SIMUnet: an open-source tool for the simultaneous fit of PDFs and SMEFT coefficients

Mark N. Costantini

High Precision for Hard Processes
Torino, 2024

In collaboration with:
PBSP collaboration: Maria Ubiali, Elie Hammou, Zahari Kassabov, Maeve Madigan, Luca Mantani, Manuel Morales Alvarado, James Moore

PBSP

UNIVERSITY OF CAMBRIDGE

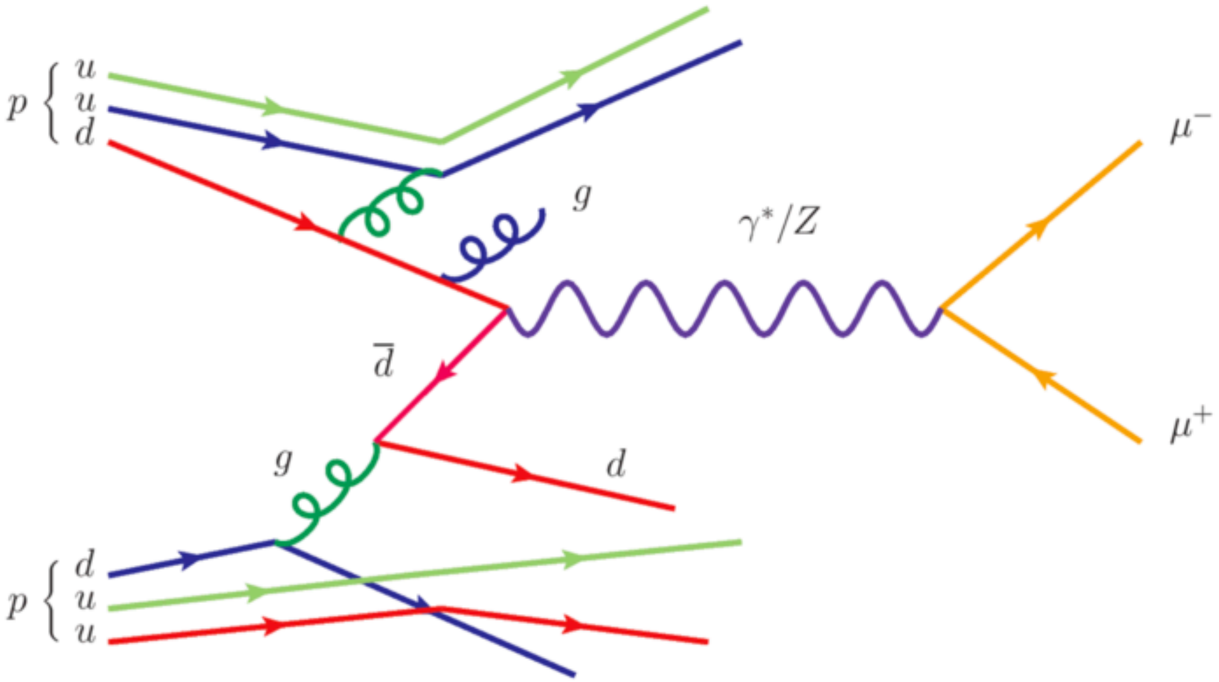
erc European Research Council

Outline

- Introduction: PDF and SMEFT global fits
- SIMUnet: methodology overview
- SIMUnet: Applications
- Conclusions / Outlook

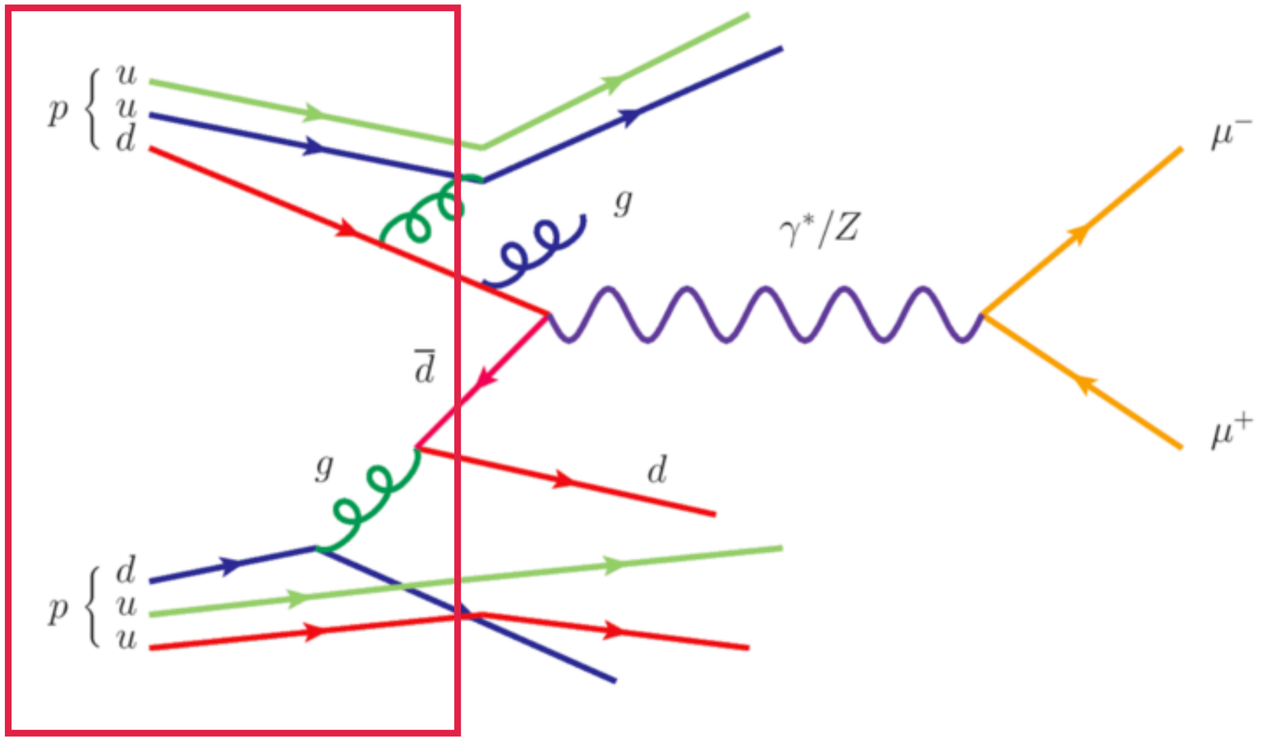
Introduction

Theoretical predictions at the LHC



- **Theory predictions** for collider experiments are obtained from the standard **factorisation formula**

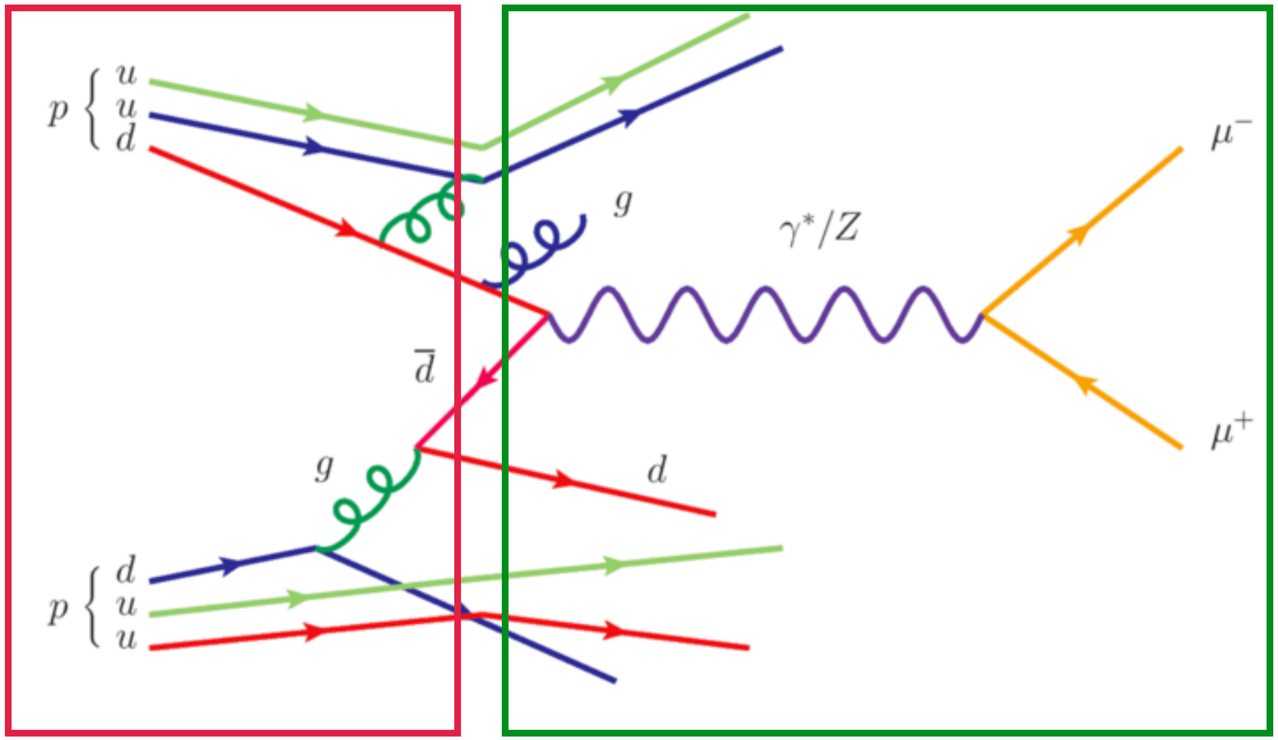
Theoretical predictions at the LHC



- **Theory predictions** for collider experiments are obtained from the standard **factorisation formula**

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

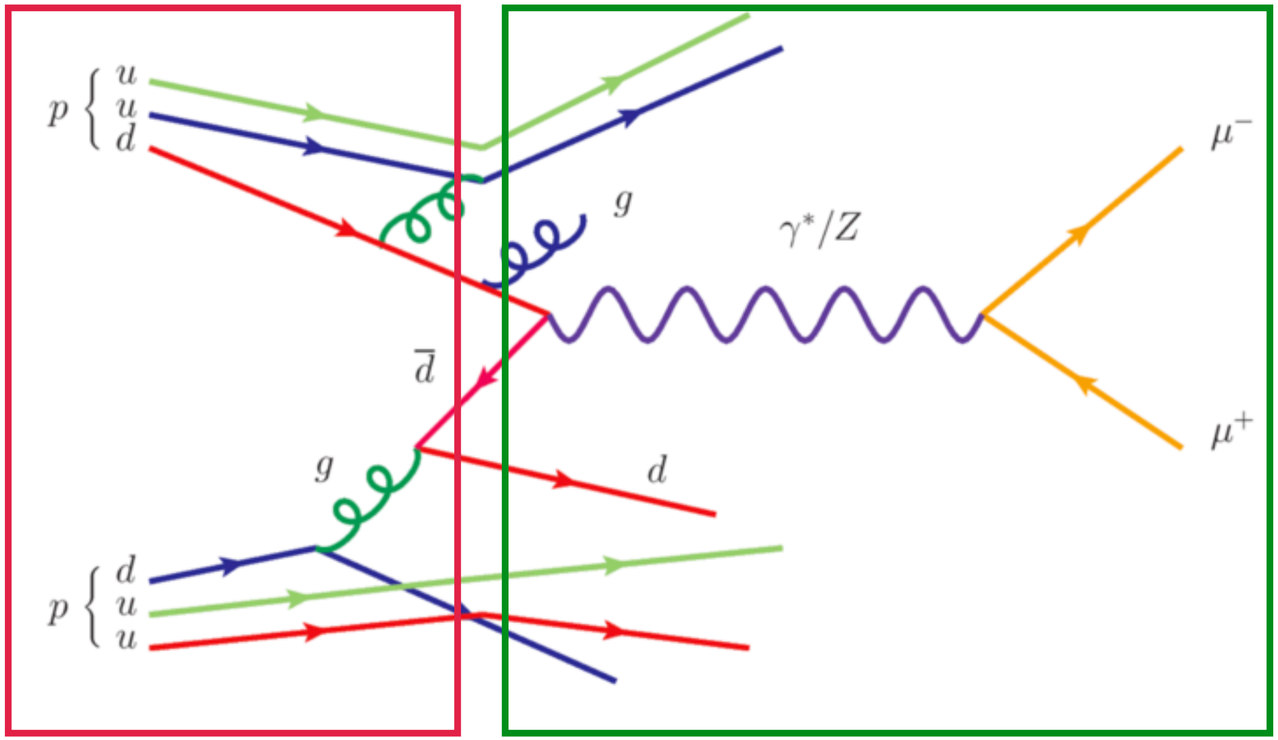
Theoretical predictions at the LHC



- **Theory predictions** for collider experiments are obtained from the standard **factorisation formula**

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

Theoretical predictions at the LHC



- **Theory predictions** for collider experiments are obtained from the standard **factorisation formula**

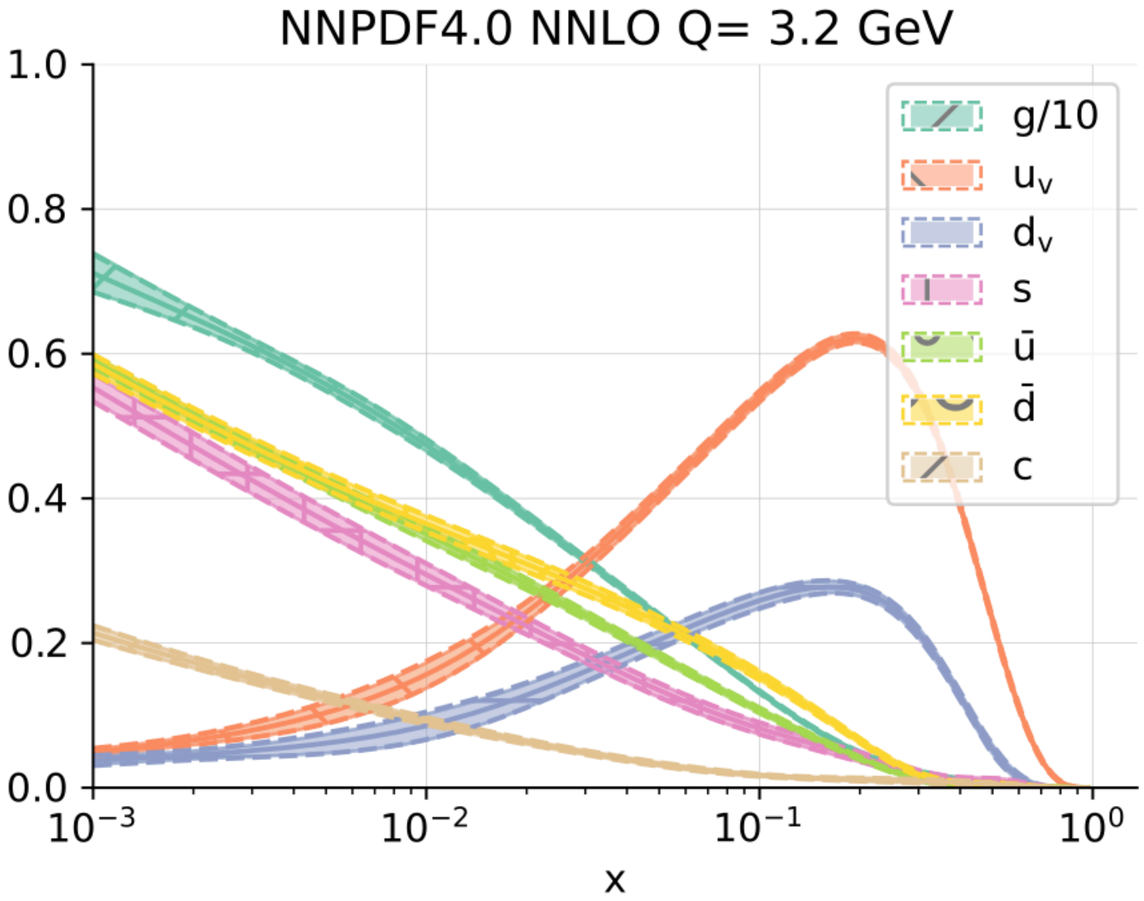
$$T(\{\theta\},\{c\}) = \text{PDFs}(\{\theta\}) \otimes \hat{\sigma}(\{c\})$$

Theory predictions are functions of :

1. Physics parameters c: e.g. α_S, m_W, **Wilson Coefficients** if we use the **SMEFT**
2. PDF parameters θ e.g. the weights of the neural network parametrising the PDF at the initial scale

PDF Parameter Fits

The common approach



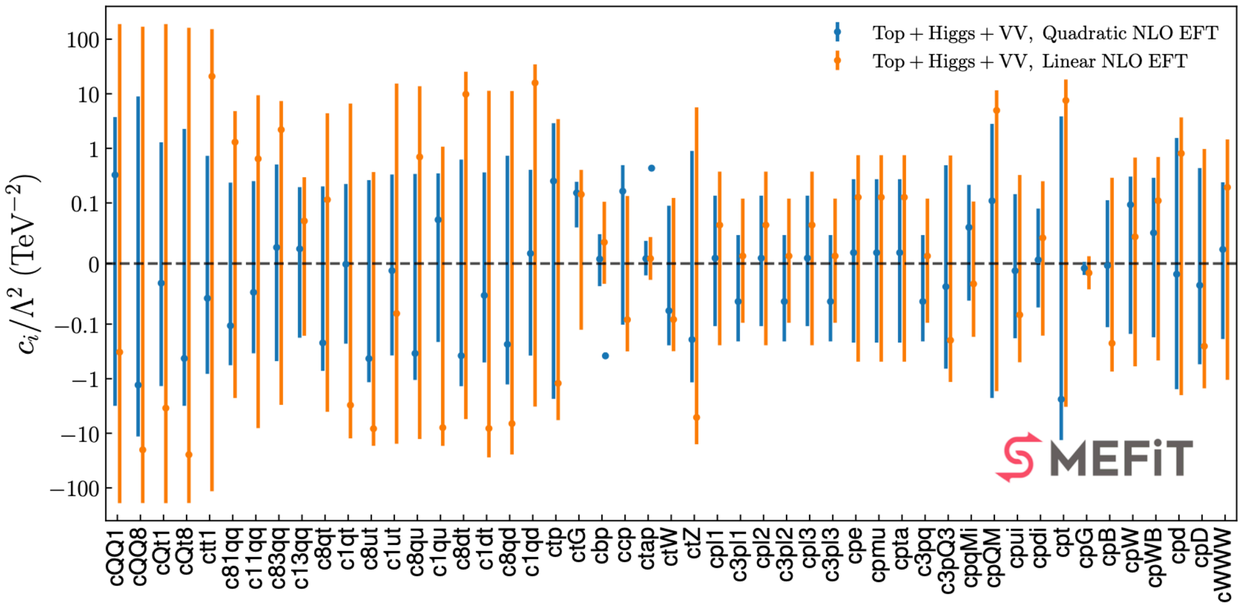
- Fix physical parameters $c = \bar{c}$
- Extract PDFs from data, with the implicit dependence $\theta = \theta(\bar{c})$

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

NNPDF40, **Ball et al, arXiv:**2109.02653

SMEFT Fits

The common approach



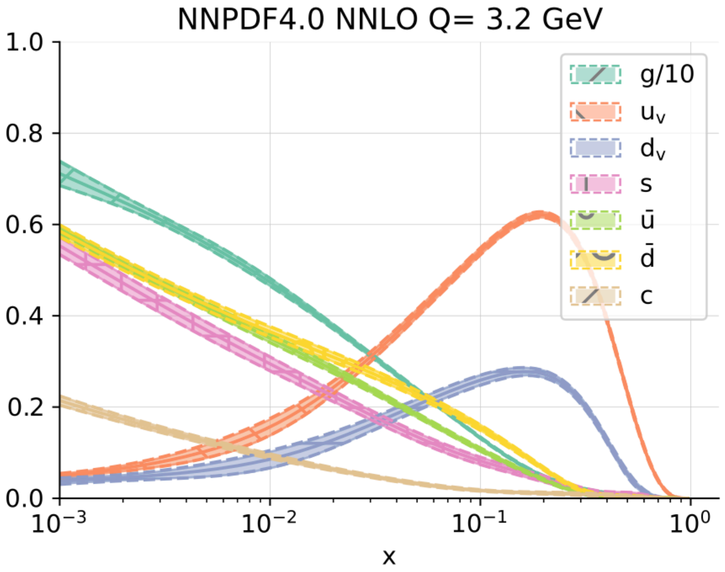
- Fix the PDFs parameters $\theta = \bar{\theta}$
- Fit physical parameters from the data with the implicit dependence $c = c(\bar{\theta})$

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

SMEFiT, Ethier et al, arXiv:2105.00006

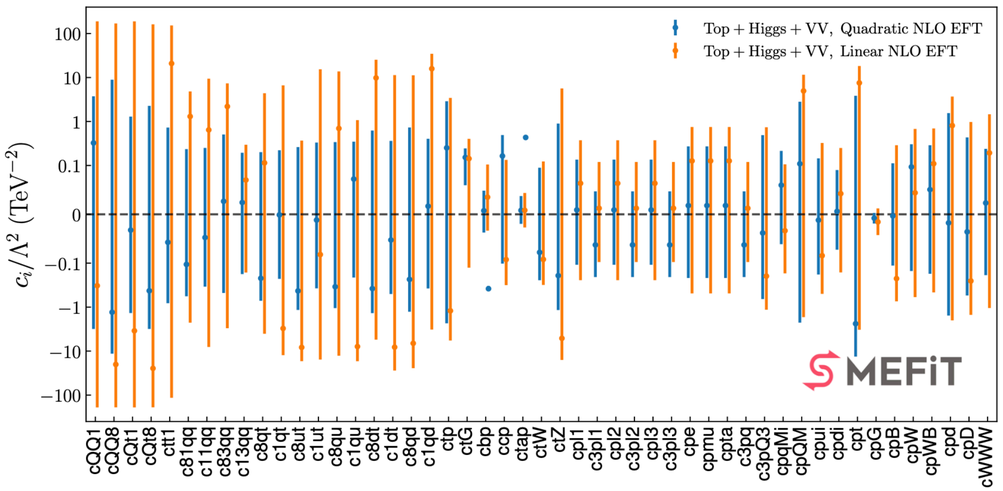
Simultaneous PDF and SMEFT Fits

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



PBSP

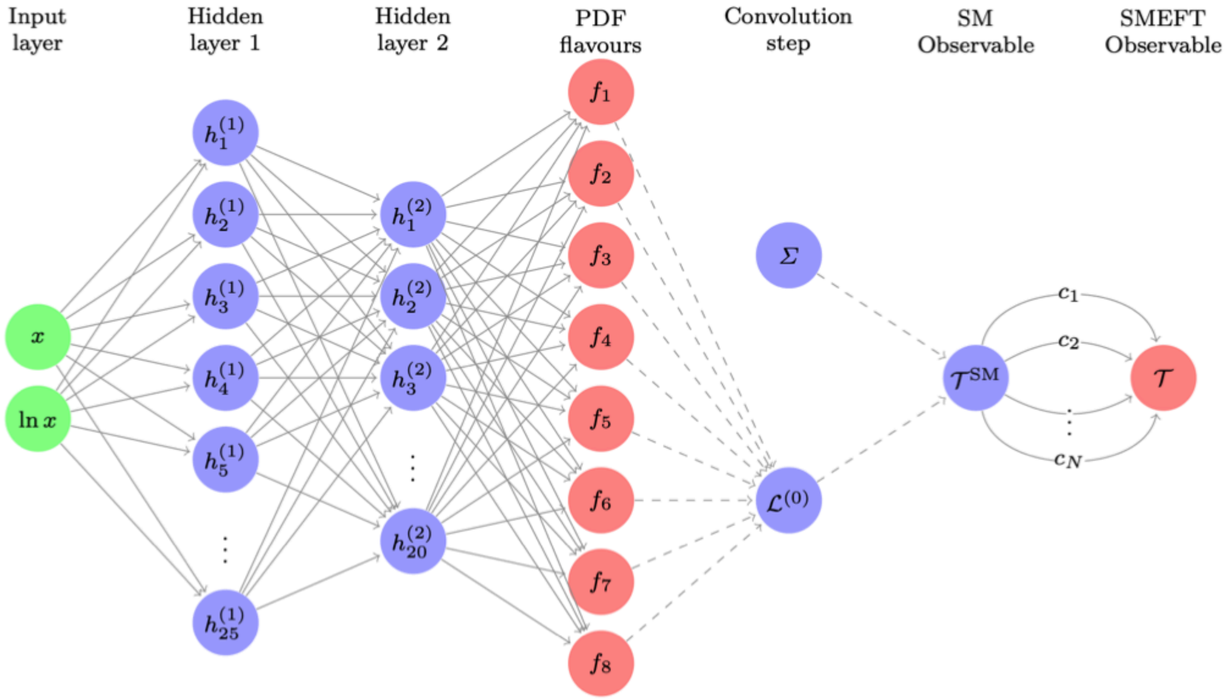
MNC, E. Hammou, L. Mantani, J, Moore, M.Morales Alvarado, M. Ubiali



SIMUnet: Methodology

SIMUnet: Methodology

https://github.com/HEP-PBSP/SIMUnet



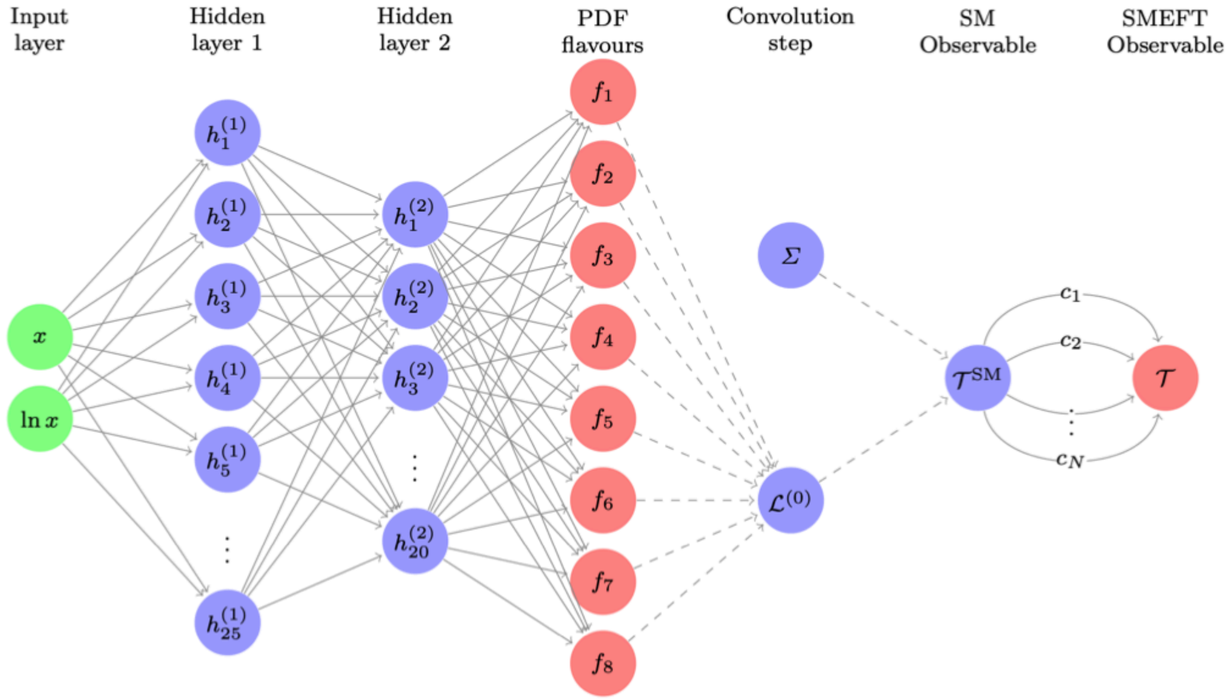
- Simultaneous fit of PDF and theory parameters (e.g. EFT WC)

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

S. Iranipour and M. Ubiali, arXiv:2201.07240
PBSP Collaboration, Costantini et al, arXiv:2404.10056

SIMUnet: Methodology

https://github.com/HEP-PBSP/SIMUnet



- Simultaneous fit of PDF and theory parameters (e.g. EFT WC)

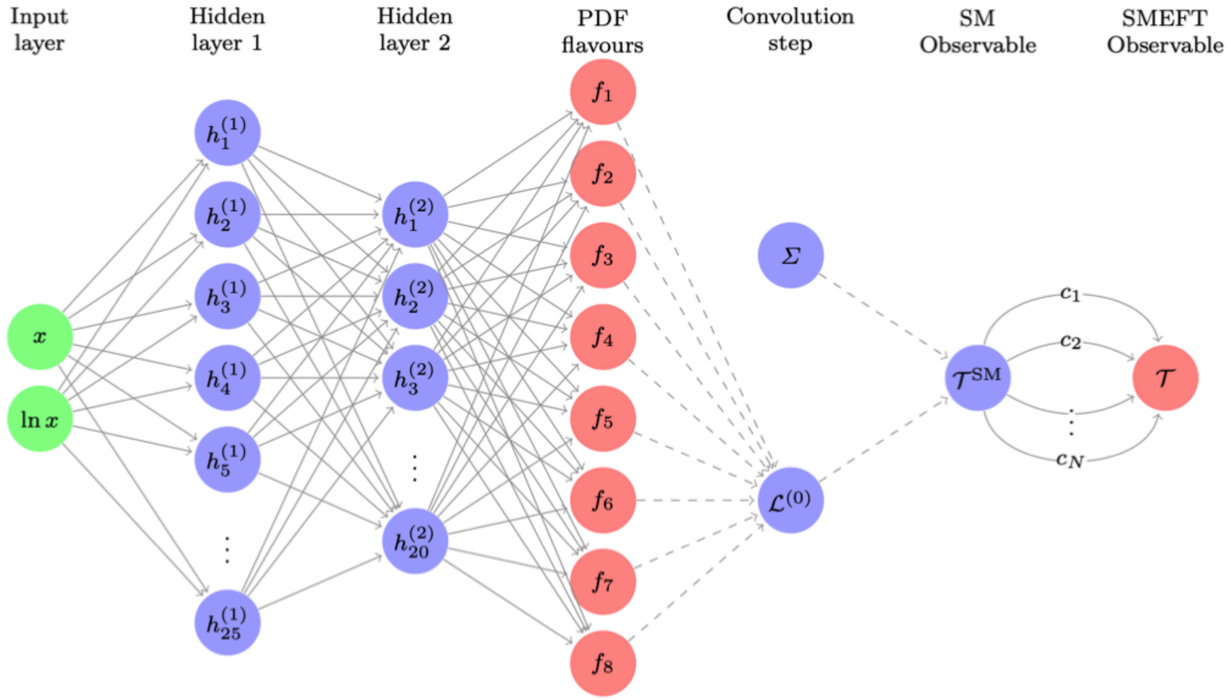
$$T(\{\theta\},\{c\}) = \text{PDFs}(\{\theta\}) \otimes \hat{\sigma}(\{c\})$$

- Extension of NNPDF40 framework to account for theory parameters dependence of partonic cross sections

S. Iranipour and M. Ubiali, arXiv:2201.07240
PBSP Collaboration, Costantini et al, arXiv:2404.10056

SIMUnet: Methodology

https://github.com/HEP-PBSP/SIMUnet



- Simultaneous fit of PDF and theory parameters (e.g. EFT WC)

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

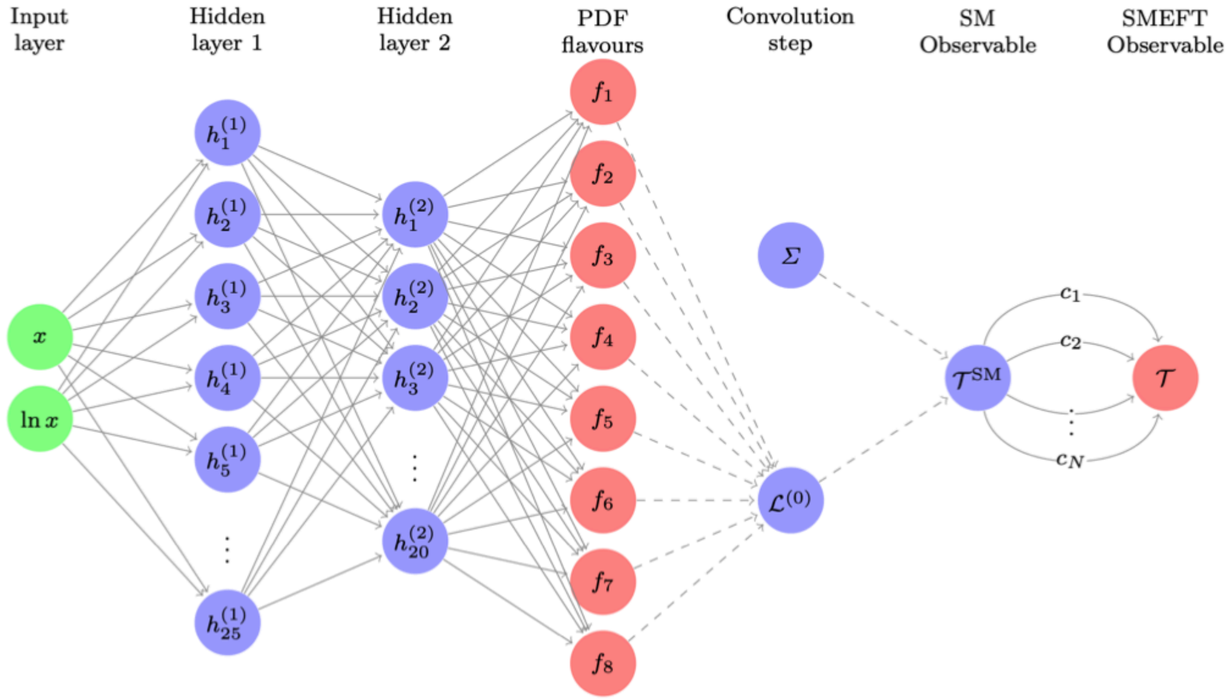
- Extension of NNPDF40 framework to account for theory parameters dependence of partonic cross sections

→ Extra combination layer to the NNPDF40 Neural Network

S. Iranipour and M. Ubiali, arXiv:2201.07240
PBSP Collaboration, Costantini et al, arXiv:2404.10056

SIMUnet: Methodology

https://github.com/HEP-PBSP/SIMUnet



- Simultaneous fit of PDF and theory parameters (e.g. EFT WC)

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

- Extension of NNPDF40 framework to account for theory parameters dependence of partonic cross sections

→ Extra combination layer to the NNPDF40 Neural Network

→ SMEFT corrections added in the form of K-factors

S. Iranipour and M. Ubiali, arXiv:2201.07240
PBSP Collaboration, Costantini et al, arXiv:2404.10056

SIMUnet: Uncertainty Quantification

- Monte Carlo replica method:

→ Generate N_{rep} data replica $d_i \sim \mathcal{N}(d, C_{\text{exp}}), \quad i \in 1, \ldots, N_{\text{rep}}$

→ Fit each replica separately so as to get a distribution in the space of PDFs and WCs

SIMUnet: Uncertainty Quantification

- Monte Carlo replica method:

  → Generate N_{rep} data replica $d_i \sim \mathcal{N}(d, C_{\text{exp}}), \;\; i \in 1, \ldots, N_{\text{rep}}$

  → Fit each replica separately so as to get a distribution in the space of PDFs and WCs

- Limitations: Linear fits only

  → Uncertainty quantification is reliable on "Linear fits" only

  → The open-source SIMUnet only allows for fits that are linear in the EFT / Theory parameter expansion

MNC, M. Madigan, L. Mantani, J. Moore, arXiv:2404.10056

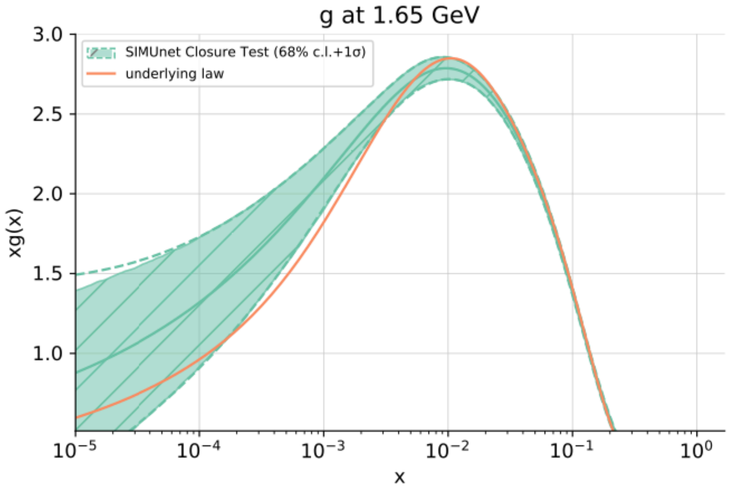
Simultaneous Closure Tests

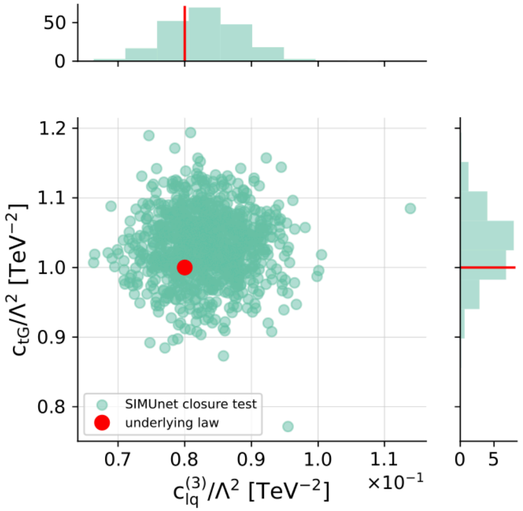
In a closure test we fit the model to fake-data that has been generated by adding some Gaussian noise to the theory predictions

$$d_{ct} = T(\theta^*, c^*) + \eta, \;\; \eta \sim \mathcal{N}(0, C_{\text{exp}})$$

→ Allows for the simultaneous fit of PDFs and Wilson coefficients within a closure test

→ Validation of the SIMUnet methodology in its ability to produce an underlying law comprising both PDFs and Wilson Coefficients





Example:

"True" PDF: `NNPDF40_nnlo_as_01180`

"True" WC : $c_{tG} = 1, \;\; c_{lq}^{(3)} = 0.08$

PBSP Collaboration, Costantini et al, arXiv:2404.10056

SIMUnet: A tool for simultaneous fits

SIMUnet released open-source with detailed documentation

https://github.com/HEP-PBSP/SIMUnet

→ PDF only fits

→ SMEFT only fits, Linear SMEFT

→ Simultaneous SMEFT and PDF fits, Linear SMEFT

```
dataset_inputs:
  - {dataset: CMSDY1D12, cfac: ['QCD', 'EWK']}
  - {dataset: ATLASTTBARTOT7TEV, cfac: [QCD], simu_fac: "EFT_NLO"}
  - {dataset: ATLAS_WHEL_13TEV, simu_fac: "EFT_NLO", use_fixed_predictions: True}
```

→ SIMUnet allows for generation of pseudo data containing NP, can be used to test robustness of SM like PDF fits in particular NP scenarios

```
dataset_inputs:
  - {dataset: CMSDY1D12, frac: 0.75, cfac: ['QCD', 'EWK'], contamination: 'EFT_LO'}
  - {dataset: CMS_HMDY_13TEV, frac: 0.75, cfac: ['QCD', 'EWK'], contamination: 'EFT_LO'}
```

PBSP Collaboration, Costantini et al, arXiv:2404.10056

SIMUnet:Applications

NP contaminated fits

Can new physics (NP) be absorbed in the PDFs?

SIMUnet allows for generation of pseudo-data containing NP

$$T = T(\{\theta\}, \{c, c_{NP}\}),$$
$$\text{Data} = T + \text{Gaussian noise}$$

Then, perform a fit assuming $c_{NP} = 0$ using the NNPDF methodology

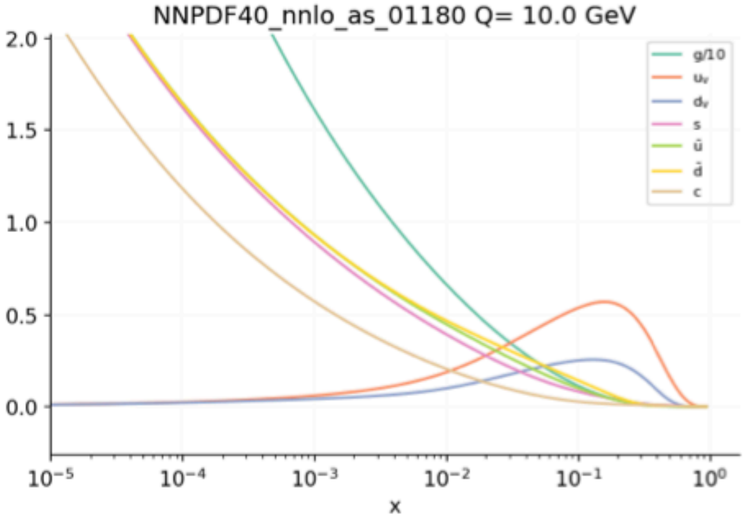
Assess whether we can mimic the modified interactions with the “wrong" PDFs!

PBSP Collaboration, Hammou et al, arXiv:2307.10370

NP contaminated fits

A case study: heavy W’

Suppose the underlying law of nature were



“Real” proton structure

$$\otimes$$

$$J_L^{a,\mu} = \sum_{f_L} \bar{f}_L T^a \gamma^\mu f_L$$

$$\mathcal{L}^{W'}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} - \frac{g^2 \hat{W}}{2m_W^2} J_L^\mu J_{L,\mu}$$

$$\hat{\sigma} = \hat{\sigma}_{SM} + \hat{\sigma}_{NP}$$

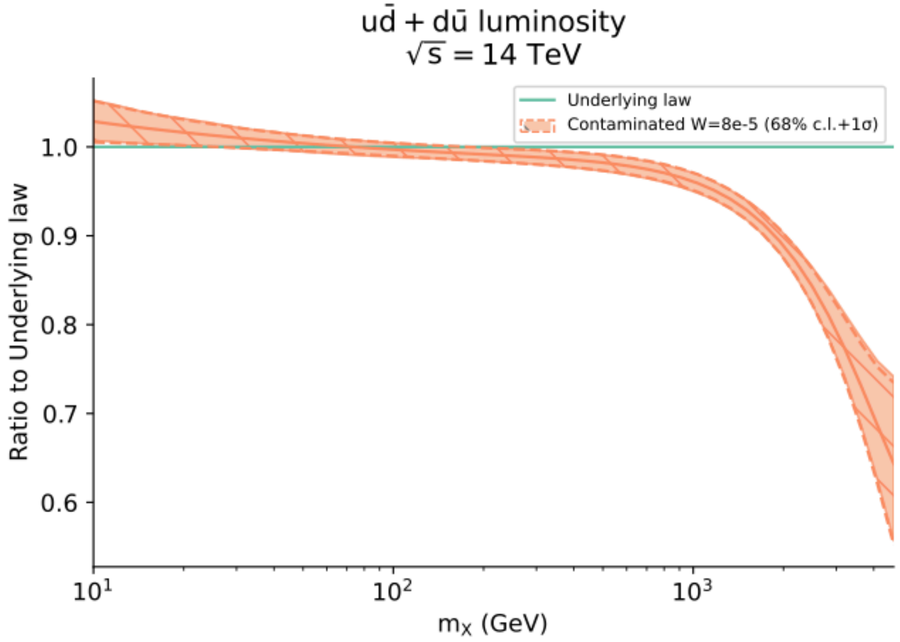
“Real” partonic cross-section

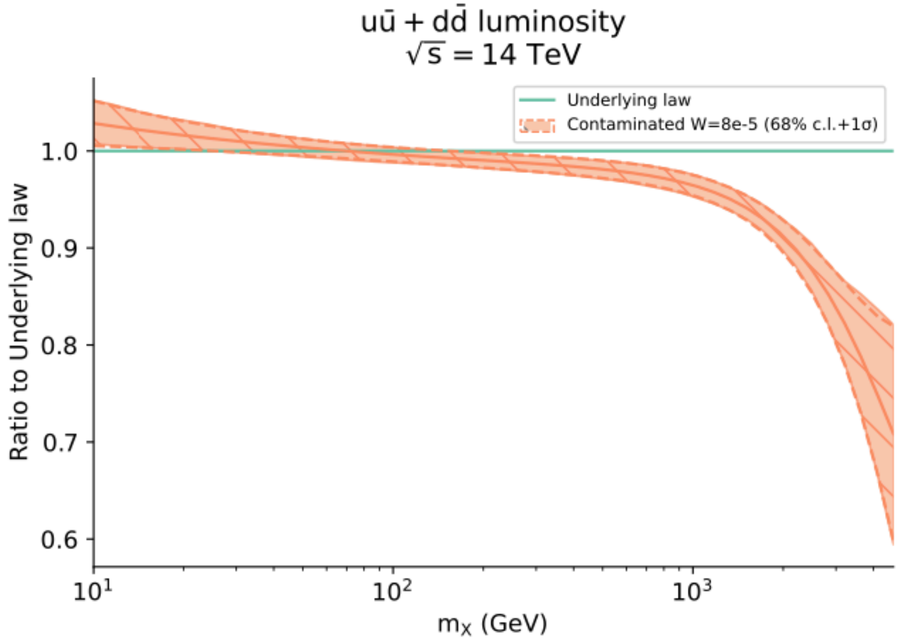
PBSP Collaboration, Hammou et al, arXiv:2307.10370

NP contaminated fits

Contaminated PDFs

Contaminated PDFs





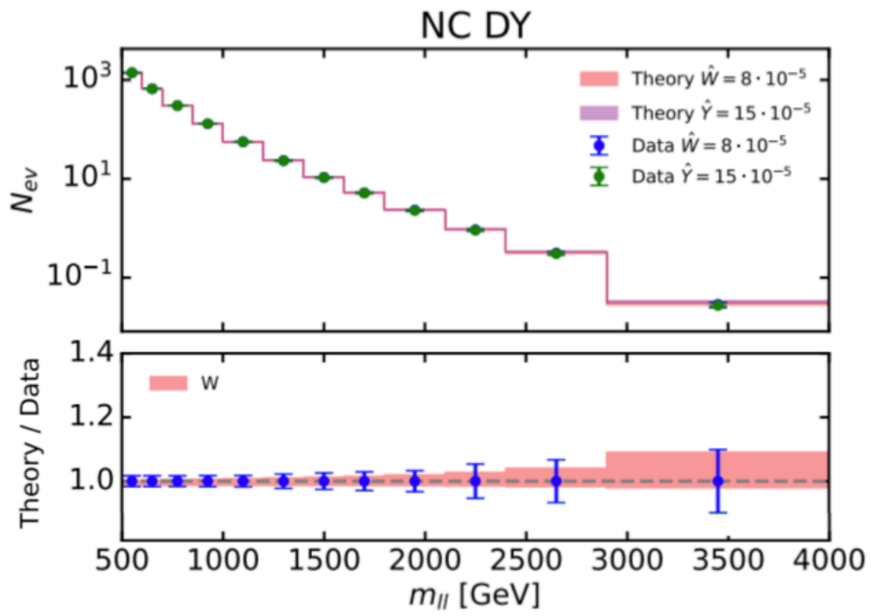
Huge shift and yet we find a good fit to the data!

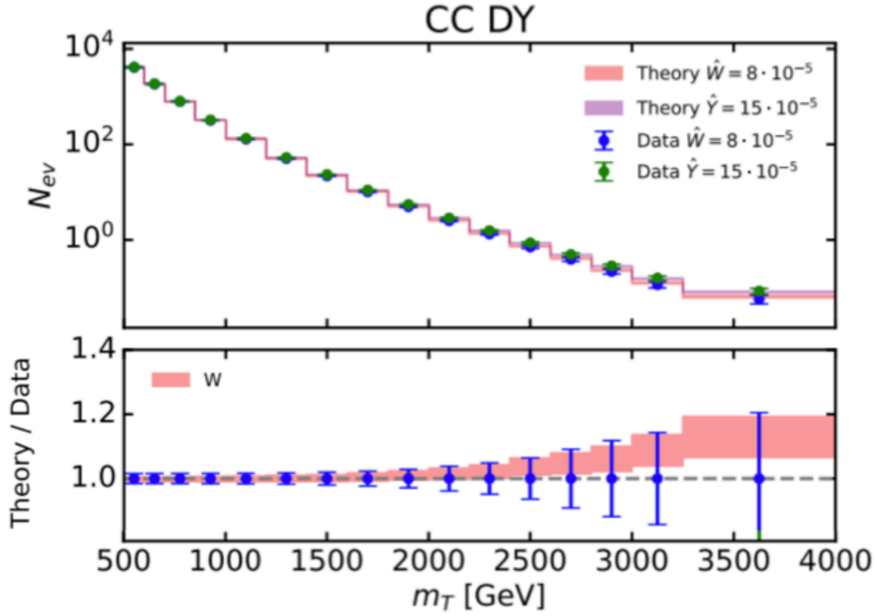
Large-x behaviour in PDFs is not constrained: especially anti-quark PDFs allow for NP absorption

PBSP Collaboration, Hammou et al, arXiv:2307.10370

NP contaminated fits

Data - Theory comparison





Data $= f_{true} \otimes \sigma_{NP}$

Theory $= f_{fit} \otimes \sigma_{SM}$

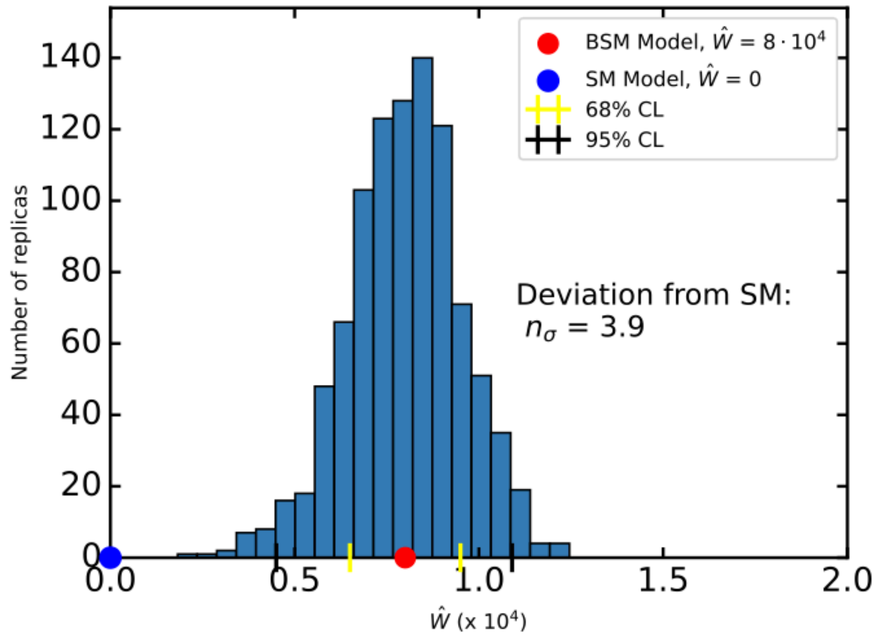
Data and Theory are compatible → PDF shift compensates for the NP effect!

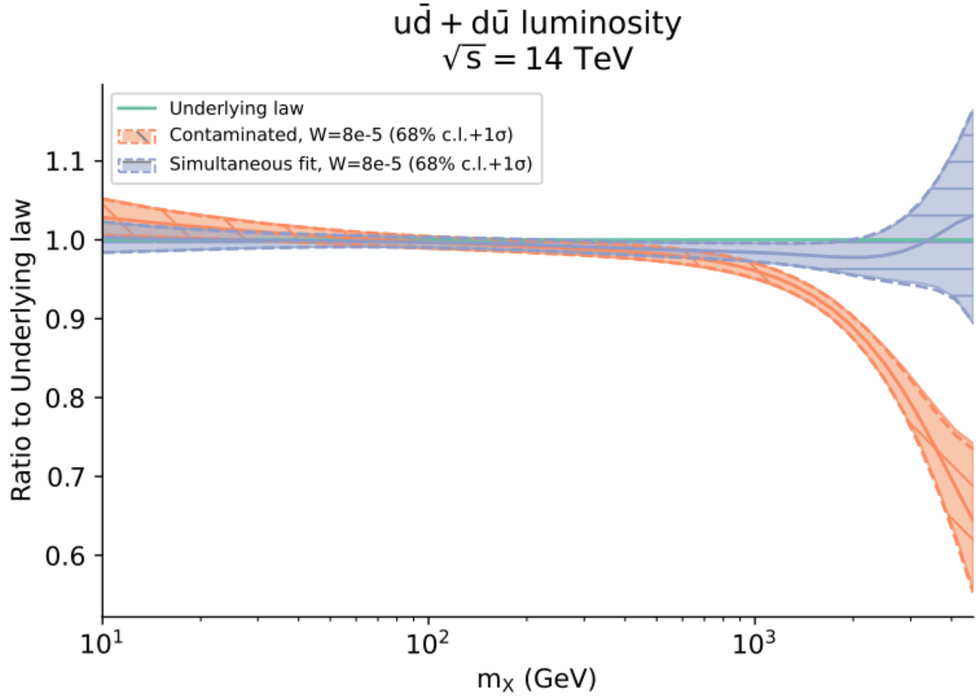
PBSP Collaboration, Hammou et al, arXiv:2307.10370

NP contaminated fits

Disentangling with a joint fit

Simultaneous fit of PDFs and W parameter





PBSP Collaboration, Hammou et al, arXiv:2307.10370

Simultaneous PDF SMEFT Fits

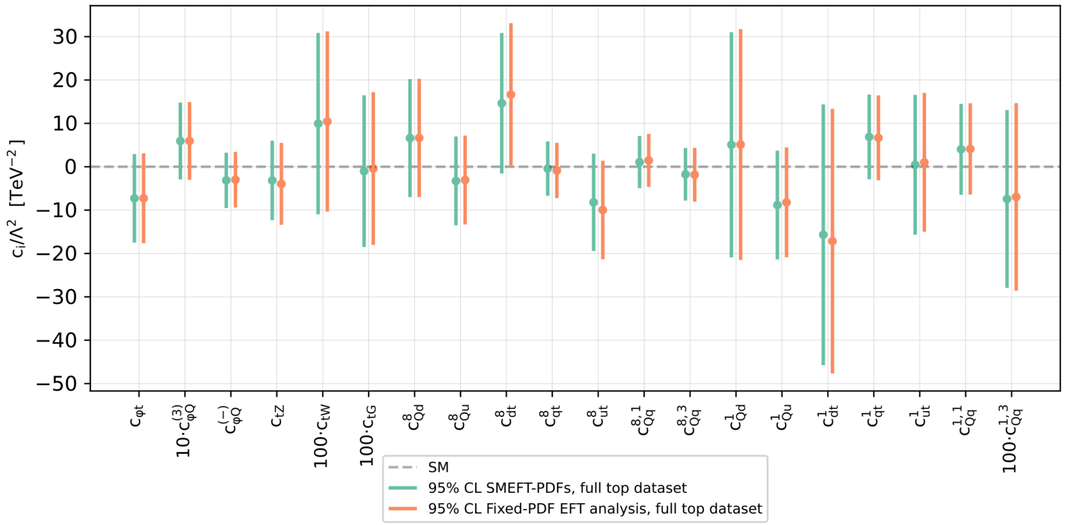
SMEFT - PDF interplay in the top sector

NNPDF40 dataset + all available top datasets based on the full Run II luminosity

Moderate effect on WC ~ 5-10%

Fixed PDF, EFT fits

SMEFT-PDF fits

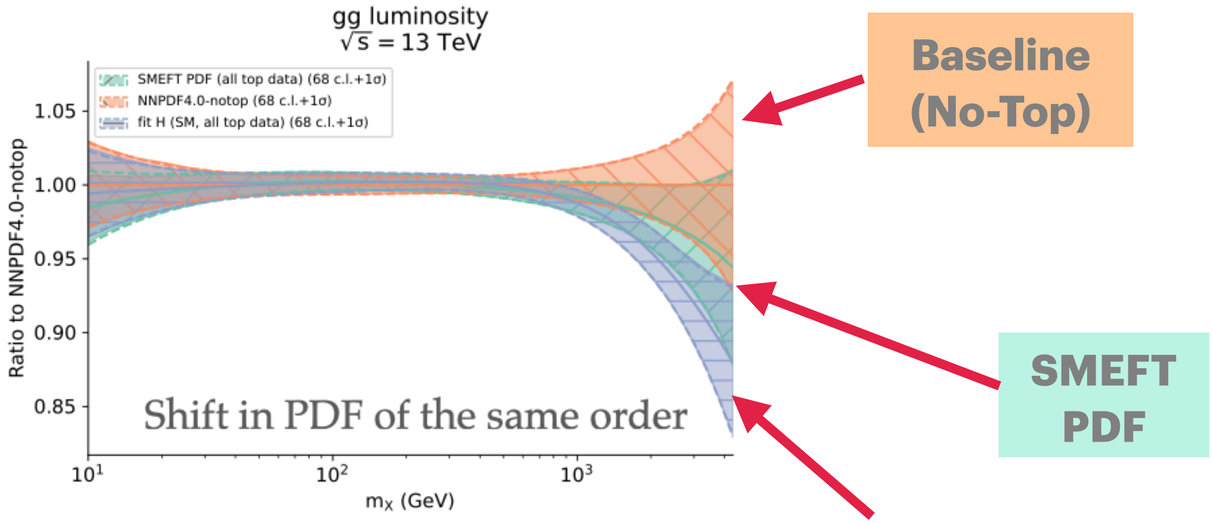


PBSP Collaboration, Kassabov et al, arXiv:2303.06159

Simultaneous PDF SMEFT Fits

SMEFT - PDF interplay in the top sector

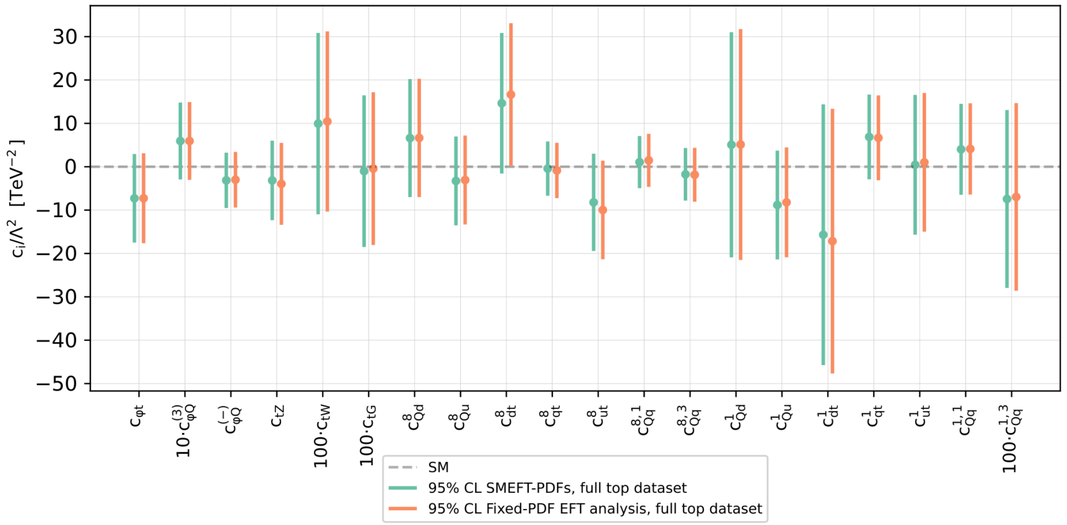
NNPDF40 dataset + all available top datasets based on the full Run II luminosity



Baseline (No-Top)

SMEFT PDF

NNPDF40, all top data

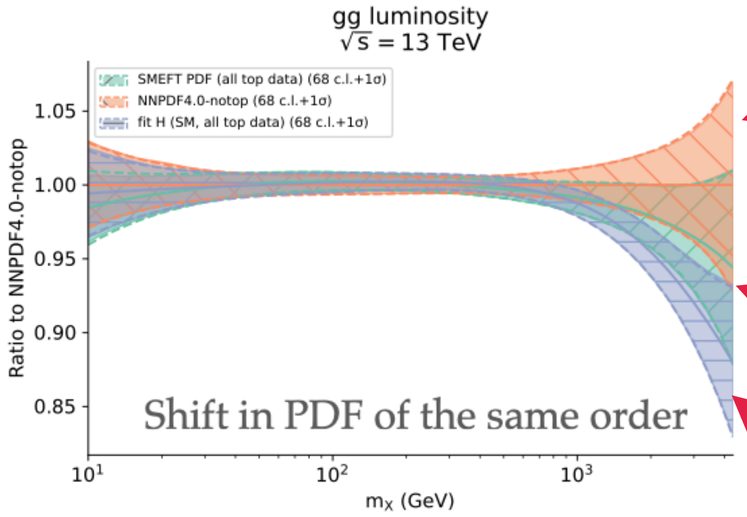


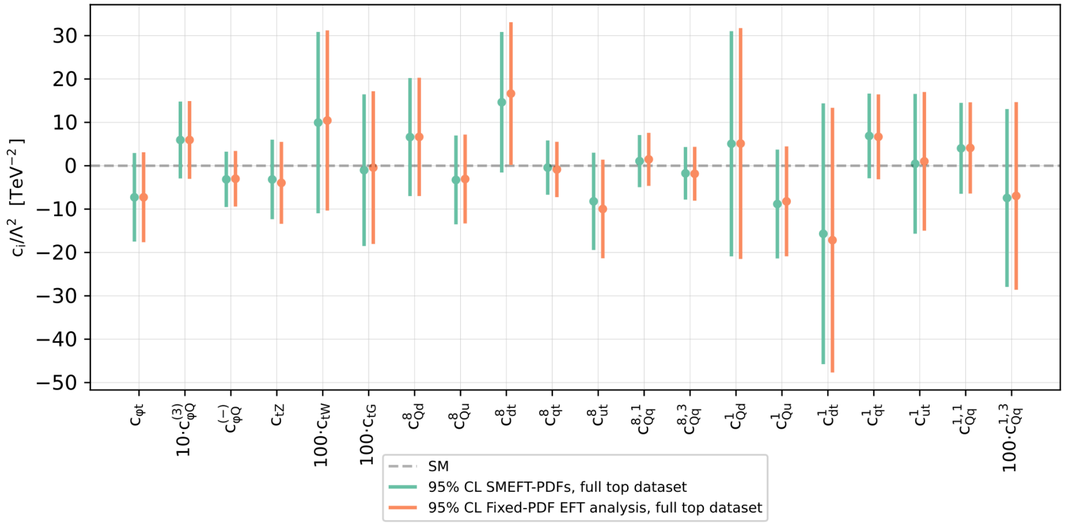
PBSP Collaboration, Kassabov et al, arXiv:2303.06159

Simultaneous PDF SMEFT Fits

SMEFT - PDF interplay in the top sector

NNPDF40 dataset + all available top datasets based on the full Run II luminosity





Baseline (No-Top)

SMEFT PDF

NNPDF40, all top data

→ Standard PDF fits are (NP contamination) safe for current data

PBSP Collaboration, Kassabov et al, arXiv:2303.06159

Summary

- SIMUnet:
  - Open-source with detailed documentation
  - Extends the NNPDF methodology to include SMEFT corrections in theoretical predictions
  - Allows for the simultaneous fit of PDFs and WCs
- SIMUnet Applications:
  - Contaminated PDFs study → need of simultaneous fits
  - PDFs and SMEFT interplay in the Top sector
- Outlook
  - Uncertainty quantification for non-linear inverse problems
  - Work in progress for new PDF / SMEFT - fitting methodology that is reliable when non-linearities are relevant

Backup

PDF-SMEFT interplay questions

Question: won't the PDF-SMEFT interplay be negligible?

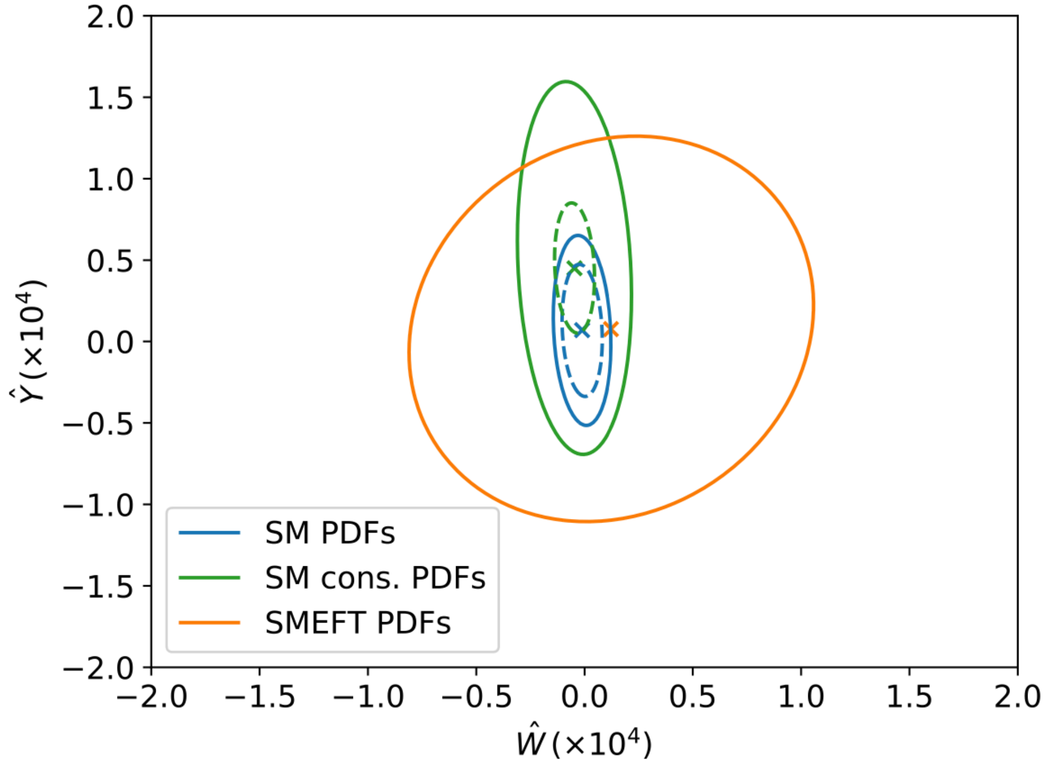
→ It was shown in Carrazza et al., 1905.05215, that interplay is very mild in the case of simultaneous extractions of four-fermion operators and PDFs using DIS-only data.

→ Similarly, it was shown in Greljo et al., 2104.02723, that interplay is mild between the $\hat{W}$, $\hat{Y}$ operators and PDFs using current DIS and DY data.

PDF-SMEFT interplay questions

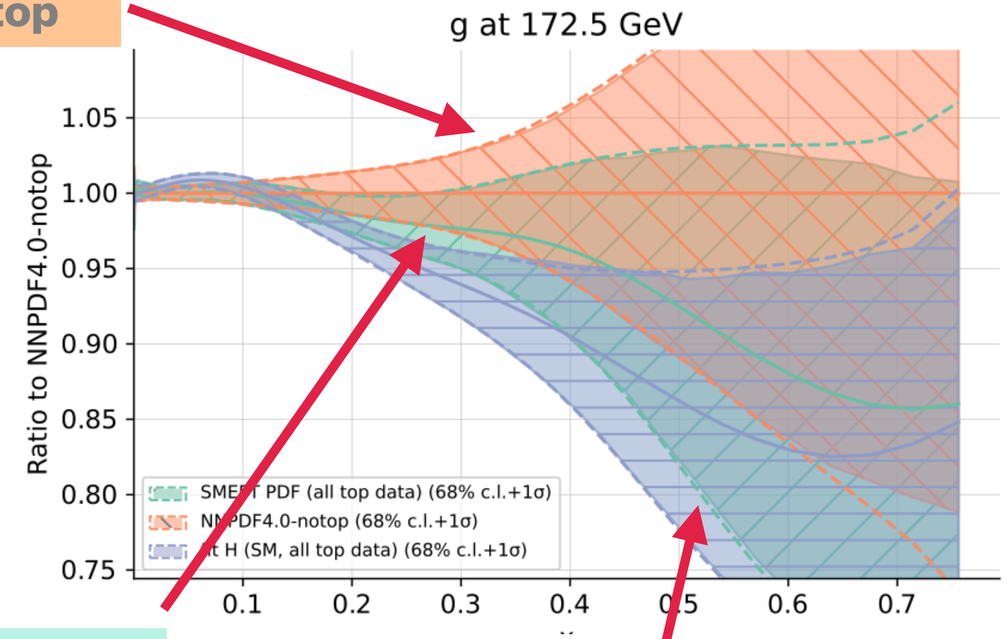
Question: won't the PDF-SMEFT interplay be negligible?

→ However, it was also shown in Greljo et al., 2104.02723, that interplay is very important between the $\hat{W}$, $\hat{Y}$ operators and PDFs when using project high luminosity DY data.



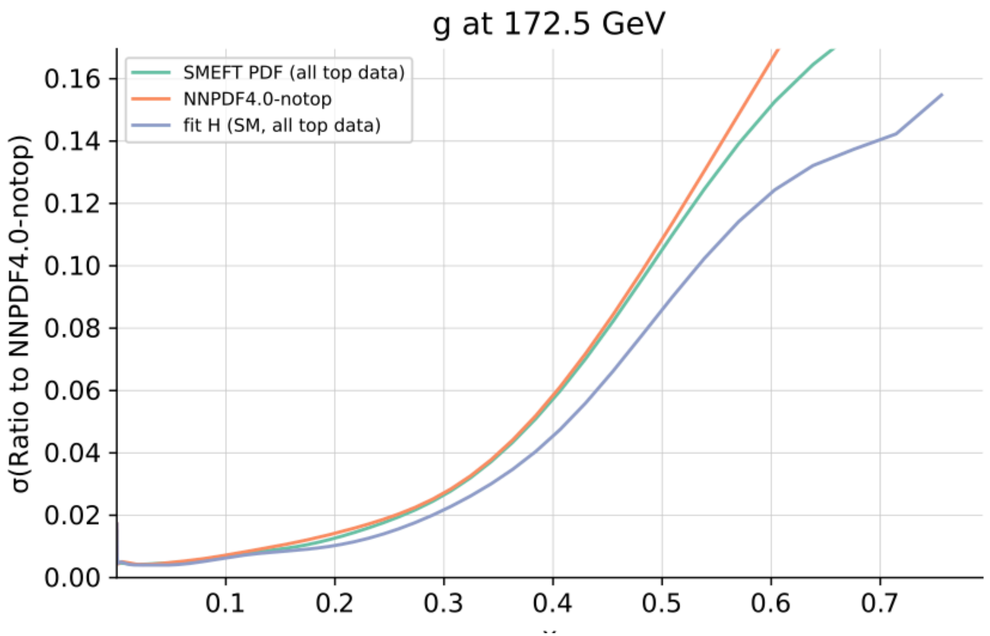
Simultaneous PDF SMEFT Fits: PDFs

NNPDF40, no-top



SMEFT PDF

NNPDF40, all top data



- Difference between SM gluon PDF and SMEFT gluon PDF

→ Larger PDF uncertainties

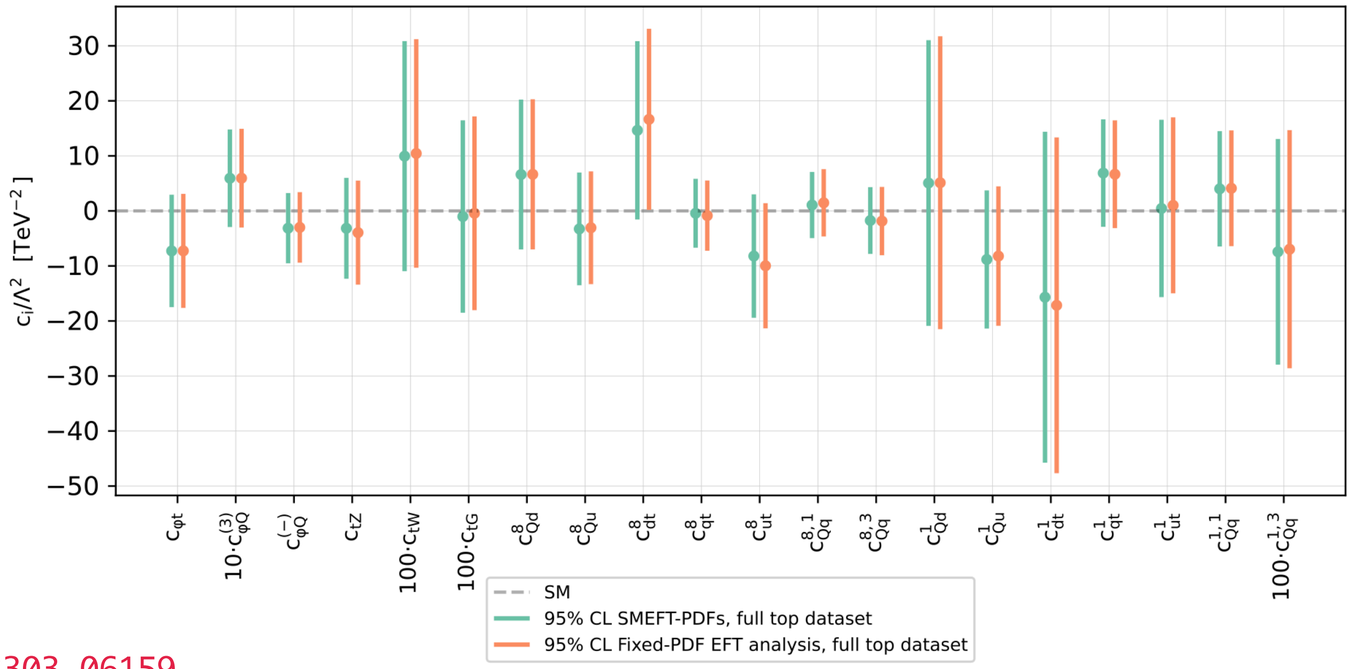
Kassabov et al, arXiv: 2303.06159

Simultaneous PDF SMEFT Fits: EFT

- Moderate impact at the level of the bounds on the EFT coefficients
- Mild shift in the central values and slight broadening of the uncertainties

Fixed PDF, EFT fits

SMEFT-PDF fits



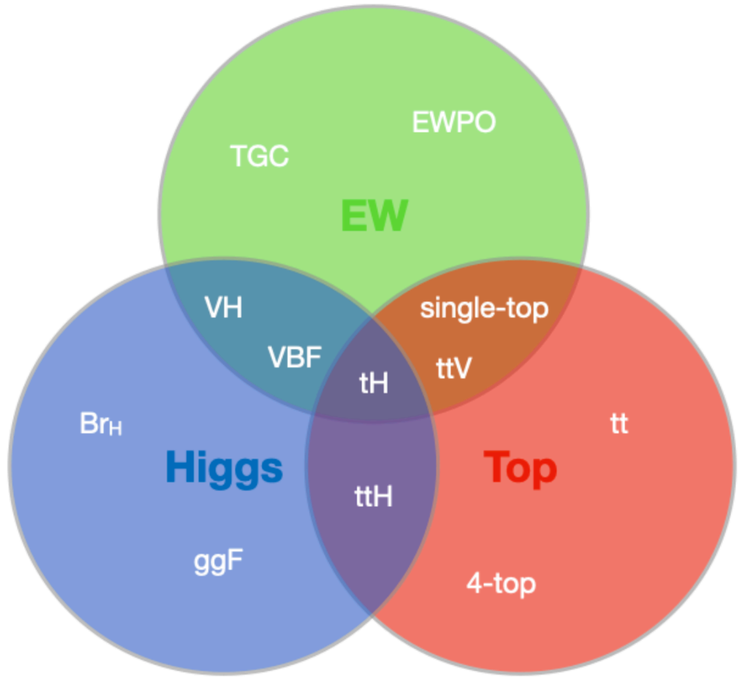
Kassabov et al, arXiv: 2303.06159

SIMUnet Release: New Data

- **EW precision observables**

→ 44 data points

Dataset Name	N_dat	Reference	Theory Tables
EWPO on Z resonance	19	hep-ex/ 0509008	SMEFiT
W branching ratio	3	1302.3415	SMEFiT
Bhabha Scattering	21	1302.3415	SMEFiT
Alpha EW	1	PDG	PDG



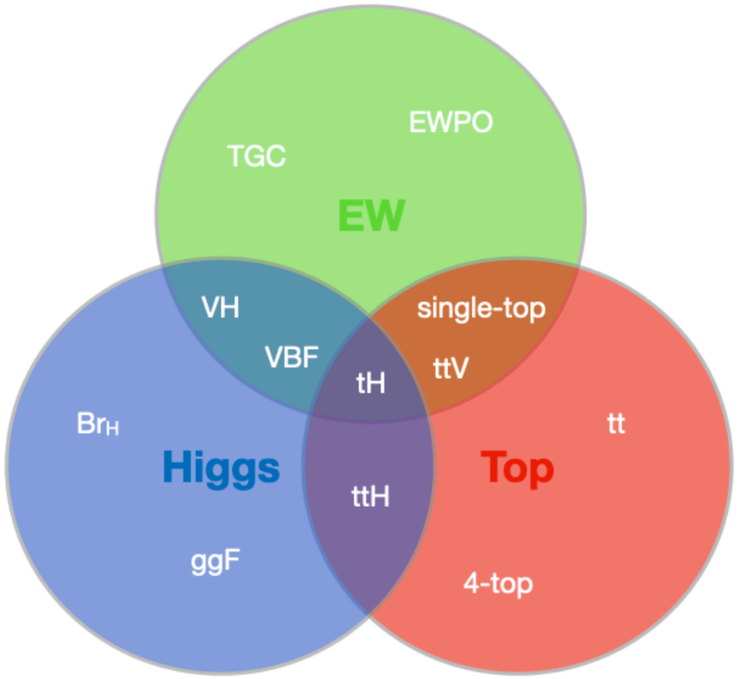
© Ken Mimasu

SIMUnet Release: New Data

- **Diboson production from LEP and LHC**

→ 82 data points

Dataset Name	N_dat	Reference	Theory Tables
LEP W- differential angular xsec	40	1302.3415	SMEFiT
ATLAS, W+W- differential xsec	13	1905.04242	SMEFiT
ATLAS, WZ transverse mass	6	1902.05759	SMEFiT
CMS, WZ transverse momentum	11	1901.03428	SMEFiT
ATLAS, Zjj azimuthal differential xsec	12	2006.15458	FitMaker



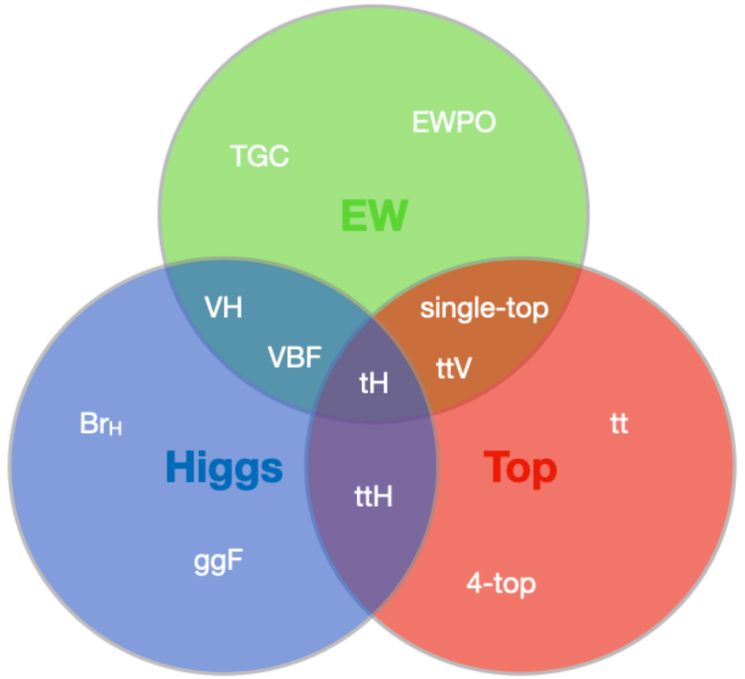
© Ken Mimasu

SIMUnet Release: New Data

- **Higgs production and decay**

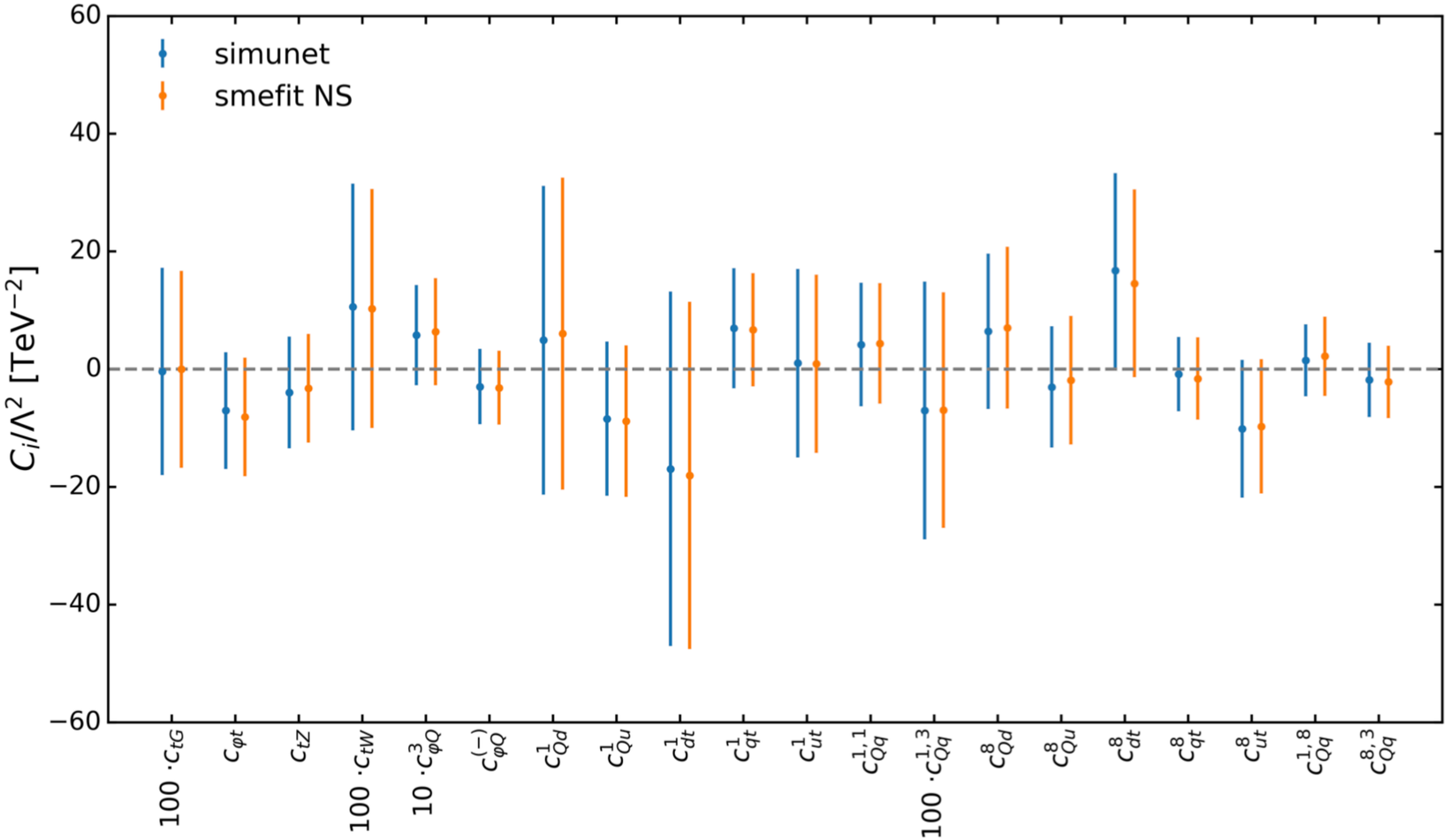
→ 73 data points from Run I and II

Dataset Name	N_dat	Reference	Theory Tables
Production and decay rates (8 TeV)	22	1606.02266	SMEFiT
Production and decay rates (13 TeV)	24	1809.10733	SMEFiT
Production xsec and branching fractions	25	1909.02845	FitMaker
Zgamma decay	1	2005.05382	FitMaker
Dimuon decay	1	2007.07830	FitMaker



© Ken Mimasu

SIMUnet: Fixed-PDF EFT fits



- Freeze weights parametrising PDFs ($\theta = \bar{\theta}$) and optimise only Wilson Coefficients c

Kassabov et al, arXiv: 2303.06159