

Toward the next CTEQ-TEA parton distributions

Marco Guzzi

Kennesaw State University, USA

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

China: A. Ablat, S. Dulat, Y. Fu, T.-J. Hou, I. Sitiwaldi

Mexico: A. Courtoy

USA: P. Nadolsky, M.G., T.J. Hobbs, J. Huston, H.-W. Lin, C. Schmidt, K. Xie, C.-P. Yuan

and other coauthors



LoopFest 2024 Tue, May 21, Southern Methodist University

Importance of high-order perturbative calculations



Large inflow of high-precision data from current (and possible, hopefully) future colliders necessitates precise and accurate description to relate experimental observations to underlying theory.

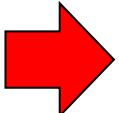
Progress is required in both multi-loop calculations and global PDF analyses:

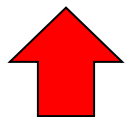
- Higgs boson & EW sector
- Fundamental parameters, coupling constants, particles masses, ...
- BSM Physics searches
-

Importance of high-order perturbative calculations

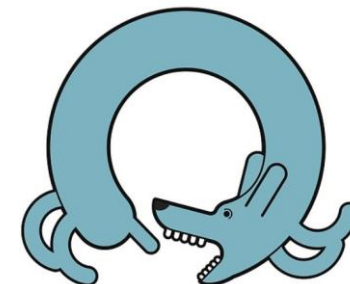


Latest developments in precision QFT and applications to hard-scattering calculations are crucial to bring PDF analyses to the next level of accuracy and precision for an improved understanding of experimental data.

Improved PDFs  Improved theory predictions to be compared to data



- Compatible high-precision experimental data
- Developments in loop calculations for DGLAP evolution (**N3LO not yet complete**) and hard scatterings
- **Fast calculations (critical for global QCD analyses)**
- Improvements in current understanding of uncertainties



Importance of high-order perturbative calculations



A lot of efforts within the CT group are going into understanding PDF uncertainties which, in turn, affect uncertainties of hadronic observables.

In particular, the impact of tolerance criteria to define PDF uncertainties.

Uncertainties related to parametric functional forms and fitting methodology must be fully accounted for and not be underestimated (see discussion in Courtoy et al., 2205.10444)

• 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)
 - ...

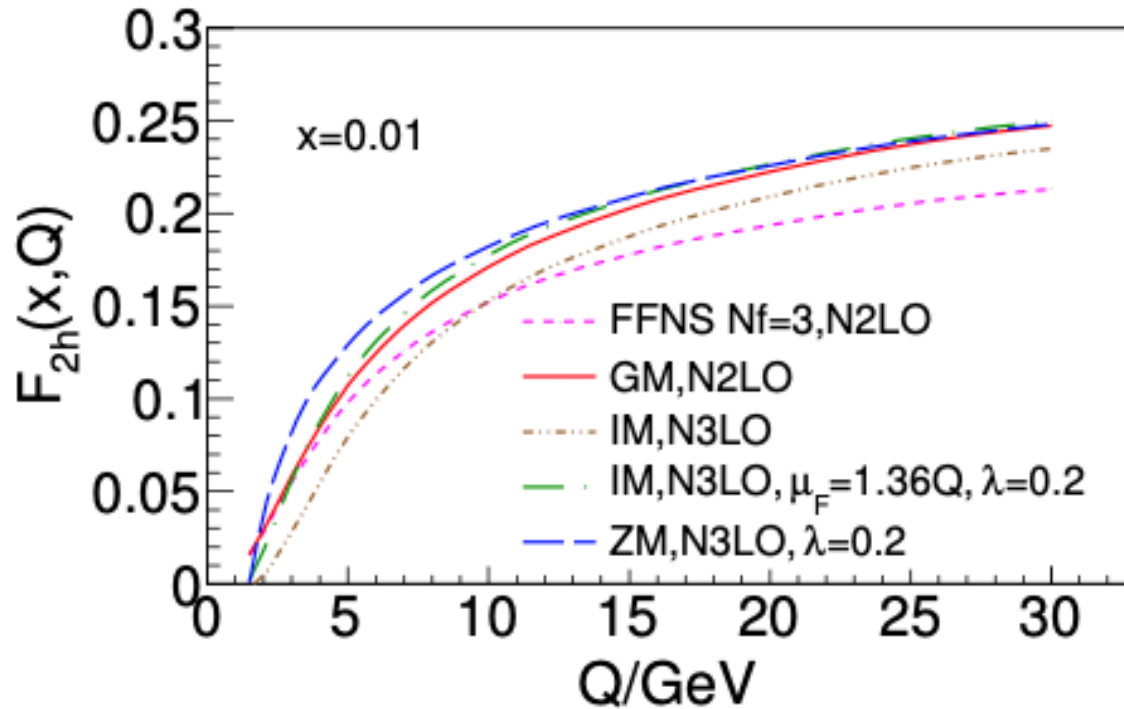
K. Xie's Talk (Xie, M.G., Nadolsky, Reina, Wackerroth)

} See A. Courtoy, DIS2024

Toward a new generation of CT202X PDFs

1. Multiple preliminary NNLO fits with LHC Run-2 (di)jet, vector boson, $t\bar{t}$ data
 - based on the selections of experiments recommended in [2305.10733](#), [2307.11153](#)
2. Work on implementation of N3LO contributions
3. Work on GMVFN schemes to have consistent HQ treatment
4. Next-generation PDF uncertainty quantification: Bézier curves, META combination, ML stress-testing, multi-Gaussian approaches, ...
5. Physics applications
 - a. QCD+QED PDFs for a neutron
 - b. PDF dependence of forward-backward asymmetry
 - c. An L2 sensitivity study using xFitter
 - d. Pion PDFs
 - e. ...

QCD cross sections @N3LO



- **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)

Boting Wang's and Keping Xie's Theses, SMU

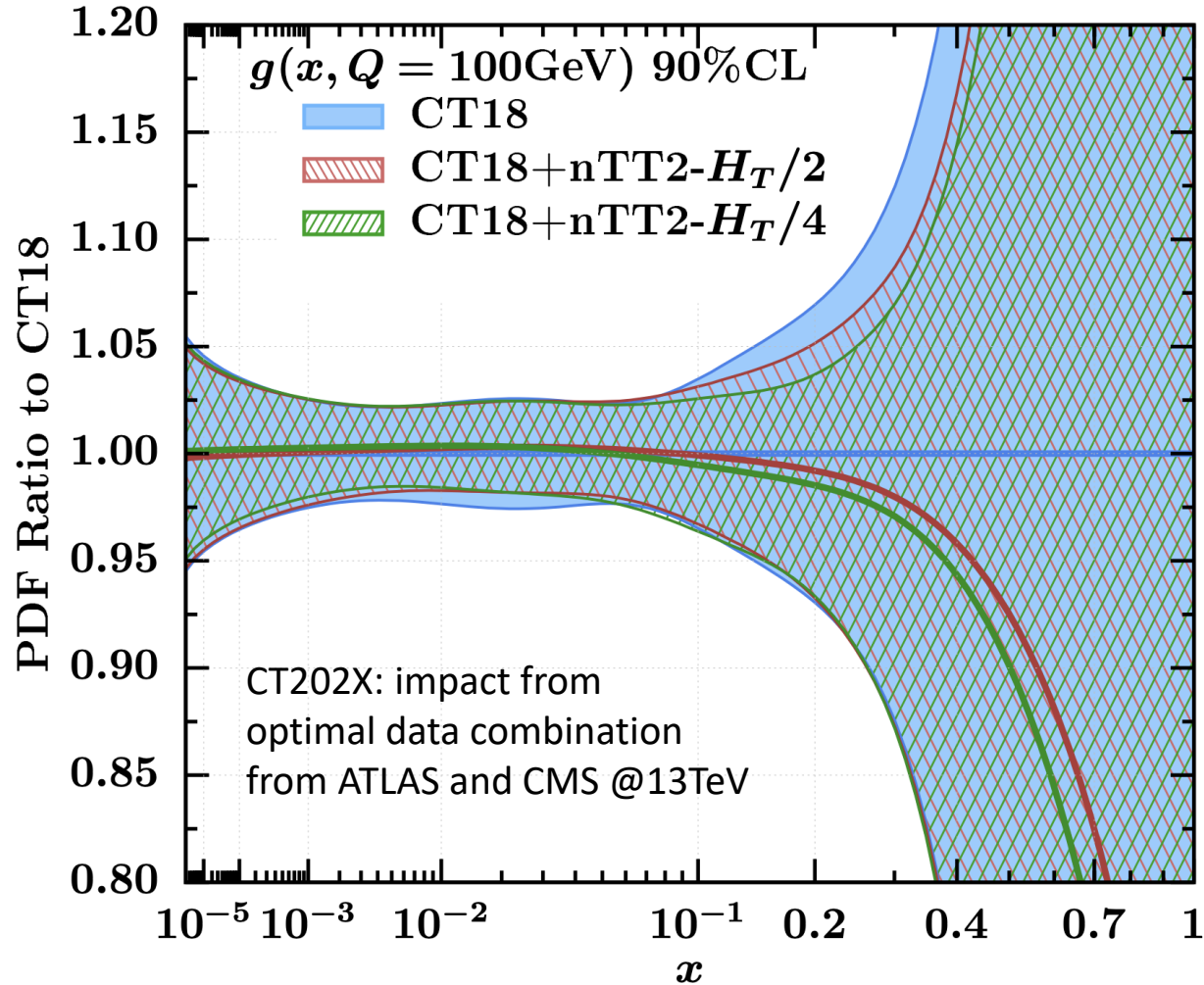
- **Imminent implementation of massive N3LO heavy-quark coefficients to obtain N3LO DIS cross sections in the SACOT-MPS General-Mass VFN scheme**

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 approx. N3LO with APFEL/APFEL++.
- Drell-Yan: Ongoing work to include N3LO DY effects using NNLO ApplFast + N3LO/N2LO K-factor tables

NNLO PDFs with new LHC data at 8 and 13 TeV

Impact of LHC 13 TeV $t\bar{t}$ production on CT PDFs (Ablat, M.G., Xie, Dulat, Hou, Sitiwaldi, Yuan, PRD109 2024; arXiv:2307.11153)



Theory predictions:

- MATRIX (Catani, Grazzini et al. PRD 2019)
- FastNNLO (Czakon, et al. 1704.08551)

Blue band: CT18NNLO 90% C.L.

Hatched bands: CT18 + new top-quark data

Green: $\mu_R = \mu_F = H_T/2$

Red: $\mu_R = \mu_F = H_T/4$

Differences related to different scale choices and are well within the CT18 PDF error band.

nTT2 baseline consists of 1D abs $t\bar{t}$ Xsec from:

- ATLAS all hadronic, $y_{t\bar{t}}$
- ATLAS lepton + jets, $\{y_{t\bar{t}}, M_{t\bar{t}}, y_{Bt\bar{t}}, H_{Tt\bar{t}}\}$ stat. comb.
- CMS dilepton, $y_{t\bar{t}}$
- CMS lepton + jets, $M_{t\bar{t}}$

NNLO PDFs with new LHC data at 8 and 13 TeV

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

$t\bar{t}$ + Drell-Yan + Jet production @NNLO

ID	Exp	N_{pt}	χ^2/N_{pt}
Drell-Yan			
215	ATLAS 5.02 TeV W,Z	27	$0.82^{+0.55}_{-0.16}$ ($1.15^{+1.22}_{-0.43}$)
211	ATLAS 8 TeV W	22	$2.42^{+2.49}_{-1.51}$ ($4.25^{+6.39}_{-3.34}$)
214	ATLAS 8 TeV Z3D	188	$1.12^{+0.46}_{-0.02}$ ($1.99^{+5.10}_{-1.85}$)
212	CMS 13 TeV Z	12	$2.48^{+4.76}_{-0.88}$ ($12.03^{+38.04}_{-21.84}$)
217	LHCb 8 TeV W	14	$1.35^{+0.59}_{-0.61}$ ($1.35^{+0.72}_{-0.64}$)
218	LHCb 13 TeV Z	16	$1.18^{+1.42}_{-0.60}$ ($1.49^{+1.74}_{-0.89}$)
13 TeV $t\bar{t}$			
521	ATLAS all-hadronic $y_{t\bar{t}}$	12	$1.06^{+0.14}_{-0.09}$ ($1.05^{+0.21}_{-0.10}$)
528	CMS dilep $y_{t\bar{t}}$	10	$1.10^{+1.08}_{-0.68}$ ($1.03^{+1.60}_{-0.74}$)
587	ATLAS lep+Jet $m_{t\bar{t}} + y_{t\bar{t}} + y_{t\bar{t}}^B + H_T^{t\bar{t}}$	34	$0.92^{+0.32}_{-0.14}$ ($0.94^{+0.59}_{-0.16}$)
581	CMS lep+jet $m_{t\bar{t}}$	15	$1.44^{+1.18}_{-0.73}$ ($1.37^{+1.86}_{-0.82}$)
Inclusive Jet			
553	ATLAS 8 IncJet	171	$1.76^{+0.20}_{-0.12}$ ($1.80^{+0.33}_{-0.16}$)
554	ATLAS 13 IncJet	177	$1.38^{+0.13}_{-0.10}$ ($1.39^{+0.20}_{-0.11}$)
555	CMS 13 IncJet	78	$1.10^{+0.24}_{-0.17}$ ($1.11^{+0.30}_{-0.16}$)

nDY

nTT

nIncJet

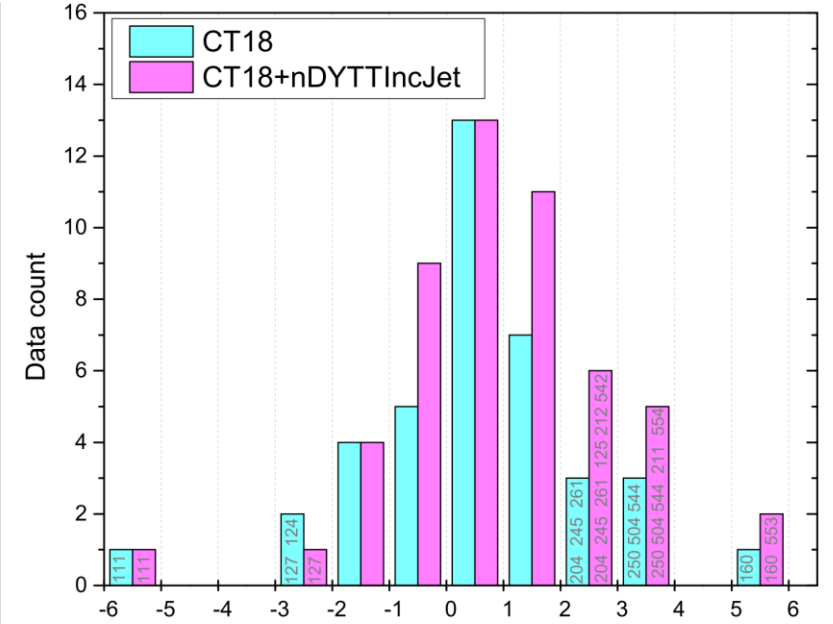
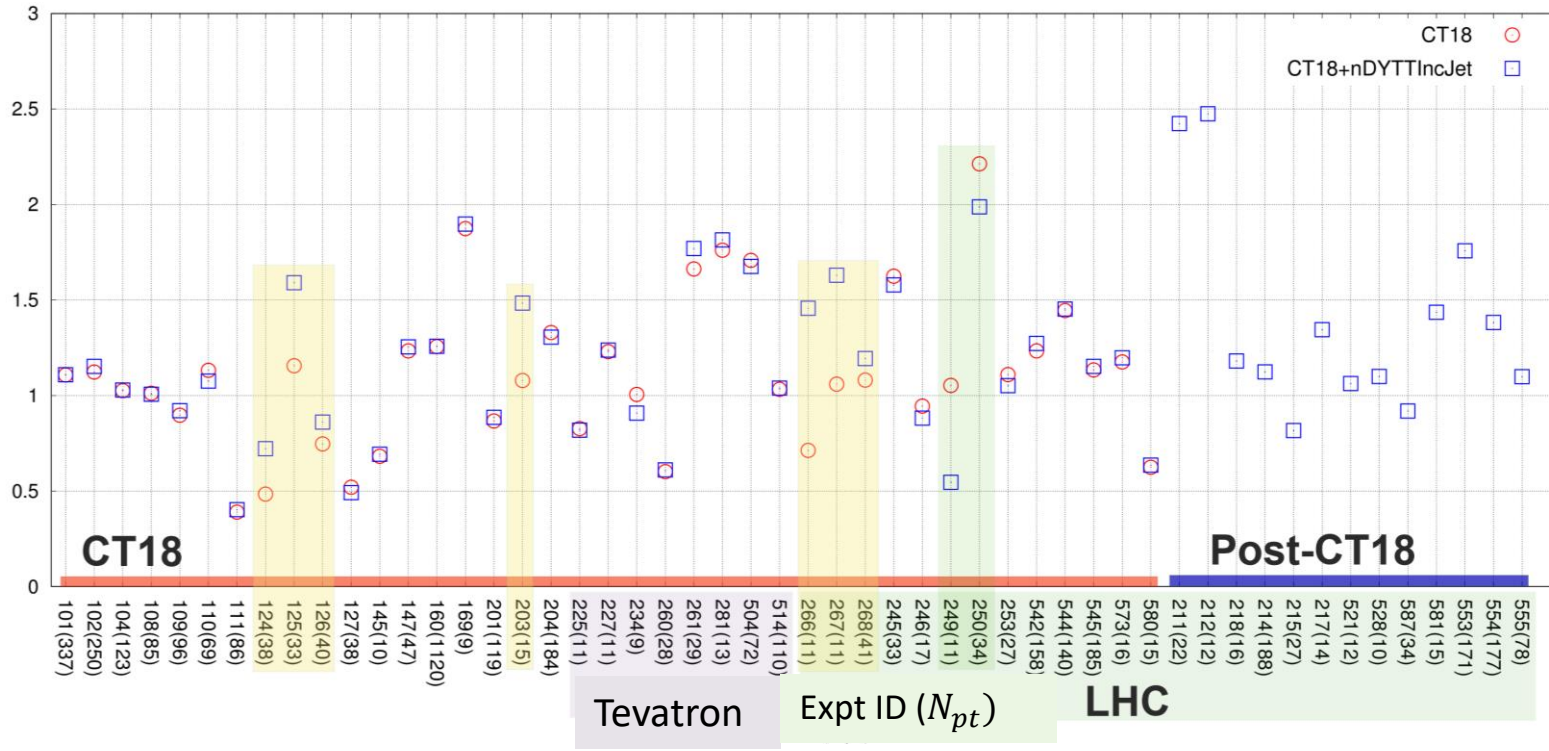
nDYTTIncJet

Fits with 1 type of new data

A fit with all 3 types

A 3-data-type fit (CT18+nDYTTIncJet)

χ^2/N_{pt}



$$S_n \approx (\chi^2 - N_{pt}) / \sqrt{2N_{pt}}$$

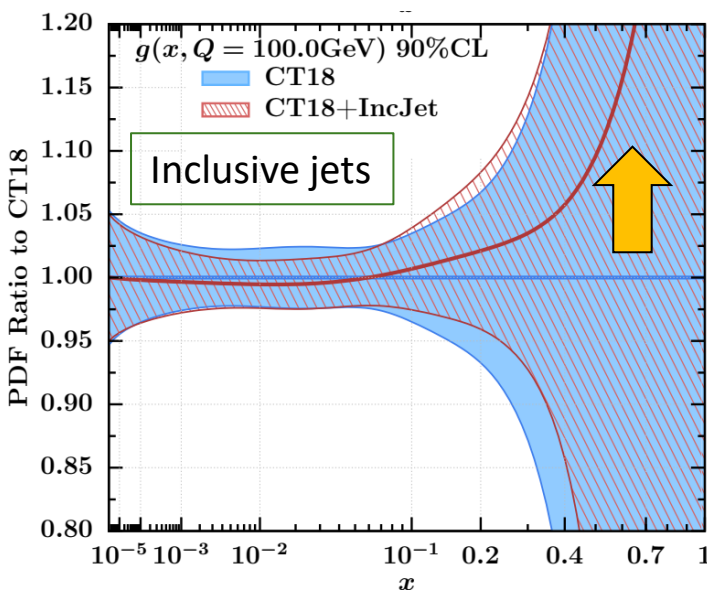
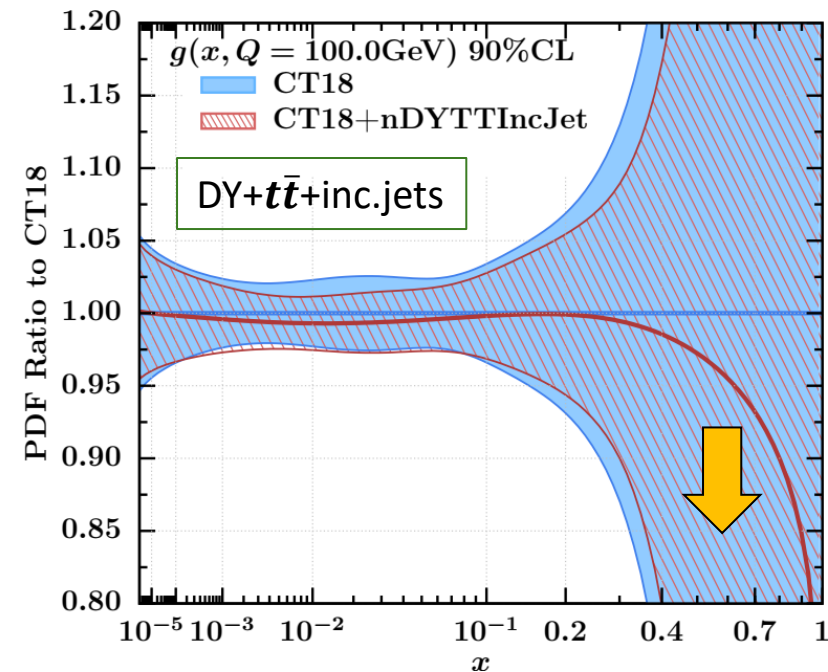
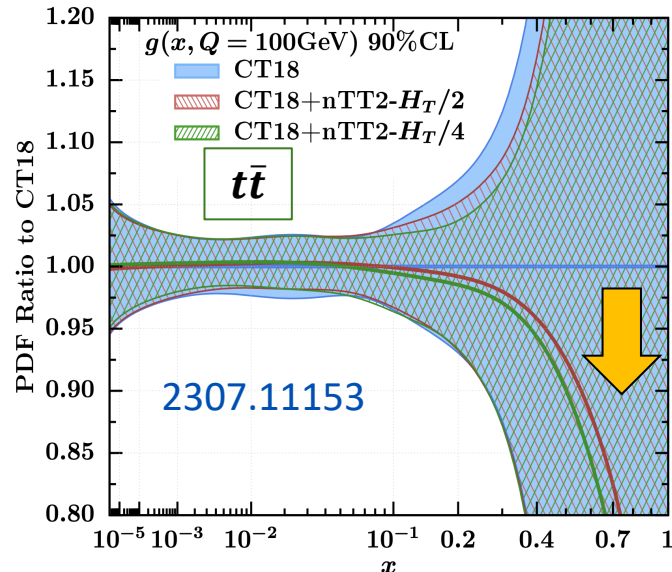
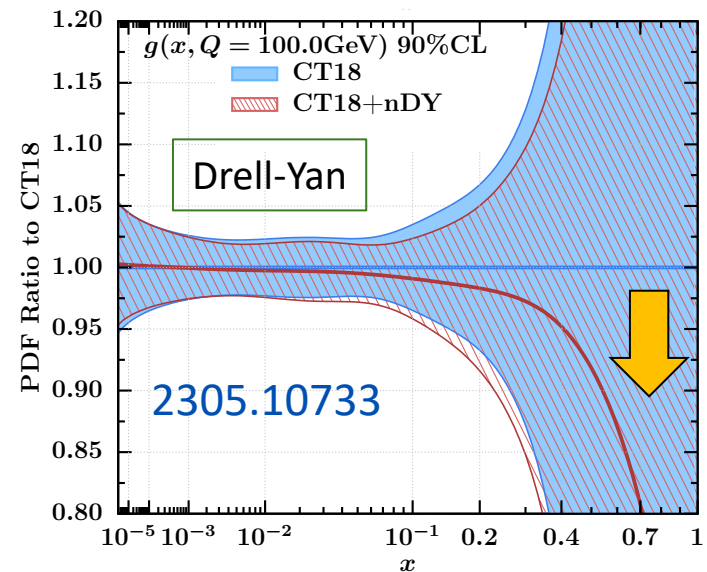
CT10, 1007.2241, PRD 2010

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 7TeV Ach), 268 (ATLAS 7TeV W, Ach).

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

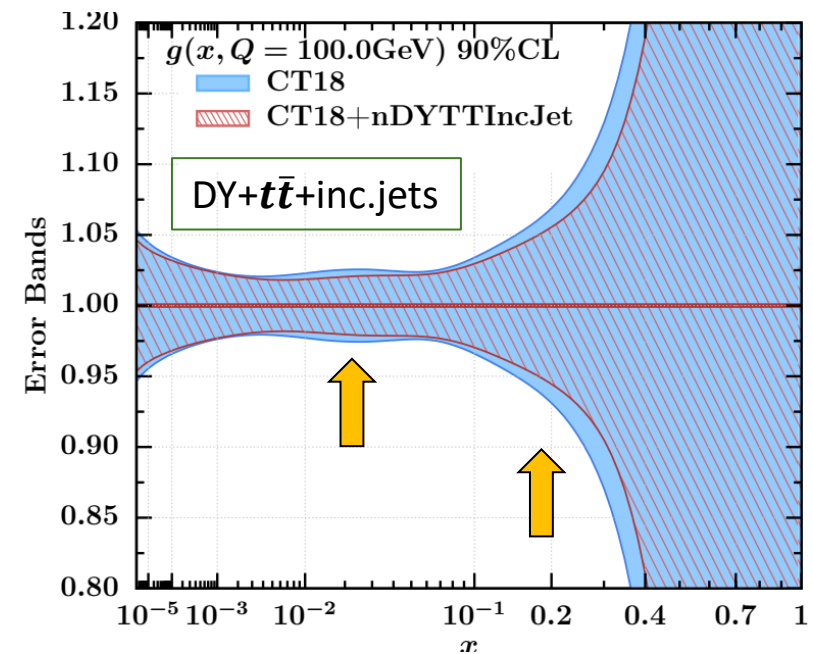
Pulls on the gluon PDF by the new data type



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.

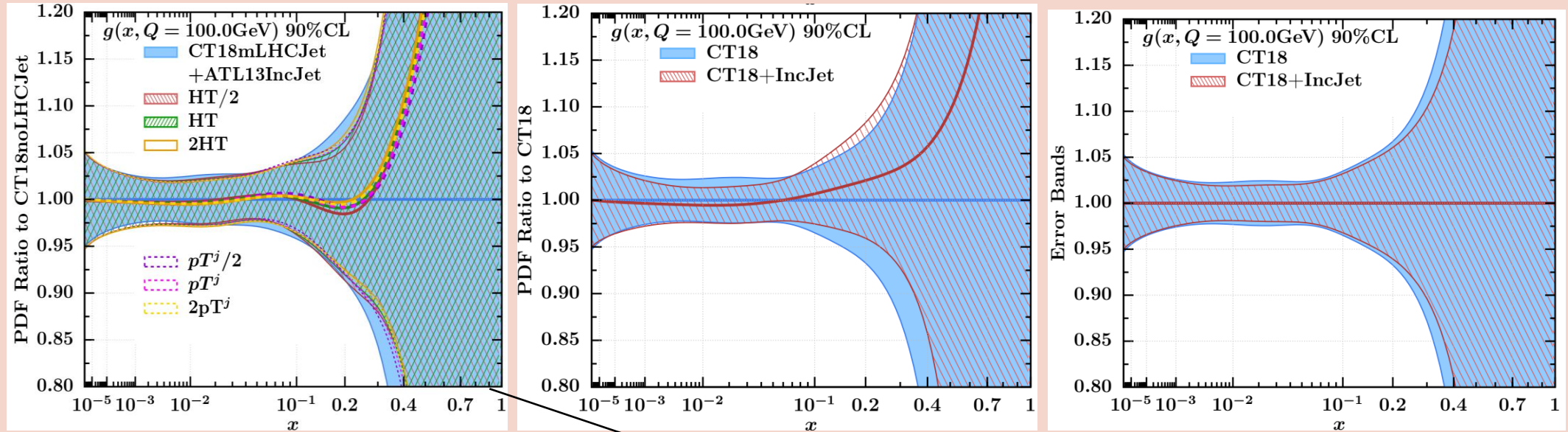
Mild changes in the gluon uncertainty

PRELIMINARY

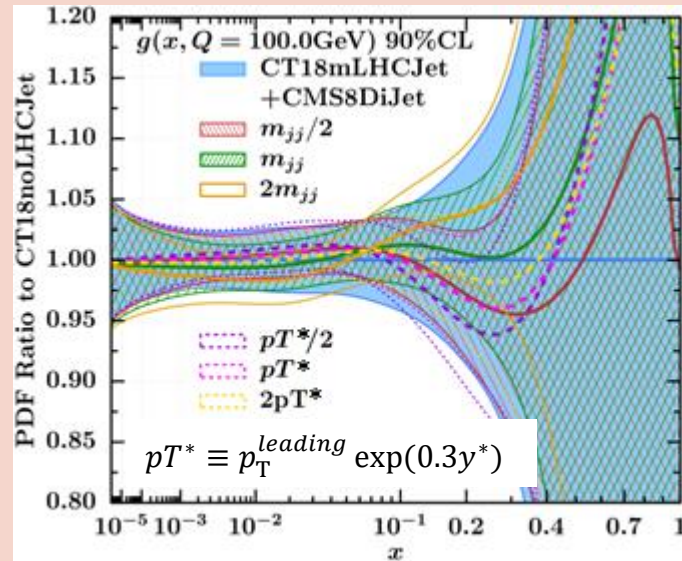


Inclusive jet vs. dijet data sets: impact on the gluon for various QCD scales

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



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



The impact of the Inc. jet data 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 .

The impact of dijet data substantially depends on scale choices, especially in the case of CMS8 TeV dijet.

PRELIMINARY

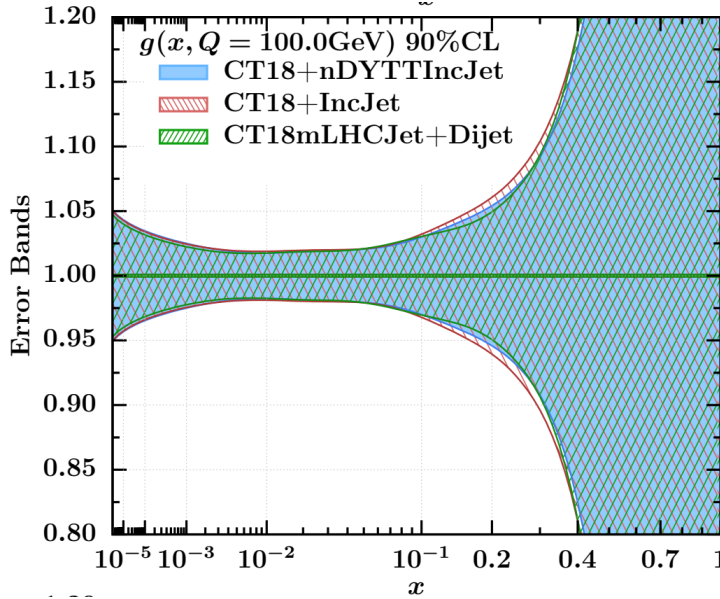
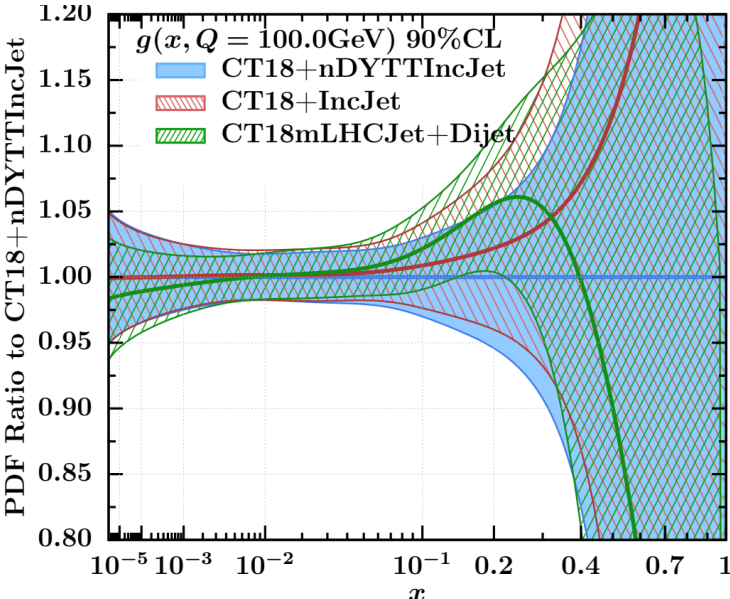
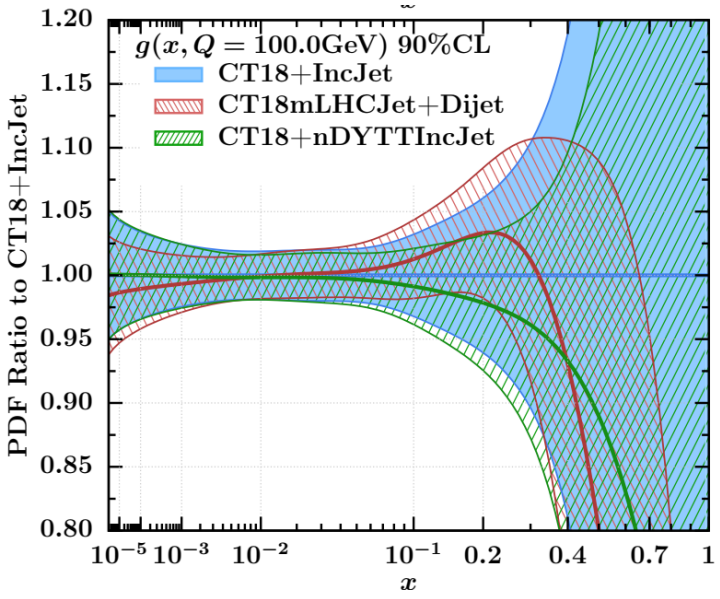
χ^2/N_{pt} for fits that add one inclusive jet or dijet data set to the CT18 (without LHC jets) baseline at a time

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



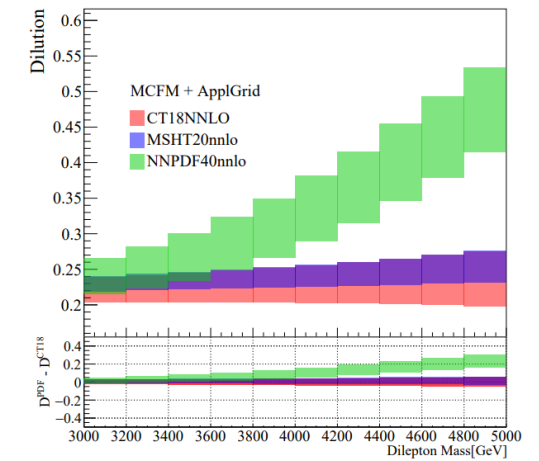
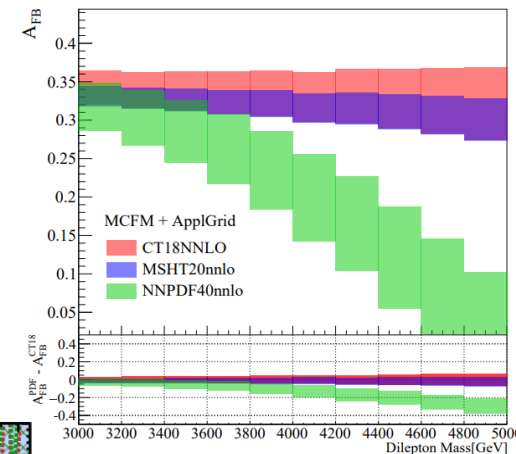
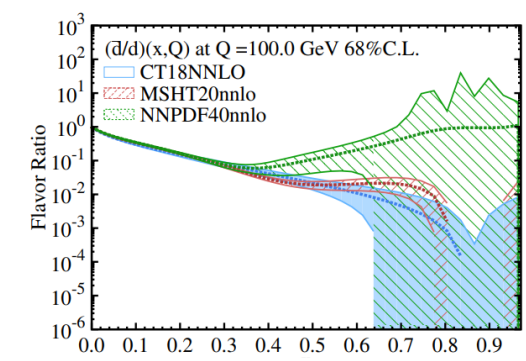
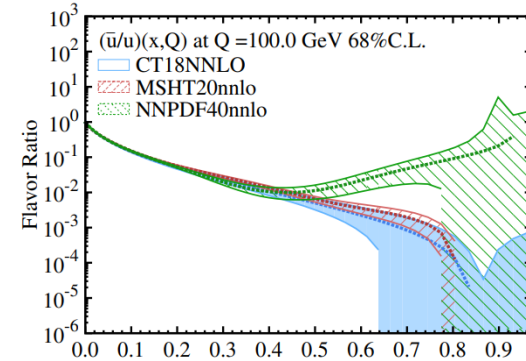
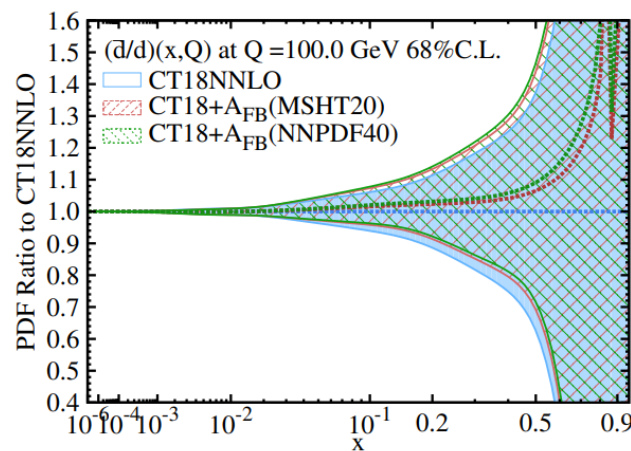
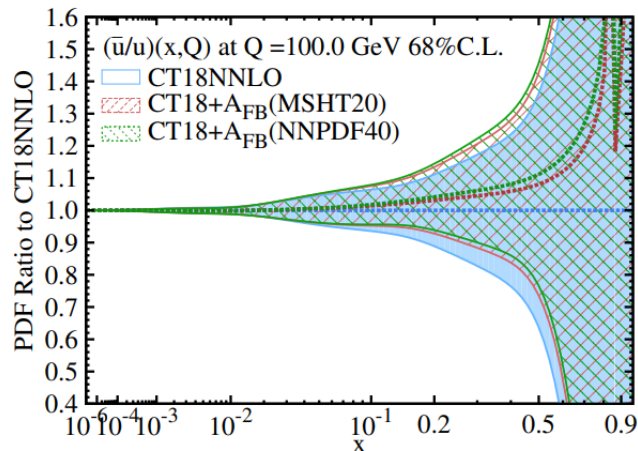
Dijet data sets tend to have larger uncertainties than inc. jets, facilitating better χ^2 for similar constraints on PDFs

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

- A_{FB} at the LHC is sensitive to the energy dilution factor D (probability of $k_q^0 < k_q^0$ 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 $x > 0.2$



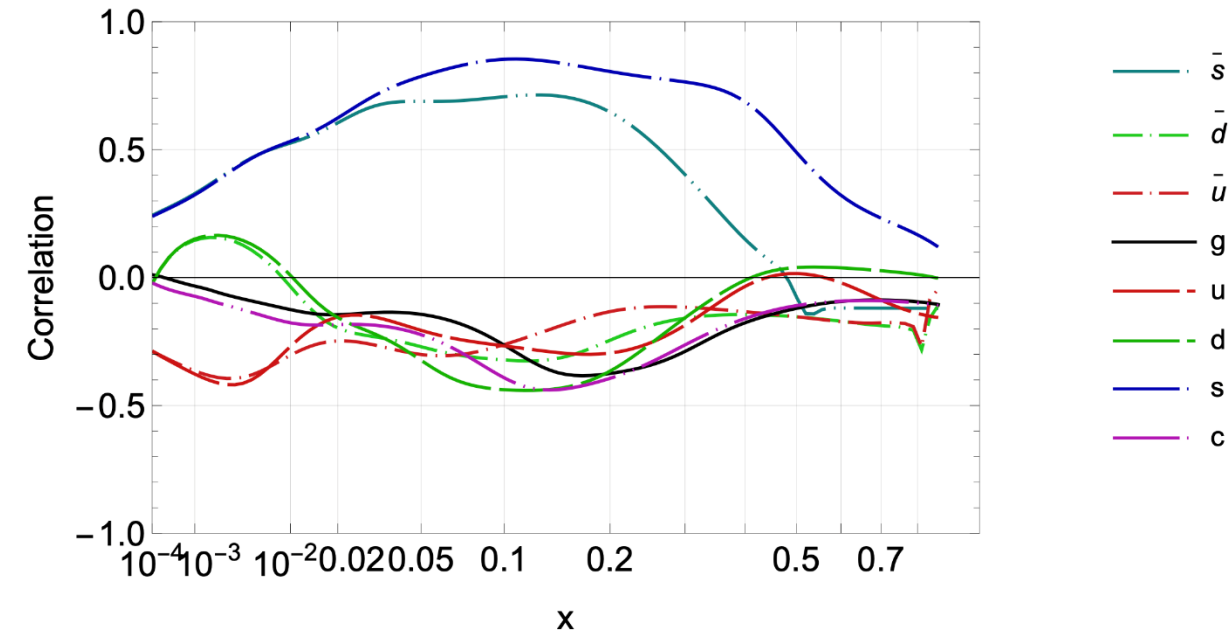
- CT18, MSHT20, and NNPDF4.0 predict very different \bar{q}/q at $x > 0.2$
 - The article quantified the potential effect of high-mass A_{FB} on large- x antiquarks
- See also NNPDF (2209.08115), Fiaschi et al. (2211.06188)

Recent CT studies of APV

- parity-violating DIS may access quark-level EW couplings
- 20 GeV JLab data may constrain nucleon PDFs
- NC SFs like F_3 (right) can inform high-x valence distributions; may indirectly constrain nucleon sea

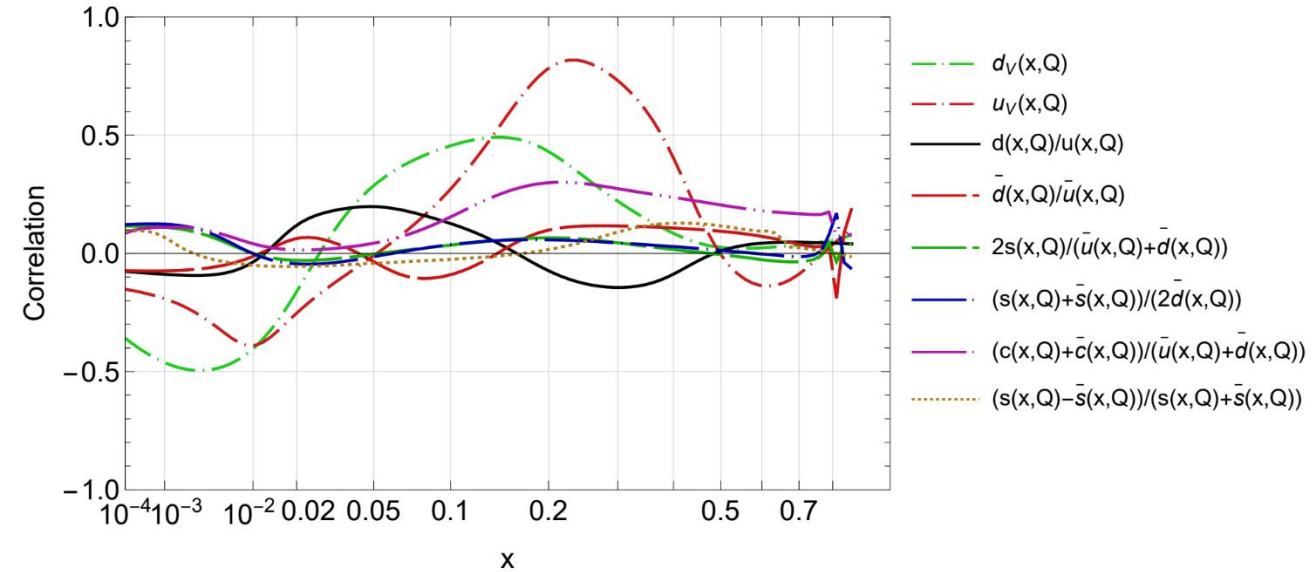
CT18As NNLO, 68% CL

$f(x, \mu^2=4 \text{ GeV}^2)$ and $(5F_2^{p,VZ}-2F_2^N)$ ($x_B=0.15, Q^2=4 \text{ GeV}^2$)



CT18As_Lat NNLO

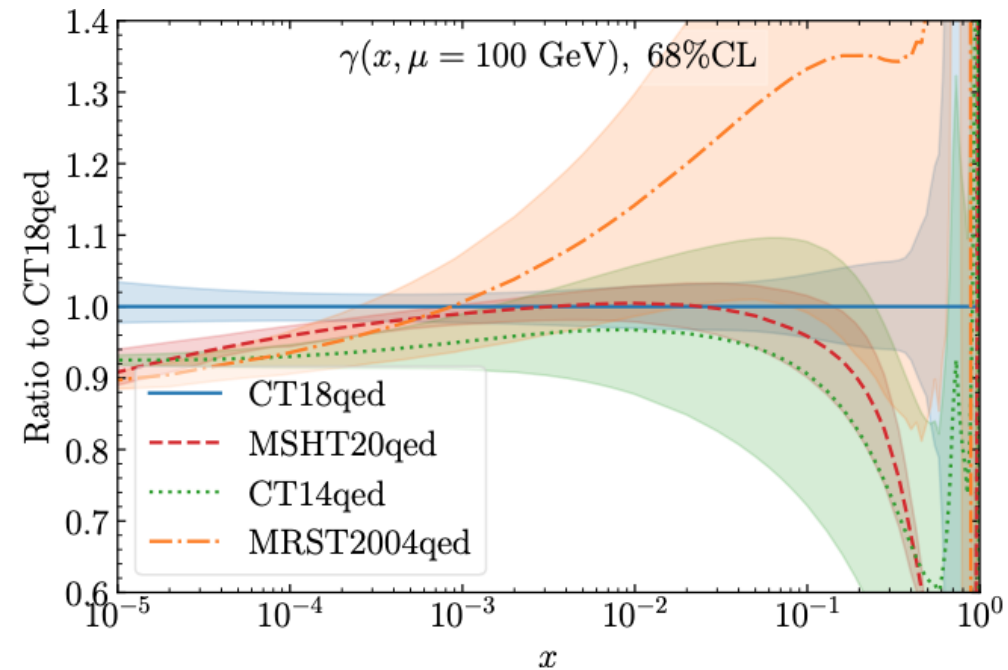
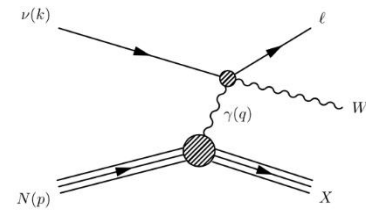
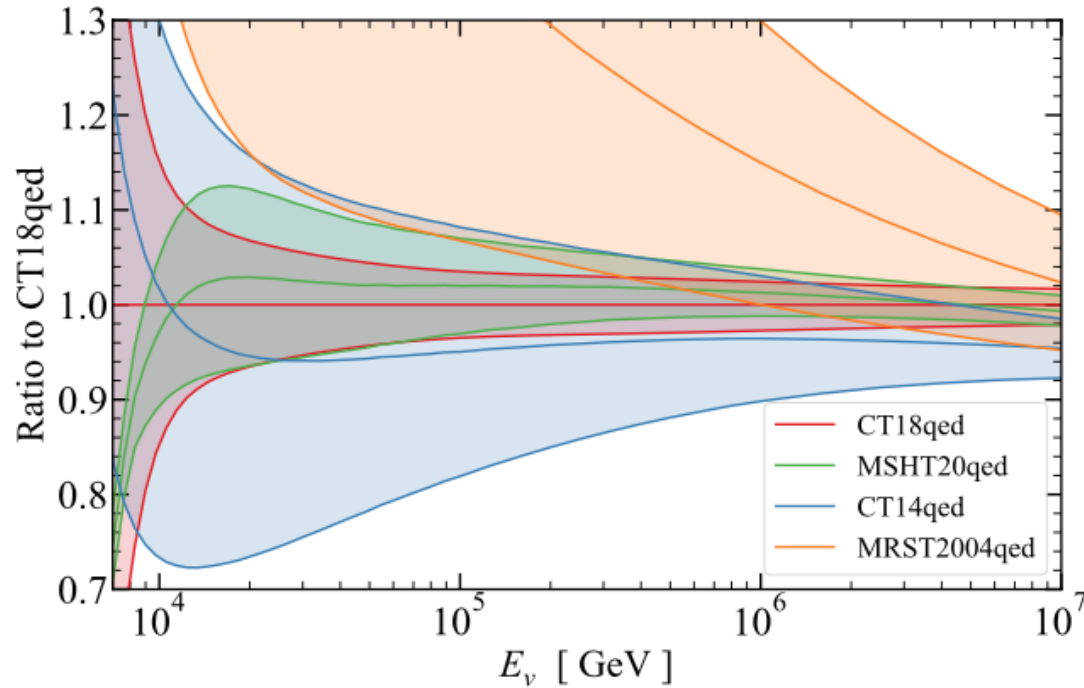
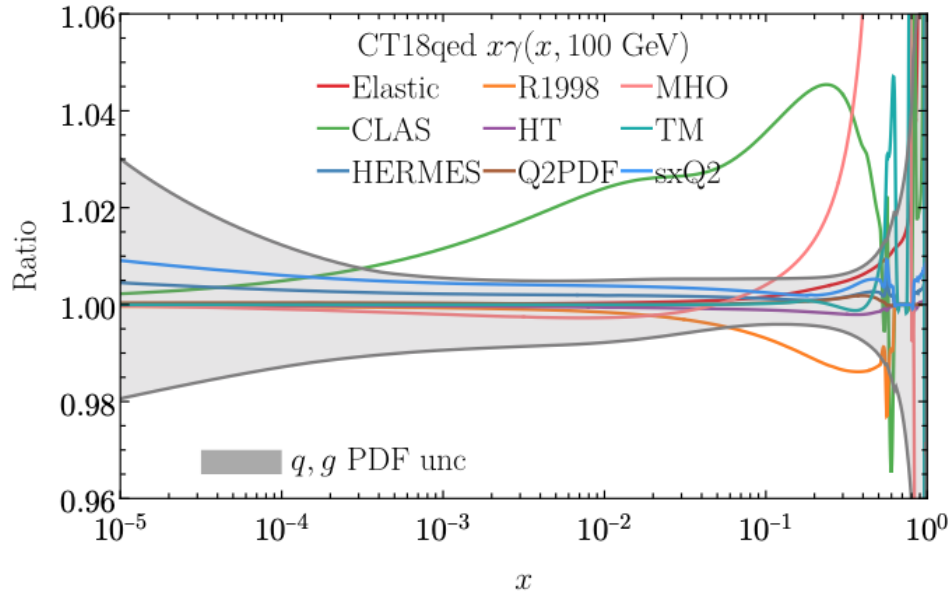
$f(x, Q^2=4 \text{ GeV}^2)$ and $F_3^{\text{ep} \rightarrow \text{eX}}$ ($x=0.25, Q^2=4 \text{ GeV}^2$)



- PVDIS may also allow constraints to strange PDF
- APV depends on interference structure functions
- combination with (isoscalar) deuteron structure function strongly correlated with $s(x)$
- needs high luminosity; control over deuteron (off-shell) corrections; thorough study of systematics

more detail in 2306.09360

Neutron's photon PDF

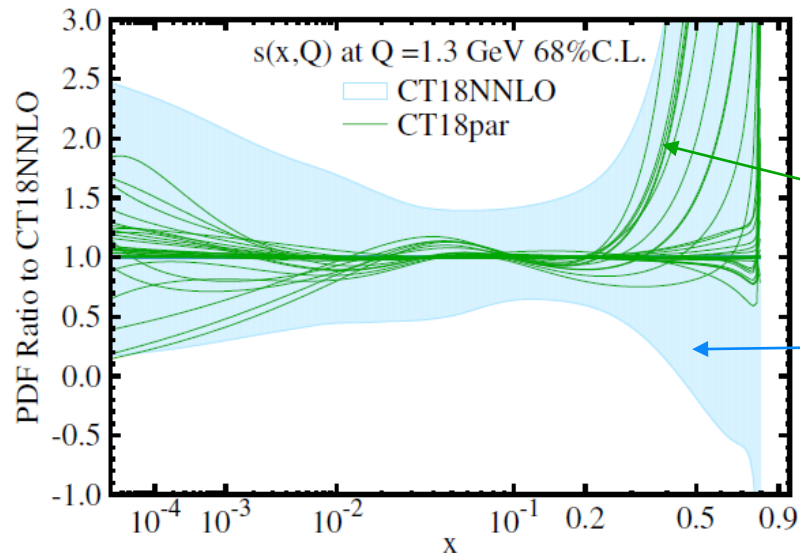


- We have determined the neutron's photon PDF using a similar methodology as for the proton one.
- The structure function is determined using pQCD at high Q^2 and HERMES and CLAS/CB data a low Q^2
- We estimated many low- Q^2 uncertainties, including the isospin symmetry violation and the QED evolution effects. We also explored implications for W-boson production, etc.
- CT18qed and MSHT20qed are in a good agreement
- In comparison to the first generation of photon PDFs, the uncertainty is significantly reduced.

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

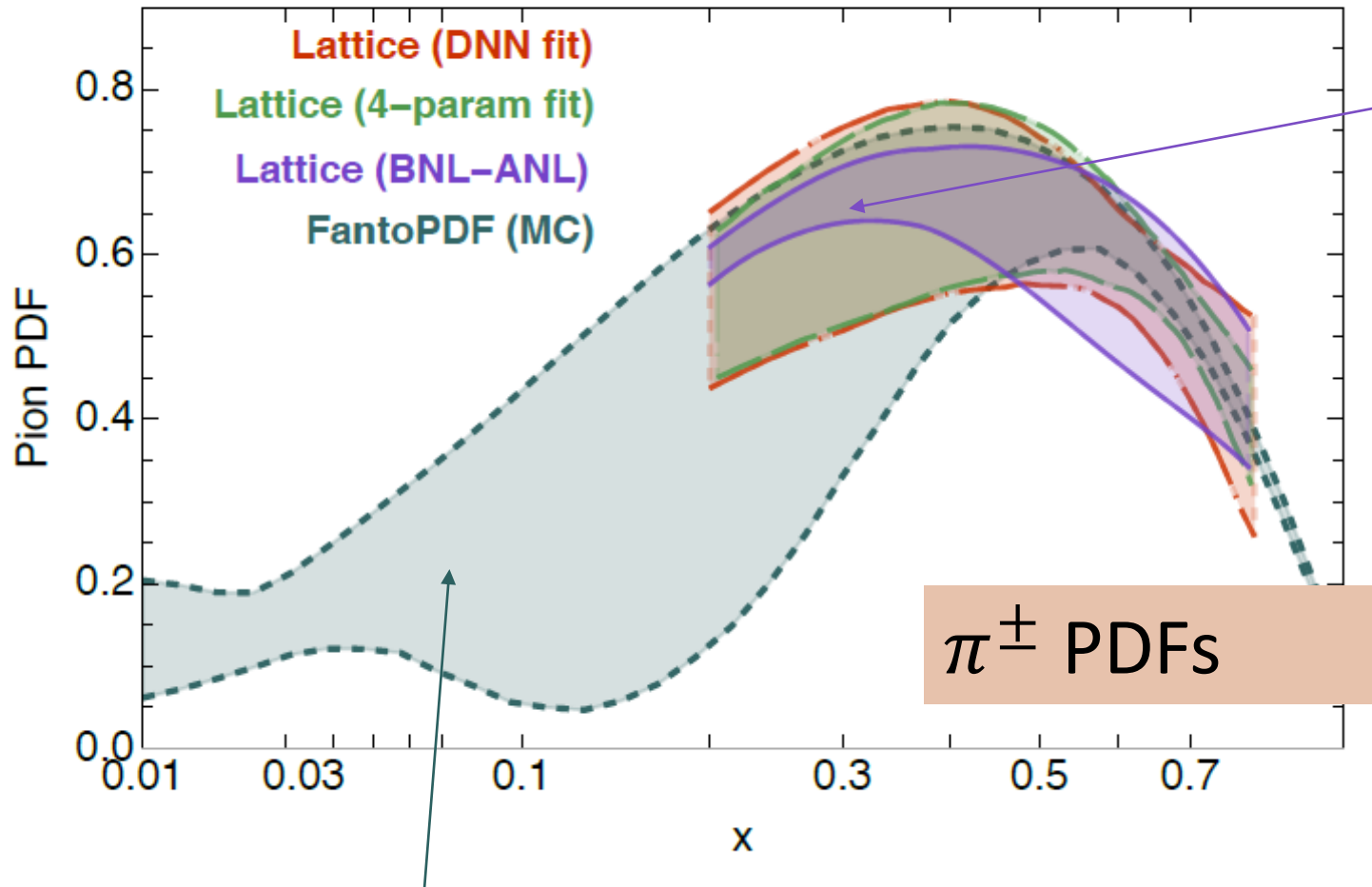
Preliminary fits explore experimental, theoretical, parametrization, methodological uncertainties

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

Fantômas + mp4lhc 2.0: pion PDFs with advanced parametrization uncertainties

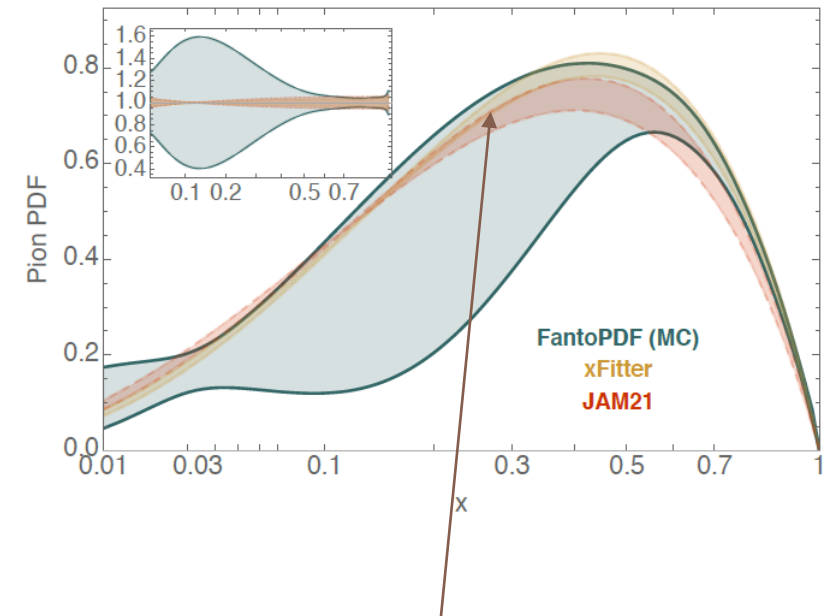
See A. Courtoy, DIS2024

$xV(x,Q)$ at $Q=2.$ GeV, 68% c.l. (band)



are the lattice uncertainties fully estimated?

$xV(x,Q)$ at $Q=1.4$ GeV, 68% c.l. (band)



without parametrization dependence

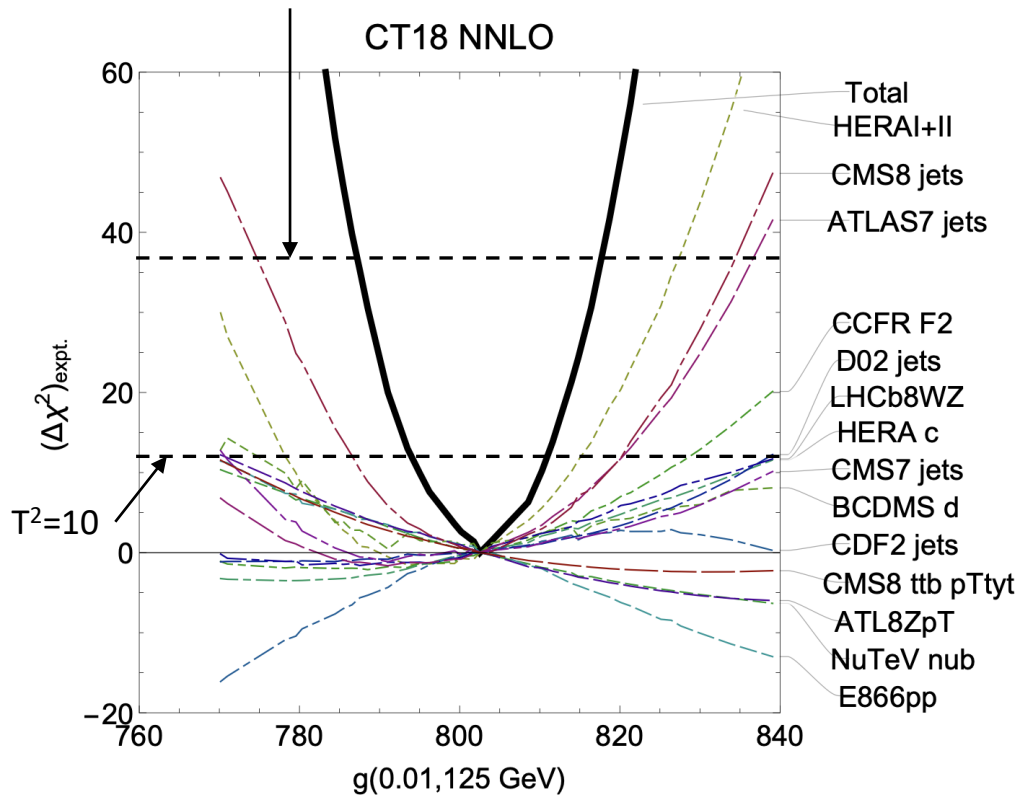
Phenomenological analysis, including the parametrization dependence
L. Kotz, A. Courtoy, M. Chavez, P. N., F. Olness, arXiv:2311.08447, accepted to PRD

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:

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

$\equiv r_{\text{posterior}} \qquad \qquad \equiv r_{\text{likelihood}} \qquad \qquad \equiv r_{\text{prior}}$

\therefore 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).

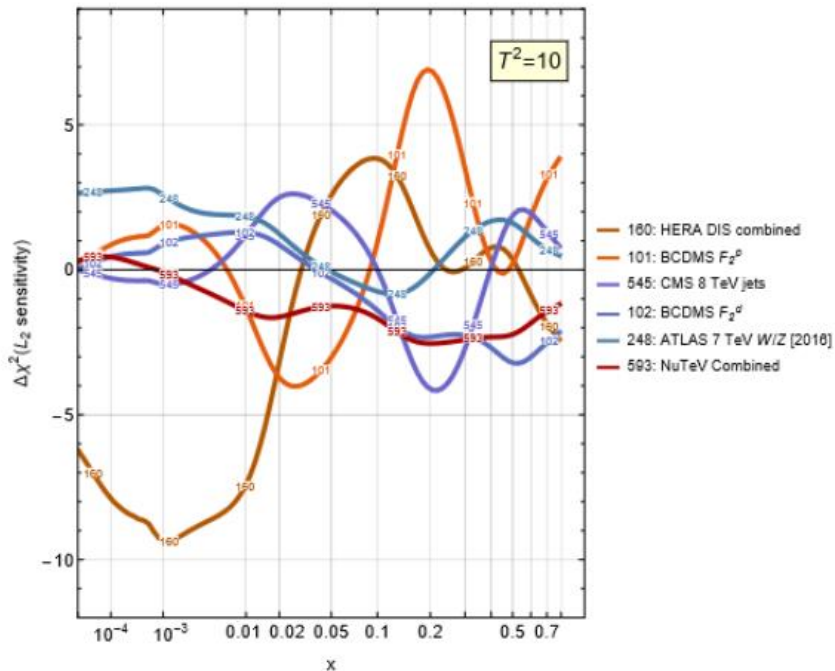
It also explains why $\Delta\chi^2 = 1$ does not capture the full uncertainty.

[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].

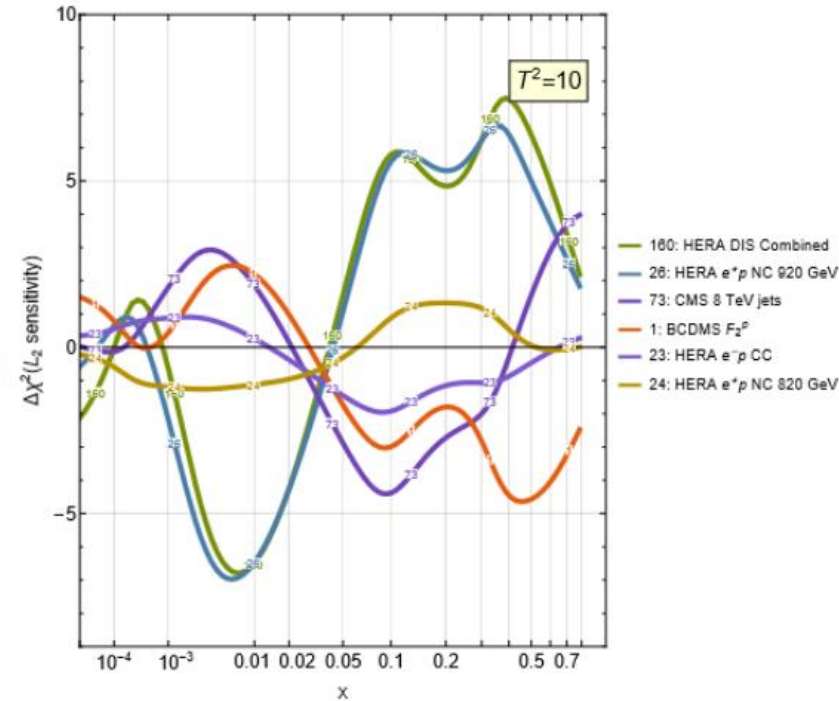
CT and MSHT use different criteria to account for the full uncertainty.

comparative study of NNLO and aN3LO PDF sensitivities

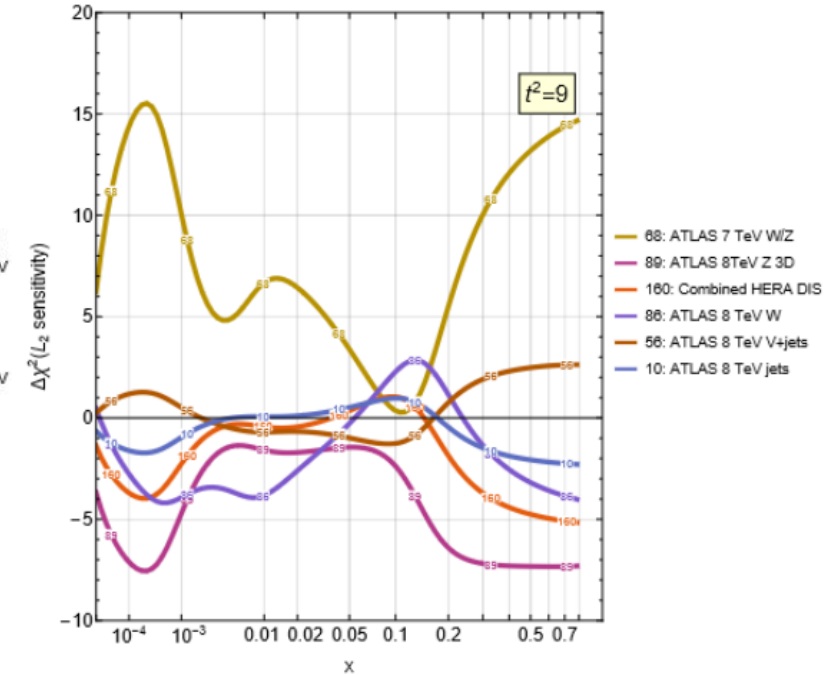
CT18' NNLO reduced
g(x, 2 GeV)



MSHT20 NNLO reduced
g(x, 2 GeV)



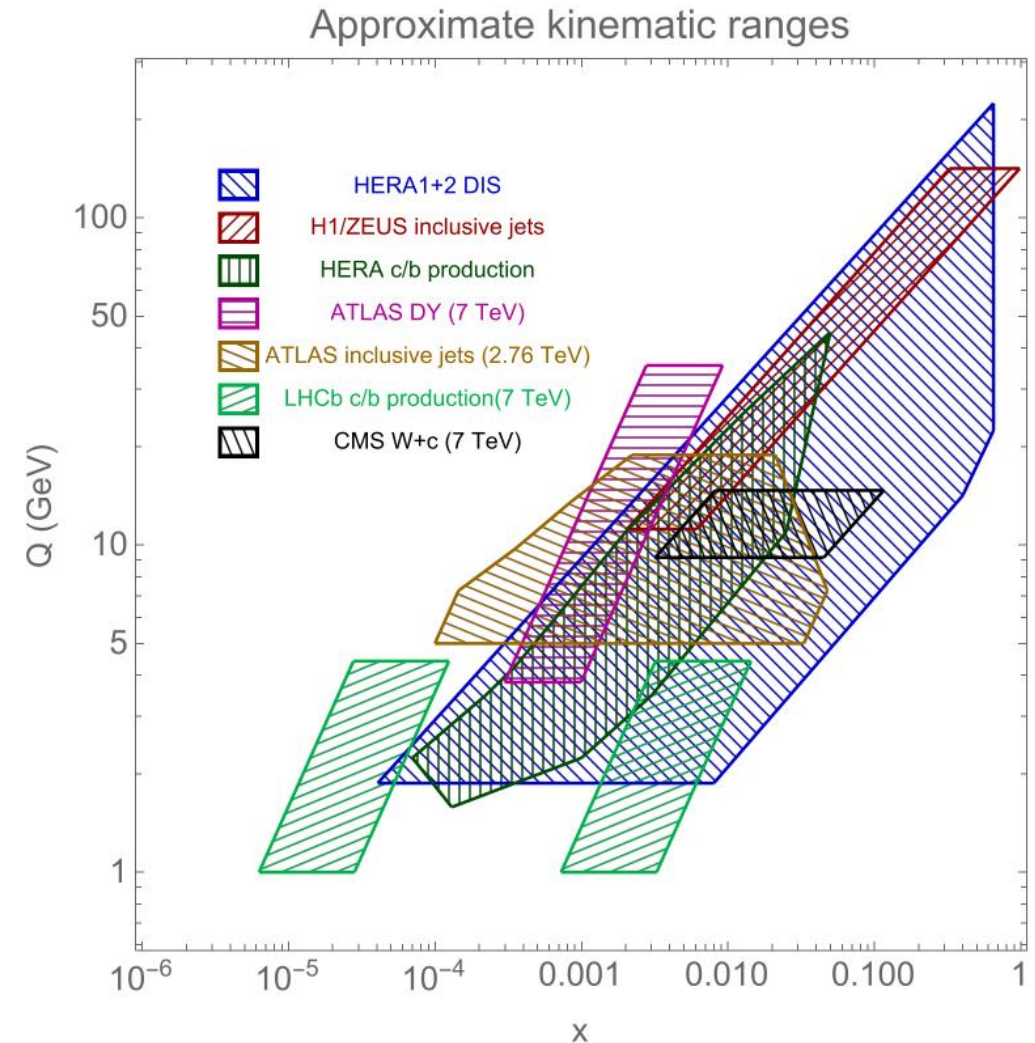
ATLAS21 NNLO
g(x, 2 GeV)



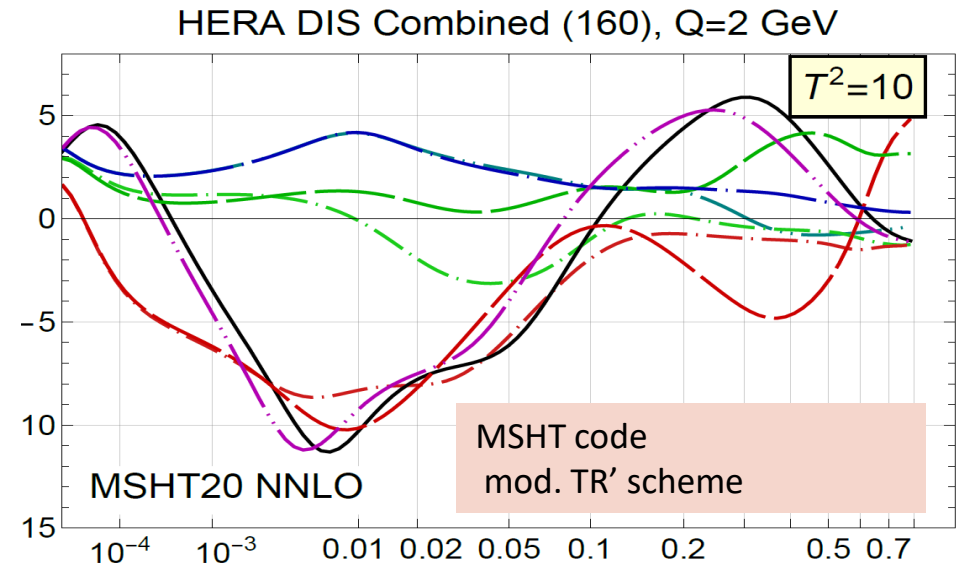
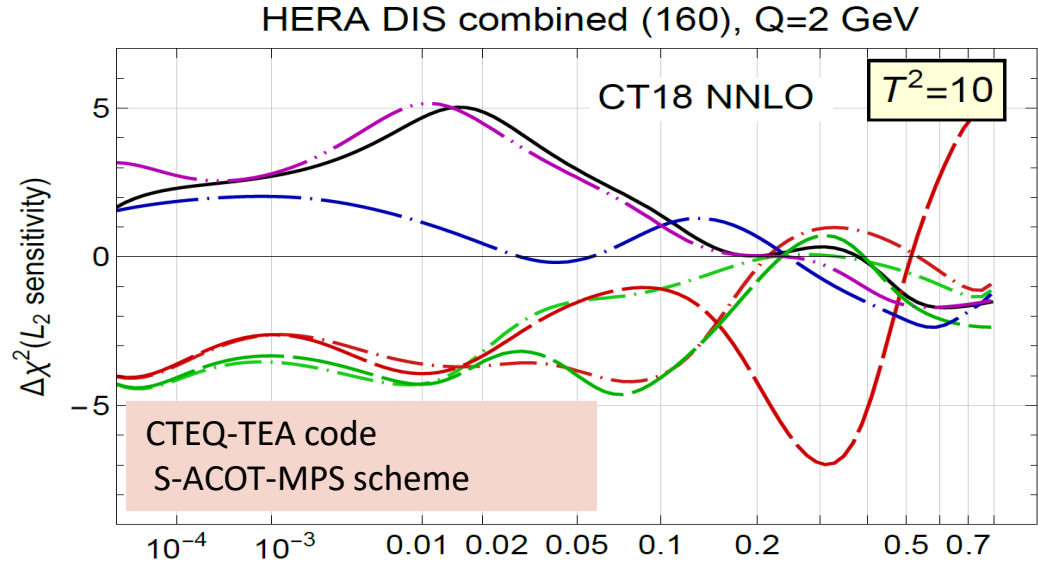
- Comparisons of strengths of constraints from individual data sets in 8 PDF analyses using the common L_2 sensitivity metric. [Definitions in the backup.]
- An interactive website (<https://metapdf.hepforge.org/L2/>) to plot such comparisons [2070 figures in total; a code L2LHAexplorer to plot L2 sensitivities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

L2 sensitivities were computed using xFitter

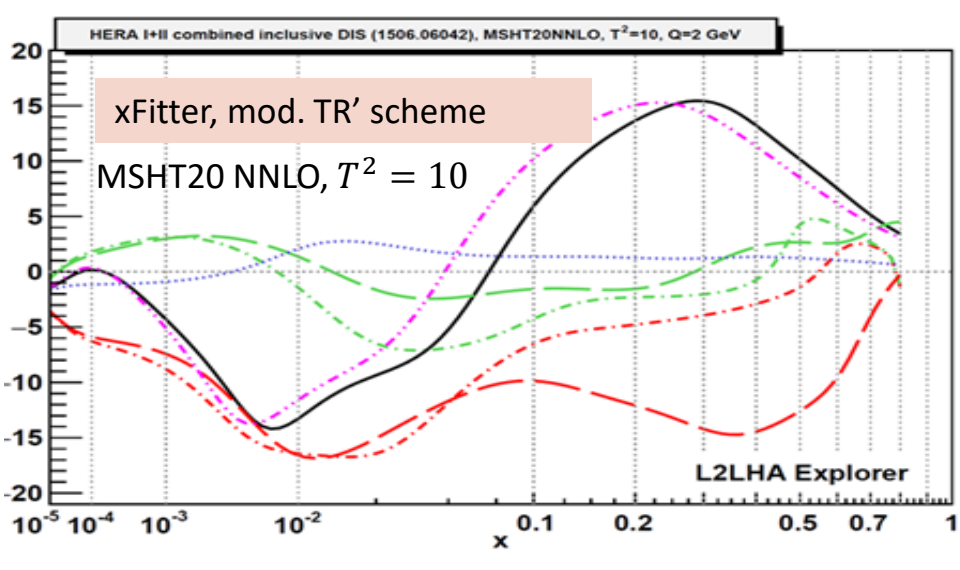
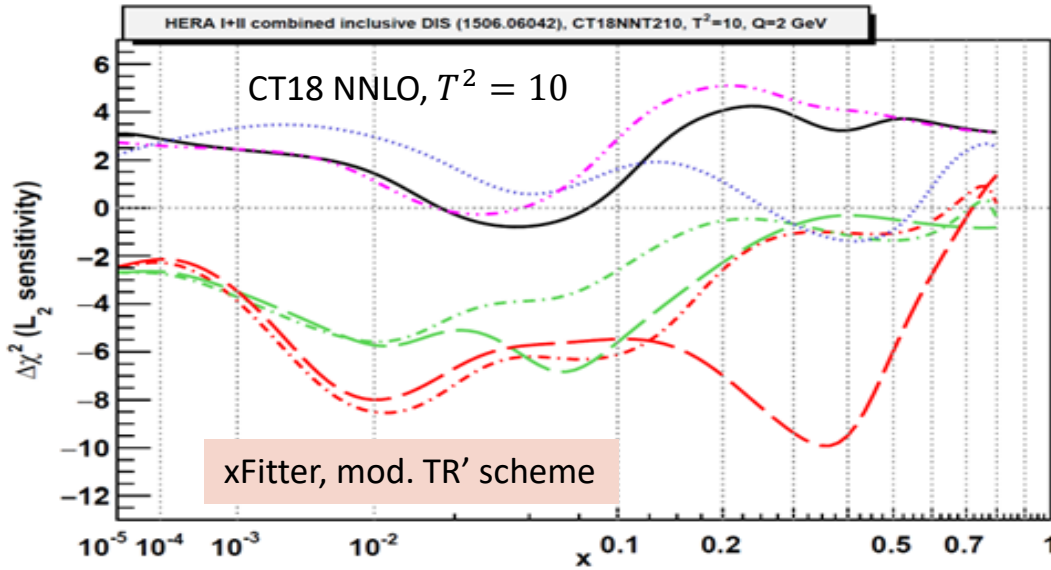
- PDF sets (NNLO, $\alpha_s(M_Z) = 0.118$, $Q = 2 \text{ GeV}$, $T^2 = 10$):
 - CT18
 - CT18As
 - MSHT20
- Data sets (included in xFitter):
 - ATLAS Drell-Yan ($\sqrt{s} = 7 \text{ TeV}$)
 - ATLAS jet production ($\sqrt{s} = 2.76 \text{ TeV}$)
 - CMS W+c production ($\sqrt{s} = 7 \text{ TeV}$)
 - H1+ZEUS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



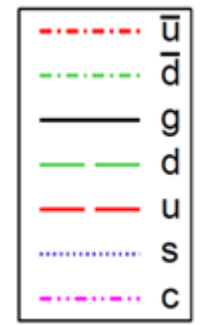
HERA I+II combined inclusive DIS [in CT18 and MSHT20]



Upper row:
From Jing et al., 2306.03918



Lower row:
From L. Kotz, 2401.11350

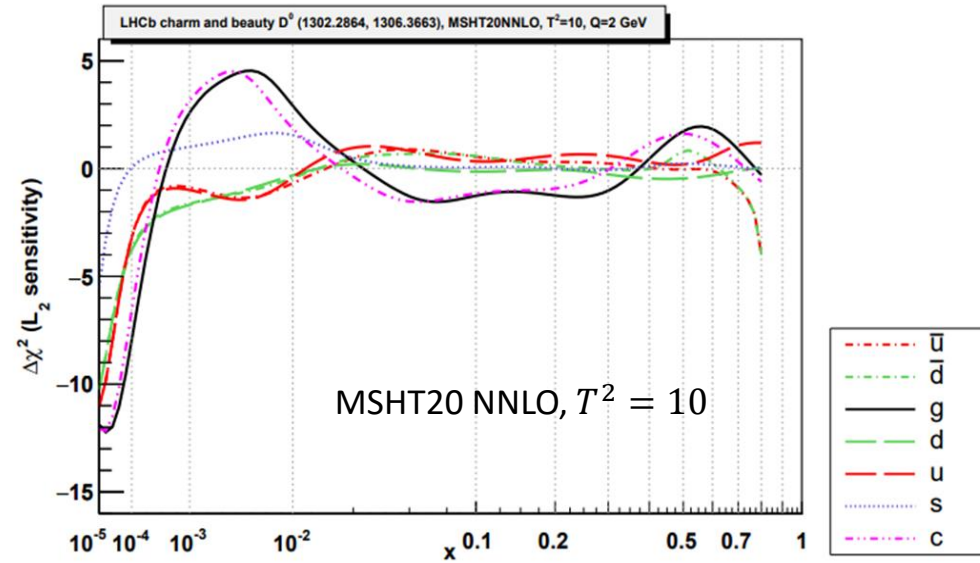
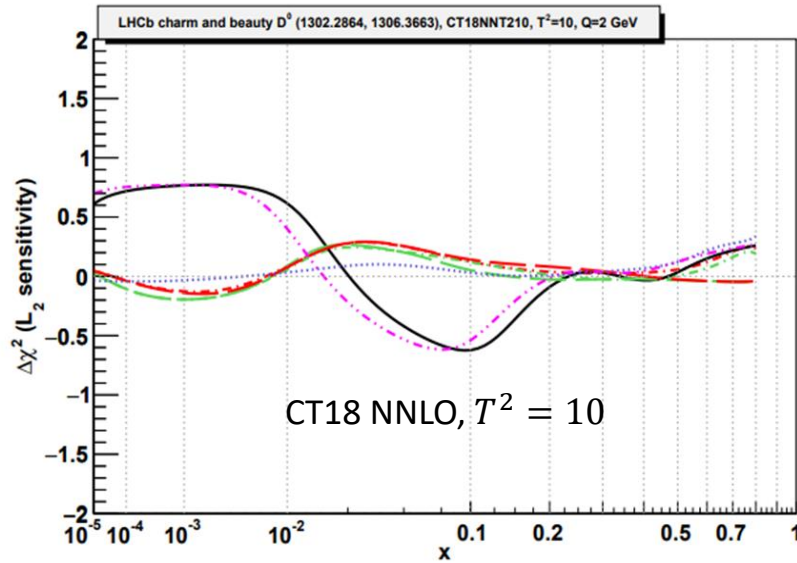


Left column: differences in χ^2 definition and heavy-quark scheme. Same PDFs and m_Q .

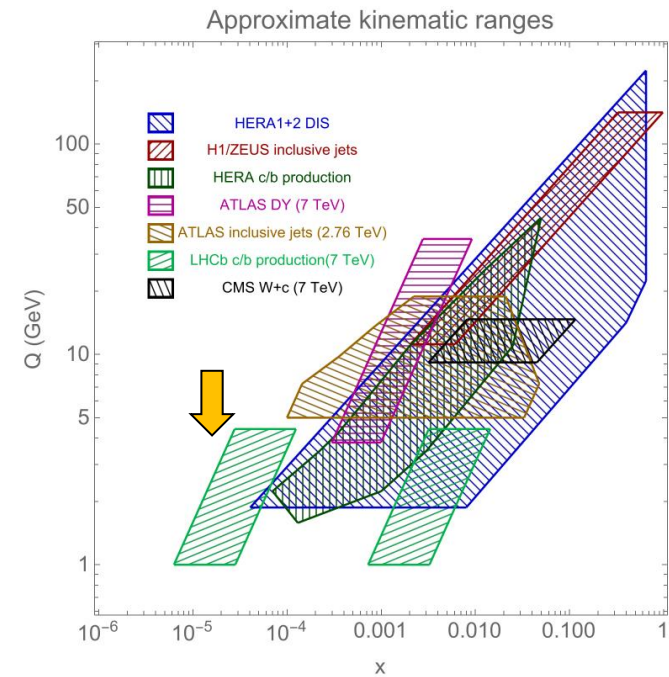
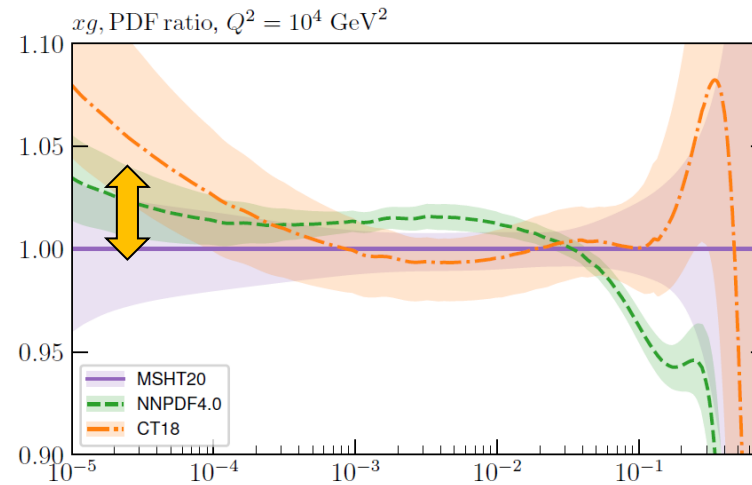
Right column: differences in χ^2 definition only. Same PDFs and m_Q .

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

From L. Kotz, 2401.11350

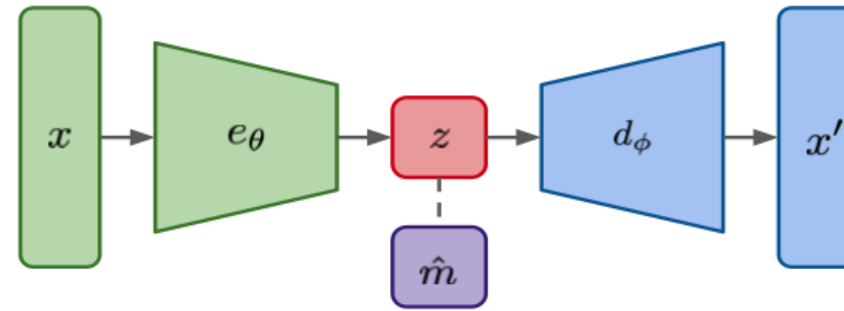
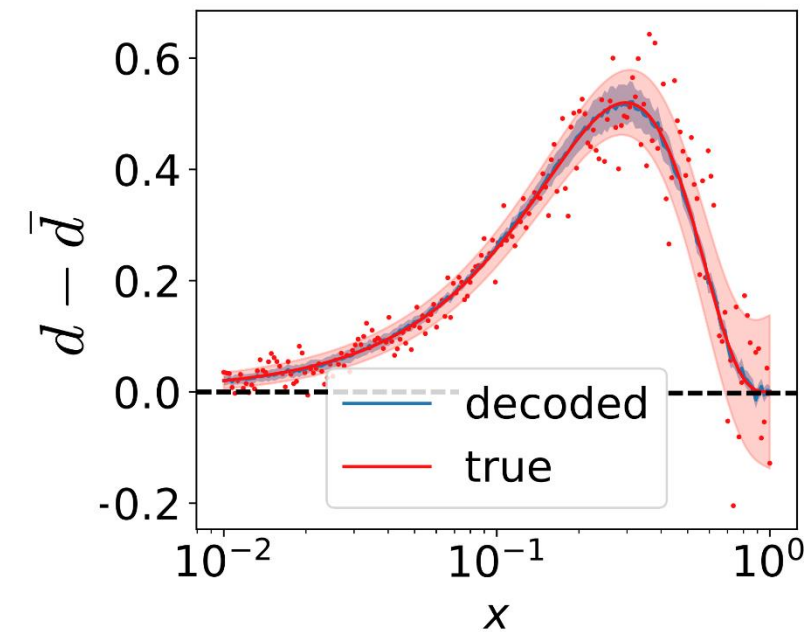


- LHCb c and b production prefers about the same gluon for CT18/CT18As.
- A larger gluon is preferred for MSHT20 at $x < 10^{-4}$.

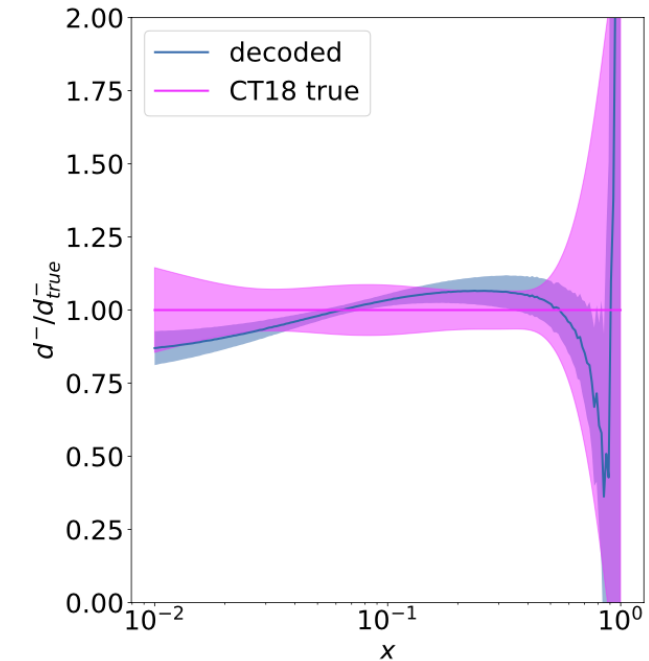


ML models for PDF generation

Kriesten and Hobbs, arXiv:2312.02278



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



- autoencoder-based models can efficiently represent PDFs in dimensionally reduced form
 - physics constraints may include PDFs' Mellin-space behavior (i.e., integrated moments)
 - through careful choice of network topology, can impose interpretable structure on latent space
 - trained models (like VAIM at right) can generatively predict PDFs from moments
 - how do AEs perform in the generic task of flexibly parametrizing PDFs?
 - where do the limits in the interpretability of this ML task lie?
- new ML tool to mutually compare PDFs, explore statistical properties (e.g., out-of-distribution behavior)

Near-future plans

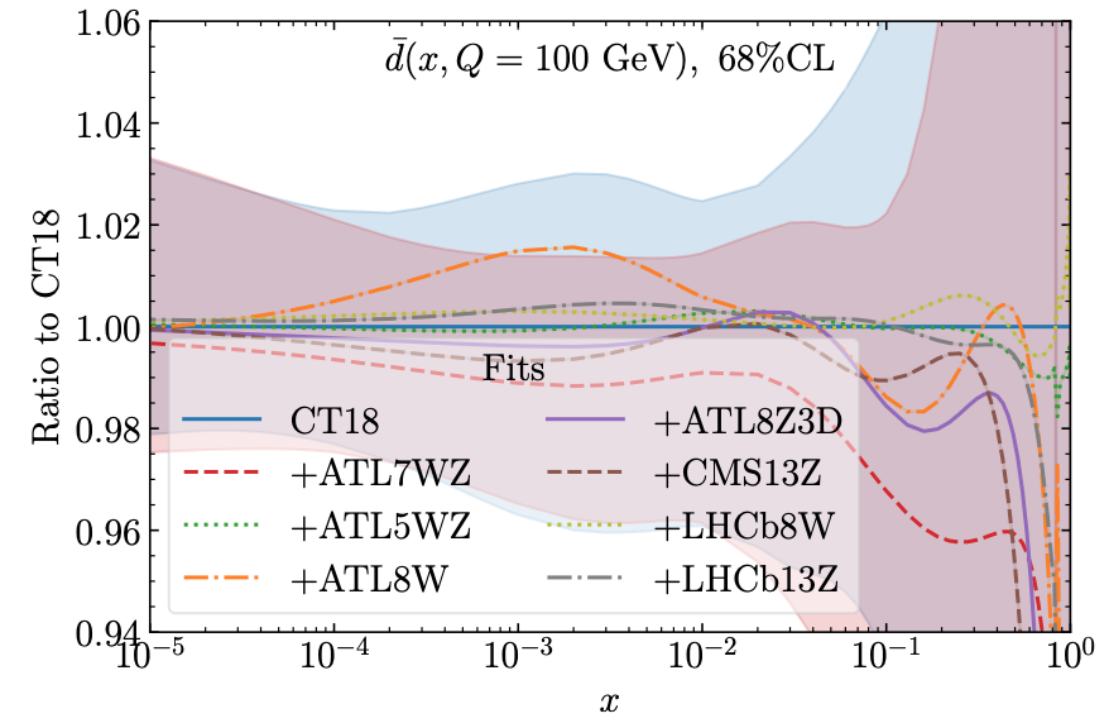
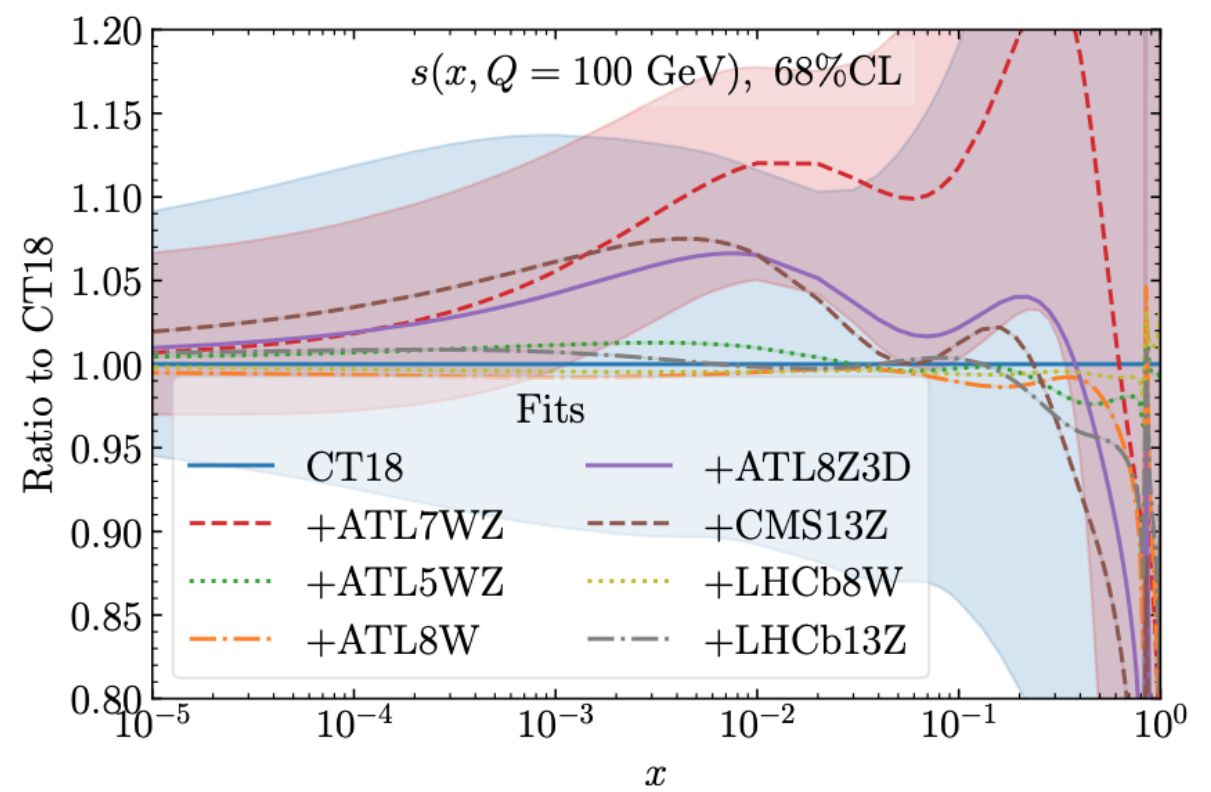
1. Final selection of experiments for NNLO PDFs planned for the next year
2. Work on N3LO contributions
3. Work on GMVFN schemes to have consistent HQ treatment
4. Next-generation PDF uncertainty quantification
5. Recent and imminent PDF releases
 - a. QCD+QED PDFs for protons and neutrons
 - b. Subtracted S-ACOT-MPS PDFs
 - c. Fantômas 1.0 pion PDFs (Hessian)
 - d. Release of the Fantômas PDF parametrization package in xFitter

Backup

Post-CT18 Drell-Yan data's impact

2305.10733 (PRD23')

ID	Experiment	N_{pt}	Ratio to CT18		
			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) data came out after the release of CT18 PDFs.
- We found that most of the nDY data sets are consistent with the ATLAS 7 WZ precision data (16') and prefer enhanced strangeness at $x \sim 0.02$
- Only one exception: ATL8W has an opposite pull on d, \bar{d}
- CMS13Z and ATL8W have a similar χ^2/N_{pt} as ATL7WZ
- The more flexible strangeness parameterization in CT18As can relax the tension, but not completely resolve it.

What is the L_2 sensitivity?

- The L_2 sensitivity is a way of visualizing the pulls of fitted experiments on the best-fit PDF $f_a(x, Q)$, for a particular parton flavor x , as a function of x and Q
 - or, when plotted for a PDF luminosity, as a function of the final-state mass M_X
- The best-fit value for a particular $f_a(x, Q)$ is determined by the sum of these pulls
- Both the L_2 and LM methods explore the parametric dependence of the χ^2 function in the vicinity of the global minimum
- The L_2 sensitivity streamlines comparisons among independent PDF analyses using **published error PDFs**
- The L_2 sensitivity has been used internally by CT (in CT18), by the PDF4LHC21 benchmarking group (to determine which data sets should be in the reduced PDF fit used for benchmarking), and now by ATLASpdf, CT, and MSHT

Tolerance hypersphere in the PDF space

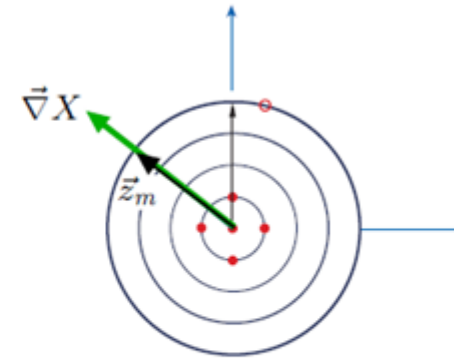
2-dim (i,j) rendition of N-dim (26) PDF parameter space

Hessian method: Pumplin et al., 2001

A symmetric PDF error for a physical observable X is given by

$$\Delta X = \vec{\nabla} X \cdot \vec{z}_m = |\vec{\nabla} X|$$

$$= \frac{1}{2} \sqrt{\sum_{i=1}^N \left(X_i^{(+)} - X_i^{(-)} \right)^2}$$



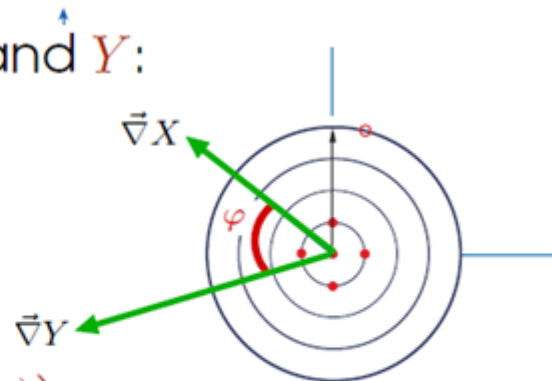
(b)
Orthonormal eigenvector basis

Correlation cosine for observables X and Y :

hep-ph/0110378

$$\cos \varphi = \frac{\vec{\nabla} X \cdot \vec{\nabla} Y}{\Delta X \Delta Y} =$$

$$\frac{1}{4\Delta X \Delta Y} \sum_{i=1}^N \left(X_i^{(+)} - X_i^{(-)} \right) \left(Y_i^{(+)} - Y_i^{(-)} \right)$$

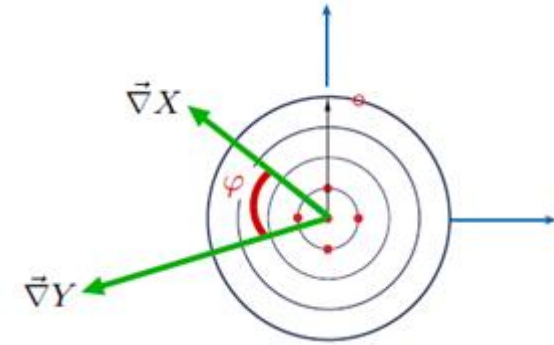


(b)
Orthonormal eigenvector basis

L_2 sensitivity, definition

$S_{f,L_2}(E)$ for experiment E is the estimated $\Delta\chi_E^2$ for this experiment when a PDF $f_a(x_i, Q_i)$ increases by the +68% c.l. Hessian PDF uncertainty

Take $X = f_a(x_i, Q_i)$ or $\sigma(f)$; $Y = \chi_E^2$ for experiment E .



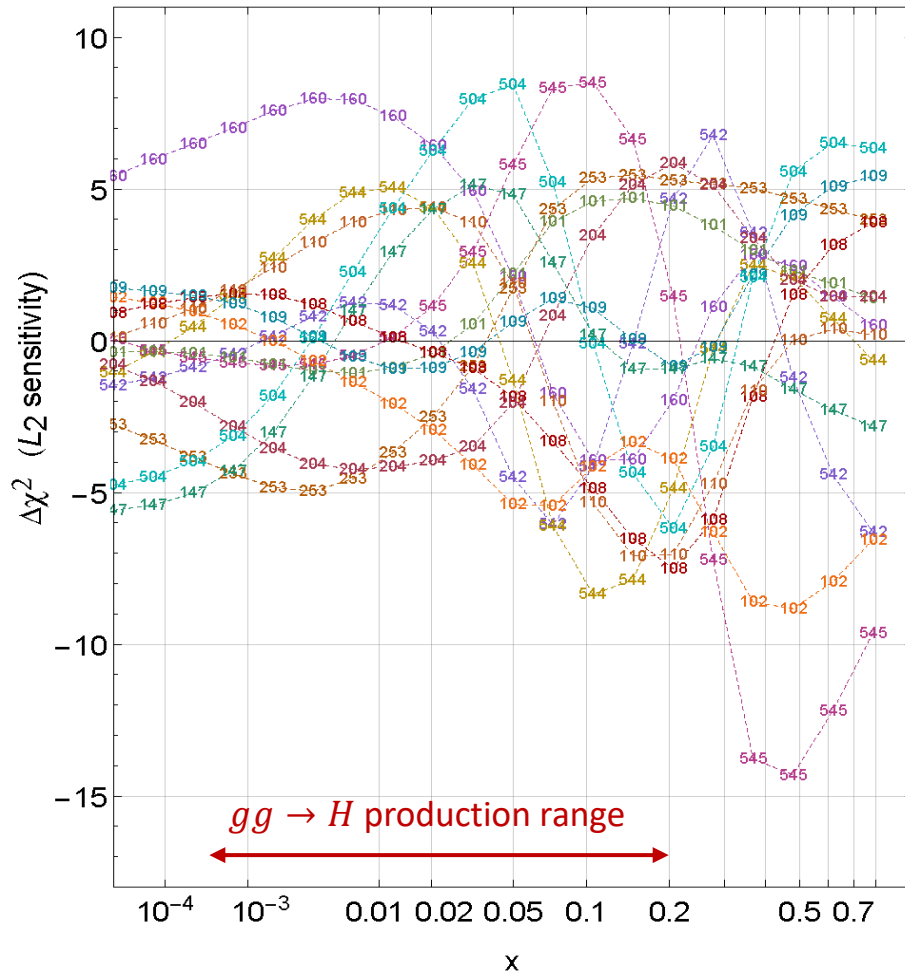
$$S_{f,L_2} \equiv \Delta Y(\vec{z}_{m,X}) = \vec{\nabla}Y \cdot \vec{z}_{m,X} = \vec{\nabla}Y \cdot \frac{\vec{\nabla}X}{|\vec{\nabla}X|} = \Delta Y \cos \varphi$$

A fast version of the Lagrange Multiplier scan of χ_E^2 along the direction of $f_a(x_i, Q_i)$!

Estimated χ^2 pulls from experiments

(L_2 sensitivity, T. J. Hobbs et al., arXiv:1904.00222)

CT18 NNLO, $g(x, 100 \text{ GeV})$



CT18 NNLO, gluon at $Q=100 \text{ GeV}$

15 core-minutes

Most sensitive experiments

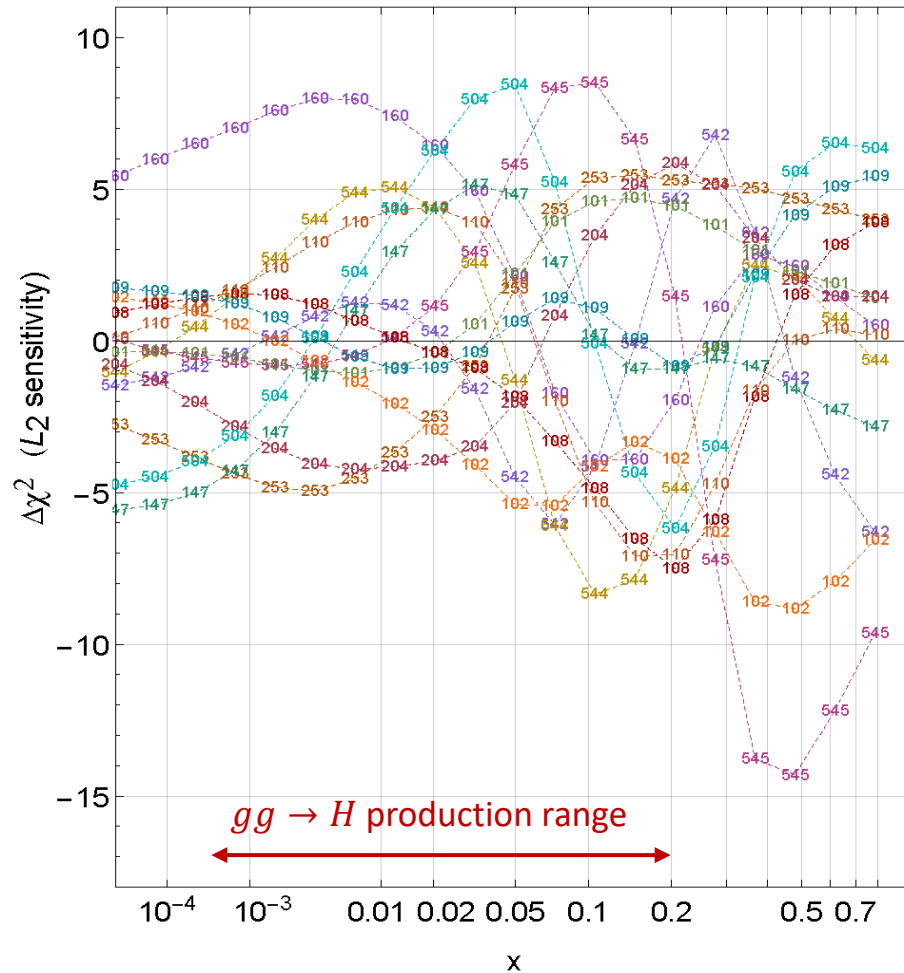
- 253--- ATL8ZpTbT
- 542--- CMS7jtR7y6T
- 544--- ATL7jtR6uT
- 545--- CMS8jtR7T
- 160--- HERAplI
- 101--- BcdF2pCor
- 102--- BcdF2dCor
- 108--- cdhswf2
- 109--- cdhswf3
- 110--- ccf2.mi
- 147--- Hn1X0c
- 204--- e866ppxf
- 504--- cdf2jtCor2

Experiments with large $\Delta\chi^2 > 0$ [$\Delta\chi^2 < 0$] pull $g(x, Q)$ in the negative [positive] direction at the shown x

Estimated χ^2 pulls from experiments

(L_2 sensitivity, T. J. Hobbs et al., arXiv:1904.00222)

CT18 NNLO, $g(x, 100 \text{ GeV})$



CT18 NNLO, gluon at Q=100 GeV

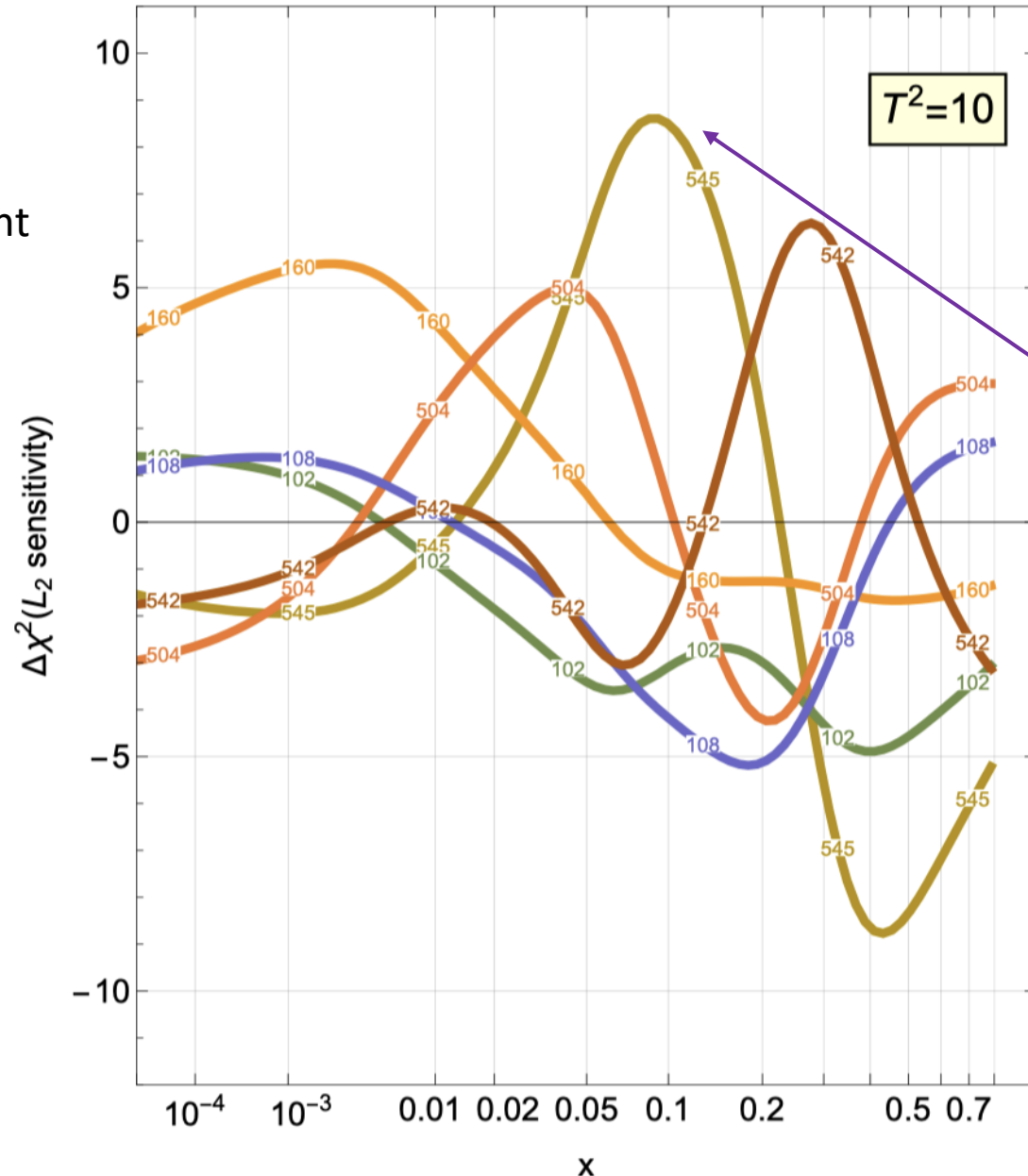
Most sensitive experiments

- 253--- ATLAS8ZpTbT
- 109--- cdhswf3
- 542--- CMS7jtR7y6T
- 110--- ccfif2.mi
- 544--- ATLAS7jtR6uT
- 147--- Hn1X0c
- 545--- CMS8jtR7T
- 204--- e866ppxf
- 160--- HERAIpII
- 504--- cdf2jtCor2
- 101--- BcdF2pCor
- 102--- BcdF2dCor
- 108--- cdhswf2

Note opposite pulls (tensions) in some x ranges between HERA I+II DIS (ID=160); CDF (504), ATLAS 7 (544), CMS 7 (542), CMS 8 jet (545) production; E866pp DY (204); ATLAS 8 Z pT (253) production; BCDMS and CDHSW DIS

CT18 NNLO
g(x, 100 GeV)

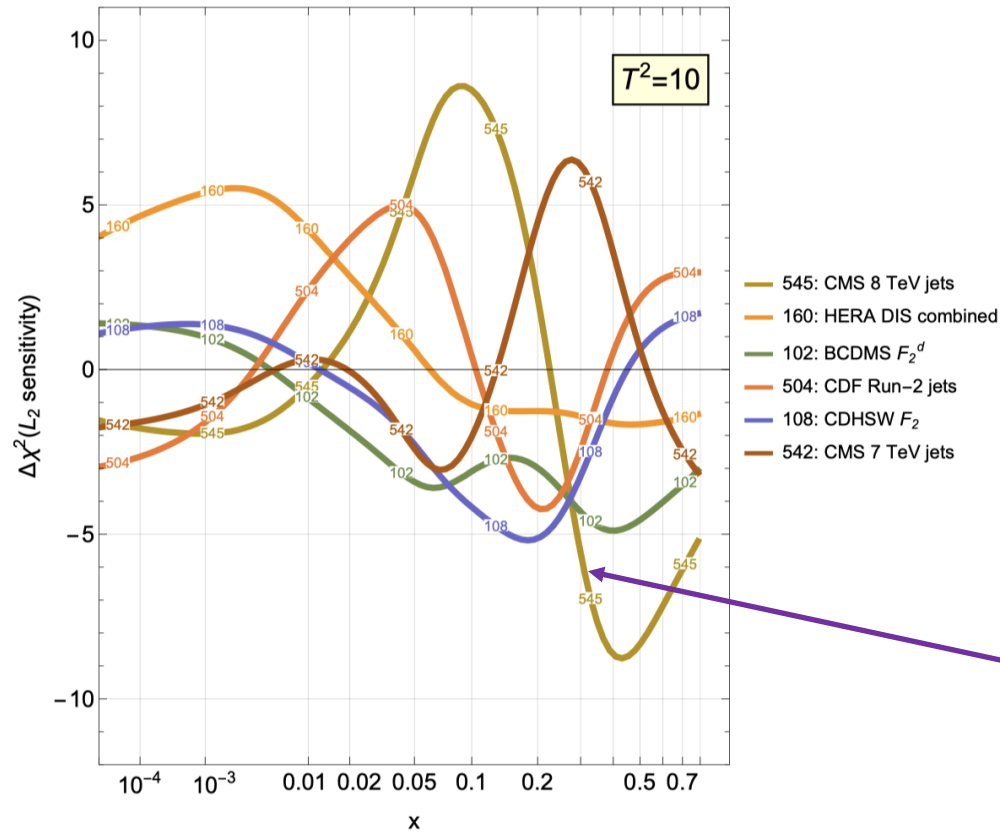
show only 6
most important
experiments



Small tolerance to stay in
the region where total χ^2
has best quadratic behavior

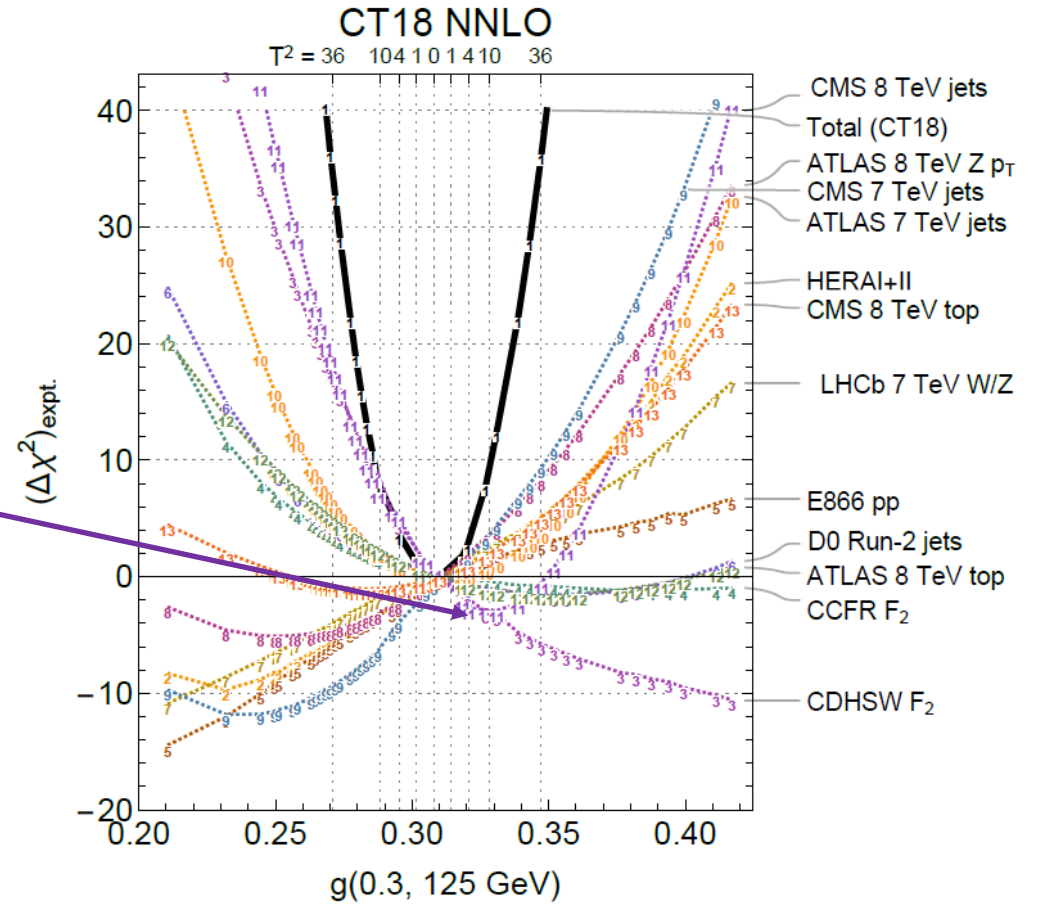
- 545: CMS 8 TeV jets
- 160: HERA DIS combined
- 102: BCDMS F_2^d
- 504: CDF Run-2 jets
- 108: CDHSW F_2
- 542: CMS 7 TeV jets

CT18 NNLO
 $g(x, 100 \text{ GeV})$

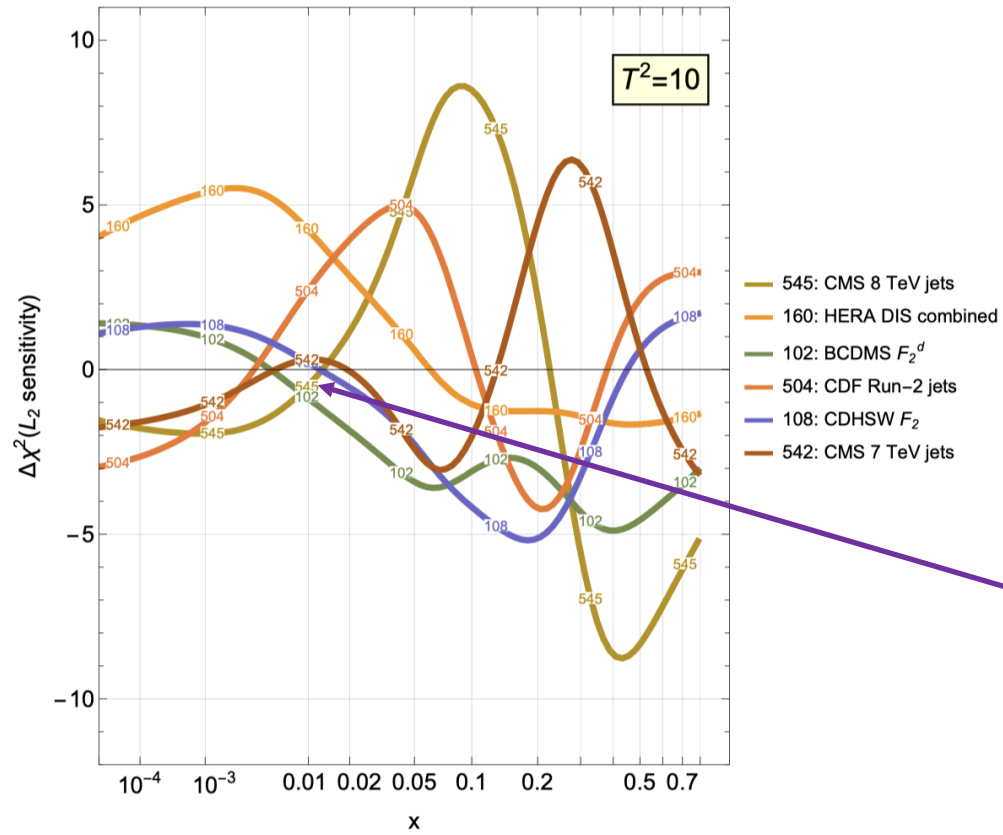


Compare to LM scans

(focus on CMS 8 TeV jets,
 IDs=545 and 11)

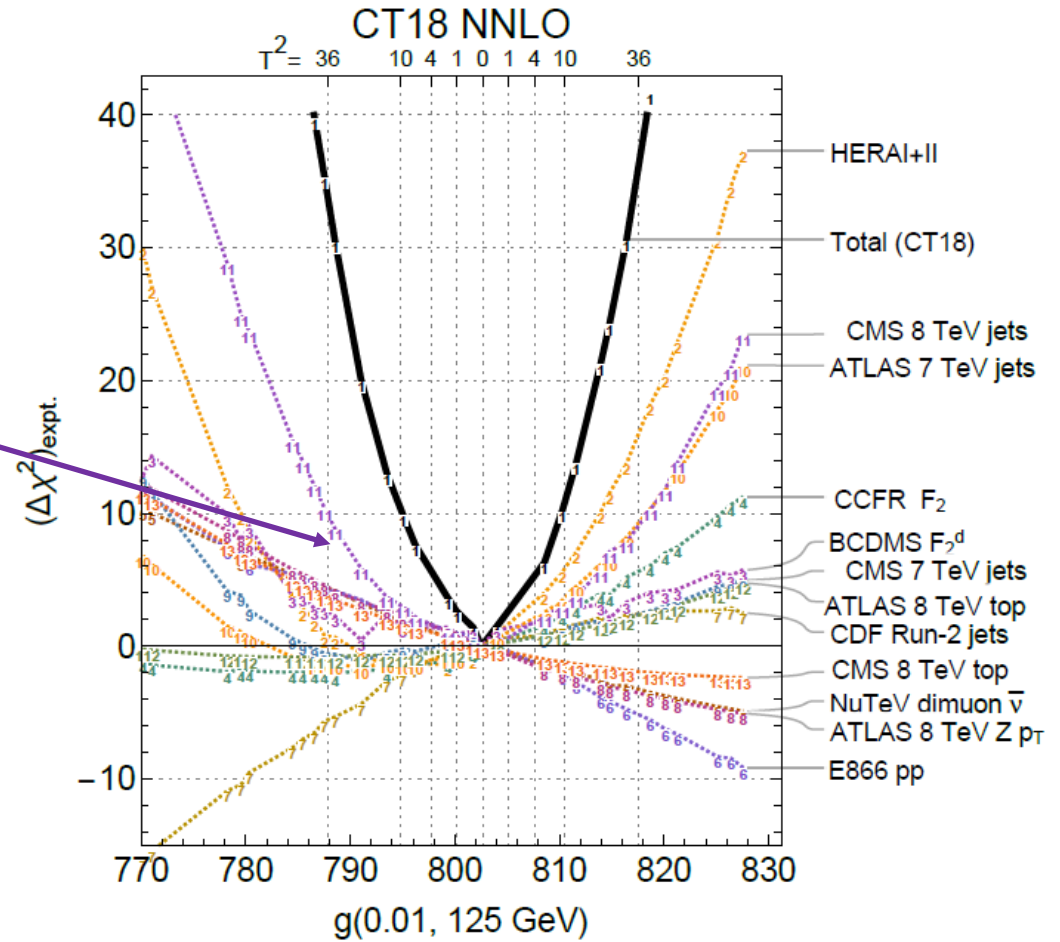


CT18 NNLO
 $g(x, 100 \text{ GeV})$



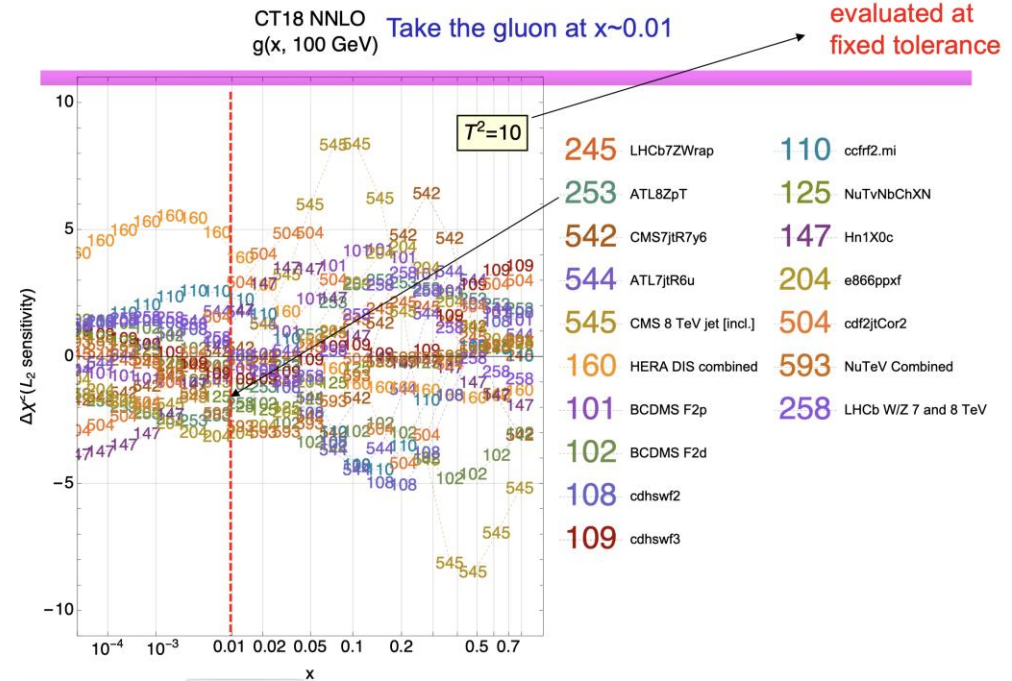
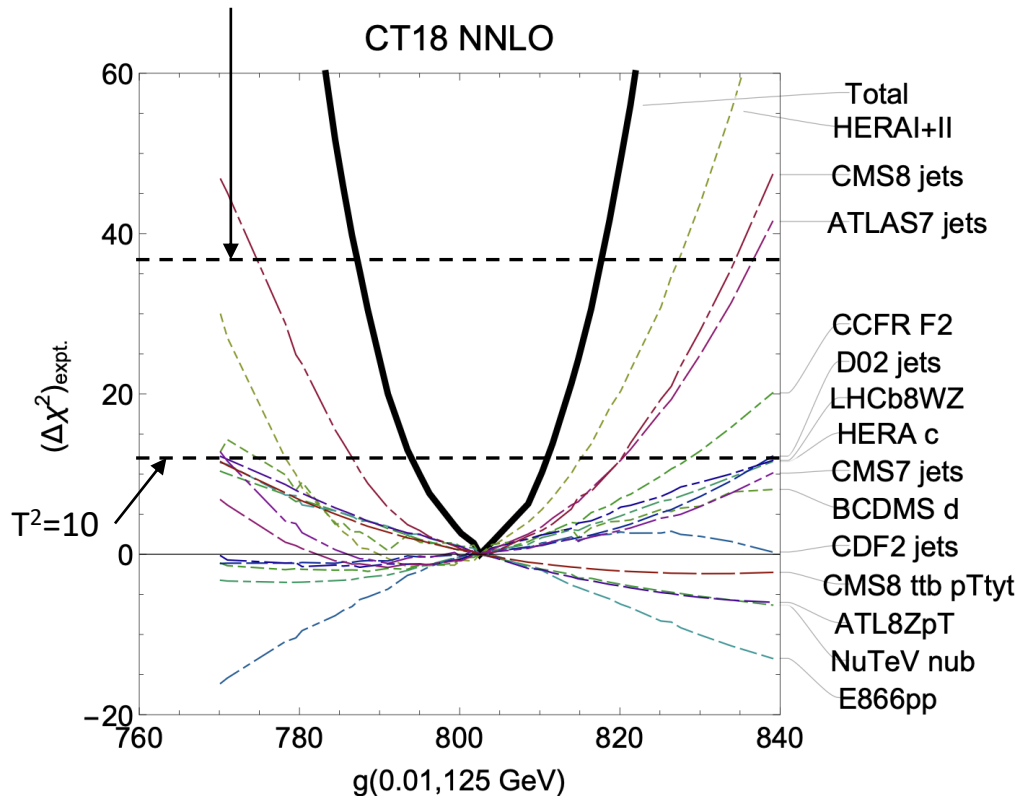
Compare to LM scans

(focus on CMS 8 TeV jets,
 IDs=545 and 11)



CT and MSHT both use a Hessian technique to determine the central PDF. By definition, this is at the best χ^2 . This is not necessarily true for NNPDF.

The uncertainty is determined by allowing an excursion from that central value. CT18 uses $\Delta\chi^2=37$ for a 68% CL error.



In a global PDF fit, there are tensions between the input data sets, by definition. These tensions are most easily demonstrated by the use of the L2 sensitivity above. Some data sets pull the gluon up at $x \sim 0.01$, some down.

The end result of the pulling is the central PDF. The PDF uncertainty reflects the size of those pulls/tensions.

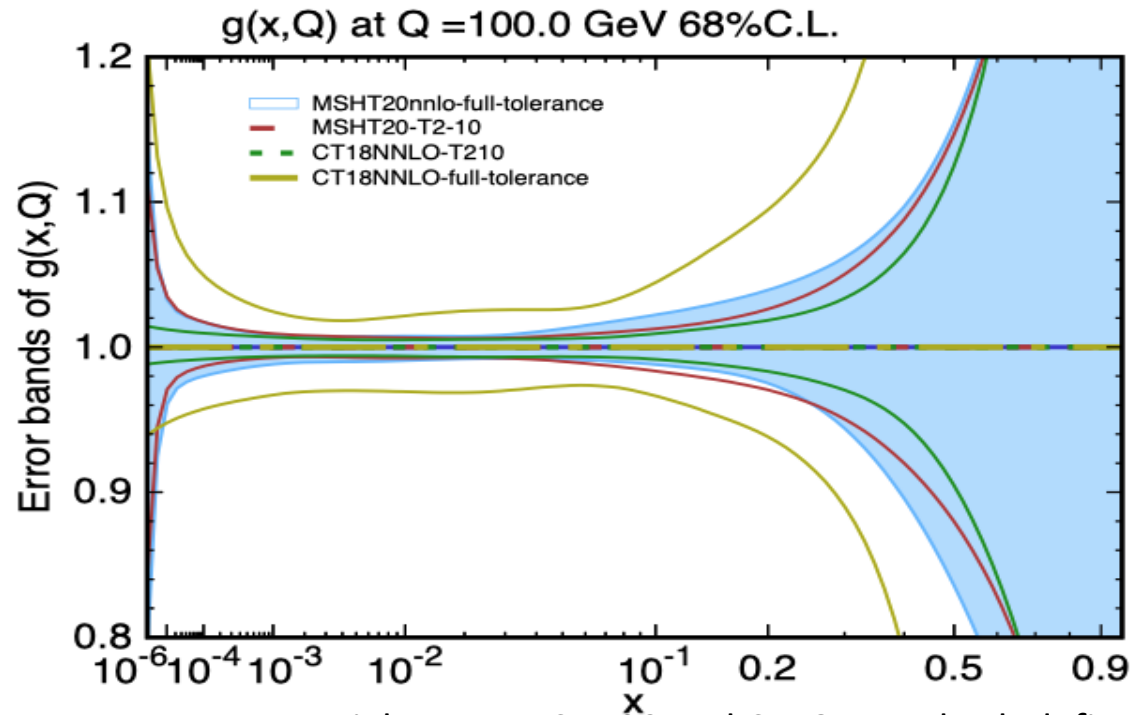
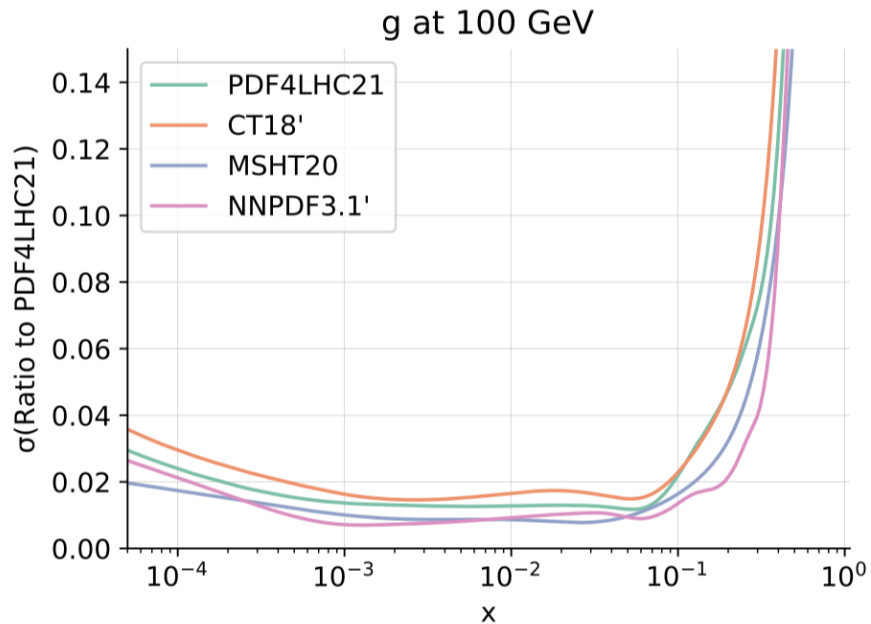
Typical χ^2/dof are of the order of 1.1 for >4000 points, or very unlikely from the pure statistical POV. $\Delta\chi^2=1$ does not capture the full uncertainty.

CT and MSHT use different criteria to define those tensions/define the uncertainty.

The PDF uncertainties for the combination in the PDF4LHC21 exercise is shown below. Same data sets used for all PDF fits.

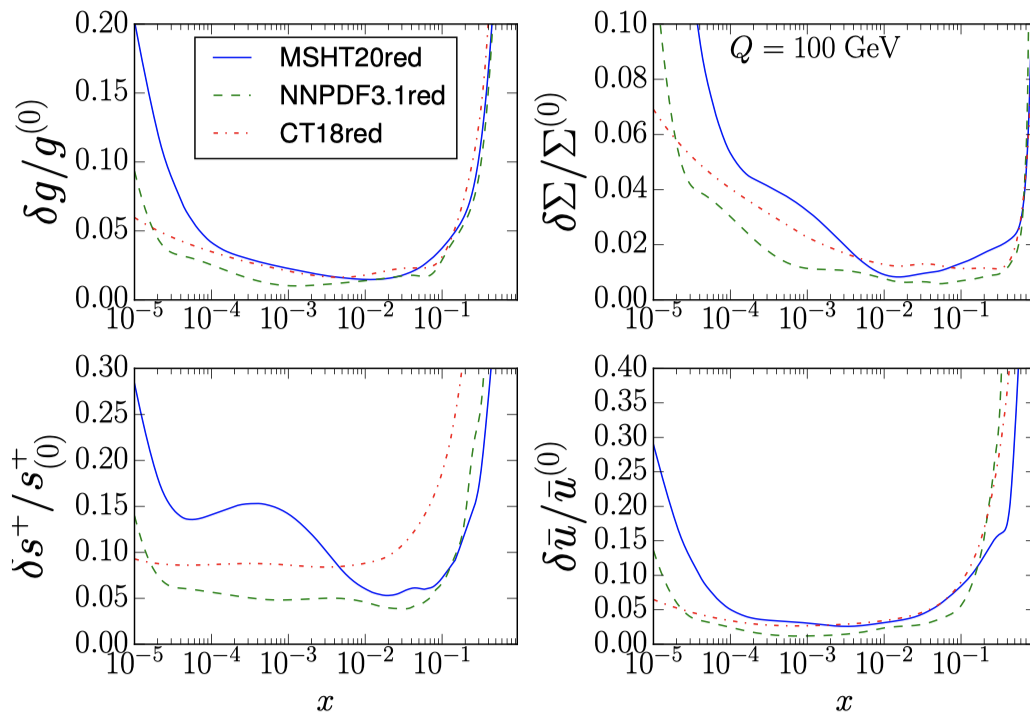
NNPDF3.1' is the smallest and CT18 is the largest, with MSHT20 in-between.

NB: MSHT20 nominally does not use a fixed tolerance, but instead cuts off an eigenvector direction when a particular experiment is badly fit. Thus, the uncertainty can be greatly affected by one experiment.



For some special cases, MSHT20 and CT18 were both defined using a $\Delta\chi^2$ of 10 (see above). The uncertainties are equivalent, as may be expected from them both using similar data sets, and in this case having the same criteria for determining the uncertainty.

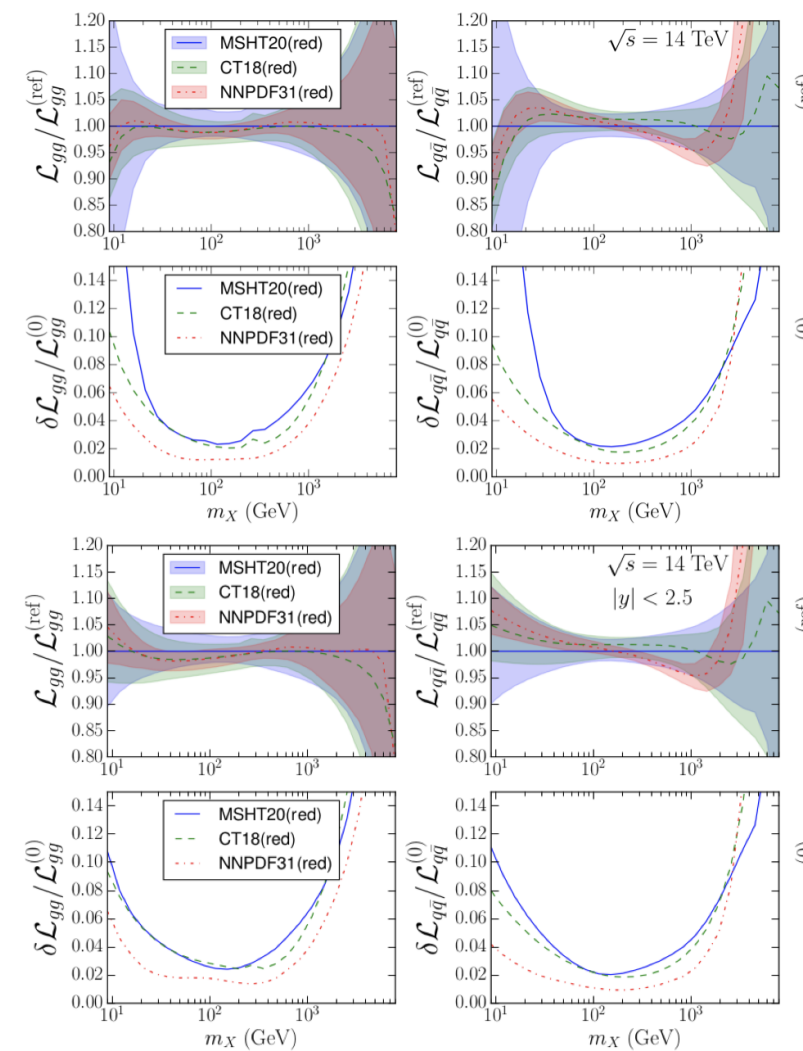
MSHT20-full-tolerance (i.e. the canonical MSHT20) in some cases has a larger uncertainty than MSHT20-T210, and in some cases smaller, indicating that the effective tolerance for the full fit is sometimes less than 10 and sometimes greater.



It is difficult to perform a directly similar comparison to NNPDF, as they don't use the Hessian formalism. However, as part of the PDF4LHC exercise, fits were carried out to a reduced data set, using similar theory parameters, in which equivalent results should be obtained, if the uncertainty criteria were the same.

CT18red and MSHT20red agree for the most part. There are fewer experiments included, so less likely for a particular experiment to truncate the uncertainty from a particular eigenvector.

NNPDF consistently has a consistently smaller uncertainty, indicative that their effective tolerance is smaller than either CT18red or MSHT20red.



This difference is even more prominent when the PDF luminosities are compared (above).