PDFs and SMEFT interplay

[Hammou et al. (PBSP + Mangano), 2307.10370, JHEP]

[Costantini et al. (PBSP), 2402.03308, Eur.Phys.J.C]

[Cole et al. (PBSP), forthcoming]

European Research Council

Established by the European Commission

Funded by
the European Union

PBSP

Elie Hammou

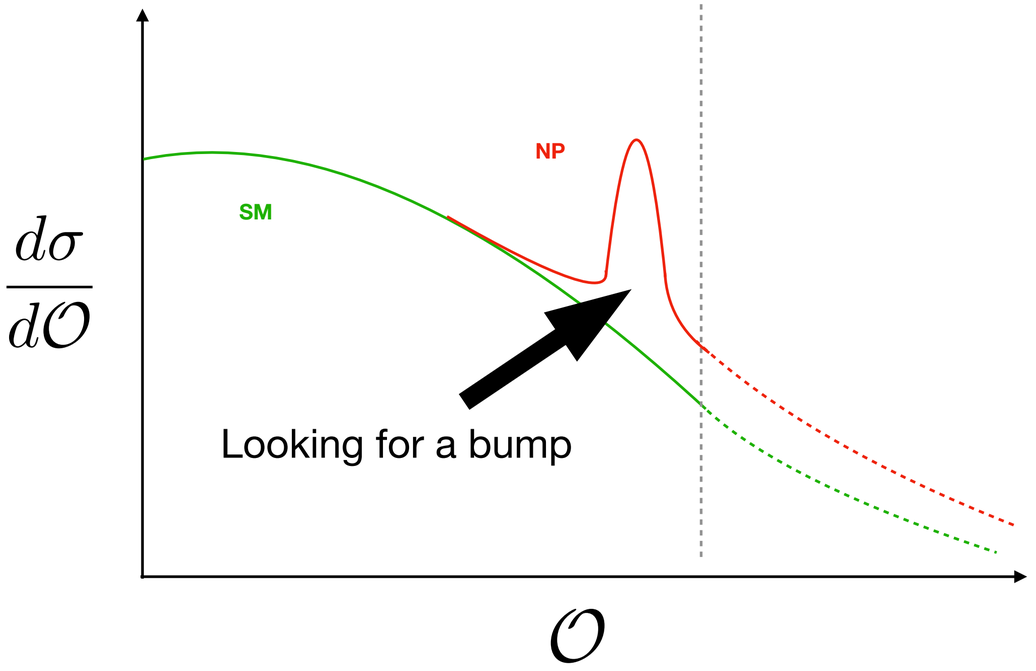
ATLAS UK PDF meeting, Queen Mary University of London, September 2025

New physics searches

Looking toward higher energy scales and indirect searches

Direct searches

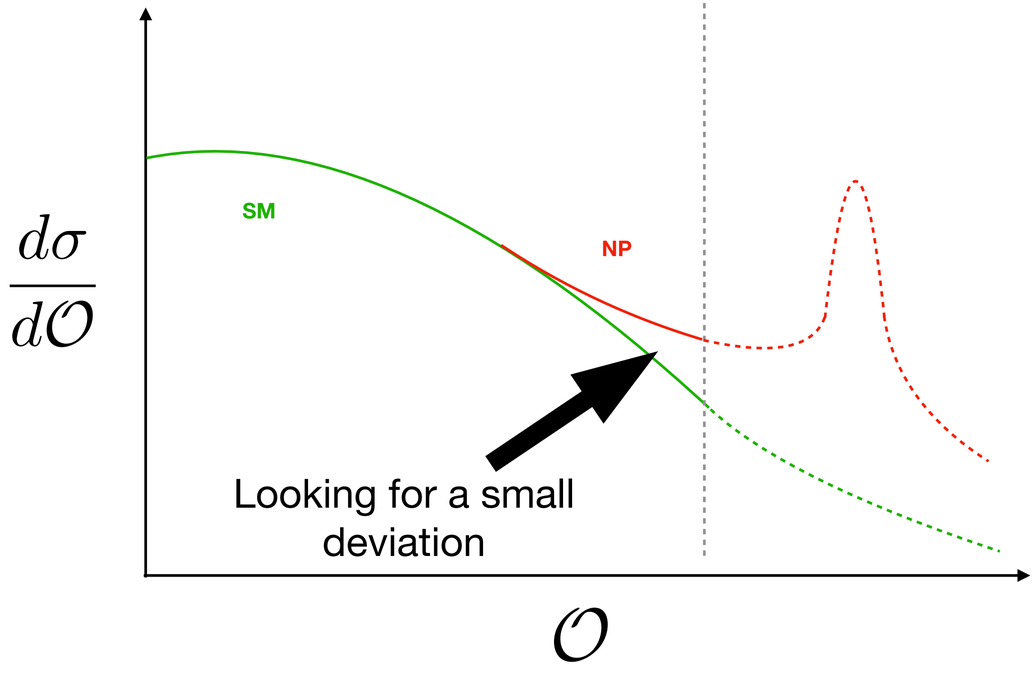
$E_{NP} < E_{collider}$



No luck so far…

Indirect searches

$E_{NP} > E_{collider}$



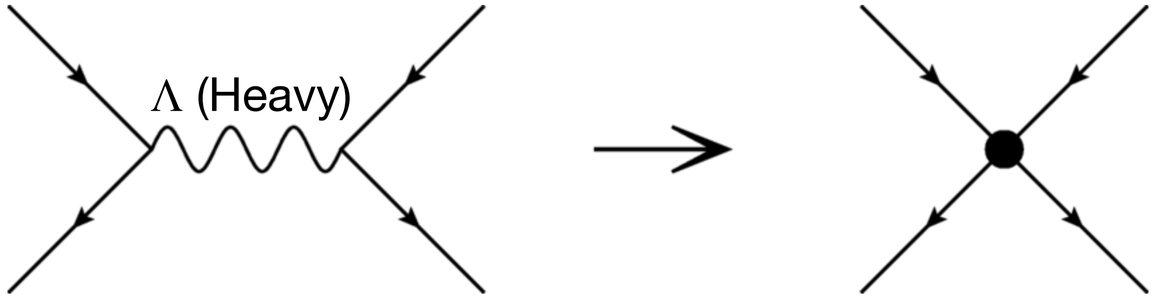
Requires precision

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]

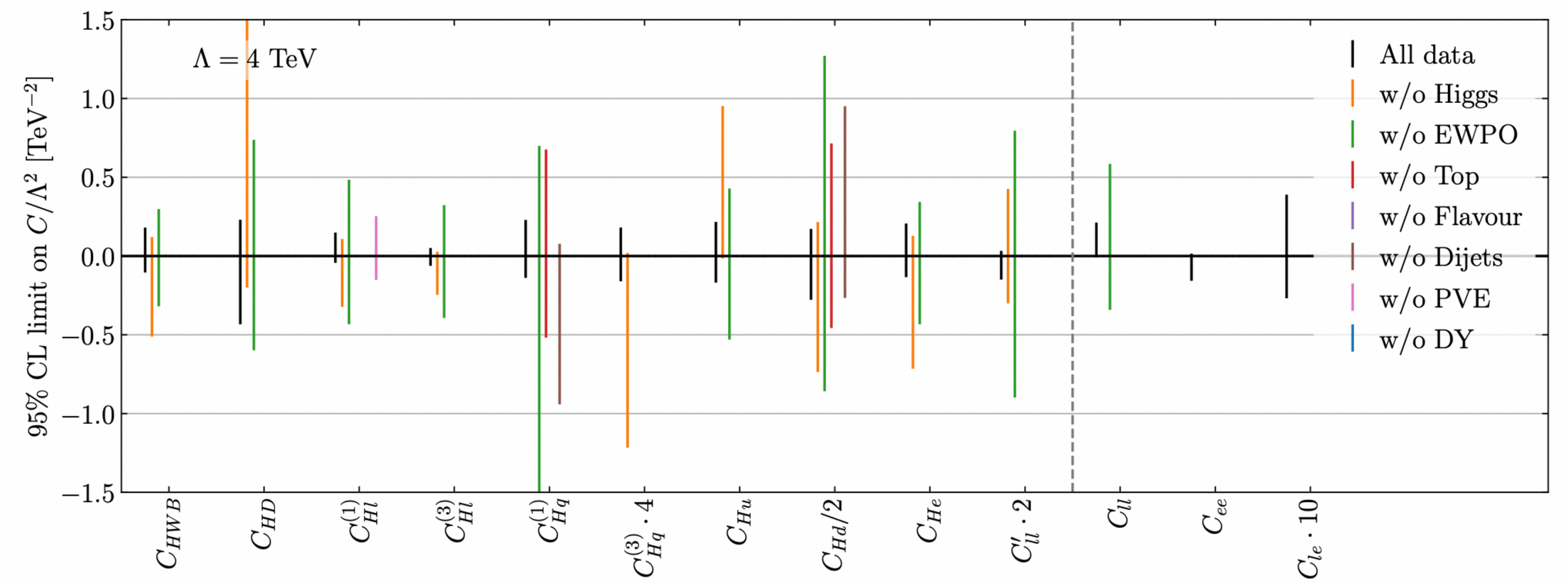
Obtain model independent Lagrangian:

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

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

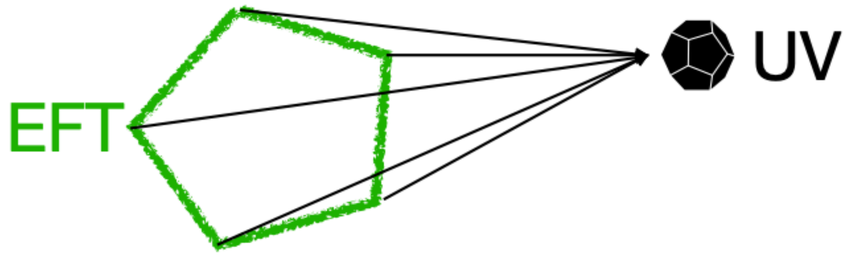
SMEFT fit from data

[SMEFiT, 2302.06660]



Can fit $\{\frac{c_i}{\Lambda^2}\}$:

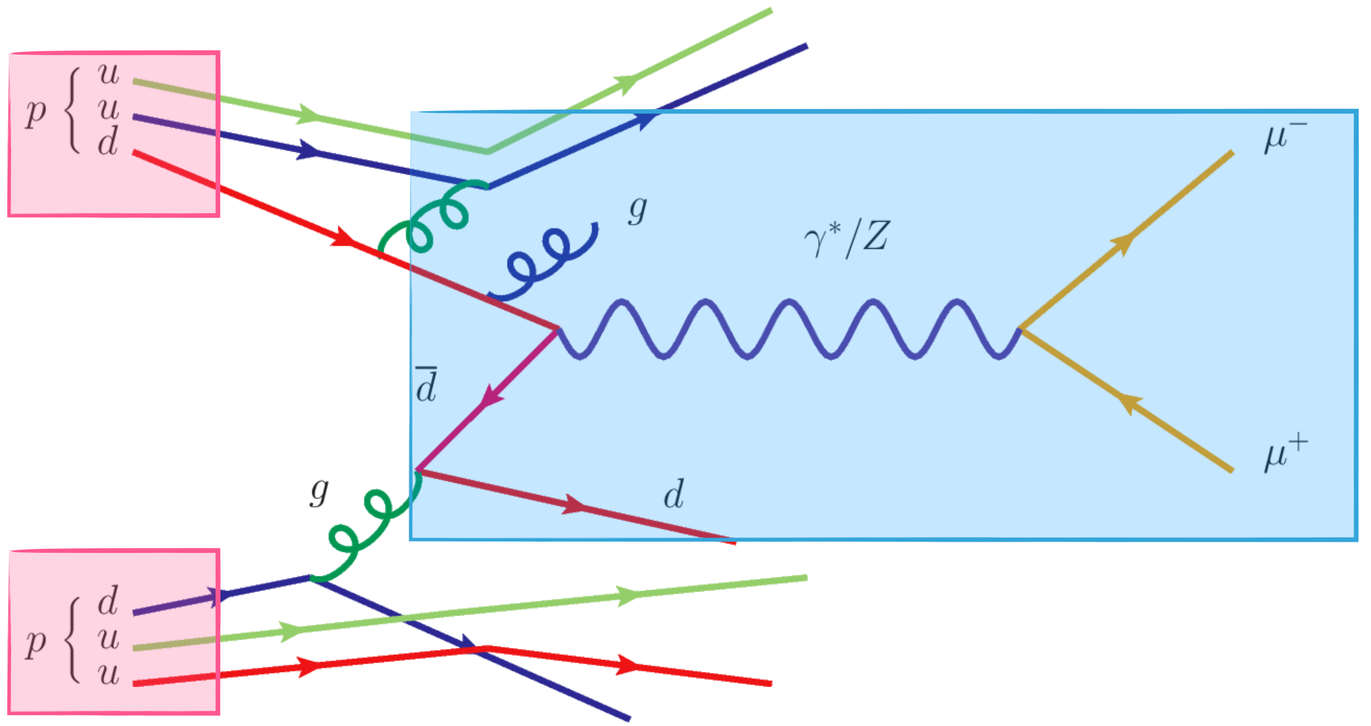
➡ can then be matched to a UV model



Hadron colliders, SMEFT and PDFs

Collinear factorization theorem

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



Parton Distribution Functions

Hard Scattering: Perturbative QCD + EW

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

Injecting new physics in HL-LHC projections

Description of the models

High-mass Drell-Yan

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

- ➡ W'^{+} / W'^{-} and W_3'
- ➡ matching to SMEFT: oblique parameter $\hat{W}$

Heavy neutral boson under $U(1)_Y : Z'$

- ➡ Impacts only neutral current DY
- ➡ matching to SMEFT: oblique parameter $\hat{Y}$

Top pair production:

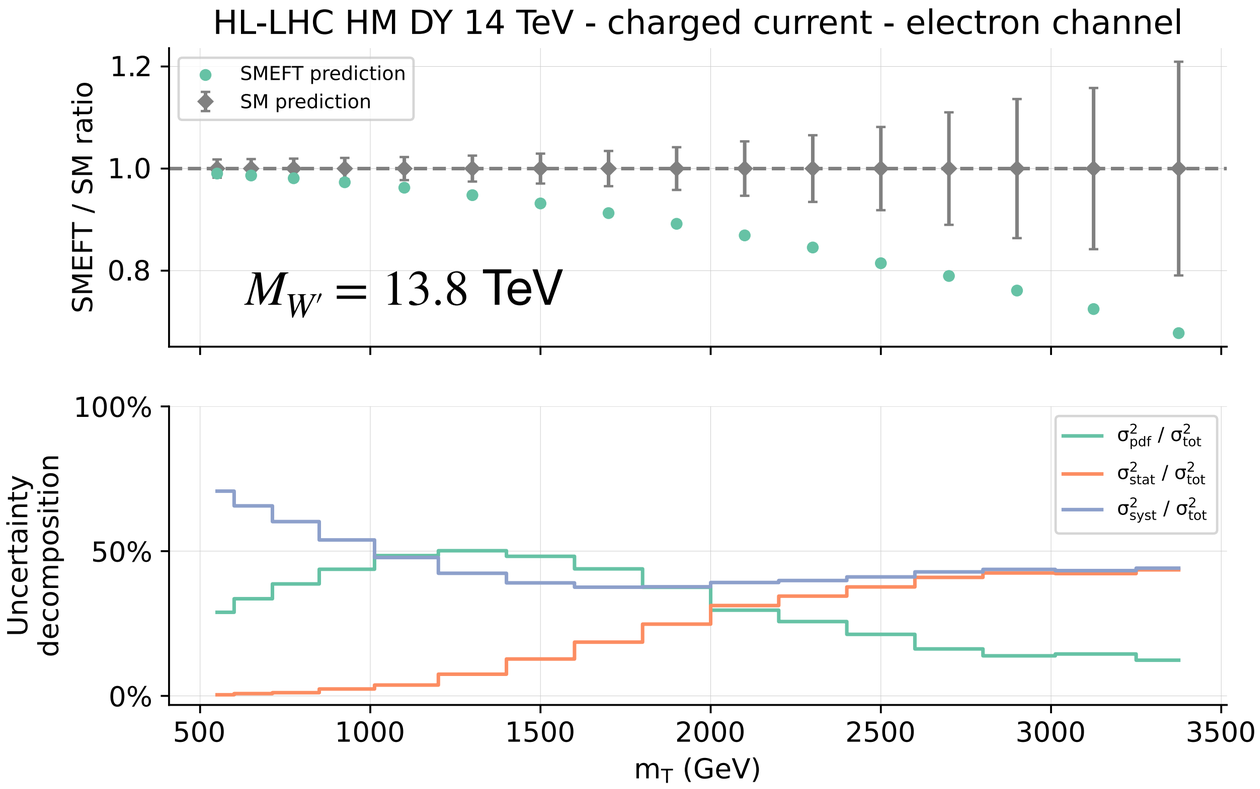
Heavy octet under $SU(3)_C : G'$

- ➡ Couples to third quark family
- ➡ matching to SMEFT: oblique parameter $\hat{Z}$

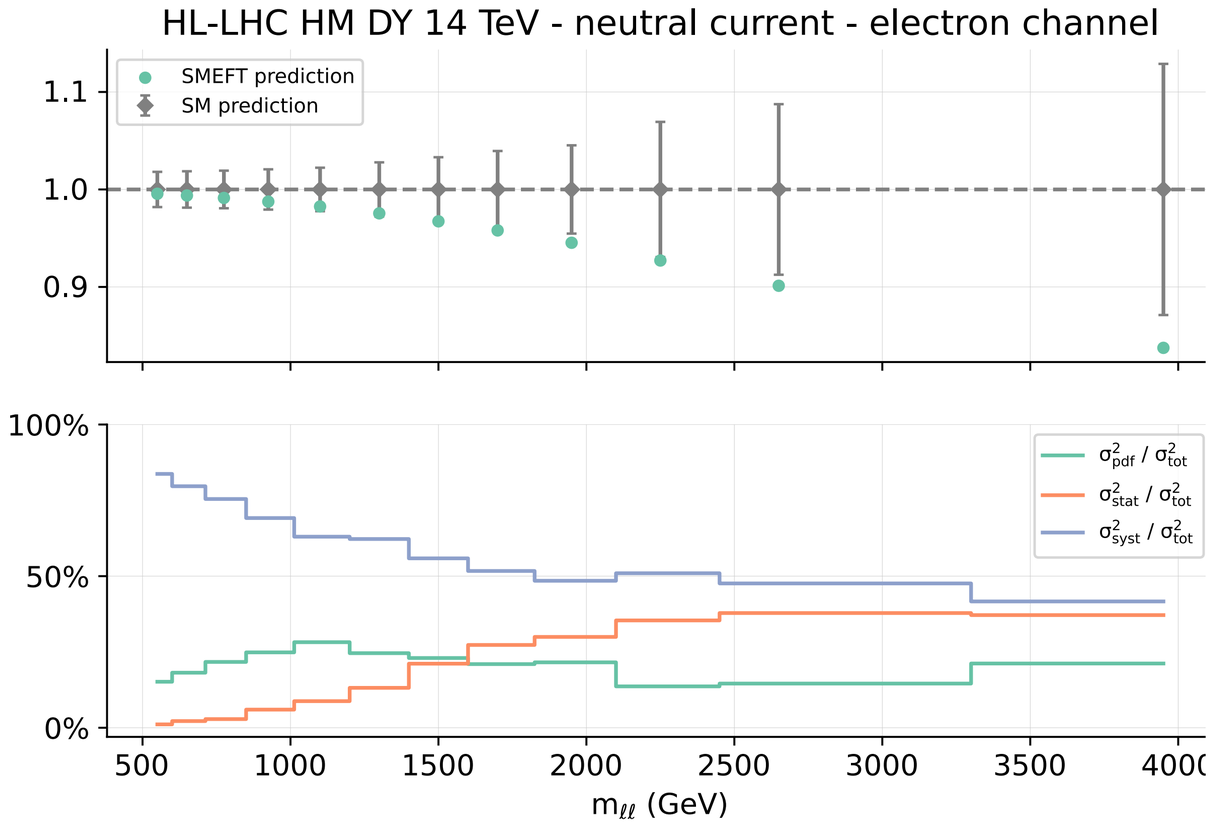
Injecting new physics in HL-LHC projections (DY)

Goal: recover it with a SMEFT fit

Charged current HMDY



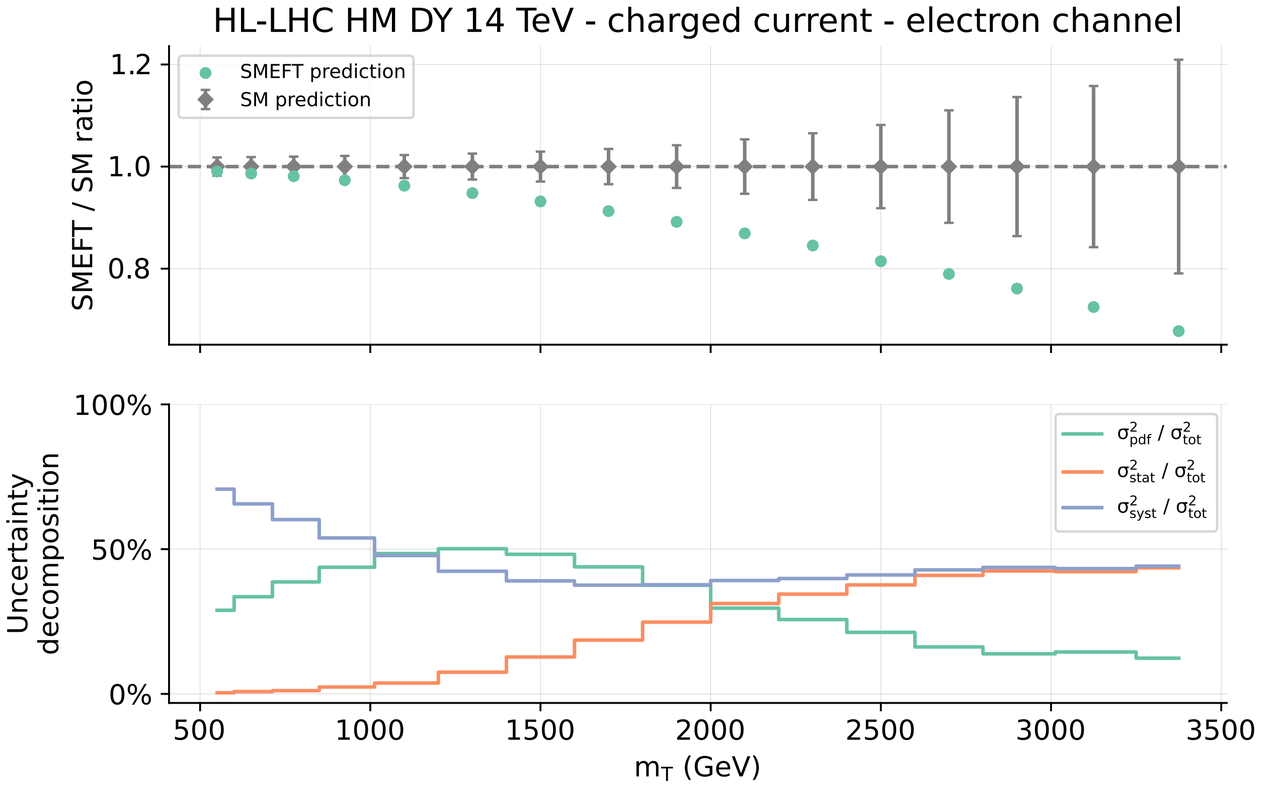
Neutral current HMDY

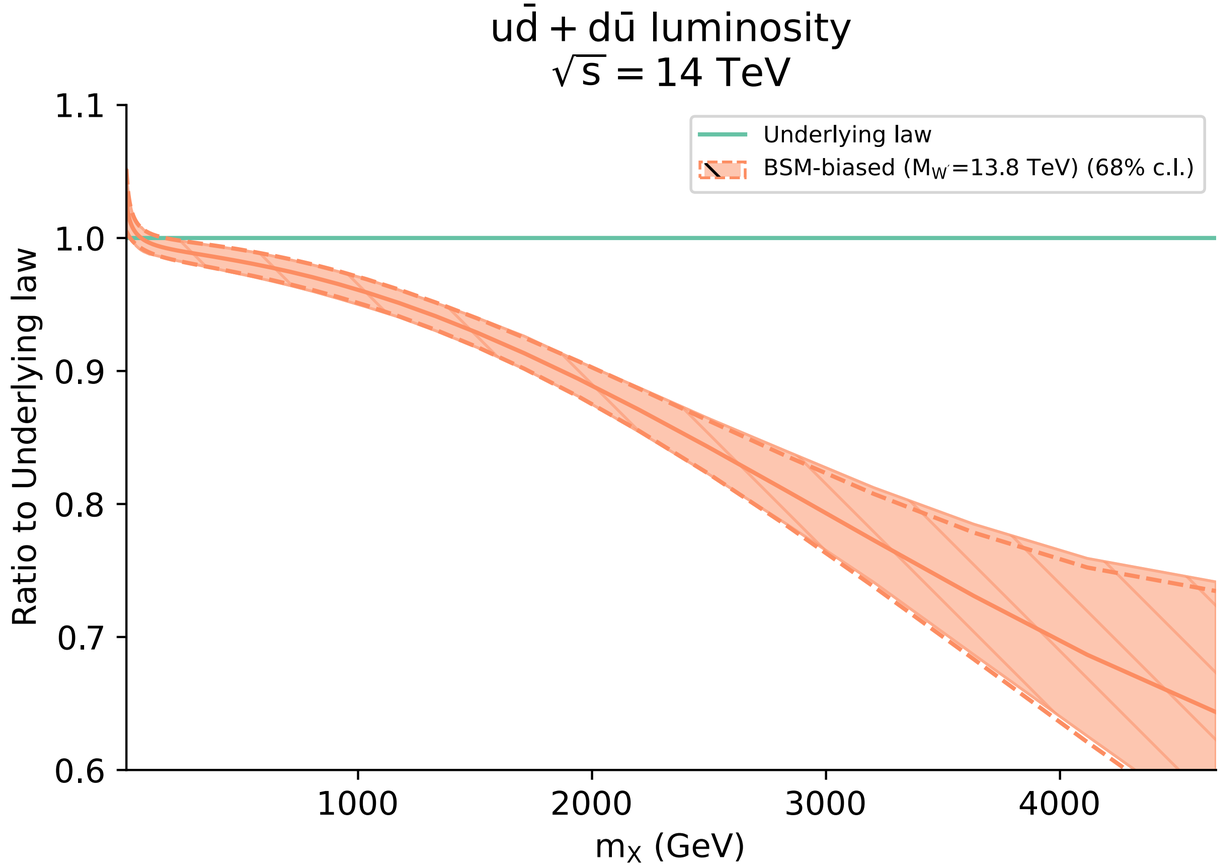


“BSM-biased” PDF in HMDY

PDFs are mimicking the SMEFT corrections

Charged current HMDY



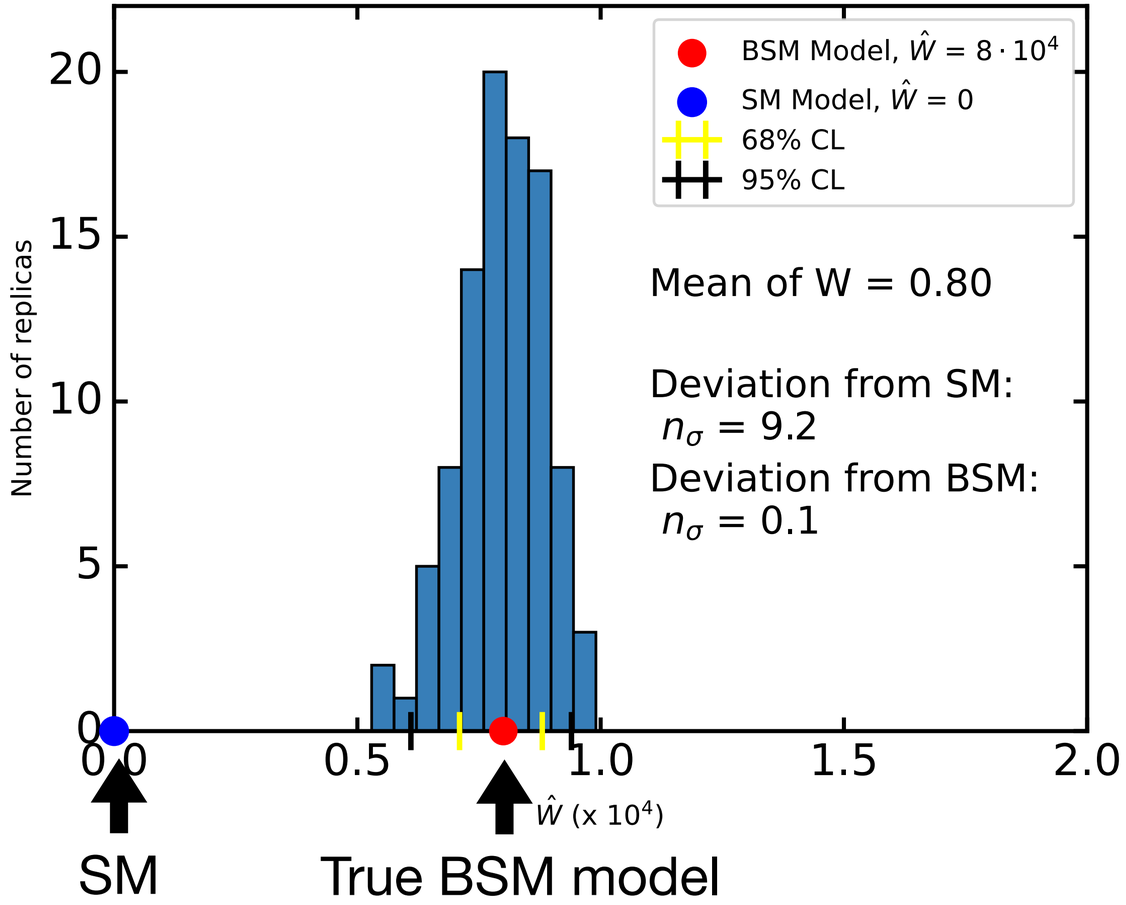


$$\mathscr{L}_{true} \otimes \hat{\sigma}_{BSM} \approx \mathscr{L}_{BSM-biased} \otimes \hat{\sigma}_{SM}$$

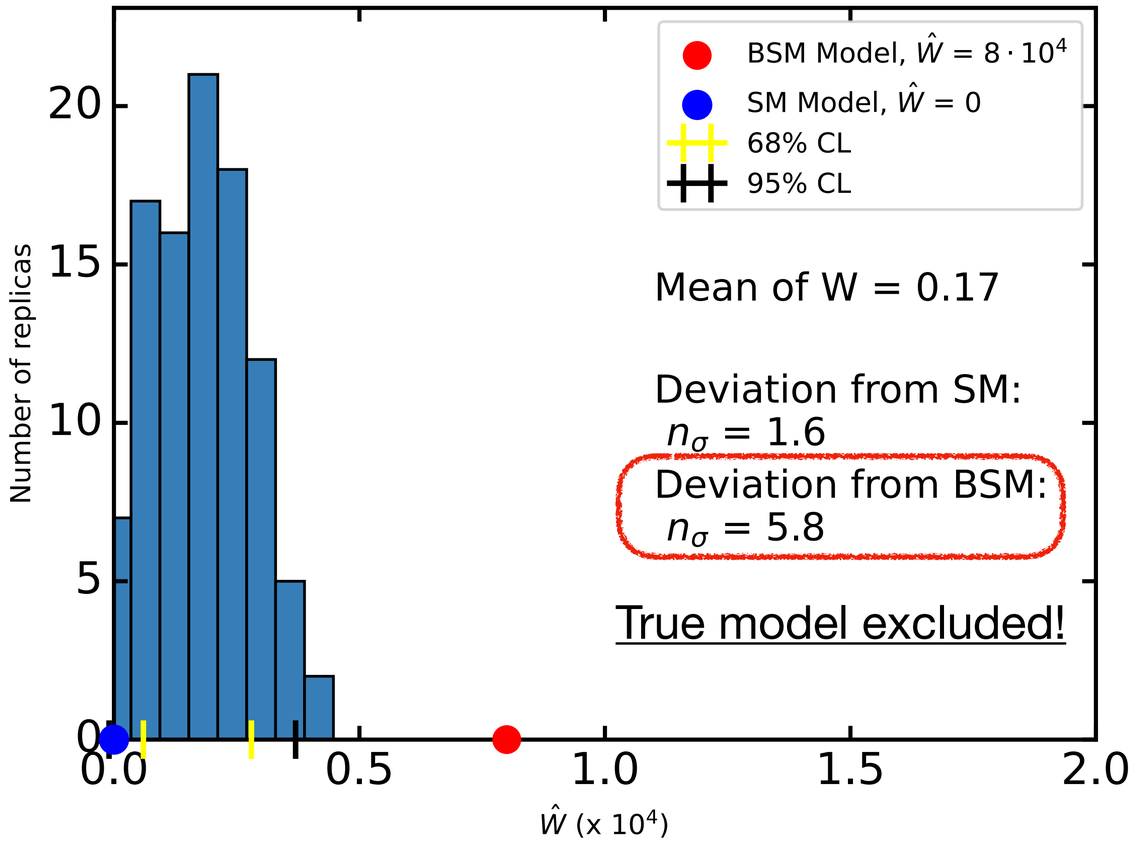
Missing new physics in Drell-Yan

Impact of the BSM-biased PDFs on SMEFT fits

SMEFT fit with true PDF

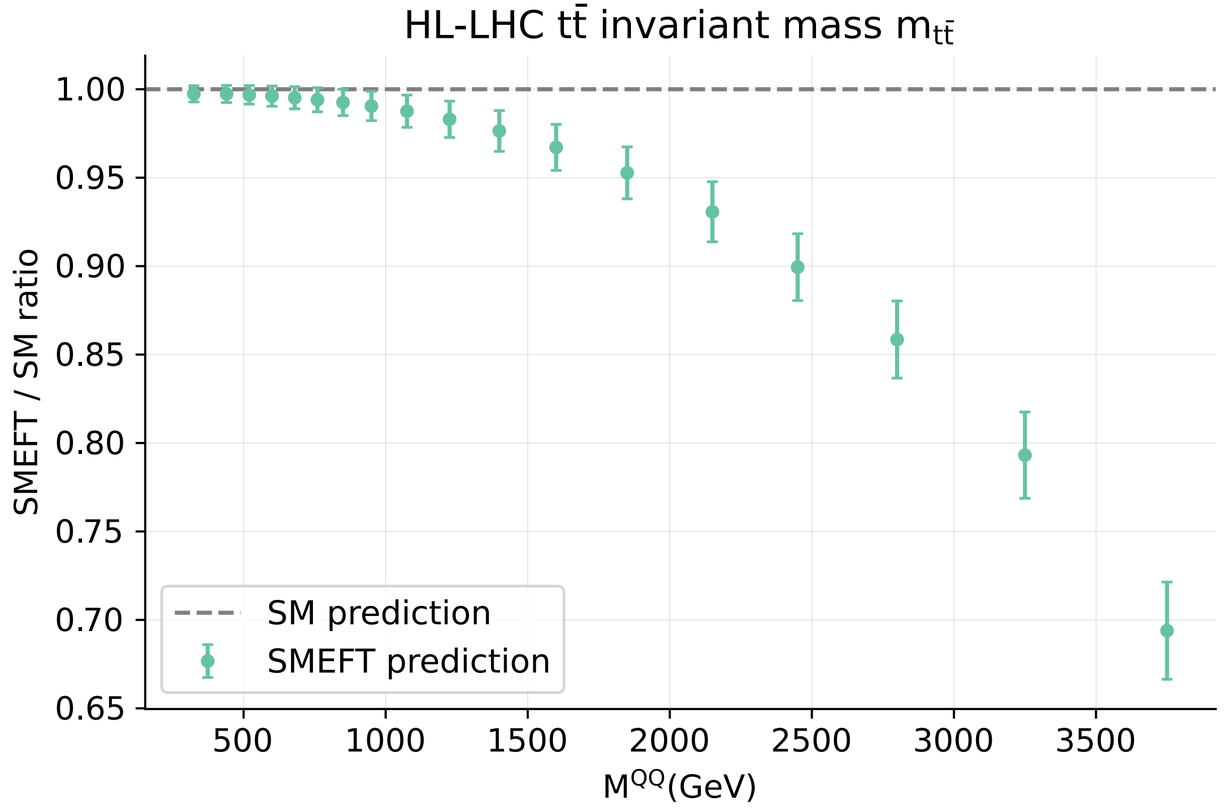


SMEFT fit with BSM-biased PDF

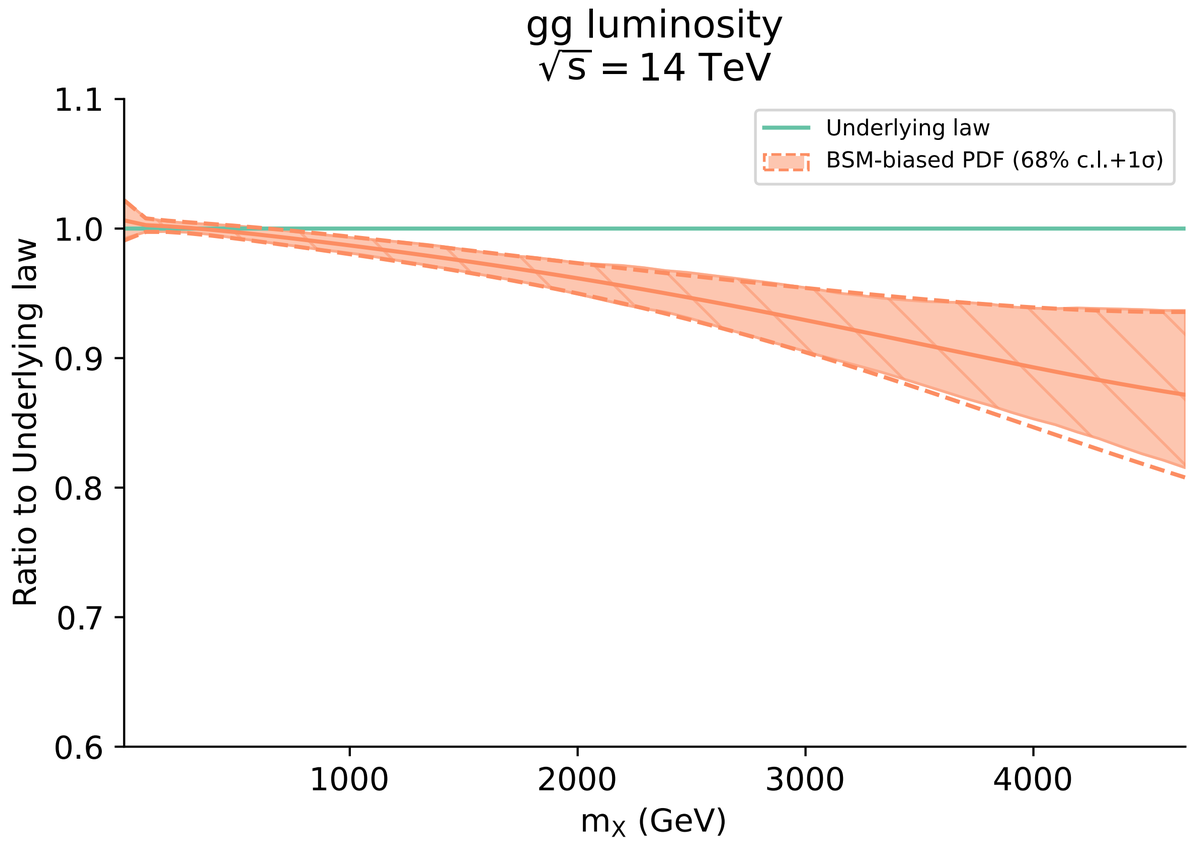


“BSM-biased” PDF in $t\bar{t}$

PDFs are partially mimicking the SMEFT corrections



$M_{G'} = 8$ TeV

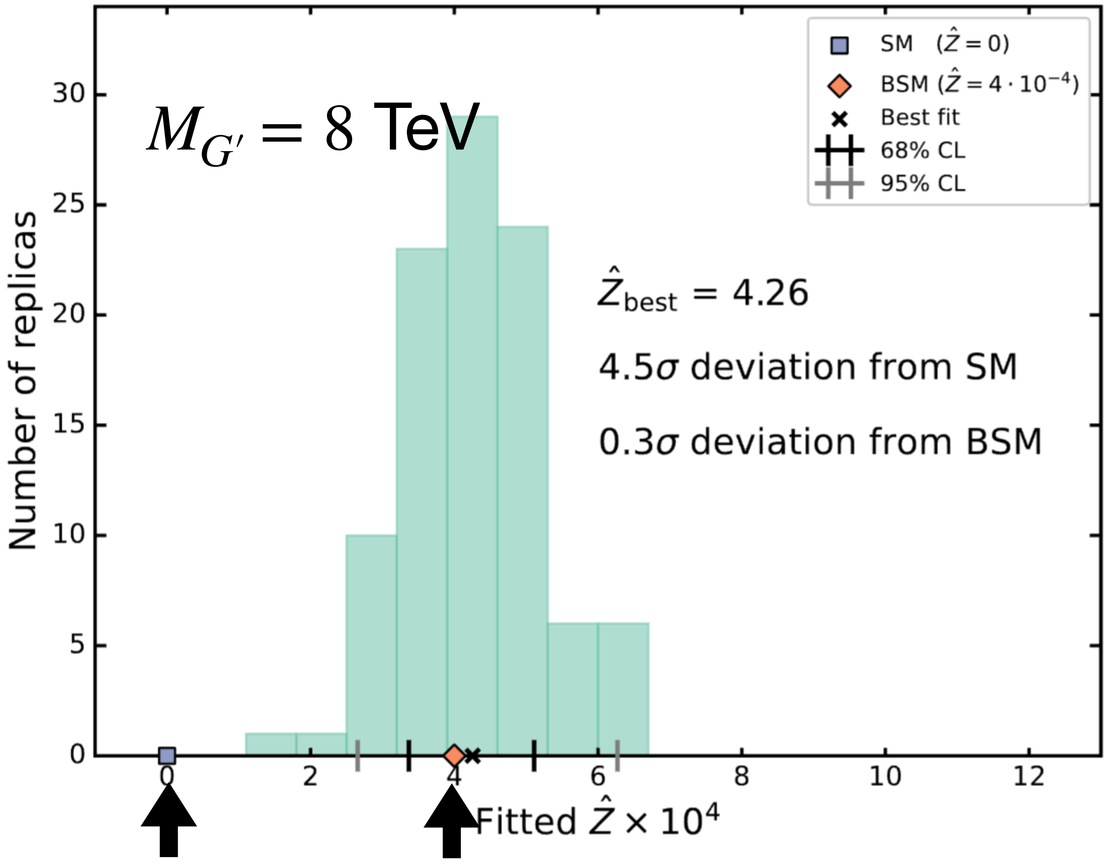


$$\mathscr{L}_{true} \otimes \hat{\sigma}_{BSM} \approx \mathscr{L}_{BSM-biased} \otimes \hat{\sigma}_{SM}$$

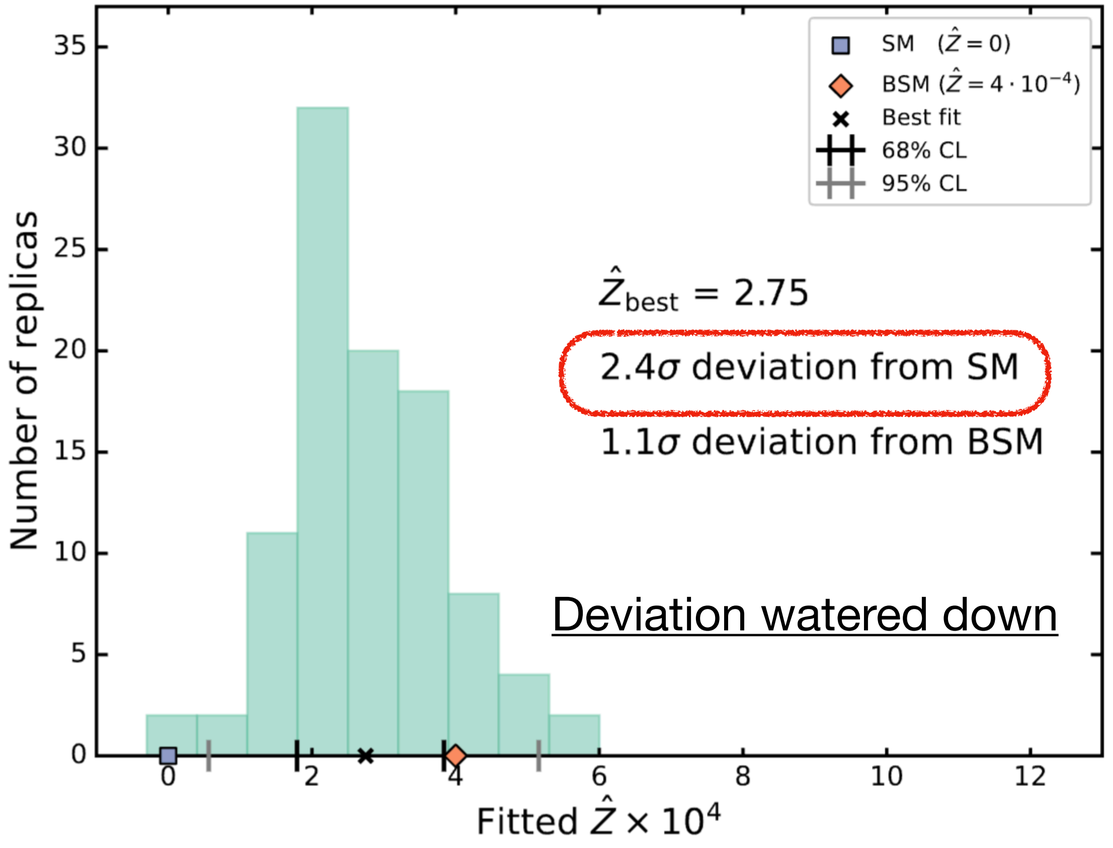
Diluting new physics deviation in the top sector

Same exercise with heavy gluon G' in ttbar at HL-LHC

SMEFT fit with true PDF



SMEFT fit with BSM-biased PDF



Separate and simultaneous fits of PDF and SMEFT

Comparison of the methodologies

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:

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

➡ $\bar{c}$

Need to make sure SM assumption is true

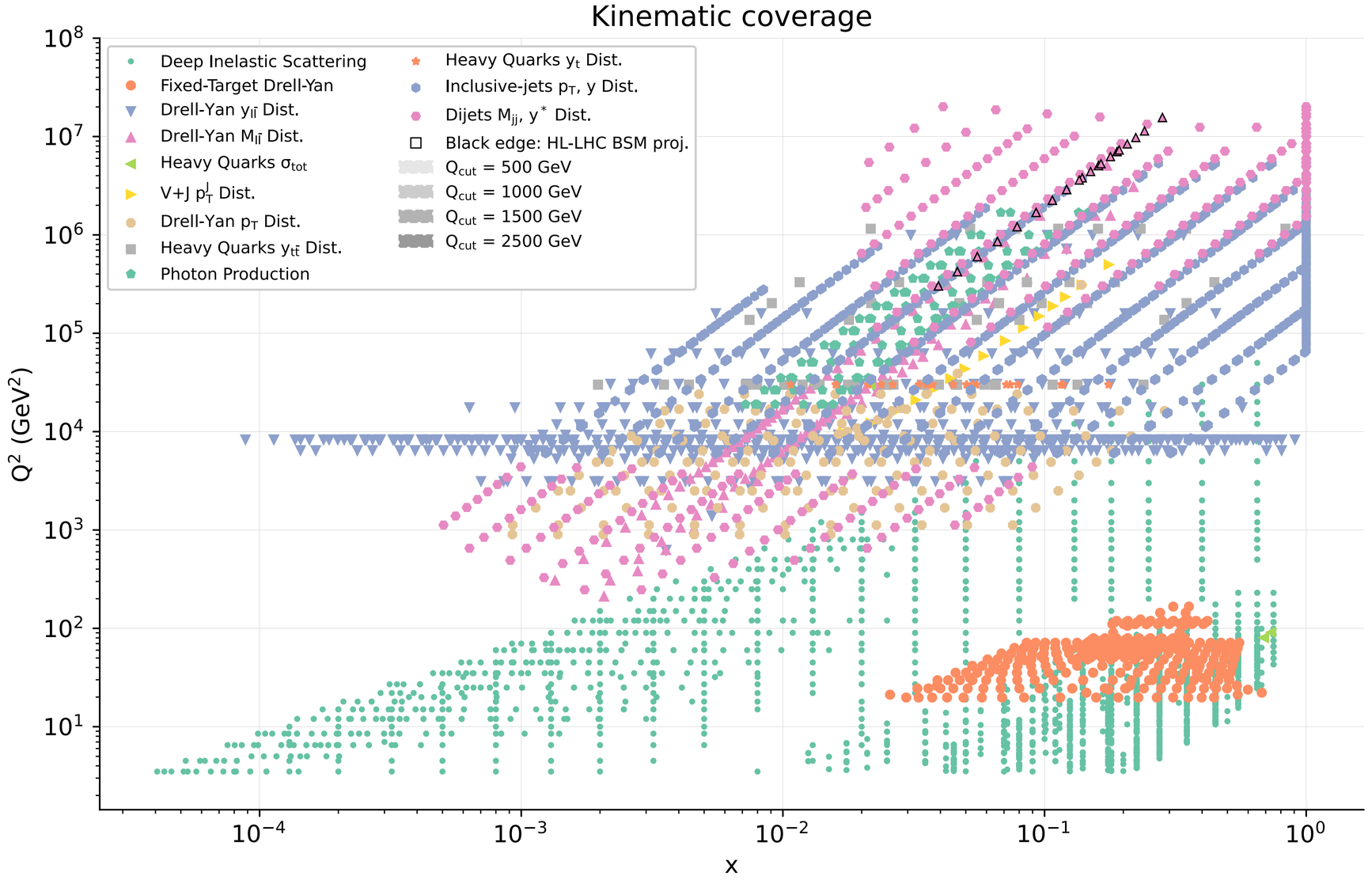
Simultaneous fits

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

➡ $\{\bar{\theta}, \bar{c}\}$

Removes assumption-based bias

Safe separate fits: “conservative” PDFs

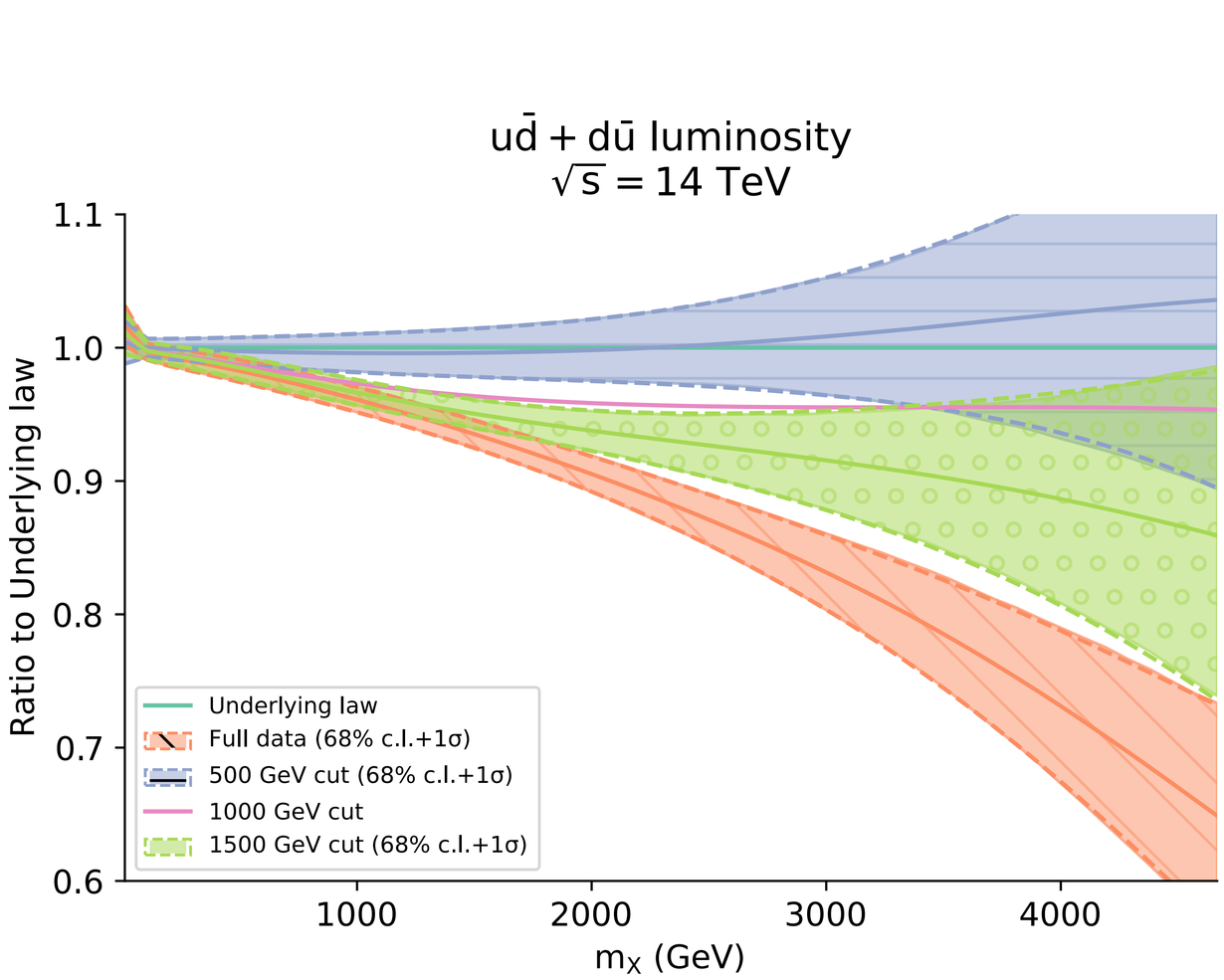


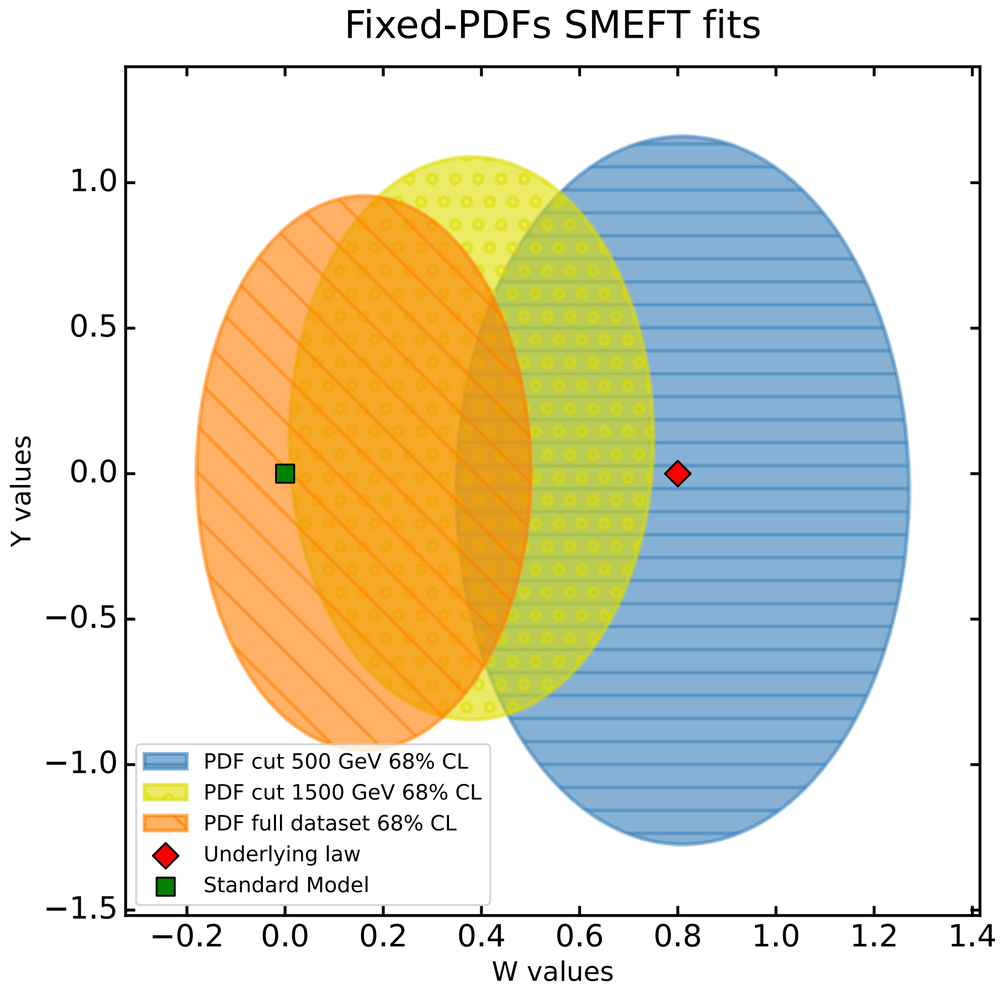
Agnostic energy cuts

“Conservative” separate fits

[PBSP, forthcoming]

Impact on PDF and SMEFT bounds for DY





Simultaneous fit of PDF and SMEFT

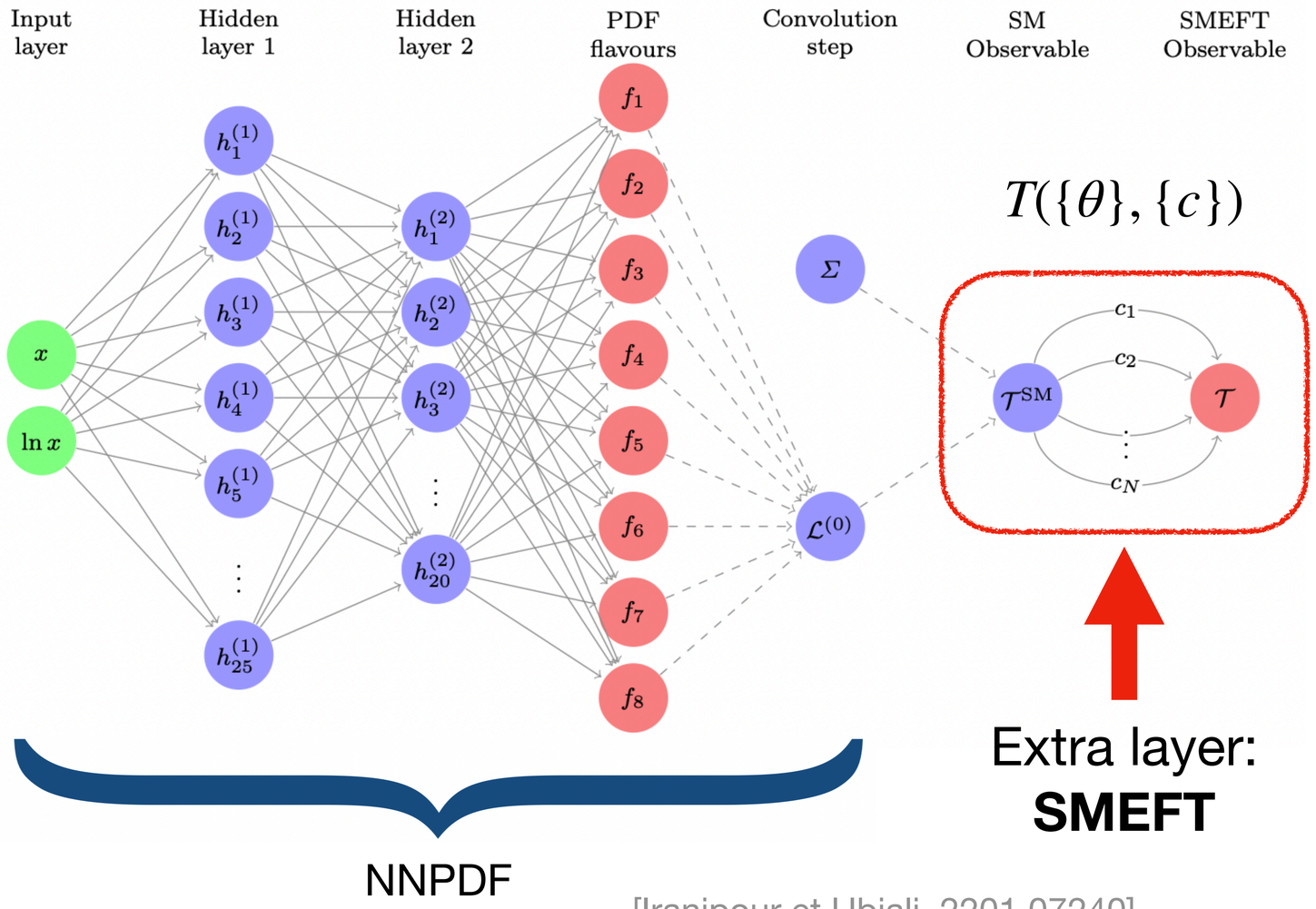
Presentation of the tool: SIMUnet

SIMUnet:

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

  [PBSP, 2402.03308]

- Fits PDFs and WC simultaneously
- Performs BSM closure tests

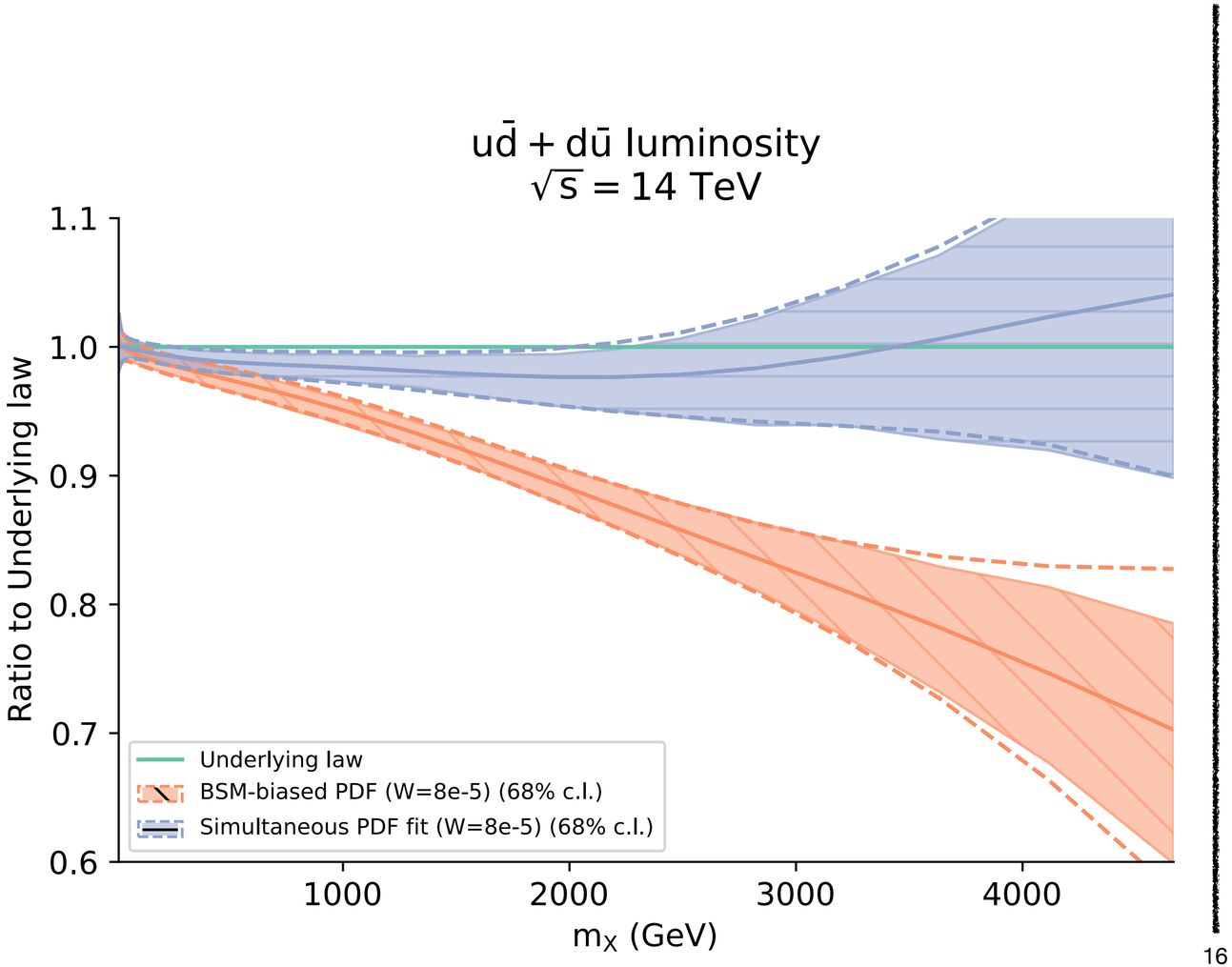


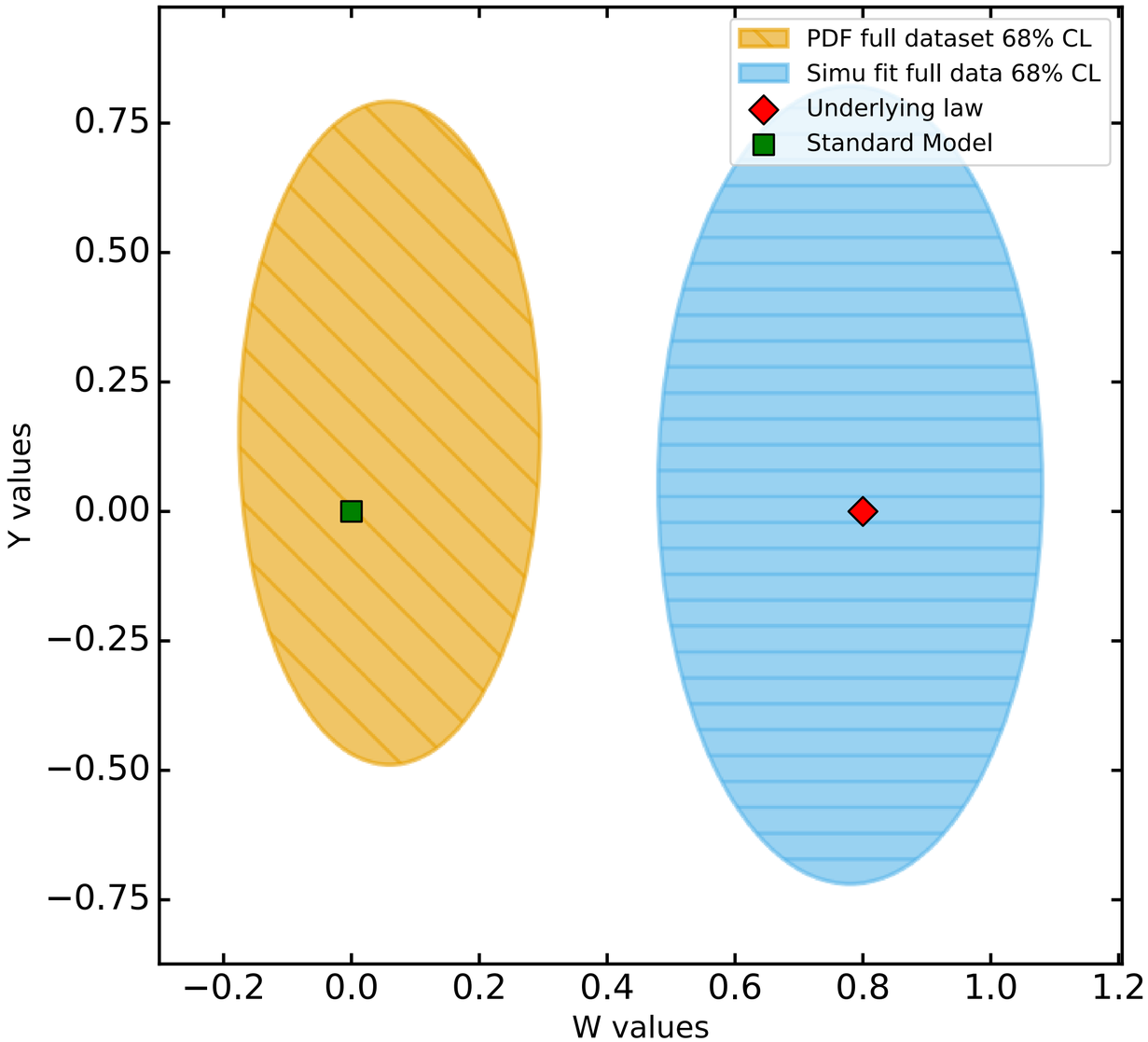
[Iranipour et Ubiali, 2201.07240]

Simultaneous fit of PDF and SMEFT

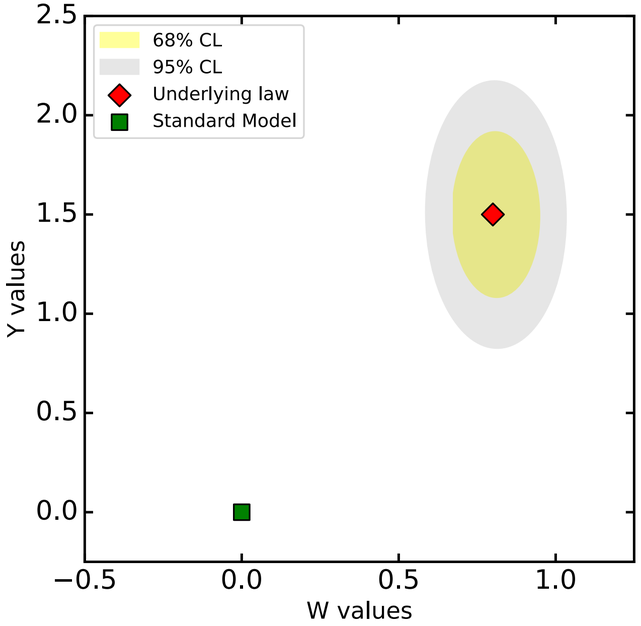
[PBSP, forthcoming]

Marginal distributions for DY





PDFs for new physics searches in HMDY

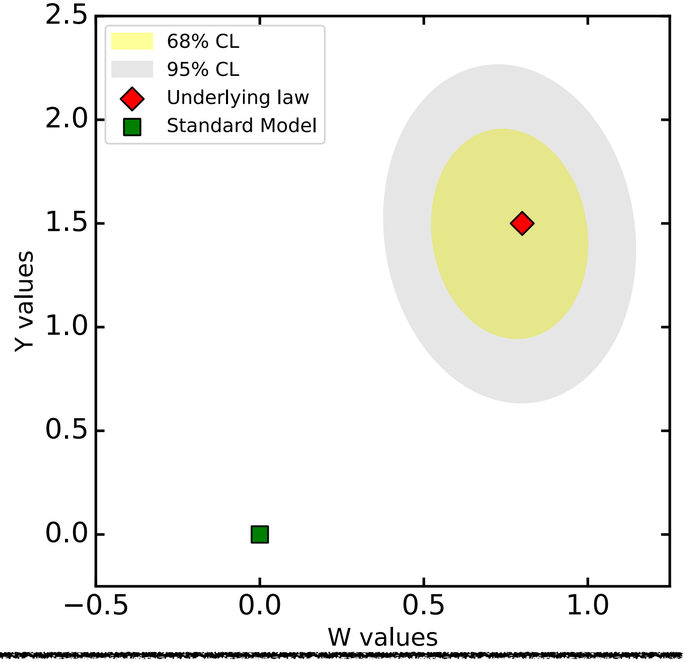


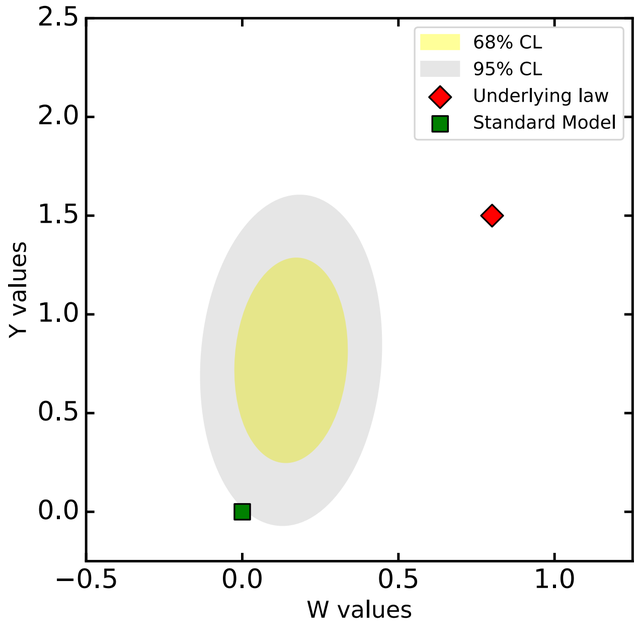
SMEFT only fit
(True PDF)

✗ Impossible

Simultaneous fit
(no PDF assumption)

✓ Doable



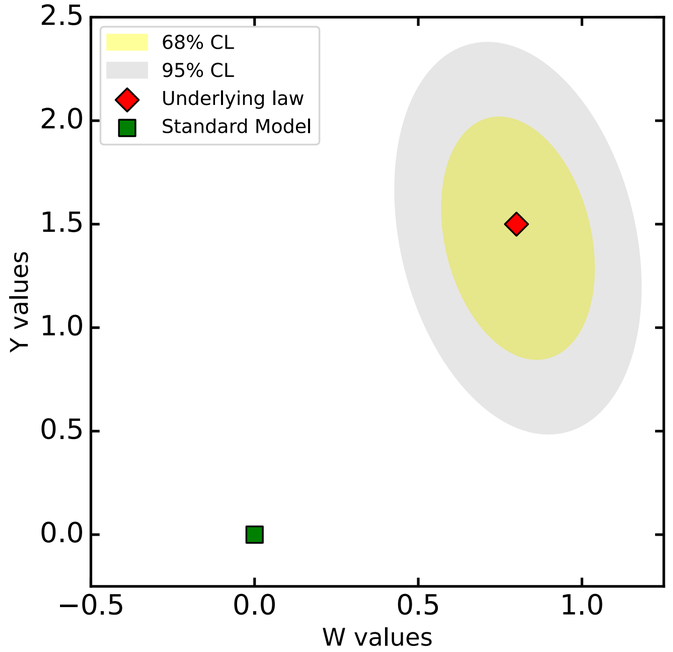


SMEFT only fit
(BSM-biased PDF)

✗ Wrong

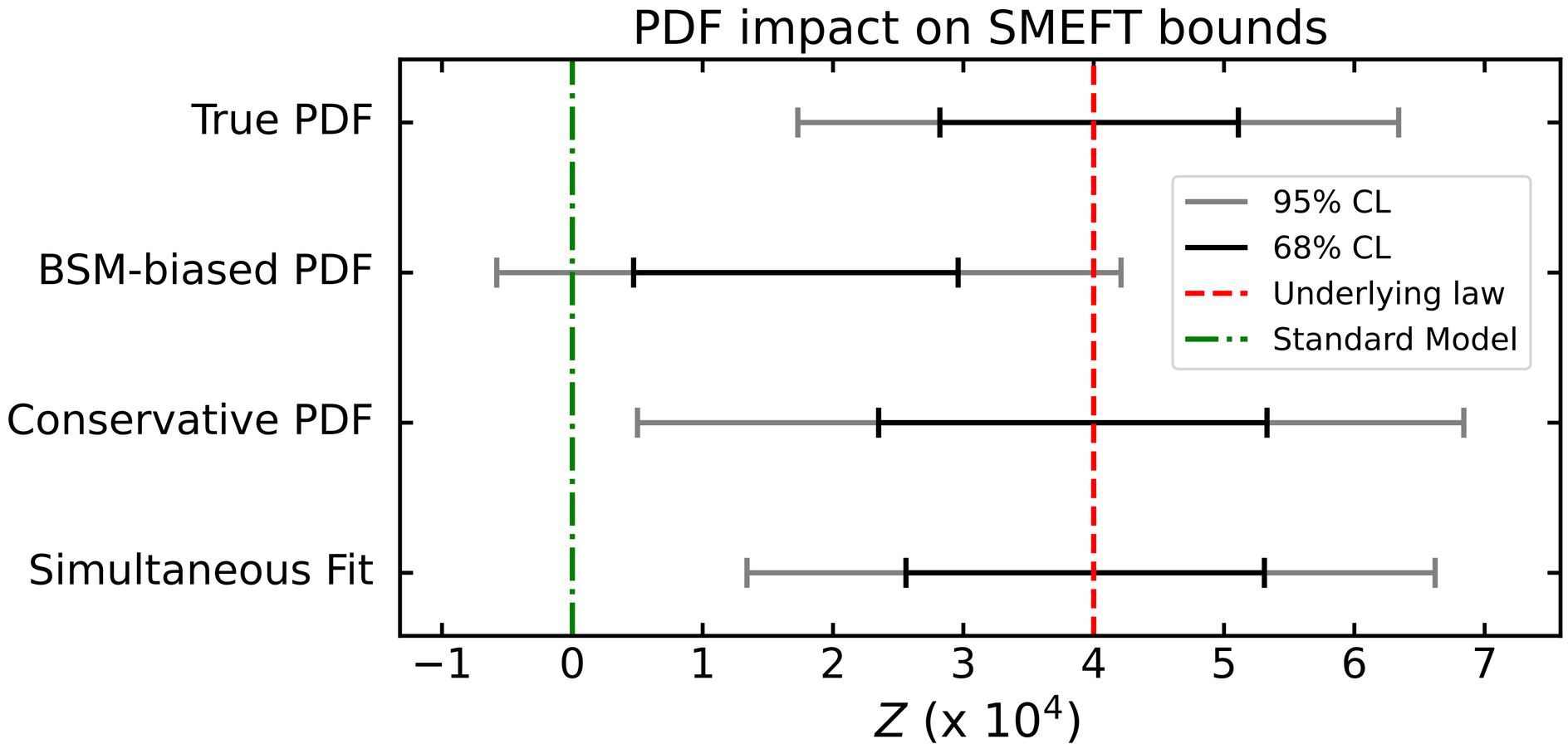
SMEFT only fit
(Conservative PDF)

✓ 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
  - Exclude true underlying law (DY) / water down BSM deviation (ttbar)
  - **Must be aware of this source of bias**
- Practical recommendations:
  - Conservative PDFs for SMEFT fits
  - Simultaneous fits of PDFs and SMEFT (**SIMUnet** tool already available)
- Ongoing work:
  - Ongoing study on real jet data [Greljo, Hammou, Merlotti, Smolkovic, Ubiali, forthcoming]
  - Developing a bayesian framework [Costantini, Moore, Mantani, Schutze, Ubiali, forthcoming (PDF)]
  - Allowing choice in PDF modelisation [Hammou, ter Hoeve, Shutze, in progress (SMEFT)]

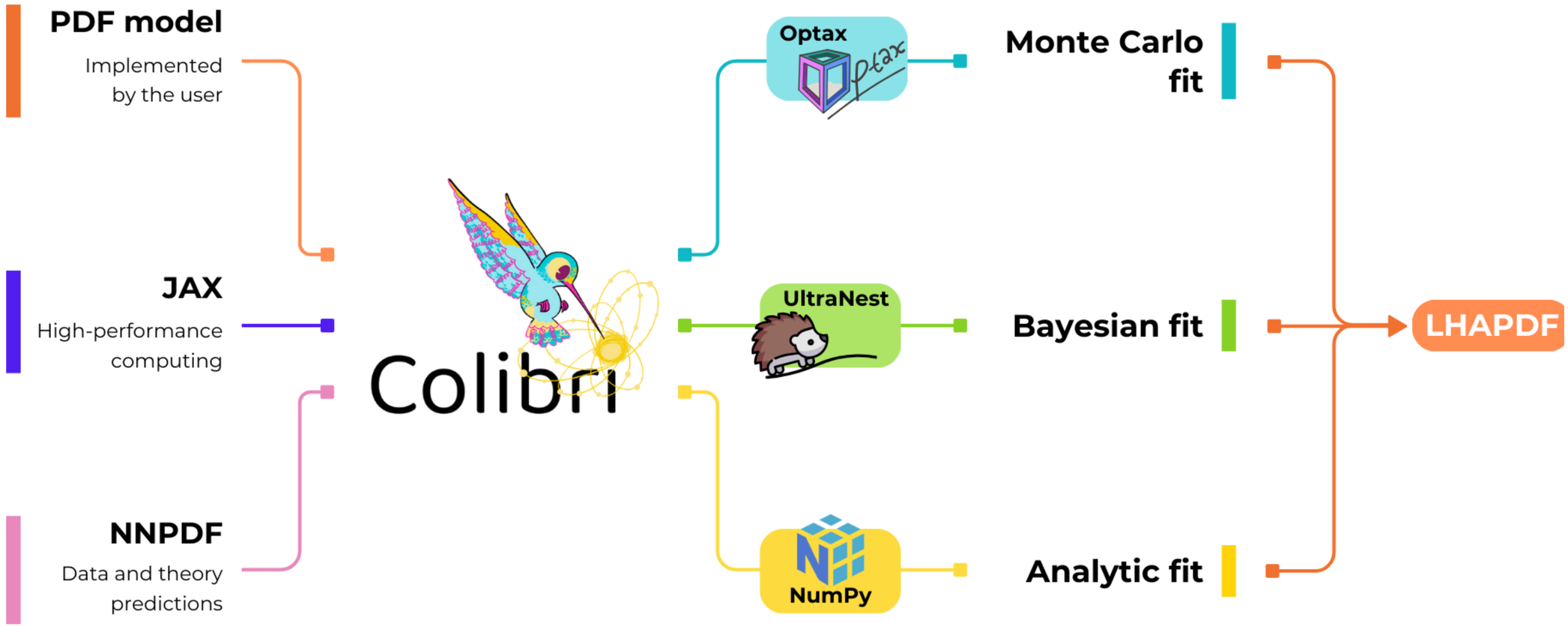
You can contact me at:
eh651@cam.ac.uk

Thank you for your attention!

Extra slides

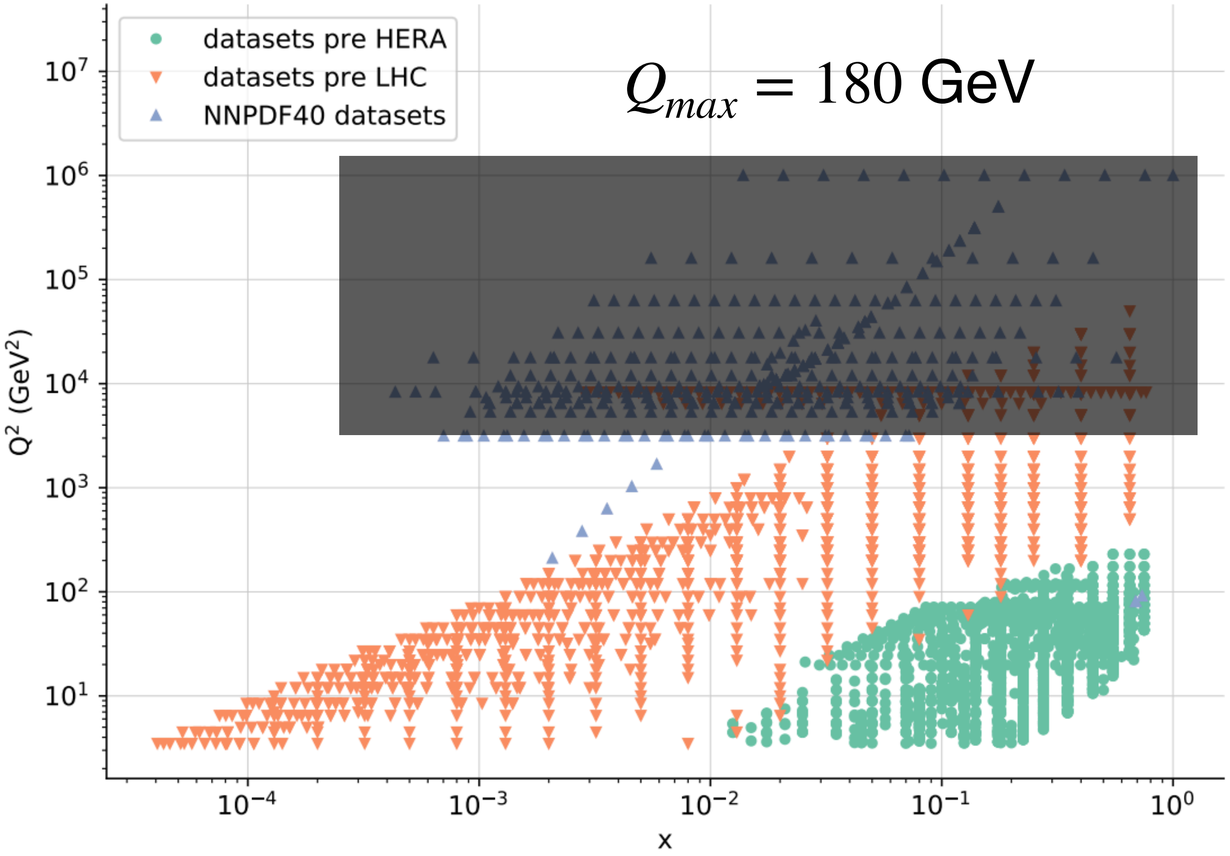
Upcoming: Colibri

Model-flexible framework for bayesian fits

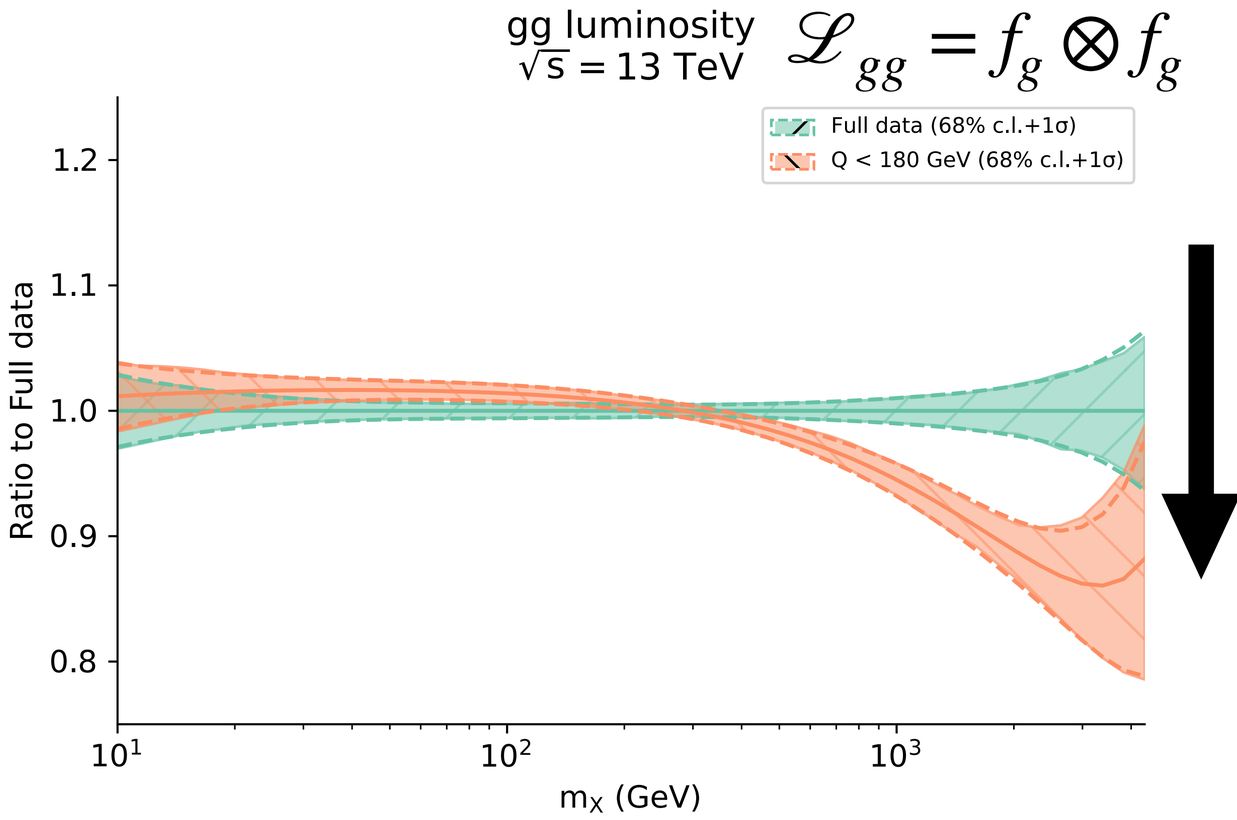


Discrepancy between low and high-energy data fits

Kinematic coverage



Impact of energy cut



Potential pitfalls of PDF fitting

- No theory constraint available (e.g. from lattice)

  ➡ Only determined from measurements

- Very flexible parametrisation by necessity

  ➡ Can mimic deviations in the observables

- Assumes SM in theory predictions

- Large-x PDF mostly constrained by high-energy measurements

  ➡ Risk to absorb BSM deviations!

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

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

New physics scenarios: W'

From UV to the SMEFT

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

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

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

Matching to the SMEFT:

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

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

New physics parameter

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

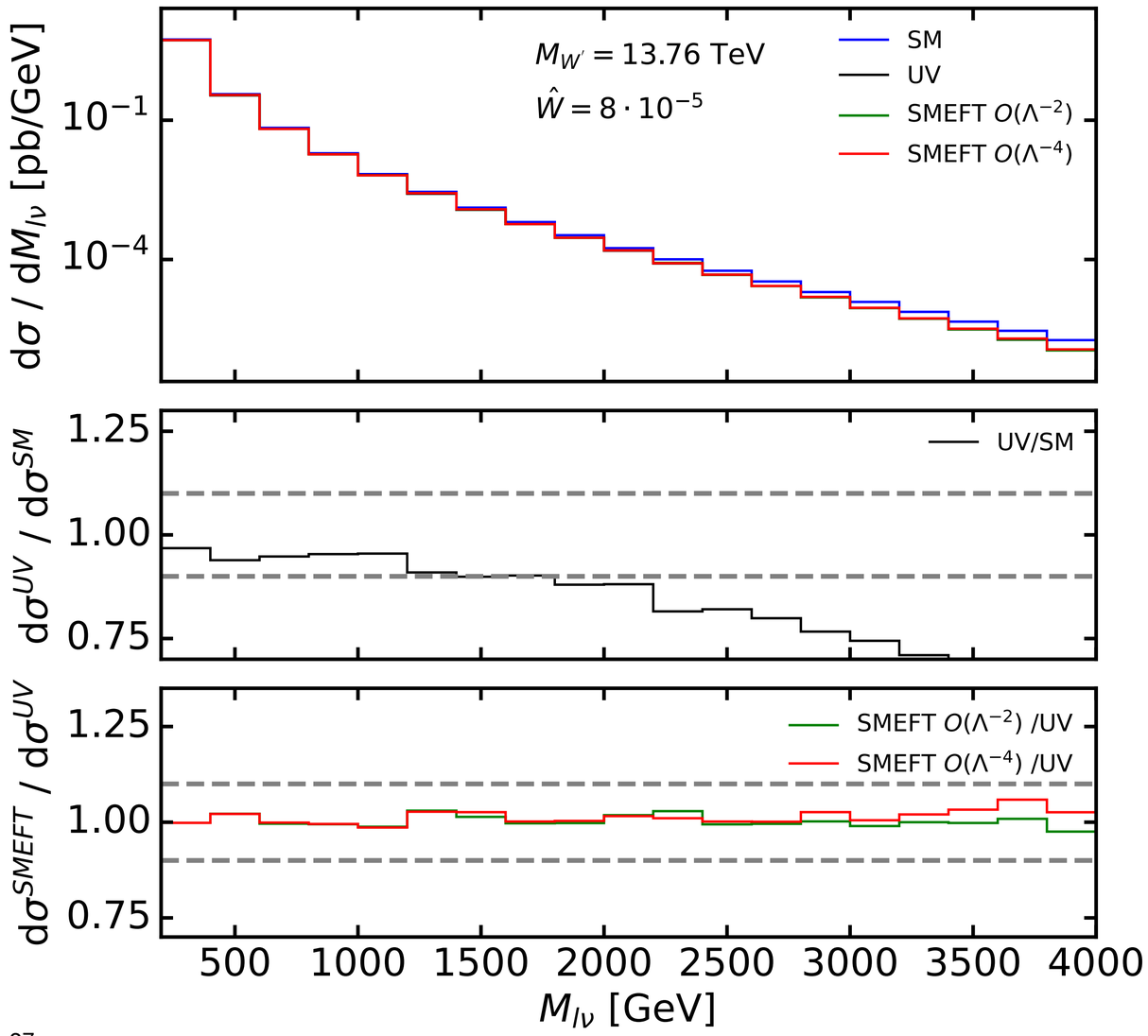
Generation of the pseudodata

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

➡ Impacts CC and NC Drell-Yan

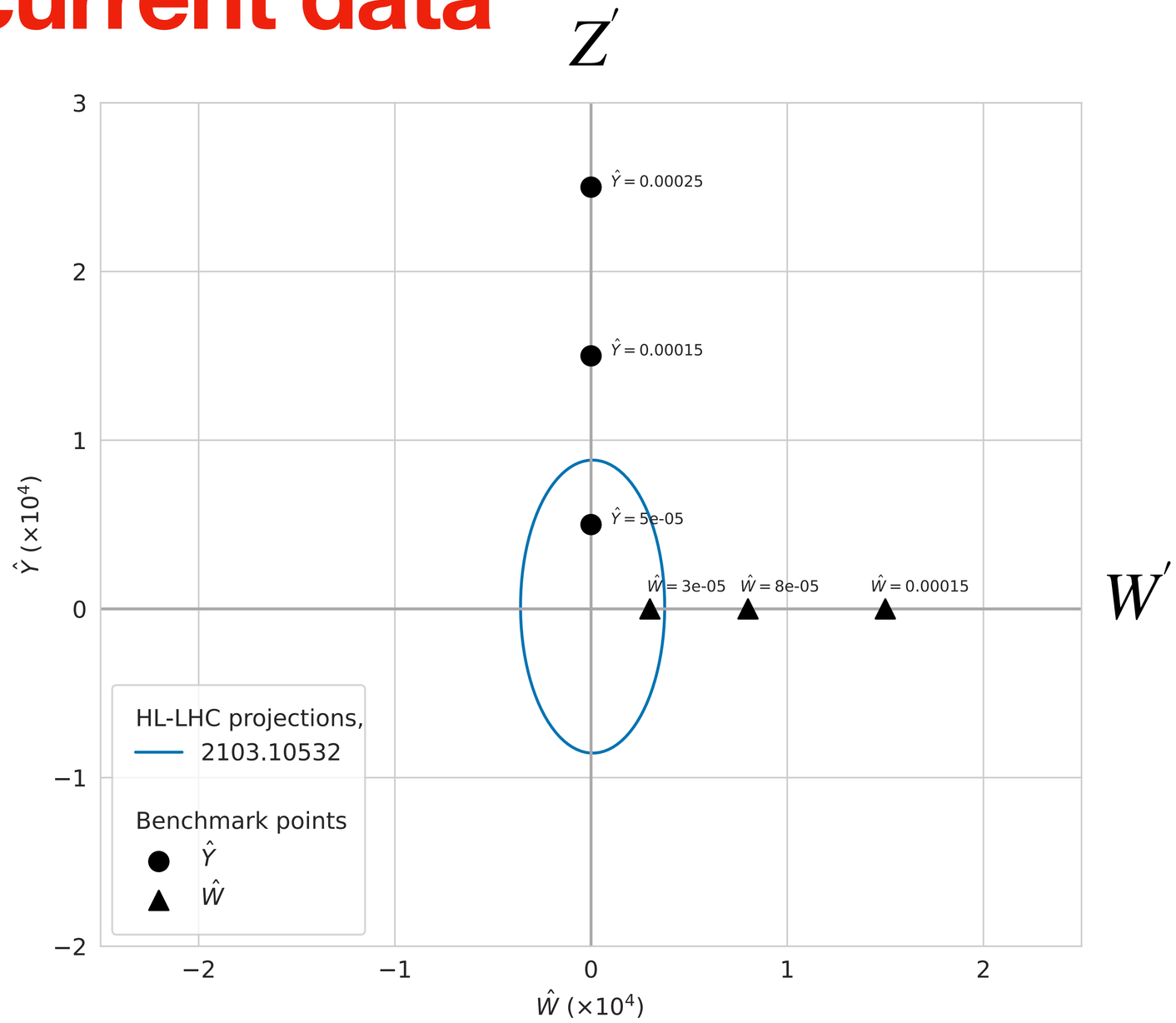
$$\hat{W} = \frac{g^2_{W'}}{g^2} \frac{m^2_W}{M^2_{W'}} \propto \frac{c}{\Lambda^2}$$

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



Constraints from current data

- New physics scenarios compared to constraints at 95% CL



Impact of contamination: fake deviations

SM predictions with:

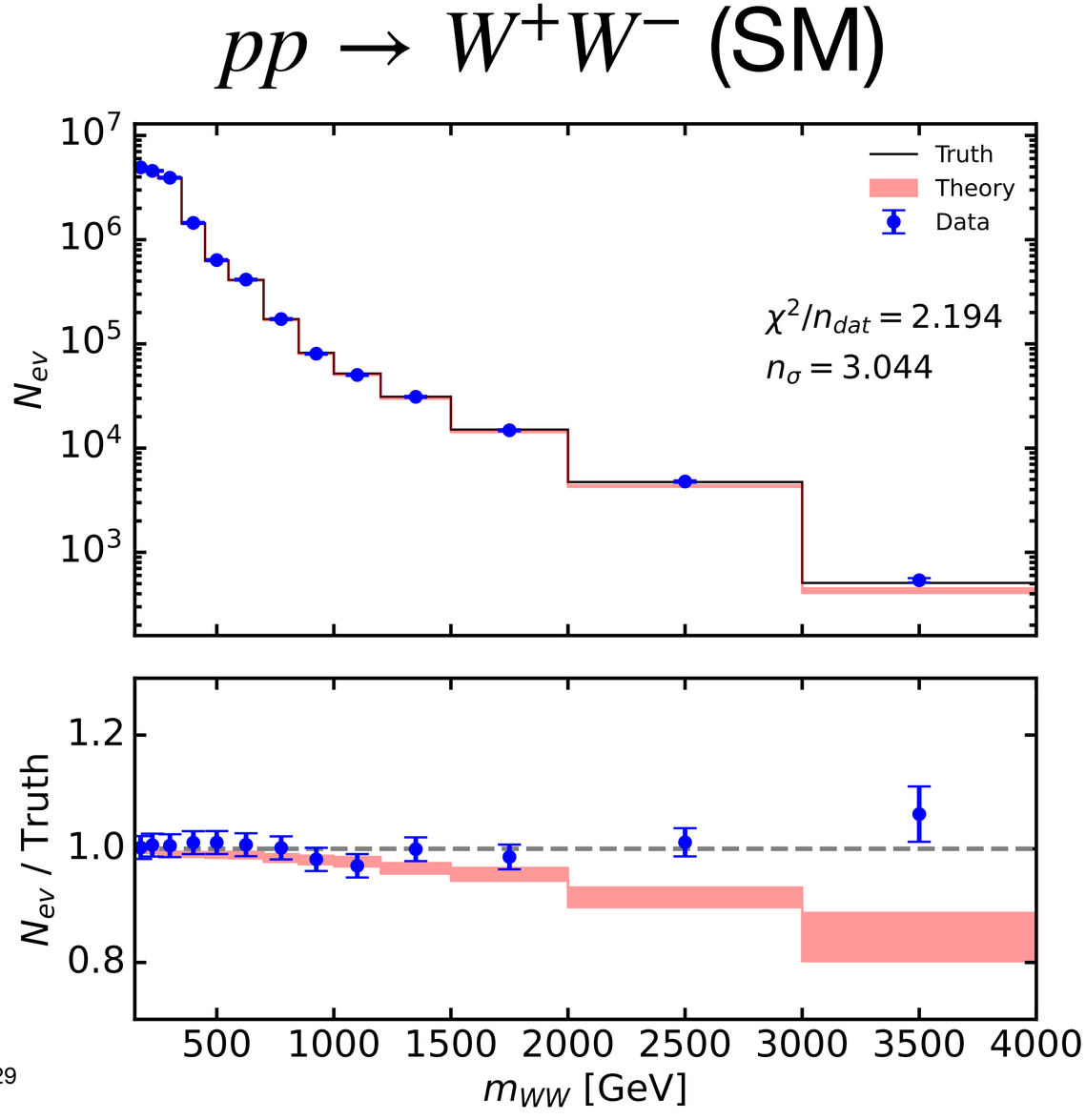
- Contaminated PDFs (red)
- True PDFs (black)

➡ Fake deviation in other sectors

Also seen in:

WH, WZ, ZH production

HL-LHC Projections



Comparing SMEFT bounds for different PDFs

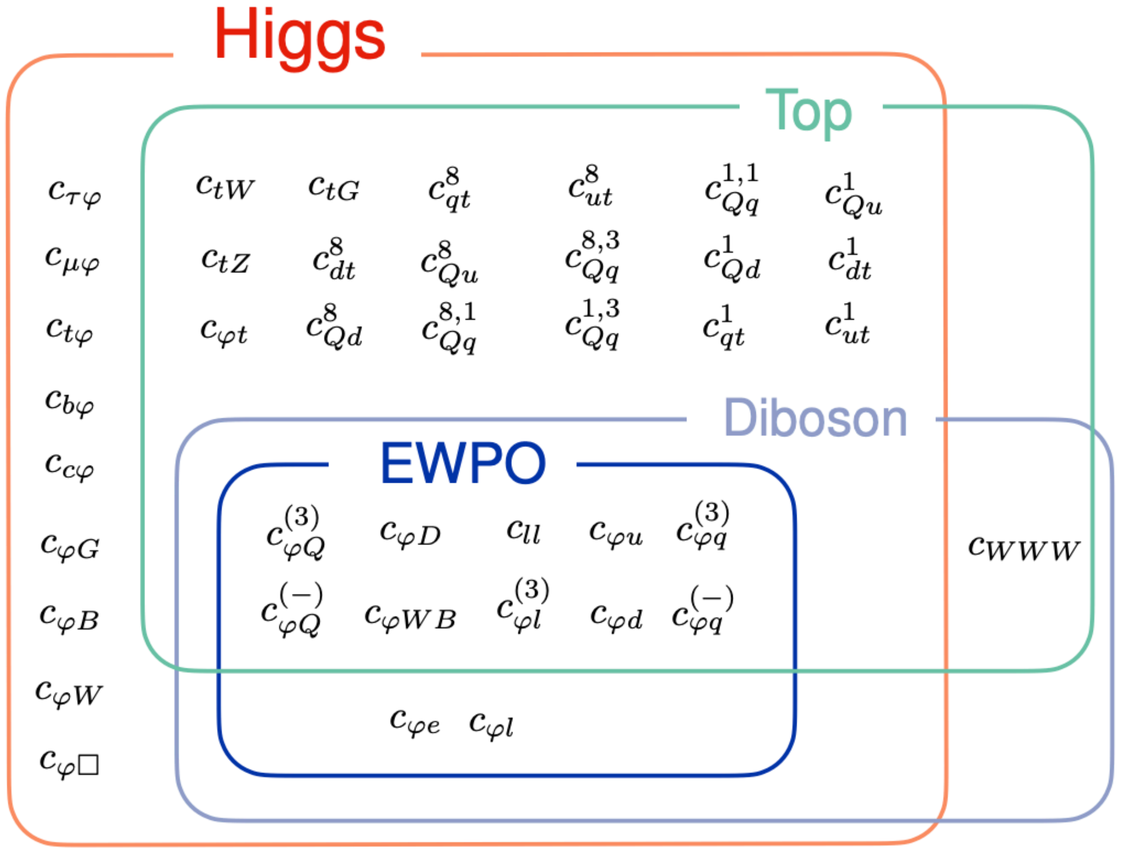
Injection of W' and Z' in Drell-Yan

Fit	$\hat{W} \times 10^5$ (UL = 0.8)		$\hat{Y} \times 10^5$ (UL = 1.5)	
	95% CI	n_σ **(SM/UL)**	**95% CI**	n_σ **(SM/UL)**
True PDF	(0.64, 0.98)	9.00 / 0.11	(0.99, 2.03)	5.50 / 0.00
BSM-biased PDF	(-0.08, 0.39)	1.34 / 5.51	(0.11, 1.44)	2.26 / 2.16
Cons. PDF	(0.50, 1.1)	5.29 / 0.03	(0.68, 2.22)	3.73 / 0.18
Simu. fit	(0.44, 1.05)	4.91 / 0.28	(0.79, 2.09)	4.39 / 0.15

Simultaneous fit of PDF and new physics

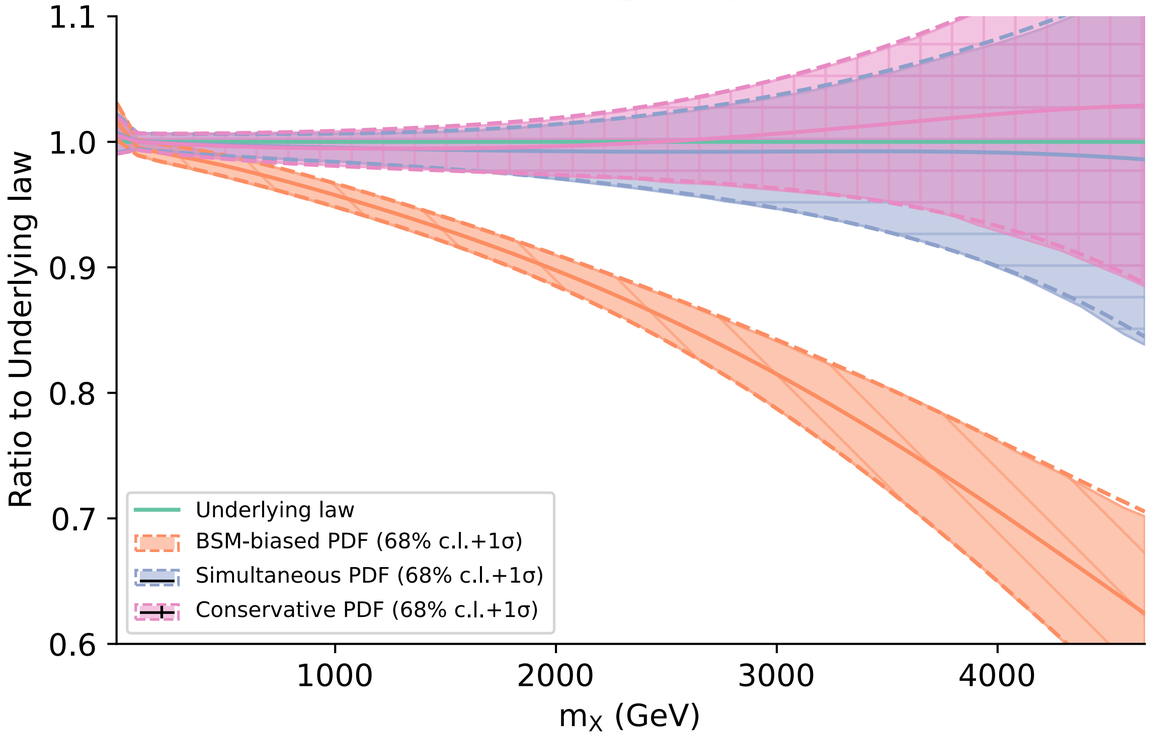
SMEFT operators implemented

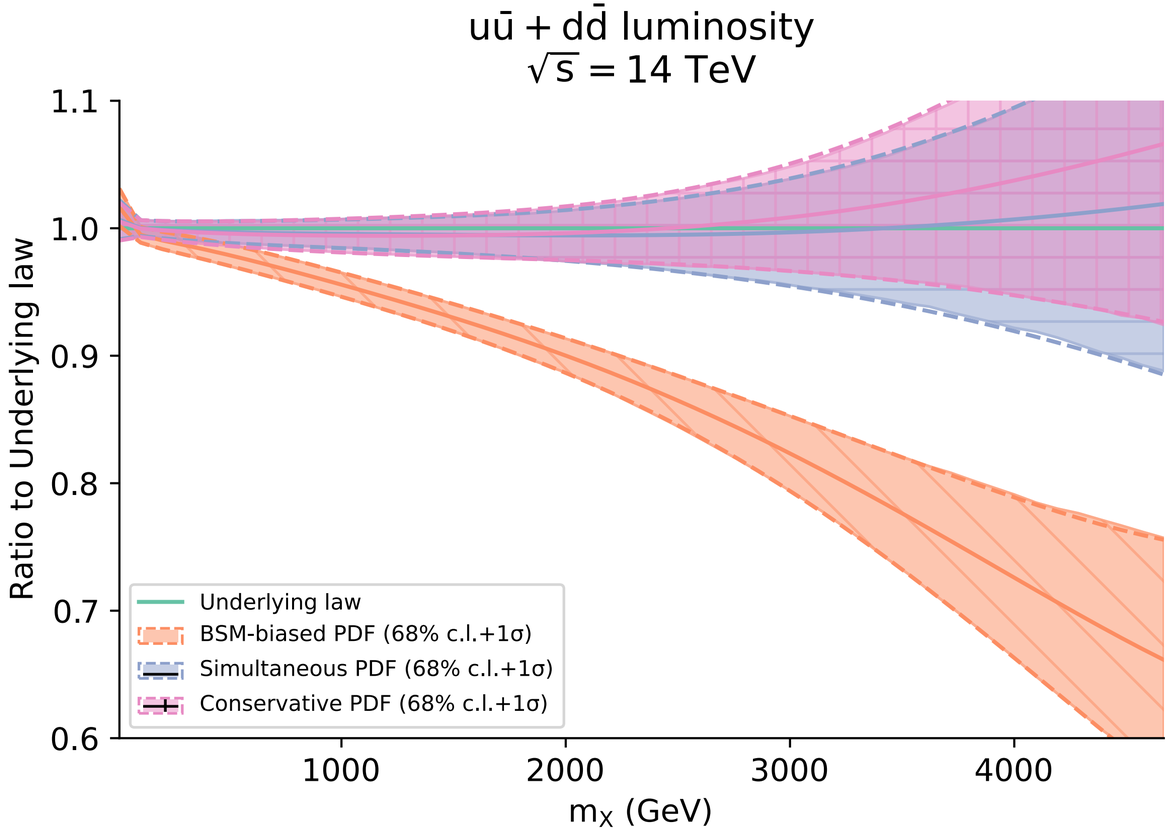
- 40 operators implemented
- Observables:
  - top sector
  - diboson
  - Higgs
  - Drell-Yan
  - EW Precision Observables



PDFs for new physics searches in HMDY

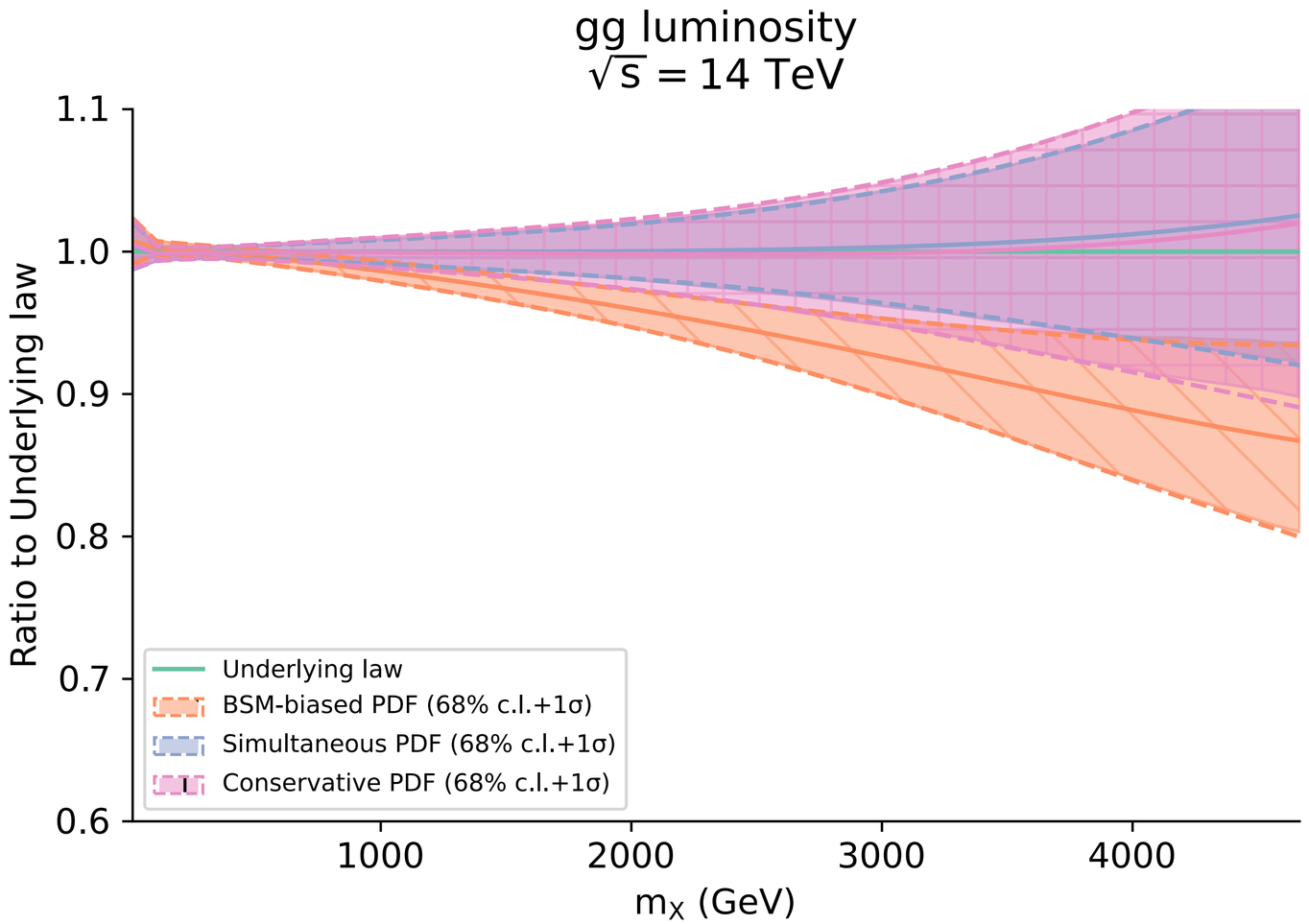
PDF luminosities



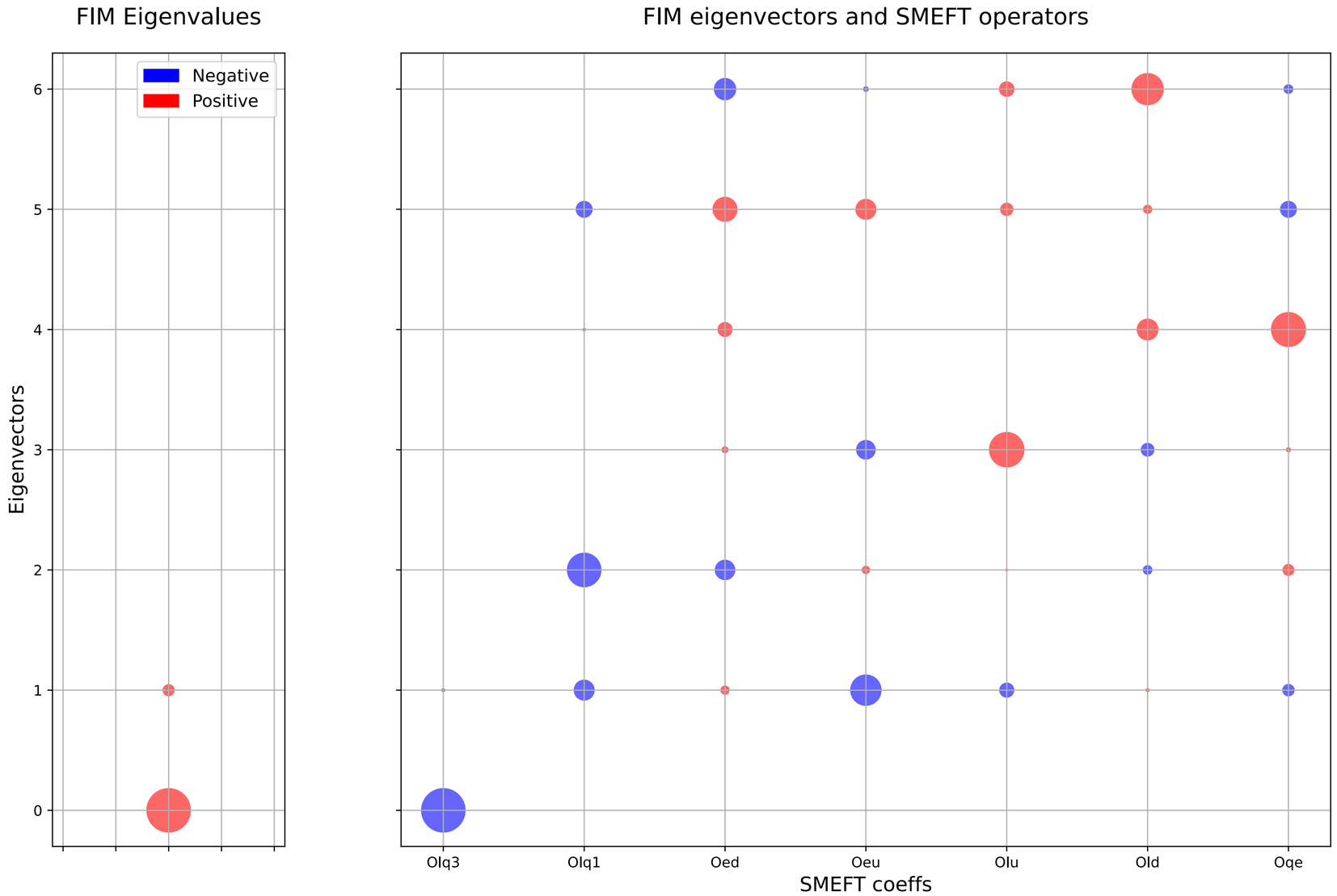


PDFs for new physics searches in $t\bar{t}$

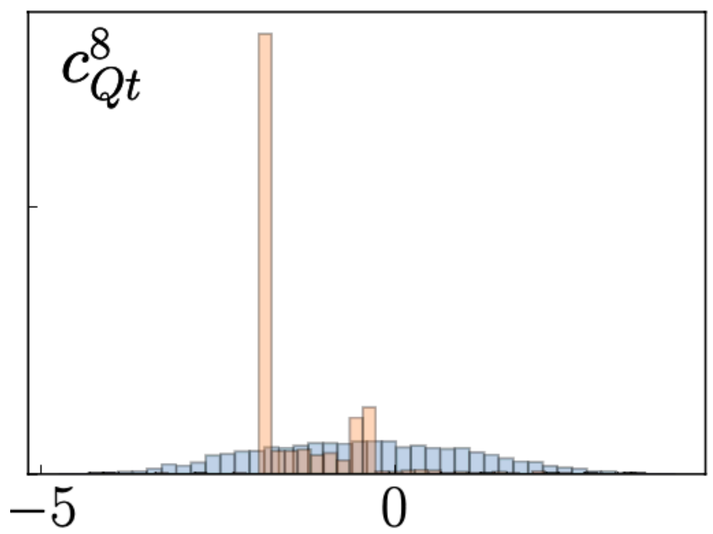
Same exercise with heavy gluon in ttbar at HL-LHC

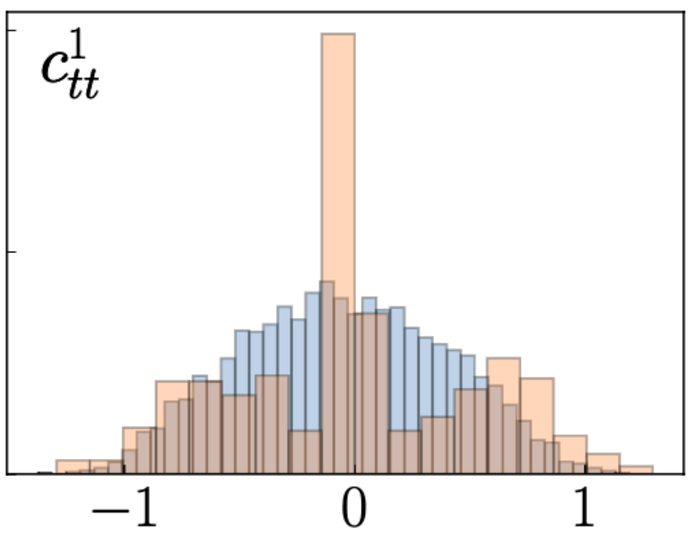


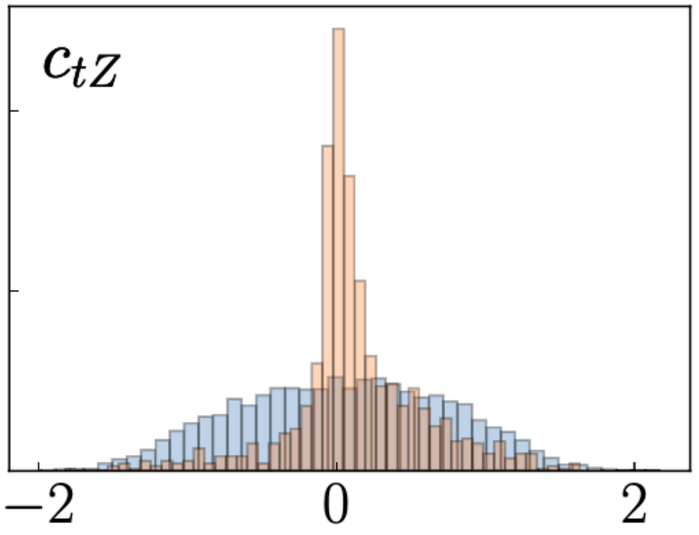
SMEFT PCA results



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, Ubiali, in progress]

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

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