European Research Council
Established by the European Commission

PBSP

MARIA UBIALI
UNIVERSITY OF CAMBRIDGE

PDF AND SMEFT INTERPLAY

LA THUILE 2024 - LES RENCONTRES DE PHYSIQUE DE LA VALLÉE D'AOSTE

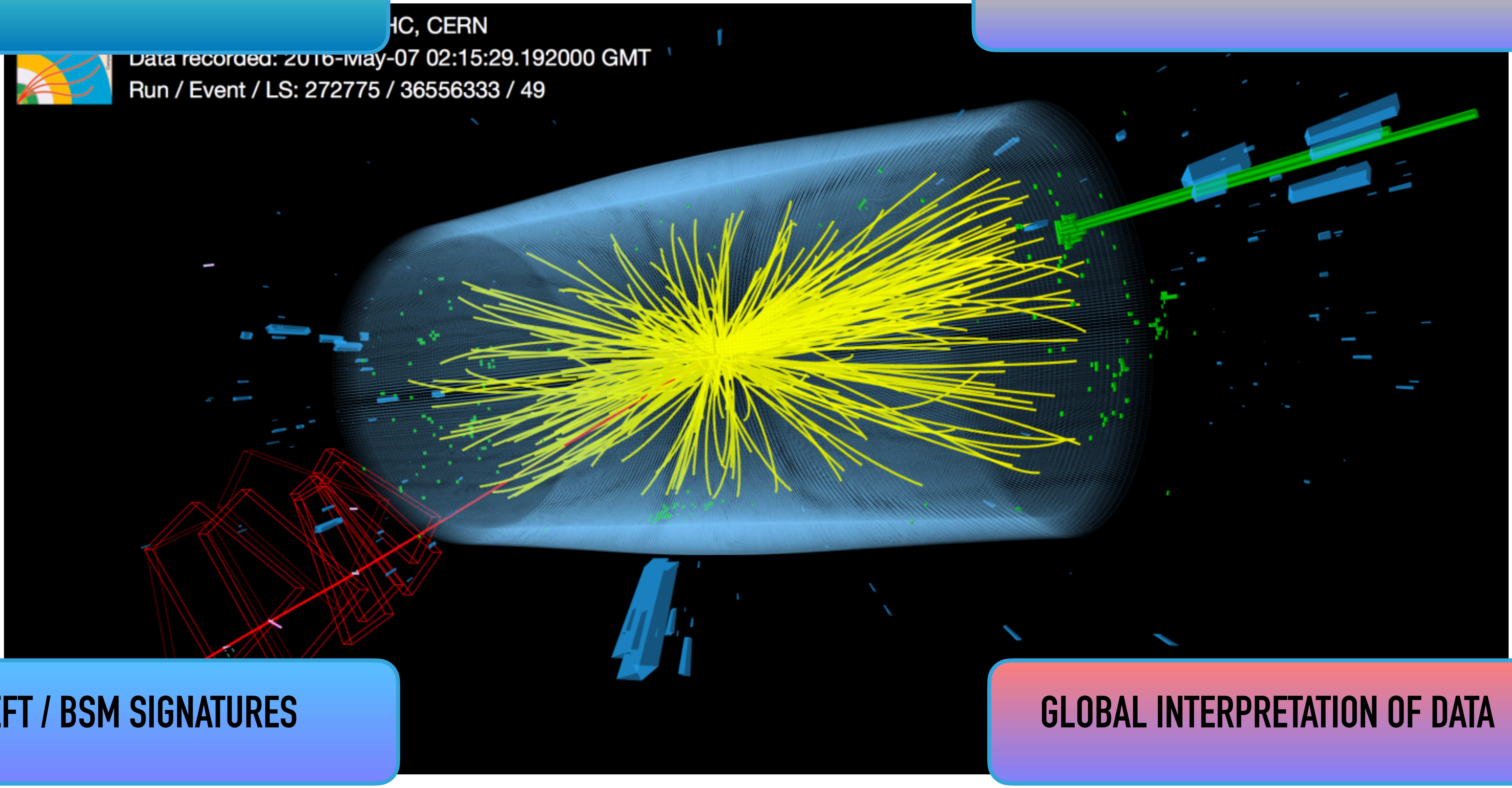
7TH MARCH 2024

OUTLINE

Part I

PDFS

PRECISION PHYSICS



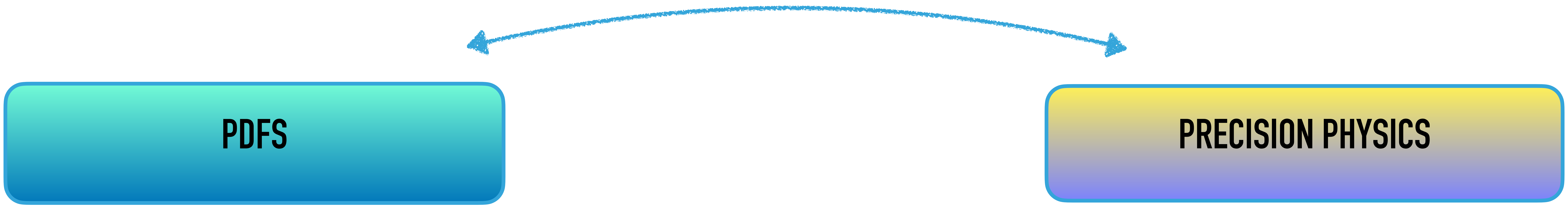
Part II

Part III

SMEFT / BSM SIGNATURES

GLOBAL INTERPRETATION OF DATA

Part III



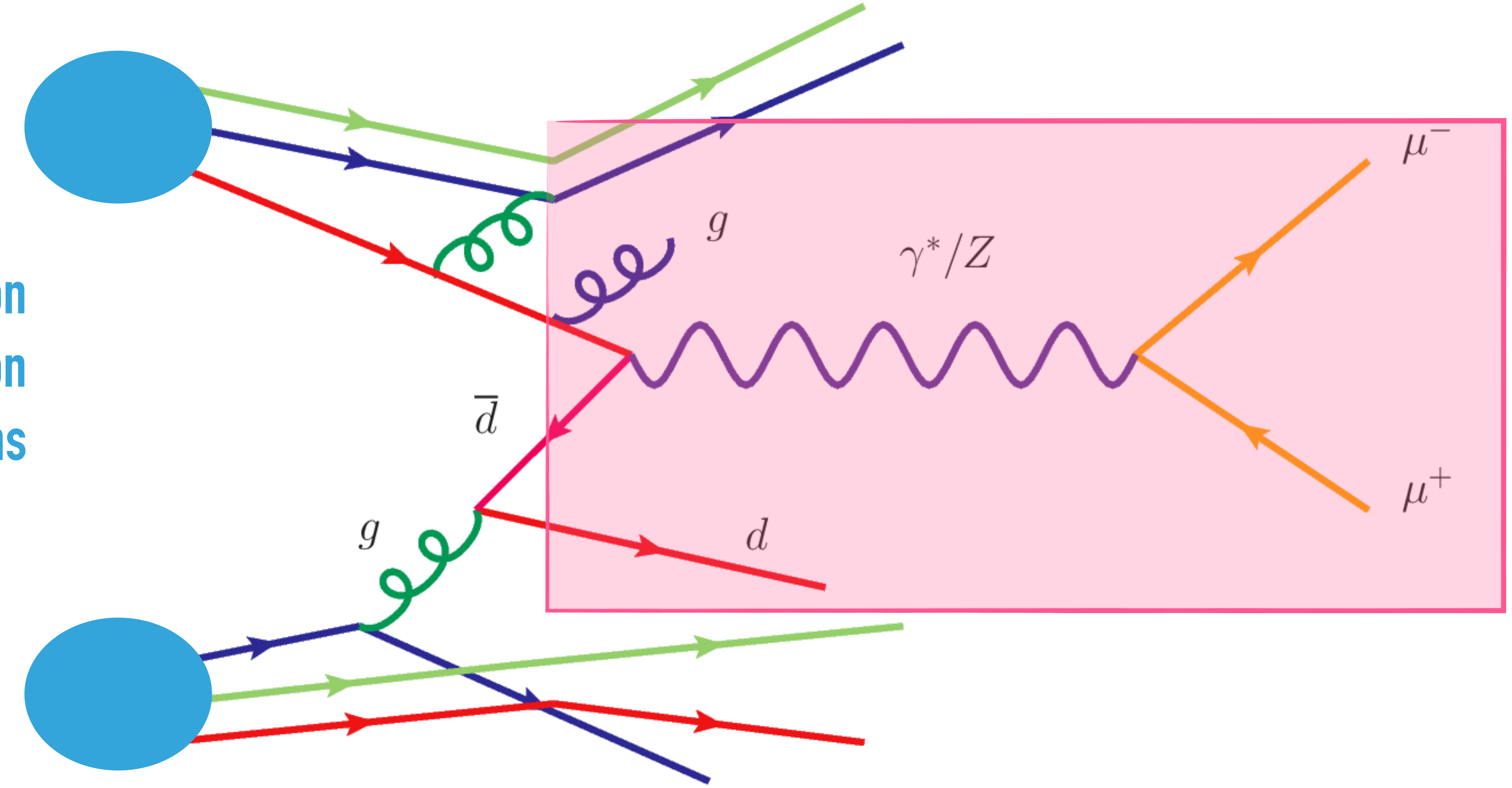
PDFS AND PRECISION PHYSICS

THEORETICAL PREDICTIONS AT THE LHC

$$\sigma^{pp\to ab} = \sum_{i,j=-n_f}^{n_f} \int dz_1\, dz_2\, f_i(z_1,\mu_F)\, f_j(z_2,\mu_F)\, \hat{\sigma}^{ij\to ab}(z_1 z_2 S, \alpha_s(\mu_R), \mu_F) + \mathcal{O}\left(\frac{\Lambda^n}{S^n}\right)$$

Collinear factorisation: separate long-distance universal information on proton structure in terms of partons (quarks, antiquark, gluons, photons…) from from short-distance parton interaction (hard scattering)

Parton Distribution Functions



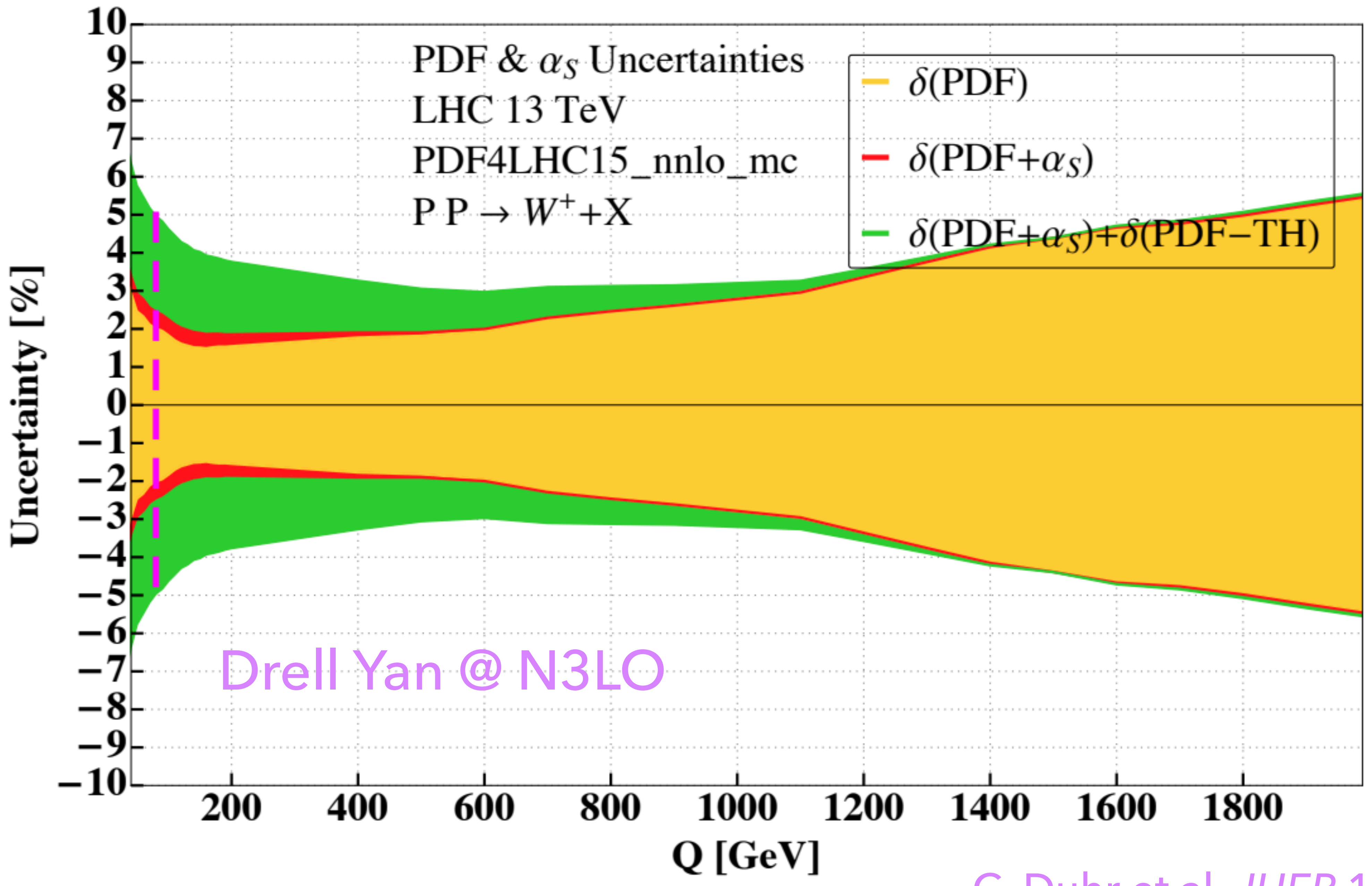
Hard Scattering: Perturbative QCD + EW

$$d\sigma^{pp\to ab} = \sum_{i,j} f_i \otimes f_j \otimes d\hat{\sigma}^{ij\to ab} + \dots$$

PDFS ARE A CRUCIAL INPUT FOR PRECISION PHYSICS ...

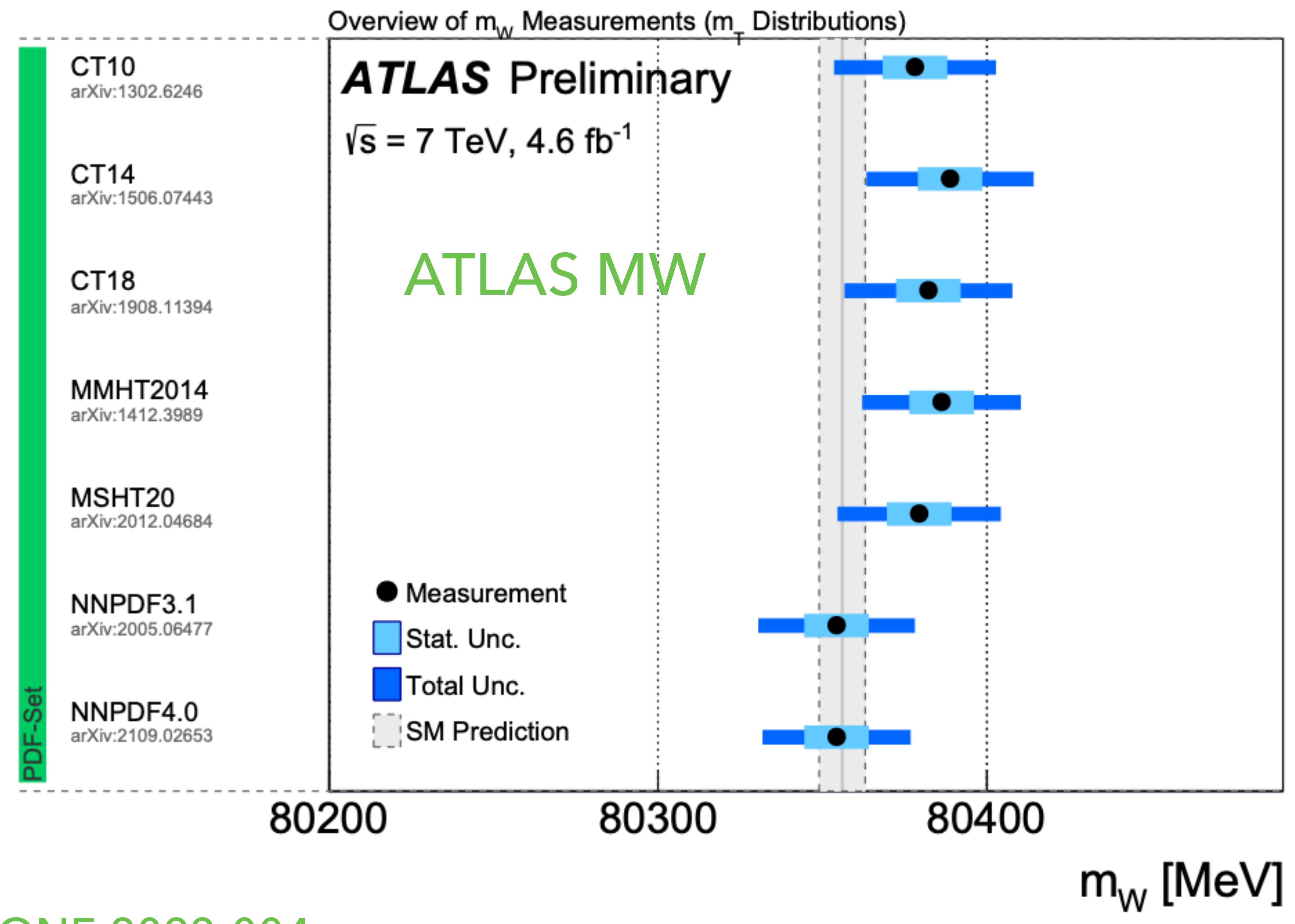
$$d\sigma^{pp\to ab} = \sum_{i,j} f_i \otimes f_j \otimes d\hat{\sigma}^{ij\to ab} + \ldots$$

#1: Theory uncertainty of SM predictions



C. Duhr et al, *JHEP* 11 (2020) 143

#2: Determination of SM parameters



ATLAS-CONF-2023-004

. . . AND FOR NEW PHYSICS SEARCHES

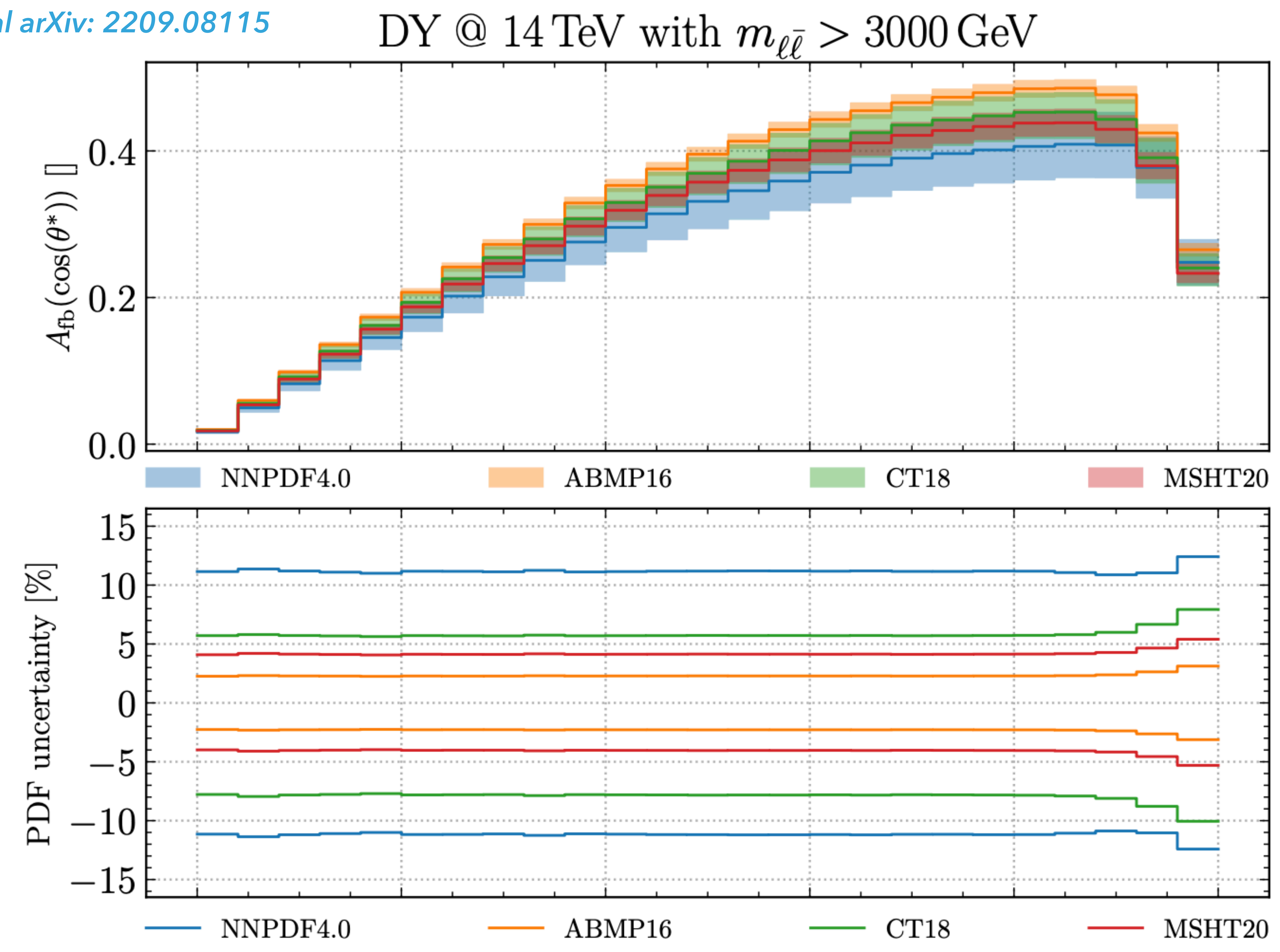
$$x \approx \frac{M}{\sqrt{S}}$$

High-mass final states

⇔

Large-x PDFs

Ball et al arXiv: 2209.08115



#1: Direct searches: heavy Z′ search in DY Forward-Backward asymmetry

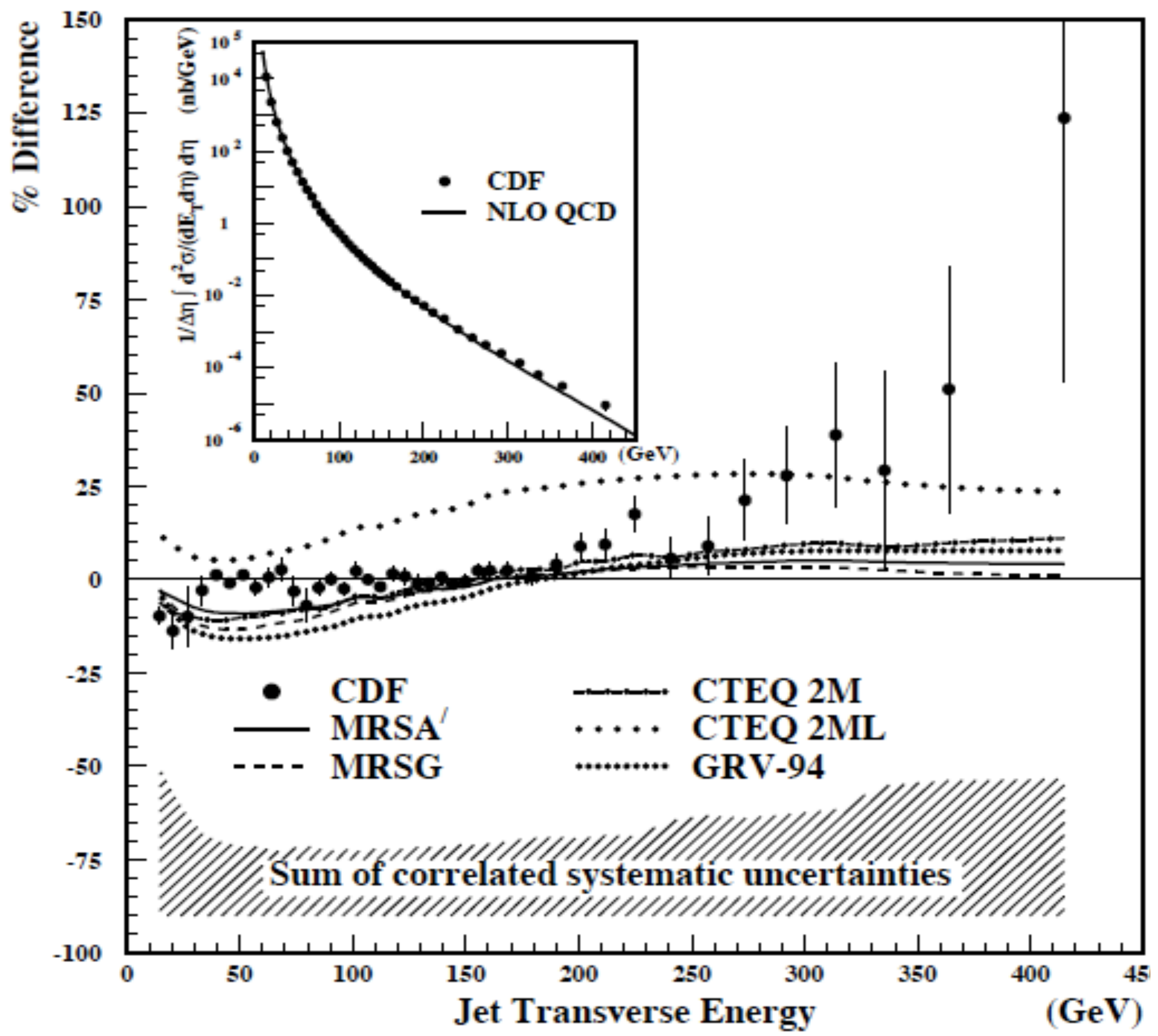
. . . AND FOR NEW PHYSICS SEARCHES

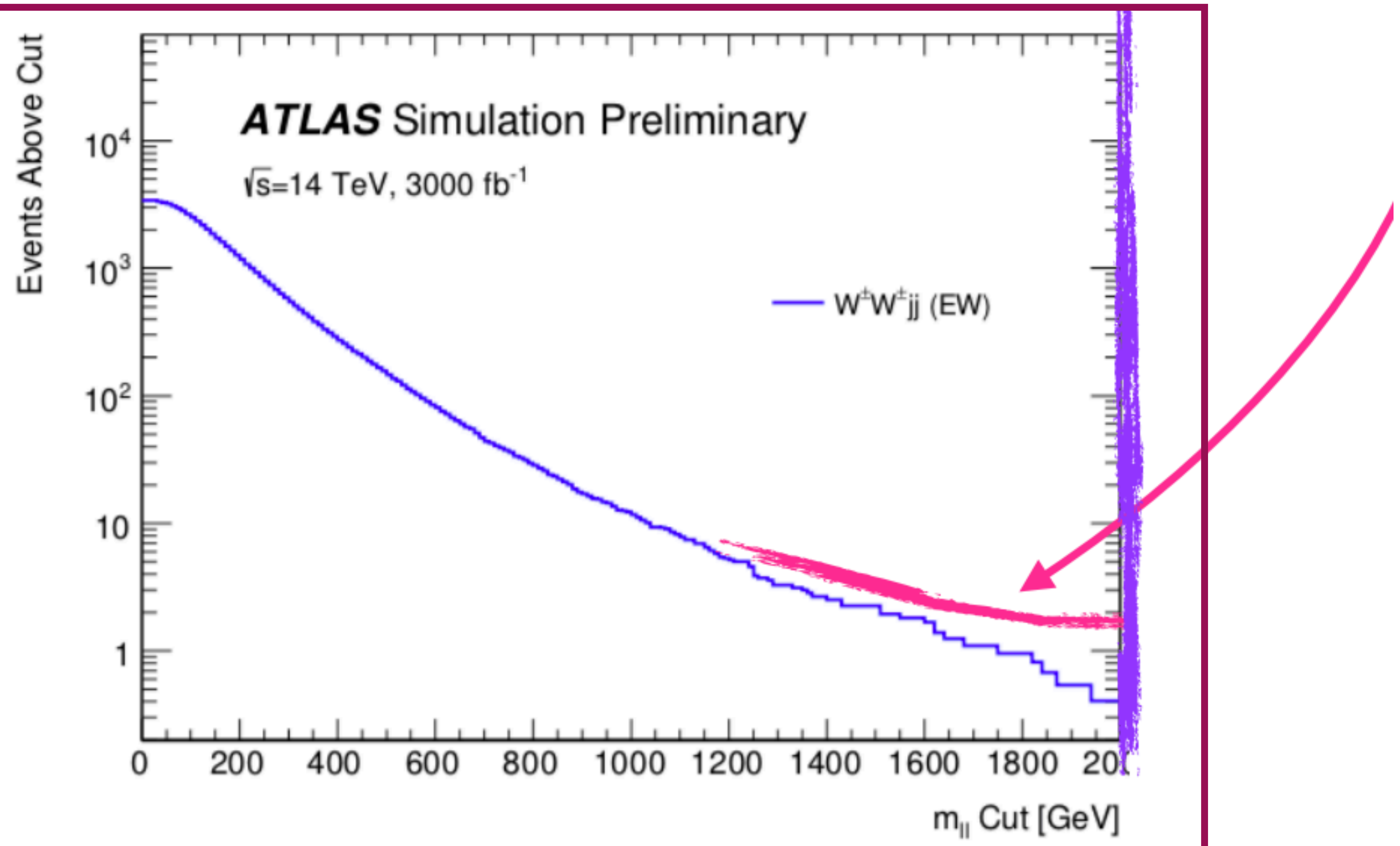
$$x \approx \frac{M}{\sqrt{S}}$$

High-mass final states

⇔

Large-x PDFs





CA Lee, HL/HE-LHC Jamboree, 1 March 2019

Discrepancy between QCD calculation and CDF jets data (1995)

At that time no information on PDF uncertainties and theory predictions strongly depended on gluon shape at x>0.1. Once data included in the CTEQ fit, discrepancy disappeared.

Deviations from SM predictions in high energy tails (>2023)

New physics or limited understanding of proton structure?

#2: Indirect searches

PARTON DISTRIBUTION FUNCTIONS $f_i(x,\mu)$

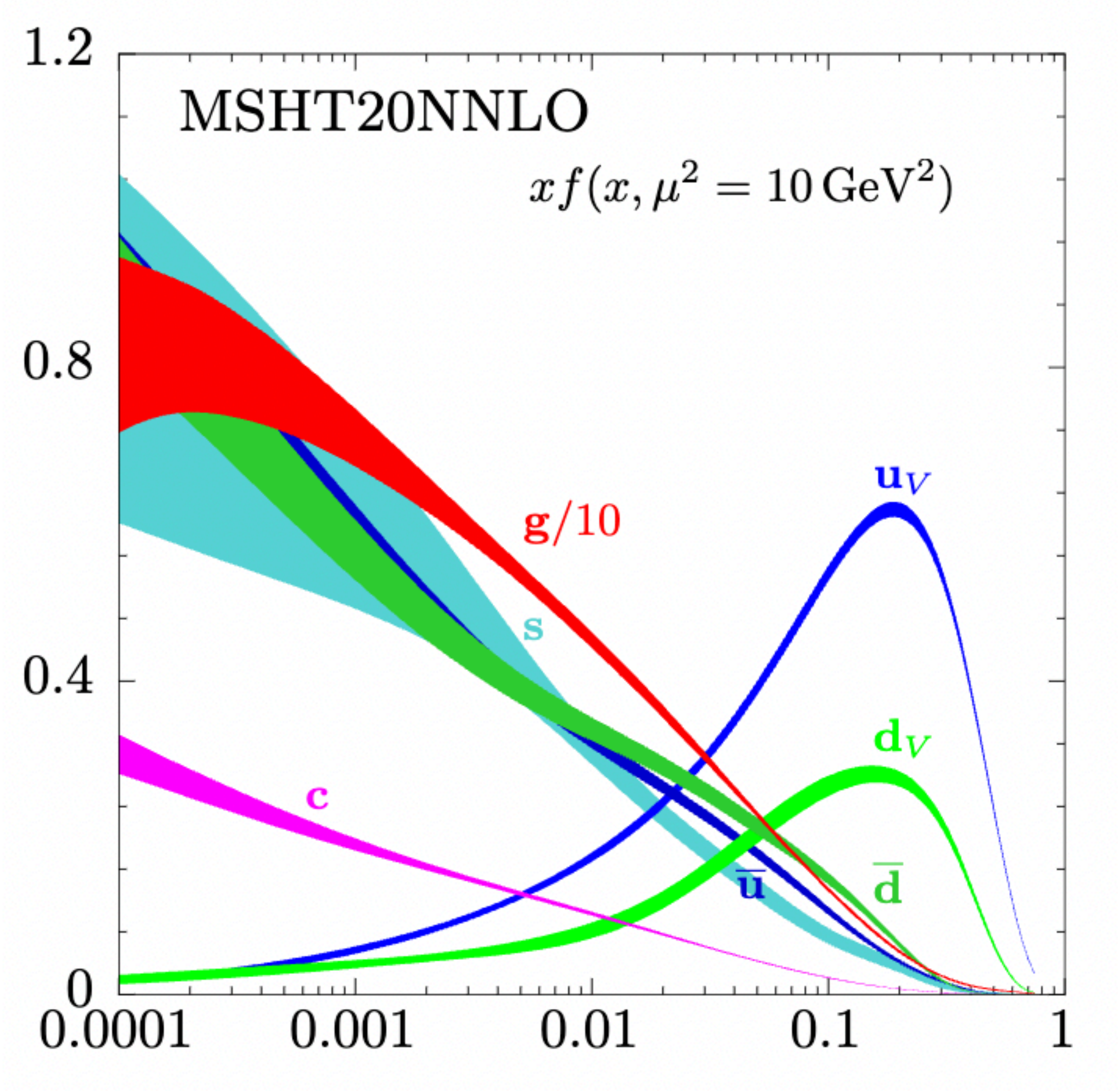
μ → Perturbative QCD

$$\frac{d}{dt}\begin{pmatrix} q_i(x,t) \\ g(x,t) \end{pmatrix} = \frac{\alpha_s(t)}{2\pi}\int_x^1 \sum_{j=q,\bar{q}} \frac{d\xi}{\xi} \begin{pmatrix} P_{ij}\left(\frac{x}{\xi},\alpha_s(t)\right) & P_{ig}\left(\frac{x}{\xi},\alpha_s(t)\right) \\ P_{gj}\left(\frac{x}{\xi},\alpha_s(t)\right) & P_{gg}\left(\frac{x}{\xi},\alpha_s(t)\right) \end{pmatrix} \otimes \begin{pmatrix} q_j(\xi,t) \\ g(\xi,t) \end{pmatrix}$$

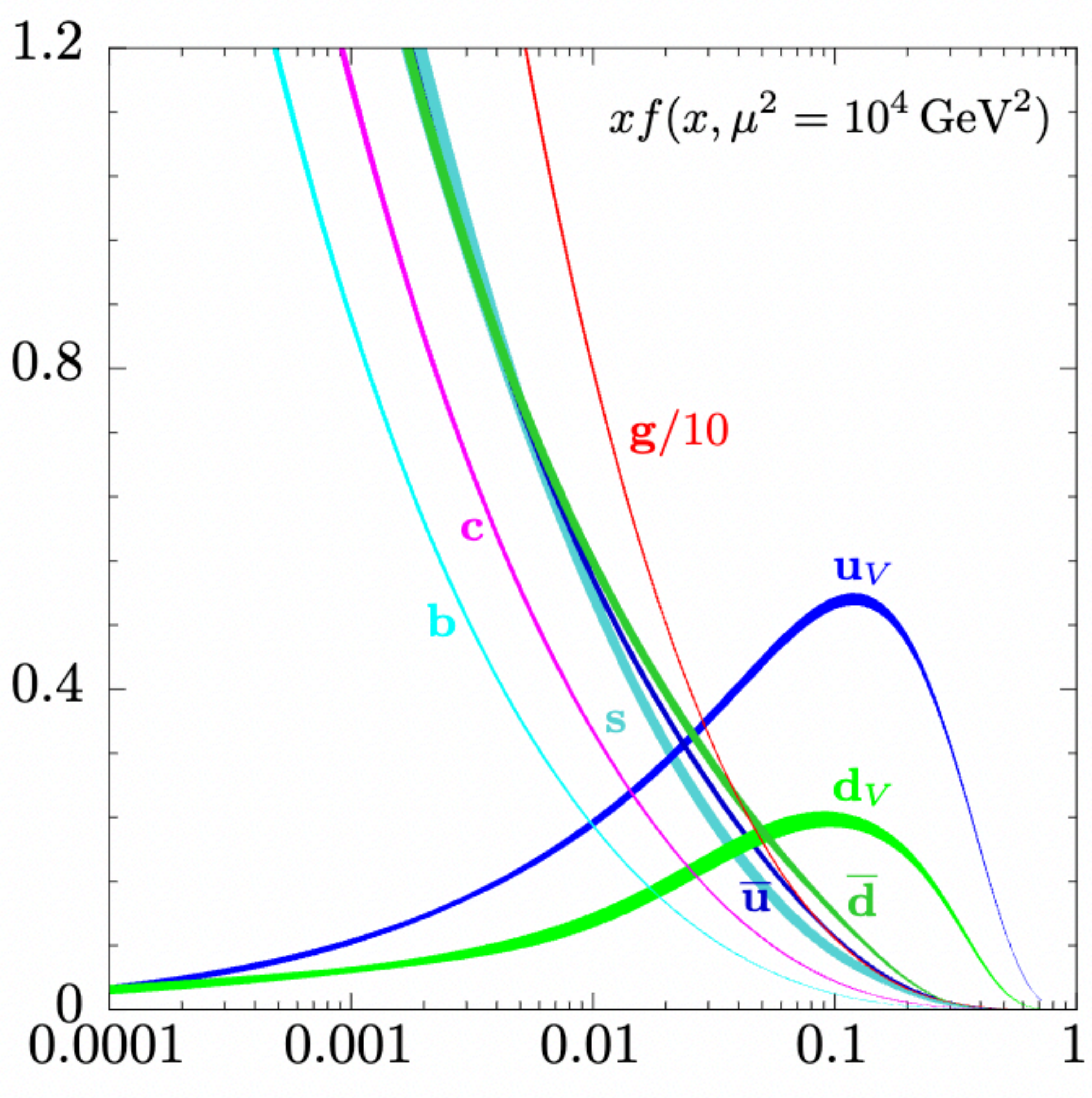
Splitting functions P_{ab} known up to approximate N3LO [Blumlein, Moch, Gehrman, von Manteufel, Sotnikov, Yang, Davies, Vogt, Bonvini, Marzani,…]

First approximate N3LO sets MSHT20aN3LO
McGowan et al 2207.04739
NNPDF40aN3LO
Ball et al 2402.18635

Hadronic scale: global fit of PDFs



DGLAP



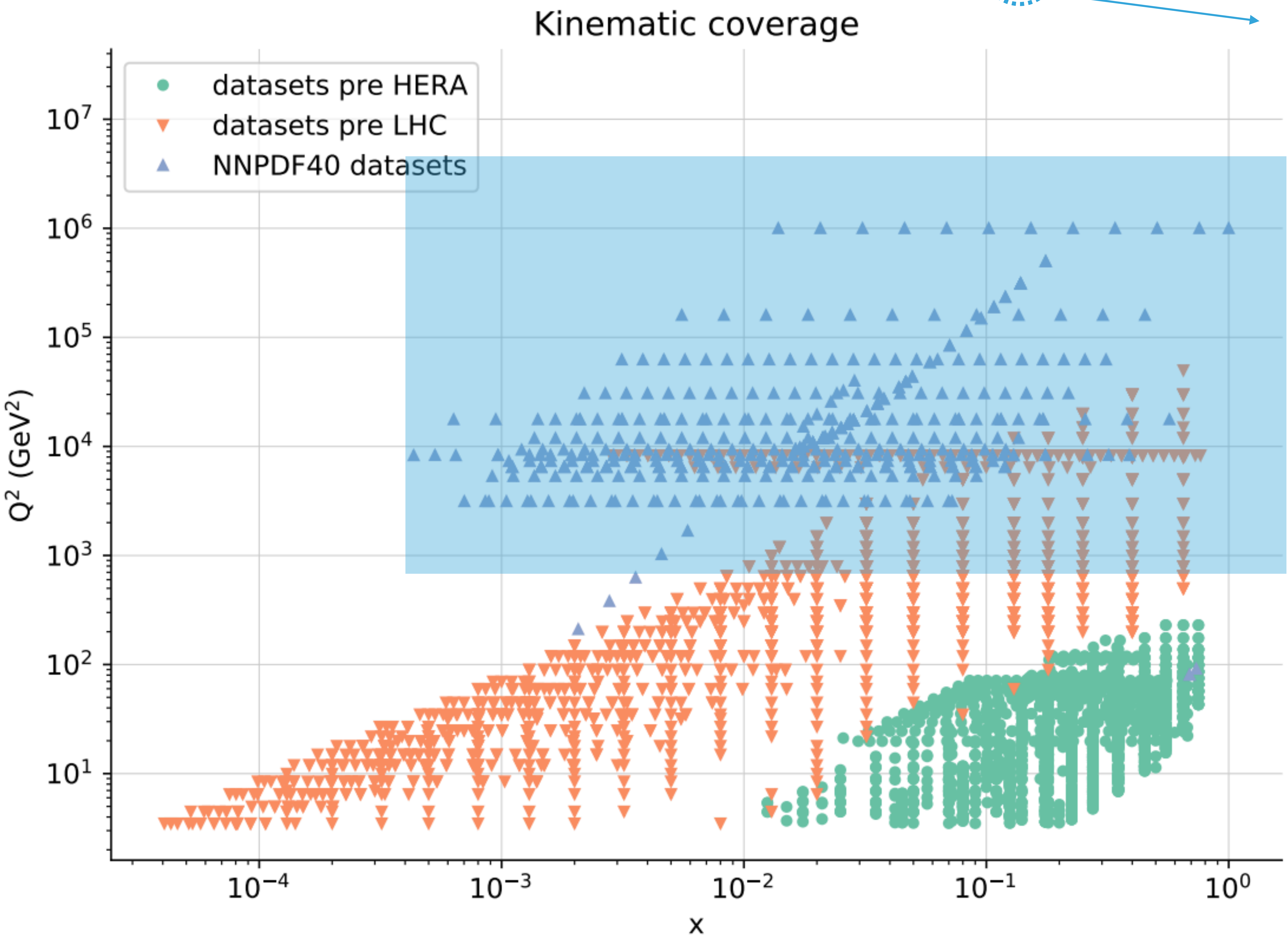
High scale: input to the LHC

PDG 2023

A WEALTH OF DATA FROM THE LHC

$$f_i(x, \mu)$$

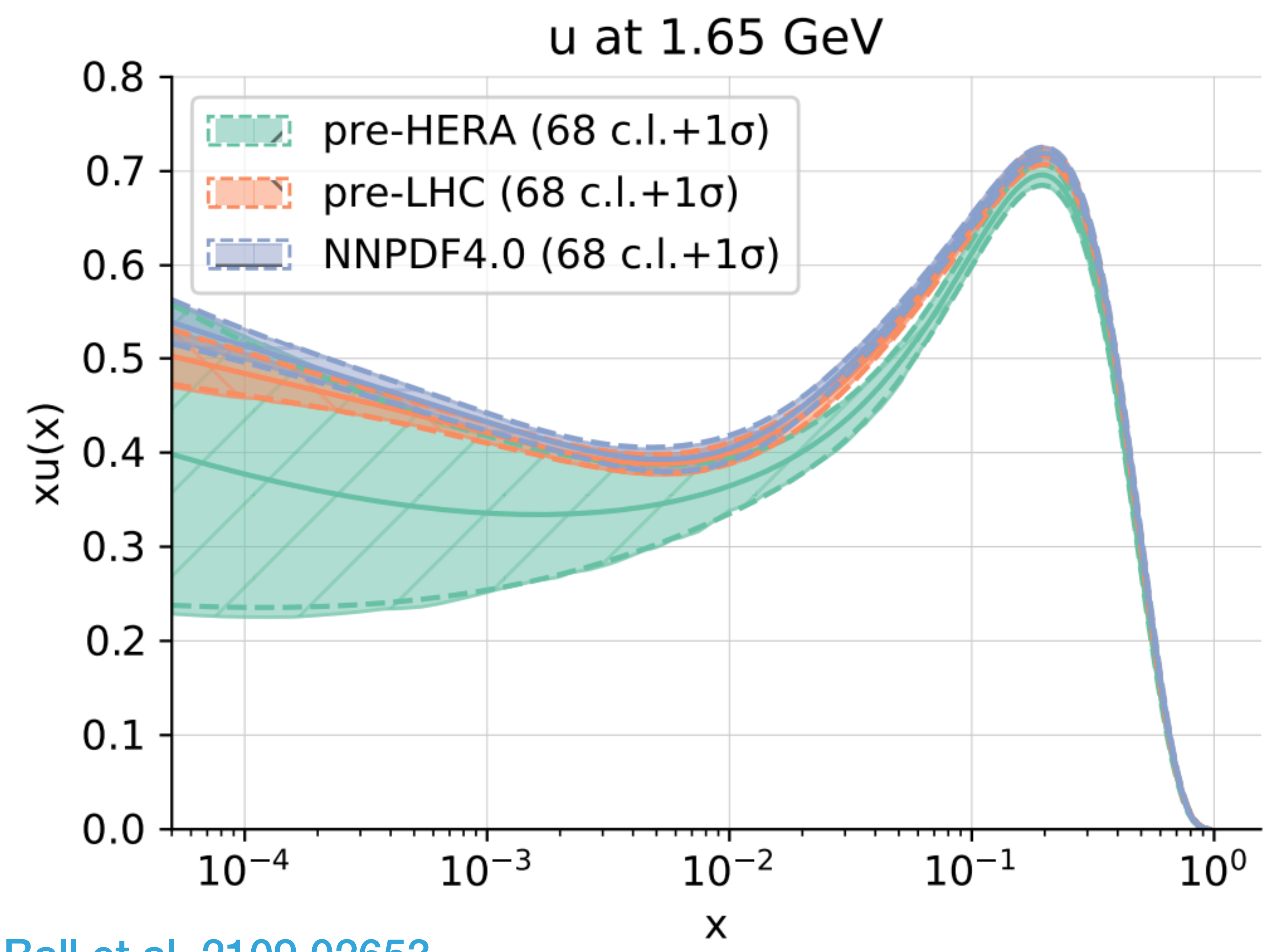
Data



NNPDF4.0: About 30% of input data are LHC data!

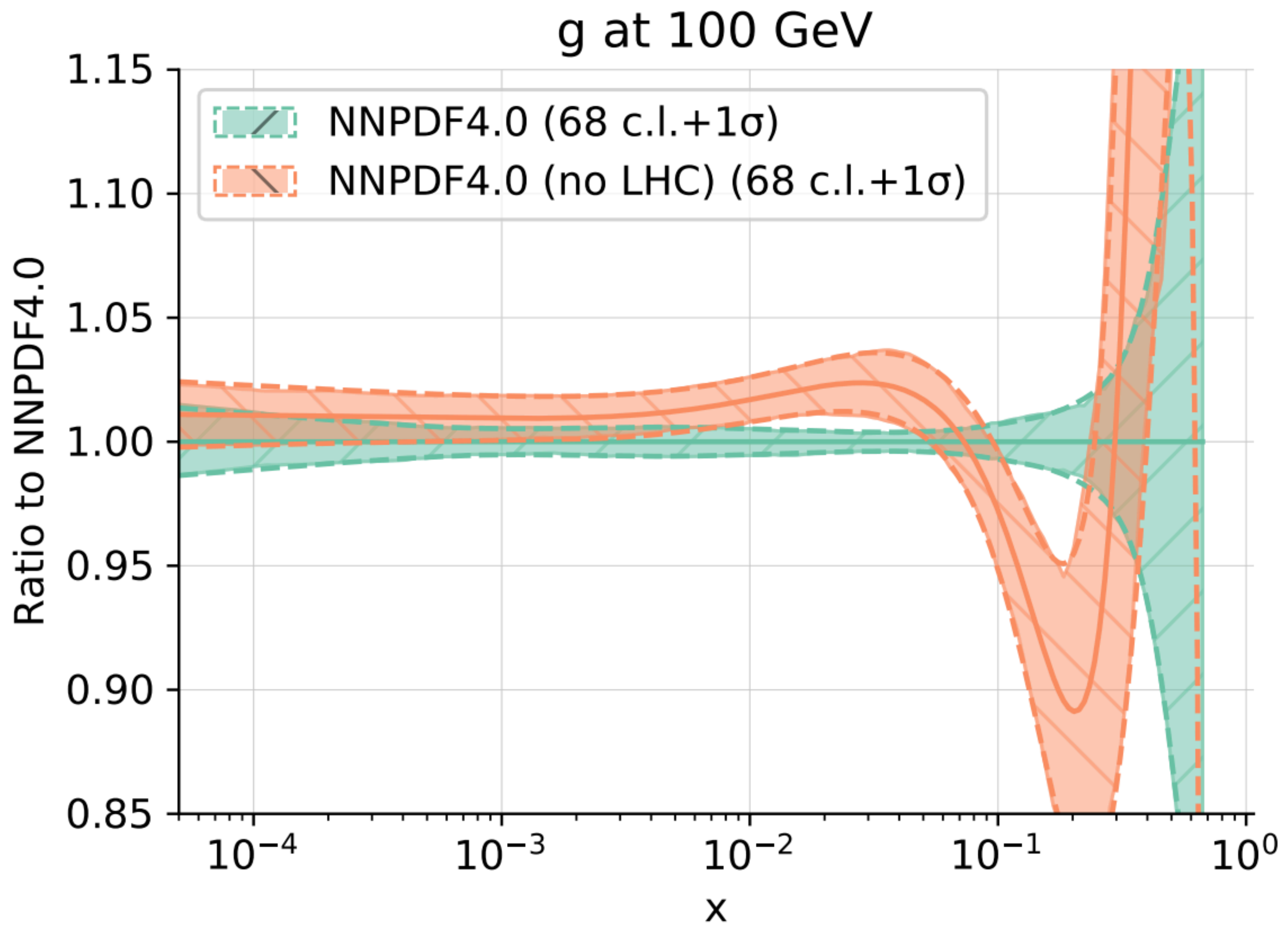
Ball et al, 2109.02653

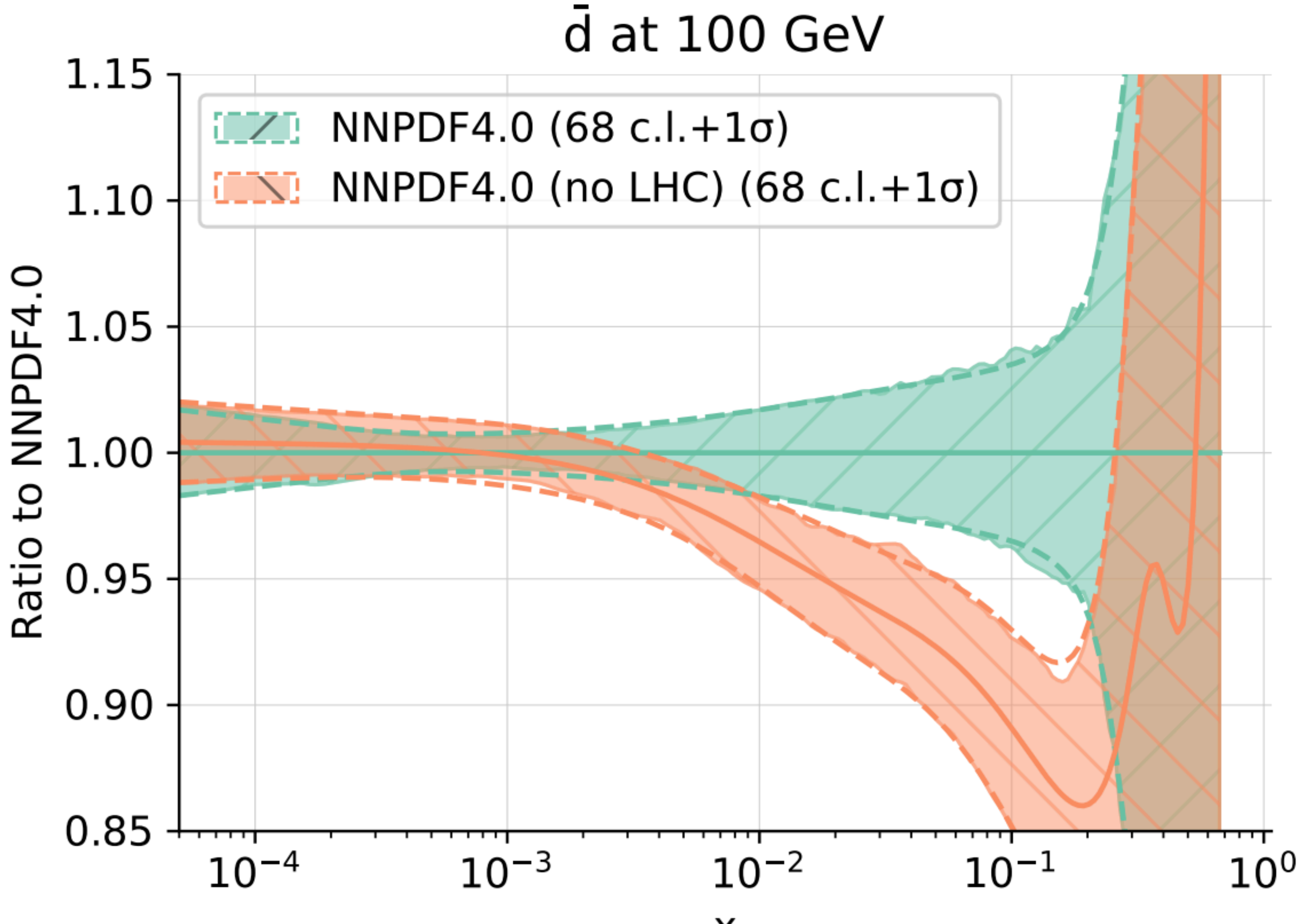
IMPACT OF LHC DATA



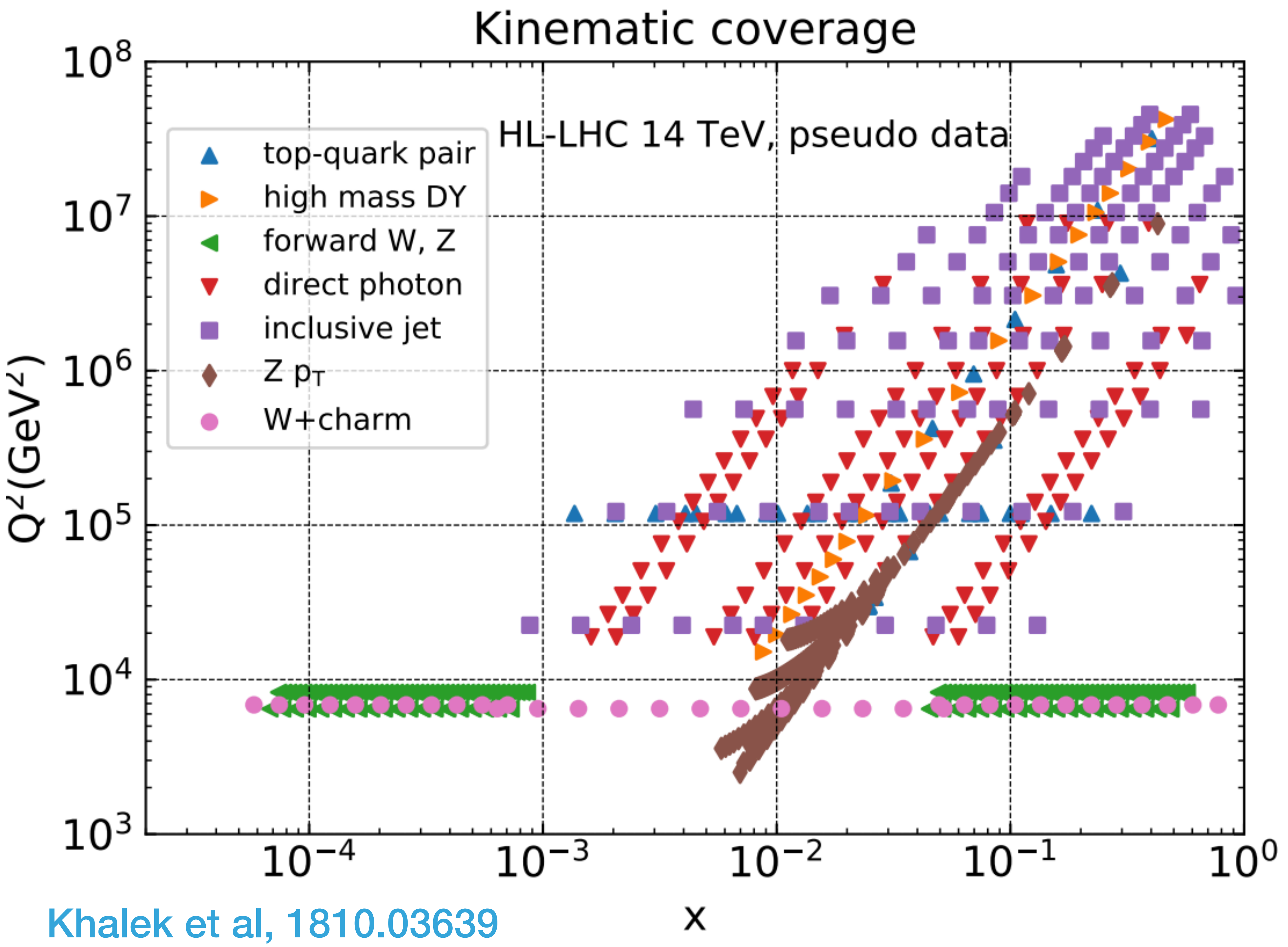
Ball et al, 2109.02653

- HERA data crucial to constrain quark valence (up and to less extent down valence) across intermediate to small x and gluon at small x
- LHC high energy data crucial to provide additional constraints to PDFs, in particular in medium- to large-x gluon and quarks.
- Mild tension with older fixed-target Drell-Yan and DIS data visible in the large-x region (especially gluon and anti-quarks)

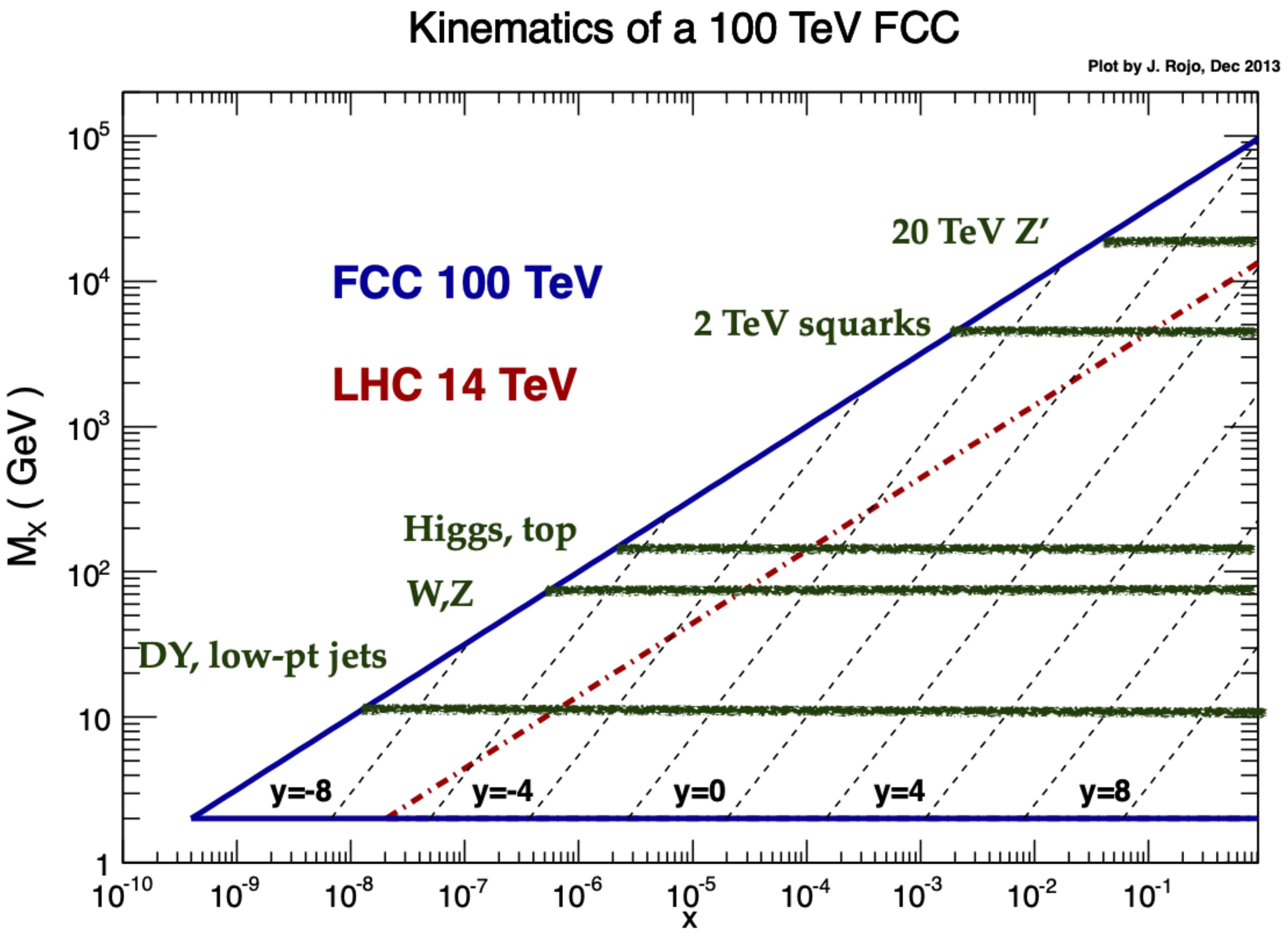




CONSTRAINTS FROM THE HL-LHC



Khalek et al, 1810.03639



M. Mangano et al 1607.01831

- Thanks to luminosity upgrade, HL-LHC will go nearly two orders of magnitude higher in Q2, populating the high-energy region, and this will allow to further constrain gluon and (anti)-quarks at large x [Khalek et al, 1810.03639]
- FCC-hh will go further up by two orders of magnitude in Q2 [Mangano et al 1607.01831]

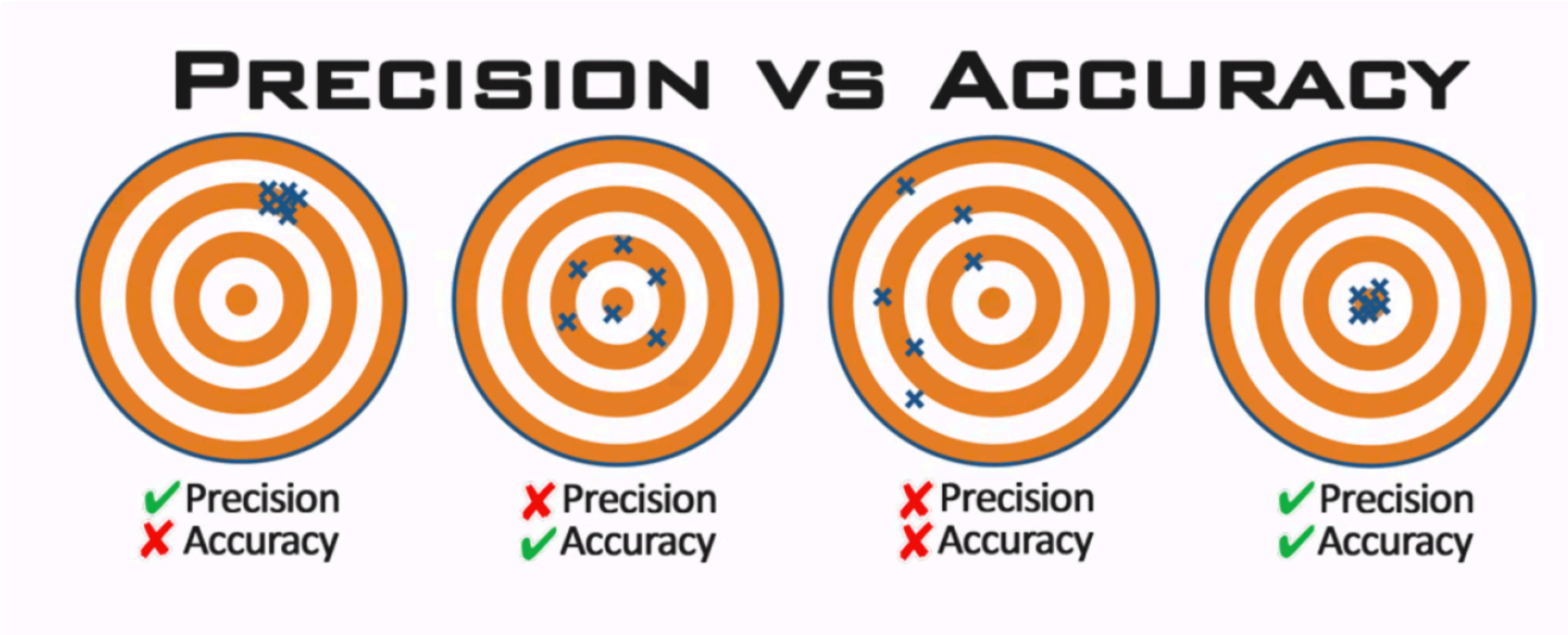
Large x ⟷ Large E and/or Large Y

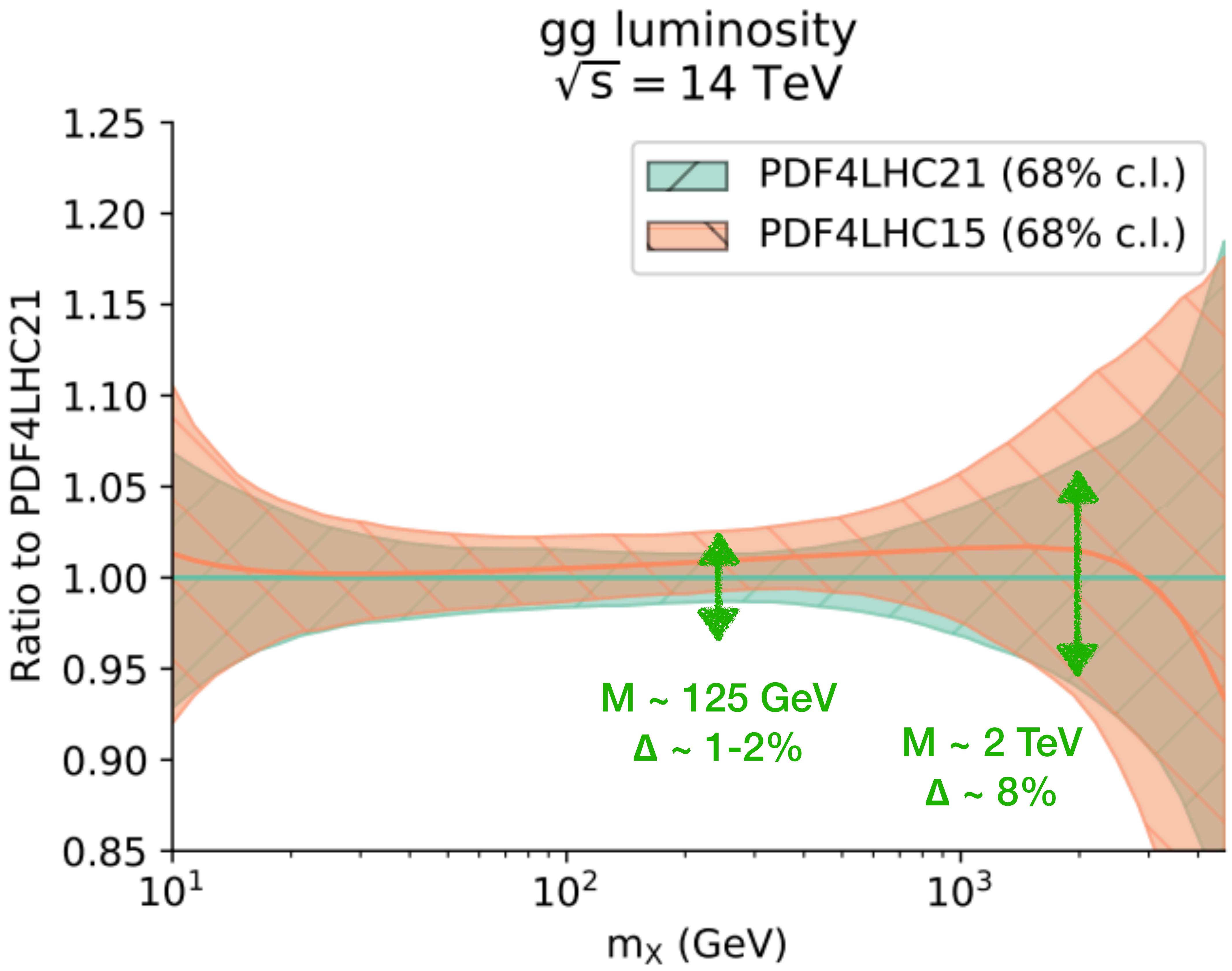
THE PRECISION VERSUS ACCURACY CHALLENGE

Now that PDFs are getting more and more precise, it is crucial for them to be accurate.

Challenges:

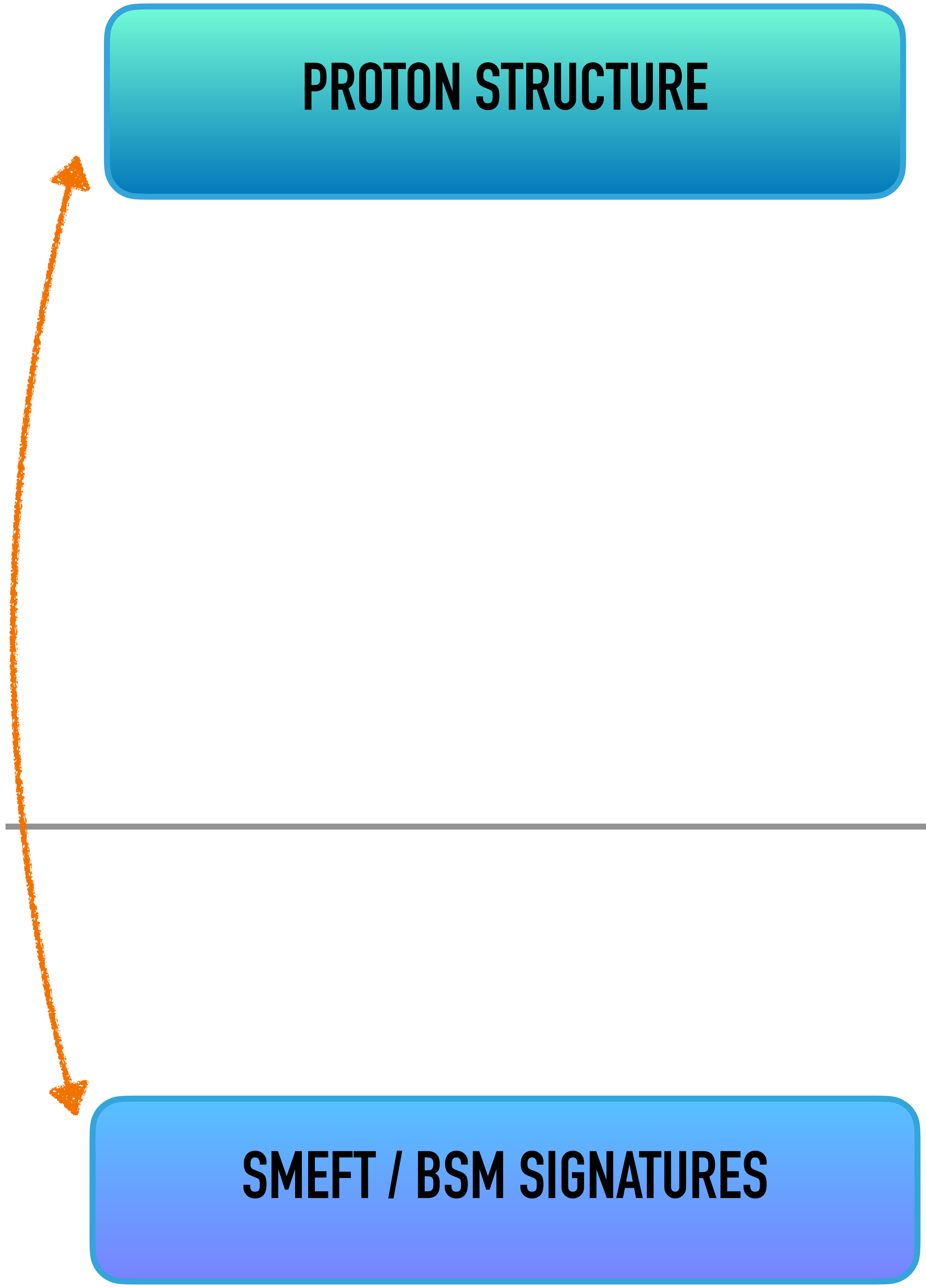
- Genuine inconsistencies in experimental inputs
- Imperfect fitting methodology
- Inaccurate theoretical framework
  - ➡ Missing higher order uncertainties & N3LO
  - ➡ Other corrections (nuclear, non-perturbative effects…)
  - ➡ BSM effects - second part of the talk





PDF4LHC21 study 2203.05506

AN UNEXPECTED INTERPLAY



SMEFT: A FRAMEWORK FOR INDIRECT DISCOVERIES

- EFT is a well-defined theoretical approach for indirect searches
- Assumption: new physics states are heavy
- Write the Lagrangian with only light SM particles, BSM effects incorporated as a momentum expansion
- SMEFT: assume SM field content and gauge symmetries (apart from accidental)

$$\mathcal{L}_{\mathrm{SMEFT}} = \mathcal{L}_{\mathrm{SM}} + \sum_i^{N_{d6}} \frac{c_i}{\Lambda^2} \mathcal{O}_i^{(6)} + \sum_j^{N_{d8}} \frac{b_j}{\Lambda^4} \mathcal{O}_j^{(8)} + \dots$$

- Full dim-6 basis of operators under SMEFT assumptions includes **2499** operators [Grzadkowski et al, arXiv:1008.4884]
- Current SMEFT fits make flavour assumptions and restricted to a few observables/sectors & reduce the number of operators.
- Huge progress in global dim-6 SMEFT fits [Ken Mimasu's talk]

EXTRACTING PARAMETERS FROM DATA

$$\chi^2 = \frac{1}{N_{\rm dat}} \sum_{i=1}^{N_{\rm dat}} (T_i(\{\theta\},\{c\}) - D_i)\,{\rm cov}^{-1}_{ij}\,(T_j(\{\theta\},\{c\}) - D_j)$$

$$T_i(\{\theta\},\{c\}) = {\rm PDFs}(\{\theta\},\{c\}) \otimes \hat{\sigma}_i(\{c\})$$

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i^{N_{d6}} \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_j^{N_{d8}} \frac{b_j}{\Lambda^4}\mathcal{O}_j^{(8)} + \dots$$

SMEFT WCs

Parameters determining PDFs at initial scale

EXTRACTING PARAMETERS FROM DATA

$$\chi^2 = \frac{1}{N_{\rm dat}} \sum_{i=1}^{N_{\rm dat}} (T_i(\{\theta\},\{c\}) - D_i)\, {\rm cov}_{ij}^{-1}\, (T_j(\{\theta\},\{c\}) - D_j)$$

$$T_i(\{\theta\},\{c\}) = {\rm PDFs}(\{\theta\},\{c\}) \otimes \hat{\sigma}_i(\{c\})$$

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i^{N_{d6}} \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_j^{N_{d8}} \frac{b_j}{\Lambda^4}\mathcal{O}_j^{(8)} + \dots$$

SMEFT WCs

Parameters determining PDFs at initial scale

✓ In a PDF fit typically

$$T_i(\{\theta\}) = {\rm PDFs}(\{\theta\},\{c=0\}) \otimes \hat{\sigma}_i(\{c=0\})$$

EXTRACTING PARAMETERS FROM DATA

$$\chi^2 = \frac{1}{N_{\rm dat}} \sum_{i=1}^{N_{\rm dat}} (T_i(\{\theta\},\{c\}) - D_i)\, {\rm cov}^{-1}_{ij}\, (T_j(\{\theta\},\{c\}) - D_j)$$

$$T_i(\{\theta\},\{c\}) = {\rm PDFs}(\{\theta\},\{c\}) \otimes \hat{\sigma}_i(\{c\})$$

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i^{N_{d6}} \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_j^{N_{d8}} \frac{b_j}{\Lambda^4}\mathcal{O}_j^{(8)} + \ldots$$

SMEFT WCs

Parameters determining PDFs at initial scale

✓ In a PDF fit typically

$$T_i(\{\theta\}) = {\rm PDFs}(\{\theta\},\{c=0\}) \otimes \hat{\sigma}_i(\{c=0\})$$

✓ In a fit of SMEFT Wilson Coefficients

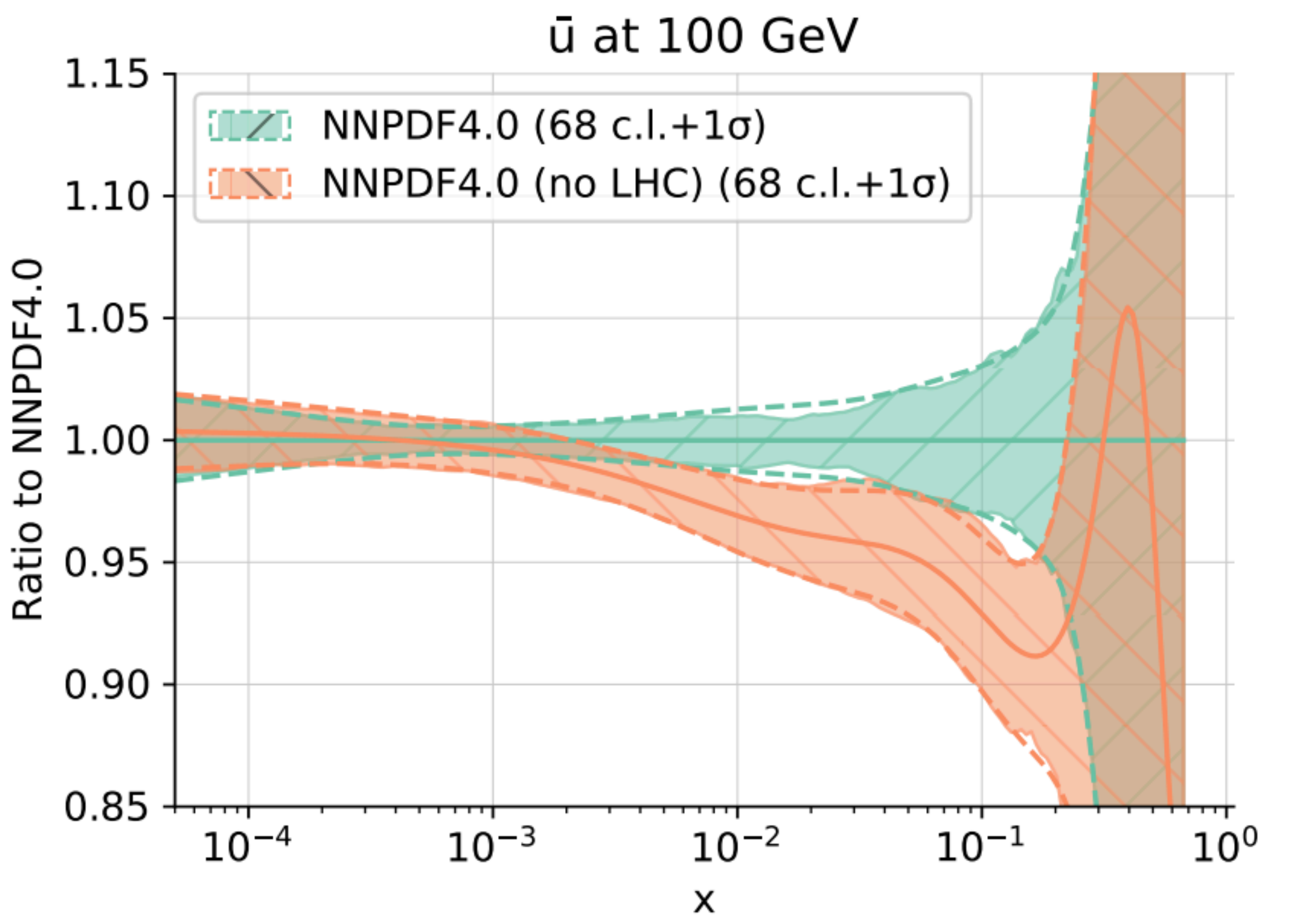
$$T_i(\{c\}) = {\rm PDFs}(\{\theta=\bar{\theta}\},\{c=0\}) \otimes \hat{\sigma}_i(\{c\})$$

PDF AND SMEFT INTERPLAY

- PDFs are low-scale quantities extracted from experimental data at all scales, without considering any potential high-scale contamination due to new physics.

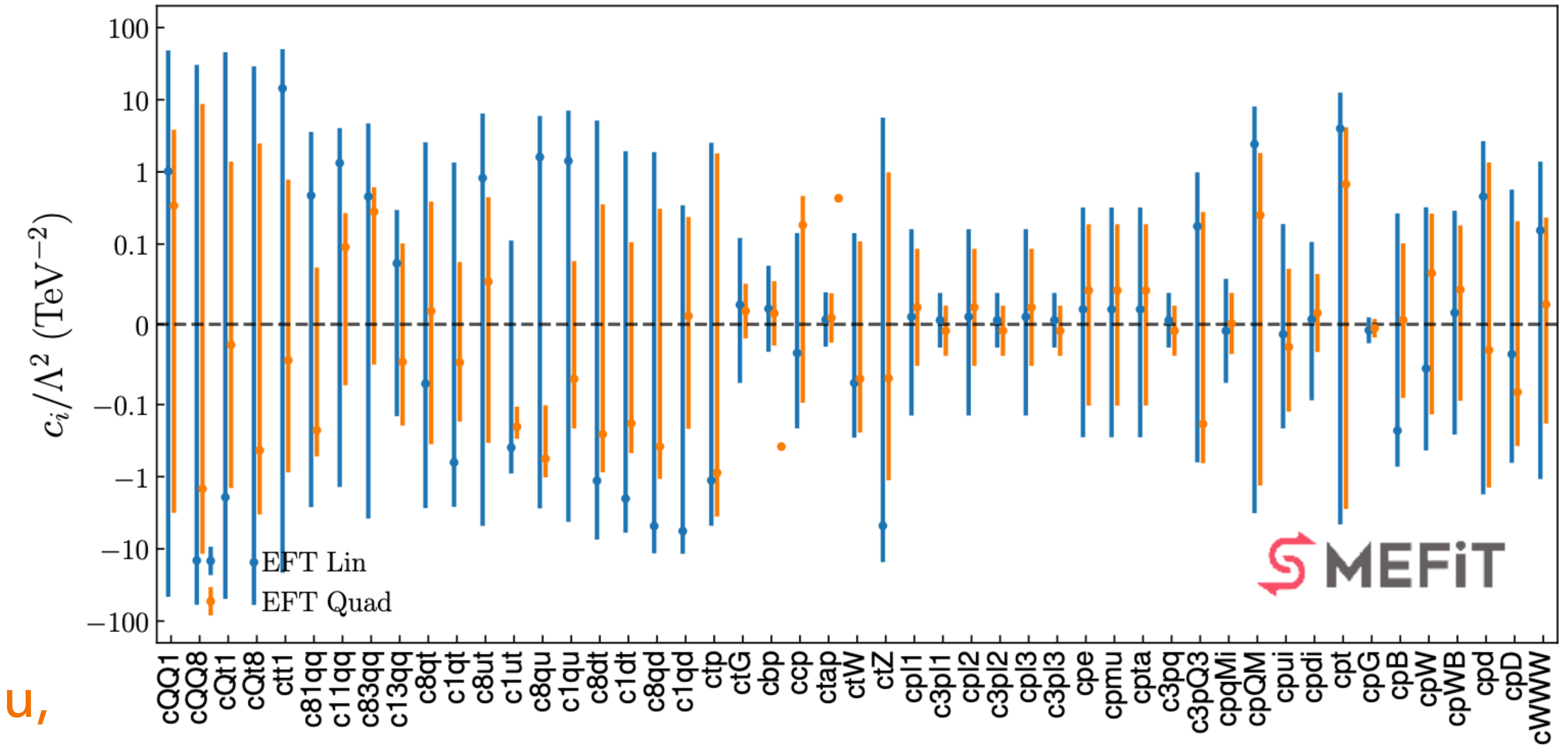
- (SM)EFT fits are performed by assuming a priori that PDFs are SM-like.

- In principle low-scale physics is separable from high-scale physics, BUT the complexity of LHC environment might well intertwine them



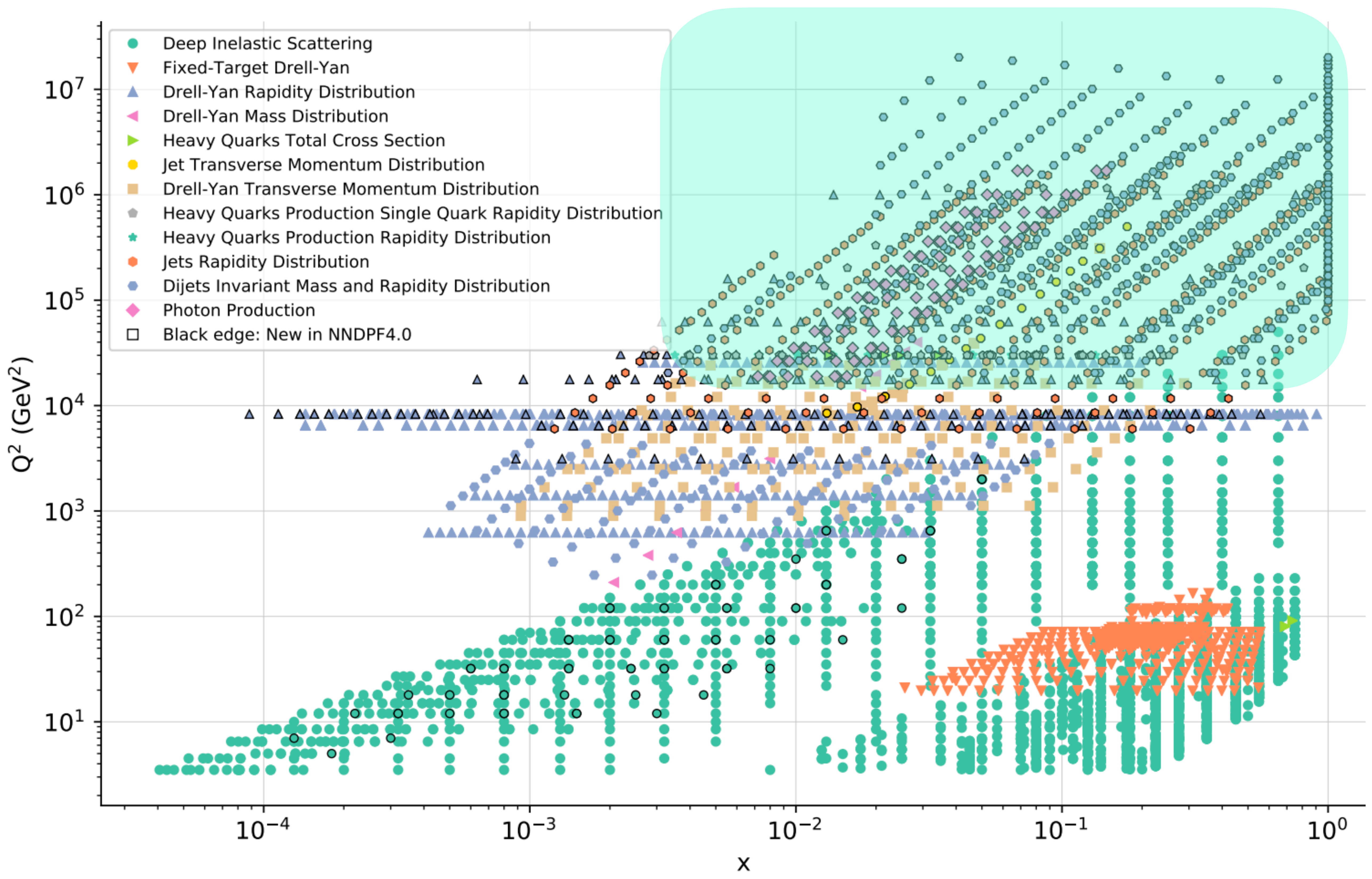
Ball et al, arXiv:2109.02653

MU, M. Costantini, E. Hammou, M. Madigan, L. Mantani, J. Moore, M. Morales Alvarado



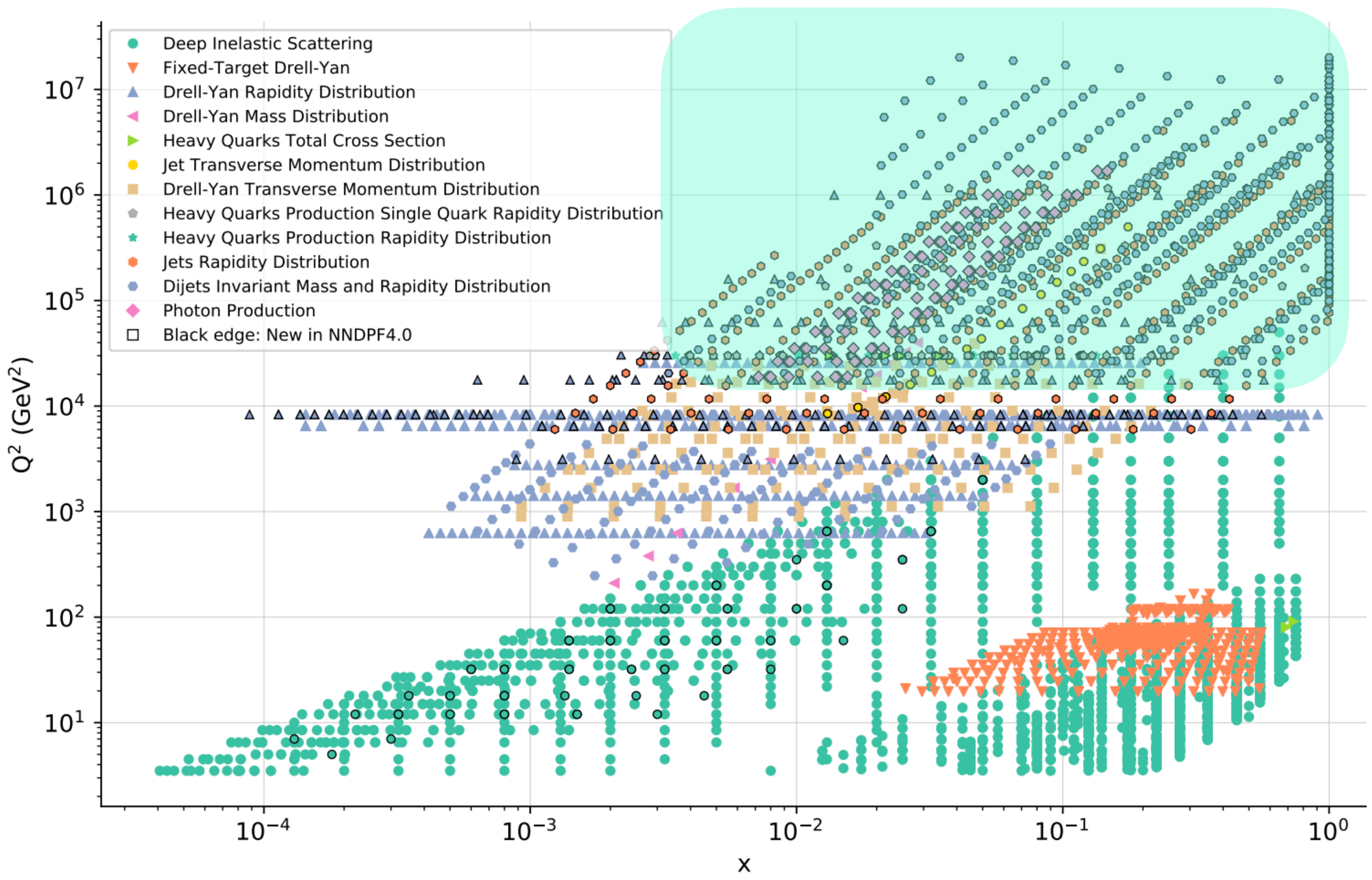
Ethier et al, arXiv: 2105.00006

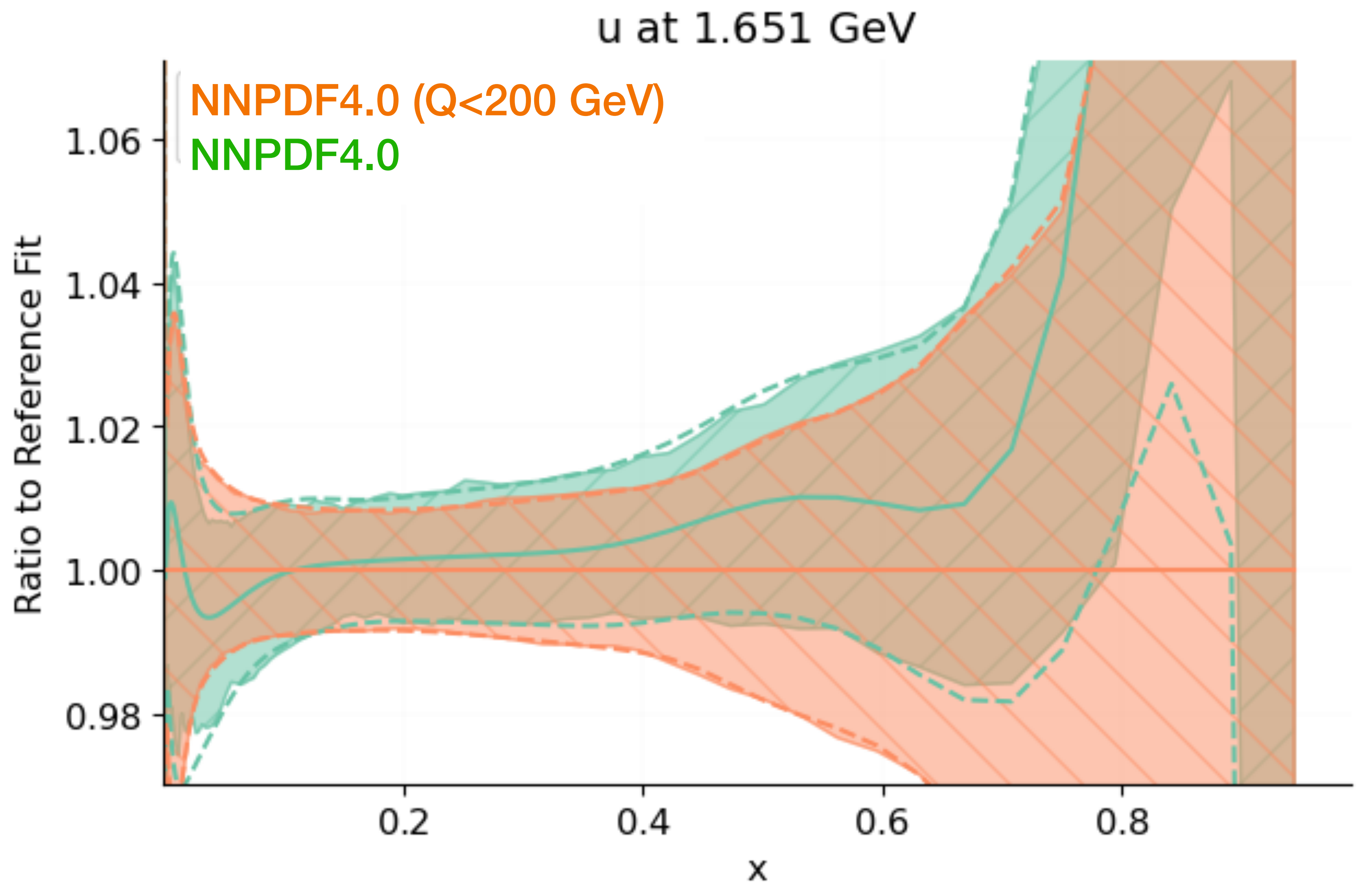
PDF AND SMEFT INTERPLAY

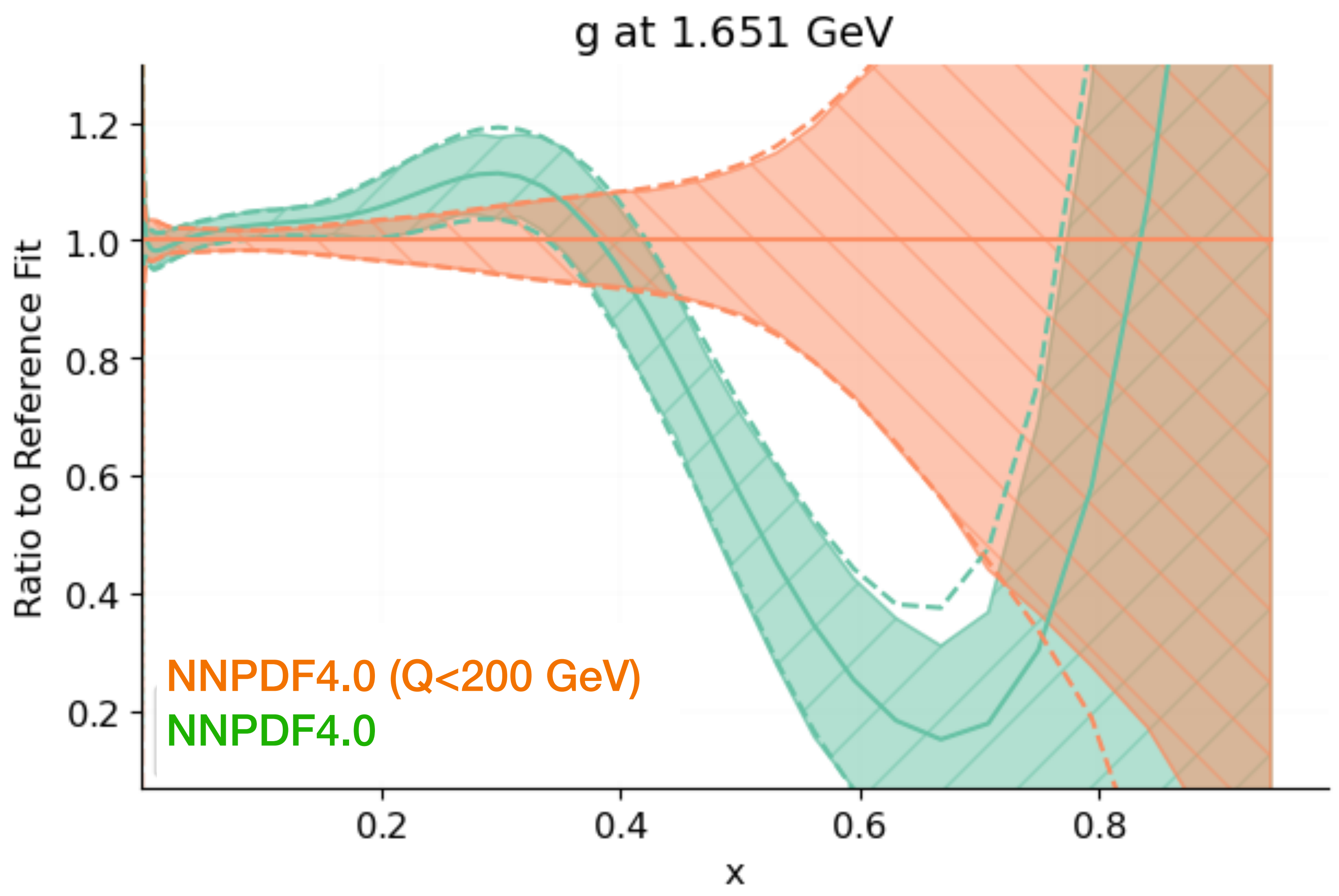


- ➡ Top pair production and single top data included in SMEFT analysis [Hartland et al 1901.05965] [Ellis et al 2012.02779]
- ➡ Drell-Yan high mass data [Farina et al 1609.08157, Torre et al 2008.12978]
- ➡ Inclusive Jets data [Alte et al 1711.07484]
- ➡ Dijets data [Krauss et al 1611.00767] [Alioli et al 1706.03068] [Bordone et al 2103.10332]
- ➡ Overlap enhanced in HL-LHC projections [Abdul Khalek et al1810.03639]

PDF AND SMEFT INTERPLAY

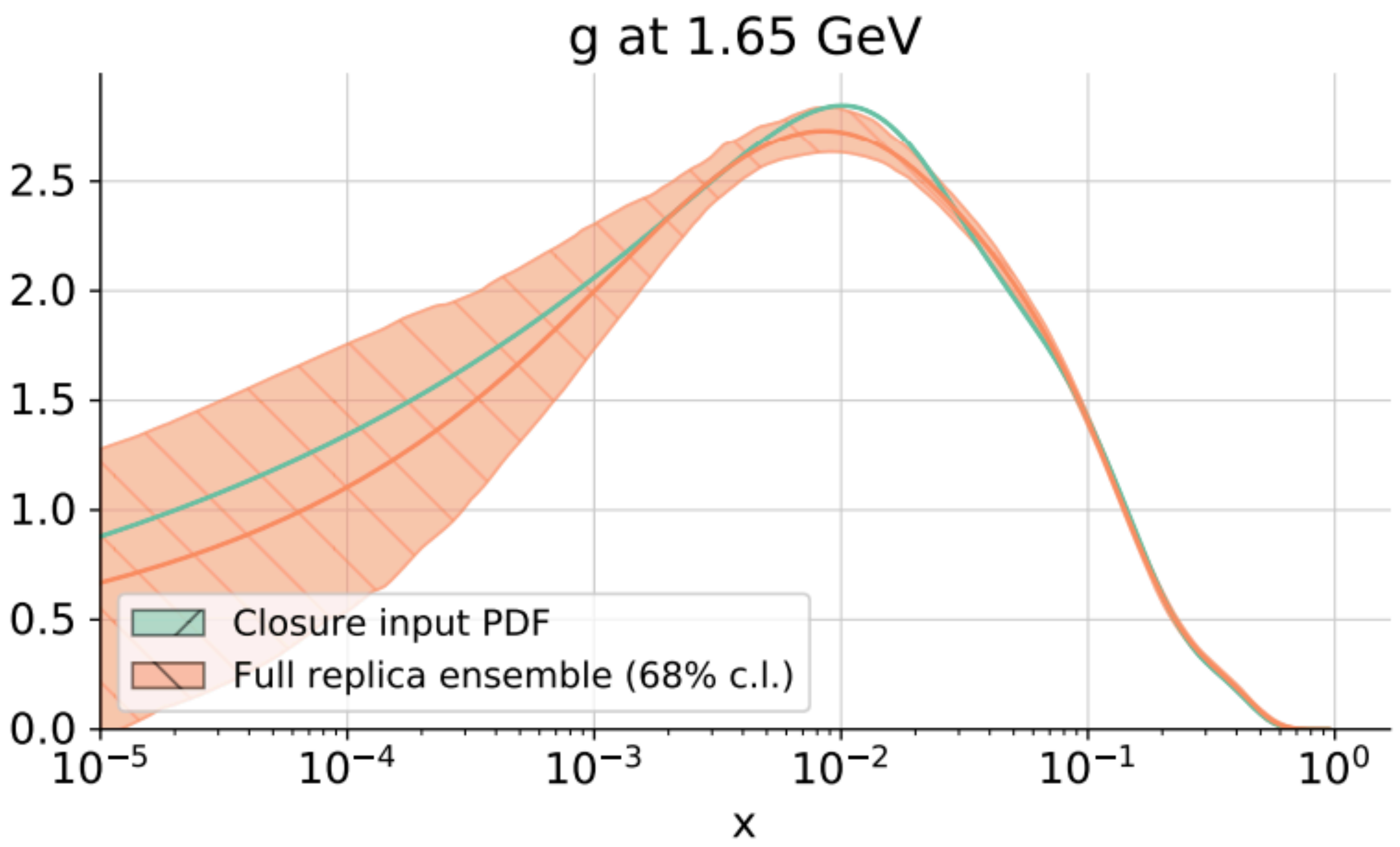






- Naive solution: remove all data above $Q > Q_{high}$
- Cutting away all high energy data would result in dramatic uncertainty increase and loss of precision!

CAN PDFS ABSORB NEW PHYSICS?



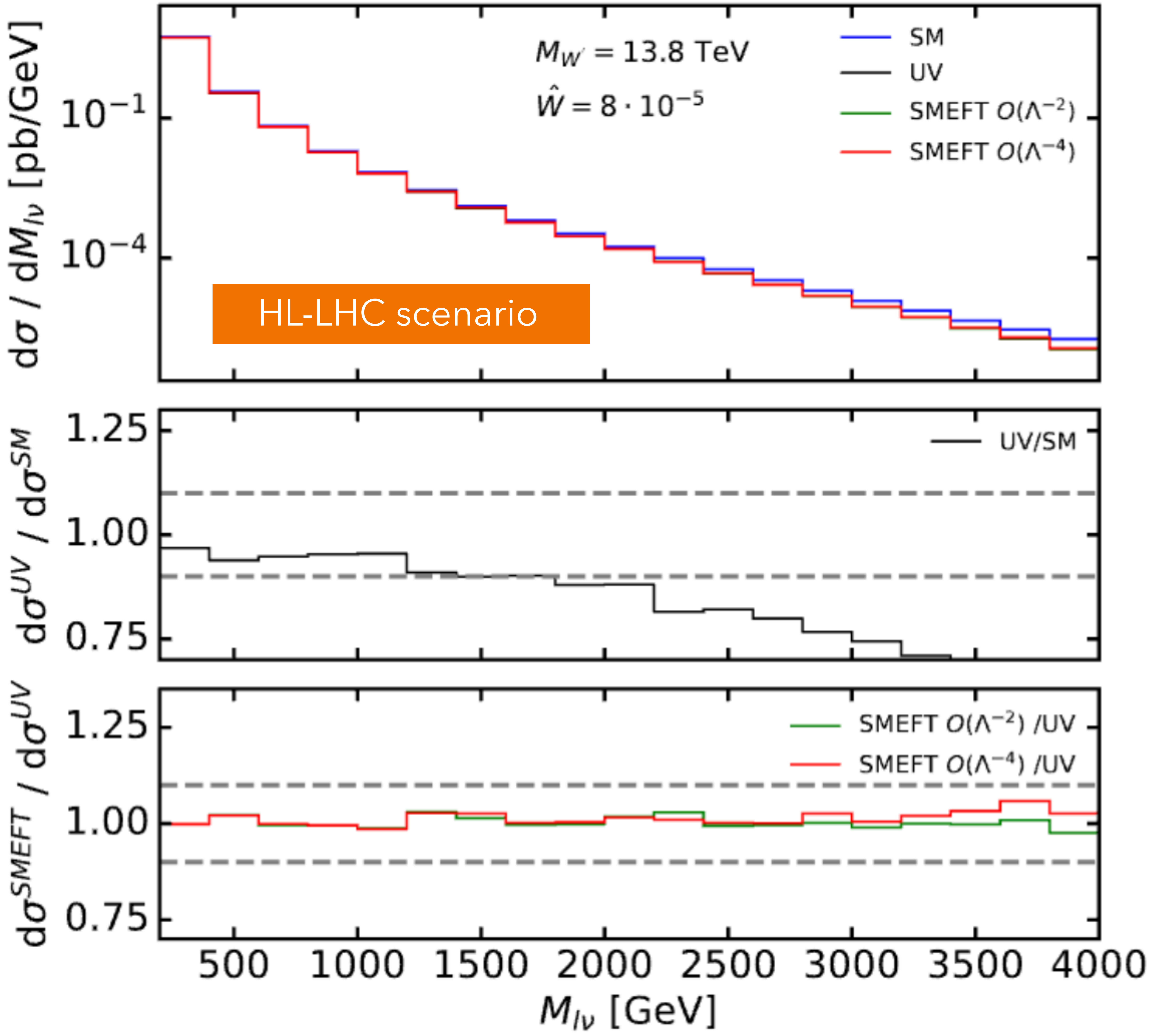
✓ NNPDF methodology routinely tested via closure test (in the data region) [Del Debbio, Giani, Wilson, 2111.05787] and future test (in the extrapolation region) [Cruz-Martinez, Forte, Nocera, 2103.08606].

✓ Closure tests assess methodology robustness and efficiency & faithfulness of uncertainty estimate.

✓ Input the "true" PDFs, generate MC data according to the "truth" with exp. uncertainty and check if what you get out of the fit corresponds to the truth

✓ **Can we test if we can "safely" add high M data using this statistical test?**

CAN PDFS ABSORB NEW PHYSICS?



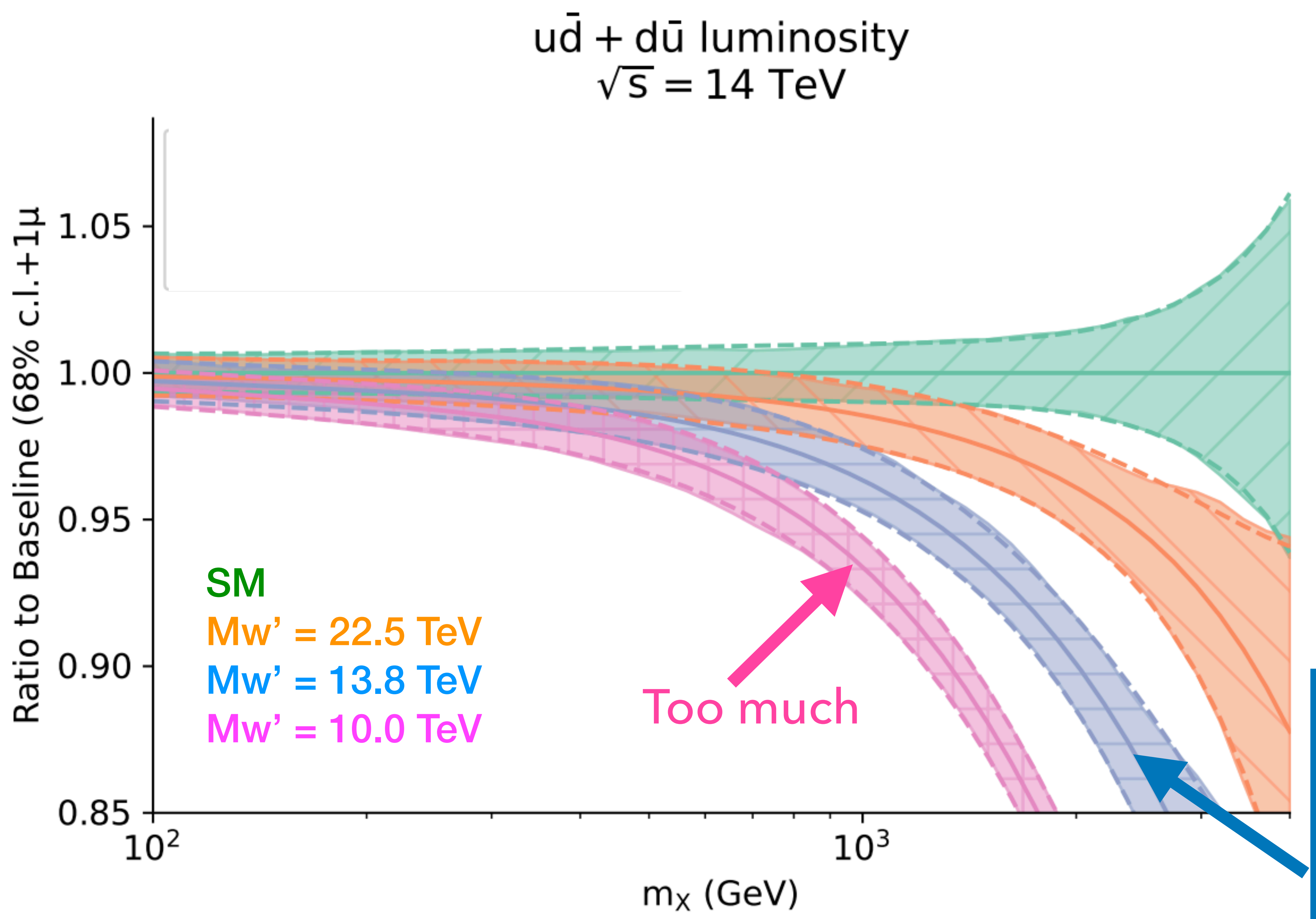
- Imagine that on top of the "true" PDFs one inject the "true" BSM model in the MC data
- Generate artificial MC data assuming "true" law of nature =
  "true" PDFs + "true" BSM model
- Fit PDFs assuming SM
- Can PDFs absorb signs of new physics?

E. Hammou, Z. Kassabov, M. Madigan,
M. Mangano, L. Mantani, J. Moore, M. Morales, MU
2307.10370

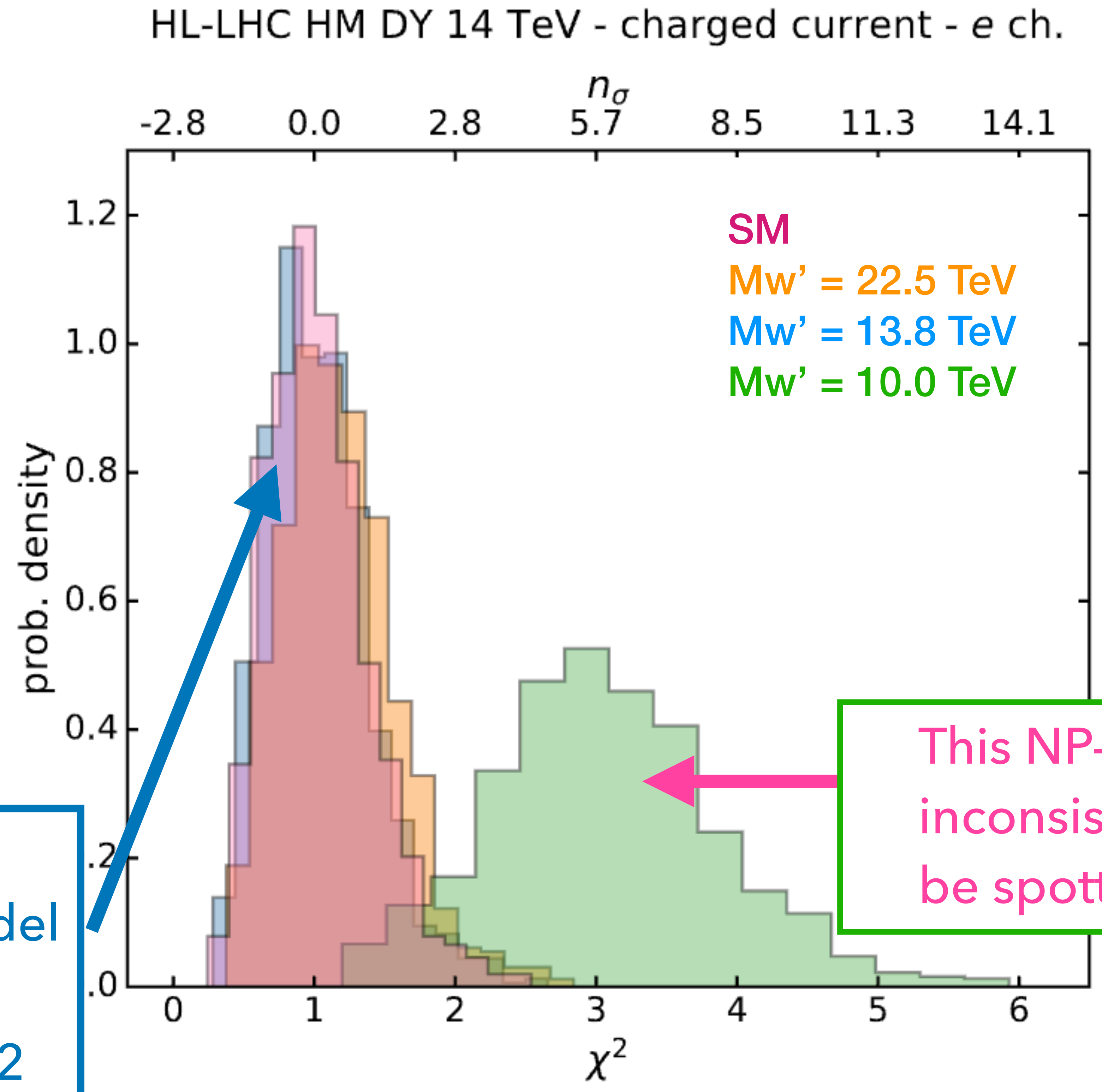
Z'	W'
Selection test: ✗	Selection test: ✓
➡ Excluded from PDF fit	➡ Included in PDF fit
No impact on PDFs	**PDFs contaminated**

CAN PDFS ABSORB NEW PHYSICS? THE W' CASE AT HL-LHC

- The fit-quality of the global fit is unchanged even with signal from Mw' = 13.8 TeV injected in all data (mostly visible in HL-LHC NC and CC Drell-Yan data)
- Once we go beyond this point ,the fit-quality deteriorates due to the HL-LHC neutral current and charged current Drell-Yan MC data.
- Already for Mw' = 13.8 TeV the qq~ luminosity shifts far beyond the PDF uncertainties because anti-quark PDFs at large-x compensate or "fit away" the effect of New Physics and we would not know in a real fit.



Max contamination allowed for this model by global fit without spoiling chi2



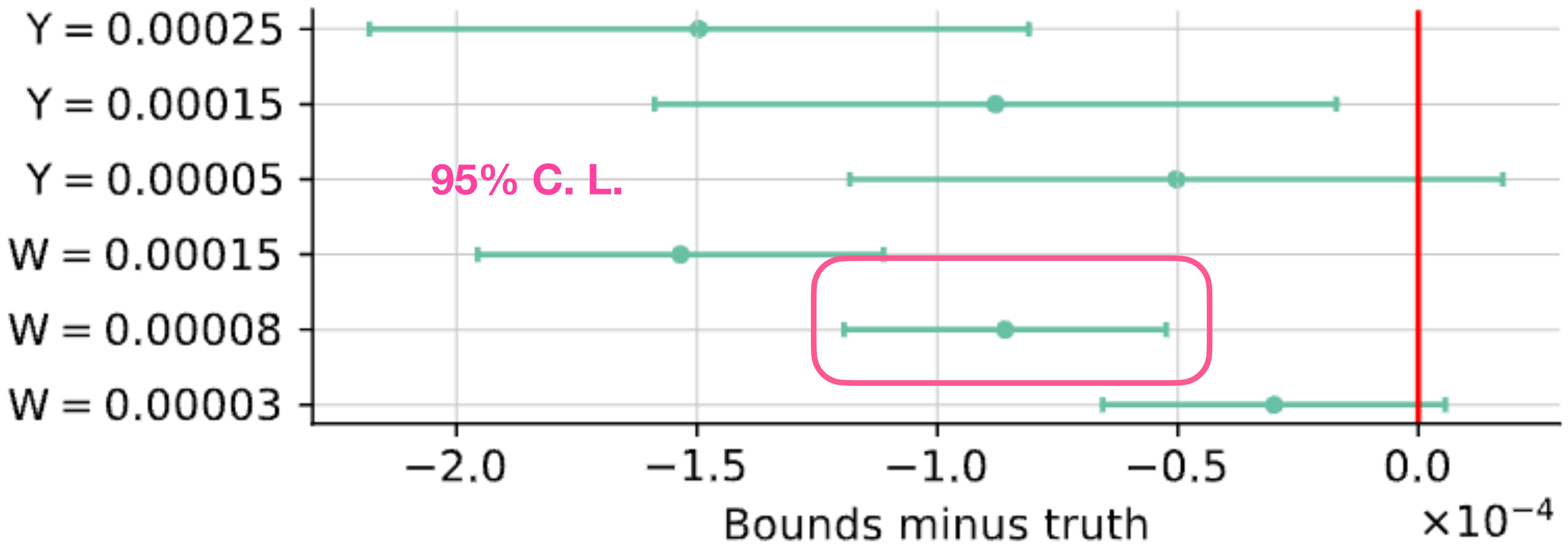
This NP-induced inconsistency would be spotted

E. Hammou, et al 2307.10370

CAN PDFS ABSORB NEW PHYSICS? THE W' CASE AT HL-LHC

- The fit-quality of the global fit is unchanged even with signal from Mw′ = 13.8 TeV injected in all data (mostly visible in HL-LHC NC and CC Drell-Yan data)
- Once we go beyond this point ,the fit-quality deteriorates due to the HL-LHC neutral current and charged current Drell-Yan MC data.
- Already for Mw′ = 13.8 TeV the qq~ luminosity shifts far beyond the PDF uncertainties because anti-quark PDFs at large-x compensate or "fit away" the effect of New Physics and we would not know in a real fit.

Consequence #1:
Would not see indirect effect of new physics as we would find SMEFT bounds compatible with the SM!

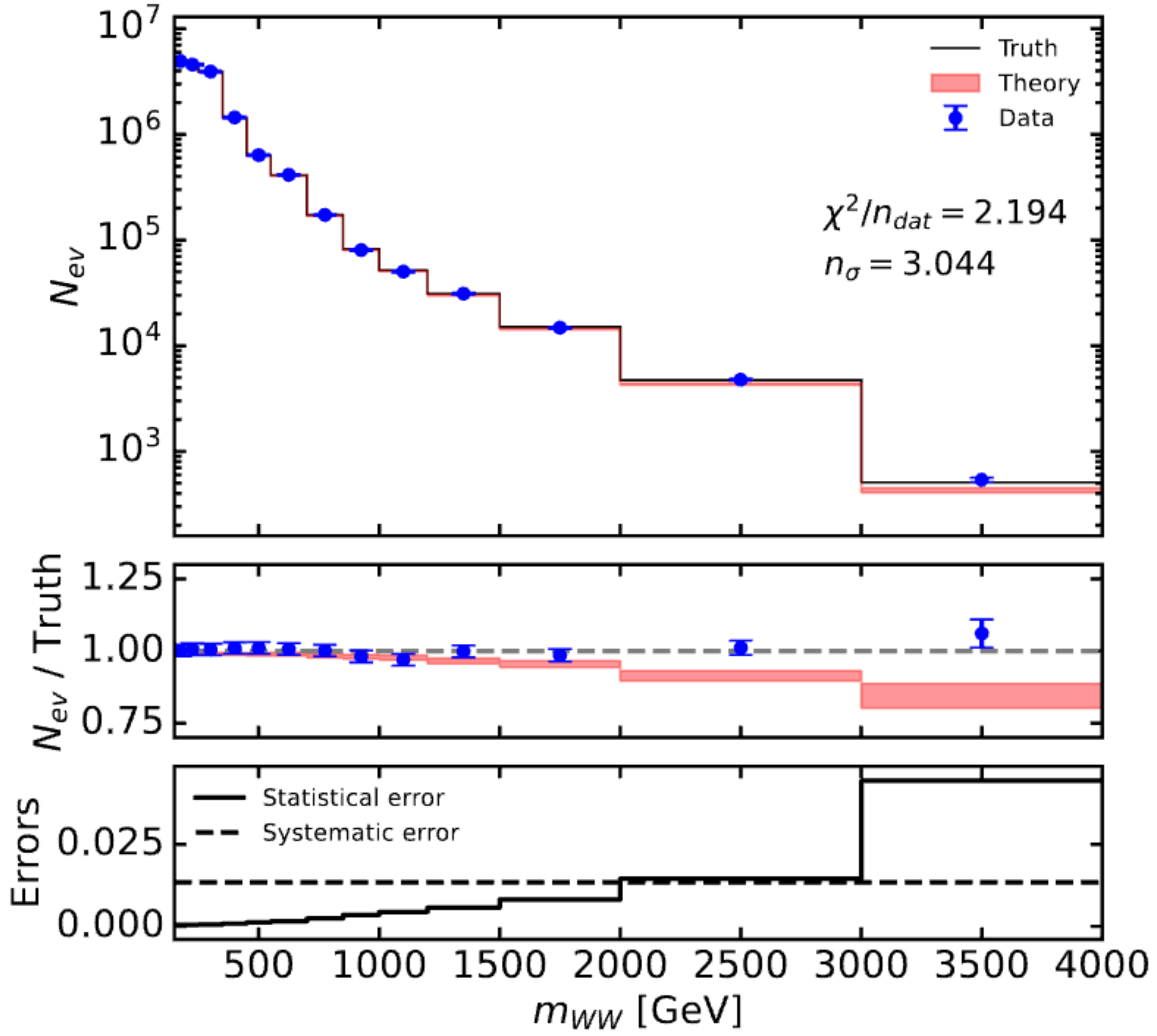


E. Hammou, et al 2307.10370

CAN PDFS ABSORB NEW PHYSICS? THE W' CASE AT HL-LHC

- The fit-quality of the global fit is unchanged even with signal from Mw' = 13.8 TeV injected in all data (mostly visible in HL-LHC NC and CC Drell-Yan data)
- Once we go beyond this point ,the fit-quality deteriorates due to the HL-LHC neutral current and charged current Drell-Yan MC data.
- Already for Mw' = 13.8 TeV the qq~ luminosity shifts far beyond the PDF uncertainties because anti-quark PDFs at large-x compensate or "fit away" the effect of New Physics and we would not know in a real fit.

Consequence #2:
Would see New Physics effects where there are none (for example in WW)

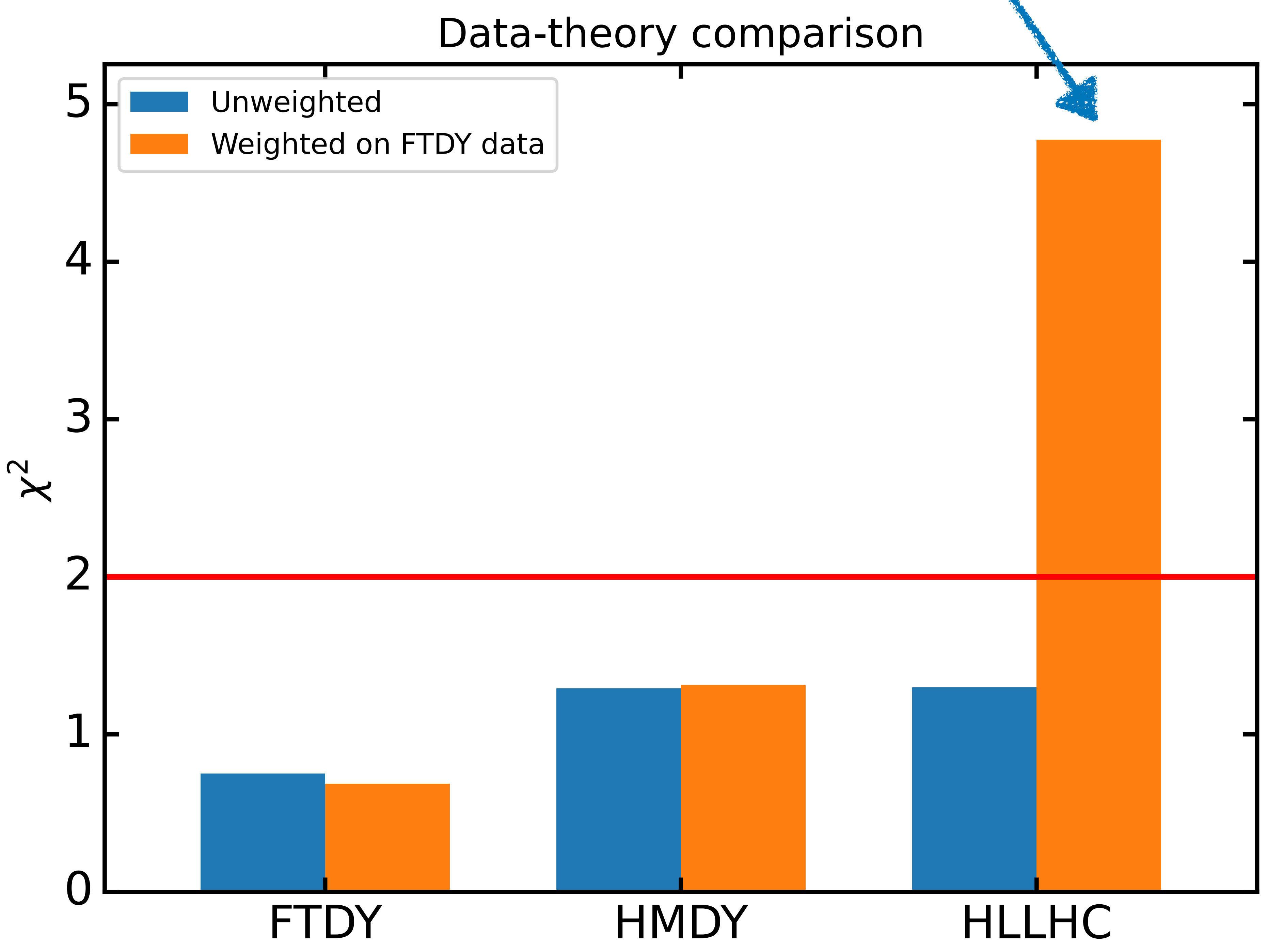


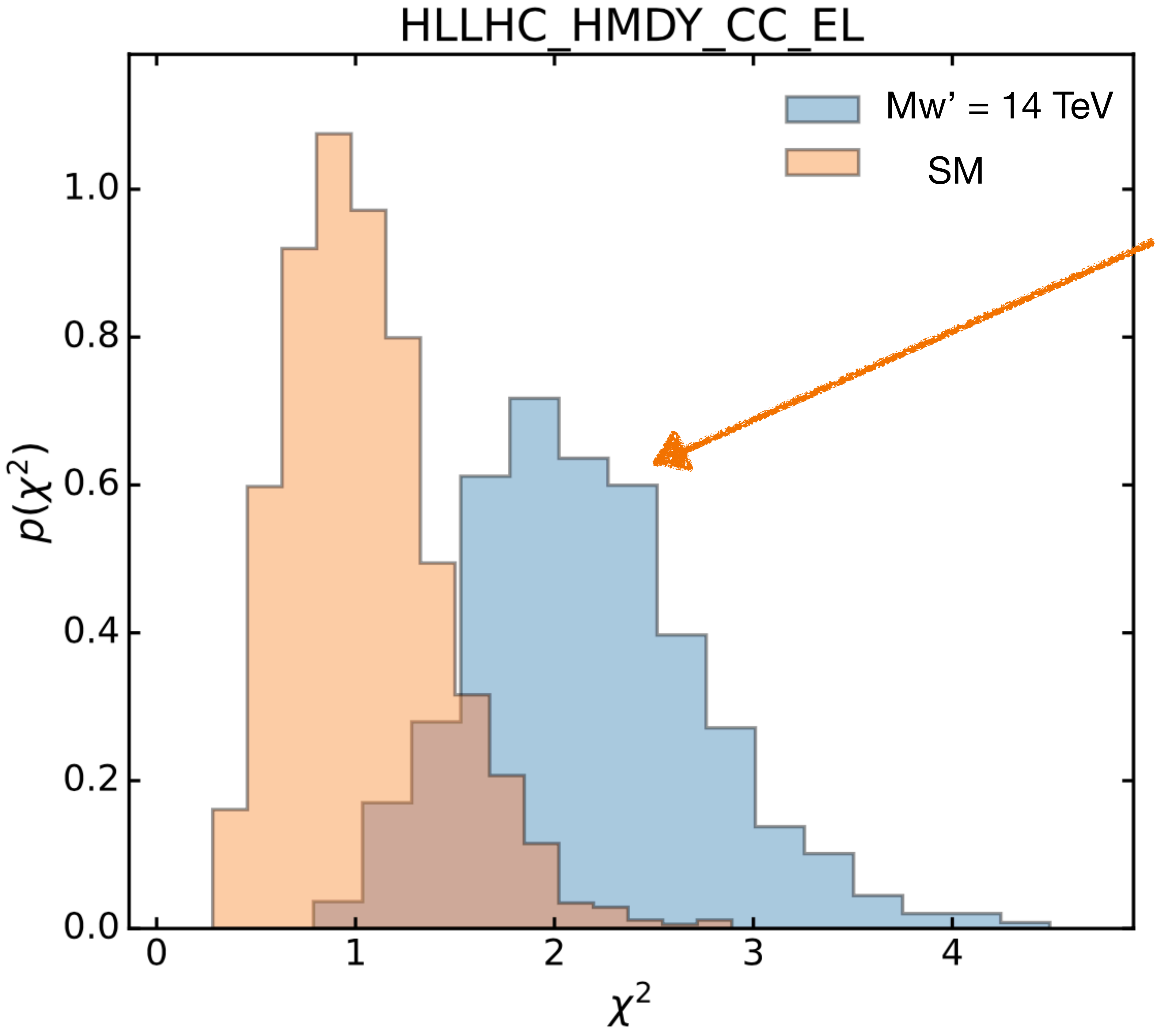
E. Hammou, et al 2307.10370

HOW TO AVOID NEW PHYSICS CONTAMINATION?

- LHCb on-shell at high rapidity data do not help as quark probed at large x, antiquark at small x
- Need more accurate low-energy/large-x constraining measurements to really disentangle such effects
- Strong motivation for synergy with ep experiments (FCC-eh, EIC and Forward Physics Facility) probing lower energies (hence no “contamination” from heavy new physics) and large-x

Low energy fixed target Drell-Yan data constrain antiquarks at large x, if they were more precise, tension with new-physics contaminated HL-LHC data would be evident!





Inclusion of FPF pseudo-data (FASERv, FASERv2, advSNDv) along with HL-LHC (DY NC and DY CC) pseudo-data would prevent new physics contaminations!

Hammou, Madigan, MU, in preparation

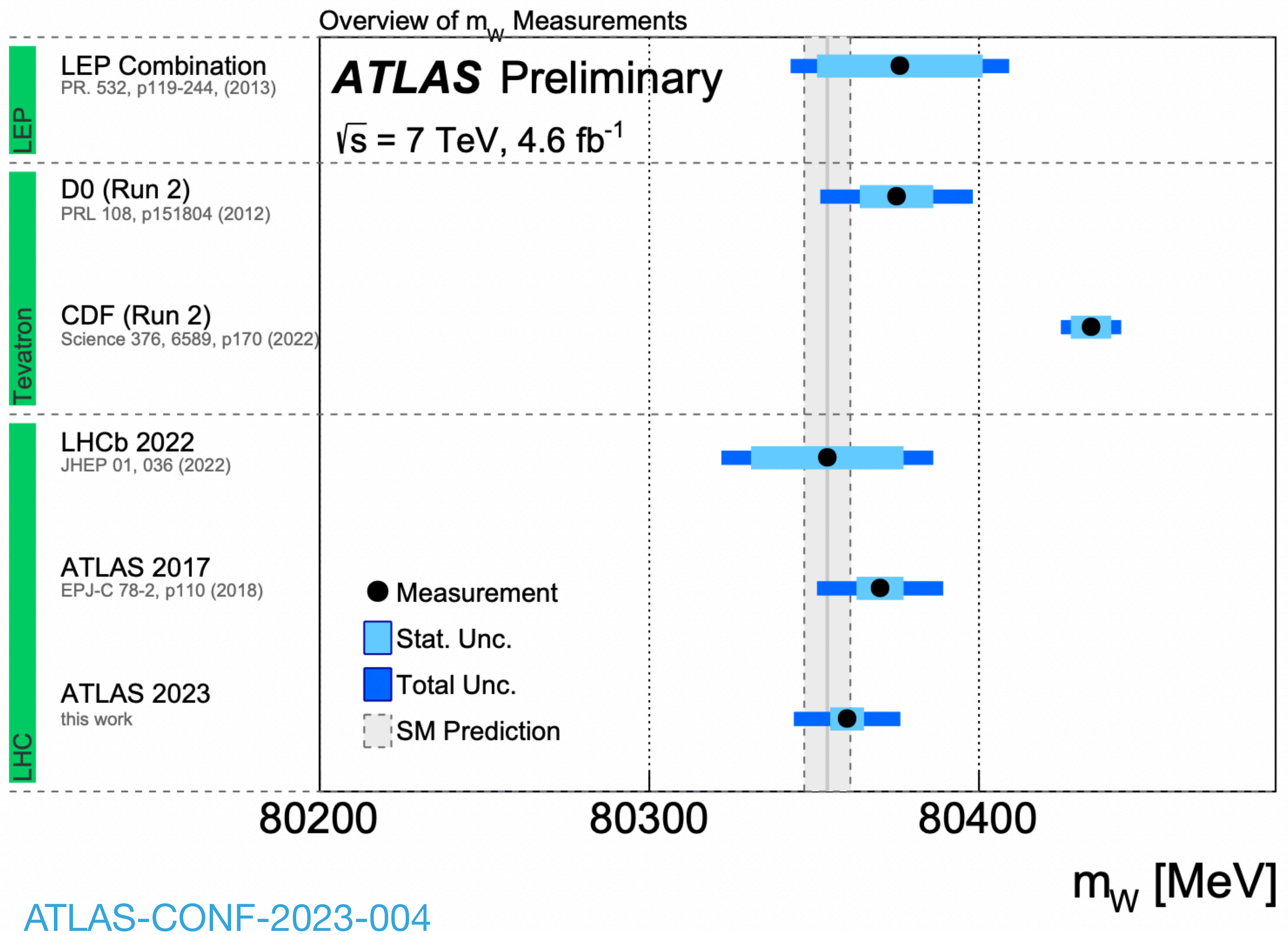
PRECISION PHYSICS

A GLOBAL INTERPRETATION OF THE LHC DATA

SMEFT / BSM SIGNATURES

GLOBAL INTERPRETATION OF DATA

GLOBAL INTERPRETATION OF LHC DATA



ATLAS-CONF-2023-004

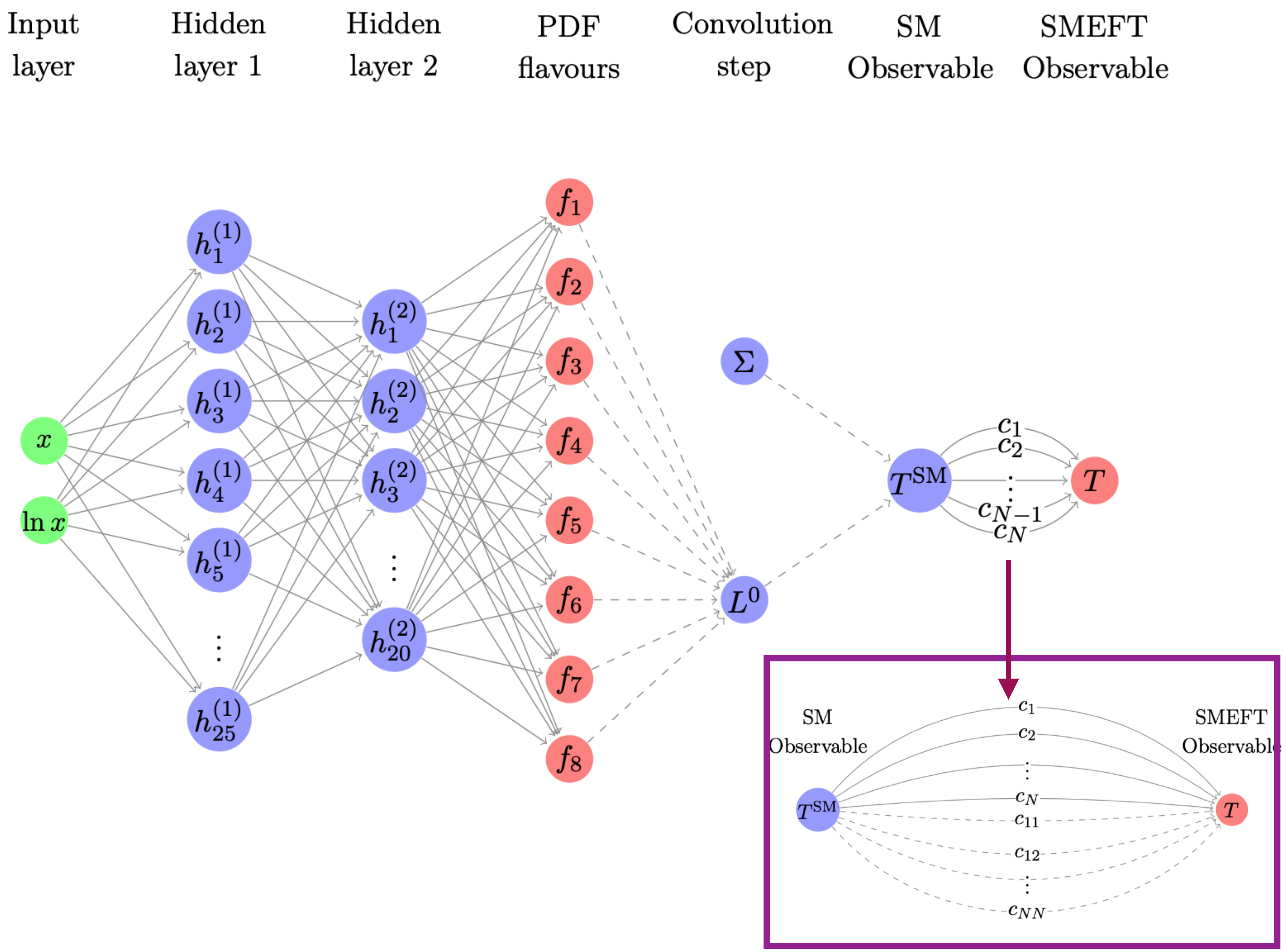
Extremely precise LHC data & advances in theoretical predictions and statistical techniques allow to extract SM (and BSM) parameters to a great level of precision

- $\alpha_s(M_z)$
- M_w
- Parton Distribution Functions
- SMEFT Wilson coefficients
- …

While huge progress made in determining each of these key ingredients of theoretical predictions from the data, not yet evident how to combine all these partial fits into a global interpretation of the LHC data. Simultaneous fits are pivotal step in this direction.

SIMUNET: A TOOL FOR SIMULTANEOUS FITS

- How to perform simultaneous fits of PDFs and SM/BSM parameters?
- SimuNET yields simultaneous fit of PDFs and SMEFT coefficients, it does not have limit in number of parameters that can be fitted alongside PDFs at the initial scale! Extendable to SM parameters!



Linear dim-6 operator

$$T(\hat{\theta}) = \Sigma(\{c_n\}) \cdot L^0(\theta) = T^{\rm SM}(\theta) \cdot \left(1 + \sum_{n=1}^{N} c_n R_{\rm SMEFT}^{(n)}\right)$$

$$T^{\rm SM}(\theta) = \Sigma^{\rm SM} \cdot L^0(\theta)$$

Quadratic dim-6 operator

$$T(\hat{\theta}) = T^{\rm SM}(\theta) \cdot \left(1 + \sum_{n=1}^{N} c_n R_{\rm SMEFT}^{(n)} + \sum_{1\leq n\leq m\leq N} c_{nm} R_{\rm SMEFT}^{(n,m)},\right)$$

$$c_n c_m$$

S. Iranipour, MU - arXiv: 2201.07240, M. Costantini, E Hammou, M. Madigan, L. Mantani, J. Moore, M. Morales, MU (arXiv: 2402.03308)

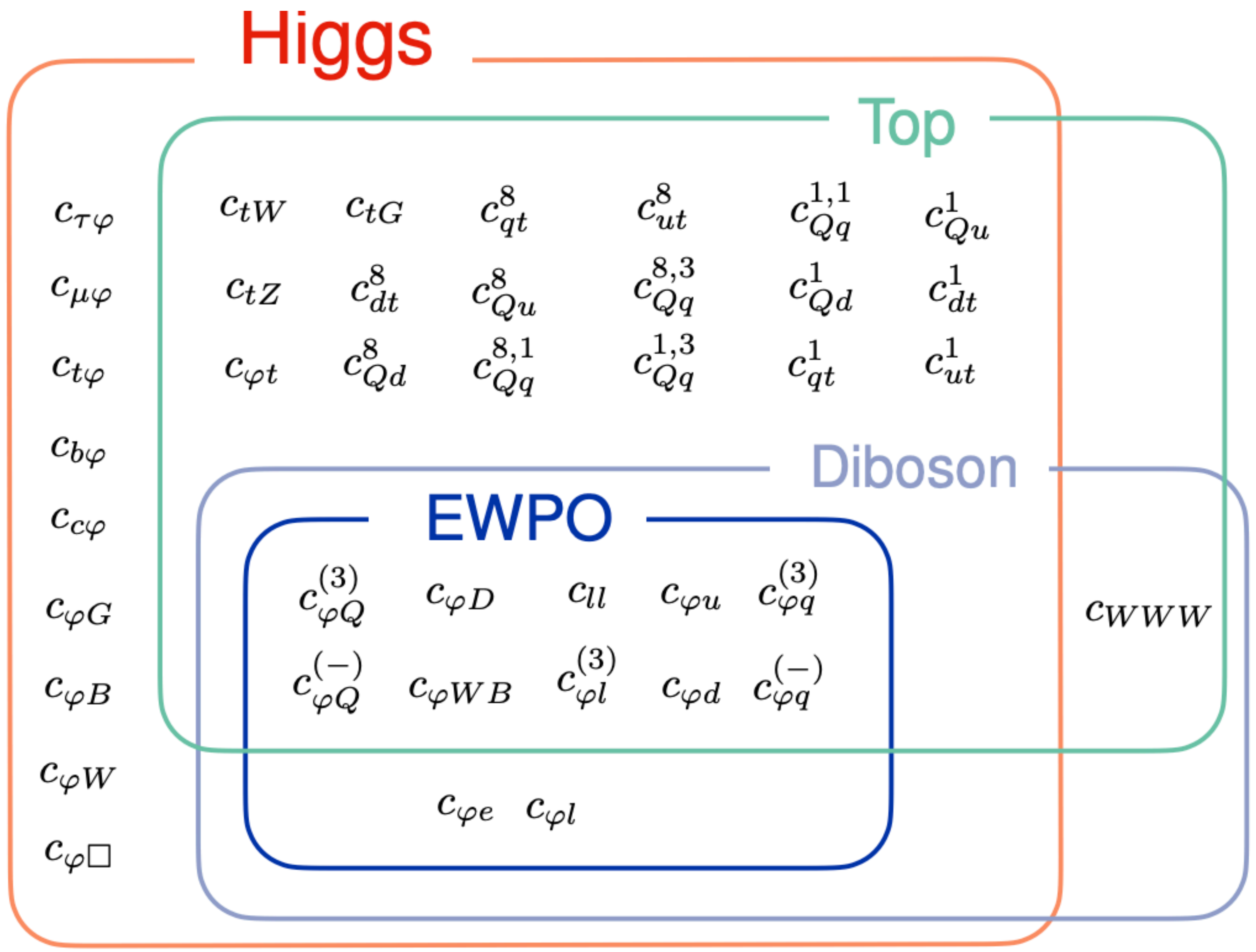
SIMUNET: A TOOL FOR SIMULTANEOUS FITS

SimuNET released open-source with detailed documentation & allow users

https://github.com/HEP-PBSP/SIMUnet

- ✓ PDF only fits
- ✓ SMEFT only fits - linear SMEFT
- ✓ Simultaneous SMEFT & PDFs - linear SMEFT

```
dataset_inputs:
- {dataset: NMC, frac: 0.75}
- {dataset: ATLASTTBARTOT7TEV, cfac: [QCD], simu_fac: "EFT_NLO"}
- {dataset: CMS_SINGLETOPW_8TEV_TOTAL, simu_fac: "EFT_NLO", use_fixed_predictions: True}
```



40 SMEFT operators can be fitted alongside PDFs

SIMUNET: A TOOL FOR SIMULTANEOUS FITS

SimuNET released open-source with detailed documentation & allow users

https://github.com/HEP-PBSP/SIMUnet

- ✓ PDF only fits
- ✓ SMEFT only fits - linear SMEFT
- ✓ Simultaneous SMEFT & PDFs - linear SMEFT

```
dataset_inputs:
- {dataset: NMC, frac: 0.75}
- {dataset: ATLASTTBARTOT7TEV, cfac: [QCD], simu_fac: "EFT_NLO"}
- {dataset: CMS_SINGLETOPW_8TEV_TOTAL, simu_fac: "EFT_NLO", use_fixed_predictions: True}
```

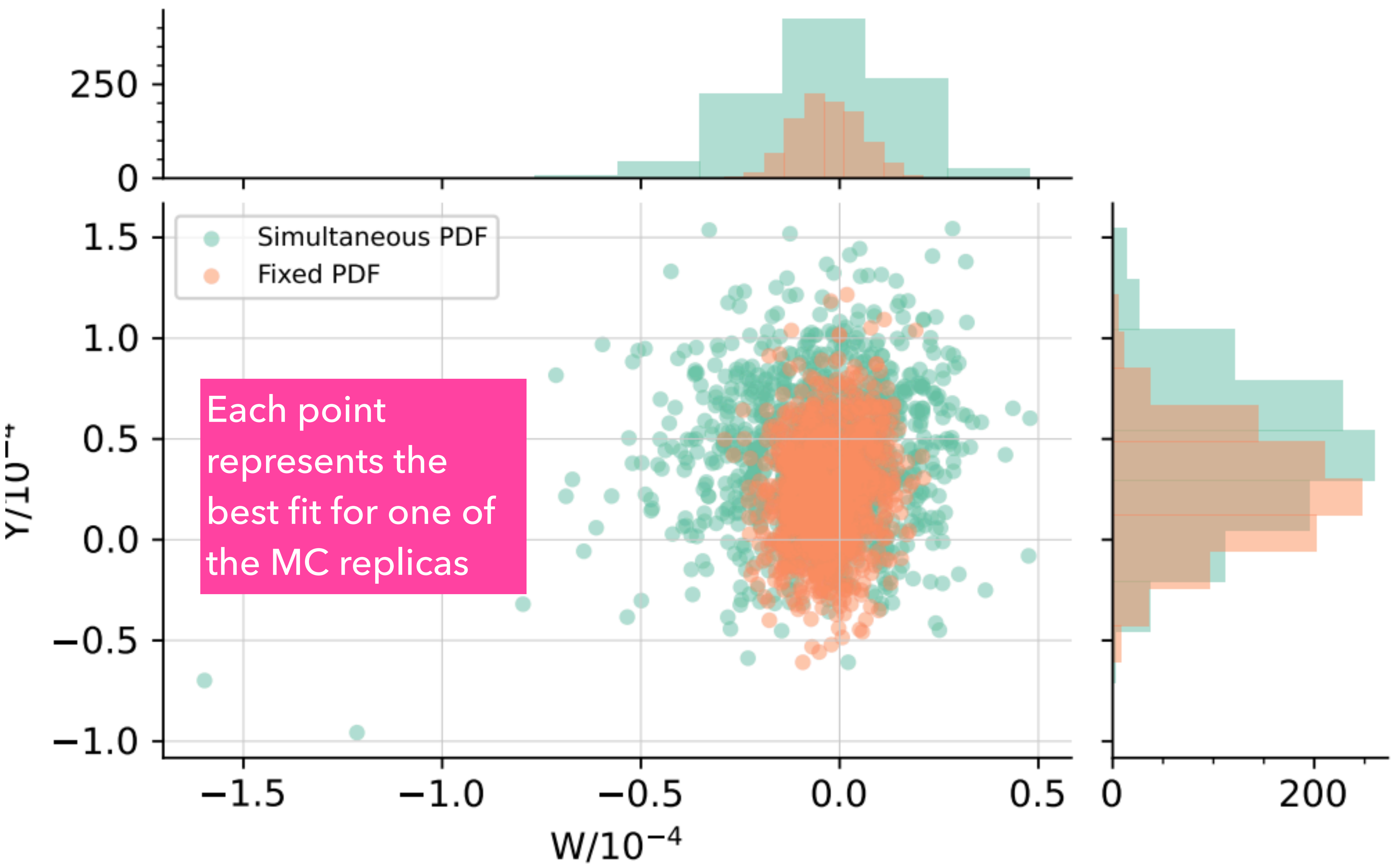
- ✓ Inject any new physics model in the data and check robustness against PDF absorbing it

```
dataset_inputs:
- {dataset: LHCB_Z_13TEV_DIELECTRON, frac: 0.75, cfac: ['QCD']}
- {dataset: CMSDY1D12, frac: 0.75, cfac: ['QCD', 'EWK'], contamination: 'EFT_LO'}
```

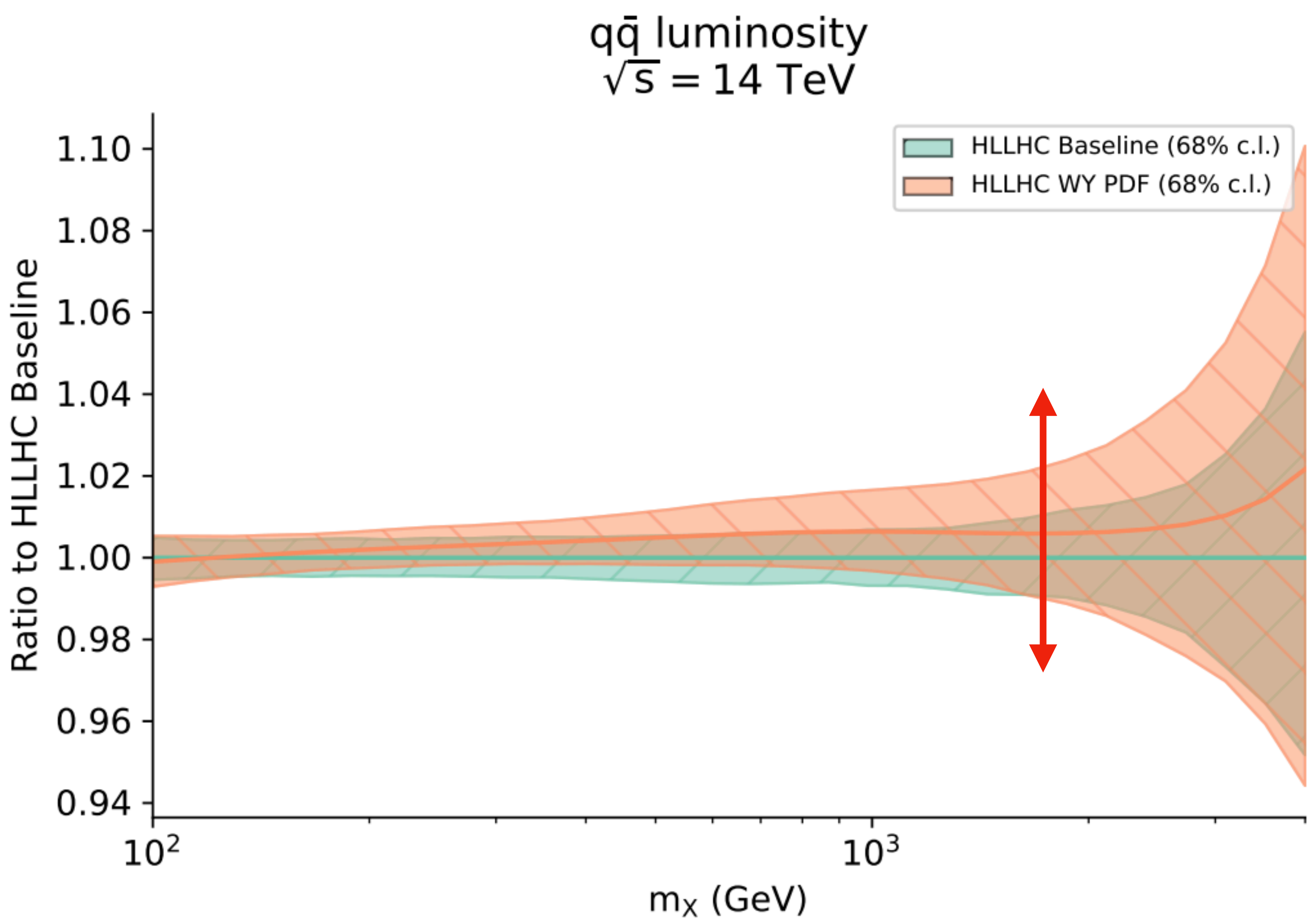
M. Costantini, E Hammou, M. Madigan, L. Mantani, J. Moore, M. Morales, MU

arXiv: 2402.03308

SIMUNET: APPLICATION TO DRELL-YAN SECTOR



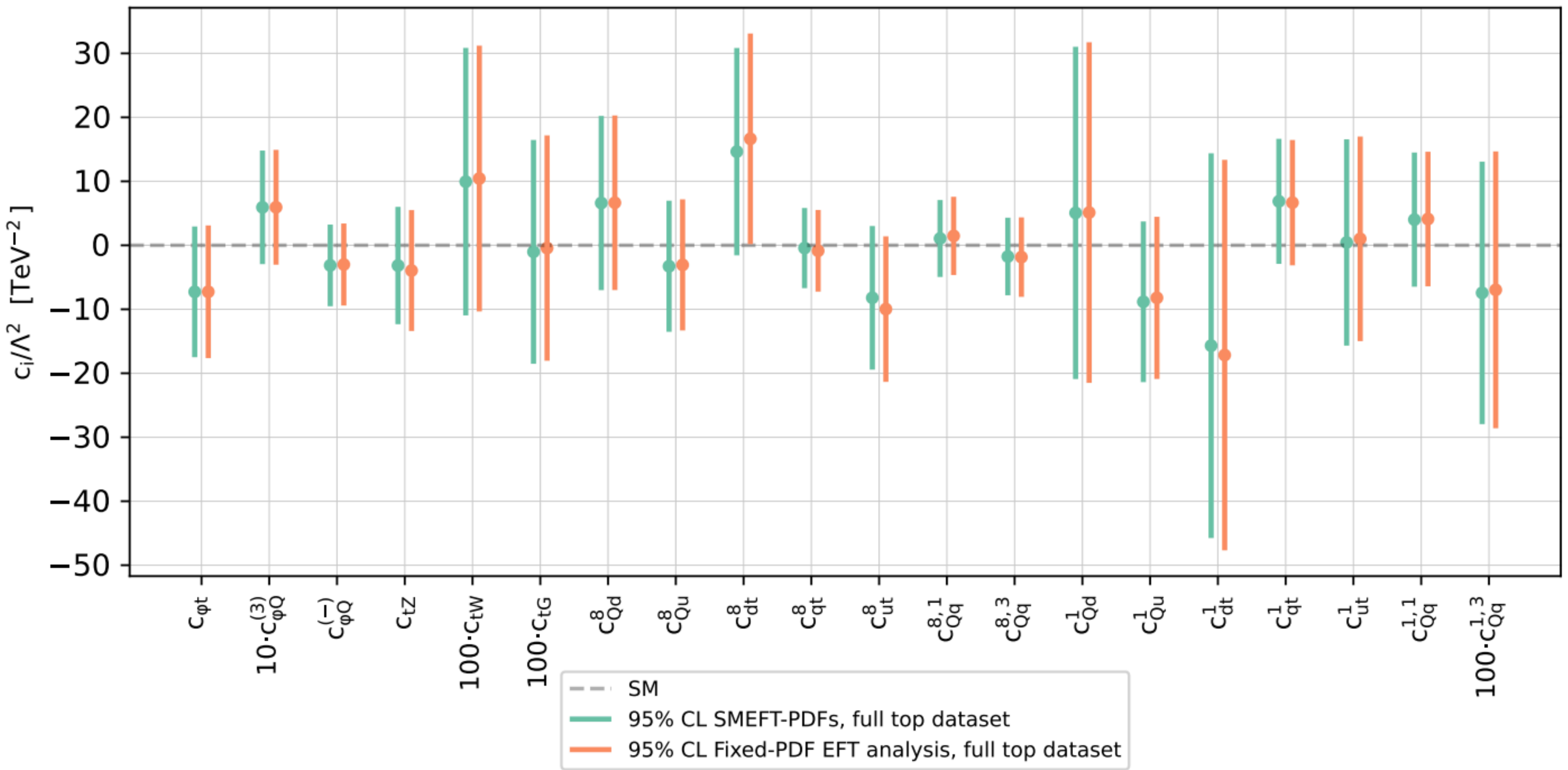
S. Iranipour, MU - arXiv: 2201.07240

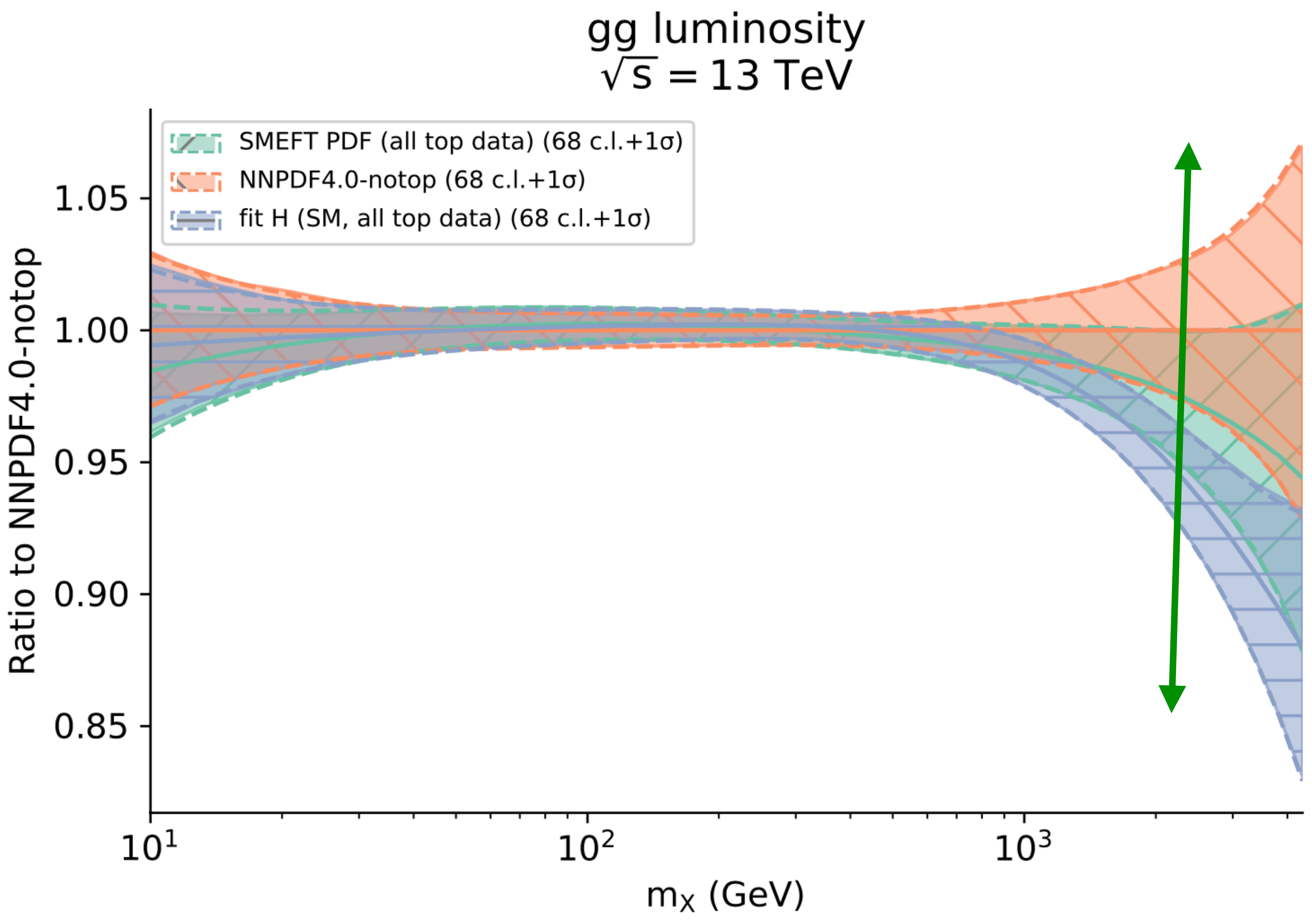


✓ Simultaneous analysis of PDFs and **Drell-Yan sector** Wilson coefficient in the context of universal parameters of DIS + DY (including HL-LHC projections) using simuNET method shows that at HL-LHC the effect of interplay becomes important: WCs bounds broaden and PDF uncertainties change significantly once SMEFT effects allowed in theory predictions entering PDF fit

[Greljo et al 2104.02723] [Iranipour, MU 2201.07240]

SIMUNET: APPLICATION TO TOP SECTOR





Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU - arXiv: 2303.06159

✓ Simultaneous analysis of PDFs and more than 20 operators in the **top sector** using simuNET method shows that WCs are stable, while PDF uncertainties broaden and PDF fitted simultaneously alongside WCs site nicely between PDF-only fit without top data and PDF fit including all top data - sign of partial NP absorption (!)

CONCLUSIONS AND OUTLOOK

- Interplay between indirect new physics searches via (SM)EFT fits and PDFs is going to be more and more relevant as we move to the High-Luminosity LHC phase

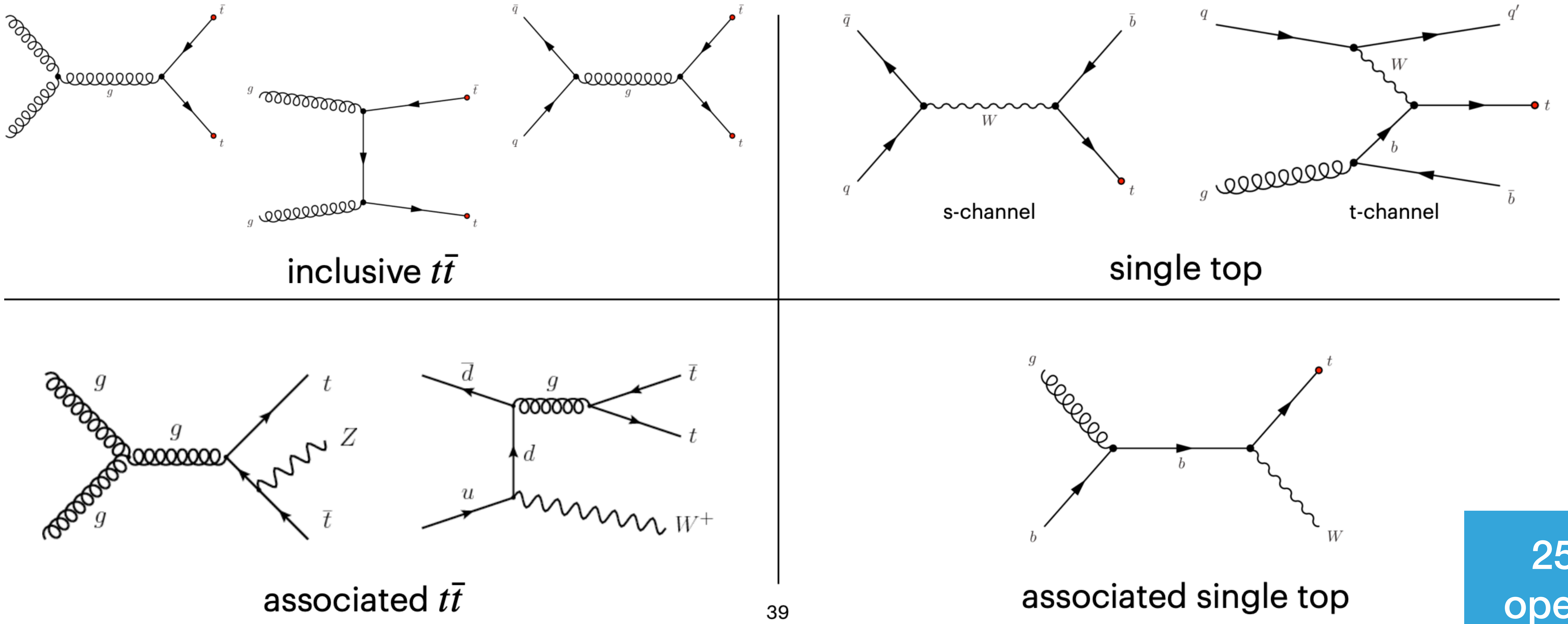
- Can PDF absorb new physics?
  ➡ Identified a naive UV scenario such that the high-mass HL-LHC invariant mass can absorb NP and quantified effect
  ➡ Important to disentangle large-x from high-energy / low-energy (SeaQuest, JLAB, EIC, FPF..)
  ➡ More realistic scenarios to explore, especially large-x gluons (jets, top)

- Simultaneous fits
  ➡General simuNET methodology for simultaneous fits linear SMEFT + PDFs for an arbitrary number of SMEFT operators has been developed
  ➡Open source code
  ➡ Effect of simultaneous fits on PDFs and SMEFT depends on the sector
  ➡ Top: shift and larger uncertainty in PDFs, unchanged SMEFT bounds
  ➡ Drell-Yan: unshifted PDFs and large uncertainties, broader SMEFT bounds
  ➡ Quadratic SMEFT corrections: need a new methodology to determine joint probability distribution beyond Monte Carlo sampling (Bayesian PDFs)

EXTRA MATERIAL

THE TOP SECTOR

- After testing methodology on small number of WC, stress-test on large SMEFT parameters space.
- Huge amount of Run II top quark data from ATLAS and CMS
- Four basic processes: inclusive tt~ and asymmetry (inclusive and differential), single top (inclusive and differential), associated ttV production, associated single top production



25 (21) dim-6 operators at the quadratic (linear) SMEFT

Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU - arXiv: 2303.06159

THE TOP SECTOR

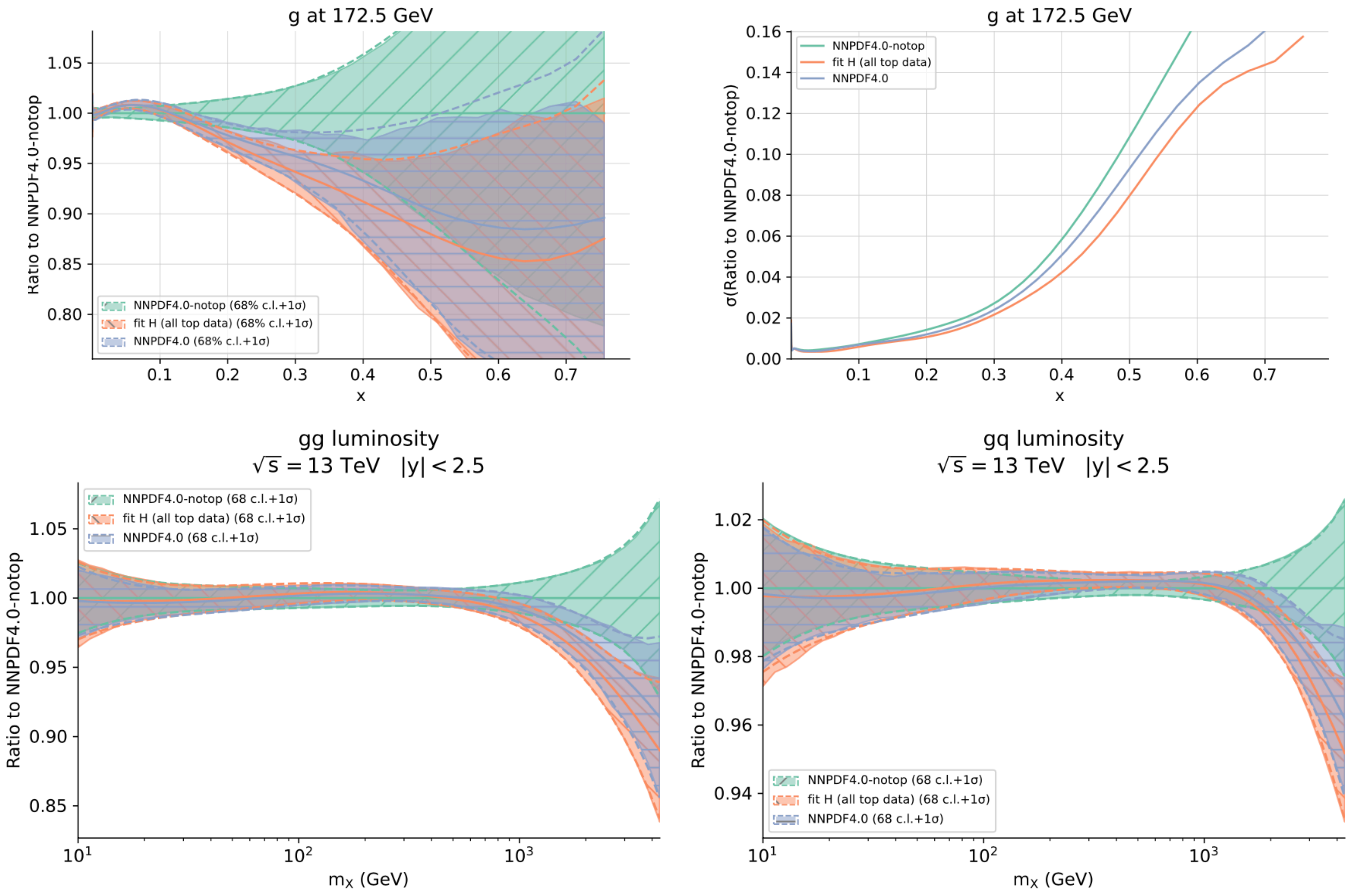
DoF	Definition (Warsaw basis)
c^1_{QQ}	$2c^{1(3333)}_{qq} - \frac{2}{3}c^{3(3333)}_{qq}$
c^8_{QQ}	$8c^{3(3333)}_{qq}$
c^1_{Qt}	$c^{1(3333)}_{qu}$
c^8_{Qt}	$c^{8(3333)}_{qu}$
c^1_{tt}	$c^{(3333)}_{uu}$
$c^{1,8}_{Qq}$	$c^{1(i33i)}_{qq} + 3c^{3(i33i)}_{qq}$
$c^{1,1}_{Qq}$	$c^{1(ii33)}_{qq} + \frac{1}{6}c^{1(i33i)}_{qq} + \frac{1}{2}c^{3(i33i)}_{qq}$
$c^{3,8}_{Qq}$	$c^{1(i33i)}_{qq} - c^{3(i33i)}_{qq}$
$c^{3,1}_{Qq}$	$c^{3(ii33)}_{qq} + \frac{1}{6}(c^{1(i33i)}_{qq} - c^{3(i33i)}_{qq})$
c^8_{tq}	$c^{8(ii33)}_{qu}$
c^1_{tq}	$c^{1(ii33)}_{qu}$
c^8_{tu}	$2c^{(i33i)}_{uu}$
c^1_{tu}	$c^{(ii33)}_{uu} + \frac{1}{3}c^{(i33i)}_{uu}$
c^8_{Qu}	$c^{8(33ii)}_{qu}$
c^1_{Qu}	$c^{1(33ii)}_{qu}$
c^8_{td}	$c^{8(33jj)}_{ud}$
c^1_{td}	$c^{1(33jj)}_{ud}$
c^8_{Qd}	$c^{8(33jj)}_{qd}$
c^1_{Qd}	$c^{1(33jj)}_{qd}$

25 (21) dim-6 operators at the quadratic (linear) SMEFT

Operator	Coefficient	Definition
$\mathcal{O}^{(1)}_{\varphi Q}$	$-\ (c^{(1)}_{\varphi Q})$	$i(\varphi^\dagger \overleftrightarrow{D}_\mu \varphi)(\bar{Q}\gamma^\mu Q)$
$\mathcal{O}^{(3)}_{\varphi Q}$	$c^{(3)}_{\varphi Q}$	$i(\varphi^\dagger \overleftrightarrow{D}_\mu \tau_I \varphi)(\bar{Q}\gamma^\mu \tau^I Q)$
$\mathcal{O}_{\varphi t}$	$c_{\varphi t}$	$i(\varphi^\dagger \overleftrightarrow{D}_\mu \varphi)(\bar{t}\gamma^\mu t)$
$\mathcal{O}_{tW}$	c_{tW}	$i(\bar{Q}\tau^{\mu\nu}\tau_I t)\tilde{\varphi} W^I_{\mu\nu} + \text{h.c.}$
$\mathcal{O}_{tB}$	$-\ (c_{tB})$	$i(\bar{Q}\tau^{\mu\nu} t)\tilde{\varphi} B_{\mu\nu} + \text{h.c.}$
$\mathcal{O}_{tG}$	c_{tG}	$i(\bar{Q}\tau^{\mu\nu} T_A t)\tilde{\varphi} G^A_{\mu\nu} + \text{h.c.}$

DoF	Definition
$c^{(-)}_{\varphi Q}$	$c^{(1)}_{\varphi Q} - c^{(3)}_{\varphi Q}$
c_{tZ}	$-\sin\theta_W c_{tB} + \cos\theta_W c_{tW}$

PDF-ONLY FIT



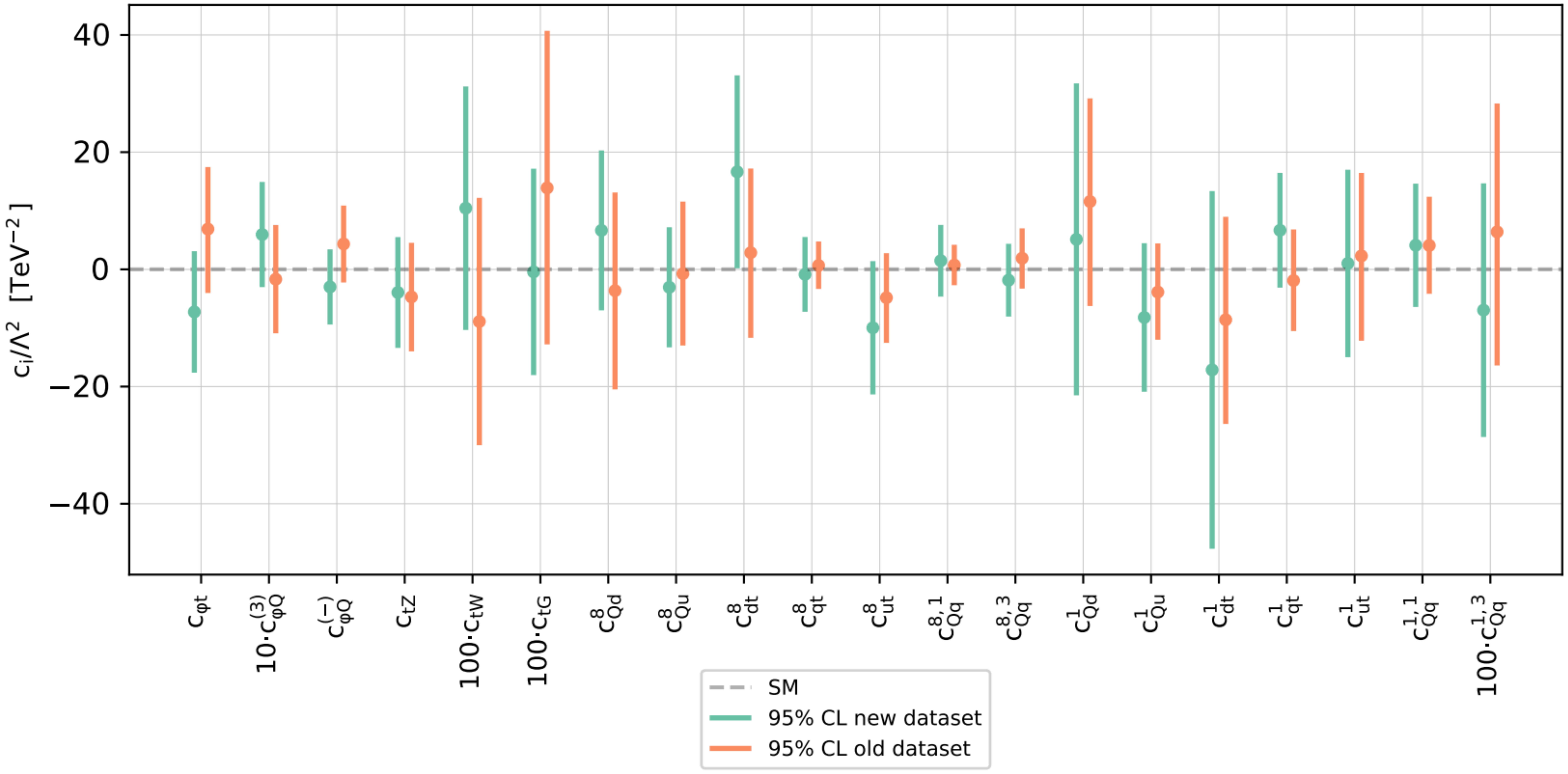
78 new datapoints in the top sector as compared to NNPDF4.0

Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU - arXiv: 2303.06159

SMEFT-ONLY FIT

$$\sigma_{\rm eft}\left(\boldsymbol{c}/\Lambda^2\right) = \sigma_{\rm SM} + \sum_{i=1}^{n_{\rm op}} \sigma_{{\rm eft},i}\frac{c_i}{\Lambda^2} \cdot$$

Linear SMEFT

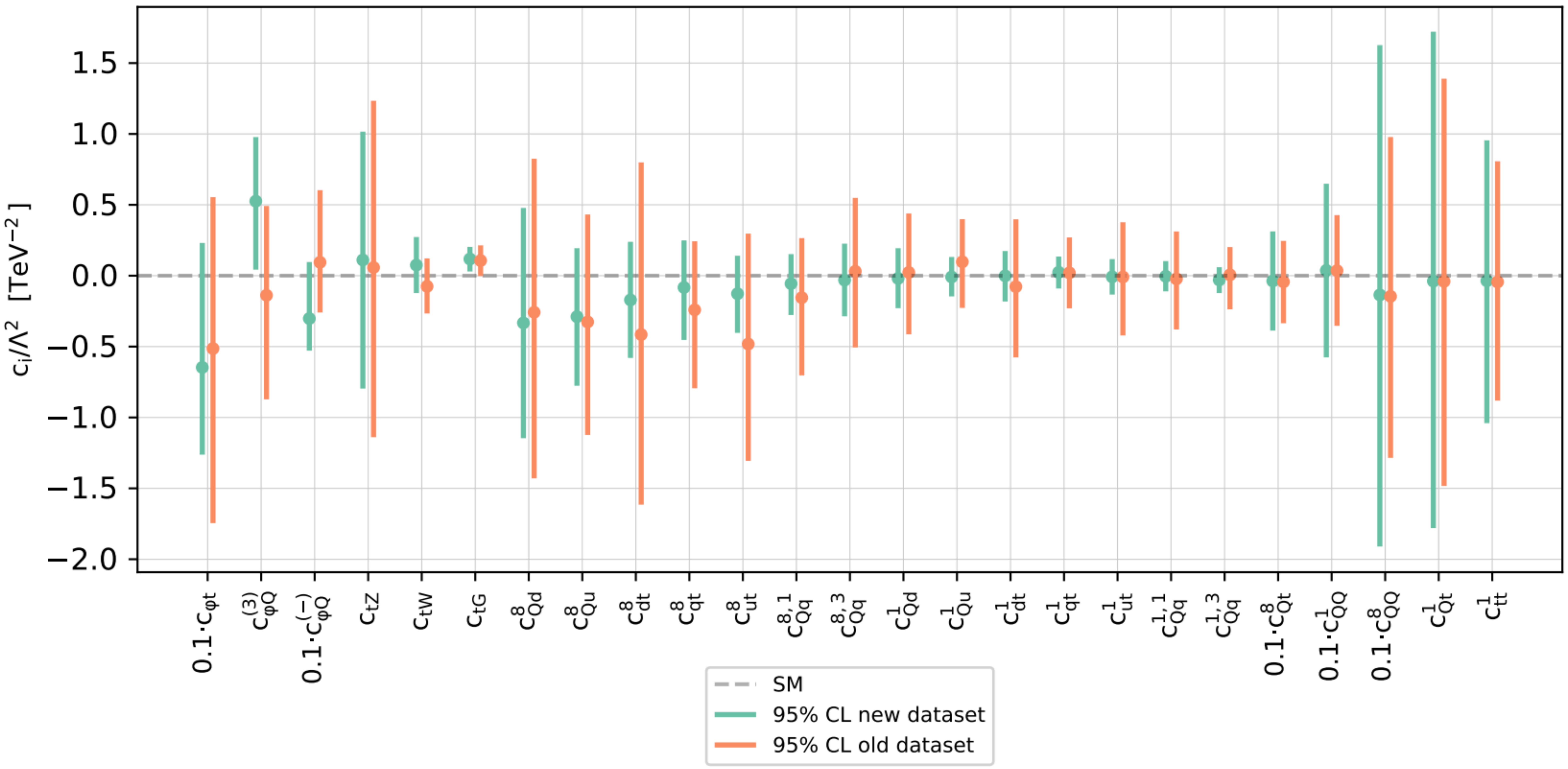


Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU - arXiv: 2303.06159

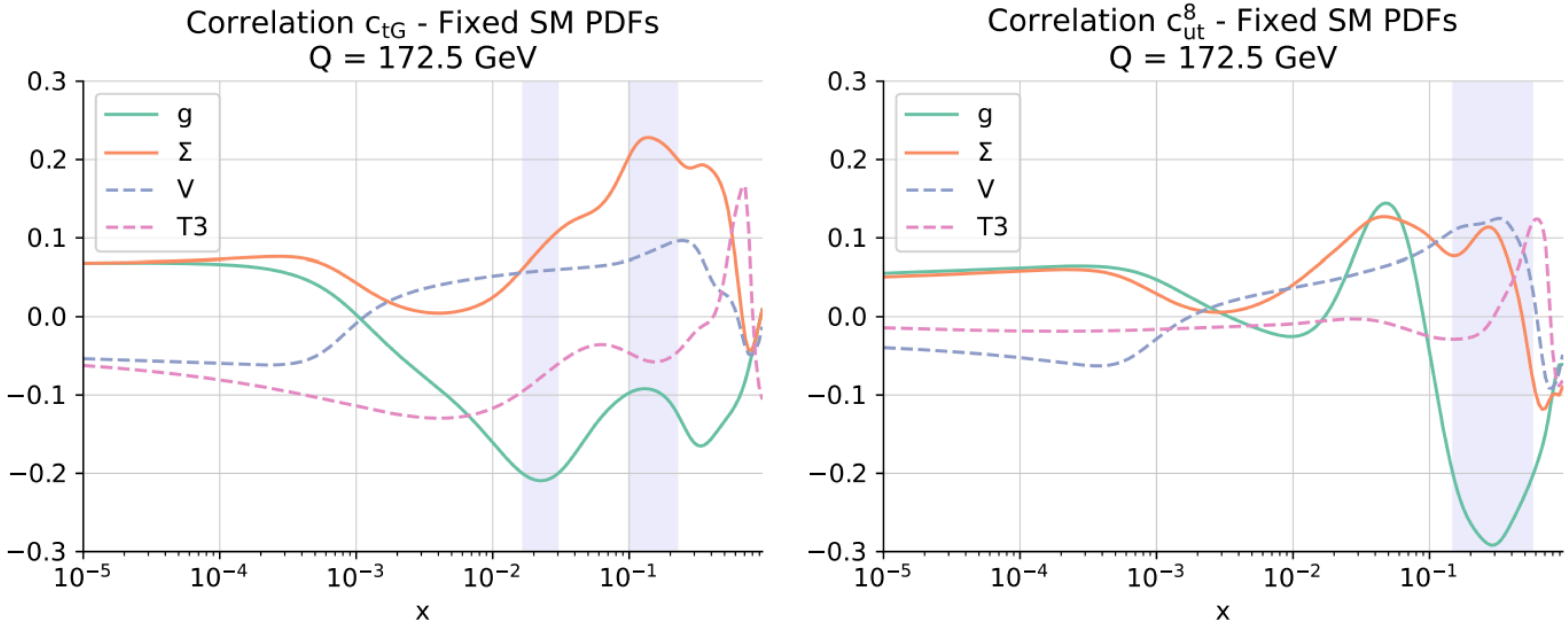
SMEFT-ONLY FIT

$$\sigma_{\rm eft}\left(\boldsymbol{c}/\Lambda^2\right) = \sigma_{\rm SM} + \sum_{i=1}^{n_{\rm op}} \sigma_{{\rm eft},i}\frac{c_i}{\Lambda^2} + \sum_{i,j=1}^{n_{\rm op}} \sigma_{{\rm eft},ij}\frac{c_i c_j}{\Lambda^4}$$

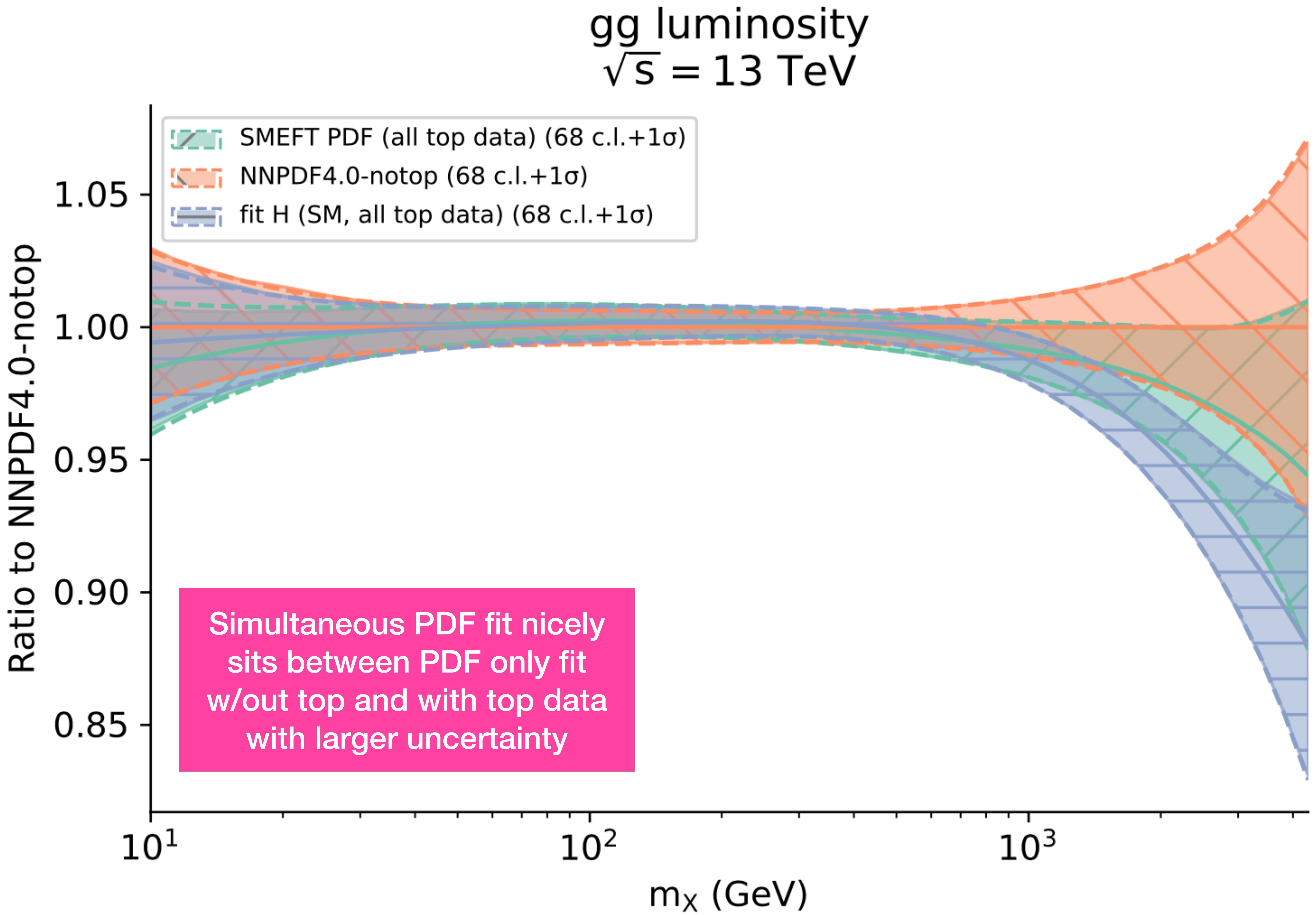
Quadratic SMEFT



SMEFT AND PDF CORRELATIONS



SIMULTANEOUS PDF-SMEFT FIT



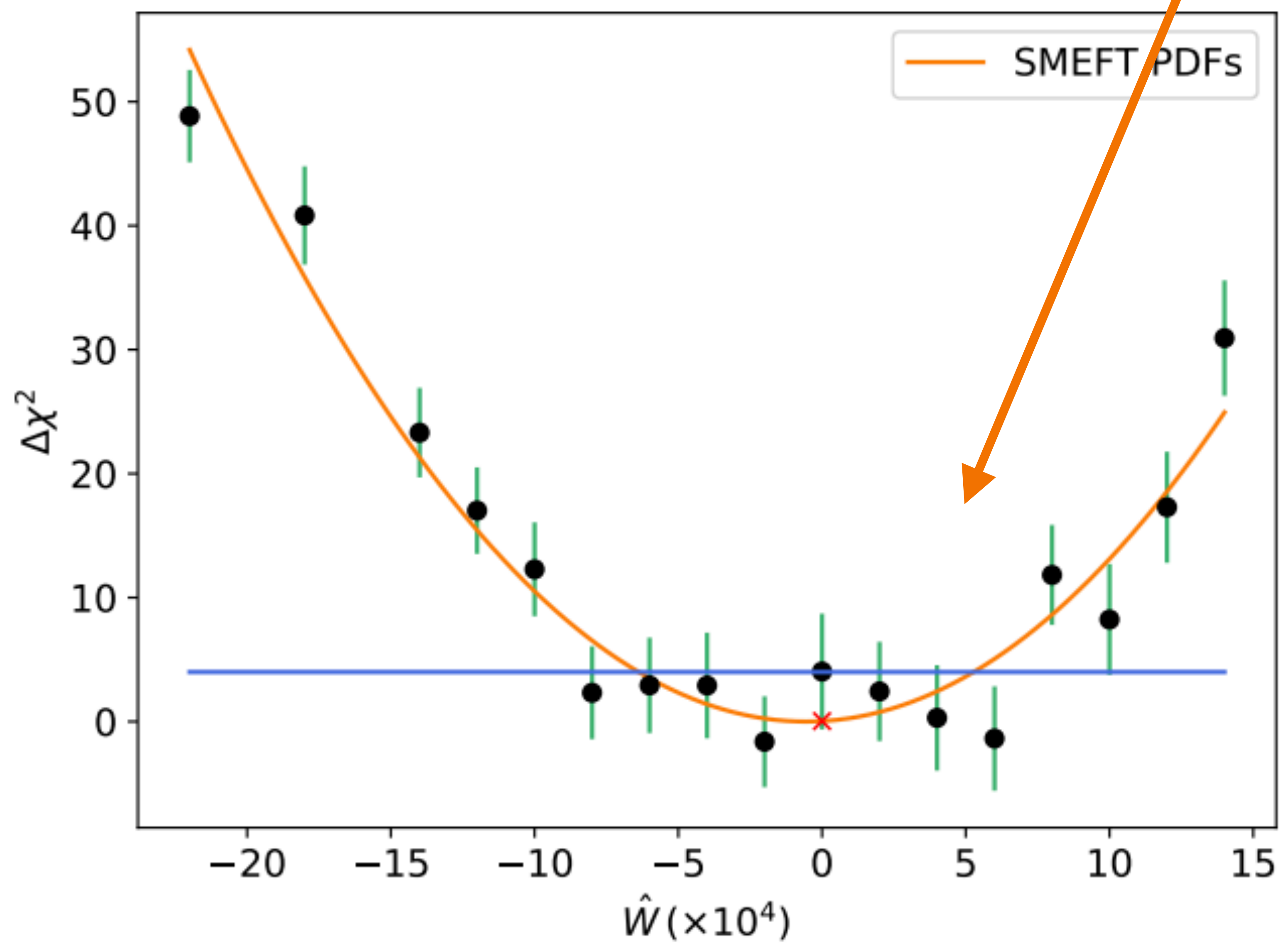
Linear SMEFT

ANALYSIS METHODOLOGY

$$T = f_1(\hat{W}) \otimes f_2(\hat{W}) \otimes \hat{\sigma}(\hat{W})$$

- We performed a similar analysis as in Torre et al, now with emphasis on PDF and their interplay with bounds on oblique operators
  [Greljo, Iranipour, Kassabov, Madigan, Moore, Rojo, MU, Voisey: 2104.02723]

- Methodology for simultaneous fit is similar to the one adopted in fits of α_S from a global fit of PDFs



Greljo et al, 2104.02723

$$\chi^2 = \frac{1}{n_{\rm dat}} \sum_{i,j=1}^{n_{\rm dat}} (D_i - T_i)\,({\rm cov}^{-1})_{ij}\,(D_j - T_j)$$

1. Take data, make theoretical predictions accounting for operator in partonic cross section with fixed SM PDFs.
2. Compute chi2 as a function of WCs (Wilson Coefficients)
3. Minimise chi2 and find best-fit and C.L.s of WCs
4. Extract bounds

$$T = f_{1,\rm SM} \otimes f_{2,\rm SM} \otimes \hat{\sigma}_{\rm BSM}$$

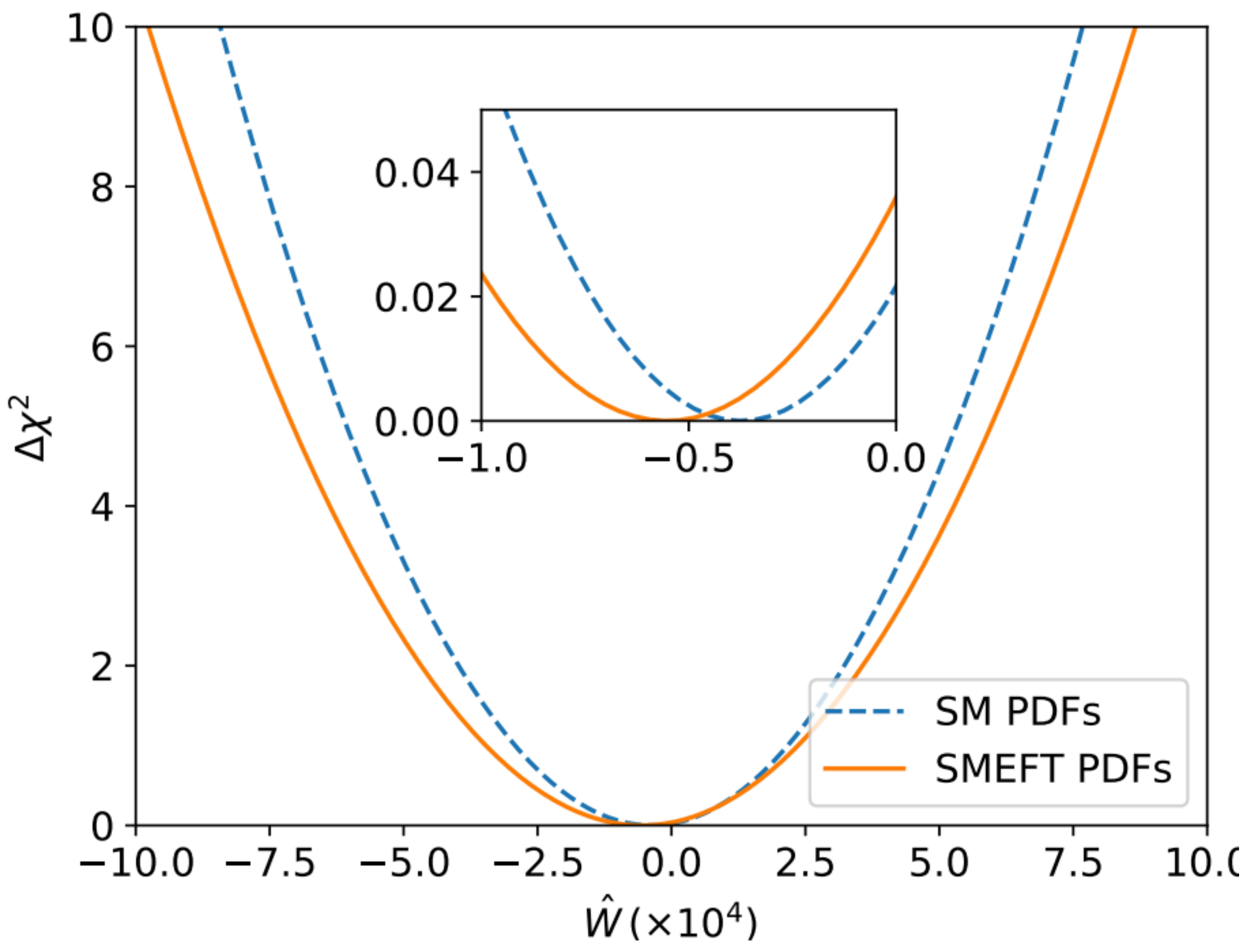
SM PDFs

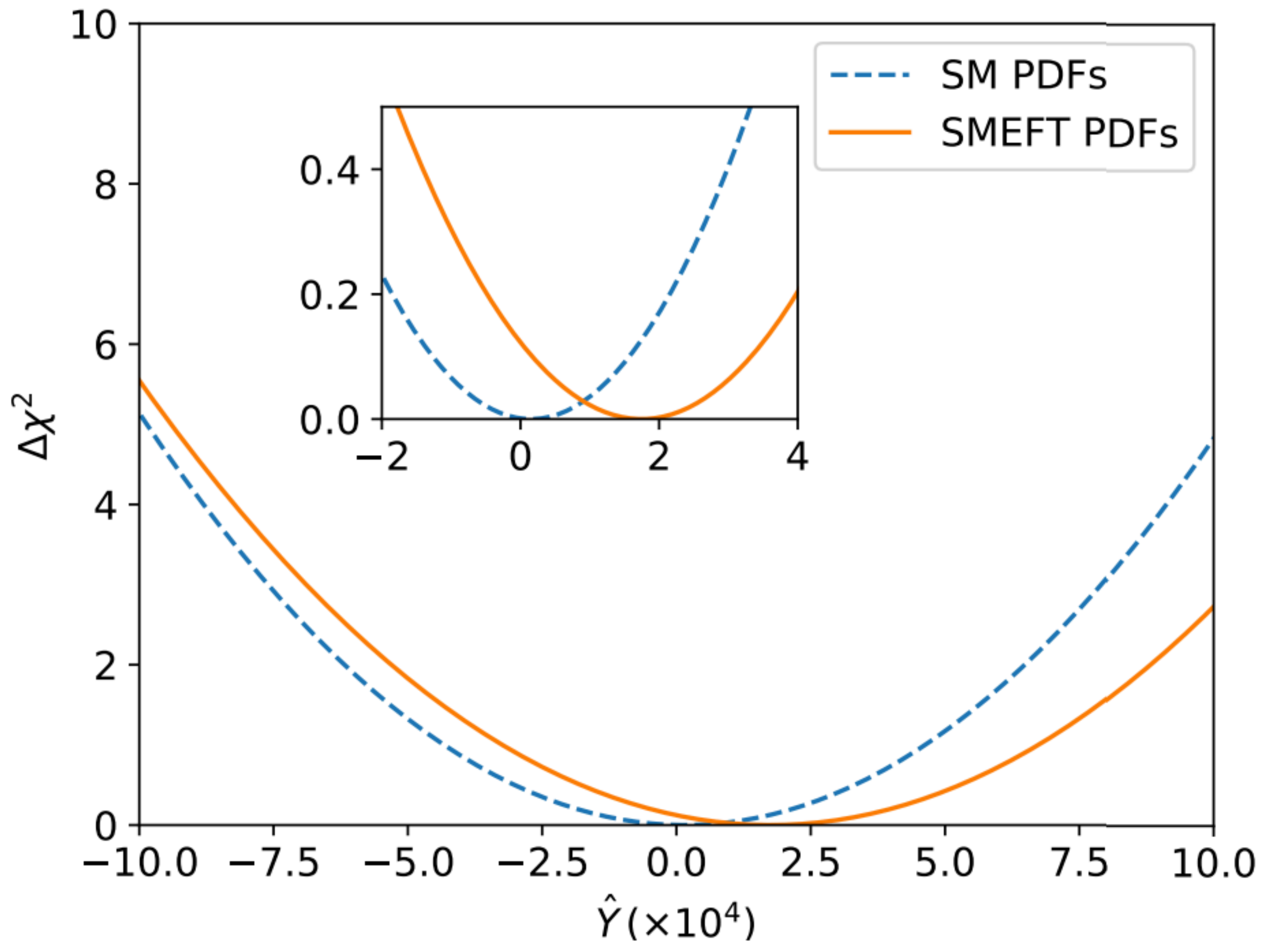
1. Take data, make theoretical predictions accounting for operator in partonic cross section and PDFs.
2. Compute chi2 as a function of WCs (Wilson Coefficients)
3. Minimise chi2 and find best-fit and C.L.s of WCs
4. Extract bounds

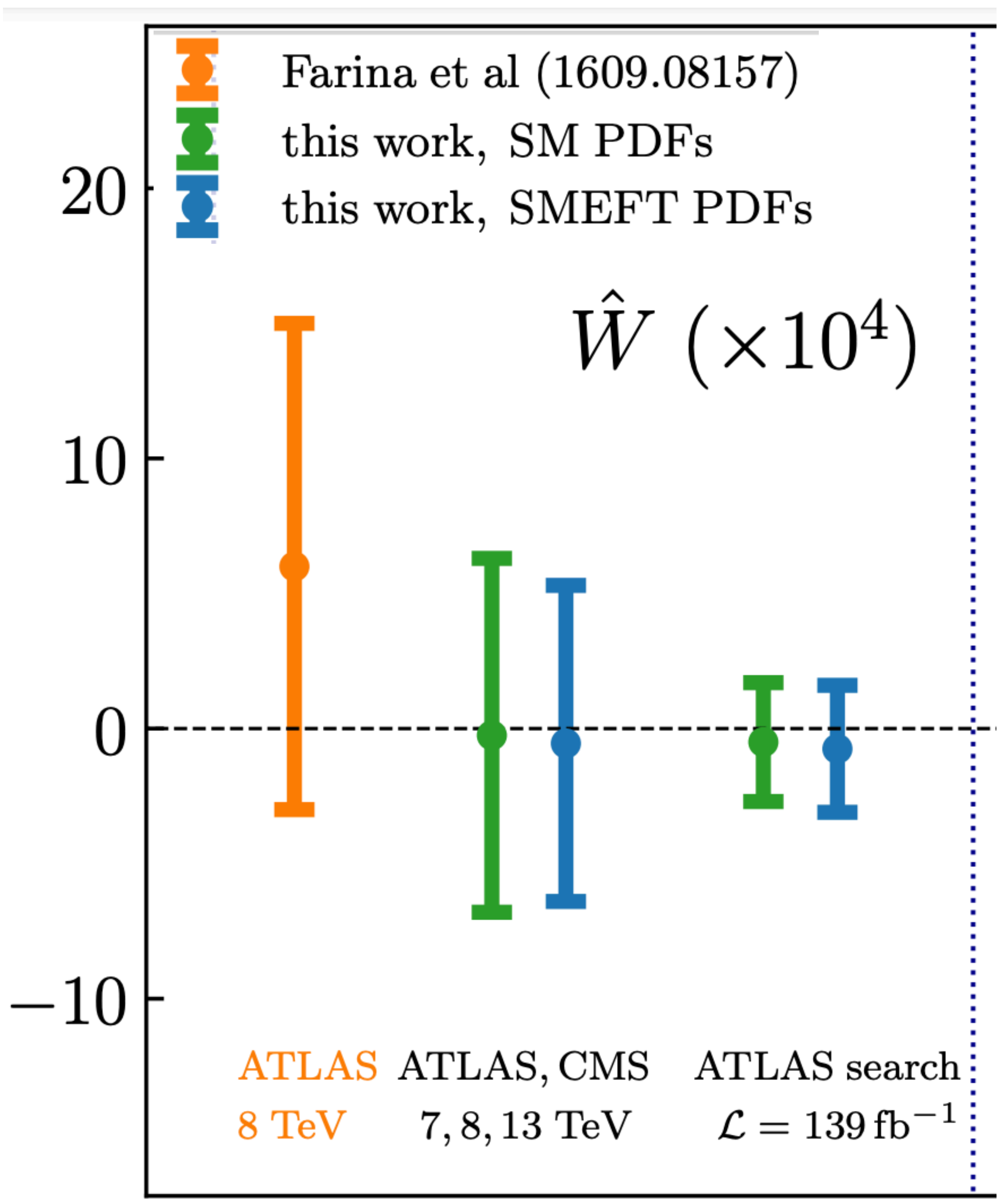
$$T = f_{1,\rm BSM} \otimes f_{2,\rm BSM} \otimes \hat{\sigma}_{\rm BSM}$$

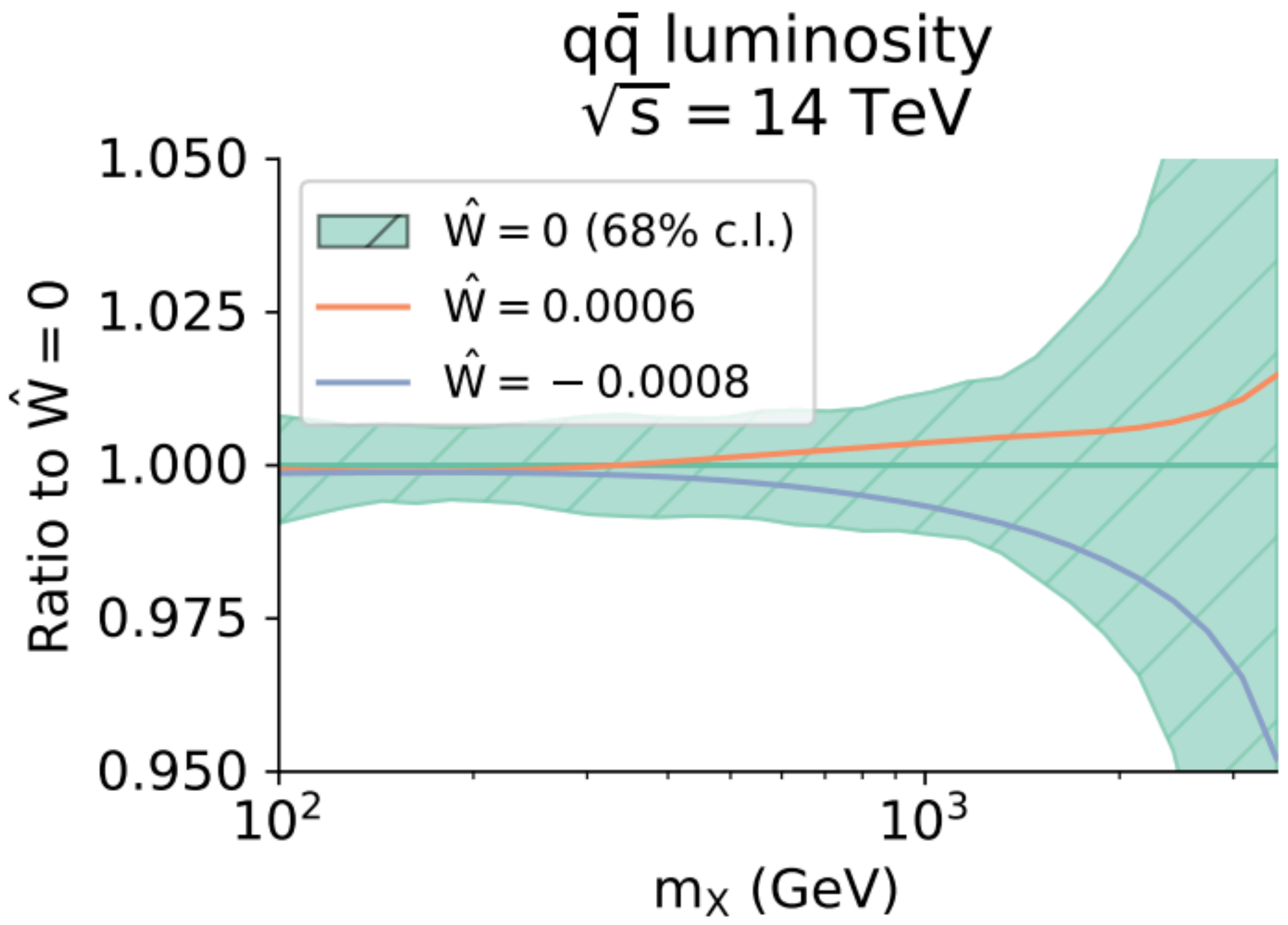
SMEFT PDFs / Simultaneous fit

INTERPLAY @ RUN I AND RUN II









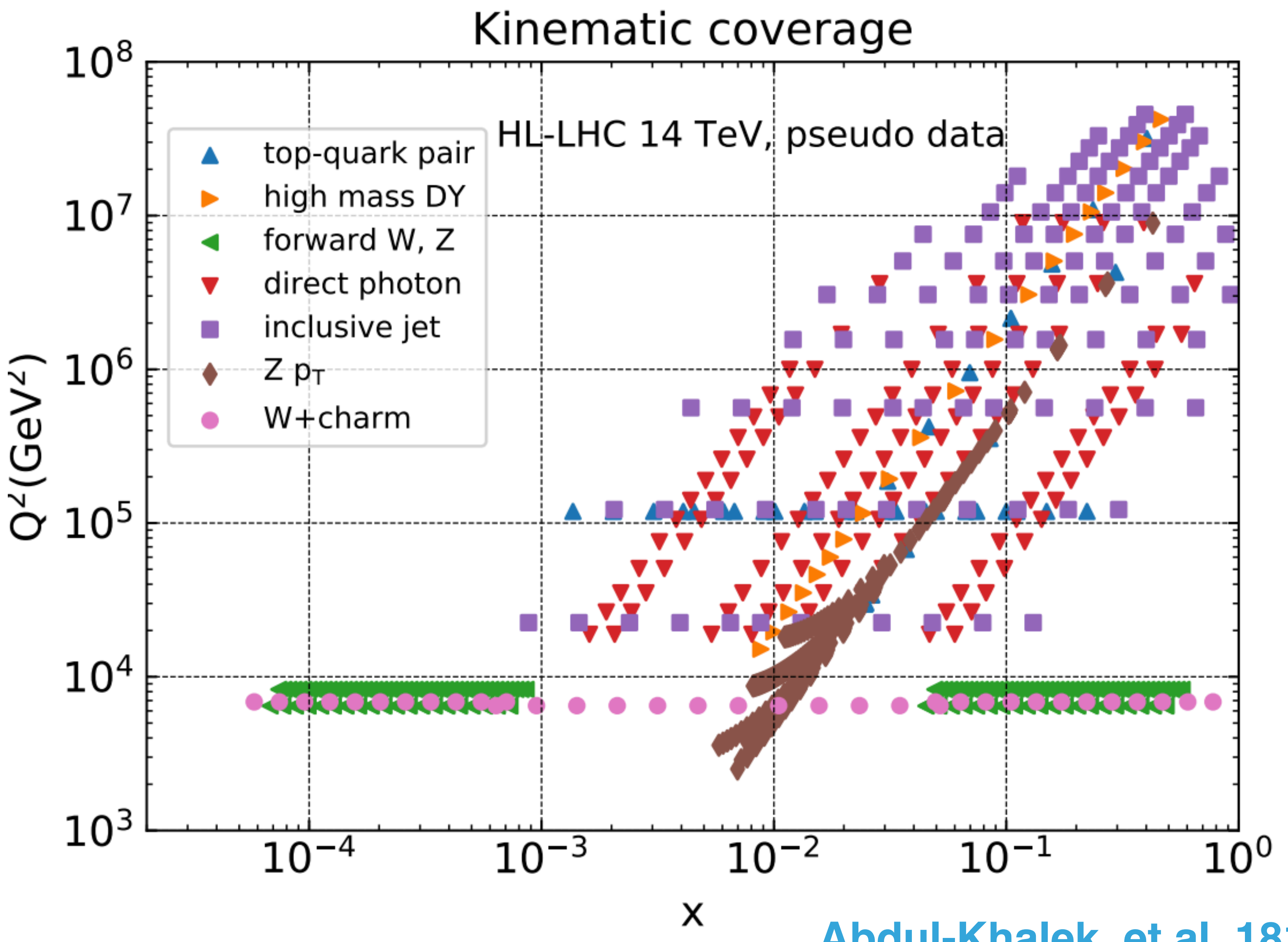
- With current data, PDFs are moderately affected by inclusion of non-zero W and Y coefficients in the fit, mostly quark-antiquark luminosity within uncertainties
- Broadening of individual bounds on W and Y once SMEFT PDFs are used (i.e. PDFs that have been fitted with consistent values of W and Y) is not negligible, but still within PDF uncertainties
- If SMEFT PDFs are used in determining bounds from ATLAS search same mild broadening (larger than PDF uncertainties)

RESULTS: DRELL-YAN DATA @HL-LHC

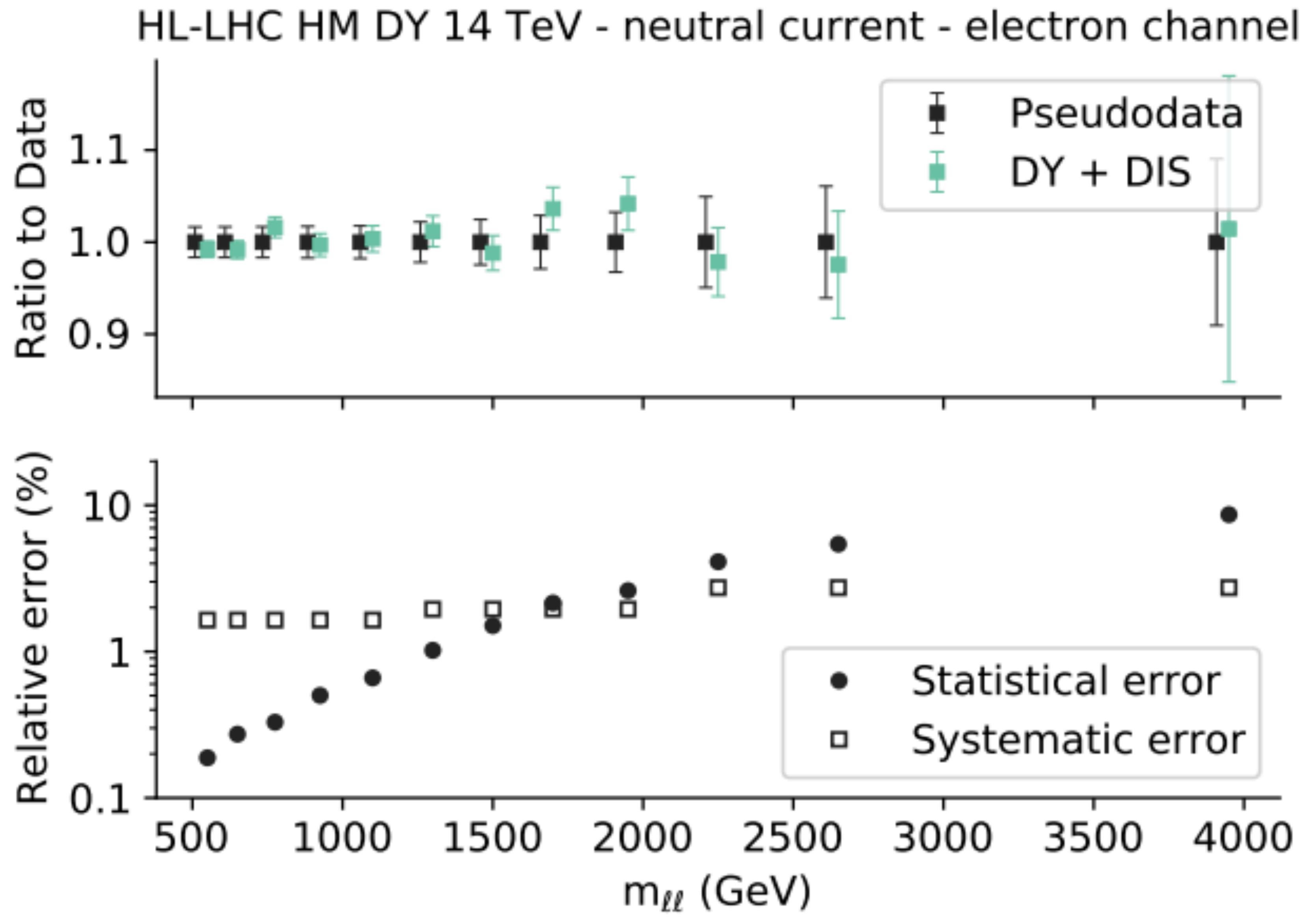
- Add HL-LHC projections for both NC and CC in PDF fit

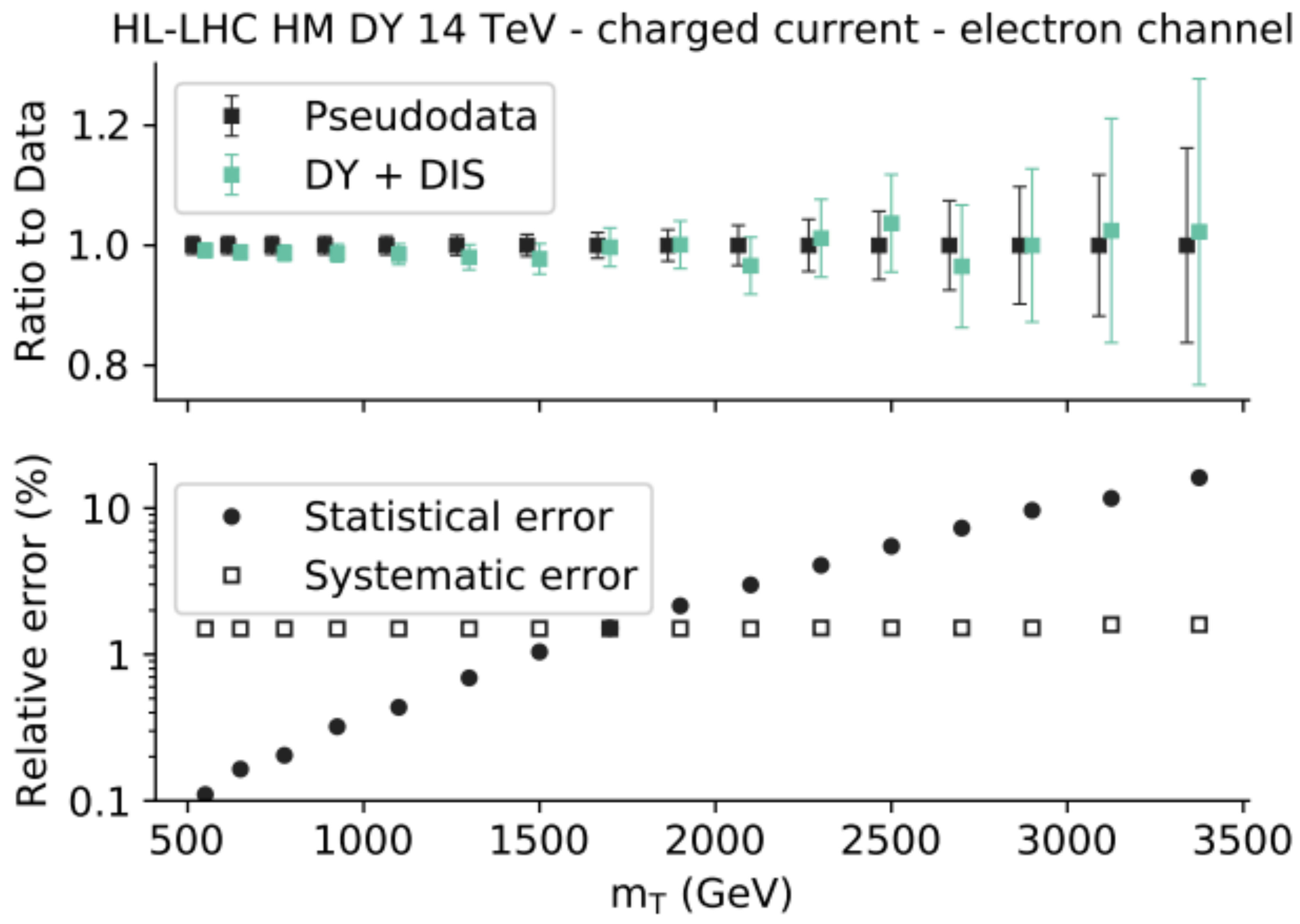
$$\sigma_i^{\rm hllhc} \equiv \sigma_i^{\rm th}\left(1+\lambda\delta_{\mathcal{L}}^{\rm exp}+r_i\delta_{{\rm tot},i}^{\rm exp}\right)\,, \qquad i=1,\ldots,n_{\rm bin}$$

$$\delta_{{\rm tot},i}^{\rm exp}=\left(\left(\delta_i^{\rm stat}\right)^2+\sum_{j=1}^{n_{\rm sys}}\left(f_{{\rm red},j}\delta_{i,j}^{\rm sys}\right)^2\right)^{1/2}$$



Abdul-Khalek et al, 1810.03639

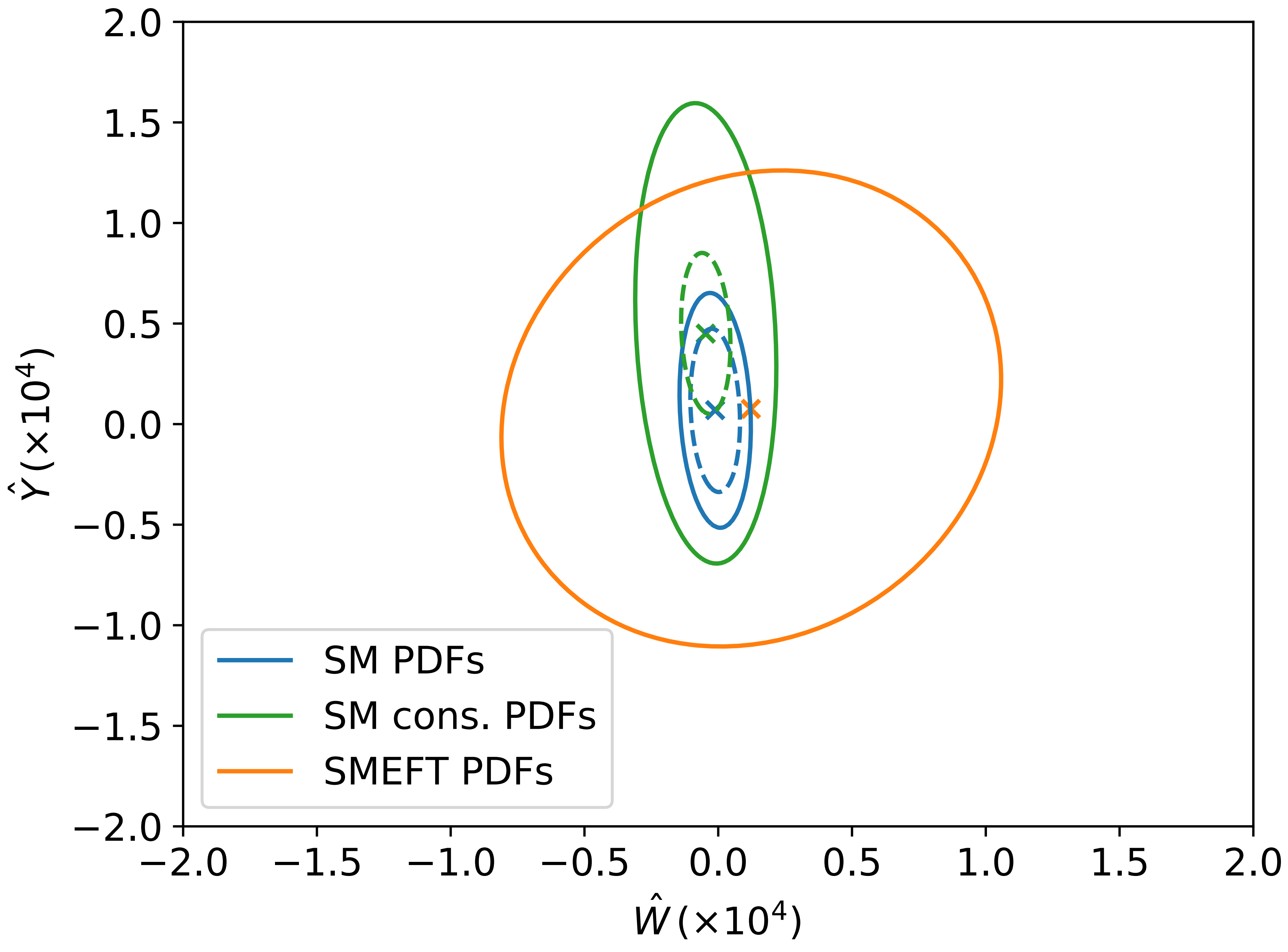


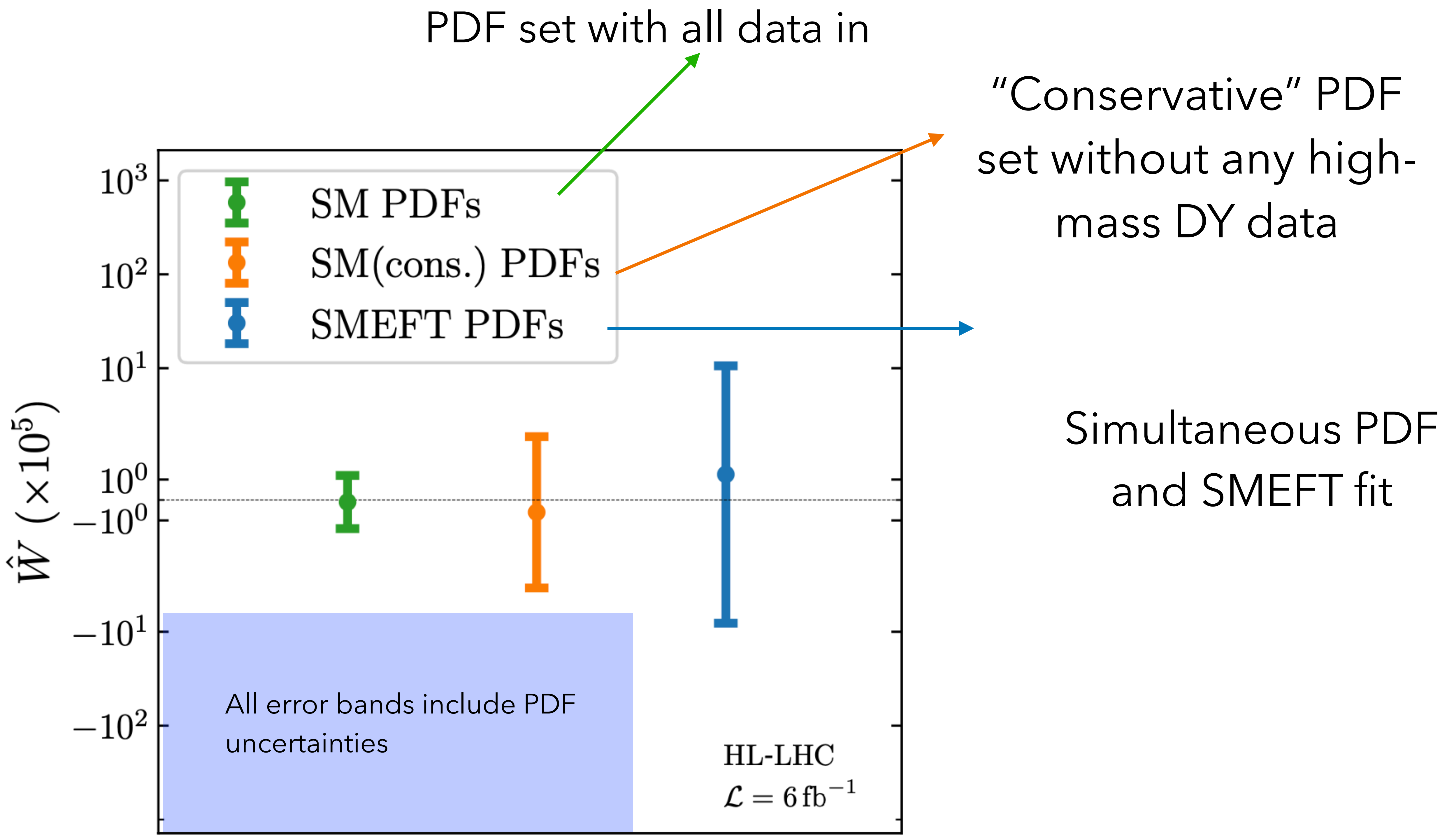


+ muon channel

INTERPLAY @ HL-LHC

- Compare Wilson coefficients bounds from HL-LHC projections assuming SM PDFs (that include NC+CC data) to the bounds on the same Wilson coefficients obtained from a simultaneous fit of PDFs and Wilson coefficients
- Not accounting for interplay (using PDFs as a black box) leads to over-constrained bounds
- PDFs do absorb effect of new physics in this case!



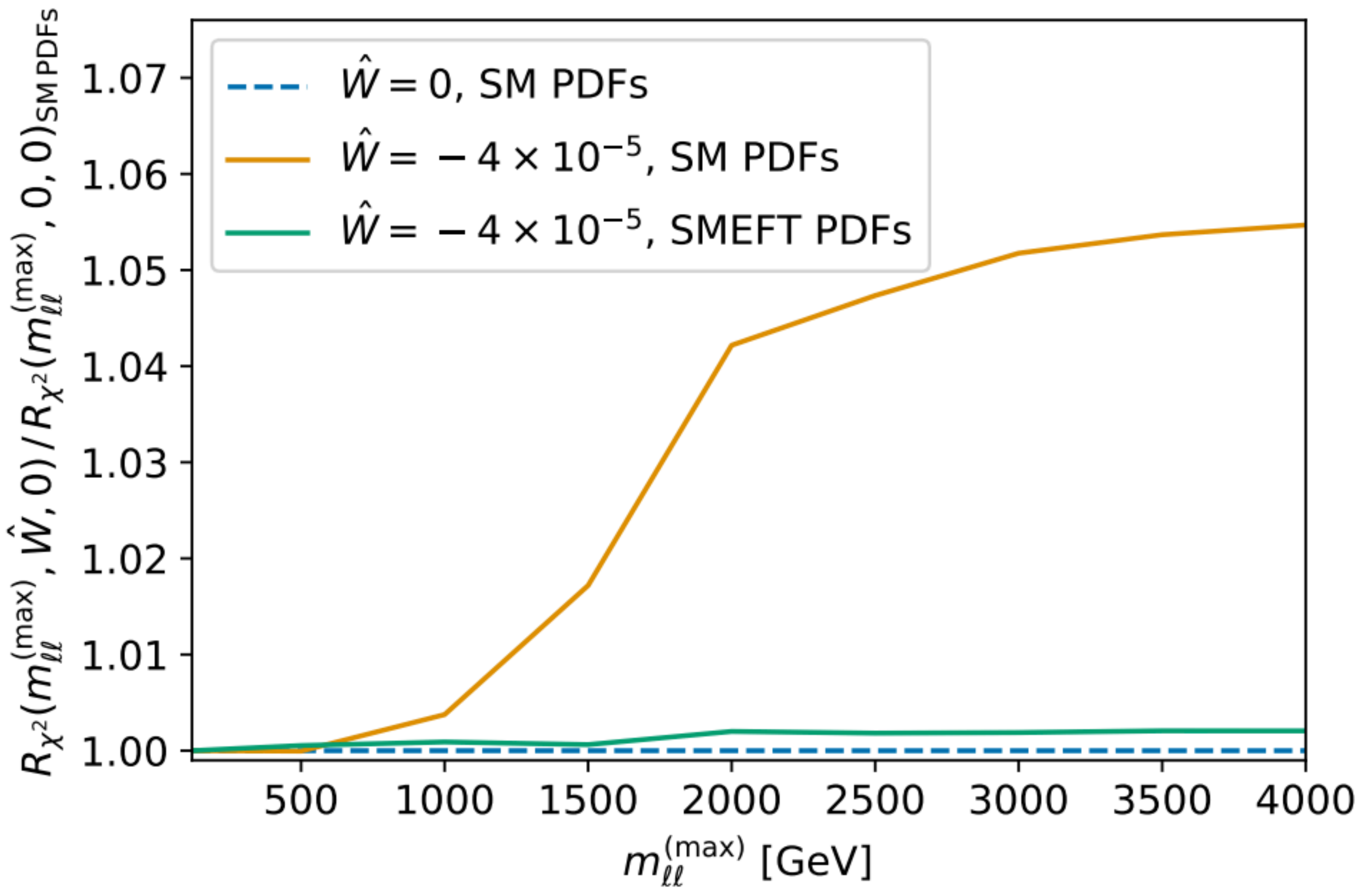


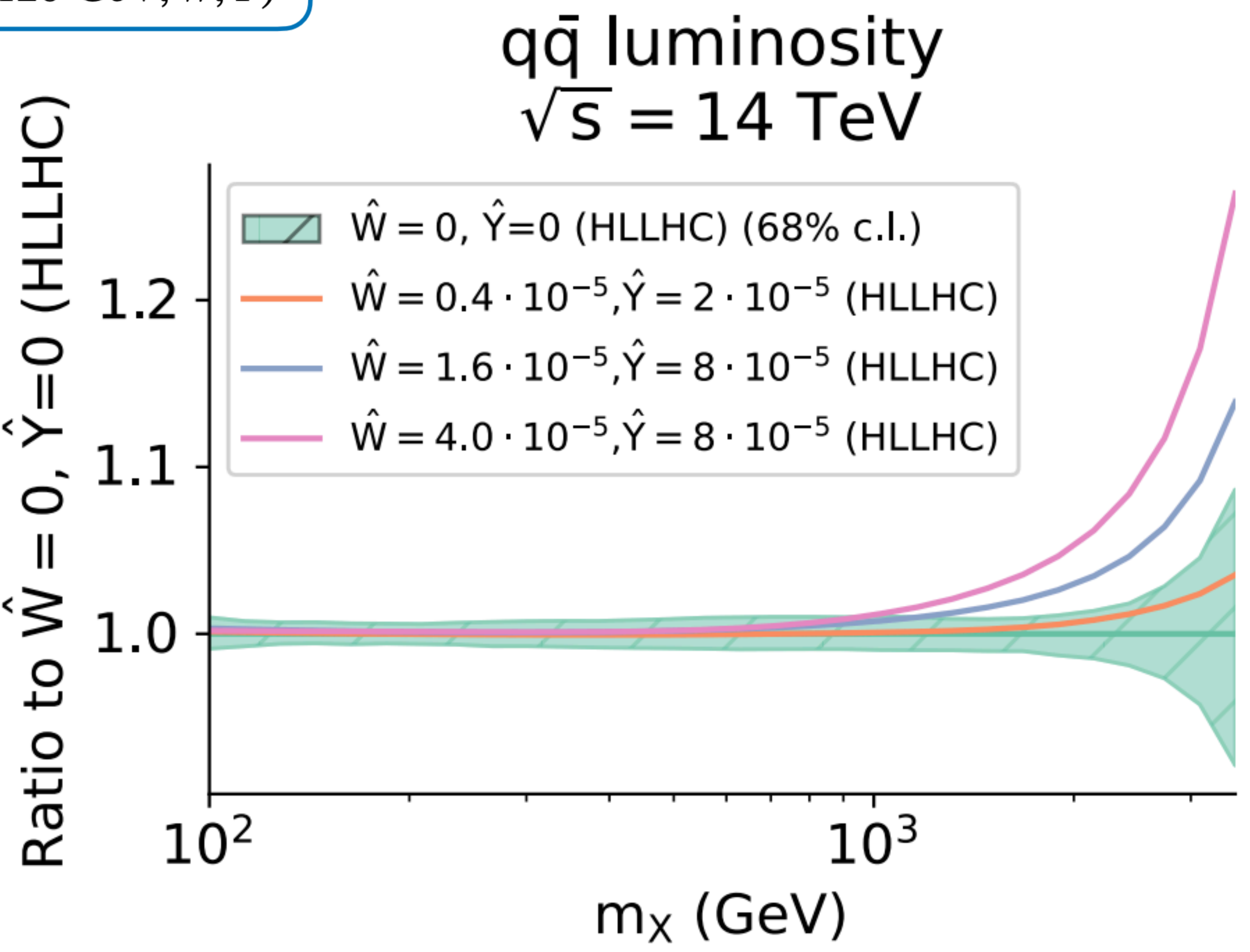
Greljo et al, 2104.02723

INTERPLAY @ HL-LHC

- Compare Wilson coefficients bounds from HL-LHC projections assuming SM PDFs (that include NC+CC data) to the bounds on the same Wilson coefficients obtained from a simultaneous fit of PDFs and Wilson coefficients
- Not accounting for interplay (using PDFs as a black box) leads to over-constrained bounds
- PDFs do absorb effect of new physics in this case!

$$R_{\chi^2}(m_{\ell\ell}^{(\max)}, \hat{W}, \hat{Y}) \equiv \frac{\chi^2(m_{\ell\ell}^{(\max)}, \hat{W}, \hat{Y})}{\chi^2(120\ \mathrm{GeV}, \hat{W}, \hat{Y})}$$





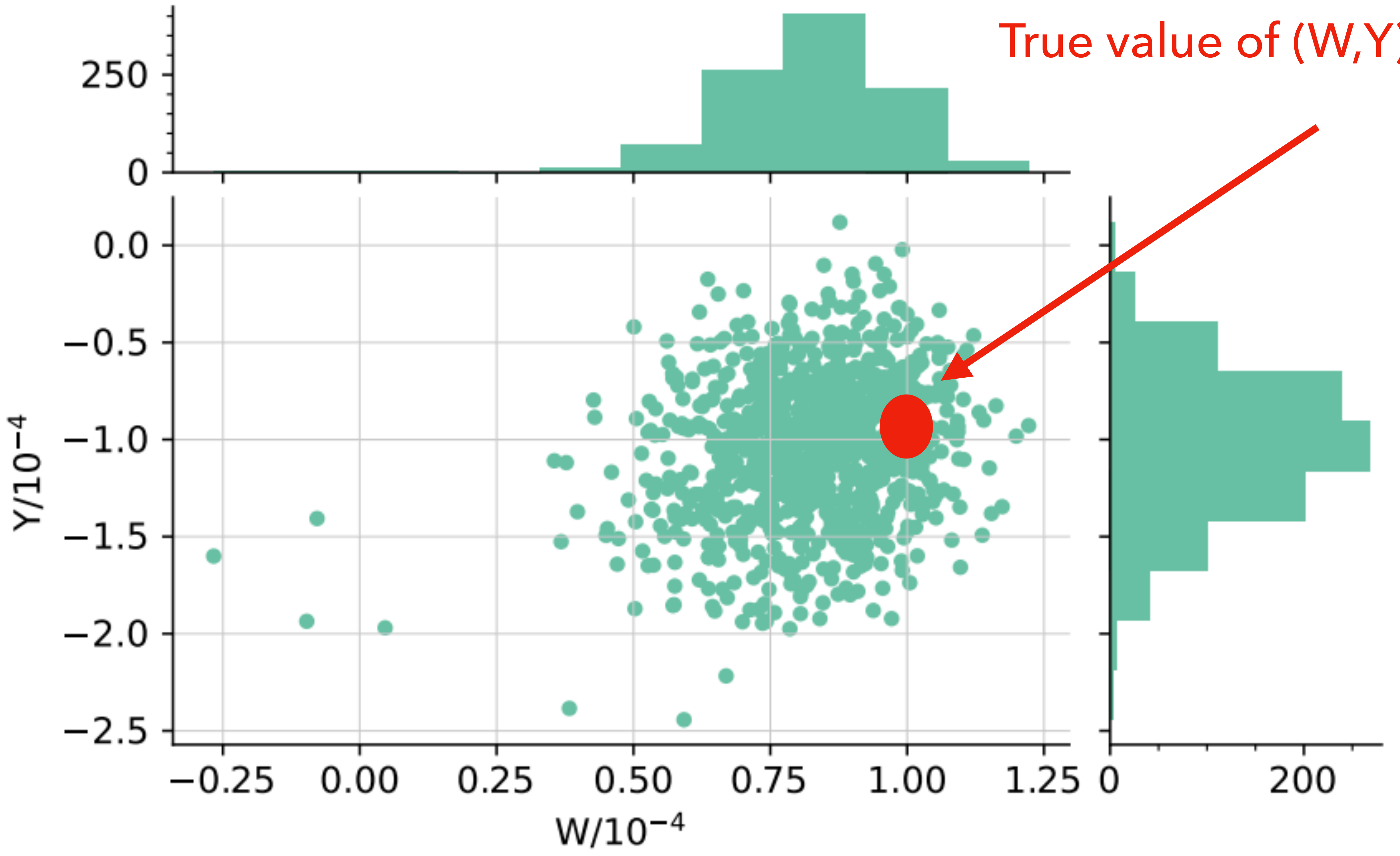
Greljo et al, 2104.02723

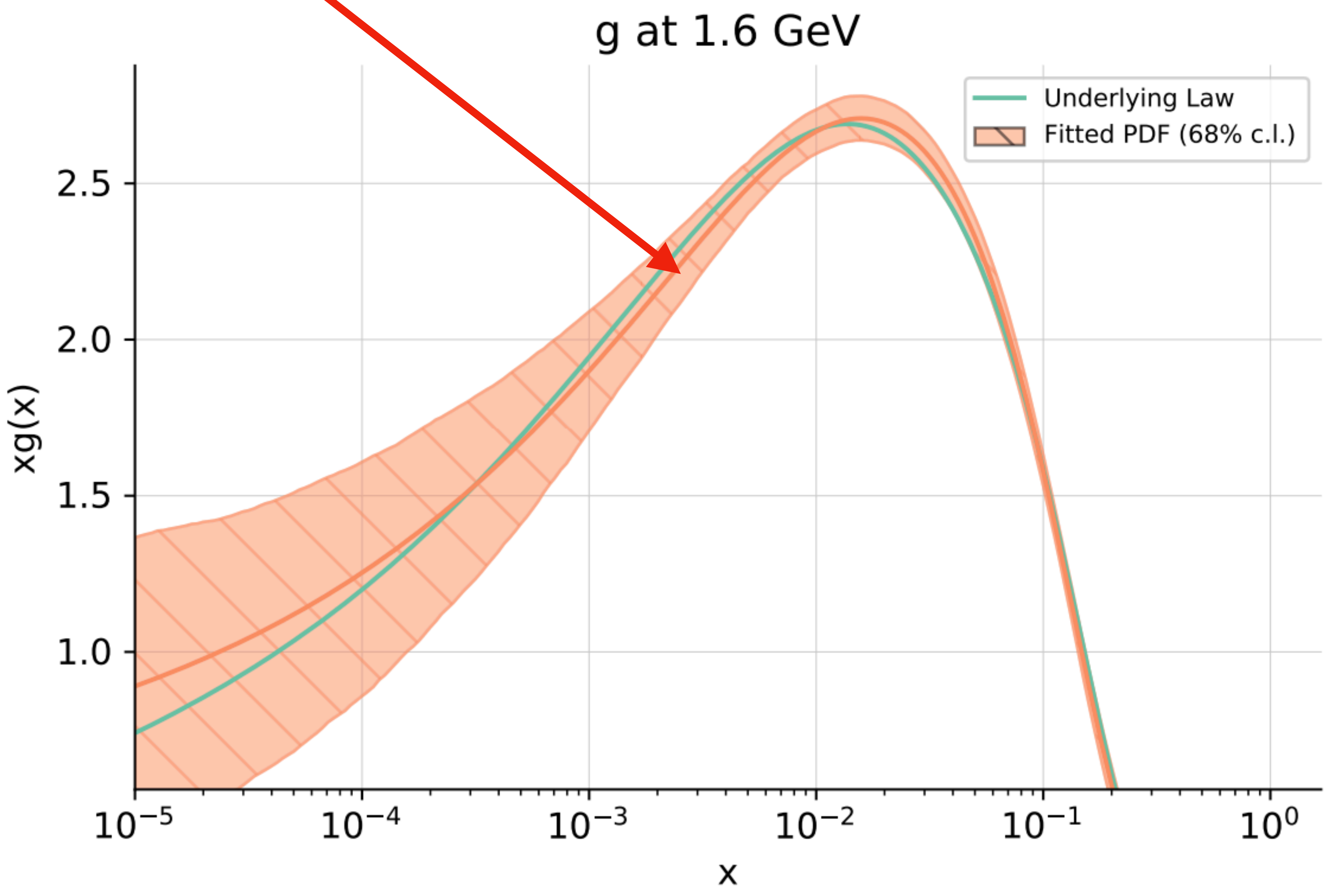
RESULTS: DRELL-YAN DATA @HL-LHC

S. Iranipour, MU - arXiv: 2201.07240

True PDFs = MMHT2020

True value of (W,Y) . 10^5 = (-10,10)





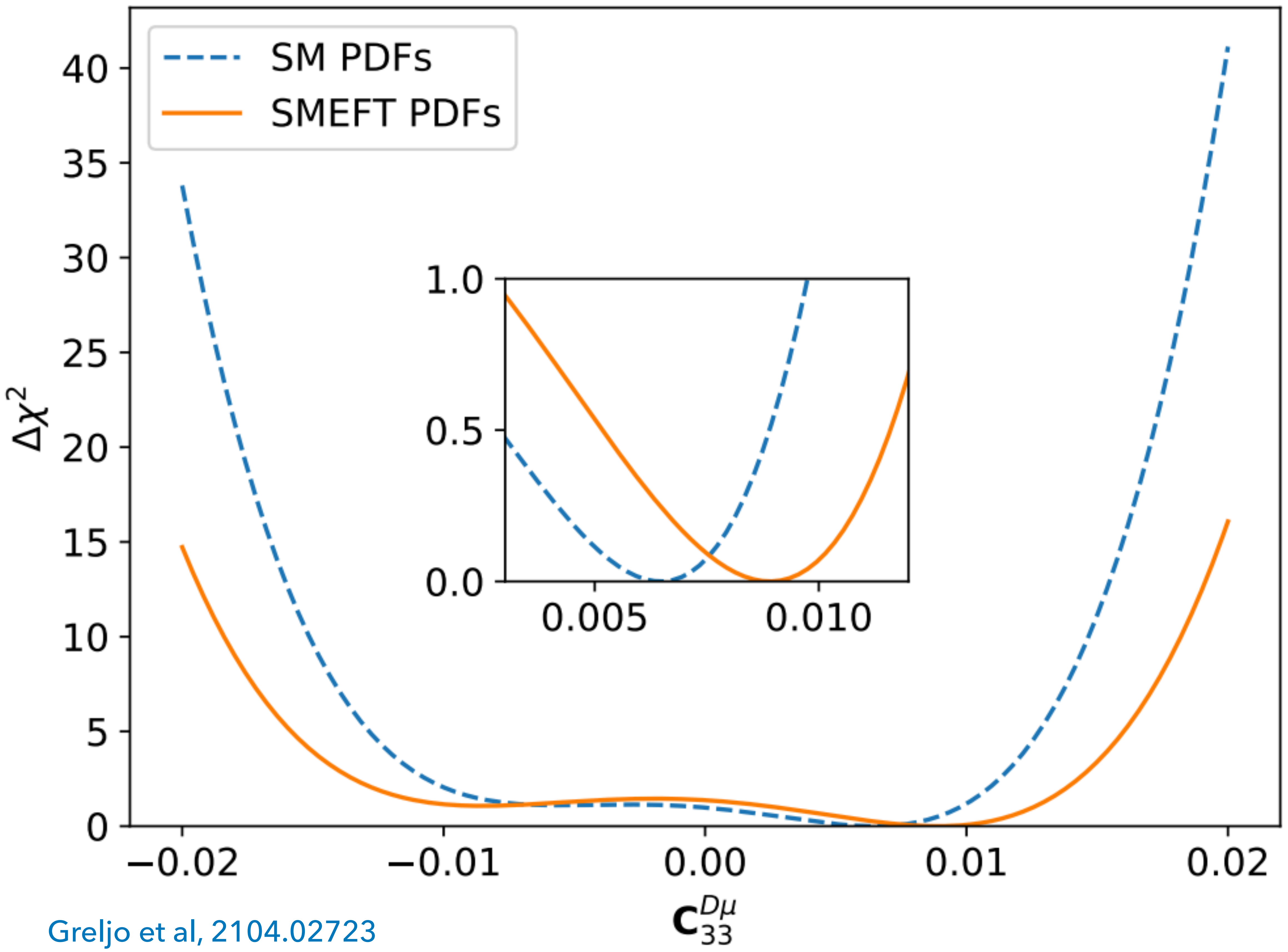
- ✓ Simultaneous analysis of PDFs and W&Y SMEFT coefficient of DIS + DY (including HL-LHC projections) using simuNET method shows that at HL-LHC the effect of interplay becomes important as WCs bounds broaden and PDFs change significantly once SMEFT effects allowed in theory predictions entering PDF fit
- ✓ Stress-tested and shown robustness with closure tests

FLAVOUR-SPECIFIC 4-FERMIONS OPERATORS COUPLING MUONS AND B-QUARKS

- Consider a scenario with a single non-zero WC among gauge invariant four-fermion operators built from the SM quark and lepton SU(2)$_L$ doublets
- If the observed deviations in R(K$^{(*)}$) due to new physics, generically expect $|C^{D\mu}_{33}| \gtrsim 0.001$

$$\mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \frac{\mathbf{C}^{D\mu}_{33}}{v^2}(\bar{d}^3_L \gamma_\mu d^3_L)(\bar{\mu}_L \gamma^\mu \mu_L)$$

	SM PDFs	SMEFT PDFs	best-fit shift	broadening
$\mathbf{C}^{D\mu}_{33} \times 10^2$ (68% CL)	$[-0.1, 1.1]$	$[-0.3, 1.2]$	0.06	25%
$\mathbf{C}^{D\mu}_{33} \times 10^2$ (95% CL)	$[-1.0, 1.2]$	$[-1.2, 1.4]$	0.06	18%

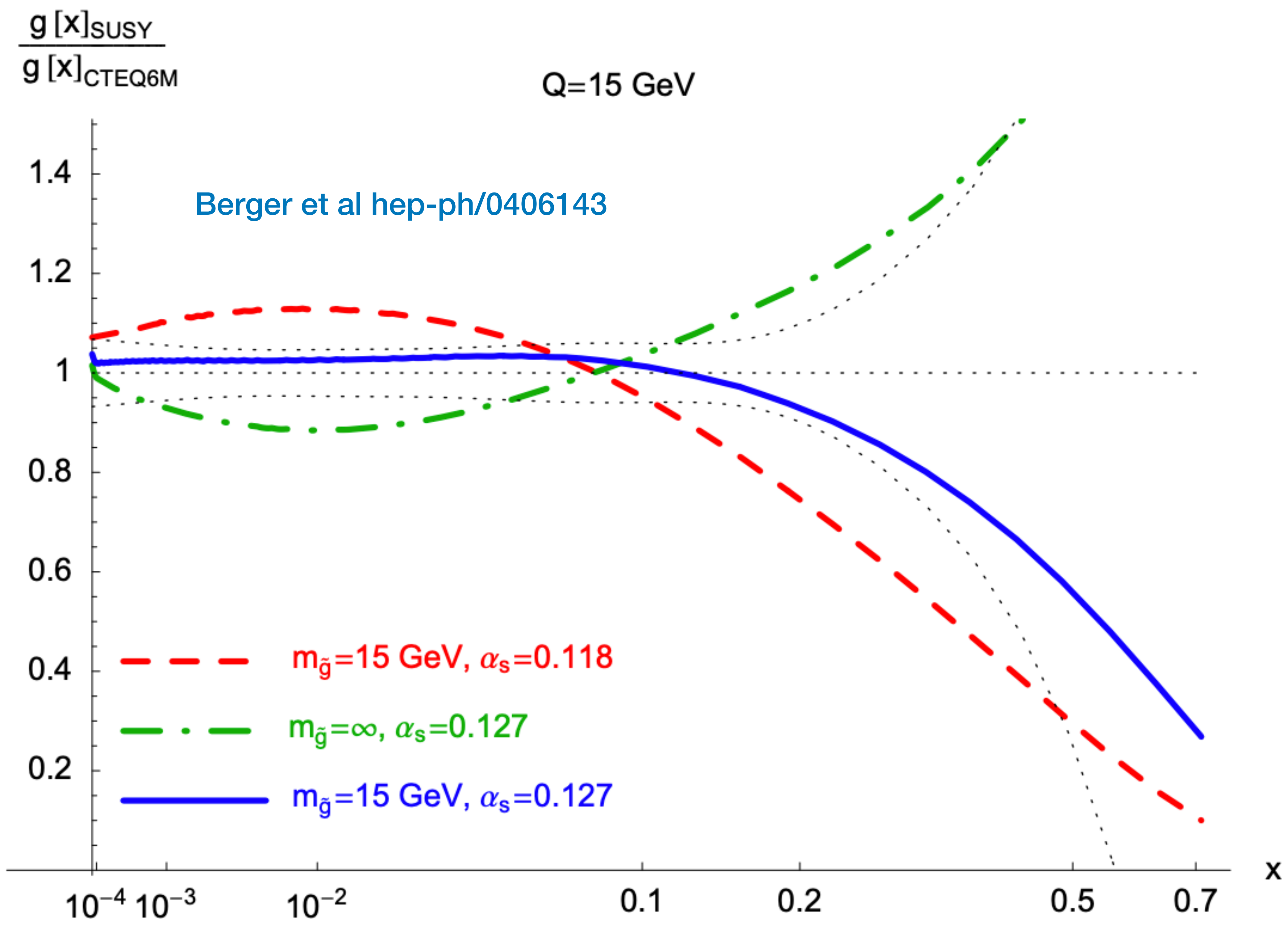


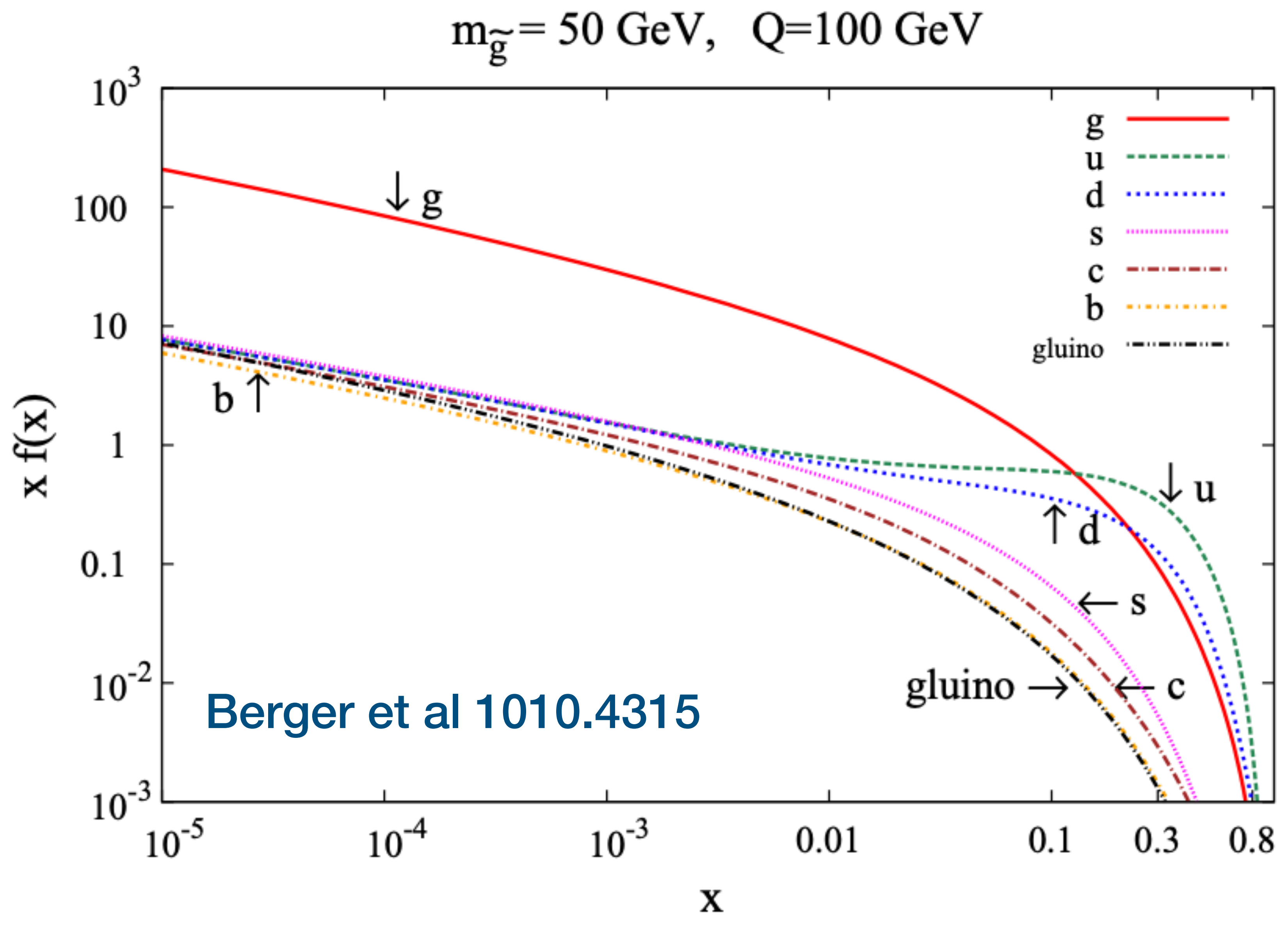
Greljo et al, 2104.02723

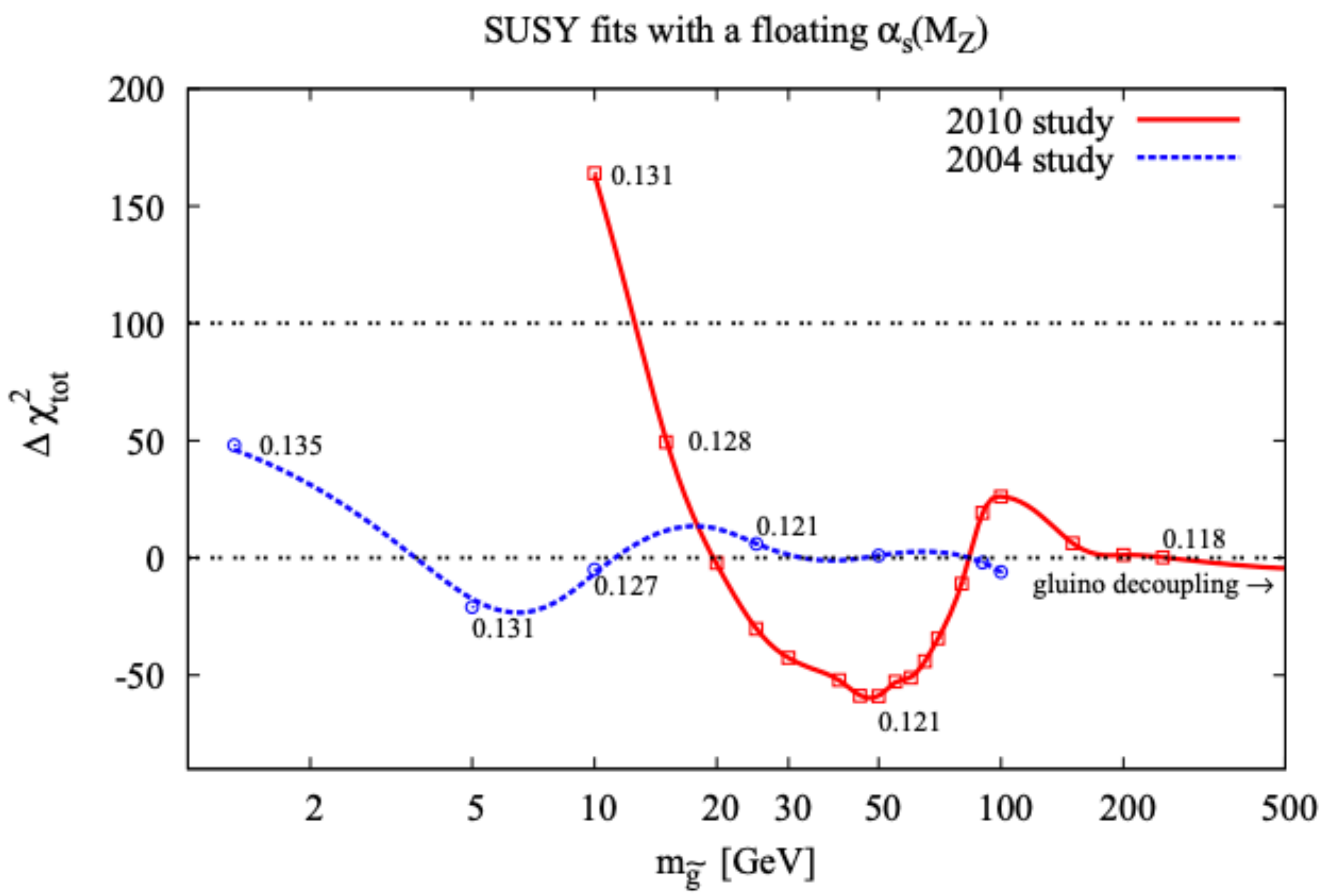
- From PDF point of view, new physics only in Drell-Yan muon data and PDF constrained by Drell-Yan electron data

- Measurements in separate leptonic final states is of utmost importance to test BSM scenarios that account for violations of Lepton Flavour Universality.

DISCOVERY THROUGH PRECISION: A DARK PHOTON EXAMPLE







- Pre-LHC studies: what is there was a light SUSY coloured partner?
- A light SUSY Parton would modify DGLAP equation and running of α_s
- Comparison to data excludes any light coloured parton on increasing mass range as more (and more precise) data are included in the global PDF fit

DISCOVERY THROUGH PRECISION: A DARK PHOTON EXAMPLE

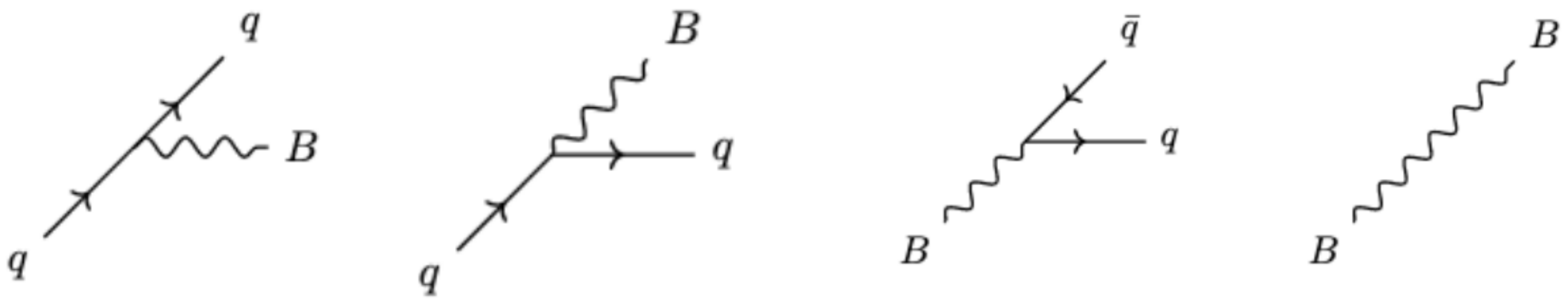
M. McCullough, J. Moore, MU, arXiv:2203.12628

- Idea: now PDFs are known very precisely, and their uncertainties will continue to reduce in the near future with the HL-LHC, could we do the same for a colourless particle too?
- If there was a lepto-phobic dark photon weakly coupled to quarks via effective Lagrangian

$$\mathcal{L}_{\rm int} = \frac{1}{3} g_B \bar{q} \not{B} q \qquad\qquad m_B \in [2,80]\ \text{GeV}$$

it would appear among the partons of the proton.

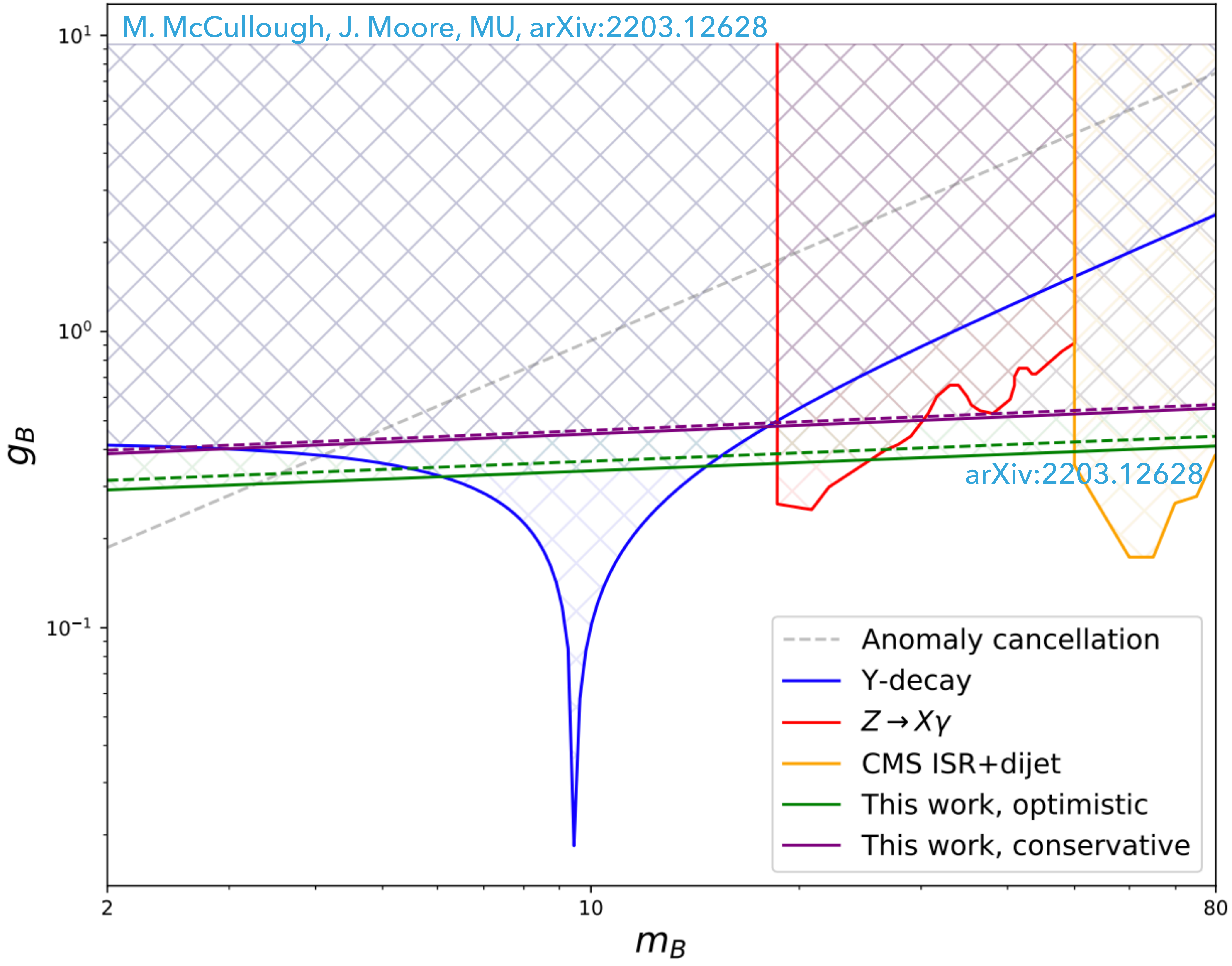
- To include the dark photon as a constituent of the proton: compute the dark photon splitting functions, and add them to DGLAP evolution. Starting from an appropriate initial-scale ansatz (dark photon generated dynamically off quarks and antiquarks at threshold) and a reference PDF set, evolve using the modified DGLAP equations



$$P_{ij} = \left(\frac{\alpha_s}{2\pi}\right) P_{ij}^{(1,0,0)} + \left(\frac{\alpha_s}{2\pi}\right)^2 P_{ij}^{(2,0,0)} + \left(\frac{\alpha_s}{2\pi}\right)^3 P_{ij}^{(3,0,0)}$$
$$+ \left(\frac{\alpha}{2\pi}\right) P_{ij}^{(0,1,0)} + \left(\frac{\alpha_s}{2\pi}\right)\left(\frac{\alpha}{2\pi}\right) P_{ij}^{(1,1,0)} + \left(\frac{\alpha}{2\pi}\right)^2 P_{ij}^{(0,2,0)}$$
$$+ \left(\frac{\alpha_B}{2\pi}\right) P_{ij}^{(0,0,1)} + \cdots ,$$

$\alpha_B \sim 0.001$

DISCOVERY THROUGH PRECISION: A DARK PHOTON EXAMPLE



- The presence of the dark Parton would modify the evolution of standard quarks and gluon.
- Interesting that dark photon effect in PDFs is dominant but was not taken into account in recent publication about evidence for dark photons in DIS data [N. T. Hunt-Smith et al arXiv:2302.11126]

Precise HL-LHC data can indirectly constrain parameter space of the dark photon in a competitive way compared to direct searches

EXTRACTING FUNCTIONS FROM DATA

$$f_i(x, \mu)$$

Data

- Given a finite number of experimental data points D want a set of functions f with errors
- Want to find a infinite-dimensional object from a finite number of information
- Mapping D into f is mathematically ill defined, nobody knows the true f, say $\bar{f}$
- Best we can do it to find best f given the data.
- In Bayesian terms

$$P(f|D) \propto P(D|f)\, P(f)$$

✓ Likelihood ? Prior

$$P(D|f) \sim \exp(-\chi^2[f]/2)$$

$$\chi^2[f] = (D - T[f])\,(\mathrm{cov})^{-1}\,(D - T[f])$$

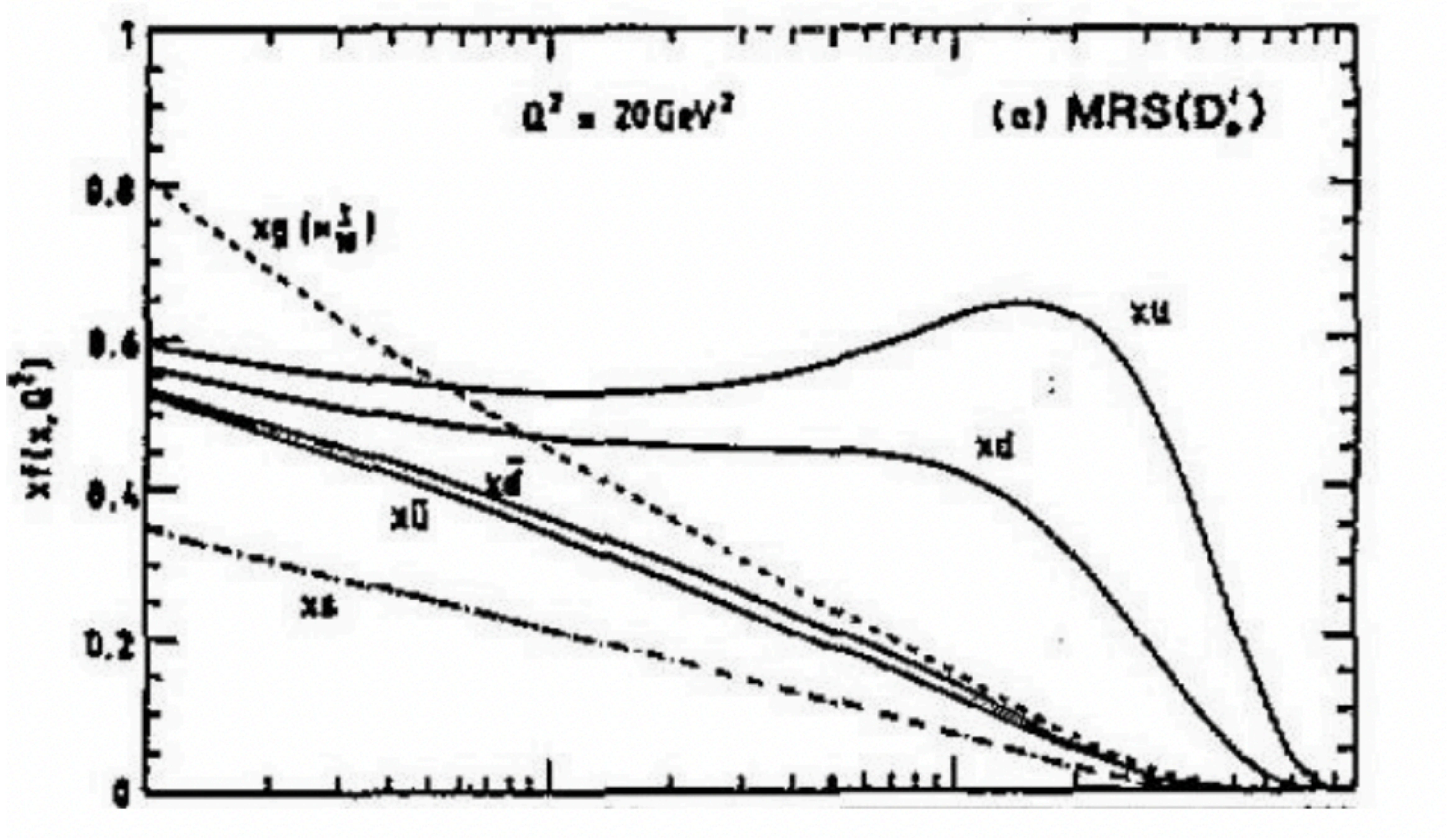
Covariance matrix: experimental errors and correlations, **cov ~ precision**

(D-T) ~accuracy

Two ways of finding the **prior**:

- Explicit parametrization (parametrical modelling)
- Non-parametrical inference (NNs, functional space sampling): use data to also infer probable ‘smoothness’ of f(x) and thus infer P(f) throughout the space of functions

EXTRACTING FUNCTIONS FROM DATA



Martins, Roberts, Stirling (1993)
20 free parameters

- Explicit parametrisation:

$$f(x,\mu_0) = a_0 x^{a_1}(1-x)^{a_2} P(x,\{a_3,...,a_{N_{\rm par}}\})$$

N_{par}: not too many, not too few

- Assume that P[f] has uniform support in this particular space of functions
- Maximise P(f|D) by maximising P(D|f) - maximum likelihood
- Gives best fit PDFs by minimising χ^2 in $\{a_i\}$ space and uncertainties by setting C.L.s on χ^2 contour

$\Delta\chi^2$ = 1 <=> 68% C.L.

<u>Problem:</u> precision << accuracy
Possible reasons: data inconsistencies or over-constraining modelling

<u>Solution:</u> inflate experimental uncertainties by a factor T such that $\Delta\chi^2 = T^2$ until (T~5-10)
precision ~ accuracy

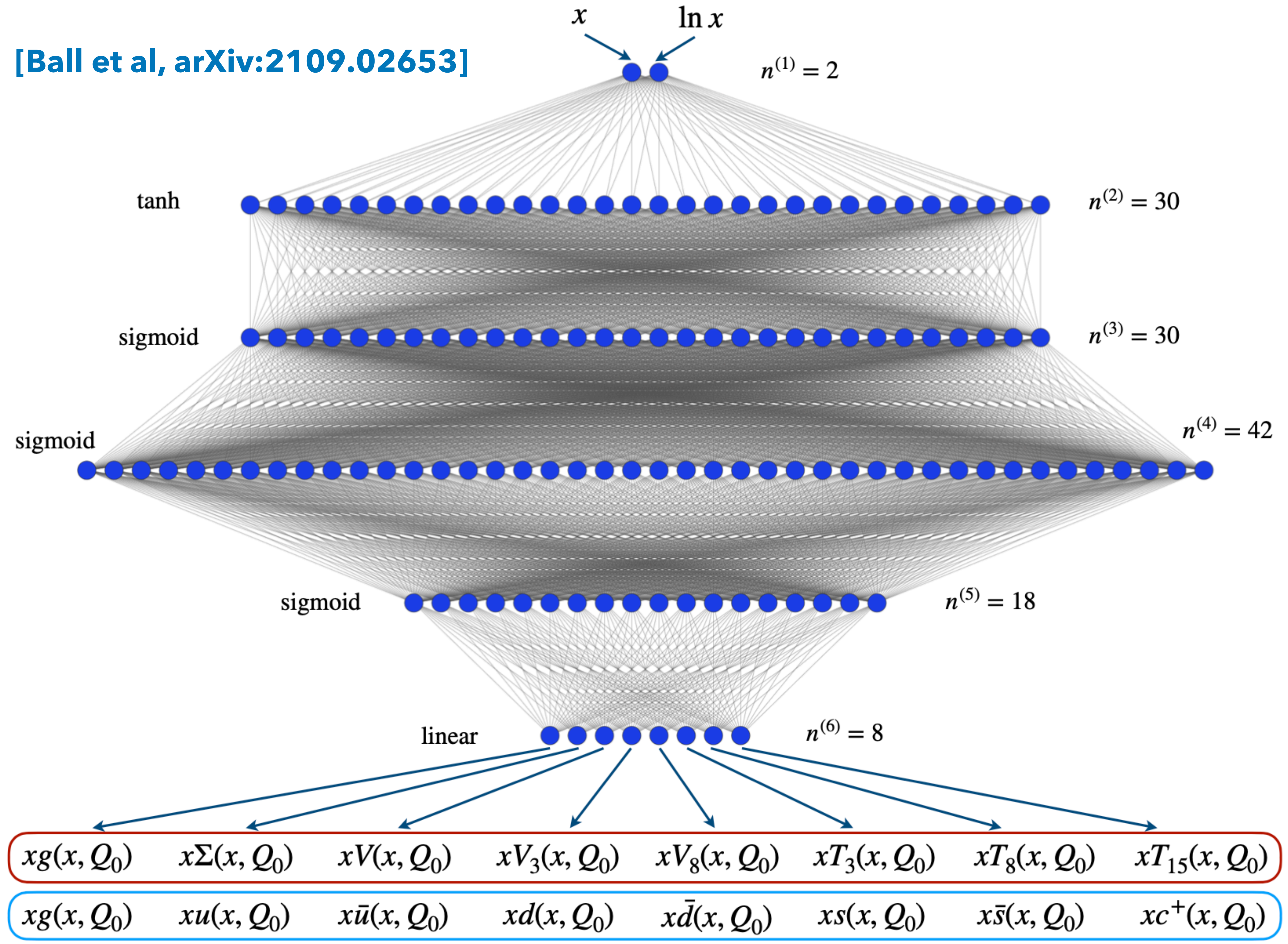
EXTRACTING FUNCTIONS FROM DATA

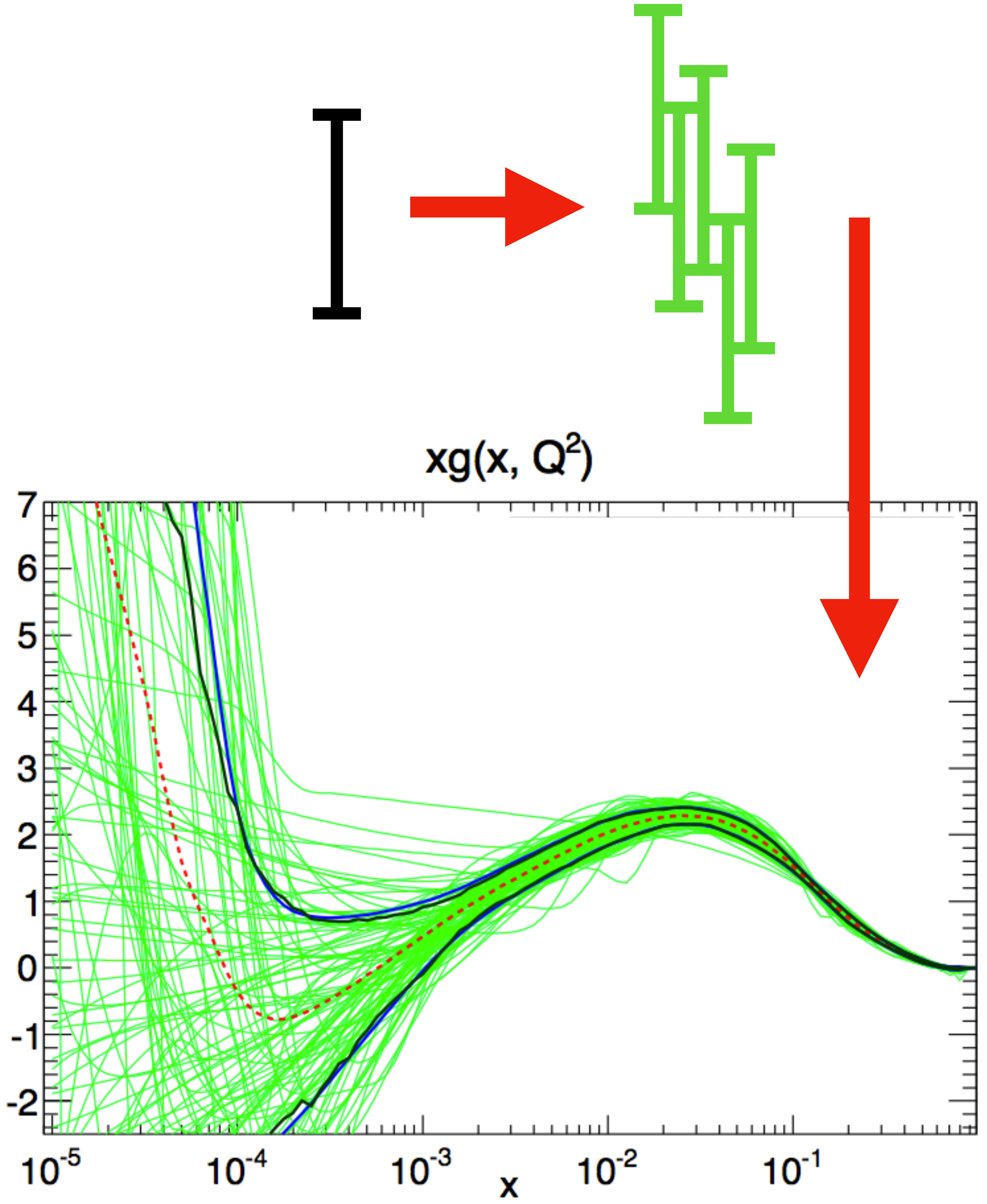
- Alternatively, choose parametrisation so large that in principle can t any conceivable f.
- One option is to use NNs [Forte, Latorre 2002]
- NNDF4.0 parametrisation involves 763 free parameters
- Hessian approach not applicable. Use Monte Carlo - or bootstrap - error propagation by importance sampling

[Giele, Kosover 1993] [Forte et al, 2006]

$$f(x,\mu_0) = x^{\alpha}(1-x)^{\beta}\mathrm{NN}(x)$$

[Ball et al, arXiv:2109.02653]





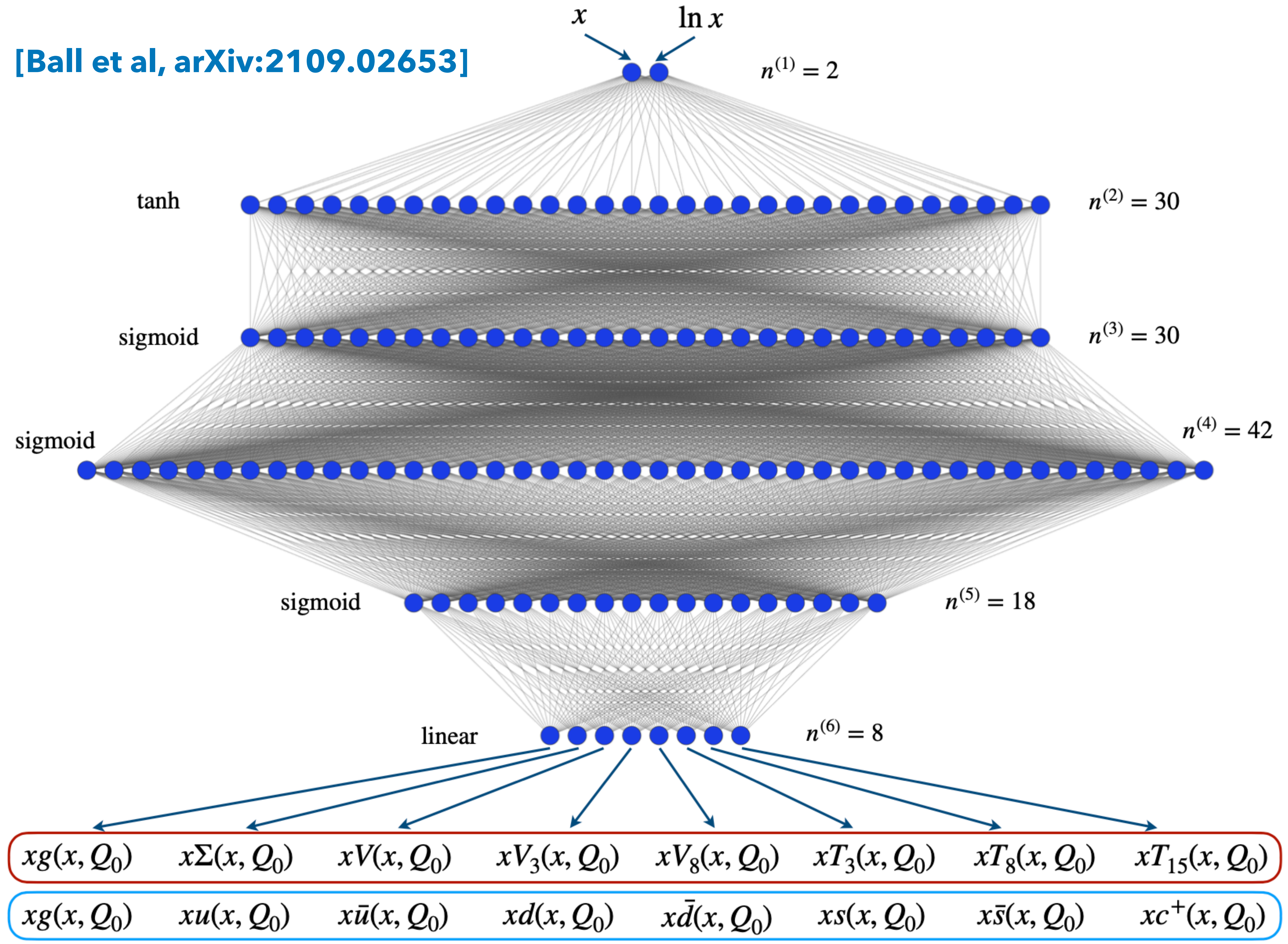
EXTRACTING FUNCTIONS FROM DATA

- Alternatively, choose parametrisation so large that in principle can t any conceivable f.
- One option is to use NNs [Forte, Latorre 2002]
- NNDF4.0 parametrisation involves 763 free parameters
- Hessian approach not applicable. Use Monte Carlo - or bootstrap - error propagation by importance sampling

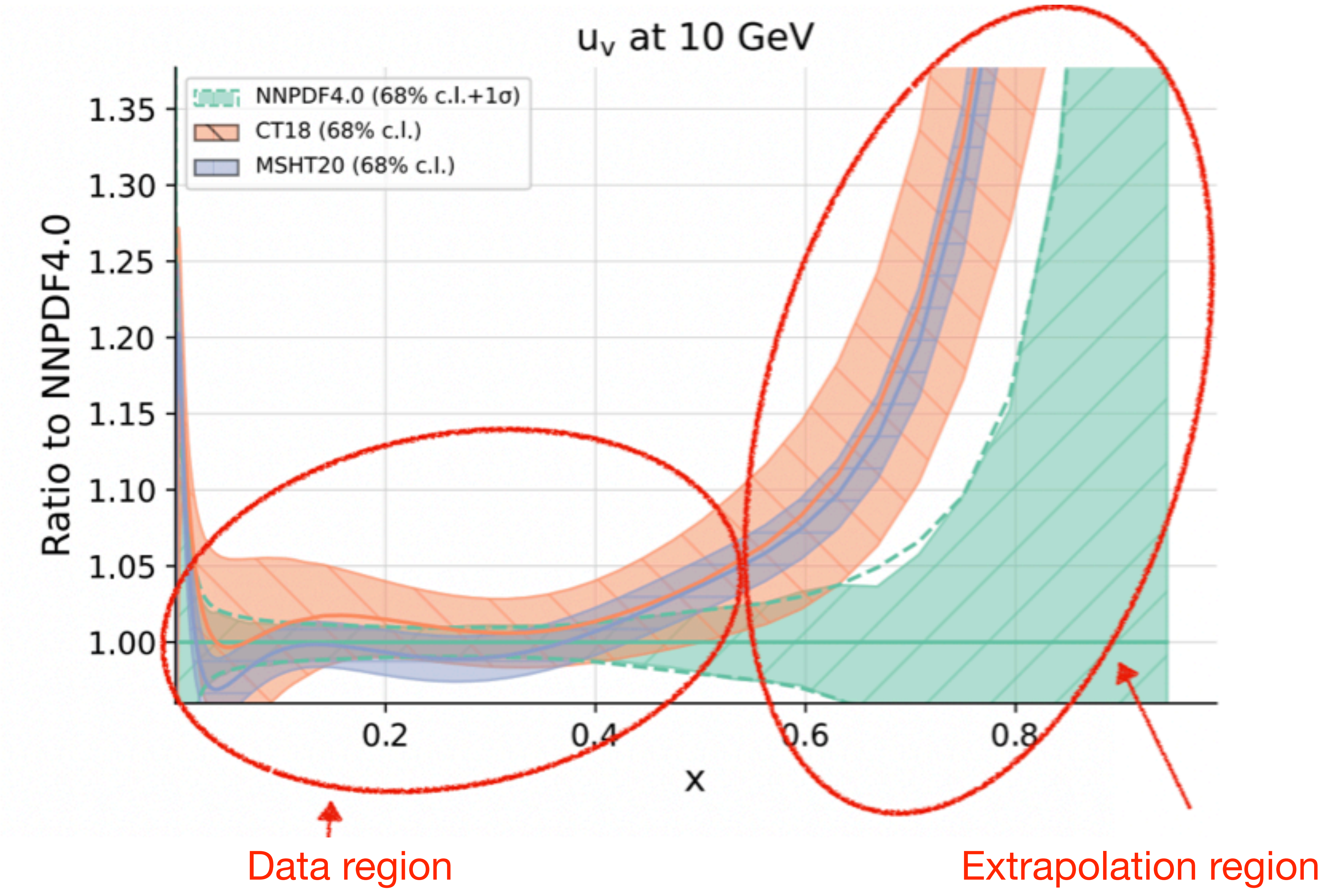
[Giele, Kosover 1993] [Forte et al, 2006]

$$f(x,\mu_0) = x^{\alpha}(1-x)^{\beta}\mathrm{NN}(x)$$

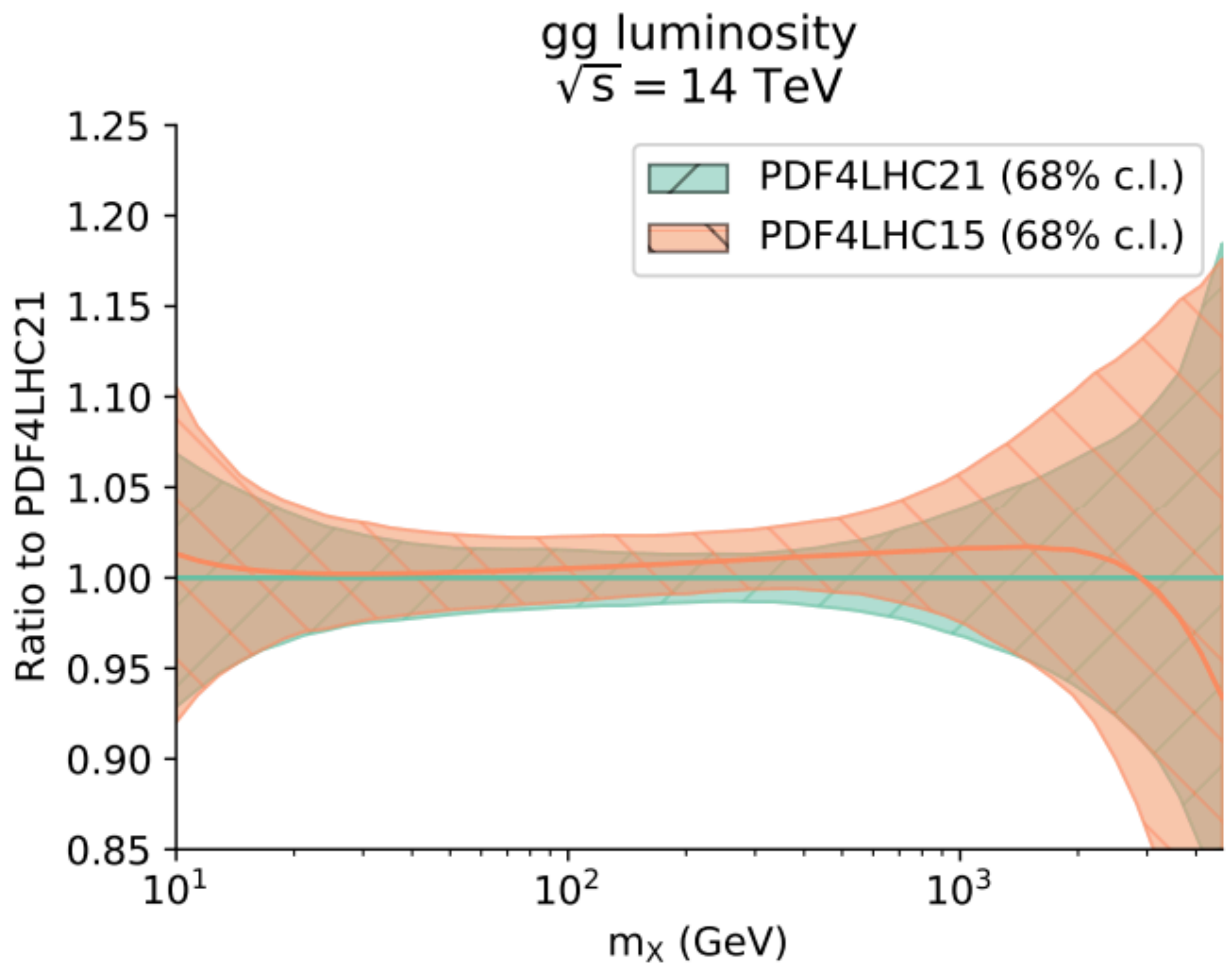
[Ball et al, arXiv:2109.02653]



NN model exhibits smaller uncertainty in data region compared to other sets but larger uncertainties in extrapolation region.

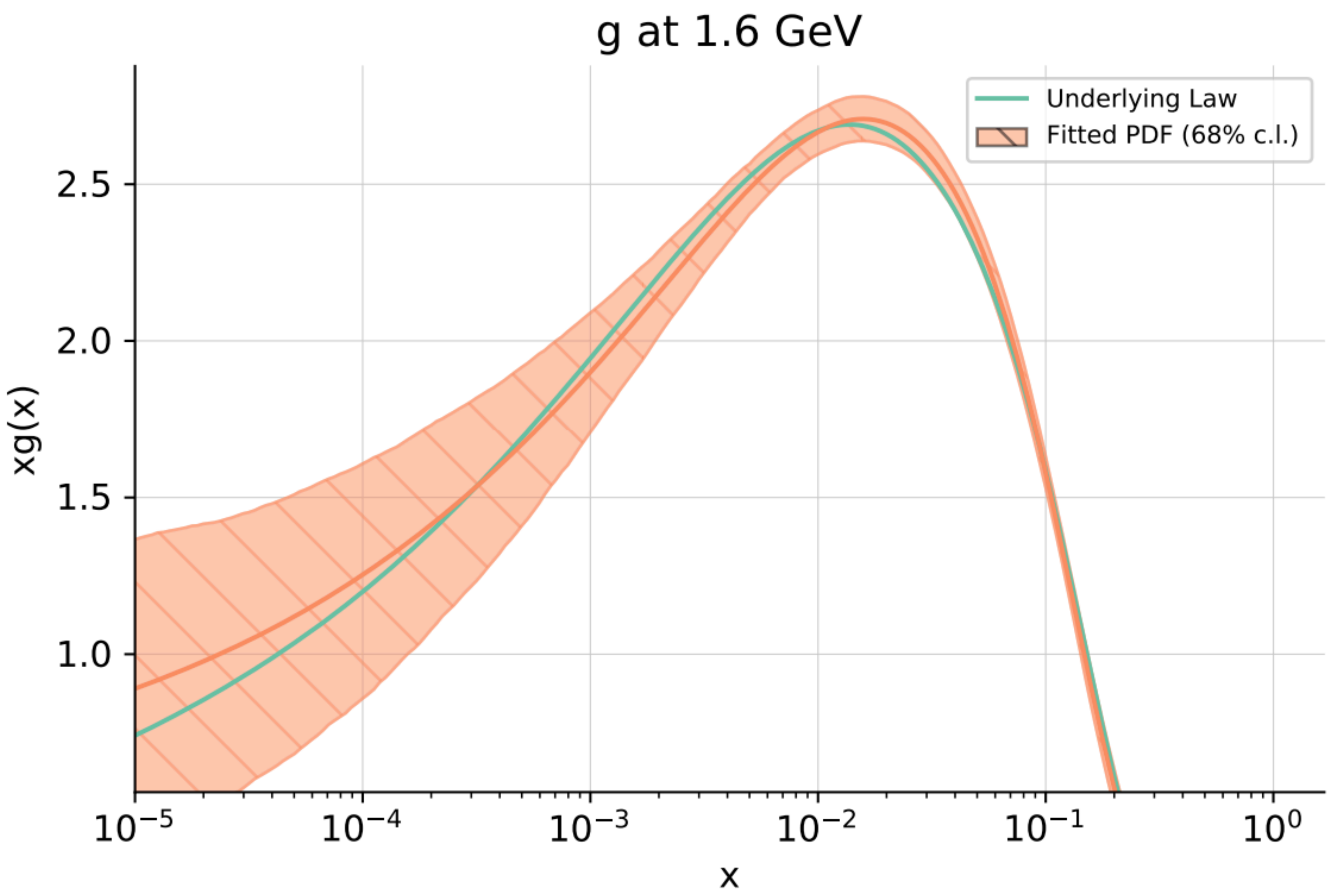


THE PRECISION VERSUS ACCURACY CHALLENGE



Now that PDFs are getting more and more precise, it is crucial for them to

- Genuine inconsistencies in data
- Imperfect fitting methodology
- Inaccurate theoretical framework
  - ➡ Missing higher order uncertainties & N3LO
  - ➡ Other corrections (nuclear, non-perturbative effects…)
  - ➡ BSM effects - second part of the talk



Methodology robustness

- Closure tests for data region: imagine we knew the law of Nature, is fitting methodology able to reproduce it? Is the uncertainty faithful?
  [Del Debbio et al, Eur.Phys.J.C 82 (2022) 4, 330]
  [Barontini et al - in progress]

Generalisation/ extrapolation

- Future tests: how well do PDF describe data that are not included in the fit (either in data or extrapolation region)? Help to discriminate among PDF sets
  [J. Cruz-Martinez et al, Acta Phys.Polon.B 52 (2021) 243]
  [M. Costantini et al - in progress]