PDFs and SMEFT

A study of the interplay of Parton Distribution Functions (PDFs) and BSM signals in global fits

Work with Maria Ubiali and her group:

[PBSP, 2307.10370, JHEP]

[PBSP, 2402.03308]

[Hammou et Ubiali, 2410.00963]

[PBSP, Forthcoming]

European Research Council

Established by the European Commission

PBSP

Funded by
the European Union

Elie Hammou, University of Cambridge

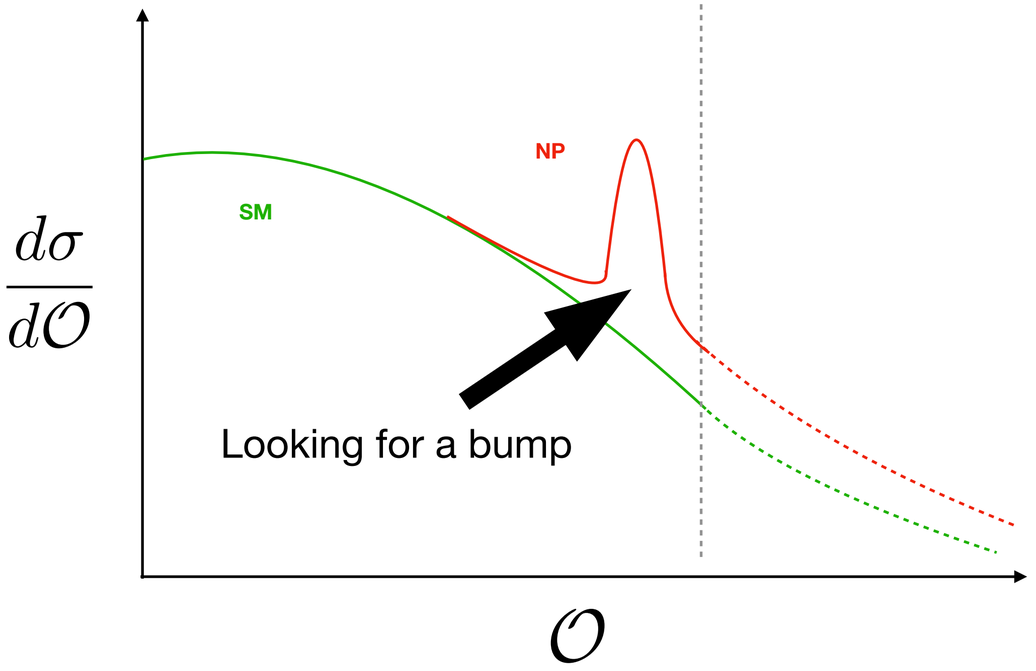
EFTs and Beyond workshop, Dec 2024, MITP Youngst@rs

New physics searches

Looking toward higher energy scales and indirect searches

Direct searches

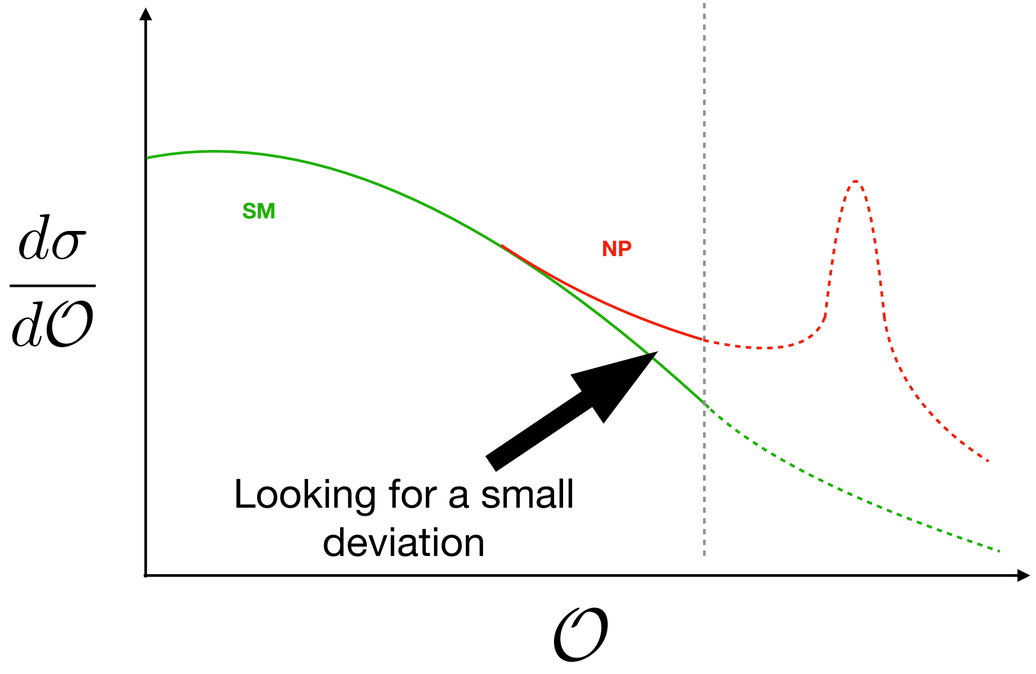
$E_{NP} < E_{collider}$



No luck so far…

Indirect searches

$E_{NP} > E_{collider}$

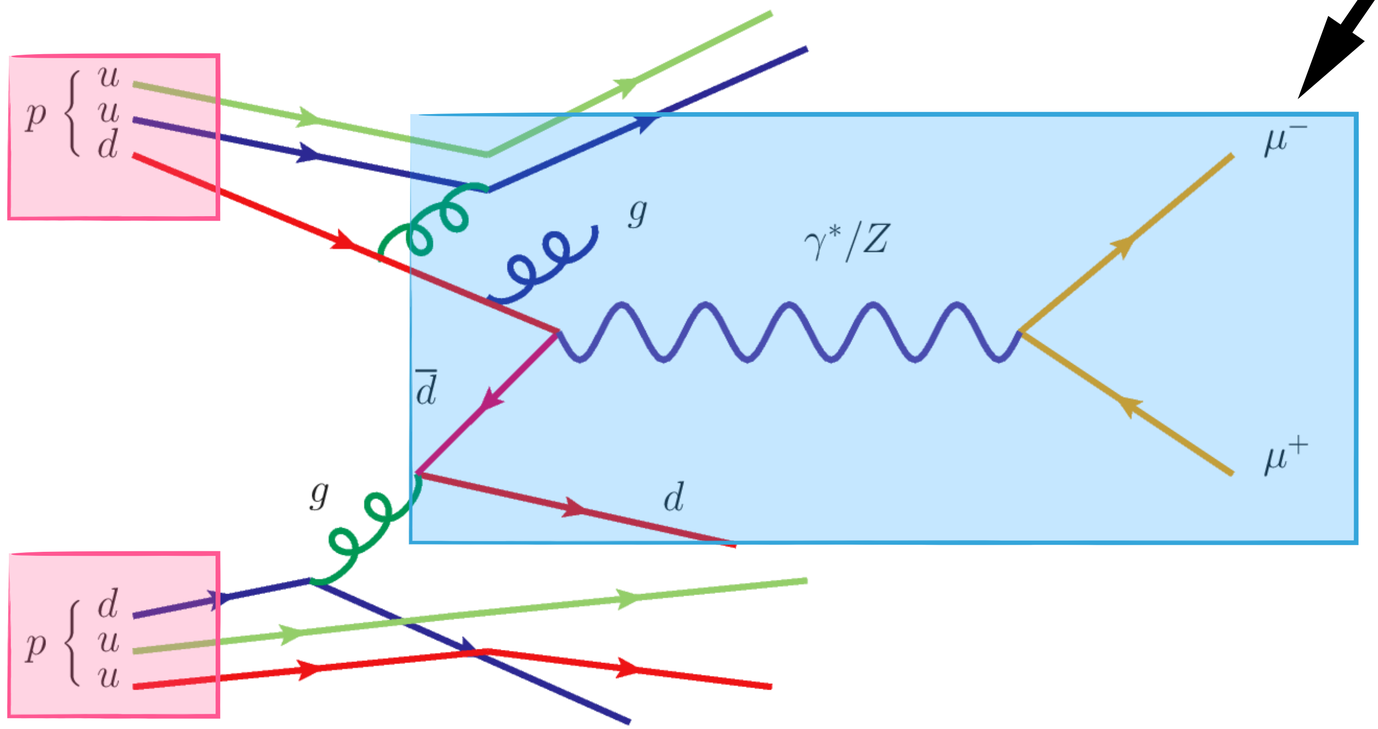


Requires precision

Hadron colliders and PDFs

Collinear factorization theorem

SMEFT corrections



Parton Distribution Functions

Hard Scattering: Perturbative QCD + EW

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

Background on Parton Distribution Functions

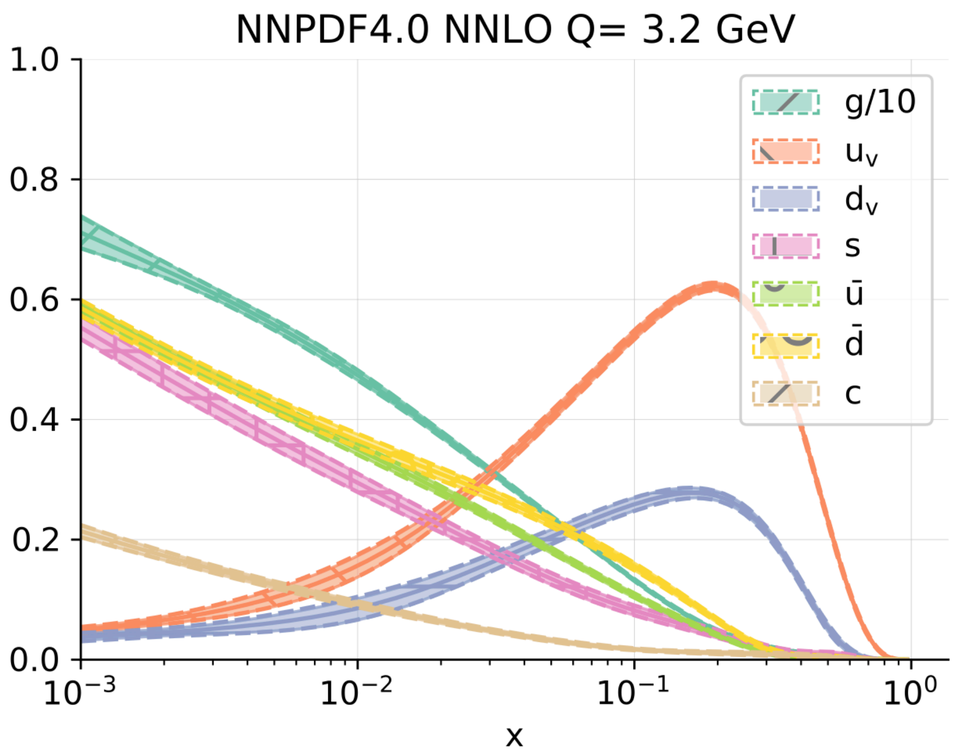
Hadron collider observable: $\sigma = \hat{\sigma} \otimes f_1 \otimes f_2$

PDFs in a nutshell:

- describe proton's partonic content
- $f(x, Q)$
- x dependance: non-perturbative QCD

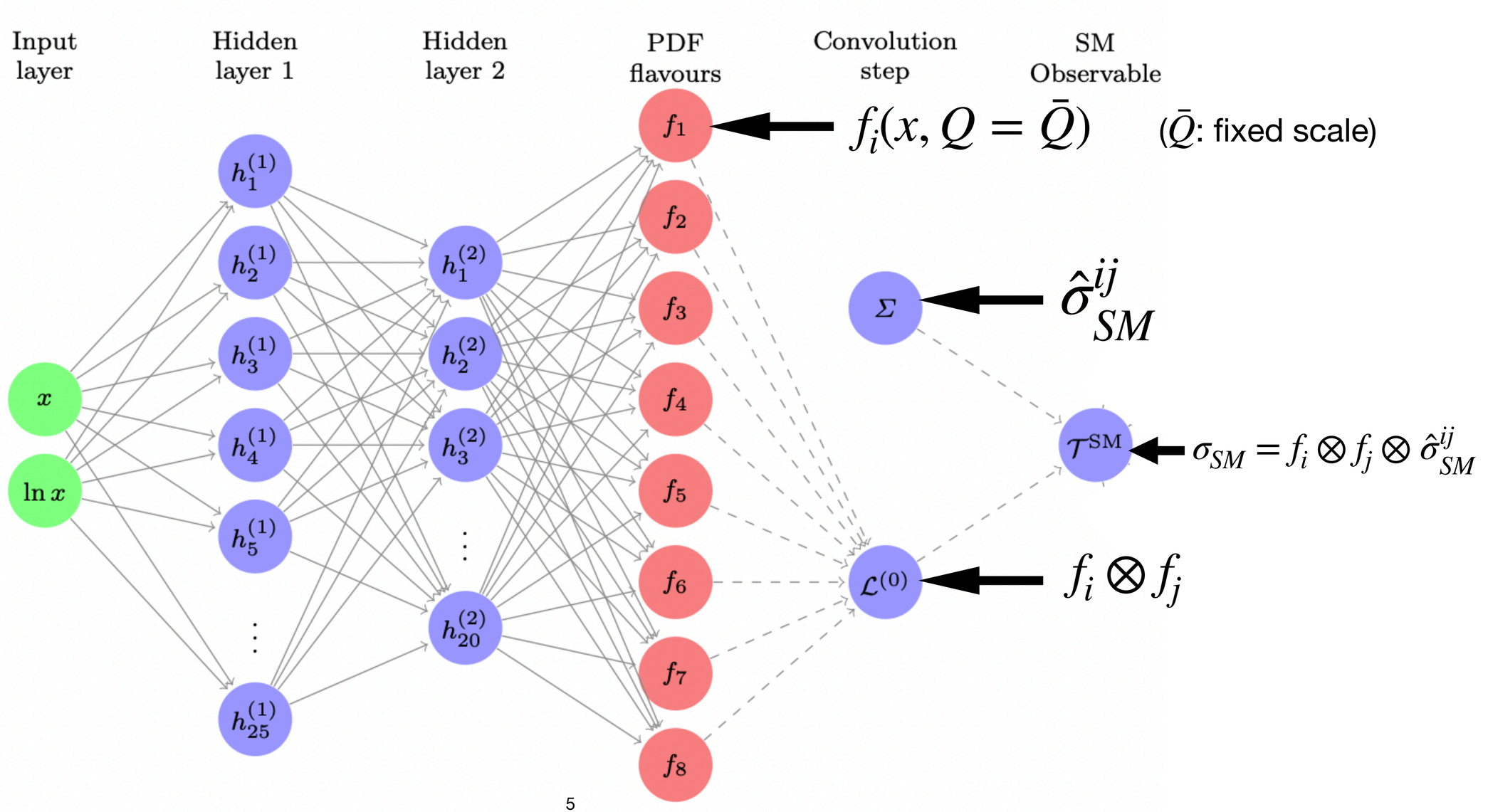
➡ **Fitted from data**

Using NNPDF methodology



[Ball et al., NNPDF4.0, 2109.02653]

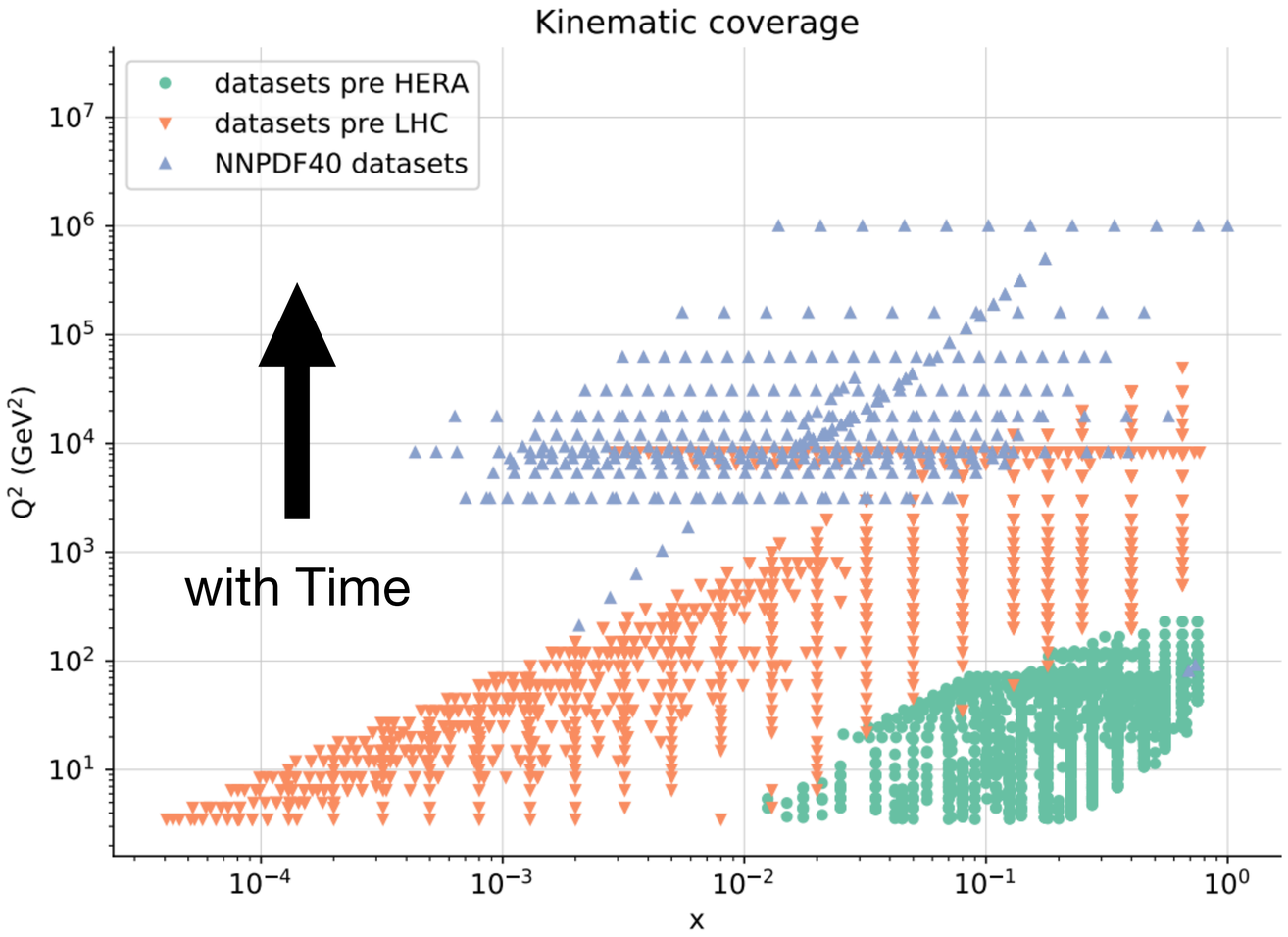
NNPDF methodology: fitting PDFs



Bjorken-x
(Momentum fraction)

Fitting PDF from data

The dataset used by NNPDF



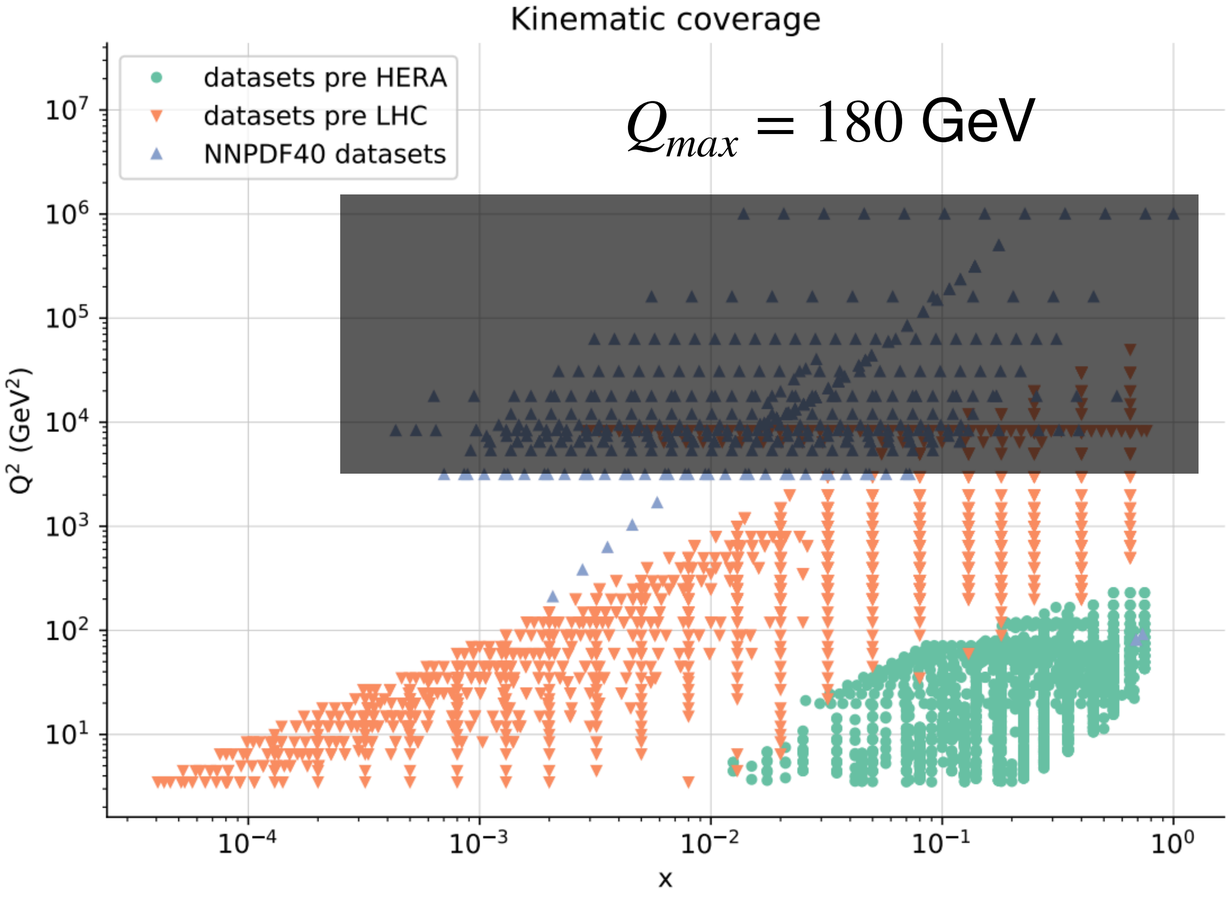
Evolution of the dataset through time:

- Moved toward higher energies
- 30% is LHC data
- More to come with HL-LHC run

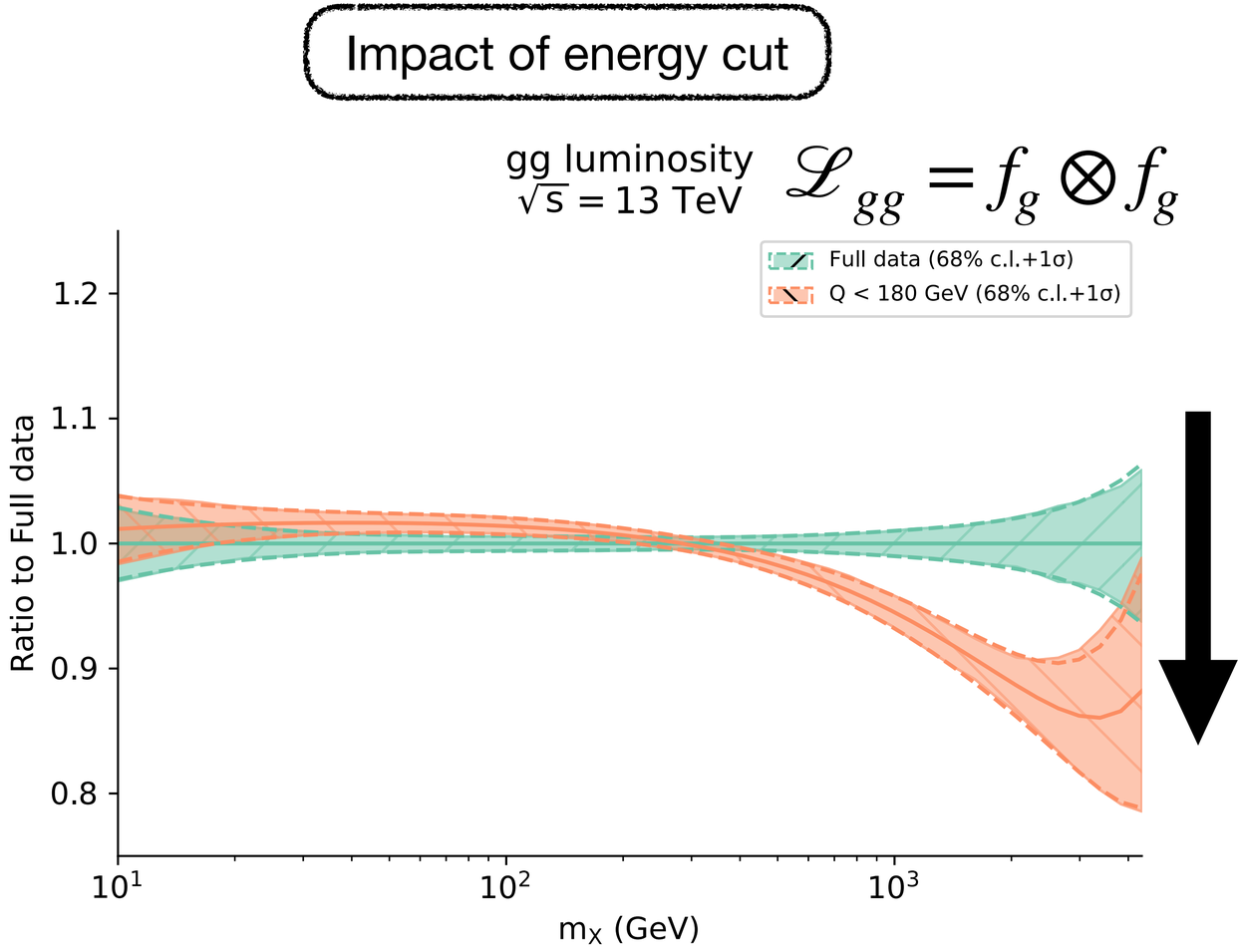
[Ball et al., NNPDF4.0, 2109.02653]

Discrepancy between low and high-energy data fits

Comparison of full data and no LHC PDF fit



Impact of energy cut



Risk of absorbing new physics in PDFs?

Methodology for risk assessment

Perform a “Contamination test”:

1. Choose a BSM model and a “true PDF” set
2. Produce BSM pseudodata
3. Fit PDFs on pseudodata assuming SM
4. Compare results with baseline PDFs (no BSM physics)

[2307.10370]

Contamination criteria:

- Incompatible with baseline
- Fit quality does not deteriorate

➡ $\chi^2 = (Dat - Th)^\top \cdot \Sigma_{cov}^{-1} \cdot (Dat - Th)$

PDF contamination:

➡ **PDFs have absorbed new physics signals**

New physics scenarios: W'

From UV to the SMEFT

Heavy triplet under $SU(2)_L$: W'

$$\mathscr{L}_{UV}^{W'} = \mathscr{L}_{SM} - \frac{1}{4} W'^{a}_{\mu\nu} W'^{a,\mu\nu} + \frac{1}{2} M^2_{W'} W'^{a}_{\mu} W'^{a,\mu} - g_{W'} W'^{a,\mu} \sum_{f_L} \bar{f}_L T^a \gamma^\mu f_L - g_{W'} (W'^{a,\mu} \varphi^\dagger T^a i D_\mu \varphi + \text{h.c.})$$

➡ Creates two charged particles: W'^{+} / W'^{-} and a neutral one: W'_3

Matching to the SMEFT:

$$\mathscr{L}_{SMEFT}^{W'} = \mathscr{L}_{SM} - \frac{g^2_{W'}}{2M^2_{W'}} J_L^{a,\mu} J^a_{L,\mu} \qquad J_L^{a,\mu} = \sum_{f_L} \bar{f}_L T^a \gamma^\mu f_L$$

→ $$\mathscr{L}_{SMEFT}^{W'} = \mathscr{L}_{SM} - \frac{g^2 \hat{W}}{2m^2_W} J_L^{a,\mu} J^a_{L,\mu} \qquad \hat{W} = \frac{g^2_{W'}}{g^2} \frac{m^2_W}{M^2_{W'}} \propto \frac{c}{\Lambda^2}$$

New physics parameter

New physics scenarios: W' $pp \to l^- \bar{\nu}$ $M_{W'} = 13.8$ TeV

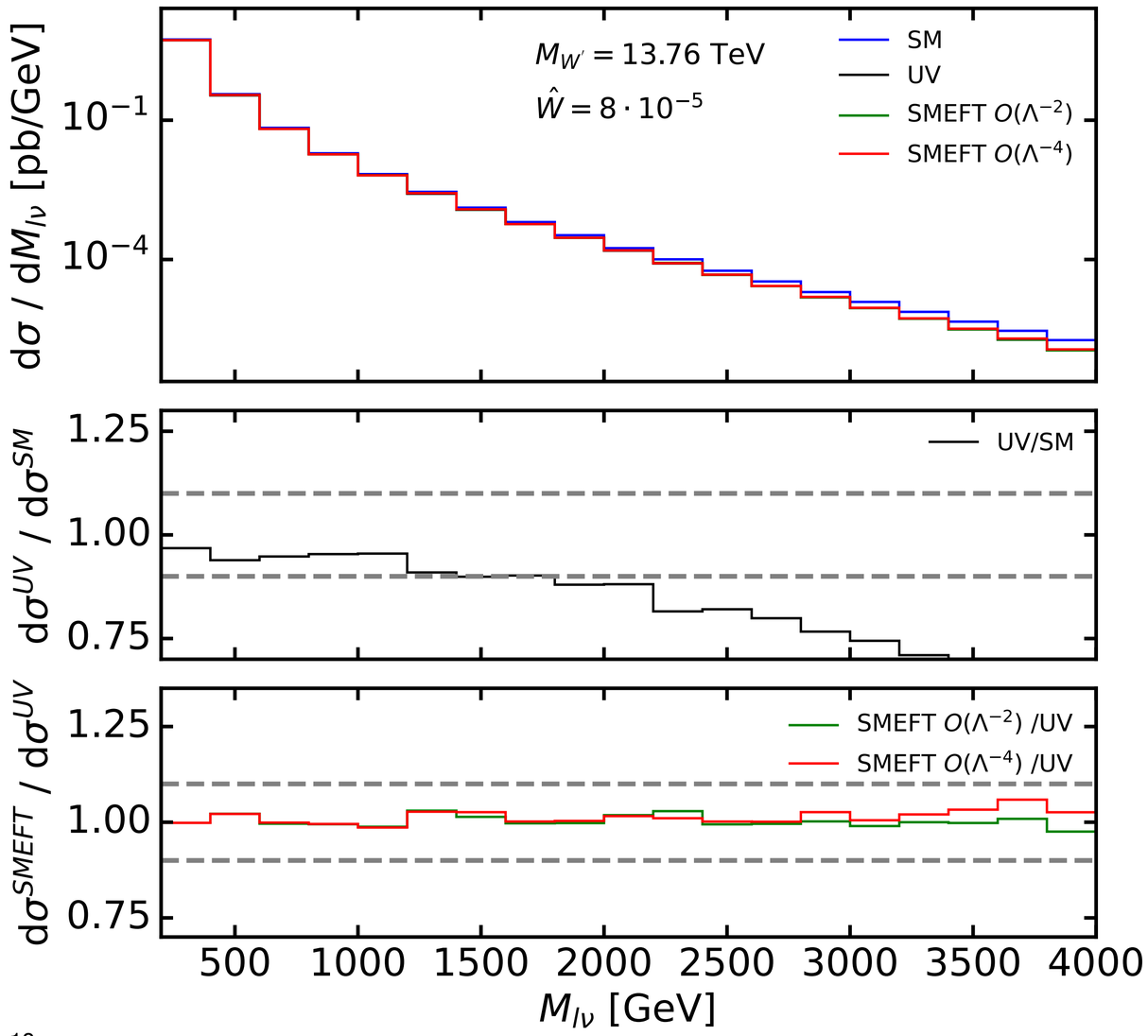
Generation of the pseudodata

$$\hat{W} \leftrightarrow M_{W'} \quad (g_{W'} = 1)$$

$$\mathscr{L}^{W'}_{SMEFT} = \mathscr{L}_{SM} - \frac{g^2 \hat{W}}{2m_W^2} J_L^{a,\mu} J^a_{L,\mu}$$

➡ Impacts CC and NC Drell-Yan

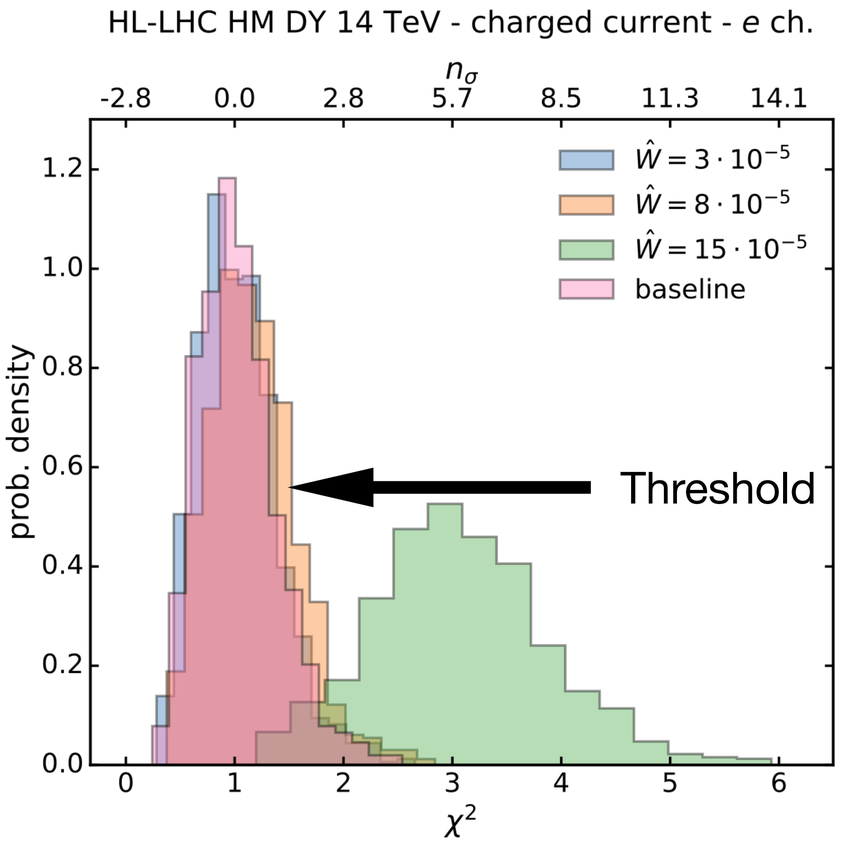
HL-LHC Projections



Impact of contamination on the PDFs

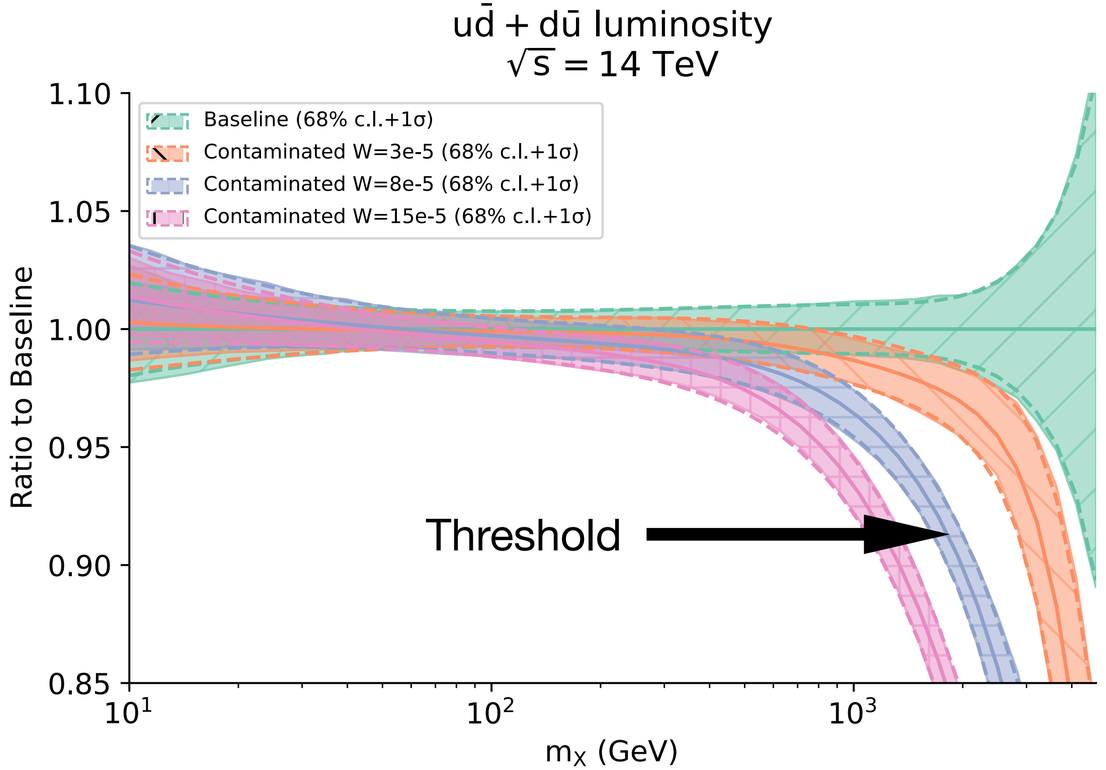
Comparison between contaminated and Baseline PDFs

Fit quality



$M_{W'} = 13.8$ TeV

Impact on PDFs

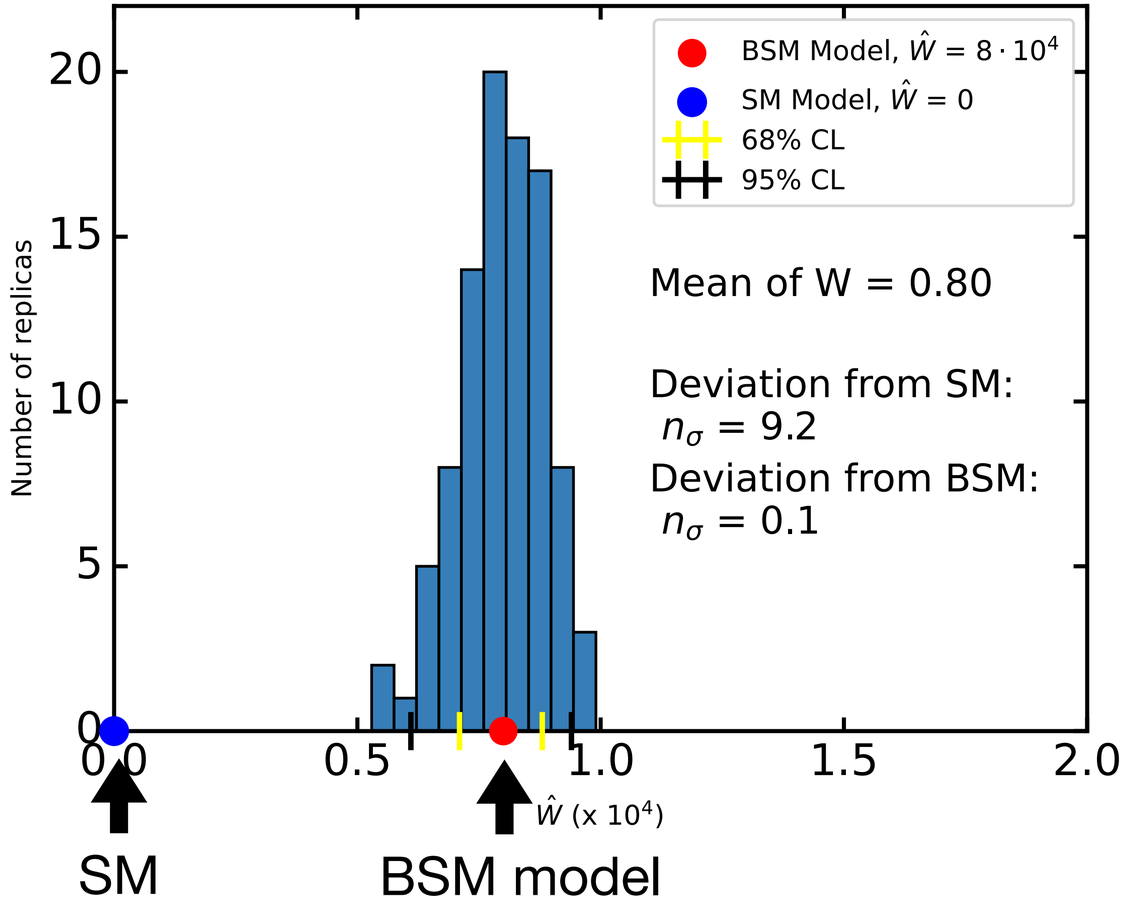


$$\sigma_{BSM}^{Data} \approx \hat{\sigma}_{SM} \otimes \mathscr{L}_{BSM}$$

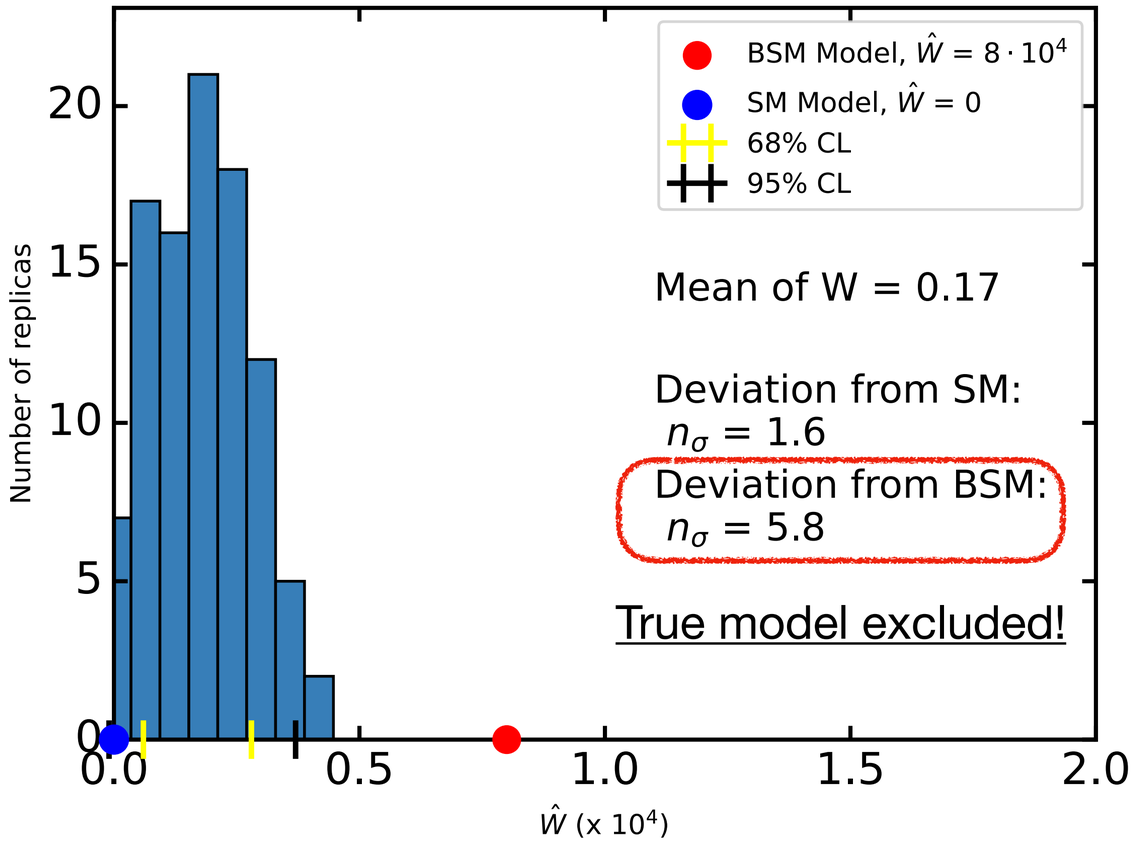
Missing new physics

Impact of the PDF contamination on SMEFT fits

SMEFT Fit with true PDF



SMEFT Fit with SMEFT PDF



Apparition of fake deviations

Impact of contamination on predictions for other sectors

Theory predictions (red band):

- BSM PDFs + SM

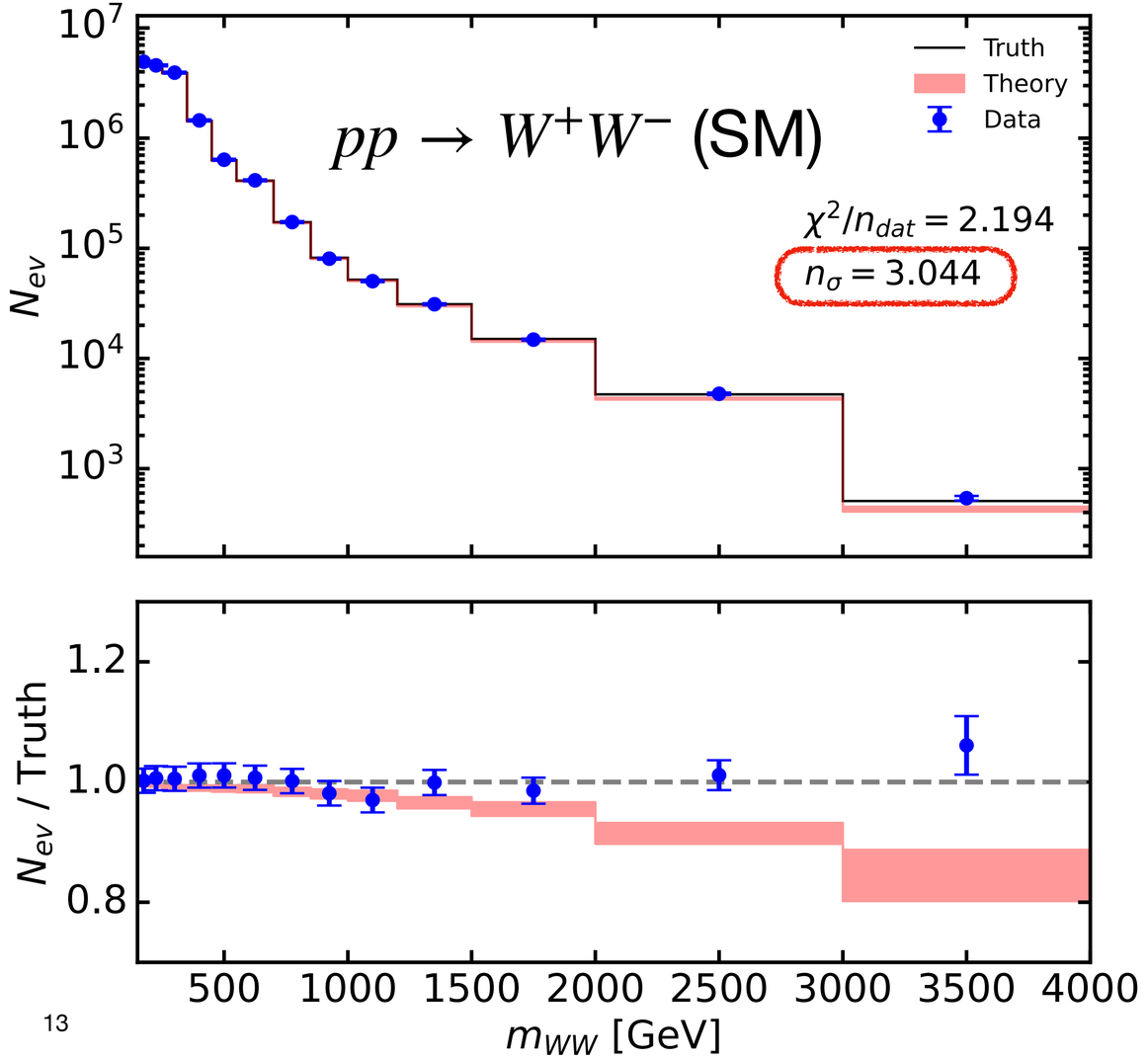
Data (blue dots):

- True PDFs + SM

➡ Fake deviation from SM

Also seen in WH, WZ, ZH production

HL-LHC Projections



PDF/BSM signals mixing: summary

- BSM data in PDF fit:
  - At best: BSM data flagged and excluded
  - At worst: BSM signal absorbed by the PDF

- Consequences of PDF contamination:
  - New physics is hidden (model can be ruled out)
  - Introduced fake deviations in other sectors

➡ Possible solution?

Simultaneous fit of PDF and new physics

Separate versus simultaneous fits

Separate fits

PDF fit:

$$T(\{\theta\}, \{c = 0\}) = \text{PDF}(\{\theta\}) \otimes \hat{\sigma}(\{c = 0\})$$

➡ $\bar{\theta}$

Assumes SM: source of bias

SMEFT fit:

$$T(\{\theta = \bar{\theta}\}, \{c\}) = \text{PDF}(\{\theta = \bar{\theta}\}) \otimes \hat{\sigma}(\{c\})$$

➡ $\bar{c}$

Risk of getting wrong value

Simultaneous fits

$$T(\{\theta\}, \{c\}) = \text{PDF}(\{\theta\}) \otimes \hat{\sigma}(\{c\})$$

➡ $\{\bar{\theta}, \bar{c}\}$

Removes assumption-based bias

Simultaneous fit of PDF and new physics

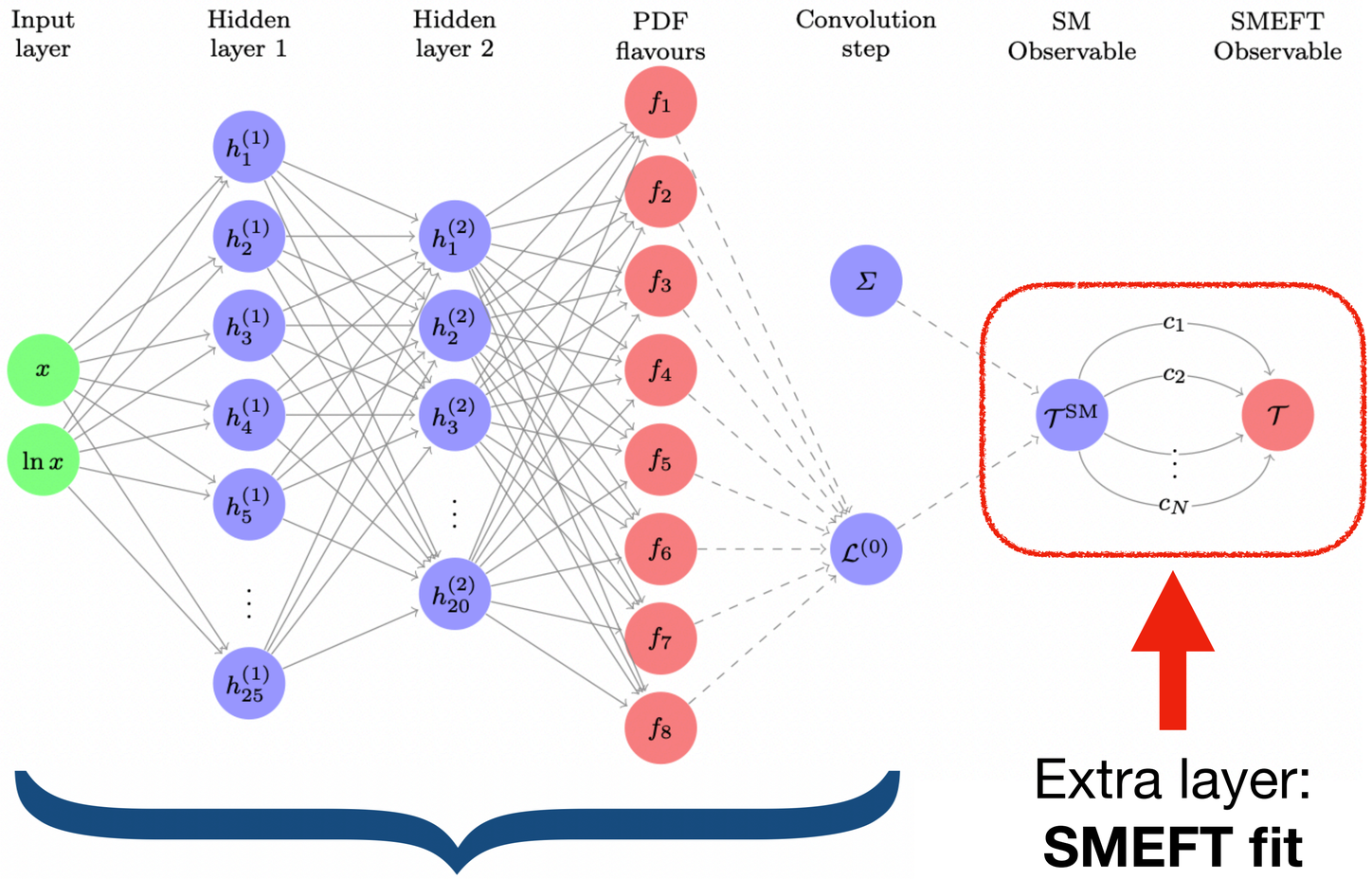
Presentation of the tool: SIMUnet

SIMUnet:

- Open-source tool: github.com/HEP-PBSP/SIMUnet

  [2402.03308]

- Fits PDFs and WC simultaneously



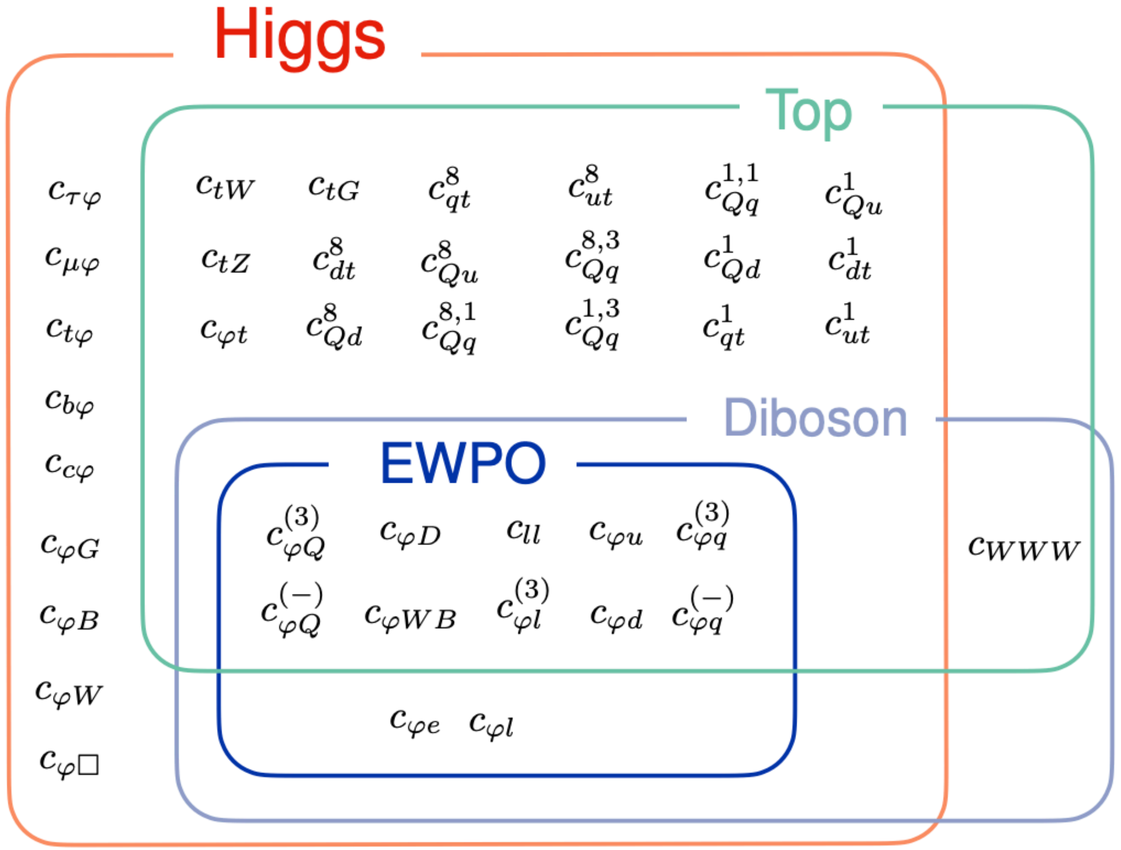
[Iranipour et Ubiali, 2201.07240]

Simultaneous fit of PDF and new physics

SMEFT operators implemented

- 40 operators implemented

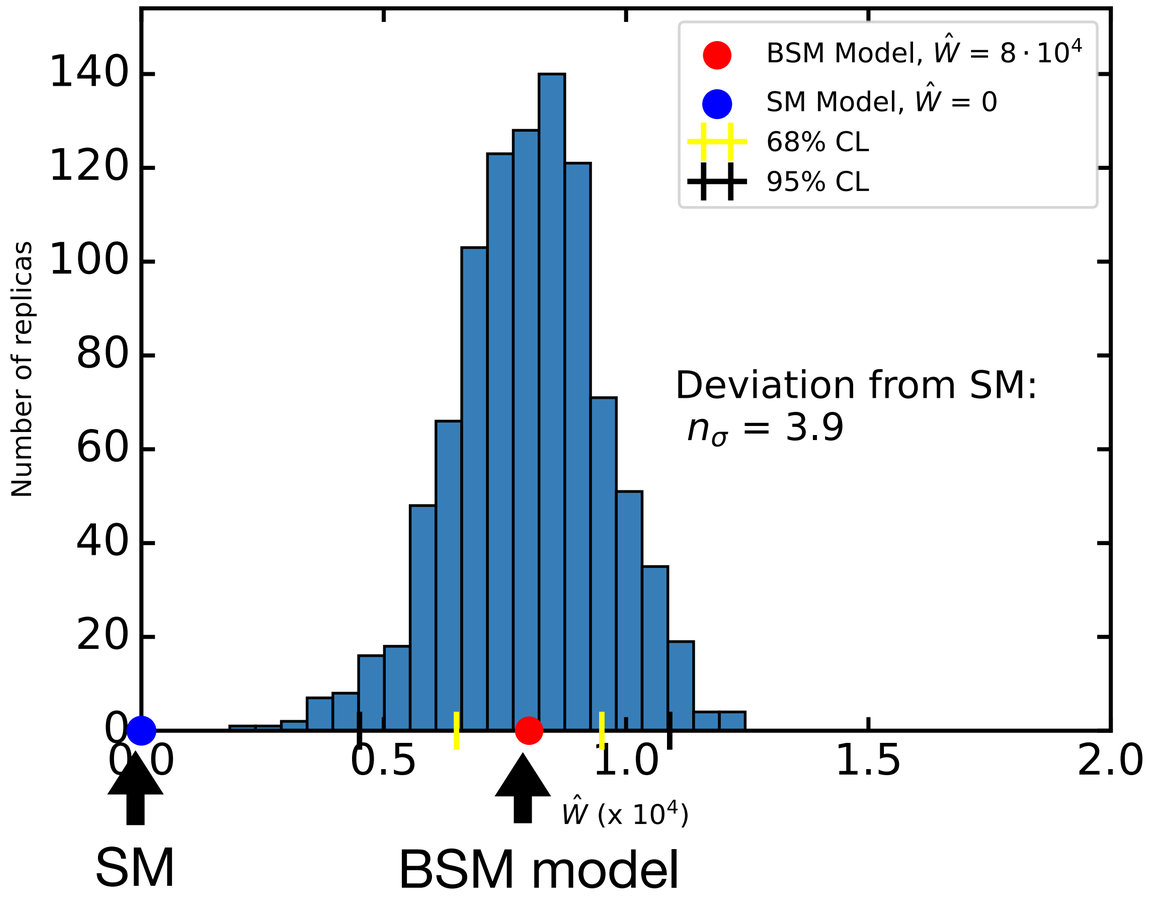
- Observables:
  - top sector
  - diboson
  - Higgs
  - Drell-Yan
  - EW Precision Observables



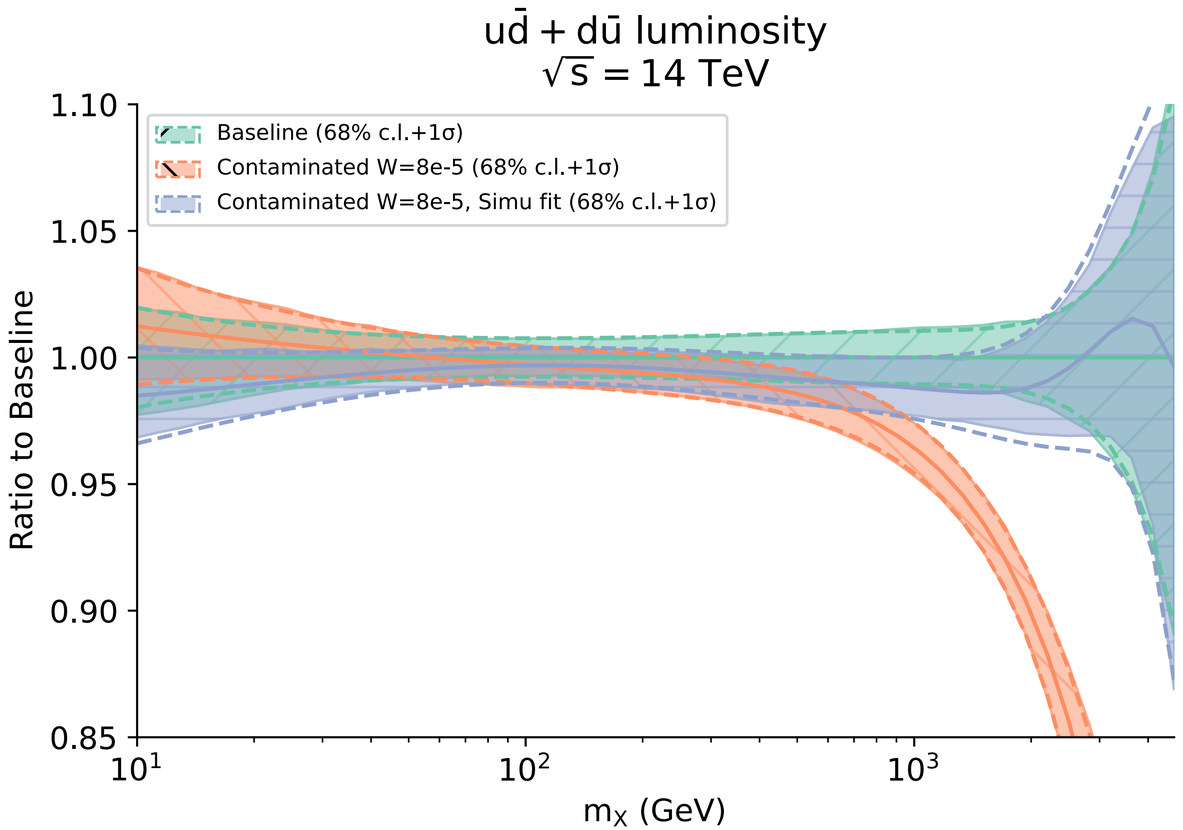
Simultaneous fit of PDF and new physics

Disentangling PDF contamination

SMEFT Fit



PDF Fit



Limits of the simultaneous fits

- Technical limits:
  - Can only fit linear SMEFT corrections (fitting method)
    - Working on an alternative bayesian method

- Fundamental limits:
  - More difficult than PDF fit
  - Need to choose SMEFT operators
  - PDF still universal?
  - Risk of fitting QCD/EWK corrections as BSM signals

[PBSP, forthcoming]

Summary and outlook

- Signs of W' got fitted away in PDF parametrisation
  - Missed new physics
  - Introduced fake deviations in other sectors
- Simultaneous fits of PDFs and SMEFT:
  - **SIMUnet** tool available
  - Seems to disentangle PDF and SMEFT

You can contact me at:
eh651@cam.ac.uk

Thank you for your attention!

Extra slides

Synergy of high and low-energy data

Adding low-energy dataset constraining the large-x region

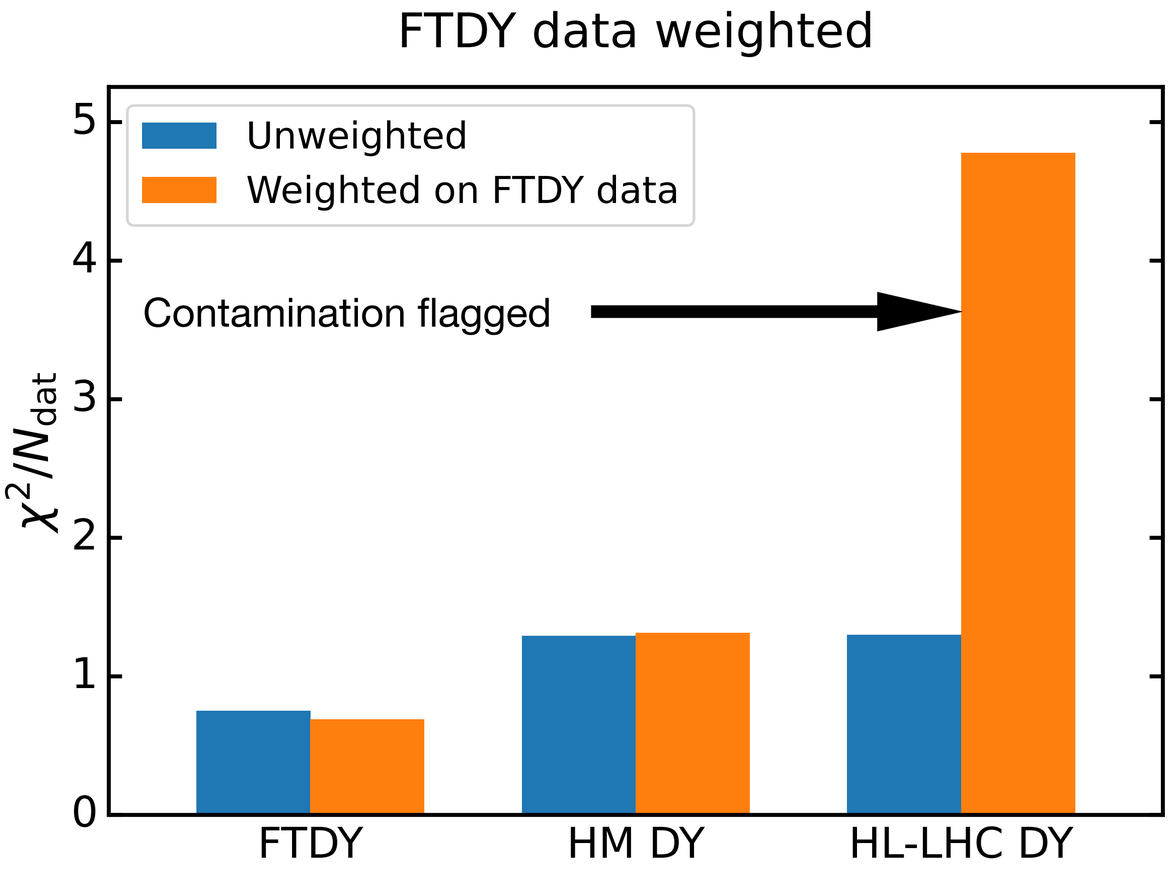
Excessive antiquark PDF flexibility in large-x region:

- ➡ Accommodates real data and BSM pseudodata
- ➡ Allows contamination

Including low-energy large-x data:

- Constraint large-x region
- Safe from BSM contamination

[Hammou et Ubiali, 2410.00963]



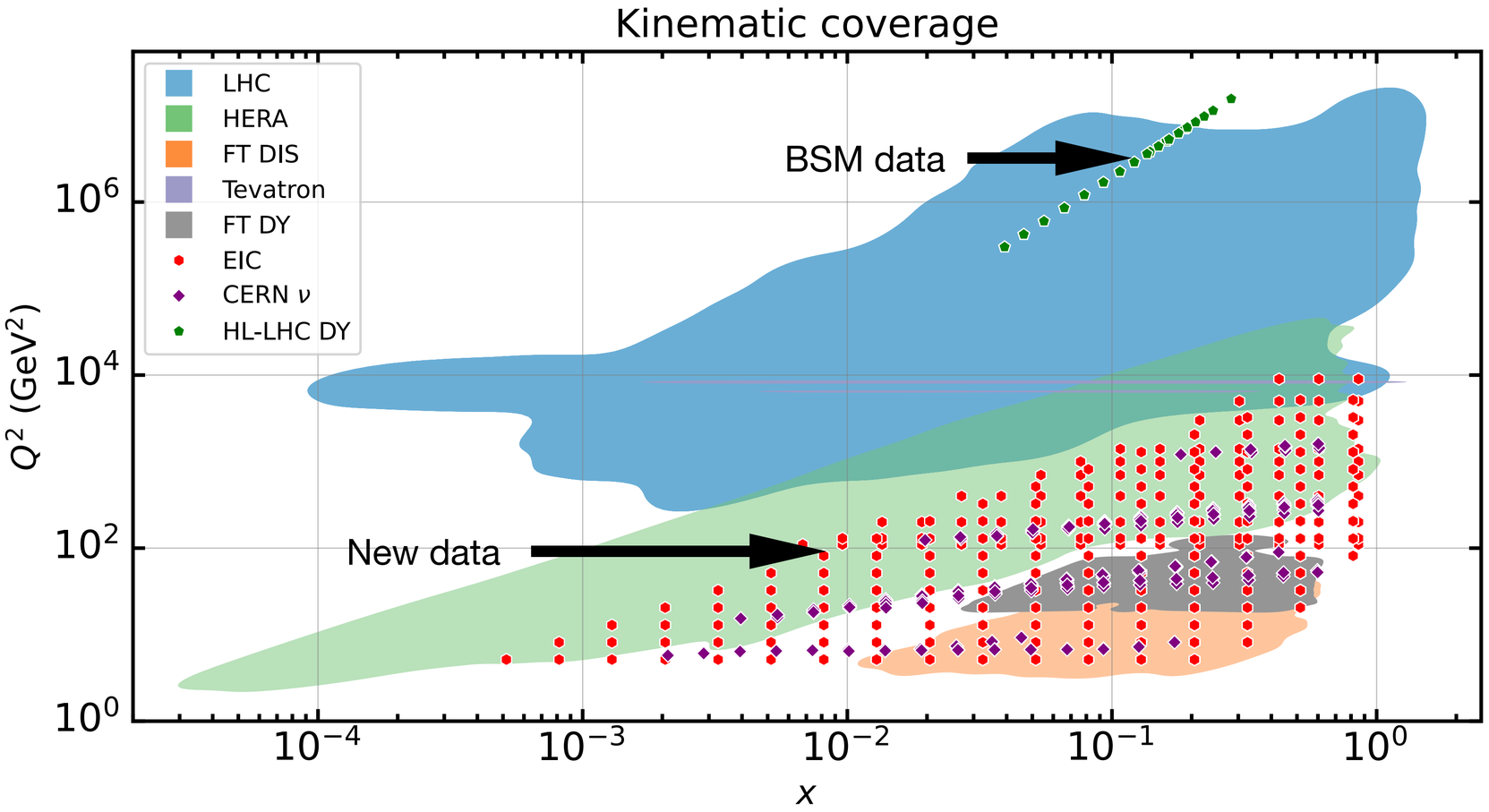
Future low energy data

Kinematic coverage

Projection data:

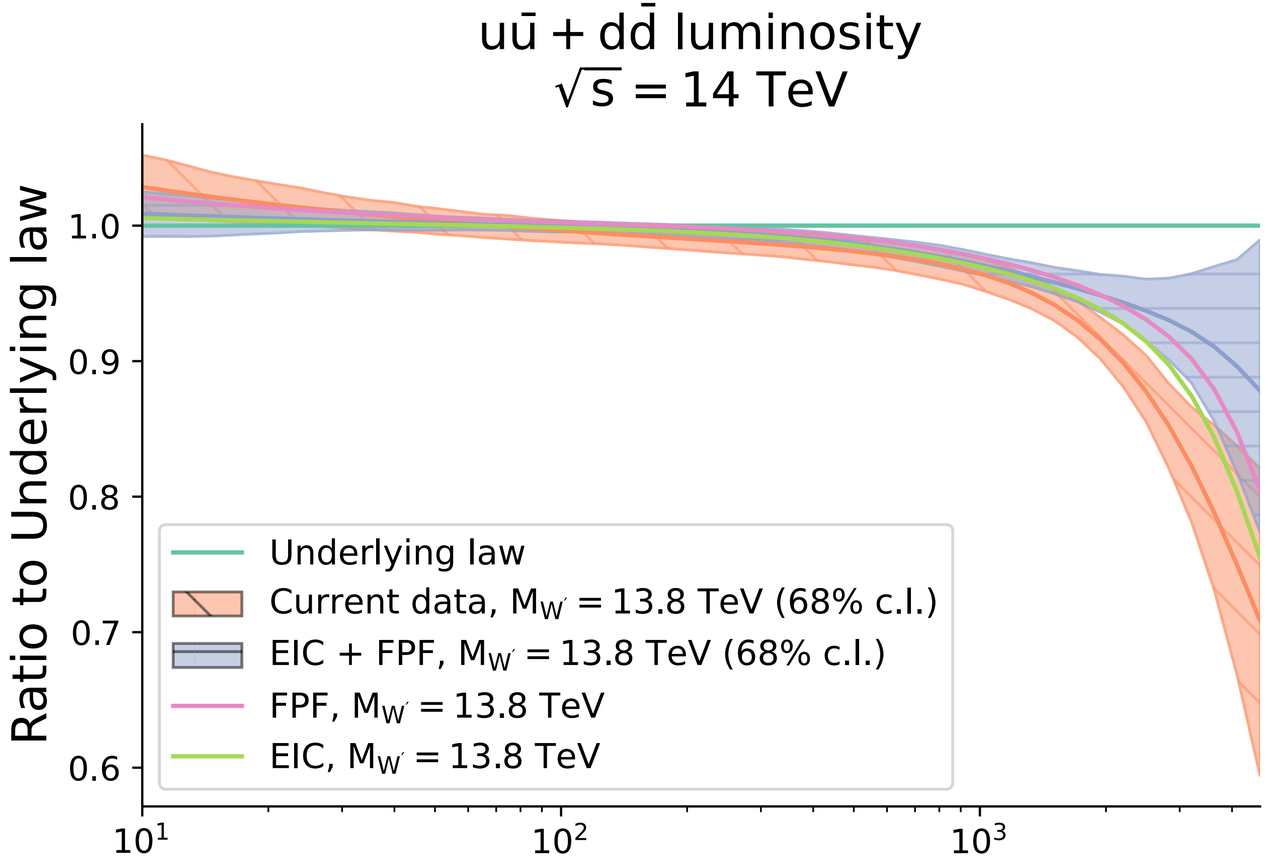
- Electron Ion Collider (EIC)
- Forward Physics Facility (FPF)

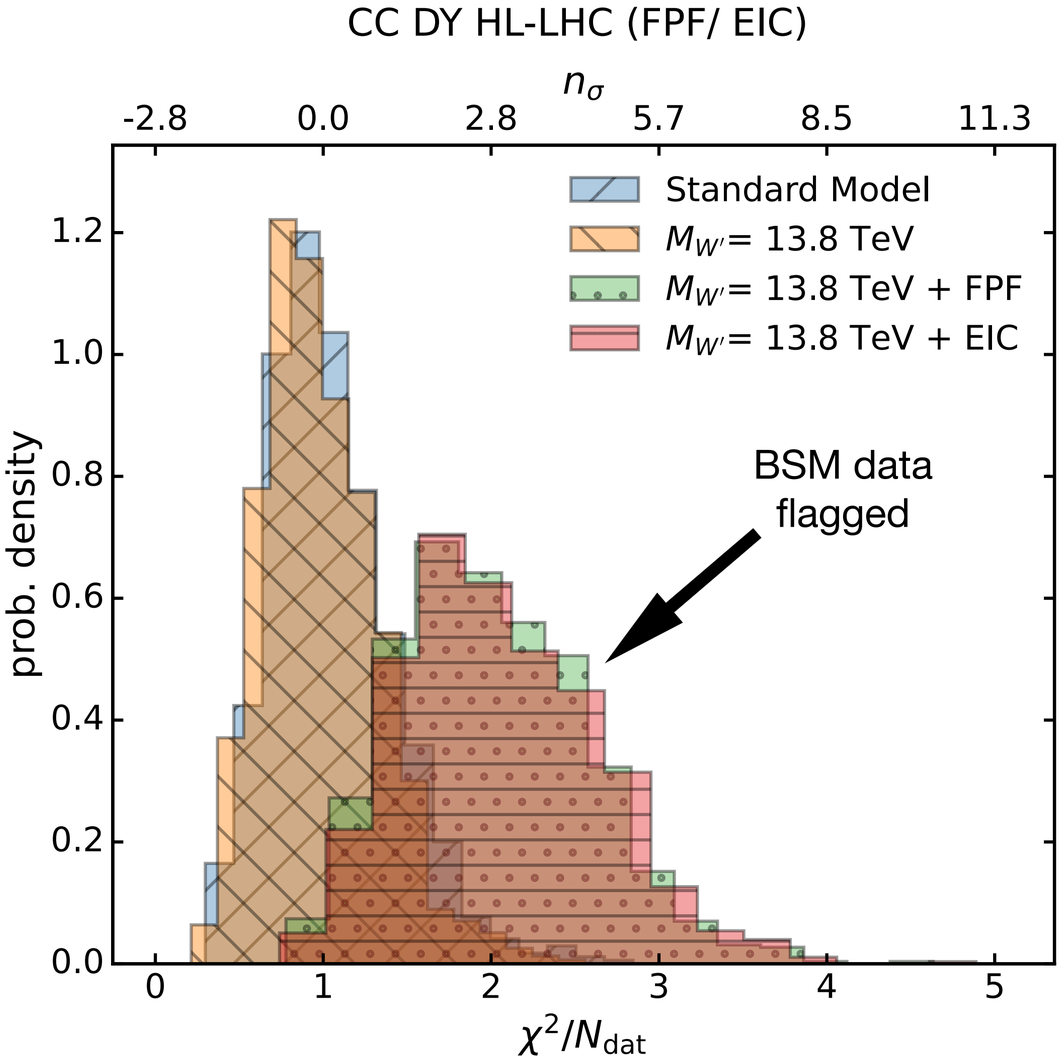
  (neutrino DIS)



Impact on the PDF contamination

Flagging the BSM data





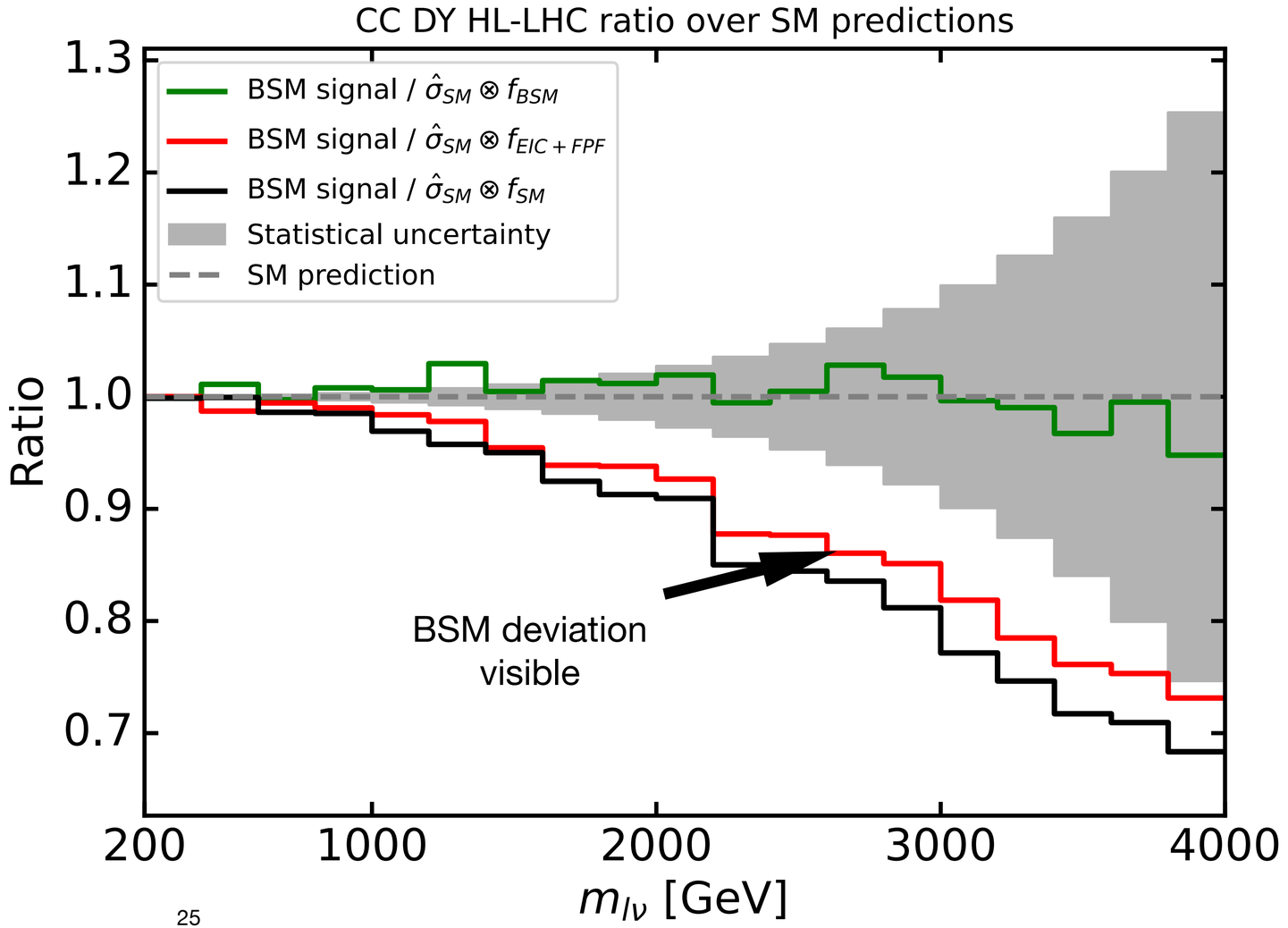
Recovering the signs of new physics

BSM data versus SM theory predictions

$$\hat{\sigma}_{BSM} \otimes \mathscr{L}_{SM} \approx \hat{\sigma}_{SM} \otimes \mathscr{L}_{BSM}$$

$$\hat{\sigma}_{BSM} \otimes \mathscr{L}_{SM} \neq \hat{\sigma}_{SM} \otimes \mathscr{L}_{EIC+FPF}$$

$M_{W'}$: 13.8 TeV



Shift of the contamination threshold

From the fit quality

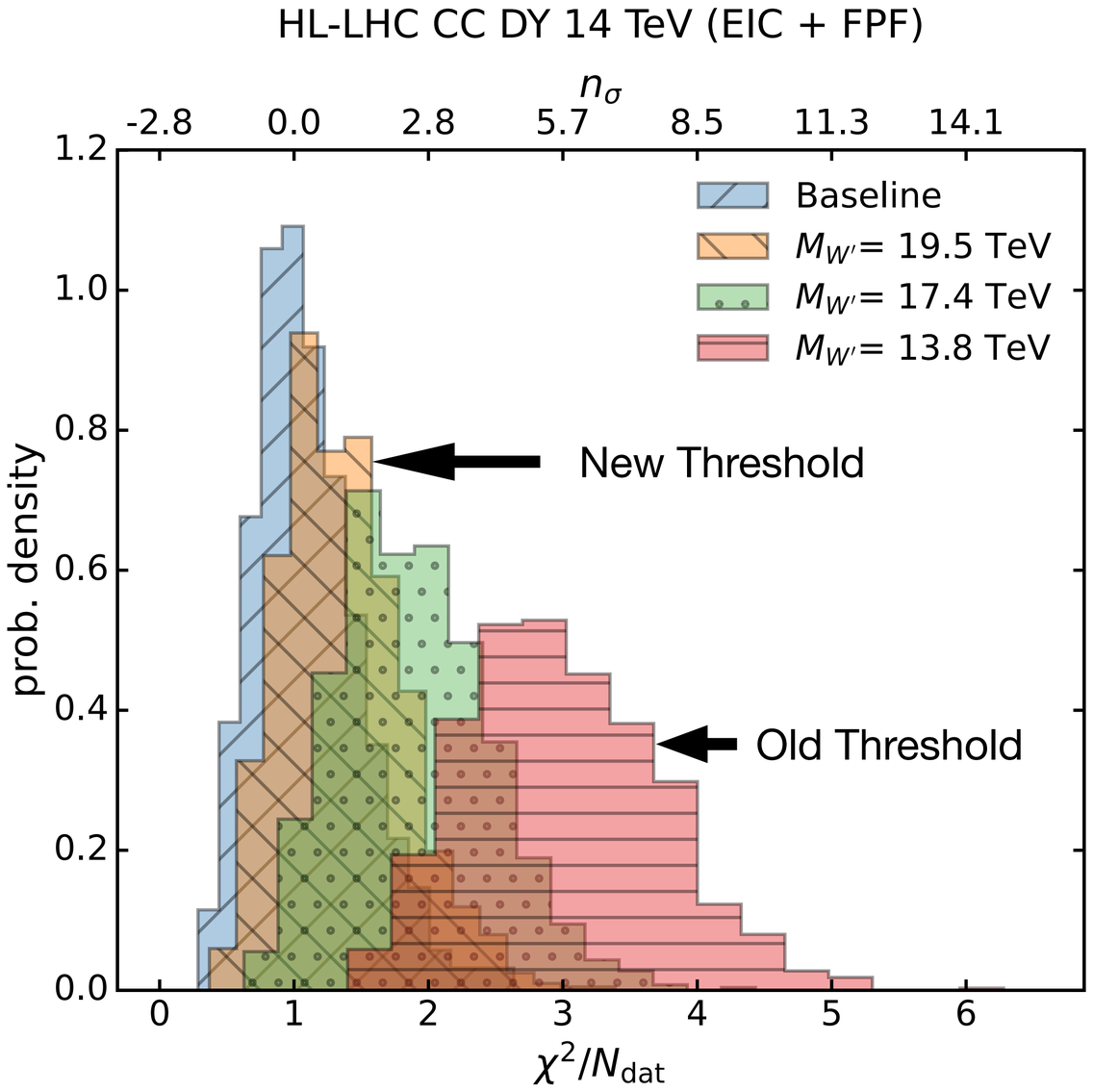
Not a complete solution:

Smaller deviations can still be absorbed

➡ risk at higher BSM mass

Reduction of the “blindspot”:

$$M_{W'} : 13.8 \rightarrow 19.5 \text{ TeV}$$



New physics scenarios: Z'

Generation of the pseudodata

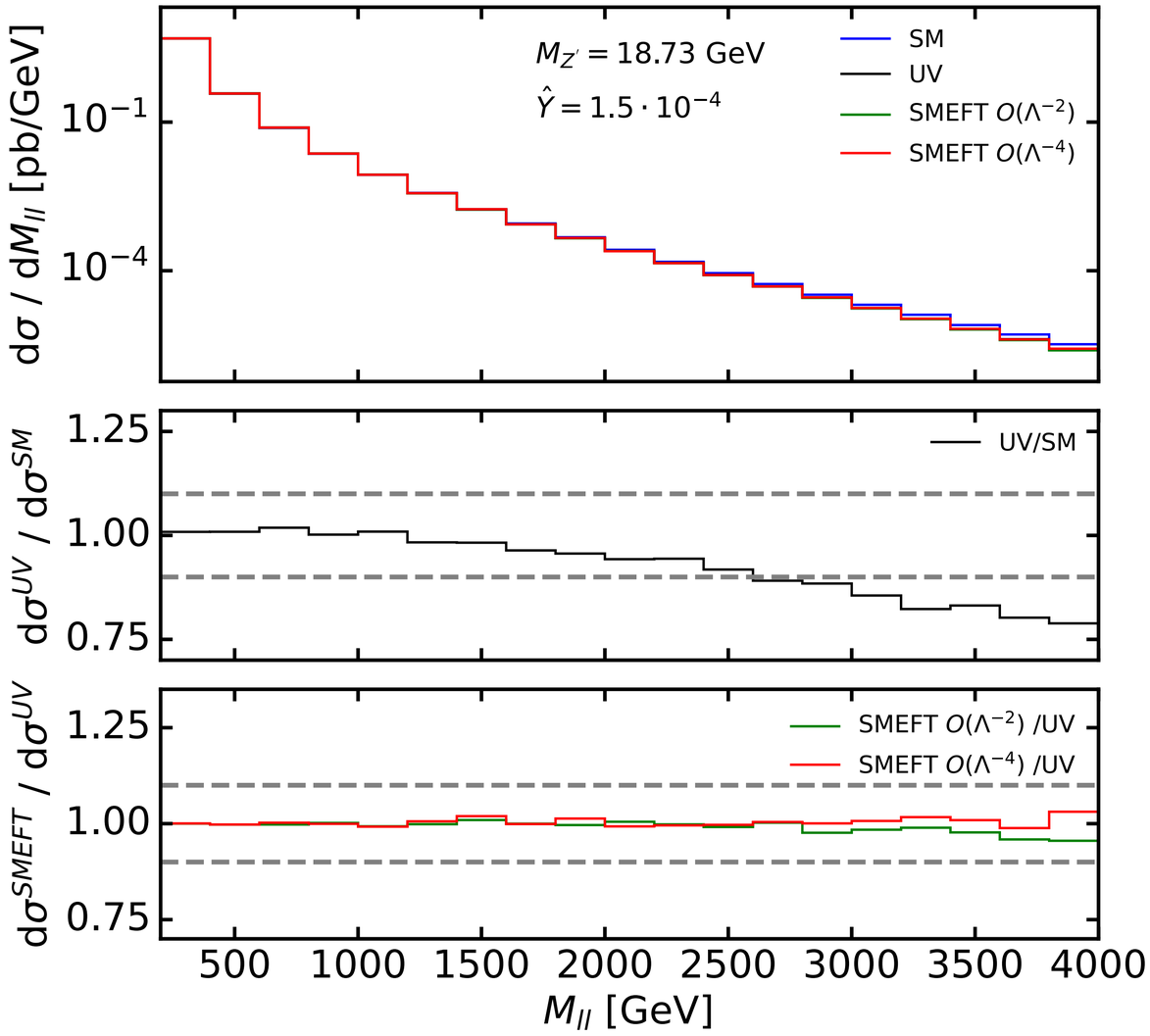
$$\mathscr{L}^{Z'}_{SMEFT} = \mathscr{L}_{SM} - \frac{g^2_{Z'}}{2M^2_{Z'}} J^{\mu}_{Y} J_{Y,\mu}$$

$$J^{\mu}_{Y} = \sum_{f} Y_f \bar{f} \gamma^{\mu} f$$

Impacts neutral current Drell-Yan processes

$$p\bar{p} \to l^+ l^-$$

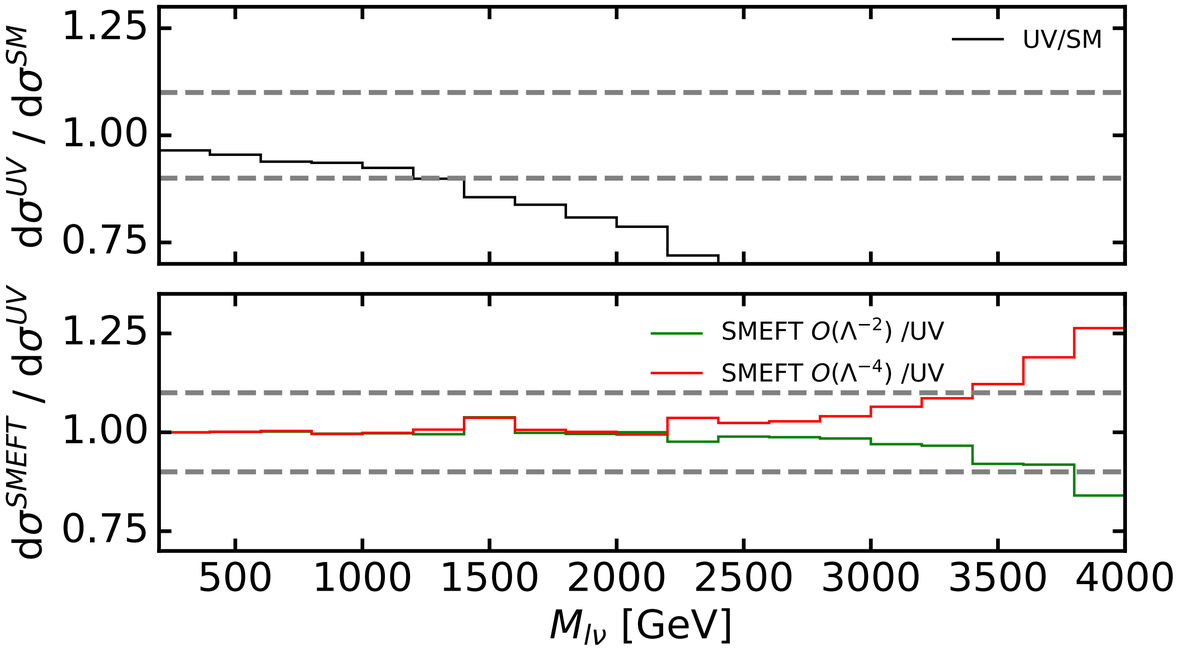
$M_{Z'} = 18.7$ TeV



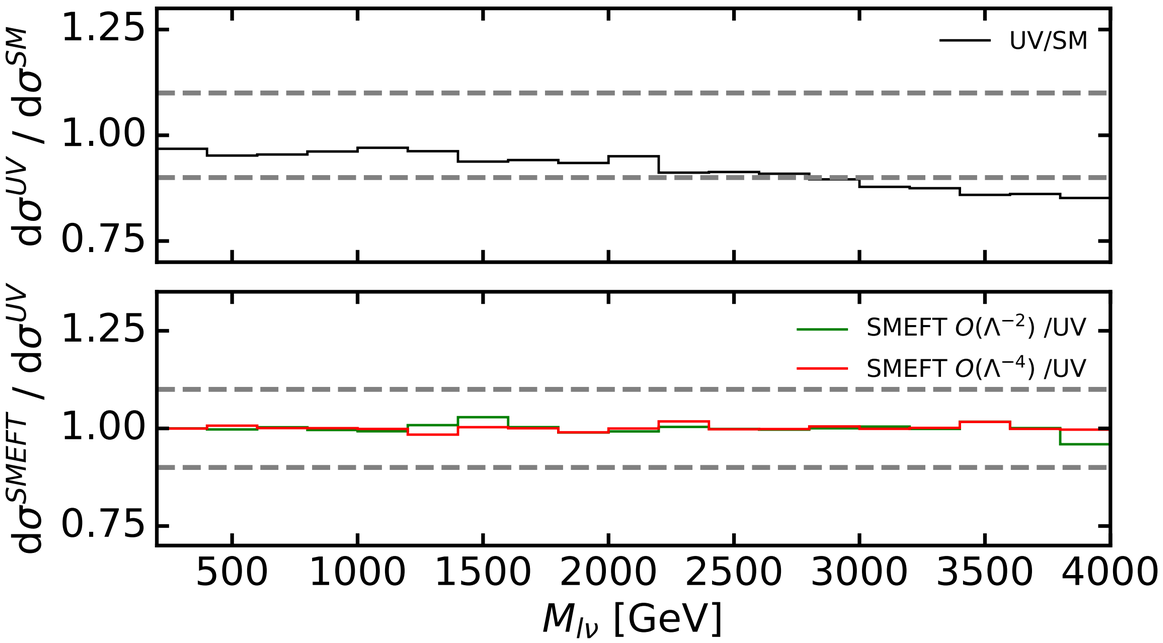
New physics scenarios: W'

Consideration of different masses

$M_{W'} = 10$ TeV



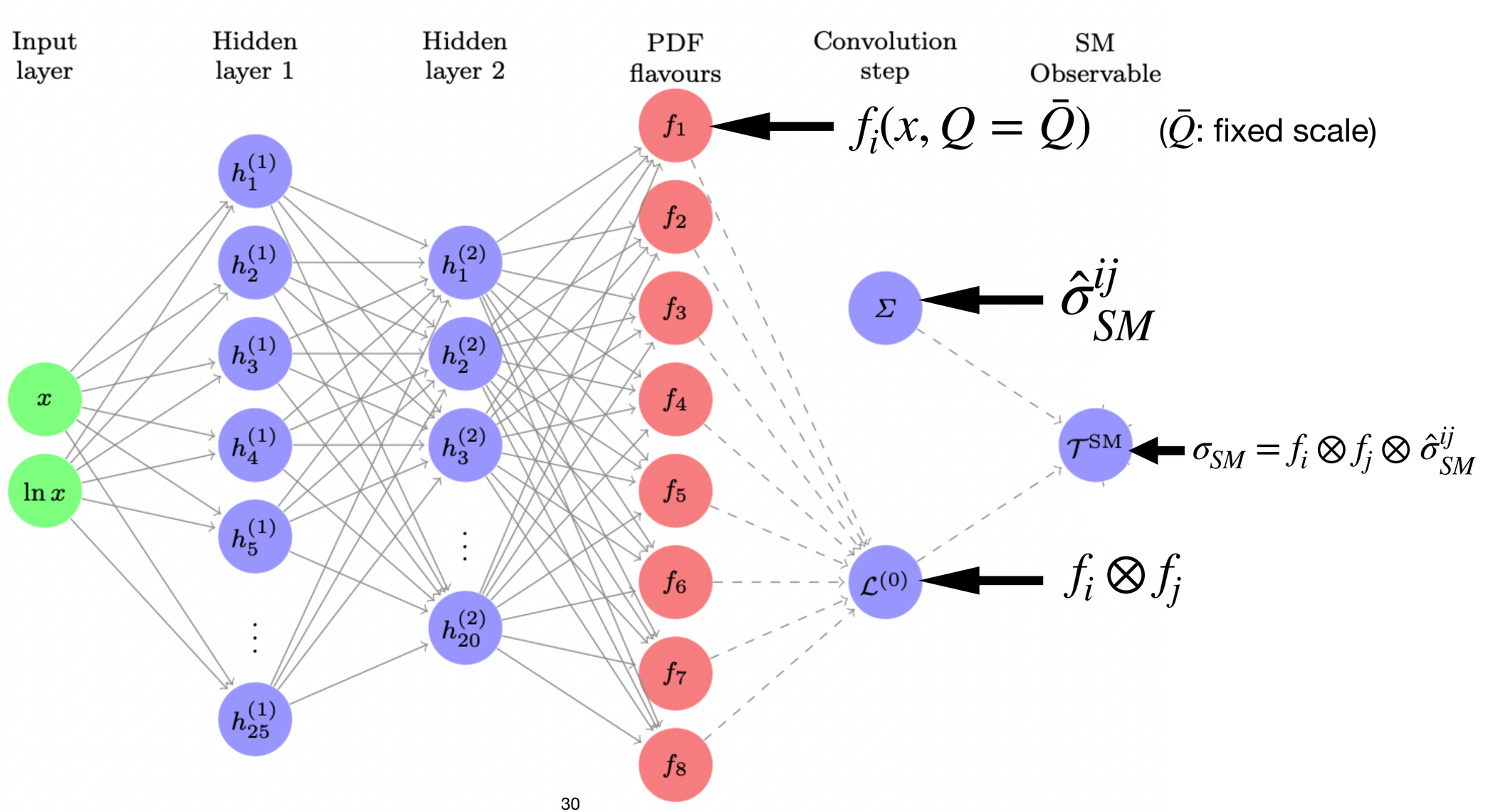
$M_{W'} = 22.5$ TeV



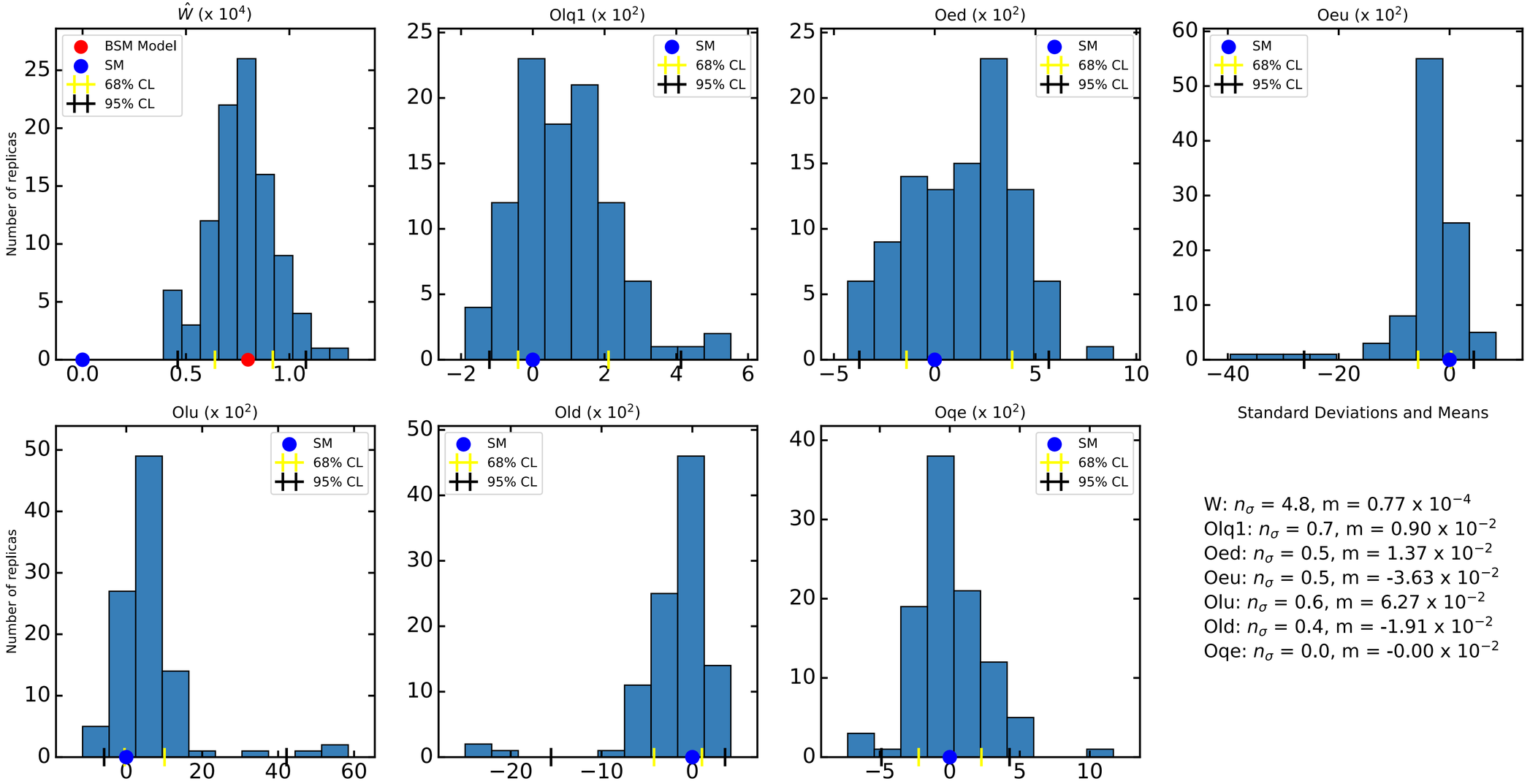
List of deviations

	HL-LHC		Stat. improved	
Dataset	$\chi^2/n_{\rm dat}$	n_σ	$\chi^2/n_{\rm dat}$	n_σ
W^+H	1.17	0.41	1.77	1.97
W^-H	1.08	0.19	1.08	0.19
W^+Z	1.08	0.19	1.49	1.20
W^-Z	0.99	-0.03	1.02	0.05
ZH	1.19	0.44	1.67	1.58
W^+W^-	2.19	3.04	2.69	4.31
VBF $\rightarrow$ H	0.70	-0.74	0.62	-0.90

NNPDF methodology

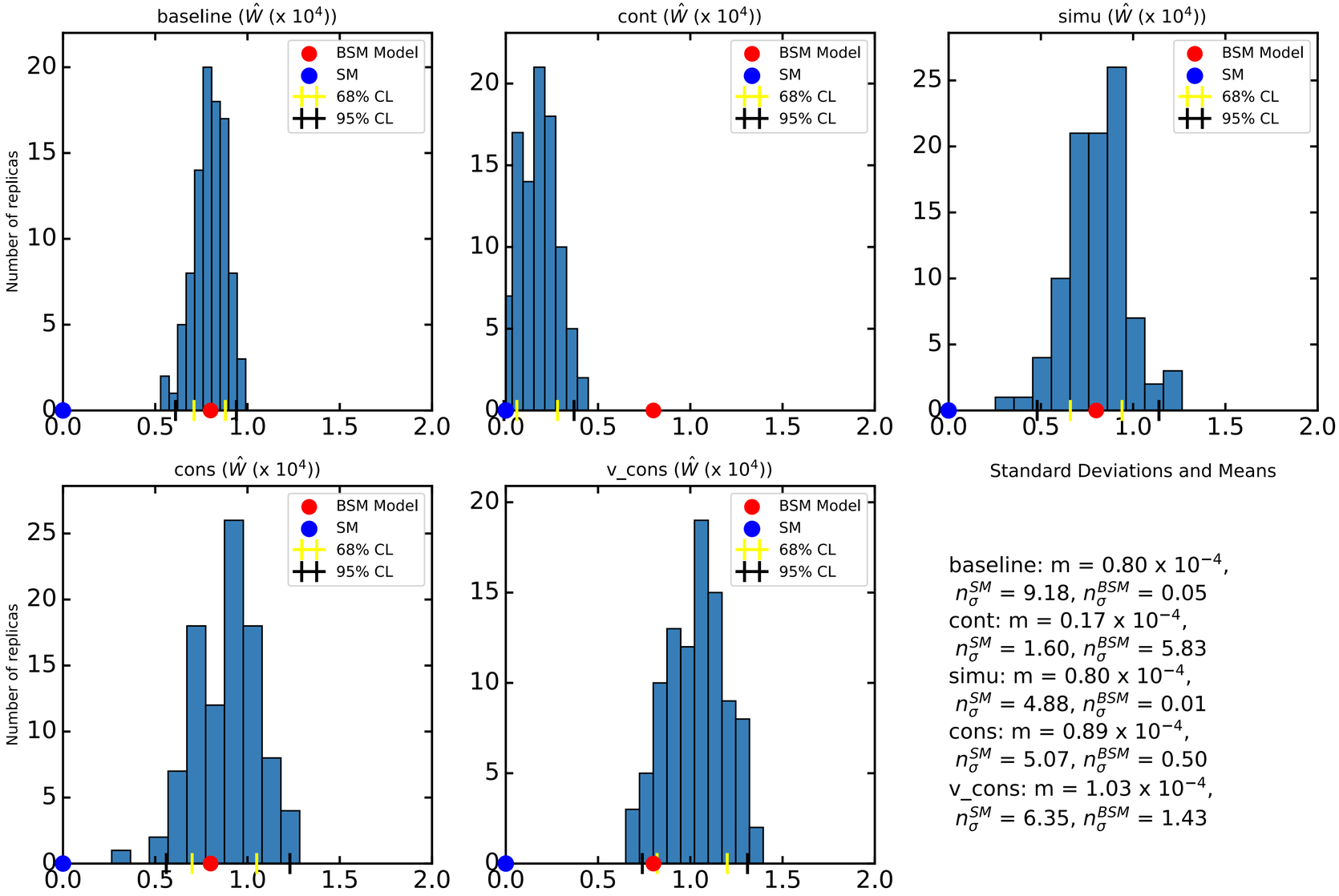


Global SMEFT fit, 4 fermions operators



W: n_σ = 4.8, m = 0.77 x 10^{-4}
Olq1: n_σ = 0.7, m = 0.90 x 10^{-2}
Oed: n_σ = 0.5, m = 1.37 x 10^{-2}
Oeu: n_σ = 0.5, m = -3.63 x 10^{-2}
Olu: n_σ = 0.6, m = 6.27 x 10^{-2}
Old: n_σ = 0.4, m = -1.91 x 10^{-2}
Oqe: n_σ = 0.0, m = -0.00 x 10^{-2}

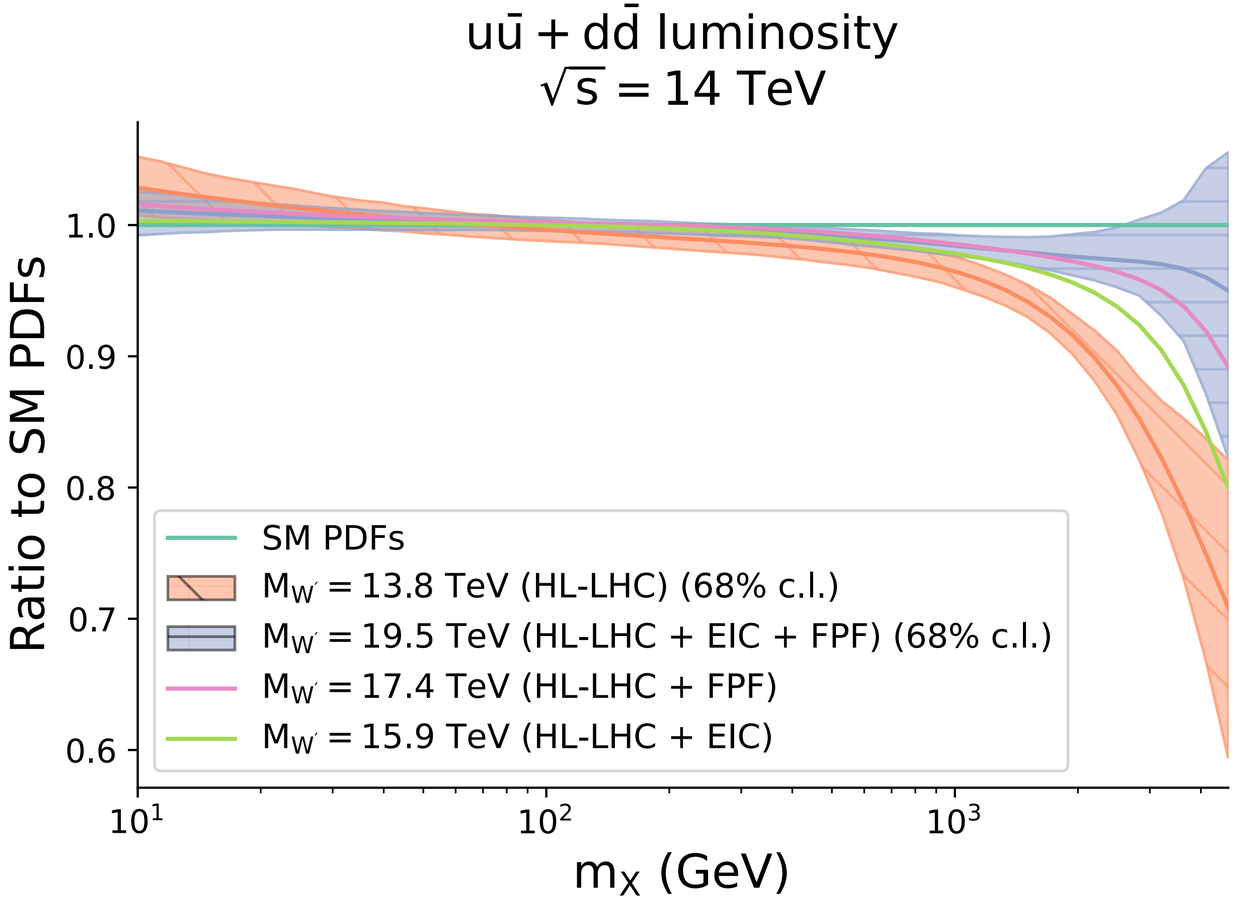
SMEFT fits with different PDFs

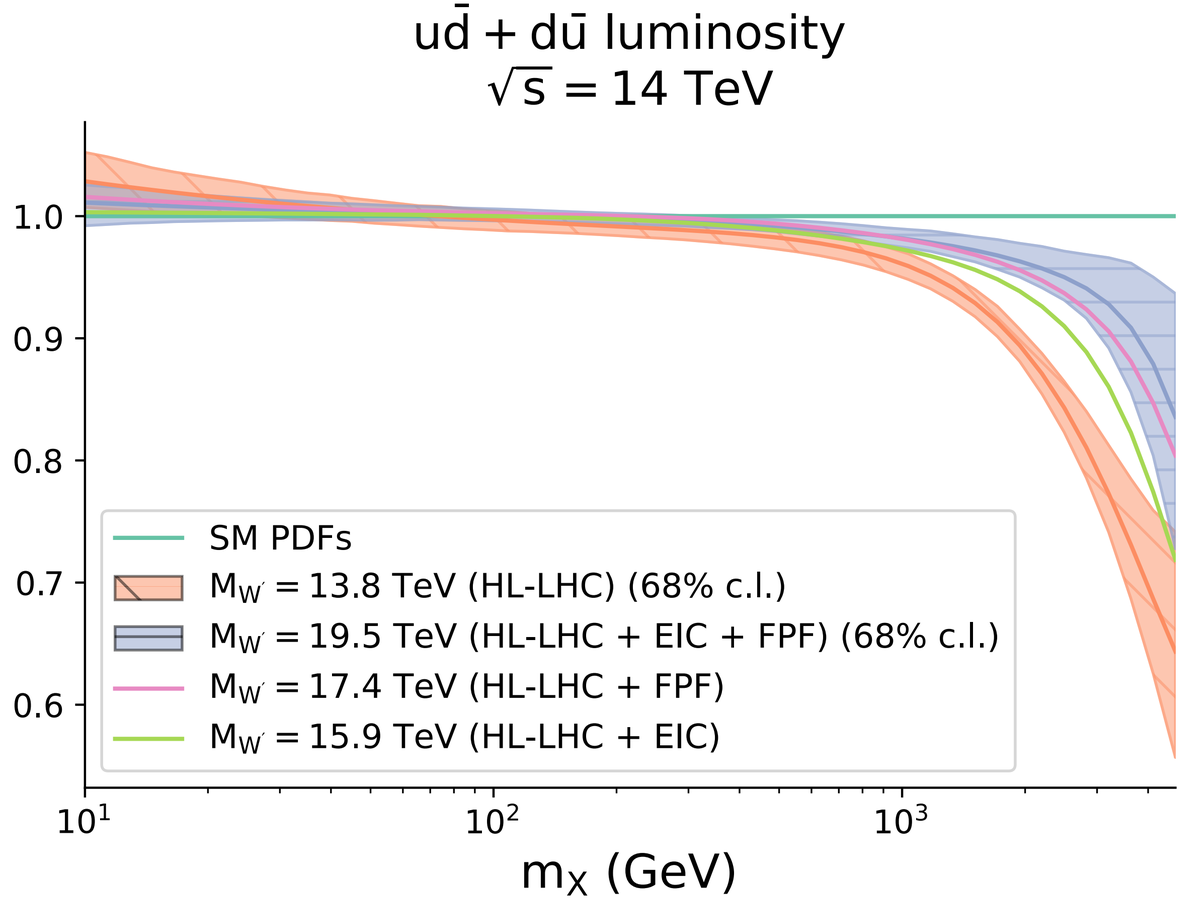


baseline: m = 0.80 x 10^{-4},
n_σ^{SM} = 9.18, n_σ^{BSM} = 0.05
cont: m = 0.17 x 10^{-4},
n_σ^{SM} = 1.60, n_σ^{BSM} = 5.83
simu: m = 0.80 x 10^{-4},
n_σ^{SM} = 4.88, n_σ^{BSM} = 0.01
cons: m = 0.89 x 10^{-4},
n_σ^{SM} = 5.07, n_σ^{BSM} = 0.50
v_cons: m = 1.03 x 10^{-4},
n_σ^{SM} = 6.35, n_σ^{BSM} = 1.43

Shift of the contamination threshold

Impact on PDF luminosities

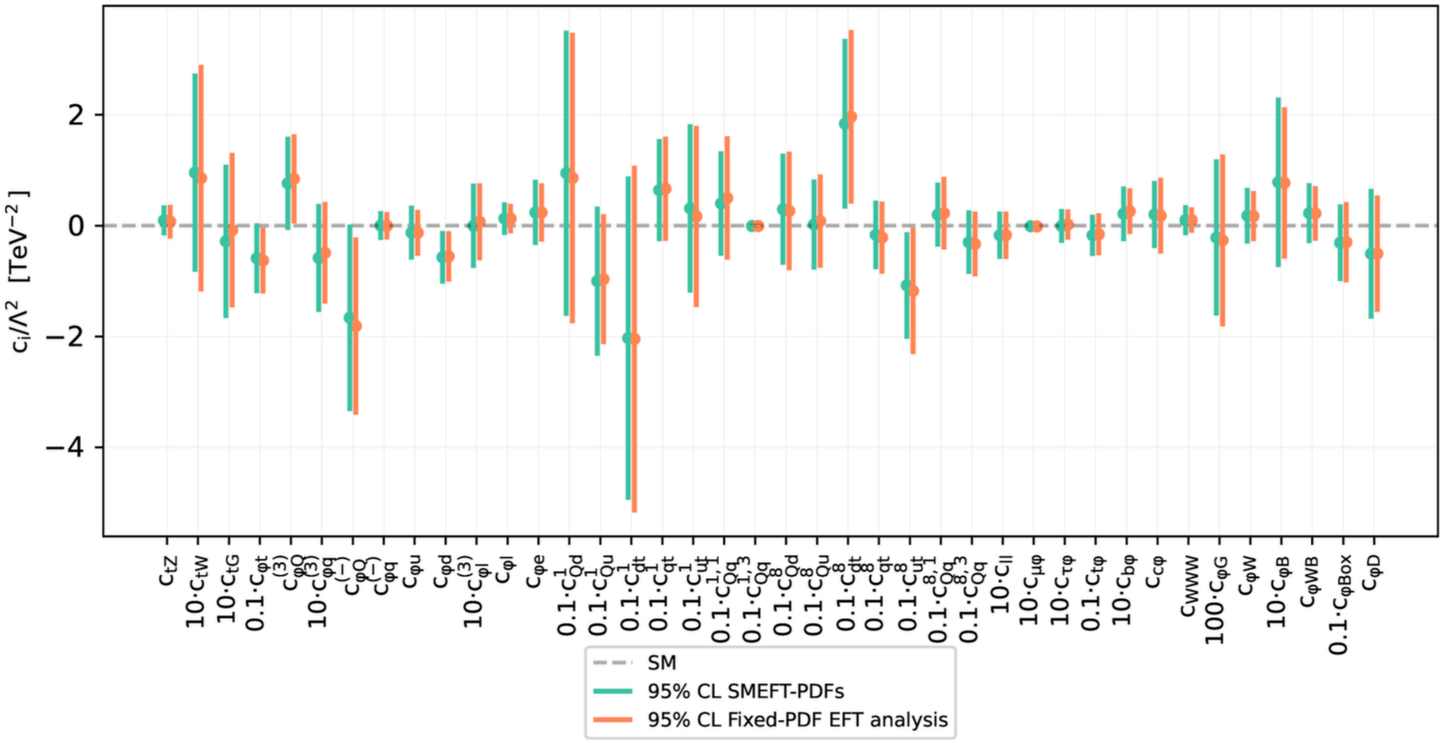




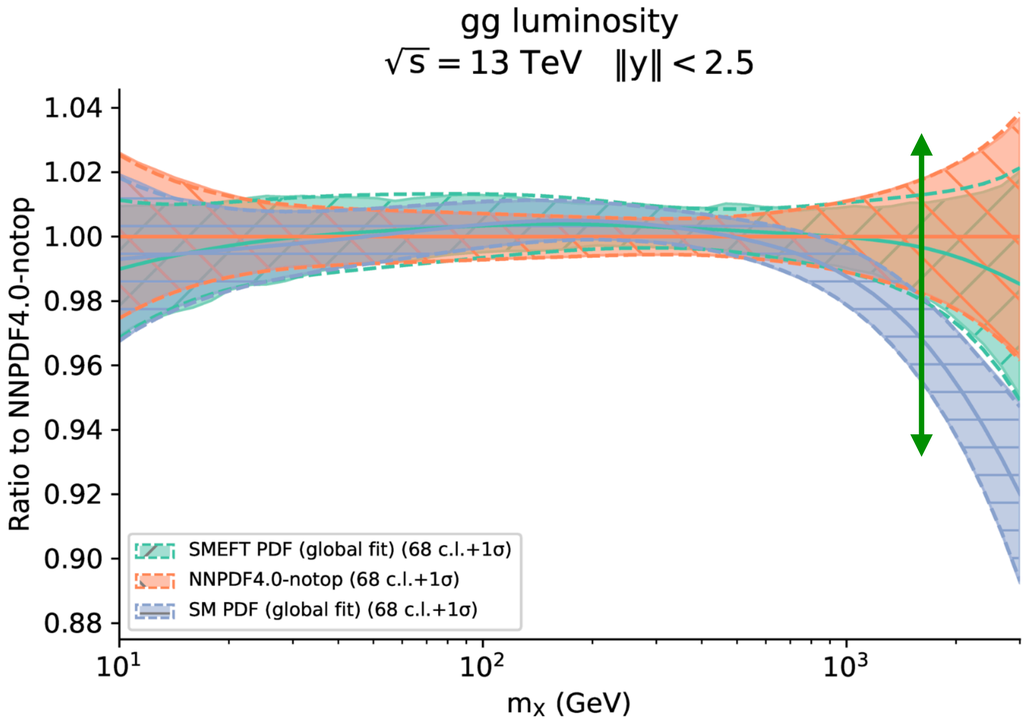
Application to the top sector (real data)

In progress

SMEFT Fit

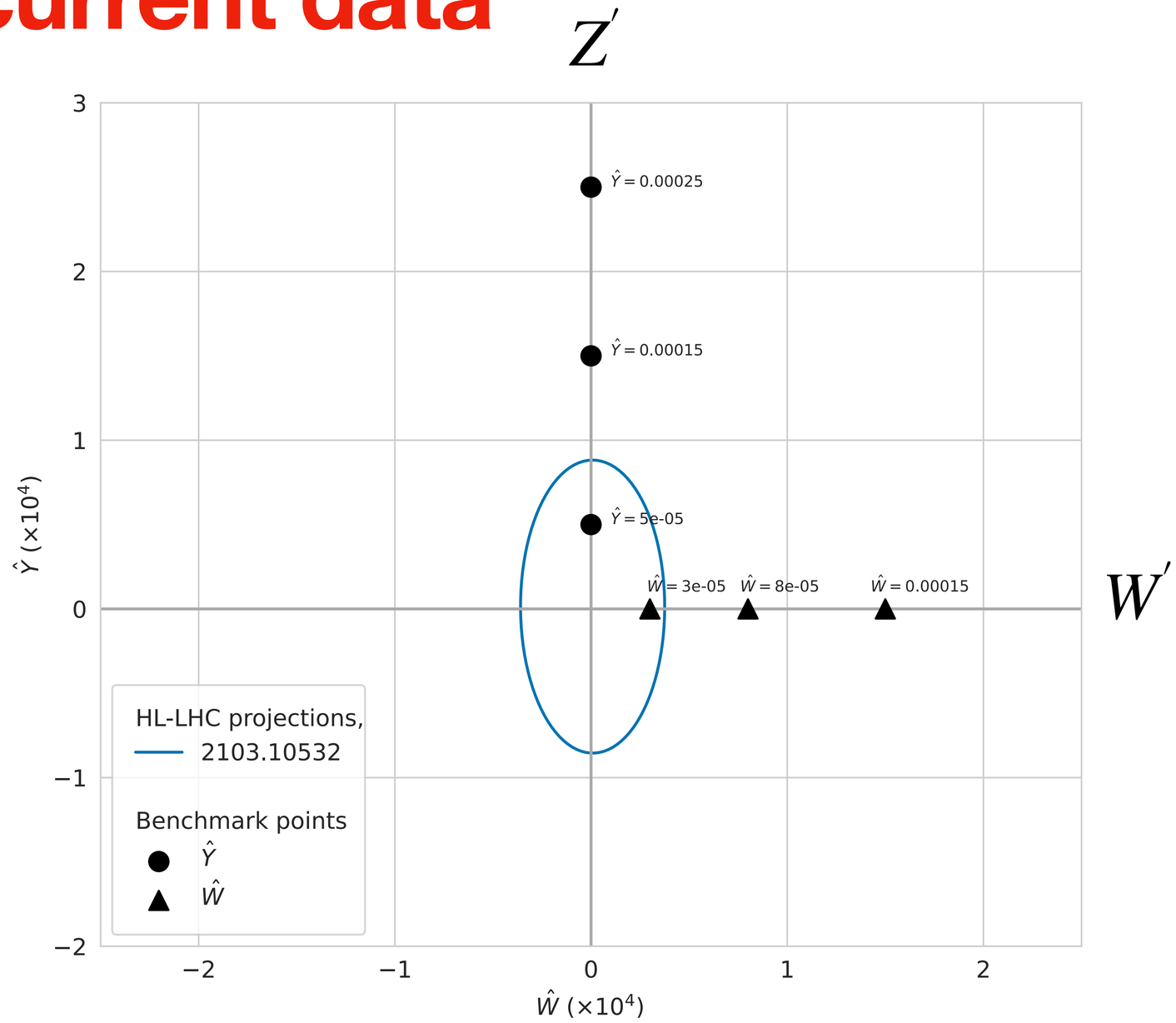


PDF Fit



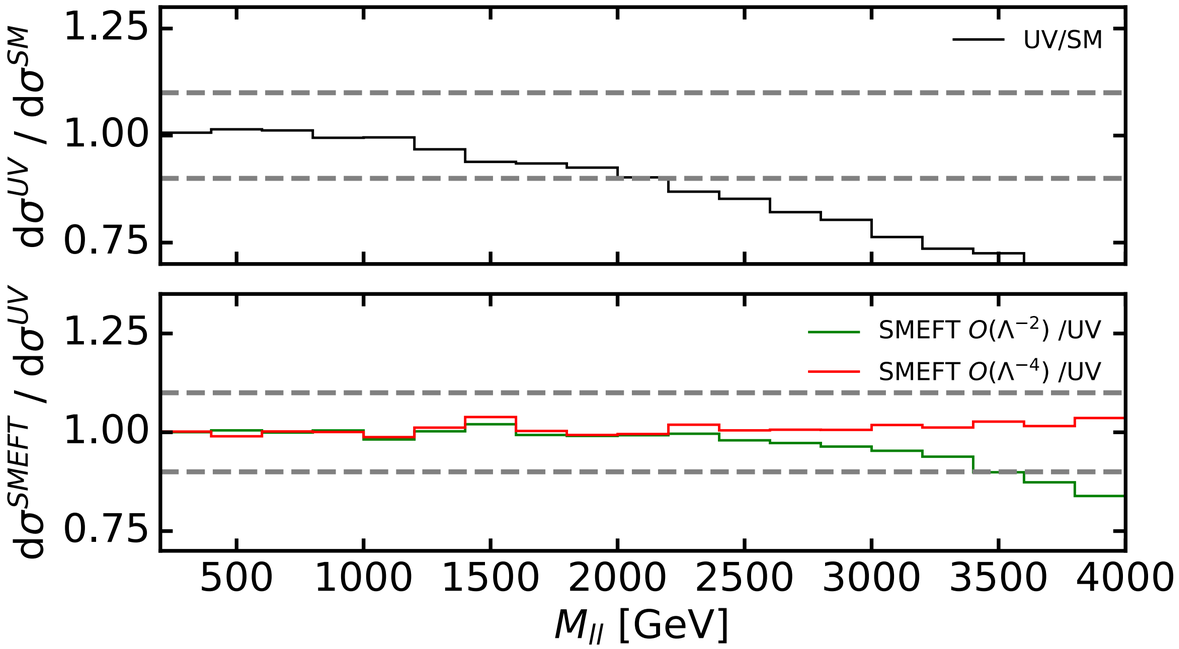
Constraints from current data

- New physics scenarios compared to constraints at 95% CL

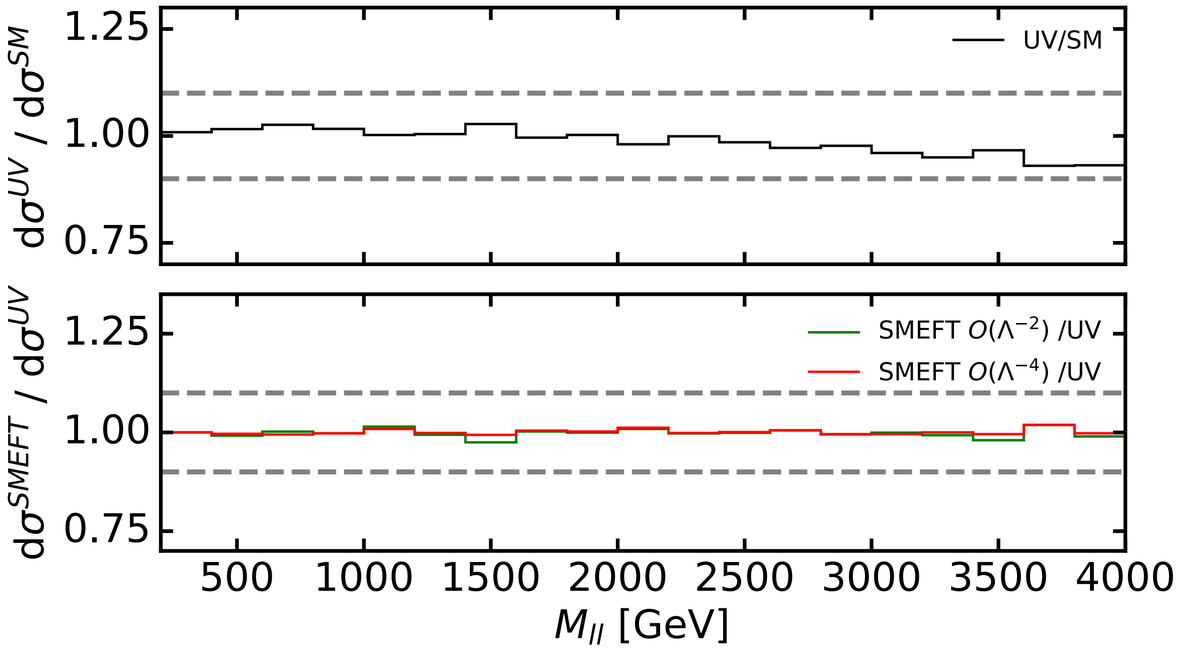


New physics scenarios: Z'

$M_{Z'} = 14.5$ TeV

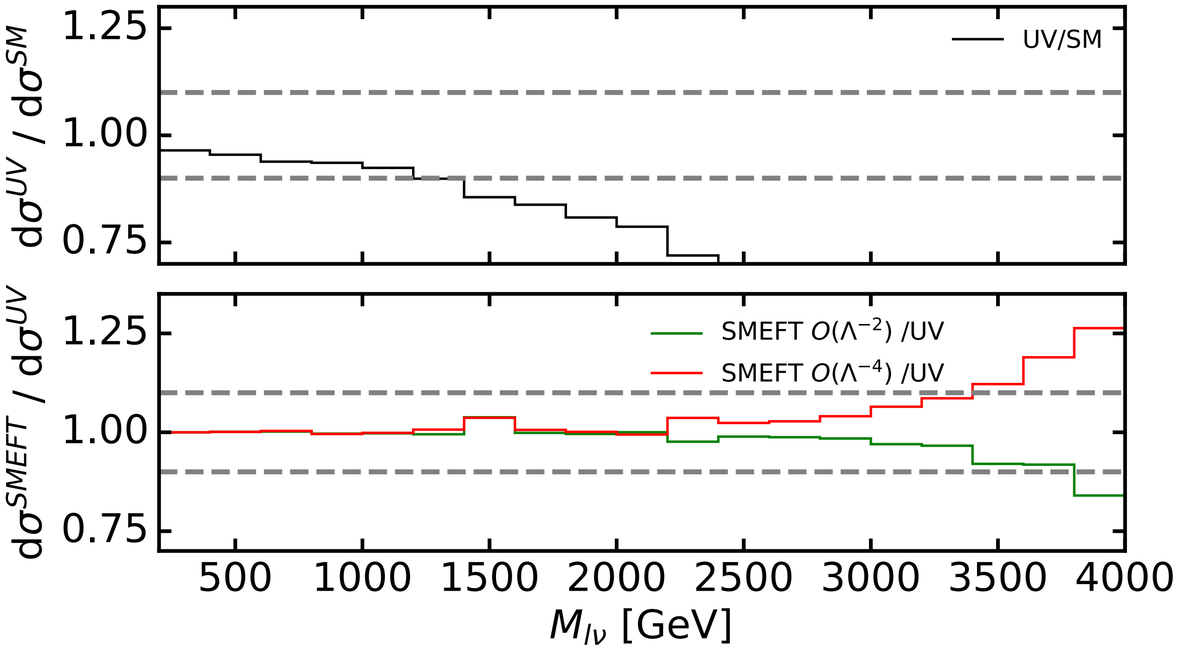


$M_{Z'} = 32.5$ TeV



New physics scenarios: W'

$M_{W'} = 10$ TeV



$M_{W'} = 22.5$ TeV