PBSP

European Research Council

Established by the European Commission

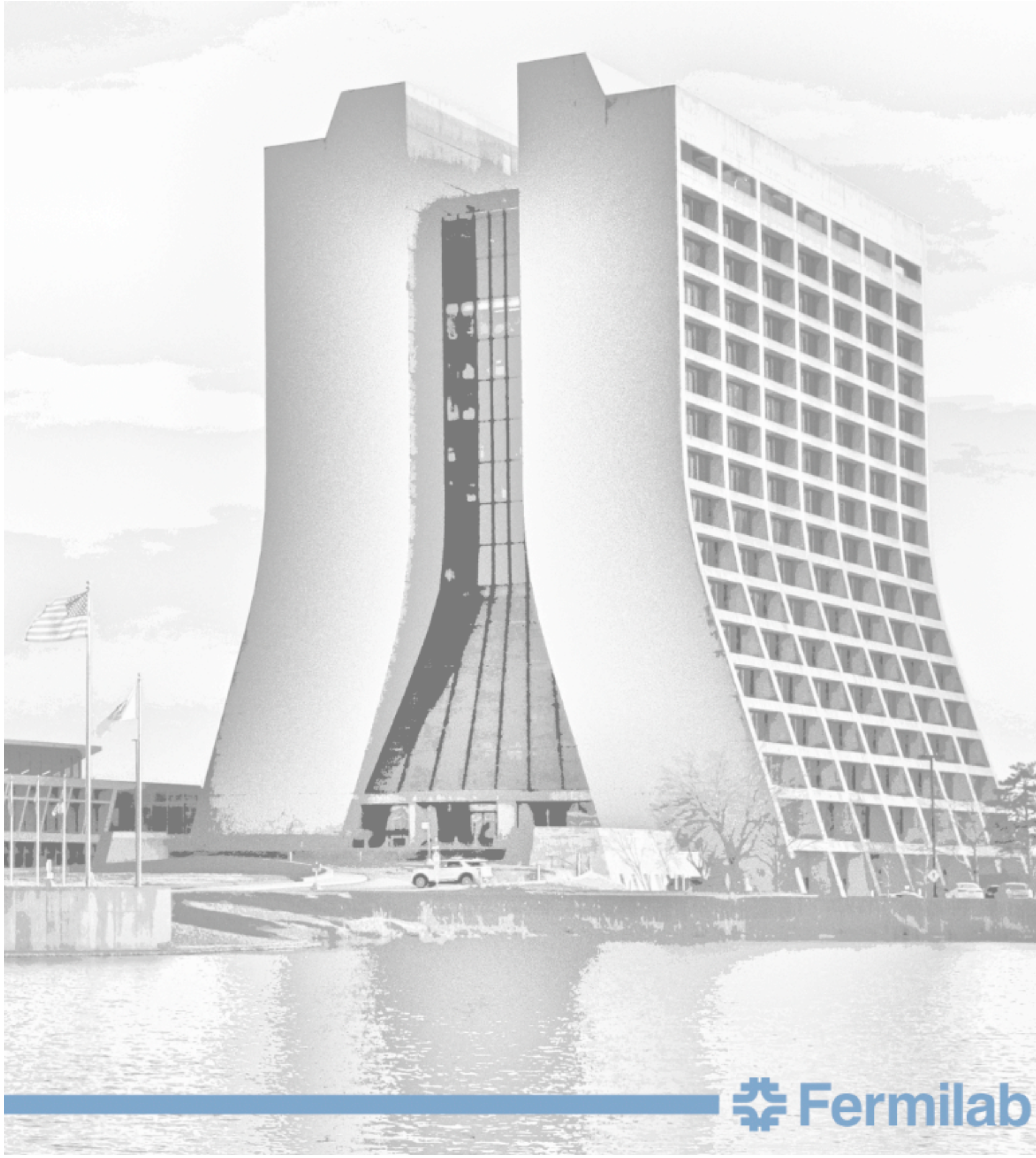
MARIA UBIALI

UNIVERSITY OF CAMBRIDGE

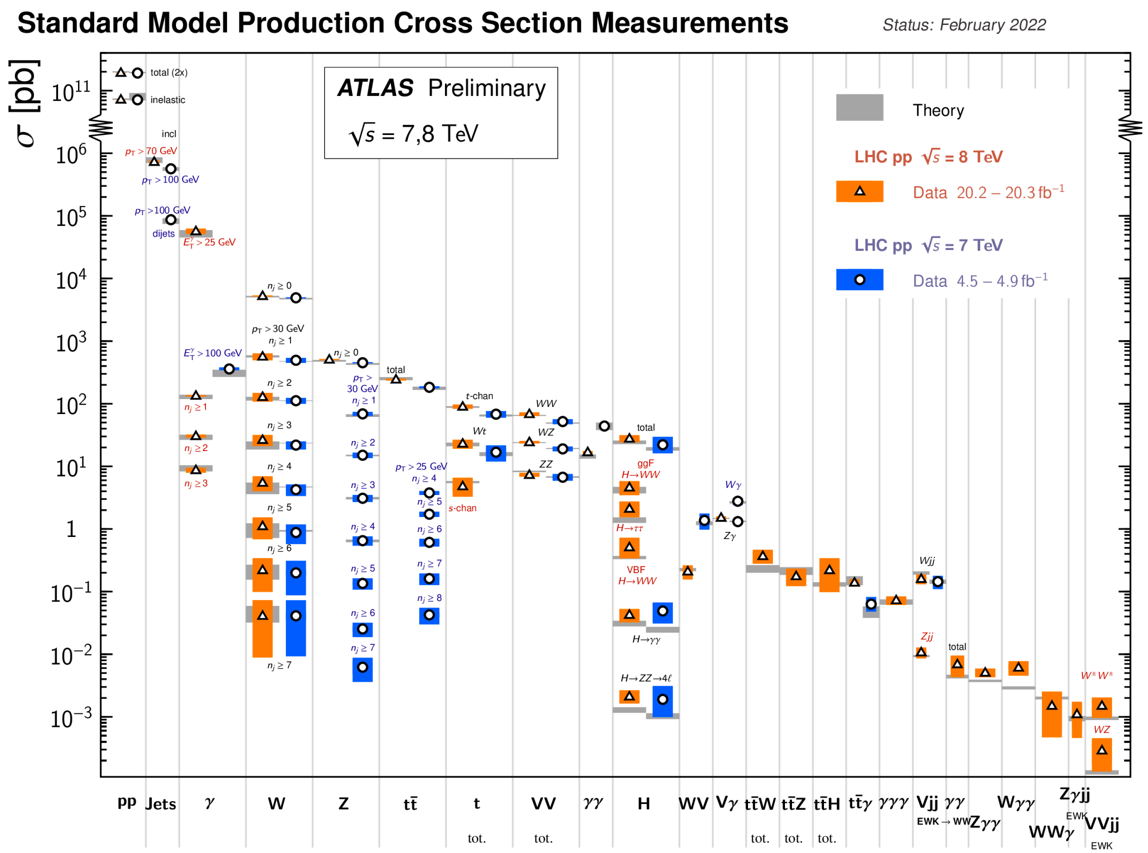
PROGRESS IN JOINT SMEFT/PDF FITS

SM@LHC2023 - FERMILAB

13TH JULY 2023



GLOBAL INTERPRETATION OF LHC DATA



ATLAS summary plots, February 2022

Extremely precise LHC data & advances in statistical techniques allow to extract SM (and BSM) parameters to a great level of precision, for example :

αs(Q) **[F. Giuli's talk],**

Parton Distribution Functions **[J. Houston's talk]**,

Higgs couplings **[yesterday's talk]**

SMEFT Wilson coefficients **[H. Mildner's talk]**

Mw **[S. Amoroso's talk]**

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

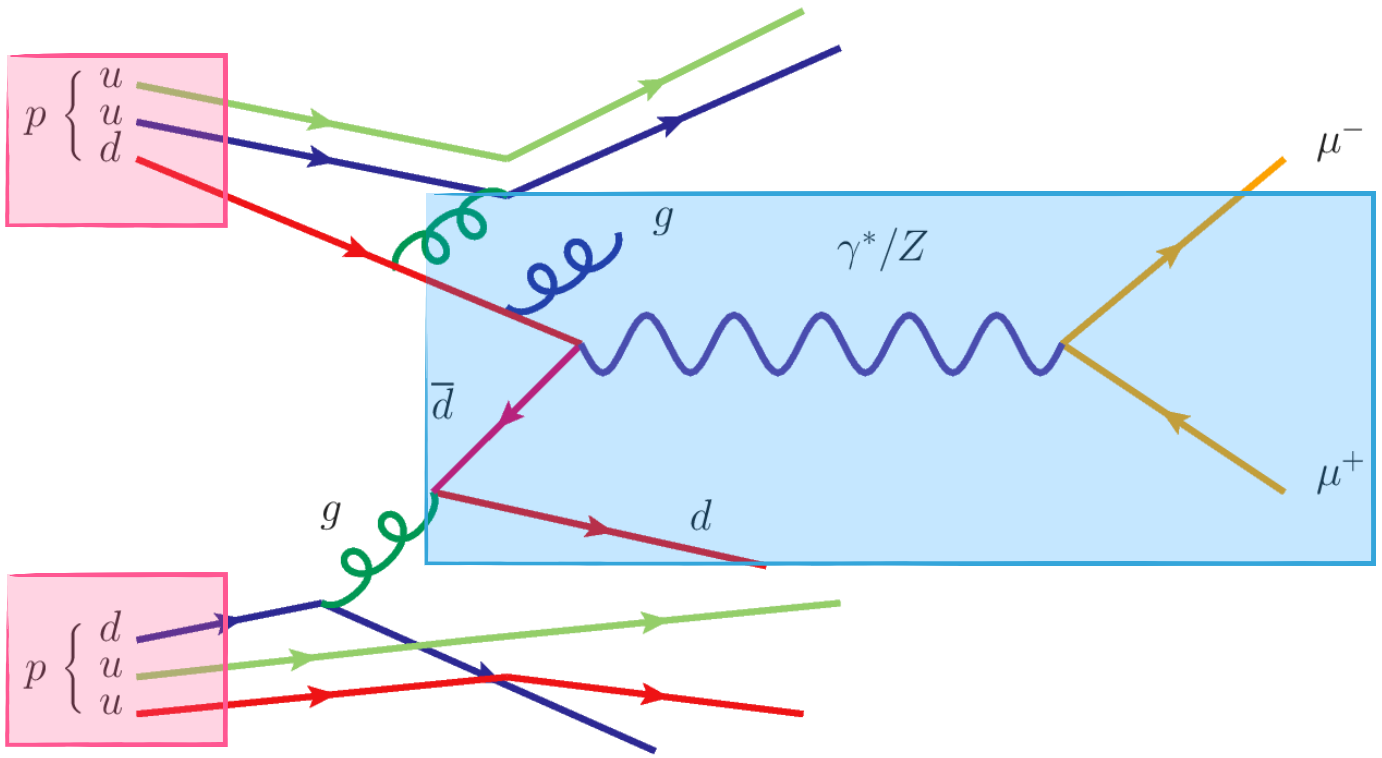
- Introduction:
  - ➡ PDF and SMEFT fits: time to study their interplay
- Simultaneous fits of PDFs and SMEFT Wilson coefficients
  - ➡ SimuNet: a tool for global simultaneous fits - S. Iranipour, MU - arXiv: 2201.07240
  - ➡ A Drell-Yan-sector analysis - A. Greljo, S. Iranipour, Z. Kassabov, M. Madigan, 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. Voice - arXiv: 2303.06159
- New physics contamination in PDF fits - E. Hammou, M. Madigan, M. Mangano, L. Mantani, J. Moore, M. Morales, MU
- Conclusions and outlook

INTRODUCTION

THEORETICAL PREDICTIONS AT THE LHC

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

Parton Distribution Functions



Hard Scattering: Perturbative QCD + EW

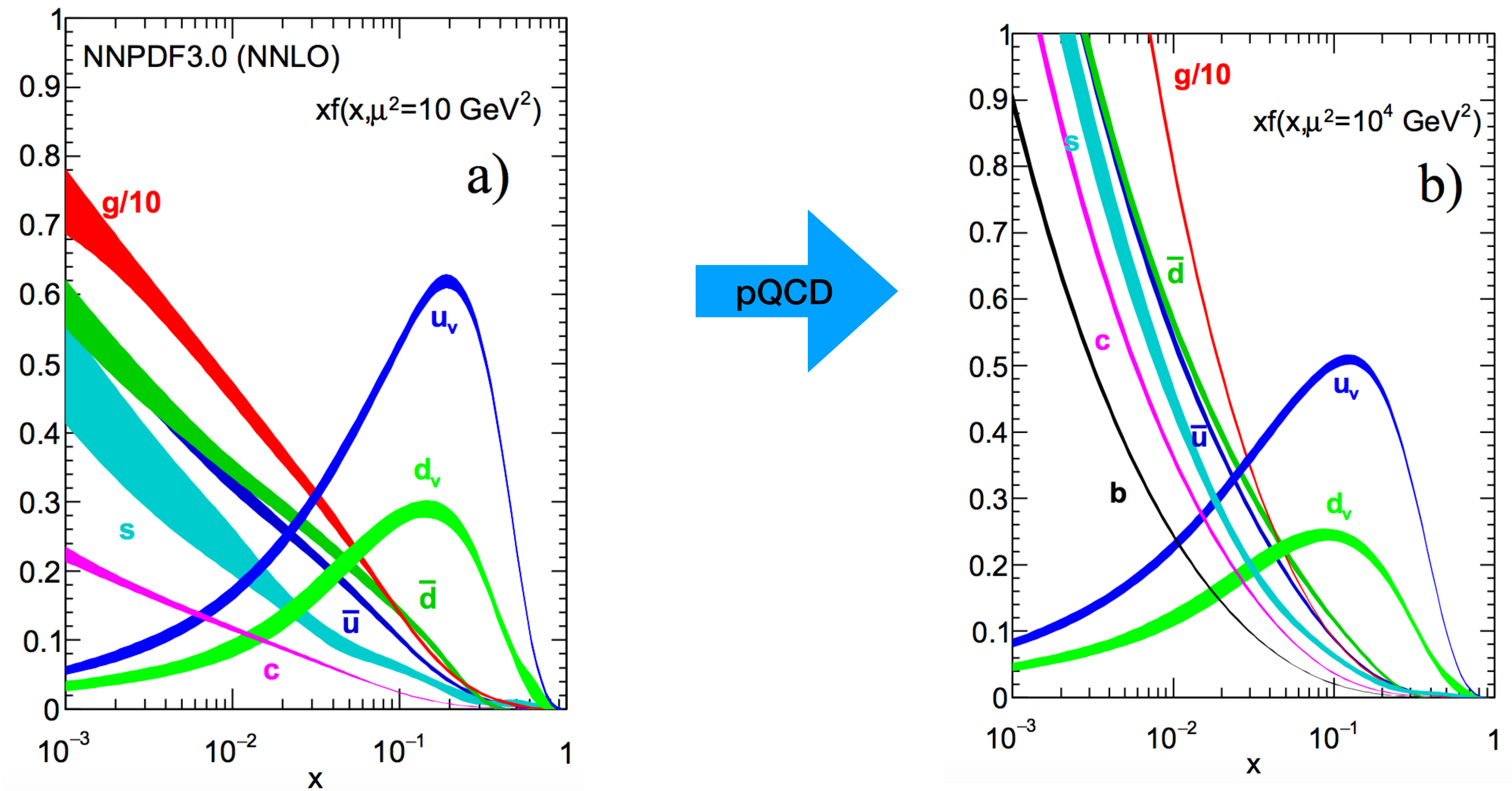
$$d\sigma^{pp\to ab} = \sum_{i,j} f_i \otimes f_j \otimes d\hat{\sigma}^{ij\to ab} + \ldots$$

PARTON DISTRIBUTION FUNCTIONS

$$f_i(x, \mu)$$

Data ← x | μ → Perturbative QCD

Hadronic scale: global fit of PDFs



High scale: input to the LHC

PDG 2016

EXTRACTING PARAMETERS FROM DATA

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

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

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

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

Parameters determining PDFs at initial scale

EXTRACTING PARAMETERS FROM DATA

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

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

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

(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

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

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

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

(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

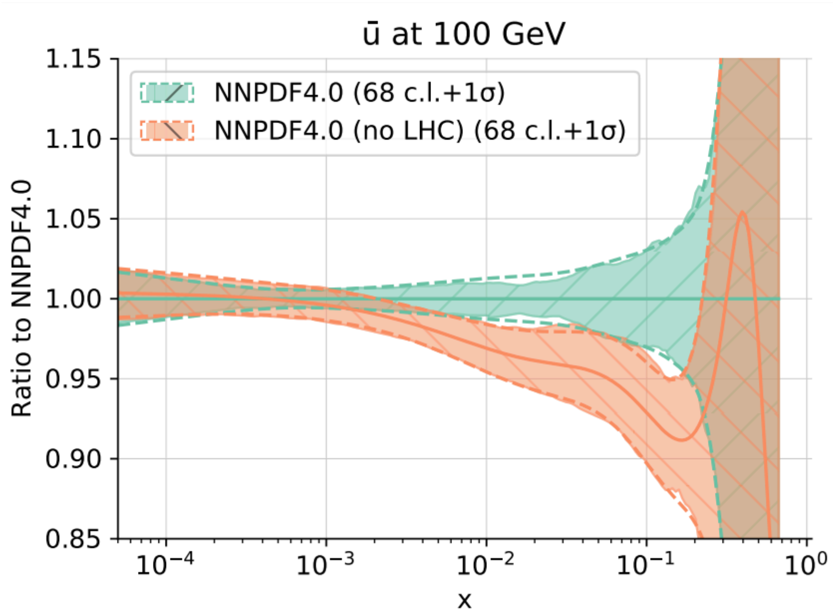
$$T_i(\{c\}) = {\rm PDFs}(\{\theta=\bar{\theta}\},\{c=0\}) \otimes \hat{\sigma}_i(\{c\})$$

PDF AND SMEFT INTERPLAY

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

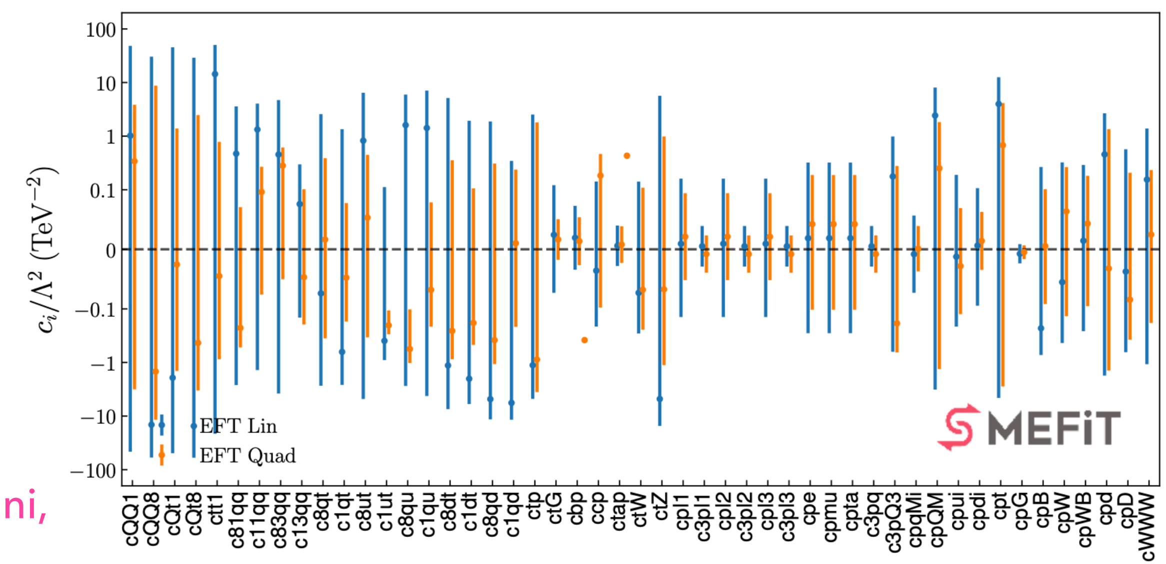
- PDFs are low-scale quantities extracted from experimental data at all scales, without considering any potential high-scale contamination due to new physics.

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



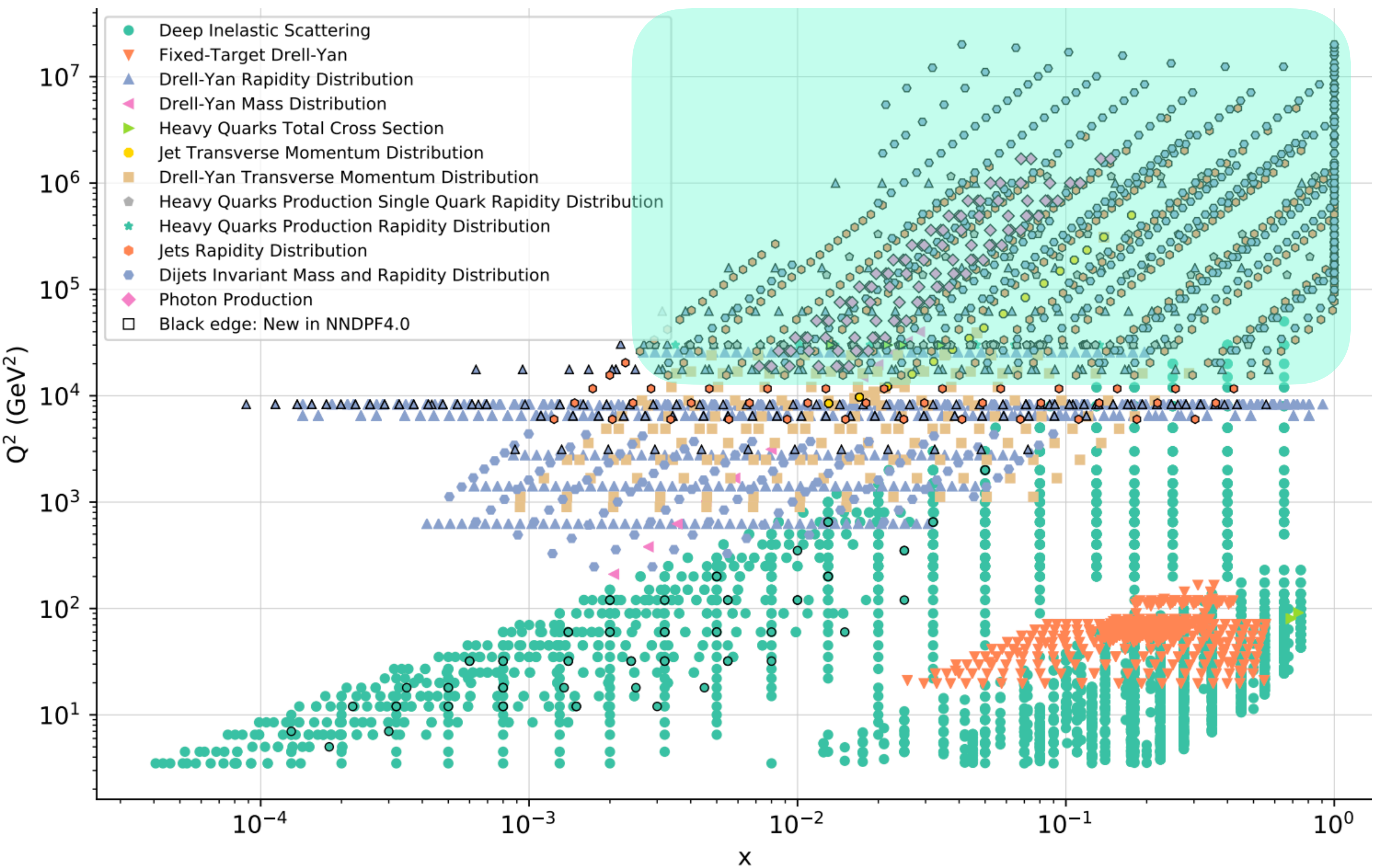
Ball et al, arXiv:2109.02653

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



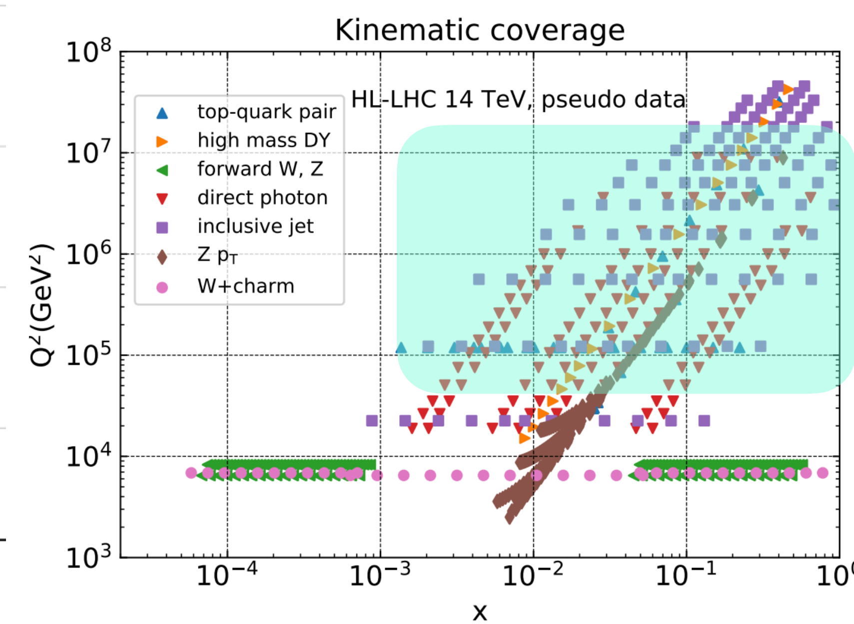
Ethier et al, arXiv: 2105.00006

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

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

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

Simultaneous fits can shed light on their interplay

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

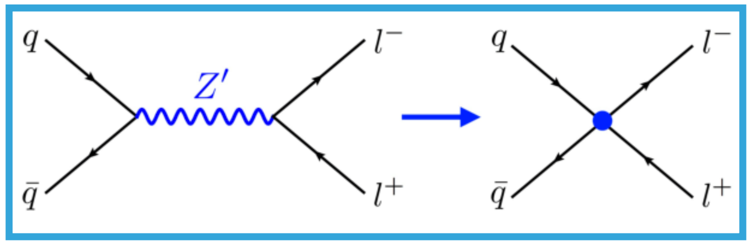
$$f(\{\theta_k\})$$

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

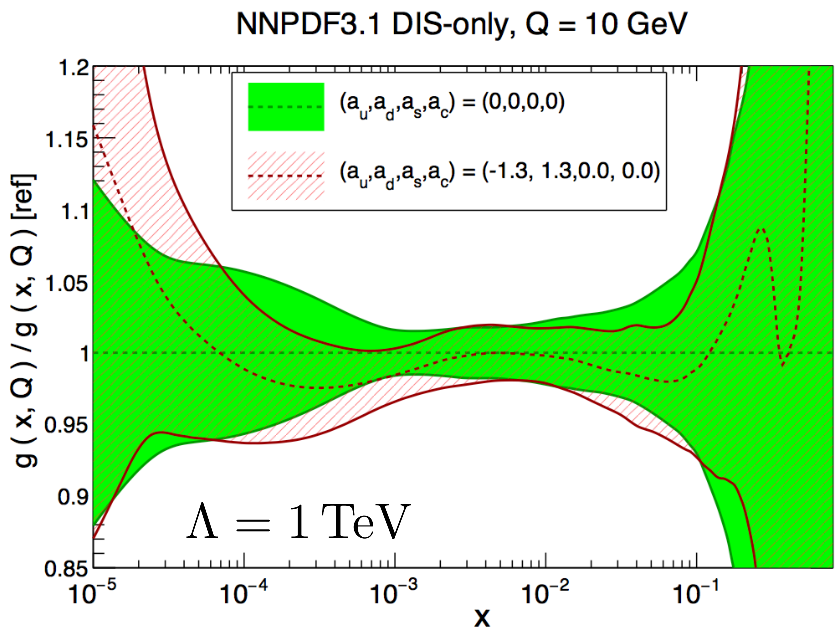
SIMULTANEOUS PDF AND SMEFT FITS

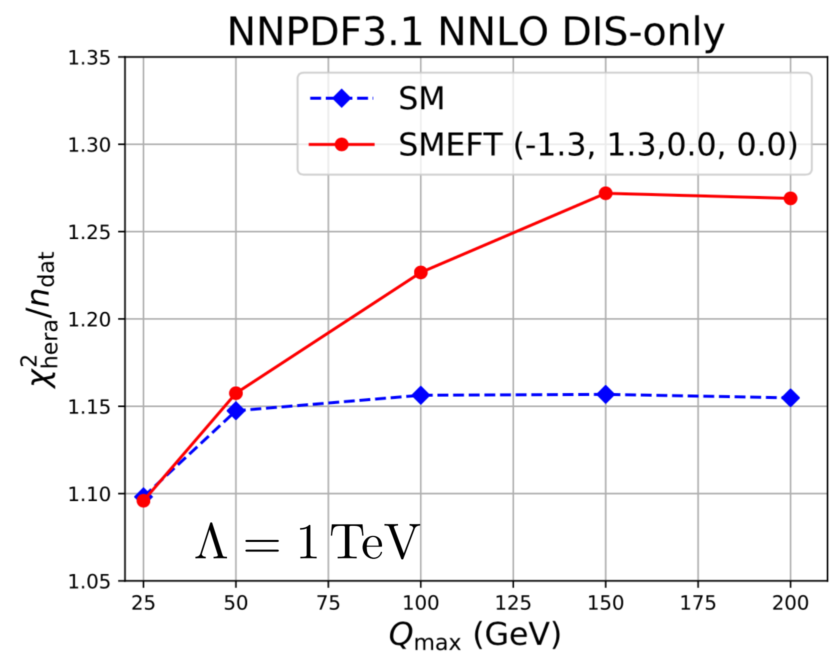
A SCAN-BASED APPROACH: A DIS CASE STUDY

- 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)



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



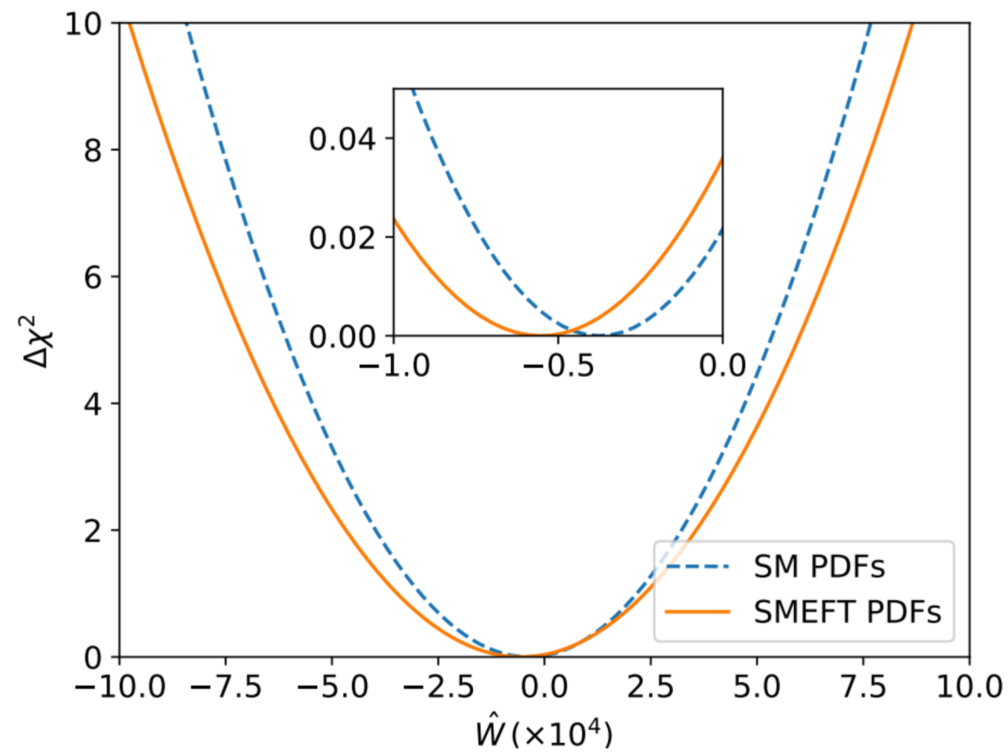


Only gluon affected by the presence of non-zero coefficients, but distortion of PDFs leads to a deterioration of data-theory agreement that scales with energy
=> A fit based on DIS data is only moderately affected by interplay and the effects of new physics can be disentangled

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]

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



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

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

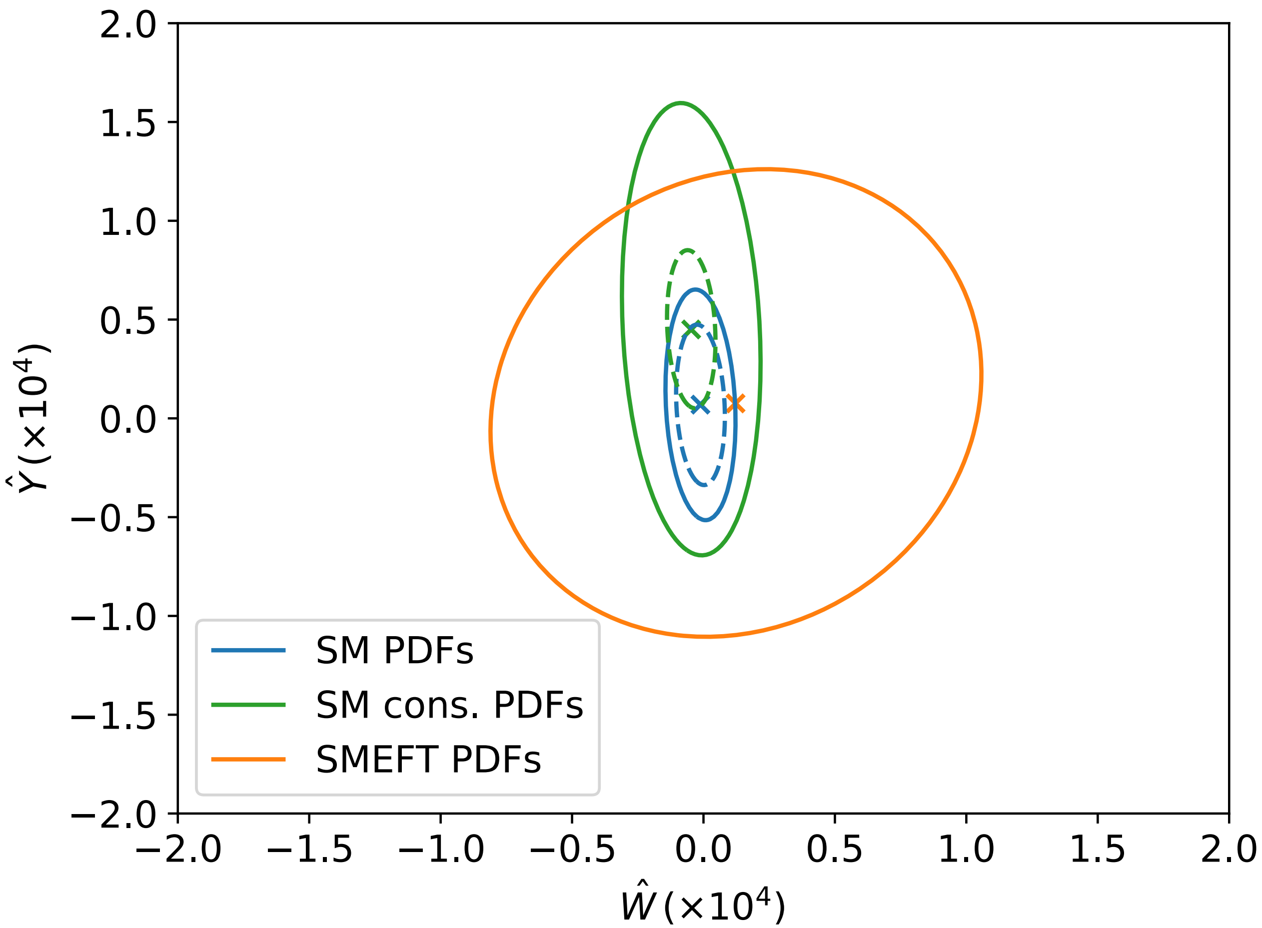
1. Take data, make theoretical predictions accounting for operator in partonic cross section and PDFs.
2. Compute chi2 as a function of WCs (Wilson Coefficients)
3. Minimise chi2 and find best-fit and C.L.s of WCs
4. Extract bounds

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

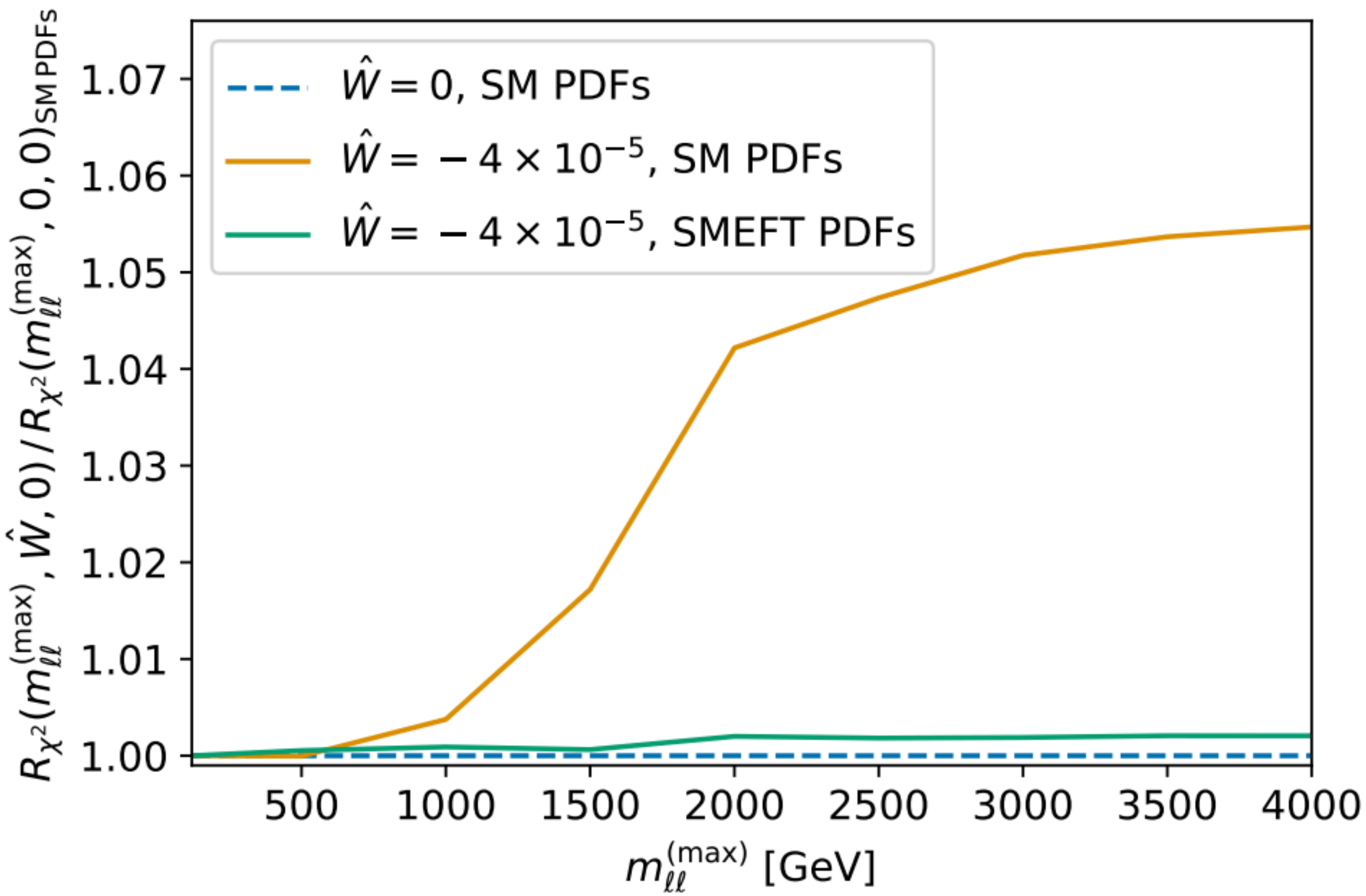
SMEFT PDFs / Simultaneous fit

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

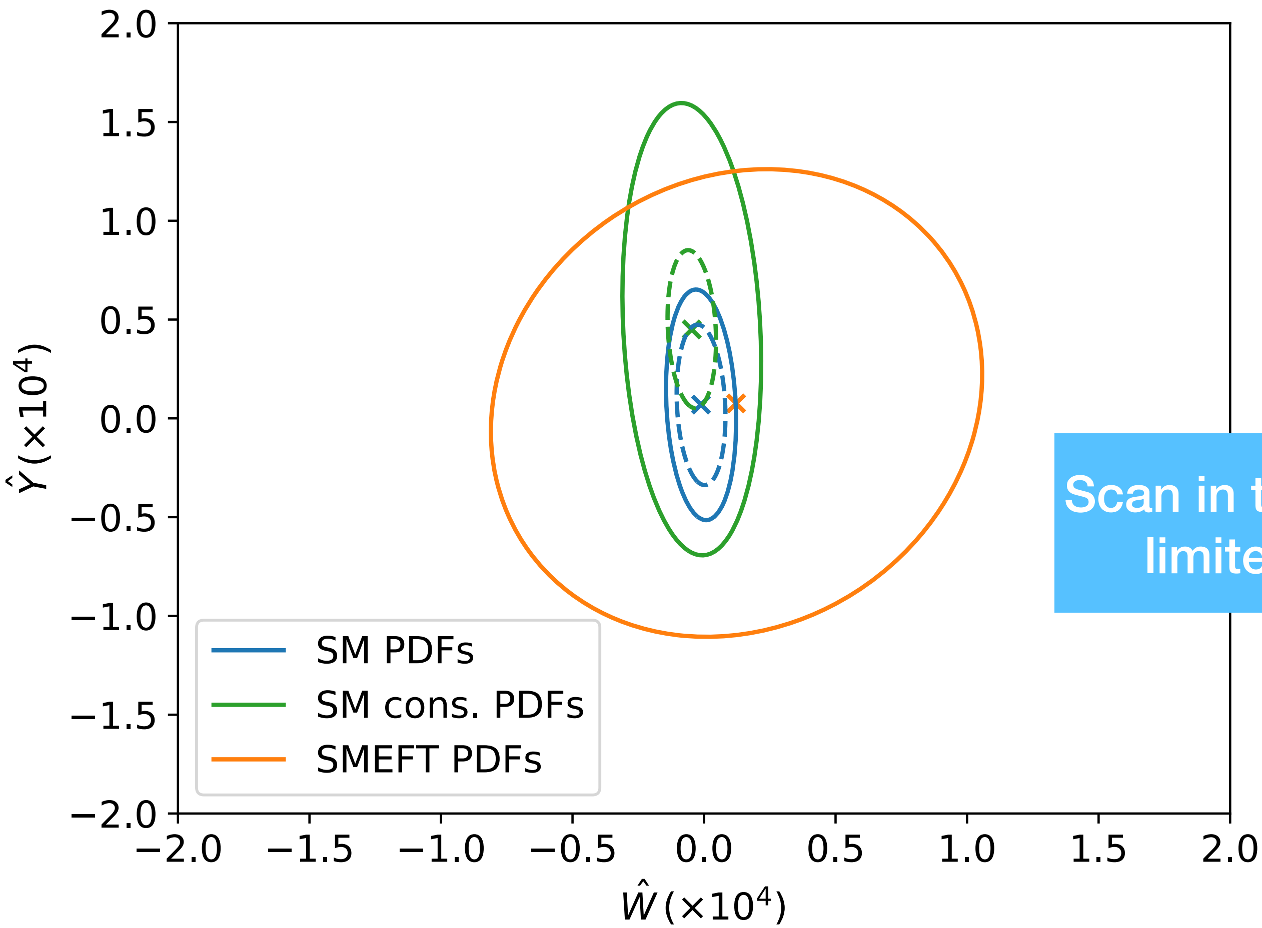


Greljo, Iranipour, Kassabov, Madigan, Moore, Rojo, MU, Voisey: 2104.02723

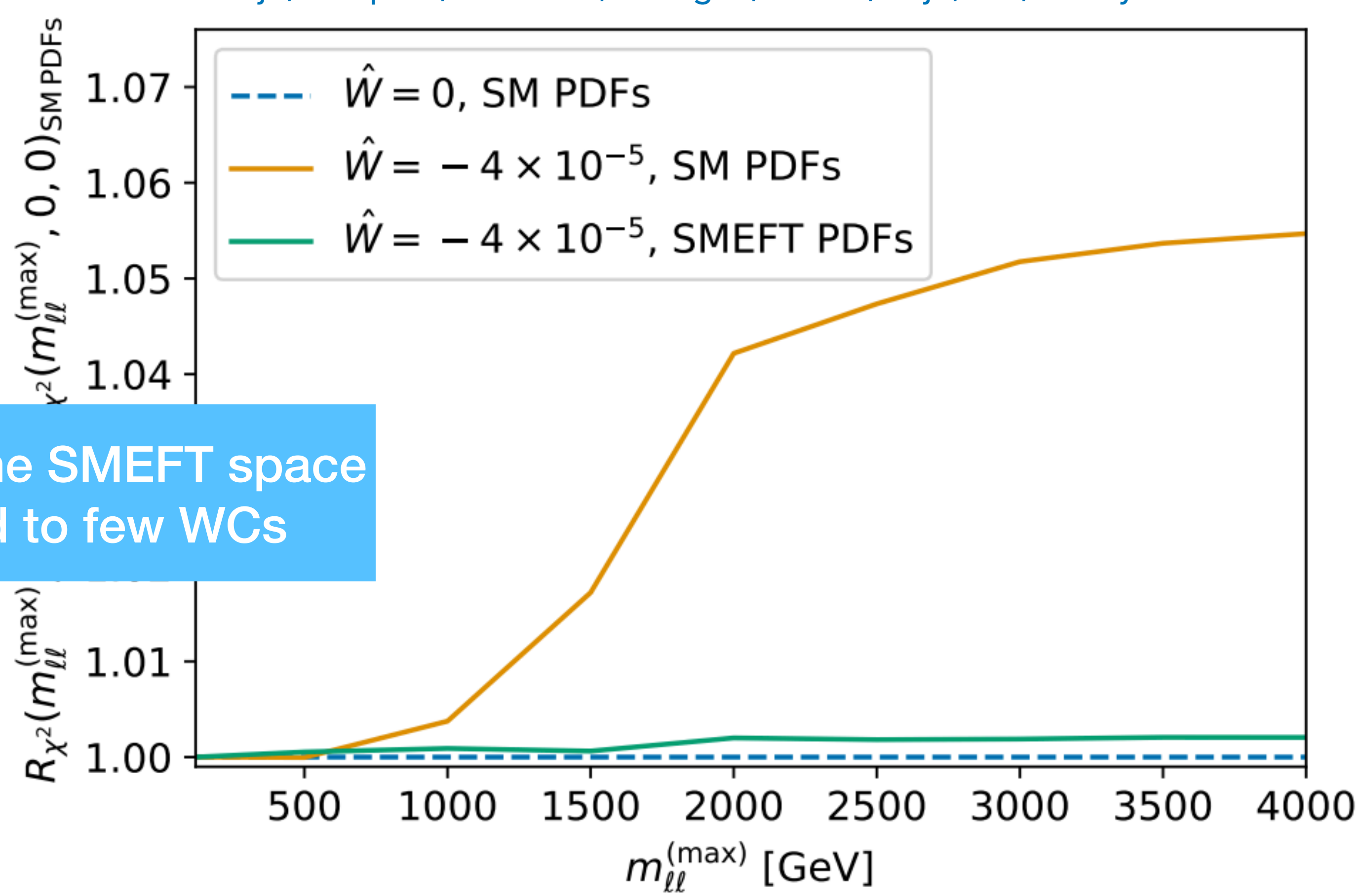


A SCAN-BASED APPROACH: A DY CASE STUDY

- 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

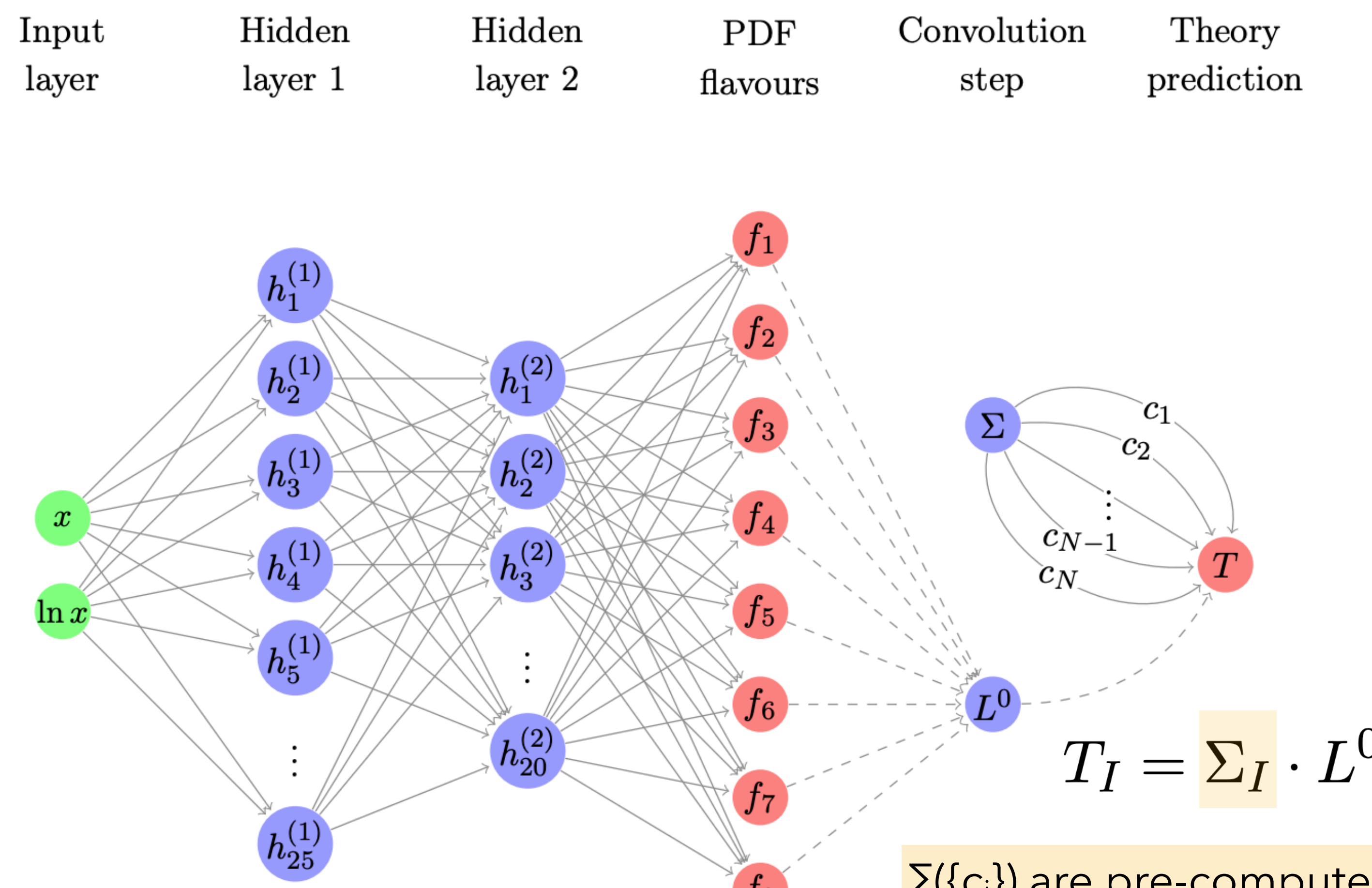
- The idea: take a PDF fit based on NNPDF4.0 methodology and make dependence of observables on physics parameters $\{c_i\}$ explicit before computing the loss function (e.g. adding SMEFT corrections, or expanding observables in terms of SM precision parameters)

- 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.



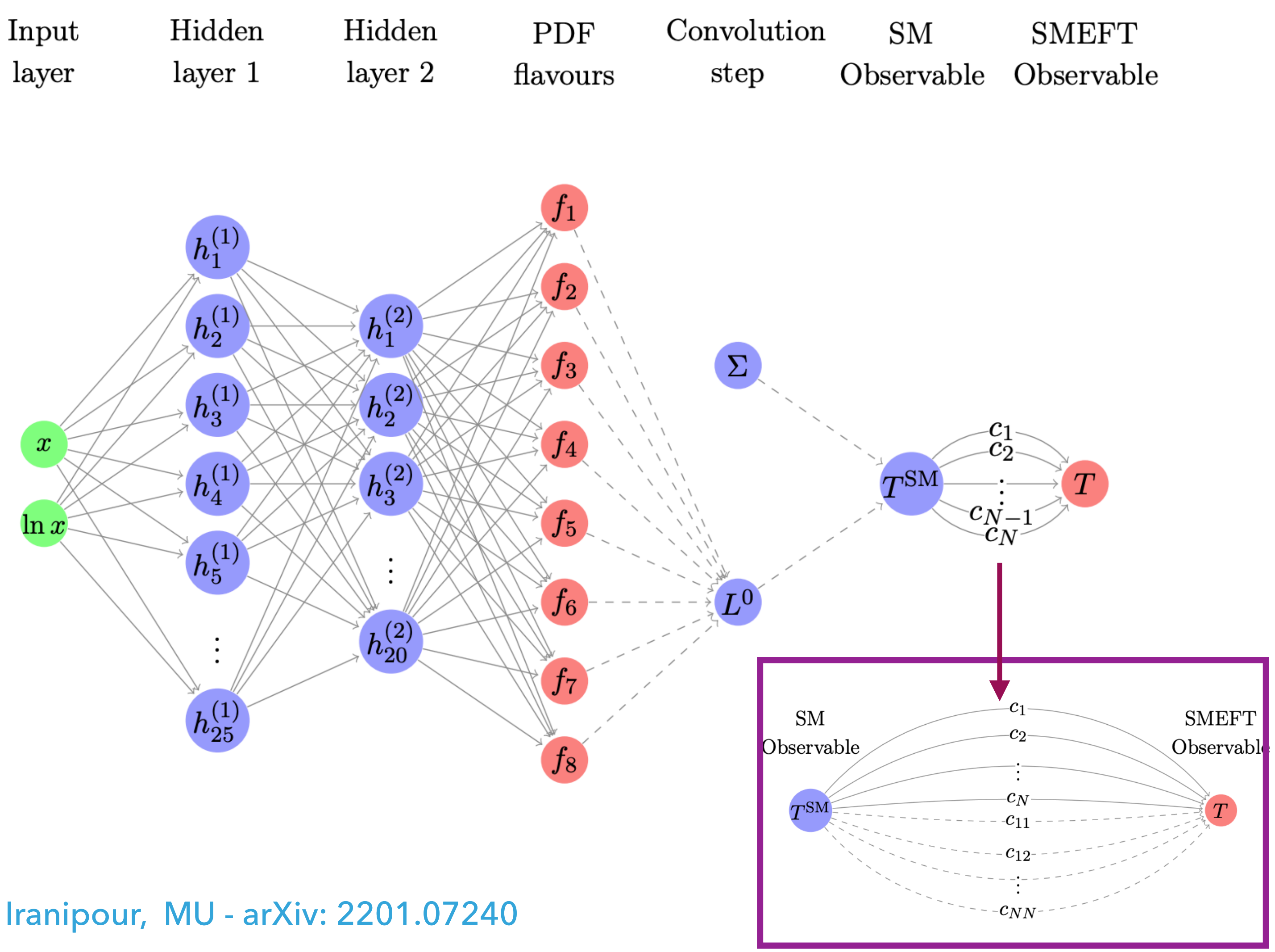
$$T_I = \Sigma_I \cdot L^0$$

$\Sigma(\{c_i\})$ 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

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

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

Quadratic dim-6 operator

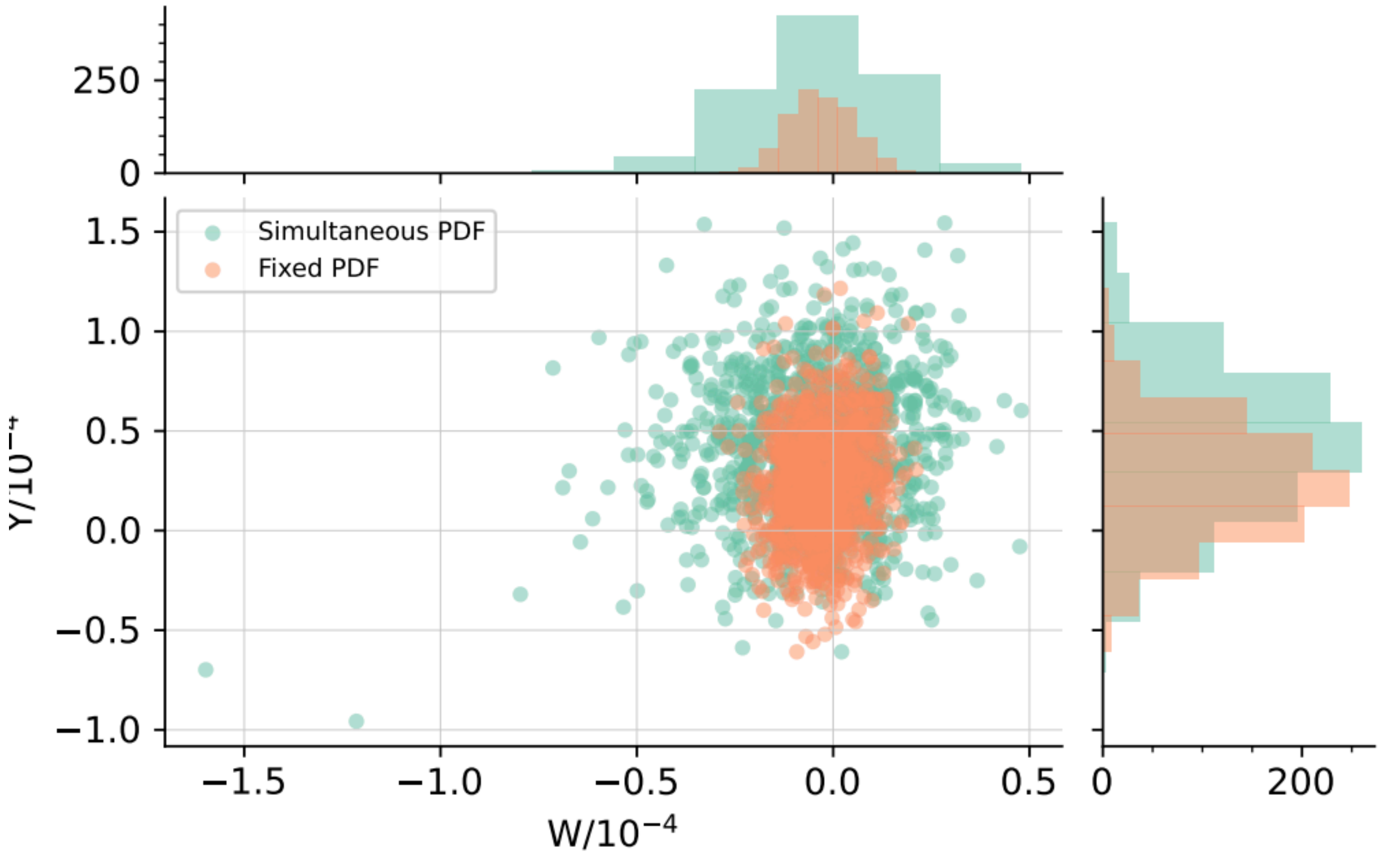
$$T(\hat{\theta}) = T^{\rm SM}(\theta) \cdot \left(1 + \sum_{n=1}^{N} c_n R^{(n)}_{\rm SMEFT} + \sum_{1\leq n\leq m\leq N} c_{nm} R^{(n,m)}_{\rm SMEFT}\,,\right)$$

$c_n c_m$

S. Iranipour, MU - arXiv: 2201.07240

RESULTS: DRELL-YAN DATA @HL-LHC

S. Iranipour, MU - arXiv: 2201.07240



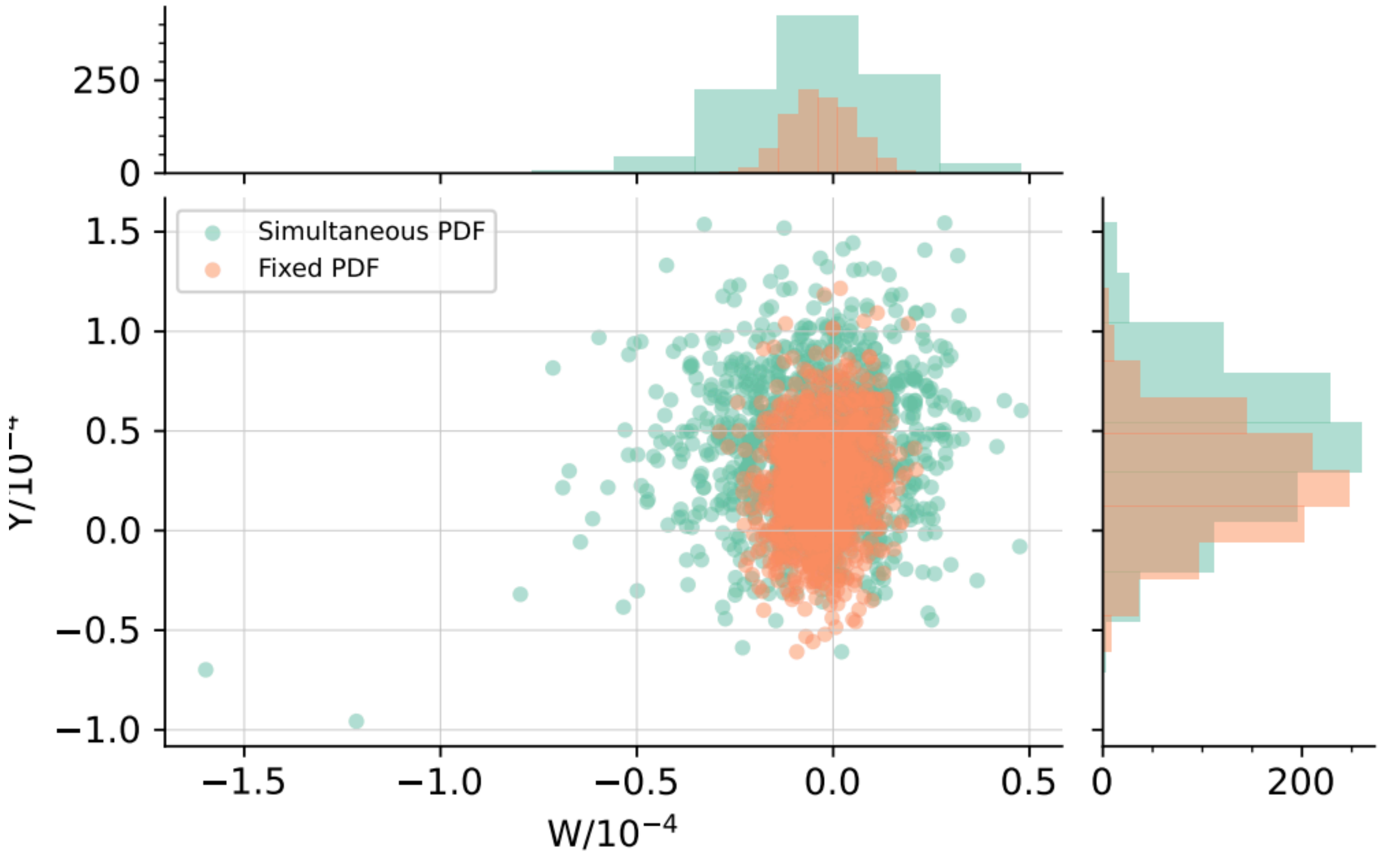
	SM PDFs	SMEFT PDFs
$W \times 10^5$ (68% CL)	$[-1.1, 0.5]$	$[-2.4, 1.5]$
$W \times 10^5$ (95% CL)	$[-2.0, 1.4]$	$[-4.3, 3.4]$
$Y \times 10^5$ (68% CL)	$[-0.4, 5.2]$	$[0.6, 8.0]$
$Y \times 10^5$ (95% CL)	$[-3.2, 8.1]$	$[-3.1, 11.7]$

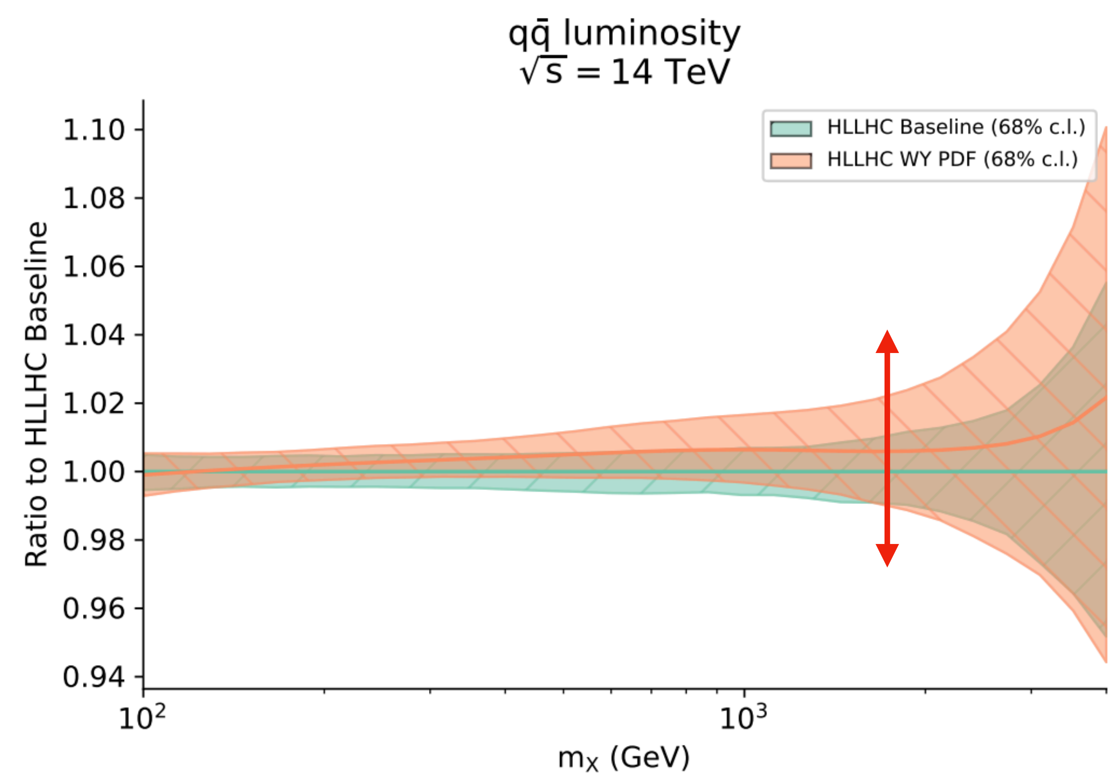
x 2.3 broadening of bounds for W
x 1.3 broadening of bounds for Y

✓ 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

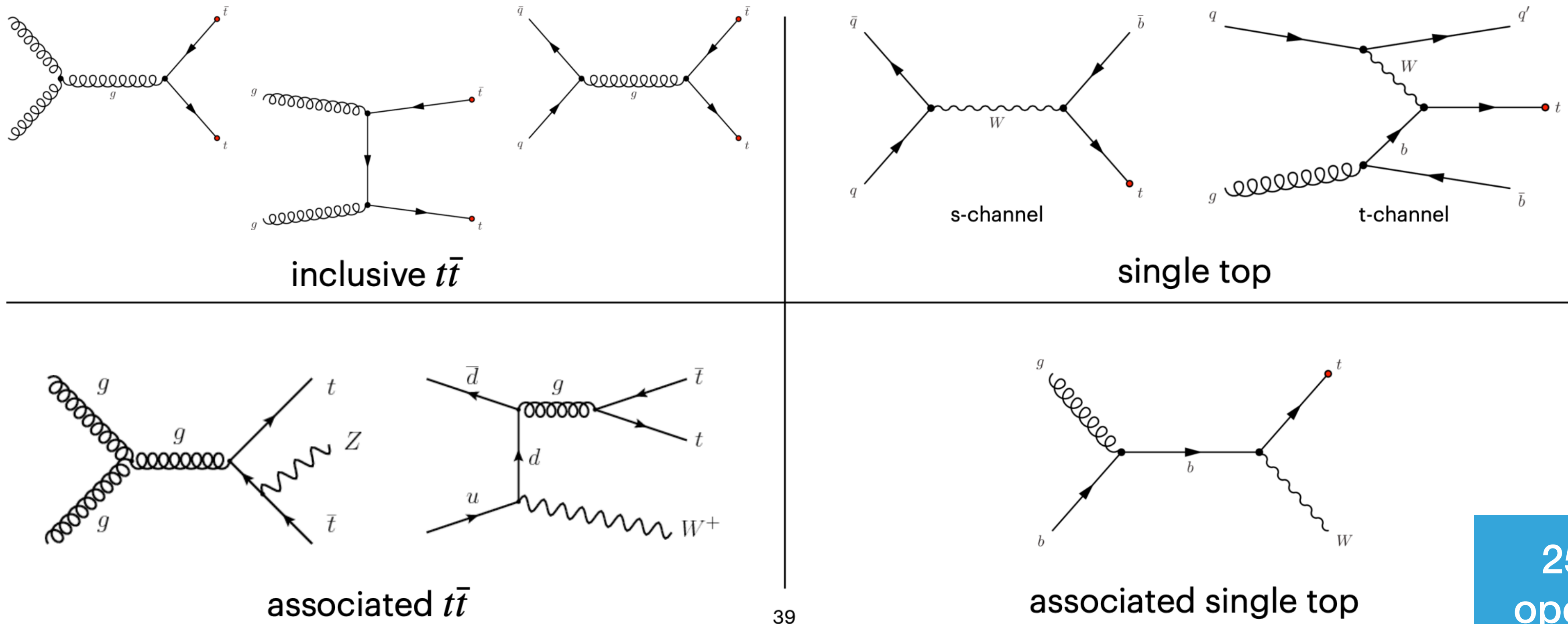




✓ Simultaneous analysis of PDFs and W&Y SMEFT coefficient of DIS + DY (including HL-LHC projections) using simuNET method shows that at HL-LHC the effect of interplay becomes important as WCs bounds broaden and PDF uncertainties change significantly once SMEFT effects allowed in theory predictions entering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

THE TOP SECTOR

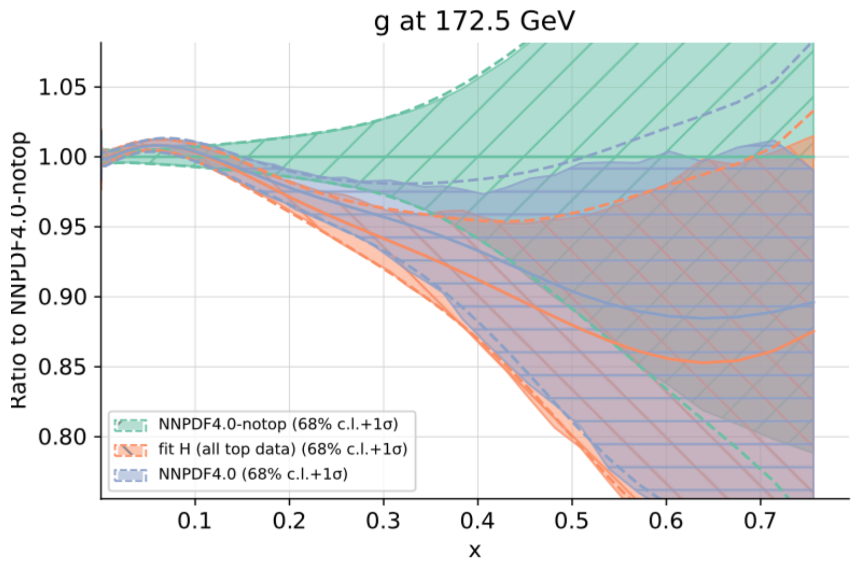
- After testing methodology on small number of WC, stress-test on large SMEFT parameters space.
- Huge amount of Run II top quark data from ATLAS and CMS
- Four basic processes: inclusive tt~ and asymmetry (inclusive and differential), single top (inclusive and differential), associated ttV production,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

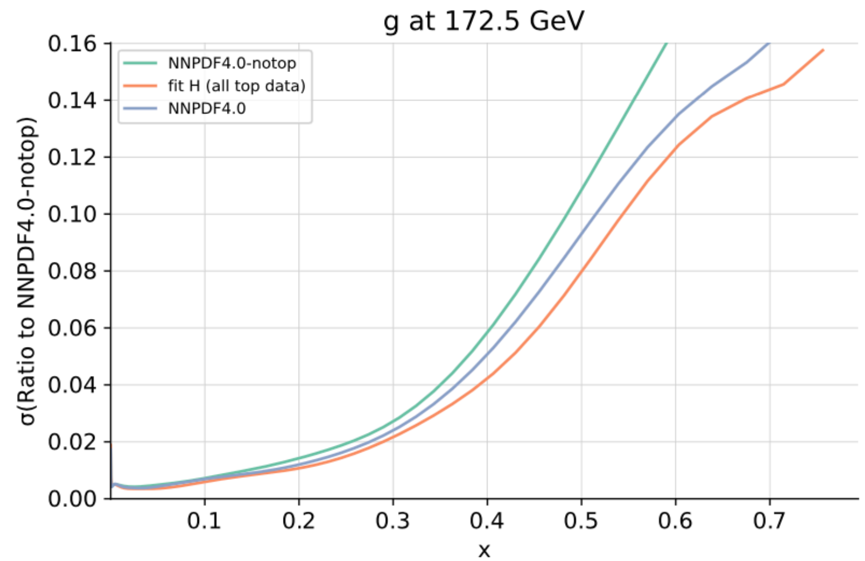


25 (21) dim-6 operators at the quadratic (linear) SMEFT

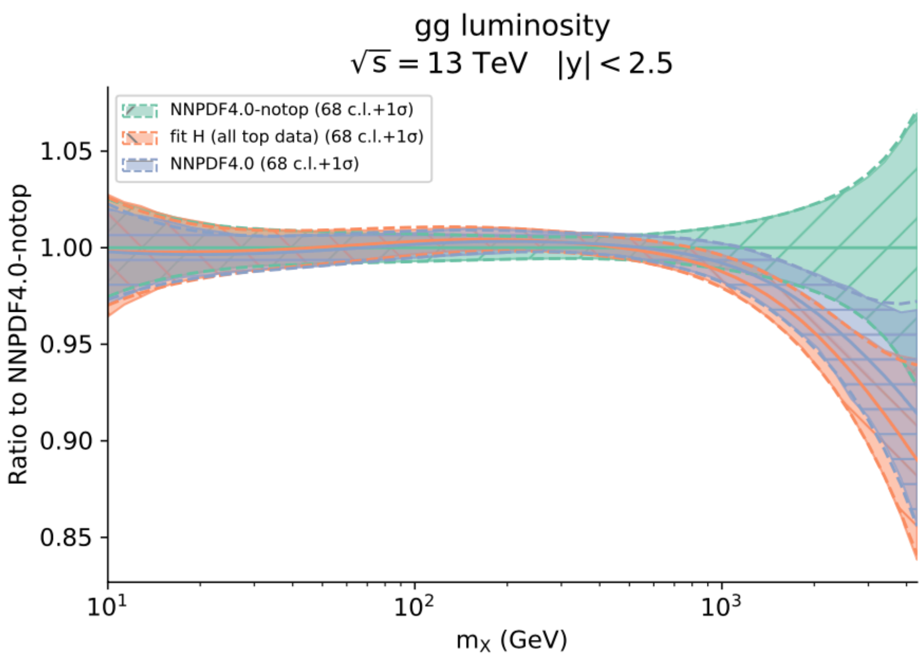
Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU - arXiv: 2303.06159

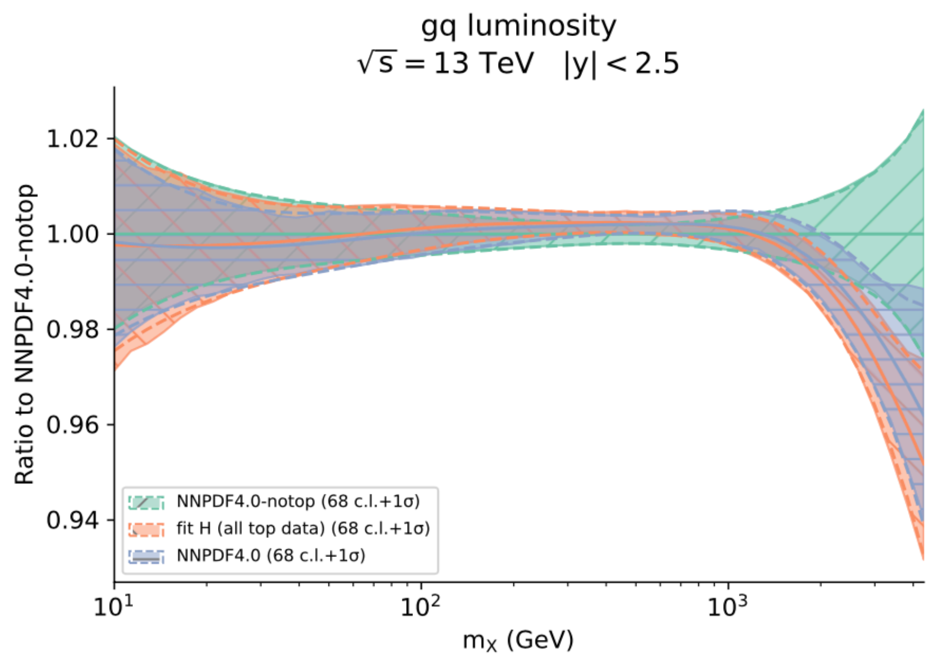
PDF-ONLY FIT





78 new datapoints compared to NNPDF4.0



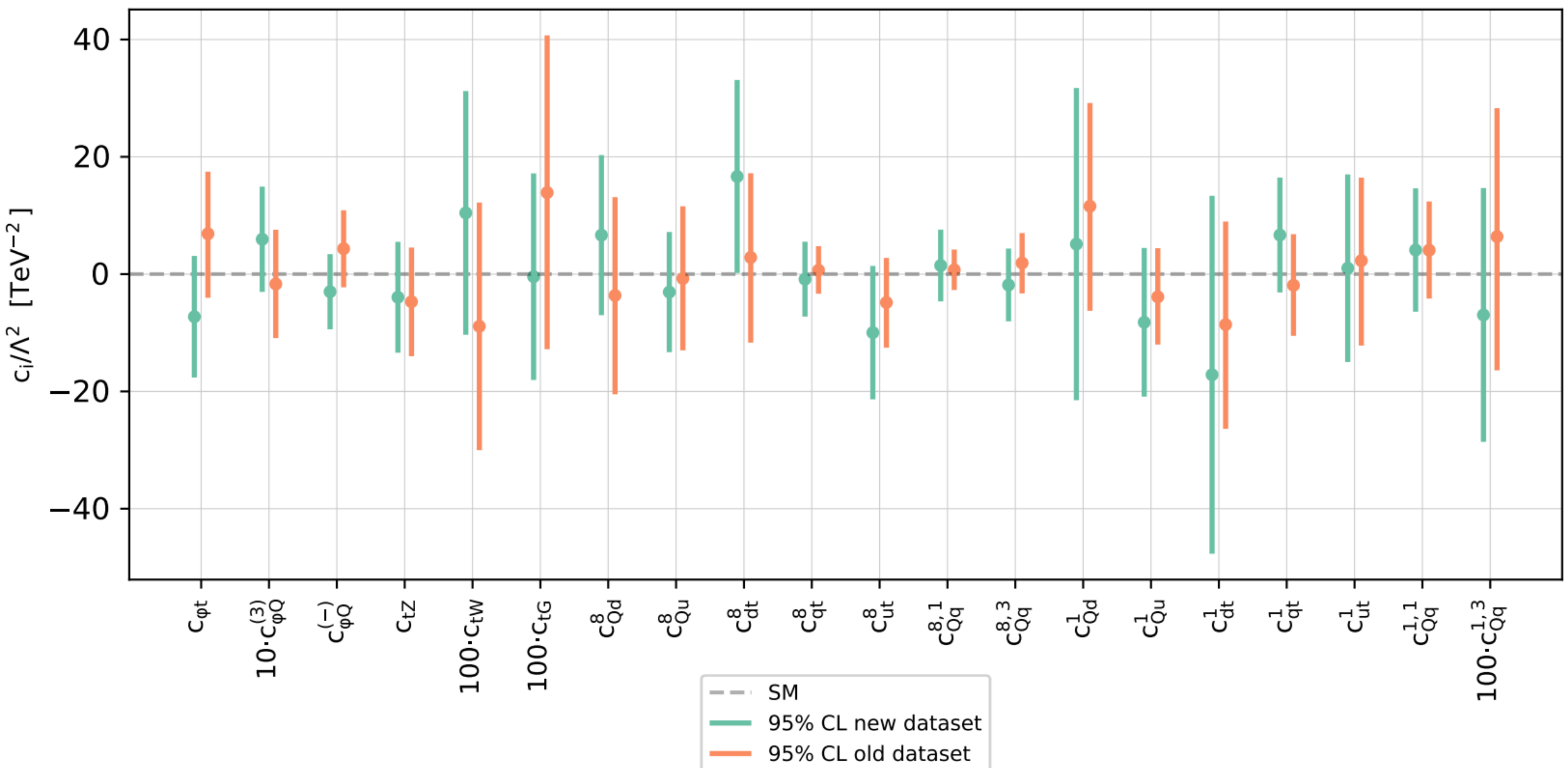


Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU - arXiv: 2303.06159

SMELT-ONLY FIT

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

Linear SMEFT

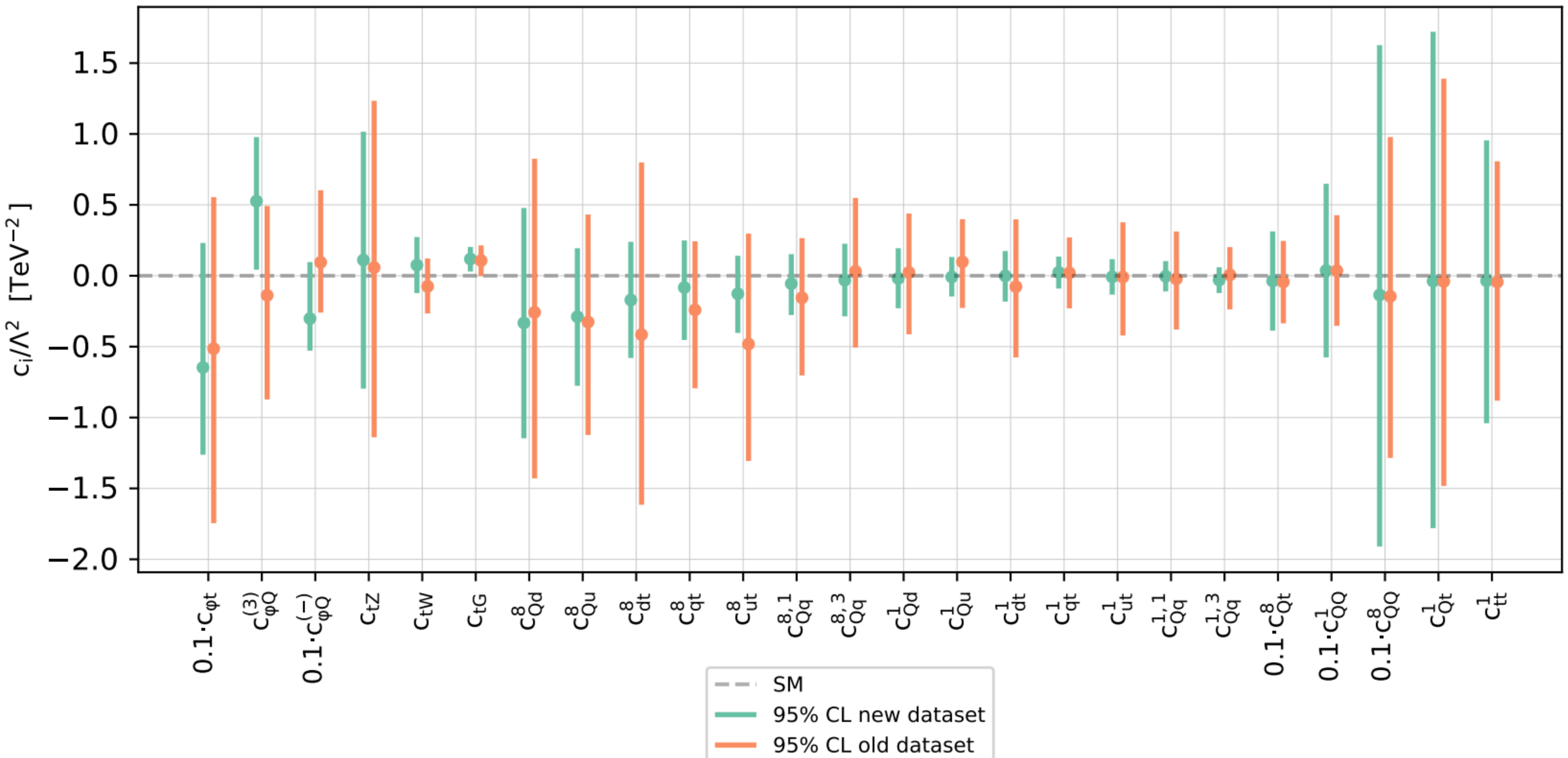


Z. Kassabov, M. Madigan, L. Mantani, J. Moore, M. Morales, J. Rojo, MU - arXiv: 2303.06159

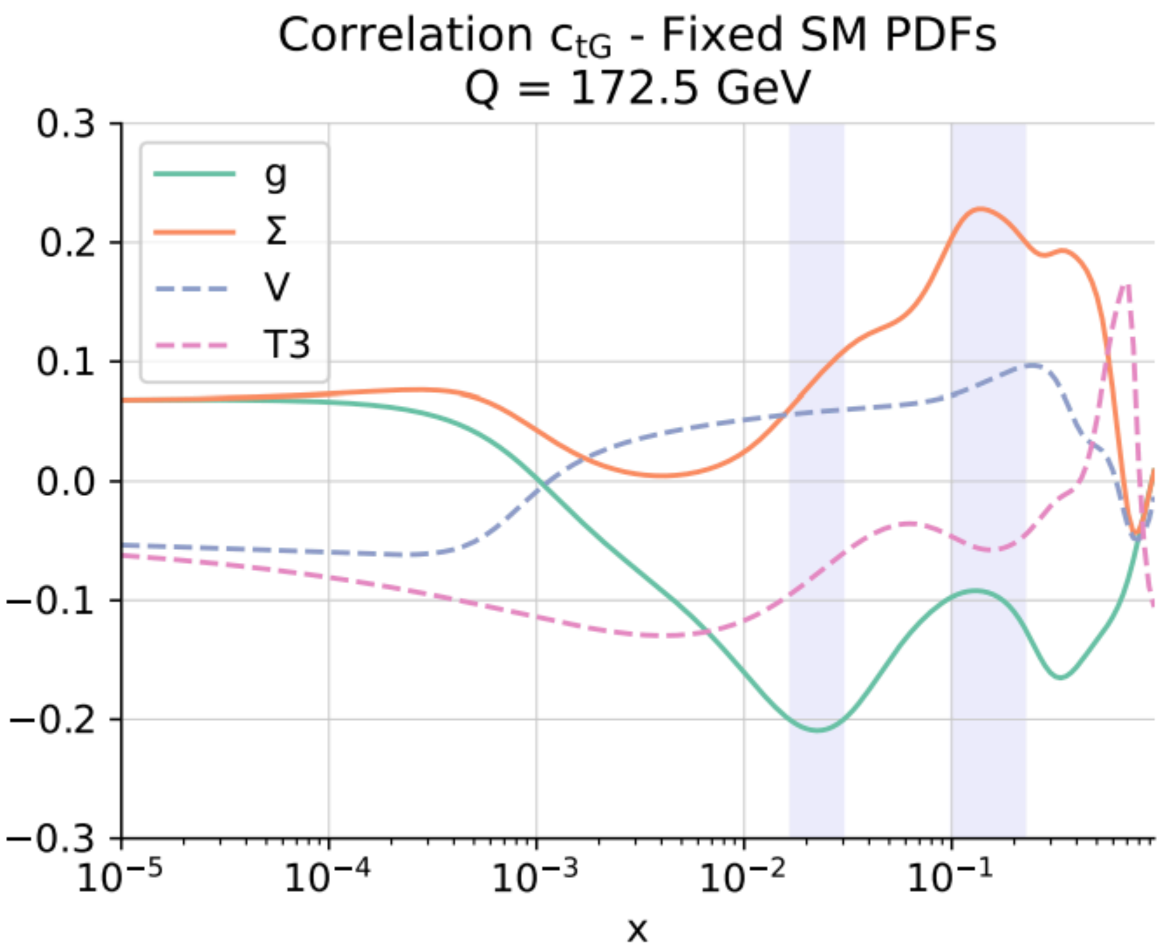
SMEFT-ONLY FIT

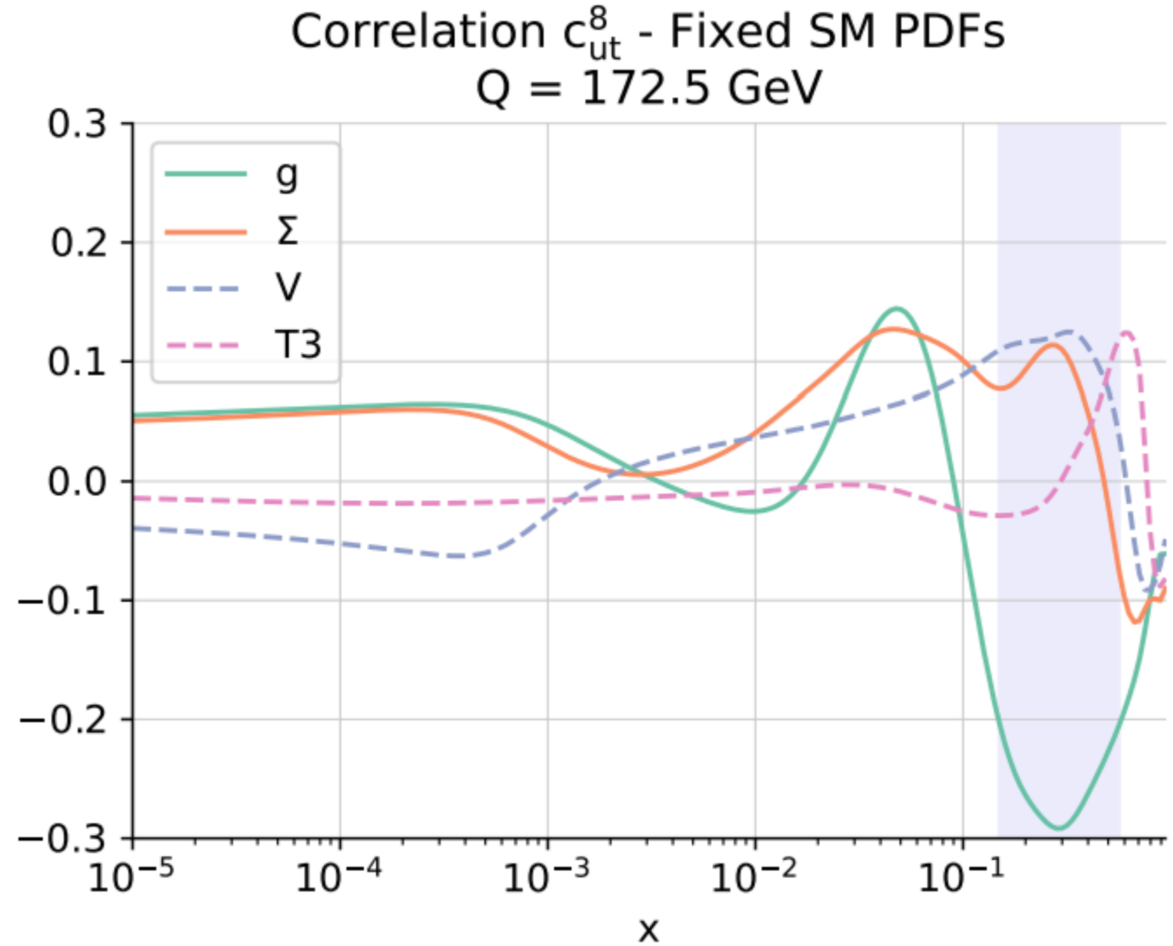
$$\sigma_{\rm eft}\left(\boldsymbol{c}/\Lambda^2\right) = \sigma_{\rm SM} + \sum_{i=1}^{n_{\rm op}} \sigma_{{\rm eft},i}\frac{c_i}{\Lambda^2} + \sum_{i,j=1}^{n_{\rm op}} \sigma_{{\rm eft},ij}\frac{c_i c_j}{\Lambda^4}$$

Quadratic SMEFT

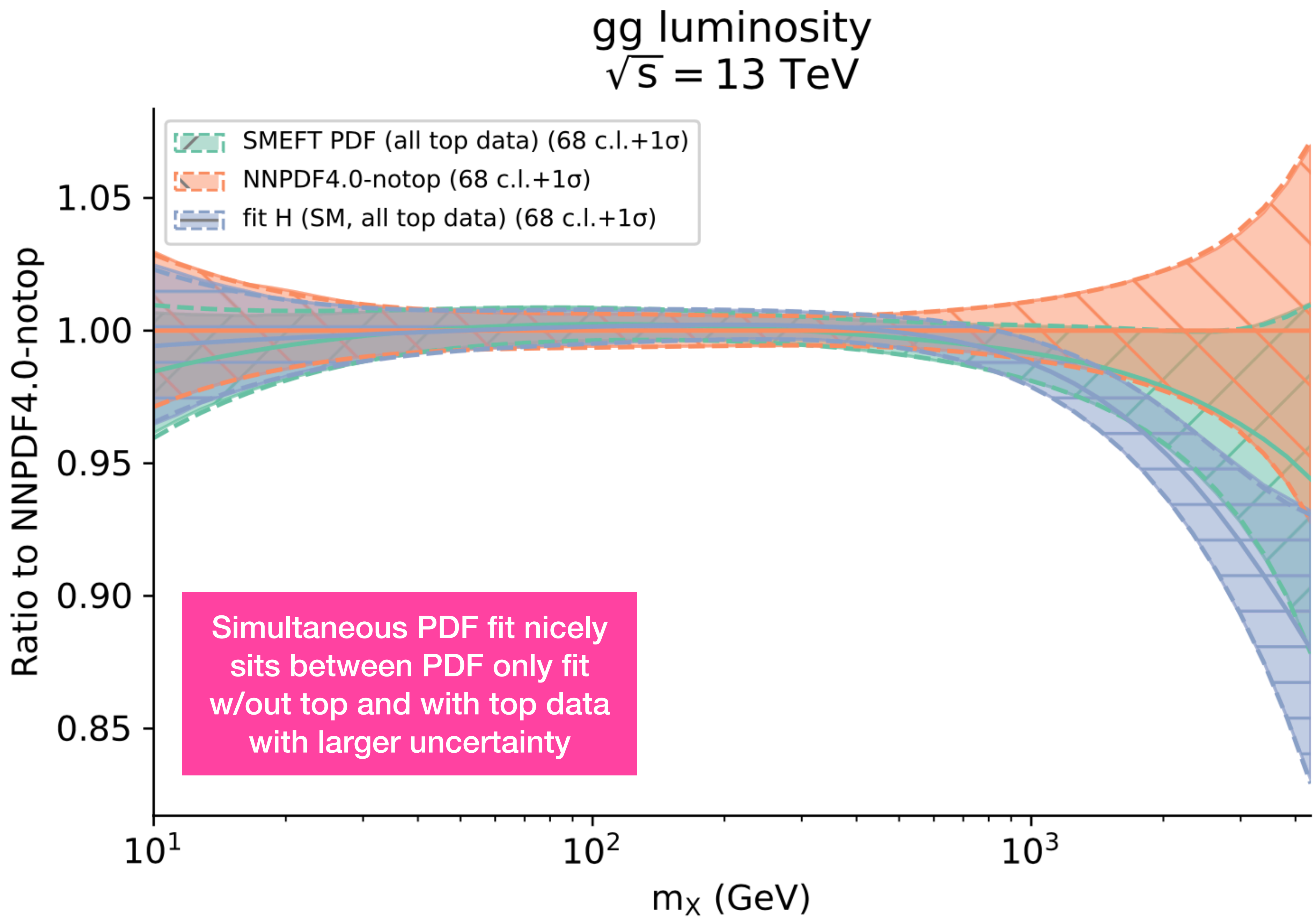


SMEFT AND PDF CORRELATIONS





SIMULTANEOUS PDF-SMEFT FIT



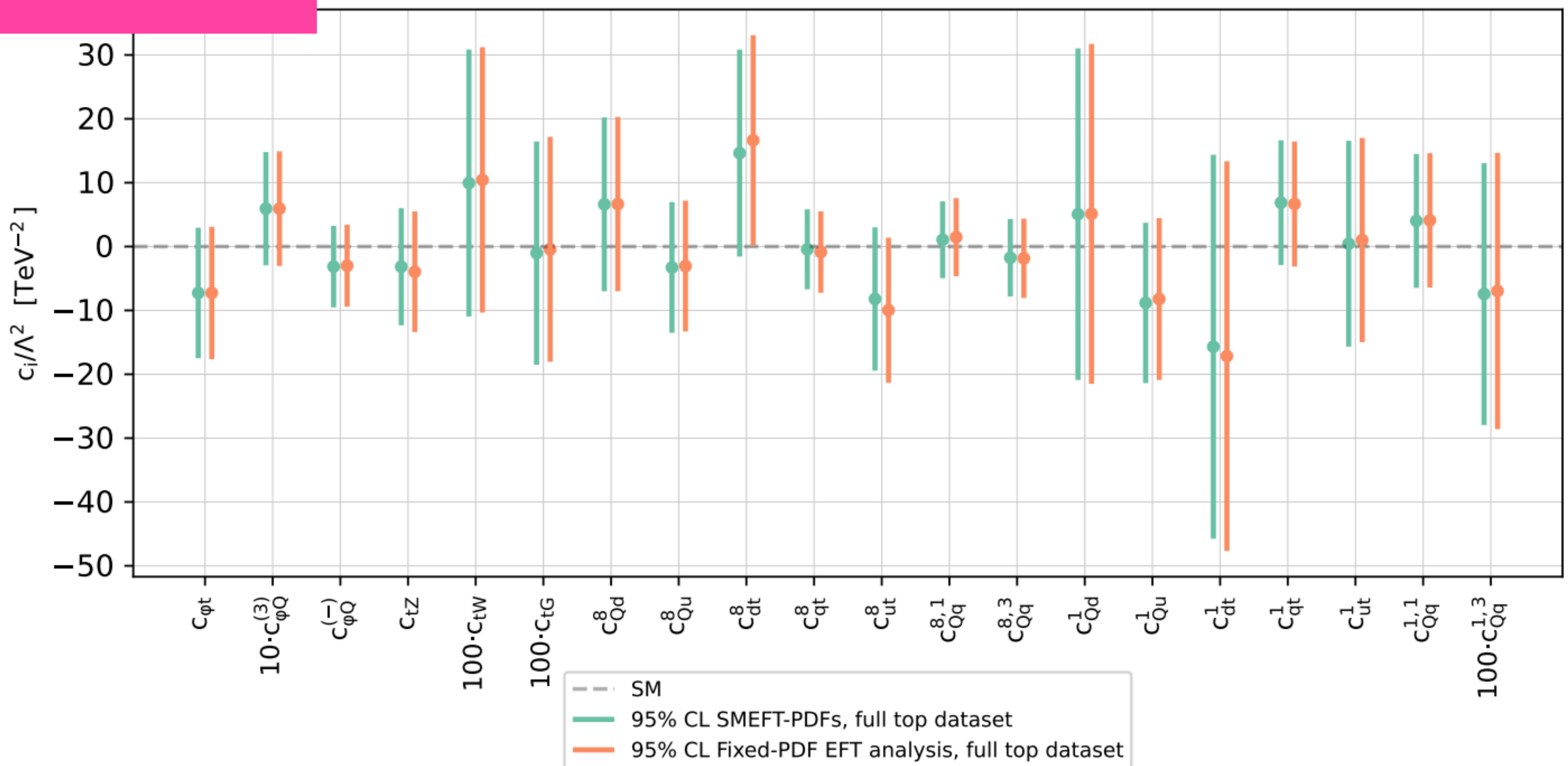
Linear SMEFT

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

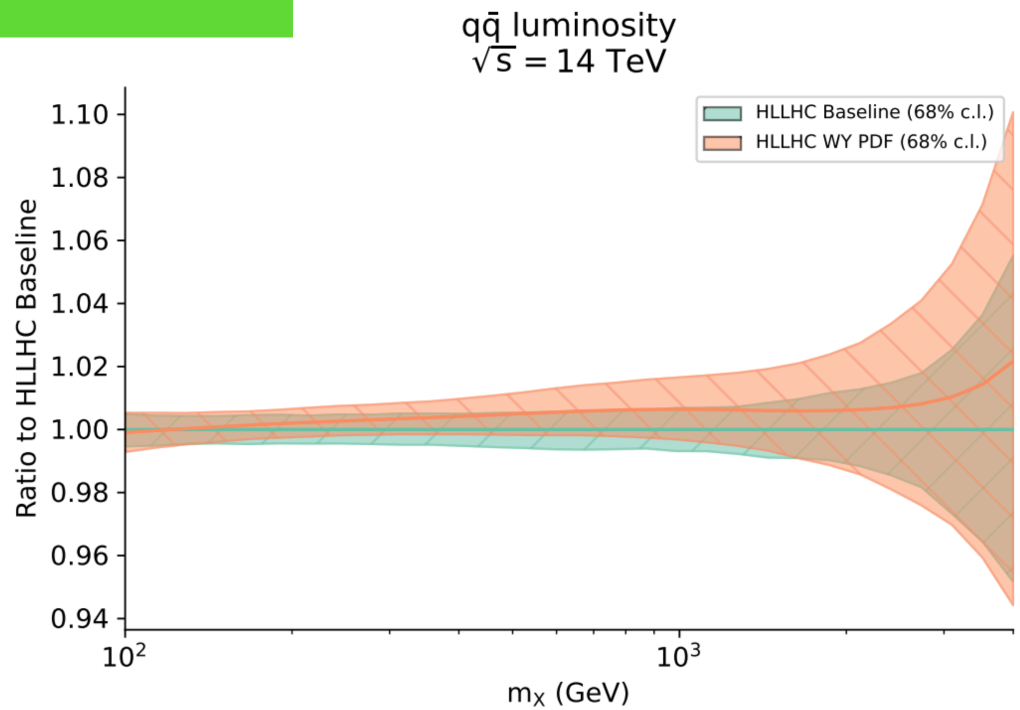
SIMULTANEOUS PDF-SMEFT FIT

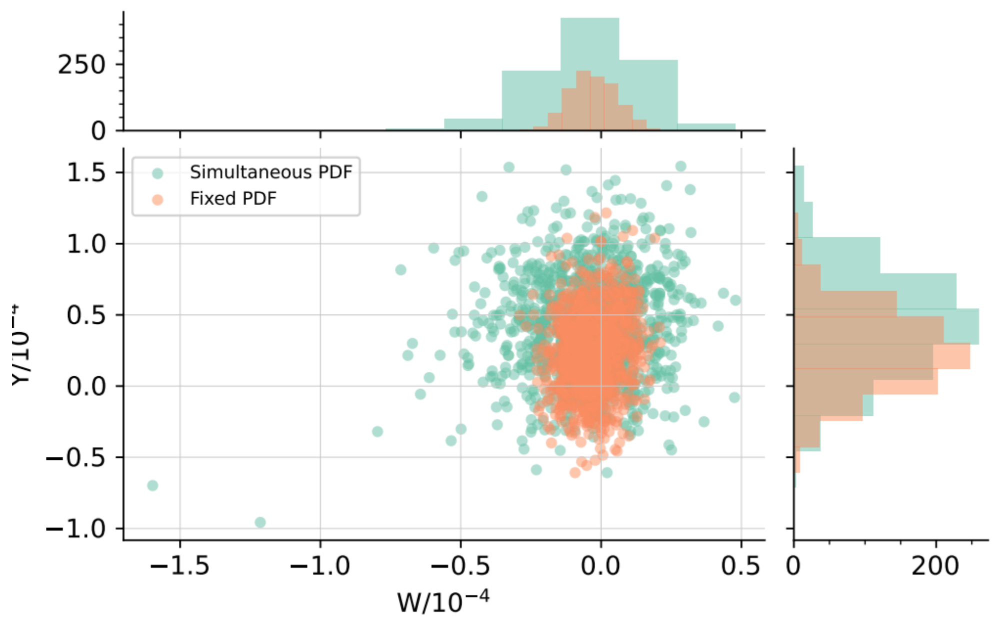
WCs stable upon PDF variations

Linear SMEFT

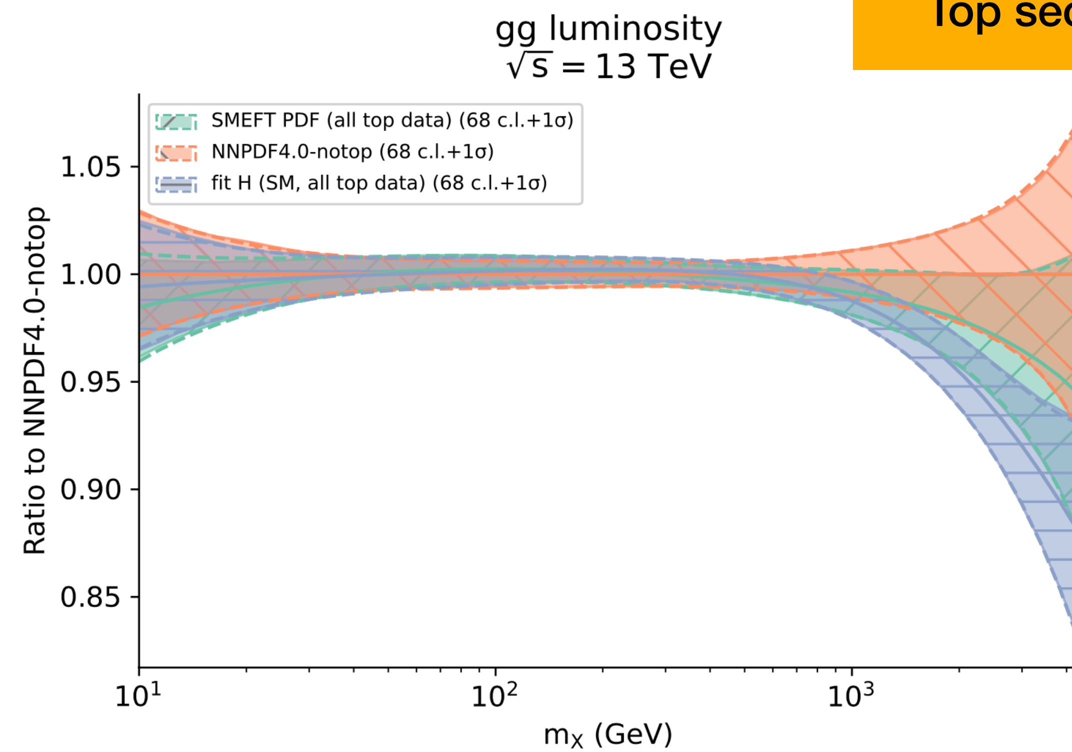


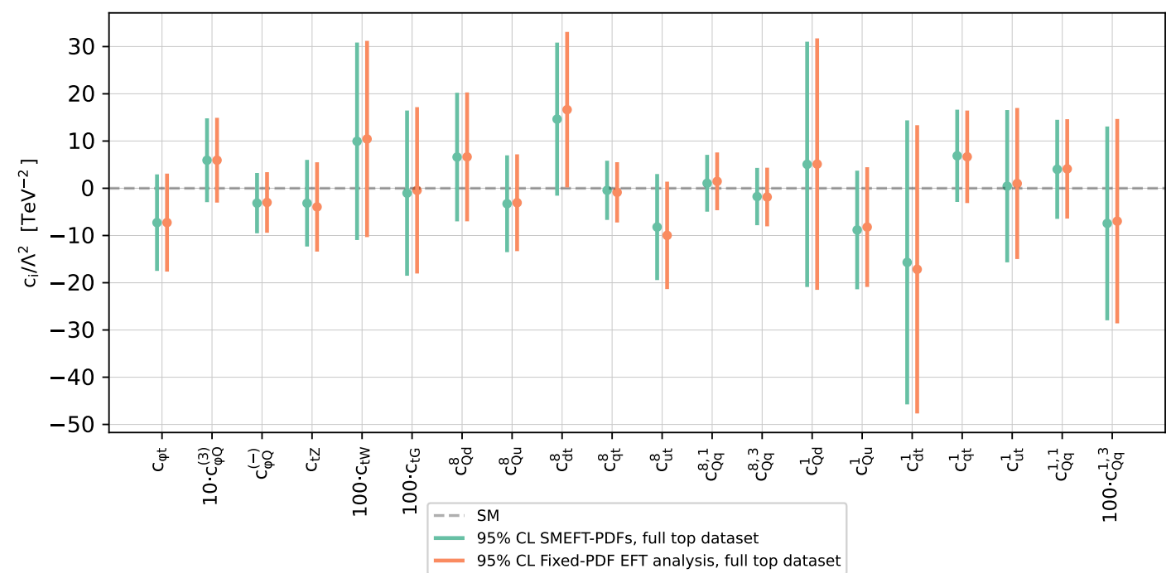
DY sector





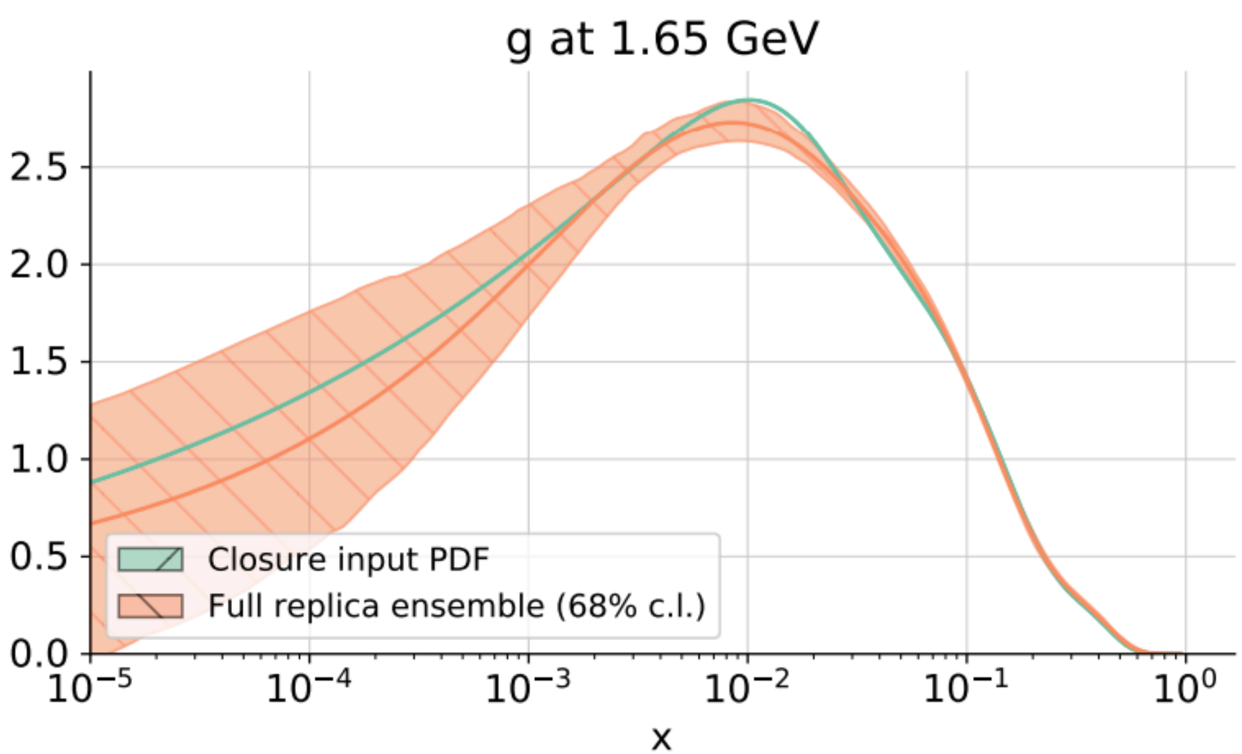
Top sector





CONTAMINATION

CAN PDFS ABSORB NEW PHYSICS?

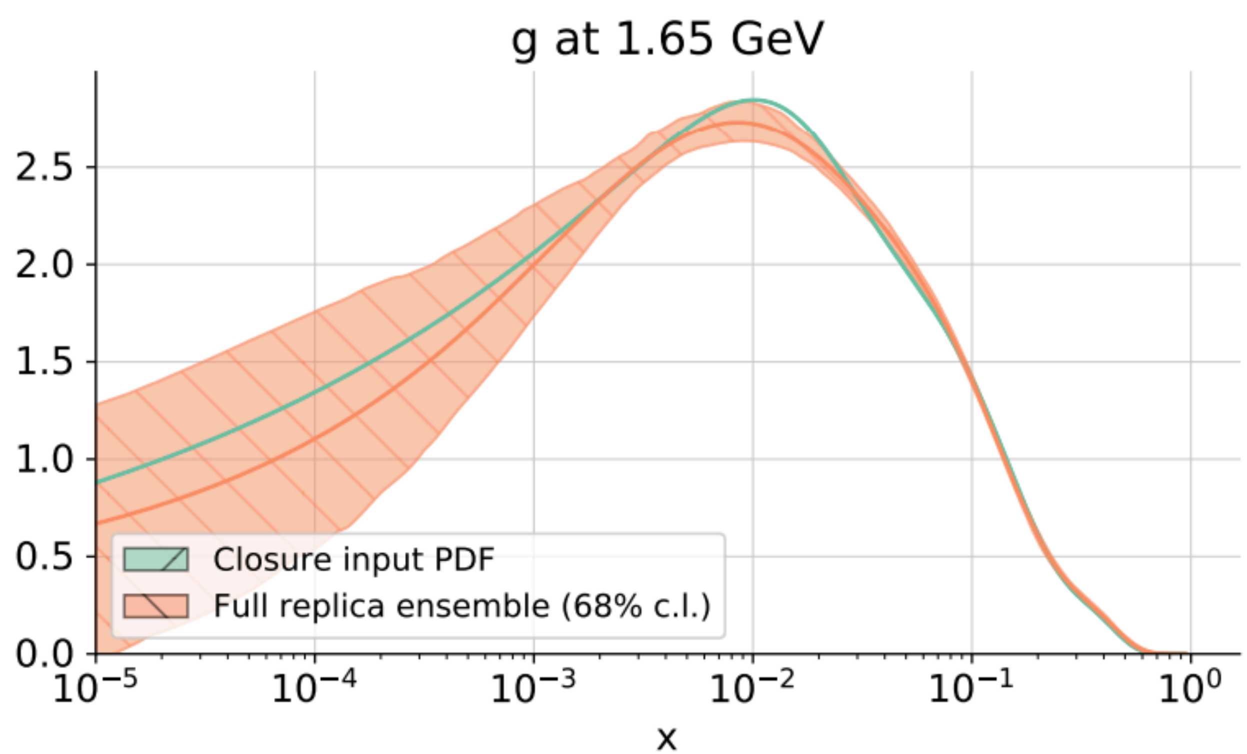


✓ NNPDF methodology routinely tested via closure test (in the data region) [Del Debbio, Giani, Wilson, 2111.05787] and future test (in the extrapolation region) [Cruz-Martinez, Forte, Nocera, 2103.08606].

✓ 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?

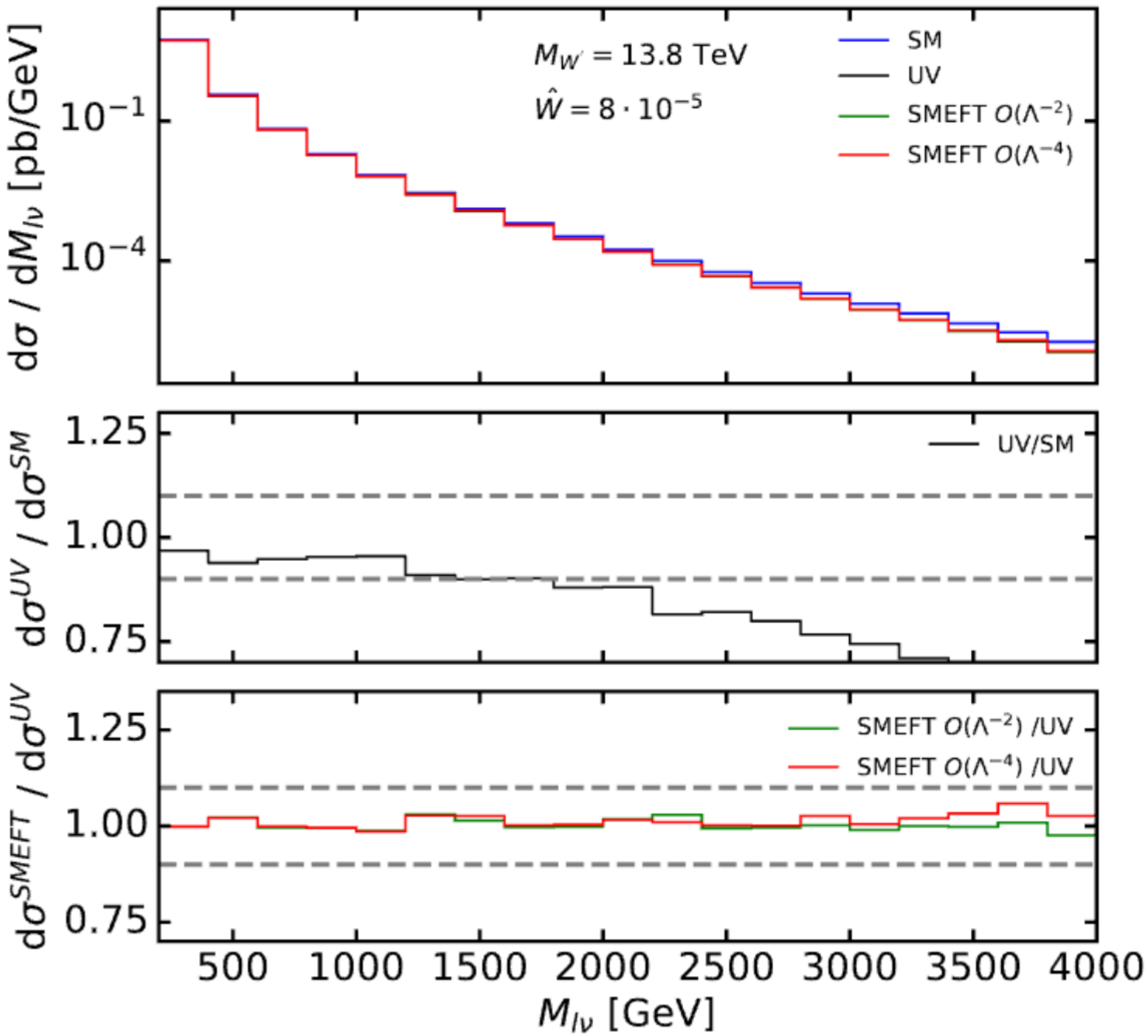
E. Hammou, Z. Kassabov, M. Madigan, M. Mangano, L. Mantani, J. Moore, M. Morales, MU 2307.xxxxx

✓ 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

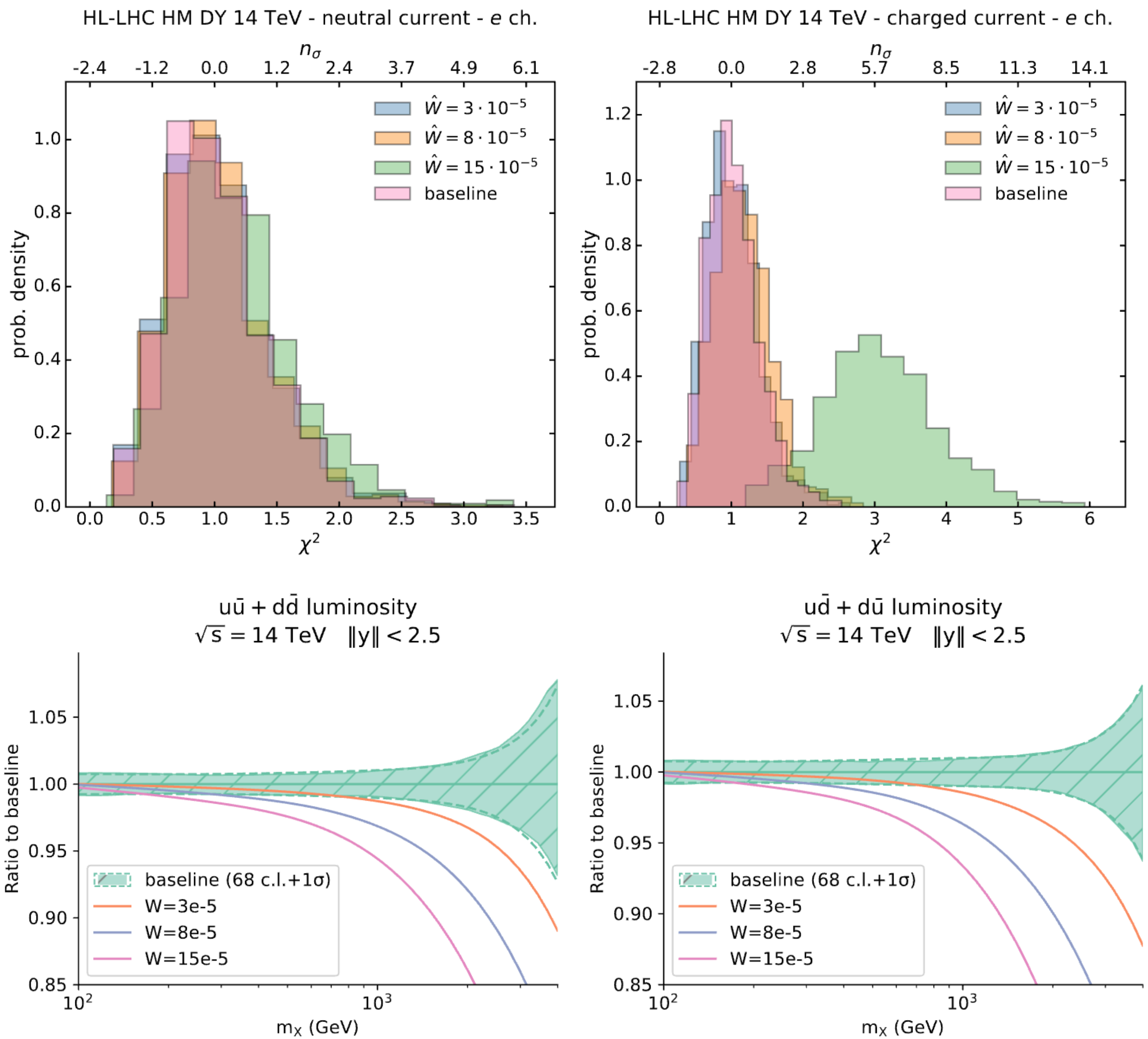


- Imagine that on top of the "true" PDFs one inject the "true" UV model in the MC data
- Generate artificial MC data assuming "true" law of nature =
  "true" PDFs + "true" UV model
- Fit PDFs assuming SM
- Can PDFs absorb signs of new physics?

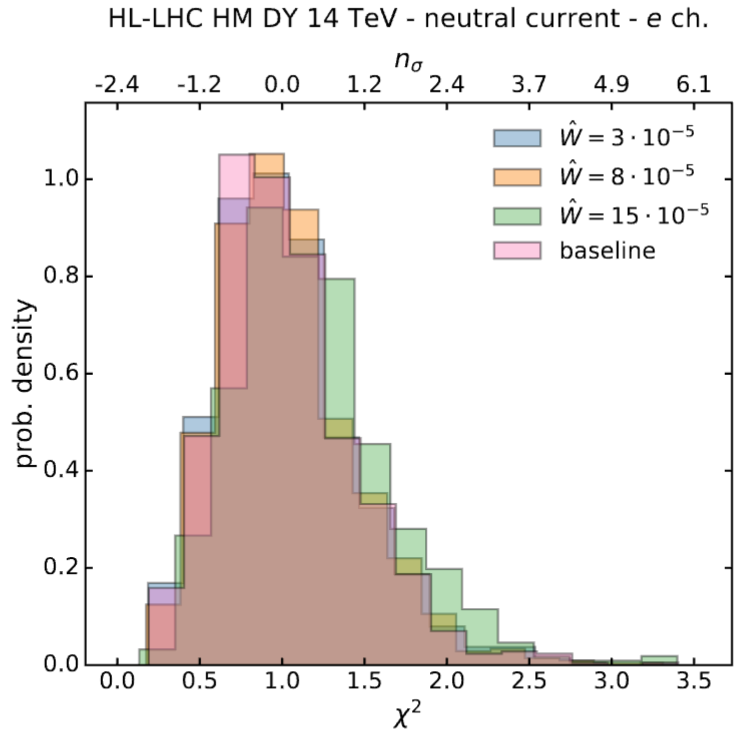
E. Hammou, Z. Kassabov, M. Madigan,
M. Mangano, L. Mantani, J. Moore, M. Morales, MU
2307.xxxxx

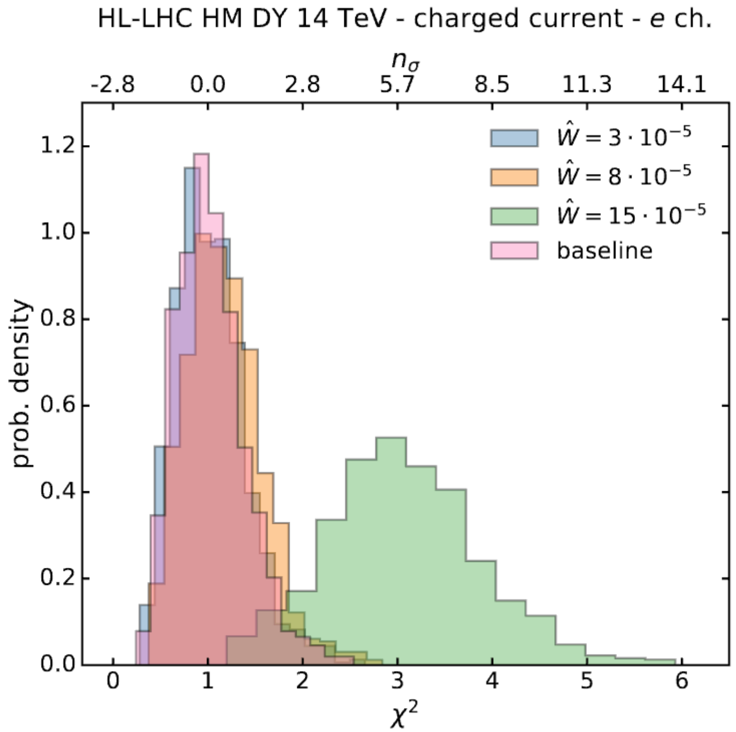
Z'	W'
Selection test: ✗	Selection test: ✓
➡ Excluded from PDF fit	➡ Included in PDF fit
No impact on PDFs	**PDFs contaminated**

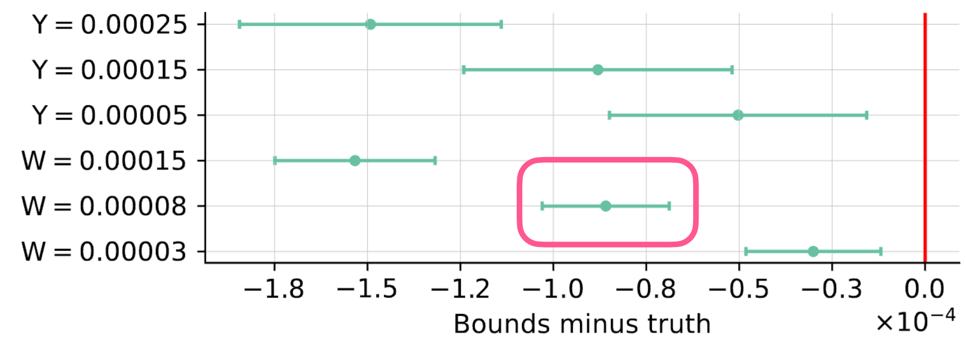
THE W' TEST-CASE

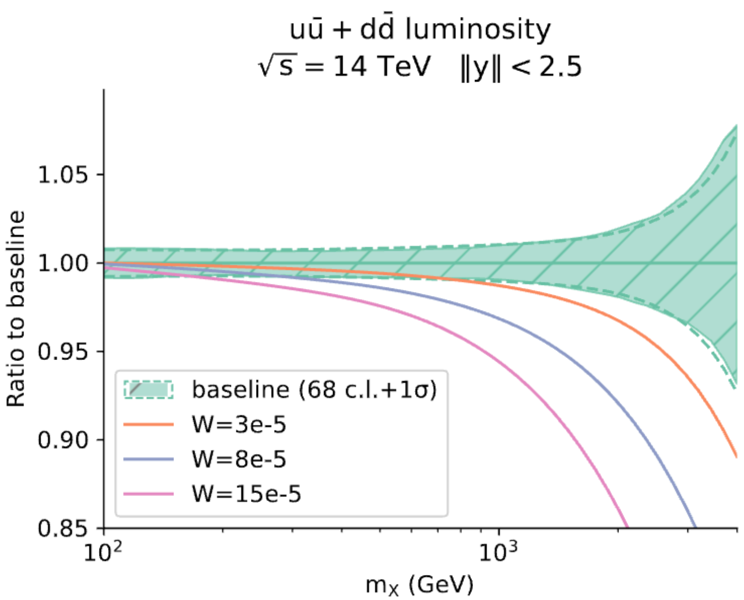


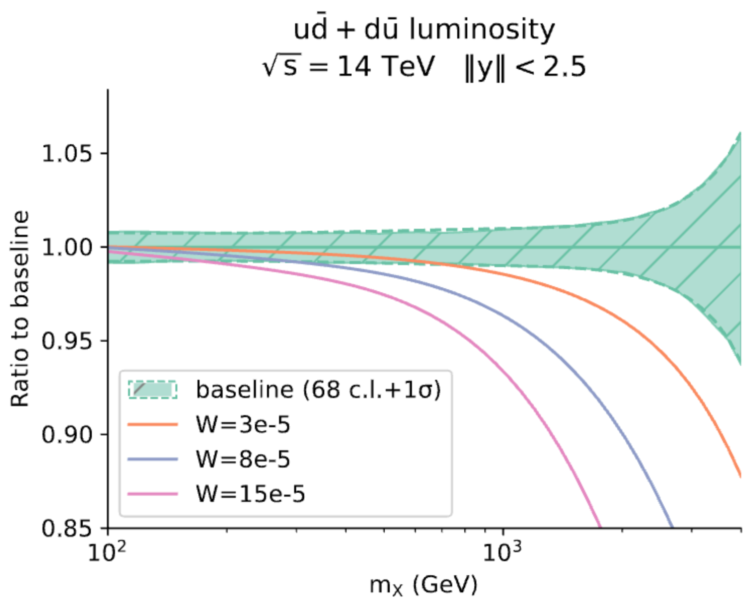
- The fit-quality of the closure test is unchanged up to W= 8e-5 (corresponding to Mw' = 13.8 TeV)
- Once we go beyond this point ,the fit-quality deteriorates due to the HL-LHC neutral current and charged current Drell-Yan MC data.
- Already for W= 8e-5 the qq~ luminosity shifts far beyond the PDF uncertainties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