The Structure of the Proton to One-Percent Accuracy: NNPDF4.0

Juan Rojo (VU Amsterdam & Theory group, Nikhef)

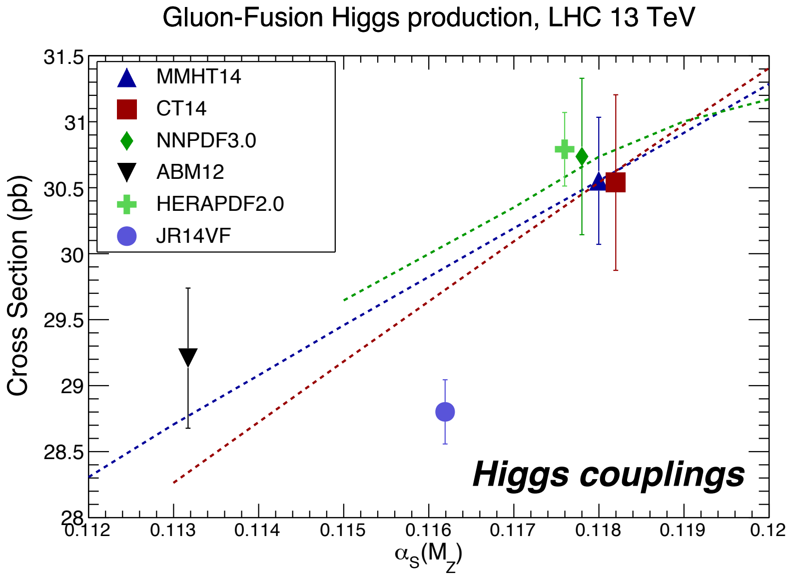
on behalf of:

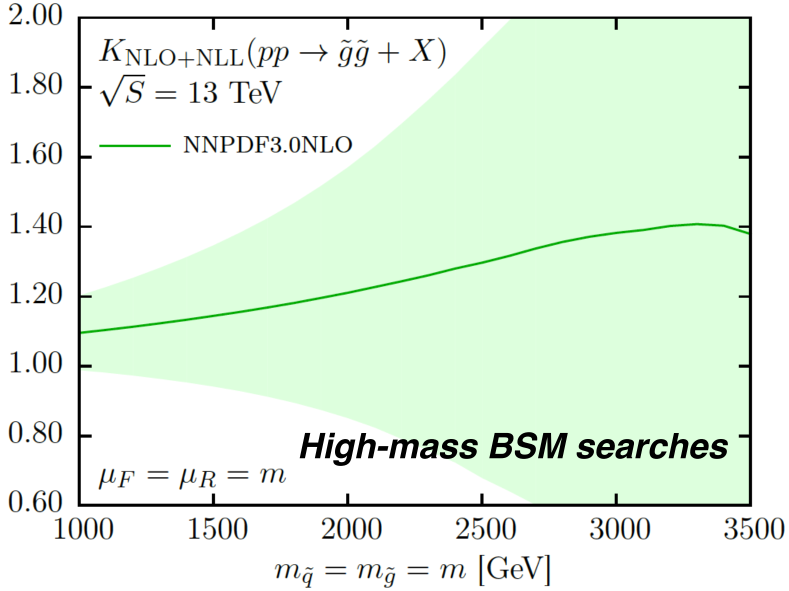
THE NNPDF COLLABORATION & N3PDF TEAM

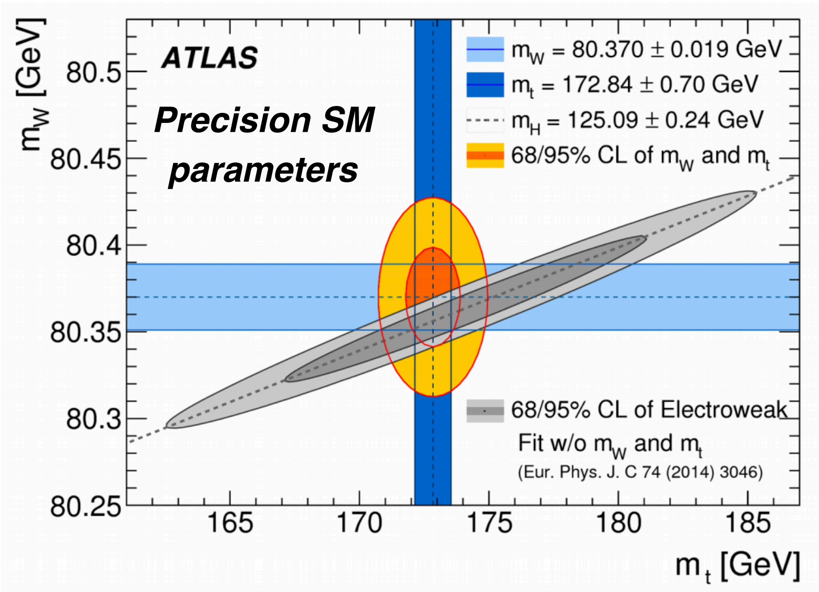
AMSTERDAM-BARCELONA-CAMBRIDGE-EDINBURGH-INFN-MILAN-NIKHEF

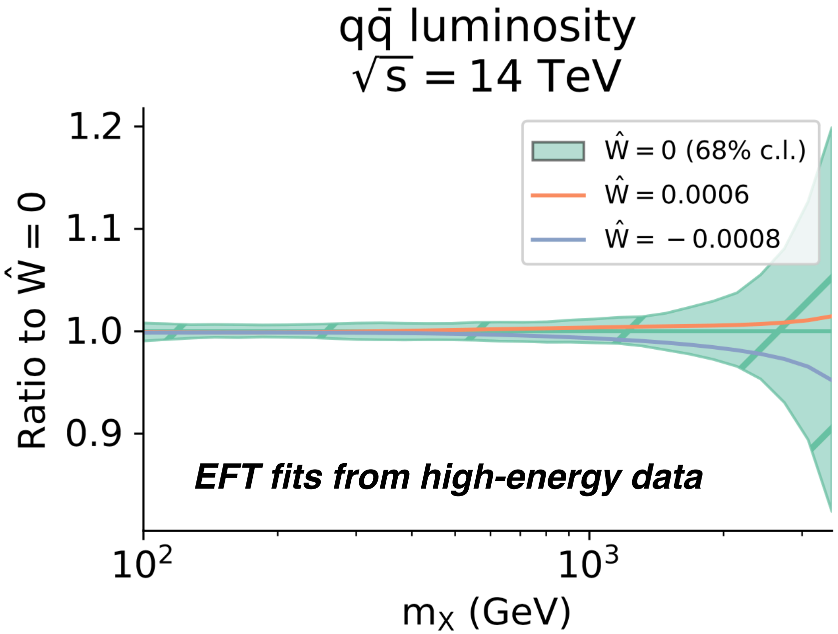
Why we need better PDFs?

PDF uncertainties: limiting factor in **theoretical interpretation** for many LHC analysis









Juan Rojo

The road towards NNPDF4.0

Collaborative progress towards extending data, theory and methodology

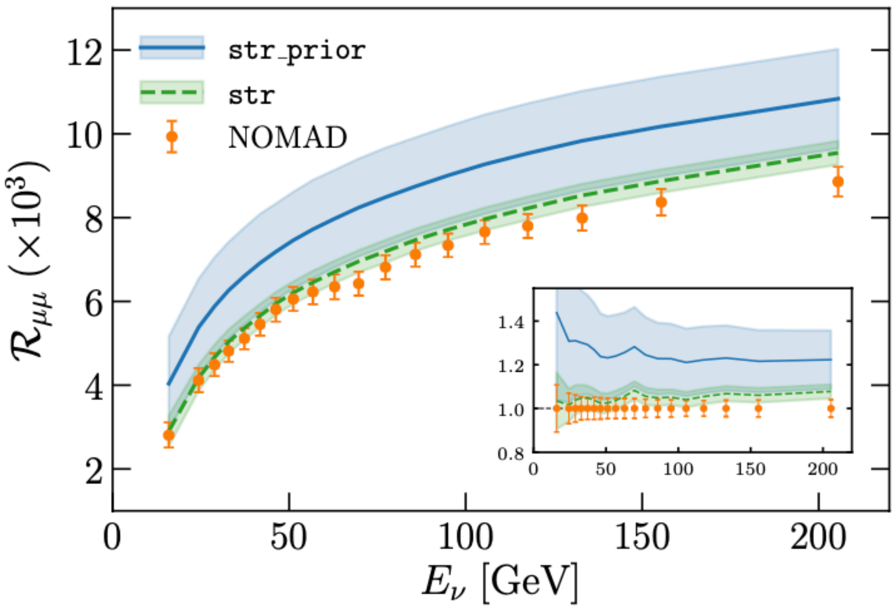
06/2017	**NNPDF3.1**		[EPJ C77 (2017) 663]
10/2017	NNPDF3.1sx: PDFs with small-x resummation		[EPJ C78 (2018) 321]
12/2017	NNPDF3.1luxQED: consistent photon PDF à la luxQED		[SciPost Phys. 5 (2018) 008]
02/2018	NNPDF3.1+ATLASphoton: inclusion of direct photon data		[EPJ C78 (2018) 470]
12/2018	NNPDF3.1alphas: α_s from a correlated-replica method		[EPJ C78 (2018) 408]
12/2018	NNPDF3.1nuc: heavy ion nuclear uncertainties in a fit		[EPJ C79 (2019) 282]
05/2019	NNPDF3.1th: missing higher-order uncertainties in a fit		[EPJ C79 (2019) 838; ibid. 931]
07/2019	Gradient descent and hyperoptimisation in PDF fits		[EPJ C79 (2019) 676]
12/2019	NNPDF3.1singletop: inclusion of single top t-channel data		[JHEP 05 (2020) 067]
05/2020	NNPDF3.1dijets: comparative study of single- and di-jets		[EPJ C80 (2020) 797]
06/2020	Positivity of $\overline{\mathrm{MS}}$ PDFs		[JHEP 11 (2020) 129]
08/2020	PineAPPL: fast evaluation of EW$\times$QCD corrections		[JHEP 12 (2020) 108]
08/2020	NNPDF3.1strangeness: assessment of strange-sensitive data		[EPJ C80 (2020) 1168]
11/2020	NNPDF3.1deu: deuteron uncertainties in a fit		[EPJ C81 (2021) 37]
03/2021	Future tests		[arXiv:2103.08606]
2021	**NNPDF4.0**		[to appear]

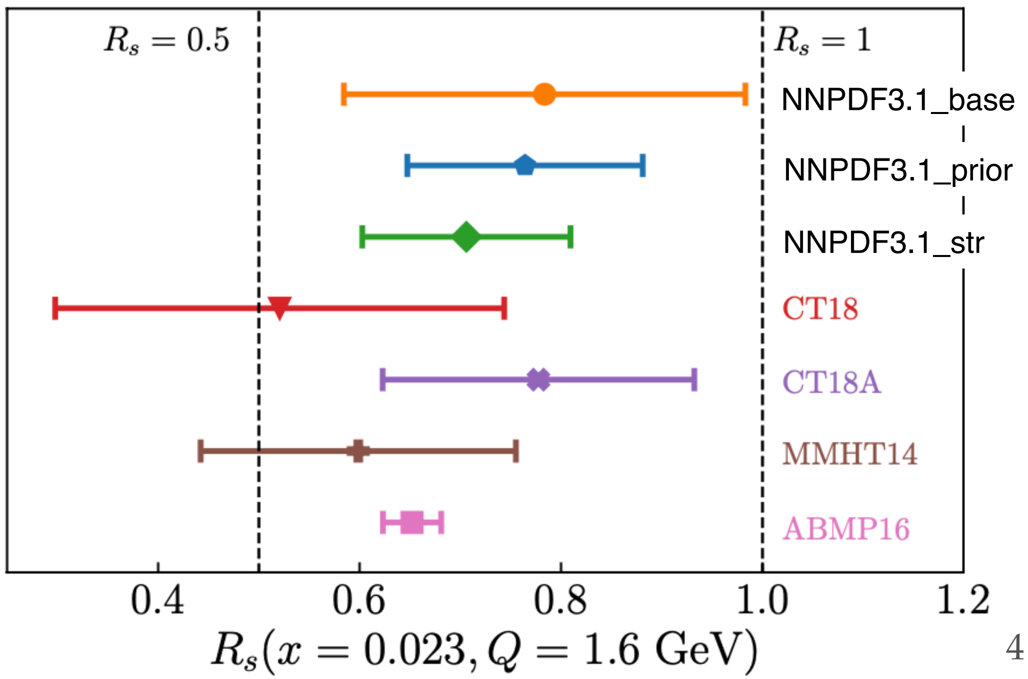
More details in the NNPDF talks at the PDF4LHC meeting 03/21

The strangest proton?

Reappraisal of the proton strangeness based combination of **all relevant experimental inputs**

Process	Dataset	$n_{\rm dat}$	$\chi^2_{\rm base}$	$\chi^2_{\rm pr}$	$\chi^2_{\rm str}$
νDIS ($\mu\mu$)		76/76/95/91/95	0.76	0.71	0.53
	NuTeV [9]	76/76/76/76/76	0.76	0.71	0.53
	NOMAD [10]	—/—/19/15/19	[9.3]	[8.8]	0.55
W, Z (incl.)		391/418/418/418/418	1.45	1.40	1.40
	ATLAS [12]	34/61/61/61/61	1.96	1.65	1.67
W+c		—/37/37/37/37	[0.73]	0.68	0.60
	CMS [17, 18]	—/15/15/15/15	[1.04]	0.98	0.96
	ATLAS [16]	—/22/22/22/22	[0.52]	0.48	0.42
W+jets	ATLAS [15]	—/32/32/32/32	[1.58]	1.18	1.18
Total		3981/4077/4096/4092/4096	1.18	1.17	1.17

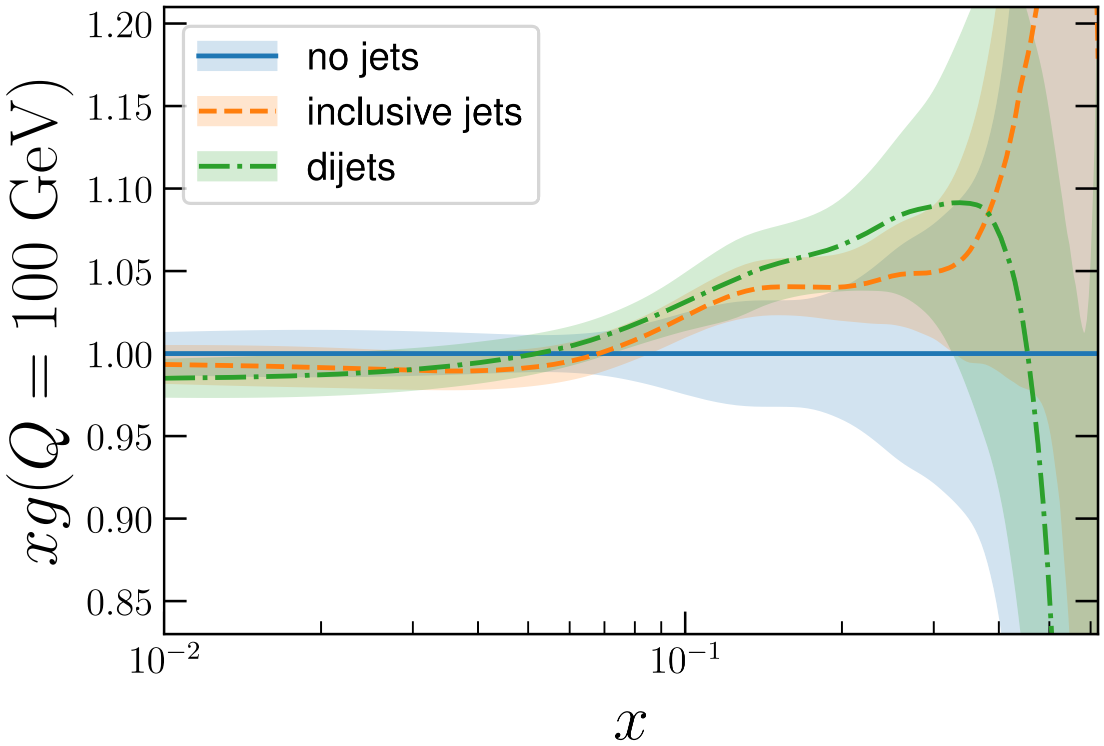




- Satisfactory simultaneous description of all datasets
- No **evidence for tension** between datasets or groups of processes
- Sizeable constraints from **NOMAD neutrino DIS** data, consistent with collider data
- Preference for a **moderately suppressed strangeness**

$$R_S(x = 0.023,\ Q = 1.6\ \text{GeV}) = 0.71 \pm 0.10$$

LHC dijets to map the gluon PDF



- ☑ With NNLO QCD (+ NLO EW) calculations, **all available dijet cross-sections** are successfully included in global fit, good agreement with data

  first use of dijets in PDF fits!

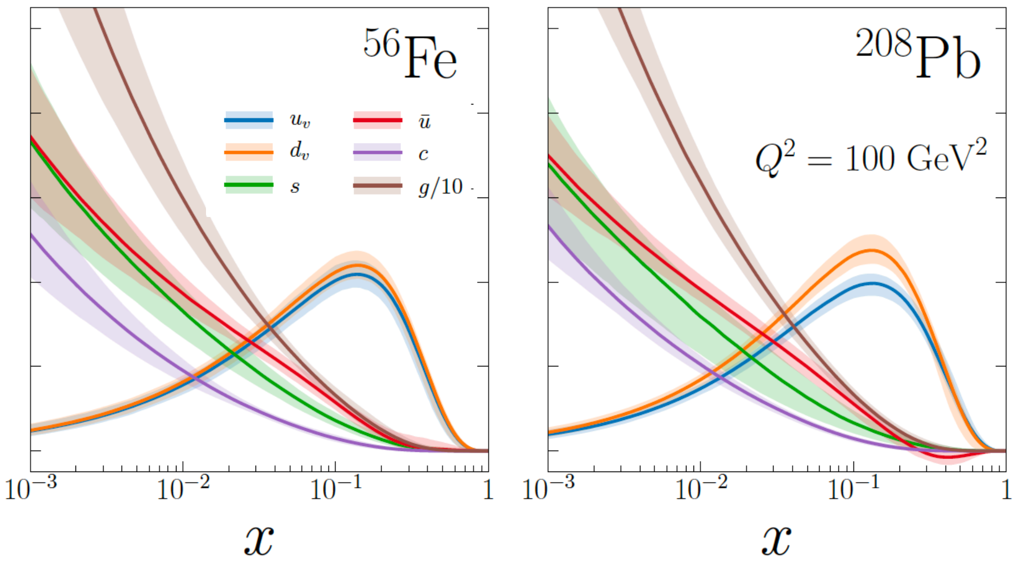
- ☑ Constrains on the gluon **qualitatively consistent** between dijet and inclusive jet observables

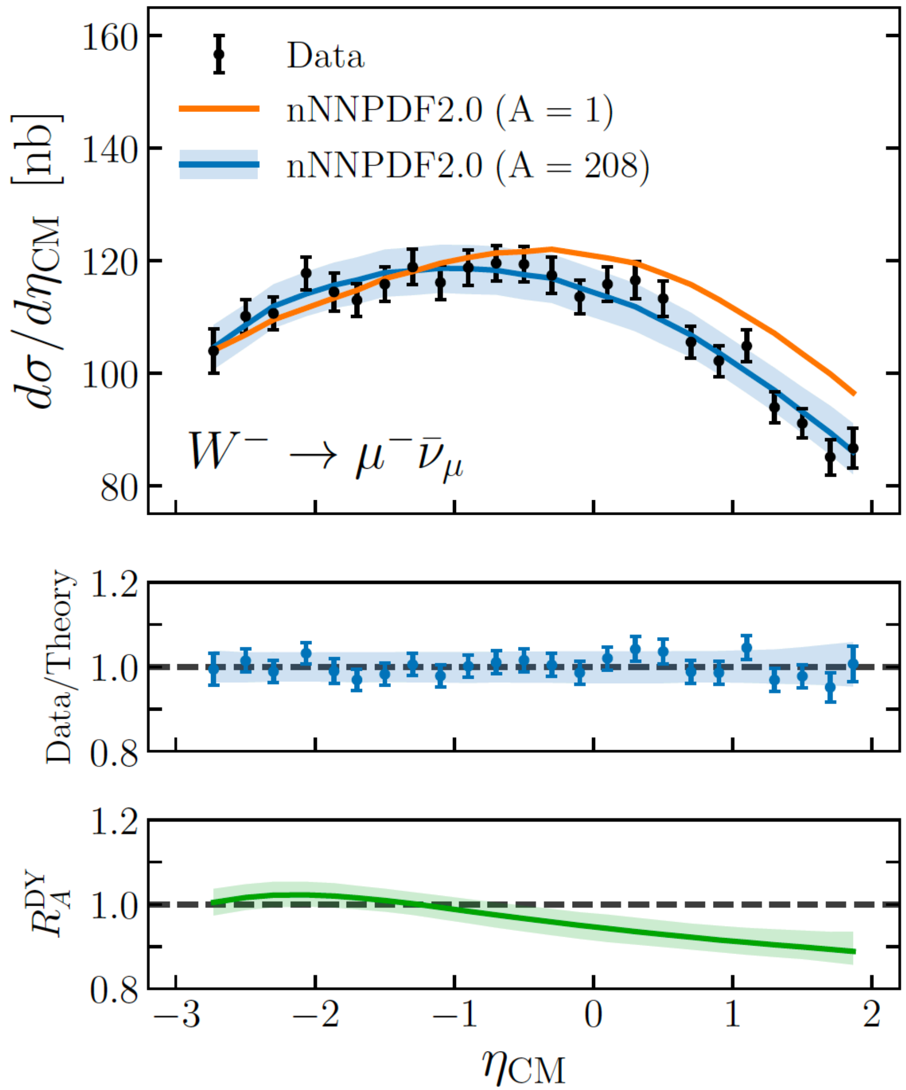
- ☑ Inclusive jets requires introducing **tailored decorrelation** models to achieve **good χ^2**

Experiment	Measurement	$\sqrt{s}$ [TeV]	$\mathcal{L}$ [fb^{-1}]	R	Distribution	$n_{\rm dat}$
ATLAS	Inclusive jets	7	4.5	0.6	$d^2\sigma/dp_T d\lvert y\rvert$	140
CMS	Inclusive jets	7	4.5	0.7	$d^2\sigma/dp_T d\lvert y\rvert$	133
ATLAS	Inclusive jets	8	20.2	0.6	$d^2\sigma/dp_T d\lvert y\rvert$	171
CMS	Inclusive jets	8	19.7	0.7	$d^2\sigma/dp_T d\lvert y\rvert$	185
ATLAS	Dijets	7	4.5	0.6	$d^2\sigma/dm_{jj} d\lvert y^*\rvert$	90
CMS	Dijets	7	4.5	0.7	$d^2\sigma/dm_{jj} d\lvert y_{\rm max}\rvert$	54
CMS	Dijets	8	19.7	0.7	$d^3\sigma/dp_{T,\rm avg} dy_b dy^*$	122

Quark flavour separation in nuclei from LHC data

- **nNNPDF2.0**: global determination of **nuclear PDFs** from heavy nuclear DIS (NC+CC) and LHC data
- Consistent fitting methodology with proton fits, NNPDF3.1 (and uncertainties) **reproduced for *A=1***
- Strong evidence for **quark shadowing at small-*x***

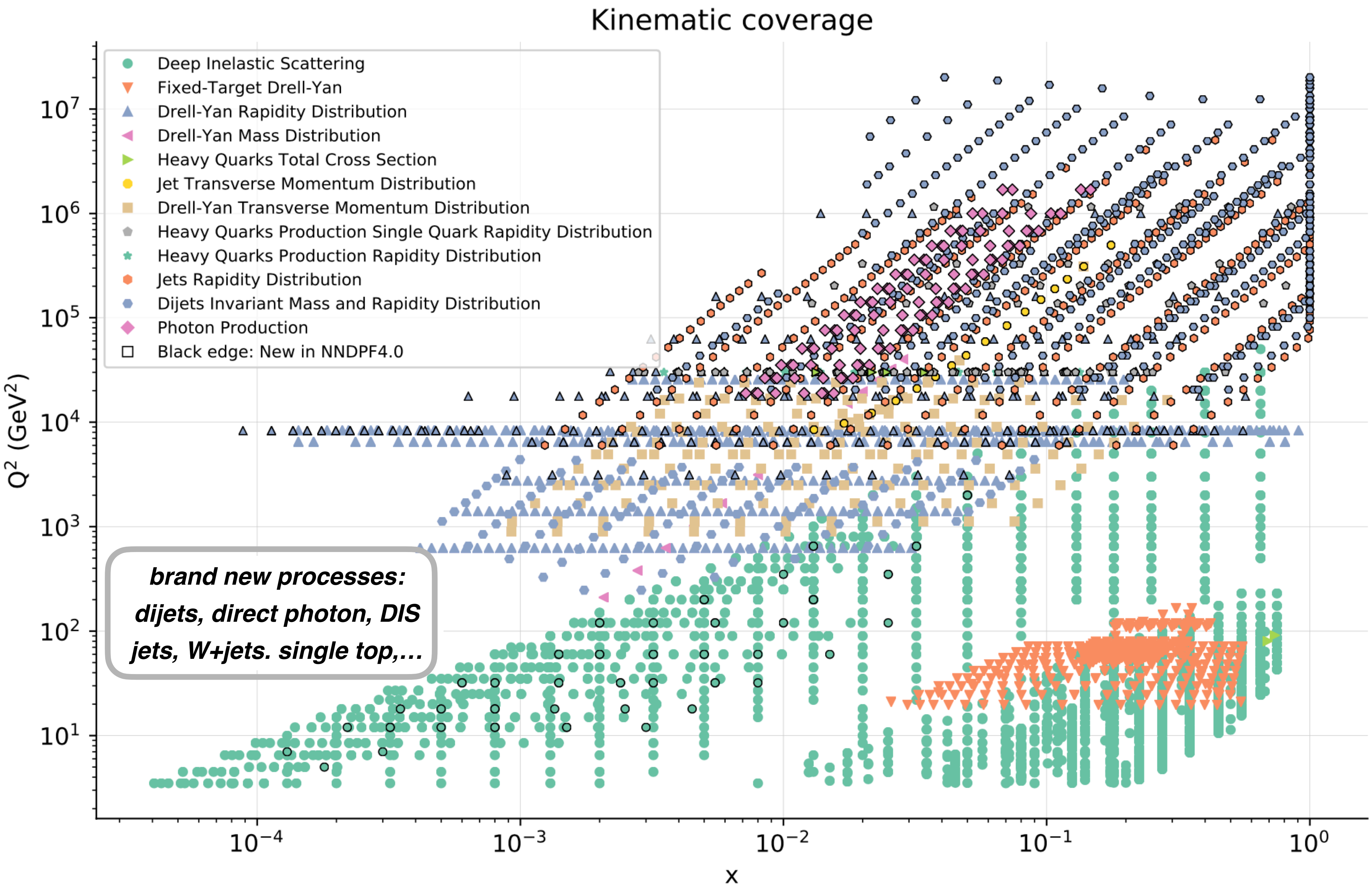




with nNNPDF2.0: consistent inclusion of **nuclear effects from neutrino DIS data** in NNPDF4.0

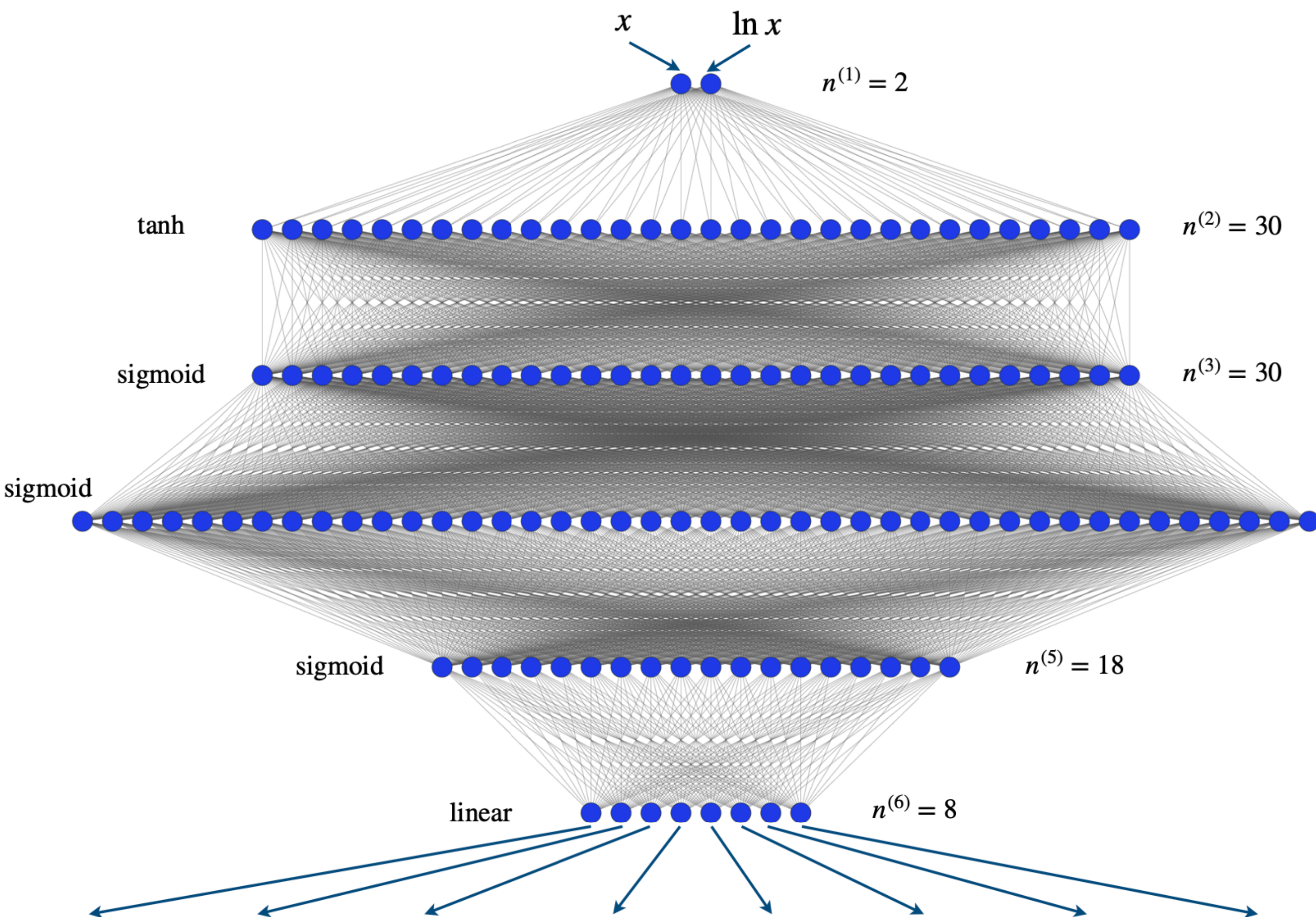
using the Ball-Pearson-Nocera approach

NNPDF4.0: data set extension



$\mathcal{O}(50)$ data sets investigated; $\mathcal{O}(400)$ data points more in NNPDF4.0 than in NNPDF3.1

Improved fitting methodology



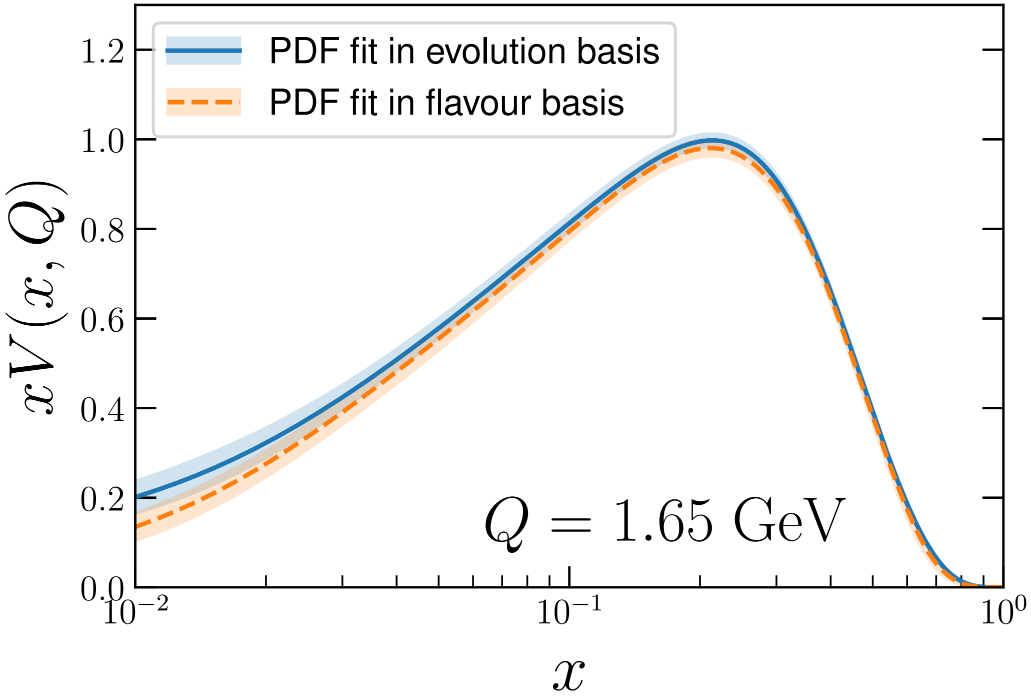
$xg(x, Q_0)$ $x\Sigma(x, Q_0)$ $xV(x, Q_0)$ $xV_3(x, Q_0)$ $xV_8(x, Q_0)$ $xT_3(x, Q_0)$ $xT_8(x, Q_0)$ $xT_{15}(x, Q_0)$ ***evolution basis***

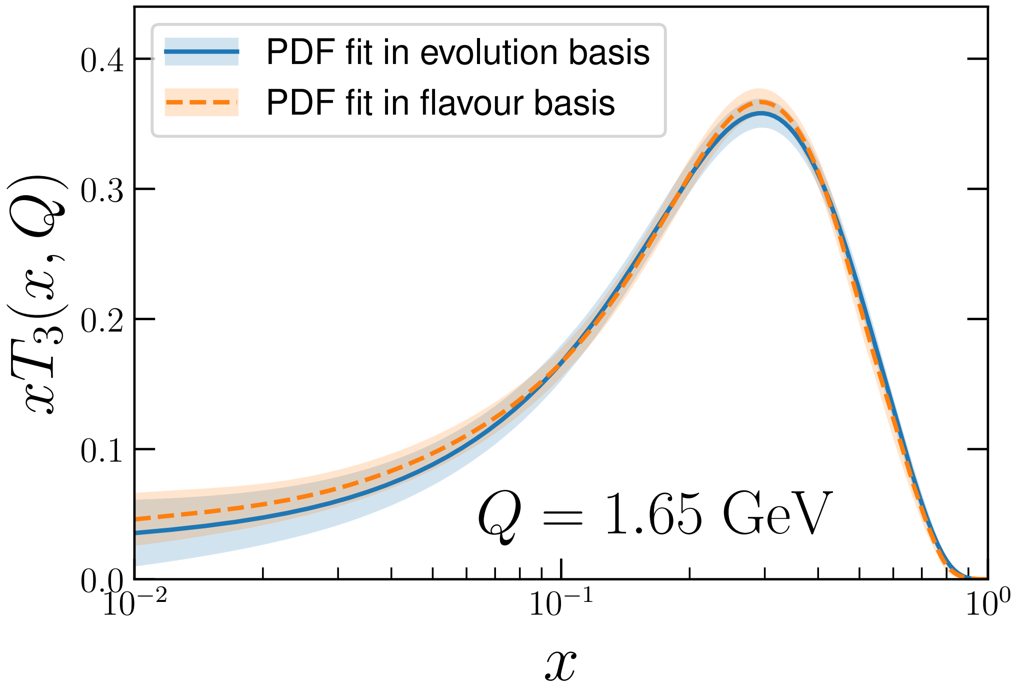
$xg(x, Q_0)$ $xu(x, Q_0)$ $x\bar{u}(x, Q_0)$ $xd(x, Q_0)$ $x\bar{d}(x, Q_0)$ $xs(x, Q_0)$ $x\bar{s}(x, Q_0)$ $xc^+(x, Q_0)$ ***flavor basis***

$$f_i(x, Q_0) = x^{-\alpha_i}(1-x)^{\beta_i}\mathrm{NN}_i(x)$$

- **Stochastic Gradient Descent** (via TensorFlow) for NN training
- Automated model **hyperparameter optimisation**
- Validation with **future tests** (forecasting new datasets) and **closure tests** (pseudo-data based on known PDFs)
- **Generative Adversarial Networks** for replica compression

Parametrisation basis independence





evolution basis PDF parametrisation:

$$xV(x, Q_0) \propto \mathrm{NN}_V(x)$$
$$xT_3(x, Q_0) \propto \mathrm{NN}_{T_3}(x)$$

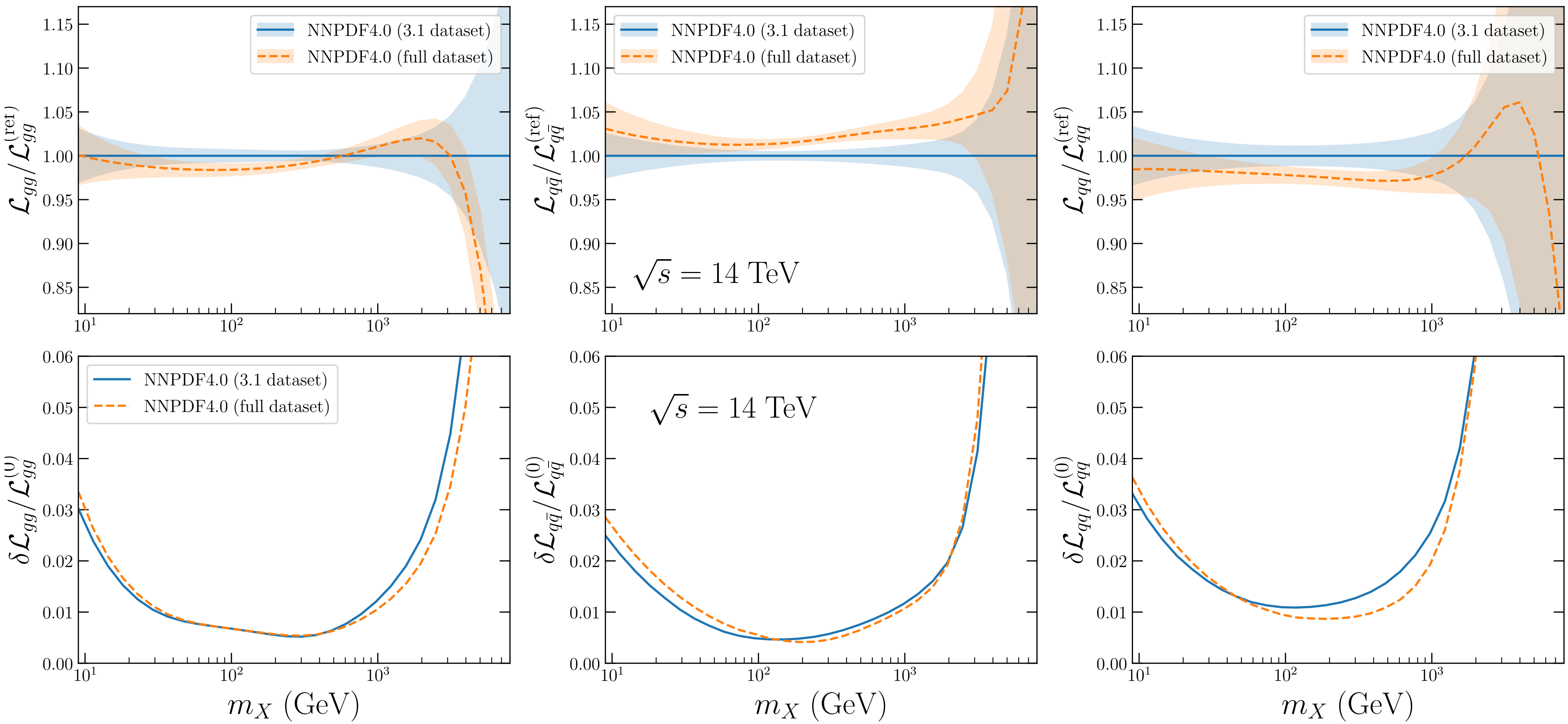
Radically different strategies to parametrize the **quark PDF flavour combinations** lead to identical results: ultimate test of **parametrisation independence**

flavour basis PDF parametrisation:

$$xV(x, Q_0) \propto \left(\mathrm{NN}_u(x) - \mathrm{NN}_{\bar{u}}(x) + \mathrm{NN}_d(x) - \mathrm{NN}_{\bar{d}}(x) + \mathrm{NN}_s(x) - \mathrm{NN}_{\bar{s}}(x)\right)$$
$$xT_3(x, Q_0) \propto \left(\mathrm{NN}_u(x) + \mathrm{NN}_{\bar{u}}(x) - \mathrm{NN}_d(x) - \mathrm{NN}_{\bar{d}}(x)\right)$$

Impact of new data

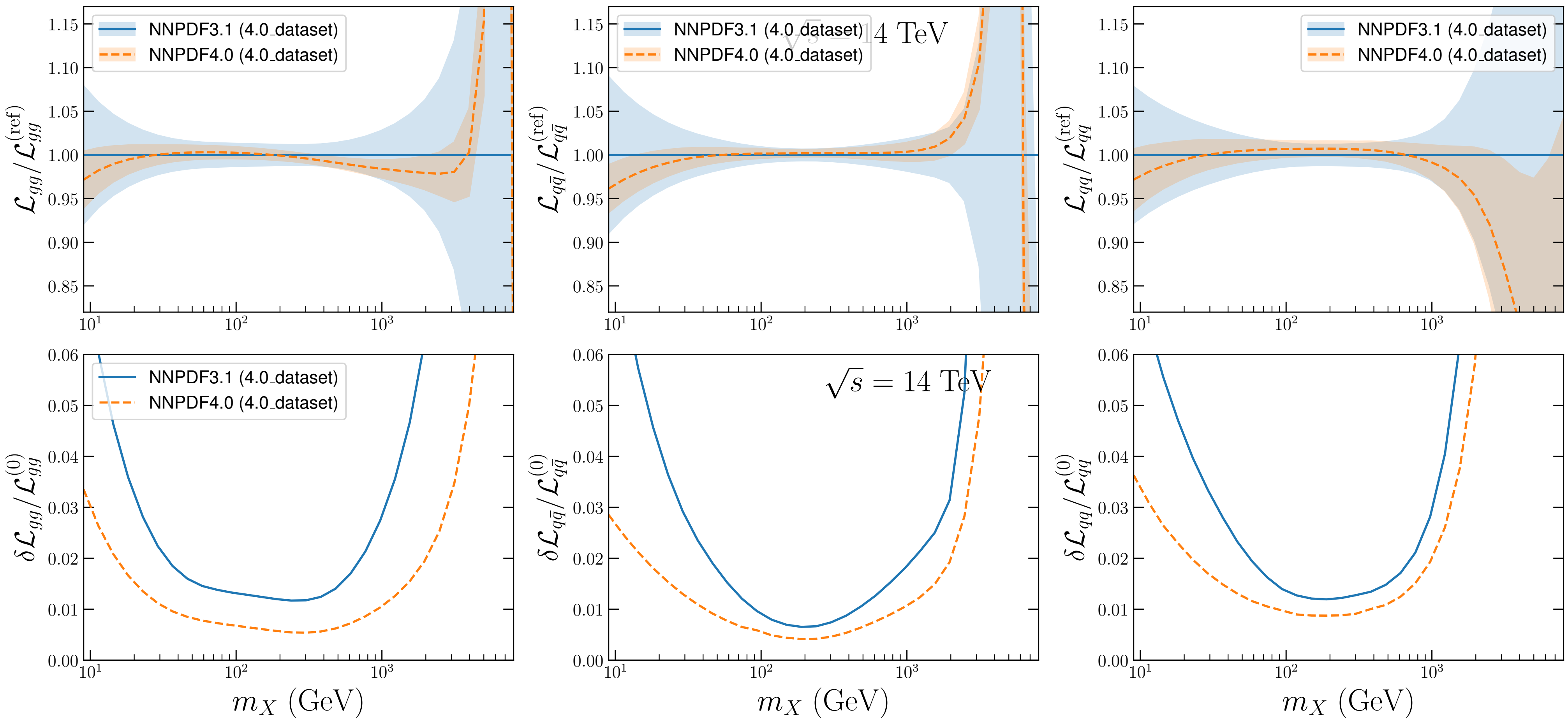
Partonic luminosities from fits based on the **same NNPDF4.0 methodology** but either with the **previous 3.1** or the **new 4.0 dataset:** good consistency, with 4.0 **more precise**



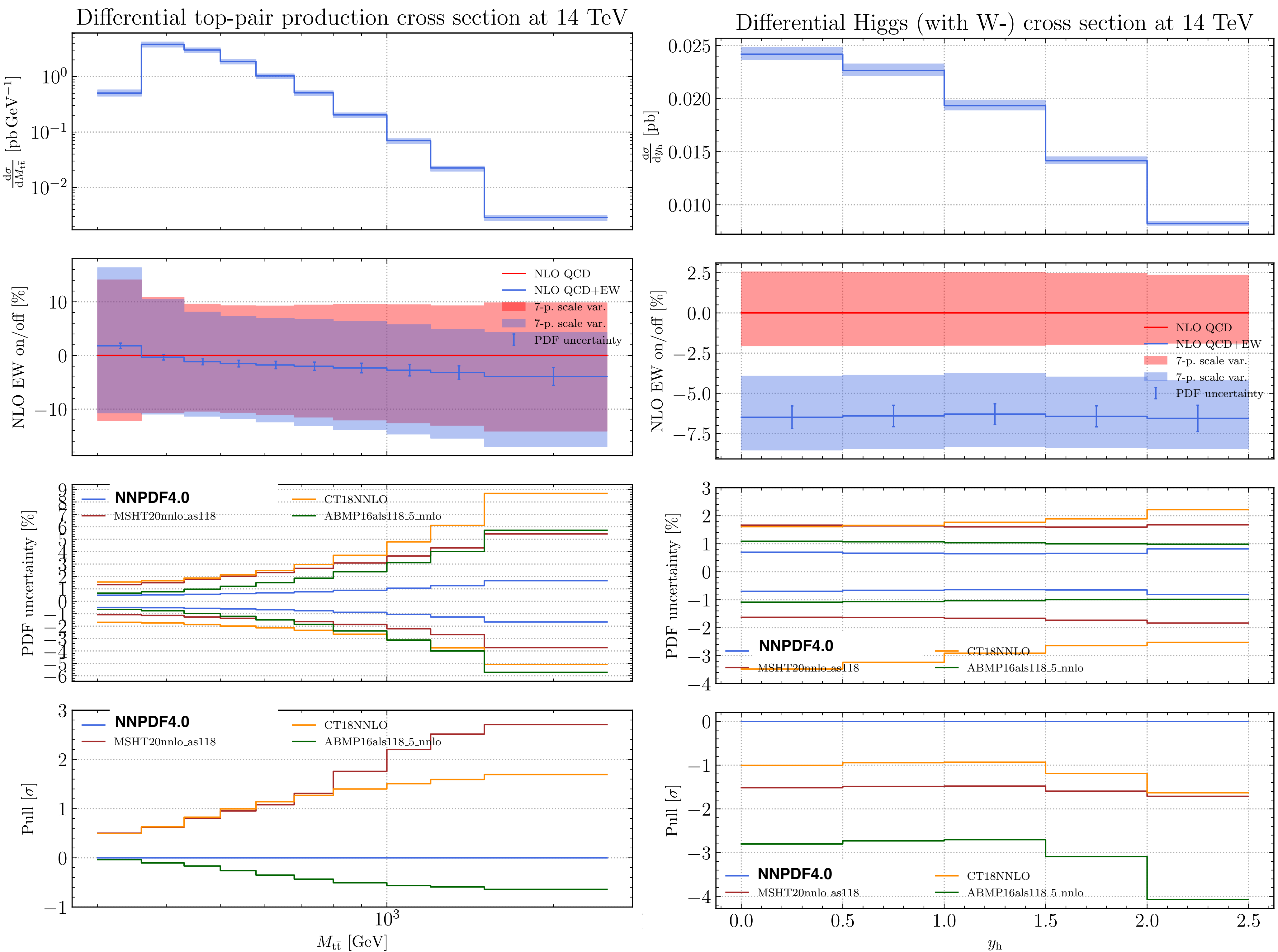
One-percent PDF uncertainties in LHC cross-sections for wide kinematic range

Impact of new fitting methodology

Partonic luminosities from fits based on the **same NNPDF4.0 dataset** but either with the **previous 3.1** or the **new 4.0 methodology:** perfect consistency, with 4.0 markedly **more accurate**

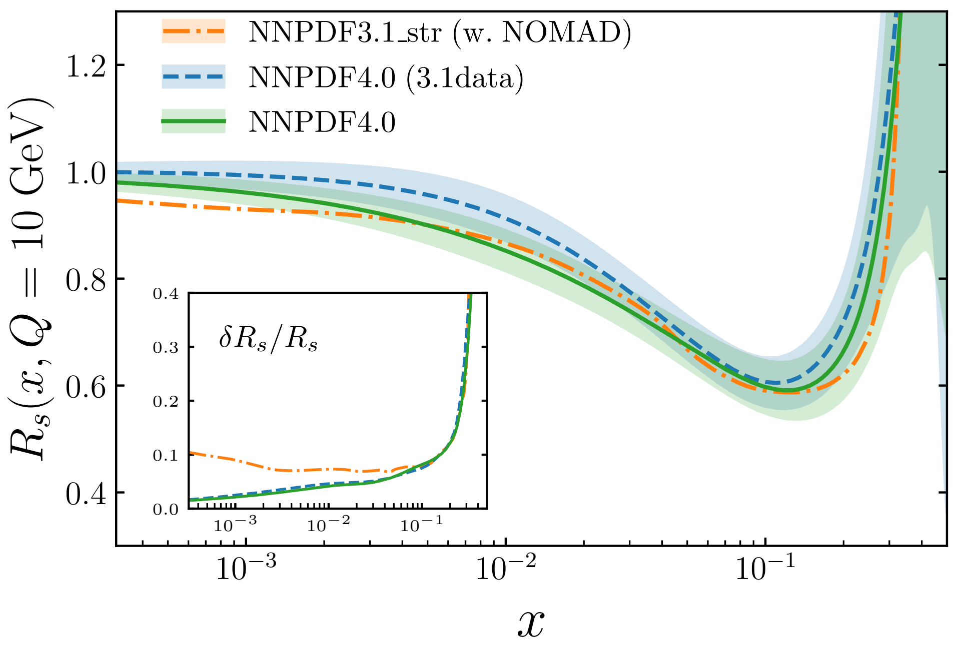


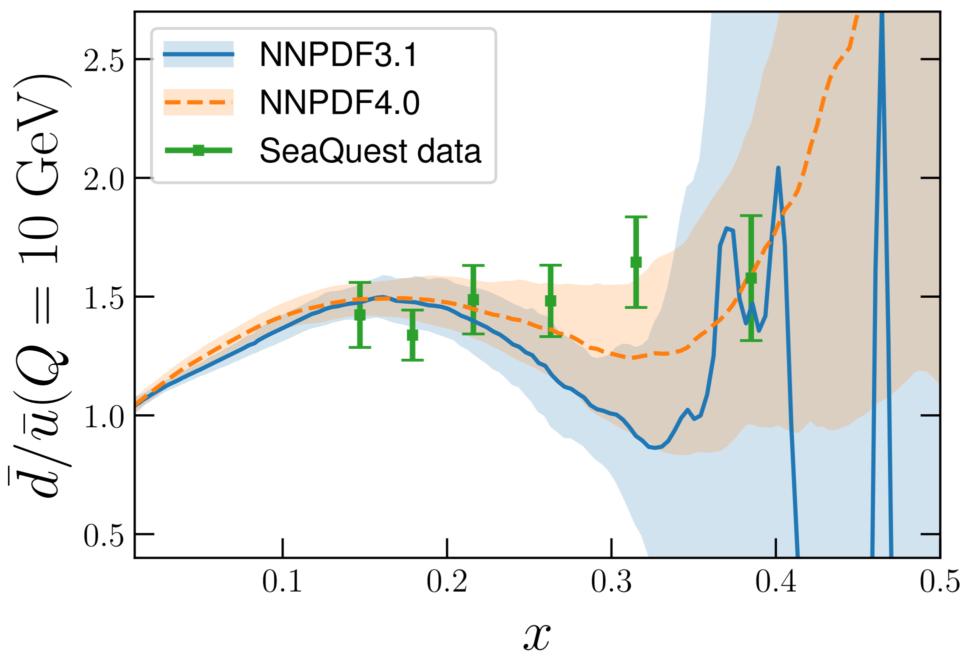
One-percent PDF uncertainties in LHC cross-sections for wide kinematic range

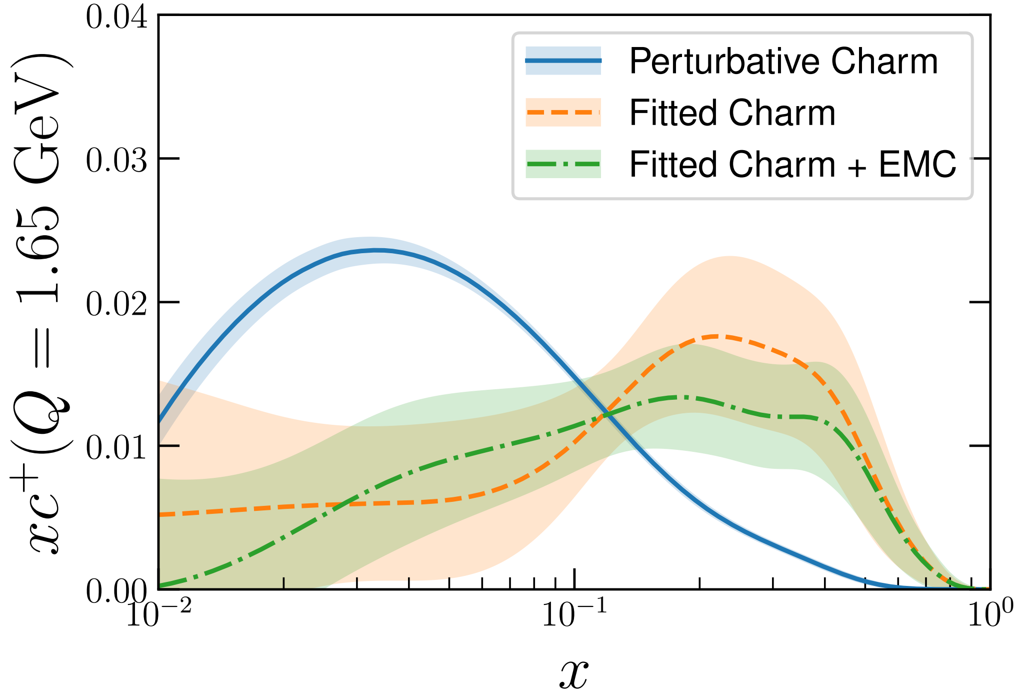


Differential top-pair production cross section at 14 TeV
$\frac{d\sigma}{dM_{t\bar{t}}}$ [pb GeV^{-1}]
NLO EW on/off [%]
NLO QCD
NLO QCD+EW
7-p. scale var.
7-p. scale var.
PDF uncertainty
PDF uncertainty [%]
NNPDF4.0
MSHT20nnlo_as118
CT18NNLO
ABMP16als118_5_nnlo
Pull [σ]
$M_{t\bar{t}}$ [GeV]
Differential Higgs (with W-) cross section at 14 TeV
$\frac{d\sigma}{dy_h}$ [pb]
y_h

From antimatter asymmetry to intrinsic charm







High-precision mapping of proton structure:

- Confirmation of **strangeness suppression factor**, full consistency with NOMAD constraints
- Excellent agreement with SeaQuest *``measurements''* of **proton antimatter asymmetry**
- Increasing evidence for **a charm component** in proton wave function

Summary and outlook

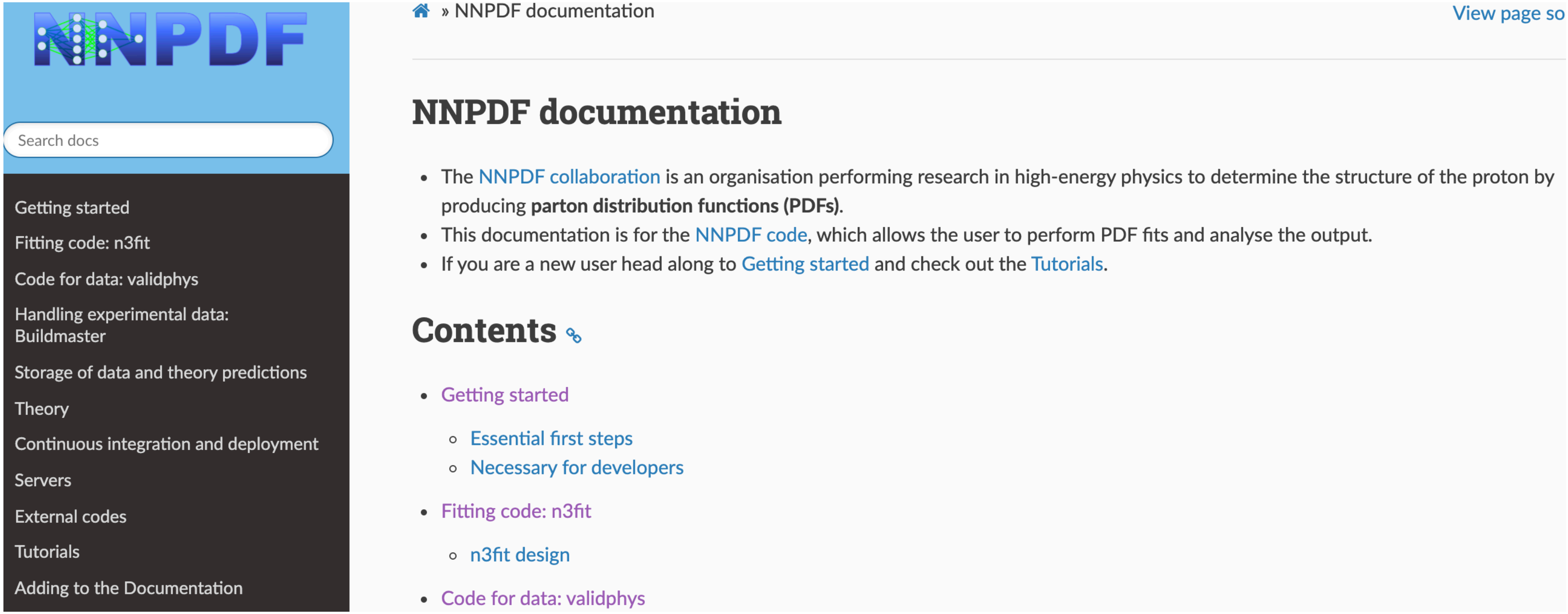
- The global NNPDF4.0 fit achieves **1% accuracy** in an unprecedentedly broad kinematic range, thanks so its **extensive dataset** combined with **deep-learning optimisation models**

- Its faithfulness in representing PDF uncertainties is completely validated by **closure tests**, **future tests**, and **parametrisation basis independence**

- In addition to its implications for **LHC precision physics**, NNPDF4.0 sheds novel light of crucial aspects of proton structure from **light antiquark asymmetries** to **intrinsic charm**

- The current level of PDF uncertainties challenge the accuracy of theoretical predictions and demand an increased effort towards the systematic inclusion in the fit of **theoretical uncertainties** (nuclear, higher orders, SM parameters, . . .) and **higher-order QCD** (including N3LO) **and EW corrections**

Summary and outlook



The **NNPDF machine learning fitting framework** will be publicly released **open source,** together with extensive documentation and user-friendly examples, at same time as upcoming paper!