PDFs related for W boson mass measurement

Jun Gao

INPAC, Shanghai Jiao Tong University

25th Mini-workshop on the frontier of LHC-W boson mass:
Uncertainties and Opportunities

CHEP, Peking University

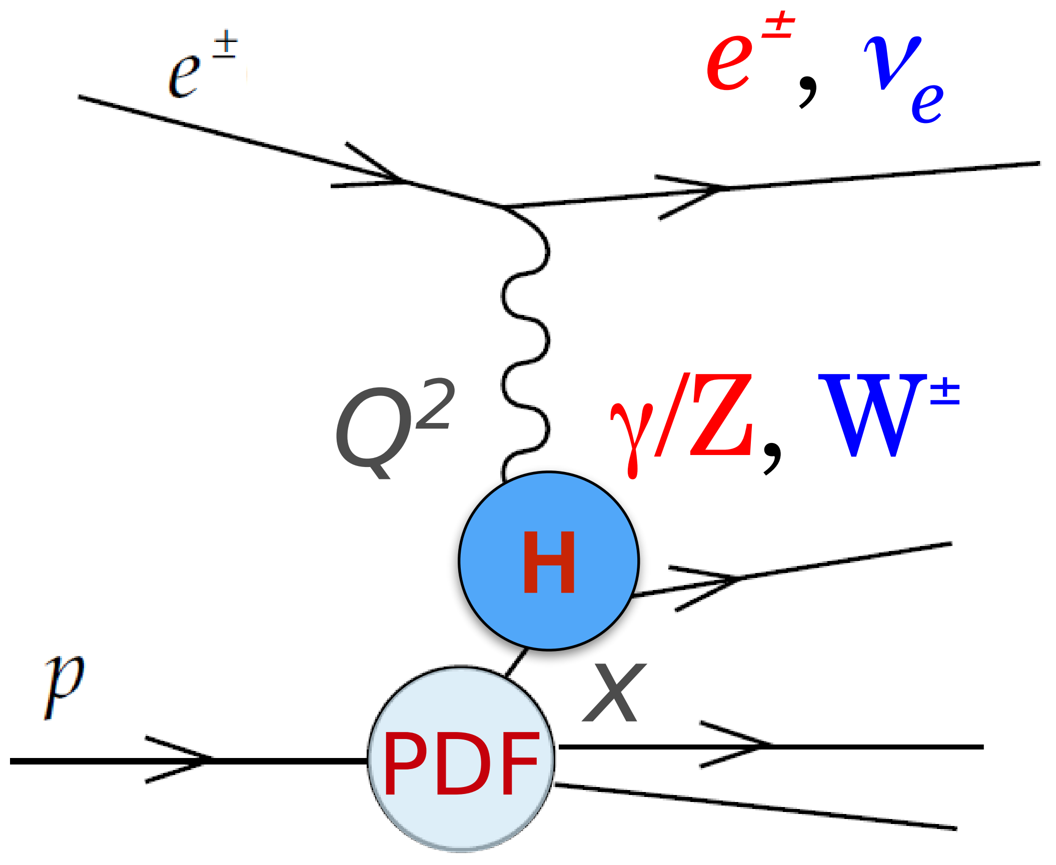
May 7, 2022

上海交通大學 Shanghai Jiao Tong University | SHANGHAI JIAO TONG UNIVERSITY Department of Physics | INPAC INSTITUTE OF NUCLEAR AND PARTICLE PHYSICS

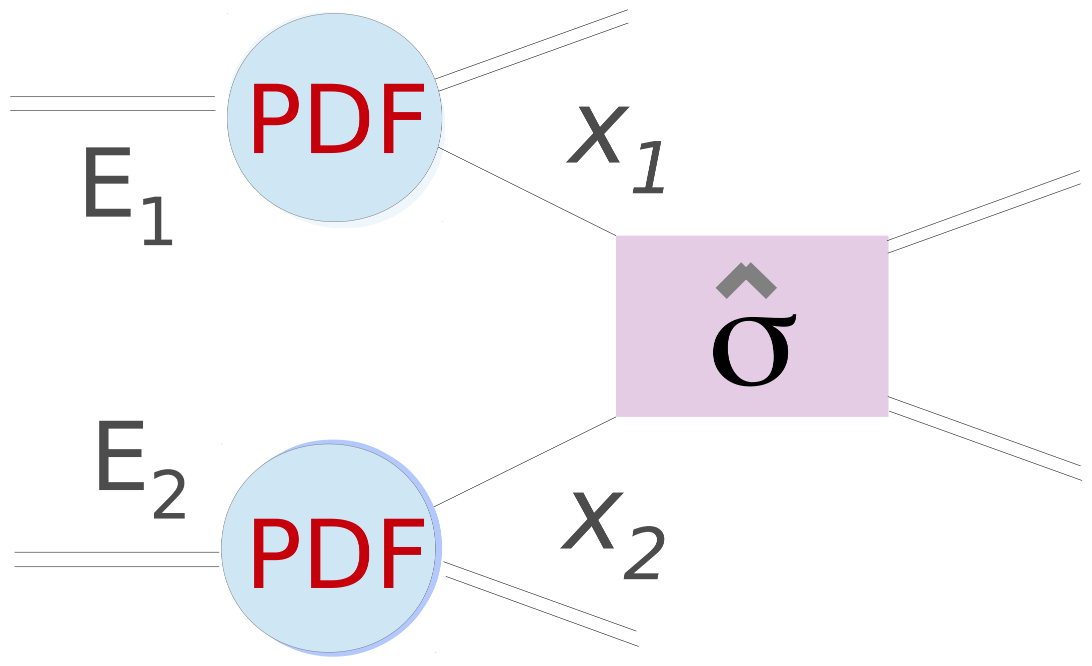
QCD collinear factorization

- QCD collinear factorization ensures universal separation of long-distance and short-distance contributions in high energy scatterings involving initial state hadrons, and enables predictions on cross sections

DIS structure functions



hadron-hadron collision



$$\sigma = \hat{\sigma} \otimes f_1 \otimes f_2$$

$$F_2(x, Q^2) = \sum_{i=q,\bar{q},g} \int_0^1 d\xi\, C_2^i(x/\xi, Q^2/\mu_r^2, {\mu_f}^2/\mu_r^2, \alpha_s(\mu_r^2)) \times f_{i/h}(\xi, \mu_f)$$

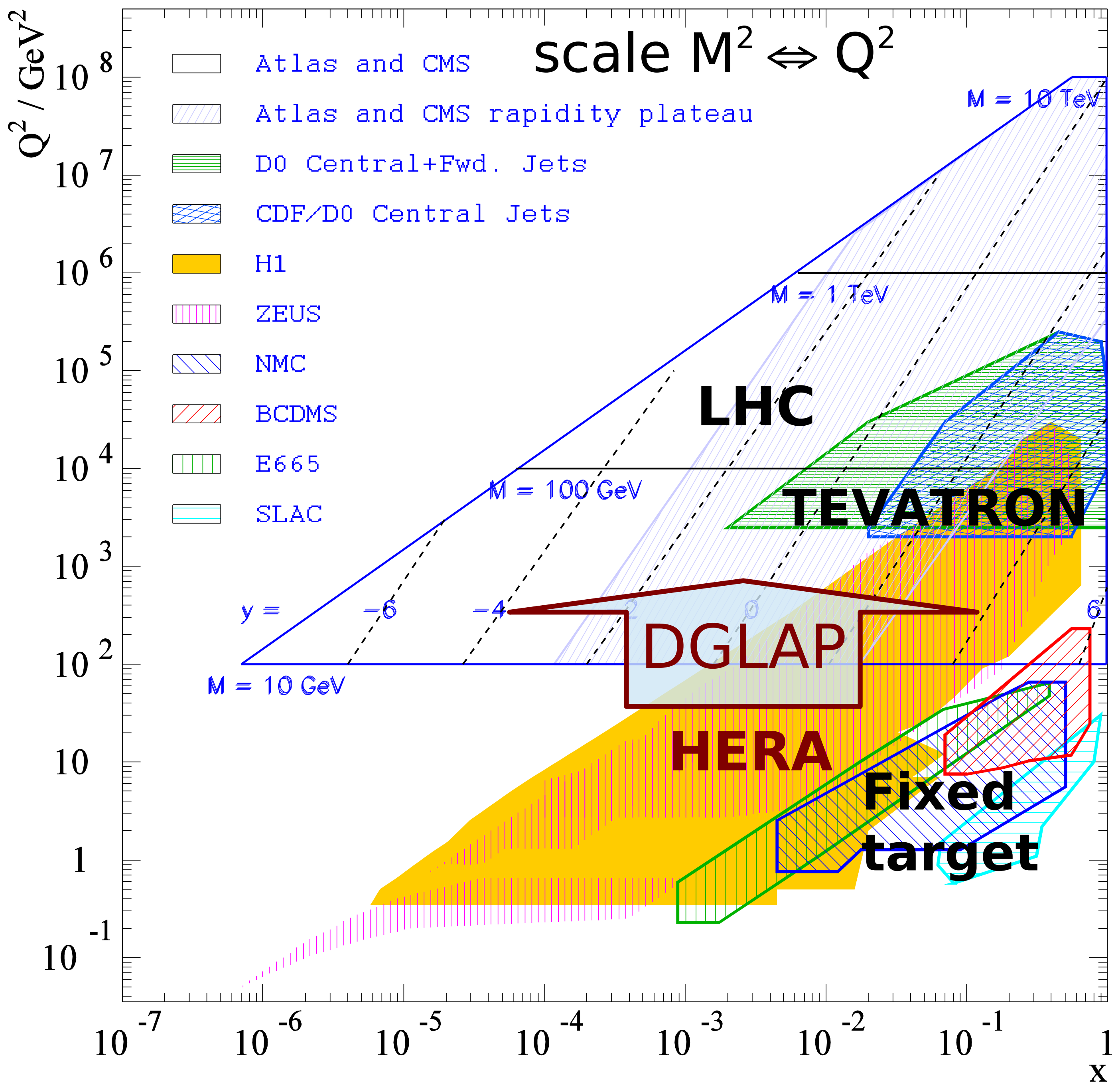
[Collins, Soper, Sterman, 1989]

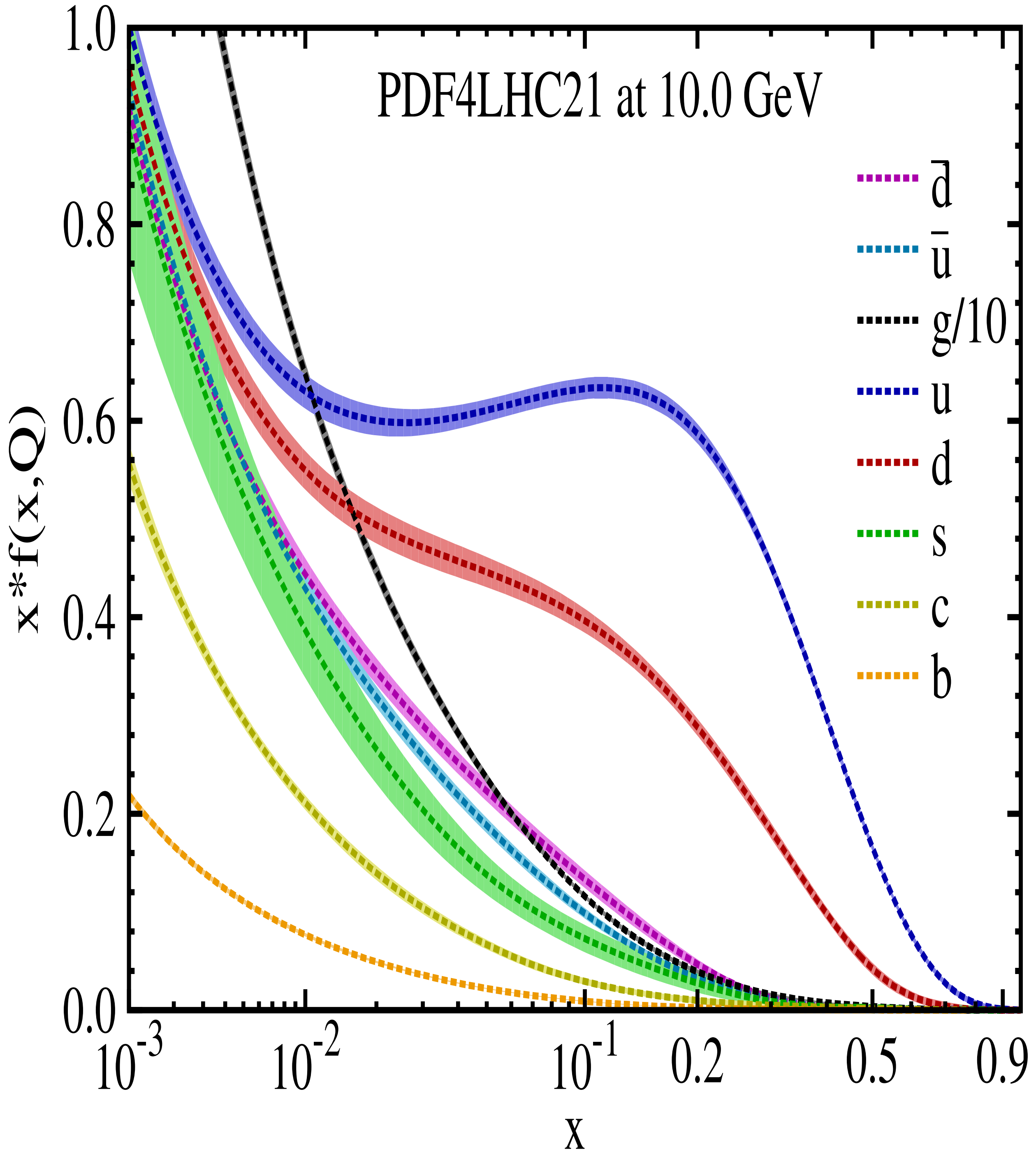
- coefficient functions, hard scattering; infrared (IR) safe, calculable in pQCD, independent of the hadron
- PDFs, reveal inner structure of hadrons; non-perturbative (NP) origin, universality, e.g. DIS vs. pp collisions
- factorization scale μ_f
- runnings of $f_{i/h}$ with μ_f are governed by the DGLAP equation

choose $\mu_f = \mu_r = Q$, thus Q dependence (scaling violation) of F_2 are mostly from PDFs and thus are predicted by the DGLAP evolution

Global analysis of PDFs

- PDFs are usually extracted from global analysis on variety of data, e.g., DIS, Drell-Yan, jets and top quark productions at fixed-target and collider experiments, with increasing weight from LHC, together with SM QCD parameters **[see 1709.04922, 1905.06957 for recent review articles]**





parameter variations

α_S(Mz)

Mc, Mb, Mt

QCD/EW corrections

nuclear corrections

EW parameters

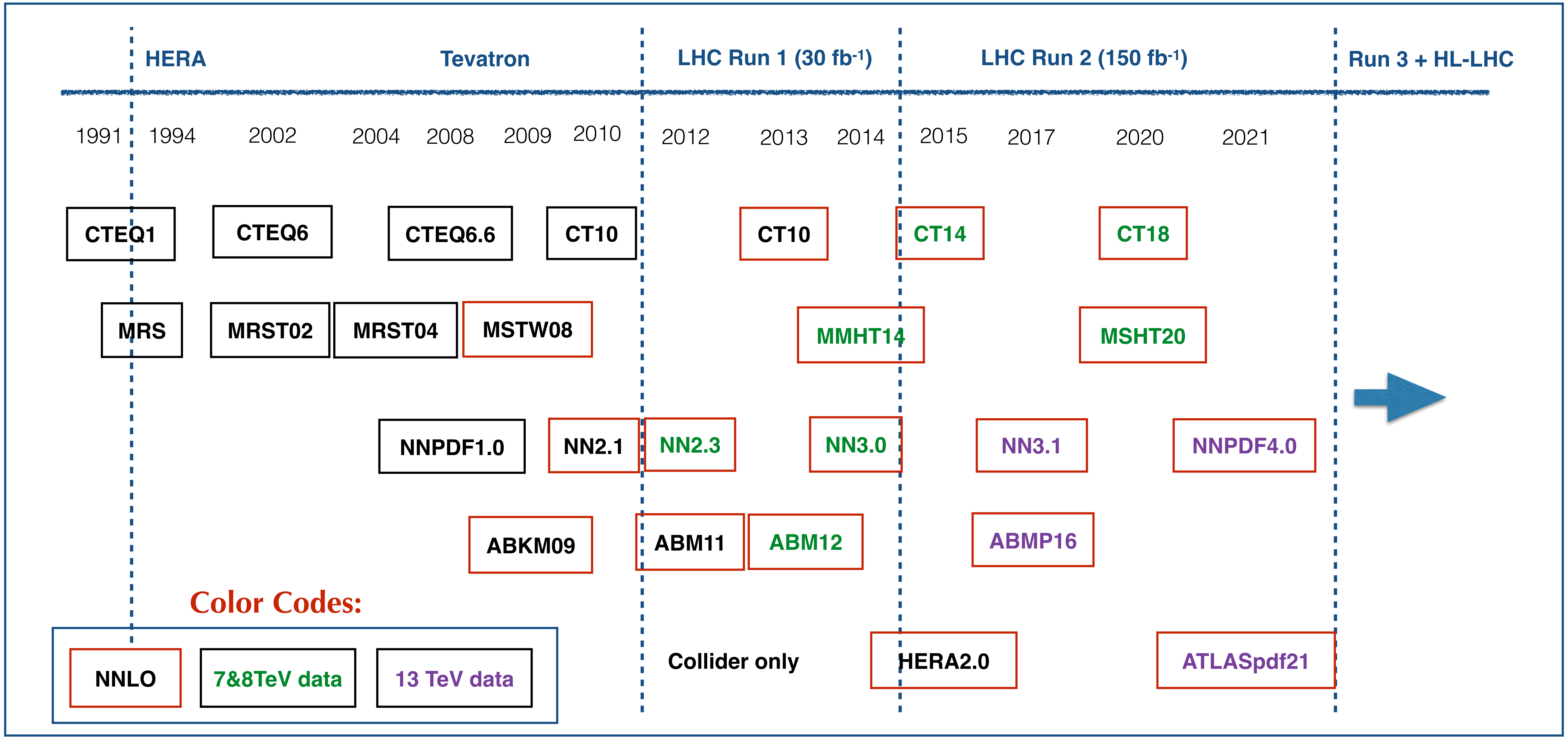
New Physics

- diversity of the analysed data are important to ensure flavor separation and to avoid theoretical/experimental bias; possible extensions to include EW parameters and possible new physics for a self-consistent determination

- alternative approach from lattice QCD simulations, for various PDF moments or PDFs directly calculated in x-space with large momentum effective theory or pseudo-PDFs **[2004.03543]**

Major analysis groups

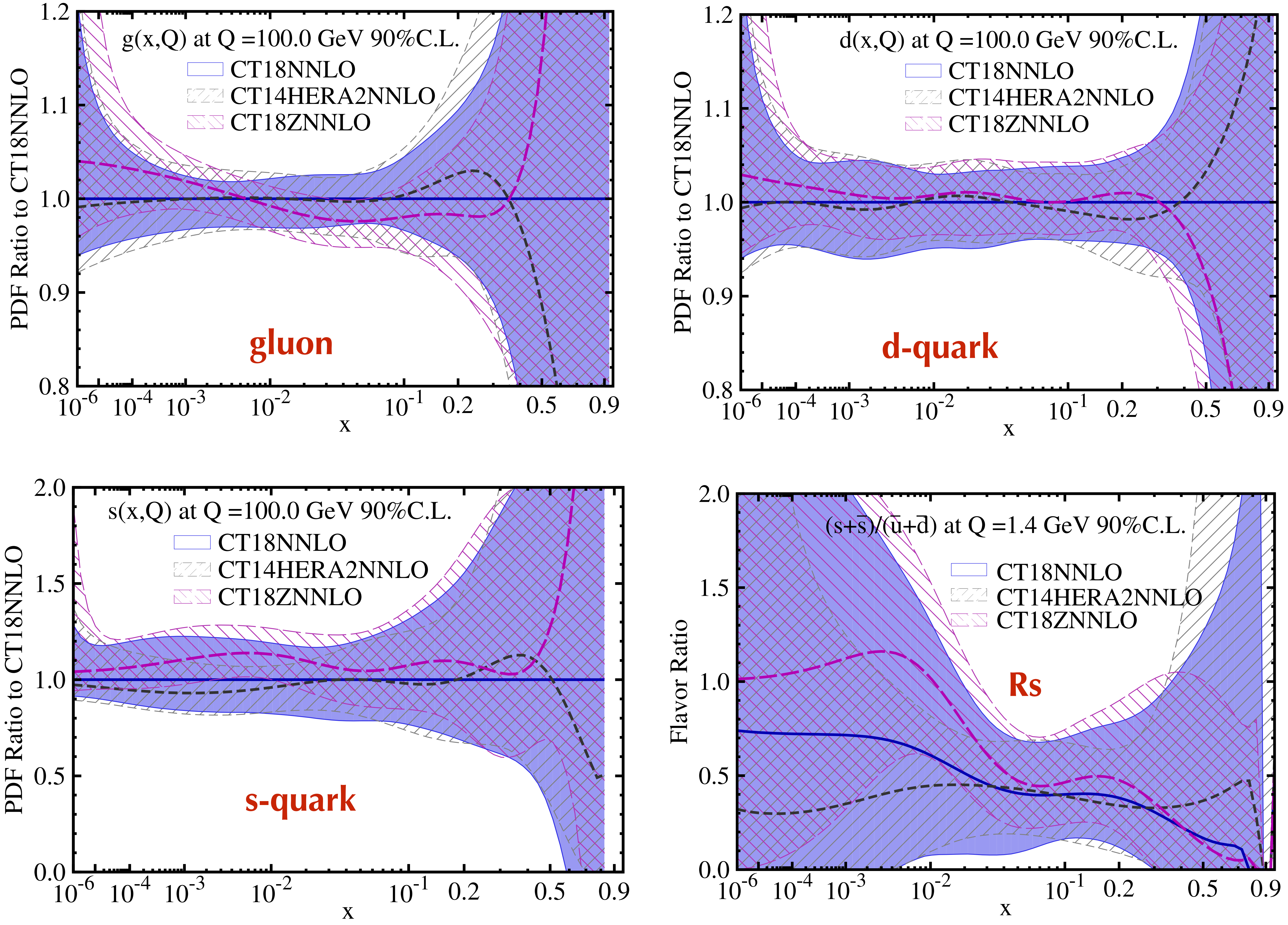
- PDFs provided by several major analysis groups (CT, MSHT, NNPDF, ABM, HERAPDF, ATLASpdf, CJ, JAM…) using slightly different heavy-quark schemes, selections of data, and methodologies



must have as many independent analyses as possible to have a faithful determination of PDFs and their uncertainties; state of the art PDFs are extracted at NNLO in QCD and with numerous LHC data

CTEQ-TEA PDFs

- CT18 PDFs show moderate reductions of PDF uncertainties due to new LHC data sets, and agree with previous CT14 within uncertainties; alternative fits CT18Z/A/X for evaluation of certain systematic effects

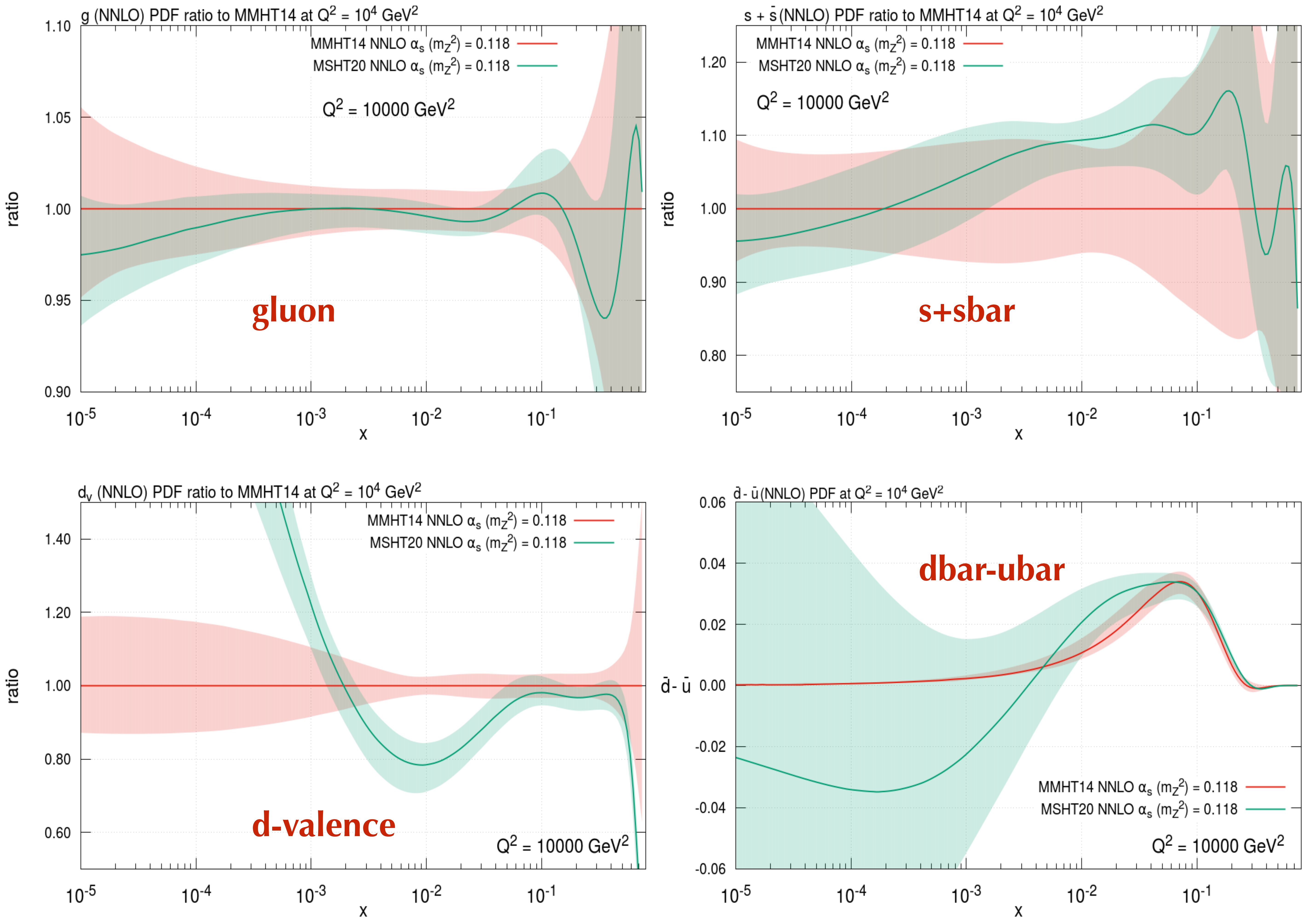


- CT18 vs CT14: gluon unc. reduced everywhere (jets, Z pT, top); d-quark unc. reduced at x~0.2 (LHCb W/Z); s-quark almost unchanged
- ATLAS 7 TeV W/Z data are not included in CT18 fit but in CT18A; CT18X uses a x-dependent scale in DIS to mimic small-x resummations
- CT18Z includes both variations, differences wrt. CT18 are most significant in s-quark and gluon/sea-quarks

[CT18, 1912.10053]

MSHT PDFs

- MSHT20 (Mass Scheme Hessian Tolerance) PDFs adopt an extended parametrization form, as comparing to MMHT14, to accomodate for newly included LHC precision data

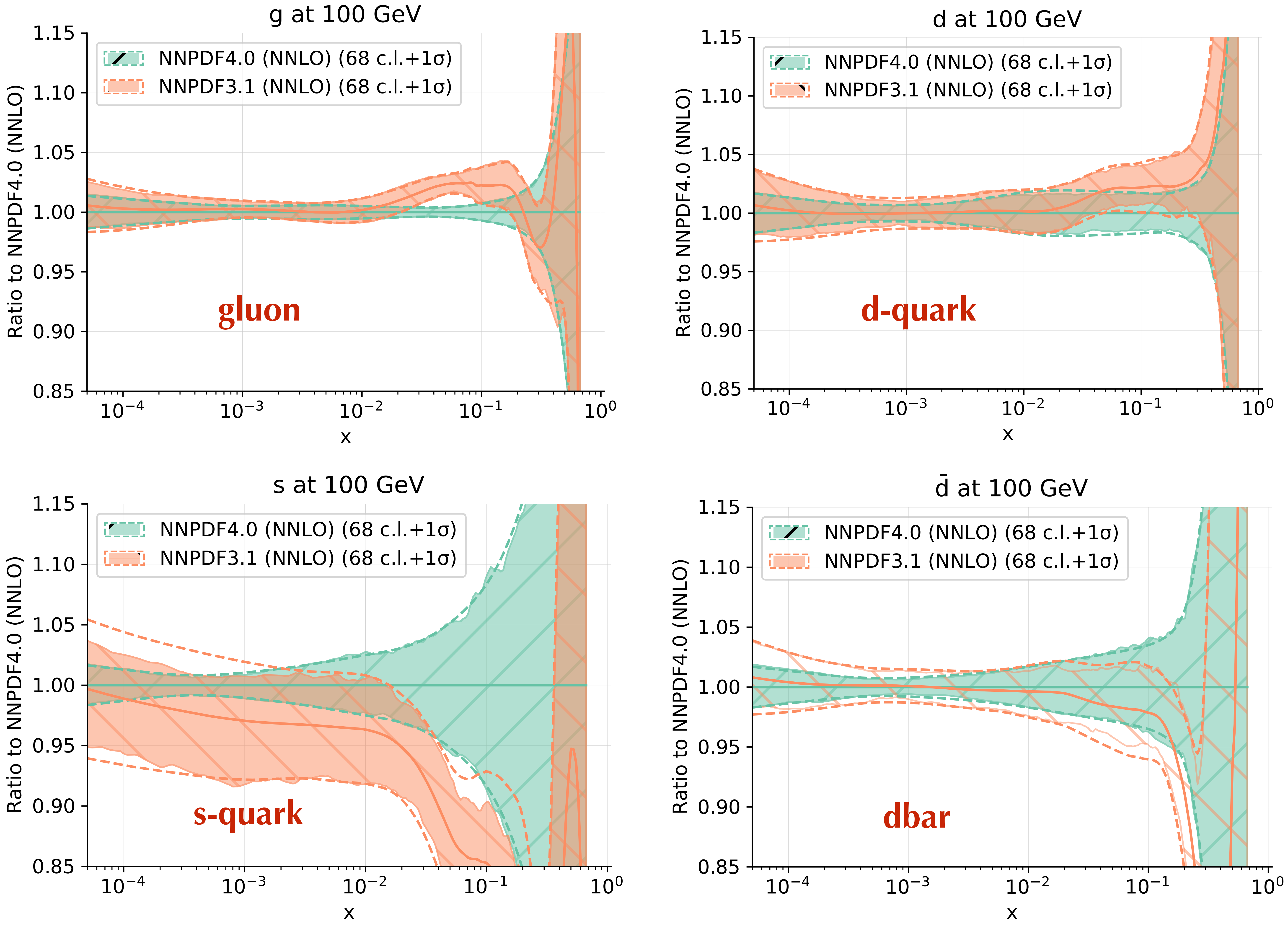


- central of gluon PDF remains mostly unchanged except for a suppression at x>0.2; moderate reduction on gluon uncertainty
- enhancement of s-quark at intermediate x region and large reduction on uncertainty, due to LHC 7 TeV W/Z data and update of dimuon theory calculations
- new parametrization allows a change of d-valence shape to better fit LHC W/Z data, and also large uncertainties of isospin asymmetry in small-x region

[MSHT20, 2012.04684]

NNPDFs

✦ NNPDF4.0 PDFs improves previous NNPDF3.1 with a major update on methodologies and a dedicated global survey and selection of available LHC data



❖ changes on parametrization and NN architecture, optimization algorithm; additional positivity and integrability constraints and post-fit selections

❖ central PDF of NNPDF4.0 is generally consistent with NNPDF3.1 except for a notable decrease of gluon PDF at x~0.1 and moderate increase of strangeness

❖ NNPDF4.0 shows PDF uncertainty of ~1-2% at data constrained region, largely reduced comparing to NNPDF3.1

[NNPDF4.0, 2109.02653]

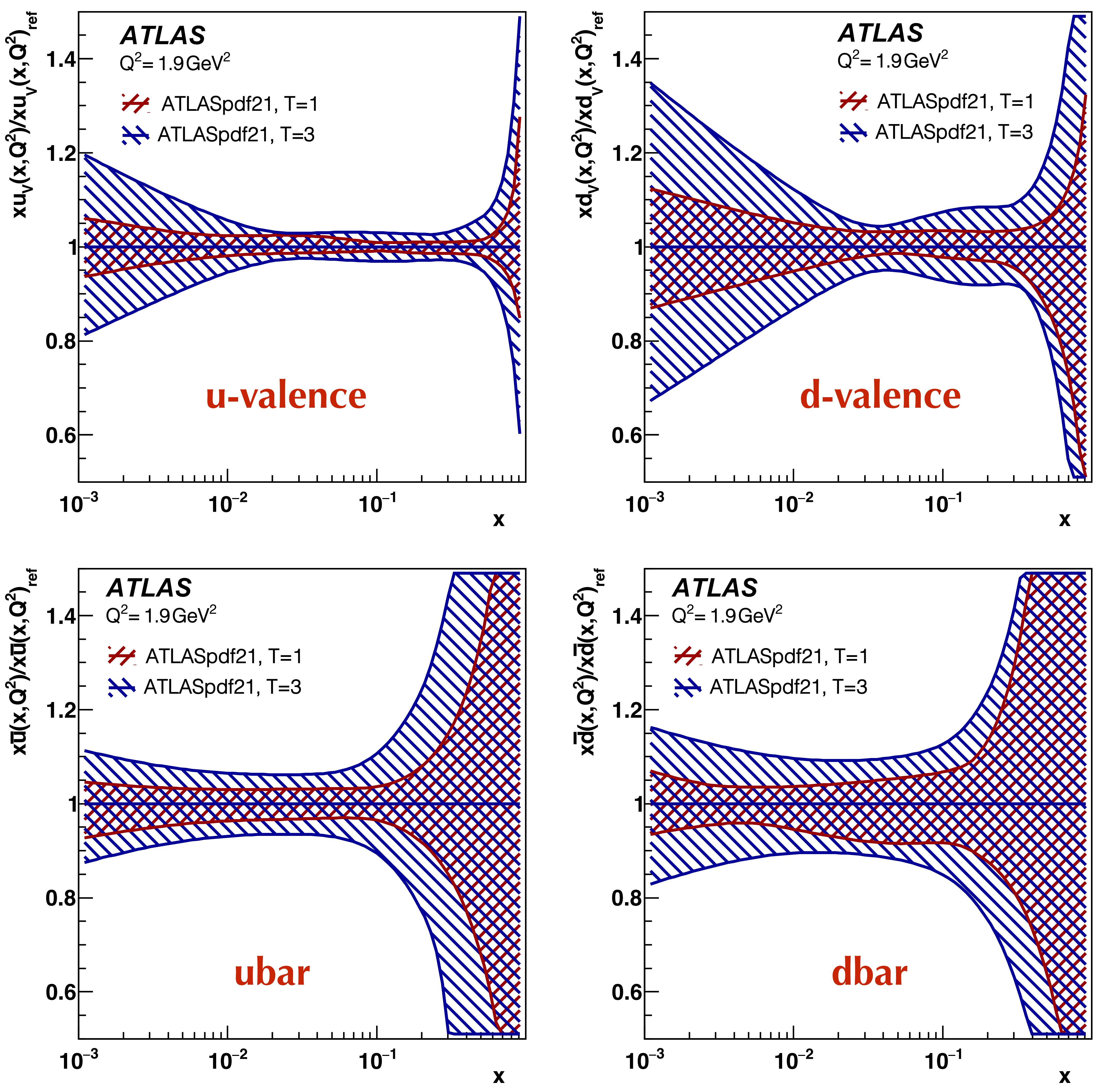
ATLAS PDFs

- ATLAS releases the most recent 2021 PDFs based on a NNLO analysis of HERA combined data and a variety of ATLAS data from 7, 8 to 13 TeV and with several new features explored

Data sets included and χ^2 **[ATLAS,2112.11266]**

Total χ^2/NDF	2010/1620
HERA χ^2/NDP	1112/1016
HERA correlated term	50
ATLAS W, Z 7 TeV χ^2/NDP	68/55
ATLAS Z/γ^* 8 TeV χ^2/NDP	208/184
ATLAS W 8 TeV χ^2/NDP	31/22
ATLAS W and Z/γ^* 7 and 8 TeV correlated term	71 = (38 + 33)
ATLAS direct γ 13/8 TeV χ^2/NDP	27/47
ATLAS direct γ 13/8 TeV correlated term	6
ATLAS V+jets 8 TeV χ^2/NDP	105/93
ATLAS $t\bar{t}$ 8 TeV χ^2/NDP	13/20
ATLAS $t\bar{t}$ 13 TeV χ^2/NDP	25/29
ATLAS inclusive jets 8 TeV χ^2/NDF	207/171
ATLAS V+jets 8 TeV and $t\bar{t}$ + jets 8,13 TeV and $R = 0.6$ inclusive jets 8 TeV correlated term	87 = (16 + 9 + 21 + 41)

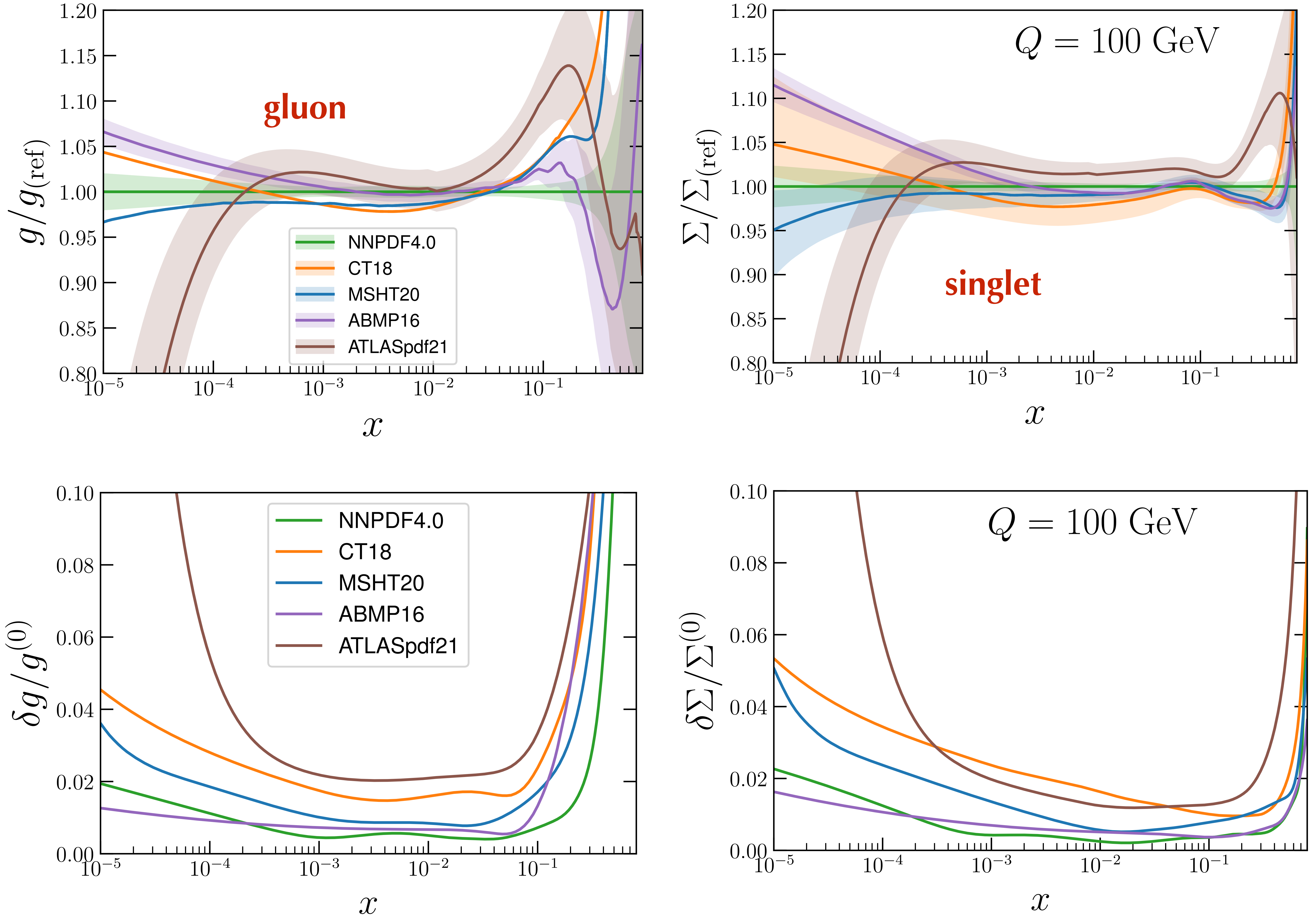
full uncertainty consists of experimental, theoretical, model, and parametrization uncertainties; evaluated either using $\Delta\chi^2$=1 or $\Delta\chi^2$=9 (a global tolerance of T=3)



PDF benchmarking

- Many ongoing efforts on comparisons and understanding of differences of up-to-date PDFs, in order to have a faithful determination of PDFs and its uncertainties

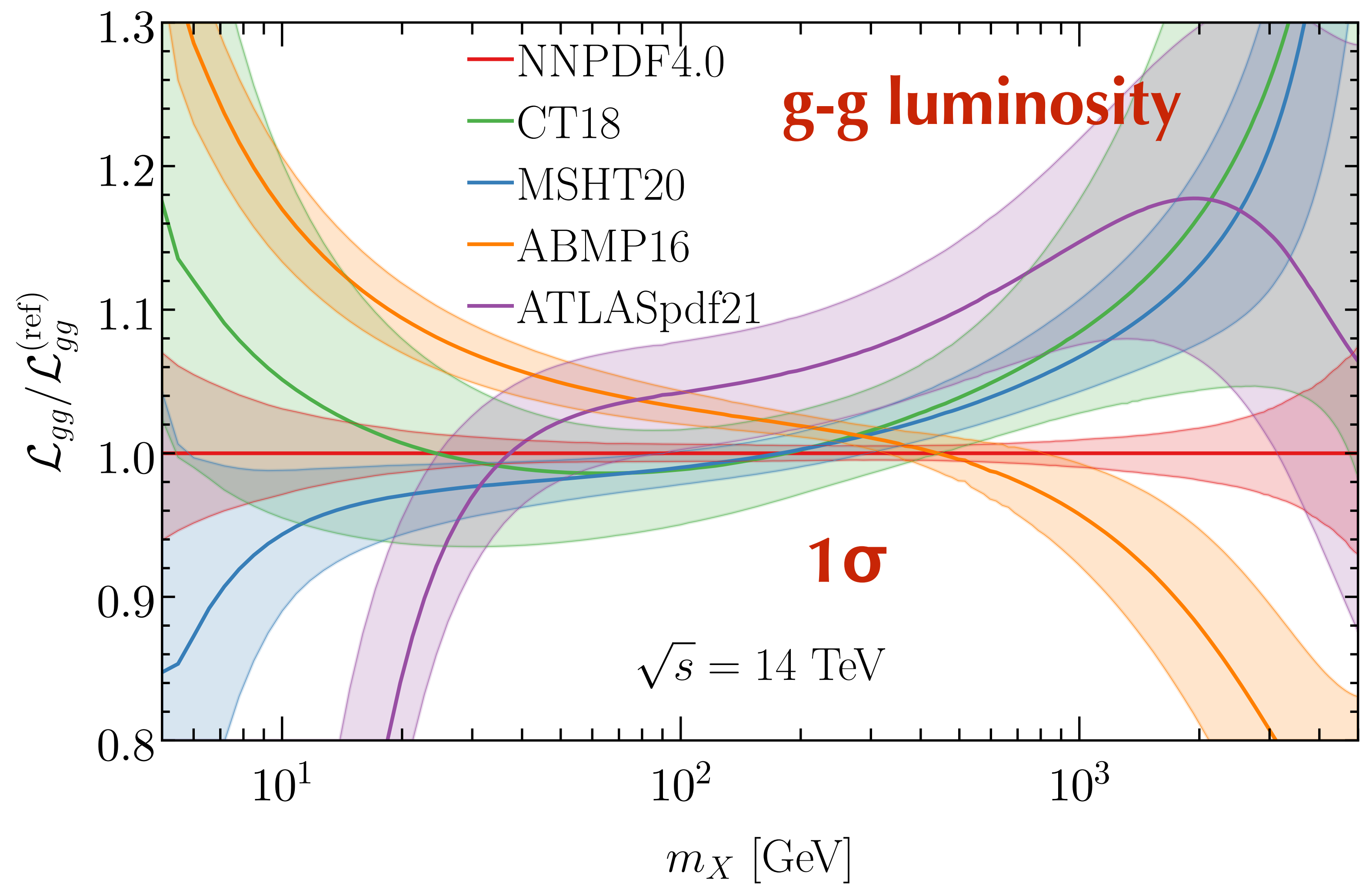
[Snowmass 2021, 2203.13923]

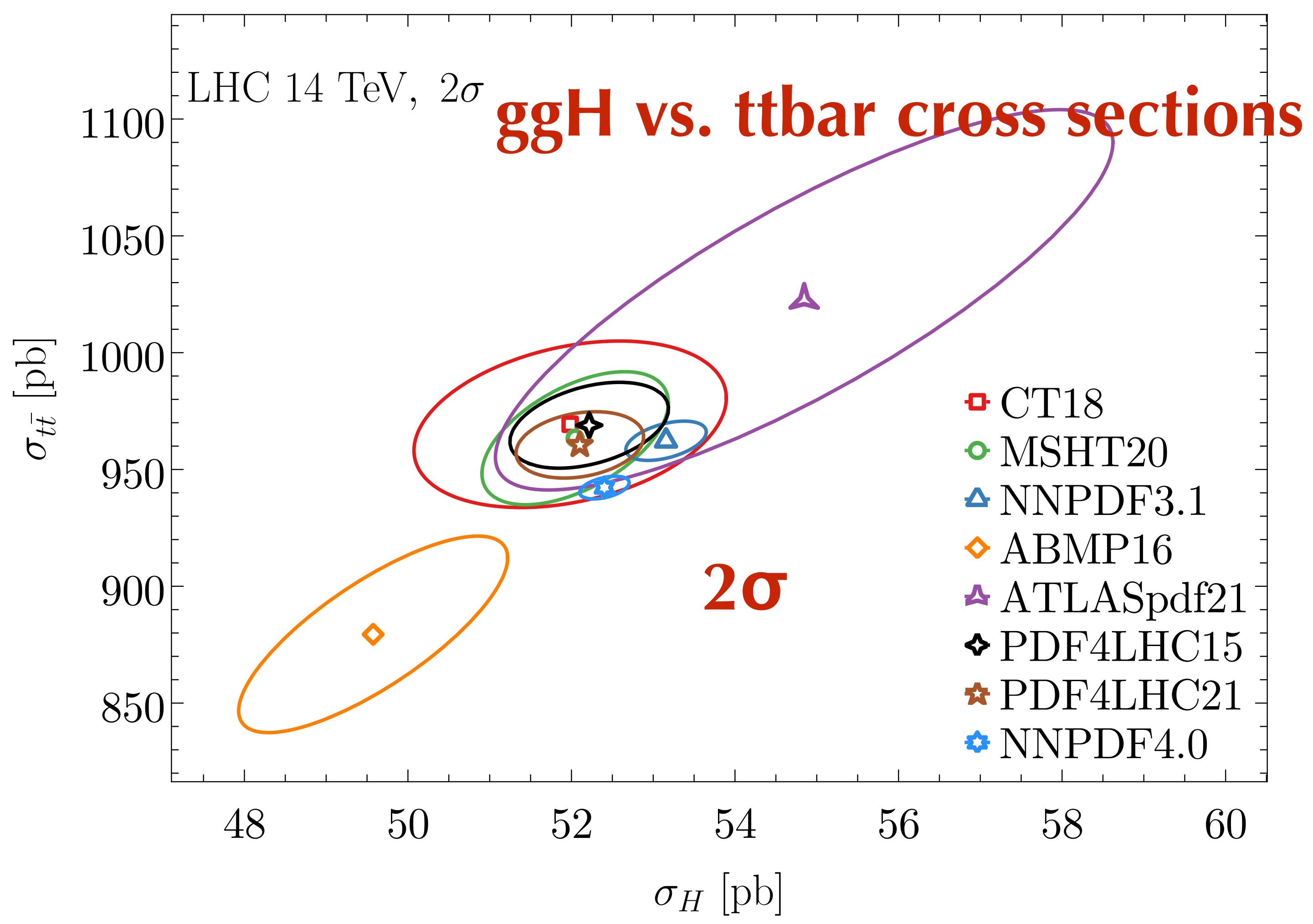


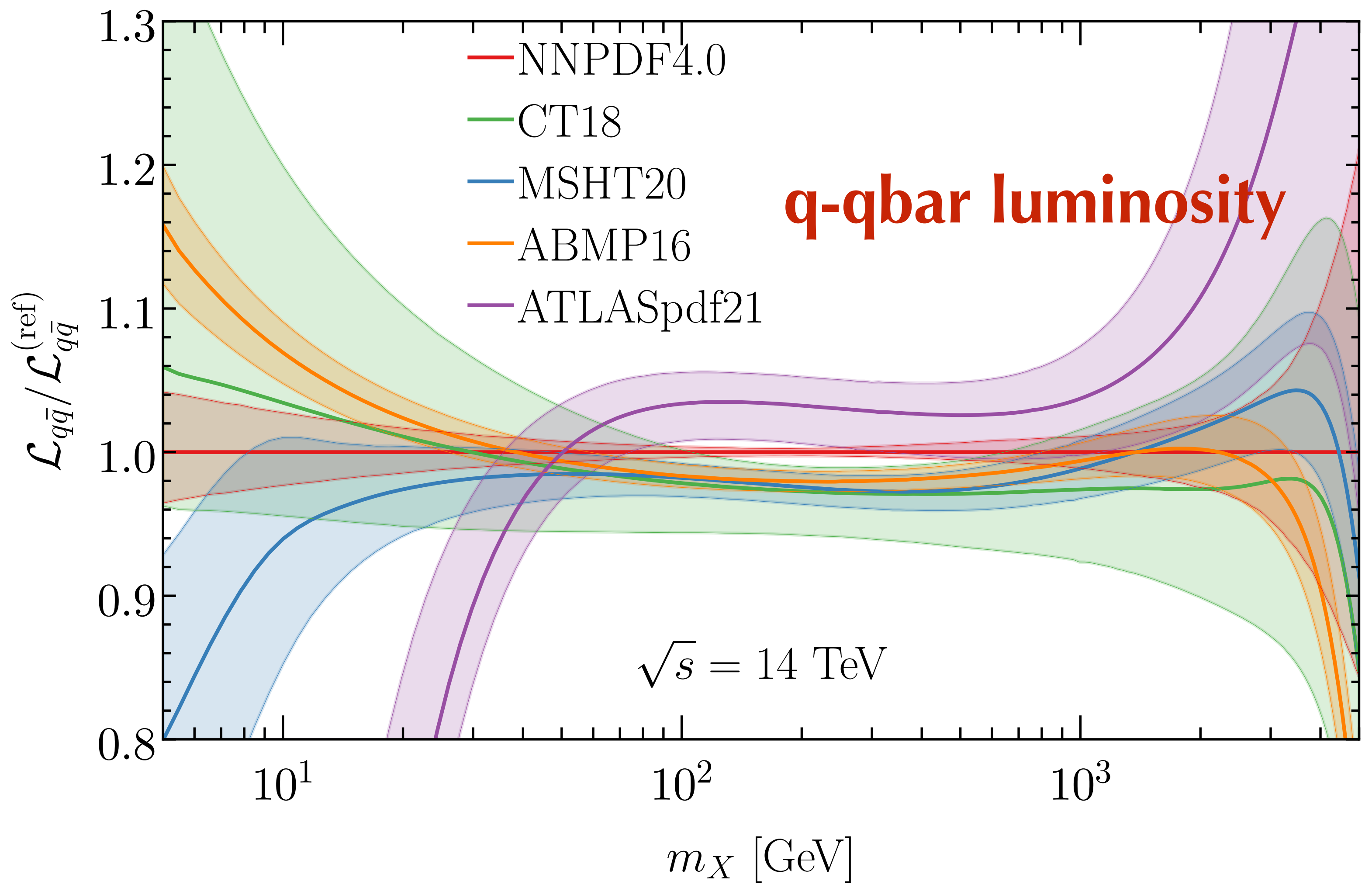
- general agreement between different groups (NN4.0, CT18, MSHT20, ABMP16, ATLAS21) over the range of x in 10^{-4} to 10^{-1} within uncertainties
- gluon: notable differences at x~0.2, with 2σ for NN vs. CT&MSHT; singlet: ATLASpdf deviate at x<10^{-4} due to Q^2>10 GeV^2 applied on HERA data, and at x>0.2 due to lack of fixed-target data
- NN and ABMP show uncertainty of ~1-2% in constrained region mostly due to methodologies; CT18 being conservative among all fits; ATLAS unc. blow up in unconstrained region

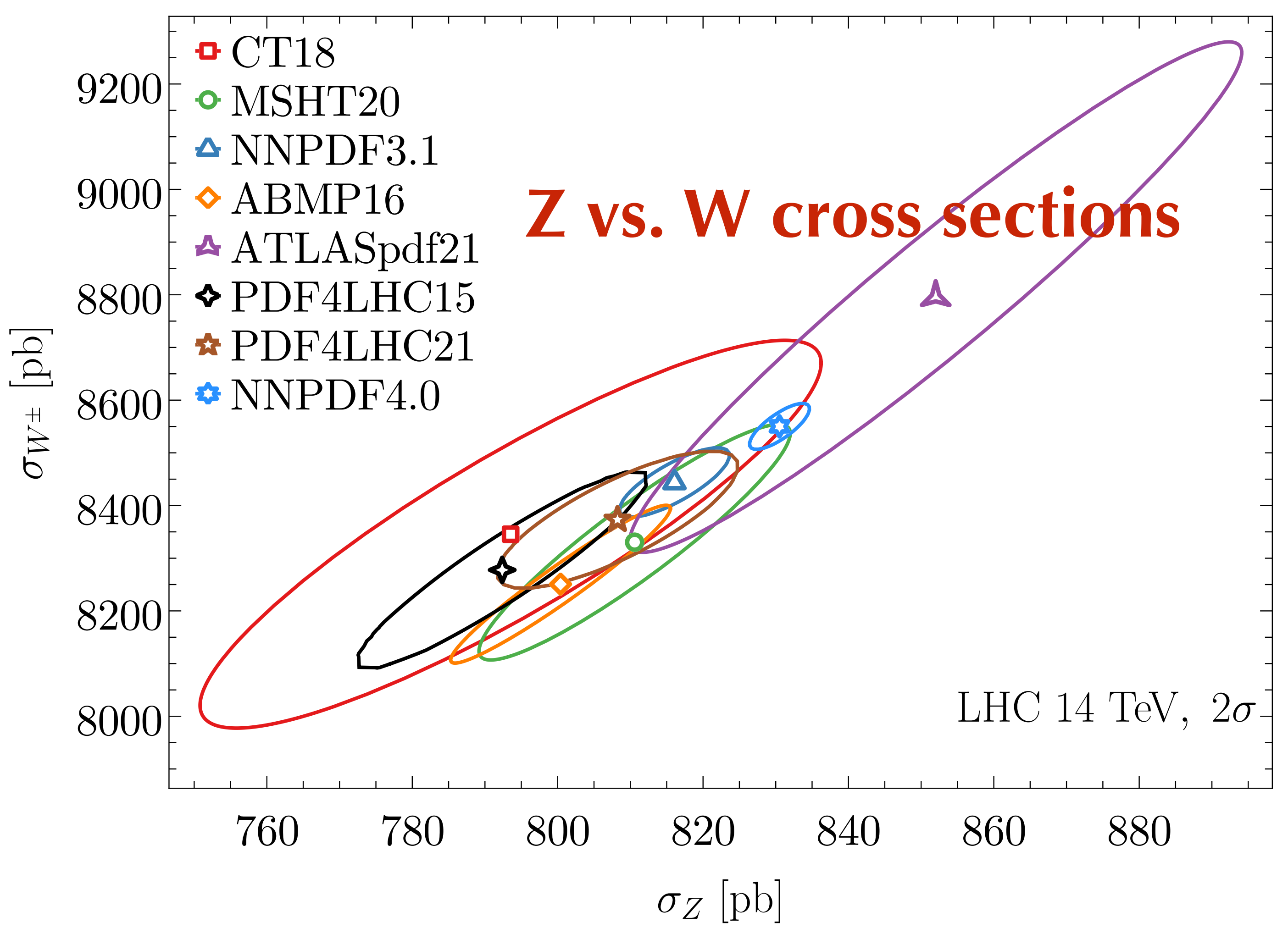
PDF benchmarking

- Spread of PDFs from different groups propagates into the parton-parton luminosity or cross sections at the LHC 14 TeV and some cases enlarged due to (anti-)correlations between different x-regions/flavors







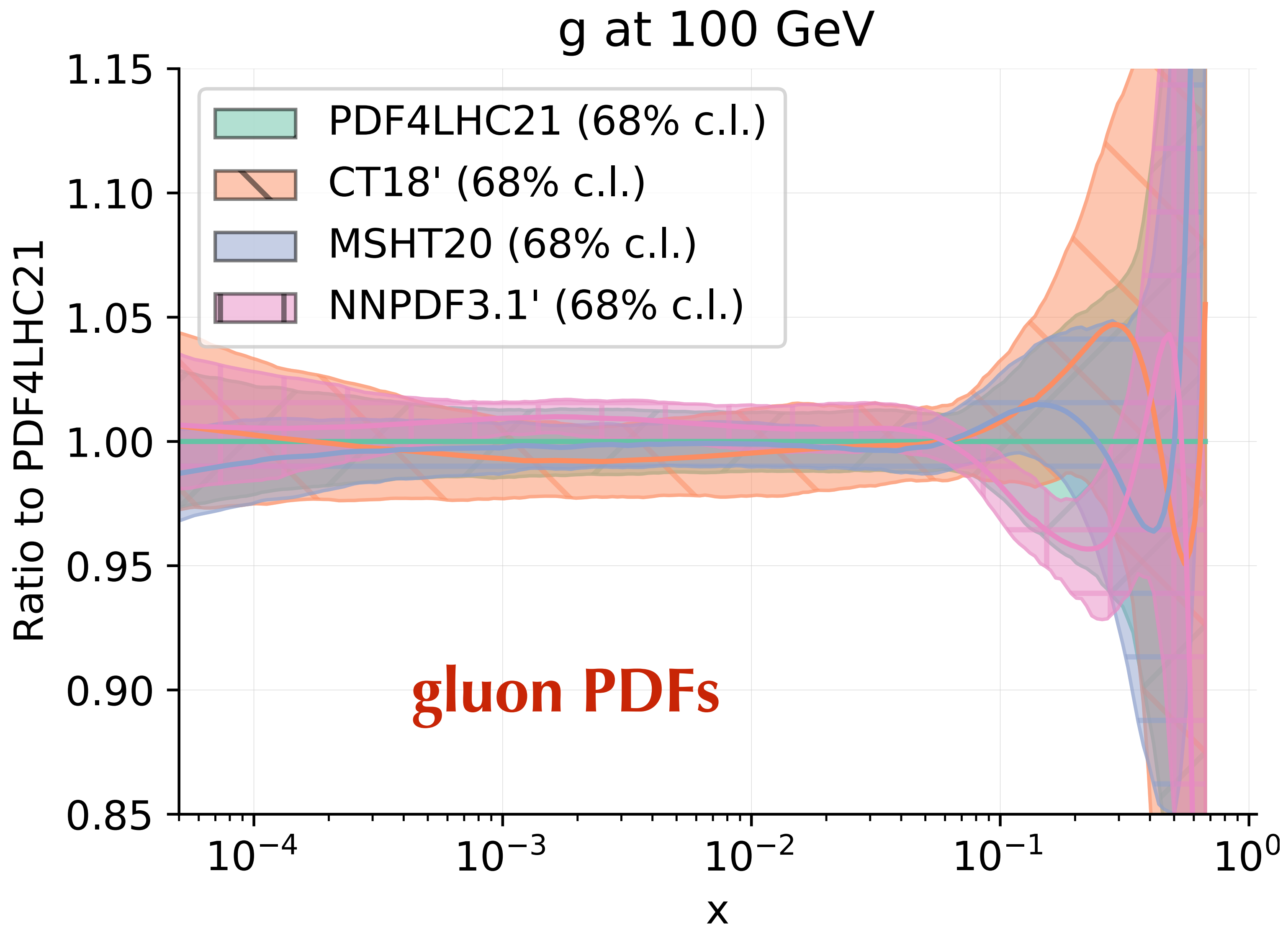


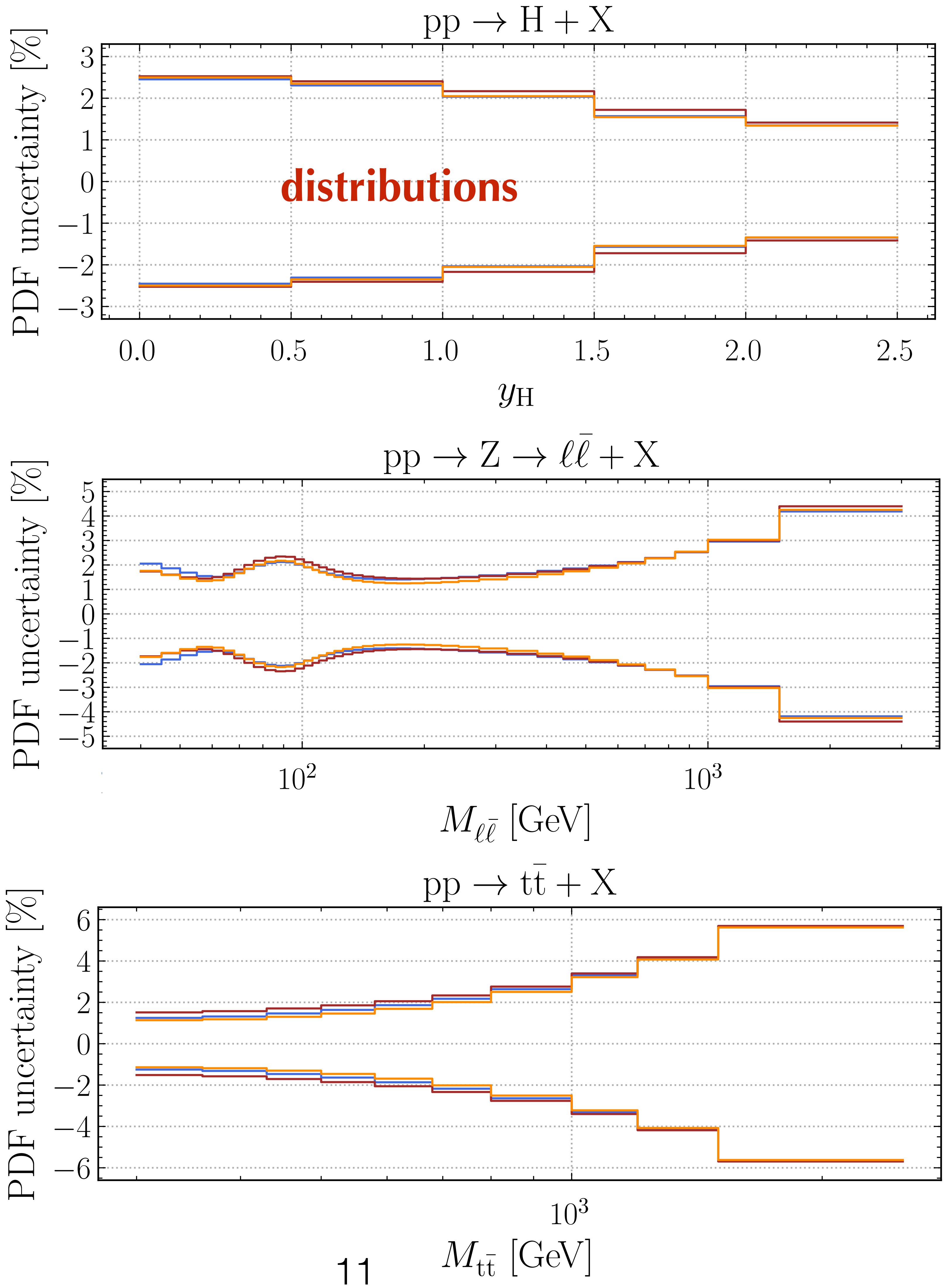
- g-g luminosity shows a spread of more than 20% in the multi-TeVs region; q-qbar luminosity agrees better in general except at a mass around 300 GeV
- 2σ error ellipse shown for cross sections of standard candle processes; NNPDF4.0 shows an uncertainty of less than 0.5% while CT18 2σ ellipse seems to cover most groups
- NNPDF4.0 and 3.1 show no overlaps on cross sections even at 2σ

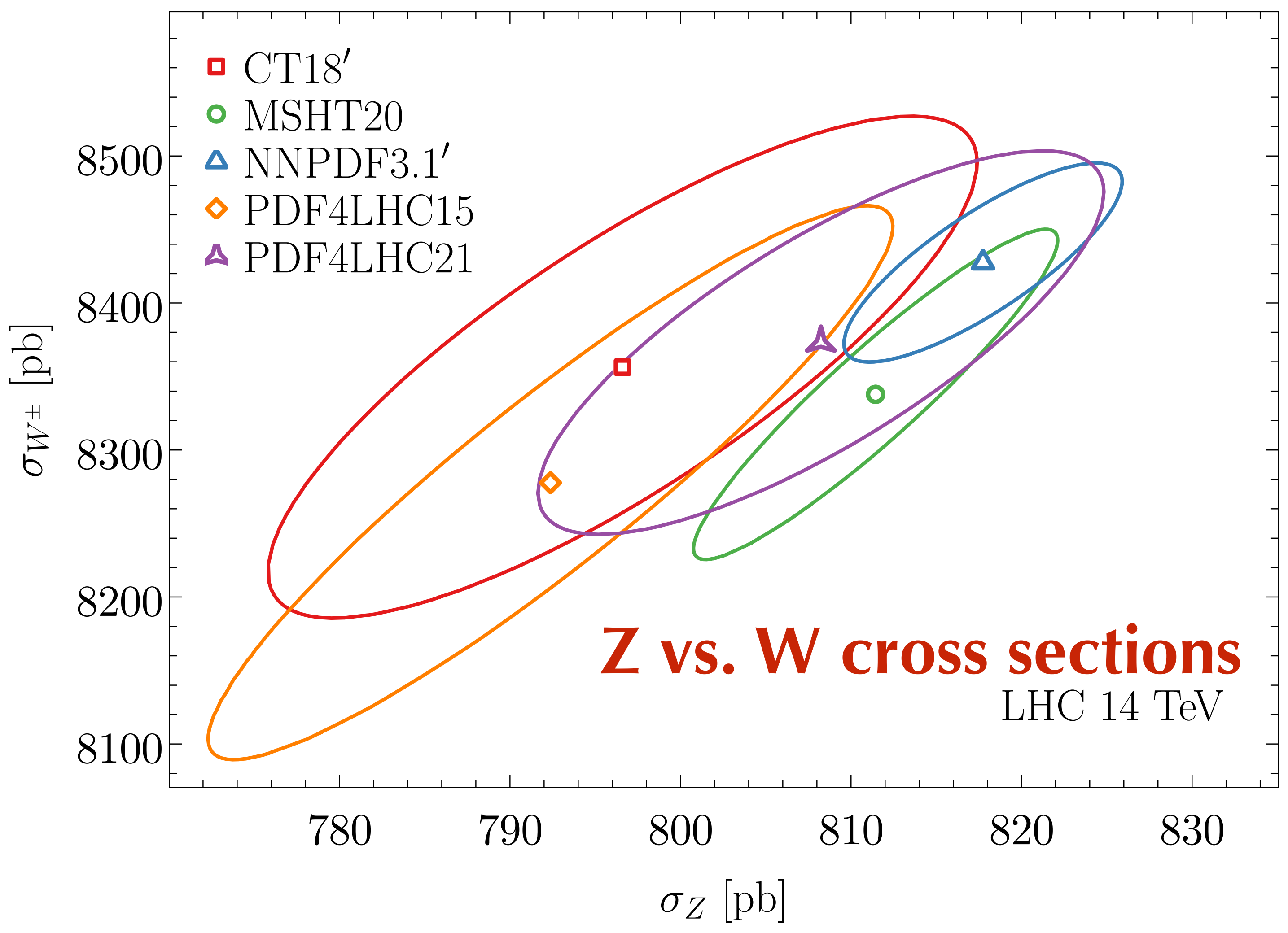
[Snowmass 2021, 2203.13923]

PDF4LHC recommendations

✦ The PDF4LHC group performs extensive benchmarks on methodologies of several groups, and presents the PDF4LHC21 PDFs, an effective combination of CT18′, MSHT20 and NNPDF3.1′, for LHC Run3 usage







- ❖ CT18′ differs slightly from CT18 by using m_c=1.4 GeV; NNPDF3.1′ differs from 3.1 by including additional jet and top-quark data
- ❖ PDF4LHC21 PDFs are presented in the form of either a MC set of 100 replica PDFs or a Hessian set of 40 PDFs
- ❖ PDF unc. at the level of 2~3% for the inclusive cross section and 5~10% for distribution at multi-TeVs region

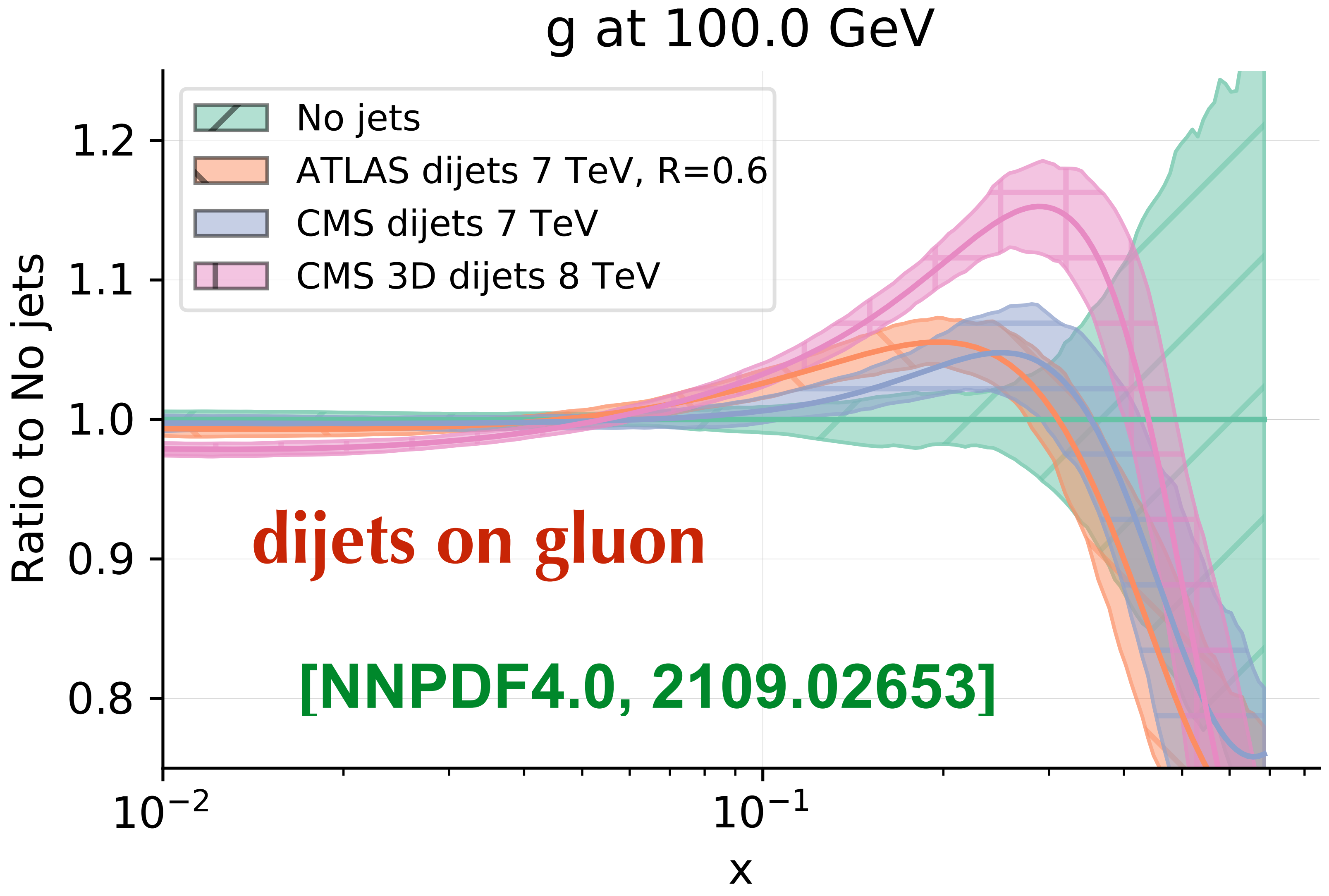
[PDF4LHC21, 2203.05506]

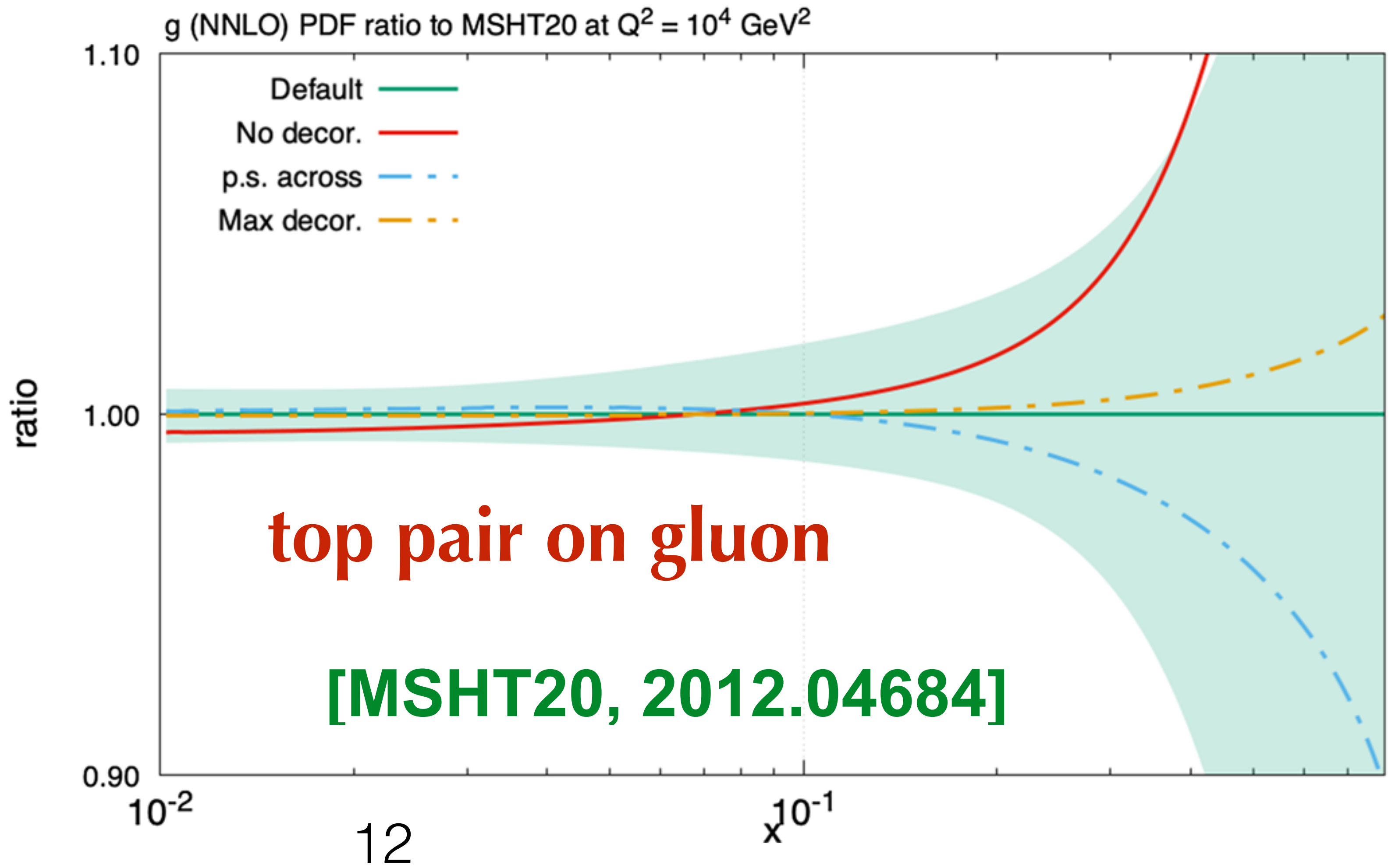
Impact of LHC data

- LHC provides measurements on a variety of PDF-sensitive standard candle processes with precision reaching a few percents; Their impact is subjected to possible tensions among different data and complications of the experimental systematic errors

χ^2 of LHC data in MSHT20

Data set	NLO	NNLO
ATLAS W^+, W^-, Z [118]	34.7/30	29.9/30
CMS W asym. $p_T > 35$ GeV [153]	11.8/11	7.8/11
CMS asym. $p_T > 25, 30$ GeV [154]	11.8/24	7.4/24
LHCb $Z \to e^+e^-$ [155]	14.1/9	22.7/9
LHCb W asym. $p_T > 20$ GeV [156]	10.5/10	12.5/10
CMS $Z \to e^+e^-$ [157]	18.9/35	17.9/35
ATLAS High-mass Drell–Yan [158]	20.7/13	18.9/13
CMS double diff. Drell–Yan [71]	222.2/132	144.5/132
Tevatron, ATLAS, CMS $\sigma_{t\bar{t}}$ [92,93]	22.8/17	14.5/17
LHCb 2015 W, Z [94,95]	114.4/67	99.4/67
LHCb 8 TeV $Z \to ee$ [96]	39.0/17	26.2/17
CMS 8 TeV W [97]	23.2/22	12.7/22
ATLAS 7 TeV jets [18]	226.2/140	221.6/140
CMS 7 TeV $W + c$ [98]	8.2/10	8.6/10
ATLAS 7 TeV high precision W, Z [20]	304.7/61	116.6/61
CMS 7 TeV jets [99]	200.6/158	175.8/158
CMS 8 TeV jets [100]	285.7/174	261.3/174
CMS 2.76 TeV jet [106]	124.2/81	102.9/81
ATLAS 8 TeV Z p_T [74]	235.0/104	188.5/104
ATLAS 8 TeV single diff $t\bar{t}$ [101]	39.1/25	25.6/25
ATLAS 8 TeV single diff $t\bar{t}$ dilepton [102]	4.7/5	3.4/5
CMS 8 TeV double differential $t\bar{t}$ [104]	32.8/15	22.5/15
CMS 8 TeV single differential $t\bar{t}$ [107]	12.9/9	13.2/9
ATLAS 8 TeV High-mass Drell–Yan [72]	85.8/48	56.7/48
ATLAS 8 TeV W [105]	84.6/22	57.4/22
ATLAS 8 TeV W + jets [103]	33.9/30	18.1/30
ATLAS 8 TeV double differential Z [73]	157.4/59	85.6/59
Total	5822.0/4363	5121.9/4363



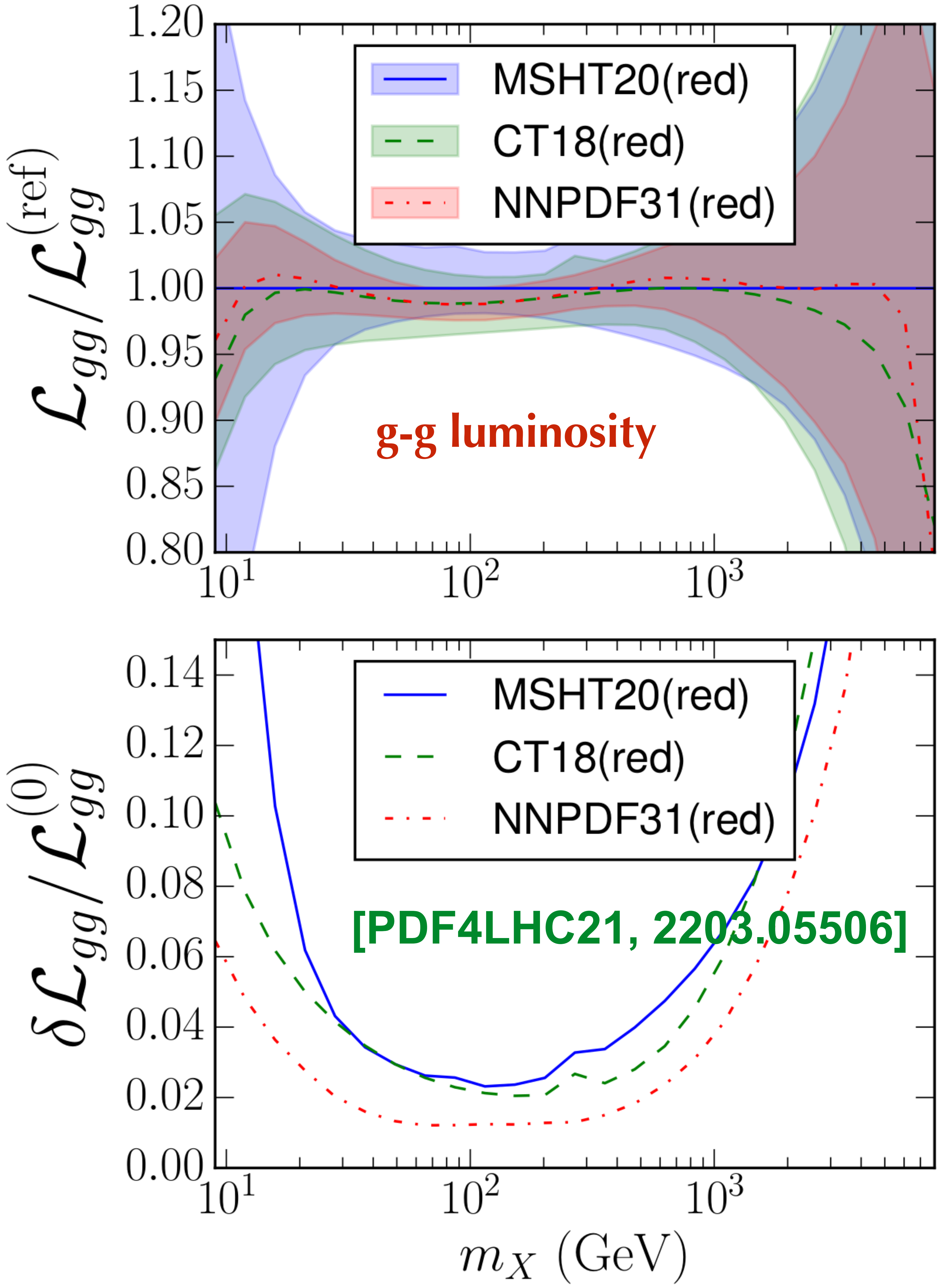


- fit quality to LHC data is moderate in general or very poor for specific data sets
- decorrelation/regularization of experimental systematics or theoretical errors are added to reach a reasonable χ^2
- appraisal and selection of LHC data become a major task

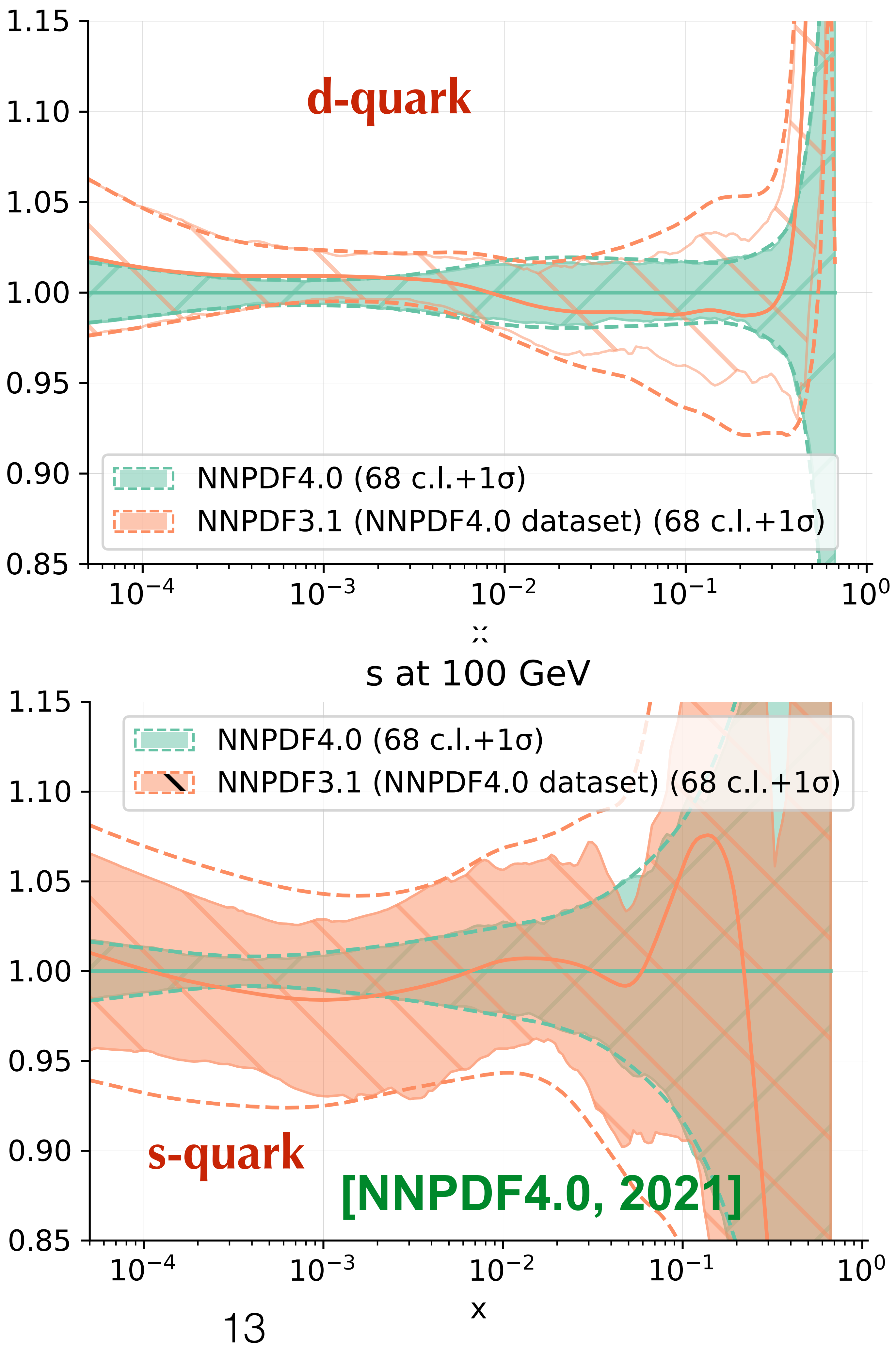
Methodology and uncertainties

- Textbook criterion "Δχ²=1" on estimation of uncertainties is not reliable in global fit, involving large data samples and degrees of freedoms; PDF unc. depends very much on methodologies including "tolerance"

PDFs from reduced fits



NNPDF methodology update



- CT uses tier1+tier2 tolerance, MSHT uses a pure dynamic tolerance, both close to a hypothesis test criterion
- NNPDF3.1 uses ML algorithm with effective tolerance that is smaller than CT and MSHT as checked explicitly from reduced fits
- substantial changes on methodologies for NN4.0 vs. NN3.1 further affect the uncertainty

Summary on PDFs

- Global analyses of parton distributions demonstrate great success of QCD and on understanding internal structures of proton, and phenomenologically become more and more prominent for electroweak precision test and searches for new physics at the (HL-)LHC

- LHC delivers plenty of PDF sensitive data with high statistics and with theory evaluated almost all at NNLO; some of the N3LO calculations are already available; however, an advance on the treatment of the LHC experimental systematics and methodologies of PDF determinations can be crucial

- With the global efforts from many groups, we are gradually approaching PDFs precision of a few percents; while LHC-independent inputs on PDFs, for instance from future DIS experiments or lattice QCD simulation with improved precisions will be highly valuable

[Snowmass 2021 PDF white paper, 2203.13923]

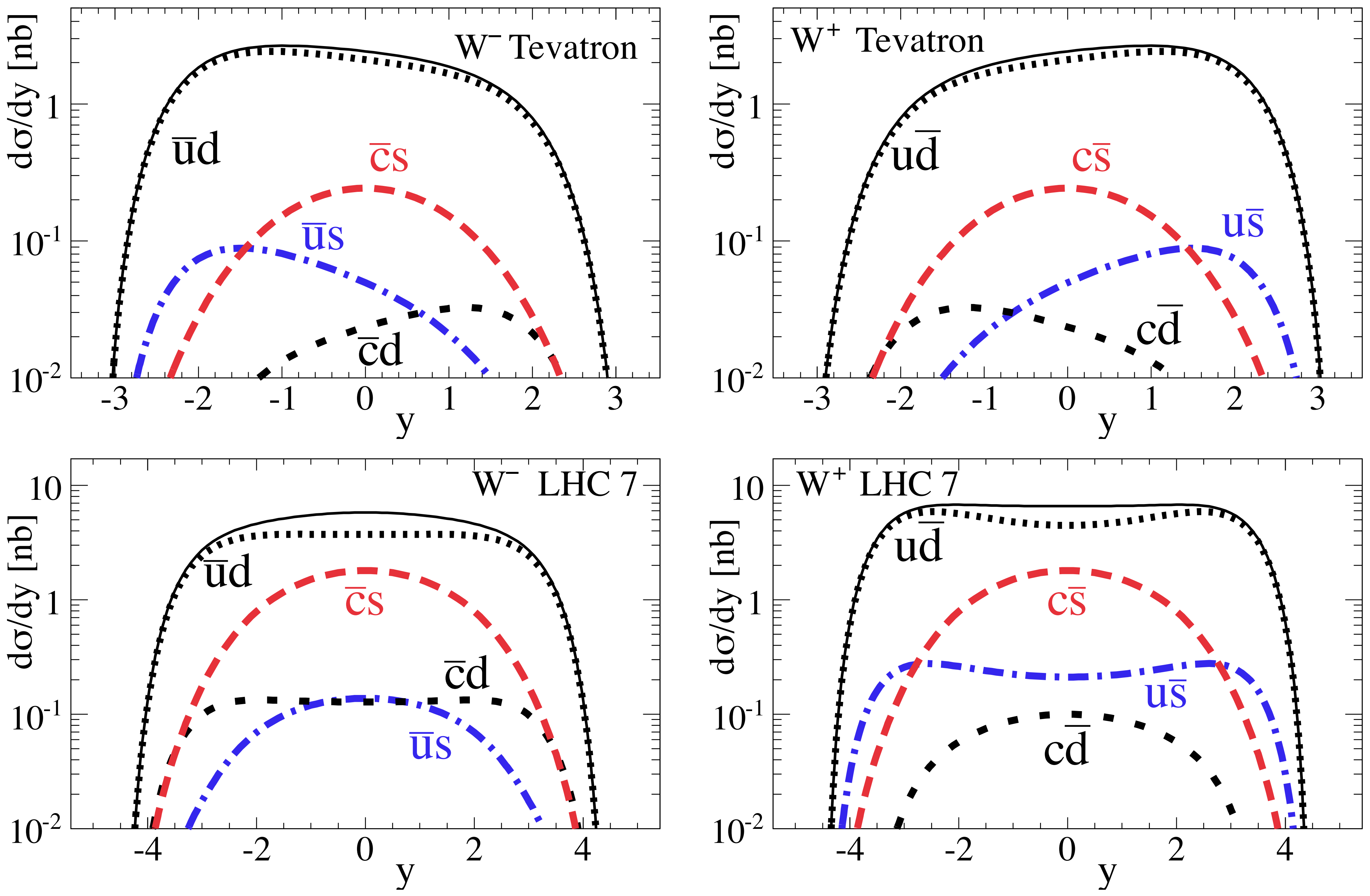
NEW TASKS in the HL-LHC ERA:

Obtain complete N2LO and N3LO predictions for PDF-sensitive processes	Improve models for correlated systematic errors	Find ways to constrain large-x PDFs without relying on nuclear targets
Develop and benchmark fast N2LO interfaces	Estimate N2LO theory uncertainties	New methods to combine PDF ensembles, estimate PDF uncertainties, deliver PDFs for applications

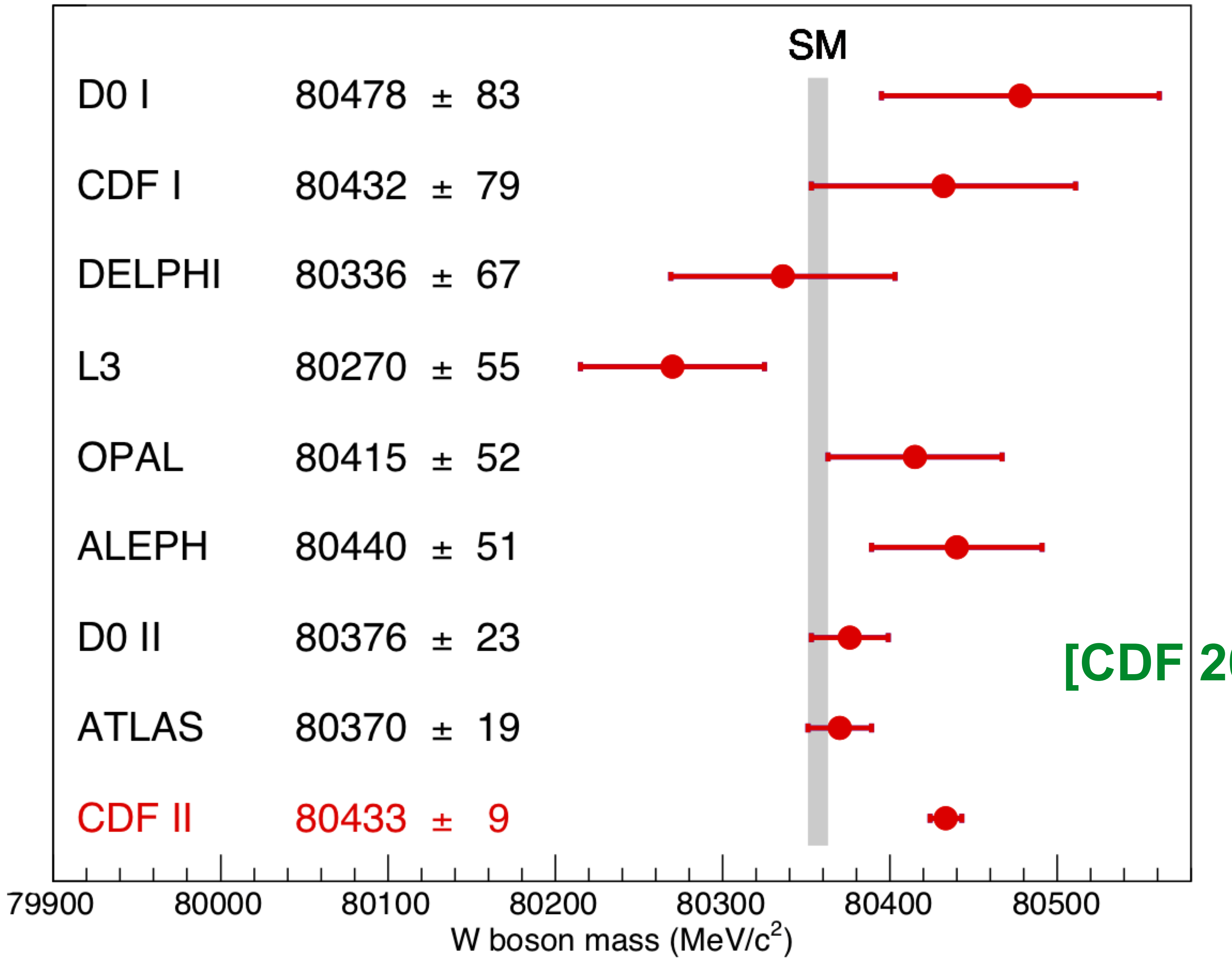
W boson mass measurement

- PDFs are key inputs for precision programs at hadron colliders, e.g., precision electroweak measurements, searches for new physics beyond the SM, especially non-resonance signatures hiding in high mass tails

W boson rapidity distribution **[1203.1290]**



W boson mass from different experiments



[CDF 2022]

Fig. 5

SM expectation: $M_W = 80{,}357 \pm 4_{\text{inputs}} \pm 4_{\text{theory}}$ (PDG 2020)

LHCb measurement : $M_W = 80{,}354 \pm 23_{\text{stat}} \pm 10_{\text{exp}} \pm 17_{\text{theory}} \pm 9_{\text{PDF}}$ *[JHEP* **2022,** 36 (2022)]

PDF unc. of CDF / ATLAS / LHCb: 3.9 / 8 / 9 MeV

W boson mass measurement

- PDFs are key inputs for precision programs at hadron colliders, e.g., precision electroweak measurements, searches for new physics beyond the SM, especially non-resonance signatures hiding in high mass tails

PDF unc. at LHCb, NNPDF3.1, CT18, MSHT20

$$m_W = 80362 \pm 23_{\rm stat} \pm 10_{\rm exp} \pm 17_{\rm theory} \pm 9_{\rm PDF}\ \text{MeV},$$
$$m_W = 80350 \pm 23_{\rm stat} \pm 10_{\rm exp} \pm 17_{\rm theory} \pm 12_{\rm PDF}\ \text{MeV},$$
$$m_W = 80351 \pm 23_{\rm stat} \pm 10_{\rm exp} \pm 17_{\rm theory} \pm 7_{\rm PDF}\ \text{MeV},$$

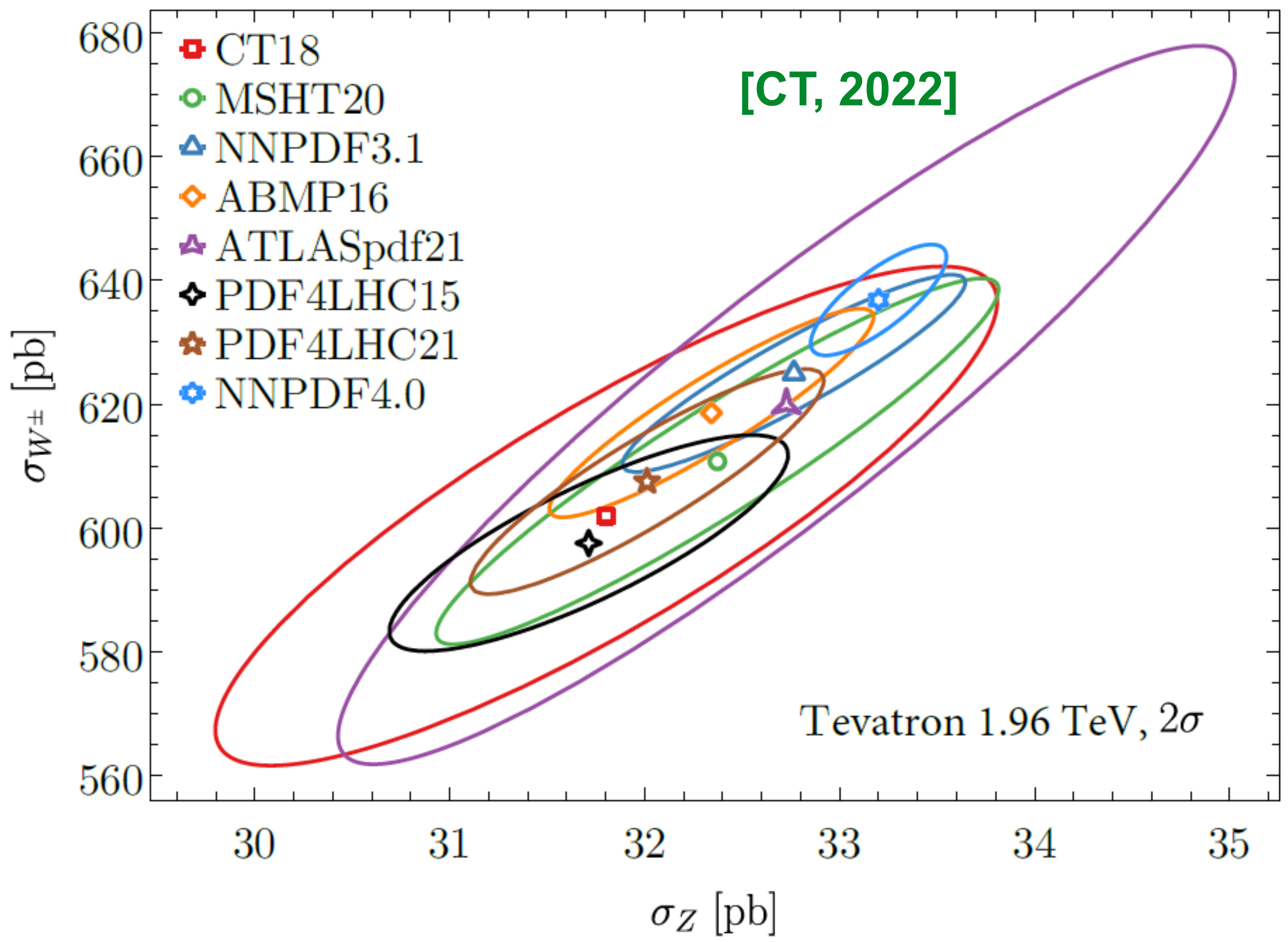
ATLAS, CT10 + 3.8 MeV (MMHT14-CT14)

W-boson charge	W^+		W^-		Combined	
Kinematic distribution	p_T^ℓ	m_T	p_T^ℓ	m_T	p_T^ℓ	m_T
δm_W [MeV]						
Fixed-order PDF uncertainty	13.1	14.9	12.0	14.2	8.0	8.7
AZ tune	3.0	3.4	3.0	3.4	3.0	3.4
Charm-quark mass	1.2	1.5	1.2	1.5	1.2	1.5

CDF, NNPDF3.1 only (3.9 MeV)

(other tested, CT18, MMHT14, +-2.1 MeV)

W/Z fiducial cross sections at Tevatron (95% C.L.)



spread of predictions from different PDFs could be much larger than the PDF unc. of a specific set
even for the same group the PDF unc. not necessarily decrease with time

Analyzing of W mass data with most UP-TO-DATE PDFs will be highly desirable

A phenomenological study

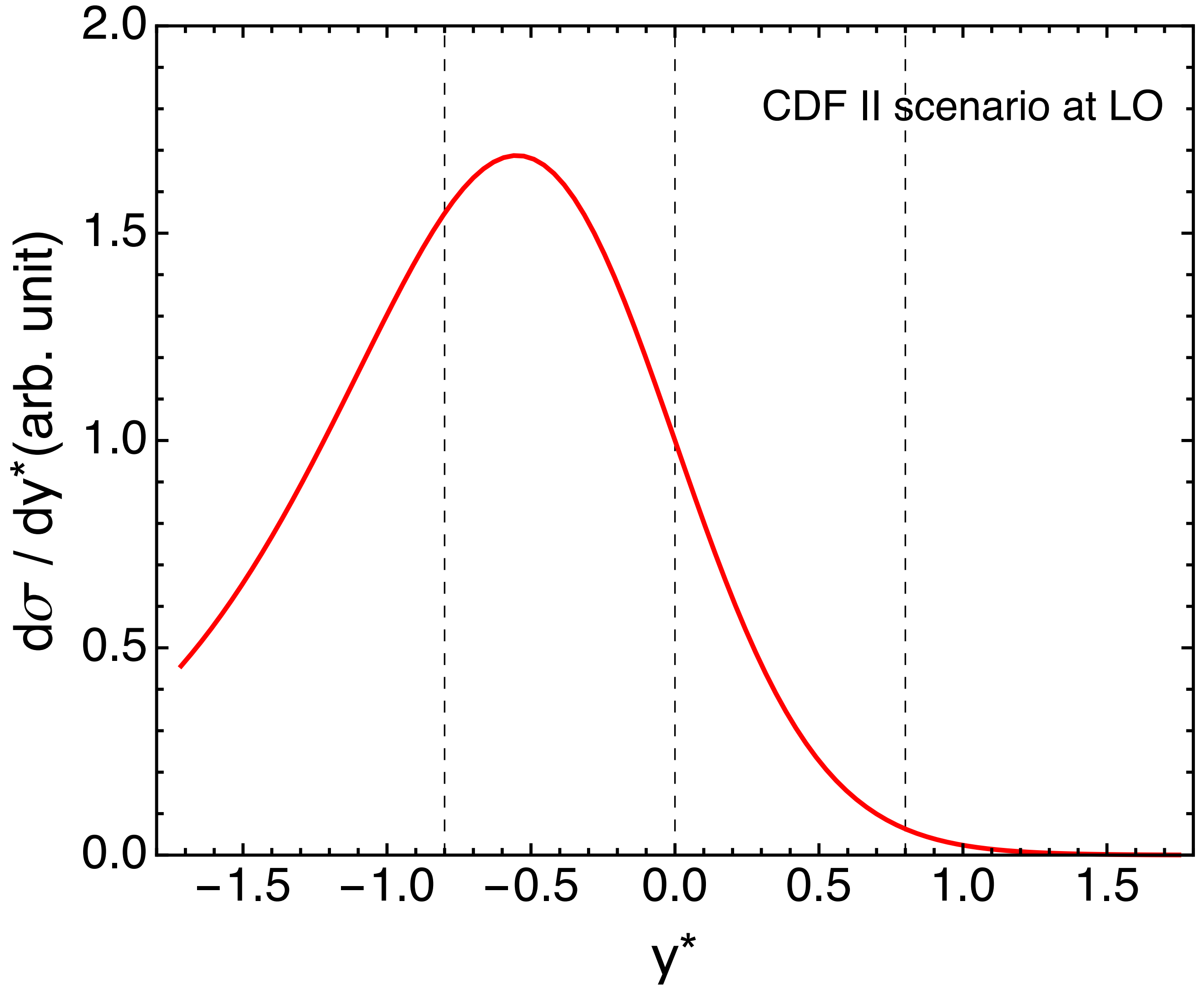
[in collaboration with K. Xie and D. Liu to appear soon; PDF unc. also studied in 2205.02788 by C.-P. Yuan+]

- Kinematics of the decayed leptons are usually factorized into several components in experimental analyses that are modeled separately; PDFs alter several of them especially rapidity and p_T of W boson

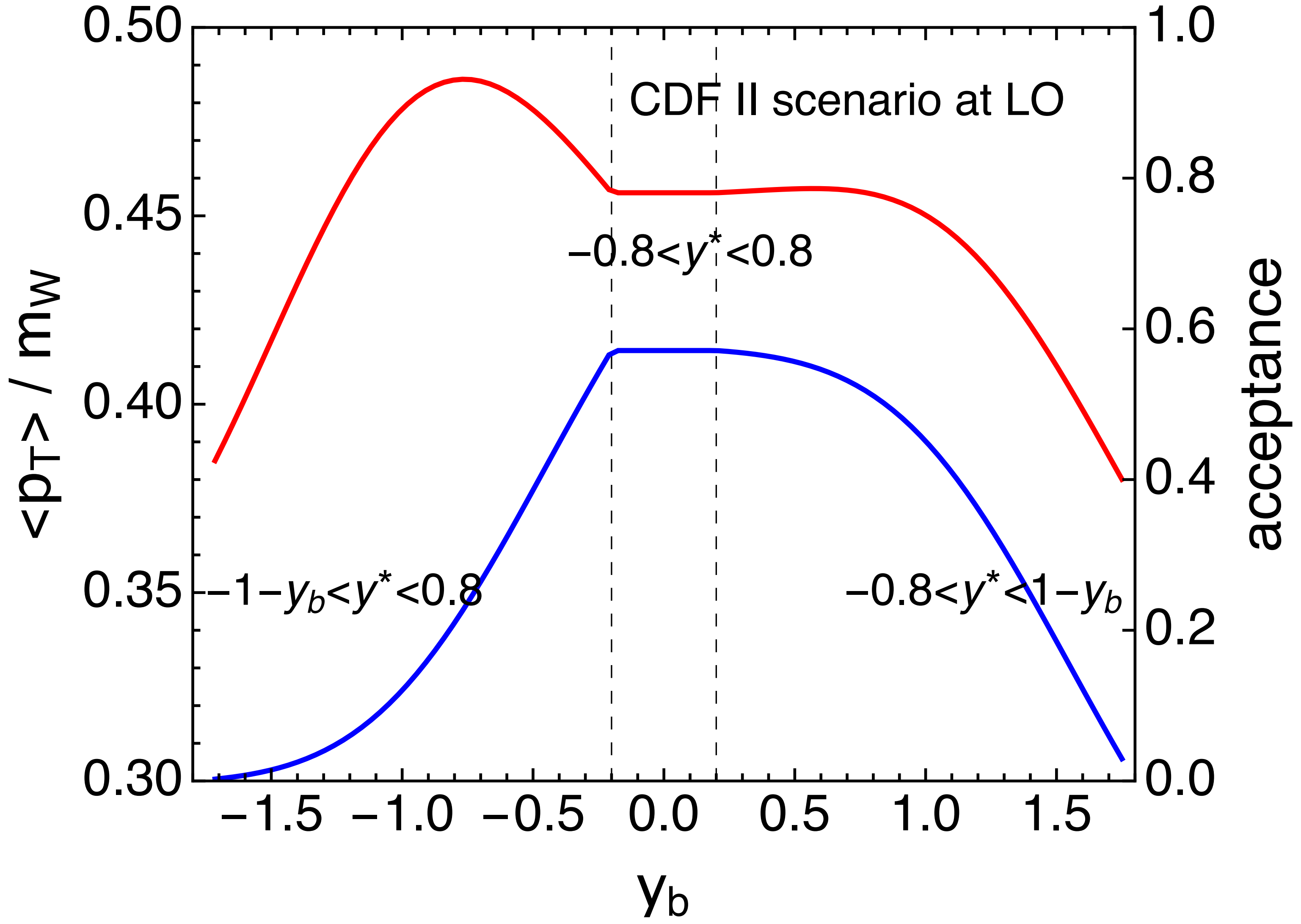
$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}p_1\mathrm{d}p_2} = \left[\frac{\mathrm{d}\sigma(m)}{\mathrm{d}m}\right]\left[\frac{\mathrm{d}\sigma(y)}{\mathrm{d}y}\right]\left[\frac{\mathrm{d}^2\sigma(p_T,y)}{\mathrm{d}p_T\mathrm{d}y}\left(\frac{\mathrm{d}\sigma(y)}{\mathrm{d}y}\right)^{-1}\right] \times \left[(1+\cos^2\theta) + \sum_{i=0}^{7} A_i(p_T,y)P_i(\cos\theta,\phi)\right],$$

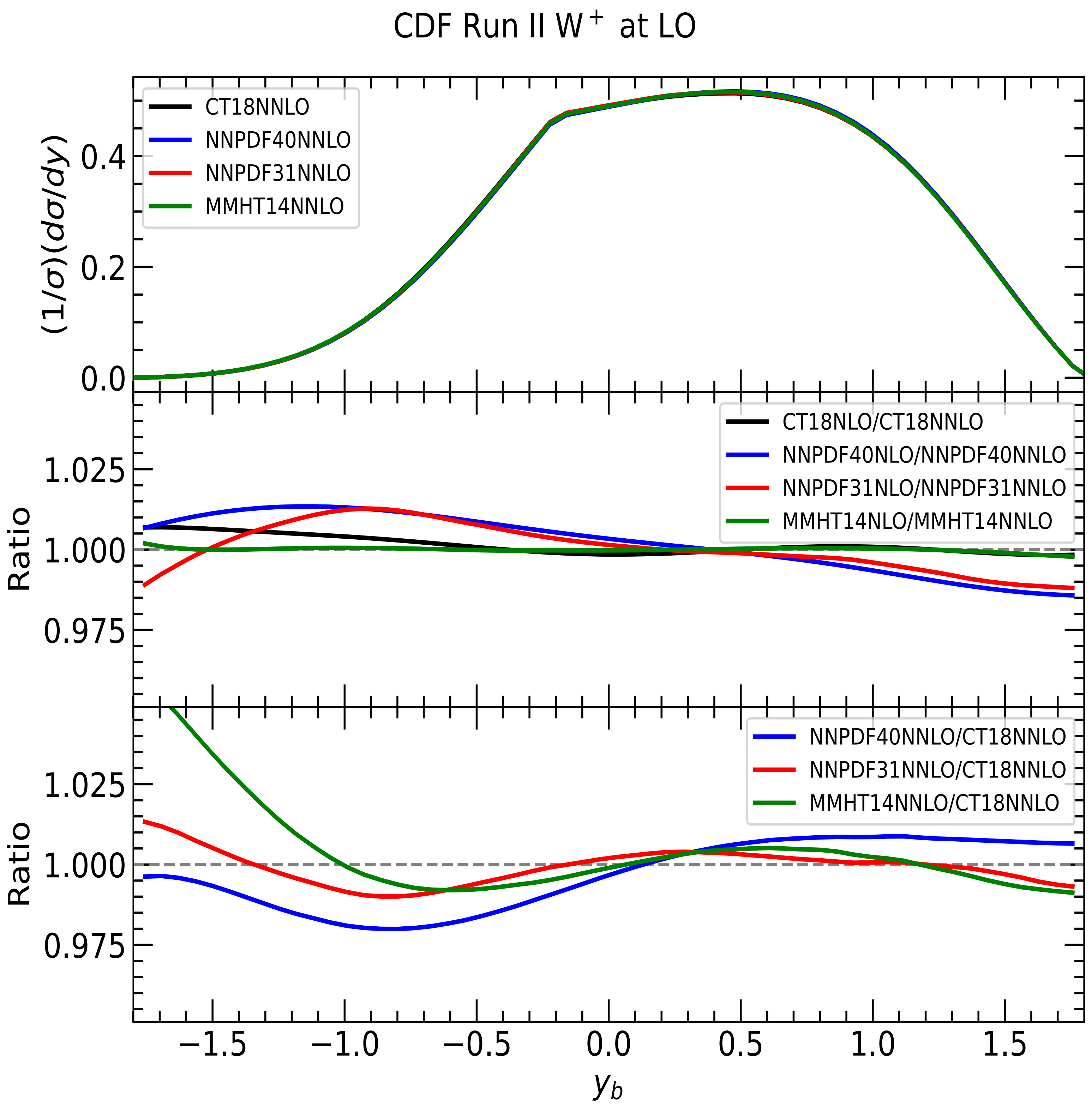
- PDF unc. on invariant mass or angular coefficients are small; on rapidity is dominant
- in principle only theoretical unc. on ratio $p_T(W)/p_T(Z)$ should be considered due to data-driven method in exp. analyses

positron rapidity dis. in COM frame



mean p_T (red) or acceptance (blue) of positron vs. y_W



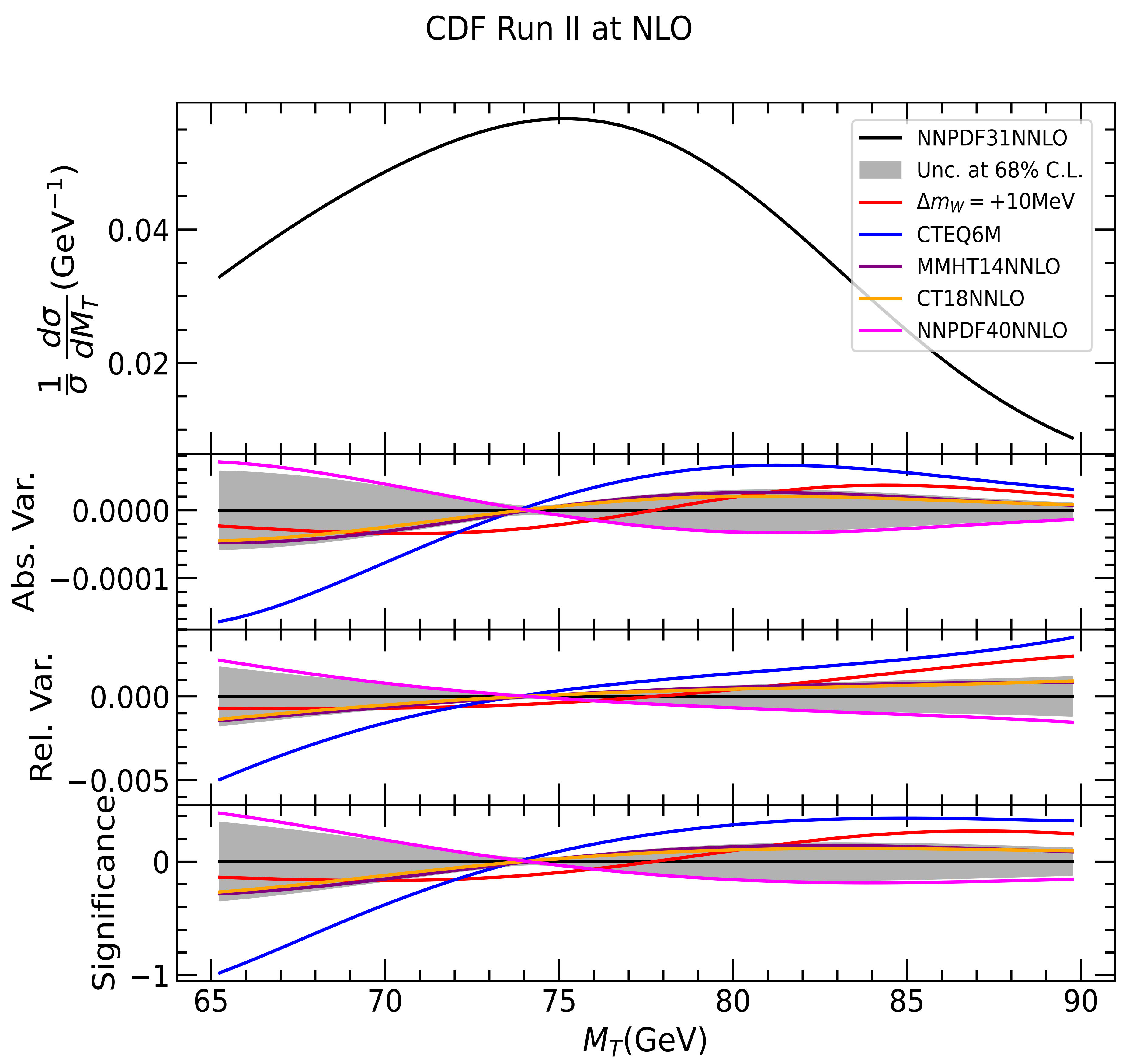


- CDF imposes $p_{T,l}$>30 GeV and $|y_l|$<1; mean p_T depends on rapidity of the W boson; while distribution of y_W is sensitive to PDFs

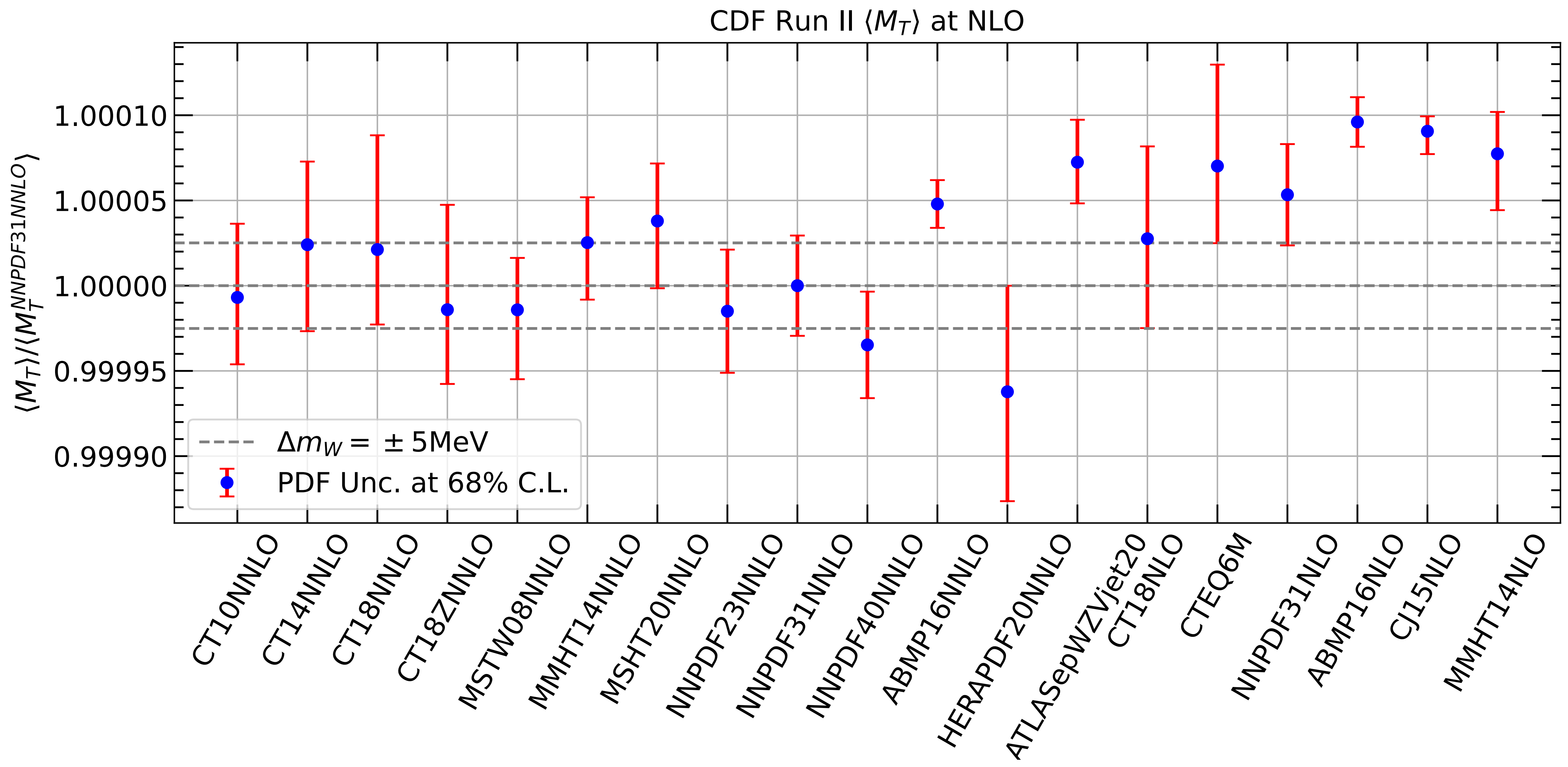
Transverse mass at CDF

- We estimate shift of extracted W boson mass induced by variation of PDFs, and the associated PDF uncertainty for a variety of PDFs, focusing on the kinematic variable of transverse mass at CDF

normalized m_T distribution
PDF var. vs. M_W var.



mean value of m_T



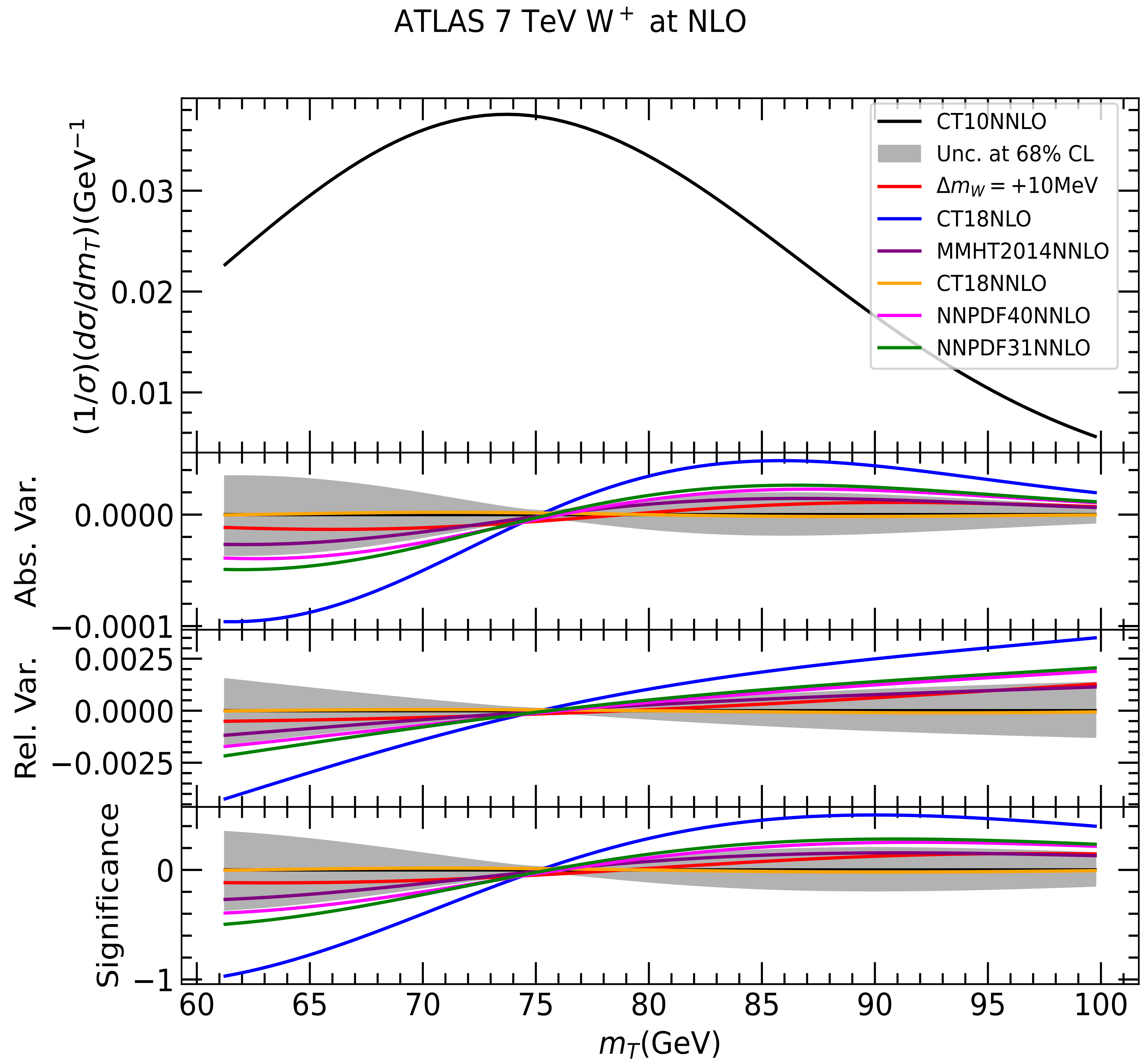
estimate shift and PDF unc. of W mass

δM_W in MeV	sta.	NNPDF3.1	CT18	MMHT2014	NNPDF4.0	MSHT2020
$\langle M_T \rangle$(LO)	–	$0^{+8.3}_{-8.3}$	$-1.0^{+8.3}_{-11.4}$	$-3.3^{+7.4}_{-4.2}$	$+7.8^{+5.1}_{-5.1}$	$-3.1^{+6.7}_{-5.7}$
χ^2 fit (LO)	8.0	$0^{+7.6}_{-7.6}$	$-1.0^{+5.4}_{-8.6}$	$-3.3^{+6.1}_{-3.0}$	$+8.0^{+3.7}_{-3.7}$	$-3.0^{+5.0}_{-4.0}$
$\langle M_T \rangle$(NLO)	–	$0^{+5.9}_{-5.9}$	$-4.2^{+8.8}_{-13.3}$	$-5.0^{+6.7}_{-5.3}$	$+6.9^{+6.2}_{-6.2}$	$-7.6^{+7.9}_{-6.7}$
χ^2 fit (NLO)	8.0	$0^{+4.2}_{-4.2}$	$-4.3^{+5.4}_{-10.1}$	$-5.1^{+4.8}_{-3.4}$	$+7.1^{+4.5}_{-4.5}$	$-7.8^{+5.7}_{-4.5}$
CDF	9.2	$0^{+3.9}_{-3.9}$	–	–	–	–

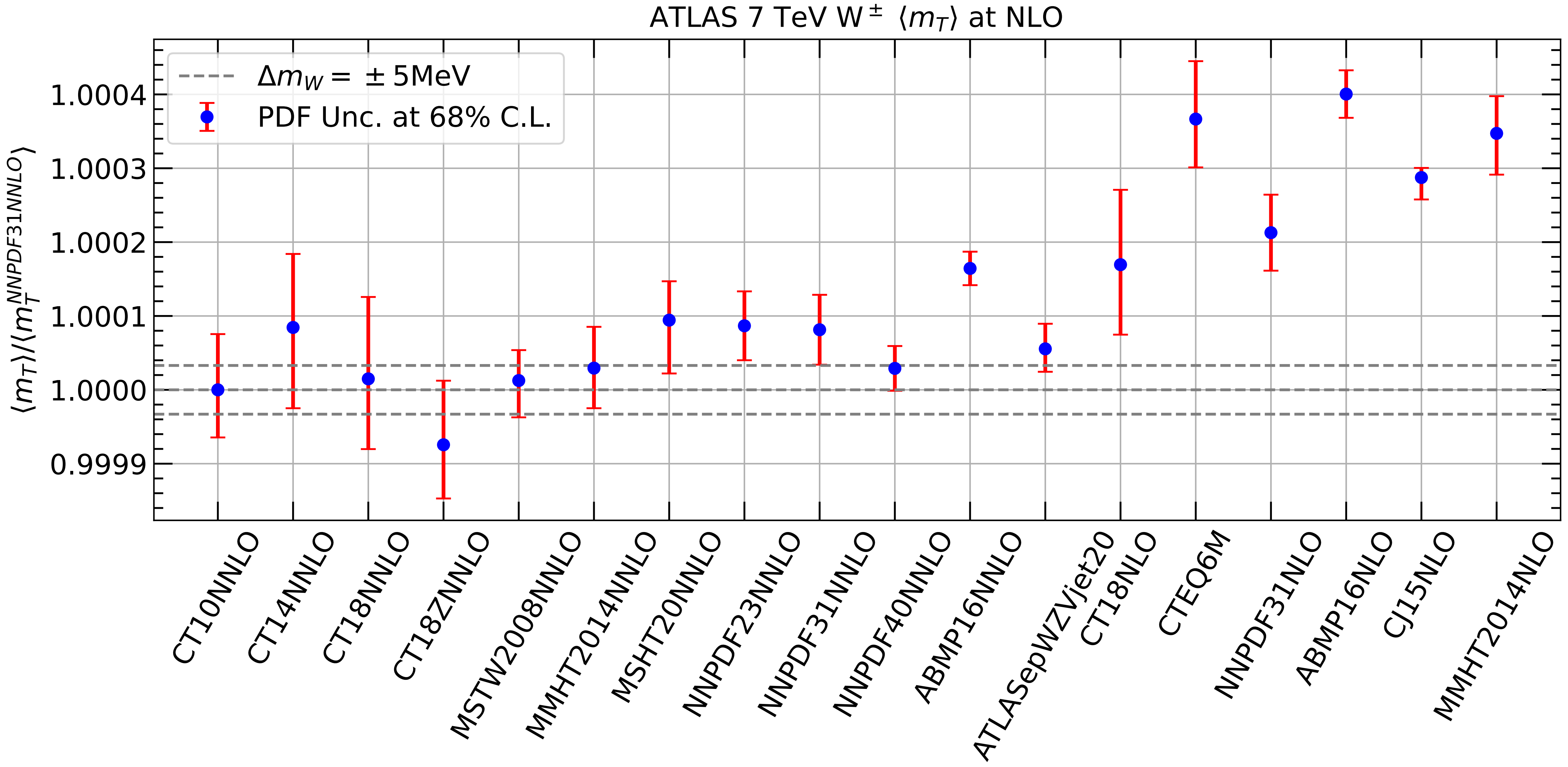
Transverse mass at ATLAS

- We estimate shift of extracted W boson mass induced by variation of PDFs, and the associated PDF uncertainty for a variety of PDFs, focusing on the kinematic variable of transverse mass at ATLAS

normalized m_T distribution
PDF var. vs. M_W var.



mean value of m_T



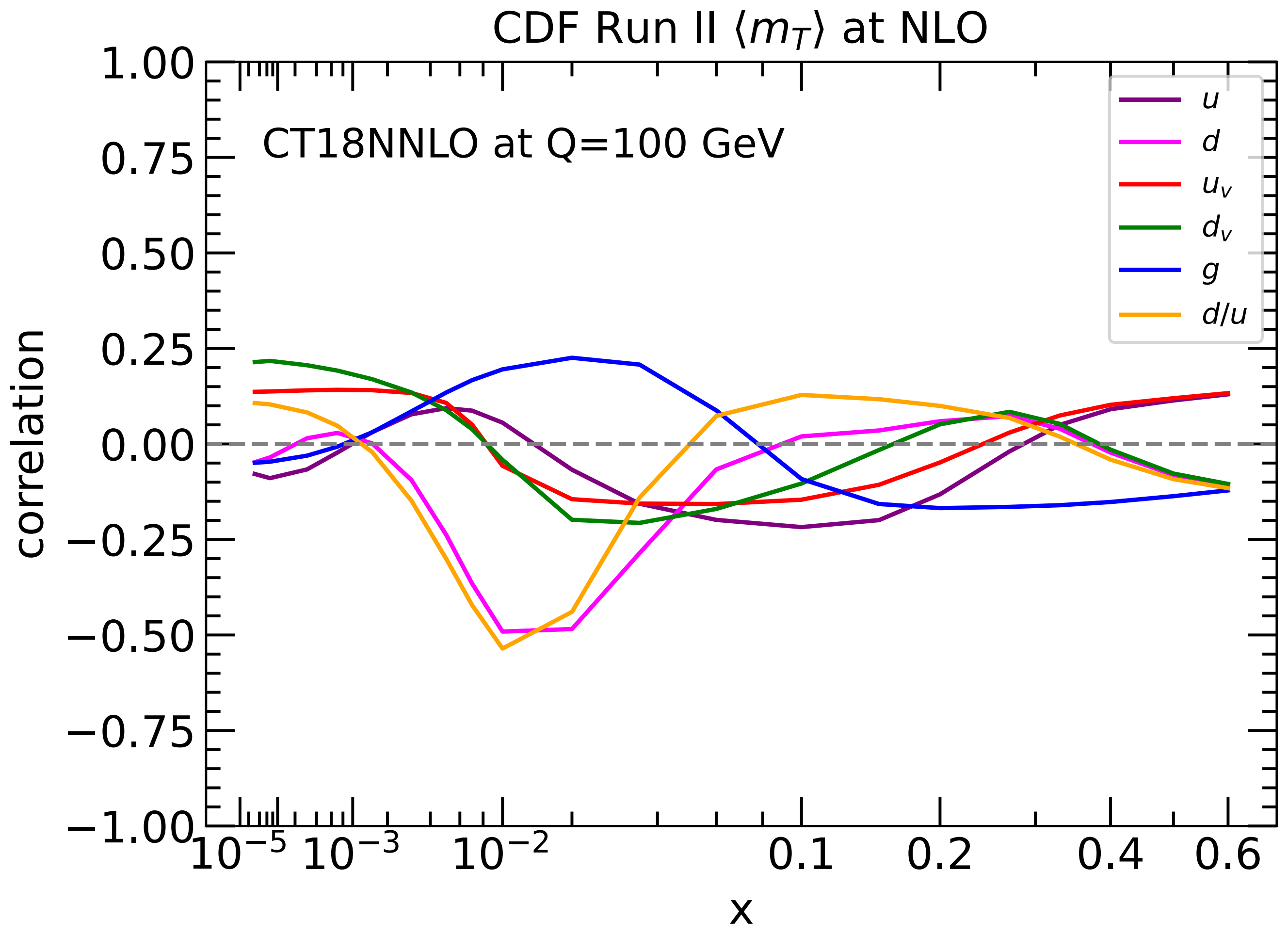
δM_W in MeV	CT10	CT18	MMHT14	NNPDF4.0	CT14	MSHT20
$W^+\ \langle M_T\rangle$ (NLO)	$0^{+12.1}_{-12.9}$	$+1.4^{+21.8}_{-20.0}$	$-10.3^{+11.6}_{-11.1}$	$-17.1^{+7.4}_{-7.4}$	$-16.2^{+23.5}_{-19.1}$	$-24.8^{+16.8}_{-11.9}$
$W^-\ \langle M_T\rangle$ (NLO)	$0^{+13.5}_{-15.2}$	$-5.7^{+14.0}_{-19.5}$	$+1.1^{+8.6}_{-10.3}$	$+7.5^{+4.9}_{-4.9}$	$-9.6^{+12.8}_{-15.3}$	$-4.5^{+8.3}_{-7.5}$
$W^\pm\ \langle M_T\rangle$ (NLO)	$0^{+9.8}_{-11.4}$	$-2.3^{+14.4}_{-16.8}$	$-4.5^{+8.2}_{-8.5}$	$-4.4^{+4.6}_{-4.6}$	$-12.8^{+16.6}_{-15.1}$	$-14.3^{+10.9}_{-8.0}$
$W^+\ \langle M_T\rangle$ (LO)	$0^{+10.8}_{-11.4}$	$-6.5^{+14.1}_{-10.0}$	$-5.7^{+8.1}_{-7.1}$	$-14.1^{+5.8}_{-5.8}$	$-4.1^{+15.0}_{-12.9}$	$-14.4^{+10.2}_{-7.3}$
$W^-\ \langle M_T\rangle$ (LO)	$0^{+8.9}_{-11.4}$	$-7.2^{+10.1}_{-12.5}$	$+3.1^{+8.3}_{-9.9}$	$+3.5^{+4.5}_{-4.5}$	$-7.0^{+6.2}_{-8.9}$	$+2.1^{+6.3}_{-4.9}$
$W^\pm\ \langle M_T\rangle$ (LO)	$0^{+5.2}_{-7.0}$	$-0.6^{+7.6}_{-7.4}$	$-1.2^{+5.3}_{-5.9}$	$-5.0^{+3.0}_{-3.0}$	$-5.6^{+8.0}_{-8.4}$	$-5.9^{+5.9}_{-4.2}$
W^+ ATLAS	$0^{+14.9}_{-14.9}$	–	–	–	–	–
W^- ATLAS	$0^{+14.2}_{-14.2}$	–	–	–	–	–
$W^\pm$ ATLAS	$0^{+7.4}_{-7.4}$	–	–	–	–	–

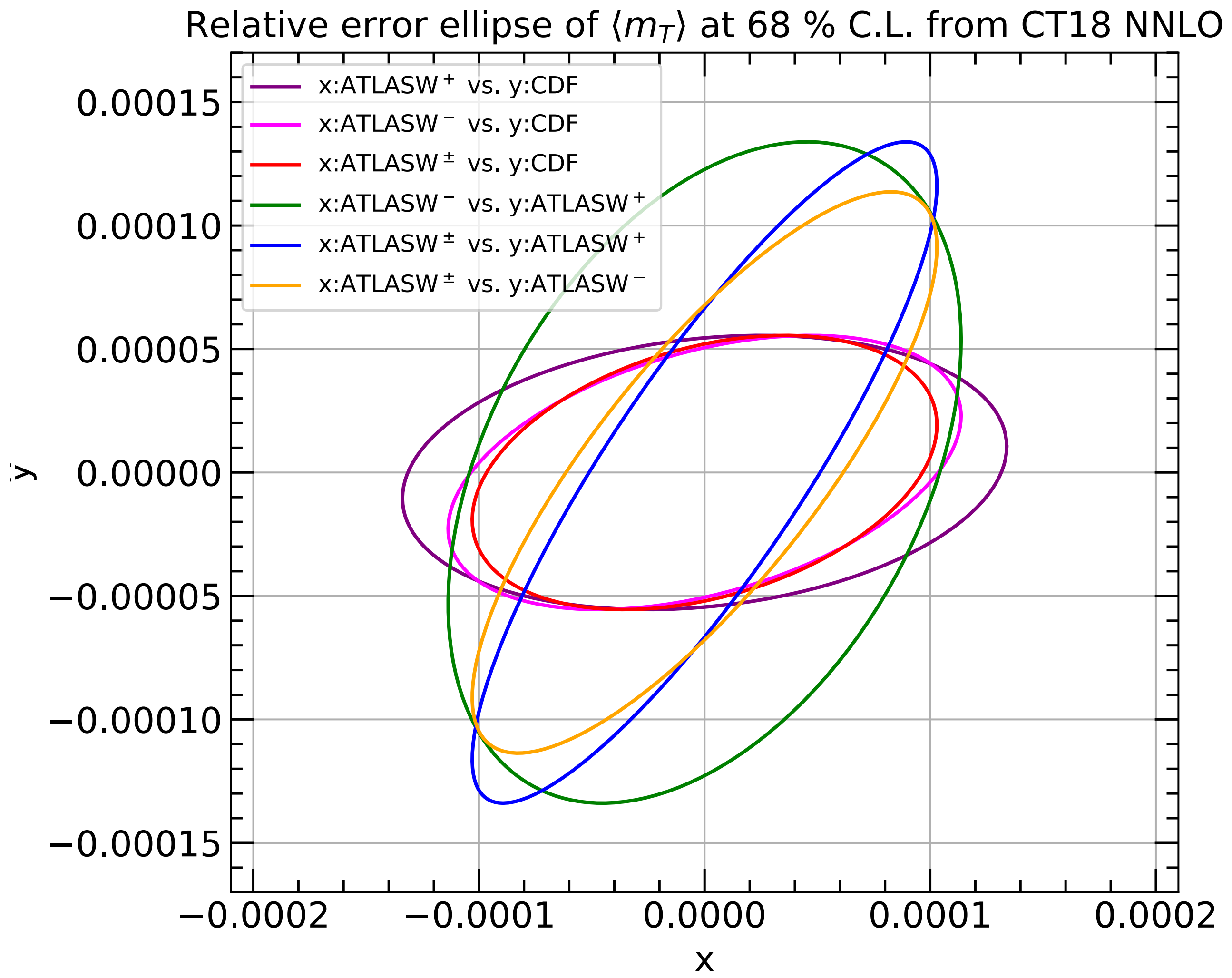
estimate shift and PDF unc. of W mass

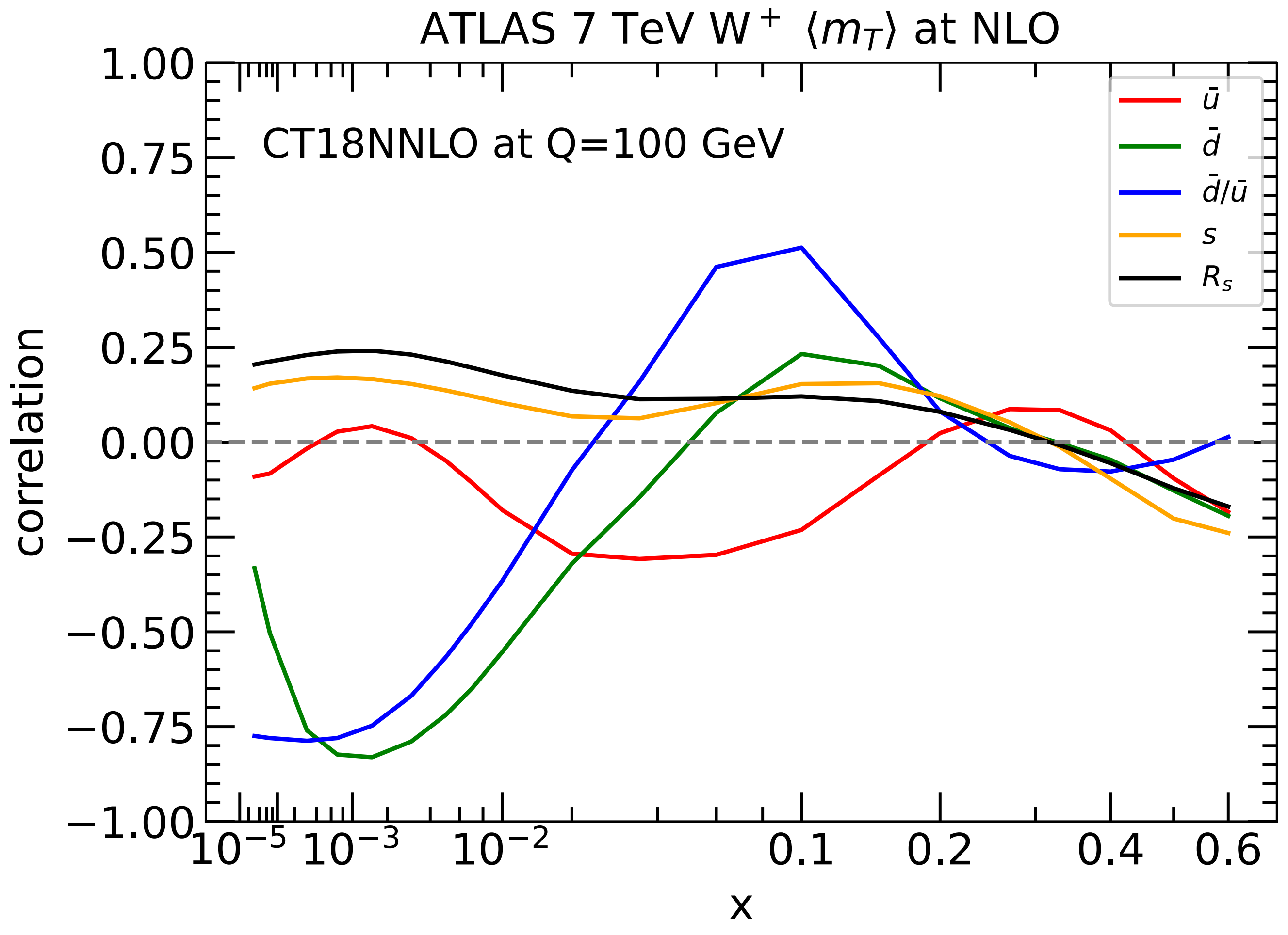
Constraints in CT18

- We carry out a series of Lagrange multiplier scans to identify the constraints on the transverse mass distribution (using mean M_T) imposed by individual data sets in the CT18 global analysis

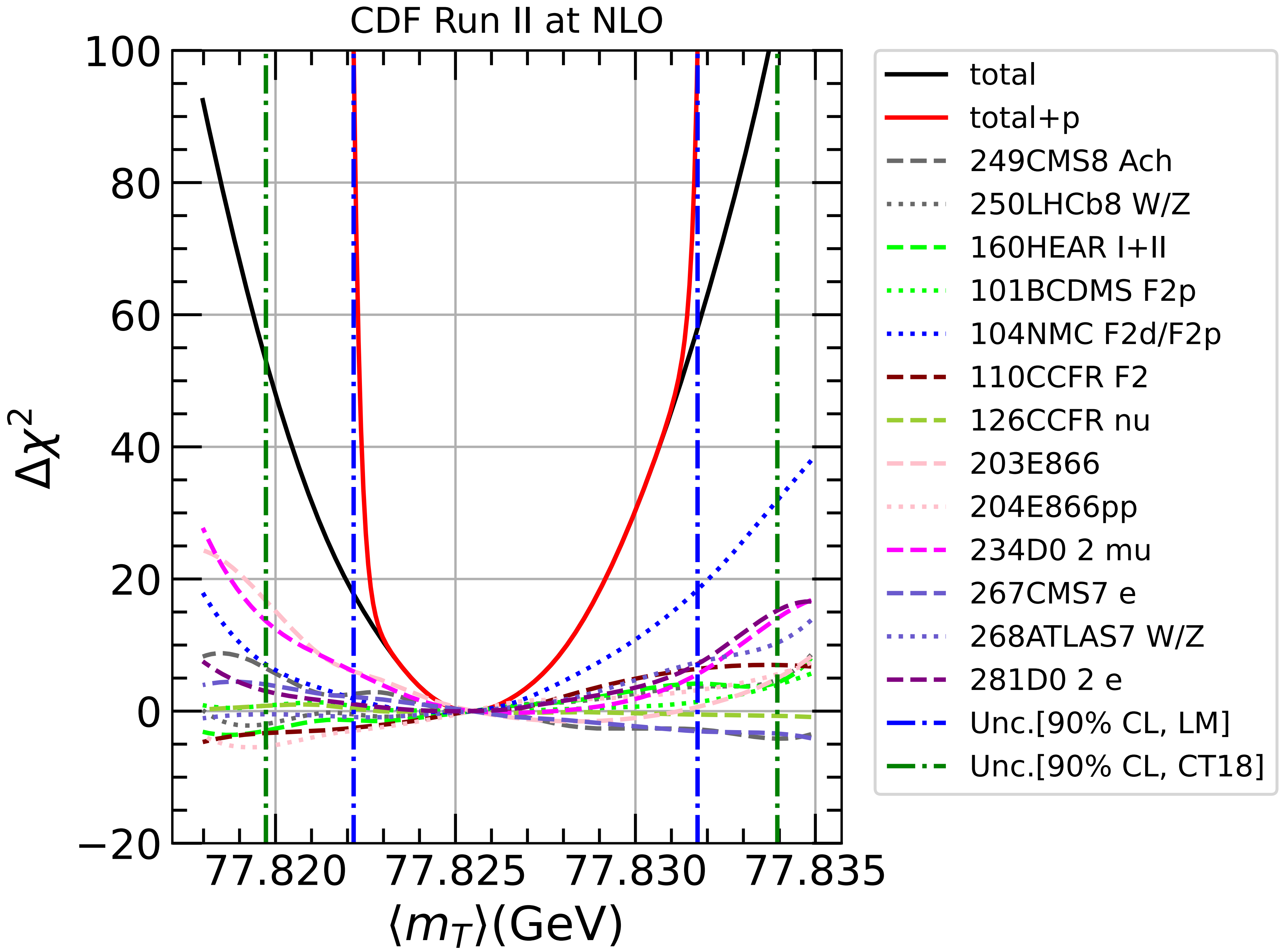
PDF induced correlations







constraints in CT18



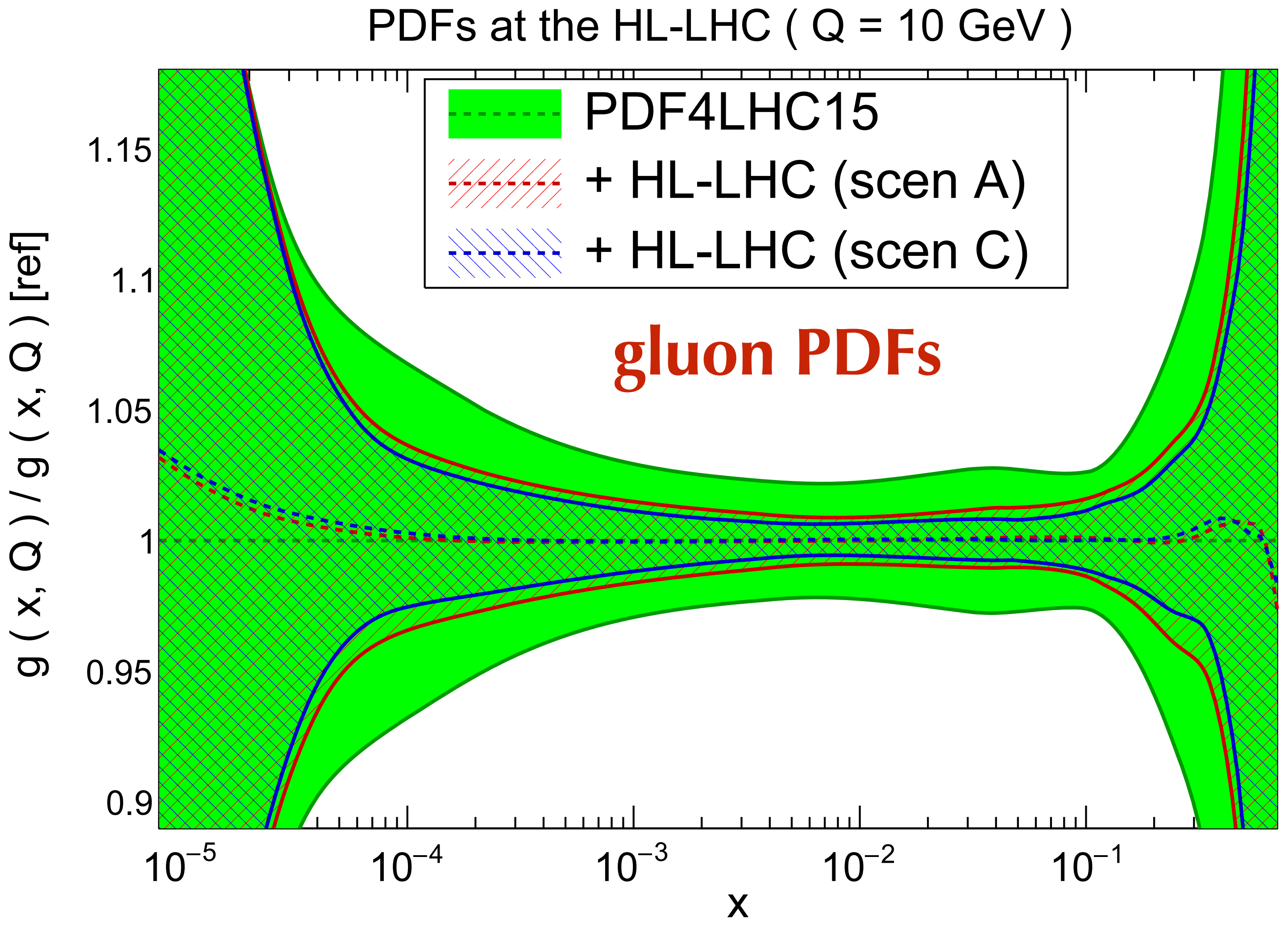
- m_T at CDF (ATLAS) is mostly sensitive to the d-quark (dbar-quark) at x~0.01(0.001); CDF and ATLAS are uncorrelated

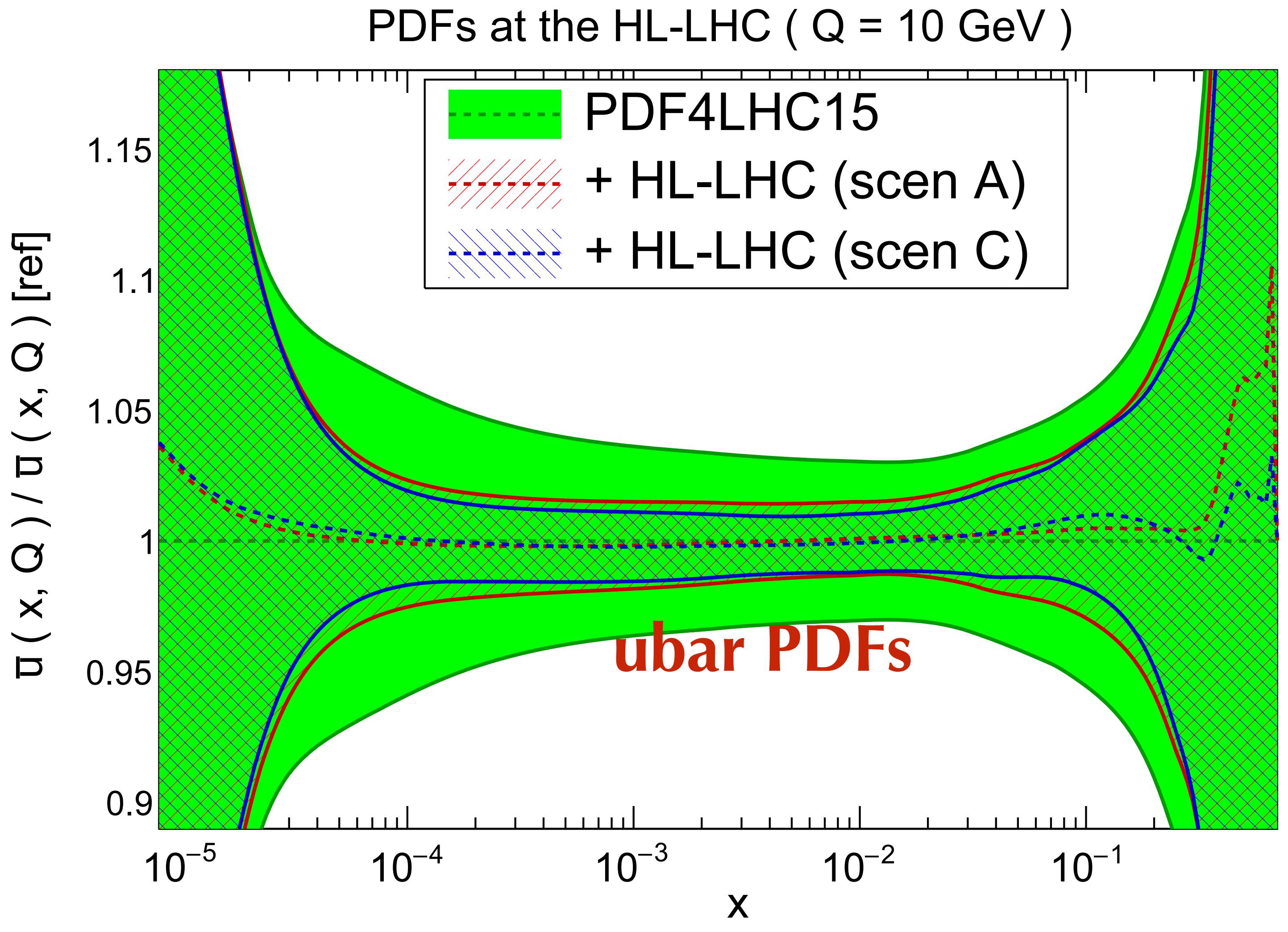
- m_T at CDF is largely constrained by the DIS and Drell-Yan data on deuteron target, the Tevatron lepton charge asymmetry data; at ATLAS also the CMS charge asymmetry data

Future prospect on PDF unc.

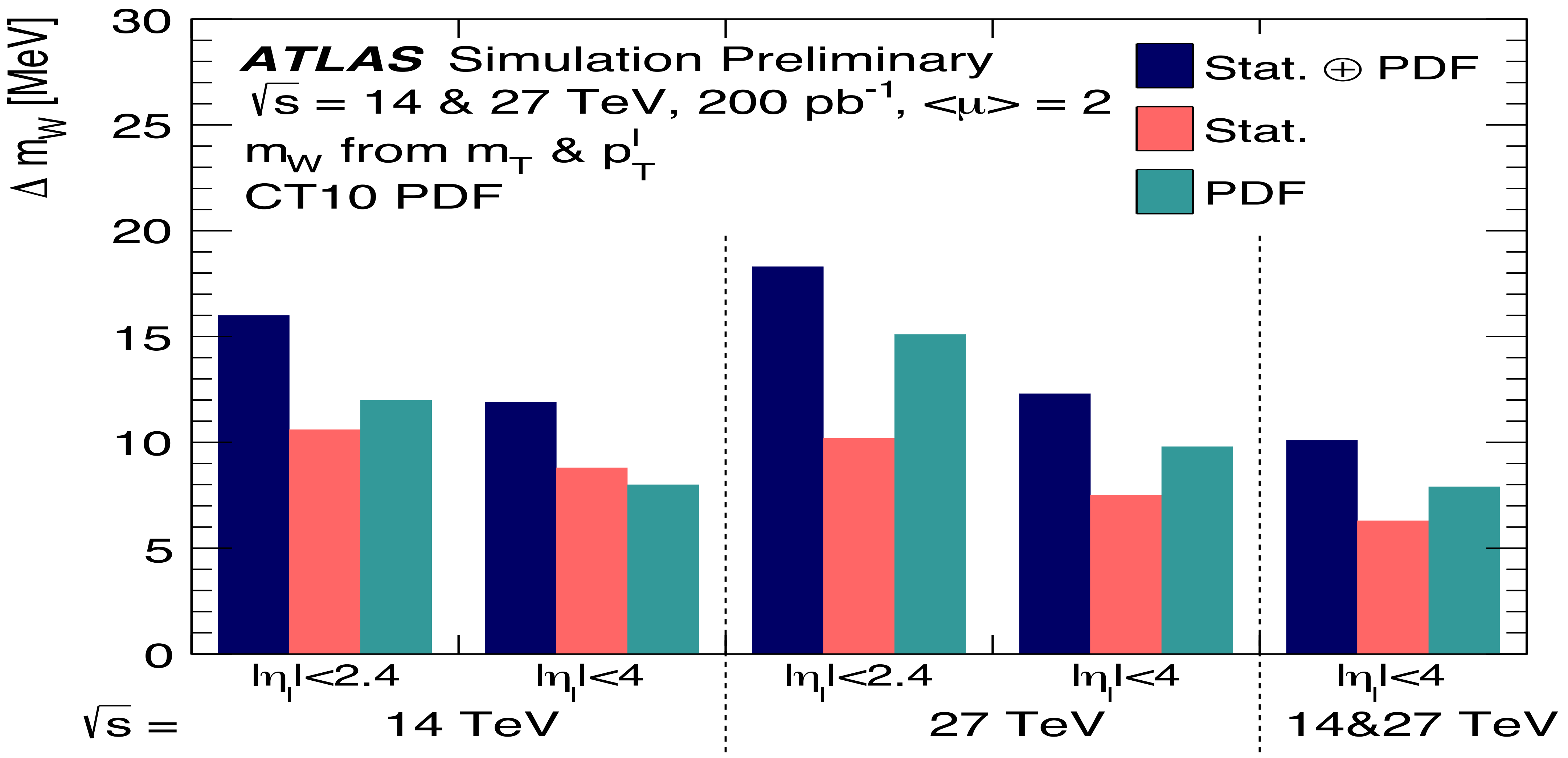
- Precision on PDFs can be further improved with upcoming data from EIC(c), HL-LHC, or ultimately if LHeC is possible; projections on M_W have been made with PDFs fitted to pseudo-data

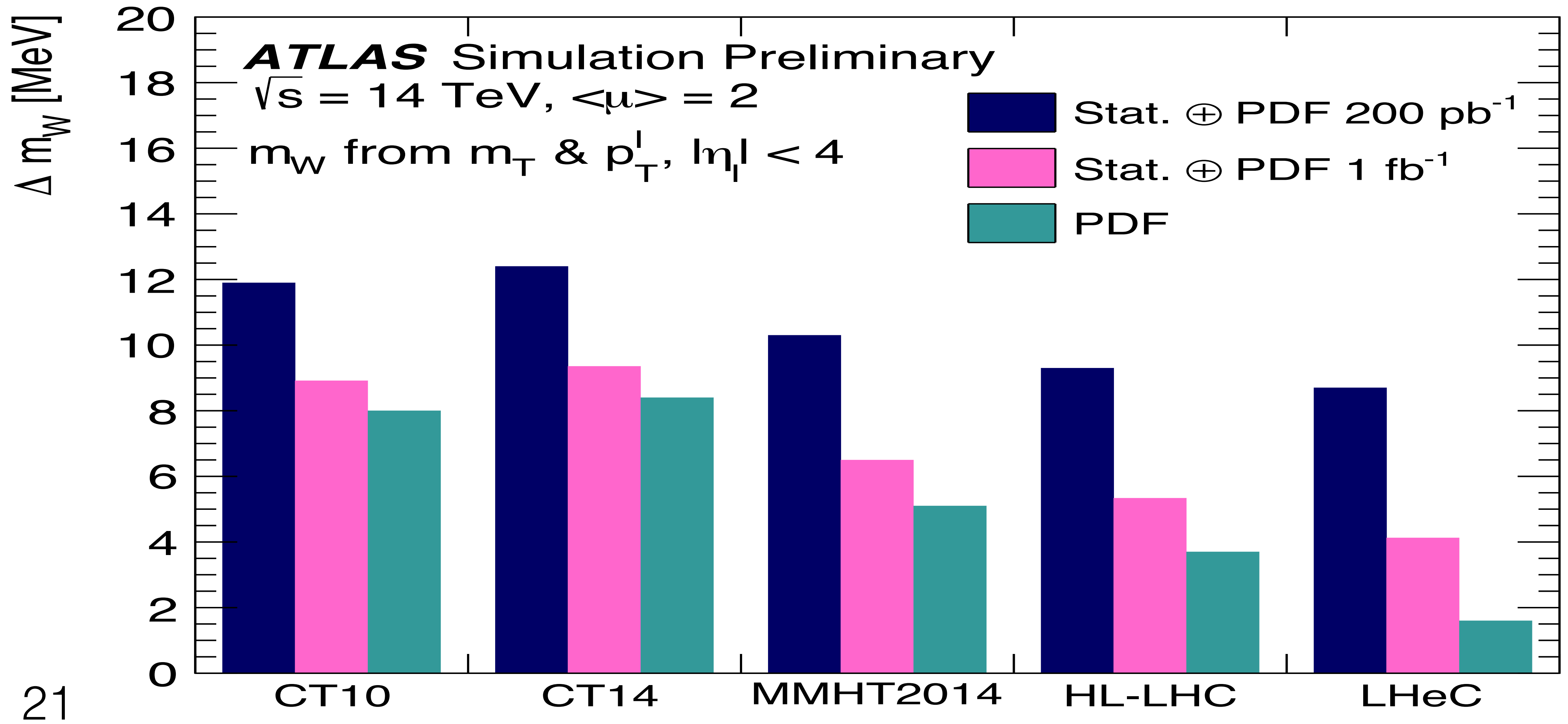
[JG+, 2018]





[SM Report, 1902.04070]

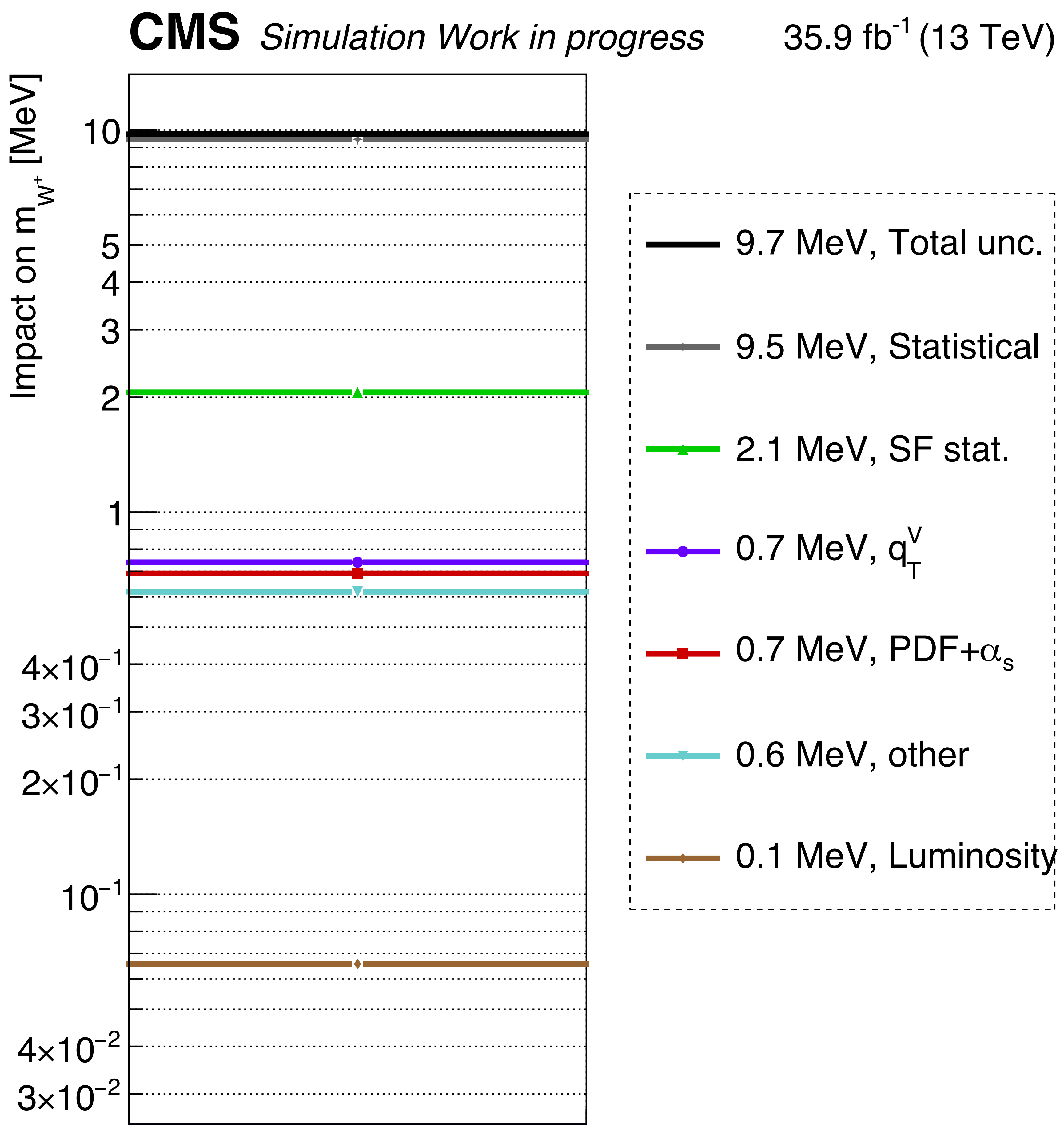


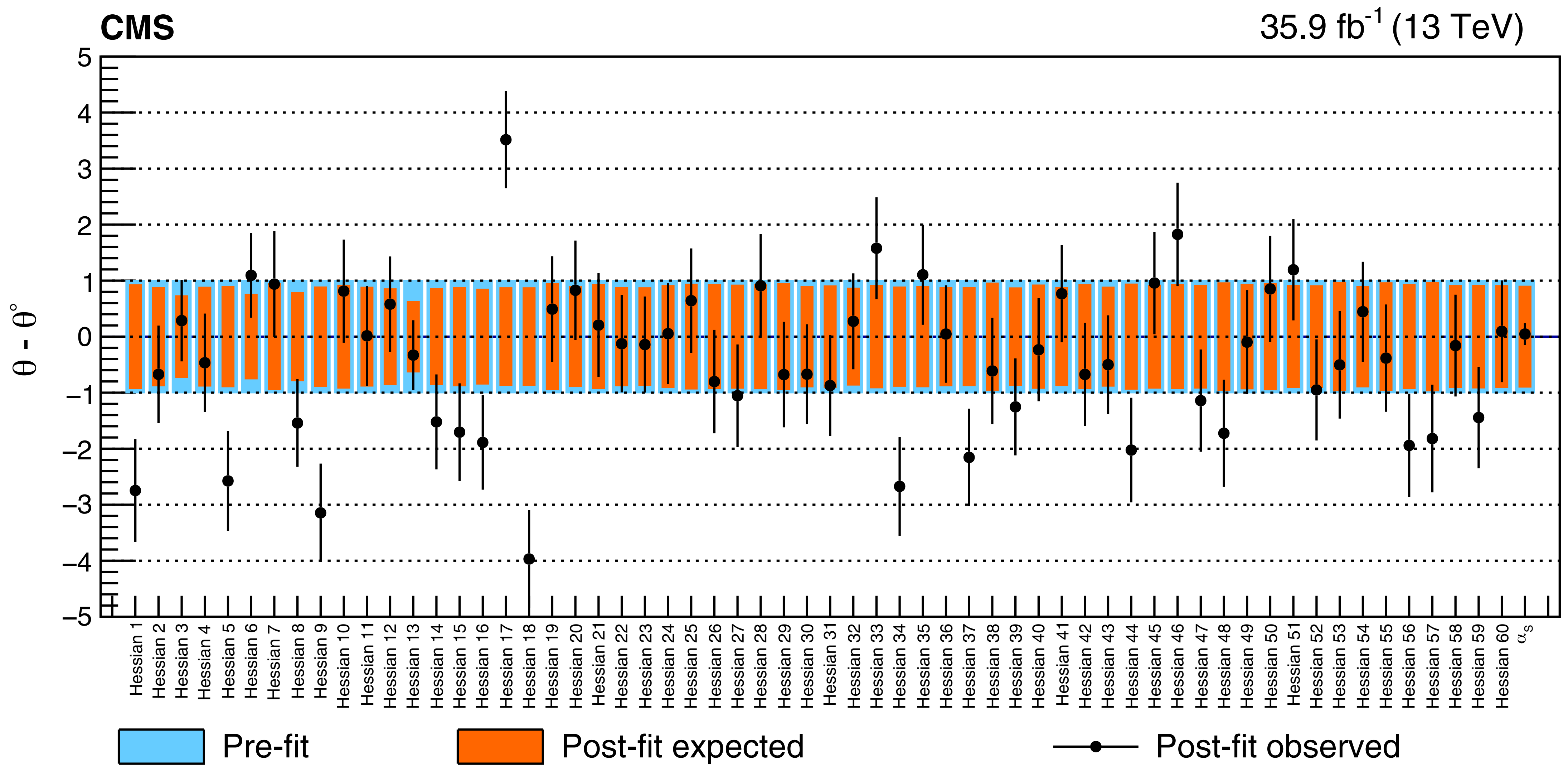


Future prospect on PDF unc.

- There are ongoing efforts from both CMS and ATLAS of using the same kinematic distributions for W mass measurement to profile or to constrain the PDFs, namely a spontaneous fit of M_W and PDFs

[CERN-THESIS-2021-100; Phys. Rev. D 102, 092012]





The template fit to m_W, fixing all the POIs and the PDF nuisance parameters to their pre-fit values and uncertainties predicts a systematic uncertainty due to the PDF on m_W of 12.7 MeV. This is compatible to the expectations [3].

The template fit to m_W, fixing all the POIs and allowing the PDF nuisance parameters to be constrained, predicts a PDF uncertainty of 3.0 MeV on m_W. This is equivalent to the procedure performed in Ref. [11] to assess the PDF constraint, but in this framework it is possible to propagate the constrained uncertainty directly to m_W.

Thank you for your attention!

Discussions