Can new physics be fitted away in the PDFs?

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

Work with Maria Ubiali and her group:

[2307.10370, JHEP]

[2402.03308]

[Forthcoming]

Elie Hammou, University of Cambridge
HEFT 2024, Bologna

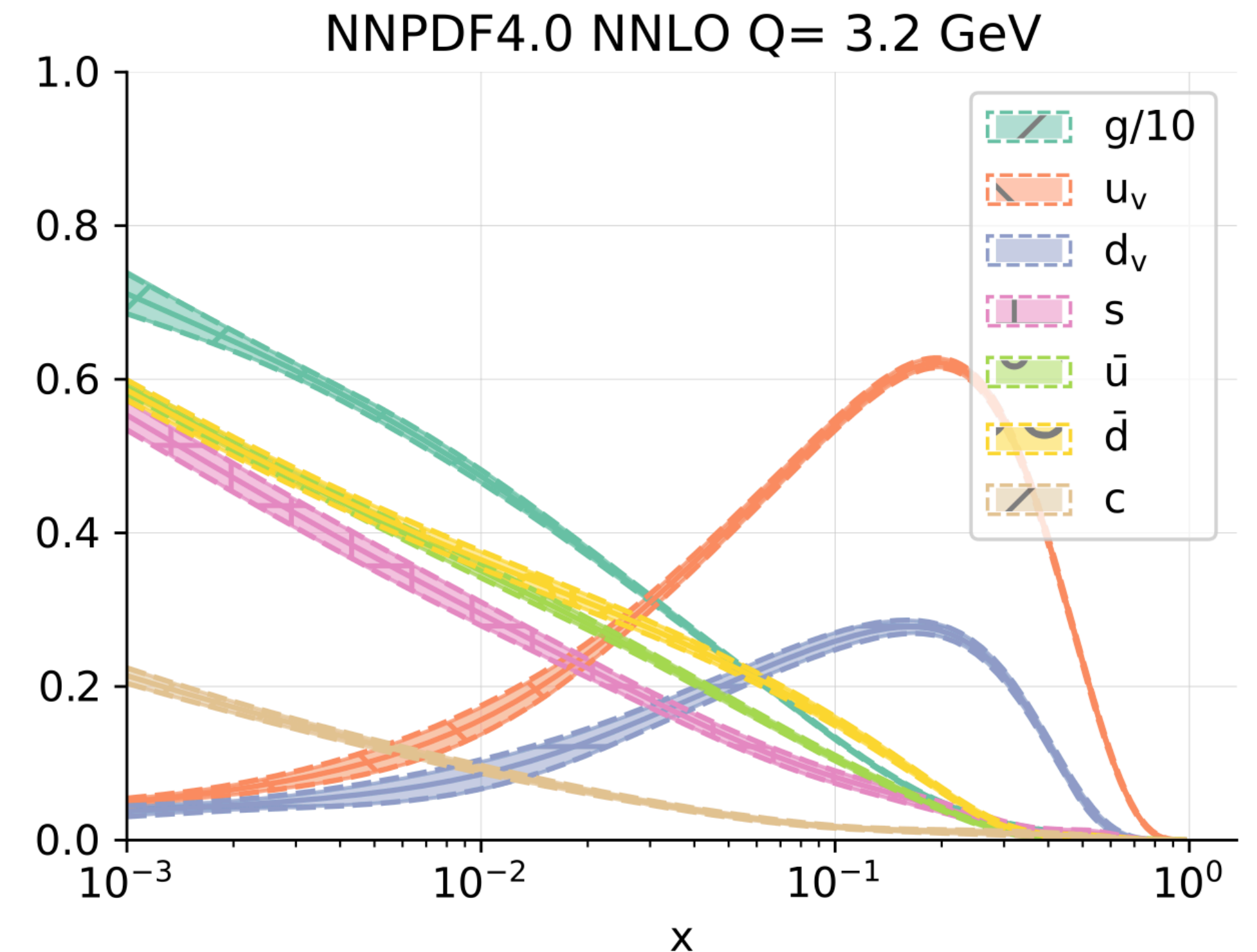# Background on Parton Distribution Functions

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

PDFs in a nutshell:

- describe proton's partonic content

- $f(x, Q)$

- x dependance: non-perturbative QCD

➡ **<u>Fitted from data</u>**

Using NNPDF methodology

NNPDF4.0 NNLO Q= 3.2 GeV

Legend:
- g/10
- u_v
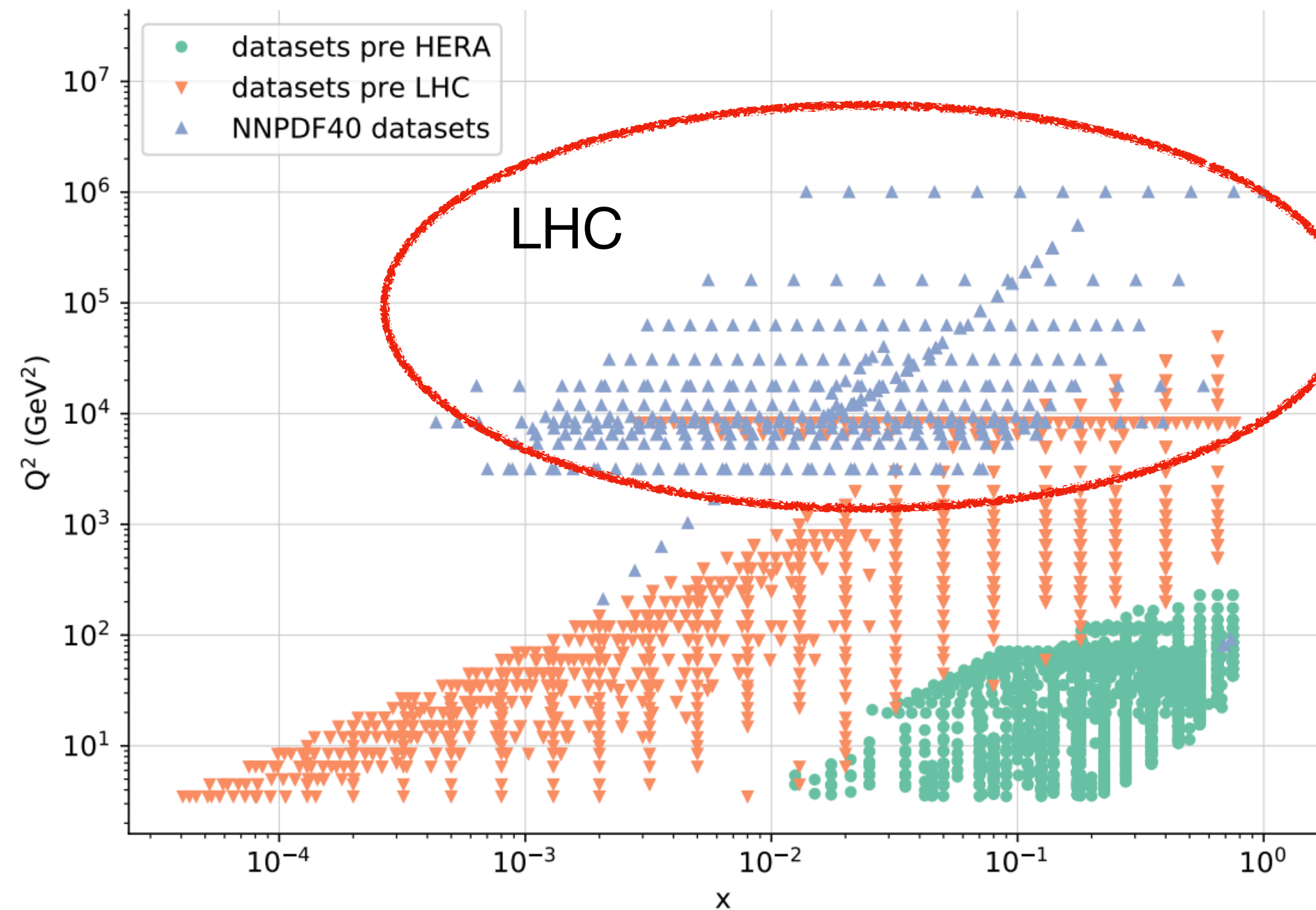- d_v
- s
- $\bar{u}$
- $\bar{d}$
- c

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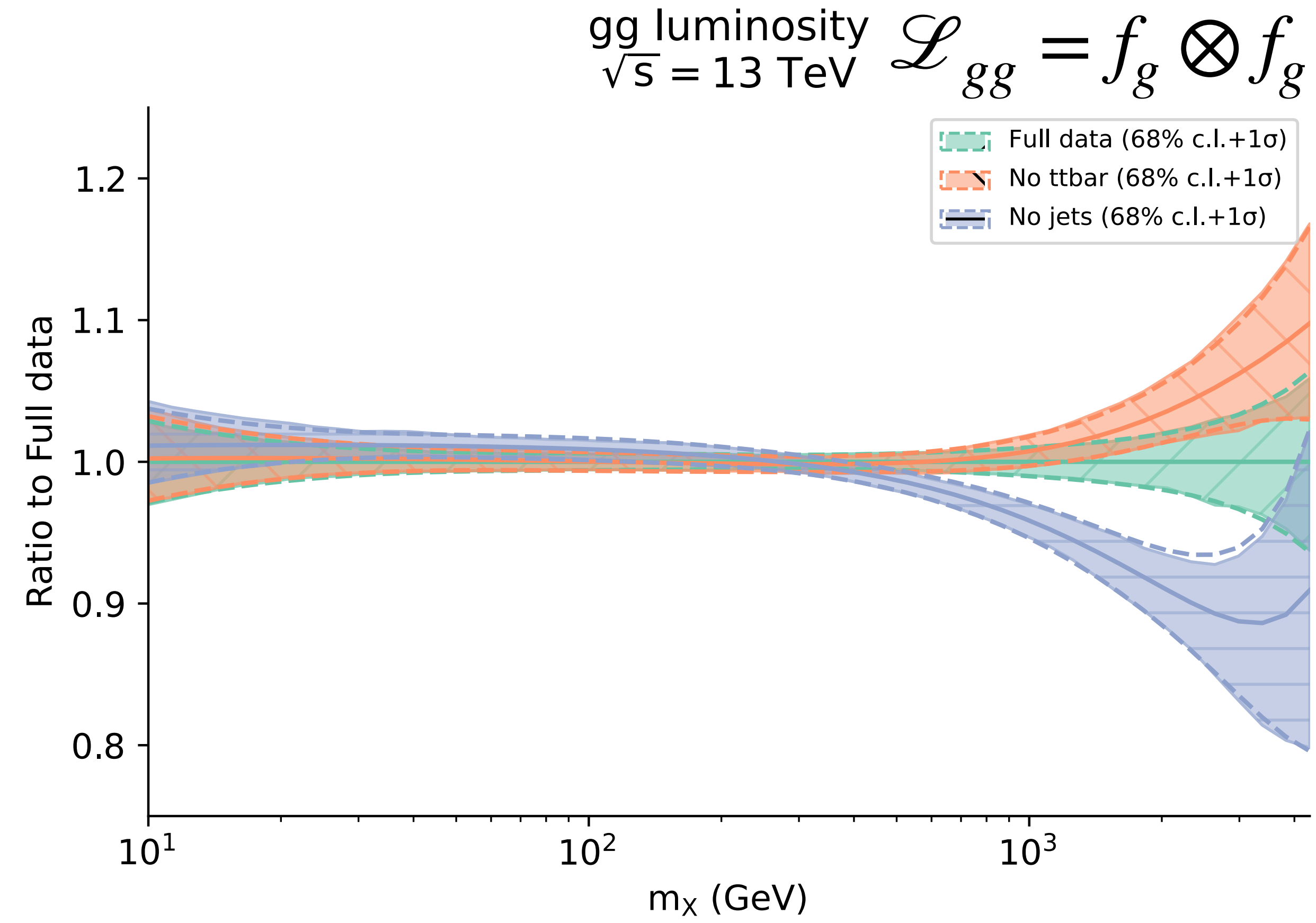
[Ball et al., NNPDF4.0, 2109.02653]

Incompatibility between top and jet data
Comparison of PDFs trained on different datasets

Full data kinematic coverage

PDFs' process dependance...

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

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

PDF contamination:

➡ **PDFs have absorbed new physics signals**

New physics scenario: W'

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

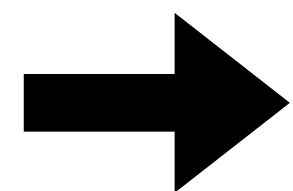## Generation of the pseudodata

$$\mathscr{L}_{UV}^{W'} = \mathscr{L}_{SM} - \frac{1}{4} W'^a_{\mu\nu} W'^{a,\mu\nu} + \frac{1}{2} M_{W'}^2 W'^a_\mu W'^{a,\mu} - g_{W'} W'^{a,\mu} \sum_{f_L} \bar{f}_L T^a \gamma^\mu f_L$$
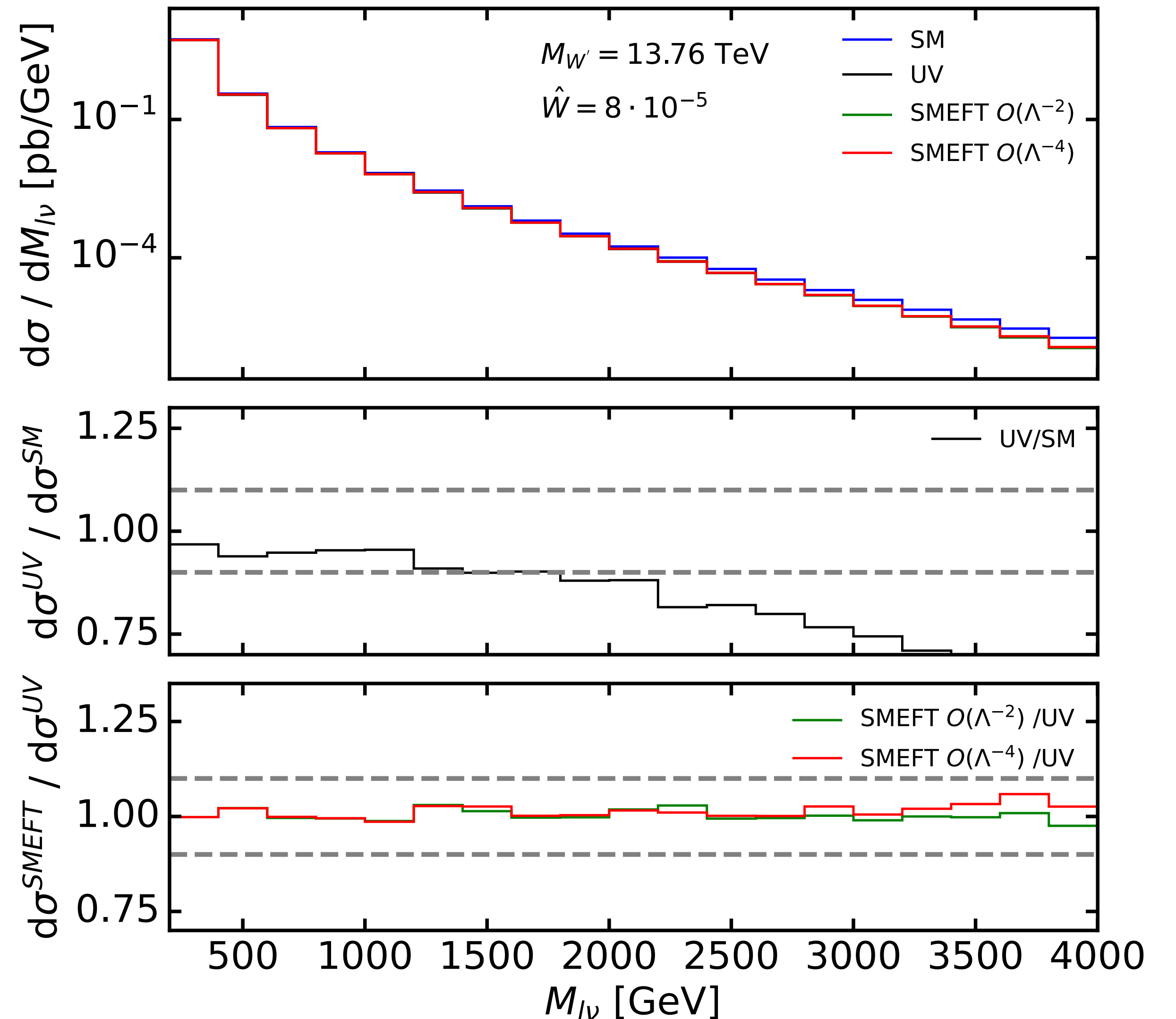
$\downarrow$

$$\mathscr{L}_{SMEFT}^{W'} = \mathscr{L}_{SM} - \frac{g_{W'}^2}{2 M_{W'}^2} J_L^{a,\mu} J_{L,\mu}^a$$

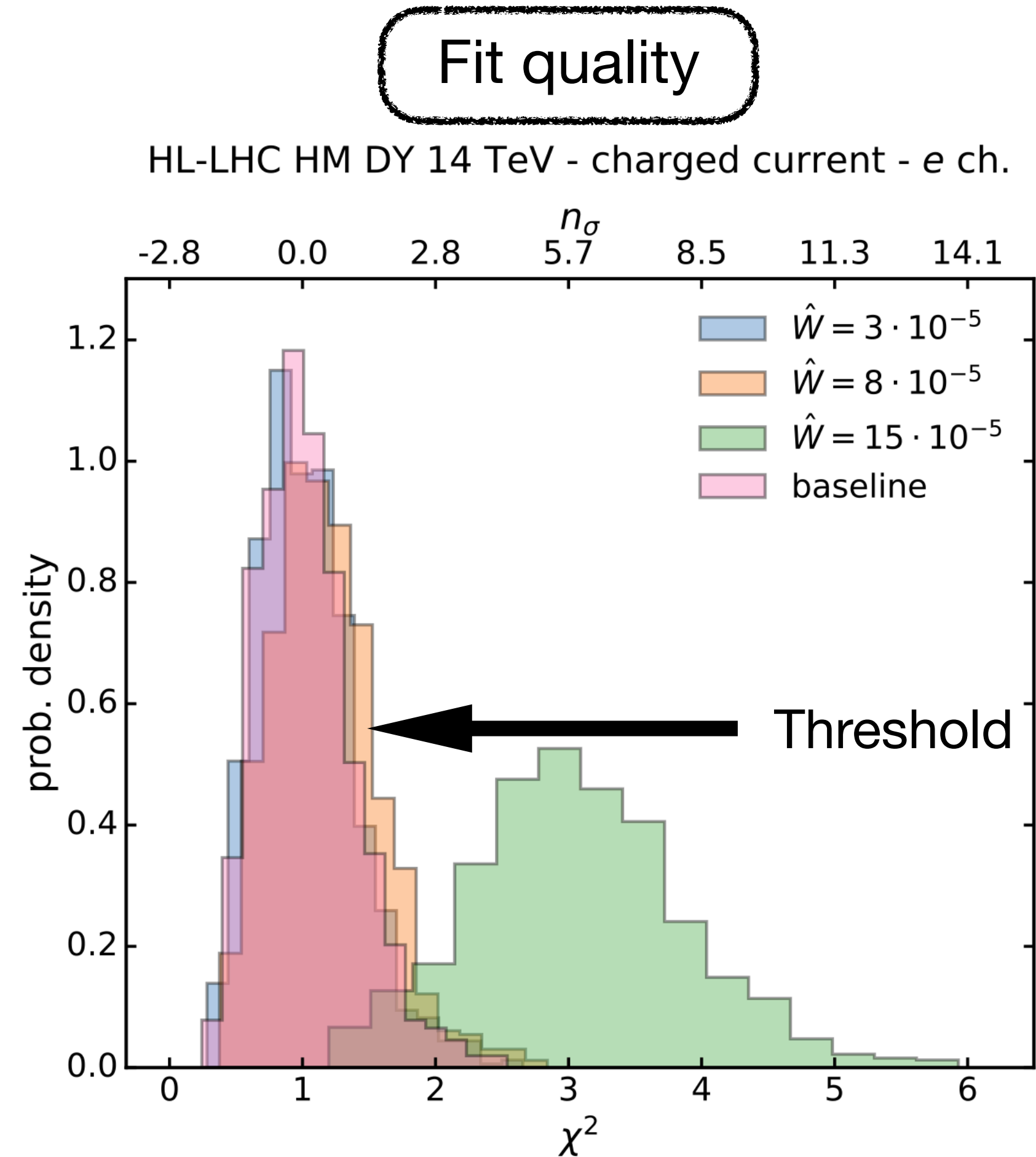$$J_L^{a,\mu} = \sum_{f_L} \bar{f}_L T^a \gamma^\mu f_L$$
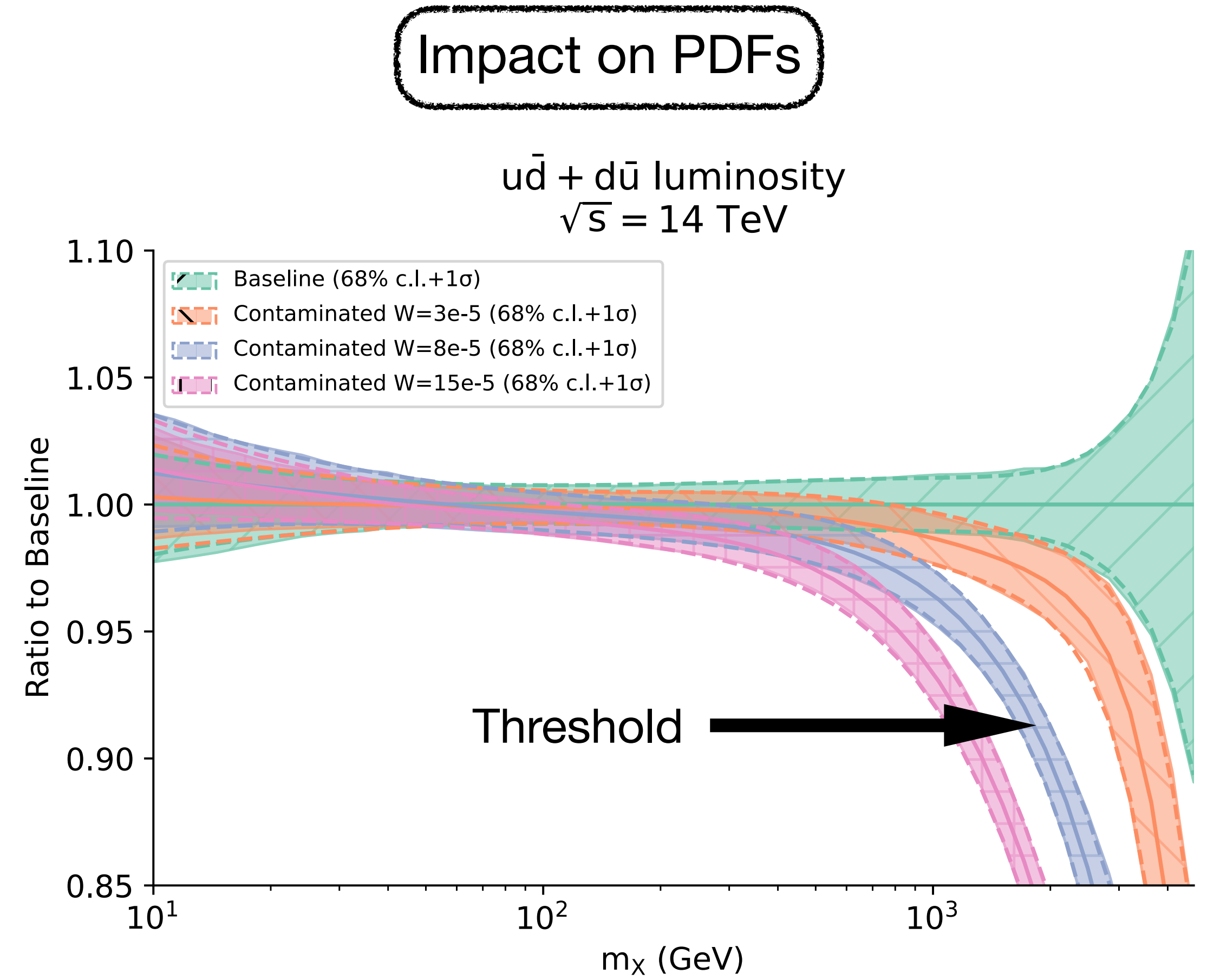
➡ Impacts Drell-Yan

HL-LHC Projections

$M_{W'} = 13.76$ TeV
$\hat{W} = 8 \cdot 10^{-5}$

SM
UV
SMEFT $O(\Lambda^{-2})$
SMEFT $O(\Lambda^{-4})$

UV/SM

SMEFT $O(\Lambda^{-2})$ /UV
SMEFT $O(\Lambda^{-4})$ /UV

5

Impact of contamination on the PDFs

Comparison between contaminated and Baseline PDFs

Fit quality

Impact on PDFs

HL-LHC HM DY 14 TeV - charged current - e ch.

$u\bar{d} + d\bar{u}$ luminosity
$\sqrt{s} = 14$ TeV
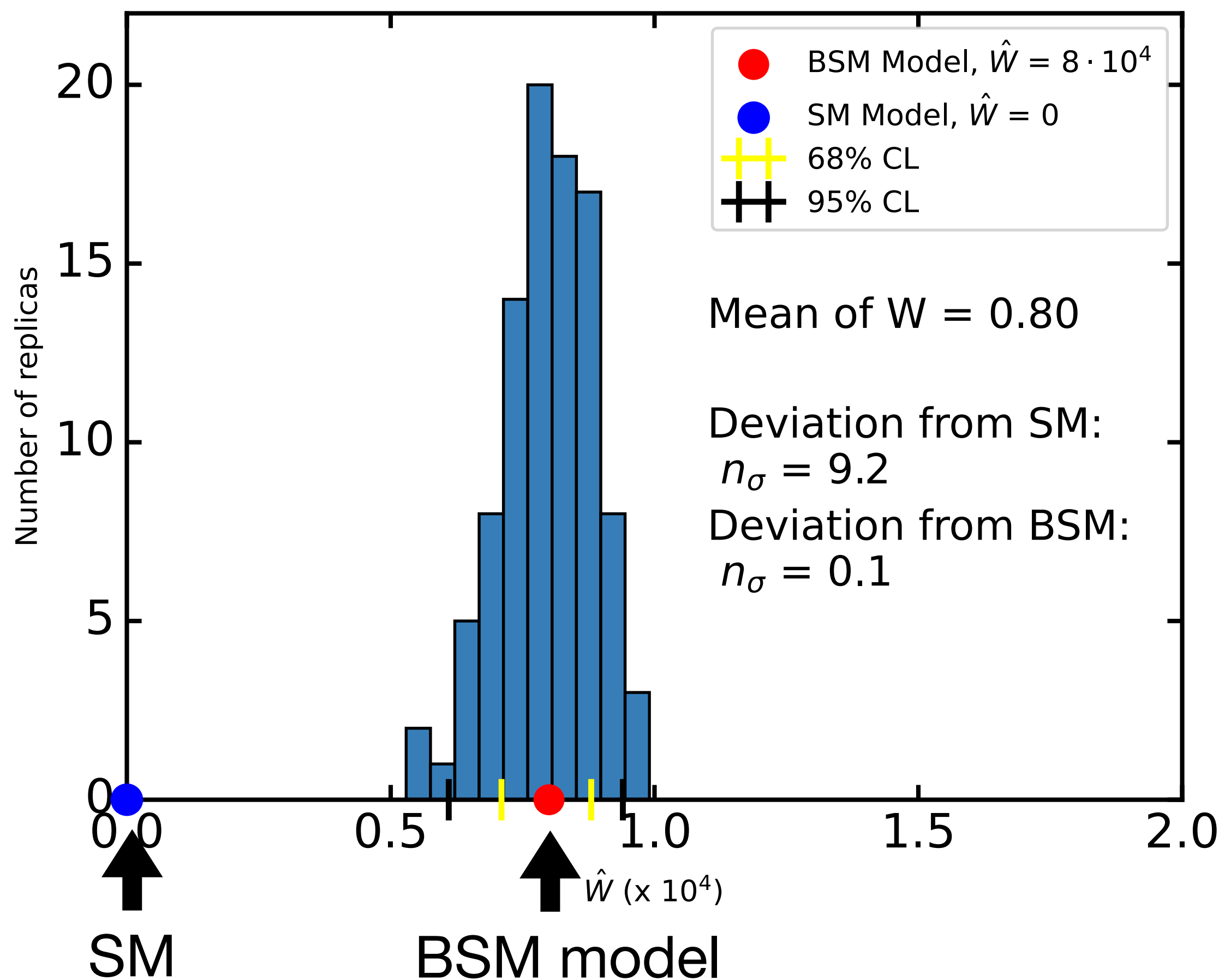
Threshold

Threshold

$M_{W'} = 13.8$ TeV

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

Missing new physics

Impact of the PDF contamination on SMEFT fits

SMEFT Fit with true PDF

- BSM Model, $\hat{W} = 8 \cdot 10^4$
- SM Model, $\hat{W} = 0$
- 68% CL
- 95% CL

Mean of W = 0.80

Deviation from SM:
$n_\sigma = 9.2$

Deviation from BSM:
$n_\sigma = 0.1$

SM   BSM model

$\hat{W}$ (x 10^4)

SMEFT Fit with contaminated PDF

- BSM Model, $\hat{W} = 8 \cdot 10^4$
- SM Model, $\hat{W} = 0$
- 68% CL
- 95% CL

Mean of W = 0.17

Deviation from SM:
$n_\sigma = 1.6$

Deviation from BSM:
$n_\sigma = 5.8$

True model excluded!

$\hat{W}$ (x 10^4)
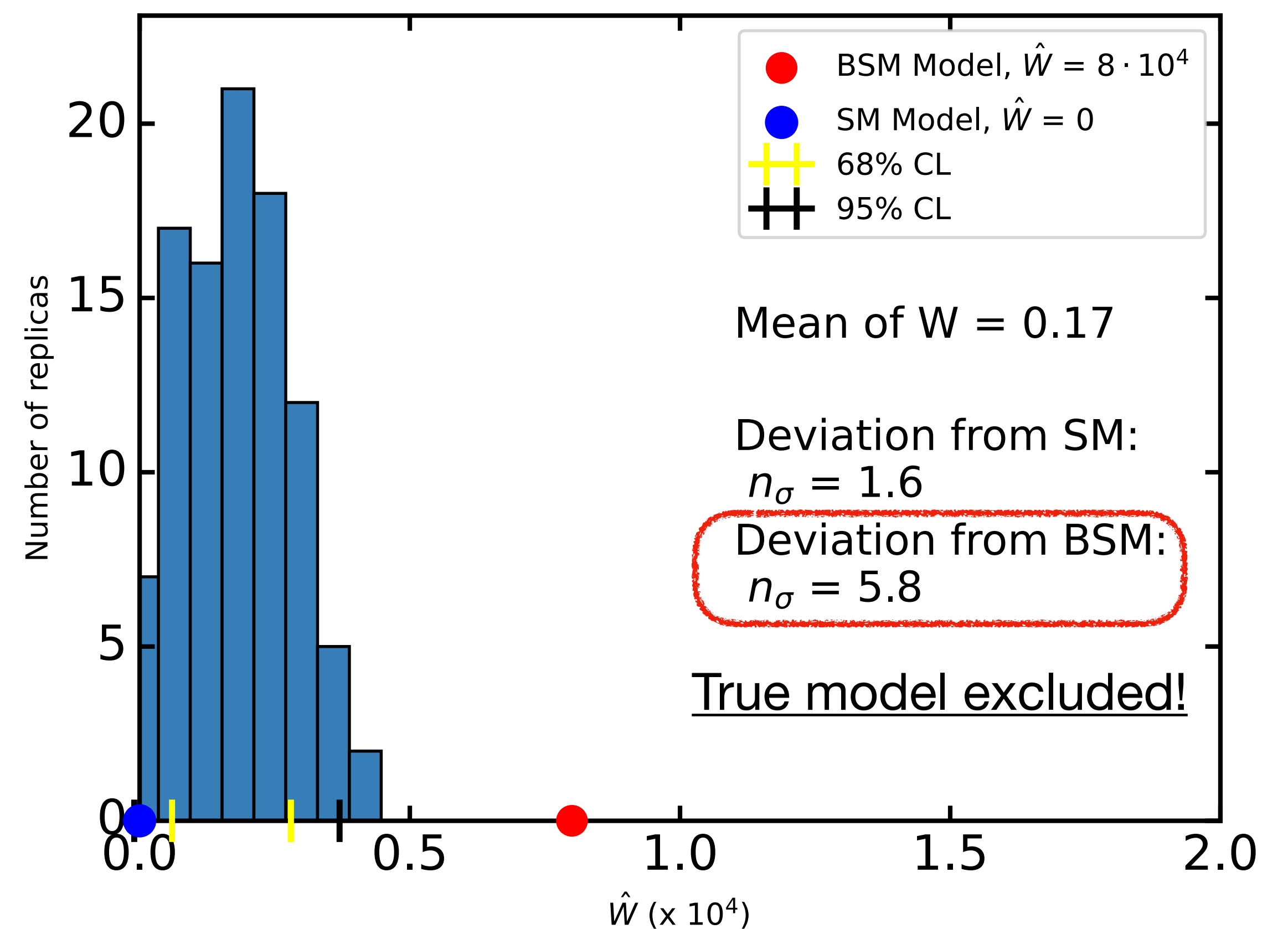
Apparition of fake deviations

Impact of contamination on predictions for other sectors
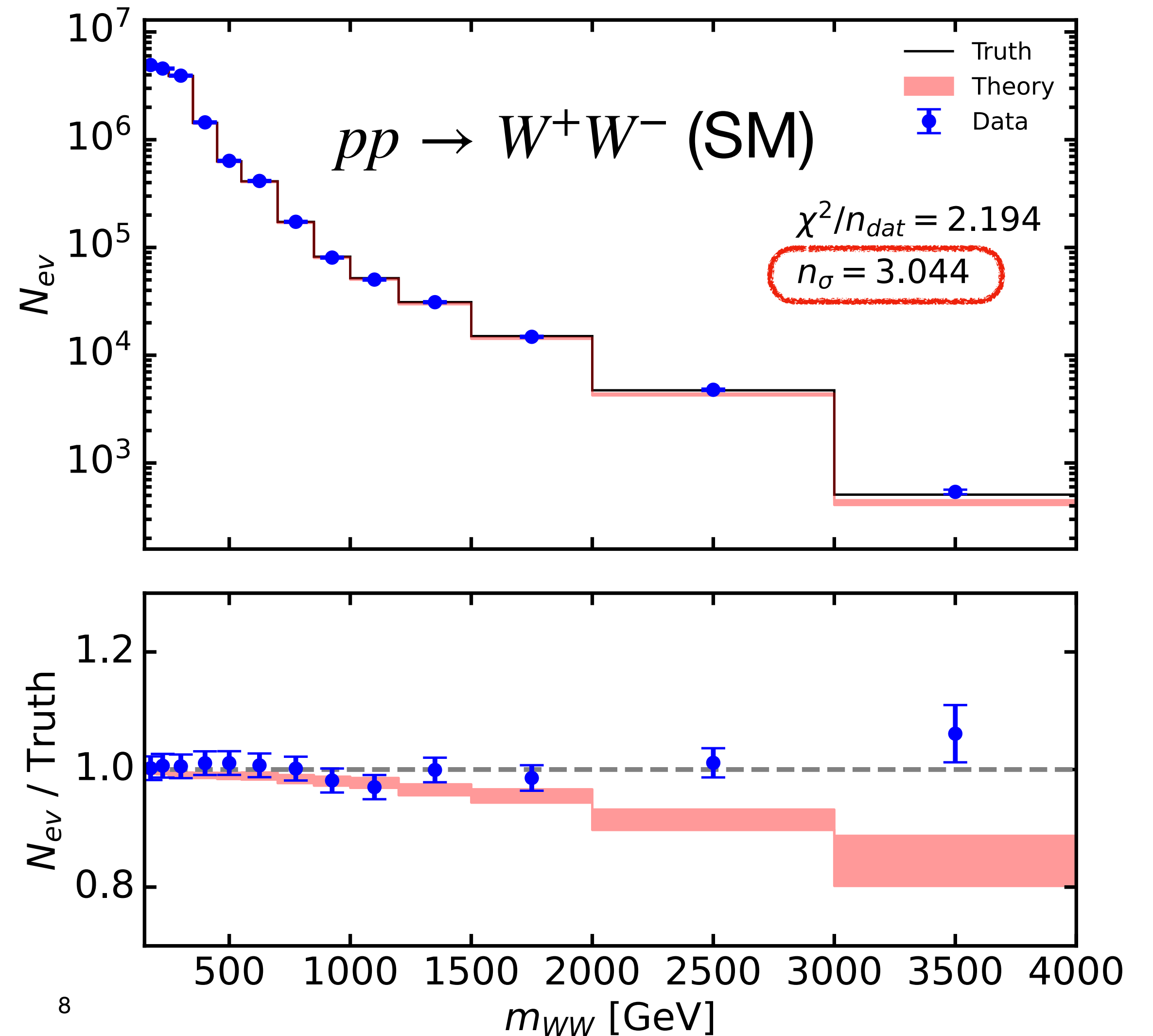
Theory predictions (red band):

- Contaminated PDFs + SM

Data (blue dots):

- True PDFs + SM

➡️ Fake deviation from SM

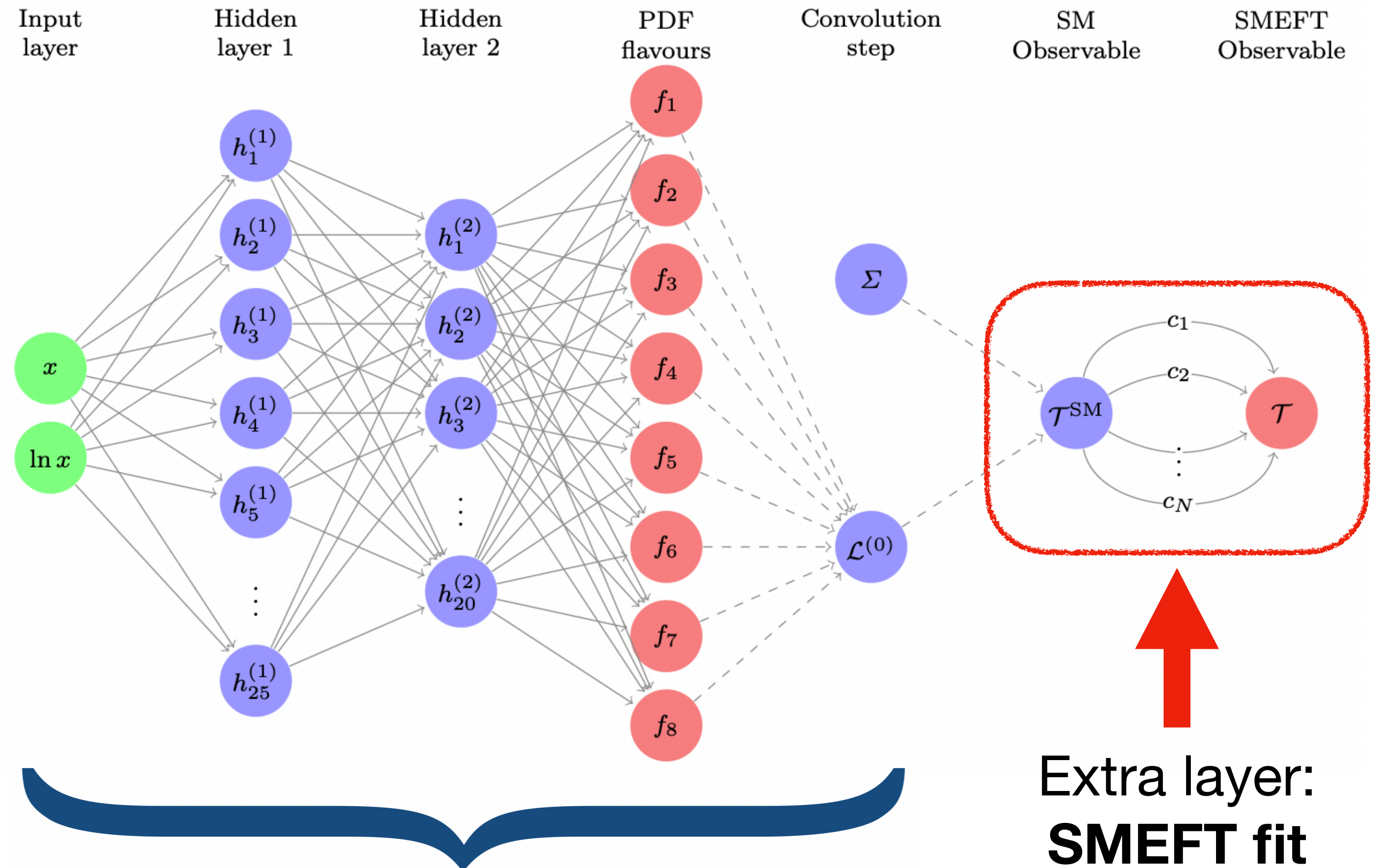Also seen in WH, WZ, ZH production

HL-LHC Projections

$pp \rightarrow W^+W^-$ (SM)

$\chi^2/n_{dat} = 2.194$

$n_\sigma = 3.044$

Truth
Theory
Data

N_{ev}

N_{ev} / Truth

m_{WW} [GeV]

Simultaneous fit of PDF and new physics

Presentation of the tool: SIMUnet

SIMUnet:

- Open-source tool:
  github.com/HEP-PBSP/SIMUnet
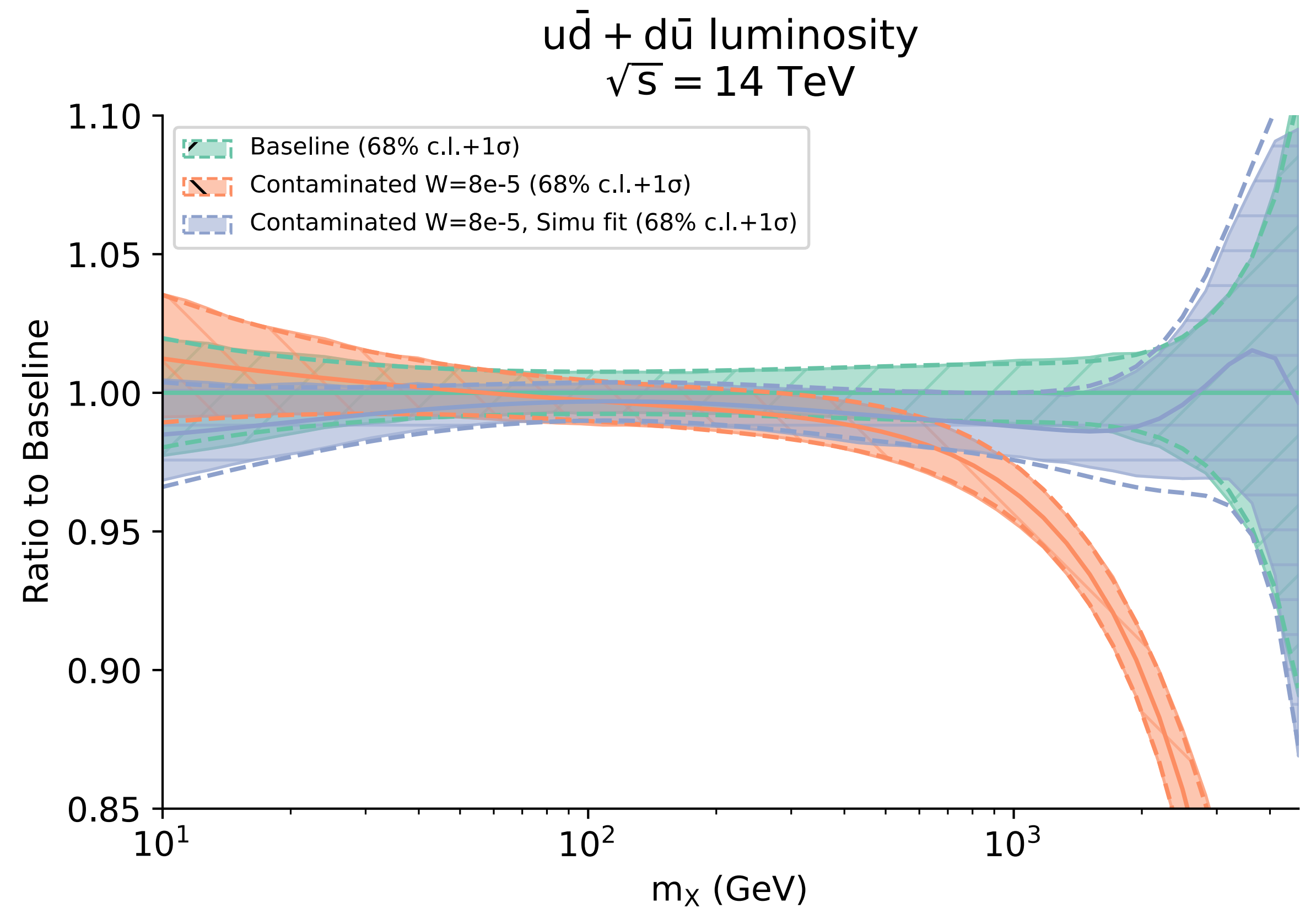
  [2402.03308]

- Fits PDFs and WC simultaneously

Input layer · Hidden layer 1 · Hidden layer 2 · PDF flavours · Convolution step · SM Observable · SMEFT Observable

NNPDF

Extra layer: **SMEFT fit**

[Iranipour et Ubiali, 2201.07240]

Simultaneous fit of PDF and new physics

Disentangling PDF contamination
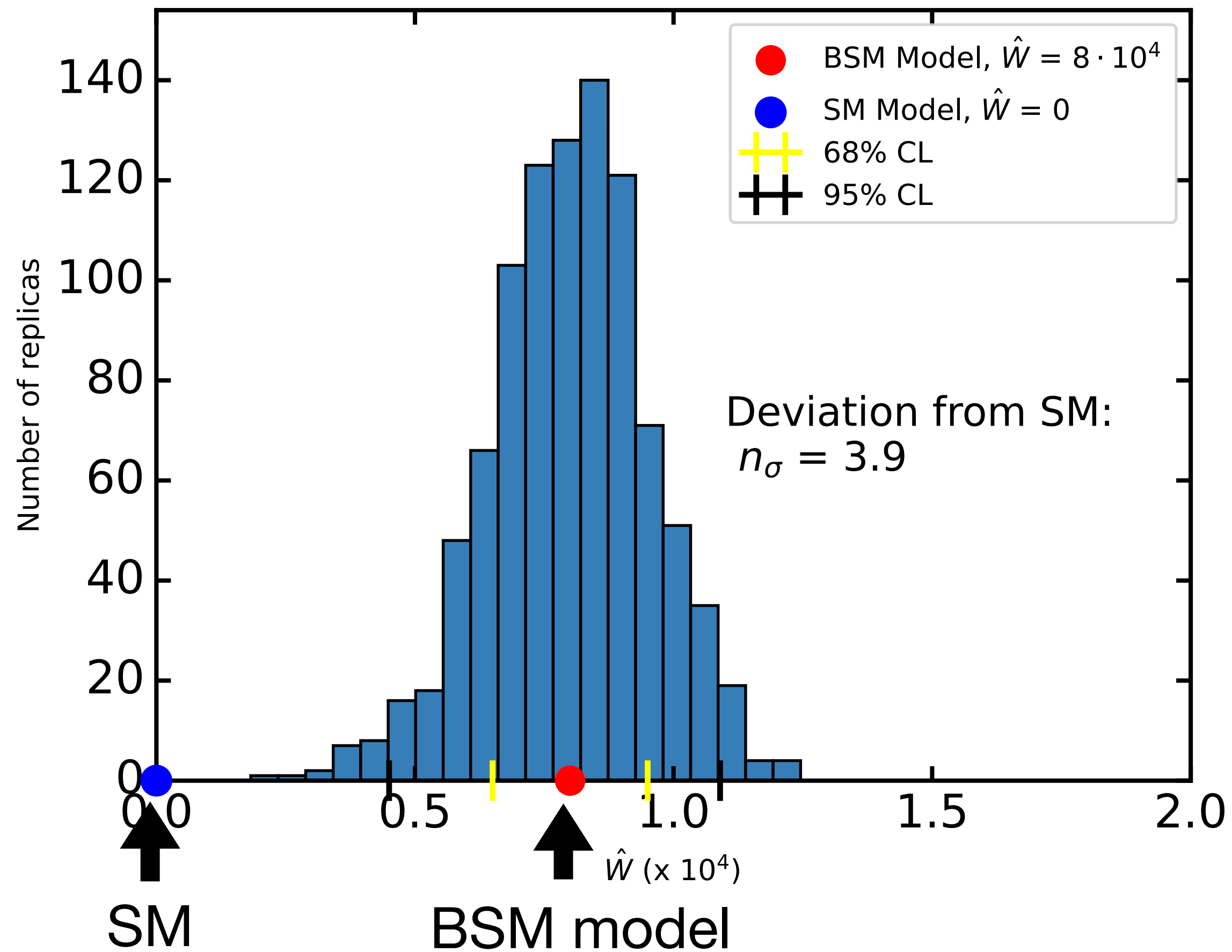
Summary

- Incompatibilities between different sectors in PDF fits

  ‣ PDF contamination?

- Signs of W' got fitted away in PDFs

  ‣ Missed new physics!

  ‣ Fake deviations in other sectors

- A solution to prevent contamination:

  ‣ Fitting simultaneously PDF and new physics: **SIMUnet** tool available
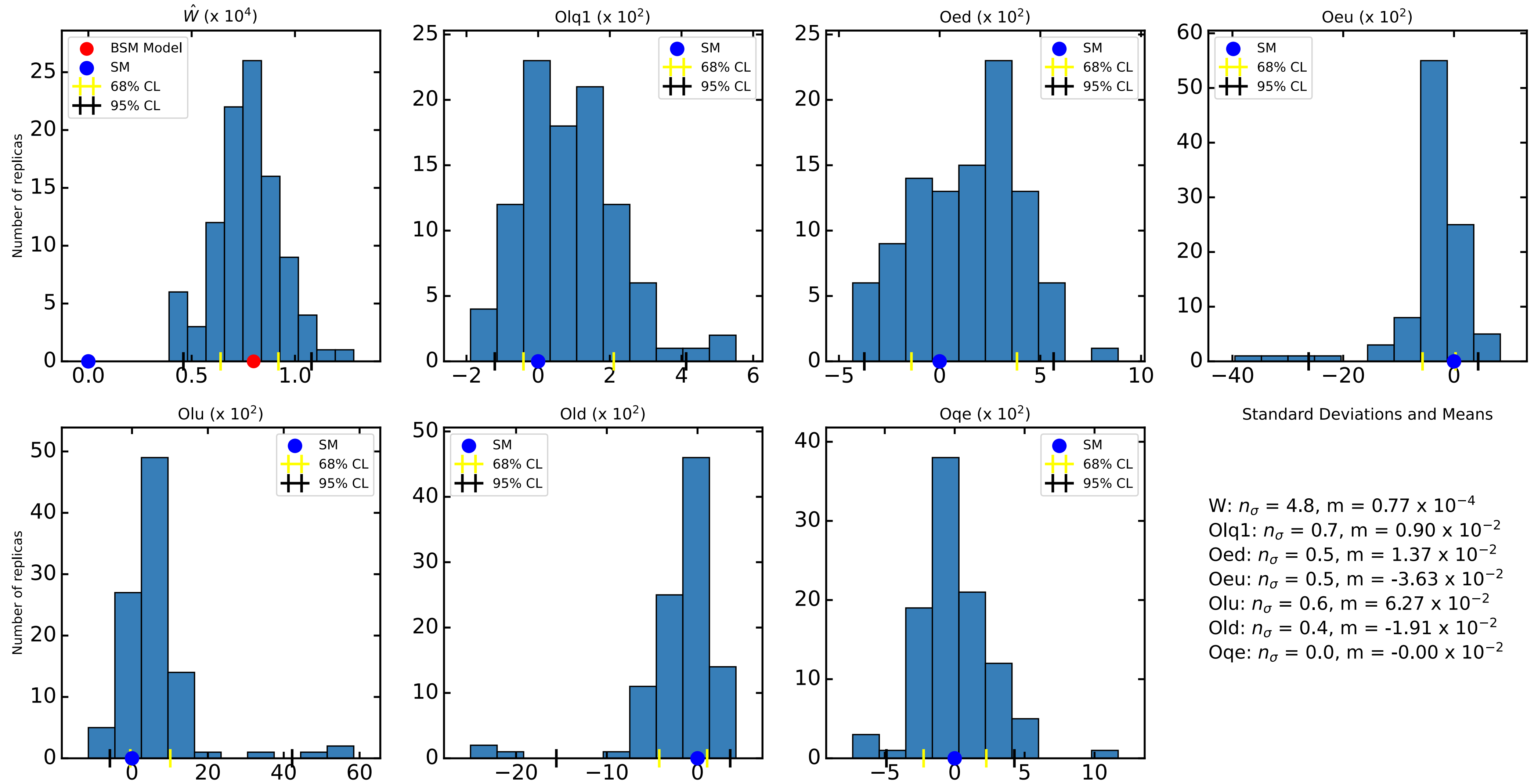
You can contact me at:
eh651@cam.ac.uk

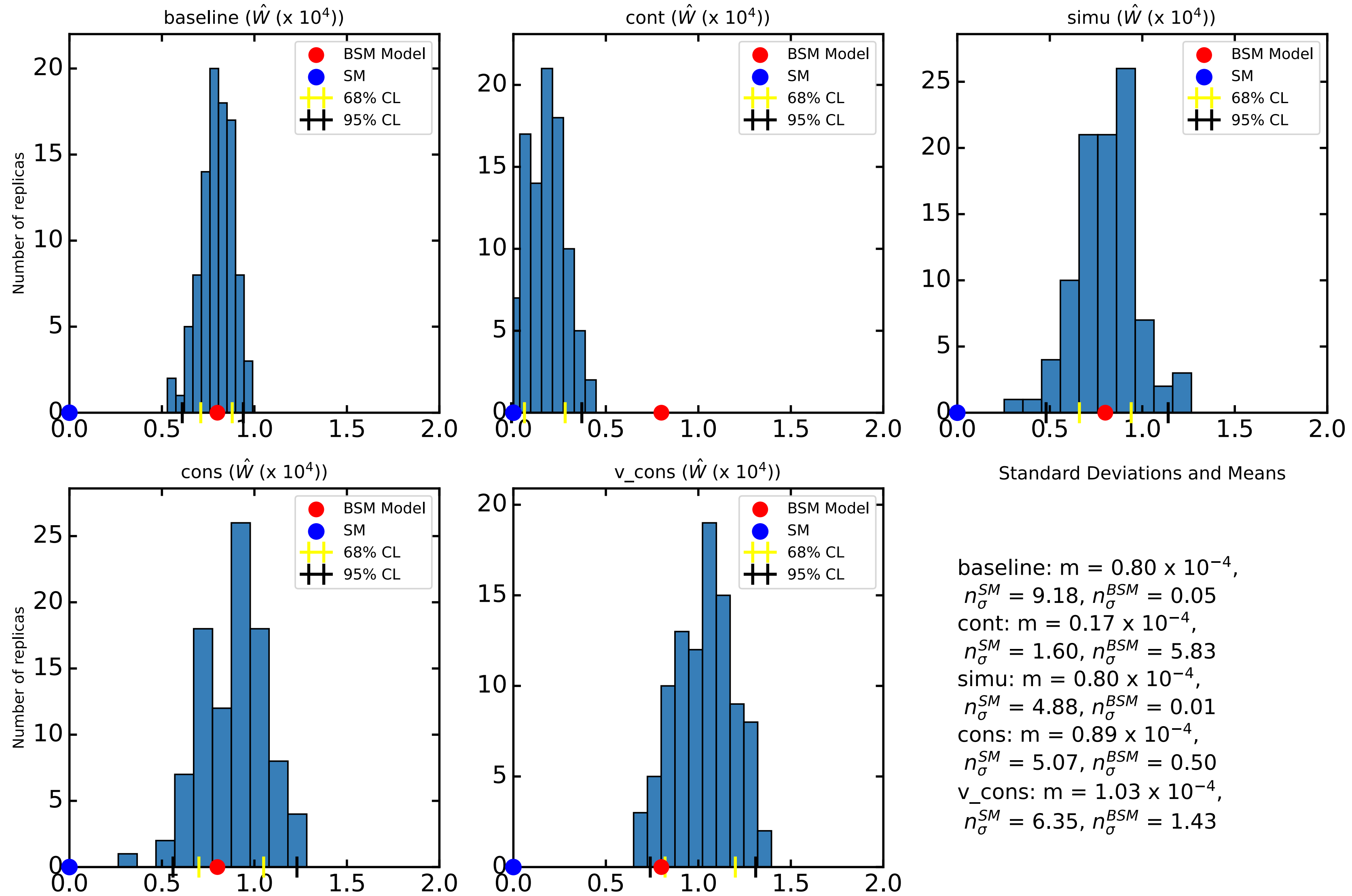**Thank you for your attention!**

Extra slides

List of deviations

Dataset	HL-LHC		Stat. improved	
	χ^2/n_{dat}	n_σ	χ^2/n_{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

Global SMEFT fit, 4 fermions operators

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

Synergy of high and low-energy data
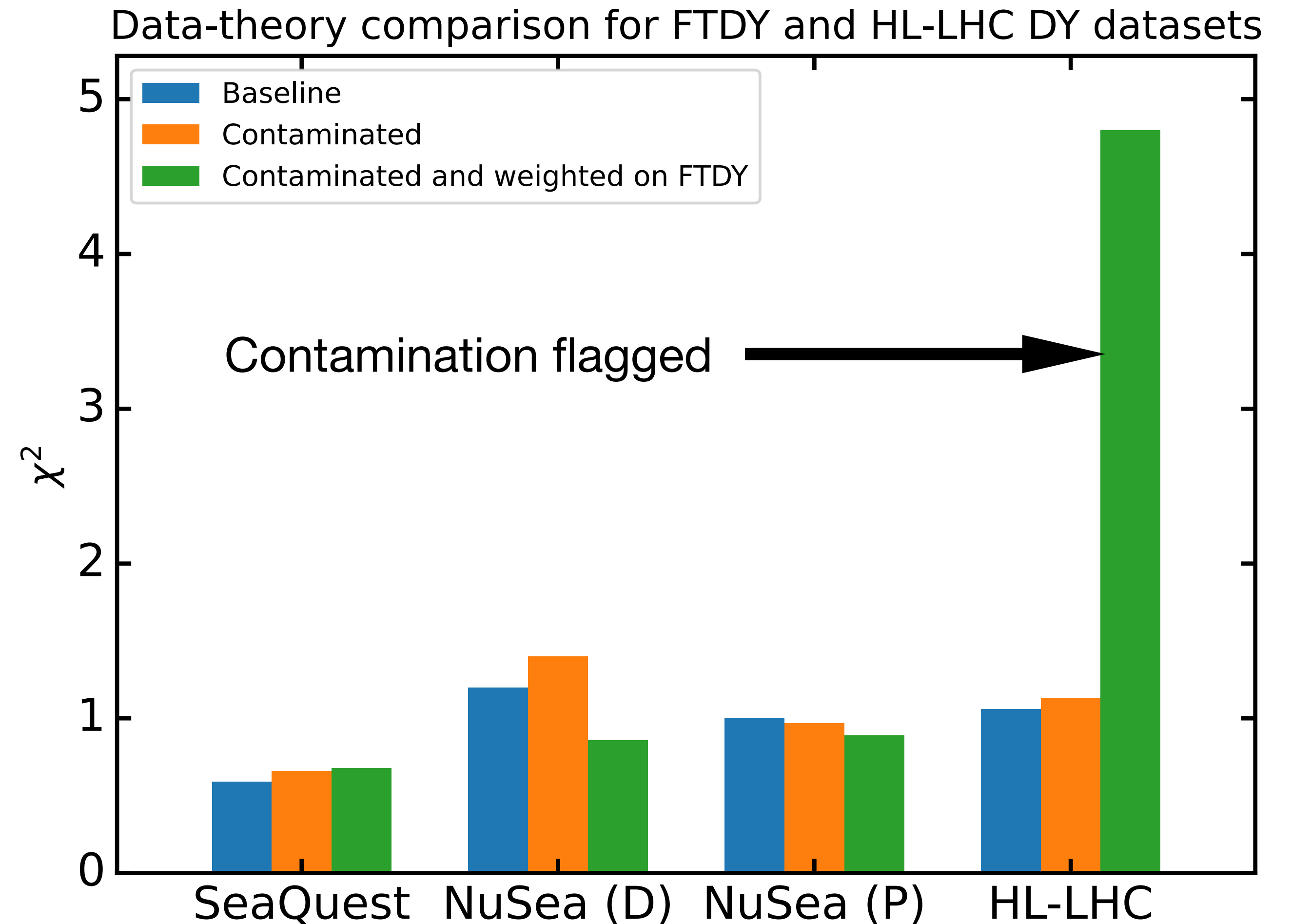## Adding low-energy dataset constraining the large-x region

Excessive antiquark PDF flexibility in large-x region:

➡ Accommodates real data and BSM pseudodata

➡ Allows contamination

Including low-energy large-x data:

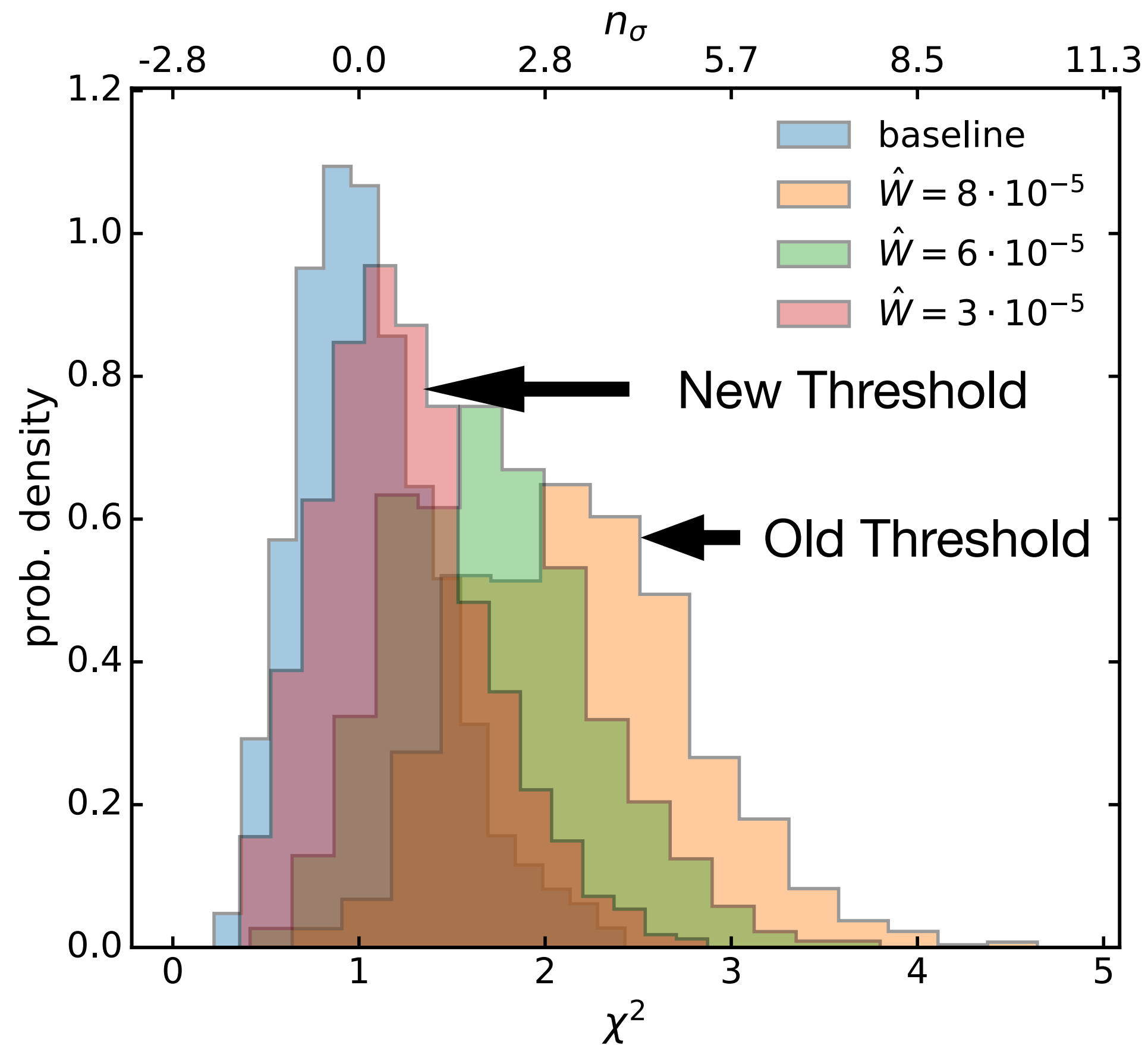- Constraint large-x region

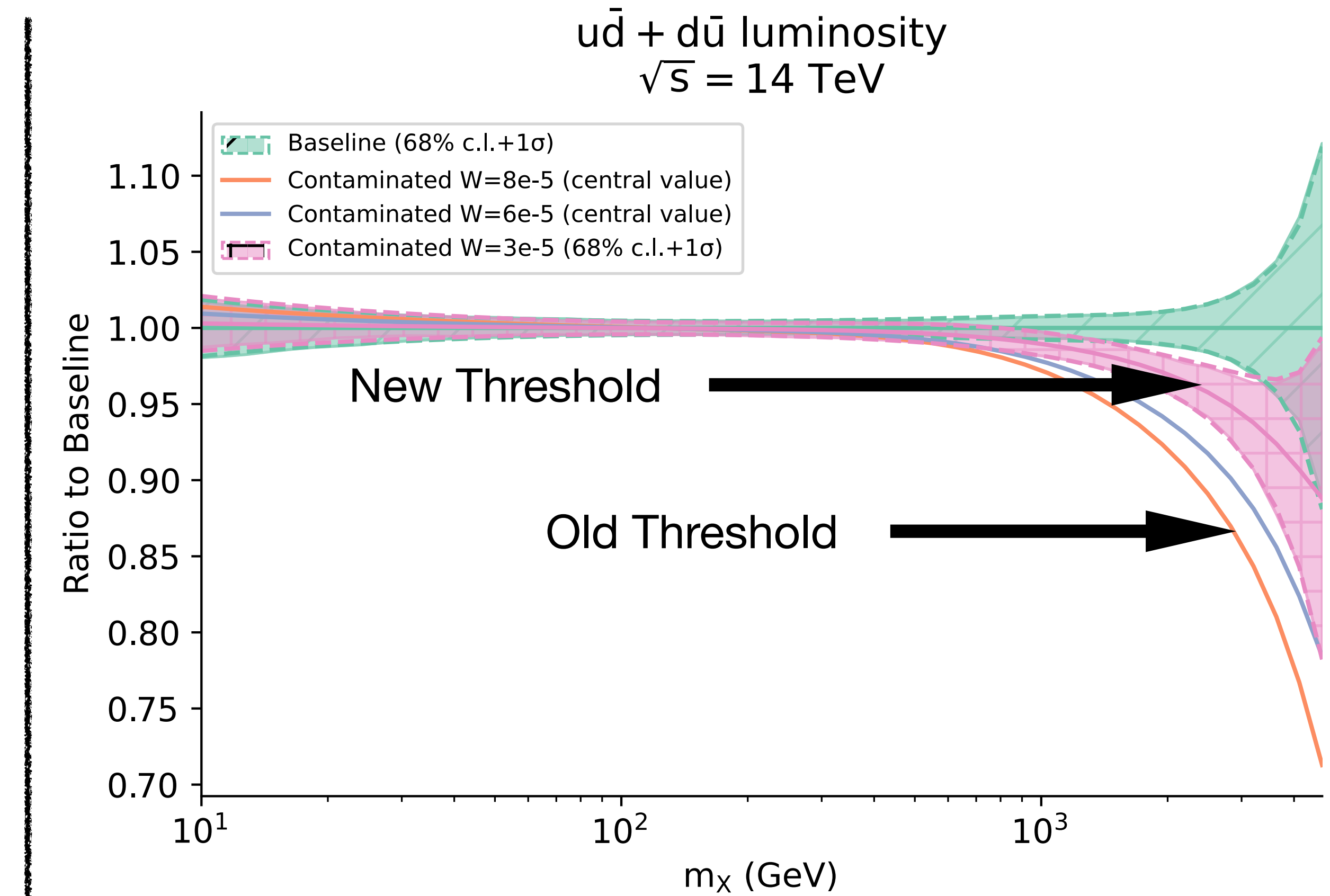- Safe from BSM contamination

[Hammou et Ubiali, incoming]

Data-theory comparison for FTDY and HL-LHC DY datasets

- Baseline
- Contaminated
- Contaminated and weighted on FTDY

Contamination flagged →

χ^2

SeaQuest   NuSea (D)   NuSea (P)   HL-LHC

Impact of FPF data on PDF contamination

Projection data from neutrino DIS at the LHC

HL-LHC HM DY 14 TeV - charged current - electron channel

Legend: baseline, $\hat{W} = 8 \cdot 10^{-5}$, $\hat{W} = 6 \cdot 10^{-5}$, $\hat{W} = 3 \cdot 10^{-5}$

New Threshold

Old Threshold

$M_{W'} : 13.8 \to 22.5$ TeV

$u\bar{d} + d\bar{u}$ luminosity
$\sqrt{s} = 14$ TeV

Legend:
- Baseline (68% c.l.+1σ)
- Contaminated W=8e-5 (central value)
- Contaminated W=6e-5 (central value)
- Contaminated W=3e-5 (68% c.l.+1σ)

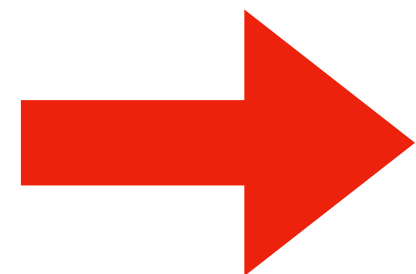New Threshold

Old Threshold

Reduces fake deviations

New physics scenarios: Z'

$$M_{Z'} = 18.7 \text{ TeV}$$
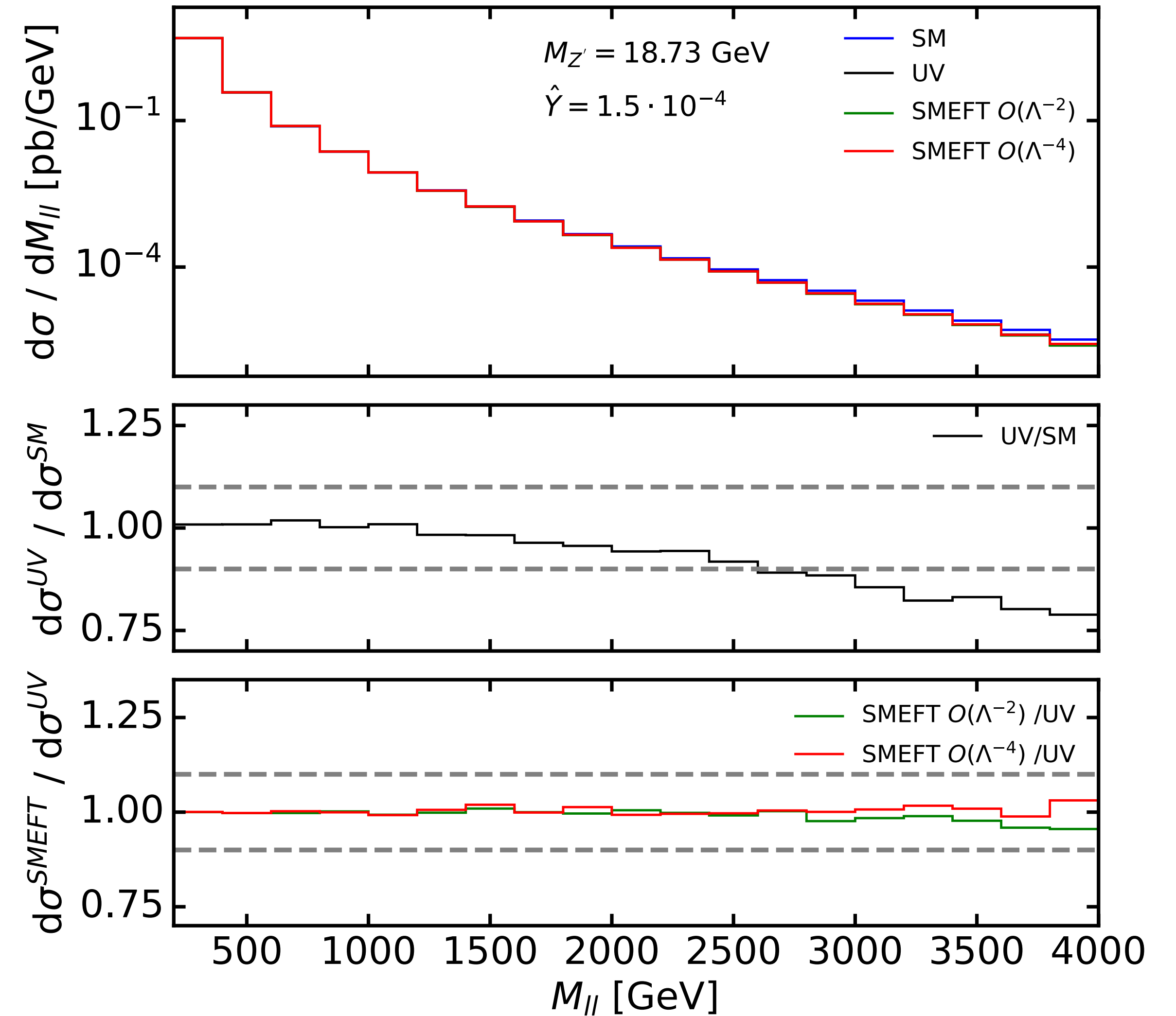
Generation of the pseudodata

$$\mathcal{L}^{Z'}_{SMEFT} = \mathcal{L}_{SM} - \frac{g^2_{Z'}}{2M^2_{Z'}} J^\mu_Y J_{Y,\mu}$$
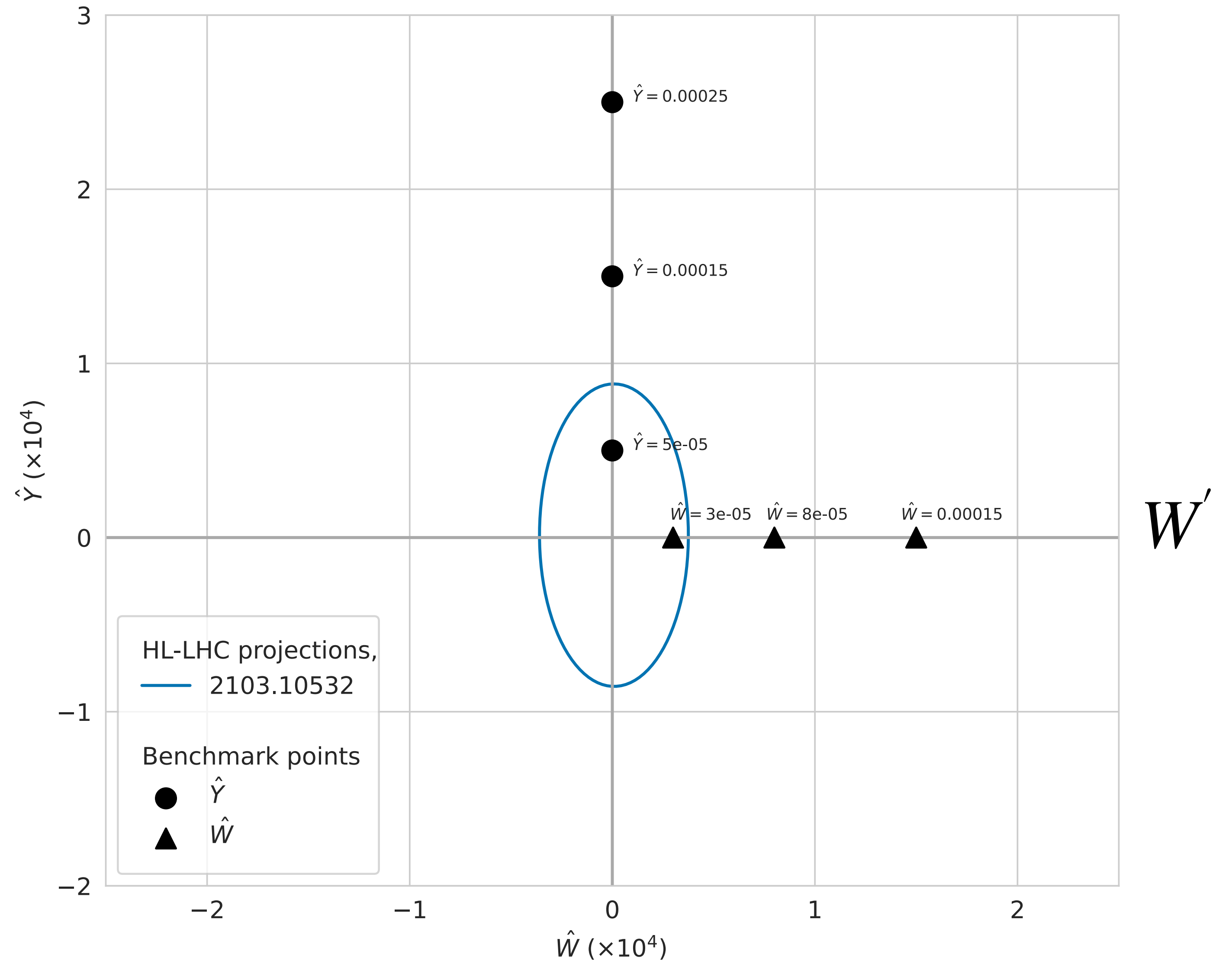
$$J^\mu_Y = \sum_f Y_f \bar{f} \gamma^\mu f$$

Impacts neutral current Drell-Yan processes

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

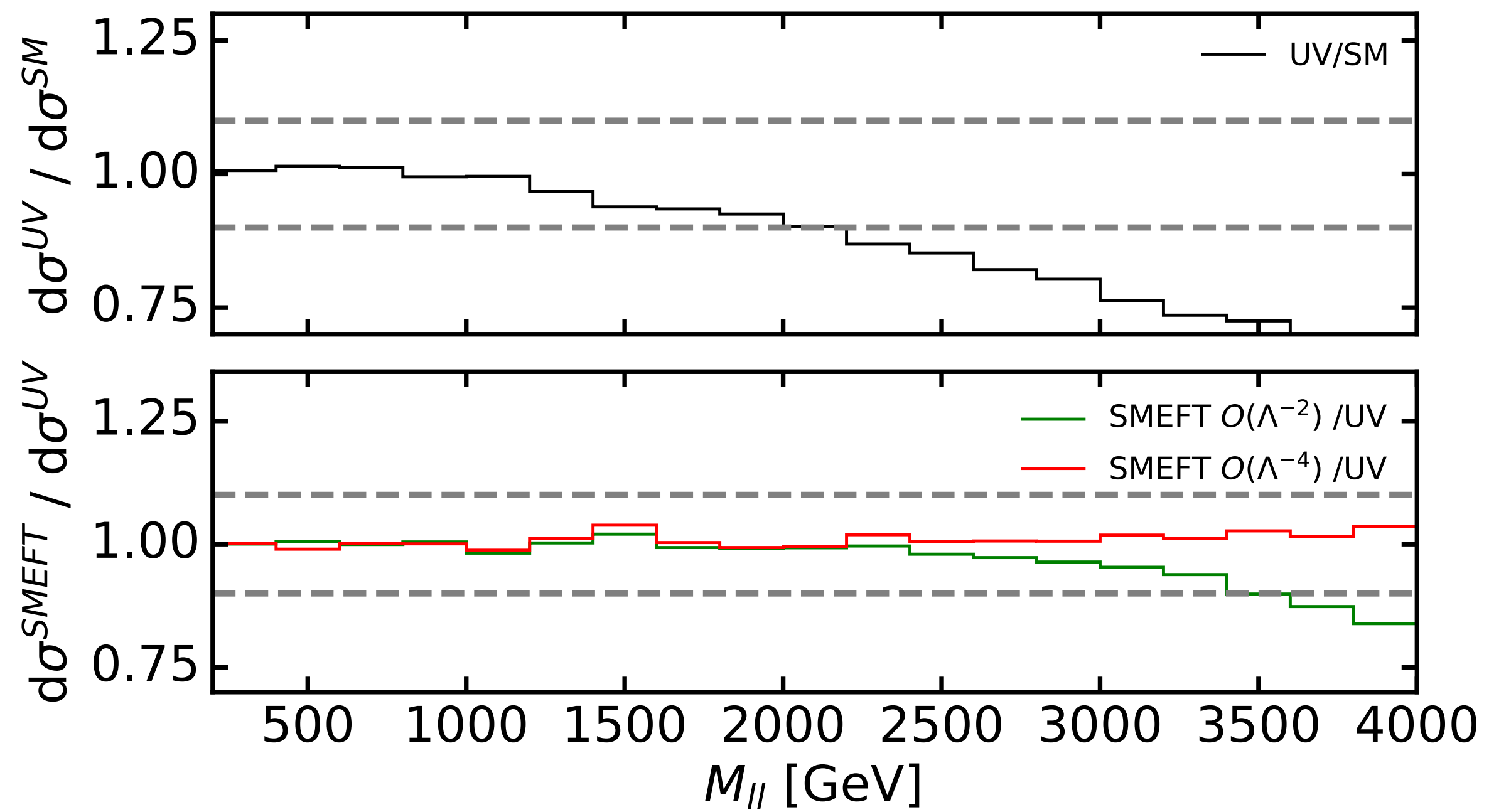# Constraints from current data

Z'
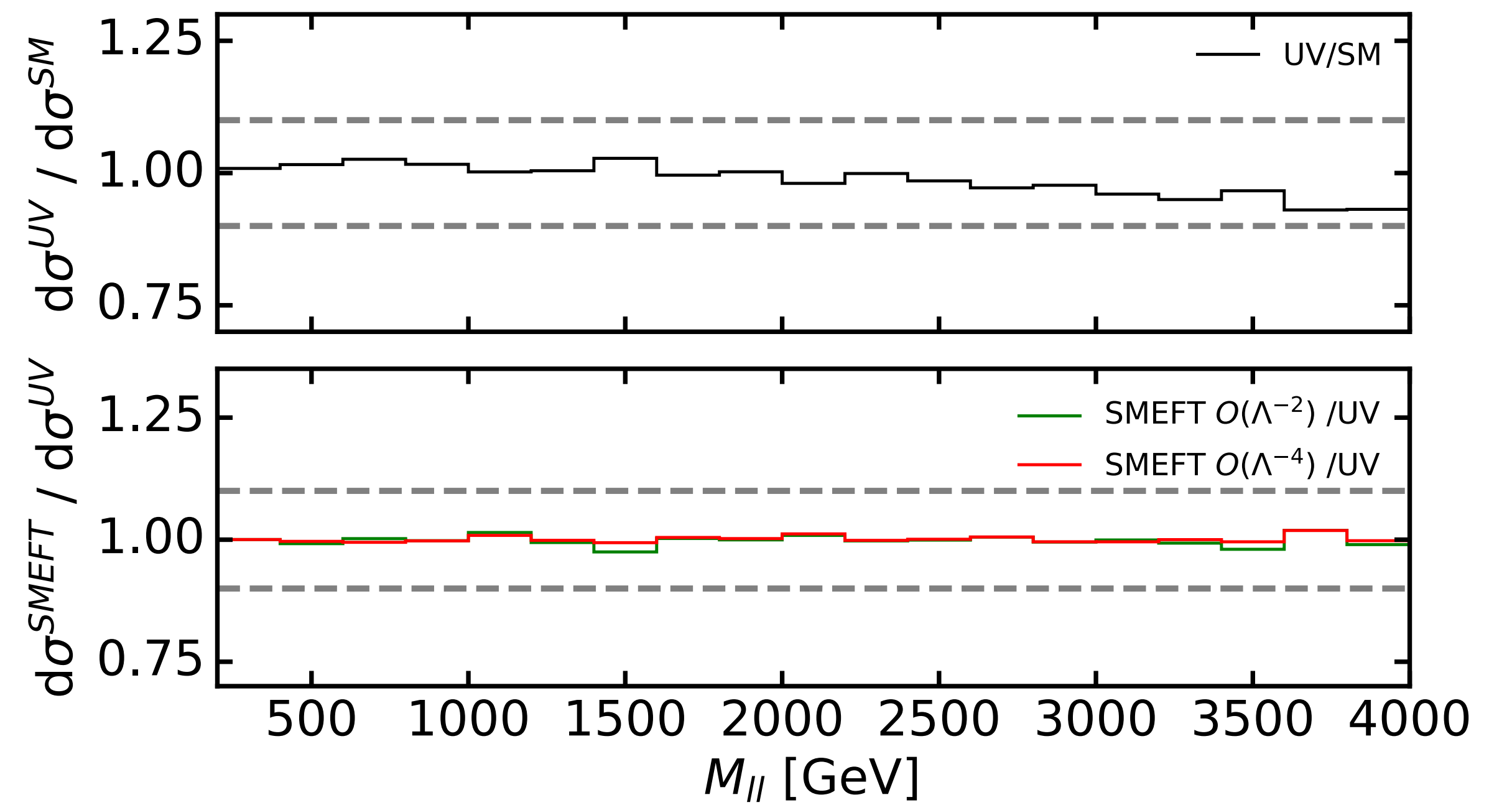
- New physics scenarios compared to constraints at 95% CL
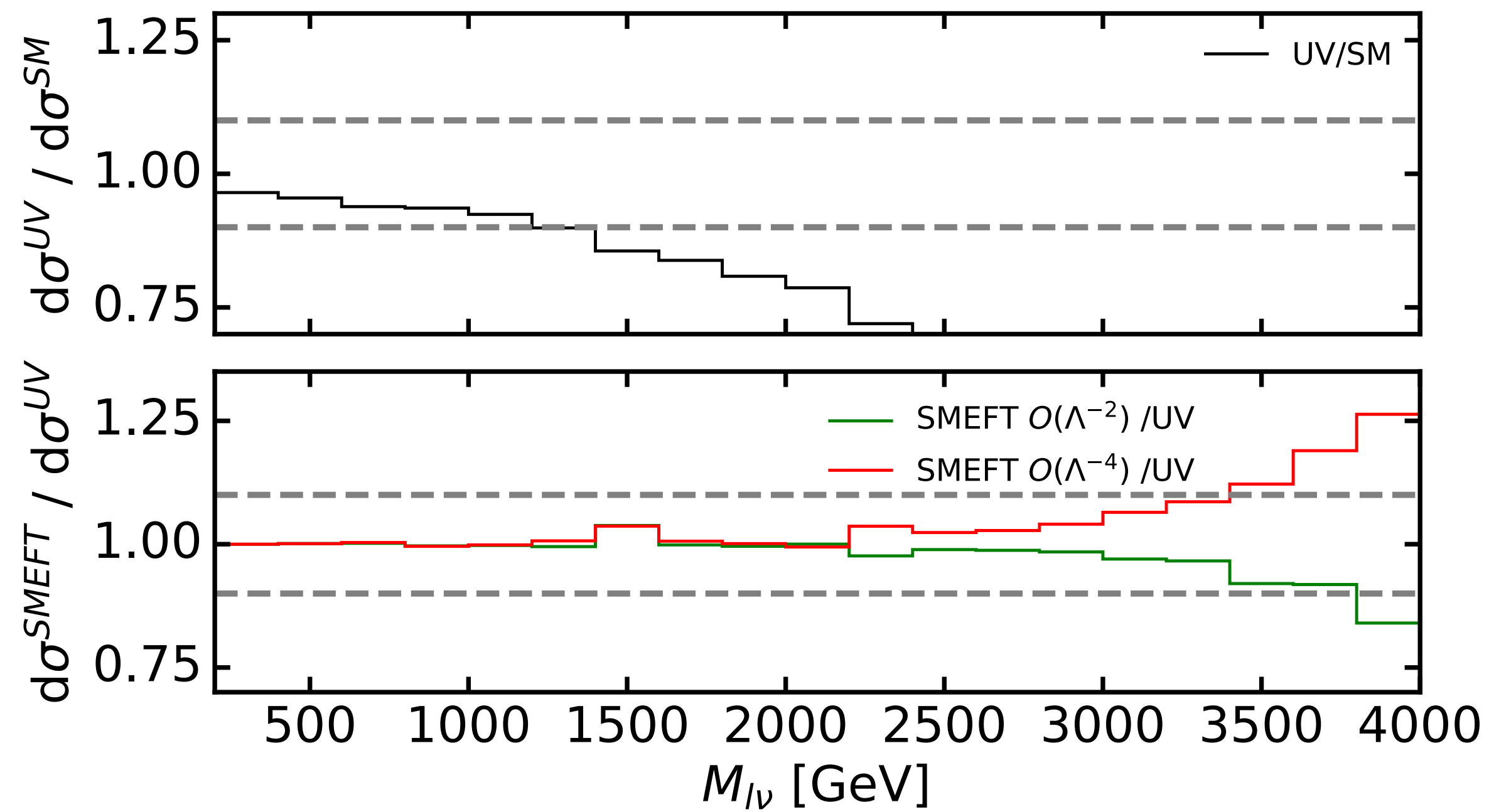
New physics scenarios: Z'

$M_{Z'} = 14.5$ TeV

$M_{Z'} = 32.5$ TeV

UV/SM

UV/SM

SMEFT $O(\Lambda^{-2})$ /UV
SMEFT $O(\Lambda^{-4})$ /UV

SMEFT $O(\Lambda^{-2})$ /UV
SMEFT $O(\Lambda^{-4})$ /UV

$\mathrm{d}\sigma^{UV} / \mathrm{d}\sigma^{SM}$

$\mathrm{d}\sigma^{SMEFT} / \mathrm{d}\sigma^{UV}$

M_{ll} [GeV]

New physics scenarios: W'