Updates from NNPDF

A determination of $\alpha_s(m_Z)$ from aN^{3}LO global PDF analysis

Roy Stegeman
The University of Edinburgh

QCD@LHC
Freiburg, 8 October 2024

NNPDF4.0 timeline

Sep 2021	NNPDF4.0	**[arXiv:2109.02653]**
Sep 2021	NNPDF4.0 (code)	**[arXiv:2109.02671]**
Oct 2022	Intrinsic charm	**[arXiv:2208.08372]**
Sep 2022	PDFs and BSM searcher (A_{FB})	**[arXiv:2209.08115]**
Nov 2023	Asymmetric intrinsic charm	**[arXiv:2311.00743]**
Jan 2024	NNPDF4.0 QED	**[arXiv:2401.08749]**
Jan 2024	NNPDF4.0 MHOU	**[arXiv:2401.10319]**
Feb 2024	NNPDF4.0 aN^{3}LO	**[arXiv:2402.18635]**
Jun 2024	NNPDF4.0 QED+aN^{3}LO	**[arXiv:2406.01779]**
Jun 2024	NNPDF4.0 for MC event generators	**[arXiv:2406.12961]**
4Q 2024	Improved hyper optimization on GPUs	**[in preparation]**
4Q 2024	Assessment of modern PDF sets on Run II data	**[in preparation]**
4Q 2024	Closure tests with inconsistent data	**[in preparation]**
4Q 2024	α_s extraction at aN^{3}LO	**[in preparation]**
2025/26	NNPDF4.1	**[in preparation]**

NNPDF4.0 timeline

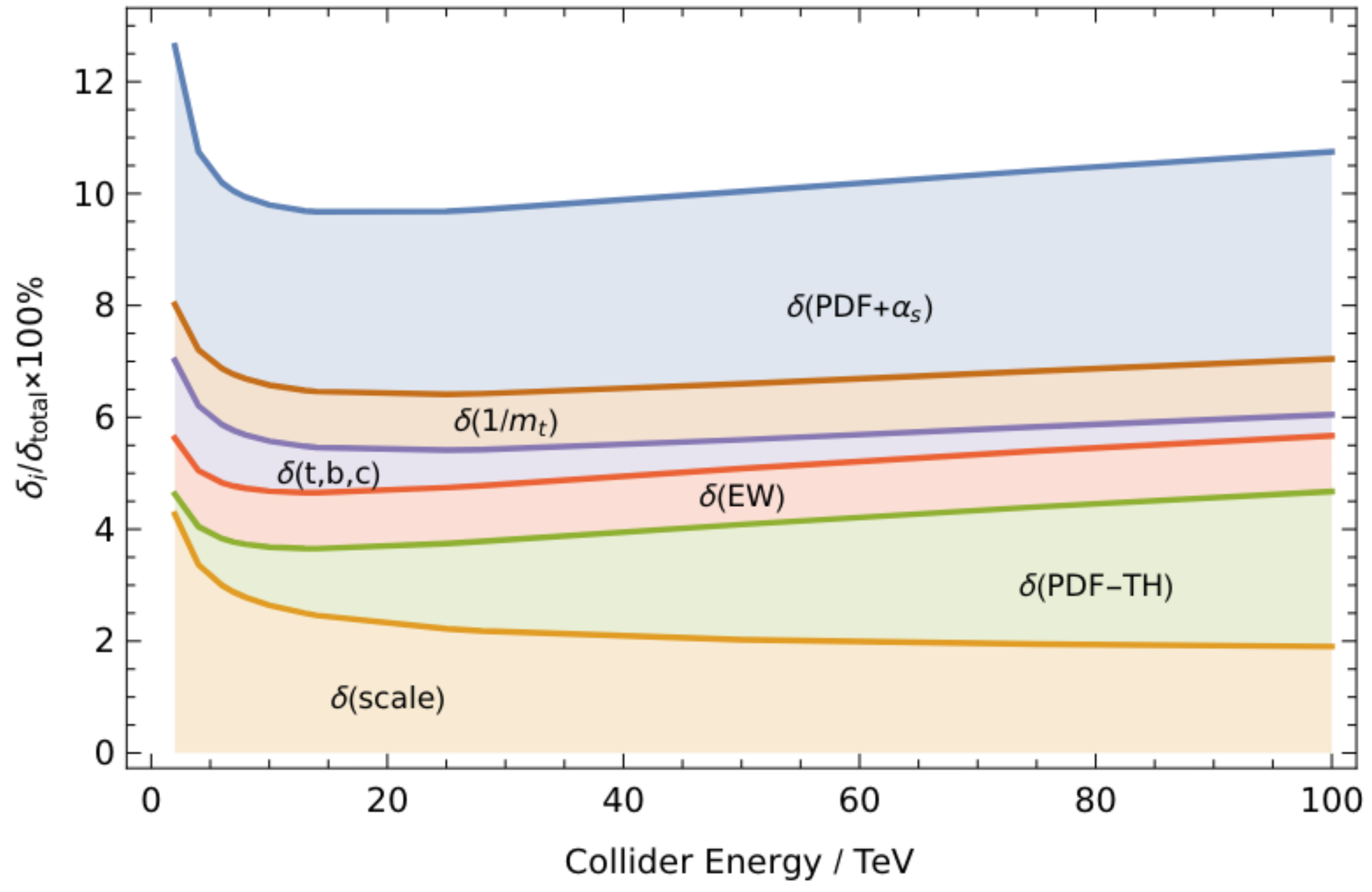
Sep 2021	NNPDF4.0	[arXiv:2109.02653]
Sep 2021	NNPDF4.0 (code)	[arXiv:2109.02671]
Oct 2022	Intrinsic charm	[arXiv:2208.08372]
Sep 2022	PDFs and BSM searcher (A_{FB})	[arXiv:2209.08115]
Nov 2023	Asymmetric intrinsic charm	[arXiv:2311.00743]
Jan 2024	**NNPDF4.0 QED**	[arXiv:2401.08749]
Jan 2024	NNPDF4.0 MHOU	[arXiv:2401.10319]
Feb 2024	**NNPDF4.0 aN^{3}LO**	[arXiv:2402.18635]
Jun 2024	**NNPDF4.0 QED+aN^{3}LO**	[arXiv:2406.01779]
Jun 2024	NNPDF4.0 for MC event generators	[arXiv:2406.12961]
4Q 2024	Improved hyper optimization on GPUs	[in preparation]
4Q 2024	Assessment of modern PDF sets on Run II data	[in preparation]
4Q 2024	Closure tests with inconsistent data	[in preparation]
4Q 2024	**α_s extraction at aN^{3}LO**	[in preparation]
2025/26	NNPDF4.1	[in preparation]

Motivation

Experiment — PDF fit — pQCD

$$\sigma = \sum_{ij} f_i \otimes f_j \otimes \hat{\sigma}_{ij}$$

- Predictions at particle colliders such as the LHC use two main ingredients:
  - Matrix elements (MEs)
  - Parton distribution functions (PDFs)
- Much progress has been made in the computation of MEs at N^{3}LO
- **PDF uncertainties are a bottleneck** for many LHC precision calculations
- Most widely used PDF sets are at NNLO and without theory uncertainties



Sources of uncertainty for inclusive Higgs production

Dulat, Lazopoulos, Mistlberger [arXiv:1802.00827]

Much progress since this plot, in particular:

- NNLO top quark corrections [arXiv:2105.04436]
- Mixed QDC-EW corrections [arXiv:2010.09451] [arXiv:2007.09813]

- Towards N^3LO PDFs [arXiv:2402.18635]
- Photon PDF
- $\alpha_s(M_Z)$ from NNPDF4.0

QCD corrections for aN³LO PDFs

A PDF fit requires several theory inputs:

- **DGLAP splitting functions** for PDF evolution

$$Q^2 \frac{df_i}{dQ^2} = P_{ij} \otimes f_j$$

- **Matching conditions** for variable flavor number schemes

$$f_i^{(n_f+1)}(x, Q^2) = A_{ij}(x, \alpha_s) f_j^{(n_f)}(x, Q^2)$$

- **Partonic coefficient functions** for the data used in the fit

Splitting Functions (information is partial)

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 $a_{H,g}^3$ [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]

E. Nocera, Workshop on Hadron Physics and Opportunities Worldwide
Dalian, China, August 2024

Approximate N^{3}LO splitting functions

$$P_{ij} = \alpha_s P_{ij}^{(0)} + \alpha_s^2 P_{ij}^{(1)} + \alpha_s^3 P_{ij}^{(2)} + \alpha_s^4 P_{ij}^{(3)}, \quad i,j = q,g$$

Complete analytic results for the N^{3}LO splitting functions are not available
Approximations are constructed from partial results

Large-n_f limit $\mathcal{O}(n_f^3)$, $P_{NS}^{(n_f^2)}$ [arXiv:1610.07477], $P_{qq,PS}^{(n_f^2)}$ [arXiv:2308.07958], $P_{qg}^{(n_f^2)}$ [arXiv:2310.01245]

Small-x limit [arXiv:1805.06460] [arXiv:2202.10362]

Large-x limit [arXiv:2205.04493], [arXiv:1911.10174], [arXiv:0912.0369]

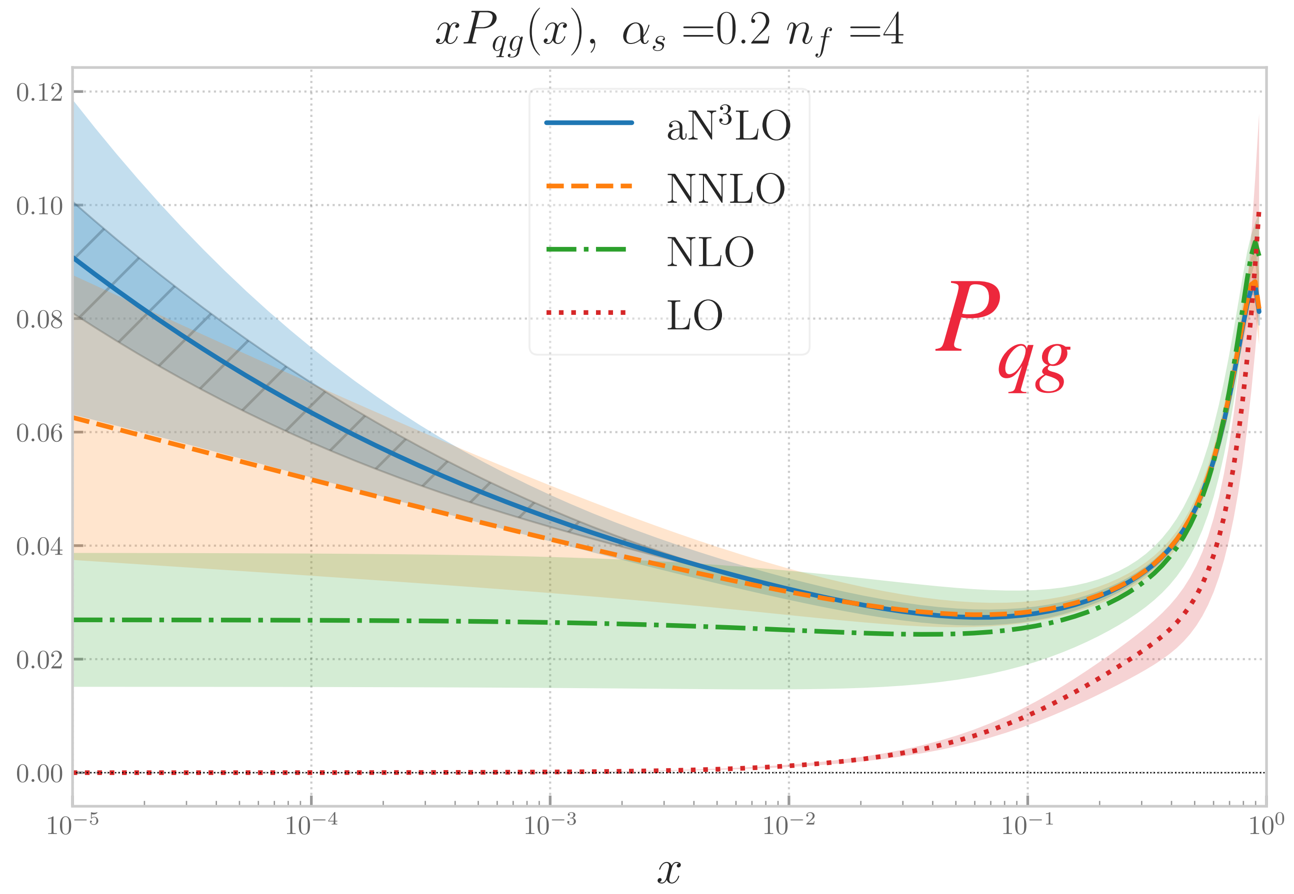
Mellin moments [arXiv:1707.08315] [arXiv:2111.15561], [arXiv:2302.07593], [arXiv:2307.04158],[arXiv:2310.05744], ([arXiv:2404.09701], not included)

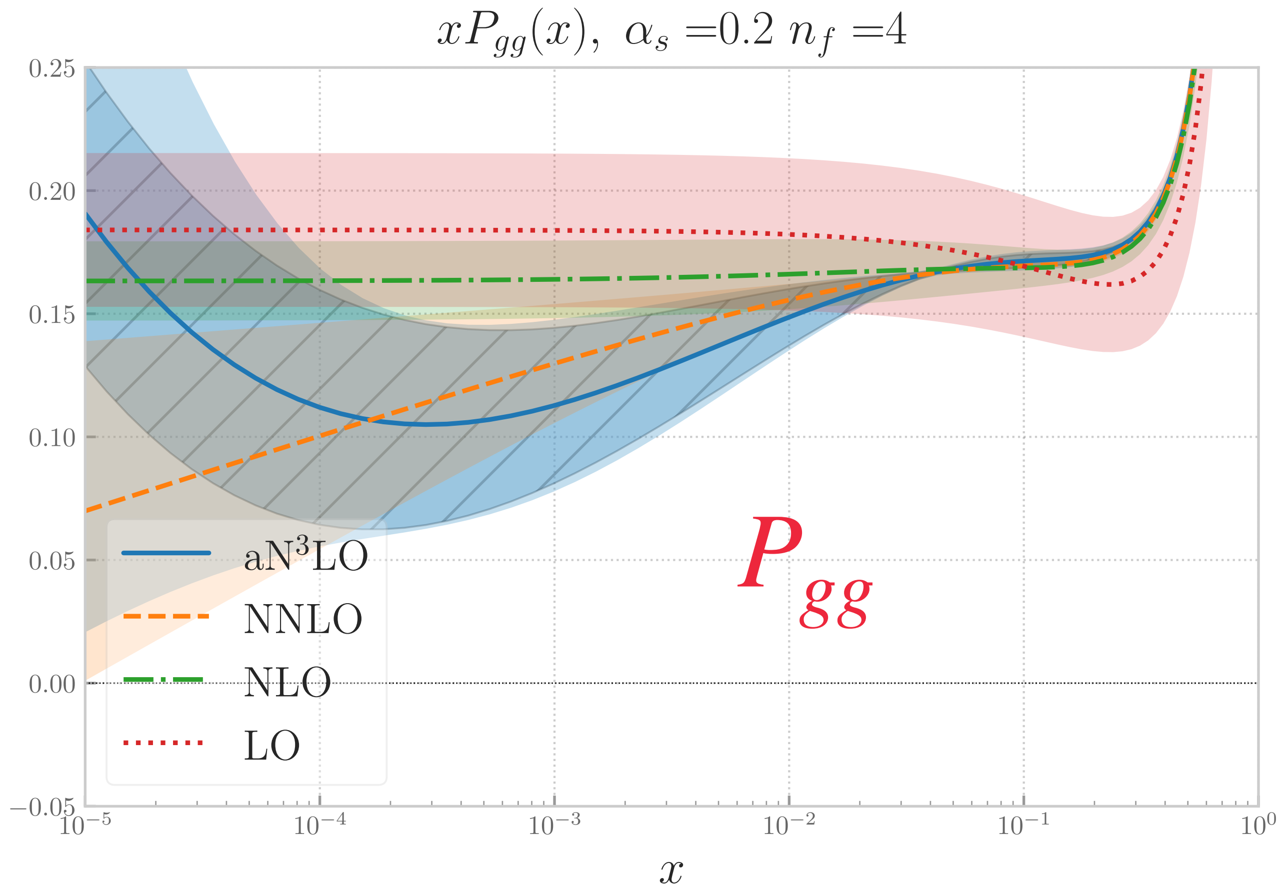
How do we use this information?

1. Parametrize $P_{ij}^{(3)}$ such that it:
   - matches the **small-x and large-x limits**
   - reproduces the **known moments**
2. Vary parametrization choices to obtain an ensemble of approximate aN3LO splitting functions
3. Using the ensemble, determine the **parametrization uncertainty**

See talk by S. Moch yesterday

Approximate N^{3}LO splitting functions

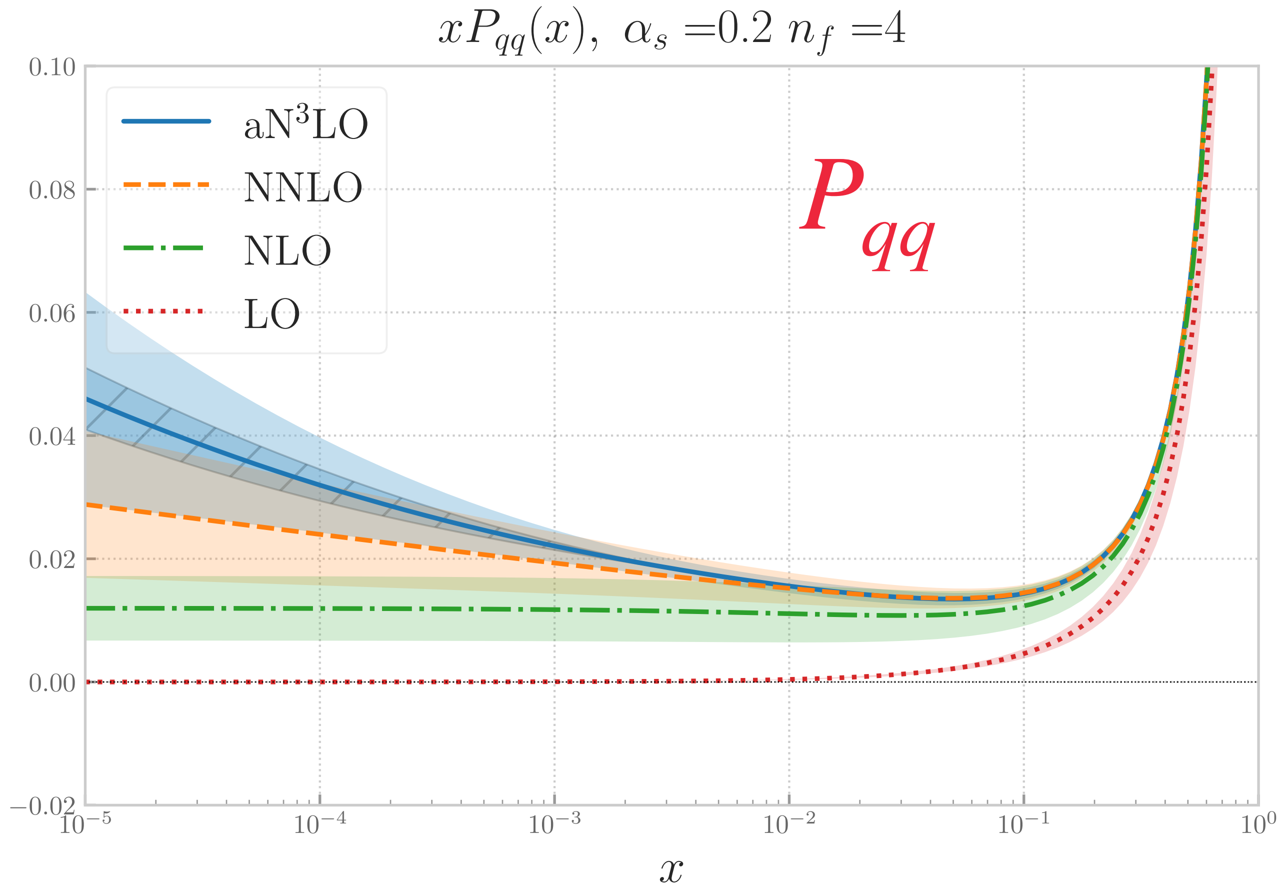


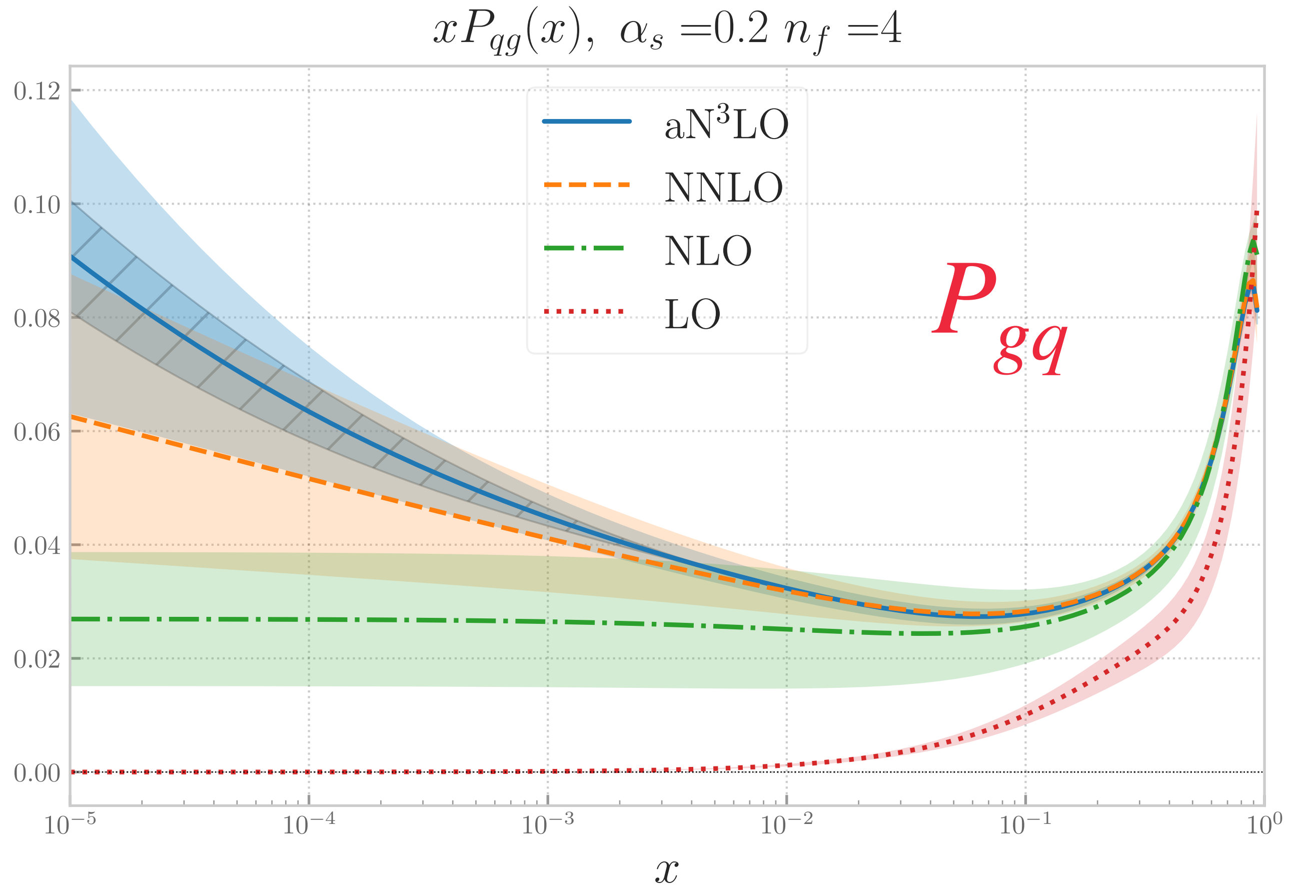


We distinguish two sources of theory uncertainty:

Incomplete Higher Order Uncertainties (IHOUs) due to parametrization of aN3LO contributions (dark band)

Missing Higher Order Uncertainties (MHOUs) due to finite perturbative expansion estimated from scale variations

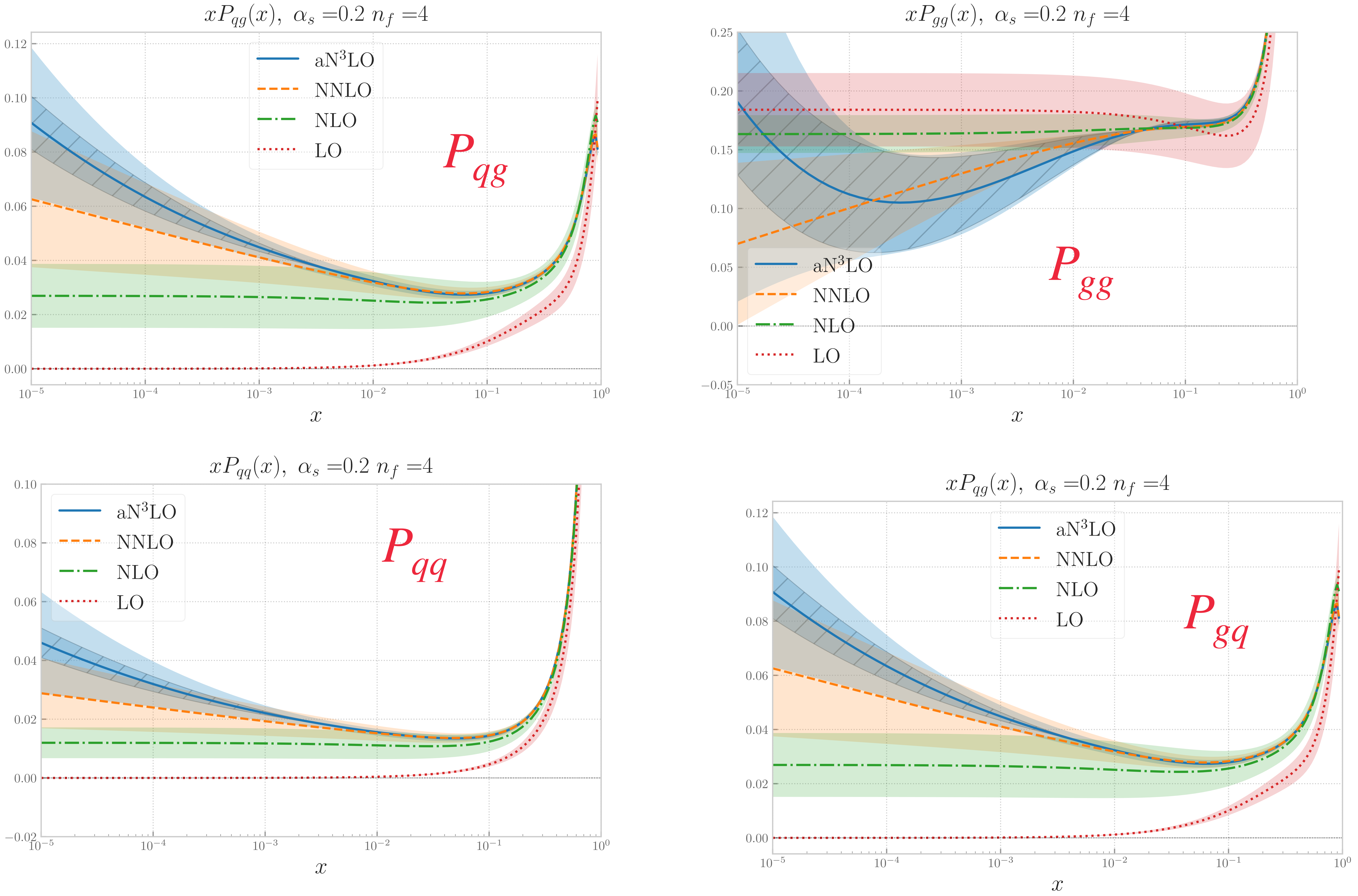




- Good **perturbative stability** within uncertainties
- Small IHOUs in large range of x

For more info see the ***Les Houches benchmark paper*** *[arXiv:2406.16188]*

Approximate N^{3}LO splitting functions



We distinguish two sources of theory uncertainty:

Incomplete Higher Order Uncertainties (IHOUs) due to parametrization of aN3LO contributions (dark band)

Missing Higher Order Uncertainties (MHOUs) due to finite perturbative expansion estimated from scale variations

- Good **perturbative stability** within uncertainties
- Small IHOUs in large range of x

Approximate does not mean poorly-known!

*For more info see the **Les Houches benchmark paper [arXiv:2406.16188]***

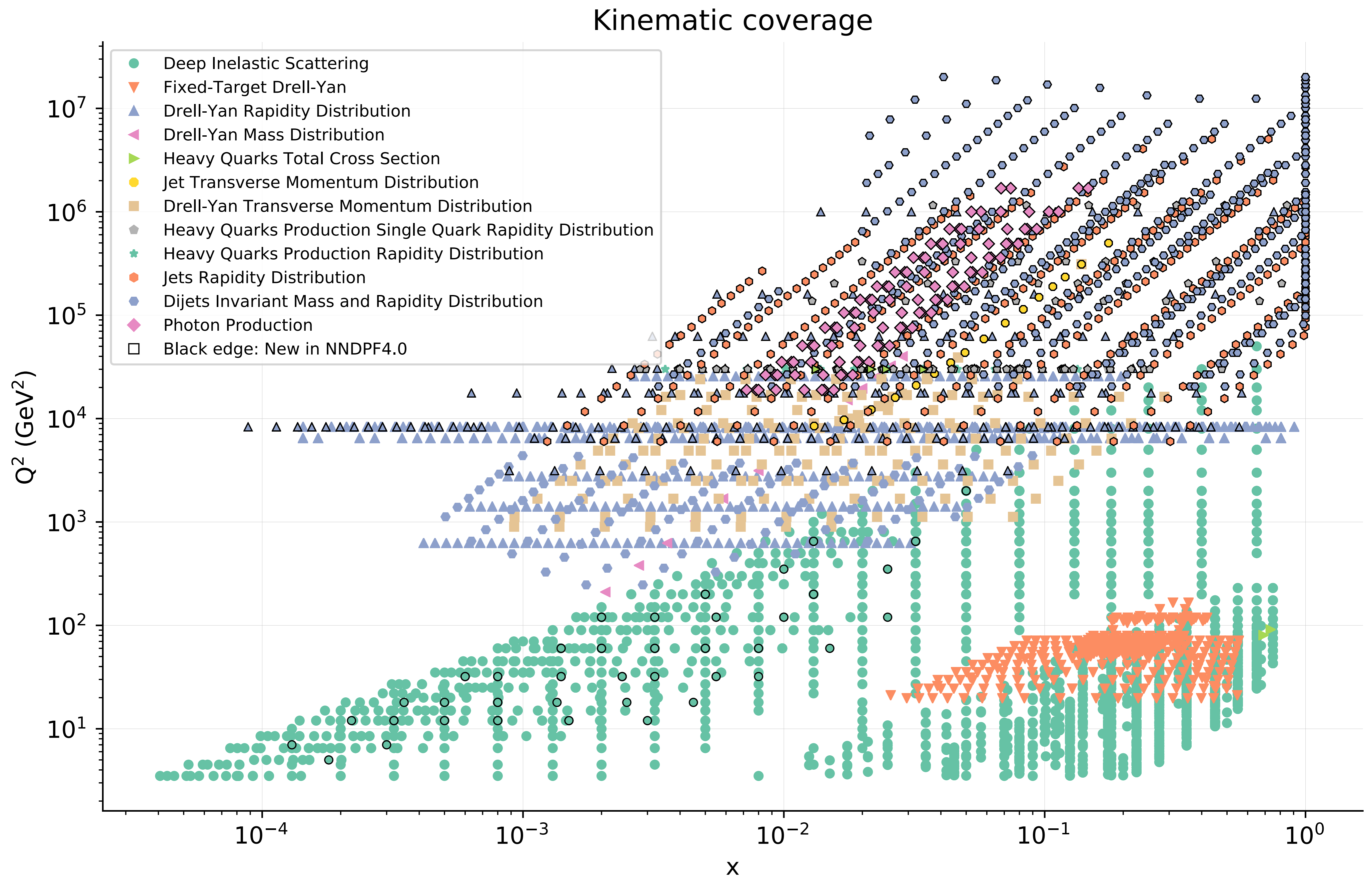
The NNPDF4.0 aN^3LO PDF set

To produce the N^3LO fit, we:

- Use **exact N^3LO** massless DIS coefficient functions
- Include **approximate N^3LO** contributions in DGLAP massive DIS and account for IHOUs
- Use NNLO renormalization scale variations to estimate **unknown N^3LO** terms for hadronic processes

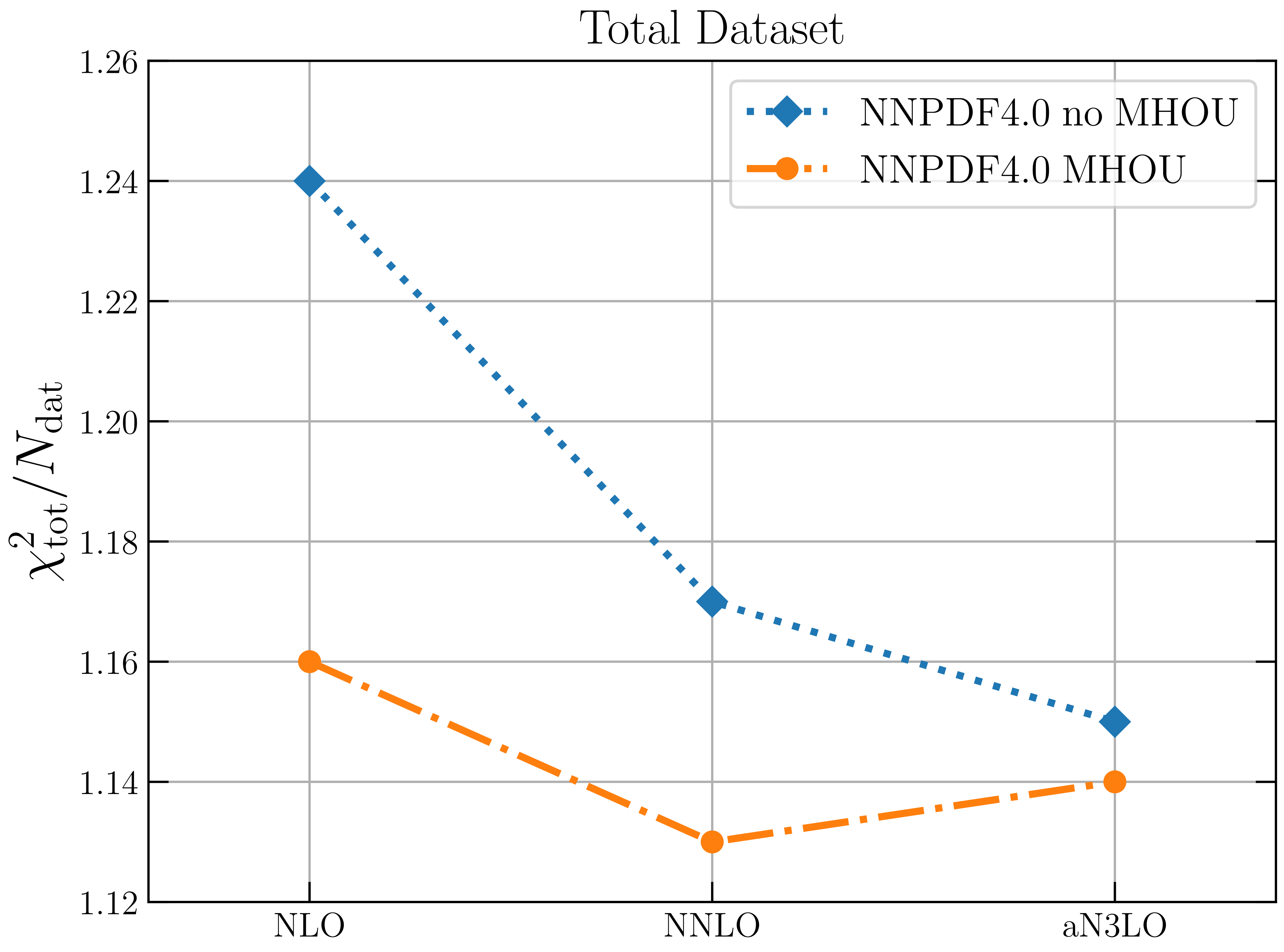
- Treat **theory uncertainties** on the equal footing with experimental uncertainties:

$$\mathrm{Cov}_{\mathrm{tot}} = \mathrm{Cov}_{\mathrm{exp}} + \mathrm{Cov}_{\mathrm{IHOU,DGLAP}} + \mathrm{Cov}_{\mathrm{IHOU,DIS}} + \mathrm{Cov}_{\mathrm{MHOU}}$$



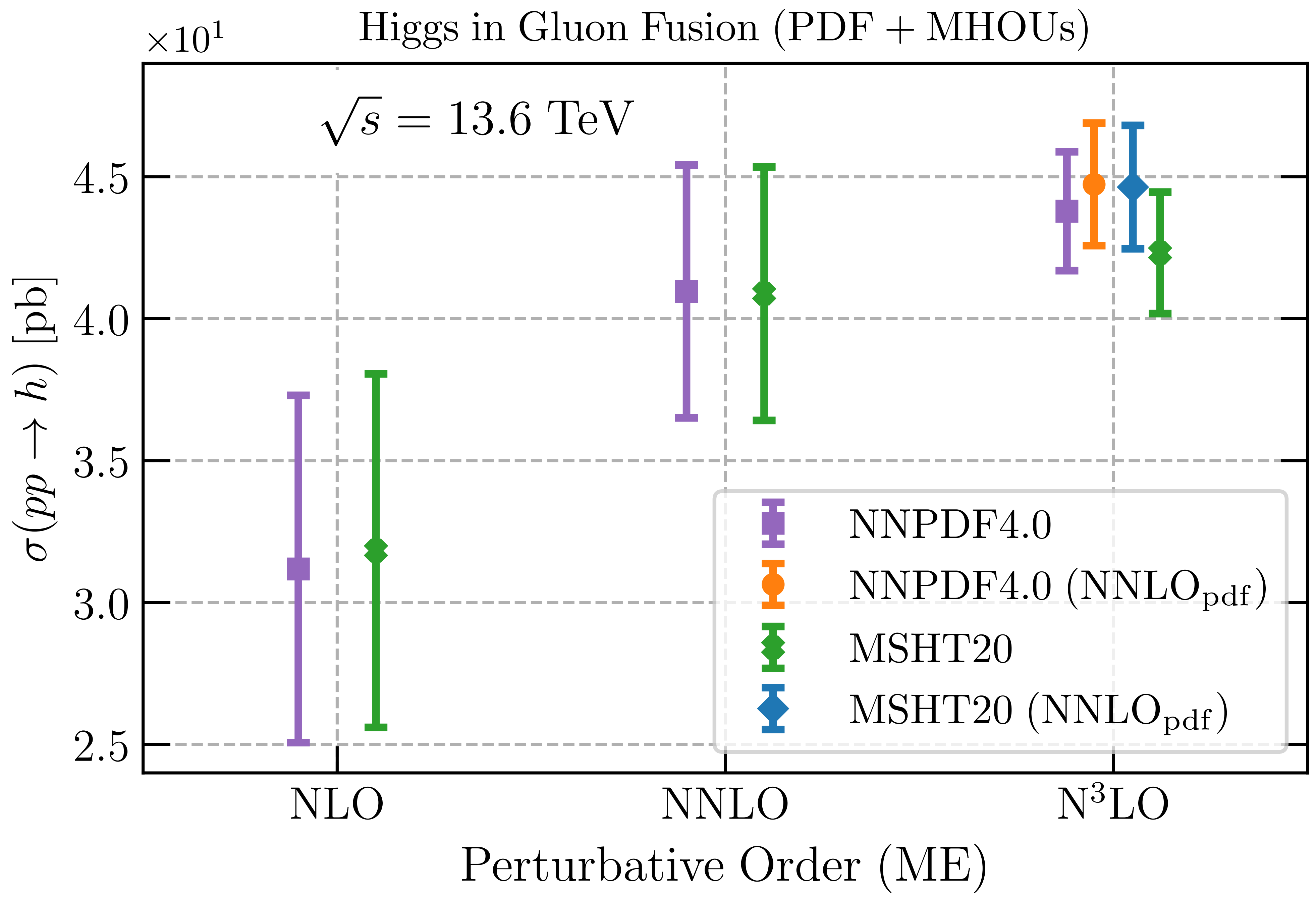
NNPDF4.0 is based on over 4000 datapoints from many processes: DIS, Jets, Top, Drell-Yan, …

Fit quality

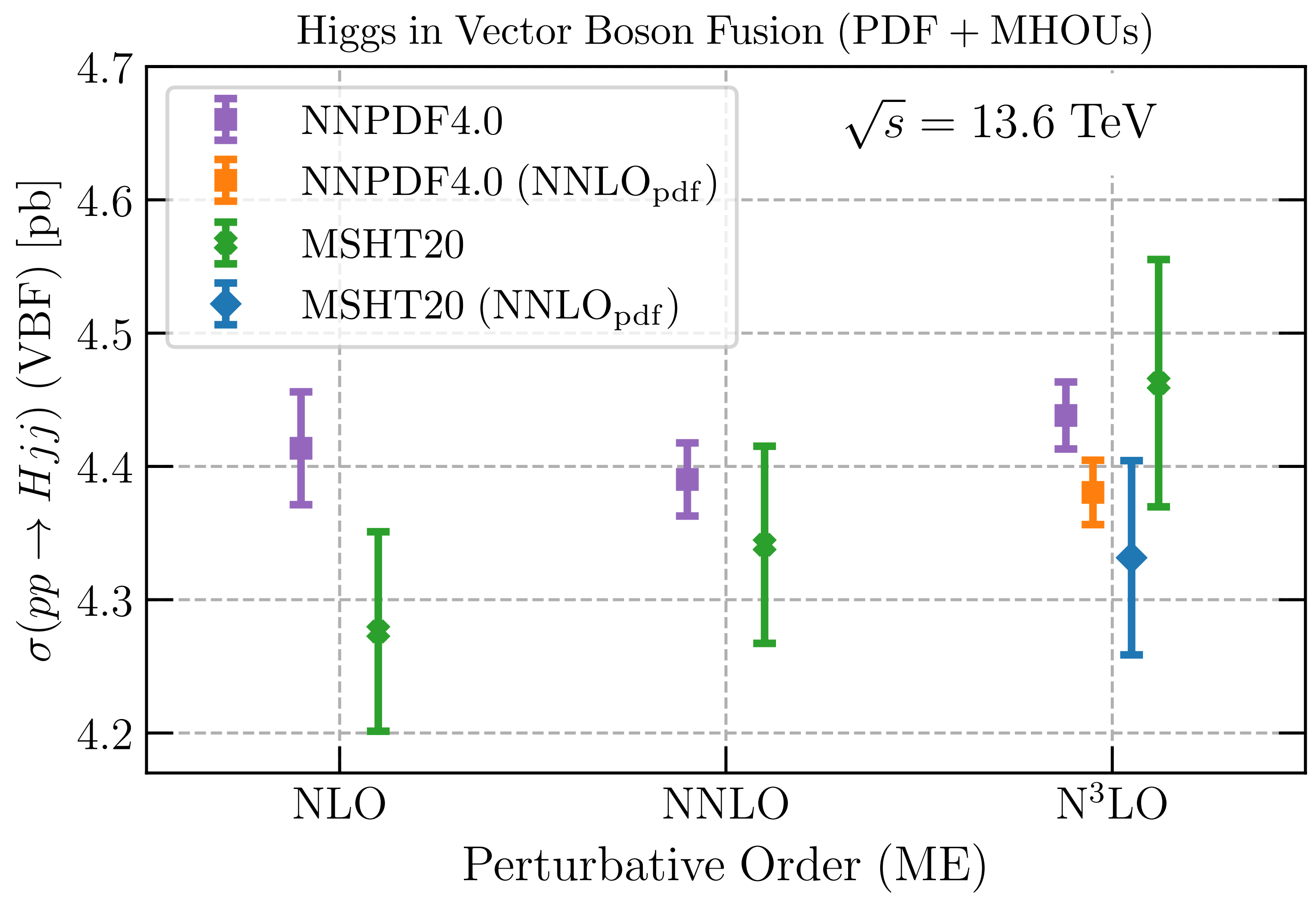


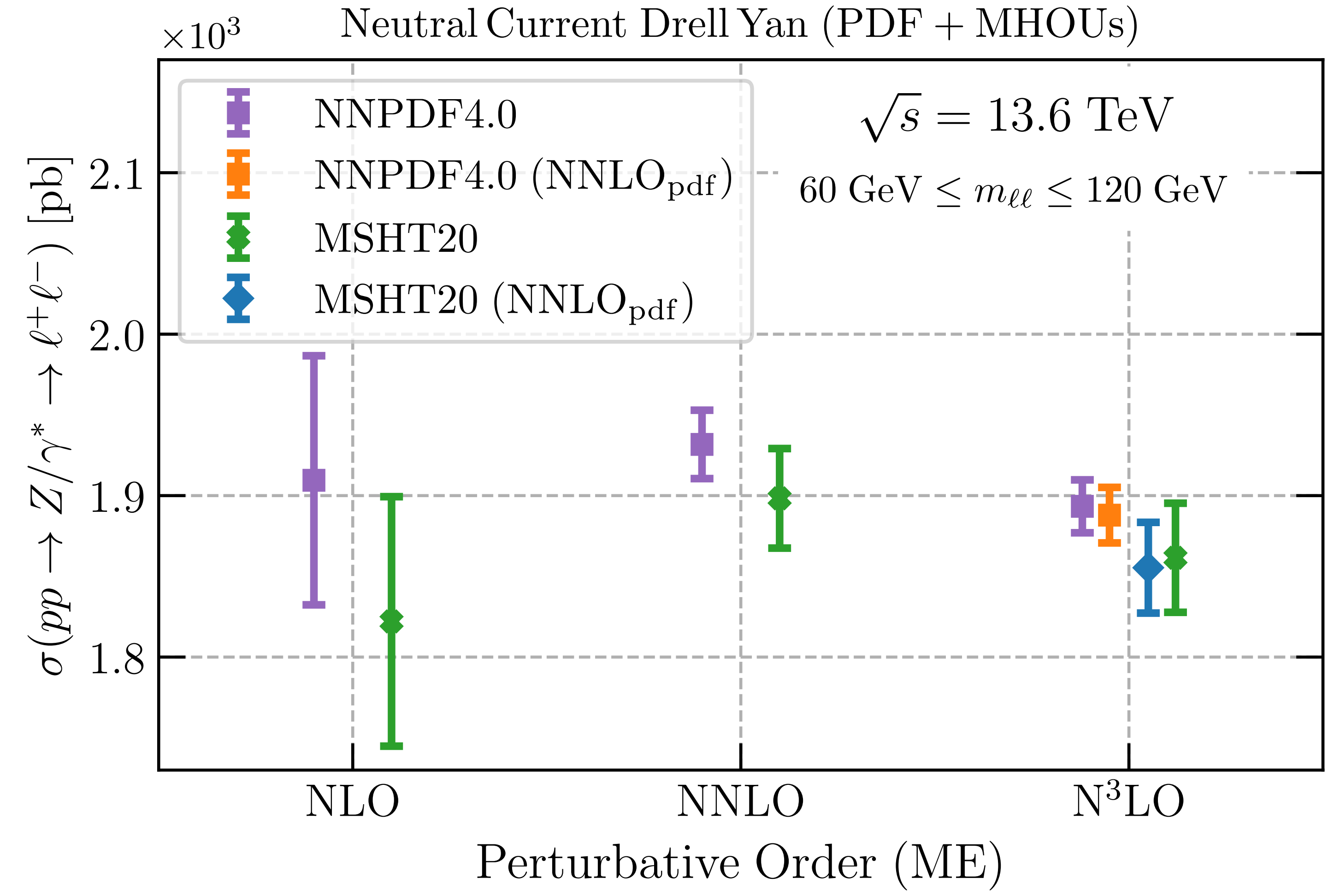
- Without MHOUs the fit improves (lower χ^2) with increasing perturbative order
- With MHOUs the fit depends only weakly on the perturbative order
- At N^3LO MHOUs have a small impact on the χ^2

Impact on LHC cross sections



N^{3}LO PDFs result in a small (~2%) suppression of the Higgs gluon fusion cross section compared to **NNLO PDFs**





Generally **good perturbative convergence** for Higgs in VBF and Drell-Yan

N^{3}LO/NNLO ratio is similar for NNPDF and MSHT **[arXiv:2406.16188]**

- Towards N^{3}LO PDFs
- Photon PDF [arXiv:2401.08749]
- $\alpha_s(M_Z)$ from NNPDF4.0

NNPDF4.0 QED

[arXiv:2401.08749]

So far we considered only QCD evolution, but $\mathcal{O}(\alpha_s^2) \approx \mathcal{O}(\alpha_{em})$

Also **photon initiated contributions** may be relevant

- Modify the DGLAP running to **account for QED corrections**:

$$P = P_{QCD} + P_{QCD\otimes QED}$$

$$P_{QCD\otimes QED} = \alpha_{em} P^{(0,1)} + \alpha_{em}\alpha_s P^{(1,1)} + \alpha_{em}^2 P^{(0,2)}$$

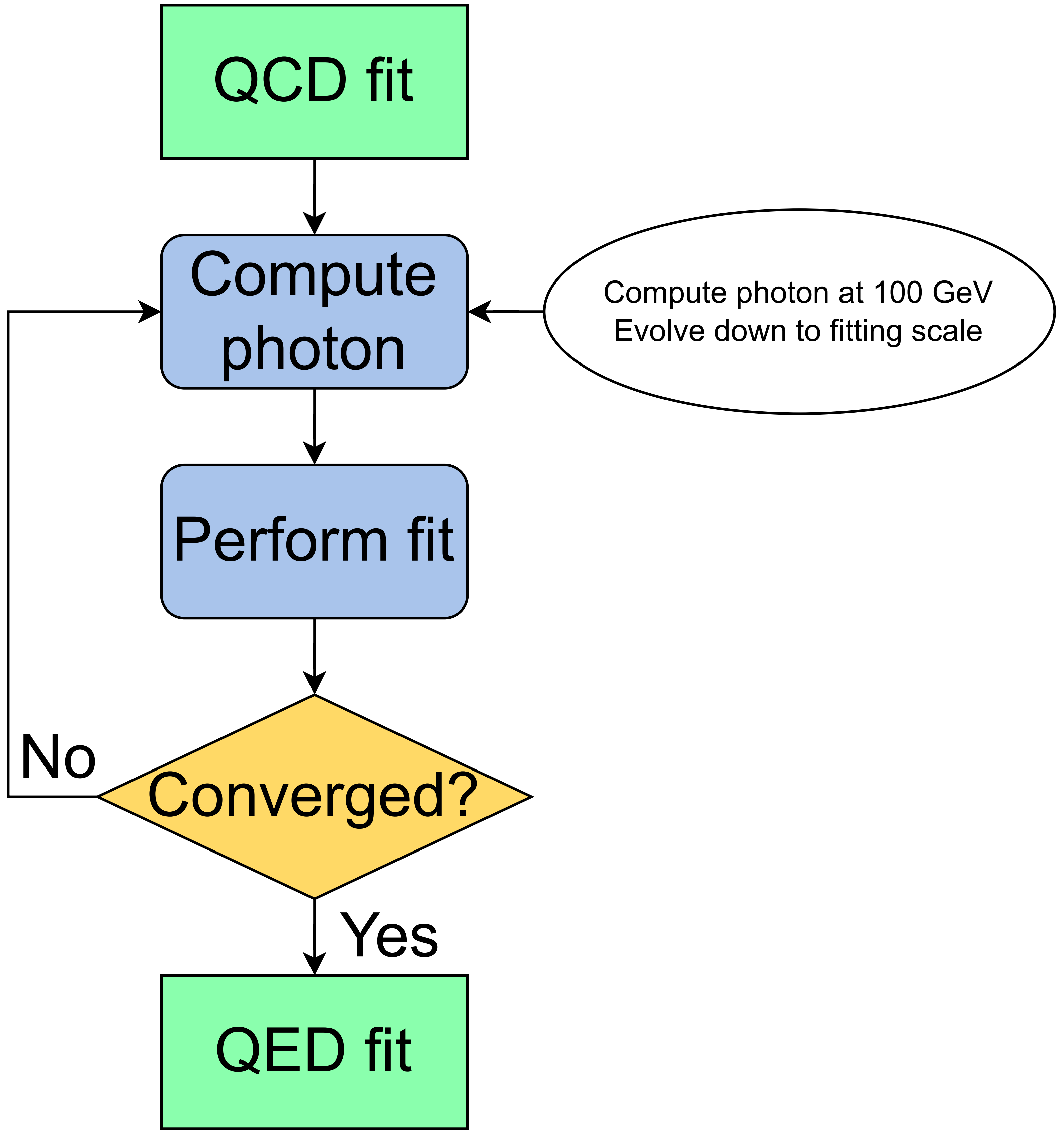
- Data does not provide strong constraints on the photon, but the **photon PDF** can be computed from DIS structure functions: Manohar, Nason, Salam, Zanderighi, [arXiv:1607.04266], [arXiv:1708.01256]

$$x\gamma\left(x,\mu^2\right) = \frac{2}{\alpha\left(\mu^2\right)} \int_x^1 \frac{dz}{z} \left\{ \int_{\frac{m_p^2 x^2}{1-z}}^{\frac{\mu^2}{1-z}} \frac{dQ^2}{Q^2} \alpha^2(Q^2) \left[-z^2 F_L\left(x/z, Q^2\right) + \left(z P_{\gamma q}(z) + \frac{2x^2 m_p^2}{Q^2} \right) F_2\left(x/z, Q^2\right) \right] - \alpha^2\left(\mu^2\right) z^2 F_2\left(x/z, \mu^2\right) \right\}$$

- The **momentum sum rule** needs to account for the photon PDF:

$$\sum_{i=q,\bar{q},g,\gamma} \int_0^1 dx x f_i\left(x, Q^2\right) = 1.$$

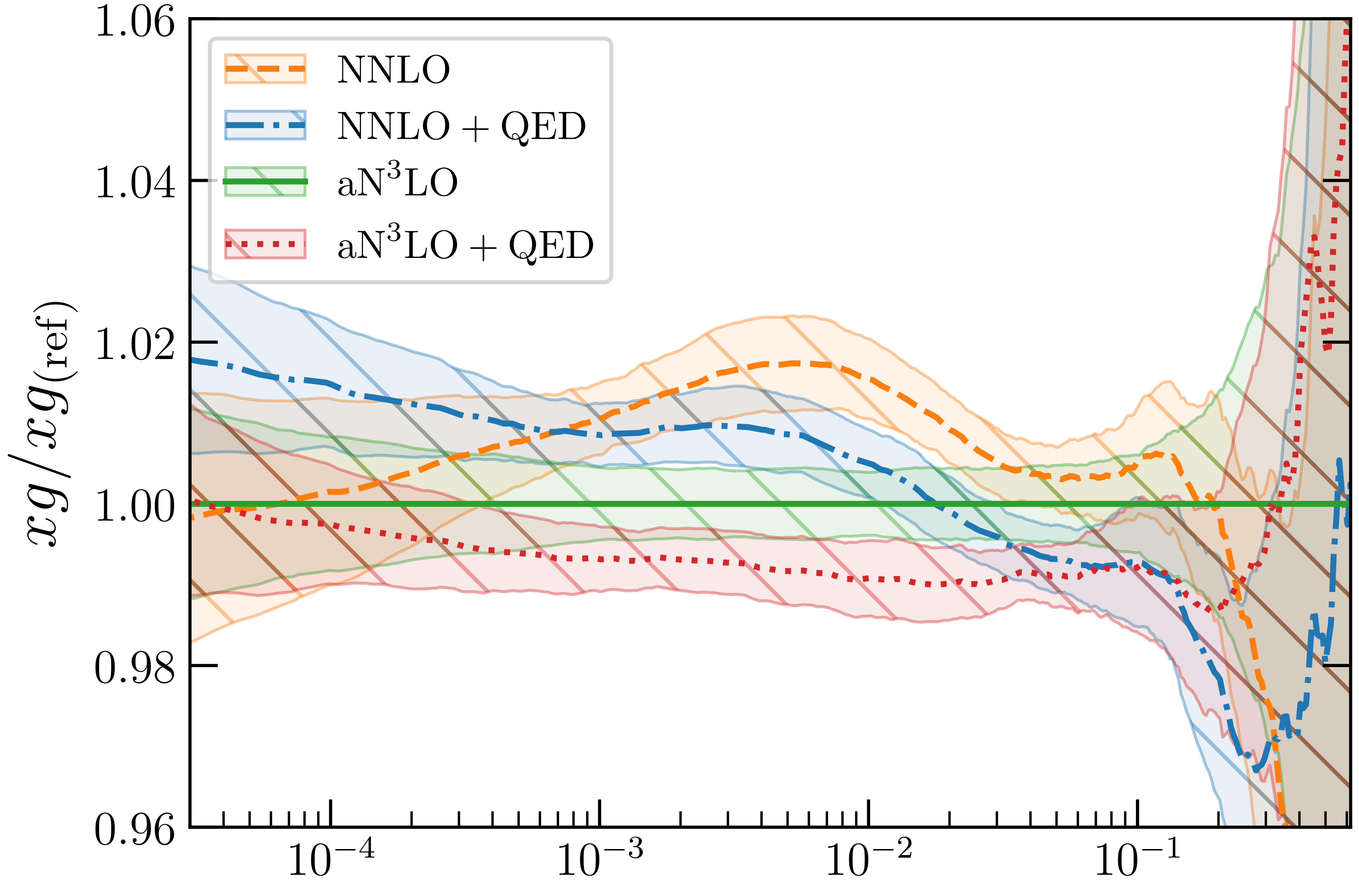
We address the **interplay between the photo and other partons** through an iterative procedure



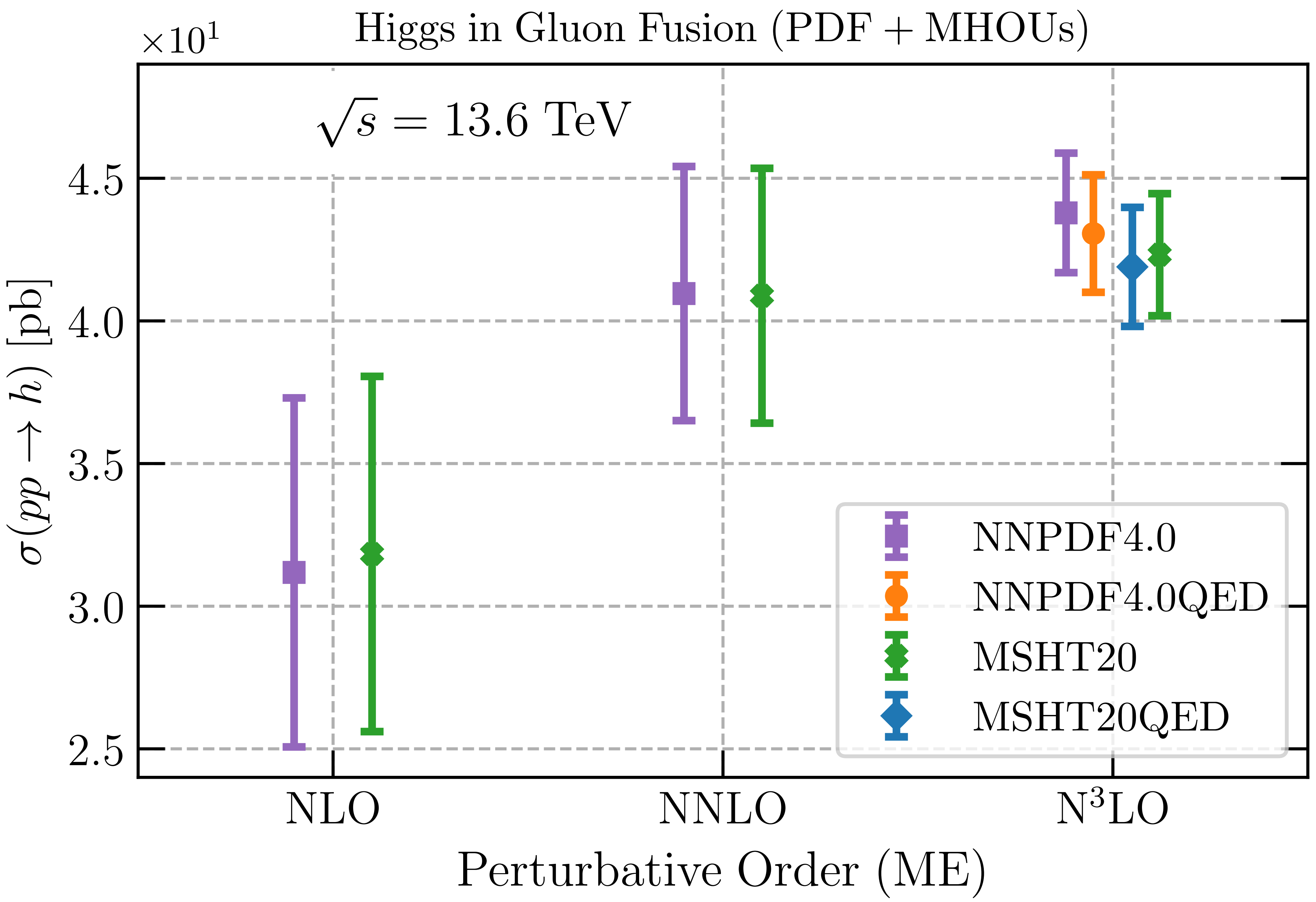
NNPDF4.0 aN³LO QED

[arXiv:2406.01779]

aN³LO+QED is the current state-of-the-art!



- Photon subtracts momentum from the gluon PDF
- QED effect has a similar magnitude as aN3LO

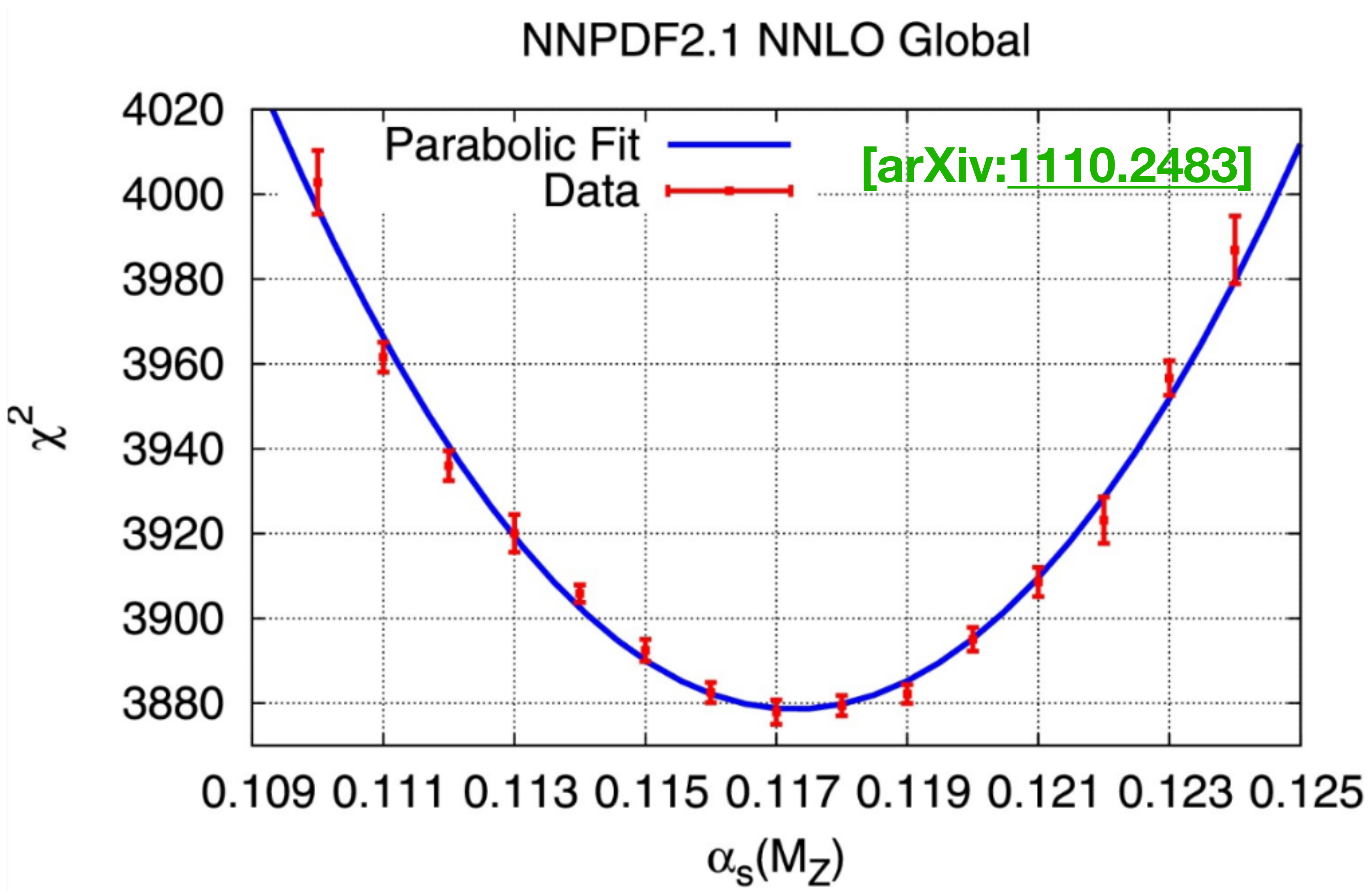


- **aN3LO+QED** result in percent level suppression compared to **pure QCD** in ggH
- Theory uncertainties related to the photon PDF are not included in any pure QCD fits

- Towards N^3LO PDFs
- Photon PDF
- $\alpha_s(M_Z)$ from NNPDF4.0 [in preparation]

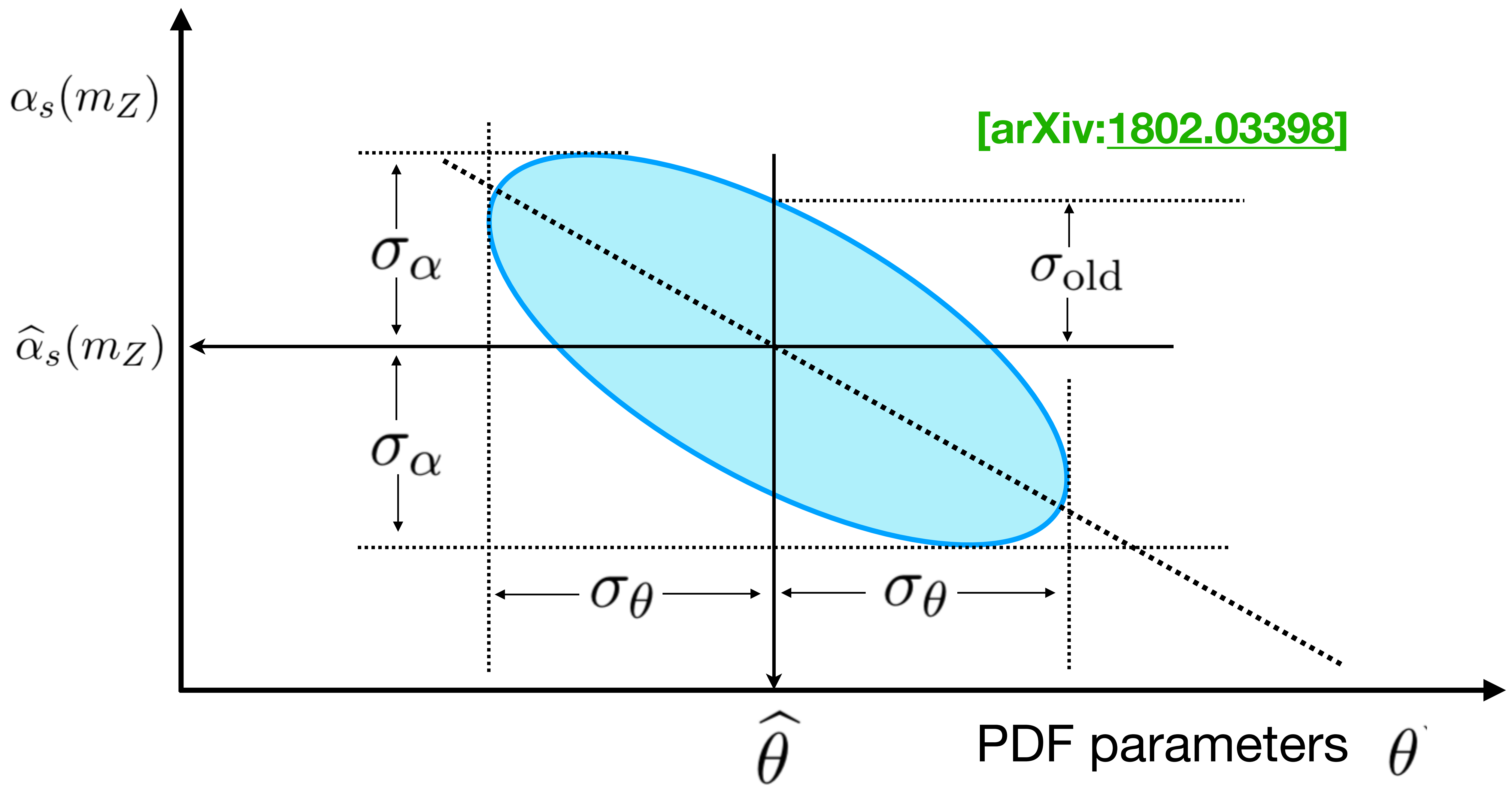
α_s from PDFs

PDFs and α_s are highly correlated so **extracting α_s from collider data requires a simultaneous determination with PDFs**



In most cases α_s is determined by extracting it from a parabolic fit to the χ^2 profile

Uncertainty is determined from $\Delta\chi^2 = 1$



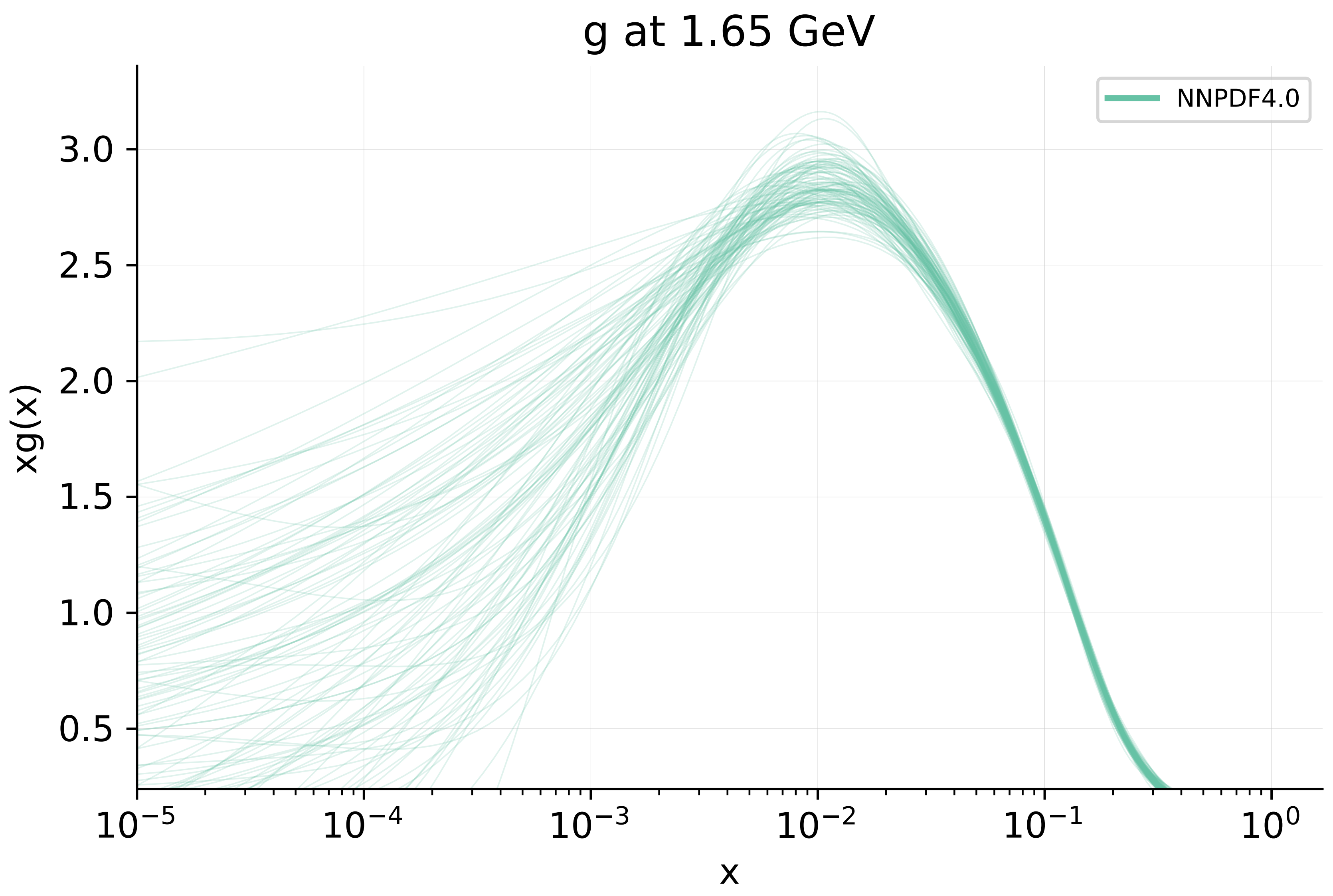
However, the usual methodology neglects correlations between α_s and the PDFs which may lead to underestimated uncertainties

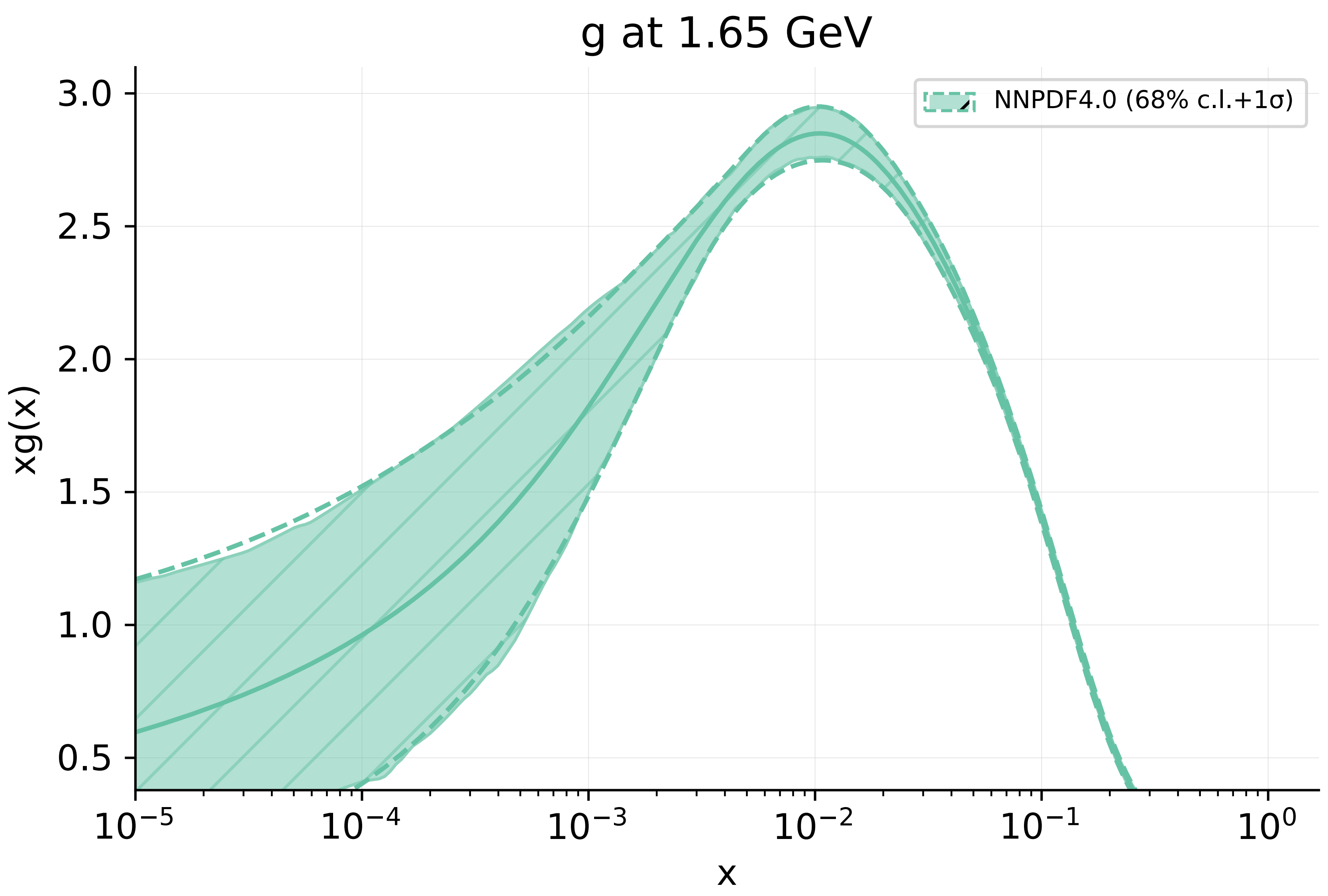
Intermezzo - how to propagate experimental uncertainty to PDFs

An NNPDF set (usually) consists of 100 PDF replicas produced as follows:

1. Assume experimental data is **defined** by a vector of central values and a covariance matrix
2. Sample this distribution to create 100 Monte Carlo replicas in data space
3. Perform a fit to each of the data replicas

➡A PDF set encoding experimental uncertainties

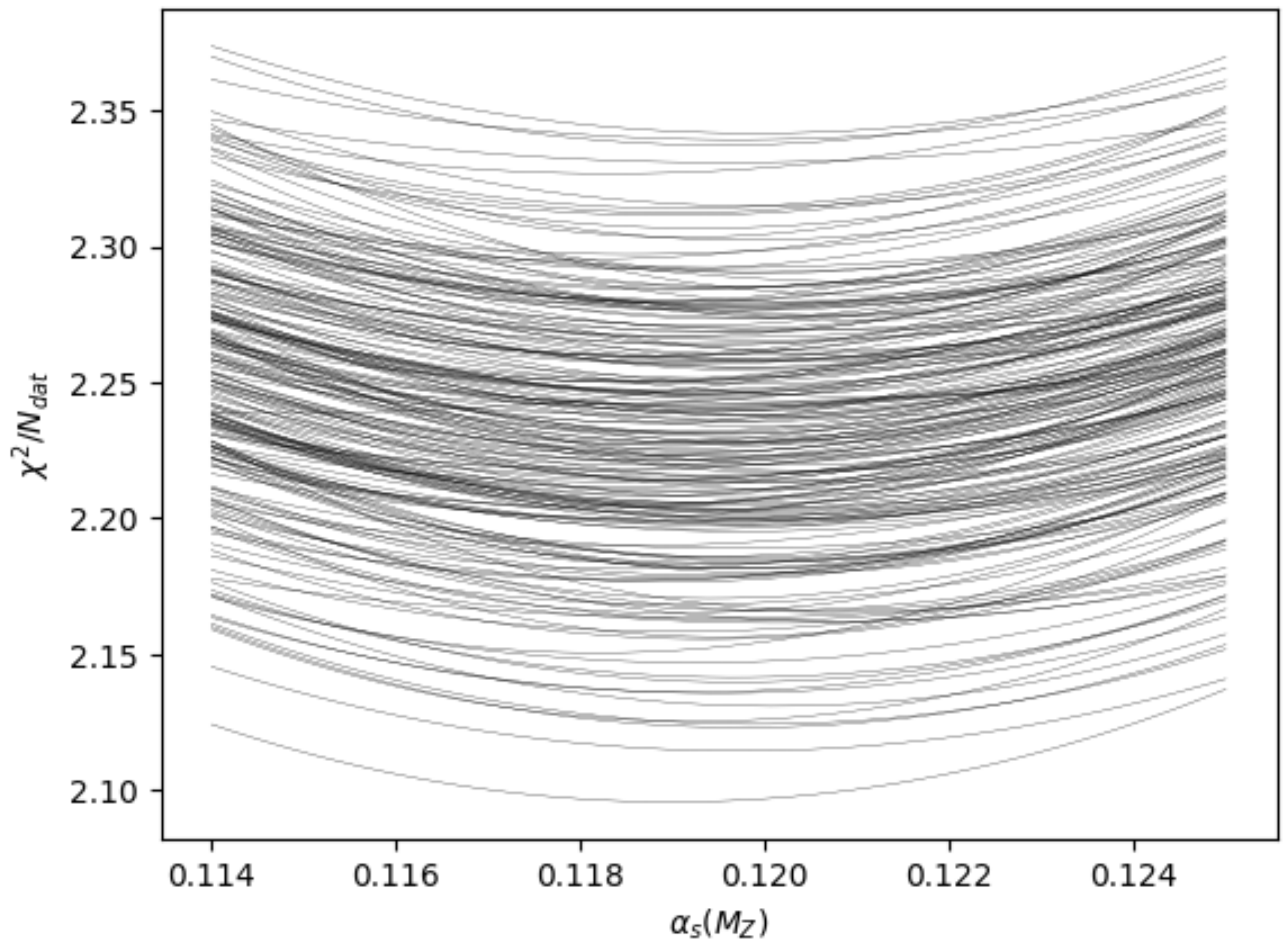




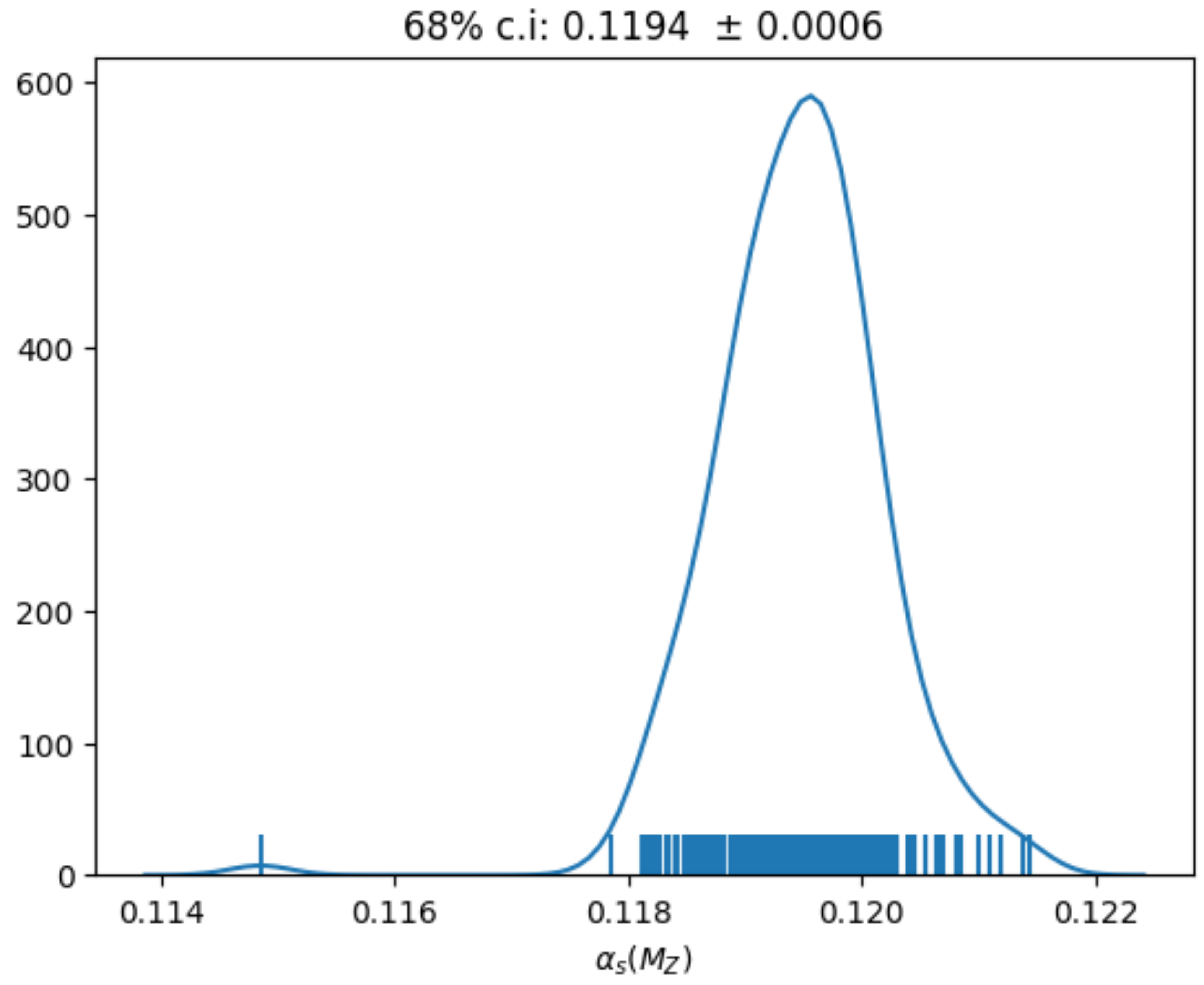
Simultaneous minimization of PDF and α_s

Correlated replicas method

Results **confirmed** by comparison to another method based on Bayesian inference!



Fit the same data replica at different values of α_s and fit a parabola for each replica …



… then look at the distribution of minima of the parabolas

Validating the methodologies

We use closure tests to validate our methodology

Basic idea: generate a global pseudo dataset from theory predictions and extract α_s from this

PDFs at input scale Q_0 $\otimes$ **Wilson coefficients + DGLAP depending on α_s** → **pseudodata** → **methodology** → **extracted α_s**

Is α_s the same (within uncertainties)?

This question is answered through the analysis of various statistical estimators

Results

If we were to determine α_s without MHOUs, it would depend on the perturbative order:

NNLO: $\alpha_s(M_Z) = 0.1204 \pm 0.0004$

aN^{3}LO: $\alpha_s(M_Z) = 0.1200 \pm 0.0003$

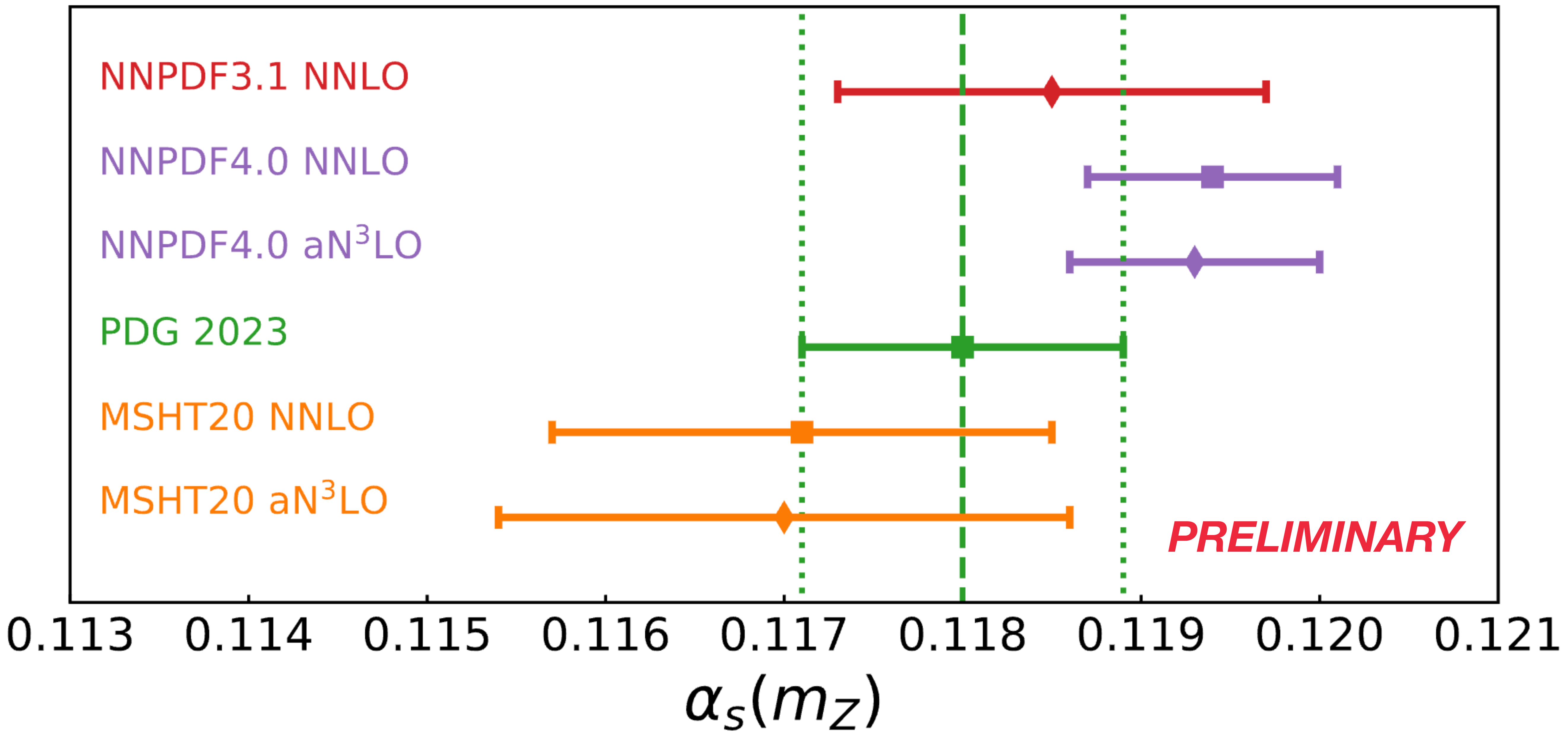
With MHOUs the result is perturbatively stable, i.e. α_s **does not depend on the perturbative order:**

NNLO: $\alpha_s(M_Z) = 0.1194 \pm 0.0007$

aN^{3}LO: $\alpha_s(M_Z) = 0.1193 \pm 0.0007$

The result is in agreement with NNPDF3.1:
$\alpha_s(M_Z) = 0.1185 \pm 0.00012$

This determination will also be updated with QED effects

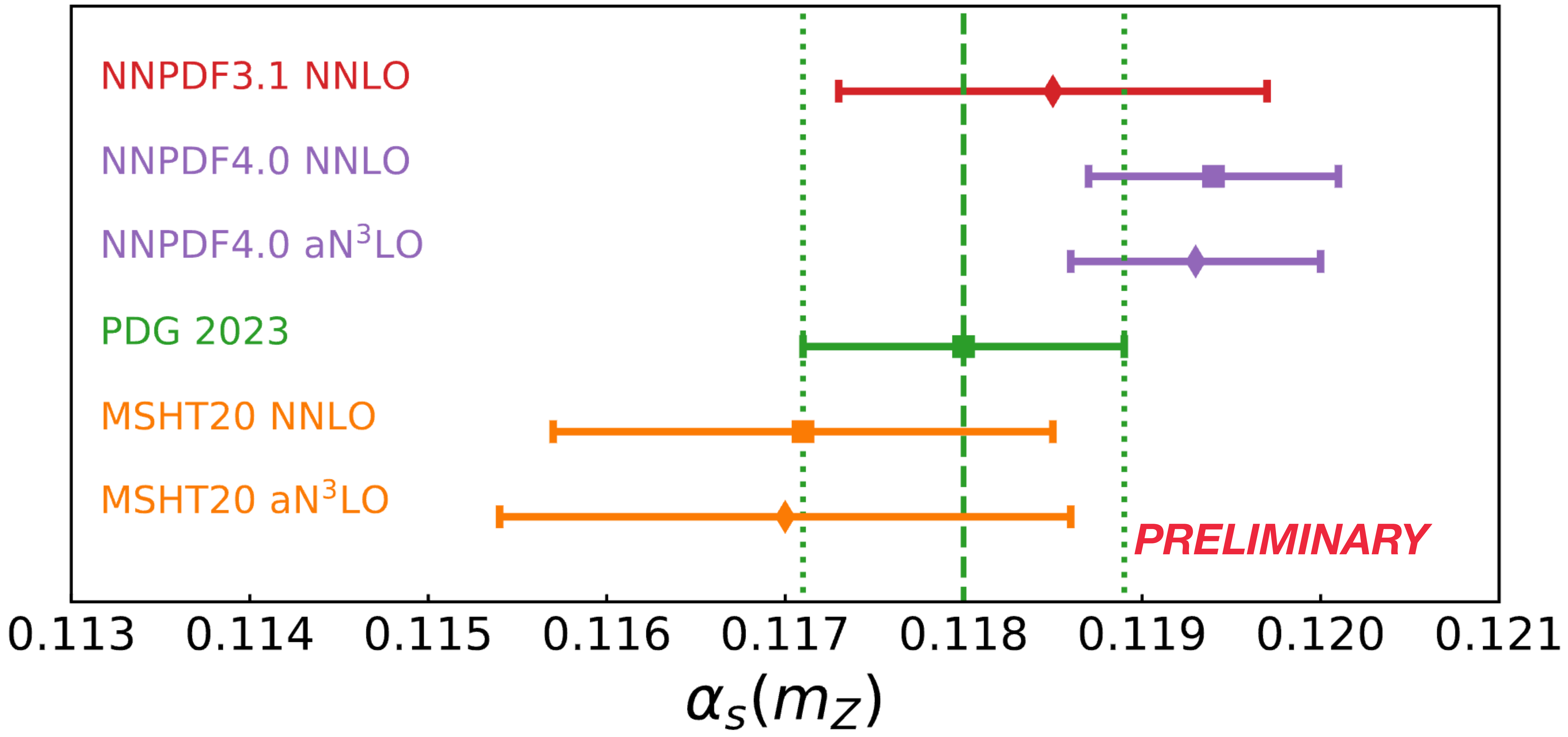


Summary and Outlook

- PDFs are a key ingredient for LHC physics
- aN3LO PDFs allow for a consistent computation of observables at N3LO. Initial results suggest good convergence for Higgs and Drell-Yan production
- SM parameters from collider data require a simultaneous determination with the PDFs
- **The extracted strong coupling constant is perturbatively stable between NNLO and aN3LO:** *PRELIMINARY*

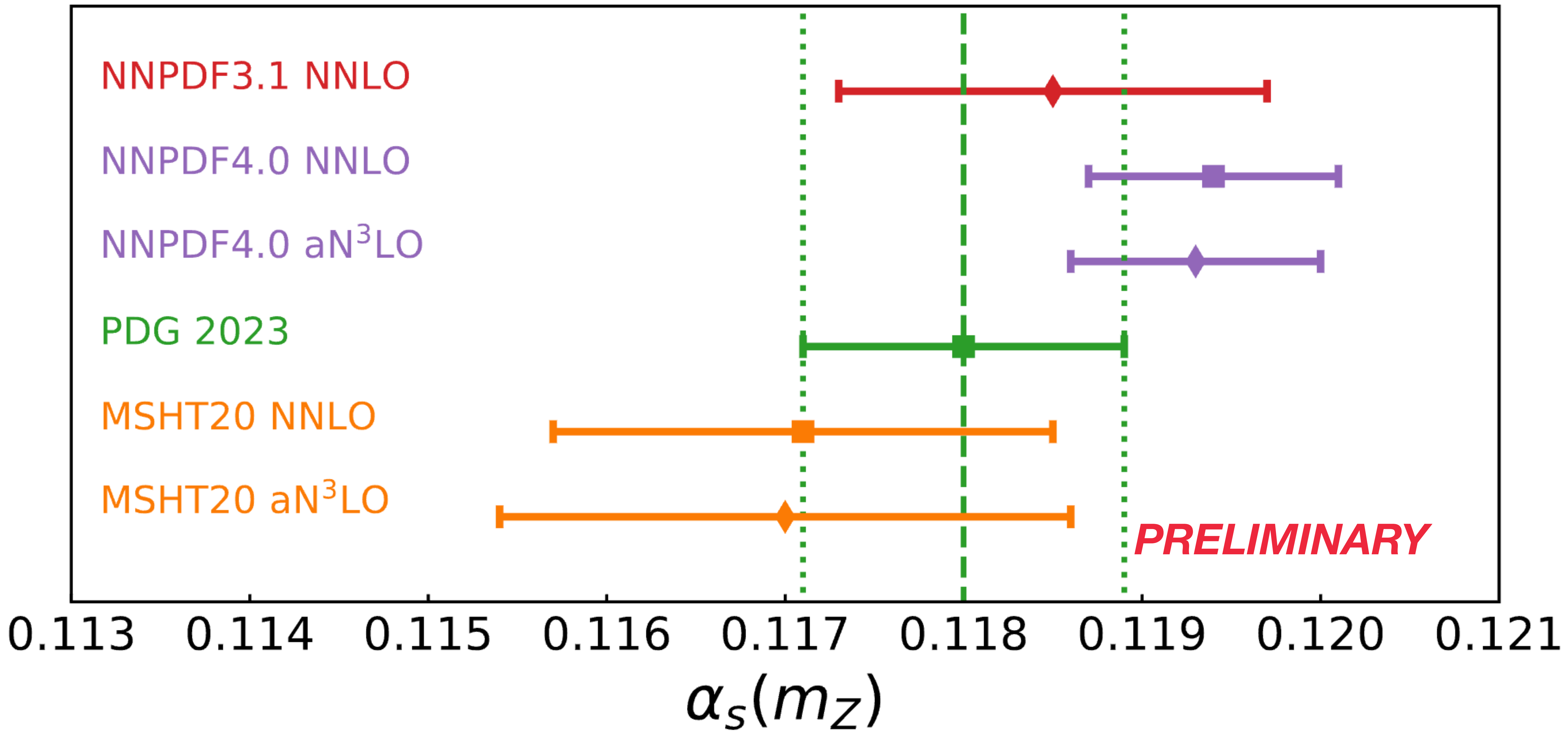
  NNLO $\alpha_s(M_Z) = 0.1194 \pm 0.0007$

  aN^{3}LO $\alpha_s(M_Z) = 0.1193 \pm 0.0007$



Summary and Outlook

- PDFs are a key ingredient for LHC physics
- aN3LO PDFs allow for a consistent computation of observables at N3LO. Initial results suggest good convergence for Higgs and Drell-Yan production
- SM parameters from collider data require a simultaneous determination with the PDFs
- **The extracted strong coupling constant is perturbatively stable between NNLO and aN3LO:** *PRELIMINARY*
  NNLO $\alpha_s(M_Z) = 0.1194 \pm 0.0007$
  aN^{3}LO $\alpha_s(M_Z) = 0.1193 \pm 0.0007$



Thank you for your attention!

Backup slides

$$\alpha_s(M_Z)$$

α_s from correlated theory uncertainties

The “correlated replicas” method is computationally costly and lacks a mathematical framework

Alternatively, α_s can be determined in a Bayesian framework from nuisance parameters: [arXiv:2105.05114]

1. Model the theory uncertainty as a shift correlated for all datapoints
   $T \to T + \Delta\alpha_s \cdot \beta$, for $\beta \equiv T(\alpha_s^+) - T(\alpha_s^-)$
   we can then write
   $$P(T \mid D, \Delta\alpha_s) \propto \exp\left(-\frac{1}{2}(T + \Delta\alpha_s \cdot \beta - D)^T \mathrm{Cov}_{EXP}^{-1}(T + \Delta\alpha_s \cdot \beta - D)\right)$$

2. Choose a prior
   $$P(\Delta\alpha_s) \propto \exp\left(-\frac{1}{2}\Delta\alpha_s^2\right)$$

3. Marginalize over $\Delta\alpha_s$ to get $P\left(T|D\right)$

4. Compute the posterior for $\Delta\alpha_s$ using the ingredients we just wrote down
   $$P(\Delta\alpha_s \mid T, D) = \frac{P(T \mid D, \Delta\alpha_s)P(\Delta\alpha_s)}{P(T \mid D)} \propto \exp\left[-\frac{1}{2}Z^{-1}(\Delta\alpha_s - \overline{\Delta\alpha_s})\right]$$

For predictions T computed using α_s^0, **the final value is**
$\alpha_s = \alpha_s^0 + \overline{\Delta\alpha_s} \pm Z$

The results agree with the correlated replicas method

N3LO

Construction of aN3LO splitting functions

Complete results for the N3LO splitting functions are not yet available, but some information exists:

- Mellin moments
- Small-x limit (threshold resummation)
- Large-x limit (BFKL resummation)
- Large-n_f limit

Idea: combine these limits to construct splitting functions at approximate N3LO

1. The approximation is constructed in Mellin space, independently for each order in n_f

$$\gamma_{ij}^{(3)} = \gamma_{ij,n_f}^{(3)} + \gamma_{ij,N\to\infty}^{(3)} + \gamma_{ij,N\to 0}^{(3)} + \gamma_{ij,N\to 1}^{(3)} + \widetilde{\gamma}_{ij}^{(3)}$$

2. The remainder term $\widetilde{\gamma}_{ij}^{(3)}$ is constructed as a linear combination of interpolating functions equal to the number of known Mellin moments
   - A function for the leading unknown large-N contribution
   - A function for each of the two leading unknown small-N contribution
   - 5 functions for the subleading small-N and large-N contributions
3. The weights of these interpolating functions are determined by equating to the known moments
4. Then, vary the subleading contributions included in the basis of interpolating functions to estimate incomplete higher order uncertainties (IHOU) on the splitting functions

Incomplete Higher-Order Uncertainties

- We construct an ensemble of $\widetilde{N}_{ij}$ different approximations to $\gamma_{ij}^{(3)}(N)$ as interpolation functions that satisfy the known limits
- We approximate its best estimate with the average

$$\gamma_{ij}^{(3)}(N) = \frac{1}{\widetilde{N}_{ij}} \sum_{k=1}^{\widetilde{N}_{ij}} \gamma_{ij}^{(3),\,(k)}(N)\,.$$

- We include the uncertainty on the average with the theory covariance matrix formalism (each instance $\gamma_{ij}^{(3),\,(k)}$ is seen as a nuisance parameter)

$$\Delta_m(ij,k) = T_m(ij,k) - \bar{T}_m \qquad \mathrm{cov}_{mn}^{(ij)} = \frac{1}{\widetilde{N}_{ij}-1} \sum_{k=1}^{\widetilde{N}_{ij}} \Delta_m(ij,k)\Delta_n(ij,k)\,.$$

- The total contribution to the theory covariance matrix is

$$\mathrm{cov}_{mn}^{\mathrm{IHOU}} = \mathrm{cov}_{mn}^{(gg)} + \mathrm{cov}_{mn}^{(gq)} + \mathrm{cov}_{mn}^{(qg)} + \mathrm{cov}_{mn}^{(qq)}$$

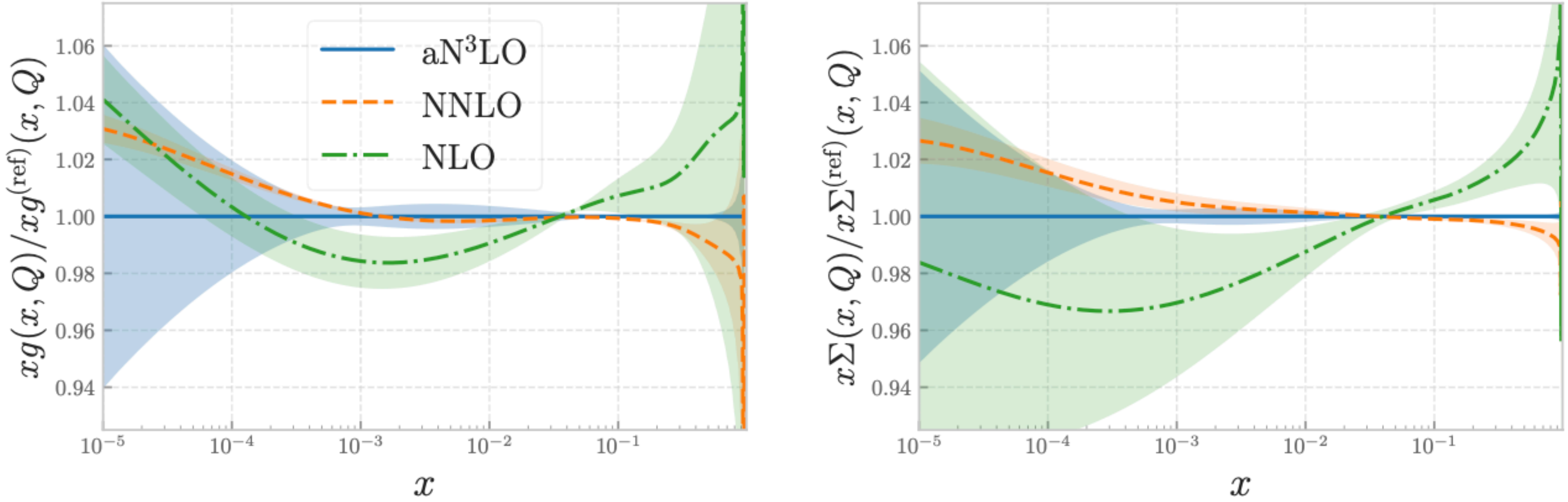
- The total theory uncertainty is the sum in quadrature of the IHOU and MHOU

$$\mathrm{cov}_{mn}^{\mathrm{tot}} = \mathrm{cov}_{mn}^{\mathrm{IHOU}} + \mathrm{cov}_{mn}^{\mathrm{MHOU}}$$

E. Nocera, Workshop on Hadron Physics and Opportunities Worldwide
Dalian, China, August 2024

aN3LO DGLAP evolution

NNPDF4.0 evolved from $Q = 1.65$ GeV to $Q = 100$ GeV



- Effects of N3LO corrections to DGLAP evolution at most percent level, except at small-x and large-x
- Good perturbative convergence

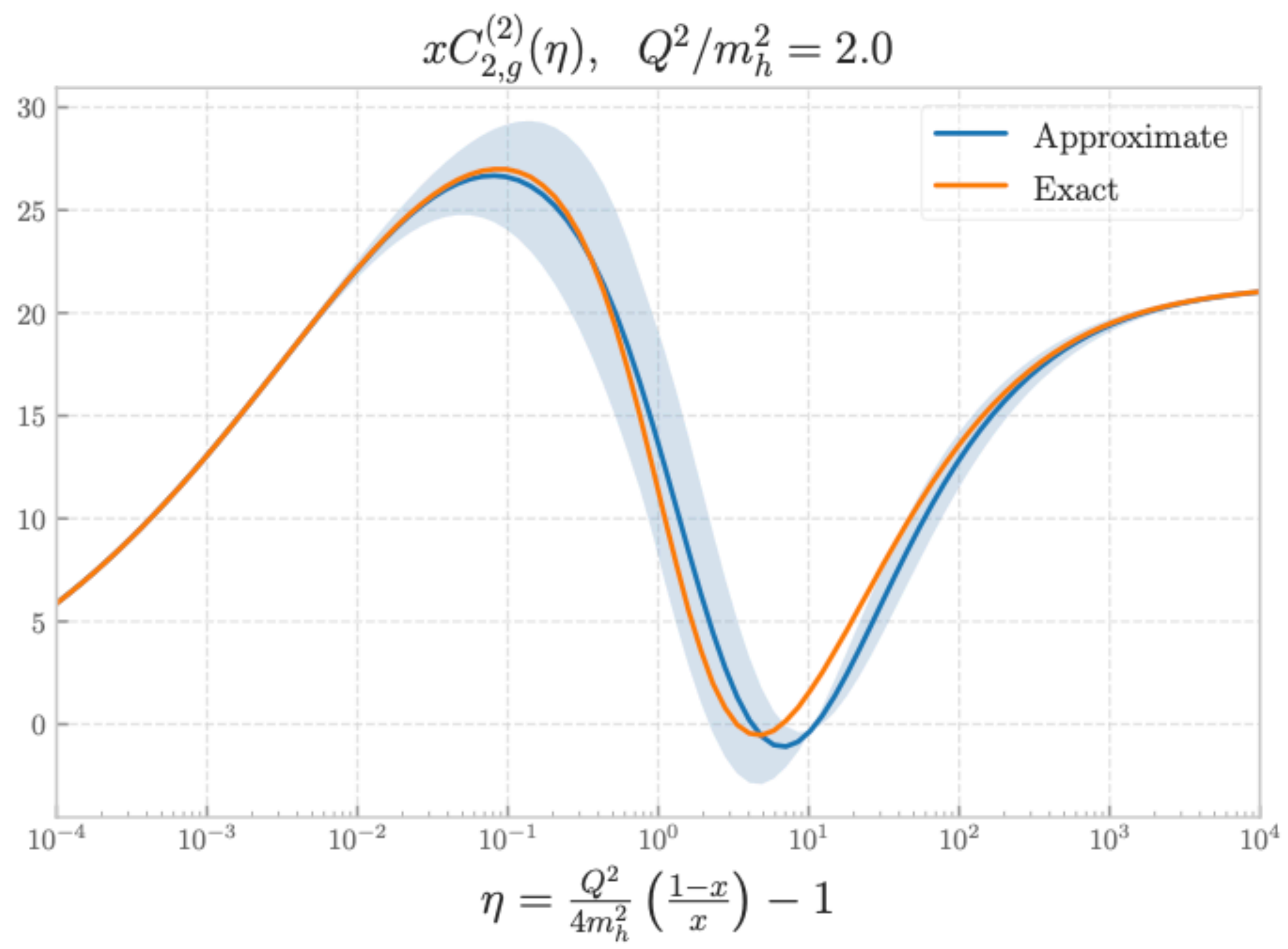
aN3LO DIS coefficients

- DIS coefficient functions are known up to N3LO in the massless limit [Larin, Nogueira, Van Ritbergen, Vermaseren, 9605317], [Moch Vermaseren Vogt, 0411112, 0504242], [Davies, Moch, Vermaseren, Vogt, 0812.4168, 1606.08907]
- Massive coefficient functions can be constructed by smoothly joining the known limits from threshold and high energy resummation, and the massless limit [Barontini, Bonvini, Laurenti, in preparation]

$$C^{(3)}(x, m_h^2/Q^2) = C^{(3),\mathrm{thr}}(x, m_h^2/Q^2) f_1(x) + C^{(3),\mathrm{asy}}(x, m_h^2/Q^2) f_2(x)$$

$$f_1(x) \xrightarrow[x\to x_{\max}]{} 1, \quad f_2(x) \xrightarrow[x\to x_{\max}]{} 0$$

$$f_1(x) \xrightarrow[x\to 0]{} 0, \quad f_2(x) \xrightarrow[x\to 0]{} 1$$

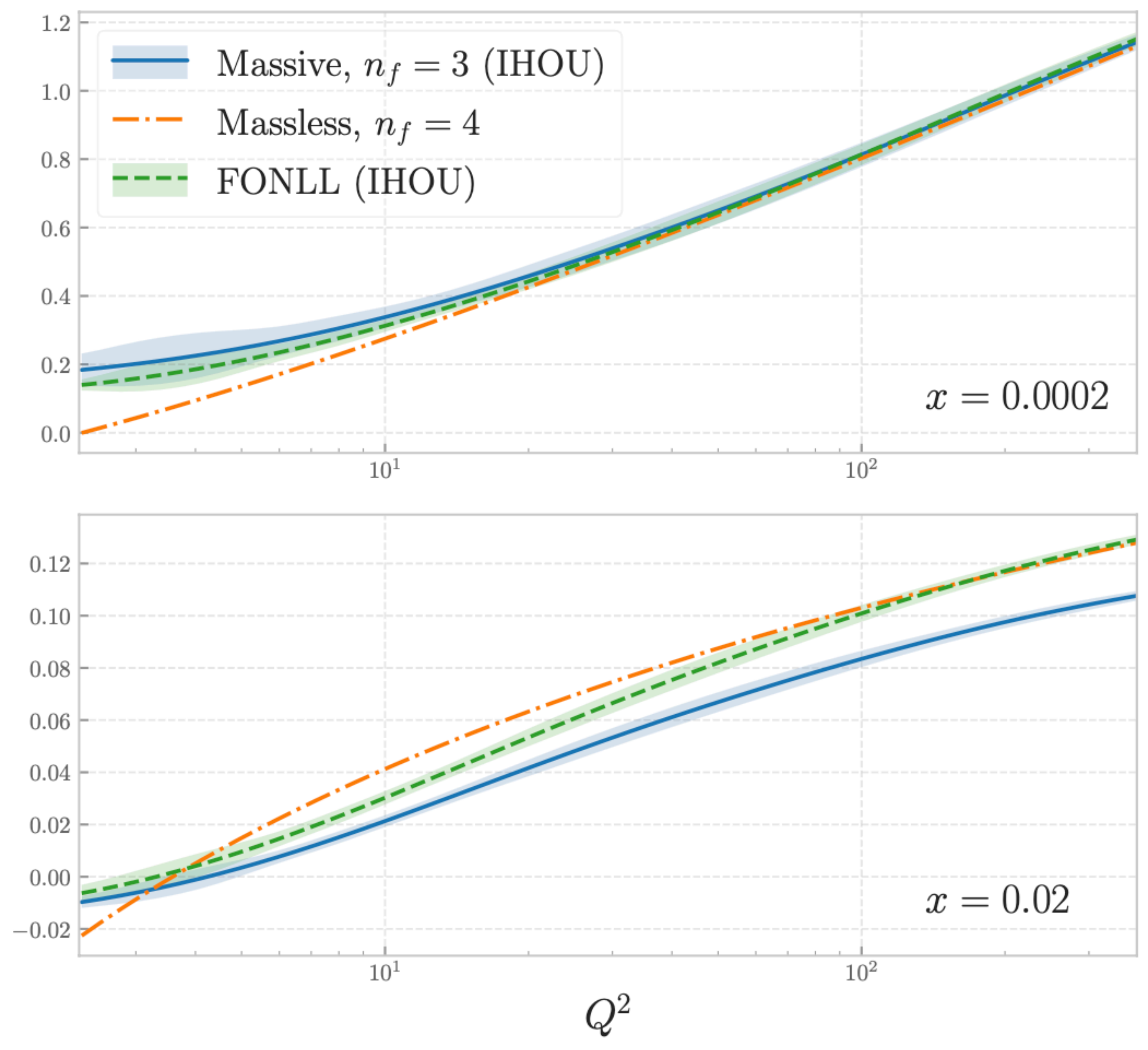


Validate the procedure at NNLO
Uncertainty band is obtained by varying interpolation functions

DIS general mass variable flavor number scheme (GM-VFNS)

- In a PDF fit different flavour number schemes are joined in a variable flavour number scheme (VFNS) to ensure reliable results from $Q^2 \sim m_h^2$ up to $Q^2 \gg m_h^2$
- The matching conditions encoding the transition between schemes have almost completely been computed up to N3LO
- The VFNS used in NNPDF is the FONLL scheme

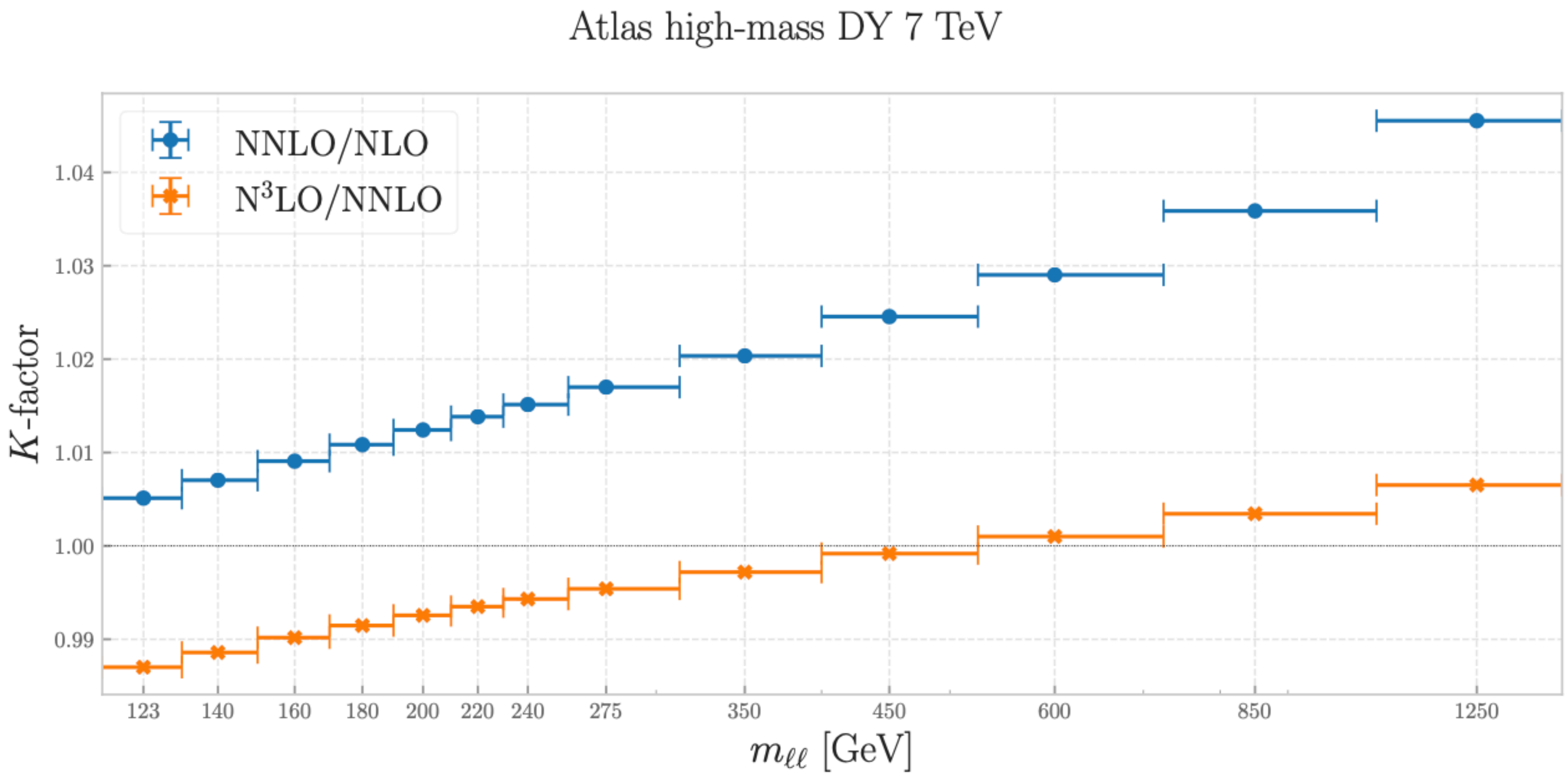
$$F_{\mathrm{FONLL}}(m_c) = F^{(n_f+1)}(m_h = 0) + F^{(n_f)}(m_c) - \lim_{m_c \to 0} F^{(n_f)}(m_h)$$



$F_2^{(c)}$

Hadronic processes

- Corrections to collider DY and W production can be included through k-factors
- N3LO effects around 1 to 2% for LHC observables
- For many processes N3LO corrections are not available, for those we introduce account for MHOU through μ_r variations



Comparison to MSHT20 aN3LO

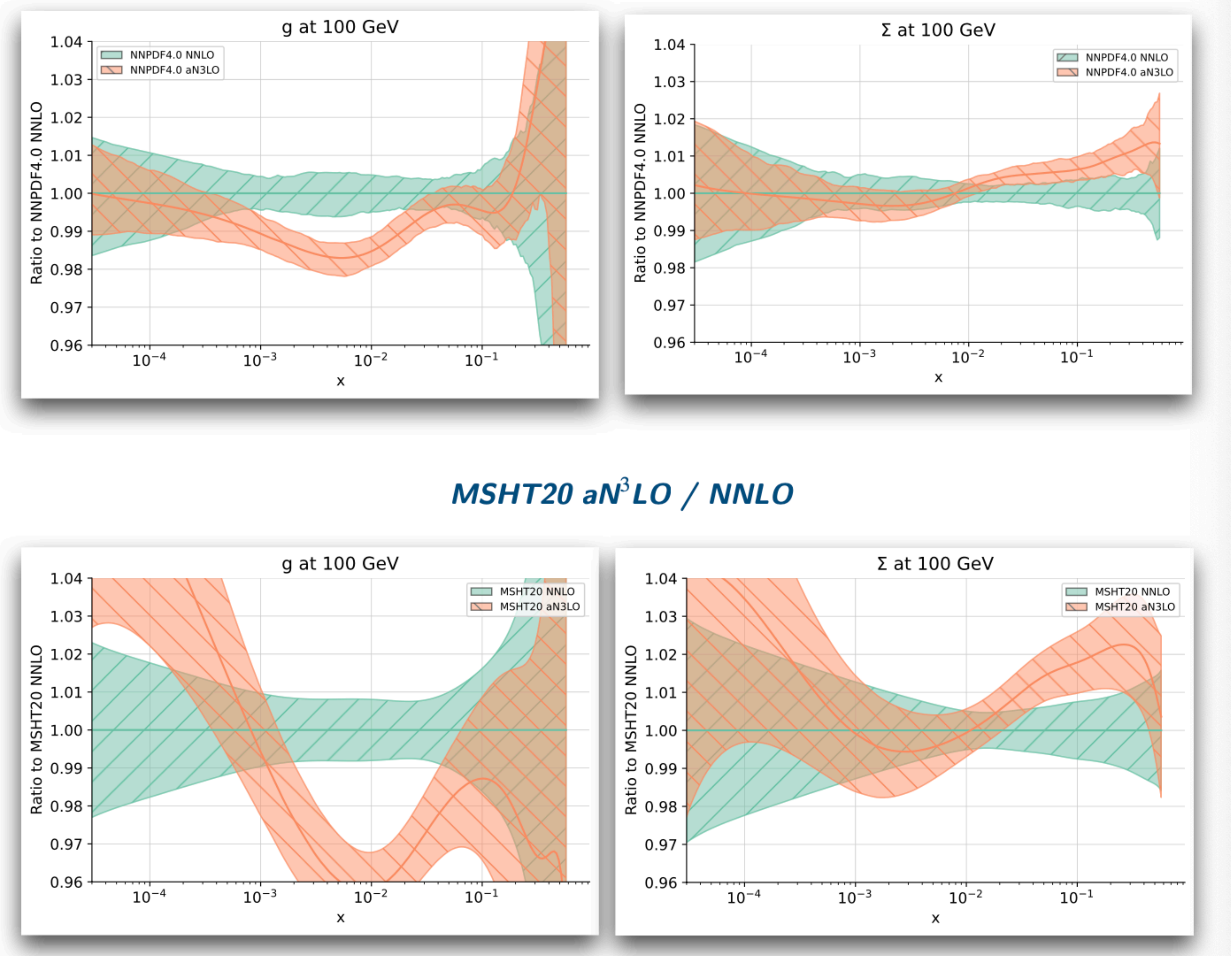
[arXiv:2207.04739]

Main differences are due to:

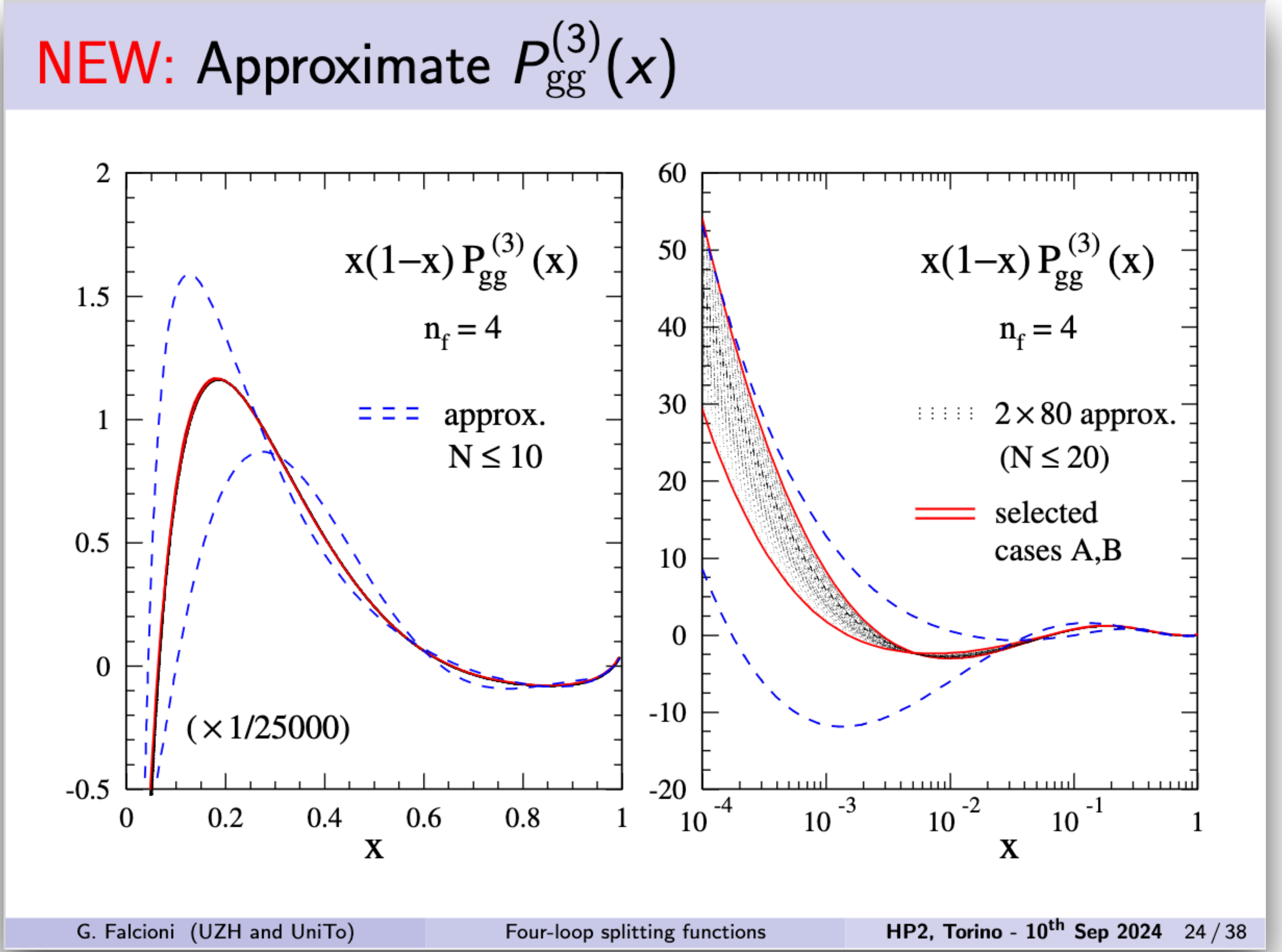
- **Mellin moments** for splitting functions computed in the last two years: MSHT has an earlier cut-off/publication date
- **DGLAP parameterization uncertainty**. NNPDF uses only prior while MSHT extracts posterior from data
- Treatment of **partonic coefficients**: DIS heavy quark schemes, hadronic k-factors
- **Fitting methodology and data**

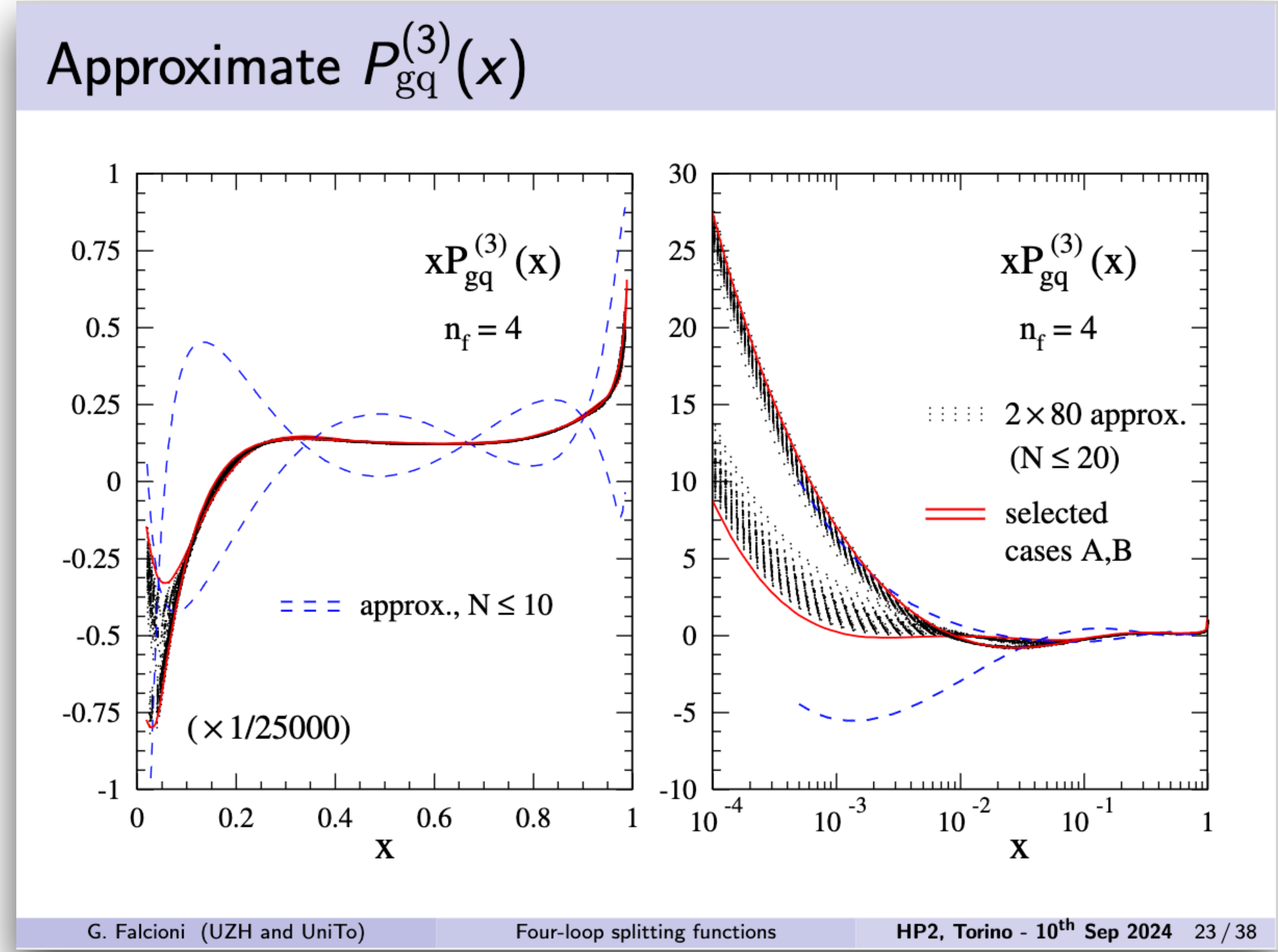
NNPDF4.0 aN^{3}LO / NNLO

MSHT20 aN^{3}LO / NNLO



G. Magni, HP2
Turin, September 2024





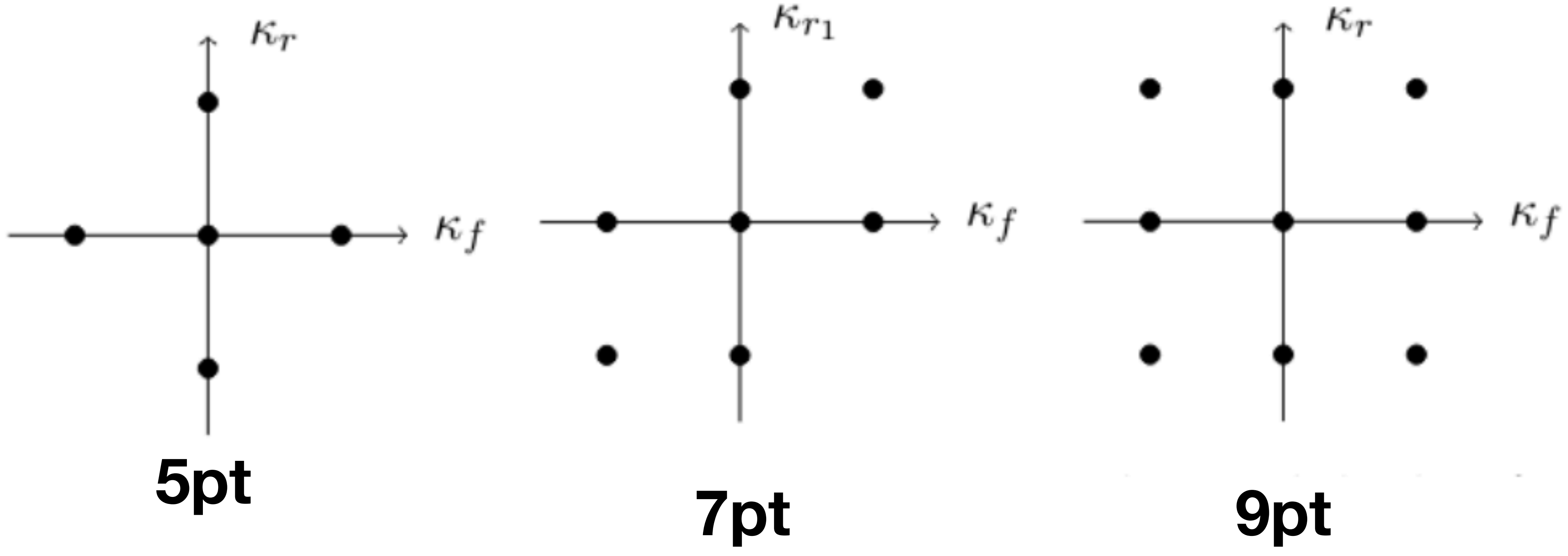
[arXiv:2404.09701]

G. Falcioni, HP2
Turin, Italy, September 2024

MHOU

Theory uncertainties in PDFs

Missing higher order uncertainties (MHOUs) are estimated through 7 point scale variations



- In a fit we minimize the χ^2:

$$P(T \mid D\lambda) \propto \exp\left(-\frac{1}{2}(T-D)^T C^{-1}(T-D)\right) \equiv \exp\left(\chi^2\right)$$

- To account for MHOUs we treat the theory covmat on the same footing as the experimental covmat: $C = C_{\text{exp}} + C_{\text{MHOU}}$

$$C_{\text{MHOU},ij} = n_m \frac{1}{V_m} \sum \left(T_i(\kappa_f, \kappa_r) - T_i(0,0)\right)\left(T_j(\kappa_f, \kappa_r) - T_j(0,0)\right)$$

Validating the MHOU covmat

The MHOU covmat is validated by comparing the **shifts from scale variations at NLO** to the known **NNLO-NLO shifts**