MARIA UBIALI
UNIVERSITY OF CAMBRIDGE

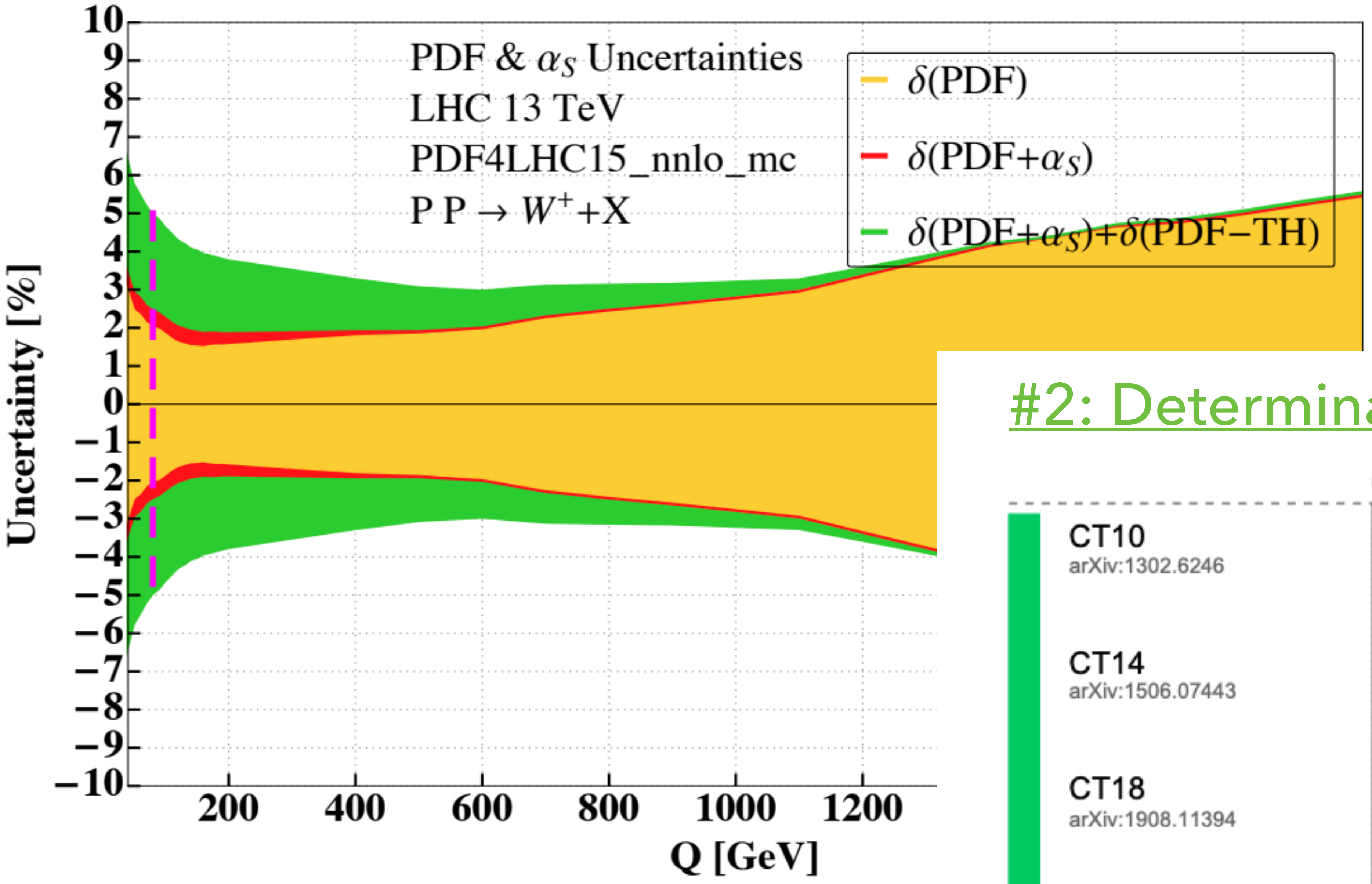
PDFS FOR PRECISION MEASUREMENTS AT THE LHC

LHCP2024, NORTHEASTERN UNIVERSITY, BOSTON, USA

4TH JUNE 2024

PDFS ARE KEY TO PRECISION PHYSICS

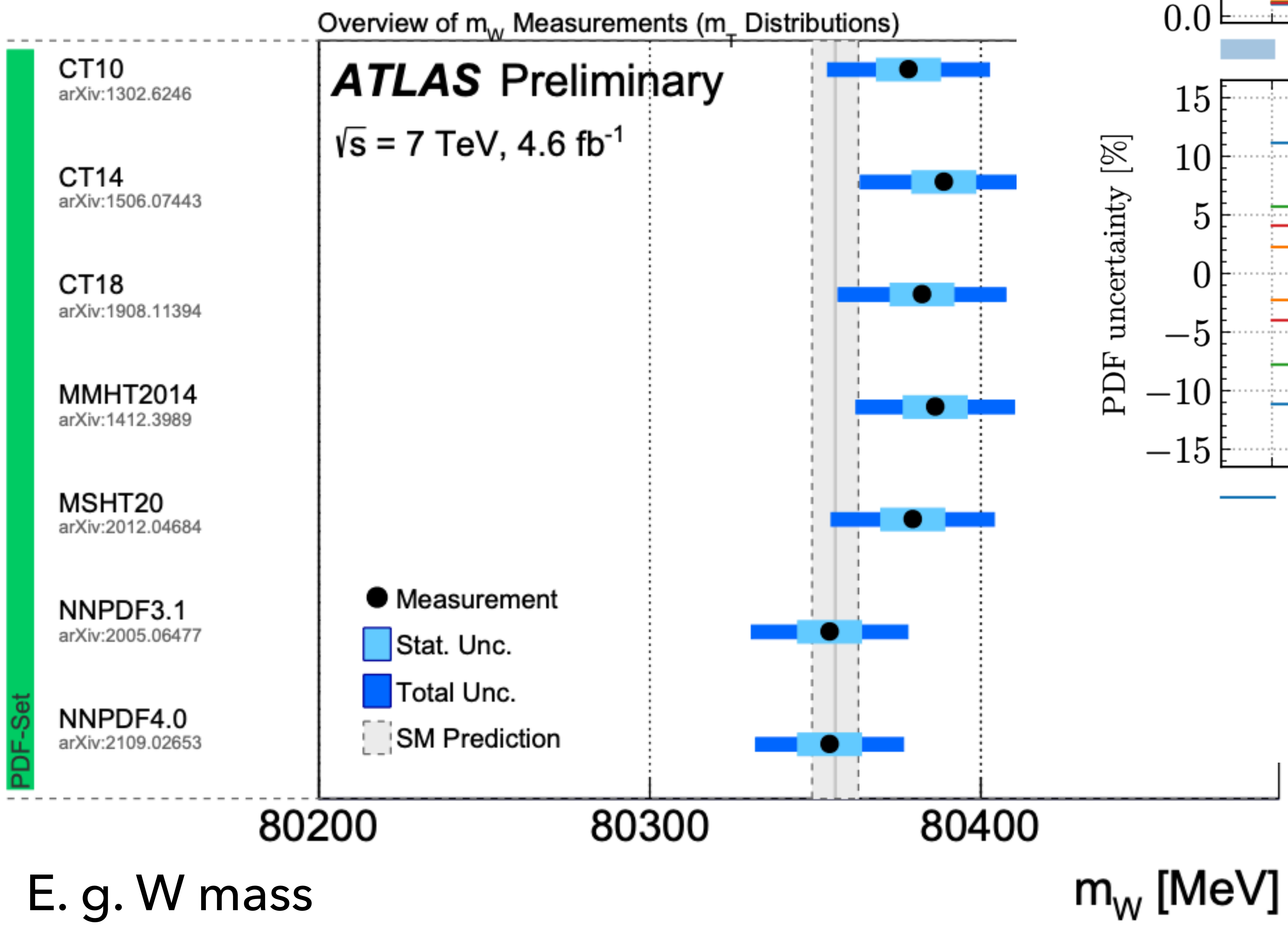
#1: Theory uncertainty of SM predictions



E. g. Drell-Yan @ N3LO

C. Duhr et al, *JHEP* 11 (2020) 143

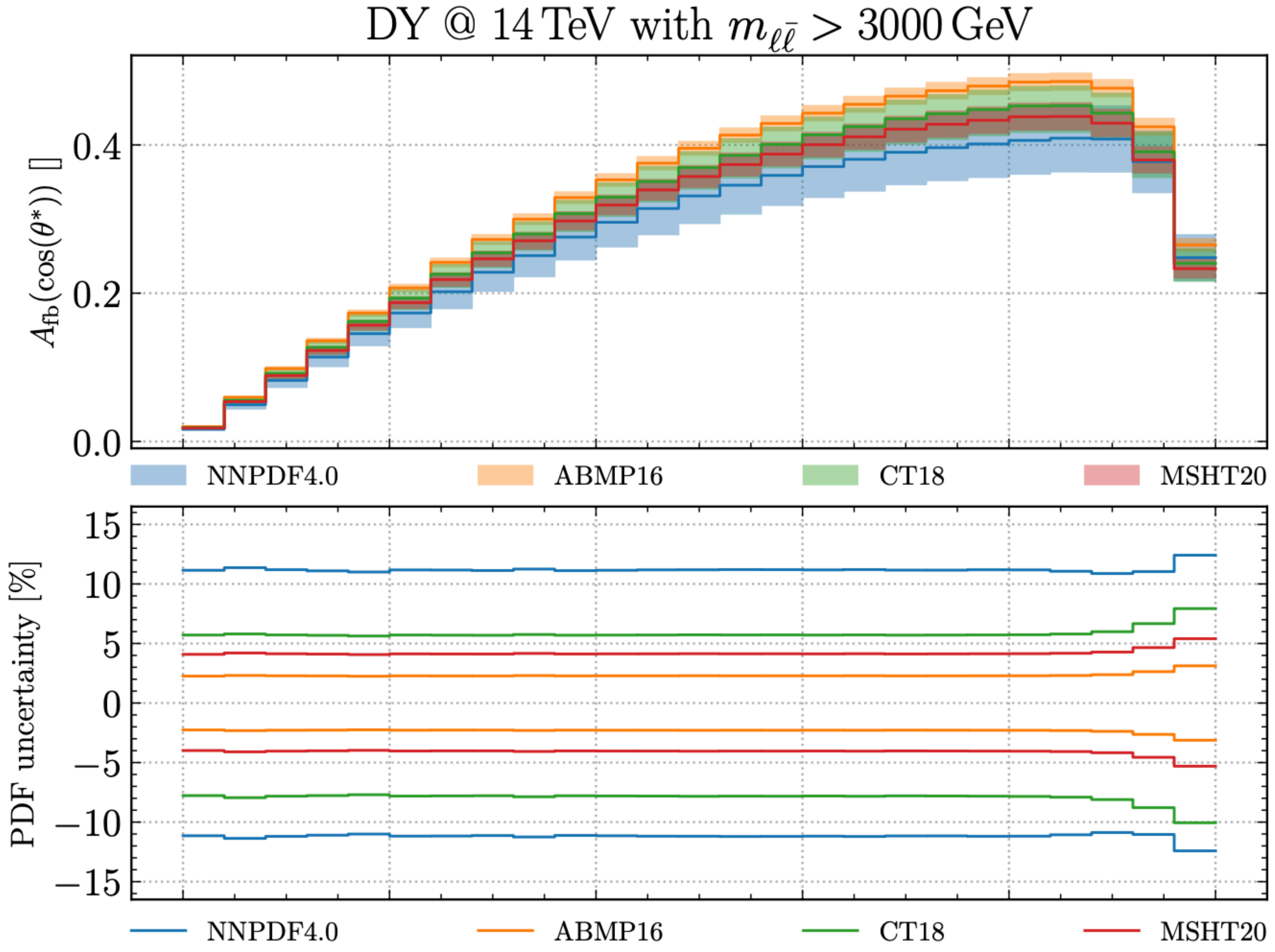
#2: Determination of SM parameters



E. g. W mass

ATLAS-CONF-2023-004

#3: NP searches



E. g. Heavy Z′ search in DY Forward-Backward asymmetry

Ball et al arXiv: 2209.08115

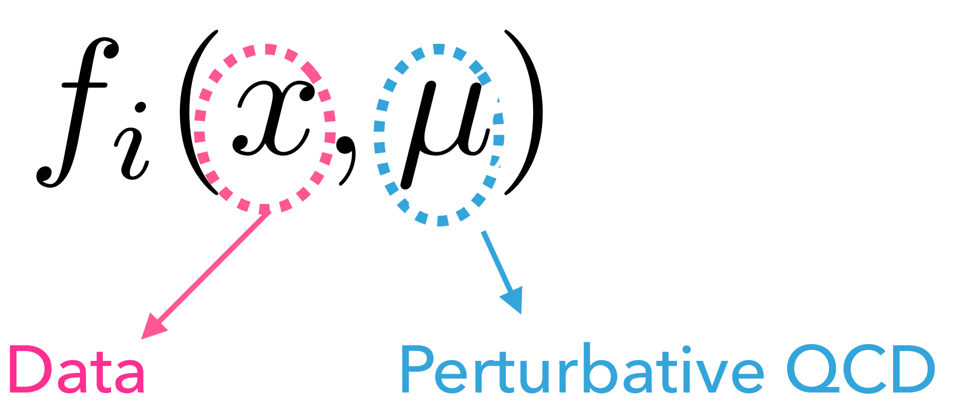
OUTLINE

➡ Part I: a short overview of recent developments in the determination of PDFs

➡ Part II: The precision vs accuracy challenge: a road-map to robustly test accuracy

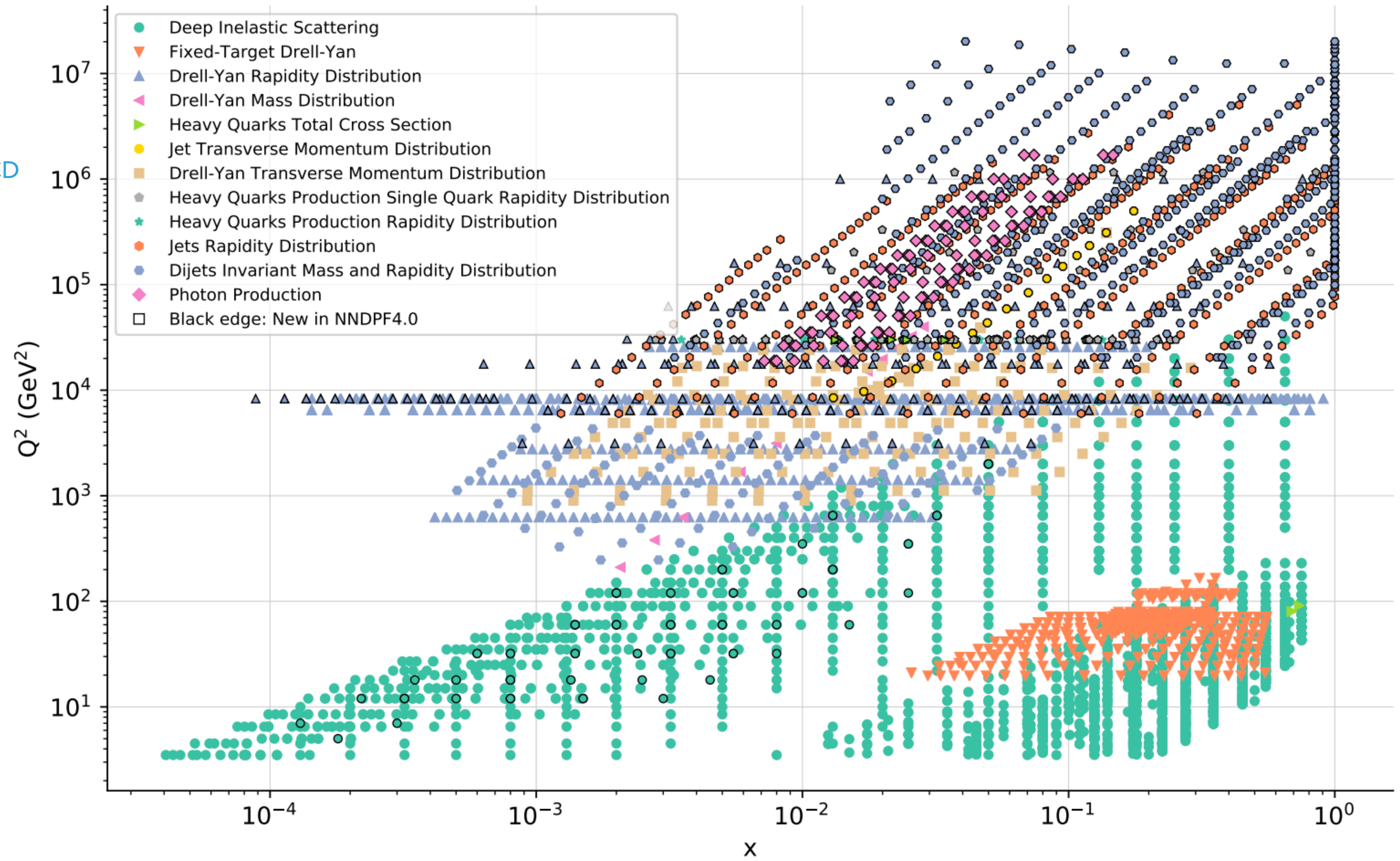
PART I: PROTON STRUCTURE AND PRECISION PHYSICS

PDF DETERMINATION



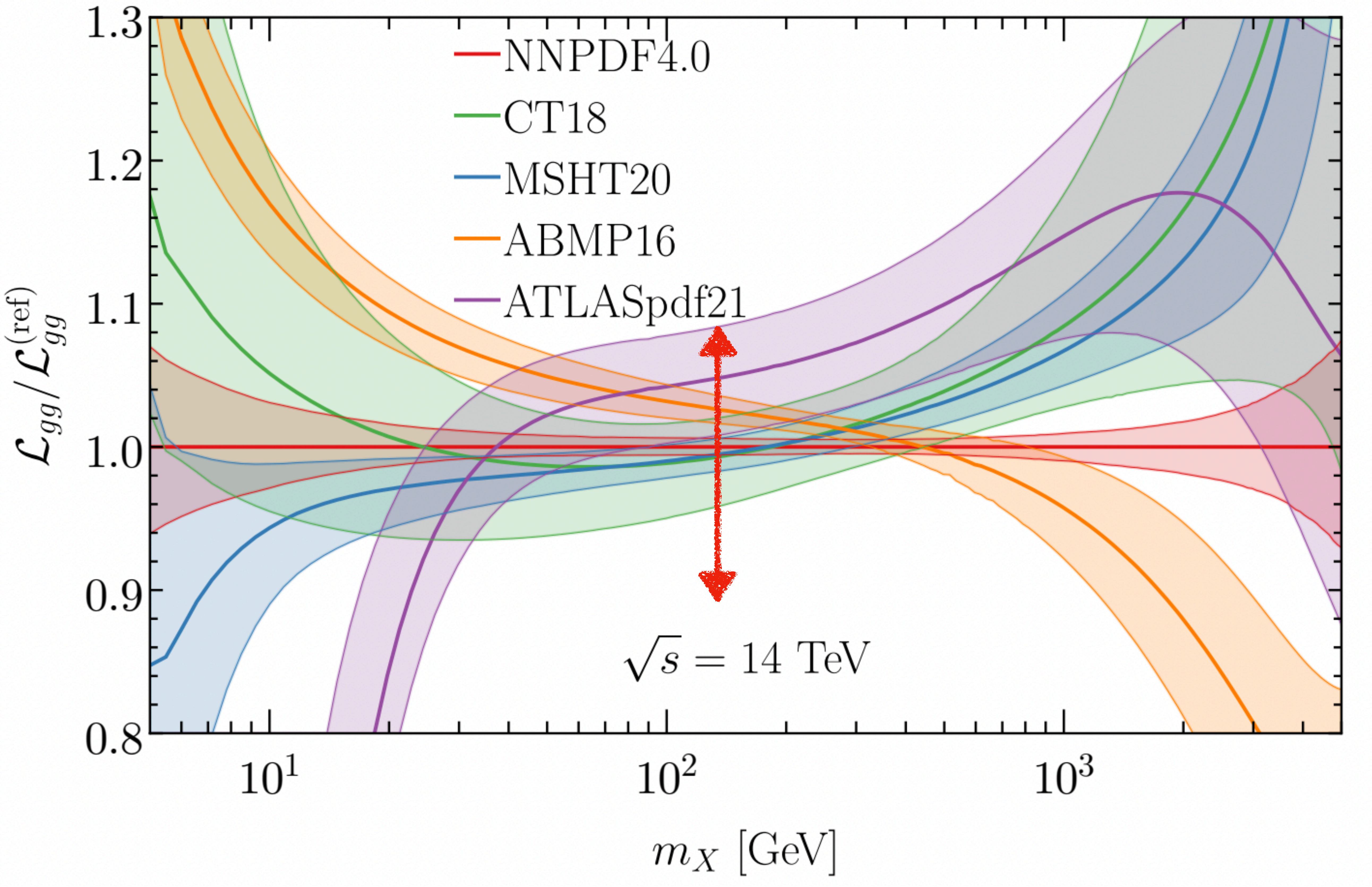
- DGLAP evolution equations fully known up to NNLO and partially at N³LO
- Most theory predictions for processes in PDF fits known at NNLO (some at N³LO)
- Precision of most data of order few %
- Mostly from correlated systematic uncertainties

Ball et al, NNPDF4.0, arXiv:2109.02653

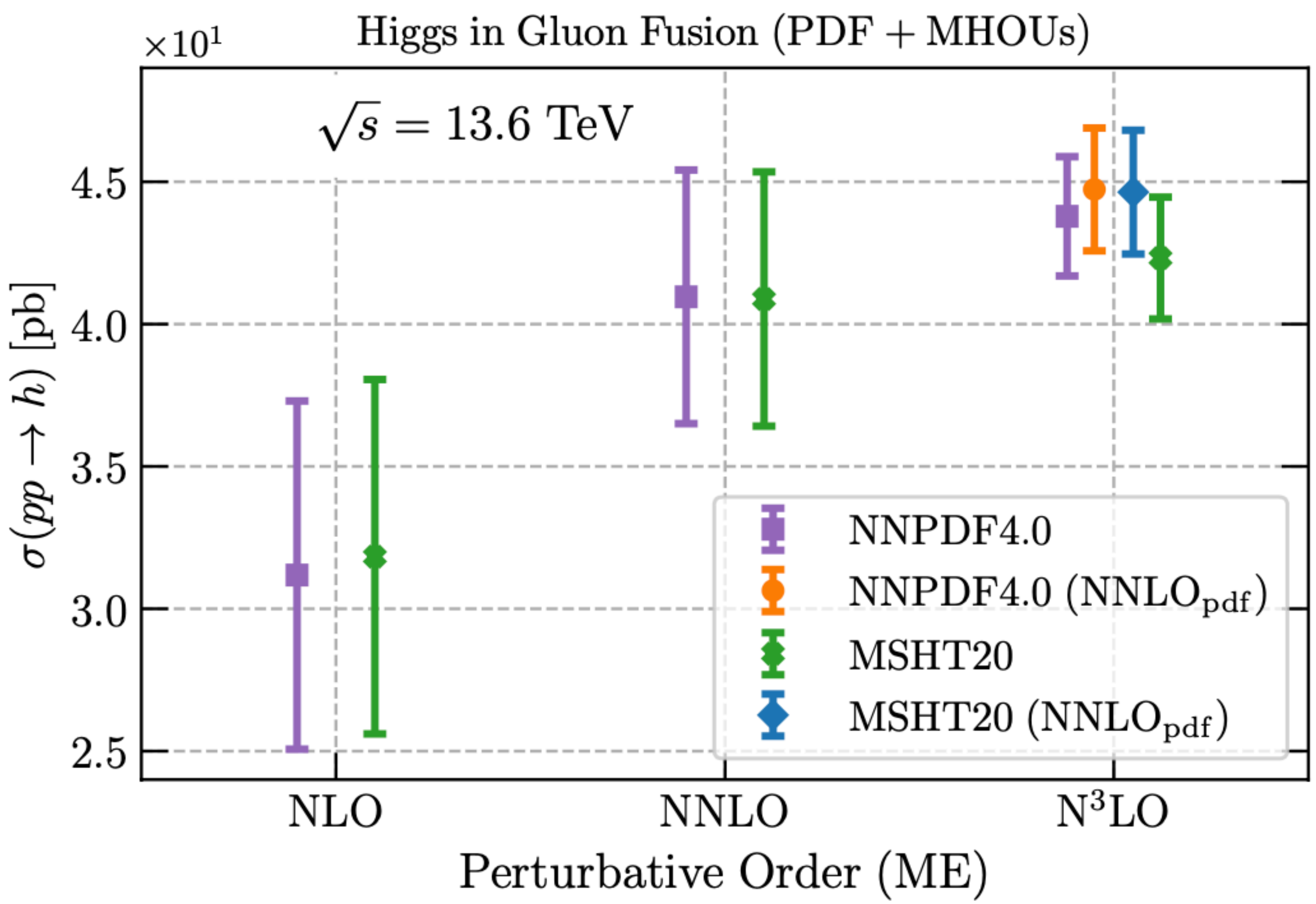


WHERE WE STAND NOW

Snowmass 2021 white paper [arXiv:2203.13923]



Ball et al 2402.18635

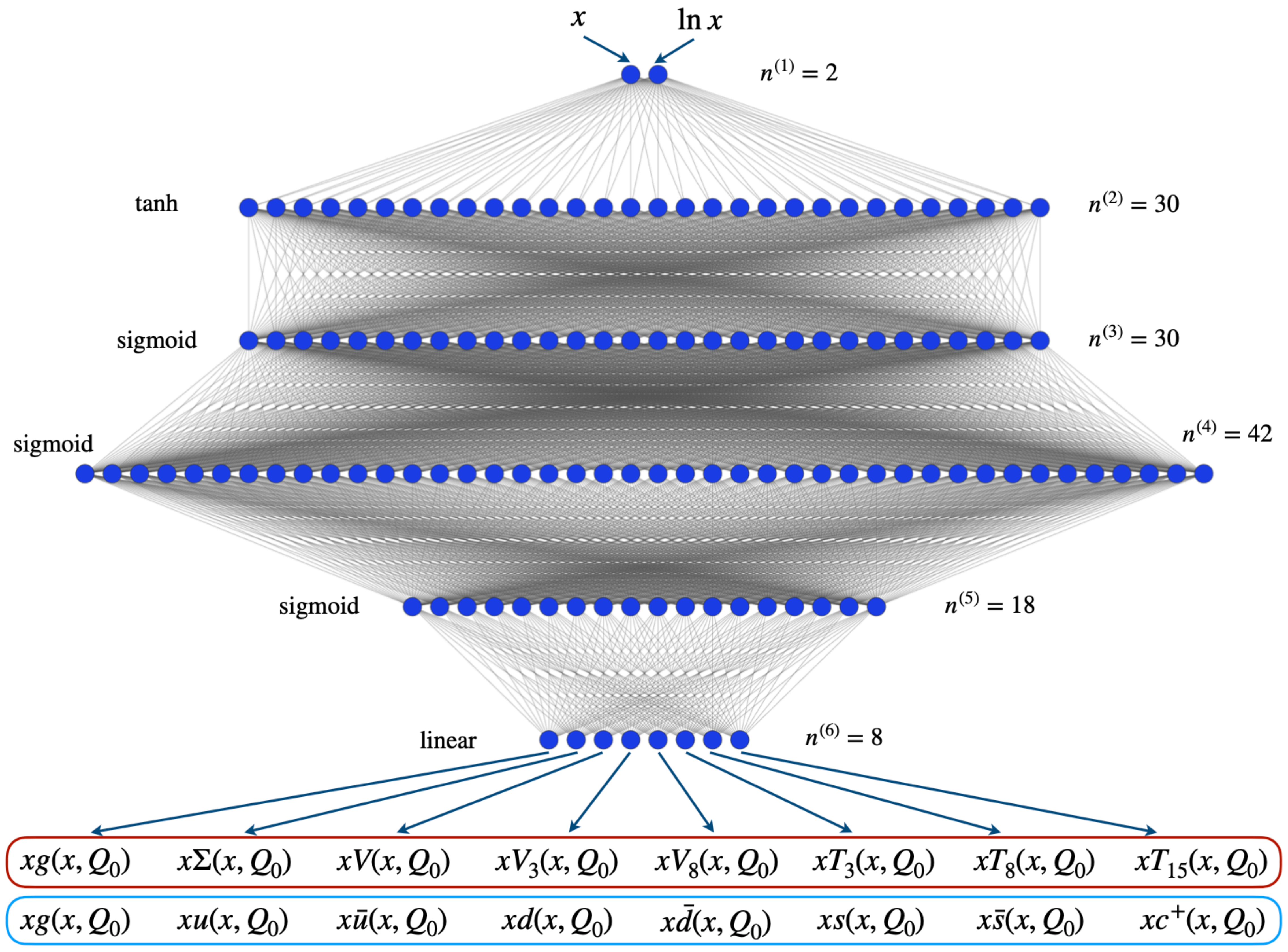


- NNLO sets: overall agreement around Higgs mass between global PDF sets with some shifts and large differences in PDF uncertainties due to differences in methodologies and datasets included.
- Approx N3LO sets: MSHT20aN3LO [arXiv:2207.04739] and NNPDF40aN3LO [arXiv:2402.18635] have been released

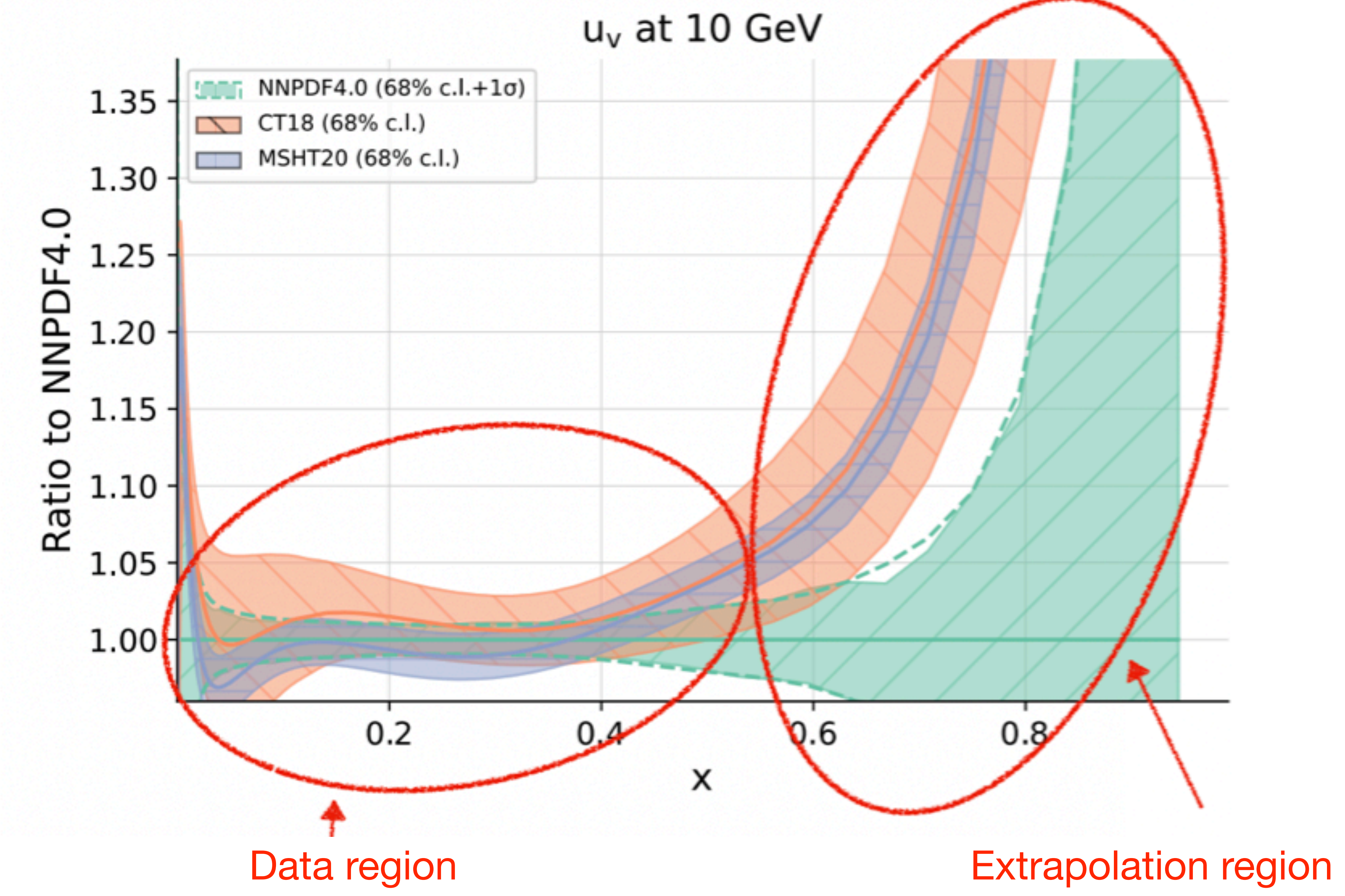
NNPDF4.0: THE ROLE OF DATA AND METHODOLOGY

- NNPDF4.0 data: large portion of LHC data included, total of O(4500) points (shift in PDFs and some uncertainty reduction)
- NNPDF4.0 methodology: single neural network to parametrise 8 independent PDF combinations (g, u, d, s, u~, d~, s~, c=c~) & optimisation strategy based on gradient descent rather than genetic algorithm (NNPDF3.1)
- Hyper-optimised methodology: scan of the hyper parameter space to find optimal minimisation settings (optimiser, initialiser, stopping patience, number of layers, learning rate, epochs, activation function) by minimising χ^2_{val}

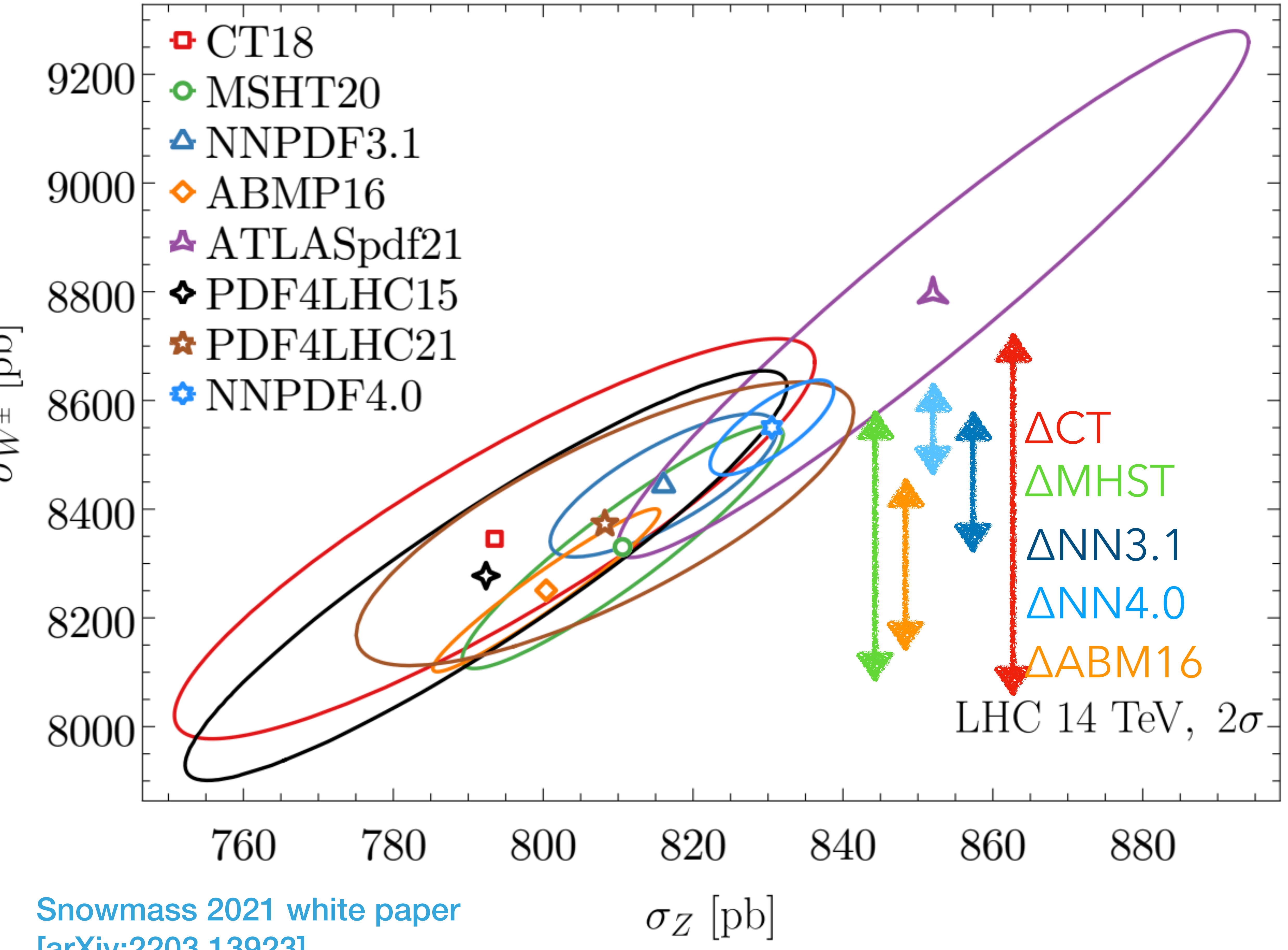
[**Carrazza et al, arXiv:1907.05075**]



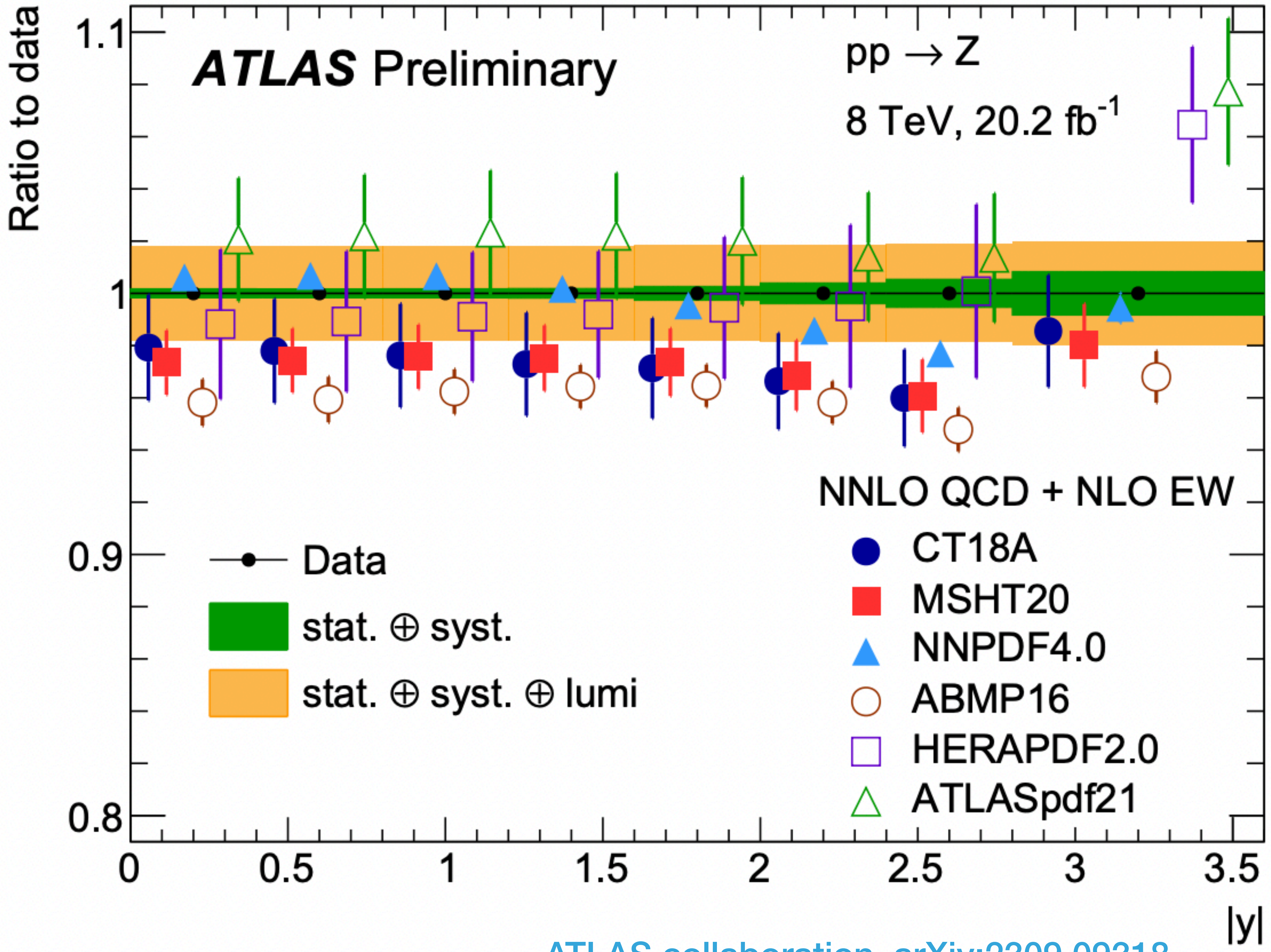
Thanks to superior methodology compared to NN31, NN40 displays smaller uncertainty than NN3.1, O(500) extra LHC data decrease uncertainties and shift PDFs.



PHENOMENOLOGICAL IMPLICATIONS



Snowmass 2021 white paper [arXiv:2203.13923]

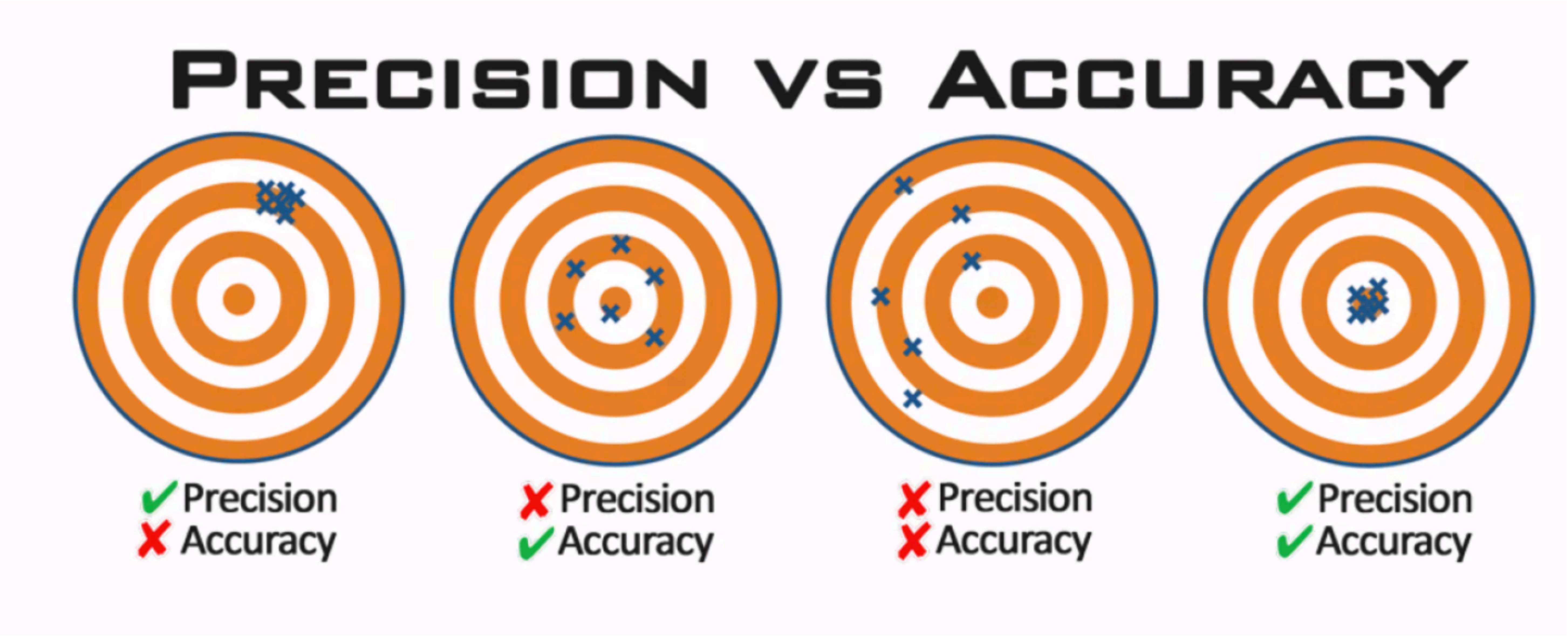


ATLAS collaboration, arXiv:2309.09318

- NNPDF4.0 has the smallest uncertainty in the data region, towards 1% precision.
- Extremely precise predictions and small PDF uncertainties for electroweak observables: implications for extraction of SM precision parameters, like α_s from Z p_T distribution and M_W from W m_T distributions [G. Bozzi's talk]
- Important to understand reason behind differences between different PDF sets & stress-test methodologies

PART II: A ROADMAP TO TEST ACCURACY

THE PRECISION VERSUS ACCURACY CHALLENGE



PDF4LHC21, 2203.05506

Dataset	N_{pt}	χ^2/N_{pt}		
		CT18	MSHT20	NNPDF3.1
BCDMS F_2^p	$329/163^{\dagger\dagger}/325^{\dagger}$	1.06	1.00	1.21
BCDMS F_2^d	$246/151^{\dagger\dagger}/244^{\dagger}$	1.06	0.88	1.10
NMC F_2^d/F_2^p	$118/117^{\dagger}$	0.93	0.93	0.90
NuTeV dimuon $\nu+\bar{\nu}$	38+33	0.79	0.83	1.22
HERAI+II	1120	1.23	1.20	1.22
E866 $\sigma_{pd}/(2\sigma_{pp})$	15	1.24	0.80	0.43
LHCb 7 TeV & 8TeV W,Z	29+30	1.15	1.17	1.44
LHCb 8 TeV $Z\to ee$	17	1.35	1.43	1.57
ATLAS 7 TeV W,Z (2016)	34	1.96	1.79	2.33
D0 Z rapidity	28	0.56	0.58	0.62
CMS 7 TeV electron A_{ch}	11	1.47	1.52	0.76
ATLAS 7 TeV W,Z(2011)	30	1.03	0.93	1.01
CMS 8TeV incl. jet	$185/174^{\dagger\dagger}$	1.03	1.39	1.30
Total N_{pt}	—	2263	1991	2256
Total χ^2/N_{pt}	—	1.14	1.15	1.20

Challenges

- Inconsistency or tension in data of experimental origin (underestimate of systematics…)
- Deficiencies in fitting methodology (data-driven parametrisation change, optimisation issues, overfitting…)
- Inaccuracy in theoretical framework
  - ➡ Missing higher order uncertainties (QCD, EW)
  - ➡ Other corrections (nuclear, higher-twist, non-perturbative effects…)
- Fitting away possible BSM signals

CLOSURE TEST: A TOOL TO TEST METHODOLOGY AND THEORY

- Closure tests for **data region**: imagine we knew the law of Nature f: is our fitting methodology able to reproduce it? Is the uncertainty faithful? Statistical validation of PDF uncertainties can be performed via closure tests.

L0 pseudo-data

$$y_0 = f + \eta$$

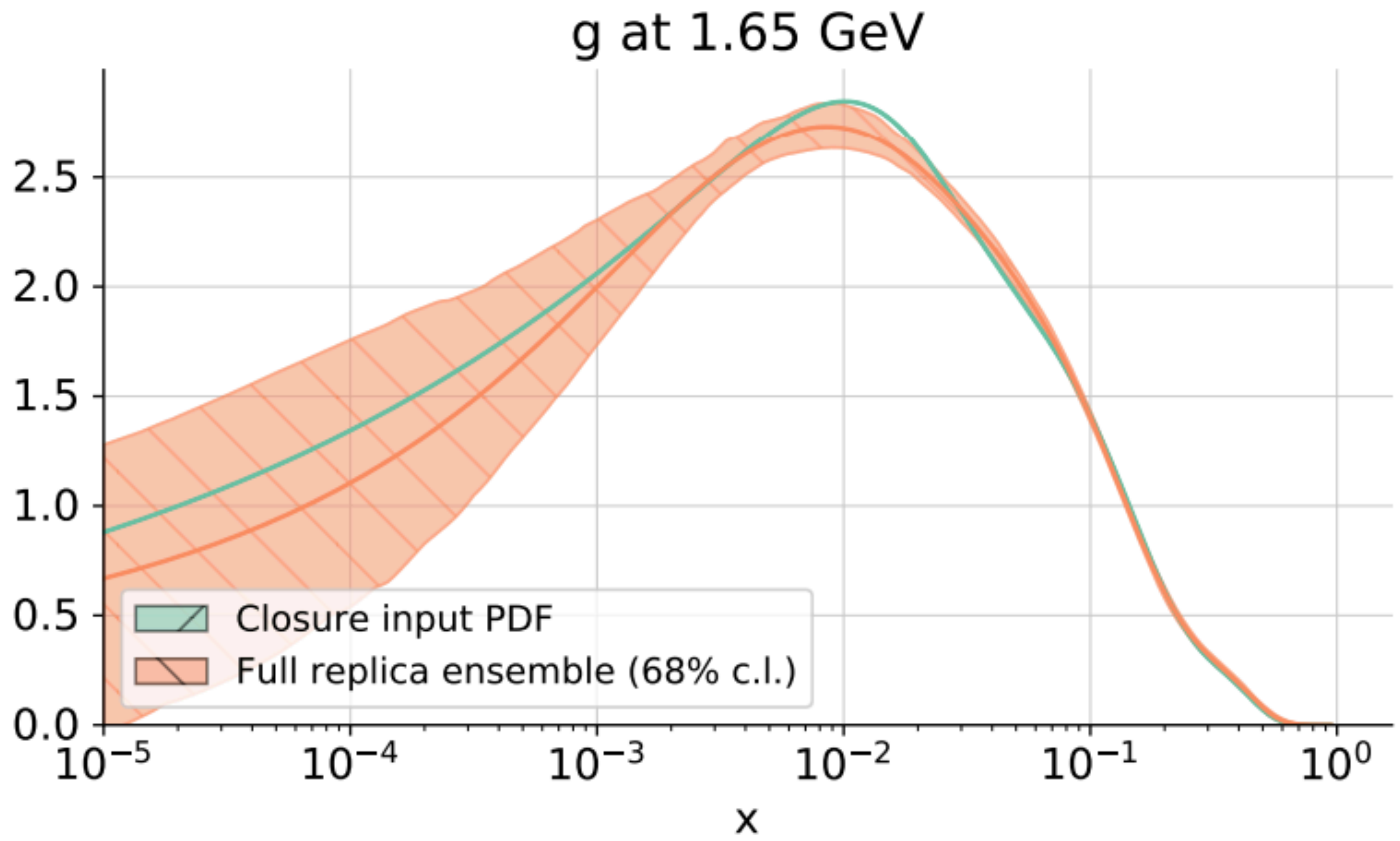
Experimental noise

Law of Nature

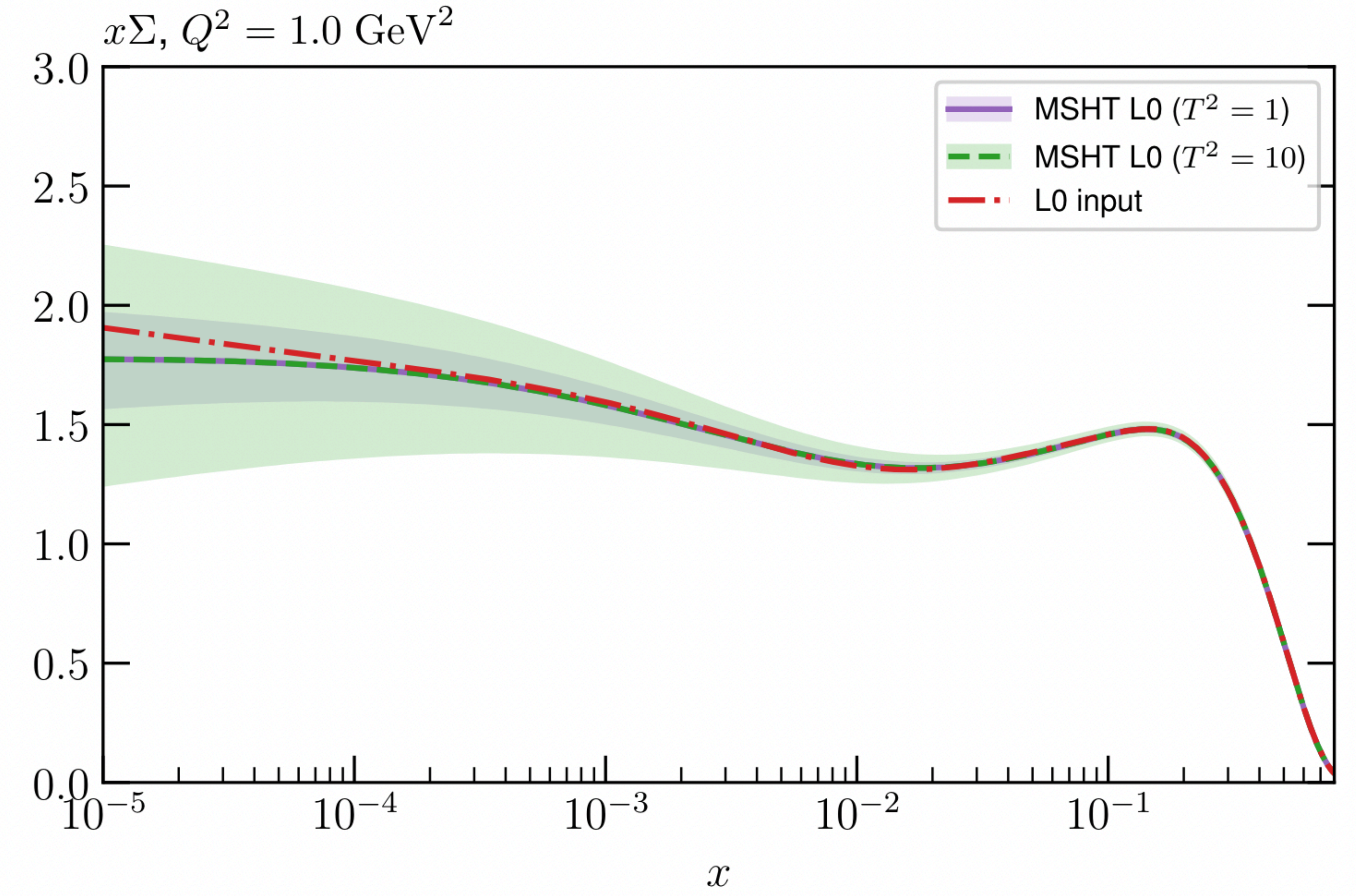
$$\hat{\sigma}_{\text{NNLO}} \otimes (f \otimes f)_{\text{"true" PDF}}$$

Test fitting methodology with consistent data (by construction)

Del Debbio et al, [arXiv: 2111.05787]



L. Harland Lang, DIS 2024



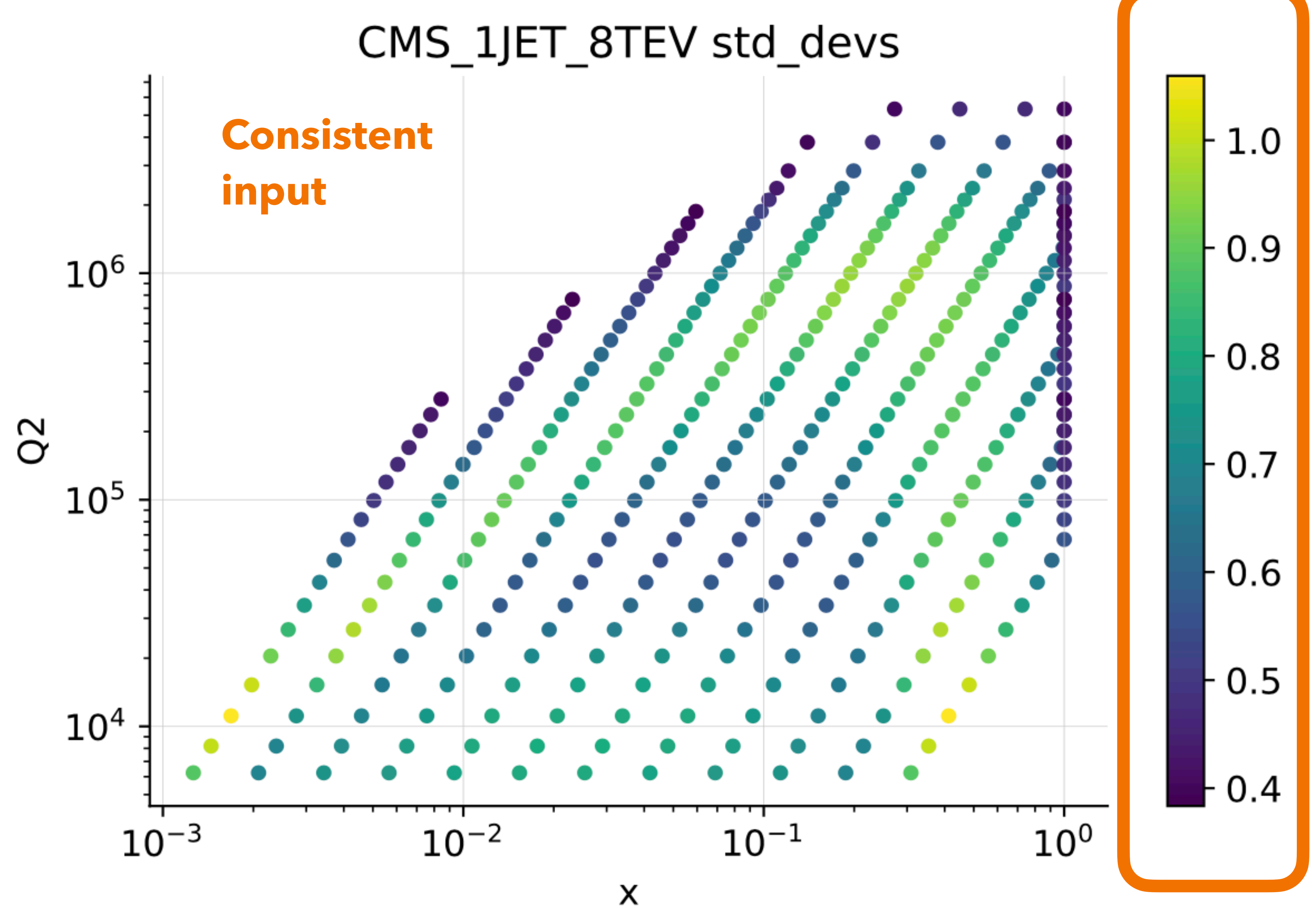
CLOSURE TEST: A TOOL TO TEST METHODOLOGY AND THEORY

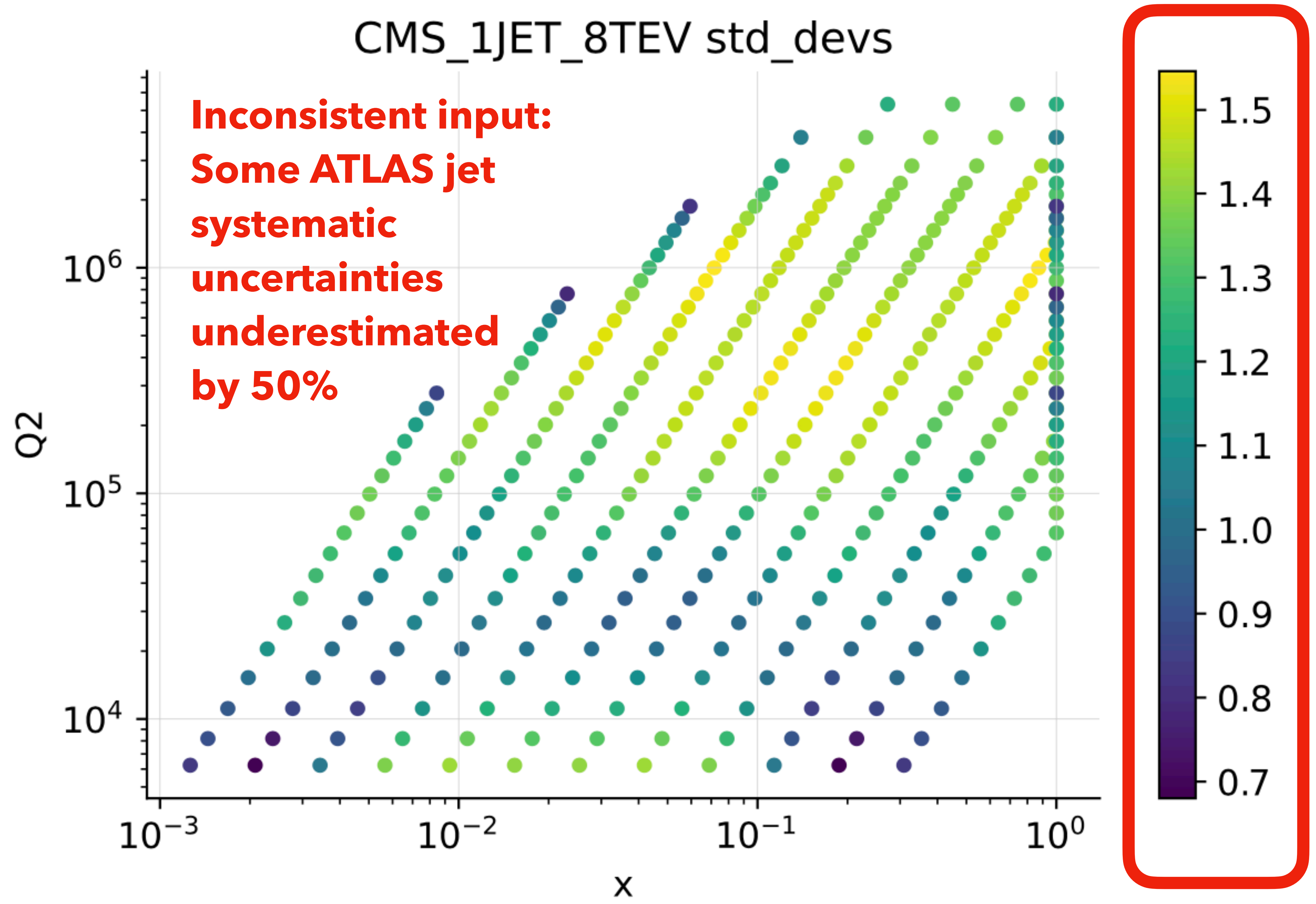
- Closure tests for data region: imagine we knew the law of Nature **f**: is our fitting methodology able to reproduce it? Is the uncertainty faithful? Statistical validation of PDF uncertainties can be performed via closure tests.
- What happens if experimentalists underestimated some systematics?

Statistical indicators of under (>1) or over (<1) estimated uncertainties allow to check effect of possible experimental inconsistencies on datasets included and on the PDFs

Test fitting methodology with inconsistent data

[Barontini, Costantini, De Crescenzo, Ubiali - in progress]



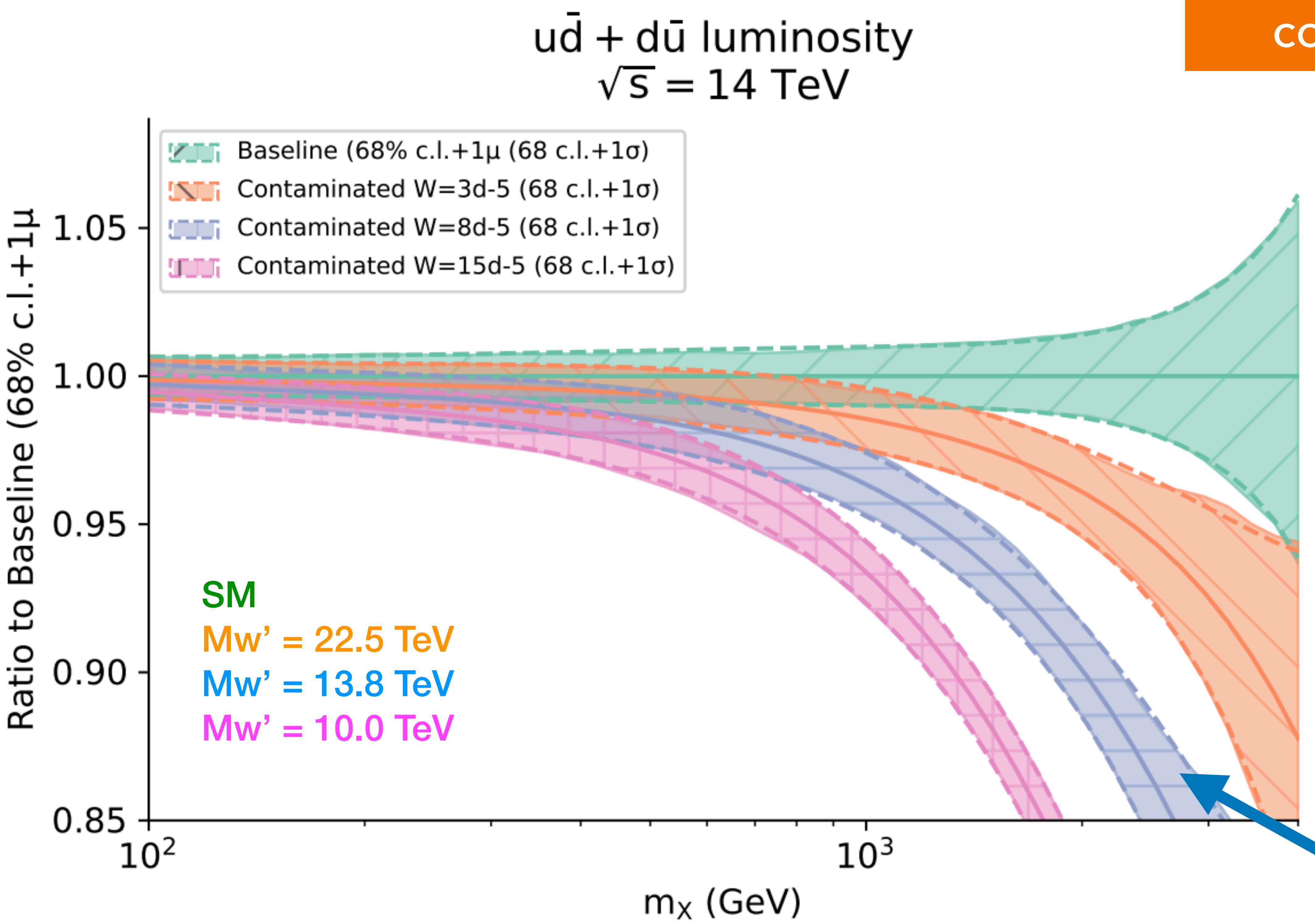


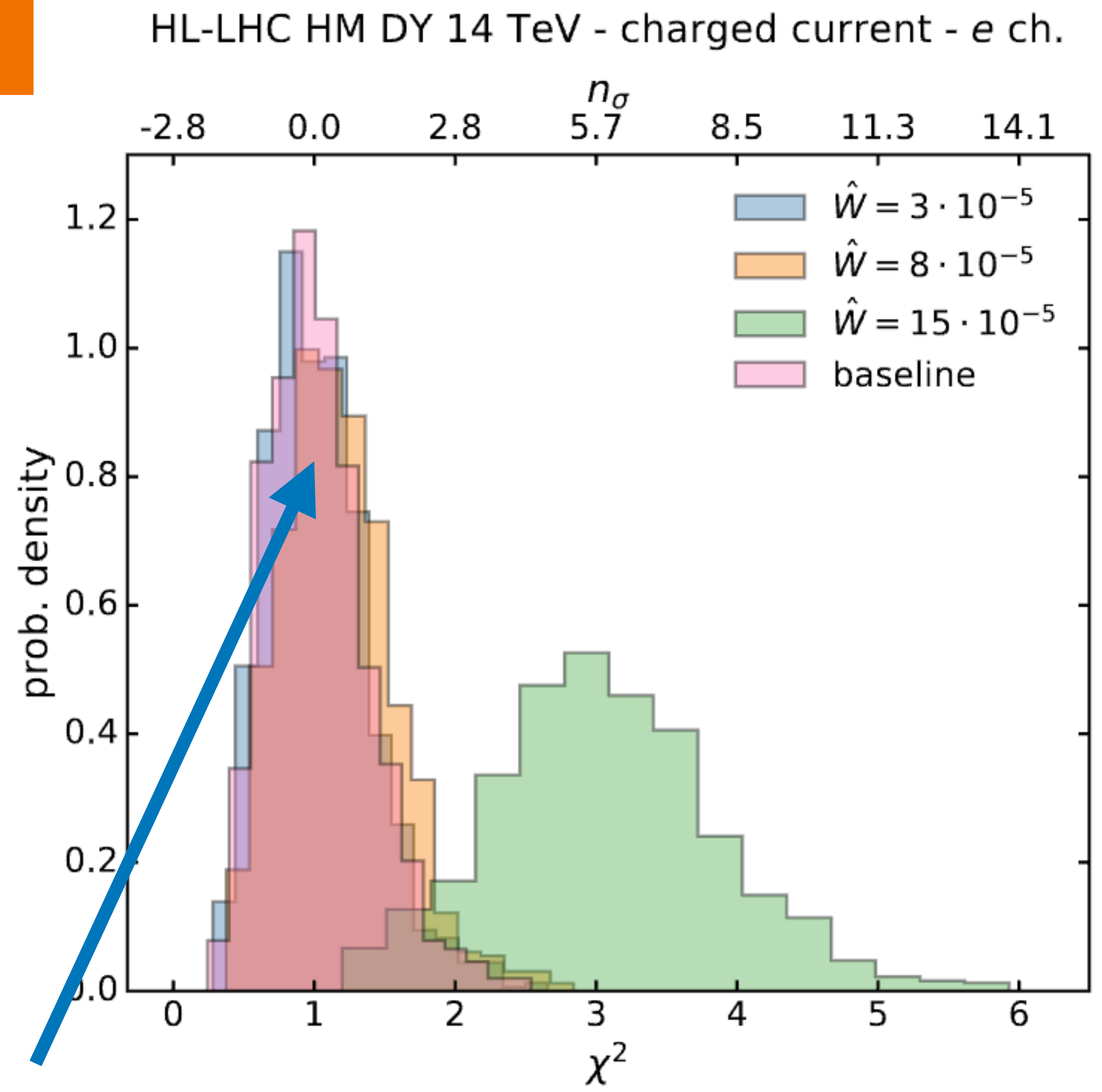
CLOSURE TEST: A TOOL TO TEST METHODOLOGY AND THEORY

- Imagine that on top of the "true" PDFs one inject the "true" NP model in the pseudo-data
- Generate HL-LHC pseudo-data assuming
  "true" law of nature = "true" PDFs + "true" UV model
- Fit PDFs assuming SM
- Can PDFs absorb signs of new physics?

Z'	W'
Selection test: ✗	Selection test: ✓
➡ Excluded from PDF fit	➡ Included in PDF fit
No impact on PDFs	**PDFs contaminated**

Test possible New Physics contamination in PDF fits





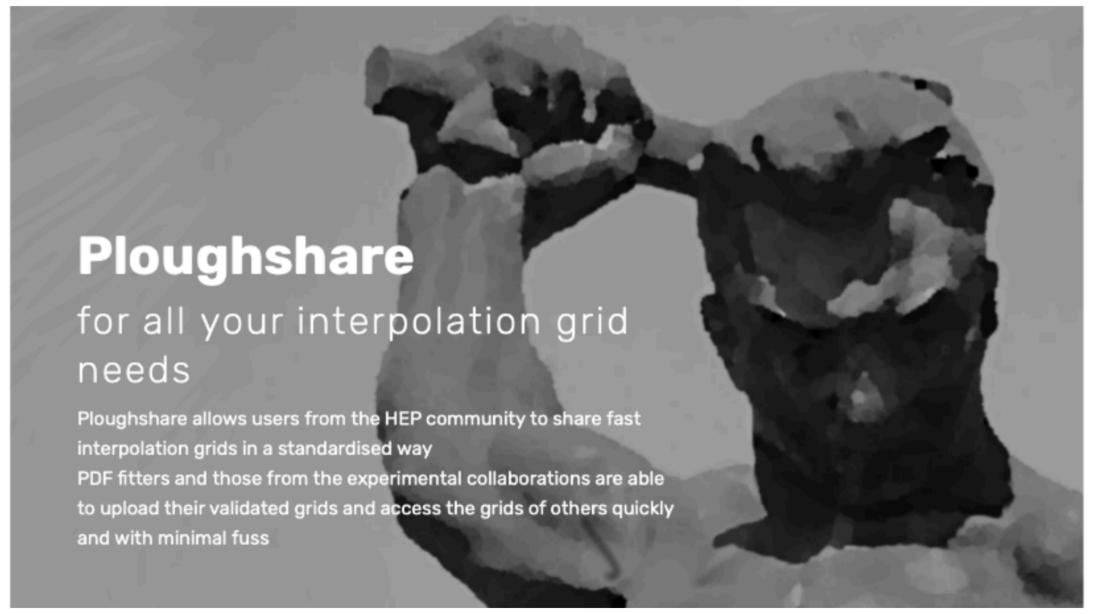
Max contamination allowed by global fit Without spoiling χ^2

Hammou, Madigan, Mangano, Mantani, Moore, Morales, MU 2307.10370

TEST GENERALISATION OF PDF AND EXTRAPOLATION

- Future tests help to discriminate among PDF sets [J. Cruz-Martinez et al, Acta Phys.Polon.B 52 (2021) 243 - on Run I]
- [Chiefa, Costantini, Cruz-Martinez et al, in progress]: test all global PDF sets agains new precise data from LHC Run I and Run II data & DIS HERA jets data.
  How well do various PDF sets describe data that are not yet included in the fit?

- Remarks:
  ➡ All results are NNLO (no k-factor approximation) thanks to PineAPPL, NNLOJET, MATRIX and Ploughshare



Carrazza et al: 2008.12789

Gehrmann-De Ridder et al: 1507.02850, 1605.04295…

Grazzini et al: 1711.06631

  ➡ NNLO PDF sets considered: PDF4LHC15, PDF4LHC21, ABMP16, CT18, CT18A, CT18Z, MSHT20, NNPDF3.1, NNPDF4.0
  ➡ The computation of χ^2 is always shown considering as uncertainties either (exp) or (exp + mho) or (exp + mho + pdf) and over the entire dataset / HEPdata entry

$$\chi^2 = \sum_{i,j=1}^{N_{\rm dat}} \left(T_i^{(0)} - D_i\right) \left(\mathrm{cov}^{-1}\right)_{ij} \left(T_j^{(0)} - D_j\right)$$

$$\chi^2_{\rm exp} \leftrightarrow \mathrm{cov} = \mathrm{cov}_{\rm exp}$$
$$\chi^2_{\rm exp+mho} \leftrightarrow \mathrm{cov} = \mathrm{cov}_{\rm exp} + \mathrm{cov}_{\rm mho}$$
$$\chi^2_{\rm exp+mho+pdf} \leftrightarrow \mathrm{cov} = \mathrm{cov}_{\rm exp} + \mathrm{cov}_{\rm mho} + \mathrm{cov}_{\rm pdf}$$

THE EXPERIMENTAL DATA IN THE TEST SET

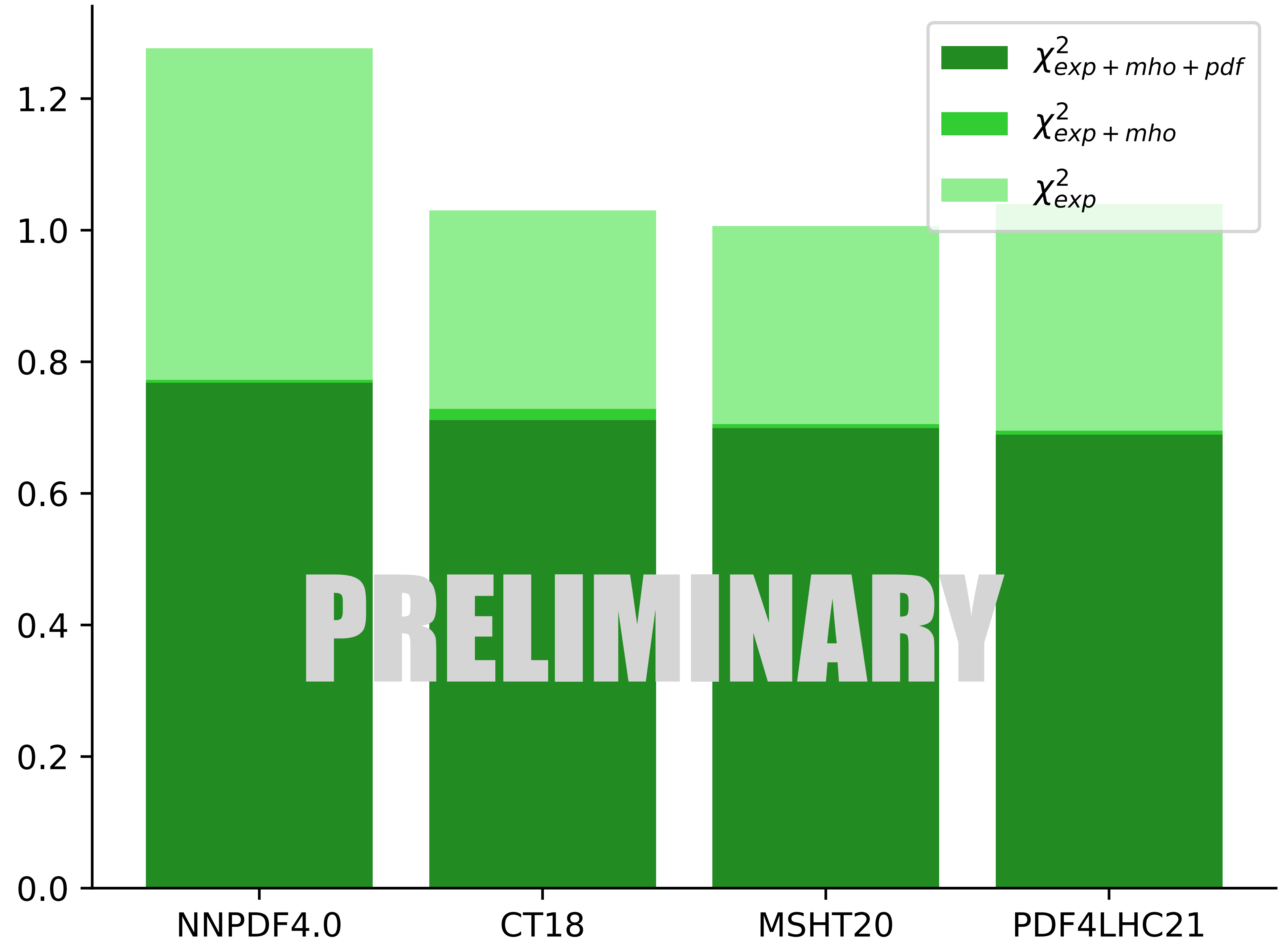
Sector	Exp.	$\sqrt{s}$ (TeV)	Channel	Observable	$\mathcal{L}$ (fb^{-1})	N_{dat}
W, Z	ATLAS	13	Z p_T spectrum	$\frac{d\sigma}{dp_T^Z}$	36.1	10
	ATLAS	8	Z incl. prod.	$\frac{d\sigma}{d\|y_{ll}\|}$	20.2	7
	CMS	8	W incl. prod.	$\frac{d\sigma_{W^\pm}}{d\eta_l}$	35.9	36
	LHCb	13	Z incl. forward prod.	$\frac{d\sigma_Z}{dy_Z}$	5.1	17
top	ATLAS	13	hadronic	$\left(\frac{1}{\sigma}\right)\frac{d\sigma}{dm_{t\bar{t}}}, \frac{d\sigma}{d\|y_{t\bar{t}}\|}, \frac{d^2\sigma}{d\|y_{t\bar{t}}\|dm_{t\bar{t}}}$	36.1	9, 12, 11
	ATLAS	13	ℓ+jets	$\left(\frac{1}{\sigma}\right)\frac{d\sigma}{dm_{t\bar{t}}}, \frac{d\sigma}{dp_{T,t}}, \frac{d\sigma}{d\|y_t\|}, \frac{d^2\sigma}{d\|y_{t\bar{t}}\|}$	36.1	9, 8, 5, 7
	CMS	13	ℓ+jets	$\left(\frac{1}{\sigma}\right)\frac{d\sigma}{dm_{t\bar{t}}}, \frac{d\sigma}{dp_{T,t}}, \frac{d\sigma}{d\|y_{t\bar{t}}\|}, \frac{d\sigma}{d\|y_t\|}, \frac{d^2\sigma}{d\|y_{t\bar{t}}\|dm_{t\bar{t}}}$	137	15, 16, 10, 11, 35
jets	ATLAS	13	incl. jet R=0.4, 0.7	$\frac{d^2\sigma}{dp_{T,j}d\|y_j\|}$	3.2	177
	ATLAS	13	di-jets R=0.4	$\frac{d^2\sigma}{dm_{jj}d\Delta y}$	3.2	136
	CMS	13	incl. jets R=0.4, 0.7	$\frac{d^2\sigma}{dp_{T,j}d\|y_j\|}$	3.2	78
DIS jets	H1	0.319	incl. jet (low q^2)	$\frac{d^2\sigma}{dq^2dpT}$	0.29	48
	H1	0.319	di-jets (low q^2)	$\frac{d^2\sigma}{dq^2d\langle pT\rangle}$	0.29	48
	H1	0.319	incl. jet (high q^2)	$\frac{d^2\sigma}{dq^2dpT}$	0.351	24
	H1	0.319	di-jets (high q^2)	$\frac{d^2\sigma}{dq^2d\langle pT\rangle}$	0.351	24
	ZEUS	0.3	incl. jet	$\frac{d^2\sigma}{dE_Tdq^2}$	0.038	30
	ZEUS	0.319	incl. jet	$\frac{d^2\sigma}{dE_Tdq^2}$	0.082	30
	ZEUS	0.319	d-jets	$\frac{d^2\sigma}{dE_Tdq^2}$	0.374	22

Chiefa et al, in progress

THE TOP SECTOR

CMS top pair production at 13 TeV, l+jet channel (137fb^{-1})

CMS collaboration [arXiv:2108.02803]

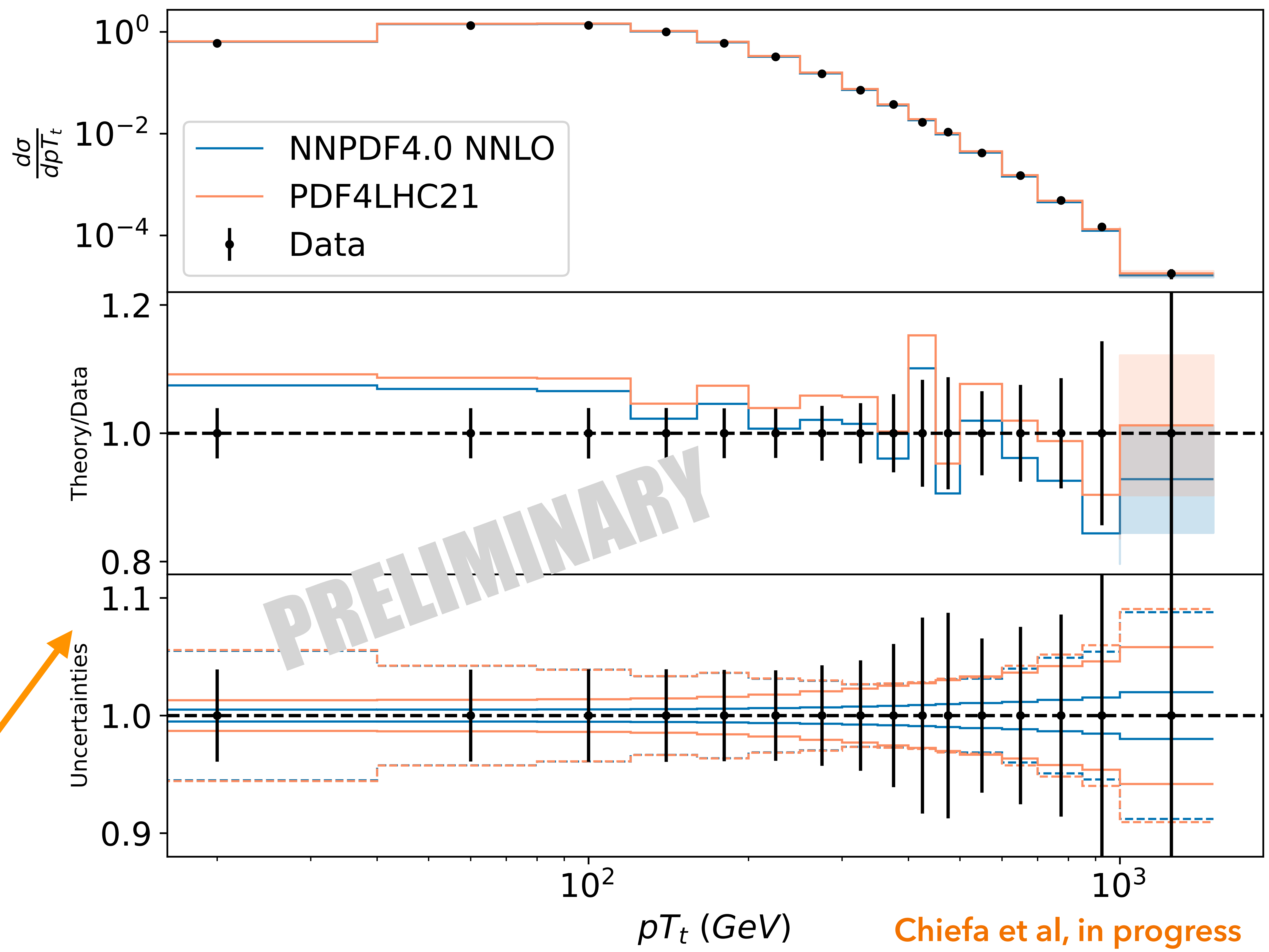


PDF set	$\frac{d\sigma}{dpT_t}$
ABMP16	0.986 (1.253)
CT18	0.712 (0.729)
CT18A	0.678 (0.686)
MSHT20	0.699 (0.705)
NNPDF3.1	0.782 (0.794)
NNPDF4.0	0.768 (0.773)
PDF4LHC15	0.709 (0.722)
PDF4LHC21	0.689 (0.695)

Absolute comparison, normalized & size of PDF (solid) and dashed (MHO) **Δ** compared to data **Δ**.

Shaded band includes MHO (9 pts variation) and PDF uncertainties added in quadrature.

CMS 13 TeV top quark pair l+j channel: $\frac{d\sigma}{dpT_t}$



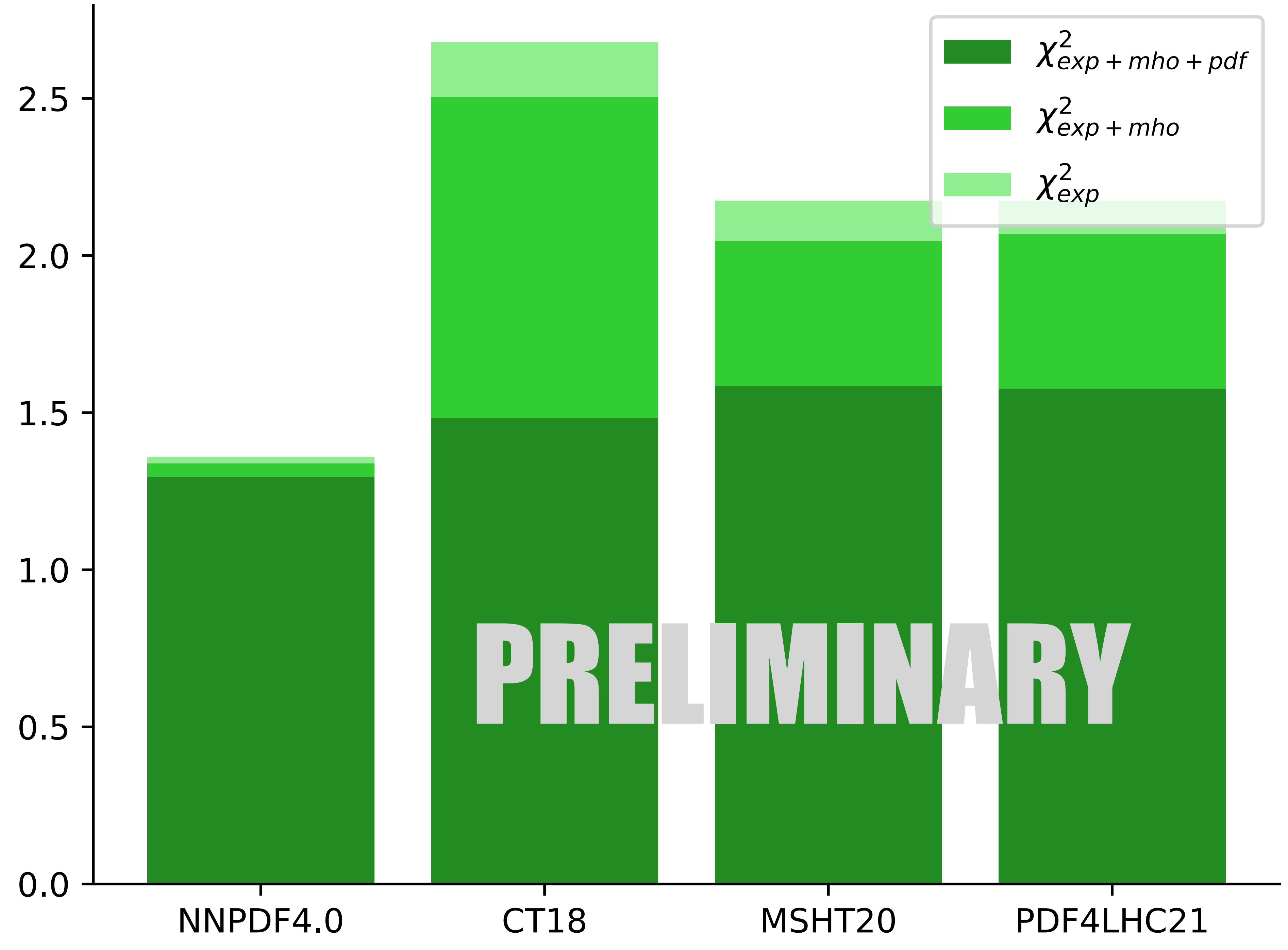
Chiefa et al, in progress

Predictions: MATRIX
Hepdata: 10.17182/hepdata.102956

THE TOP SECTOR

ATLAS top pair production at 13 TeV, l+jet channel, (36.1fb[-1])

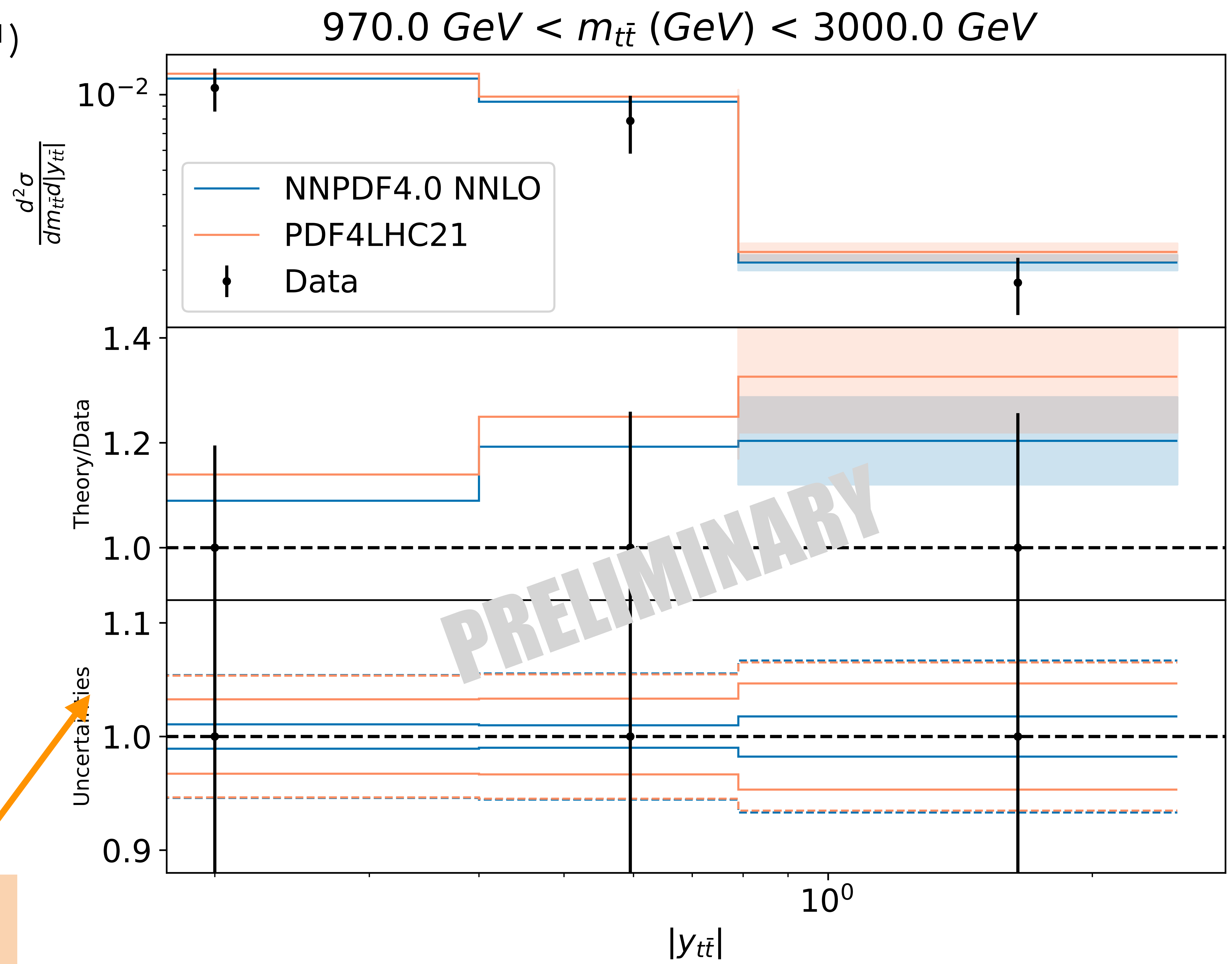
ATLAS collaboration [arXiv:1908.07305]



PDF set	$\frac{d^2\sigma}{dm_{t\bar{t}}d\lvert y_{t\bar{t}}\rvert}$
ABMP16	0.832 (0.835)
CT18	1.483 (2.504)
CT18A	1.489 (2.347)
MSHT20	1.585 (2.046)
NNPDF3.1	1.200 (1.244)
NNPDF4.0	1.297 (1.338)
PDF4LHC15	1.298 (2.088)
PDF4LHC21	1.577 (2.068)

Absolute comparison, normalized & size of PDF (solid) and dashed (MHO) **Δ** compared to data **Δ**.

Shaded band includes MHO (9 pts variation) and PDF uncertainties added in quadrature.



Chiefa et al, in progress

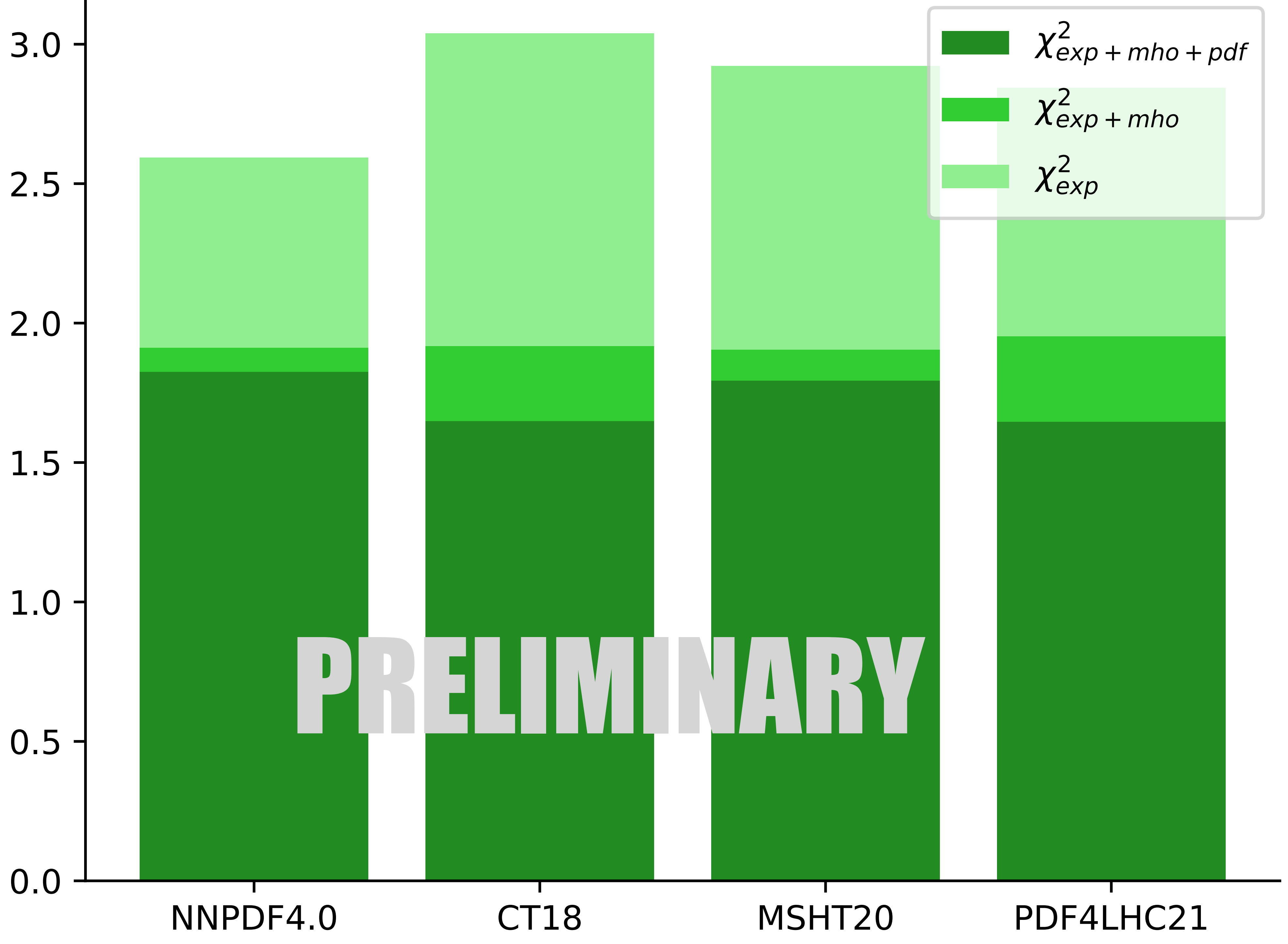
Predictions: MATRIX

Hepdata: 10.17182/hepdata.95758

THE JET SECTOR

ATLAS di-jet cross-sections at 13 TeV (3.2fb-1), R=0.4, anti-kT

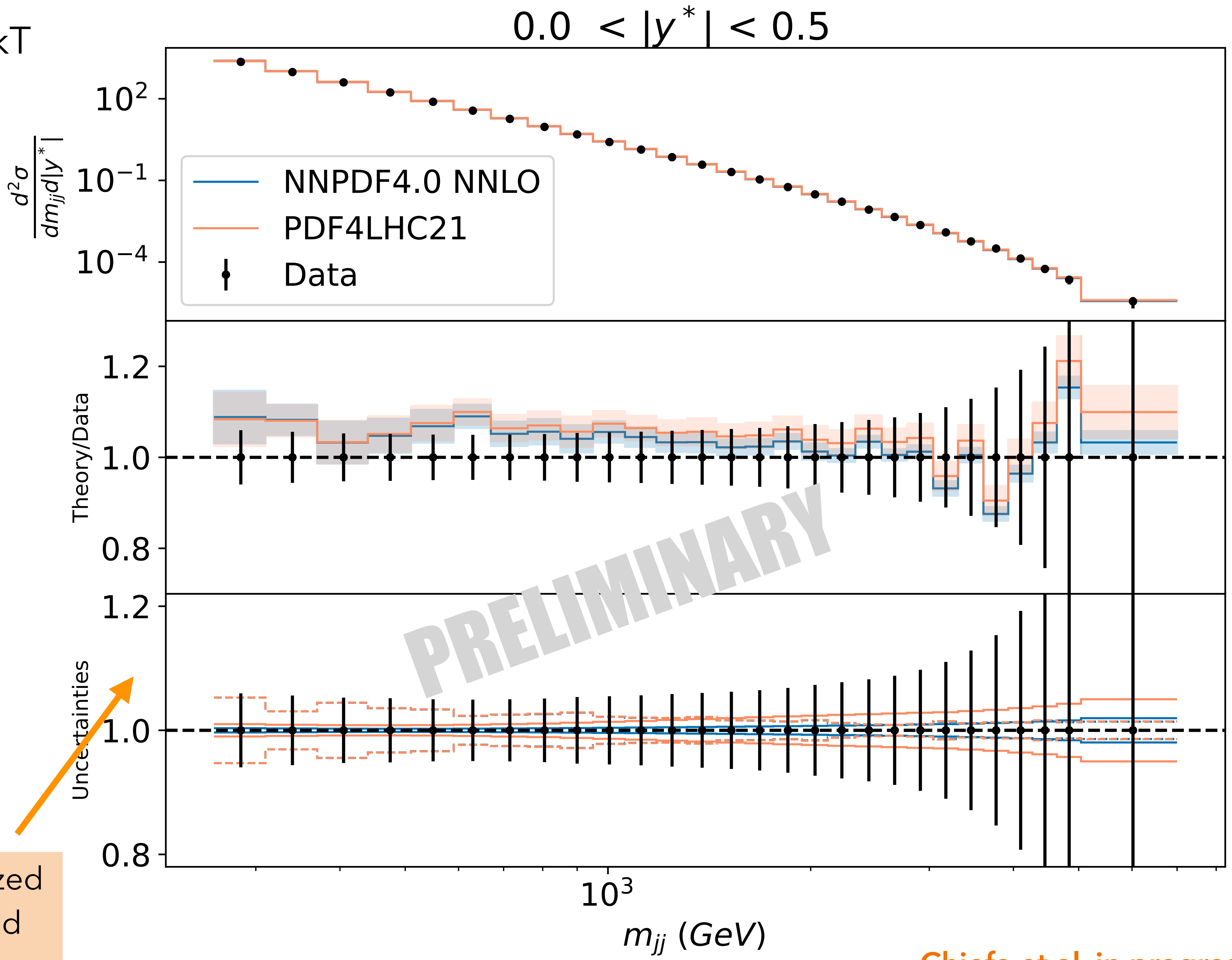
ATLAS collaboration [arXiv:1711.02692]



PDF set	
ABMP16	1.72 (2.12)
CT18	1.65 (1.92)
CT18A	1.67 (1.91)
CT18Z	1.61 (**1.88**)
MSHT20	1.79 (1.91)
‘ NNPDF3.1	1.78 (2.03)
NNPDF4.0	1.82 (1.91)
PDF4LHC15	**1.59** (1.91)
PDF4LHC21	1.65 (1.95)

Absolute comparison, normalized & size of PDF (solid) and dashed (MHO) **Δ** compared to data **Δ**.

Shaded band includes MHO (9 pts variation) and PDF uncertainties added in quadrature.



Chiefa et al, in progress

Leading color NNLO correction

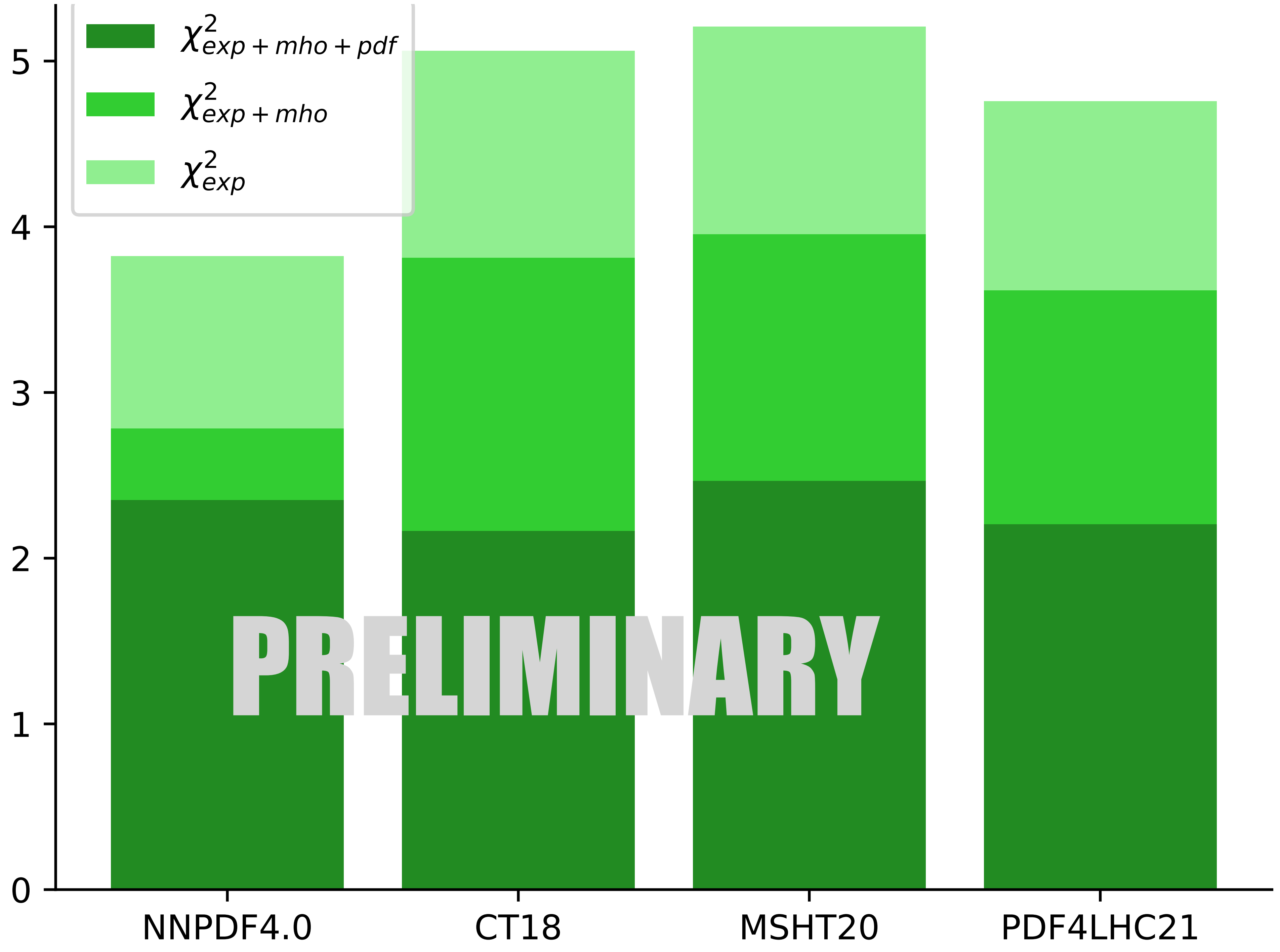
Predictions: NNLOJET (plougshare)

Hepdata: 10.17182/hepdata.79952

THE JET SECTOR

CMS inclusive jet cross-sections at 13 TeV (33.5fb-1) anti-kT, R=0.7

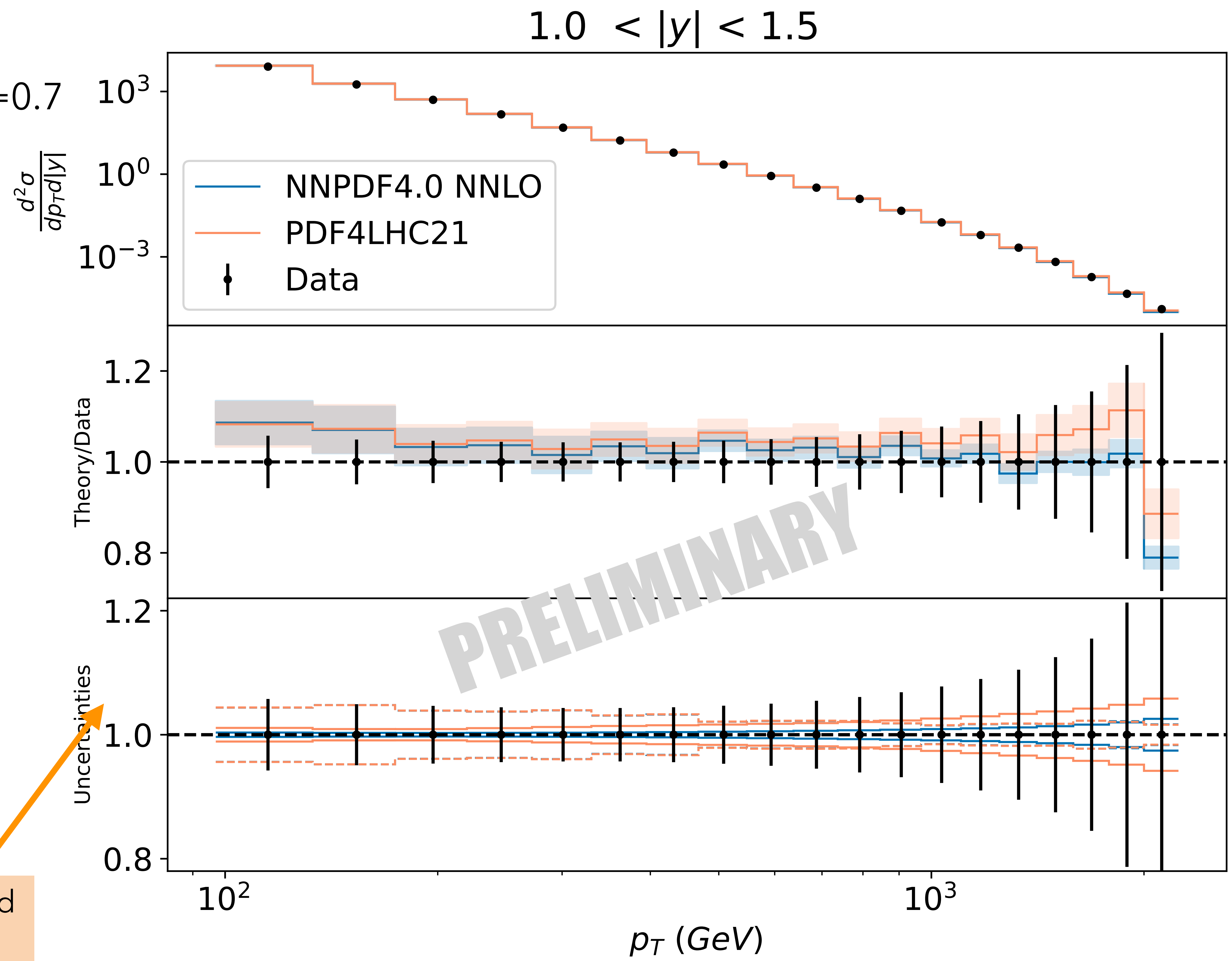
CMS collaboration [arXiv:2111.10431]



PDF set	
ABMP16	2.77 (3.71)
CT18	2.16 (3.81)
CT18A	2.16 (3.77)
CT18Z	**2.06** (3.34)
MSHT20	2.47 (3.96)
‘ NNPDF3.1	2.52 (3.88)
NNPDF4.0	2.35 (**2.78**)
PDF4LHC15	2.15 (3.80)
PDF4LHC21	2.21 (3.62)

Absolute comparison, normalized & size of PDF (solid) and dashed (MHO) **Δ** compared to data **Δ**.

Shaded band includes MHO (9 pts variation) and PDF uncertainties added in quadrature.



Chiefa et al, in progress

Leading color NNLO correction

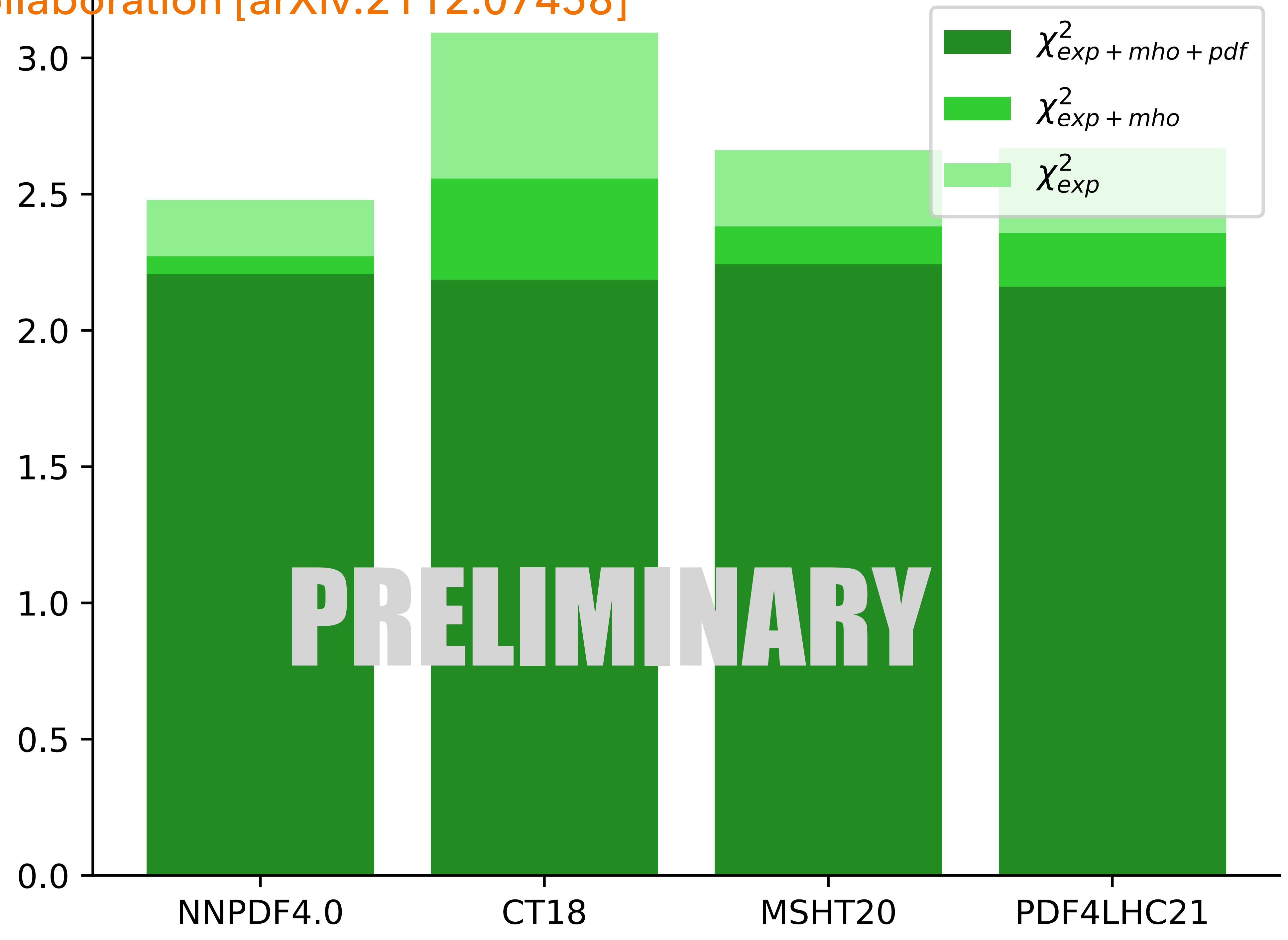
Predictions: NNLOJET (plougshare)

Hepdata: 10.17182/hepdata.115022.v2

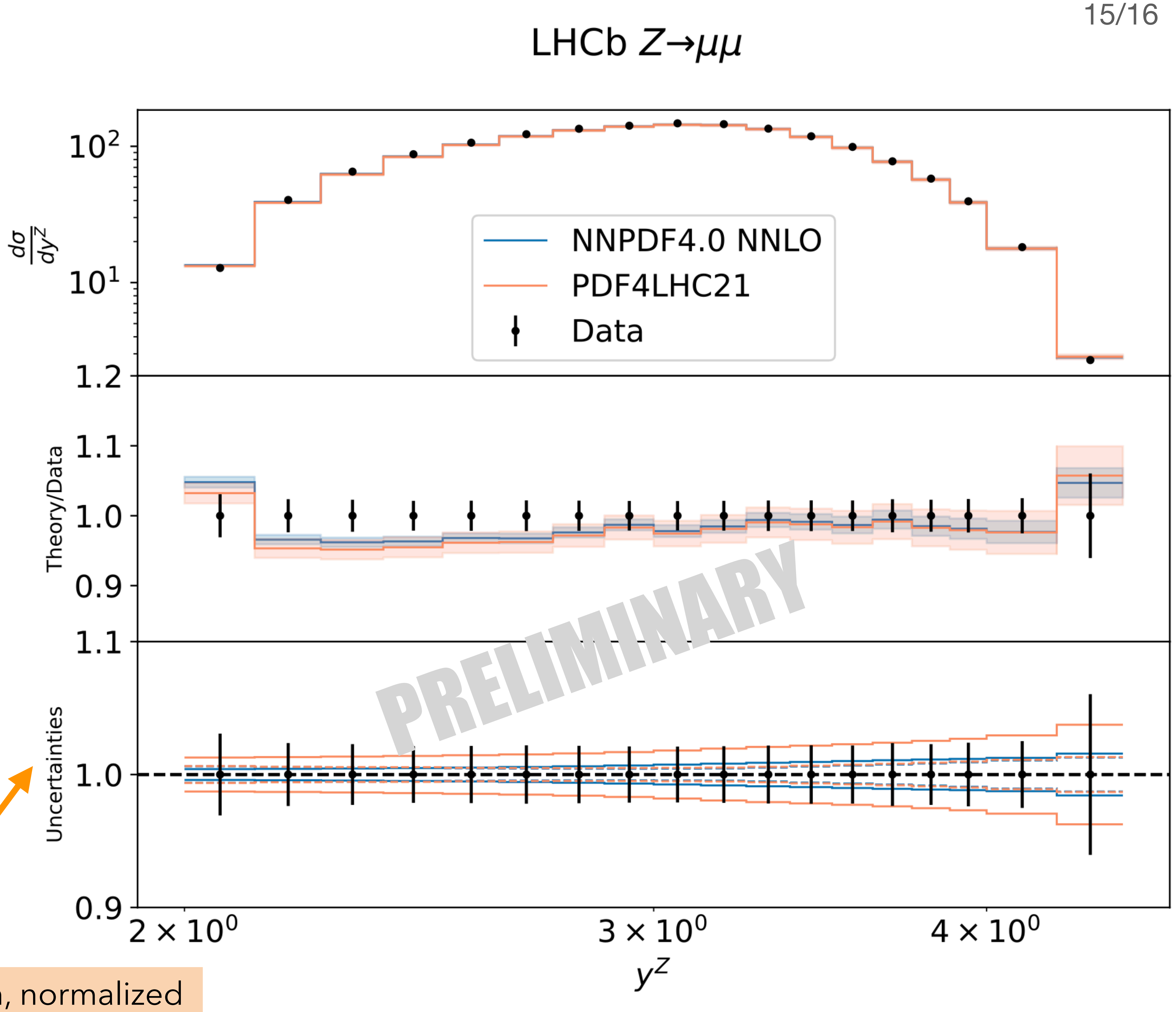
THE DRELL-YAN SECTOR

LHCb Z forward production at 13 TeV (5.1fb-1)

LHCb collaboration [arXiv:2112.07458]



PDF set	LHCb $\frac{d\sigma}{dy^Z}$ 13 TeV
ABMP16	2.22 (**2.26**)
CT18	2.19 (2.56)
CT18A	2.25 (2.49)
MSHT20	2.24 (2.38)
NNPDF3.1	2.49 (2.73)
NNPDF4.0	2.21 (2.27)
PDF4LHC15	2.17 (2.59)
PDF4LHC21	**2.16** (2.36)



Absolute comparison, normalized & size of PDF (solid) and dashed (MHO) **Δ** compared to data **Δ**.

Shaded band includes MHO (9 pts variation) and PDF uncertainties added in quadrature.

Chiefa et al, in progress

Predictions: MATRIX
Hepdata: 10.17182/hepdata.102956

CONCLUSIONS AND OUTLOOK

- In an era of precision at LHC, need precise and accurate PDFs
- Tools like closure tests are crucial to test methodology robustness, effects of possible experimental inconsistencies and even possible effects of new physics in the high energy tails
- Progress possible thanks to automated tools for fast computation of observables and sharing (NNLOJET, MATRIX, PineAPPL, PloughShare, APPLgrid, FastNLO) and public PDF tools (XFitter, NNPDF)
- Comparing performance of broad range of PDF sets on unseen data we see that, while NNPDF4.0 PDF uncertainties are smaller, the data-theory description is similar to that of other PDF set
- Predictions for NNPDF4.0 seem to fall close the data, hinting that the improvement in precision (smaller uncertainties than NNPDF3.1) came with an improvement in accuracy
- Oscillations far from the data can have big MHO uncertainties, luckily the era of N3LO PDFs is starting now!

EXTRA MATERIAL

NNPDF4.0: THE ROLE OF LHC DATA

Data set	NNPDF3.1	NNPDF4.0	ABMP16	CT18	MSHT20
CMS W asym. 7 TeV ($\mathcal{L} = 36$ pb^{-1})	✗	✗	✗	✗	✓
CMS Z 7 TeV ($\mathcal{L} = 36$ pb^{-1})	✗	✗	✗	✗	✓
CMS W electron asymmetry 7 TeV	✓	✓	✗	✓	✓
CMS W muon asymmetry 7 TeV	✓	✓	✓	✓	✗
CMS Drell-Yan 2D 7 TeV	✓	✓	✗	(✓)	✓
CMS Drell-Yan 2D 8 TeV	(✓)	✗	✗	✗	✗
CMS W rapidity 8 TeV	✓	✓	✓	✓	✓
CMS W, Z p_T 8 TeV ($\mathcal{L} = 18.4$ fb^{-1})	✗	✗	✗	(✓)	✗
CMS Z p_T 8 TeV	✓	✓	✗	(✓)	✗
CMS $W + c$ 7 TeV	✓	✓	✗	(✓)	✓
CMS $W + c$ 13 TeV	✗	✓	✗	✗	(✓)
CMS single-inclusive jets 2.76 TeV	✓	✗	✗	✗	✓
CMS single-inclusive jets 7 TeV	✓	(✓)	✗	✓	✓
CMS dijets 7 TeV	✗	✓	✗	✗	✗
CMS single-inclusive jets 8 TeV	✗	✓	✗	✓	✓
CMS 3D dijets 8 TeV	✗	(✓)	✗	✗	✗
CMS σ_{tt}^{tot} 5 TeV	✗	✓	✗	✗	✗
CMS σ_{tt}^{tot} 7, 8 TeV	✓	✓	✗	✗	✗
CMS σ_{tt}^{tot} 8 TeV	✗	✗	✗	✗	✓
CMS σ_{tt}^{tot} 5, 7, 8, 13 TeV	✗	✗	✓	✗	✗
CMS σ_{tt}^{tot} 13 TeV	✓	✓	✓	✗	✗
CMS $t\bar{t}$ lepton+jets 8 TeV	✓	✓	✗	✗	✓
CMS $t\bar{t}$ 2D dilepton 8 TeV	✗	✓	✗	✓	✓
CMS $t\bar{t}$ lepton+jet 13 TeV	✗	✓	✗	✗	✗
CMS $t\bar{t}$ dilepton 13 TeV	✗	✓	✗	✗	✗
CMS single top $\sigma_t + \sigma_{\bar{t}}$ 7 TeV	✗	✓	✓	✗	✗
CMS single top R_t 8, 13 TeV	✗	✓	✓	✗	✗
CMS single top 13 TeV	✗	✗	✗	✗	(✓)

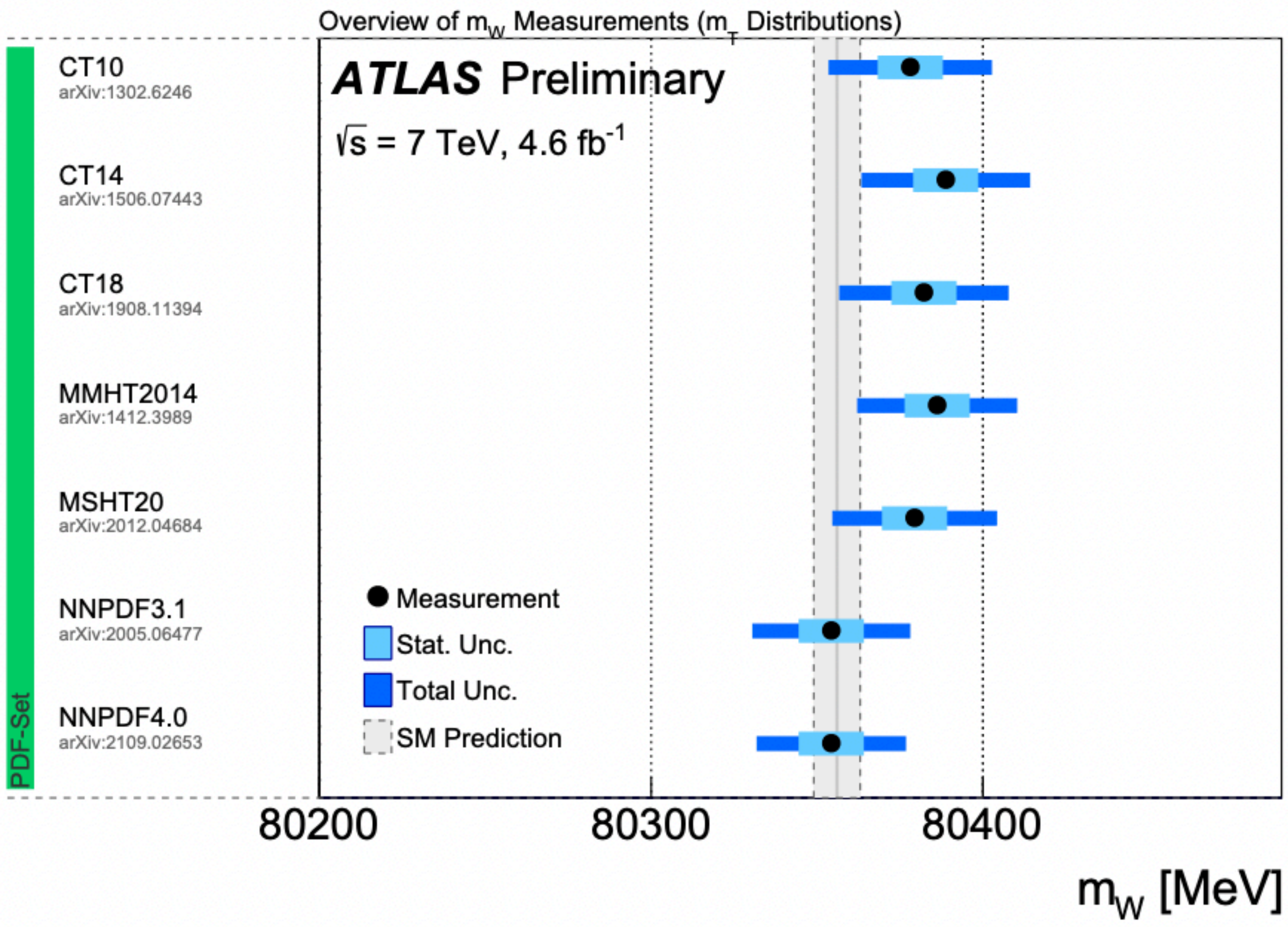
Data set	NNPDF3.1	NNPDF4.0	ABMP16	CT18	MSHT20
ATLAS W, Z 7 TeV ($\mathcal{L} = 35$ pb^{-1})	✓	✓	✓	✓	✓
ATLAS W, Z 7 TeV ($\mathcal{L} = 4.6$ fb^{-1})	✓	✓	✗	(✓)	✓
ATLAS low-mass DY 7 TeV	✓	✓	✗	(✓)	✗
ATLAS high-mass DY 7 TeV	✓	✓	✗	(✓)	✓
ATLAS W 8 TeV	✗	(✓)	✗	✗	✓
ATLAS DY 2D 8 TeV	✗	✓	✗	✗	✓
ATLAS high-mass DY 2D 8 TeV	✗	✓	✗	(✓)	✓
ATLAS $\sigma_{W,Z}$ 13 TeV	✗	✓	✓	✗	✗
ATLAS W+jet 8 TeV	✗	✓	✗	✗	✓
ATLAS Z p_T 7 TeV	(✓)	✗	✗	(✓)	✗
ATLAS Z p_T 8 TeV	✓	✓	✗	✓	✓
ATLAS $W + c$ 7 TeV	✗	✓	✗	(✓)	✗
ATLAS σ_{tt}^{tot} 7, 8 TeV	✓	✓	✓	✗	✗
ATLAS σ_{tt}^{tot} 7, 8 TeV	✗	✗	✓	✗	✗
ATLAS σ_{tt}^{tot} 13 TeV ($\mathcal{L} = 3.2$ fb^{-1})	✓	✗	✓	✗	✗
ATLAS σ_{tt}^{tot} 13 TeV ($\mathcal{L} = 139$ fb^{-1})	✗	✓	✗	✗	✗
ATLAS σ_{tt}^{tot} and Z ratios	✗	✗	✗	✗	(✓)
ATLAS $t\bar{t}$ lepton+jets 8 TeV	✓	✓	✗	✓	✓
ATLAS $t\bar{t}$ dilepton 8 TeV	✗	✓	✗	✗	✓
ATLAS single-inclusive jets 7 TeV, R=0.6	✓	(✓)	✗	✓	✓
ATLAS single-inclusive jets 8 TeV, R=0.6	✗	✓	✗	✗	✗
ATLAS dijets 7 TeV, R=0.6	✗	✓	✗	✗	✗
ATLAS direct photon production 8 TeV	✗	(✓)	✗	✗	✗
ATLAS direct photon production 13 TeV	✗	✓	✗	✗	✗
ATLAS single top R_t 7, 8, 13 TeV	✗	✓	✓	✗	✗
ATLAS single top diff. 7 TeV	✗	✓	✗	✗	✗
ATLAS single top diff. 8 TeV	✗	✓	✗	✗	✗

- NNPDF4.0 based on larger set of data from LHC Run I and Run II
- Open-source public code **[arXiv:2109.02671]**
- O(4500) data points and first time inclusion of di-jets preferred over inclusive jets based on perturbative stability **[Khalek et al, 2005.11327]**

PHENOMENOLOGICAL IMPLICATIONS

ATLAS-CONF-2023-004

PDF-Set	p_T^ℓ [MeV]	m_T [MeV]	combined [MeV]
CT10	$80355.6^{+15.8}_{-15.7}$	$80378.1^{+24.4}_{-24.8}$	$80355.8^{+15.7}_{-15.7}$
CT14	$80358.0^{+16.3}_{-16.3}$	$80388.8^{+25.2}_{-25.5}$	$80358.4^{+16.3}_{-16.3}$
CT18	$80360.1^{+16.3}_{-16.3}$	$80382.2^{+25.3}_{-25.3}$	$80360.4^{+16.3}_{-16.3}$
MMHT2014	$80360.3^{+15.9}_{-15.9}$	$80386.2^{+23.9}_{-24.4}$	$80361.0^{+15.9}_{-15.9}$
MSHT20	$80358.9^{+13.0}_{-16.3}$	$80379.4^{+24.6}_{-25.1}$	$80356.3^{+14.6}_{-14.6}$
NNPDF3.1	$80344.7^{+15.6}_{-15.5}$	$80354.3^{+23.6}_{-23.7}$	$80345.0^{+15.5}_{-15.5}$
NNPDF4.0	$80342.2^{+15.3}_{-15.3}$	$80354.3^{+22.3}_{-22.4}$	$80342.9^{+15.3}_{-15.3}$

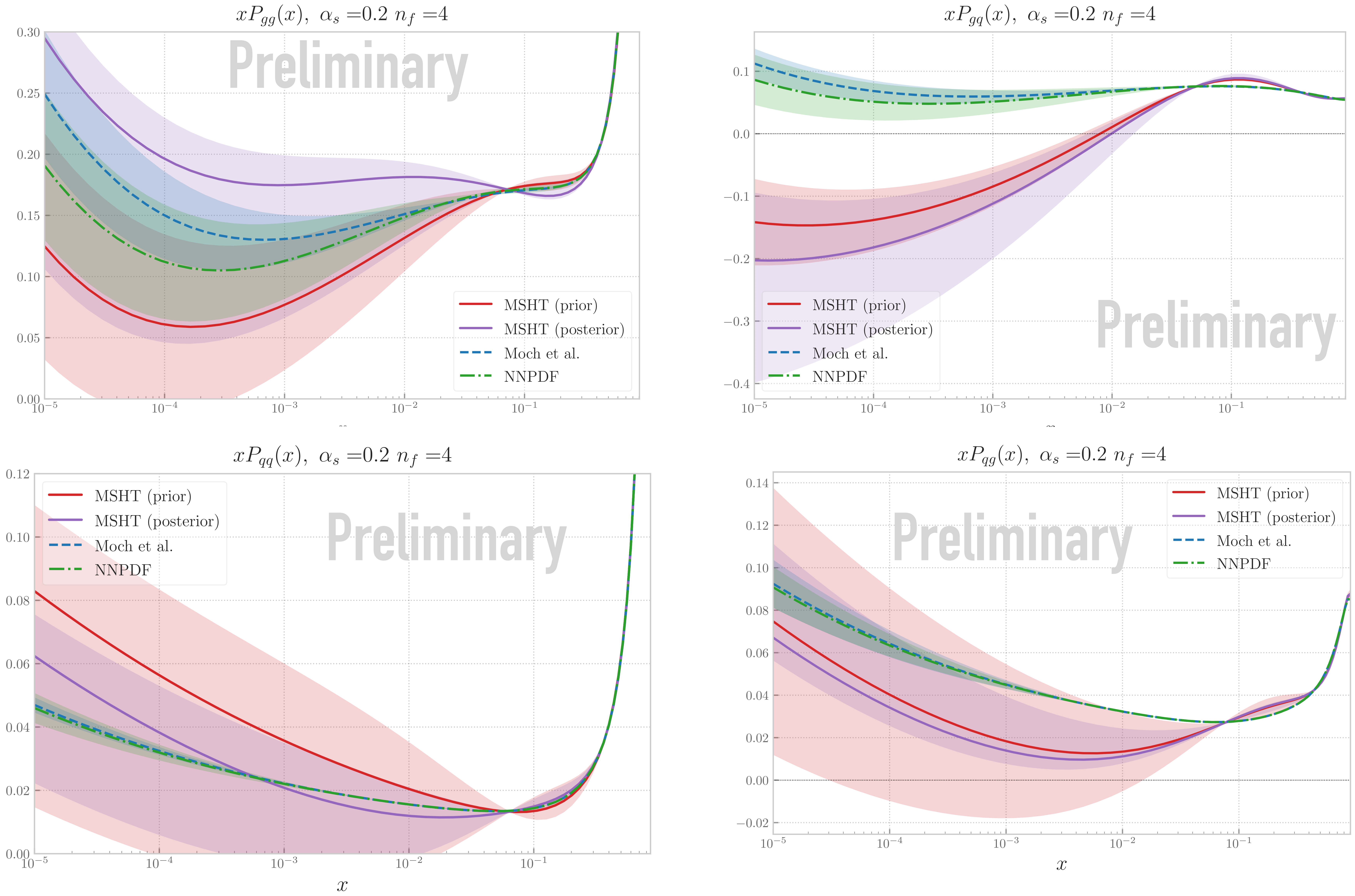


- Different PDF uncertainties yield slightly different total uncertainties.
- Differences of ~18 MeV between central values obtained with NNPDF4.0 and CT18 comparable (larger) than measurement uncertainty
- Important to confirm results, understand reason behind the differences & stress-test methodologies

#1 – APPROXIMATE N3LO PDFS

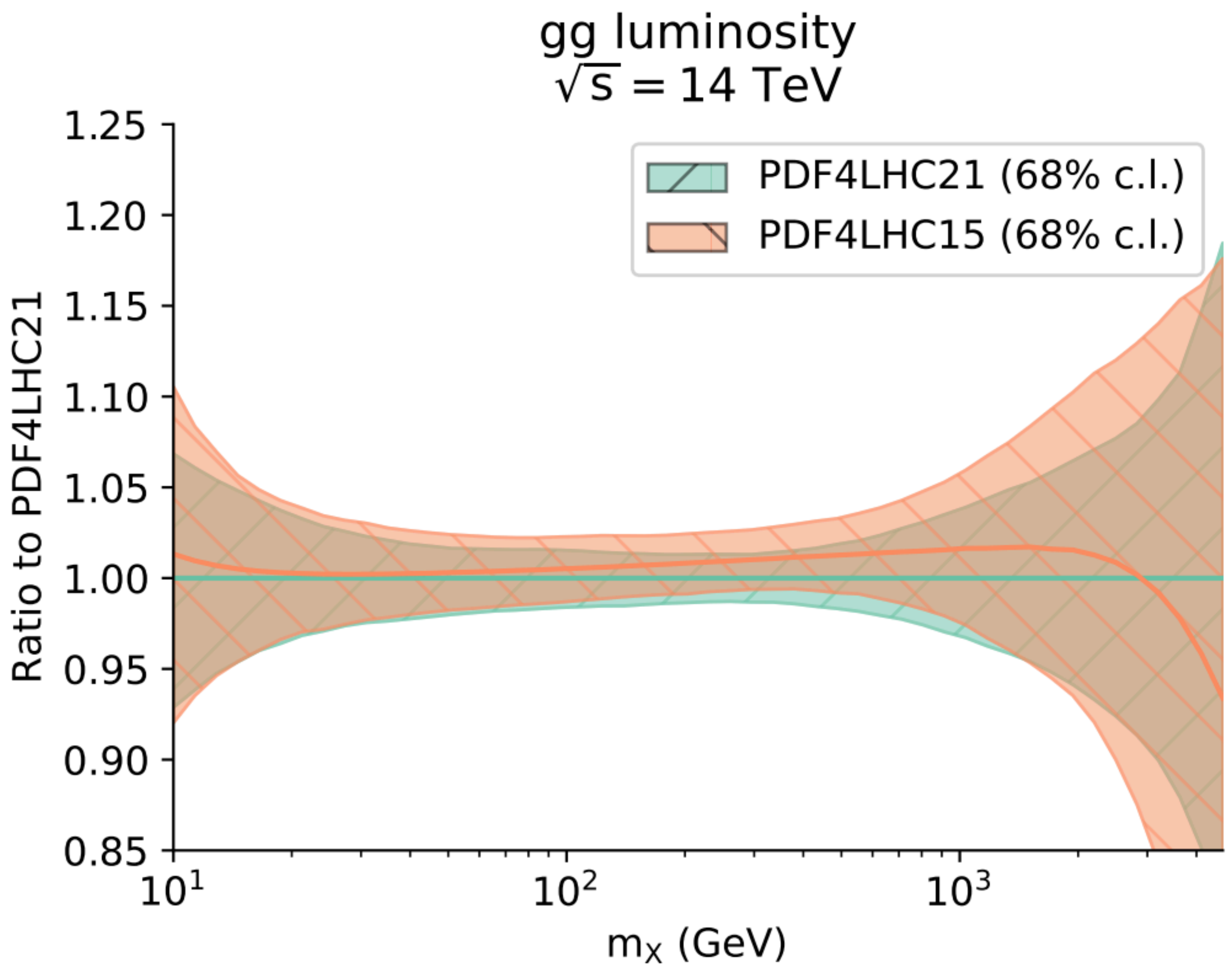
MSHT prior = pre-fit variation
MSHT posterior = MSHTaN3LO
NNPDF = NNPDF40aN3LO
Moch at al = theory papers

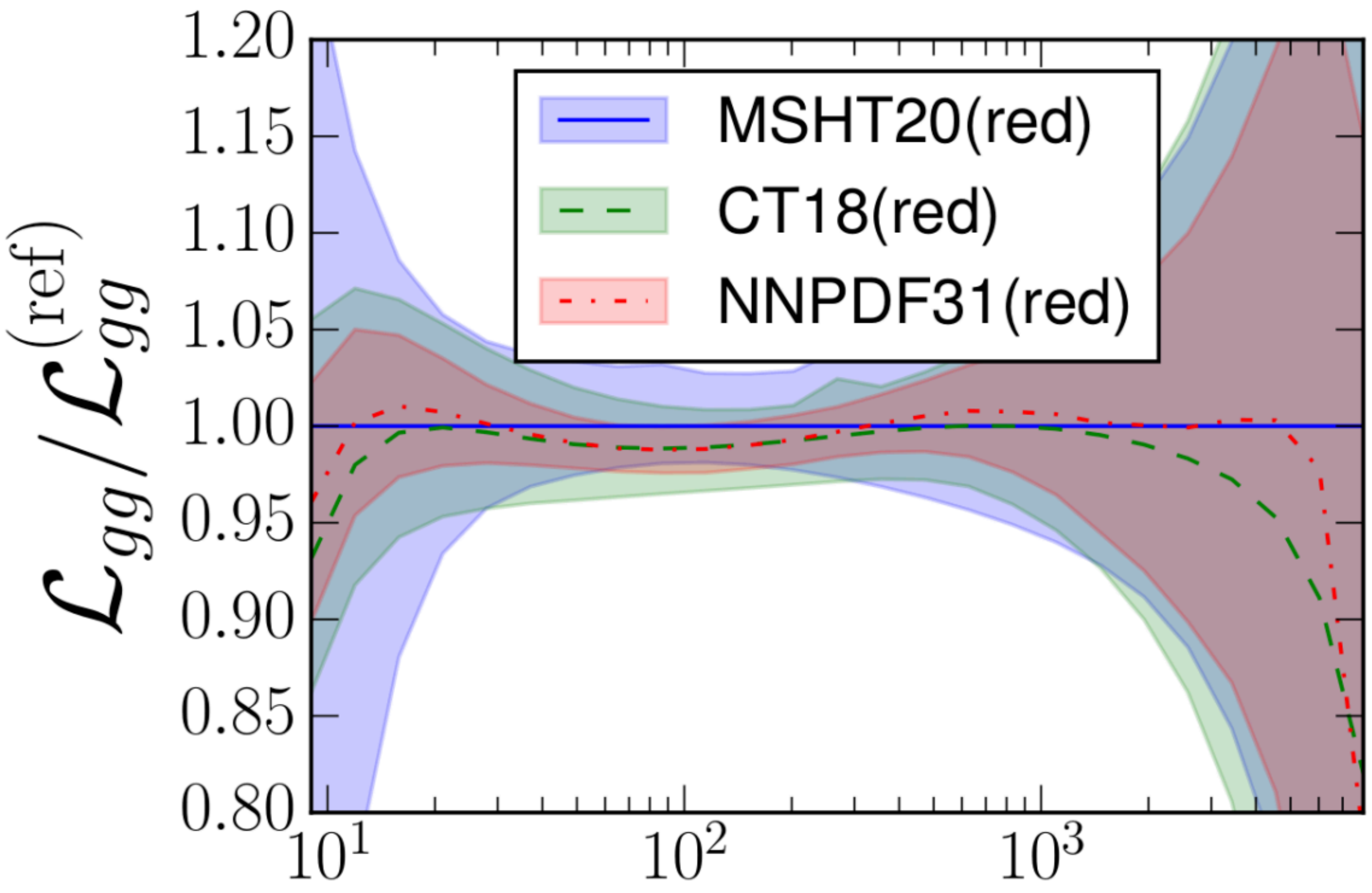
- Different known ingredients included and different ways of parametrising sub-leading contributions produces different candidates (IHOU)
- Benchmarking exercise started before Summer in Les Houches to check the impact of the aN3LO splitting functions
- Write-up in preparation for Les Houches proceedings



THE PDF4LHC21 COMBINATION

PDF4LHC WG, arXiv: 2203.05506





- In recent updates from global PDF fitting collaborations (NNPDF3.1, CT18, MSHT20, ABMP16) effect of LHC data driving PDF uncertainties down ➜ PDF4LHC21 combination of NNPDF31′, CT18′ and MSHT20′ smaller uncertainties than PDF4LHC15
- The increased precision of the data and their strong correlation demands methodological improvements (e. g. more flexible parametrisation that shifts PDFs outside nominal error band - MSHT) and shows some tension among (CT18 vs CT18Z)
- Benchmark among NNPDF3.1, MSHT20 and CT18: overall agreement, which improves once common dataset is used, differences in uncertainties with ΔCT18 ≳ ΔMSHT20 ≳ ΔNNPDF31
- NNPDF4.0 [arXiv:2109.02653], MSHT20aN3LO [arXiv:2207.04739] state-of-the-art post PDF4LHC21 sets

#1 – APPROXIMATE N3LO PDFS

➡ Several ingredients required to perform N^{3}LO PDF fits, many available some missing [S. Moch's talk]

➡ Only publicly available aN^{3}LO PDF determination is from the MSHT collaboration [McGowan et al, arXiv:2207.04739]

➡ NNPDF has presented preliminary aN^{3}LO results and paper is to appear soon

MSHT

- All available theory input at the time of publication included (impact of new ingredients being explored)
- Incomplete N3LO terms added as variation in the prior and estimated by fitting nuisance parameters to the data (hence **posterior determined by fitting data**)
- No MHOU associated with NNLO contributions

NNPDF

- More theory inputs published in between and included, in particular 6(1) extra momentum for P_{qg}, P_{qq} (P_{gq}, P_{gg}), some terms in the large n_f limit, sub-leading small-x and large-x terms.
- **Only theory inputs** and their variations added to an additional theory covariance matrix associated with incomplete missing higher orders (IHOU)
- MHOU associated with NNLO included via theory covariance matrix

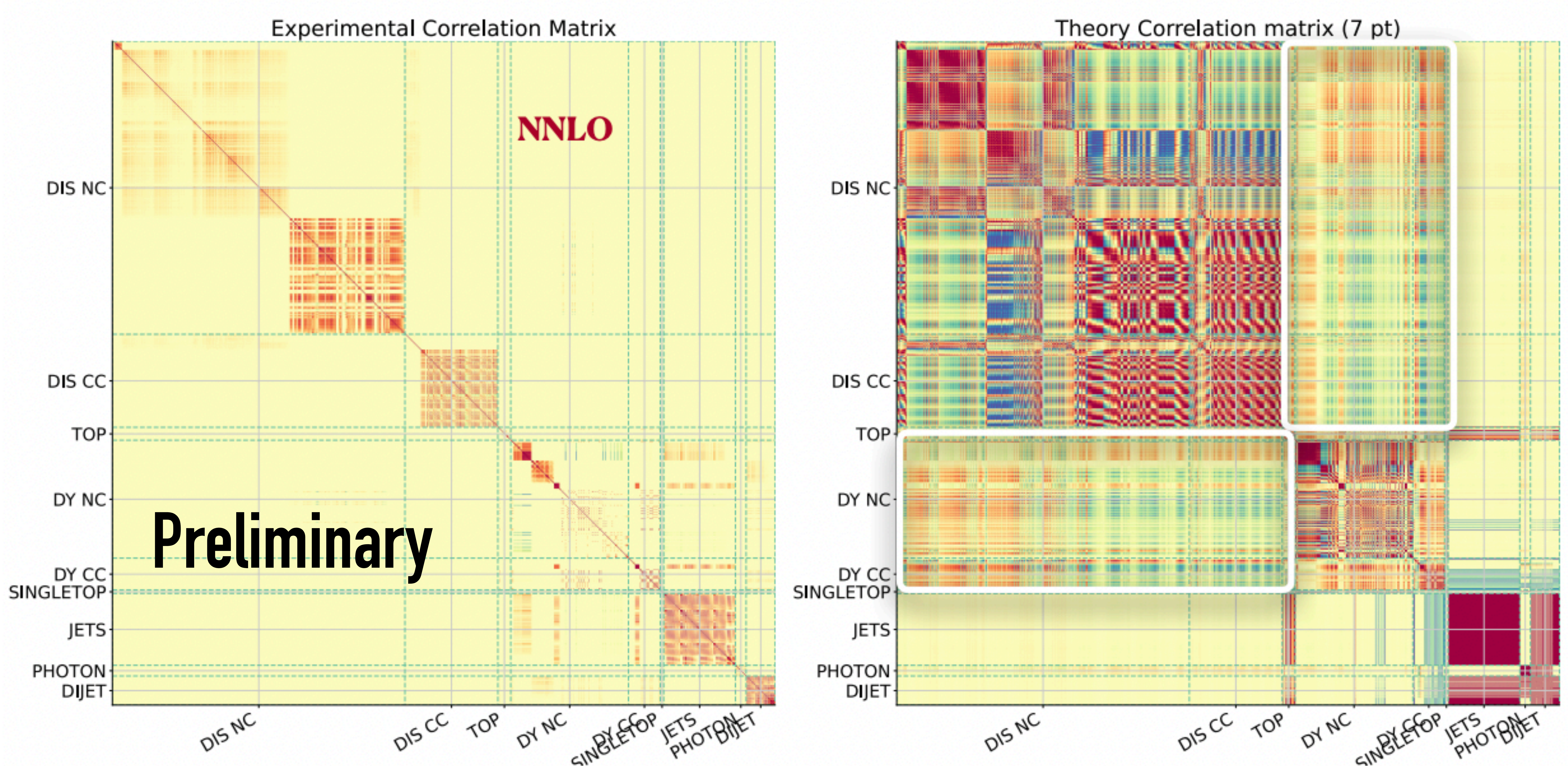
#2 – MISSING HIGHER ORDER UNCERTAINTIES IN PDF FITS

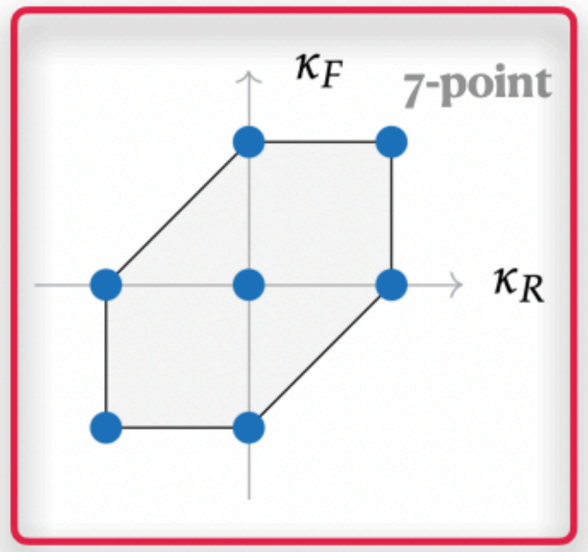
➡PDF uncertainties only include experimental component: how to include uncertainty associated to truncation of perturbative expansion in theoretical predictions used in PDF fits?

➡One option is to construct a theory covariance matrix from scale-varied cross sections and combine it with the experimental covariance matrix

$$\chi^2 = \sum_{m,n=1}^{N} (d_m - t_m)(\mathrm{cov}_{\mathrm{exp}} + \mathrm{cov}_{\mathrm{th}})^{-1}_{mn}(d_n - t_n)$$

- Assumptions: experimental and theoretical errors independent and Gaussian

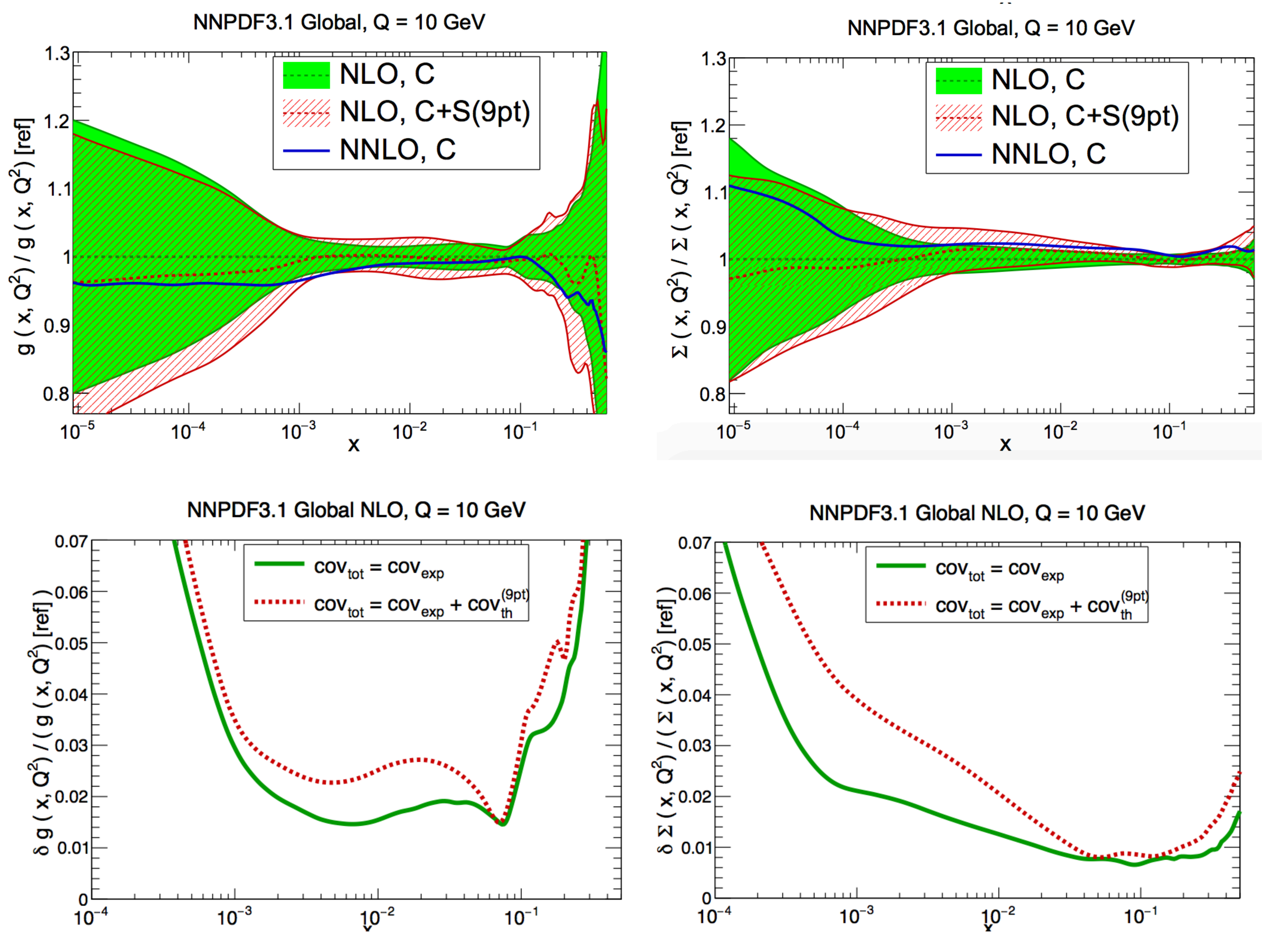




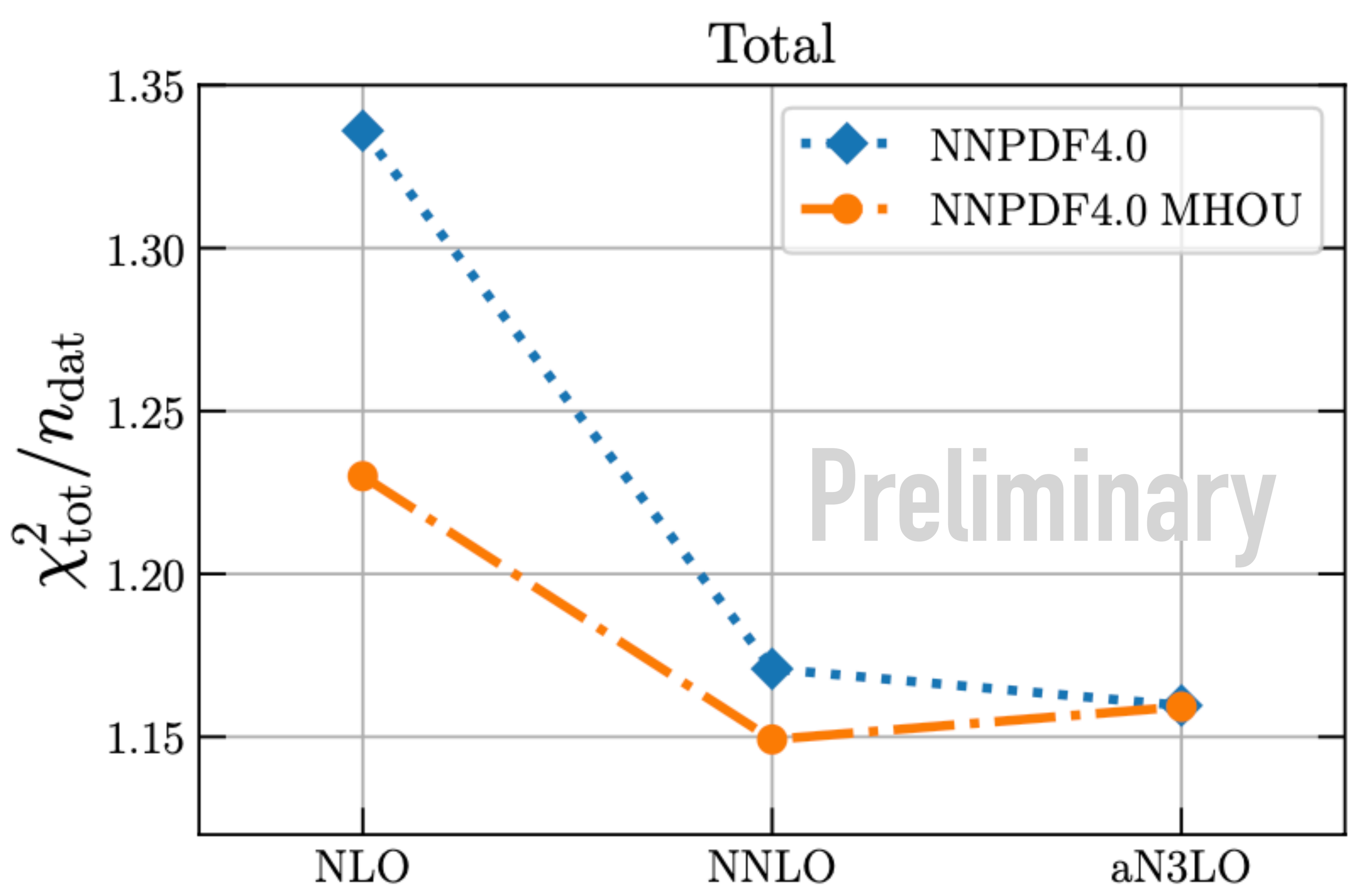
7-pts variation

[NNPDF: 1906.10698]
NNLO paper in preparation

#2 – MISSING HIGHER ORDER UNCERTAINTIES IN PDF FITS



Abdul Khalek et al *Eur.Phys.J.C* 79 (2019) 11



➡ NNPDF40NNLO with MHOU and NNPDF40aN3LO will be published soon

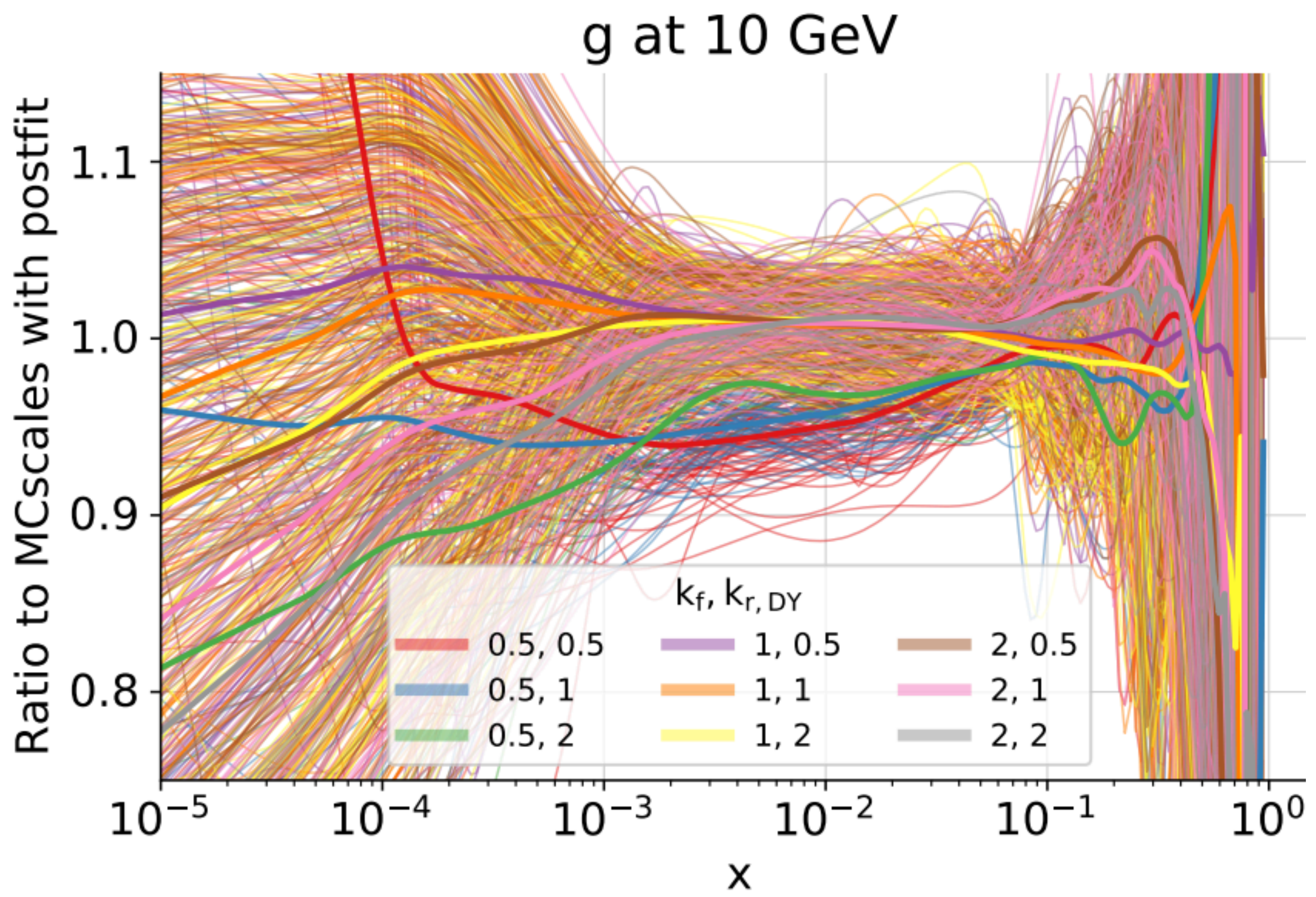
$$Cov_{ij} = Cov_{ij,EXP} + Cov_{ij,MHOU} + Cov_{ij,IHOU}$$

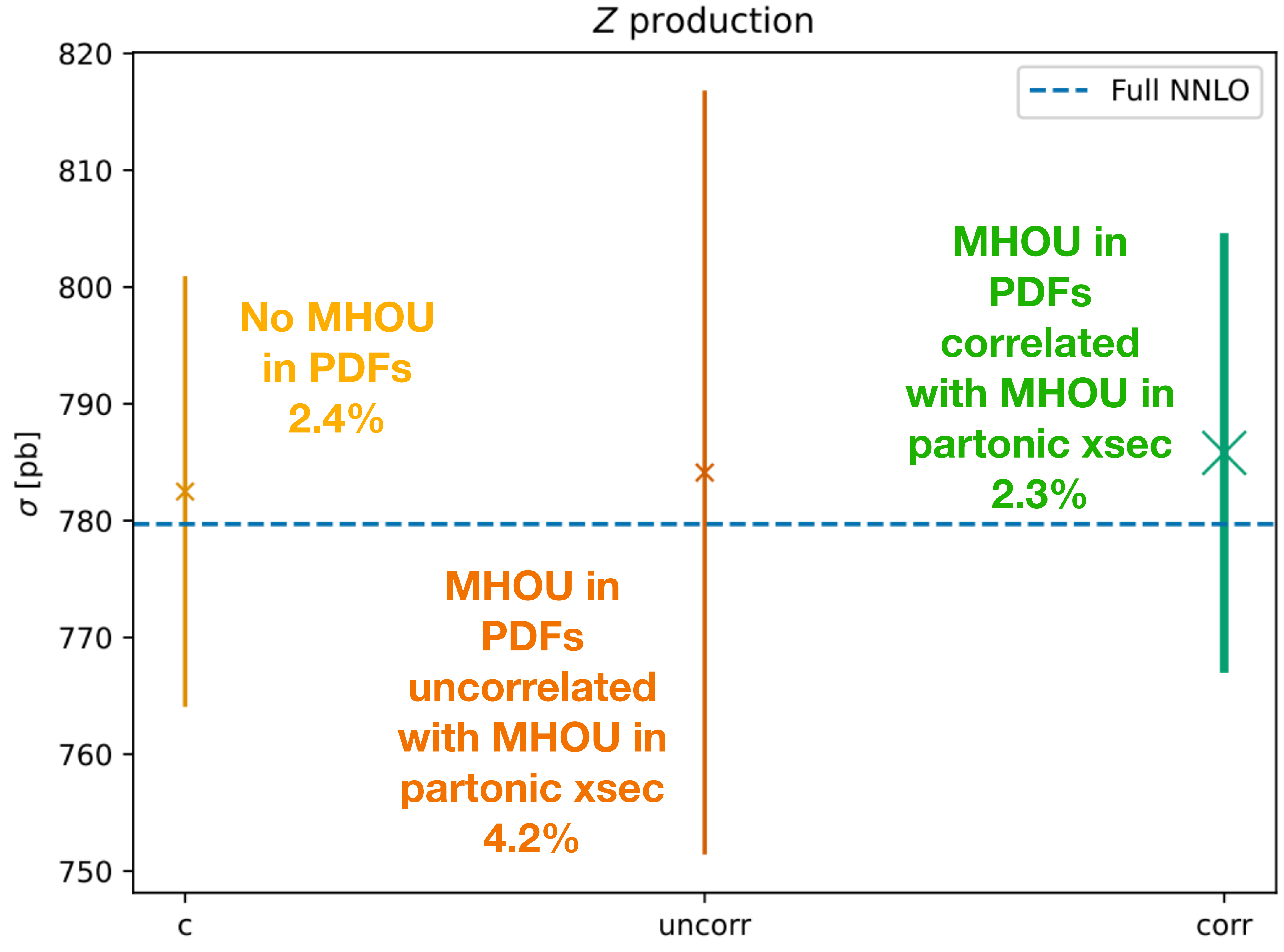
➡ Missing correlation between scale variation in PDF fits and hard cross sections [Harland-Lang, Thorne Eur.Phys.J.C 79 (2019) 3, 225] [Ball, Pearson arXiv:2105.05114]

#2 – MISSING HIGHER ORDER UNCERTAINTIES IN PDF FITS

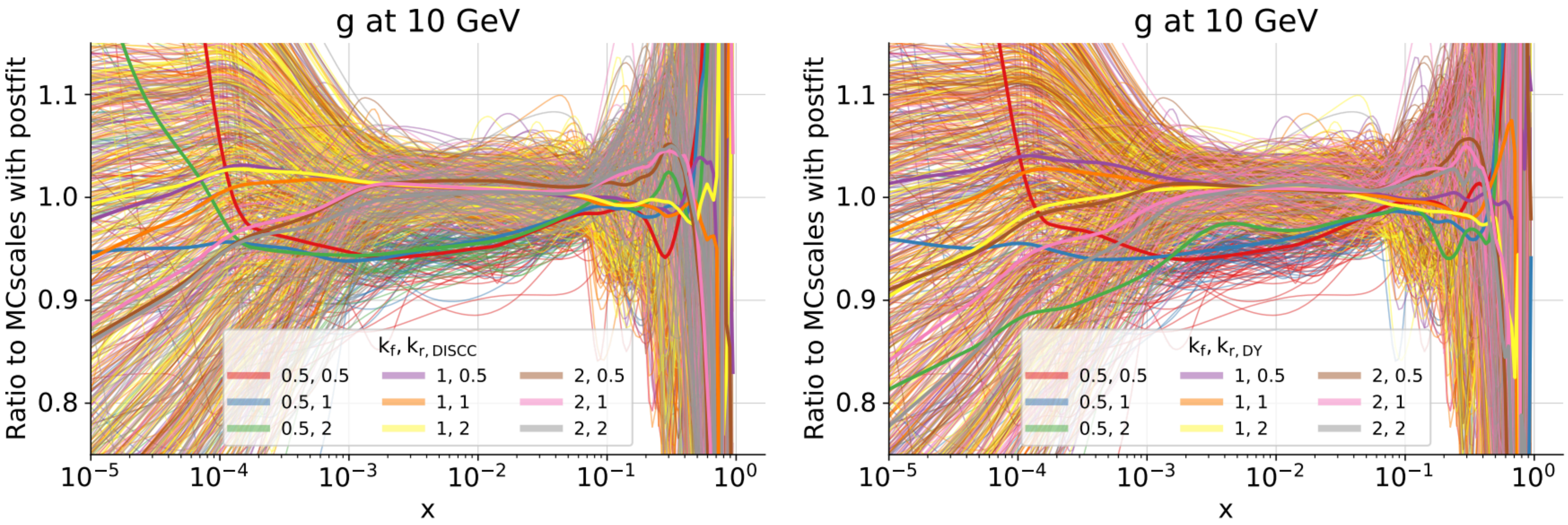
- ➡ Alternatively, MCscales [Kassabov et al, 2207.07616]: renormalisation and factorisation scales treated as free parameters of the fixed-order theory, that induce an uncertainty on the theory predictions included in a PDF fit & need to be propagated.
- ➡ Joint sampling of experimental uncertainty (propagated to PDF uncertainty by MC sampling) by specifying a suitable prior probability distribution of all possible scale choices & a-posteriori criterion based on agreement with the data.
- ➡ Can compute full PDF+SCALE uncertainty in cross sections by matching the scales in the hard cross section computation with the scales in the MCscale PDF set: correlation fully taken into account

Kassabov et al: 2207.07616]





THE MCSCALES APPROACH



- ✓ Main idea of MCscales: the renormalisation and factorisation scales are free parameters of the fixed-order theory, that induce an uncertainty on the theory predictions included in a PDF fit & need to be propagated
- ✓ Joint sampling of experimental uncertainty (propagated to PDF uncertainty by MC sampling) by specifying a suitable prior probability distribution of all possible scale choices & a-posteriori criterion based on agreement with the data.

$$P\left(k_f = \xi_f, k_{r_1} = \xi_1, \ldots, k_{r_{N_\mathrm{p}}} = \xi_{N_\mathrm{p}}\right) = P(\omega)$$

with $\omega \in \Omega = \{(\xi_f, \xi_1, \ldots, \xi_{N_\mathrm{p}})\ \forall\ \xi_f, \xi_1, \ldots, \xi_{N_\mathrm{p}} \in \Xi\}$

3^{1+N_p} elements, with N_p = 5, p=DIS NC, DIS CC, DY, JET, TOP

Choose prior = choose P($\boldsymbol{\omega}$)

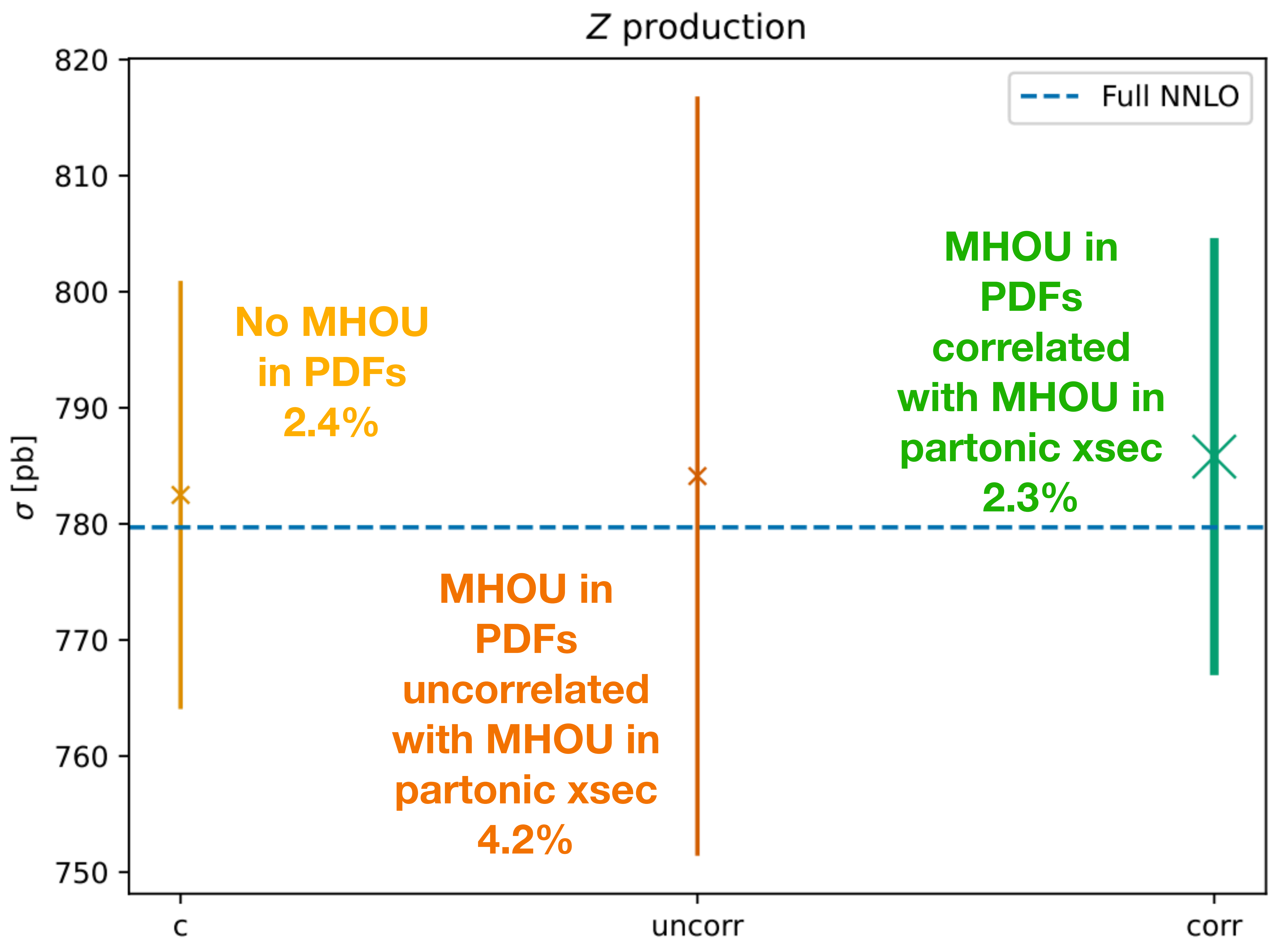
Posterior

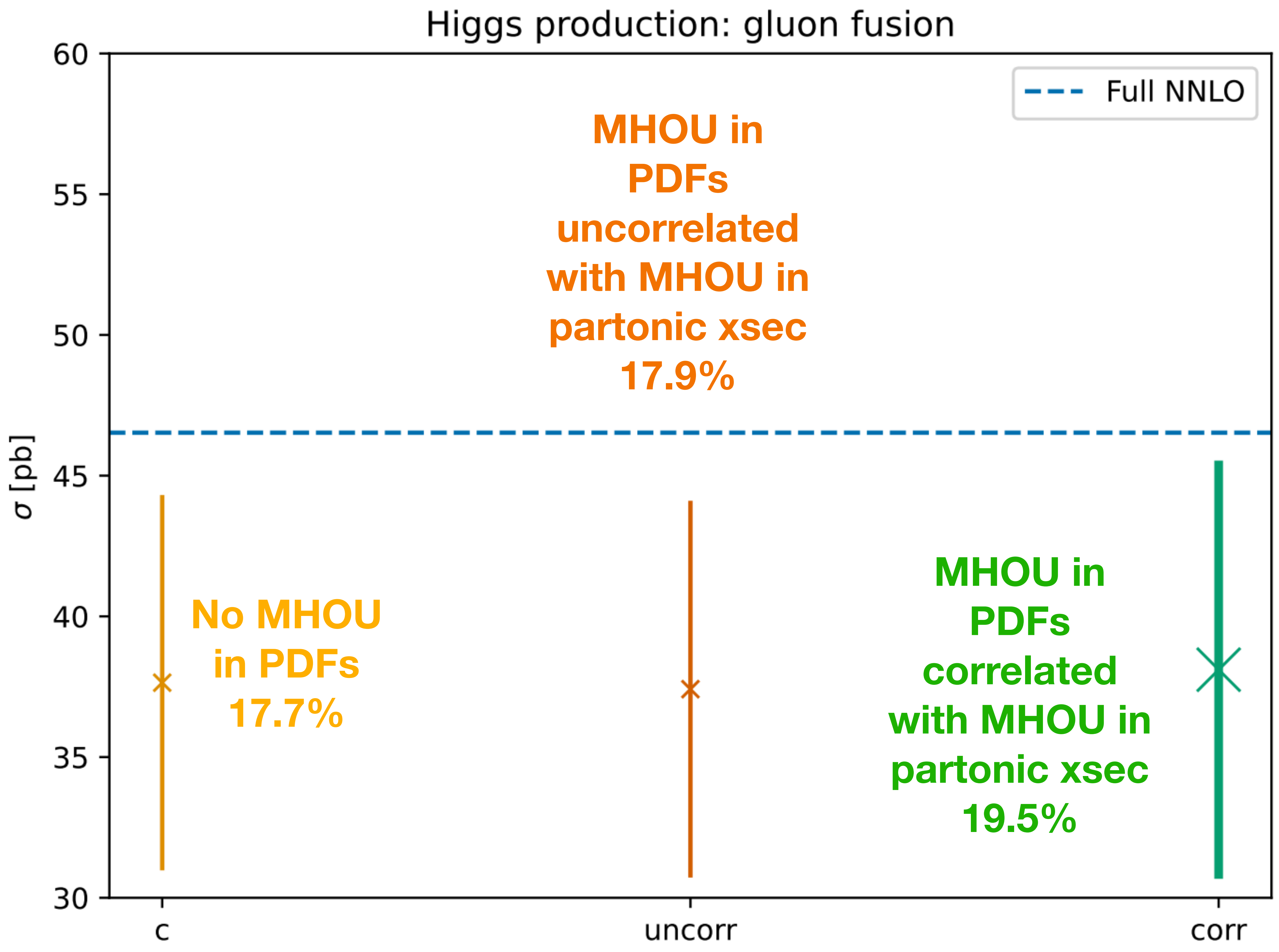
$$\chi^2_n > \langle\chi^2\rangle_{n\,|\,\omega^{(n)}=\{1,\ldots,1\}} + 4\,\mathrm{std}(\chi^2)_{n\,|\,\omega^{(n)}=\{1,\ldots,1\}}$$

THE MCSCALES APPROACH

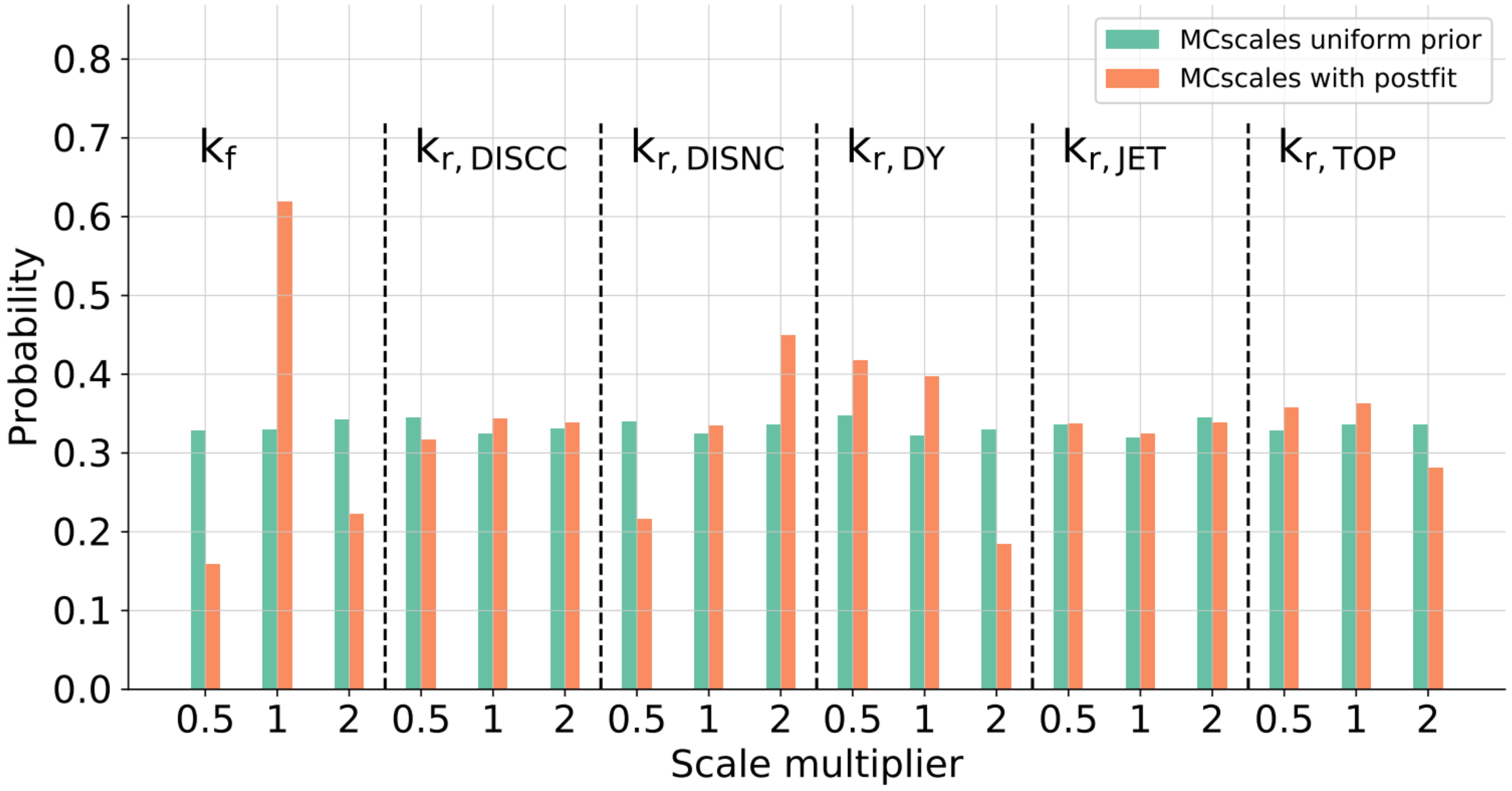
√Can compute full PDF+SCALE uncertainty in cross sections at NLO by matching the scales in the hard cross section computation with the scales in the MCscale PDF set: correlation fully taken into account

$$\left\{\sigma_n = \hat{\sigma}_p(k_f^{(n)}, k_{r_p}^{(n)}) \otimes f_n(k_f^{(n)}, k_{r_p}^{(n)}) \ \ \forall n = 1, \ldots, N\right\}$$





THE MCSCALES APPROACH

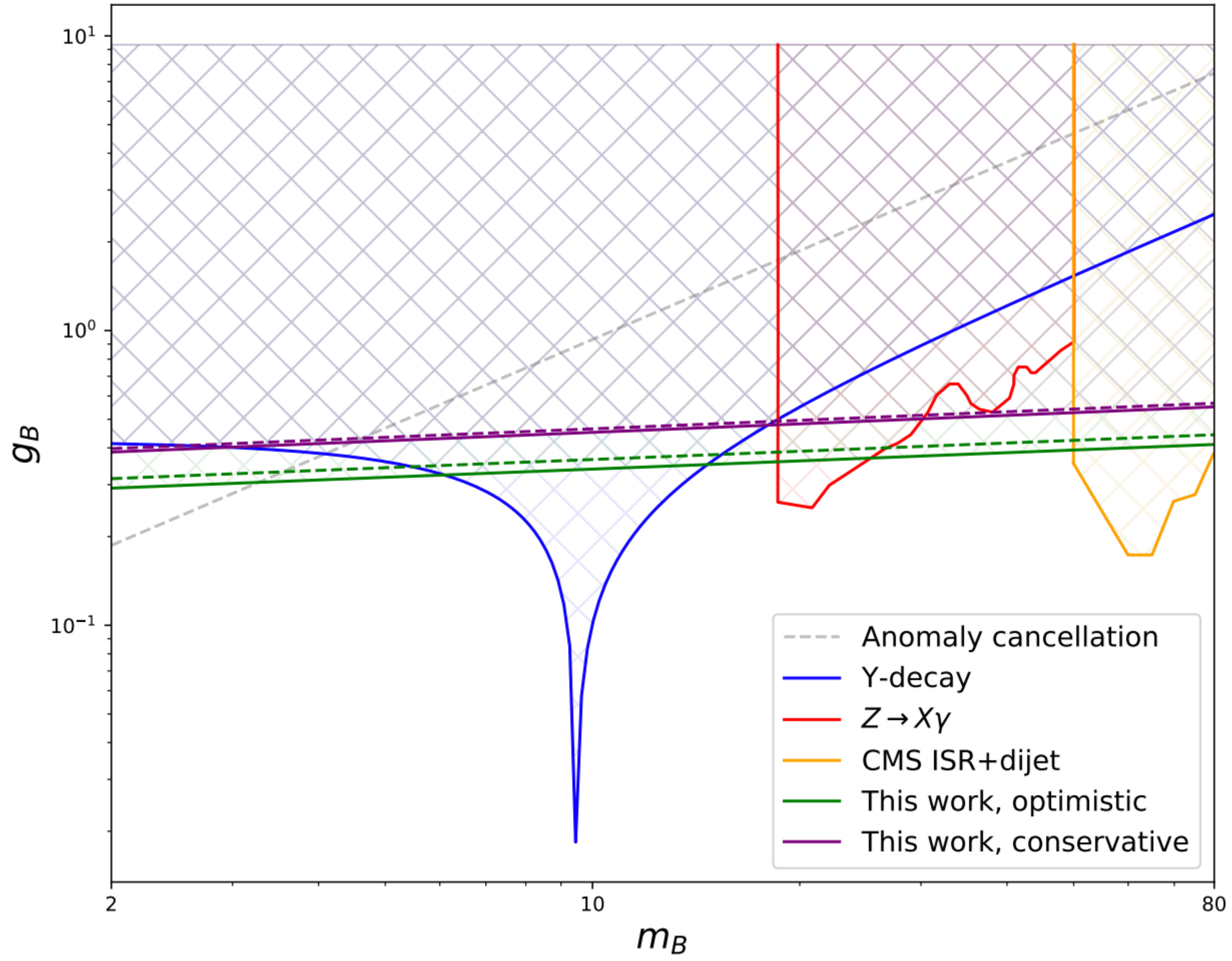


- ✓ Can look at the distribution of each of the scales over replicas.
- ✓ Flat distribution for the MCscales uniform prior.
- ✓ After applying postfit observe preference for central factorisation scale.
- ✓ Each process affected in a different way.

Scale multipliers	Process	Preferred values
(k_f, k_r)	DIS CC	$(1,1)$
	DIS NC	$(1,2)$
	DY	$(1,1)$
	Jets	$(1,\frac{1}{2})$
	Top	$(1,1)$

DARK PHOTON

M. McCullough, J. Moore, MU, arXiv:2203.12628



- If there was a lepto-phobic dark photon weakly coupled to quarks, it would appear among the partons of the proton.

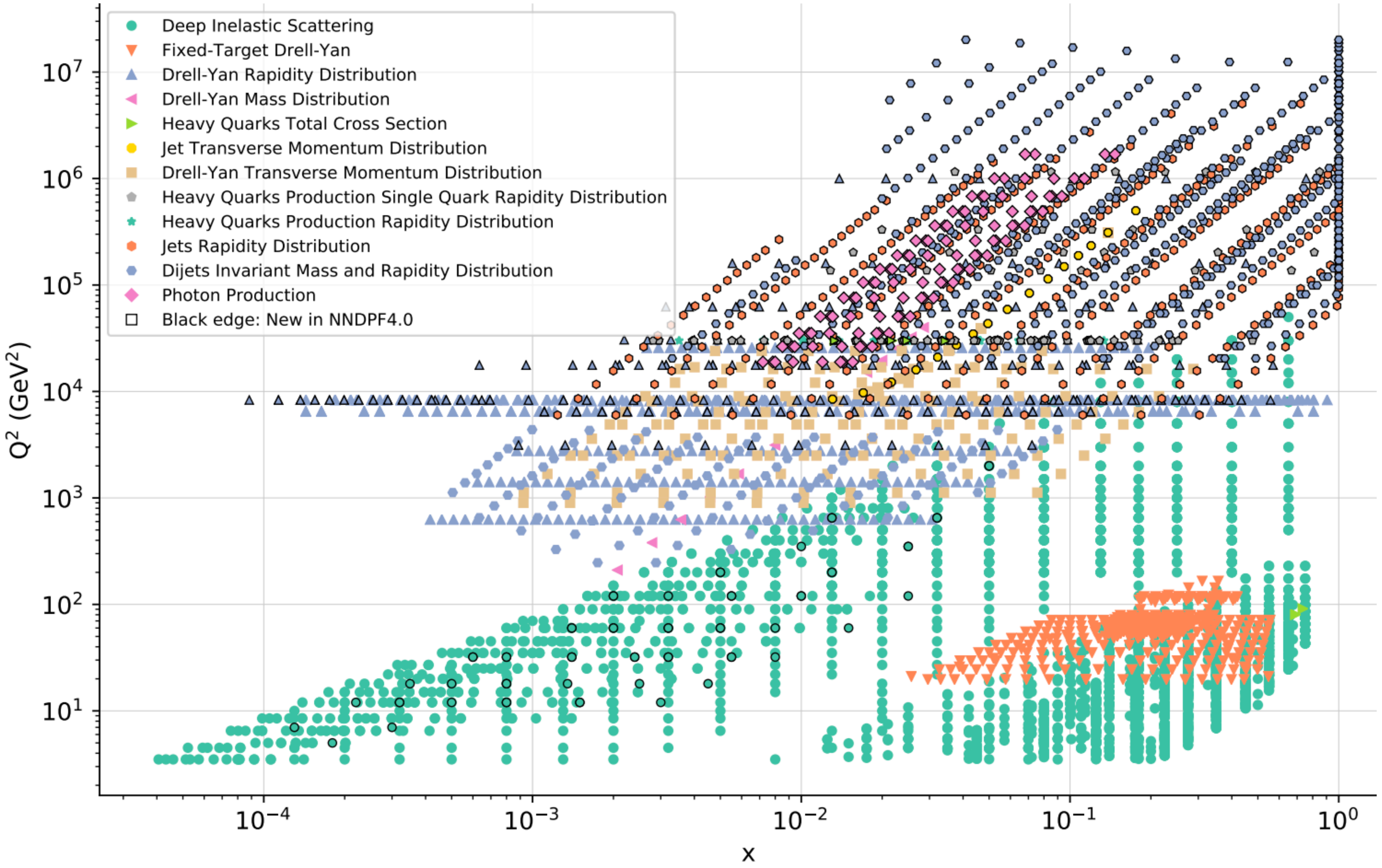
$$\mathcal{L}_{\rm int} = \frac{1}{3} g_B \bar{q} \not{B} q$$

- The presence of the dark Parton would modify the evolution of standard quarks and gluon.

- Precise LHC data can indirectly constrain parameter space of the dark photon in a competitive way compared to direct searches

#3 – PDF AND NEW PHYSICS INTERPLAY

$$\chi^2 = \frac{1}{N_{\rm dat}} \sum_{i=1}^{N_{\rm dat}} (T_i(\{\theta\},\{c\}) - D_i)\, {\rm cov}^{-1}_{ij}\, (T_j(\{\theta\},\{c\}) - D_j)$$



NNPDF4.0, Ball et al, arXiv: 2109.02653

$$T_i(\{\theta\},\{c\}) = {\rm PDFs}(\{\theta\},\{c\}) \otimes \hat{\sigma}_i(\{c\})$$

(B)SM parameters: $\alpha_s(M_Z)$, M_W, θ_W ,SMEFT WCs…..

Parameters determining PDFs at initial scale

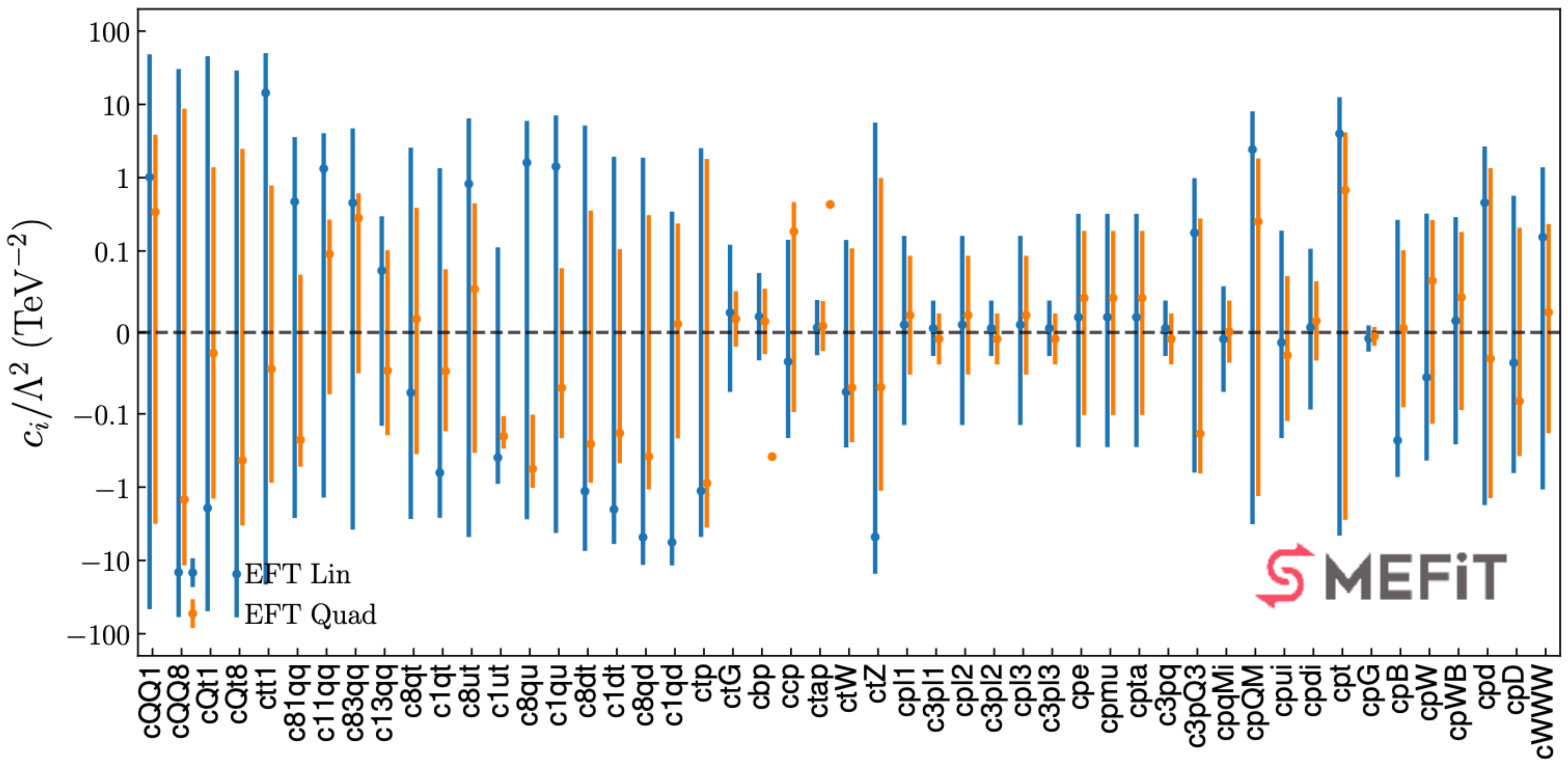
✓ In a PDF fit typically

$$T_i(\{\theta\}) = {\rm PDFs}(\{\theta\},\{c=\bar{c}\}) \otimes \hat{\sigma}_i(\{c=\bar{c}\})$$

#3 – PDF AND NEW PHYSICS INTERPLAY

$$\chi^2 = \frac{1}{N_{\rm dat}} \sum_{i=1}^{N_{\rm dat}} (T_i(\{\theta\},\{c\}) - D_i)\, {\rm cov}^{-1}_{ij}\, (T_j(\{\theta\},\{c\}) - D_j)$$

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i \frac{c_i}{\Lambda^2} \mathcal{O}_i^{(6)} + \cdots$$



$$T_i(\{\theta\},\{c\}) = {\rm PDFs}(\{\theta\},\{c\}) \otimes \hat{\sigma}_i(\{c\})$$

(B)SM parameters: $\alpha_s(M_z)$, M_w, θ_w ,SMEFT WCs…..

Parameters determining PDFs at initial scale

✓ In a PDF fit typically

$$T_i(\{\theta\}) = {\rm PDFs}(\{\theta\},\{c=\bar{c}\}) \otimes \hat{\sigma}_i(\{c=\bar{c}\})$$

✓ In a fit of (B)SM parameters

$$T_i(\{c\}) = {\rm PDFs}(\{\bar{\theta}\},\{\bar{c}\}) \otimes \hat{\sigma}_i(\{c\})$$

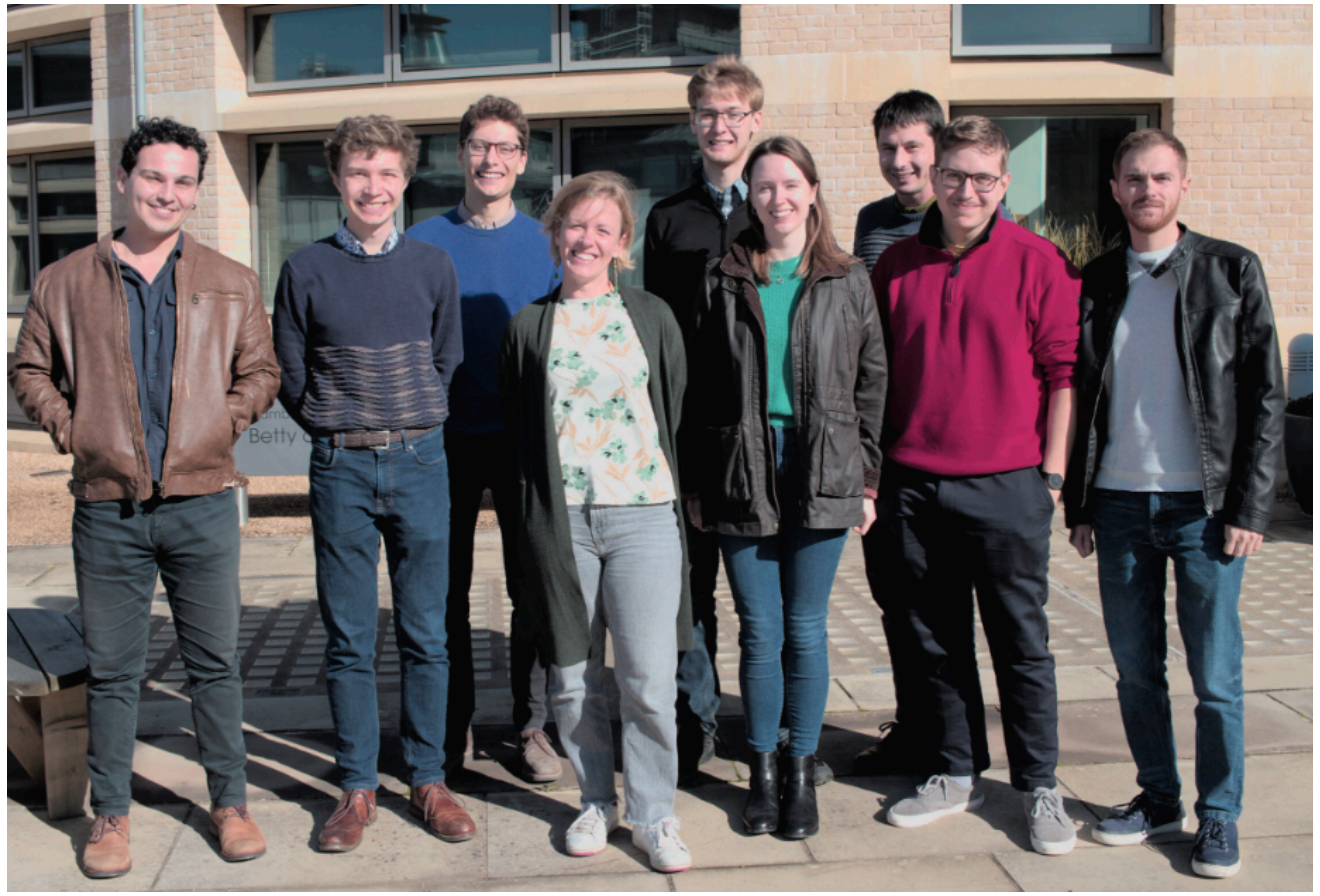
Ethier et al, arXiv: 2105.00006

#3 - PDF AND NEW PHYSICS INTERPLAY

- From the point of view of PDF fits:
  - ➡ How to make sure that new physics effects are not inadvertently fitted away in a PDF fit?
- From the point of view of SMEFT fits:
  - ➡ Should I make sure I am using a clean set of PDFs in a SMEFT analysis? How to define it? Is it enough?
  - ➡ How would the bounds change if I was consistently using PDFs that include in the fit the same operators that I am fitting?

PBSP

MU + Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales Alvarado, E. Hammou, M. Costantini

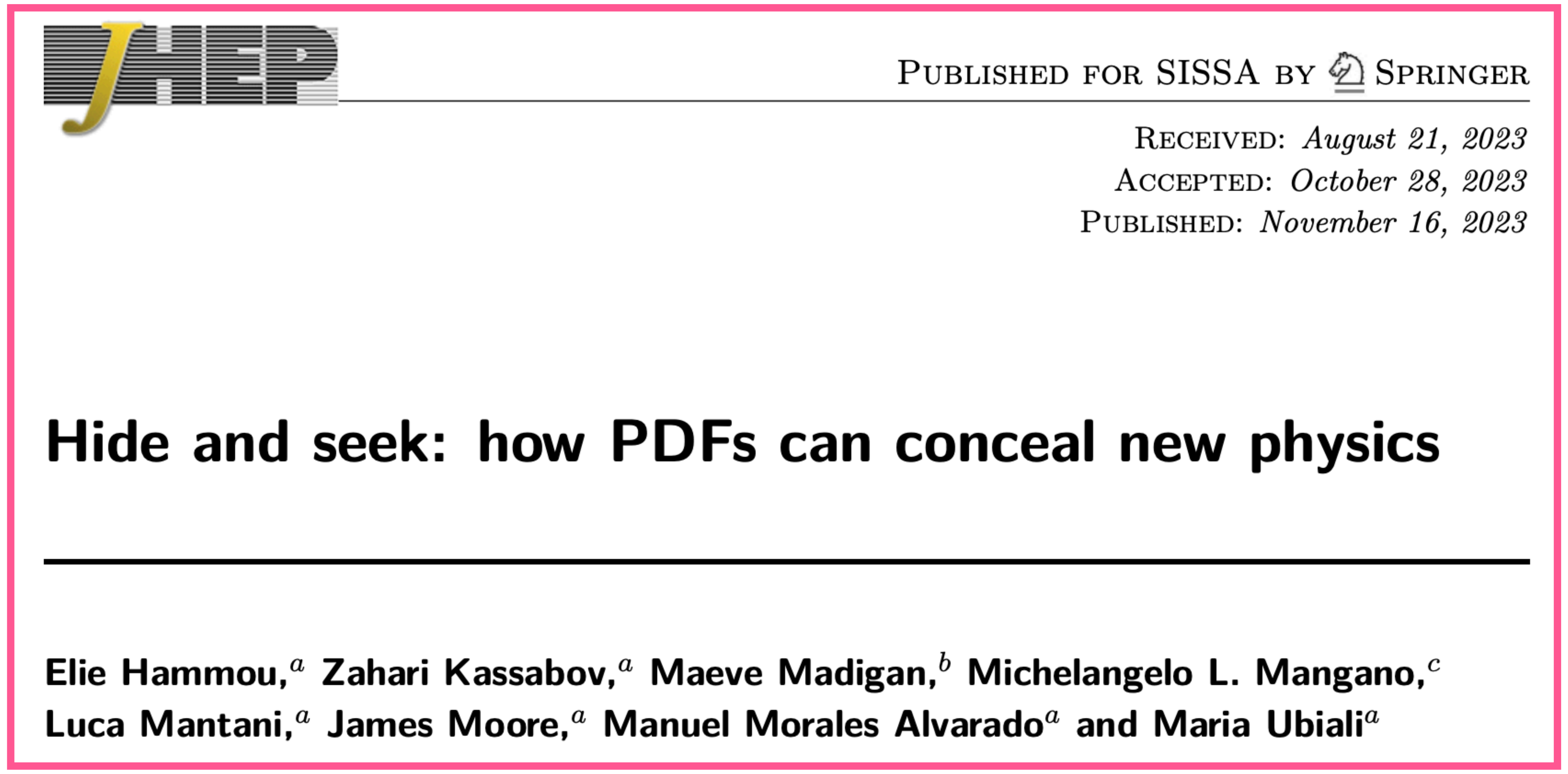


#3 – PDF AND NEW PHYSICS INTERPLAY

- From the point of view of PDF fits:
  - ➡ How to make sure that new physics effects are not inadvertently fitted away in a PDF fit?

✦ With current experiments, at the HL-LHC found a model involving a new heavy W′ that would affect high energy Drell-Yan tails. If there was a new heavy W′ PDF fits would have same quality but PDFs (hence theory predictions at LHC) would be significantly different. Large-x PDFs would absorb New Physics in this case!

✦ Need orthogonal input from low-energy experiments!

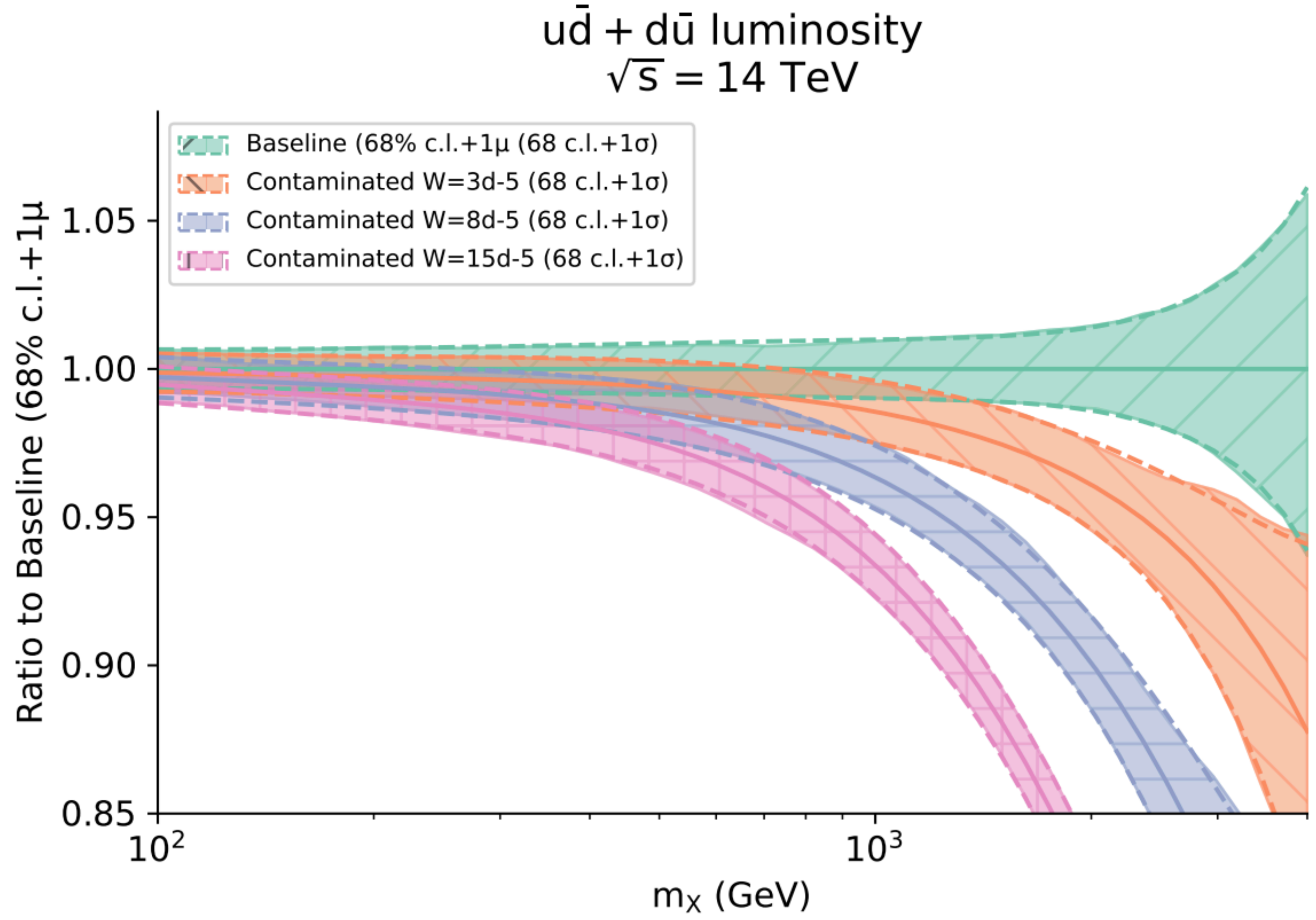

JHEP

Published for SISSA by Springer

Received: *August 21, 2023*
Accepted: *October 28, 2023*
Published: *November 16, 2023*

Hide and seek: how PDFs can conceal new physics

Elie Hammou,[a] Zahari Kassabov,[a] Maeve Madigan,[b] Michelangelo L. Mangano,[c] Luca Mantani,[a] James Moore,[a] Manuel Morales Alvarado[a] and Maria Ubiali[a]



#3 – PDF AND NEW PHYSICS INTERPLAY

- From the point of view of PDF fits:
  - ➡ How to make sure that new physics effects are not inadvertently fitted away in a PDF fit?
- From the point of view of SMEFT fits:
  - ➡ Should I make sure I am using a clean set of PDFs in a SMEFT analysis? How to define it? Is it enough?
  - ➡ How would the bounds change if I was consistently using PDFs that include in the fit the same operators that I am fitting?

$$d\sigma^{pp\to ab} = \sum_{i,j} f_i \otimes f_j \otimes d\hat{\sigma}^{ij\to ab} + \dots$$

Simultaneous fits of SMEFT and PDFs

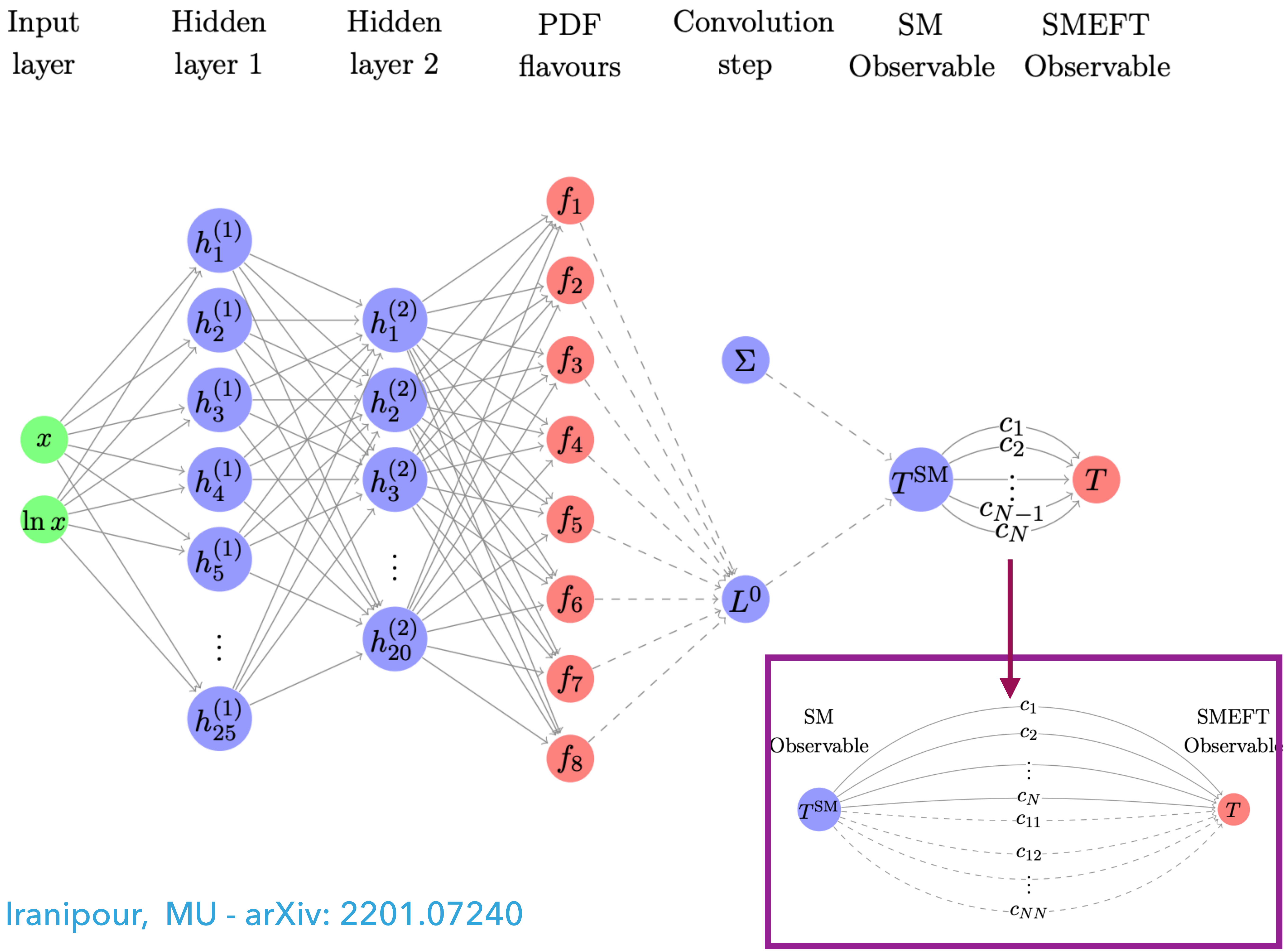
$$T(\{\theta_k\},\{c_i\})$$

$$f(\{\theta_k\})$$

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \cdots$$

#3 – PDF AND NEW PHYSICS INTERPLAY

- How to perform simultaneous fits?
- SimuNET [soon available open-source] yields simultaneous fit of PDFs and SMEFT coefficients, it does not have limit in number of parameters that can be fitted alongside PDFs at the initial scale!



S. Iranipour, MU - arXiv: 2201.07240

Linear dim-6 operator

$$T(\hat{\theta}) = \Sigma(\{c_n\}) \cdot L^0(\theta) = T^{\rm SM}(\theta) \cdot \left(1 + \sum_{n=1}^{N} c_n R^{(n)}_{\rm SMEFT}\right)$$

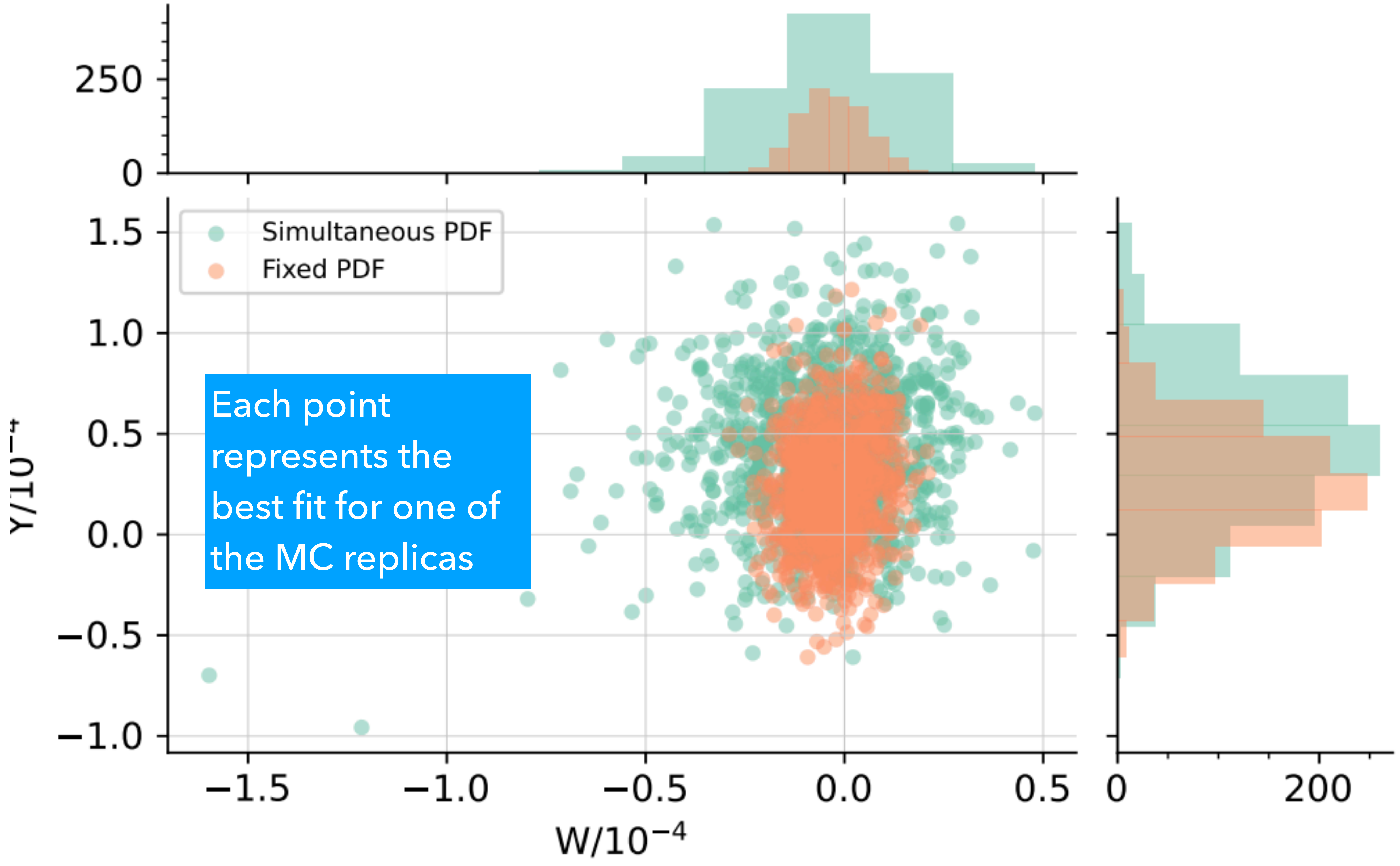
$$T^{\rm SM}(\theta) = \Sigma^{\rm SM} \cdot L^0(\theta)$$

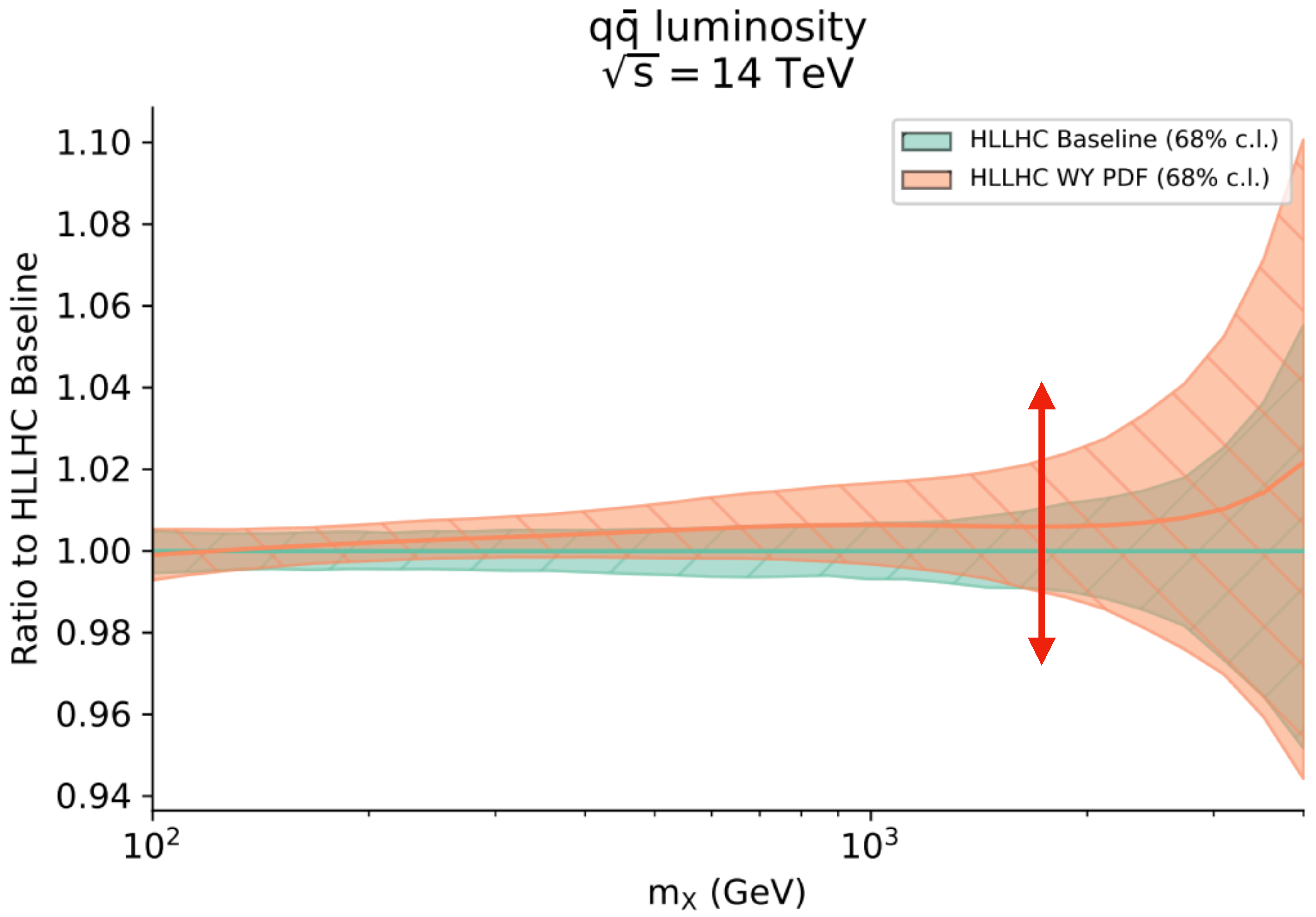
Quadratic dim-6 operator

$$T(\hat{\theta}) = T^{\rm SM}(\theta) \cdot \left(1 + \sum_{n=1}^{N} c_n R^{(n)}_{\rm SMEFT} + \sum_{1 \leq n \leq m \leq N} c_{nm} R^{(n,m)}_{\rm SMEFT}\,,\right)$$

$$c_n c_m$$

#3 – PDF AND NEW PHYSICS INTERPLAY

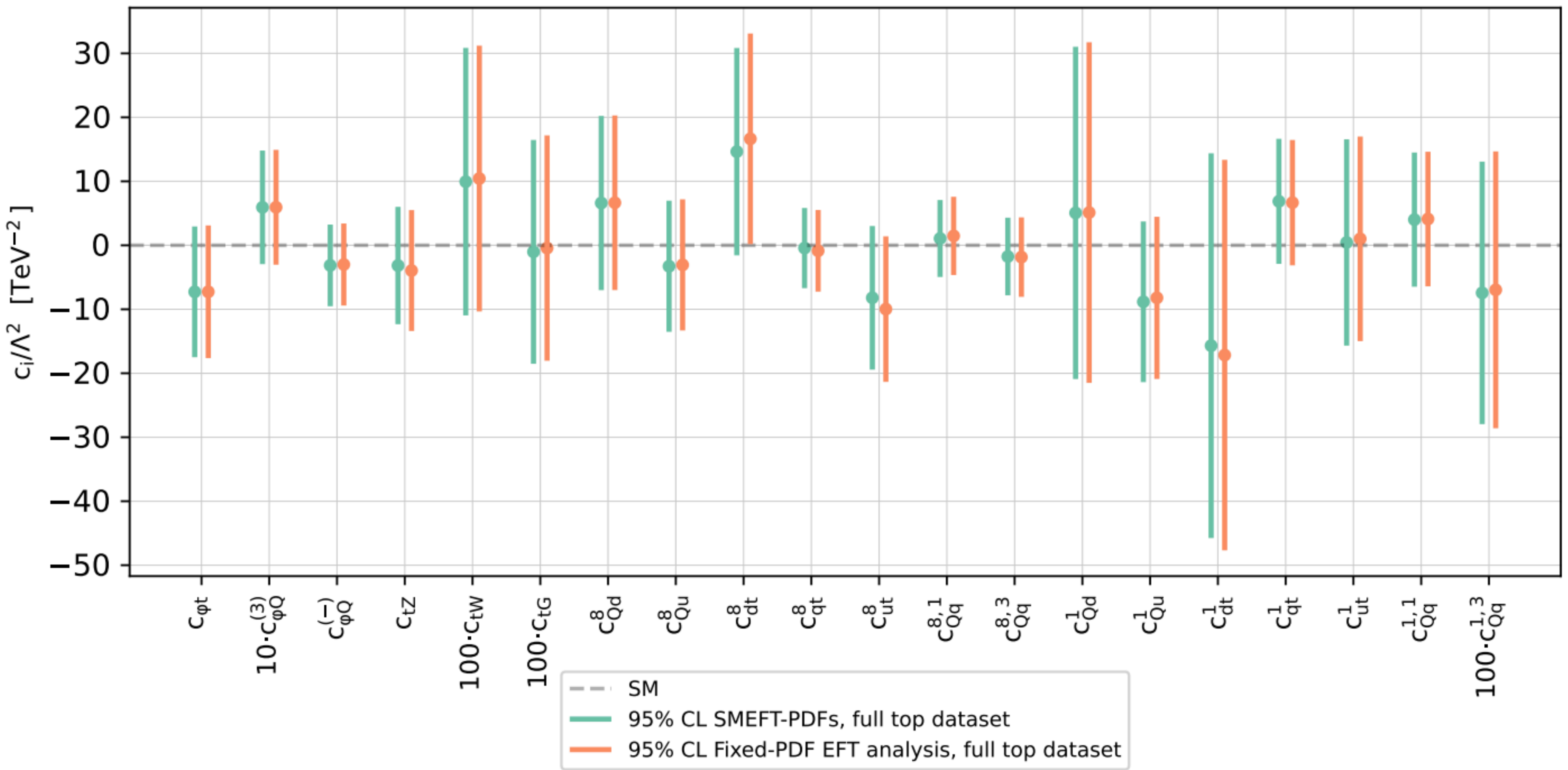


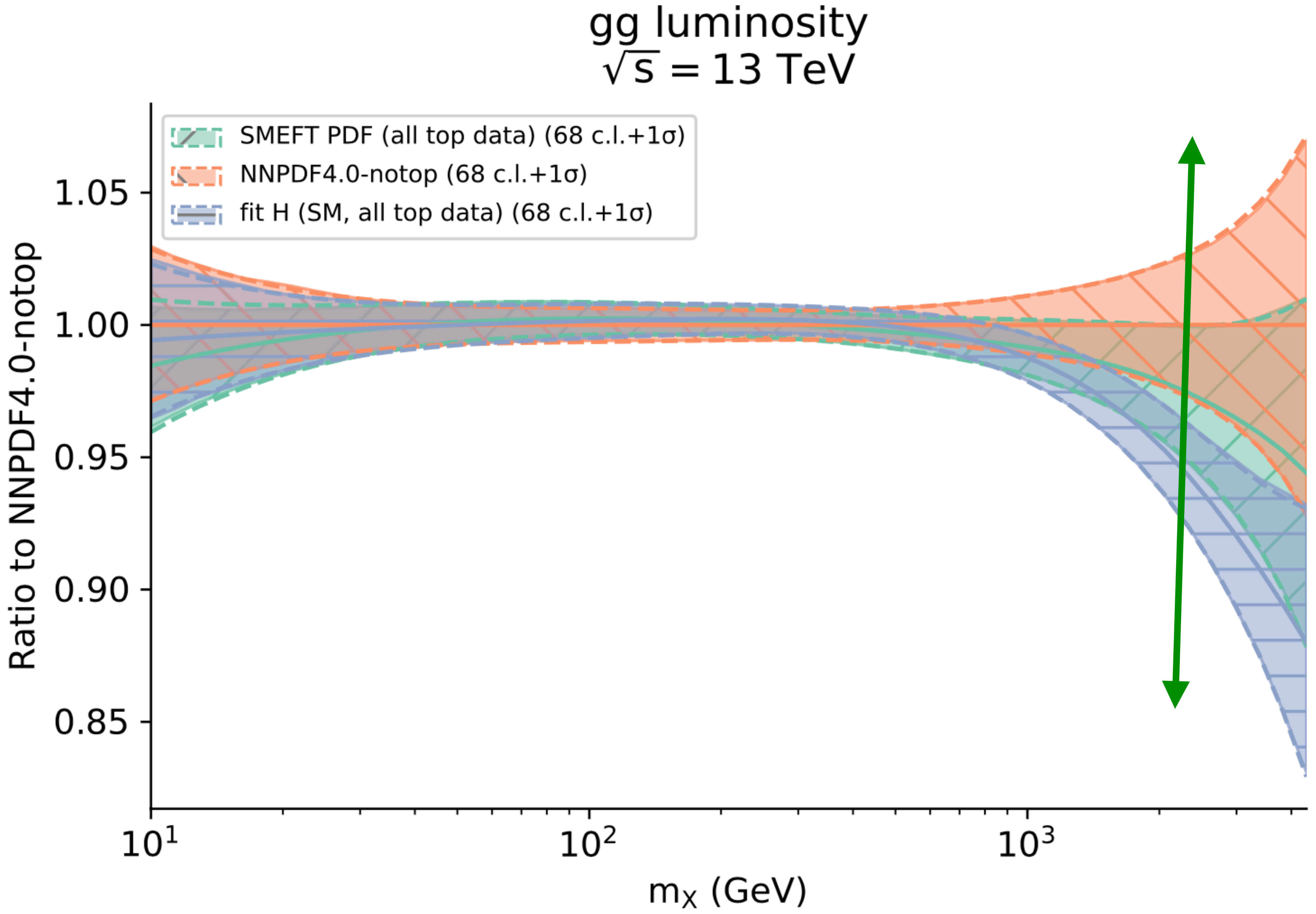


S. Iranipour, MU - arXiv: 2201.07240

✓ Simultaneous analysis of PDFs and **Drell-Yan sector** Wilson coefficient of DIS + DY (including HL-LHC projections) using simuNET method shows that at HL-LHC the effect of interplay becomes important as WCs bounds broaden and PDF uncertainties change significantly once SMEFT effects allowed in theory predictions entering PDF fit

#3 - PDF AND NEW PHYSICS INTERPLAY





Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU - arXiv: 2303.06159

✓ Simultaneous analysis of PDFs and more than 20 operators in the **top sector** using simuNET method shows that WCs are stable, while PDF uncertainties broaden and PDF fitted simultaneously alongside WCs site nicely between PDF-only fit without top data and PDF fit including all top data