

MARIA UBIALI **UNIVERSITY OF CAMBRIDGE**

RECENT UPDATES FROM NNPDF

LHC EW WG GENERAL MEETING - CERN





European Research Council

Established by the European Commission

11TH JULY 2024

Part I: Highlights of recent progress from the NNPDF collaboration



Part II: The precision vs accuracy challenge: a road-map to robustly test accuracy

PART I: RECENT PROGRESS FROM NNPDF

🗹 Sep 2021:

NNPDF4.0

(paper & open-source code)

M Aug 2022:

Intrinsic charm

Sept 2022: PDFs & BSM

searches (A_{FB} high-mass)

Mov 2023: IC asymmetry study





Sep 2021:

NNPDF4.0

(paper & open-source code)

M Aug 2022: Intrinsic charm

🗹 Jan 2024: NNPDF4.0 QED

M Jan 2024: NNPDF4.0 MHOUs

Sept 2022: PDFs & BSM **searches** (A_{FB} high-mass)

Mov 2023: IC asymmetry study

I Feb 2024: NNPDF4.0 aN3LO **V** June 2024: NNPDF4.0 aN3LO & QED & MHOU

M June 2024: NNPDF4.0 for MC event generators



🗹 Sep 2021:	🗹 Aug 2022:	V
NNPDF4.0	Intrinsic charm	
(paper & open-source code)		

🗹 Jan 2024:	🗹 Jan 2024:	🗹 F
NNPDF4.0 QED	NNPDF4.0 MHOUs	Ν

WIP: Implications of NNPDF4.0 for LHC processes WIP: Closure test with

inconsistent experimental data

Sept 2022: **PDFs & BSM searches** (A_{FB} high-mass) ✓ Nov 2023: ICasymmetry study

⁻eb 2024: NNPDF4.0 **aN3LO** June 2024:NNPDF4.0 aN3LO &QED & MHOU

June 2024:NNPDF4.0 for MCevent generators

WIP: Towards NNPDF4.1



- Improved ML hyper-parameter optimisation from parallel replica training on GPUs [WIP]
- Fixed functional forms for Hessian fits in NNPDF [WIP]
- \rightarrow Determination of strong coupling simultaneously with PDFs at aN³LO [WIP]
- Determination of higher twist corrections [WIP]
- Updated NNPDF polarised fits and EIC projections [WIP]
- Simultaneous SMEFT and PDF fits [Costantini et al 2402.03308] [Iranipour, MU 2201.07240]
- Study of possible BSM contamination effects in PDF fits [Hammou et al 2307.10370] ➡ ...
- WIP: Implications of NNPDF4.0 for LHC processes

WIP: Closure test with

inconsistent experimental data

Beyond neural networks: PDFs from Bayesian inference [Candido et al 2404.07573, Costantini et al 2404.10056]

WIP: Towards NNPDF4.1





NNPDF40 AN3LO: SPLITTING FUNCTIONS

• Approximate parametrisation for the N³LO splitting functions satisfying known **exact results and limits**



Good perturbative consistency within uncertainties

LO, NLO, NNLO: MHOU (µ_F)

N³LO: MHOU (μ_F) + IHOUs (dark)

- Estimate **Incomplete Higher Order Uncertainties** (IHOUs) by varying interpolating functions connecting known limits
- WIP: dedicated LH benchmark paper on N3LO splitting functions and PDFs





NNPDF40 AN3LO: IMPACT ON PDF EVOLUTION



Evolution of fixed PDF boundary condition from Q=1.65 GeV to Q=100 GeV

- Effects of N³LO corrections to DGLAP evolution < 1% except at small-x and large-x
- Excellent perturbative
 convergence of PDF
 evolution, may be improved
 with small-x or large-x
 resummations



NNPDF40 AN3LO: FIT SETTINGS

- Same methodology, dataset, and pipeline for theory calculations as in NNPDF4.0 MHOU & QED sets
- Produce fit variants with and without theory uncertainties (using the theory covariance matrix)

$$\Delta_i(\rho_f, \rho_r) \equiv T_i(\rho_f, \rho_r) - T_i(0, 0),$$
$$S_{ij} = n_m \sum_{V_m} \Delta_i(\rho_f, \rho_{r_i}) \Delta_j(\rho_f, \rho_{r_j}),$$

- The theory covariance matrix includes contributions from **MHOUs** (μ_F and μ_R variations) and **IHOUs**
- variations to account for missing K-factors



Shift wrt central theory on the physical observables due to theory variations (e.g. scales)

Theory covariance matrix: combine all shifts, keeping into account their correlations

• Hadronic data is fitted using aN³LO evolution and NNLO matrix elements, supplemented by MHOUs associated to μ_R







- Without MHOUs, the **x² improves with the perturbative accuracy** of the PDF fit
- With MHOUs, the **x**² becomes **feebly dependent on the perturbative accuracy**
- At aN³LO impact of MHOUs is small (also at PDF level) but non negligible
- **N³LO corrections** required for perturbative convergence at the PDF fit level!



PERTURBATIVE CONVERGENCE AND MHOUS



Good perturbative convergence

- Impact of N³LO corrections moderate, specially for the quark luminosities
- For the gluon-gluon luminosity, NNPDF4.0 finds a **small suppression** around Higgs mass (2% effects)
- Impact of MHOUs is not negligible even at N³LO, both in terms of central values and uncertainties
- Motivates inclusion of exact N³LO calculations for hadronic processes in global PDF fits (e.g. Drell-Yan production, already available)





NNPDF40 AN3LO: PHENOMENOLOGICAL IMPACT



- N³LO PDF corrections to **Higgs in gluon fusion small**, with a 1.5% suppression wrt NNLO PDFs
- Good perturbative convergence at N³LO also for **quark-initiated processes**



<u>NEW: COMBINED NNPDF AN3LO + MHOU +QED PDF SETS</u>

Fit ID	Perturbative accuracy	Theory cov. mat.
NNPDF40_nnlo_as_01180_qed_mhou	NNLO _{QCD} NLO _{QED}	MHOU _{7pt}
NNPDF40_an3lo_as_01180_qed	aN ³ LO _{QCD} ⊗NLO _{QED}	IHOU+MHOU _{3pt}
NNPDF40_an3lo_as_01180_qed_mhou	aN ³ LO _{QCD} ⊗NLO _{QED}	IHOU+MHOU _{7pt}



Barontini et al arXiv:2406.01779

- Qualitative impact of QED corrections on quark and gluon is the same in the NNLO and aN3LO fits.
- Largest impact associated to the gluon, where QED effects lead to overall decrease of ~1%
- For the quark PDFs, the impact of QED effects in the aN3LO fit is small but consistent with the results the corresponding NNLO fits





NEW: NNPDF40 LO, NLO AND NNLO SETS FOR MC EVENT GENERATORS



J. Cruz-Martines et al arXiv:2406.12961

 NNPDF4.0MC PDFs satisfy the requirements of event generators (nonnegative down to $Q \sim 1$ GeV, smooth extrapolation to very small-x and Q, fast growing gluon at small-x, photon PDF included and perturbatively generated heavy quark PDFs) at LO, NLO, NNLO.

 Available with various settings to be matched to different Monte Carlo event generators

)130_qed	[27]	$QCD_{LO} \otimes QED_{LO} TRN (1.0 GeV)$	$g, q_i, \bar{q}_i > 0 \ (1 \text{ GeV})$	0.130
01180	[32]	QCD_{LO} TRN (1.65 GeV)	$g, q_i, \bar{q}_i > 0 \ (1.65 \text{ GeV})$	0.118
_as_01180	[32]	QCD_{LO} TRN (1.65 GeV)	$g, q_i, \bar{q}_i > 0$ (1 GeV)	0.118
s_01180	t.w.	QCD_{LO} TRN (1.0 GeV)	$g, q_i, \bar{q}_i > 0$ (1 GeV)	0.118
s_01180_qed	t.w.	$QCD_{LO} \otimes QED_{LO} EXA (1.0 GeV)$	$g, q_i, \bar{q}_i > 0 \ (1 \text{ GeV})$	0.118
01180	[32]	QCD{NLO} TRN (1.65 GeV)	$g, q_i, \bar{q}_i > 0 \ (\sqrt{5} \text{ GeV})$	0.118
n_as_01180	[32]	QCD_{NLO} TRN (1 GeV)	$g, q_i, \bar{q}_i > 0 \ (\sqrt{5} \text{ GeV})$	0.118
as_01180	t.w.	QCD_{NLO} TRN (1 GeV)	$g, \Sigma > 0$ (1 GeV)	0.118
			$q_i, \bar{q}_i > 0 \ (\sqrt{5} \text{ GeV})$	







PART II: A ROADMAP TO TEST ACCURACY

THE PRECISION VERSUS ACCURACY CHALLENGE



Challenges

- Inconsistency or tension in data of experimental origin (underestimate of systematics...)
- Inaccuracy in theoretical framework
 - Missing higher order uncertainties (QCD, EW)
 - Other corrections (nuclear, higher-twist, non-perturbative effects...)
- Fitting away possible BSM signals

Deficiencies in fitting methodology (data-driven parametrisation change, optimisation issues, overfitting...)

10/19

CLOSURE TEST: A TOOL TO TEST METHODOLOGY AND THEORY

 Closure tests for <u>data region</u>: imagine we knew the law of Nature **f**: is our fitting methodology able to reproduce it? Is the uncertainty faithful? Statistical validation of PDF uncertainties can be performed via closure tests.



11/19

CLOSURE TEST: A TOOL TO TEST METHODOLOGY AND THEORY

- Closure tests for <u>data region</u>: imagine we knew the law of Nature f: is our fitting methodology able to reproduce it? Is the uncertainty faithful? Statistical validation of PDF uncertainties can be performed via closure tests.
- What happens if experimentalists underestimated some systematics? Example: Build in the closure test an complete underestimate of ATLAS jets systematics, check effects on CMS jets observables (if out-of-sample) and gluon PDFs



Diagnostic tools: determine statistical indicators such as ratio bias-variance of under (>1) or over (<1) estimated uncertainties allow to check effect of experimental inconsistencies on datasets included and on the PDFs





<u>CLOSURE TEST: A TOOL TO TEST METHODOLOGY AND THEORY</u></u>

- Imagine that on top of the "true" PDFs one inject the "true" NP model in the pseudo-data
- Generate HL-LHC pseudo-data assuming "true" law of nature = "true" PDFs + "true" UV model
- Fit PDFs assuming SM
- Can PDFs absorb signs of new physics?



Hammou, Madigan, Mangano, Mantani, Moore, Morales, MU 2307.10370



Test possible New Physics contamination in PDF fits

HL-LHC HM DY 14 TeV - charged current - e ch.



Max contamination allowed by global fit Without spoiling X²



TEST GENERALISATION OF PDF AND EXTRAPOLATION

- test all global PDF sets agains new precise data from LHC Run I and Run II data & DIS HERA jets data. How well do various PDF sets describe data that are <u>not yet</u> included in the fit?
- Remarks:
 - → All results are NNLO (no k-factor approximation) thanks to PineAPPL, NNLOJET, MATRIX and Ploughshare



Carrazza et al: 2008.12789

Gehrmann-De Ridder et al: 1507.02850, 1605.04295...

and over the entire dataset / HEPdata entry

$$\chi^{2} = \sum_{i,j=1}^{N_{\text{dat}}} \left(T_{i}^{(0)} - D_{i} \right) \left(\text{cov}^{-1} \right)_{ij} \left(T_{j}^{(0)} - D_{j} \right)$$

• Future tests help to discriminate among PDF sets [J. Cruz-Martinez et al, Acta Phys.Polon.B 52 (2021) 243 - on Run I] [Chiefa et al, in progress]:







NNLO PDF sets considered: PDF4LHC15, PDF4LHC21, ABMP16, CT18, CT18A, CT18Z, MSHT20, NNPDF3.1, NNPDF4.0 \rightarrow The computation of X² is always shown considering as uncertainties either (exp) or (exp + mho) or (exp + mho + pdf)

$$\chi^2_{
m exp} \leftrightarrow
m cov =
m cov_{
m exp}$$

 $\chi^2_{
m exp+mho} \leftrightarrow
m cov =
m cov_{
m exp} +
m cov_{
m mho}$
 $\chi^2_{
m exp+mho+pdf} \leftrightarrow
m cov =
m cov_{
m exp} +
m cov_{
m mho} +
m cov_{
m pd}$









THE EXPERIMENTAL DATA IN THE TEST SET

Sector	Exp.	$\sqrt{\mathrm{s}}$ (TeV)	Channel	Observable	$\mathcal{L} \; (\mathrm{fb}^{-1})$	$\mathbf{N}_{ ext{dat}}$
W, Z	ATLAS	13	Z p_T spectrum	$rac{d\sigma}{dp_T^Z}$	36.1	10
	ATLAS	8	Z incl. prod.	$rac{d\sigma}{d y_{ll} }$	20.2	7
	CMS	8	W incl. prod.	$rac{d\sigma_{W^{\pm}}}{d\eta_l}$	35.9	36
	LHCb	13	Z incl. forward prod.	$rac{d\sigma_Z}{dy_Z}$	5.1	17
top	ATLAS	13	hadronic	$\left(rac{1}{\sigma} ight) rac{d\sigma}{dm_t ar{t}}, \ rac{d\sigma}{d y_t ar{t} }, \ rac{d^2\sigma}{d y_t ar{t} dm_t ar{t}}$	36.1	9,12,11
	ATLAS	13	$\ell + \mathrm{jets}$	$\left(\frac{1}{\sigma}\right)\frac{d\sigma}{dm_{t\bar{t}}}, \ \frac{d\sigma}{dp_{T,t}}, \ \frac{d\sigma}{d y_t }, \ \frac{d^2\sigma}{d y_{t\bar{t}} }$	36.1	9,8,5,7
	CMS	13	$\ell + \mathrm{jets}$	$\left(\frac{1}{\sigma}\right)\frac{d\sigma}{dm_{t\bar{t}}},\ \frac{d\sigma}{dp_{T,t}},\ \frac{d\sigma}{d y_{t\bar{t}} },\ \frac{d\sigma}{d y_{t\bar{t}} },\ \frac{d\sigma}{d y_{t\bar{t}} }dm_{t\bar{t}}$	137	15, 16, 10, 11, 35
jets	ATLAS	13	incl. jet $R=0.4, 0.7$	$rac{d^2\sigma}{dp_{T,j}d y_j }$	3.2	177
	ATLAS	13	di-jets R=0.4	$rac{d^2\sigma}{dm_{jj}d\Delta y}$	3.2	136
	CMS	13	incl. jets R=0.4, 0.7	$\frac{\frac{d^2\sigma}{dp_{T,j}d y_j }}{\frac{d^2\sigma}{dp_{T,j}d y_j }}$	3.2	78
DIS jets	H1	0.319	incl. jet (low q^2)	$rac{d^2\sigma}{dq^2dpT}$	0.29	48
	H1	0.319	di-jets (low q^2)	$rac{d^2\sigma}{dq^2d\langle pT angle}$	0.29	48
	H1	0.319	incl. jet (high q^2)	$rac{d^2\sigma}{dq^2dpT}$	0.351	24
	H1	0.319	di-jets (high q^2)	$rac{d^2\sigma}{dq^2d\langle pT angle}$	0.351	24
	ZEUS	0.3	incl. jet	$rac{d^2\sigma}{dE_Tdq^2}$	0.038	30
	ZEUS	0.319	incl. jet	$rac{d^2\sigma}{dE_Tdq^2}$	0.082	30
	ZEUS	0.319	d-jets	$rac{d^2\sigma}{dE_T dq^2}$	0.374	22

Chiefa et al, in progress





THE TOP SECTOR

ATLAS top pair production at 13 TeV, I+jet channel, (36.1fb⁻¹) ATLAS collaboration [arXiv:1908.07305]



 $\chi^2_{exp + mho + pdf}$ $\chi^2_{exp + mho}$

ABMP16	$0.832 \ (0.835)$
CT18	1.483 (2.504)
CT18A	1.489(2.347)
MSHT20	$1.585\ (2.046)$
NNPDF3.1	$1.200\ (1.244)$
NNPDF4.0	$1.297\ (1.338)$
PDF4LHC15	$1.298\ (2.088)$
PDF4LHC21	$1.577 \ (2.068)$

Absolute comparison, normalized & size of PDF (solid) and dashed (MHO) Δ compared to data Δ. Shaded band includes MHO (9 pts variation) and PDF uncertainties added in quadrature.



16/19

THE JET SECTOR

CMS inclusive jet cross-sections at 13 TeV (33.5fb-1) anti-kT, R=0.7 CMS collaboration [arXiv:2111.10431]



2.15(3.80)2.21(3.62)

PDF4LHC15

PDF4LHC21

variation) and PDF uncertainties added in quadrature.



Leading color NNLO correction Predictions: NNLOJET (plougshare) Hepdata: 10.17182/hepdata.115022.v2



THE DRELL-YAN SECTOR

LHCb Z forward production at 13 TeV (5.1fb-1) LHCb collaboration [arXiv:2112.07458]





CONCLUSIONS AND OUTLOOK

- In an era of precision at LHC, need precise and accurate PDFs
- NNPDF: lots of progress in several main and side projects, NNPDF4.0 aN3LO, NNPDF4.0 MHOU, towards NNPDF4.1. Stay tuned!
- How to make progress on the precision vs accuracy challenge?
 - 1. Public codes ensure reproducibility
 - These would assess faithfulness of central values and uncertainties of each PDF fits.
 - generalisation and extrapolation of PDFs.
 - 4. frontiers that should be explored from multiple angles

NNPDF4.0 QED, NNPDF4.0 for MC event generators, methodological studies, EW corrections, more data

2. Closure tests (now explored also by MSHT collaboration): a coordinated effort of PDF fitting collaboration should help converging on agreed set of tests (Level 0, Level 1, Ratio Bias to Variance). 3. Tests on comprehensive set of data not yet in the global PDF fits are only possible discriminant of

Effects of possible experimental inconsistencies and even possible effects of new physics in the high energy tails, definition of conservative PDF sets, simultaneous PDFs and SM parameter fits the new

THANK YOU FOR YOUR ATTENTION





EXTRA MATERIAL

THE MCSCALES APPROACH



✓ Main idea of MCscales: the renormalisation and factorisation scales are free parameters of the fixed-order theory, that induce an uncertainty on the theory predictions included in a PDF fit & need to be propagated ✓ Joint sampling of experimental uncertainty (propagated to PDF uncertainty by MC sampling) by specifying a suitable prior probability distribution of all possible scale choices & a-posteriori criterion based on agreement with the data.

$$P\left(k_{f} = \xi_{f}, k_{r_{1}} = \xi_{1}, \dots, k_{r_{N_{p}}} = \xi_{N_{p}}\right) = P(\omega)$$

with $\omega \in \Omega = \{(\xi_{f}, \xi_{1}, \dots, \xi_{N_{p}}) \forall \xi_{f}, \xi_{1}, \dots, \xi_{N_{p}} \in \Xi\}$
 $3^{1+N_{p}}$ elements, with N_p = 5, p=DIS NC, DIS C

Choose prior = choose $P(\omega)$ Posterior $\chi_n^2 > \left\langle \chi^2 \right\rangle_{n \, | \, \omega^{(n)} = \{1, \dots, 1\}} + 4 \operatorname{std}(\chi^2)_{n \, | \, \omega^{(n)} = \{1, \dots, 1\}}$

C, DY, JET, TOP





THE MCSCALES APPROACH

✓ Can compute full PDF+SCALE uncertainty in cross sections at NLO by matching the scales in the hard cross section computation with the scales in the MCscale PDF set: correlation fully taken into account



$$\left\{\sigma_n = \hat{\sigma}_p(k_f^{(n)}, k_{r_p}^{(n)}) \otimes f_n(k_f^{(n)}, k_{r_p}^{(n)}) \; \forall n = 1, \dots, N\right\}$$





THE MCSCALES APPROACH



✓ Can look at the distribution of each of the scales over replicas.

✓ Flat distribution for the MCscales uniform prior.

 ✓ After applying postfit observe preference for central factorisation scale.

✓Each process affected in a different way.

Scale multipliers	Dragona	Droformad re
Scale multipliers	Process	Preferred va
(k_f,k_r)	DIS CC	(1,1)
	DIS NC	(1,2)
	DY	(1,1)
	Jets	$(1, \frac{1}{2})$
	Top	$(1,\overline{1})$

