Proton PDFs: connections between small and large x

Synergies between the EIC and the LHC

Emanuele R. Nocera
Università degli Studi di Torino and INFN, Torino

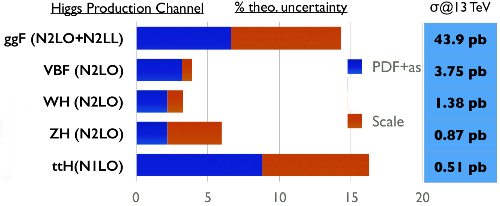
14 December 2023

PDFs at the LHC

$$\sigma(Q^2,\tau,\mathbf{k}) = \sum_{ij}\int_{\tau}^{1}\frac{dz}{z}\mathcal{L}_{ij}(z,Q^2)\hat{\sigma}_{ij}\left(\frac{\tau}{z},\alpha_s(Q^2),\mathbf{k}\right) \quad \mathcal{L}_{ij}(z,Q^2) = (f_i^{h_1}\otimes f_j^{h_2})(z,Q^2)$$

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

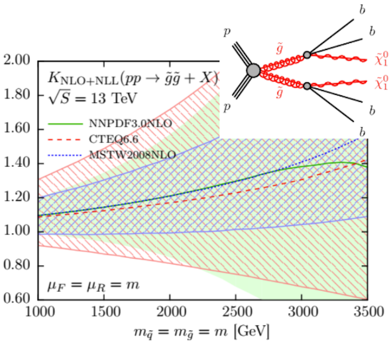
Precision



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

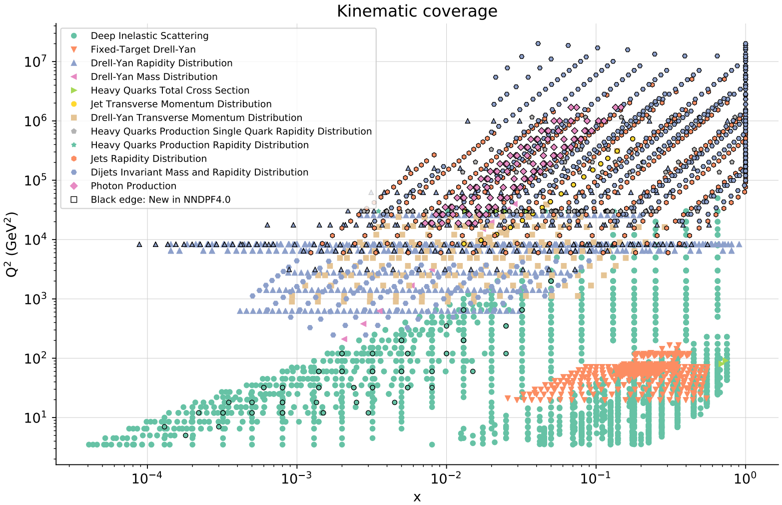
[Plot from the CERN Yellow Report 2016]

Discovery



[EPJC 76 (2016) 53]

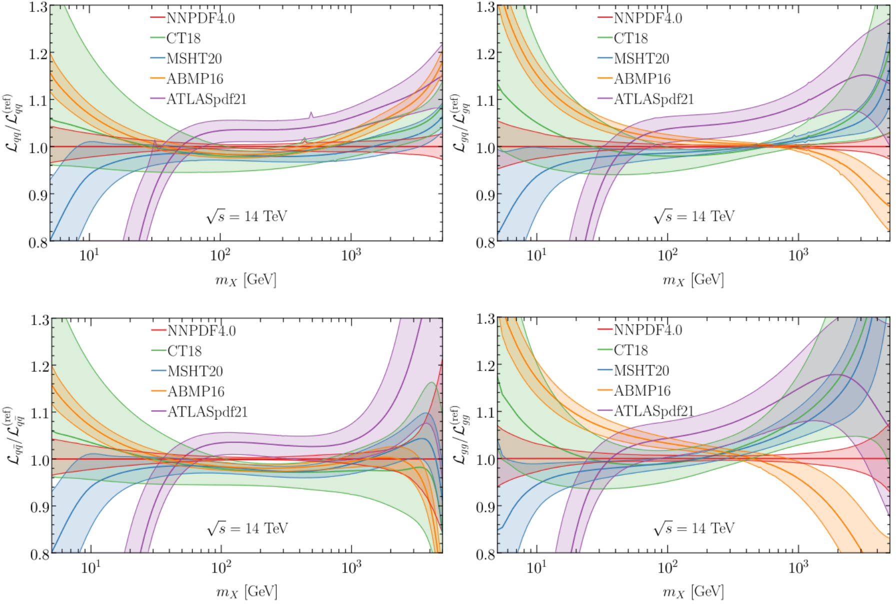
Experimental data



The inverse problem of PDF determination is addressed by parametric regression

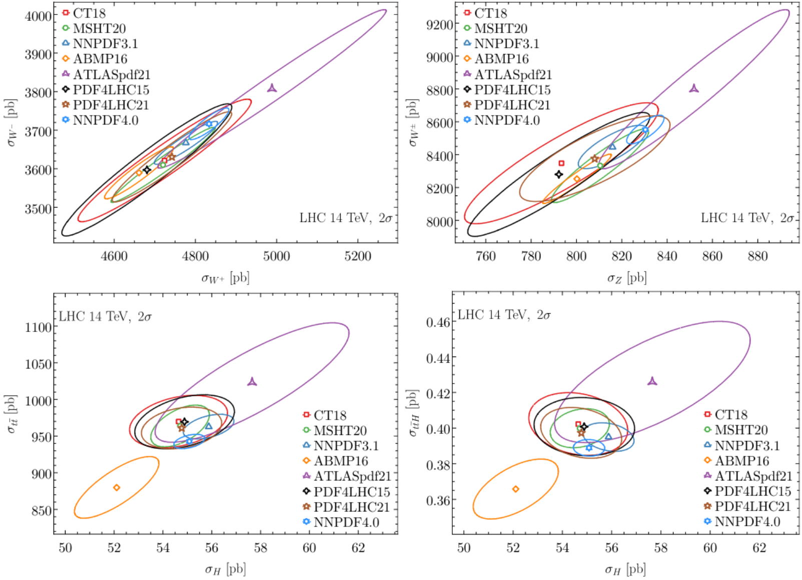
NNPDF4.0 (NNLO) $\quad N_{\rm dat} = 4618 \quad \chi^2/N_{\rm dat} = 1.16$

Comparing PDF sets



[Acta Phys.Polon.B 53 (2022) 12]

Making predictions with PDFs



[Acta Phys.Polon.B 53 (2022) 12]

Validation of PDF uncertainties

Data region: closure tests

Fit PDFs to pseudodata generated assuming a known underlying law

Define bias and variance

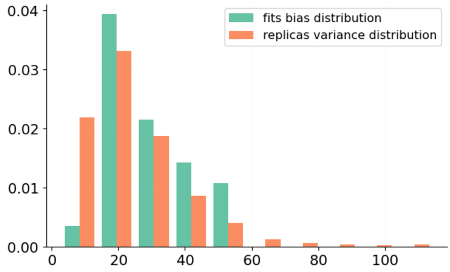
bias difference of central prediction and truth

variance uncertainty of replica predictions

If PDF uncertainty faithful, then

E[bias] = variance

25 fits, 40 replicas each



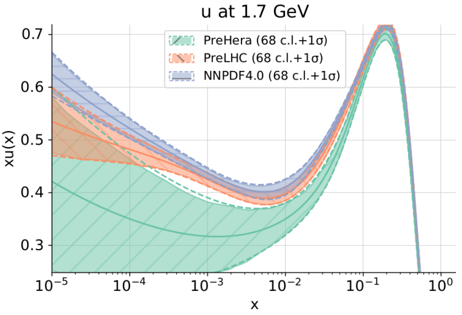
[EPJ C77 (2017) 663; EPJ C82 (2022) 330]

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

Data region: closure tests

Fit PDFs to pseudodata generated assuming a known underlying law

Define bias and variance

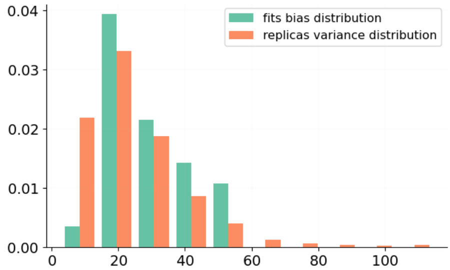
bias difference of central prediction and truth

variance uncertainty of replica predictions

If PDF uncertainty faithful, then

E[bias] = variance

25 fits, 40 replicas each



[EPJ C77 (2017) 663; EPJ C82 (2022) 330]

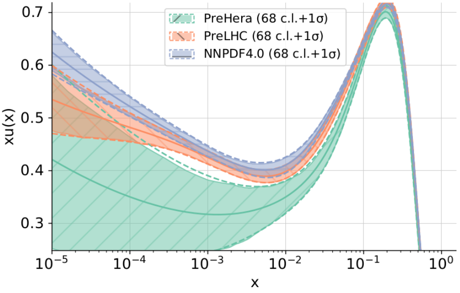
Extrapolation regions: future test

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

u at 1.7 GeV

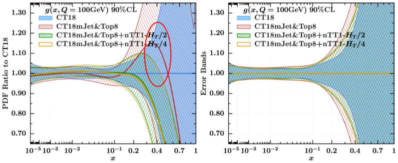


[Acta Phys.Polon. B52 (2021) 243]

1. Data

Gluon

Global fit without jet and $t\bar{t}$ data vs CT18NNLO



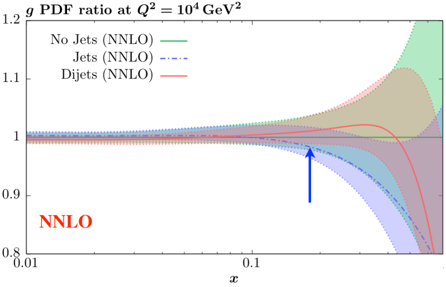
[M. Guzzi, PDF4LHC Nov. 2023]

Various processes (included in all PDF sets)
Z p_T, jets, di-jets, $t\bar{t}$

Largest impact of jets/di-jets at large x

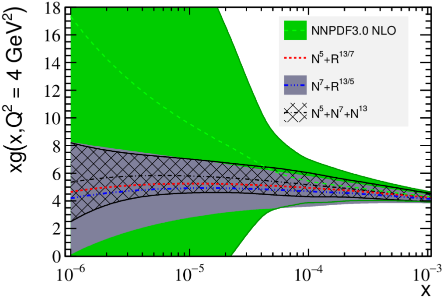
Di-jets preferred over single-inclusive jets

Forward charm production impacts small x potentially crucial for UHE neutrino-nucleus cross section measurements



[L. Harland-Lang, PDF4LHC Nov. 2023]

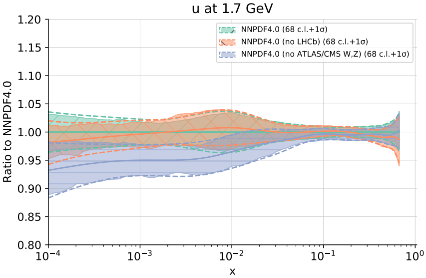
[See also EPJ C80 (2020) 797]

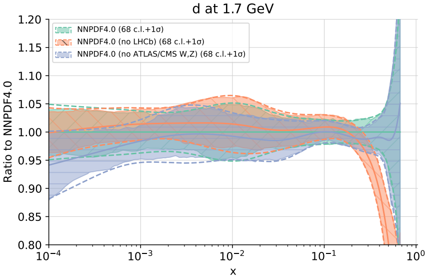


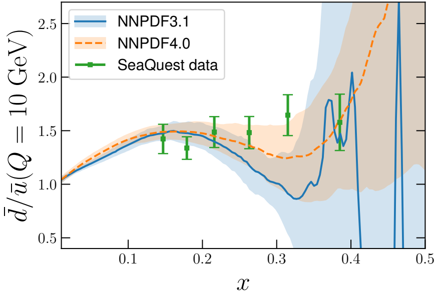
[PRL 118 (2017) 072001]

[See also arXiv:2303.13607]

Quark flavour separation







Relative impact of ATLAS/CMS/LHCb gauge boson production

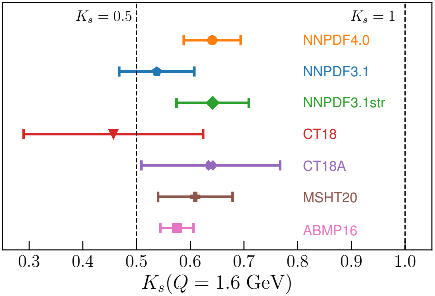
LHCb is at forward rapidity

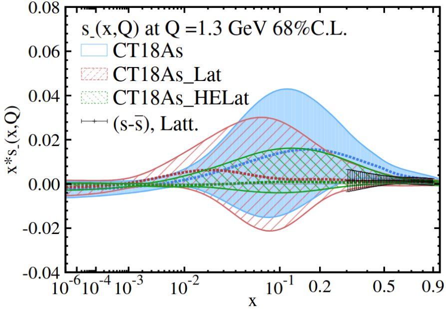
New constraint on $\bar{d}/\bar{u}$ ratio from SeaQuest

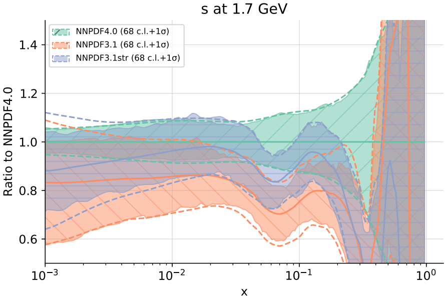
[Nature 590 (2021) 561]

Studied by CT, MSHT, NNPDF, ABMP
Some tension with NuSea found

Strange







Good consistency of K_s across PDF sets

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

Effect of data and nuclear uncertainties

ATLAS W, Z and W+jet data enhance s

NOMAD data reduce uncertainties

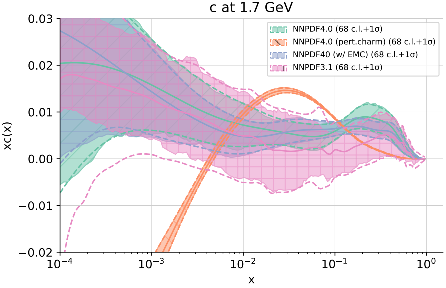
nuclear uncertainties accommodate data sets

Useful input from lattice QCD

[EPJ C80 (2020) 1168; PRD 107 (2023) 076018]

[See also PRD 91 (2015) 094002]

Charm



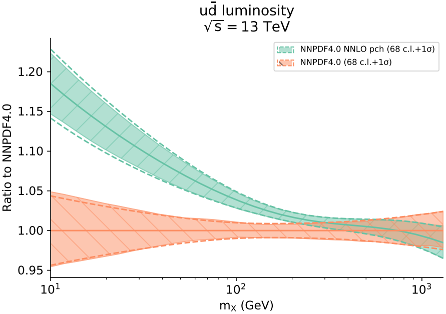
Perturbative charm alters the flavour decomposition and deteriorates the fit

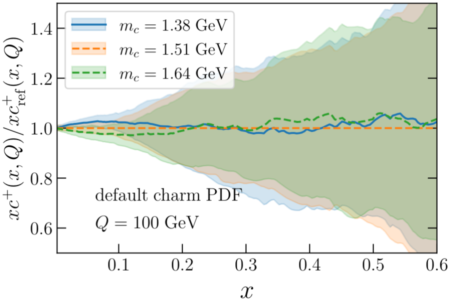
$$\chi^2_{\text{fitted charm}} = 1.17 \to \chi^2_{\text{pert. charm}} = 1.19$$

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

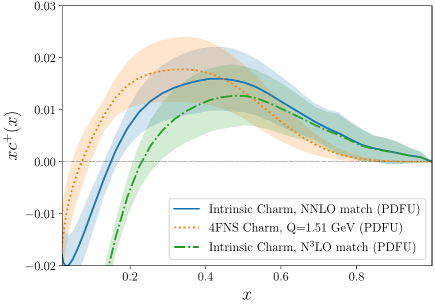
fitting charm reduces the dependence from m_c

[EPJ C76 (2016) 647; C77 (2017) 663; C82 (2022) 428]





Intrinsic Charm



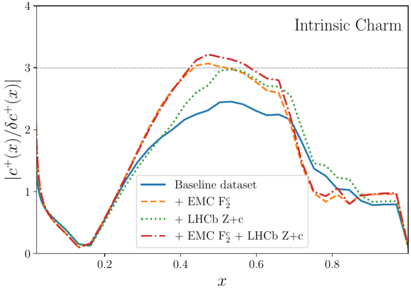
Evolve results backwards (below m_c) with N^{3}LO matching

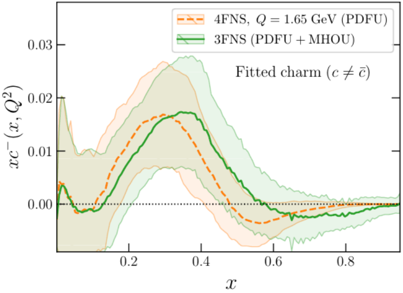
Evidence of intrinsic charm and of $c - \bar{c}$ shape compatible with models

[Nature 608 (2022) 483; arXiv:2311.00743]

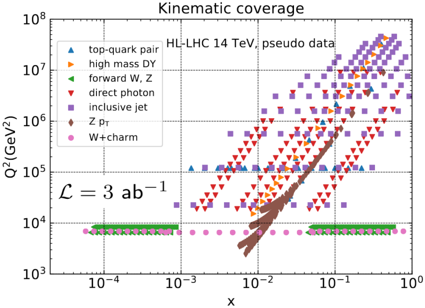
Evidence enhanced by EMC F_2^c and $Z + D$

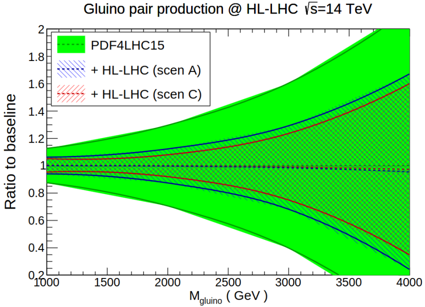
Challenged by CT18 [PLB 843 (2023) 137975]

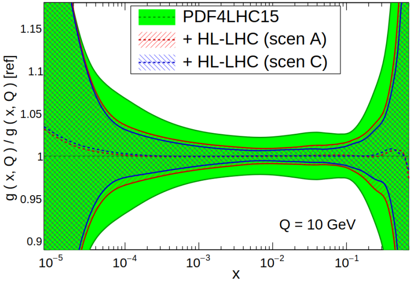


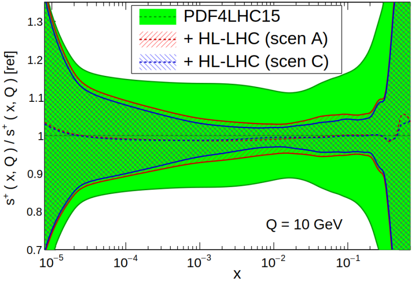


Impact of future data: HL-LHC



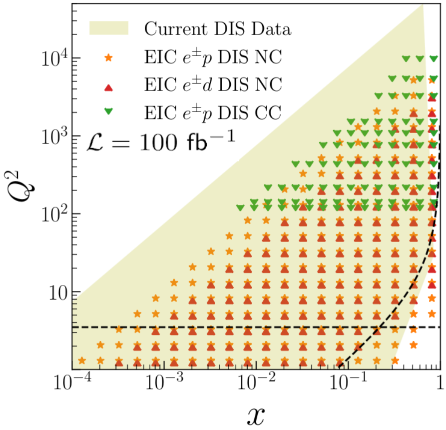


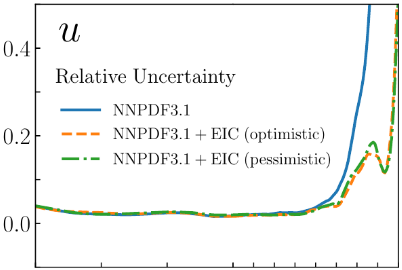




[EPJ.C 78 (2018) 962]

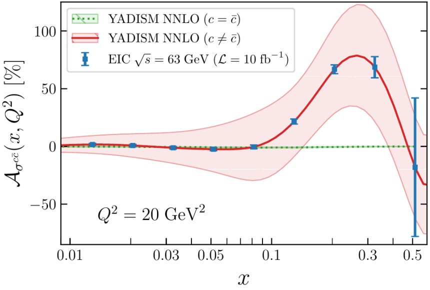
Impact of future data: EIC





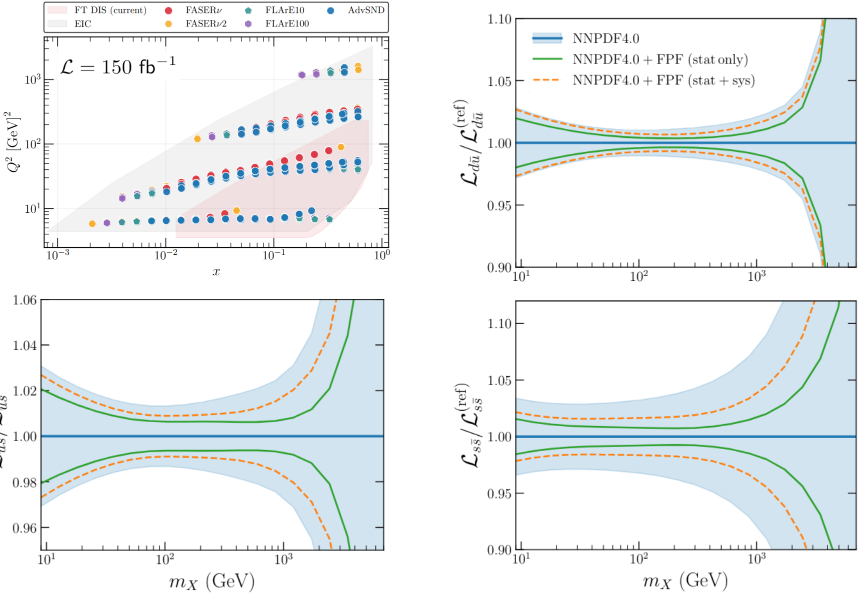
$E_\ell \times E_p$ [GeV]: 18× 275; 10× 100; 5× 100

$$\mathcal{A}_{\sigma^{c\bar{c}}} = \frac{\sigma^c_{\rm red} - \sigma^{\bar{c}}_{\rm red}}{\sigma^{c\bar{c}}_{\rm red}}$$



[PRD 103 (2021) 096005; see also arXiv:; arXiv:2311.00743]

Impact of future data: FPF



[arXiv:2309.09581; see T. Mäkelä's talk]

2. Theory

N^{3}LO QCD corrections in PDF determination

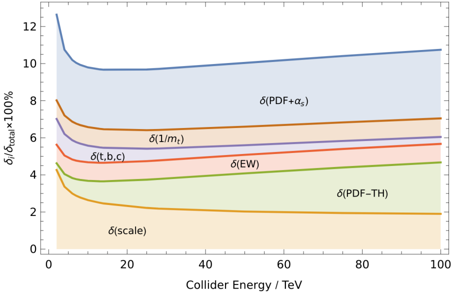
NNLO is the precision frontier for PDF determination

N3LO is the precision frontier for partonic cross sections

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

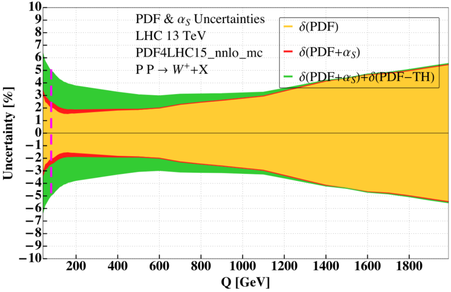
$$\hat{\sigma} = \alpha_s^p \hat{\sigma}_0 + \alpha_s^{p+1}\hat{\sigma}_1 + \alpha_s^{p+2}\hat{\sigma}_2 + \mathcal{O}(\alpha_s^{p+3}) \qquad \delta(\mathrm{PDF-TH}) = \frac{1}{2}\left|\frac{\sigma^{(2)}_{\mathrm{NNLO-PDFs}} - \sigma^{(2)}_{\mathrm{NLO-PDFs}}}{\sigma^{(2)}_{\mathrm{NNLO-PDFs}}}\right|$$

Higgs production in gluon-gluon fusion



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

W^+ boson production in CC Drell-Yan



[JHEP 11 (2020) 143]

N^3LO QCD corrections in PDF determination [See also G. Falcioni's talk]

Splitting Functions

Singlet (P_{qq}, P_{gg}, P_{gq}, P_{qg})

– large-n_f limit [NPB 915 (2017) 335; arXiv:2308.07958]

– small-x limit [JHEP 06 (2018) 145]

– large-x limit [NPB 832 (2010) 152; JHEP 04 (2020) 018; JHEP 09 (2022) 155]

– 5 (10) lowest Mellin moments [PLB 825 (2022) 136853; ibid. 842 (2023) 137944; ibid. 846 (2023) 138215]

Non-singlet ($P_{NS,v}$, $P_{NS,+}$, $P_{NS,-}$)

– large-n_f limit [NPB 915 (2017) 335; arXiv:2308.07958]

– small-x limit [JHEP 08 (2022) 135]

– large-x limit [JHEP 10 (2017) 041]

– 8 lowest Mellin moments [JHEP 06 (2018) 073]

DIS structure functions (F_L, F_2, F_3)

– DIS NC (massless) [NPB 492 (1997) 338; PLB 606 (2005) 123; NPB 724 (2005) 3]

– DIS CC (massless) [Nucl.Phys.B 813 (2009) 220]

– massive from parametrisation combining known limits and damping functions [NPB 864 (2012) 399]

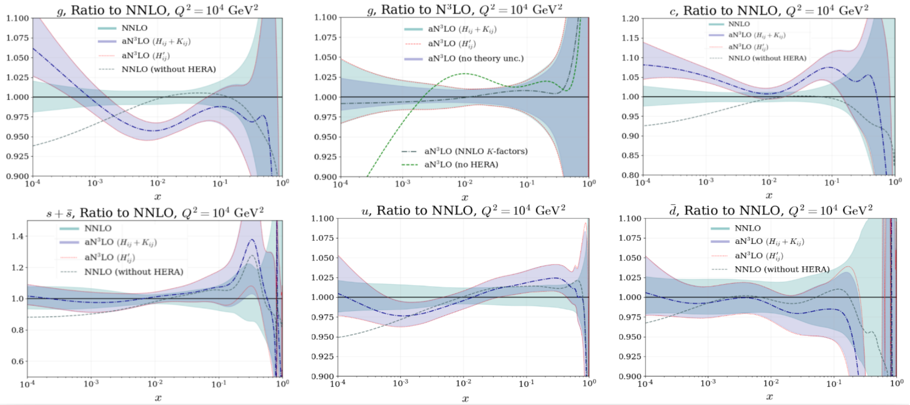
PDF matching conditions

– all known except for $a^3_{H,g}$ [NPB 820 (2009) 417; NPB 886 (2014) 733; JHEP 12 (2022) 134]

Coefficient functions for other processes

– DY (inclusive) [JHEP 11 (2020) 143]; DY (y differential) [PRL 128 (2022) 052001]

aN^{3}LO PDFs — MSHT



[EPJ C83 (2023) 185; see also T. Cridge's talk]

3-5% correction on the gluon PDF at $x \sim 10^{-2}$

larger charm PDF (perturbatively generated)

inclusion of theory uncertainties may inflate PDF uncertainties at small x

inclusion of aN^{3}LO corrections generally improve the χ^2 of HERA and LHC jets

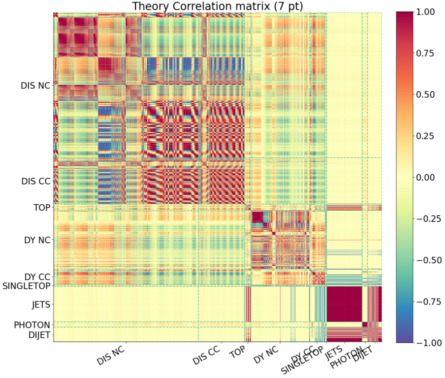
Theory uncertainties in PDF determination

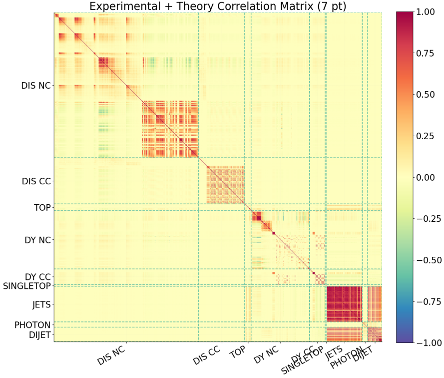
Assuming that theory uncertainties are (a) Gaussian and (b) independent from experimental uncertainties, modify the figure of merit to account for theory errors

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

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

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





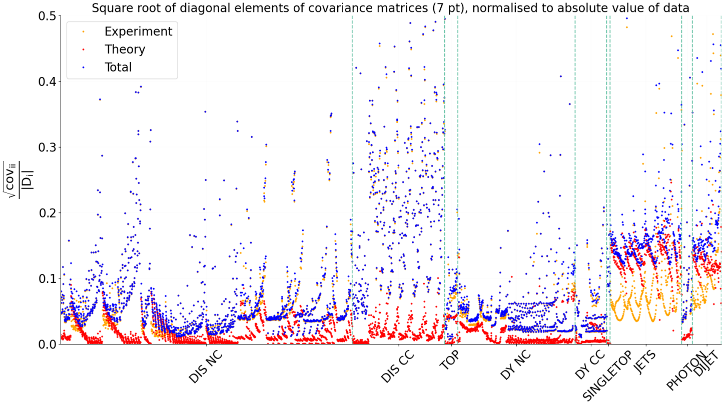
Theory uncertainties in PDF determination

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

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

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

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



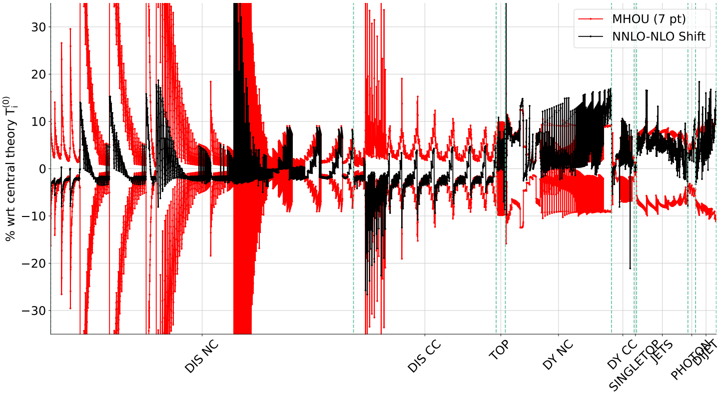
Theory uncertainties in PDF determination

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

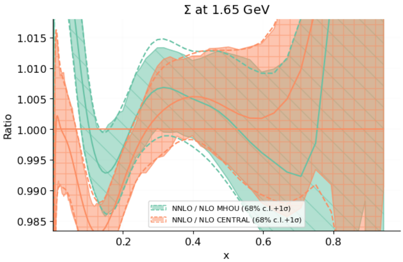
$$\chi^2 = \sum_{i,j}^{N_{\rm dat}} (D_i - T_i)({\rm cov}_{\rm exp} + {\rm cov}_{\rm th})^{-1}_{ij}(D_j - T_j);\ ({\rm cov}_{\rm th})_{ij} = \frac{1}{N}\sum_k^N \Delta_i^{(k)}\Delta_j^{(k)};\ \Delta_i^{(k)} \equiv T_i^{(k)} - T_i$$

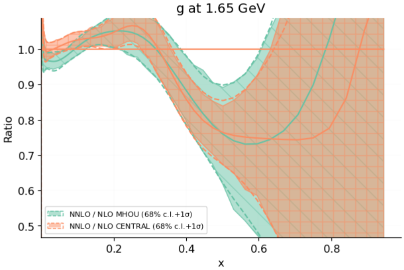
Problem reduced to estimate the th. cov. matrix, *e.g.* in terms of nuisance parameters

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



Theory uncertainties in PDF determination PRELIMINARY





Faster perturbative convergence when MHOU are incorporated into PDFs

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: $\chi^2/N_{\rm dat} = 1.21$ exp+th: $\chi^2/N_{\rm dat} = 1.20$

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

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

What happens at N3LO?

Incomplete higher order uncertainties

Approximate N^{3}LO splitting functions as

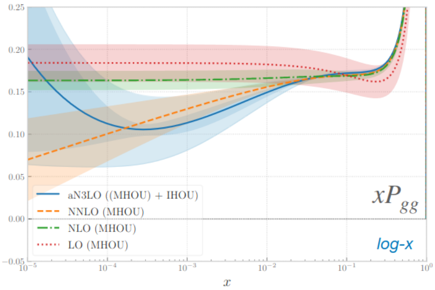
$$\gamma_{ij}^{(3)} = \gamma_{ij,n_f}^{(3)} + \gamma_{ij,N\to\infty}^{(3)} + \gamma_{ij,N\to 0}^{(3)} + \tilde{\gamma}_{ij}^{(3)}$$

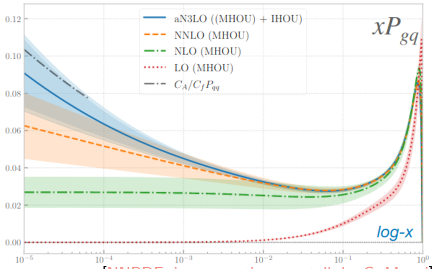
Parametrise $\tilde{\gamma}_{ij}^{(3)} = \sum_l a_{ij}^{(l)} G_l(N)$

– G_1 for the leading unknown large-N term
– G_2 for the leading unknown small-N term
– 3 or 8 G_l for the sub-leading unknown small- and large-N contributions
– vary the functions G_l to generate a variety of approximations and estimate IHOU
– determine the coefficients $a_{ij}^{(l)}$ with known moments and momentum conservation

Adopted basis function for $\tilde{\gamma}_{qg}^{(3)}$

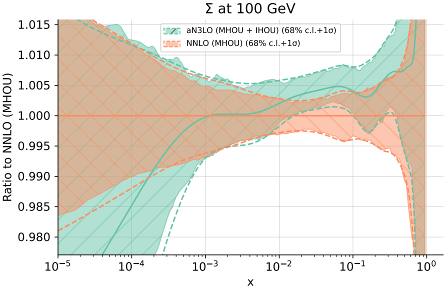
$G_1(N)$	$\mathcal{M}[(1-x)\ln^2(1-x)]$
$G_2(N)$	$-\frac{1}{(N-1)^2} + \frac{1}{N^2}$
$G_3(N)$	$\frac{1}{N^4}$, $\frac{1}{N^3}$, $\mathcal{M}[(1-x)\ln(1-x)]$ $\mathcal{M}[(1-x)^2\ln(1-x)^2]$, $\frac{1}{N-1}-\frac{1}{N}$, $\mathcal{M}[(1-x)\ln(x)]$
$G_4(N)$	$\mathcal{M}[(1-x)(1+2x)]$, $\mathcal{M}[(1-x)x^2]$, $\mathcal{M}[(1-x)x(1+x)]$, $\mathcal{M}[(1-x)]$

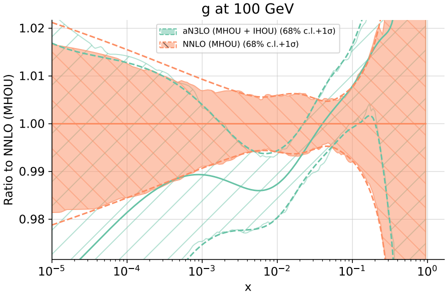




[NNPDF, in preparation; see talk by G. Magni]

aN^{3}LO PDFs — NNPDF PRELIMINARY





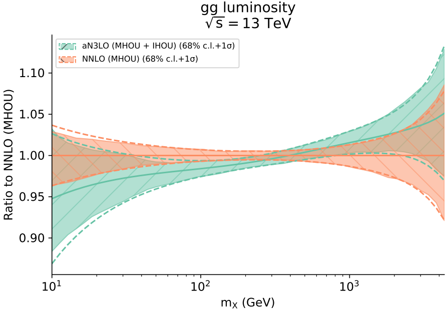
IHOU incorporated into an independent covariance matrix where nuisance parameters are averaged over parametrisation variations

$\chi^2/N_{\rm dat} = 1.20$ (NNLO (MHOU))

$\chi^2/N_{\rm dat} = 1.19$ (aN^{3}LO (MHOU+IHOU))

PDFs only affected at small x

largest effect: 2% suppression in $\mathcal{L}_{gg}$ around the Higgs mass



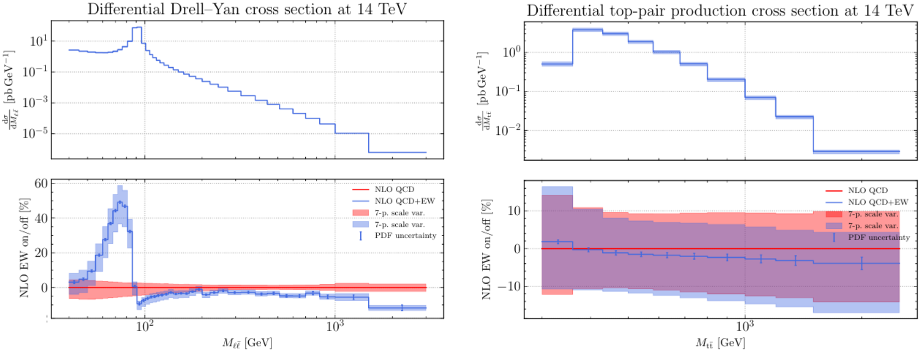
NLO EW corrections in PDF determination

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]

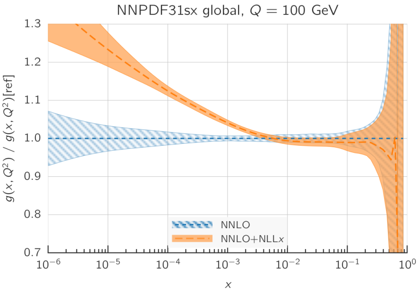
Fast interpolation grids: PineAPPL

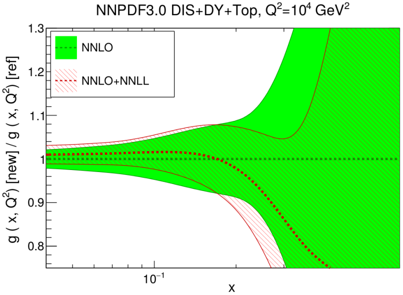
[JHEP 12 (2020 108]

Careful scrutiny of data

(no FSR nor photon-initiated subtraction)

Beyond fixed-order accuracy





small x: $\frac{1}{x}\ln^k x$
high-energy gluon emission: single logs

large x: $\left(\frac{\ln^k(1-x)}{(1-x)}\right)_+$
soft gluon emission: double logs

Large logs $\alpha_s \ln \sim 1$ spoil the convergence of the perturbative series

PDFs with threshold resummation [JHEP 1509 (2015) 191] (only DIS, DY Z/γ, total $t\bar{t}$ + evol.)
suppression in PDFs partially or totally compensates enhancements in partonic cross-sections
accuracy of the resummed fit competitive with the fixed-order fit, except for the large-x gluon

PDFs with high-energy (BFKL) resummation [EPJ C78 (2018) 321] (only DIS + evol.)
Resummed PDFs enhanced at small x, uncertainties reduced, fit quality improves
Large effects for future colliders, or b production at LHC
High-densitiy effects modelled in CT18X; similar outcome on PDFs and fit quality

Fitting away New Physics

DIS [PRL 123 (2019) 132001]

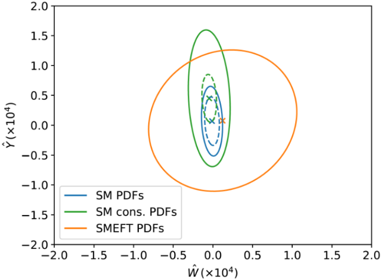
DY tails [JHEP 07 (2021) 122]

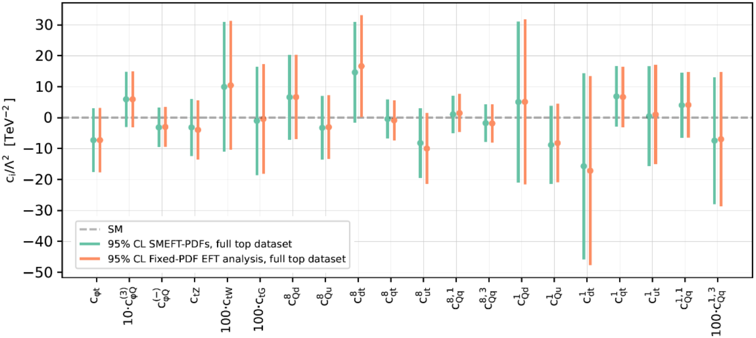
DIS/DY [JHEP 08 (2022) 088]

Jet/top [JHEP 05 (2023) 003]

Jets [JHEP 02 (2022) 142]

Many more analyses by ATLASs, CMS, . . .





3. Conclusions

Summary

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

LHC measurements are being instrumental to reduce PDF uncertainties to few percent.

This is not enough. Good complementarity with other planned facilities.

The goal of achieving PDF determinations accurate to 1% opens up some challenges.

Understand the interplay between data, theory, and methodology into PDF uncertainties.

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 precision phenomenology.

LHC measurements are being instrumental to reduce PDF uncertainties to few percent.

This is not enough. Good complementarity with other planned facilities.

The goal of achieving PDF determinations accurate to 1% opens up some challenges.

Understand the interplay between data, theory, and methodology into PDF uncertainties.

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