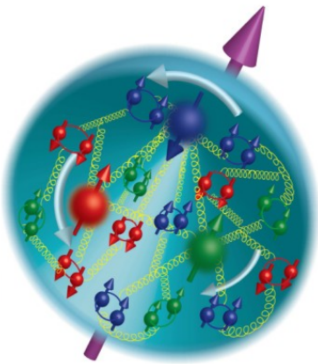
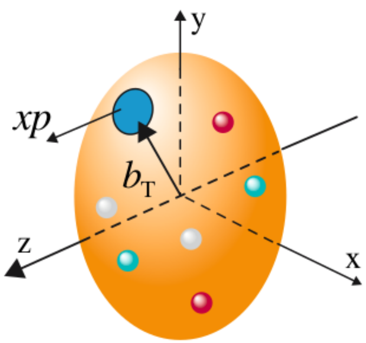
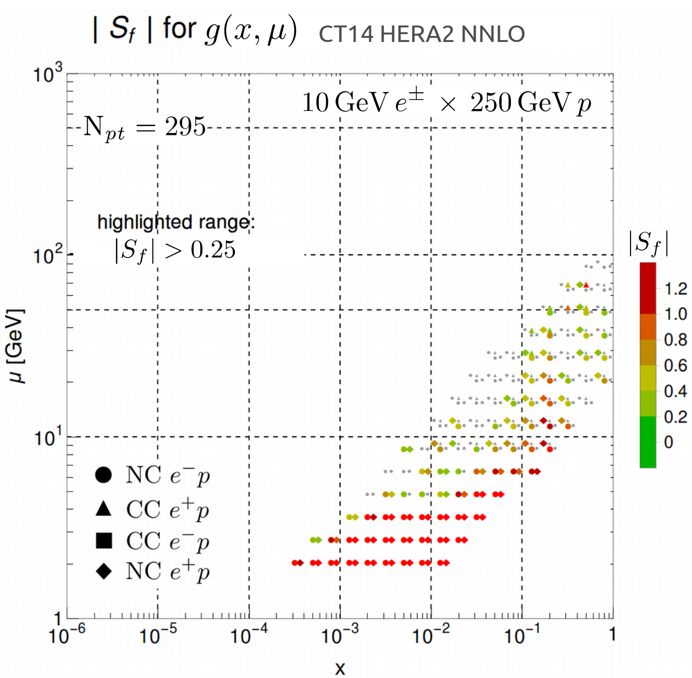
Update – ongoing PDF studies for EIC IRG

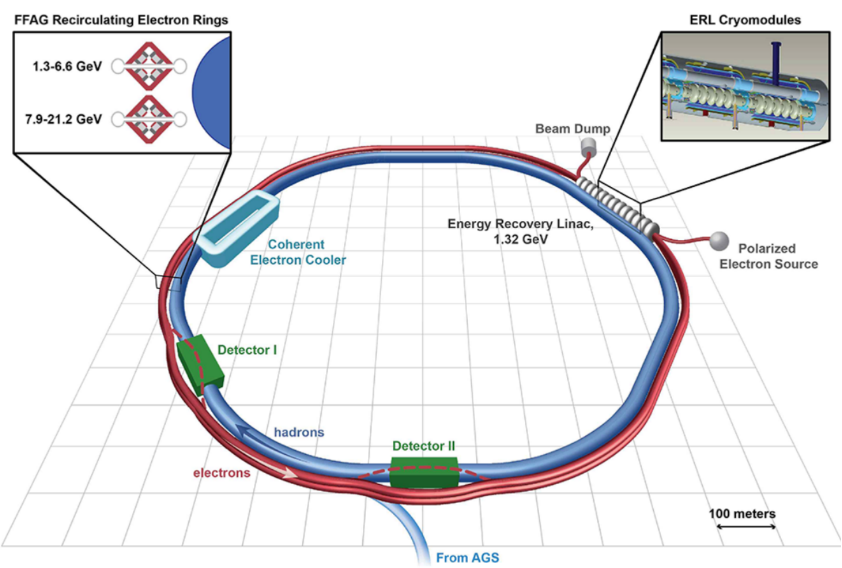
Tim Hobbs, JLab EIC Center & CTEQ@SMU

May 21st 2020







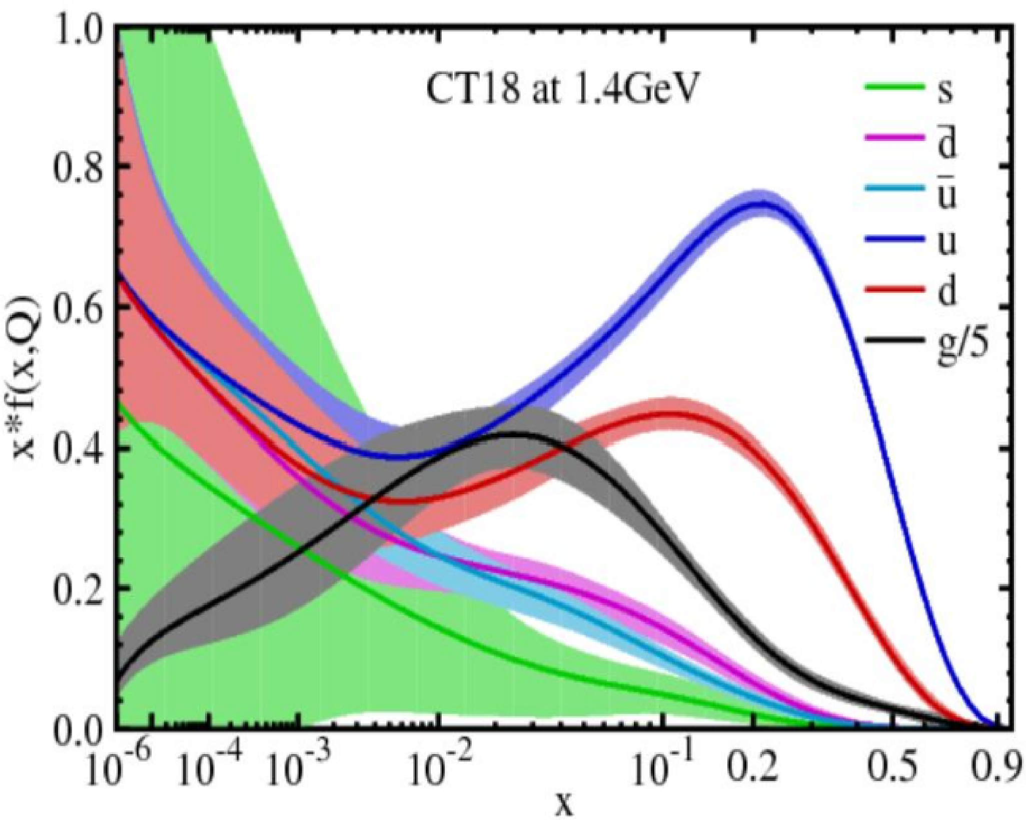


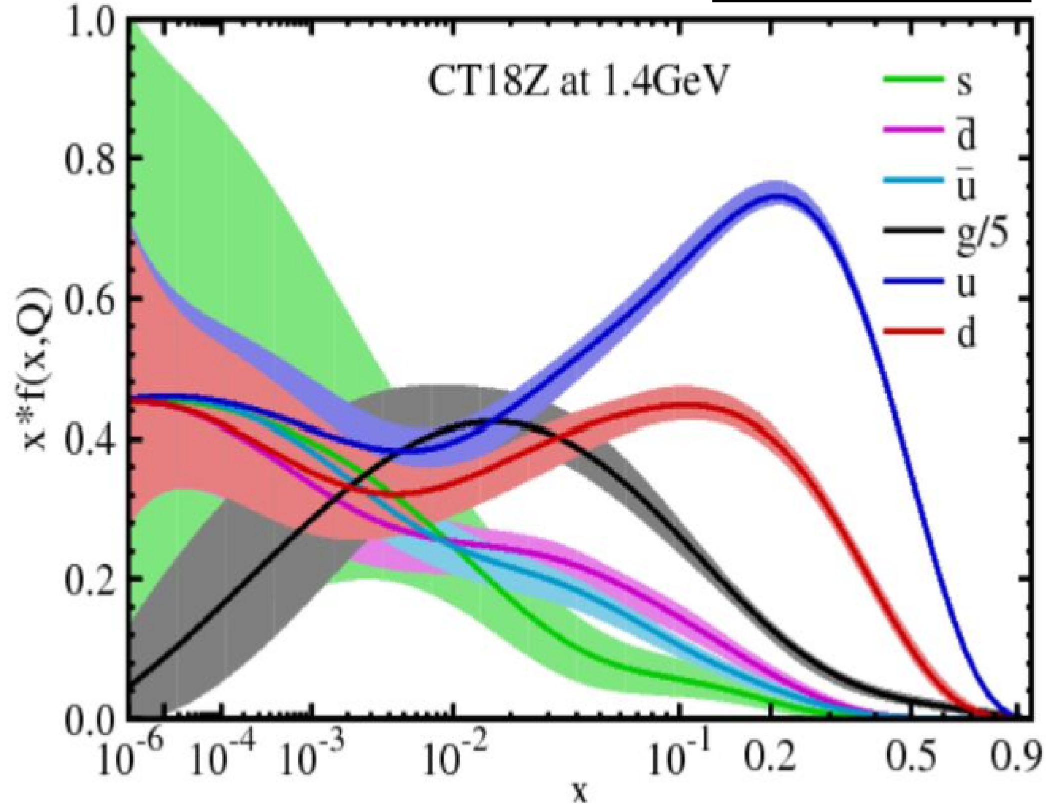
C T E Q

CT18 parton distributions

- focus remains on PDF impacts of inclusive and heavy-flavor tagged EIC data

arxiv:[1912.10053]





- advocate fast methods in addition to event-level studies and eventual fitting

[Hessian profiling with ePump and PDFSense]

developments: (i) ongoing fits; (ii) event-level studies; (iii) theory predictions

we have a dedicated effort to explore PDF impacts of the EIC

Collaborators and consultants

Pavel Nadolsky Fred Olness Bo-Ting Wang Steve Sekula	Southern Methodist
Sayipjamal Dulat	Xinjiang Univ.
C.-P. Yuan	Michigan State
Migual Arratia	UC Riverside
Alberto Accardi Yulia Furletova	Jefferson Lab

- the goal: use recently-developed tools for PDF global analyses to examine the PDF pulls of EIC pseudodata

1803.02777
1806.07950
1904.00022
1907.00988
2001.07862

tools and EIC apps

- needed for the Yellow Report Initiative: <u>quick, unambiguous</u> PDF impact metrics

→ in turn, these can be incorporated into the YR workflow:

iteratively, **machine design → simulation → physics**

→ speed, simplicity, clarity can help ensure robust and timely convergence

high-energy EIC pseudodata

- reach in center-of-mass energy, $20 \leq \sqrt{s} \leq \underline{140\,\mathrm{GeV}}$

→ luminosities 2-3 decades greater than at HERA

→ á la HERA, the combination of precision & kinematic coverage provide constraining 'lever arm' on QCD evolution

→ QCD evolution: (**high x, low Q**) ↔ (**low x, high Q**)

- as a generic scenario, we consider here the simulated impact of a machine with: $10\,\mathrm{GeV}\,e^{\pm}\ \mathrm{on}\ 250\,\mathrm{GeV}\,p \quad (\sqrt{s} = 100\,\mathrm{GeV})$

~year of data-taking

$$\left\{\begin{array}{l} \mathcal{L} = 100\,\mathrm{fb}^{-1}\ e^{-}\ \text{pseudodata} \\ \mathcal{L} = 10\,\mathrm{fb}^{-1}\ e^{+}\ \text{pseudodata} \end{array}\right.$$

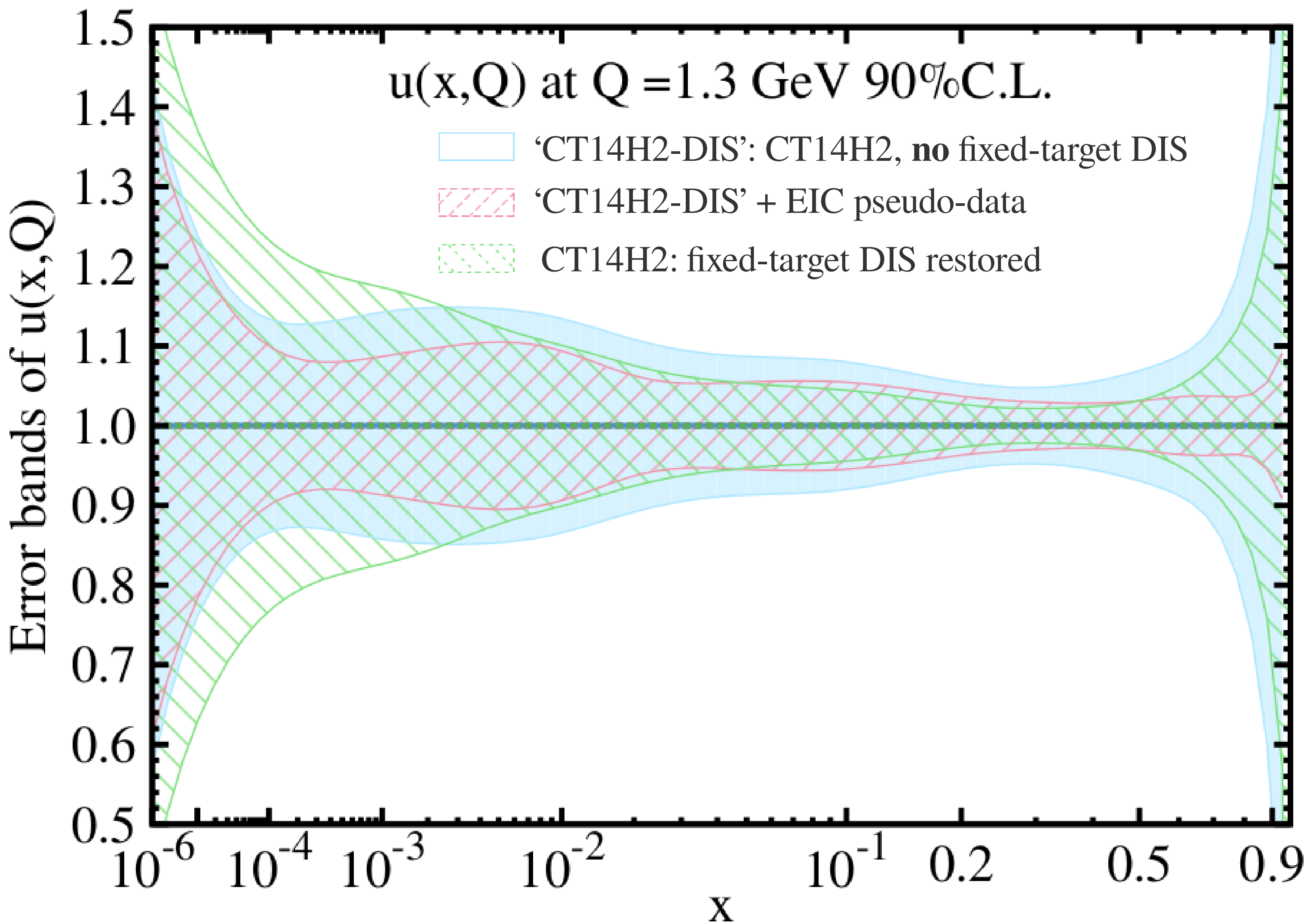
NC/CC

→ currently, <u>proton scattering only</u>

- generated based on CT14$_{\mathrm{HERA2}}$ NNLO PDF fit

Hessian profiling [ePump] for EIC impacts on PDF errors

ePump: Schmidt, Pumplin, and Yuan; PRD98 (2018) no.9, 094005

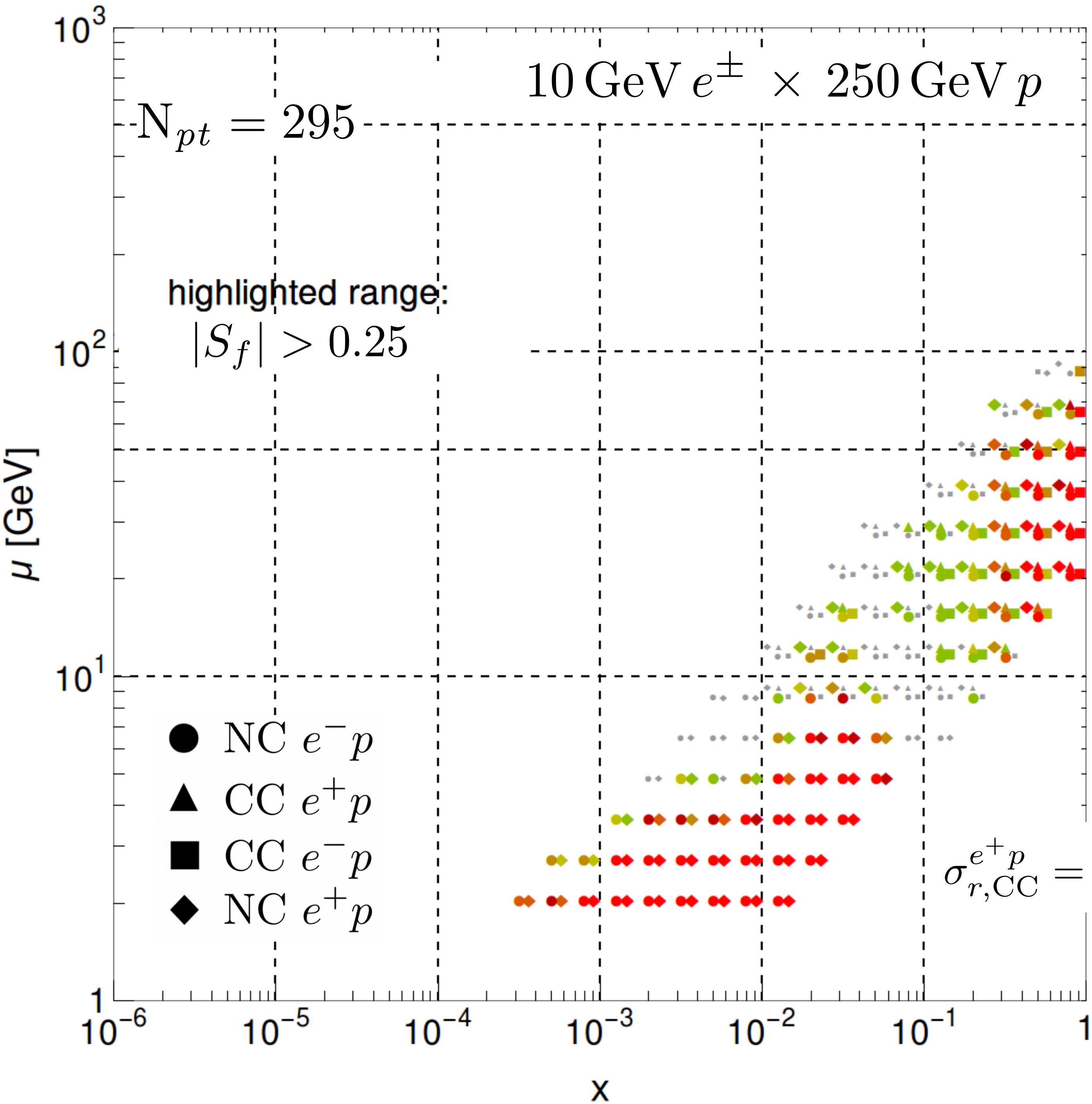


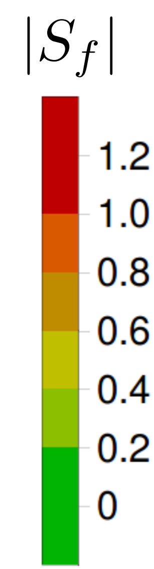
- EIC pseudodata supersede fixed-target DIS information in CT fits
- reweighting strongly depends on tolerance; CT-like tolerance used here

→ complementary approaches welcome!

$|S_f|$ for $d(x,\mu)$ CT14 HERA2 NNLO

Phys.Rev. D98 (2018) 094030





- note: with `PDFSense`, already employing kinematical maps for PDF impacts; can extend to measurement space

- an EIC affords **strong sensitivities without a nuclear target**; here, at both very high and very low x

for $x \to 1$

$$\sigma_{r,\mathrm{CC}}^{e^+p} = \frac{Y_+}{2} W_2^+ \mp \frac{Y_-}{2} x W_3^+ - \frac{y^2}{2} W_L^+$$

$$\simeq [1-y]^2\, x(d+s)$$

$|S_f|$ for $d(x,\mu)$ CT14 HERA2 NNLO

Phys.Rev. D98 (2018) 094030

- note: with PDFSense, already employing kinematical maps for

- <u>in development</u>: these maps can eventually be interfaced with bin-by-bin fast stat. uncertainty estimates

→ for 1 fb^{-1}; can recover stat. uncertainties for desired lumi.

x

	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
5	3 697 277	2 004 394	1 335 307	984 580	771 579	629 786	529 176	454 253	396 474	2 346 683	1 063 724	526 723	250 610	108 060	39 648	11 096	1838	80
10	563 217	331 880	228 507	171 010	134 964	110 518	92 967	79 792	69 577	407 584	179 488	86 233	39 818	16 648	5906	1588	250	10
15	70 225	94 233	67 300	51 929	41 850	34 770	29 552	25 555	22 409	132 685	58 477	27 828	12 692	5232	1826	481	74	3
20	0	19 037	25 757	20 153	16 537	13 972	12 040	10 528	9314	56 377	25 241	12 009	5451	2231	771	201	30	1
25	0	0	3446	8360	7677	6504	5655	4994	4459	27 743	12 740	6097	2766	1128	388	100	15	1
30	0	0	0	308	2630	3441	2978	2626	2353	15 001	7098	3432	1562	637	218	56	8	0
35	0	0	0	0	0	364	1260	1524	1360	8704	4232	2074	949	387	133	34	5	0
40	0	0	0	0	0	0	0	187	555	5324	2653	1320	608	249	85	22	3	0
45	0	0	0	0	0	0	0	0	0	2818	1731	873	406	167	57	15	2	0
50	0	0	0	0	0	0	0	0	0	1207	1172	596	280	116	40	10	1	0
55	0	0	0	0	0	0	0	0	0	421	822	418	198	82	28	7	1	0
60	0	0	0	0	0	0	0	0	0	44	594	301	143	60	21	5	1	0
65	0	0	0	0	0	0	0	0	0	0	280	222	105	44	15	4	1	0
70	0	0	0	0	0	0	0	0	0	0	91	169	79	33	12	3	0	0
75	0	0	0	0	0	0	0	0	0	0	5	116	61	25	9	2	0	0
80	0	0	0	0	0	0	0	0	0	0	0	45	48	20	7	2	0	0
85	0	0	0	0	0	0	0	0	0	0	0	7	38	16	5	1	0	0
90	0	0	0	0	0	0	0	0	0	0	0	0	18	13	4	1	0	0
95	0	0	0	0	0	0	0	0	0	0	0	0	3	10	4	1	0	0
100	0	0	0	0	0	0	0	0	0	0	0	0	0	5	3	1	0	0

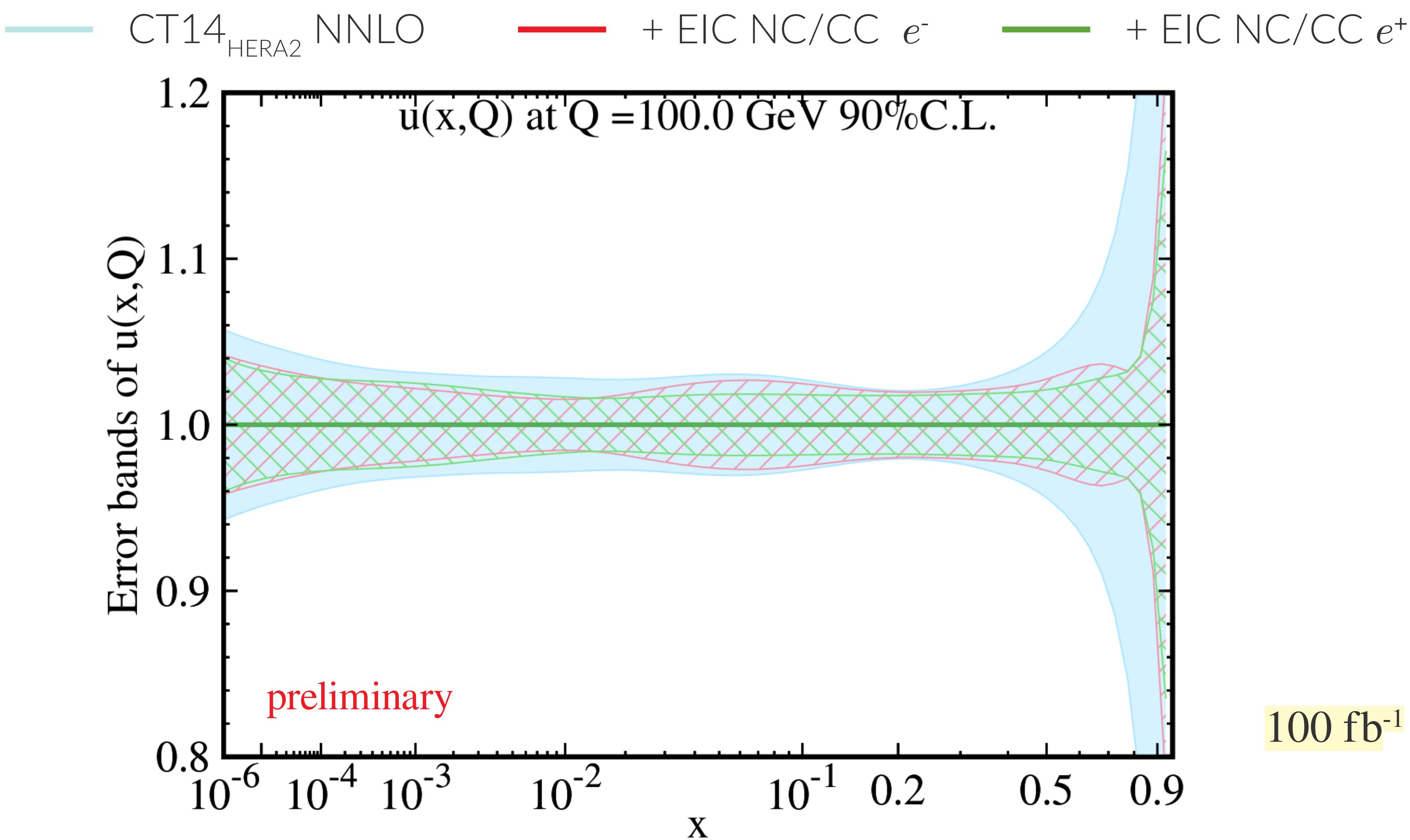
Q [GeV]

for NC e- p at $\sqrt{s}$ = 140 GeV

x

(i) direct fits of EIC pseudo-data

- in addition to fast impact methods, direct fitting is available for end-stage analyses
- results align closely with fast methods; *i.e.*, **`PDFSense`** and **`ePump`**



100 fb^{-1}

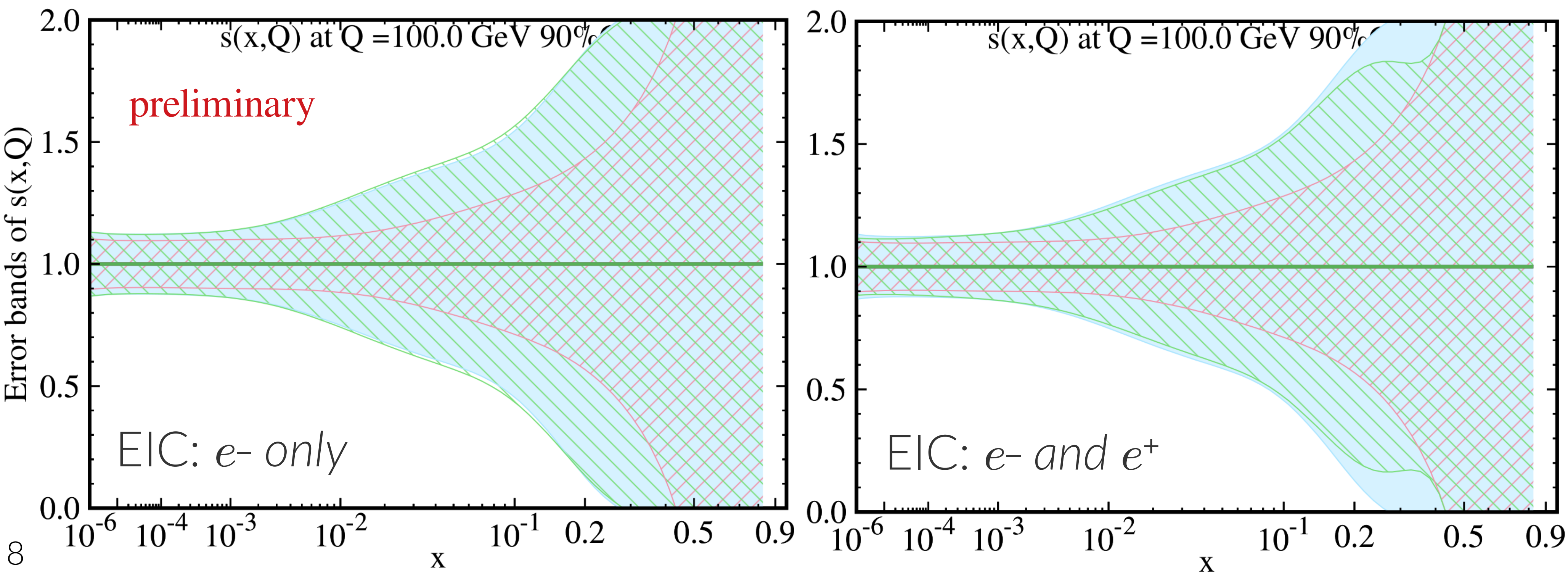
EIC pseudo-data fitted against legacy CT data sets

- beyond fitting EIC over present data, comparisons with legacy data are instructive
- in CT, νFe dimuon production (NuTeV, CCFR) are important constraints on $s(x,Q)$
- especially without e^+, EIC CC inclusive DIS data struggle to compete

→ some high-x pull → inputs from tagged CC data helpful

CT14$_{\text{HERA2}}$ NNLO (*no* dimuon) no dimon + EIC CC DIS (inclusive)

CT14$_{\text{HERA2}}$ NNLO (*with* dimuon)



Unraveling PDFs' flavor dependence is challenging; multiple EIC channels/processes needed

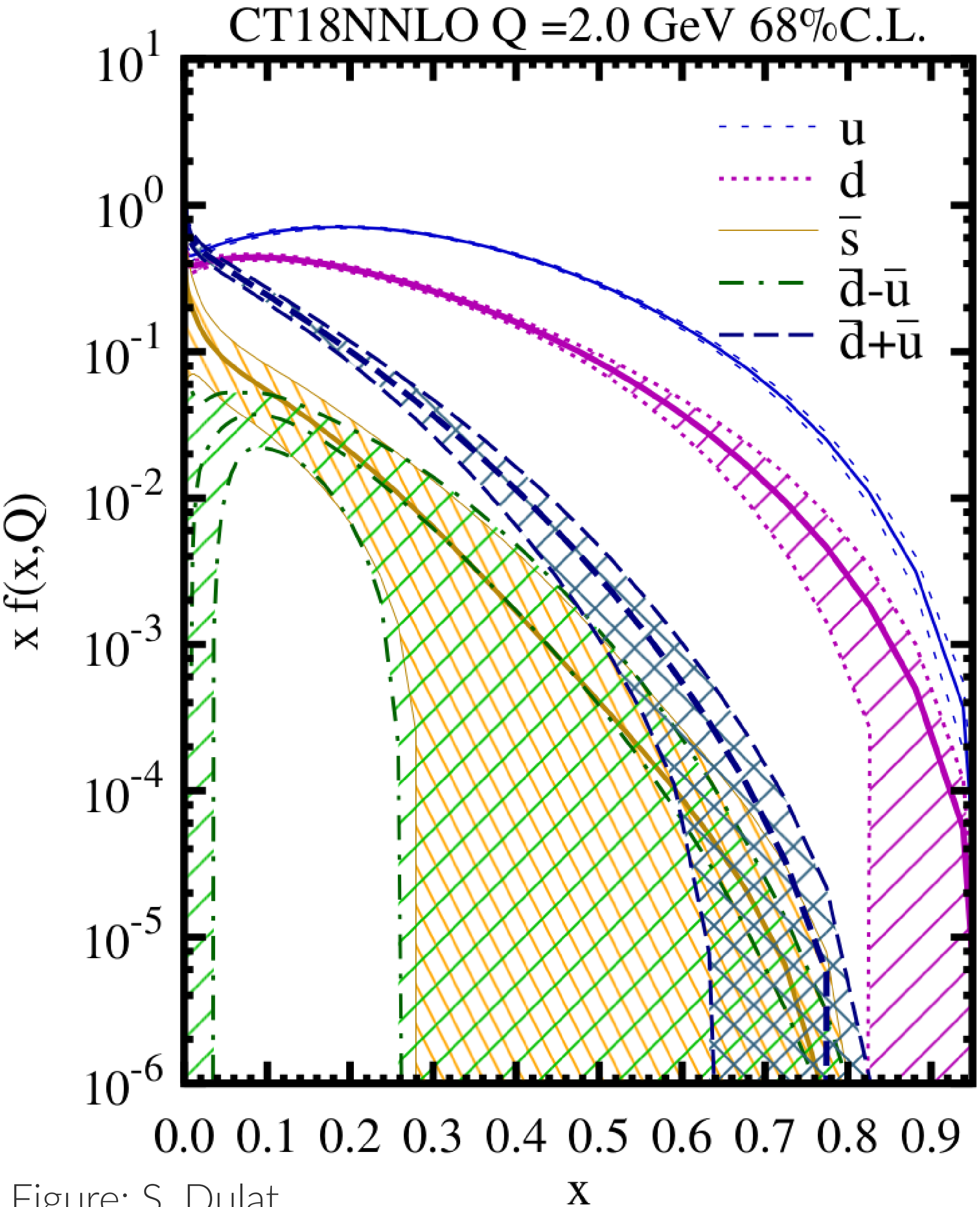


Figure: S. Dulat

note PDFs' different orders-of-mag.!

NC DIS: sensitivity to d-type quarks ¼ that of u-type

$$\sigma \propto \frac{4}{9}(u_+ + c_+) + \frac{1}{9}(d_+ + s_+ + b_+)$$

CC DIS: lower accuracy (1/10 lumi.)

high x (>0.1) inclusive DIS

→ u-quark dominates

→ d-quark ½ of u, but harder to access in NC DIS (above)

→ $\bar{d}+\bar{u}$ ~ few percent of u

…1% error on u → 50-100% error on $\bar{d}+\bar{u}$

→ for x~0.1,

$$s \approx \bar{s} \approx \bar{d} - \bar{u} < 0.1(\bar{d}+\bar{u})$$

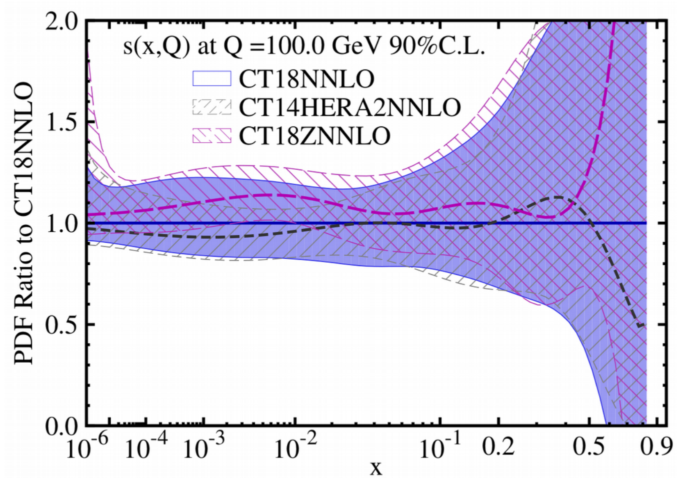
→ at x>0.5, no separation for $\bar{u}, \bar{d}, \bar{s}$

(ii) case study: accessing nucleon strangeness in (tagged) CC processes?

→ in CT, main inputs are from νDIS on heavy nuclei (nuclear corrs. relevant)

→ of interest: the strange suppression ratio, $R_s = \dfrac{s + \bar{s}}{\bar{u} + \bar{d}}$

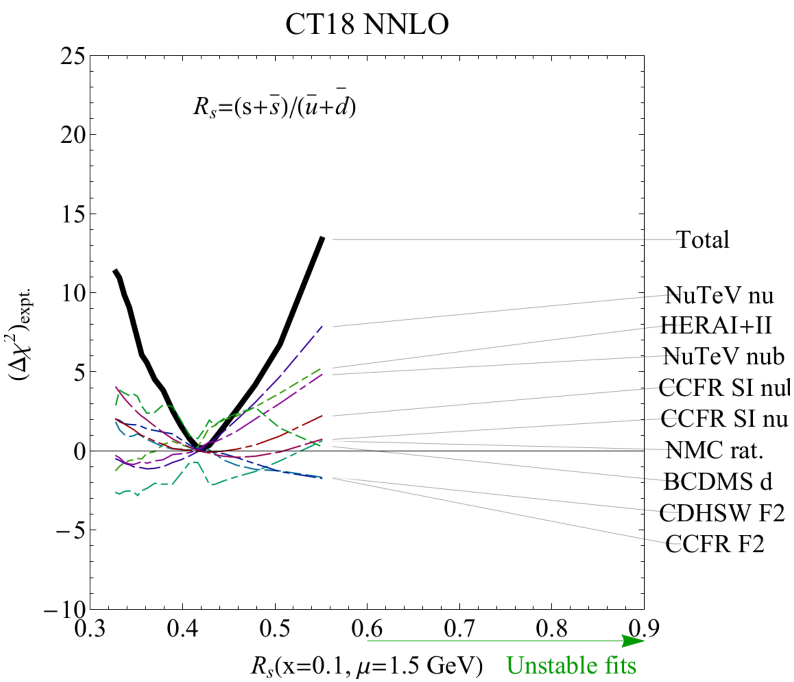
→ typical QCD fits find R_s ~ 0.5; ATLAS W/Z production favors R_s ~ 1

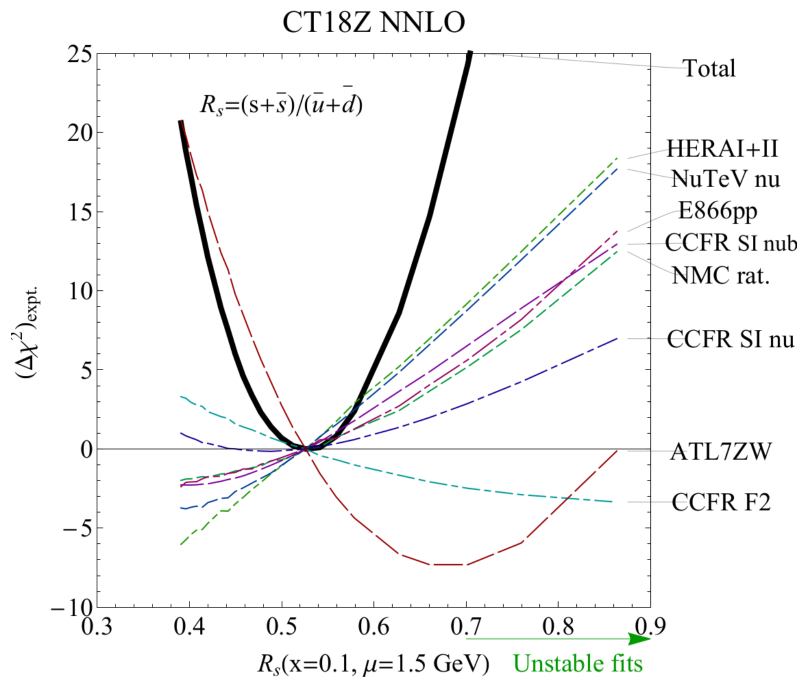


→ *question*: can CC charm jet production off <u>proton</u> distinguish small from large R_s?

- Lagrange multiplier scans of R_s in CT18 analysis:

arXiv: 1912.10053 [hep-ph]



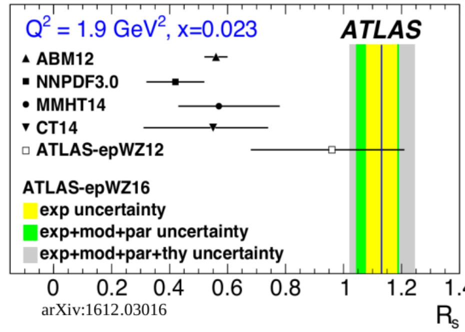


<u>theory inputs:</u> for

$x = 0.1, \ \ \mu = 1.5\,\mathrm{GeV}$

* $R_s = 0.325$ (CT18)

* $R_s = 0.863$ (CT18Z)

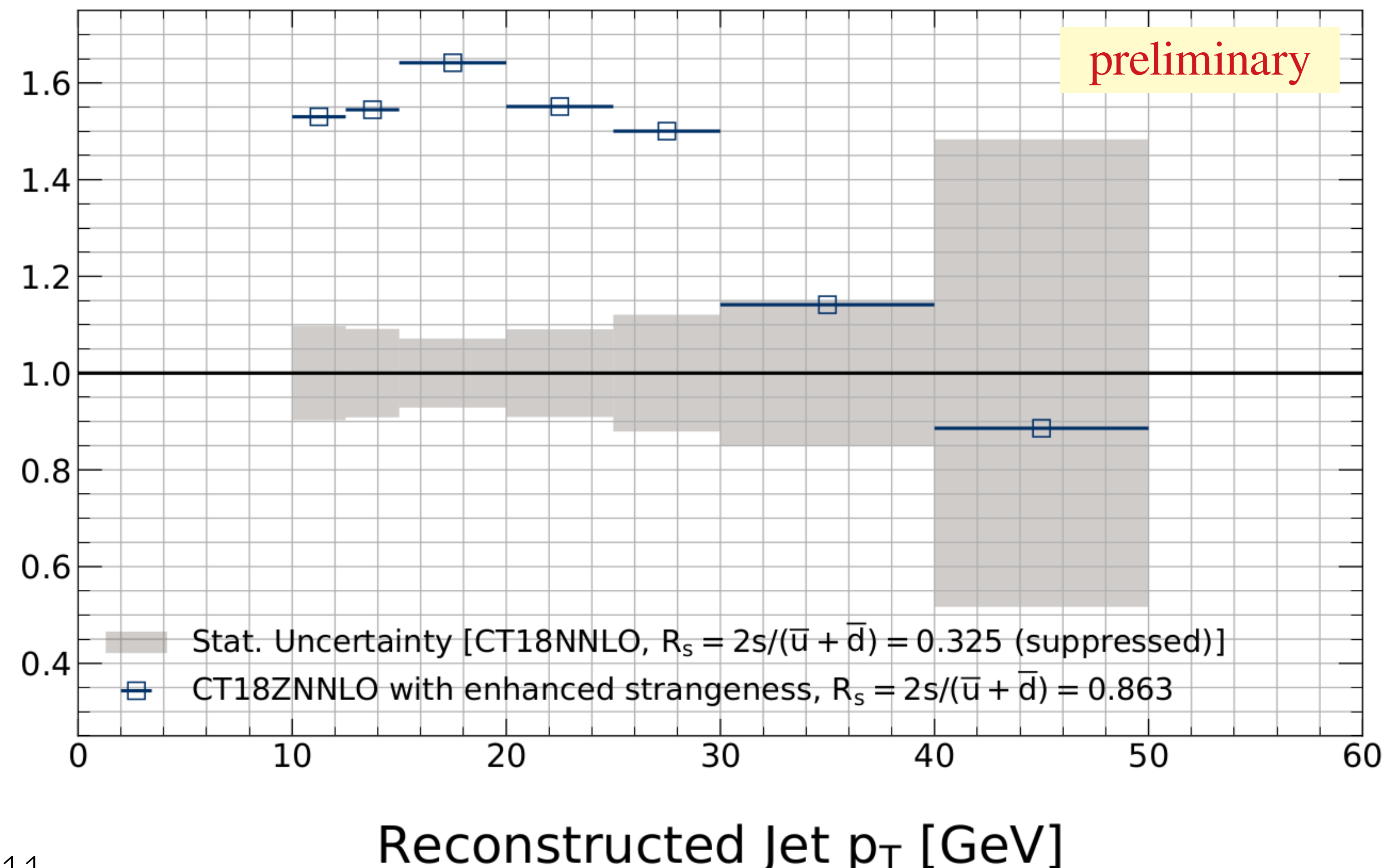


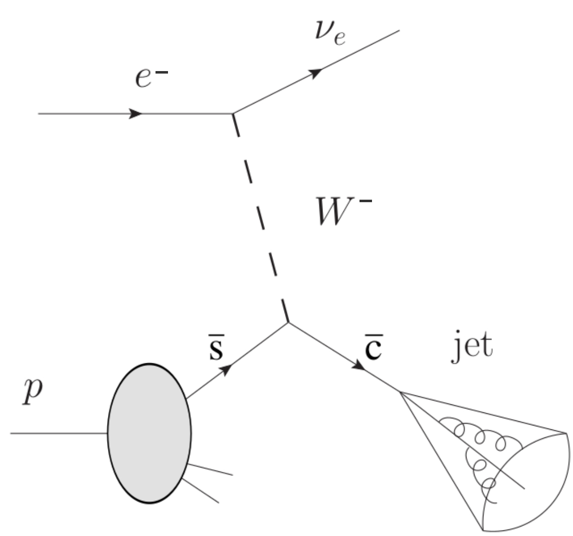
sensitivity of CC charm jet production to strange suppression

see talks: S. Sekula (Inclusive/SIDIS/Jets-HQ Joint) & M. Arratia

- 100 fb^{-1} CC DIS (10M simulated events), at 10x275 GeV (e^- on p); Q^2 > 100 GeV^2

- **even assuming conservative charm-tagging efficiency, event-level discrimination potential is substantial, relative to statistical uncertainties**





final-state tagging will provide a critical lever arm for flavor separation

(iii) input theory grids from modern CT global fits

- CT-based theory grids are currently in development
- for preliminary testing and benchmarking, we have initially made available:

$$F_2^{\gamma}\ (900);\ \ F_L^{\gamma}\ (901)$$

$$F_2^{W^-}\ (930);\ \ F_L^{W^-}\ (931);\ \ F_3^{W^-}\ (932)$$

1.234100E-04 1.507330E-0 ... 112530E-04 7.465860E-04 9.118820E-04
1.360370E-03 1.661560E-0 ... 737950E-03 8.229750E-03 1.005180E-02
1.499560E-02 1.831560E-0 ... 427360E-02 9.071800E-02 1.108030E-01
1.446650E-01 1.546380E-0 ... 465970E-01 2.635970E-01 2.817690E-01
3.219580E-01 3.441540E-0 ... 488120E-01 5.611440E-01 5.737530E-01
5.998290E-01 6.133070E-0 ... 165310E-01 7.326320E-01 7.490950E-01
7.831390E-01 8.007370E-0 ... 559400E-01 8.623030E-01 8.687150E-01
8.816800E-01 8.882350E-0 ... 355070E-01 9.424620E-01 9.494700E-01
9.636400E-01 9.708050E-0
1.295000E+00 1.298750E+0 ... 396640E+00 4.001450E+00 4.750000E+00
7.070720E+00 8.758190E+0 ... 468640E+01 7.507240E+01 1.047120E+02
2.143800E+02 3.152120E+0
900 901 930 931 932
7.02993E-01 2.68314E-02 3.34988E+00 9.58268E-01 -3.26391E-01
7.07730E-01 2.67969E-02 3.37103E+00 9.63083E-01 -3.29891E-01
9.52684E-01 3.41008E-02 4.34177E+00 1.23546E+00 -5.49899E-01
1.22029E+00 5.09282E-02 5.60916E+00 1.50347E+00 -6.28726E-01
1.53696E+00 8.47902E-02 7.10935E+00 1.83204E+00 -7.01457E-01
1.91427E+00 1.37909E-01 8.86982E+00 2.22522E+00 -7.64617E-01
2.36476E+00 2.12566E-01 1.09098E+01 2.68973E+00 -8.11648E-01
2.90192E+00 3.10932E-01 1.32566E+01 3.23596E+00 -8.40145E-01
3.53866E+00 4.35331E-01 1.59289E+01 3.87542E+00 -8.45608E-01
4.28932E+00 5.88655E-01 1.89499E+01 4.61887E+00 -8.26580E-01
5.17781E+00 7.71996E-01 2.23491E+01 5.47507E+00 -7.83729E-01
6.28751E+00 1.01494E+00 2.65775E+01 6.61457E+00 -7.07546E-01
7.57983E+00 1.27277E+00 3.13436E+01 7.84127E+00 -6.19615E-01
9.06879E+00 1.56691E+00 3.67284E+01 9.26204E+00 -5.13664E-01
1.07719E+01 1.89701E+00 4.28131E+01 1.09051E+01 -3.96581E-01
1.27040E+01 2.26086E+00 4.96774E+01 1.27910E+01 -2.77037E-01
1.48814E+01 2.65660E+00 5.74157E+01 1.49432E+01 -1.67317E-01
1.73151E+01 3.07804E+00 6.61715E+01 1.74206E+01 -8.68745E-02
2.00128E+01 3.51927E+00 7.61371E+01 2.03048E+01 -7.35058E-02
2.29994E+01 3.97423E+00 8.74733E+01 2.35942E+01 -2.04320E-01
2.62961E+01 4.43726E+00 1.00317E+02 2.71613E+01 -6.30205E-01
2.99477E+01 4.90898E+00 1.15002E+02 3.08598E+01 -1.65132E+00
3.40065E+01 5.38836E+00 1.31991E+02 3.43119E+01 -3.82499E+00
3.86088E+01 5.88410E+00 1.51426E+02 3.66519E+01 -7.77026E+00
4.40040E+01 6.42754E+00 1.73169E+02 3.74007E+01 -1.32902E+01
5.05488E+01 7.10023E+00 1.97291E+02 3.74439E+01 -1.88855E+01
6.95231E-01 2.73262E-02 3.30746E+00 9.29410E-01 -3.23339E-01
6.99809E-01 2.73087E-02 3.32808E+00 9.34205E-01 -3.26732E-01
9.35754E-01 3.54460E-02 4.27011E+00 1.20284E+00 -5.42529E-01

→ based on CT18 NNLO

→ other targets (*e.g.*, deuteron), boson/SF combinations can be generated as needed

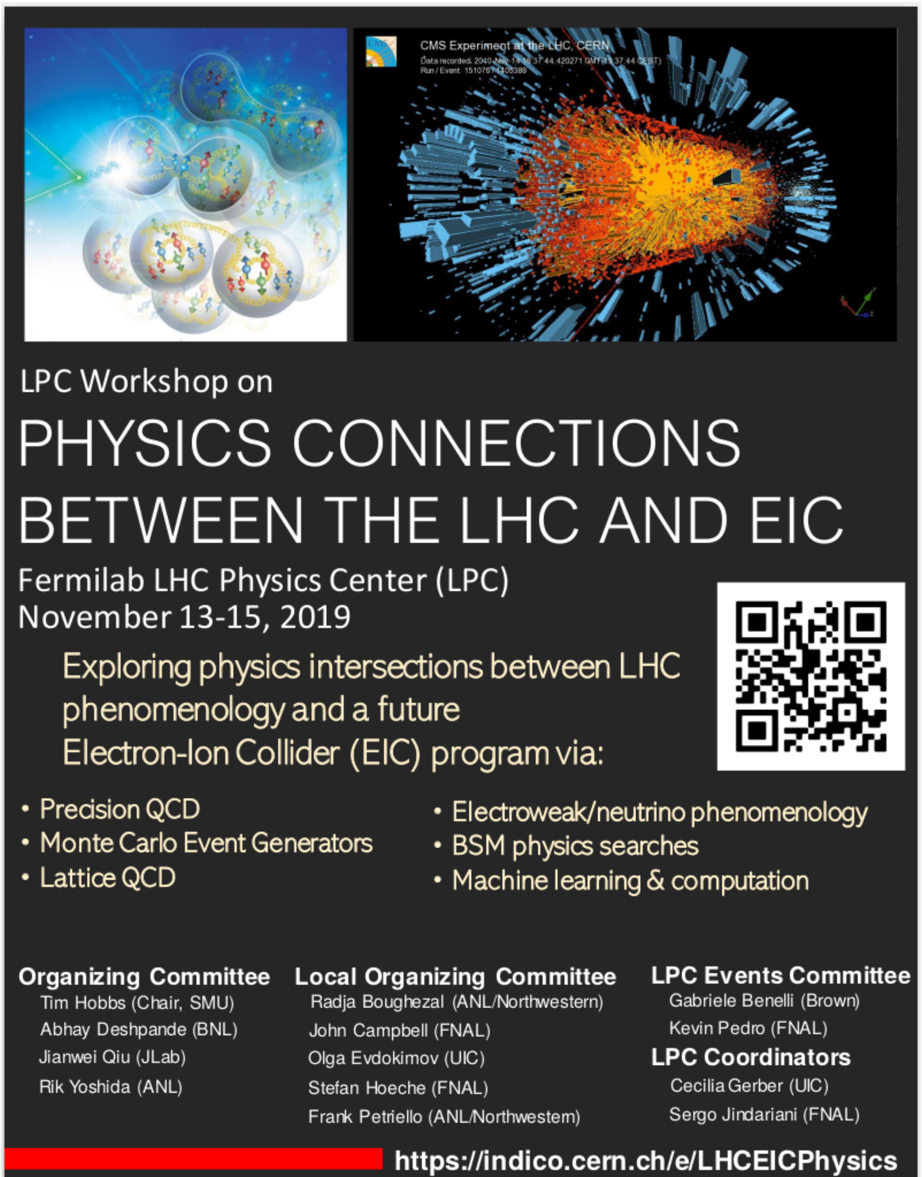
→ different theory settings also quickly done

- pQCD choices [$O(\alpha_s)$, m_c, ...]
- alternate fits:
  - → CT18X (modified DIS scale)
  - → CT18Z (ATLAS W/Z; s-PDF)

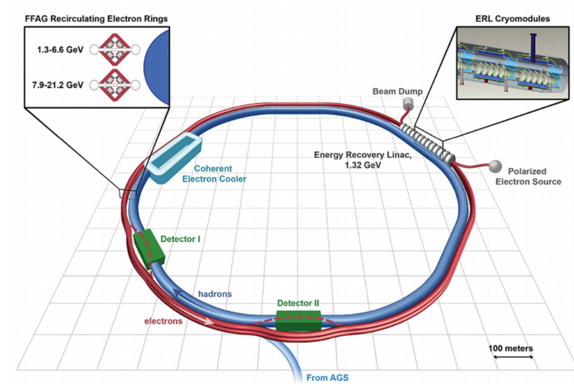
meanwhile: the Snowmass kickoff overlaps with this meeting…

- the PDF/pQCD focuses inside the YR IRG (as well as HQ/SIDIS) are central to Snowmass objectives [which has comparable timescale]

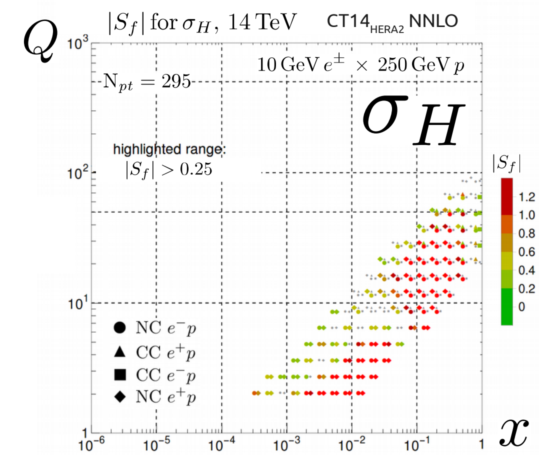
a community is coalescing around this topic



BNL/eRHIC design



arXiv: 1907.00988



EIC pseudodata sensitivity

- **as a machine for precision QCD, this will have serious implications for Energy Frontier efforts**

  e.g., PDF improvements → SM precision

- the EIC will intersect multiple Snowmass frontiers/topics

  → coordination within the EICUG needed

full program available:

https://indico.cern.ch/e/LHCEICPhysics

conclusions, future steps

- **post-Temple updates**
  - → validation of fast methods with direct pseudo-data fits
    - → reproduce fast-method conclusions
    - → e^{+} data play a vital role in extending EIC impact

      also relevant for CSV, $s \neq \bar{s}$
  - → initial examination of event-level sensitivity of CC DIS charm jet prod.
    - → provides another avenue to unravel light-quark sea: $s(x), R_s(x)$

see talks: S. Sekula (Inclusive/SIDIS/Jets-HQ Joint) & M. Arratia

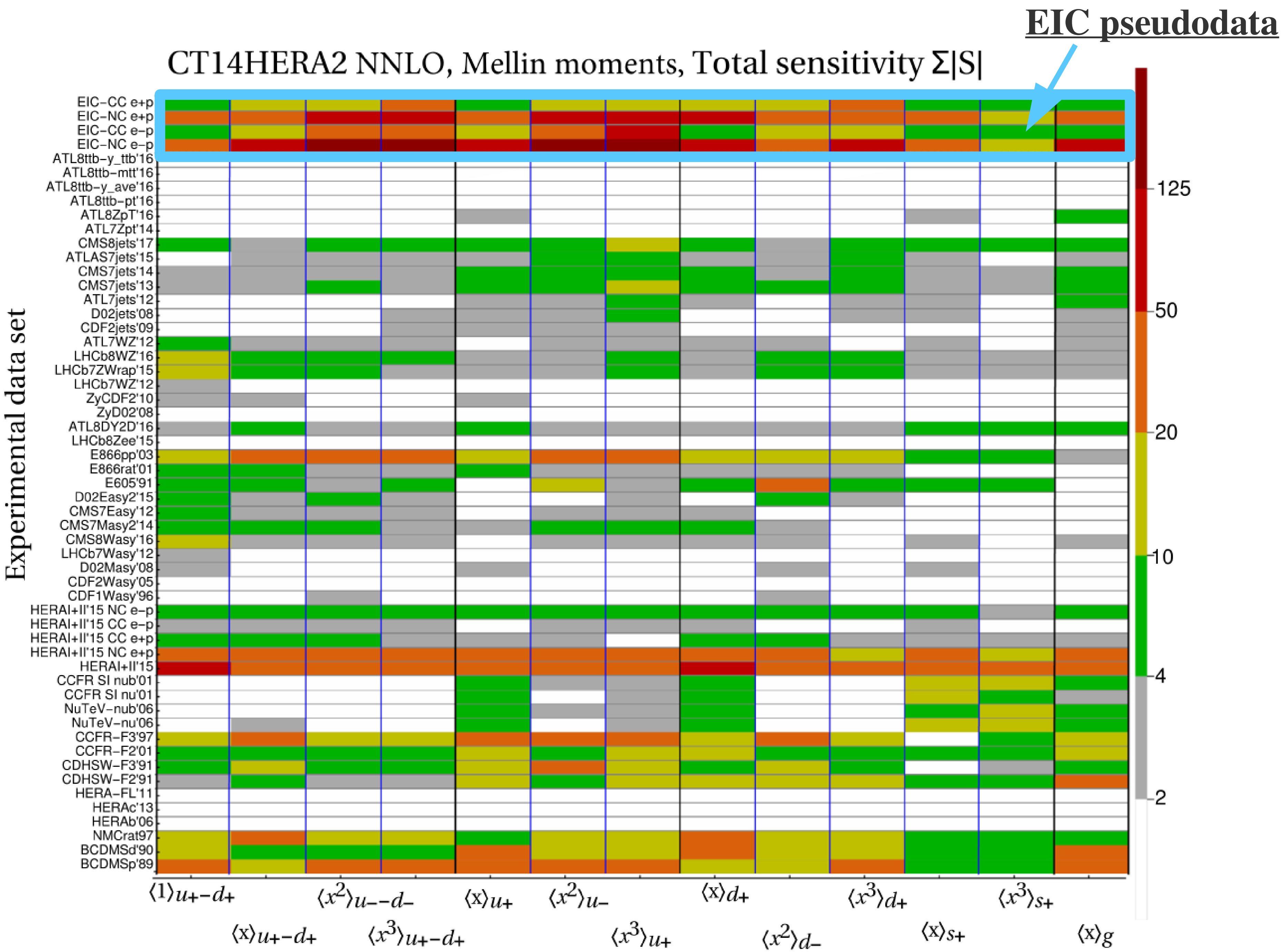
  - → input theory grids according to CT are in production; initial grids released
    - → extension/rapid updates possible based on need

- going forward: build on these studies by examining dependence on systematics

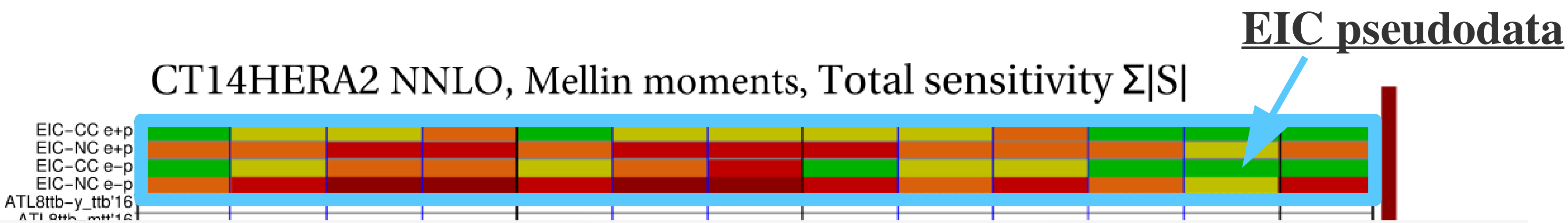
THANKS!

supplementary material

sensitivities can be aggregated for direct comparisons of exps



sensitivities can be aggregated for direct comparisons of exps



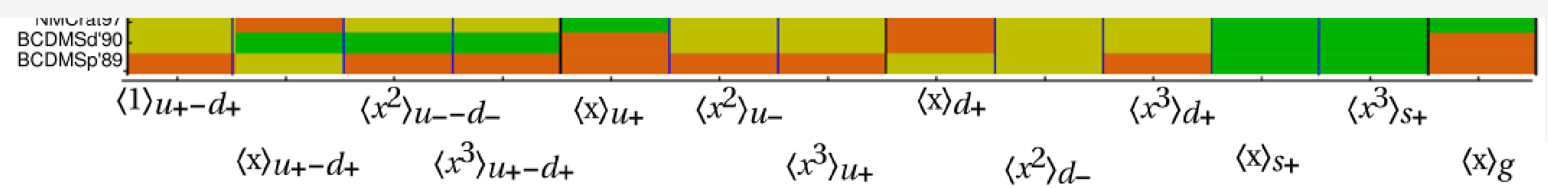
`PDFSense`* furnishes:

→ quick evaluations of pulls of individual data

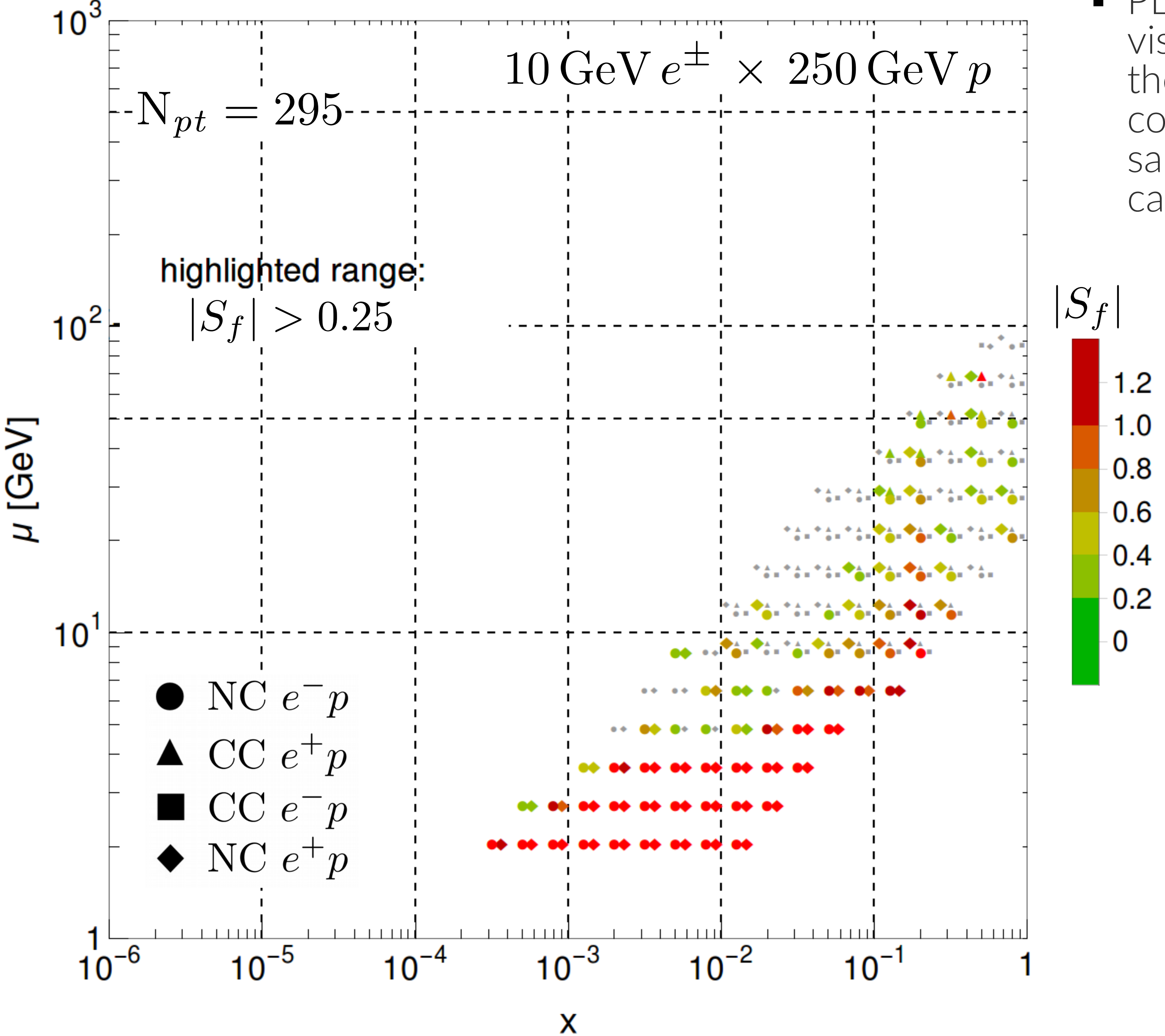
→ info. less tied to PDF parametrization(s)

→ aggregate scores to compare experiments

*powerful in combination with direct fitting/reweighting



$|S_f|$ for $g(x,\mu)$ CT14 HERA2 NNLO

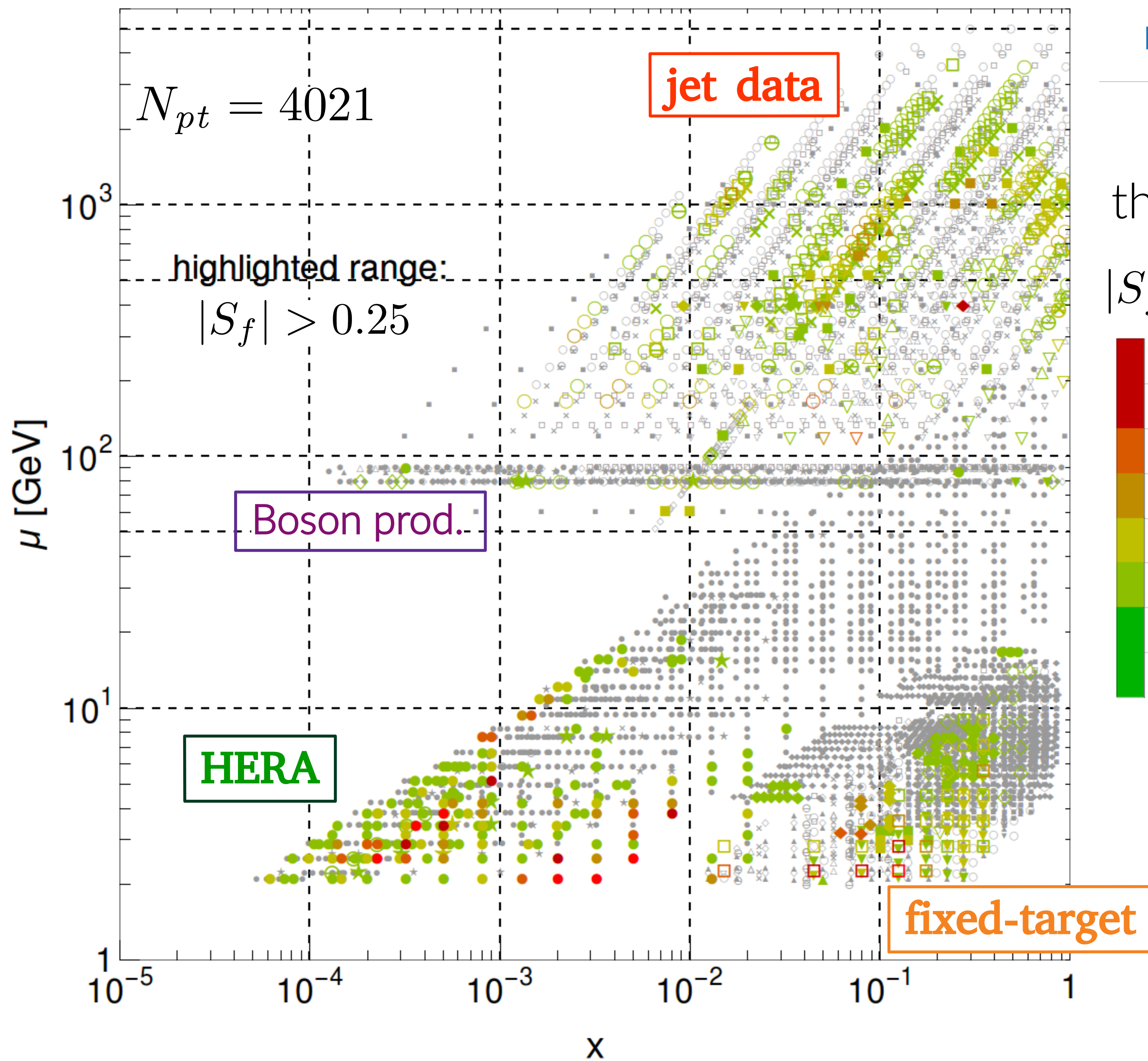


- PDFSense quickly visualizes the pulls of the EIC pseudodata, comparing on the same scale as for the candidate CT18 exps.

- an EIC will provide a sensitive probe to the gluon distribution – especially at low x
  $$x \gtrsim 3 \times 10^{-4}$$

- these constraints arise from high statistics neutral current data on $\sigma_{r,\mathrm{NC}}^{e^{\pm}p}$

$|S_f|$ for $g(x,\mu)$, CT14 HERA2 NNLO

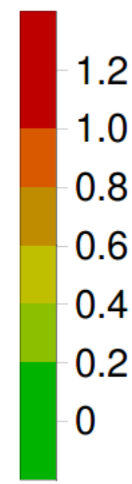


II) visualizing impacts with PDFSense

Phys.Rev. D98 (2018) 094030

the PDF *sensitivity*

$|S_f|$



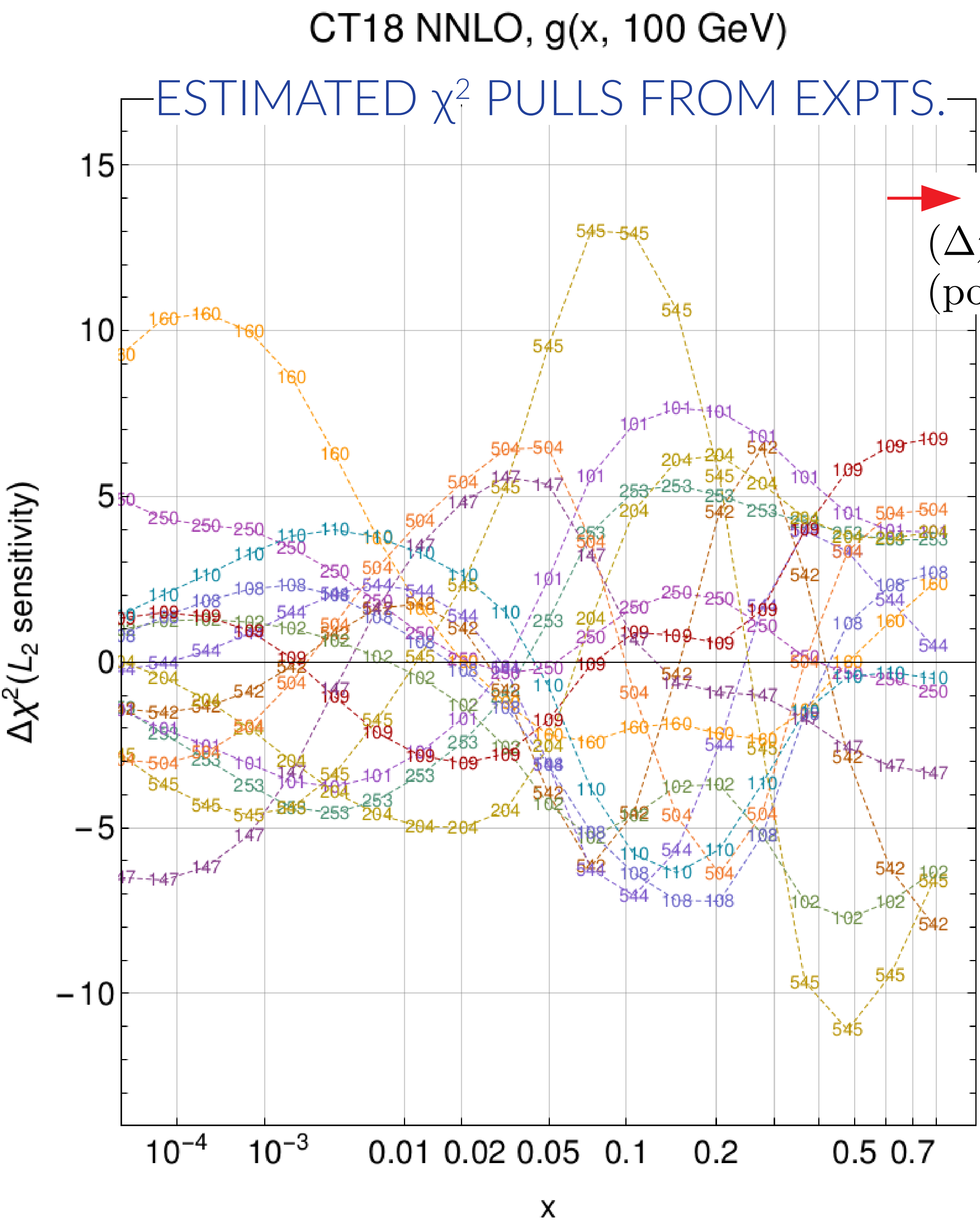
(magnitude of PDF pull of each datum)

→ measurements with strong PDF correlations <u>AND</u> high precision have high $|S_f|$

- used to identify high-impact data for CT18

TJH, Wang, Nadolsky, Olness;
PRD100 (2019) 094040

ESTIMATED χ^2 PULLS FROM EXPTS.



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

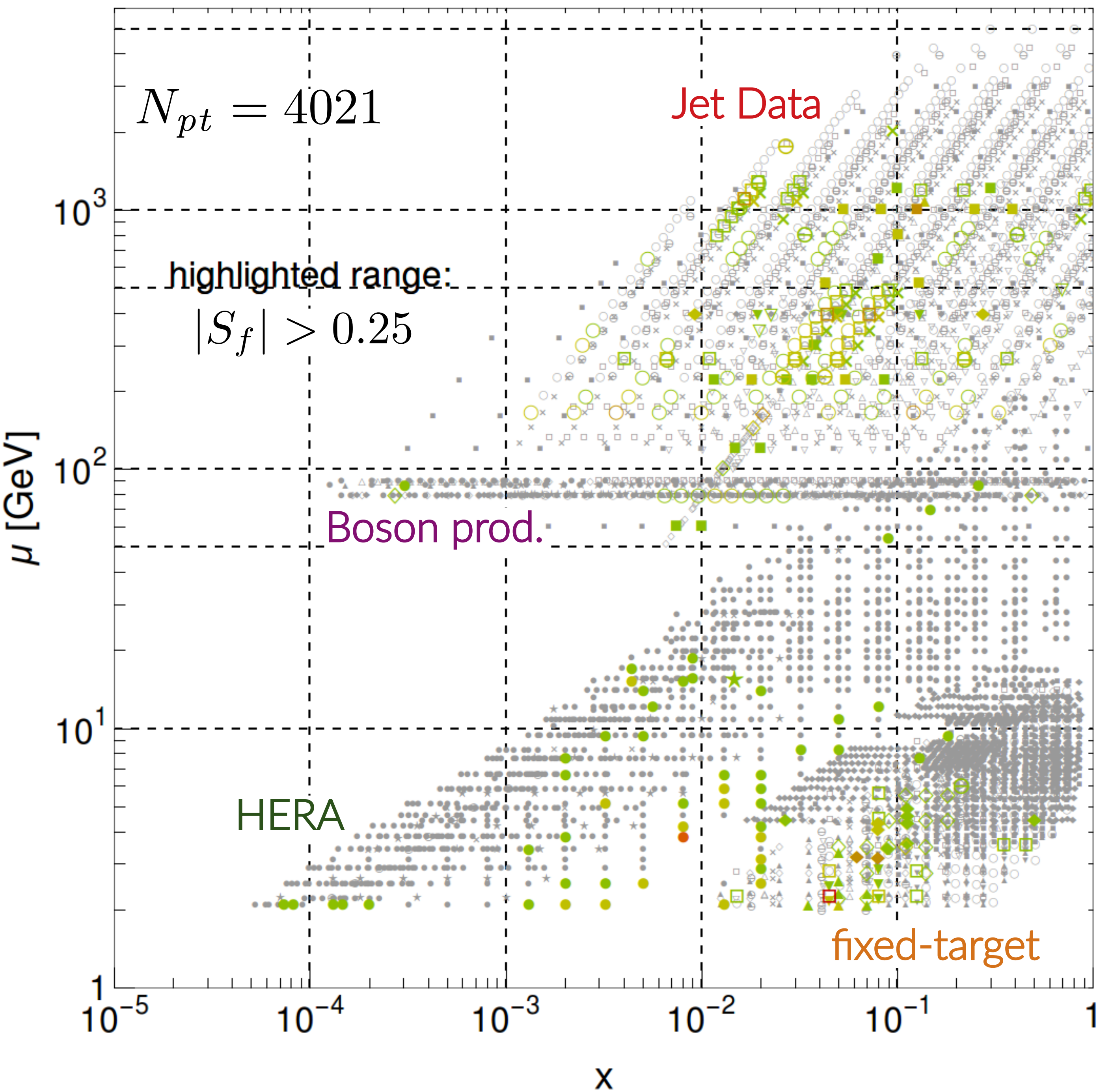
stronger (anti-) correlation

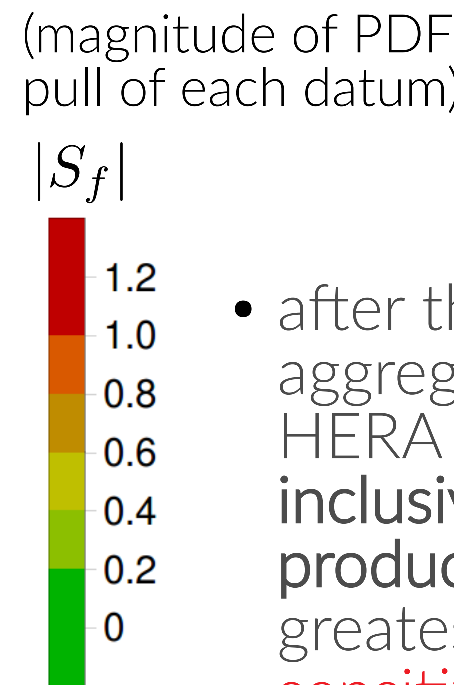
$$S_{f,\,L_2} \sim \mathrm{Corr}[f_a, \chi^2_E]$$

| S_f | for σ_{H^0} 14 TeV, CT14$_{\mathrm{HERA2}}$ NNLO

B.-T. Wang, TJH, S. Doyle, J. Gao, T.-J. Hou, P. M. Nadolsky, F. I. Olness

Phys.Rev. D98 (2018) 094030

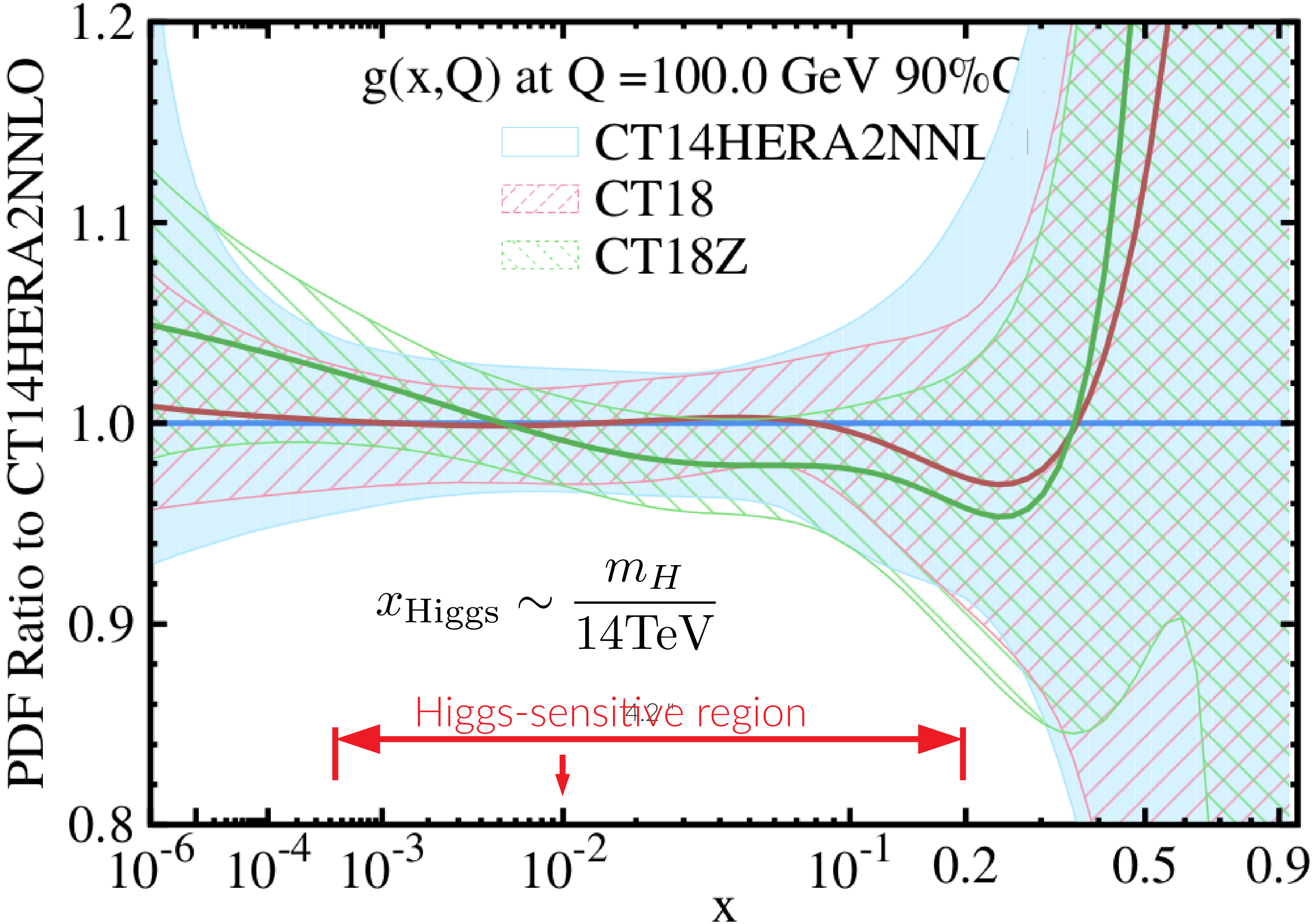




- after the aggregated HERA data, **inclusive jet production** – greatest total sensitivity!

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

LHC Run-1 gluon PDF impact in CT14 → CT18(Z)

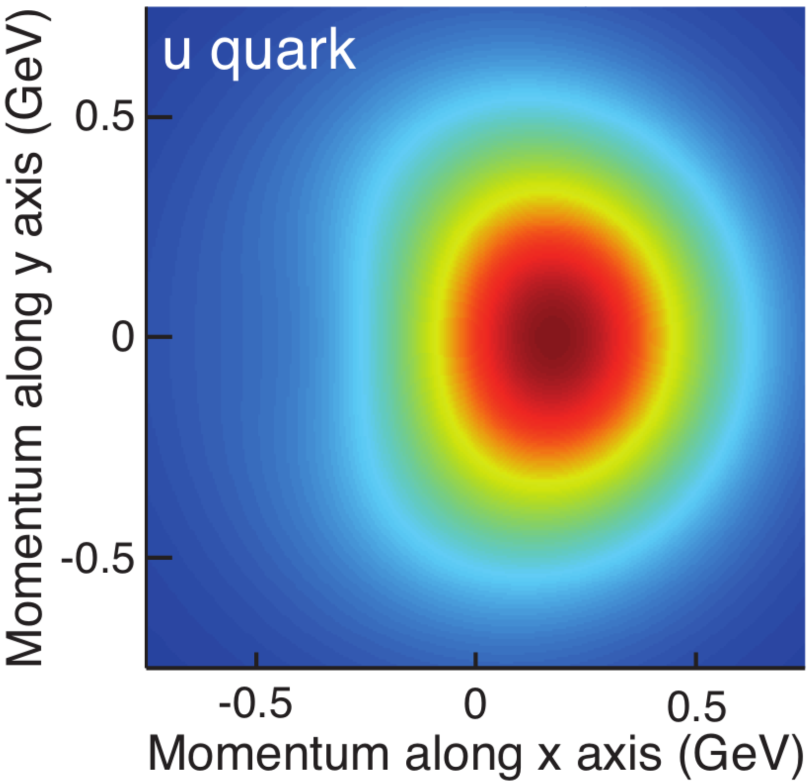


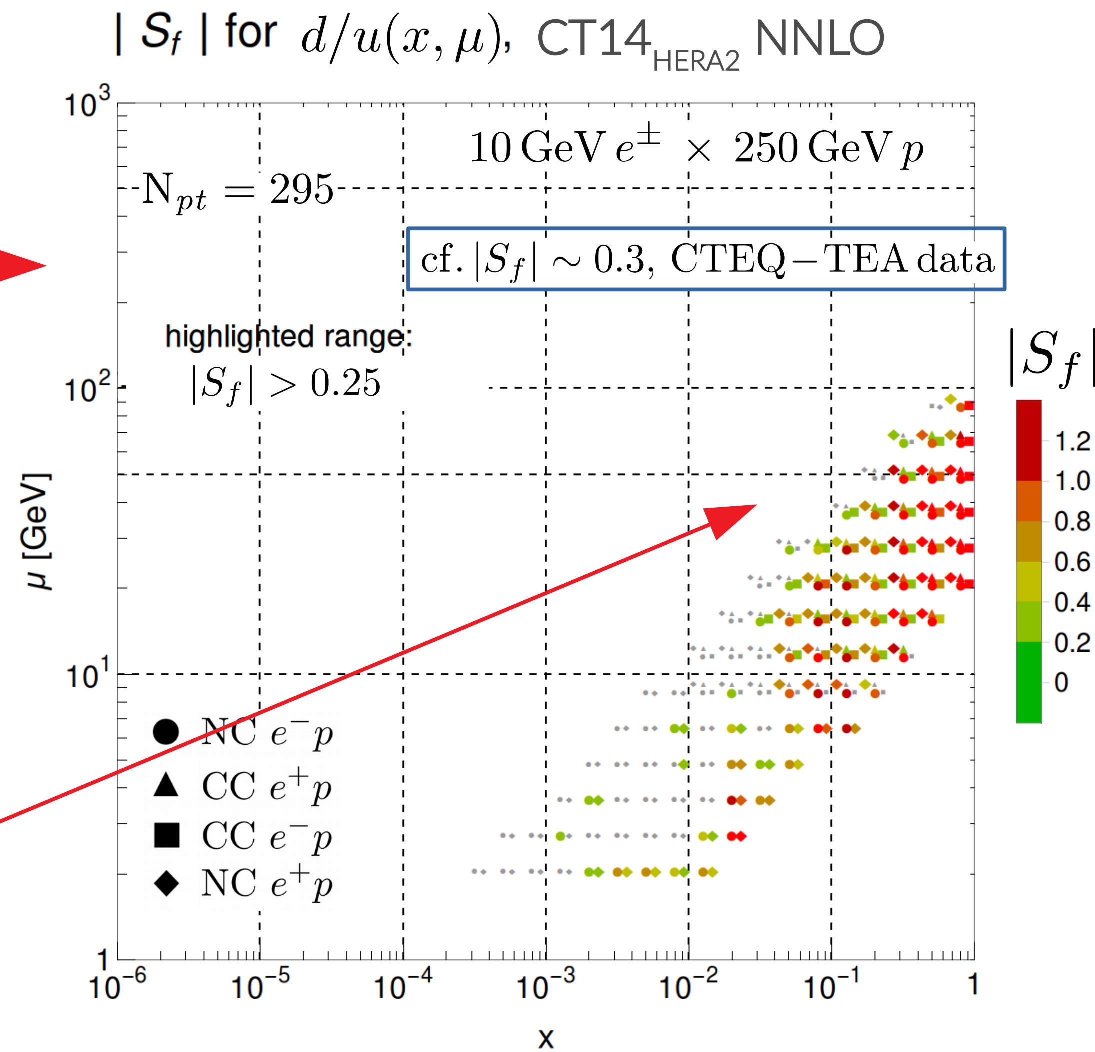
- while LHC Run-1 data drive important PDF improvements, including for the gluon at high-, low-x, the effect is relatively incremental

the EIC tomography program will deliver high-precision DIS

- by measuring the nucleon's multi-dimensional wave function with high precision, the EIC will hugely constrain proton collinear structure

Accardi et al., EPJA52 (2016) no.9, 268.

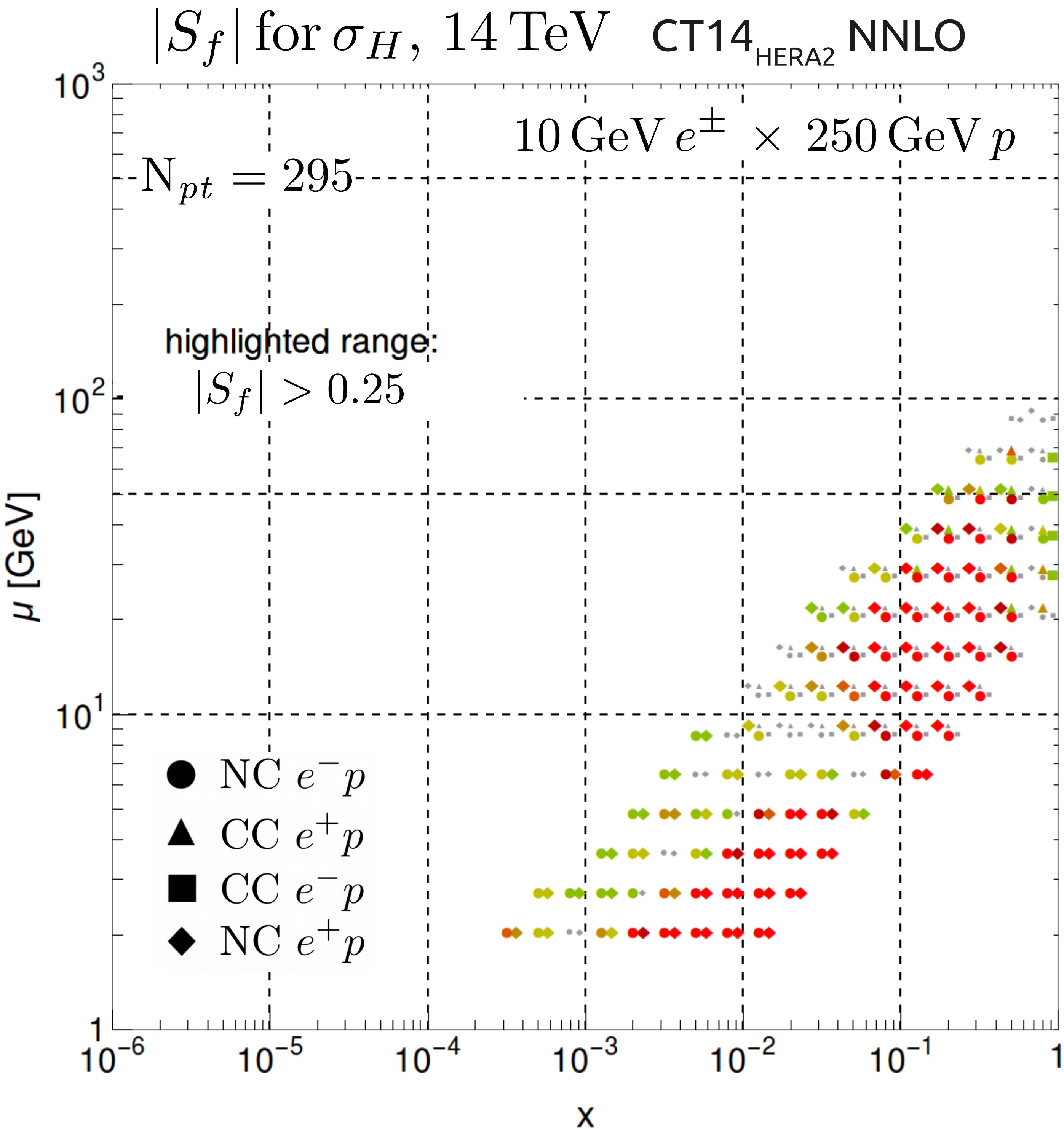


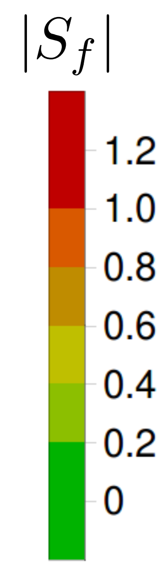


PROJECTED IMPACT OF EIC PSEUDO-DATA VERY LARGE – RED SIMULATED MEASUREMENTS

- DIS cross sections from EIC will supercede the bulk of fixed-target information in contemporary QCD fits; provide an 'anchor-point' to resolve systematic PDF tensions

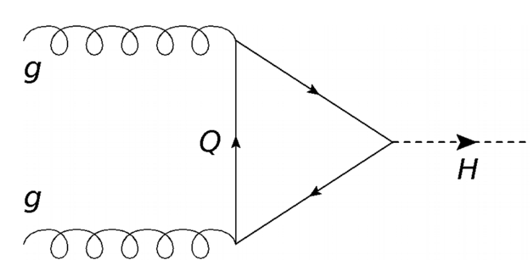
strong predicted impact on the Higgs sector





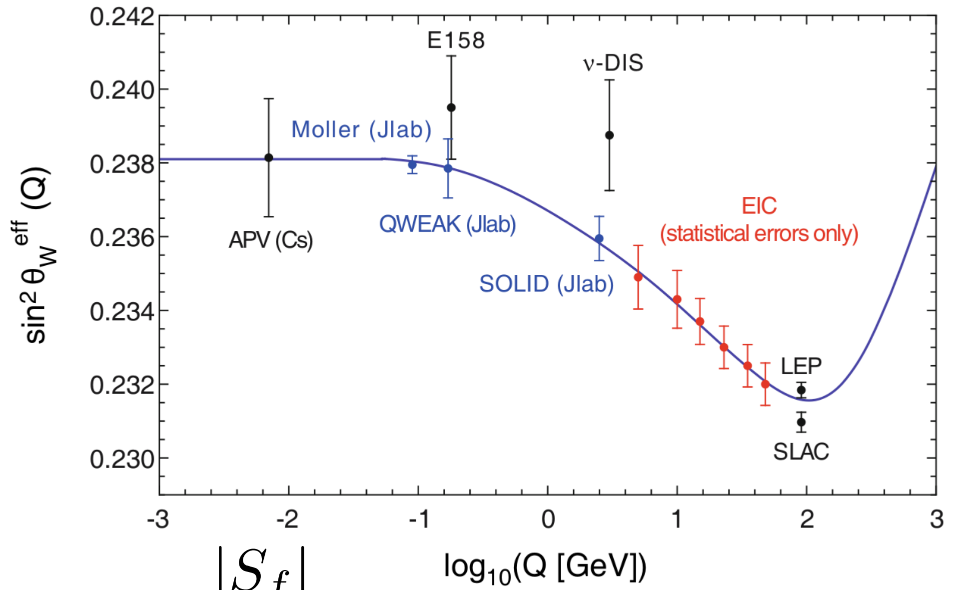
- the impact of an EIC upon the theoretical predictions for inclusive Higgs production arises from a very broad region of the kinematical space it can access

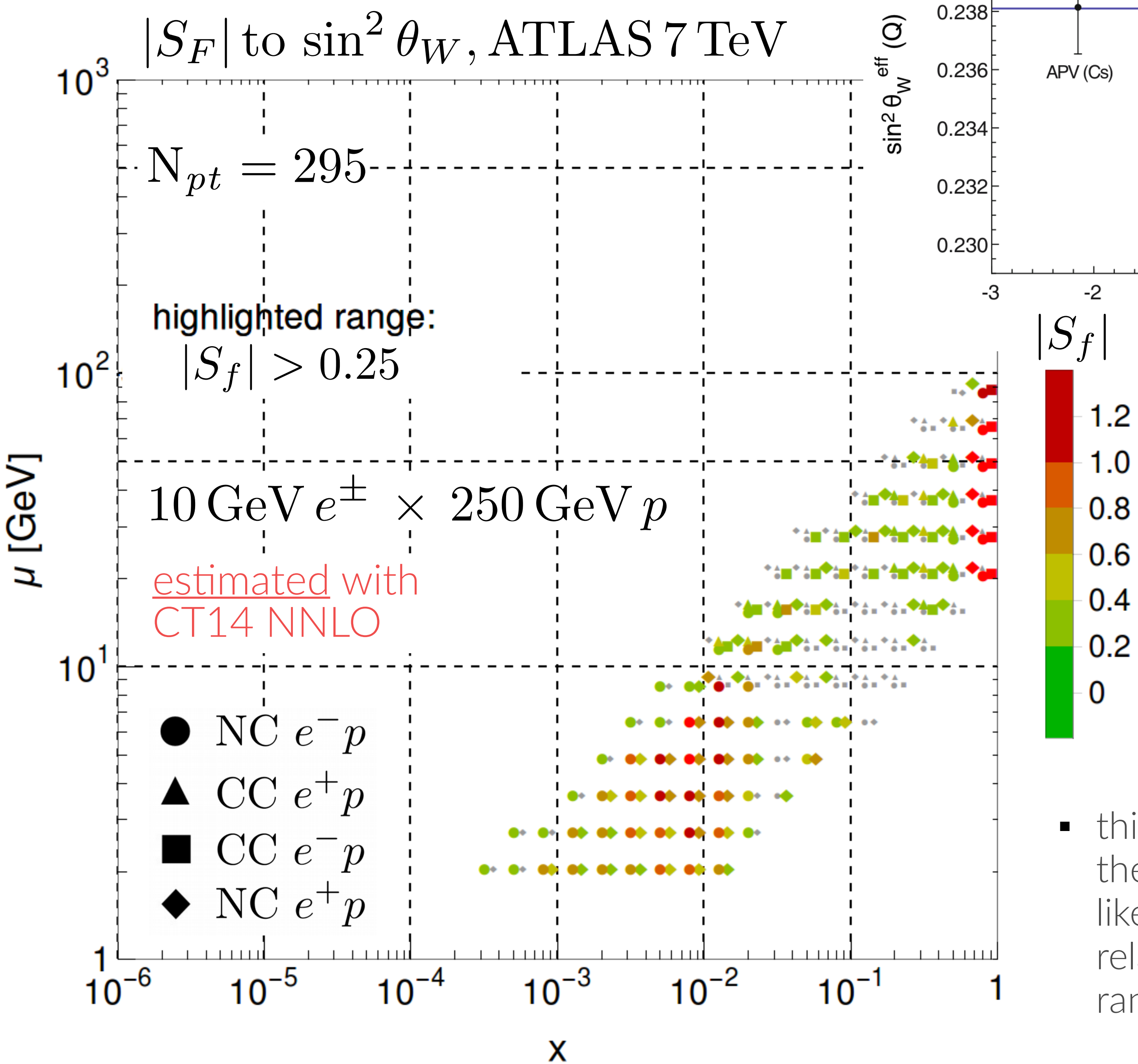
- impact rather closely tied to that of the integrated gluon PDF:



an EIC will probe EW parameters and New Physics!

Accardi et al., EPJA**52**, 268 (2016).





- observe a pronounced sensitivity to the Weinberg angle, especially low and high x, **even at**

$$\mathcal{L} = 100\mathrm{fb}^{-1}$$

- this corresponds closely to the kinematics at which EIC is likely to measure A^{PV} — relatively large Q^2 and in the x range $0.2 \lesssim x \lesssim 0.5$

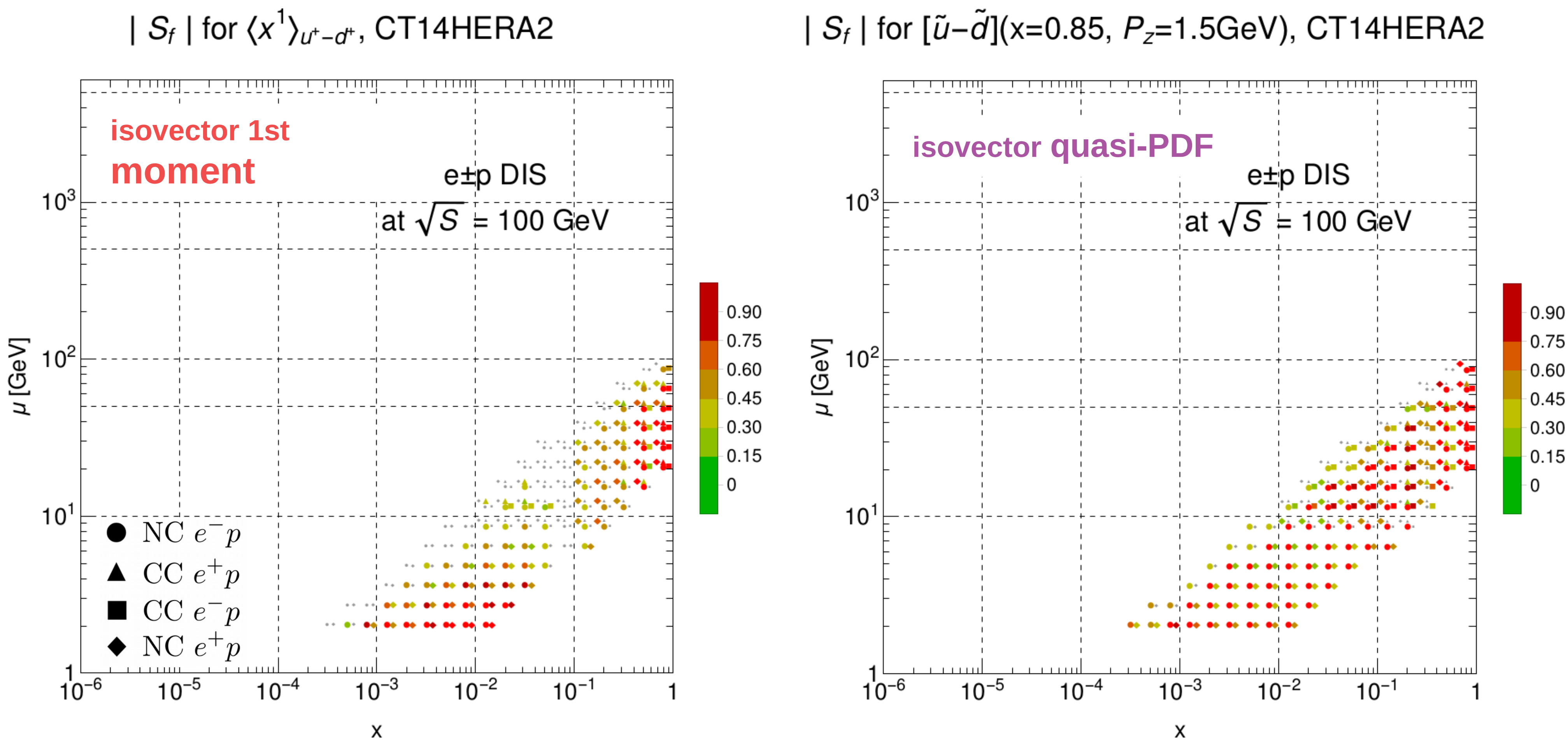
An EIC would drive lattice phenomenology

arXiv:1904.00022 [hep-ph]

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

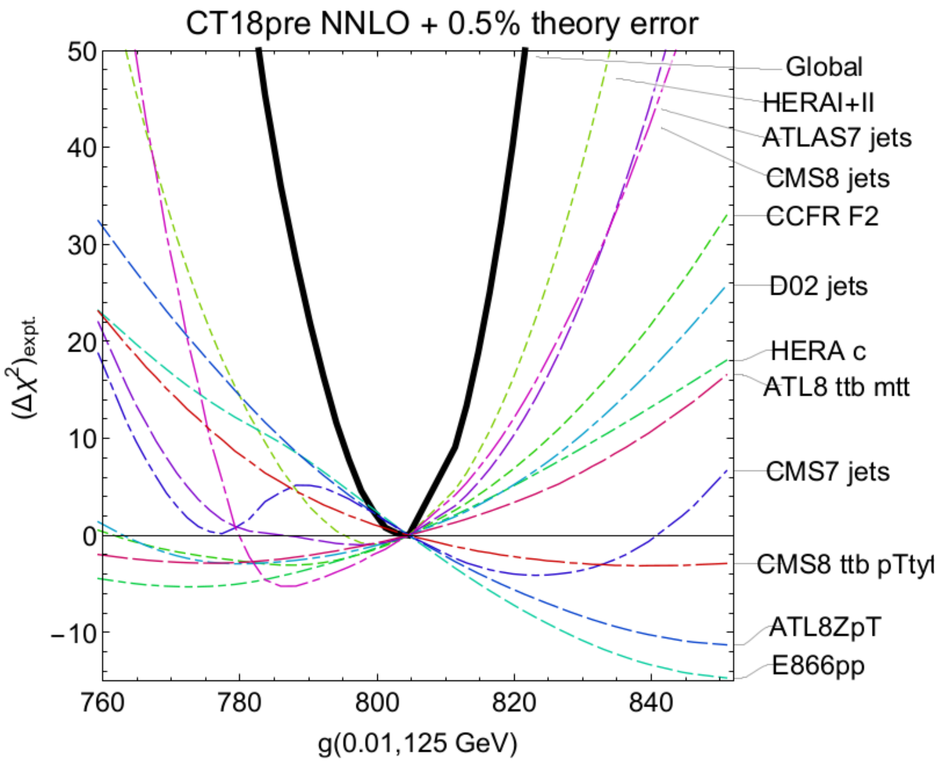
- Many of the experiments most sensitive to PDF Mellin moments and qPDFs involve nuclear targets → eA data from EIC would sharpen knowledge of nuclear corrections

$$\langle x^n \rangle_{q,g} = \int dx\, x^n f_{q,g}(x, \mu = 2\,\mathrm{GeV}) \qquad \widetilde{q}(x, P_z, \widetilde{\mu}) = \int dy\, Z\left(\frac{x}{y}, \frac{\Lambda}{P_z}, \frac{\mu}{P_z}\right) q(y,\mu) + \mathcal{O}\left(\frac{\Lambda^2_{\mathrm{QCD}}}{P_z^2}, \frac{M^2}{P_z^2}\right)$$



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

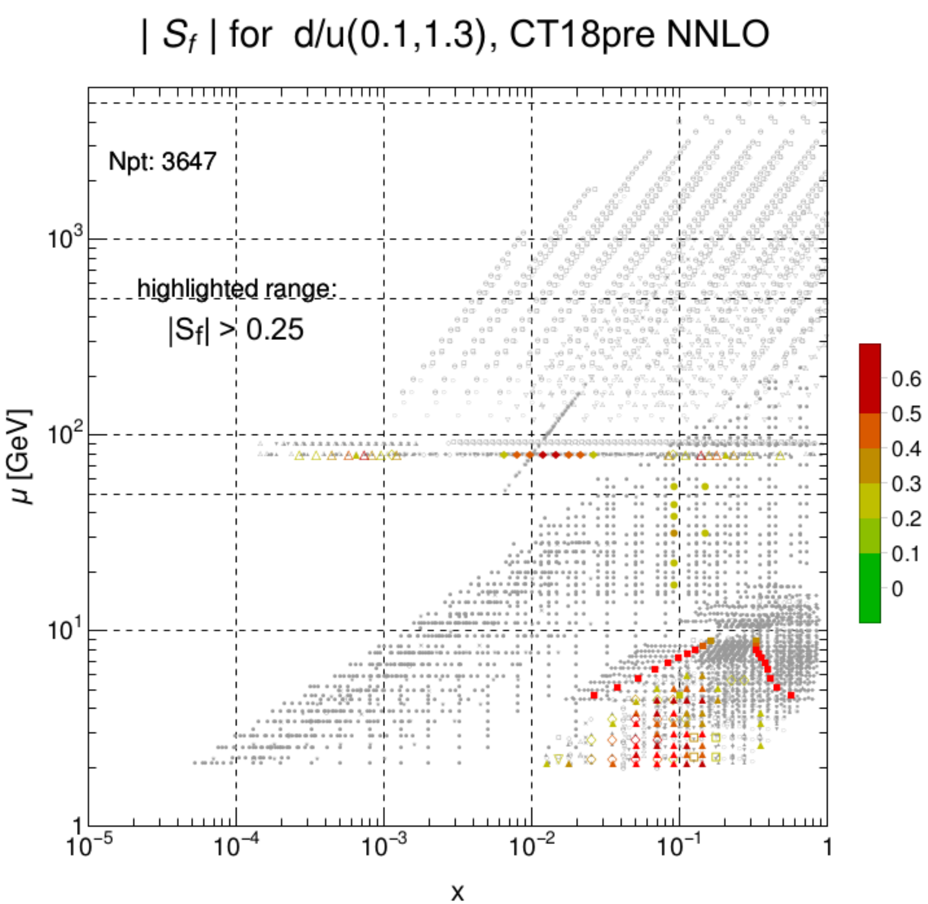
...for the gluon PDF in the Higgs region, $g(0.01, m_H)$

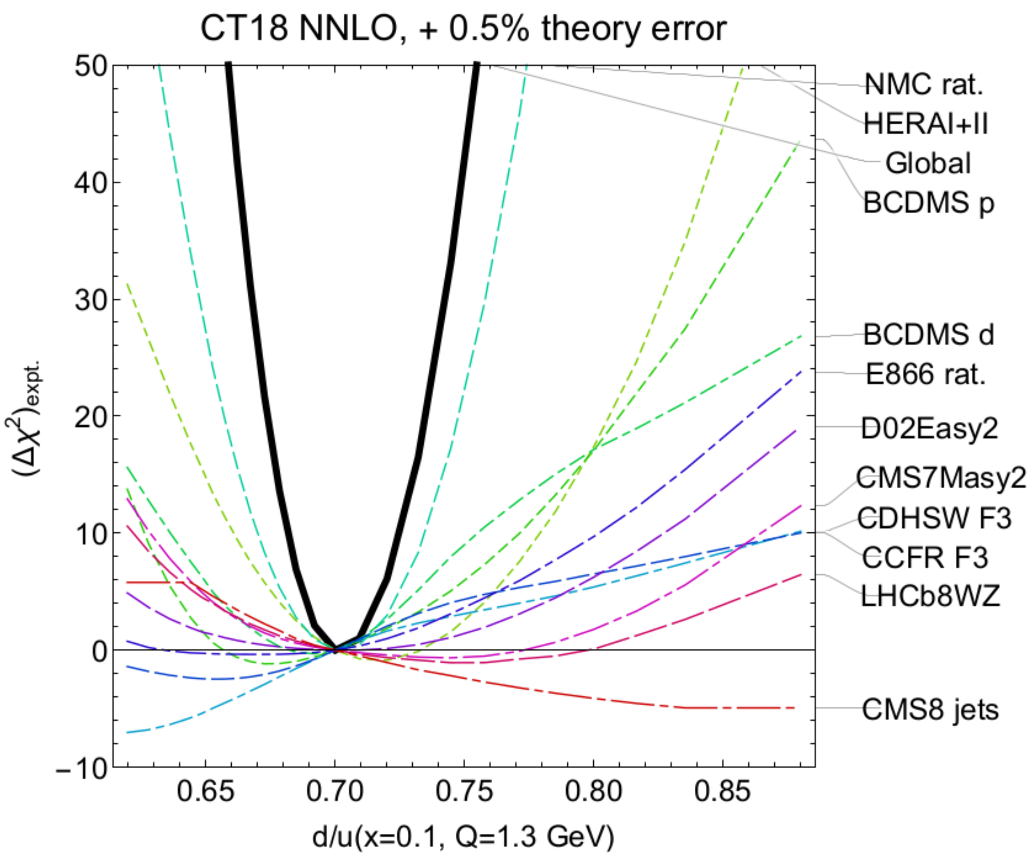


$g(x=0.01, \mu=125$ GeV$)$		
PDFSense		LM scan
CT14HERA2	CT18pre	CT18pre
HERAI+II'15	HERAI+II'15	HERAI+II'15
CMS8jets'17	CMS8jets'17	CMS8jets'17
CMS7jets'14	CMS7jets'14	ATL8ZpT'16
ATLAS7jets'15	E866pp'03	E866pp'03
E866pp'03	ATLAS7jets'15	ATLAS7jets'15
BCDMSd'90	BCDMSd'90	CCFR-F2'01
CCFR-F3'97	BCDMSp'89	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]

PDFSense predictions can be validated against actual fits





- PDFSense successfully predicts the highest impact data sets *before* fitting, as shown in this illustration for the large x PDF ratio d/u
- Lagrange Multiplier scans provide an independent test of which datasets most drive the global fit in connection with specific PDFs

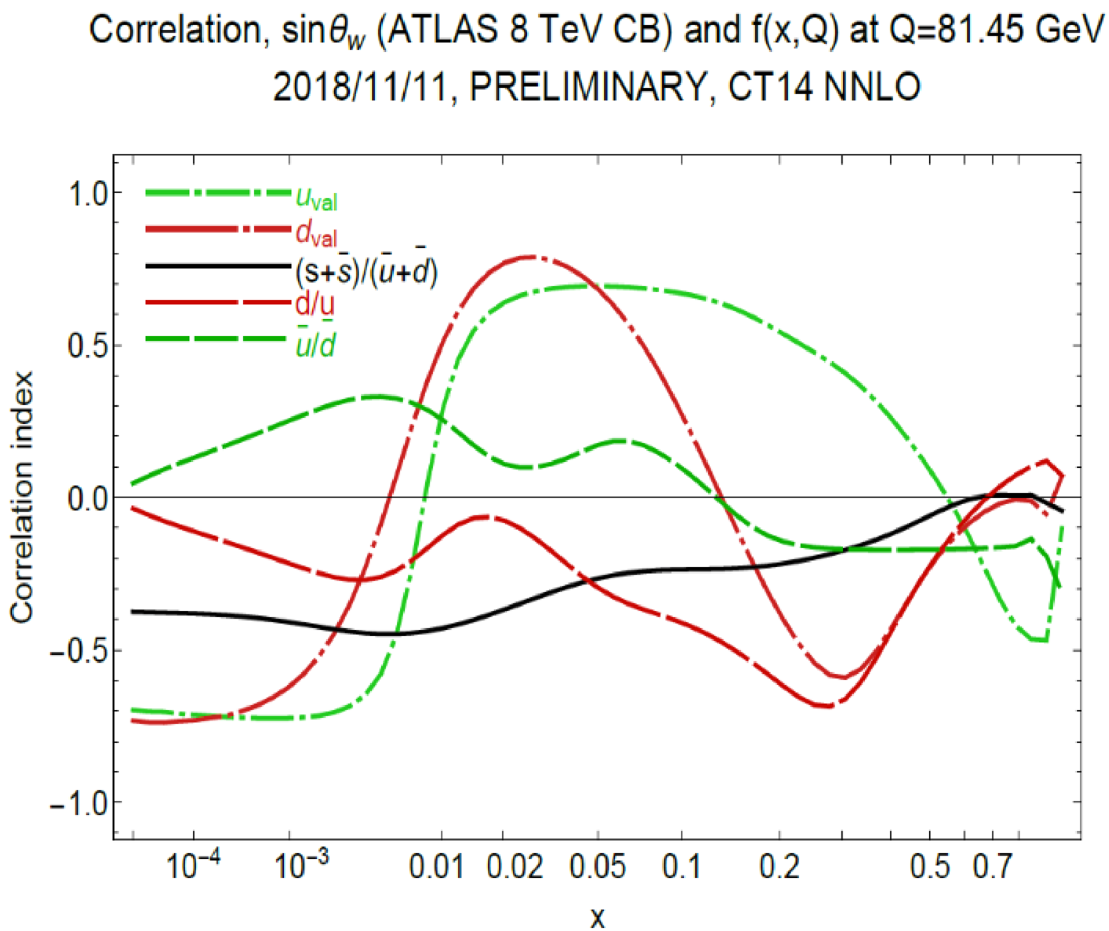
HERA and fixed-target (BCDMS, NMC) data are dominant!

$\sin^2\theta_W$ (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 $0.005 \lesssim x \lesssim 0.2$

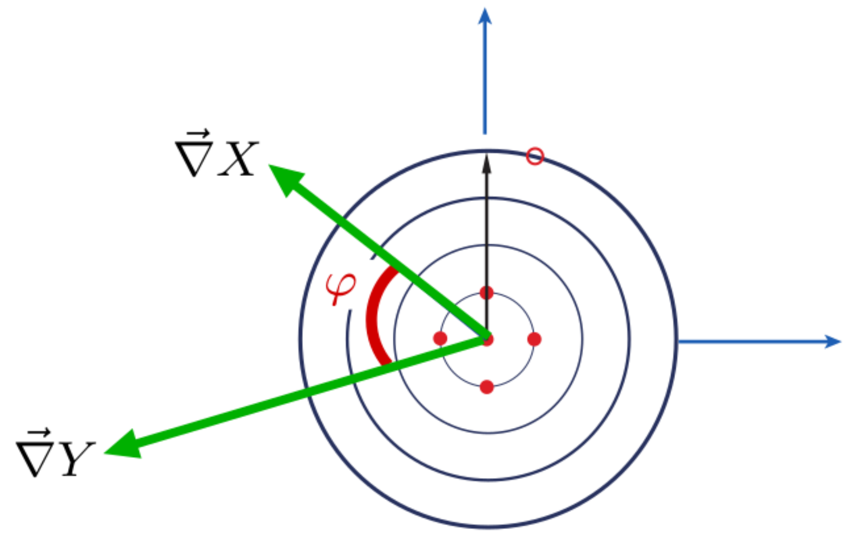
weak correlations with $\bar{u}$, $\bar{d}$, $\bar{s}$, g

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}

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

the *L2 sensitivity*



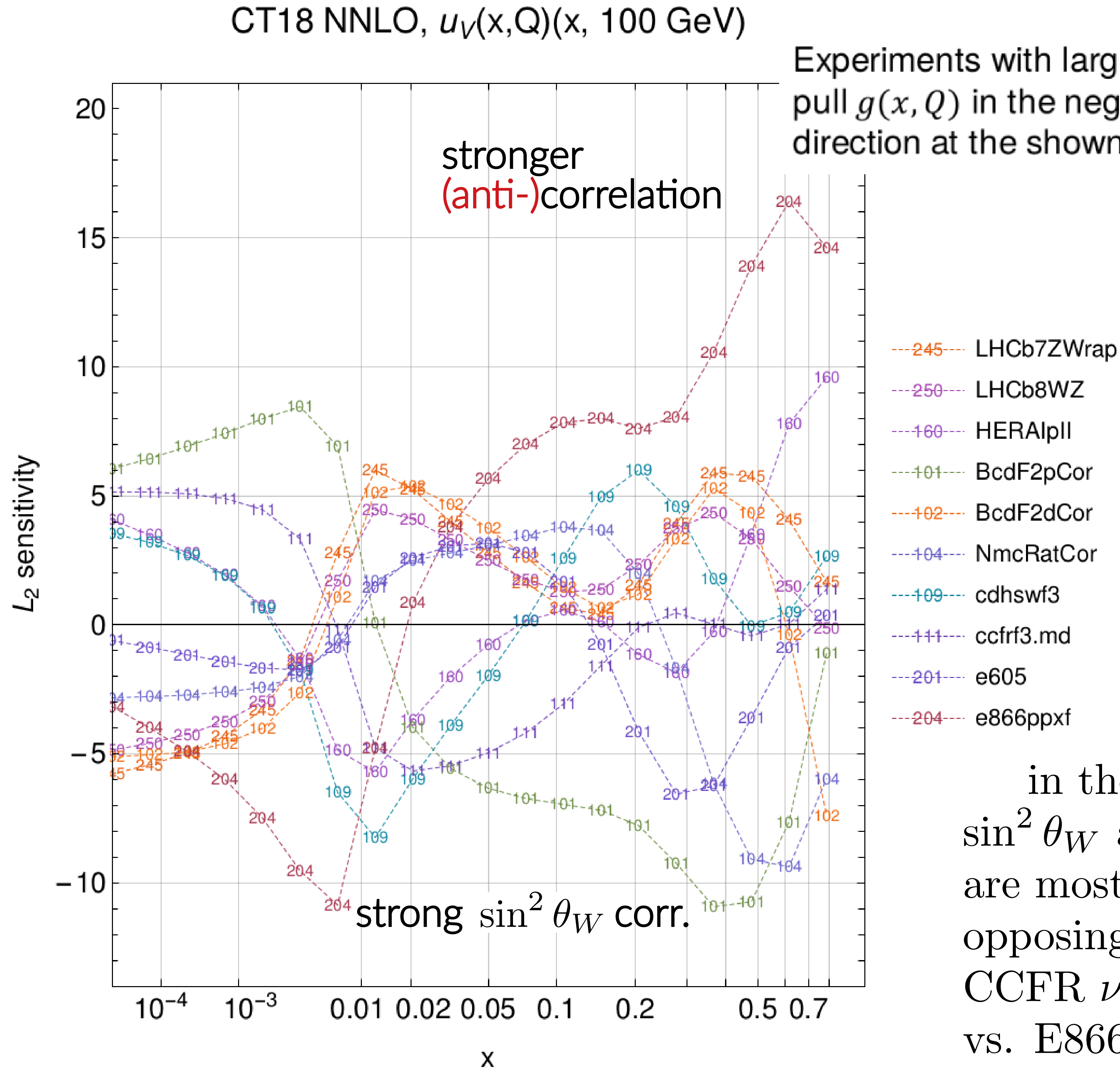
L_2 sensitivity. Take $X = f_a(x_i, Q_i)$ or $\sigma(f)$; $Y = \chi^2_E$ for experiment E. Find $\Delta Y(\vec{z}_{m,X})$ for the displacement $|\vec{z}_{m,X}| = 1$ along the direction $\vec{\nabla}X / \left|\vec{\nabla}X\right|$ (corresponding to $\Delta\chi^2_{tot} = T^2$ and $X(\vec{z}) = X(0) + \Delta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

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

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

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

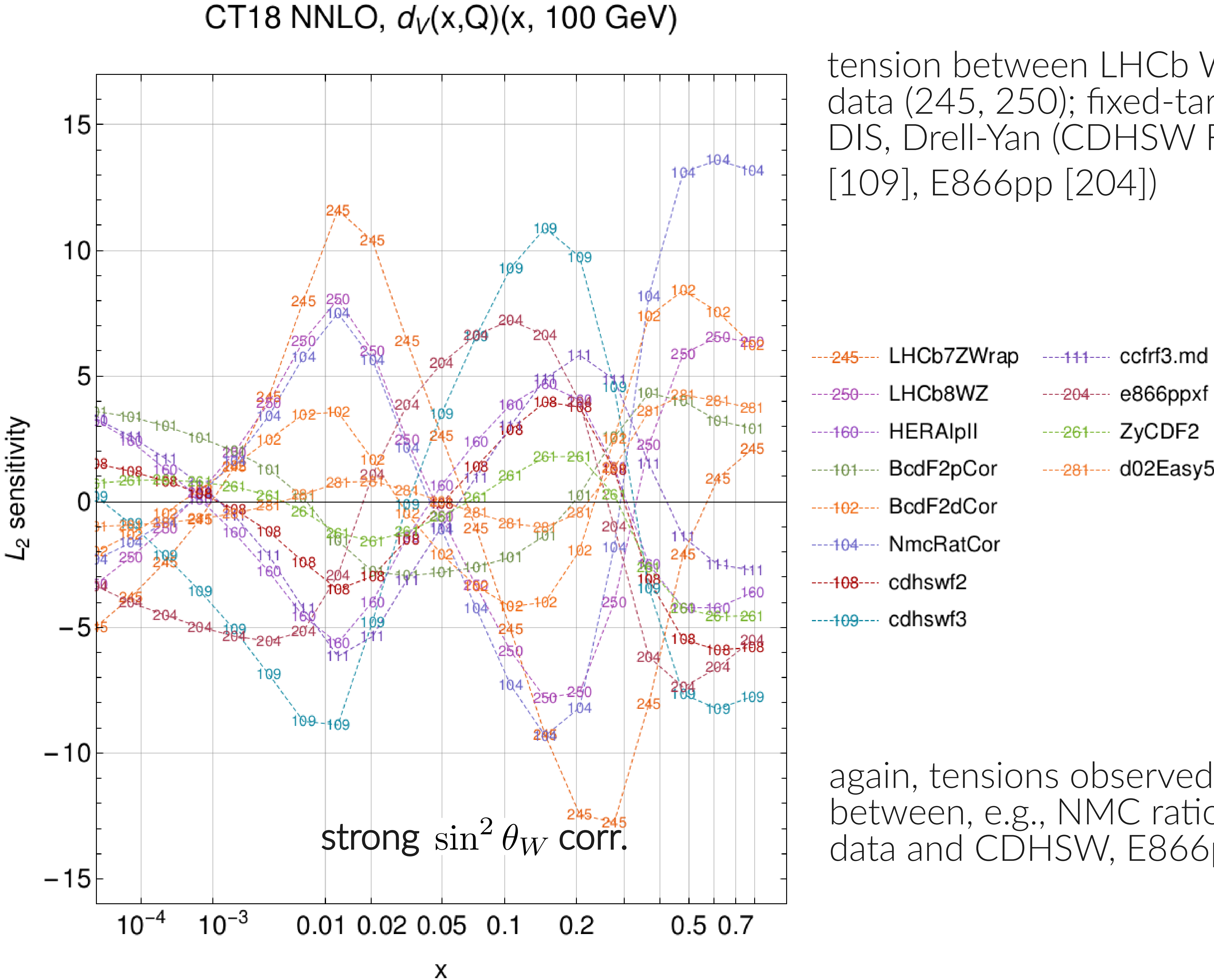
...extent to which total χ^2_E of specific expts. correlates with x-dep. of PDFs



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

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

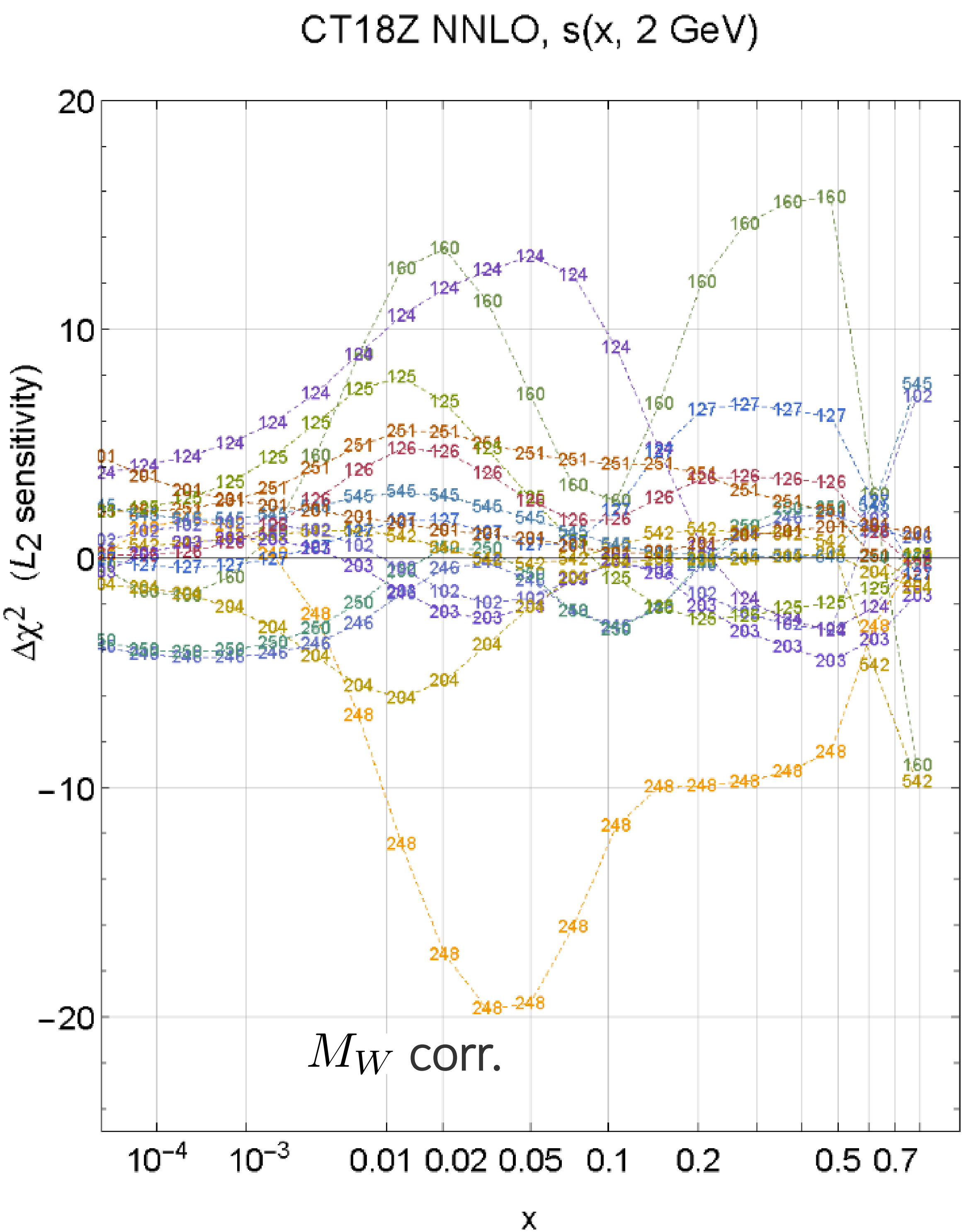
in the region where $\sin^2\theta_W$ and u_v are most correlated, opposing pulls from CCFR νDIS, BCDMS vs. E866 pp, NMC rat.



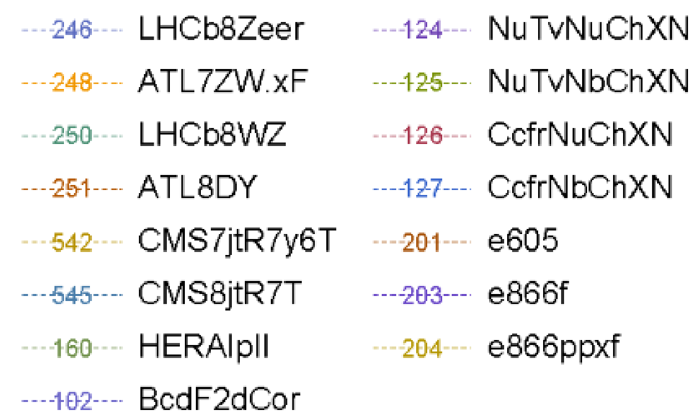
tension between LHCb W/Z data (245, 250); fixed-target DIS, Drell-Yan (CDHSW F_3 [109], E866pp [204])

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

L_2 sensitivity, strangeness: CT18Z



Most sensitive experiments



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!

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

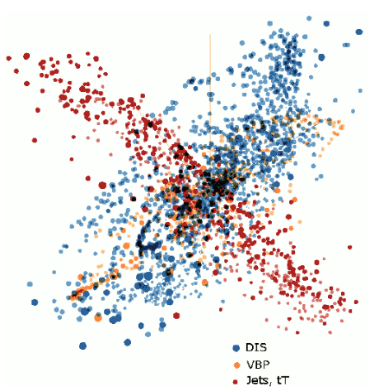
$$r_i(\vec{a}) = \frac{1}{s_i}\left(T_i(\vec{a}) - D_{i,sh}(\vec{a})\right)$$

these result in systematic shifts to data central values:

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

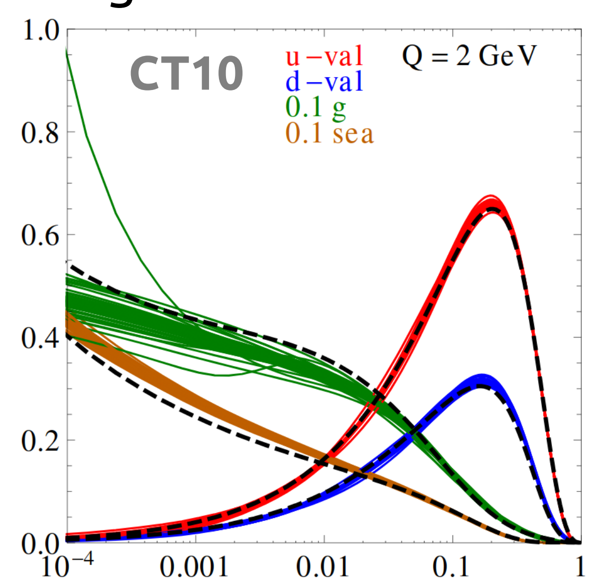
- a 56-dimensional parametric basis $\vec{a}$ is obtained by diagonalizing the Hessian matrix H determined from χ^2 (following a 28-parameter fit)

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

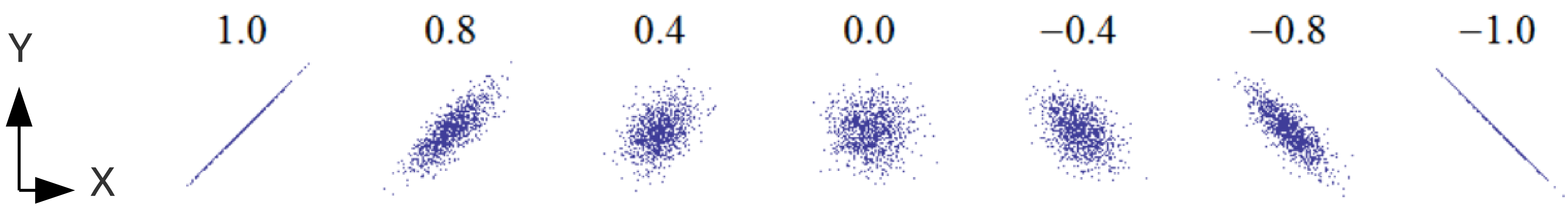


a brief statistical aside, ii

- **… 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



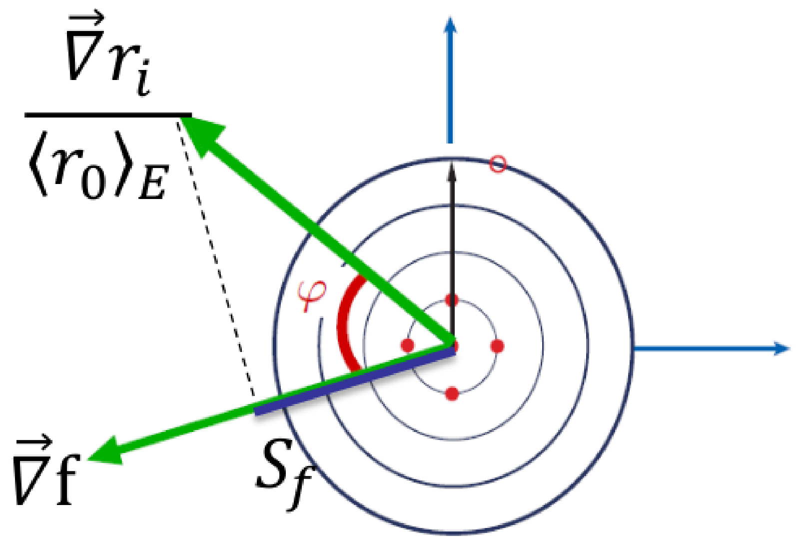
[X,Y] are exactly (anti-)correlated at the far (right) left above.

- 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.

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

x_i : parton mom. fraction

μ_i : factorization scale

deeply-inelastic scattering:

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

hadron-hadron collisions:

$AB \to CX$

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

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: $Q = \sqrt{(p_T^Z)^2 + (M_Z)^2},\ \ y_C = y_Z$