PROCEEDINGS OF SCIENCE

PDFs and Top Physics

R.S. Thorne*
Department of Physics and Astronomy,
University College London, WC1E 6BT, UK
E-mail: robert.thorne@ucl.ac.uk

I present the results from the recent PDF4LHC study, and the resulting new recommendation for combining PDFs sets for LHC calculations. In order to put this into context I summarise continuing developments in PDFs. This includes improvements and recent updates of particular PDF sets due to theory improvements and a variety of new data sets, including most of the up-to-date LHC data. I will emphasise particular issues relevant for top physics.

8th International Workshop on Top Quark Physics
14-18 September, 2015
Ischia, Italy

*Speaker.

PoS(TOP2015)053

1. Recent PDF Updates - effect and treatment of LHC data

Each group has produced updates including new data, often including data from the LHC. The recent analysis from the ABM group, ABM12 [1], now includes more HERA cross-section data, and vector boson production data from ATLAS, CMS and LHCb. The PDF sets are determined together with α_S, whose value comes out to be $\alpha_S(m_Z^2) = 0.1132$ at NNLO. Top quark pair production data from the LHC is investigated, but not included in the default PDFs. Its inclusion tend to raise the high-x gluon and $\alpha_S(m_Z^2)$ a little, the precise details depending on the top quark mass (and mass renormalization scheme) used. ABM PDF sets currently give the best fit to the ratio of t-channel single top to single anti-top production [2, 3], as seen in Fig. 1, which is a constraint on u/d.

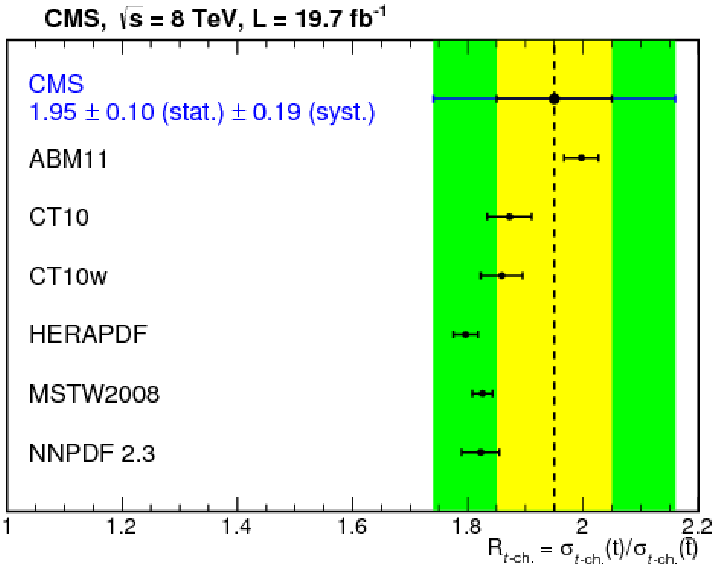

Figure 1: The CMS measurement of the $t/\bar{t}$ ratio in t-channel production, figure from [2].

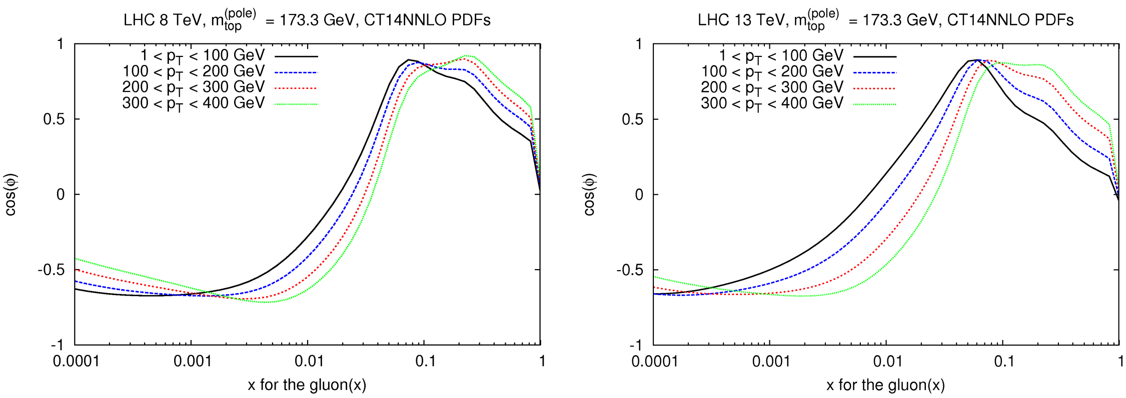

Figure 2: The correlation between top pair production in different bp_T bins and the gluon, figures from [4].

The CT14 PDF sets [4] have been made recently available at NLO, NNLO, and also at LO. These sets include a variety of LHC data sets as well as the most recent D0 data on electron charge asymmetry. The PDFs also use an updated parametrization based on Bernstein polynomials which peak at a specific x. LHC inclusive jet data are included at NLO and also in the NNLO fit. The main change in the PDFs as compared to CT10 is a softer high-x gluon, a smaller strange quark (partially due to correction of the charged current DIS cross section code) and the details of the flavour decomposition, e.g. $\bar{u}/\bar{d}$ and the high-x valence quarks. CT14 does not fit top quark production data but does make comparisons, e.g. the correlation of top pair production with the gluon are shown in Fig.2.

The NNPDF3.0 PDFs [5] are the recent major update within the NNPDF framework. As new data they include HERA inclusive structure function Run II data from H1 and ZEUS (before their combination), more recent ATLAS, CMS and LHCb data on gauge boson production and inclusive jets, and W+charm and top quark pair production. A subset of jet data is included at NNLO using an approximate NNLO treatment. The full set of data fit is illustrated in Fig. 3. The NNPDF3.0 fitting procedure has been tuned using a closure test, i.e, by generating pseudo-data based on an assumed underlying set of PDFs. One verifies in this case that the output of the fitting procedure is consistent with the a priori known answer. As a by-product, one can investigate directly the origin of PDF uncertainties. The minimization has been optimized based on the closure test. The NNPDF3.0 PDFs display moderate changes in comparison to NNPDF2.3: specifically somewhat smaller uncertainties and a noticeable change in the gluon-gluon luminosity which is mainly due to the change in methodology.

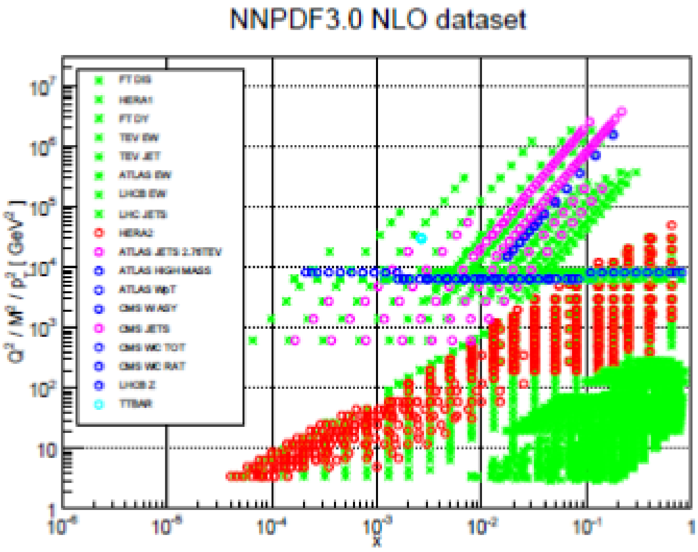

Figure 3: The data included in the NNPDF3.0 analysis, figure from [5].

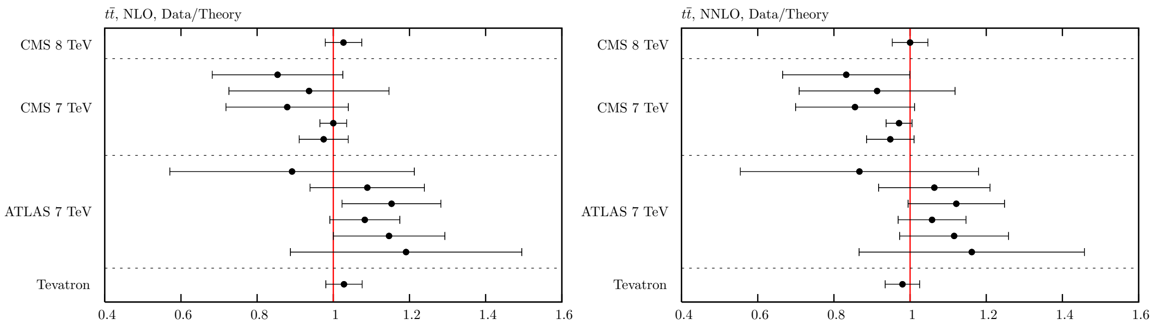

Figure 4: The MMHT fit to $\sigma_{t\bar{t}}$ data, figures from [6].

The MSTW group is renamed MMHT due to a change in personnel. The MMHT2014 PDFs [6] incorporate the improved parametrization and deuteron corrections in the MMSTWW study [7], and also a change in the heavy flavour scheme, and a change in the branching fraction $B_\mu = B(D \rightarrow \mu)$ used in the determination of the strange quark from $\nu N \rightarrow \mu\mu X$ data. The updated analysis includes new data: the combined HERA structure function data, improved Tevatron lepton asymmetry data, vector boson and inclusive jet data from the LHC (though LHC jet data is not included at NNLO), and top pair cross section data from the Tevatron and LHC. No PDFs change dramatically in comparison to MSTW2008[8], with the most significant changes being the shift in the small-x valence quarks already observed in the MMSTWW study, a slight increase in the central value of the strange quark to help the fit to LHC data, and a much expanded uncertainty on the strange distribution. The PDFs are made available with 25 eigenvector pairs for $\alpha_S(m_Z^2) = 0.118$ and 0.120 at NLO and 0.118 at NNLO. However, $\alpha_S(m_Z^2)$ is also determined by the NLO and NNLO fits and values of $\alpha_S(m_Z^2) = 0.1201$ and 0.1172 respectively are found, in good agreement with the world average. A dedicated study of the uncertainties in the determination of $\alpha_S(m_Z^2)$ in the MMHT2014 analysis has been presented in [9].

MMHT fit to data on $\sigma_{t\bar{t}}$ from the Tevatron (combined cross section measurement from D0 and CDF), and all published data from ATLAS and CMS for 7TeV and one point at 8TeV. They use $m_t = 172.5$ GeV with an error of 1 GeV and with χ^2 penalty applied. The predictions and the fit are good, with the NLO fit preferring masses slightly below $m_t = 172.5$ GeV and NNLO masses slightly above, see Fig. 4. The fit quality to $\sigma_{t\bar{t}}$ data alone is very sensitive to m_t and $\alpha_S(M_Z^2)$ interplay [9]. In the NLO fit the inclusive $t\bar{t}$ cross section data used does not constrain any PDF eigenvectors. Nearly constrains eigenvector number 29 and 31, both of which correspond to a decreased gluon at high x only. 31 is primarily constrained by CDF jet data. In the NNLO fit the inclusive $t\bar{t}$ cross section constrains one eigenvector, number 29 and (nearly) 41. Both correspond to increased gluon at high x only. The eigenvectors are shown in Fig. 5

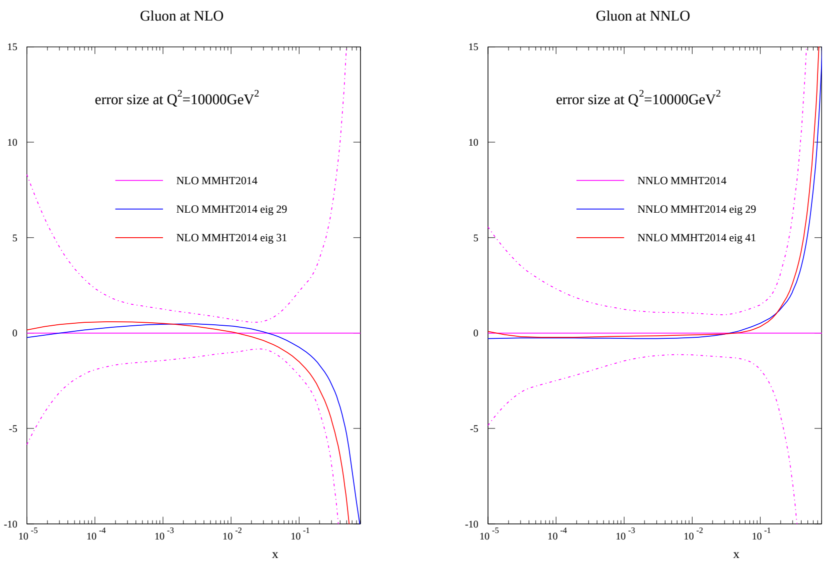

Figure 5: Eigenvectors constraints from top cross section data for MMHT.

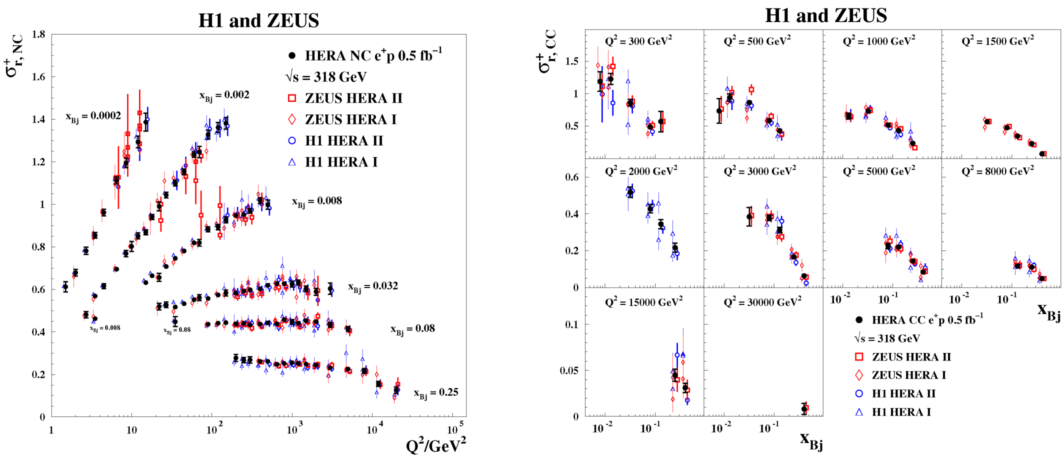

Figure 6: Neutral(left) and charged (right) current data from the final HERA combination, figure from [11].

The data for $t\bar{t}$ differential distributions are not currently used in PDF determinations as they did not meet cut-off dates for data inclusion and also had missing NNLO corrections which may be important. In comparison with existing PDFs at NLO the $y_{\bar{t}t}$ distribution tends to be very good, but the p_t distribution off in shape, while $m_{\bar{t}t}$ is somewhere in between). It is interesting to see the NNLO corrections [10] improve the comparison to the p_T distribution markedly.

Since these updates a HERA combination of all inclusive structure function measurements from Runs I and II has been presented [11], and included in the HERAPDF2.0 set. The improved data can be seen in Fig. 6. The resulting HERAPDF set has considerably reduced uncertainties, and a much improved constraint on flavour decomposition at moderate and high x due to the difference between neutral current e^+ and e^- cross sections, and to much more precise charged current data. The running at different energies gives sensitivity to $F_L(x,Q^2)$ which constrains the gluon.

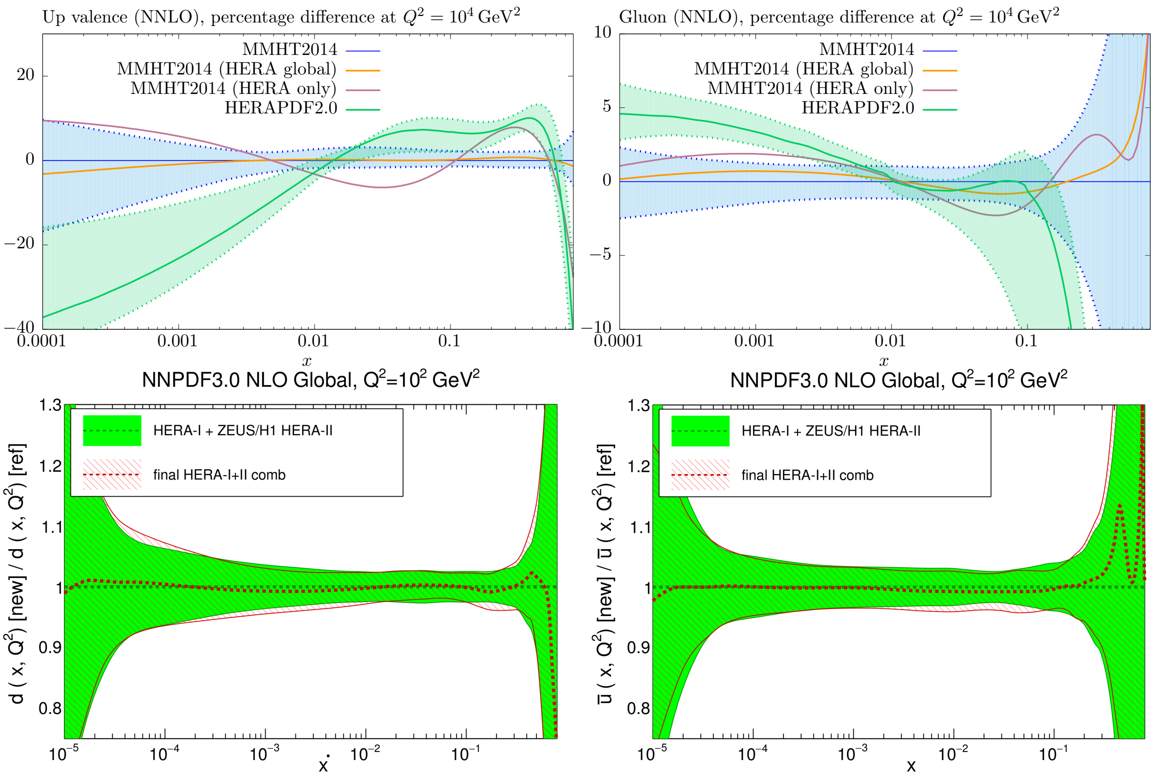

Figure 7: MMHT (top) (figures from [12]) and NNPDF (bottom) (figures from [14]) PDFs with the inclusion of the final HERA combined data.

These HERA combined data have now been included in global fits. Good fits, with little deterioration for other data are obtained for both MMHT [12, 13] and NNPDF [14]. These also result in small changes in the central PDFs and uncertainties, as shown in Fig. 7. Hence, there is no imperative to provide immediate further updates since these will appear soon due to new LHC data.

The comparison between the most recent versions of the different PDF sets is shown for the gluon and up quark in the upper of Fig. 8. There is now excellent agreement between CT14, MMHT2014 and NNPDF3.0, much better than in the previous versions of these PDF sets, but there is still some significant differences in central values and uncertainty between the other PDF sets. The comparison of PDF luminosities is shown also shown in the lower of Fig. 8. The gg

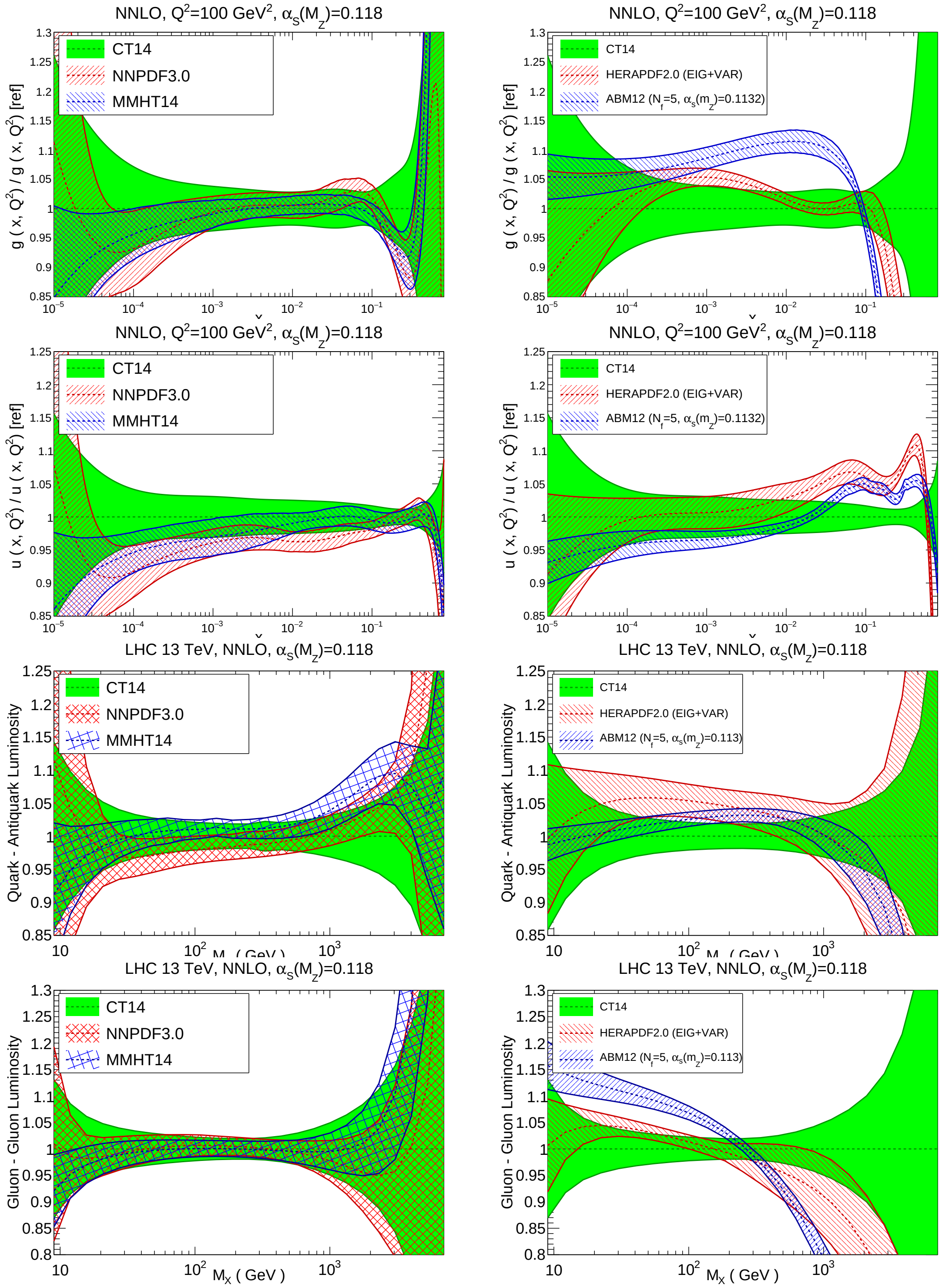

Figure 8: The comparison of different PDFs (top two plots) and parton luminosities (lower two plots), figures from [15].

luminosity now in almost perfect agreement for the three "global" sets, but some variation is seen in quark (antiquark) luminosities.

2. Combination of PDF sets

It is not obvious how to combine different "Hessian" PDF sets. However, it is now known how to generate "random" PDF sets directly from the representation in terms of eigenvectors [16]

$$F(\mathscr{S}_k) = F(S_0) + \sum_j \left[F(S_j^{\pm}) - F(S_0)\right] |R_{jk}| \tag{2.1}$$

Hence, one can combine different PDF sets either at PDF level or predictions. The latter is shown using the last round of global PDFs for the Higgs cross section in Fig. 9, and can be applied to the PDFs at a particular x and Q^2 value in the same manner.

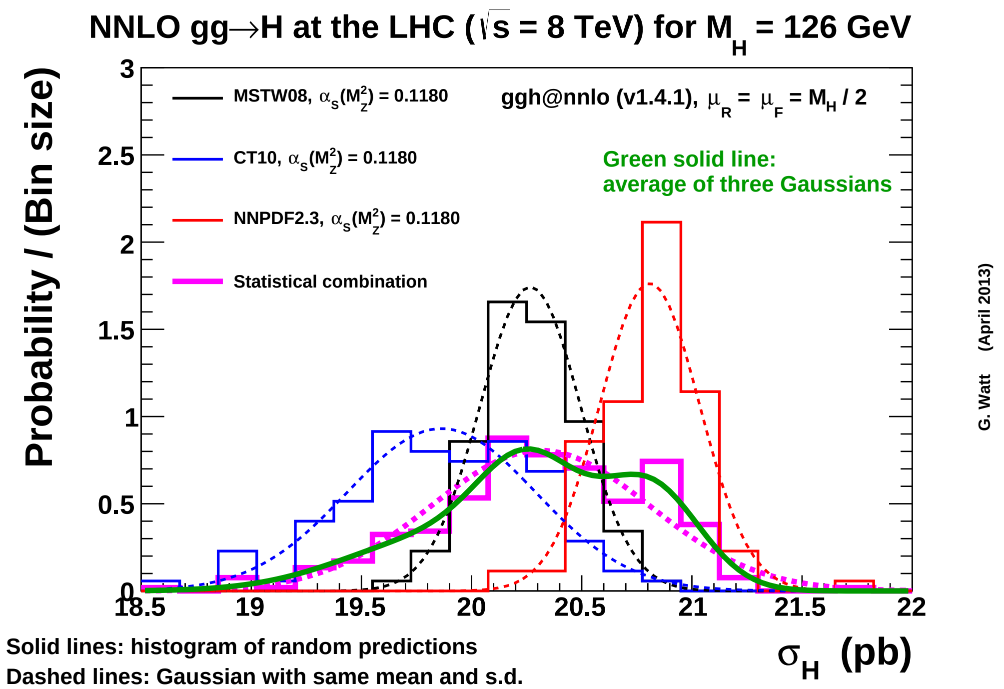

Figure 9: Combination of distributions for $\sigma_{gg\to H}$ (plot by G. Watt. [17]).

The application to the combination of the CT14, MMHT2014 and NNPDF3.0 PDFs is shown in Fig. 10. It works well if the PDFs are fairly compatible - both in central value and uncertainty - giving the mean of the central values and a spread which combines the individual PDF uncertainties and the variation in the PDFs. Following this initial development the Meta-PDF approach [18] subsequently showed how refit the combination in terms of a large number Monte Carlo PDFs to a functional form, and hence convert the combination to Hessian set with a relatively small number of eigenvector sets. Further developments showed how to compress the Monte Carlo set to a smaller number [19] and how to use the Monte Carlo sets in the combination as a basis for an extremely precise Hessian representation (MC-H) [20].

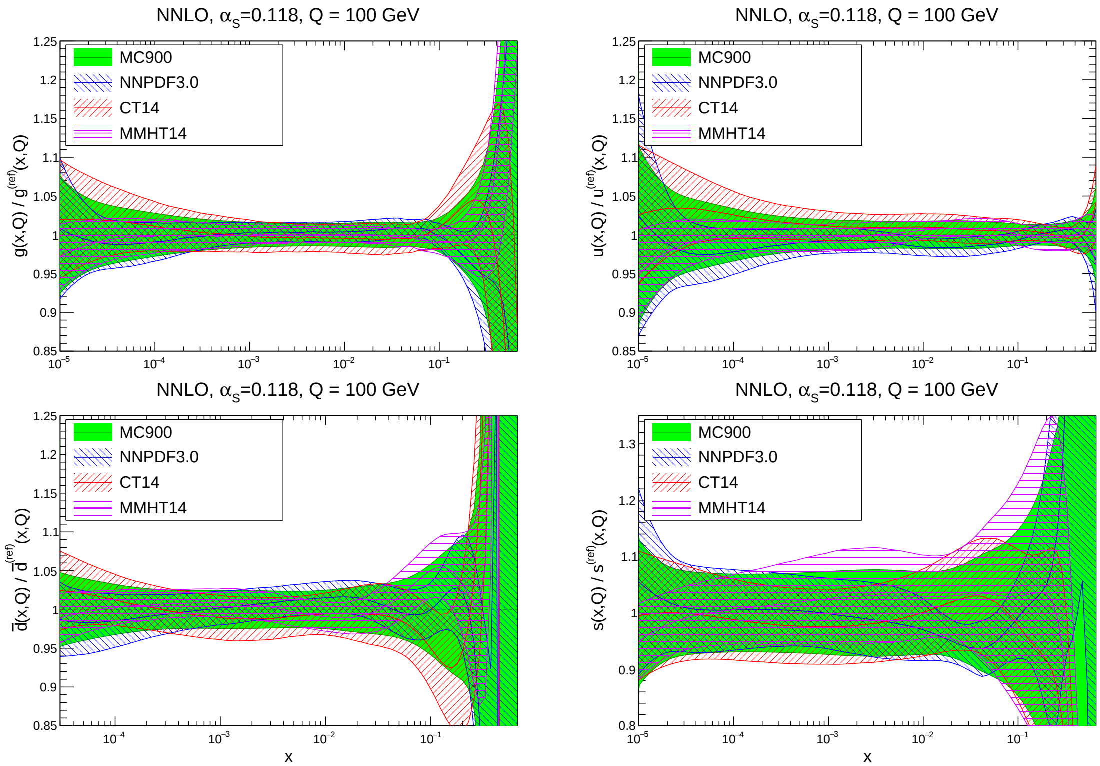

Figure 10: The combination of 300 randomly distributed sets of each of the CT14, MMHT2014 and NNPDF3.0 PDF sets, figures from [15].

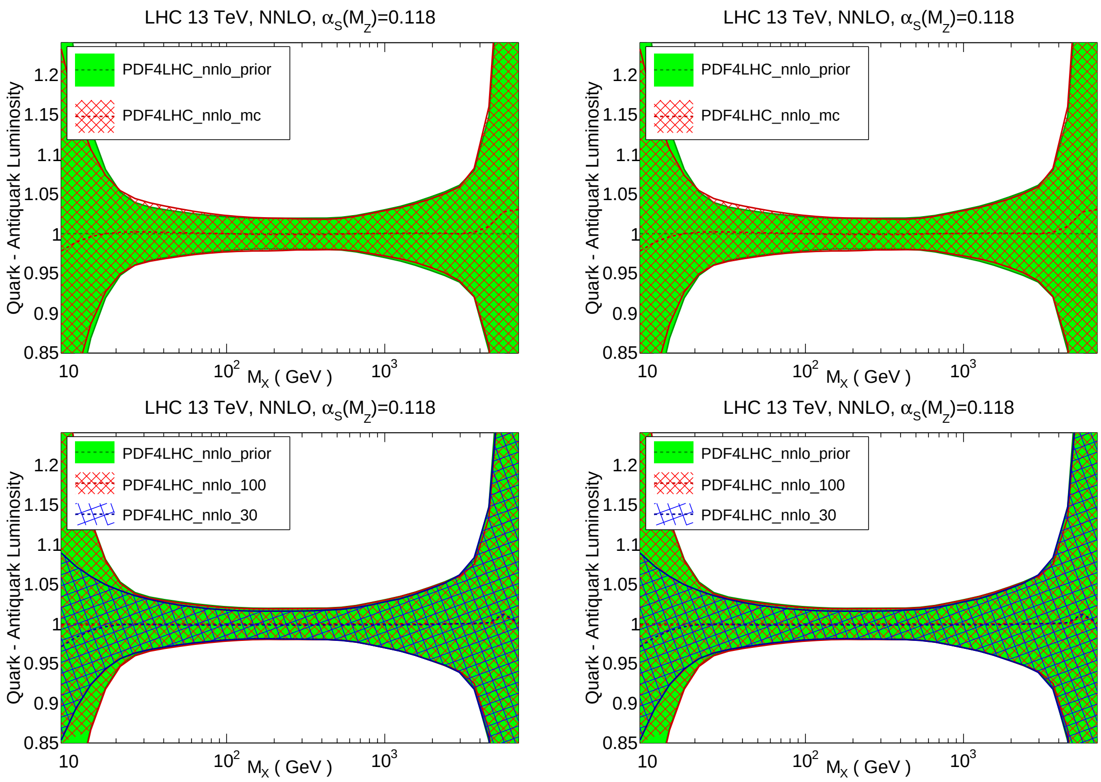

Figure 11: Comparison of PDF luminosities for Monte Carlo compression (top) and Hessian compression (bottom), figures from [15].

3. The New PDF4LHC Prescription

The improved agreement of the global PDF sets and the means of combining them in a more statistically robust fashion allows for an update in the previous PDF4LHC prescription [21] for combining PDFs when a single prediction representing a reasonable average prediction and quite conservative uncertainty is required. The sets entering into the combination must satisfy requirements, i.e. be compatible for combination, and at present CT14, MMHT2014 and NNPDF3.0 are included. It has been agreed that this should be for the common value of the coupling $\alpha_S(M_Z^2) = 0.118$. The recommendation now allows the use of a single combined PDF set in either Monte Carlo or Hessian form [15]: Monte Carlo - A set of PDF replicas is delivered, where the mean is the central value and the standard deviation the uncertainty; Hessian - a central set and eigenvectors representing orthogonal sources of uncertainty are delivered, and the uncertainty obtained by summing each uncertainty source in quadrature. In each case a single combined set at both $\alpha_S(M_Z^2) = 0.1165$ and $\alpha_S(M_Z^2) = 0.1195$ is provided to give the $\alpha_S(M_Z^2)$ uncertainty (i.e. $\Delta\alpha_S(M_Z^2) = 0.0015$) to be added in quadrature with other uncertainties.

Three different options are provided along with suggestions for when they should be used:
PDF4LHC15-mc: A compressed **Monte Carlo** set with $N_{\rm rep} = 100$ [19]. Contains non-gaussian features – important for searches at high masses (high x). See Fig. 11 for the compressed set compared to the full 900 starting PDFs.
PDF4LHC15-30: A symmetric **Hessian** set with $N_{\rm eig} = 30$. (Meta-PDF approach [18].) This has good precision and is useful for many experimental needs and when using nuisance parameters
PDF4LHC15-100: A symmetric **Hessian** set with $N_{\rm eig} = 100$ (MC-H) [20]. This has optimal precision if running time is not a problem or extreme accuracy needed. See Fig. 11 for the both Hessian sets compared to the full 900 starting PDFs.

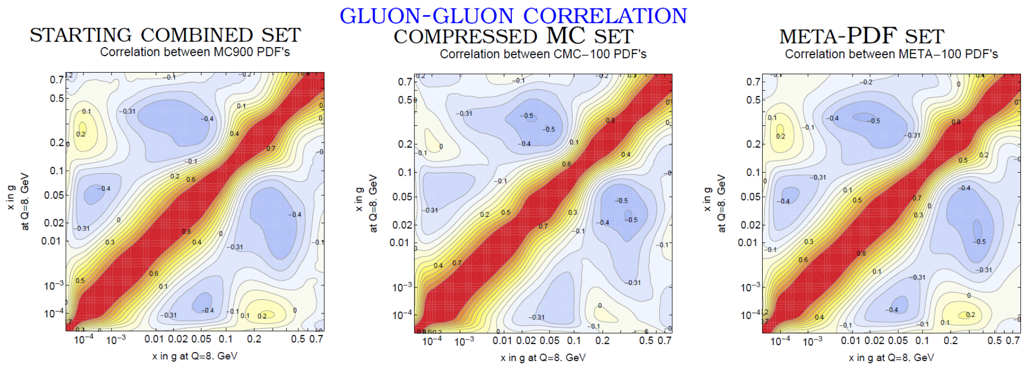

Figure 12: Comparison of PDF correlations from various means of combination, figures from [15]

PDF correlations are maintained by the compression in all cases. An example is shown in Fig. 12. The results for cross sections using all the compressed sets for LHC quantities work, at worst, quite well, even in more extreme regions of kinematics, see Fig. 13.

Finally, it is important to note that the PDF4LHC prescription is meant for assessment of the PDF uncertainty in searches, discovery, acceptance corrections ... (e.g. Higgs, Susy). When comparing theory predictions to experiment in well-determined standard model processes, e.g. jets, W,Z distributions, it is recommended to use the individual PDF sets. Other than the very first

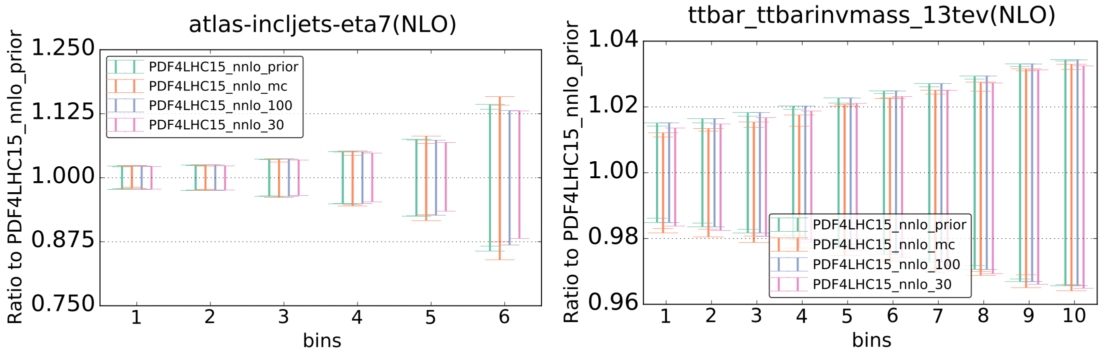

Figure 13: Examples of differential cross sections using each means of combination, figures from [15].

measurements at new energies processes such as top pair cross sections, differential distributions etc will tend to fall into the latter category, especially when real precision is reached.

Acknowledgements

I would like to thank the members of the MMHT collaboration and of the PDF4LHC working group for discussions. This work is supported partly by the London Centre for Terauniverse Studies (LCTS), using funding from the European Research Council via the Advanced Investigator Grant 267352. I thank the STFC for support via grant awards ST/J000515/1 and ST/L000377/1.

References

[1] S. Alekhin, J. Blumlein and S. Moch, Phys. Rev. D **89** (2014) 5, 054028 doi:10.1103/PhysRevD.89.054028 [arXiv:1310.3059 [hep-ph]].

[2] V. Khachatryan *et al.* [CMS Collaboration], JHEP **1406** (2014) 090 doi:10.1007/JHEP06(2014)090 [arXiv:1403.7366 [hep-ex]].

[3] G. Aad *et al.* [ATLAS Collaboration], Phys. Rev. D **90** (2014) 11, 112006 doi:10.1103/PhysRevD.90.112006 [arXiv:1406.7844 [hep-ex]].

[4] S. Dulat *et al.*, Phys. Rev. D **93** (2016) no.3, 033006 doi:10.1103/PhysRevD.93.033006 [arXiv:1506.07443 [hep-ph]].

[5] R. D. Ball *et al.* [NNPDF Collaboration], JHEP **1504** (2015) 040 doi:10.1007/JHEP04(2015)040 [arXiv:1410.8849 [hep-ph]].

[6] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, Eur. Phys. J. C **75** (2015) 5, 204 doi:10.1140/epjc/s10052-015-3397-6 [arXiv:1412.3989 [hep-ph]].

[7] A. D. Martin, A. J. T. M. Mathijssen, W. J. Stirling, R. S. Thorne, B. J. A. Watt and G. Watt, Eur. Phys. J. C **73** (2013) 2, 2318 doi:10.1140/epjc/s10052-013-2318-9 [arXiv:1211.1215 [hep-ph]].

[8] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, Eur. Phys. J. C **63** (2009) 189 doi:10.1140/epjc/s10052-009-1072-5 [arXiv:0901.0002 [hep-ph]].

[9] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, Eur. Phys. J. C **75** (2015) 9, 435 doi:10.1140/epjc/s10052-015-3630-3 [arXiv:1506.05682 [hep-ph]].

[10] M. Czakon, D. Heymes and A. Mitov, Phys. Rev. Lett. **116** (2016) no.8, 082003 doi:10.1103/PhysRevLett.116.082003 [arXiv:1511.00549 [hep-ph]].

[11] H. Abramowicz *et al.* [H1 and ZEUS Collaborations], Eur. Phys. J. C **75** (2015) 580 doi:10.1140/epjc/s10052-015-3710-4 arXiv:1506.06042 [hep-ex].

[12] R. S. Thorne, L. A. Harland-Lang, A. D. Martin and P. Motylinski, PoS EPS **-HEP2015** (2015) 505 [arXiv:1508.06621 [hep-ph]].

[13] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, arXiv:1601.03413 [hep-ph], Eur. Phys. J. C in press.

[14] J. Rojo, PoS EPS **-HEP2015** (2015) 506 [arXiv:1508.07731 [hep-ph]].

[15] J. Butterworth *et al.*, J. Phys. G **43** (2016) 023001 doi:10.1088/0954-3899/43/2/023001 [arXiv:1510.03865 [hep-ph]].

[16] G. Watt and R. S. Thorne, JHEP **1208** (2012) 052 doi:10.1007/JHEP08(2012)052 [arXiv:1205.4024 [hep-ph]].

[17] G. Watt. 2013. talk presented at PDF4LHC meeting, 17th April, CERN, Geneva. http://indico.cern.ch/event/244768/contribution/5.

[18] J. Gao and P. Nadolsky, JHEP **1407** (2014) 035 doi:10.1007/JHEP07(2014)035 [arXiv:1401.0013 [hep-ph]].

[19] S. Carrazza, J. I. Latorre, J. Rojo and G. Watt, Eur. Phys. J. C **75** (2015) 474 doi:10.1140/epjc/s10052-015-3703-3 [arXiv:1504.06469 [hep-ph]].

[20] S. Carrazza, S. Forte, Z. Kassabov, J. I. Latorre and J. Rojo, Eur. Phys. J. C **75** (2015) 8, 369 doi:10.1140/epjc/s10052-015-3590-7 [arXiv:1505.06736 [hep-ph]].

[21] M. Botje *et al.*, arXiv:1101.0538 [hep-ph].

PoS(TOP2015)053