CT25 nucleon PDFs for precision QCD phenomenology

Keping Xie

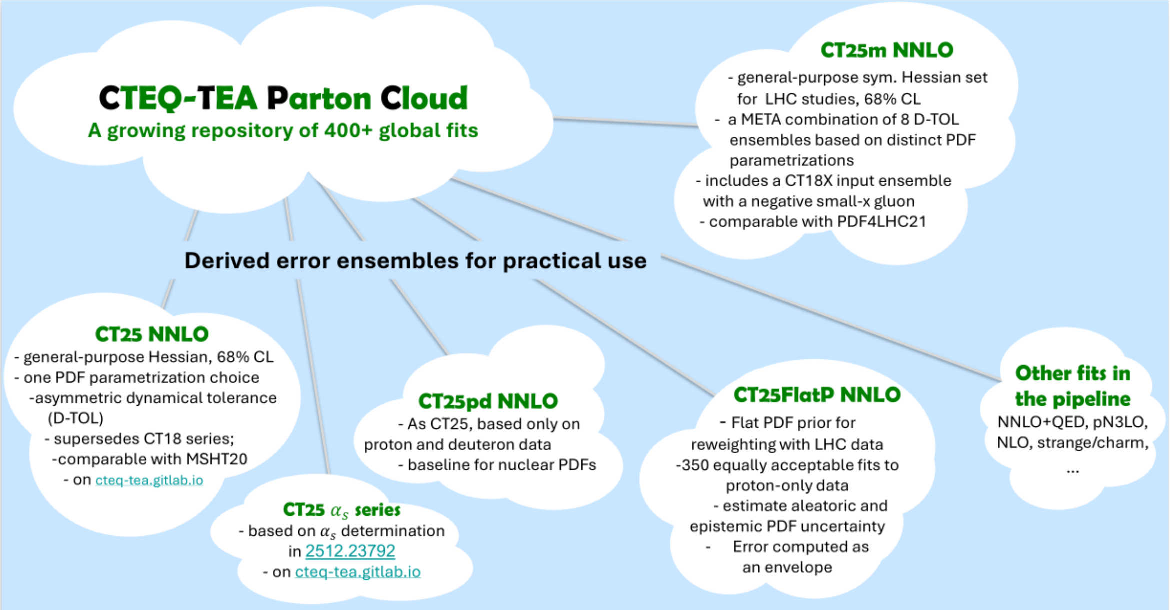
李政道研究所
TSUNG-DAO LEE INSTITUTE

6th LHCb Workshop @ SCNU
May 24, 2026

On behalf of the CTEQ-TEA (Tung Et Al.) Collaboration:
A. Ablat, A. Courtoy, S. Dulat, Y. Fu, M. Guzzi, T.J. Hobbs, J. Huston, E. Lopez, K. Mohan, P. Nadolsky, M. Ponce Chavez, T. Sharma, D. Stump, C.-P. Yuan

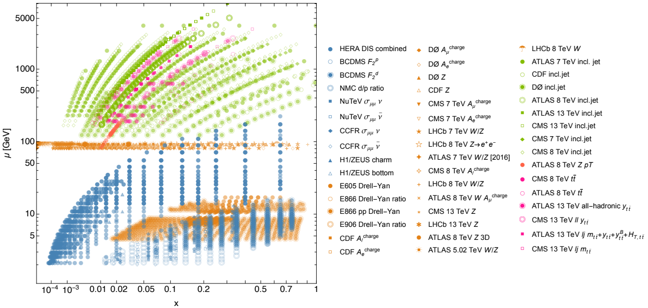
CT25 PDFs in the nucleon

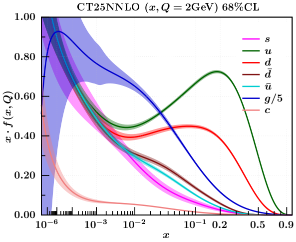
- A wide-ranging study on PDFs for NNLO and partial N^3LO QCD calculations.
- Improved predictions for the large-x region ($x > 0.1$).
- CT25 supersedes the CT18 family of PDFs.
- Four PDF error ensembles based on more than 300 exploratory fits with varied PDF parametrizations and methodological assumptions.
- Allows customization of error PDFs for concrete applications.
- Moved away from a single tolerance prescription.
- Multiple uncertainty sources are estimated:
  - statistical/systematic,
  - experimental/theoretical/parametrization,
  - data-selection dependence.

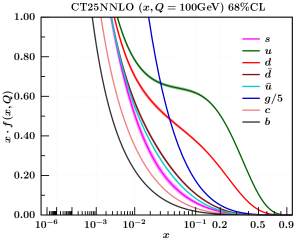

CTEQ-TEA Parton Cloud
A growing repository of 400+ global fits
CT25m NNLO
- general-purpose sym. Hessian set for LHC studies, 68% CL
- a META combination of 8 D-TOL ensembles based on distinct PDF parametrizations
- includes a CT18X input ensemble with a negative small-x gluon
- comparable with PDF4LHC21
Derived error ensembles for practical use
CT25 NNLO
- general-purpose Hessian, 68% CL
- one PDF parametrization choice
-asymmetric dynamical tolerance (D-TOL)
- supersedes CT18 series;
-comparable with MSHT20
- on cteq-tea.gitlab.io
CT25pd NNLO
- As CT25, based only on proton and deuteron data
- baseline for nuclear PDFs
CT25FlatP NNLO
- Flat PDF prior for reweighting with LHC data
-350 equally acceptable fits to proton-only data
- estimate aleatoric and epistemic PDF uncertainty
- Error computed as an envelope
Other fits in the pipeline
NNLO+QED, pN3LO, NLO, strange/charm, ...
CT25 α_s series
- based on α_s determination in 2512.23792
- on cteq-tea.gitlab.io

CT25 NNLO

- General-purpose Hessian set at 68% CL.
- Uses asymmetric dynamical tolerance.
- Supersedes CT18; comparable with MSHT20.
- Includes 13 precision LHC data sets at:
  - 5.02 TeV,
  - 8 TeV,
  - 13 TeV.
- Includes SeaQuest Drell–Yan data.
- Based on:
  - arXiv:2305.10733 (Drell-Yan),
  - arXiv:2307.11153 (Top-quark pair),
  - arXiv:2412.00350 (jets).







Drell-Yan data in CT18(Z) global analyses

ID	Expt.	$N_{\rm pt}$	χ^2	$\chi^2/N_{\rm pt}$	S_E
201	E605DY	119	103.4(102.4)	0.9(0.9)	-1.0(-1.1)
203	E866 $\sigma_{pd}/(2\sigma_{pp})$	15	16.1(17.9)	1.1(1.2)	0.3(0.6)
204	E866 $Q^3 d^2\sigma_{pp}/(dQdx_F)$	184	244(240)	1.3(1.3)	2.9(2.7)
225	CDF1Z $A(e)$	11	9.0(9.3)	0.8(0.8)	-0.3(-0.2)
227	CDF2W $A(e)$	11	13.5(13.4)	1.2(1.2)	0.6(0.6)
234	D∅2W $A(\mu)$	9	9.1(9.0)	1.0(1.0)	0.2(0.1)
260	D∅2Z $y_{\ell\ell}$	28	16.9(18.7)	0.6(0.7)	-1.7(-1.3)
261	CDF2Z $y_{\ell\ell}$	29	48.7(61.1)	1.7(2.1)	2.2(3.3)
266	CMS7W $A(\mu)$	11	7.9(12.2)	0.7(1.1)	-0.6(0.4)
267	CSM7W $A(e)$	11	4.6(5.5)	0.4(0.5)	-1.6(-1.3)
268	ATL7WZ$_{(2012)}$	41	44.4(50.6)	1.1(1.2)	0.4(1.1)
281	D∅2W $A(e)$	13	22.8(20.5)	1.8(1.6)	1.7(1.4)
			New LHC data in CT18(Z)		
245	LHCb7WZ(μ)	33	53.8(39.9)	1.6(1.2)	2.2(0.9)
246	LHCb8Z(e)	17	17.7(18.0)	1.0(1.1)	0.2(0.3)
248	ATL7WZ$_{(2016)}$	34	287.3(88.7)	8.4(2.6)	13.7(4.8)
249	CMS8W $A(\mu)$	11	11.4(12.1)	1.0(1.1)	0.2(0.4)
250	LHCb8WZ(μ)	34	73.7(59.4)	2.1(1.7)	3.7(2.6)
253	ATL8ZpT	27	30.2(28.3)	1.1(1.0)	0.5(0.3)

New LHC Drell-Yan data in CT25

Boson	$\sqrt{s}$	Lumi	Observable	Ref.
ATLAS				
W, Z	2.76	4.0 pb^{-1}	$\sigma^{\mathrm{fid,tot}}$	1907.03567
W, Z	13	81.0 pb^{-1}	σ^{fid}	1603.09222
W, Z	5.02	25.0 pb^{-1}	$(\eta_\ell, y_{\ell\ell})$	1810.08424
Z	8	20.2 fb^{-1}	$(m_{\ell\ell}, y_{\ell\ell})$	1710.05167
$W \to \mu\nu$	8	20.2 fb^{-1}	η_μ	1904.05631
Z	13	36.1 fb^{-1}	$p_T^{\ell\ell}$	1912.02844
CMS				
Z	13	2.8 fb^{-1}	$m_{\ell\ell}$	1812.10529
Z	13	35.9 fb^{-1}	(y, p_T, ϕ^*)	1909.04133
W	13	35.9 fb^{-1}	σ^{fid}, $y_W, (\eta_\ell, p_T^\ell)$	2008.04174
LHCb				
$W \to e\nu$	8	2.0 fb^{-1}	η_e	1608.01484
Z	13	294 pb^{-1}	σ^{fid}, (y, p_T, ϕ^*)	1607.06495
$Z \to \mu\mu$	13	5.1 fb^{-1}	σ^{fid}, (y, p_T, ϕ^*)	2112.07458

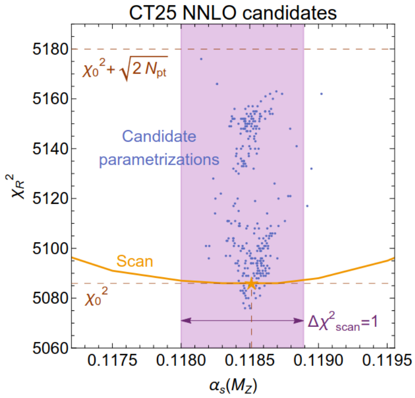
(Pseudo-)Rapidity distributions for DY (~ 300 data points) newly included in CT25.

Drell-Yan data fits

ID	Experiment	$N_{\rm pt}$	$\chi^2/N_{\rm pt}$					
			CT18	CT18A	CT18As	ATLASpdf21	MSHT20	NNPDF4.0
215	ATL5WZ	27	0.89	0.70	0.70	–	–	–
211	ATL8W	22	2.75	2.94	2.79	1.41	2.61	[3.50]
214	ATL8Z3D	188	1.14	1.13	1.17	1.13(184)	1.45(59)	1.22(60)
212	CMS13Z	12	2.45	2.02	1.73	–	–	–
216	LHCb8W	14	1.41	2.02	1.73	–	–	–
213	LHCb13Z	16	1.24	0.98	0.82	–	–	–
248	ATL7WZ	34	2.59	2.51	2.31	1.24(55)	1.91(61)	1.67(61)
Total 3994/3953/3959 points			1.20	1.20	1.19	–	–	–

- The global fitted results can be deduced from the individual fits.
- The tension between the ATL8W and ATL7WZ can be relaxed (but not completely resolved) with a more flexible strangeness parameterization.
- With CT18As, the impact on strangeness is minimal, but on $d(\bar{d})$ remains

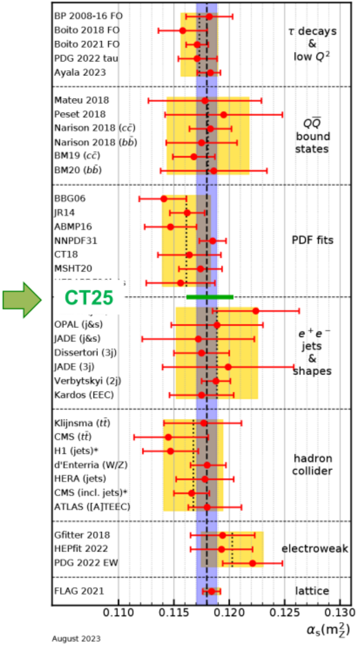
α_s determination



Statistical Method	Eq.	$\delta\alpha_s(M_Z)$
Global Tolerance	6.1	$0.1184^{+0.0026}_{-0.0028}$
Dynamical Tolerance	6.2	$0.1184^{+0.0024}_{-0.0012}$
BHM	5.3	$0.1181^{+0.0022}_{-0.0023}$
GMM	5.5	$0.1183^{+0.0023}_{-0.0023}$
Average		$0.1183^{+0.0023}_{-0.0020}$

- A combination of 4 uncertainty estimation methods: Global Tolerance, Dynamical Tolerance, Bayesian Hierarchical Mdel (BHM), Gaussian Mixture Model (GMM) [2512.23792]
- Tests of uncertainty quantification.
- Consistent with PDG average.

α_s c



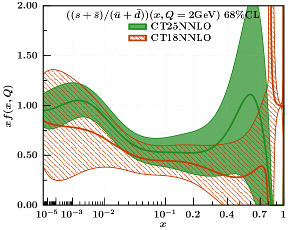
Statistical Method	Eq.	$\delta\alpha_s(M_Z)$
Global Tolerance	6.1	$0.1184^{+0.0026}_{-0.0028}$
Dynamical Tolerance	6.2	$0.1184^{+0.0024}_{-0.0012}$
BHM	5.3	$0.1181^{+0.0022}_{-0.0023}$
GMM	5.5	$0.1183^{+0.0023}_{-0.0023}$
Average		$0.1183^{+0.0023}_{-0.0020}$

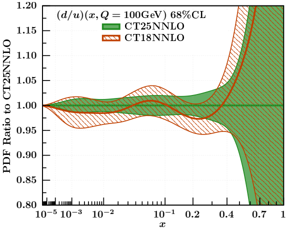
- A … methods: Global Tolerance, Dynamical Tolerance,
  Ba… an Mixture Model (GMM) [2512.23792]
- Te…
- Cc…

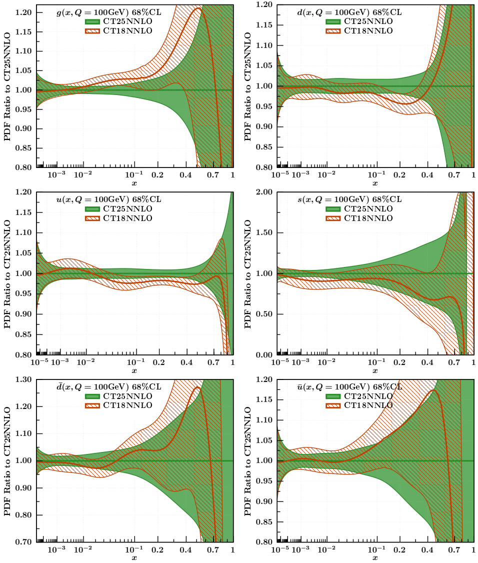
$$\alpha_s(m_Z^2) = 0.1180 \pm 0.0009 \qquad \text{(PDG 2023 average)}.$$

CT25 vs CT18 PDFs

- Error bands consistent within uncertainties
- CT25 uncertainties are smaller because of including new data sets and new definition of tolerance (D-TOL)
- Choices in the data baseline make CT25 closer to CT18Z
- Remove Inclusive neutrino-nucleus DIS (CCFR, CDHSW)
- LHC jet data sets $\rightarrow$ softer gluon at $x > 10^{-2}$
- LHC DY data $\rightarrow$ larger $s(x)$

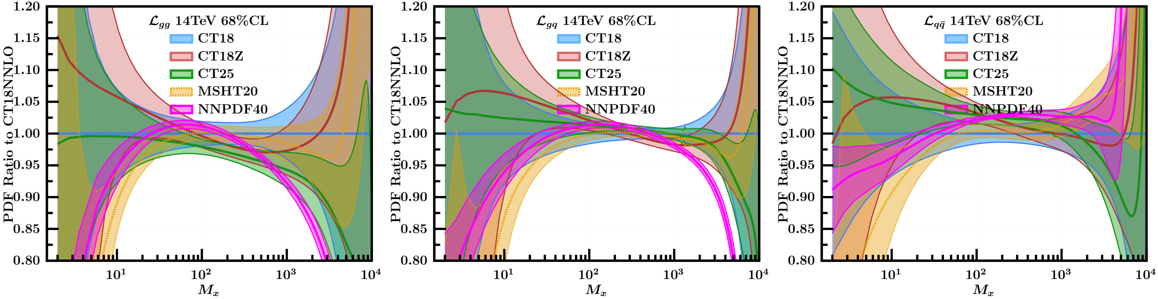






Parton–parton luminosities

- CT25 NNLO is largely consistent with:
  - MSHT20,
  - NNPDF4.0 NNLO.
- Noticeable differences remain in several luminosity channels.
- CT25 NNLO uncertainties are of similar size to MSHT20 NNLO.



Distances of CT25 PDFs to other PDF sets

(a) 2 GeV

PDF flavor	CT18	NNPDF4.0	MSHT20
$u+\bar{u}$	0.010	0.018	0.013
$d+\bar{d}$	0.022	0.050	0.039
$s+\bar{s}$	0.297	0.408	0.144
g	0.081	0.093	0.058

(b) 100 GeV

PDF flavor	CT18	NNPDF4.0	MSHT20
$u+\bar{u}$	0.009	0.023	0.012
$d+\bar{d}$	0.025	0.071	0.047
$s+\bar{s}$	0.128	0.415	0.092
g	0.053	0.119	0.081

TABLE VI: Distance measure D for CT18, NNPDF4.0 and MSHT20 with respect to CT25.

$$D = \frac{1}{x_{max} - x_{min}} \int_{x_{min}}^{x_{max}} \frac{|f_1(x) - f_2(x)|}{|f_1(x)| + |f_2(x)|}$$

(a) 2 GeV

PDF flavor	CT18	NNPDF4.0	MSHT20
$u+\bar{u}$	0.828	1.547	1.157
$d+\bar{d}$	0.545	0.681	0.629
$s+\bar{s}$	0.827	1.375	0.383
g	0.748	0.956	0.684

(b) 100 GeV

PDF flavor	CT18	NNPDF4.0	MSHT20
$u+\bar{u}$	0.733	1.875	1.073
$d+\bar{d}$	0.452	0.920	0.590
$s+\bar{s}$	0.601	1.602	0.350
g	0.680	1.180	0.850

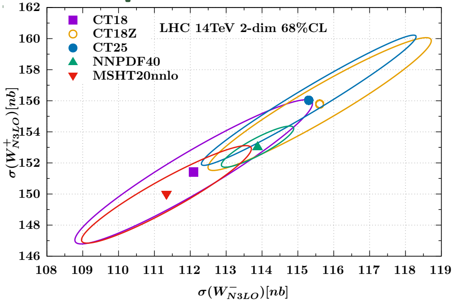
TABLE VII: Distance measure D_{err} for CT18, NNPDF4.0 and MSHT20 with respect to CT25.

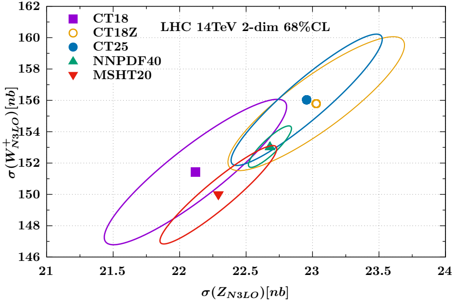
$$D_{err} = \frac{1}{x_{max} - x_{min}} \int_{x_{min}}^{x_{max}} \frac{|f_1(x) - f_2(x)|}{\sqrt{\sigma_1^2(x) + \sigma_2^2(x)}}$$

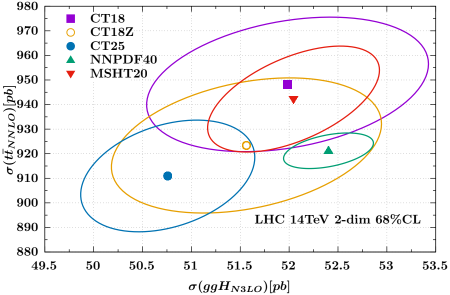
$x_{min} = 10^{-4}, x_{max} = 0.7$

Green, yellow, red: closest, intermediate, furthest among the three

CT25 predictions for LHC cross sections



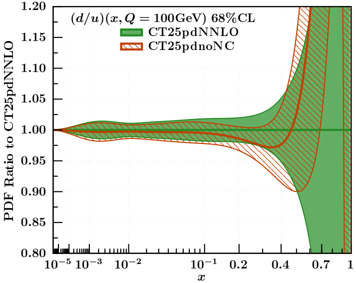


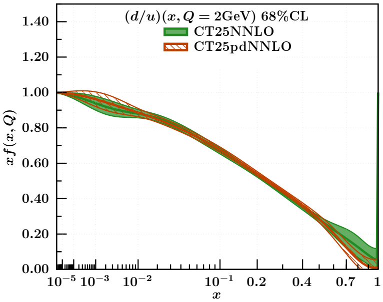


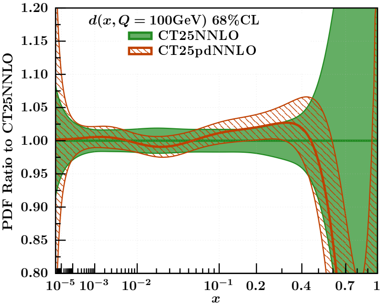
- The central of CT25 is closed to CT18
- The error of CT25 is comparable to MSHT20
- NNPDF40 error is smaller
- Uncertainty quantification is under debate: PDF4LHC benchmark

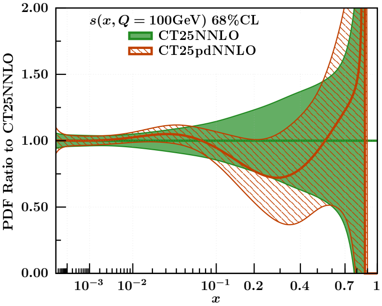
CT25pd NNLO

- CT25pd is based only on proton and deuteron data.
- Baseline for nuclear PDFs
- A nucleon-binding correction to deuteron DIS data from the CJ-CT study [2102.01107]



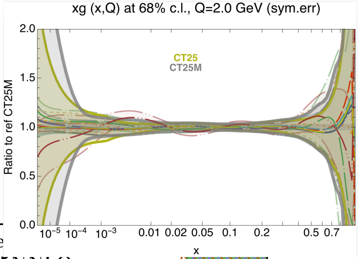


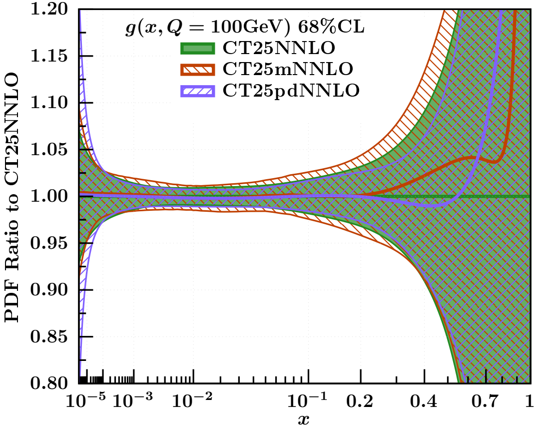


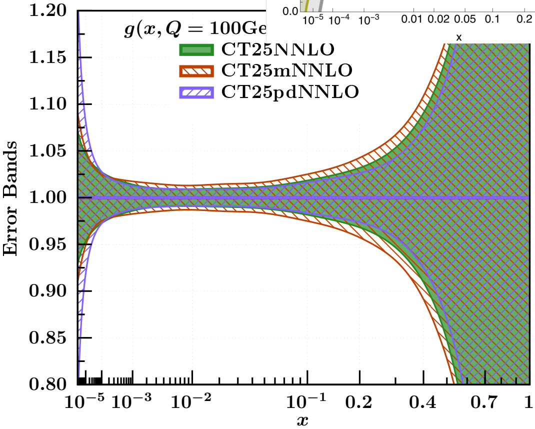


CT25m NNLO

- General-purpose symmetric Hessian set for LHC studies.
- 68% CL uncertainties.
- META combination of 8 D-TOL ensembles.
- Includes a CT18X ensemble with negative small-x gluon.
- Designed to be comparable with PDF4LHC21.

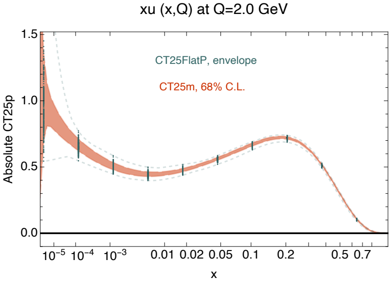


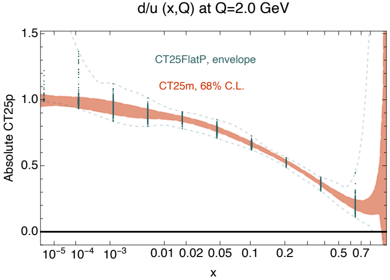


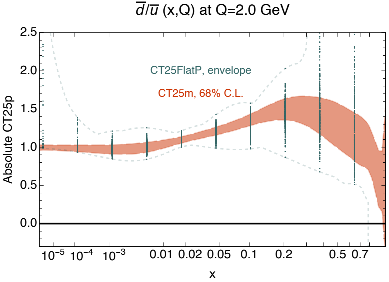


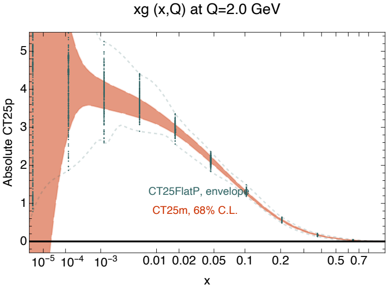
CT25FlatP NNLO

- For reweighting with LHC data
- 350 equally acceptable fits to proton-only data
- estimate aleatoric and epistemic PDF uncertainties
- Error computed as an envelope



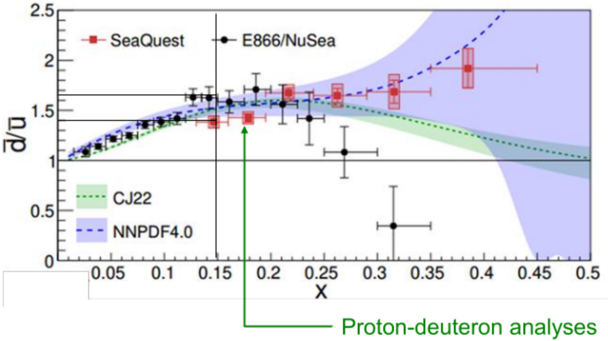






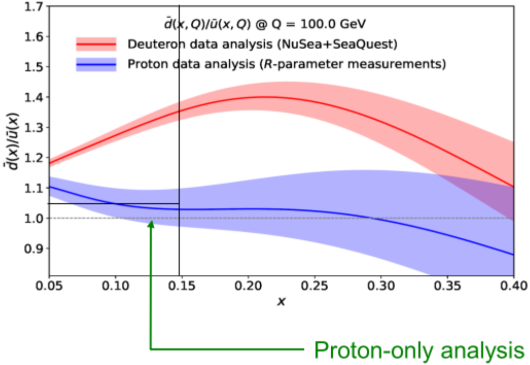
How big is the asymmetry of $\bar{d}/\bar{u}$?

Final E906/SeaQuest $\sigma_{pd}/(2\sigma_{pp})$ DY ratio [2512.17564]



Note: the E906 $\bar{d}/\bar{u}$ at $x > 0.1$, which differs considerably among PDF groups

CMS, forward-backward asymmetry of Drell-Yan [2510.08941, See Yang's talk]

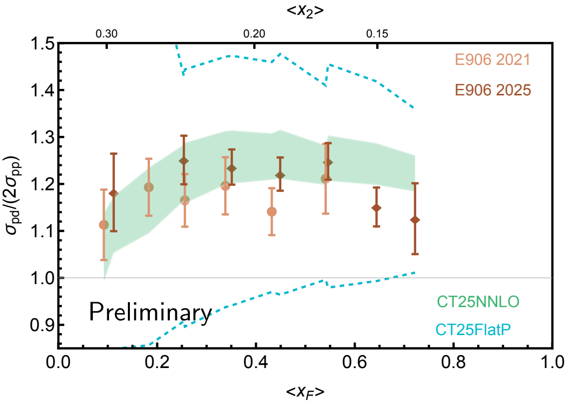


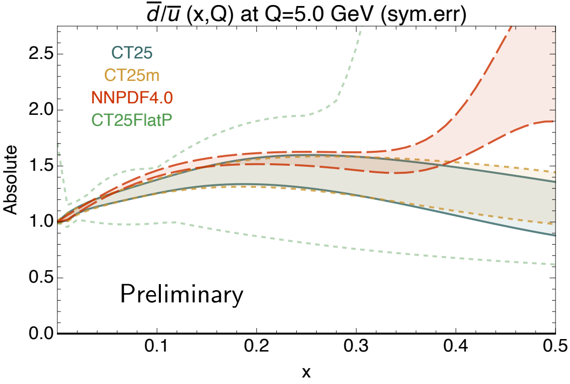
- Uncertainty bands are displayed for $\Delta\chi^2 = 1$ with one PDF parametrization choice
- LHCb is expected to provide some answers.

$\bar{d}/\bar{u}$ analysis with CT25FlatP

- CT25FlatP is not biased by E906 data.
- CT25 includes E906 (2021).
- Postdiction for E906 (2025) – excellent agreement with both E906 data sets

Combined E906 (2021 + 2025)





Intrinsic charm puzzle

Common assumption in PDF fits:

- The proton contains intrinsic up, down, strange (anti-)quarks, but no but no intrinsic charm quarks
- But it does not need to be so!

An intrinsic charm component predicted by many models.

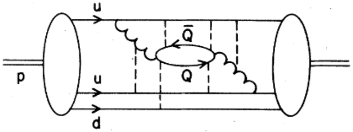
THE INTRINSIC CHARM OF THE PROTON

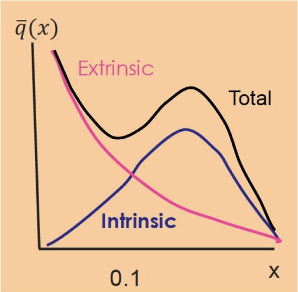
S.J. BRODSKY [1]
Stanford Linear Accelerator Center, Stanford, California 94305, USA

and

P. HOYER, C. PETERSON and N. SAKAI [2]
NORDITA, Copenhagen, Denmark

Received 22 April 1980





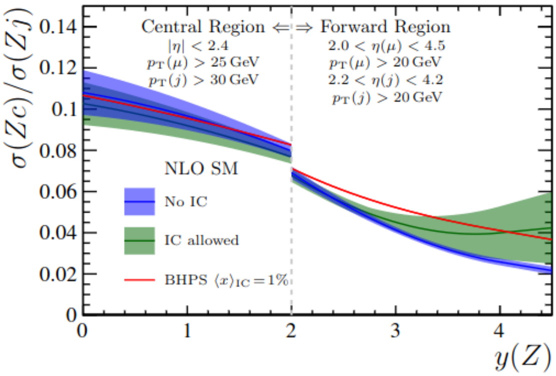
BPHS model: $|p\rangle = \mathscr{P}_{3q}|uud\rangle + \mathscr{P}_{5q}|uudc\bar{c}\rangle + \cdots$

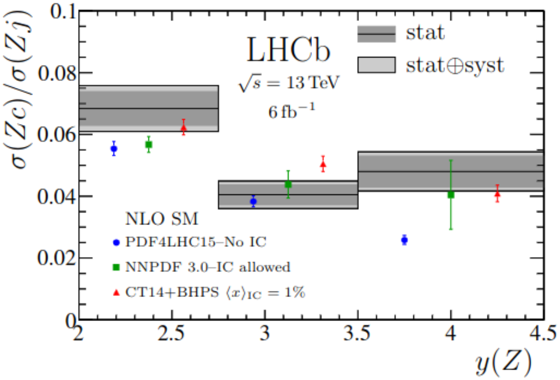
$$c^{(n_f=4)}(x, Q > m_c) \simeq c^{(n_f=4)}_{\rm pert} + c^{(n_f=4)}_{\rm intr}$$

We need to disentangle perturbative from intrinsic components.

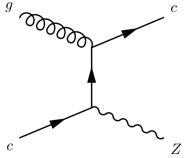
Recent data give unexpectedly large cross-sections for charmed particle production at high x_F in hadron collisions. This may imply that the proton has a non-negligible uudc$\bar{c}$ Fock component. The interesting consequences of such a hypothesis are explored.

LHCb measurement of $Z+c$ production

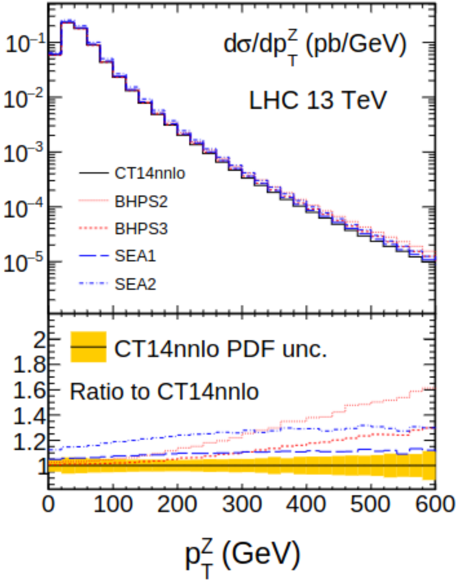


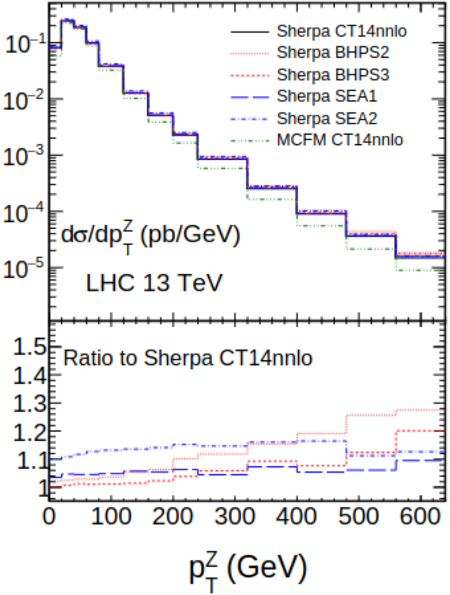


- $Z+c$ is expected to be sensitive the charm PDF.
- In the central region, the perturbative charm hide the intrinsic charm.
- The forward region, the intrinsic charm can have more chance to show up.



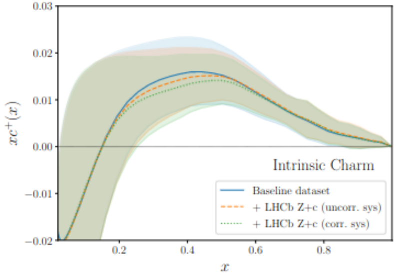
CT14IC exploration

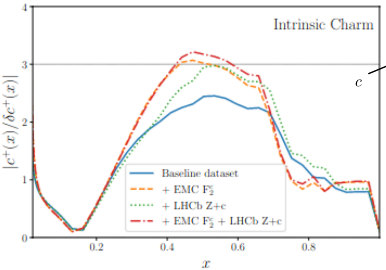


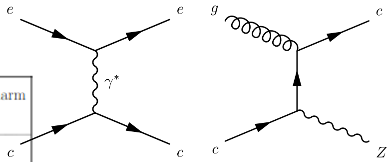


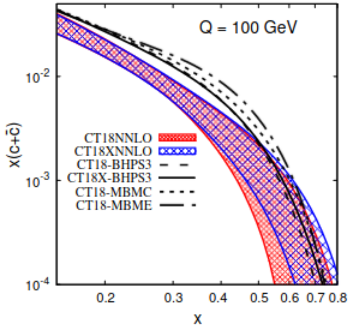
- We have pointed out the sensitivity of $Z + c$ measurement [Guzzi, KX et al., 1707.00657]
- Parton-Shower (final-state radiation) can dilute the sensitivity

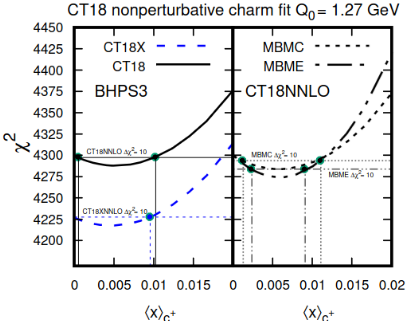
CT18 and NNPDF debates





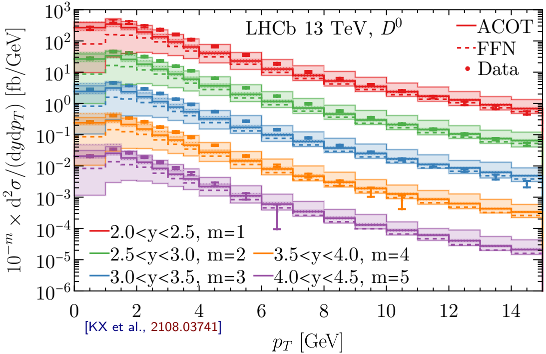




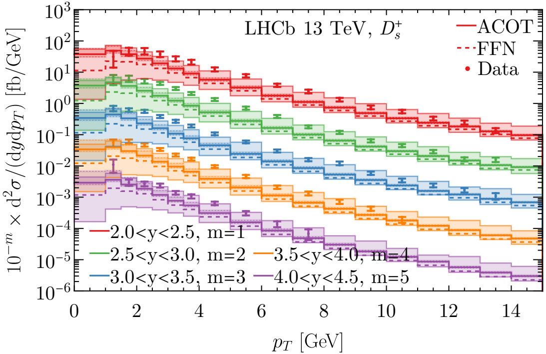


- NNPDF found a 3σ discovery of intrinsic charm PDF [Nature, 2208.08372]
- We re-do a more comprehensive IC analysis carefully, and found 1σ significance [Guzzi, Hobbs, KX et al., CT18FC, 2211.01387]
- Quantification of error bands [Nadolsky, KX et al., 2205.10444]
- FPF and Muon Colliders have a great potential

Charm production at LHCb

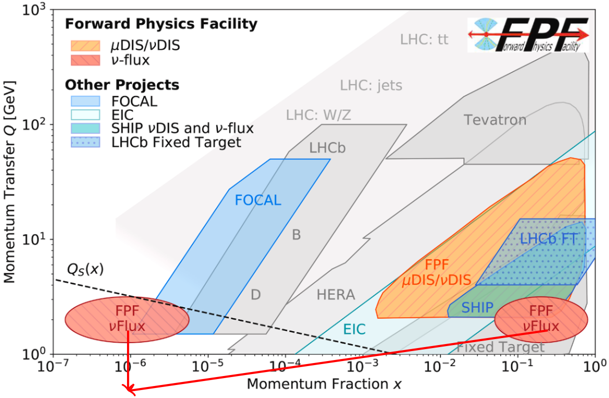


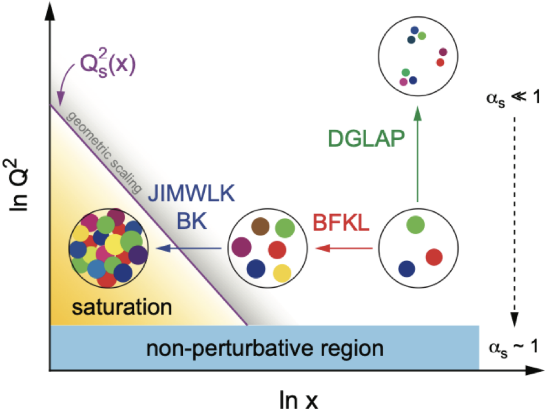
[KX et al., 2108.03741]



- Charm production in the forward region are sensitive to both small and large x charm and gluon PDFs.
- Intrinsic charm's effect mainly in the large x region.
- Both the LHCb and the FASER measurement can provide probe to the gluon at small x and intrinsic charm at large x.

PDFs at Small-x



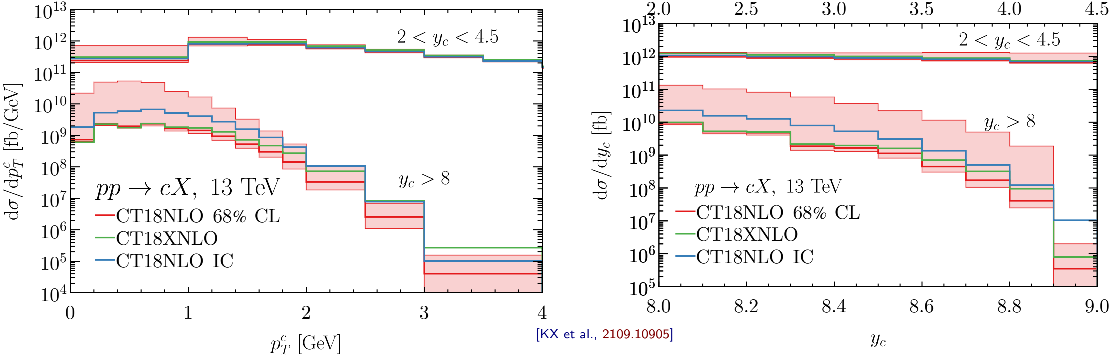


- LHCb/FPF probe parton momentum fractions at both large and small x

$$x_1 \sim \frac{m_c}{\sqrt{s}} e^{-y} \sim 10^{-5} - 10^{-6} \text{ and } x_2 \sim \frac{m}{\sqrt{s}} e^{+y} \sim 1$$

- Small-x physics: : resummation vs saturation

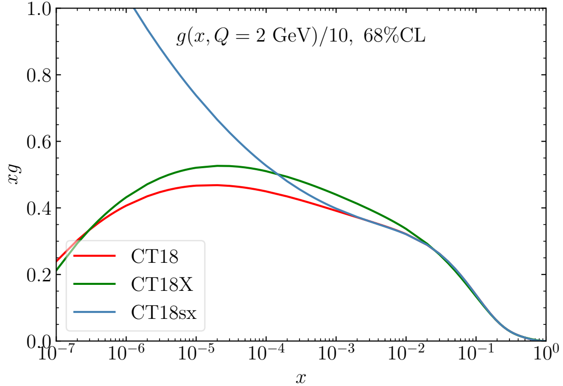
Impacts on the far forward charm production (FPF)

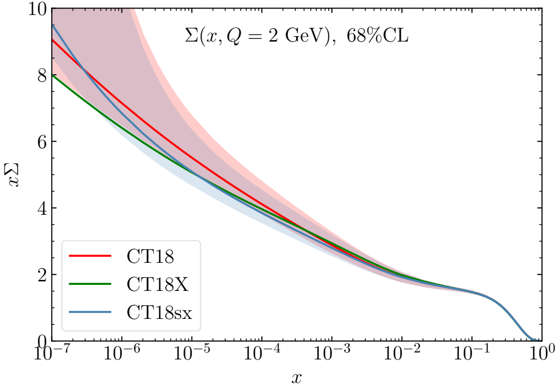


[KX et al., 2109.10905]

- Both the saturation and intrinsic charm models enhance the charm production, especially in the forward region.
- The description of the LHCb D meson data can in turn constrain the gluon and charm PDFs.
- It can be directly applied to the charmed hadron production relevant to FPF.

CT18sx vs CT18X





- Fitted data is only viable down to $x \sim 10^{-5}$
- We take the extrapolation (and evolution)
- At extremely small x, the BFKL ressummation of $\log(x)$ (CT18sx [2108.06596]) and saturation (CT18X [1912.10053]) effects.

Conclusions

- CT25 provides a major update of the CT18 global PDF analysis.
- Includes extensive new LHC precision data.
- Introduces flexible uncertainty estimation methods.
- Dynamic tolerance and Bayesian methods provide more realistic uncertainty estimates.
- 4 specialized PDF ensembles are available: CT25, CT25pd, CT25m, CT25FlatP
- Public releases include:
  - Fantômas module in xFitter: Bézier polynomial parametrizations for PDFs https://gitlab.com/cteq-tea/public/fantomas4qcd
  - CTEQ–TEA Parton Cloud: repository of best-fit PDF ensembles, cteq-tea.gitlab.io
  - DragonWell (pending): open-source pN3LO fitting code based on CTEQ-TEA methodology

FANTÔMAS4QCD

DRAGONWELL

Conclusions

- CT25 provides a major update of the CT18 global PDF analysis.
- Includes extensive new LHC precision data.
- Introduces flexible uncertainty estimation methods.
- Dynamic tolerance and Bayesian methods provide more realistic uncertainty estimates.
- 4 specialized PDF ensembles are available: CT25, CT25pd, CT25m, CT25FlatP
- Public releases include:
    - Fantômas module in xFitter: Bézier polynomial parametrizations for PDFs https://gitlab.com/cteq-tea/public/fantomas4qcd
    - CTEQ–TEA Parton Cloud: repository of best-fit PDF ensembles, cteq-tea.gitlab.io
    - DragonWell (pending): open-source pN3LO fitting code based on CTEQ-TEA methodology

FANTÔMAS4QCD

DRAGONWELL

Thanks for your attention!