European Research Council

Established by the European Commission

PBSP

MARIA UBIALI

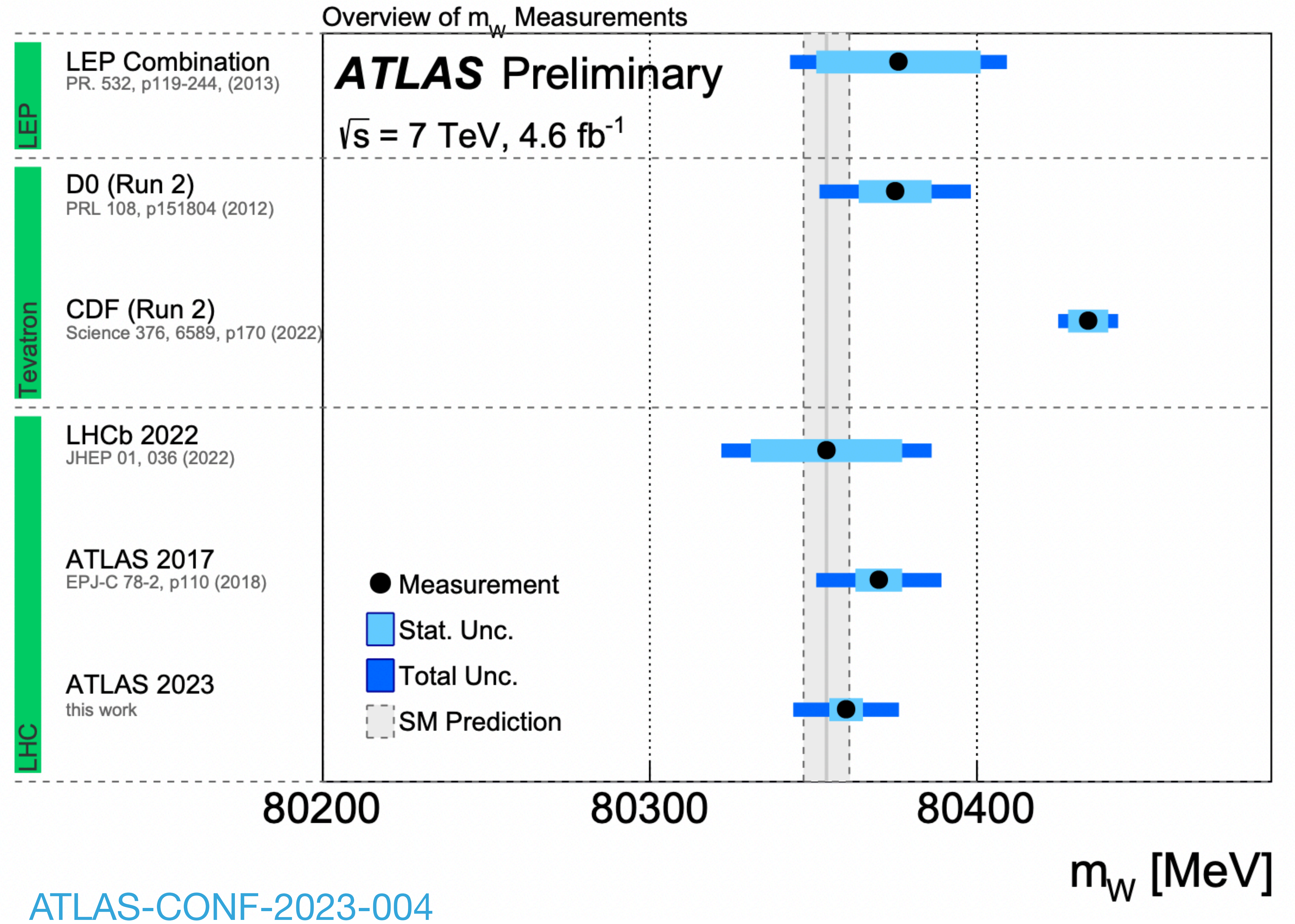
UNIVERSITY OF CAMBRIDGE

INTERPLAY BETWEEN PDF FITS AND NEW PHYSICS

INFN LABORATORI NAZIONALI DI FRASCATI

26TH OCTOBER 2023

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.

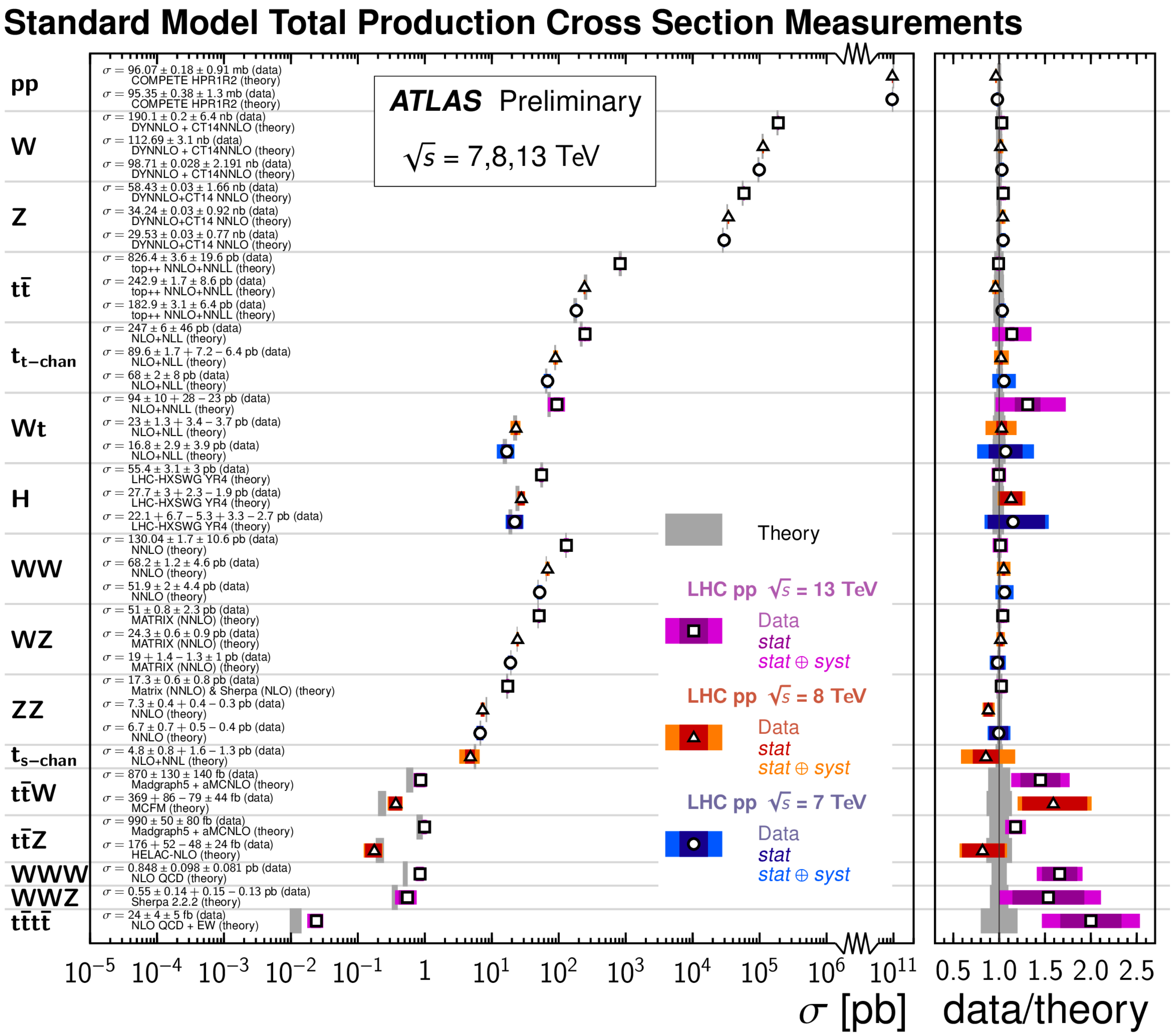
OUTLINE

- Introduction:
  - ➡ Precision for new physics (A dark photon example) M. McCullough, J. Moore, MU - arXiv:2203.12628
  - ➡Progress in PDF fits and SMEFT fits
  - ➡ PDF and SMEFT fits: time to study their interplay
- Simultaneous fits of PDFs and SMEFT Wilson coefficients
  - ➡ SimuNet: a tool for global simultaneous fits - S. Iranipour, MU - arXiv: 2201.07240
  - ➡ A Drell-Yan-sector analysis - A. Greljo, S. Iranipour, Z. Kassabov, M. Madigan, J. Moore, J. Rojo, MU - arXiv: 2104.02723
  - ➡ A global top-sector analysis - Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU, C. Voisy - arXiv: 2303.06159
- Can PDFs absorb New Physics? - E. Hammou, M. Madigan, M. Mangano, L. Mantani, J. Moore, M. Morales, MU - arXiv:2307.10370
- Conclusions and outlook

INTRODUCTION

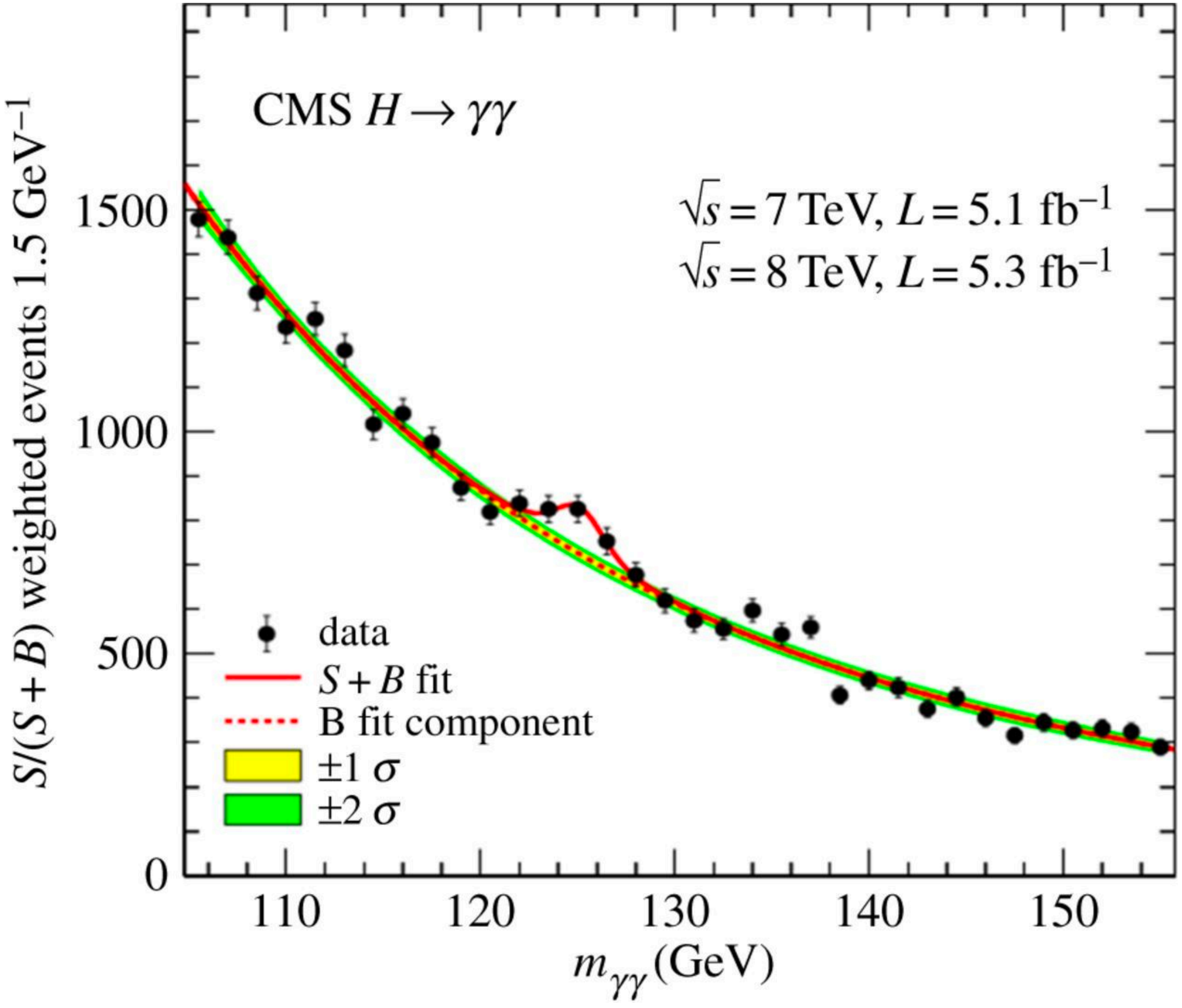
A NEW PRECISION ERA AT THE LHC

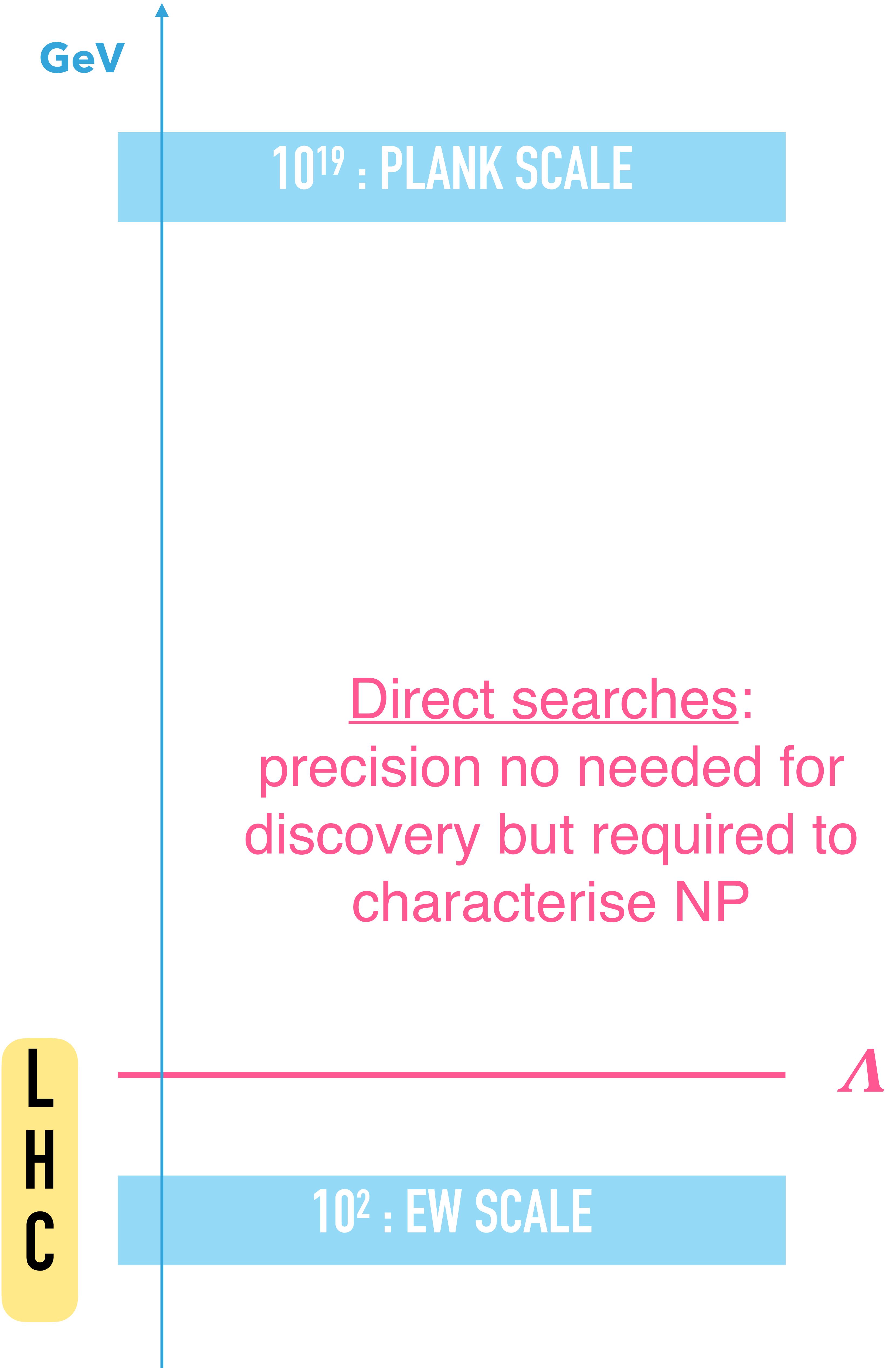
- Impressive range of precise cross section measurements and huge progress in theoretical calculations
- Overall stunning agreement between data and SM predictions although we know that SM is necessarily incomplete



LHC paradigm: discovery → discovery through precision

DISCOVERY THROUGH PRECISION

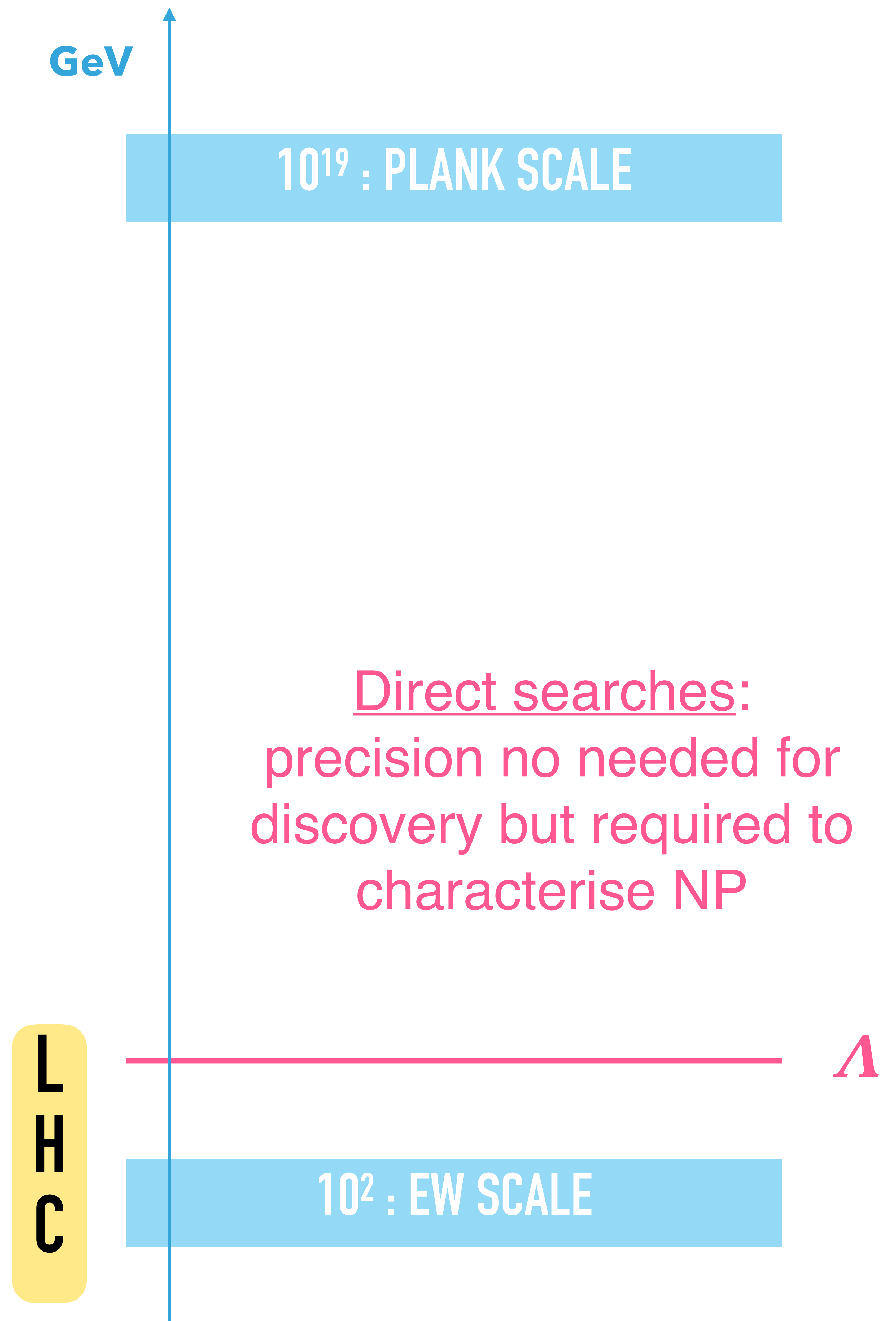




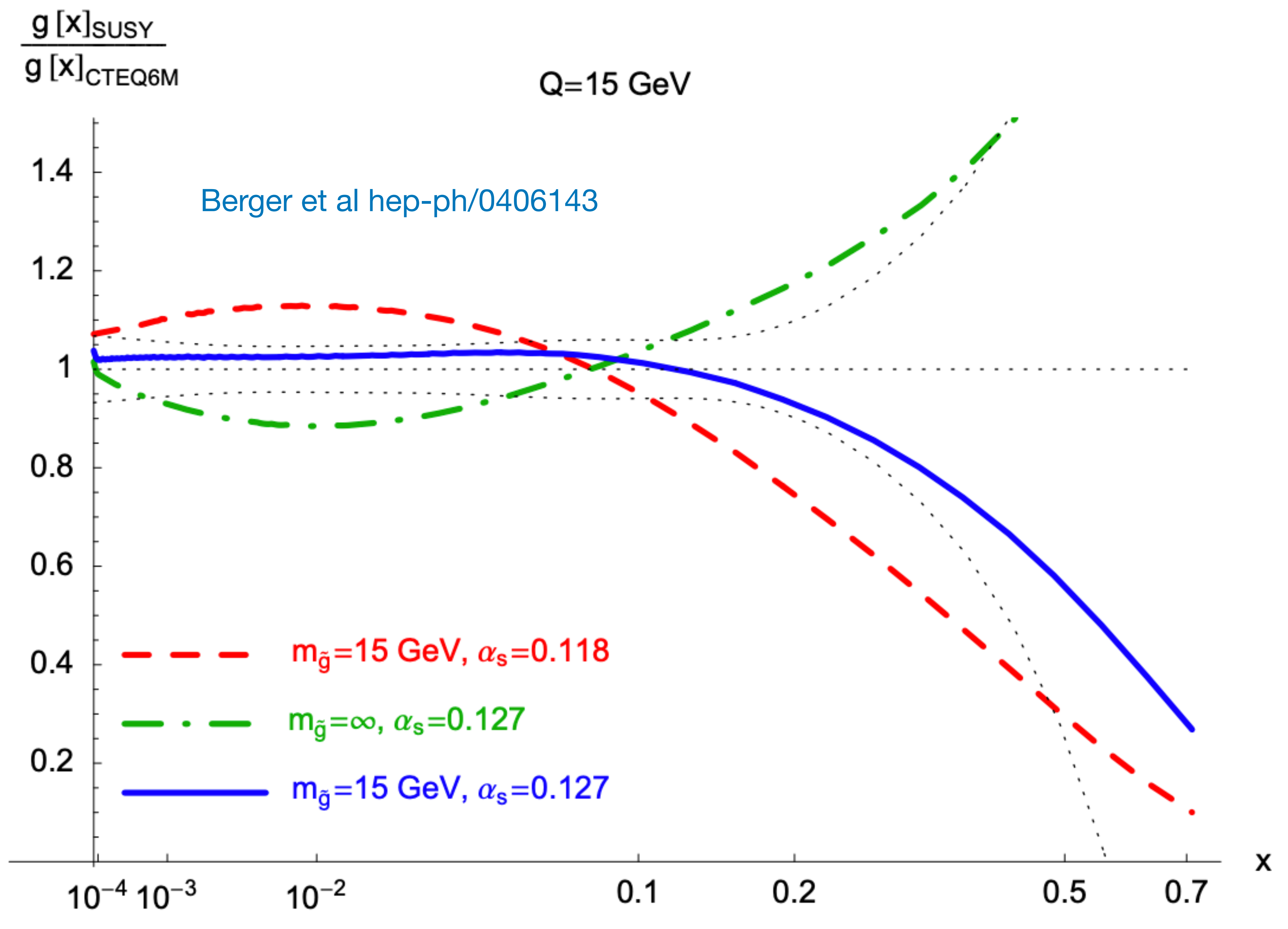
Direct searches:
precision no needed for discovery but required to characterise NP

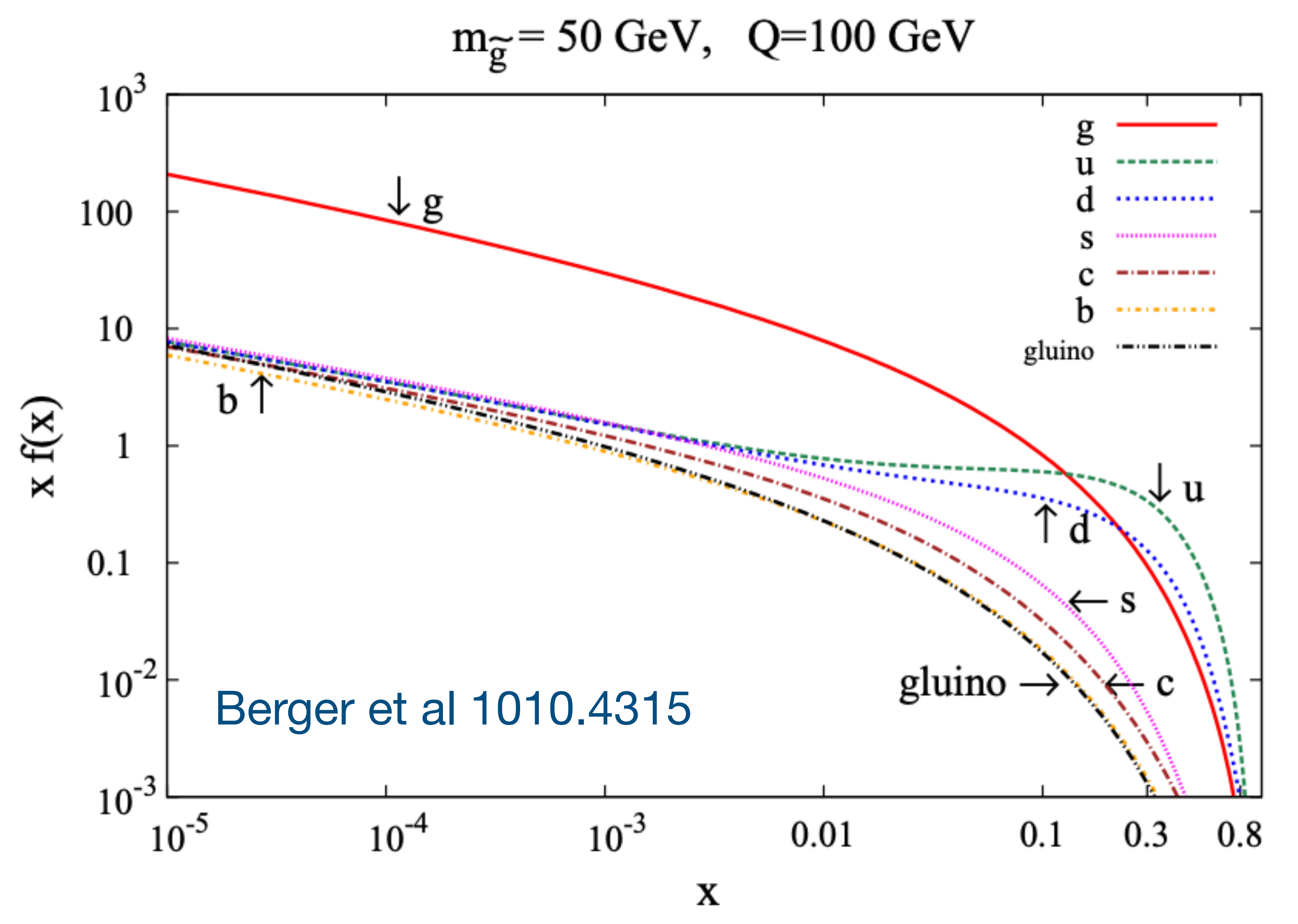
DISCOVERY THROUGH PRECISION

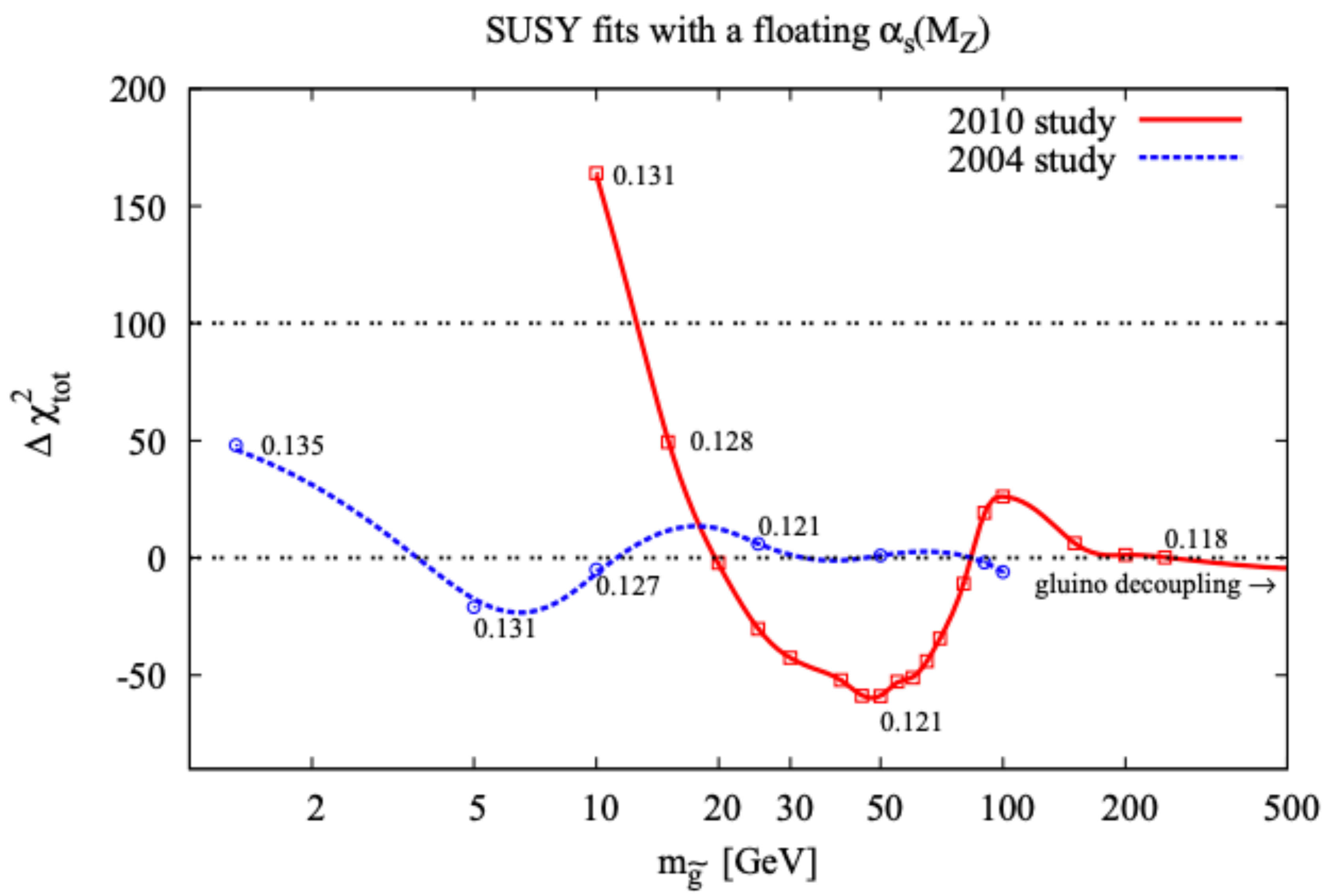
What if new physics was very weakly coupled to the SM?



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

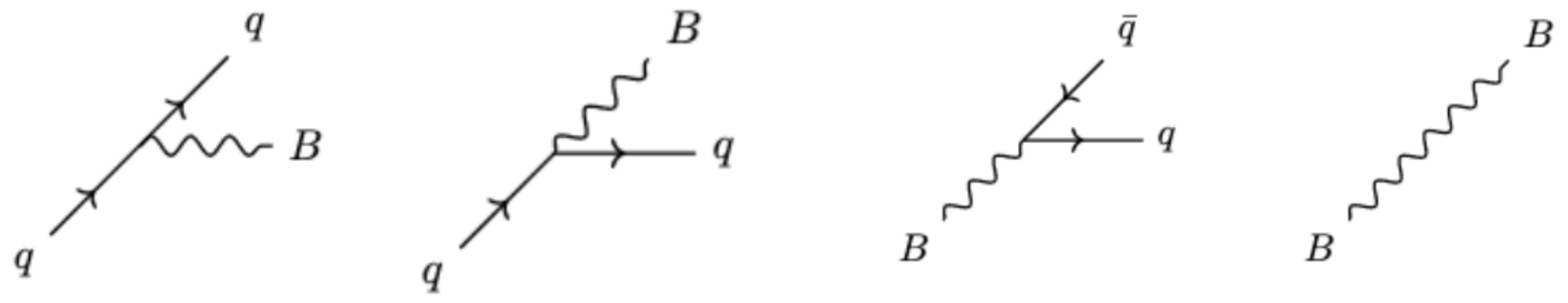
M. McCullough, J. Moore, MU, arXiv:2203.12628

- 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

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

it would appear among the partons of the proton.

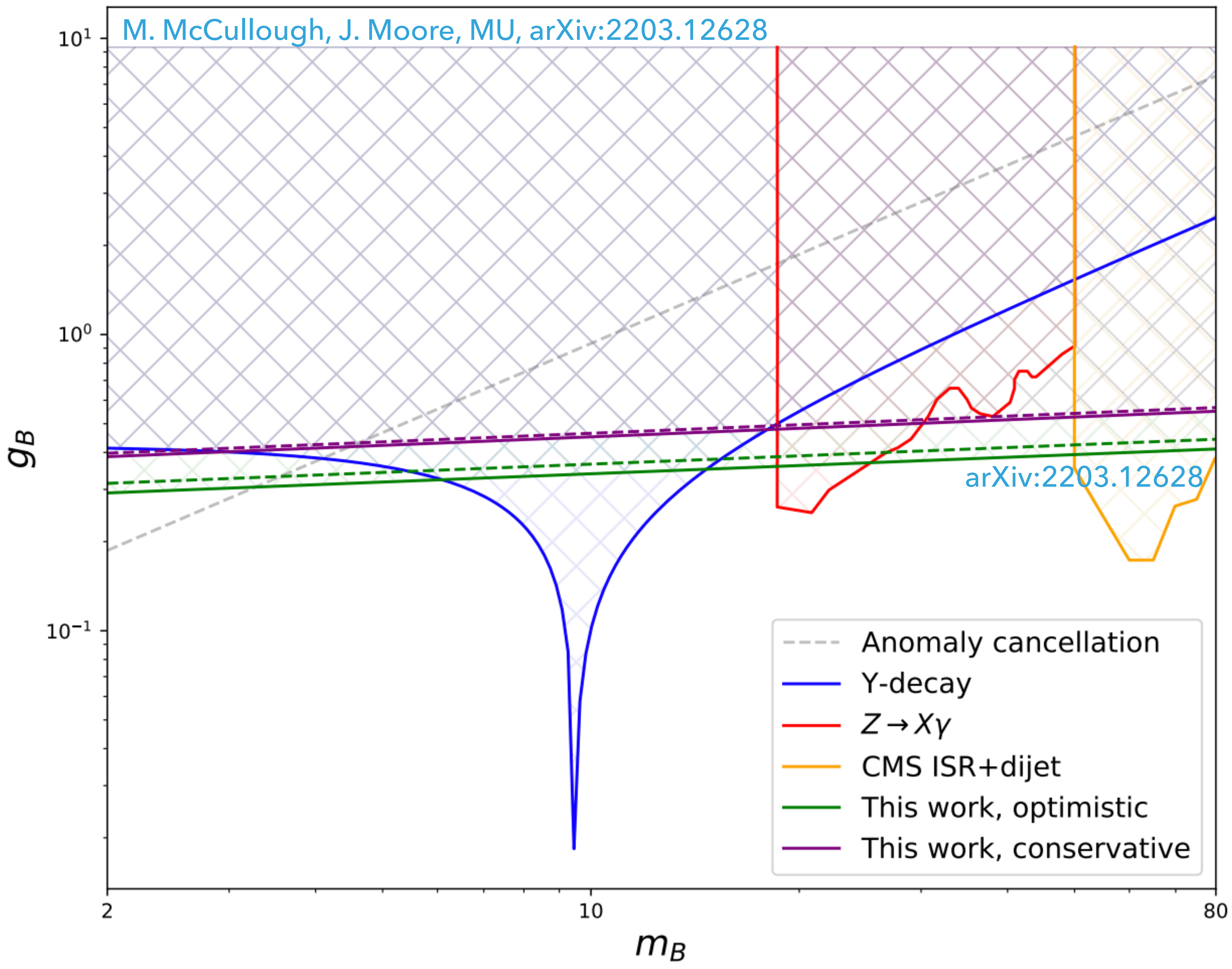
- To include the dark photon as a constituent of the proton: compute the dark photon splitting functions, and add them to DGLAP evolution. Starting from an appropriate initial-scale ansatz (dark photon generated dynamically off quarks and antiquarks at threshold) and a reference PDF set, evolve using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

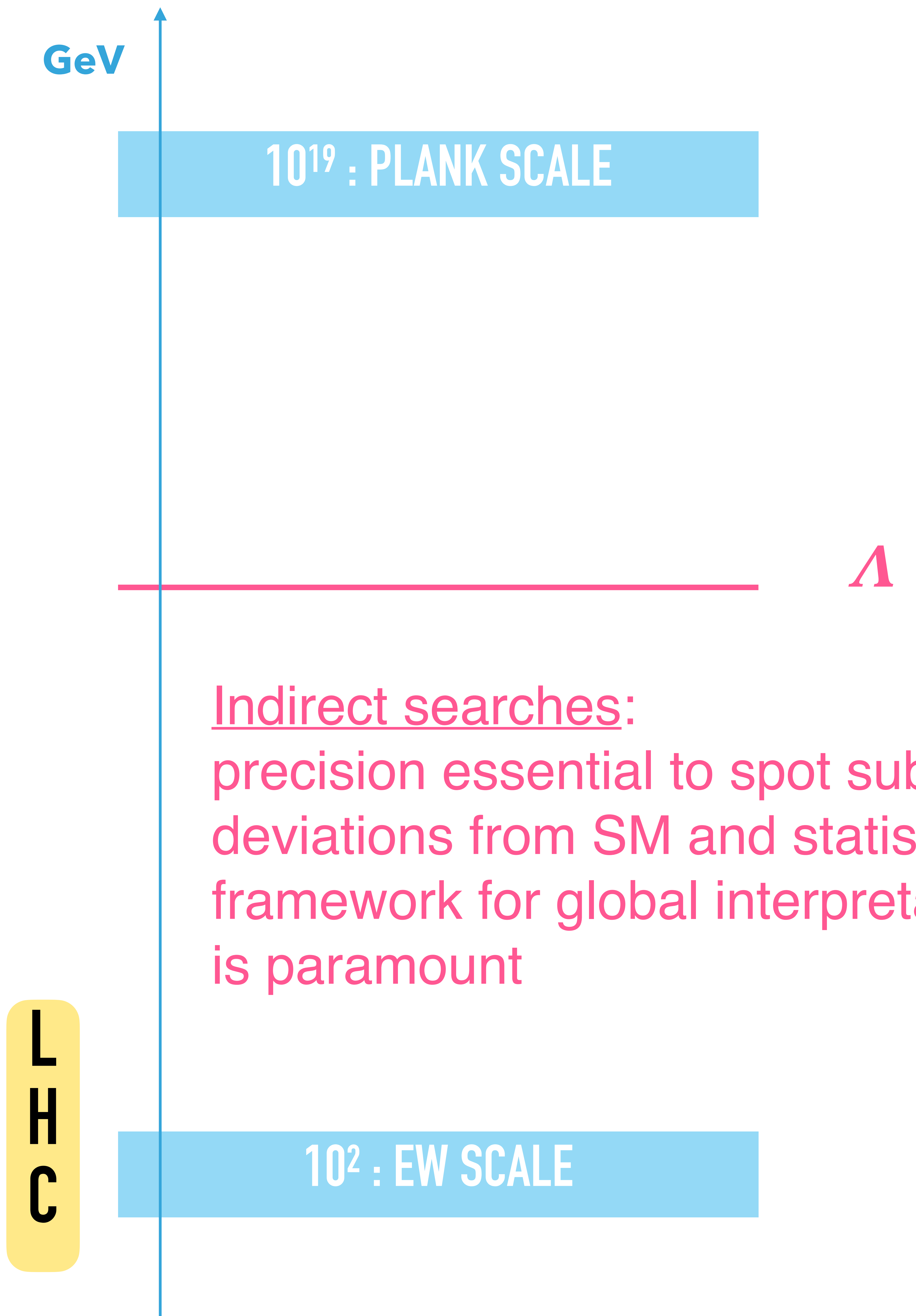


- The presence of the dark Parton would modify the evolution of standard quarks and gluon.
- Interesting that dark photon effect in PDFs is dominant but was not taken into account in recent publication about evidence for dark photons in DIS data [N. T. Hunt-Smith et al arXiv:2302.11126]

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

DISCOVERY THROUGH PRECISION

What if new physics lives at a much higher scale than the one that the LHC can reach?



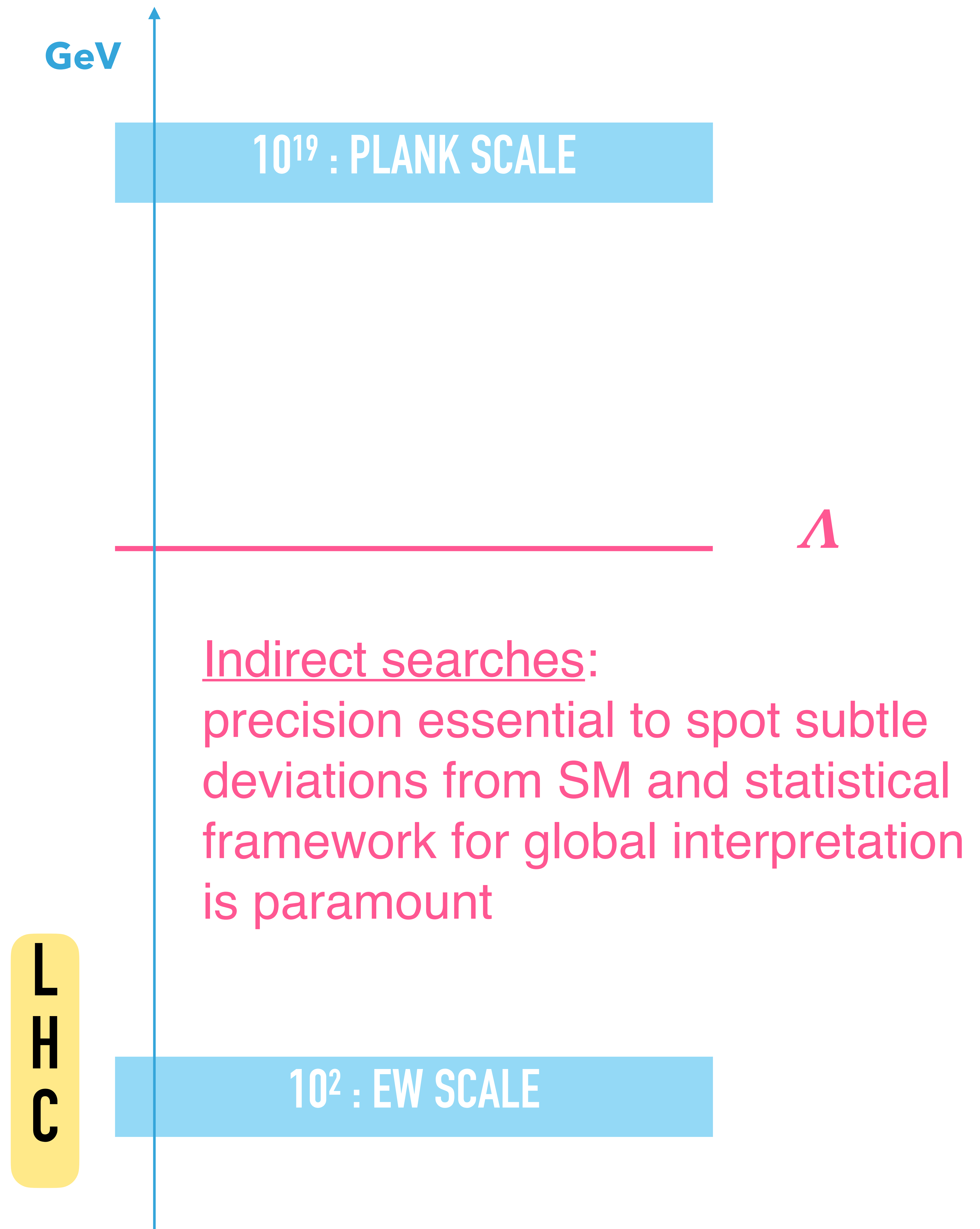
Indirect searches:
precision essential to spot subtle deviations from SM and statistical framework for global interpretation is paramount

DISCOVERY THROUGH PRECISION

- 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 can be incorporated as a momentum expansion
- SMEFT: assume SM field content and gauge symmetries (apart from accidental)

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

- Full dim-6 basis of operators under SMEFT assumptions includes **2499** operators **[Grzadkowski et al, arXiv:1008.4884]**



LOOKING FOR PATTERNS

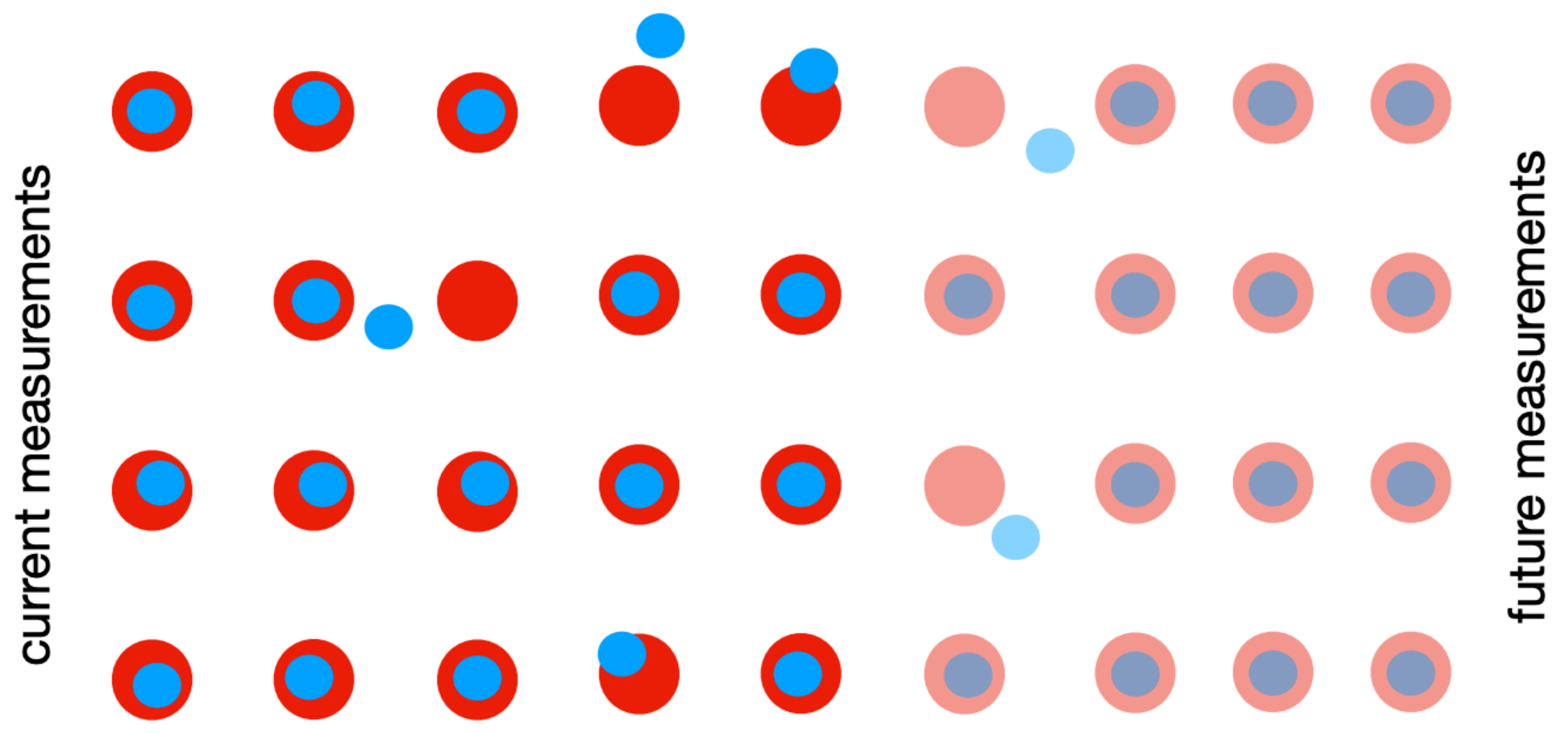
Image credit: F. Maltoni

$$\Delta \mathrm{Obs}_n = \mathrm{Obs}_n^{\mathrm{EXP}} - \mathrm{Obs}_n^{\mathrm{SM}} = \frac{1}{\Lambda^2}\sum_i a_{n,i}^{(6)}(\mu)\, c_i^{(6)}(\mu) + \mathcal{O}\left(\frac{1}{\Lambda^4}\right)$$

Experimental measurements

SM predictions

Coefficients parametrising the deviations due to BSM physics



LOOKING FOR PATTERNS

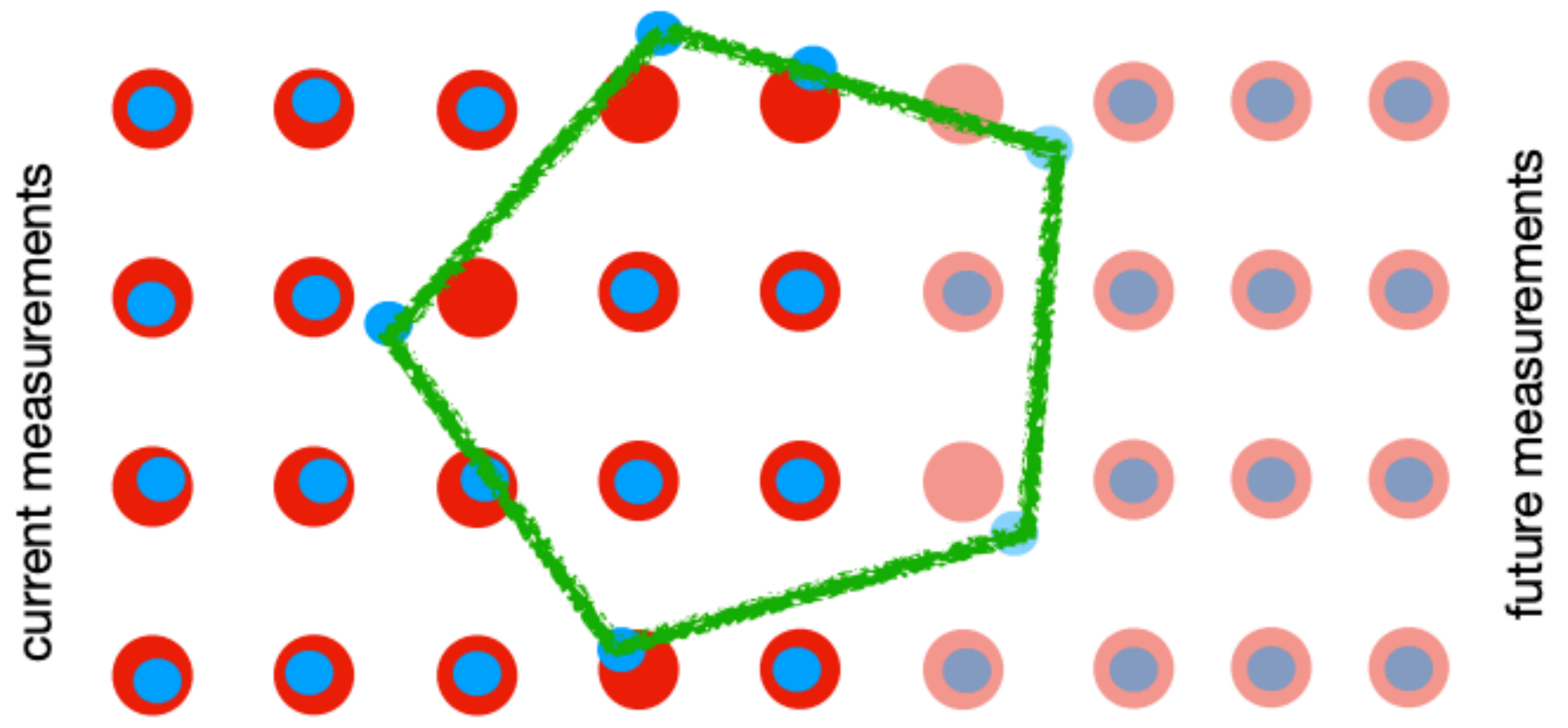
Image credit: F. Maltoni

$$\Delta \mathrm{Obs}_n = \mathrm{Obs}_n^{\mathrm{EXP}} - \mathrm{Obs}_n^{\mathrm{SM}} = \frac{1}{\Lambda^2} \sum_i a_{n,i}^{(6)}(\mu)\, c_i^{(6)}(\mu) + \mathcal{O}\left(\frac{1}{\Lambda^4}\right)$$

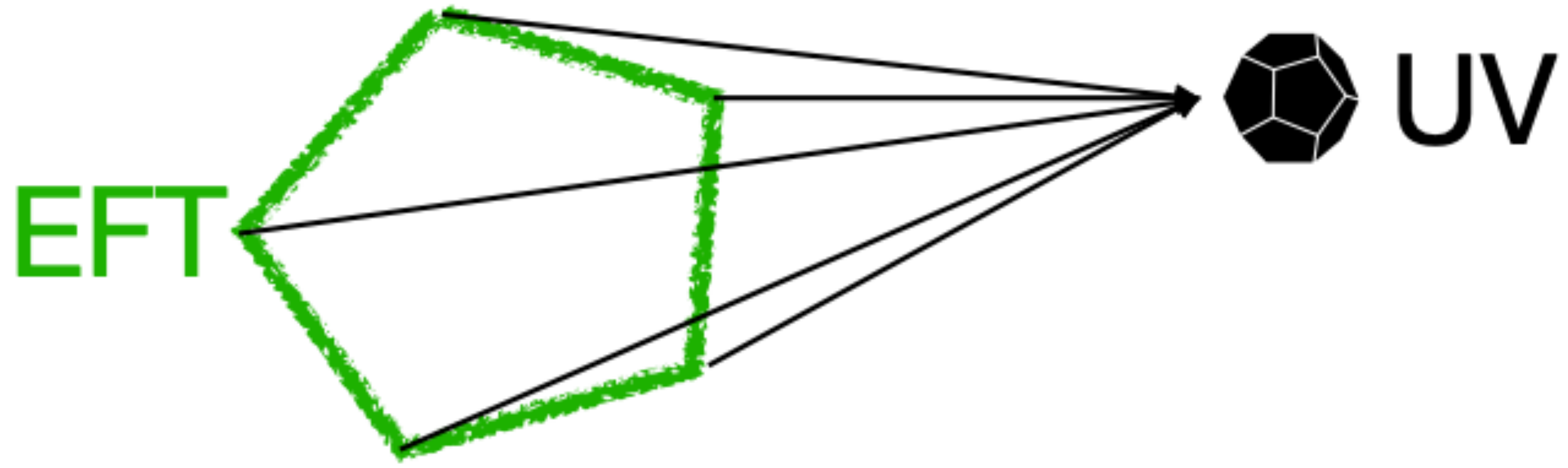
Experimental measurements

SM predictions

Coefficients parametrising the deviations due to BSM physics



As more precise data and SM theoretical predictions become available, we can identify patterns of small differences induced by new physics and from there deduce what is the new model that causes a given pattern!



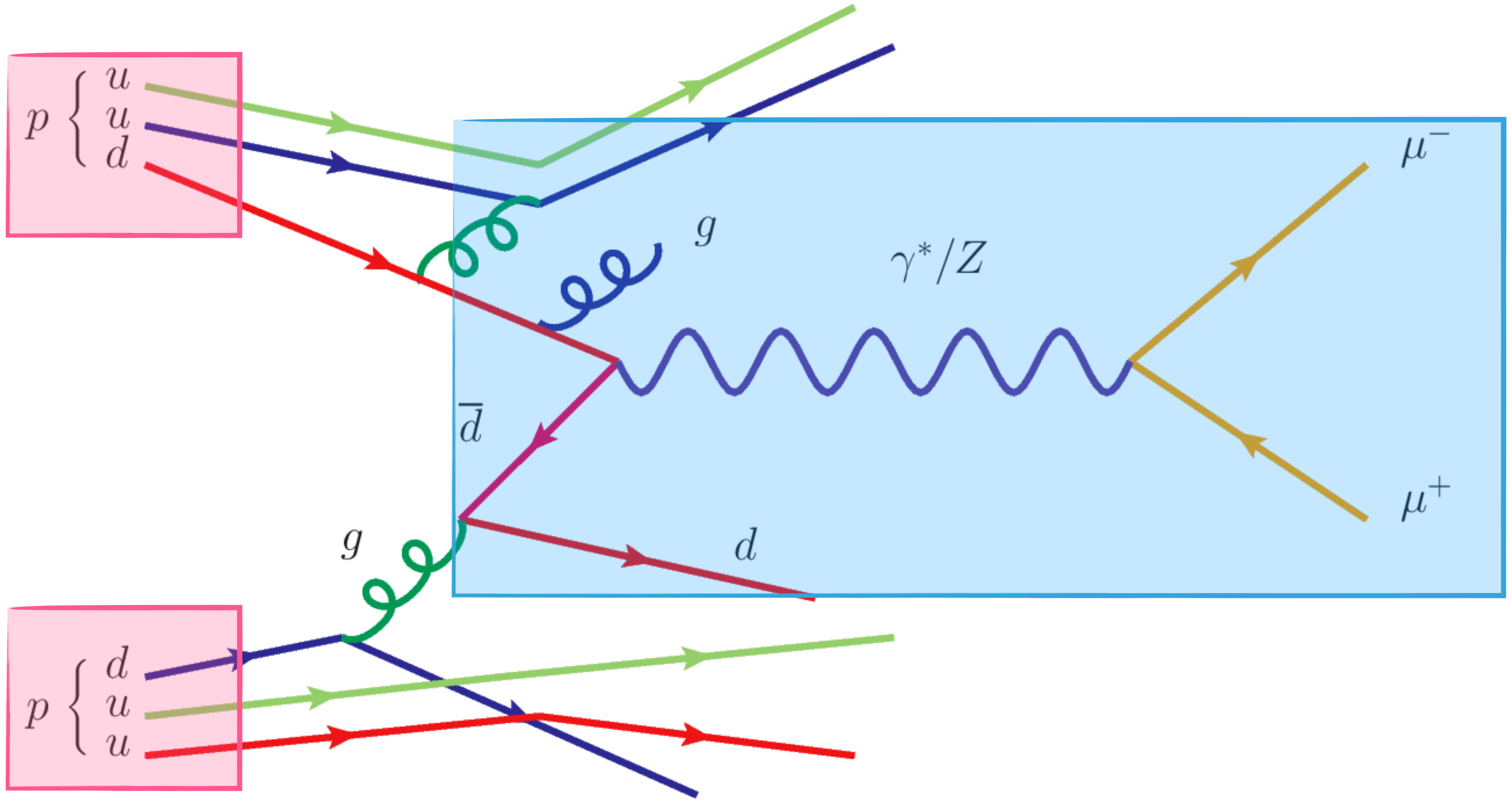
PDF FITS AND SMEFT FITS

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 quarks and gluons (partons) from from short-distance parton interaction (hard scattering)

Parton Distribution Functions



Hard Scattering: Perturbative QCD + EW

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

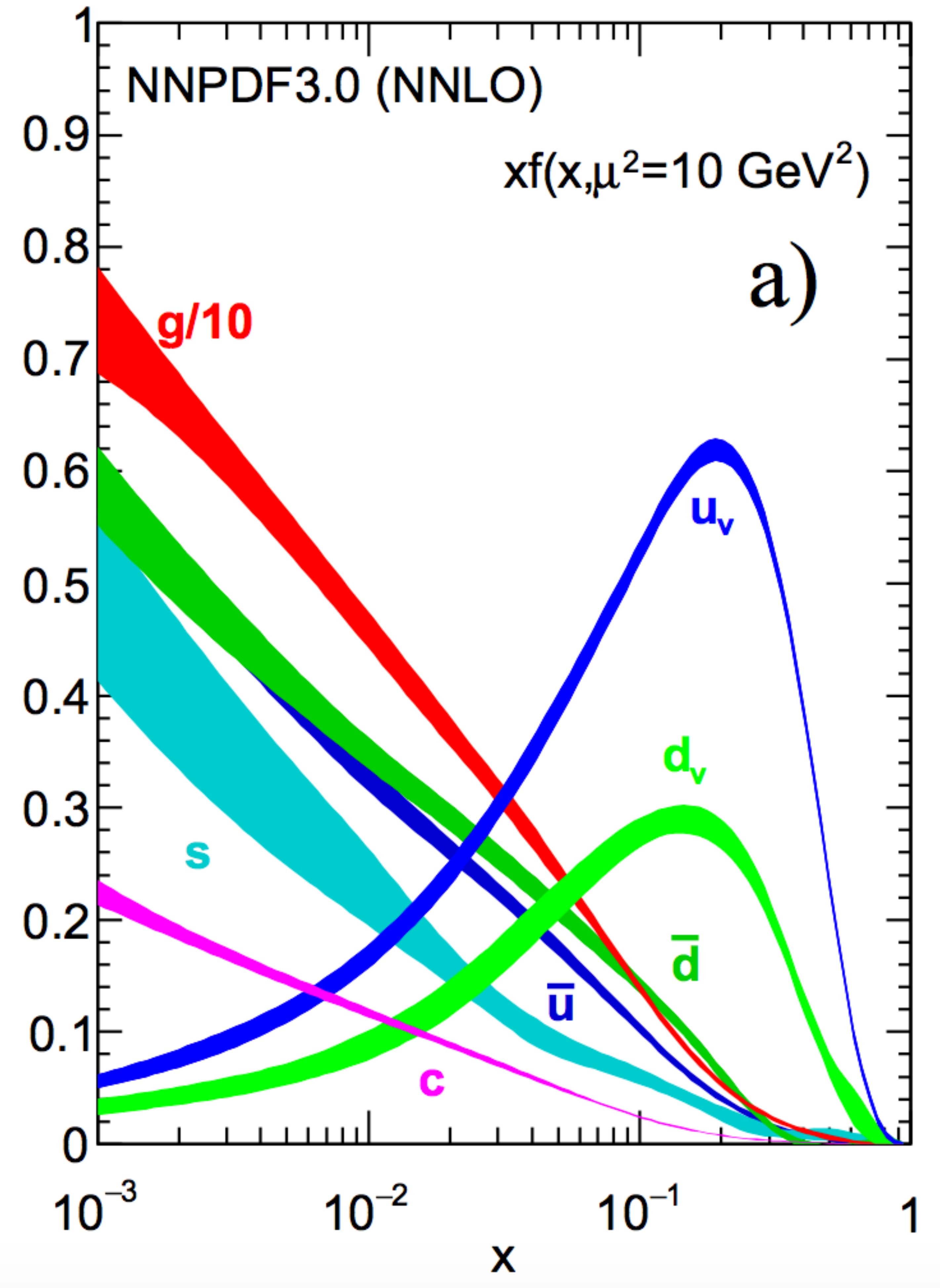
PARTON DISTRIBUTION FUNCTIONS

$$f_i(x, \mu)$$

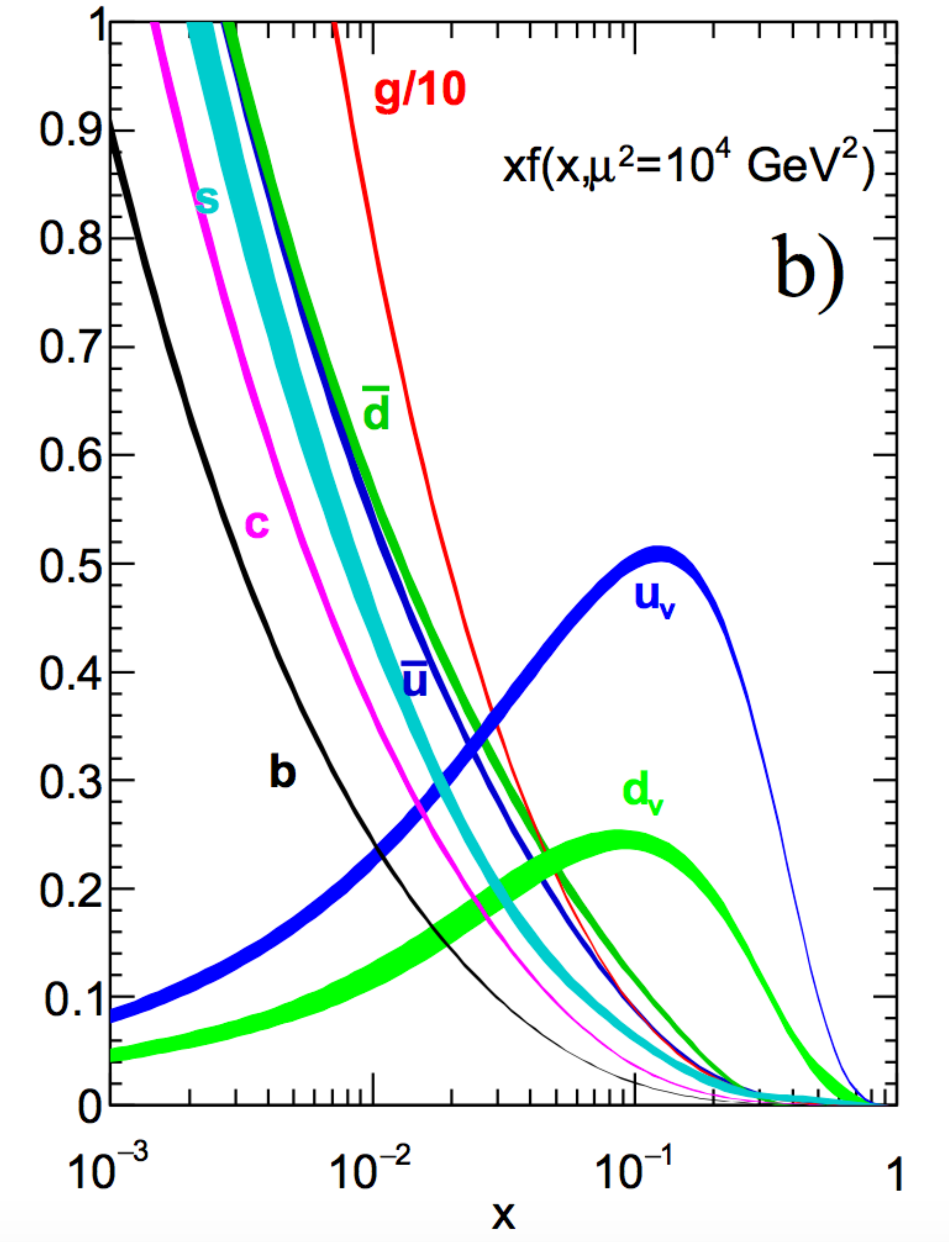
Data ← x

μ → Perturbative QCD

Hadronic scale: global fit of PDFs



pQCD →



High scale: input to the LHC

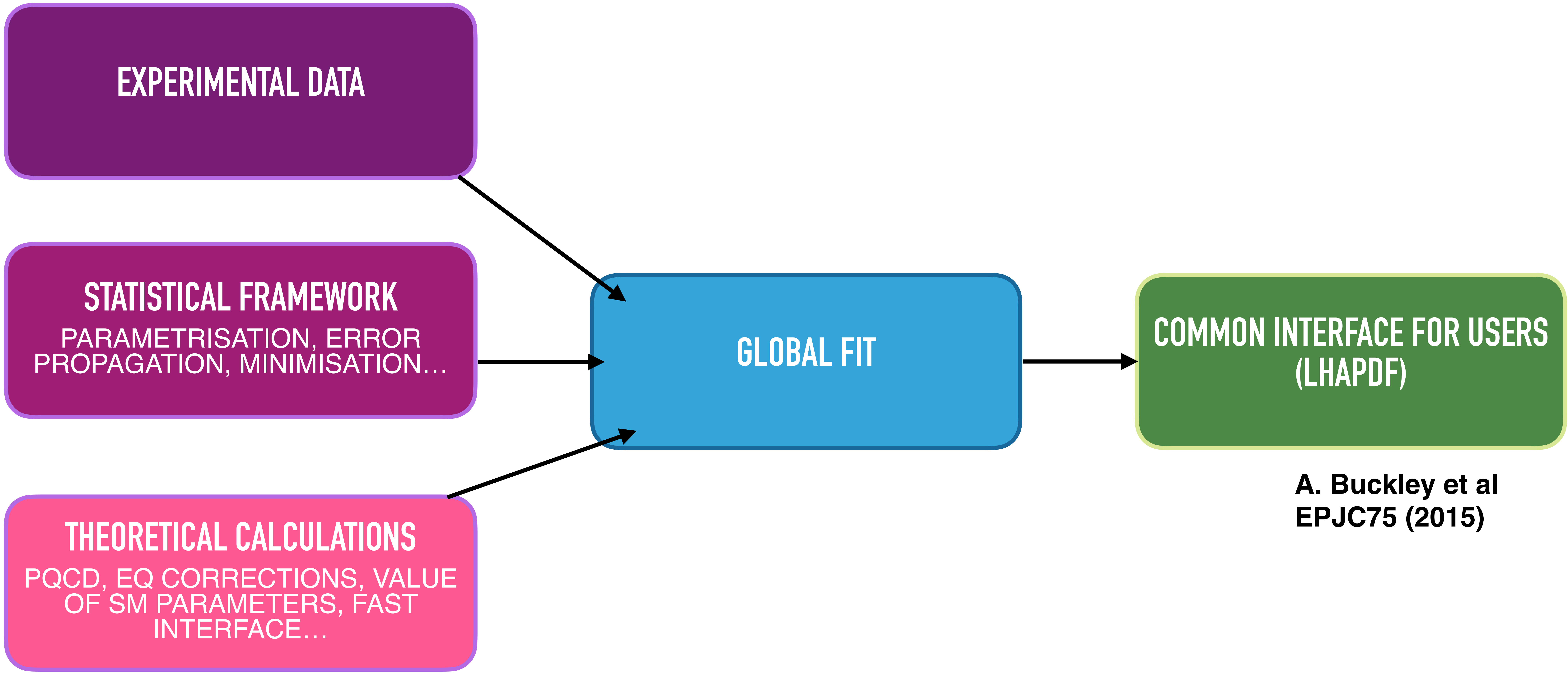
PDG 2016

PARTON DISTRIBUTION FUNCTIONS

$$f_i(x, \mu)$$

Data

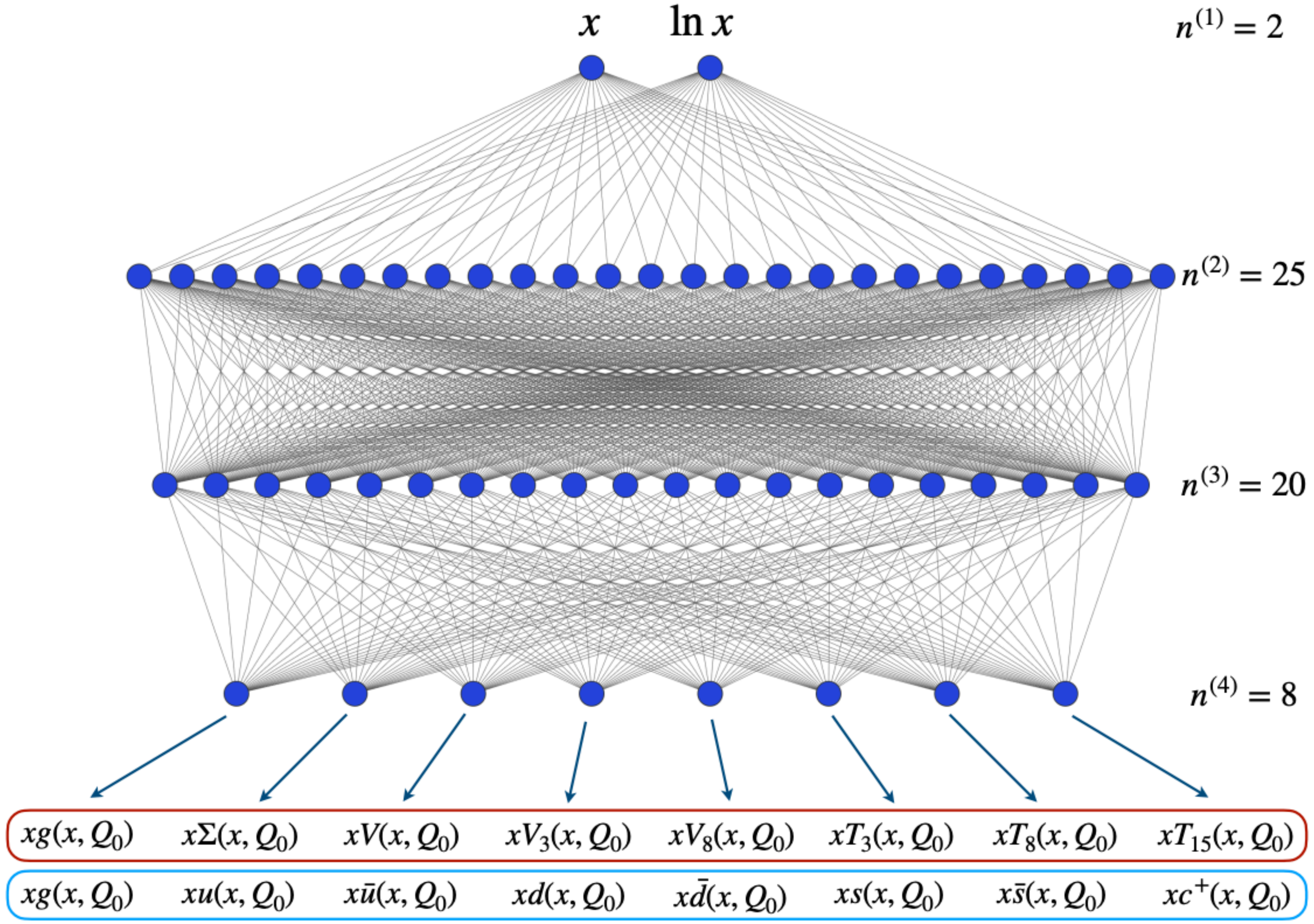
Perturbative QCD

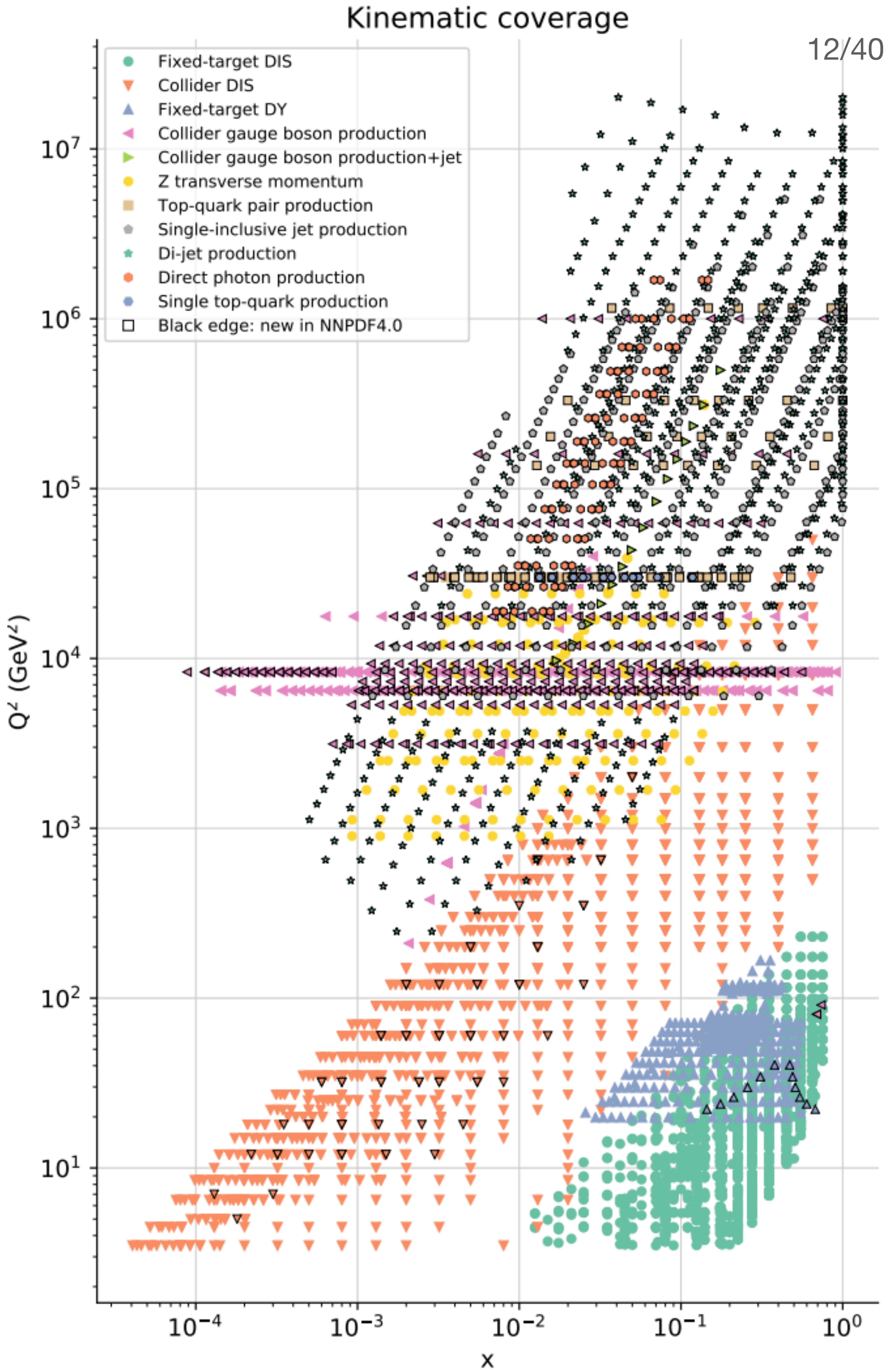


A. Buckley et al EPJC75 (2015)

THE NNPDF4.0

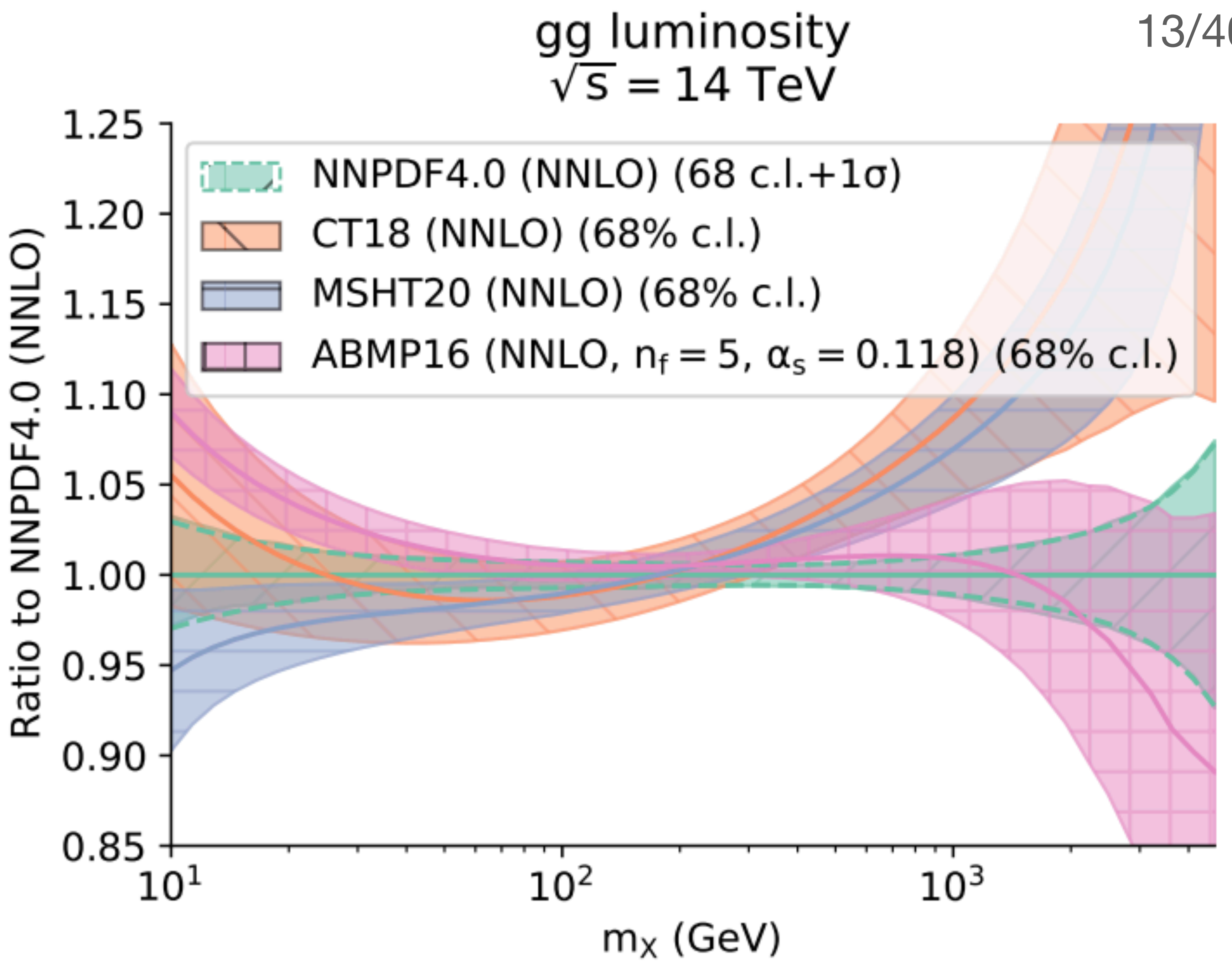
- NNPDF4.0: most recent global PDF set from NNPDF based on very large set of data from LHC: O(5000) data points.
- New methodology based on hyper-parameter optimisation validated by closure tests and future tests. [Ball et al, arXiv:2109.02653]
- Open-source code [Ball et al. arXiv:2109.02671]

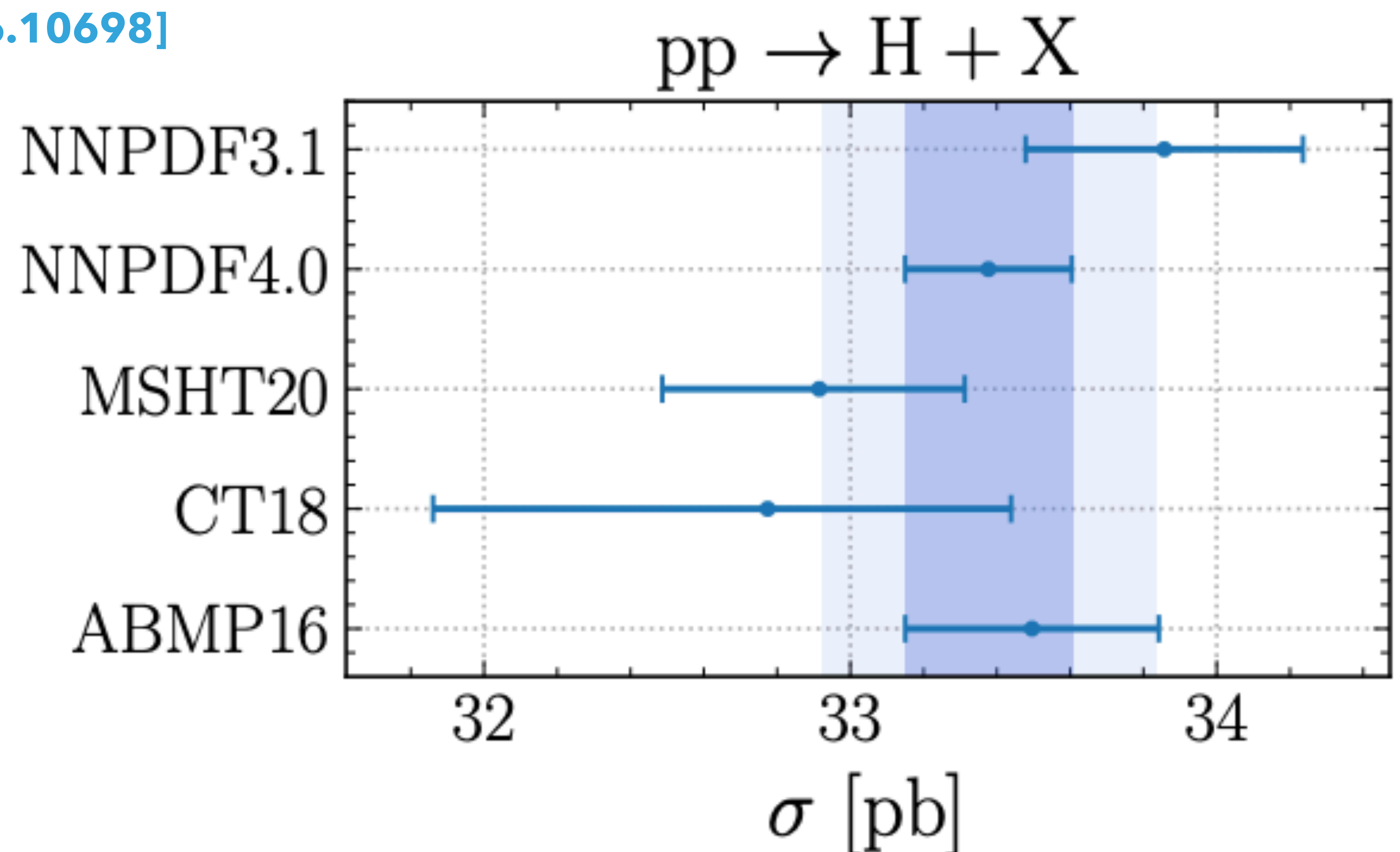


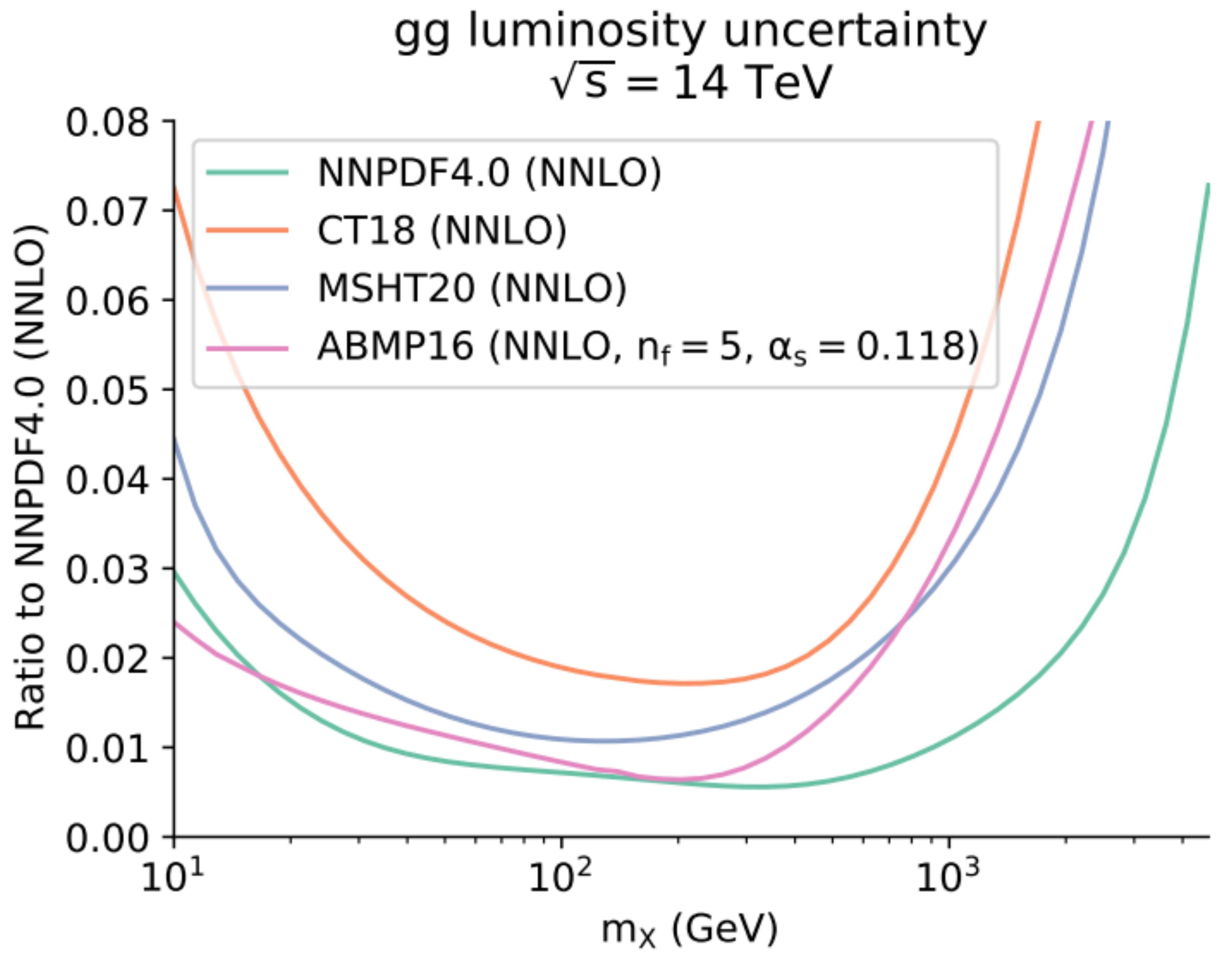


THE NNPDF4.0

- NNPDF4.0: most recent global PDF set from NNPDF based on very large set of data from LHC: O(5000) data points.
- New methodology based on hyper-parameter optimisation validated by closure tests and future tests. [Ball et al, arXiv:2109.02653]
- Open-source code [Ball et al. arXiv:2109.02671]
- Parton luminosity uncertainties down to 1-2% in many regions.
- At such level of precision crucial to account for all theoretical uncertainties, starting from missing higher order uncertainty. [Abdul Khalek et al, arXiv:1906.10698]

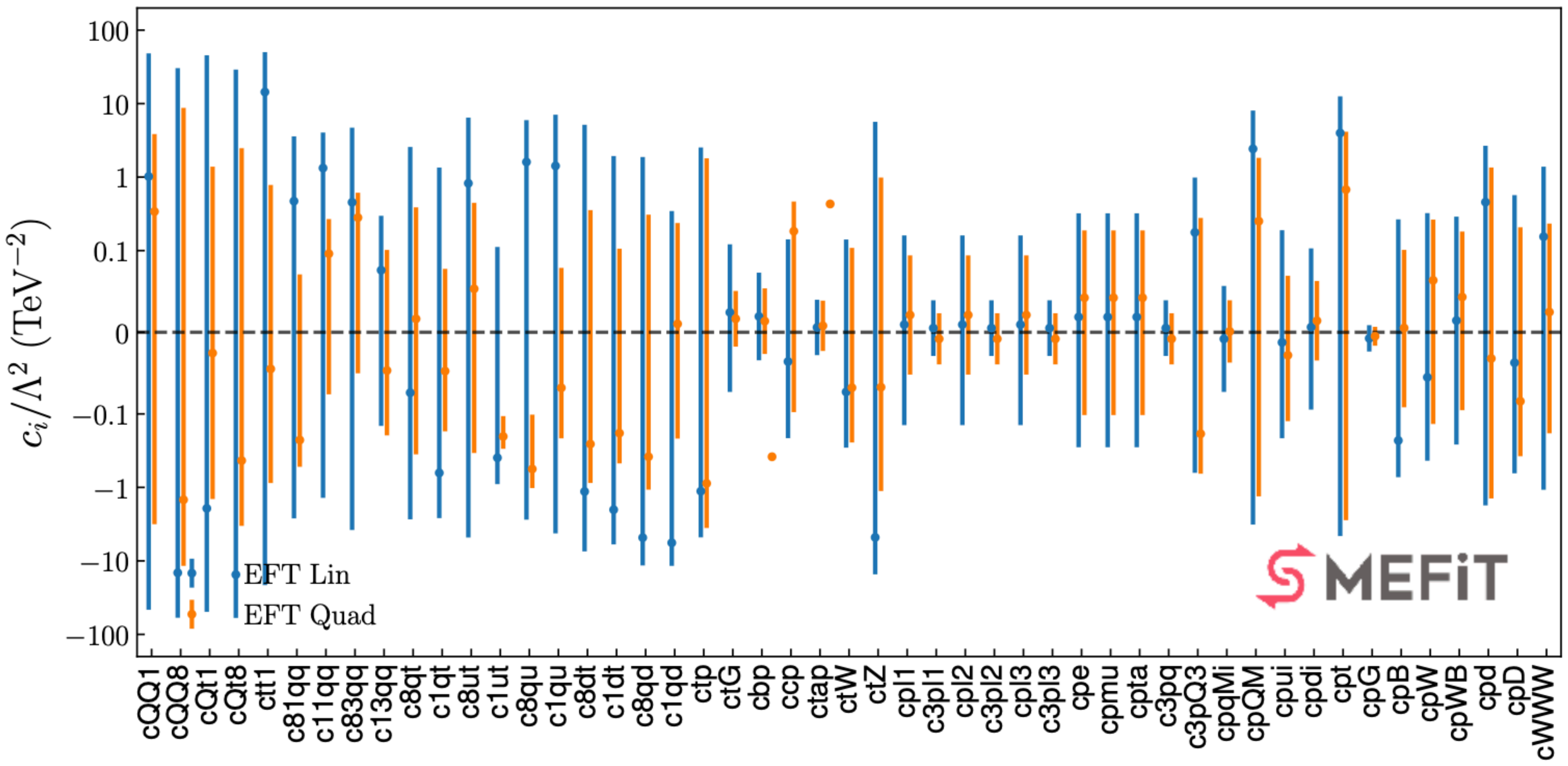






RECENT GLOBAL SMEFT FITS

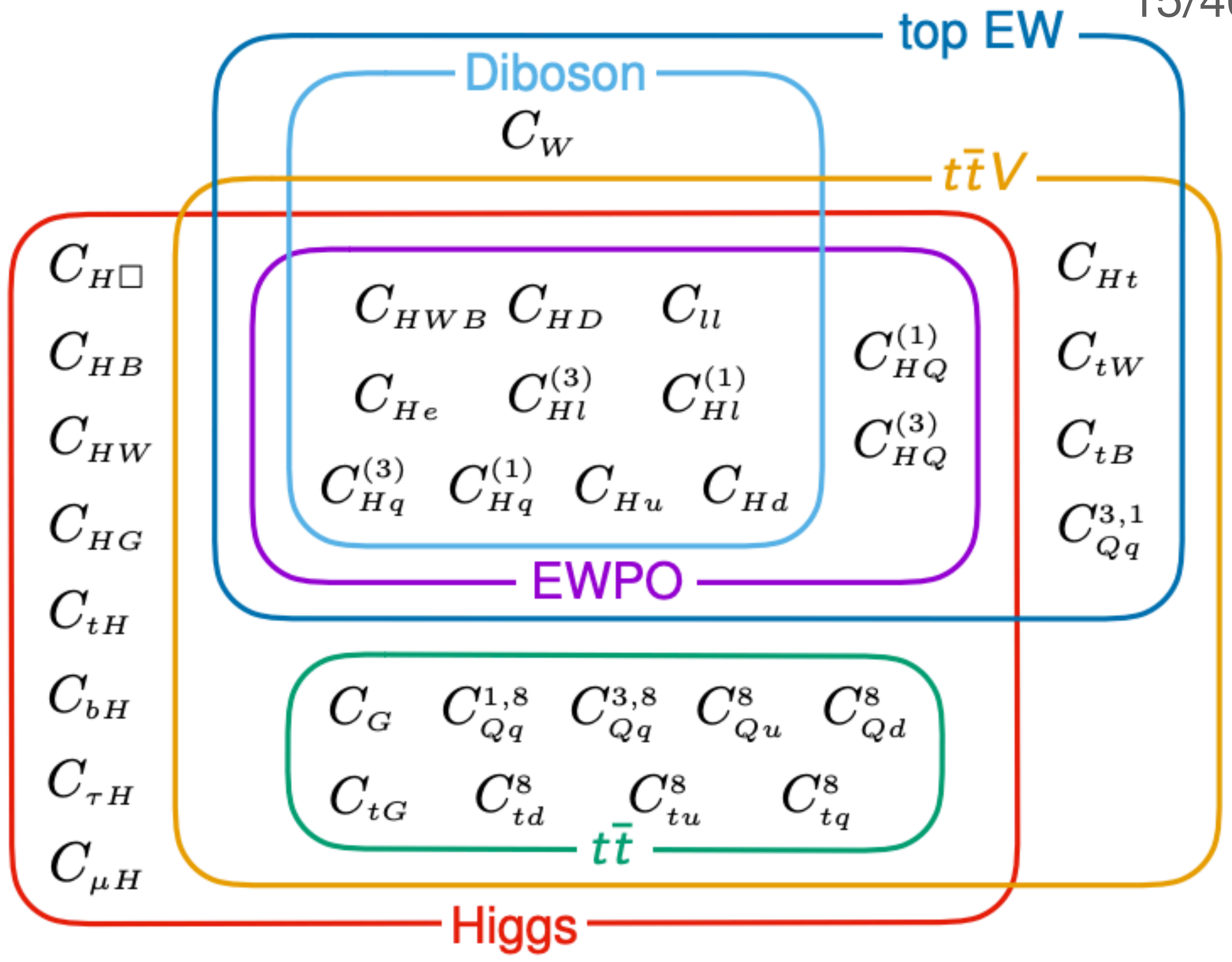
- Current SMEFT fits make flavour assumptions and restricted to a few observables/sectors & reduce the number of operators.
- **SMEFiT**: SMEFT fit based on Monte Carlo technique for propagation of experimental uncertainty [Hartland et al, arXiv:1901.05965]
- Global dim-6 SMEFT fit of Higgs, diboson, and top quark production and decay measurements (36 independent Wilson coefficients, including linear and quadratic contributions and NLO QCD corrections to SMEFT) [Either et al, arXiv:2105.00006]

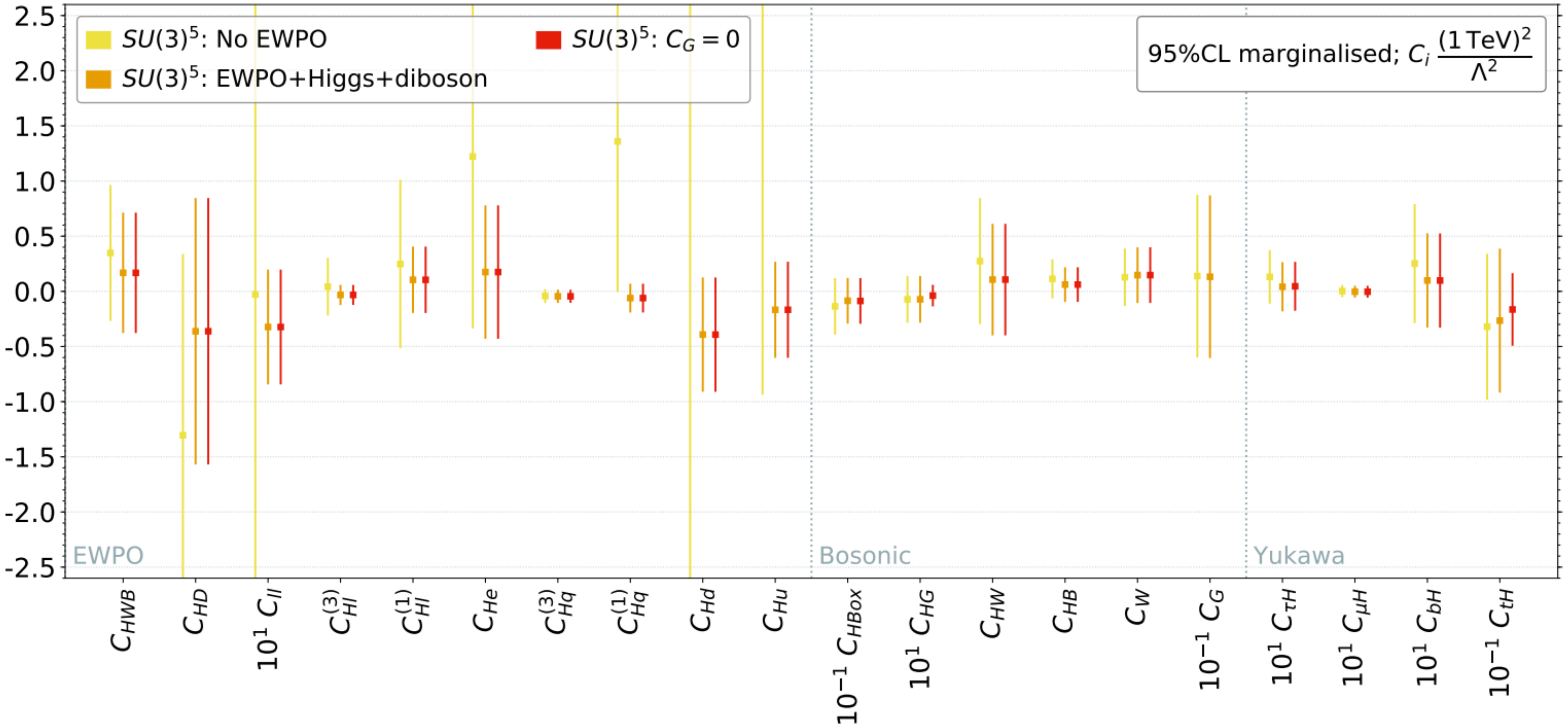


Either et al, arXiv:2105.00006

RECENT GLOBAL SMEFT FITS

- **FitMaker**: Global dim-6 SMEFT fit of Higgs, diboson, electroweak precision and top quark production and decay measurements (linear contributions and LO QCD corrections to SMEFT) [Ellis et al, arXiv:2012.02779]





Ellis et al, arXiv:2012.02779

EXTRACTING PARAMETERS FROM DATA

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

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

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

(B)SM parameters: $\alpha_s(M_z)$, M_w, θ_w ,**SMEFT WCs**……

Parameters determining PDFs at initial scale

EXTRACTING PARAMETERS FROM DATA

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

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

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

(B)SM parameters: $\alpha_s(M_z)$, M_w, θ_w ,**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

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

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

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

(B)SM parameters: $\alpha_s(M_z)$, M_w, θ_w ,**SMEFT WCs**……

Parameters determining PDFs at initial scale

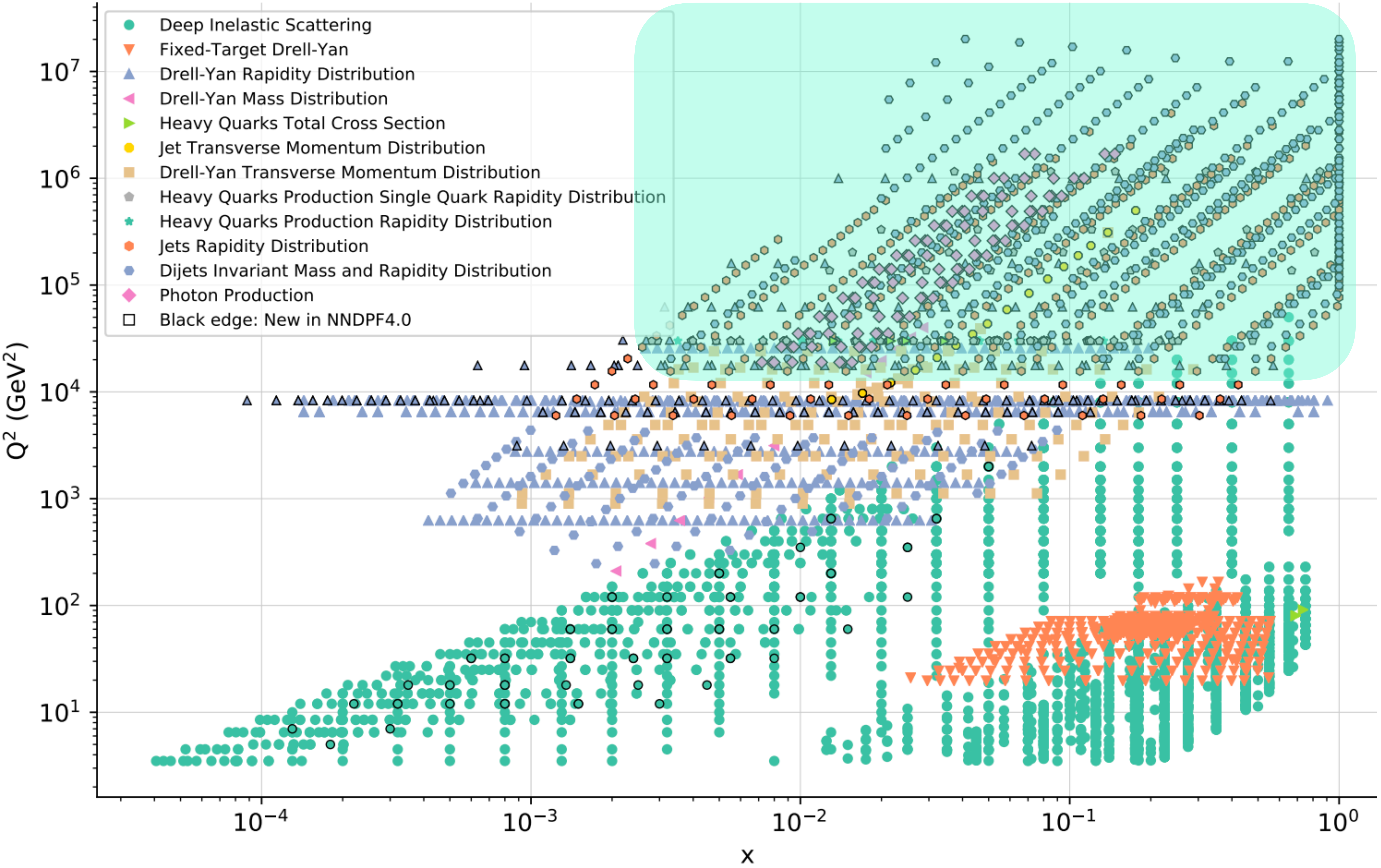
✓ In a PDF fit typically

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

✓ In a fit of SMEFT Wilson Coefficients

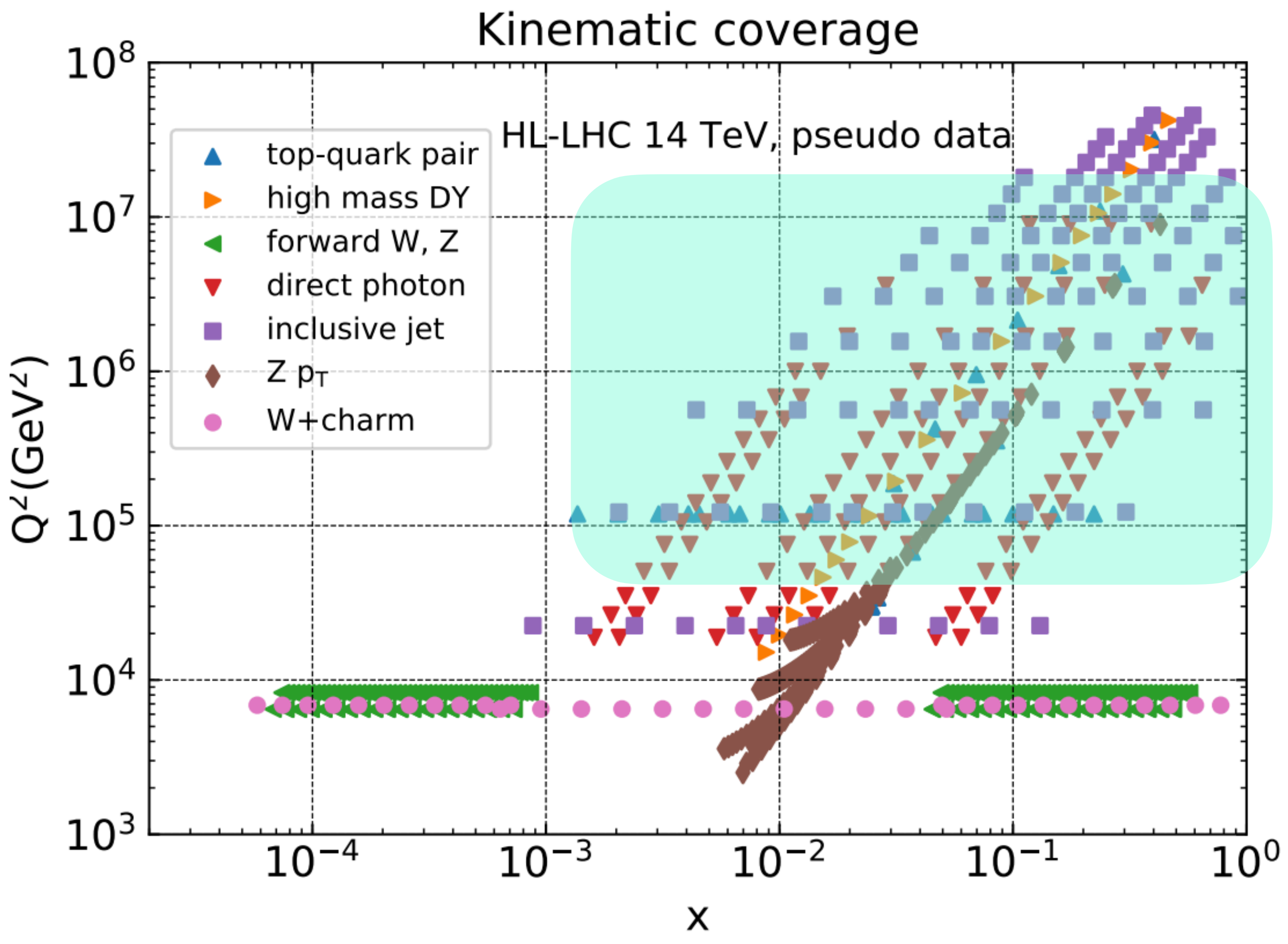
$$T_i(\{c\}) = {\rm PDFs}(\{\theta=\bar{\theta}\},\{c=0\}) \otimes \hat{\sigma}_i(\{c\})$$

PDF AND SMEFT INTERPLAY



Ball et al, arXiv:2109.02653

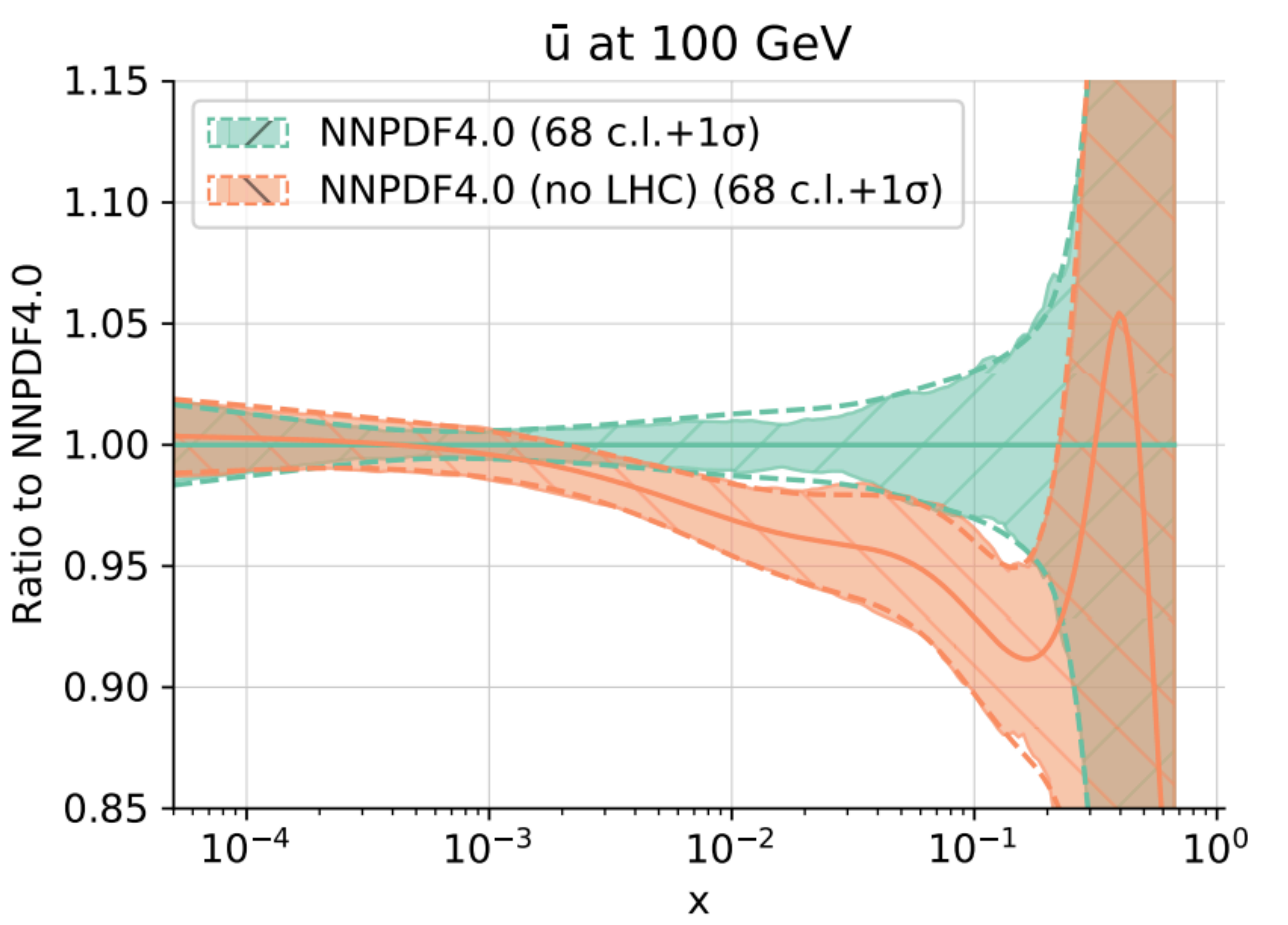
- ➡ Top pair production and single top data included in SMEFT analysis [Hartland et al 1901.05965] [Ellis et al 2012.02779]
- ➡ Dijets data in [Bordone et al 2103.10332] [Alioli et al 1706.03068]
- ➡ Drell-Yan data in [Farina et al 1609.08157, Torre et al 2008.12978]
- ➡ Jets and dijets [Alte et al 1711.07484]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



Abdul Khalek et al,1810.03639

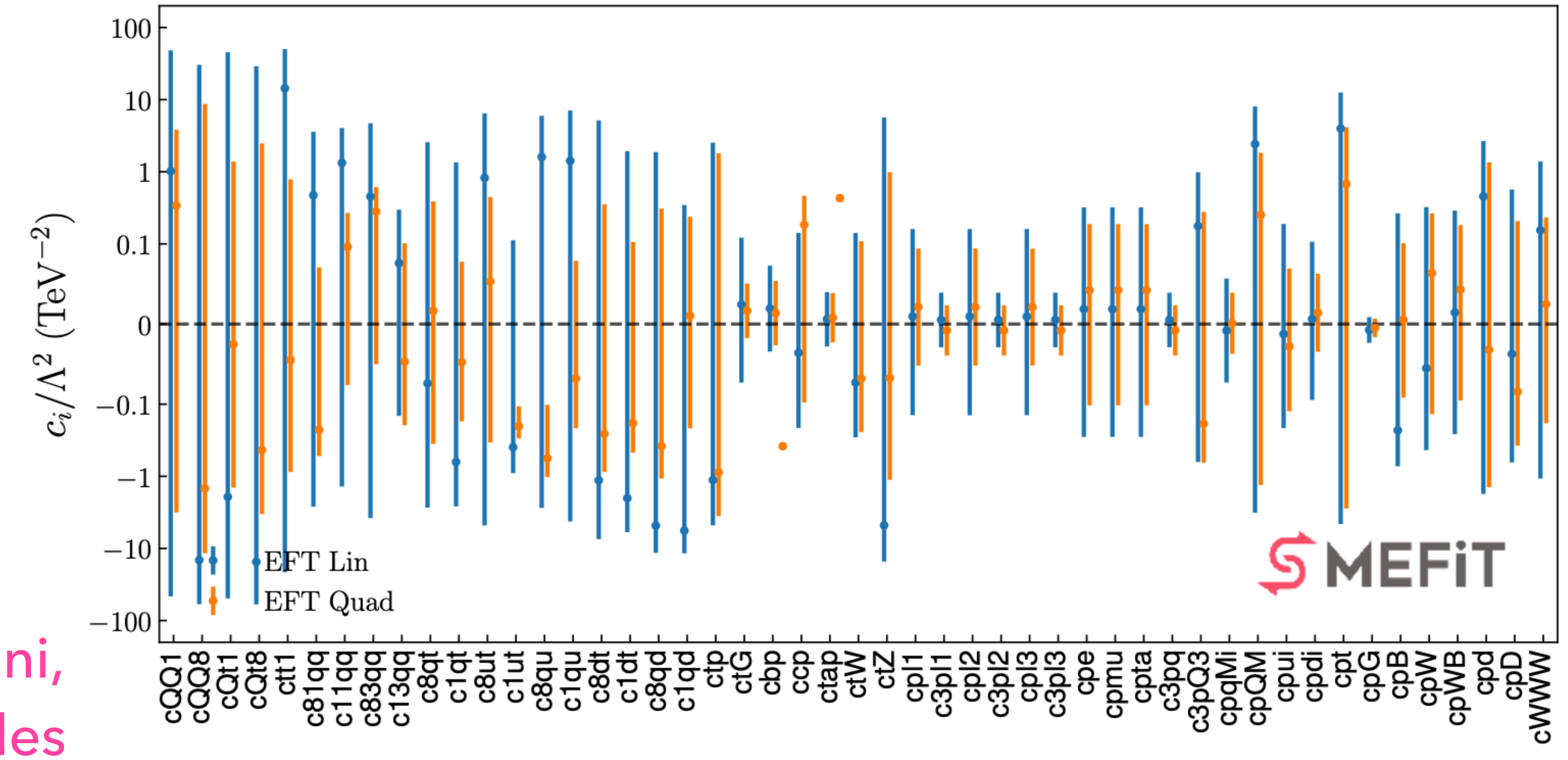
PDF AND SMEFT INTERPLAY

- 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 (PI) M. Madigan, L. Mantani, J. Moore (postdocs), M. Morales Alvarado, E. Hammou, M. Costantini (PhD students)



Ethier et al, arXiv: 2105.00006

A FEW COMPELLING QUESTIONS

- From the point of view of PDF fits:
  - ➡ How to make sure that new physics effects are not inadvertently fitted away in a PDF fit?
- From the point of view of SMEFT fits:
  - ➡ Should I make sure I am using a clean set of PDFs in a SMEFT analysis? How to define it? Is it enough?
  - ➡ How would the bounds change if I was consistently using PDFs that include in the fit the same operators that I am fitting?

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

$$f(\{\theta_k\}) \qquad \mathcal{L}_{\rm SMEFT} = \mathcal{L}_{\rm SM} + \sum_i \frac{c_i}{\Lambda^2}\mathcal{O}_i^{(6)} + \cdots$$

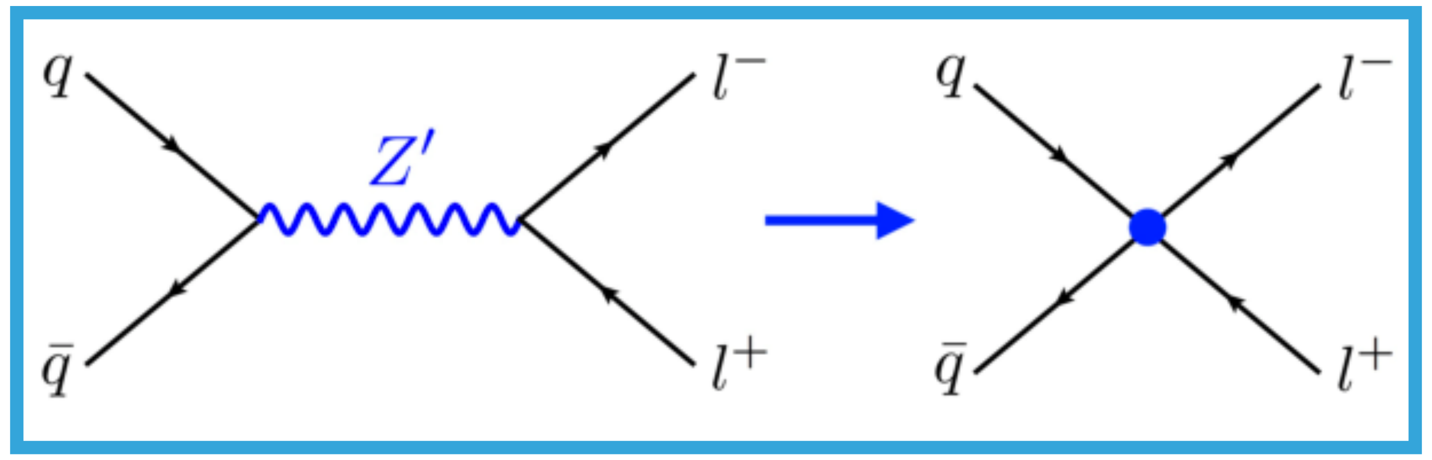
Simultaneous fits can shed light on their interplay

$$T(\{\theta_k\},\{c_i\})$$

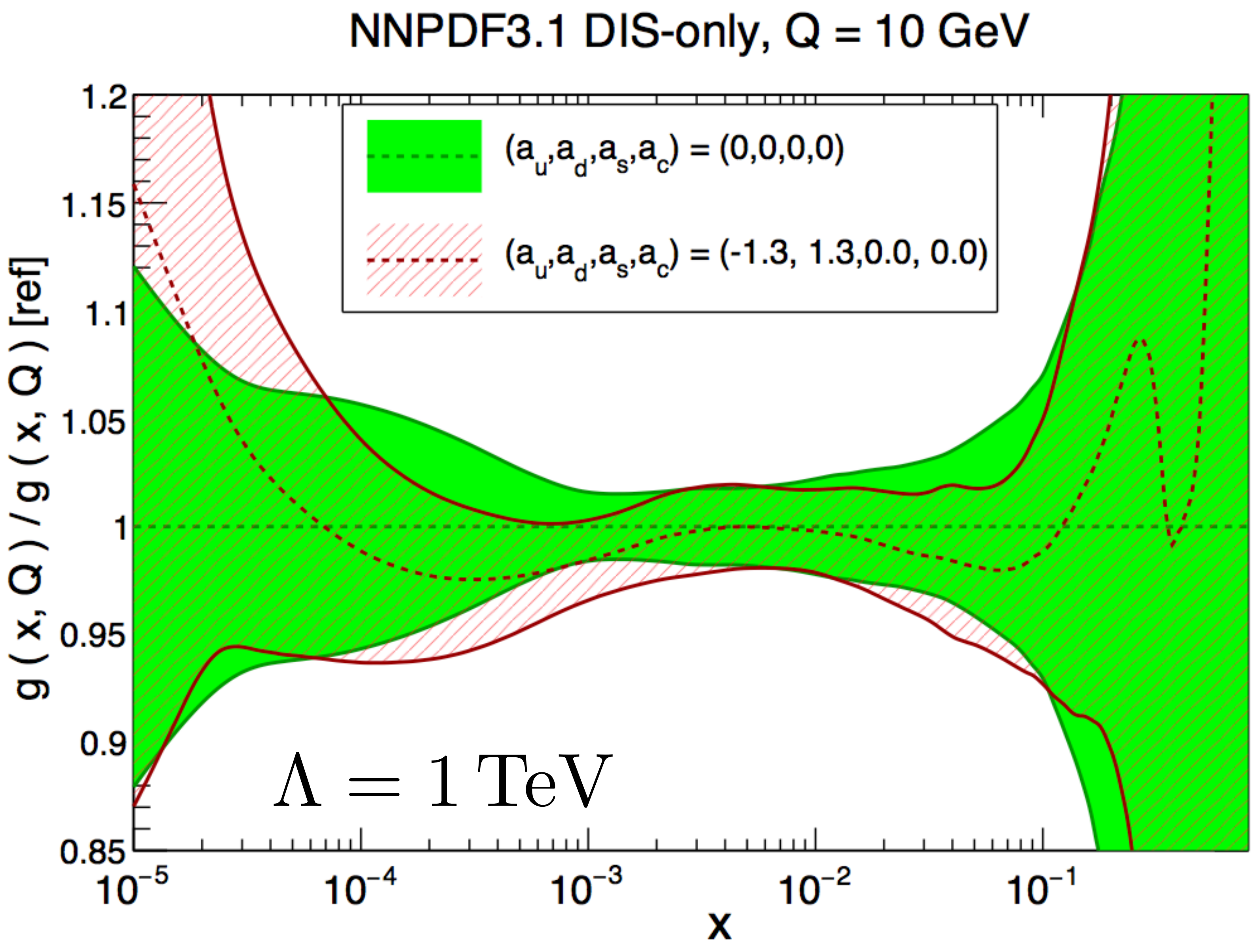
SIMULTANEOUS PDF AND SMEFT FITS

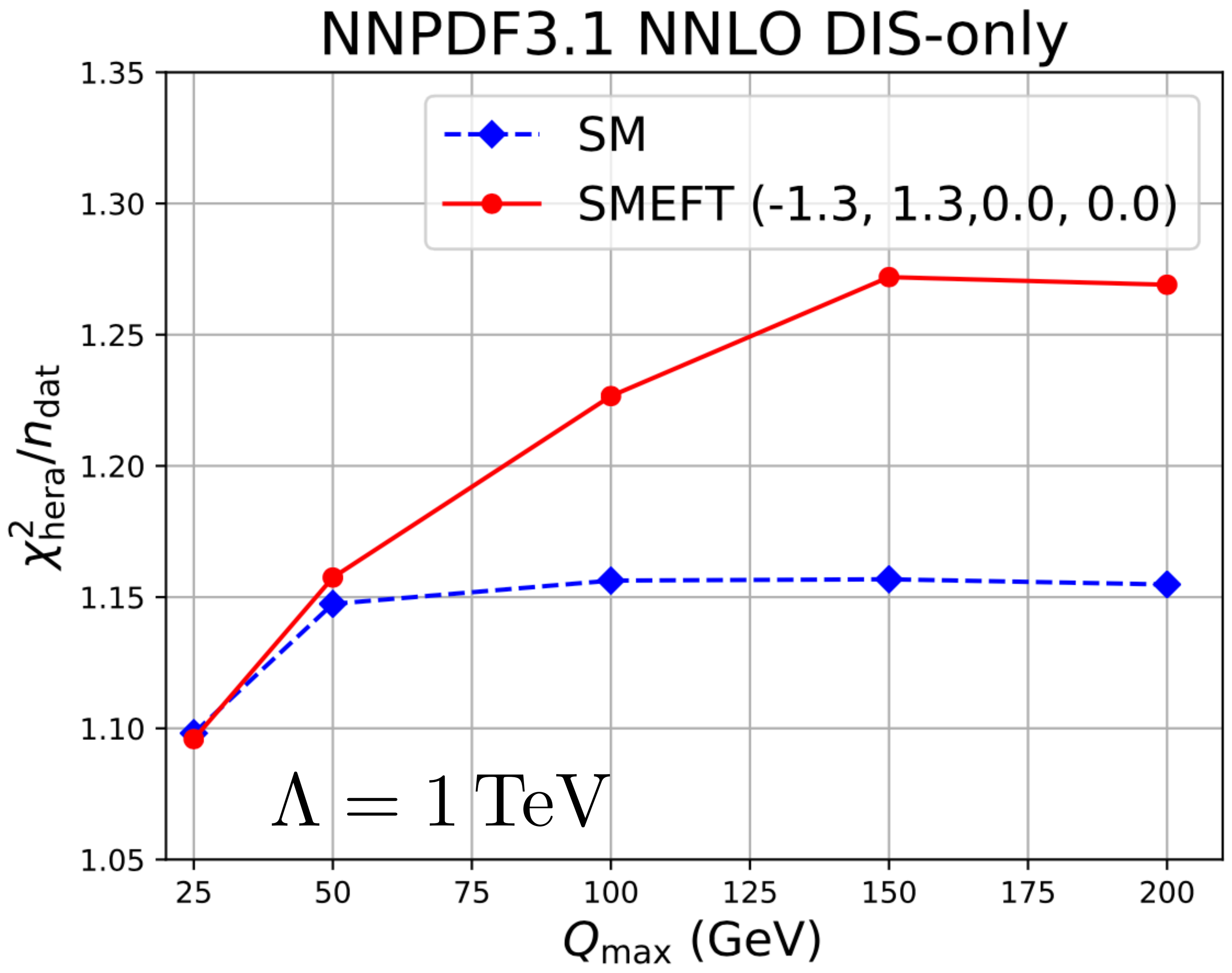
A SCAN-BASED APPROACH: A DIS CASE STUDY

- First study of interplay in case of DIS data
  [Carrazza, Degrande, Iranipour, Rojo, MU, Phys.Rev.Lett. 123 (2019) 13, 132001]
- Simple scenario, only right-handed 4F operators, lepton flavour blind, quark flavours split to evade strong LEP constraints
- N PDF fits in N points of 4D operator space, fits based on DIS only data (Q ≲ 200 GeV for HERA data)



$$\mathcal{O}_{lu} = (\bar{l}_R \gamma^\mu l_R)(\bar{u}_R \gamma_\mu u_R)$$
$$\mathcal{O}_{lc} = (\bar{l}_R \gamma^\mu l_R)(\bar{c}_R \gamma_\mu c_R)$$
$$\mathcal{O}_{ld} = (\bar{l}_R \gamma^\mu l_R)(\bar{d}_R \gamma_\mu d_R)$$
$$\mathcal{O}_{ls} = (\bar{l}_R \gamma^\mu l_R)(\bar{s}_R \gamma_\mu s_R)$$





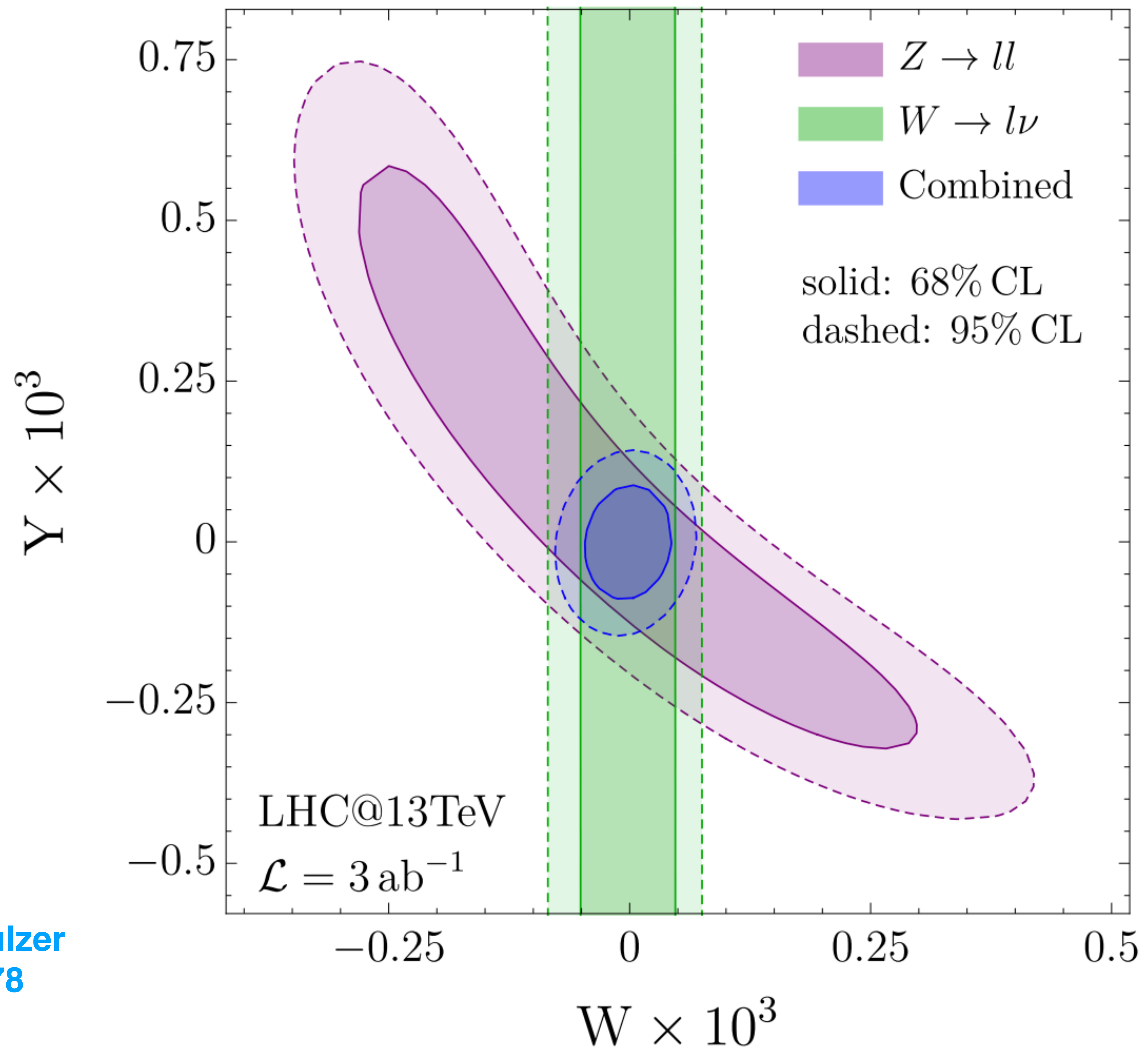
Only gluon affected by the presence of non-zero coefficients, but distortion of PDFs leads to a deterioration of data-theory agreement that scales with energy
=> A fit based on DIS data is only moderately affected by interplay and the effects of new physics can be disentangled

A SCAN-BASED APPROACH: A DY CASE STUDY

- Focus on effect of oblique operators (Y and W) on high-energy Drell-Yan invariant-mass tails
  [Greljo, Iranipour, Kassabov, Madigan, Moore, Rojo, MU, Voisey: 2104.02723]

	universal form factor ($\mathcal{L}$)	contact operator ($\mathcal{L}'$)
W	$-\frac{W}{4m_W^2}(D_\rho W^a_{\mu\nu})^2$	$-\frac{g_2^2 W}{2m_W^2} J_L{}^a_\mu J_L{}^\mu_a$
Y	$-\frac{Y}{4m_W^2}(\partial_\rho B_{\mu\nu})^2$	$-\frac{g_1^2 Y}{2m_W^2} J_Y{}_\mu J_Y{}^\mu$

- Oblique parameters encapsulate effect of universal new physics
- Linear combinations of four-fermion operators
- W and Y effect grows with energy

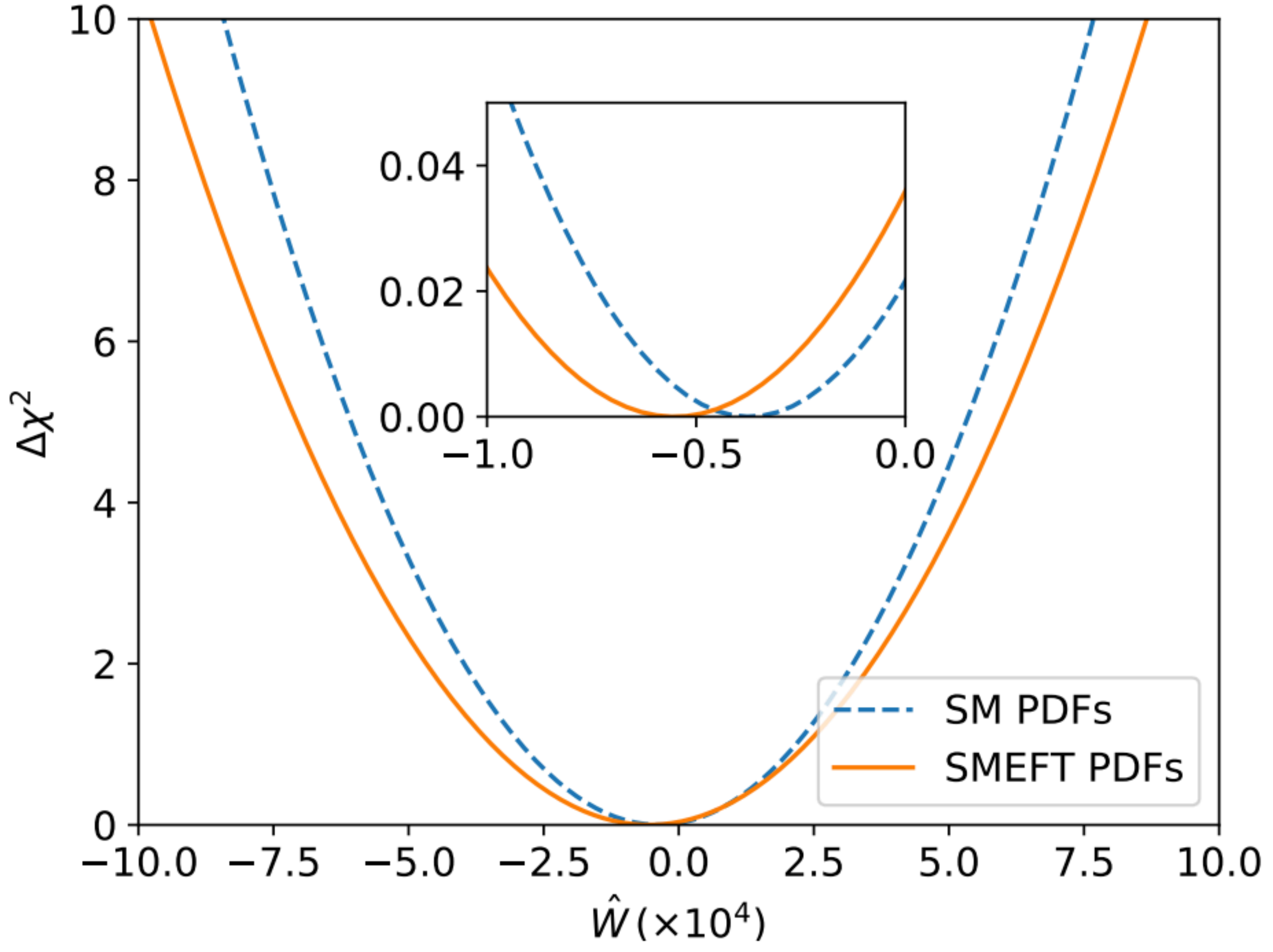


Ricci, Torre, Wulzer
arXiv:2008.12978

A SCAN-BASED APPROACH: A DY CASE STUDY

- Focus on effect of oblique operators (Y and W) on high-energy Drell-Yan invariant-mass tails
  [Greljo, Iranipour, Kassabov, Madigan, Moore, Rojo, MU, Voisey: 2104.02723]

- Scan on BPs in the (Y,W) space
- Run I & Run II high-mass neutral current DY data: little effect



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

SM PDFs

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

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

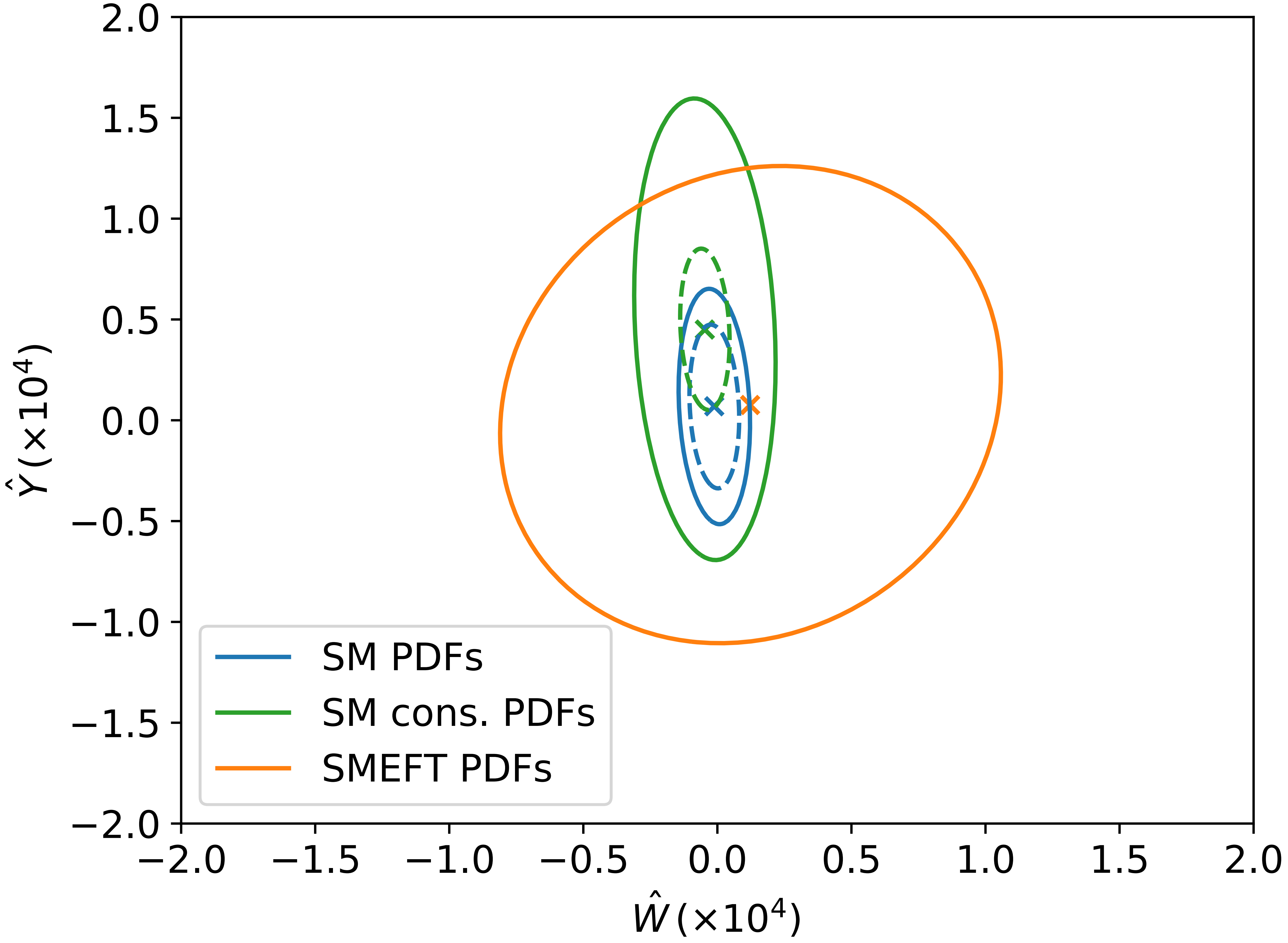
SMEFT PDFs / Simultaneous fit

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

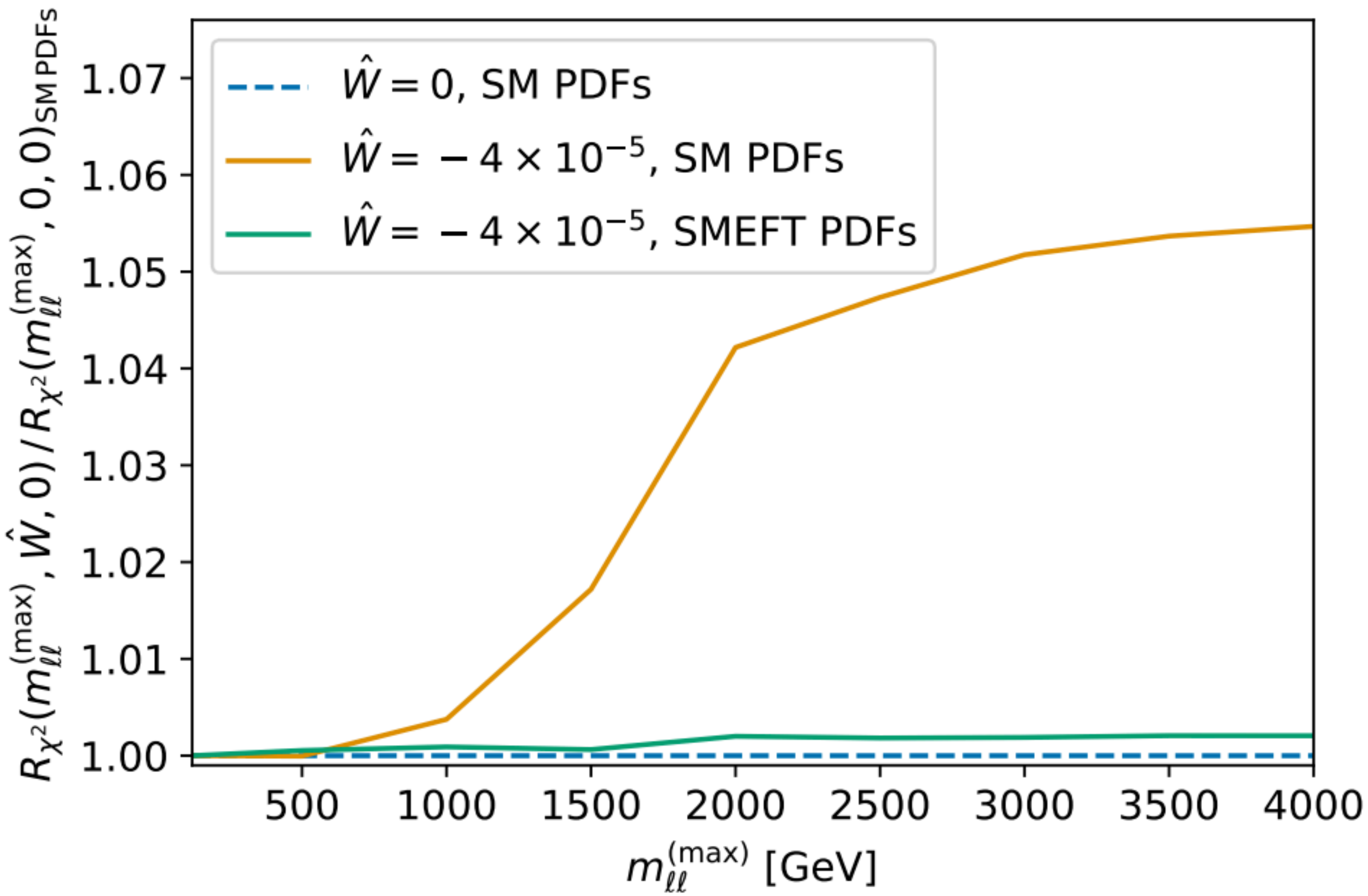
$$T = f_{1,\rm BSM} \otimes f_{2,\rm BSM} \otimes \hat{\sigma}_{\rm BSM}$$

A SCAN-BASED APPROACH: A DY CASE STUDY

- Compare Wilson coefficients bounds from HL-LHC projections including neutral and charged current Drell-Yan 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

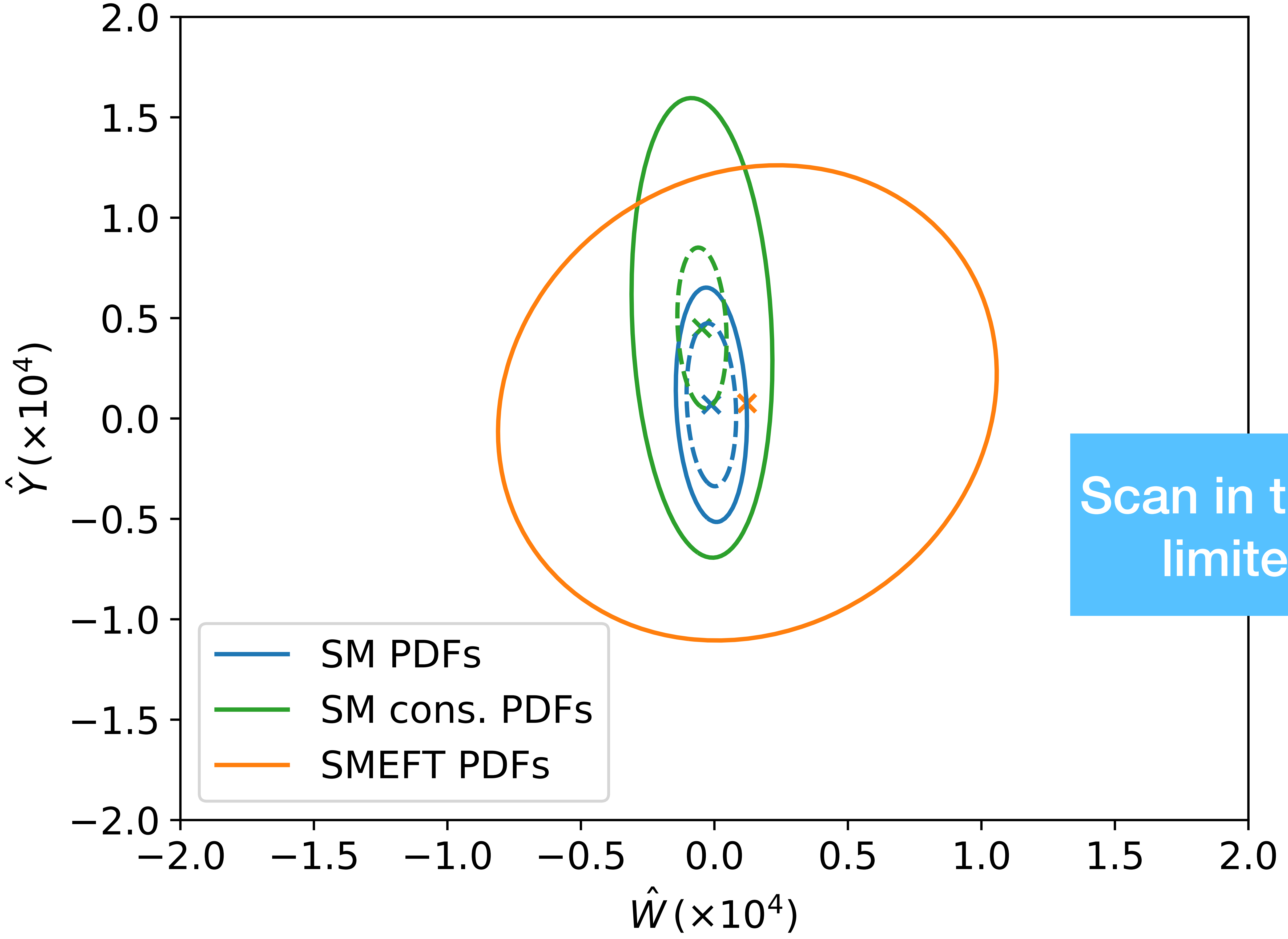


Greljo, Iranipour, Kassabov, Madigan, Moore, Rojo, MU, Voisey: 2104.02723

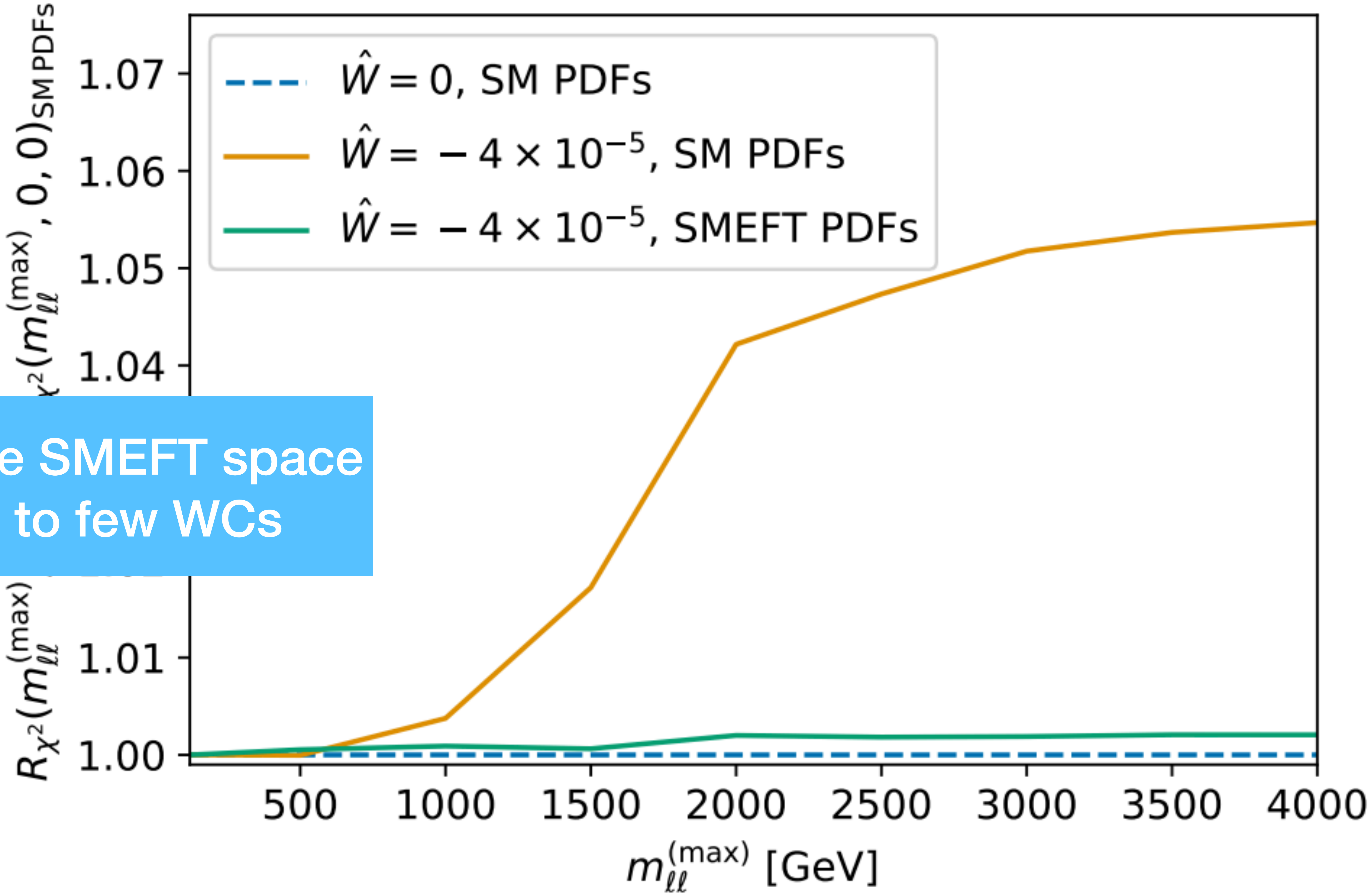


A SCAN-BASED APPROACH: A DY CASE STUDY

- Compare Wilson coefficients bounds from HL-LHC projections including neutral and charged current Drell-Yan 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



Greljo, Iranipour, Kassabov, Madigan, Moore, Rojo, MU, Voisey: 2104.02723



Scan in the SMEFT space limited to few WCs

SIMUNET: A DEEP-LEARNING BASED SIMULTANEOUS FIT

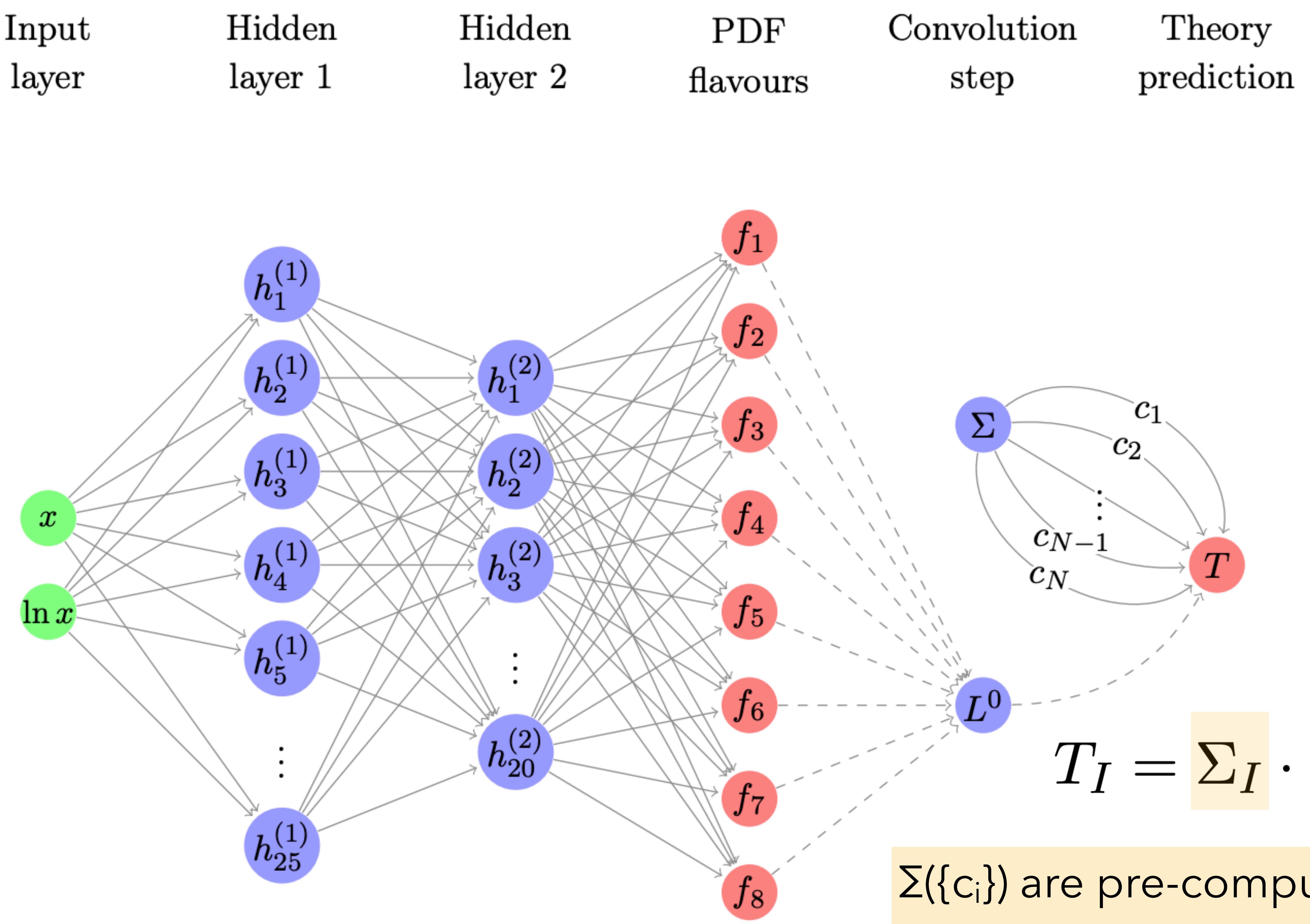
- The idea: take a PDF fit based on NNPDF4.0 methodology and make dependence of observables on physics parameters {c_i} explicit before computing the loss function (e.g. adding SMEFT corrections, or expanding observables in terms of SM precision parameters)

- Perform minimisation of loss function over

$$\hat{\theta} = \theta \bigcup \{c_i\}$$

by adding new layer to the deep neural network used in NNPDF4.0

- Can expand dependence on c_i beyond linear terms in T (up to generic power in polynomial expansion) by adding non-trainable edges
- Can be done both for SM parameters and SMEFT coefficients.



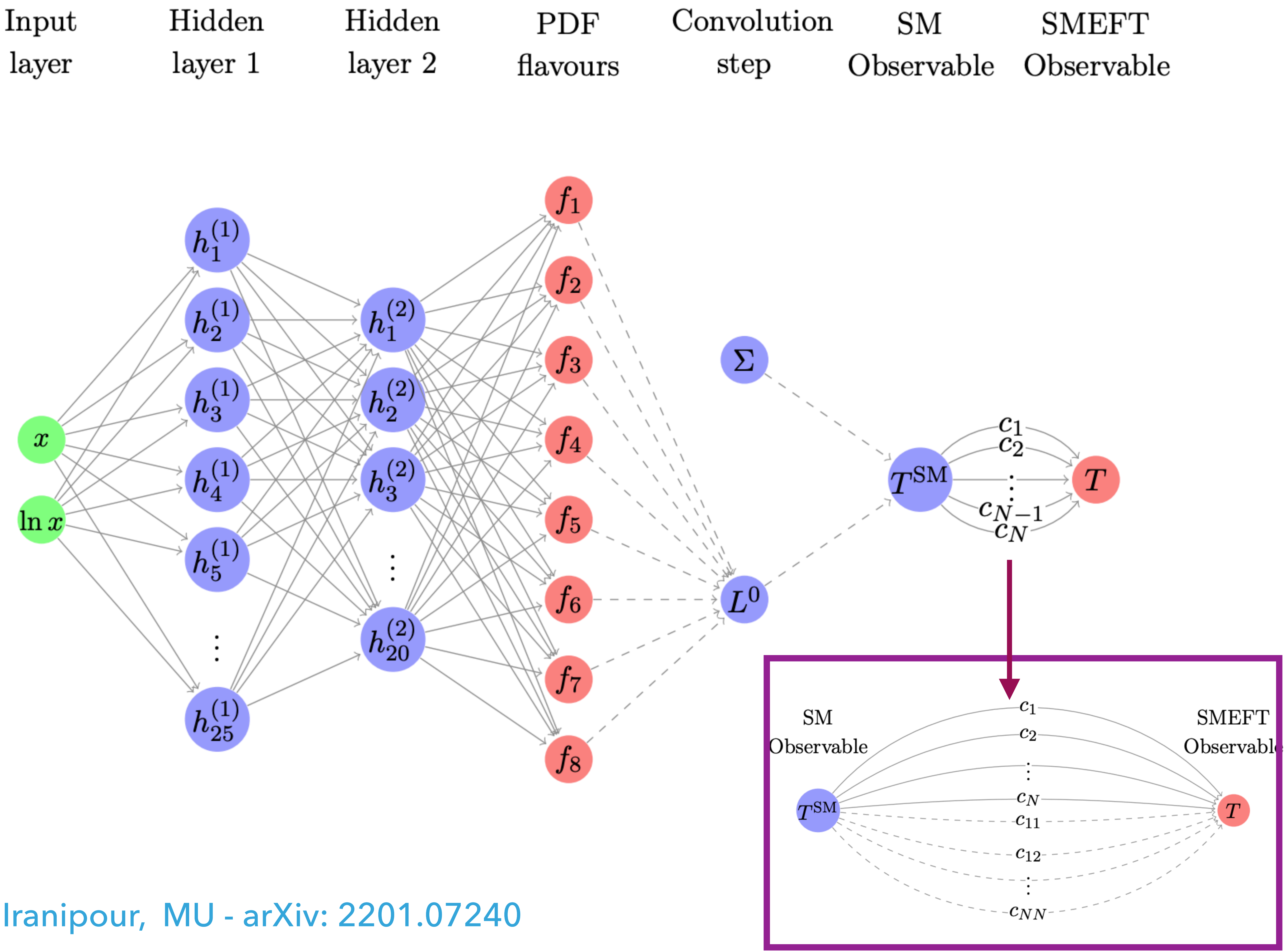
$$T_I = \Sigma_I \cdot L^0$$

Σ({c_i}) are pre-computed tables for fast interface accounting for PDF evolution and part. xsec

S. Iranipour, MU - arXiv: 2201.07240

THE SIMUNET ANALYSIS

- SimuNET yields a truly simultaneous fit, rather than a scan in benchmark point in WC space and it does not have limit in number of parameters that can be fitted alongside PDFs at the initial scale!



Linear dim-6 operator

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

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

Quadratic dim-6 operator

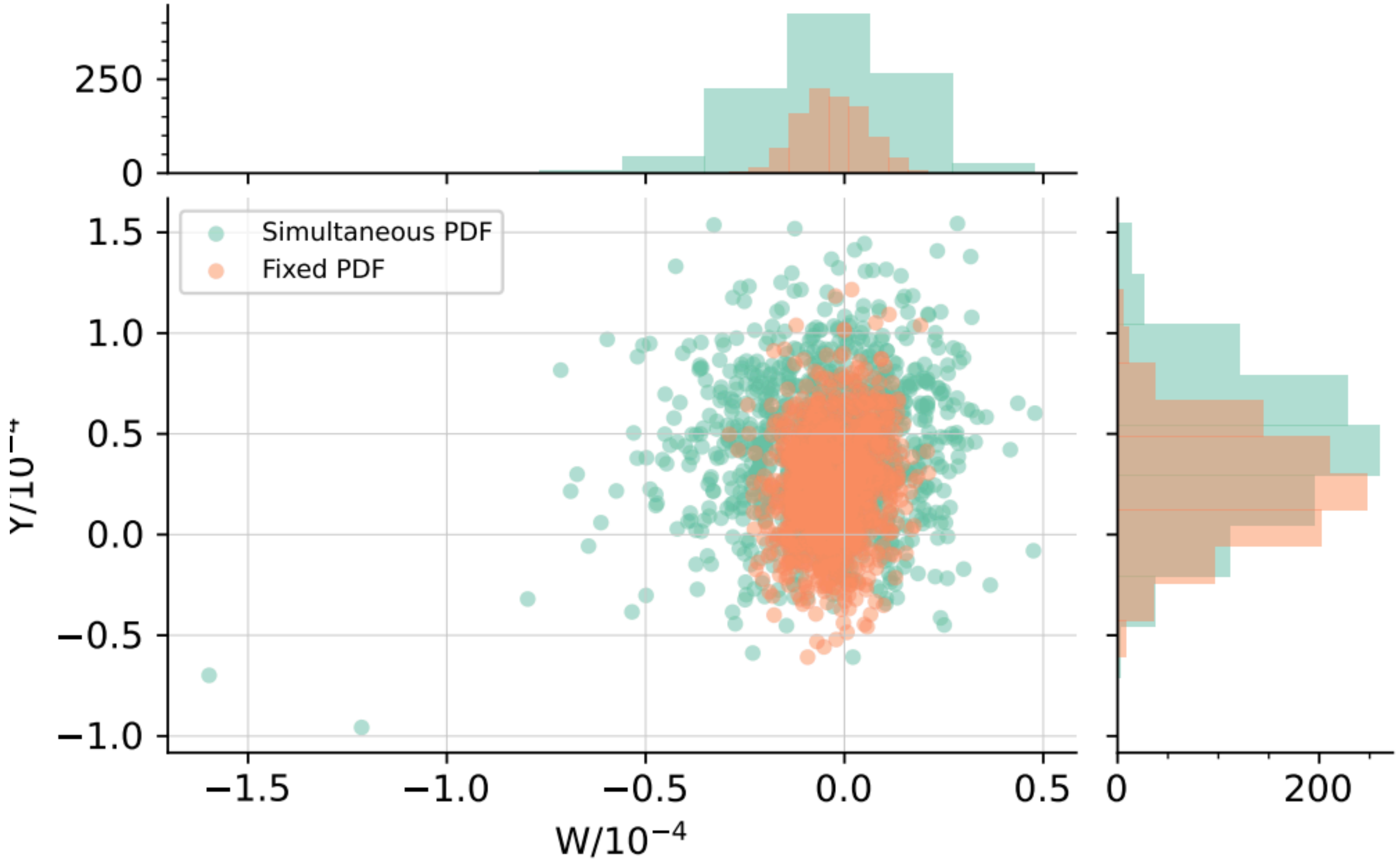
$$T(\hat{\theta}) = T^{\text{SM}}(\theta) \cdot \left(1 + \sum_{n=1}^{N} c_n R_{\text{SMEFT}}^{(n)} + \sum_{1 \le n \le m \le N} c_{nm} R_{\text{SMEFT}}^{(n,m)}, \right)$$

$c_n c_m$

S. Iranipour, MU - arXiv: 2201.07240

RESULTS: DRELL-YAN DATA @HL-LHC

S. Iranipour, MU - arXiv: 2201.07240



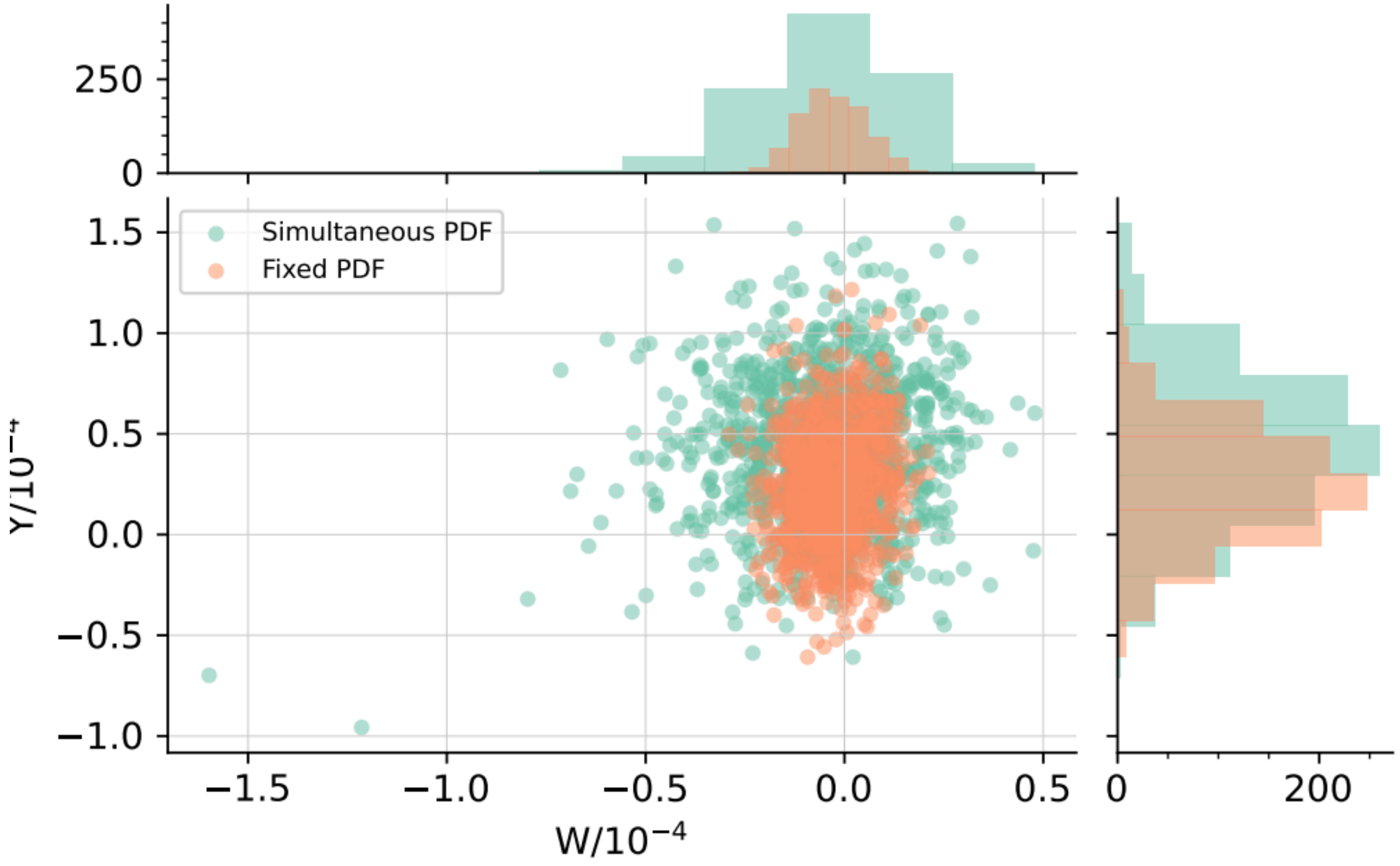
	SM PDFs	SMEFT PDFs
$W \times 10^5$ (68% CL)	$[-1.1, 0.5]$	$[-2.4, 1.5]$
$W \times 10^5$ (95% CL)	$[-2.0, 1.4]$	$[-4.3, 3.4]$
$Y \times 10^5$ (68% CL)	$[-0.4, 5.2]$	$[0.6, 8.0]$
$Y \times 10^5$ (95% CL)	$[-3.2, 8.1]$	$[-3.1, 11.7]$

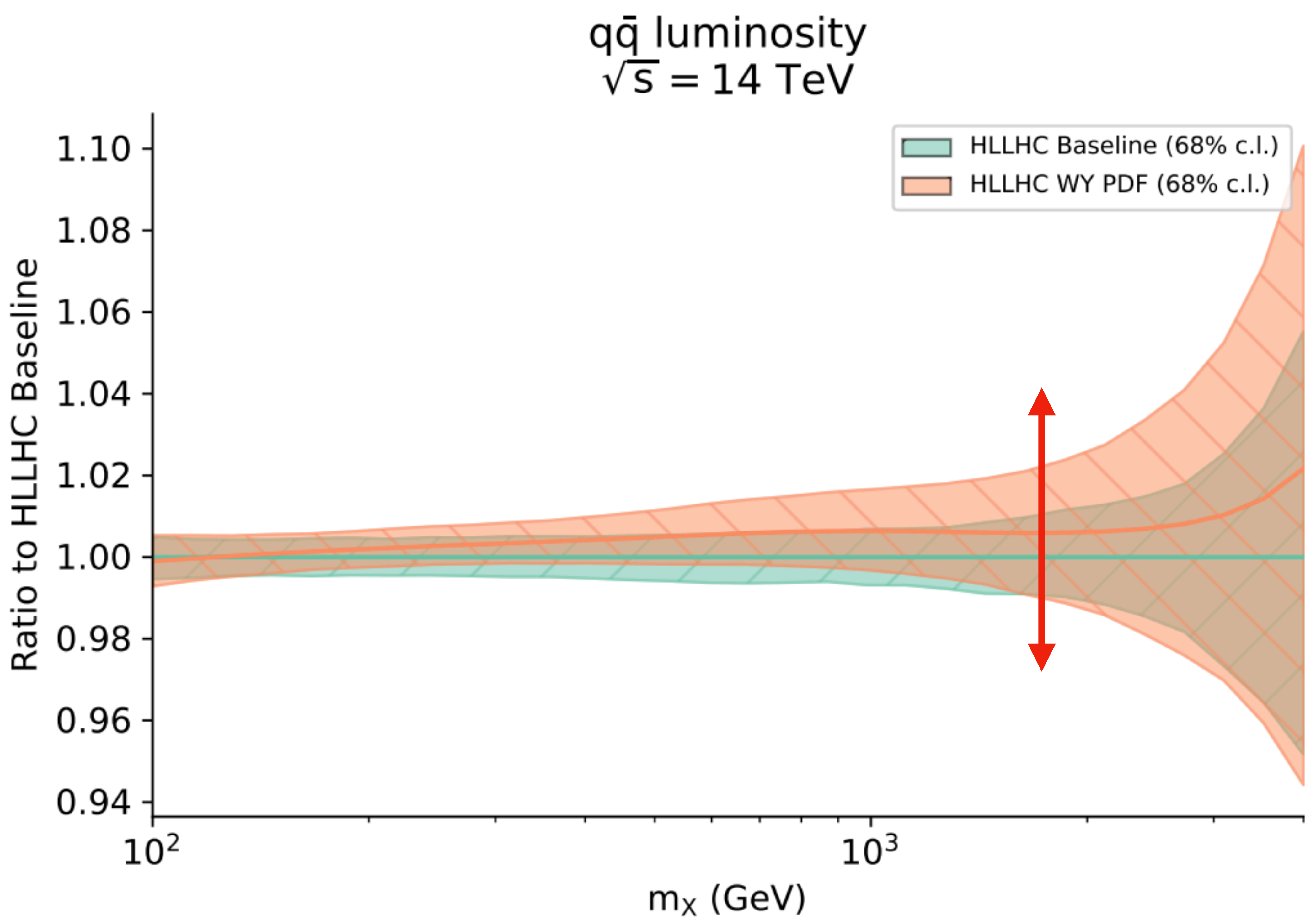
x 2.3 broadening of bounds for W
x 1.3 broadening of bounds for Y

✓ 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 PDF uncertainties change significantly once SMEFT effects allowed in theory predictions entering PDF fit

RESULTS: DRELL-YAN DATA @HL-LHC

S. Iranipour, MU - arXiv: 2201.07240

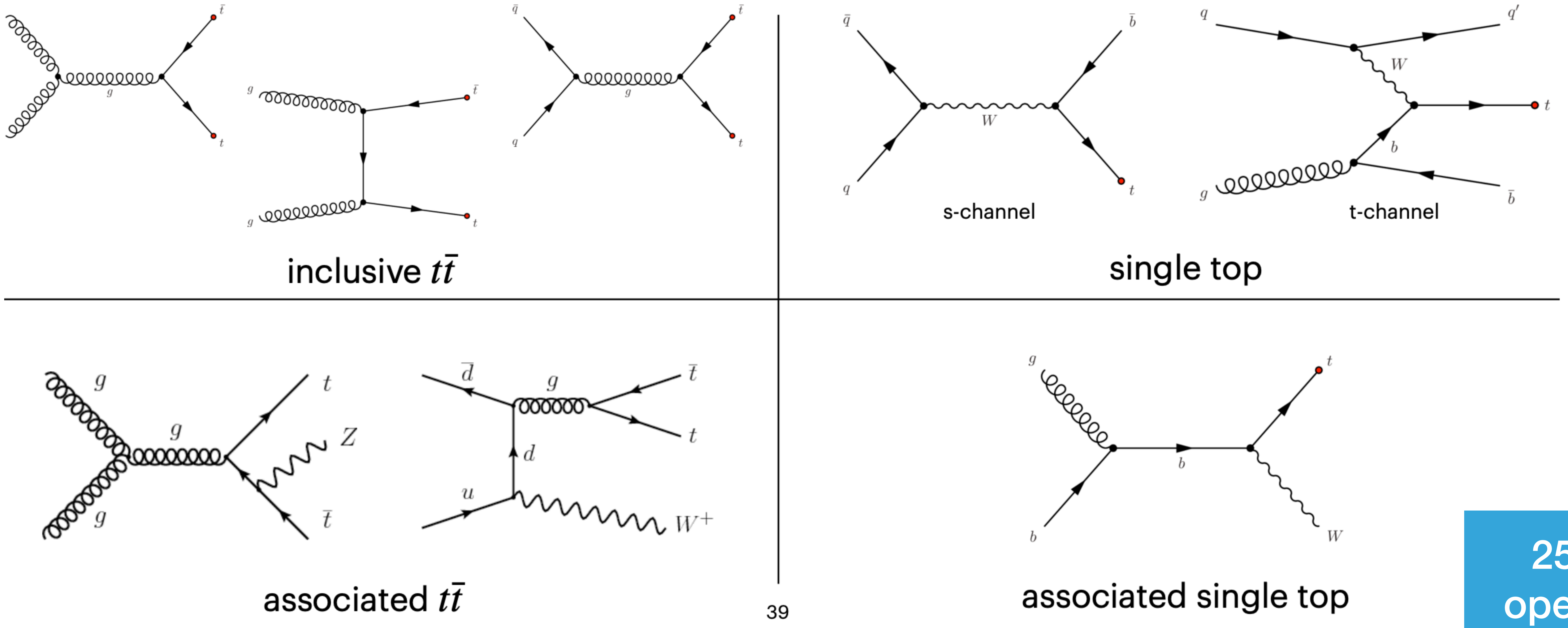




✓ Simultaneous analysis of PDFs and W&Y SMEFT coefficient of DIS + DY (including HL-LHC projections) using simuNET method shows that at HL-LHC the effect of interplay becomes important as WCs bounds broaden and PDF uncertainties change significantly once SMEFT effects allowed in theory predictions entering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

THE TOP SECTOR

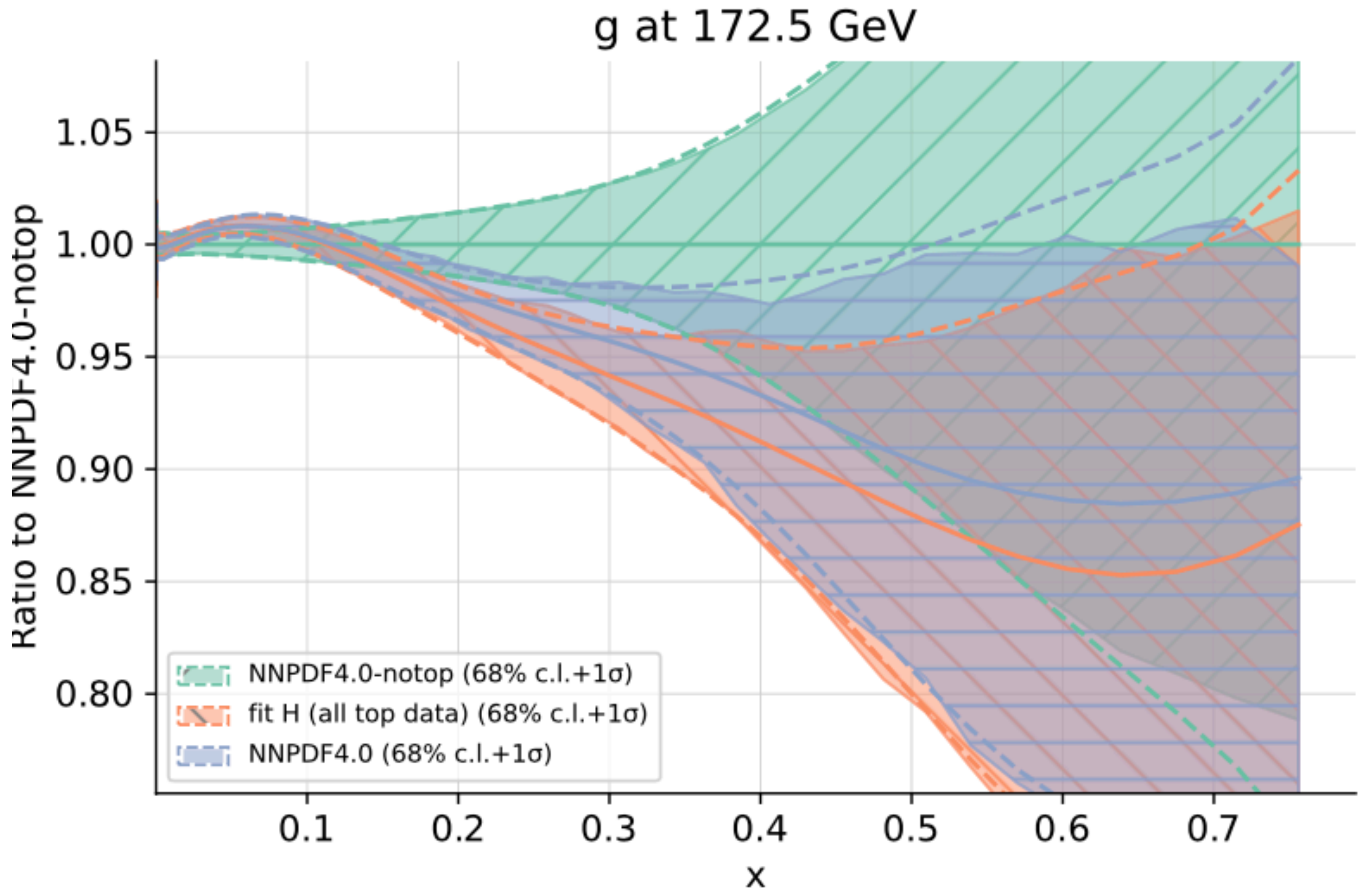
- After testing methodology on small number of WC, stress-test on large SMEFT parameters space.
- Huge amount of Run II top quark data from ATLAS and CMS
- Four basic processes: inclusive tt~ and asymmetry (inclusive and differential), single top (inclusive and differential), associated ttV production,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

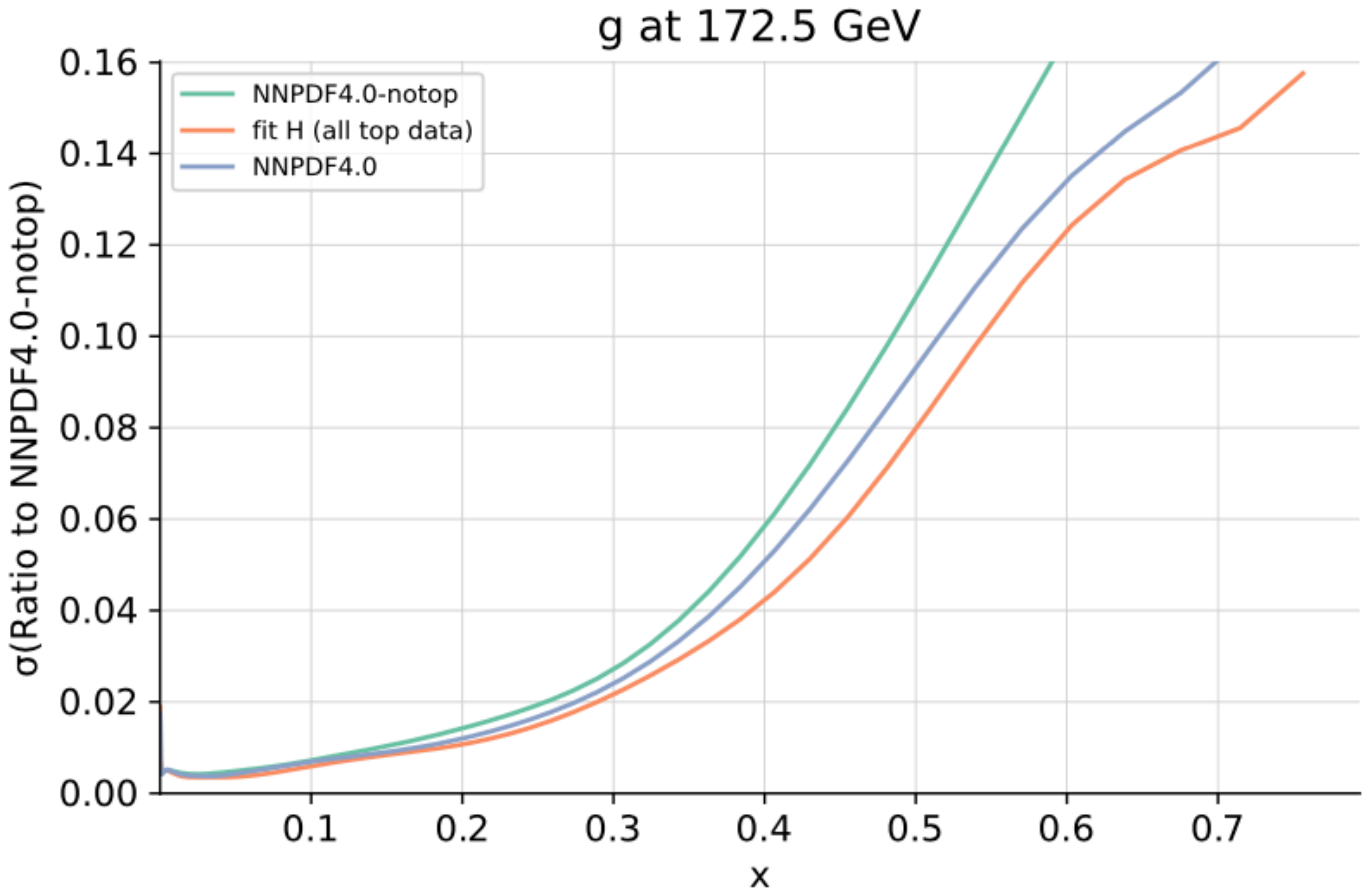


25 (21) dim-6 operators at the quadratic (linear) SMEFT

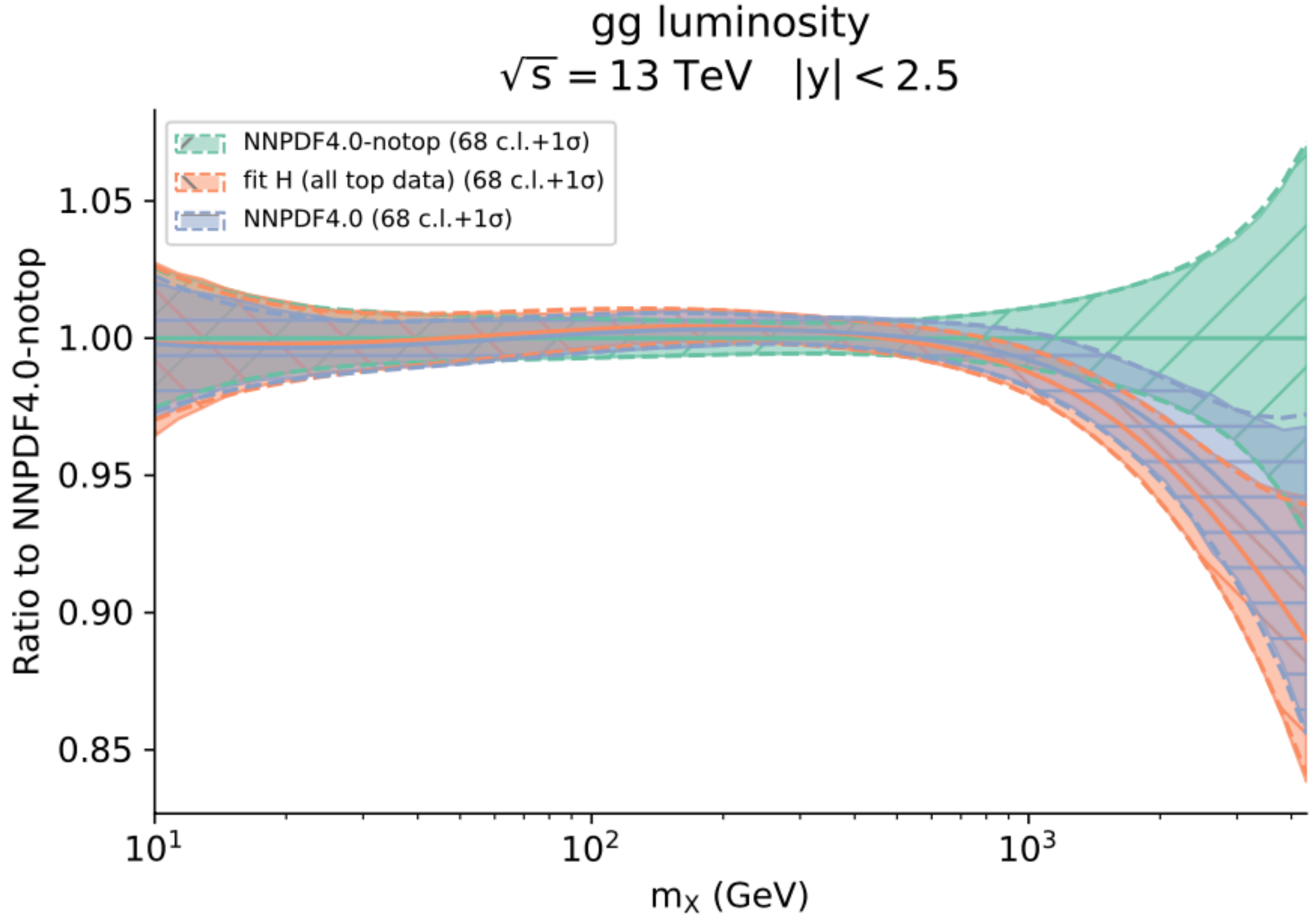
Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU - arXiv: 2303.06159

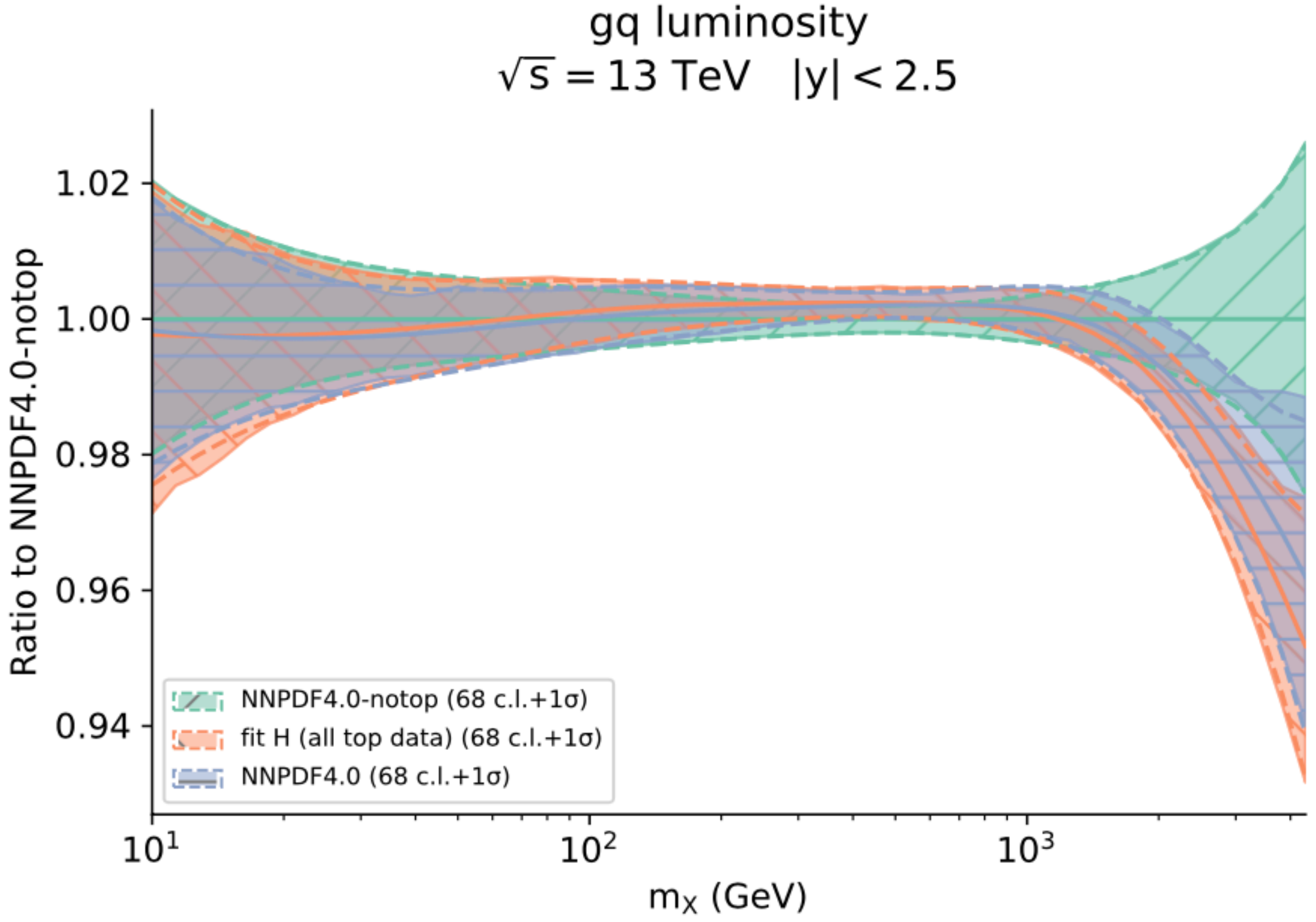
PDF-ONLY FIT





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



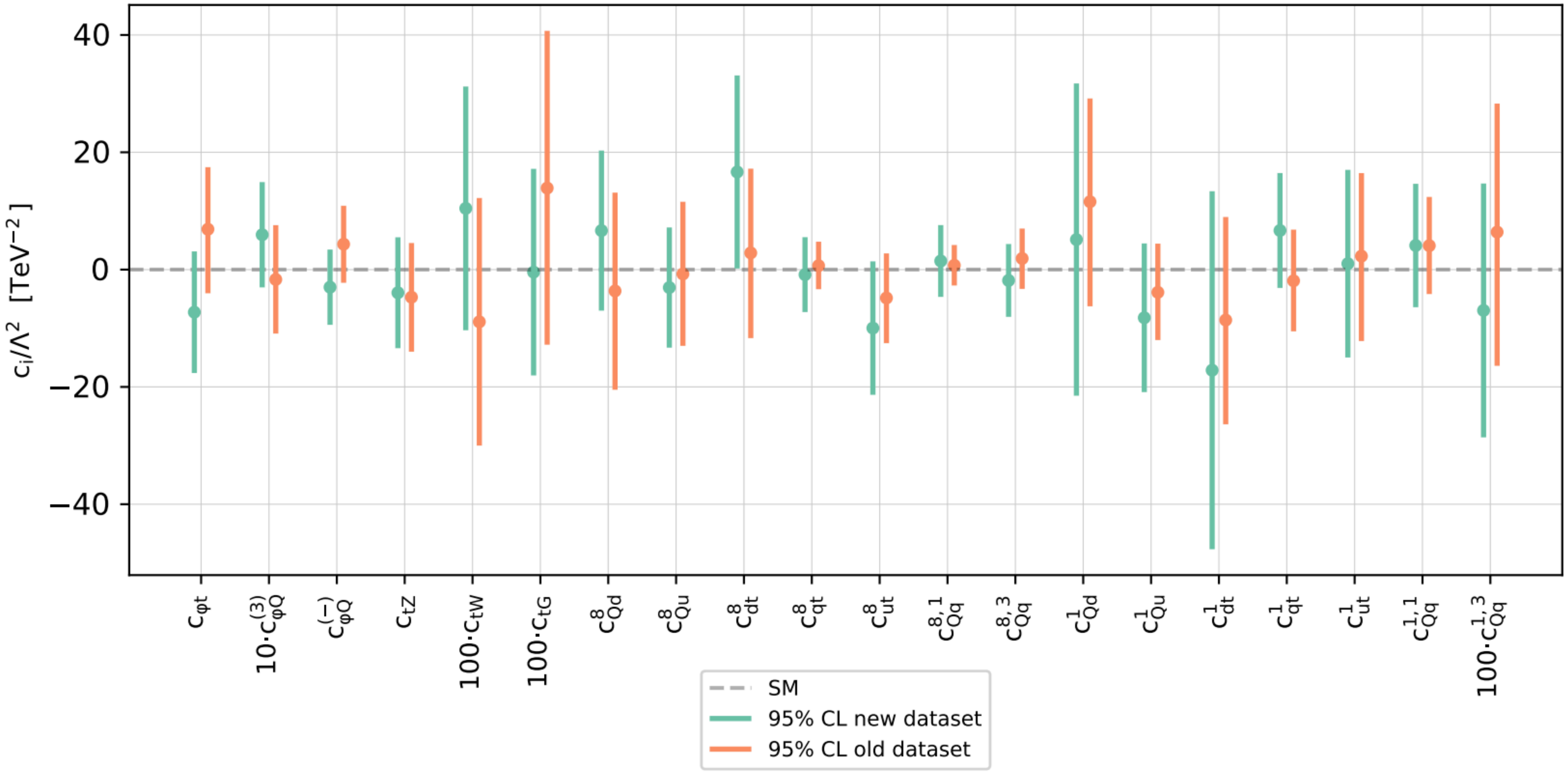


Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU - arXiv: 2303.06159

SMEFT-ONLY FIT

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

Linear SMEFT

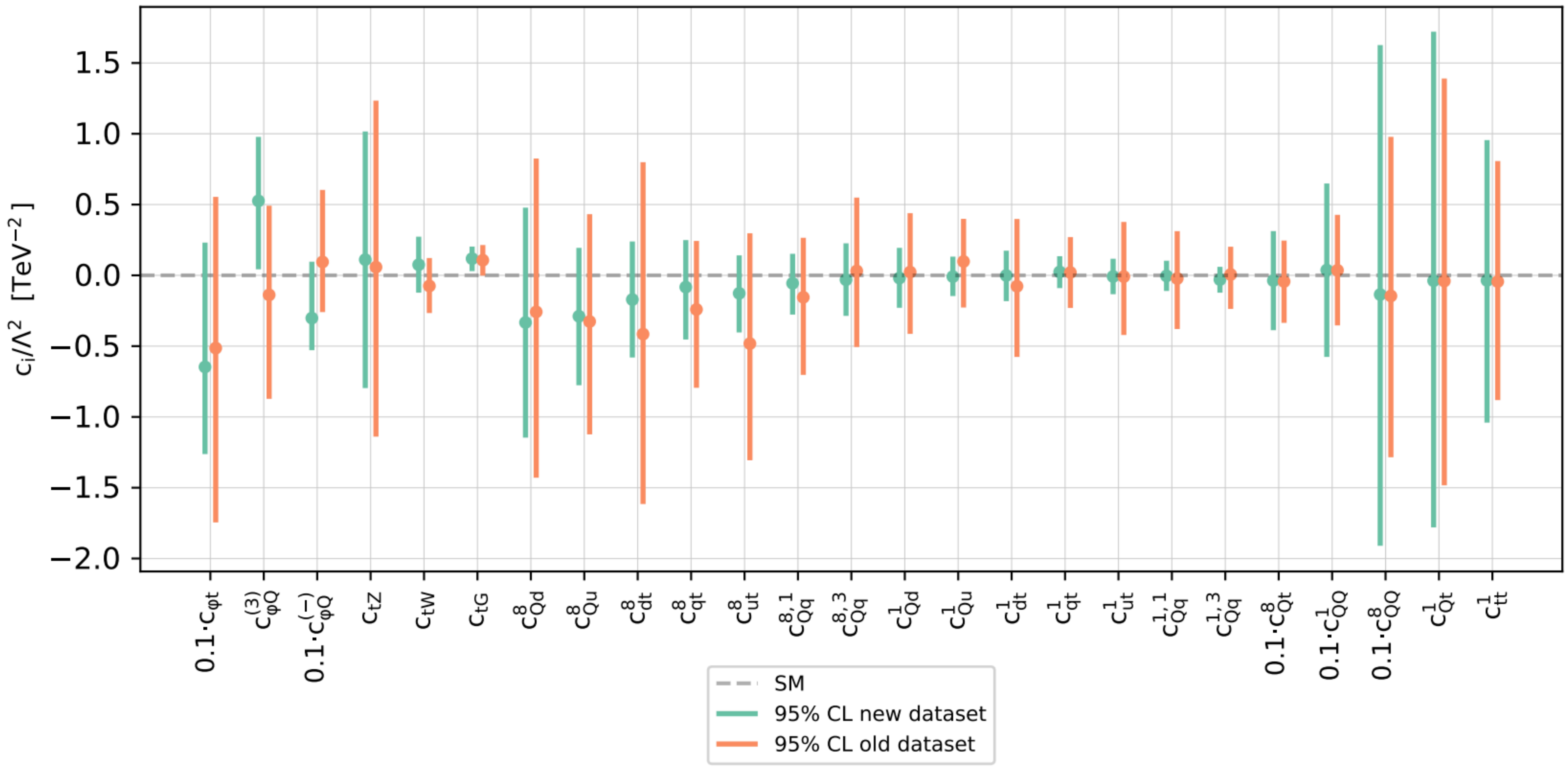


Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU - arXiv: 2303.06159

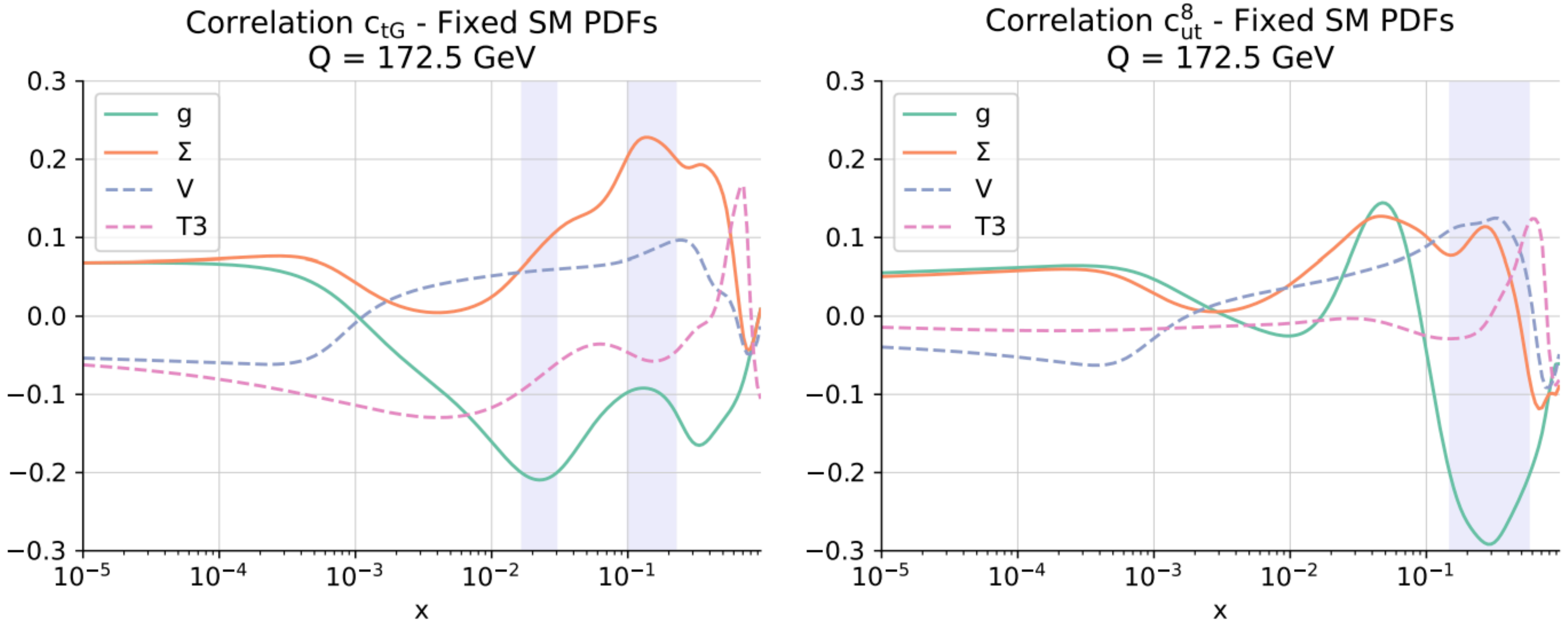
SMEFT-ONLY FIT

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

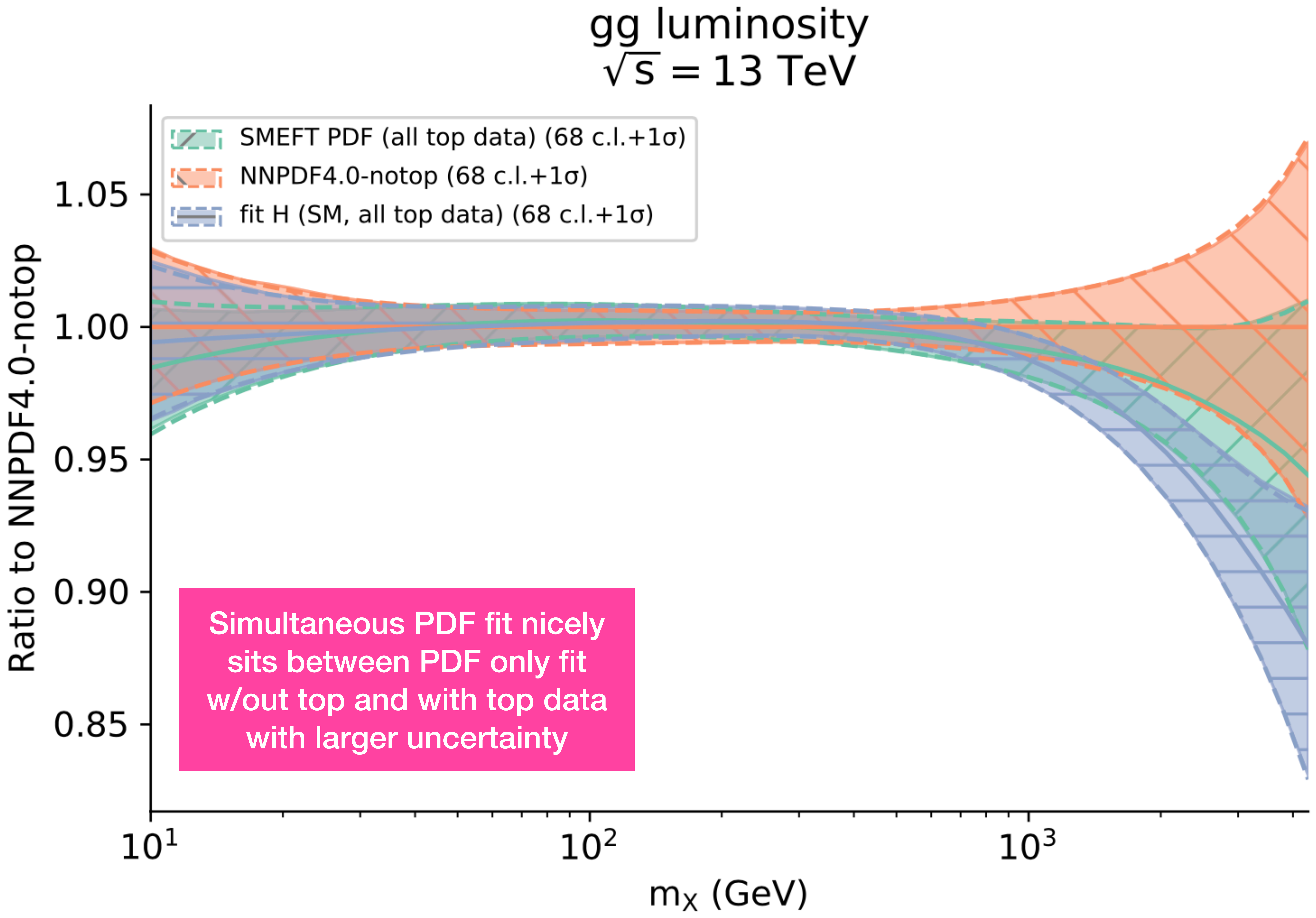
Quadratic SMEFT



SMEFT AND PDF CORRELATIONS



SIMULTANEOUS PDF-SMEFT FIT



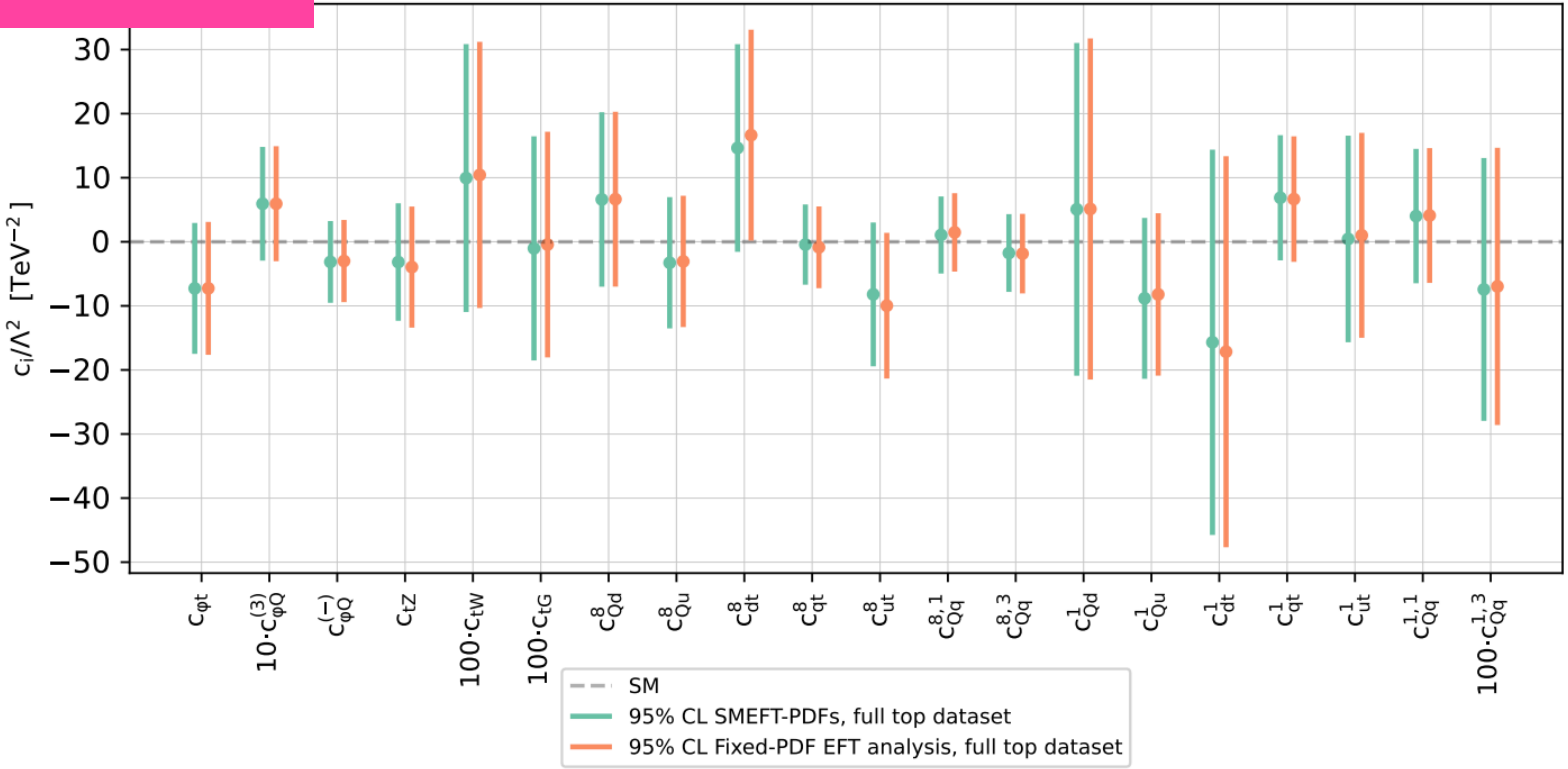
Linear SMEFT

Simultaneous PDF fit nicely sits between PDF only fit w/out top and with top data with larger uncertainty

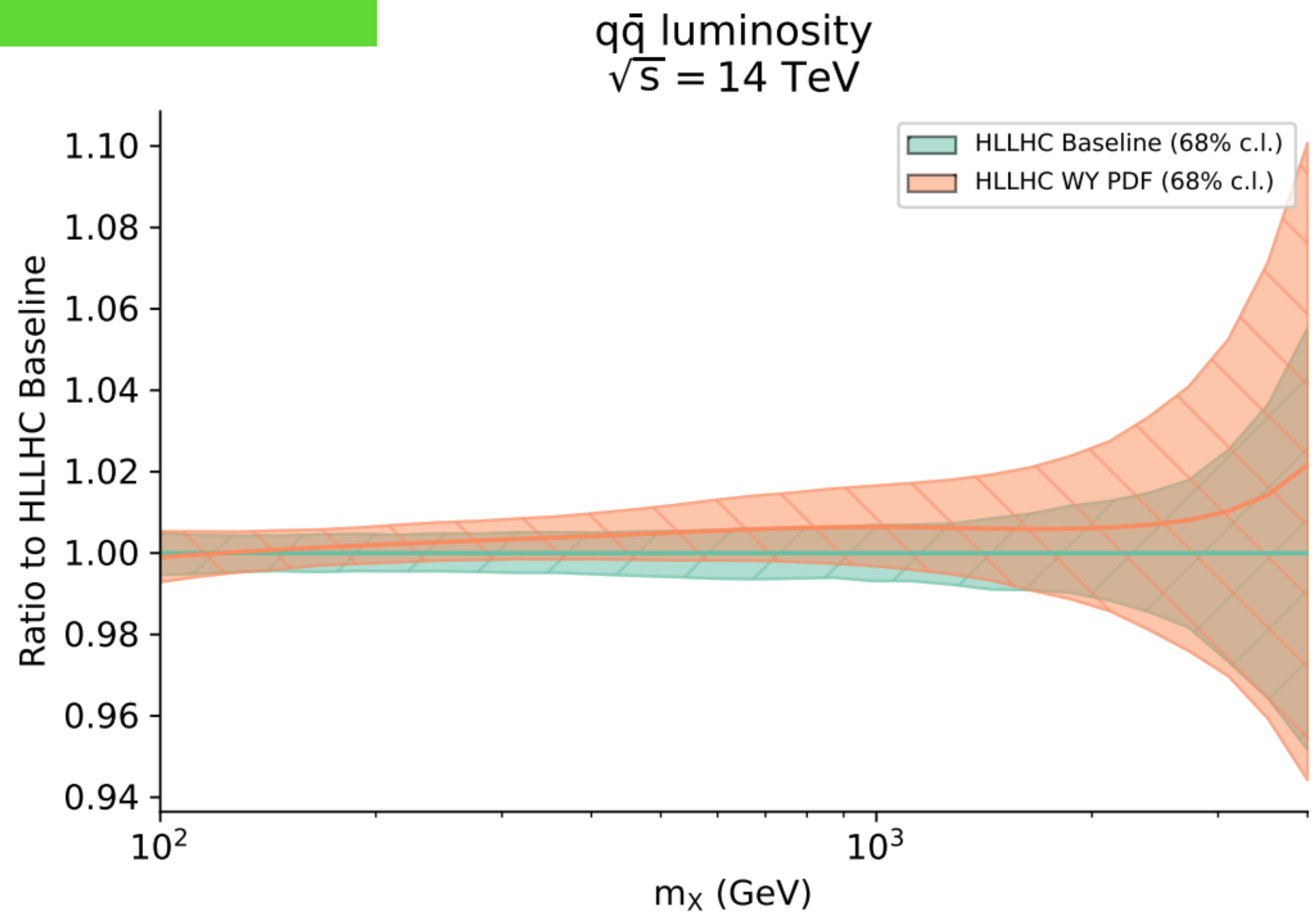
SIMULTANEOUS PDF-SMEFT FIT

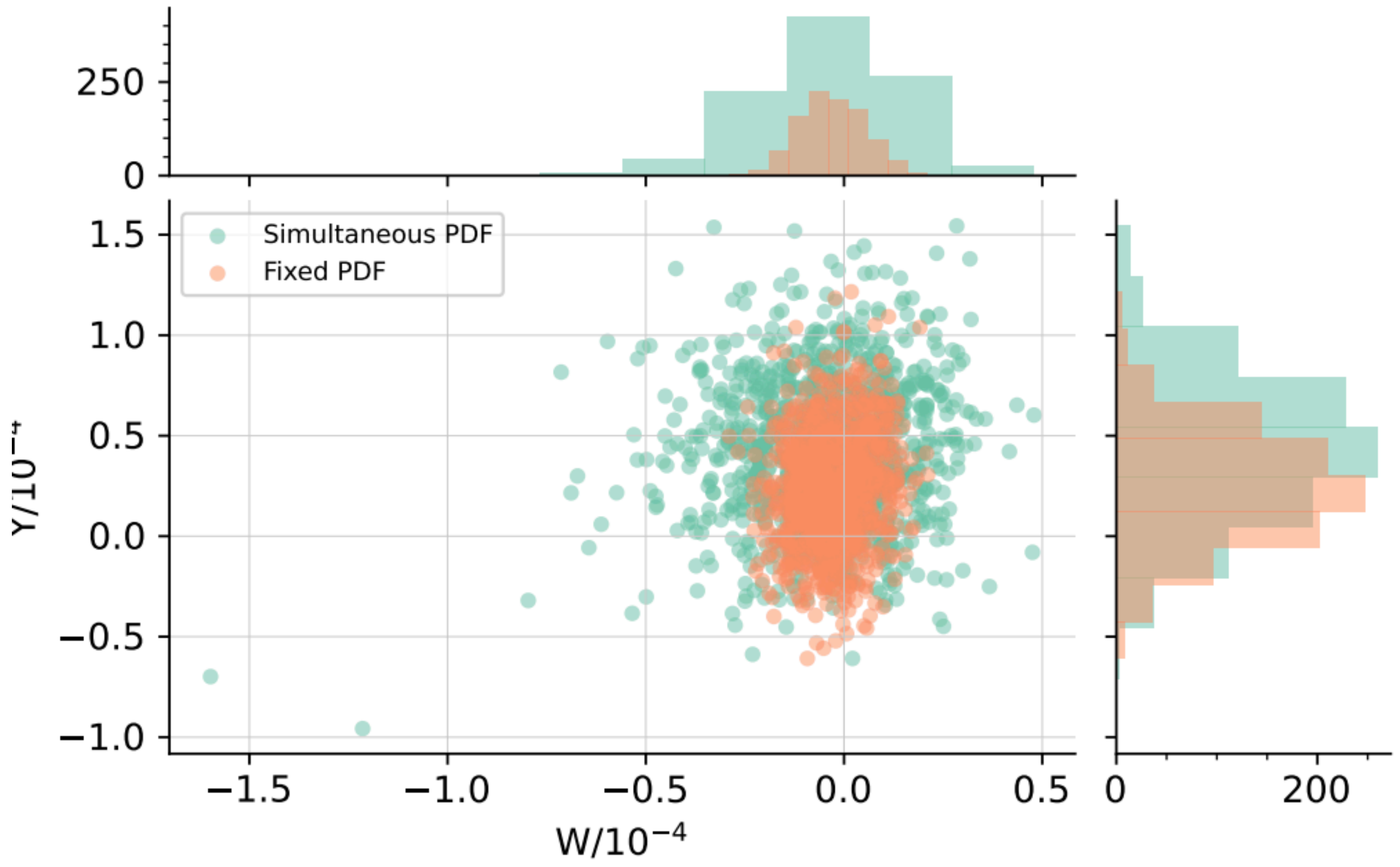
WCs stable upon PDF variations

Linear SMEFT

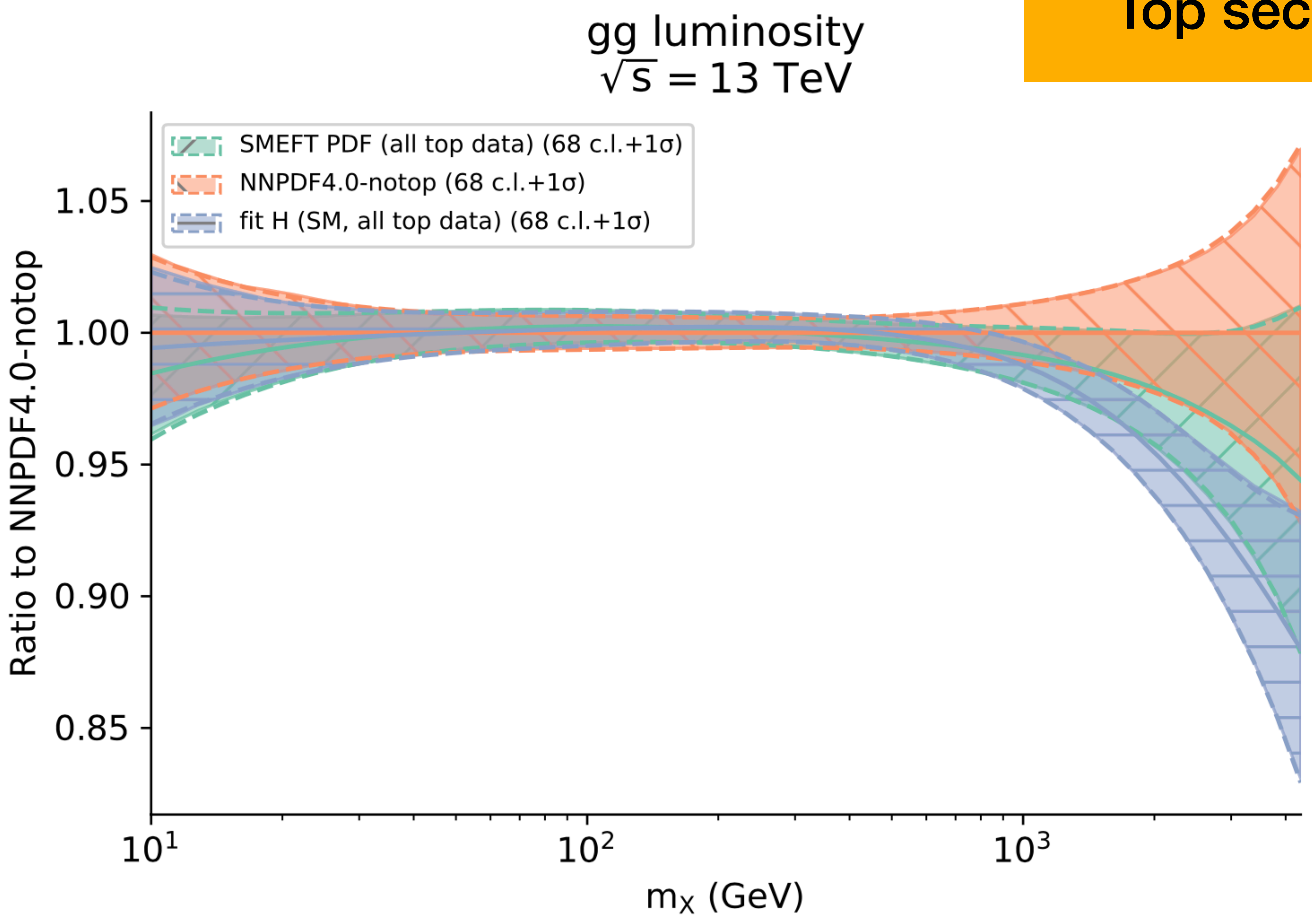


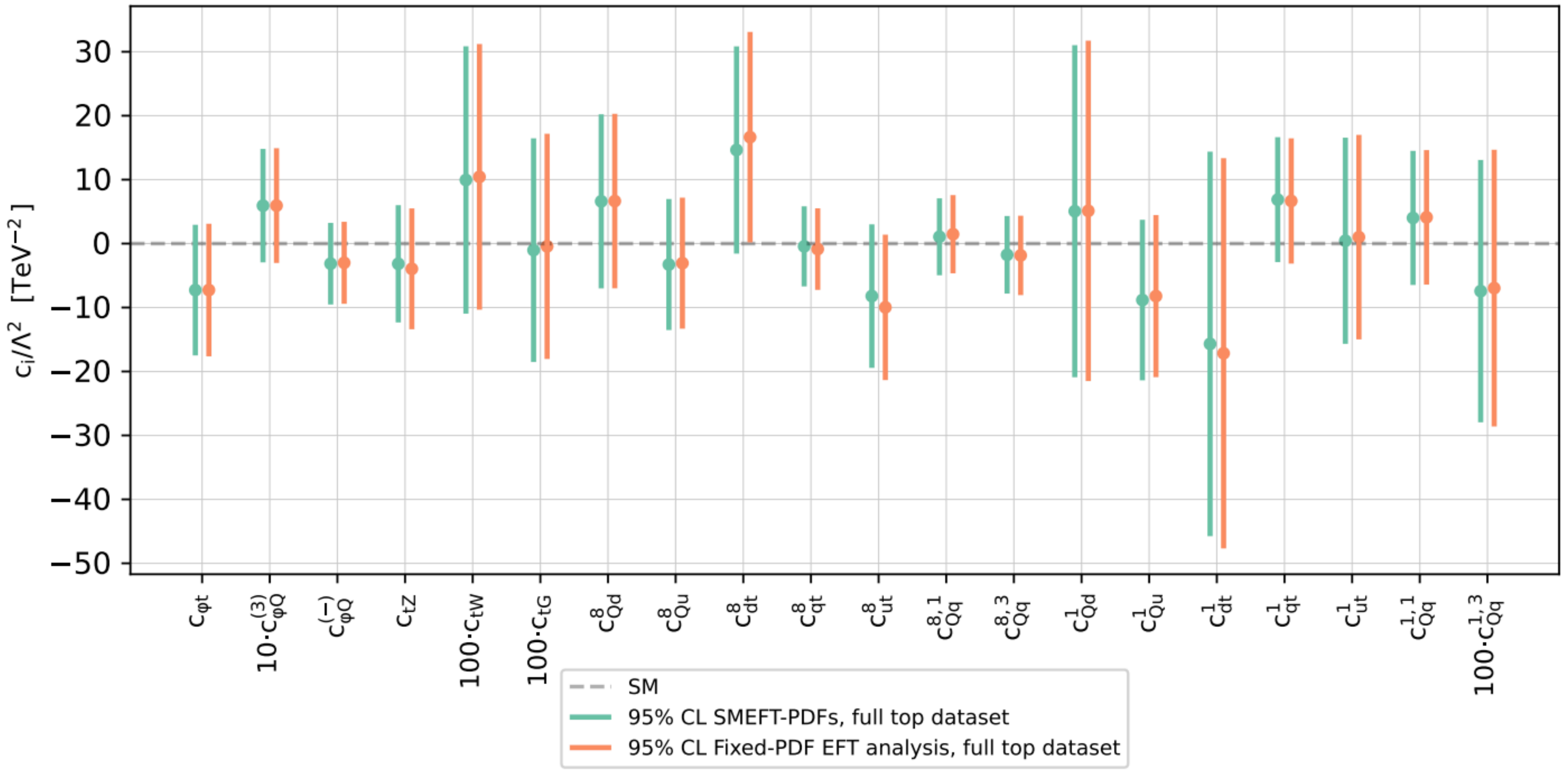
DY sector





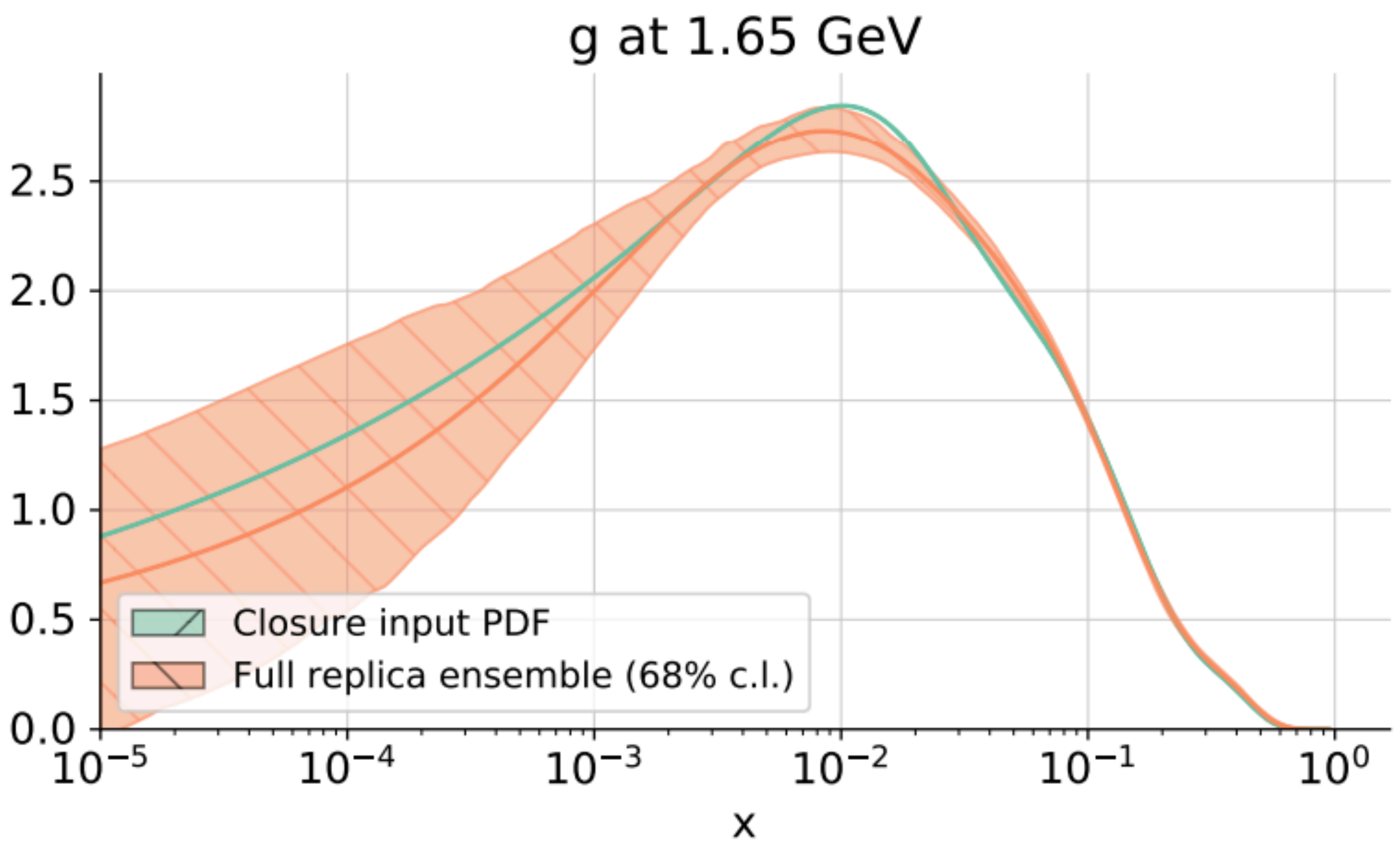
Top sector





CAN PDFS ABSORB NEW PHYSICS?

CAN PDFS ABSORB NEW PHYSICS?

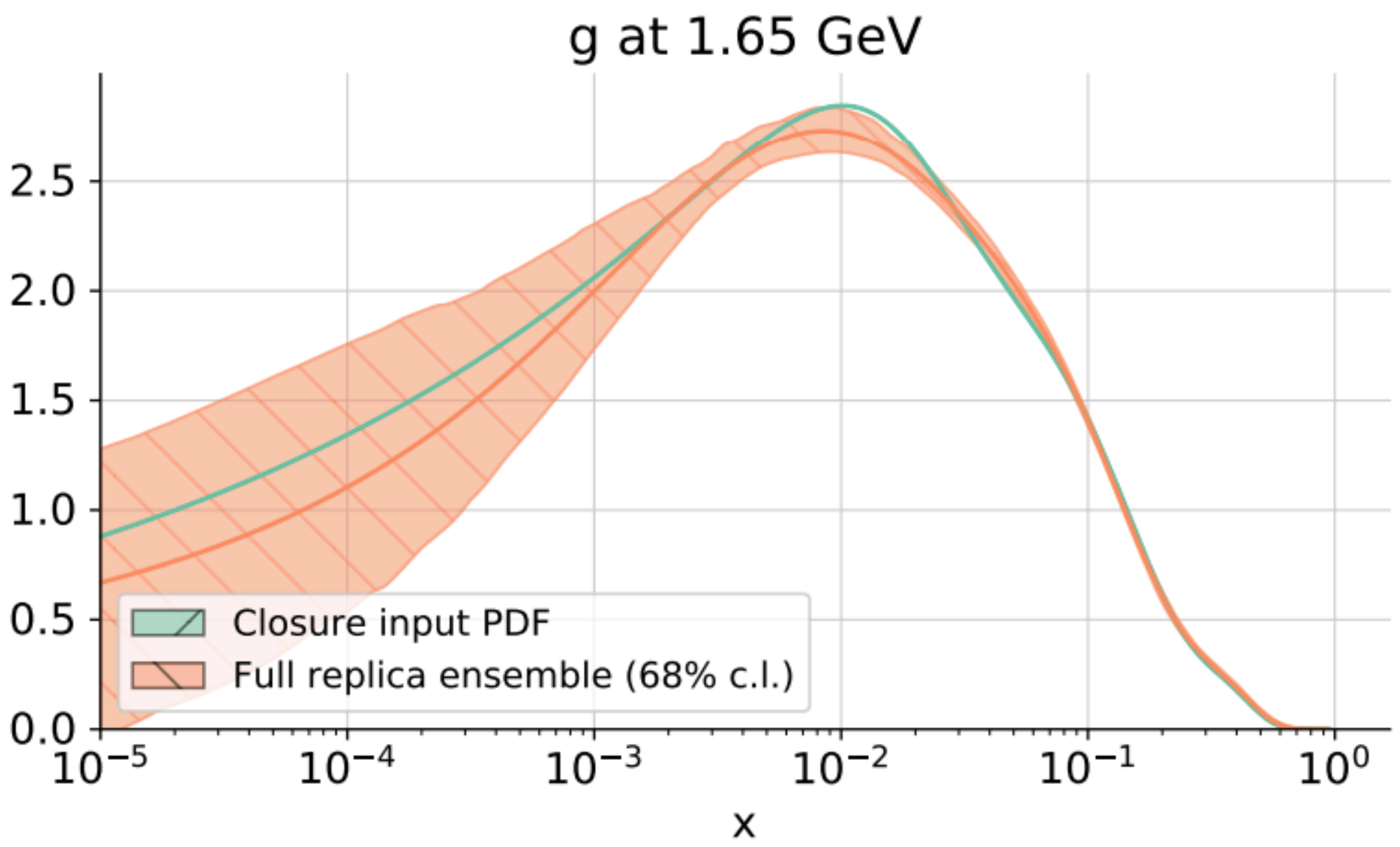


✓ NNPDF methodology routinely tested via closure test (in the data region) [Del Debbio, Giani, Wilson, 2111.05787] and future test (in the extrapolation region) [Cruz-Martinez, Forte, Nocera, 2103.08606].

✓ 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 PDFS ABSORB NEW PHYSICS?

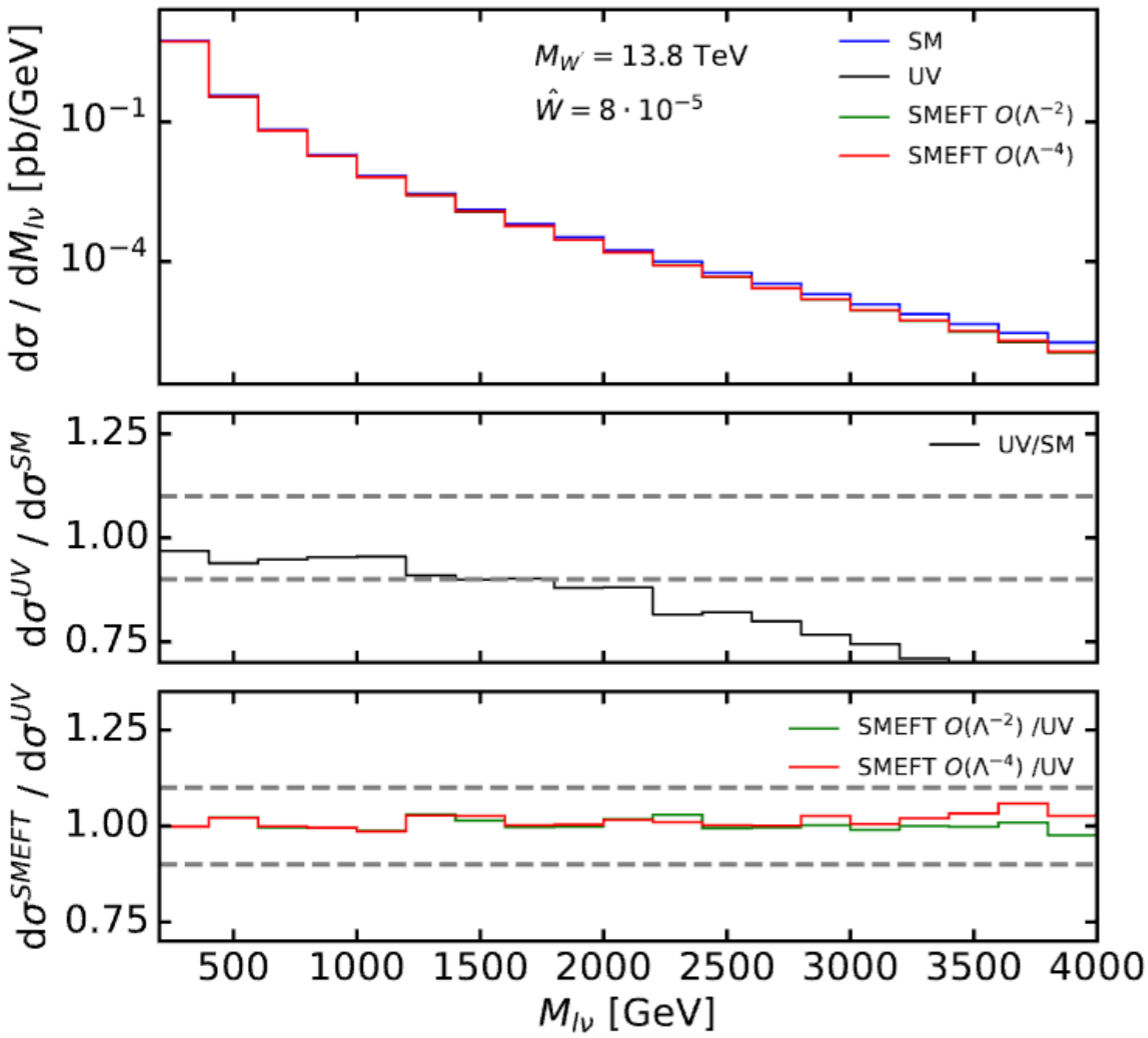


- Imagine that on top of the "true" PDFs one inject the "true" UV model in the MC data
- Generate artificial MC data assuming "true" law of nature =
  "true" PDFs + "true" UV model
- Fit PDFs assuming SM
- Can PDFs absorb signs of new physics?

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

✓ 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

A Z' AND W' TEST-CASE

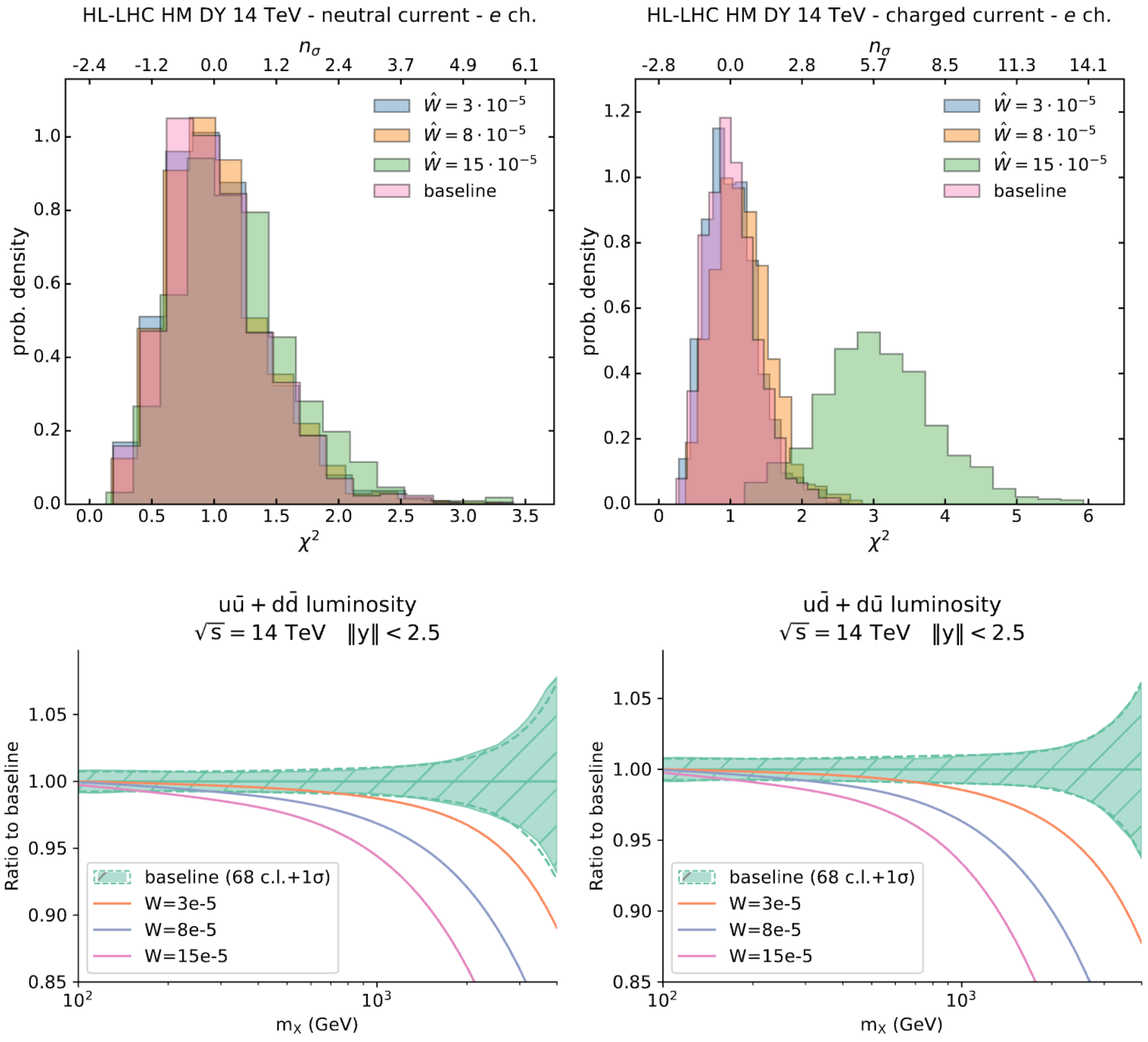


- Imagine that on top of the "true" PDFs one inject the "true" UV model in the MC data
- Generate artificial MC data assuming "true" law of nature =
  "true" PDFs + "true" UV model
- Fit PDFs assuming SM
- Can PDFs absorb signs of new physics?

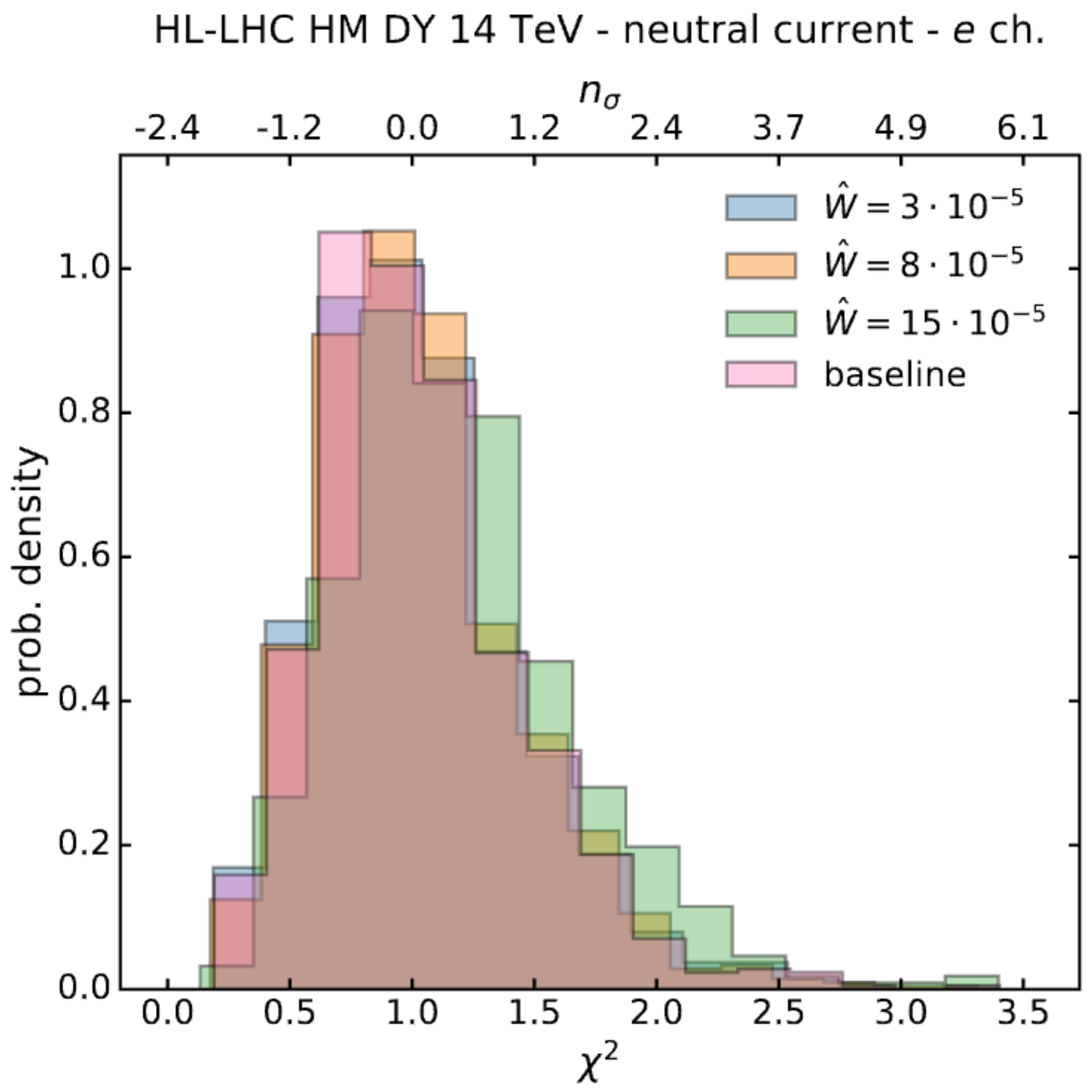
E. Hammou, Z. Kassabov, M. Madigan, M. Mangano, L. Mantani, J. Moore, M. Morales, MU
2307.10370

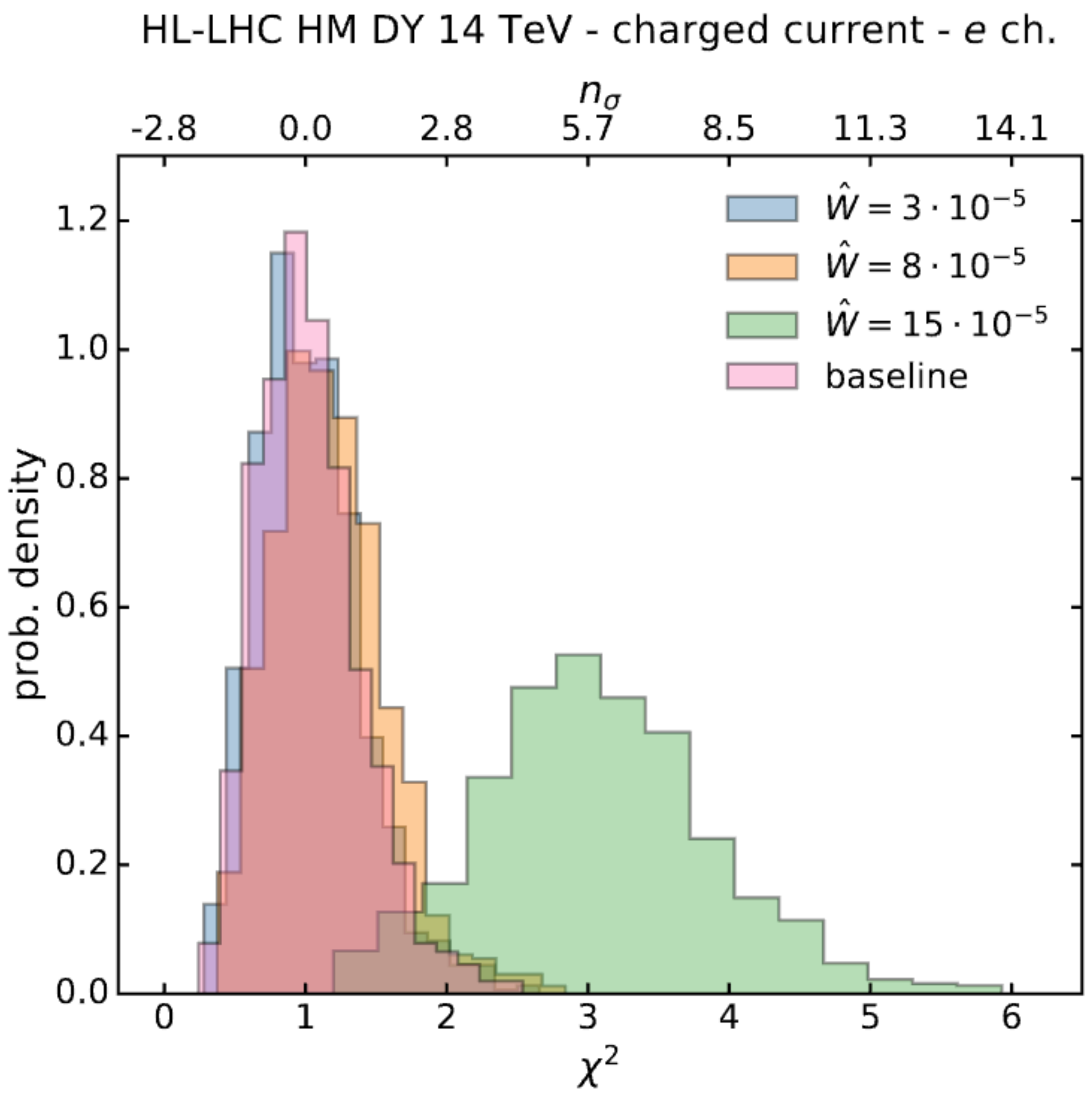
Z'	W'
Selection test: ✗	Selection test: ✓
➡ Excluded from PDF fit	➡ Included in PDF fit
No impact on PDFs	**PDFs contaminated**

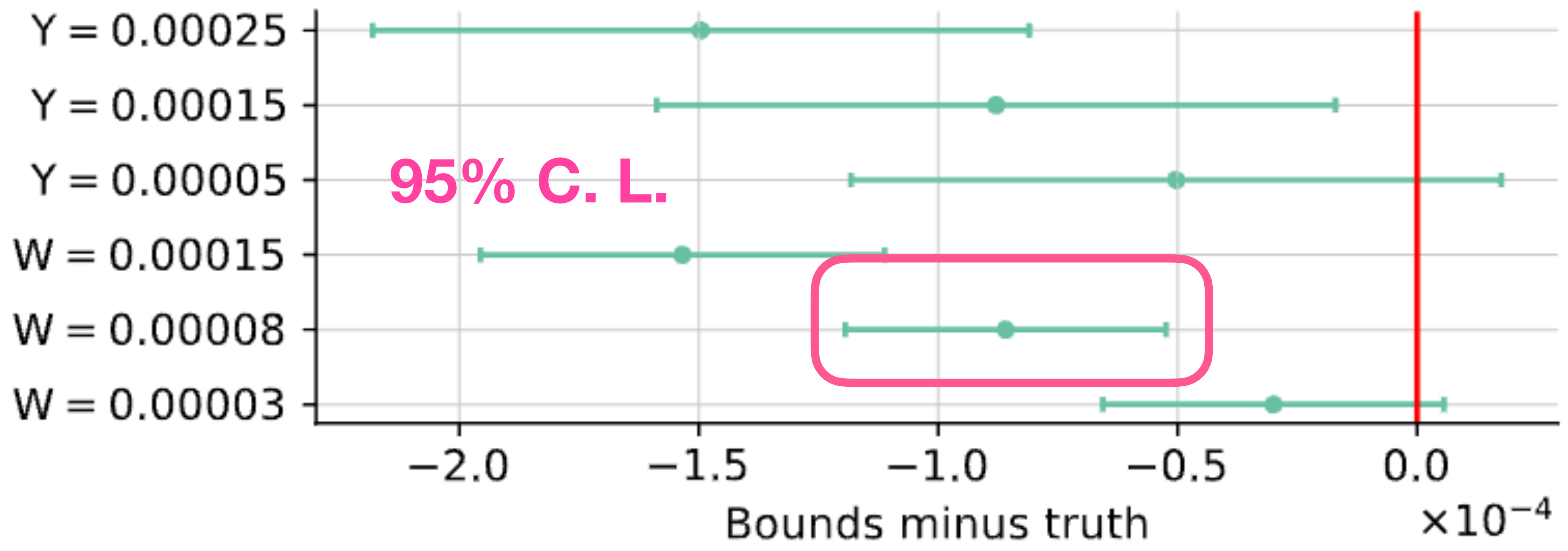
THE W' TEST-CASE

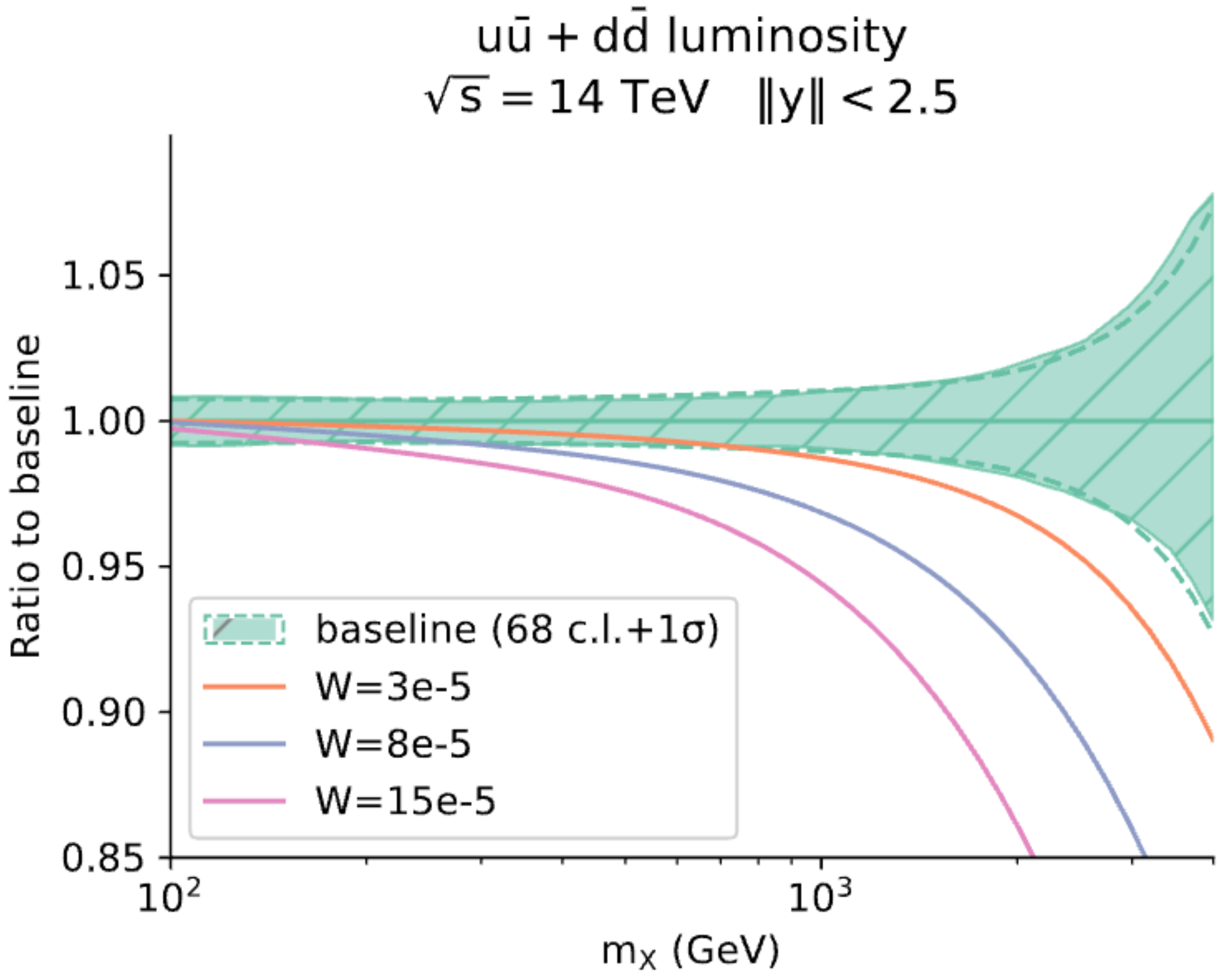


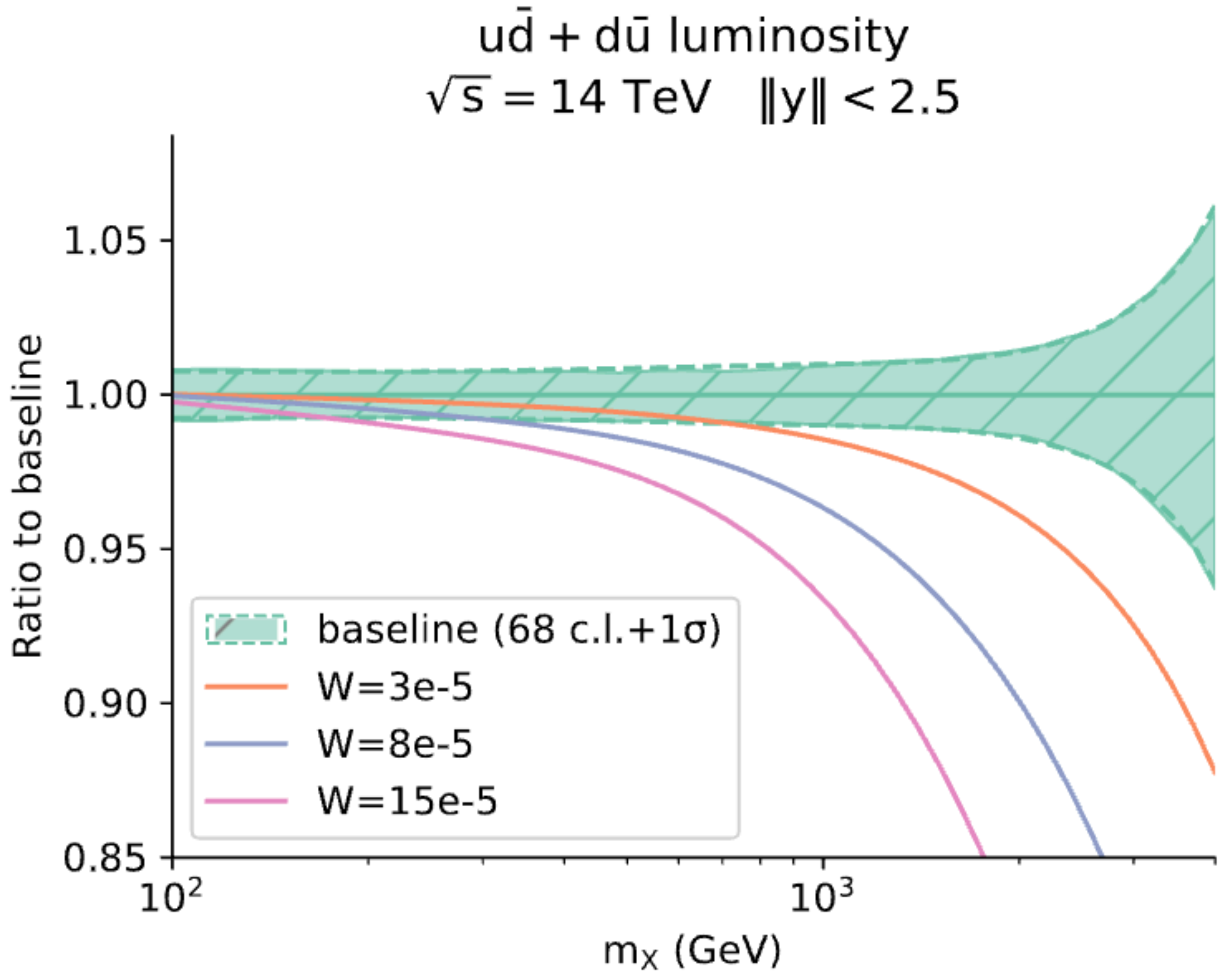
- The fit-quality of the closure test is unchanged up to W= 8e-5 (corresponding to Mw' = 13.8 TeV)
- Once we go beyond this point ,the fit-quality deteriorates due to the HL-LHC neutral current and charged current Drell-Yan MC data.
- Already for W= 8e-5 the qq~ luminosity shifts far beyond the PDF uncertainties 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.
- What are the consequences of such contamination?



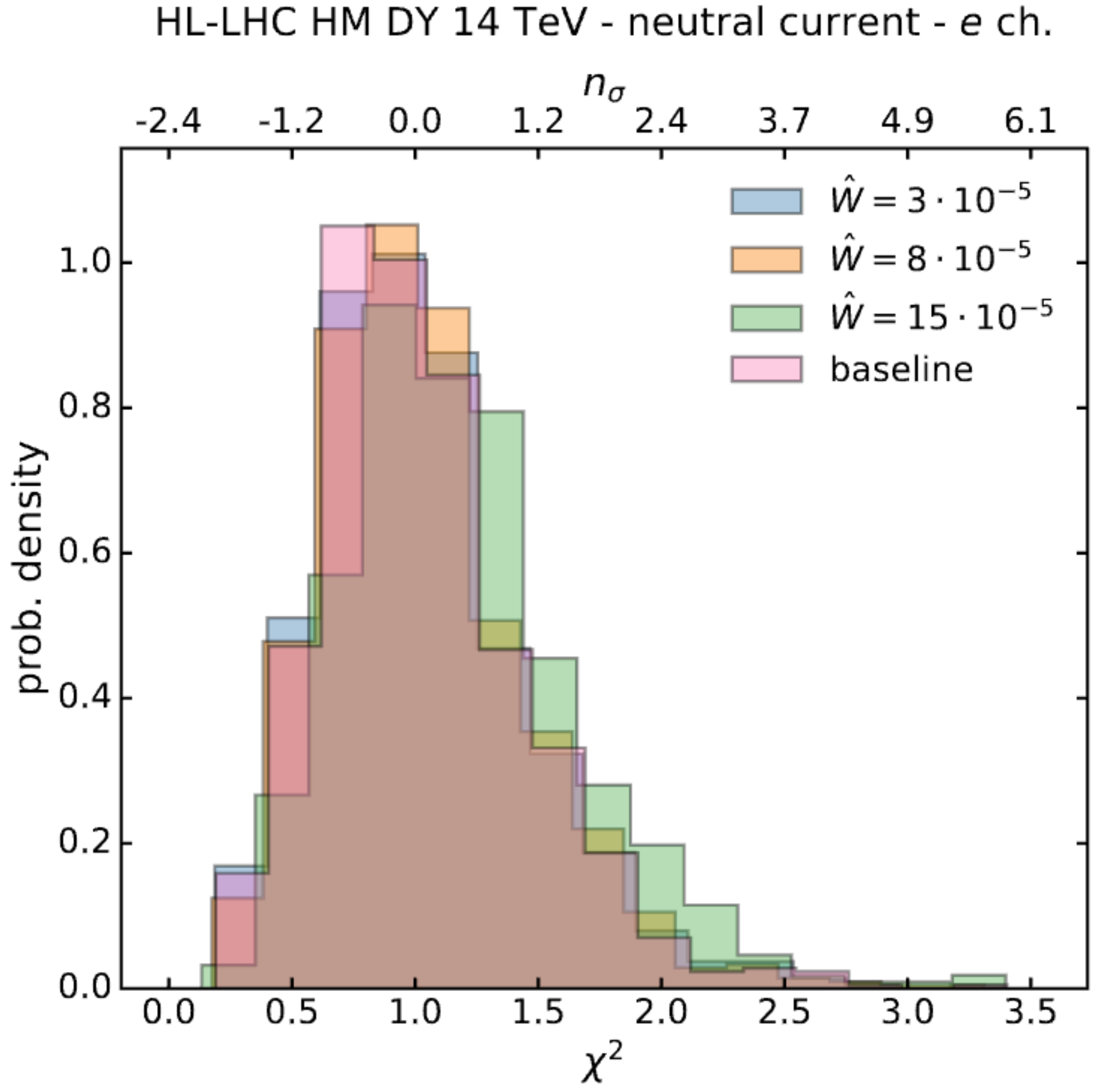


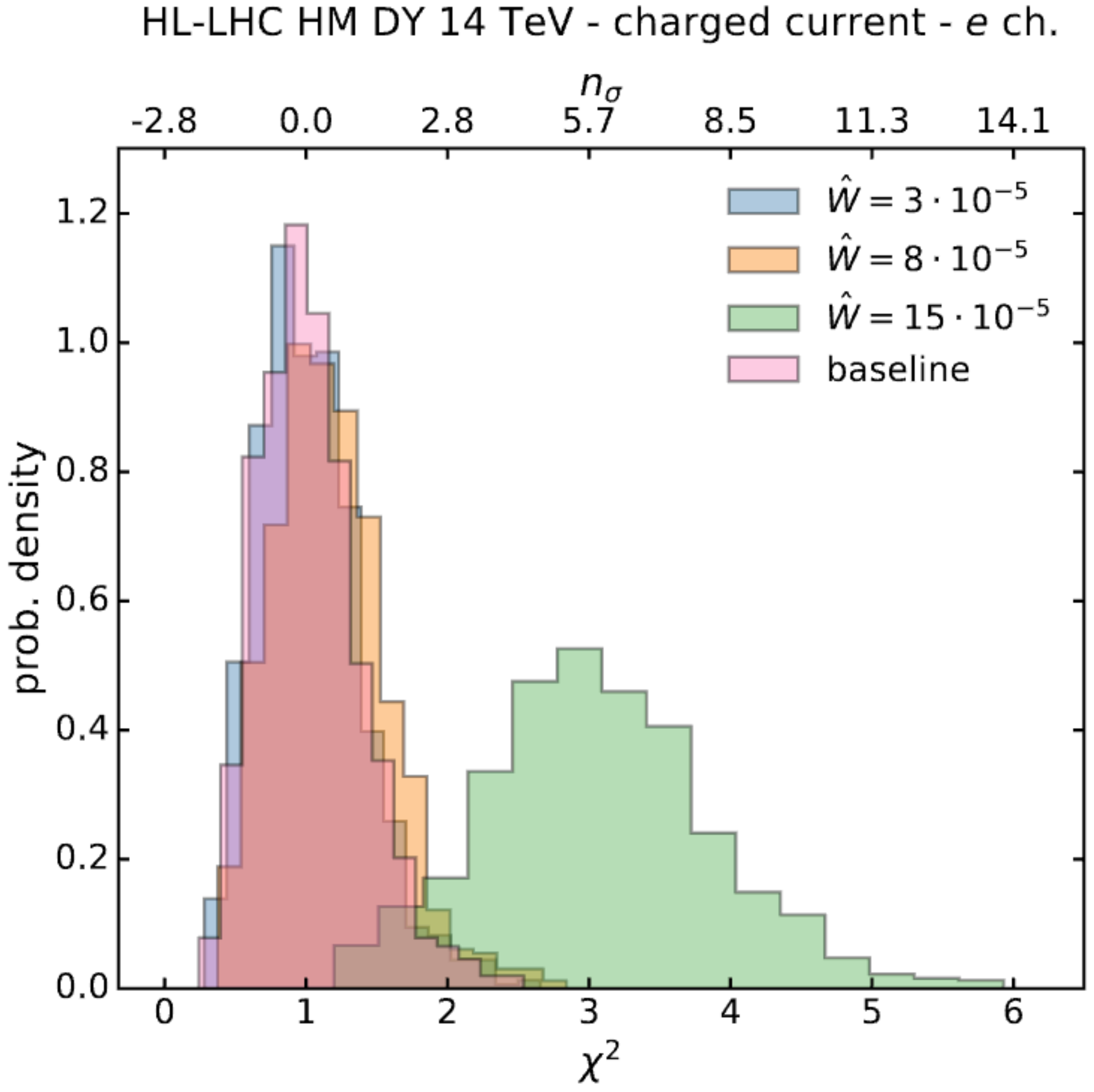




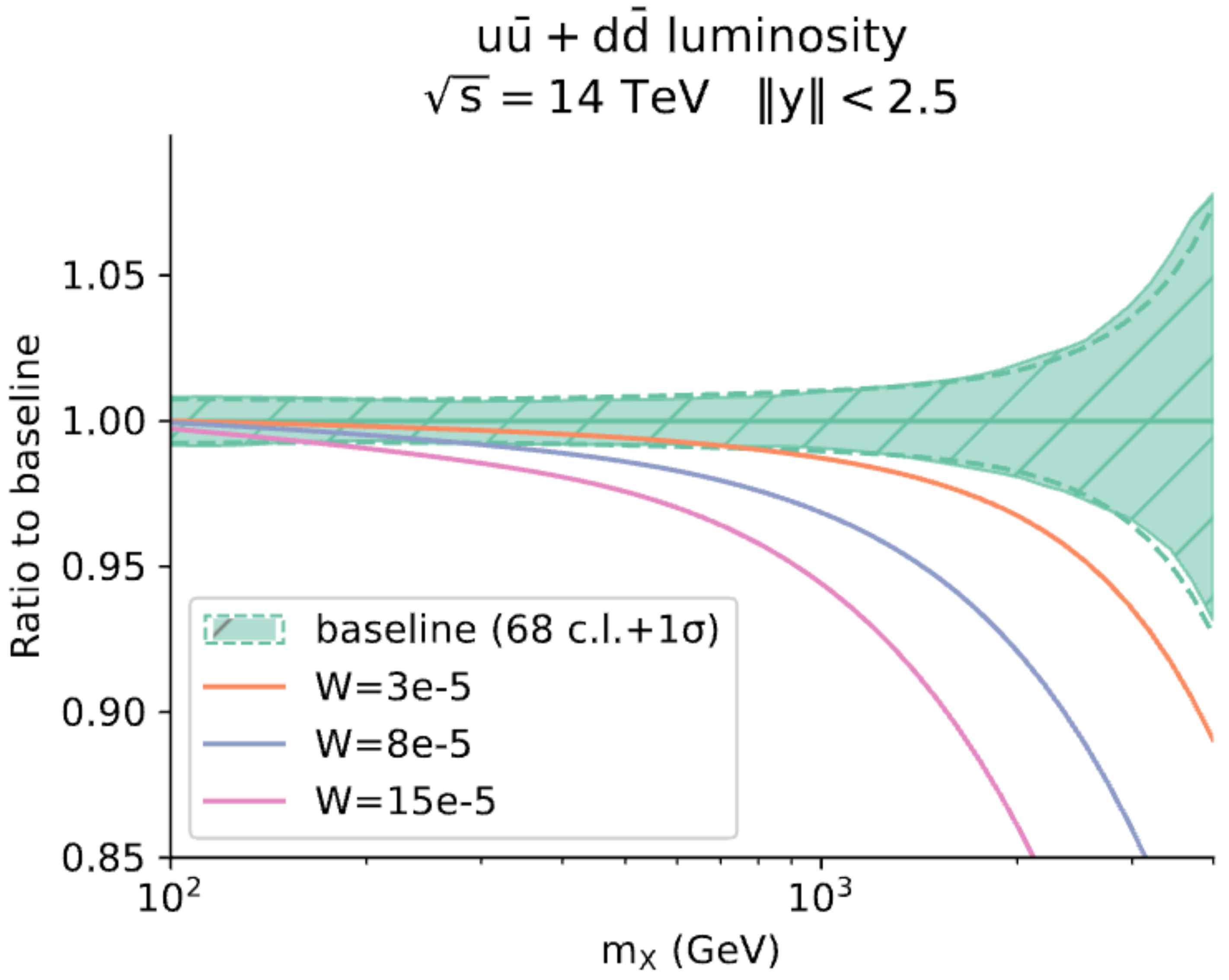


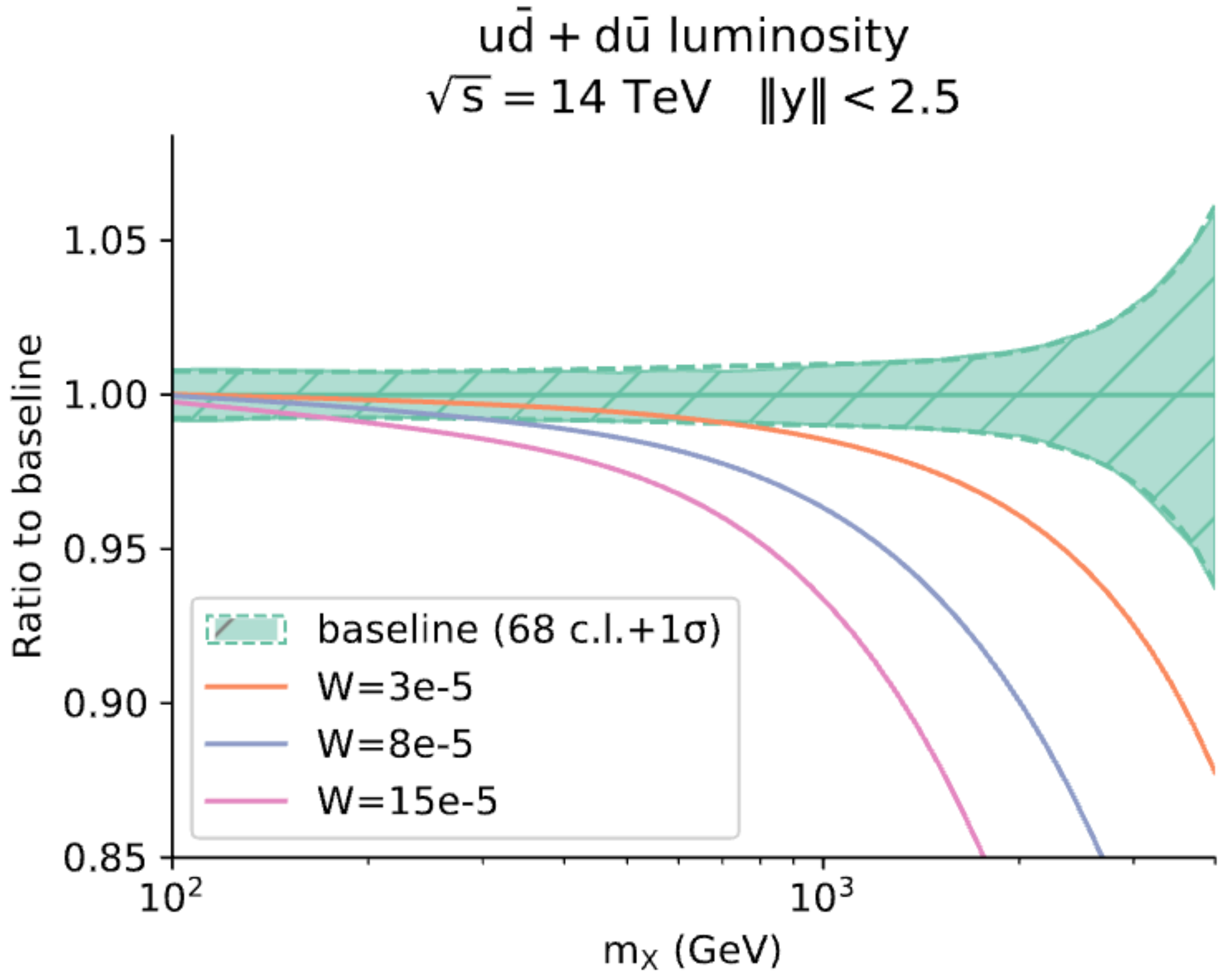
Consequence #1:
Would not see indirect effect of new physics as we would find bound for What in [-3e-5,3e-5] missing its true 8e-5 value

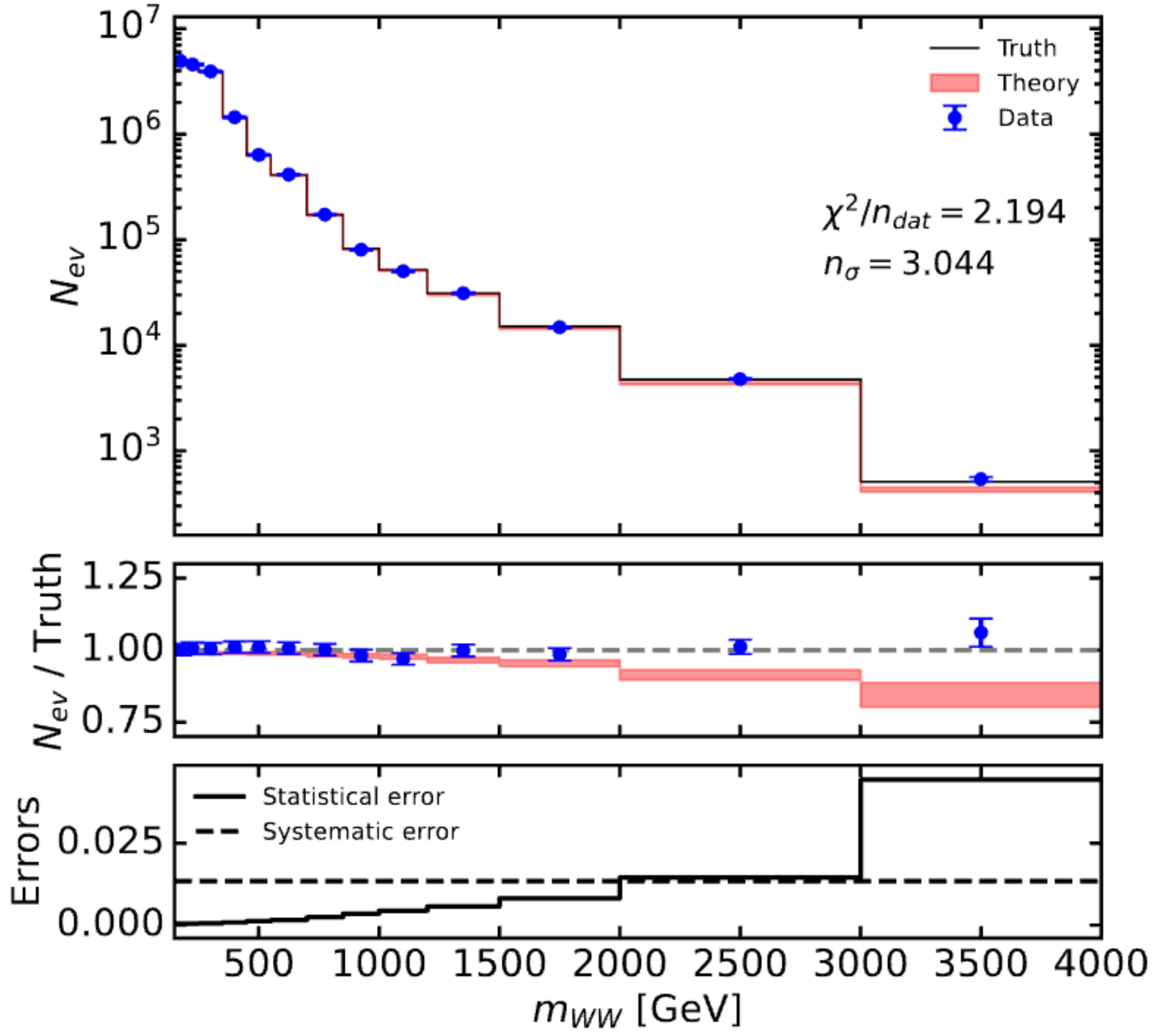




Consequence #2:
Would see New Physics effects where there are none (for example in WW)



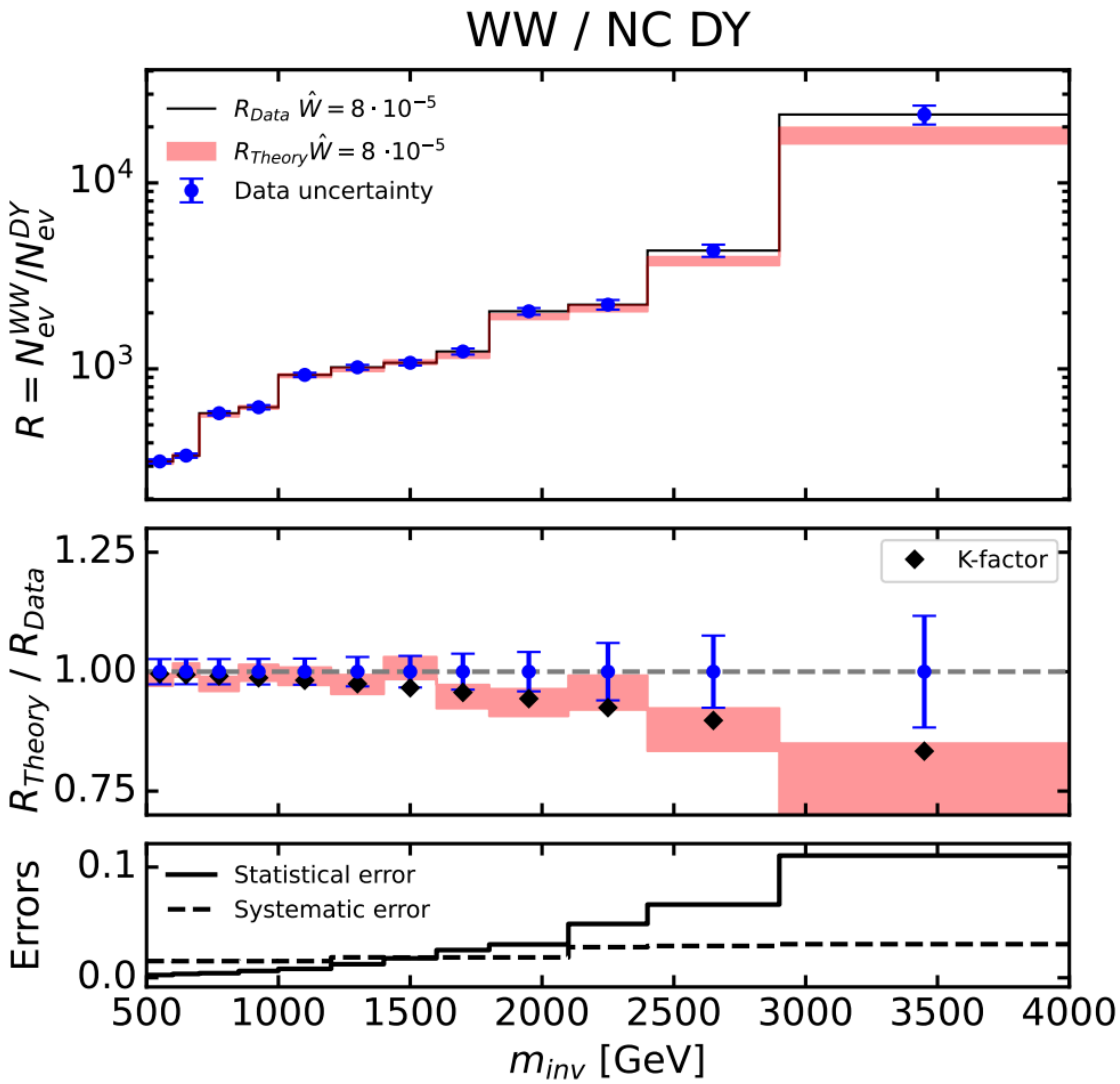




HOW TO AVOID NEW PHYSICS CONTAMINATION?

- PDF independent observable ratios would shed some light
- There are several soft indicators that there is something wrong with the PDFs but none conclusive
- Need more accurate low-energy/large-x constraining measurements to really disentangle such effects
- Strong motivation for SeaQuest, EIC precision data, FASERnu… for example

Alekhin et al, 2306.01918
Cruz-Martinez et al, 2309.09581



CONCLUSIONS AND OUTLOOK

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

- Simultaneous fits
  - ➡ Scan in SMEFT space helps assessing PDF and SMEFT interplay but limited as one includes more than a few operators A. Greljo et al,2201.07240 - see also approach by J. Gao et al, 2211.01094
  - ➡ General simuNET methodology for simultaneous fits linear SMEFT + PDFs for an arbitrary number of SMEFT operators has been developed S. Iranipour et al,2201.07240
  - ➡ Code will be soon public and EWPO+Higgs will be included
  - ➡ Effect of simultaneous fits on PDFs and SMEFT depends on the sector
  - ➡ Top: shift and larger uncertainty in PDFs, unchanged bounds Kassabov et al, 2303.06159
  - ➡ Drell-Yan: unshifted PDFs and large uncertainties, broader bounds S. Iranipour et al,2201.07240
  - ➡ Quadratic corrections: beyond MC sampling Kassabov et al, 2303.06159

- Can PDF absorb new physics?
  - ➡ Identified a UV scenario such that the high-mass HL-LHC invariant mass can absorb.
  - ➡ Important to disentangle large-x from high-energy / low-energy (SeaQuest, JLAB, EIC, FASERnu…)

THANK YOU FOR YOUR ATTENTION!

EXTRA MATERIAL

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

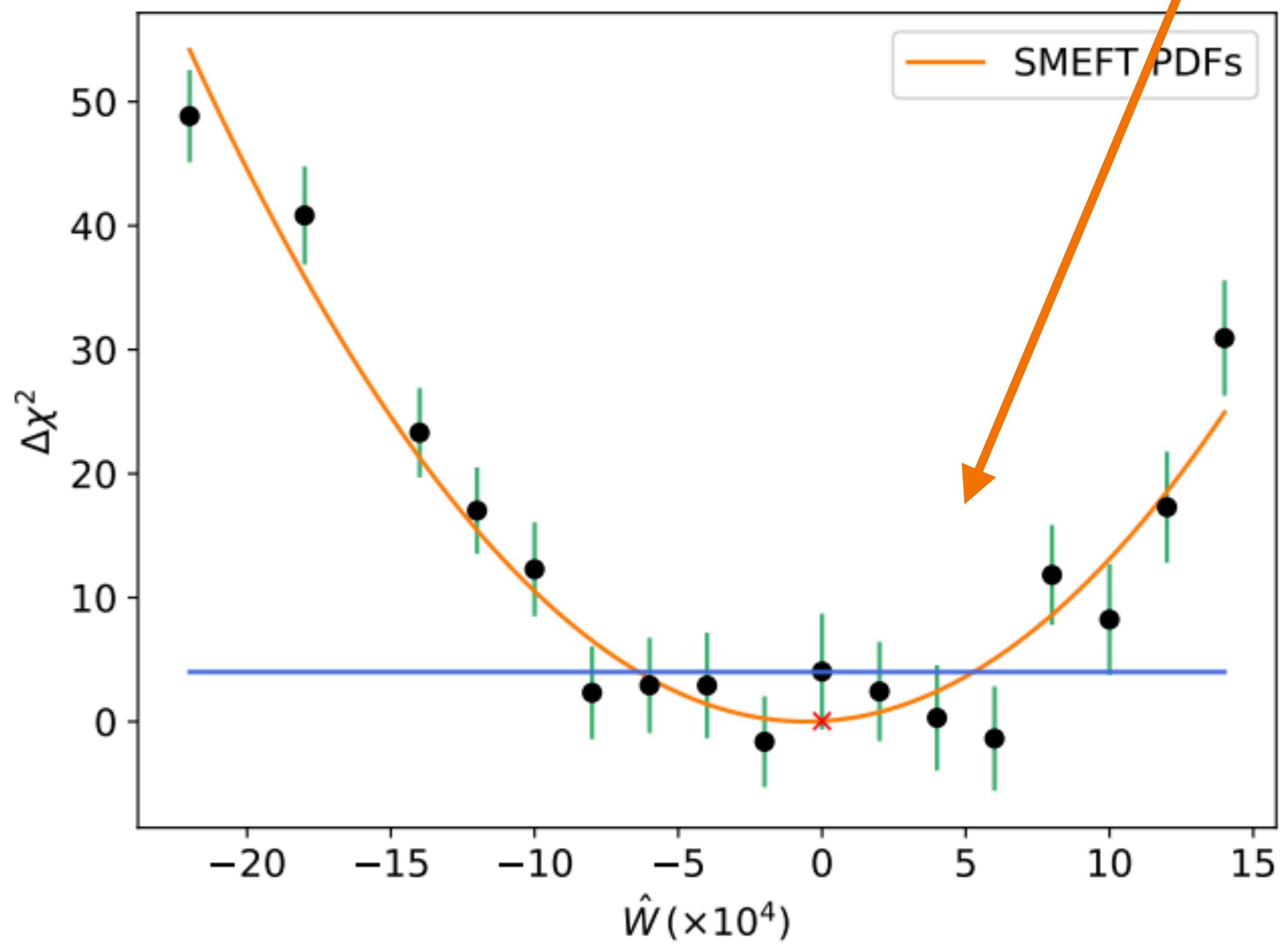
Operator	Coefficient	Definition
$\mathcal{O}^{(1)}_{\varphi Q}$	$-\ (c^{(1)}_{\varphi Q})$	$i(\varphi^\dagger \overleftrightarrow{D}_\mu \varphi)(\bar{Q}\gamma^\mu Q)$
$\mathcal{O}^{(3)}_{\varphi Q}$	$c^{(3)}_{\varphi Q}$	$i(\varphi^\dagger \overleftrightarrow{D}_\mu \tau_I \varphi)(\bar{Q}\gamma^\mu \tau^I Q)$
$\mathcal{O}_{\varphi t}$	$c_{\varphi t}$	$i(\varphi^\dagger \overleftrightarrow{D}_\mu \varphi)(\bar{t}\gamma^\mu t)$
$\mathcal{O}_{tW}$	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}	$-\sin\theta_W c_{tB} + \cos\theta_W c_{tW}$

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

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

SM PDFs

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

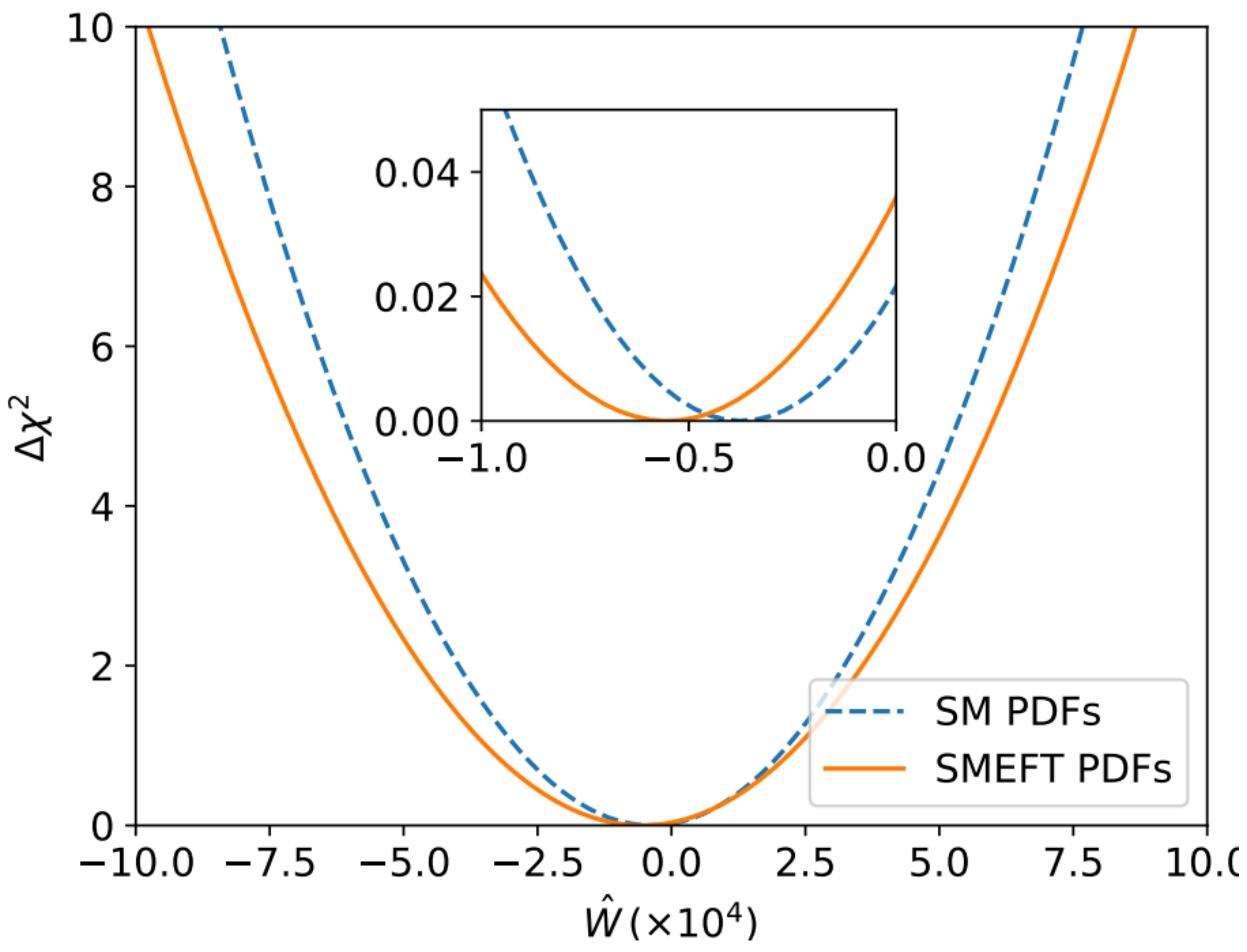
$$T = f_{1,\rm SM} \otimes f_{2,\rm SM} \otimes \hat{\sigma}_{\rm BSM}$$

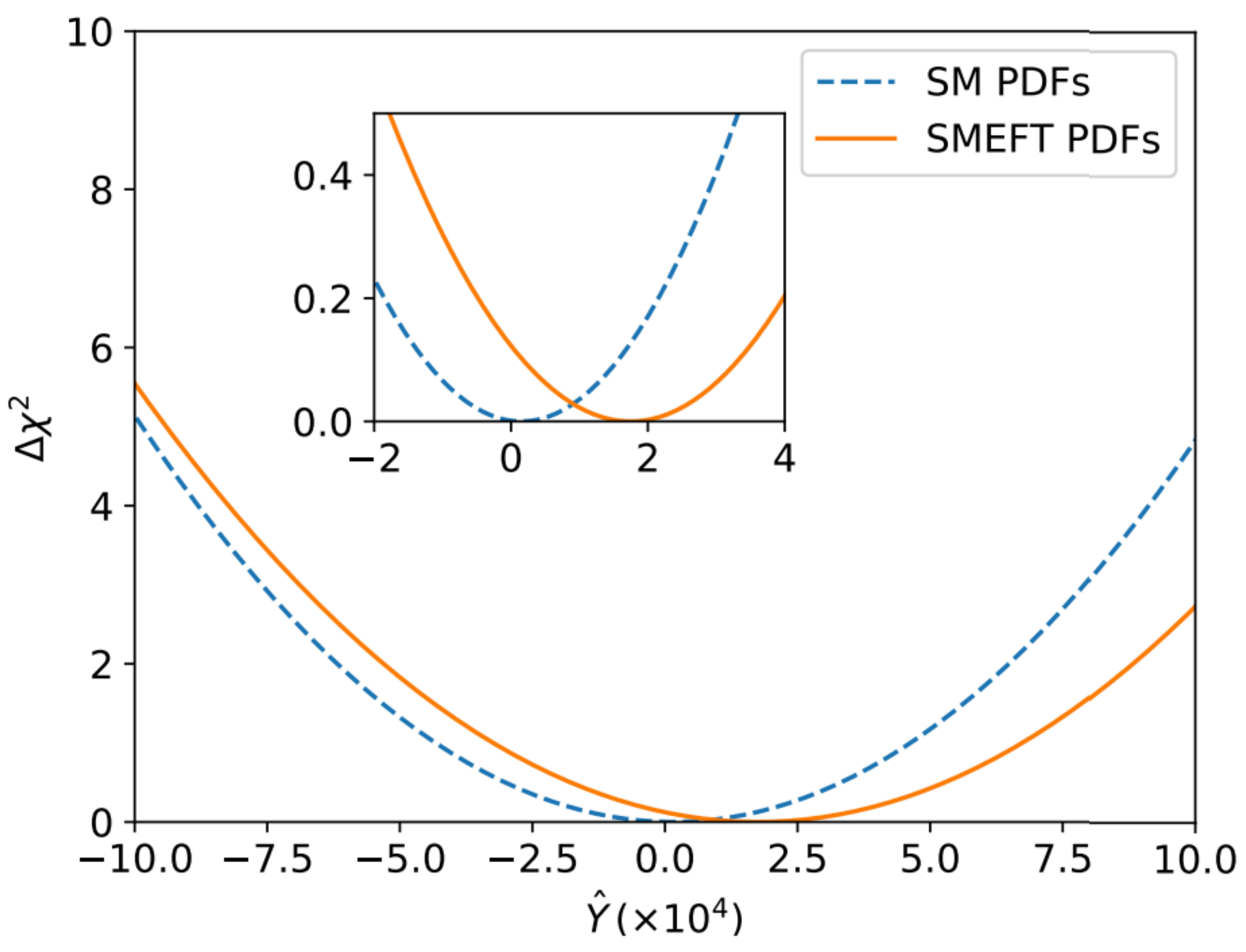
SMEFT PDFs / Simultaneous fit

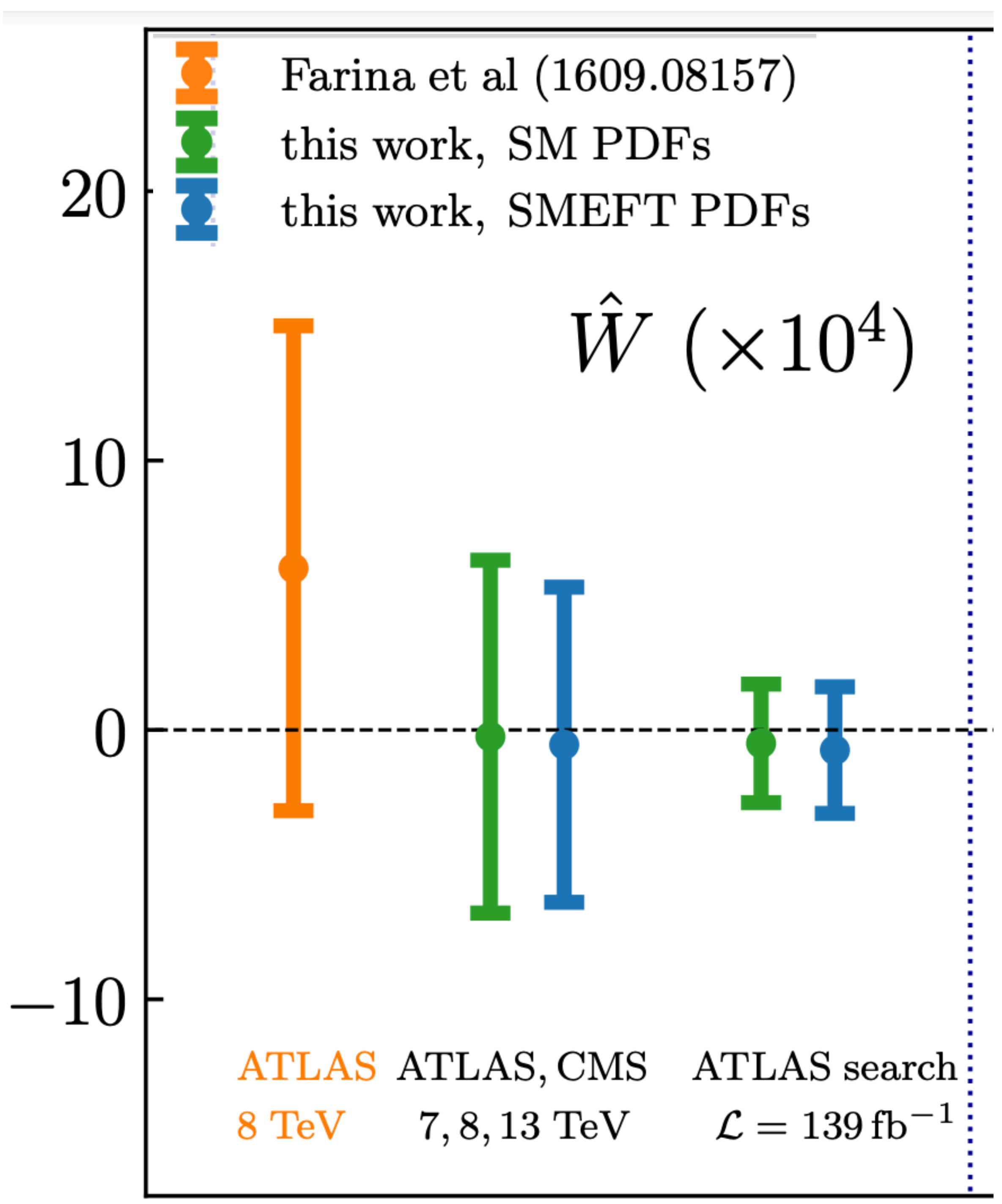
1. Take data, make theoretical predictions accounting for operator in partonic cross section and PDFs.
2. Compute chi2 as a function of WCs (Wilson Coefficients)
3. Minimise chi2 and find best-fit and C.L.s of WCs
4. Extract bounds

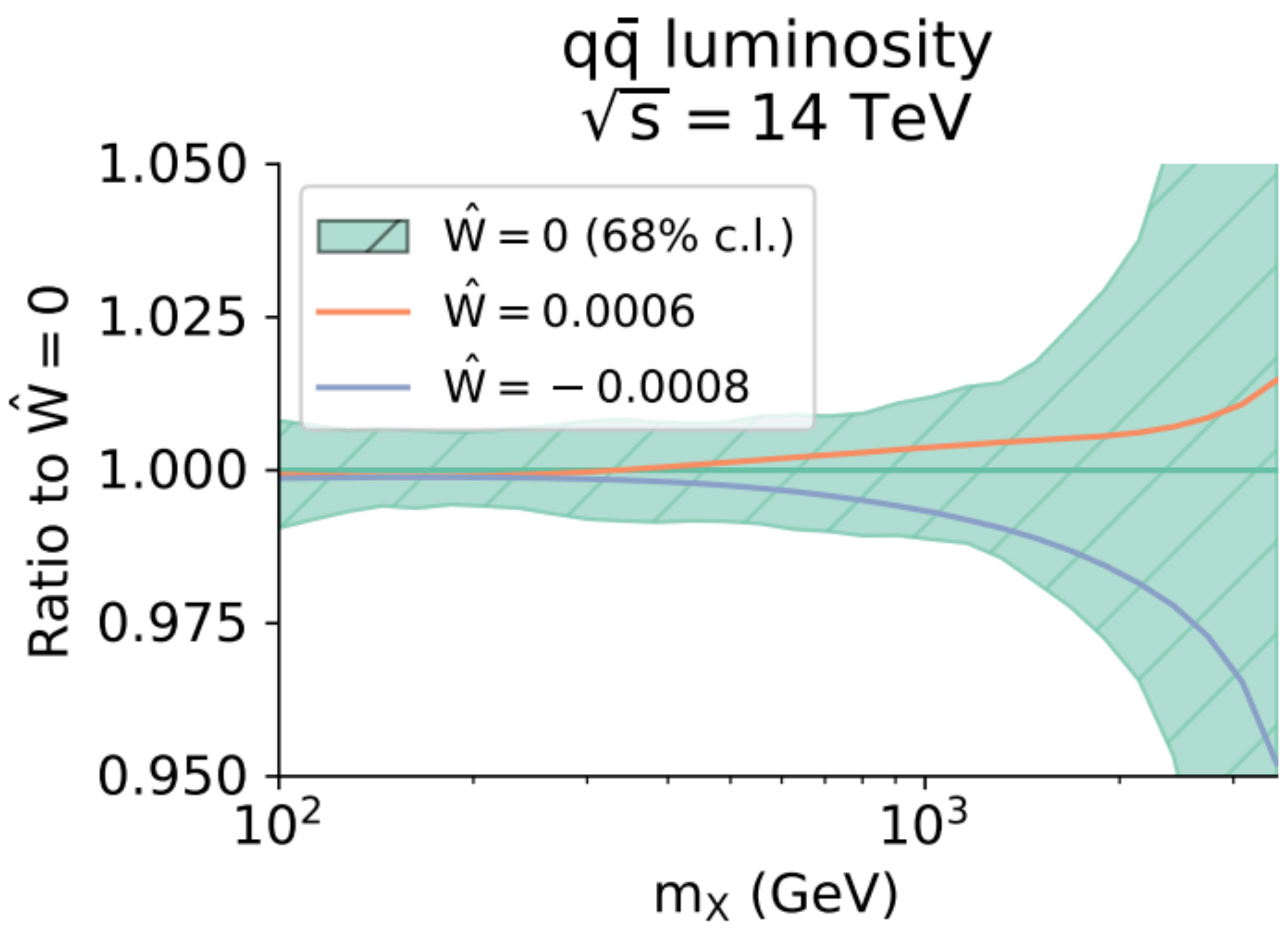
$$T = f_{1,\rm BSM} \otimes f_{2,\rm BSM} \otimes \hat{\sigma}_{\rm BSM}$$

INTERPLAY @ RUN I AND RUN II









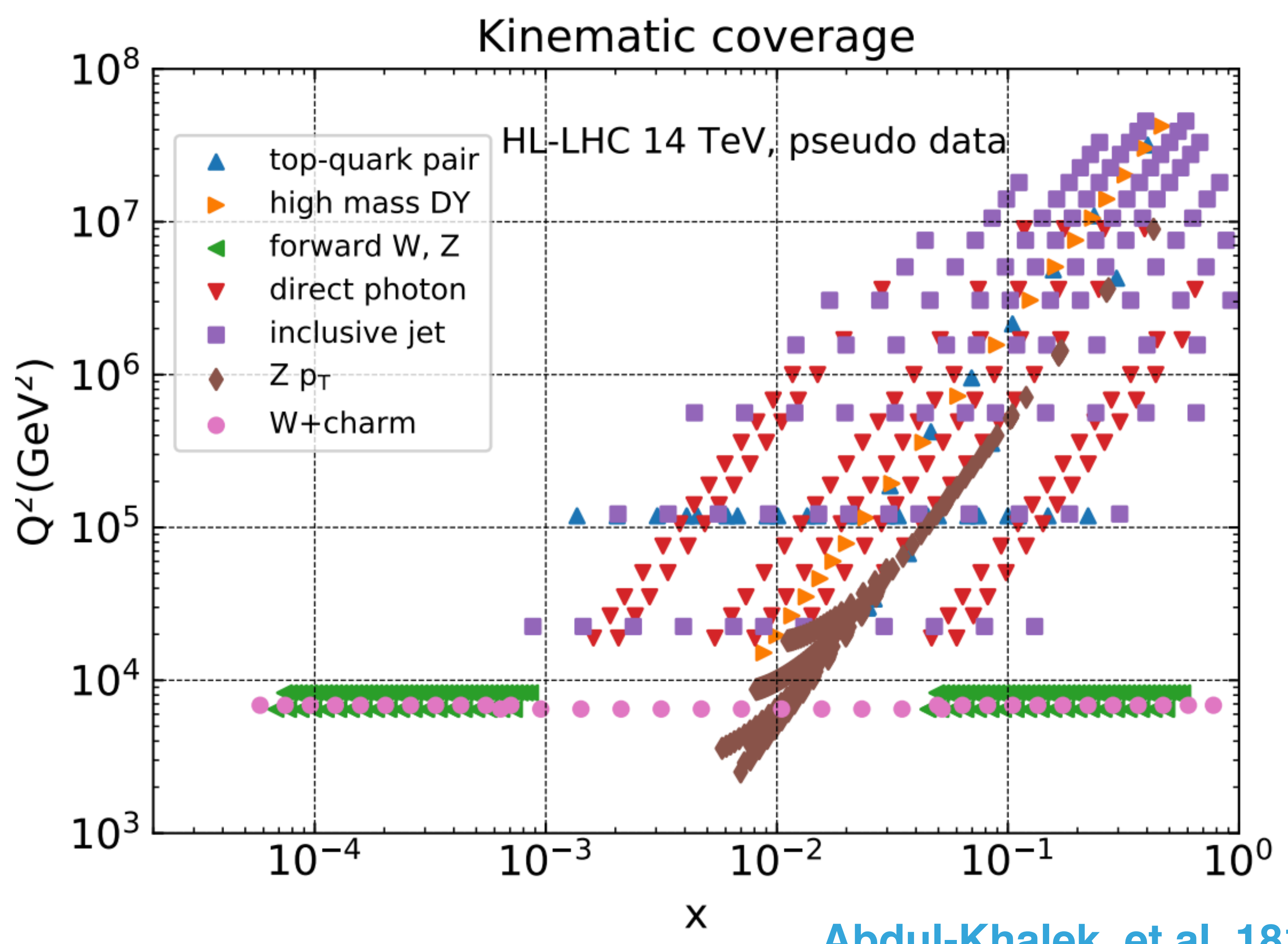
- With current data, PDFs are moderately affected by inclusion of non-zero W and Y coefficients in the fit, mostly quark-antiquark luminosity within uncertainties
- Broadening of individual bounds on W and Y once SMEFT PDFs are used (i.e. PDFs that have been fitted with consistent values of W and Y) 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

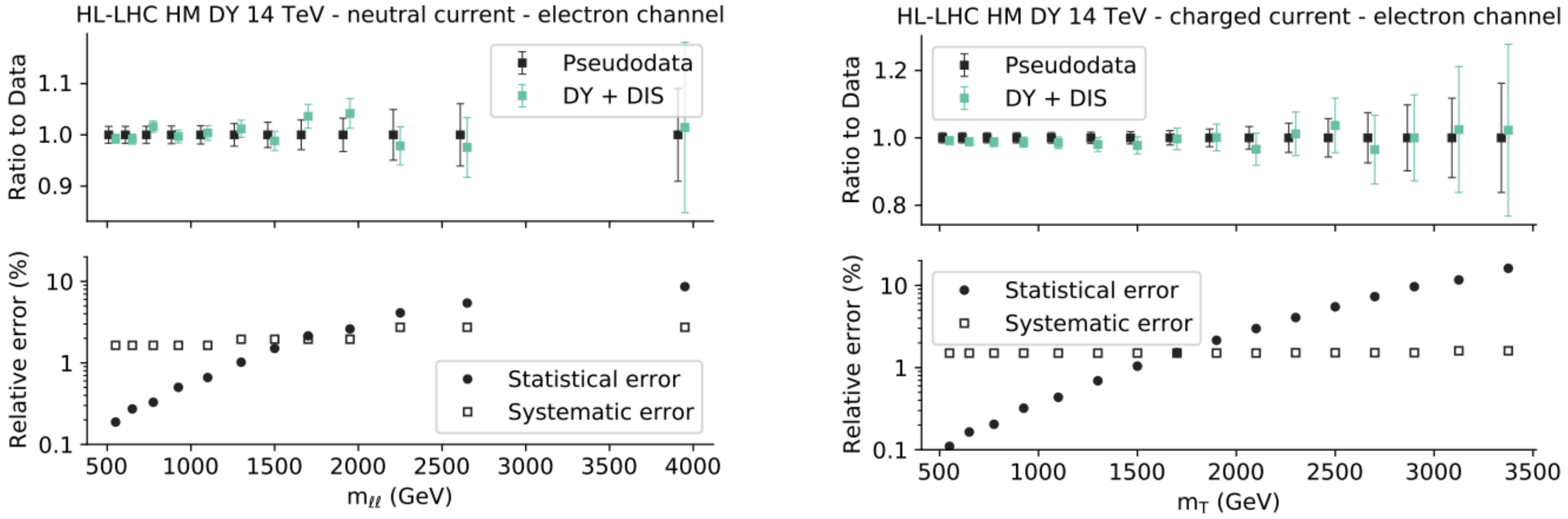
- Add HL-LHC projections for both NC and CC in PDF fit

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

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



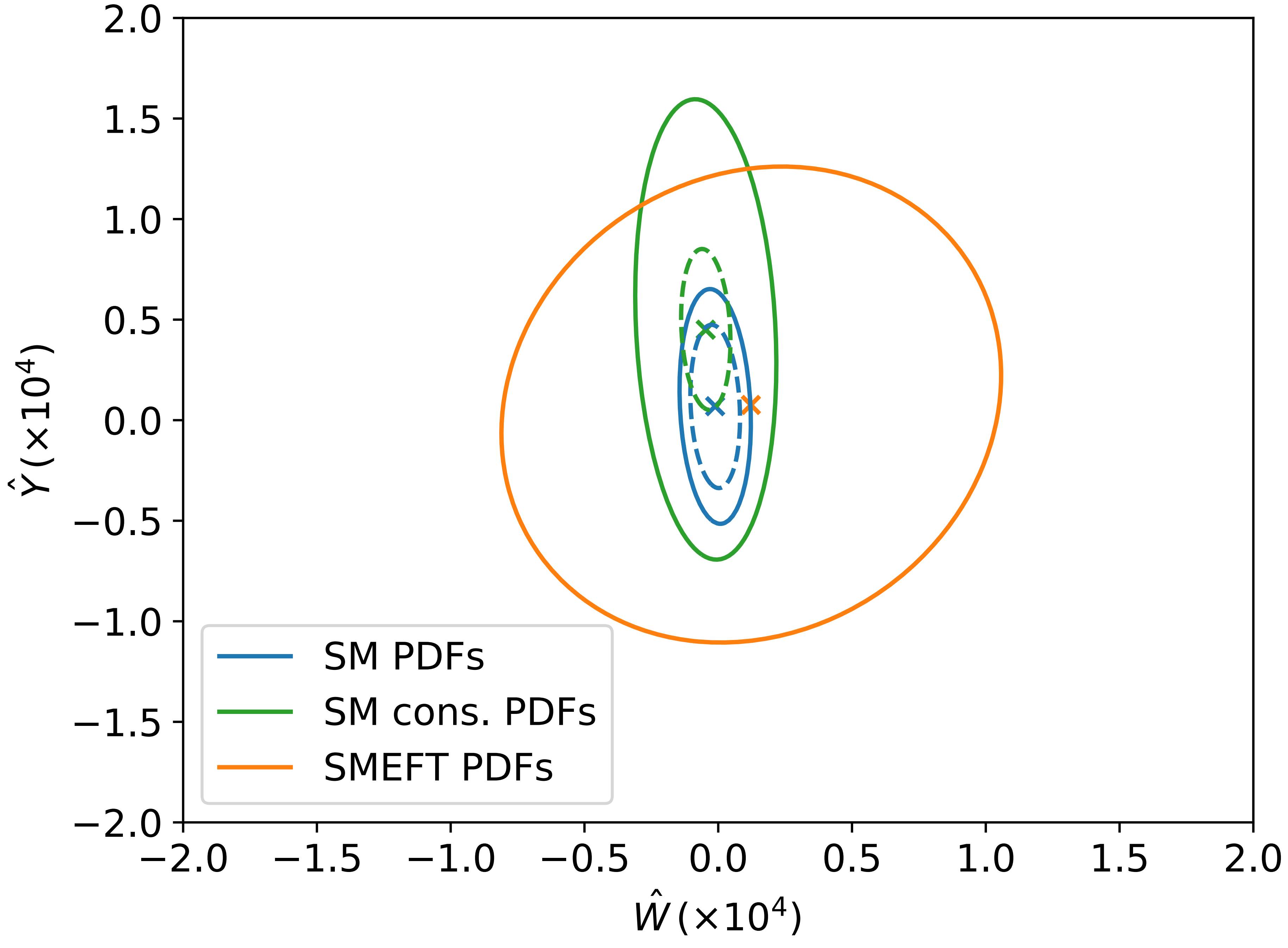
Abdul-Khalek et al, 1810.03639

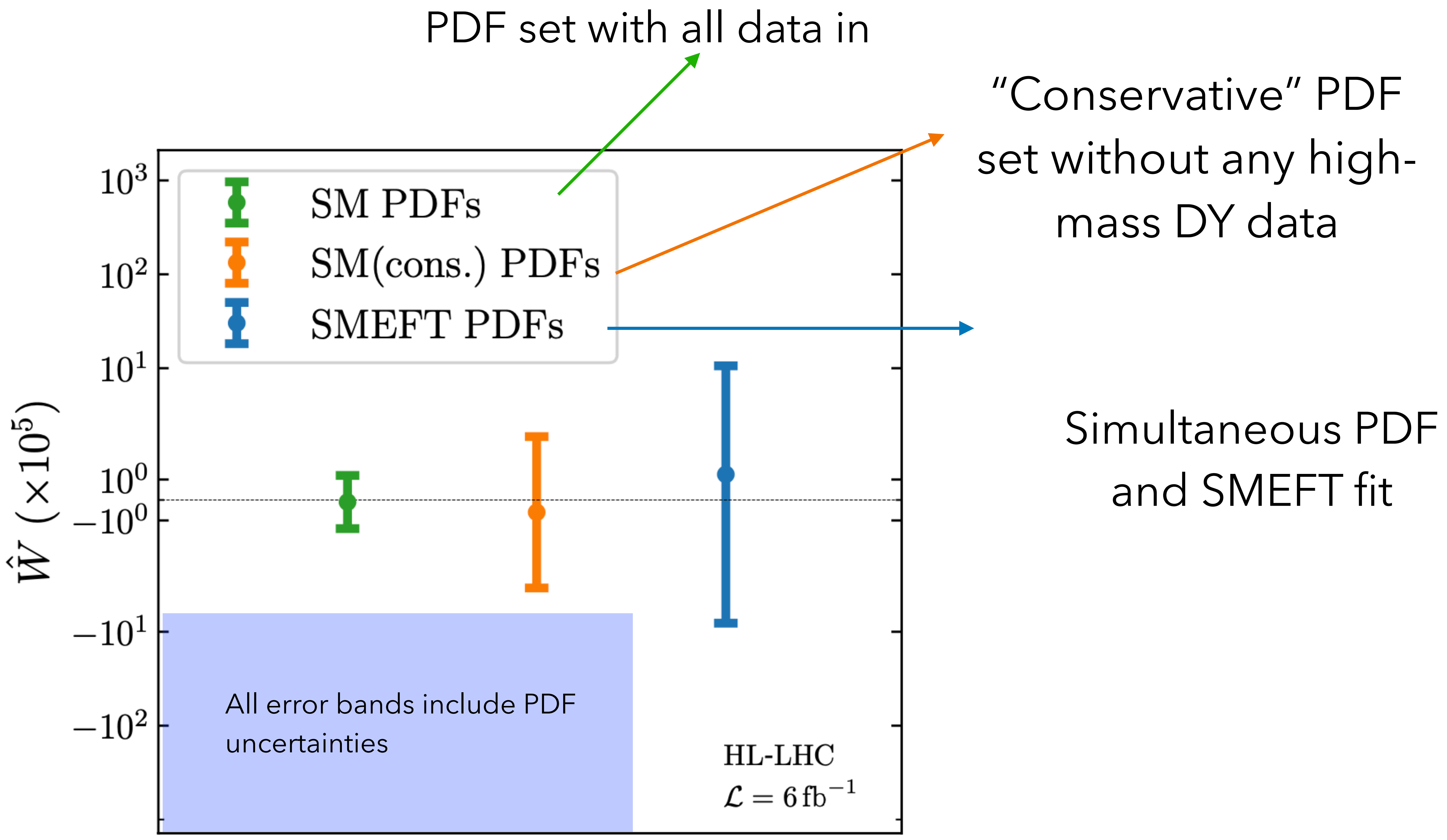


+ muon channel

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!



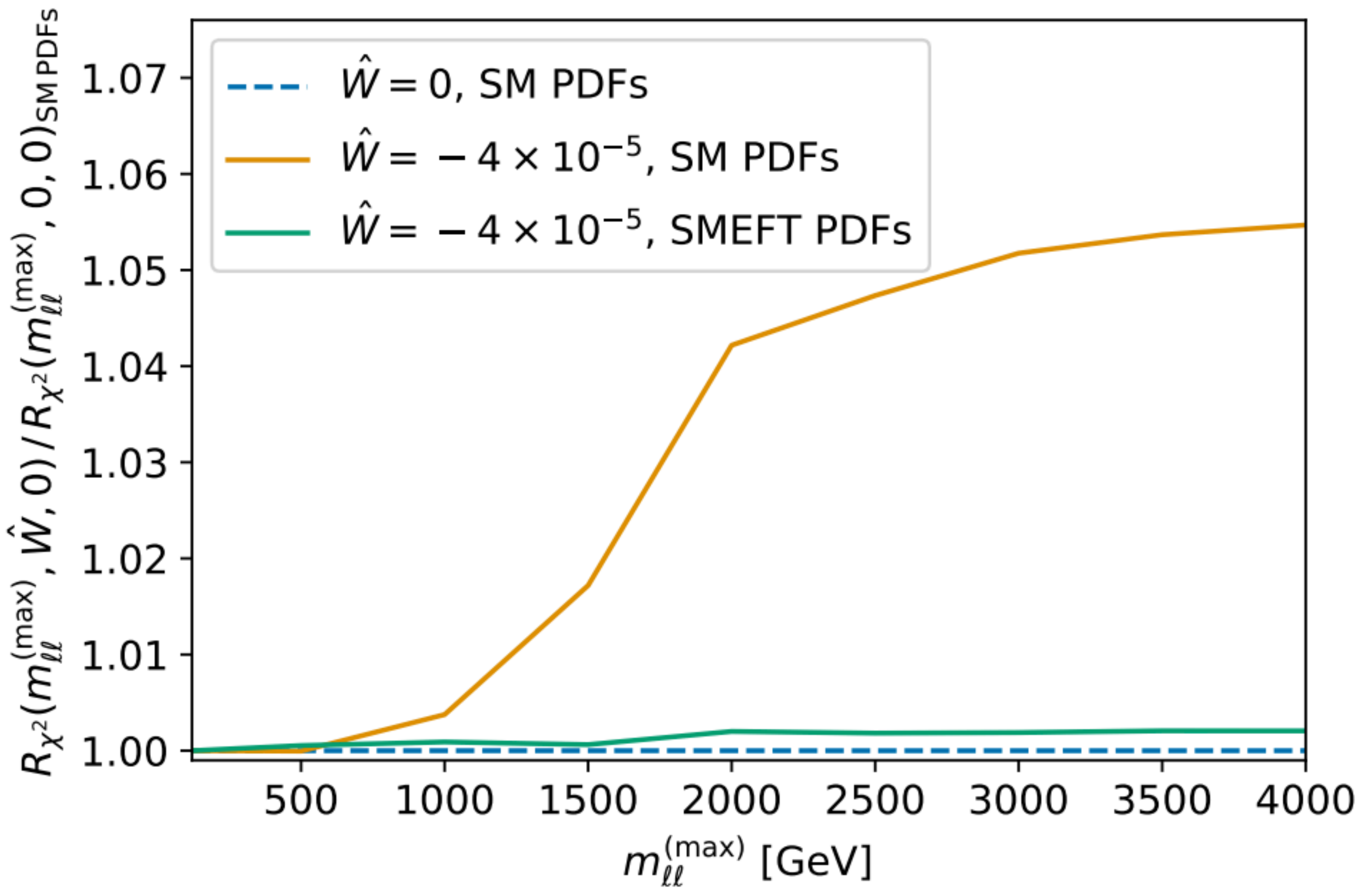


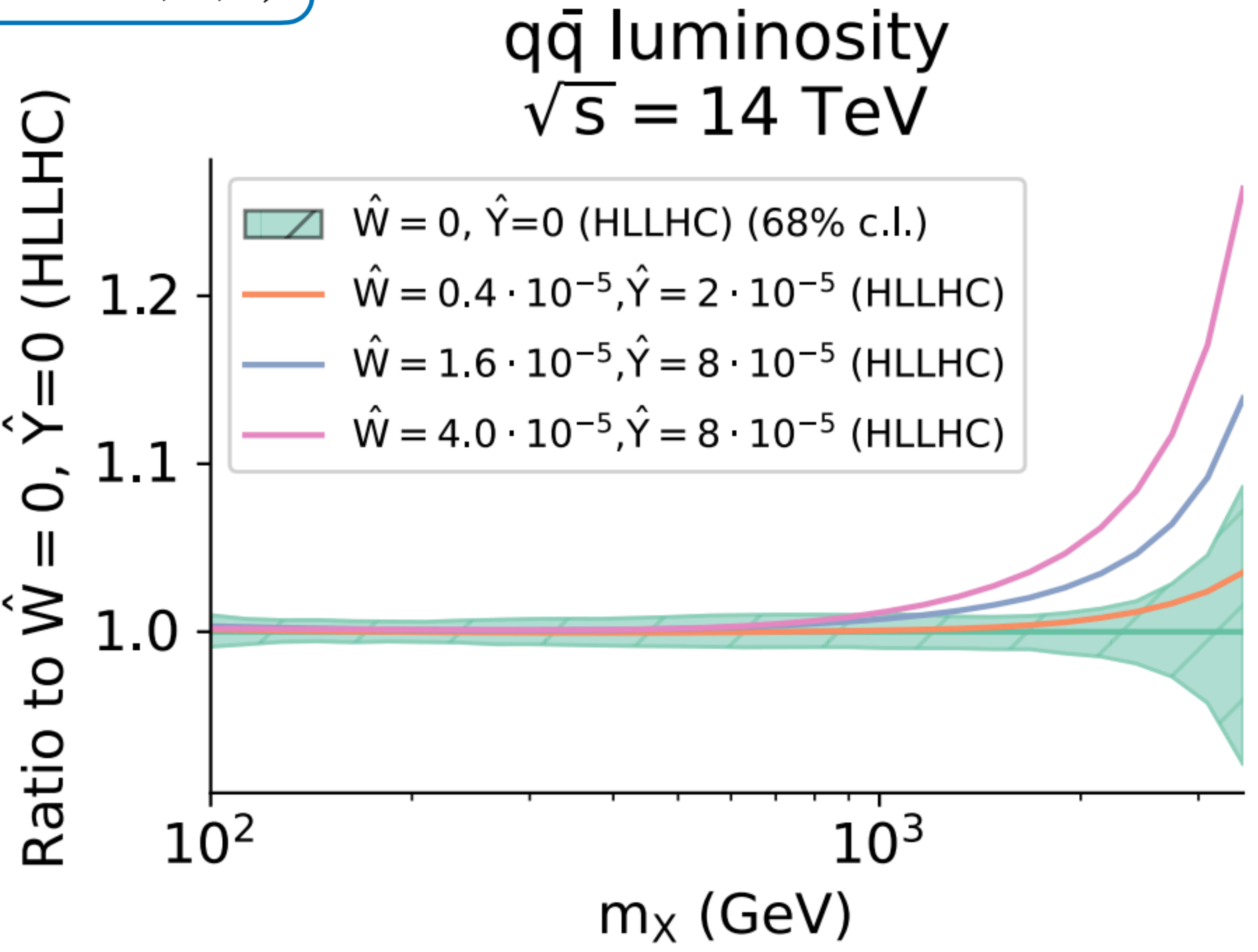
Greljo et al, 2104.02723

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!

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

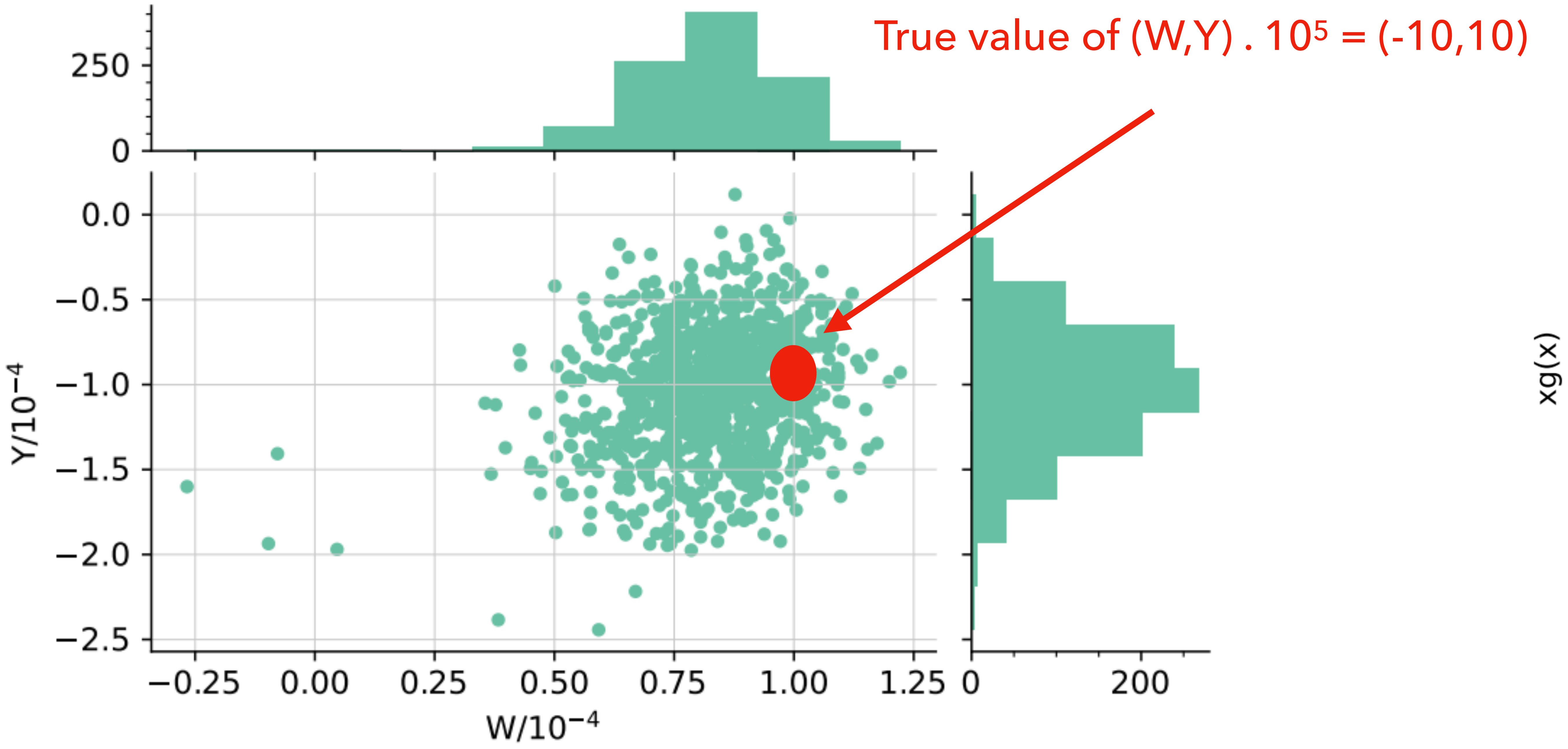


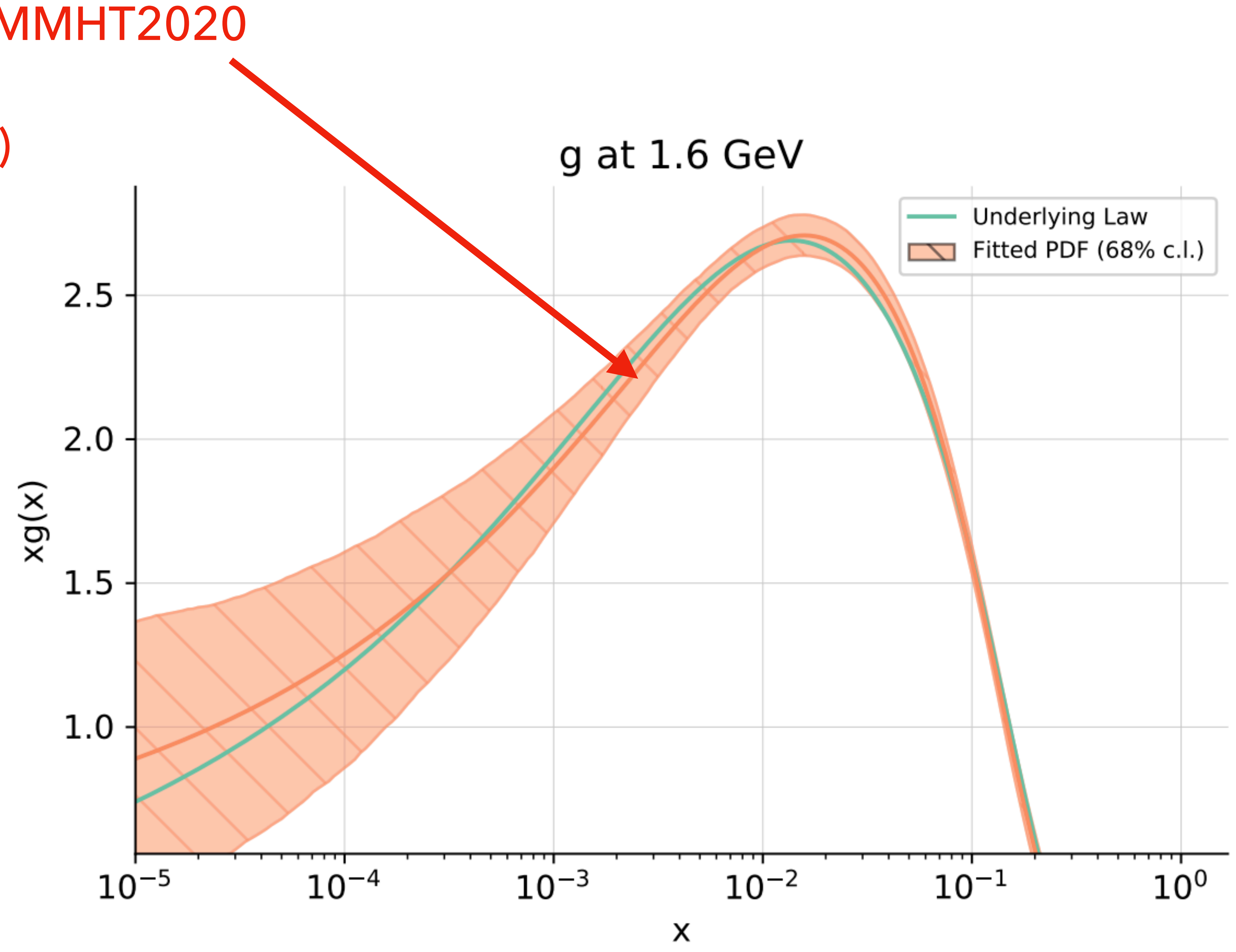


Greljo et al, 2104.02723

RESULTS: DRELL-YAN DATA @HL-LHC

S. Iranipour, MU - arXiv: 2201.07240





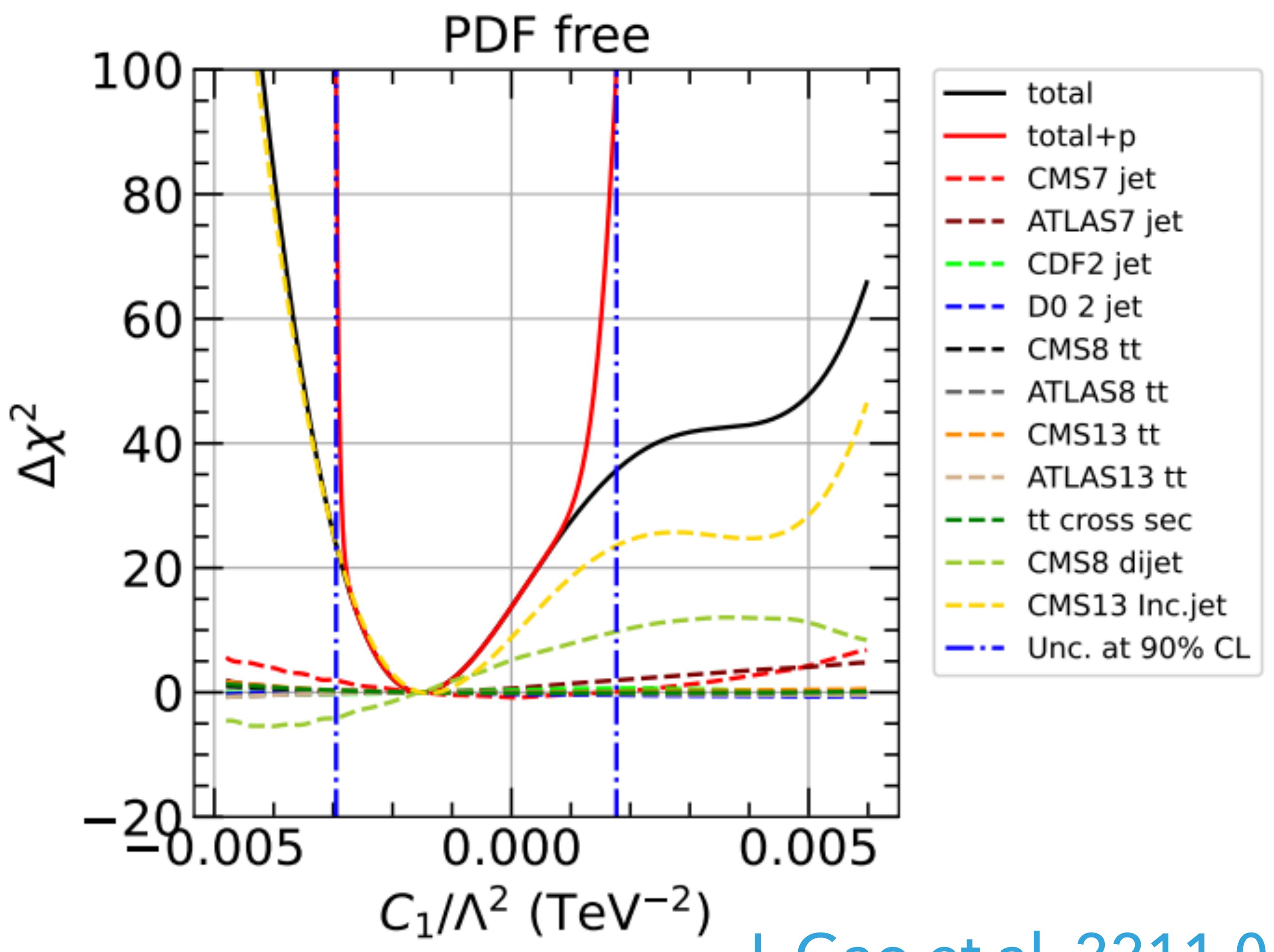
- ✓ Simultaneous analysis of PDFs and W&Y SMEFT coefficient of DIS + DY (including HL-LHC projections) using simuNET method shows that at HL-LHC the effect of interplay becomes important as WCs bounds broaden and PDFs change significantly once SMEFT effects allowed in theory predictions entering PDF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

THE TOP SECTOR

- We extend previous analysis to 20 operators in the top sector, including all available unfolded top data that constrain PDFs and/or SMEFT operators: tt~ inclusive and differential/double-differential cross sections at 5, 7, 8 and 13 TeV, tt~ asymmetries at 8 an 13 TeV, W-helicity fractions at 8 and 13 TeV, tt~V associated production at 8 TeV and 13 TeV, tt~tt~ and tb~tb~ production total cross sections, single top t-channel and s-channel production and associated single top and vector boson production.

- Improved simuNET algorithm allows to include both PDF dependent and independent measurements.

- Strategy (Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU) :
  - ➡ Add all available top-sector data and assess impact on PDFs
  - ➡ Add all available top-sector data and fit SMEFT coefficients
  - ➡ Add all available top-sector data in simultaneous fit of PDFs and SMEFT to assess interplay and correlations
  - ➡ Verify results via closure tests



J. Gao et al, 2211.01094

- Note that this is different from analysis by J. Gao et al both in terms of approach (individual vs marginalised, scan versus fit) and in number of operators (4 versus 20) and datasets & interesting to compare.

PRIOR PROBABILITY IN PDF FITS

- ✓ PDF fitting example of inverse problem: aim to find a posterior probability of ***f*** given the data ***D***.

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

- ✓ Parametrization of PDFs: finite-dimensional problem.

$$f(x) \approx \tilde{f}(x,\theta) \in \mathcal{F}$$

- ✓ The posterior probability for the parametrization depends on both the figure of merit that maximises the data likelihood given the parameters and on prior probability ***H.***

$$p(\theta|D,\mathcal{H}) \propto p(D|\theta,\mathcal{H})\,p(\theta|\mathcal{H}) = \exp(-\mathcal{L}(\theta,D))\,p(\theta|\mathcal{H})$$

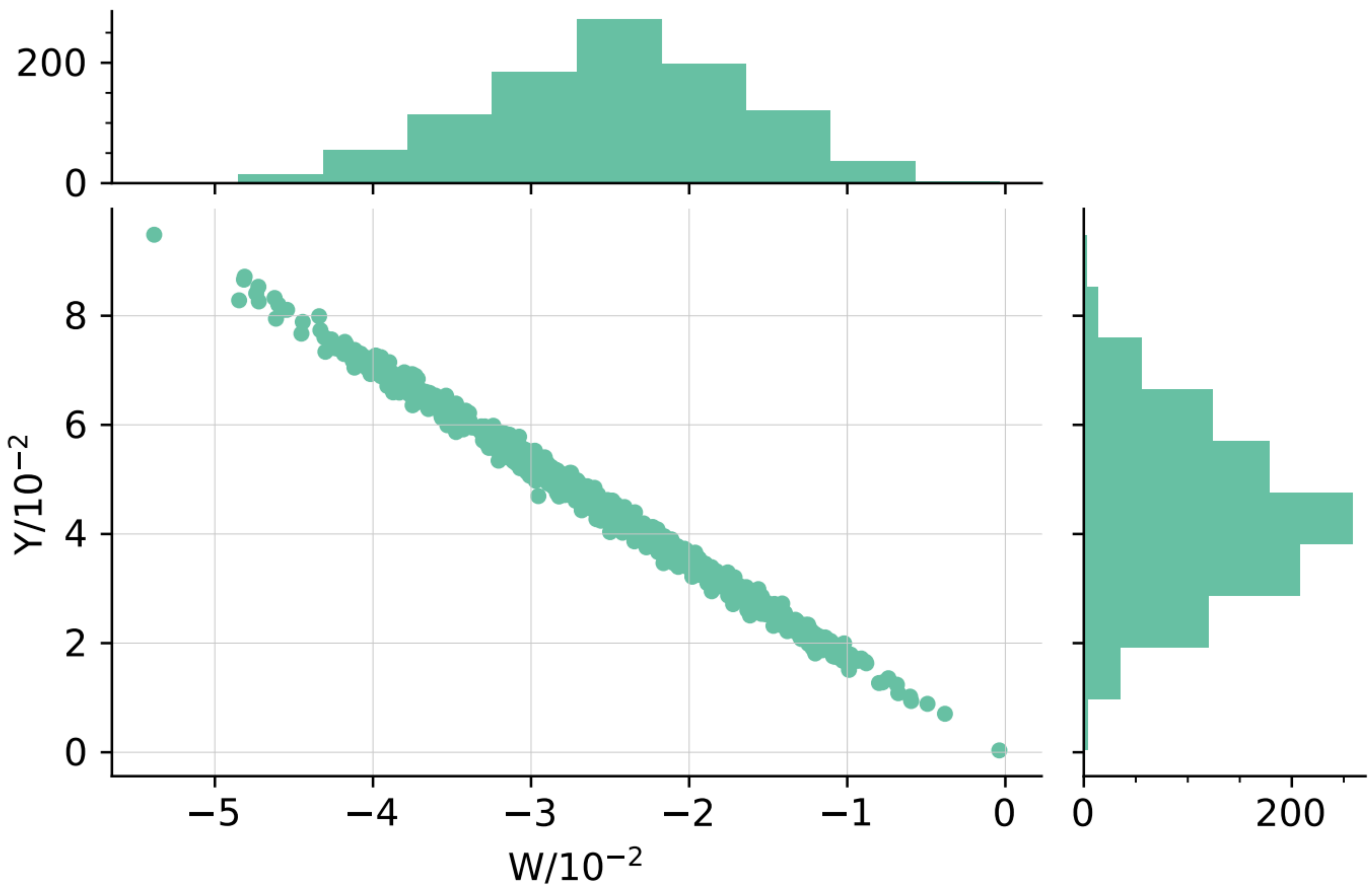
Prior: functional form, integrability, positivity, sum rules, behaviour at small-x and large-x…

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

$${\rm cov}_{ij} \equiv {\rm cov}^{t_0}_{ij} = \left(\sum_{l=1}^{N-N_{\rm norm}} \sigma_{i,l}\sigma_{j,l}\right) T_i T_j + \left(\sum_{m=1}^{N_{\rm norm}} \sigma_{i,m}\sigma_{j,m}\right) T_i^{(0)} T_j^{(0)}$$

Ball et al, arXiv:0912.2276

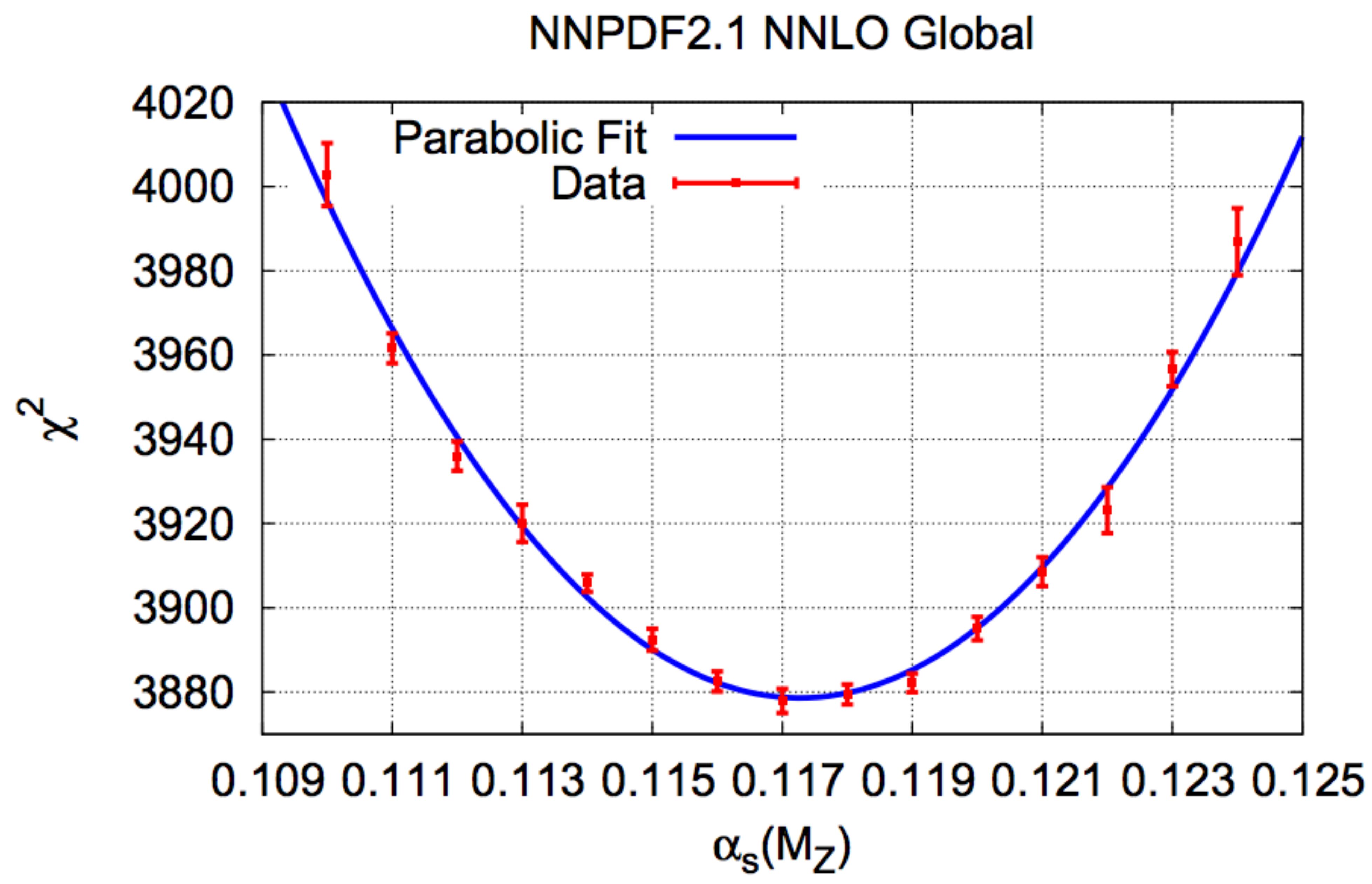
RESULTS: DRELL-YAN DATA @RUN1 AND RUN2



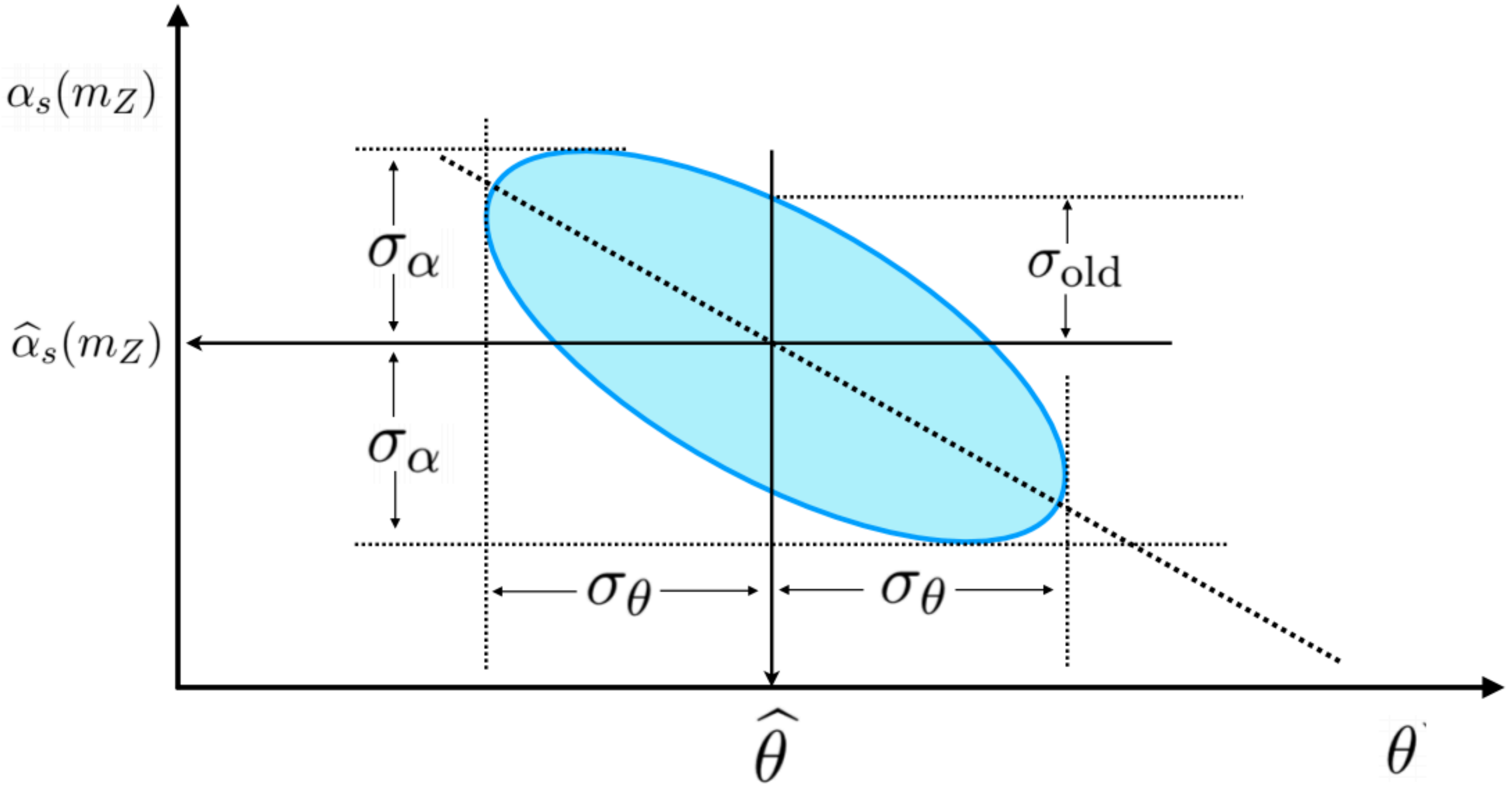
- ✓ Simultaneous analysis confirms results of previous study based on scan on benchmark points in the SMEFT space: with current data effect is not-negligible but small compared to PDF uncertainties
- ✓ Methodology able to find flat direction in W-Y parameter space
- ✓ To eliminate it, need Drell-Yan charged current data

PDFs AND α_s

- ➡ PDFs and α_s strongly correlated (PDF evolution with the scale and hard cross sections)
- ➡ Cleanest determinations of α_s from processes that do not require knowledge of the PDFs
- ➡ A determination of α_s jointly with the PDFs has advantage that it is driven by the combination of many experimental measurements from several different processes.



Ball et al, 1110.2483

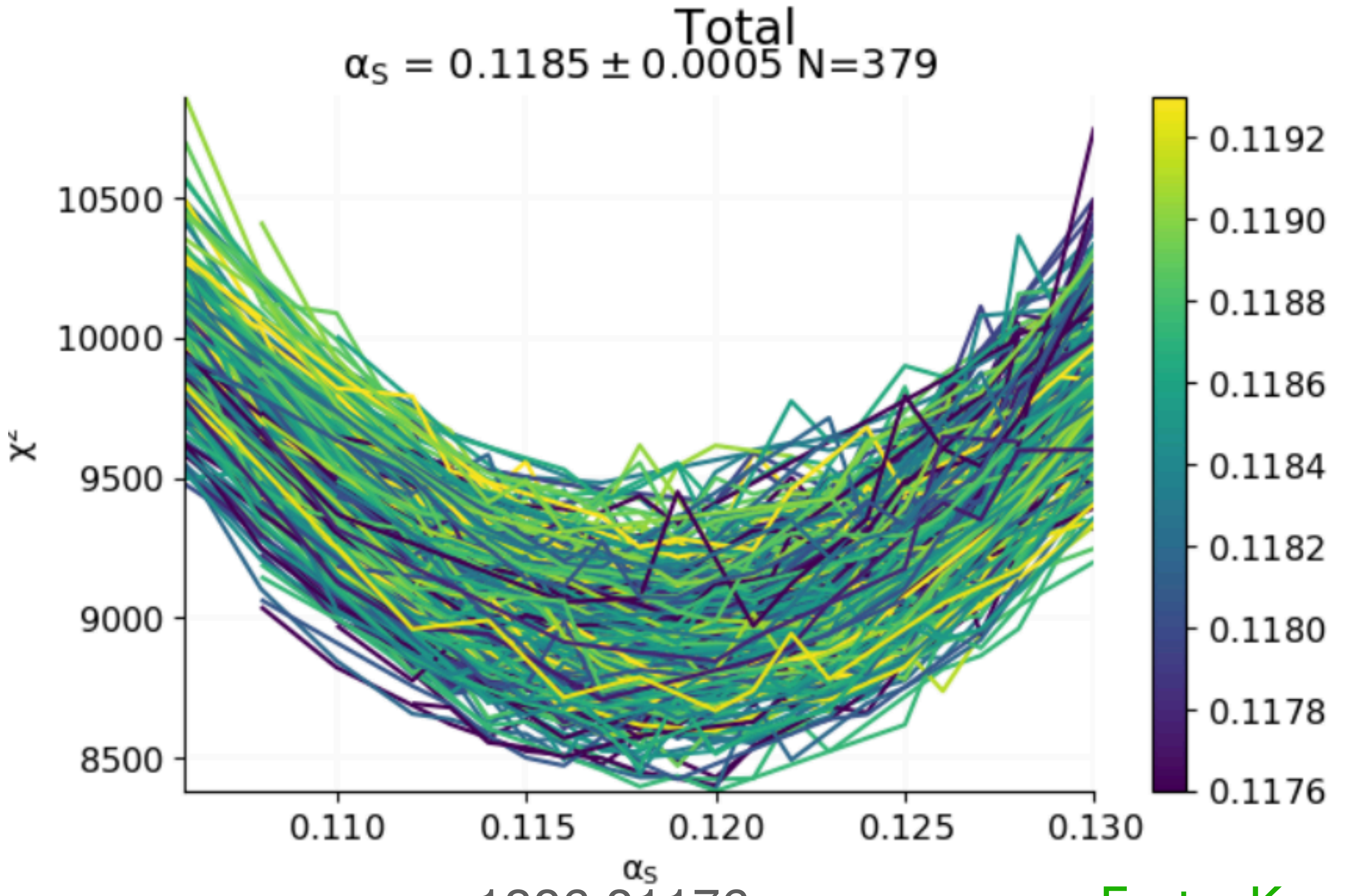


- ➡ Early determinations involve a scan over α_s and ignored PDF and α_s correlation in the fit
- ➡ Recent simultaneous determination of PDF and α_s using correlated replica method
- ➡ Many determination of α_s from analyses of specific LHC processes have been published recently (from tt~, Z and W production, jets)
- ➡ How reliable are such partial determination of α_s ?

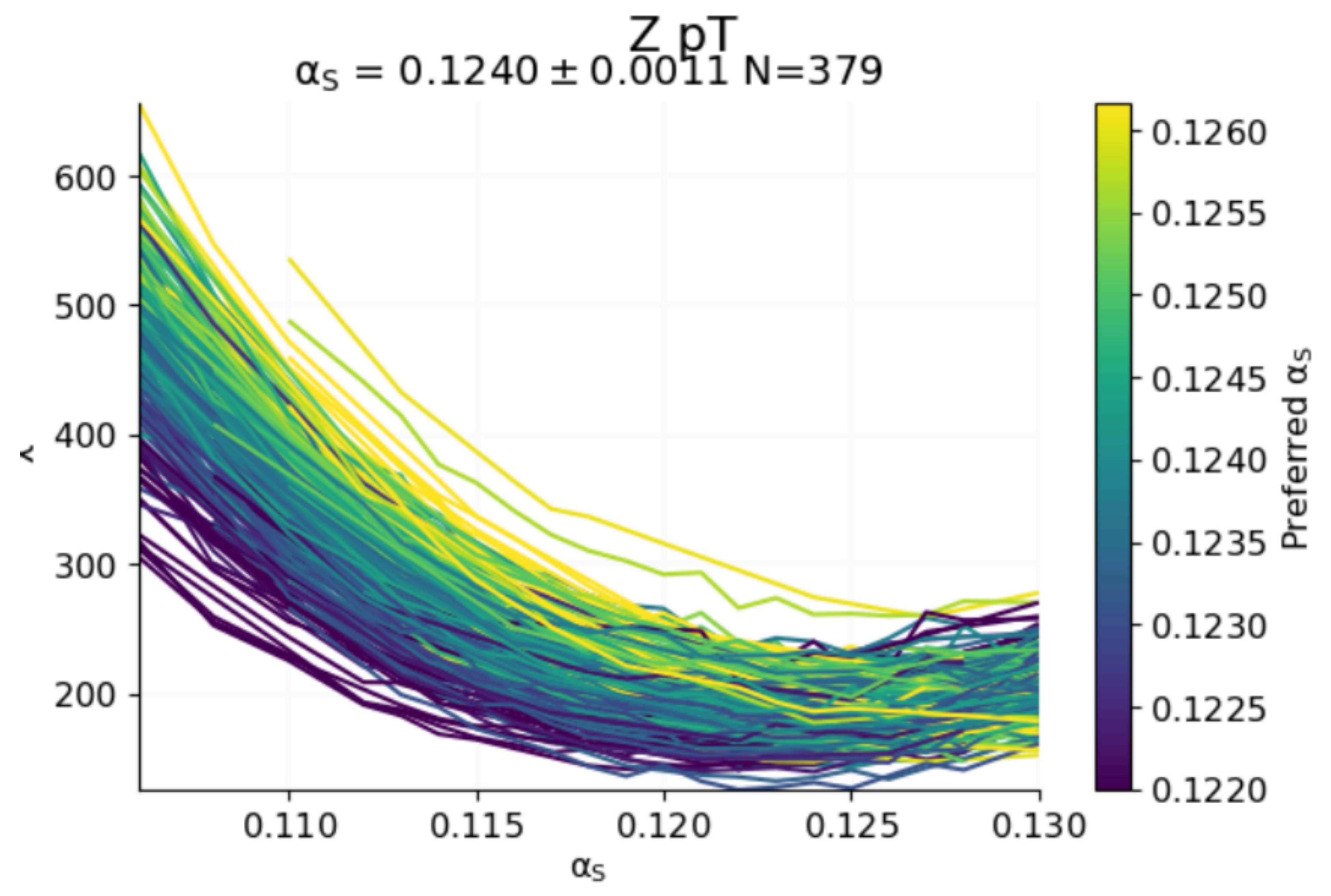
Ball, Carrazza, Del Debbio, Forte, Kassabov, Rojo, Slade, MU 1802.03398

SIMULTANEOUS FITS FOR SM PARAMETERS

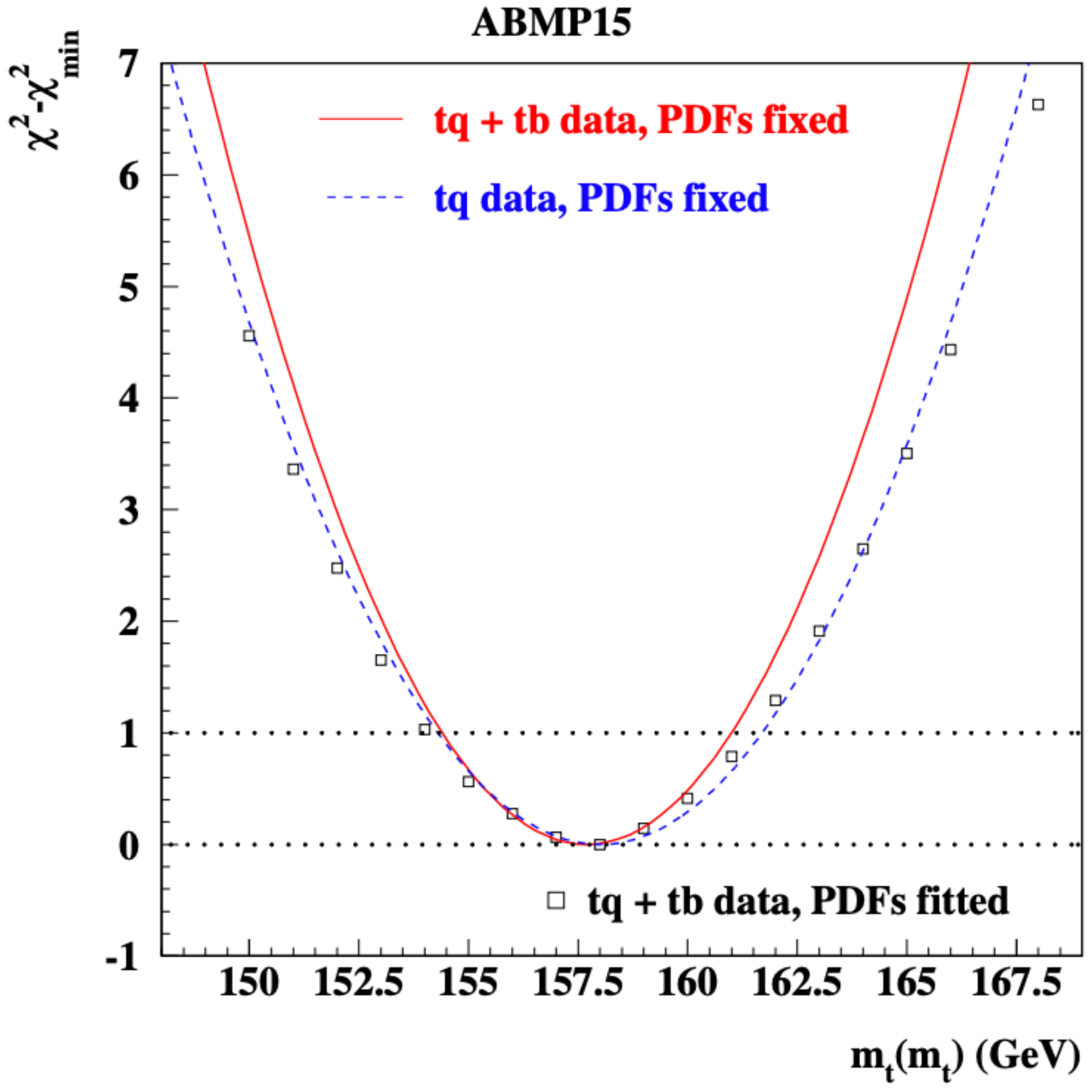
- Given the strong correlation between PDFs of the proton and α_s, a non simultaneous determination of α_s along with the PDFs from LHC processes might yield misleading results



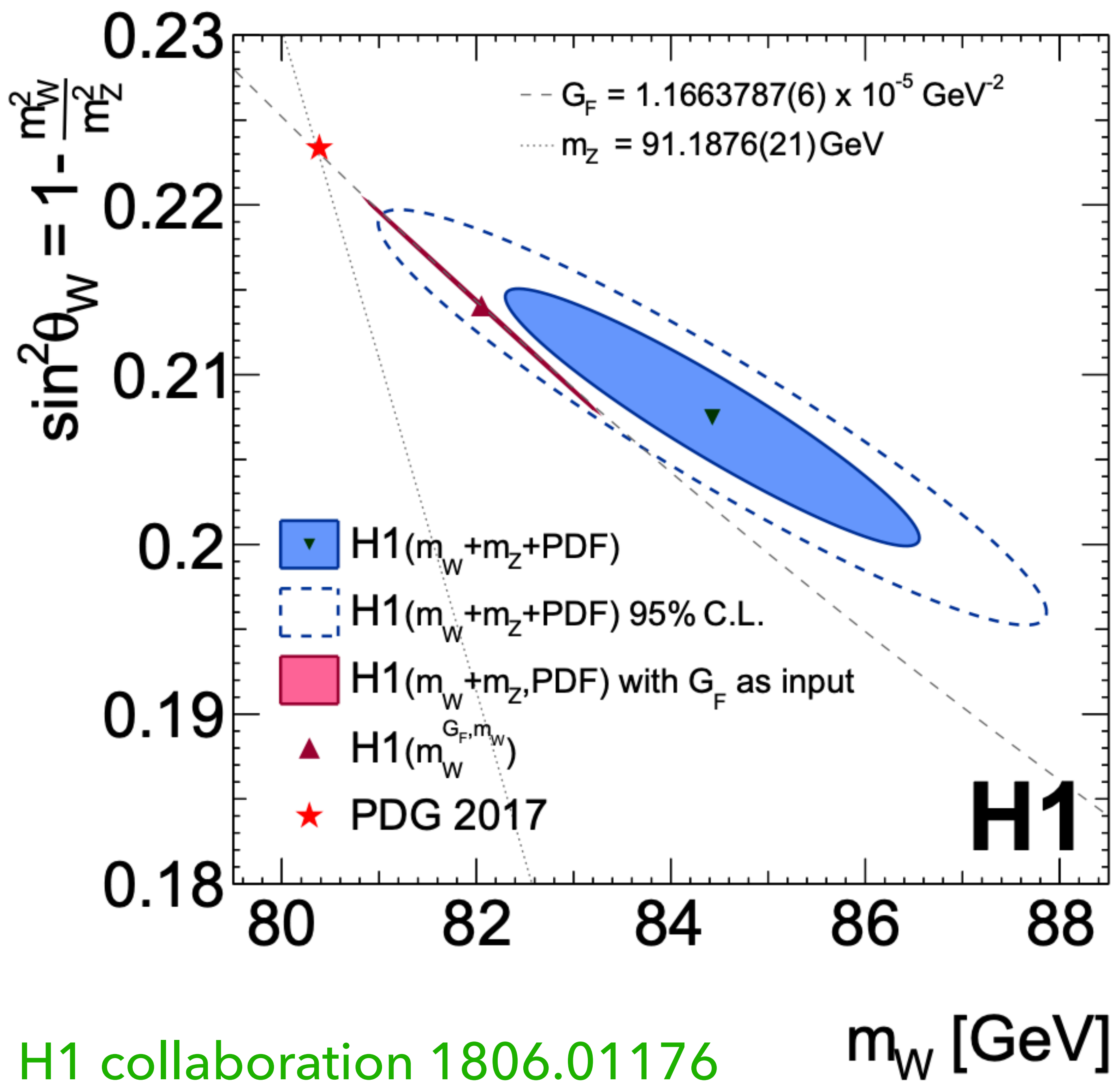
1806.01176



Forte, Kassabov 2001.04986



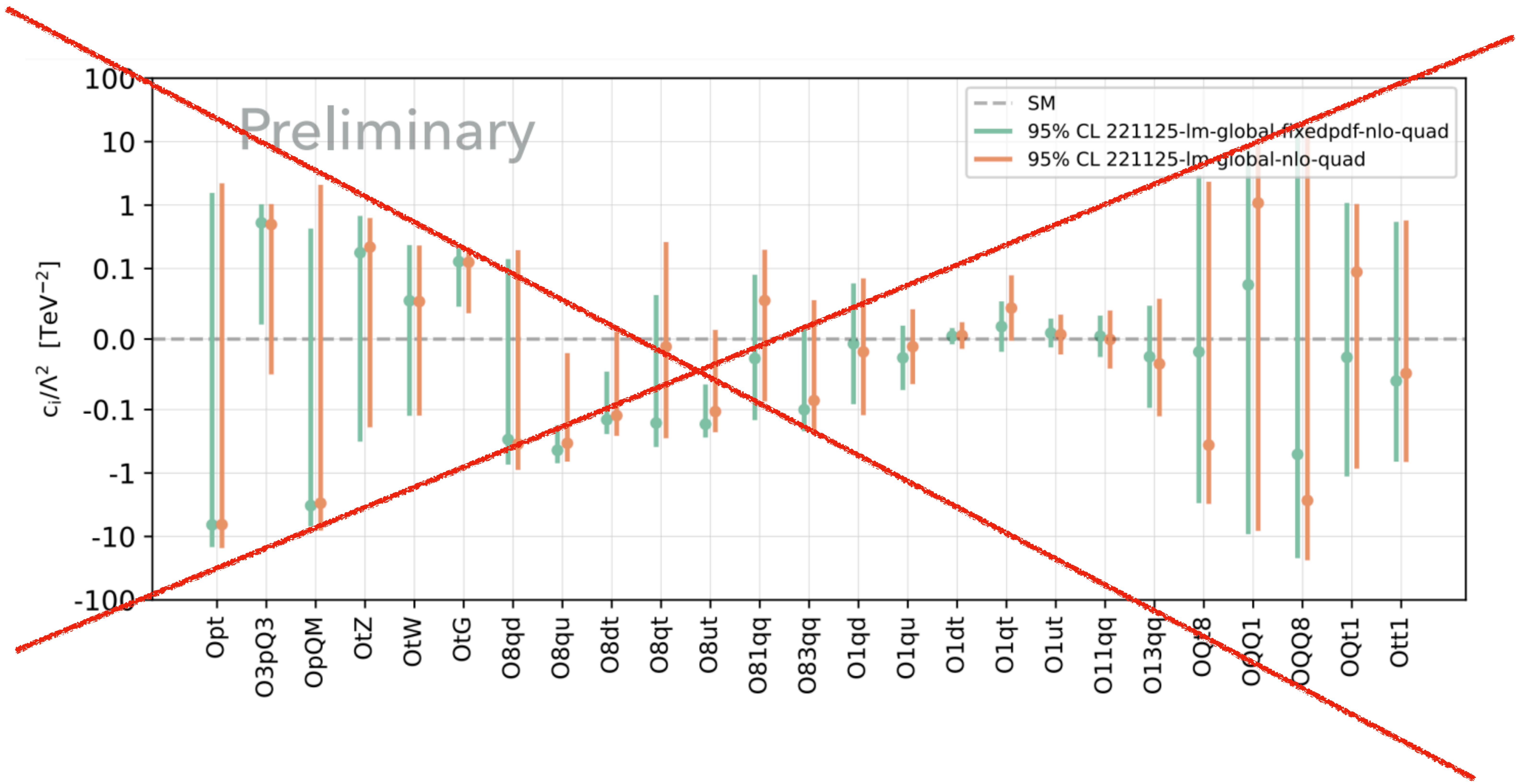
Alekhin, Moch, Their 1608.05212

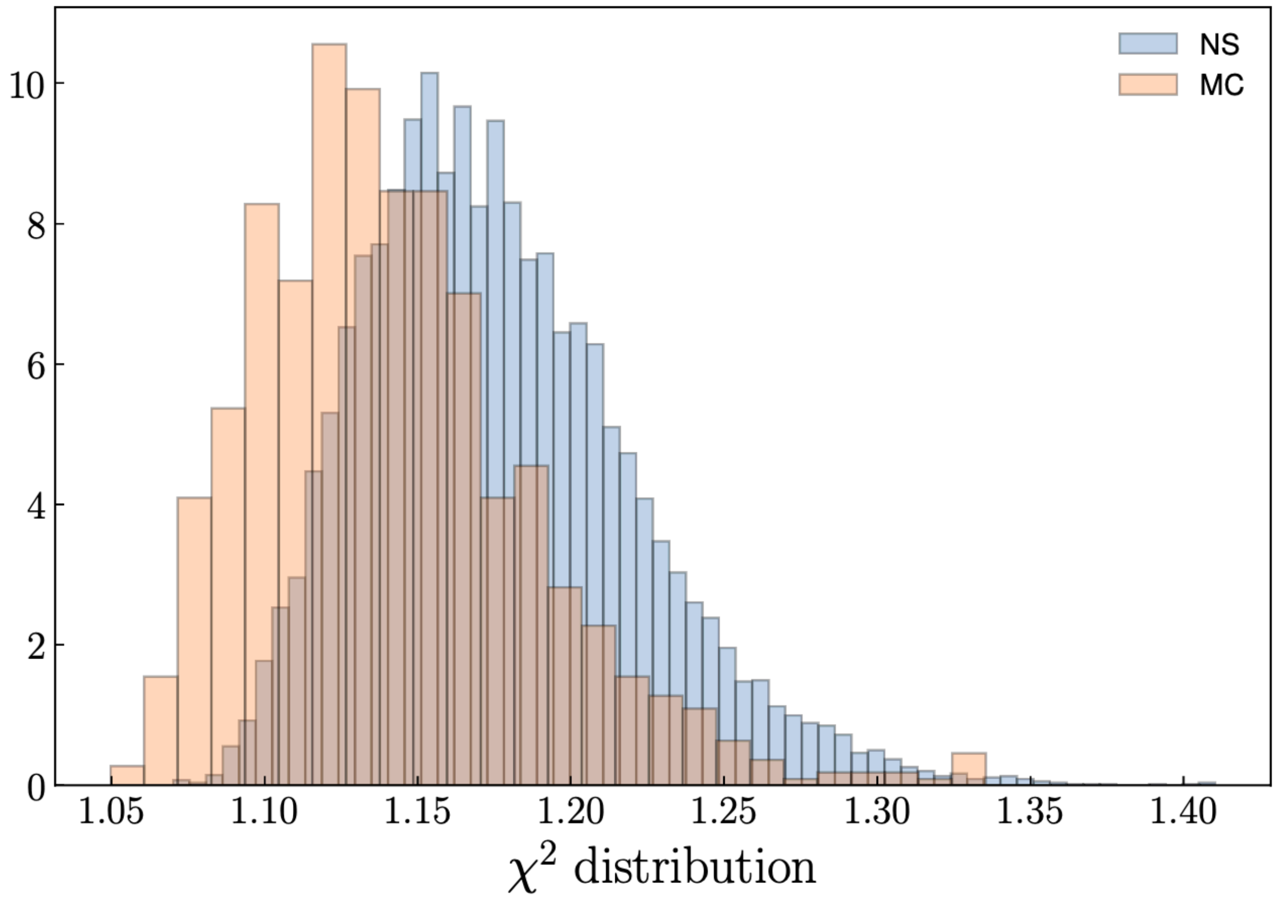


H1 collaboration 1806.01176

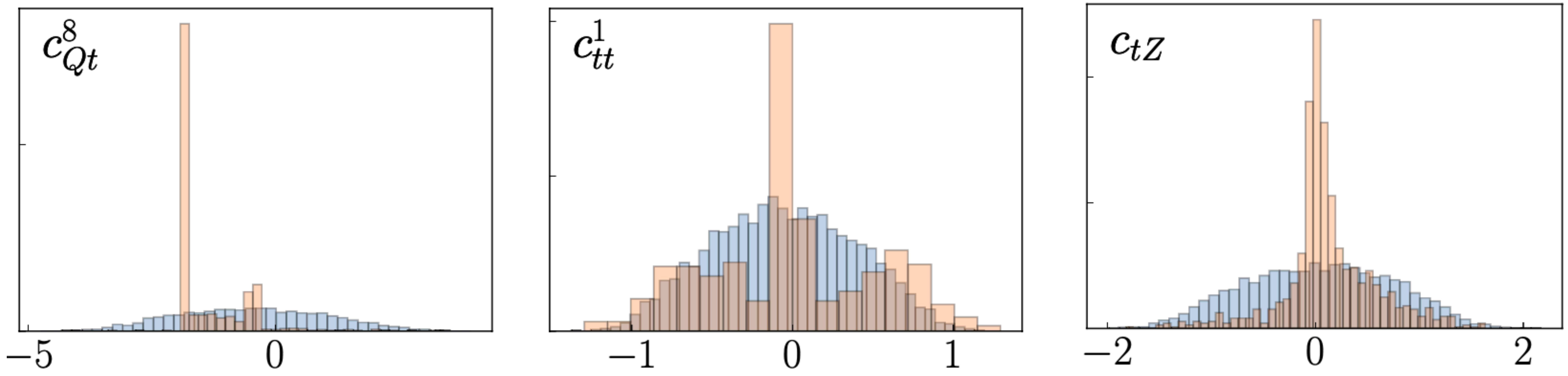
- Correlation of PDFs and the EW parameters or m_t weaker than in the case of α_s, but the very high accuracy which is sought suggests that the effect of simultaneous determination is not negligible
- Similar considerations for fits of polarised/ unpolarised PDFs, proton/nuclear PDFs or PDFs and FFs (universal fits)

Quadratic SMEFT



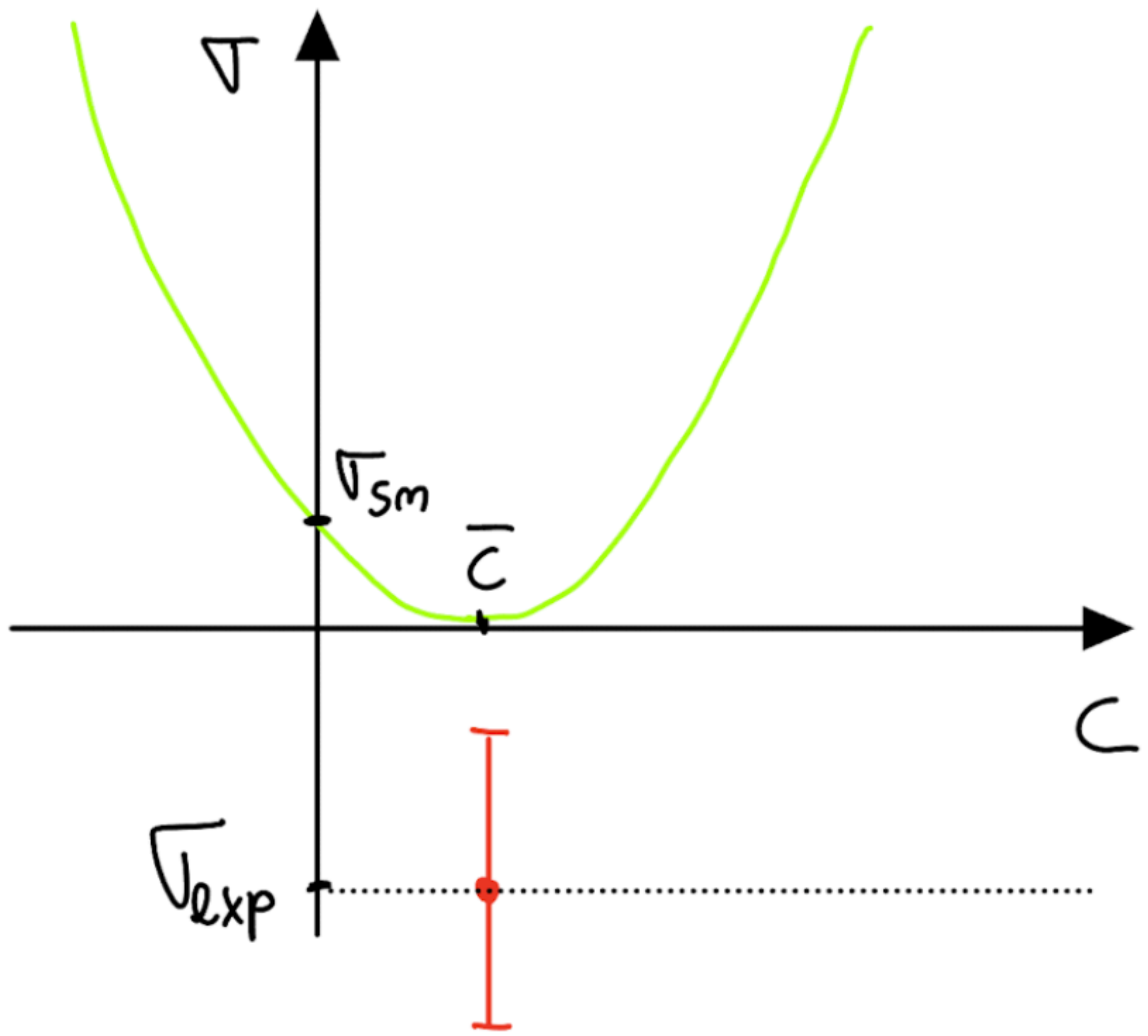


In the quadratic fit observed disagreement between MC method and Bayesian method. Very different posterior (hence different CLs)

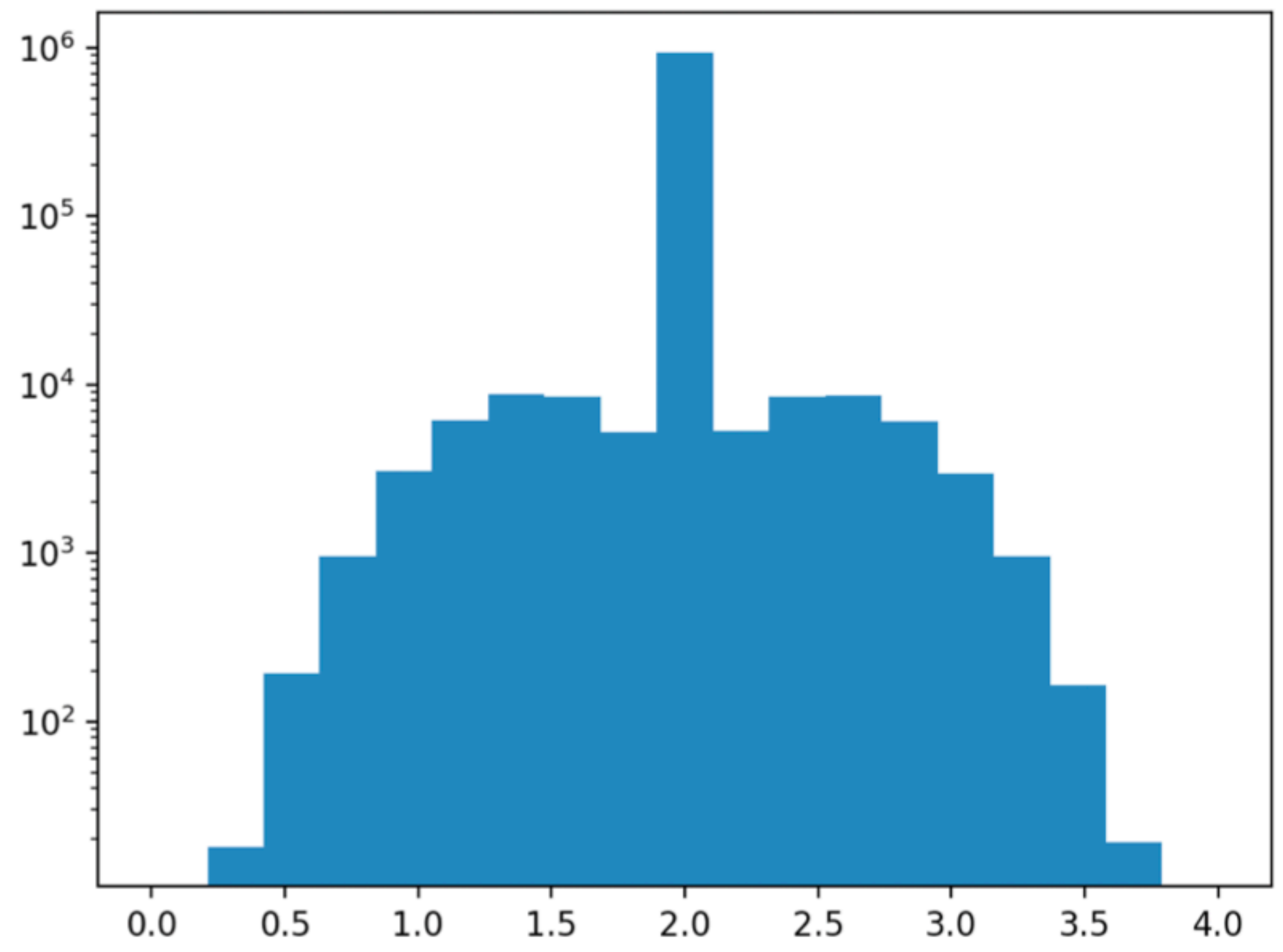


Let's consider a simple scenario: 1 operator, 1 datapoint

$$\chi^2 = \frac{(\sigma(c) - \sigma_{exp})^2}{\delta\sigma^2} \qquad \Delta\chi^2 = \chi^2 - \chi_{min} = 1 \rightarrow [c_-, c_+]$$



Computed bounds completely wrong:
the spike dominates



$$d \sim N(t(c), \sigma^2) \qquad t(c) = t^{\mathrm{SM}} + c t^{\mathrm{lin}} + c^2 t^{\mathrm{quad}} \; (\text{with } t^{\mathrm{quad}} > 0)$$

Bayesian

$$\mathbb{P}(c|d) \propto \mathbb{P}(d|c)\mathbb{P}(c)$$

P(c) uniform prior, P(d|c) Gaussian

$$\mathbb{P}(c|d) \propto \exp\left(-\frac{1}{2\sigma^2}\left(d - t(c)\right)^2\right)$$

Highest density intervals to compute 100 * α % C. L.

$$\int_{\{c:\mathbb{P}(c|d)>p(\alpha)\}} \mathbb{P}(c|d)\; dc = \alpha;$$

MC

$$d^{(1)}, \ldots, d^{(N_{\mathrm{rep}})}$$

$$c^{(i)} = \arg\min_c \chi^2(c, d^{(i)}) = \arg\min_c \left(\frac{(d^{(i)} - t(c))^2}{c^2}\right)$$

$$P_{f(X)}(y) = \int_{-\infty}^{\infty} dx \; P_X(x)\delta(y - f(x))$$

$$P_{c^{(i)}}(c) \propto \delta\left(c + \frac{t^{\mathrm{lin}}}{2t^{\mathrm{quad}}}\right)\int_{-\infty}^{t_{\min}} dx \; \exp\left(-\frac{1}{2\sigma^2}(x-d)^2\right) + \frac{2}{|2ct^{\mathrm{quad}} + t^{\mathrm{lin}}|}\exp\left(-\frac{1}{2\sigma^2}(d - t(c))^2\right)$$

↑ Spike

t_{lin} >> t_{quad} => $t_{min} \to -\infty$ then the posterior agree, else they do not