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

XXVIII International Workshop on Deep-Inelastic Scattering and Related Subjects

Emanuele R. Nocera
School of Physics and Astronomy, The University of Edinburgh

on behalf of the NNPDF Collaboration

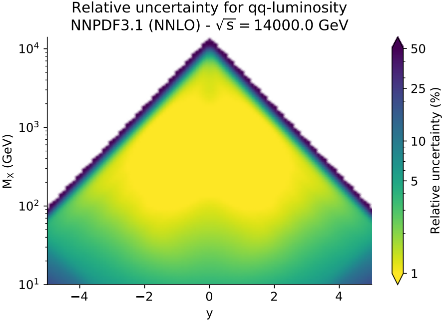
April 13, 2021

From NNPDF3.1 to NNPDF4.0

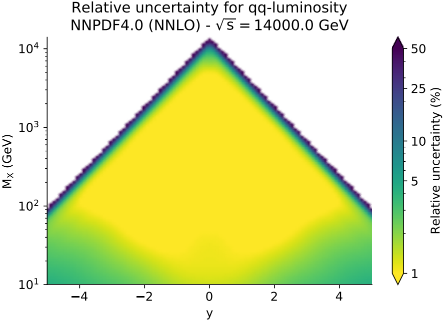
$$\mathcal{L}_{ij}(M_X, y, \sqrt{s}) = \frac{1}{s} \sum_{i,j} f_i\left(\frac{M_X e^{y}}{\sqrt{s}}, M_X\right) f_j\left(\frac{M_X e^{-y}}{\sqrt{s}}, M_X\right)$$

SINGLET

NNPDF3.1 (NNLO)



NNPDF4.0 (NNLO)



Steady progress towards 1% relative uncertainties on $\mathcal{L}_{ij}$ on a broad kinematic range

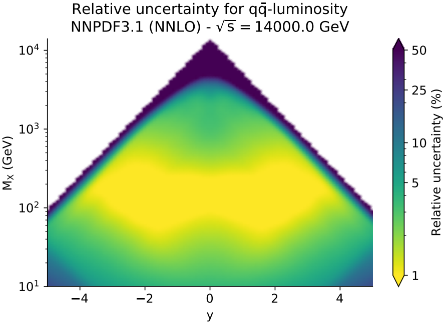
How are we getting there?

From NNPDF3.1 to NNPDF4.0

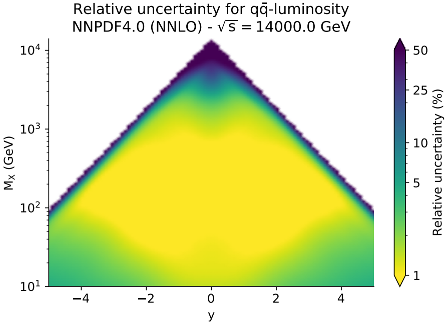
$$\mathcal{L}_{ij}(M_X, y, \sqrt{s}) = \frac{1}{s} \sum_{i,j} f_i \left(\frac{M_X e^{y}}{\sqrt{s}}, M_X \right) f_j \left(\frac{M_X e^{-y}}{\sqrt{s}}, M_X \right)$$

SINGLET

NNPDF3.1 (NNLO)



NNPDF4.0 (NNLO)



Steady progress towards 1% relative uncertainties on $\mathcal{L}_{ij}$ on a broad kinematic range

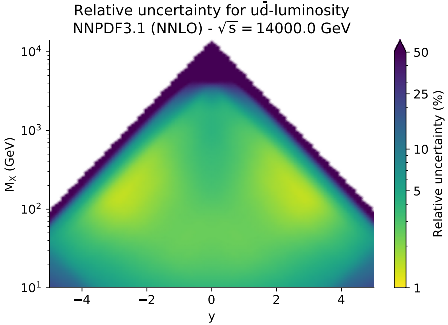
How are we getting there?

From NNPDF3.1 to NNPDF4.0

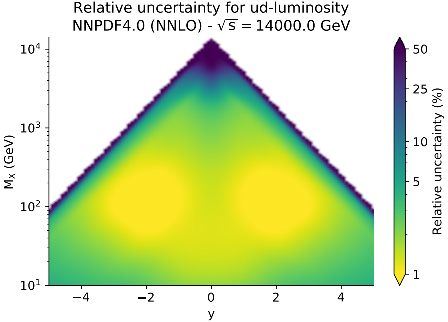
$$\mathcal{L}_{ij}(M_X, y, \sqrt{s}) = \frac{1}{s}\sum_{i,j} f_i\left(\frac{M_X e^{y}}{\sqrt{s}}, M_X\right) f_j\left(\frac{M_X e^{-y}}{\sqrt{s}}, M_X\right)$$

FLAVOURS

NNPDF3.1 (NNLO)



NNPDF4.0 (NNLO)



Steady progress towards 1% relative uncertainties on $\mathcal{L}_{ij}$ on a broad kinematic range

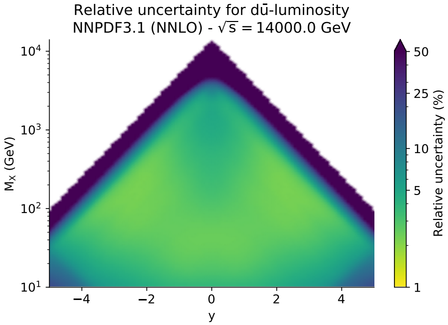
How are we getting there?

From NNPDF3.1 to NNPDF4.0

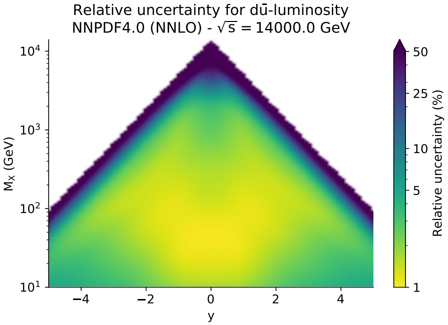
$$\mathcal{L}_{ij}(M_X, y, \sqrt{s}) = \frac{1}{s}\sum_{i,j} f_i\left(\frac{M_X e^{y}}{\sqrt{s}}, M_X\right) f_j\left(\frac{M_X e^{-y}}{\sqrt{s}}, M_X\right)$$

FLAVOURS

NNPDF3.1 (NNLO)



NNPDF4.0 (NNLO)



Steady progress towards 1% relative uncertainties on $\mathcal{L}_{ij}$ on a broad kinematic range

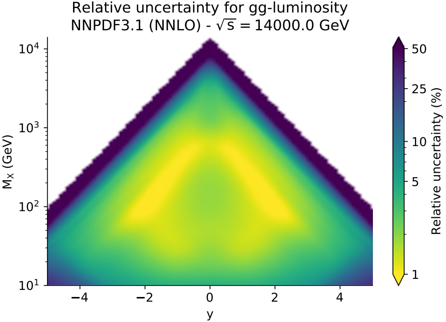
How are we getting there?

From NNPDF3.1 to NNPDF4.0

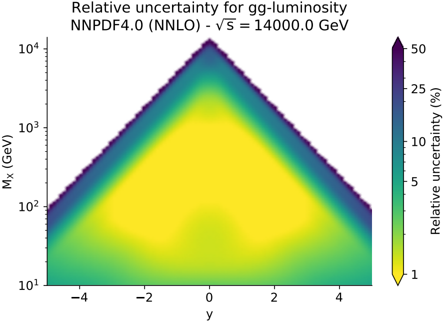
$$\mathcal{L}_{ij}(M_X, y, \sqrt{s}) = \frac{1}{s}\sum_{i,j} f_i\left(\frac{M_X e^{y}}{\sqrt{s}}, M_X\right) f_j\left(\frac{M_X e^{-y}}{\sqrt{s}}, M_X\right)$$

GLUON

NNPDF3.1 (NNLO)



NNPDF4.0 (NNLO)



Steady progress towards 1% relative uncertainties on $\mathcal{L}_{ij}$ on a broad kinematic range

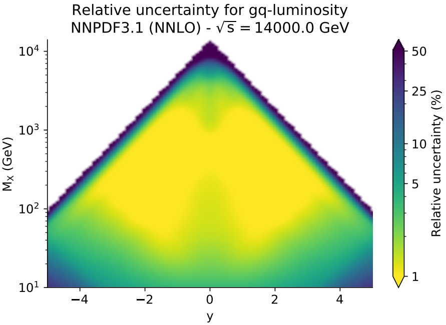
How are we getting there?

From NNPDF3.1 to NNPDF4.0

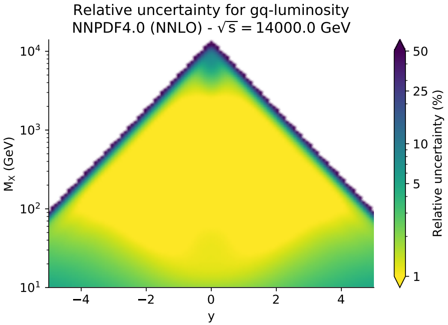
$$\mathcal{L}_{ij}(M_X, y, \sqrt{s}) = \frac{1}{s}\sum_{i,j} f_i\left(\frac{M_X e^{y}}{\sqrt{s}}, M_X\right) f_j\left(\frac{M_X e^{-y}}{\sqrt{s}}, M_X\right)$$

GLUON

NNPDF3.1 (NNLO)



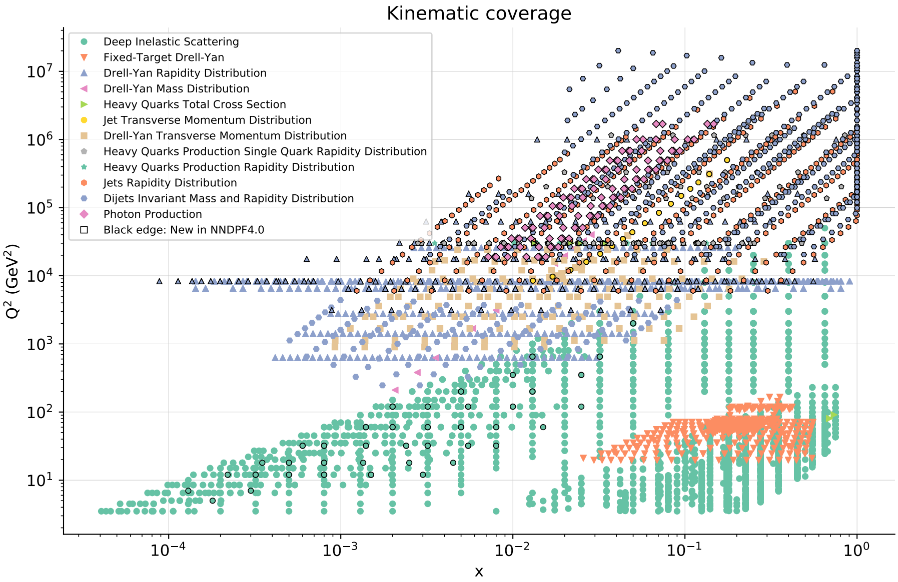
NNPDF4.0 (NNLO)



Steady progress towards 1% relative uncertainties on $\mathcal{L}_{ij}$ on a broad kinematic range

How are we getting there?

NNPDF4.0: data set extension



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

NNPDF4.0: new data sets

Process	Experiment	Description	Reference
DIS	HERA	Combined reduced c and b cross sections	[EPJ C78 (2018) 473]
	NOMAD*	$\mathcal{R}_{\mu\mu}(E) = \sigma_{\mu\mu}(E)/\sigma_{\rm CC}(E)$	[NPB 876 (2013) 339]
DY	ATLAS	W, Z central/forward rapidity distr., **7 TeV**	[EPJ C77 (2017) 367]
	ATLAS	$\{m_{\ell\ell}, \lvert y_{\ell\ell}\rvert\}$ Z high-mass distribution, **8 TeV**	[JHEP 08 (2016) 009]
	ATLAS	$\{m_{\ell\ell}, \lvert y_{\ell\ell}\rvert\}$ Z distribution, **8 TeV**	[JHEP 12 (2017) 059]
	ATLAS	W rapidity distr., **8 TeV**	[EPJ C79 (2019) 760]
	ATLAS	W and Z total cross section, **13 TeV**	[PLB 759 (2016) 601]
	LHCb	y_Z distribution, $2e$ and 2μ, **13 TeV**	[JHEP 09 (2016) 136]
W+c	ATLAS†	$\lvert\eta^\ell\rvert$ distribution **7 TeV**	[JHEP 05 (2014) 068]
	CMS†	$\lvert\eta^\mu\rvert$ distribution **13 TeV**	[EPJ C79 (2019) 269]
single-jet	ATLAS	$\{p_T, \lvert y\rvert\}$ distribution, **8 TeV**	[JHEP 09 (2017) 020]
$t\bar{t}$	CMS	total inclusive cross section, **5 TeV**	[JHEP 03 (2018) 115]
	CMS	normalised $\{m_{t\bar{t}}, y_t\}$ distribution, **8 TeV**	[EPJ C77 (2017) 459]
	CMS	normalised y_t distribution (dilepton), **13 TeV**	[JHEP 02 (2019) 149]
	CMS	normalised y_t distribution (lepton+jet), **13 TeV**	[PRD 97 (2018) 112003]
single top	ATLAS	R_t **7, 8, 13 TeV**	[JHEP 04 (2017) 086]
	ATLAS	normalised y_t and $y_{\bar{t}}$ distributions, **7, 8 TeV**	[PRD 90 (2014) 112006; EPJ C77 (2017) 531]
	CMS	$t+\bar{t}$ cross section, **7 TeV**	[JHEP 12 (2012) 035]
	CMS	R_t **8, 13 TeV**	[JHEP 06 (2014) 090; PLB 772 (2017) 752]
W+jet	ATLAS	p_T distribution, **8 TeV**	[JHEP 05 (2018) 077]
isolated photon	ATLAS	$\{E_T^\gamma, \lvert\eta^\gamma\rvert\}$ distribution, **13 TeV**	[PLB 770 (2017) 473]
di-jets	ATLAS	$\{m_{12}, y^*\}$ distribution **7 TeV**	[JHEP 05 (2014) 059]
	CMS	$\{m_{12}, \lvert y_{\rm max}\rvert\}$ distribution **7 TeV**	[PRD 87 (2013) 112002]
	CMS	$\{p_{\rm T,avg}, y_b, y^*\}$ distribution **8 TeV**	[EPJ C77 (2017) 746]
DIS+jets	H1*	Single- and di-jet differential distributions	[EPJ C75 (2015) 65; C77 (2017) 215]

*Not in baseline fit; studied via reweighting

†Only NLO fit

Processes highlighted in red correspond to processes NOT in NNPDF3.1

NNPDF4.0: theoretical and methodological features

- Refined theoretical framework [EPJ C79 (2019) 282; EPJ C81 (2021) 37; EPJ C80 (2020) 1168;]
  $\rightarrow$ nuclear uncertainties for both deuteron and heavy nuclei included by default
  $\rightarrow$ NNLO charm-quark massive corrections implemented (a bug in the NLO corrected)
  $\rightarrow$ EW corrections not included to ensure consistency with data, but carefully checked
  $\rightarrow$ charm PDF parametrised on the same footing as other PDFs
- Improved implementation of PDF properties [JHEP 11 (2020) 129]
  $\rightarrow$ extended positivity constraints for light quark/antiquark and gluon PDFs
  $\rightarrow$ extended integrability constraints of non-singlet light quark PDF combinations
- New PDF parametrisation and optimisation [EPJ C79 (2019) 676]
  $\rightarrow$ single neural network to parametrise eight independent PDF combinations
  $\rightarrow$ check of the independence of the results from the chosen parametrisation basis
  $\rightarrow$ new optimisation strategy based on gradient descent rather than genetic algorithms
  $\rightarrow$ scan of the hyperparameter space to find the optimal minimisation settings
- Complete statistical validation of PDF uncertainties [Acta Phys.Polon. B52 (2021) 243]
  $\rightarrow$ (multi-)closure tests to validate PDF uncertainties in the data region
  $\rightarrow$ future tests to check the sensibleness of PDF uncertainties in extrapolation regions
- More efficient compression tool for PDF set delivery [arXiv:2104.04535]

[See also talks by R.L. Pearson later today and by C. Schwan and F. Heckhorn tomorrow.]

NNPDF4.0: Fit quality – NNLO

Data set	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$
Fixed-target DIS	1881	1.10
HERA	1208	1.21
σ_c	37	2.11
σ_b	26	1.48
Fixed-target Drell-Yan	189	1.00
CDF	28	1.31
D0	37	1.00
ATLAS	621	1.18
Drell-Yan, 7, 8, 13 TeV	153	1.32
W+jet, 8 TeV	32	1.15
single top, 7, 8, 13 TeV	14	0.36
di-jets, 7 TeV	90	1.93
jets, 8 TeV	171	0.61
top pair, 7, 8, 13 TeV	16	2.30
Zp_T, 8 TeV	92	0.86
direct photon, 13 TeV	53	0.72
CMS	411	1.40
Drell-Yan, 7, 8 TeV	154	1.34
single top, 7, 8, 13 TeV	3	0.43
di-jets, 7 TeV	54	1.67
di-jets, 8 TeV	122	1.50
top pair, 5, 7, 8 TeV	29	0.84
top pair, 13 TeV	21	0.67
Zp_T, 8 TeV	28	1.42
LHCb	116	1.53
Total	4491	1.17

Overall good description of the data sets

Two exceptions:
HERA σ_c and ATLAS top pair

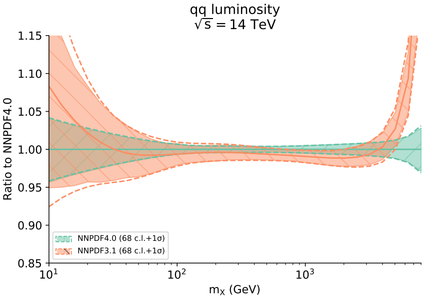
Weighted fits analysis:
in case of HERA σ_c:
lack of small-x resummation

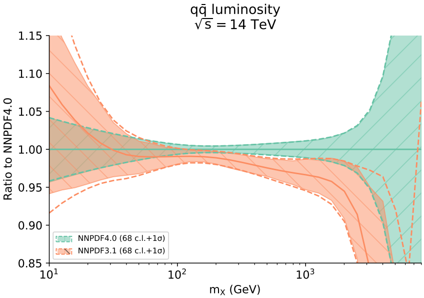
in case of ATLAS top pair:
slight tension with (di-jet) data sets
poor fit if all distributions are included
normalised rapidity distributions retained
although their $\chi^2/N_{\rm dat}$ of order 3
CMS top pair data almost insensitive to all this

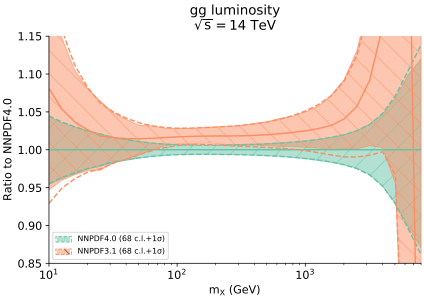
General remark:
as statistical uncertainties become smaller
a good control of systematic uncertainties
and their correlations becomes fundamental
to interpret the sensibleness of the fit

All results in the sequel are obtained at NNLO

From NNPDF3.1 to NNPDF4.0







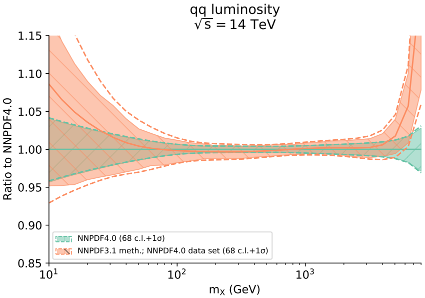
methodology / data set (N_{dat})	NNPDF3.1	NNPDF4.0
NNPDF3.1 (4093)	**1.19**	1.12
NNPDF4.0 (4491)	1.25	**1.17**

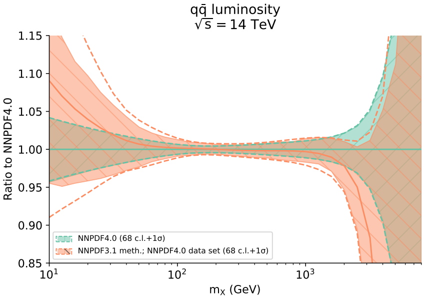
<u>Consistency</u> between PDF sets

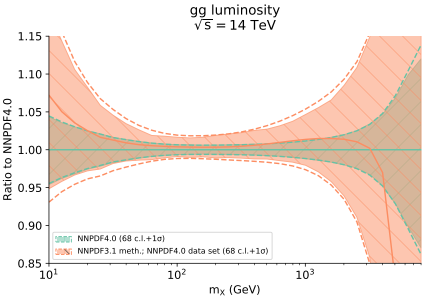
NNPDF4.0 <u>more precise</u>
(combination of data set and methodology)

NNPDF4.0 <u>more accurate</u>
(superiority of the NNPDF4.0 methodology)

From NNPDF3.1 to NNPDF4.0







methodology / data set ($N_{\rm dat}$)	NNPDF3.1	NNPDF4.0
NNPDF3.1 (4093)	1.19	1.12
NNPDF4.0 (4491)	**1.25**	**1.17**

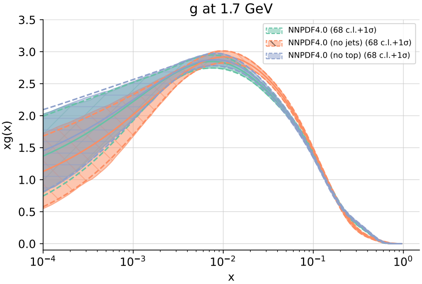
Consistency between PDF sets

NNPDF4.0 more precise
(combination of data set and methodology)

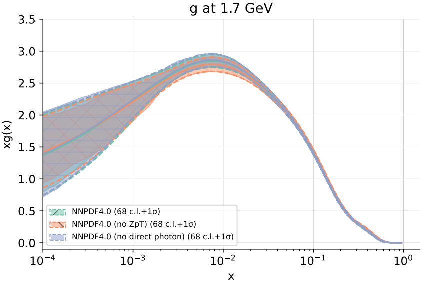
NNPDF4.0 more accurate
(superiority of the NNPDF4.0 methodology)

The gluon PDF: impact of data

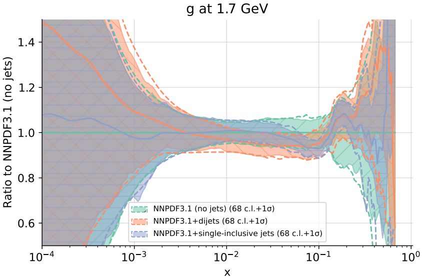
jet and $t\bar{t}$ data



Zp_T and direct photon data



di-jets vs single-inclusive jets

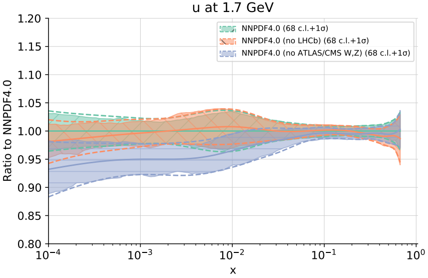


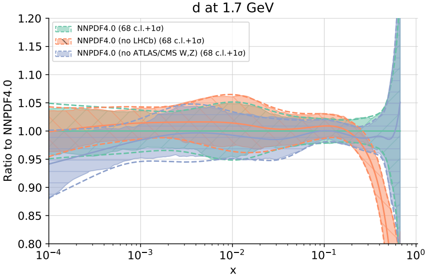
Hierarchical impact of different data sets
di-jet measurements have the largest pull
$t\bar{t}$ and Zp_T measurements have a comparatively small pull, which is consistent with the global fit
direct photon measurements almost immaterial

Inclusion of di-jet measurements is preferred over single-inclusive jet measurements
given their greater theoretical accuracy and the avoidance of decorrelation models
For details, see [EPJ C80 (2020) 8]

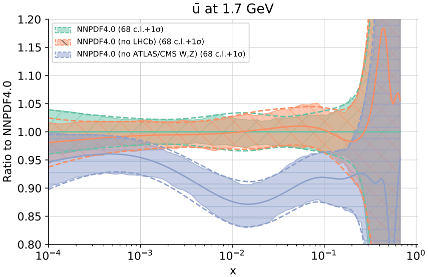
Quark flavour decomposition: impact of data

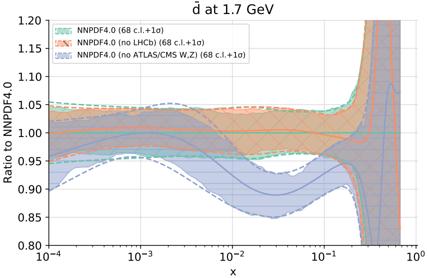
Quarks



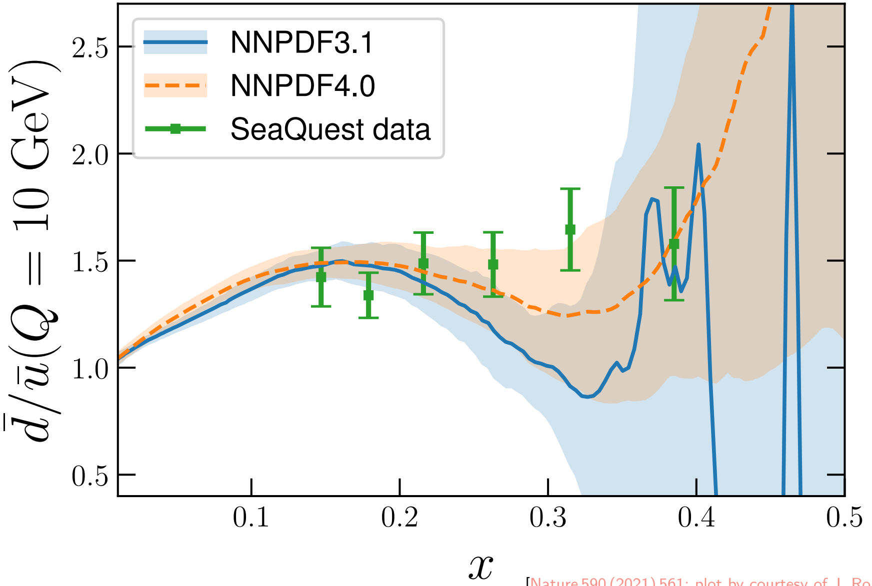


Antiquarks



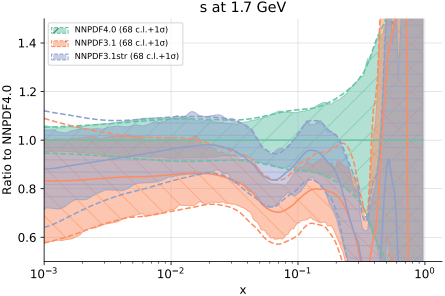


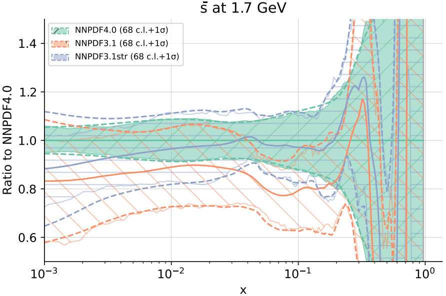
Sea quark asymmetry: SeaQuest

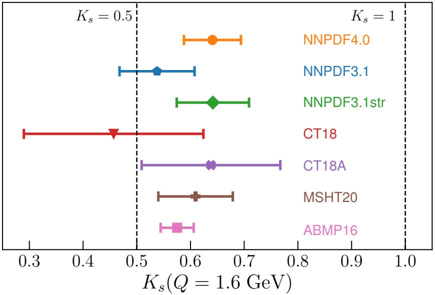


[Nature 590 (2021) 561; plot by courtesy of J. Rojo]

The strange PDF: impact of data







Enhanced s and $\bar{s}$ PDFs w.r.t. NNPDF3.1

effect of ATLAS W, Z and W+jet data

Good consistency with NNPDF3.1str

no nuclear uncertainties in NNPDF3.1str

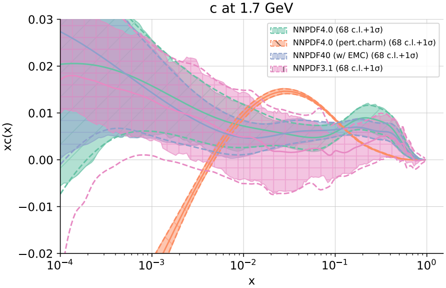
no NOMAD data in NNPDF4.0

Good consistency of K_s across PDF sets

$$K_s(Q^2) = \frac{\int_0^1 dx[s(x,Q^2)+\bar{s}(x,Q^2)]}{\int_0^1 dx[\bar{u}(x,Q^2)+\bar{d}(x,Q^2)]}$$

See also [EPJ C80 (2020) 1168]

Impact of theory: perturbative vs fitted charm

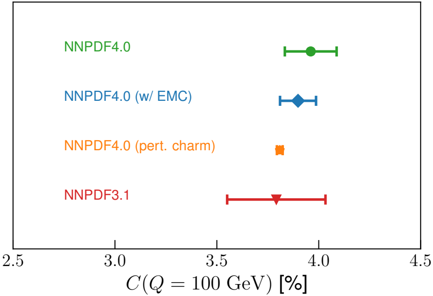


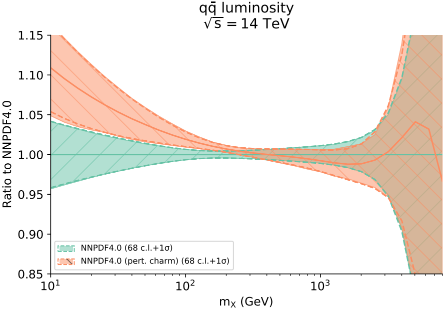
Striking evidence of intrinsic charm
even w/o EMC F_2^c data

Perturbative charm alters the flavour decomposition and deteriorates the fit

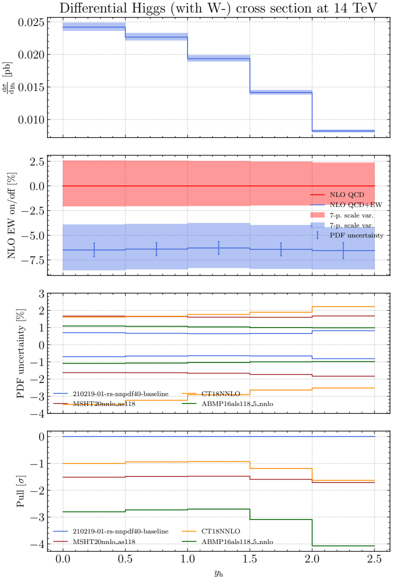
$$\chi^2_{\text{fitted charm}} = 1.17 \rightarrow \chi^2_{\text{pert. charm}} = 1.19$$

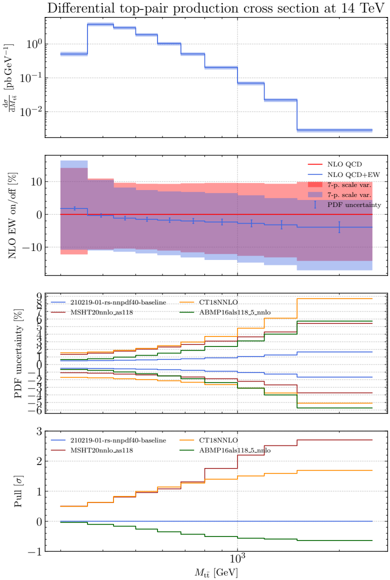
mainly due to a worsening
of the LHC W, Z and top pair data sets





NNPDF4.0: implications for LHC phenomenology





[Plots by courtesy of C. Schwan]

Conclusions

NNPDF4.0 is the next generation parton set of the NNPDF family.

It achieves 1% accuracy in an unprecedentedly broad kinematic range by consistently improving the previous NNPDF3.1 parton set.

This result builds upon an extensive LHC data set combined with deep-learning optimisation models.

Its faithfulness in representing PDF uncertainties is completely validated by closure tests.

1% PDF uncertainties challenge the accuracy of theoretical predictions and demand an increasing effort towards the systematic inclusion in the fit of theoretical uncertainties (nuclear, higher orders, physical parameters, ...) and higher-order QCD and EW corrections.

The **NNPDF code** used to produce the NNPDF4.0 parton set **will be made publicly available** with its documentation.

Conclusions

NNPDF4.0 is the next generation parton set of the NNPDF family.

It achieves 1% accuracy in an unprecedentedly broad kinematic range
by consistently improving the previous NNPDF3.1 parton set.

This result builds upon an extensive LHC data set
combined with deep-learning optimisation models.

Its faithfulness in representing PDF uncertainties is completely validated by closure tests.

1% PDF uncertainties challenge the accuracy of theoretical predictions
and demand an increasing effort towards the systematic inclusion in the fit of
theoretical uncertainties (nuclear, higher orders, physical parameters, ...)
and higher-order QCD and EW corrections.

The **NNPDF code** used to produce the NNPDF4.0 parton set
will be made publicly available with its documentation.

Thank you

Individual data sets: χ^2 breakdown

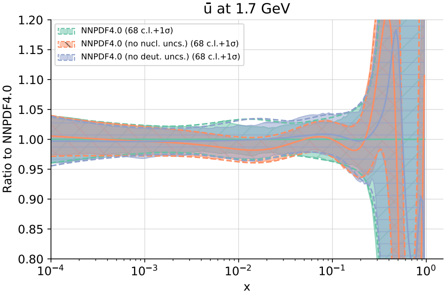
fit \ data set	ATLAS jets	CMS jets	ATLAS top	CMS top	ATLAS Zp_T	CMS Zp_T	ATLAS dir. phot.	total
NNPDF4.0	1.06	1.55	2.29	0.77	0.86	1.41	0.71	1.17
(no jets)	[1.71]	[3.70]	1.54	1.00	0.86	1.35	0.72	1.14
(no top)	1.08	1.57	[3.51]	[0.91]	0.86	1.43	0.74	1.18
(no Zp_T)	1.08	1.57	2.30	0.76	[0.99]	[1.41]	0.69	1.14
(no dir. phot.)	1.06	1.55	2.30	0.77	0.86	1.42	[0.71]	1.18

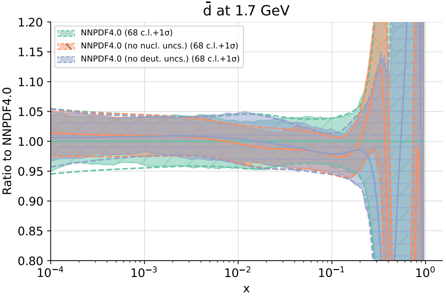
fit \ data set	ATLAS 2j	CMS 2j	ATLAS 1j (7 TeV)	ATLAS 1j (8 TeV)	CMS 1j	Zp_T	top	total
NNPDF4.0	1.93	1.56	[1.28] [3.42]*	0.61 [2.82]*	[1.31]	0.99	1.17	1.17
(single-jets instead of di-jets)	[2.41]	[2.68]	1.23 [3.36]*	0.85 [3.10]*	1.07	0.99	1.19	1.14

*No decorrelation model

fit \ data set	FT DIS	HERA	FT DY	Tevatron	ATLAS W, Z	CMS W, Z	LHCb	single top	total
NNPDF4.0	1.10	1.21	1.00	1.14	1.28	1.33	1.54	0.37	1.17
(no LHCb)	1.08	1.21	0.97	1.27	1.34	1.35	[2.60]	0.34	1.16
(no ATLAS/CMS W, Z)	1.05	1.20	0.85	1.02	[2.14]	[1.36]	1.39	0.37	1.11
DIS-only	1.03	1.21	[1.40]	[1.22]	[4.15]	[3.83]	[2.96]	[0.33]	1.10

Quark flavour separation: nuclear uncertainties

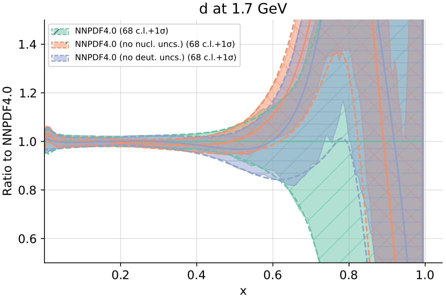




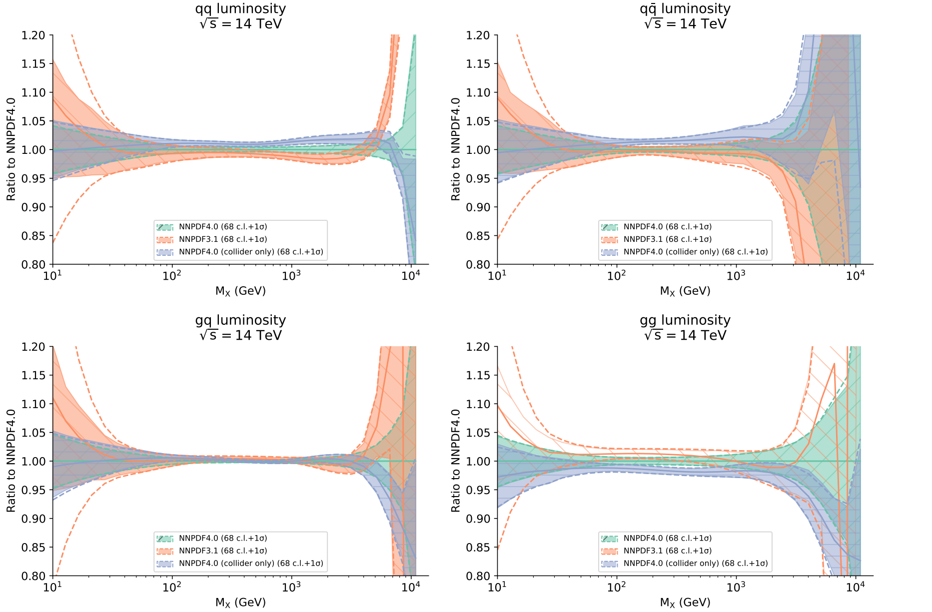
Effect of nuclear uncertainties relevant at large x
to reconcile FT DIS with LHC DY data

$\chi^2_{\rm tot} = 1.17 \to \chi^2_{\rm tot} = 1.26$ (no nucl. uncs.)
$\chi^2_{\rm LHCb} = 1.54 \to \chi^2_{\rm tot} = 1.76$ (no nucl. uncs.)

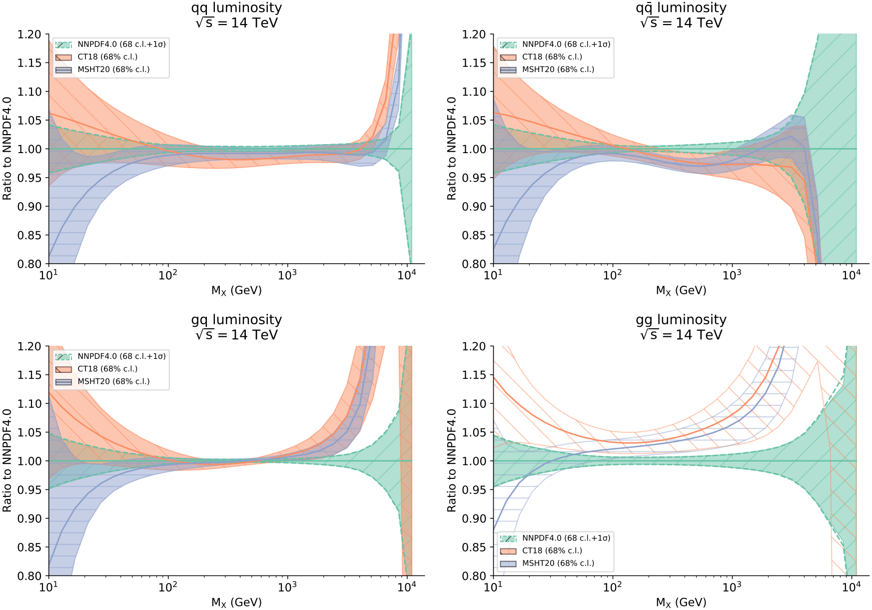
The bulk of the effect is due to nuclear uncertainties for heavy nuclei
deuteron uncertainties have a comparatively smaller effect at inermediate values of x



NNPDF4.0: parton luminosities



NNPDF4.0: parton luminosities

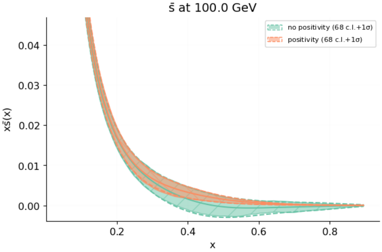


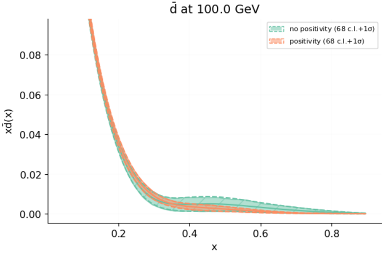
Positivity - Implementation

Quarks, anti-quarks and gluon $\overline{MS}$ PDFs q_k have to be positive: we add a term in the χ^2 penalizing negative distributions

$$\chi^2_{tot} = \chi^2_{exp} + \sum_k \chi^2_{k,\text{pos}} \,,$$

$$\chi^2_{k,pos} = \Lambda_k \sum_i \Theta\left(-q_k\left(x_i, Q^2\right)\right) \,, \quad \text{with} \quad \Theta(t) = \begin{cases} t & \text{if } \; t > 0 \\ 0 & \text{if } \; t < 0 \end{cases} \,.$$





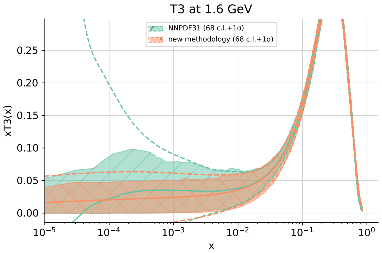
Integrability

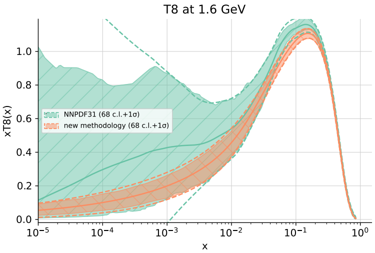
In order to satisfy valence and Gottfried sum rules the distributions $q_k = V, V_3, V_8, T_3, T_8$ have to be integrable at small-x

$$\lim_{x\to 0} x q_k\left(x, Q_0^2\right) = 0\,.$$

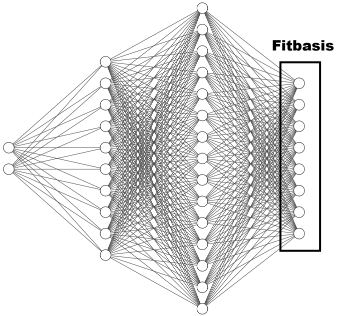
Similarly to what done for positivity, we add to the total χ^2 a penalty of the form

$$\chi^2_{k,integ} = \Lambda_k \sum_i \left[x_i\, q_k\left(x_i, Q^2\right)\right]^2\,.$$





Fitbasis



Flavour basis:

$g, u, \bar{u}, d, \bar{d}, s, \bar{s}, c$

Evolution basis:

$g, \Sigma, V, V_3, V_8, T_3, T_8, T_{15}$

- independently on the basis choice the same physical constraints have to be satisfied: positivity and integrability
- NNPDF4.0 will be hyper-optimized in the evolution basis
- the final results should not depend on the details of the methodology $\rightarrow$ fitbasis independence studies

Key differences with respect to the 3.1 methodology

NNPDF 3.1 code

- → **Genetic Algorithm optimizer**
- → One network per flavour
- → Physical constraints imposed independently of optimization
- → Preprocessing fixed per each of the replicas
- → C++ monolithic codebase
- → In-house Machine Learning optimization framework
- → Fitting times of up to various days

↓

Fit parameters manually chosen (manual optimization of hyperparameters)

NNPDF 4.0 code

- → **Gradient Descent optimization**
- → One network for all flavours
- → Physical constraints integrated in the optimization
- → Preprocessing can be fitted within replicas
- → Python object oriented codebase
- → Freedom to use external libraries (default: TensorFlow)
- → Results available in less than an hour

↓

Fit parameters chosen automatically (hyperparameter scan)

Beyond the PDF fit: fitting the methodology

The main objective of NNPDF is to minimize choices that can bias the PDF:

- ✗ Functional form $\longrightarrow$ Neural Networks
- ✗ However: NN are defined by set of parameters!

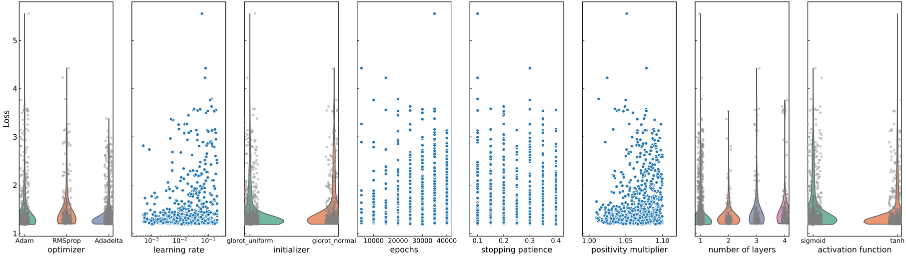
Humans are good at recognising patterns but selecting the best set of parameters is a slow process and systematic success is not guaranteed

To overcome this selection problem we implement a hyperparameter scan: let the computer decide automatically

- ✓ Scan over thousands of hyperparameter combinations
- ✓ Define a reward function to grade the model
- ✓ Check the generalization power of the model

Hyperparameter scan

Each blue dot corresponds to a fit of a different set of hyperparameters:



Thousands of fits for the hyperoptimization algorithm to choose:

- ✓ Optimizer
- ✓ Learning Rate
- ✓ Initializer
- ✓ Epochs
- ✓ Stopping Patience
- ✓ Positivity Multiplier
- ✓ Number of Layers
- ✓ Activation Function

Hyperoptimization: reward and generalization

If we use as hyperoptimization target the χ^2 of the fitted data, we risk finding the hyperparameter set that better overfits.

We avoid this problem by adopting ***k*-folding**:

- Divide the data into k sets.
- Leave one set out and fit the $k-1$ sets left.
- Optimize the average χ^2 of the k non-fitted sets.

Example of function to hyperoptimize:

$$\text{Loss}(optimizer\_name,\ depth\_of\_network) = \frac{1}{k}\sum_{k}^{i}\frac{\chi_i^2}{N_i}$$

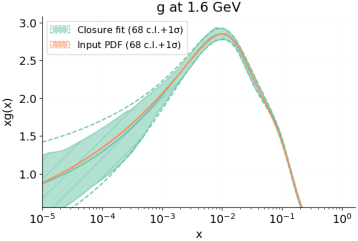
Where we are computing the χ^2 for the data that did not enter the fit. This ensures that the methodology can accommodate well even data that has never been seen by the fit.

Closure Tests

Fit replicas to pseudodata in usual way

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{f} + \boldsymbol{\eta} + \boldsymbol{\epsilon} \\ &= \boldsymbol{z} + \boldsymbol{\epsilon}, \end{aligned} \tag{1}$$

where $\boldsymbol{\eta} \sim \mathcal{N}(0, C)$ and $\boldsymbol{\epsilon} \sim \mathcal{N}(0, C)$ are sampled independently.
Use predictions from an input PDF as proxy for $\boldsymbol{f}$.



Example closure fit and input PDF.

Allows testing of methodology, if the input assumptions hold.

For example:

Bias: difference between central prediction and true observable

Variance: uncertainty of replica predictions

Bias is a stochastic variable. If PDF uncertainty is faithful then

$$\mathbf{E}_{\eta}[\text{bias}] = \text{variance} \tag{2}$$

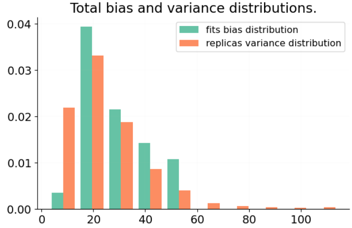
High demand on resources - made feasible with next generation fitting code.

Preliminary results

Compare first moments:

	$\sqrt{\mathbf{E}_\eta[\text{bias}]/\mathbf{E}_\eta[\text{variance}]}$
Total	1.11 ± 0.5

Alternatively look at the respective distributions

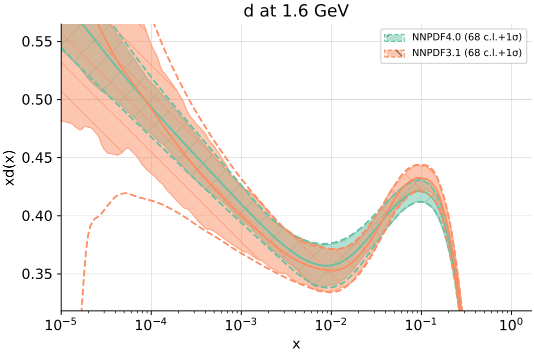


Bias distribution sampled with 25 fits, 40 replicas each.

How can we future-proof the methodology?

Do we trust our errorbands?

The smaller error bands in the NNPDF4.0 fits are driven both by the increased amount of data and the improved methodology. But there are still kin. regions not covered by data!



Ideally: design an experiment for the regions not covered by fitted-data!

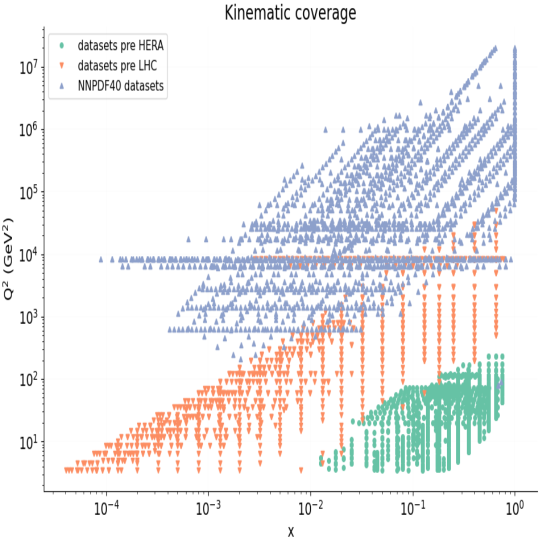
Problem: we want the results before 2050...

Fig: Other valid and certified future-testing methods

Solution: chronologically ordered subsets of data to test unseen regions, we named this "future tests".

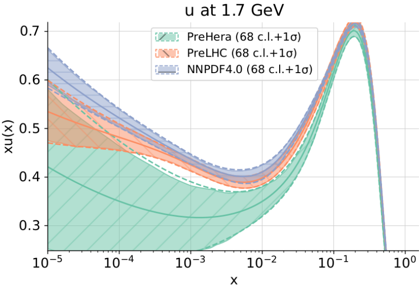
Future tests

for more information see arxiv:2103.08606



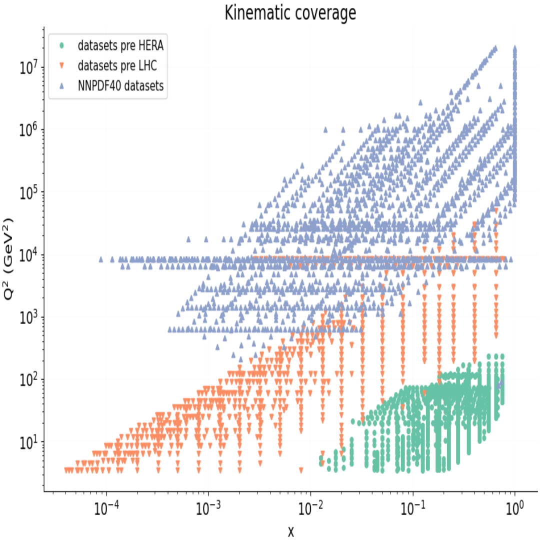
χ^2/N (only exp. covmat)

(dataset)	NNPDF4.0	pre-LHC	pre-Hera
pre-HERA	1.09	1.01	0.90
pre-LHC	1.21	1.20	**23.1**
NNPDF4.0	1.29	**3.30**	**23.1**



Future tests

for more information see arxiv:2103.08606



χ^2/N (exp. and PDF covmat)

(dataset)	NNPDF4.0	pre-LHC	pre-Hera
pre-HERA			0.86
pre-LHC		1.17	**1.22**
NNPDF4.0	1.12	**1.30**	**1.38**