Parton Distributions for the Precision Era at the LHC

Keping Xie

Michigan State University

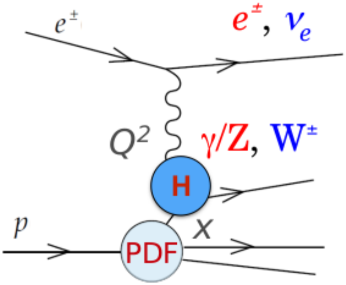
LPC Physics Forum
October 26, 2023

On behalf of the CTEQ-TEA (Tung Et Al.) Collaboration

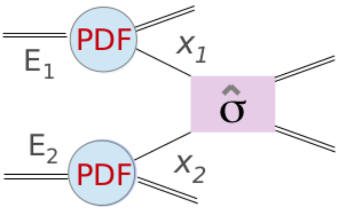
Outline

QCD collinear factorization

DIS structure functions



hadron-hadron collision



The QCD collinear factorization theorem separates long-distance and short-distance contributions in high-energy scatterings involving initial-state hadrons, and gives the predictions as

$$F_a(x,Q^2) = \sum_{i=q,\bar{q},g} C_a^i\left(\frac{x}{\xi}, \frac{Q^2}{\mu_R^2}, \frac{\mu_F^2}{\mu_R^2}, \alpha_s(\mu_R^2)\right) \otimes f_{i/h}(\xi,\mu_F^2)$$
$$+\mathcal{O}(\Lambda_{\rm QCD}^2/Q^2).$$

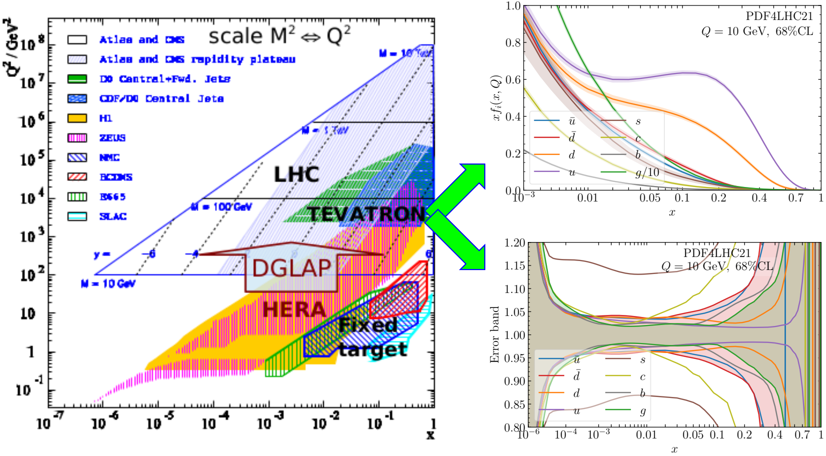
[Collins, Soper, Sterman, 1989]

- C_a^i: Coefficient functions, hard scattering, infrared (IR) safe, calculable in pQCD, independent of hadrons
- $f_{i/h}$: PDFs, reveal inner structure of hadrons, non-perturbative (NP) origin, universal, *e.g.*, for DIS and pp collisions.
- μ_F factorization scale
- PDFs run in terms of the DGLAP equation

Normally, we choose $\mu_R = \mu_F = Q$, thus, Q dependence (scaling violation) of F_2 mainly resulted from PDFs, predicted by the DGLAP equation.

Global analysis of PDFs

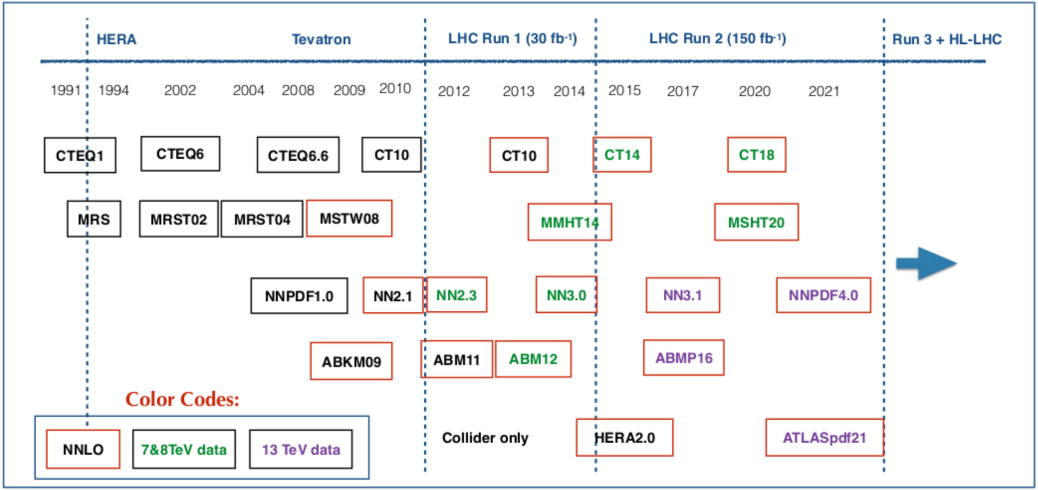
- PDFs are usually extracted from global analyses on variety of experimental data, *e.g.*, DIS, Drell-Yan, jets and top quark production at fixed-target and collider experiments, with increasing impact from the LHC, as well as the precise SM parameters. [See 1709.04922,1905.06957 for recent reviews.]



- Diversity of the analyzed data are important to separate flavors and to avoid theoretical and experimental bias.
- Can be extended to include EW ($\sin^2\theta_W$) and new physics (SMEFT) parameters simultaneously.
- Alternative extracted from lattice QCD simulations, through quasi- or pseudo-PDFs.

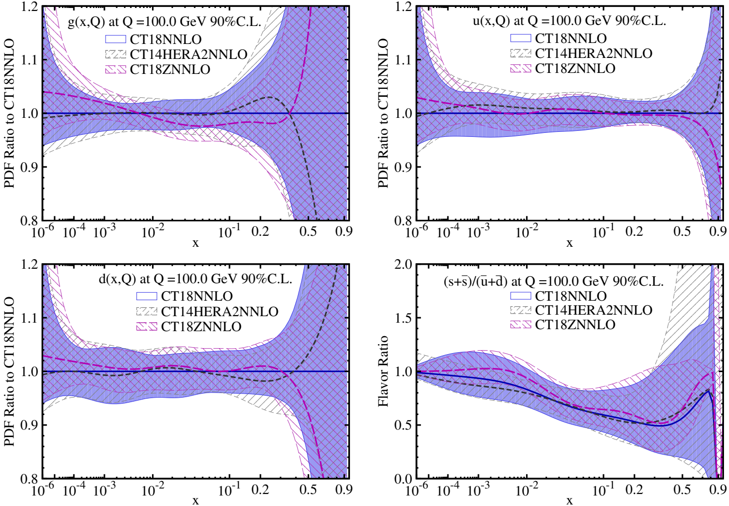
Major PDF analysis groups

- PDFs provided by several major analysis groups, including, CT, MSHT, NNPDF, ABM, HERAPDF, ATLASpdf, CJ, JAM, etc, with slightly different heavy-quark schemes, selections of data, and methodologies. [Jun Gao, DIS 2022]



CTEQ-TEA PDFs

- CT18 PDFs give moderate reductions of PDF uncertainties due to new LHC data sets, agreeing with previous CT14 within uncertainties.
- Alternative CT18A/X/Z fits to deal with small-x saturation and the ATLAS 7 TeV W/Z data

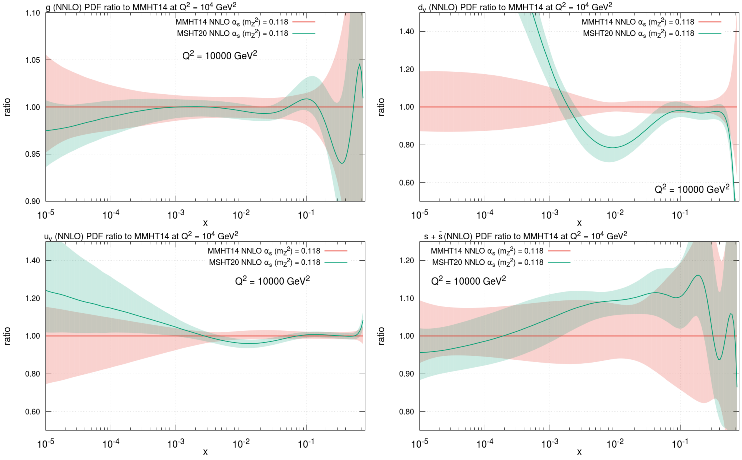


- CT18 vs CT14: gluon unc. reduced everywhere (jets, Z p_T, top), d-quark unc. reduced around $x \sim 0.2$ (LHCb W/Z), s-quark almost unchanged.
- CT18A include ATLAS 7 TeV W/Z as a separate fit due to the tension with other data.
- CT18X include the small-x saturation effect
- CT18Z accumulate both effects, differing from CT18 mostly in gluon and s-quark PDFs

[KX et al., CT18, 1912.10053]

MSHT PDFs

- MSHT20 (Mass Scheme Hessian Tolerance) PDFs adopted an extended parametrization form, as comparison to th MMHT14, to accommodate newly included LHC precision data

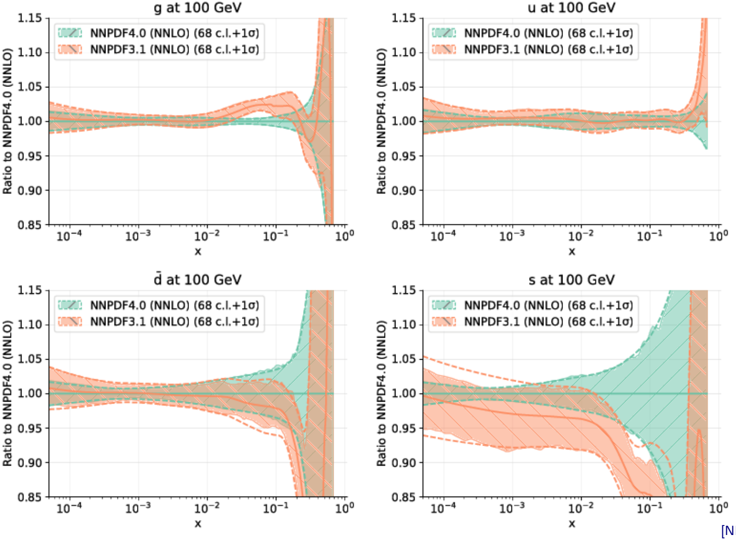


- central of gluon PDF remains mostly unchanged except for a suppression at $x > 0.2$; moderate reduction on gluon uncertainty
- enhancement of s-quark at intermediate x region and large reduction on uncertainty, due to ATLAS 7 TeV W/Z data and update of dimuon theory calculations
- new parametrization 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

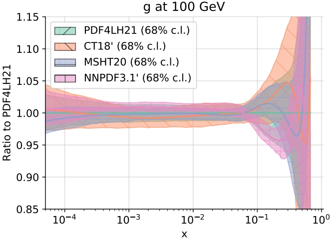


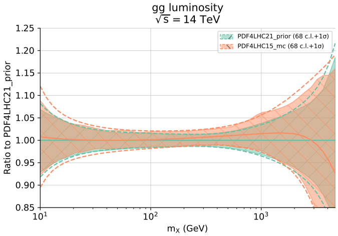
- changes on parametrization and NN architecture, optimization algorithm; additional positivity and integrability constraints and post-fit selections
- central PDF of NNPDF4.0 is generally consistent with NNPDF3.1 except for a notable decrease of gluon PDF at $x \sim 0.1$ and moderate increase of strangeness
- NNPDF4.0 shows PDF uncertainty of $\sim\ 1-2\%$ at data constrained region, largely reduced comparing to NNPDF3.1

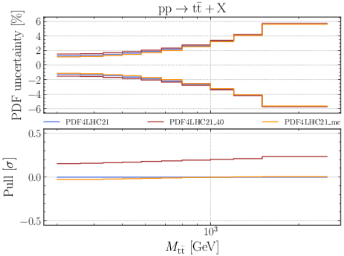
[NNPDF4.0, 2109.02653]

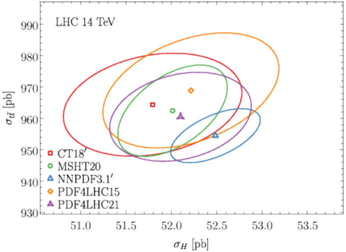
PDF4LHC21 combinations

- The PDF4LHC group performs extensive benchmarks on methodologies of several groups, and present the PDF4LHC21 PDFs as an effective combination of CT18’, MSHT20, and NNPDF3.1’, as a recommendation for LHC Run3 usage







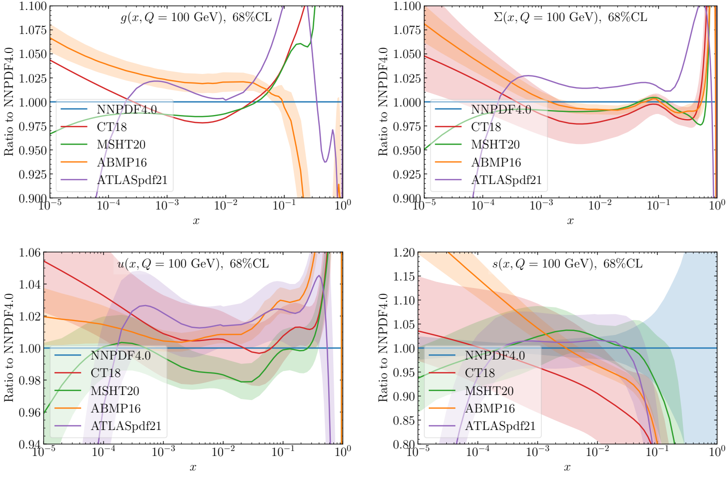


- CT18’ differs slightly from CT18 by using $m_c = 1.4$ GeV, NNPDF3.1’ differs frm 3.1 by including additional jet and top-quark data
- PDF4LHC21 PDFs are presented in the form of either a MC set of 100 replicas or Hessian 40 sets.
- PDF unc. at the level of $2 \sim 3\%$ for the inclusive cross sections, which increase up to $5 \sim 10\%$ in the multi-TeV region

[KX et al., PDF4LHC21, 2203.05506]

PDF benchmarking

- Many efforts on comparisons and understanding of differences of up-to-date PDFs, in order to have a faithful determination of PDFs and its uncertainties [KX et al., Snowmass 2021, 2203.13923]

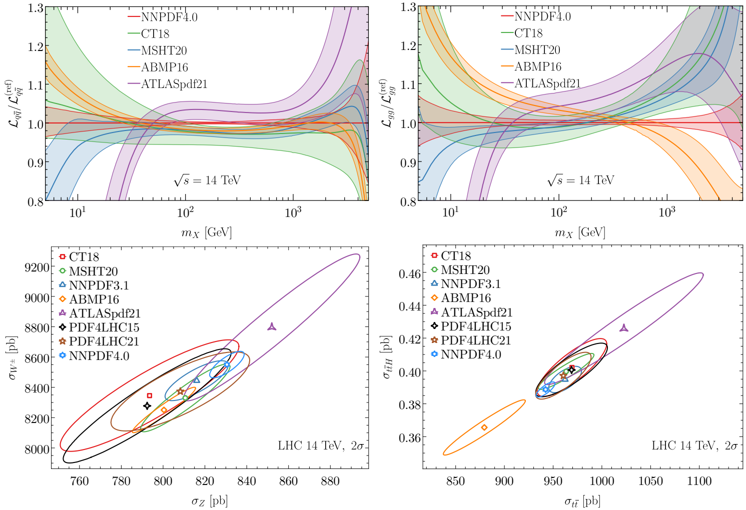


- general agreement among different groups over range $x \sim 10^{-4} - 10^{-1}$ within uncertainties
- gluon notable difference at $x \sim 0.2$ with 2σ for NNPDF4.0 vs CT18 and MSHT20.
- singlet ATLASpdf21 deviate at $x < 10^{-4}$ due to $Q^2 > 10$ GeV^2 applied on HERA data and $x > 0.2$ due to lack of fixed-target data
- NNPDF and ABMP show uncertainty of $1 - 2\%$ in constrained region, mostly due to methodologies; CT18 being conservative among all fits

Phenomenological implications

- 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

[KX et al., Snowmass 2021, 2203.13923]

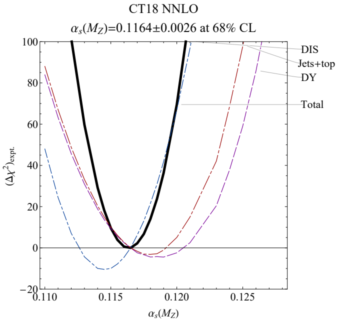


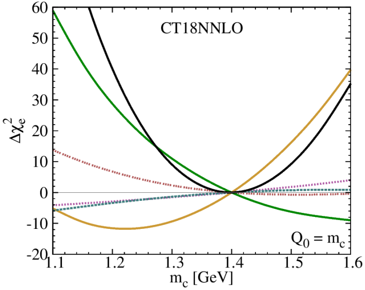
- gg luminosity shows a spread of more then 20% in the multi-TeV region; $q\bar{q}$ agrees better in general, exceptat a mass around 300 GeV.
- 2σ error ellipses show cross sections of standard candle processes; NNPDF4.0 shows an uncertainty of less than 1%; CT18 2σ ellipse cover most groups

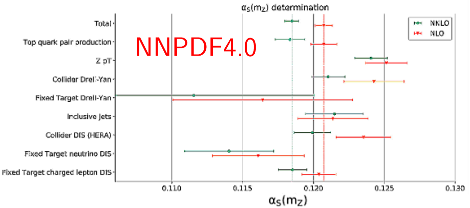
Outline

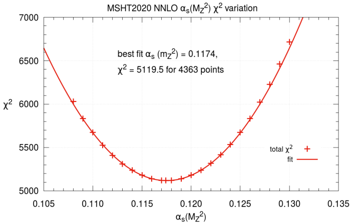
QCD parameter determinations

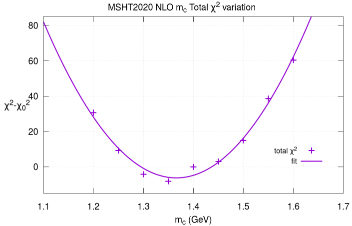
- From global analysis of PDFs, one can also extract QCD parameters, including the strong coupling [KX et al., Snowmass 2021, 2203.08271] at NNLO with compatible precision, and the heavy-quark pole masses.







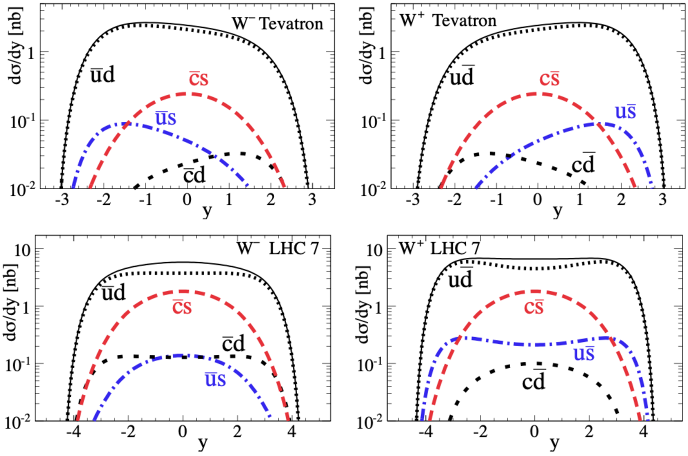




- The 3 most recent global fits gives $\alpha_s(M_Z) = 0.1164, 0.1174, 0.1185$ at NNLO for CT18, MSHT20, and NNPDF4.0, respectively.
- Very mild sensitivity to $m_c = 1.35$ GeV for both CT18Z and MSHT20.

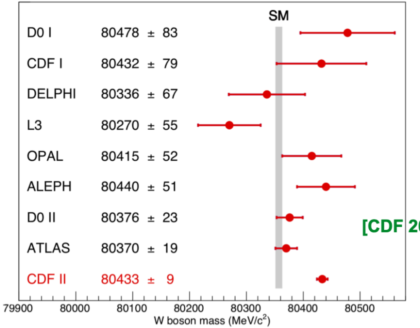
W-boson mass measurement

- PDFs are key inputs for precision programs at hadron colliders, e.g., direct measurements on the W boson mass and the weak mixing angle



[ATLAS 2018]

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
Parton shower μ_F with heavy-flavour decorrelation	5.0	6.9	5.0	6.9	5.0	6.9
Parton shower PDF uncertainty	3.6	4.0	2.6	2.4	1.0	1.6
Angular coefficients	5.8	5.3	5.8	5.3	5.8	5.3
Total	15.9	18.1	14.8	17.2	11.6	12.9

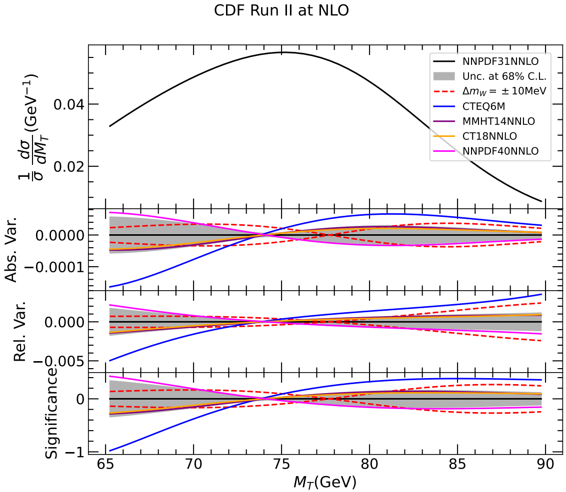


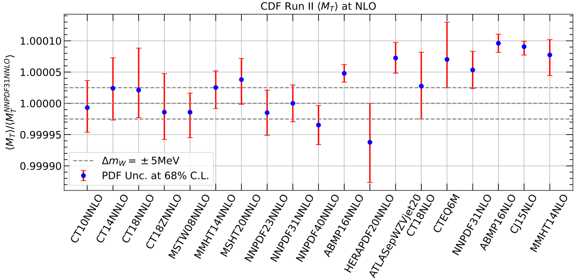
[CDF 2022]

Fig. 5

PDF impact on the m_W measurement

- 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





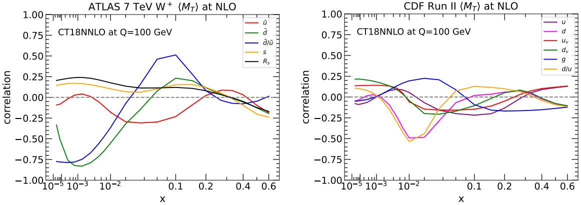
The shift and PDF unc. of W mass

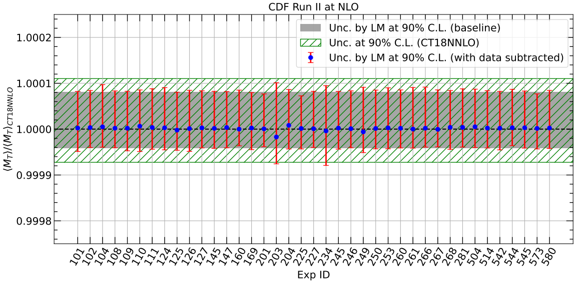
δM_W in MeV	sta.	NNPDF3.1	CT18	MMHT2014	NNPDF4.0	MSHT2020
$\langle M_T\rangle$(LO)	–	$0^{+8.3}_{-8.3}$	$-1.0^{+8.3}_{-11.4}$	$-3.3^{+7.4}_{-4.2}$	$+7.8^{+5.1}_{-5.1}$	$-3.1^{+6.7}_{-5.7}$
χ^2 fit (LO)	8.0	$0^{+7.6}_{-7.6}$	$-1.0^{+5.4}_{-8.6}$	$-3.3^{+6.1}_{-3.0}$	$+8.0^{+3.7}_{-3.7}$	$-3.0^{+5.0}_{-4.0}$
$\langle M_T\rangle$(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	$0^{+3.9}_{-3.9}$	–	–	–	–

[Gao, Liu, KX, 2205.03942]

PDF variations cannot explain the CDF discrepancy

- 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



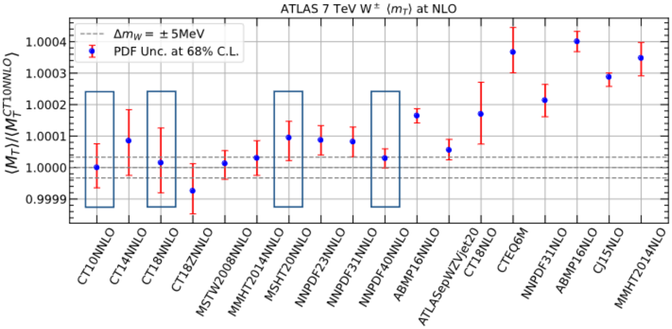


- m_T at CDF (ATLAS) is mostly sensitive to the d-quark (dbar-quark) at x 0.01(0.001); CDF and ATLAS are largely uncorrelated
- m_T at CDF is largely constrained by the DIS and Drell-Yan data on deuteron target, the Tevatron lepton charge asymmetry data; at ATLAS also the CMS charge asymmetry data

[Gao, Liu, KX, 2205.03942]

Transverse mass at ATLAS

- We further investigate PDF variations on transverse mass distribution focusing on the ATLAS 7 TeV measurement; note the transverse momentum distribution has a relatively larger weight in ATLAS



[Gao, Liu, KX, 2205.03942]

- 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

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

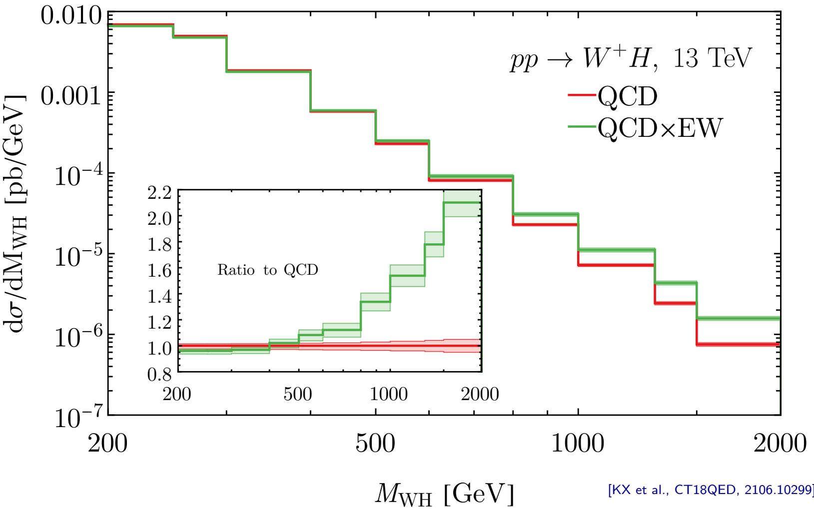
$$m_W = 80362 \pm 23_{\text{stat}} \pm 10_{\text{exp}} \pm 17_{\text{theory}} \pm 9_{\text{PDF}}\ \text{MeV},$$
$$m_W = 80350 \pm 23_{\text{stat}} \pm 10_{\text{exp}} \pm 17_{\text{theory}} \pm 12_{\text{PDF}}\ \text{MeV},$$
$$m_W = 80351 \pm 23_{\text{stat}} \pm 10_{\text{exp}} \pm 17_{\text{theory}}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

ATLAS, CT10 + 3.8 MeV (MMHT14-CT14)

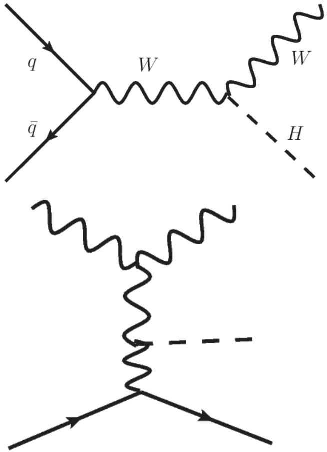
W-boson charge	W^+		W^-		Combined	
Kinematic distribution	p_T^ℓ	m_T	p_T^ℓ	m_T	p_T^ℓ	m_T
δm_W [MeV]						
Fixed-order PDF uncertainty	13.1	14.9	12.0	14.2	8.0	8.7
AZ tune	3.0	3.4	3.0	3.4	3.0	3.4
Charm-quark mass	1.2	1.5	1.2	1.5	1.2	1.5

Electroweak corrections become vita

- At high mass tail, the EW corrections and photon PDF $\gamma(x,\mu^2)$ becomes important

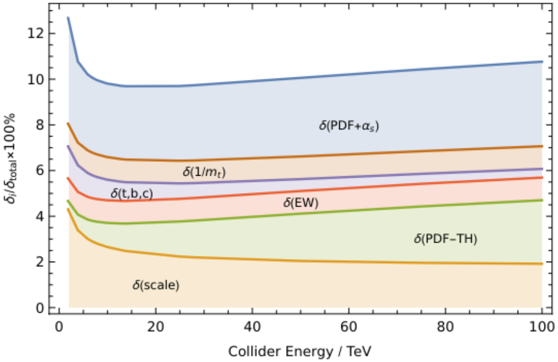


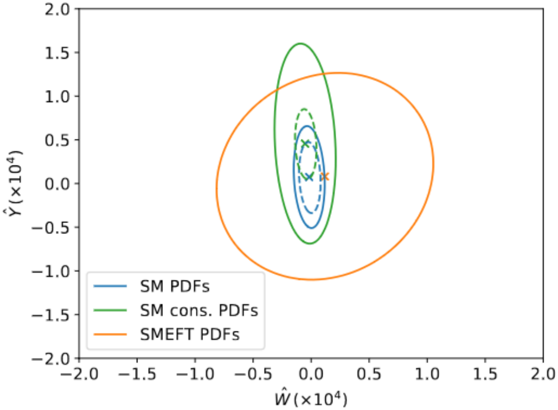
[KX et al., CT18QED, 2106.10299]



PDF impacts on Higgs, BSM, and SMEFT

- PDF uncertainties becomes one of the main theoretical obstacles to explore Higgs and SMEFT as well as other BSM physics





[KX et al., Snowmass 2021, 2203.13923]

Outline

Post-CT18 LHC Drell-Yan data

[KX et al., 2305.10733]

ID	Experiment	N_{pt}	χ^2/N_{pt}				
			Pre-fit†	ePump†	CT18	CT18A	CT18As
215	ATLAS 5.02 TeV W, Z	27	1.15	0.96	1.07	0.74	0.71
211	ATLAS 8 TeV W	22	4.96	2.98	2.46	2.72	2.49
214	ATLAS 8 TeV Z 3D	188	1.95	1.18	1.16	1.13	1.14
212	CMS 13 TeV Z	12	9.24	2.93	2.75	1.89	2.02
216	LHCb 8 TeV W	(16)14	(3.48)1.52	(3.24)1.45	(2.81)1.33	(1.89)1.45	(3.00)1.52
	LHCb 13 TeV Z	18	0.89	0.88	0.99	0.92	0.90
213	LHCb 13 TeV $Z \to \mu^+\mu^-$	(18)16	(2.39)1.27	(2.33)1.17	(2.55)1.12	(2.49)1.12	(2.28)0.87

ID	Experiment	N_{pt}	χ^2/N_{pt}					
			CT18	CT18A	CT18As	ATLASpdf21	MSHT20	NNPDF4.0
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	1.41	2.61	[3.50]
214	ATLAS 8 TeV Z 3D†	188	1.12	1.14	1.18	1.13(184)	1.45(59)	1.22(60)
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	1.24(55)	1.91(61)	1.67(61)
Total 3994/3953/3959 points			1.20	1.20	1.19	–	–	–

Post-CT18 top-quark-pair data

[KX et al., 2307.11153]

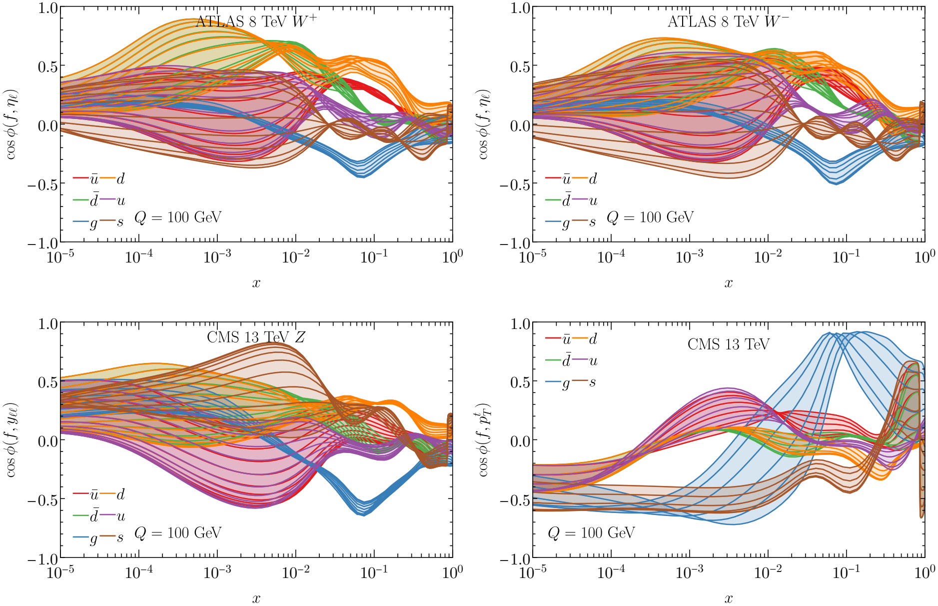
Scale				$H_T/2$	$H_T/2$	$H_T/4$	$H_T/4$
ID#	**Data set**	$N_{\rm pt}$	CT18	nTT1	nTT2	nTT1	nTT2
573	CMS 8 TeV 2D norm. (p_T, y)	16	1.2	1.2	1.2	1.1	1.1
580	ATLAS 8 TeV $p_T^t, m_{t\bar{t}}$ comb.	15	0.6	0.7	0.7	0.7	0.7
521	ATLAS 13 TeV all-hadronic $y_{t\bar{t}}$	12	-	1.0	1.0	1.1	1.1
528	CMS 13 TeV dilepton $y_{t\bar{t}}$	10	-	0.8	0.8	0.5	0.7
532	ATLAS 13 TeV l+j $y_{t\bar{t}}$	10	-	0.7	-	0.8	-
587	ATLAS 13 TeV l+j comb.	34	-	-	0.7	-	1.1
581	CMS 13 TeV l+j $m_{t\bar{t}}$	15	-	1.1	1.1	1.6	1.7

Other types of LHC data are being checking in progress

- Jet data: inclusive jet, di-jet
- high-p_T Z and Z+jet
- $W+c$, $Z+c$
- ...

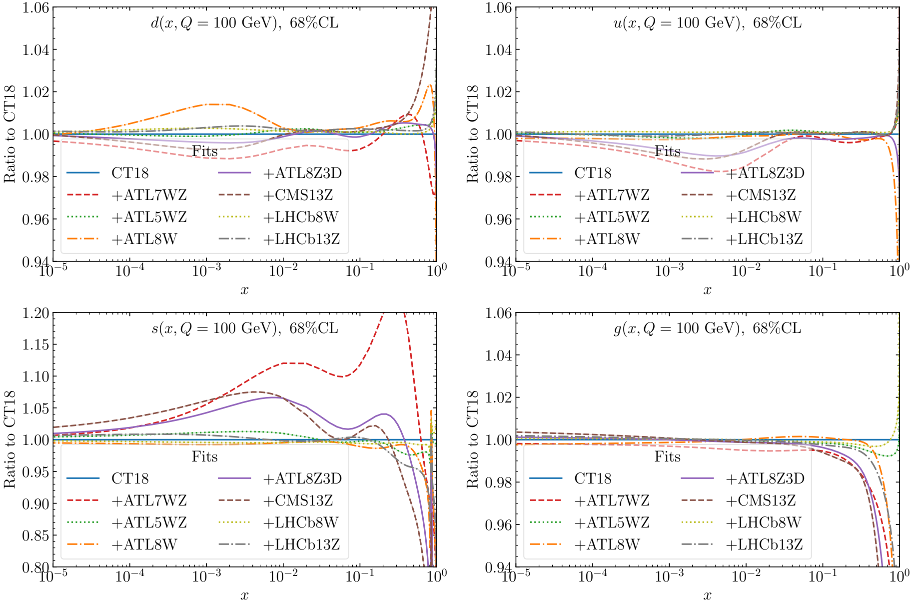
Sensitivity to PDFs

[KX et al., 2305.10733, 2307.11153]



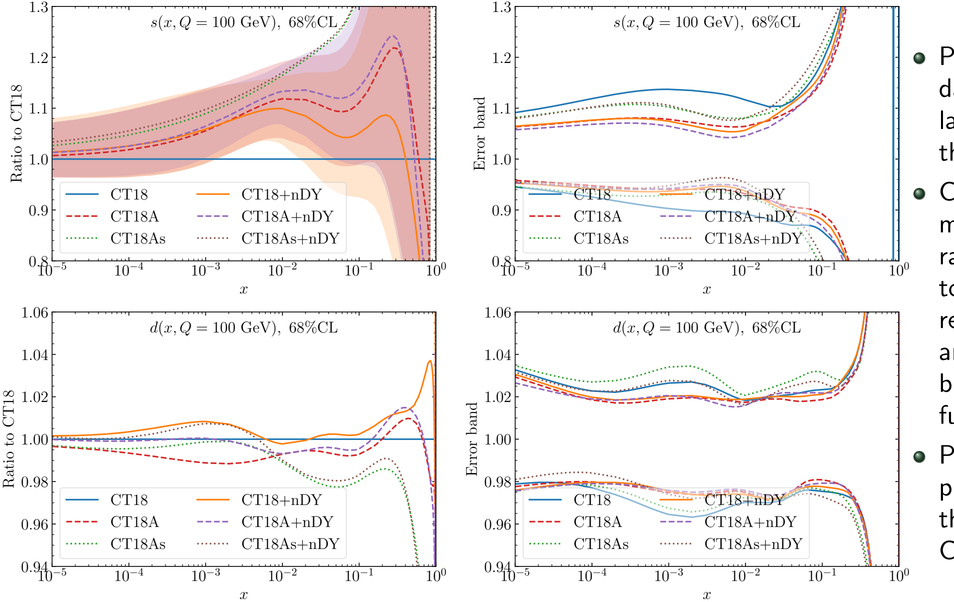
- $W^+(u\bar{d})$ is sensitive to $\bar{d}$, and small-x $\bar{d}$ and d are correlated ($g \to d\bar{d}$)
- $W^-(d\bar{u})$ is sensitive to d
- Z is sensitive to g and s
- $t\bar{t}$ is sensitive to large-x g

Impacts of Drell-Yan data on the CT18 PDFs



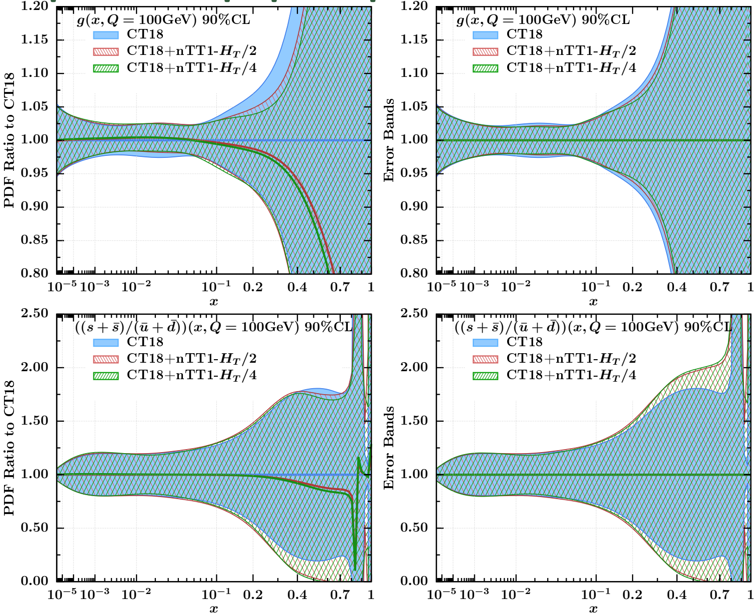
- Post-CT18 Drell-Yan data are consistent to pull s (mainly Z) and u
- An exception ATL8W give an opposite pull to $d(\bar{d})$
- Pull on gluon is very mild

Simultaneous fits



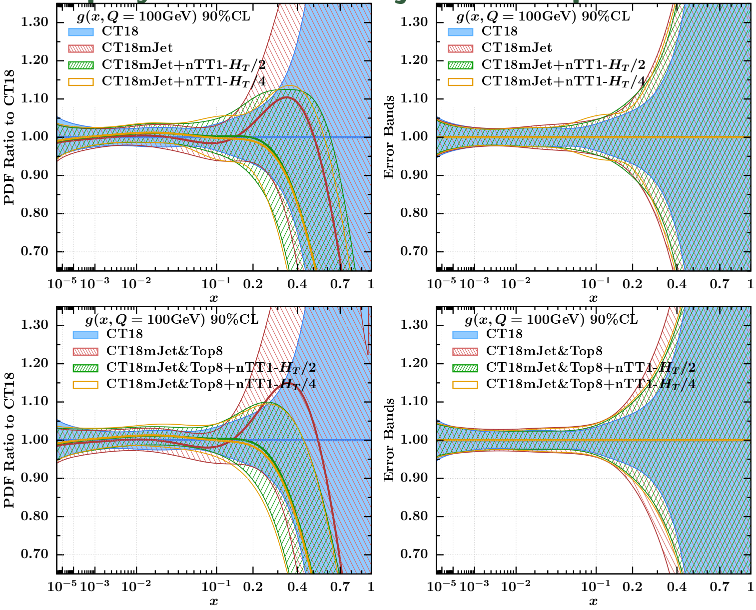
- Post-CT18 DY data pull s larger and shrink the error band
- CT18As add more flexible parameterization to s PDF (to relax tension), and allows s to be pulled further away
- Post-CT18 DY pull CT18 to the opposite to CT18A(ATL7WZ)

Impact of top-quark-pair data on the CT18 PDFs



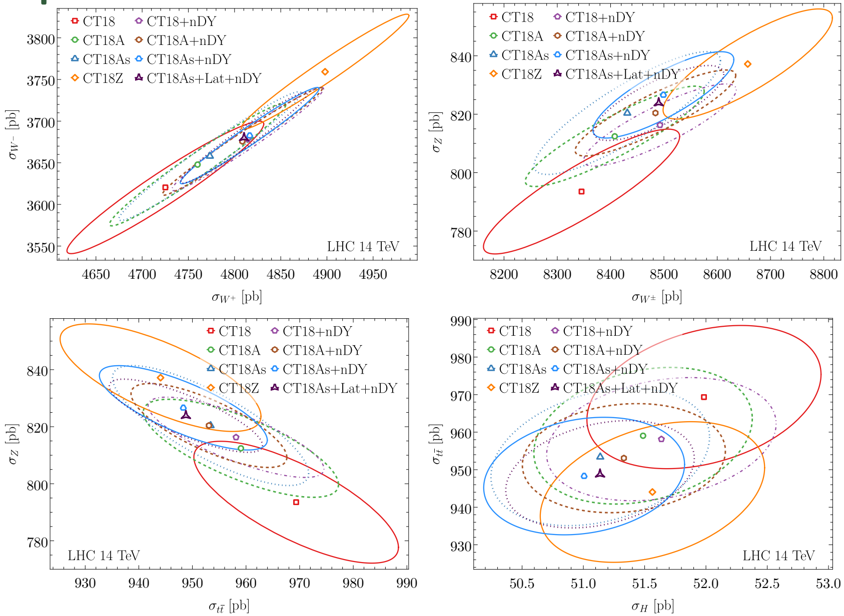
- $t\bar{t}$ data soften g PDF at large x
- PDF error bands are shrunk mildly
- $R_s = (s+\bar{s})/(\bar{u}+\bar{d})$ error band can be enlarged due to the sum rules.

Interplay between the jet and top data



- $t\bar{t}$ and jet data pull the large-x gluon to the same direction (softer)
- The $t\bar{t}$ impact is smaller than jet, mainly driven by the smaller number of data points

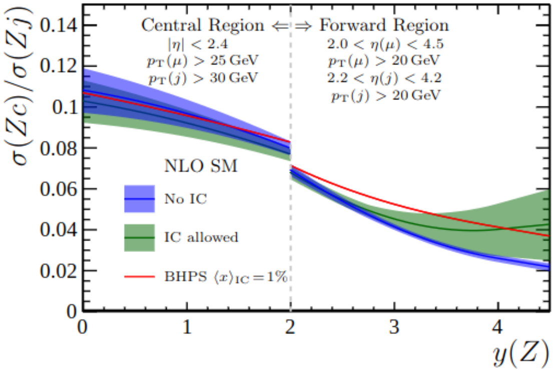
Implications to the future measurements

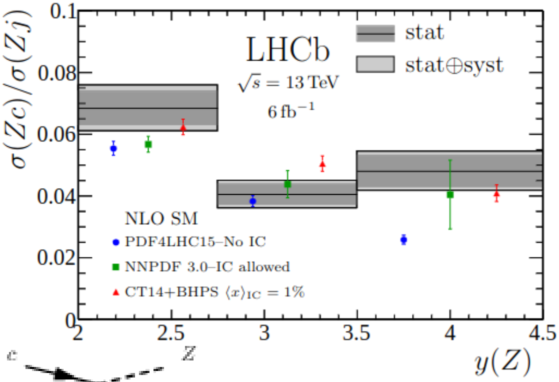


- Post-CT18 DY data pull the prediction of CT18 to the CT18Z direction
- The semi-major axis $(\sigma_Z + \sigma_{W^\pm})$ pull is driven by gluon PDF
- The semi-minor axis $(\sigma_Z / \sigma_{W^\pm})$ pull is identified as strangeness
- The error bands get a reduction, reflecting the constraining power

Outline

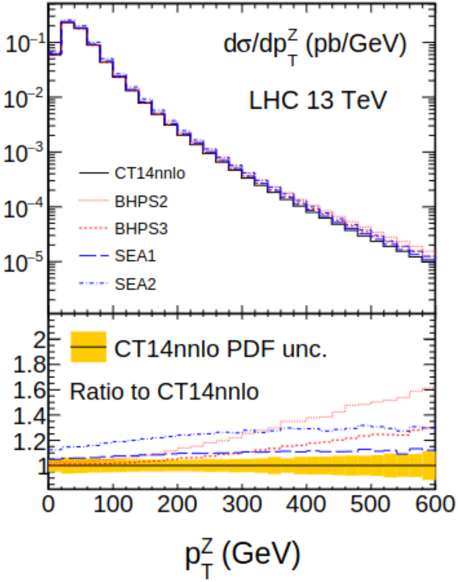
Intrinsic charm and $Z + c$ production

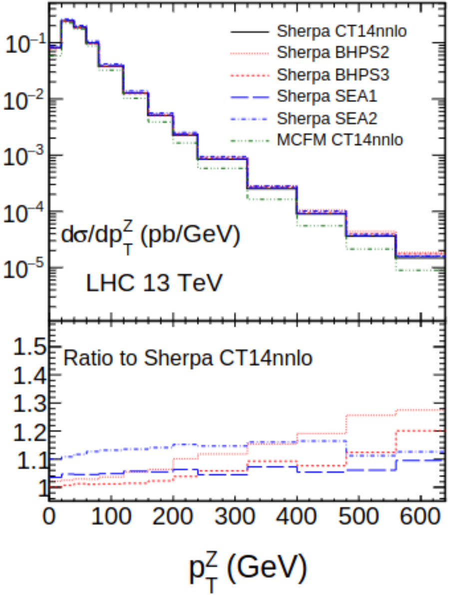




- $Z + c$ production directly probes charm PDF
- The forward measurement give sensitivity to a large $x(\sim Q/\sqrt{s}e^y)$, where intrinsic charm (valence-like) shows up
- CMS measurement [1711.02143] didn't find enough significance, while LHCb measurement [2109.08084] prefers intrinsic charm component

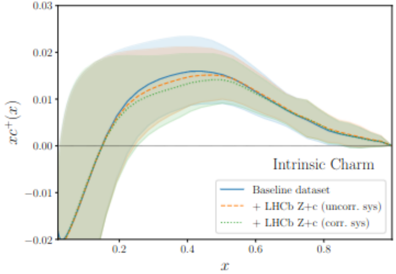
CT14 exploration

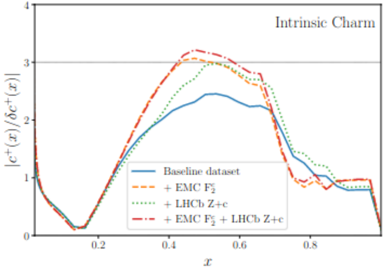


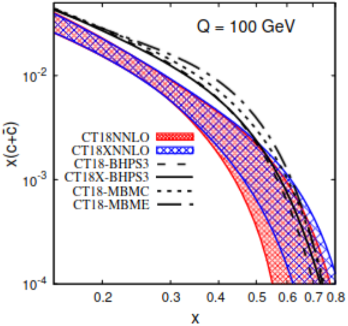


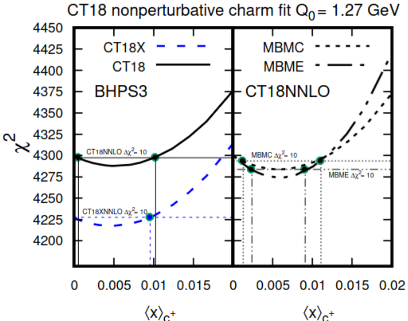
- We have pointed out the sensitivity of $Z + c$ measurement [Xie et al., 1707.00657]
- Parton-Shower (final-state radiation) can dilute the sensitivity

NNPDF's discovery and CT's reaction



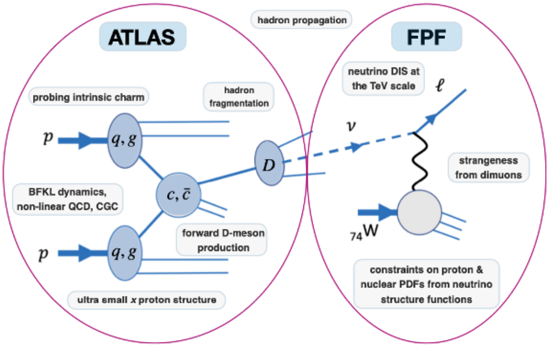


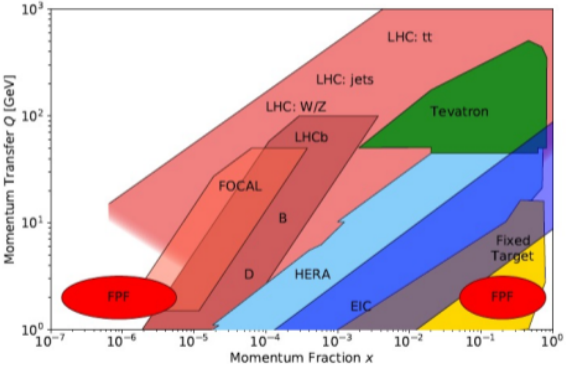




- NNPDF found a 3σ discovery of intrinsic charm PDF [Nature, 2208.08372]
- We re-do a more comprehensive IC analysis, and still found the significance is about 1σ [LX et al., 2211.01387]
- Driven by NNPDF's error band [KX et al., 2205.10444]
- See Nadolsky's talks, LPC and SM@LHC2023

Forward Physics Facility

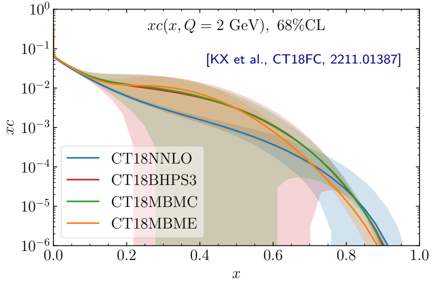


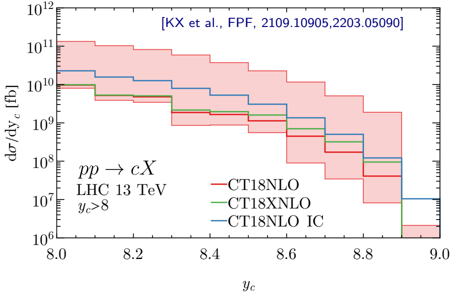


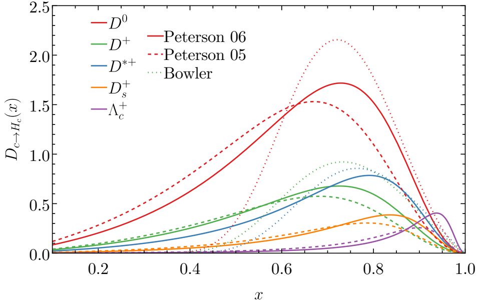
The neutrino measurement in the ongoing FASER and future FPF probe both large and small Bjoken x.

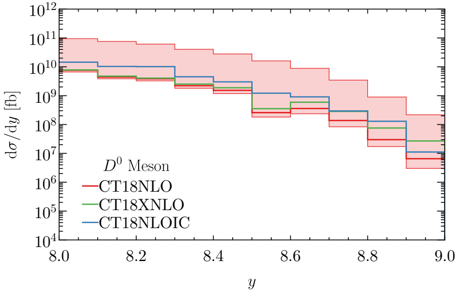
[KX et al., FPF, 2203.05090]

The charm production at the FPF

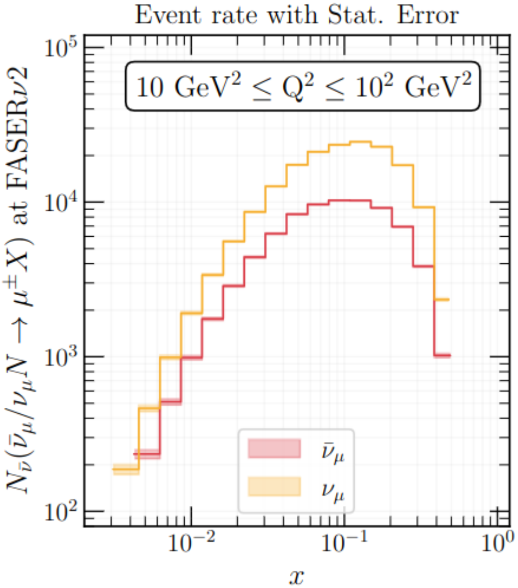


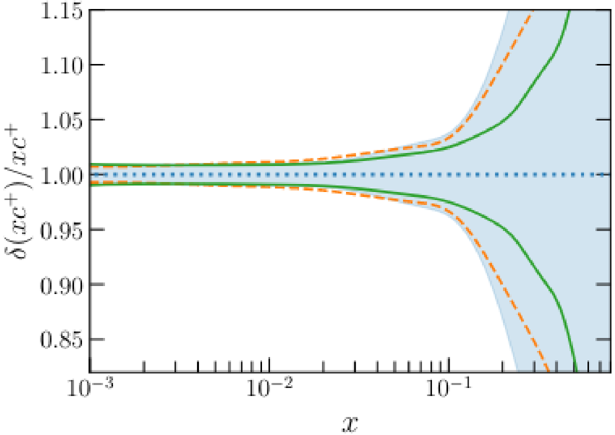






LHC (FPF) as a neutrino-ion collider





Neutrino cross section measurement constrains the charm and gluon PDFs [NNPDF, 2309.09581]

Towards (a)N3LO PDFs

- MSHT20aN3LO PDFs [2207.04739]
- NNPDF preliminary PDFs [Giacomo Magni, DIS2023]
- Ongoing aN3LO PDF benchmarking

Many challenges

- N3LO calculations (Les Houches wish list [2207.02122])
- Tolerance puzzle
- Theoretical uncertainties (scale)

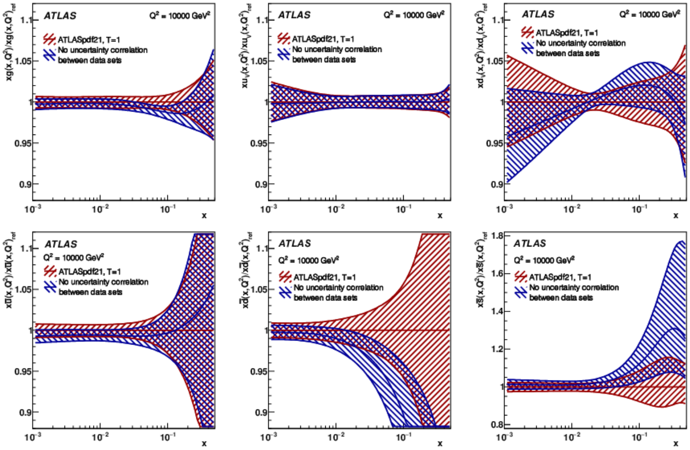
Not a conclusion

- 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.
- The PDF uncertainty becomes one of the main obstacles, M_W measurement, Higgs cross section
- Fitting current and future LHC precision data (Drell-Yan, $t\bar{t}$, jet) data become challenging
- Further developments, including both theoretical and experimental, are essential
- Intrinsic charm component of proton is still a puzzle, and future data are critical

Thanks for your attention!

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, 13 TeV with several new features explored



- Consistent pull from 7, 8 TeV inclusive W, Z data on strangeness
- W/Z+jet data stabilize sea quarks at large x closed to preferences from fixed-target experiments
- impact of correlation of experimental uncertainties across different data sets, and of scale variations are investigated

[ATLASpdf21, 2112.11266]