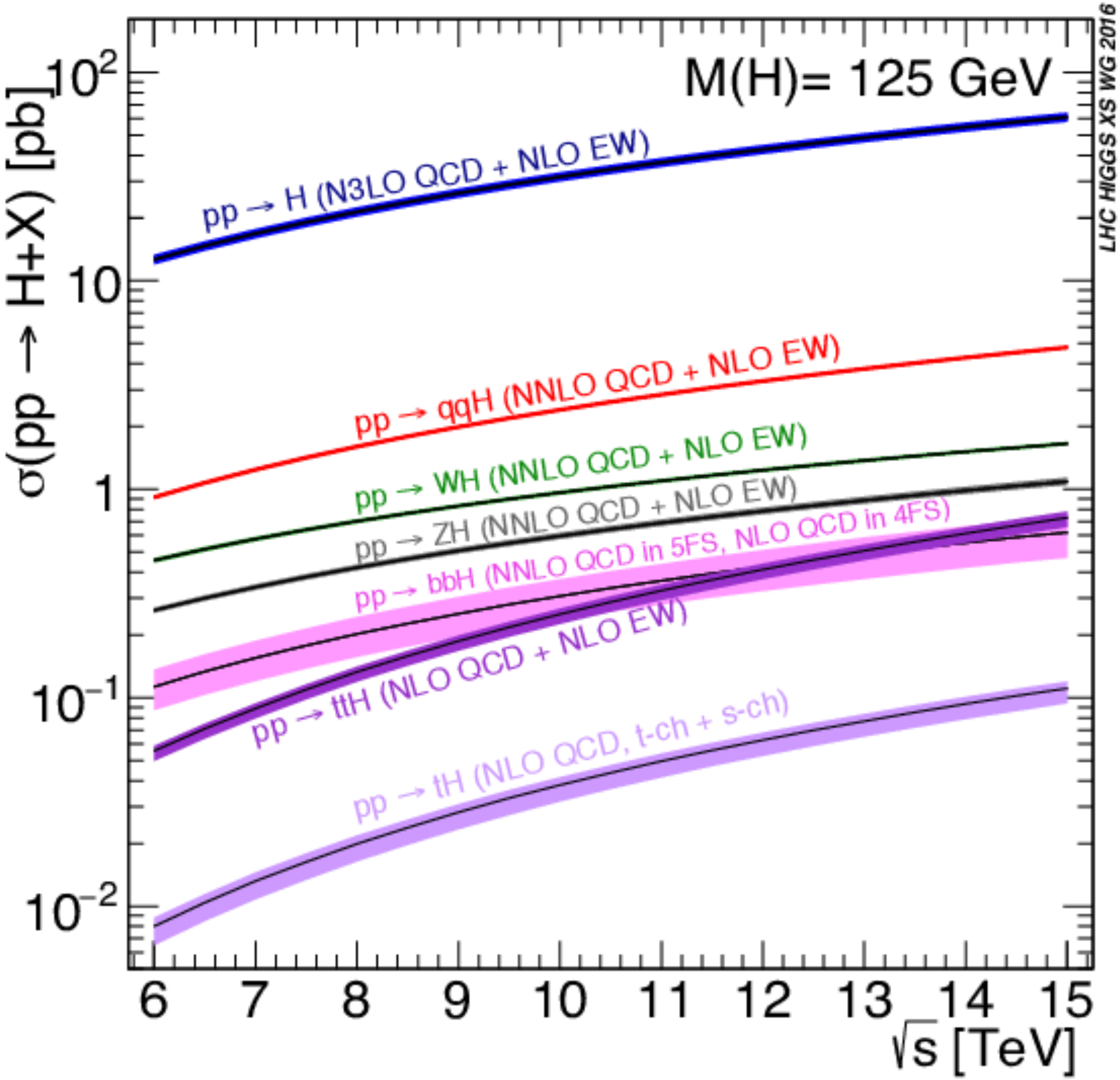
PDFs for Higgs Boson Production at the LHC

Juan Rojo, VU Amsterdam & Nikhef

Higgs Hunting, Paris, 15th July 2025

PDFs and the Higgs Boson

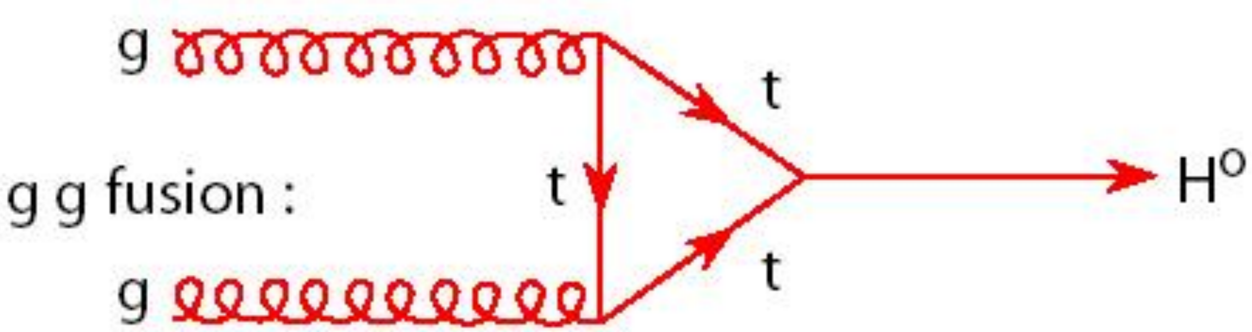
Higgs production and PDFs



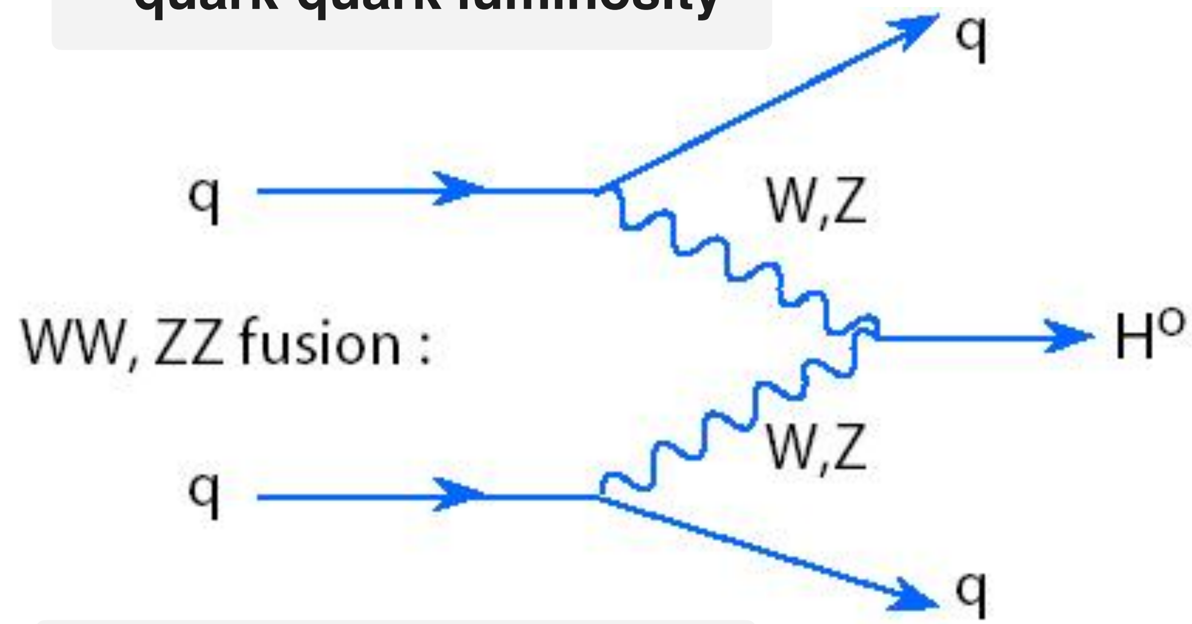
Higgs production and PDFs

Higgs production cross-sections depend sensitively on the **input PDFs**

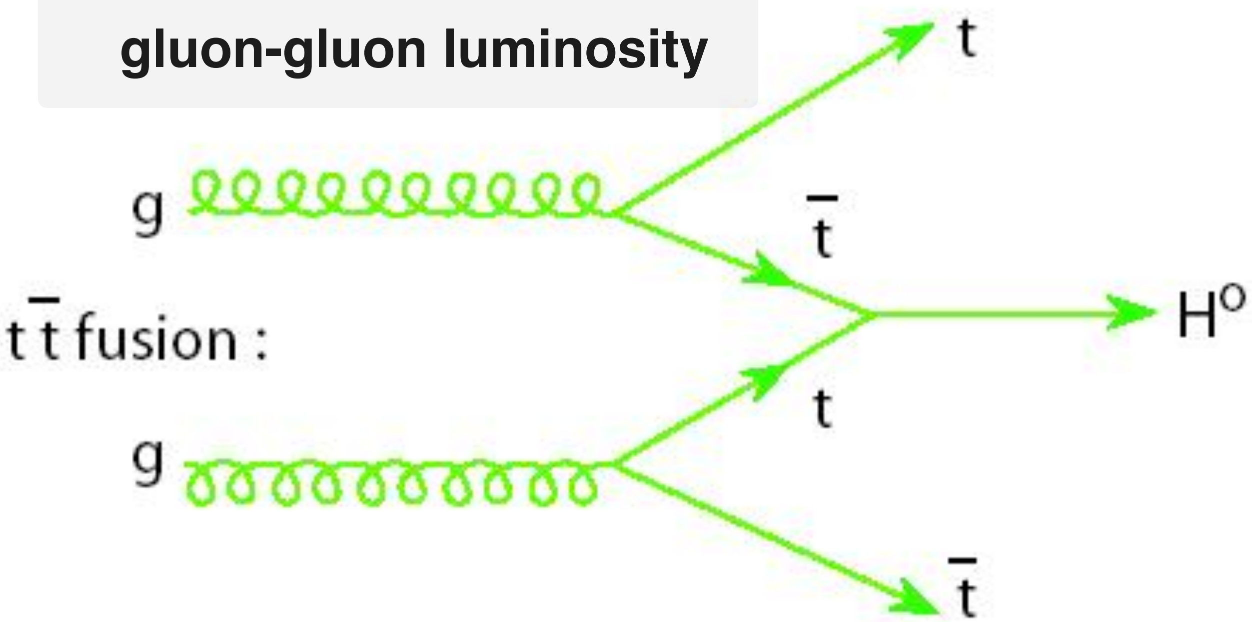
gluon-gluon luminosity



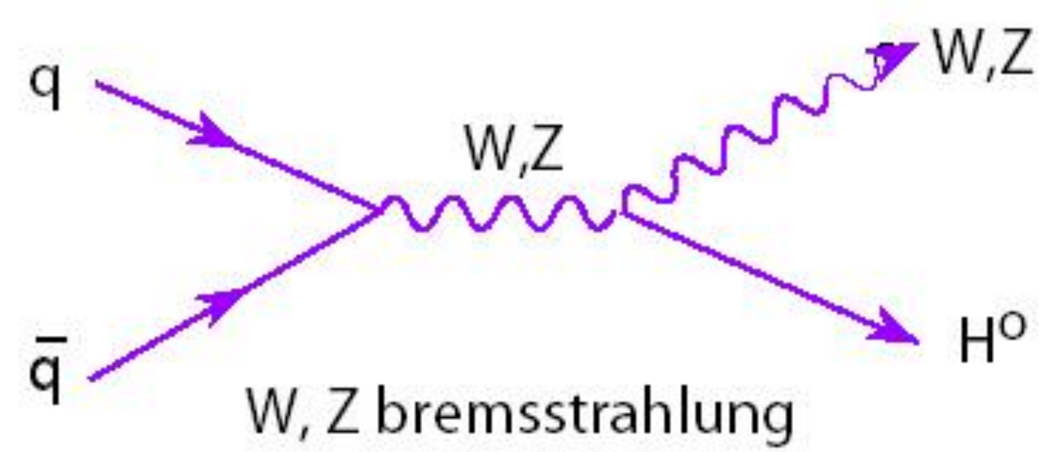
quark-quark luminosity



gluon-gluon luminosity



quark-antiquark luminosity

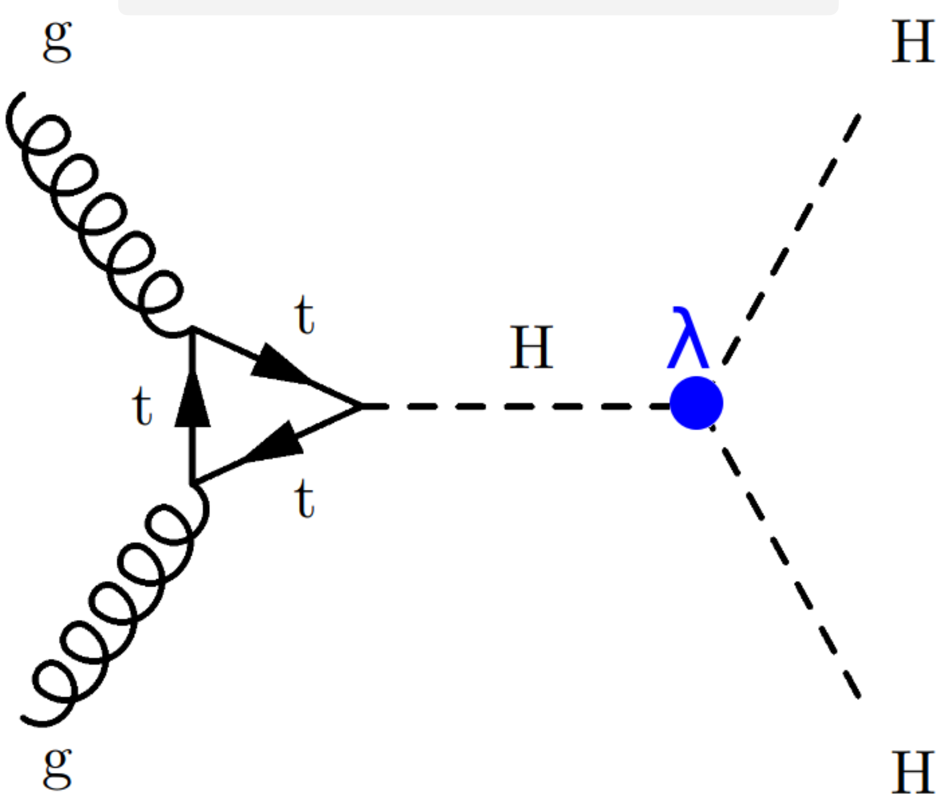


NNLO PDFs standard since many years, **aN³LO PDFs** represent now accuracy frontier

Higgs production and PDFs

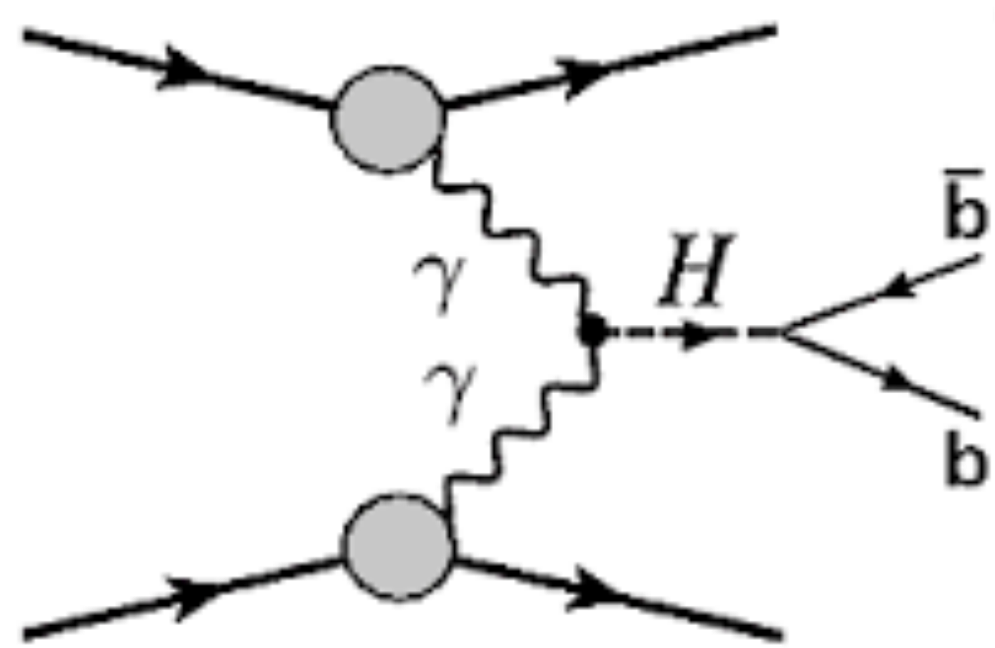
Higgs production cross-sections depend sensitively on the **input PDFs**

gluon-gluon luminosity



Higgs pair production

photon-photon luminosity



Photon-initiated Higgs production relevant once QED and EW effects accounted for

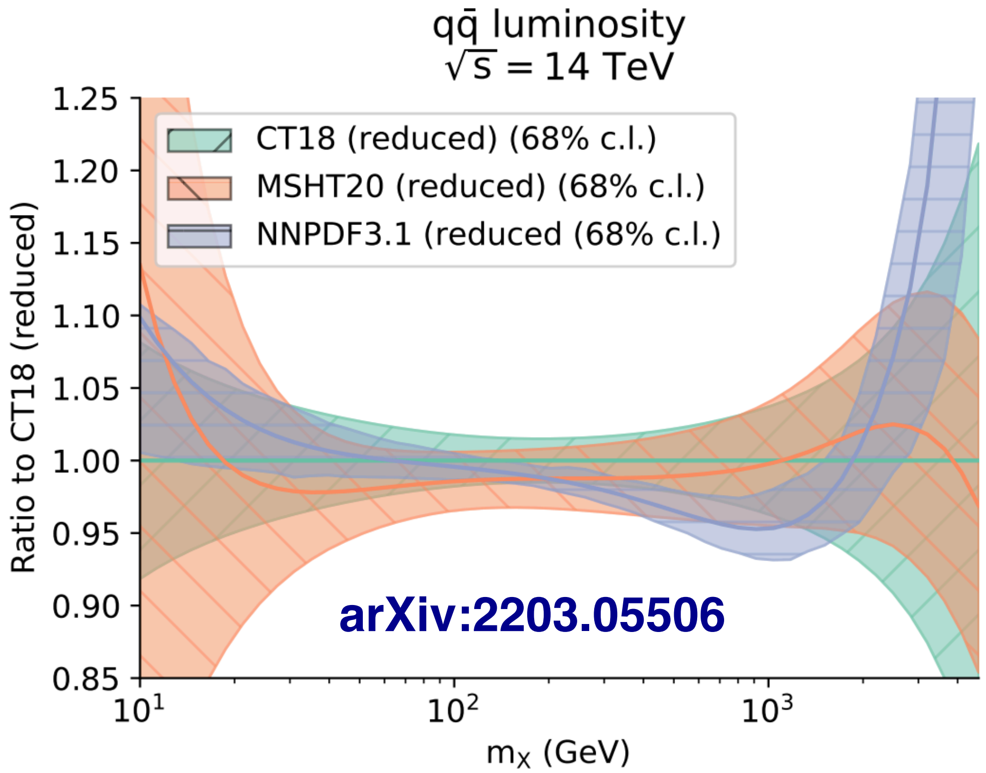
Improving our understanding of PDFs and reducing their uncertainties is essential for accurate **Higgs boson characterisation** e.g. through SMEFT interactions

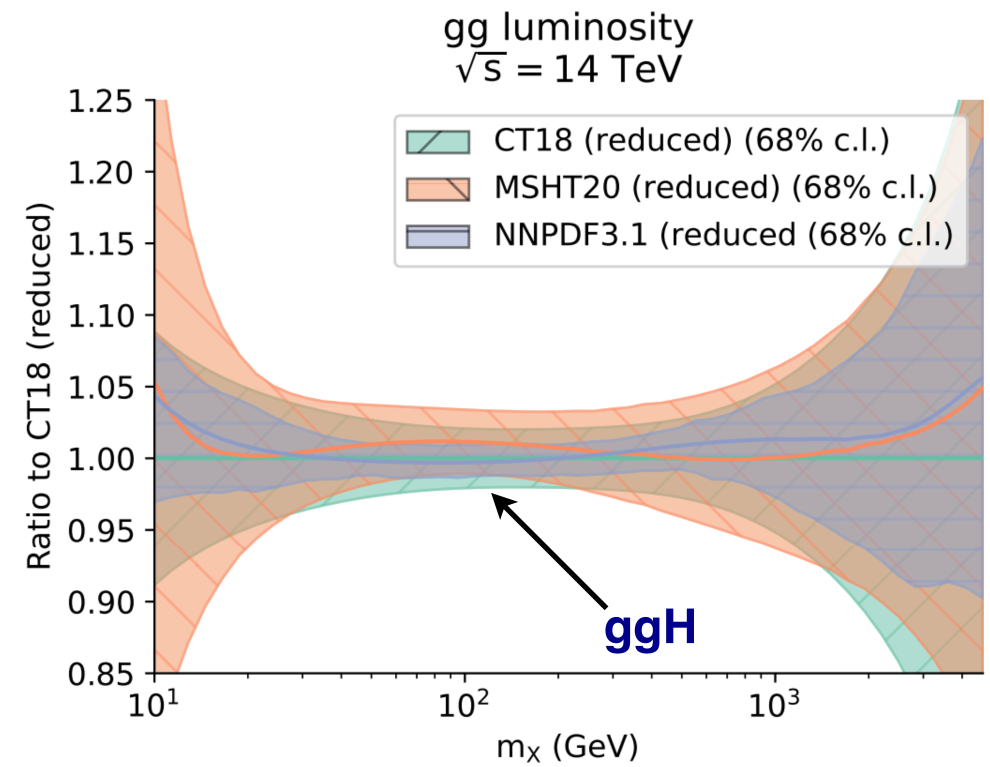
Progress in global PDF fits

	NNPDF4.0	CT18	MSHT20	ABMP16
Dataset	Global	Global	Global	Global (no jet data)
Heavy quark treatment	GM-VFN & fitted charm	GM-VFN & perturbative chram	GM-VFN & perturbative chram	FFN & perturbative charm
Perturbative accuracy	aN^3LO_{QCD} + NLO_{QED} & MHOUs	$NNLO_{QCD}$ + LO_{QED}	aN^3LO_{QCD} + NLO_{QED} & MHOUs	$NNLO_{QCD}$
Methodology	Neural Networks & Monte Carlo replicas	Fixed parametrisation & Hessian (w. tolerance)	Fixed parametrisation & Hessian (w. tolerance)	Fixed parametrisation & Hessian (no tolerance)

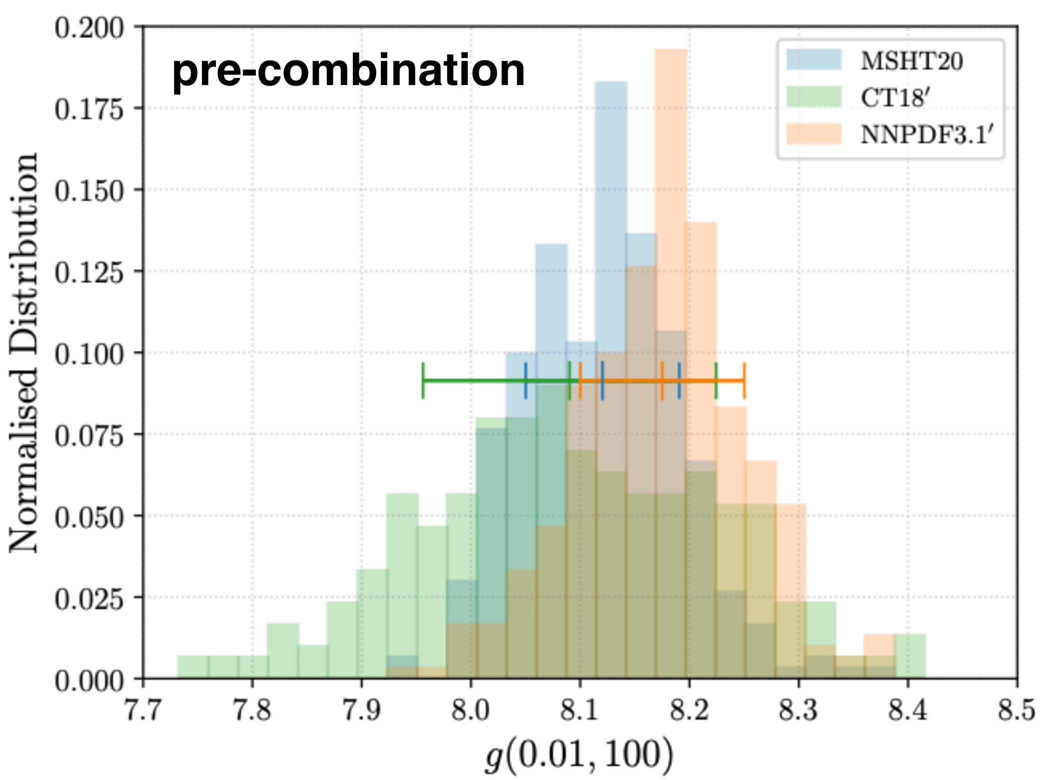
The PDF4LHC21 benchmark & combination

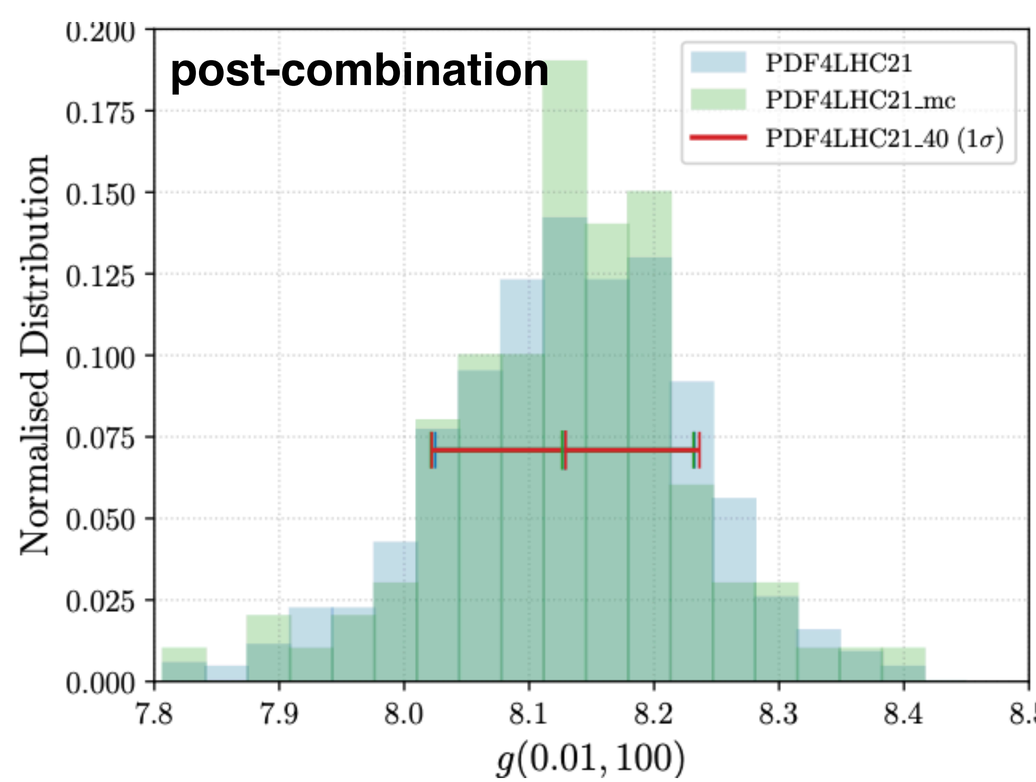
- CT18, MSHT20, NNPDF3.1 fitted to **common dataset**: excellent agreement within errors



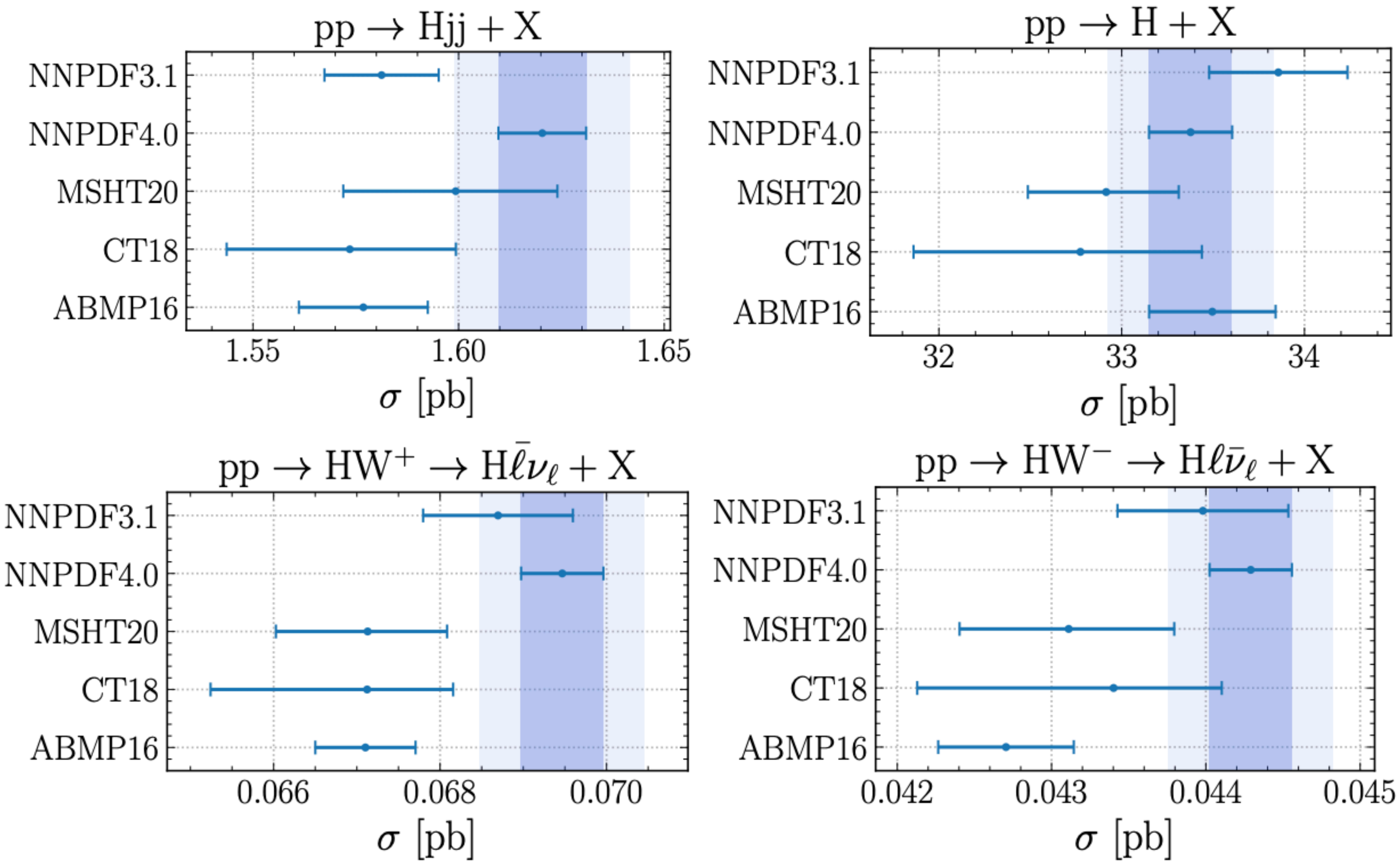


- **PDF4LHC21**: combination of CT18, MSHT20, NNPDF3.1 NNLO PDFs with coherent theory settings, each group using their preferred input dataset**. Current HXSWG baseline**

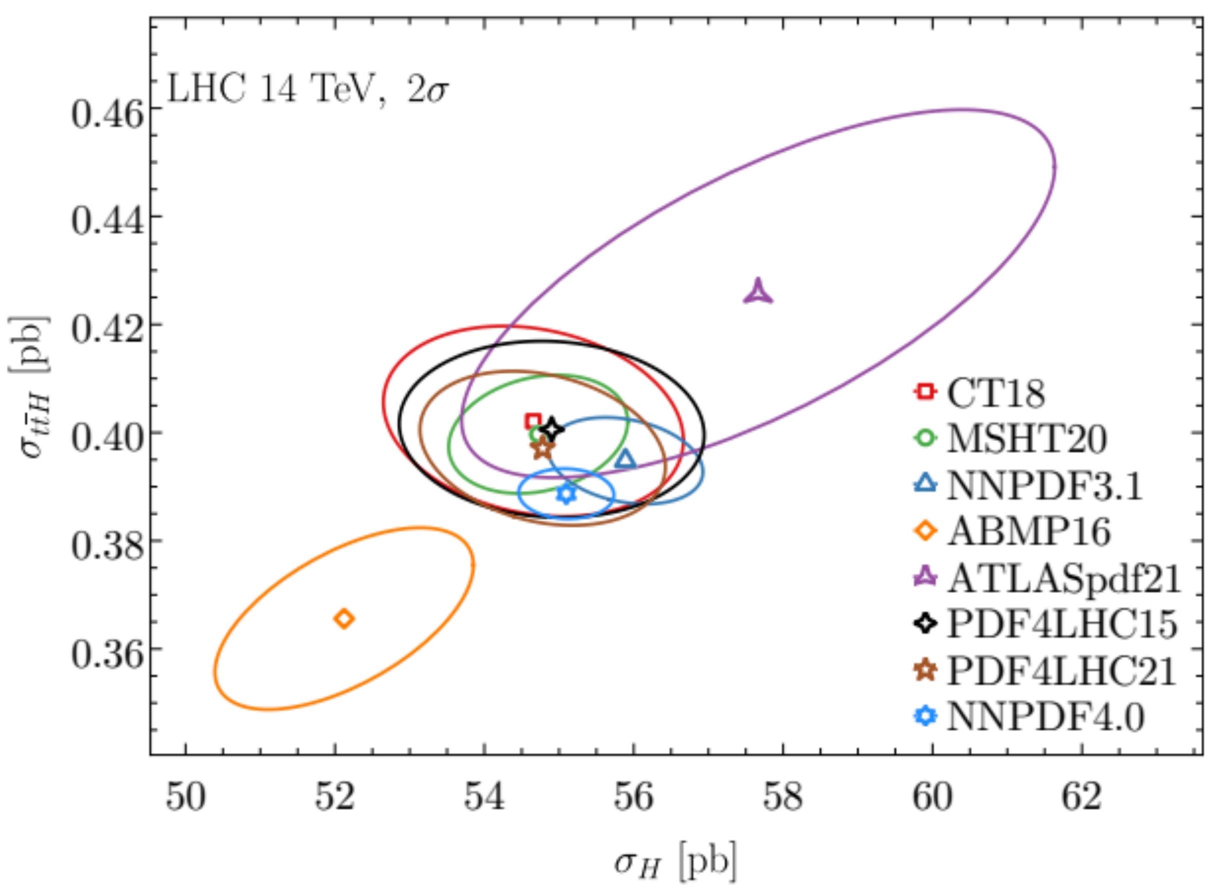


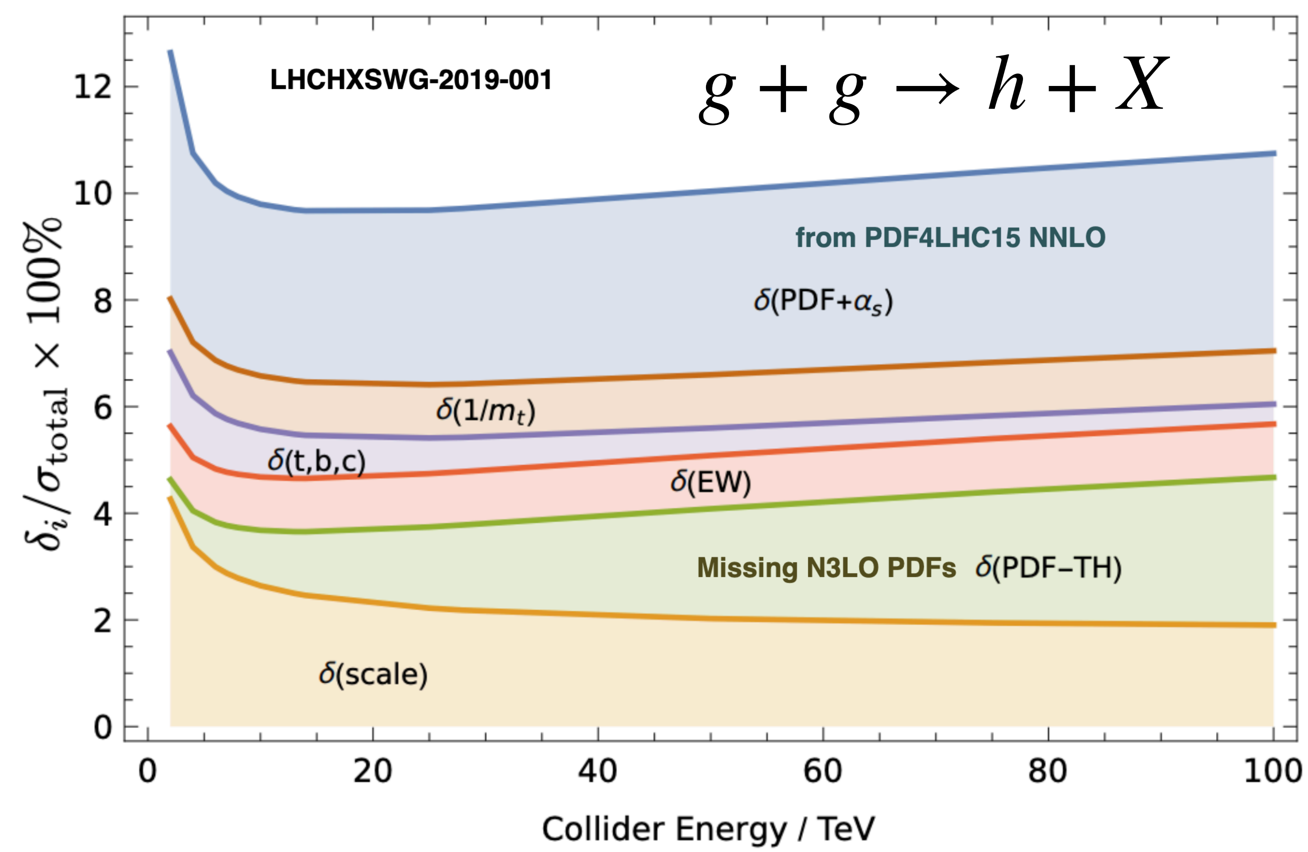


Implications for Higgs cross-sections

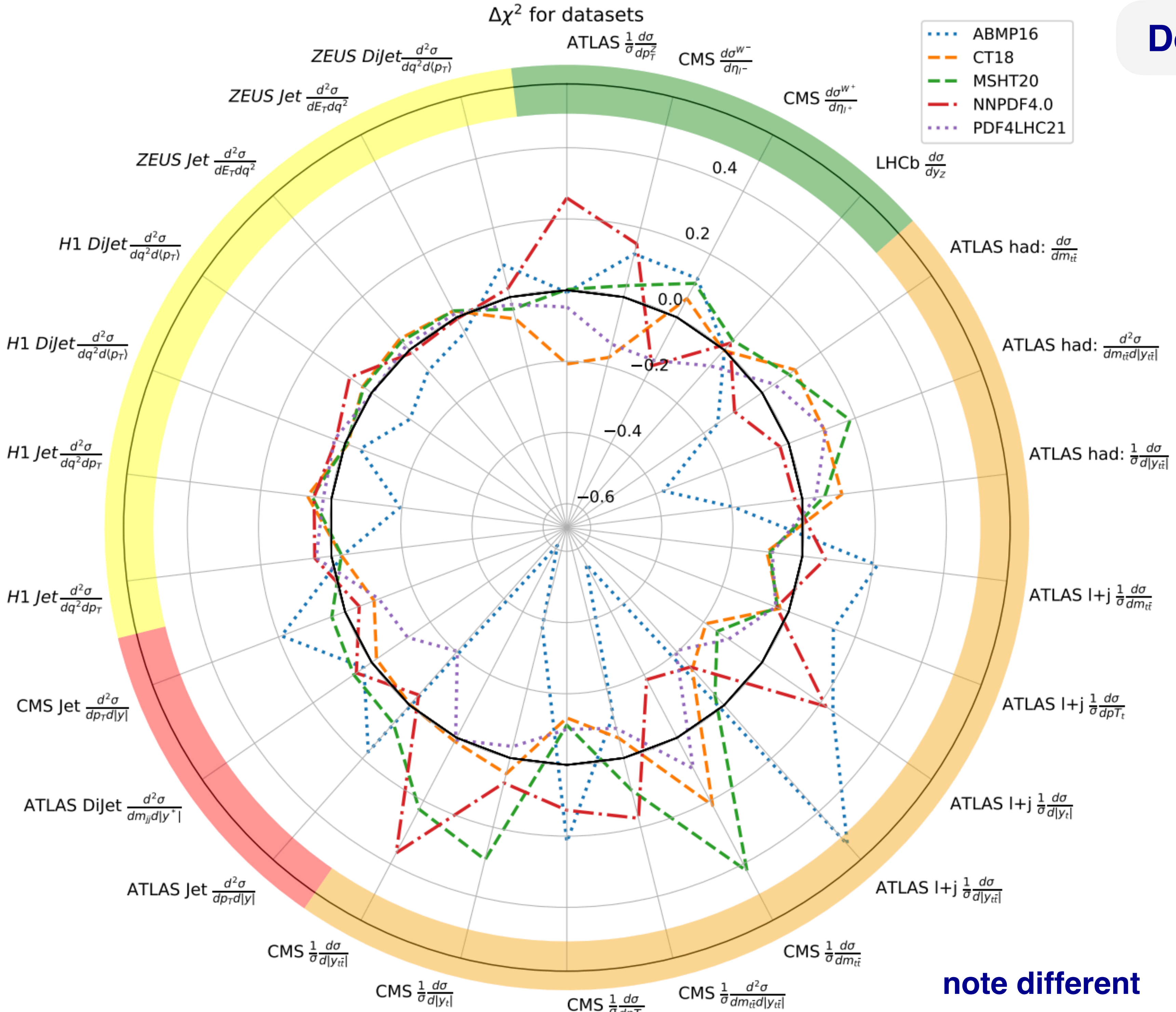


- Despite recent progress, differences between PDF sets remain both in **central values and in uncertainties** for Higgs cross-sections
- Differences arising from **input dataset and theory** as well as fitting methodologies
- **PDF+α_s uncertainties** dominate the LHCHXSWG estimate for ggF





PDF validation with LHC Run 2 data

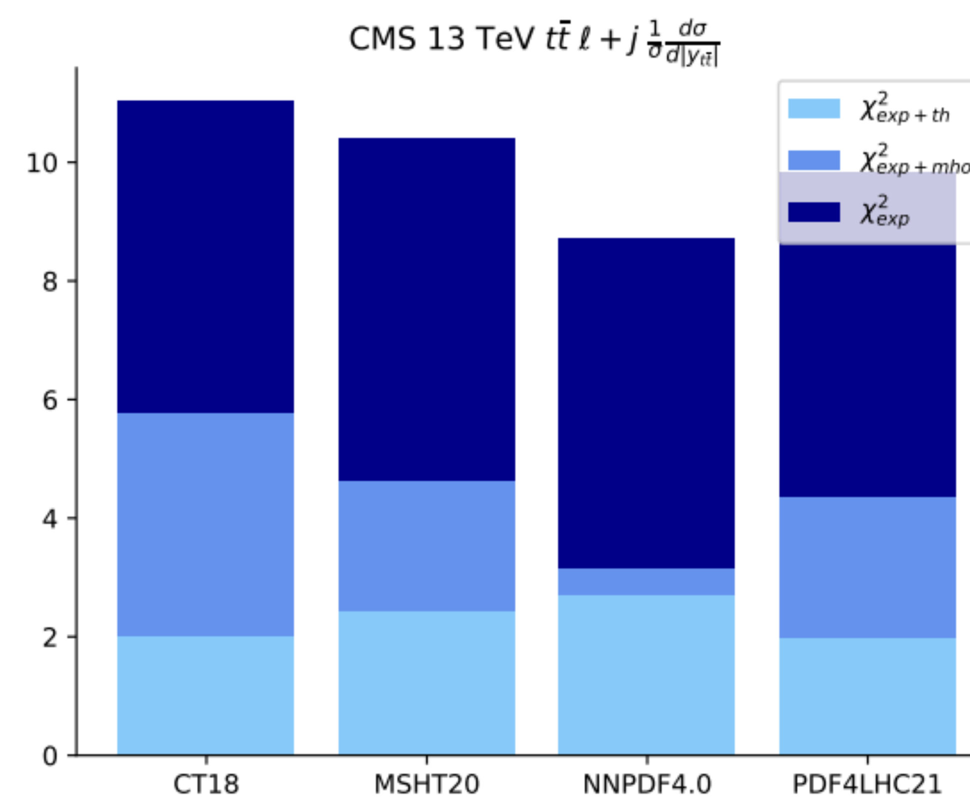


note different contribution of PDF errors

$$\Delta\chi^{2(i)} = \frac{\chi^{2(i)}_{\rm exp+th} - \left\langle \chi^2_{\rm exp+th}\right\rangle_{\rm pdfs}}{\left\langle \chi^2_{\rm exp+th}\right\rangle_{\rm pdfs}},$$

Do LHC data favour specific PDF sets?

- Quantitatively assess the **quality of data description** for different PDFs
- Account for all relevant sources of **experimental and theory** errors
- No single preferred / disfavoured PDF set: **more data needed**



arXiv:2501.10359

The Path to PDFs at N^{3}LO

Why PDFs at aN^{3}LO accuracy?

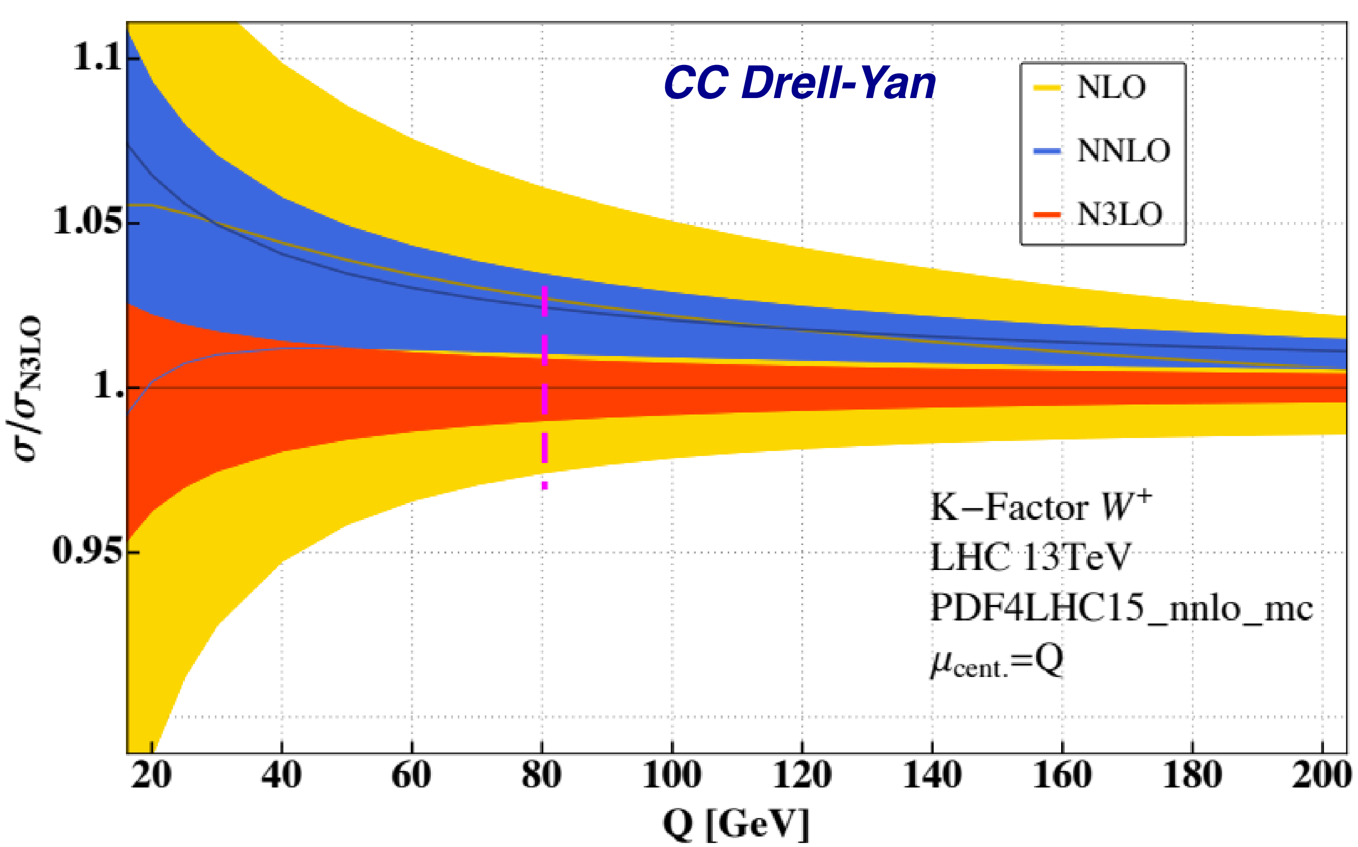
- **HXSWG YR4**: Perturbative mismatch between **partonic matrix elements** (accurate at N^{3}LO) and **PDFs** (accurate at NNLO) in core Higgs production processes, including gluon fusion

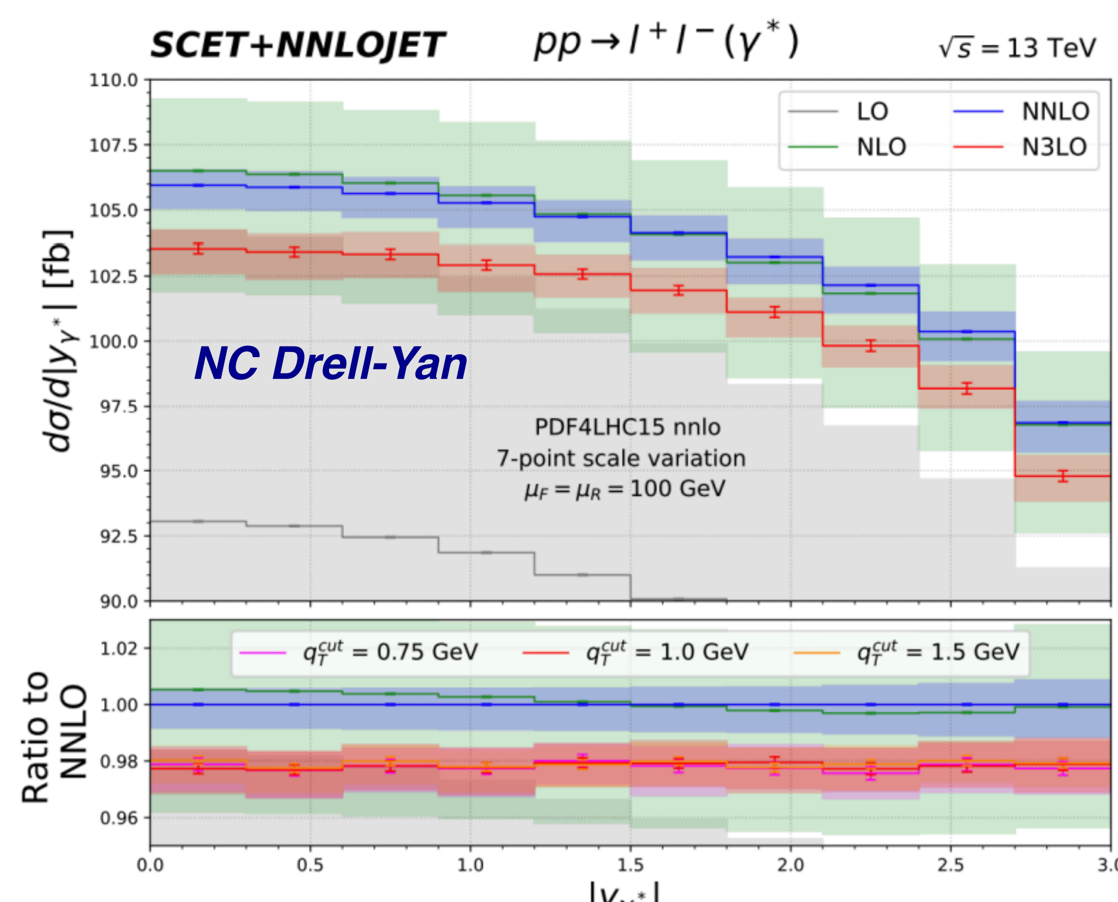
- Impact of this mismatch estimated to be **0.9% (ggF)**, **0.5% (VBF)**, **0.2% (hW)** (from NLO → NNLO)

$$\Delta^{\rm app}_{\rm NNLO} \equiv \frac{1}{2}\left|\frac{\sigma^{\rm NNLO}_{\rm NNLO-PDF} - \sigma^{\rm NNLO}_{\rm NLO-PDF}}{\sigma^{\rm NNLO}_{\rm NNLO-PDF}}\right|$$

Is this estimate accurate enough?

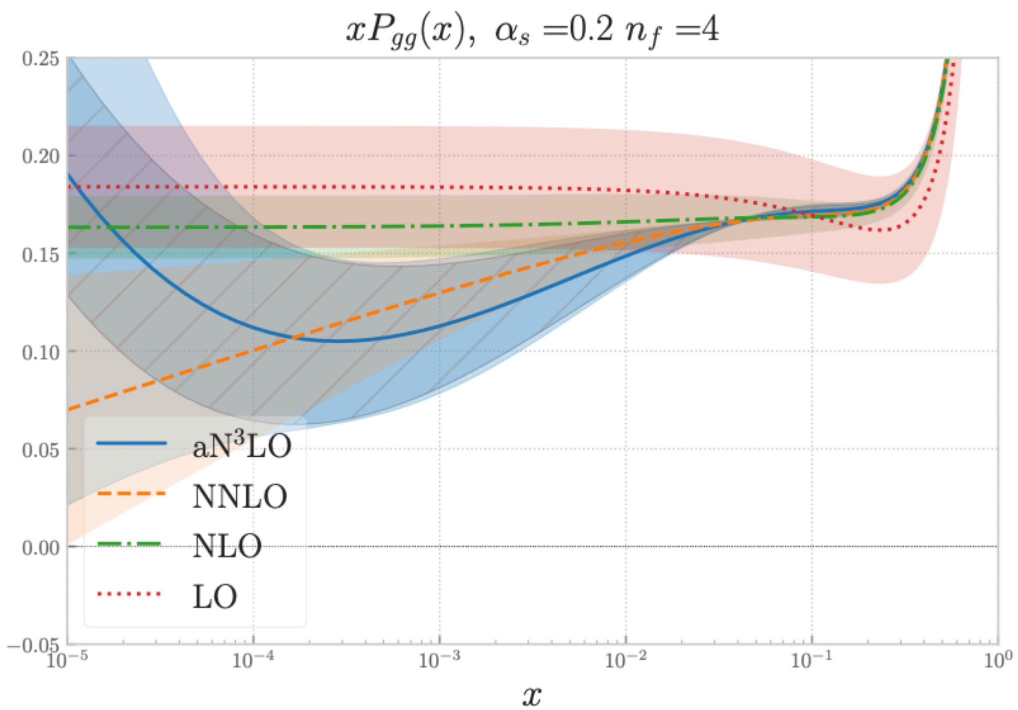
- **Perturbative convergence** of N^{3}LO calculations sub-optimal - due to **missing N^{3}LO PDFs**?

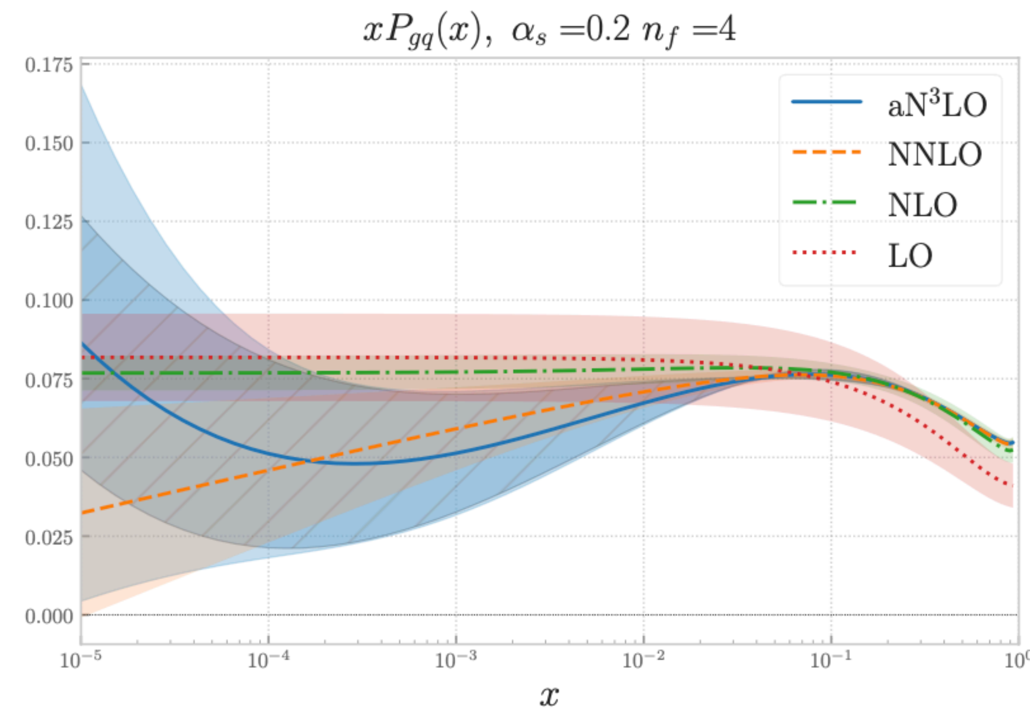




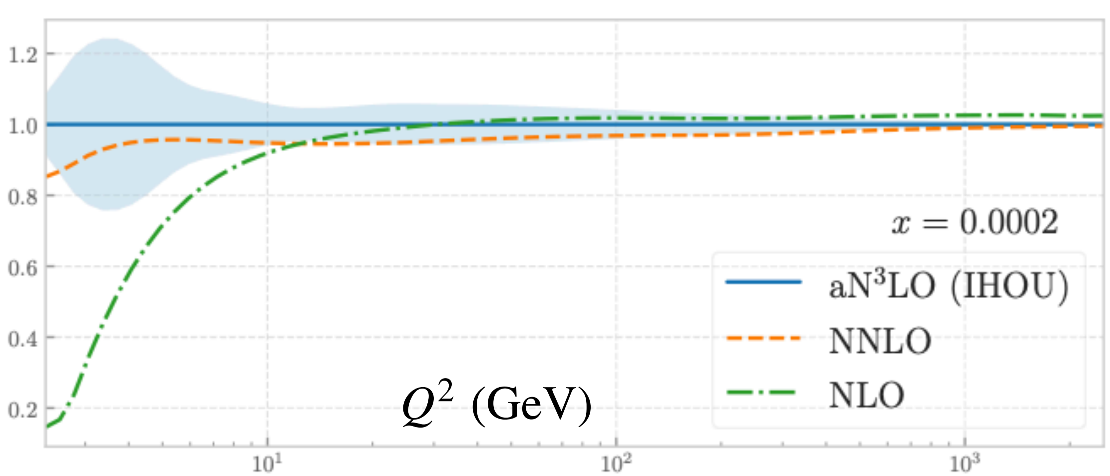
NNPDF4.0 at aN³LO accuracy

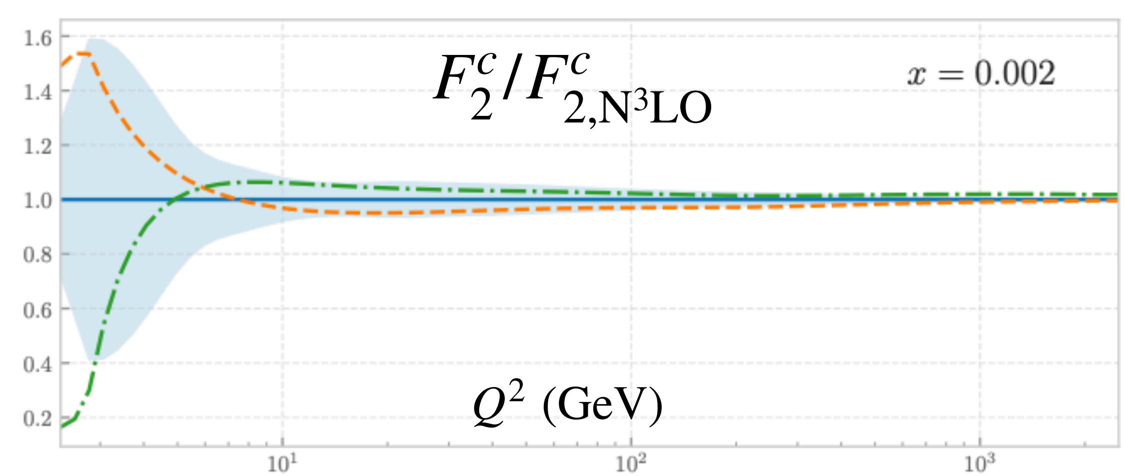
- Approximate parametrisation for the N³LO splitting functions satisfying known **exact results and limits**





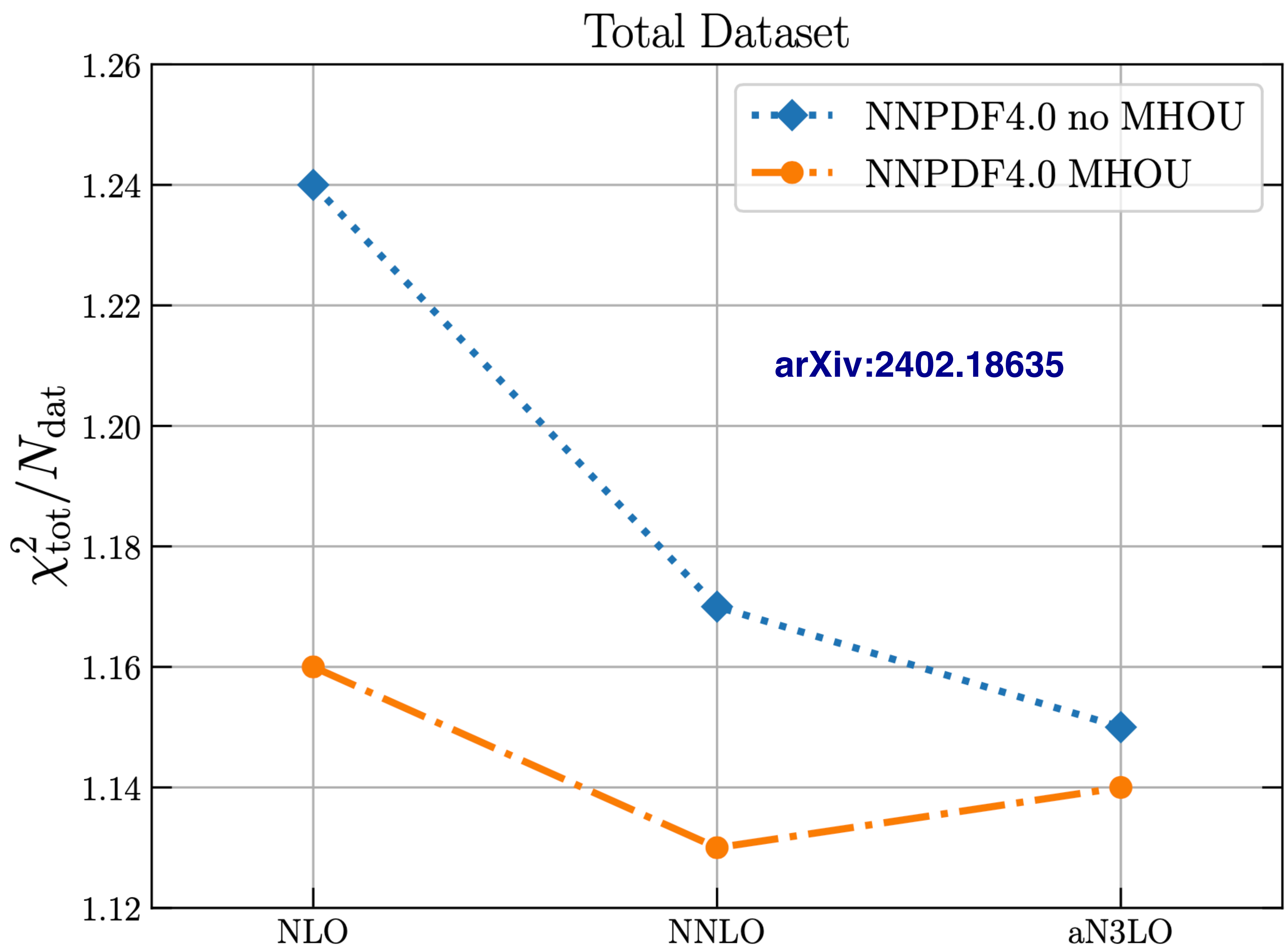
- Exact (approximate) massless (massive) deep-inelastic **coefficient functions and heavy quark matching coefficients** at N³LO accuracy & extension of the **FONLL general-mass scheme at N³LO**





- Theory covariance matrix includes contributions from **MHOUs** (μ_F and μ_R variations) and **IHOUs**
- Hadronic data fitted **using aN³LO evolution and NNLO matrix elements**, supplemented by MHOUs associated to μ_R variations to account for missing N³LO *K*-factors

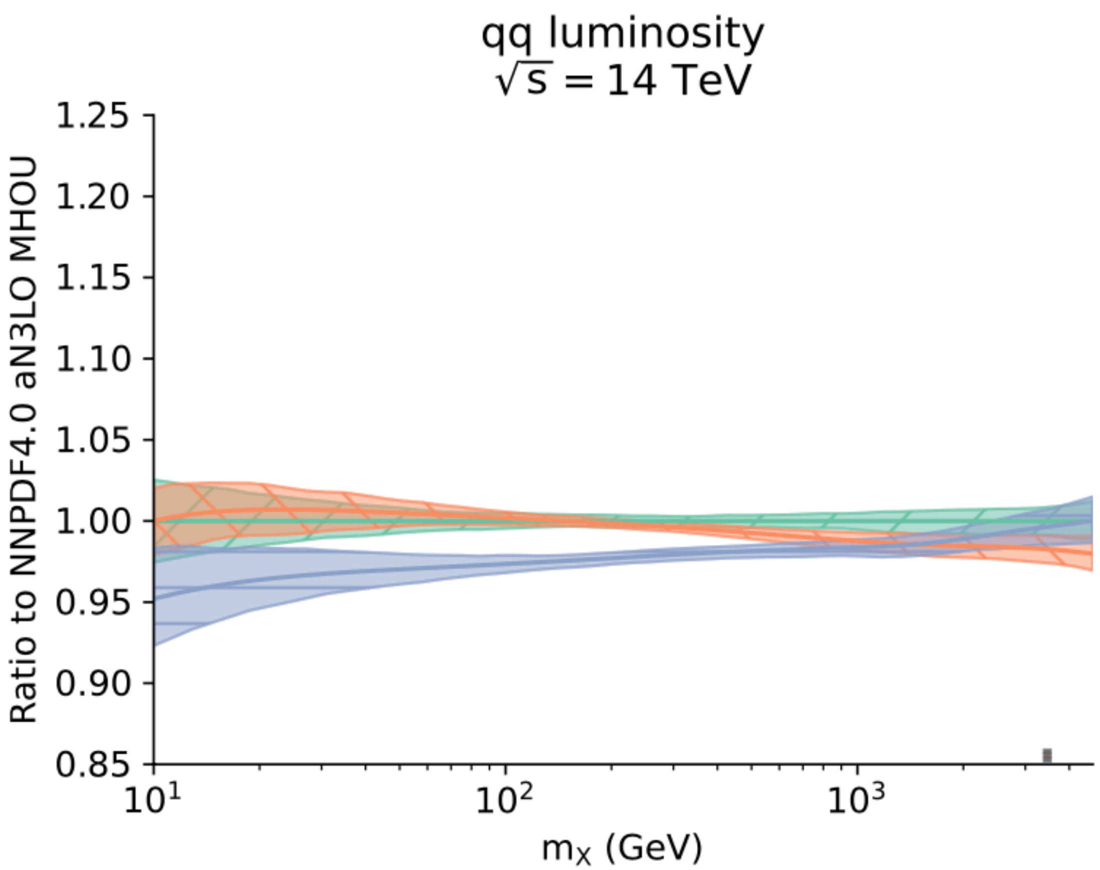
Fit quality

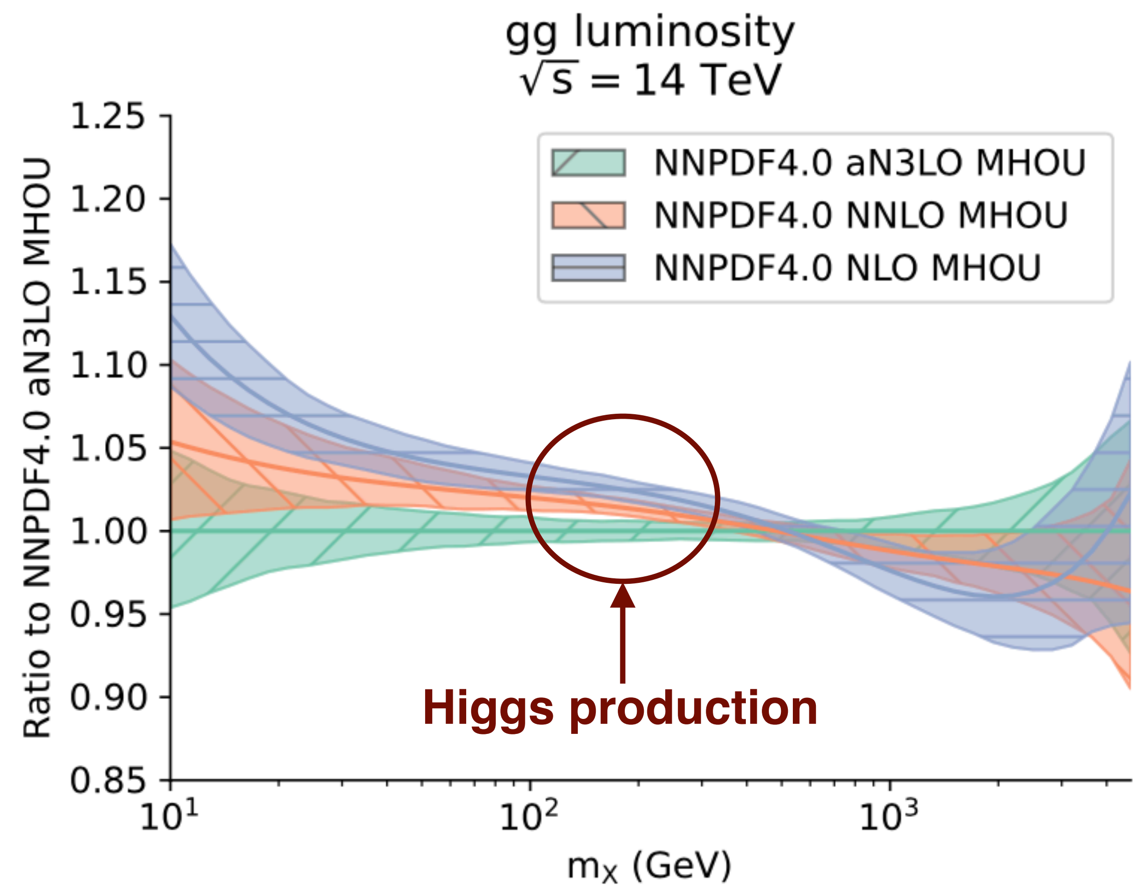


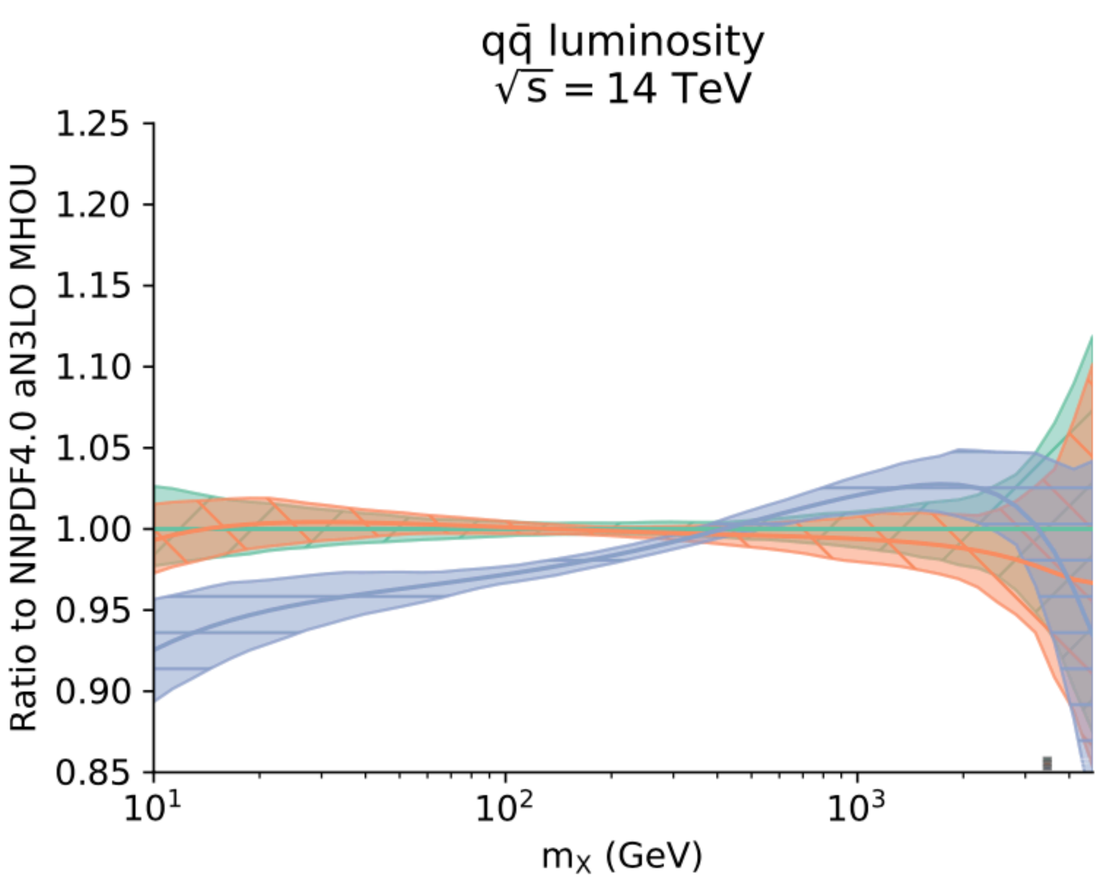
- Without MHOUs, the **χ^2 improves with the perturbative accuracy** of the PDF fit
- With MHOUs, the **χ^2** becomes **feebly dependent on the perturbative accuracy**
- At aN^{3}LO impact of MHOUs is small (also at PDF level) but non negligible

N^{3}LO corrections required for perturbative convergence at the PDF fit level

Perturbative convergence

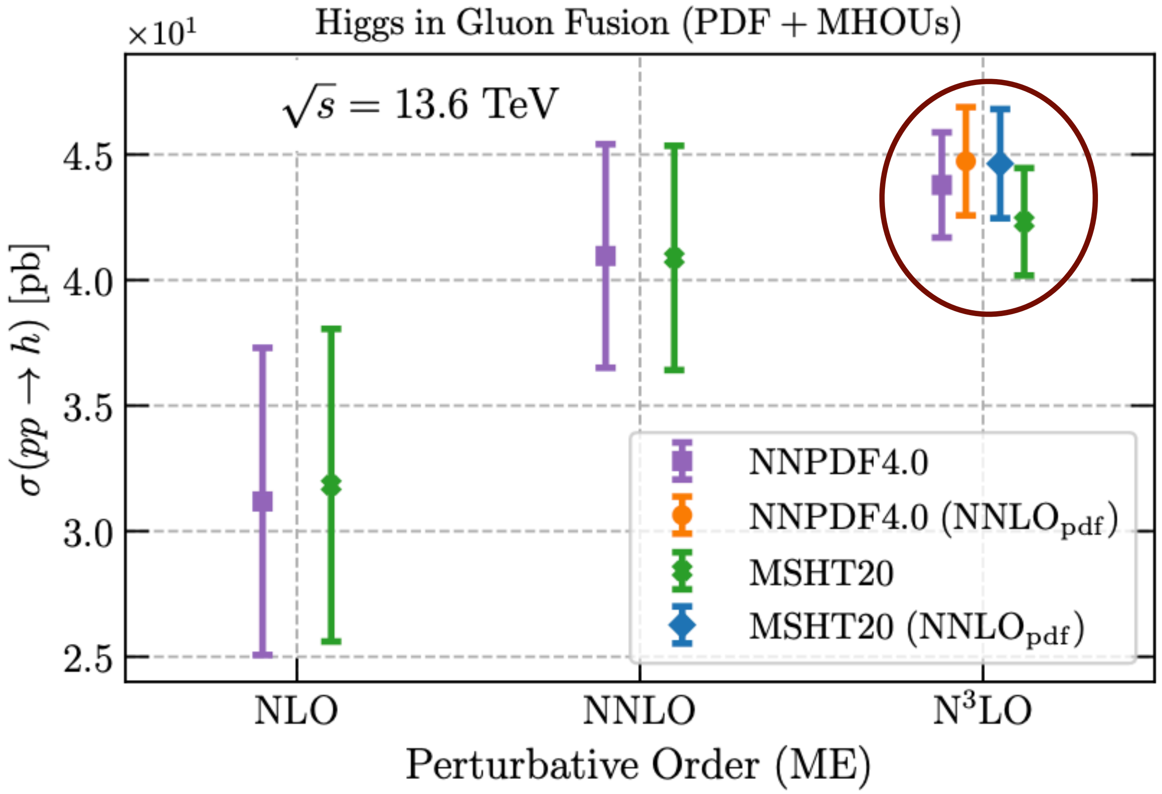


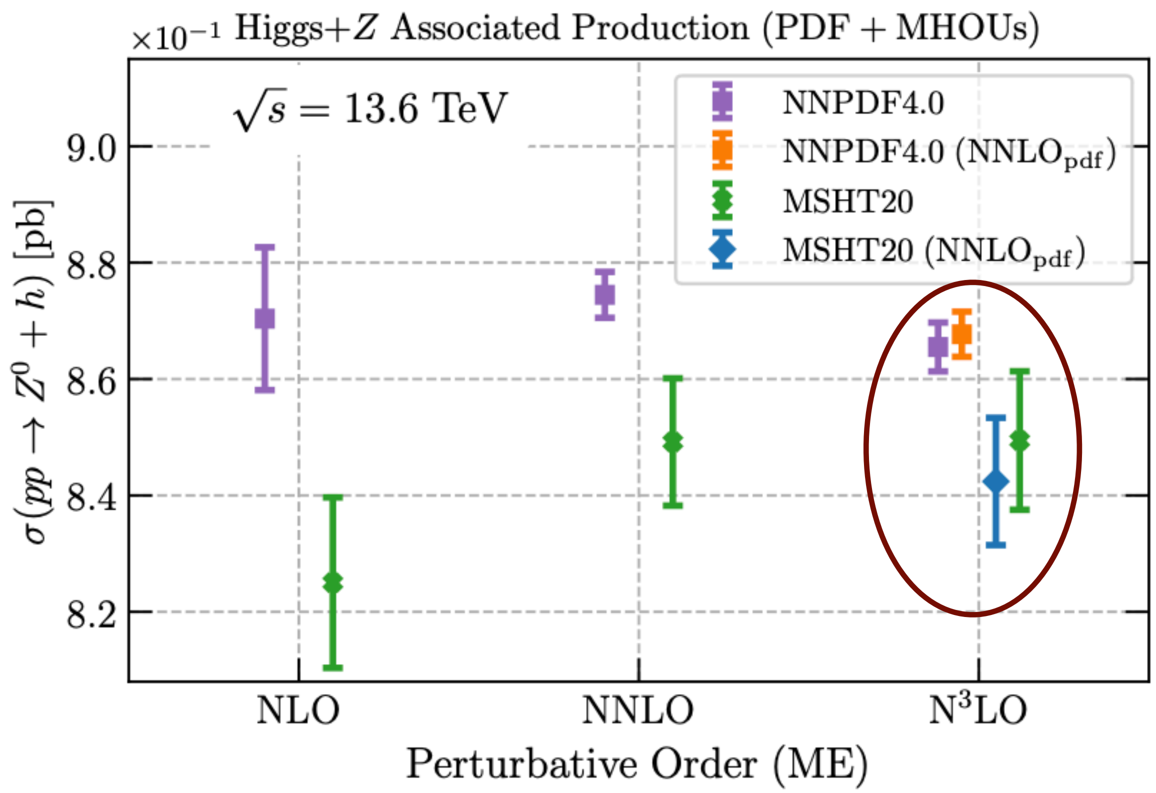


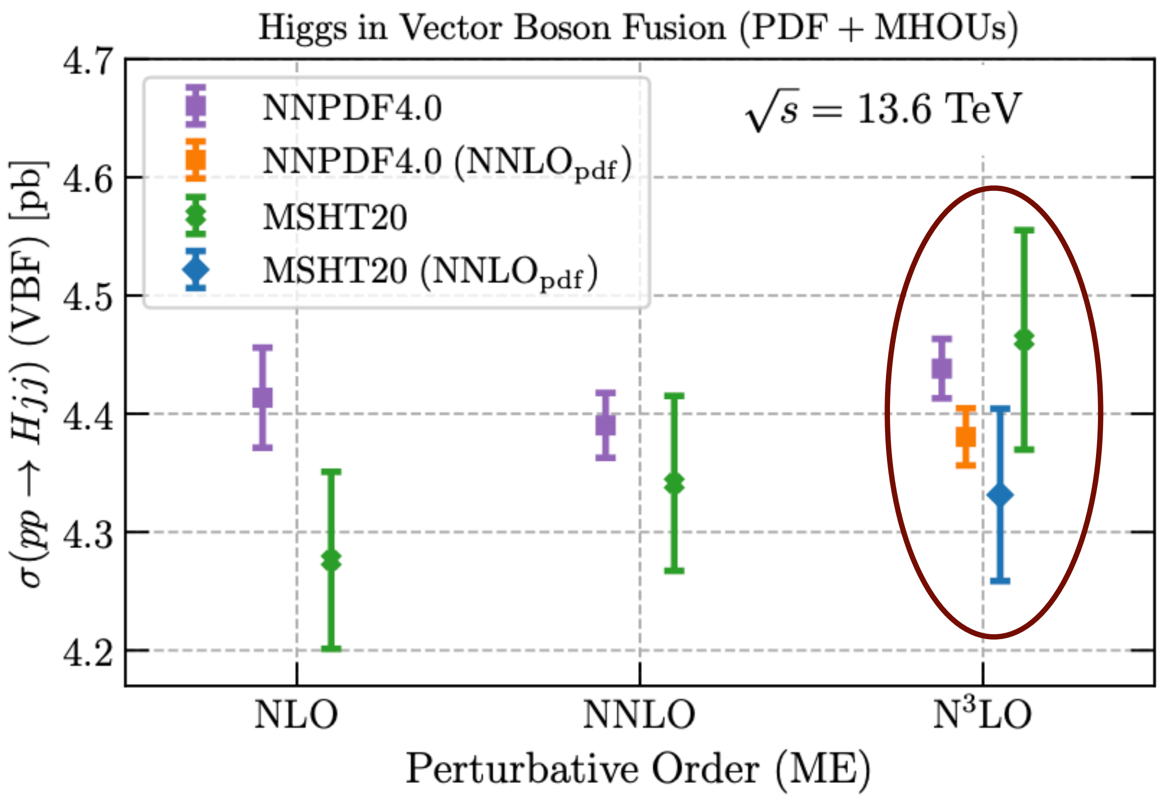


- **Good perturbative convergence**
- Impact of N^3LO corrections moderate but **not negligible**: impact on LHC phenomenology
- e.g. for the gluon-gluon luminosity, **suppression** around Higgs mass (2% effect)

Higgs production at N^3LO accuracy





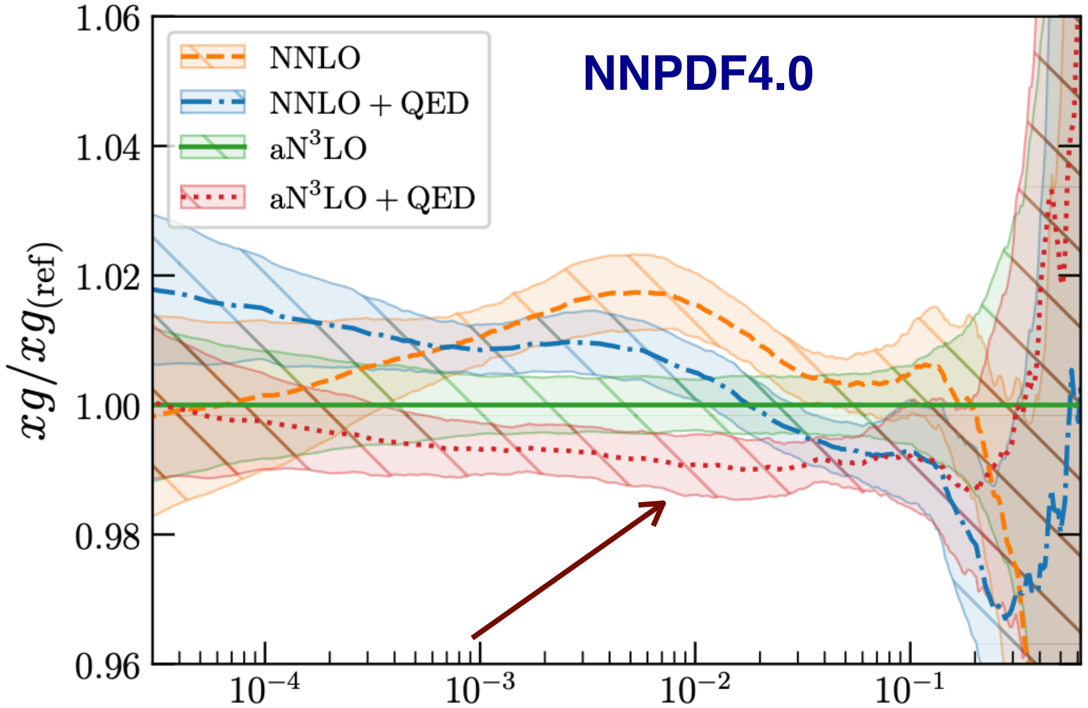


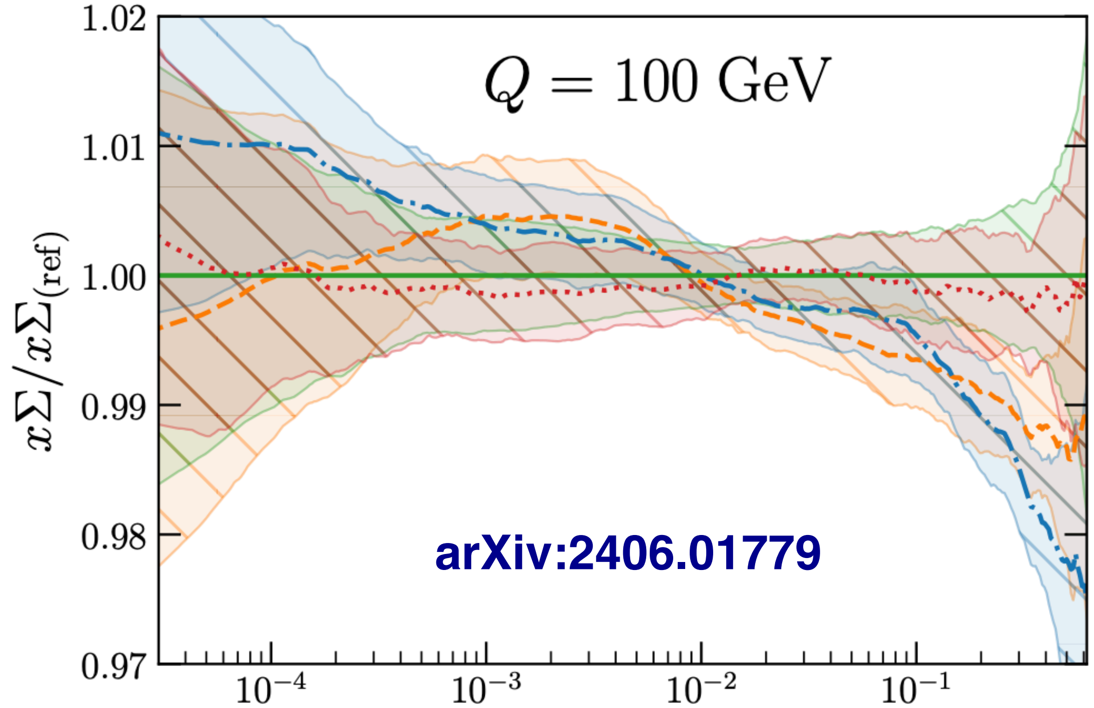
Compare with inconsistent calculation with NNLO PDFs

- N^3LO PDF corrections to **Higgs in gluon fusion:** 1.5% suppression wrt NNLO PDFs
- N^3LO corrections improve agreement between NNPDF4.0 and MSHT20 for ***hZ***
- **Higgs VBF:** large corrections when compared to the small N^3LO scale error

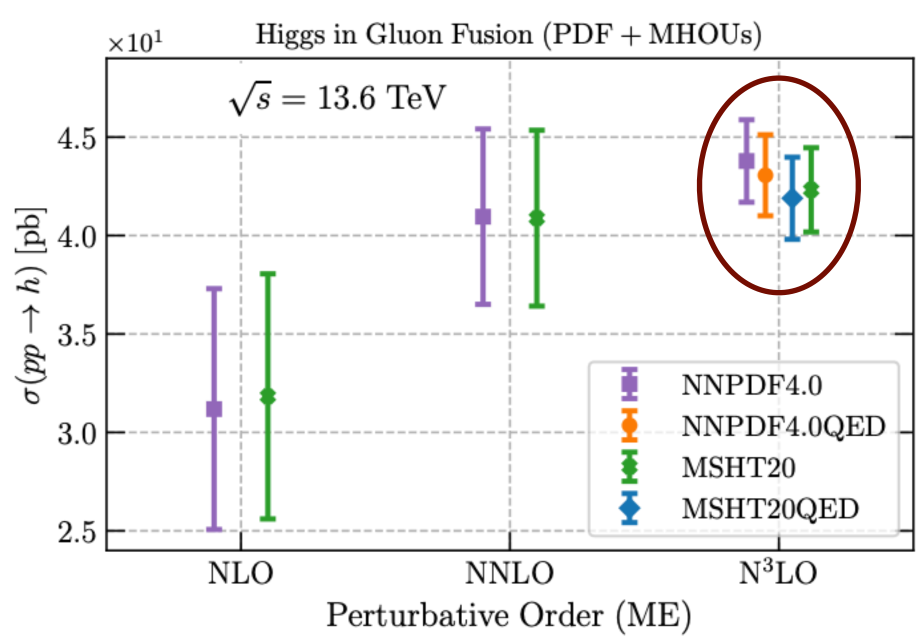
NNPDF4.0 aN^{3}LO with QED effects

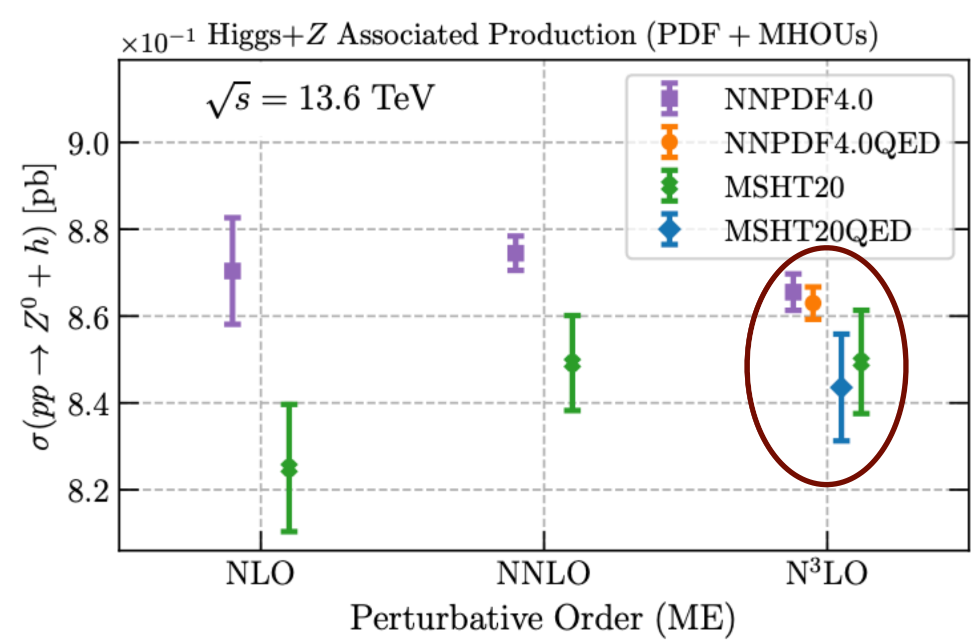
- PDFs with **QED corrections** and **photon PDF** key for accurate LHC phenomenology
- Higgs cross-sections may receive sizeable photon-initiated contributions
- QED effects **suppress the gluon by up to 1%** due to photon PDF ``eating up'' proton momentum





QED effects **decrease both ggF and hV cross-sections** (for fixed PDF)





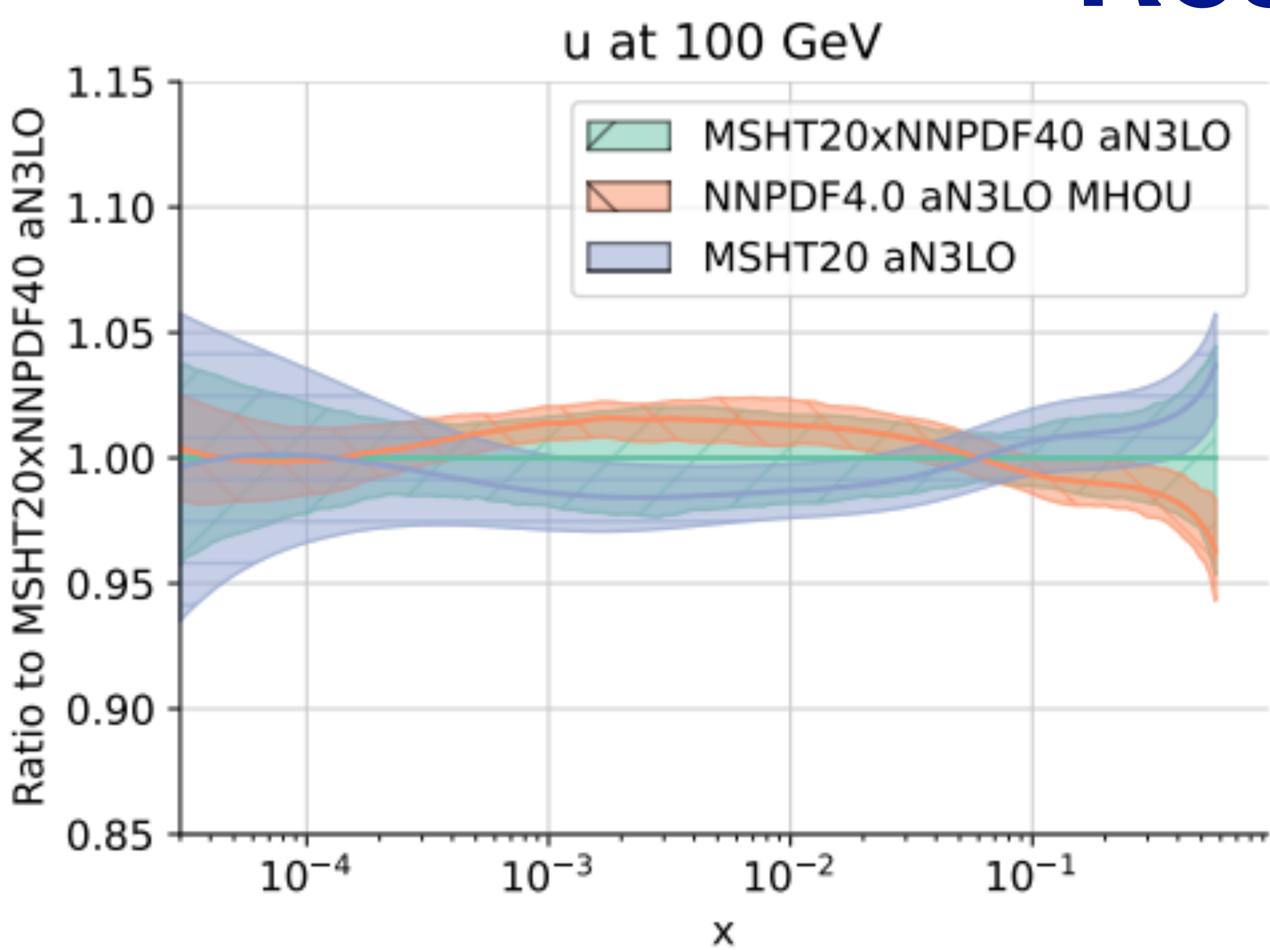
LHCHXSWG baseline PDFs neglect QED effects

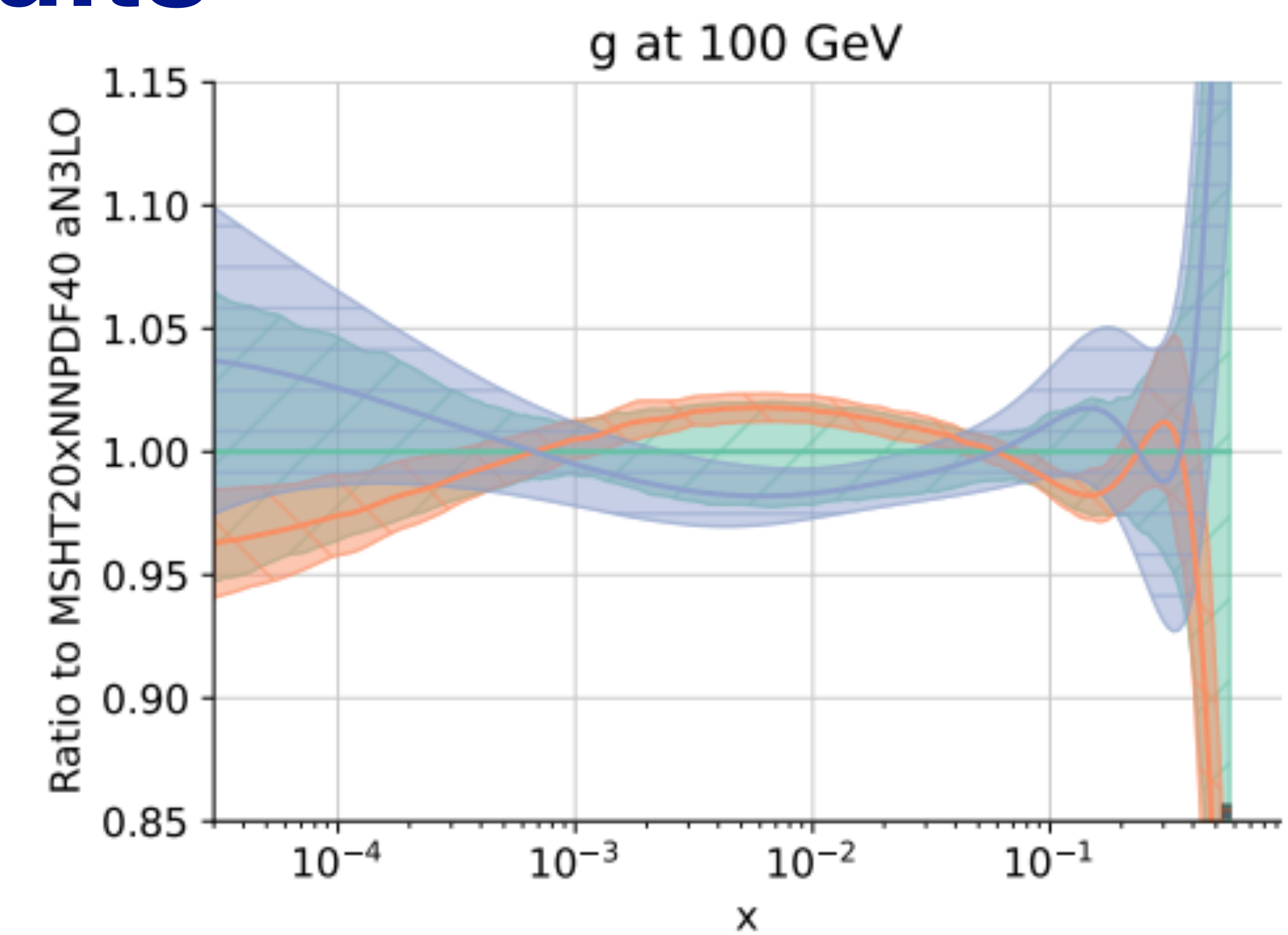
MSHT20 + NNPDF4.0 aN³LO combination

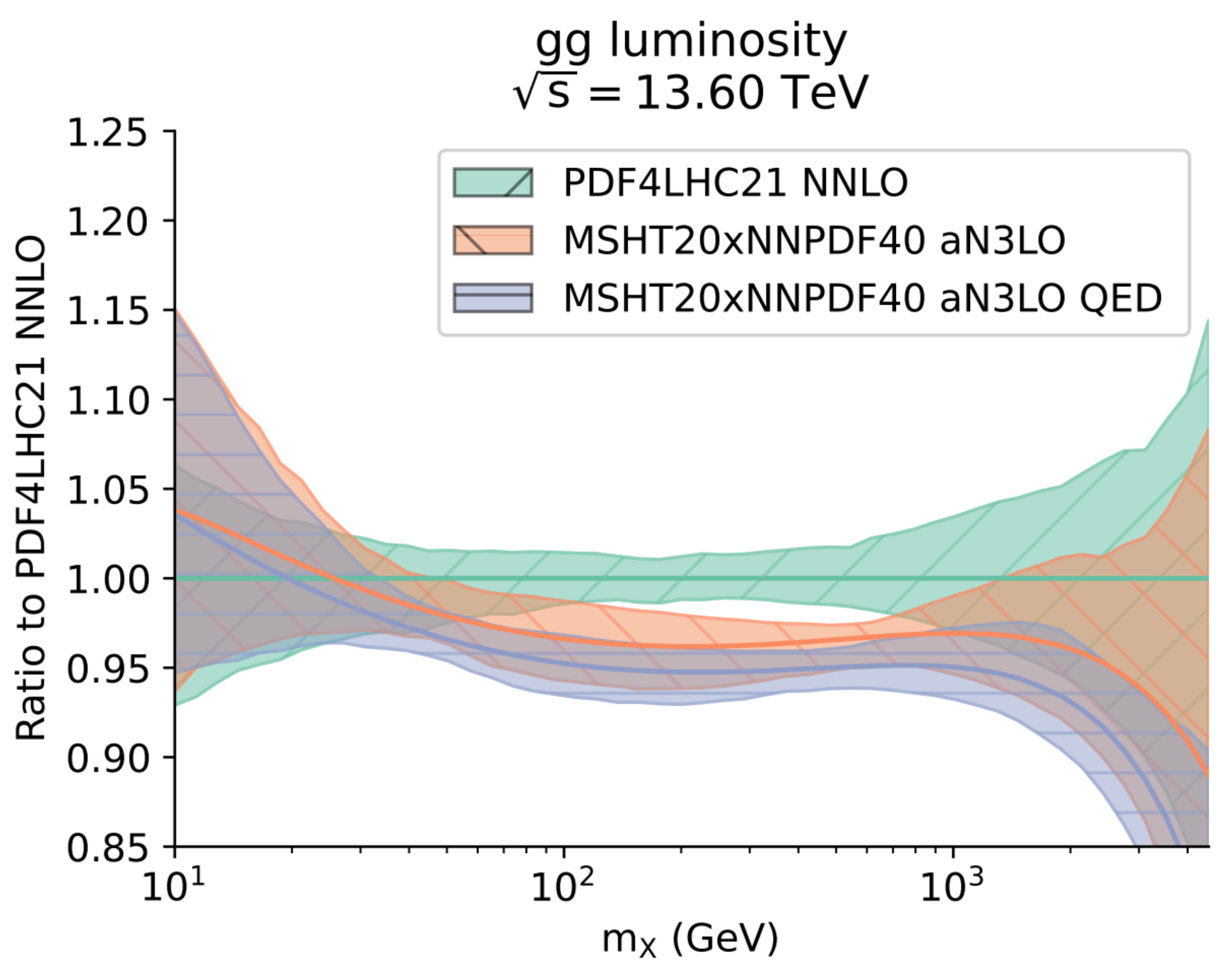
- Same approach as **PDF4LHC21**: $N_{\rm rep} = 100$ replicas of MSHT20 (from native Hessian) combined with $N_{\rm rep} = 100$ replicas of NNPDF4.0 **arXiv:2511.05373**
- Both for **aN³LO and aN³LO+QED variants**, together with NNLO and NNLO+QED as baseline

	PDF set	pert. order (PDF)
reference →	PDF4LHC21_mc	$\mathrm{NNLO_{QCD}}$
new combined sets →	MSHT20xNNPDF40_nnlo	$\mathrm{NNLO_{QCD}}$
	MSHT20xNNPDF40_nnlo_qed	$\mathrm{NNLO_{QCD}} \otimes \mathrm{NLO_{QED}}$
	MSHT20xNNPDF40_an3lo	$\mathrm{aN^3LO_{QCD}}$
	MSHT20xNNPDF40_an3lo_qed	$\mathrm{aN^3LO_{QCD}} \otimes \mathrm{NLO_{QED}}$
inputs →	NNPDF40_an3lo_as_01180_mhou	$\mathrm{aN^3LO_{QCD}}$
	NNPDF40_an3lo_as_01180_qed_mhou	$\mathrm{aN^3LO_{QCD}} \otimes \mathrm{NLO_{QED}}$
	MSHT20an3lo_as118	$\mathrm{aN^3LO_{QCD}}$
	MSHT20qed_an3lo	$\mathrm{aN^3LO_{QCD}} \otimes \mathrm{NLO_{QED}}$

Results

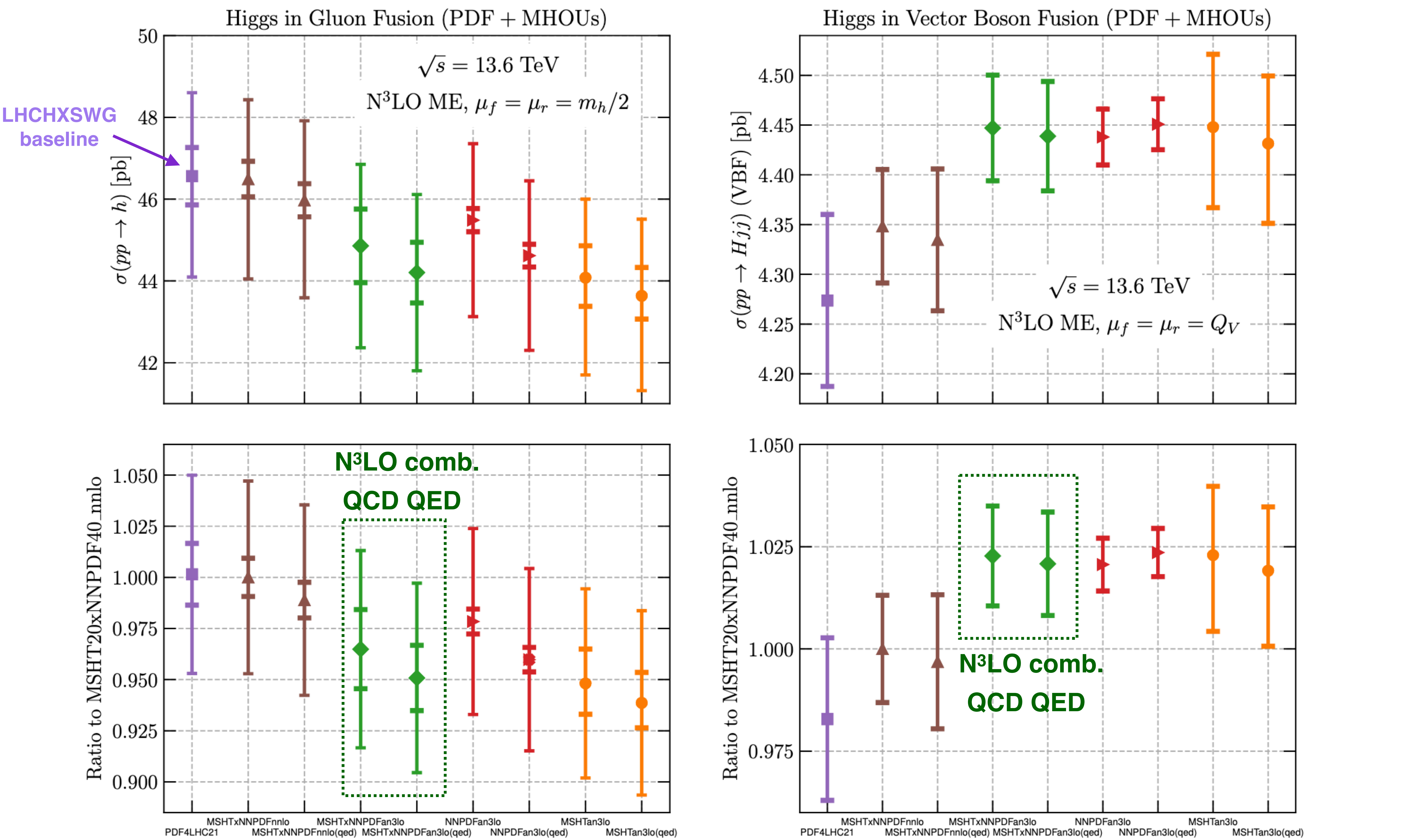






- Unweighted combination, no attempt to minimise differences between the two sets
- Bulk of differences between MSHT20 and NNPDF4.0 **already present at NNLO**
- Differences between N^{3}LO combination and PDF4LHC21 large for *gg* and *qq* lumis

Implications for Higgs physics



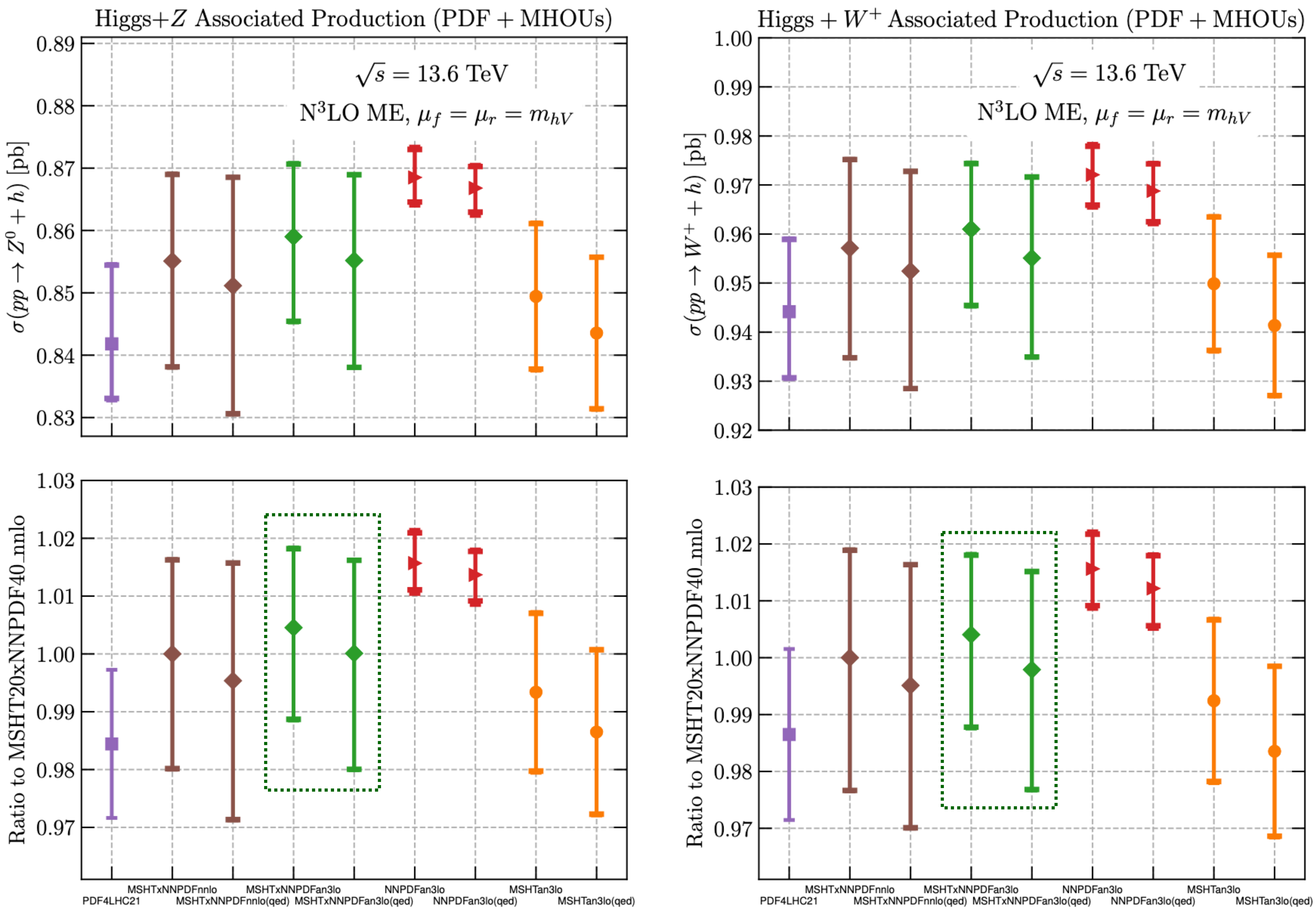
aN³LO (+QED) PDF corrections: **-3.5% (-5%)**

PDF4LHC21 close to **NNLO combination**

aN³LO (+QED) PDF corrections: **+2.5% (+2.5%)**

aN³LO combination: +1.8% higher than PDF4LHC21

Implications for Higgs physics



Impact of aN³LO & QED PDF corrections at the **few-permille level** for hV

Impact of different NNLO PDF combination: **up to +1.5%**

N^3LO effects: LHCXSWG estimates vs exact

- **HXSWG YR4**: Perturbative mismatch between **partonic matrix elements** (accurate at N^3LO) and **PDFs** (accurate at NNLO) in core Higgs production processes, including gluon fusion
- Impact of this mismatch estimated to be **0.9% (ggF)**, **0.5% (VBF)**, **0.2% (hW)** (from NLO → NNLO)

$$\Delta^{\rm app}_{\rm NNLO} \equiv \frac{1}{2}\left|\frac{\sigma^{\rm NNLO}_{\rm NNLO-PDF}-\sigma^{\rm NNLO}_{\rm NLO-PDF}}{\sigma^{\rm NNLO}_{\rm NNLO-PDF}}\right| \qquad \Delta^{\rm exact}_{\rm NNLO} \equiv \left|\frac{\sigma^{\rm N^3LO}_{\rm N^3LO-PDF}-\sigma^{\rm N^3LO}_{\rm NNLO-PDF}}{\sigma^{\rm N^3LO}_{\rm N^3LO-PDF}}\right| .$$

Exact shift due to N^3LO PDFs

exact shift

	ggF	VBF-h
$\Delta^{\rm exact}_{\rm NNLO}$ (NNPDF4.0)	2.2%	1.3%
$\Delta^{\rm exact}_{\rm NNLO}$ (MSHT20)	5.3%	2.3%
$\Delta^{\rm exact}_{\rm NNLO}$ (combination)	3.3%	2.3%
$\Delta^{\rm app}_{\rm NNLO}$ (NNPDF4.0)	0.2%	0.2%
$\Delta^{\rm app}_{\rm NNLO}$ (MSHT20)	1.4%	1.3%
$\Delta^{\rm app}_{\rm NNLO}$ (combination)	0.9%	0.5%

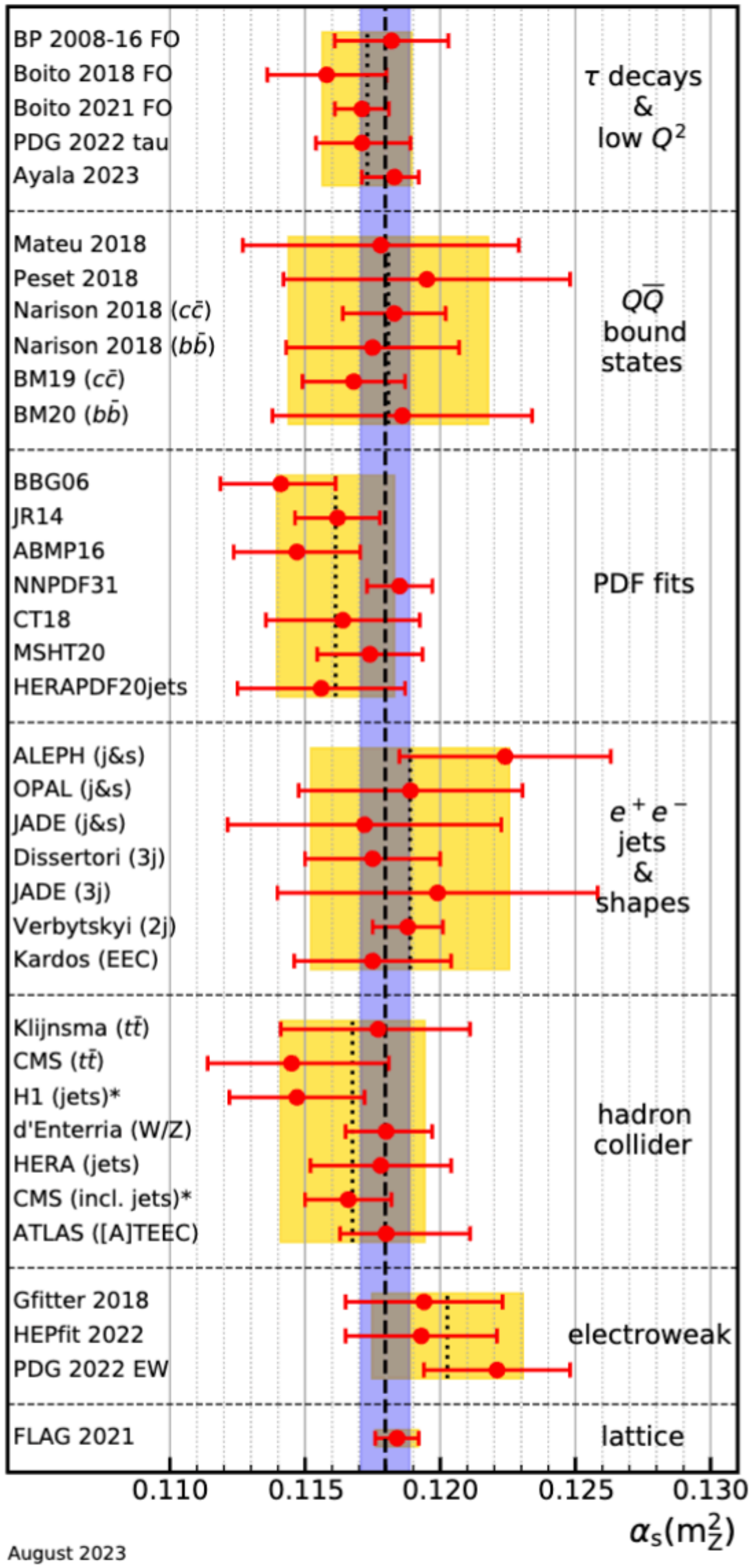
approx estimate

- LHCHXSWG estimates of aN^3LO PDF effects **underestimate true shift**
- LHCHXSWG chooses to use PDF4LHC21 for **YR5,** hence **neglecting** ``known" large corrections to Higgs xsecs due to N^3LO and QED effects

Is this the best choice?

The strong coupling from a aN^{3}LO global PDF fit

The strong coupling & Higgs physics



- The precise determination of $\alpha_s(m_Z)$is crucial for theoretical predictions for **Higgs production and decay**

- We carried out a first $\alpha_s(m_Z)$ extraction from a global PDF based on same **accuracy of state-of-the-art Higgs cross-section calculations**: $\mathrm{aN^3LO_{QCD}} \otimes \mathrm{NLO_{QED}}$

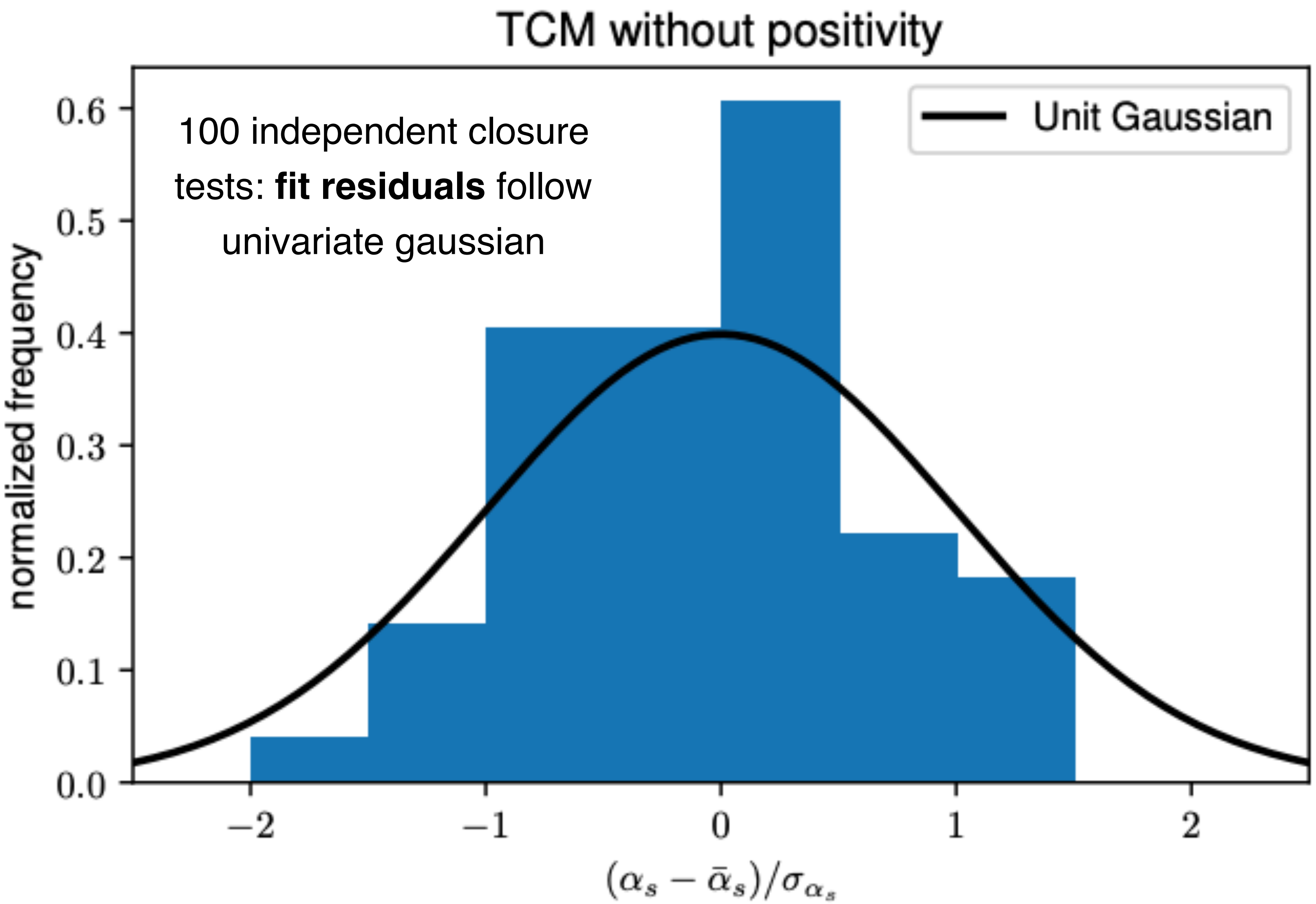
Process	σ (pb)	$\delta\alpha_s(\%)$	PDF $+\alpha_s(\%)$	Scale(%)
ggH	49.87	$\pm$ 3.7	-6.2 +7.4	-2.61 + 0.32
ttH	0.611	$\pm$ 3.0	$\pm$ 8.9	-9.3 + 5.9

Partial width	intr. QCD	para. m_q	para. α_s
$H \to b\bar{b}$	$\sim 0.2\%$	1.4%	0.4%
$H \to c\bar{c}$	$\sim 0.2\%$	4.0%	0.4%
$H \to gg$	$\sim 3\%$	$< 0.2\%$	3.7%

D. d'Enterria, ESPPU Open Symposium 2025

Closure tests validation

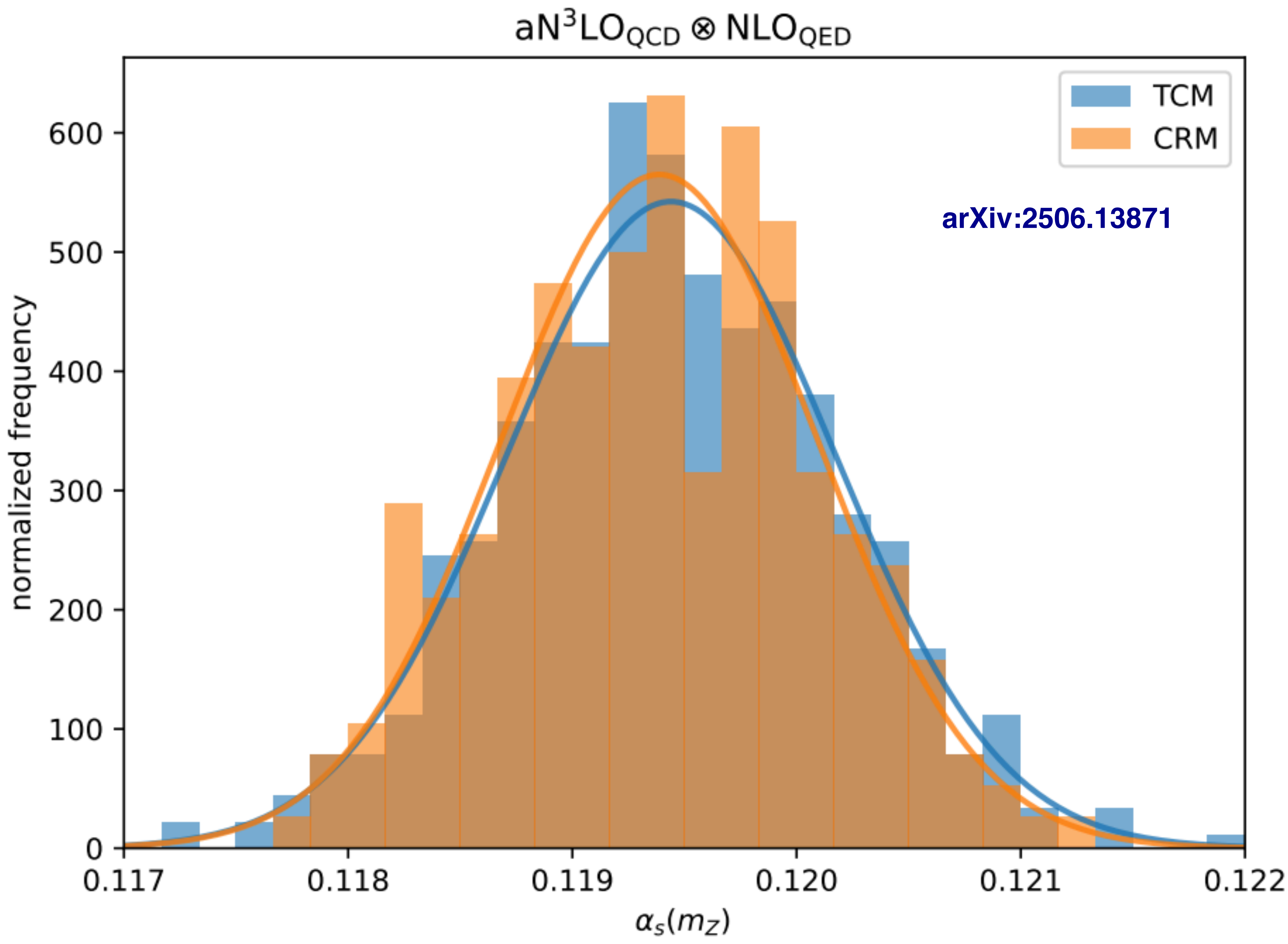
- Generate **synthetic data** based on a given value of $\alpha_s(m_Z)$
- Verify we reproduce this ground truth by **two independent fitting methodologies,** one bayesian (Theory Covariance Matrix) and the other frequentist (Correlated Replica Method)



- Identified **plausible** methodologies that **fail the closure test!** For example, varying the value of $\alpha_s(m_Z)$ in the t_0 covariance matrix leads to $\alpha_s(m_Z) = 0.1195$ (for $\bar{\alpha}_s = 0.118$) in the fit

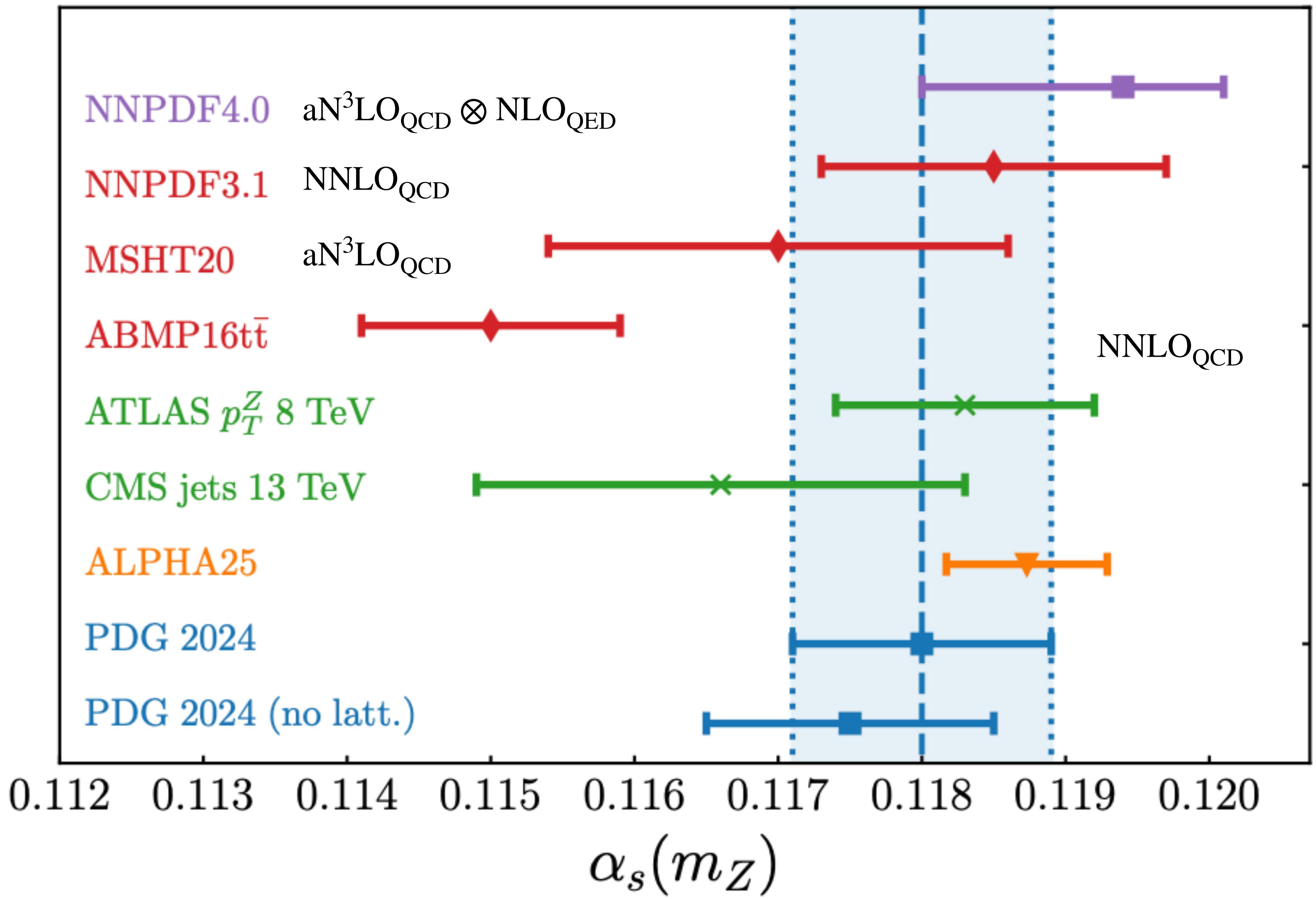
Validation with synthetic data essential to identify biases

Results



Consistent results with two fully independent methodologies

Results



$0.1194^{+0.0007}_{-0.0014}$

0.1185 ± 0.0012

Most precise calculation from **lattice QCD**

- Total uncertainty is **0.9%** (includes MHOUs and δm_{top}), consistent with PDG and lattice QCD
- Neglecting MHOUs, QED effects, and aN³LO corrections shifts result by **0.7%, 0.3%**, and **0.6%**
- Large weight of LHC data in fit; methodological bias identified and corrected via closure tests

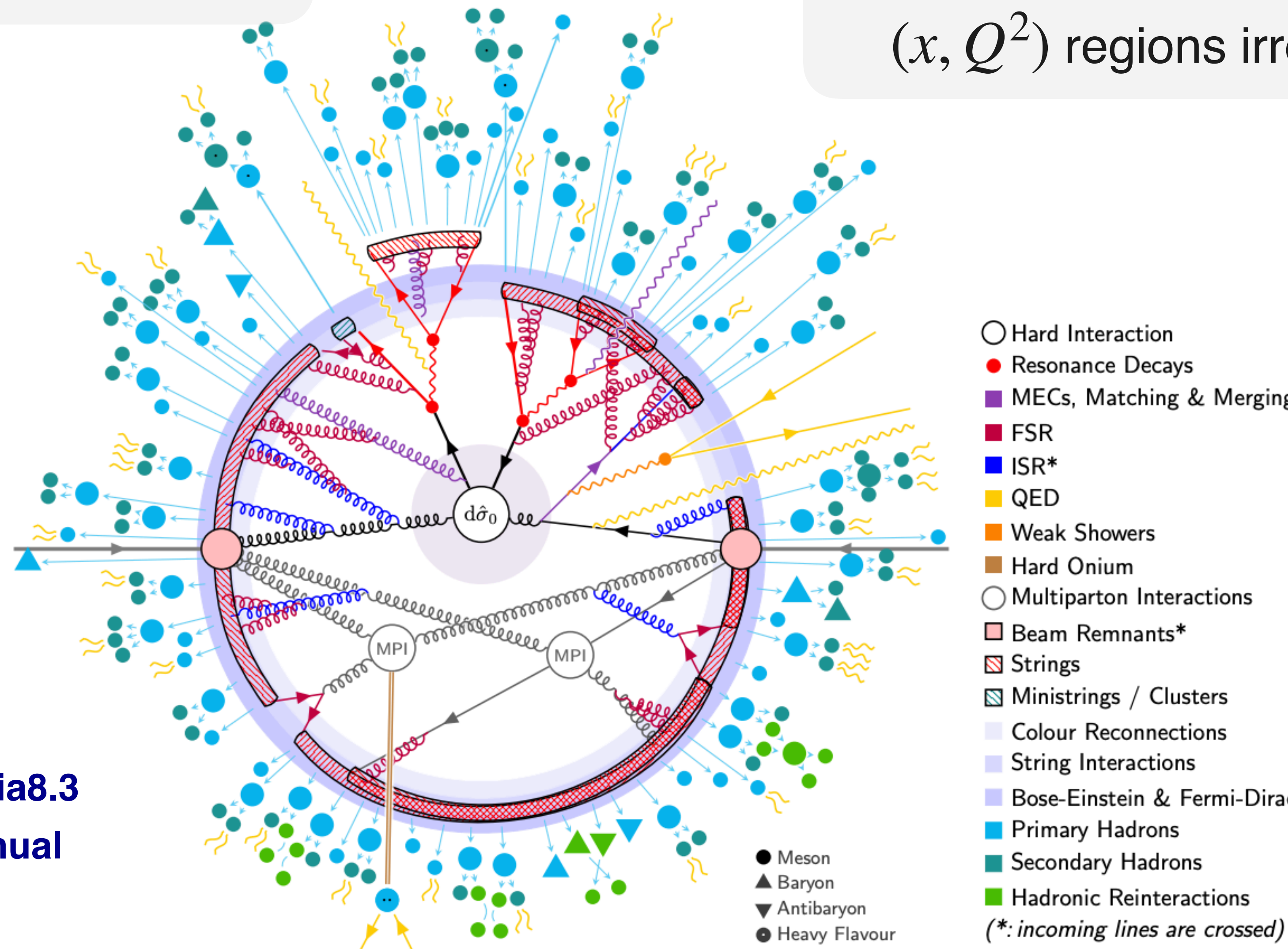
PDFs for (N)NLO Monte Carlo Generators

PDFs & Event Generators

Why **regular PDF sets** are sometimes sub-optimal when used within event generators?

ISR showers require positive-definite PDFs down to $Q \sim 1$ GeV

MC integration & sampling requires smooth, numerically stable PDFs even in (x, Q^2) regions irrelevant for pheno



Pythia8.3 manual

Modelling of UE & MPI demand smooth extrapolation to very small-*x* & gluon PDF raising sufficiently fast

Simulation of **QED showers & photon-initiated processes** demands fits with QED effects included

PDFs & Event Generators

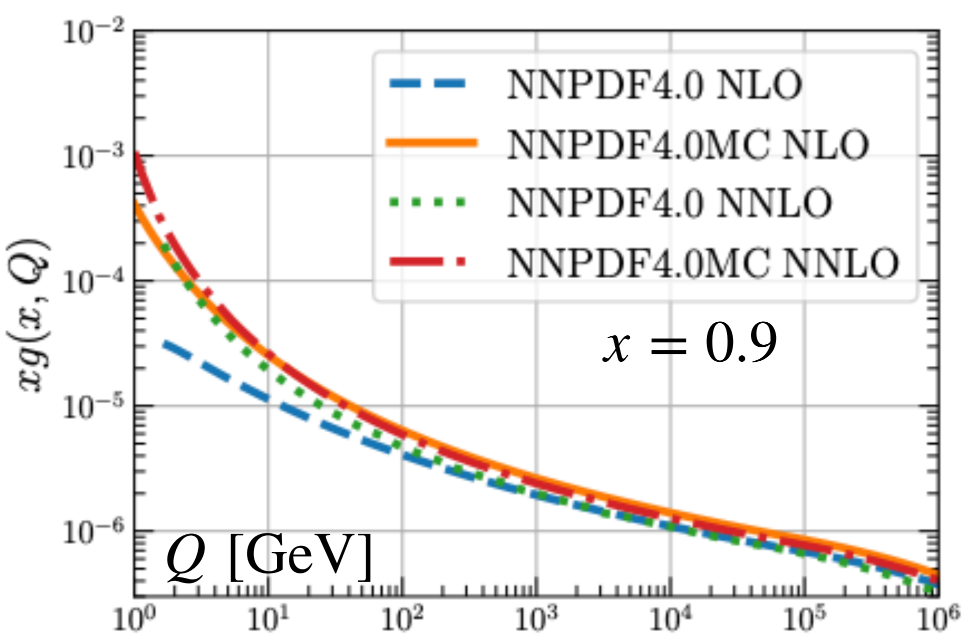
The **NNPDF4.0MC PDFs** satisfy these requirements at LO as well as **NLO and NNLO**

Satisfactory **NNLO** χ^2, only small worsening wrt baseline PDFs

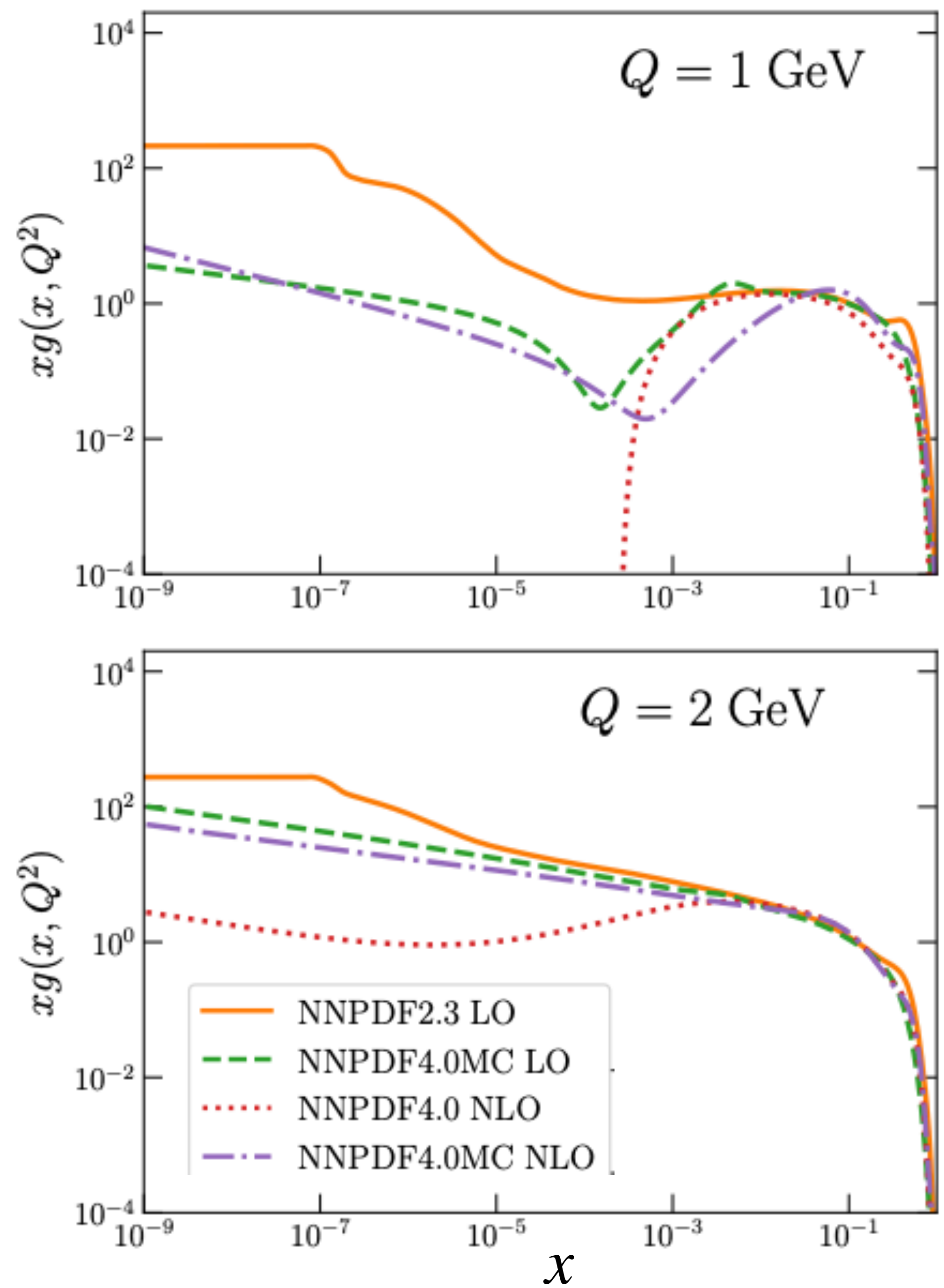
Dataset by process group	NLO					NNLO				
	n_{dat}	QCD		QCD+QED		n_{dat}	QCD		QCD+QED	
		BL	MC	BL	MC		BL	MC	BL	MC
DIS NC	1953	1.35	1.37	1.38	1.54	2110	1.22	1.30	1.22	1.29
DIS CC	988	0.91	0.92	0.94	0.95	989	0.90	0.89	0.90	0.89
DY NC	669	1.58	1.84	1.67	2.04	736	1.20	1.30	1.22	1.33
DY CC	197	1.38	1.56	1.40	1.61	157	1.45	1.55	1.47	1.57
Top pairs	66	2.40	2.14	2.51	2.47	64	1.27	1.16	1.31	1.27
Single-inclusive jets	356	0.82	0.88	0.83	0.93	356	0.94	1.01	0.93	1.00
Dijets	144	1.51	1.55	1.56	1.62	144	2.01	2.01	1.94	1.93
Photon	53	0.57	0.60	0.64	0.74	53	0.76	0.67	0.74	0.68
Single top	17	0.36	0.36	0.38	0.36	17	0.37	0.38	0.39	0.40
Total	4443	1.28	1.30	1.30	1.44	4626	1.16	1.22	1.17	1.22

arXiv:2406.12961

Numerically stable in extrapolation regions

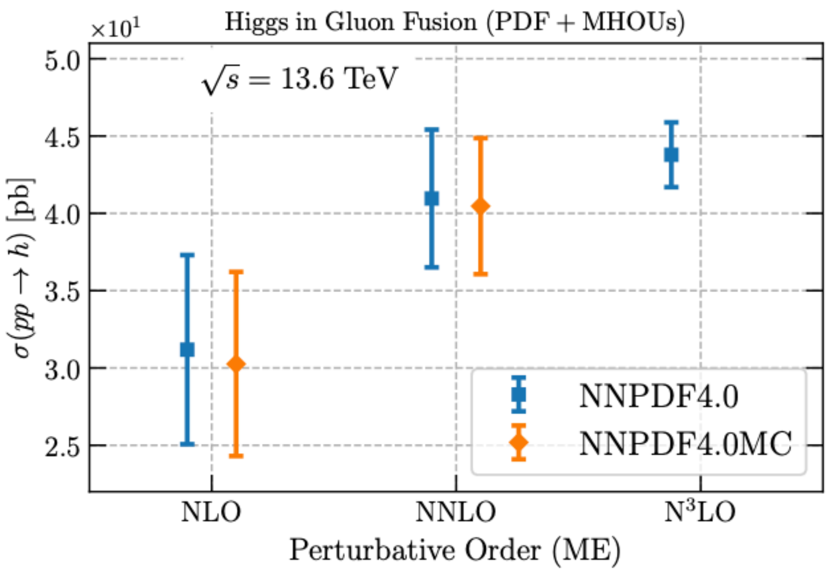


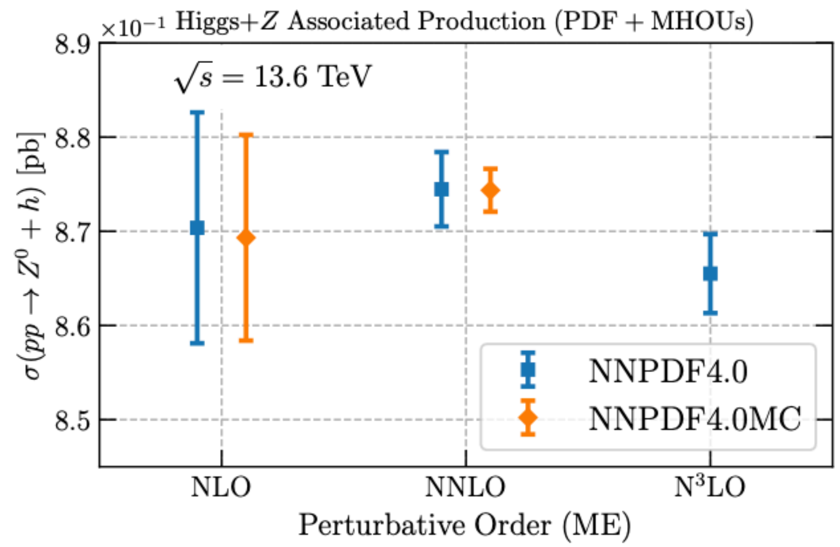
Positive, steeply rising **small-*x* gluon**

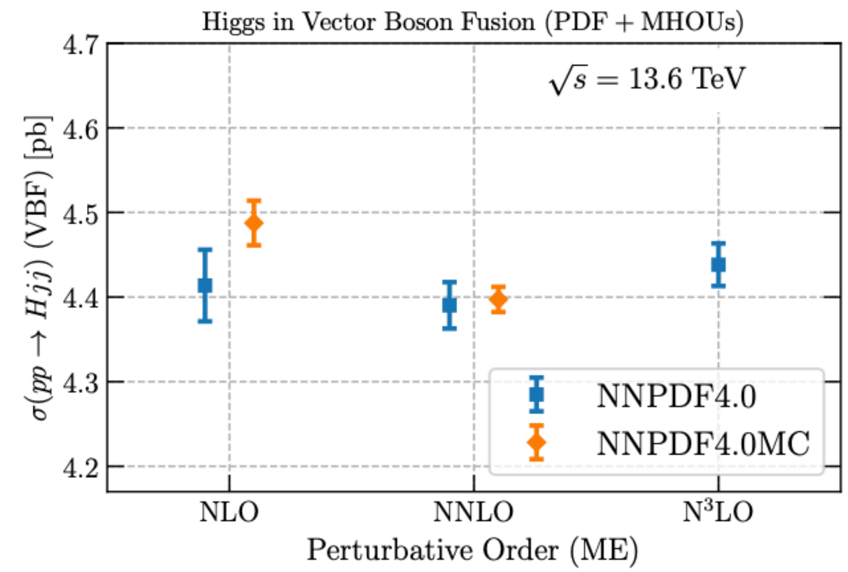


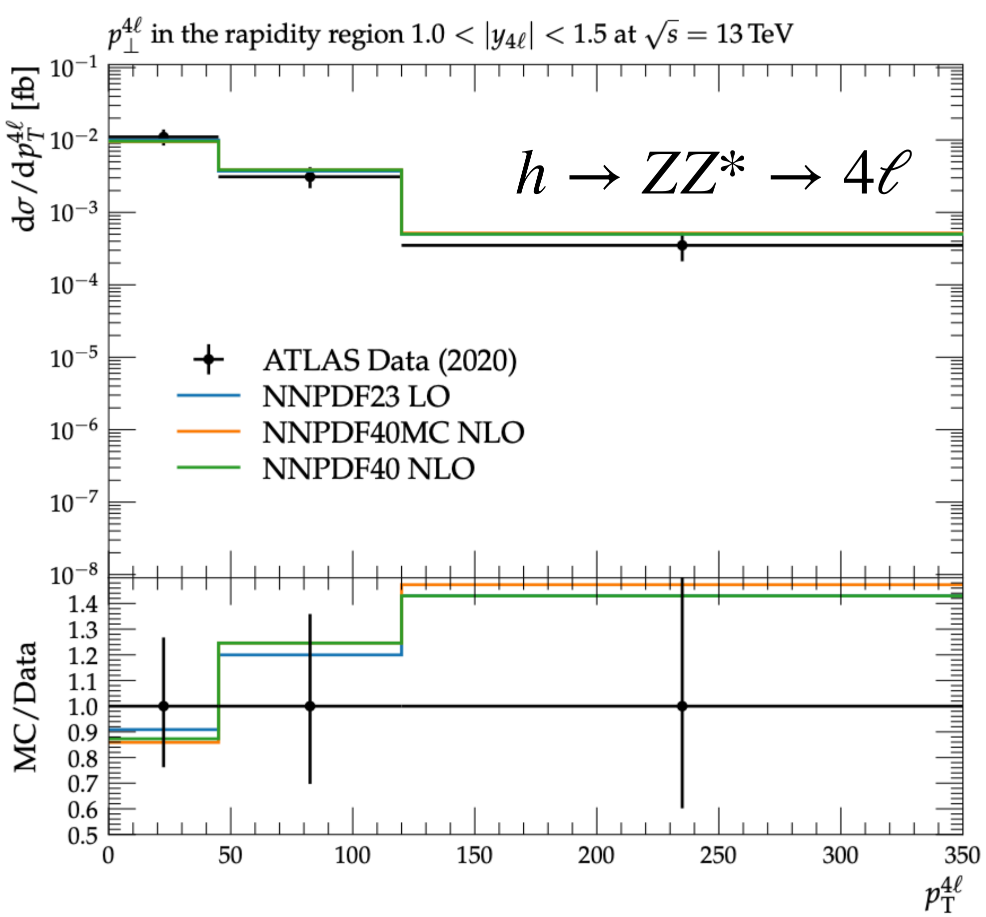
NNPDF4.0MC & Higgs Physics

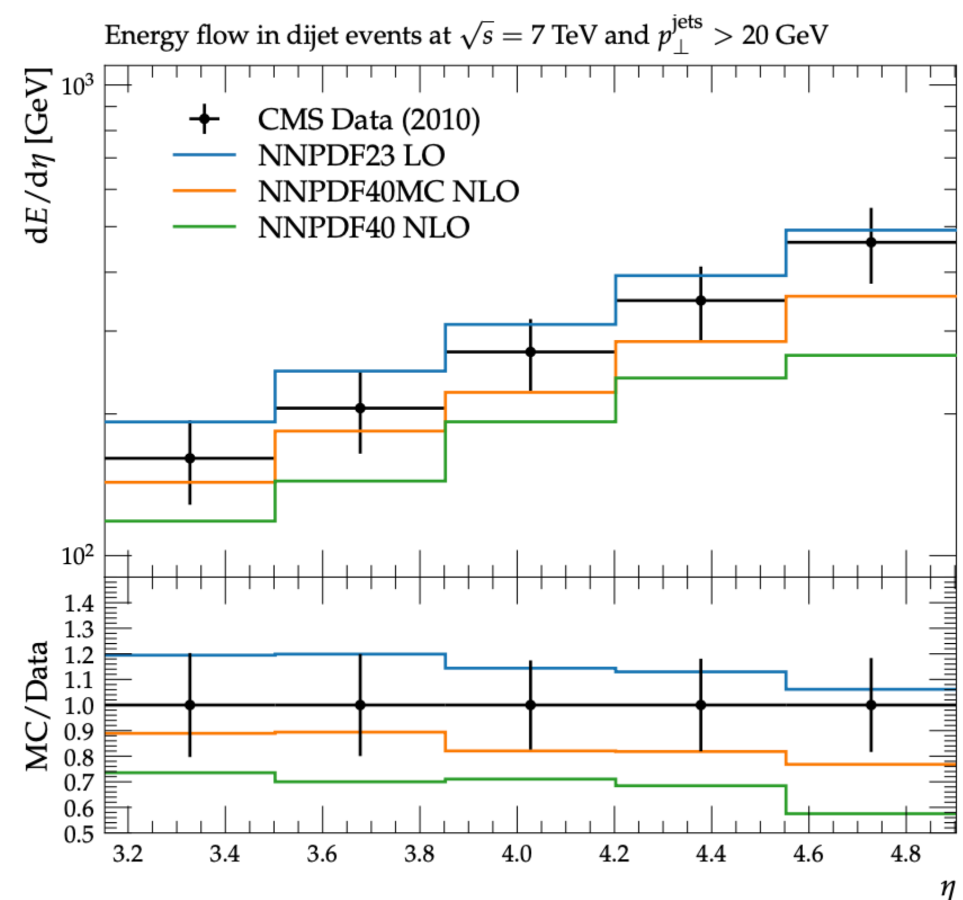
- For **Higgs production cross-sections**, MC PDFs variants close to regular (N)NLO PDFs











- For **soft-QCD processes** MC tuning is required

NNPDF4.0MC enables simultaneous description of **both hard and soft QCD process** relevant for Higgs physics

Summary and outlook

- Improving our understanding of PDFs is **essential for Higgs physics**
- Despite recent progress, differences between PDF sets remain both in **central values and in uncertainties** for many Higgs cross-sections
- The NNPDF4.0 aN^3LO+QED determination enables **consistent N^3LO calculations** of Higgs cross-sections while accounting for QED corrections and the photon PDF.
- The combination of MSHT20 and NNPDF4.0 aN^3LO (QED) PDFs leads to **large shifts** for ggF and VBF cross-sections as compared to PDF4LHC21: how to deal with these?
- First determination of the strong coupling from aN^3LO+QED calculations and validated with closure tests: **agrees with the PDG average** and latest lattice QCD results
- PDFs tailored for **NLO and NNLO Monte Carlo generators** suitable for the exclusive description of Higgs production and decay

Summary and outlook

- Improving our understanding of PDFs is **essential for Higgs physics**

- Despite recent progress, differences between PDF sets remain both in **central values and in uncertainties** for many Higgs cross-sections

- The NNPDF4.0 aN^{3}LO+QED determination enables [illegible] **ations** of Higgs cross-sections while accounting for QED co[illegible] PDF.

- The combination of M[illegible] aN^{3}LO (QED) PDFs leads to **large shifts** for ggF and VBF cr[illegible] to PDF4LHC21: how to deal with these?

- First dete[illegible]ation of the strong coupling from aN^{3}LO+QED calculations and validated with closure tests: **agrees with the PDG average** and latest lattice QCD results

- PDFs tailored for **NLO and NNLO Monte Carlo generators** suitable for the exclusive description of Higgs production and decay

Thanks for your attention