PDFs for the LHC
A M Cooper-Sarkar
DIS 2011

PDF4LHC comparisons 2010 updated to 2011

Updates: CT10, NNPDF2.1, HERAPDF1.5, ABM11
Treatment of heavy quarks: NNPDF2.1, HERAPDF, MSTW, ABM
Value of $\alpha_s(M_Z)$: NNPDF, HERAPDF, MSTW, CT/CTEQ
Does DIS data like low $\alpha_s(M_Z)$? --- HERAPDF1.6+jets

The year of NNLO: CTEQ, NNPDF, HERAPDF join MSTW, JR, ABKM

The Tevatron Higgs limits: is it essential to use PDFs which are fitted to Tevatron jets?
Is it essential to use a global fit? Look at NNLO predictions NOT at NLO
Global fits:
Doubts about deuterium corrections, and doubts about fixed target F2

Early LHC W/Z data – some visible impact on PDFs– (NNPDF, HERAPDF)
Early LHC jet data
In 2010 the PDF4LHC group considered:
- MSTW08
- CTEQ66
- HERAPDF1.0
- NNPDF2.0
- ABKM09
- GJR08

Overall disagreement ~8% in W, Z cross-sections

The PDF4LHC recommendation was to take the envelope of the NNPDF, MSTW, CTEQ predictions --even this may not be enough!
Why these disagreements?

**Firstly** groups use different values of $\alpha_S(M_Z)$, the effect of this can be seen on the figures. A common value would bring some of the predictions into better agreement.

HERAPDF, NNPDF, CTEQ and MSTW provide PDFs at a series of $\alpha_S(M_Z)$ values.

**Secondly** groups have different ways of accounting for heavy quark production. And use different values of the heavy quark mass. Within any chosen scheme a change of quark mass from 1.4 to 1.65 GeV can change the W/Z cross-sections by ~2.5%.

HERAPDF, NNPDF, MSTW provide PDFs at different mc values.

**Thirdly**, different groups use different input data sets, e.g. the data used by CTEQ and MSTW are very similar and they do NOT include the latest most accurate HERA data which are used in HERAPDF1.0. Not only are these new (2009) data more accurate they also have a different normalisation. This accounts for ~2.5% upward shift of the HERAPDF prediction.

CT, NNPDF, ABM update to include these data.

Fourthly there are some differences in philosophy regarding choices of PDF parametrisation and theoretical/model prejudices which are imposed.

Let's look more closely at some of these points...
Heavy quark production.: there are two extremes-
Use only 3 massless parton flavours and calculate exact ME’s for heavy quark production (FFN method)- wrong at high scale since \( \ln(Q^2/m_c^2) \) terms not resummed
Consider all partons as massless except that charm and beauty turn on abruptly at their kinematic thresholds (ZMVFN) – WRONG at low scale near these thresholds
A GMVFN (General-mass Variable Flavour Number Scheme) is supposed to give us the best of both worlds..
BUT there are different ways to do this...ACOT, Thorne, FO-NLL and there are tunable scale choices

NNLO differences are not so large...

And a second question is
What value of the charm mass should be used?
H1 and ZEUS have also combined charm data recently.

And these data show a preference for a charm mass

$$m_c = 1.57 \pm 0.02 \text{ GeV}$$

If the standard Thorne VFN is used.
But the value of the preferred charm mass depends on the heavy quark scheme. Each scheme can be used to predict the W and Z cross-sections at the LHC as a function of charm mass parameter.

If a fixed value of $m_c$ is used then the spread is considerable (~7%) - but if each prediction is taken at its own optimal mass value the spread is dramatically reduced (~2%) even when a Zero-Mass (ZMVFN) approximation has been used.

The PDFs MSTW08, CTEQ6.6, NNPDF2.0 do NOT use charm mass parameters at their optimal values - and explains part of their differences.

Now MSTW, NNPDF and HERAPDF provide a series of $m_c, m_b$ values. But what is the mass that is being used? The pole mass?

The running mass has been measured independently would it be better to use this? ABM have considered this in the FFN scheme.
The schemes considered above all use the pole mass with VFN fits. AB(K)M have always used FFN. ABM now use the running mass.

The NNLO(approx.) FFNS ABM predictions based on the running mass definition are in nice agreement with the new HERA data.

\[ m_c(m_c) = 1.27 \pm 0.08 \text{ GeV (PDG '10)} \]

\[ m_c(m_c) = 1.18 \pm 0.06 \text{ GeV (incl. } F_2 + \text{PDG)} \]

The HERA data prefer \( m_c(m_c) \) close to the PDG value.

\[ \text{(courtesy of K. Lipka)} \]
NNPDF2.0 has been updated to NNPDF2.1 using FONLL VFN
CTEQ6.6 to CT10, ABKM09 to ABM11, HERAPDF1.0 to HERAPDF1.5

The use of the VFN scheme puts NNPDF2.1 closer to MSTW,
CT10 and CTEQ6.6 are very similar, HERAPDF1.5 is a little higher than 1.0 for W⁺,Z
CMS and ATLAS data agree well with all predictions.
Are ratios better predicted than the cross-sections?

- The W/Z ratio is YES - ~1% spread – this is obvious when you think about the quark flavours which enter
- The W+/W- ratio is still NOT ~>5% and this shows up more strongly in the W and lepton asymmetries - this is a u-valence – d-valence difference in a previously unmeasured region of x

Is there an ATLAS/CMS discrepancy?
W/Z production depends on $q$-$\bar{q}$ luminosity

Update with NNPDF2.1, CT10 and with HERAPDF1.5
Features of the updates HERAPDF1.5: update of data AND fit

Uses preliminary HERA I+II data combination

Gives increased precision at high-\(x\)
Features of the updates: NNPDF2.1

Effect of adding HERA-I combined data

Features of the updates: CT10

Features of the updates: ABM11

Compare NNPDF2.1 and CT10 to MSTW08
Considering the value of $\alpha_s(M_Z)$ – very important for Higgs

The table shows the values of $\alpha_s(M_Z)$ used by the PDF groups at NLO and at NNLO plus the values from more recent studies.

HERAPDF, NNPDF, CTEQ and MSTW all provide PDFs using a series of $\alpha_s(M_Z)$ values which can be used to evaluate $\alpha_s(M_Z)$ uncertainty.

It is established that PDF and $\alpha_s(M_Z)$ uncertainty may be combined in quadrature.

<table>
<thead>
<tr>
<th></th>
<th>ABKM</th>
<th>CTEQ</th>
<th>GJR</th>
<th>HERA</th>
<th>NNPDF</th>
<th>MSTW</th>
</tr>
</thead>
<tbody>
<tr>
<td>NLO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_s(M_Z)$ in PDF</td>
<td>0.118</td>
<td>0.118</td>
<td>0.1135</td>
<td>0.1176</td>
<td>0.119</td>
<td>0.120</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NNLO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha_s(M_Z)$ 2011</td>
<td></td>
<td></td>
<td>0.1202</td>
<td>0.1191</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\pm 0.0019$</td>
<td>$\pm 0.0006$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|      |        |        |        |        |        |        |
|      |        |        |        |        |        |        |
|      |        |        |        |        |        |        |
|      |        |        |        |        |        |        |
|      |        |        |        |        |        |        |
|      |        |        |        |        |        |        |
|      |        |        |        |        |        |        |
There is an idea that DIS data prefer lower values of $\alpha_s(M_Z)$.

- MSTW challenge this – this is basically only BCDMS, HERA data prefer higher values.
- NNPDF say DIS only value is 0.1177 rather than 0.1191.
- NNPDF study shows that the variation of $\chi^2$ for HERA data is very flat.

Now we have heard from HERA themselves.

They have used their preliminary HERA I+II data combination and they agree that the $\chi^2$ dependence on is very flat – and the model dependence is large.

BUT if the DIS jet data are added there is a different story AND alphas is larger $\alpha_s(M_Z) = 0.1202 \pm 0.0019$ (excluding scale error)

\[ \chi^2 - \chi^2_{\text{min}} \]

\[ \alpha_s \text{ scan} \]

--- HERAPDF1.5f

--- HERAPDF1.6

\[ \alpha_s(M_Z) \]

\[ 0.114 \quad 0.116 \quad 0.118 \quad 0.12 \quad 0.122 \quad 0.124 \quad 0.126 \]
HERAPDF has updated from their standard parametrisation HERAPDF1.5 to a more flexible parametrisation HERAPDF1.5f and then added jets HERAPDF1.6.

\[ \alpha_s(M_Z) = 0.1202 \pm 0.0019 \]
Gluon-gluon luminosity is important for Higgs production

CT10 and NNPDF2.1 are both closer to MSTW than CTEQ6.6 and NNPDF2.0

Note how the larger value of $\alpha_s(M_Z)$ brings better agreement at low scale
Large spread in cross-sections $> 15\%$. For Higgs there is MSTW/NNPDF(2.0 or 2.1) Vs CTEQ/CT10/HERA/ABKM.
Strong dependence on $\alpha_s(M_Z)$ so use of a common value would help.

But for Higgs we had better consider NNLO
Until this year there were fewer NNLO PDFs: MSTW, ABKM, JR and HERAPDF1.0 (for two values of $\alpha_s(M_Z)$ without error band)

arXiv:1011.6259 from ABM and JR gives a comprehensive comparison
NNLO predictions for t-tbar and Higgs

The PDF4LHC recommendation at NNLO is to use the MSTW08 result but increase the errors by ~2 (to reflect the NLO envelope).

However the other NNLO PDF predictions are consistently below those of MSTW- partly due to $\alpha_s(M^2_Z)$
This has been used to challenge the Tevatron Higgs exclusion region

Why was MSTW08 favoured ?

• Because it is a global fit. Are global fits actually the best?- see criticisms of the use of deuterium target data and NMC F2 ..next slides.

• Because it fits Tevatron jet data well- determining high-x gluon BUT…AB(K)M and HERAPDF also describe Tevatron jet data

However, the MAIN development of 2011 is more NNLO PDFs
Criticisms of the use of fixed target data

NMC data are usually used as $F_2$. To get this some correction must be made for FL but the corrections made were based on old theory and old prejudices.

ABM find considerable differences in their fits if they use $F_2$ rather than $\sigma$.

But both NNPDF and MSTW say this effect is negligible. However they work with fixed $\alpha_s(M_Z)$. ABM have $\alpha_s(M_Z)$ free in the fit.

<table>
<thead>
<tr>
<th>$\alpha_s(M_Z)$</th>
<th>$\alpha_s(M_Z)$ with $\sigma_{\text{NMC}}$</th>
<th>$\alpha_s(M_Z)$ with $F_2^{\text{NMC}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NLO</td>
<td>0.1179(16)</td>
<td>0.1195(17)</td>
</tr>
<tr>
<td>NNLO</td>
<td>0.1135(14)</td>
<td>0.1170(15)</td>
</tr>
<tr>
<td>NNLO + $F_L$ at $O(\alpha_s^3)$</td>
<td>0.1122(14)</td>
<td>0.1171(14)</td>
</tr>
</tbody>
</table>

NMC may have an impact on $\alpha_s(M_Z)$ determinations from PDF fits. Only affects LHC Higgs XS if PDG $\alpha_s(M_Z) = 0.1184 \pm 0.0007$ value not trusted.
Accardi et al. have been considering corrections for the deuterium targets. They conclude that the uncertainties are larger than is usually accounted for and of course this uncertainty affects the $d$-quark.

They also conclude that this can affect the determination of the high-$x$ gluon when Tevatron jet data are fitted at the same time as deuterium fixed target data.

High $p_T$ jet production receives contributions from $q\bar{q}, qg$ and $gg$. Since the $u$ quark is well known the variation on the $d$ from deuterium corrections is compensated by an anti-correlated shift in the gluon density.

The high-$x$ uncertainty on the $d/u$ ratio due to deuteron wave function and off shell effects as well as PDF uncertainty.
How well are Tevatron jet data described by non-global fits?

HERAPDF1.5 $\chi^2/dp = 176/76$ for CDF and 245/110 for D0 for the central PDF

However this ignores the error band of the fit. If these data are included in an NLO fit we get $\chi^2/dp = 113/76$ and 157/110 resp.

The resulting PDF is at the edge of HERAPDF1.5 (68%CL) error bands

But the real question is how well does an NNLO fit describe these data?

For HERAPDF1.5 NNLO the description is MUCH better $\chi^2/dp=72/76$ for CDF even for the central PDF.
And how well can HERAPDF describe Tevatron W/Z data?

\[
\chi^2 = \frac{27}{28} \quad \chi^2 = \frac{16}{28} \quad \chi^2 = \frac{21}{13} \quad \chi^2 = \frac{25}{11}
\]

**Before fitting Tevatron W/Z data**

HERAPDF1.5 central PDF gives a good description of Tevatron W/Z data even before fitting.

**After fitting Tevatron W/Z data**

Fitting results in a PDF which is within the error bands.

However PDF uncertainties on the d-valence quark are much reduced.
The year of NNLO: HERAPDF1.5 NNLO

Same data set as for HERAPDF1.5(f) and use of flexible parametrisation as for 1.5f.

The NNLO PDFs differ from NLO in a similar manner as for MSTW08 (same heavy quark treatment).

The low-\(x\) gluon is more uncertain than at NLO - NNLO DGLAP is NOT a better fit to low-\(x, Q^2\) data.

HERAPDF1.5 NNLO has a harder high-\(x\) gluon than 1.0.

Both the g-g and the q-qbar NNLO luminosities of HERAPDF1.5 are closer to MSTW than HERAPDF1.0 hence Higgs predictions also closer.

PDFs are available in an \(\alpha_s(M_Z)\) series, but the standard value is 0.1176.
NNPDF use the FONLL scheme to treat heavy quarks in VFN, HERAPDF use Thorne’s scheme, ABM and JR use FFN formulations—there are differences between these schemes.

NNLO CT is also nearly ready—it will use SACOT.
And now we have LHC W/Z data - we have already seen the total cross-sections and ratios. Let us look at rapidity distributions both for lepton asymmetry and Z\(\nu\) production.

These data are well described by almost all the PDFs, but what improvement if any can they bring if included in the fits?
Improvement in HERAPDFs from adding CMS W-asymmetry by re-fitting

HERAPDF uncertainties after including Tevatron W/Z data

Improvement in NNPDFs from adding CMS W-asymmetry by re-weighting

HERAPDF uncertainties after including Tevatron W/Z data + CMS W-asymmetry

Improvement in u and d-valence at low-x

Improvement in u and d at low-x
And how well is LHC jet data described?

Jet data will also soon be discriminating for PDFs.

The PDFs that fit the Tevatron jets best are not necessarily those that fit the LHC jets best. The mixture of q-q, q-g, g-g induced jets is different.

HERAPDF1.5 is doing the best job at LHC.
PDF4LHC comparisons 2010 updated to 2011
Updates: CT10, NNPDF2.1, HERAPDF1.5(6) all NLO and NNPDF2.5, HERAPDF1.5 ABM11 at NNLO

Treatment of heavy quarks: still some differences in GMVFN schemes
NNPDF2.1, HERAPDF, MSTW provide series of $m_c, m_b$ values
ABM uses FFN and running mass
Value of $\alpha_s(M_Z)$: NNPDF, HERAPDF, MSTW, CT/CTEQ favour larger values but provide PDFs for a series of $\alpha_s(M_Z)$ values
ABM/GJR favour a smaller value
Does DIS data like low $\alpha_s(M_Z)$? NO!
Does including Tevatron jet data in ABM fits raise $\alpha_s(M_Z)$? also NO.

The Higgs controversy: is it essential to fit Tevatron jets to make a meaningful prediction? NO-but in fact both ABM and HERAPDF can describe these data well. (Particularly at NNLO)
Is it essential to use a global fit? NO and there are doubts about fixed target deuterium corrections and doubts about the use of fixed target F2 data.
BUT HERAPDF1.5 NNLO is much closer to MSTW2008 NNLO than 1.0 AND NNPDF2.5NNLO is close to MSTW2008 if the same $\alpha_s(M_Z)$ is used AND ABM11 Higgs cross-section is larger if Tevatron jet data are included: The Tevatron Higgs limits are probably safe.

Early LHC W/Z data – have some visible impact on PDFs– (NNPDF, HERAPDF)
extras
IS NLO DGLAP applicable for the low-x, Q² part of the kinematic plane?

Before combined HERA-1

CTEQ say the do not confirm this tension, but when HERA combine their low energy run data the low x, Q² part of the data is not so well fit and the gluon which results from imposing harder Q² cuts or Q² > 0.5x⁻⁰.³ cut is steeper—this seems NOT to be solved by NNLO

Not much effect at LHC W/Z scale—no sign of unusual behaviour
HERAPDF plus Tevatron W/Z plus LHC W/Z fit χ² for the LHC W/Z data

The new fit PDFs do not move (much)outside HERAPDF1.5 error bands

Only the CMS asymmetry data lead to any substantial further improvement in PDF uncertainties