Mapping the impact of present and future experiments on hadron structure knowledge

T.J. Hobbs, SMU/CTEQ

Sept. 11th 2018

ECT*: Mapping Parton Distribution Amplitudes and Functions

arXiv:1803.02777

http://metapdf.hepforge.org/PDFSense/

…with **Bo-Ting Wang**, S. Doyle, J. Gao, T.-J. Hou,
Pavel Nadolsky, and F. Olness

owing to the nonperturbative nature of QCD, efforts to directly compute PDFs have generally relied on various effective prescriptions

→ Dyson-Schwinger/Bethe-Salpeter methodology

→ MIT Bag Model

→ light-front wave functions (example briefly described here)
[nonperturbative charm]

→ novel Lattice QCD calculations based in the LaMET approach

select examples!

- -

the most general phenomenological scheme, however, is **QCD global analysis**
(**see talk by P. Nadolsky**)

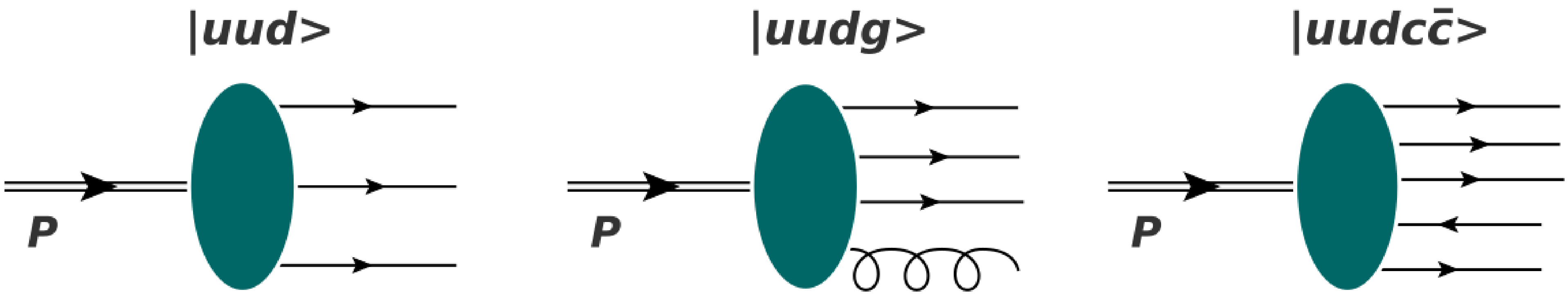
- driven by ***experimental data*** ; knowledge of pQCD; parametrization flexibility
- understanding the relationship between empirical data and resulting PDF predictions therefore becomes a very non-trivial problem

can we predict *a priori* how data might impact PDFs *without fitting* ?

(this talk!)

i) introductory example: the problem of modeling the proton's **nonperturbative charm**

→ or, how a lack of empirical constraints means that even simple models have large uncertainties



$\rightarrow$ original models possessed *scalar* vertices...

- Brodsky et al. (1980):

$$P(p \to uudc\bar{c}) \sim \left[M^2 - \sum_{i=1}^{5} \frac{k_{\perp i}^2 + m_i^2}{x_i} \right]^{-2}$$

$\rightarrow$ produces *intrinsic* PDF, $c^{\mathrm{IC}}(x) = \bar{c}^{\mathrm{IC}}(x)$

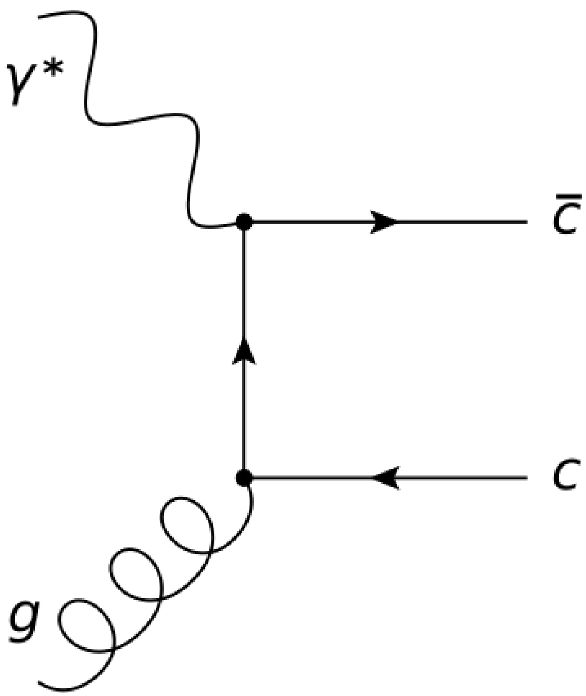
- Blümlein (2015):

$$\tau_{life} = \frac{1}{\sum_i E_i - E} = \frac{2P}{\left(\sum_{i=1}^{5} \frac{k_{\perp i}^2 + m_i^2}{x_i} - M^2 \right)} \Bigg|_{\sum_j x_j = 1} \quad \textbf{vs.} \quad \tau_{int} = \frac{1}{q_0}$$

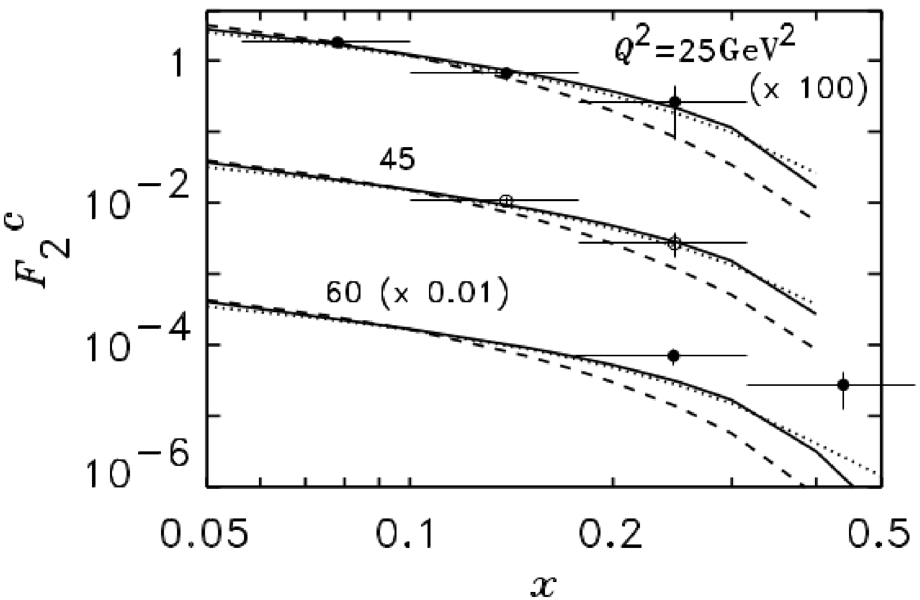
$\rightarrow$ comparison constrains $x - Q^2$ space over which IC is observable

this is in contrast to the usual pQCD treatment of the charm PDF

- $c(x, Q^2 \leq m_c^2) = \bar{c}(x, Q^2 \leq m_c^2) = 0$



F. M. Steffens, W. Melnitchouk and A. W. Thomas, Eur. Phys. J. C **11**, 673 (1999) [hep-ph/9903441].



- *intermediate* Q^2:

$$F^c_{2,\,\mathrm{PGF}}(x, Q^2) \;=\; \frac{\alpha_s(\mu^2)}{9\pi} \int_x^{z'} \frac{dz}{z}\, C^{\mathrm{PGF}}(z, Q^2, m_c^2) \cdot xg\left(\frac{x}{z}, \mu^2\right)$$

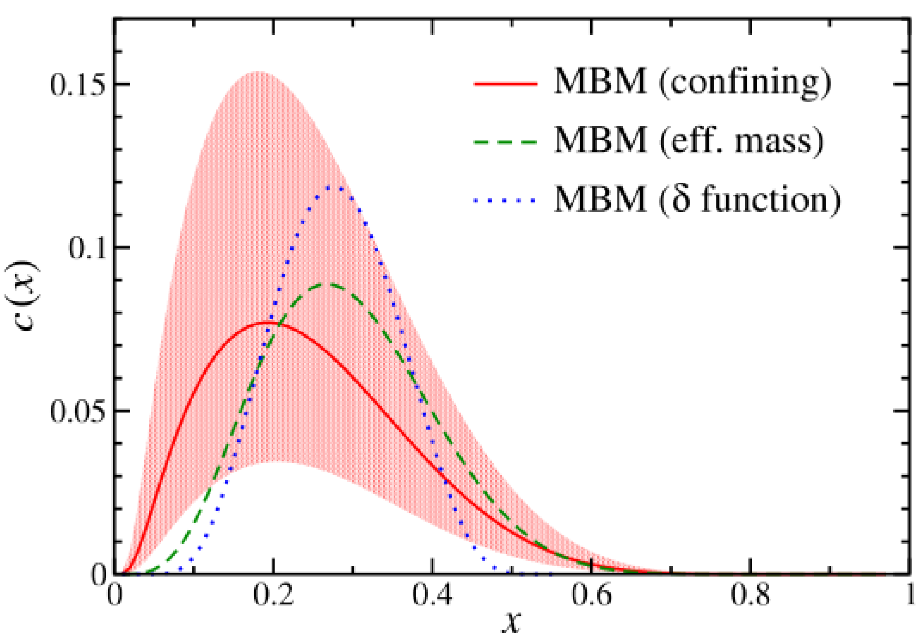
- *high* Q^2:

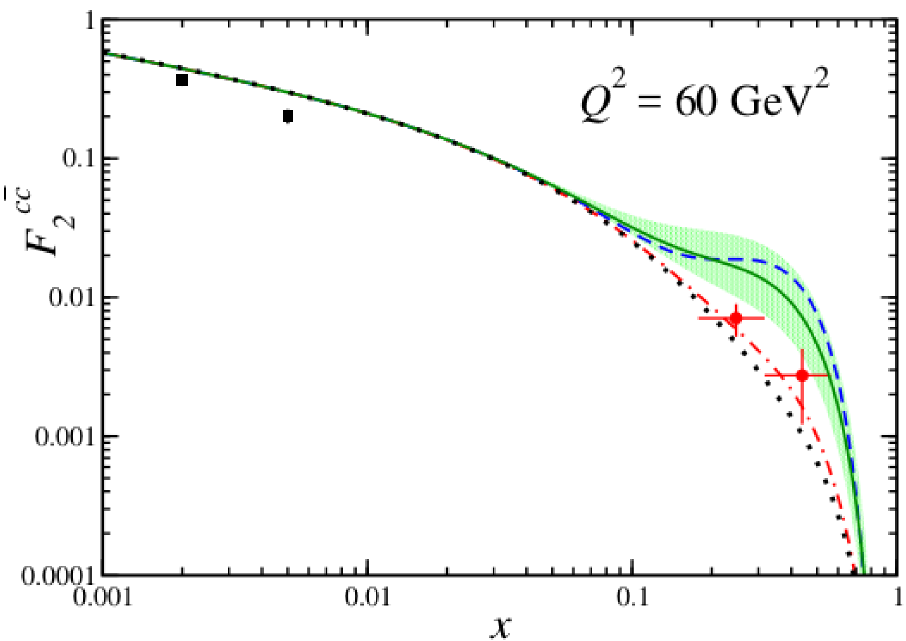
massless **DGLAP** (i.e., *variable flavor-number* schemes) typical model calculations (here, MBM) have large uncertainties:

- tune **universal** cutoff $\Lambda = \hat{\Lambda}$ to fit **ISR** $pp \to \Lambda_c X$ collider data

multiplicities, **momentum sum**: [TJH, Londergan, Melnitchouk; PRD89 (2014).]

$$\langle n \rangle^{(\mathrm{charm})}_{MB} = 2.40\% \, {}^{+2.47}_{-1.36}; \qquad P_c := \langle x \rangle_{\mathrm{IC}} = 1.34\% \, {}^{+1.35}_{-0.75}$$





$$F_2^{c\bar{c}}(x, Q^2) = \frac{4x}{9}\left[c(x,Q^2) + \bar{c}(x,Q^2)\right]$$

$\to$ evolve to EMC scale, $Q^2 = 60$ GeV^2

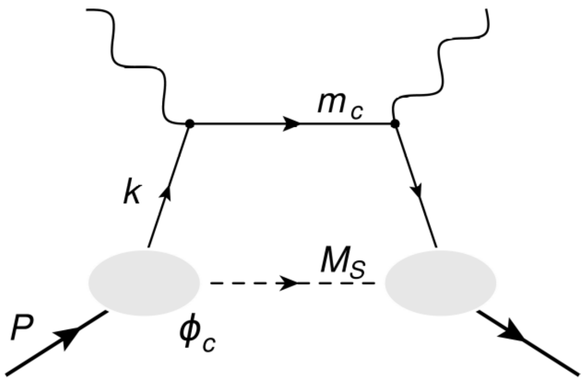
low-x H1/ZEUS data check *massless* **DGLAP** evolution

build a **light-front** model for the IC PDF

- use a scalar spectator picture of the 5-quark Fock state:

$$|\Psi_P^\lambda(P^+,\mathbf{P}_\perp)\rangle = \frac{1}{16\pi^3}\sum_{q=c,\bar{c}}\int \frac{dx\,d^2\mathbf{k}_\perp}{\sqrt{x(1-x)}}$$

TJH, Alberg, Miller; PRD**96** (2017) no.7, 074023.



$$\times\,\psi^\lambda_{q\lambda_q}(x,\mathbf{k}_\perp)\,|q;xP^+,x\mathbf{P}_\perp+\mathbf{k}_\perp\rangle$$

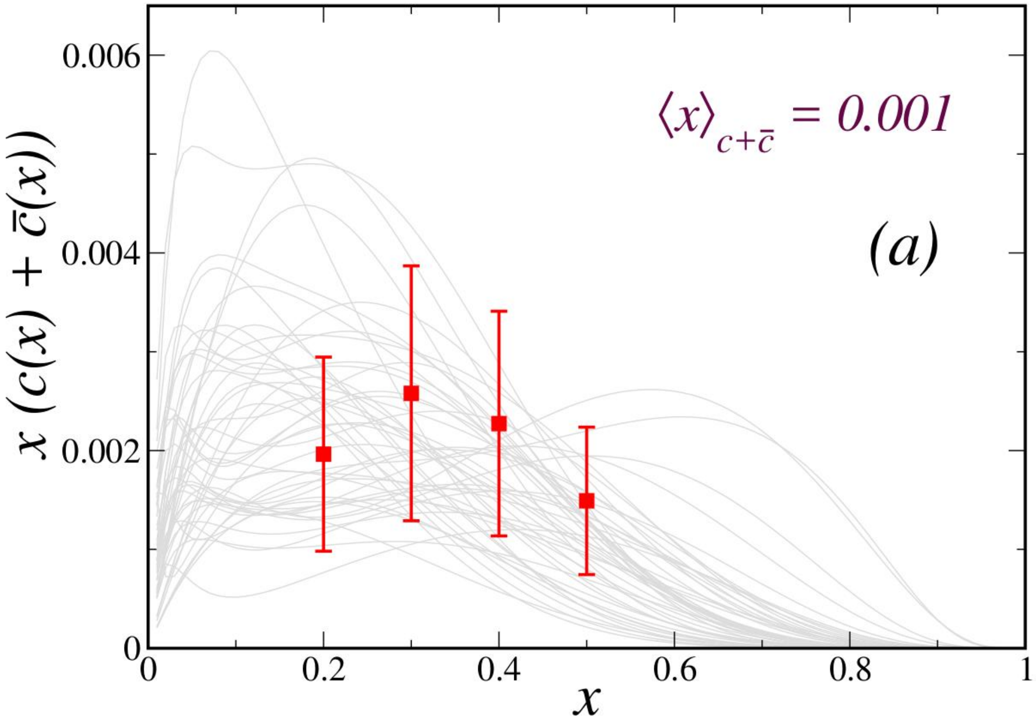
$$F_2^{c\bar{c}}(x,Q^2=m_c^2) = \frac{4x}{9}\left(c(x)+\bar{c}(x)\right)$$

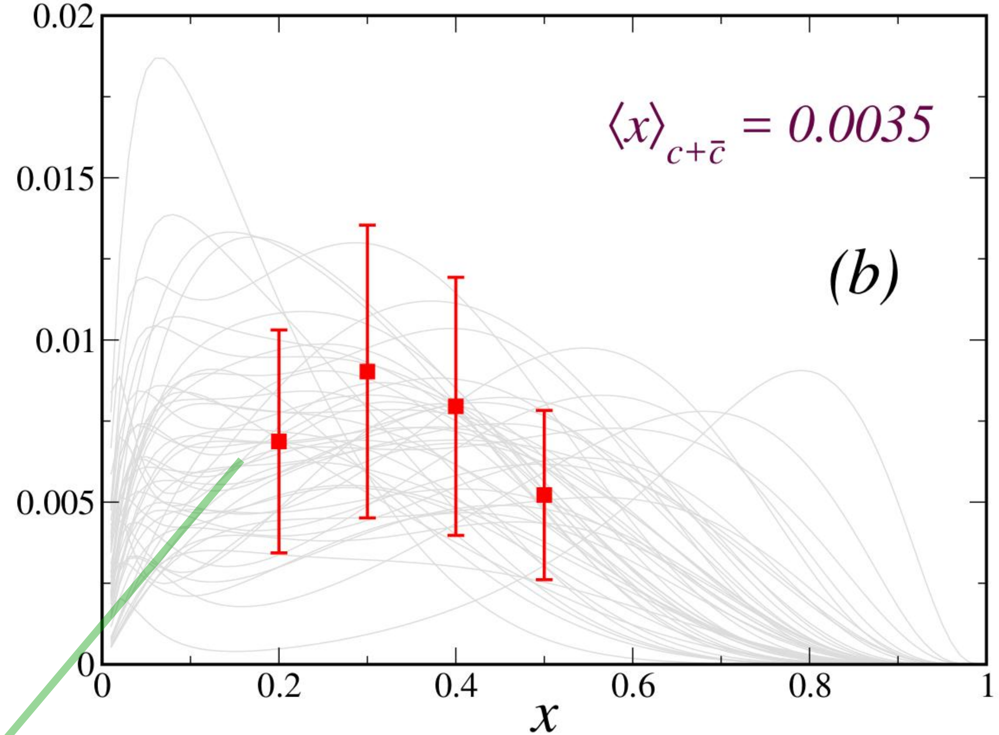
$$c(x) = \frac{1}{16\pi^2}\int\frac{dk_\perp^2}{x^2(1-x)}\left[\frac{k_\perp^2+(m_c+xM)^2}{(M^2-s_{cS})^2}\right]\left|\phi_c(x,k_\perp^2)\right|^2$$

use a power-law (γ=3) covariant vertex function,

$$\phi_c(x,k_\perp^2) = \sqrt{g_c}\left(\frac{\Lambda_c^2}{t_c-\Lambda_c^2}\right)^\gamma$$

$$\begin{cases} s_{cS}\left(x,k_\perp^2\right) = \dfrac{1}{x\left(1-x\right)}\left(k_\perp^2+\left(1-x\right)m_c^2+x\,M_S^2\right) & \text{invariant mass} \\ t_c\left(x,k_\perp^2\right) = \dfrac{1}{1-x}\left(-k_\perp^2+x\left[(1-x)M^2-M_S^2\right]\right) & \text{covariant } k^2 \end{cases}$$





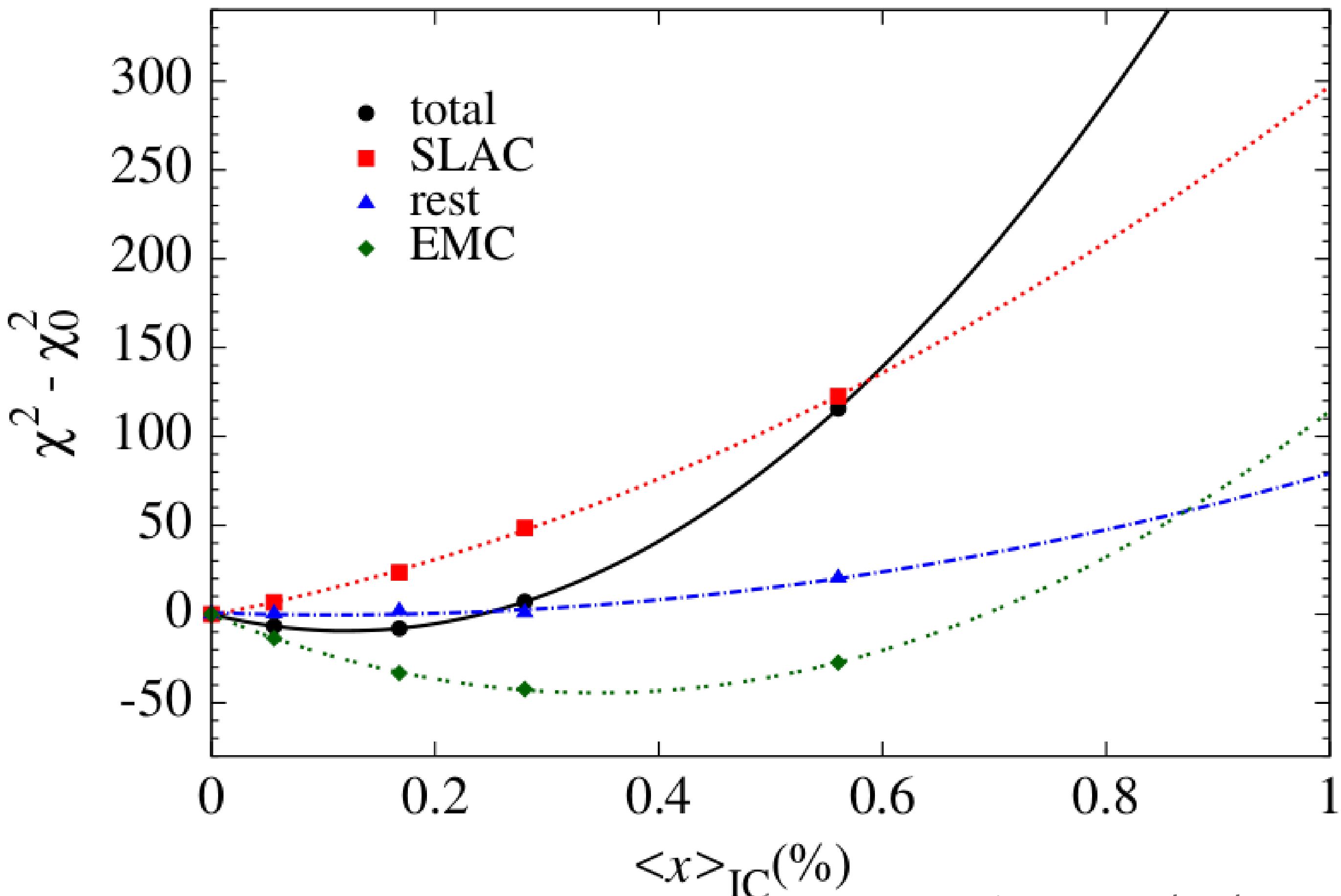
- we constrain the model with hypothetical **pseudo-data** (taken from the \`confining' MBM) of a given $\langle x \rangle_{\rm IC} \pm 50\%$

(input data normalizations are inspired by the just-described global analysis)

$$\begin{cases} \langle x \rangle_{\rm IC} = 0.001 & \text{[upper limit tolerated by the full fit/dataset]} \\ \langle x \rangle_{\rm IC} = 0.0035 & \text{[central value preferred by EMC data alone]} \end{cases}$$

- **in general, pseudo-data permit a very broad set of intrinsic charm PDF behaviors**

very little data (**other than EMC**) suggest the presence of IC

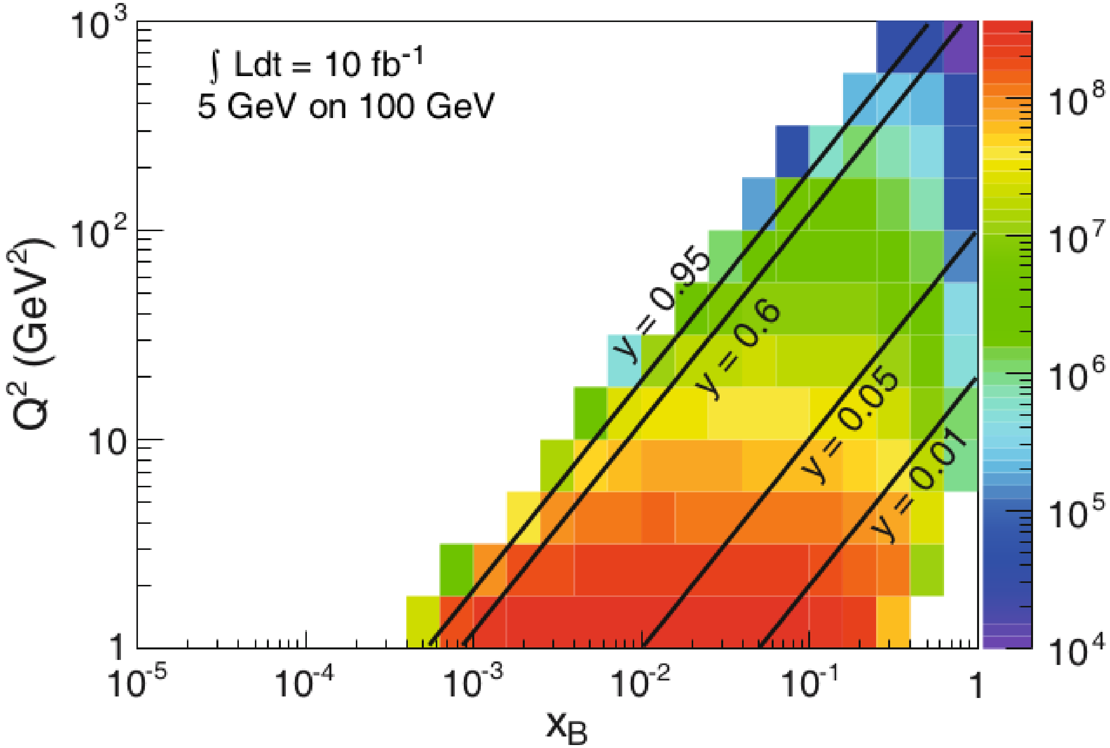


Jimenez-Delgado, TJH, Londergan, Melnitchouk, PRL**114** (2015) no.8, 082002.

EMC alone: $\langle x \rangle_{\rm IC} = 0.3 - 0.4\%$

+ SLAC/'REST': $\boldsymbol{\langle x \rangle_{\rm IC} = 0.13 \pm 0.04\%}$

...but $F_2^{c\bar{c}}$ poorly fit — $\chi^2 \sim 4.3$ per datum!



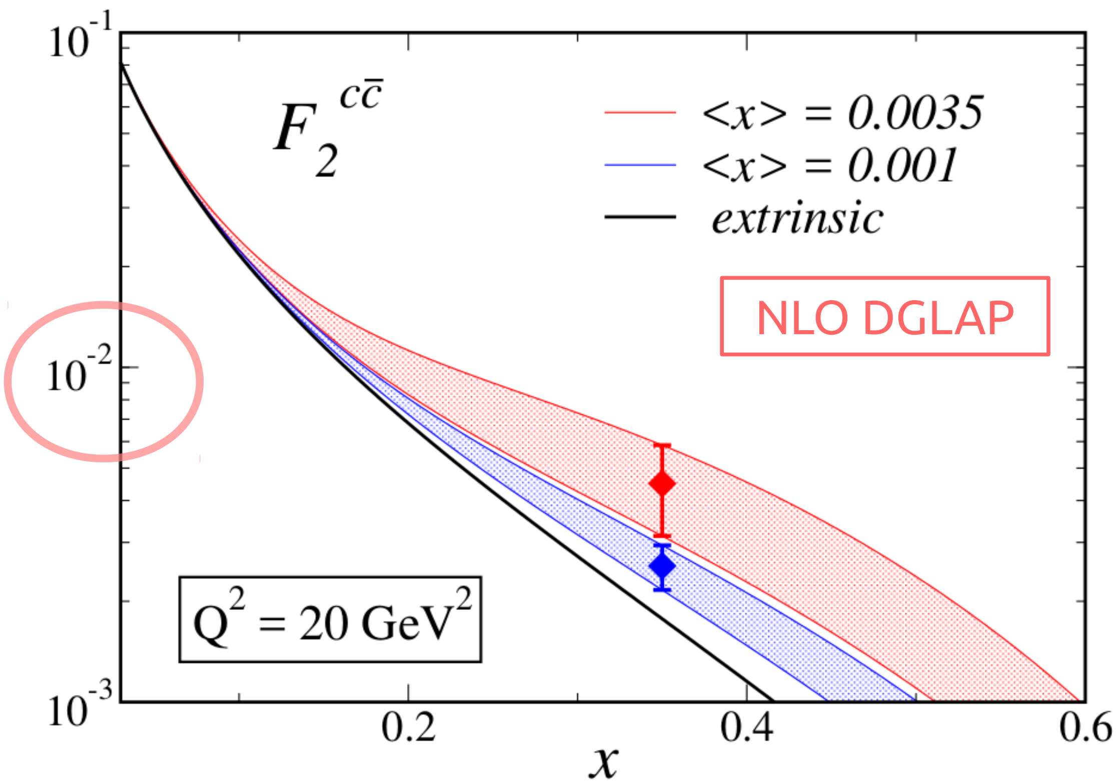
EIC Whitepaper, Eur. Phys. J. A (2016) **52**: 268

- e.g., JLEIC-like scenario:

$$\sqrt{s} = 45\,\mathrm{GeV}$$

- a definitive measurement would simply **reprise the EMC observation of $F^{c\bar{c}}_2$**

- still, considerable precision will be needed to be sensitive at the necessary level



which measurements/data are likely to give the needed information to improve the charm PDF?

… and, more broadly,

can we anticipate the effect of experiments on the PDFs?

ii) **PDFSense**: a tool for mapping the impact of the world's data

arXiv:1803.02777

http://metapdf.hepforge.org/PDFSense/

Bo-Ting Wang, S. Doyle, J. Gao, T.-J. Hou, Pavel Nadolsky, and F. Olness

question: can we reliably anticipate the impact on knowledge of nucleon structure (here, the **PDFs**) of experiments that have not yet been performed? (…like an EIC…)

or

given our current theoretical and PDF knowledge how hard must experimentalists work to push the envelope?

→ what, how well, and where should experimentalists measure?

(… and where should we concentrate our efforts as theorists!)

- the most robust means of determining PDFs from experimental data is the technique of **QCD global analysis**

→ **BUT:** performing full QCD fits, especially in the absence of actual data, is a time-consuming and arduous task

→ as an alternative, we have a new technology **PDFSENSE** for quickly assessing the sensitivity of data to experimental (pseudo-)data

arXiv:1803.02777

→ **this framework is ideally suited to examining the possible constraints from a future EIC**

hadron PDFs from **QCD global fits** and mapping the **expected impact** of empirical data

- we want to determine from global analyses:

$$f_{q/p}(x,\mu^2) = \int \frac{d\xi^-}{4\pi} e^{-i\xi^- k^+} \langle p\,|\overline{\psi}(\xi^-)\gamma^+ \mathcal{U}(\xi^-,0)\psi(0)\,|p\rangle$$

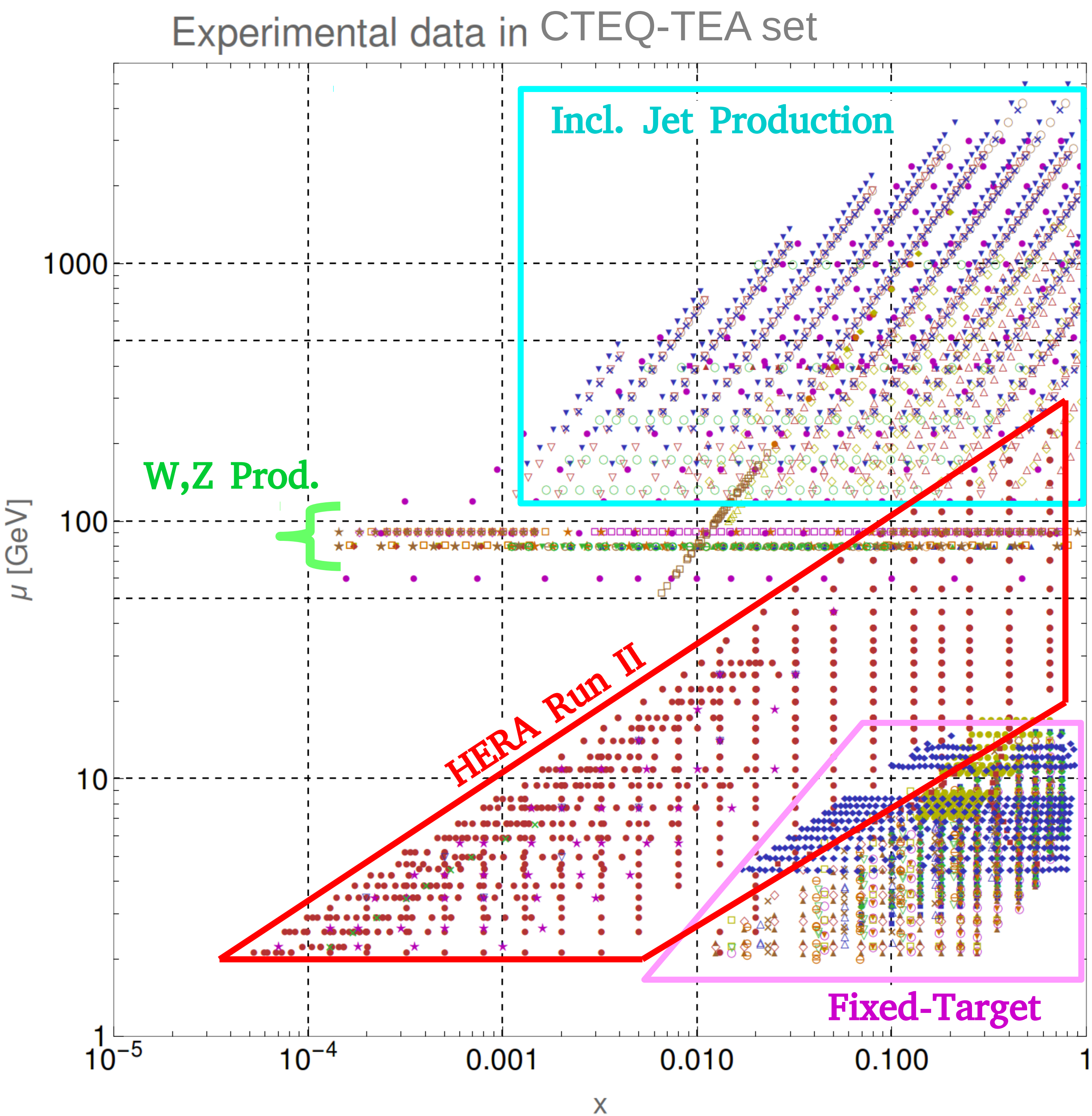
- in practice, this is achieved by leveraging our knowledge of pQCD and experimental data to constrain a flexible PDF parametrization:

  (See yesterday's talk by Pavel.)

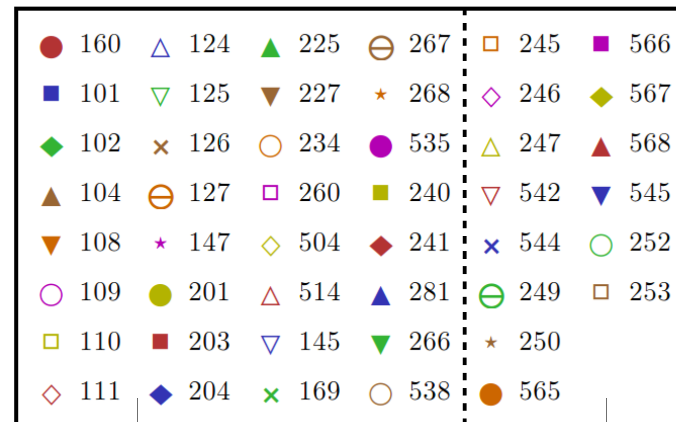
$$f_{q/h}(x,Q_0^2) = a_{q_0} x^{a_{q_1}} (1-x)^{a_{q_2}}\, P[x,\{a_{q\,n-3}\}]$$

- this is, however, an extremely challenging undertaking!
  - → global fits are time consuming/computationally expensive
  - → **what are the highest impact data**? …and what theory is required?

- **might we instead identify data with the greatest sensitivity to PDF physics *before* fitting and direct global analyses accordingly?**



the problem : modern global PDF analyses must contend with LARGE data sets involving many physical processes and channels!



individual experiments contributing to CT14 analysis of HERA Run II ("CT14H2")

new, unfitted datasets from, e.g., LHC at 7, 8 TeV

- Can we weigh the influence of datasets **WITHOUT** each time performing a full global analysis? **. . . we could then predict the impact of unfitted data and guide fits . . .**

Experiments in the CT14 HERA2 fit

ID#	Experimental dataset		N_{pt}
101	BCDMS F_2^p	[58]	337
102	BCDMS F_2^d	[59]	250
104	NMC F_2^d/F_2^p	[60]	123
108	CDHSW F_2^p	[61]	85
109	CDHSW F_3^p	[61]	96
110	CCFR F_2^p	[62]	69
111	CCFR xF_3^p	[63]	86
124	NuTeV $\nu\mu\mu$ SIDIS	[53]	38
125	NuTeV $\bar{\nu}\mu\mu$ SIDIS	[53]	33
126	CCFR $\nu\mu\mu$ SIDIS	[54]	40
127	CCFR $\bar{\nu}\mu\mu$ SIDIS	[54]	38
145	H1 σ_r^b (57.4 pb^{-1})	[64] [65]	10
147	Combined HERA charm production (1.504 fb^{-1})	[52]	47
160	HERA1+2 Combined NC and CC DIS (1 fb^{-1})	[6]	1120
169	H1 F_L (121.6 pb^{-1})	[66]	9

ID#	Experimental dataset		N_{pt}
201	E605 DY	[67]	119
203	E866 DY, $\sigma_{pd}/(2\sigma_{pp})$	[68]	15
204	E866 DY, $Q^3 d^2\sigma_{pp}/(dQdx_F)$	[69]	184
225	CDF Run-1 $A_e(\eta^e)$ (110 pb^{-1})	[70]	11
227	CDF Run-2 $A_e(\eta^e)$ (170 pb^{-1})	[71]	11
234	DØ Run-2 $A_\mu(\eta^\mu)$ (0.3 fb^{-1})	[72]	9
240	LHCb 7 TeV W/Z muon forward-η Xsec (35 pb^{-1})	[73]	14
241	LHCb 7 TeV W $A_\mu(\eta^\mu)$ (35 pb^{-1})	[73]	5
260	DØ Run-2 Z $d\sigma/dy_Z$ (0.4 fb^{-1})	[74]	28
261	CDF Run-2 Z $d\sigma/dy_Z$ (2.1 fb^{-1})	[75]	29
266	CMS 7 TeV $A_\mu(\eta)$ (4.7 fb^{-1})	[76]	11
267	CMS 7 TeV $A_e(\eta)$ (0.840 fb^{-1})	[77]	11
268	ATLAS 7 TeV W/Z Xsec, $A_\mu(\eta)$ (35 pb^{-1})	[78]	41
281	DØ Run-2 $A_e(\eta)$ (9.7 fb^{-1})	[79]	13
504	CDF Run-2 incl. jet ($d^2\sigma/dp_T^j dy_j$) (1.13 fb^{-1})	[49]	72
514	DØ Run-2 incl. jet ($d^2\sigma/dp_T^j dy_j$) (0.7 fb^{-1})	[50]	110
535	ATLAS 7 TeV incl. jet ($d^2\sigma/dp_T^j dy_j$) (35 pb^{-1})	[80]	90
538	CMS 7 TeV incl. jet ($d^2\sigma/dp_T^j dy_j$) (5 fb^{-1})	[81]	133

New experiments in the CT17pre fit

1. LHCb 7 TeV Z/W muon rapidity 1505.07024
2. LHCb 8 TeV Z rapidity 1503.00963
3. CMS 8 TeV W lept. asymmetry 1603.01803
4. LHCb 8 TeV Z/W muon rapidity 1511.08039
5. ATLAS 7 TeV Z p_T 1512.02192
6. CMS incl. jet 7 TeV, R=0.7 1406.0324
7. ATLAS incl. jet at 7 TeV, R=0.6 1410.8857
8. CMS incl. jet at 8 TeV, R=0.7 1609.05331
9. ATLAS 8 TeV $t\bar{t}$ p_T 1511.04716
10. ATLAS 8 TeV $t\bar{t}$ $m_{t\bar{t}}$ 1511.04716
11. CMS 8 TeV $t\bar{t}$ $d^2\sigma/dp_{Tt}dy_t$ 1703.01630

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

$$\sigma(AB \to W/Z+X) = \sum_n \alpha_s^n(\mu_R^2) \sum_{a,b} \int dx_a dx_b$$

$$\times\, f_{a/A}(x_a,\mu^2)\, \hat{\sigma}^{(n)}_{ab\to W/Z+X}\left(\hat{s},\, \mu^2, \mu_R^2\right) f_{b/B}(x_b,\mu^2)$$

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*** :

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

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

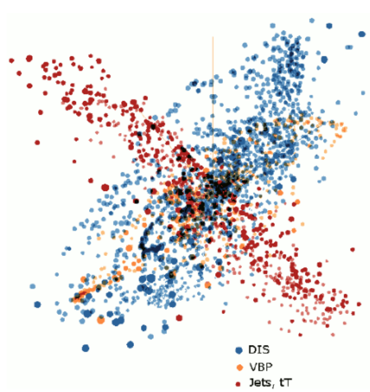
$$r_i(\vec{a}) = \frac{1}{s_i}\left(T_i(\vec{a}) - D_{i,sh}(\vec{a})\right)$$

these result in systematic shifts to data central values:

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

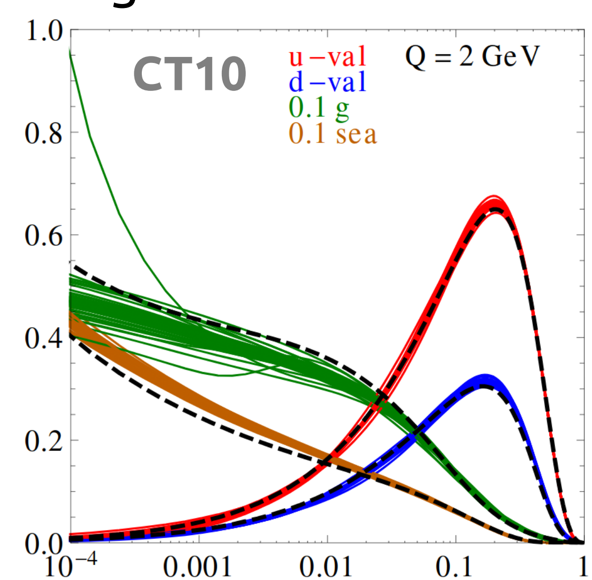
- a 56-dimensional parametric basis $\vec{a}$ is obtained by diagonalizing the Hessian matrix H determined from χ^2 (following a 28-parameter fit)

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

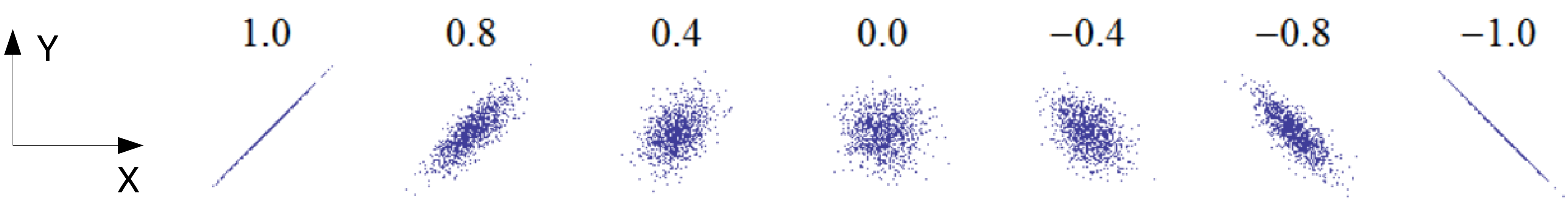


a brief statistical aside, ii

- … 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



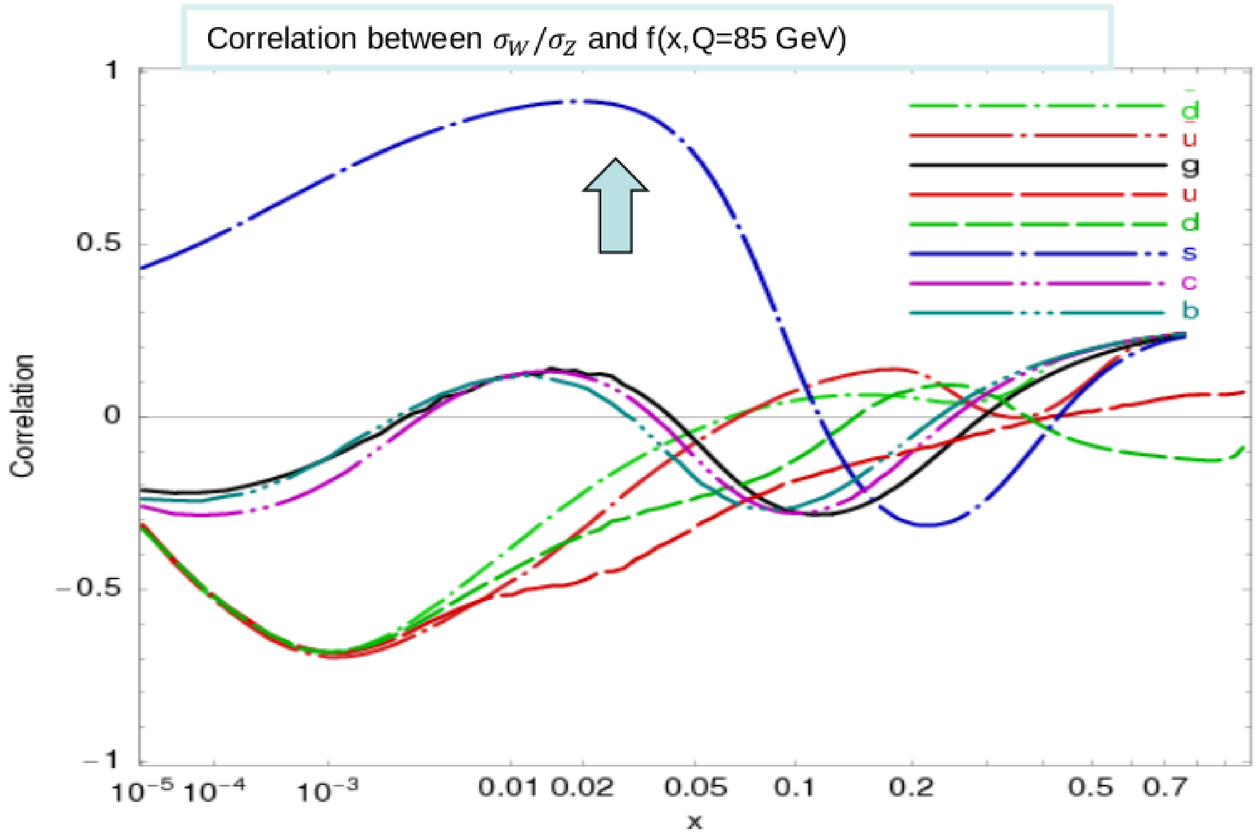
[X,Y] are exactly (anti-)correlated at the far (right) left above.

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

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

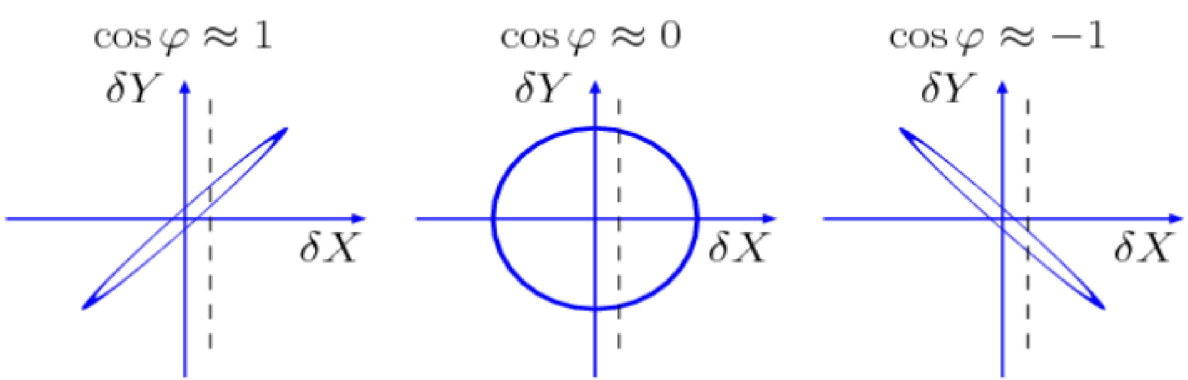
...we may turn to the Pearson correlations between PDFs and δ_i, but we first note

Correlations carry useful, but limited information



CTEQ6.6 [arXiv:0802.0007]: $\cos\varphi > 0.7$ shows that the ratio σ_W/σ_Z at the LHC must be sensitive to the strange PDF $s(x,Q)$

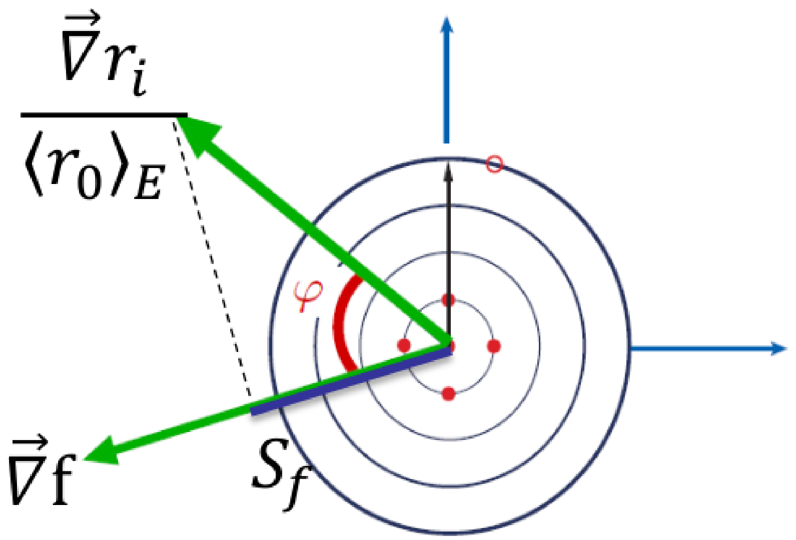
$\cos\varphi \approx \pm 1$ suggests that a measurement of X **may** impose tight constraints on Y



But, Corr[X,Y] between **theory** cross sections X and Y does not tell us about **experimental** uncertainties

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.

2^{nd} 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

$$x_i \;:\; \text{parton mom. fraction}$$

$$\mu_i \;:\; \text{factorization scale}$$

deeply-inelastic scattering:

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

hadron-hadron collisions:

$AB \to CX$

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

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

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

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

etc...

iii) predicting the impact of LHC Run 1 on collinear PDFs

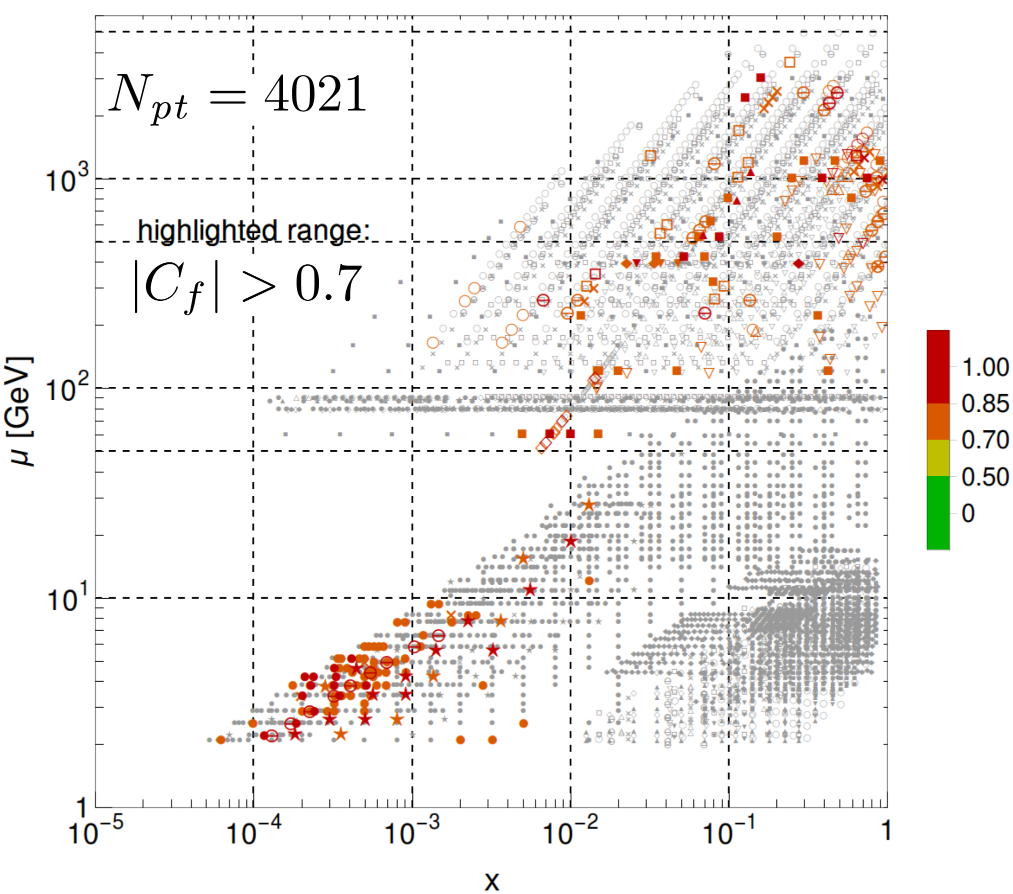
the sensitivity reveals a richer landscape than the correlation!

$$C_f(x_i, \mu_i) \equiv \mathrm{Corr}[f(x_i, \mu_i), r_i]$$

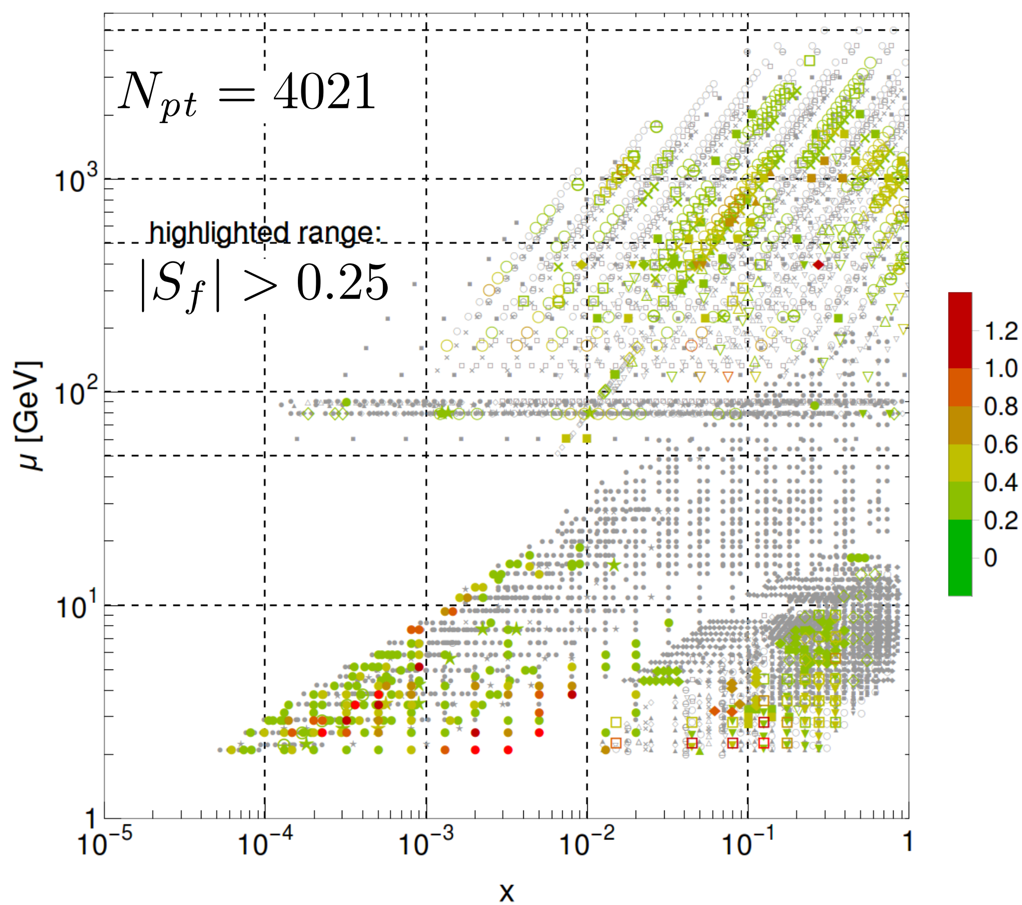
$$S_f(x_i, \mu_i) \equiv \frac{\delta^{(\mathrm{PDF})} r_i}{\sqrt{\frac{1}{N}\sum_{i=1}^{N} r_i^2}} C_f(x_i, \mu_i)$$

- broad outlays of the experimental parameter space are shown to be sensitive, but are **missed by the correlation**

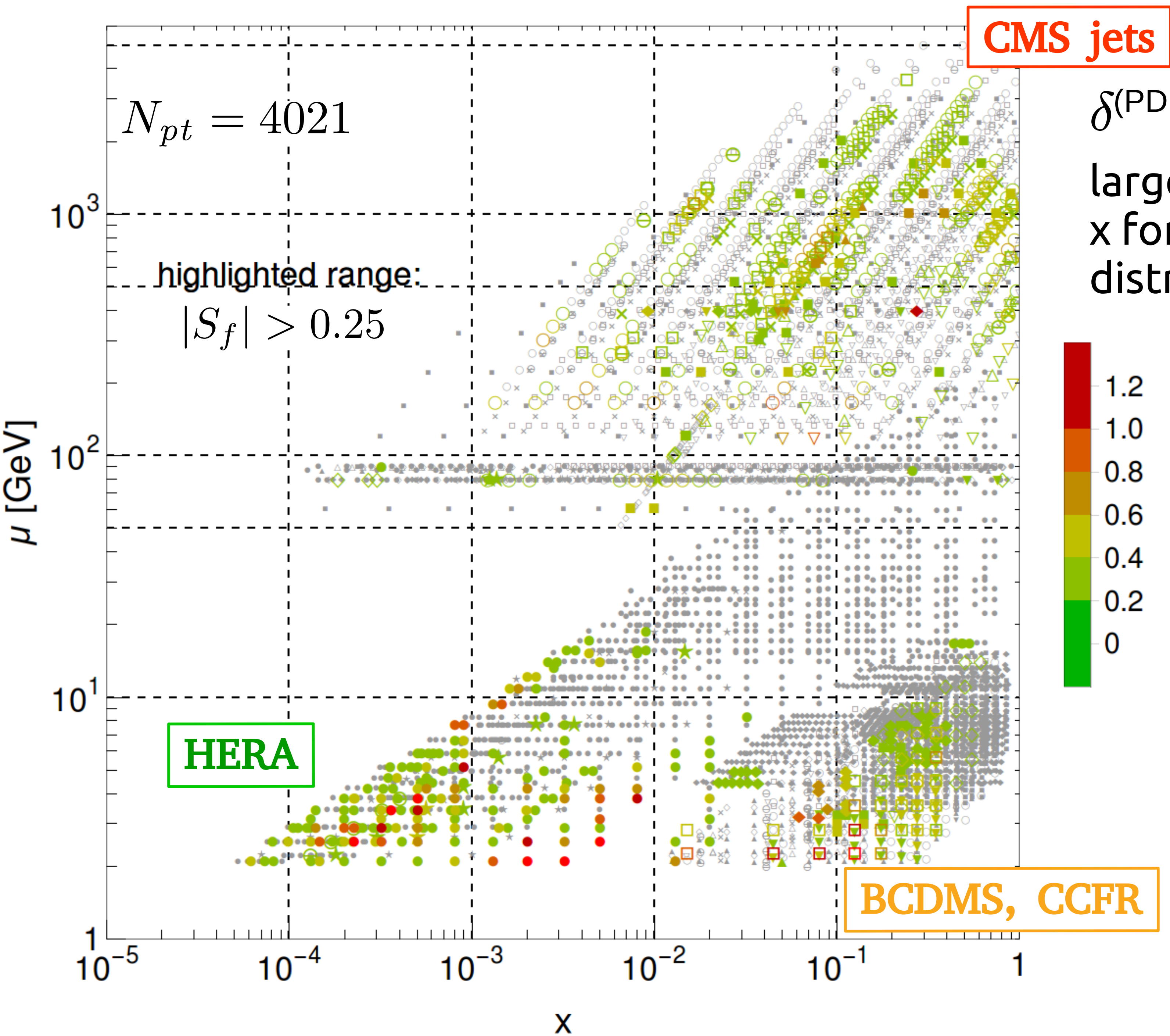
| C_f | for g(x,μ), CT14HERA2NNLO



| S_f | for g(x,μ), CT14HERA2NNLO



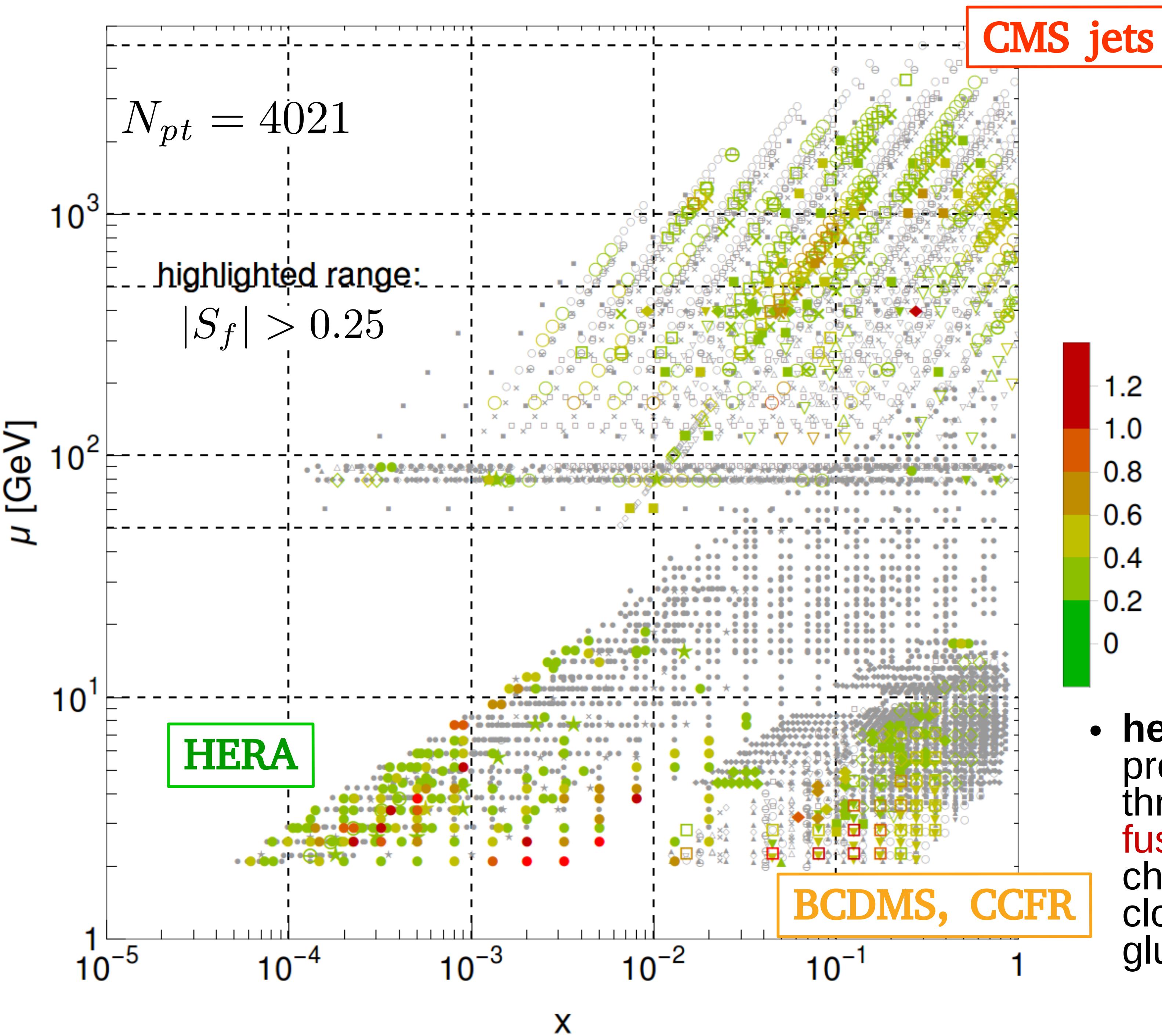
| S_f | for g(x,μ), CT14HERA2NNLO

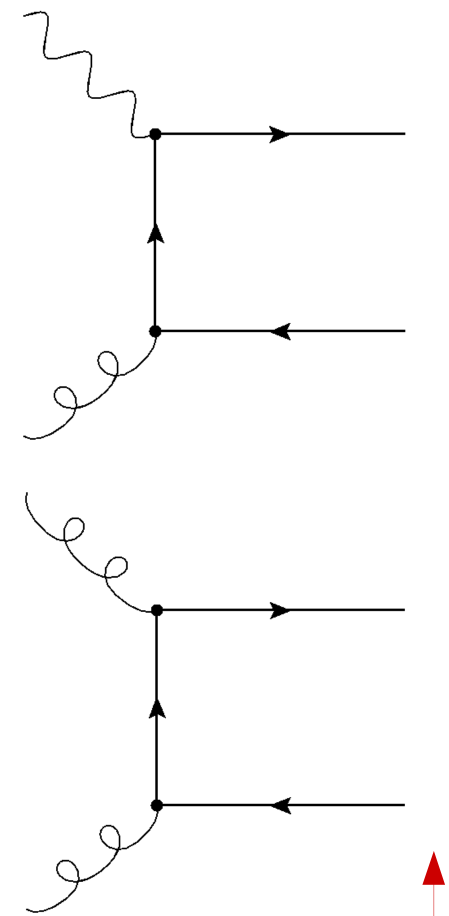


$\delta^{(PDF)} r_i$ relatively large at high and low x for the gluon distribution

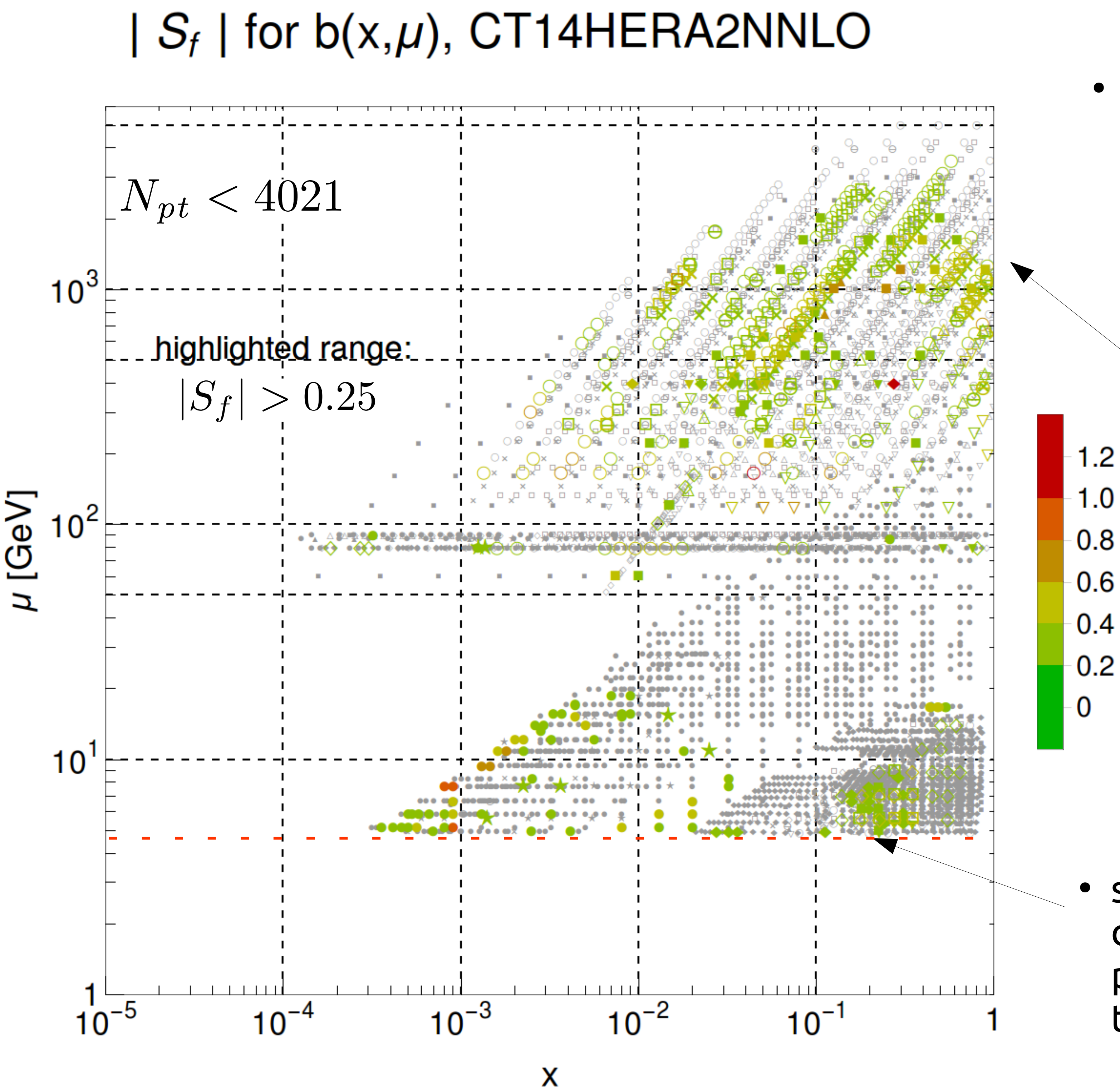
- measurements at high/low x are therefore most sensitive; esp. Run 2 **HERA** points and **CMS** jet production at 7, 8 TeV

$| S_f |$ for $c(x,\mu)$, CT14HERA2NNLO



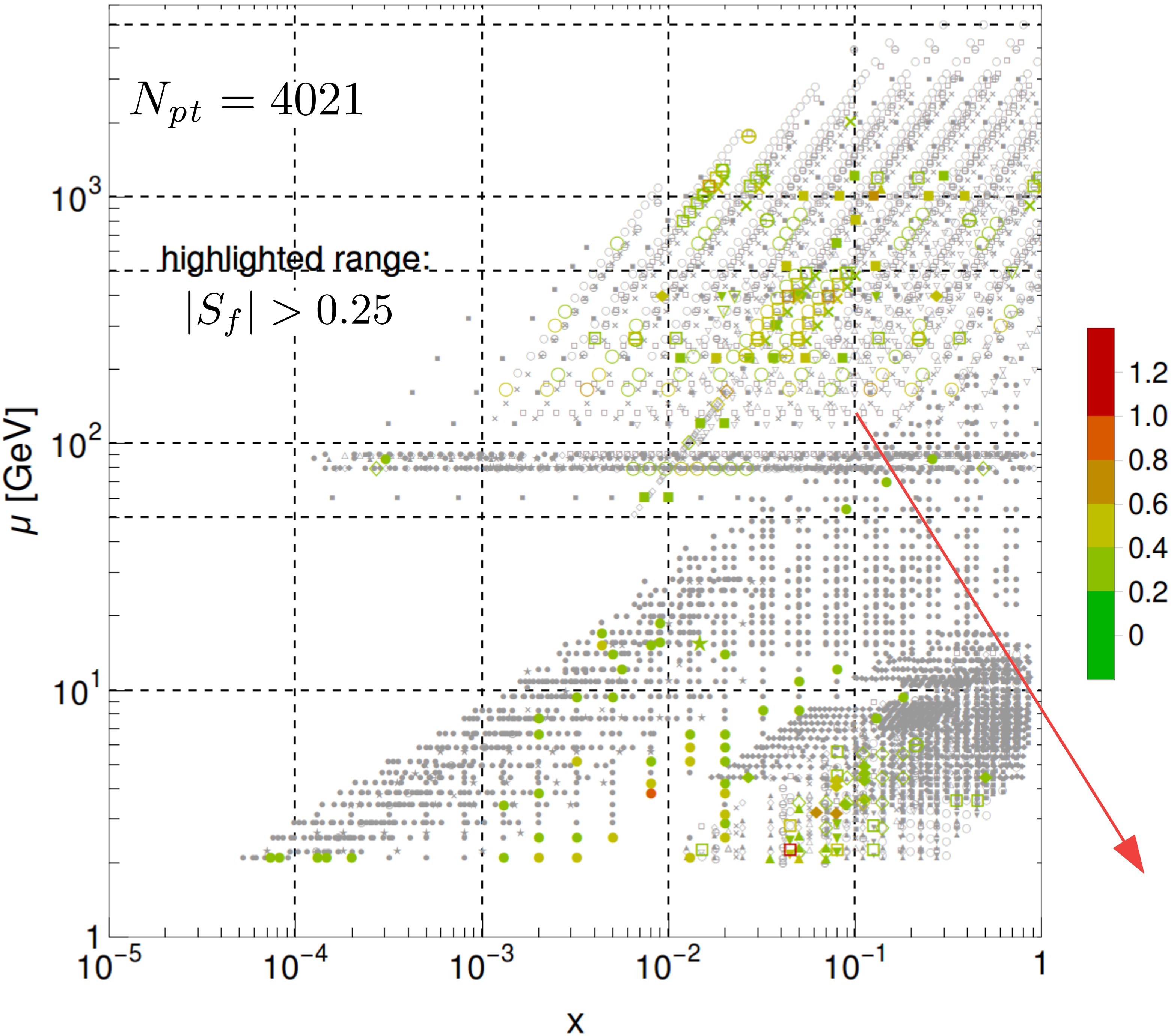


- **heavy quark** production proceeds through boson fusion diagrams: charm sensitivities closely track the gluon plot



- sensitivities broadly follow the pattern set by the charm quark distributions
- especially strong sensitivity at large x from inclusive jet production: **7 TeV and 8 TeV CMS points**
- sensitivities are only defined above the parton-level threshold, $\mu > m_b$

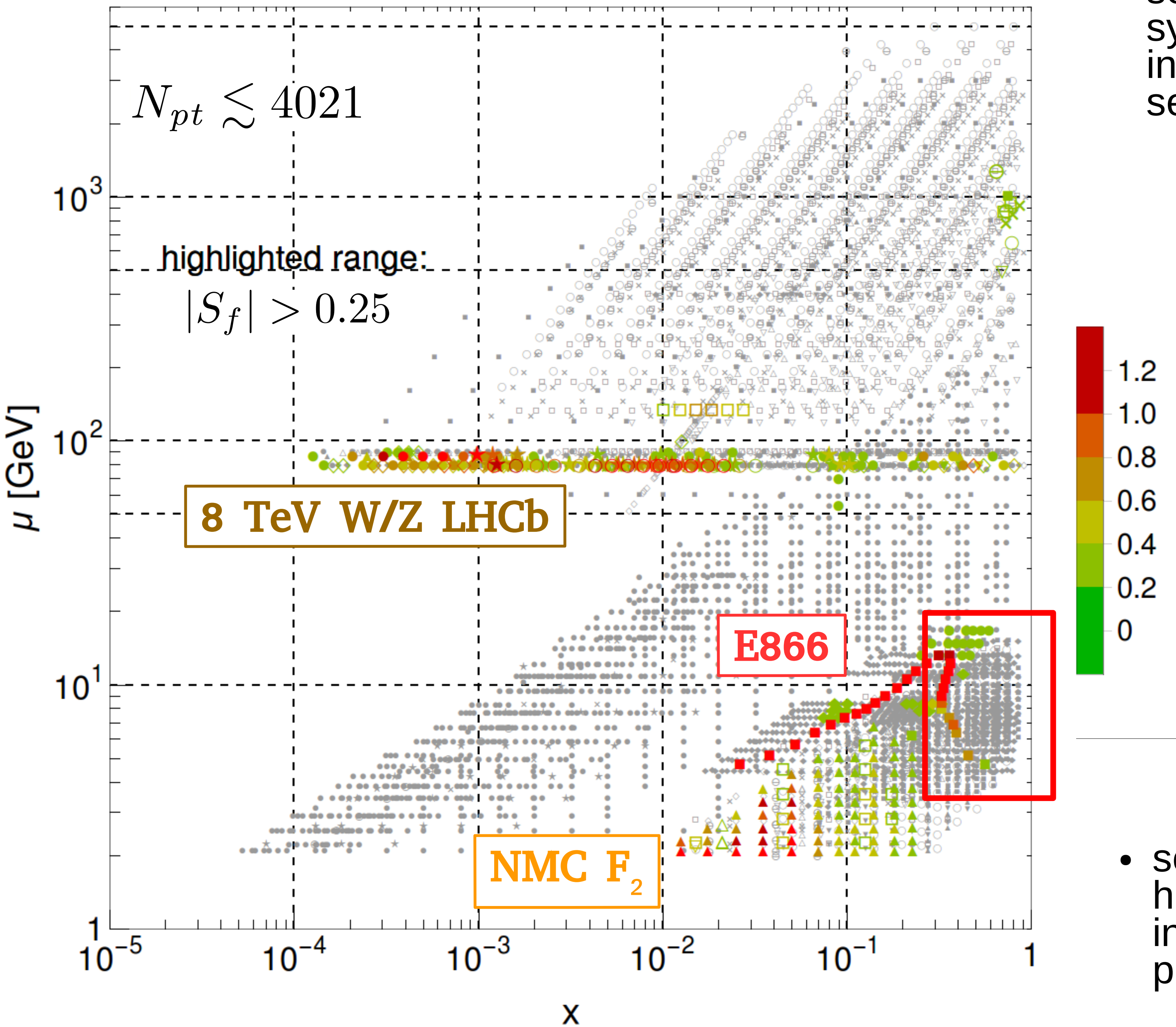
$|S_f|$ for σ_{H^0} 14 TeV, CT14HERA2NNLO



- several processes (high p_T Z prod., top prod.) have been suggested as providing leverage on the Higgs cross section

- in fact, we find **inclusive jet production** to have the broadest overall sensitivity!

| S_f | for $\bar{d}/\bar{u}(x,\mu)$, CT14HERA2NNLO



- PDF ratio is sensitive to flavor symmetry breaking in the light quark sea
- the large E866 sensitivity *degrades* at larger x

this is a prime motivation for higher x DY measurements at **E906** (SeaQuest)

- some contribution at high x from CMS inclusive jet production

Ranking tables

… to assess the impact of separate experiments

No.	Exp. ID	N_{pt}	$\sum_f \lvert S_f^E \rvert$		$\langle \sum_f \lvert S_f^E \rvert \rangle$	$\lvert S_{\bar d}^E \rvert$	$\langle \lvert S_{\bar d}^E \rvert \rangle$	$\lvert S_{\bar u}^E \rvert$	$\langle \lvert S_{\bar u}^E \rvert \rangle$	$\lvert S_g^E \rvert$	$\langle \lvert S_g^E \rvert \rangle$	$\lvert S_u^E \rvert$	$\langle \lvert S_u^E \rvert \rangle$	$\lvert S_d^E \rvert$	$\langle \lvert S_d^E \rvert \rangle$	$\lvert S_s^E \rvert$	$\langle \lvert S_s^E \rvert \rangle$
			Rankings														
1	160	1120.	620.	**HERA**				**A**	3	**A**	3	**A**	3	B		C	
2	**545**	185	232.	**CMS jets 8**			3	C	3	B	2					C	3
3	111	86	218.	CCF3 F3p			**1**	C	**1**		3	B	**1**	C	2		
4	**542**	158	194.	**CMS jets 7**			3	C	3	B	2					C	3
5	101	337	184.	BCDMS F2p				C		C		B	3	C			
6	104	123	169.	NMC			2					C	2	B	2		
7	102	250	141.	BCDMS F2d						C	3	C	3	C	3		
8	109	96	115.	CDHSW			2	C	2		3	C	2	C	3		
9	201	119	113.	E605			2	C	2				3				
41	**247**	8	5.84	Z pT 7 TeV			3		3		3		3		3		
42	169	9	3.99	HERA F_L							2						
43	**567**	7	3.9	$t\bar{t}$							2						
44	227	11	3.7	CDF WASY (2005)											3		
45	**568**	5	3.4								2						
46	**566**	5	3.19	$t\bar{t}$							2						
47	145	10	1.14	HERA b													

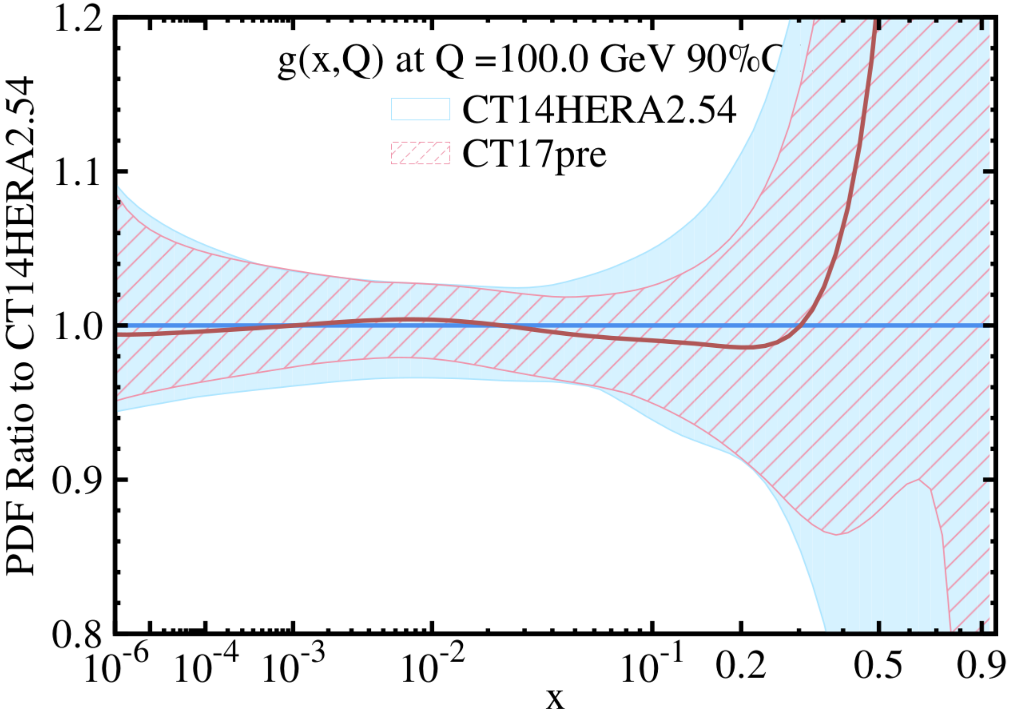
Experiments are listed in the descending order of the summed sensitivities to $\bar{d}, \bar{u}, g, u, d, s$

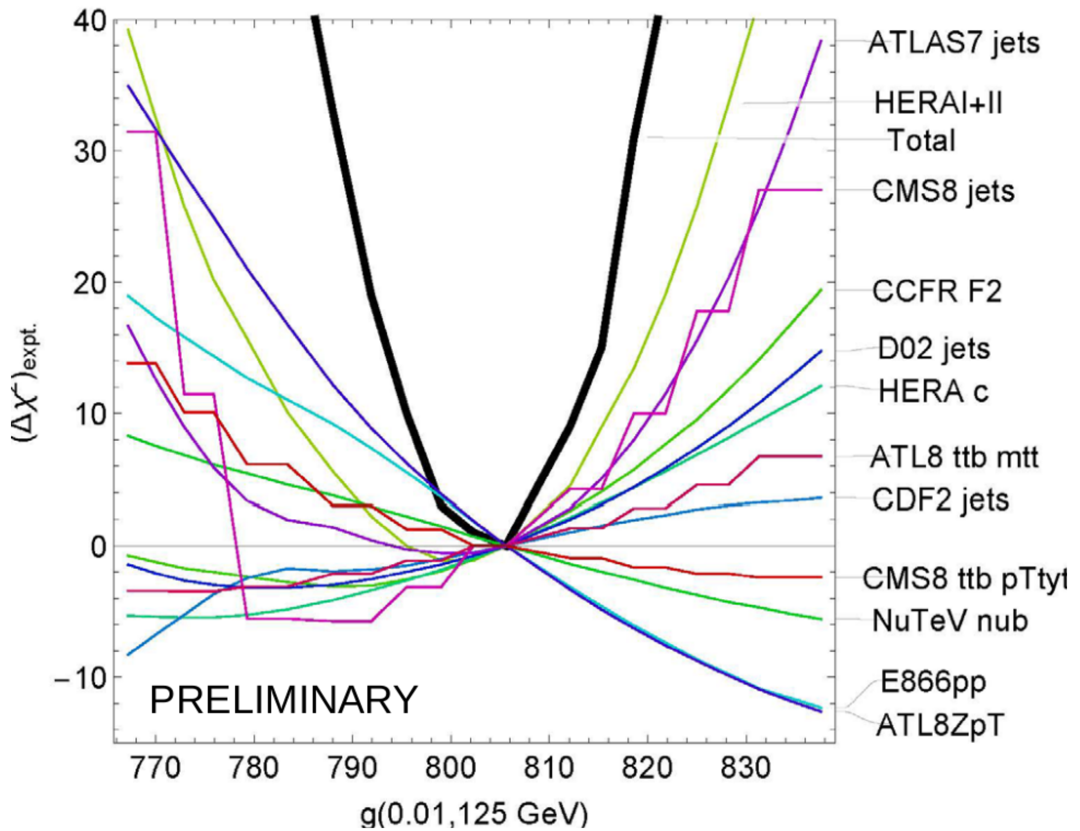
For each flavor, A and 1 indicate the strongest total sensitivity and strongest sensitivity per point

C and 3 indicate marginal sensitivities; low sensitivities are not shown

Good per-point S_f , small N_{pt}

PDFSense predictions can be validated against actual fits





- the incorporation of new LHC Run I data imposes modest constraints on the gluon PDF in the Higgs region
- Lagrange Multiplier scans provide an independent test of which datasets most drive the global fit in connection with specific PDFs

→ again, inclusive jet production data are dominant!

- what can we conclude from the application of this tool to the expanded 'CTEQ-TEA' dataset?

→ the importance of inclusive jet production – especially at CMS

→ the strength of the constraints provided by data depend importantly on the detailed treatment of systematic (correlated errors)

…e.g., 'de-correlating' error sources like jet energy scales in the CMS data can modestly diminish their expected impact

→ this is similarly true of theoretical errors (e.g., Monte Carlo uncertainties)

- **now bench-marking upcoming CT fits against these predictions**

(as well as complementary PDF reweighting methods – e.g., ePump)

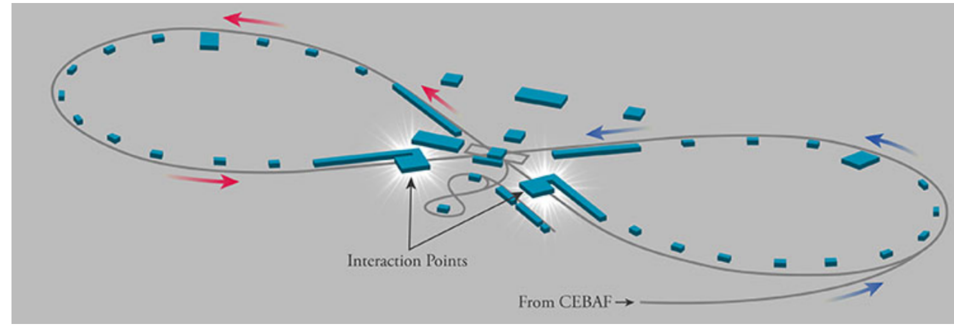
Schmidt, Pumplin and Yuan; arXiv:1806.07950.

iv) ...but what about *future experiments*

...<u>the EIC</u>?

(a representative sample of typical results...)

may apply these techniques to **nucleon PDFs** at **JL-EIC** !



- crucial point: proposals specific to eRHIC, JL-EIC, etc., differ in various respects – these differences lead to nontrivial consequences for the resulting PDF sensitivities; for now, take JL-EIC as an archetypal EIC scenario

- **as a characteristic example, we consider the highest energy [unpolarized] pseudodata:**

$$\big(10\,\mathrm{GeV}\;\, e^{-}\;\mathrm{on}\;100\,\mathrm{GeV}\;p\big)$$

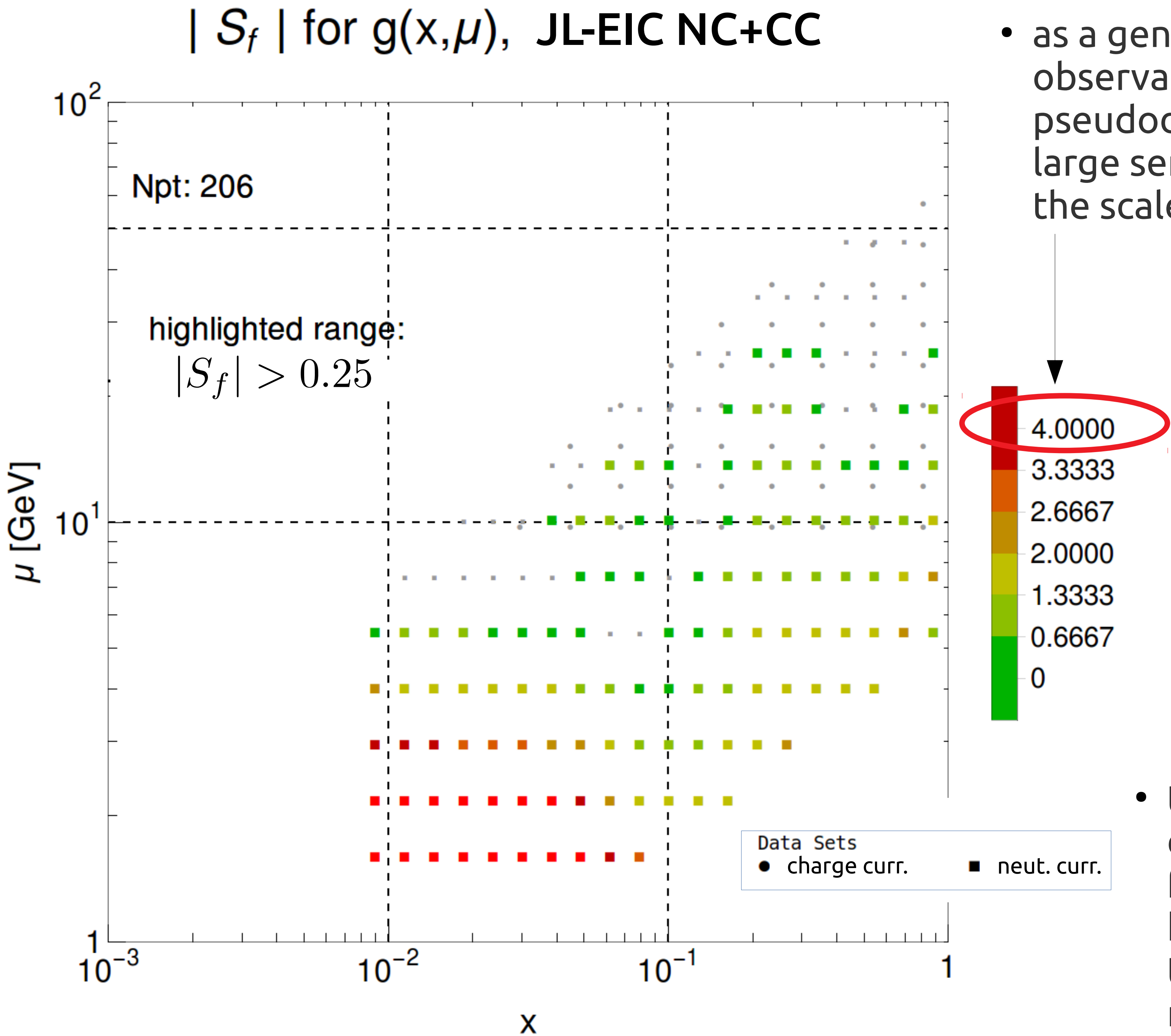
$$\ldots e^{+}p \;\text{pseudodata under analysis}\ldots$$

- NB: events were generated for an assumed integrated luminosity $\mathcal{L} = 100\,\mathrm{fb}^{-1}$

…thanks to Alberto Accardi, Yulia Furletova for generating the pseudodata

- require binning scheme; need MC-generated statistics, putative systematics

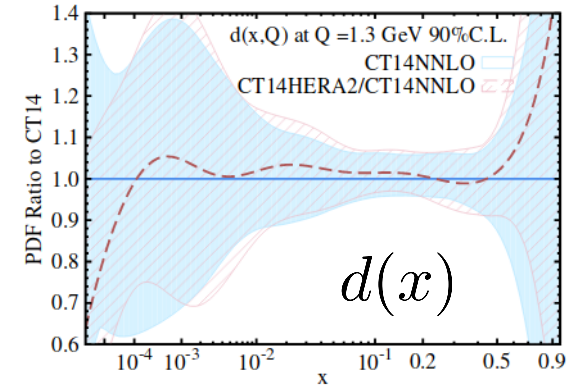
→ can then fluctuate about CT14; compute residuals; PDF correlations; sensitivities



- as a general observation, the JL-EIC pseudodata enjoy very large sensitivities (note the scale below!)
- cf. the sensitivities typical of the most important LHC Run I information: (S_f ~ **0.6 – 1.0**)
- these precision data straddle the few GeV QCD-hadron transition region!

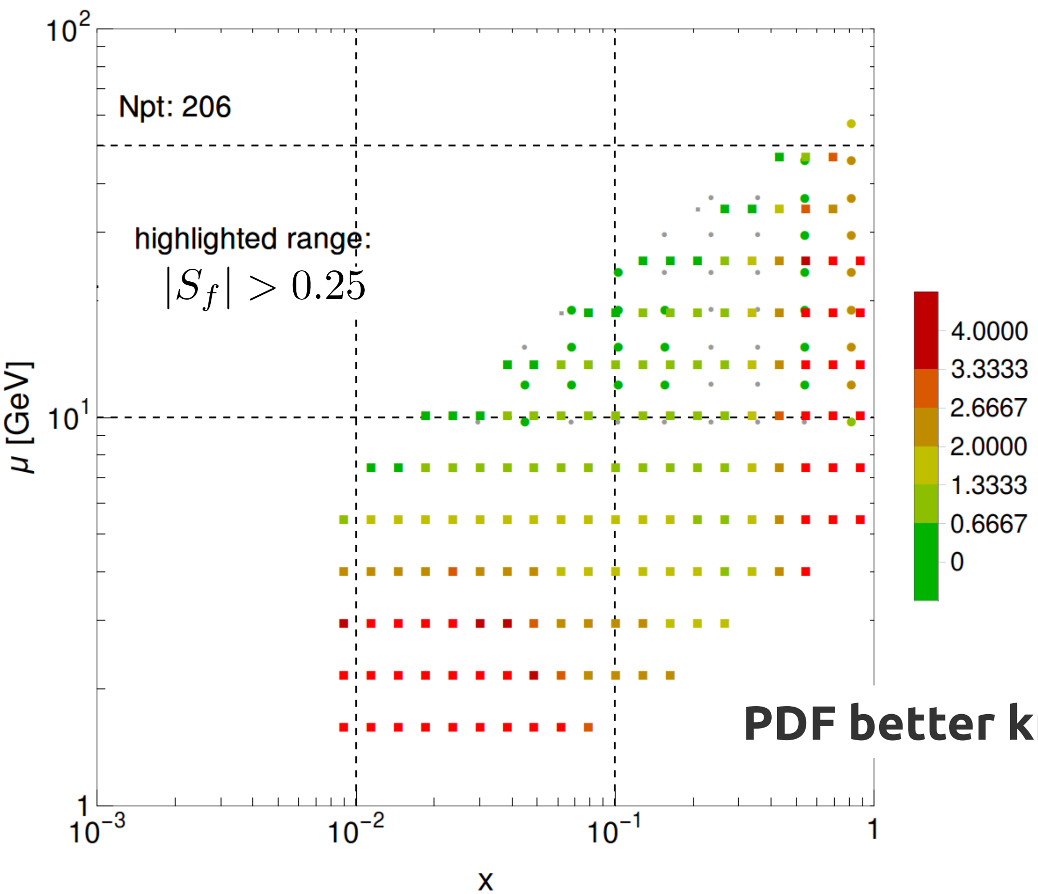
- JL-EIC's high sensitivity extends to the SU(2) light quark sector, and should yield highly constraining information over the coverage region for

$$u(x,\mu),\ d(x,\mu)\ \rightarrow\ d/u(x,\mu)$$

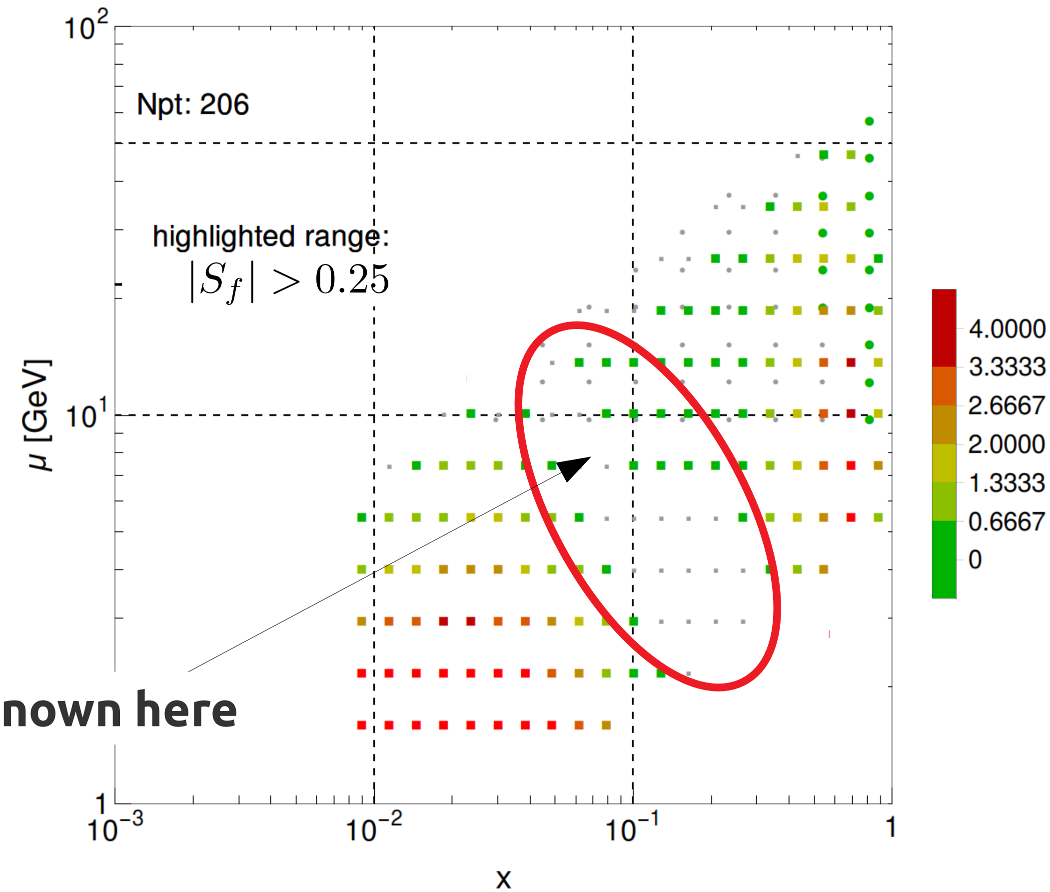


- NC cross sections yield tightest constraints at lowest and highest x; significant CC sensitivity to u-quark at high x

| S_f | for u(x,μ), **JL-EIC NC+CC**

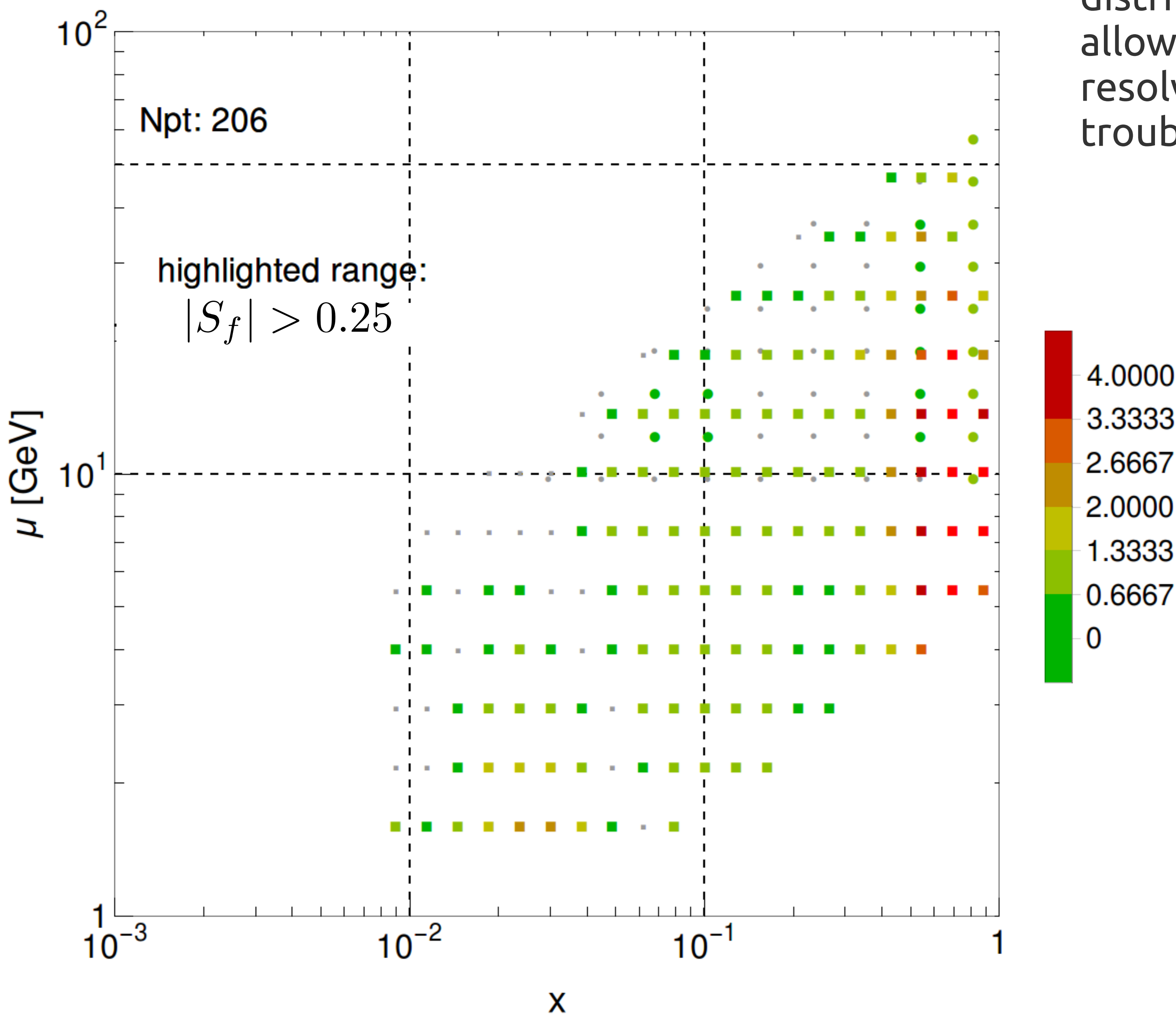


| S_f | for d(x,μ), **JL-EIC NC+CC**



PDF better known here

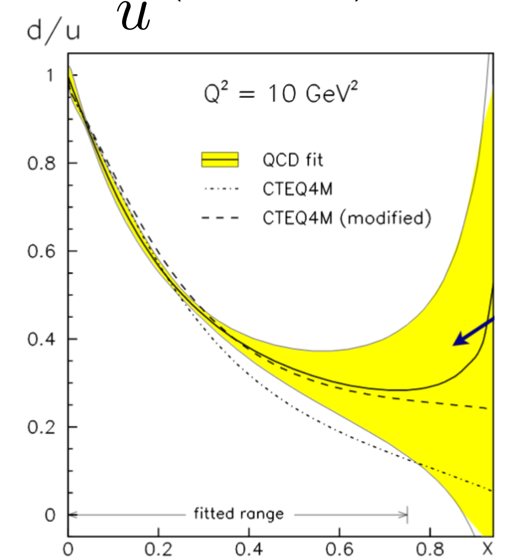
$|S_f|$ for d/u(x,mu), **JL-EIC NC+CC**



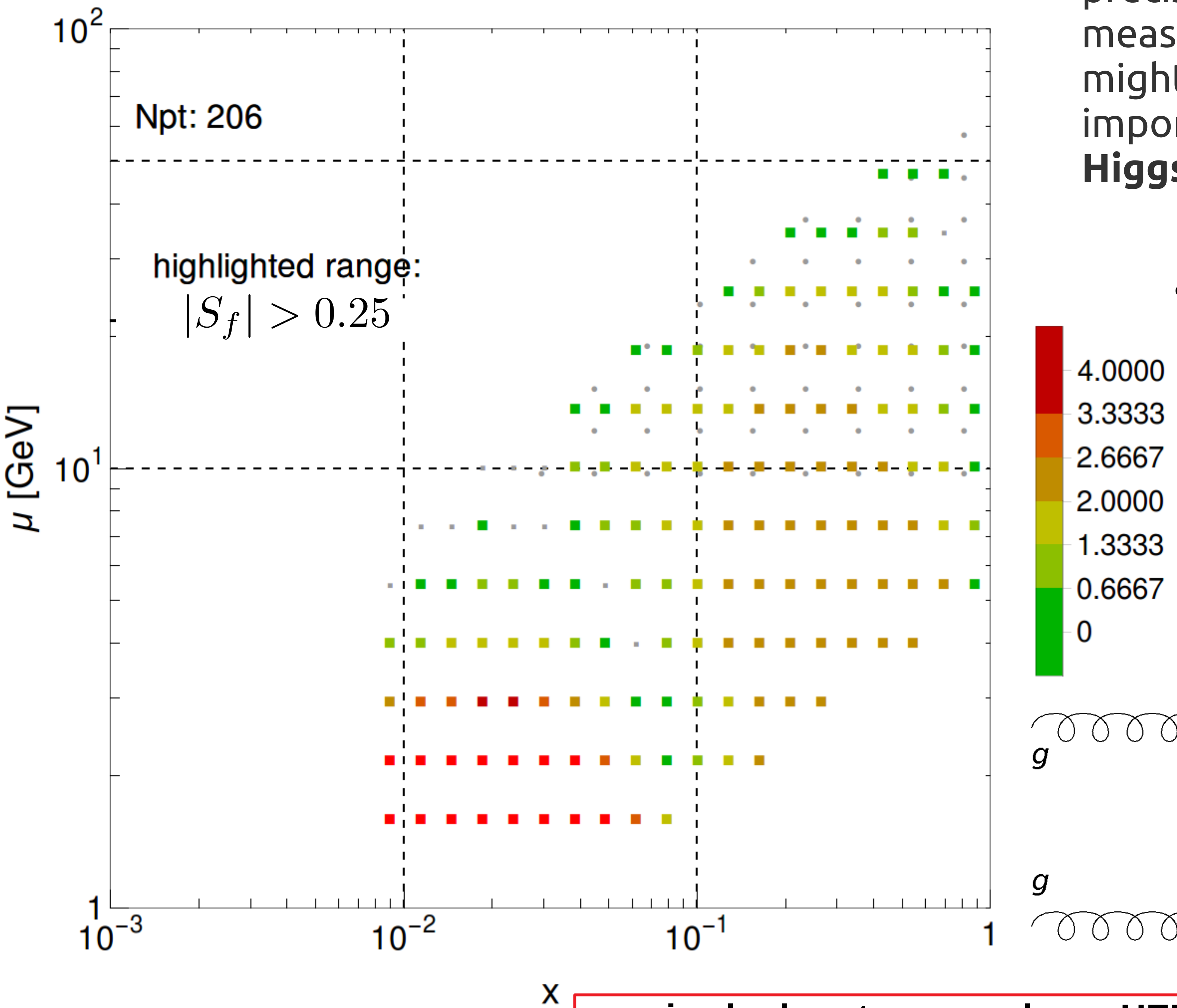
- sensitivity to light quark distributions should allow EIC to finally resolve the troublesome d/u ratio

  - will enlighten and benefit from EIC's advanced tomography (e.g., what are the diquark correlations in the nucleon?)

$$\frac{d}{u}(x \to 1) =?$$

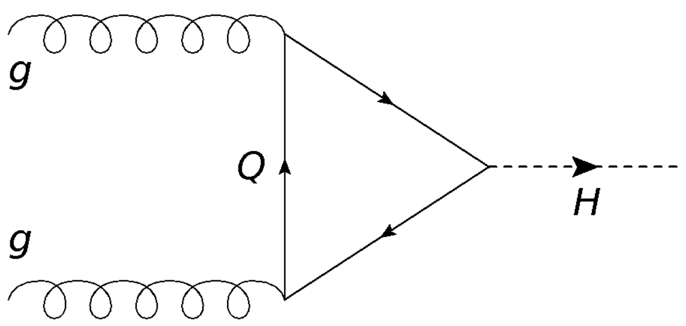


$|S_f|$ for σ_{H^0} 14 TeV, **JL-EIC NC+CC**

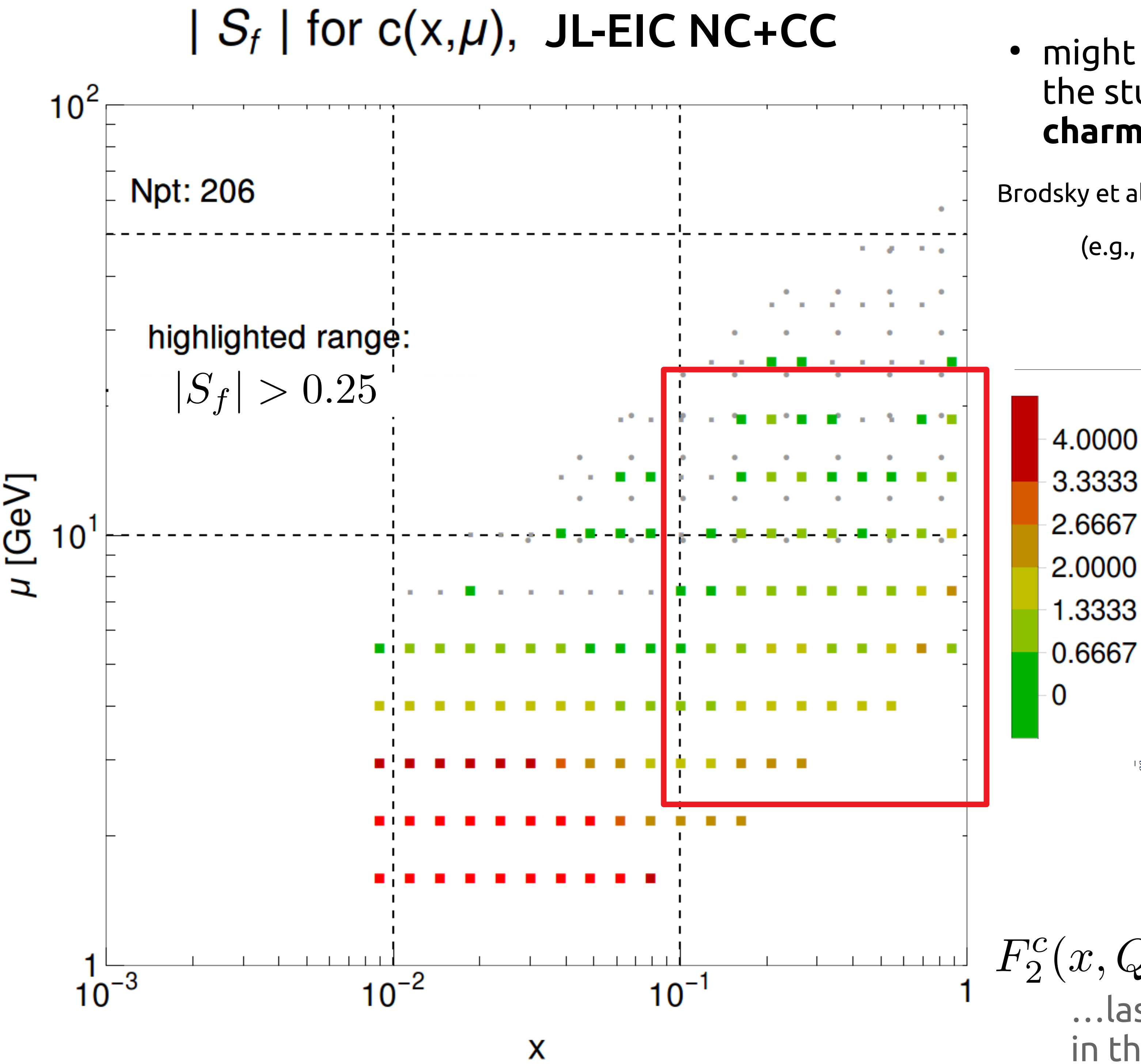


- precise NC measurements at JL-EIC might even be an important input to LHC **Higgs phenomenology**!

- while this might seem remarkable, it is a direct consequence of the dominant role gluon fusion plays in Higgs production:



precise hadron tomography ↔ HEP discovery potential



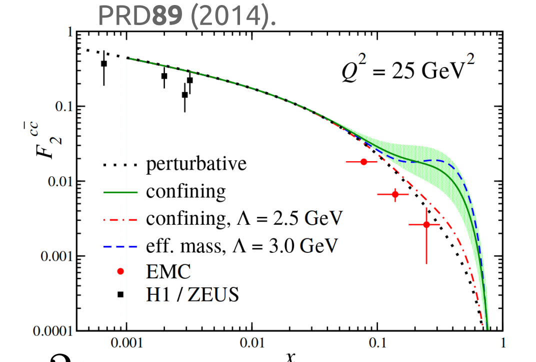
- might an EIC *finally* resolve the stubborn **intrinsic charm** issue?

Brodsky et al., Phys. Lett. **B**93, 451 (1980).

(e.g., see talk by Elke Aschenauer)

- strong sensitivity to the high x, small Q region favored by models of valence-like intrinsic charm:

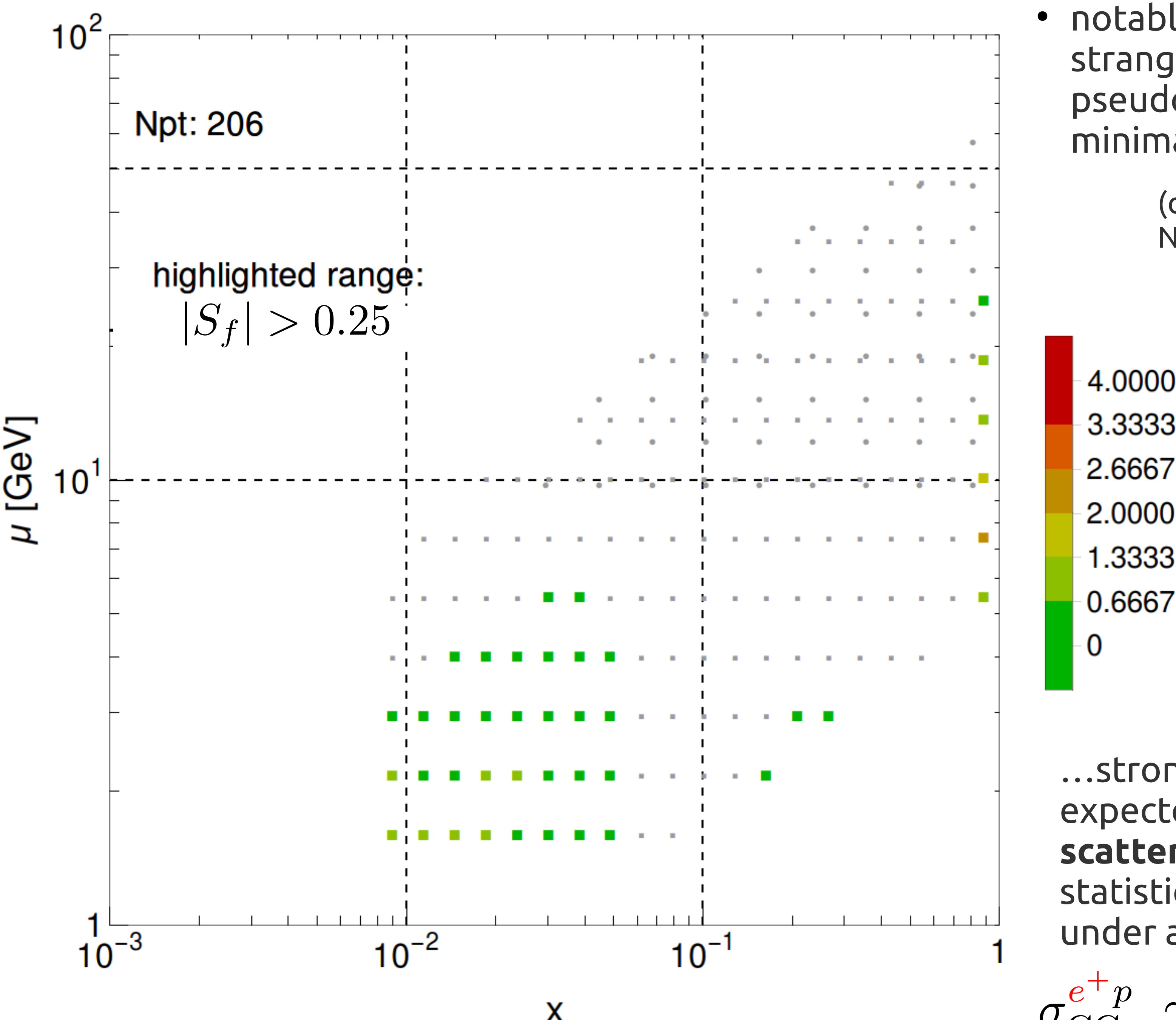
TJH, Londergan, Melnitchouk; PRD**89** (2014).



$F_2^c(x, Q^2)$

…last measured by EMC in the early 1980s!

$|S_f|$ for $s(x,\mu)$, **JL-EIC NC+CC**



- notably, the sensitivity to strangeness of these pseudodata is relatively minimal

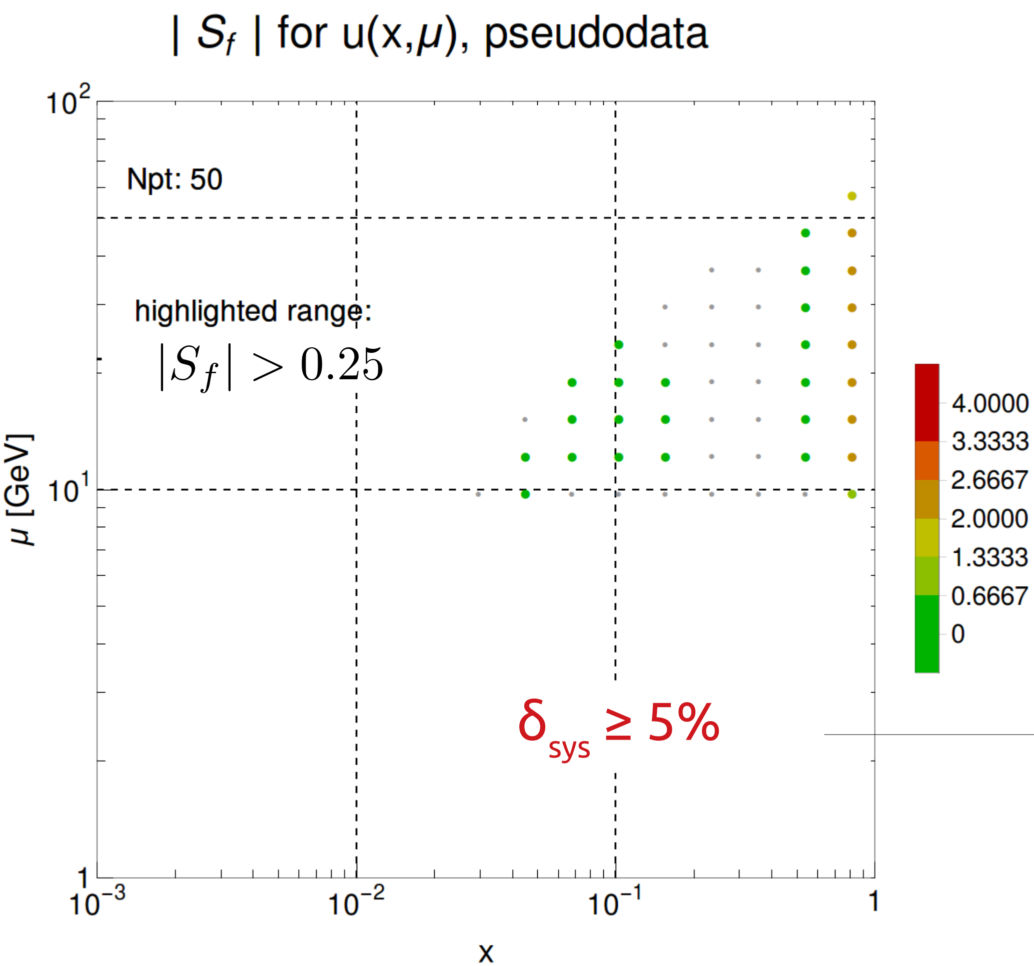
(despite some significant NC contributions at low x)

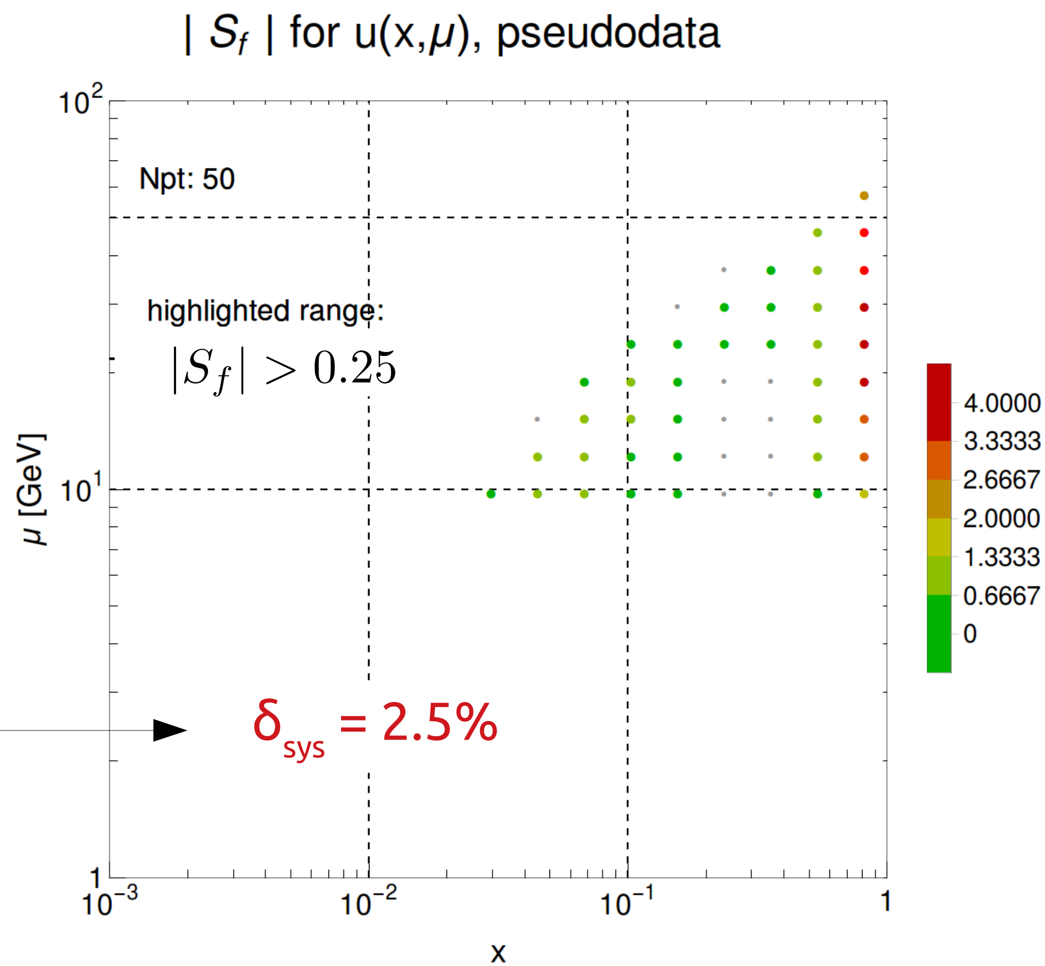
- then again, this is not shocking: flavor separation expected to come from CC data, but these pseudodata are **electron-proton**

…stronger constraints are expected for **positron scattering**! (although statistics tend to be lower – under analysis now…)

$$\sigma^{e^+p}_{\mathrm{CC,r}} \simeq [1-y]^2\, x(d+s)$$

- N.B.: with $\delta_{sys} \geq 5\%$, sensitivities in the **charge-current channel** generally fall below our significance threshold





$\delta_{sys} \geq 5\%$ → $\delta_{sys} = 2.5\%$

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

$$s_i^2 = \delta_{stat}^2 + \delta_{sys,uncor.}^2$$

- while improvements in systematic uncertainties can modestly sharpen the sensitivity (above), statistics limitations guarantee these enhancements *in isolation* yield asymptotically diminishing returns

→ being explored now…

v) in progress: weighing the impact of EIC scenarios against other proposed experiments

(to maximize the physics impact of the future EIC program)

Phys. Rept. 522 (2013) 239-255.

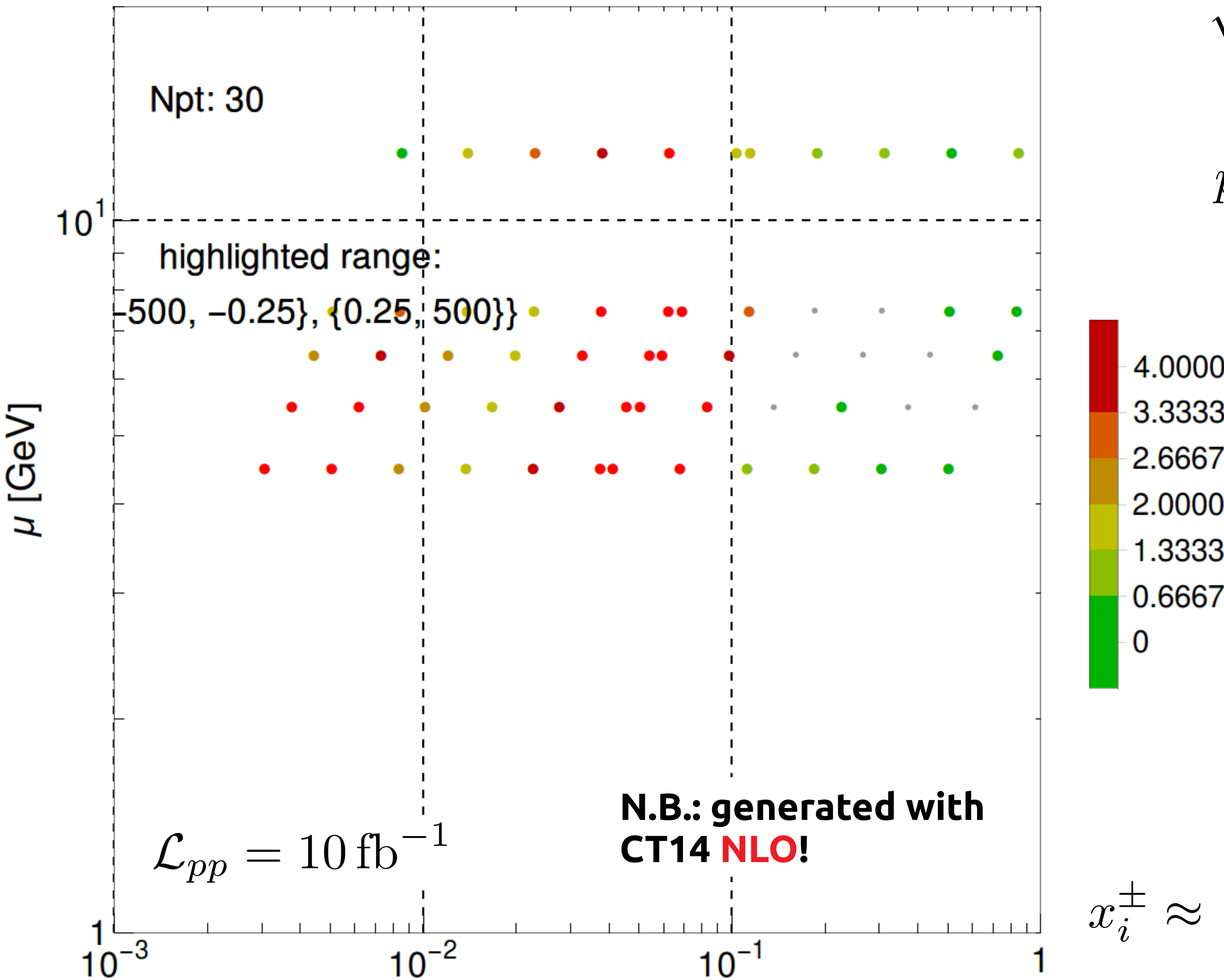
AFTER | S_f | for $g(x,\mu)$, AFTER@CERN

$\frac{d\sigma_{DY}}{dy^{c.m.}_{\mu\mu}}$ in bins of definite $M_{\mu\mu}$

$\sqrt{s} = 115\,\mathrm{GeV}$

$\mu = \mu_F = M_{\mu\mu}$

$p_T > 1.2\,\mathrm{GeV}$



$\mathcal{L}_{pp} = 10\,\mathrm{fb}^{-1}$

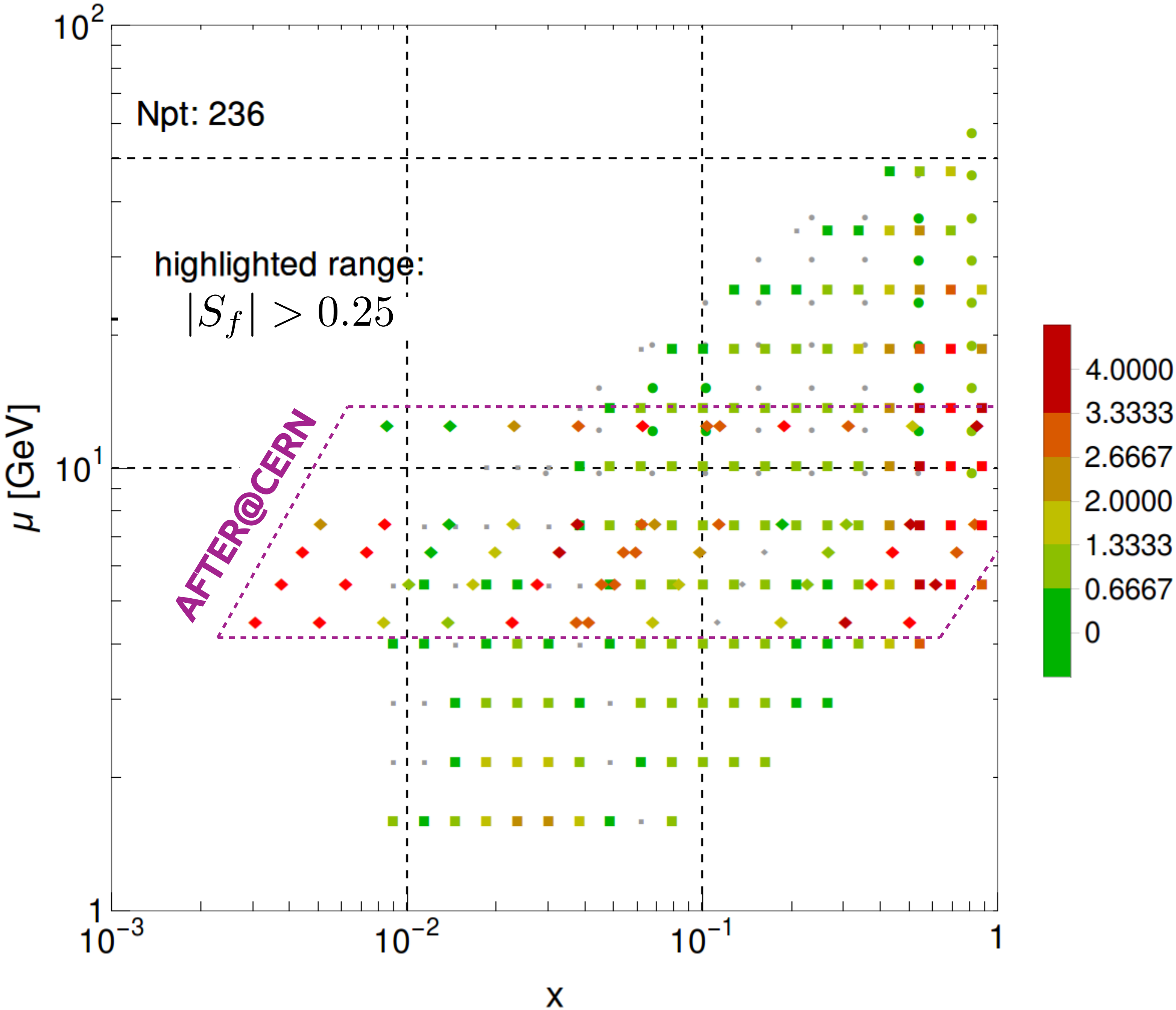
N.B.: generated with CT14 NLO!

- **density of liquid H fixed target → high statistics**
- aimed at high x extractions of heavy quark and quarkonium prod.; pp, pd, pA collisions; spin physics, etc., etc. --- important intersection with EIC motivation!

$$x_i^{\pm} \approx \frac{M_{\mu\mu}}{\sqrt{s}} \exp(\pm y_{\mu\mu})\Big|_i$$

thanks to Olek Kusina, Ingo Schienbein for fluctuated pseudodata

| S_f | for d/u(x,mu), **JL-EIC NC+CC; AFTER**



- while there is substantial overlap in kinematical coverage, there is significant **complementarity** between sensitive regions probed by AFTER@CERN and JL-EIC-like proposals

- e.g., JL-EIC shows strong sensitivity to d/u at high x over wide range of fact. scales; AFTER@CERN probes low x ratio to high precision and can constrain the evolved high-x distribution

another comparison: 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 $x \simeq 10^{-6}$

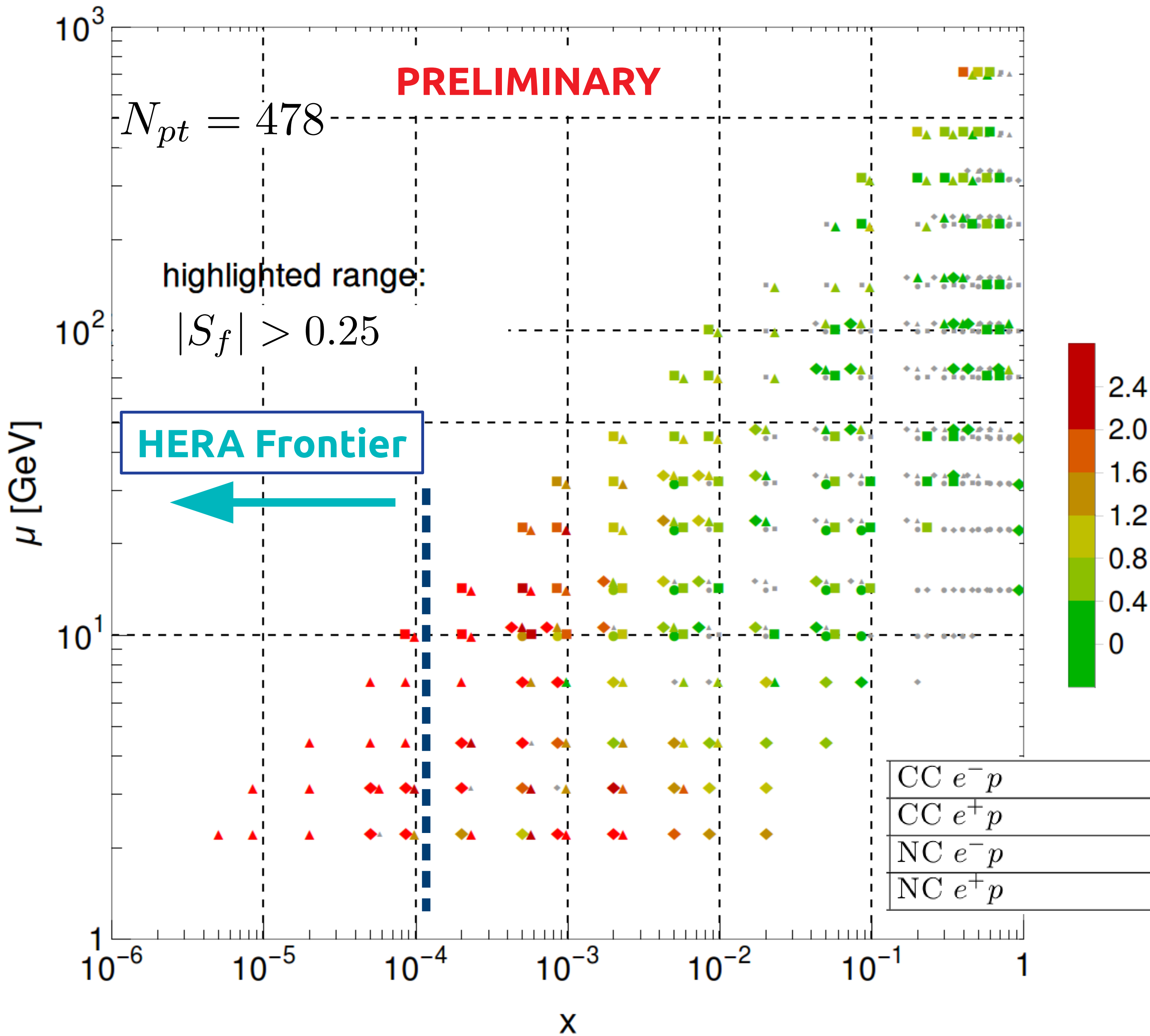
  → explore the dynamics of gluon saturation; greatly improve PDF precision; perform SM tests; and many other physics goals

- 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 CT14 prediction according to putative LHeC uncorrelated errors – based on **10 fb^{-1}** of data from a hypothetical ~year of data-taking

| S_f | for g(x,μ), CT14H2



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

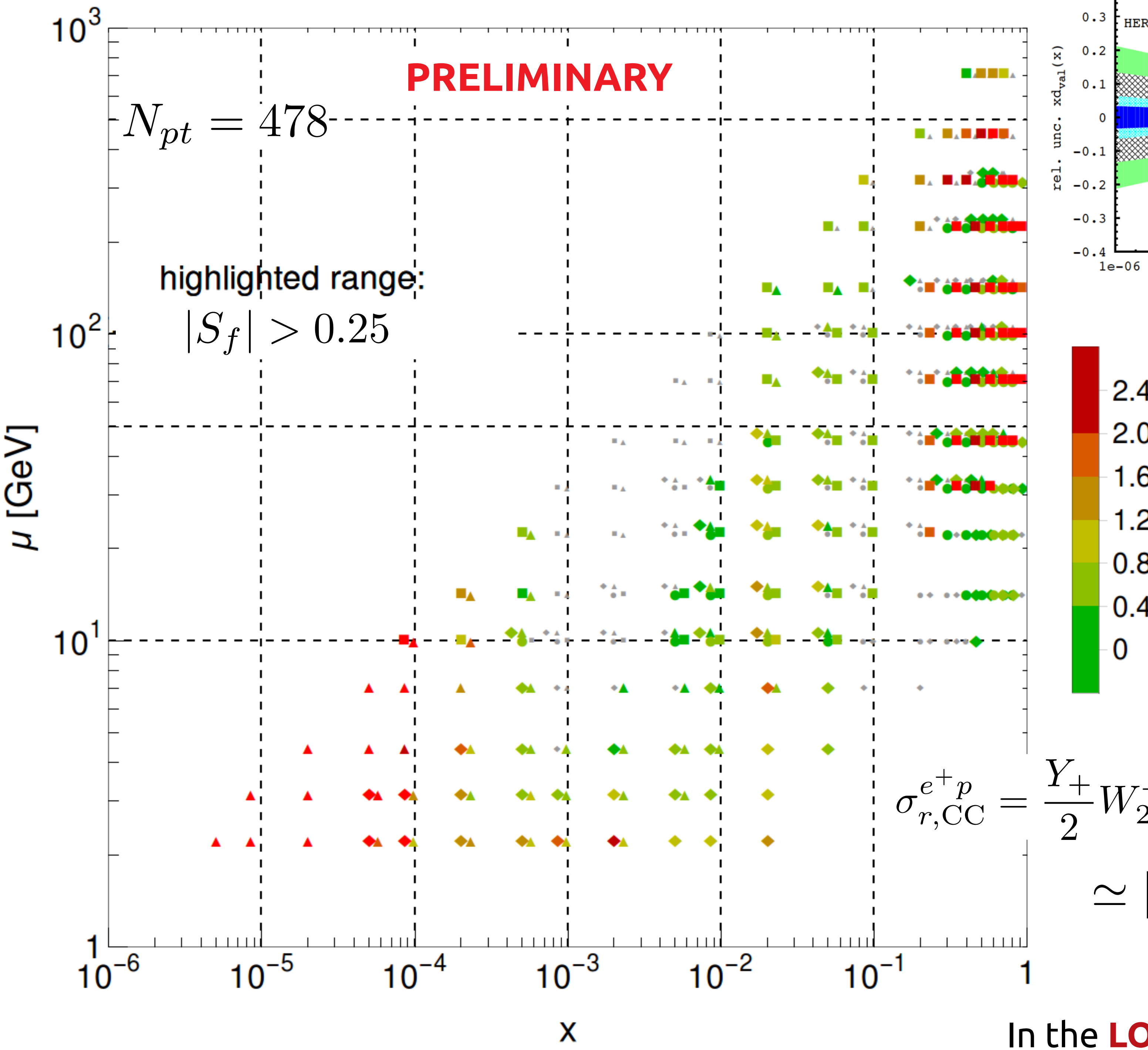
  however, there are stringent constraints in general for

  $x \lesssim 10^{-2}$

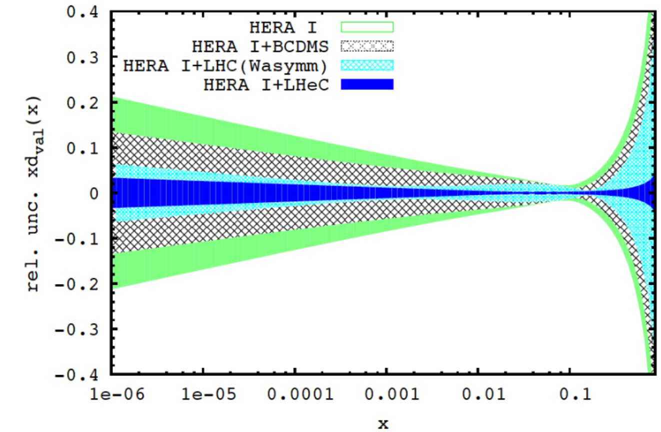
  ...and high $x, \ \mu$

	N_d	
CC e^-p	93	circle
CC e^+p	109	square
NC e^-p	128	diamond
NC e^+p	148	triangle

$| S_f |$ for d(x,μ), CT14H2



EIC Whitepaper:1206.2913



- The **strongest sensitivities** are realized for those regions with the **weakest current PDF constraints**; here, very high and very low x

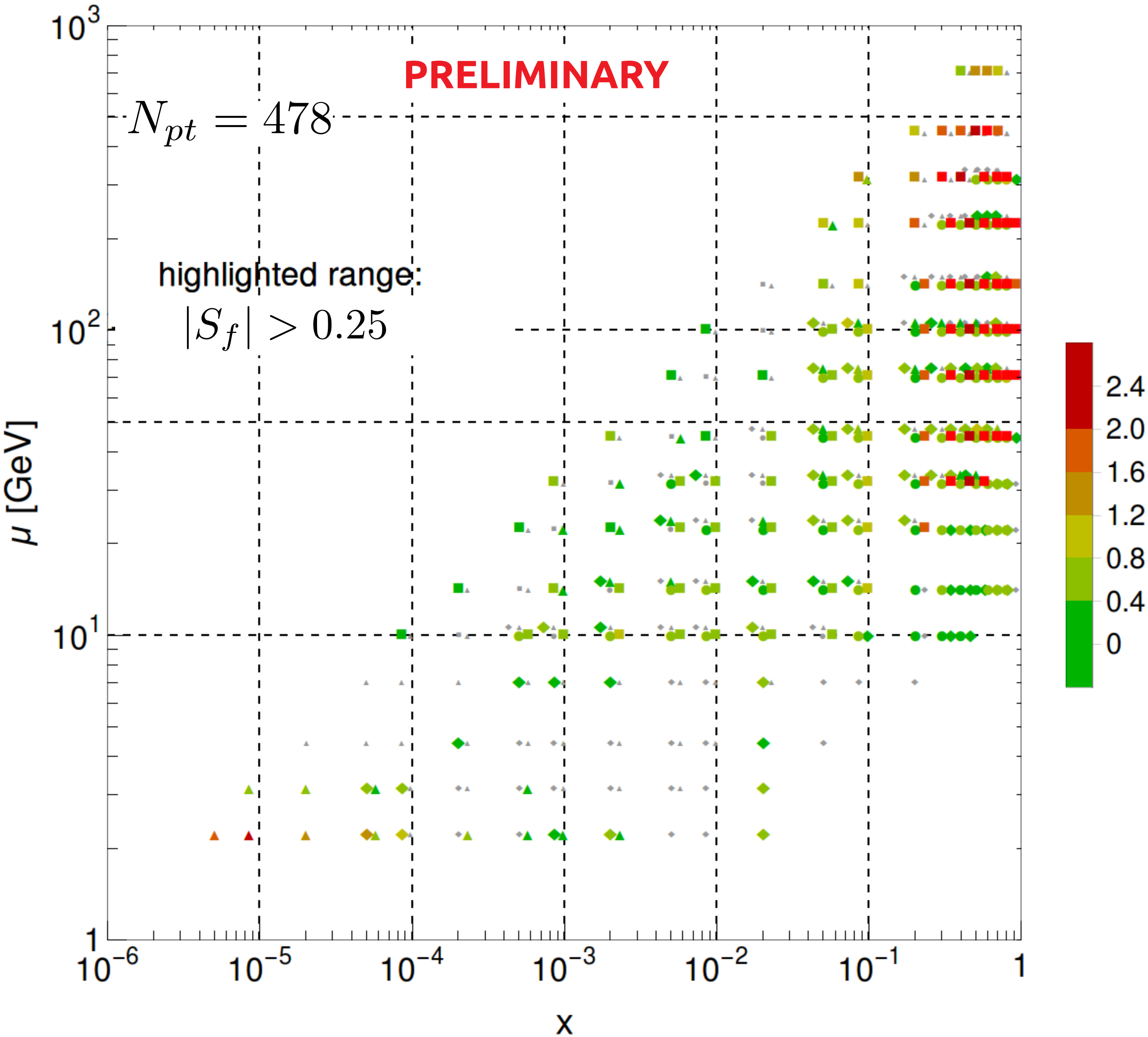
for $x \to 1$

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

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

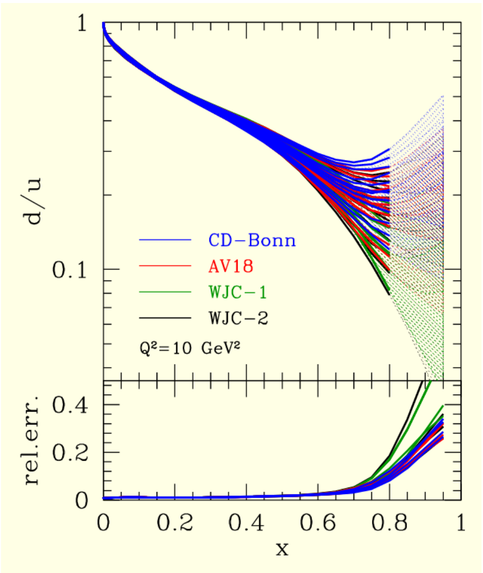
In the **LO quark-parton model**

$|S_f|$ for d/u(x,mu), CT14H2

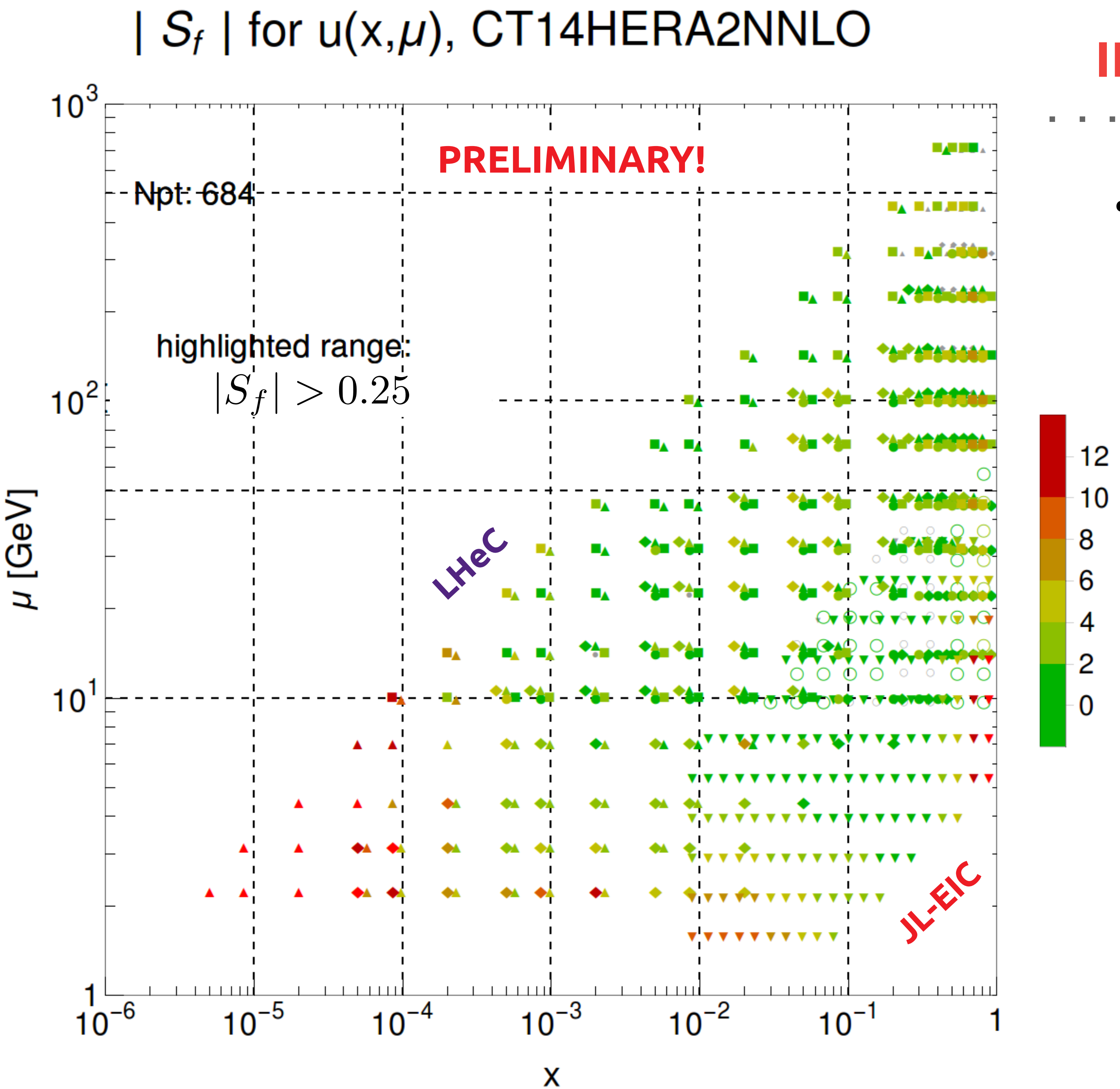


- LHeC's high luminosity may give it a reach to high enough x to help resolve the stubborn d/u question

…without a nuclear target…



J.F. Owens, PAVI11



IN PROGRESS

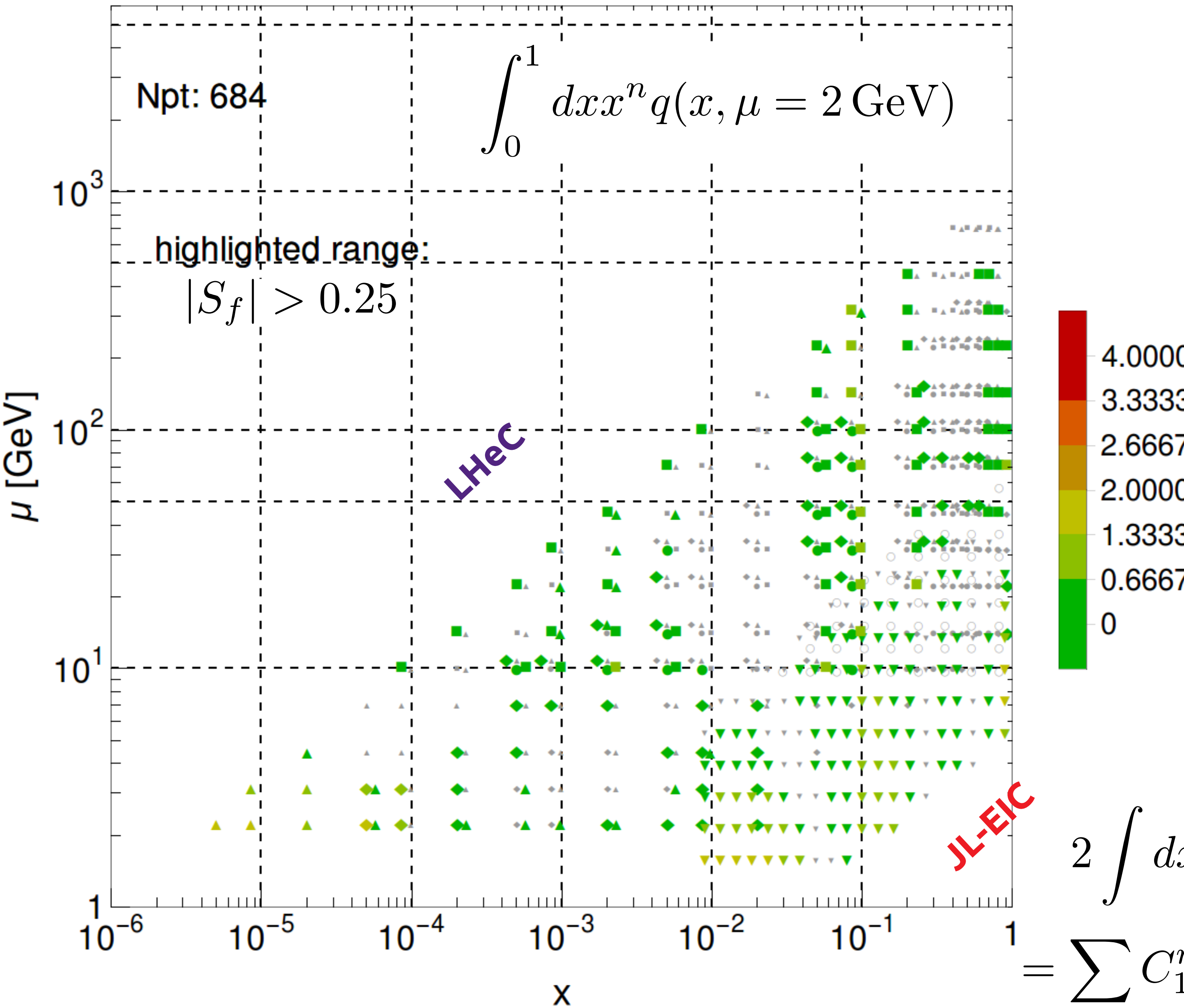
- **eventual goal:** direct, 'apples-to-apples' comparisons of the kinematical phase space coverage and sensitivity of proposed experimental programs – e.g., JL-EIC vs. LHeC, etc.

- in this illustration, we compare a presumed 10-year run of JL-EIC (in one bin of beam energy!) to **approximately** statistics-rescaled LHeC pseudo-data

vi) epilogue: how might this picture relate to efforts to compute hadron structure on the lattice?

$| S_f |$ for $\langle x^0 \rangle (u_+ - d_+)$, CT14HERA2NNLO

LATTICE PDF **Community Whitepaper:**
Prog. Part. Nucl. Phys. 100 (2018) 107-160.

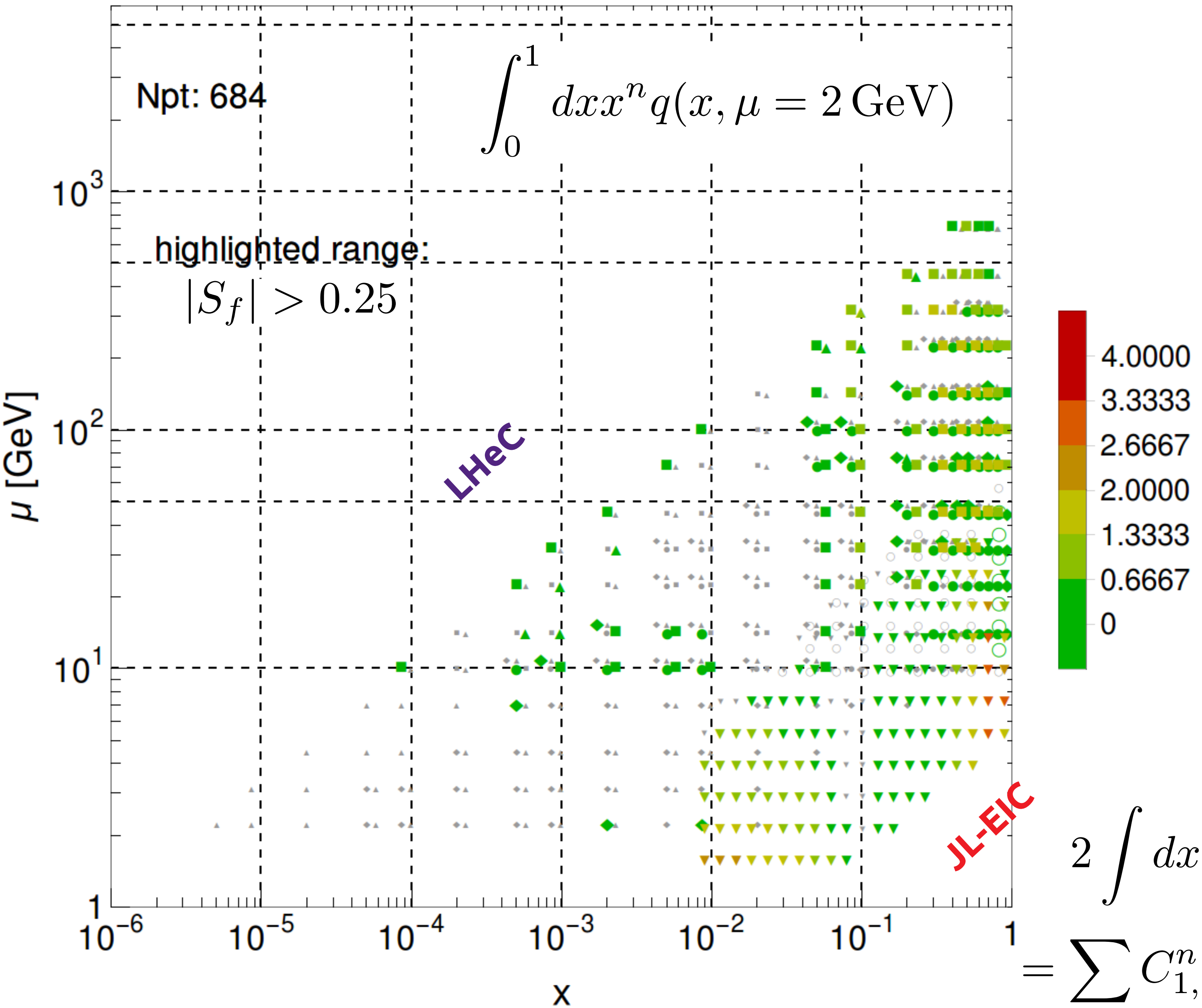


- might EIC constrain knowledge of lattice-calculable PDF Mellin moments?
- JL-EIC possesses successively larger sensitivities at high x as the order of the Mellin moments increases!
- **there will be a future synergy between LQCD and HEP phenom.**

$$2 \int dx x^{n-1} F_1(x, Q^2) = \sum_a C^n_{1,a}(\mu^2) v^n_a(\mu^2) \Big|_{\mu^2 = Q^2}$$

$|S_f|$ for $\langle x^1 \rangle(u_+ - d_+)$, CT14HERA2NNLO

LATTICE PDF **Community Whitepaper:**
Prog. Part. Nucl. Phys. 100 (2018) 107-160.

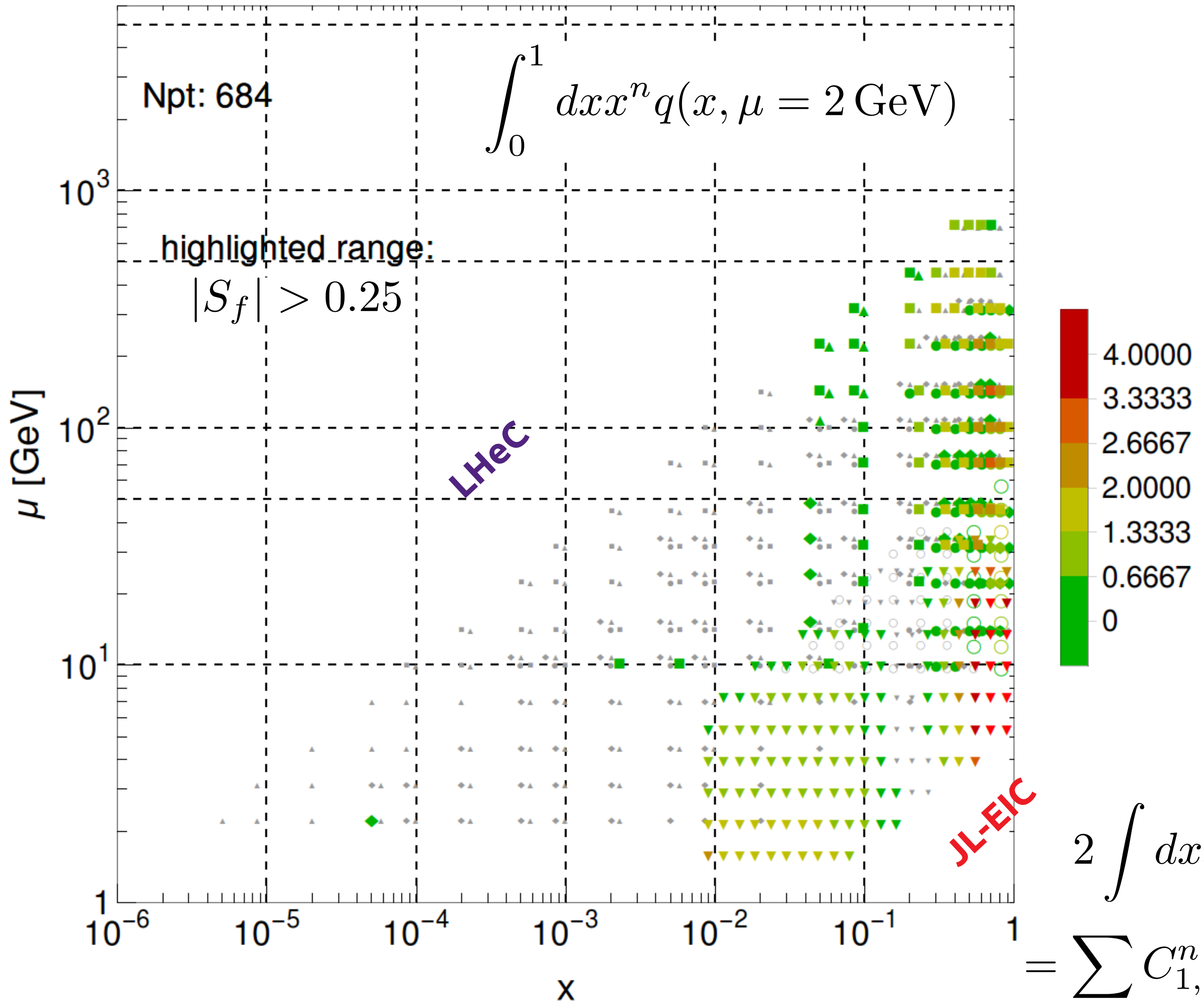


- might EIC constrain knowledge of lattice-calculable PDF Mellin moments?
- JL-EIC possesses successively larger sensitivities at high x as the order of the Mellin moments increases!
- **there will be a future synergy between LQCD and HEP phenom.**

$$2 \int dx x^{n-1} F_1(x, Q^2) = \sum_a C^n_{1,a}(\mu^2) v^n_a(\mu^2)\big|_{\mu^2 = Q^2}$$

$| S_f |$ for $\langle x^2 \rangle (u_+ - d_+)$, CT14HERA2NNLO

LATTICE PDF **Community Whitepaper:**
Prog. Part. Nucl. Phys. 100 (2018) 107-160.



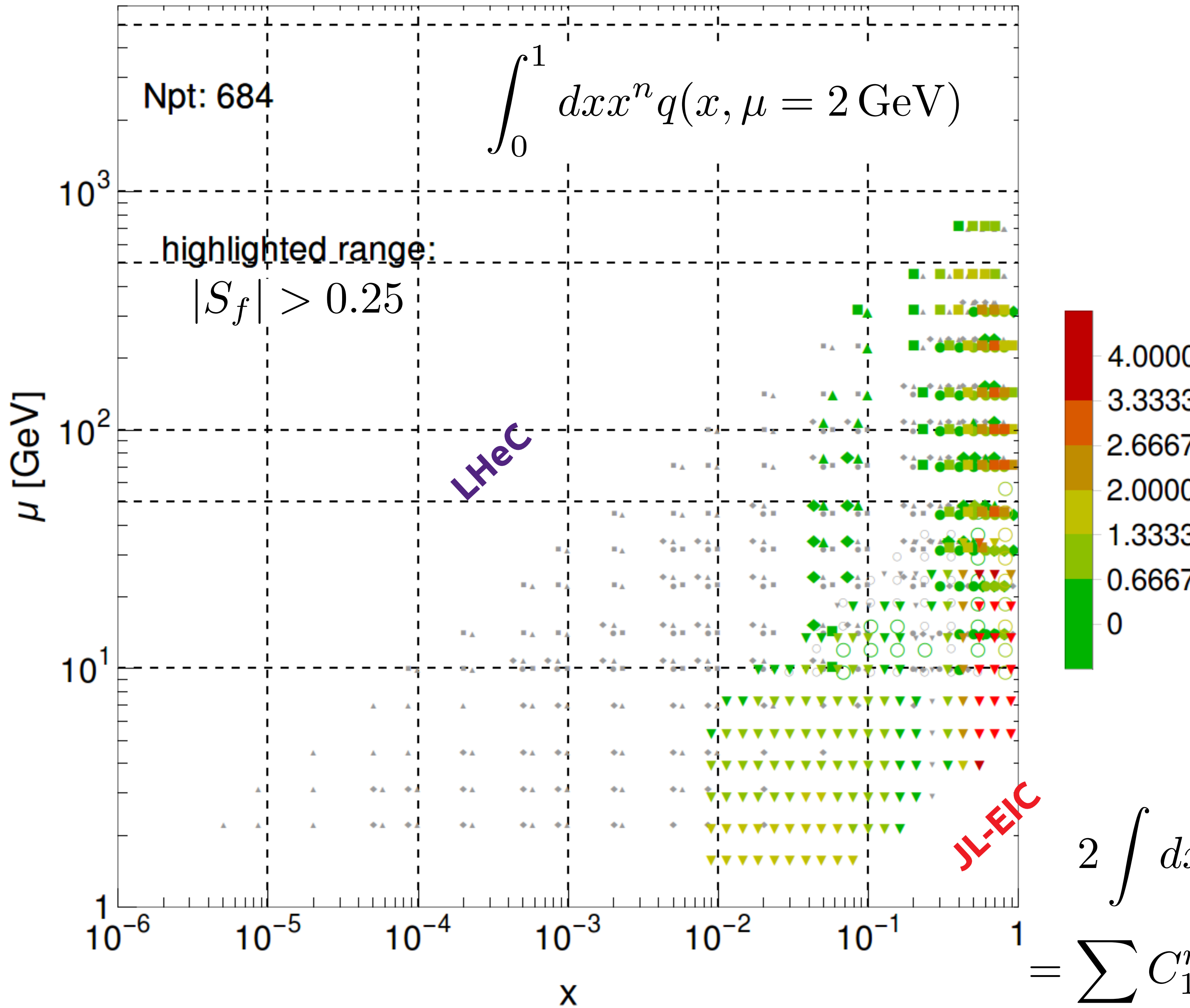
- might EIC constrain knowledge of lattice-calculable PDF Mellin moments?
- JL-EIC possesses successively larger sensitivities at high x as the order of the Mellin moments increases!
- **there will be a future synergy between LQCD and HEP phenom.**

$$2 \int dx x^{n-1} F_1(x, Q^2)$$

$$= \sum_a C^n_{1,a}(\mu^2) v^n_a(\mu^2) \Big|_{\mu^2 = Q^2}$$

| S_f | for $\langle x^3\rangle(u_+ - d_+)$, CT14HERA2NNLO

LATTICE PDF **Community Whitepaper:**
Prog. Part. Nucl. Phys. 100 (2018) 107-160.

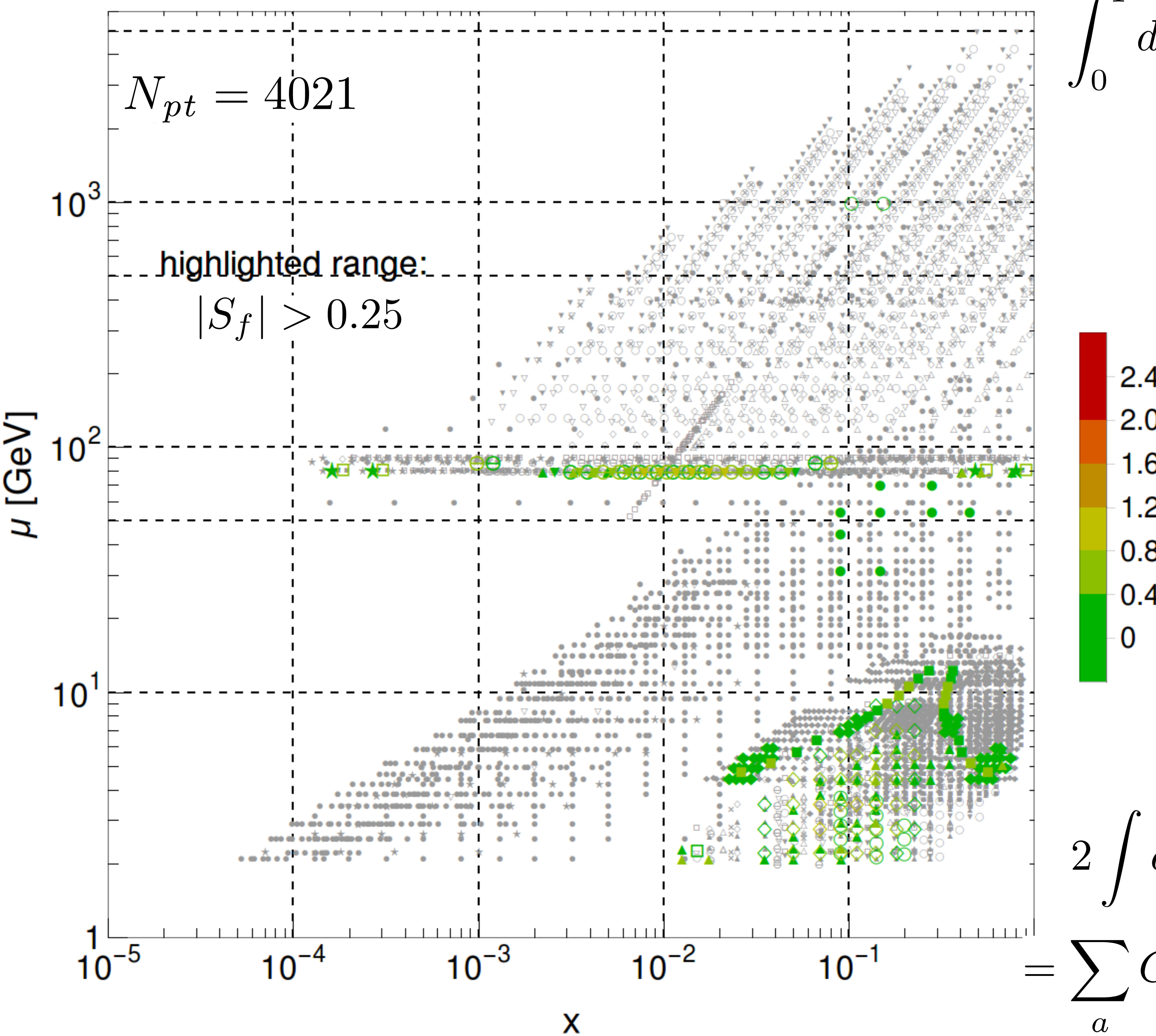


- might EIC constrain knowledge of lattice-calculable PDF Mellin moments?
- JL-EIC possesses successively larger sensitivities at high x as the order of the Mellin moments increases!
- **there will be a future synergy between LQCD and HEP phenom.**

$$2\int dx x^{n-1} F_1(x, Q^2)$$

$$= \sum_a C^n_{1,a}(\mu^2) v^n_a(\mu^2)\big|_{\mu^2=Q^2}$$

| S_f | for <x^1>u+−d+, CT14HERA2NNLO



$$\int_0^1 dx x^n q(x, \mu = 2\,\mathrm{GeV})$$

- **this can be compared to the picture obtained with the full CTEQ-TEA expanded dataset**

- **there will be a future synergy between LQCD and HEP phenom.**

$$2 \int dx x^{n-1} F_1(x, Q^2)$$

$$= \sum_a C^n_{1,a}(\mu^2) v^n_a(\mu^2) \Big|_{\mu^2 = Q^2}$$

conclusions and **future directions**

- **PDFSENSE** allows a thorough exploration of the PDF impact of current and potential data ideal for investigating possible EIC scenarios

  → has been used to identify and validate LHC-driven expansions of the CTEQ dataset

  → pseudodata sensitivities (**JL-EIC**, AFTER@CERN, LHeC); only need estimated stat./sys. uncertainties in bins of definite x and Q^2

  → a future EIC can be expected to hugely improve collinear PDFs

  → now working to relate the specifics of EIC proposals to their expected impact – especially in the setting of other possible experiments

- Moreover, we can compute sensitivities to many PDF-derived quantities: e.g., physical cross sections; PDF combinations; **lattice-calculable moments**, …

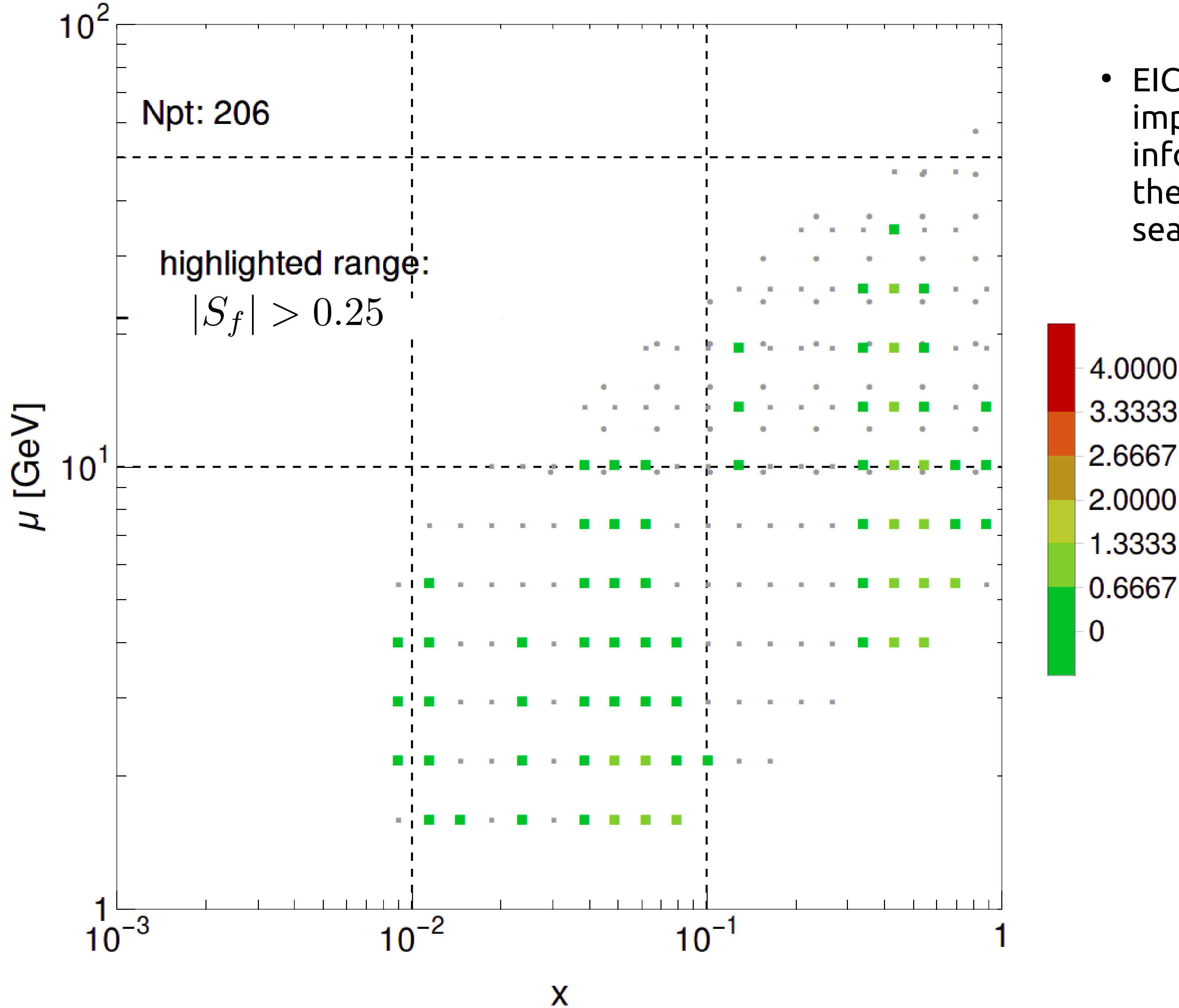
  PDFLattice whitepaper: arXiv:1711.07916

  → **other interesting quantities** to motivate EIC (TMDs, nPDFs, …)?

thanks!

additional material

$|\, S_f \,|$ for $\overline{d}/\overline{u}(x,\mu)$, JL-EIC NC+CC



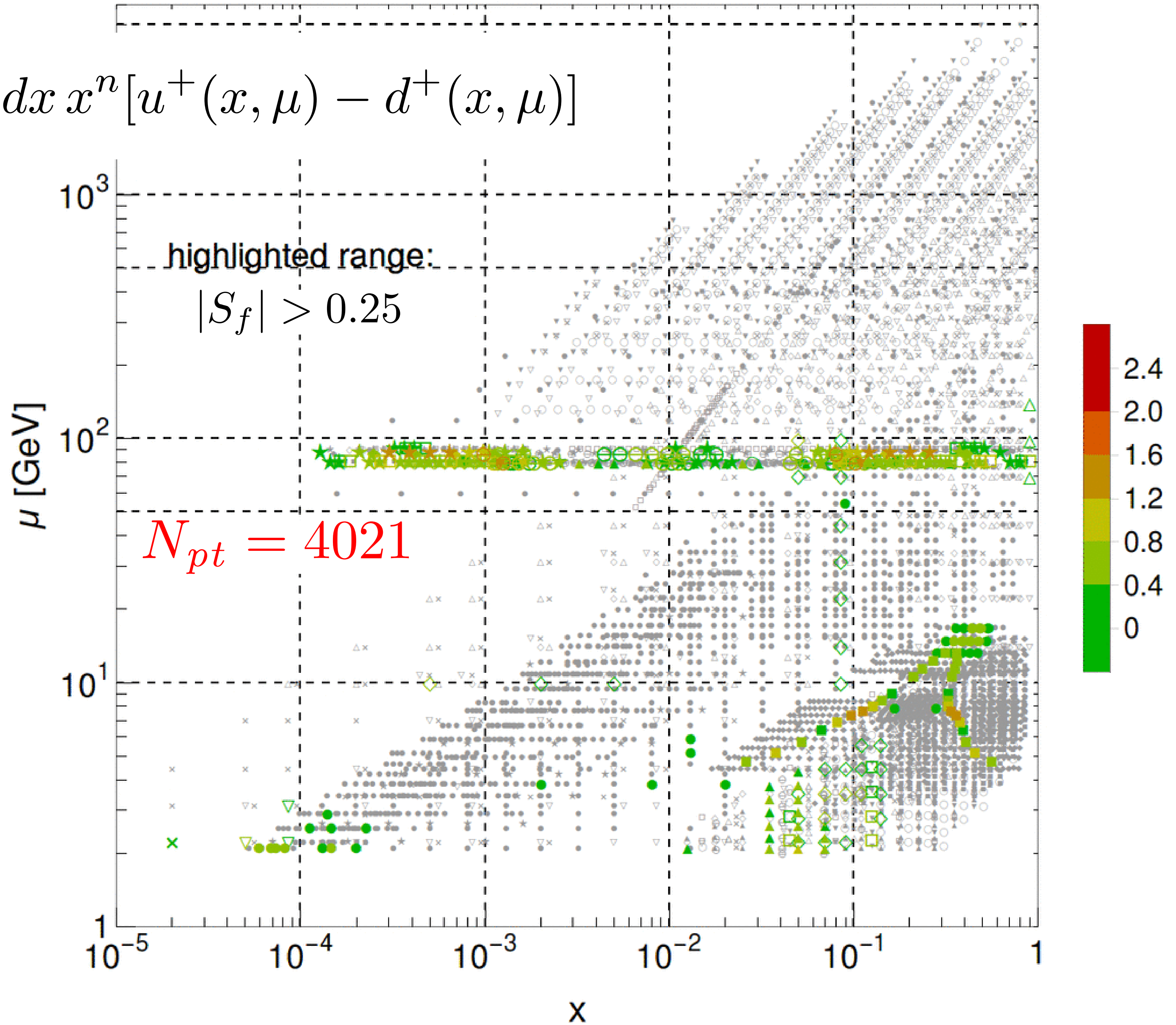
- EIC can provide important information on the light quark sea asymmetry

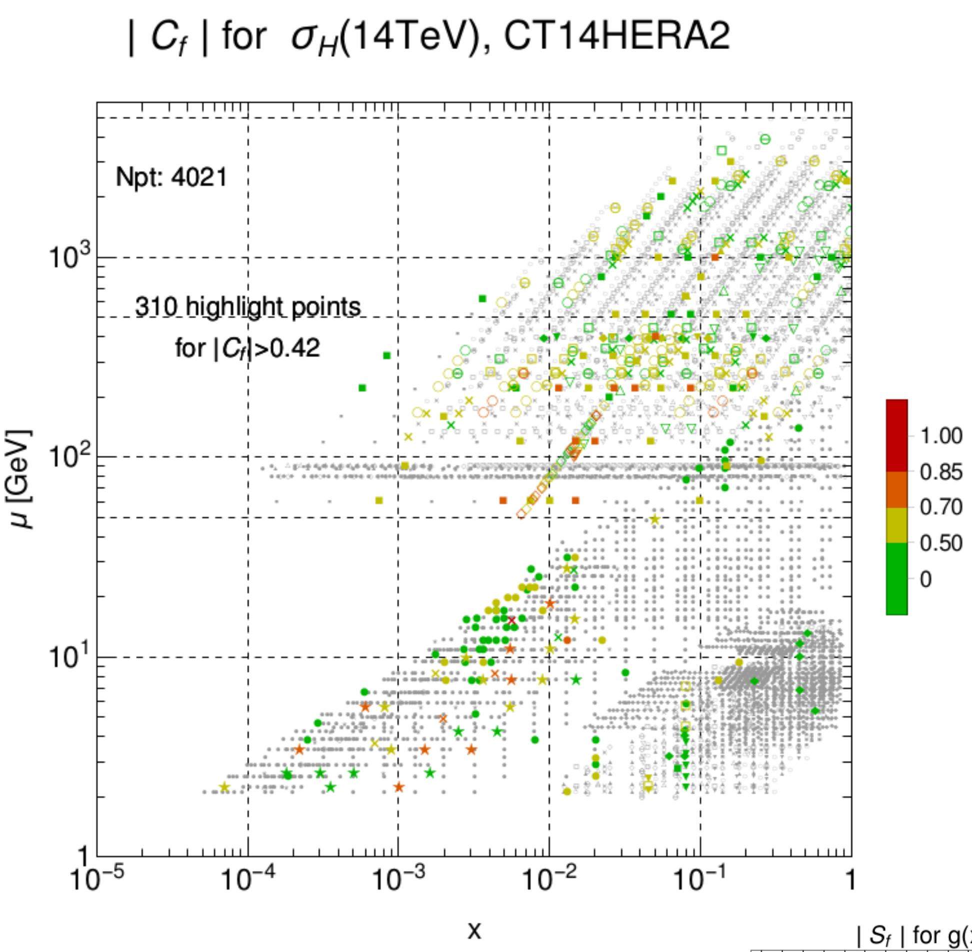
| S_f | for <x^-1>u+-d+, CT14HERA2NNLO

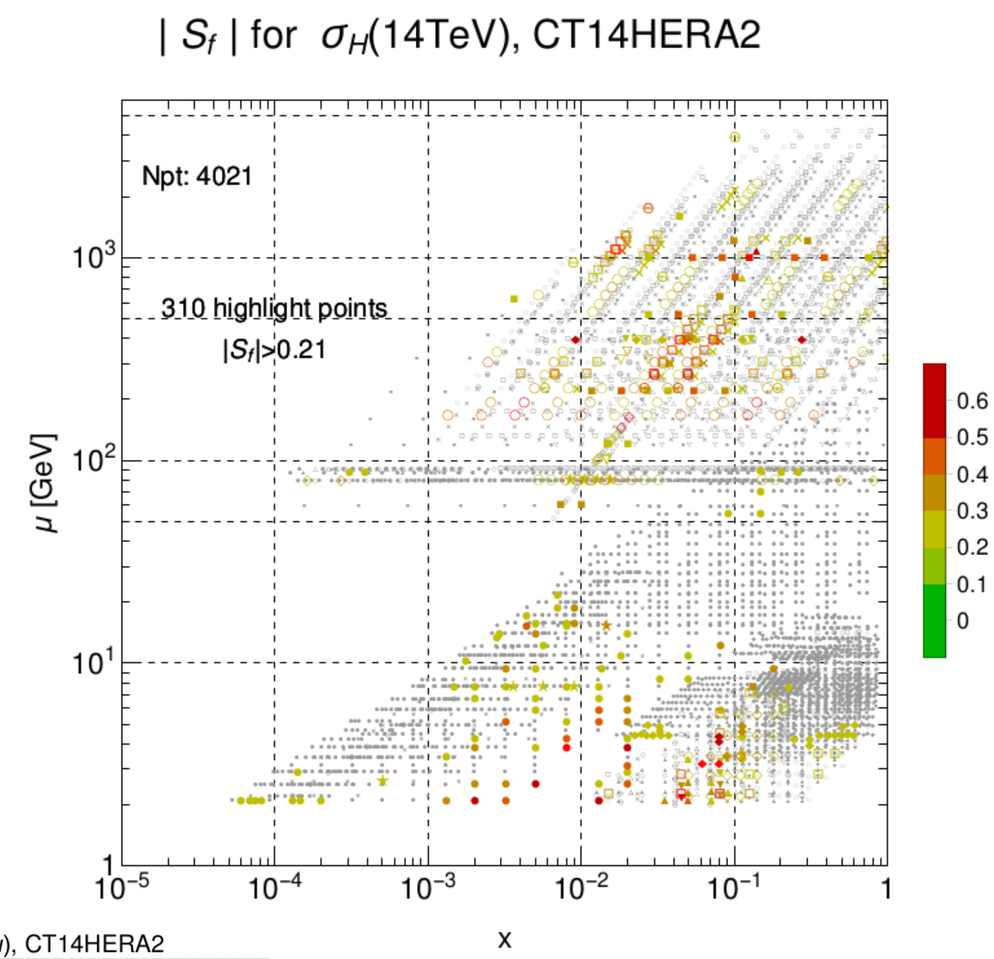
$$M^n_{u^+-d^+} = \int_0^1 dx\, x^n [u^+(x,\mu) - d^+(x,\mu)]$$

- **1st several PDF moments can be computed on the lattice; knowledge of sensitivities suggests where lattice calculations might constrain PDFs**

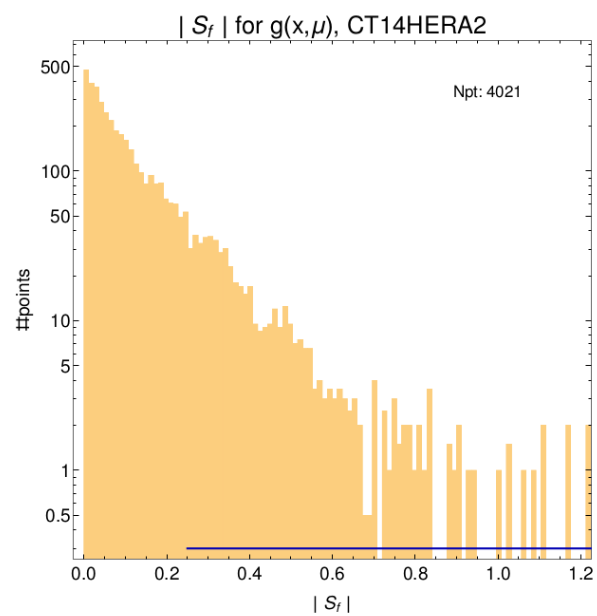
- **higher moments shift sensitivities *rightward***



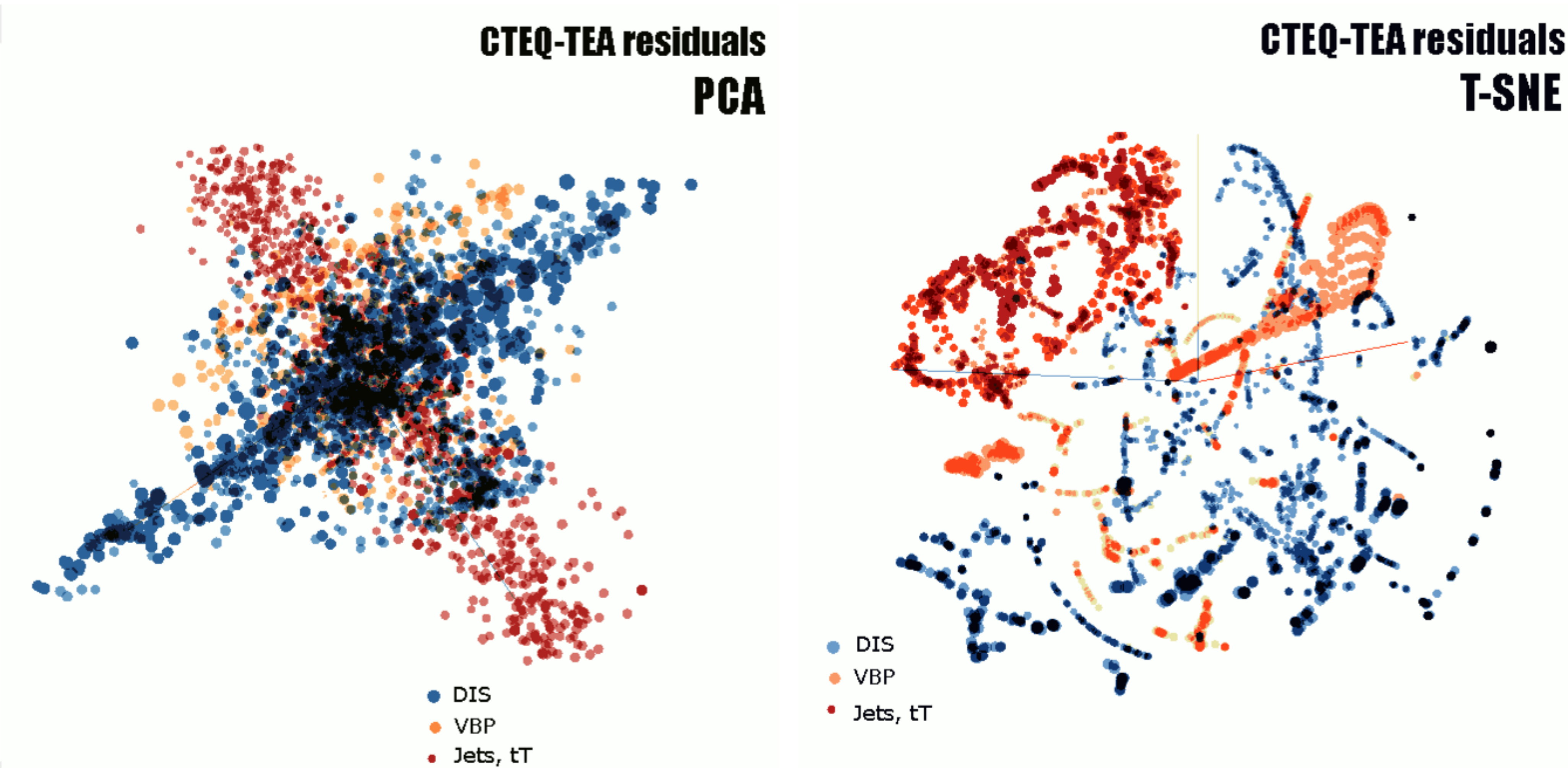




a typical sensitivity distribution;
in this case, for the gluon PDF:

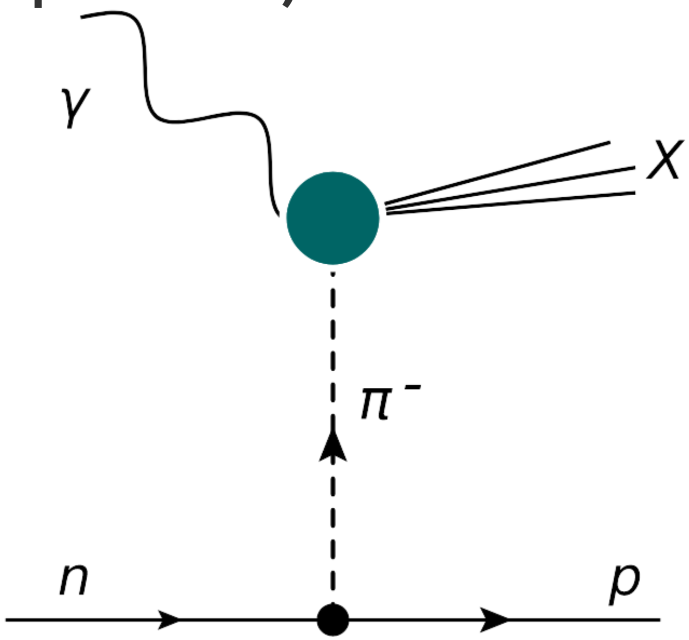


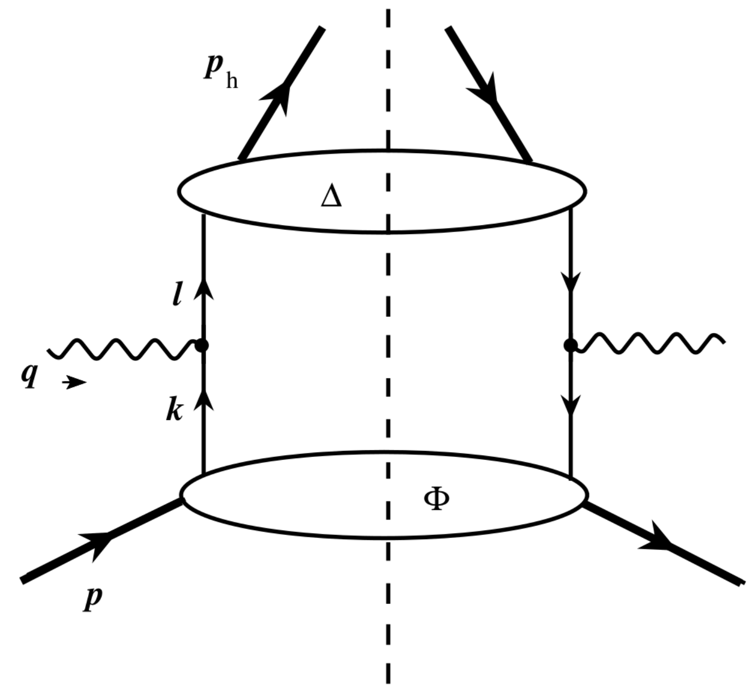
machine learning techniques applied to ensembles of residuals



going forward: PDFSense for meson structure at EIC

- **the PDFSense framework is of sufficient generality that the same approach might also be applied to EIC-relevant processes involving, e.g., $q_\pi(x, \mu^2)$**
  - **spectator-tagging for forward produced baryons (TDIS, leading neutron production):**
  - **standard semi-inclusive pion or kaon production in DIS:**



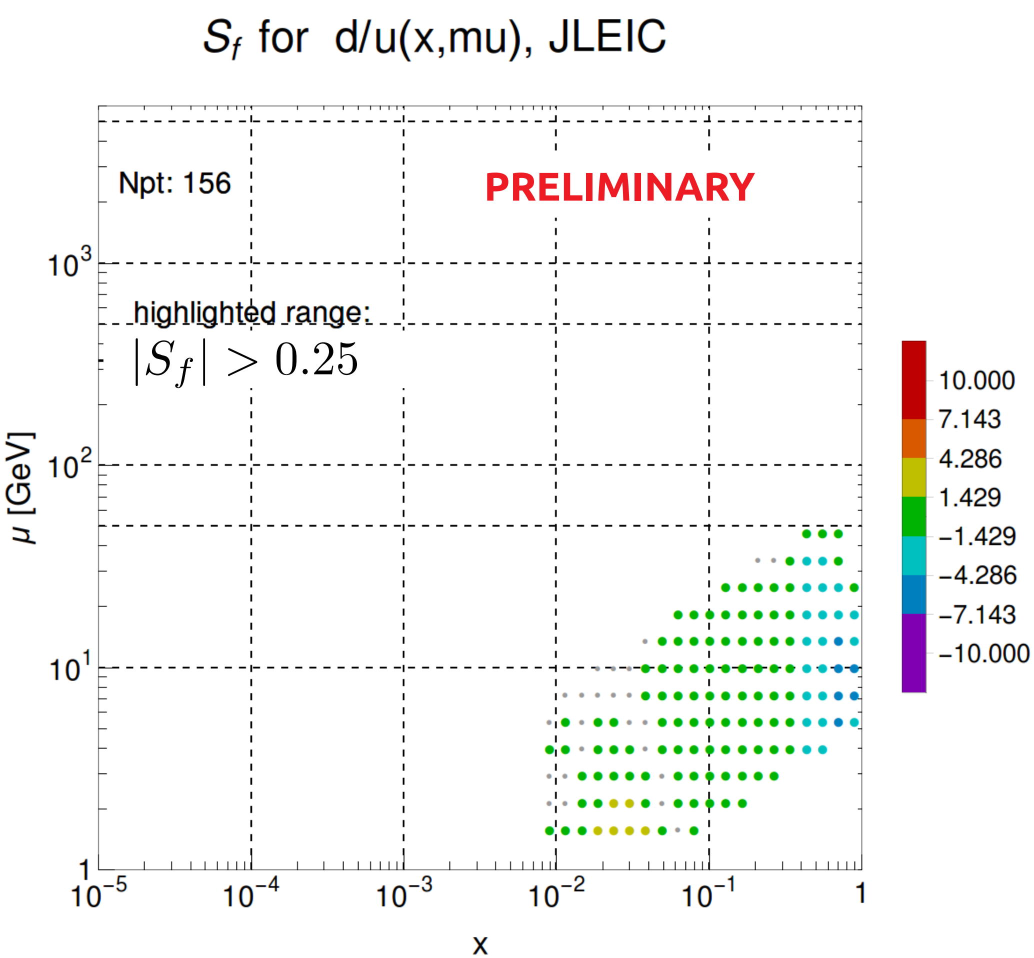


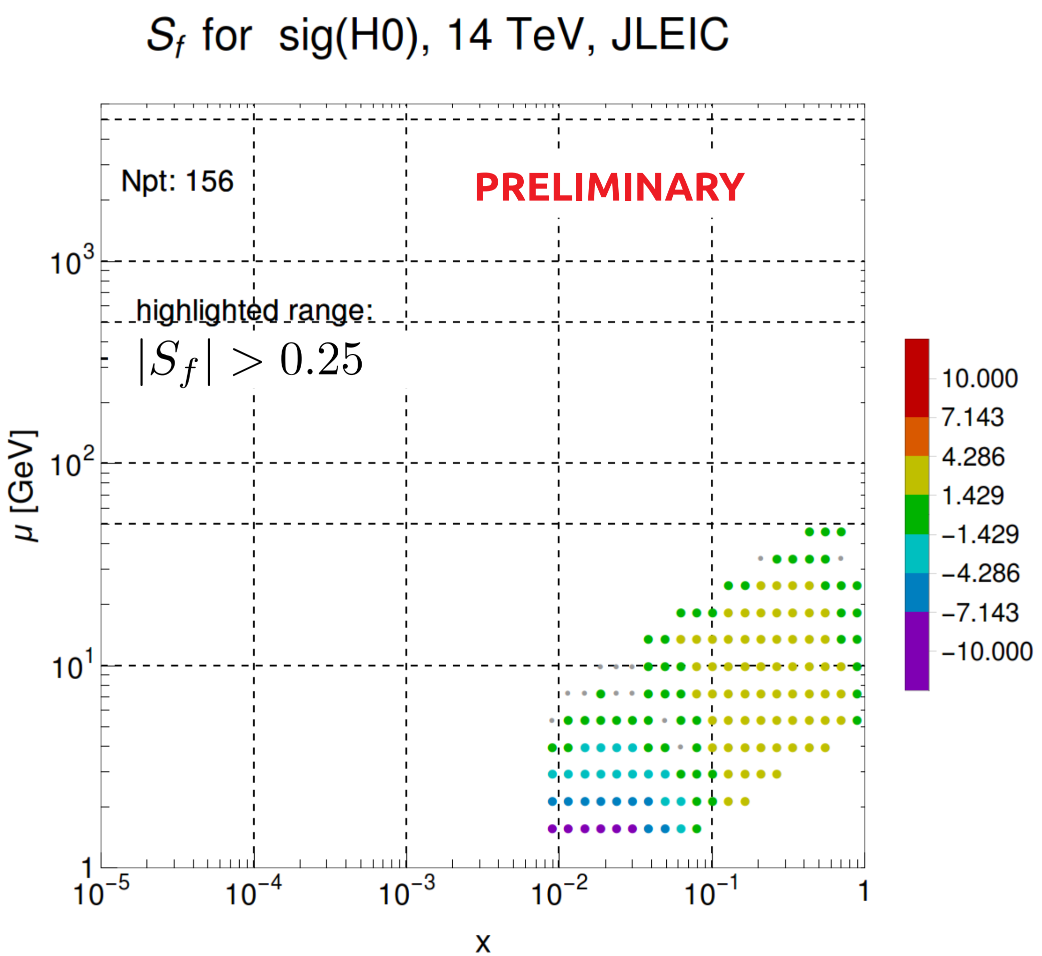
- **EIC projections for measurements of these processes may be constructed along the lines of the EW cross sections shown in previous slides; with this information, we might explore optimal EIC run scenarios for meson PDF extraction**

JL-EIC will have a powerful impact on both hadron structure and HEP phenomenology

$$d(x,\mu),\ \ d/u(x,\mu)$$

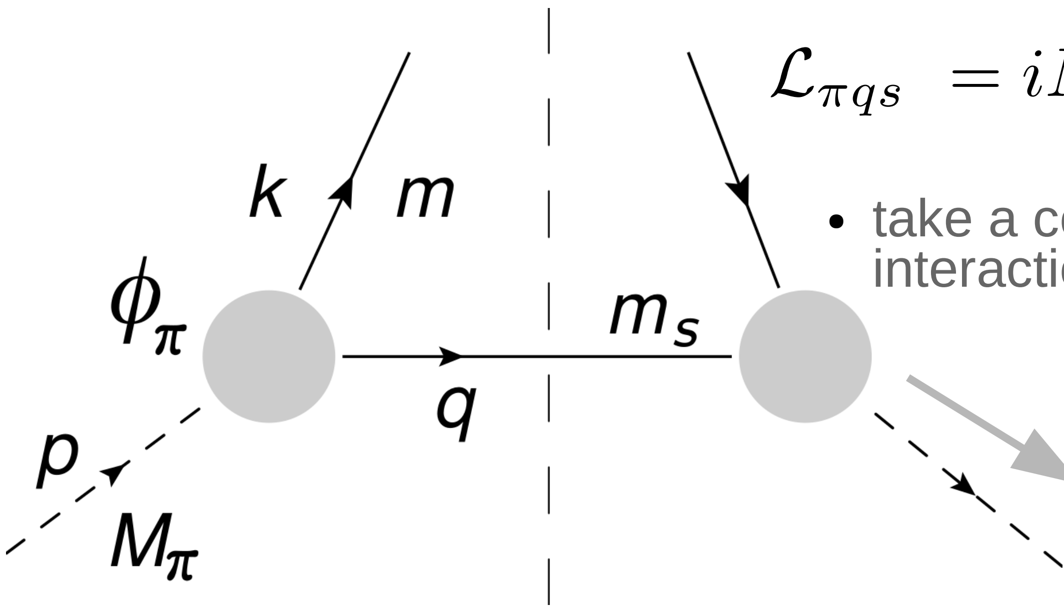
- **concerning higher energy processes, JL-EIC will impose tight constraints on the Higgs phenomenology**





the “exact” pion light-front PDF via a constituent quark model

- first evaluate the LF pion valence PDF using a minimal model that couples the pion to its constituent quarks



$$\mathcal{L}_{\pi q s} = i N_\pi^{1/2}\, \overline{\psi}_q \gamma_5 \varphi_\pi \psi_s + \text{h.c.}$$

- take a covariant vertex factor for the quark-pion interaction consistent with high x DSE results:

$$q_v(x) \sim (1-x)^2$$

$$\left[\phi_\pi(k^2)\right]^2 \equiv \left[\Lambda_\pi^2 \Big/ (k^2 - \Lambda_\pi^2)\right]^{3/2}$$

$$q_\pi^{\mathrm{LF}}(x) = \frac{N_\pi}{2\,(2\pi)^4} \int dk^+ dk^- d^2k_\perp \left(\frac{1}{2p^+}\right) \delta\left(x - \frac{k^+}{p^+}\right)$$

$$\times\, tr\Big(\gamma_5\,(\not{k} + m)\,\gamma^+ (\not{k} + m)\gamma_5(-\not{q} + m_s)\Big)\; 2\pi\,\delta\big(q^2 - m_s^2\big) \left[\frac{\phi_\pi(k^2)}{(k^2 - m^2)}\right]^2$$

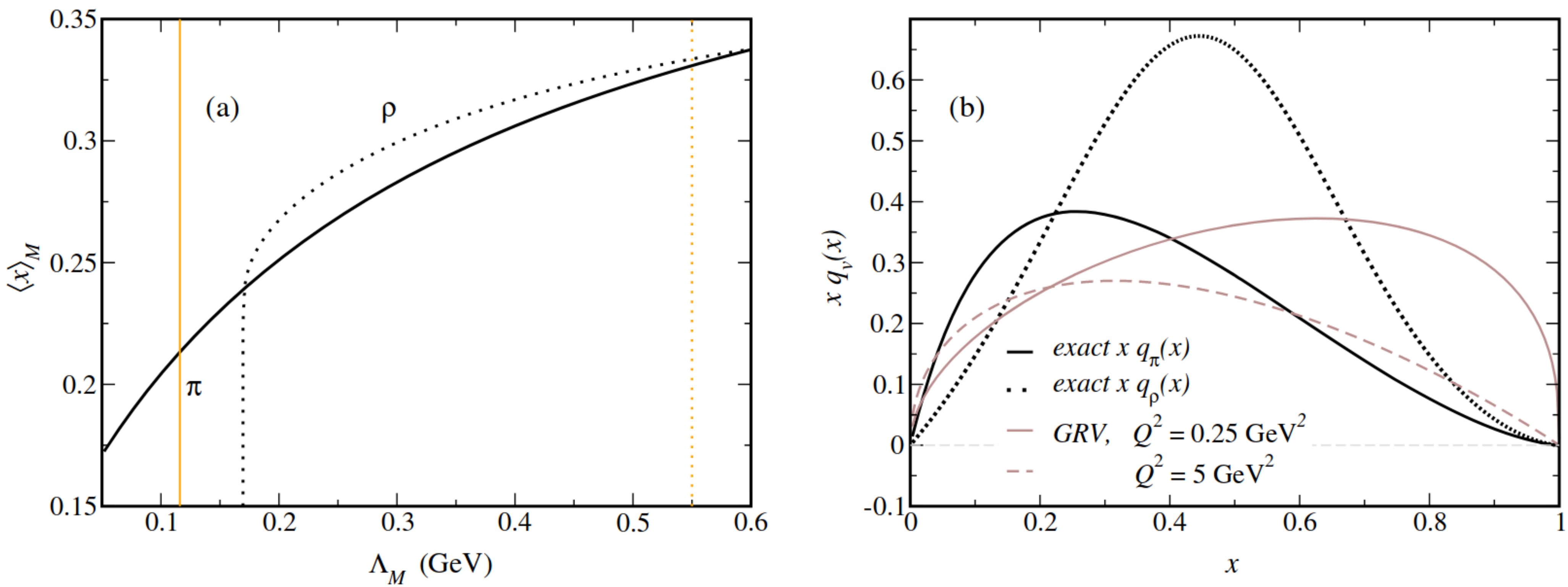
$$\boxed{q = p - k}$$

$$q_\pi^{\mathrm{LF}}(x) = \frac{N_\pi}{8\pi^2} \int \frac{dk_\perp^2}{x^2(1-x)} \left\{k_\perp^2 + \big(x m_s + (1-x)\,m\big)^2\right\} \left[\frac{\phi_\pi(t_\pi)}{(M_\pi^2 - \hat{s})}\right]^2$$

determining the pion SF model parameters

- for the pion, masses can be fixed to physical or constituent values:

$$M_\pi = 0.139\,\mathrm{GeV},\;\; m = M/3 \approx 0.33\,\mathrm{GeV}$$



$$\langle x \rangle_\pi = \int_0^1 dx\, x q_\pi^{\mathrm{LF}}(x) = 0.214(15)\left({}^{+12}_{-9}\right)^* \longrightarrow \Lambda_\pi = 0.116\,\mathrm{GeV}$$

*LQCD 1st moment calculation: Abdel-Rehim et al., PRD**92**, 114513 (2015).

- the overall strength is set by a **normalization condition** such that the model is then completely determined

$$N_\pi = 1 \Big/ \int dx\, q_\pi^{\mathrm{LF}}(x)$$

the corresponding pion quasi-PDF may then be found:

$$\widetilde{q}_\pi(x,p_z) \;=\; \frac{N_\pi}{(2\pi)^4}\int dk^0 dk_z\, d^2k_\perp \left(\frac{1}{2p_z}\right)\delta\left(x-\frac{k_z}{p_z}\right)$$

$$\times\; tr\Big(\gamma_5\,(\not{k}+m)\,\gamma^z(\not{k}+m)\gamma_5(-\not{q}+m_s)\Big)\;2\pi\,\delta\big(q^2-m_s^2\big)\left[\frac{\phi_\pi(k^2)}{(k^2-m^2)}\right]^2$$

- now, integrating delta functions introduces explicit dependence on p_z —

$$\delta\left(q^2-m_s^2\right) \;=\; \frac{1}{2\left(p^0-k^0\right)}\,\delta\left(p^0-k^0-\sqrt{m_s^2+k_\perp^2+(1-x)^2p_z^2}\right)$$

$$\widetilde{q}_\pi(x,p_z) \;=\; \frac{N_\pi}{4\pi^2}\int\frac{dk_\perp^2}{2(1-x)\mu_s}\;\Big\{2x\Big(mm_s+(\tilde{q}\cdot\tilde{k}_\pi)\Big)+(m^2-\tilde{k}_\pi^2)\,(1-x)\Big\}$$

$$\times\left[\frac{\phi_\pi(\tilde{k}_\pi^2)}{(M_\pi^2+m_s^2-m^2+2(1-x)\big(1-\mu_\pi\,\mu_s\big)}\right]^2$$

(the main result for the pion quasi-PDF)

$$\mu_\pi \;\equiv\; \sqrt{1+\frac{M_\pi^2}{p_z^2}}$$

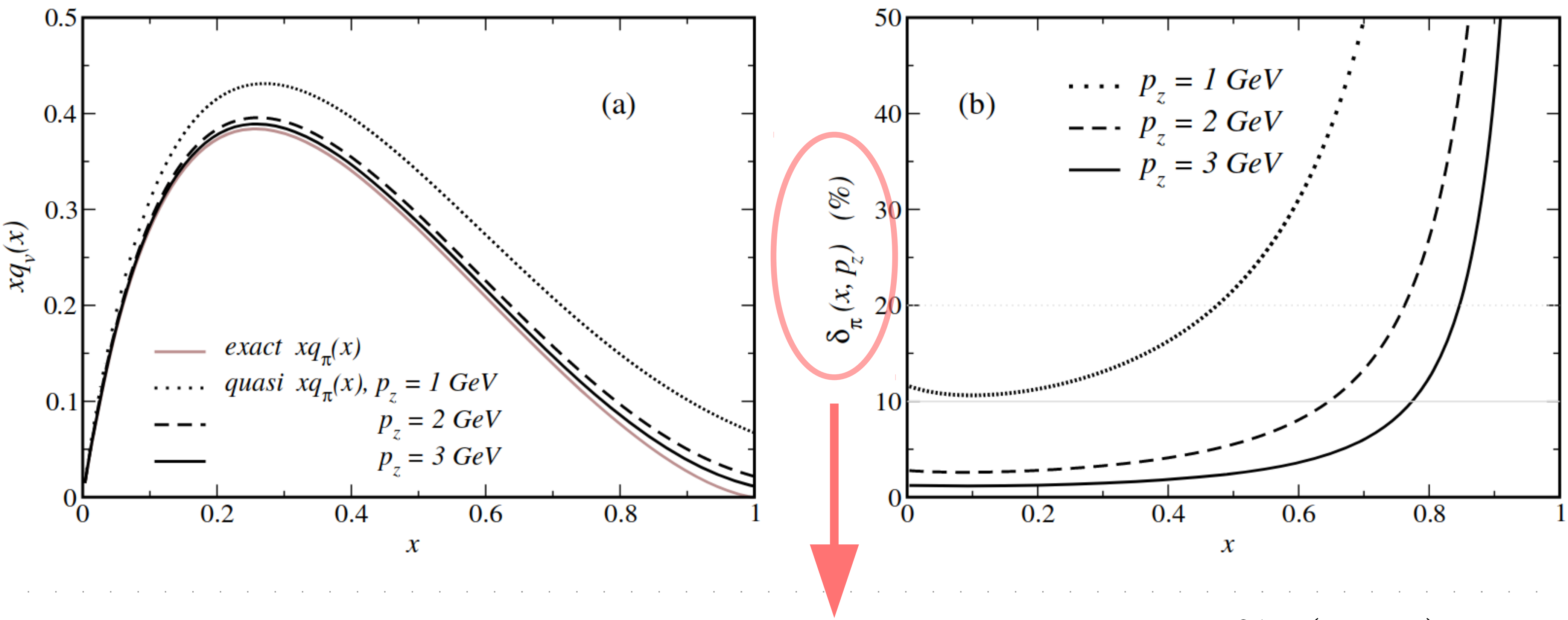
$$\mu_s \;\equiv\; \sqrt{1+\frac{m_s^2+k_\perp^2}{(1-x)^2p_z^2}}$$

$$\tilde{k}_\pi^2 \;=\; M_\pi^2+m_s^2+2\,(1-x)\big(1-\mu_\pi\,\mu_s\big)\,p_z^2$$

→ compare π quasi-/PDFs for several p_z

- we observe the expected behavior: at infinite boost, meson quasi-PDFs match onto the exact result,

$$\lim_{p_z \to \infty} \widetilde{q}_M(x, p_z) = q_M^{\mathrm{LF}}(x)$$

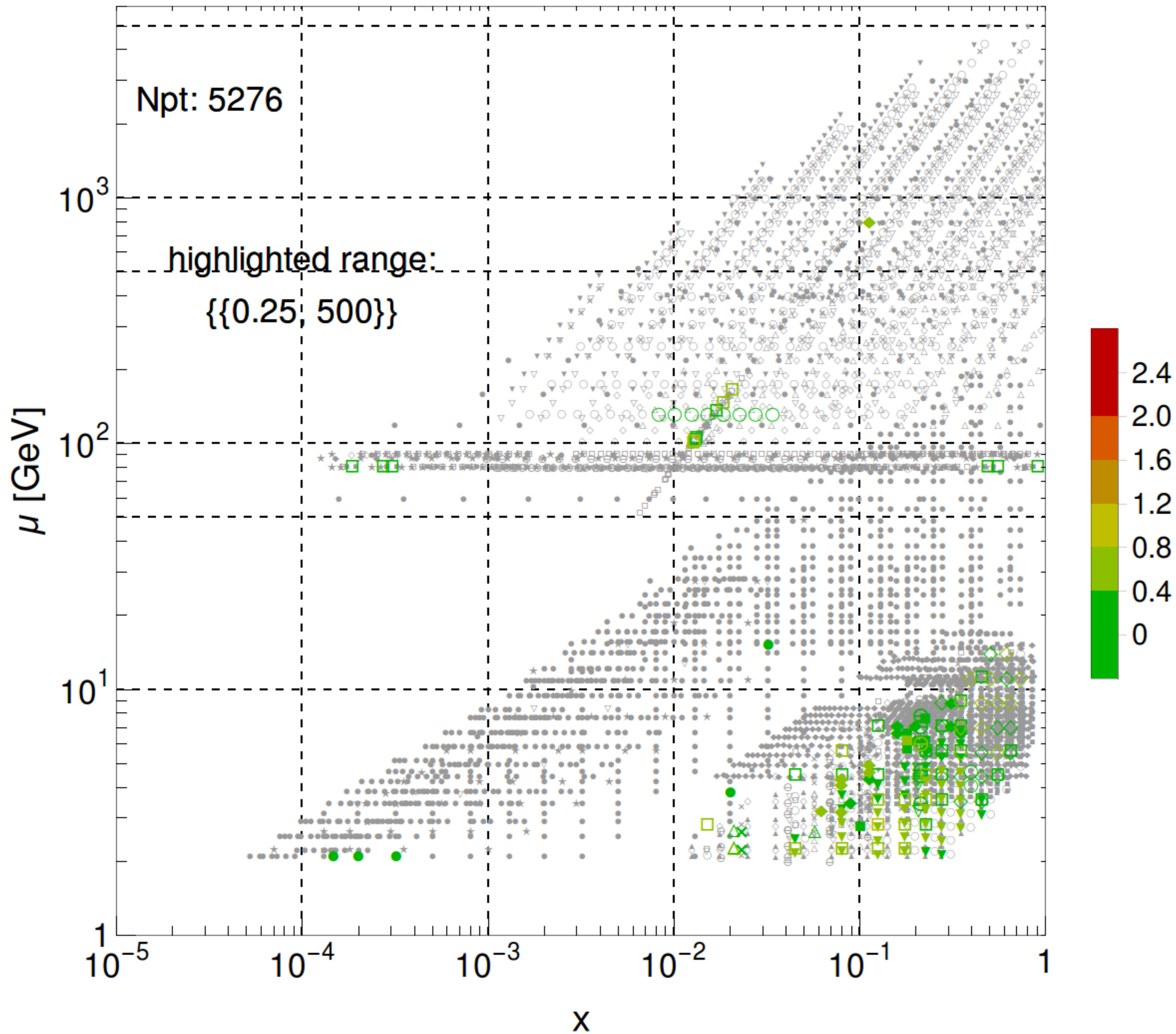


- away from this limit, we compute the LaMET deviations from the LF PDF:

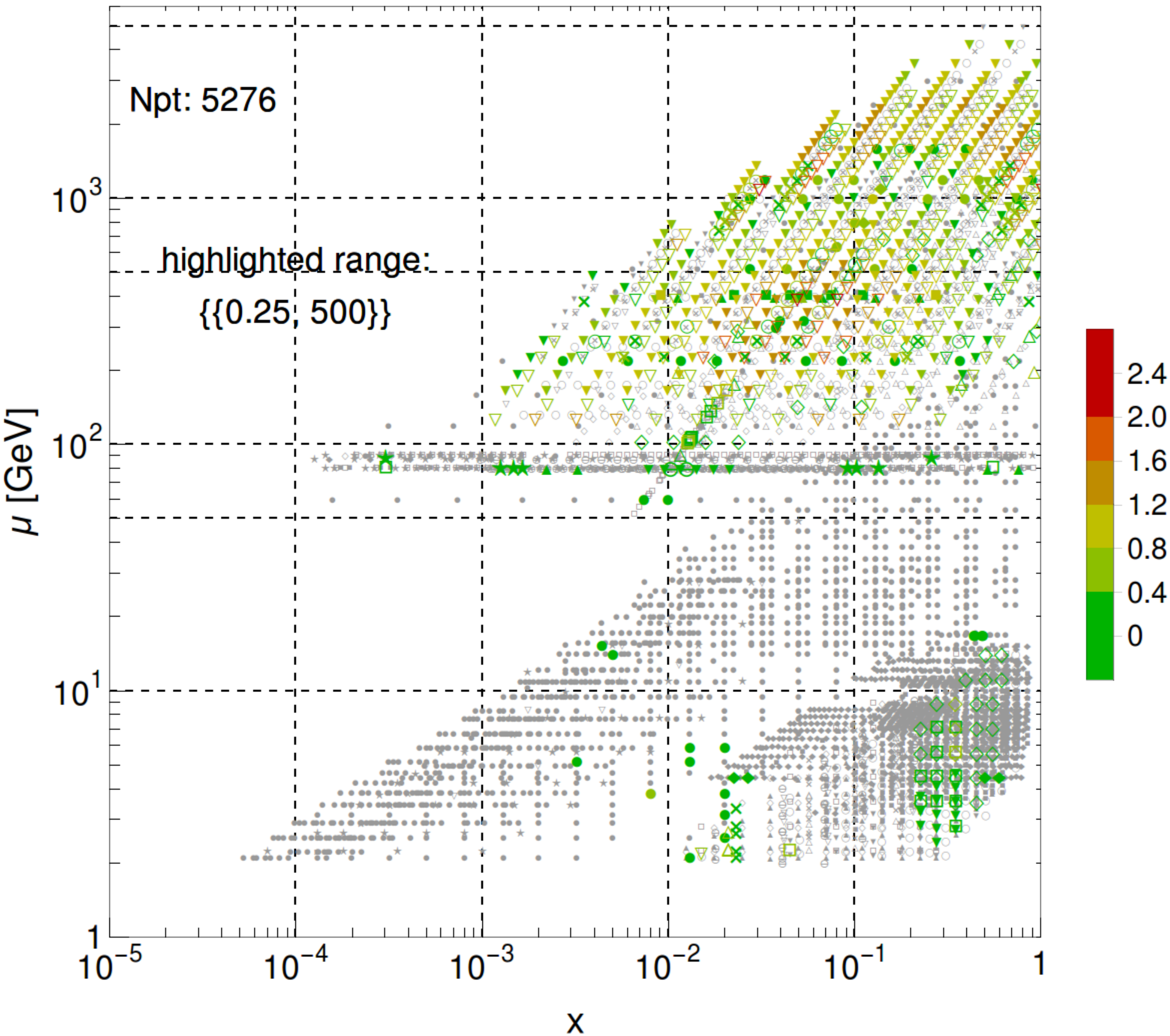
$$\delta_M(x, p_z) \equiv \frac{\widetilde{q}_M(x, p_z)}{q_M^{\mathrm{LF}}(x)} - 1$$

→ even at fairly modest p_z these corrections can be $\lesssim 10\%$!

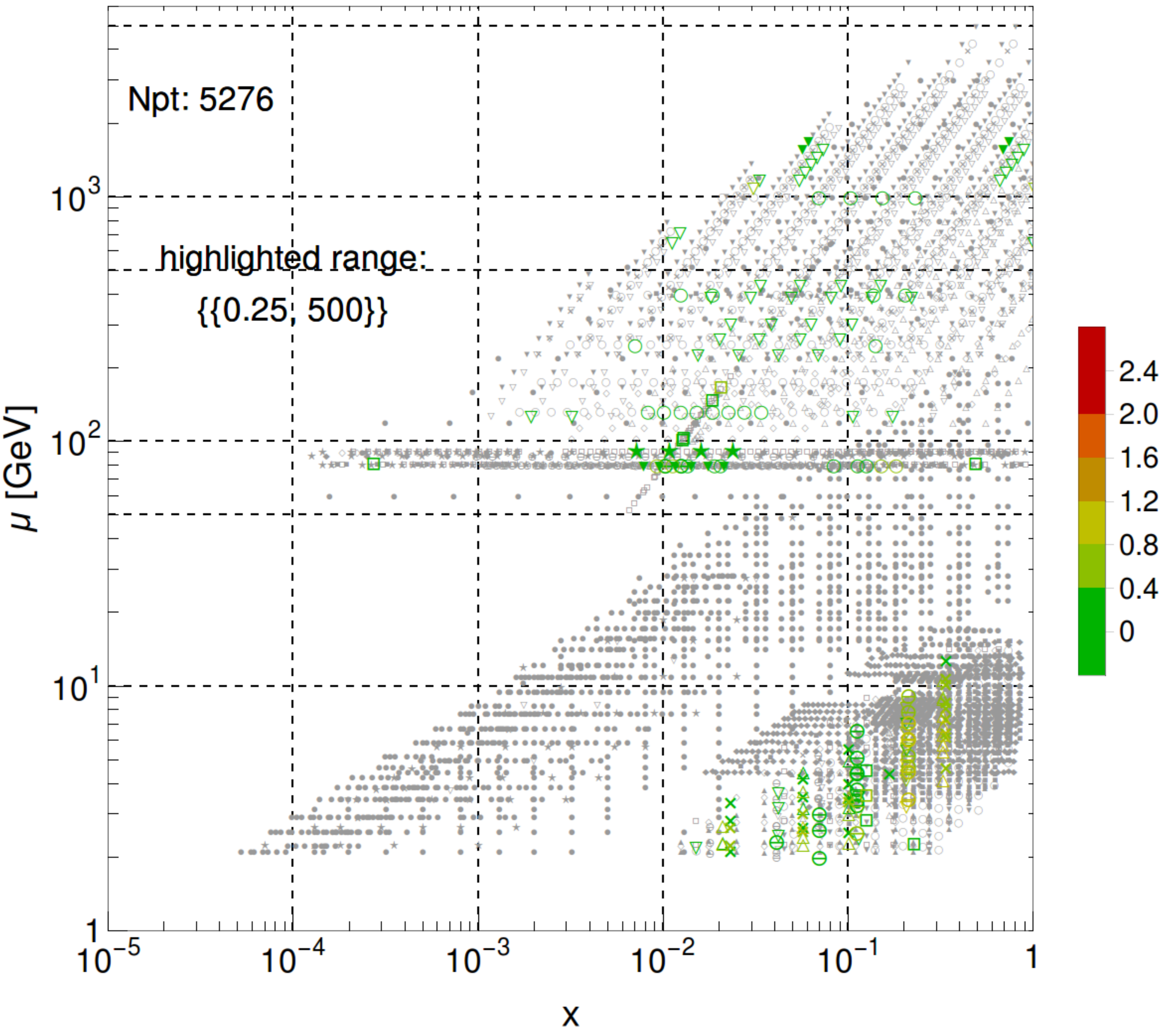
| S_f | for <x^1>g, CT14HERA2NNLO



| S_f | for <x^3>g, CT14HERA2NNLO



| S_f | for <x^1>s+, CT14HERA2NNLO



| S_f | for <x^3>s+, CT14HERA2NNLO