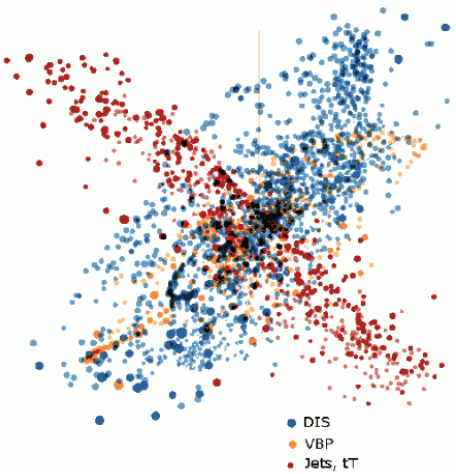
Impact of QCD and PDF uncertainties on Standard Model Measurements

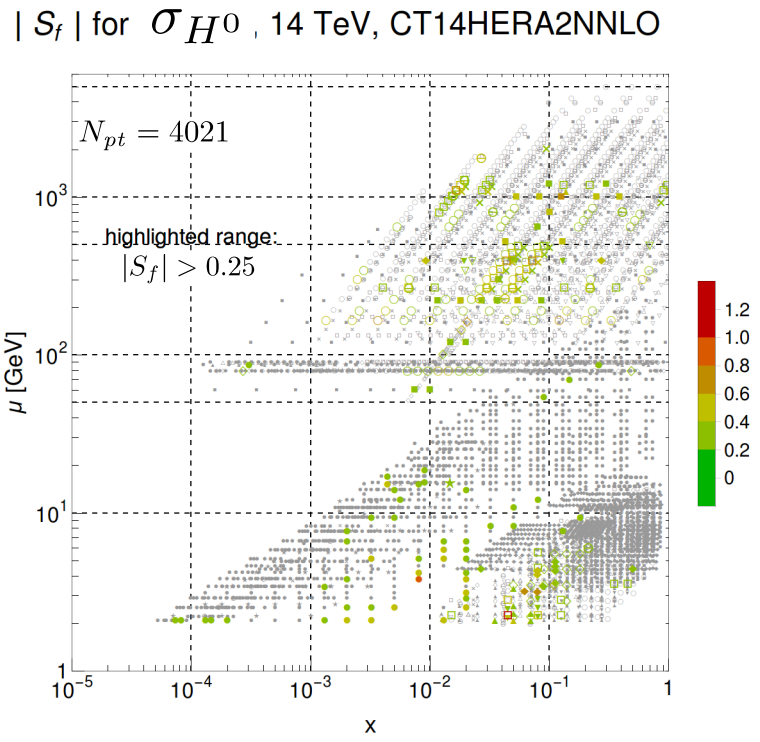
a theory/PDF analysis perspective.

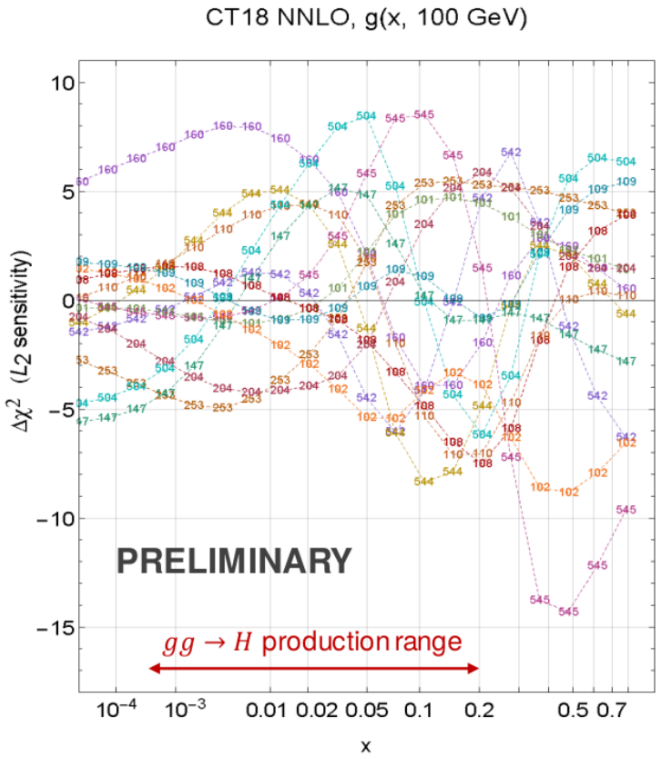
Tim Hobbs, SMU/CTEQ and JLab EIC Center

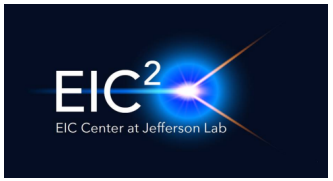
July 17th 2019

Collaboration with Bo-Ting Wang, Pavel Nadolsky, Fred Olness
...and CTEQ-TEA...



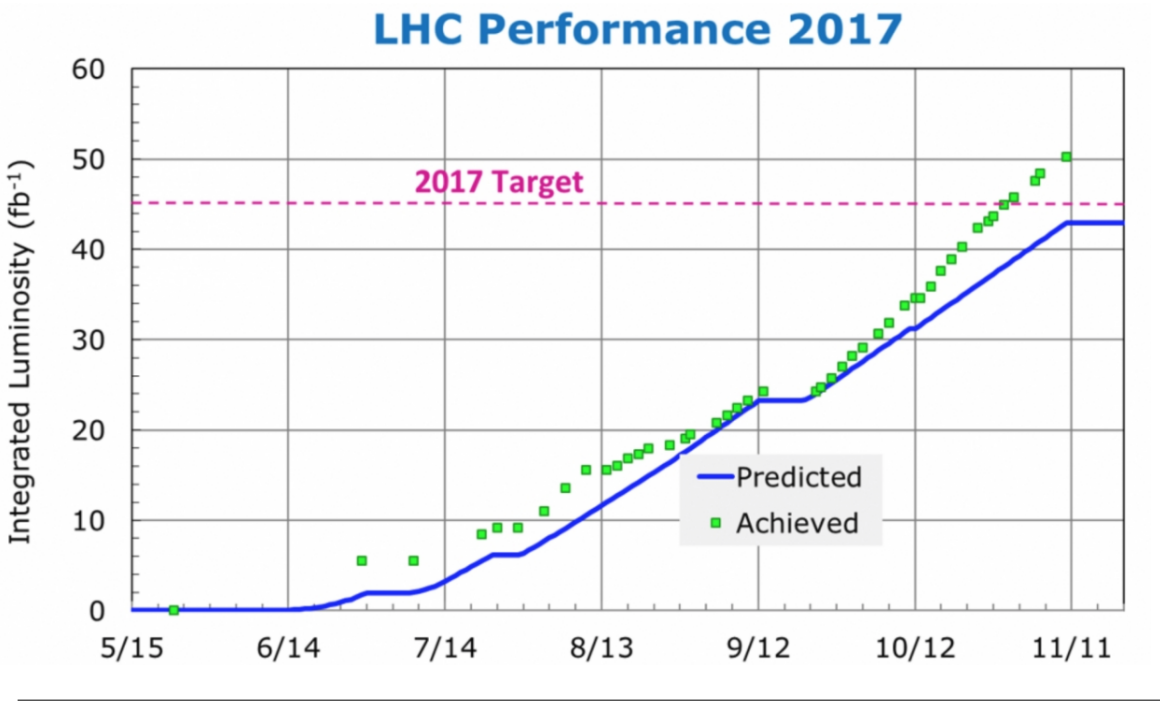






QCD@LHC '19. Buffalo, NY; July 15-19, 2019.

LHC has accumulated copious high-precision data



"data deluge..."

this data is an opportunity, but also a challenge:

→ PDF uncertainties are a **double-edged sword**

(i.e., quantities for which PDF uncertainties are large are also PDF-sensitive, and can be constraining)

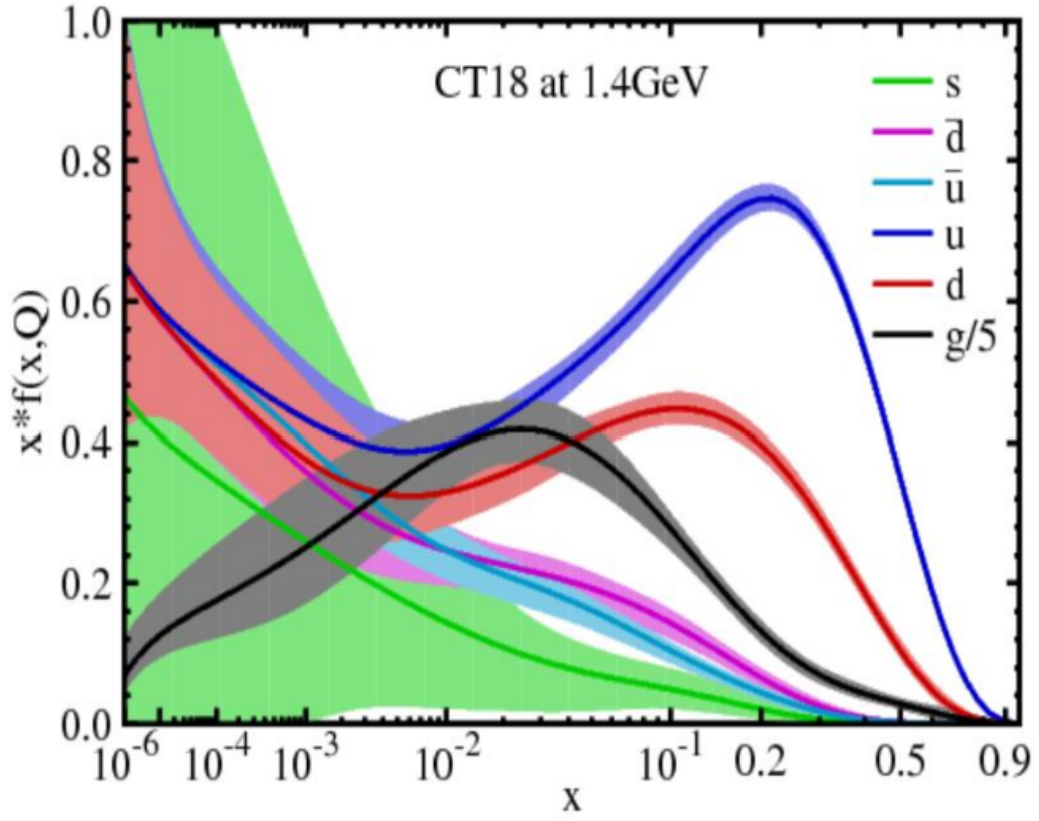
this talk (mainly): $\sin^2\theta_W$, M_W, σ_H, ...

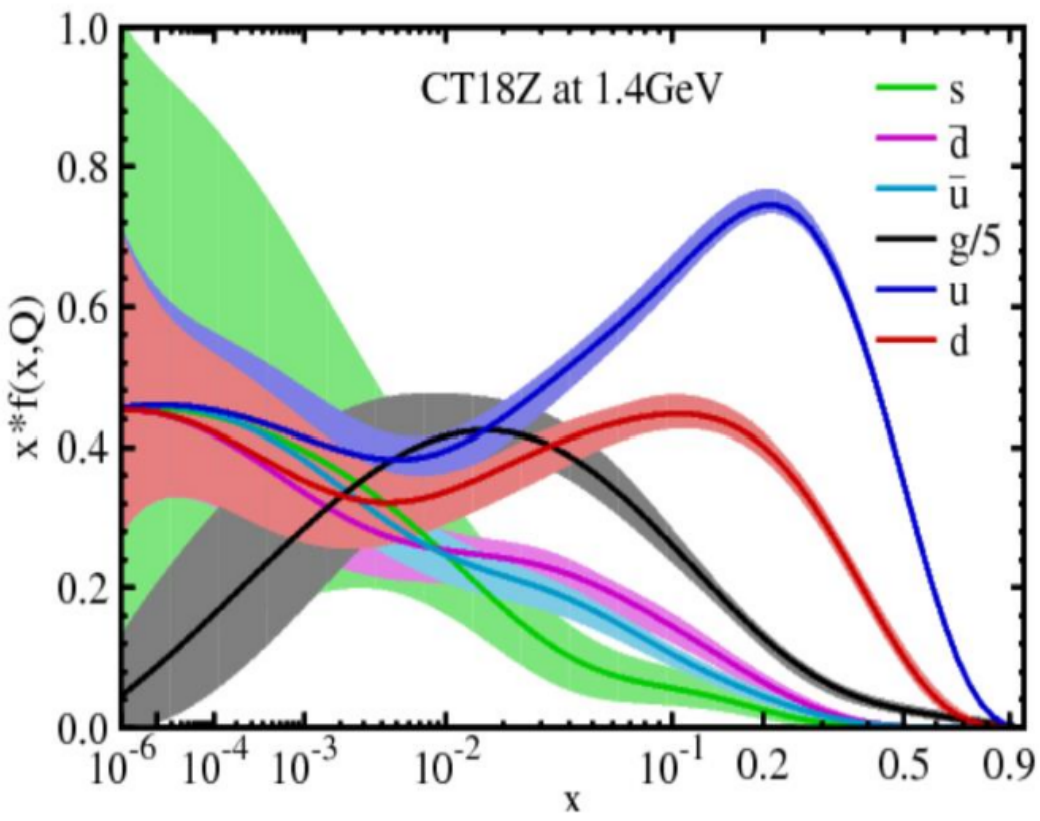
→ to reach precision objectives for HL-LHC, it will be crucial to understand pulls/tensions in modern PDF analyses

CT18 parton distributions

PDF analyses are challenging! (theoretically, computationally, statistically, …)

Four PDF ensembles: CT18 (default), A, X, and Z

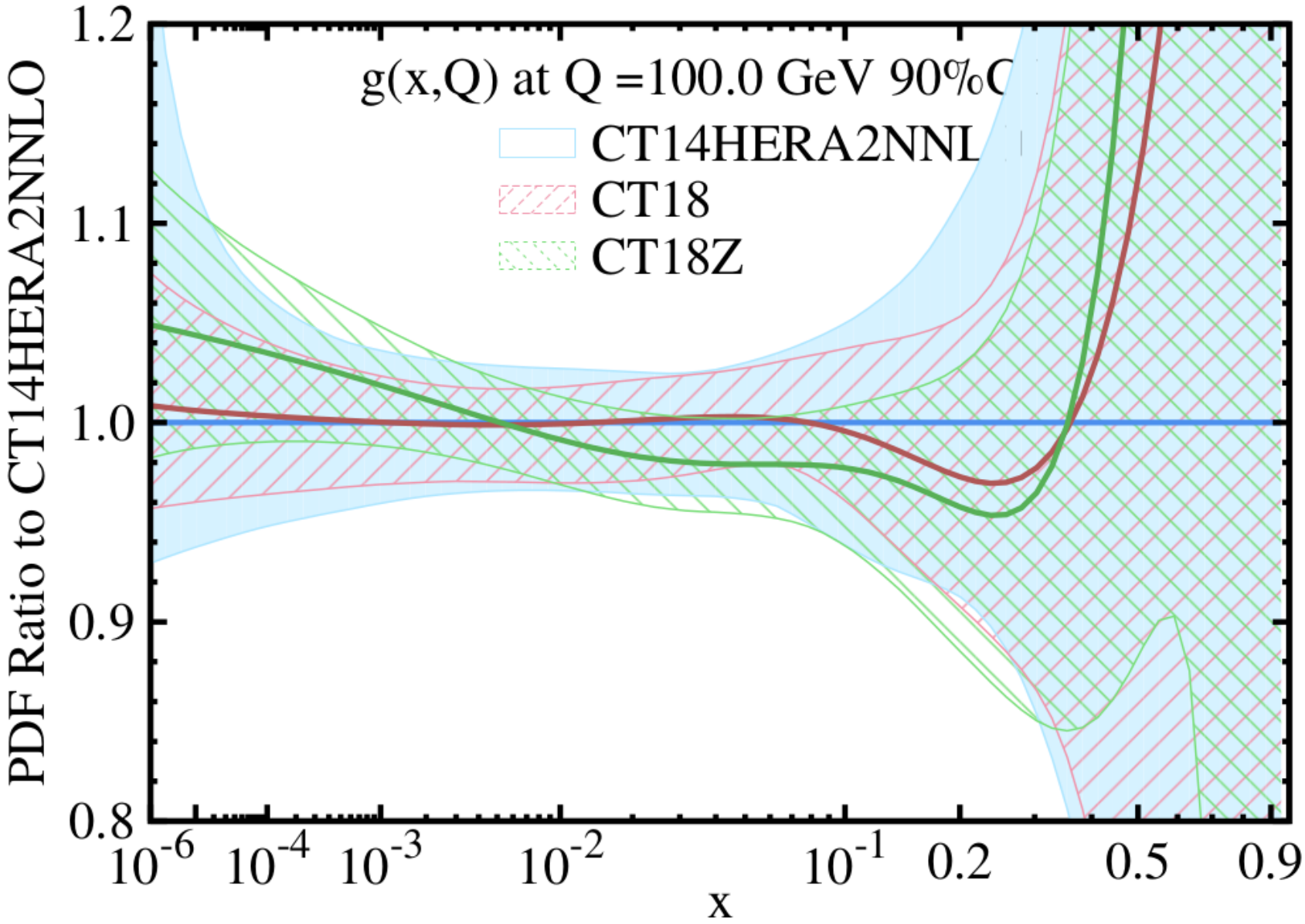




CT18Z has enhanced gluon and strange PDFs at $x \sim 10^{-4}$, and reduced light-quark PDFs at $x < 10^{-2}$. The CT18Z fit is performed so as to maximize the differences from CT18 PDFs, while preserving about the same goodness-of-fit as for CT18. CT18A and CT18X include some features of CT18Z

(See talk by P. Nadolsky, July 15th.)

understanding LHC implications for PDF fits (and vice versa)



→ LHC Run-1 data drive important PDF improvements, including for the gluon at high-, low-x

→ BUT: performing/comparing fits is insufficient for understanding PDF sensitivities

PDFSense: an impact-study framework complementary to reweighing approaches

(See ePUMP talk by C. Schmidt.)

→ **quickly** evaluate a HEP data set according to its PDF sensitivity, S_f

("**correlation 2.0**"): an easy-to-compute **figure-of-merit** for data point sensitivity to PDFs in the presence of experimental errors

→ map these sensitivities in a kinematical parameter space to quantify/visualize origin and interplay of pulls and tensions

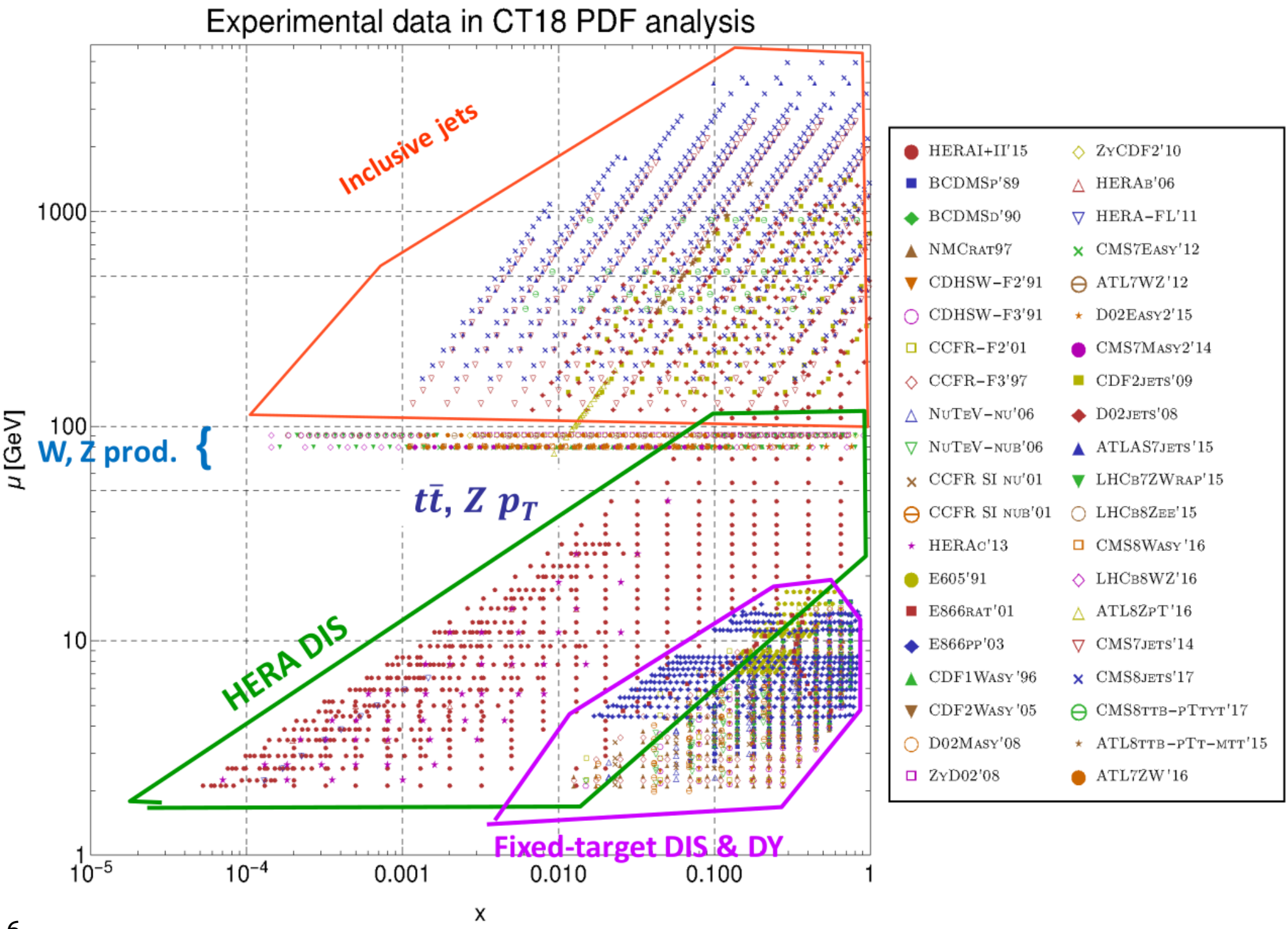
an expanding set of calculations:

TODAY

1. Mapping the sensitivity of hadronic experiments to nucleon structure
B.-T. Wang, T.J. Hobbs, S. Doyle, J. Gao, T.-J. Hou, P. M. Nadolsky, F. I. Olness
Phys.Rev. D98 (2018) 094030

2. The coming synergy between lattice QCD and high-energy phenomenology
T.J. Hobbs, Bo-Ting Wang, Pavel Nadolsky, Fredrick Olness
arXiv:1904.00222

3. Mapping the PDF sensitivity of future facilities (HL-LHC, LHeC, and EIC)
T.J. Hobbs, Bo-Ting Wang, Pavel Nadolsky, Fredrick Olness
arxiv:1907.00988
(more detailed work in preparation)


Experimental data in CT18 PDF analysis
Inclusive jets
W, Z prod.
$t\bar{t}$, Z p_T
HERA DIS
Fixed-target DIS & DY
μ [GeV]
x
1000
100
10
1
10^{-5}
10^{-4}
0.001
0.010
0.100
1
HERAI+II'15
BCDMSP'89
BCDMSD'90
NMCRAT97
CDHSW-F2'91
CDHSW-F3'91
CCFR-F2'01
CCFR-F3'97
NUTEV-NU'06
NUTEV-NUB'06
CCFR SI NU'01
CCFR SI NUB'01
HERAC'13
E605'91
E866RAT'01
E866PP'03
CDF1WASY'96
CDF2WASY'05
D02MASY'08
ZYD02'08
ZYCDF2'10
HERAB'06
HERA-FL'11
CMS7EASY'12
ATL7WZ'12
D02EASY2'15
CMS7MASY2'14
CDF2JETS'09
D02JETS'08
ATLAS7JETS'15
LHCB7ZWRAP'15
LHCB8ZEE'15
CMS8WASY'16
LHCB8WZ'16
ATL8ZPT'16
CMS7JETS'14
CMS8JETS'17
CMS8TTB-PTTYT'17
ATL8TTB-PTT-MTT'15
ATL7ZW'16

1. New LHC datasets for CT18

1. 245 1505.07024 LHCb Z (W) muon rapidity at 7 TeV(applgrid)
2. 246 1503.00963 LHCb 8 TeV Z rapidity (applgrid);
3. 249 1603.01803 CMS W lepton asymmetry at 8 TeV (applgrid)
4. 250 1511.08039 LHCb Z (W) muon rapidity at 8 TeV(applgrid)
5. 253 1512.02192 ATLAS 7 TeV Z pT (applgrid)
6. 542 1406.0324 CMS incl. jet at 7 TeV with R=0.7 (fastNLO)
7. 544 1410.8857 ATLAS incl. jet at 7 TeV with R=0.6 (applgrid)
8. 545 1609.05331 CMS incl. jet at 8 TeV with R=0.7 (fastNLO)
9. 565 1511.04716 ATLAS 8 TeV tT pT diff. distributions (fastNNLO)
10. 567 1511.04716 ATLAS 8 TeV tT mtT diff. distributions (fastNNLO)
11. 573 1703.01630 CMS 8 TeV tT (pT , yt) double diff. distributions (fastNNLO)

12. 248 1612.03016 ATLAS 7 TeV Z and W rapidity (applgrid)->CT18Z
    - also uses a special small-x factorization scale, charm mass m_c=1.4 GeV
    - serious changes in PDFs, so warrants a separate PDF

`PDFSense` guided the selection and implementation of these data for CT18

how is this done? … examine and map the sensitivities based on knowledge of Hessian **PDFs** and **residual** variations

- *idea* : study the **statistical correlation** between PDFs and the quality of the fit at a measured data point(s); fit quality encoded in a (**T**heory) – (shifted **D**ata) ***r****esidual* :

$$r_i(\vec{a}) = \frac{1}{s_i}\left(T_i(\vec{a}) - D_{i,sh}(\vec{a})\right)$$

s_i : uncorrelated uncert. | $\vec{a}$: PDF parameters

- for each data point, calculate the correlation (and sensitivity) between the residual and PDFs based on Hessian variation over PDF replicas,

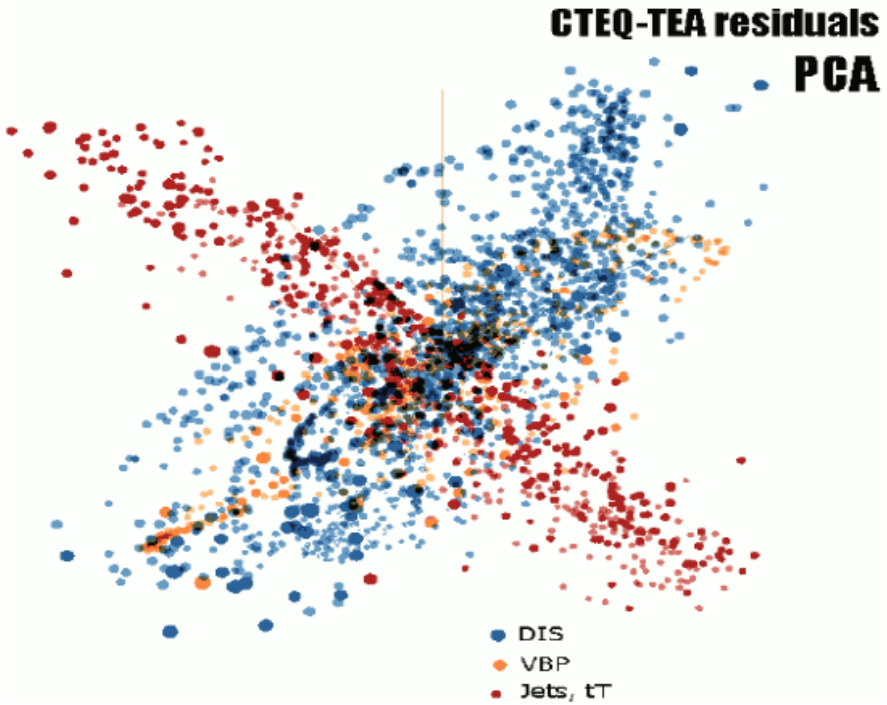
$$C_f \equiv \mathrm{Corr}[f(x_i,\mu_i), r_i(x_i,\mu_i)] \longrightarrow S_f \equiv \frac{\Delta r_i}{\langle r_0\rangle_E} C_f$$

- map these in kinematical parameter space where, e.g., for DIS,

$$x_i \approx x_{Bi} \qquad \mu_i \approx Q|_i$$

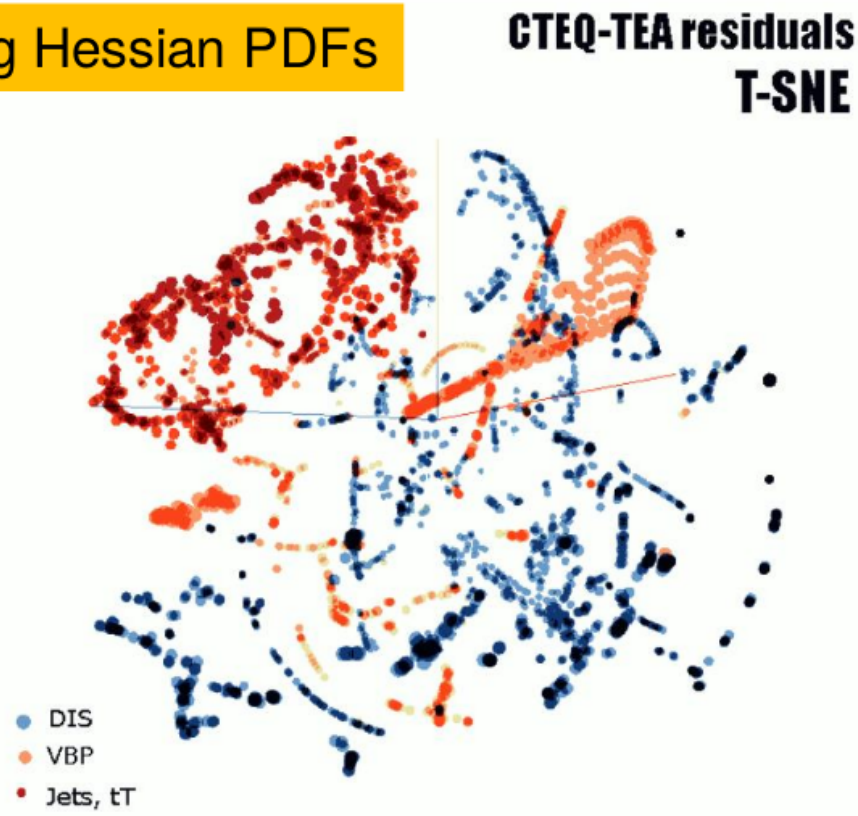
Vectors of data point residuals…

… carry detailed information about sensitivity of individual experimental data points to PDFs; can be studied using statistical packages (TensorFlow, Mathematica,…)



Principal Component Analysis (PCA) visualizes the 56-dim. manifold by reducing it to 10 dimensions (à la META PDFs)

Using Hessian PDFs



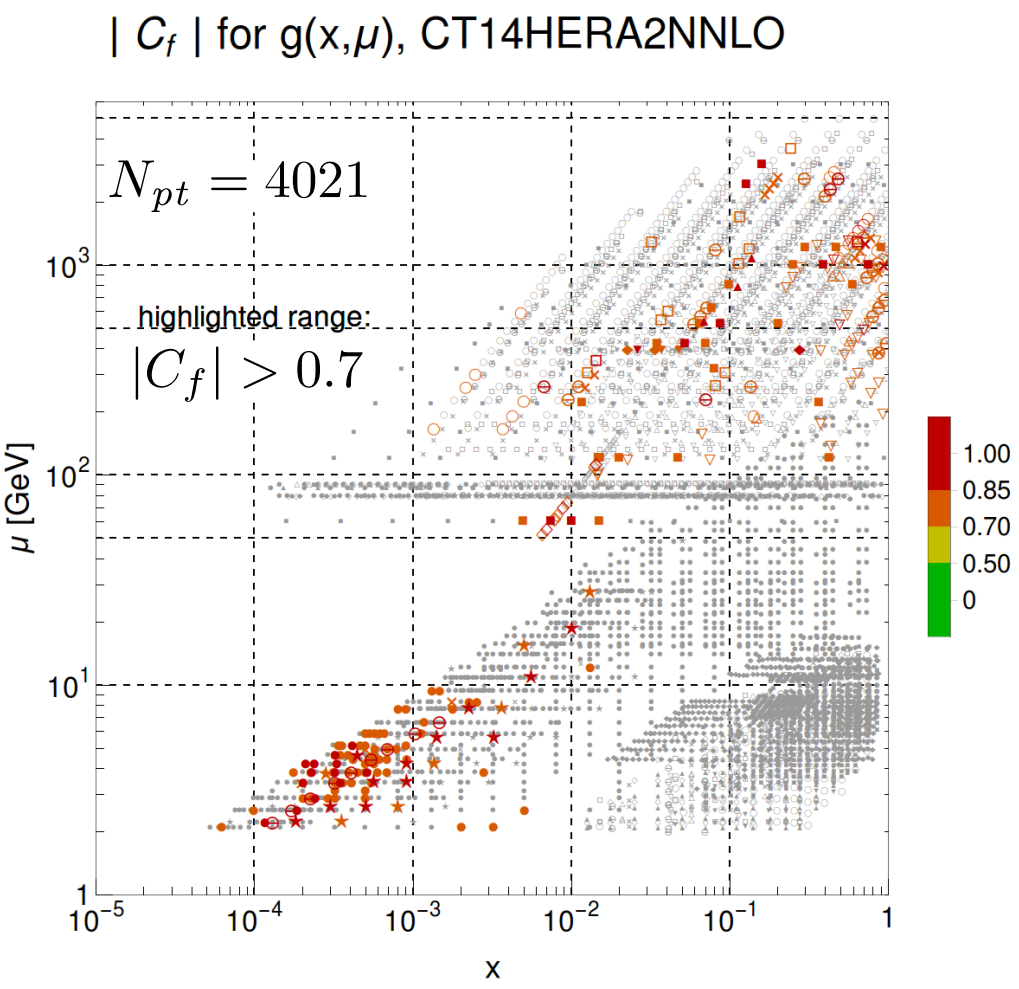
Similarly, t-distributed Stochastic Neighbor embedding (t-SNE) distinguishes PDF pulls by process

(See talk by J. Thaler, July 15th.)

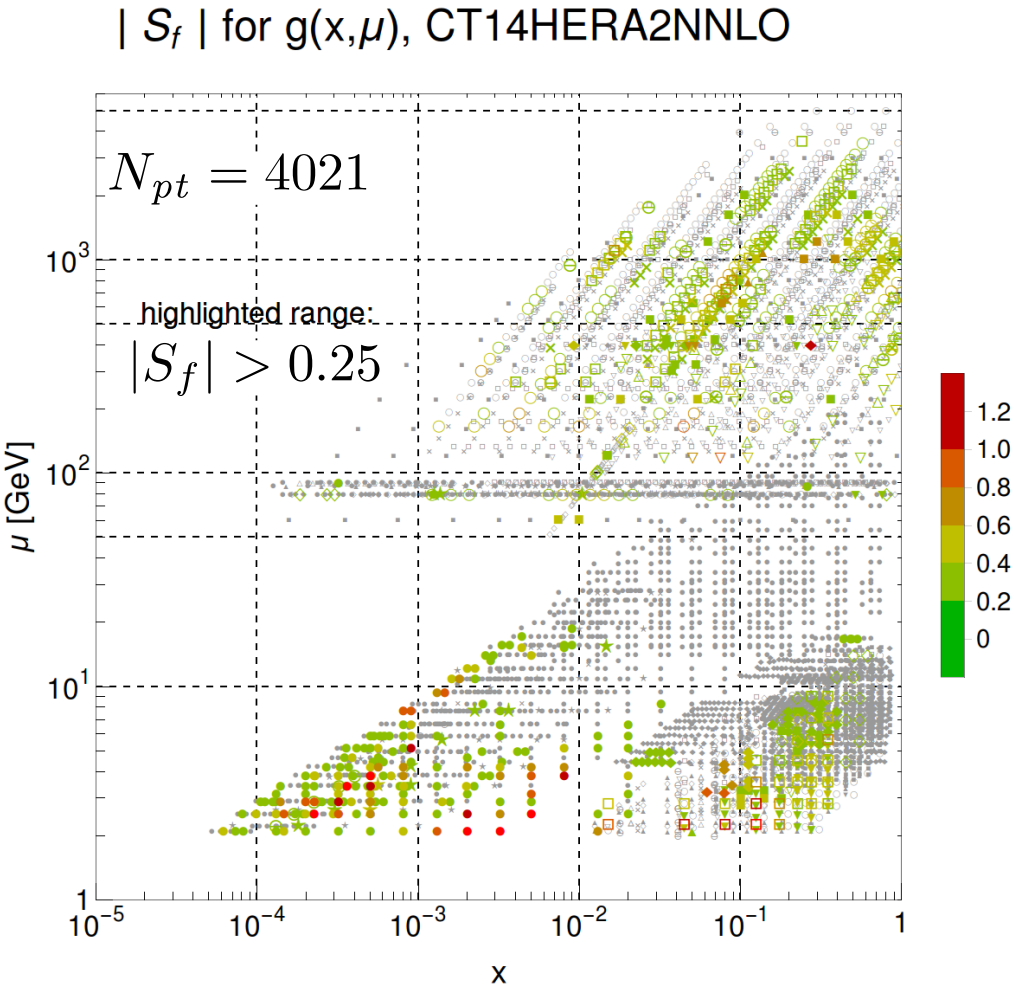
the sensitivity reveals a richer landscape than the correlation!

e.g., the substantial impact of inclusive jet data on gluon PDF is identified by S_f

$C_g(x_i, \mu_i)$



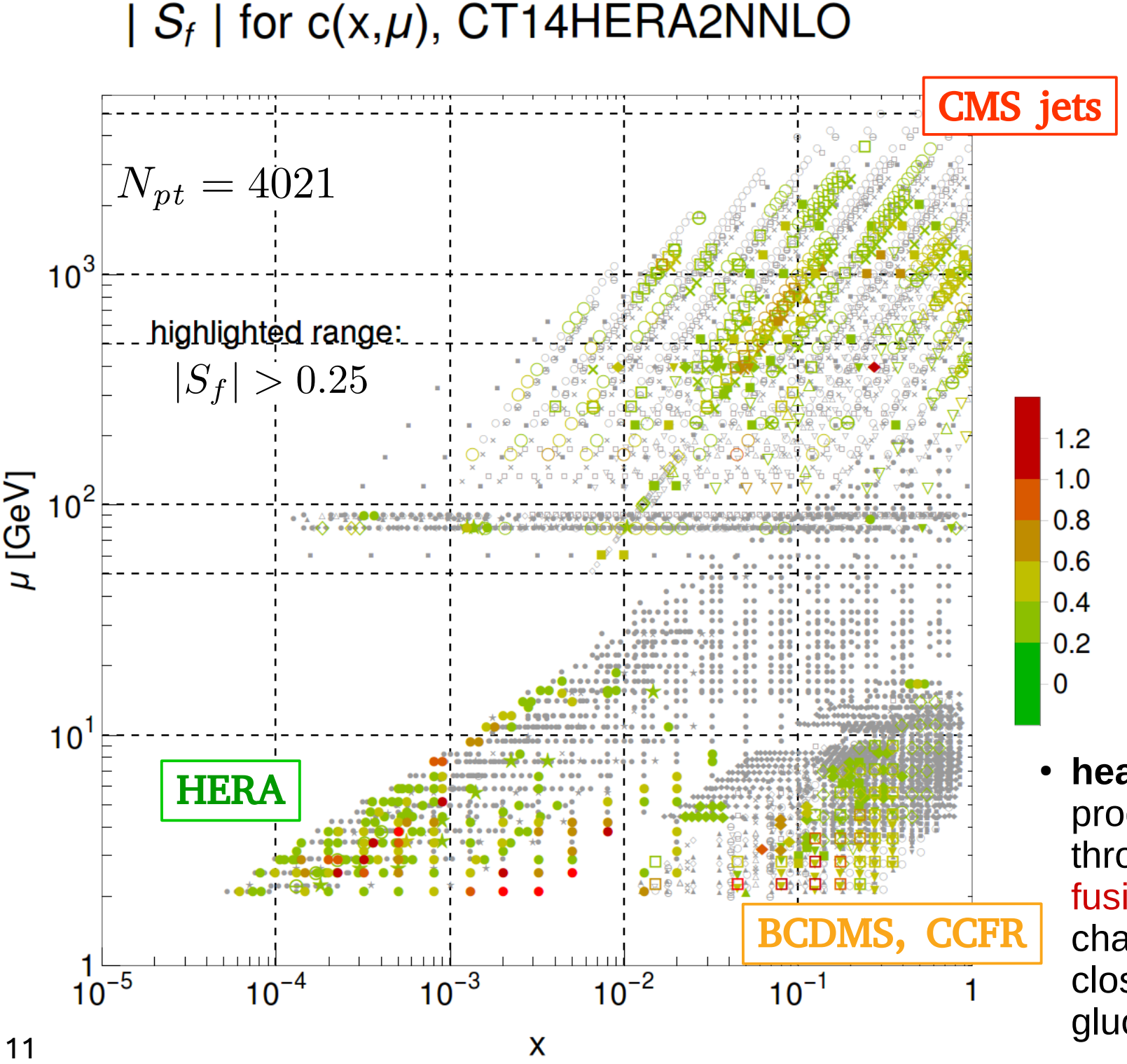
$S_g(x_i, \mu_i)$

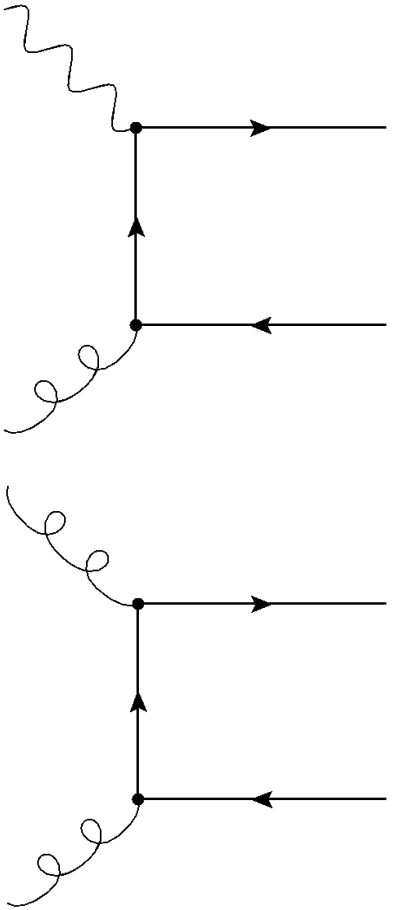


sensitivity maps reveal:

(i) highest impact data points for specific PDFs/observables
(ii) tensions/complementarities among data sets
(iii) physical origin of PDF constraints

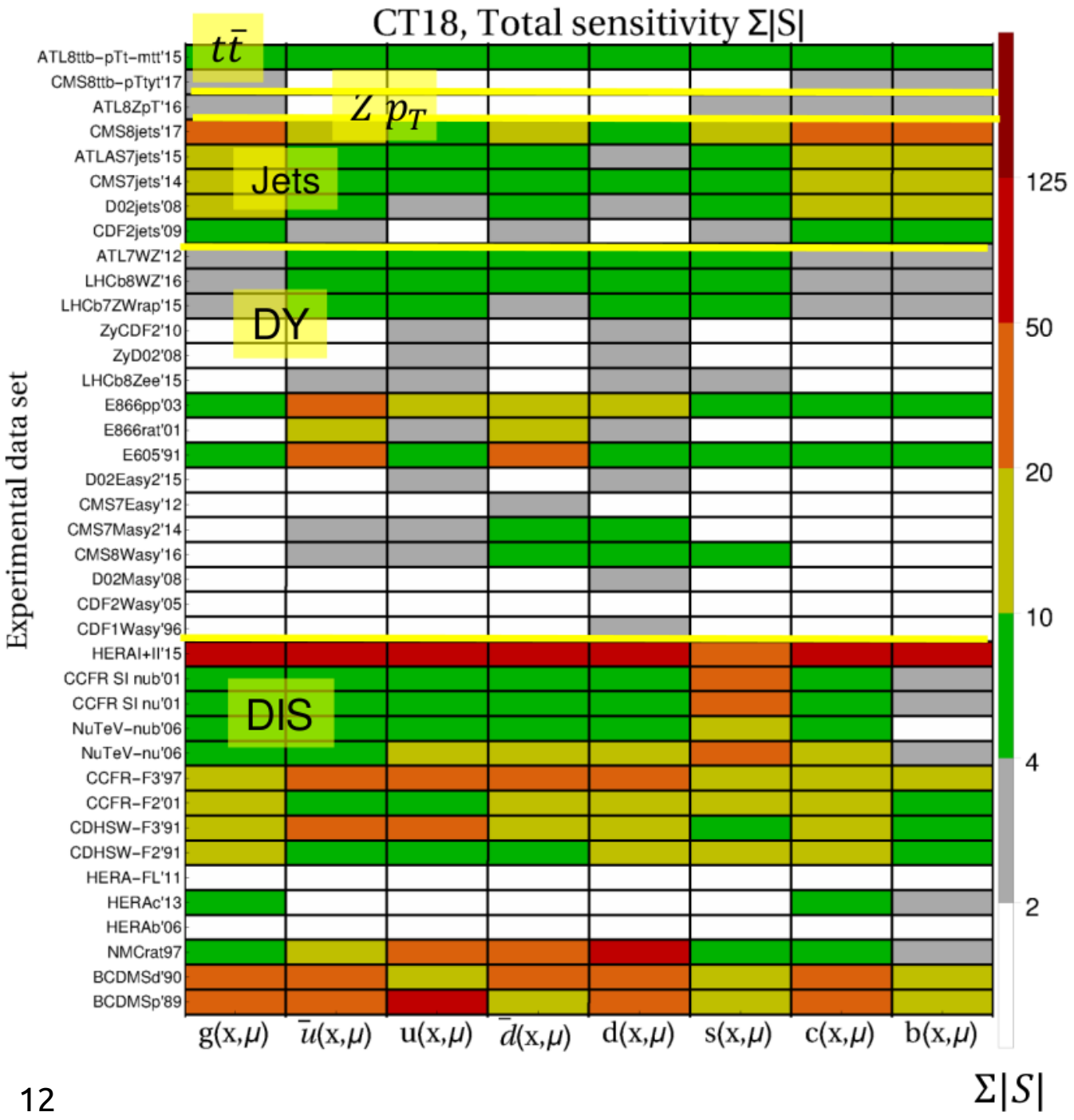
$|S_f|$ for c(x,μ), CT14HERA2NNLO





- **heavy quark** production proceeds through boson fusion diagrams: charm sensitivities closely track the gluon plot

Sensitivity of hadronic experiments to PDFs



Total sensitivity to $f_a(x_i, \mu_i)$, **summed over data points**

$$\sum_{\boldsymbol{points}} |\boldsymbol{S_{f,i}}|$$

Computed using the **PDFSense** code [arXiv:1803.02777]

we use this technology to examine
implications for HEP phenomenology at LHC

esp., EW pheno., Higgs physics, g(x), …

→ Mapping the sensitivity of hadronic experiments to nucleon structure
B.-T. Wang, T.J. Hobbs, S. Doyle, J. Gao, T.-J. Hou, P. M. Nadolsky, F. I. Olness
Phys.Rev. D98 (2018) 094030

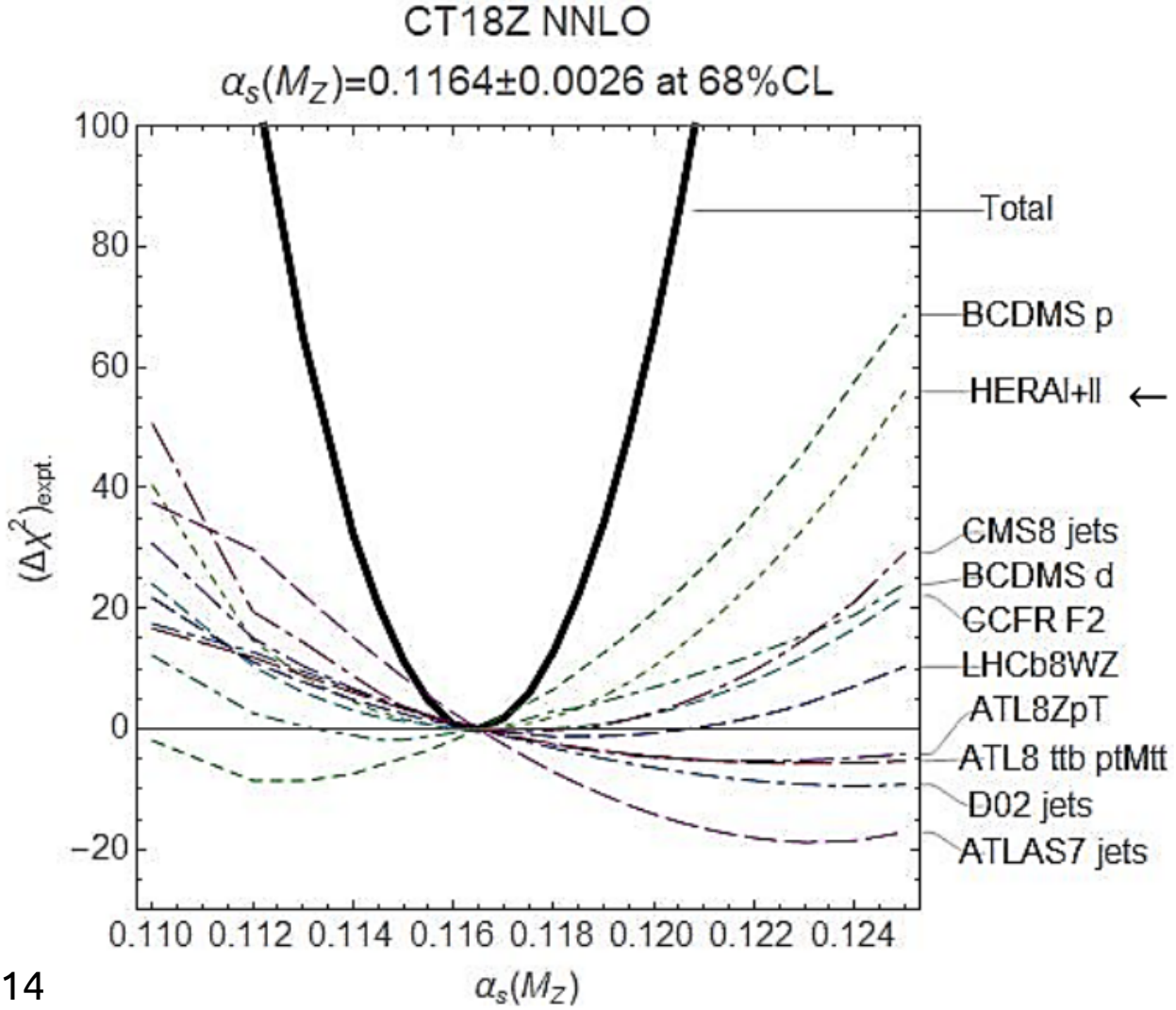
L2 sensitivity details in

The coming synergy between lattice QCD and high-energy phenomenology
T.J. Hobbs, Bo-Ting Wang, Pavel Nadolsky, Fredrick Olness
arXiv:1904.00222v2 (coming soon)

Lagrange Multiplier (LM) Scans: $\alpha_s(M_Z)$

The LM scan technique is introduced in **Stump et al., Phys.Rev. D65 (2001) 014012**

☺ **Detailed dependence of χ^2** ☹ **slow; refitting on a supercomputing cluster**



← flatter than in CT18

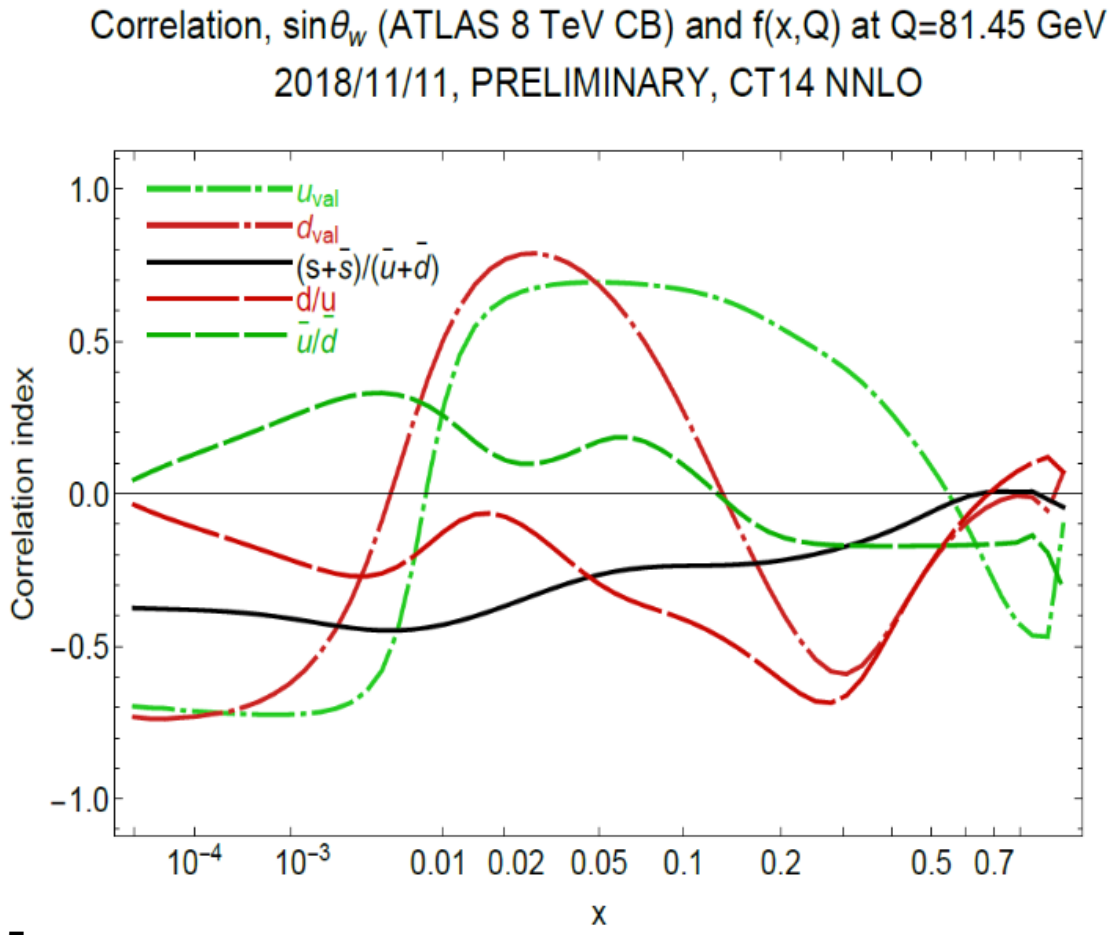
$\alpha_s(m_Z)$ from global fit closer to 0.117 than to 0.118

$\sin^2\theta_W$ (and, eventually, M_W)

...as a follow-on to Alesandro's EW-focused overview:

important PDF correlations for the ATLAS extraction of $\sin^2\theta_W$

Example: $\sin^2\theta_{weak} \equiv s2w$ measured by ATLAS 8 TeV



Strongest correlations of s2w with u_{val}, d_{val} at $0.005 \lesssim x \lesssim 0.2$

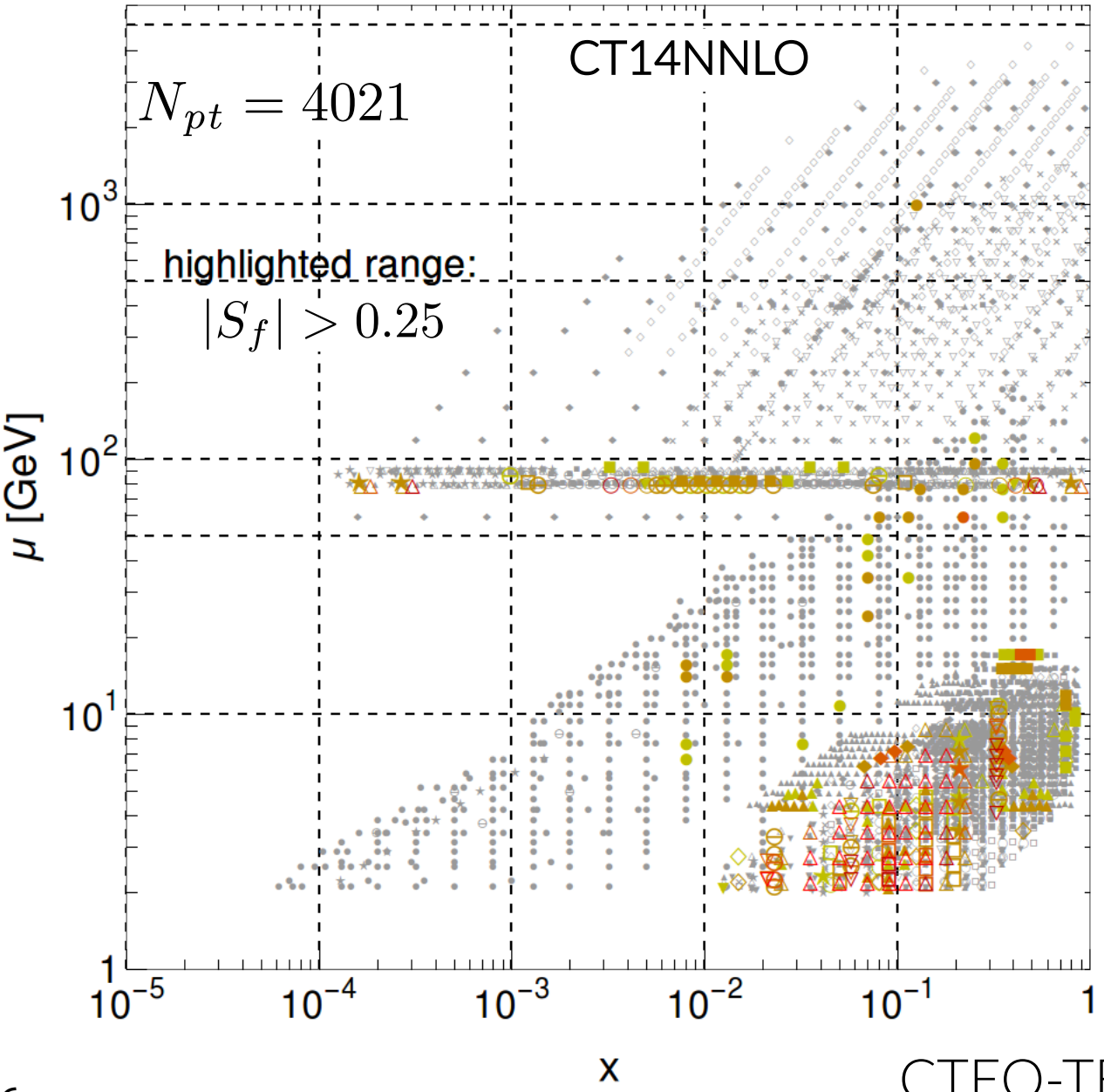
weak correlations with $\bar{u}$, $\bar{d}$, $\bar{s}$, g

u_{val}, d_{val} changed between CT10 and CT14 **[1506.07433, Sec. 2B]**

It is instructive to explore the data pulls on u_{val}, d_{val}

$\sin^2\theta_W$

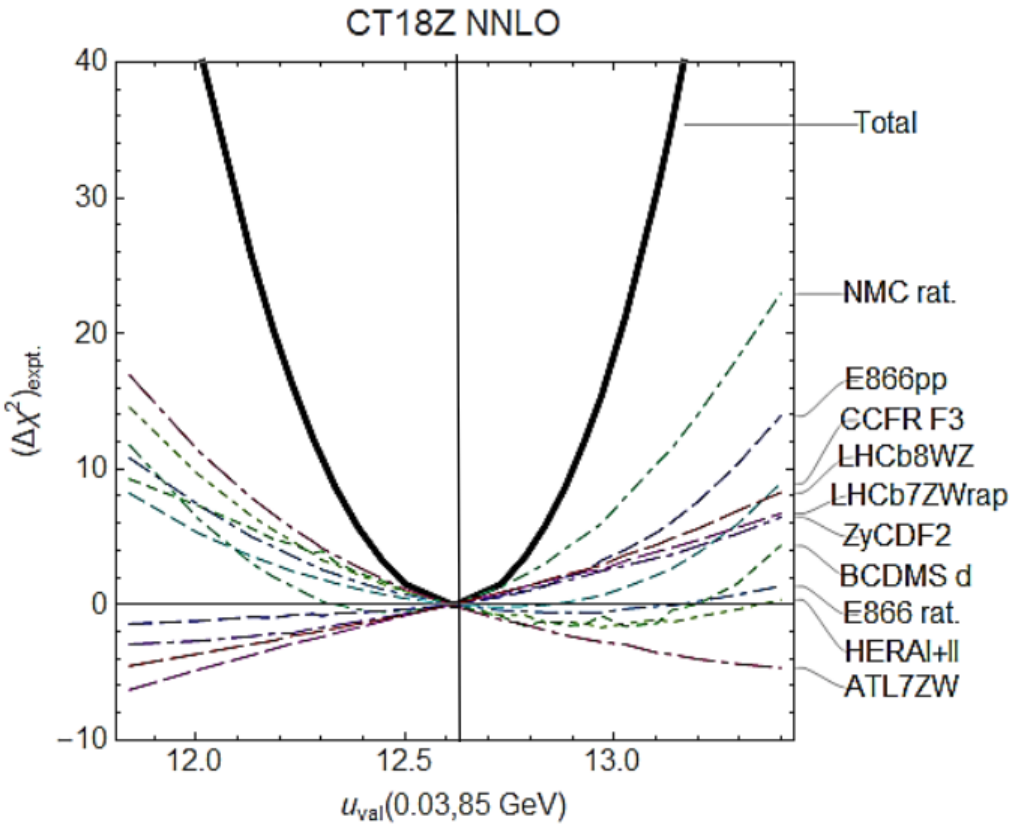
PDF sensitivity of $\sin^2\theta_W$ from 7 TeV ATLAS data



- **combined HERA1 DIS [most sensitive]**
- CCFR νp DIS $F_{3,2}$
- BCDMS $F_2^{p,d}$
- NMC ep, ed DIS
- CDHSW νA DIS
- NuTeV $\nu A \to \mu\mu X$
- CCFR $\nu A \to \mu\mu X$
- E866 $pp \to \ell^+\ell^- X$
- ATLAS 7 TeV W/Z (35 pb^{-1})

CTEQ-TEA sensitivities to $\sin^2\theta_W$

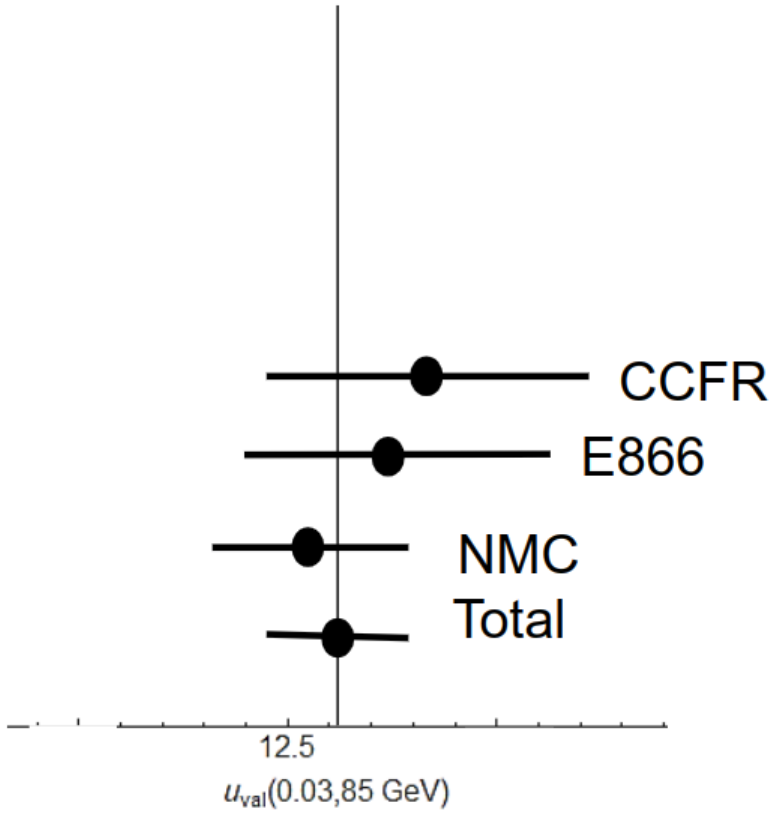
… using a Lagrange Multiplier scan…



PRELIMINARY

… or using residuals for replicas

[errors and correlations; most replicas are not good fits]

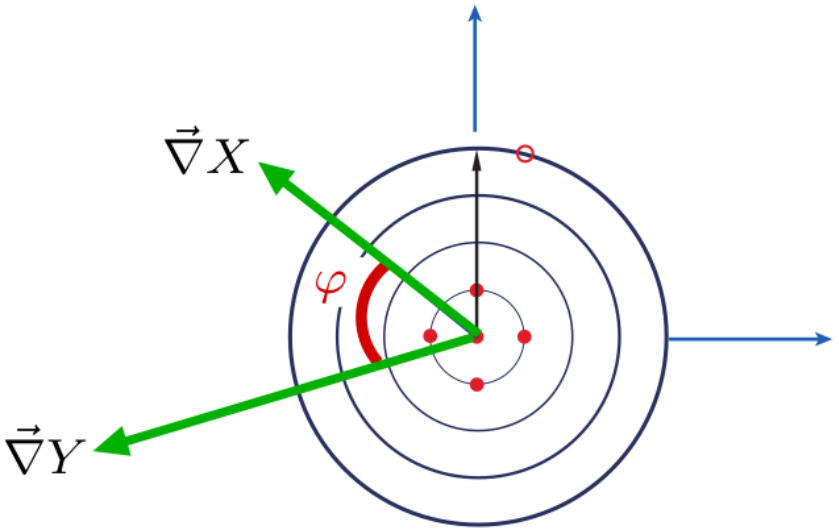


Can repeat for s2w, M_W,…

rather than the costly LM scans, we can examine a "cheaper" measure which yields comparable information

---

the *L2 sensitivity*



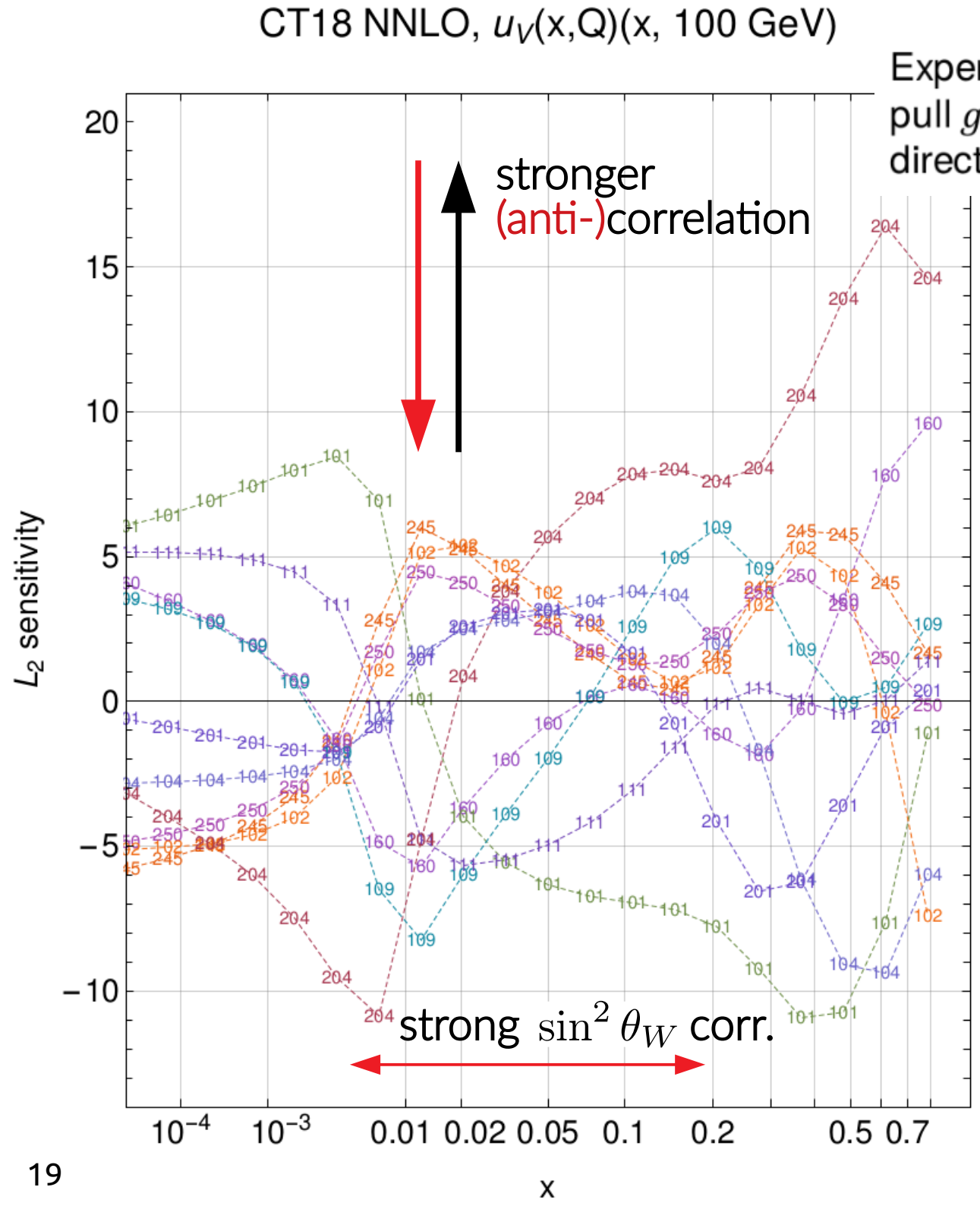
L_2 sensitivity. Take $X = f_a(x_i, Q_i)$ or $\sigma(f)$; $Y = \chi^2_E$ for experiment E. Find $\Delta Y(\vec{z}_{m,X})$ for the displacement $|\vec{z}_{m,X}| = 1$ along the direction $\vec{\nabla}X / \left|\vec{\nabla}X\right|$ (corresponding to $\Delta\chi^2_{tot} = T^2$ and $X(\vec{z}) = X(0) + \Delta X$):

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

$$= \Delta Y \cos\varphi$$

or, $\sim \mathrm{Corr}[f_a, \chi^2_E]$

...extent to which total χ^2_E of specific expts. correlates with x-dep. of PDFs

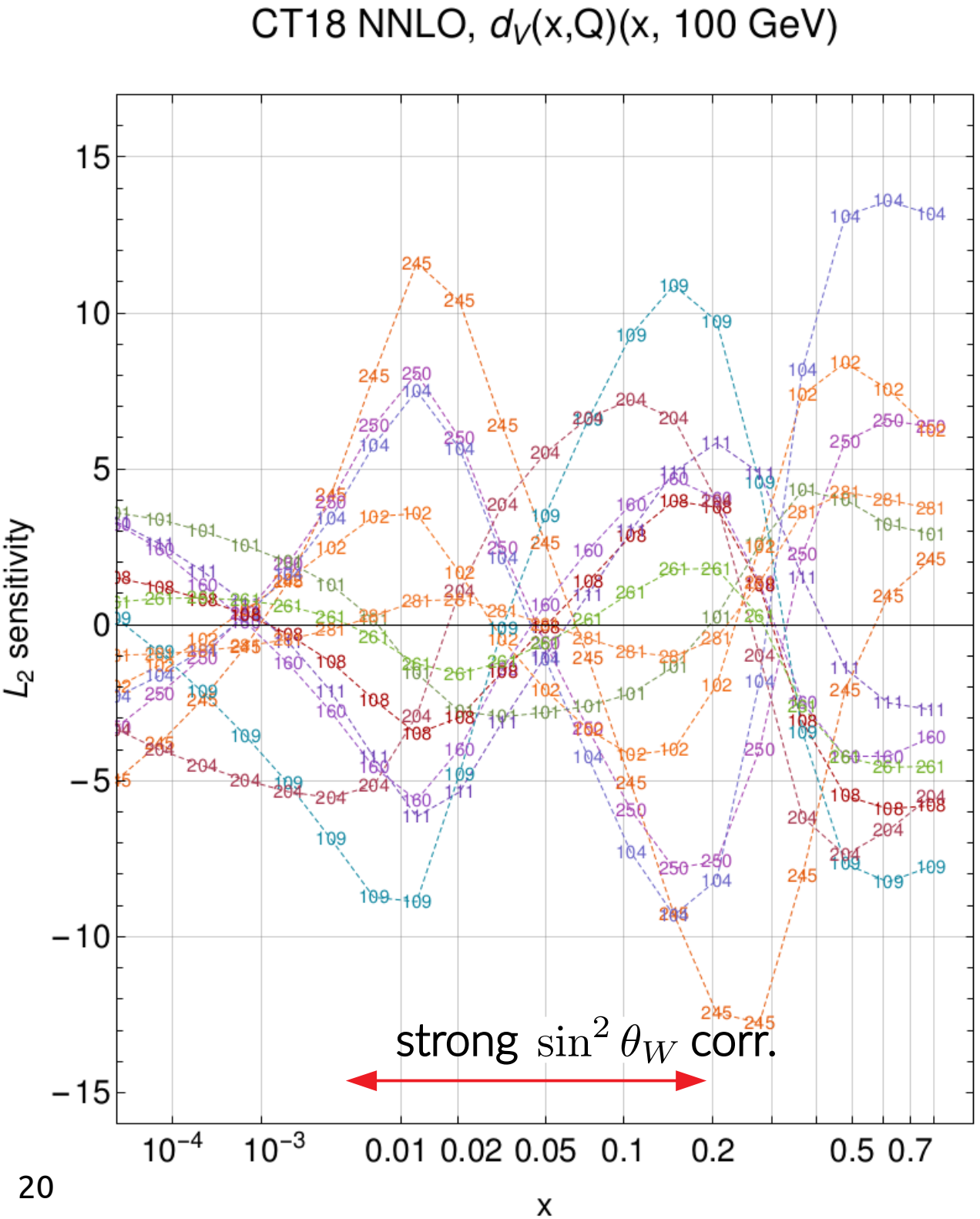


Experiments with large $\Delta\chi^2 > 0$ $[\Delta\chi^2 < 0]$ pull $g(x, Q)$ in the negative [positive] direction at the shown x

245 LHCb7ZWrap
250 LHCb8WZ
160 HERAIpII
101 BcdF2pCor
102 BcdF2dCor
104 NmcRatCor
109 cdhswf3
111 ccfrf3.md
201 e605
204 e866ppxf

tension between LHCb W/Z data (245, 250); fixed-target DIS, Drell-Yan (CDHSW F_3 [109], E866pp [204])

→ in the region where $\sin^2\theta_W$ and u_v are most correlated, opposing pulls from CCFR νDIS, BCDMS vs. E866 pp, NMC rat.



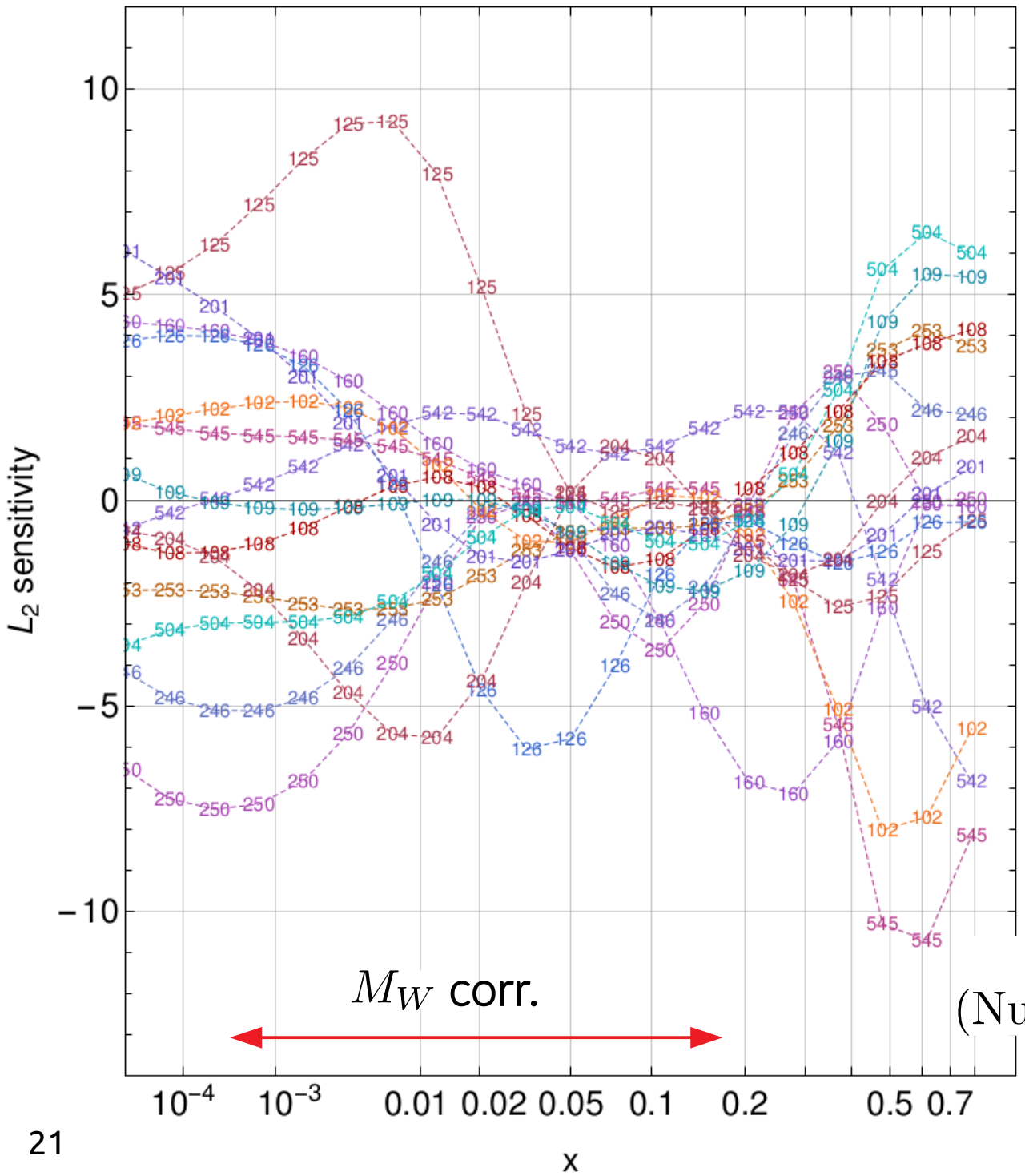
tension between LHCb W/Z data (245, 250); fixed-target DIS, Drell-Yan (CDHSW F_3 [109], E866pp [204])

245 LHCb7ZWrap
250 LHCb8WZ
160 HERAIpII
101 BcdF2pCor
102 BcdF2dCor
104 NmcRatCor
108 cdhswf2
109 cdhswf3
111 ccfrf3.md
204 e866ppxf
261 ZyCDF2
281 d02Easy5

again, tensions observed between, e.g., NMC ratio data and CDHSW, E866pp

M_W

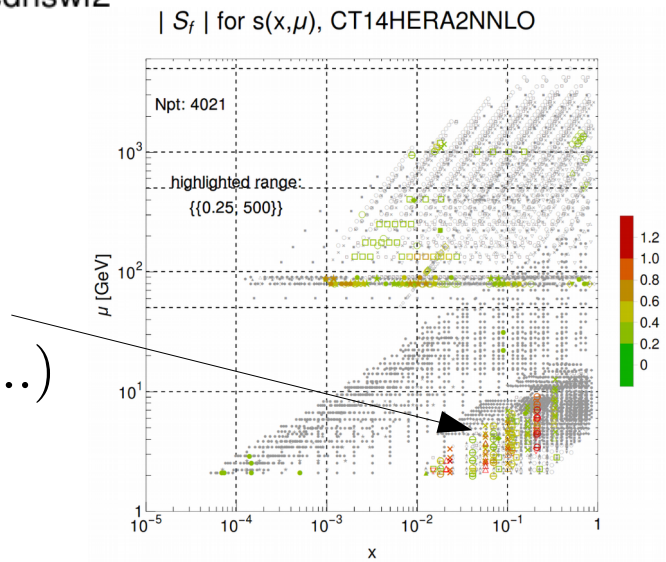
CT18 NNLO, s(x, 100 GeV)

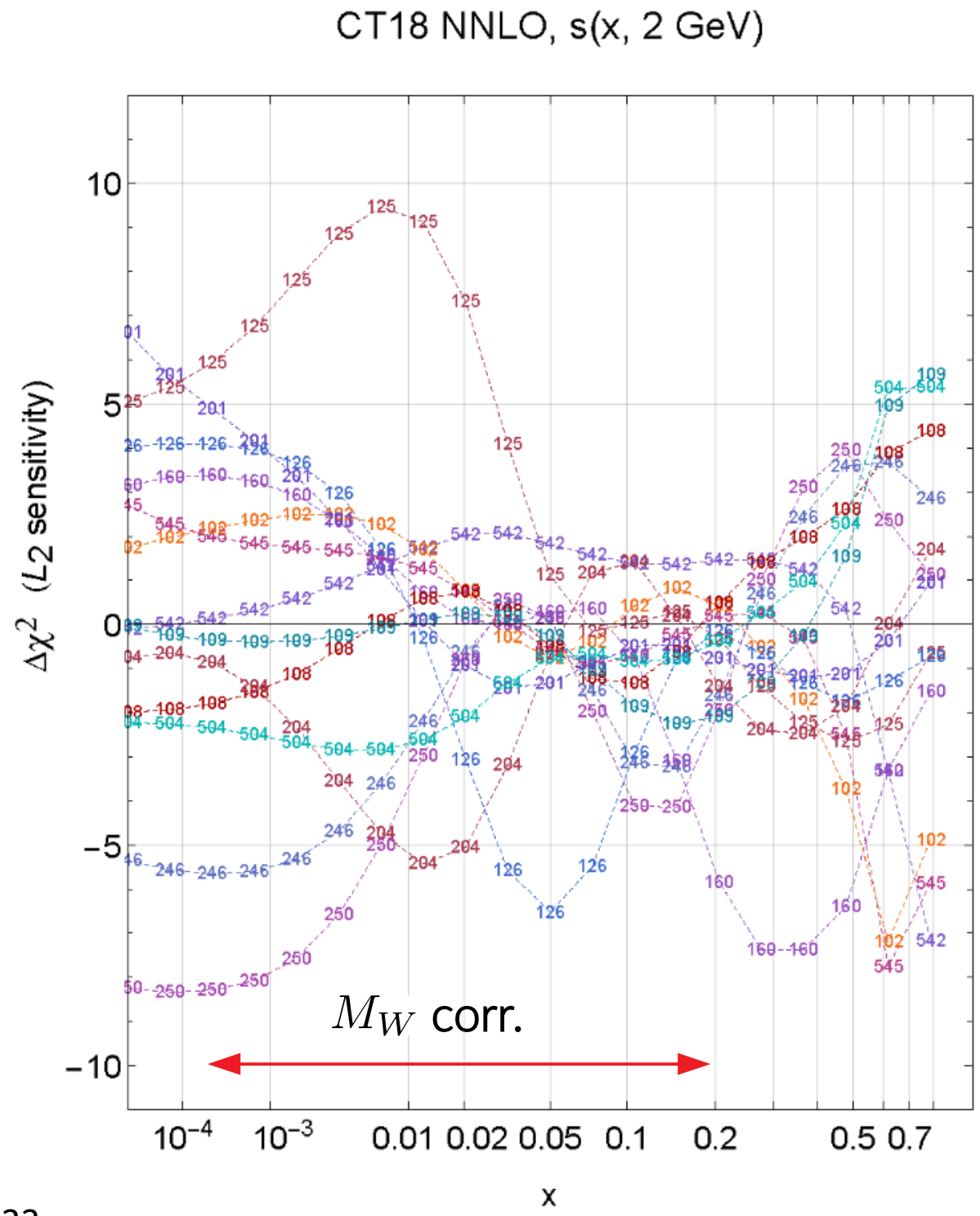


...this analysis can be extended to M_W, extractions of which are dependent upon s(x), through Z-calibration

246 LHCb8Zeer	109 cdhswf3
250 LHCb8WZ	125 NuTvNbChXN
253 ATL8ZpTbT	126 CcfrNuChXN
542 CMS7jtR7y6T	201 e605
545 CMS8jtR7T	204 e866ppxf
160 HERAIpII	504 cdf2jtCor2
102 BcdF2dCor	
108 cdhswf2	

νDIS (NuTeV, ...)



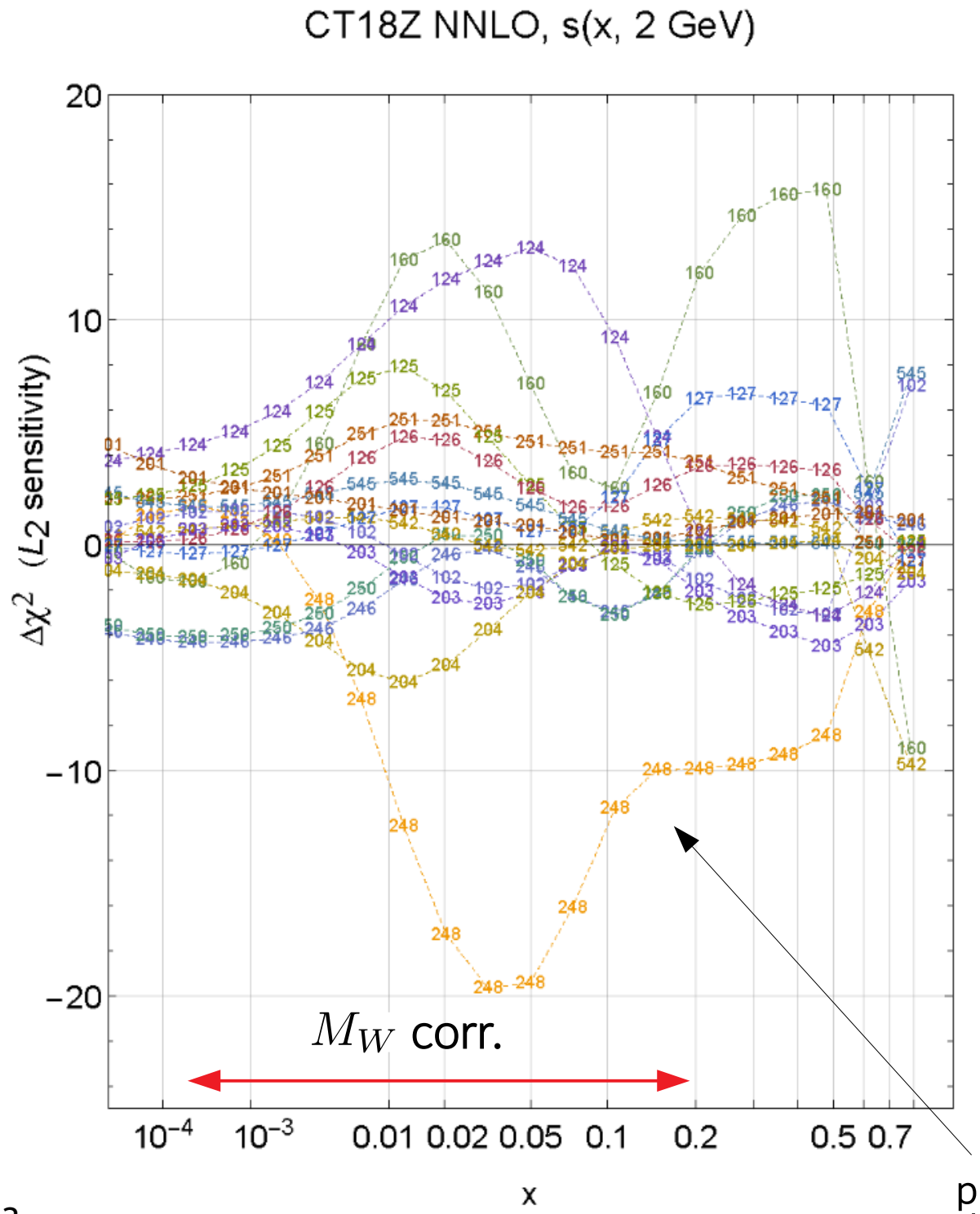


L_2 sensitivity, strangeness: CT18

Most sensitive experiments

246 LHCb8Zeer	125 NuTvNbChXN
250 LHCb8WZ	126 CcfrNuChXN
542 CMS7jtR7y6T	201 e605
545 CMS8jtR7T	204 e866ppxf
160 HERAIpII	504 cdf2jtCor2
102 BcdF2dCor	
108 cdhswf2	
109 cdhswf3	

A tension trend between DIS (HERA I+II, CCFR, NuTeV) and Drell-Yan (ATLAS 7 Z/W, LHCb W/Z, E866 pp, …) experiments



L_2 sensitivity, strangeness: CT18Z

Most sensitive experiments

246	LHCb8Zeer	124	NuTvNuChXN
248	ATL7ZW.xF	125	NuTvNbChXN
250	LHCb8WZ	126	CcfrNuChXN
251	ATL8DY	127	CcfrNbChXN
542	CMS7jtR7y6T	201	e605
545	CMS8jtR7T	203	e866f
160	HERAIpII	204	e866ppxf
102	BcdF2dCor		

A tension trend between DIS (HERA I+II, CCFR, NuTeV) and Drell-Yan (ATLAS 7 Z/W, LHCb W/Z, E866 pp, …) experiments

pronounced effect of ATLAS 7 TeV Z/W data!

σ_H crucially limited by PDF and α_s uncertainties

- there remains considerable dependence (as large as ~13%) upon PDF paramatrization and running coupling

→ the situation is such that precision in Higgs phenom. is significantly **PDF-limited**

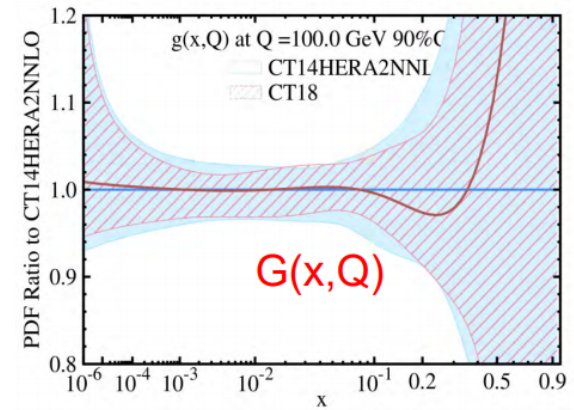
Accardi et al., EPJC**76**, 471 (2016).

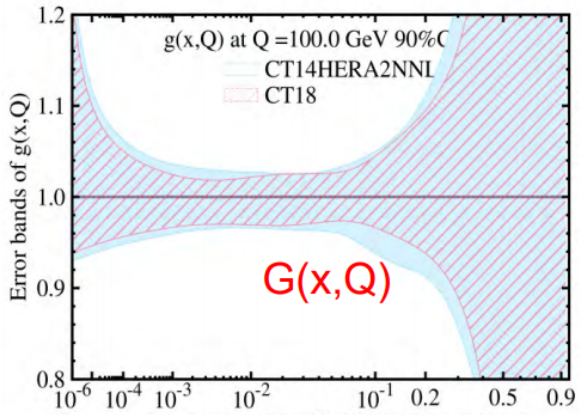
PDF sets	$\sigma(H)^{\mathrm{NNLO}}$ (pb) nominal $\alpha_s(M_Z)$	$\sigma(H)^{\mathrm{NNLO}}$ (pb) $\alpha_s(M_Z) = 0.115$	$\sigma(H)^{\mathrm{NNLO}}$ (pb) $\alpha_s(M_Z) = 0.118$
ABM12 [2]	39.80 ± 0.84	41.62 ± 0.46	44.70 ± 0.50
CJ15 [1][a]	$42.45^{+0.43}_{-0.18}$	$39.48^{+0.40}_{-0.17}$	$42.45^{+0.43}_{-0.18}$
CT14 [3][b]	$42.33^{+1.43}_{-1.68}$	$39.41^{+1.33}_{-1.56}$ (40.10)	$42.33^{+1.43}_{-1.68}$
HERAPDF2.0 [4][c]	$42.62^{+0.35}_{-0.43}$	$39.68^{+0.32}_{-0.40}$ (40.88)	$42.62^{+0.35}_{-0.43}$
JR14 (dyn) [5]	38.01 ± 0.34	39.34 ± 0.22	42.25 ± 0.24
MMHT14 [6]	$42.36^{+0.56}_{-0.78}$	$39.43^{+0.53}_{-0.73}$ (40.48)	$42.36^{+0.56}_{-0.78}$
NNPDF3.0 [7]	42.59 ± 0.80	39.65 ± 0.74 (40.74 ± 0.88)	42.59 ± 0.80
PDF4LHC15 [8]	42.42 ± 0.78	39.49 ± 0.73	42.42 ± 0.78

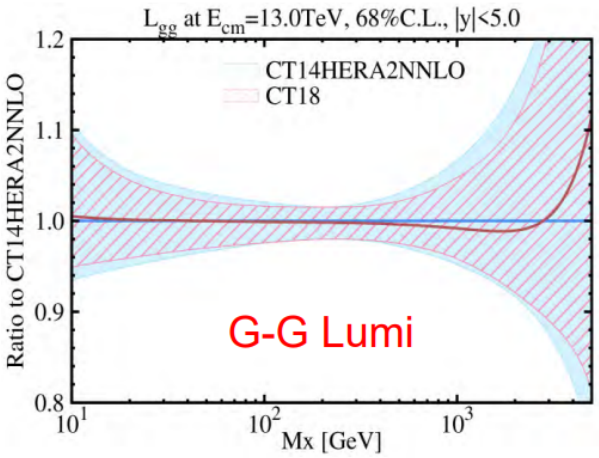
$$\sigma_H \text{ at NNLO and } \sqrt{s} = 13\,\text{TeV};\ \mu_F = \mu_R = m_H$$

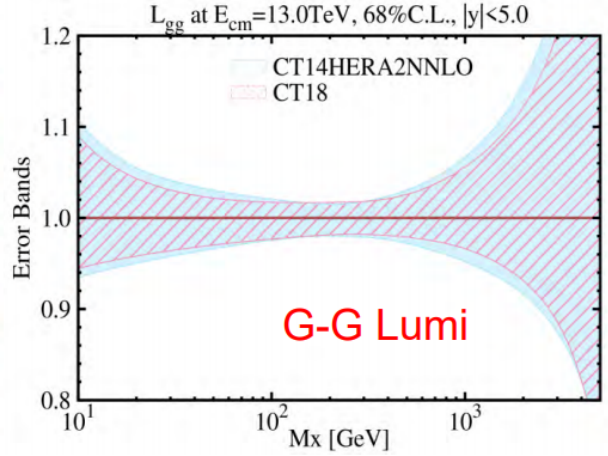
→ enhancing the discovery potential in the Higgs sector will require improving these uncertainties!

CT14 → CT18 modestly shifts Higgs cross sections and slightly reduces PDF uncertainties







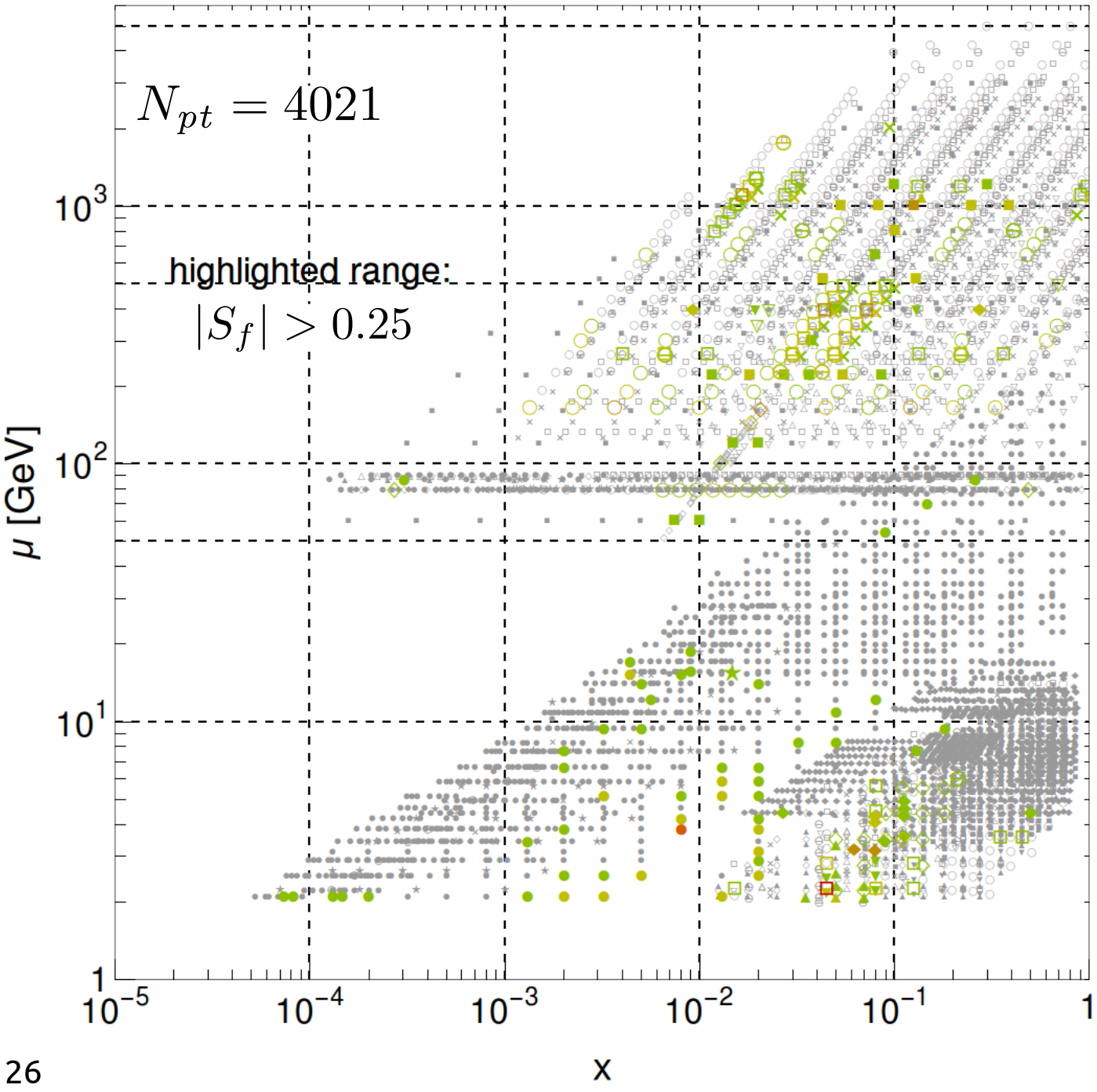


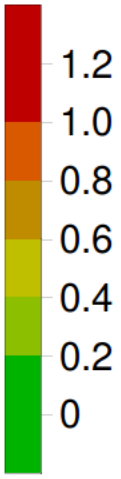
	7 TeV	
	σ(gg-h)	$\delta\sigma$ sym(90%C.L.)
CT14NNLO	14.67	0.46
CT18	14.57	0.44
	8 TeV	
	σ(gg-h)	$\delta\sigma$ sym(90%C.L.)
CT14NNLO	18.70	0.57
CT18	18.45	0.55
	13 TeV	
	σ(gg-h)	$\delta\sigma$ sym(90%C.L.)
CT14NNLO	42.78	1.32
CT18	42.43	1.26
	14 TeV	
	σ(gg-h)	$\delta\sigma$ sym(90%C.L.)
CT14NNLO	48.23	1.50
CT18	47.91	1.42

PDF induced errors (at 90% CL) are reduced by about 5% as compared to CT14 predictions.

can we disentangle elements of the global analysis responsible for these improvements?

$|S_f|$ for σ_{H^0} 14 TeV, CT14HERA2NNLO



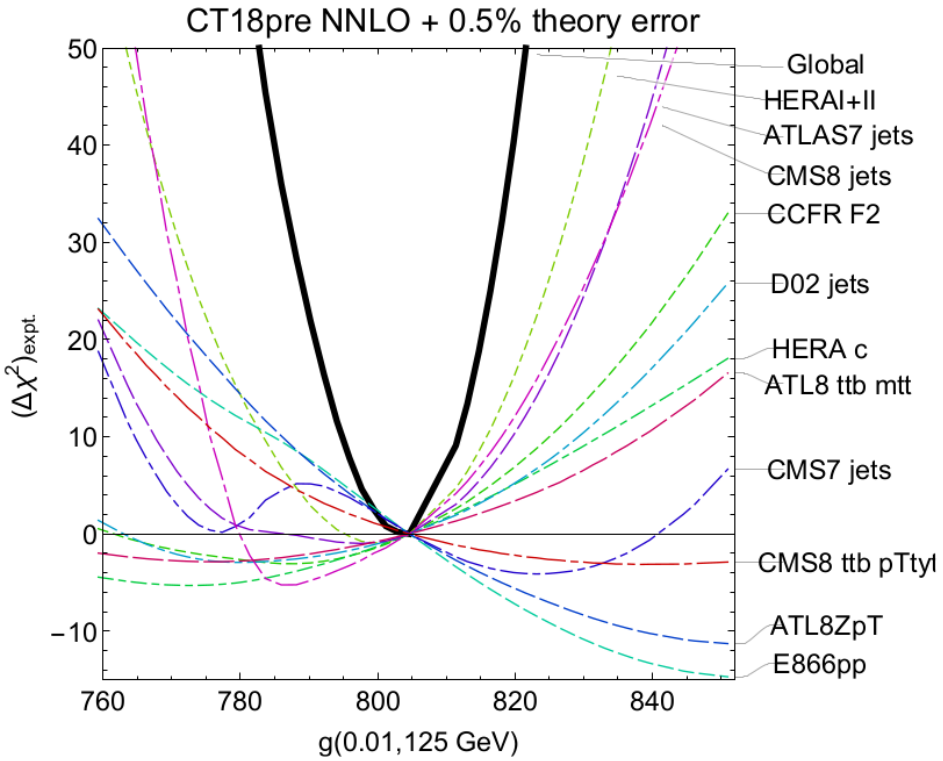


- after the aggregated HERA data, **inclusive jet production** have greatest total sensitivity!

→ large correlations for **E866**, **BCDMS**, **CCFR**, **CMS WASY**, **Z p_T** and **ttbar** production, but smaller numbers of highly-sensitive points

we use the Higgs region g(x) to validate PDFSense

...for the gluon PDF in the Higgs region, $g(0.01, m_H)$

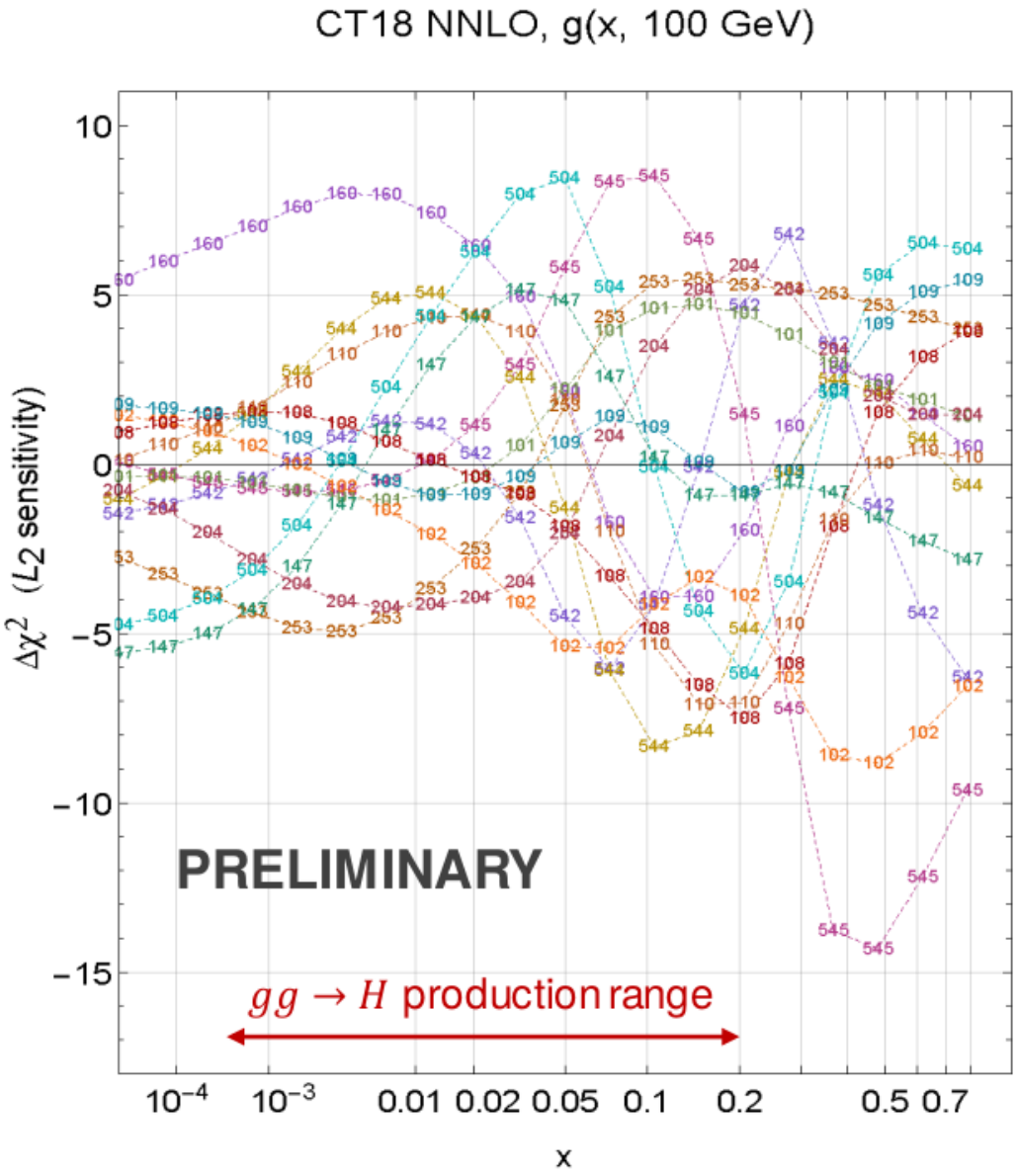


$g(x=0.01, \mu=125$ GeV)		
PDFSense		LM scan
CT14HERA2	CT18pre	CT18pre
HERAI+II'15	HERAI+II'15	HERAI+II'15
CMS8jets'17	CMS8jets'17	CMS8jets'17
CMS7jets'14	CMS7jets'14	ATL8ZpT'16
ATLAS7jets'15	E866pp'03	E866pp'03
E866pp'03	ATLAS7jets'15	ATLAS7jets'15
BCDMSd'90	BCDMSd'90	CCFR-F2'01
CCFR-F3'97	BCDMSp'89	D02jets'08
D02jets'08	D02jets'08	HERAc'13
NMCrat'97	NMCrat'97	NuTeV-nub'06
BCDMSp'89	CDHSW-F2'91	CCFR-F3'97

- PDFSense identifies the most sensitive experiments with high confidence and in accord with other methods such as the LM scans. It works the best when the uncertainties are nearly Gaussian, and experimental constraints agree among themselves [arXiv:1803.02777]

Estimated χ^2 pulls from experiments

(L_2 sensitivity, arXiv:1904.00222, v. 2)



CT18 NNLO, gluon at Q=100 GeV

Most sensitive experiments

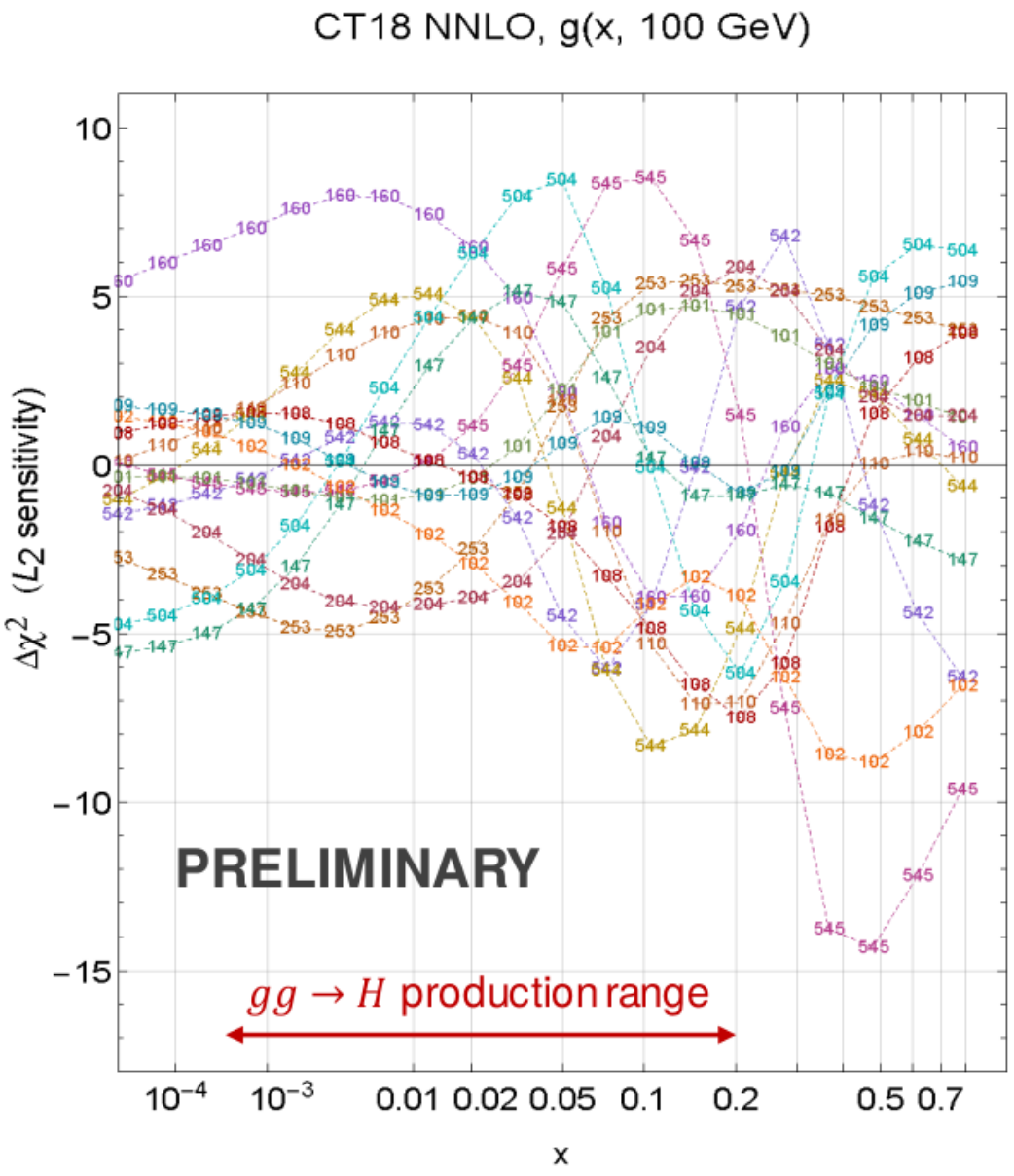
253	ATL8ZpTbT	109	cdhswf3
542	CMS7jtR7y6T	110	ccfrf2.mi
544	ATL7jtR6uT	147	Hn1X0c
545	CMS8jtR7T	204	e866ppxf
160	HERAIpII	504	cdf2jtCor2
101	BcdF2pCor		
102	BcdF2dCor		
108	cdhswf2		

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 using CT18 Hessian PDFs

Estimated χ^2 pulls from experiments

(L_2 sensitivity, arXiv:1904.00222, v. 2)



CT18 NNLO, gluon at Q=100 GeV

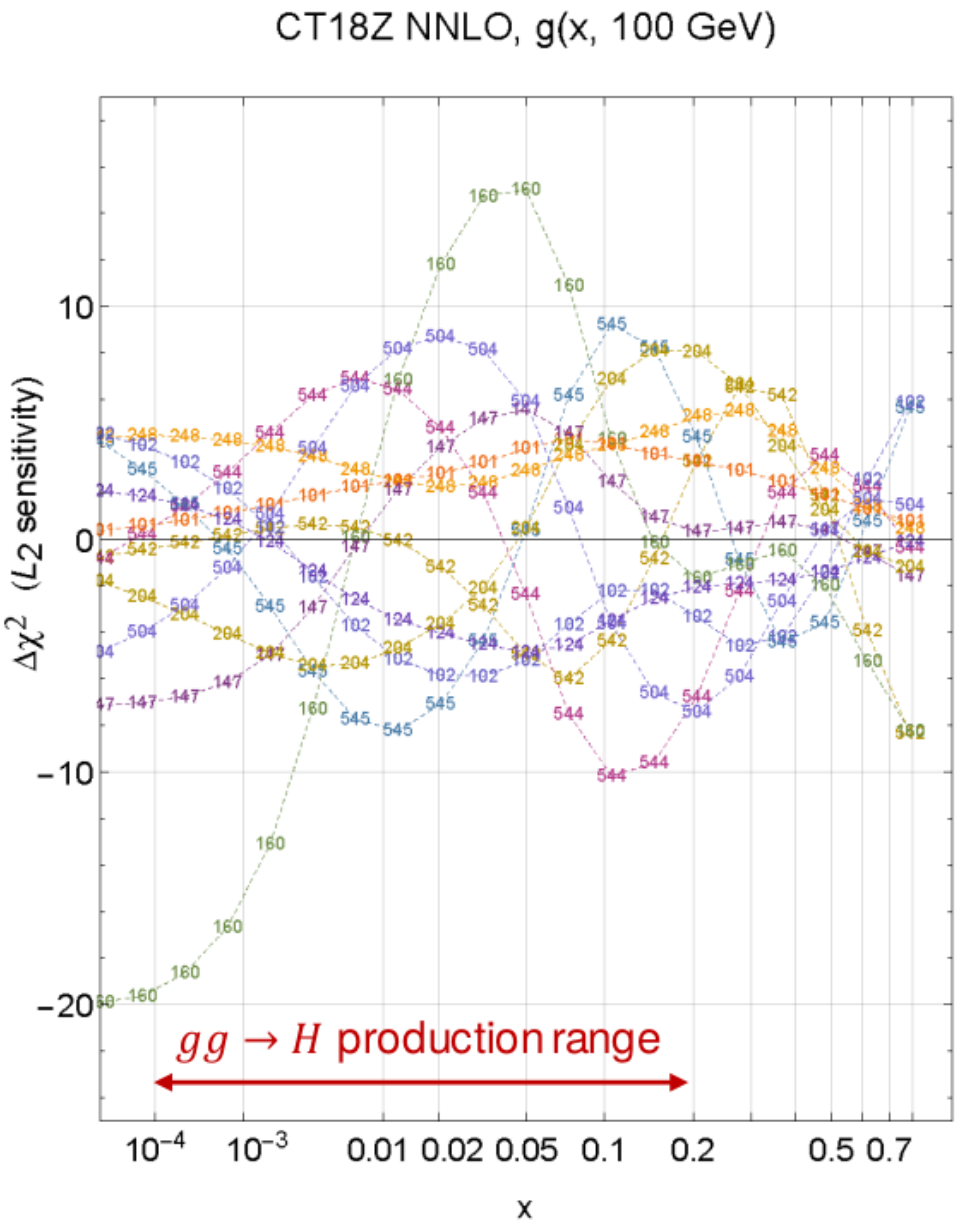
Most sensitive experiments

253	ATL8ZpTbT	109	cdhswf3
542	CMS7jtR7y6T	110	ccfrf2.mi
544	ATL7jtR6uT	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

Estimated χ^2 pulls from experiments

(L_2 sensitivity, arXiv:1904.00222, v. 2)



Same for CT18Z NNLO

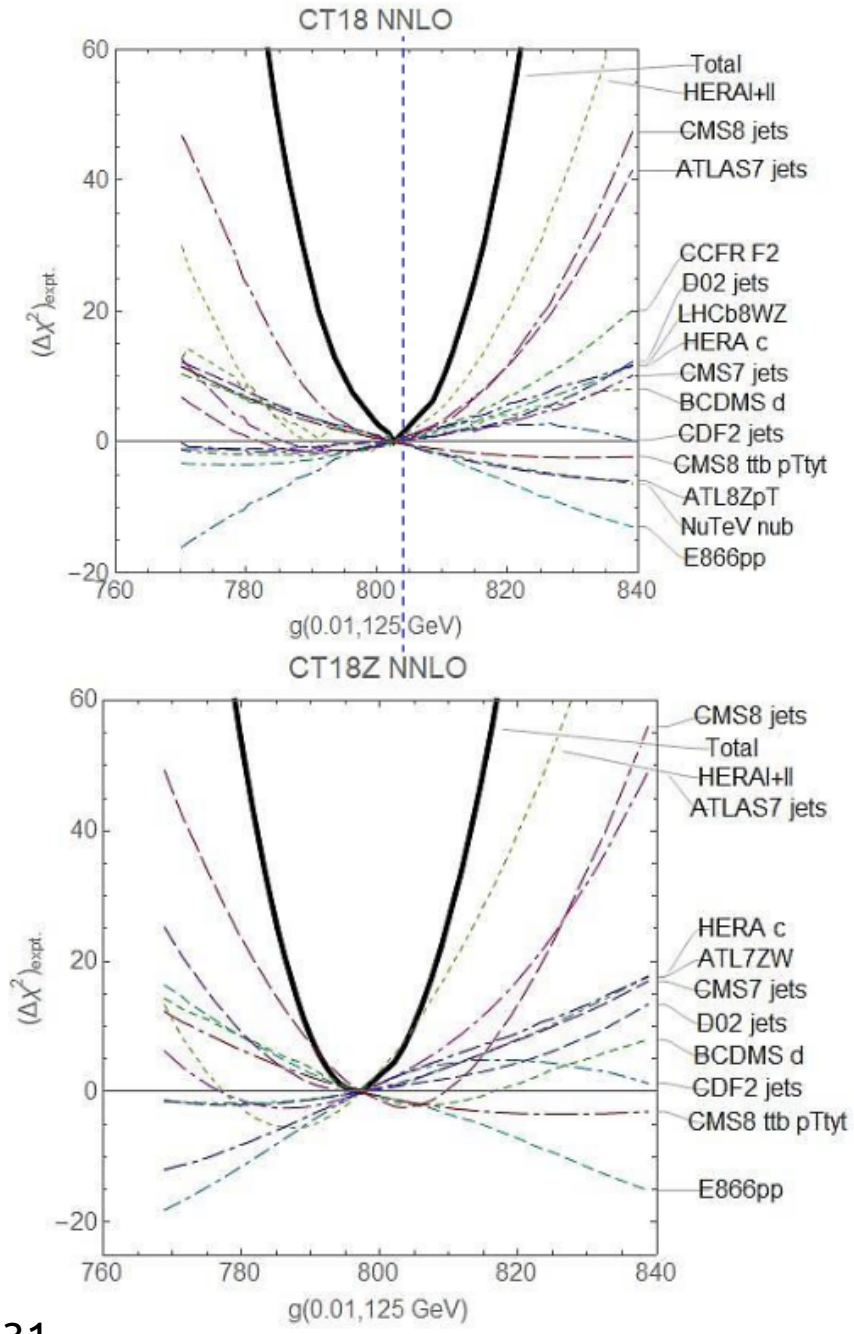
Most sensitive experiments

248 ATL7ZW.xF	147 Hn1X0c
542 CMS7jtR7y6T	204 e866ppxf
544 ATL7jtR6uT	504 cdf2jtCor2
545 CMS8jtR7T	
160 HERAIpII	
101 BcdF2pCor	
102 BcdF2dCor	
124 NuTvNuChXN	

Constraints from HERA I+II DIS (ID=160) follow a different trend from CT18 NNLO

PRELIMINARY

Lagrange Multiplier scan: $g(0.01, 125\ GeV)$



Upper row: CT18

- HERAI+II data set provides the dominant constraint, followed by ATLAS, CDF2, CMS, D02 jet production, HERA charm,…
- $t\bar{t}$ double-diff. cross sections provide weaker constraints

Lower row: CT18Z

- CT18Z: a 1% lower NNLO gluon in the Higgs production region than for CT14/CT18

  (a consequence of using a "saturation" factorization scale in NNLO DIS for CT18Z)

BSM (?)

more speculatively, BSM searches and SMEFT represent possibilities for analogous sensitivity analyses

---

- SMEFiT analyses fit a generic parametrization of BSM physics,

e.g., van Beek, Nocera, Rojo, Slade, arXiv:1906.05296

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i^{N_{\rm op}} \frac{c_i}{\Lambda^2} \mathcal{O}_i^{(6)}$$

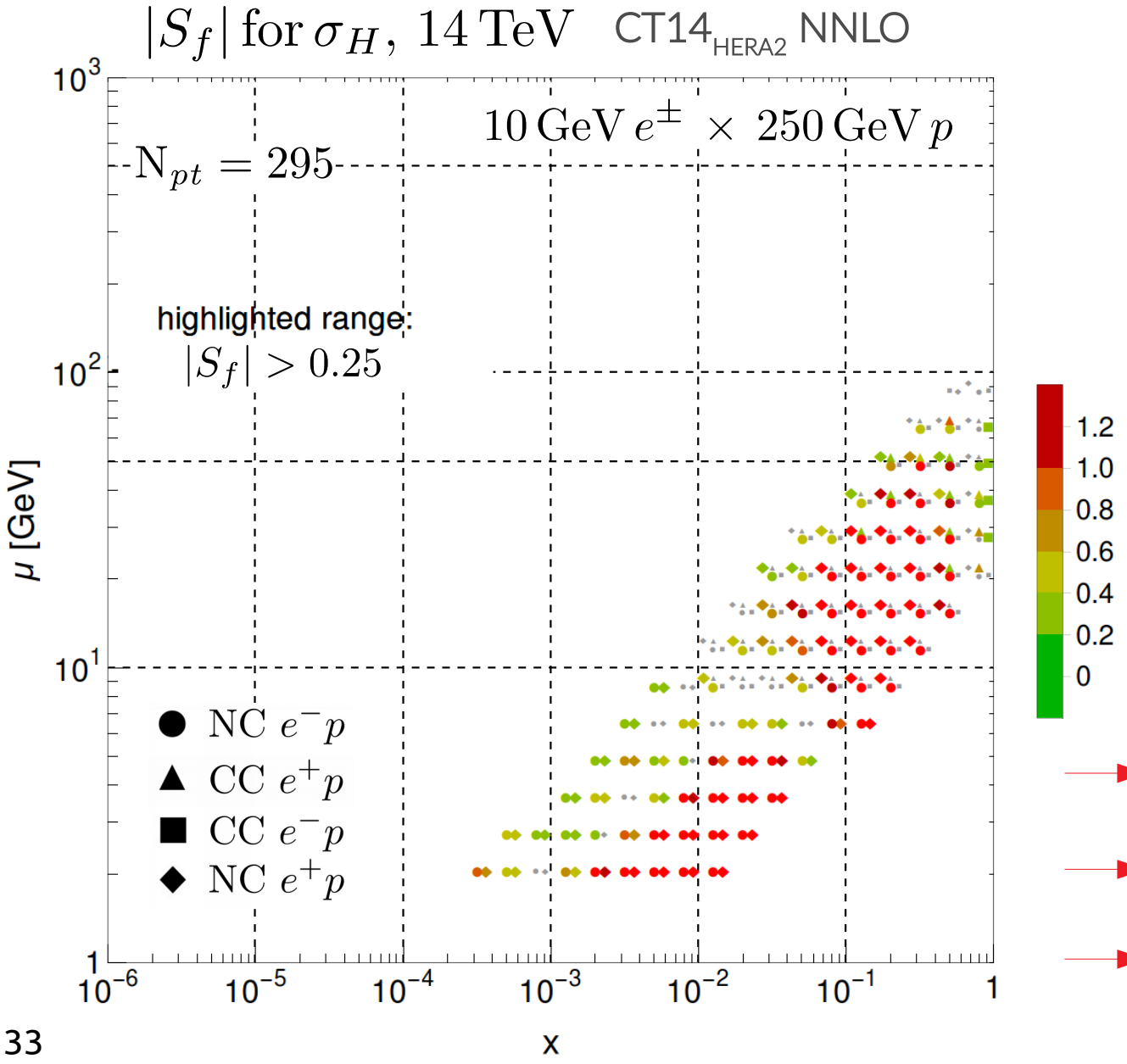
dependent upon Wilson cofficients for dim-6 operators:

$$\{ c_i^{(k)} \}, \quad i = 1, \ldots, N_{\rm op} \,, \quad k = 1, \ldots, N_{\rm rep}$$

→ with error replicas in the space of $\{c_i^{(k)}\}$, one could analyze data pulls on SMEFiT ingredients à la `PDFSense`

the future: toward precision PDFs

→ a systematic approach to reconcile LHC experiments will be imperative to reducing PDF uncertainties for the precision EW era at **HL-LHC**



- very high luminosity DIS colliders, like the electron-ion collider (EIC) at JLab/BNL or LHeC could provide a new generation of very precise data in the fixed-target region

→ 'cleaner' leptonic probes

→ free of nuclear effects

→ OR, understand nuclear medium (for CCFR, NuTeV, ...)

concluding observations

- `PDFSense` produces kinematic maps and summary charts of data-point sensitivities to PDFs and PDF-dependent quantities; readily extendable

  → fast; public code: http://metapdf.hepforge.org/PDFSense

- PDF uncertainties of HEP observables can be opportunties for sensitive PDF measurements

- heightened constraining power (on $\sin^2\theta_W$, σ_H, ...) will require smoothing data tensions that are now being explored

  → $\sin^2\theta_W$ driven by u_{val}, d_{val}, which, e.g., E866pp, BCDMS have opposing pulls

  a community effort to understand systematic tensions needed

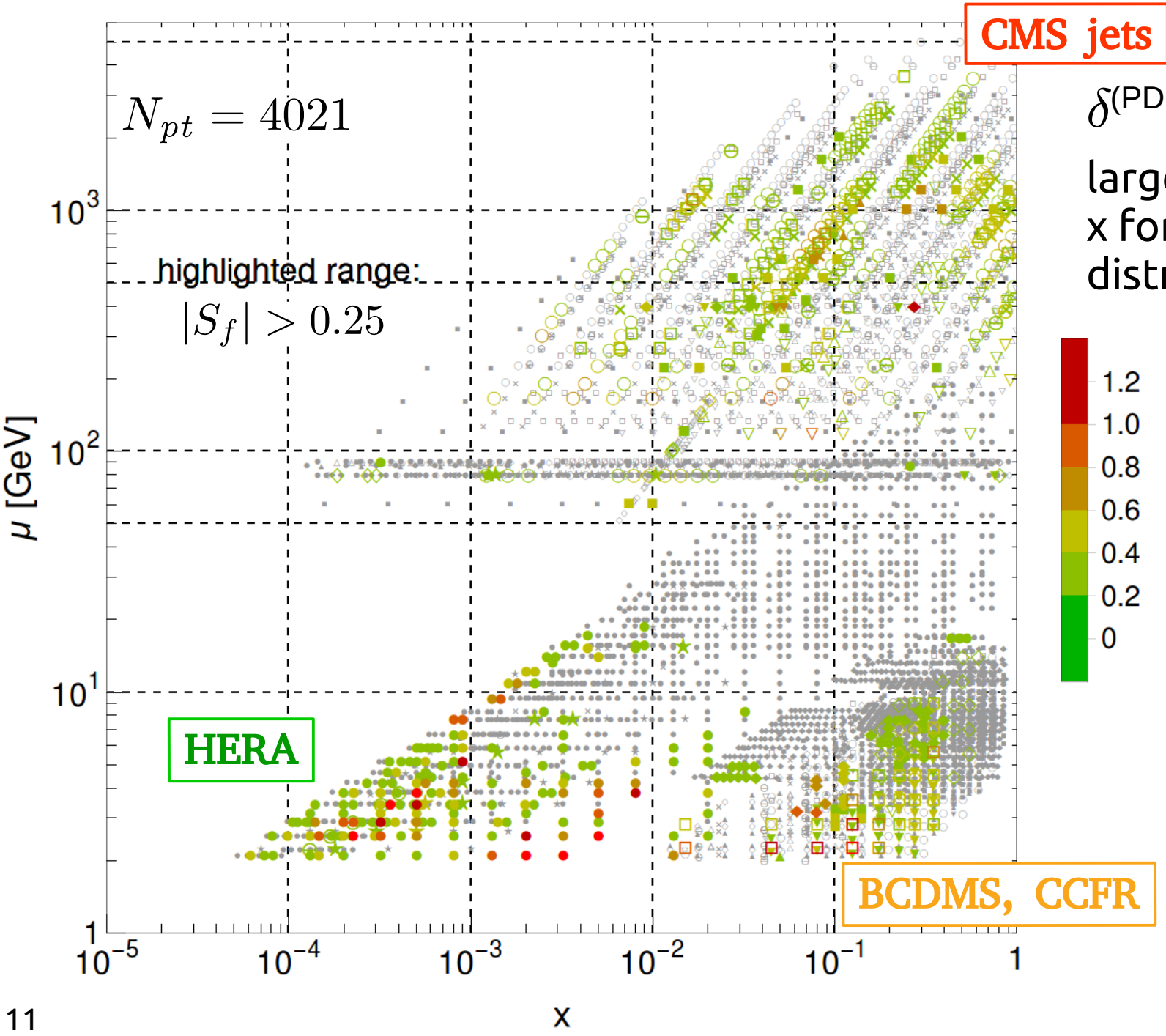
  → we find important improvements σ_H from incl. jet prod.; resolution of tensions will further tighten constraints

- dealing with these issues will be essential for HL-LHC program

THANKS!!

supplemental material

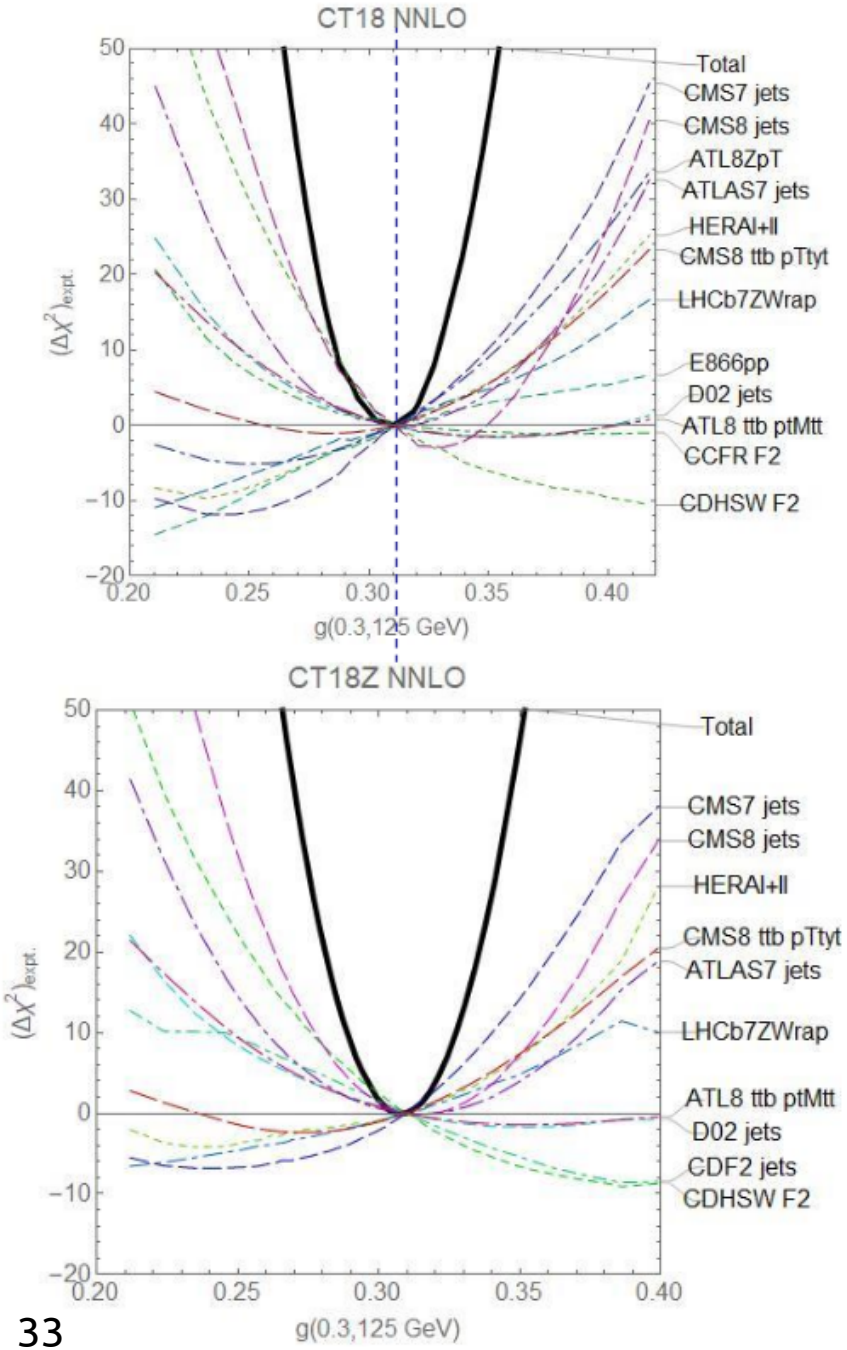
$|S_f|$ for g(x,μ), CT14HERA2NNLO



$\delta^{(\mathrm{PDF})} r_i$ relatively large at high and low x for the gluon distribution

- measurements at high/low x are therefore most sensitive; esp. Run 2 **HERA** points and **CMS** jet production at 7, 8 TeV

Lagrange Multiplier scan: $g(0.3, 125\ GeV)$



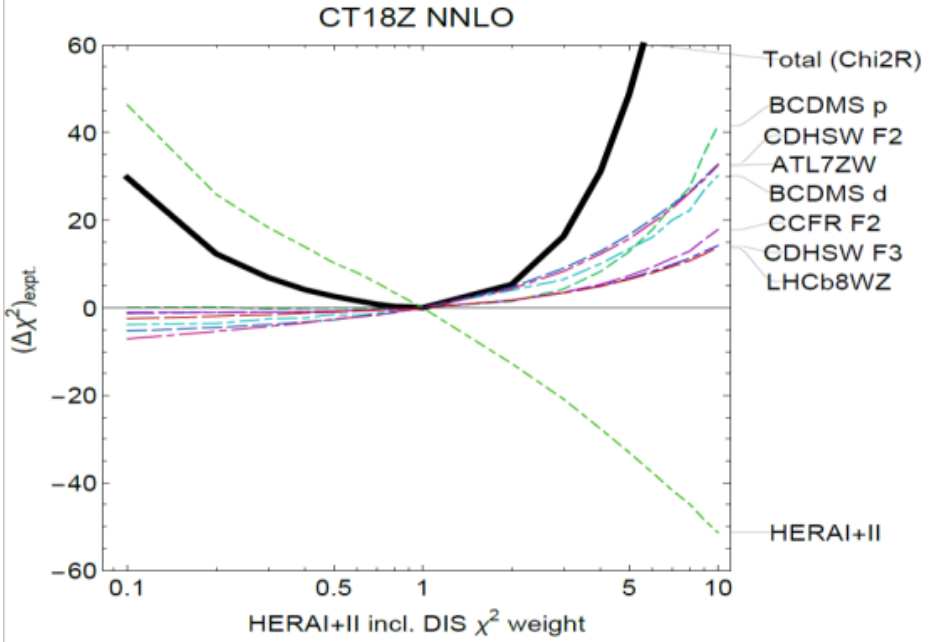
Upper/lower rows: CT18/CT18Z

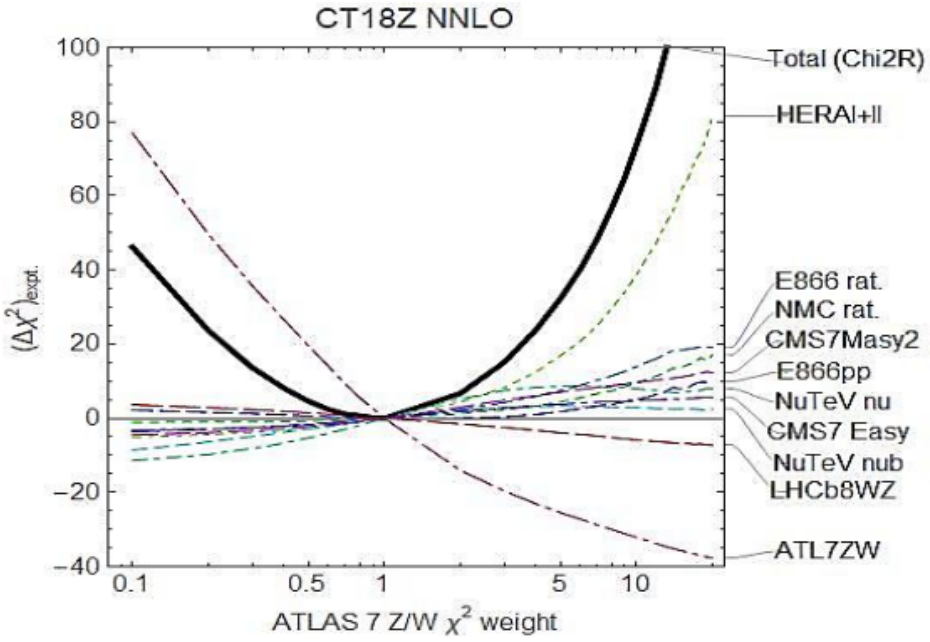
Good overall agreement. But observe opposite pulls from ATLAS7/CMS7 jet production and CMS8 jet production

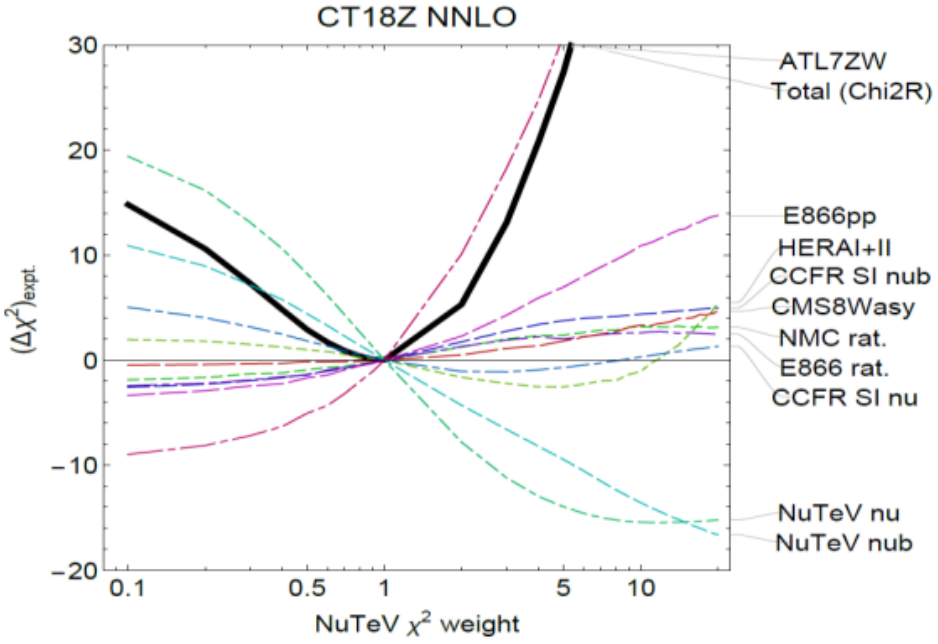
Similarly, ATLAS $t\bar{t}$ distributions $\mathrm{d}^2\sigma/(dp_{T,t}dm_{t\bar{t}})$ and CMS $t\bar{t}$ distributions $\mathrm{d}^2\sigma/(dp_{T,t}dy_{t,ave})$ at 8 TeV impose weak opposite pulls

Constraints from ATLAS 8 Z p_T production data are moderate and still affected by NNLO scale uncertainty

LM scans on χ^2 weights of HERA I+II, ATLAS 7 Z/W, and NuTeV data





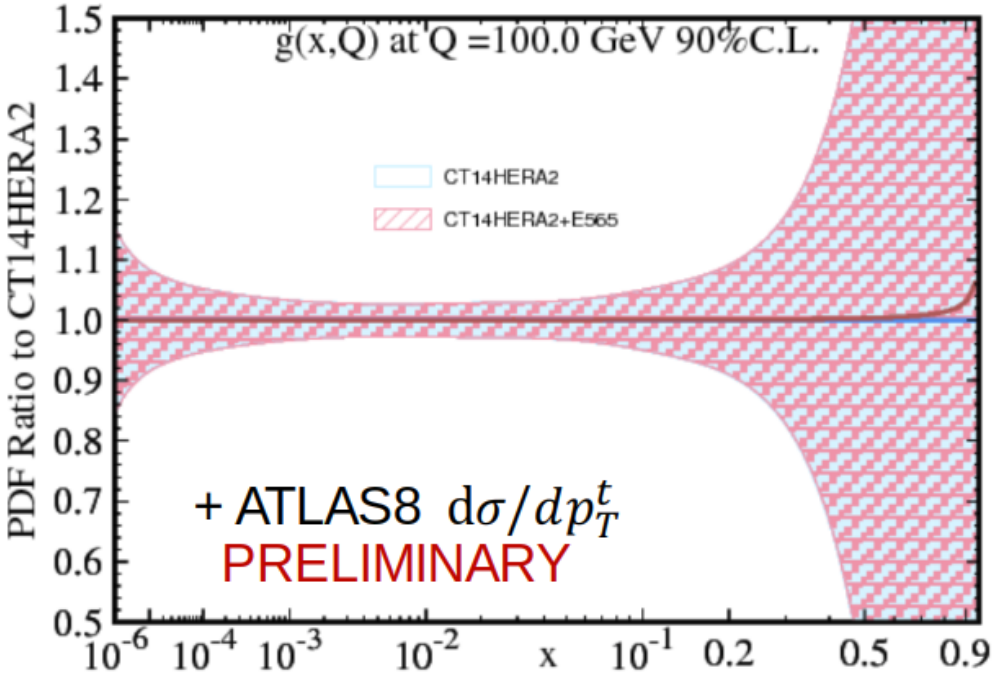


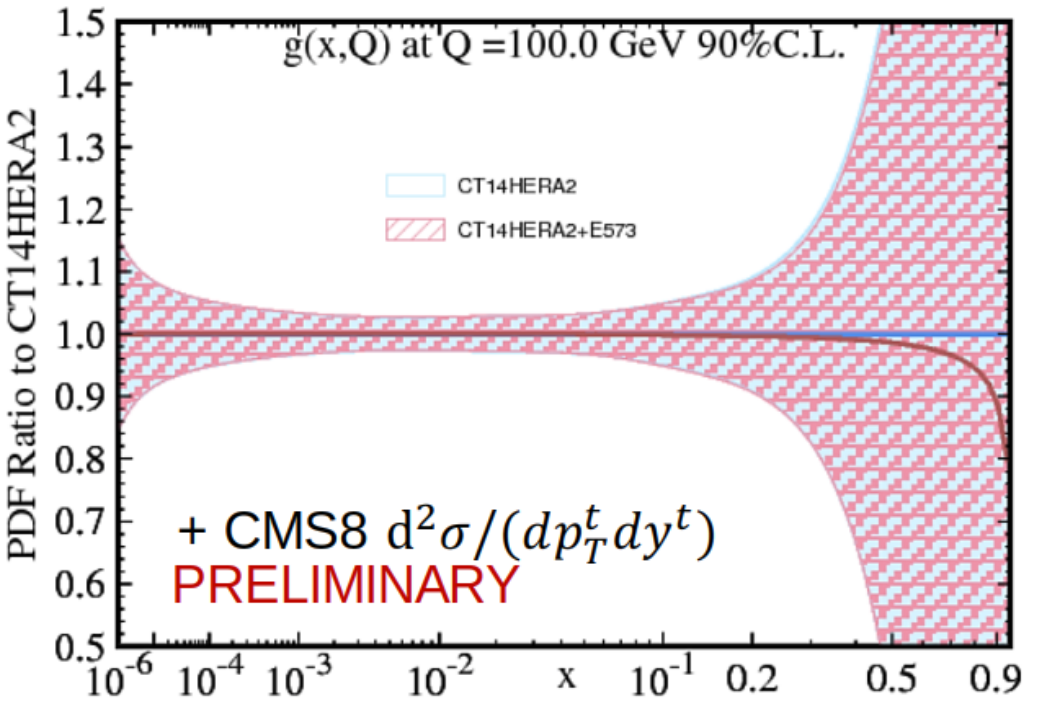
Fits with varied weights and LM scans reveal a disagreement between important DIS [primarily HERA, CCFR, NuTeV,…] and DY [primarily ATL7ZW, E866, LHCb8WZ,…] experiments. This is more pronounced for large$-x$ gluon as well as strangeness.

PDF sensitivity of $t\bar{t}$

ePump estimates of impact of $t\bar{t}$ data

When the $t\bar{t}$ data were added to the CT14HERA2 NNLO data set according to the **ePump** Hessian reweighting method (CT14nn+EXXX), no significant change in the PDF uncertainty was observed





FastNNLO calculation from **Csakon, Mitov, et al**.

Slight reduction in gluon PDF uncertainty at $x \sim 0.2$; weaker than for the other groups because of also including jet production data

Top-antitop production experiments have **strong** sensitivity per data point, offer a **novel** independent measurement of the gluon PDF in several channels

2nd aside: **kinematical matchings**

- **residual-PDF correlations and sensitivities are evaluated at parton-level kinematics determined according to leading-order matchings with physical scales in measurements**

deeply-inelastic scattering:

$$\mu_i \approx Q|_i\,,\ x_i \approx x_B|_i$$

x_i : parton mom. fraction

μ_i : factorization scale

hadron-hadron collisions:

$AB \to CX$

$$\mu_i \approx Q|_i\,,\ x_i^{\pm} \approx \frac{Q}{\sqrt{s}}\exp(\pm y_C)\Big|_i$$

single-inclusive jet production: $Q = 2p_{Tj},\ y_C = y_j$

$t\bar{t}$ pair production: $Q = m_{t\bar{t}},\ y_C = y_{t\bar{t}}$

etc...

$d\sigma/dp_T^Z$ measurements: $Q = \sqrt{(p_T^Z)^2 + (M_Z)^2},\ y_C = y_Z$

the goal is to **quantify the strength of the constraints** placed on a particular set of PDFs by both individual and aggregated measurements *without direct fitting*

- for single-particle hadroproduction of gauge bosons at, *e.g.*, LHC, factorization gives

$$\sigma(AB \to W/Z+X) = \sum_n \alpha_s^n(\mu_R^2) \sum_{a,b} \int dx_a dx_b \times f_{a/A}(x_a, \mu^2)\, \hat{\sigma}^{(n)}_{ab\to W/Z+X}(\hat{s}, \mu^2, \mu_R^2)\, f_{b/B}(x_b, \mu^2)$$

PDFs determined by fits to data; *e.g.*, "CT14H2"

pQCD matrix elements – specified by theoretical formalism in a given fit

- *idea* : study the **statistical <u>correlation</u>** between PDFs and the quality of the fit at a measured data point(s); fit quality encoded in a (**T**heory) – (shifted **D**ata) ***residual*** :

$$r_i(\vec{a}) = \frac{1}{s_i}\left(T_i(\vec{a}) - D_{i,sh}(\vec{a})\right)$$

s_i : uncorrelated uncert. | $\vec{a}$: PDF parameters

a brief statistical aside, i

- the CTEQ-TEA global analysis relies on the **Hessian formalism** for its error treatment

$$\chi^2_E(\vec{a}) = \sum_{i=1}^{N_{pt}} r_i^2(\vec{a}) + \sum_{\alpha=1}^{N_\lambda} \overline{\lambda}_\alpha^2(\vec{a})$$

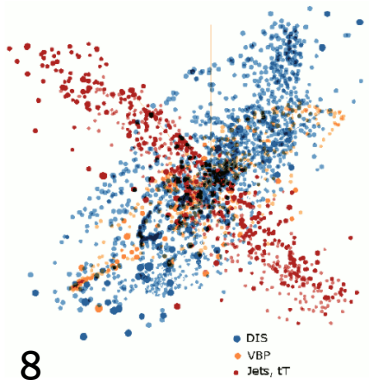
→ nuisance parameters to handle correlated errors

$$r_i(\vec{a}) = \frac{1}{s_i}\left(T_i(\vec{a}) - D_{i,sh}(\vec{a})\right)$$

these result in systematic shifts to data central values:

$$D_i \to D_{i,sh}(\vec{a}) = D_i - \sum_{\alpha=1}^{N_\lambda} \beta_{i\alpha}\overline{\lambda}_\alpha(\vec{a})$$

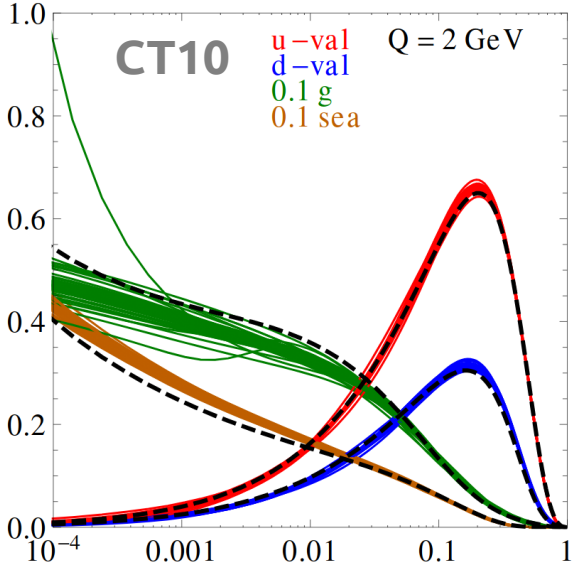
- a 56-dimensional parametric basis $\vec{a}$ is obtained by diagonalizing the Hessian matrix H determined from χ^2 (following a 28-parameter fit)



→ use this basis to compute 56-component "normalized" residuals :

$$\delta^{\pm}_{i,l} \equiv \left(r_i(\vec{a}^{\pm}_l) - r_i(\vec{a}_0)\right) / \langle r_0 \rangle_E$$

where $\langle r_0 \rangle_E \equiv \sqrt{\frac{1}{N_{pt}} \sum_{i=1}^{N_{pt}} r_i^2(\vec{a}_0)}$



a brief statistical aside, i [a]

error replicas

$$\sigma^{i}_{\mathrm{DY}}(\vec{a}) \rightarrow \sigma^{i,(1)}_{\mathrm{DY}},\ \sigma^{i,(2)}_{\mathrm{DY}},\ \vdots,\ \sigma^{i,(2n)}_{\mathrm{DY}}$$

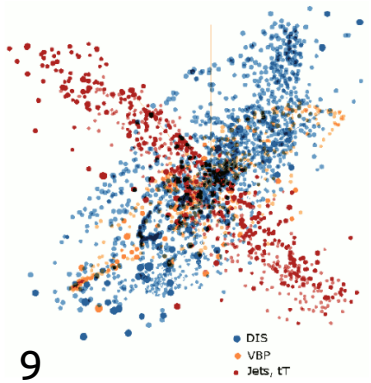
$$\begin{pmatrix} r^{i,(1)} \sim \sigma^{i,exp.}_{\mathrm{DY}} - \sigma^{i,(1)}_{\mathrm{DY}} \\ r^{i,(2)} \sim \sigma^{i,exp.}_{\mathrm{DY}} - \sigma^{i,(2)}_{\mathrm{DY}} \\ \vdots \\ r^{i,(2n)} \sim \sigma^{i,exp.}_{\mathrm{DY}} - \sigma^{i,(2n)}_{\mathrm{DY}} \end{pmatrix}$$

correlated?

$$\begin{pmatrix} g^{(1)}(x_i,\mu_i) \\ g^{(2)}(x_i,\mu_i) \\ \vdots \\ g^{(2n)}(x_i,\mu_i) \end{pmatrix}$$

PDF prediction for i^{th} empirical measurement kinematically matched to a specific (x_i, μ_i)

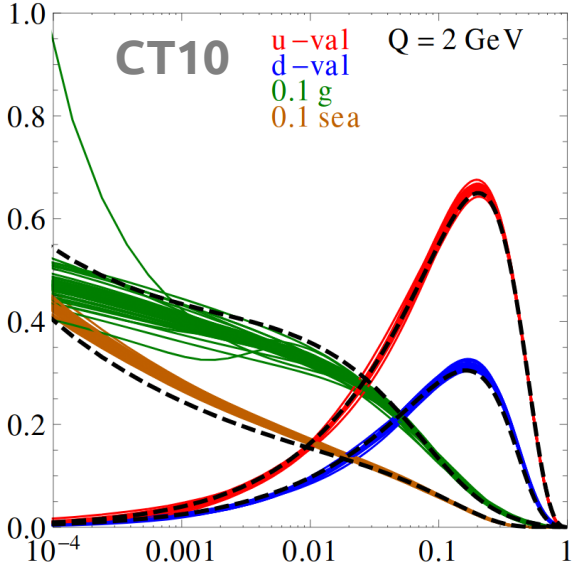
- a 56-dimensional parametric basis $\vec{a}$ is obtained by diagonalizing the Hessian matrix H determined from χ^2 (following a 28-parameter fit)



use this basis to compute 56-component "normalized" residuals :

$$\delta^{\pm}_{i,l} \equiv \left(r_i(\vec{a}^{\pm}_l) - r_i(\vec{a}_0)\right) / \langle r_0 \rangle_E$$

where $\langle r_0 \rangle_E \equiv \sqrt{\frac{1}{N_{pt}} \sum_{i=1}^{N_{pt}} r_i^2(\vec{a}_0)}$

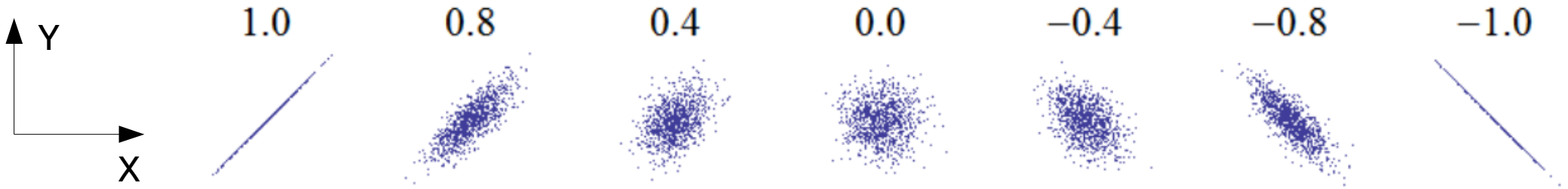


a brief statistical aside, ii

- … but how does the behavior of these residuals relate to the fitted PDFs and their uncertainties?

→ for example, how does the PDF uncertainty (at specific x, μ) correlate with the residual associated with a theoretical prediction at the same x, μ?

examine the Pearson correlation over the 56-member PDF error set between a PDF of given flavor and the residual



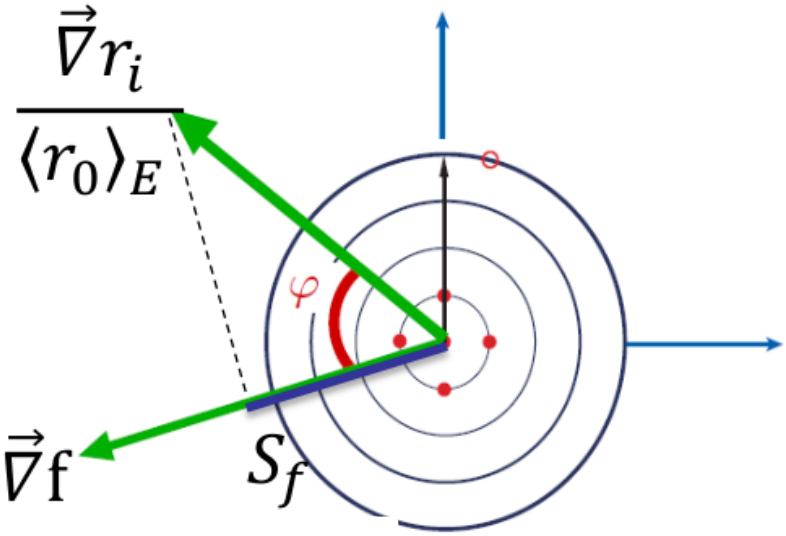
[X,Y] are exactly (anti-)correlated at the far (right) left above.

- we may then evaluate correlations between arbitrary PDF-derived quantities over the ensemble of error sets ([X,Y] may be PDFs, cross sections, residuals,…):

$$\mathrm{Corr}[X,Y] = \frac{1}{4\Delta X \Delta Y}\sum_{j=1}^{N}(X_j^+ - X_j^-)(Y_j^+ - Y_j^-) \qquad \Delta X = \frac{1}{2}\sqrt{\sum_{j=1}^{N}(X_j^+ - X_j^-)^2}$$

Correlation C_f and sensitivity S_f

The relation of data point i on the PDF dependence of f can be estimated by:



- $C_f \equiv \mathrm{Corr}[\rho_i(\vec{a})), f(\vec{a})] = cos\varphi$

$\vec{\rho}_i \equiv \vec{\nabla} r_i / \langle r_0 \rangle_E$ -- gradient of r_i normalized to the r.m.s. average residual in expt E;

$$\left(\vec{\nabla} r_i\right)_k = \left(r_i(\vec{a}_k^+) - r_i(\vec{a}_k^-)\right)/2$$

$$\mathrm{Corr}[X, Y] = \frac{1}{4\Delta X \Delta Y} \sum_{j=1}^{N} (X_j^+ - X_j^-)(Y_j^+ - Y_j^-)$$

C_f is **independent** of the experimental and PDF uncertainties. In the figures, take $|C_f| \gtrsim 0.7$ to indicate a large correlation.

- $S_f \equiv |\vec{\rho}_i| cos\varphi = C_f \frac{\Delta r_i}{\langle r_0 \rangle_E}$ -- projection of $\vec{\rho}_i(\vec{a})$ on $\vec{\nabla} f$

S_f is proportional to $cos\varphi$ and the ratio of the PDF uncertainty to the experimental uncertainty. We can sum $|S_f|$.
In the figures, take $|S_f| > 0.25$ to be significant.

independent of some fitting/reweighting ambiguities – e.g., the tolerance

more tomorrow on LHC implications

today: **what about future experiments?**

e.g., HL-LHC, LHeC, and beyond…

Mapping the PDF sensitivity of future facilities (HL-LHC, LHeC, and EIC)
T.J. Hobbs, Bo-Ting Wang, Pavel Nadolsky, Fredrick Olness
arxiv:1907.00988
(more detailed work in preparation)

a high energy EIC, **LHeC**

- an electron-proton (or electron-ion) collider to achieve **high luminosities** $\gtrsim 1000$ times that of HERA

  → access a wide range of x, including $x \simeq 10^{-6}$

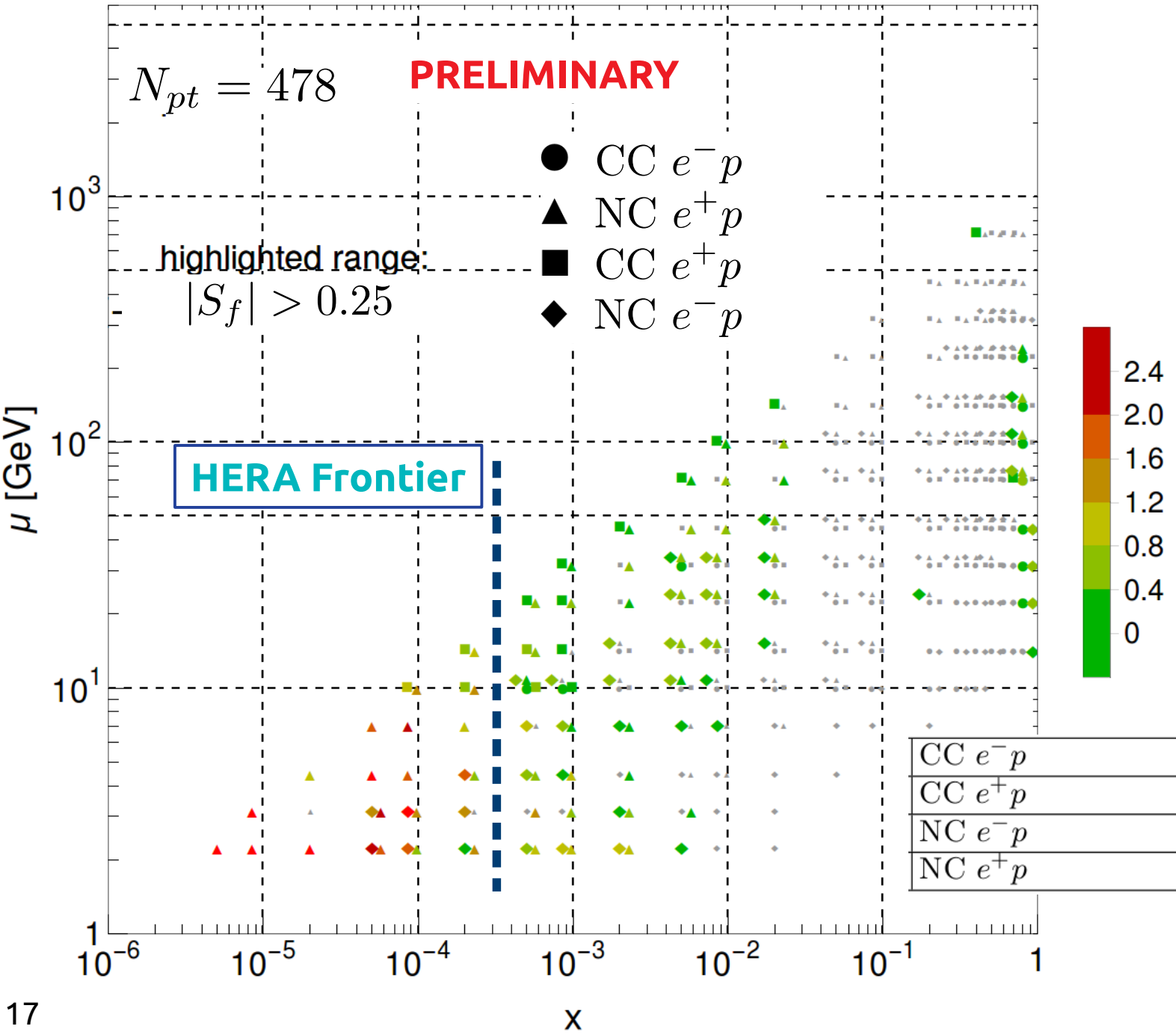
  → explore the dynamics of gluon saturation; greatly improve PDF precision; perform SM tests; and many other physics goals

- can perform a sensitivity analysis of Monte Carlo generated ep reduced NC/CC cross sections (Klein & Radescu, LHeC-Note-2013-002 PHY)

$$60\,\mathrm{GeV}\ e^{\pm}\ \mathrm{on}\ 1\ \mathrm{or}\ 7\,\mathrm{TeV}\ p$$

- to minimize the impact of large χ^2 of unfitted data (especially at low x), we study the sensitivities for **fluctuated data** – i.e., pseudodata randomly fluctuated about the PDF4LHC15 NNLO prediction according to putative LHeC uncorrelated errors – based on **100 fb^{-1}** of data

$|S_f|$ for g(x,μ), PDF4LHC15 NNLO



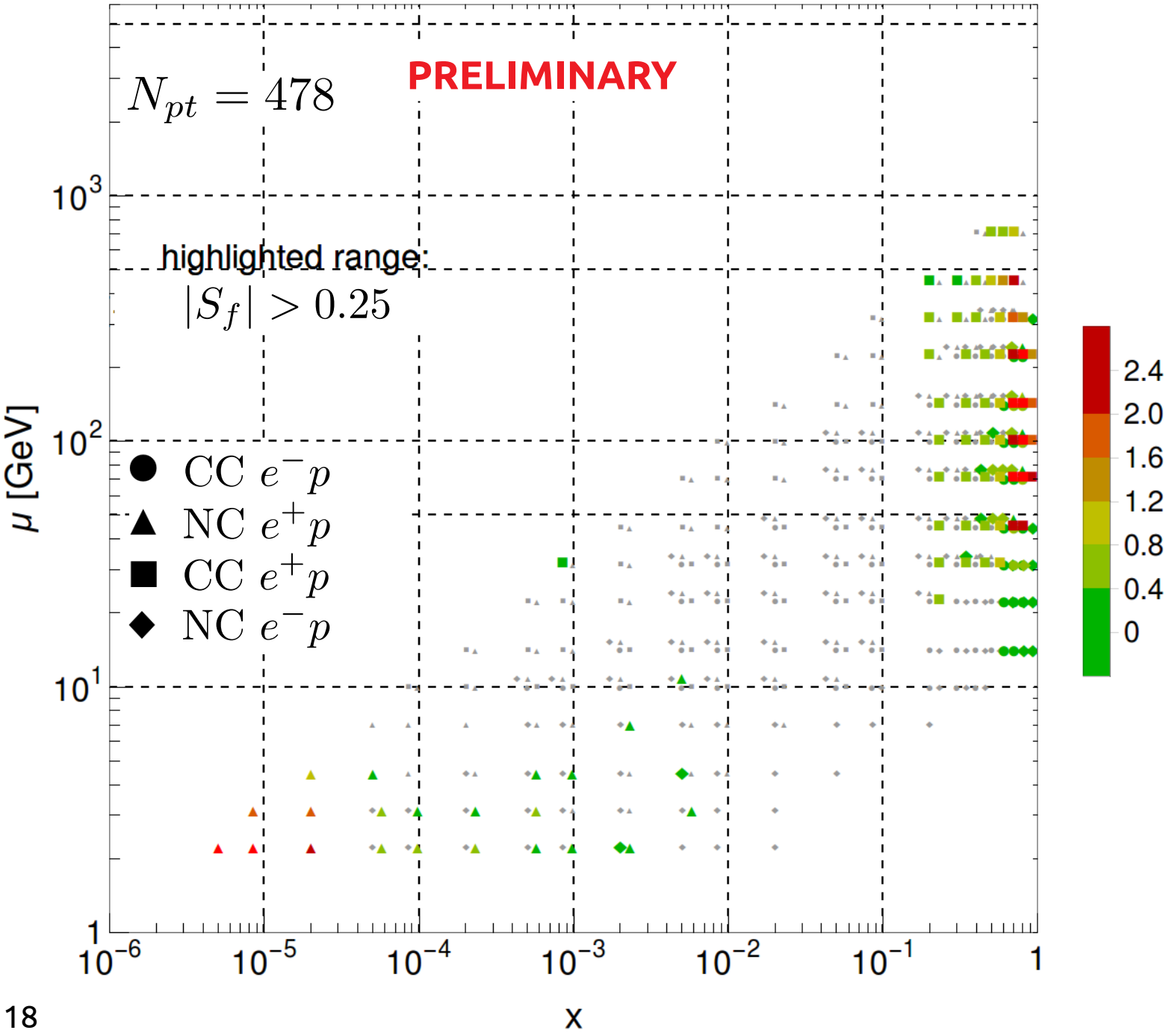
- **very strong sensitivities along the frontier of the HERA Run II data, $x \lesssim 10^{-4}$**

however, there are important constraints in general for $x \lesssim 10^{-2}$

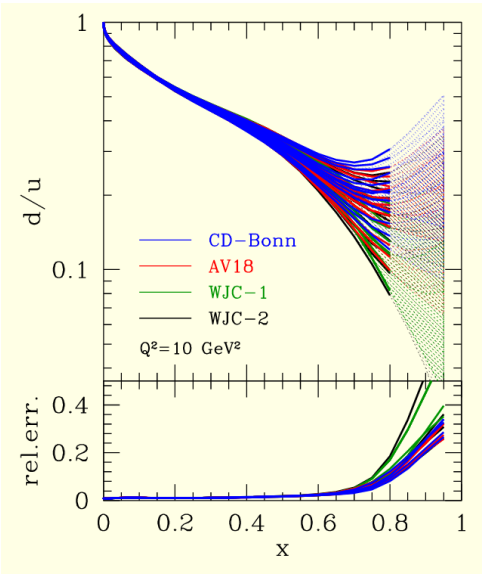
...and high x, μ

	N_{pt}	
CC e^-p	93	circle
CC e^+p	109	square
NC e^-p	128	diamond
NC e^+p	148	triangle

$|S_f|$ for d/u(x,mu), PDF4LHC15 NNLO



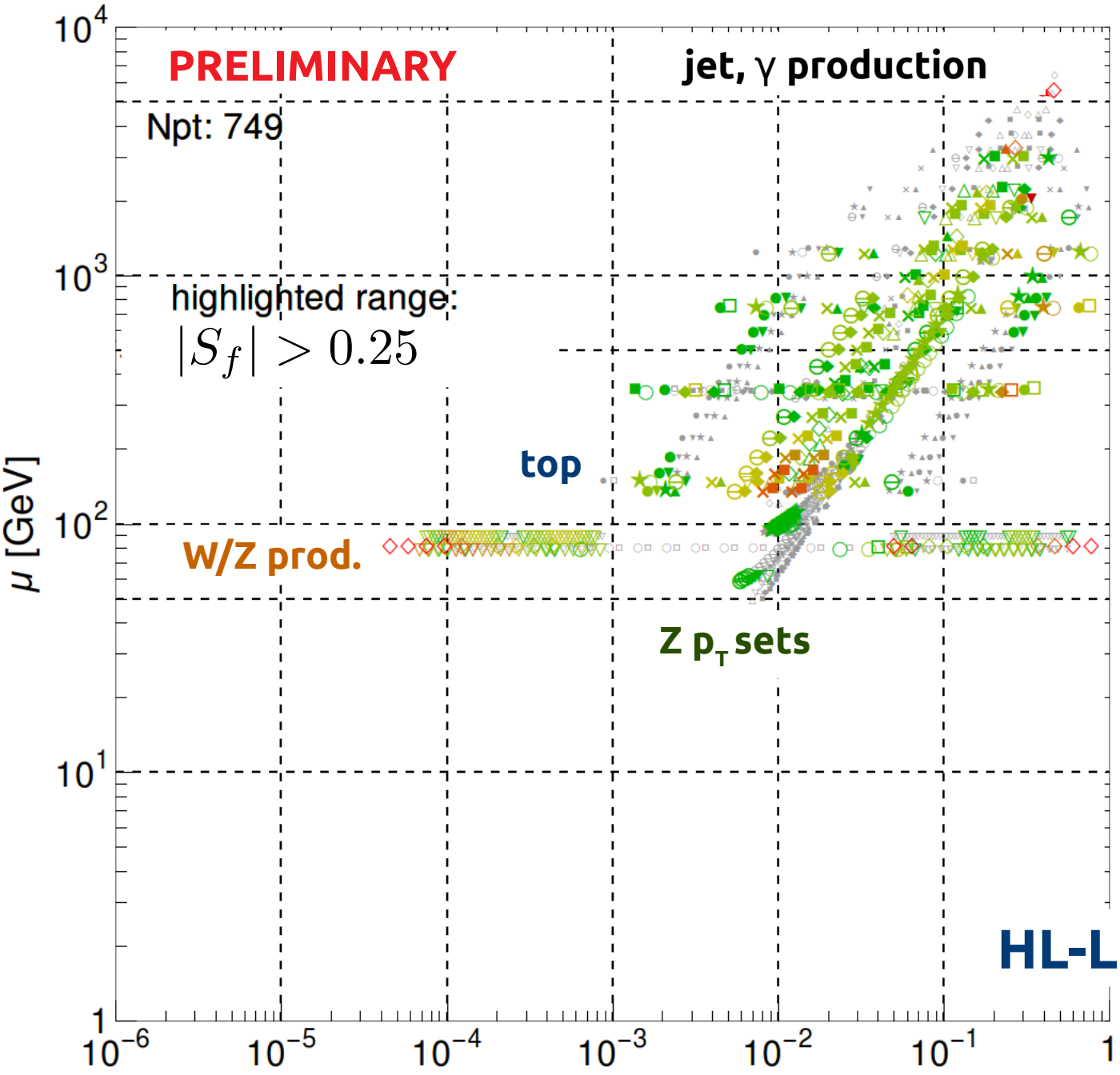
- LHeC's high luminosity may give it a reach to high enough x to help resolve the stubborn d/u question
  - …without a nuclear target…



J.F. Owens, PAVI11

| S_f | for g(x,μ), PDF4LHC15 NNLO

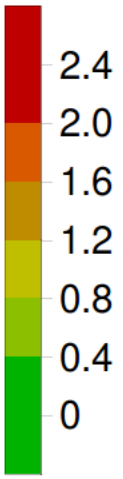
What about **HL-LHC?**



- **very <u>large</u> integrated luminosities:**

$\mathcal{L} = 3\ ab^{-1}$ (CMS/ATLAS)

$\mathcal{L} = 0.3\ ab^{-1}$ (LHCb)



Process
$Z\ p_T$
high-mass Drell-Yan
top quark pair
W+charm (central)
W+charm (forward)
Direct photon
Forward W, Z
Inclusive jets

Many thanks to **Jun Gao** for the pseudodata!

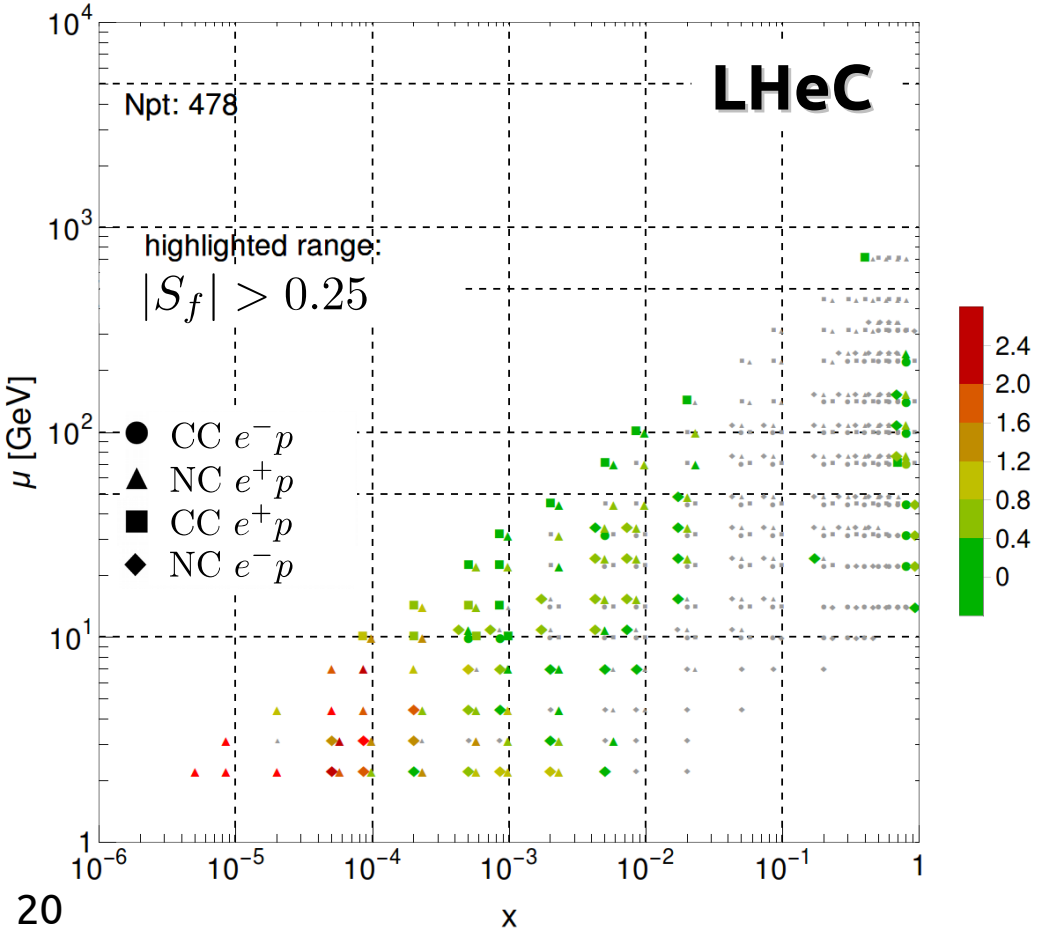
now directly compare the LHeC vs. HL-LHC flavor sensitivities*

*noting the much larger integrated luminosity of the HL-LHC pseudo-data

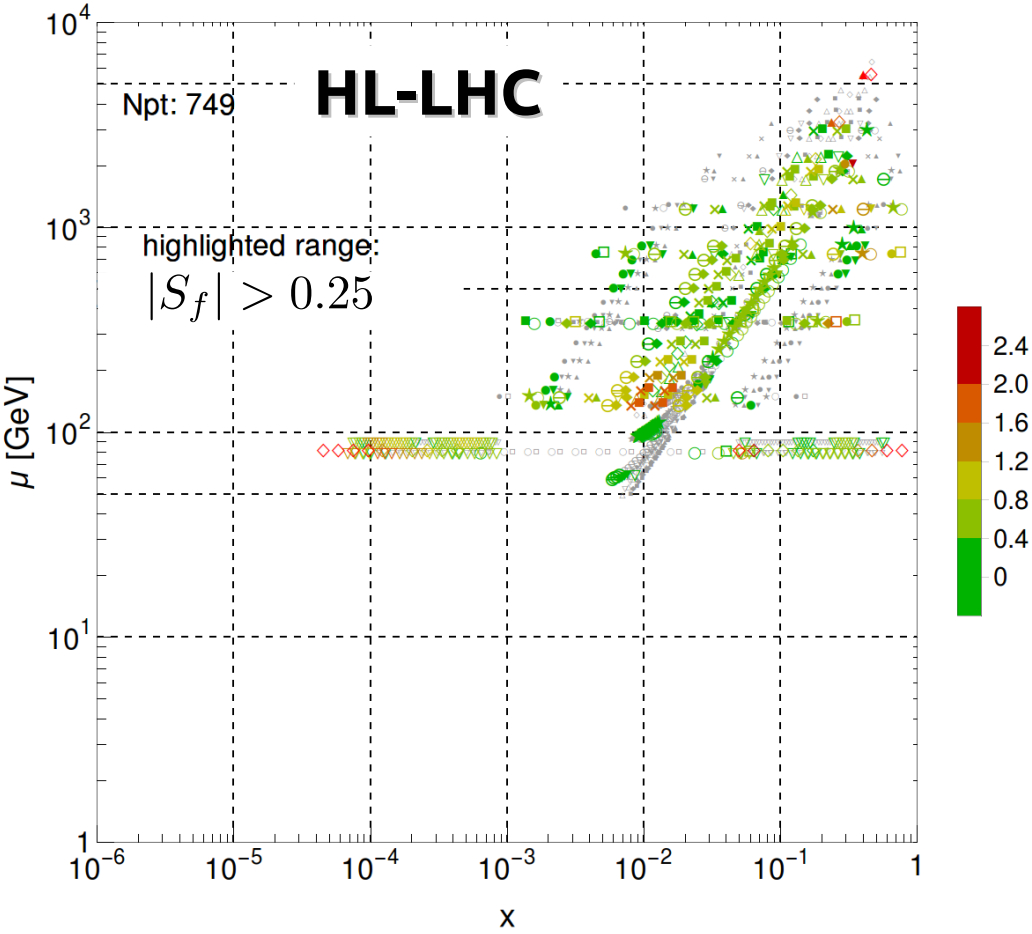
→ we compare both kinematic regions of especially strong sensitivity, and the aggregated impact of each experiment:

$$|S_g^{\mathrm{LHeC}}| = 151.4 \quad < \quad |S_g^{\mathrm{HL-LHC}}| = 244.6$$

| S_f | for g(x,μ), PDF4LHC15 NNLO



| S_f | for g(x,μ), PDF4LHC15 NNLO

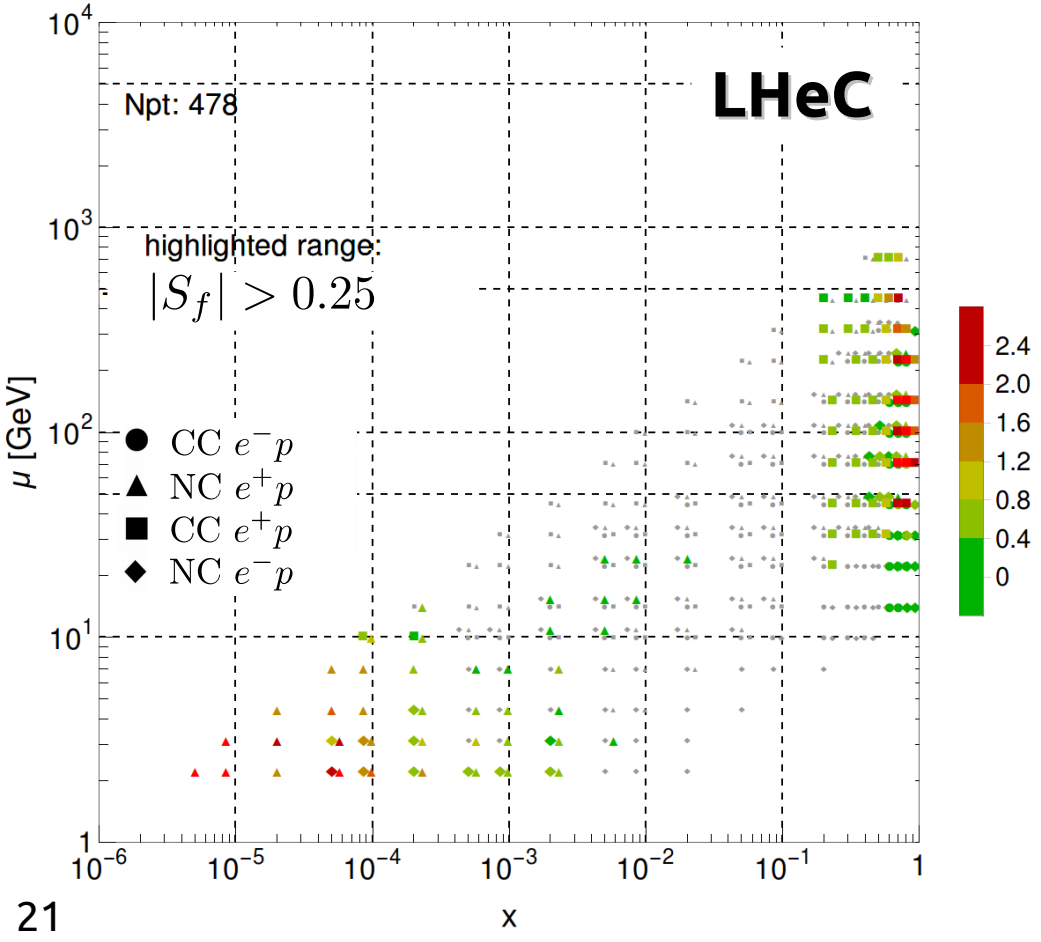


even with a small (pseudo)data set, LHeC enjoys strong sensitivity to down-type distributions!

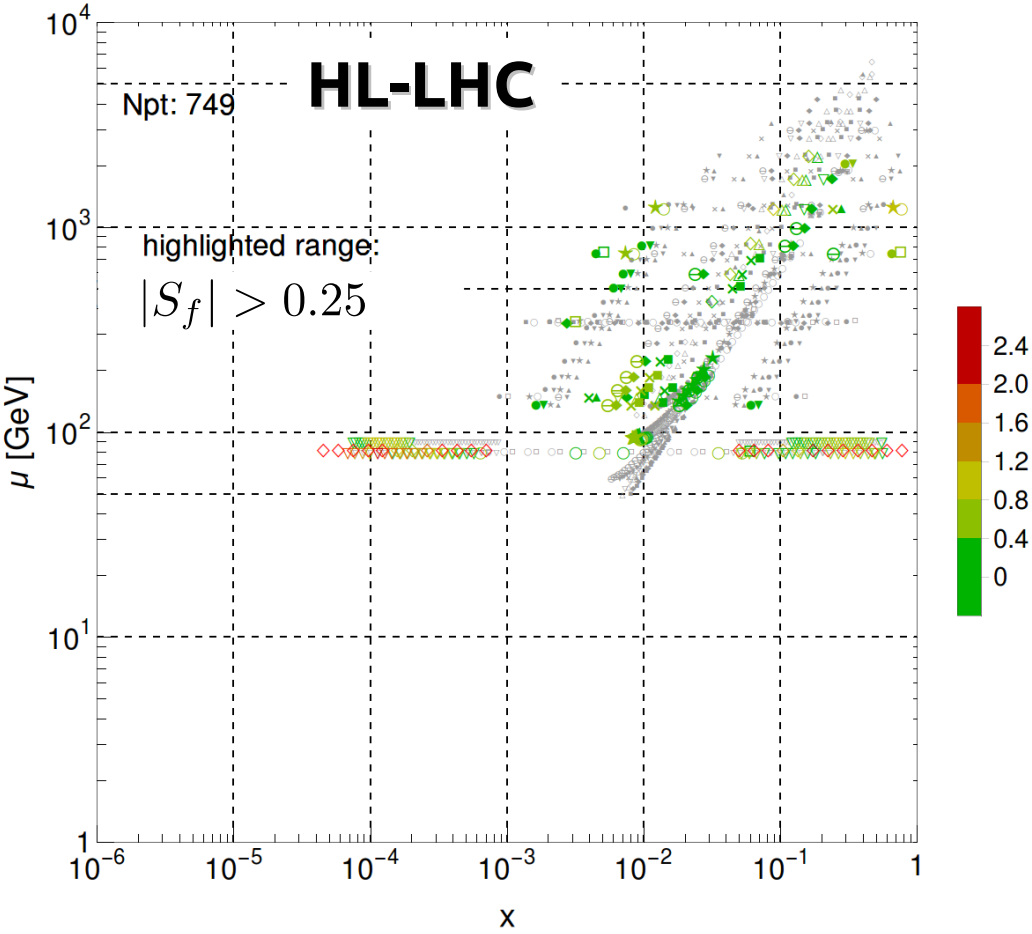
→ especially in a fashion complementary to HL-LHC, at very high/low x

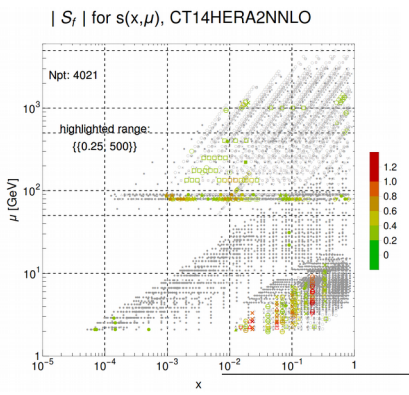
$$|S_d^{\mathrm{LHeC}}| = 214.3 \quad > \quad |S_d^{\mathrm{HL-LHC}}| = 170.8$$

| S_f | for d(x,μ), PDF4LHC15 NNLO



| S_f | for d(x,μ), PDF4LHC15 NNLO



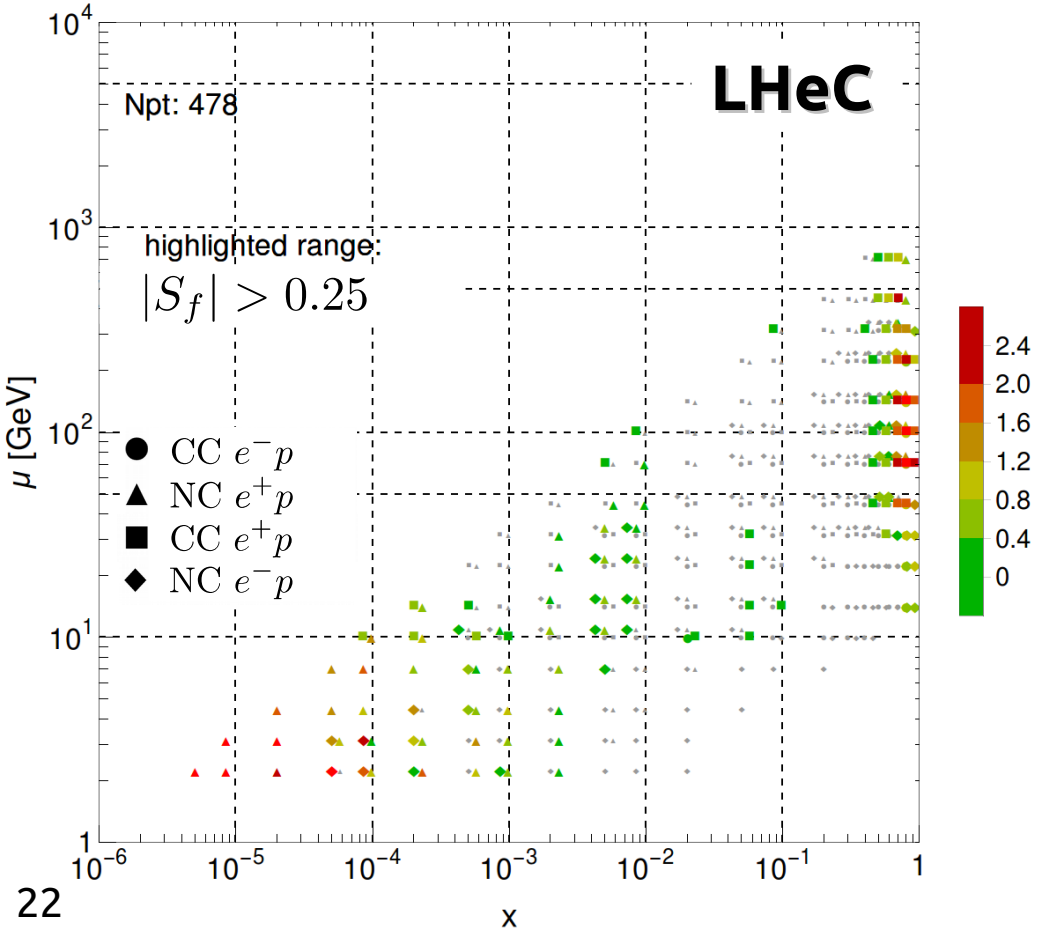


→ LHeC especially portends significantly heightened knowledge of nucleon strangeness

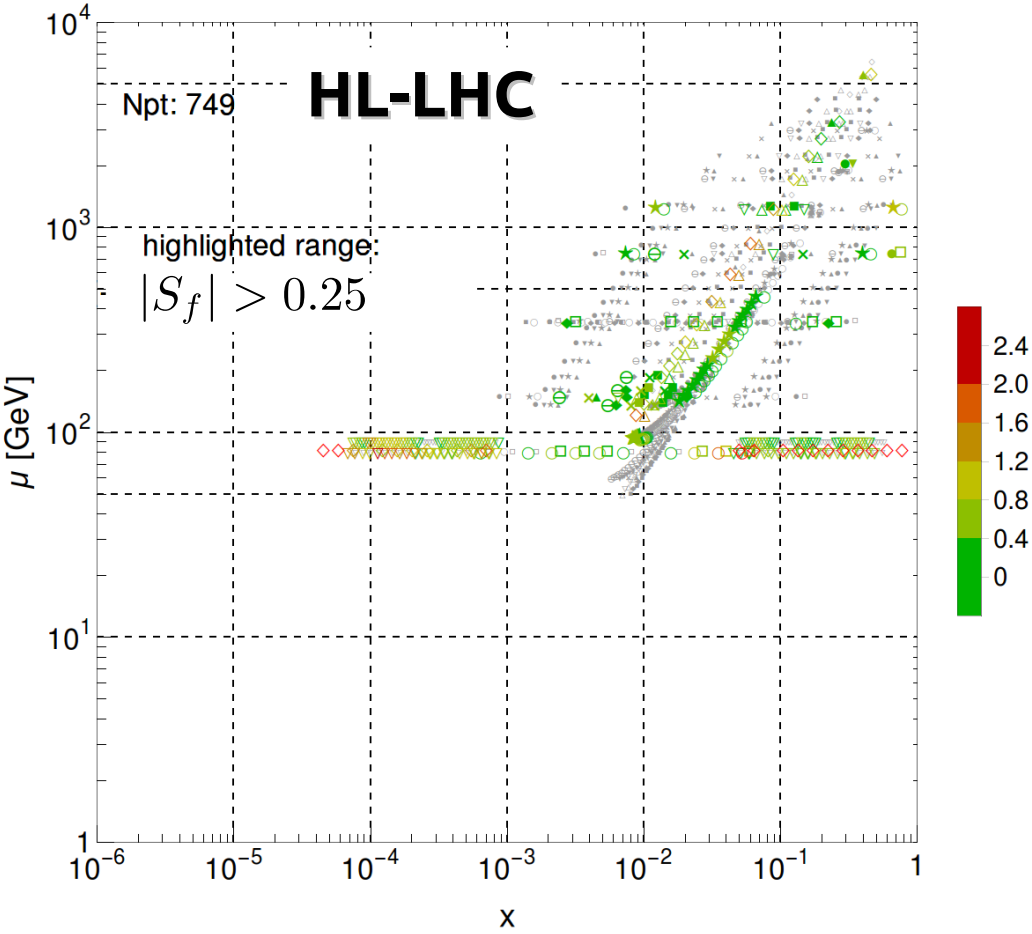
- CTEQ-TEA constraints come primarily through older fixed-target data and Tevatron data (and LHC Run I)

$$|S_s^{\mathrm{LHeC}}| = 214.1 \quad \gtrsim \quad |S_s^{\mathrm{HL-LHC}}| = 184.8$$

| S_f | for s(x,μ), PDF4LHC15 NNLO



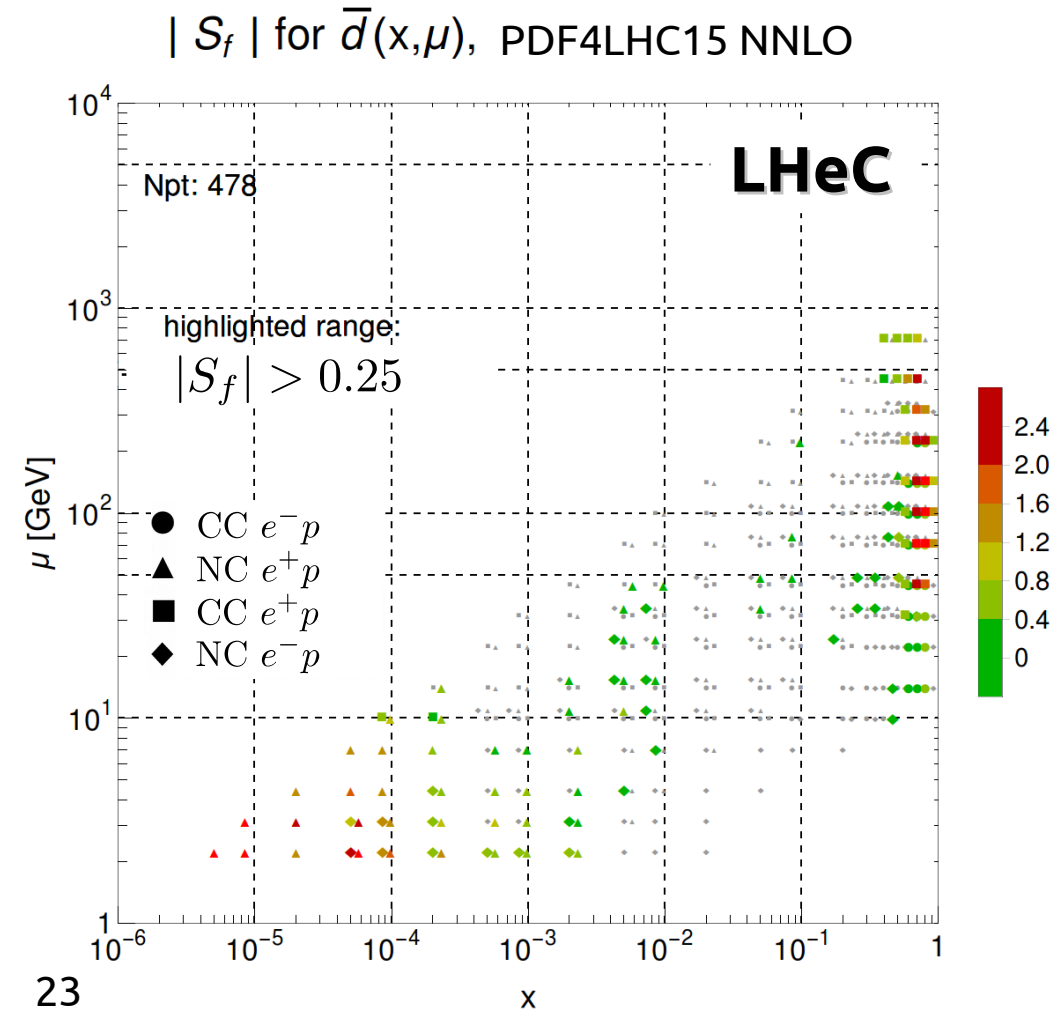
| S_f | for s(x,μ), PDF4LHC15 NNLO

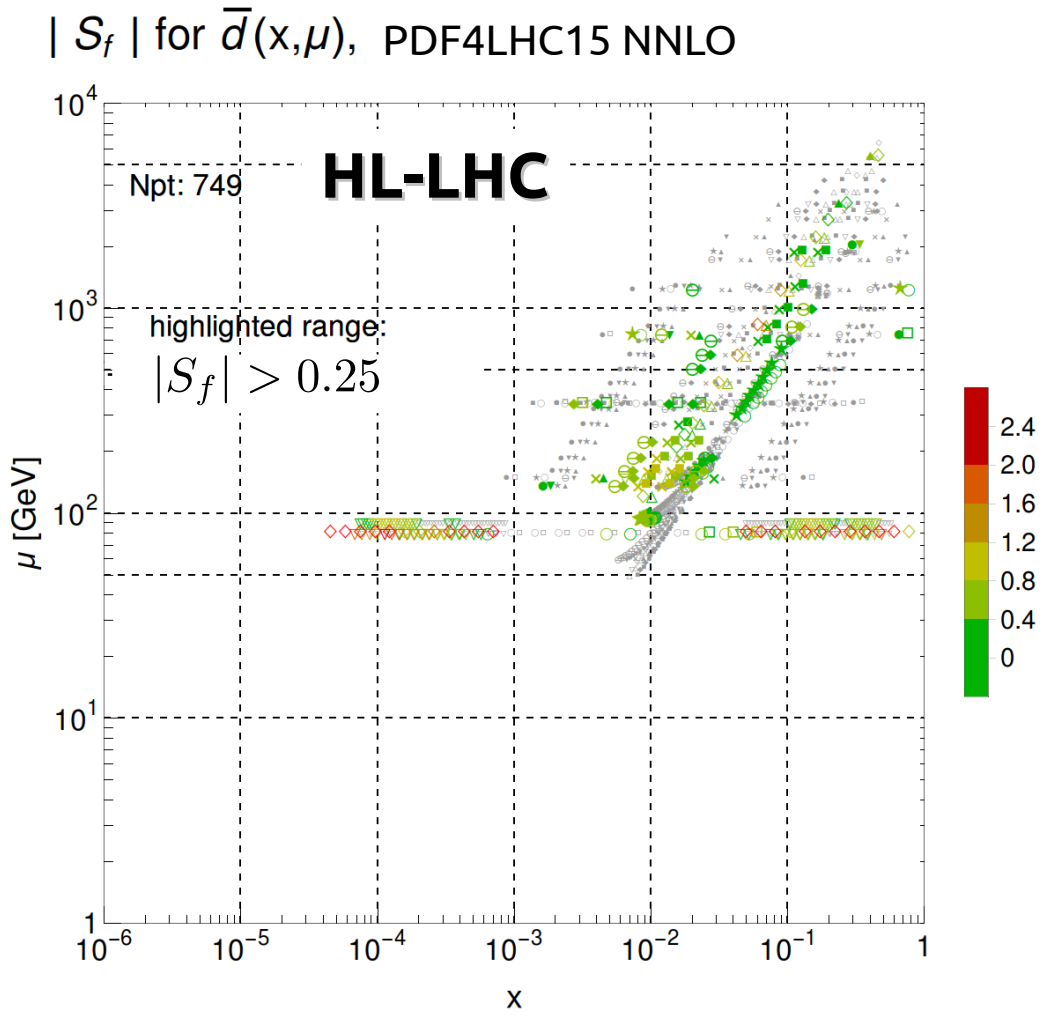


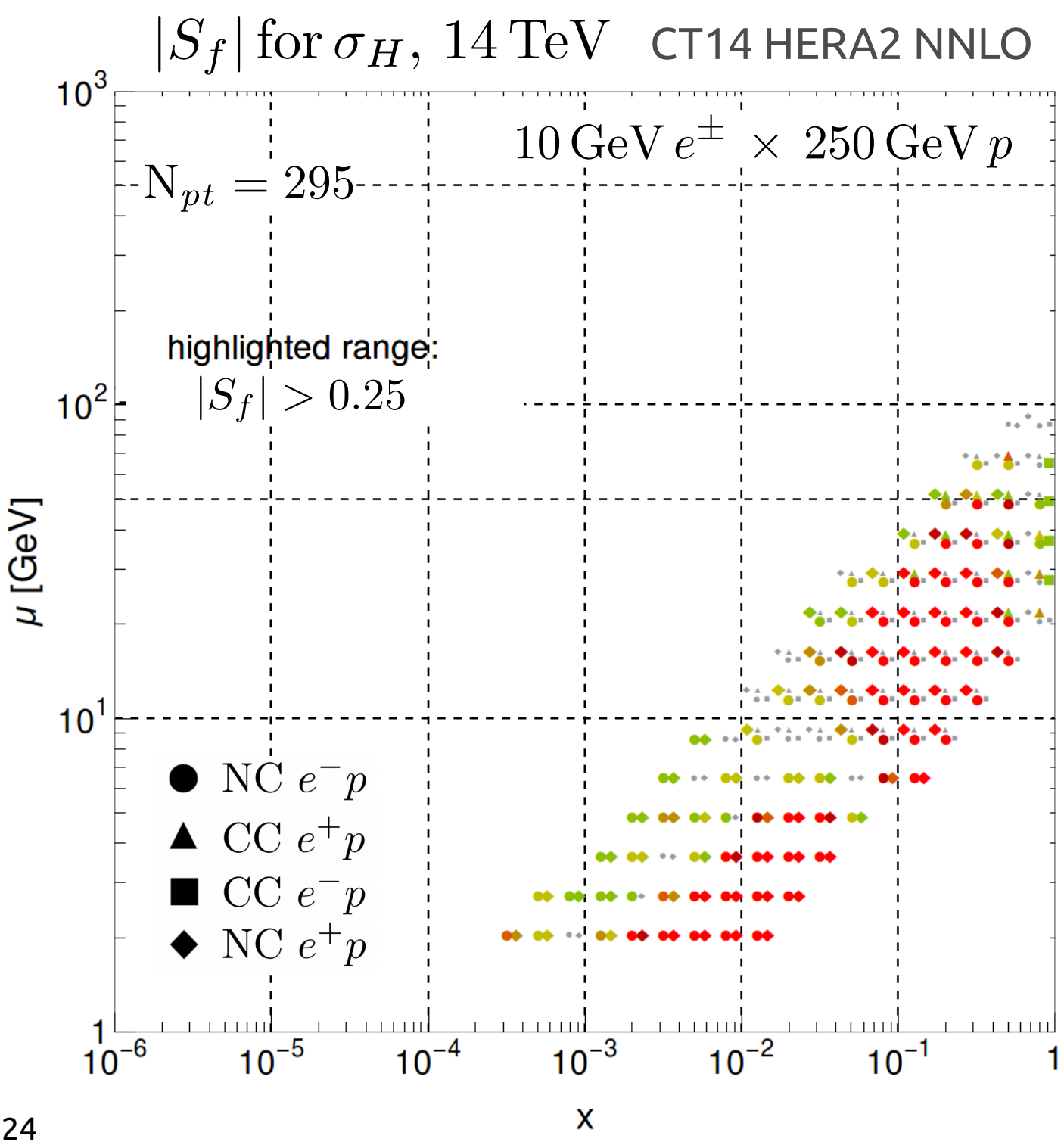
in the SU(2) quark sea, the LHeC 100 fb^{-1} set imposes constraints of magnitude comparable to HL-LHC

…these again predominate at the extrema of x

$$|S_{\bar{d}}^{\mathrm{LHeC}}| = 192.7 \quad \sim \quad |S_{\bar{d}}^{\mathrm{HL-LHC}}| = 199.4$$



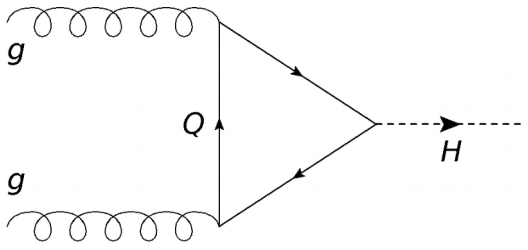




epilogue: a US-based EIC will have important HEP consequences, e.g., on Higgs physics

- the impact of an EIC upon the theoretical predictions for inclusive Higgs production arises from a very broad region of the kinematical space it can access

- impact rather closely tied to that of the integrated gluon PDF:



- lattice QCD calculations continue to improve and will be increasingly useful as inputs into QCD global analyses

PDF-Lattice whitepaper – Lin et al., PPNP100, 107 (2018); arXiv:1711.07916.

- the PDF-Lattice relationship will be *synergistic* :

→ PDF phenomenologists deliver improving benchmarks to challenge the Lattice

↕

→ Lattice calculations for PDF Mellin **moments** and **quasi-PDFs** can be theoretical priors for QCD global fits

PDFSense analysis – Hobbs, Wang, Nadolsky and Olness, arXiv:1904.00022.

- moments from lattice can help unravel PDF flavor dependence, constrain phenom. PDFs:

(i) $$\langle x^n \rangle_q = \int_0^1 dx\ x^n \left[q(x) + (-1)^{n+1}\, \overline{q}(x) \right] \ \rightarrow \langle x^{1,3,\ldots} \rangle_{q^+},\ \langle x^{2,4,\ldots} \rangle_{q^-}$$

$$\mu_F = \mu^{\text{lat.}} = 2\,\text{GeV}$$

- lattice can also now compute x-dependent quantities – the quasi-PDFs (qPDFs):

(ii) $$\widetilde{q}(x, P_z, \widetilde{\mu}) = \int_{-\infty}^{\infty} \frac{dz}{4\pi} e^{ixP_z z} \langle P | \overline{\psi}(z) \gamma^z U(z,0) \psi(0) | P \rangle$$

(i) PDF moments can be evaluated on the QCD lattice

- in lattice gauge theory, the accessible moments are C-odd/even combinations,

$$\langle x^n \rangle_q = \int_0^1 dx\; x^n \left[q(x) + (-1)^{n+1}\, \overline{q}(x) \right]$$

$$\begin{cases} \langle x^n \rangle_{q^+} = \langle x^n \rangle_q \quad \text{for} \quad n = 2\ell - 1 \\ \langle x^n \rangle_{q^-} = \langle x^n \rangle_q \quad \text{for} \quad n = 2\ell \end{cases} \qquad \ell \in \mathbb{Z}^+$$

- the PDF moments are related to hadronic matrix elements of twist-2 operators:

PDF Mellin moments

$$\frac{1}{2} \sum_s \langle p, s | \mathcal{O}^q_{\{\mu_1, \cdots, \mu_{n+1}\}} | p, s \rangle = 2 v_q^{n+1} \left[p_{\mu_1} \cdots p_{\mu_{n+1}} - \text{traces} \right]$$

$$\mathcal{O}^q_{\{\mu_1, \cdots, \mu_{n+1}\}} = \left(\frac{i}{2} \right)^n \bar{q}(x) \gamma_{\mu_1} \overleftrightarrow{D}_{\mu_2} \cdots \overleftrightarrow{D}_{\mu_{n+1}} q(x)$$

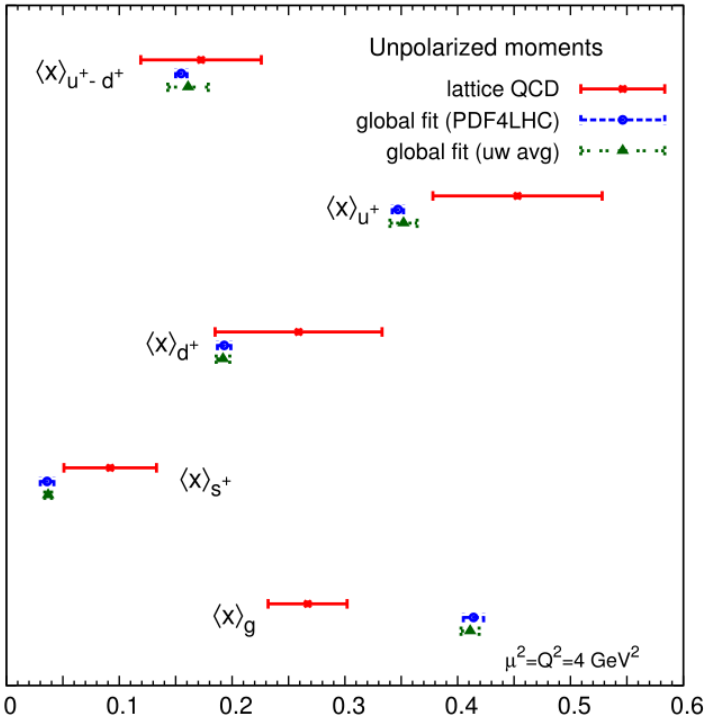
gauge-covariant derivatives

the status of lattice QCD calculations

PDF-Lattice whitepaper – Lin et al., Prog. Part. Nucl. Phys. **100**, 107 (2018).

Mom.	Collab.	Ref.	N_f	Status	Disc [fm]	QM	FV	Ren	ES
$\langle x \rangle_{u^+-d^+}$	ETMC 15	[263]	2+1+1	P	0.06, 0.08	–	■, ★	★, ★	■, ★
	ETMC 15	[263]	2	P	0.06–0.09	–	○	★	■
	RQCD 14	[251]	2	P	0.06–0.08	–	○	★	○

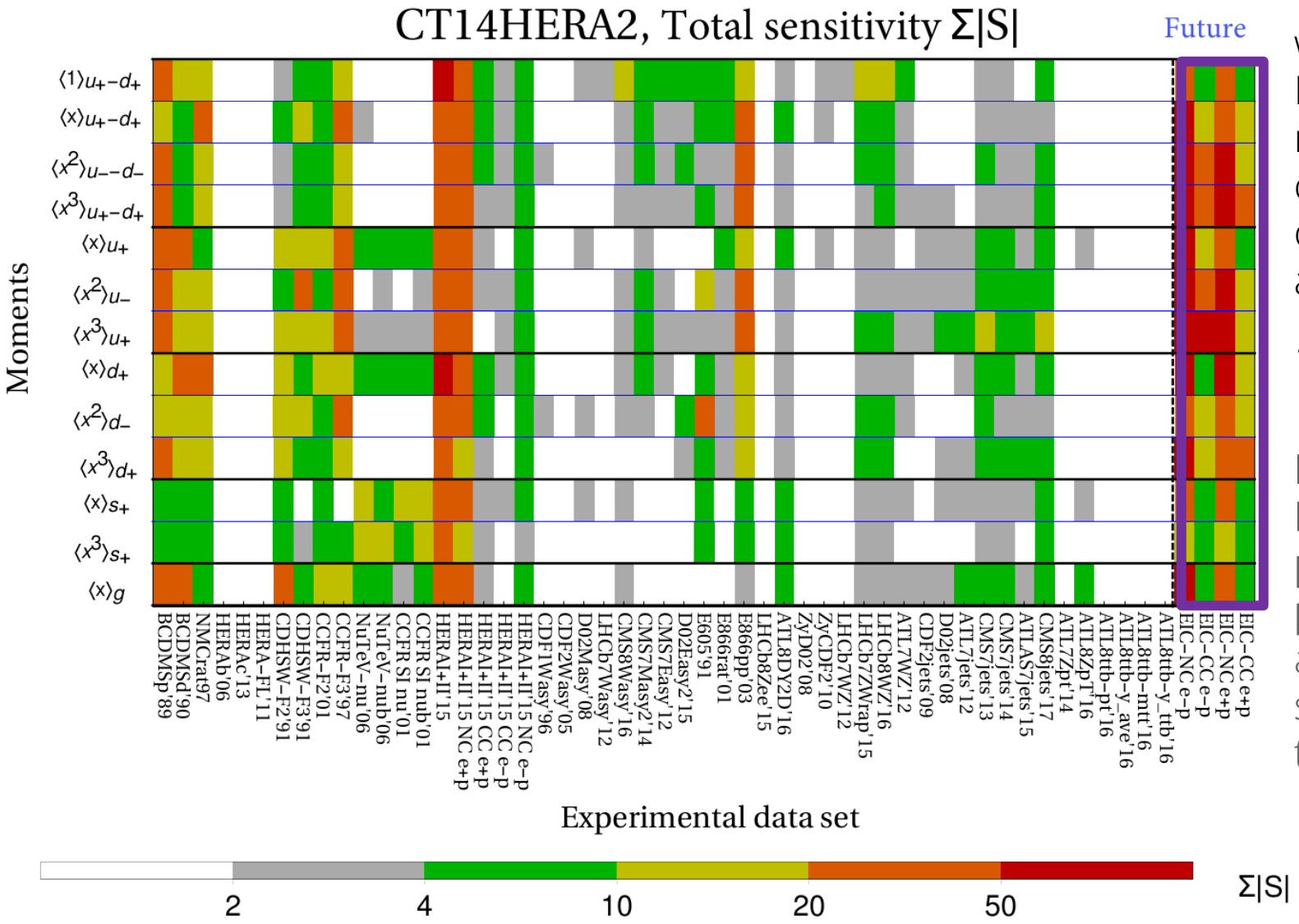
Mom.	Collab.	Ref.	N_f	Status	Disc	QM	FV	Ren	ES	
$\langle x^2 \rangle_{u^--d^-}$	LHPC and SESAM 02	[279]	2	P	■	■	■	○	■	0.145(69)
	QCDSF 05	[93]	0	P	■	■	■	★	■	0.083(17)
	LHPC and SESAM 02	[279]	0	P	■	■	■	○	■	0.090(68)



- depending upon flavor and order, lattice extractions of Mellin moments have varying status (above, FLAG evaluations)
  - → e.g., the first isovector moment has been computed by numerous groups
  - → but the second, by relatively few
- systematic lattice effects are similarly widely varied

however, improvements are being made rapidly!

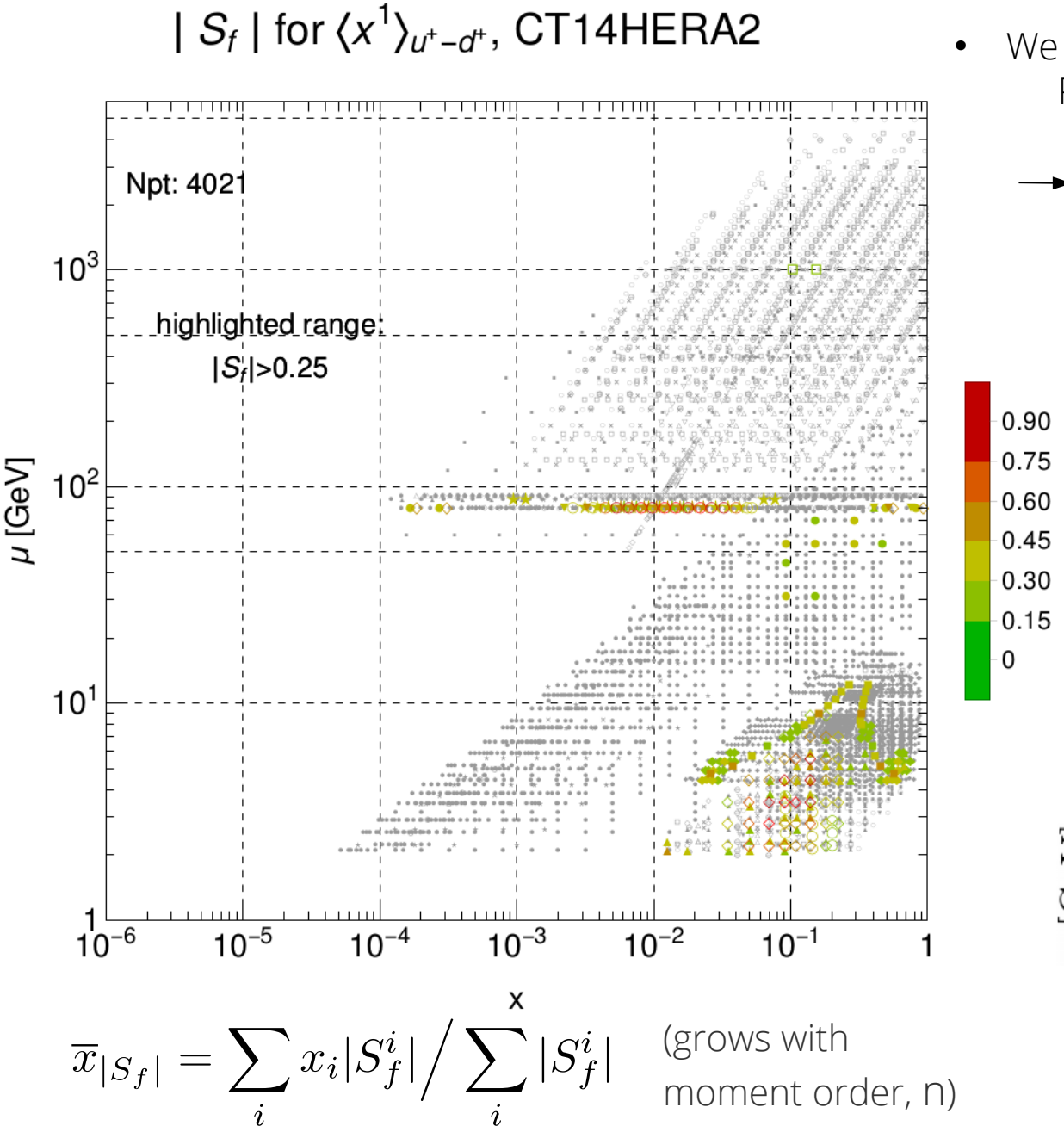
sensitivity to Mellin moments



we show Mellin moments computable on the lattice at the scale $\mu_F = 2\,\mathrm{GeV}$

HERA, BCDMS, NMC, E866 DY pair production are most sensitive to the moments

sensitivity maps: isovector moments



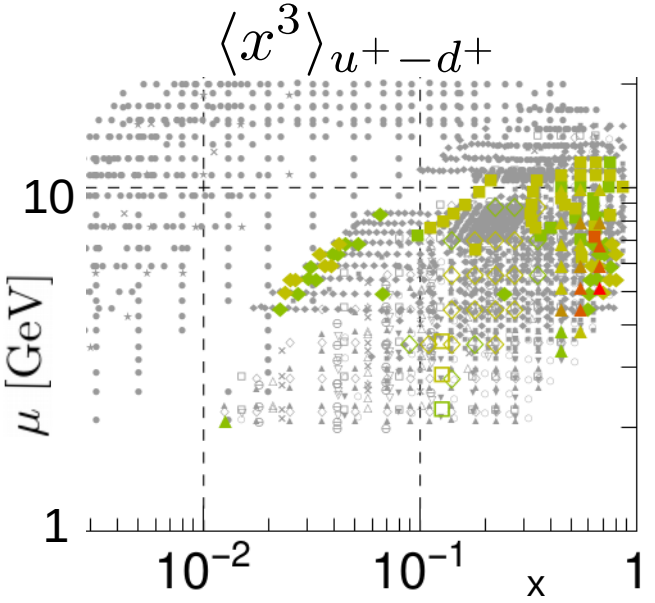
$$\overline{x}_{|S_f|} = \sum_i x_i |S_f^i| \Big/ \sum_i |S_f^i|$$

(grows with moment order, n)

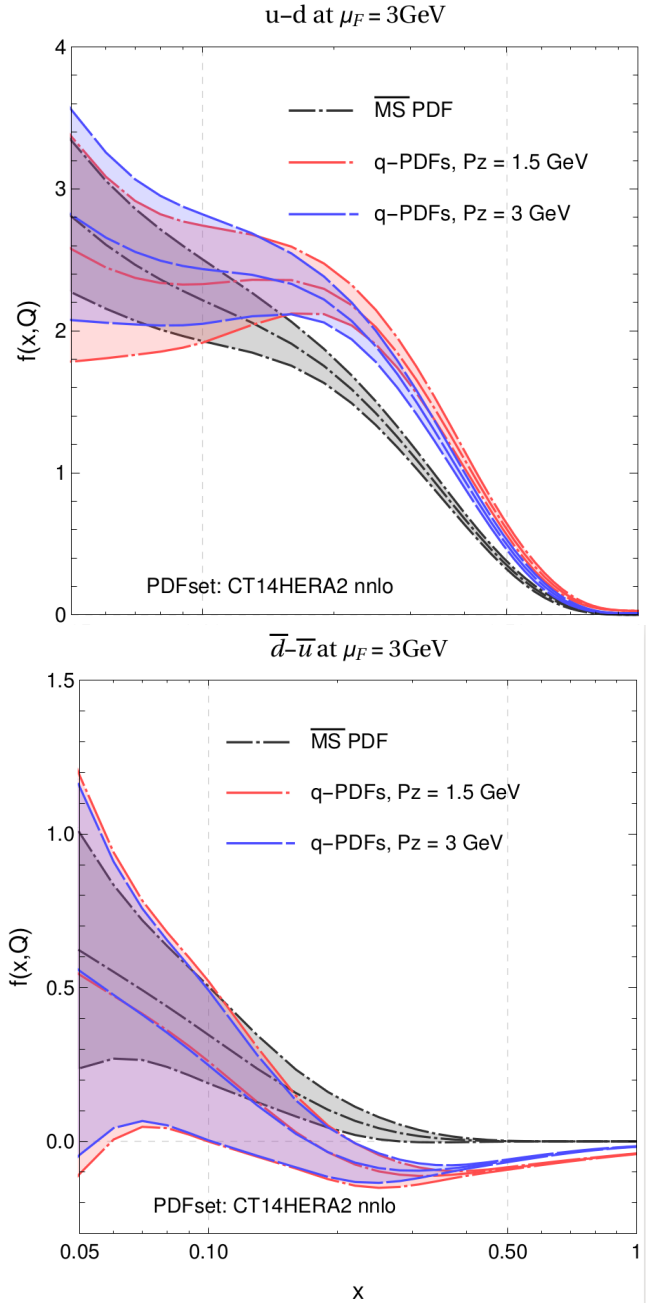
- We focus on **isovector** (u-d) PDF combinations

→ on the lattice, these are more readily computed since flavor non-singlet combinations do not receive disconnected insertions

- Moments of higher order are constrained by higher x_i fixed-target data:

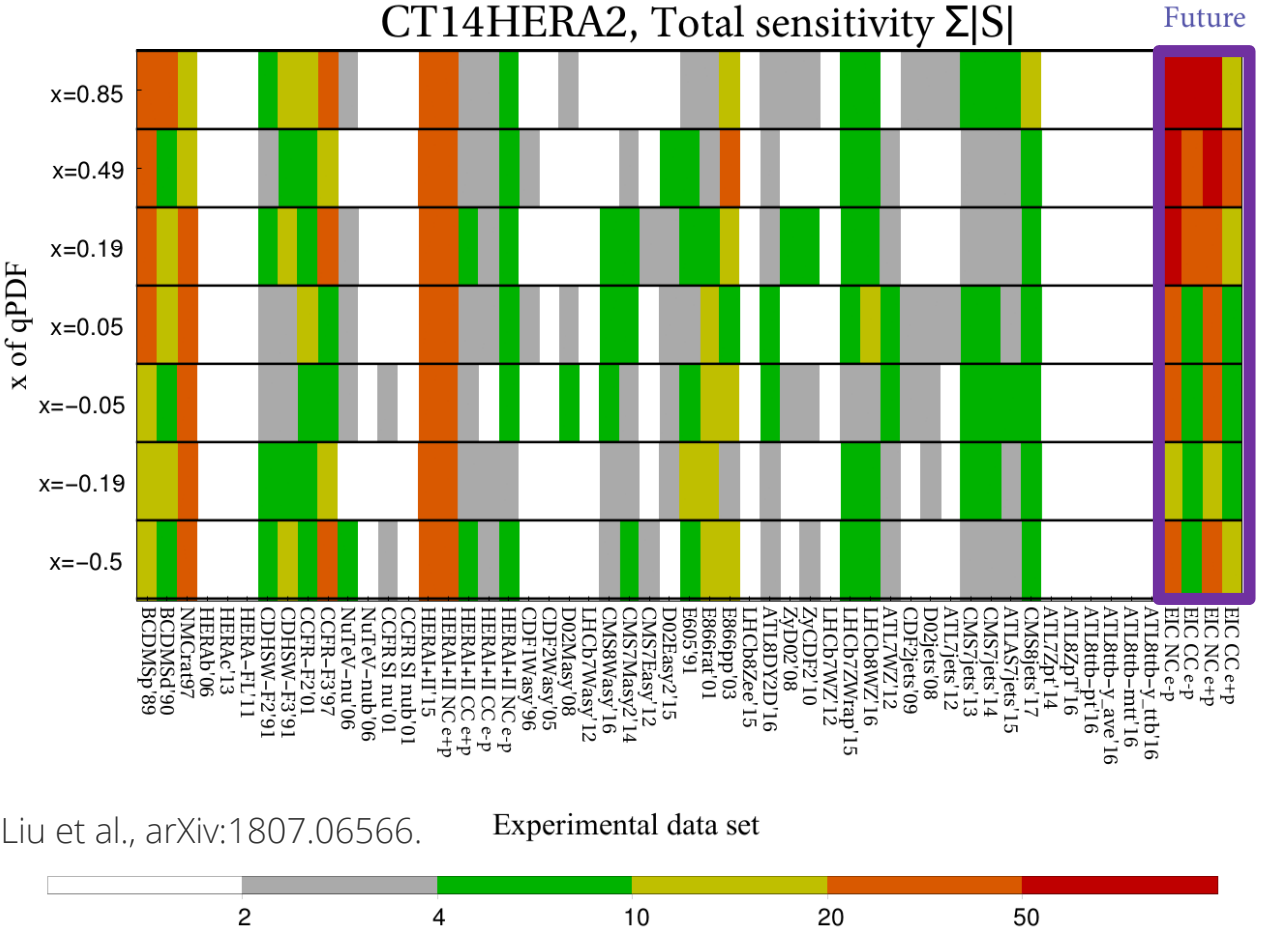


total sensitivity to matched quasi-PDFs



$$\overline{q}(x) = -q(-x) \;\rightarrow\; q(x) \equiv \begin{cases} u(x) - d(x), & x > 0 \\ \bar{d}(|x|) - \bar{u}(|x|), & x < 0 \end{cases}$$

$$P_z = 1.5\,\text{GeV};\;\; \mu_F = 3\,\text{GeV}$$

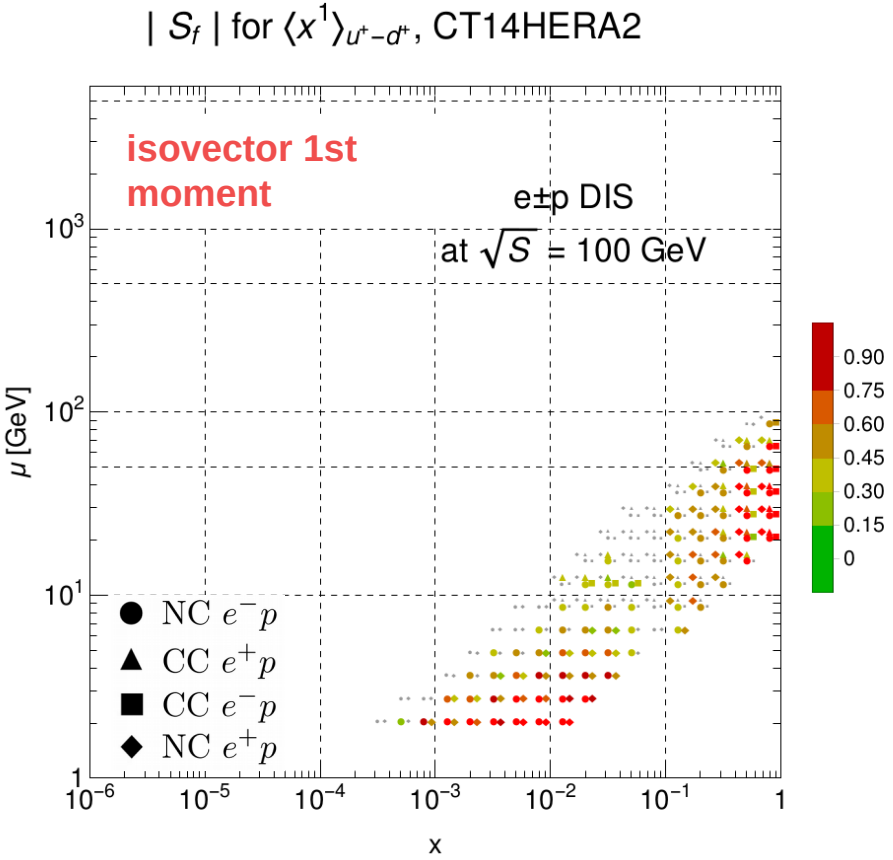


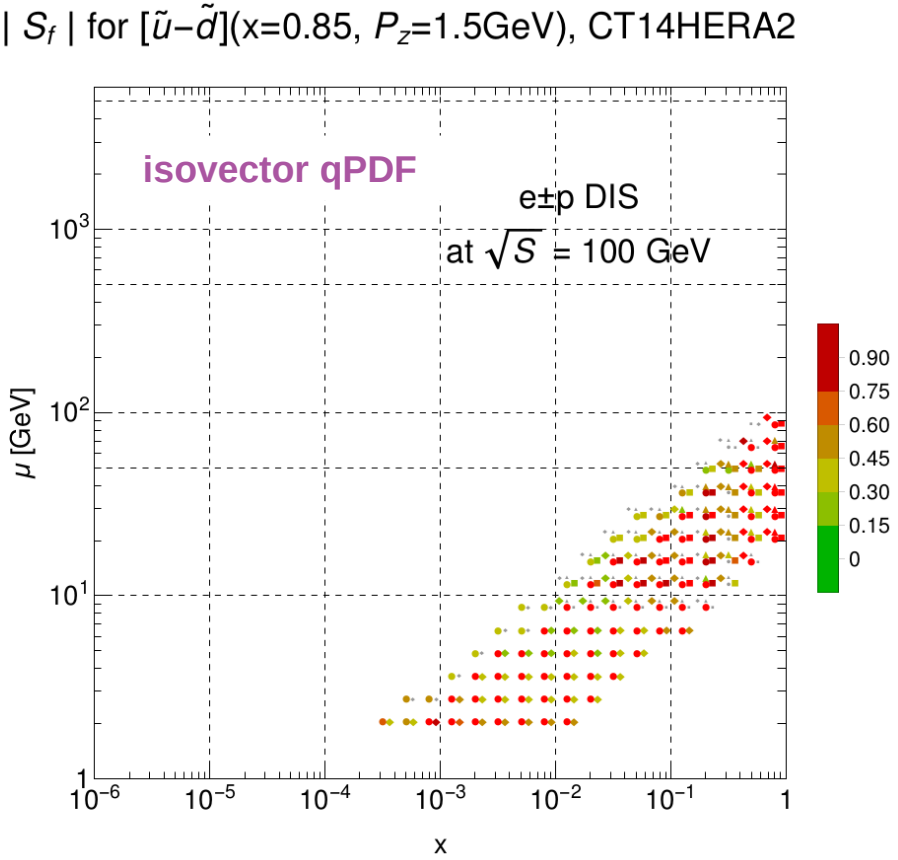
Liu et al., arXiv:1807.06566.

An EIC would drive lattice phenomenology

- A high-luminosity lepton-hadron collider will impose very tight constraints on many lattice observables; below, the isovector first moment and qPDF

- Many of the experiments most sensitive to PDF Mellin moments and qPDFs involve nuclear targets → eA data from EIC would sharpen knowledge of nuclear corrections





→ "EIC impact on tomography and HEP phenom.," Session **C09**: Sat., April 13th – 1:30pm