Realizing the next-generation CTEQ-TEA PDFs

QCD@LHC 2024, Freiburg

Tim Hobbs

Argonne National Laboratory, USA

with CTEQ-TEA (Tung Et. Al.) working group

Asia: A. Ablat, S. Dulat, T.-J. Hou, I. Sitiwaldi

North America: A. Courtoy, M. Guzzi, TJH,
J. Huston, H.-W. Lin, P. Nadolsky, K. Xie, C.-P. Yuan

and other coauthors

recent preprints:

arXiv: 2408.04020

arXiv: 2408.11131

Argonne NATIONAL LABORATORY

CTEQ

toward a new generation of CT202X PDFs

1. Development, implementation, testing of H.O. theory: N3LO QCD, …
2. Multiple preliminary NNLO fits with LHC Run-2 (di)jet, vector boson, $t\bar{t}$ data
   - based on selected experiments recommended in 2305.10733, 2307.11153
3. Next-generation PDF uncertainty quantification: Bézier curves, META combination, ML PDF modeling, multi-Gaussian approaches, …
4. Physics applications
   a. Higgs+HQ production (DY); QED-corrected PDFs and EW precision; UHE neutrinos
   b. PDF dependence of forward-backward asymmetry
   c. An L2 sensitivity study using xFitter
   d. Pion PDFs
   e. …

• RESEARCH PROJECTS AND RESULTS •

https://cteq-tea.gitlab.io/

- CTEQ-TEA publications from INSPIRE
- LHAPDF grids for parton distributions
  - CT18 (N)NLO, CT18 QED, CT18 FC, …
  - Subtracted heavy-quark PDFs in the S-ACOT-MPS scheme
- Public codes
  - ePump (Hessian updating for PDFs with tolerance > 1)
  - LHAexplorer (fast surveys of data using L2 sensitivities)
  - Fantômas (Bezier parametrizations)
  - mp4lhc/mcgen (MC PDFs, combination of PDFs)
  - …

CT18up: enhanced precision LHAPDF grids (2023)

on https://cteq-tea.gitlab.io/project/00pdfs/

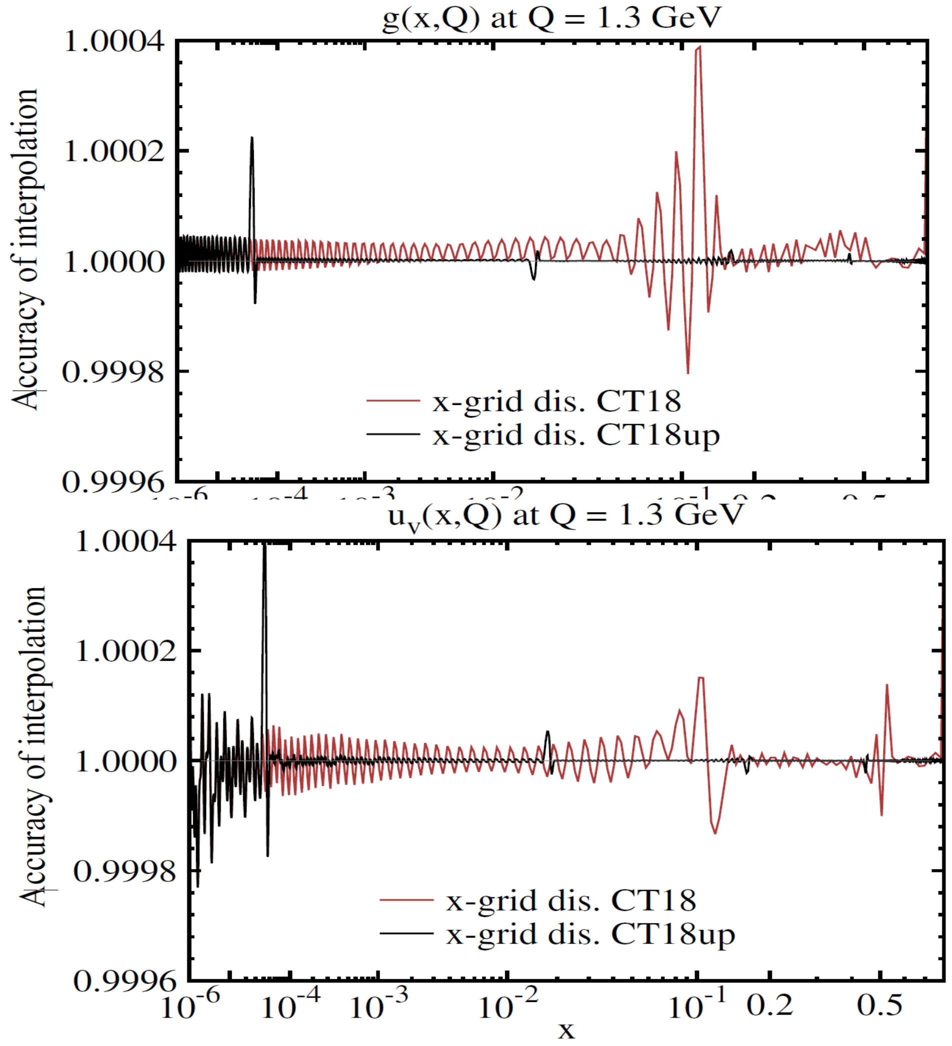
arXiv: 2408.04020

- CT18, A, X, Z NNLO PDFs (2019 edition) presented as LHAPDF grids with a 1.9x higher number of x and Q nodes
- same PDFs as in LHAPDF library; even more precise interpolation at $10^{-4} \leq x \leq 1$
- recommended for high-mass and precision calculations; 2019 grids ok in other cases

Numbers of x, Q nodes in LHAPDF grids

intervals in Q	CT18	CT18up
$[Q_0, m_c]$	2	4
$[m_c, m_b]$	8	11
$[m_b, m_t]$	14	18
$[m_t, Q_{\max}]$	13	16
Total	37	49

intervals in x	CT18	CT18up	intervals in x	CT18	CT18up
$[10^{-10}, 10^{-9}]$	1	1	$[0.1, 0.2]$	7	18
$[10^{-9}, 10^{-8}]$	11	11	$[0.2, 0.3]$	6	16
$[10^{-8}, 10^{-7}]$	12	12	$[0.3, 0.4]$	5	12
$[10^{-7}, 10^{-6}]$	11	11	$[0.4, 0.5]$	3	13
$[10^{-6}, 10^{-5}]$	12	12	$[0.5, 0.6]$	6	15
$[10^{-5}, 10^{-4}]$	11	15	$[0.6, 0.7]$	6	12
$[10^{-4}, 10^{-3}]$	12	23	$[0.7, 0.8]$	8	11
$[10^{-3}, 10^{-2}]$	11	23	$[0.8, 0.9]$	14	17
$[10^{-2}, 0.1]$	12	40	$[0.9, 1]$	15	38
			Total	161	300



N3LO PDF analyses: necessary pieces

Component		Availability
Splitting functions		Partial N3LO
Hard cross sections	• DIS, light flavors	Full N3LO
	• NC DIS, heavy flavors	Full N3LO (Blümlein et al.), not yet in fitting codes
	• Vector boson production	Full N3LO for some processes, fixed N3LO/NLO K-factor tables
	• CC DIS, jet, $t\bar{t}$ production	N2LO
	• $pp \to W + c$, $pp \to Z + b$, $\mathrm{pp} \to b$	NLO (massive); NNLO (ZM)

Looking forward to including all components **exactly and fully** to reduce the QCD scale uncertainty and guarantee the N3LO accuracy in the near future.

CTEQ-TEA and other groups include some N3LO contributions in their fitting codes: recent progress of MSHT and NNPDF in partial N3LO (aN3LO) fits

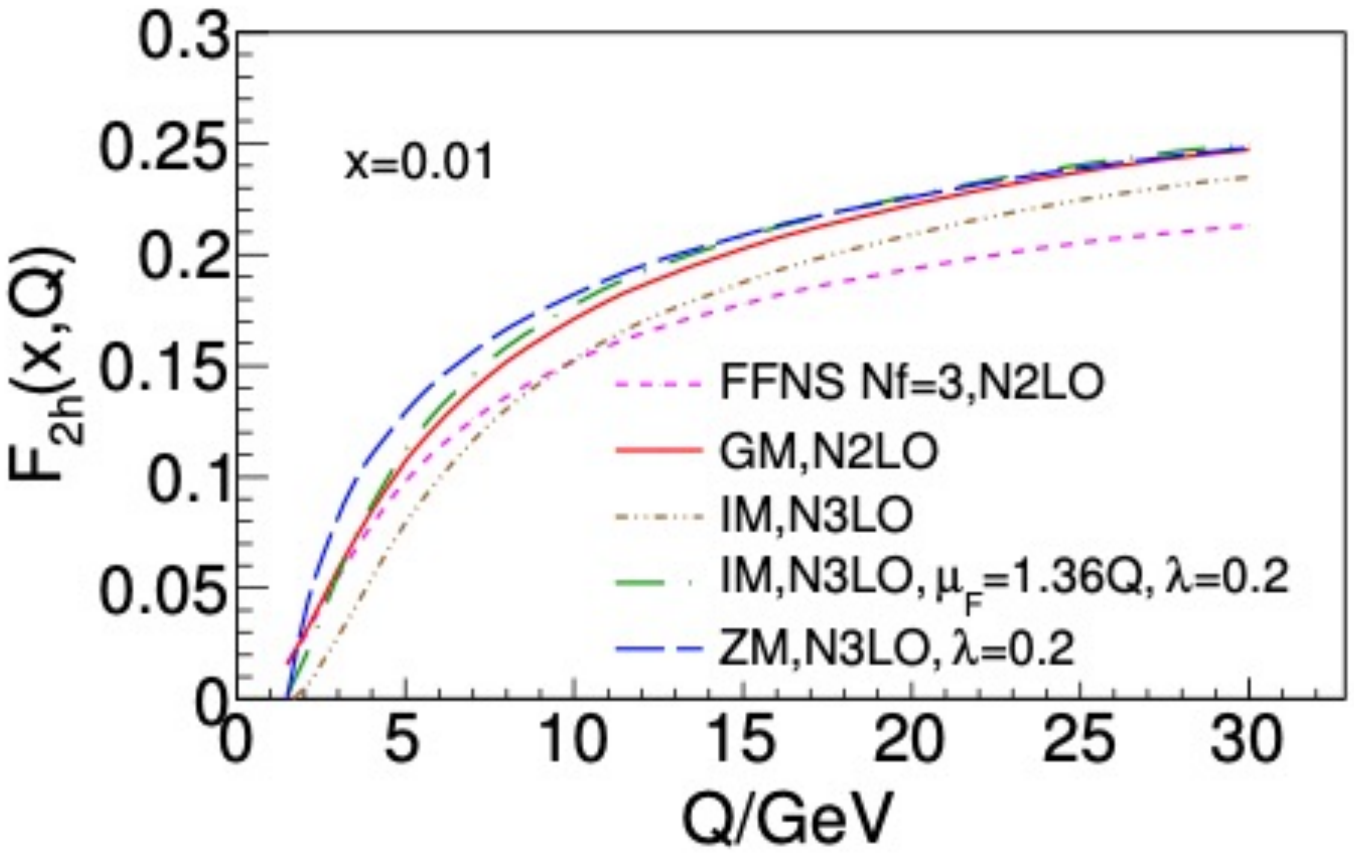
These partial N3LO calculations mostly agree with N2LO within their scale dependence

For $gg \to H^0$ production, the aN3LO-N2LO difference is comparable to other effects due to the remaining scale dependence, selection of experiments, treatment of systematic uncertainties

(cf. downward shift in central $gg \to H^0$ prediction at aN3LO for MSHT)

Work in progress

QCD cross sections @N3LO



relevant discussion: 2107.00460

(Gao, TJH, Nadolsky, Sun, Yuan)

- **DIS:** The CTEQ-TEA code implements complete flavor decompositions of DIS SFs at N3LO using approximate zero-mass Wilson coefficients with a rescaling variable (the **Intermediate-Mass VFN scheme**, cf. the figure)

Bowen Wang, Keping Xie PhD theses (SMU)

- **Imminent implementation** of massive N3LO heavy-quark coefficients to obtain N3LO DIS cross sections in **SACOT-MPS General-Mass VFN scheme**

Factorization schemes	Mass dependence in the FC terms	Mass dependence of the FE and subtraction terms	Introduce heavy-quark PDFs at large Q
FFN	Exact	N/A	no
ZM	None	None	yes
IM	Approximate	Approximate	yes
GM	Exact	Approximate	yes

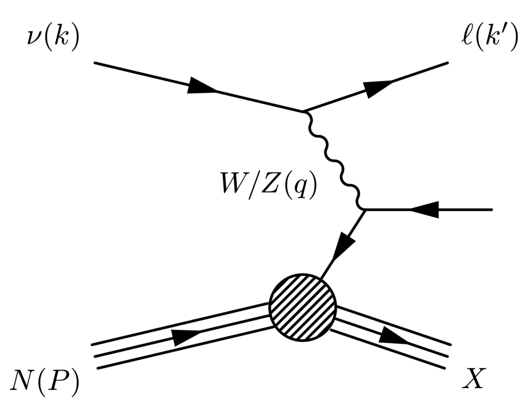
- **DGLAP evolution** is performed at N3LO with APFEL/APFEL++.
- **Drell-Yan:** forthcoming work to include N3LO DY effects using NNLO ApplFast + N3LO/N2LO K-factor tables

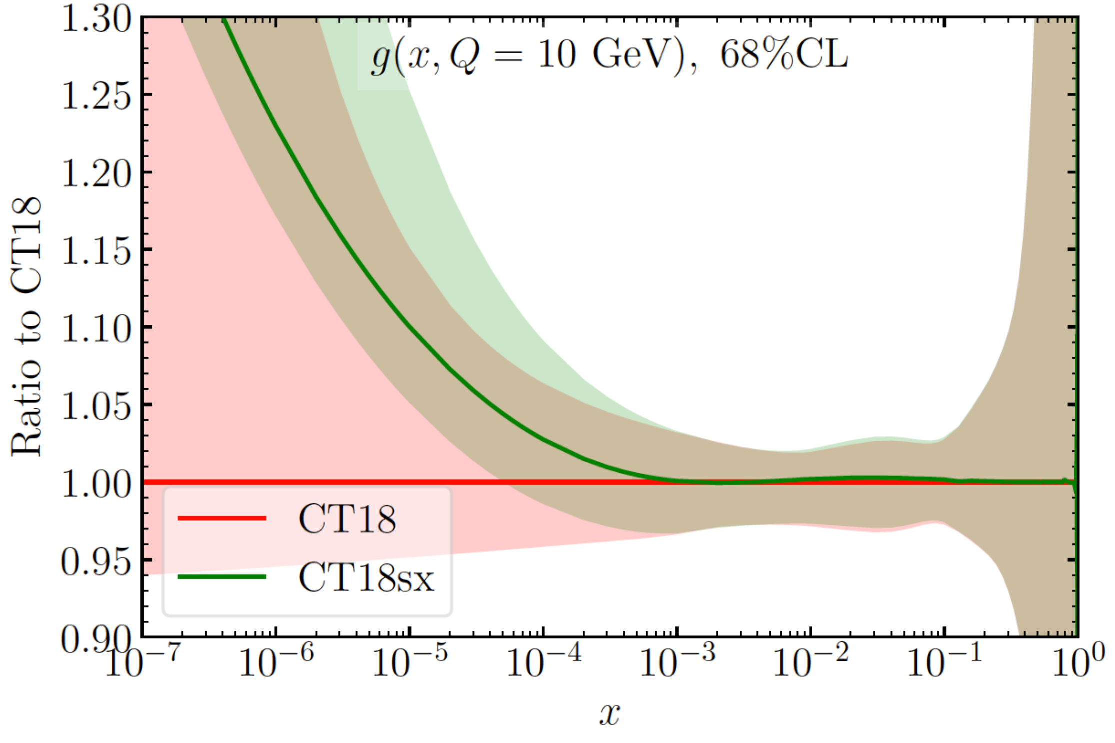
aside, predictions for ultra-high energy (UHE) neutrinos (DIS)

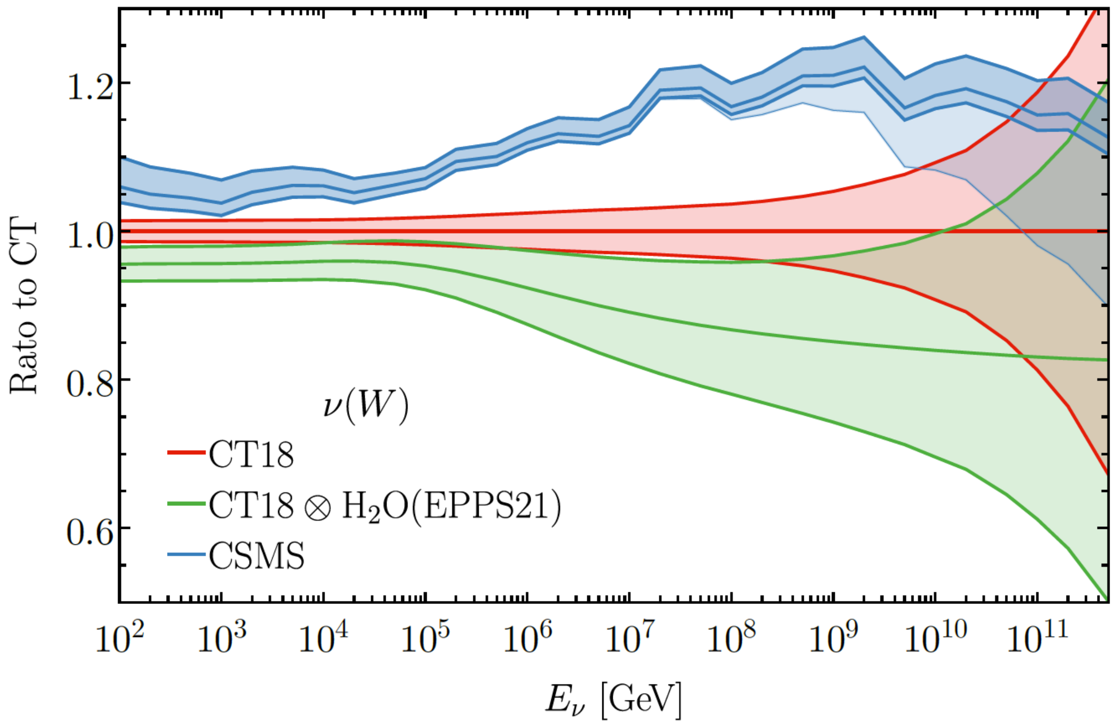
- just-published (June '24) CT predictions for UHE neutrino-nuclear scattering

[Xie, Gao, TJH, Stump, Yuan; PRD109, 113001]

$$\frac{d^2\sigma^{\nu(\bar{\nu})}}{dxdy} = \frac{G_F^2 m_N E_\nu}{\pi(1+Q^2/M_{W,Z}^2)^2}\left[\frac{y^2}{2}2xF_1 + \left(1 - y - \frac{m_N xy}{2E_\nu}\right)F_2 \pm y\left(1-\frac{y}{2}\right)xF_3\right]$$



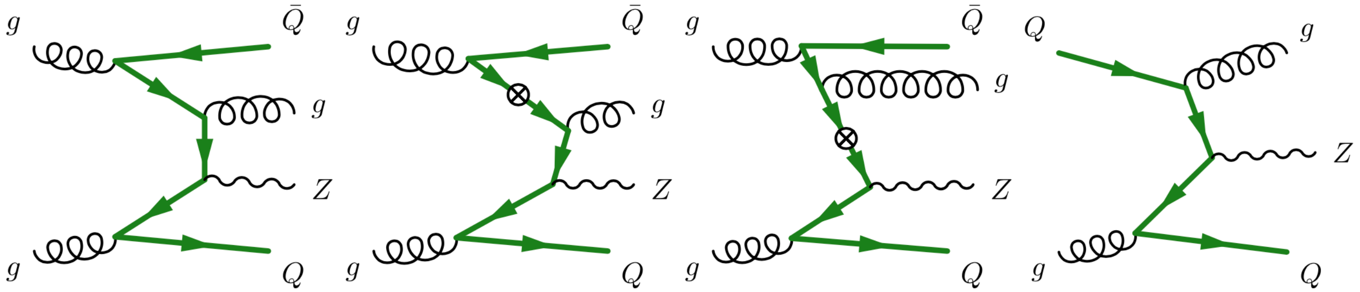




- explored N3LO (ZM) contributions; variations in HQ schemes and low-*x* resummation effects
- knowledge of perturbative treatment, low-x PDF uncertainties nuclear corrections relevant for UHE precision

A GMVFN scheme for Z production in association with a HQ at hadron colliders

M. Guzzi, P. Nadolsky, L. Reina, D. Wackeroth, K. Xie; (arXiv:2410.03876 --- Tuesday, Oct 8 [**today**!])

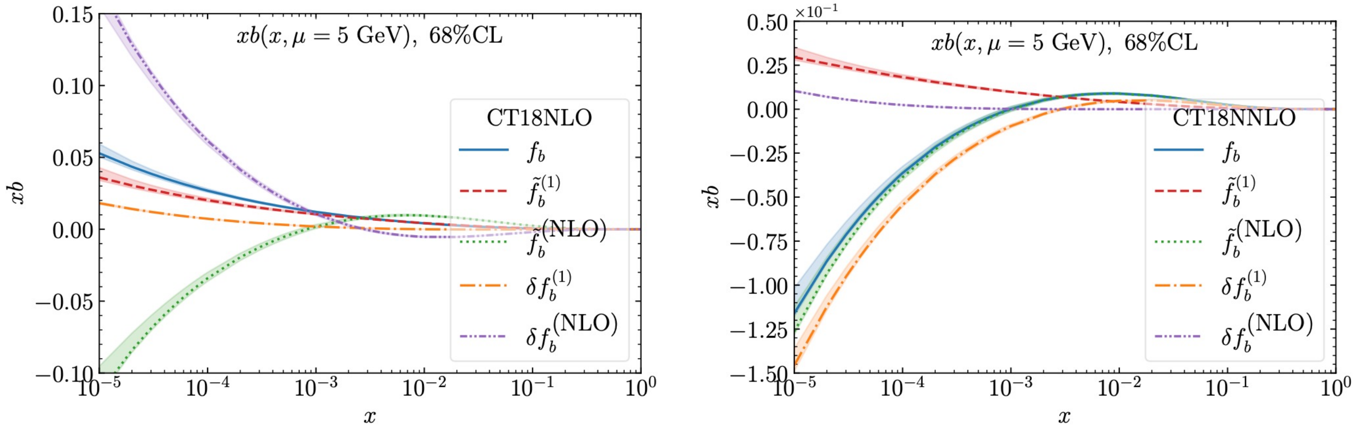


$$d\sigma_{\rm FC}^{\rm NLO} = f_g \triangleright \left[a_s^2\, d\hat{\sigma}^{(2)}_{gg\to ZQ\bar{Q}} + a_s^3\, d\hat{\sigma}^{(3)}_{gg\to ZQ\bar{Q}(g)}\right] \triangleleft f_g$$
$$+ \sum_{i=q,\bar{q}} f_i \triangleright \left[a_s^2\, d\hat{\sigma}^{(2)}_{q\bar{q}\to ZQ\bar{Q}} + a_s^3\, d\hat{\sigma}^{(3)}_{q\bar{q}\to ZQ\bar{Q}(g)}\right] \triangleleft f_{\bar{i}}$$
$$+ a_s^3 \sum_{i=q,\bar{q}} \left[f_g \triangleright d\hat{\sigma}^{(3)}_{gq\to ZQ\bar{Q}(q)} \triangleleft f_i + f_i \triangleright d\hat{\sigma}^{(3)}_{qg\to ZQ\bar{Q}(q)} \triangleleft f_g\right]$$

$$d\sigma_{\rm FE} - d\sigma_{\rm sub} = a_s(f_Q - \tilde{f}_Q^{(\rm NLO)}) \triangleright H^{(1)}_{Qg} \triangleleft g$$
$$+ a_s^2(f_Q - \tilde{f}_Q^{(1)}) \triangleright \left[H^{(2)}_{Qg} \triangleleft g + \sum_{i=q,\bar{q}} H^{(2)}_{Qq} \triangleleft f_i\right] + (\text{exch.})$$
$$= a_s\, \delta f_Q^{(\rm NLO)} \triangleright H^{(1)}_{Qg} \triangleleft g + a_s^2\, \delta f_Q^{(1)} \triangleright \left[H^{(2)}_{Qg} \triangleleft g + \sum_{i=q,\bar{q}} H^{(2)}_{Qq} \triangleleft f_i\right] + (\text{exch.}).$$

Methodology to streamline implementation of massive-quark radiative contributions in calculations with a variable number of active partons in proton-proton collisions.

It introduces subtraction and residual HQ PDFs to implement calculations in the ACOT/S-ACOT factorization schemes in various processes at high order in pQCD.



Subtraction PDFs $\tilde{f}_Q^{(1)} = a_s[A^{(1)}_{Qg} \triangleleft g], \;\; \tilde{f}_Q^{(2)} = a_s^2 \sum_{i=g,q,\bar{q}} [A^{(2)}_{Qi} \triangleleft f_i]$

$$\tilde{f}_Q^{(\rm NLO)}(x,\mu) \equiv \tilde{f}_Q^{(1)} + \tilde{f}_Q^{(2)}$$

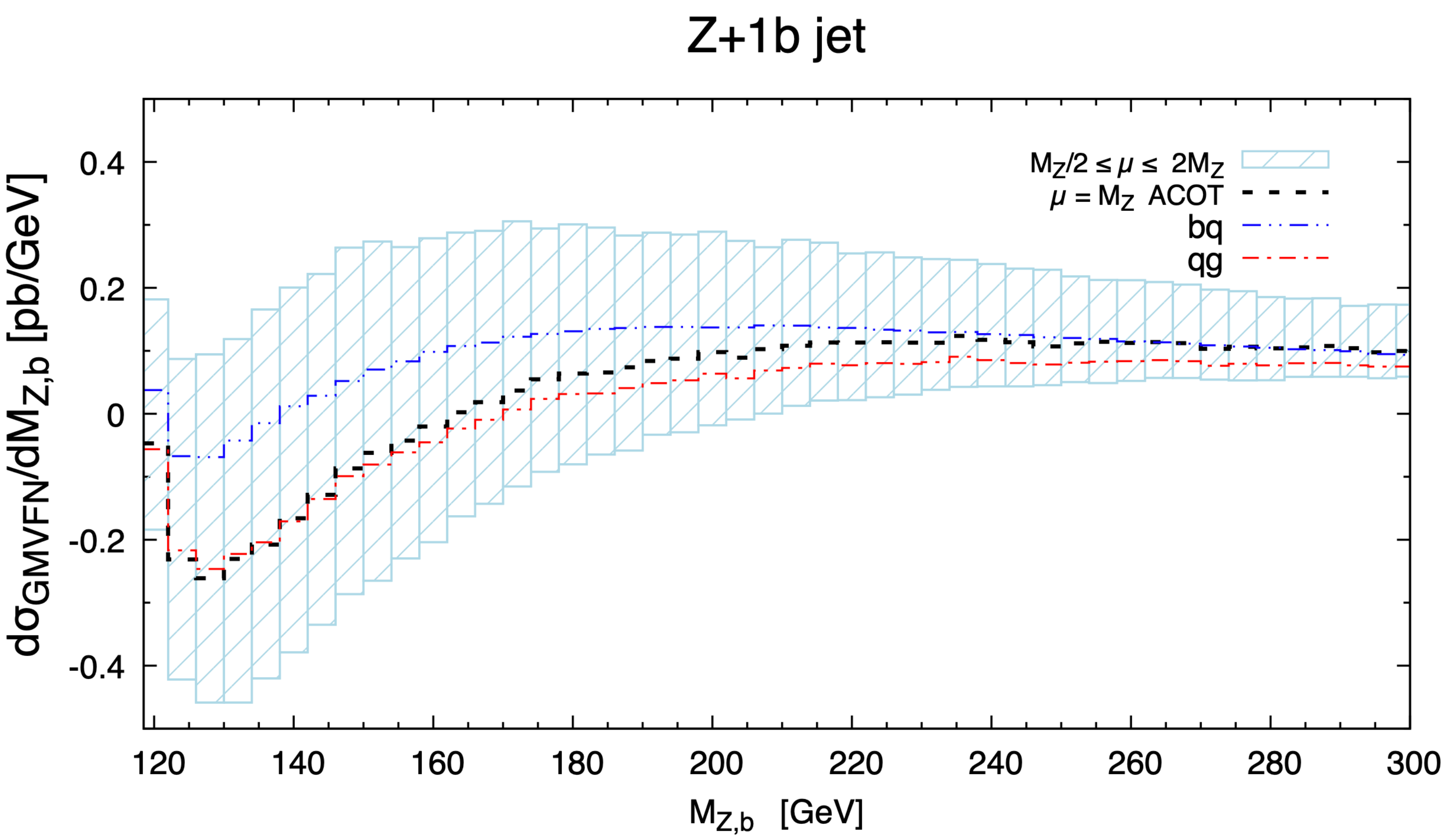
Residual PDFs

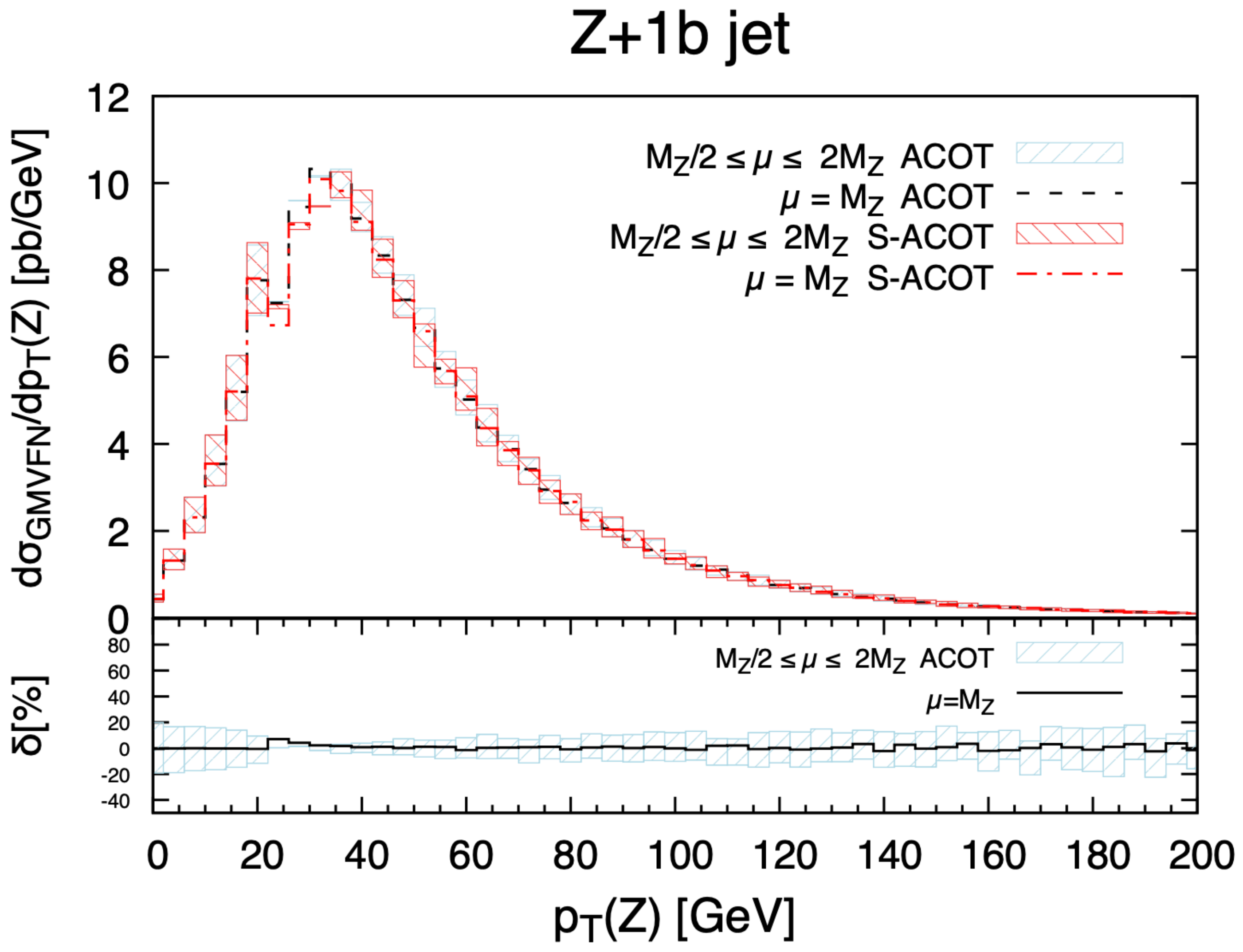
$$\delta f_Q^{(1)} = f_Q - \tilde{f}_Q^{(1)}, \qquad \delta f_Q^{(\rm NLO)} = f_Q - \tilde{f}_Q^{(\rm NLO)}$$

RESULTS

FE and FC terms *vs* the ACOT prediction

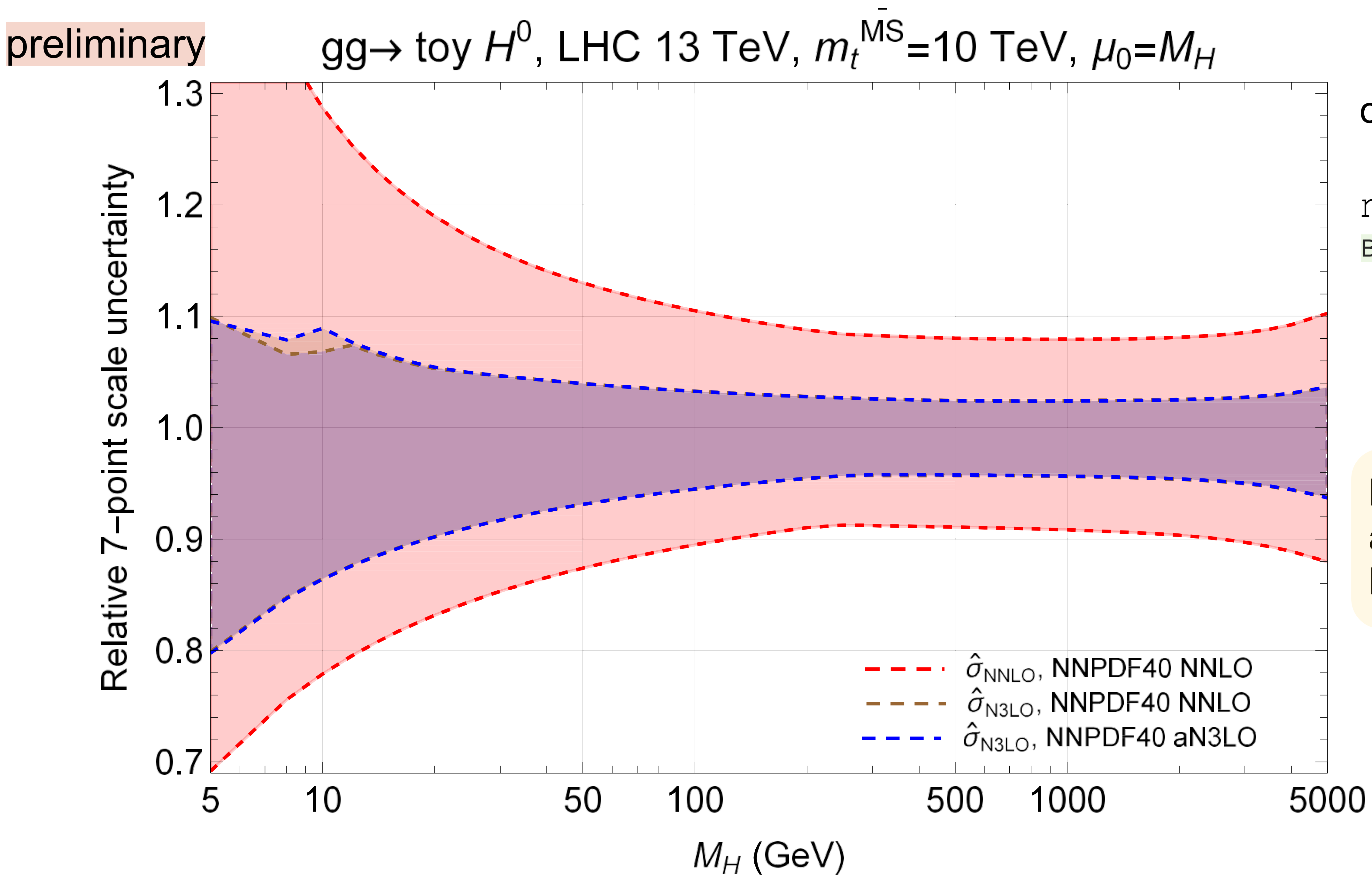
- qg channel appears first at NLO → larger relative scale uncertainty (cf. gg)





- representative results: matching of (S-)ACOT calculation to FC/FE terms proceeds smoothly
  - achieved through subtraction terms; sensitive to PS integration, applied cuts

what about (a)N3LO? total cross sections (Higgs prod. scale uncertainties)



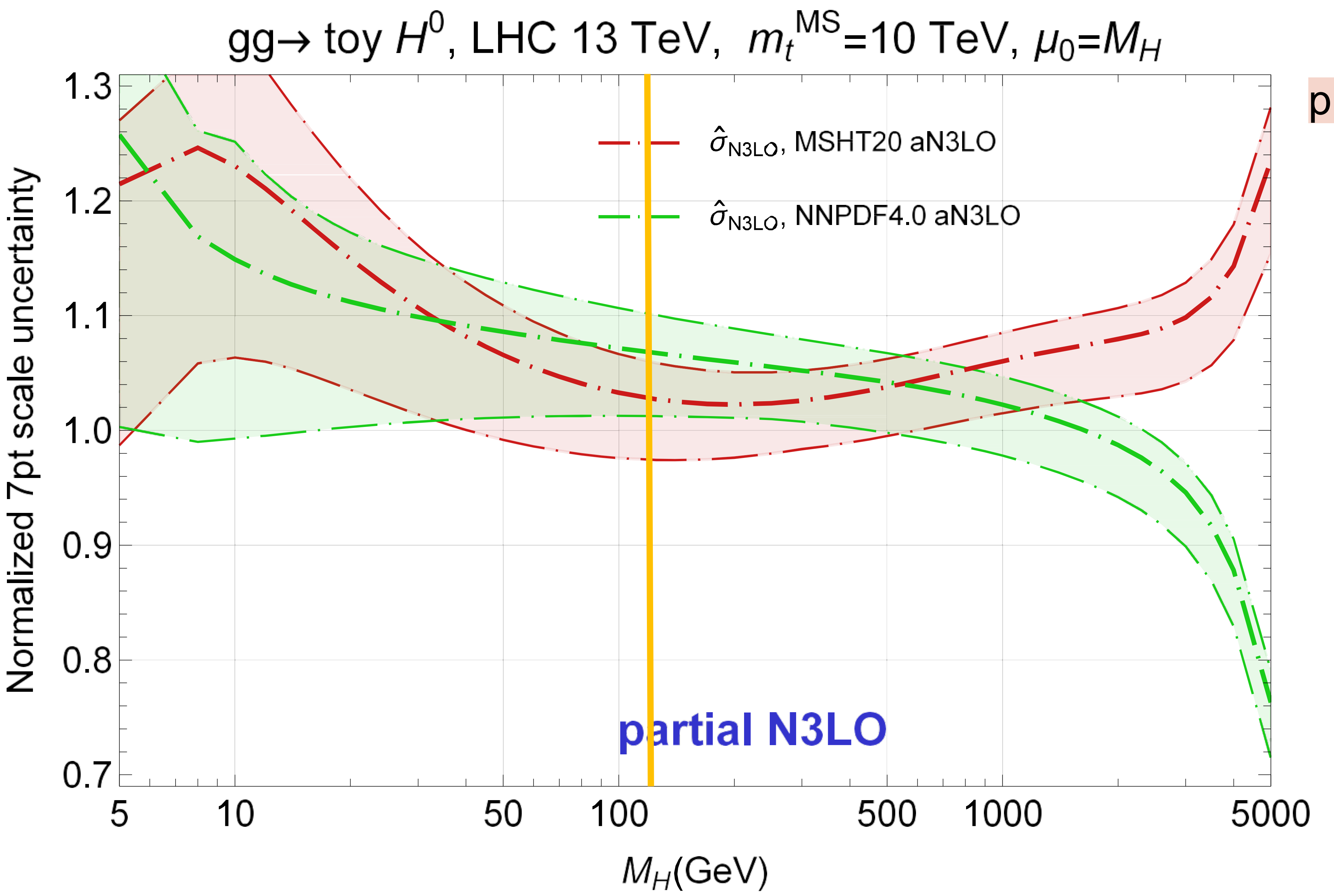
calculations with Max Ponce:

`n3loxs`

Baglio, Duhr, Mistlberger, Szafron (2209.06138)

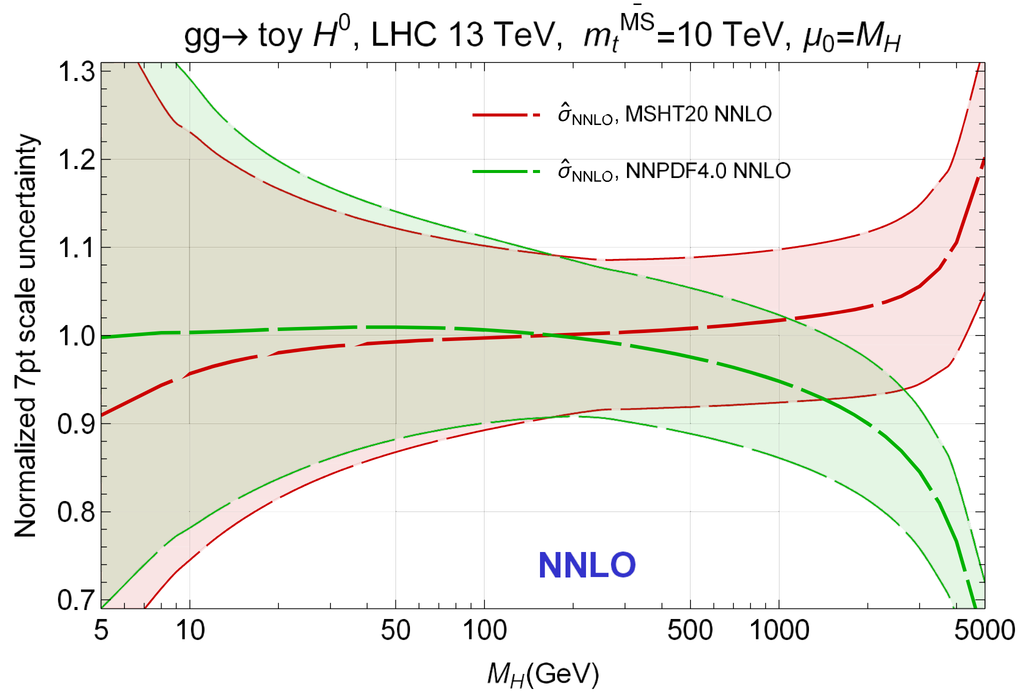
N3LO scale uncertainty is about the same with either NNLO or aN3LO PDFs

At $M_H \approx 10$ GeV, more variability due to the $b\bar{b}$ mass threshold



preliminary

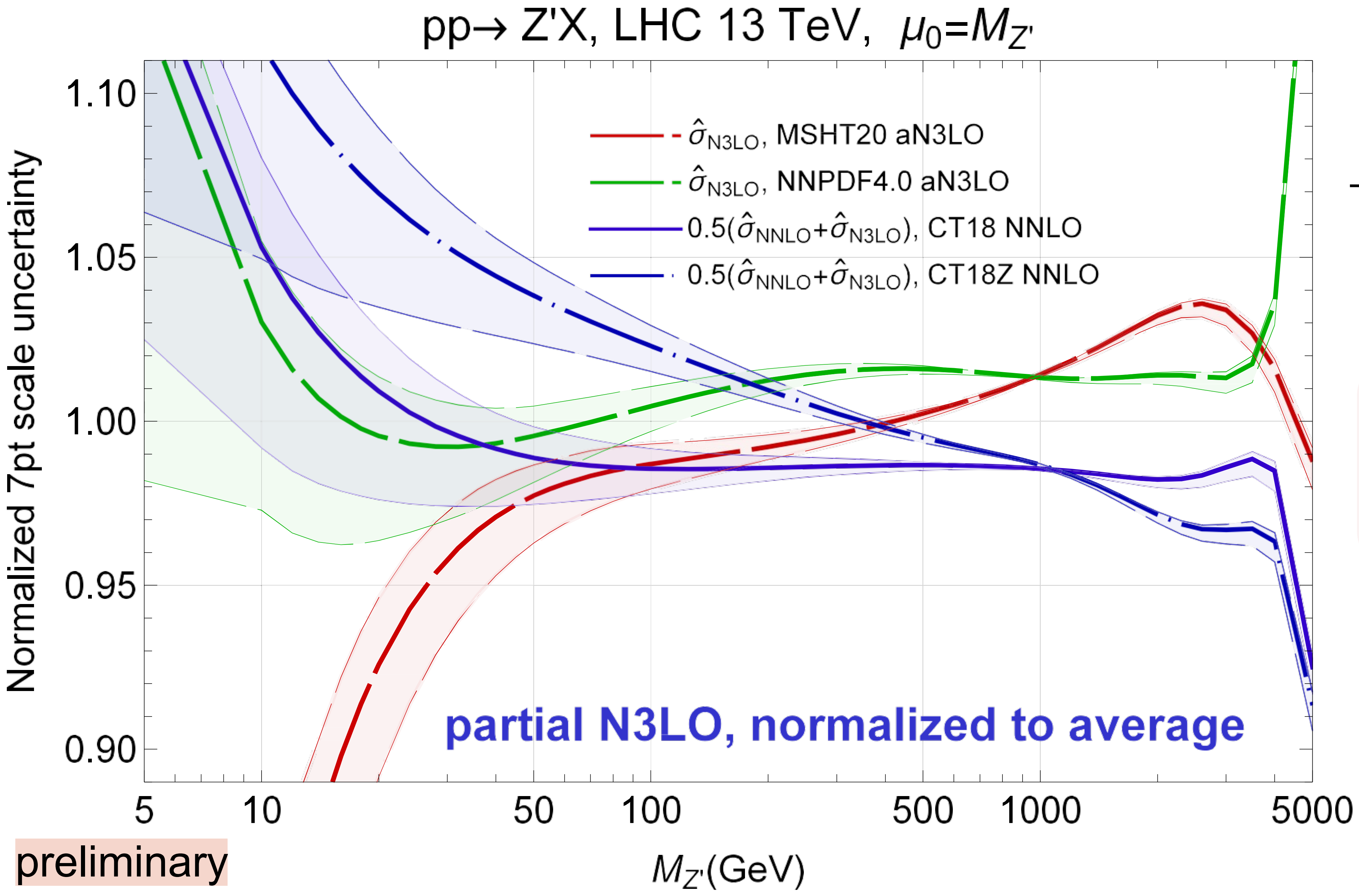
cf.:



Enlarged $g(x)$ at $x < 10^{-3}$

Weaker agreement at $M_H = 125$ GeV than at NNLO

Persistent differences at $x > 0.1$ reflect tensions in fitted data



preliminary

analogous calculation: Z' total cross section

- other PDF uncertainties > (N3LO-NNLO) differences: *e.g.*, PDF priors, modeling of systematics, …

→ don't automatically decrease at NNLO+

Increased MSHT-NNPDF aN3LO mismatch at $M_Z' < 50$ GeV

Good agreement among the groups at 50-500 GeV

Differences persist for $M_{Z'} > 1$ TeV, even as central PDFs refitted at different orders; may possibly reduce with new data

related: ongoing benchmarking of QED corrections at NNLO

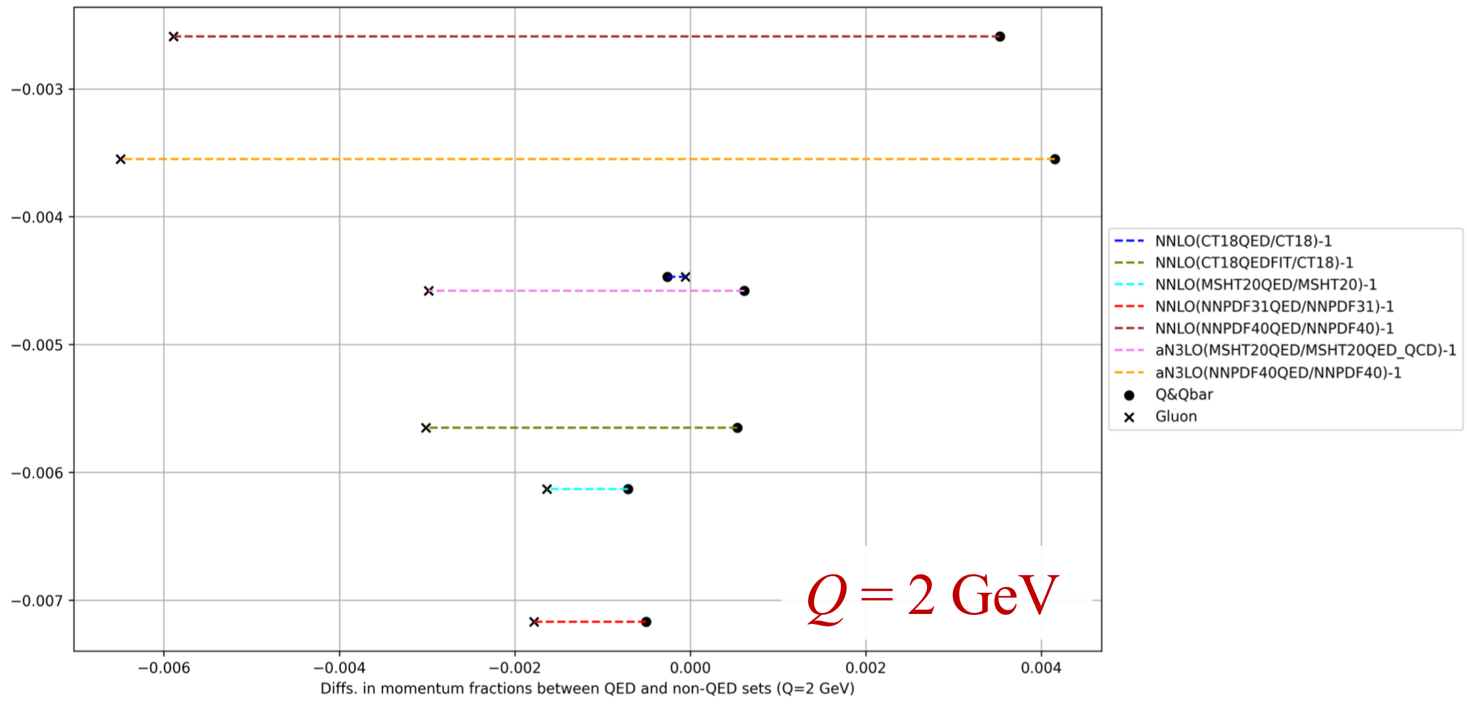
preliminary

$pp \to W^+ H$ *e.g.*, Higgsstahlung

- qualitative point: can relate total cross section shifts (vertical) to variations in QED-corrected PDF moments (abscissa)

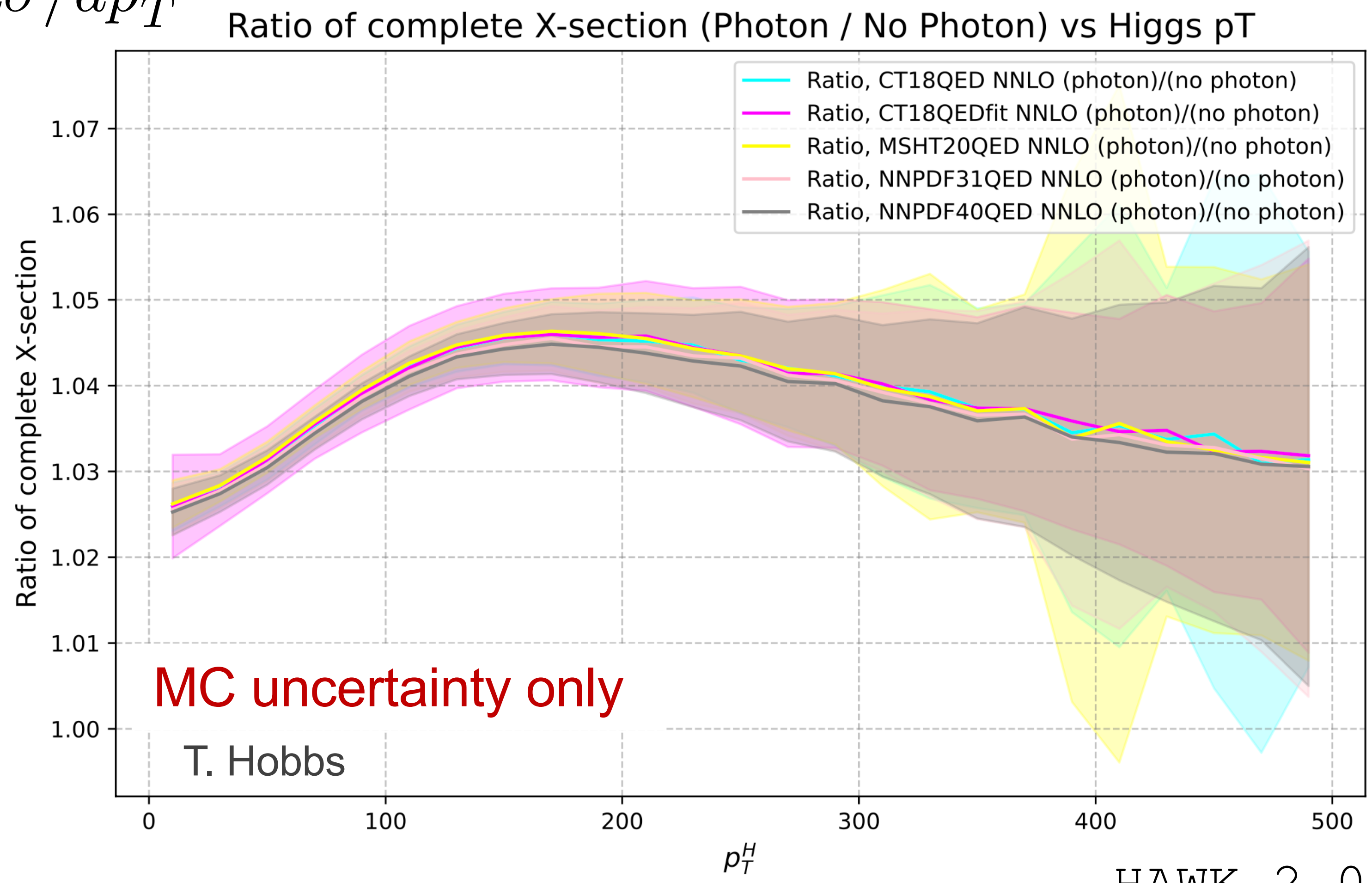
$[\sigma^{\rm QED}/\sigma^{\rm non-QED}] - 1$

Max Ponce



n3loxs $\langle f_q^{\rm QED} \rangle - \langle f_q^{\rm non-QED} \rangle$

$d\sigma/dp_T^H$



HAWK-2.0

- differential distributions at NLO EW: minimal PDF dependence in ratio of cross section with/out photon-initiated contributions; ~2% PDF variations in absolute cross sections

NNLO fits with new data at 8 and 13 TeV

arXiv: 2408.11131

χ^2/N_{pt} for CT18+new data (CT18 in parentheses) NNLO fits; 68% CL

(fits with 1 new process, 'nProcces')

ID	Experiment	N_{pt}	CT18+nDYTTIncJet (CT18)
	Drell-Yan pair production		
211	ATLAS 8 TeV W	22	$2.42^{+2.49}_{-1.51}$ ($4.25^{+6.39}_{-3.34}$)
212	CMS 13 TeV Z	12	$2.48^{+4.76}_{-0.88}$ ($12.03^{+38.04}_{-21.84}$)
214	ATLAS 8 TeV Z3D	188	$1.12^{+0.46}_{-0.02}$ ($1.99^{+5.10}_{-1.85}$)
215	ATLAS 5.02 TeV W,Z	27	$0.82^{+0.55}_{-0.16}$ ($1.15^{+1.22}_{-0.43}$)
217	LHCb 8 TeV W	14	$1.35^{+0.59}_{-0.61}$ ($1.35^{+0.72}_{-0.64}$)
218	LHCb 13 TeV Z	16	$1.18^{+1.42}_{-0.60}$ ($1.49^{+1.74}_{-0.89}$)
	$t\bar{t}$ production at 13 TeV		
521	ATLAS all-hadronic $y_{t\bar{t}}$	12	$1.06^{+0.14}_{-0.09}$ ($1.05^{+0.21}_{-0.10}$)
528	CMS dilepton $y_{t\bar{t}}$	10	$1.10^{+1.08}_{-0.68}$ ($1.03^{+1.60}_{-0.74}$)
581	CMS lepton+jet $m_{t\bar{t}}$	15	$1.44^{+1.18}_{-0.73}$ ($1.37^{+1.86}_{-0.82}$)
587	ATLAS lepton+jet $m_{t\bar{t}} + y_{t\bar{t}} + y^{B}_{t\bar{t}} + H^{t\bar{t}}_{T}$	34	$0.92^{+0.32}_{-0.14}$ ($0.94^{+0.59}_{-0.16}$)
	Inclusive jet production		
553	ATLAS 8 TeV IncJet	171	$1.76^{+0.20}_{-0.12}$ ($1.80^{+0.33}_{-0.16}$)
554	ATLAS 13 TeV IncJet	177	$1.38^{+0.13}_{-0.10}$ ($1.39^{+0.20}_{-0.11}$)
555	CMS 13 TeV IncJet	78	$1.10^{+0.24}_{-0.17}$ ($1.11^{+0.30}_{-0.16}$)

nDY

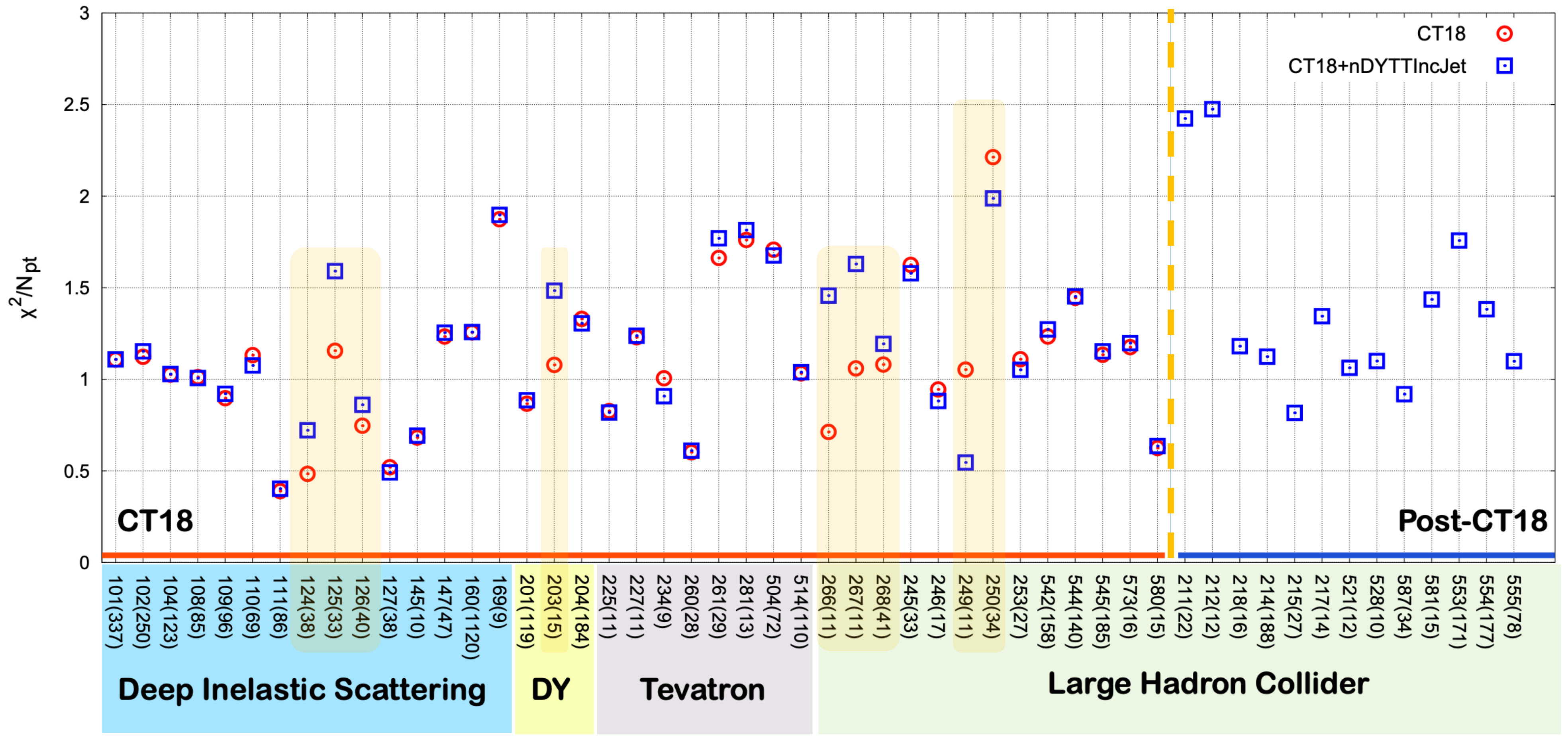
nTT

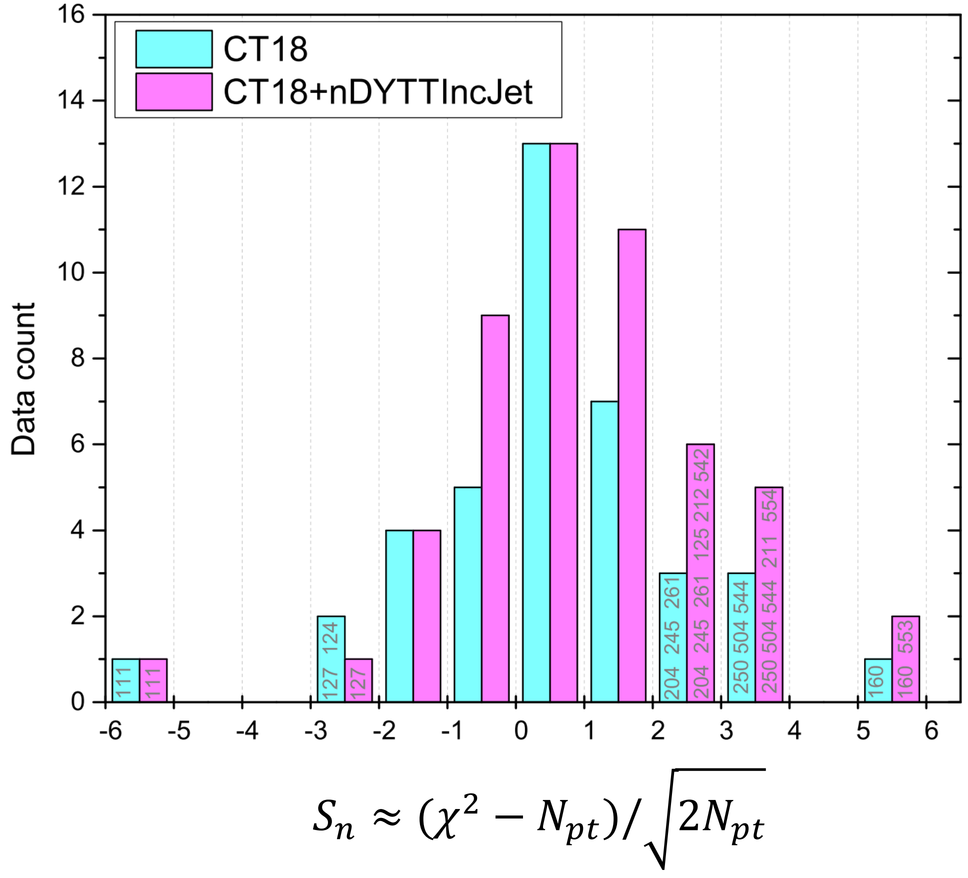
nIncJet

nDYTTIncJet

(combined fit)

CT18 baseline vs. CT18+nDYTTIncJet NNLO





The most precise new experiments tend to have an elevated χ^2/N_{pt}, in the same pattern as observed for CT18

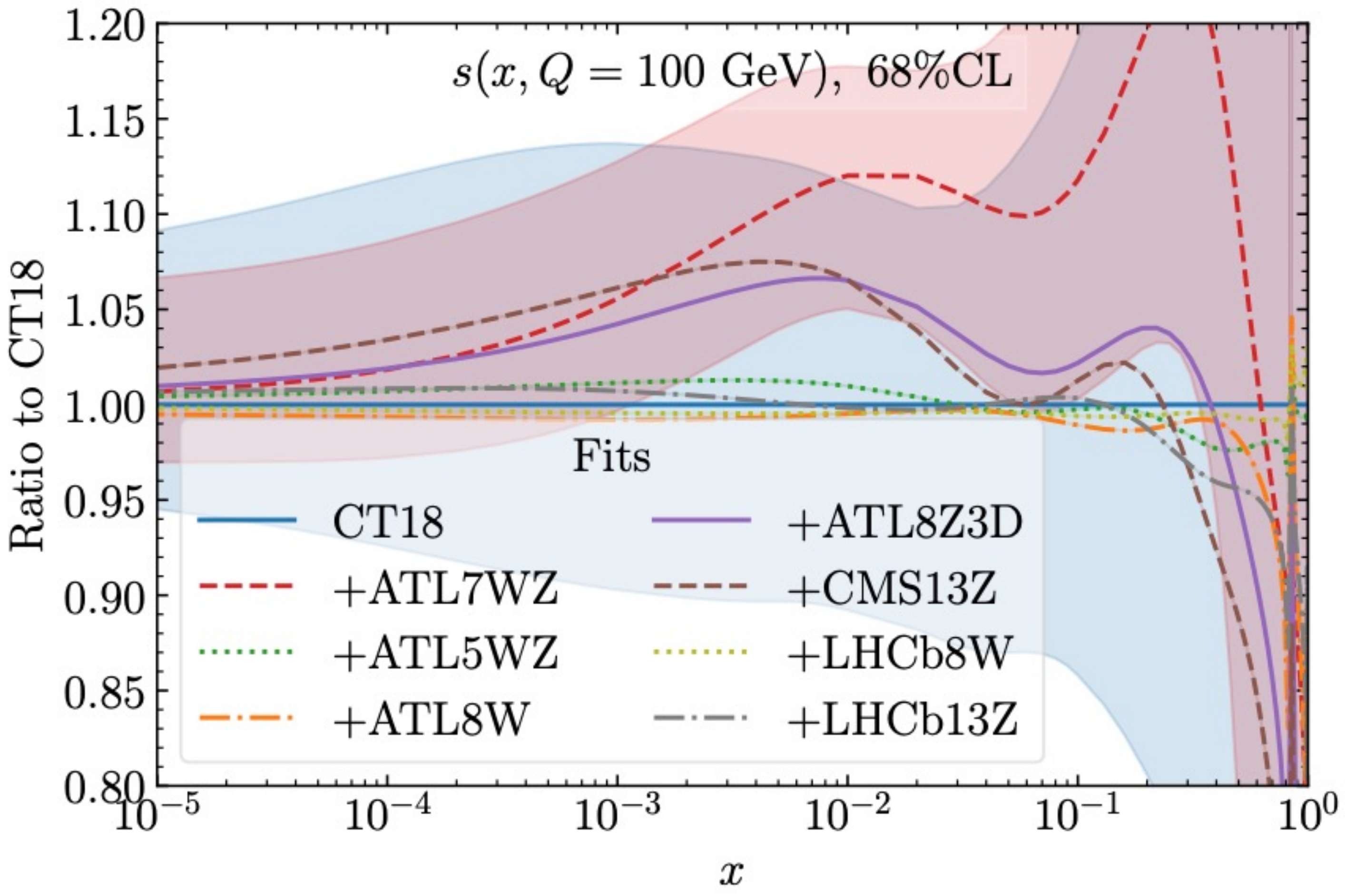
χ^2/N_{pt} increases for experiments 124 and 125 (NuTeV), 126 and 127 (CCFR) and 203 (E866 DY), 266 and 267 (CMS 7 TeV A_{ch}), 268 (ATLAS 7 TeV W, A_{ch}).

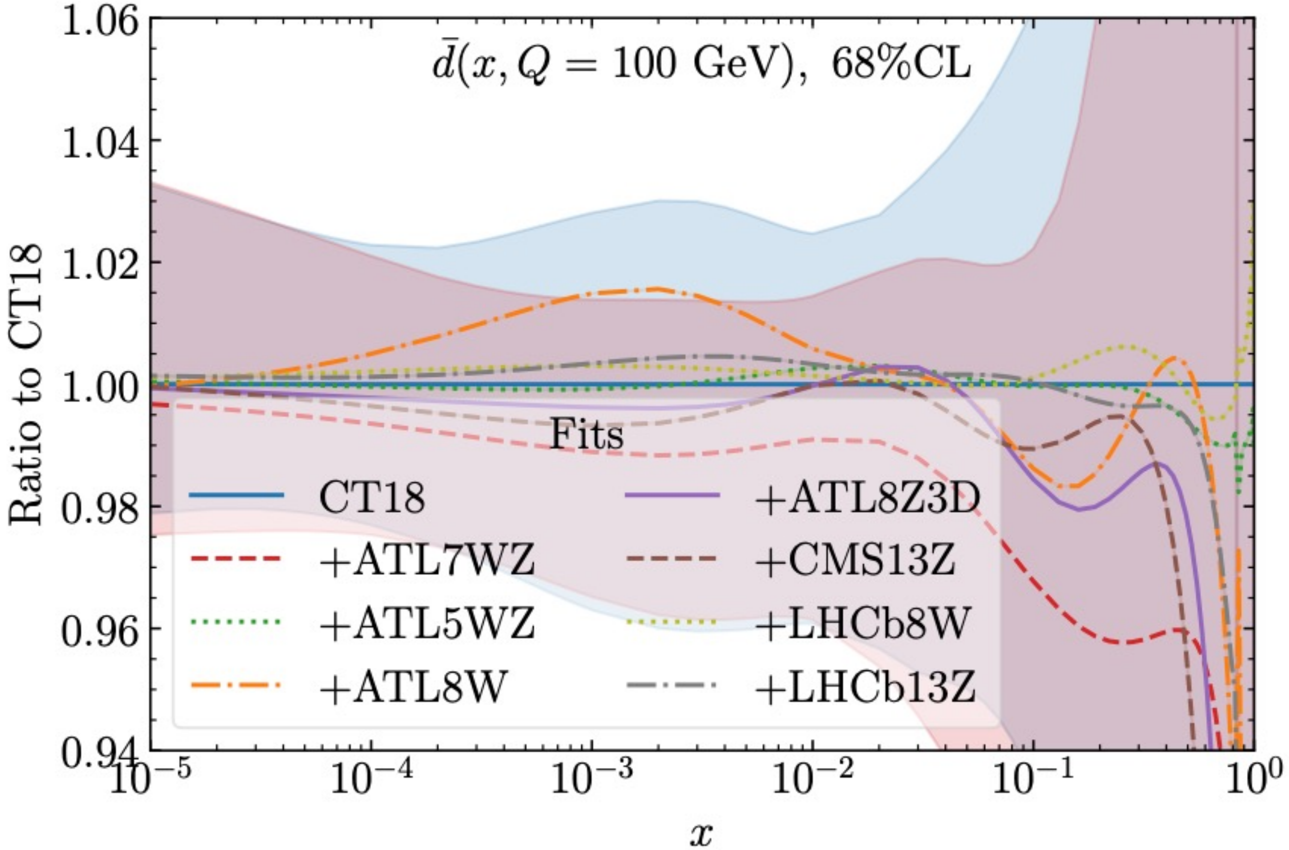
χ^2/N_{pt} decreases for experiments 249 (CMS 8 TeV A_{ch}), 250 (LHCb 8 TeV W/Z)

PDF impact: Post-CT18 Drell-Yan

2305.10733 (PRD23')

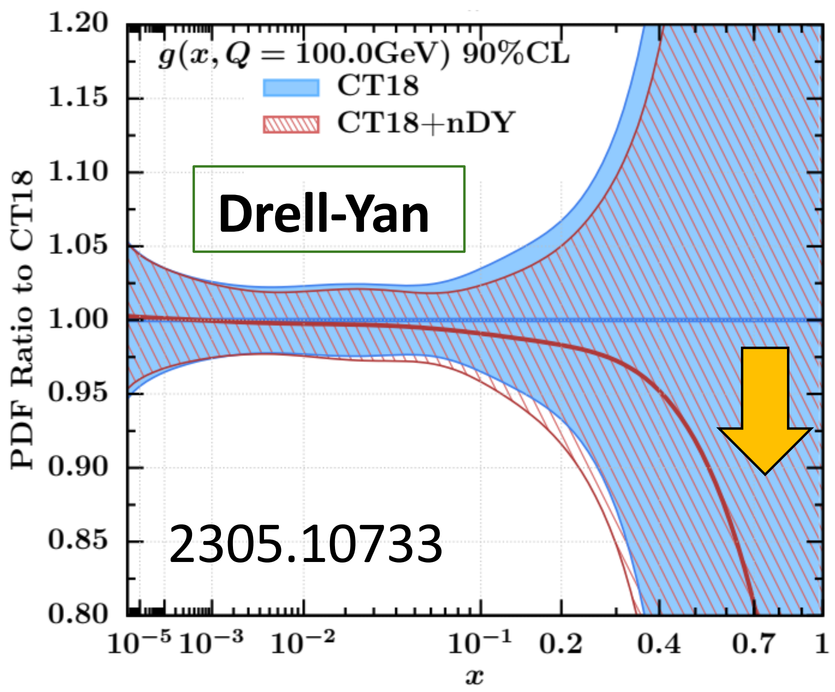
ID	Experiment	N_{pt}	CT18	CT18A	CT18As
215	ATLAS 5.02 TeV W, Z	27	0.81	0.71	0.71
211	ATLAS 8 TeV W	22	2.45	2.63	2.51
214	ATLAS 8 TeV Z 3D†	188	1.12	1.14	1.18
212	CMS 13 TeV Z	12	2.38	2.03	2.71
216	LHCb 8 TeV W	14	1.34	1.36	1.43
213	LHCb 13 TeV Z	16	1.10	0.98	0.83
248	ATLAS 7 TeV W, Z	34	2.52	2.50	2.30
Total 3994/3953/3959 points			1.20	1.20	1.19

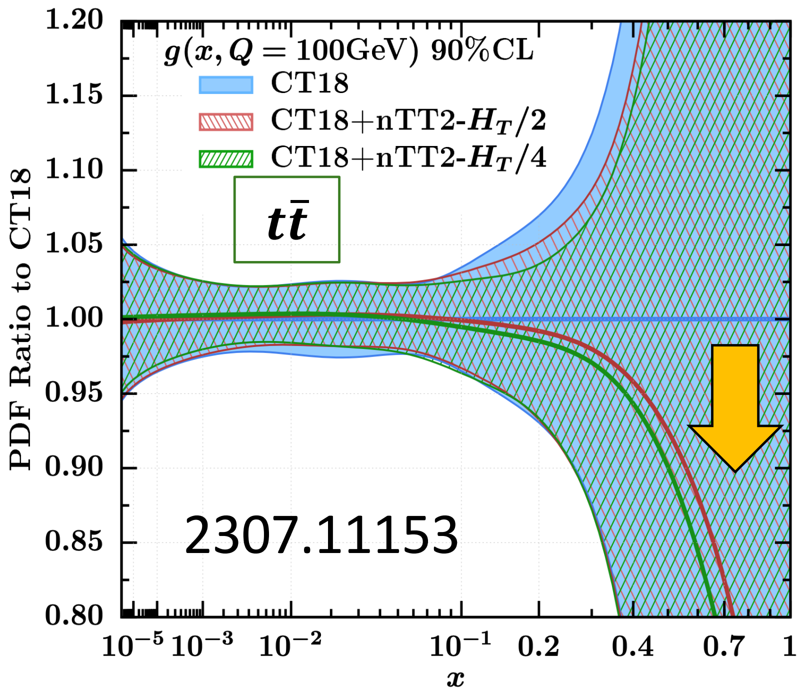


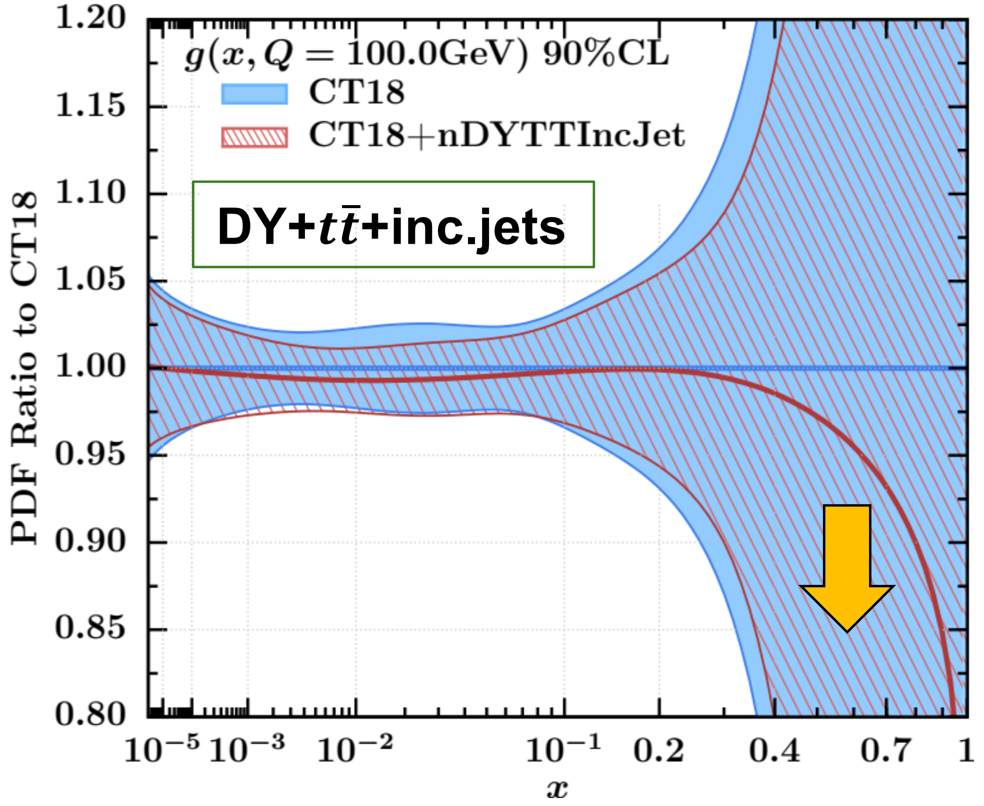


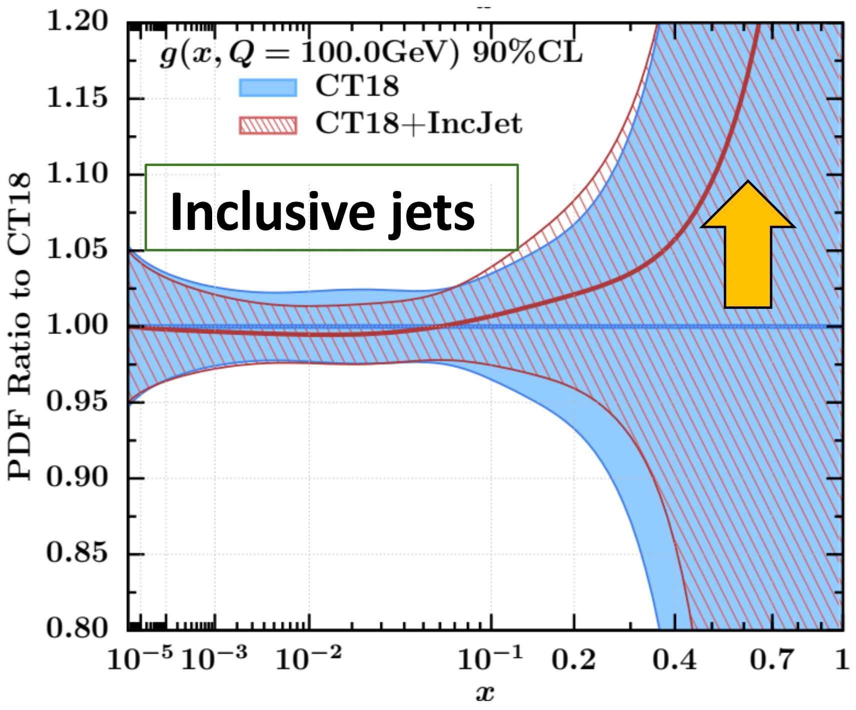
- many new Drell-Yan (nDY) after CT18 main release
- most nDY data sets consistent with ATLAS 7 WZ precision data (16'); prefer enhanced strangeness at $x \sim 0.02$, but with somewhat smaller enhancement
- one exception: ATL8W has opposing pull on $d, \bar{d}$
- CMS13Z and ATL8W have a similar χ^2/N_{pt} as ATL7WZ
- more flexible strangeness [CT18As] reduces (but does not resolve) tension

Pulls on gluon PDF by new data type



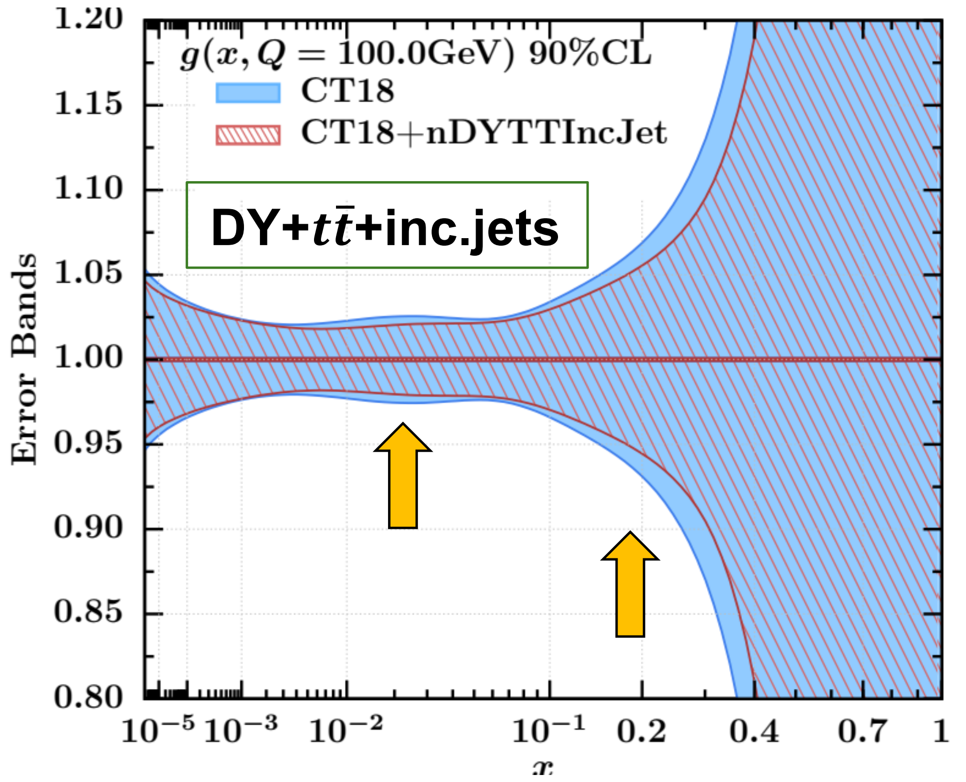






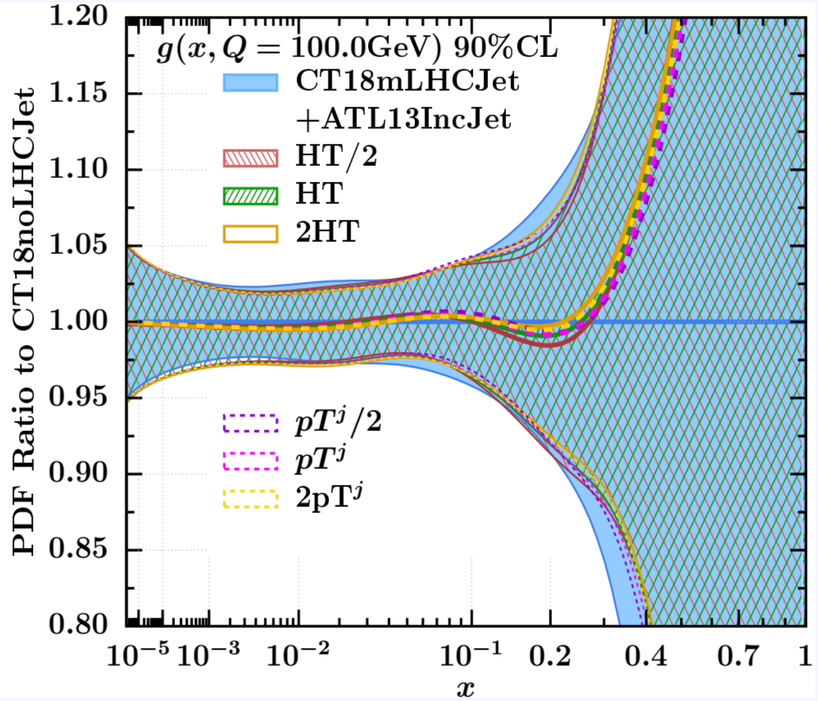
After including DY, $t\bar{t}$, and inc. jet data simultaneously, we get a softer gluon. Note that new DY and $t\bar{t}$ data favor a softer gluon, new inc. jet data prefer a harder gluon.

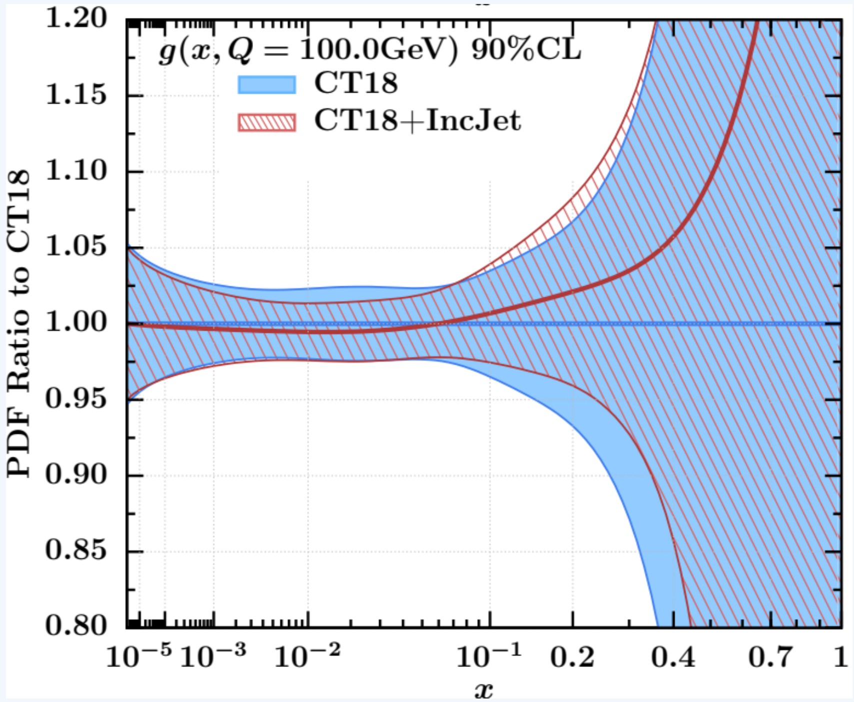
Mild changes in the gluon uncertainty

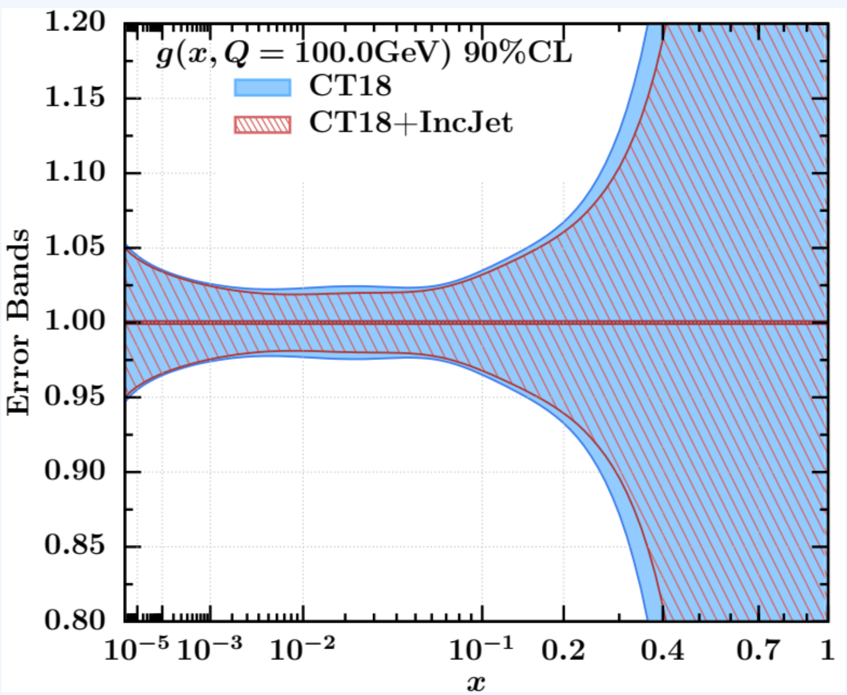


inclusive jet vs. dijet data sets: impact on gluon, various QCD scales

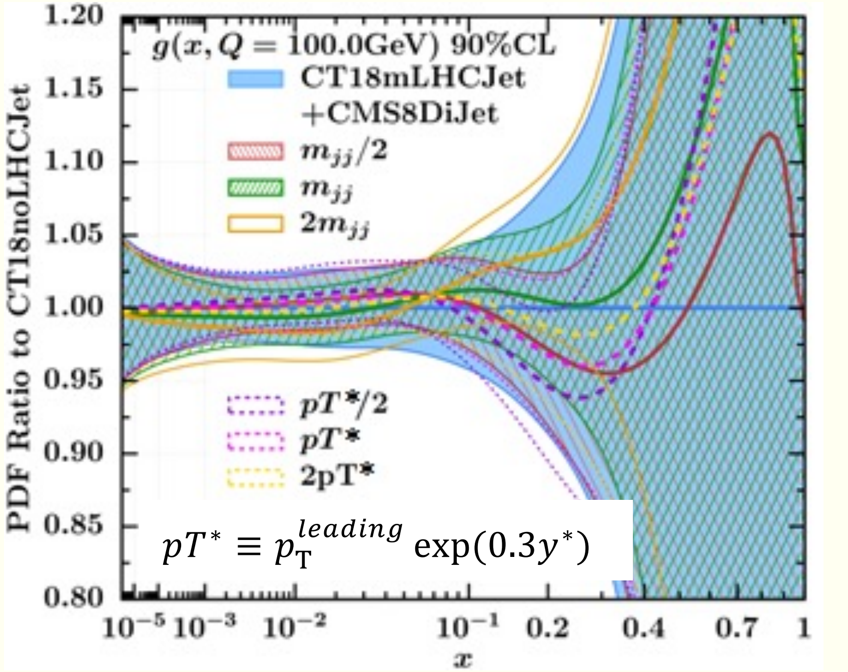
+ inclusive jets: small scale dependence, a harder $g(x,Q)$







+ dijets: significant scale, dependence, varied pulls on $g(x,Q)$



- Inclusive jet impact on $g(x,Q)$ is relatively independent of the scale choice. The final fit uses $\mu_{R,F} = p_T^j$, giving better χ^2
- **Dijet PDF impact substantially depends on scale choices**, especially for CMS8 TeV DiJet

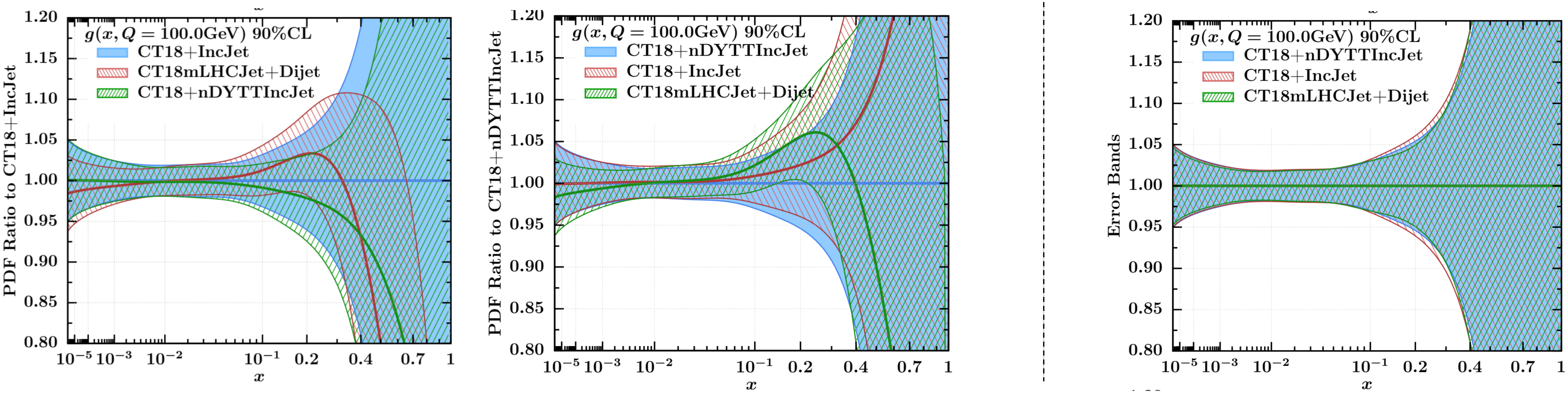
χ^2/N_{pt} for fits that add one inclusive jet or dijet data set to the CT18 (without LHC jets) baseline at a time

Inclusive jets		χ^2/N_{pt} using $\mu_{R,F} \propto HT$ or p_T^j					
Experiment	N_{pt}	$HT/2$	HT	$2HT$	$p_T^j/2$	p_T^j	$2p_T^j$
ATL8IncJet	171	1.7	1.74	1.87	1.75	1.66	1.7
ATL13IncJet	177	1.42	1.36	1.4	1.52	1.31	1.28
CMS13IncJet	78	1.2	1.16	1.2	1.08	1.09	1.1
Dijets		χ^2/N_{pt} using $\mu_{R,F} \propto HT$ or $p_T^* = p_T^j \exp(0.3y^*)$					
Experiment	N_{pt}	$M_{jj}/2$	M_{jj}	$2M_{jj}$	$p_T^*/2$	p_T^*	$2p_T^*$
ATL7DiJet	90	0.81	0.79	0.87			
CMS7DiJet	54	1.55	1.55	1.63			
CMS8DiJet	122	0.95	1.2	1.9	1.25	1	1.01
ATL13DiJet	136	0.9	0.87	0.93			

Dijet data are dominated by the CMS 8 TeV dataset

Dijet data sets tend to have larger uncertainties than inc. jets, facilitating better χ^2 for similar constraints on PDFs

PDFs from fits with inclusive jet and dijet data



- ❑ dijet data sets tend to have larger uncertainties than inc. jets, facilitating better χ^2 for similar constraints on PDFs

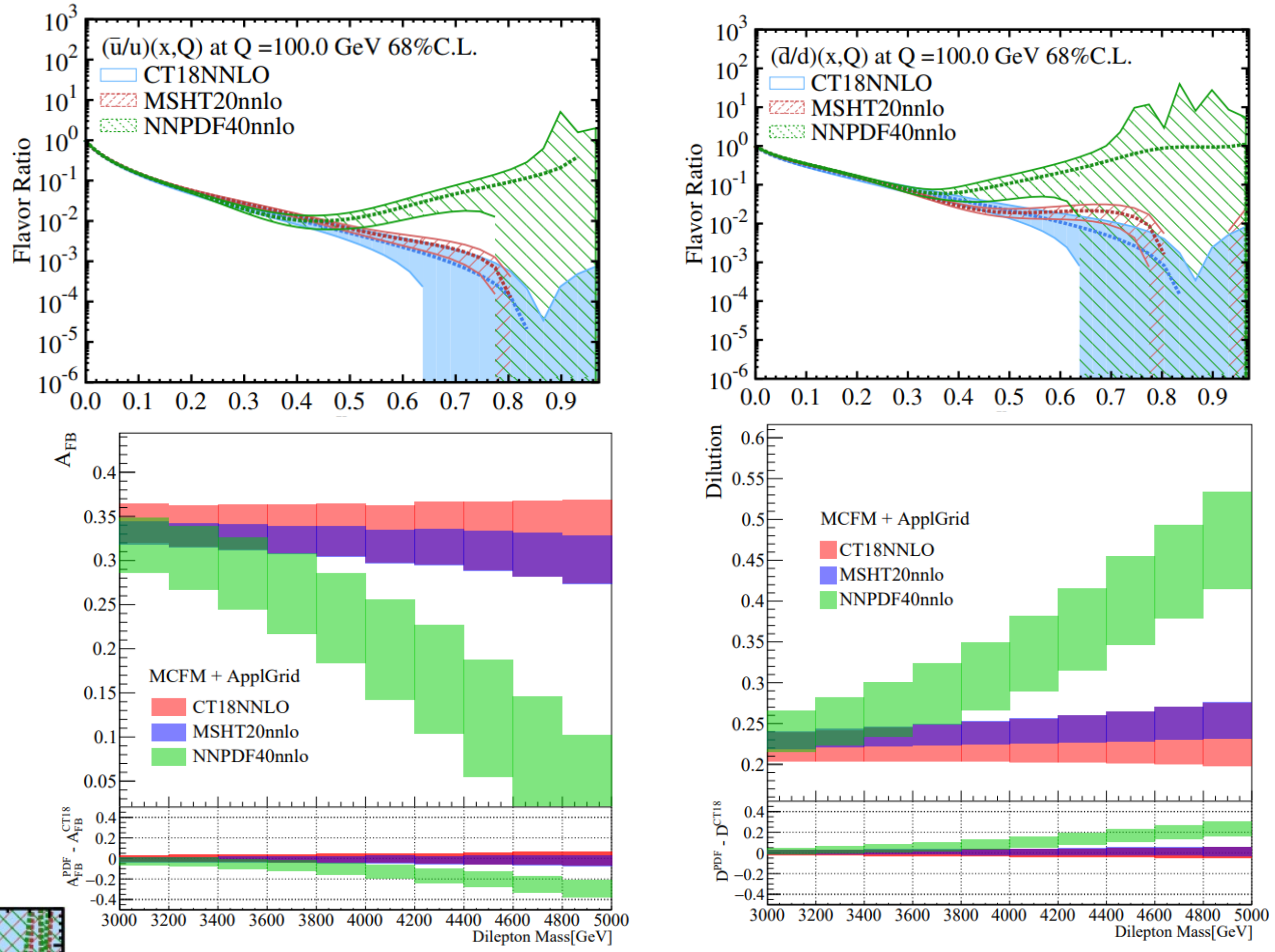
Impact of A_{FB} in the high-mass Drell-Yan process

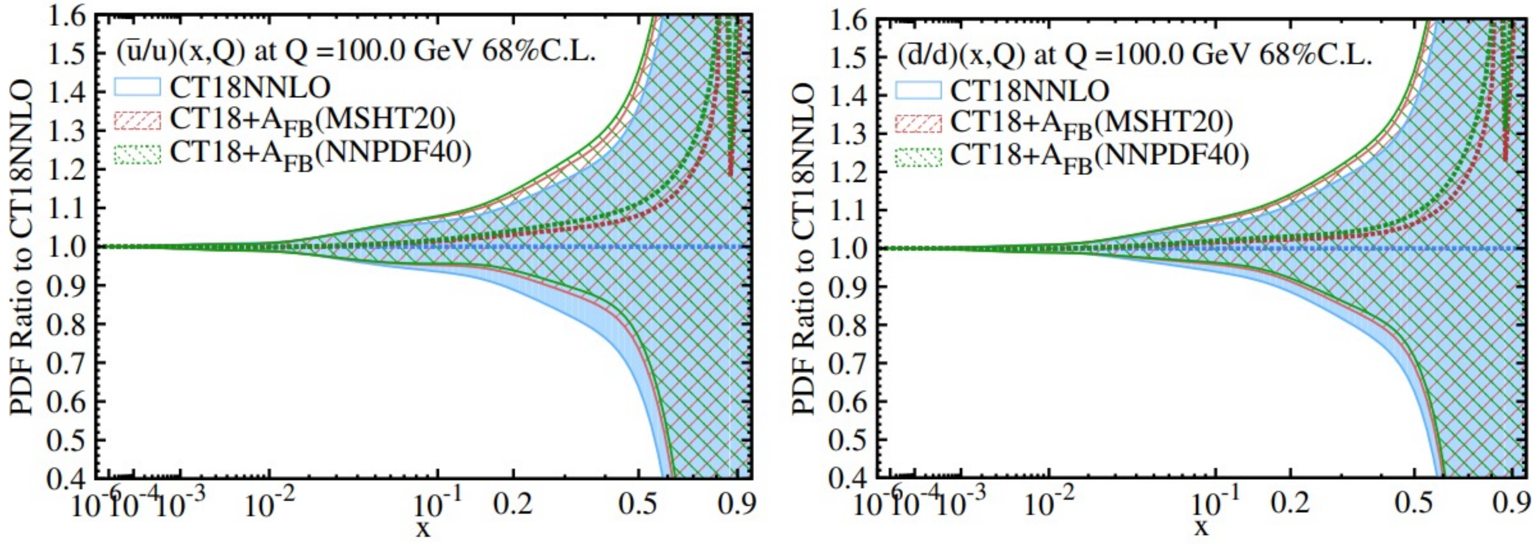
Y. Fu et al., 2307.07839
C. Willis et al., 1809.09481

- A_{FB} at the LHC is sensitive to the energy dilution factor D (probability of $k_q^0 < k_{\bar{q}}^0$ in the Collins-Soper frame)
- $A_{FB}^h = \frac{N_F^h - N_B^h}{N_F^h + N_B^h} \approx (1-2D)A_{FB}^q$
- A_{FB} at high invariant mass region probes $\bar{u}/u$, $\bar{d}/d$ at $x > 0.2$

relevant also for BSM searches; cf. PDF-BSM work

Gao, Gao, TJH, Liu, Shen, 2211.01094





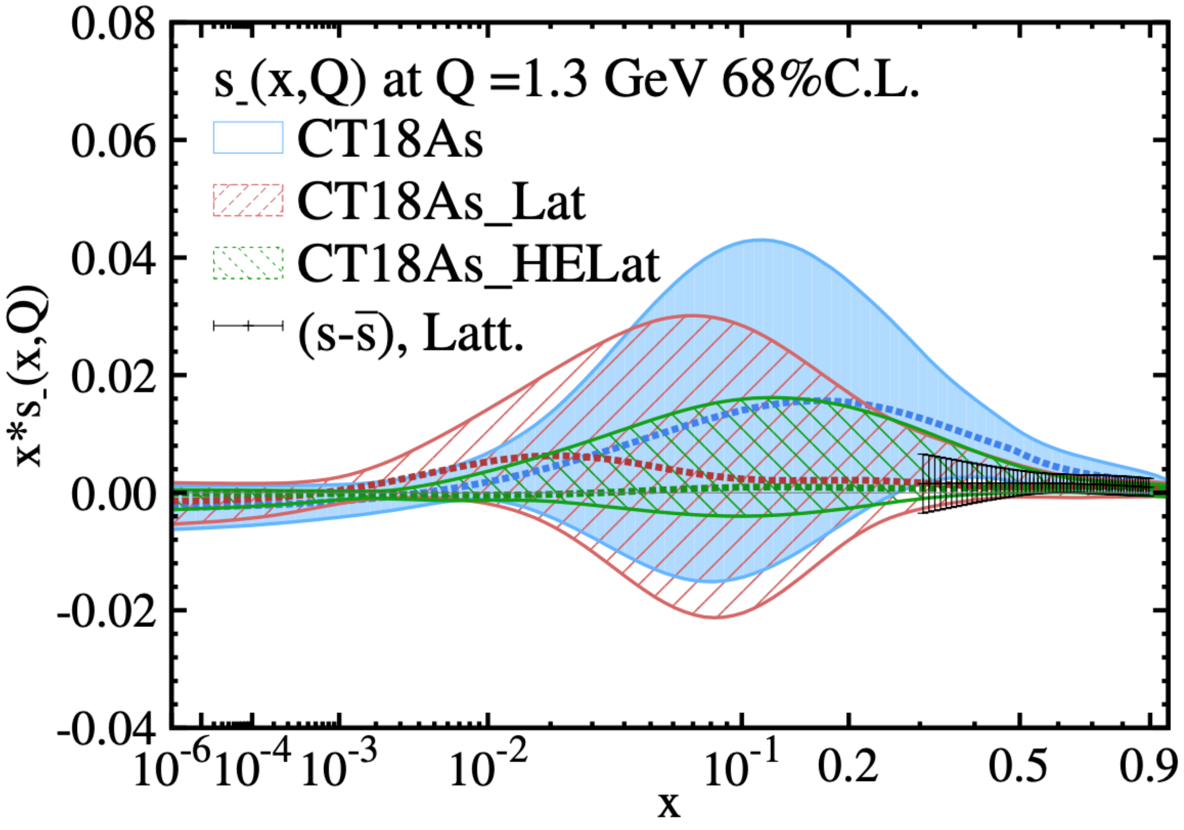
- CT18, MSHT20, and NNPDF4.0 predict very different $\bar{q}/q$ at $x > 0.2$
- The article quantified the potential effect of high-mass A_{FB} on large-x antiquarks

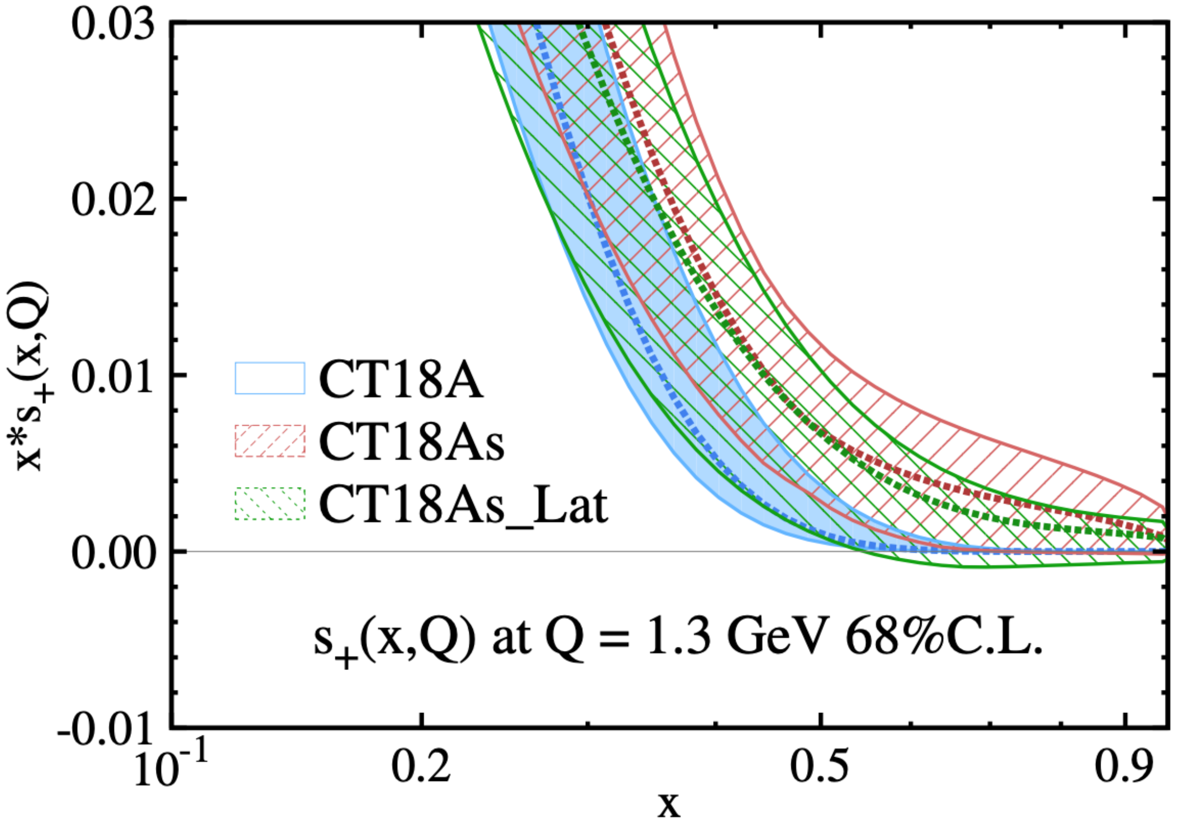
See also NNPDF (2209.08115), Fiaschi et al. (2211.06188)

recent CT studies on PDF-lattice connection(s)

- possible lattice QCD constraints to PDFs: *e.g.*, information on nucleon strangeness asymmetry
  - implement LM constraints from LaMET-based lattice data for $s - \bar{s}$; yield `CT18As_Lat` fit

Hou, Lin, Yan, Yuan, PRD107 (2023) 7, 076018





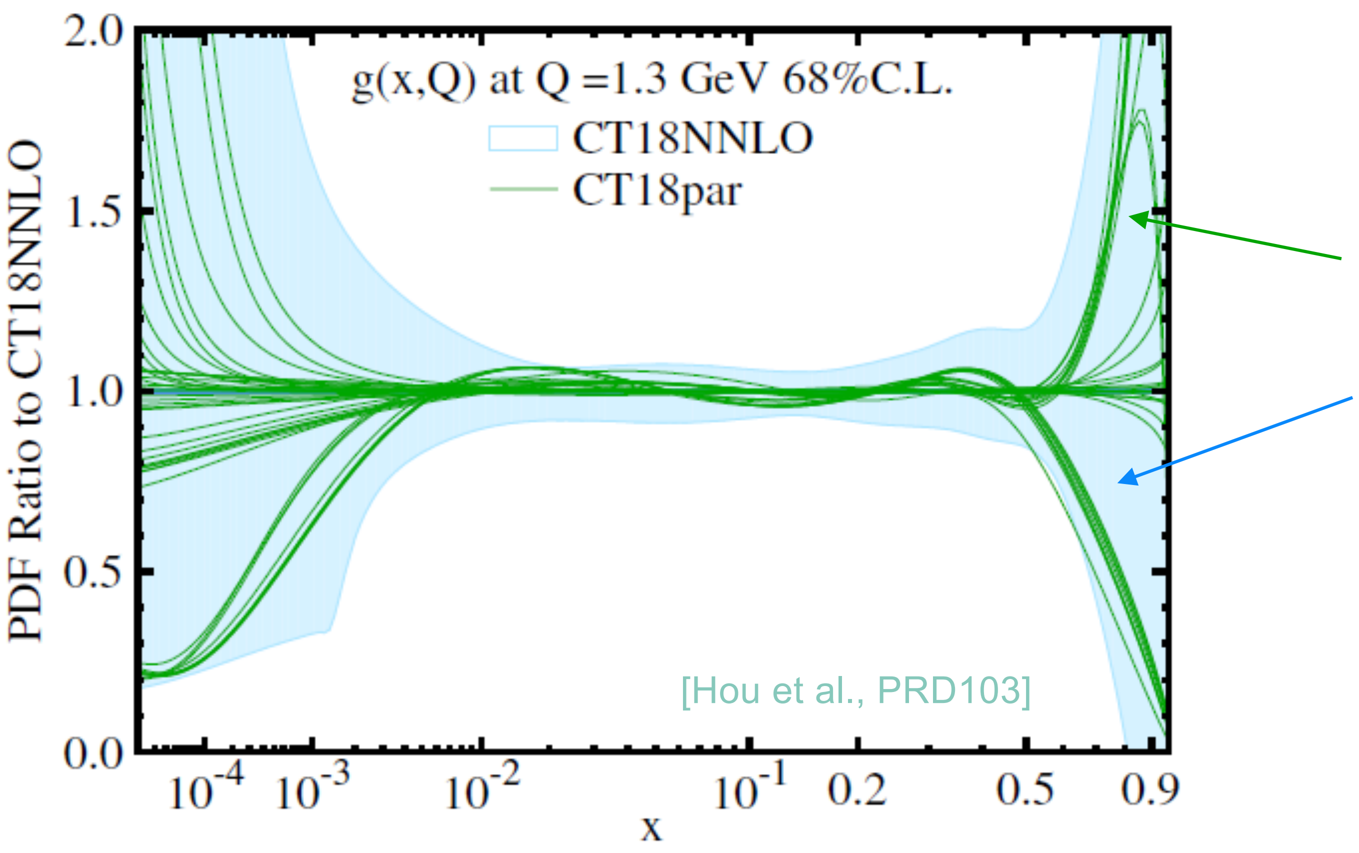
- significant reduction in asymmetric strange uncertainty; EW pheno implications (W/Z production; ν interactions; F_3 DIS SFs)

other CT-related studies --- *e.g.*, 1904.00022 (L_2 sensitivity method) TJH, Wang, Nadolsky, Olness

Taming PDF uncertainties in CT202X PDFs

Several efforts to refine PDF uncertainty quantification:

- understand conceptual underpinnings of the multivariate inverse problem. Much can be learned from non-HEP statistics applications
- suppress aleatory and perturbative uncertainties (e.g., from higher-order contributions)
- comprehensively estimate epistemic uncertainties (e.g., due to the PDF parametrization forms)



CT approach: "Bayesian exploration with Gaussian emulation"

preliminary PDFs for alternative parametrizations

final uncertainty with one parametrization

Preliminary fits explore experimental, theoretical, parametrization, methodological uncertainties

The final Hessian error set (50-60) approximates the total uncertainty due to the above factors.

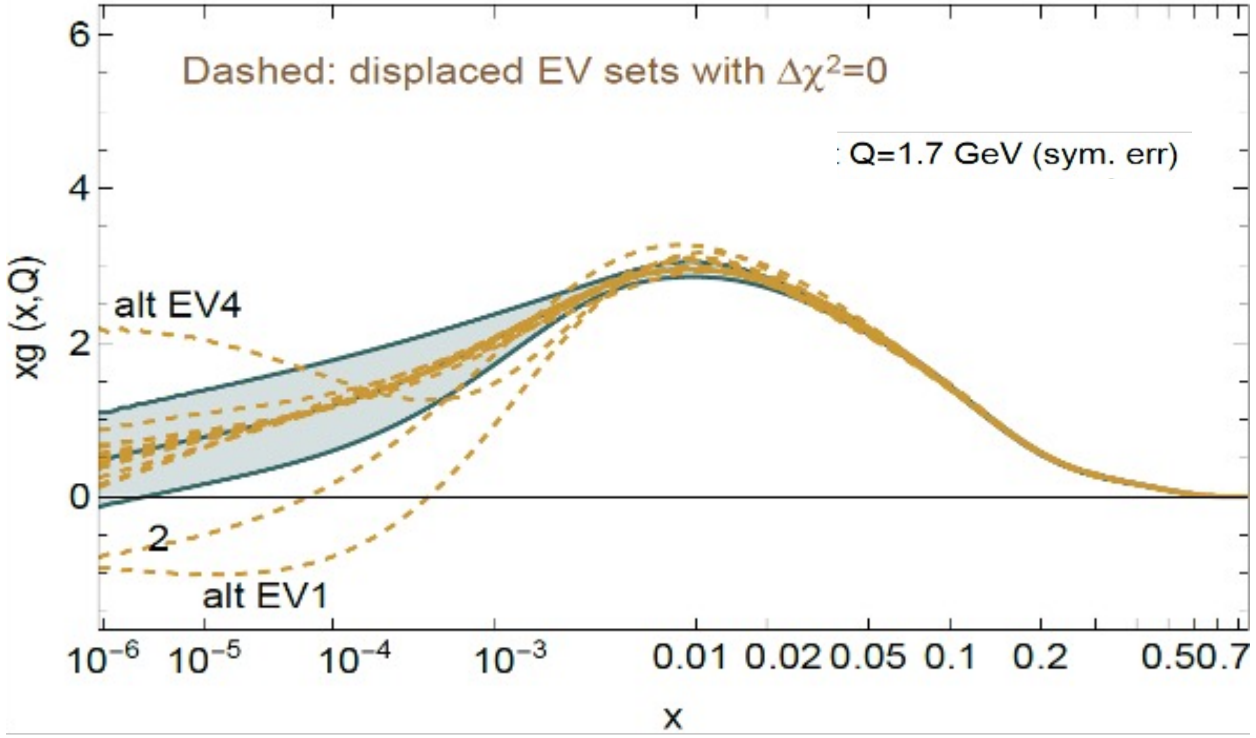
CT uncertainty quantification (UQ) developments

- CT quantifies parametrization choice as a crucial source of epistemic uncertainty

(justified by studies of the need for representative sampling of PDF model space)

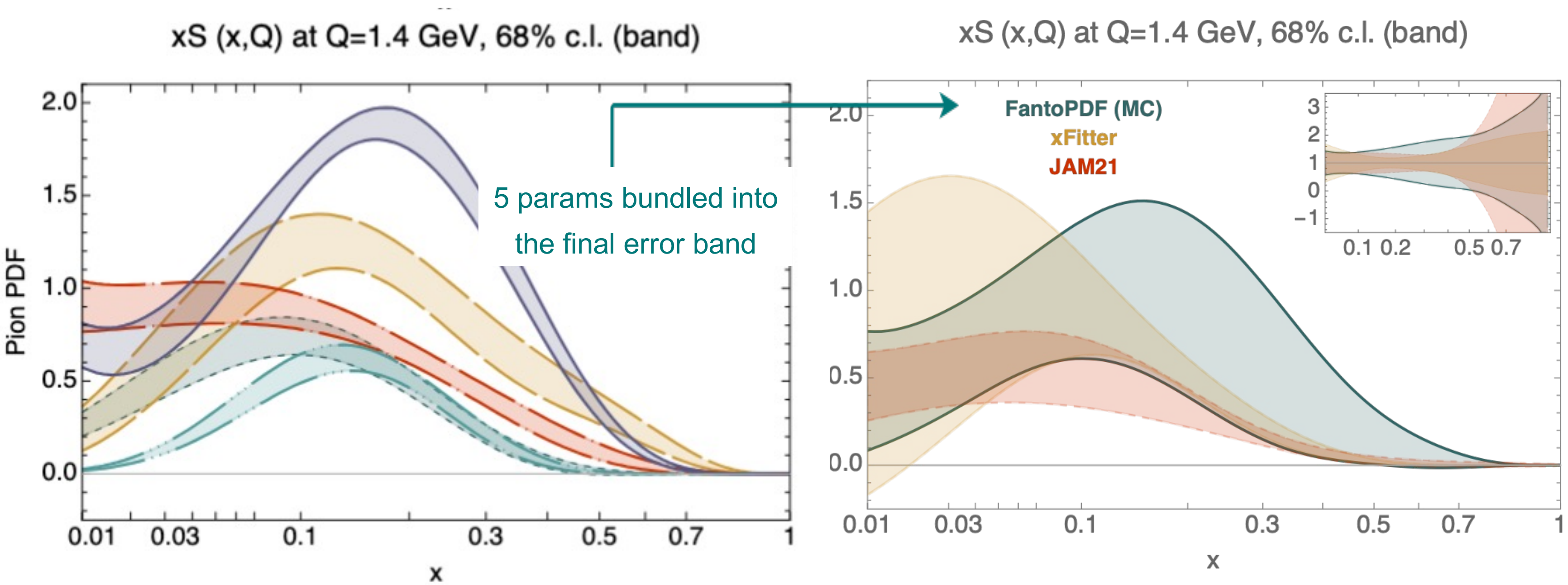
[Courtoy et al., PRD107, 2205.10444]

PDF with / out criteria for representative sampling



- complementary to parallel ML-based studies (more later)

studies PDF parameterization dependence using Bézier curves: Fantômas4QCD— first physics application for the pion PDFs

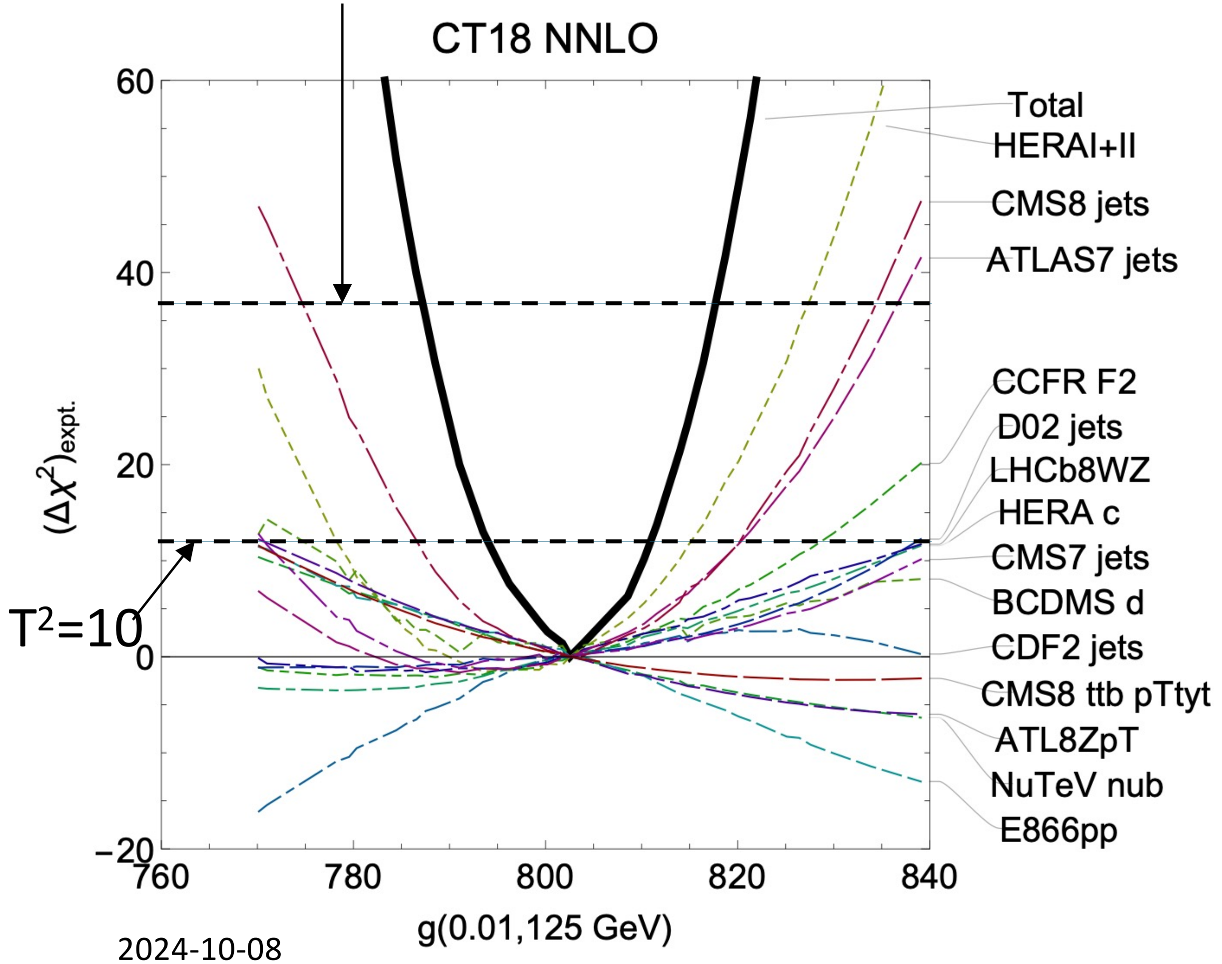


[Kotz et al., PRD109, 2311.08447]

J. Huston et al., a study of tolerances in progress (cf. backup)
L2 sensitivity: Jing et al., 2306.03918

CT and MSHT both use analytic minimization to determine the central PDF (by definition, at their best χ^2). This is different for the Monte-Carlo method of NNPDF.

The uncertainty is determined by allowing an excursion from that central value. For a 68% CL error on average, CT18 uses $\Delta\chi^2 \lesssim 37$. For MSHT it is closer to $\Delta\chi^2 \approx 10$.



Conceptually, uncertainties based on χ^2 are traced to the likelihood-ratio test:

$$\frac{P(T_2|D)}{P(T_1|D)} = \frac{P(D|T_2)}{P(D|T_1)} \times \frac{P(T_2)}{P(T_1)}$$

$\equiv r_{\rm posterior}$ $\equiv r_{\rm likelihood}$ $\equiv r_{\rm prior}$

∴ If two PDFs T_1, T_2 with the same priors have the same $\chi^2 = -2\ln P(D|T_i)$, they have the same confidence level

This fundamental Bayesian test justifies the technique of Lagrange Multiplier scans (on the left) as well as its fast approximation called "L2 sensitivity" (next slide).

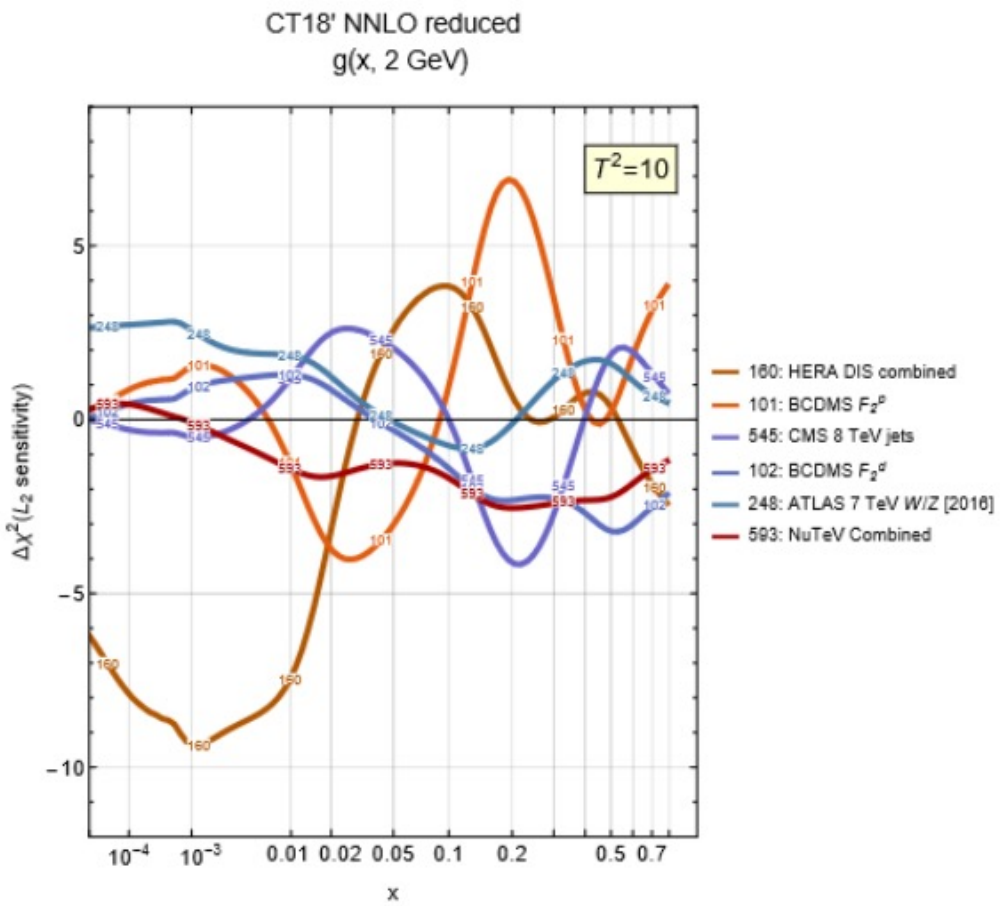
It also explains why $\Delta\chi^2 = 1$ does not capture the full uncertainty.

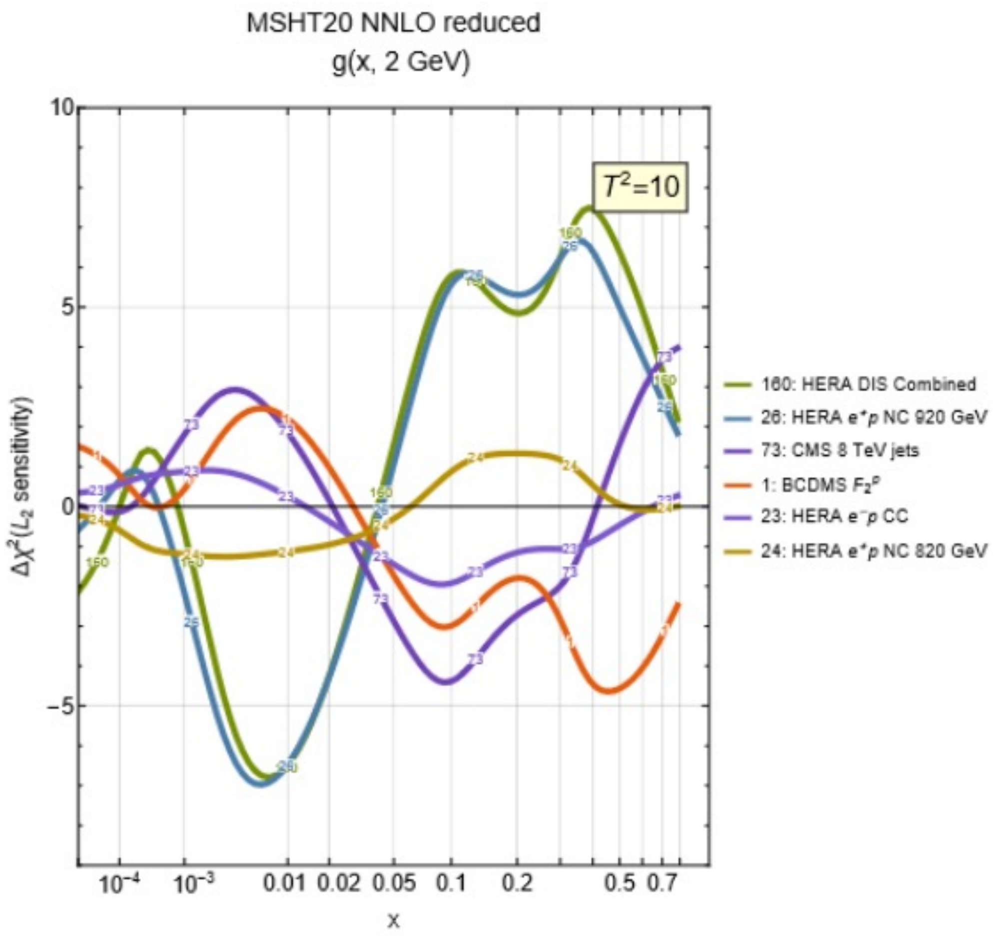
[Many typical χ^2/dof are >1.1 for >4000 points, or very unlikely from the pure statistical fluctuations. They reflect tensions among the experiments. In addition, the choice of PDF parametrization forms may change the PDFs without changing the χ^2].

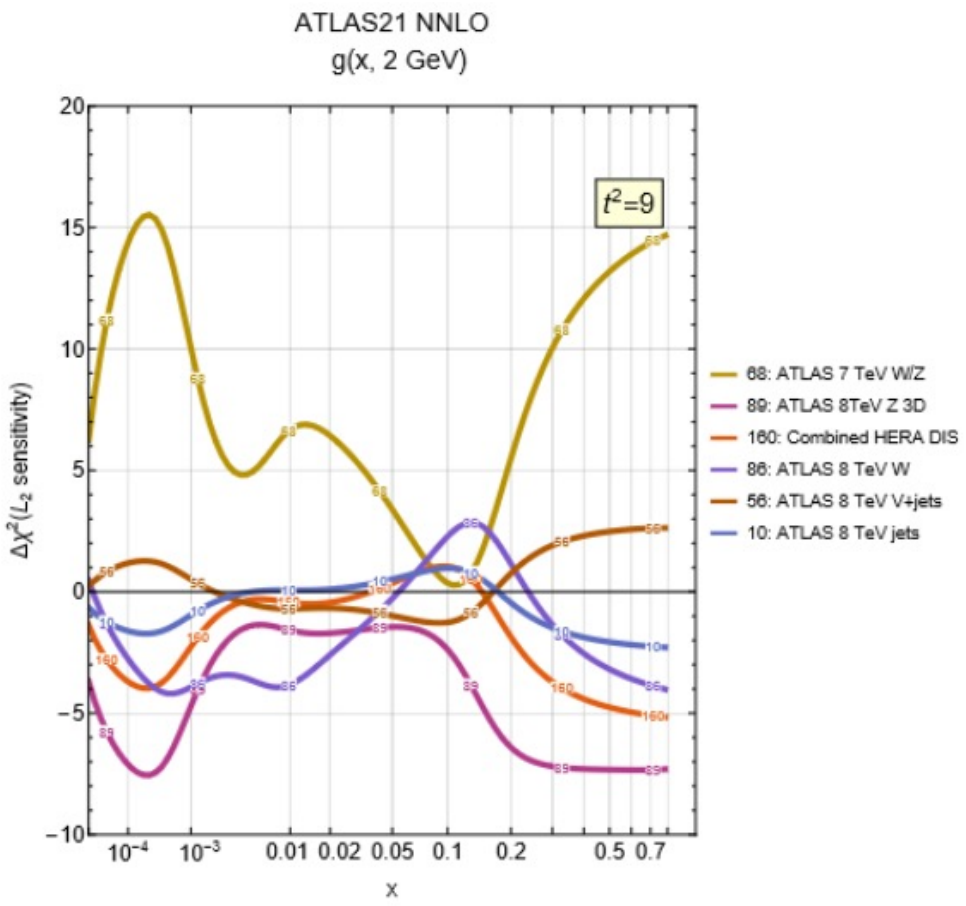
CT and MSHT use different criteria to account for the full uncertainty.

ATLAS, CT, MSHT comparative study: NNLO, aN3LO PDF sensitivities

Jing et al., arXiv:2306.03918







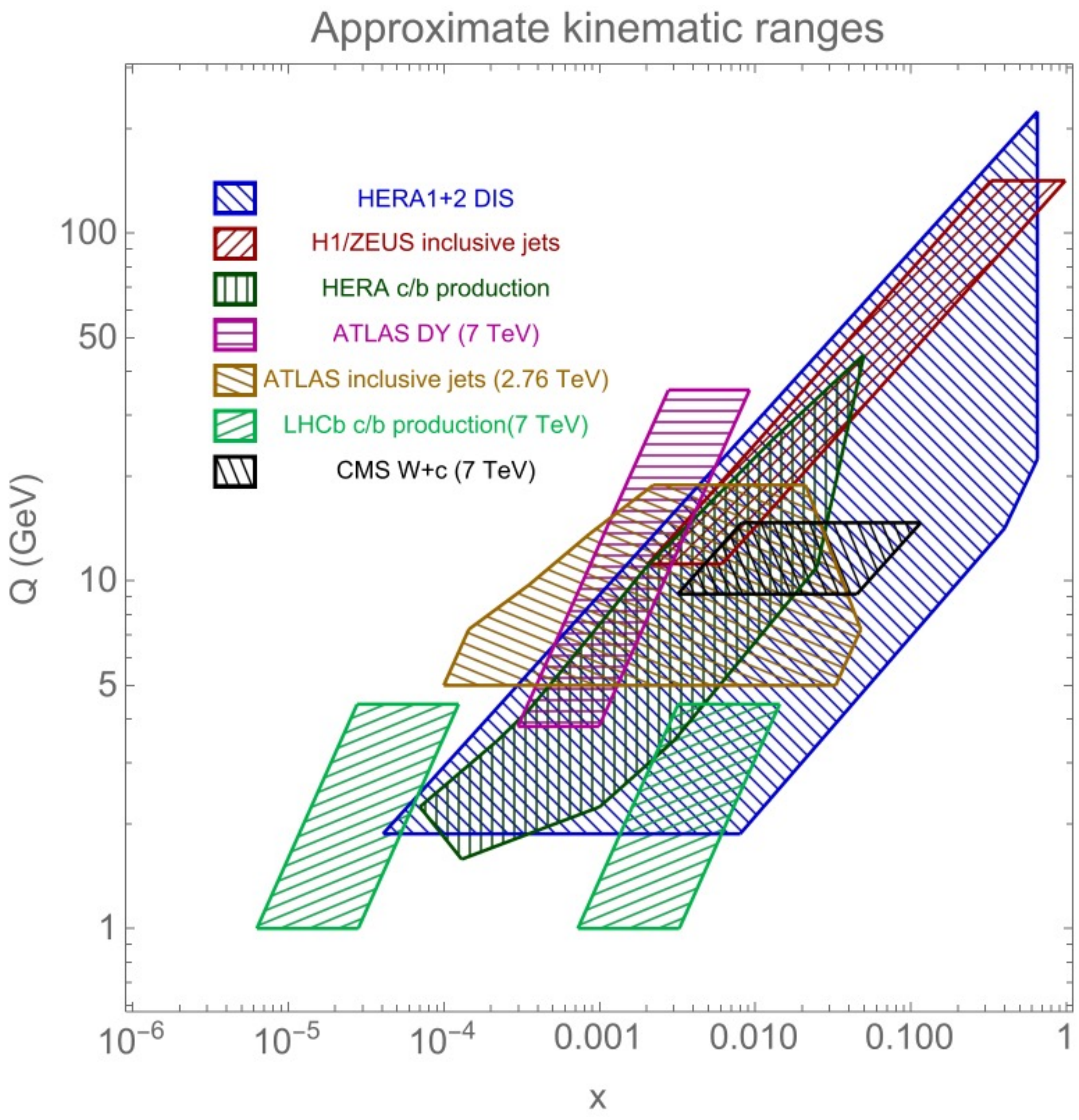
- comparisons of strengths of constraints from individual data sets in 8 PDF analyses using the common L_2 sensitivity metric. [Definitions in the backup.]
- interactive website (https://metapdf.hepforge.org/L2/) to plot such comparisons [2070 figures in total; a code L2LHAexplorer to plot L2 sensitivities for LHAPDF grids]

L. Kotz, 2401.11350

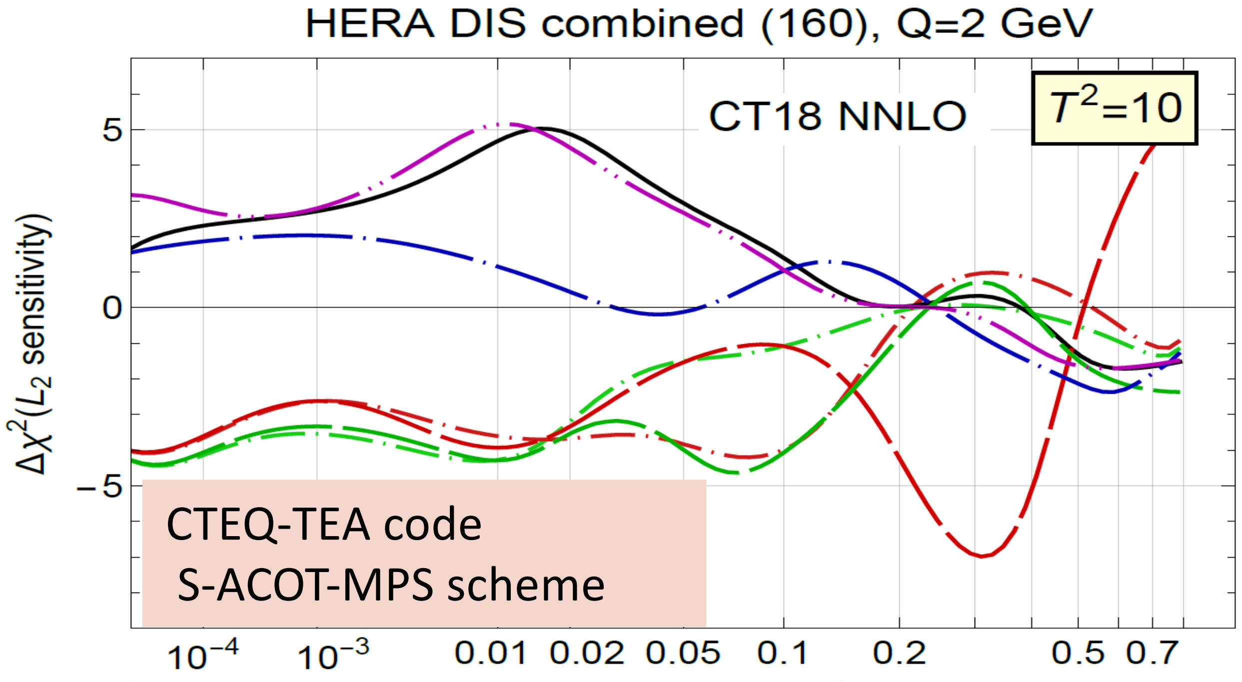
xFitter+L2LHAexplorer

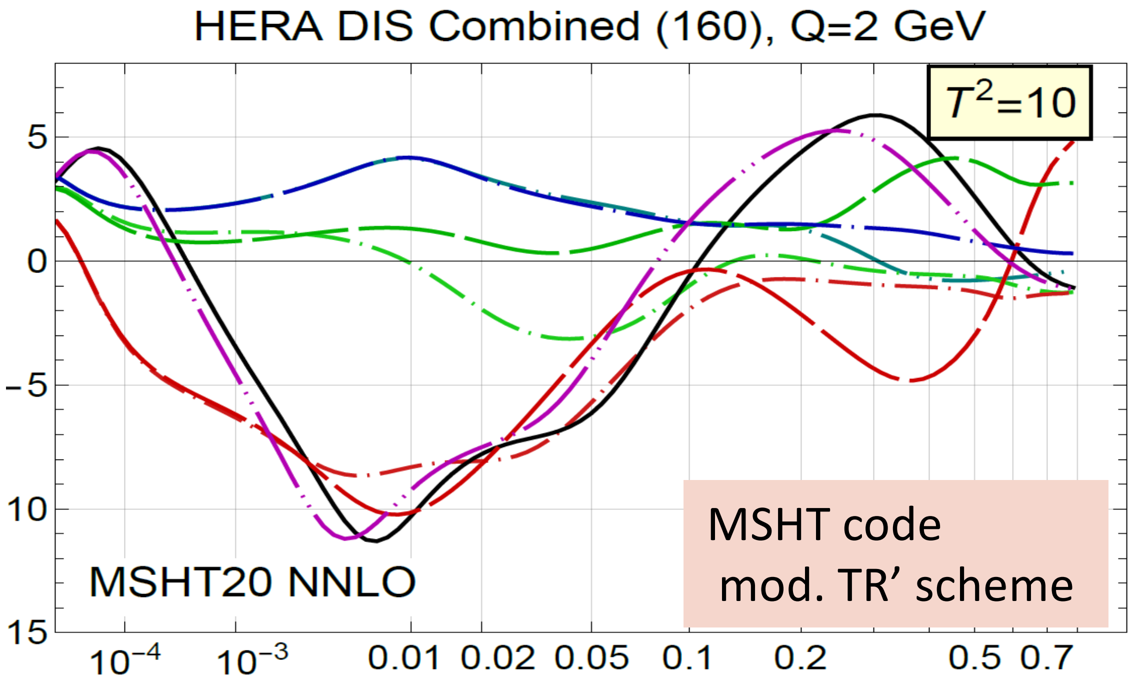
L2 sensitivities were computed using xFitter

- PDF sets (NNLO, $\alpha_s(M_Z) = 0.118$, $Q = 2\ \mathrm{GeV}$, $T^2 = 10$):
  - CT18
  - CT18As
  - MSHT20
- Data sets (included in xFitter):
  - ATLAS Drell-Yan ($\sqrt{s} = 7\ \mathrm{TeV}$)
  - ATLAS jet production ($\sqrt{s} = 2.76\ \mathrm{TeV}$)
  - CMS W+c production ($\sqrt{s} = 7\ \mathrm{TeV}$)
  - H1+ZEUS combined *c* and *b* production
  - H1 jet production
  - HERA I+II DIS
  - LHCb *c* and *b* production ($\sqrt{s} = 7\ \mathrm{TeV}$)
  - ZEUS jet production

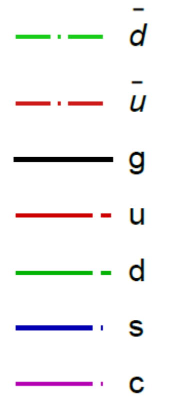


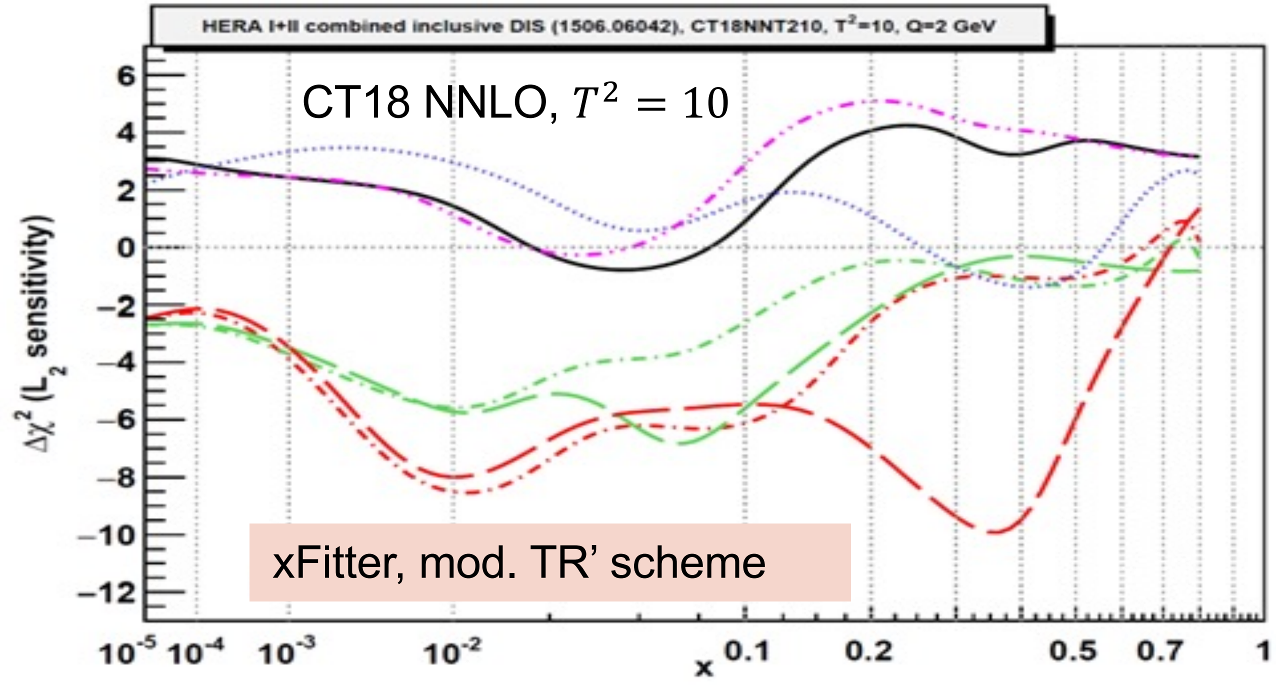
HERA I+II combined inclusive DIS [in CT18 and MSHT20]

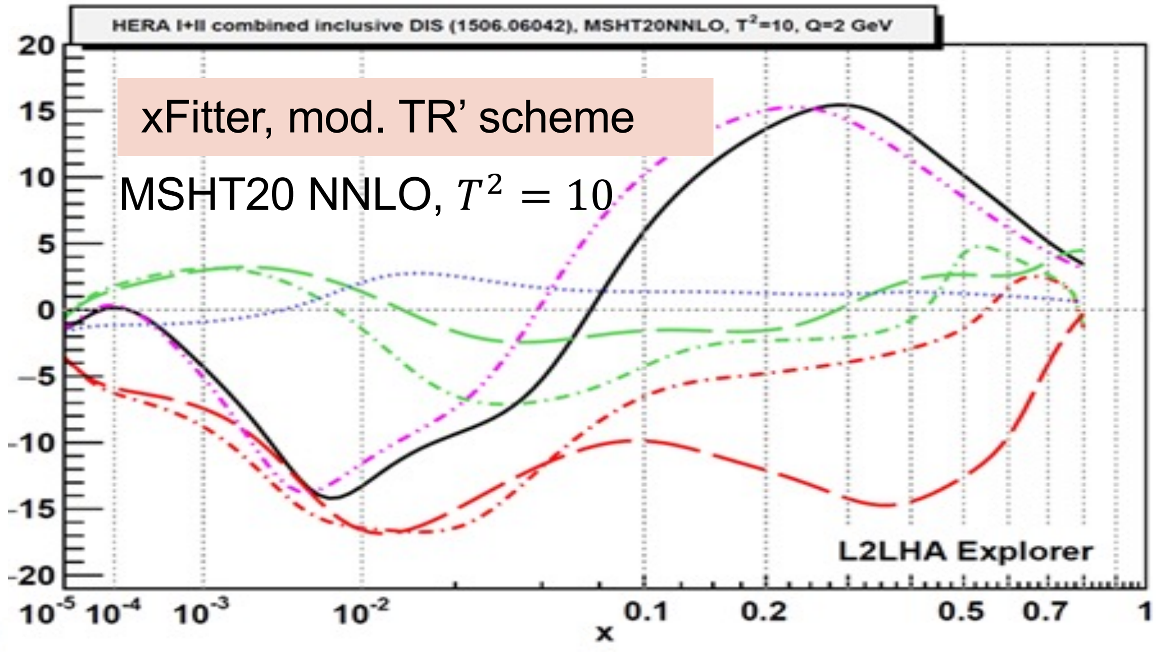




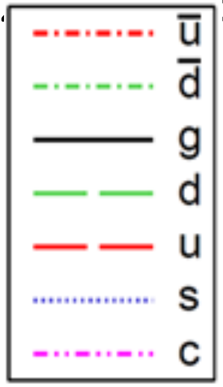
Upper row:
From Jing et al.,2306.03918







Lower row:
From L. Kotz,
2401.11350

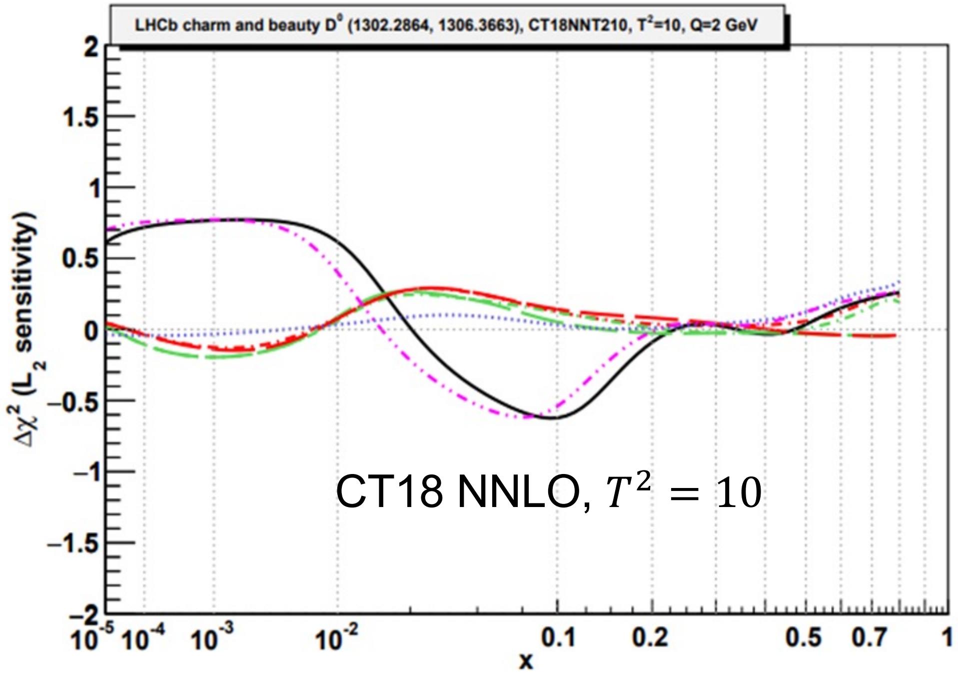


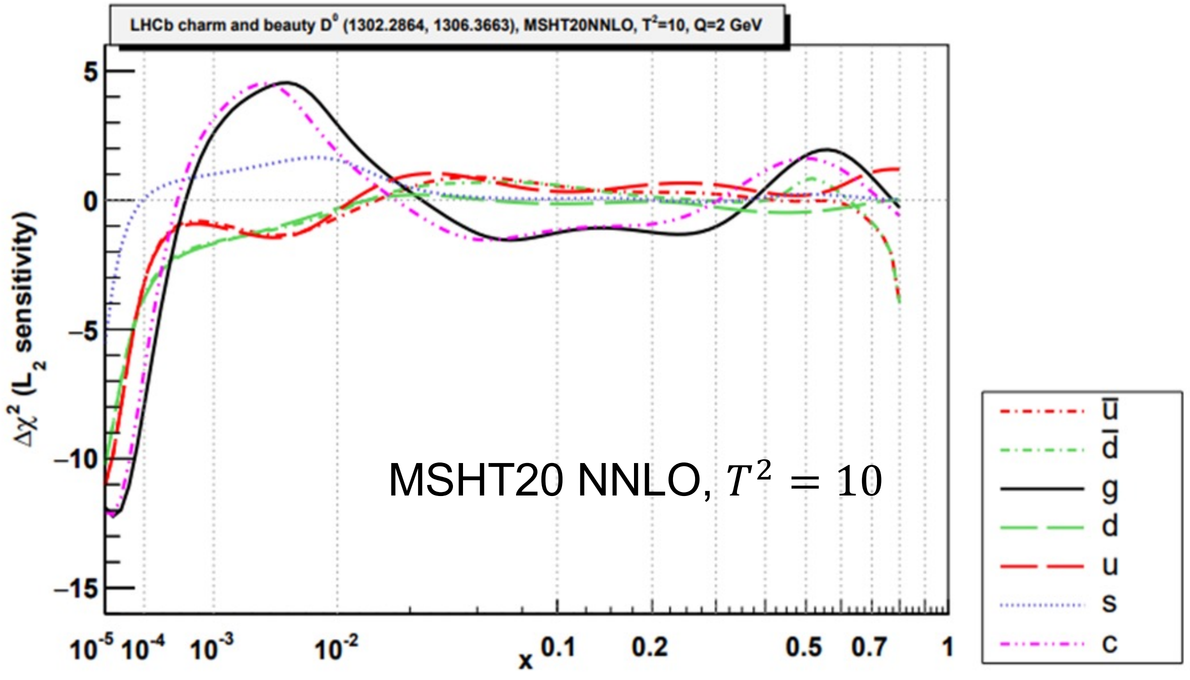
Left column: differences in χ^2 definition and heavy-quark scheme. Same PDFs and m_Q.

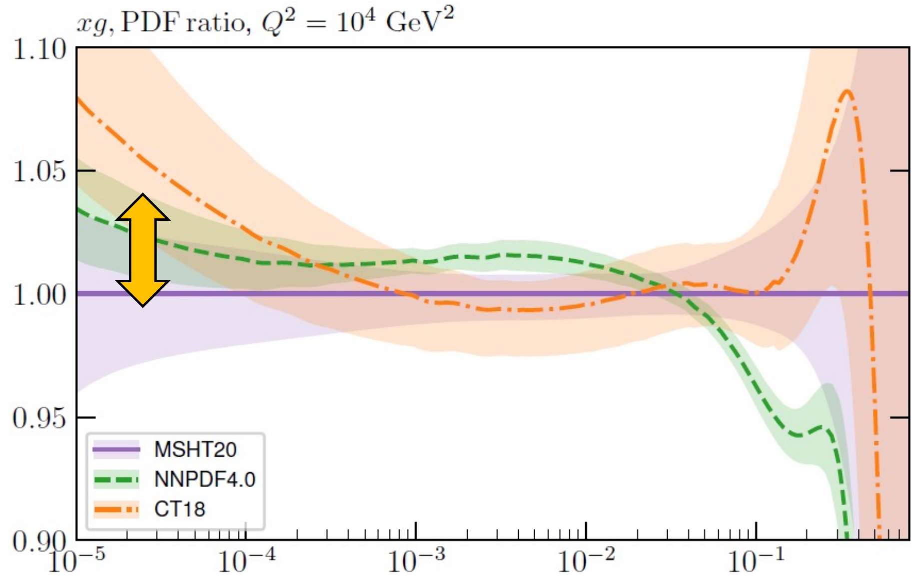
Right column: differences in χ^2 definition only. Same PDFs and m_Q.

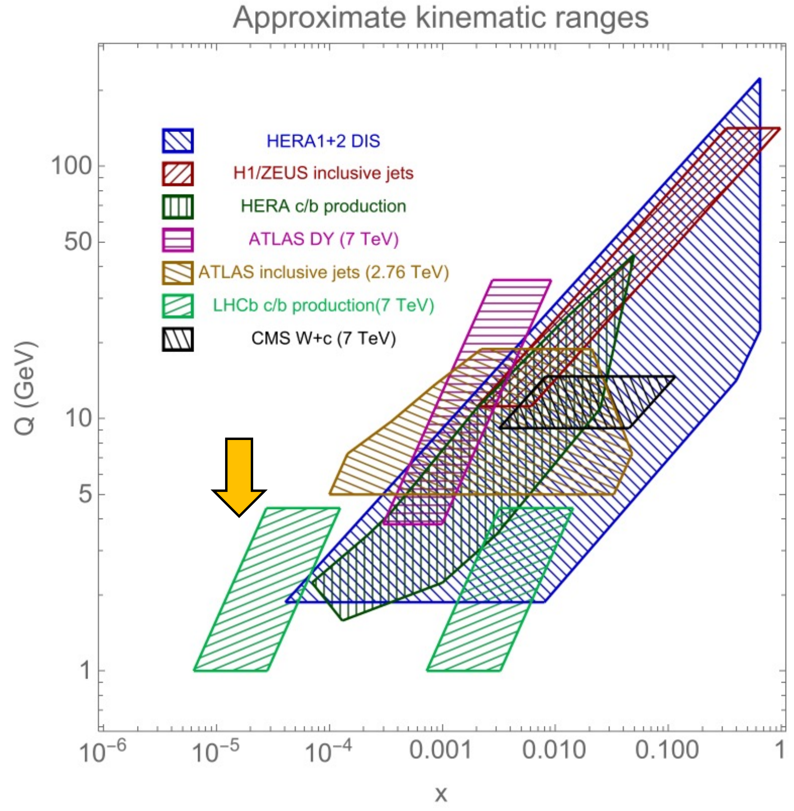
LHCb c and b @7 TeV; $p_T^{\text{meson}} \geq 2$ GeV [Not in CT18 or MSHT20]

From L. Kotz, 2401.11350





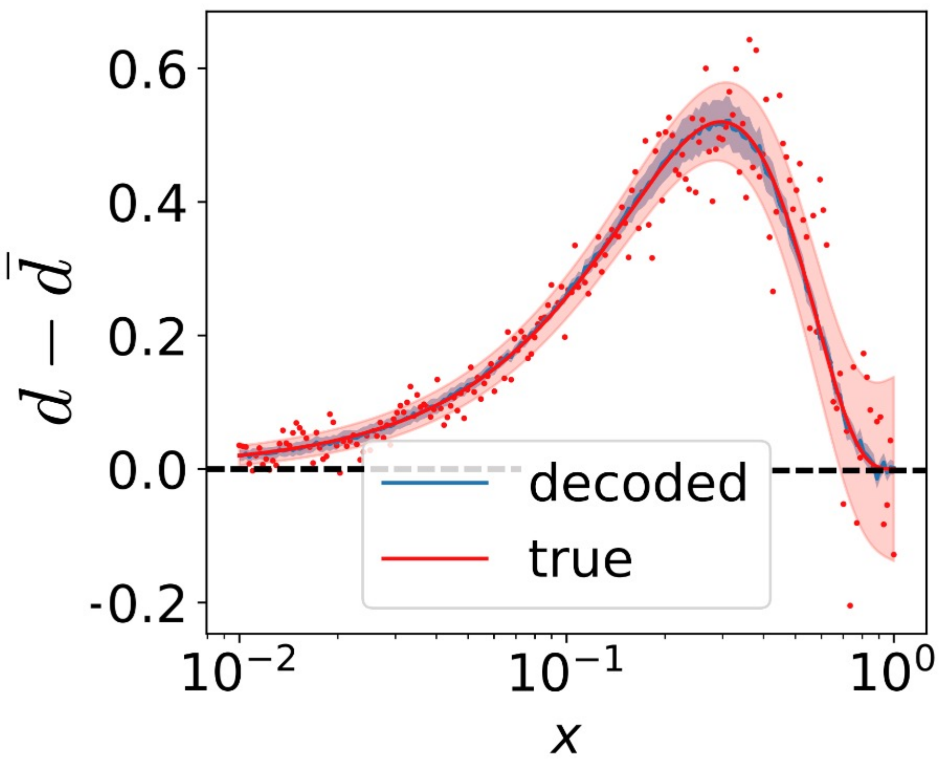


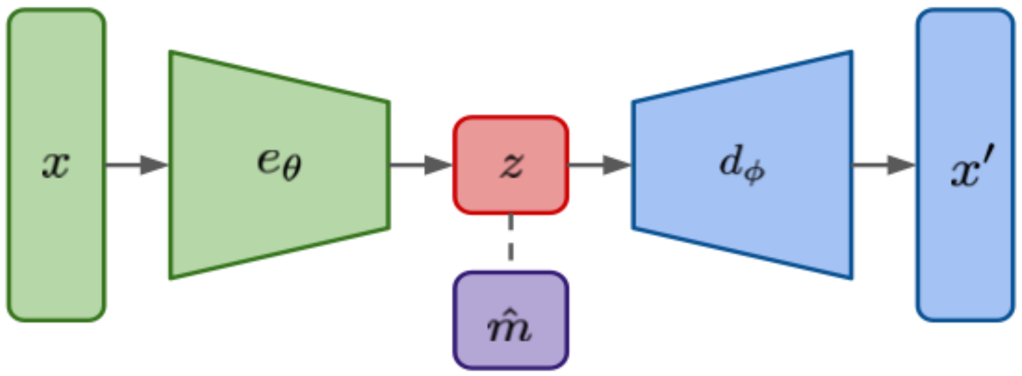


- LHCb c and b production prefers about the same gluon for CT18/CT18As.
- A larger gluon is preferred for MSHT20 at $x < 10^{-4}$.

ML models for PDF generation

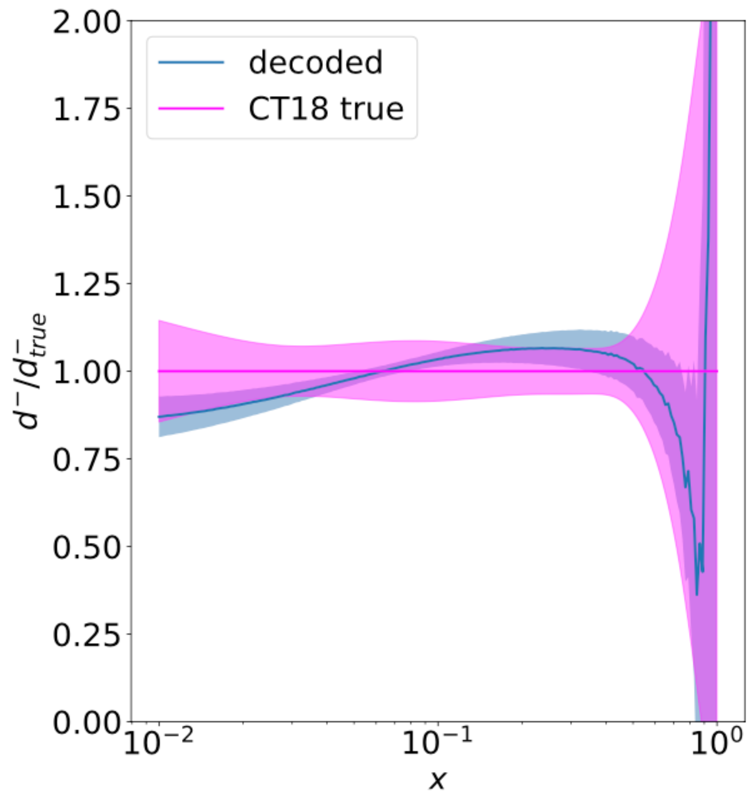
Kriesten and TJH, arXiv: 2312.02278





$$\mathcal{L} = \|x - d_\phi(e_\theta(x))\|_2^2 + \|z - \hat{m}\|_2^2$$

PDFdecoder, package available



- autoencoder-based models can efficiently represent PDFs in dimensionally reduced form
- through careful choice of network topology, can impose interpretable structure on latent space
- physics constraints may include PDFs' Mellin-space behavior (i.e., integrated moments)
- trained models (like VAIM at right) can generatively predict PDFs from moments

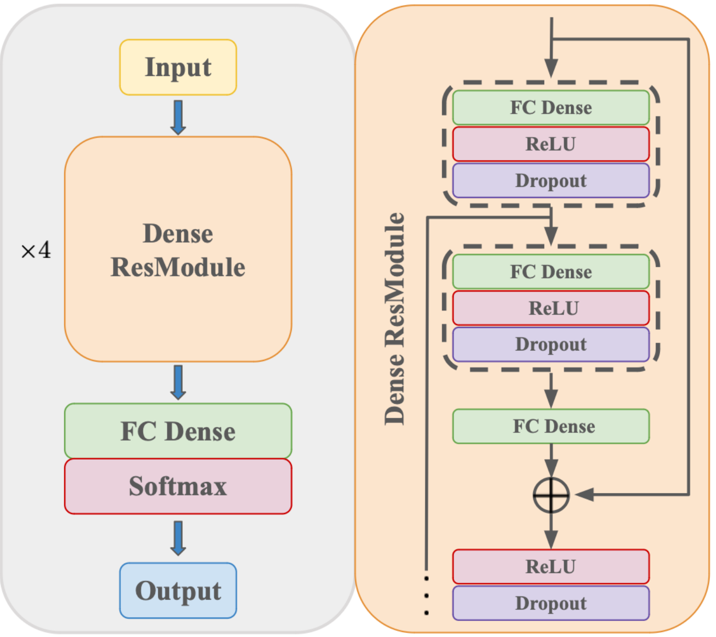
→ new ML tool to mutually compare PDFs, explore statistical properties (e.g., out-of-distribution behavior)

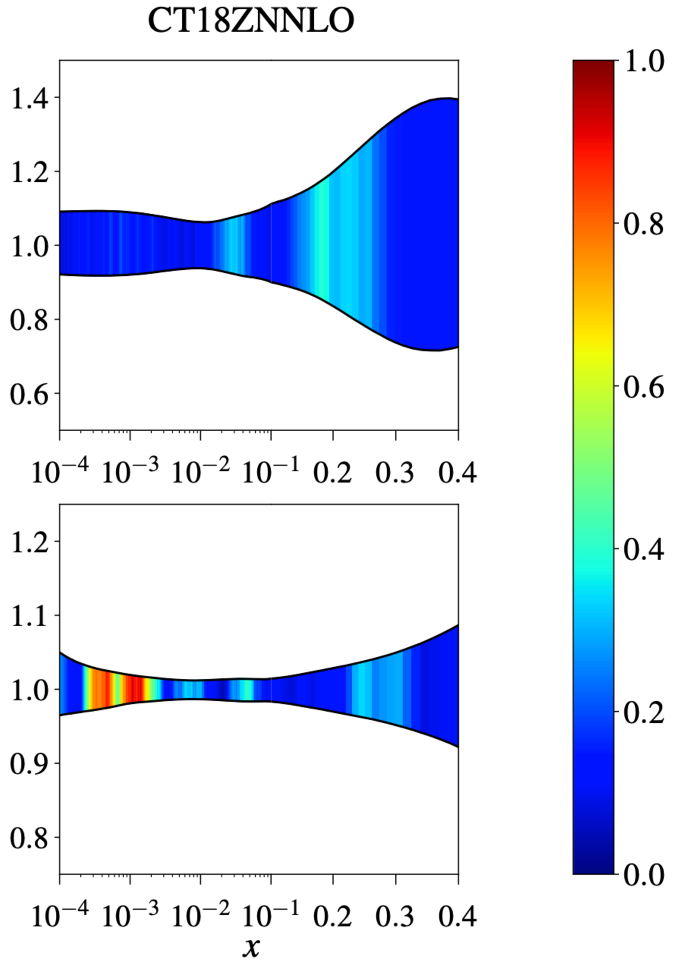
additional AI4PDF developments: explainability, evidential learning

see talk: **Brandon Kriesten**, Monday afternoon

- classify PDFs by theoretical model, parent fit
- guided backprop: link *x*-dependent features PDF to classification

Kriesten, Gomprecht, TJH, arXiv: 2407.03411

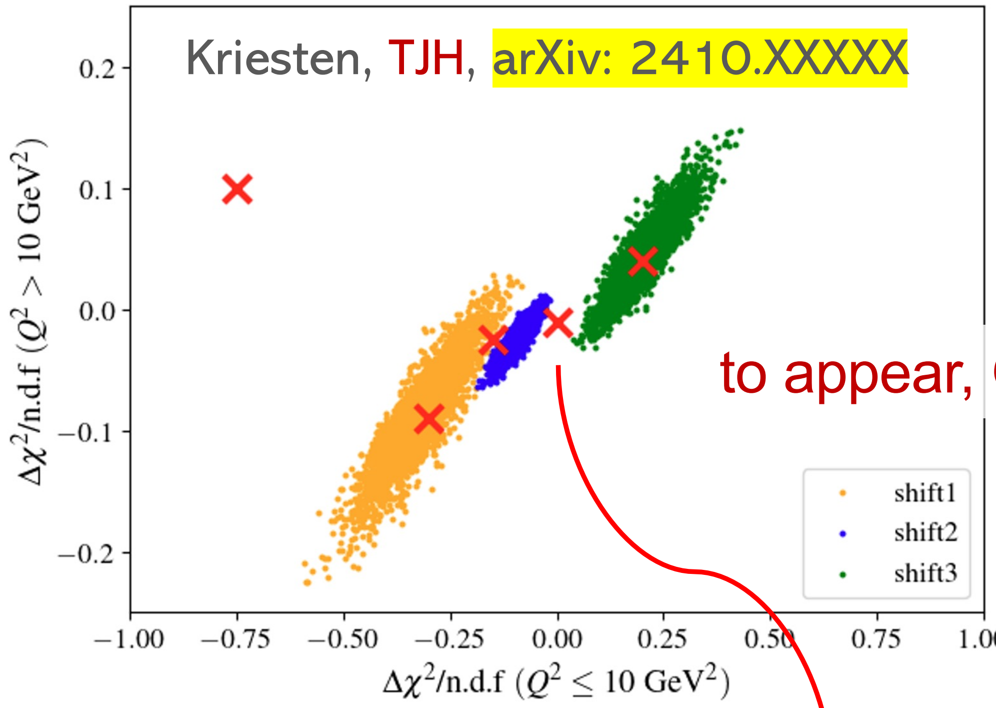




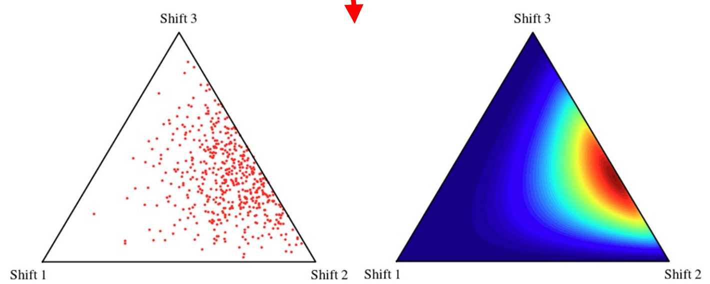
XAIPDF, package available

- train on CT18 PDF replicas alongside nonstandard interactions; quantify aleatoric, epistemic uncertainties in neutrino scattering

Kriesten, TJH, arXiv: 2410.XXXXX



to appear, Oct 2024



Recent progress, near-future plans

1. Final selection of experiments for NNLO PDFs planned for the next year
2. Work on N3LO contributions and implementations
3. Next-generation PDF uncertainty quantification
4. Recent and imminent PDF releases
   a. QCD+QED PDFs for protons and neutrons
   b. Subtracted S-ACOT-MPS PDFs
   c. Fantômas 1.0 pion PDFs (Hessian)
   d. Release of the Fantômas PDF parametrization package in xFitter
   e. Release of PDFdecoder, XAIPDF, and related ML packages