Disentangling PDFs and signs of new physics

Based on:

- [PBSP, 2307.10370, JHEP]
- [Hammou et Ubiali, 2410.00963, PRD]
- [PBSP, forthcoming]

Elie Hammou

Theory Day, Nijmegen, Oct 2025

Talk overview

1. Background on SMEFT and PDFs
2. Risk assessment: absorbing signs of new physics in PDFs [PBSP, 2307.10370, JHEP]
3. Solution 1: Bridging “blind spot” in dataset [Hammou et Ubiali, 2410.00963, PRD]
4. Solution 2: Simultaneous fit of PDFs and SMEFT [PBSP, forthcoming]
5. Conclusions and future prospects

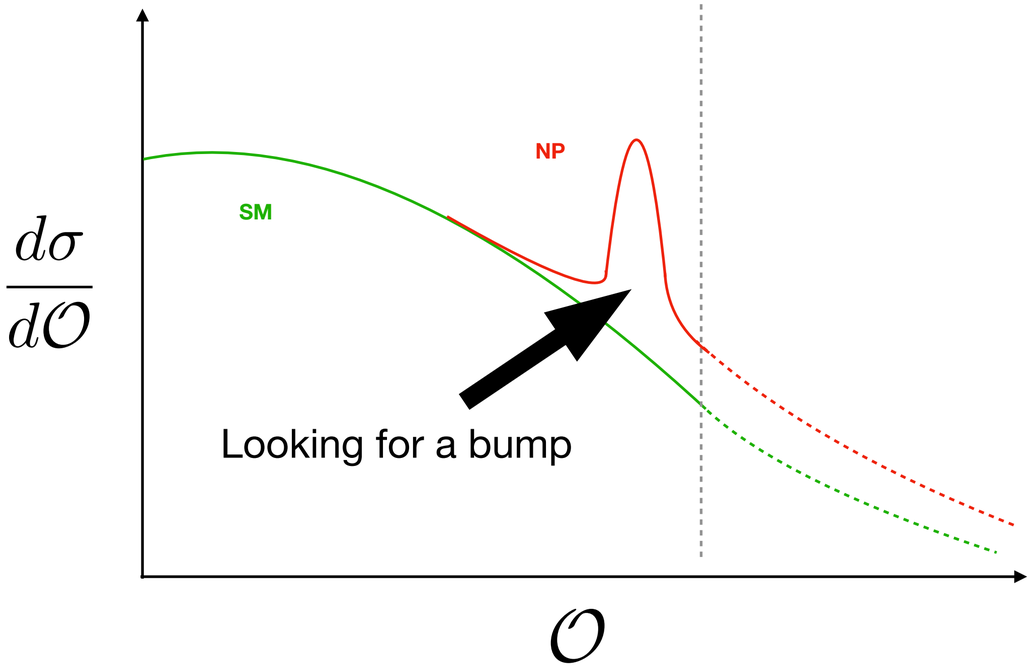
Background on SMEFT and PDFs

New physics searches

Looking toward higher energy scales and indirect searches

Direct searches

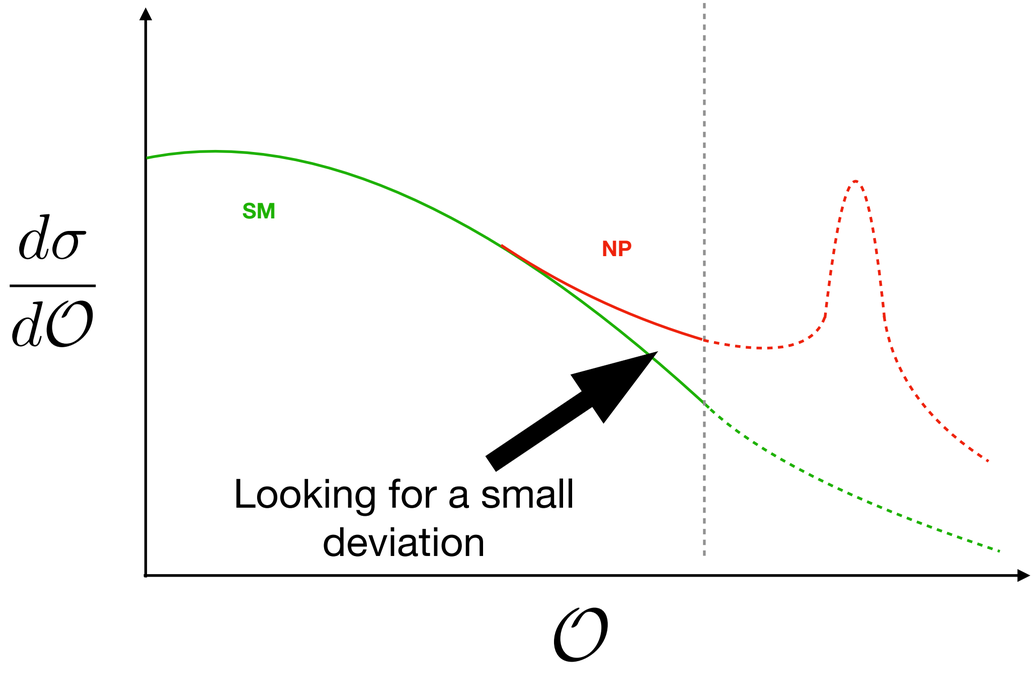
$E_{NP} < E_{collider}$



No luck so far…

Indirect searches

$E_{NP} > E_{collider}$



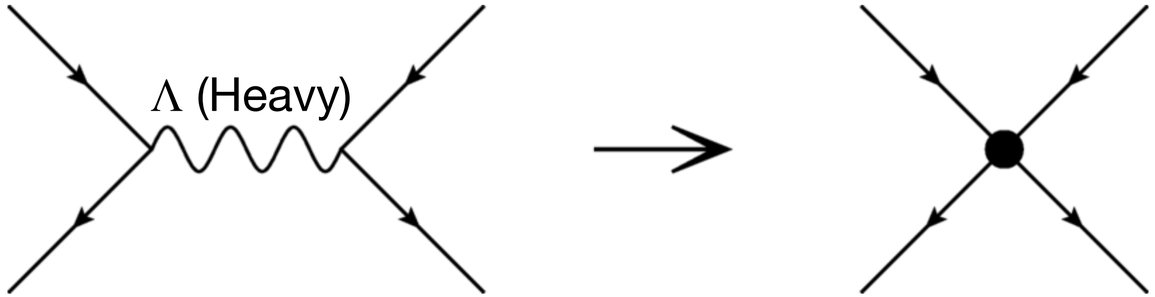
Requires precision

Indirect searches and Effective Field Theories

The Standard Model EFT (SMEFT)

Integrate heavy fields out:

UV Model



EFT Model

[10.1007/s10773-021-04723-1]

Obtain model independent Lagrangian:

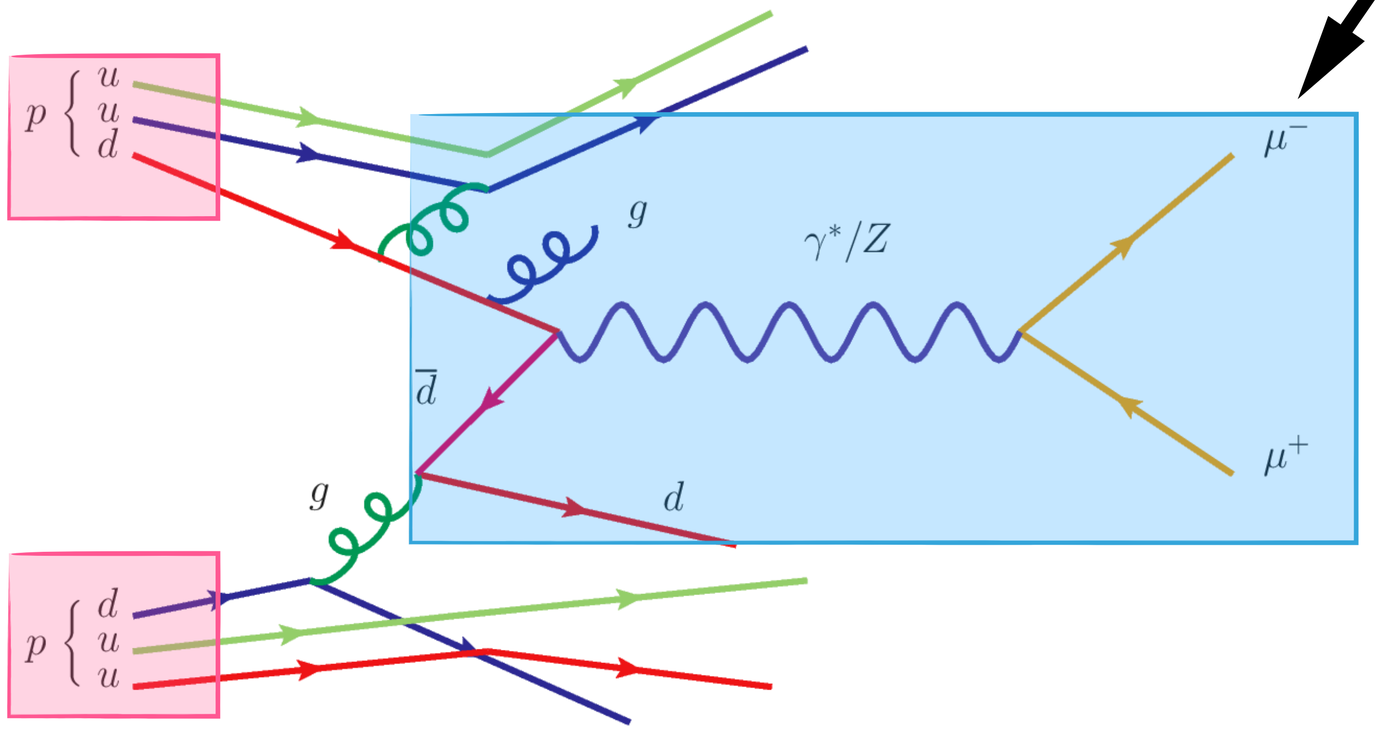
$$\mathscr{L}^{\text{UV}} = \mathscr{L}^{\text{SM}} + \mathscr{L}^{\text{Heavy}} \longrightarrow \mathscr{L}^{\text{SMEFT}} = \mathscr{L}^{\text{SM}} + \sum_i \frac{c_i}{\Lambda^2} \mathcal{O}_i^{(6)} + \dots$$

- Dim 6 EFT operators with SM fields: $\mathcal{O}_i^{(6)}$
- Wilson coefficients fittable from data: $\frac{c_i}{\Lambda^2}$

Hadron colliders and PDFs

Collinear factorization theorem

SMEFT corrections



Parton Distribution Functions

Hard Scattering: Perturbative QCD + EW

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

PDFs overview

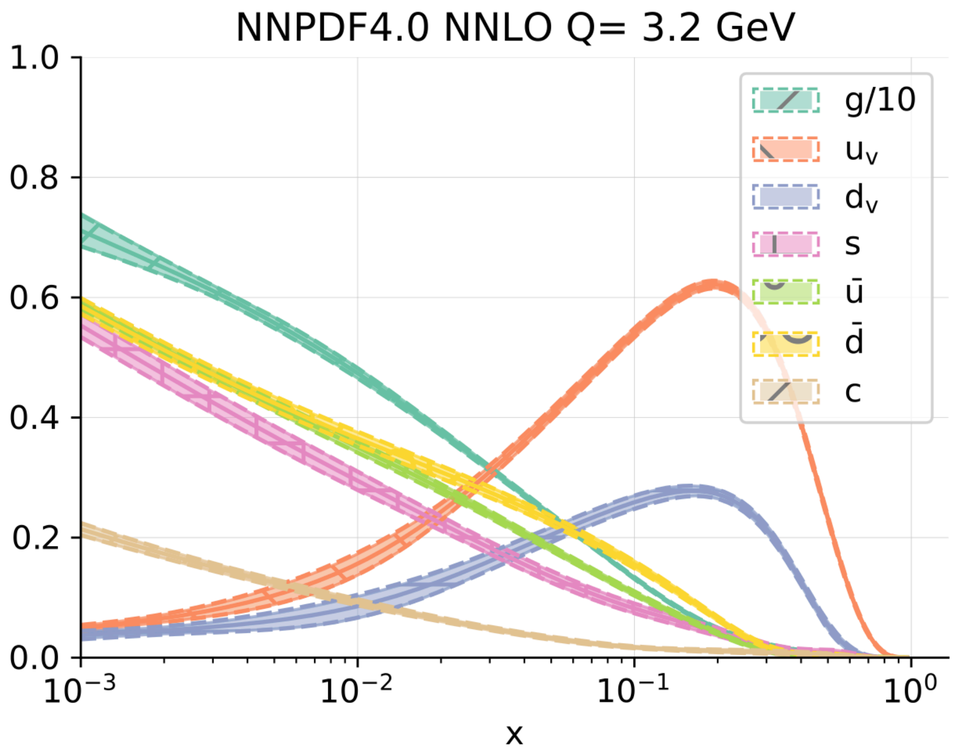
Hadron collider observable: $\sigma = f_1 \otimes f_2 \otimes \hat{\sigma}$

PDFs in a nutshell:

- describe proton in terms of partonic content
- $f(x, Q)$
- Q dependance: DGLAP equation ✓
- x dependance: non-perturbative QCD ✗

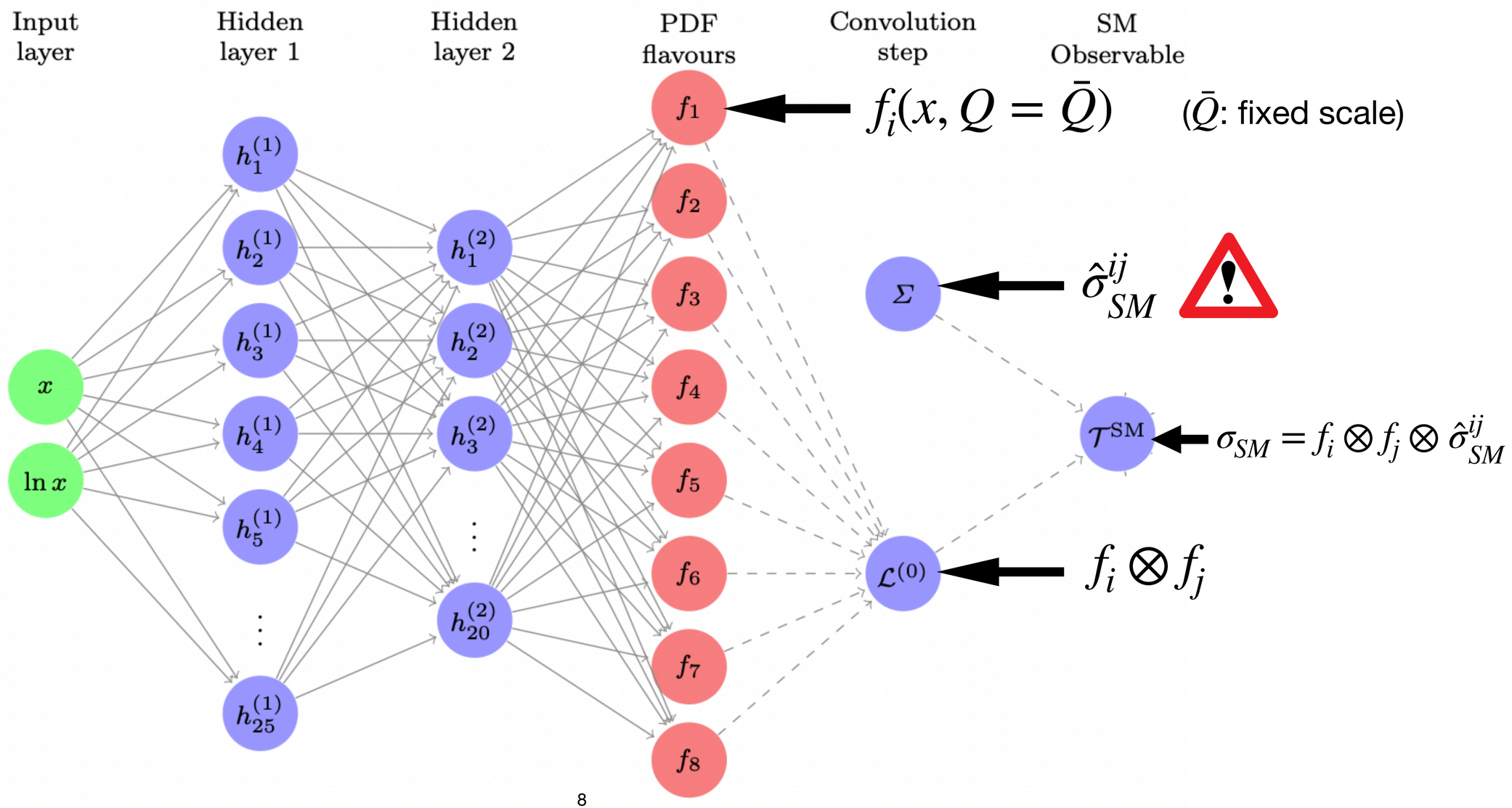
➡ **Fitted from data**

Using NNPDF methodology



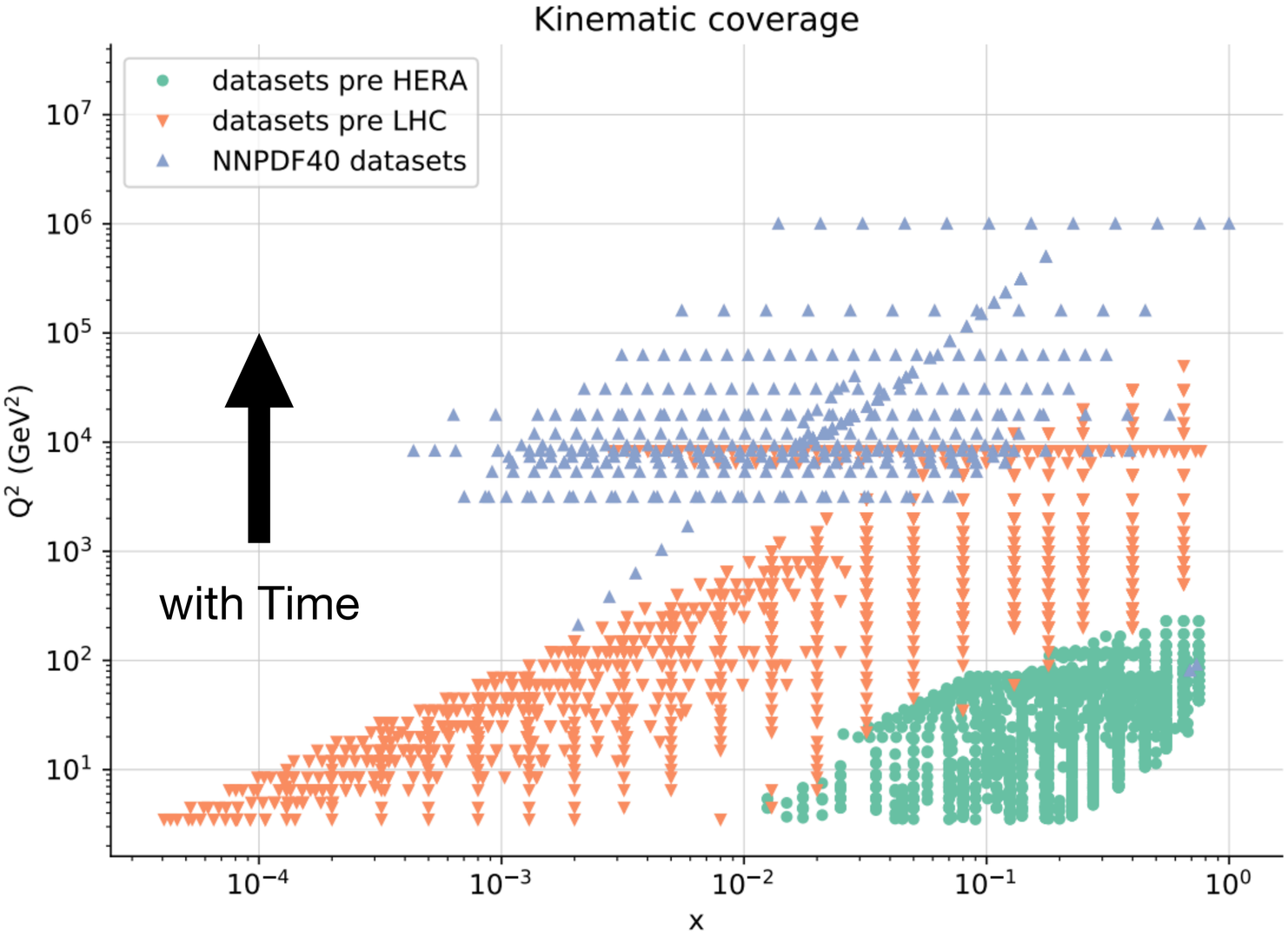
[Ball et al., NNPDF4.0, 2109.02653]

NNPDF fitting methodology



Bjorken-x
(Momentum fraction)

NNPDF dataset



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

Evolution of the dataset through time:

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

Risk of absorbing BSM deviations in PDFs

Risk assessment

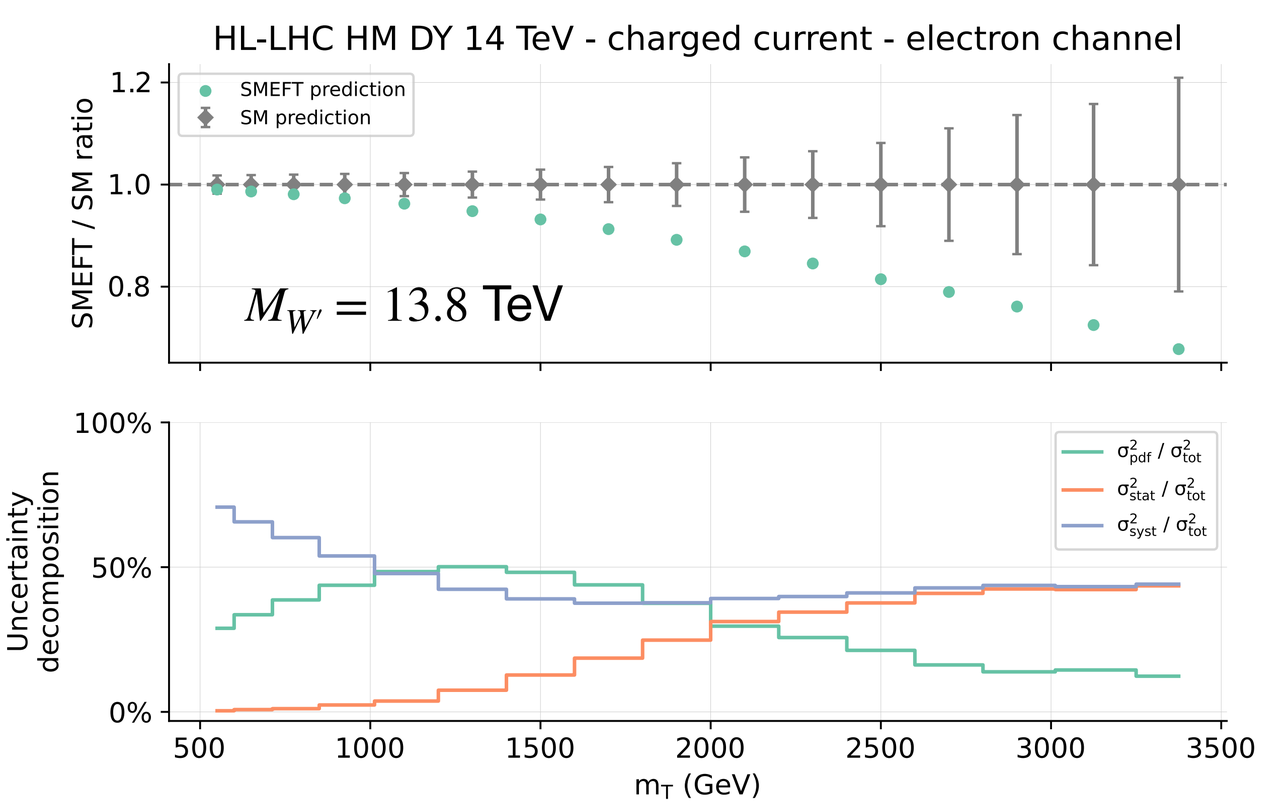
Absorbing signs of new physics in the PDFs

[PBSP, 2307.10370, JHEP]

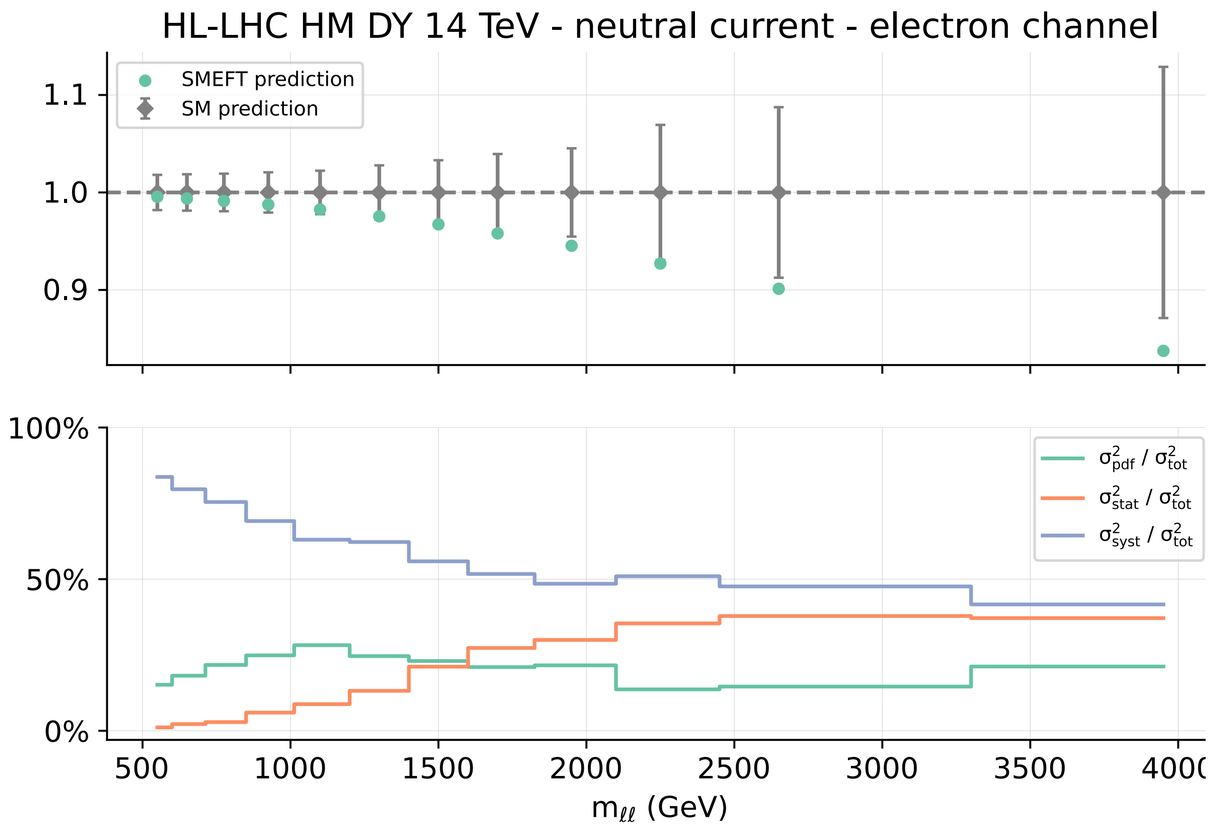
Injecting new physics in HL-LHC projections (DY)

Goal: recover it with a SMEFT fit

Charged current HMDY



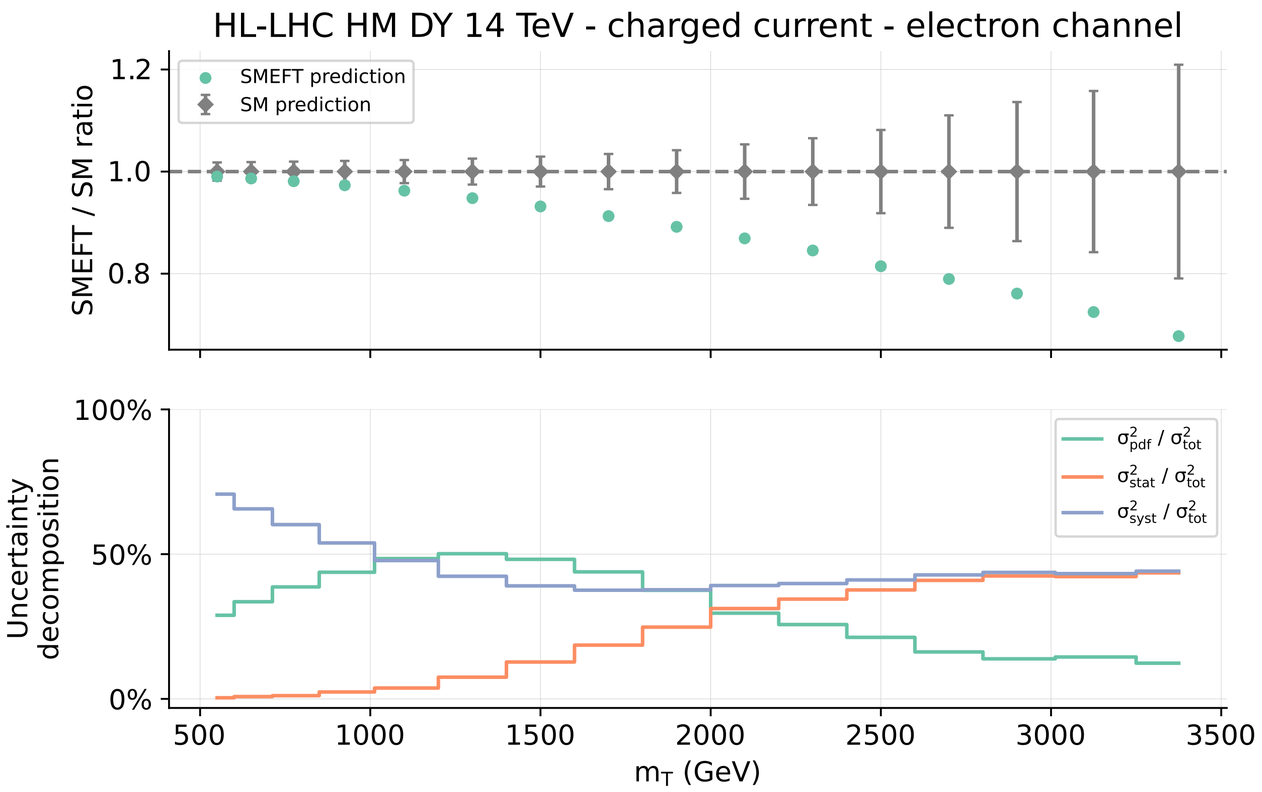
Neutral current HMDY

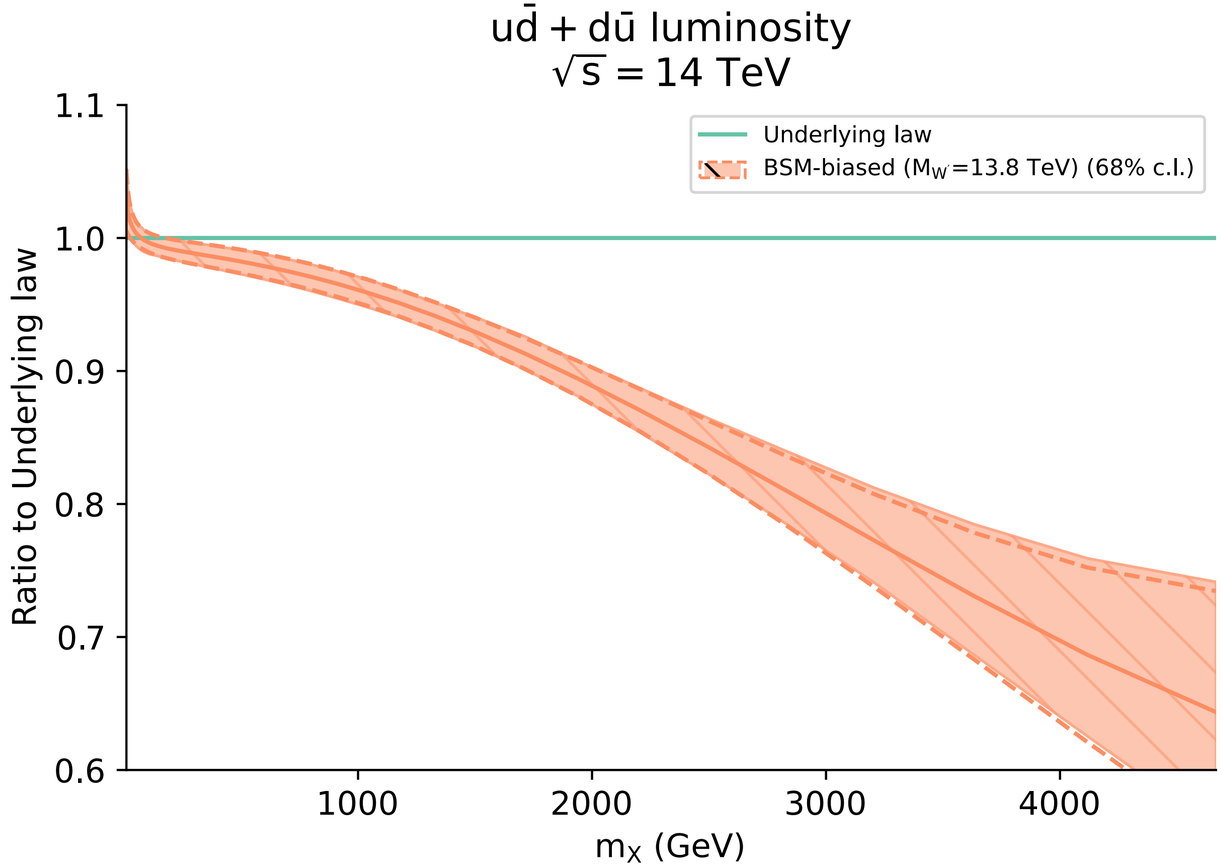


“BSM-biased” PDF in HMDY

PDFs are mimicking the SMEFT corrections

Charged current HMDY



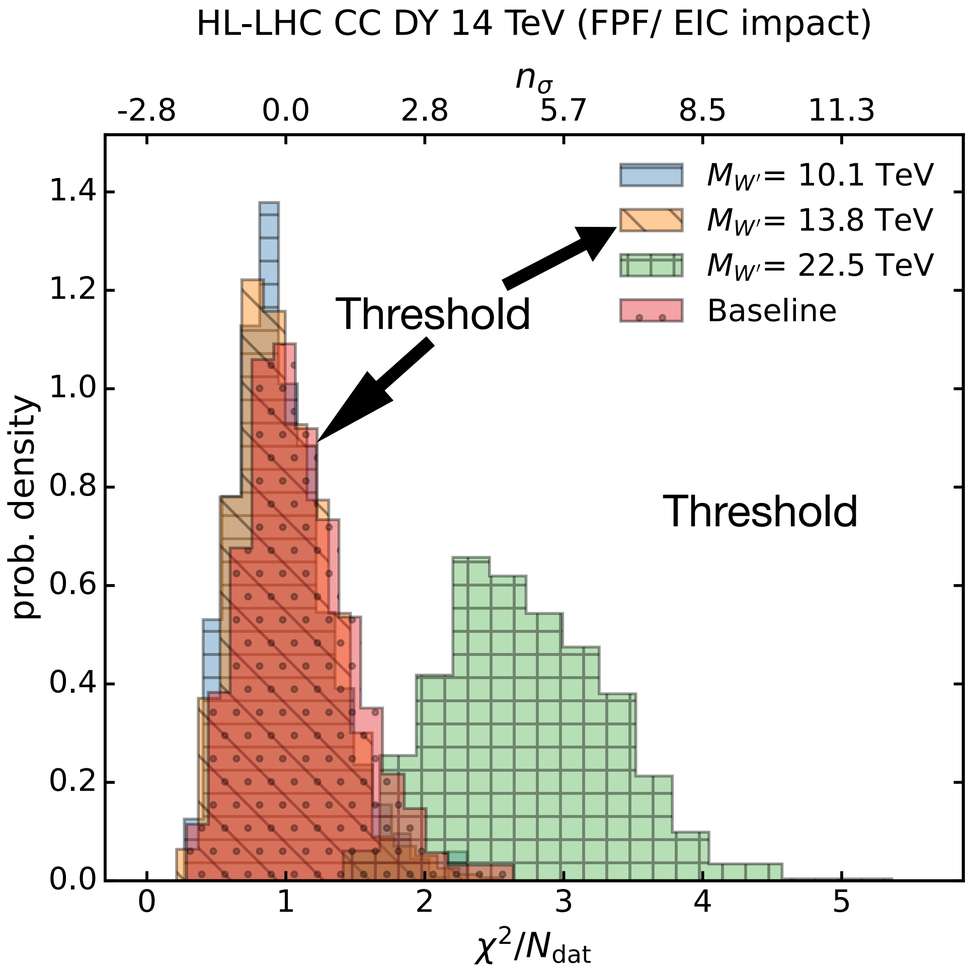


$$\mathscr{L}_{true} \otimes \hat{\sigma}_{BSM} \approx \mathscr{L}_{BSM-biased} \otimes \hat{\sigma}_{SM}$$

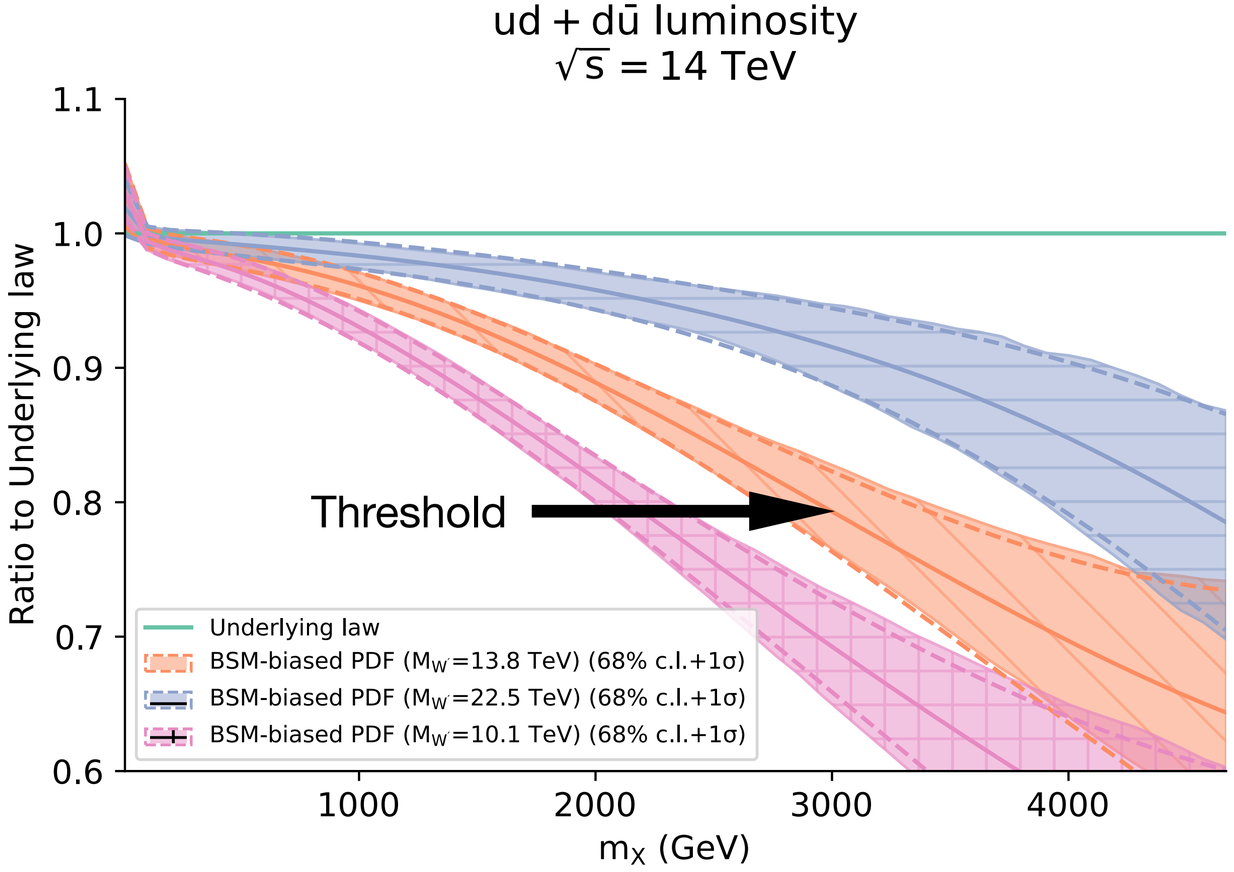
Maximal new physics absorption

Comparison between BSM and baseline fit qualities

Fit quality



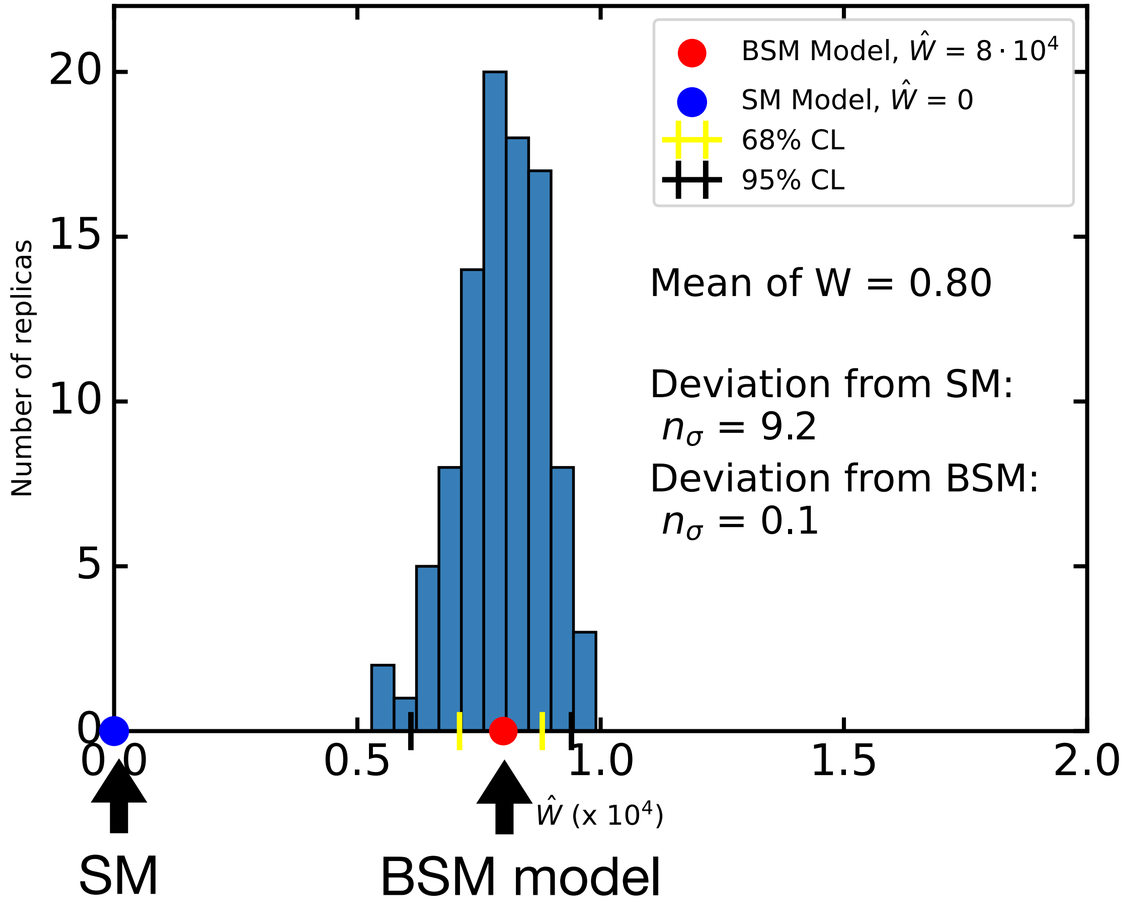
Impact on PDFs



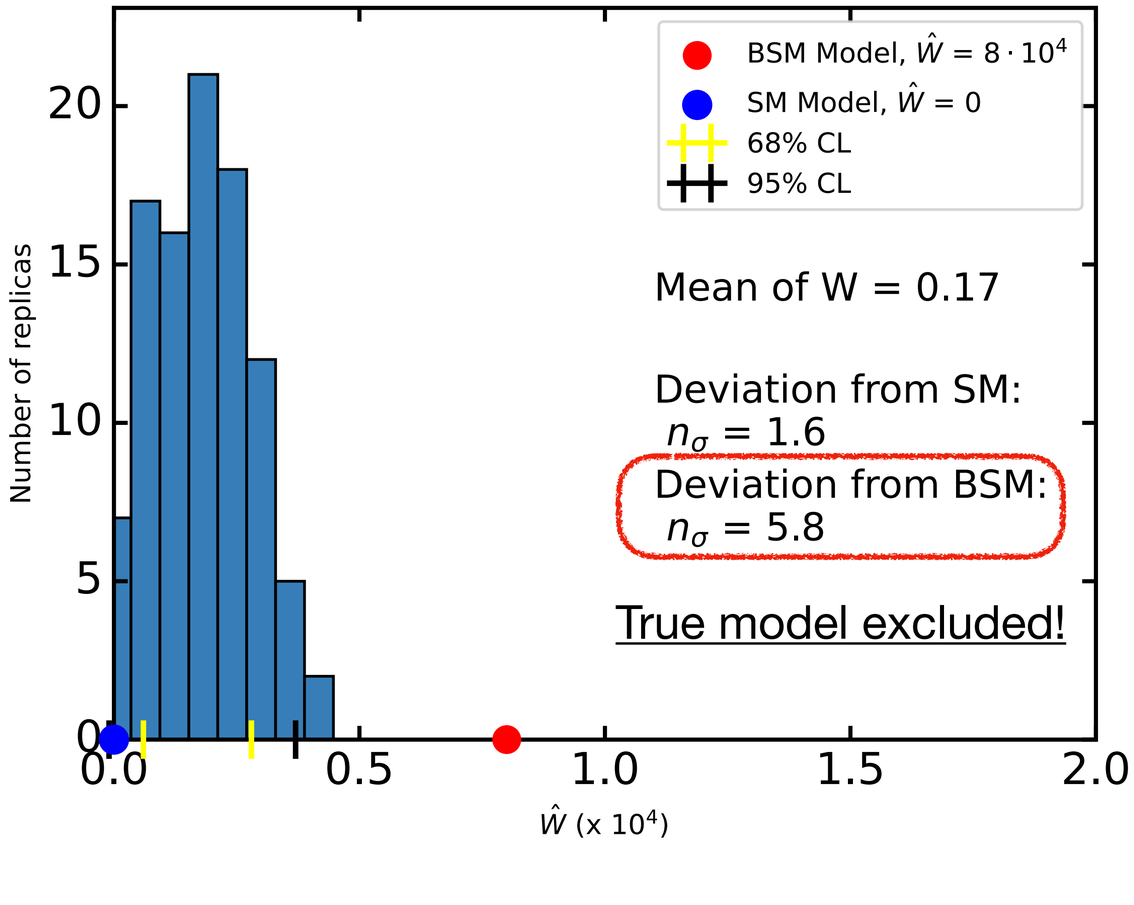
Missing new physics

Impact of the NP absorption in PDFs on SMEFT fits

SMEFT fit with true PDF



SMEFT fit with BSM-biased PDF



First solution

Bridging dataset “blindspot”

[Hammou et Ubiali, 2410.00963, accepted in PRD]

Future low energy observables

Presentation of the future DIS programmes

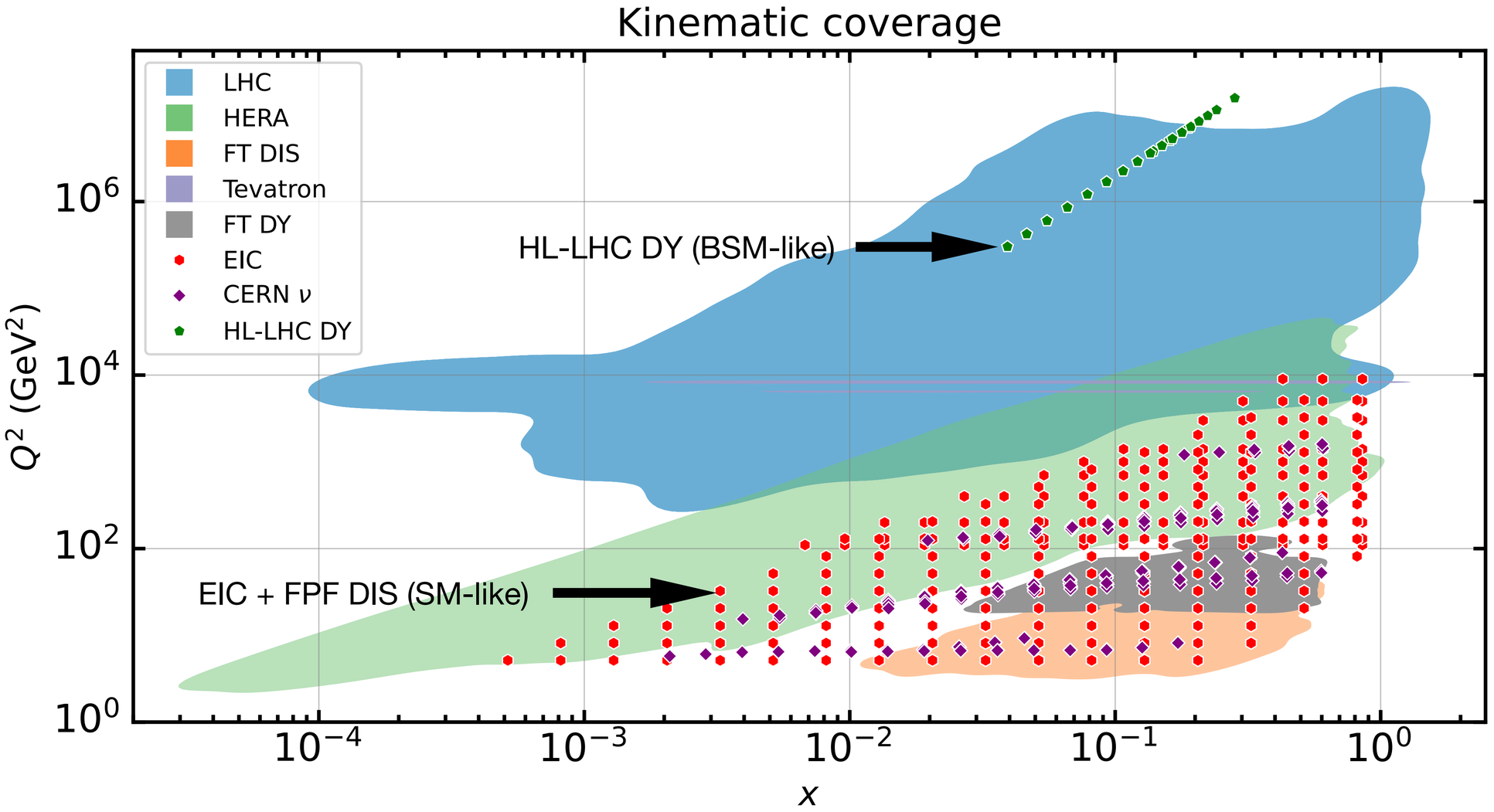
Electron Ion Collider

- e^+/e^- projectiles
- proton, deuteron and heavy ions targets
- Hosted in Brookhaven
- Planned for 2030s
- Probes large-x, low-energy

Forward Physics Facility

- “Neutrino Ion collider” at the LHC
- $\nu/\bar{\nu}$ projectiles from proton beam
- proton, neutron and other nuclear targets
- FASERν and SND@LHC already running
- Proposed expansion for HL-LHC run (FASERν2 , AdvSND, FLArE)
- Probes large-x, low-energy
- Constrain large-x antiquarks

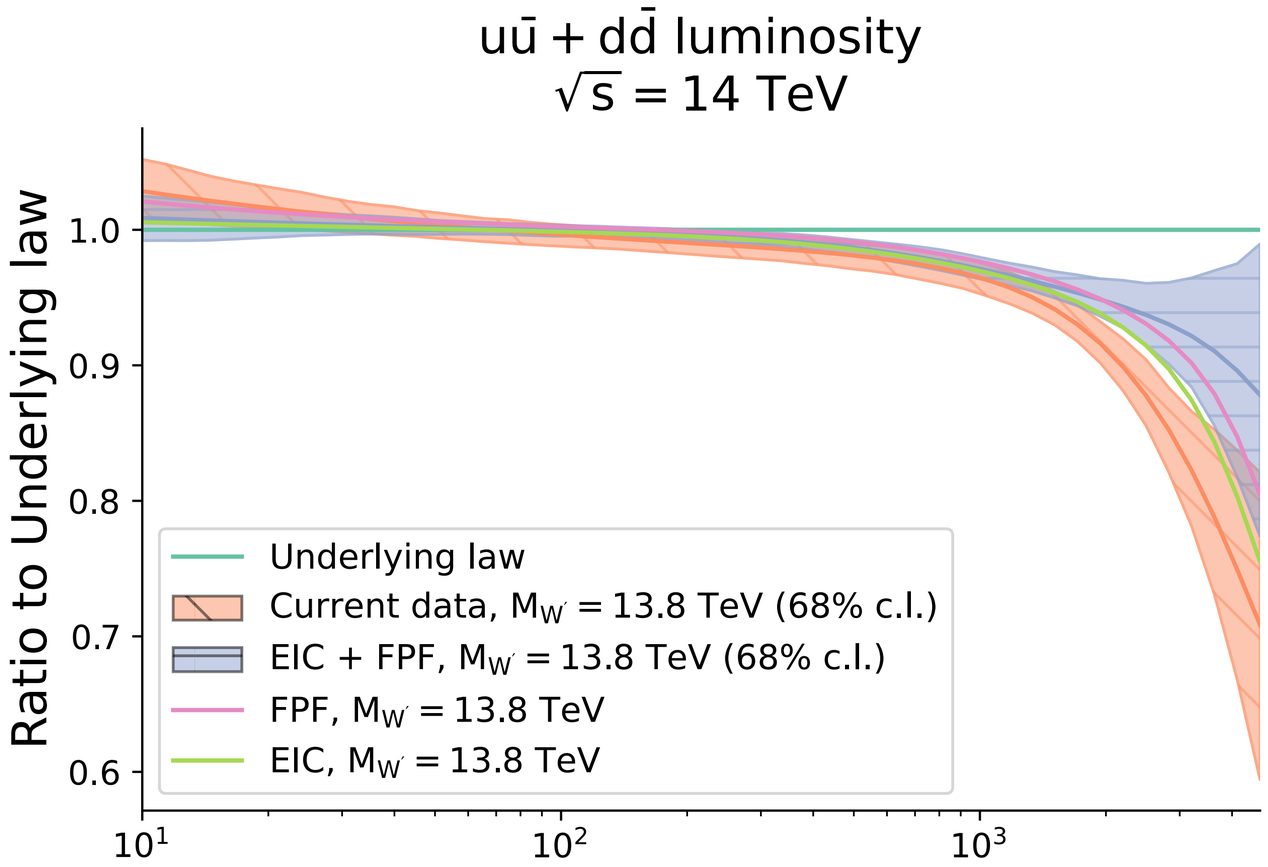
EIC, FPF and HL-LHC kinematics



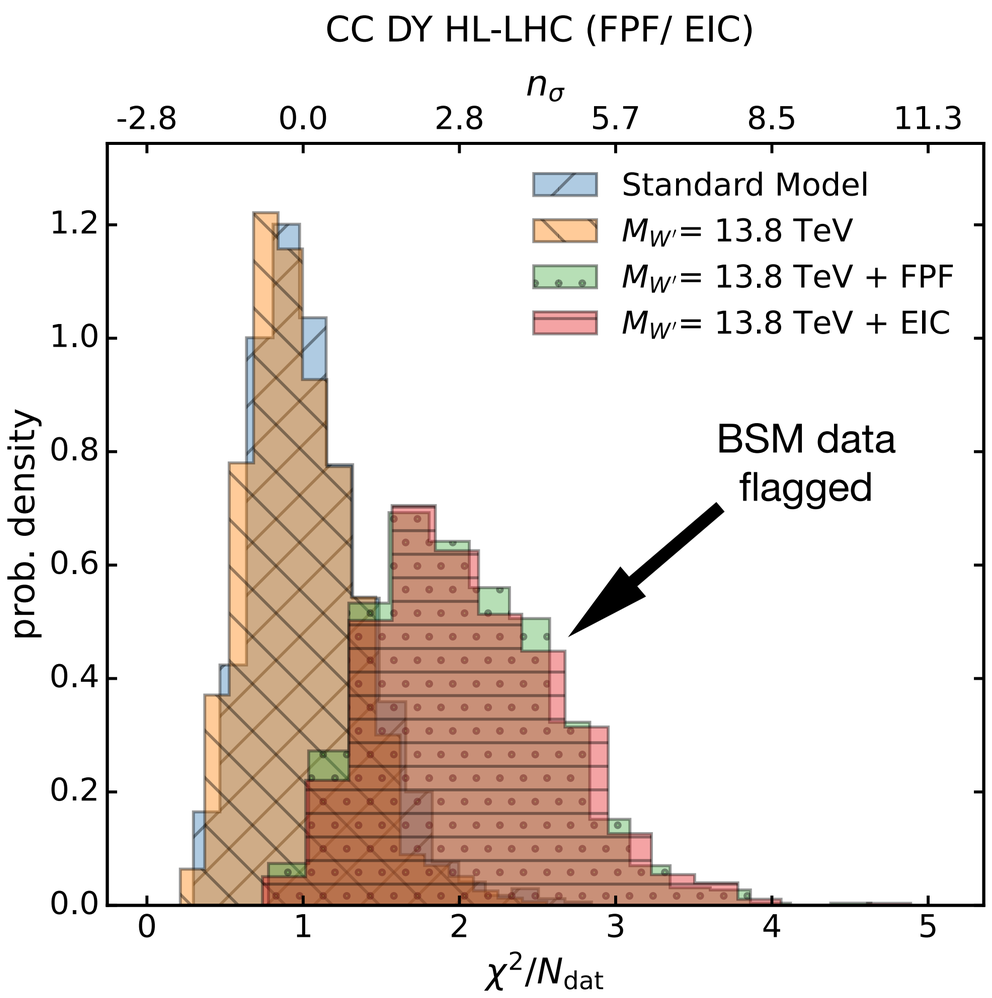
Impact on the PDF contamination

Flagging the BSM data

Pull on PDFs



Impact on fit quality



Shift of the contamination threshold

From the fit quality

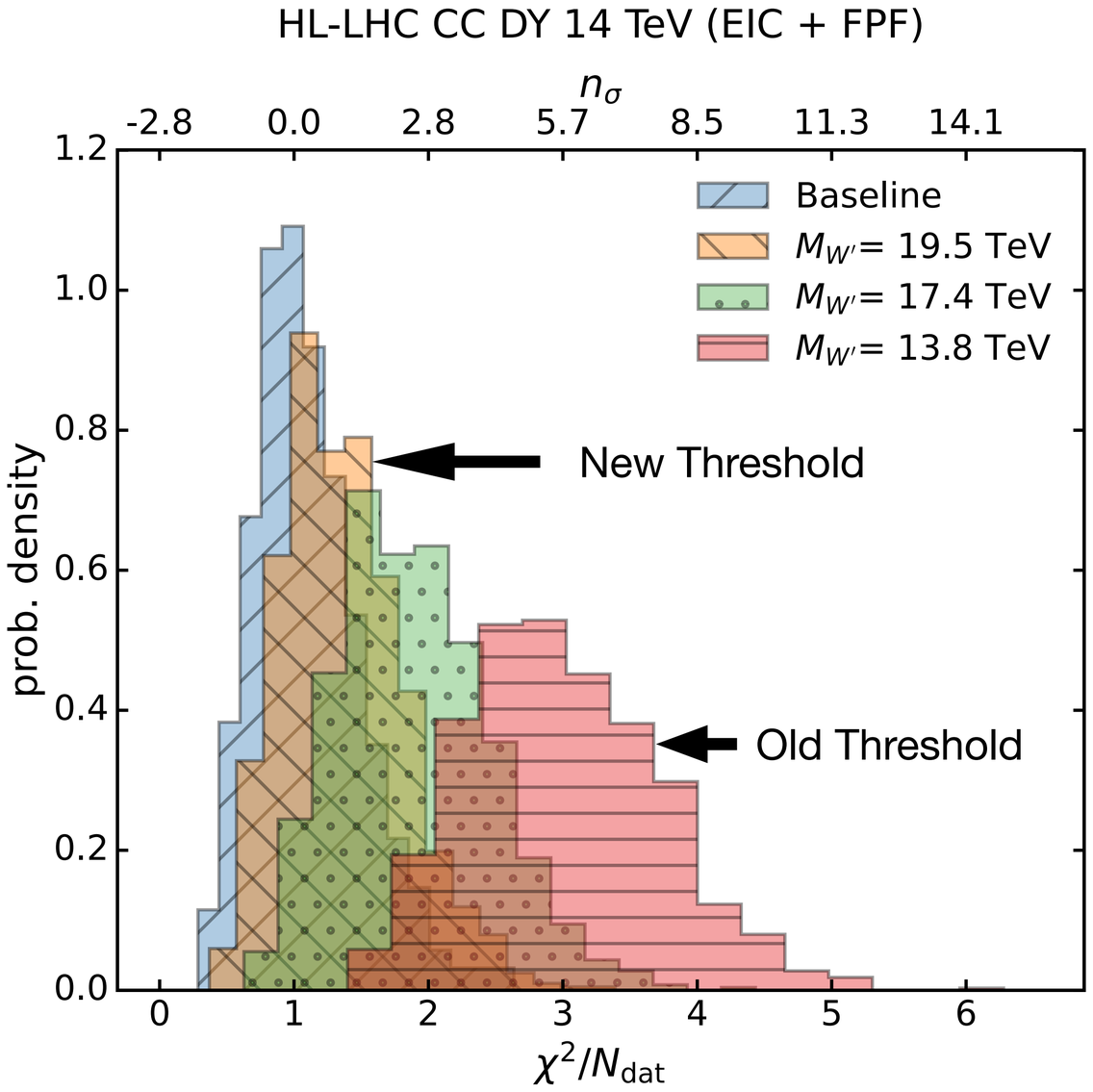
Not a complete solution:

Smaller deviations can still be absorbed

➡ risk at higher BSM mass

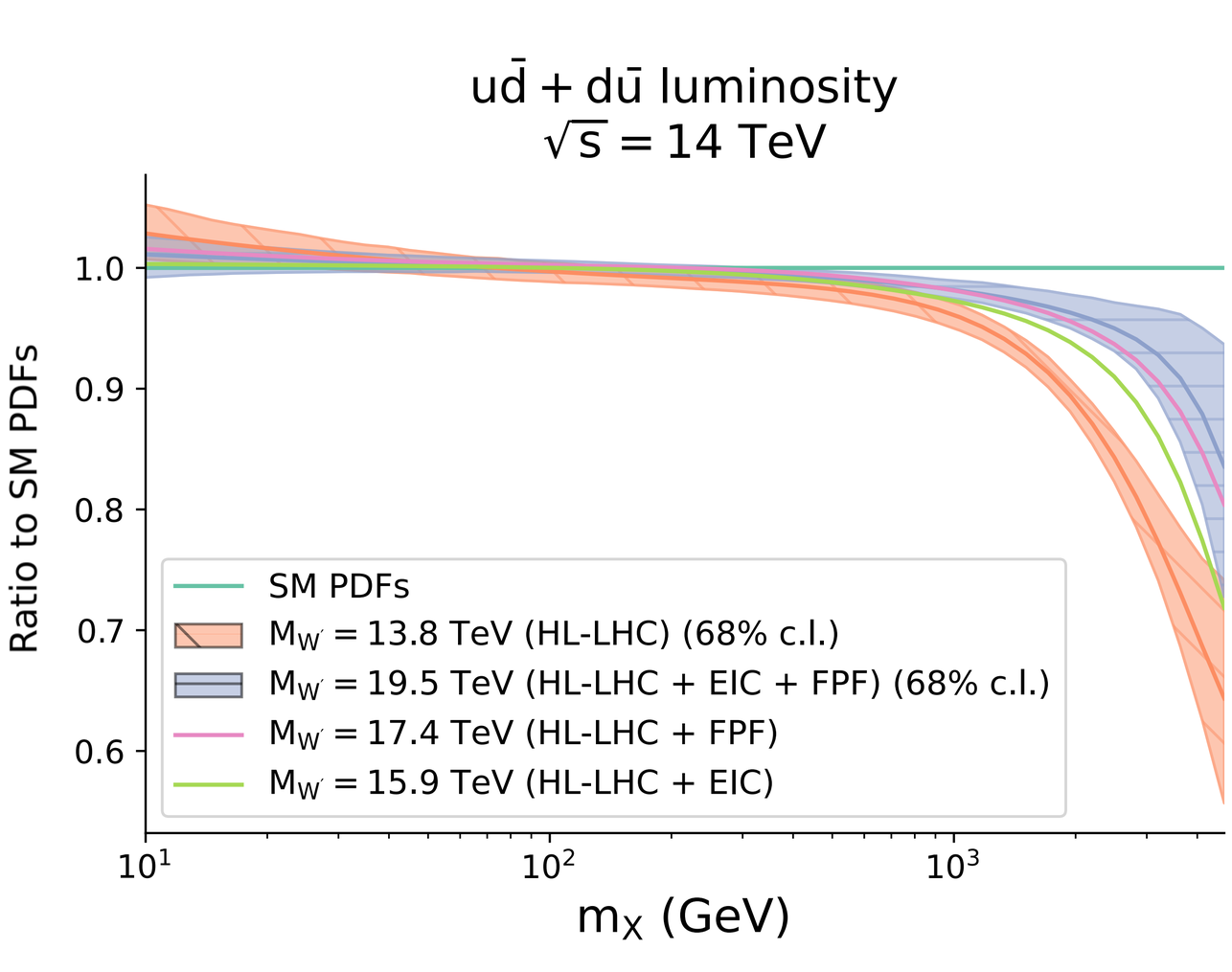
Reduction of the “blindspot”:

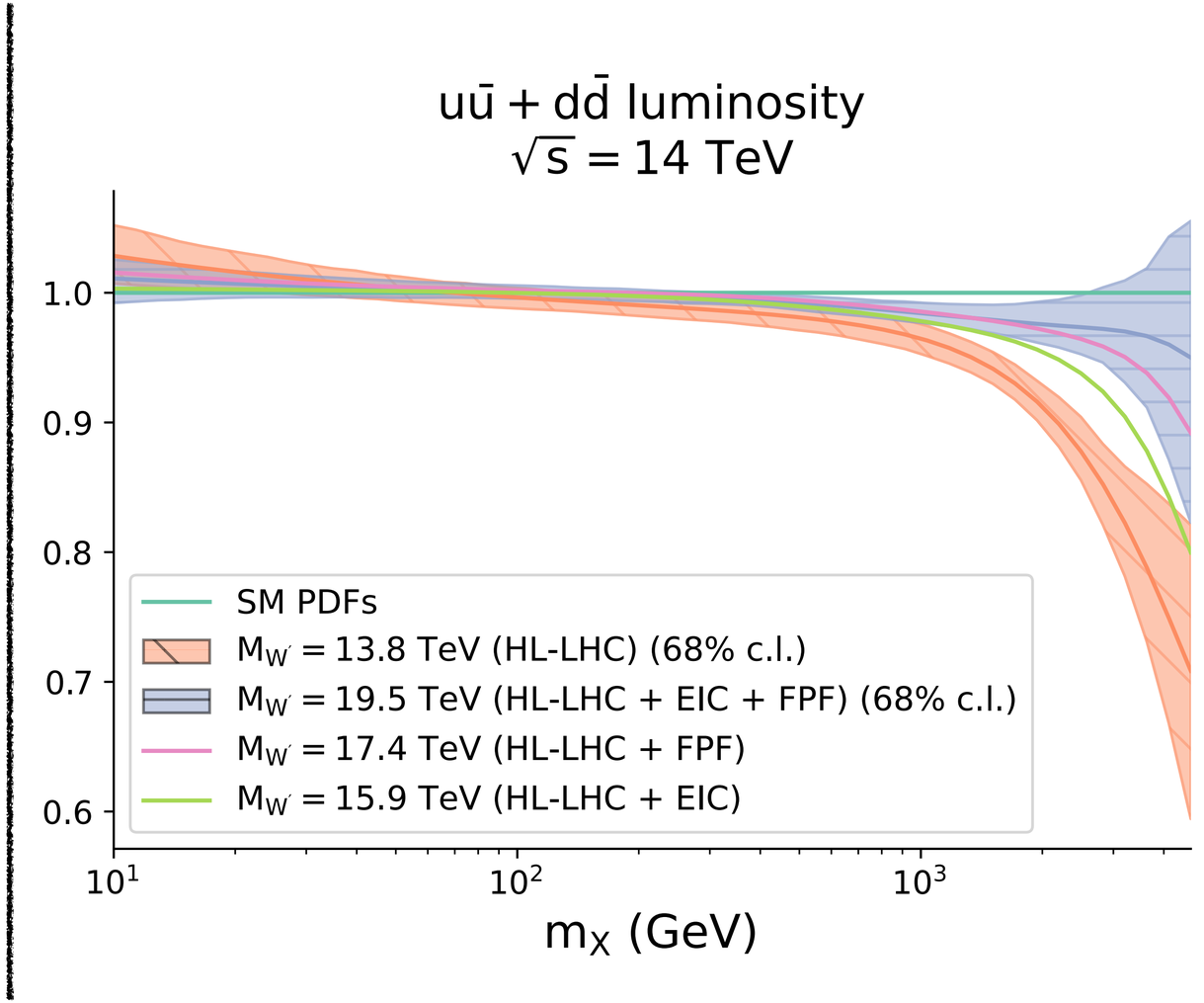
$$M_{W'} : 13.8 \rightarrow 19.5 \text{ TeV}$$



Shift of the contamination threshold

Impact on PDF luminosities





Second solution

Simultaneous fits of PDFs and SMEFT

[PBSP, 2402.03308, Eur.Phys.J.C]

[PBSP, forthcoming]

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:

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

➡ $\bar{c}$

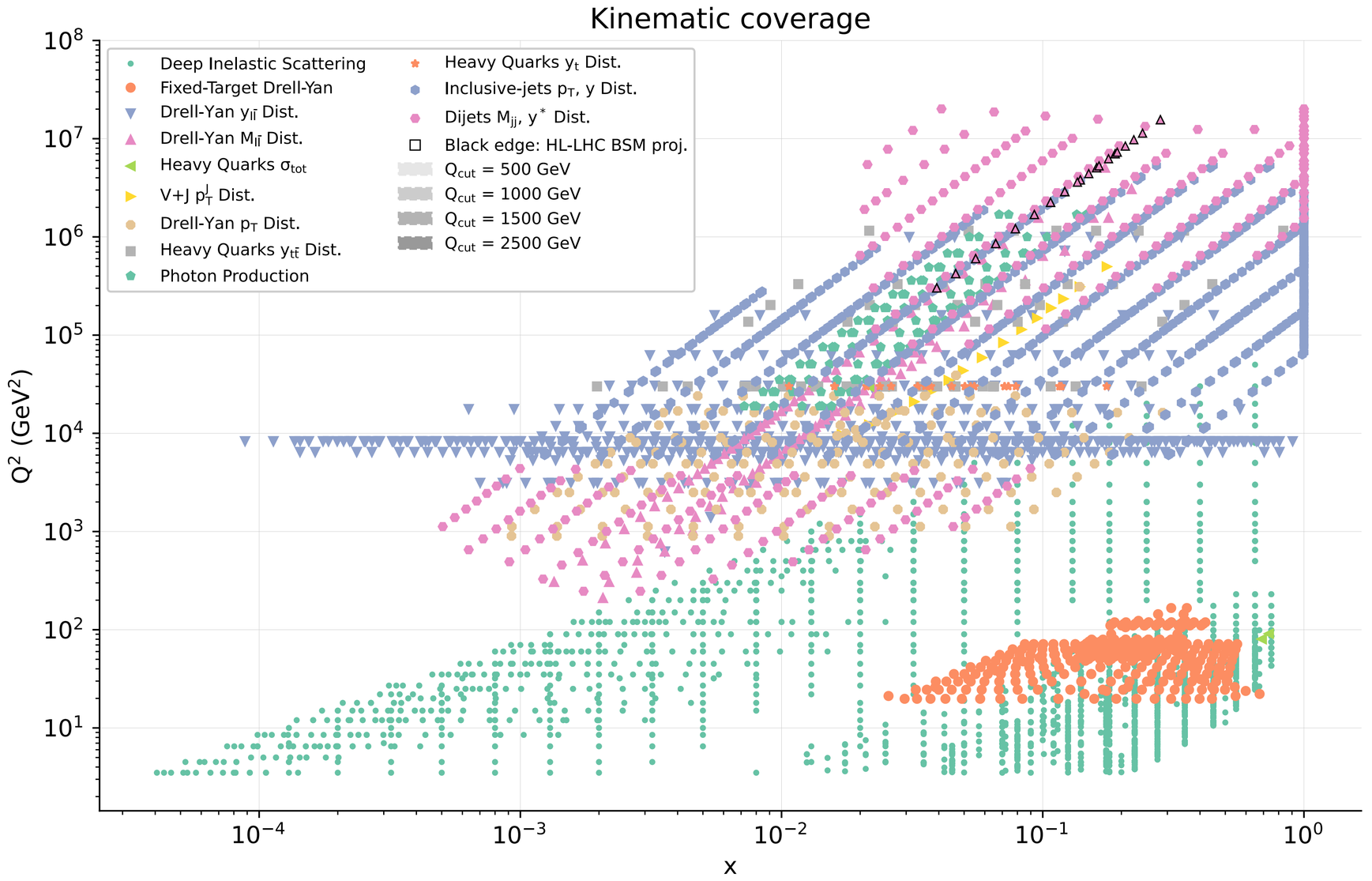
Simultaneous fits

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

➡ $\{\bar{\theta}, \bar{c}\}$

Removes assumption-based bias

Safe separate fits: “conservative” PDFs

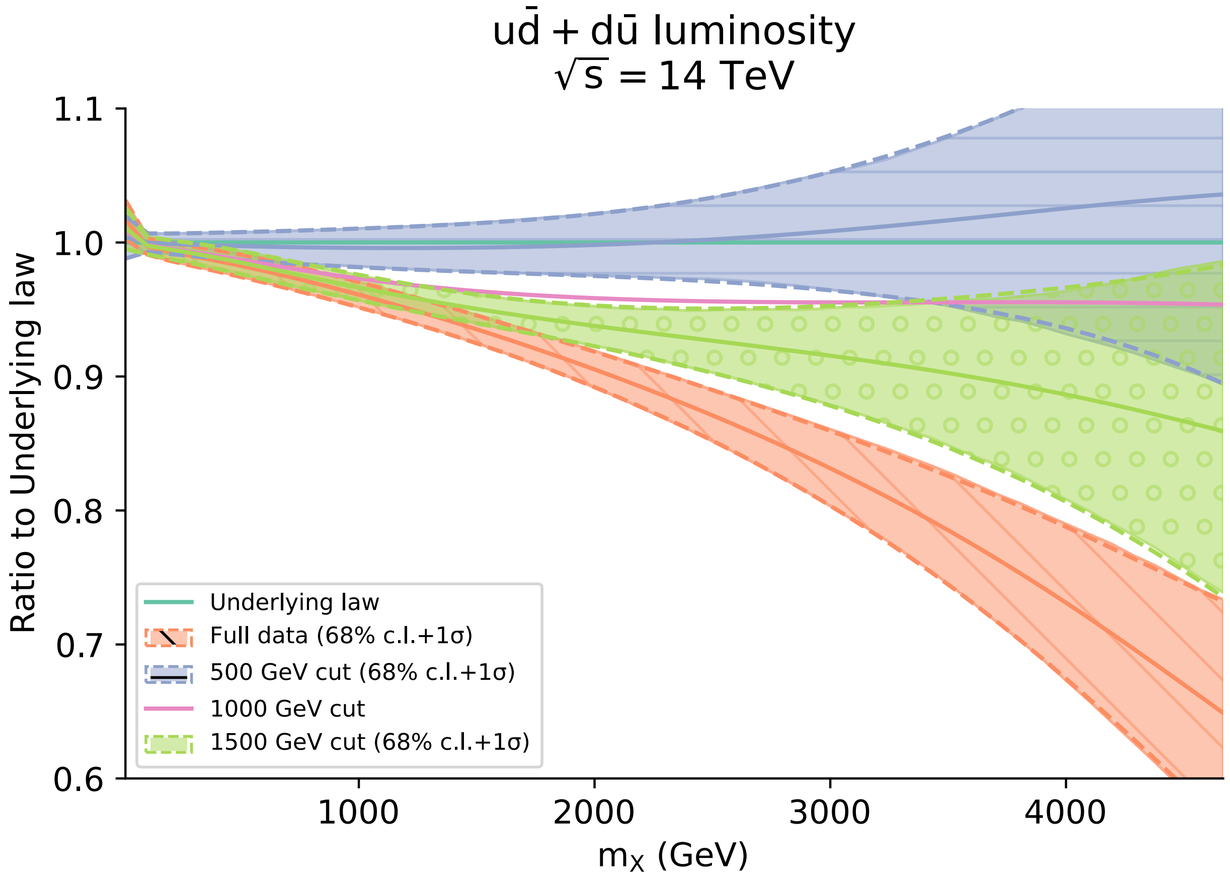


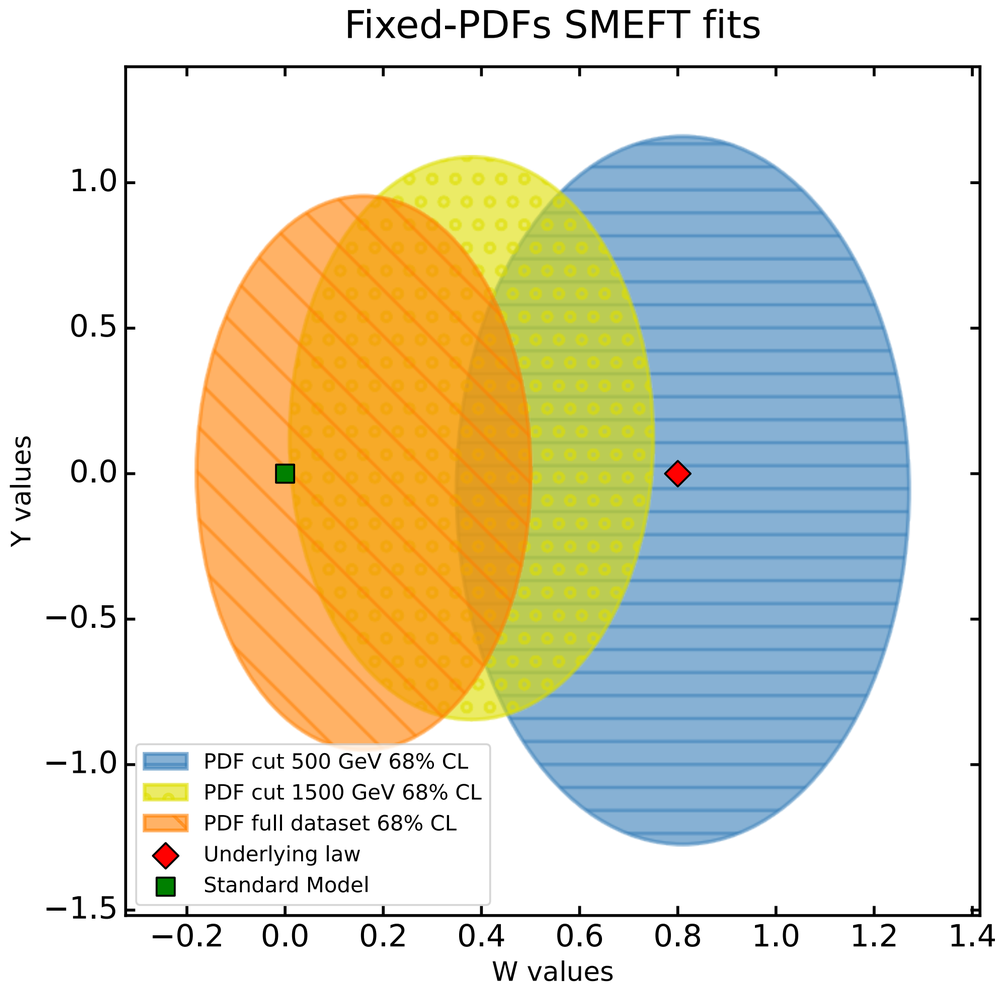
Agnostic energy cuts

“Conservative” separate fits

[PBSP, forthcoming]

Impact on PDF and SMEFT bounds for DY





Simultaneous fit of PDF and new physics

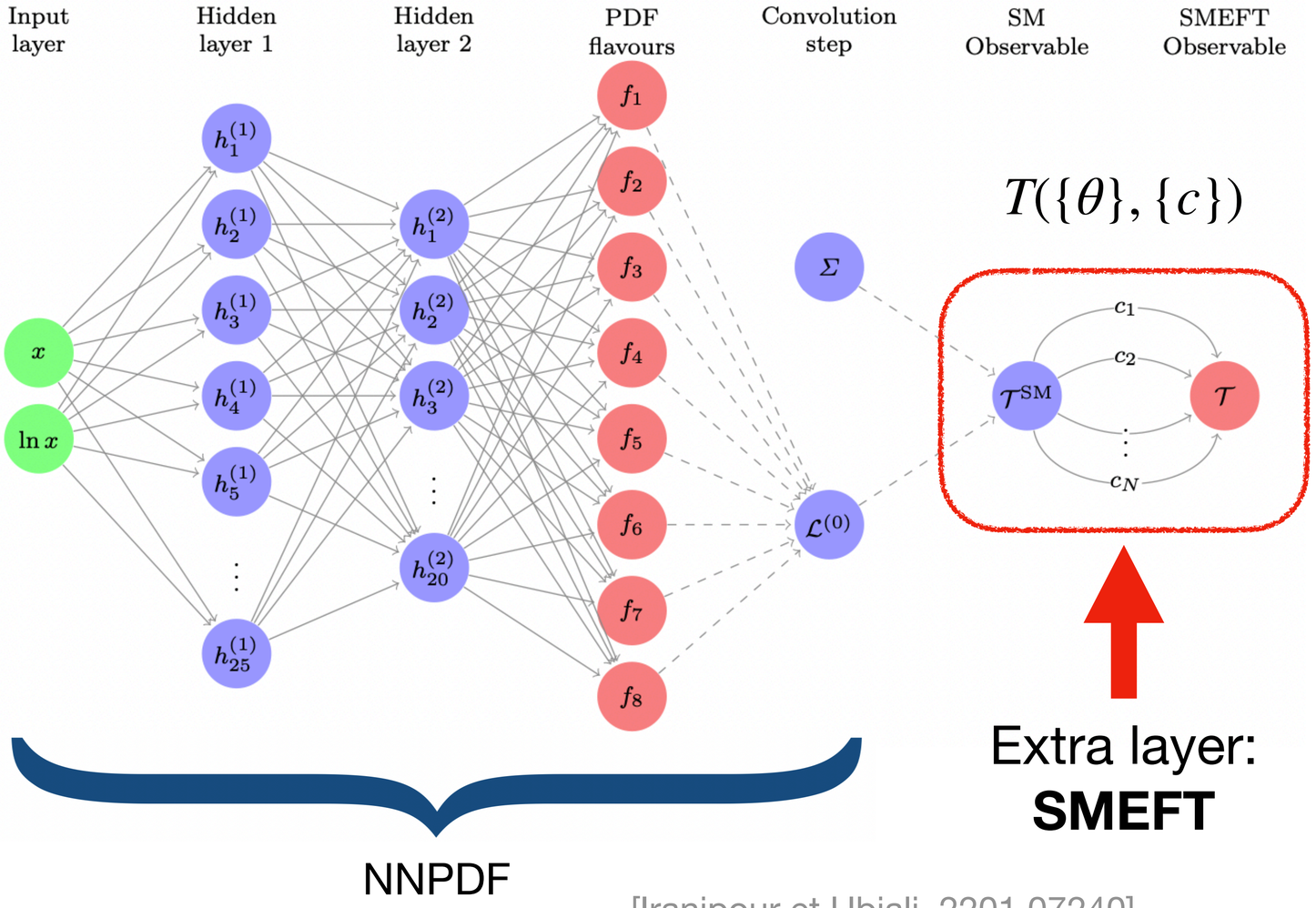
Presentation of the tool: SIMUnet

SIMUnet:

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

  [PBSP, 2402.03308]

- Fits PDFs and WC simultaneously
- Performs contaminated PDF fits

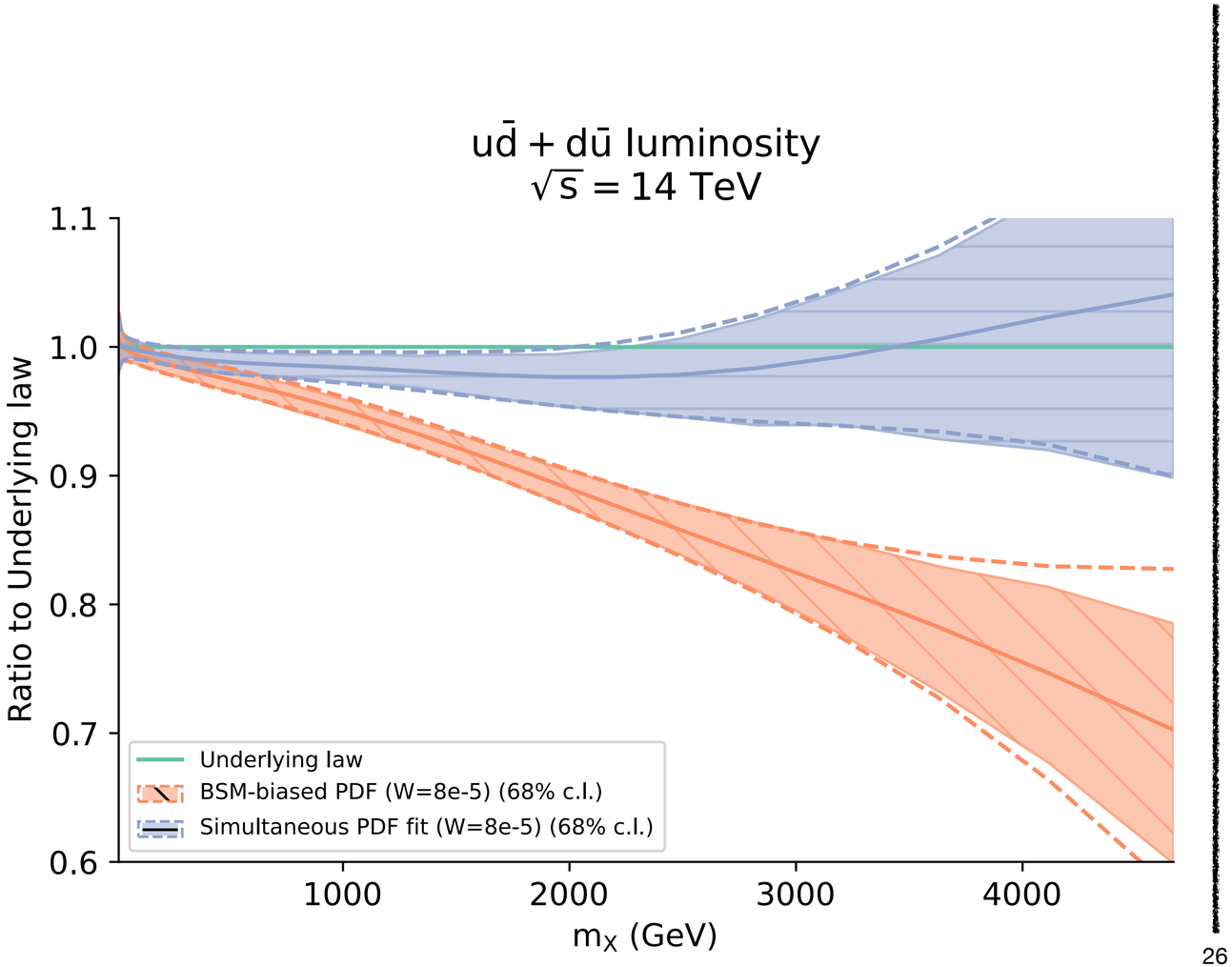


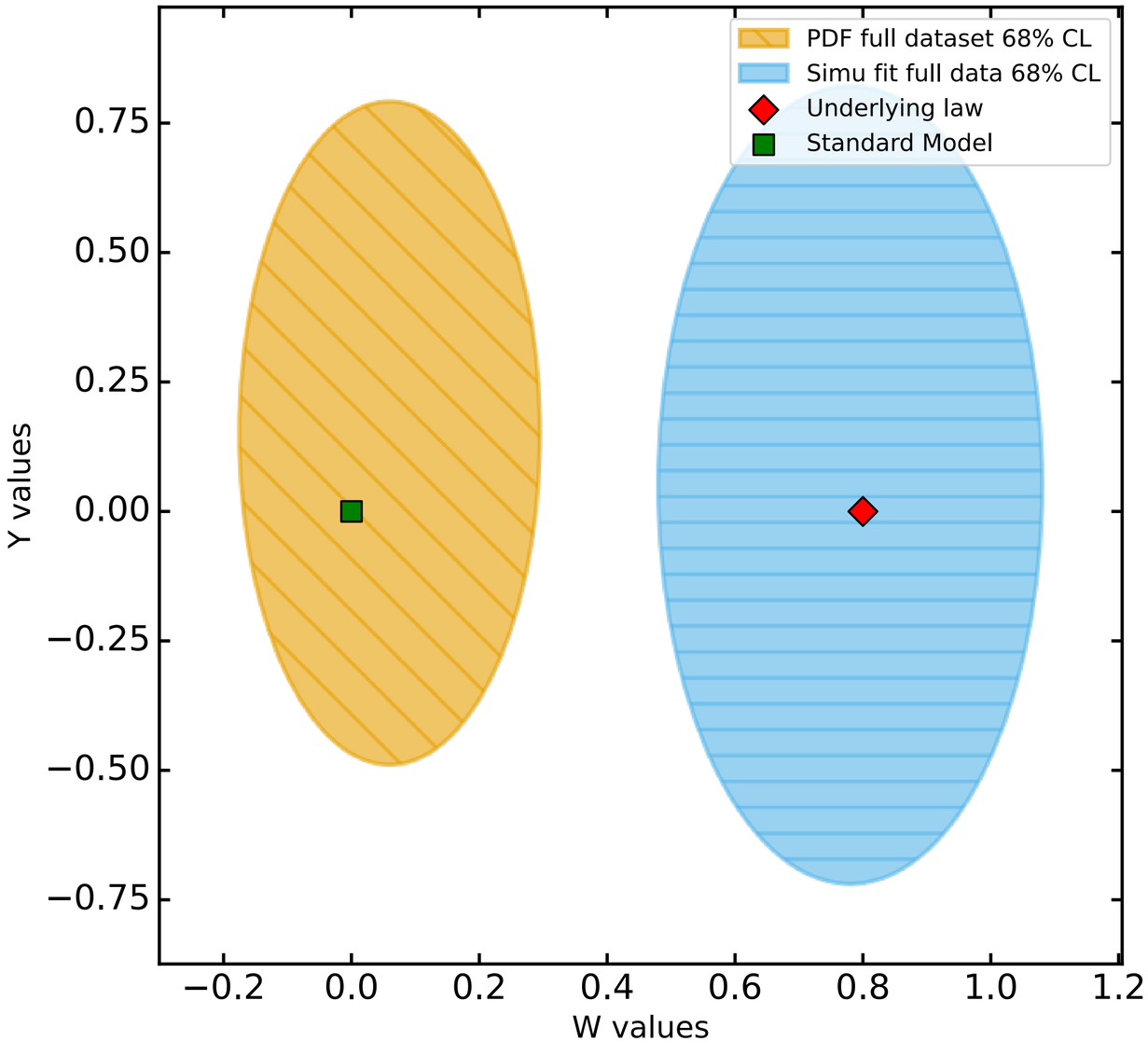
[Iranipour et Ubiali, 2201.07240]

Simultaneous fit of PDF and SMEFT

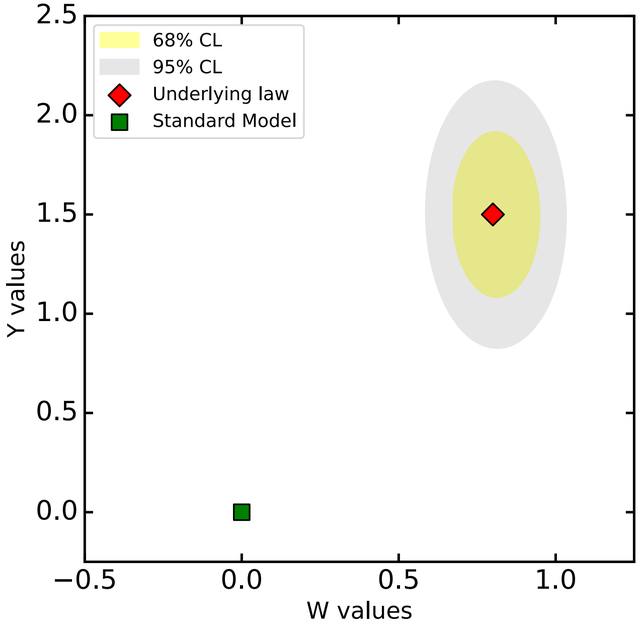
[PBSP, forthcoming]

Marginal distributions for DY





PDFs for new physics searches in HMDY

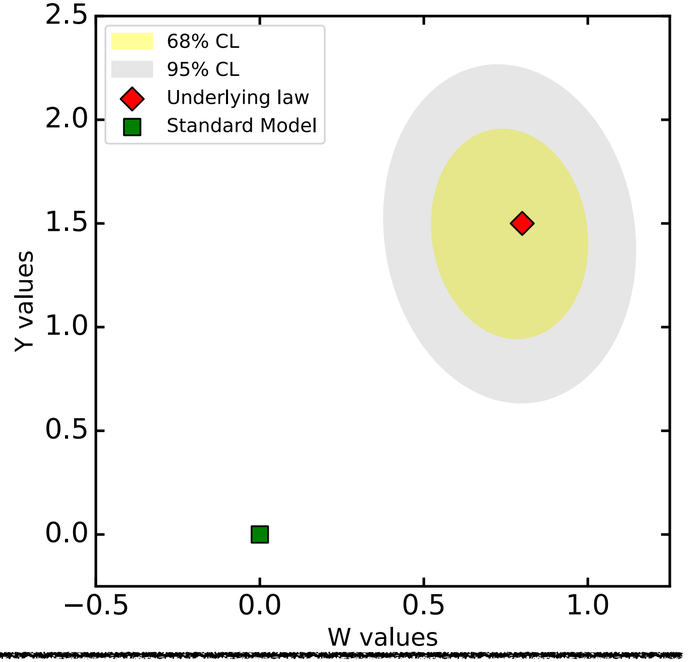


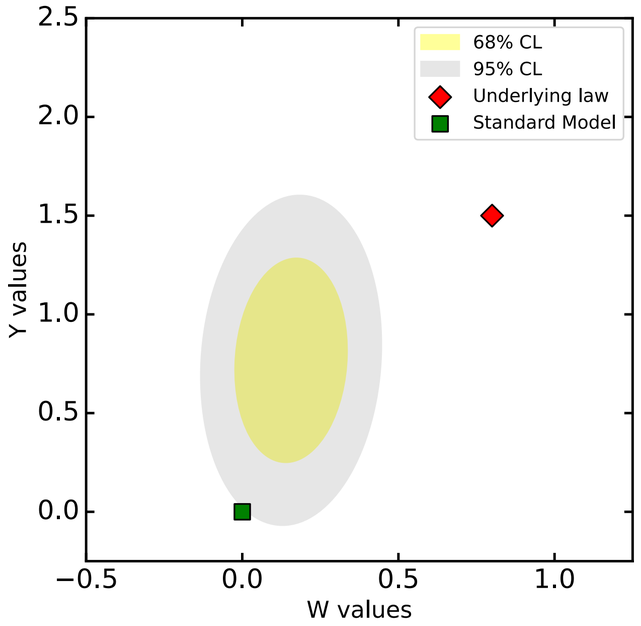
SMEFT only fit
(True PDF)

✗ Impossible

Simultaneous fit
(no PDF assumption)

✓ Doable



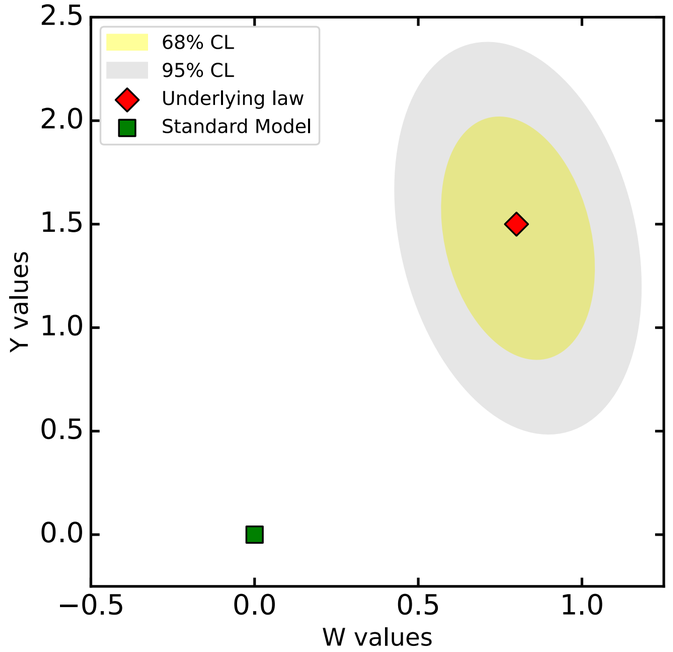


SMEFT only fit
(BSM-biased PDF)

✗ Wrong

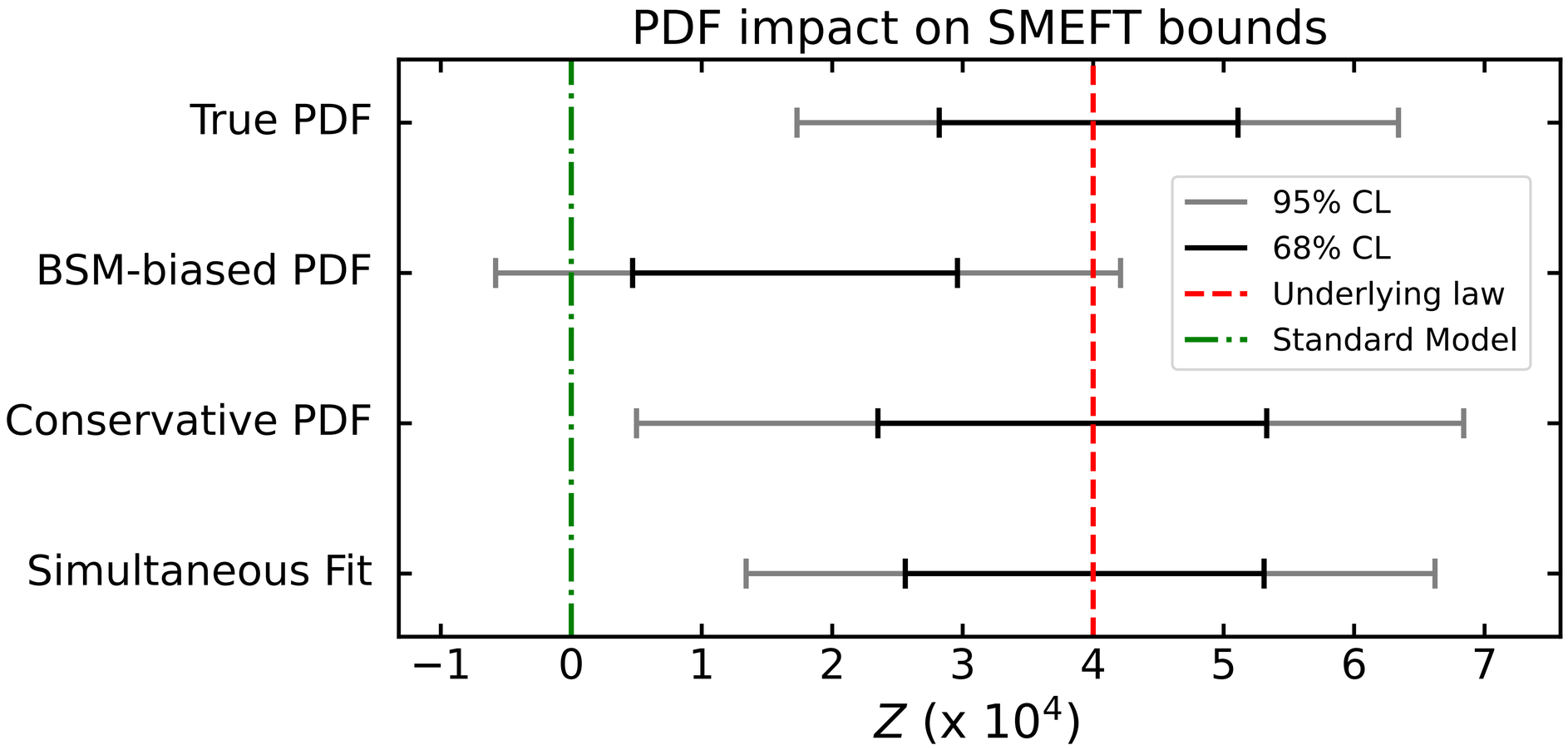
SMEFT only fit
(Conservative PDF)

✓ Doable



PDFs for new physics searches in $t\bar{t}$

Same exercise with heavy gluon in ttbar at HL-LHC



Comparing conservative and simultaneous fits

Conservative separate fits

Pros:

- Easier
- Less parameters per fit

Cons:

- Difficult to figure out optimal cutoff (manageable)
- **Cannot use precise high-energy observables to constrain PDFs**

Simultaneous fits

Pros:

- Entire dataset constrains PDF and SMEFT
- High-energy observables constrain PDF

Cons:

- More parameters -> more uncertainty (manageable)
- **Risk to absorb SM error as SMEFT signal**

Summary and outlook

- Signs of new physics fitted away in PDF parametrisation
  - Missed new physics
  - Exclude true underlying law
- Bridge the dataset “blind spot”:
  - Add precise large-x low-energy datasets into fits: FPF + EIC
  - Pushes the PDF/BSM mixing threshold toward higher energies
- Simultaneous fits of PDFs and SMEFT:
  - Fitting simultaneously PDF and new physics: **SIMUnet** tool already available
  - Ongoing study on real jet data [Greljo, Hammou, Merlotti, Smolkovic, Ubiali, forthcoming]
  - Developing a bayesian framework [Costantini, Moore, Mantani, Schutze, Ubiali, forthcoming (PDF)] [Hammou, ter Hoeve, Shutze, in progress (SMEFT)]

You can contact me at:
ehammou@nikhef.nl

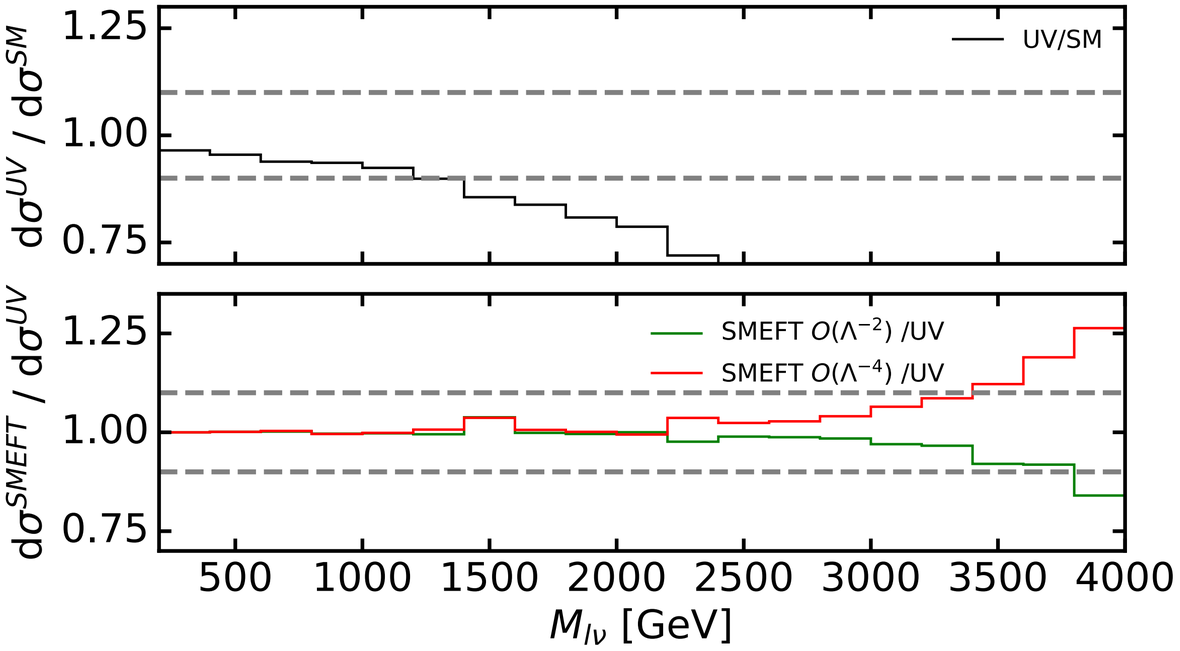
Thank you for your attention!

Extra slides

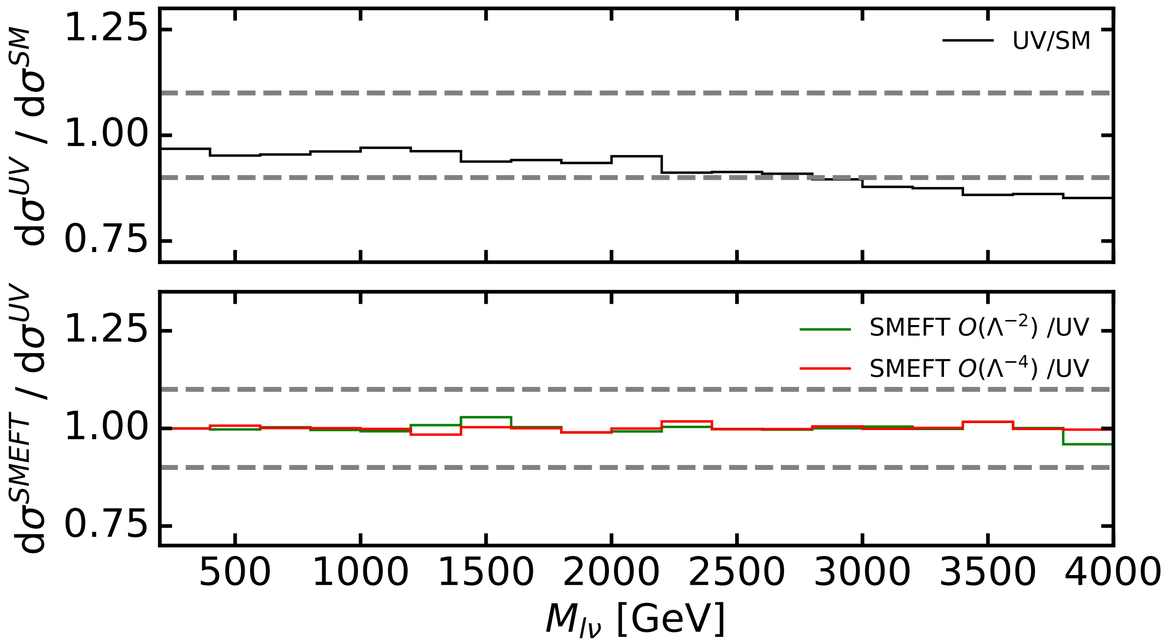
New physics scenarios: W'

Consideration of different masses

$M_{W'} = 10$ TeV



$M_{W'} = 22.5$ TeV



PDF fitting: selection criteria

Exclusion of incompatible datasets (NNPDF criteria)

Two criteria:

$$\chi^2 = (D - T_{SM})^T \cdot V_{cov}^{-1} \cdot (D - T_{SM})$$

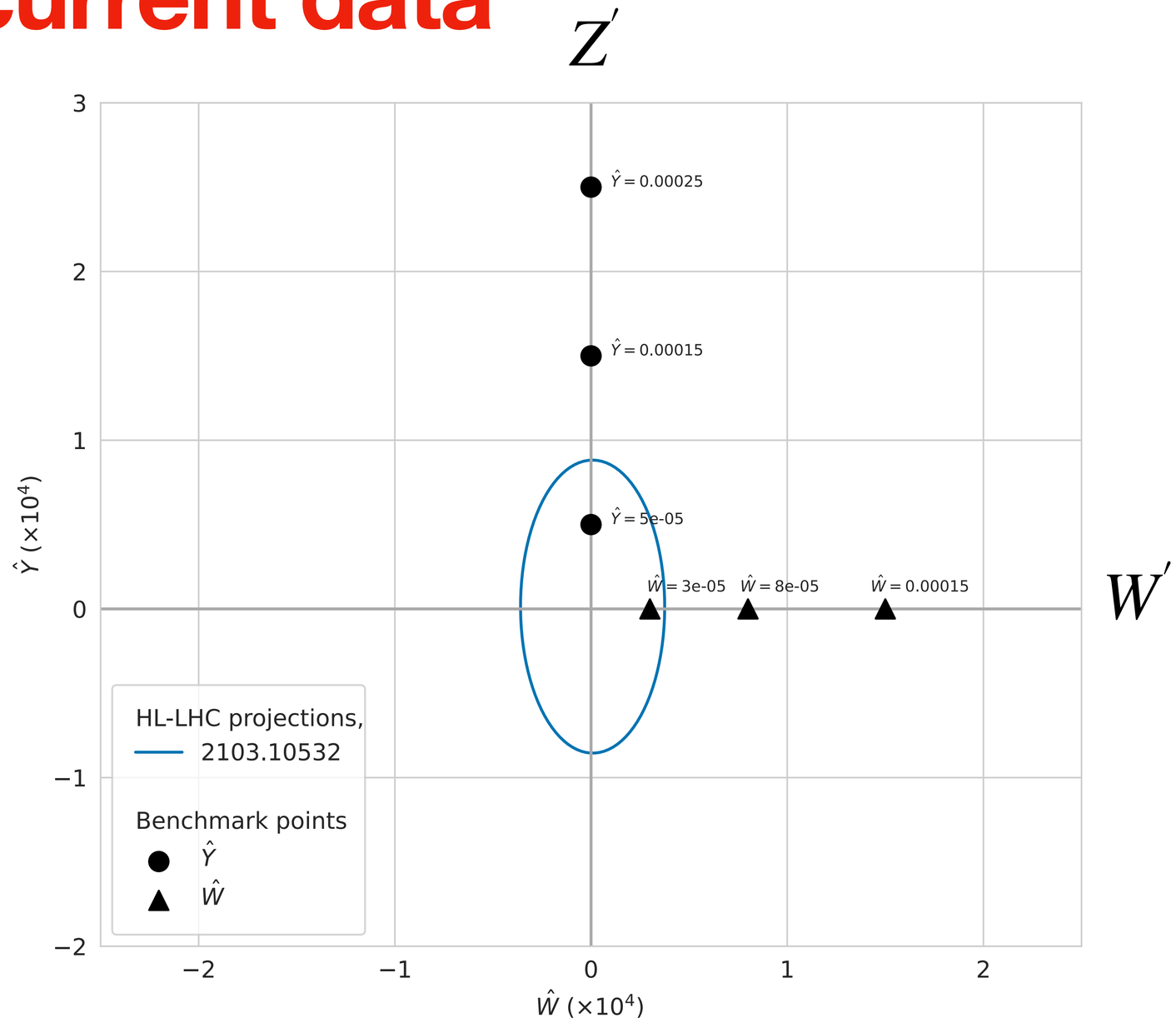
- χ^2-statistics:
  - $\frac{\chi^2}{n_{dat}} > 1.5$ → excluded

- n_σ standard deviation:
  - $n_\sigma > 2$ → excluded

$$n_\sigma = \frac{\chi^2 - 1}{\sigma_{\chi^2}}$$

Constraints from current data

- New physics scenarios compared to constraints at 95% CL



Impact of contamination: fake deviations

SM predictions with:

- Contaminated PDFs (red)
- True PDFs (black)

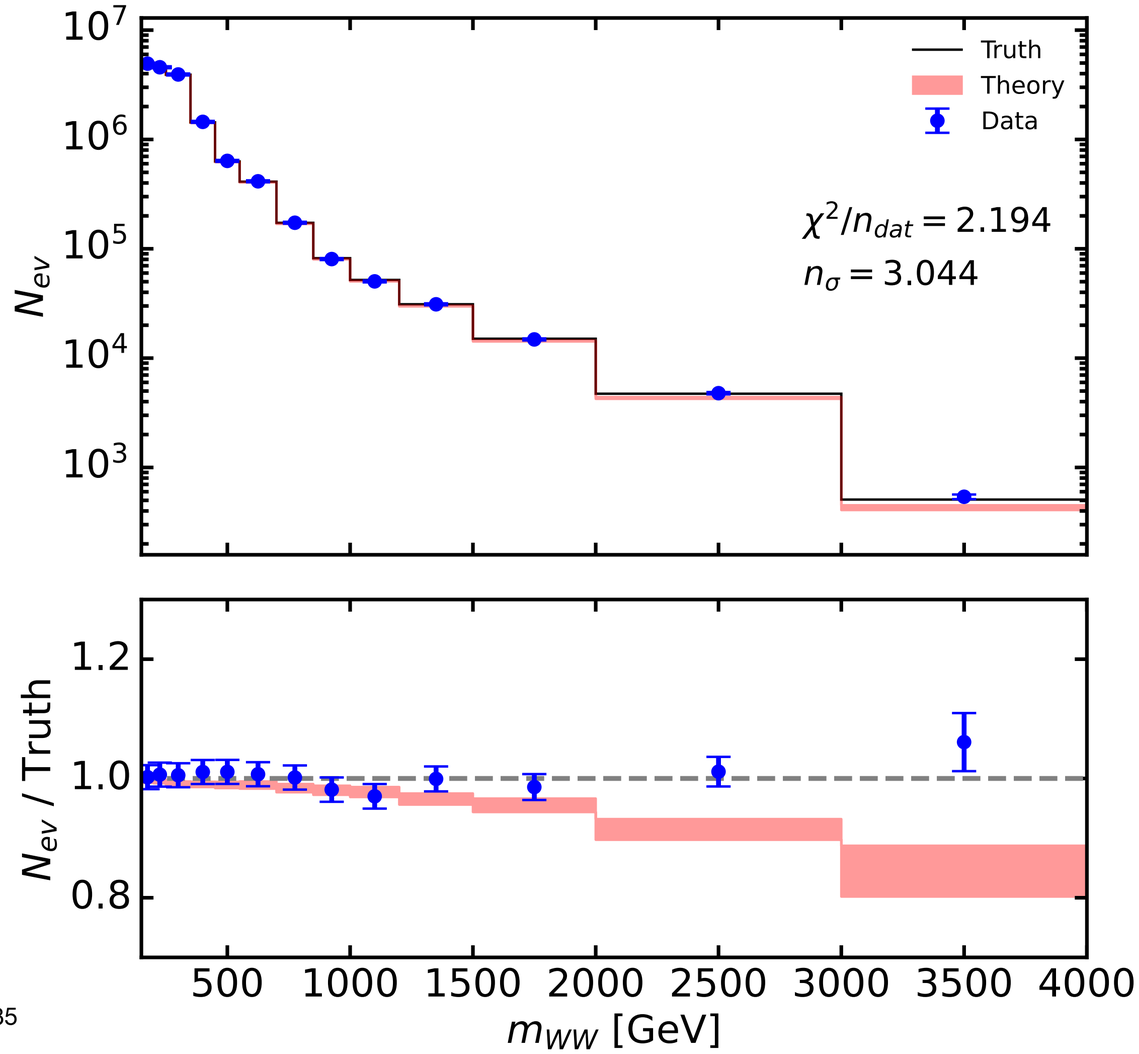
➡ Fake deviation in other sectors

Also seen in:

WH, WZ, ZH production

HL-LHC Projections

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

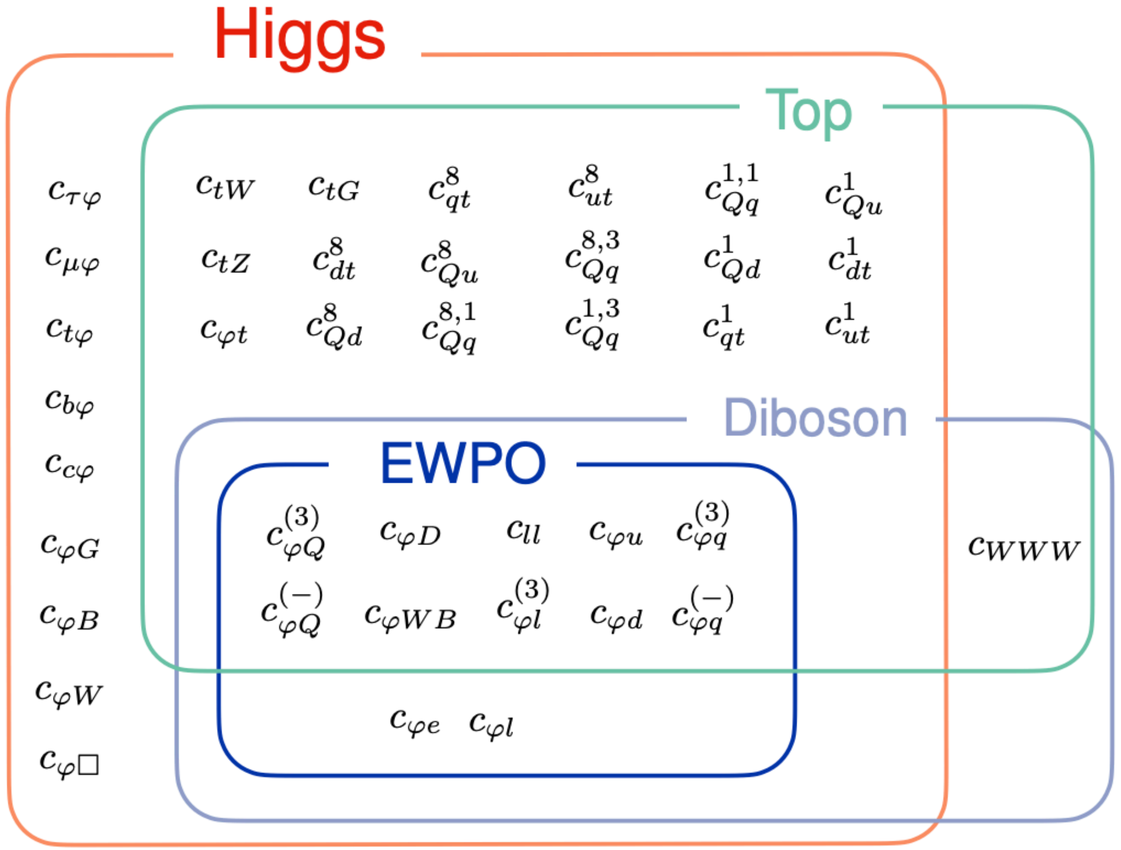


Simultaneous fit of PDF and new physics

SMEFT operators implemented

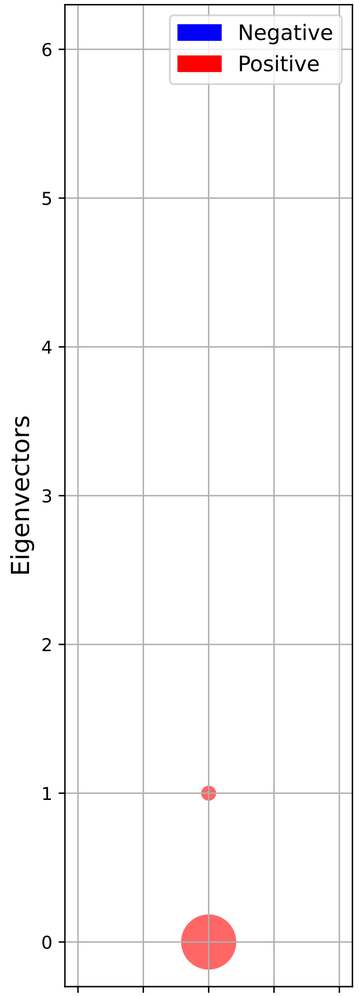
- 40 operators implemented

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

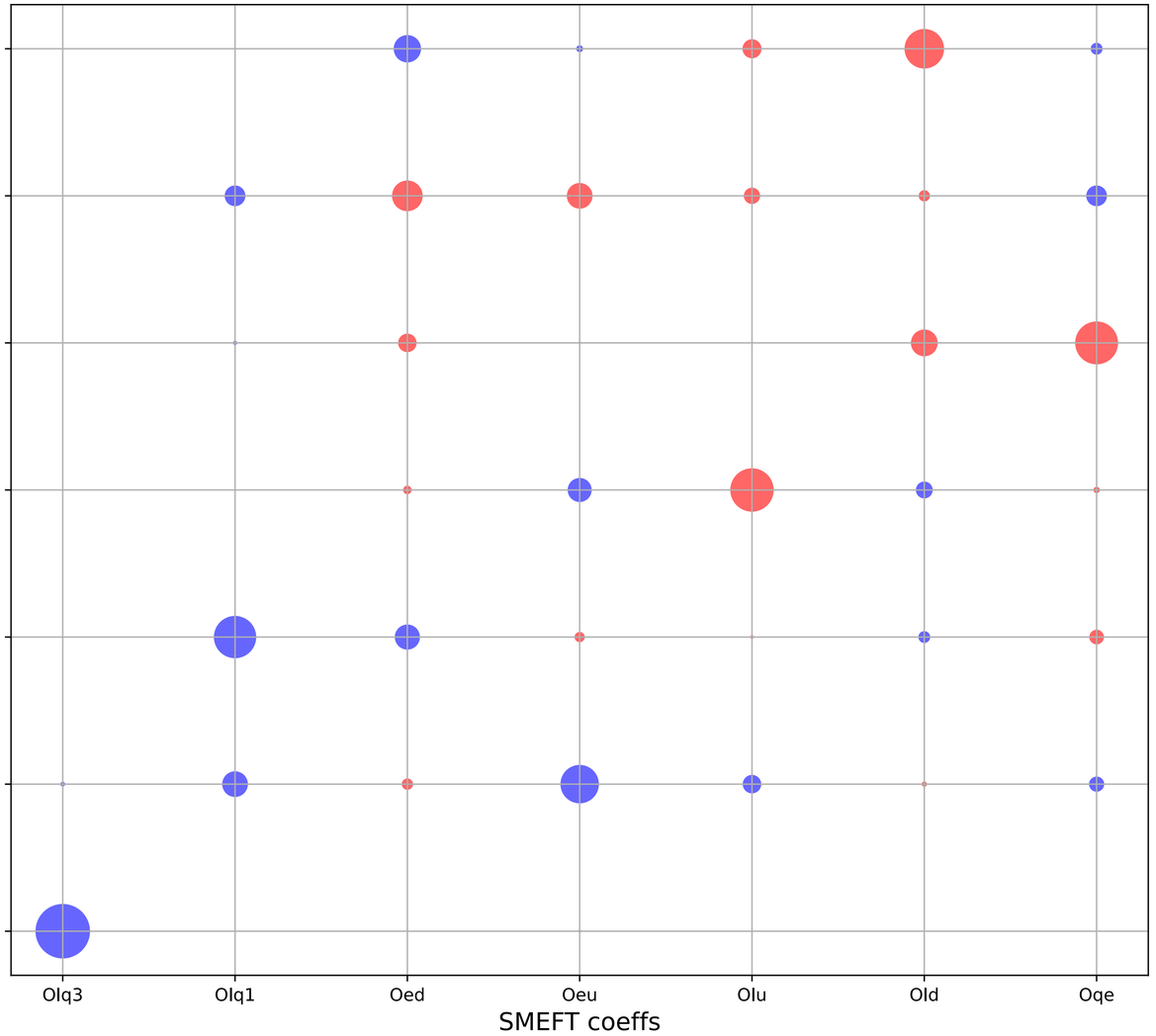


SMEFT PCA results

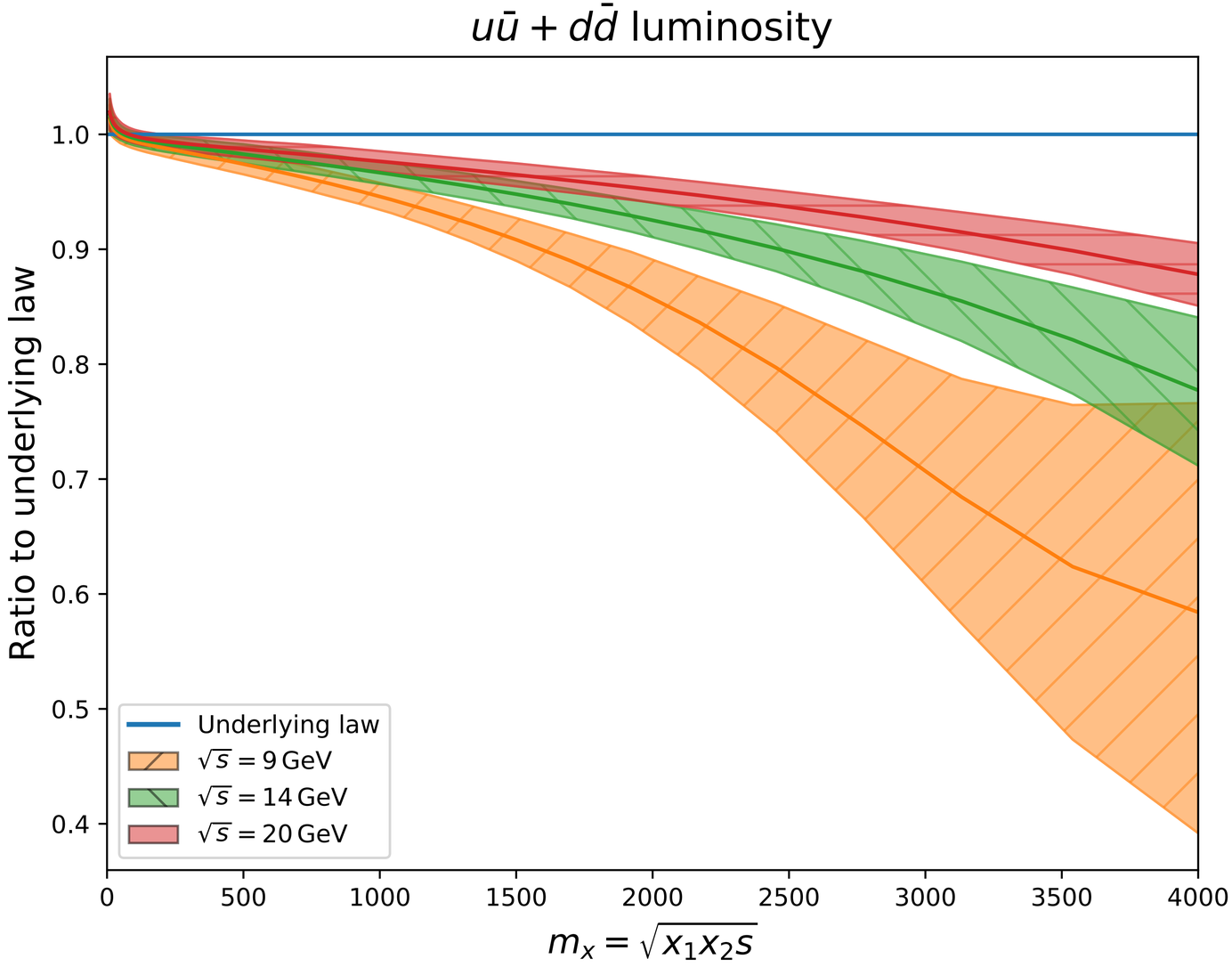
FIM Eigenvalues



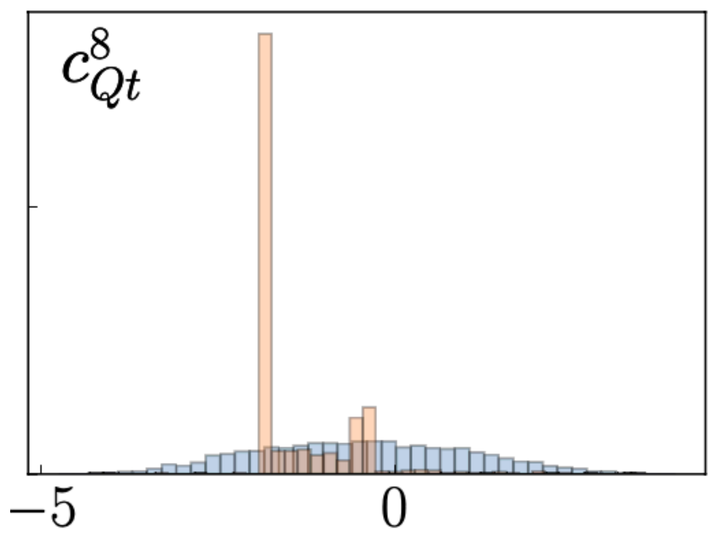
FIM eigenvectors and SMEFT operators

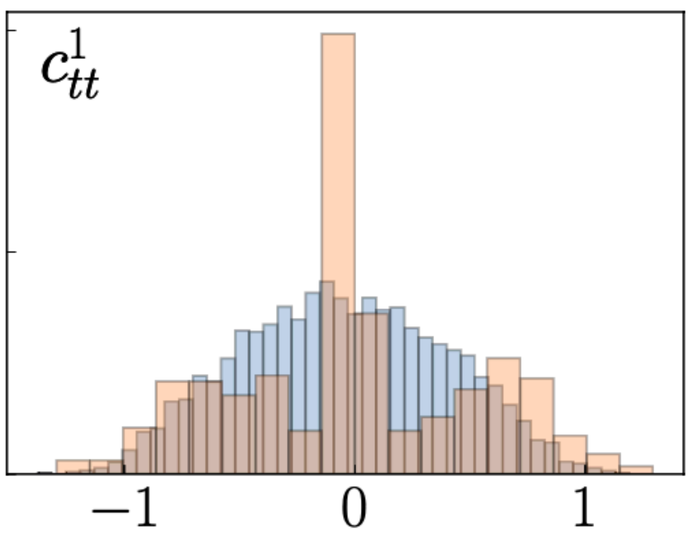


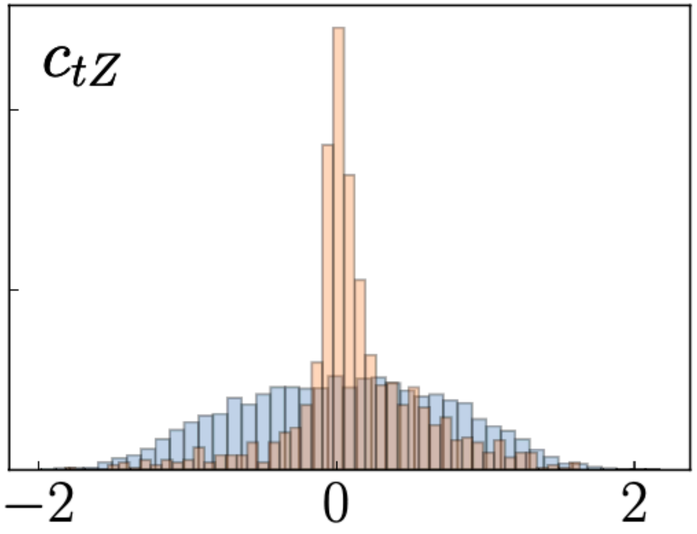
Centre-of-mass comparison



BEYOND MONTE CARLO AND NEURAL NETWORKS







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

- Study of MC versus Bayesian method based on nested sampling for PDF fits and SMEFT fits [Costantini, Madigan, Mantani, Moore arXiv:2404.10056]

- Towards a general Bayesian methodology for simultaneous fits [Costantini, Mantani, MU, in progress]

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

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