Uncertainties on MSHT PDFs.

Robert Thorne

September 22nd 2022



University College London

MSHT PDFs - major update [1].

Theoretical Procedures

As before use a general mass variable flavour scheme based on the Thorne-Roberts scheme, using the "optimal" choice of parameters for smoothness near threshold.

Use deuteron and heavy nuclear corrections. Former fit using 4 parameter model, as in MMHT14 and latter use same corrections (arXiv:1112.6324) as MMHT14 with additional penalty-free freedom of order 1%.

Fit data with systematics uncertainties using either nuisance parameters if possible (preferred method) or with the correlation matrix provided. Now also use statistical correlations whenever provided.

Fit to absolute cross sections in preference to normalized if both available.

Extension of parameterisation.



Illustration of precision possible with increasing n, sea-like (left) and valence-like (right) (where pseudo-data for x > 0.01). Using n = 6 would lead to much better than 1% precision.

For most PDFs n = 4 default for MMHT2014 – 36 parameters.

Now extend to n = 6 – total of 51 parton parameters.

When determining uncertainties go from 25 eigenvector pairs to 32.

New LHC data fit.

Include all our recent LHC data updates in the fit at NNLO (for default $\alpha_S(M_Z^2) = 0.118$).

	no. points	NNLO χ^2
D0 W asymmetry	14	12.0
$\sigma_{t\bar{t}}$ Tevatron +CMS+ATLAS 7, 8TeV	17	14.5
LHCb 7+8 TeV $W + Z$	67	99.4
LHCb 8 TeV e	17	26.2
CMS 8 TeV W	22	12.7
ATLAS 7 TeV jets $R = 0.6$	140	221.6
CMS 7 TeV $W + c$	10	8.6
ATLAS 7 TeV W, Z	61	116.6
CMS 7 TeV jets $R = 0.7$	158	175.8
ATLAS 8 TeV Zp_T	104	188.5
CMS 8 TeV jets	174	261.3
ATLAS 8 TeV $tar{t} ightarrow l+{ m j}$ single-diff	25	25.6
ATLAS 8 TeV $t\bar{t} ightarrow l^+ l^-$ single-diff	5	3.5
ATLAS 8 TeV high-mass Drell-Yan	48	56.7
ATLAS 8 TeV $W^{+,-}$ + jet	32	18.1
CMS 8 TeV $(d\sigma_{tar{t}}/dp_{T,t}dy_t)/\sigma_{tar{t}}$	15	22.5
ATLAS 8 TeV W^+, W^-	22	57.4
CMS 2.76 TeV jets	81	102.9
CMS 8 TeV $t\bar{t} y_t$ distribution	9	13.2
ATLAS 8 TeV double differential Z	59	85.6
total	4363	5122

Fit quality generally good. Relatively poor χ^2 values for some sets all observed by other groups. Fit to some data now very poor at NLO.

Changes in MSHT PDFs



Most significant in d_V (parameterisation and new LHC data) and strange quark.



Uncertainties mainly brought down by new data, but some uncertainty increases due to parameterization changes, e.g. d_V and $(\bar{d}-\bar{u}) \rightarrow (\bar{d}/\bar{u})$ at small x.

Changes in PDF uncertainties downwards but also in central values.

Gluon



Down



0.85

0.80

10⁻³

10⁻²

х

10⁻¹

10⁻²

х

10⁻¹

0.85

0.80

10⁻³

Theoretical Uncertainties - Higher Orders (J. McGowan, T. Cridge, L. Harland-Lang, RT)

Leading source of uncertainties is from from Missing Higher Orders in perturbation theory. Numerous sources of this for e.g structure functions, i.e. splitting functions

$$\boldsymbol{P}(x,\alpha_s) = \alpha_s \boldsymbol{P}^{(0)}(x) + \alpha_s^2 \boldsymbol{P}^{(1)}(x) + \alpha_s^3 \boldsymbol{P}^{(2)}(x) + \alpha_s^4 \boldsymbol{P}^{(3)}(x) + \dots ,$$

but also heavy flavour transition matrix elements and cross-sections (coefficient functions)

$$F_2(x,Q^2) = \sum_{\alpha \in \{H,q,g\}} \left(C_{q,\alpha}^{\mathrm{VF}, n_f+1} \otimes A_{\alpha i} (Q^2/m_h^2) \otimes f_i^{n_f}(Q^2) + C_{H,\alpha}^{\mathrm{VF}, n_f+1} \otimes A_{\alpha i} (Q^2/m_h^2) \otimes f_i^{n_f}(Q^2) \right),$$

Current knowledge is up to NNLO, with higher orders unknown.

Already progress in calculating features at N^3LO [2-13].

Uncertianties as Nuisance Parameters.

Theoretical uncertainties in PDFs have been addressed via scale variations [14,15].

Do we need to wait for a full description of the next order to be able to use the knowledge we have? Usual probability distribution

$$P(T|D) \propto \exp\left(-\frac{1}{2}(T-D)^T H_0(T-D)\right)$$

Can attempt to parameterise the higher order effects with a nuisance parameter defined by a prior probability distribution [16], see also [17].

Allow the fit to move these $N^{3}LO$ parameters (with a penalty attached to ensure we stay close to the behaviour already known).

$$T' + (\theta - t)u = T + tu + (\theta - t)u.$$

Defining $\theta' = \theta - t$ and

HP2 2022 - Sep. 2022

$$P(\theta') = \frac{1}{\sqrt{2\pi\sigma_{\theta'}}} \exp(-\theta'^2/2\sigma_{\theta'}^2).$$

$$P(T|D\theta) \propto \exp\left(-\frac{1}{2}(T+tu+\frac{(\theta-t)}{\sigma_{\theta'}}u-D)^T H_0(T+tu+\frac{(\theta-t)}{\sigma_{\theta'}}u-D)\right)$$
$$P(T'|D\theta') \propto \exp\left(-\frac{1}{2}(T'+\frac{\theta'}{\sigma_{\theta'}}u-D)^T H_0(T'+\frac{\theta'}{\sigma_{\theta'}}u-D)\right)$$

Overall we obtain

$$P(T'|D) \propto \int d\theta \exp\left(-\frac{1}{2}\left[(T' + \frac{\theta'}{\sigma_{\theta'}}u - D)^T H_0(T' + \frac{\theta'}{\sigma_{\theta'}}u - D) + {\theta'}^2/\sigma_{\theta'}^2\right]\right).$$

With these alterations, we follow the same practice as set out in the MSHT20 NNLO PDF fit - the exact same global fit is done to approximate $N^{3}LO$ (a $N^{3}LO$).

N³LO - What do we know?

Zero-mass structure function $N^{3}LO$ coefficient functions are known [2].

Some knowledge of leading terms in the small x and large x regime [3-12], e.g.

$$\boldsymbol{P}_{qg}^{(3)}(x) \to \frac{C_A^3}{3\pi^4} \left(\frac{82}{81} + 2\zeta_3\right) \frac{1}{2} \frac{\ln^2 1/x}{x} + \rho_{qg} \frac{\ln 1/x}{x},$$

Some numerical constraints (Low-integer Mellin moments) [3-12].

Intuition from lower orders and expectations from perturbation theory.

Very little about many crosssections (K-factors).



Splitting Functions at aN³LO (Note – small bug corrected.)

 N_m Mellin moments [2-6] can be used as constraints to define

$$F(x) = \sum_{i=1}^{N_m} A_i f_i(x) + f_e(x).$$

Choose a set of relevant functions and solve for A_i .

Introduce a degree of freedom a, interpreted as a nuisance parameter allowed to vary in a PDF fit, $f_e(x) \to f_e(x, a)$. In our treatment it is the coefficient of the most divergent unknown small-x term, e.g. for $P_{aa}^{(3)}(x)$

$$f_1(x) = \frac{1}{x} \quad \text{or} \quad \ln^4 x \quad \text{or} \quad \ln^3 x \quad \text{or} \quad \ln^2 x,$$

$$f_2(x) = \ln x,$$

$$f_2(x) = 1 \quad \text{or} \quad x \quad \text{or} \quad x^2,$$

$$f_3(x) = \ln^4(1-x) \quad \text{or} \quad \ln^3(1-x) \quad \text{or} \quad \ln^2(1-x) \quad \text{or} \quad \ln(1-x),$$

$$f_e(x, \rho_{qg}) = \frac{C_A^3}{3\pi^4} \left(\frac{82}{81} + 2\zeta_3\right) \frac{1}{2} \frac{\ln^2 1/x}{x} + \rho_{qg} \frac{\ln 1/x}{x}.$$

Resulting splitting functions



Uncertainty largest at small x. Best fit largely compatible with best estimate.

Transition Matrix Elements at aN³LO

Following the same general procedure as for the splitting functions.



K-factors at aN³LO

Parameterise the N³LO K-factor as a superposition of both NNLO and NLO K-factors.

Allows the fit to decide on a shape (based on the shapes of preceding orders) and an overall magnitude.

$$K(y) = 1 + \frac{\alpha_s}{\pi} D(y) + \left(\frac{\alpha_s}{\pi}\right)^2 E(y) + \left(\frac{\alpha_s}{\pi}\right)^3 F(y) + \mathcal{O}(\alpha_s^4)$$
$$K^{\text{N}^3\text{LO/LO}} = K^{\text{NNLO/LO}} \left(1 + \alpha_s^3 \hat{a}_1 \frac{\mathcal{N}^2}{\pi} D + \alpha_s^3 \hat{a}_2 \frac{\mathcal{N}}{\pi^2} E\right).$$

Hence default is no correction at $N^{3}LO$.

Correlated K-factors for each of the 5 processes: DY, Top, Jets (or Dijets), Zp_T and vector boson jets and Dimuon.

 \hat{a}_1, \hat{a}_2 could be included as correlated with PDF parameters (incl. other N³LO theory parameters) or as completely decorrelated from the inclusive DIS process.

Global Fit Quality at aN³LO

We see a reduction in χ^2 from NNLO across all datasets ($\Delta \chi^2 = -160$ for 20 extra parameters).

ATLAS 8 TeV $Z p_T$ [18] sees a huge reduction in χ^2 .

This is a similar reduction found at NNLO when HERA datasets were not included [1].

In the aN³LO fit, we also see an improvement in the fit to HERA data.

Dataset	N _{pts}	χ^2	$\Delta \chi^2$ from		
			NNLO		
HERA $ep F_2^{charm}$	79	143.7	+11.4		
NMC/BCDMS/SLAC/HERA F_L	57	45.6	-22.9		
HERA e^+p CC	39	49.7	-2.3		
HERA e^-p CC	42	64.9	-5.3		
HERA e^+p NC 820 GeV	75	84.3	-5.6		
HERA e^-p NC 460 GeV	209	247.7	-0.6		
HERA e^+p NC 920 GeV	402	474.0	-38.7		
HERA e^-p NC 575 GeV	259	248.5	-14.5		
HERA e^-p NC 920 GeV	159	243.0	-1.4		
ATLAS W^+ , W^- , Z	30	30.0	+0.1		
CMS double diff. Drell-Yan	132	133.2	-11.3		
LHCb 2015 W, Z	67	103.2	+3.8		
ATLAS 7 TeV jets	140	215.9	-5.6		
ATLAS 7 TeV high prec. W, Z	61	119.3	+2.7		
CMS 7 TeV jets	158	186.8	+11.0		
ATLAS 8 TeV $Z p_T$	104	108.4	-80.0		
CMS 8 TeV jets	174	271.3	+10.0		
ATLAS 8 TeV High-mass DY	48	62.8	+5.7		
ATLAS 8 TeV $W + jets$	30	18.8	+0.7		
ATLAS 8 TeV W	22	53.0	-4.4		
CMS 2.76 TeV jet	81	109.8	+6.9		
DY data Total	864	1069.4	-18.5		
Top data Total	71	75.1	-4.2		
Jets data Total	739	963.6	+21.5		
p_T Jets data Total	144	138.0	-77.2		
Dimuon data Total	170	125.0	-1.2		
DIS data Total	2375	2580.9	-90.8		
Total	4363	4961.2	-160.1		

The overall χ^2 follows the general trend one may expect from perturbation theory.

	LO	NLO	NNLO	N ³ LO	
$\chi^2_{N_{pts}}$	2.57	1.33	1.17	1.14	

Evidence that including $aN^{3}LO$ has reduced tensions between small and large-x.

 χ^2 reduction is mostly due to new theory, not just from K-factors included in fit.

Average penalty for included 20 aN³LO parameters is 0.46.

Low- Q^2 Coefficient			
$c_{q}^{\text{NLL}} = -3.868$	0.004	$c_q^{\text{NLL}} = -5.837$	0.844
Transition Matrix Elements			
$a_{Hg} = 12214.000$	0.601	$a_{qq,H}^{NS} = -64.411$	0.001
$a_{gg,H} = -1951.600$	0.857		
Splitting Functions			
$ ho_{qq}^{NS}=0.007$	0.000	$\rho_{gq} = -1.784$	0.802
$ ho_{qq}^{PS} = -0.501$	0.186	$\rho_{gg} = 19.245$	3.419
$\rho_{qg} = -1.754$	0.015		
K-factors			
$DY_{NLO} = -0.307$	0.094	$DY_{NNLO} = -0.230$	0.053
$Top_{NLO} = 0.041$	0.002	$Top_{NNLO} = 0.651$	0.424
$Jet_{NLO} = -0.300$	0.090	$Jet_{NNLO} = -0.691$	0.478
$p_T \text{Jets}_{\text{NLO}} = 0.583$	0.339	$p_T \text{Jets}_{\text{NNLO}} = -0.080$	0.006
$Dimuon_{NLO} = -0.444$	0.197	$Dimuon_{NNLO} = 0.922$	0.850
N ³ LO Penalty Total	9.262 / 20	Average Penalty	0.463
		Total	4961.2 / 4363
		$\Delta \chi^2$ from NNLO	-160.1

The PDFs at aN³LO compared to NNLO.



Gluon and quarks larger at small x.

The PDFs at aN³LO compared to NNLO - detail.

The gluon is enhanced at small-x due to the large logarithms present at higher orders.

Charm receives a sizeable contribution from $A_{Hg}^{(3)}$ at high*x* and the gluon at small *x* involved in convolution.



Light quarks enhanced slightly at high x.

Correlated and uncorrelated K- 1.025 factors show consistent uncertainty1.000 predictions. 0.975

Strange quark enhanced a little at higher x compared to NNLO.

aN³LO follows more closely the NNLO fit to non- HERA datasets at high-x – reduced tension between small-x HERA data and other data.



The PDFs at $aN^{3}LO$ with theoretical uncertainty.



Light quark uncertainty enhanced slightly at low x.

Correlated and uncorrelated K- 1.025 factors show consistent uncertainty1.000 predictions. 0.975

Strange quark uncertainty a little at higher x.



$lpha_S(M_Z^2)$ and m_c at aN^3LO

Both and show a quadratic behaviour around their respective minima.

Best fit is $\alpha_S(M_Z^2) = 0.1170$.

MSHT20 NNLO: $\alpha_S(M_Z^2) = 0.1174 \pm 0.0013$.

MSHT20 NLO: $\alpha_S(M_Z^2) = 0.120 \pm 0.0015.$

Best fit $m_c \sim 1.45$ GeV.

Both these results suggest that the fit is preferring a slight suppression of the PDFs, particularly the enhanced gluon and charm.

The aN³LO $\alpha_S(M_Z^2)$ result overlaps with the NNLO world average within uncertainties.





Resulting K-factors – Drell Yan Processes.



Predict a 1% decrease in the DY K-factors from NNLO.

In agreement with recent results found using NNLO PDFs with aN3LO cross section[19].

Resulting K-factors – Top Quark Processes.

Top K-factors see an overall increase in magnitude, consistent with recent results[20].

Results show a marginally better fit overall.



K-factor for CMS 8 TeV single diff. shown here.

Resulting K-factors – $Z p_T$ Processes.



Data prefers large N³LO corrections, but χ^2 improvement dramatic just with PDF change.

HP2 2022 - Sep. 2022

Higgs predictions at N³LO with Theoretical Uncertainty.

Good agreement between NNLO and $aN^{3}LO$ for gluon fusion (top).

Cancellation between N³LO cross section and PDFs not automatic.

Less cancellation for VBF (bottom).

However variation between orders is smaller for VBF cross-section.



Usage of aN³LO PDFs.

If (N³LO) cross-sections are known use (aN³LO) PDFs and their associated theoretical uncertainties.

For DIS processes, using the $(aN^{3}LO)$ PDF set is advised with use of $aN^{3}LO$ $(aN^{3}LO)$ coefficient functions.

For any of the other 5 processes included in the fit (which we fit K-factors for), we provide the full details of these fitted aN^3LO K-factors.

For processes not included in the fit, the change of the aN³LO compared to the NNLO PDFs should be taken as representative of the potential theoretical uncertainty in the NNLO PDFs.

Dijet data at aN³LO – preliminary.

Fit quality to dijet data at NNLO shows an improvement from inclusive jet data.

Particularly better fit to $Z p_T$ data, slightly worse fit to top data.

Fit quality is also better when fitting to dijet data at aN^3LO , and dijet data improves at aN^3LO , unlike inclusive jets.

Fit quality to all other data (incl. $Z~p_T$ and top datasets) becomes marginally better $\Delta\chi^2\sim-20$

	N.	$\chi^2/N_{\rm pts}$			N.	$\chi^2/N_{\rm pts}$	
	1 pts	NNLO	aN ³ LO		1 pts	NNLO	aN ³ LO
ATLAS 7 TeV jets	140	1.58	1.54	ATLAS 7 TeV dijets	90	1.05	1.12
CMS 7 TeV jets	158	1.11	1.18	CMS 7 TeV dijets	54	1.43	1.39
CMS 8 TeV jets	174	1.50	1.56	CMS 8 TeV dijets	122	1.04	0.83
Total	472	1.39	1.43	Total	266	1.12	1.04

Change in gluon with inclusive 1.2 jets \rightarrow dijets similar at 1.15 aN³LO as at NNLO. 1.1 1.05

No significant change in uncertainty.



Conclusions

Many updates from PDF groups in the recent past. Include large but varying amounts of LHC data – starting to have a very significant impact on PDF extractions.

Theory catching up for precision data, e.g NNLO jets, differential top, Z, Wp_T ... More data leads to possible improvements in parameterisation.

Uncertainties generally come down, but not always – in some regions just more realistic.

Approximate $N^{3}LO$ PDFs are available and we encourage their use.

Available as LHAPDF grids at www.hep.ucl.ac.uk/msht/ (see publication for usage instructions).

Full information is available in the article 2207.04739

Provide an intuitive and controllable way to include theoretical uncertainties into PDFs.

Stay tuned for further developments regarding dijets (and SeaQuest).

References

- [1] S. Bailey et. al., Eur. Phys. J. C 81 341 (2021).
- [2] J. Vermaseren, A. Vogt, and S. Moch, Nuclear Physics B, 724, 3182 (2005)
- [3] S. Moch, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, Journal of High Energy, 1653, Physics, 2017, (2017)
- [4] A. Vogt et al., PoS LL2018, 050 (2018), 1808.08981
- [5] S. Moch, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, (2021), 2111.15561
- [6] S. Moch, B. Ruijl, T. Ueda, J. A. M. Vermaseren, and A. Vogt, Journal of High Energy, 1664, Physics, 2017, (2017)
- [7] I. Bierenbaum, J. Blumlein, and S. Klein, Nuclear Physics B, 820, 417 (2009)
- [8] M. Bonvini and S. Marzani, Journal of High Energy Physics, 2018, (2018)
- [9] J. Ablinger et al., Nucl. Phys. B, 886, 733 (2014), 1406.4654.
- 10] J. Ablinger et al., Nuclear Physics B, 890, 48151 (2015)
- [11] J. Ablinger et al., Nuclear Physics B, 882, 263288 (2014)
- [12] H. Kawamura, N. A. Lo Presti, S. Moch, and A. Vogt, Nucl. Phys. B, 864, 399 (2012),1689
- [13] J. Blumlein et al., PoS, QCDEV2017, 031 (2017), 1711.07957
- [14] NNPDF coll. Eur. Phys. J. C 79 838 (2019)
- [15] Z. Kassabov, M. Ubiali and C Voisey, 2207.07616
- [16] F. Tackmann, SCET 2019 Workshop (2019)
- [17] R. D. Ball and R. L. Pearson, The European Physical Journal C, 81, (2021)
- [18] H. Kawamura, N. Lo Presti, S. Moch, and A. Vogt, Nuclear Physics B, 864, 399468, 1682, (2012).
- [19] X. Chen et al., (2021), 2107.09085
- [20] N. Kidonakis, 2109.14102

Back-up

Comparisons to other PDFs. Some big differences.

Gluon and Strange



Up and Down



HP2 2022 - Sep. 2022

PDF evolution at N³LO.



35



Effect of each individual N³LO change.







HP2 2022 - Sep. 2022

Resulting K-factors – Inclusive Jet Processes.



N³LO corrections relatively small – change shape as well as normalization.

A selection of other references not directly mentioned but used for these results:

- [] G. Altarelli and G. Parisi, Nucl. Phys. B, 126, 298 (1977)
- [] E. G. Floratos, D. A. Ross, and C. T. Sachrajda, Nucl. Phys. B, 152, 493 (1979)
- [] A. Gonzalez-Arroyo and C. Lopez, Nucl. Phys. B, 166, 429 (1980)
- [] W. Furmanski and R. Petronzio, Phys. Lett. B, 97, 437 (1980)
- [] E. G. Floratos, C. Kounnas, and R. Lacaze, Nucl. Phys. B, 192, 417 (1981)
- [] S. Moch, J. Vermaseren, and A. Vogt, Nuclear Physics B, 688, 101134 (2004)
- [] A. Vogt, S. Moch, and J. Vermaseren, Nuclear Physics B, 691, 129181 (2004)
- [] M. Buza, Y. Matiounine, J. Smith, and W. L. van Neerven, The European Physical, 1668, Journal C, 1, 301320 (1998).
- [] M. Buza, Y. Matiounine, J. Smith, and W. van Neerven, Nuclear Physics B, 485, 1670, 420456 (1997).
- [] S. Catani, M. Ciafaloni, and F. Hautmann, Nucl. Phys. B, 366, 135 (1991).
- [] . Laenen and S.-O. Moch, Phys. Rev. D, 59, 034027 (1999), hep-ph/9809550