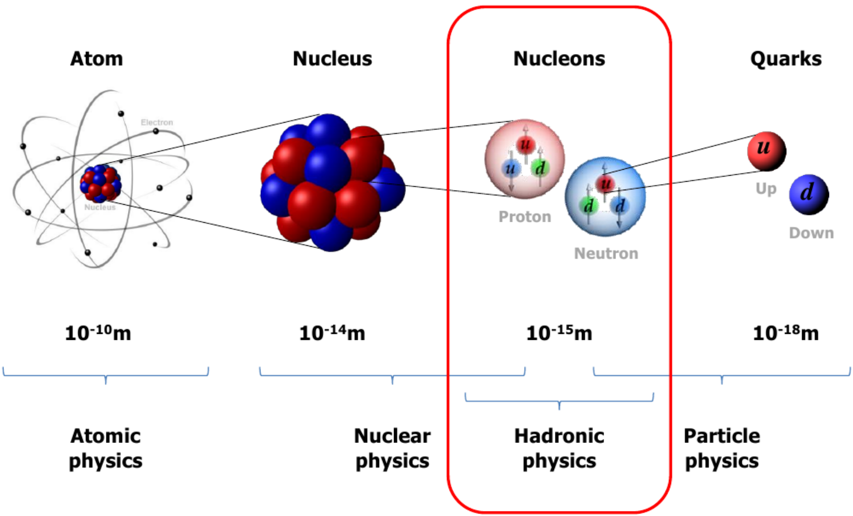
Parton Distribution Functions for Discovery at Next Generation Colliders

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

17 November 2021

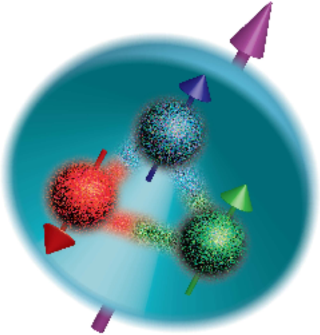
Nuclear, Hadronic and Particle Physics

Nucleons make up all nuclei, hence most of the visible matter in the Universe

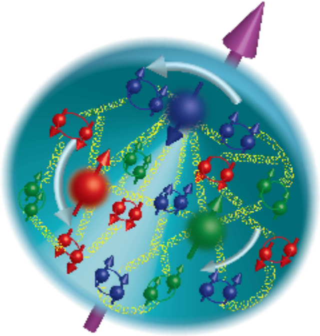


QCD, a theory for nucleon structure

Quark Model



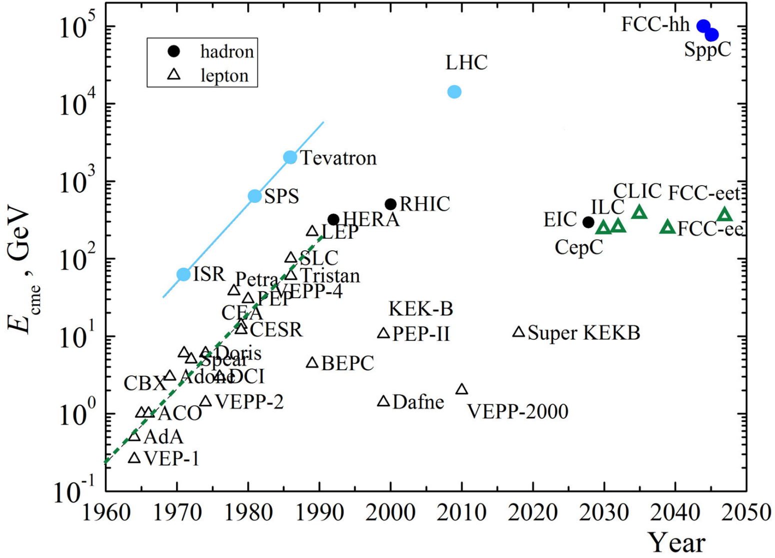
QCD Model



three non-relativistic quarks $\xleftrightarrow[\text{factorisation,evolution}]{\text{QCD}}$ indefinite number of relativistic quarks and gluons

Nucleons can be prepared in definite spin states

Colliders: past, present, future



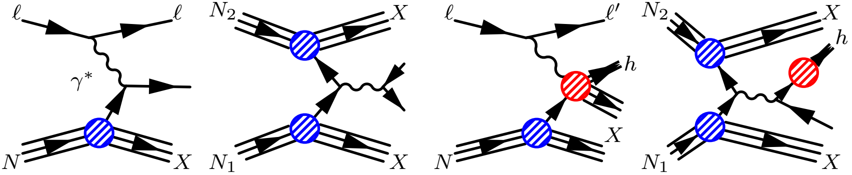
$\mathcal{L}^{HL-LHC} = 3\ \text{ab}^{-1}$ $\mathcal{L}^{EIC} = 100\ \text{fb}^{-1}$

[Reviews in Physics 6 (2021) 100053]

Theoretical framework

1. Factorisation of physical observables

$$\mathcal{O}_I = \sum_{i=q,\bar{q},g} C_{Ii}(y,\alpha_s(\mu^2)) \otimes f_i(y,\mu^2) + \text{p.s. corrections} \qquad f \otimes g = \int_x^1 \frac{dy}{y} f\left(\frac{x}{y}\right) g(y)$$



$\ell + N \to \ell + X$
deep-inelastic scattering (DIS)

$N_1 + N_2 \to A + X$
(inclusive) production in proton-proton scattering (pp)

$\ell + N \to \ell' + h + X$
semi-inclusive deep-inelastic scattering (SIDIS)

$N_1 + N_2 \to h + X$
high-p_T hadron production in proton-proton collisions (pp)

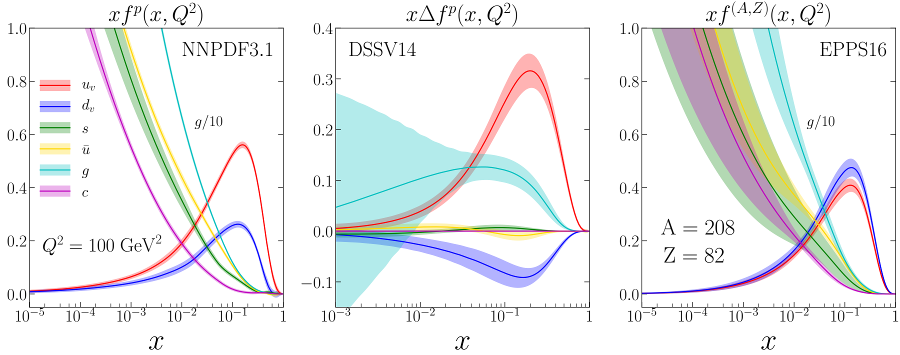
2. Perturbative expansion of coefficient functions

$$C_{Ii}(y,\alpha_s) = \sum_{k=0} a_s^k C_{Ii}^{(k)}(y), \qquad a_s = \alpha_s/(4\pi)$$

3. Perturbative (DGLAP) evolution of PDFs

$$\frac{\partial}{\partial \ln \mu^2} f_i(x,\mu^2) = \sum_j^{n_f} \int_x^1 \frac{dz}{z} P_{ji}\left(z,\alpha_s(\mu^2)\right) f_j\left(\frac{x}{z},\mu^2\right) \qquad P_{ji}(z,\alpha_s) = \sum_{k=0} a_s^{k+1} P_{ji}^{(k)}(z)$$

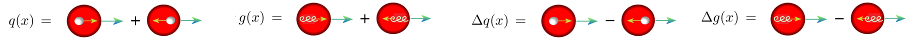
Parton Distributions



[Figure taken from Ann.Rev.Nucl.Part.Sci. 70 (2020) 43]

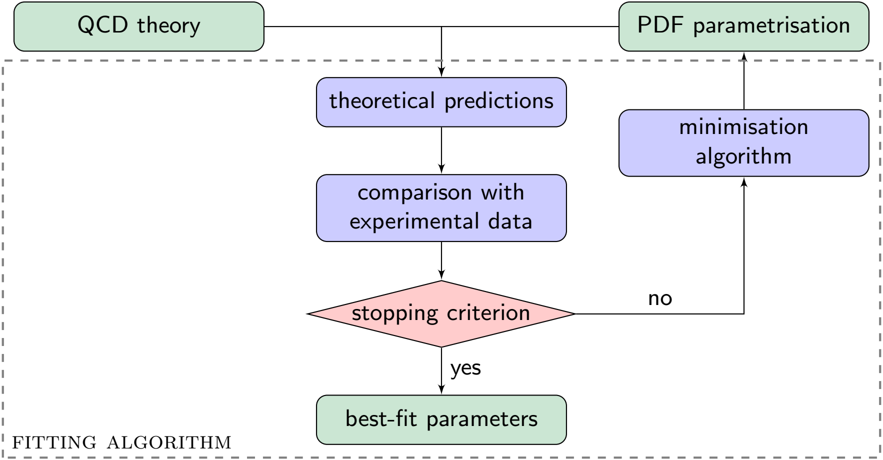
The densities of partons $f = q, \bar{q}, g$ with momentum fraction x

$$f^p(x) \equiv f^\uparrow(x) + f^\downarrow(x) \qquad \Delta f^p(x) \equiv f^\uparrow(x) - f^\downarrow(x)$$



$$f^{(A,Z)}(x) = \frac{Z}{A} f^p(x) + \frac{A-Z}{A} f^n(x)$$

A global PDF determination: the underlying strategy



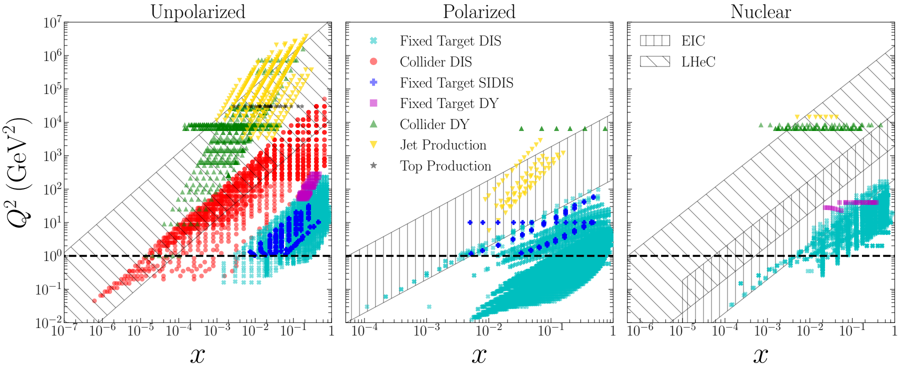
Assume a reasonable PDF parametrisation

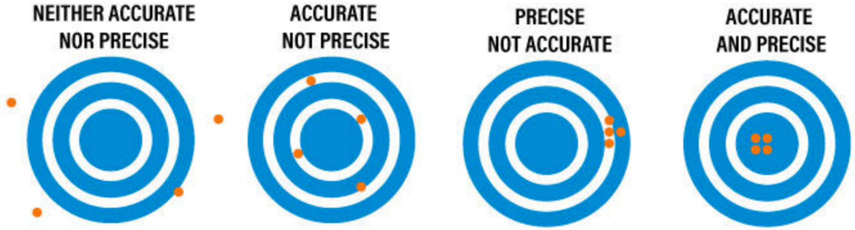
Obtain theoretical predictions for various processes and compare predictions to data

Determine the best-fit parameters via minimisation of a proper figure of merit (*e.g.* χ^2)

Self-validate PDF's accuracy and precision

Data, Accuracy and Precision





Proton PDFs

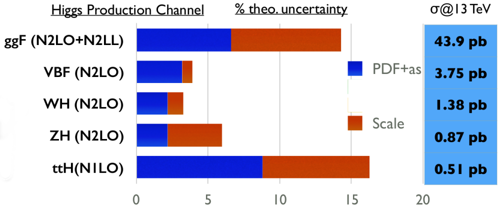
PDF uncertainty is often the dominant source of uncertainty in LHC cross sections

Higgs boson characterisation

Determination of SM parameters, such as the mass of the W boson

Searches for beyond SM physics at large invariant mass of the final state

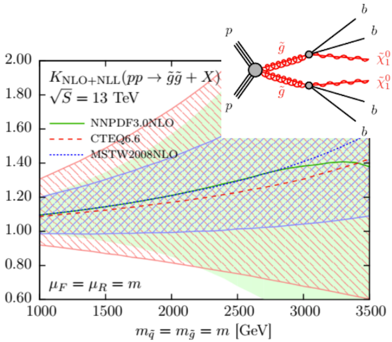
Precision



Channel	$m_{W^+} - m_{W^-}$ [MeV]	Stat. Unc.	Muon Unc.	Elec. Unc.	Recoil Unc.	Bckg. Unc.	QCD Unc.	EW Unc.	PDF Unc.	Total Unc.
$W \to e\nu$	−29.7	17.5	0.0	4.9	0.9	5.4	0.5	0.0	24.1	30.7
$W \to \mu\nu$	−28.6	16.3	11.7	0.0	1.1	5.0	0.4	0.0	26.0	33.2
Combined	−29.2	12.8	3.3	4.1	1.0	4.5	0.4	0.0	23.9	28.0

[Plot from the CERN Yellow Report 2016]

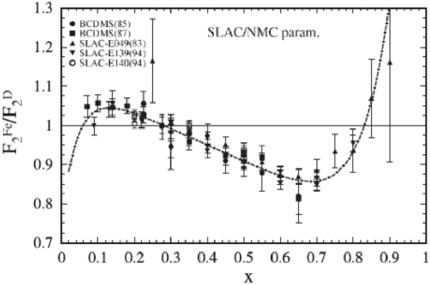
Discovery



[EPJC 76 (2016) 53]

Nuclear and Polarised PDFs

Nuclear PDFs



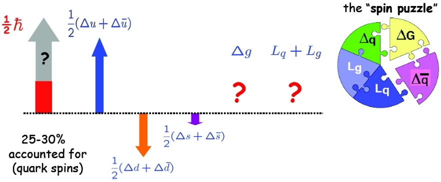
nuclei do not behave as a simple incoherent superposition of protons and neutrons

nPDFs enter theoretical predictions of signal and background events at high-energy neutrino observatories such as KM3NET and IceCube

search for exotic forms of QCD matter, such as the gluon-dominated Color Glass Condensate

interplay with proton PDFs, given the data used

Polarised PDFs



$$\mathcal{J}(\mu^2) = \sum_f \left\langle P; S | \hat{J}^z_f(\mu^2) | P; S \right\rangle$$

$$\frac{1}{2}\Delta\Sigma(\mu^2) + \Delta G(\mu^2) + \mathcal{L}_q(\mu^2) + \mathcal{L}_g(\mu^2)$$

$$\Delta\Sigma(\mu^2) = \sum_{q=u,d,s} \int_0^1 [\Delta q(x,\mu^2) + \Delta\bar{q}(x,\mu^2)]$$

$$\Delta G(\mu^2) = \int_0^1 dx \Delta g(x,\mu^2)$$

1. Some selected NNPDF results

The NNPDF methodology in a nutshell

1. Neural network parametrisation of PDFs
   - redundant and flexible parametrisation, $\mathcal{O}(200)$ parameters
   - requires a proper minimisation algorithm and stopping criterion

$\Rightarrow$ **reduce the theoretical bias due to the parametrisation**

2. Monte Carlo propagation of errors
   - generate experimental data replicas assuming multi-Gaussian probability distribution
   - validate against experimental data to determine the sample size

$\Rightarrow$ **no need to rely on linear error propagation**

PDF replicas are equally probable members of a statistical ensemble
which samples the probability density $\mathcal{P}[f_i]$ in the space of PDFs

$$\langle\mathcal{O}\rangle = \int \mathcal{D}f_i \mathcal{P}[f_i]\mathcal{O}[f_i]$$

Expectation values for observables are Monte Carlo integrals

$$\langle\mathcal{O}[f_i(x,Q^2)]\rangle = \frac{1}{N_{\text{rep}}}\sum_{k=1}^{N_{\text{rep}}}\mathcal{O}[f_i^{(k)}(x,Q^2)]$$

and similarly for uncertainties, correlations, etc.

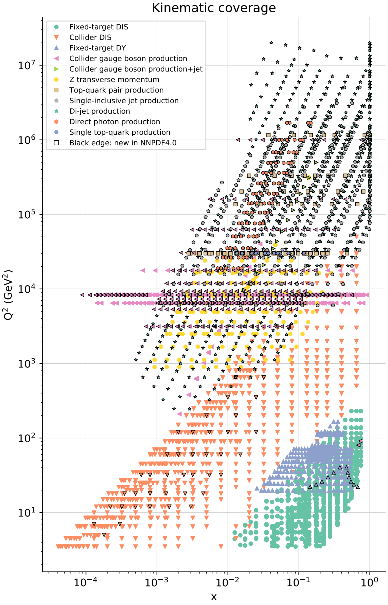
Proton PDFs: NNPDF4.0 [2109.02653]

- Refined theoretical framework [EPJ C79 (2019) 282; EPJ C81 (2021) 37; EPJ C80 (2020) 1168]
  → nuclear uncertainties for both deuteron and heavy nuclei included by default
  → NNLO charm-quark massive corrections implemented (a bug in the NLO corrected)
  → EW corrections not included to ensure consistency with data, but carefully checked
  → charm PDF parametrised on the same footing as other PDFs
- Improved implementation of PDF properties [JHEP 11 (2020) 129]
  → extended positivity constraints for light quark/antiquark and gluon PDFs
  → extended integrability constraints of non-singlet light quark PDF combinations
- New PDF parametrisation and optimisation [EPJ C79 (2019) 676]
  → single neural network to parametrise eight independent PDF combinations
  → check of the independence of the results from the chosen parametrisation basis
  → new optimisation strategy based on gradient descent rather than genetic algorithms
  → scan of the hyperparameter space to find the optimal minimisation settings
- Complete statistical validation of PDF uncertainties [Acta Phys.Polon. B52 (2021) 243]
  → (multi-)closure tests to validate PDF uncertainties in the data region
  → future tests to check the sensibleness of PDF uncertainties in extrapolation regions
- Open source fitting code [EPJC 81 (2021) 958]

https://nnpdf.mi.infn.it/nnpdf-open-source-code/

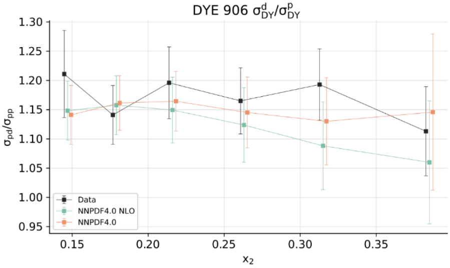
Proton PDFs: NNPDF4.0 [2109.02653]

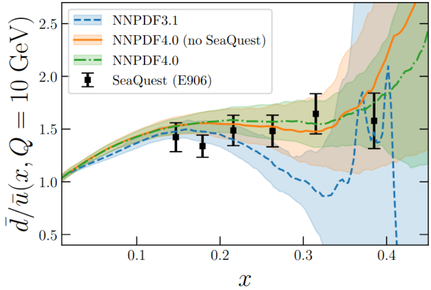
Data set	$N_{\rm dat}$	$\chi^2/N_{\rm dat}$
Fixed-target DIS	1881	1.10
HERA	1208	1.21
σ_c	37	2.11
σ_b	26	1.48
Fixed-target Drell-Yan	189	1.00
CDF	28	1.31
D0	37	1.00
ATLAS	621	1.18
Drell-Yan, 7, 8, 13 TeV	153	1.32
W+jet, 8 TeV	32	1.15
single top, 7, 8, 13 TeV	14	0.36
di-jets, 7 TeV	90	1.93
jets, 8 TeV	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



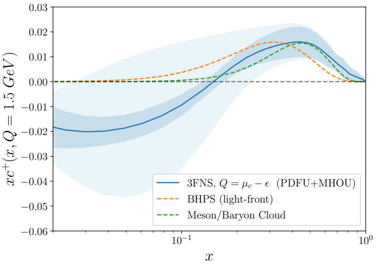
Proton PDFs at the LHC

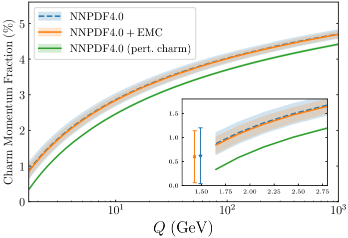
The $\bar{d}/\bar{u}$ asymmetry in the proton [2109.02653]





Charm in the proton [NNPDF, to appear]

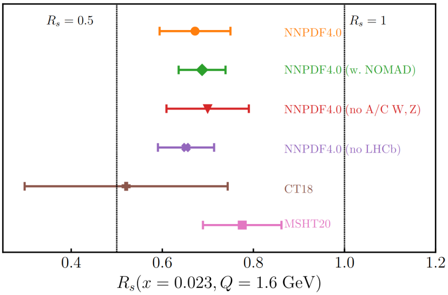




$$[c](Q) \equiv \int_0^1 dx\, x[c(x,Q) + \bar{c}(x,Q)]$$

Proton PDFs at the LHC

Strange in the proton [EPJ C80 (2020) 1168]



Satisfactory description of all datasets
no evidence for tensions

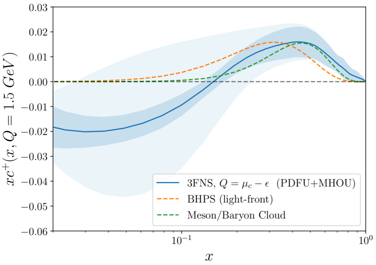
Sizeable constraint from NOMAD data
consistent with collider data

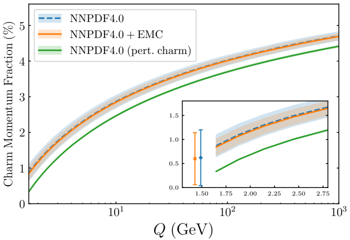
Moderate suppression of strange PDF

Good consistency of R_s across PDF sets

$$R_s(x, Q^2) = \frac{s(x,Q^2)+\bar{s}(x,Q^2)}{\bar{u}(x,Q^2)+\bar{d}(x,Q^2)}$$

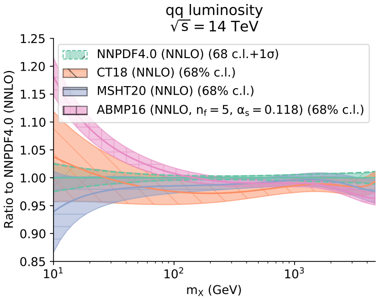
Charm in the proton [NNPDF, to appear]

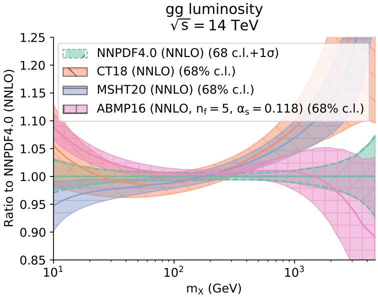




$$[c](Q) \equiv \int_0^1 dx\, x[c(x,Q) + \bar{c}(x,Q)]$$

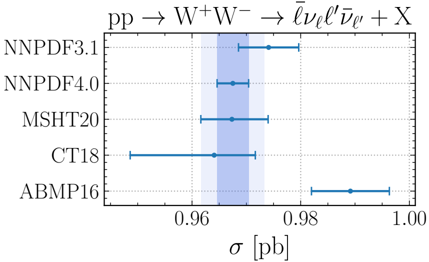
Proton PDFs at the LHC

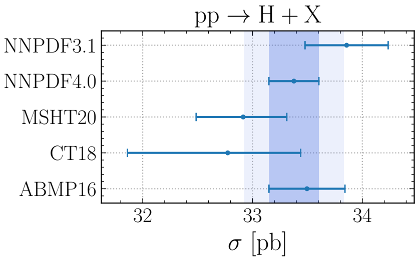




$$\mathcal{L}_{ij}(M_X, \sqrt{s}) = \frac{1}{s}\int_\tau^1 \frac{dx}{x} f_i(x, M_X) f_j(\tau/x, M_X)$$

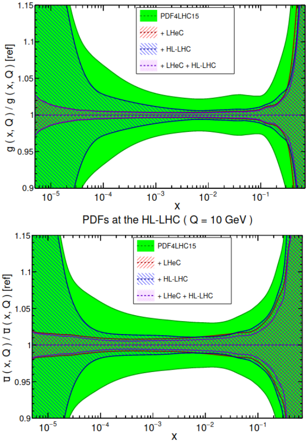
$$\tau = M_X^2/s$$





Proton PDFs at the HL-LHC and at the EIC

HL-LHC [SciPost Phys. 7 (2019) 051]



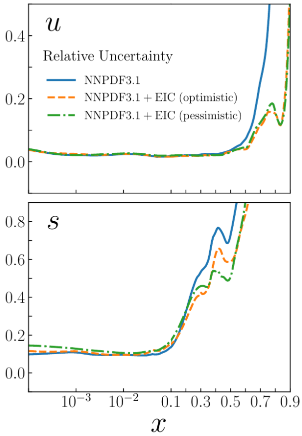
Impact determined from pseudodata

$e^{\pm}p$ (LHeC); various processes (HL-LHC)

$\sqrt{s} = 7$ TeV (LHeC); $\sqrt{s} = 14$ TeV (HL-LHC)

$\mathcal{L} = 1$ ab^{-1} (LHeC); $\mathcal{L} = 3$ ab^{-1} (HL-LHC)

EIC [PRD 103 (2021) 096005]



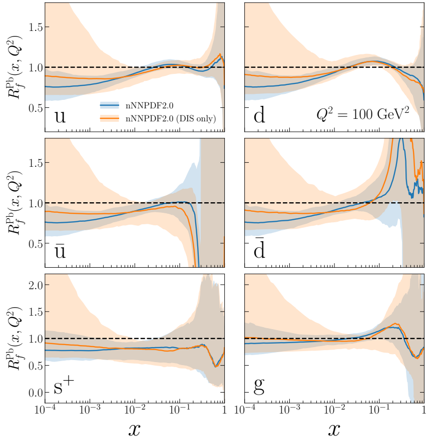
Impact determined from pseudodata

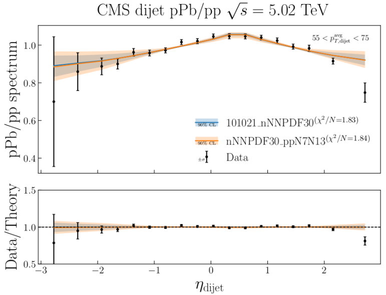
$e^{\pm}p$ (NC and CC DIS); $e^{\pm}d$ (NC DIS)

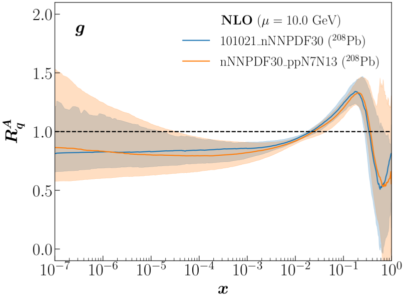
$E_\ell \times E_p$ [GeV]: 18× 275; 10× 100; 5× 100

$\mathcal{L} = 100$ fb^{-1}; $\sigma_u = 1.5/2.3\%$; $\sigma_c = 2.5/4.3\%$

Nuclear PDFs at the LHC

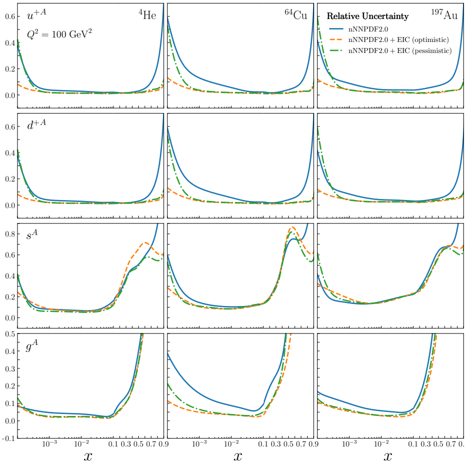






nNNPDF2.0 [JHEP 09 (2020) 183]
NC and CC DIS, LHC W, Z
nNNPDF3.0 NNPDF, in preparation]
LHC γ, dijets, D^0

Nuclear PDFs at the EIC

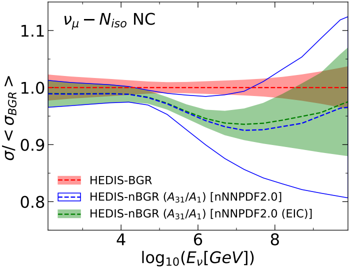


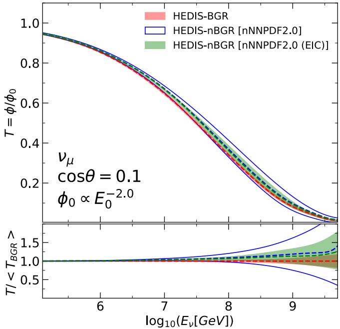
Impact from pseudodata [PRD 103 (2021) 096005]

e^-N NC DIS, N=^{4}He, ^{12}C, ^{40}Ca, ^{64}Cu, ^{197}Au

$E_\ell \times E_p$ [GeV]: 18× 100; 10× 110; 5× 41

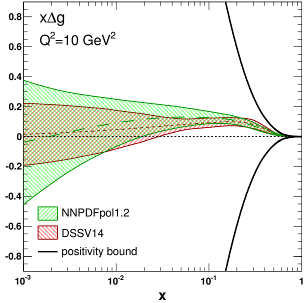
$\mathcal{L} = 10$ fb^{-1}; $\sigma_u = 1.5/2.3\%$; $\sigma_c = 2.5/4.3\%$

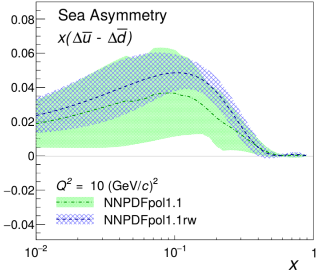




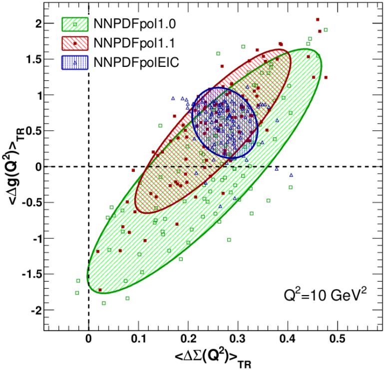
Polarised PDFs at RHIC and at the EIC

NNPDFpol1.1 [NPB 887 (2014) 276]





NNPDFpolEIC [PLB 728 (2014) 524]



$Q^2 = 10$ GeV^2	$\int_{10^{-3}}^{1} dx\, \Delta\Sigma$	$\int_{10^{-3}}^{1} dx\, \Delta g$
NNPDFpol1.0	$+0.23 \pm 0.15$	-0.06 ± 1.12
NNPDFpol1.2	$+0.25 \pm 0.10$	$+0.49 \pm 0.75$
NNPDFpolEIC	$+0.24 \pm 0.04$	$+0.49 \pm 0.25$

quarks and antiquarks $\sim 20\% - 30\%$
gluons $\sim 70\%$ OAM $\sim 0\%$

2. Methodological challenges

Data inconsistency: tensions between data sets

Give more weight to a data set p

$\chi^2 \rightarrow \chi^2 + w\chi^2_p \qquad w = N_{\rm dat}/N_{\rm set}$

Refit: the total χ^2 will increase
Which data sets get worse? How much?

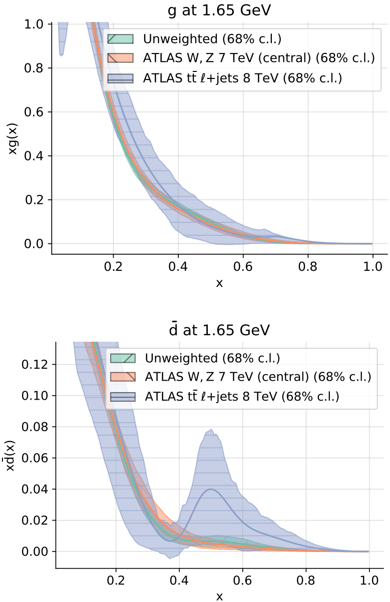
Refit: the data set χ^2_p will decrese
Self-consistency? Inconsistency?

ATLAS W, Z 7 TeV; $N_{\rm dat} = 46$; $w = 102$
ATLAS $t\bar{t}$ 8 TeV; $N_{\rm dat} = 6$; $w = 786$

Can improve the quality of the dataset BUT
description of other datasets deteriorates
unnatural PDF shapes appear

Fit quality for D0 el. asy. remains poor

Data set	baseline	wgt.	Total
ATLAS W, Z 7 TeV	1.86	1.23	1.18
ATLAS $t\bar{t}$ 8 TeV	1.86	1.32	1.47
Total	1.17		



Data inconsistency: experimental correlations

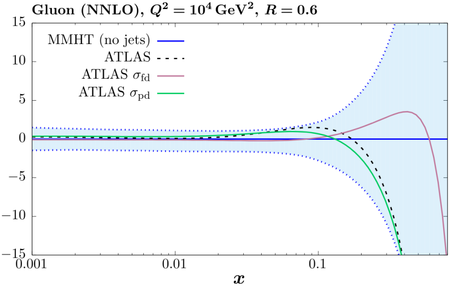
Single inclusive jet data from ATLAS 7 TeV

default correlations: terrible χ^2
(correlations across rapidity bins)

decorrelation models: improve the fit a lot

$n_{\rm dat}$	default	part. decorr.	full decorr.
140	1.89	1.28	0.83

no significant effect on the extracted gluon
similar gluon irrespective of the rapidity bin



[EPJ C78 (2018) 248; EPJ C80 (2020) 797]

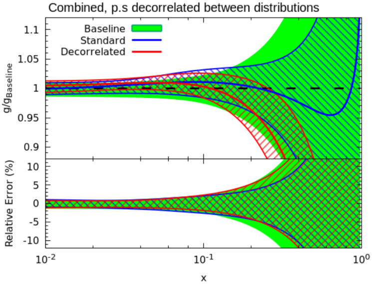
Top pair production from ATLAS 8 TeV

default correlations: terrible χ^2
(correlations across different spectra)

decorrelation models: improve the fit a lot

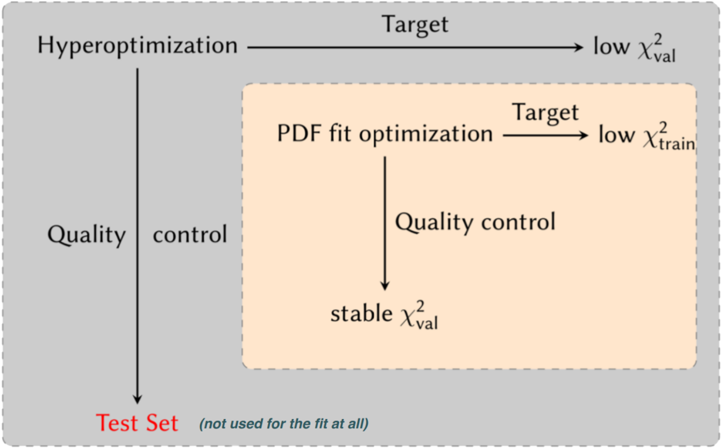
$n_{\rm dat}$	default	stat. uncorr.	p.s. uncorr
25	7.00	3.28	1.80

appreciable effect on the extracted gluon
different gluon depending on the top spectrum



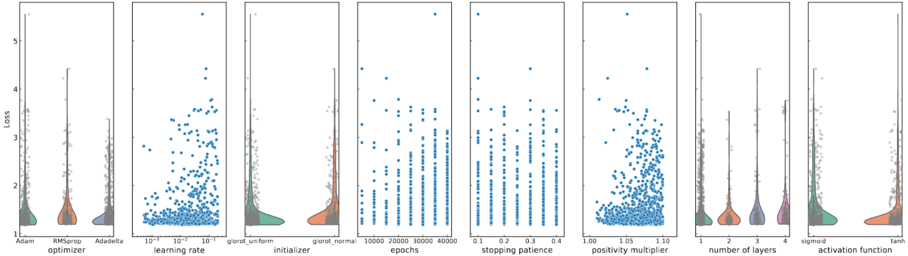
[EPJ C80 (2020) 1; Les Houches proceedings, 2019]

Fitting the methodology



Compare to a Test Set (new set of data previously not used at all)
Who picks the Test Set? Automatic generalisation based on K foldings
Divide the data into n representative sets, fit $n-1$ sets and use n-th set as test set
Hyperoptimise on mean and standard deviation of $\chi^2_{\text{test},i}$, $i = 1 \ldots n$

Hyperoptimisation



SCAN parameter space; **OPTIMISE** $\chi^2_{\rm val}$; **BAYESIAN UPDATING**

Hyperoptimisation requires to define a reward (or loss) function to grade each model

This is different from the cost function (optimised separately for each model)

cost function: $C = E_{\rm tr}$ reward function: $R = \frac{1}{2}(E_{\rm val} + E_{\rm test})$

In a hyperparameter scan one compares the performance of hundreds parameter combinations

Some parameters are discrete (type of minimiser), other are continuous (learning rate)

One should visualise which parameters are relevant and which parameters are immaterial

The *violin* plots are the KDE-reconstructed probability distributions for the hyperparameters

Hyperoptimisation successfully validated in closure tests and in future tests

Validation of PDF uncertainties

Data region: closure tests

Fit PDFs to pseudodata generated assuming a known underlying law

Define bias and variance

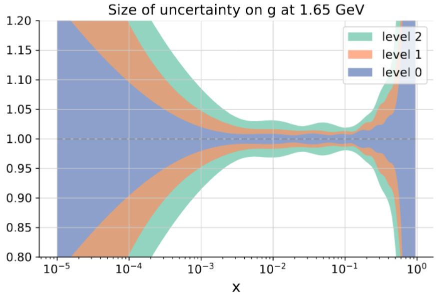
bias difference of central prediction and truth

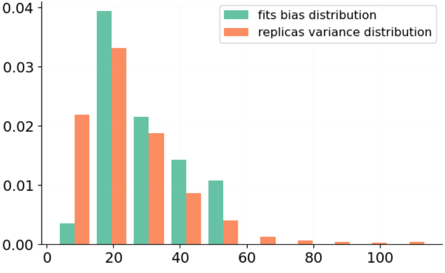
variance uncertainty of replica predictions

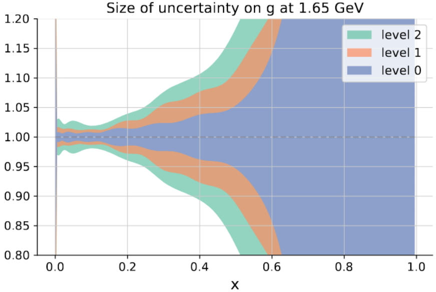
If PDF uncertainty faithful, then

E[bias] = variance

25 fits, 40 replicas each

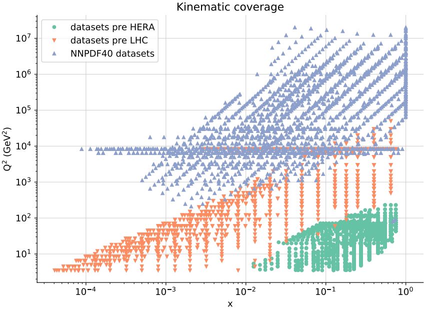






[EPJ C77 (2017) 663; arXiv:2111.05787]

Validation of PDF uncertainties

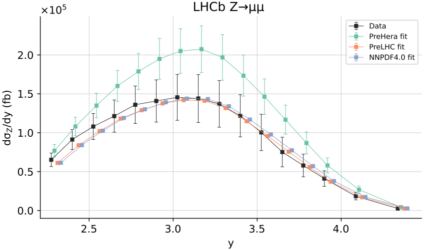


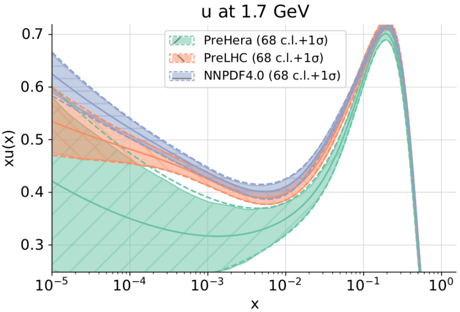
Extrapolation regions: future test

Test PDF uncertainties on data sets not included in a given PDF fit that cover unseen kinematic regions

Data set	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

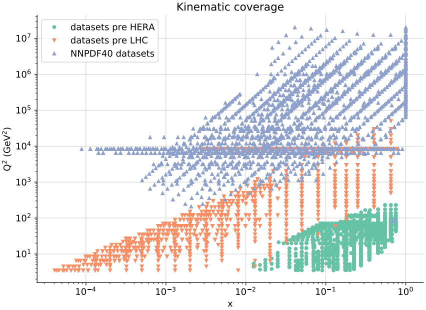
Only exp. cov. matrix





[Acta Phys.Polon. B52 (2021) 243]

Validation of PDF uncertainties

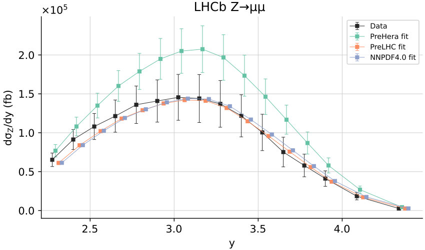


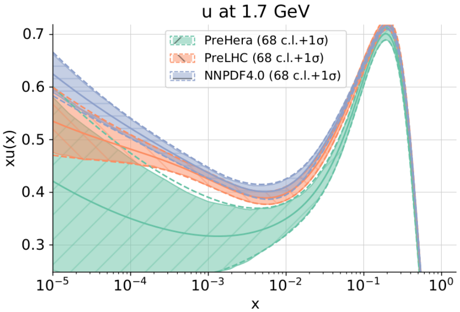
Extrapolation regions: future test

Test PDF uncertainties on data sets
not included in a given PDF fit
that cover unseen kinematic regions

Data set	NNPDF4.0	pre-LHC	pre-HERA
pre-HERA			0.86
pre-LHC		1.17	1.22
NNPDF4.0	1.12	1.30	1.38

Exp+PDF cov. matrix

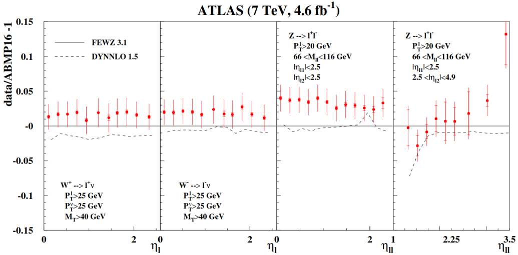




[Acta Phys.Polon. B52 (2021) 243]

Benchmarks

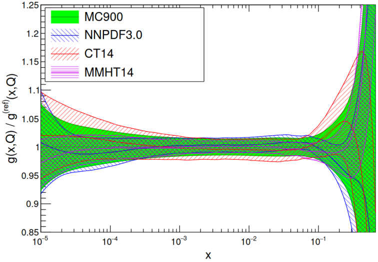
Benchmark of the theory

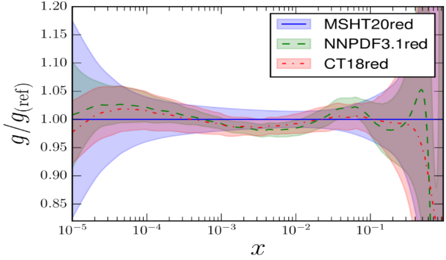


Be careful about the use of different NNLO codes for DY production in particular when experiments use non-optimal fiducial cuts [arXiv:2104.02400]

NNLO corrections usually implemented via K-factors
NNLOJet/ApplFast provide NNLO lookup tables for a limited set of data

Benchmark of PDF sets





[PDF4LHC15 combination, JPG 43 (2016) 023001]

[PDF4LHC21 benchmark, ongoing]

3. Theoretical challenges

Theory uncertainties in PDF determination

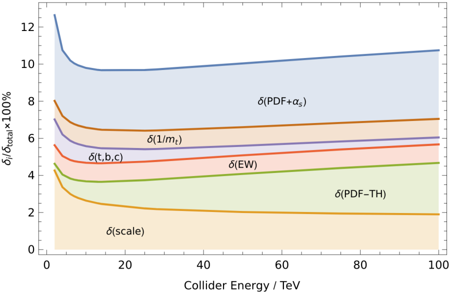
NNLO is the precision frontier for (unpolarized) PDF determination

N3LO is the precision frontier for partonic cross sections at the LHC

Mismatch between perturbative order of partonic cross sections and accuracy of PDFs is becoming a significant source of uncertainty

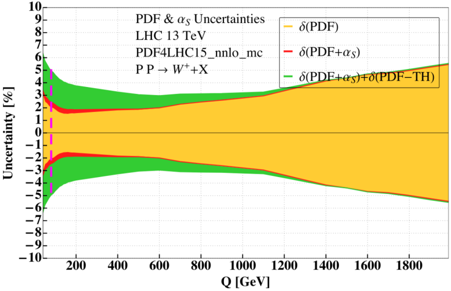
$$\hat{\sigma} = \alpha_s^p \hat{\sigma}_0 + \alpha_s^{p+1}\hat{\sigma}_1 + \alpha_s^{p+2}\hat{\sigma}_2 + \mathcal{O}(\alpha_s^{p+3}) \qquad \delta(\mathrm{PDF}-\mathrm{TH}) = \frac{1}{2}\left|\frac{\sigma^{(2)}_{\mathrm{NNLO-PDFs}} - \sigma^{(2)}_{\mathrm{NLO-PDFs}}}{\sigma^{(2)}_{\mathrm{NNLO-PDFs}}}\right|$$

Higgs production in gluon-gluon fusion



[CERN Yellow Rep.Monogr. 7 (2019) 221]

W^+ boson production in CC Drell-Yan



[JHEP 11 (2020) 143]

Theory uncertainties in PDF determination

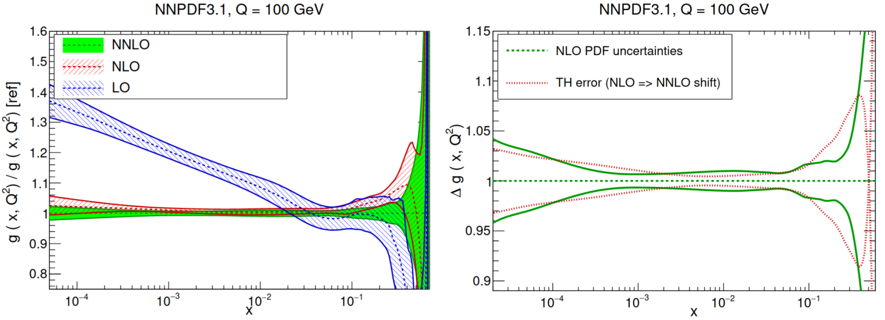
NNLO is the precision frontier for (unpolarized) PDF determination

N3LO is the precision frontier for partonic cross sections at the LHC

Mismatch between perturbative order of partonic cross sections and accuracy of PDFs is becoming a significant source of uncertainty

$$\hat{\sigma} = \alpha_s^p \hat{\sigma}_0 + \alpha_s^{p+1}\hat{\sigma}_1 + \alpha_s^{p+2}\hat{\sigma}_2 + \mathcal{O}(\alpha_s^{p+3}) \qquad \delta(\mathrm{PDF}-\mathrm{TH}) = \frac{1}{2}\left|\frac{\sigma^{(2)}_{\mathrm{NNLO-PDFs}} - \sigma^{(2)}_{\mathrm{NLO-PDFs}}}{\sigma^{(2)}_{\mathrm{NNLO-PDFs}}}\right|$$

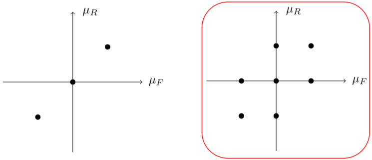
Perturbative stability and uncertainty of the gluon PDF



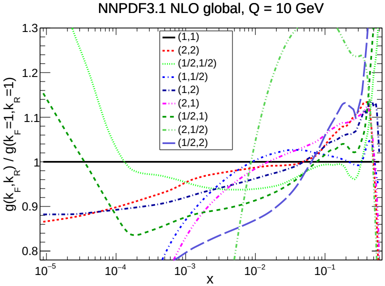
[EPJ C77 (2017) 663]

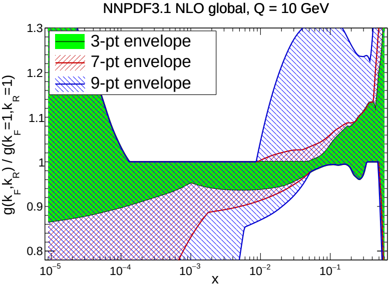
Fits with varied scales

Standard technique to estimate MHOU:
vary scales by 2 and $1/2$, and compute observables for various scale combinations



HXSWG recommendation





Useful for estimating MHOUs in PDFs but want to include them in fitting methodology, and in a way that is also applicable to other theoretical uncertainties

[EPJ C79 (2019) 931]

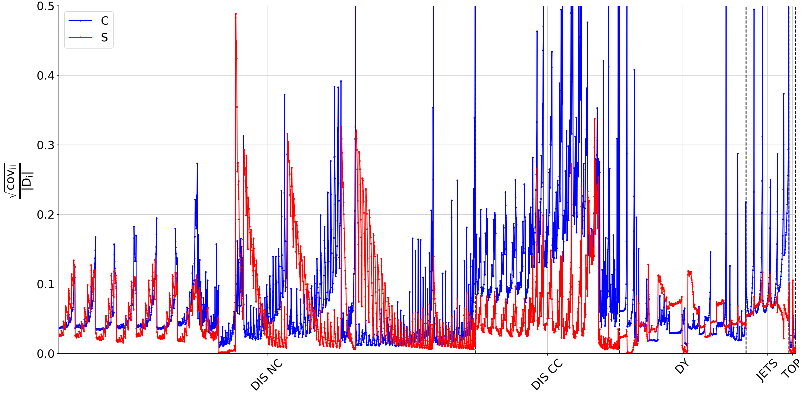
Nuisance parameters

Assuming that theory uncertainties are (a) Gaussian and (b) independent from experimental uncertainties, modify the figure of merit to account for theory errors

$$\chi^2 = \sum_{i,j}^{N_{\rm dat}} (D_i - T_i)(C+S)^{-1}_{ij}(D_j - T_j);\ ({\rm cov}_{\rm th})_{ij} = \frac{1}{N}\sum_k^N \Delta_i^{(k)}\Delta_j^{(k)};\ \Delta_i^{(k)} \equiv T_i^{(k)} - T_i$$

Problem reduced to estimate the th. cov. matrix, *e.g.* in terms of nuisance parameters

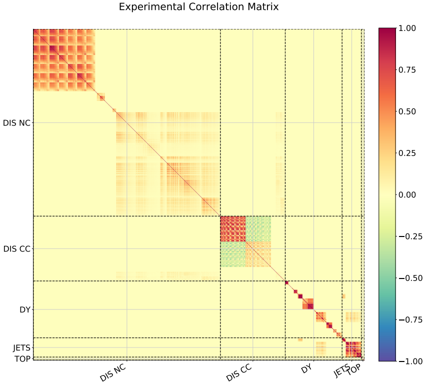
$$\Delta_i^{(k)} = T_i(\mu_R, \mu_F) - T_i(\mu_{R,0}, \mu_{F,0});\ \text{vary scales in } \tfrac{1}{2} \le \frac{\mu_F}{\mu_{F,0}}, \frac{\mu_R}{\mu_{R,0}} \le 2$$

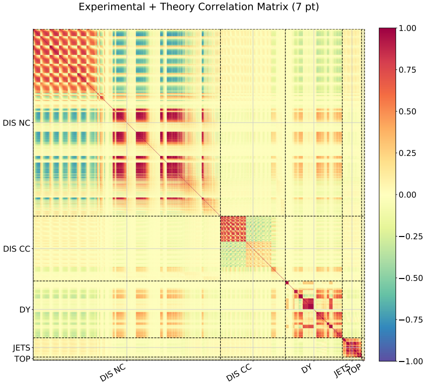


A theory covariance matrix for MHOU

One process
$$S_{ij}^{(9\mathrm{pt})} = \frac{1}{4}\{\Delta_i^{+0}\Delta_j^{+0} + \Delta_i^{-0}\Delta_j^{-0} + \Delta_i^{0+}\Delta_j^{0+} + \Delta_i^{0-}\Delta_j^{0-} + \Delta_i^{++}\Delta_j^{++} + \Delta_i^{+-}\Delta_j^{+-} + \Delta_i^{-+}\Delta_j^{-+} + \Delta_i^{--}\Delta_j^{--}\}$$

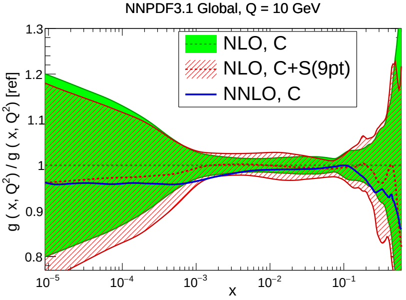
Two processes
$$S_{i_1 j_2}^{(7\mathrm{pt})} = \frac{1}{24}\{2(\Delta_{i_1}^{+0}+\Delta_{i_1}^{++}+\Delta_{i_1}^{+-})(\Delta_{j_2}^{+0}+\Delta_{j_2}^{++}+\Delta_{j_2}^{+-})+2(\Delta_{i_1}^{-0}+\Delta_{i_1}^{-+}+\Delta_{i_1}^{--})(\Delta_{j_2}^{-0}+\Delta_{j_2}^{-+}+\Delta_{j_2}^{--})\}+3(\Delta_{i_1}^{0+}+\Delta_{i_1}^{0-})(\Delta_{j_2}^{0+}+\Delta_{j_2}^{0-})\}$$

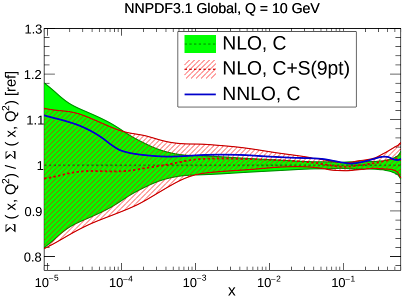




[EPJ C79 (2019) 931; ibid. 838]

Impact on Parton Distributions





PDF uncertainty increase encapsulates NLO-NNLO shift

Overall (rather small) increase in uncertainties

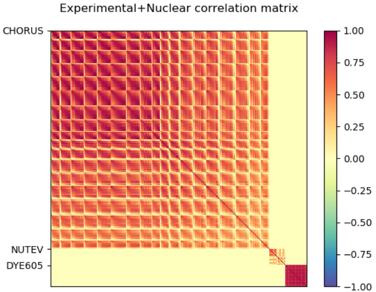
Increase in PDF uncertainties due to replica generation
is counteracted by extra correlations in fitting minimisation

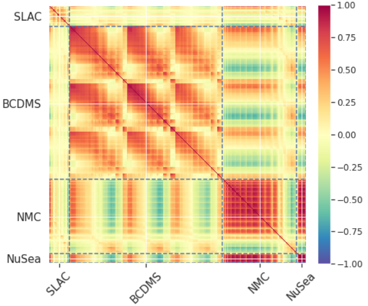
Tensions relieved: improvement in χ^2
exp only: $\chi^2/N_{\rm dat} = 1.139$ exp+th: $\chi^2/N_{\rm dat} = 1.110$

Data whose theoretical descrition is affected by large scale uncertainties
are deweighted in favour of more perturbatively stable data

[EPJ C79 (2019) 931; ibid. 838]

Theory uncertainty in PDF determination

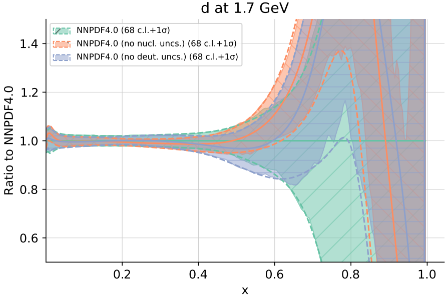




Effect of nuclear uncertainties relevant at large x
to reconcile FT DIS with LHC DY data

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

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



[EPJ C79 (2019) 282; EPJ C81 (2021) 37]

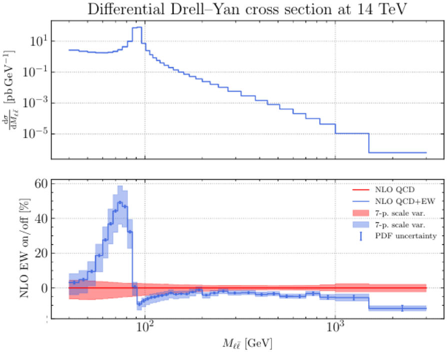
NLO EW corrections in PDF determination

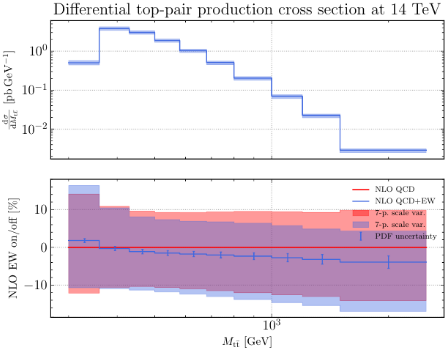
If we aim to PDF accurate to 1% NLO EW corrections do matter

especially as higher invariant mass and transverse momentum regions are accessed

Different approaches taken in general-purpose PDF fits

NLO EW K-factors (MSHT20); no NLO EW corrections by default (NNPDF4.0)





QED corrections in DGLAP evolution
[Com.Phys.Commun. 185 (2014) 1647]

Photon PDF
[PRL 117 (2016) 242002; JHEP 12 (2017) 046]

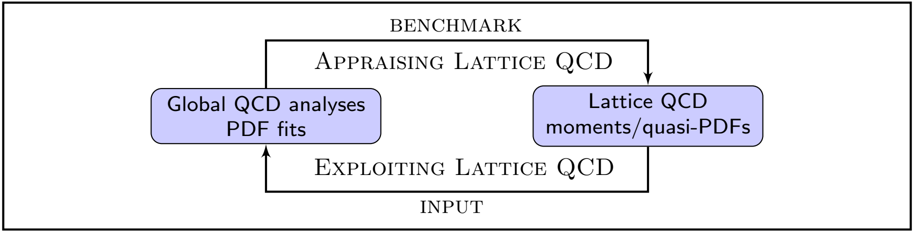
Photon PDF fits à la LuxQED
[SciPost Phys. 5 (2019) 1; JHEP 79 (2019) 10]

Automation of NLO EW corrections
[JHEP 07 (2018) 185]

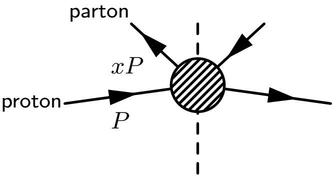
Fast interpolation grids: PineAPPL
[JHEP 12 (2020 108]

Careful scrutiny of data
(no FSR nor photon-initiated subtraction)

Input from Lattice QCD



[Prog.Part.Nucl.Phys. 100 (2018) 107; ibid. 121 (2021) 103908]



collinear transition of a massles proton h
into a massless parton i
with fractional momentum x
local OPE $\Longrightarrow$ lattice formulation

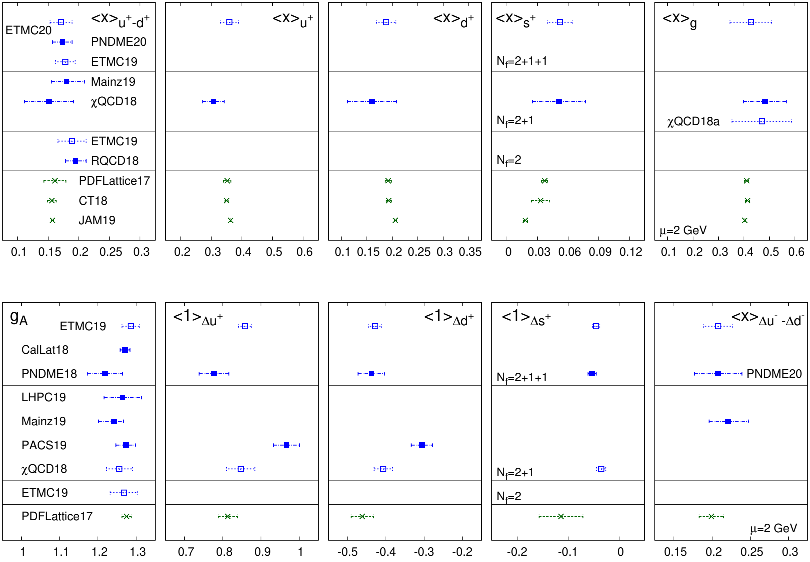
$$q(x) = \frac{1}{4\pi}\int dy^- e^{-iy^- xP^+}\langle P|\bar{\psi}(0,y^-,\mathbf{0}_\perp)\gamma^+\psi(0)|P\rangle$$

$$\Delta q(x) = \frac{1}{4\pi}\int dy^- e^{-iy^- xP^+}\langle P,S|\bar{\psi}(0,y^-,\mathbf{0}_\perp)\gamma^+\gamma^5\psi(0)|P,S\rangle$$

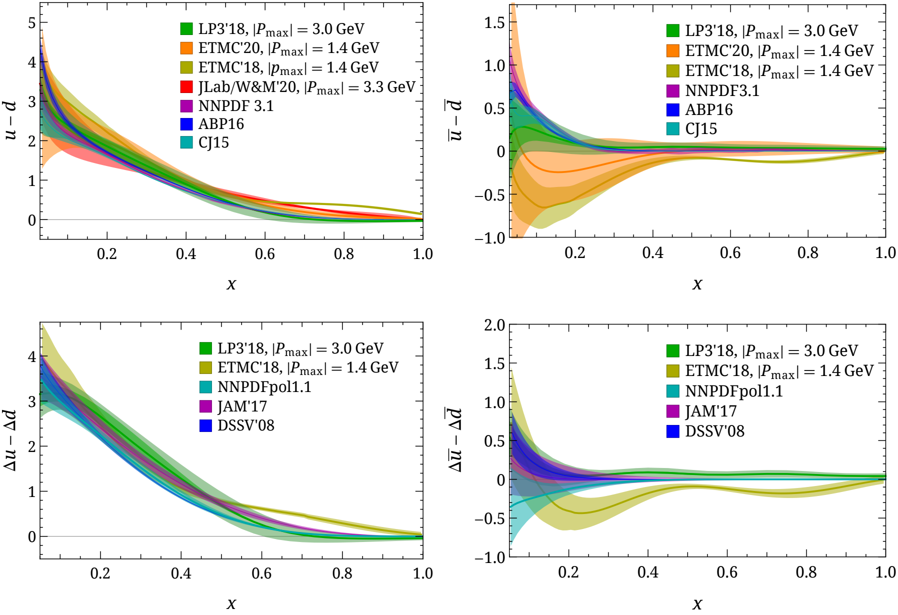
$$y=(y^+,y^-,\mathbf{y}_\perp)\,,\qquad y^+=(y^0+y^z)/\sqrt{2}\,,\qquad y^-=(y^0-y^z)/\sqrt{2}\,,\qquad \mathbf{y}_\perp=(v^x,v^y)$$

Moments of Parton Distributions

$$\langle 1\rangle_f = \int_0^1 dx f \qquad \langle x\rangle_f = \int_0^1 x dx f$$



Quasi-PDFs and Pseudo-PDFs



Can New Physics hide in the Proton?

Replace the SM $\longrightarrow$ SMEFT Lagrangian

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i^{N_{\rm d6}} \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_j^{N_{\rm d8}} \frac{c_j}{\Lambda^4}\mathcal{O}_j^{(8)}$$

assumption: new physic states are heavy
the Lagrangian contains only light SM particles

BSM effects included as a momentum expansion

based on all SM symmetries

number of couplings reduced by symmetries

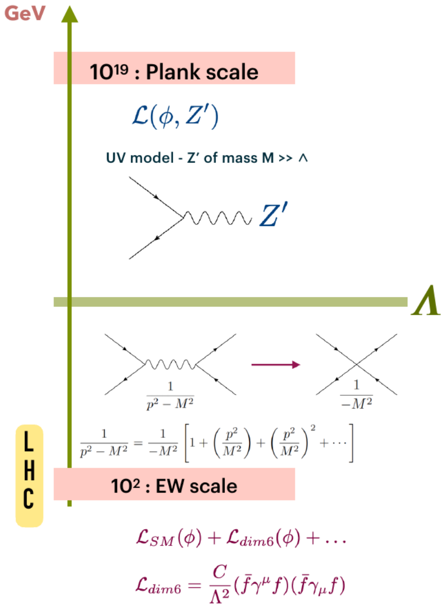
Can Wilson coefficients be determined from LHC data?

Can BSM effects be reabsorbed into PDFs?

$$\sigma_{\rm LHC} = f_i^{\rm SM} \otimes f_j^{\rm SM} \otimes \tilde{\sigma}_{\rm SM}$$

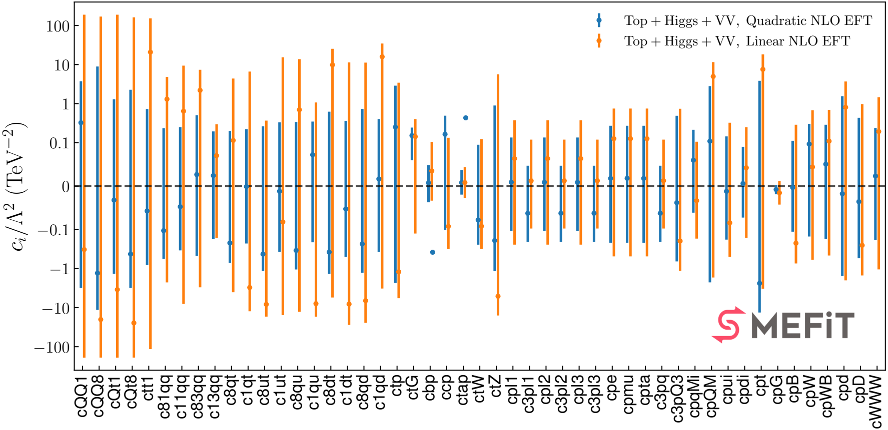
$$\sigma_{\rm LHC} = f_i^{\rm SMEFT} \otimes f_j^{\rm SMEFT} \otimes \tilde{\sigma}_{\rm SM} \times K$$

$$K = 1 + \sum_m^{N_{\rm d6}} c_m \frac{\kappa_m}{\Lambda^2} + \sum_{m,n}^{N_{\rm d6}} c_m c_n \frac{\kappa_{mn}}{\Lambda^4}$$



Can New Physics hide in the Proton?

Can Wilson coefficients be determined from LHC data?



[JHEP 04 (2019) 100; JHEP 11 (2021) 089]

Can BSM effects be reabsorbed into PDFs?

[PRL 123 (2019) 132001; JHEP 07 (2021) 122; Maria Ubiali's seminar in March 2021]

4. To conclude

Summary

A precise and accurate determination of PDFs is key to do discovery.

Collider measurements are reducing PDF uncertainties to few percent.

This opens up some challenges.

Understand experimental systematic uncertainties and their correlations.

Refine the theoretical accuracy of a PDF determination.

Represent theory uncertainties into PDF uncertainties.

Deploy a robust fitting methodology and good statistical tests of it.

Benchmark efforts may benefit from public releases of PDF codes and inputs.

Summary

A precise and accurate determination of PDFs is key to do discovery.

Collider measurements are reducing PDF uncertainties to few percent.

This opens up some challenges.

Understand experimental systematic uncertainties and their correlations.

Refine the theoretical accuracy of a PDF determination.

Represent theory uncertainties into PDF uncertainties.

Deploy a robust fitting methodology and good statistical tests of it.

Benchmark efforts may benefit from public releases of PDF codes and inputs.

Thank you