chi^2 \sim 106/61$ with scales $\mu_{R,F} = M_{W/Z}/2$, with noticeable improvement in $W^+$.

- Clear tension with existing global fit data; increase in $\chi^2$ for CMS $Z/\gamma^*$, dimuon data and others.
- If fit with new LHC data there is an improvement of $\Delta\chi^2 = -10$, -> consistent with other new LHC data (with exception of CMS 7 TeV W+c).
- So different datasets pull in different directions.
- Note also new CMS 13 TeV W+c data doesn't favour large $s + \bar{s}$, not added to "MMHT19".

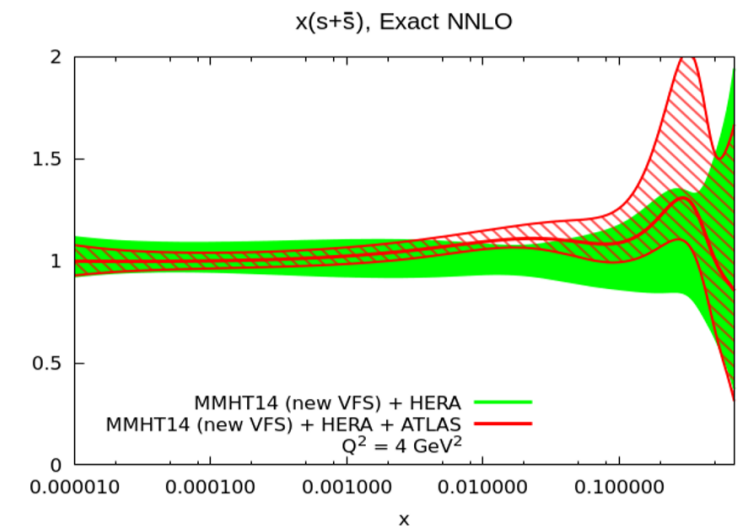*[See Thorne, Harland-Lang, Martin 1708.00047 for more details.]

| Dataset | $\Delta\chi^2$ relative to MMHT14 | |
|---|---|---|
| | With Atlas W,Z data added | With Atlas W,Z data given larger weight |
| CMS double differential $Z/\gamma^*$ | 17 | 24 |
| CCFR/NuTeV dimuon | 10 | 13 |
| E866 Drell-Yan asymmetry | 0 | 4 |
| Fixed Target DIS | 16 | 12 |
| CDF W-asymmetry | 5 | 6 |
| CDF differential $Z/\gamma^*$ | -1 | 4 |
| Hera combined | 2 | 14 |
| Total MMHT14 + Hera comb | 55 | 93 |

# High precision Atlas data

- Effect on PDFs: Large increase in $s + \bar{s}$, and a reduction in its uncertainty.

- Imposes a 25% reduction on the Branching ratio of the charm meson to muon to maintain the CCFR/NuTeV cross-section, however given uncertainty is 10-15% this is ok, but not ideal.

- No change in the down quark contribution to the CCFR/NuTeV cross-section as the down quark is well constrained by DIS and other existing data.

- We find $R_s \sim 0.83 \pm 0.15$, at $x = 0.023$, $Q^2 = 1.9$ GeV$^2$.

- Significant change also in $u_V - d_V$, reduces at low $x$, whilst sea generally increases.

- Tension may be mitigated by NNLO corrections to dimuon production, which are negative and largest at small $x$ (Berger et al, 1601.05430 and 1710.04258).
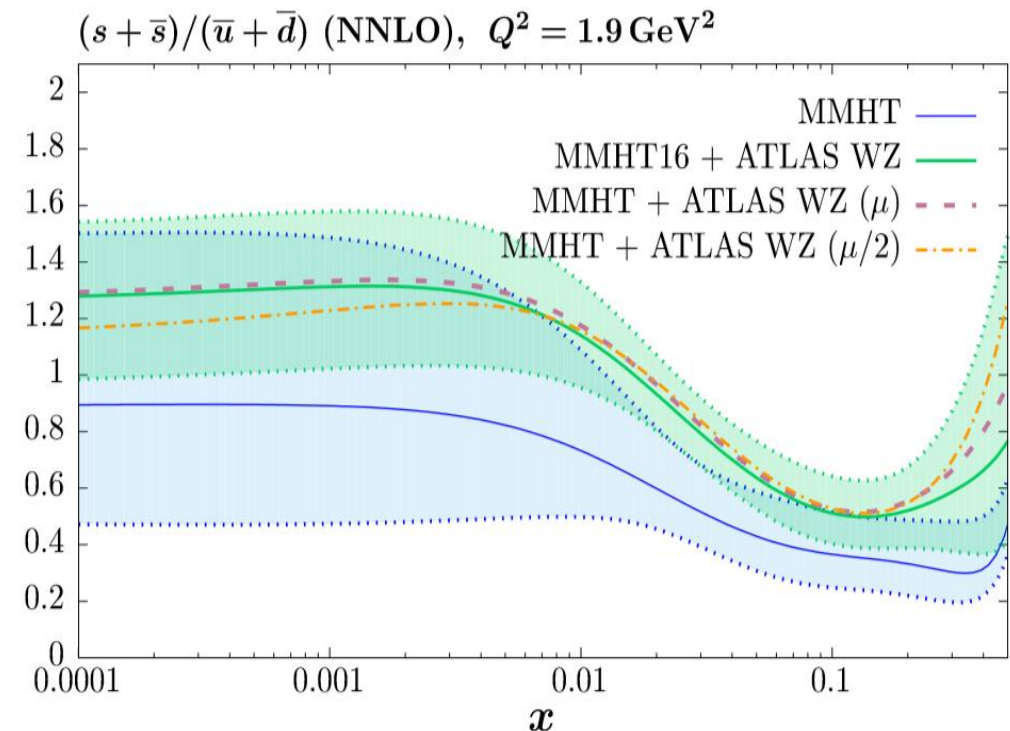


$(s+\bar{s})/(\overline{u}+\overline{d})$ (NNLO), $Q^2 = 1.9\,\mathrm{GeV}^2$

MMHT
MMHT16 + ATLAS WZ
MMHT + ATLAS WZ ($\mu$)
MMHT + ATLAS WZ ($\mu/2$)

[See Thorne, Harland-Lang, Martin 1708.00047 for more details.]

# High precision Atlas data

- Only small effect on strange PDF of adding the full NNLO dimuon corrections with the increase in the strange reduced.

- Without NNLO corrections at top, with NNLO corrections bottom. Green is MMHT14 + HERA, red is MMHT14+HERA+ATLASWZ.
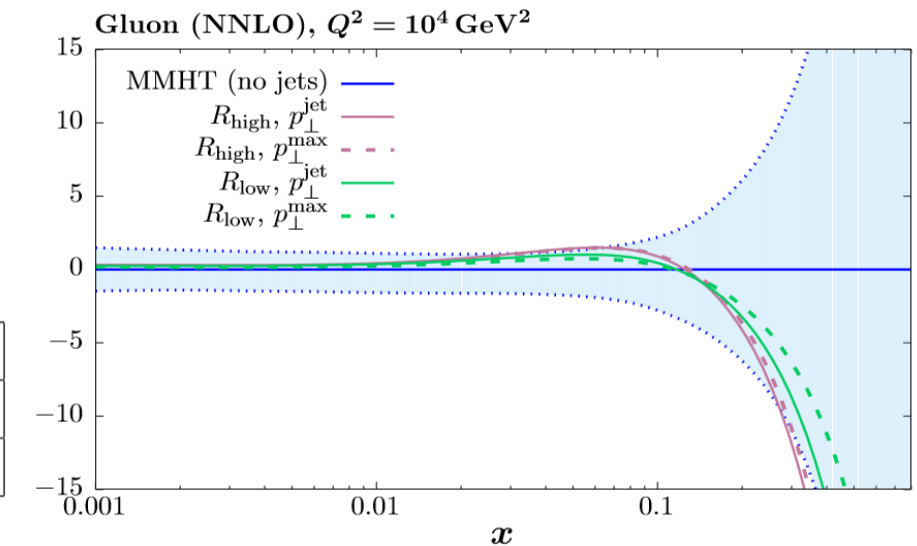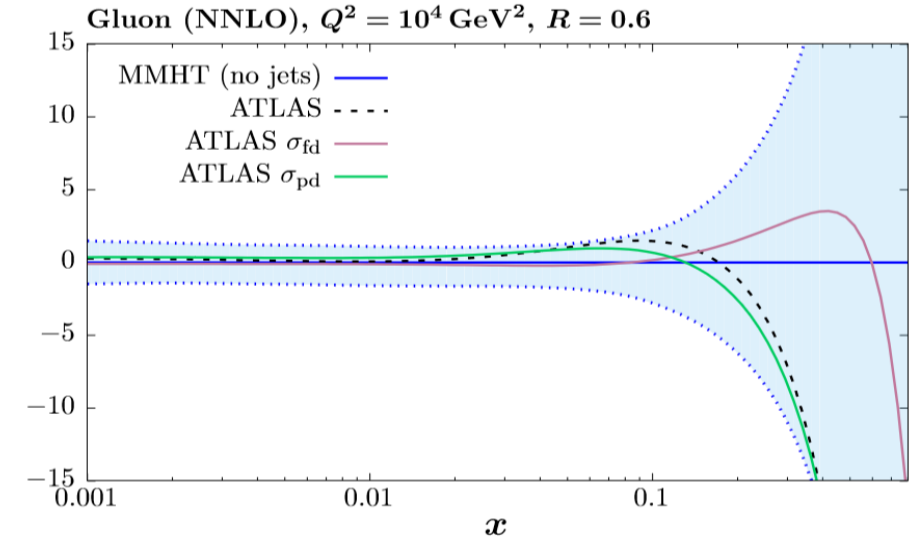
# Atlas 7TeV jets data

- Atlas data dominated by a large number (71) of individual correlated errors.

- If we decorrelate these systematic errors between rapidity bins, whilst leaving them correlated within each bin, a good fit for all rapidity bins can be found.

- Can be more precise – the shift required for each correlated systematic between data and theory is very large for two systematics (21, 62), indicating a clear tension.

- These correspond to a multi-jet balance asymmetry and the jet energy scale close by jets.

- Decorrelating just these between rapidity bins we can see the improvement of the $\chi^2$ per data point is significant and the data/theory is also then visibly better.



$$\frac{1}{6} \sum_{\text{bin } i} \sum_{\text{bin } j} (r_i - r_j)^2$$
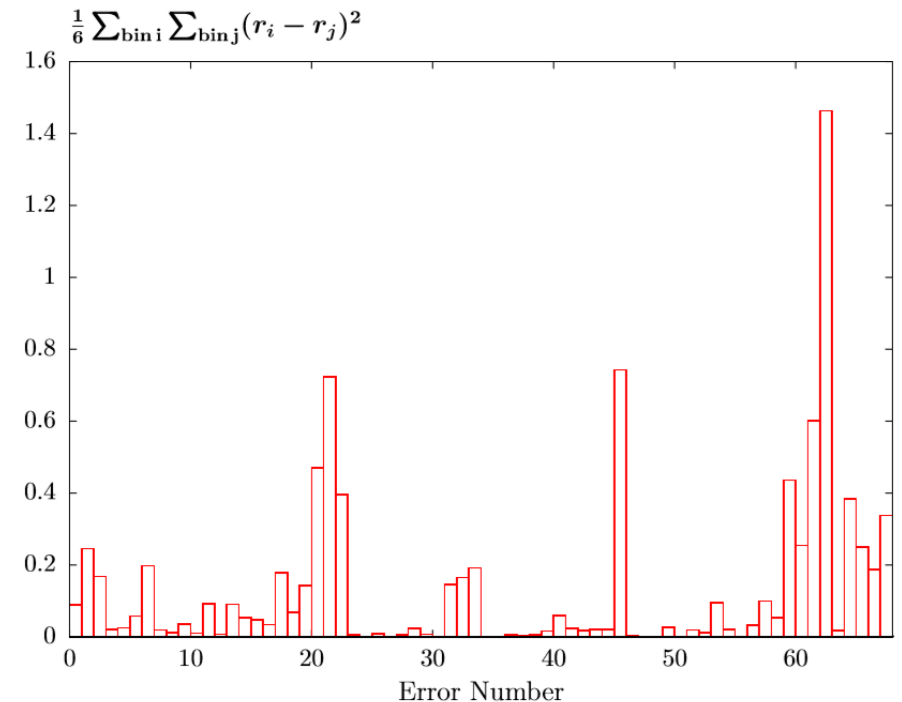
# Atlas(+CMS) 7TeV jets data

- Quantitative effect independent of jet radius (R=0.4 or 0.6) and jet scale $p_T^j$ or $p_T^{max}$ (latter is $p_T$ of leading jet).

- In contrast, CMS 7 TeV data can be fit well without need for decorrelations of errors.

- No tension between Atlas and CMS 7 TeV jet data or relative to other included datasets: $\Delta\chi^2 \sim 5$ global increase upon inclusion.

- Overall PDF effect is to soften the gluon at high $x$, and reduce the uncertainty bands there.

- Effect on gluon pdf stable with respect to R and scale choices.

|           | ATLAS         | ATLAS, $\sigma_{pd}$ | ATLAS, $\sigma_{fd}$ |           | CMS           |
|-----------|---------------|----------------------|----------------------|-----------|---------------|
| $R = 0.4$ | 350.8 (333.7) | 183.1 (170.7)        | 128.4 (122.2)        | $R = 0.5$ | 191.7 (163.4) |
| $R = 0.6$ | 304.0 (264.0) | 178.8 (148.9)        | 128.9 (115.7)        | $R = 0.7$ | 200.1 (175.2) |



Gluon (NNLO), $Q^2 = 10^4\,\mathrm{GeV}^2$, $R = 0.6$

- MMHT (no jets)
- ATLAS
- ATLAS $\sigma_{\mathrm{fd}}$
- ATLAS $\sigma_{\mathrm{pd}}$



Gluon (NNLO), $Q^2 = 10^4\,\mathrm{GeV}^2$

- MMHT (no jets)
- $R_{\mathrm{high}}, p_\perp^{\mathrm{jet}}$
- $R_{\mathrm{high}}, p_\perp^{\mathrm{max}}$
- $R_{\mathrm{low}}, p_\perp^{\mathrm{jet}}$
- $R_{\mathrm{low}}, p_\perp^{\mathrm{max}}$

# $\sigma_{t\bar{t}}$ data added

- NNLO K-factors calculated using top++.

- Fit quality is very good for initial data added.

- Drives small changes in top pole mass at NLO and NNLO and in $\alpha_s$ at NNLO:

| | MMHT14 | MMHT14 + $\sigma_{t\bar{t}}$ data |
|---|---|---|
| $m_t^{pole,\,NLO}$ | 171.7 GeV | 170.2 GeV |
| $m_t^{pole,\,NNLO}$ | 174.2 GeV | 173.4 GeV |
| $\alpha_s(M_Z^2)^{NLO}$ | 0.120 | 0.120 |
| $\alpha_s(M_Z^2)^{NNLO}$ | 0.117 | 0.118 |

- However ATLAS fully differential data at 8TeV and 13TeV suffers poor descriptions.
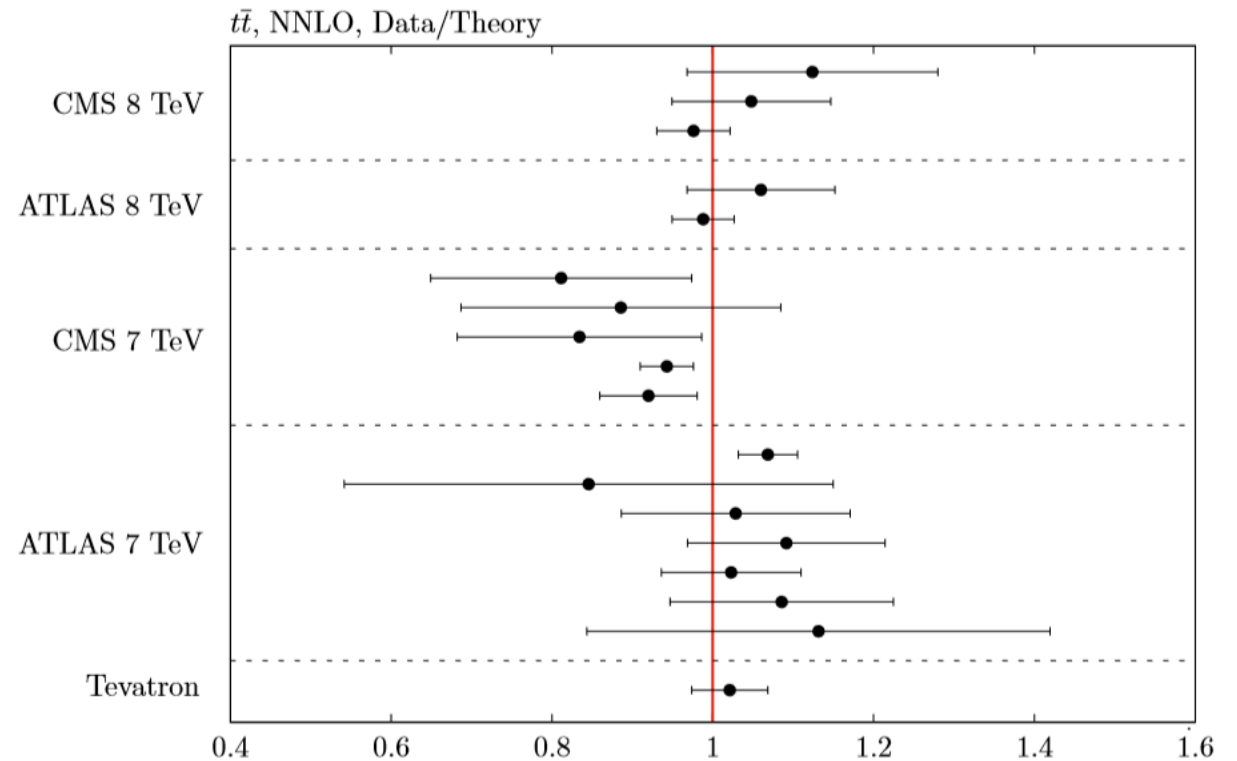
[See Thorne, Harland-Lang, Martin 1708.00047 for more details.]

# Effect of Atlas jet data on gluon

- Main effect of ATLAS jet data is it pulls the gluon lower at large $x$

- Effect of decorrelation is small on gluon PDF.
- Results independent of form of partial decorrelation.
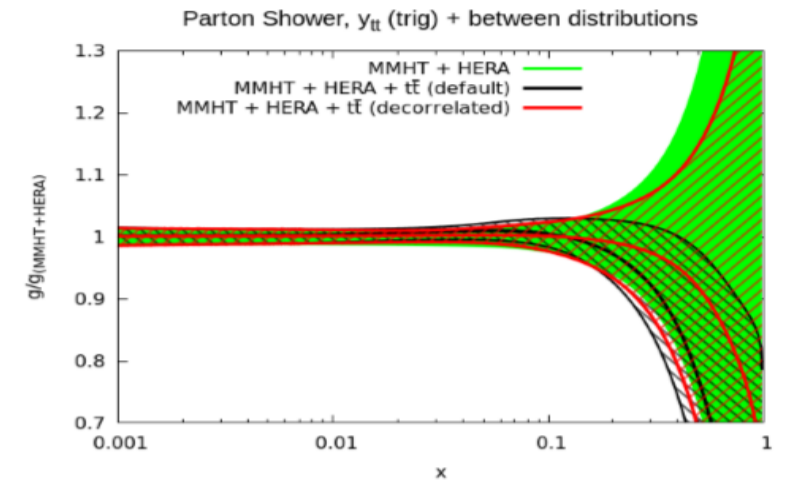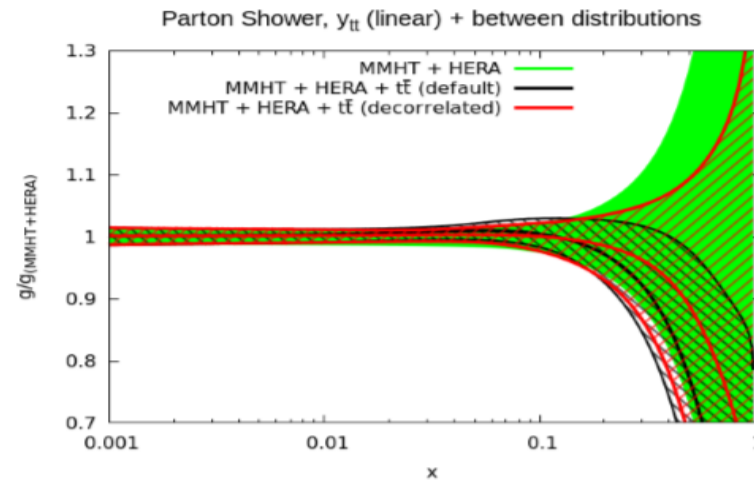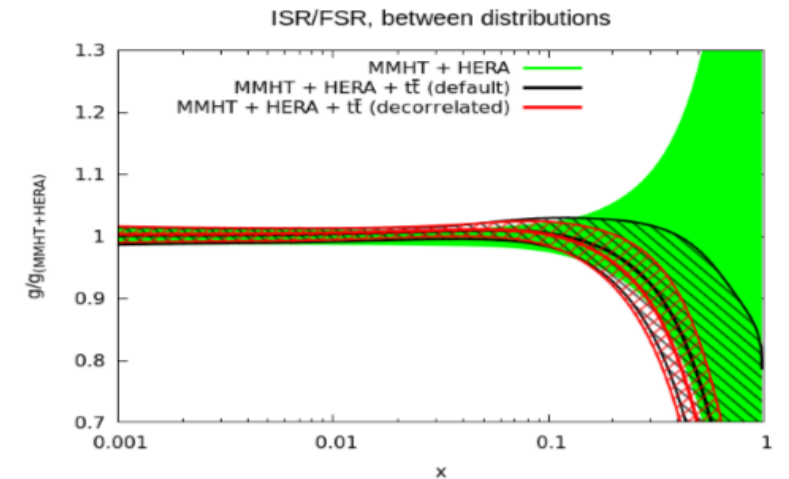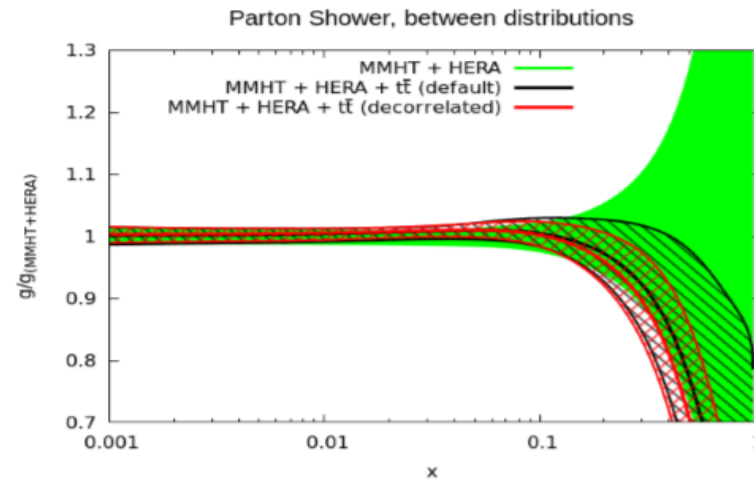
# Effect of Atlas jet data on gluon

- Shown for extended gluon parametrisation (albeit $\bar{d} - \bar{u}$ parametrised here with 6 Chebyshevs rather than $\bar{d}/\bar{u}$ with 6 Chebyshevs).

- Pulls gluon lower at low $x$ and higher at moderate $x$, whilst for all but full decorrelation it is lower at very high $x$.



g (NNLO) percentage change from MMHT14 at $Q^2 = 100\text{GeV}^2$

MMHT14
new prelim no Atlas jets
new prelim full decor Atlas jets
new prelim partial decor Atlas jets
new prelim no decor Atlas jets
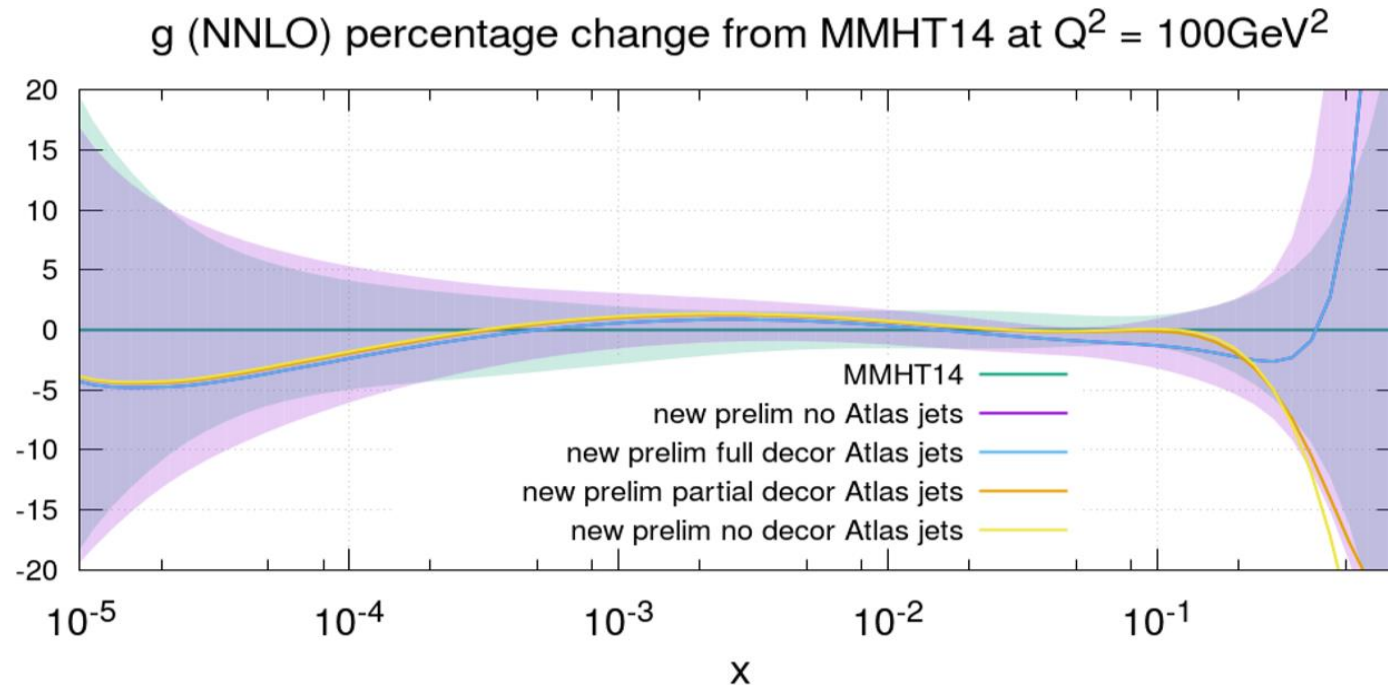
# D0 electron asymmetry

- Final D0 electron asymmetry data from Tevatron (released 2015), $p\bar{p} \rightarrow W + X \rightarrow e\upsilon + X$. [1412.2862]

- Precise lepton charge asymmetry data, measured as a function of $\eta^e$ in 5 kinematics bins in electron transverse energy and missing transverse energy.

- Good agreement between prediction (MMHT14) and refit, little change in pdf central values.

- Slightly undershoots data at high $\eta^e$, implies slightly smaller down quark, but this is not favoured by other data.

- Possibility of greater pdf constraints by using the same data as a $W$ asymmetry [1312.2895].
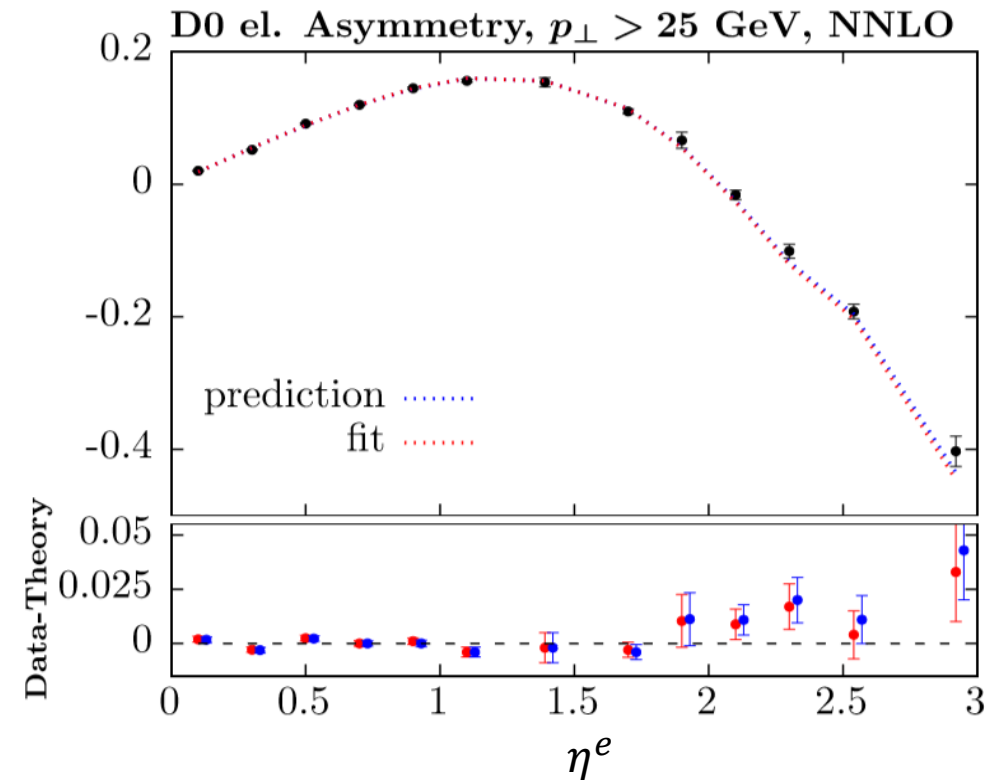
# D0 $W$ asymmetry

- In $p\bar{p}$ collisions $W^+$ is dominantly produced by $u\bar{d}$ and $W^-$ is dominantly produced by $d\bar{u}$.

- Given $u$ PDF $> d$ PDF this means $W^{+/-}$ is produced favourably in the proton/antiproton direction (+ve/-ve rapidity).

- However when observing lepton charge asymmetry this $W$ production asymmetry is convoluted with the V-A structure of the $We\upsilon$ vertex, which ensures the positron/electron is emitted in direction opposite to the $W^+/W^-$, therefore at large rapidities this overwhelms the $W$ asymmetry causing a turnover in the distribution.

- Leptons at a specific rapidity originate from a wide range of $W$ rapidities, and therefore from a wide range of parton x values, diluting the impact of these asymmetries when determining PDFs

- Mapping back to $W^+/W^-$ requires PDFs to determine $W$ asymmetry.

- However PDF errors are small whilst statistical errors are large (particularly at high rapidity).

- Therefore addition of PDF errors compensated for by reduced statistical errors from not washing out the asymmetry.

- Therefore you can get more precise constraints by viewing as a $W$ asymmetry.
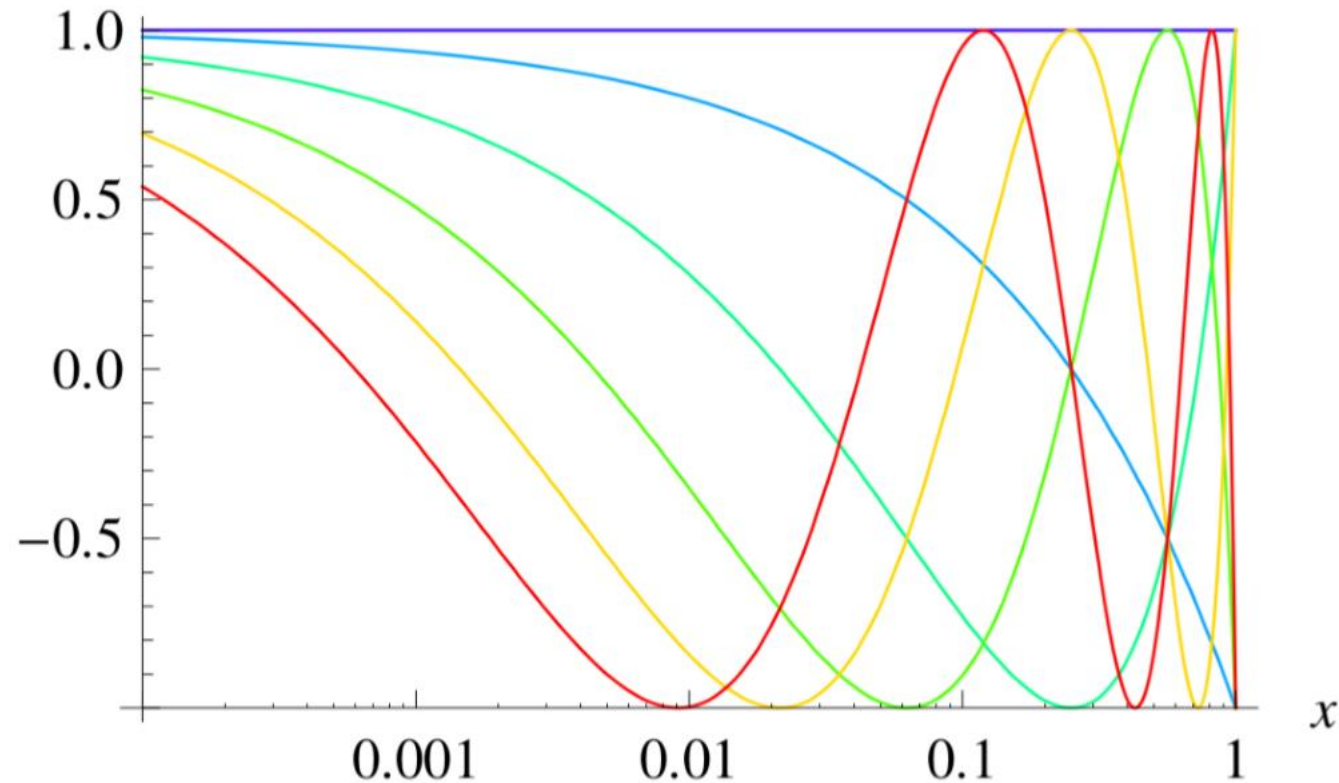
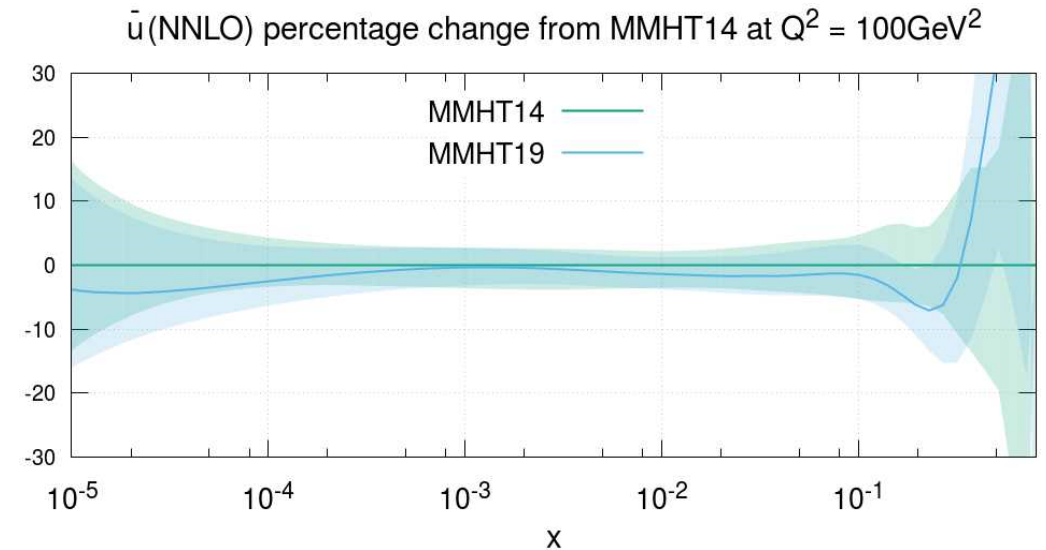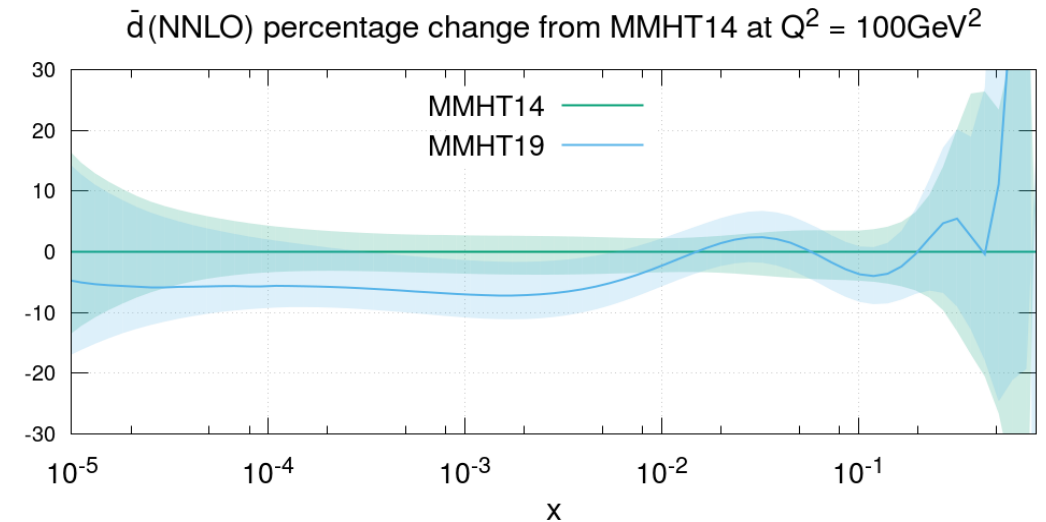[See 1412.2862 by D0 collaboration and for more details.]
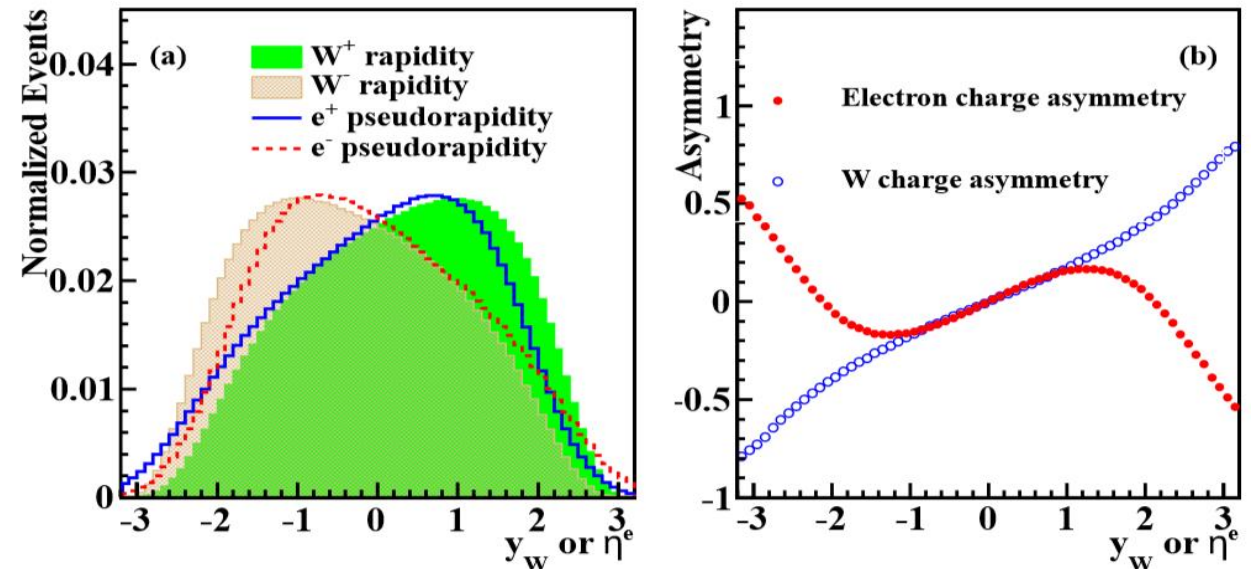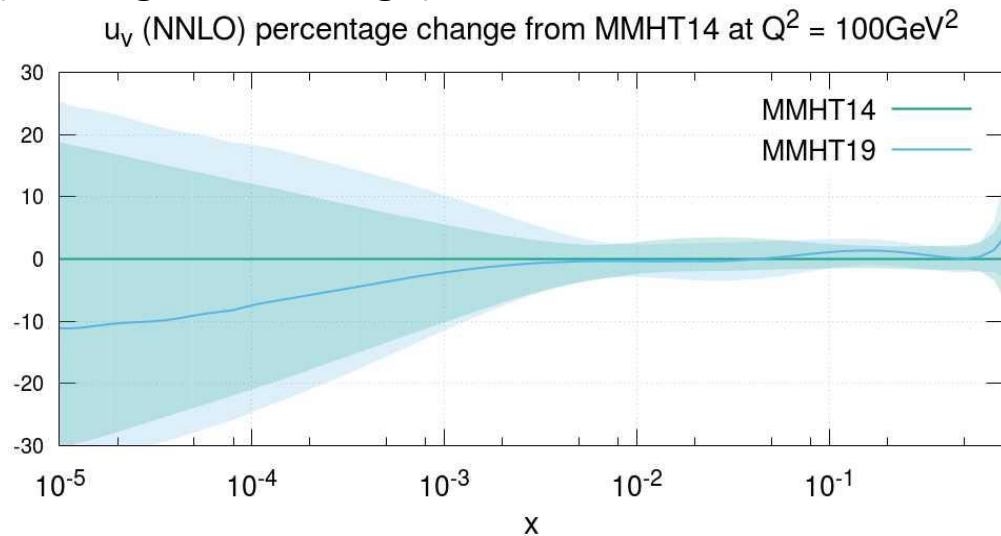
# Chebyshev polynomials

$$T_i(y(x) = 1 - 2\sqrt{x})$$

- All tend to 1 as $x \to 0$, odd polynomials then tend to -1 as $x \to 1$ whilst even polynomials tend to 1 as $x \to 1$.

# Extending PDF parametrisation – $\bar{d}, \bar{u}, u_v$

- $\bar{d}$ reduced at low-moderate $x$, occurs as $\bar{s}$ driven higher by Atlas W,Z 7 TeV data and charge-weighted data ensures total must be fixed so $\bar{d}$ reduces.

- Same effect in $\bar{u}$, but reduced as charge-weighted.

- $\bar{d}$ and $\bar{u}$ harder at high $x$ due to harder gluon at high $x$.

- $u_V$ shows little changes, slightly reduced at low $x$, again increased low $x$ uncertainties with extended parametrisation. Small bump at $x{\sim}0.2$, seen in HERAPDF (although not as large).



$\bar{d}$(NNLO) percentage change from MMHT14 at $Q^2 = 100 GeV^2$



$u_V$ (NNLO) percentage change from MMHT14 at $Q^2 = 100 GeV^2$



$\bar{u}$(NNLO) percentage change from MMHT14 at $Q^2 = 100 GeV^2$

# Extending PDF parametrisation - $g$

- Gluon nearer input is lower at low-moderate $x$ but eventually higher at very low $x$, shows behaviour similar to arXiv:1902.11125, where there is a $\log(x)$ type parametrisation for gluon at small x.

- Gluon at high $Q^2$ generally higher at moderate $x$ and smaller at low $x$, due to inclusion of further LHC jet data.

- Increased uncertainty at very high $x$ clearly visible as a result of enhanced parametrisation.

- Also gluon is harder at very high $x$ due to enhanced parametrisation and also some effect of CMS jet data.

- Atlas jet data not included in these plots.

g (NNLO) at $Q^2$ = 3GeV$^2$

g (NNLO) percentage change from MMHT14 at $Q^2$ = 100GeV$^2$
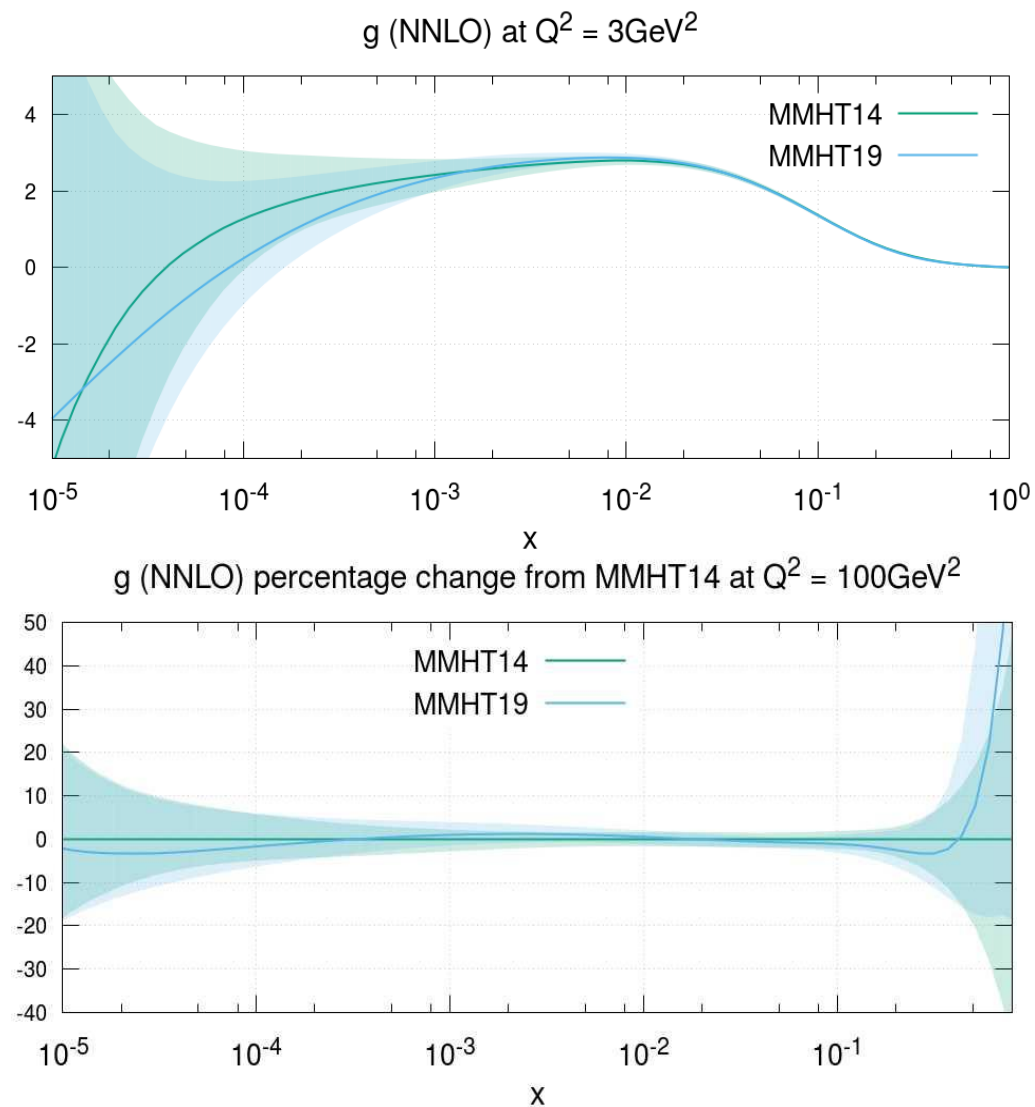
# MMHT2015QED – Global Fit and Effect on $\alpha_s$

- Global fit increase in $\chi^2$ due to QED evolution and corrections to structure functions but reduces after refitting:

| Change in $\chi^2$ due to QED evolution compared to MMHT14+HERA I+II | | | |
|---|---|---|---|
| NLO before fit | NLO after fit | NNLO before fit | NNLO after fit |
| 4180 (+41) | 4151(+12) | 3574 (+42) | 3539 (+7) |

Numbers in brackets show the change due to inclusion of QED.

- Allowing $\alpha_s$ to be free has no significant effect on the quality of the global fit ($\Delta\chi^2 < 1$).

- Given $q \rightarrow q\gamma$ emission reduces the singlet at high $x$, to obtain the same best fit PDFs at high scales expect QED inclusion to reduce $\alpha_s$ to counteract this additional emission.

- MRST QED set found that $\alpha_s(M_Z^2)$ remained the same as in pure QCD, because NMC and HERA data preferred a larger value for the gluon at small $x$ and so would increase $\alpha_s$, pulling $\alpha_s(M_Z^2)$ back up and cancelling the expected reduction in $\alpha_s$.

- MMHT2015QED sees:
  - NLO: no change in $\alpha_s(M_Z^2)$ at NLO upon inclusion of QED effects.
  - NNLO: $\alpha_s(M_Z^2)$ reduces from 0.1181 to 0.1180 upon inclusion of QED effects, therefore a small reduction.
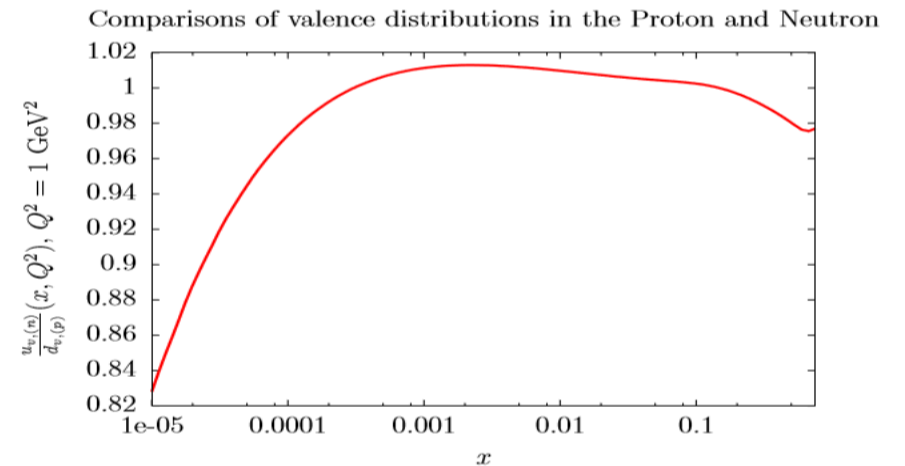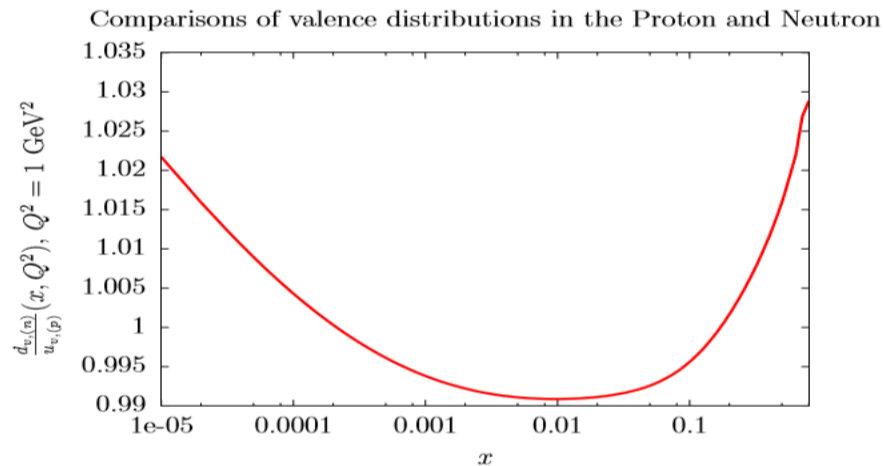
[See Harland-Lang, Martin, Nathvani, Thorne 1907.02750 for more details.]

# MMHT2015QED – Photon PDF + QED DGLAP

- QED-corrected neutron PDFs also provided, necessary for consistent fit to deuteron and nuclear fixed target data from $vN$ DIS -> Potential impacts for nuclear PDFs here.

- QED isospin violation effects reduce the NuTeV $sin^2\theta_W$ anomaly.

- With **isospin violation** one obtains an addition term, also with **strange sea momentum asymmetry** there is a further term, these can ensure the anomaly is in agreement with SM within errors:

$$R^- \approx \frac{1}{2} - \sin^2\theta_W + \left(1 - \frac{7}{3}\sin^2\theta_W\right)\frac{\int_0^1 \mathrm{d}x\; x\left[(u_v^p - d_v^n) - (d_v^p - u_v^n)\right]}{2\int_0^1 \mathrm{d}x\; x(u_v + d_v)} - \left(1 - \frac{7}{3}\sin^2\theta_W\right)\frac{\int_0^1 \mathrm{d}x\; x(s - \bar{s})}{\int_0^1 \mathrm{d}x\; x(u_v + d_v)}.$$

- For the valence distributions the magnitude of isospin violation is a few percent.

- Significant at low and high x:



Comparisons of valence distributions in the Proton and Neutron

[See Harland-Lang, Martin, Nathvani, Thorne 1907.02750 for more details.]

# MMHT2015QED – Photon PDF + QED DGLAP

- Momentum carried by partons in proton as function of $Q^2$, shows effect of QCD and QED evolution.
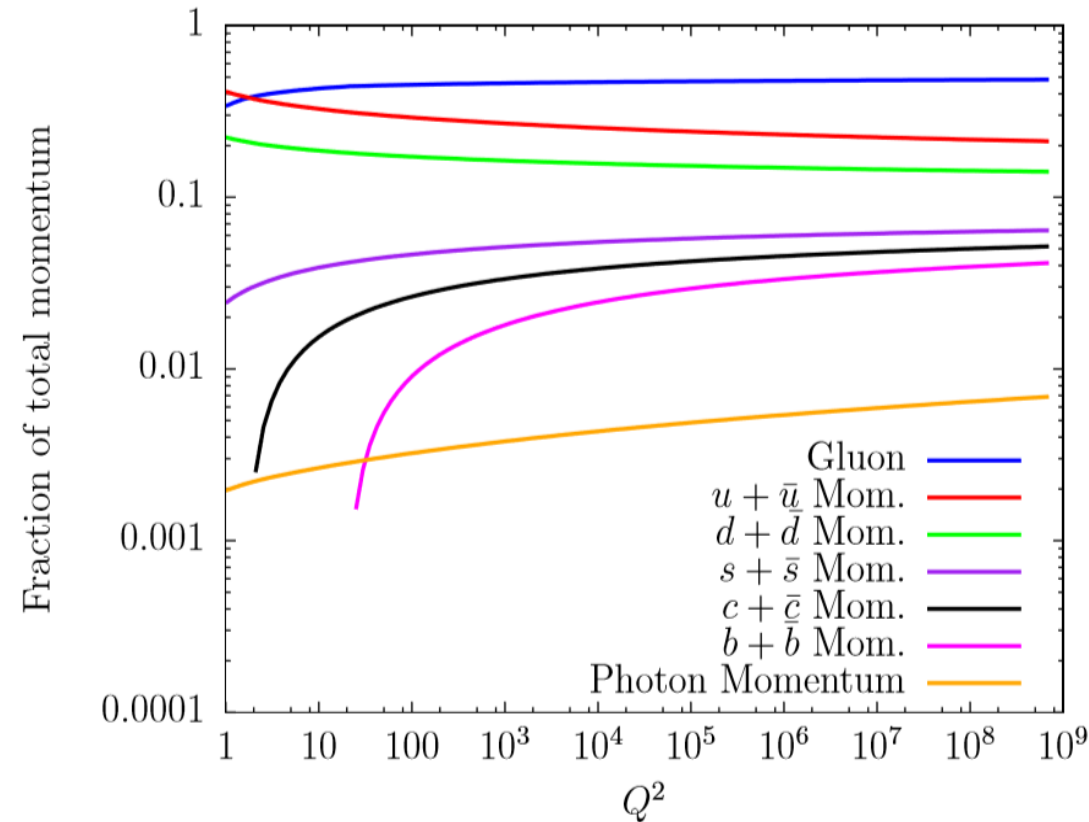


[See Harland-Lang, Martin, Nathvani, Thorne 1907.02750 for more details.]

# MMHT2015QED – Effect on High mass DY

- See effects of inclusion of our photon PDF set on double differential $(m_{ll}, \eta)$ lepton pair production in high mass Drell-Yan data from the LHC, $116 GeV < m_{ll} < 1500 GeV$ to avoid Z contribution.

- Expect effects of QED for this process to be non-negligible, in particular the photon-initiated (PI) contributions.
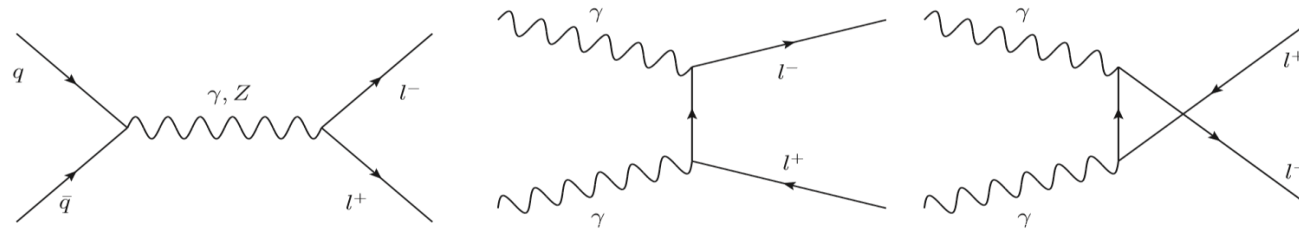


$\mathcal{O}(\alpha)$ PI contributions

- Most sensitive to $\gamma^{(inel)}(x, Q^2)$ due to high scale.

- QED DGLAP causes decrease in dominant $q\bar{q}$ contribution due to reduced quark and antiquark PDFs at high $x$.

- QED PI contributions increase cross-section across all bins as new channel for lepton pair production opens up.

- As $\gamma(x, Q^2)$ is largest at low $x$, particularly at high scales $Q^2 > 10^4$ GeV², we obtain enhanced $\sigma$ at low and intermediate rapidity bins for all mass bins. The effect at high $\eta$ is outweighed by the DGLAP reduction.

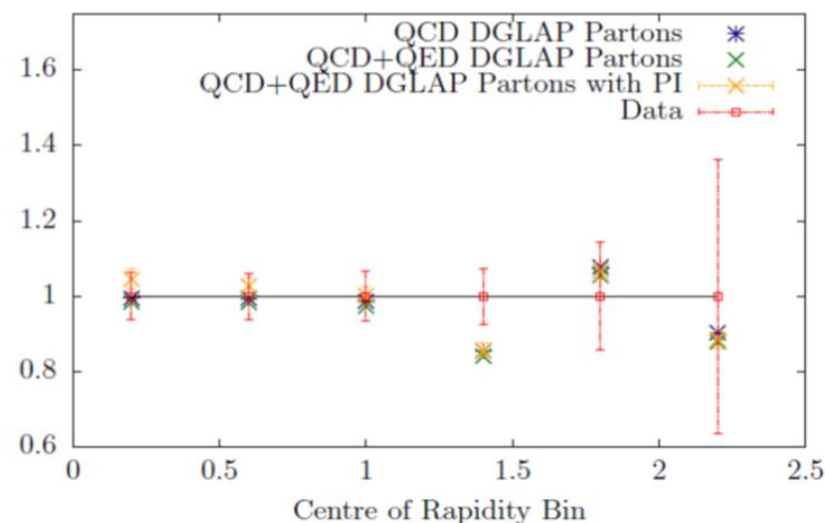- Therefore need to include both photon PDF and QED DGLAP for correct predictions.

[See Harland-Lang, Martin, Nathvani, Thorne 1907.02750 for more details.]

# MMHT2015QED – Photon PDF + QED DGLAP
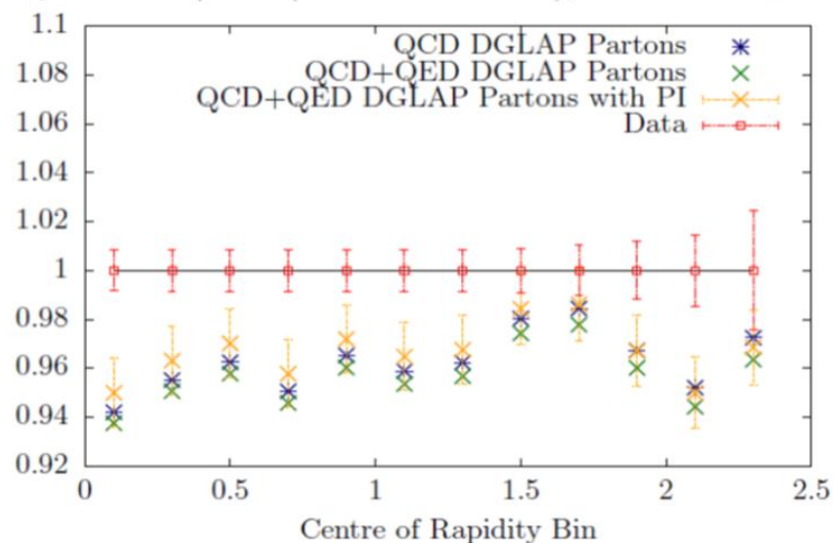


- QED DGLAP reduces cross-section in all $\eta$ and $m_{ll}$ bins, whilst QED PI processes increase it, particularly at low $\eta$.

- This has been for no global refit upon adding this data, refitting causes no further improvement to the fit $\chi^2/N_{points} \sim 65/48$, -> this data offers no meaningful constraint on the photon.

- Nonetheless obtain very mildly reduced PDF uncertainties upon refitting, shown to suggest mild preference for QED DGLAP evolution -> may see stronger hints with more accurate data.

[See Harland-Lang, Martin, Nathvani, Thorne 1907.02750 for more details.]