

The Top Quark Legacy of the LHC Run II for PDF and SMEFT Analyses

Maeve Madigan
Heidelberg University

based on 2303.06159: Z. Kassabov, MM, L. Mantani, J. Moore, M. Morales Alvarado, J. Rojo, M. Ubiali



**UNIVERSITÄT
HEIDELBERG**
ZUKUNFT
SEIT 1386



Phenomenology Symposium 2023
University of Pittsburgh

Global SMEFT interpretations

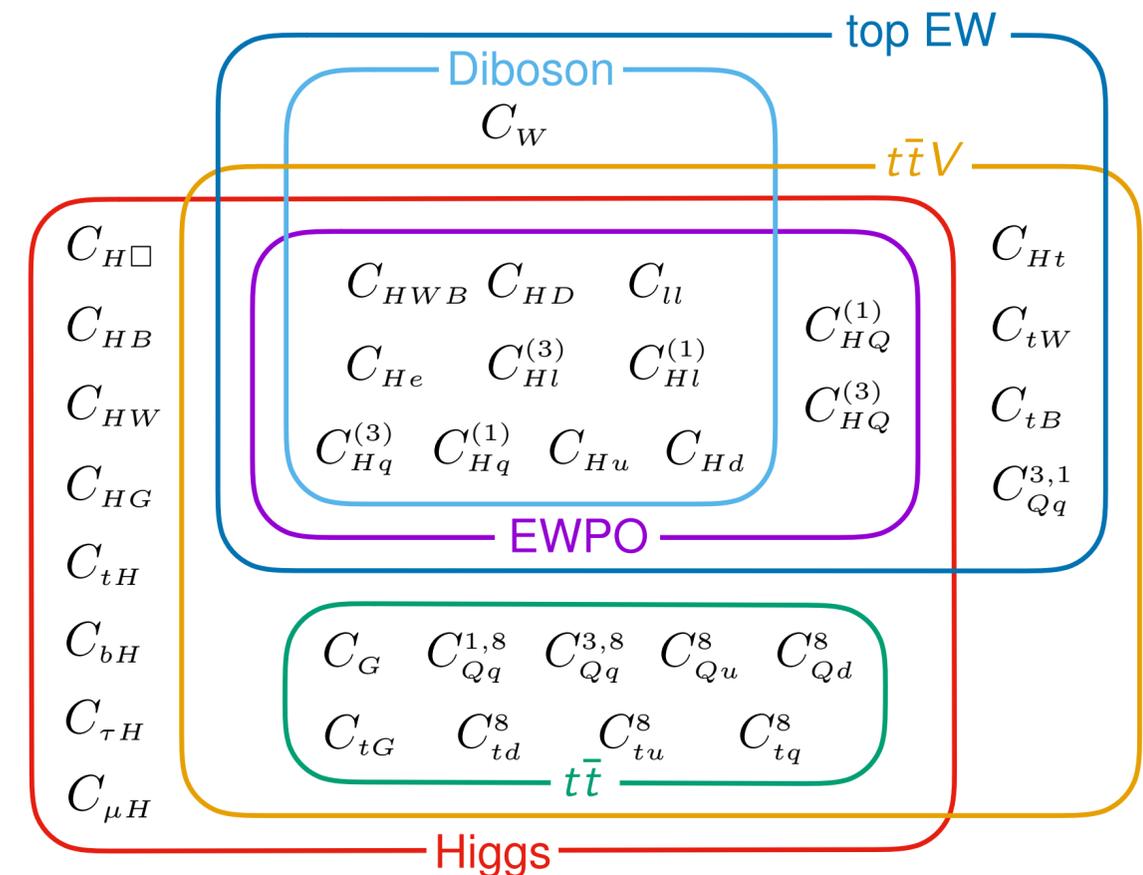
The SMEFT: a powerful framework for capturing deviations from the SM:

$$\mathcal{L}_{\text{SMEFT}} = \mathcal{L}_{\text{SM}} + \frac{C^{(5)}}{\Lambda} \mathcal{O}^{(5)} + \sum_i \frac{C_i^{(6)}}{\Lambda^2} \mathcal{O}_i^{(6)} + \dots$$

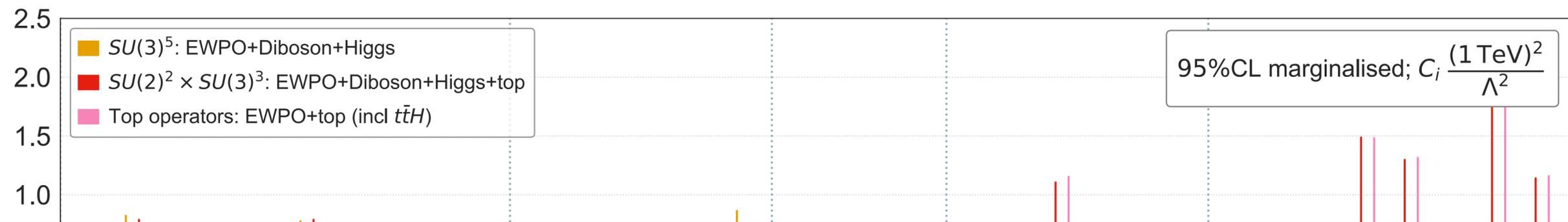
The SMEFT framework connects different sectors of observables measured at the LHC.

We need to take a **global approach**, including as many datasets as possible.

→ Model-independent interpretation of BSM physics in LHC data



2012.02779, J. Ellis, MM, K. Mimasu, V. Sanz, T. You



Global SMEFT fits

Higgs, diboson and electroweak precision data

J. Ellis et. al, 1803.03252

E. da Silva Almeida et. al, 1812.01009:

A. Biekötter et. al, 1812.07587

A. Falkowski et. al, 1911.07866

Top data

I. Brivio et. al, 1910.03606:

N. Hartland et. al, 1901.05965:

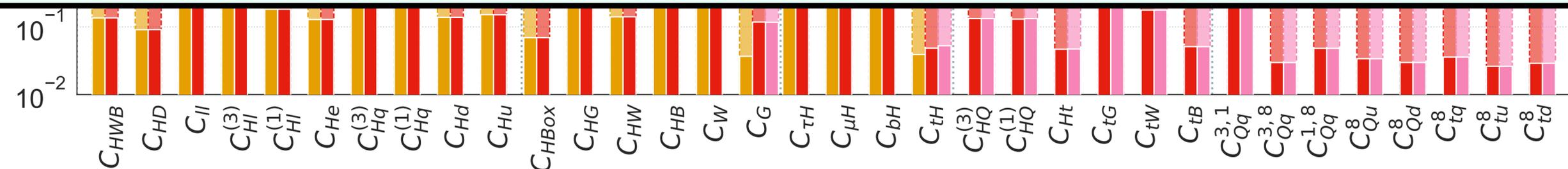
+ many others....

Higgs, diboson and top data

J. Ethier et. al, 2105.00006

Higgs, diboson, top and electroweak precision data

J. Ellis et. al, 2012.02779



Inputs and assumptions

Global SMEFT fits are dependent on many inputs and assumptions:

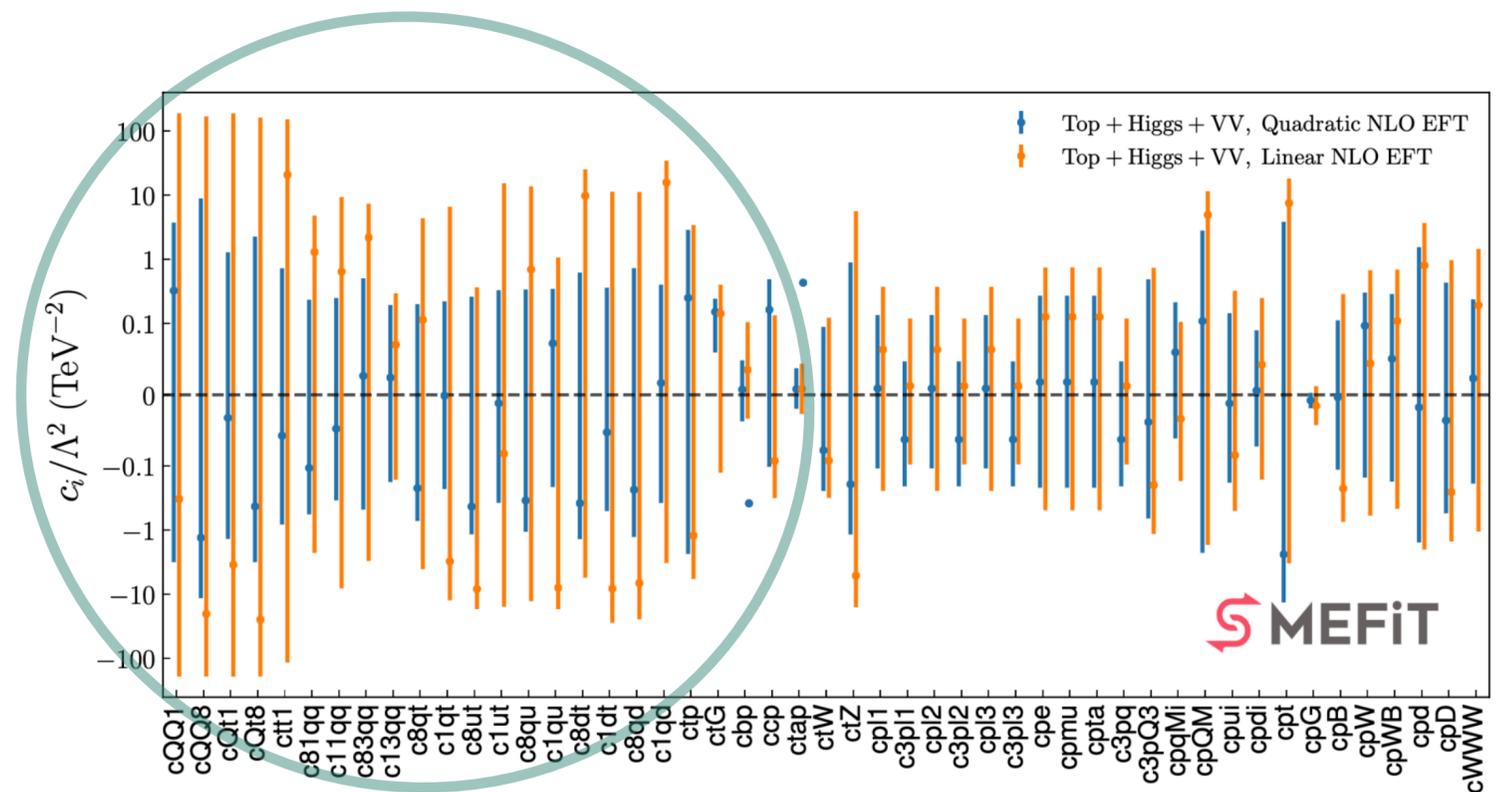
- SMEFT flavour symmetry
- Electroweak input scheme: $\{\alpha_{EW}, m_Z, G_F\}$

Inputs and assumptions

Global SMEFT fits are dependent on many inputs and assumptions:

- SMEFT flavour symmetry
- Electroweak input scheme: $\{\alpha_{EW}, m_Z, G_F\}$
- Inclusion of $\mathcal{O}(\Lambda^{-4})$ contributions:

$$\sigma = \sigma_{SM} + \frac{C}{\Lambda^2} \sigma_{lin} + \frac{C^2}{\Lambda^4} \sigma_{quad}$$



J. Ethier et. al, 2105.00006

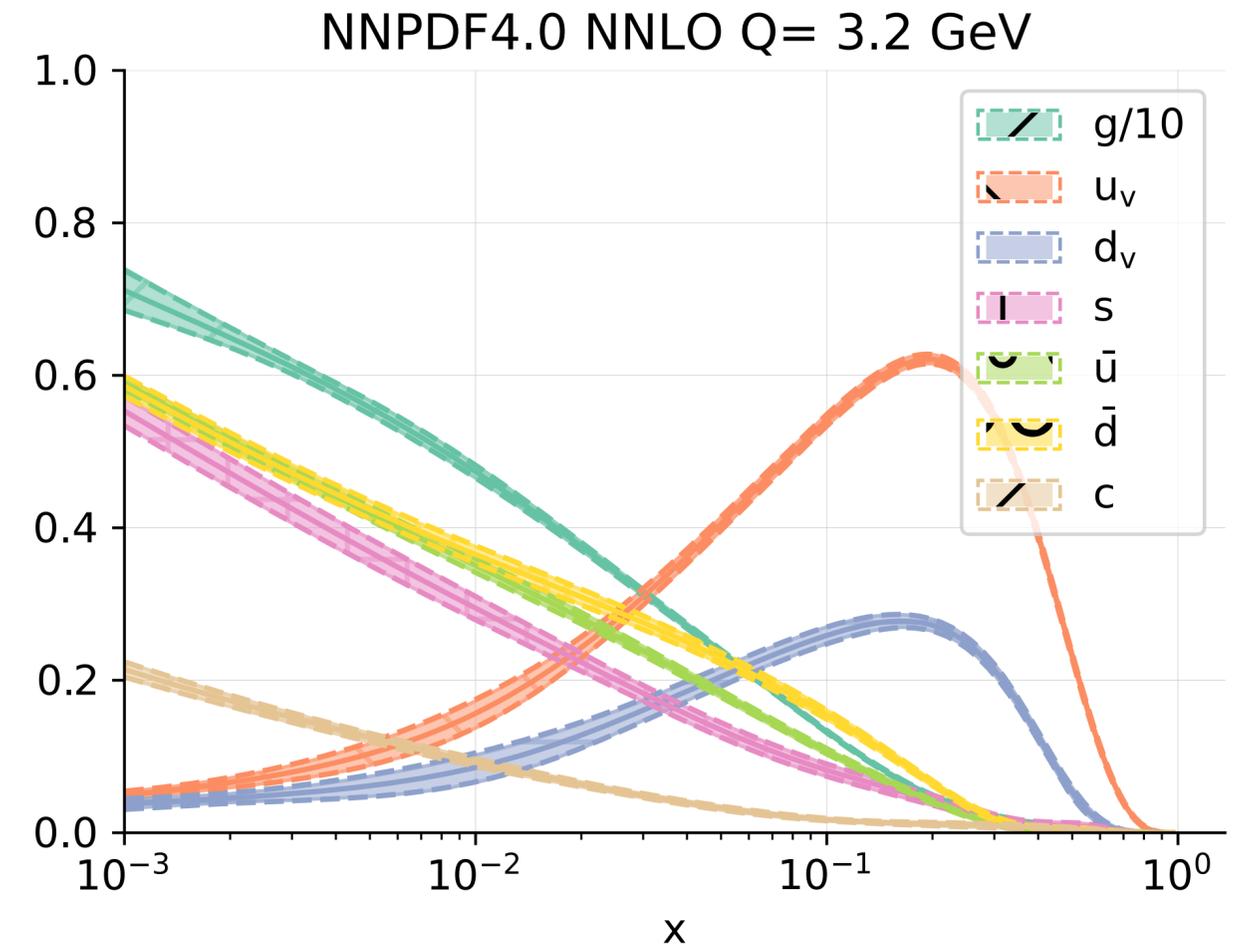
- important for SMEFT fits in the top sector
- non-Gaussian likelihood; methodologically more challenging

Inputs and assumptions

Global SMEFT fits are dependent on many inputs and assumptions:

$$f_q(x, Q^2)$$

- SMEFT flavour symmetry
- Electroweak input scheme: $\{\alpha_{EW}, m_Z, G_F\}$
- Inclusion of $\mathcal{O}(\Lambda^{-4})$ contributions
- Choice of likelihood
-
- **Parton distribution functions**



Ball et. al, NNPDF4.0, 2109.02653

Parton distribution functions

PDFs are an input to SMEFT fits: $\sigma_{\text{SMEFT}}(C) = f_1 \otimes f_2 \otimes \hat{\sigma}_{\text{SMEFT}}(C)$

But PDFs are found assuming the SM: $\sigma = f_1 \otimes f_2 \otimes \hat{\sigma}_{\text{SM}}$

↙ *'Standard Model PDFs'*

Parton distribution functions

PDFs are an input to SMEFT fits: $\sigma_{\text{SMEFT}}(C) = f_1 \otimes f_2 \otimes \hat{\sigma}_{\text{SMEFT}}(C)$

But PDFs are found assuming the SM: $\sigma = f_1 \otimes f_2 \otimes \hat{\sigma}_{SM}$

↙ *'Standard Model PDFs'*

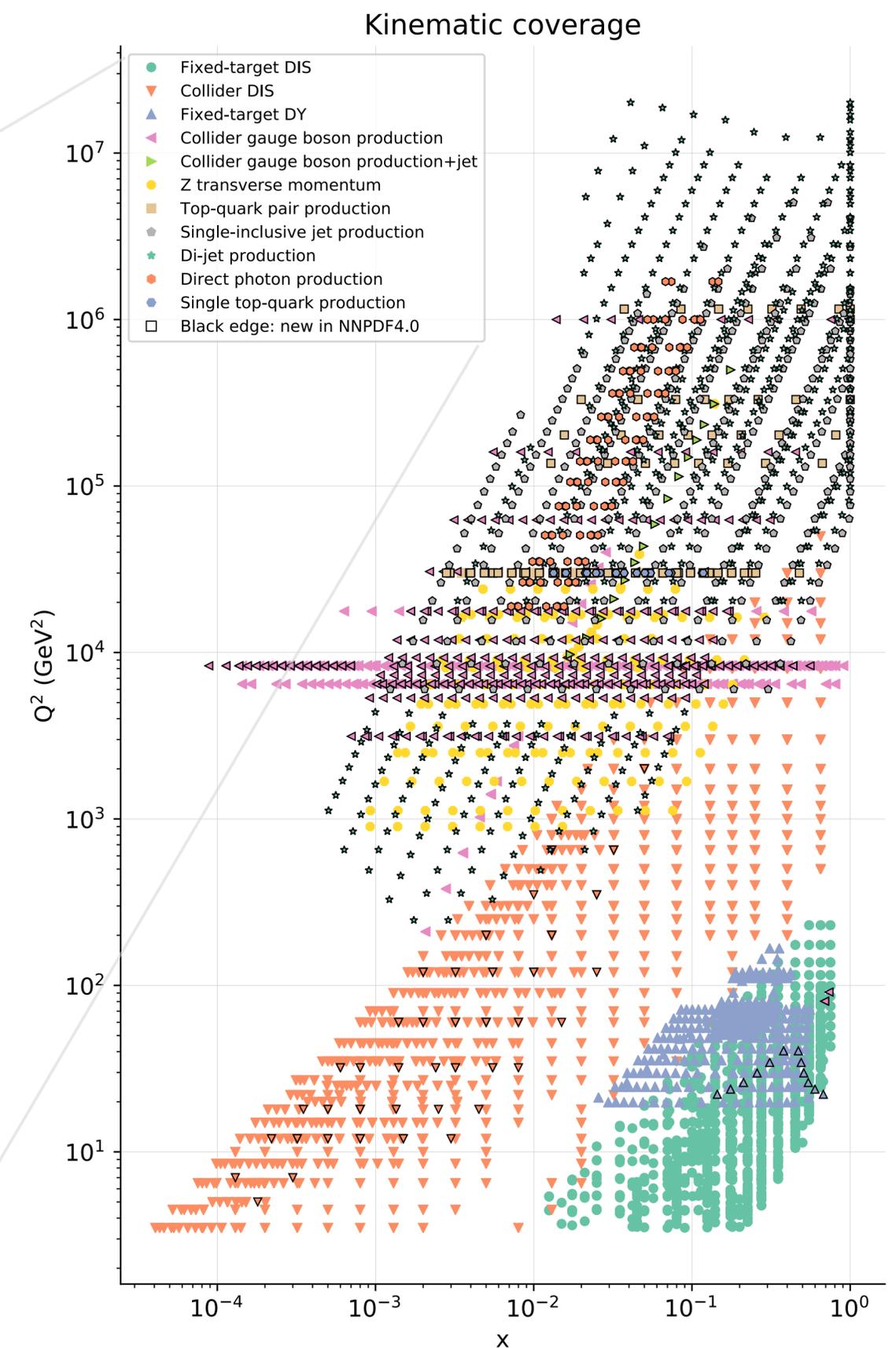
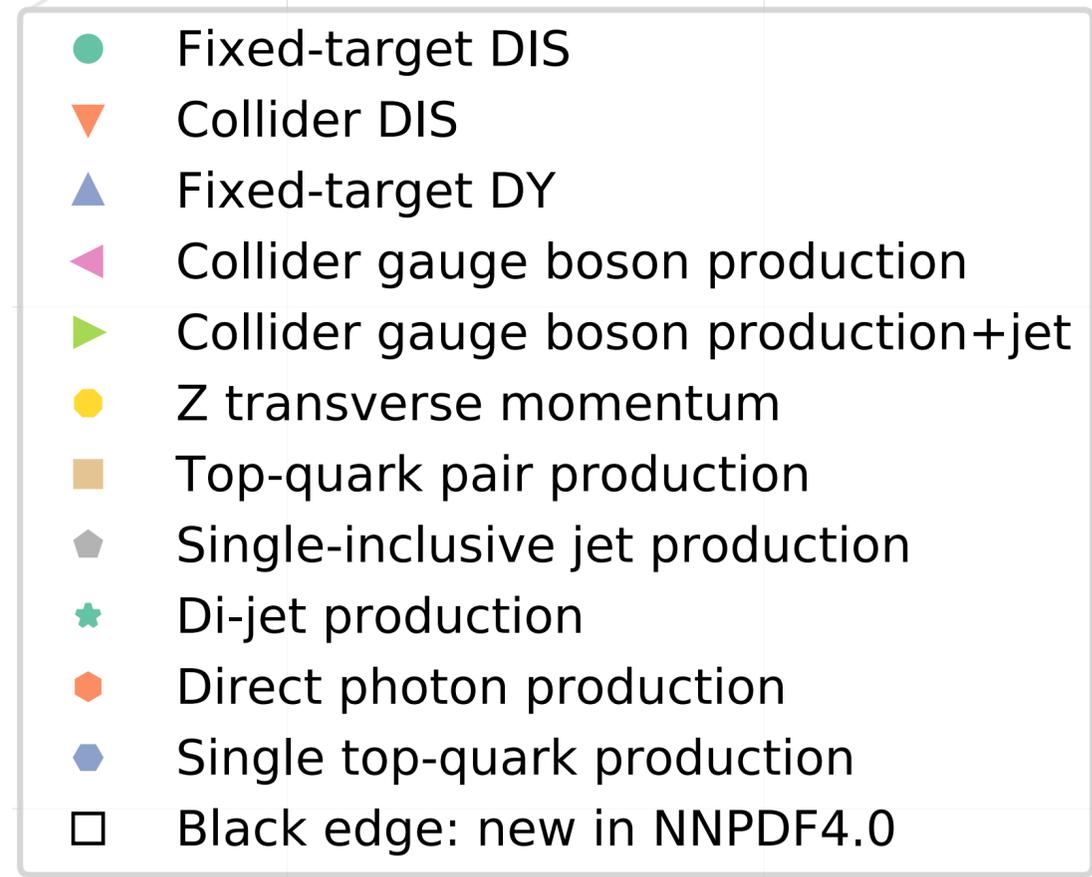
= an inconsistency in our theoretical predictions

Data overlap

Often the data used in PDF fits are also used in EFT fits.

This overlap will grow as we take the global approach to constraining the SMEFT.

Data included in NNPDF4.0, [2109.02653]:

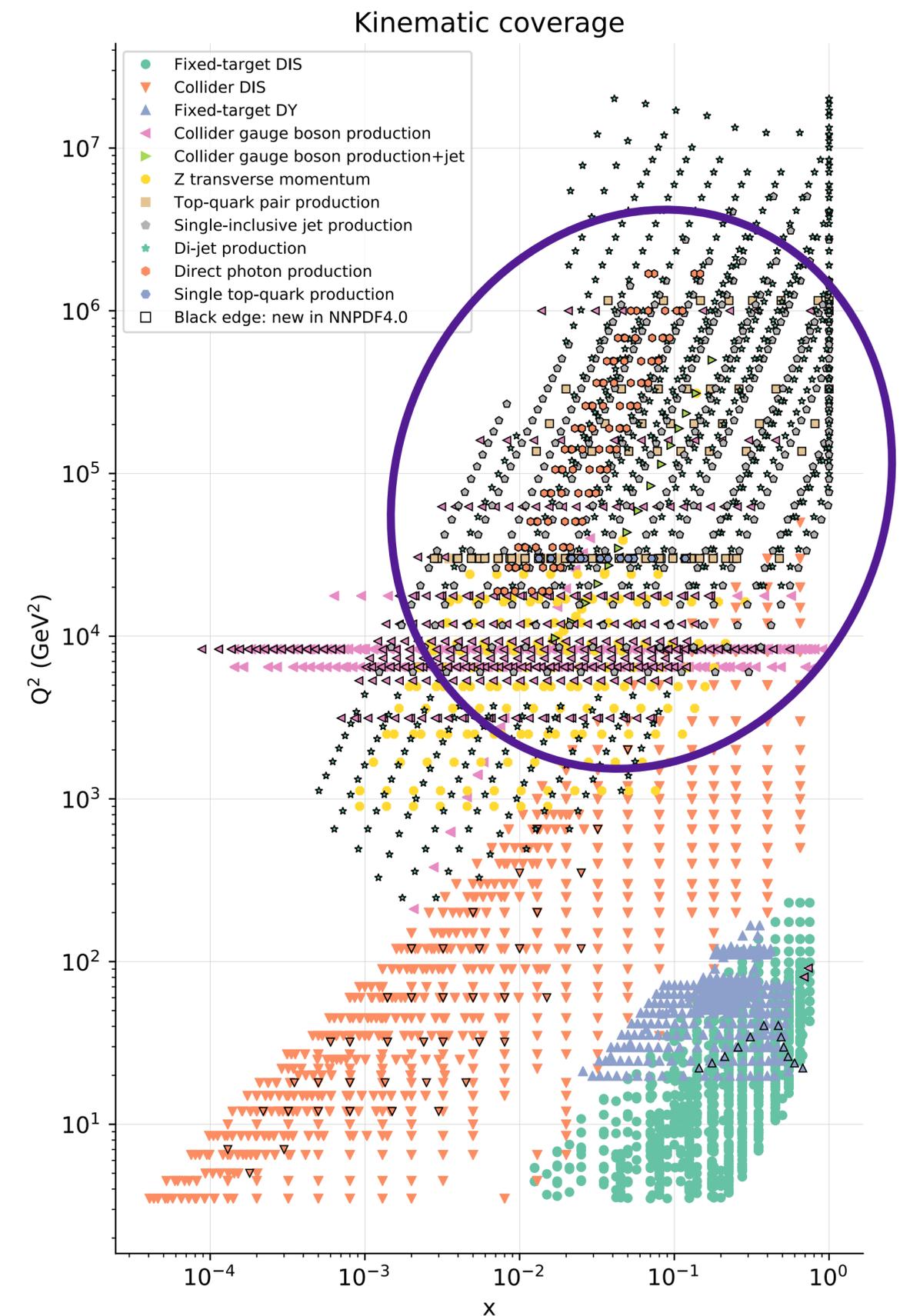
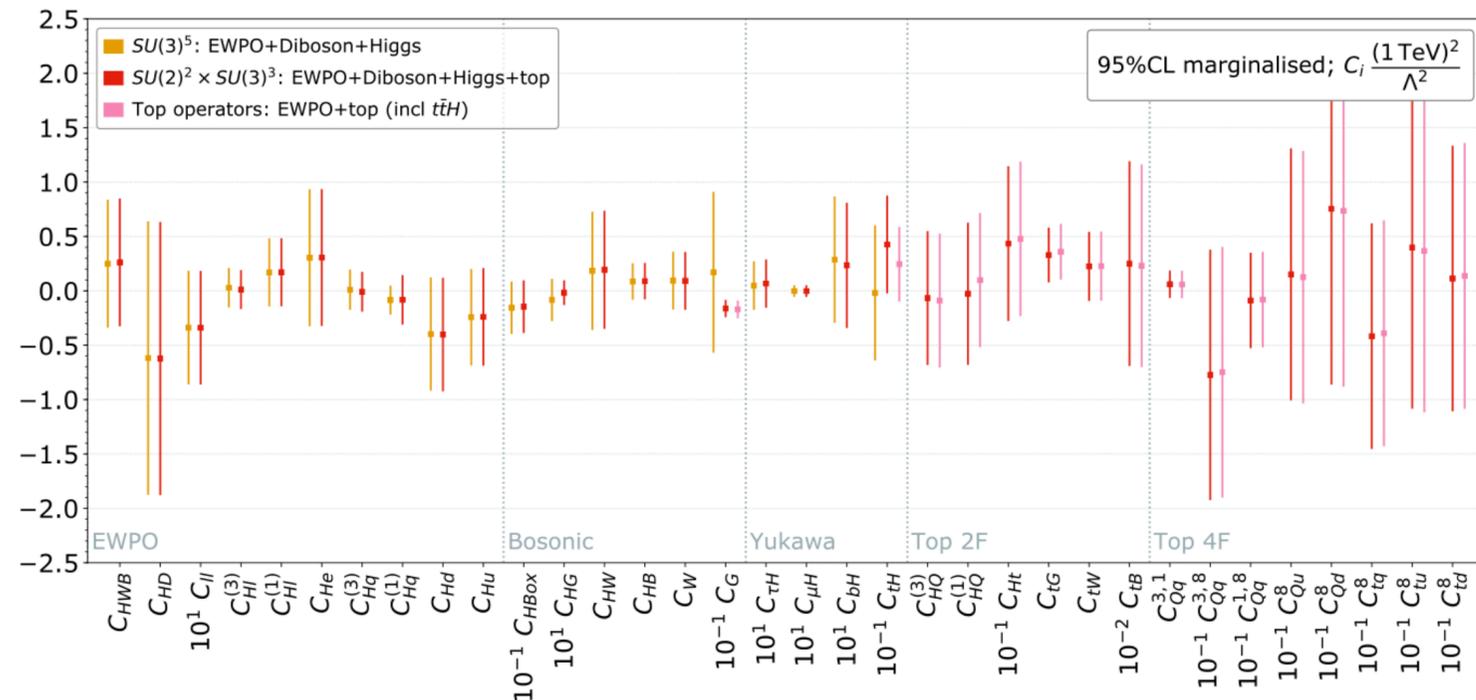


Data overlap

Often the data used in PDF fits are also used in EFT fits.

This overlap will grow as we take the global approach to constraining the SMEFT.

- ▶ e.g. Top quark data used to fit the SMEFT in the global fit of *2012.02779, J. Ellis, MM, K. Mimasu, V. Sanz, T. You*



How do the constraints on the **SMEFT** change if we perform a consistent simultaneous determination of the PDFs and SMEFT?

How do the constraints on the **SMEFT** change if we perform a consistent simultaneous determination of the PDFs and SMEFT?

How do the **PDFs** change if we perform a consistent simultaneous determination of the PDFs and SMEFT?

How do the constraints on the **SMEFT** change if we perform a consistent simultaneous determination of the PDFs and SMEFT?

How do the **PDFs** change if we perform a consistent simultaneous determination of the PDFs and SMEFT?

Could we be absorbing signs of new physics into the PDFs?

Simultaneous PDF-SMEFT determinations

Deep Inelastic Scattering data

- proof of concept: PDF and SMEFT effects can be disentangled
- no significant interplay found

Carrazza et al.: PRL 123 (2019) 13, 132001

DIS + Drell-Yan data

- including high-mass DY tails

- sensitive to both quark/antiquark PDFs and SMEFT 4F operators
- significant potential for interplay at HL-LHC

Greljo et. al 2104.02723

Top quark data

2303.06159 by Z. Kassabov, MM, L. Mantani, J. Moore, M. Morales Alvarado, J. Rojo, M. Ubiali

See also 2201.06586, 2211.01094



PDF-EFT interplay in the top sector

Top quark data provides important constraints on the large- x region of the gluon PDF.

This impact is largely driven by **top quark pair production** cross sections and differential distributions.

e.g. Czakon et. al, 1303.7215, 1611.08609, 1912.08801

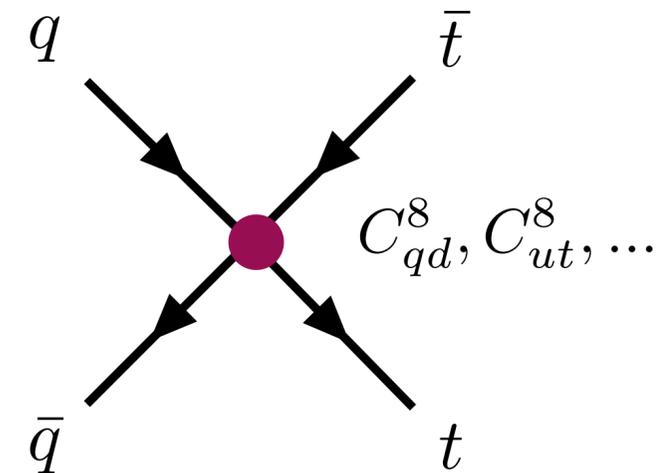
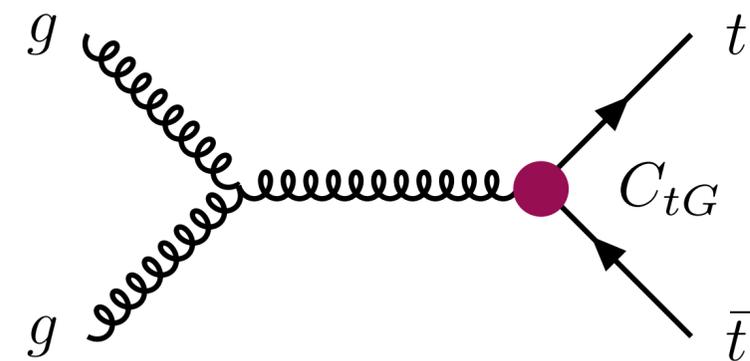
PDF-EFT interplay is most likely to occur between:

- The gluon PDF
- Dimension-6 operators which contribute to $t\bar{t}$ production:

C_{tG}

$C_{qd}^8, C_{qu}^8, C_{dt}^8,$
 $C_{ut}^8, C_{qt}^8, C_{qq}^{8,3}, C_{qq}^{8,1}$

$C_{qd}^1, C_{qu}^1, C_{dt}^1,$
 $C_{ut}^1, C_{qt}^1, C_{qq}^{1,3}, C_{qq}^{1,1}$



PDF-EFT interplay in the top sector

Data: 175 datapoints including LHC Run II measurements of $t\bar{t}$ incl. A_C and W_{hel} , $t\bar{t}V$, single top, tW , $t\bar{t}t\bar{t}$, $t\bar{t}b\bar{b}$

Theory:

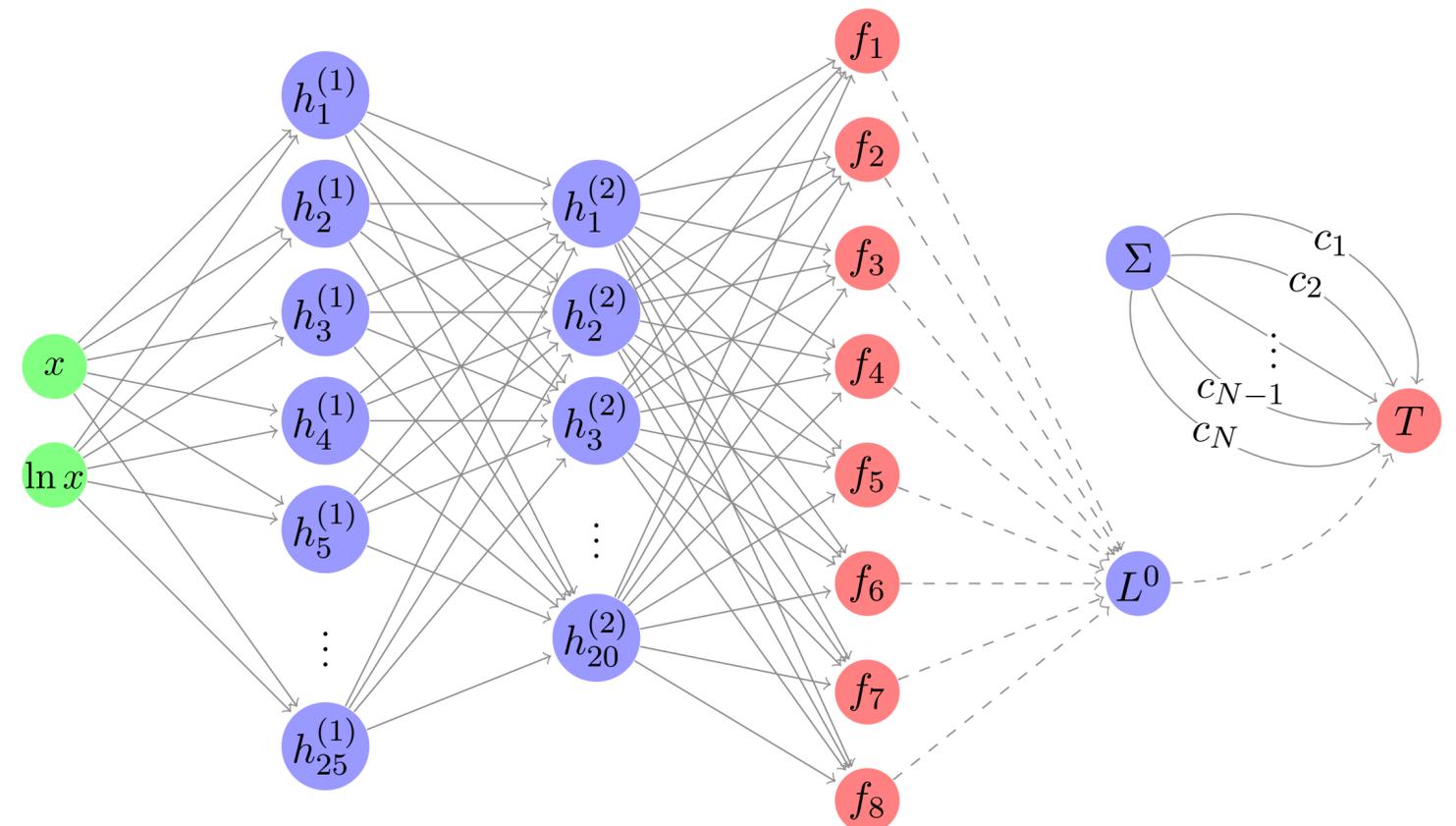
SM: NNLO QCD

SMEFT: Apply k-factors calculated at NLO with SMEFT@NLO

Methodology: SIMUnet

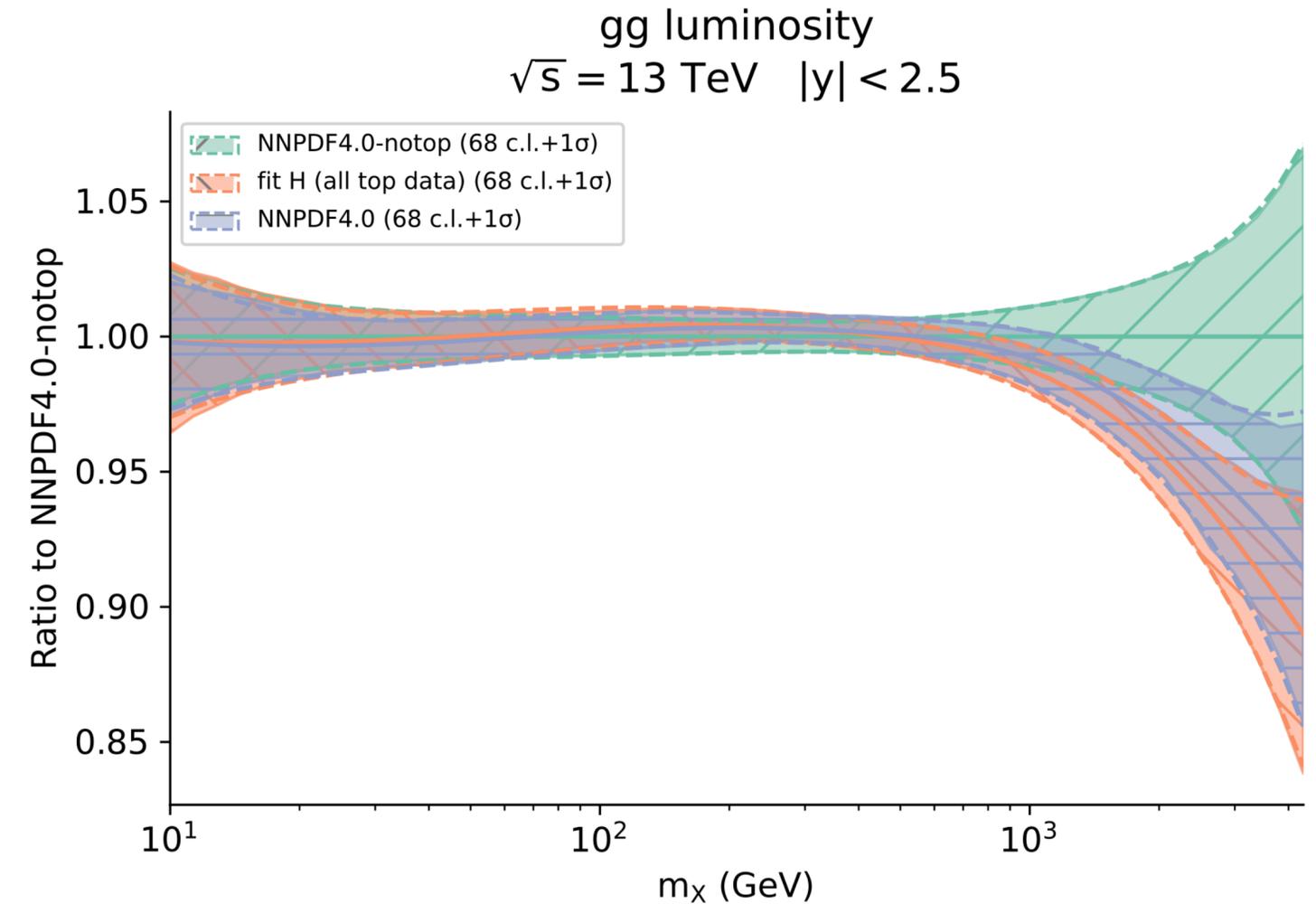
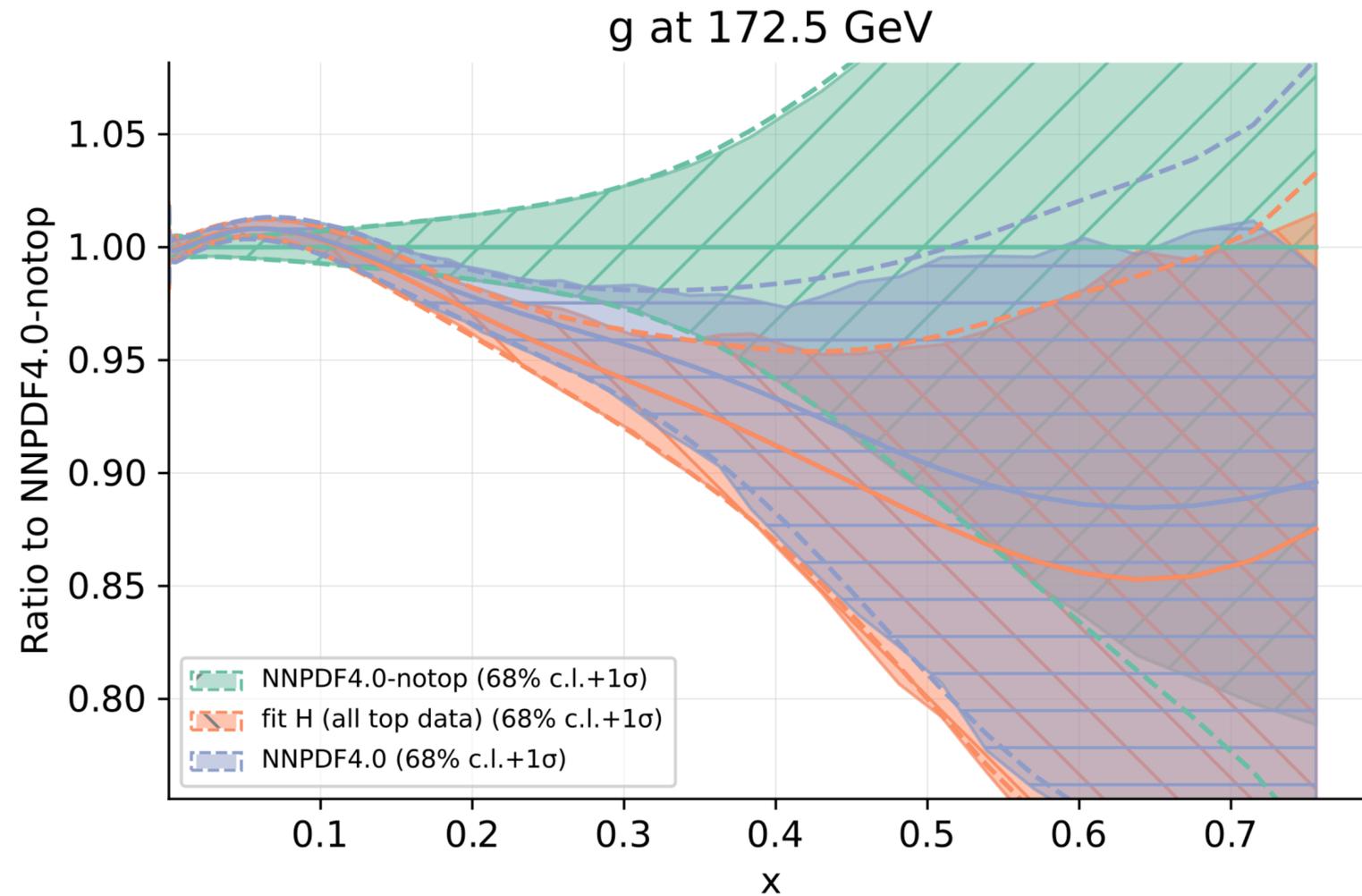
S. Iranipour, M. Ubiali, 2201.07240

Propagates uncertainties from data to NN parameters using the Monte Carlo replica method



Impact of top data on PDF fits

Fixed SM PDFs - no SMEFT effects included:

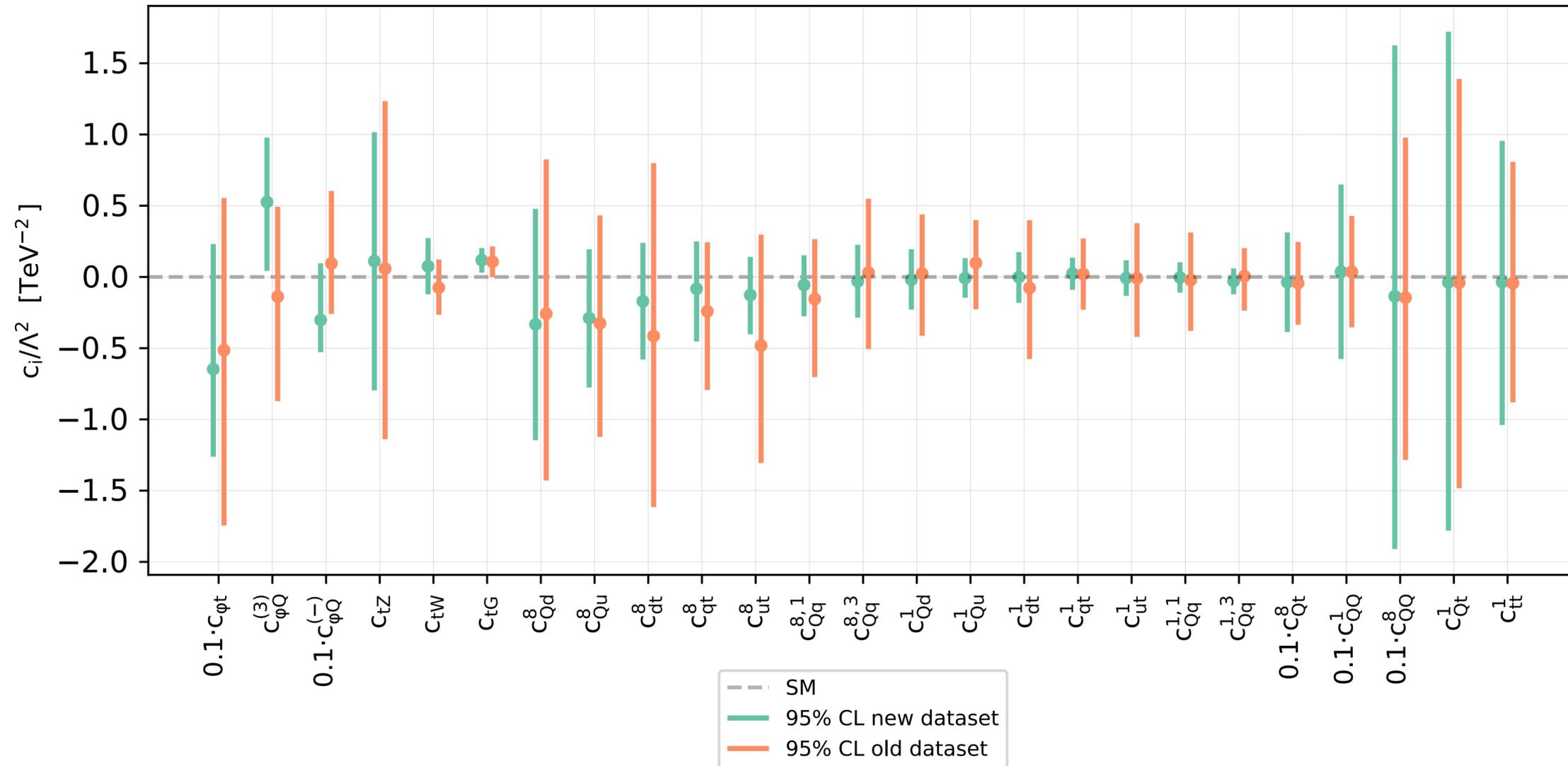


Top quark data impacts the large- x gluon PDF:

- shift in central values relative to NNPDF4.0 and no-top PDF fit (compatible within uncertainties)
- uncertainty is reduced

Impact of top data on SMEFT fits

Input PDFs are SM PDFs (NNPDF4.0)



SMEFT constraints at quadratic order are compared to those obtained from the dataset used by *J. Ethier et. al 2105.00006*

Increase in high- $m_{t\bar{t}}$ $t\bar{t}$ measurements leads to better constraints on 4-fermion operators

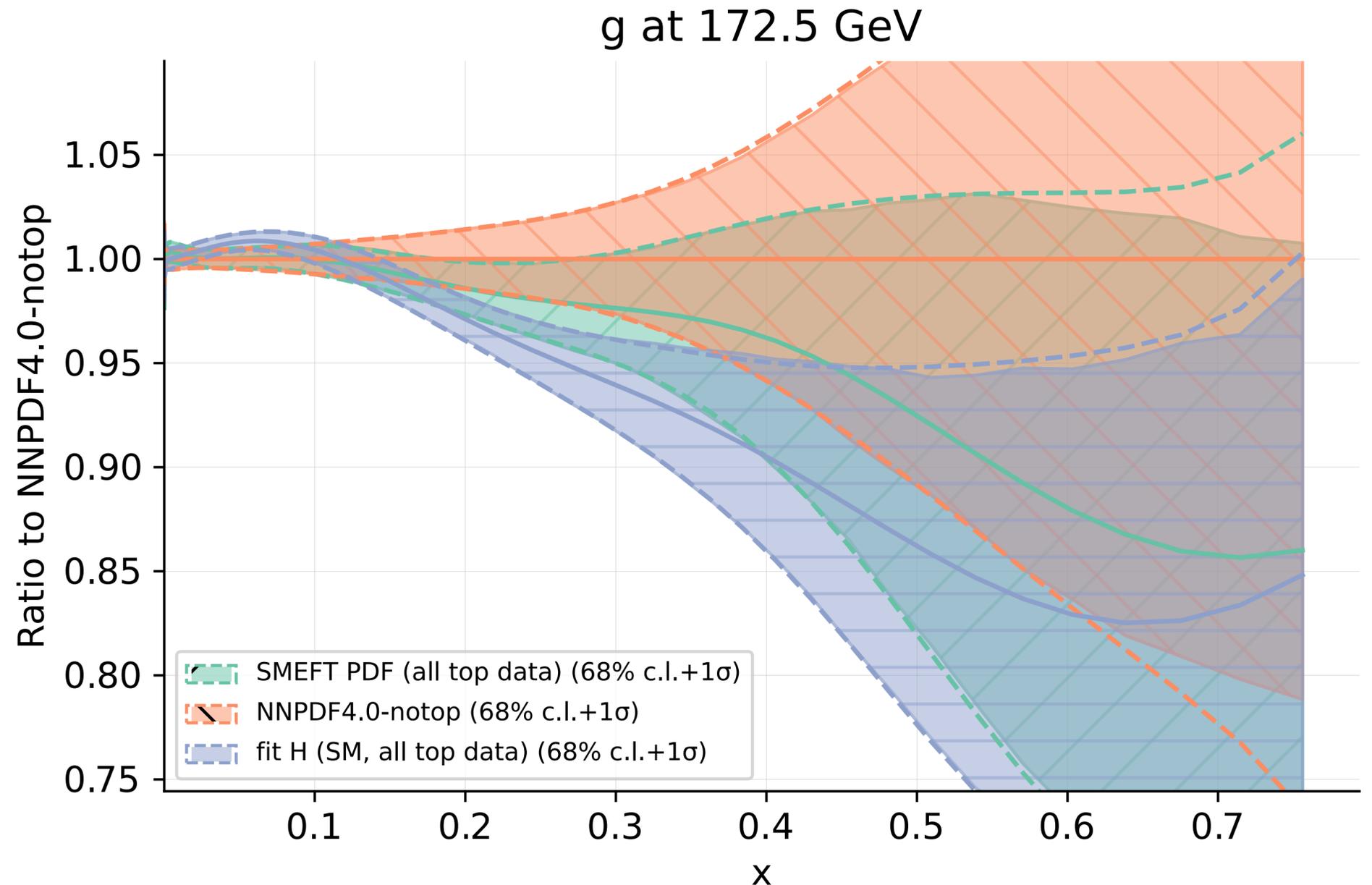
SMEFT constraints shown here are produced using the SMEFiT public code 2302.06660

Simultaneous PDF-SMEFT determination

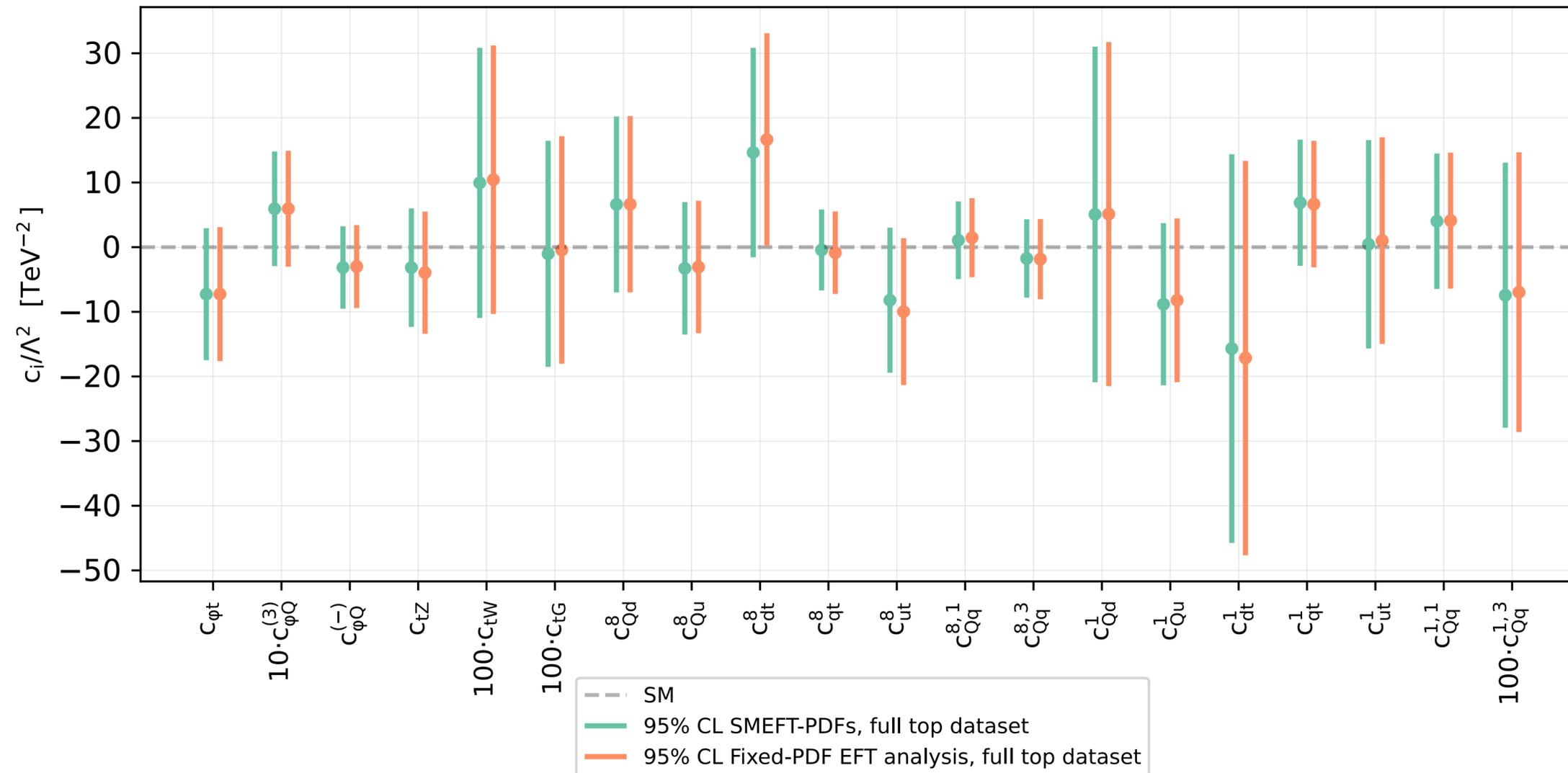
A simultaneous fit shows better agreement with the no-top fit:

- the impact of top data is diluted by the inclusion of the SMEFT

Uncertainties increase relative to the SM all top data PDF fit, reflecting the increase in fitted parameters



Simultaneous PDF-SMEFT determination



Linear SMEFT constraints are unchanged - PDF-EFT interplay is negligible in top data for **linear SMEFT fits**

We uncovered a problem with the Monte Carlo replica method used to propagate uncertainties in the SIMUnet methodology, preventing quadratic SMEFT fits - **work in progress will investigate this in more detail.**

Conclusions

Simultaneous PDF and SMEFT fits allow us to disentangle PDF-EFT interplay and assess our potential to bias and over-constrain SMEFT interpretations.

▸ this is particularly important as we move towards Run III and the HL-LHC; see *Greljo et. al 2104.02723*

- In top data, mild interplay is observed between the PDFs and SMEFT **at linear order**.
- Work in progress will assess this interplay **at quadratic order** in the SMEFT.

See *2303.06159* and <https://www.pbsp.org.uk/topproject/> for more details and results!

Conclusions

Thank you for listening!

Simultaneous PDF and SMEFT fits allow us to disentangle PDF-EFT interplay and assess our potential to bias and over-constrain SMEFT interpretations.

▸ this is particularly important as we move towards Run III and the HL-LHC; see *Greljo et. al 2104.02723*

- In top data, mild interplay is observed between the PDFs and SMEFT **at linear order**.
- Work in progress will assess this interplay **at quadratic order** in the SMEFT.

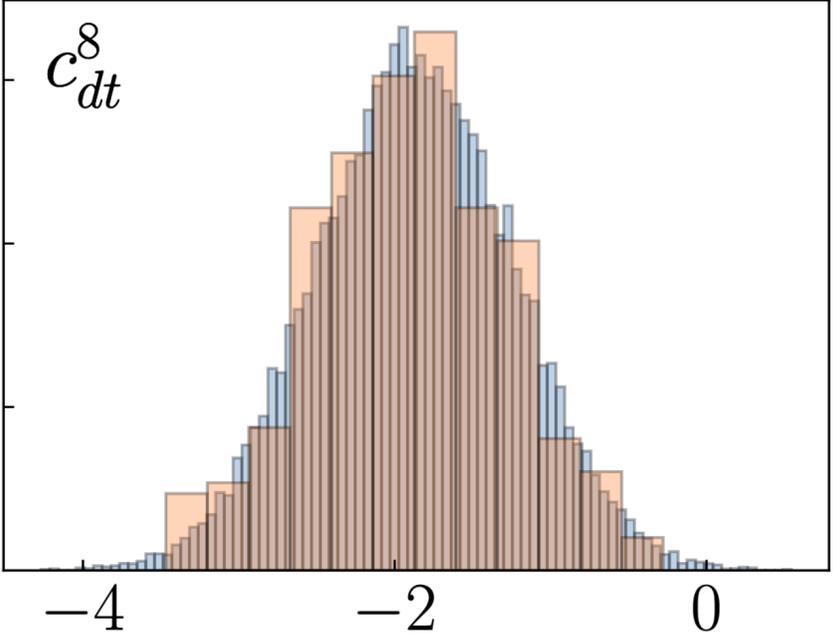
See *2303.06159* and <https://www.pbsp.org.uk/topproject/> for more details and results!

Backup

The Monte Carlo Replica Method

Consider fitting 1 Wilson coefficient c to 1 datapoint σ_{exp} : define $\chi^2 = \frac{(\sigma(c) - \sigma_{\text{exp}})^2}{\delta\sigma^2}$

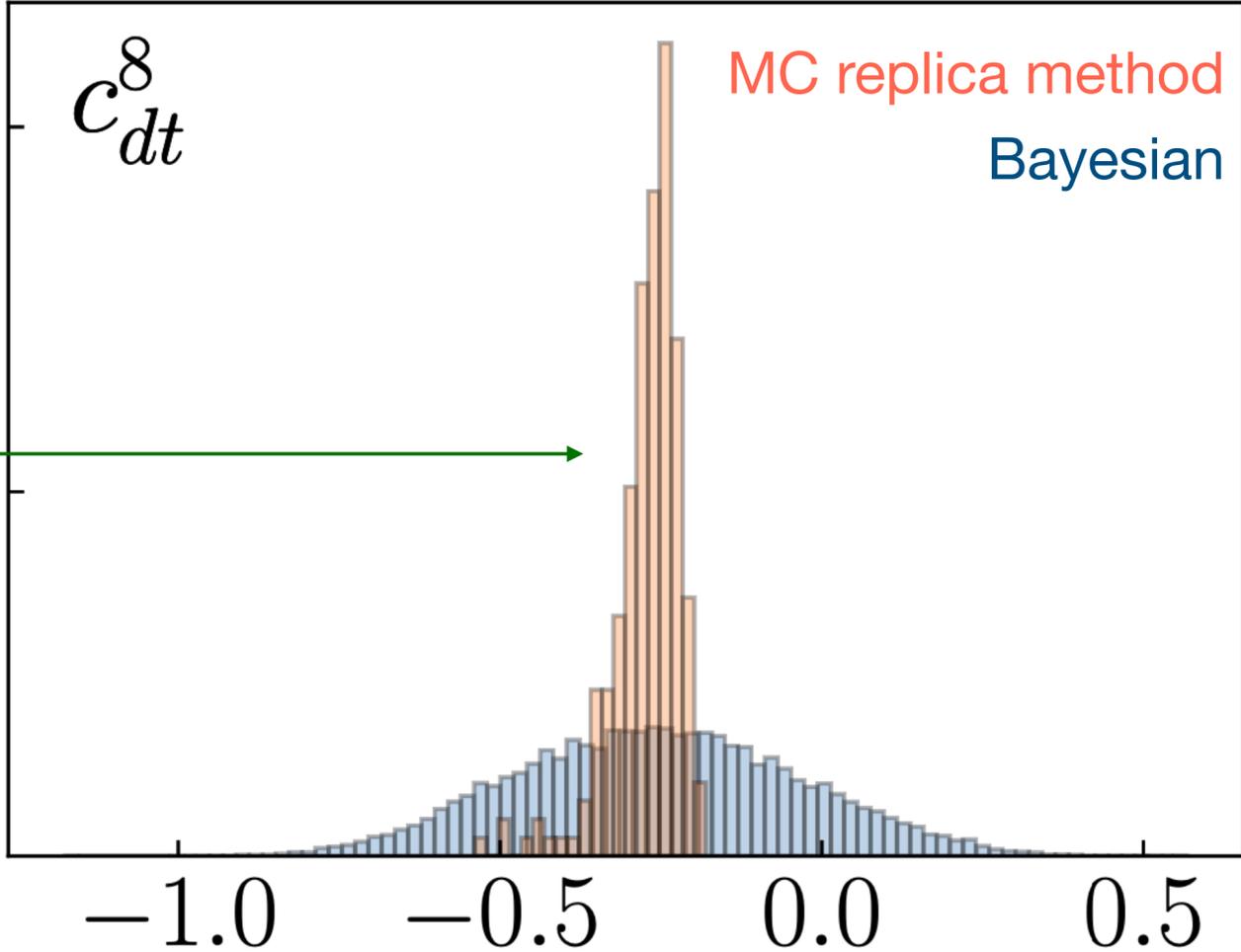
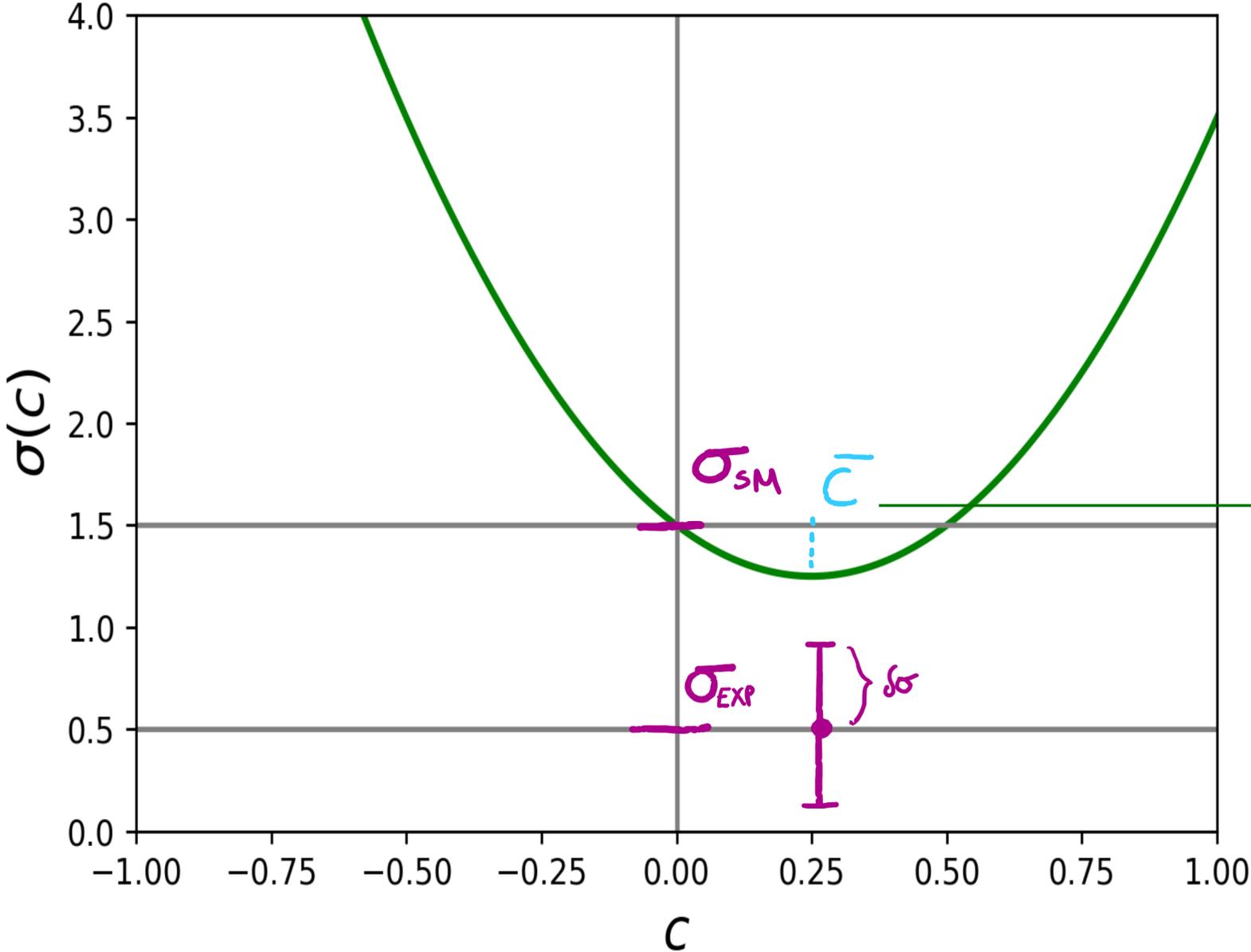
- 1. Resample: $\tilde{\sigma}_{\text{exp}} \sim \mathcal{N}(\sigma_{\text{exp}}, \Sigma)$
- 2. Minimise: $\bar{c} = \arg \min_c \frac{(\sigma(c) - \tilde{\sigma}_{\text{exp}})^2}{\delta\sigma^2}$
- 3. Treat the sample $\{\bar{c}\}$ as a sample from the Bayesian posterior $p(c|D)$



The Monte Carlo Replica Method

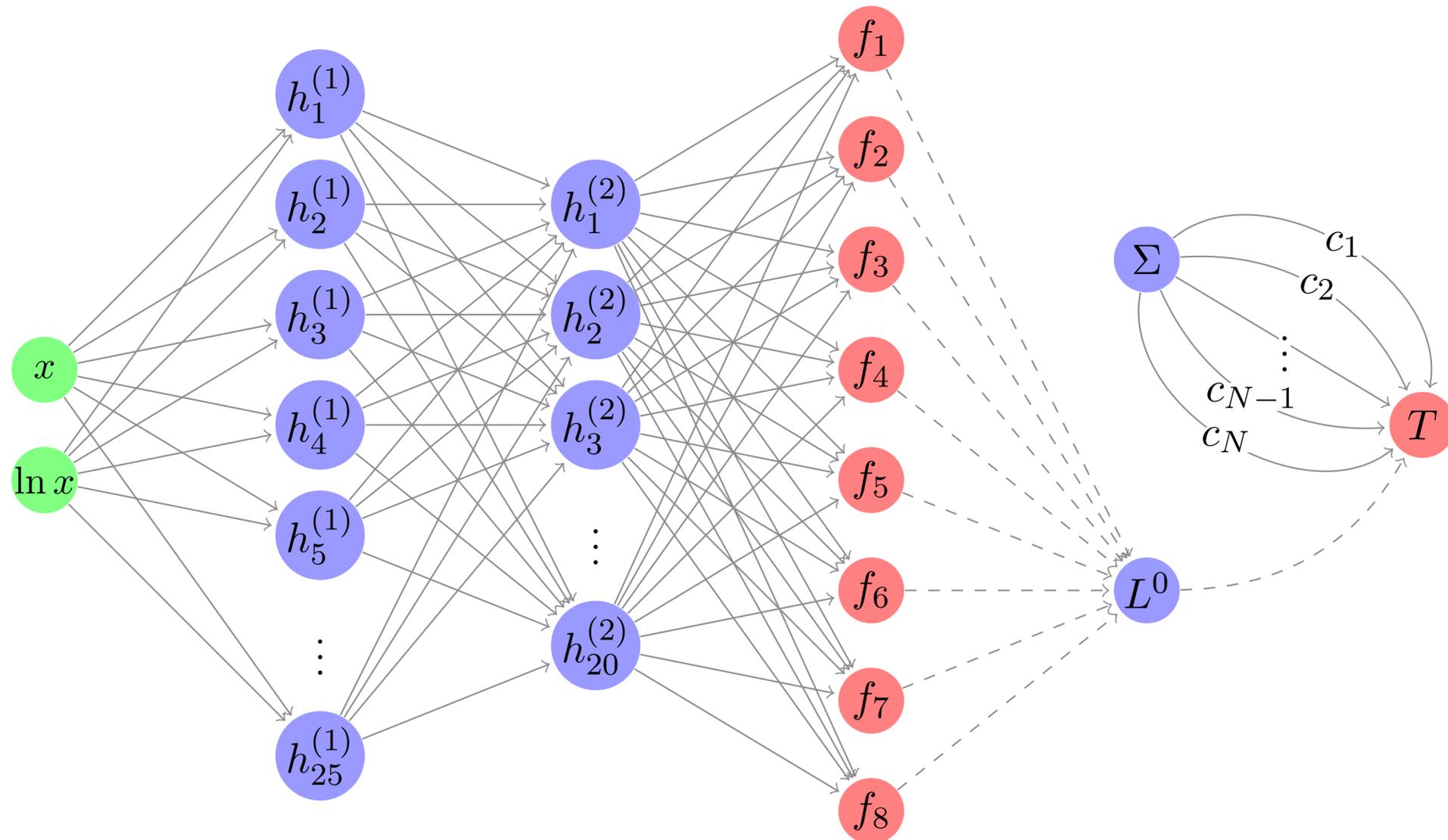
Problem: in the presence of a quadratic theory, often the minimum χ^2 will be given by the same \bar{c} .

➔ leads to an artificial spike in the distribution



PDF-EFT interplay in the top sector

Simultaneous PDF and SMEFT fits of **many Wilson coefficients** are made possible by the SIMUnet methodology *S. Iranipour, M. Ubiali, 2201.07240*



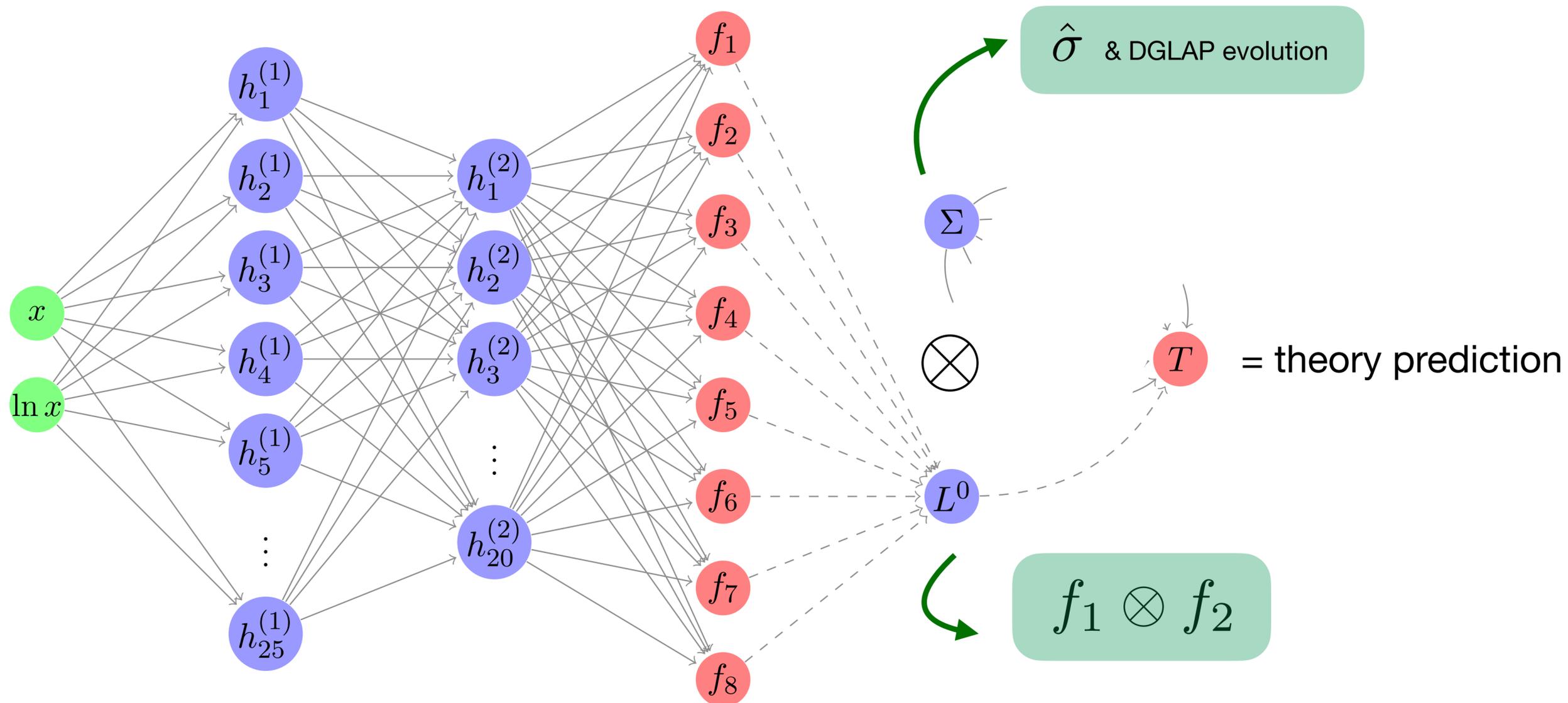
Our fit incorporates 25 operators from the top sector, following a top-specific flavour symmetry *1802.07237*

→ towards a global SMEFT & PDF fit

PDF-EFT interplay in the top sector

How does it work?

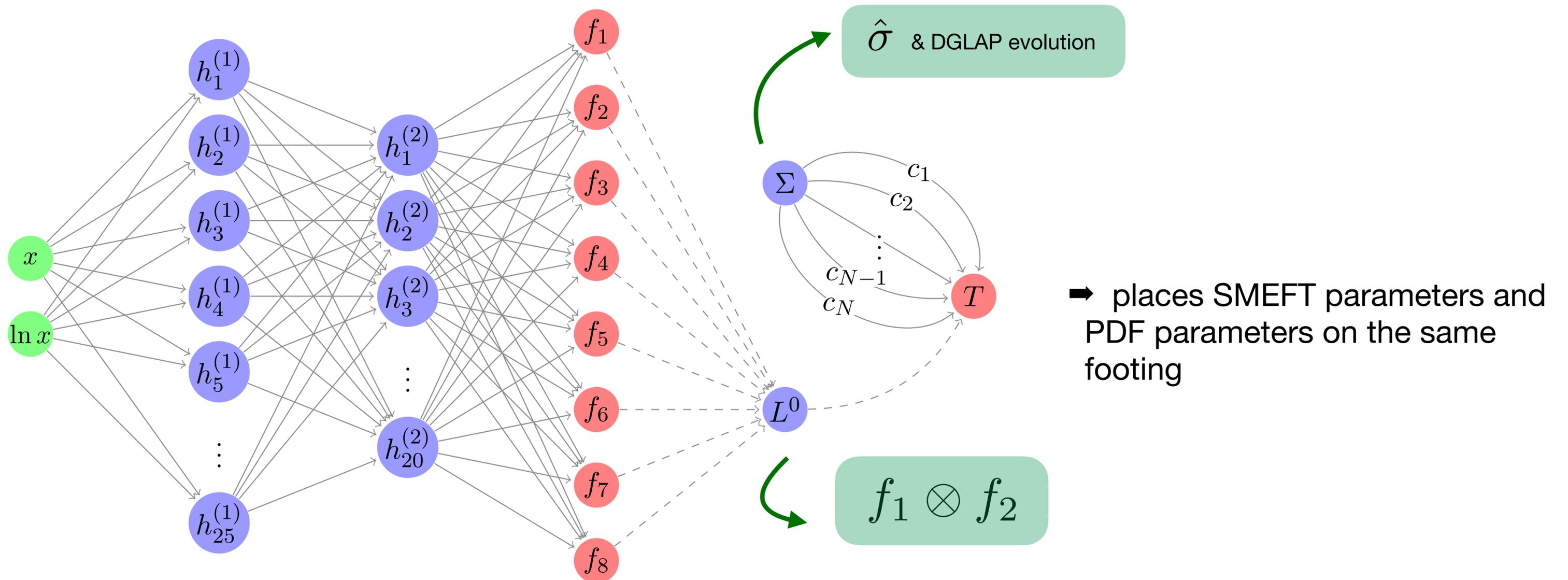
S. Iranipour, M. Ubiali, 2201.07240



PDF-EFT interplay in the top sector

How does it work?

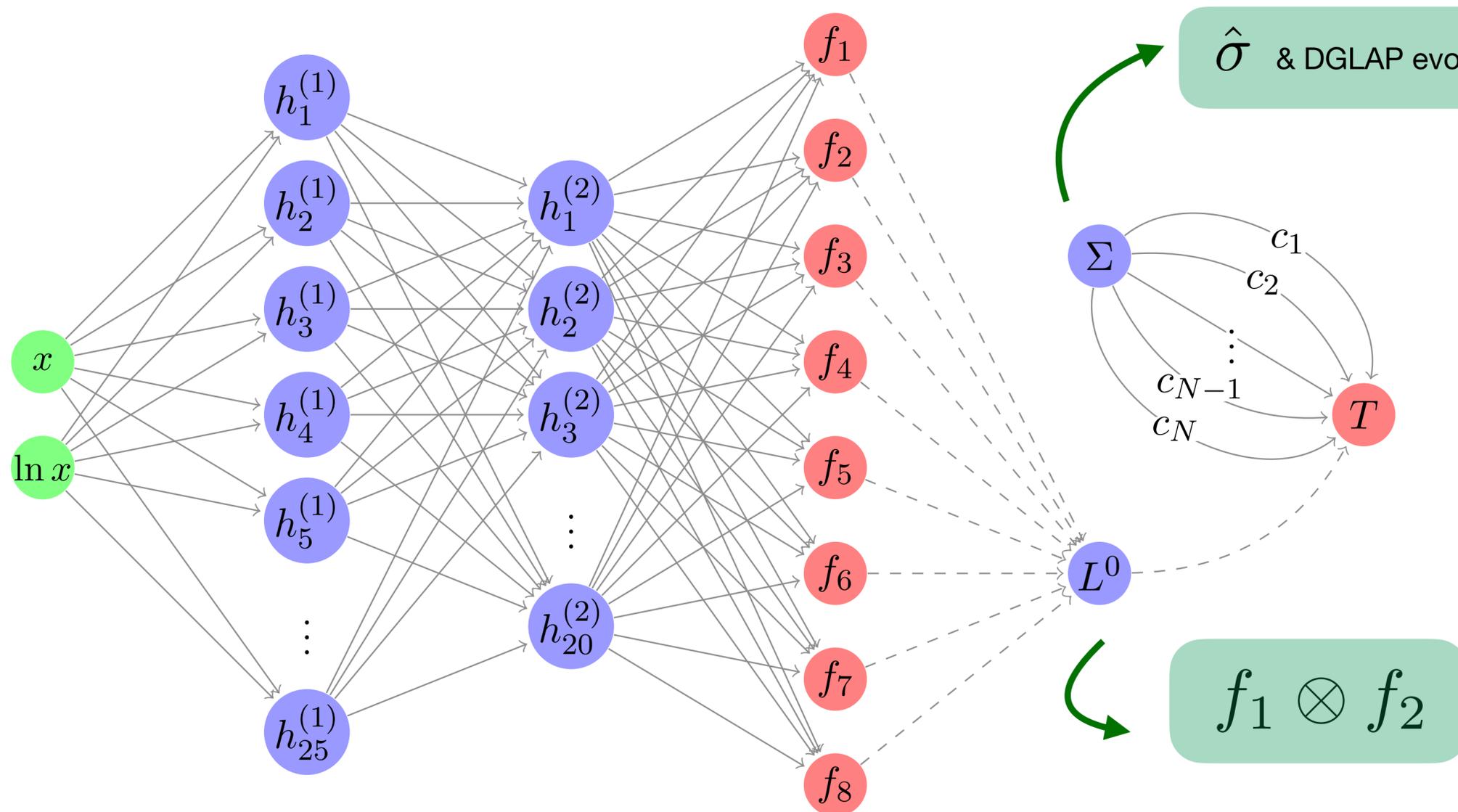
S. Iranipour, M. Ubiali, 2201.07240



PDF-EFT interplay in the top sector

How does it work?

S. Iranipour, M. Ubiali, 2201.07240

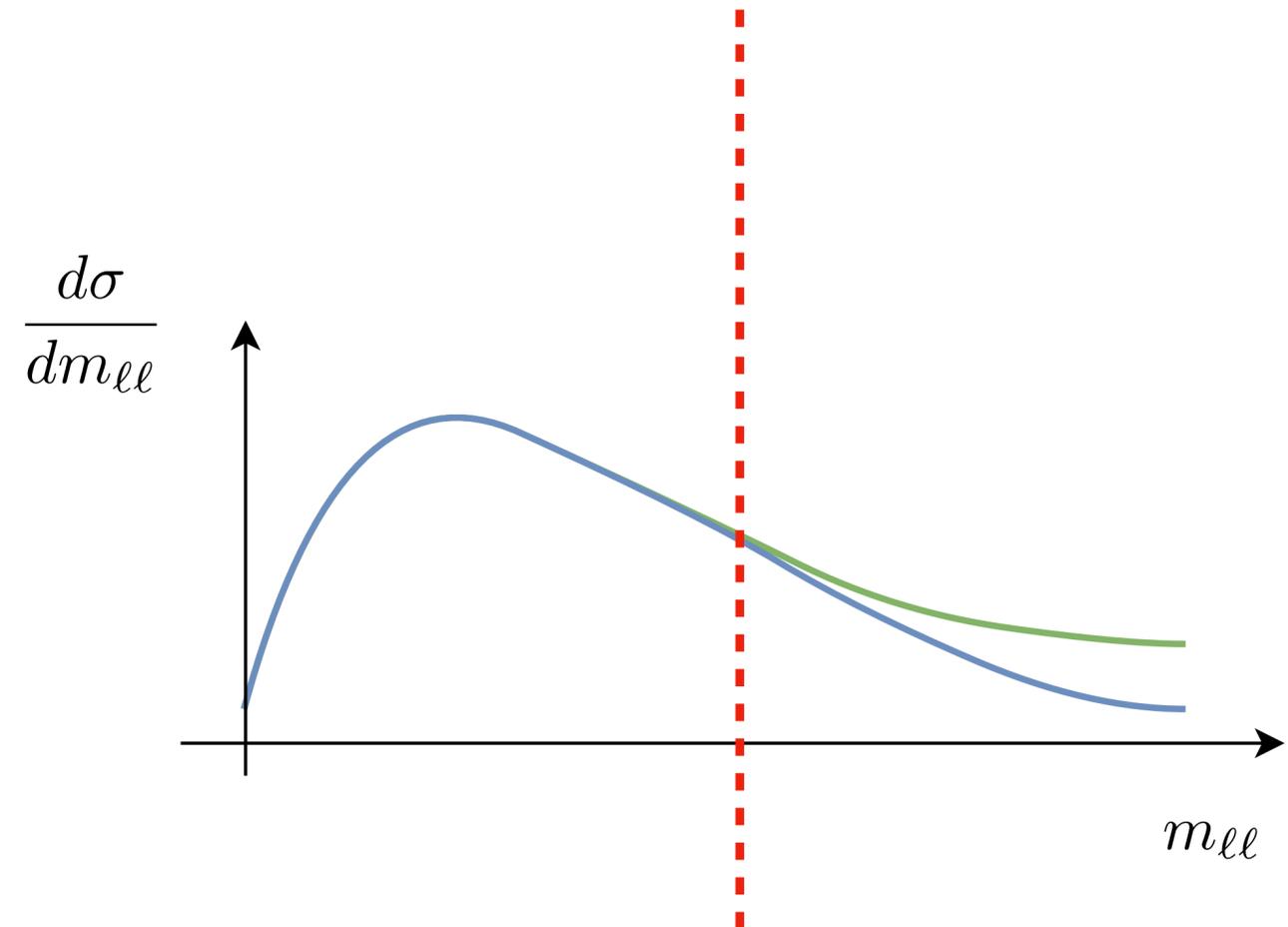


- Much more efficient than a grid scan
- Capable of handling **more SMEFT coefficients**
- Already benchmarked for the W,Y parameters in high mass DY

Conservative PDFs

Could we improve the SM PDF fits by removing the high-mass data from PDF fits?

- not in the spirit of global fits
- still have a theoretical inconsistency due to SM assumptions
- **but** much easier than doing a simultaneous PDF-SMEFT fit

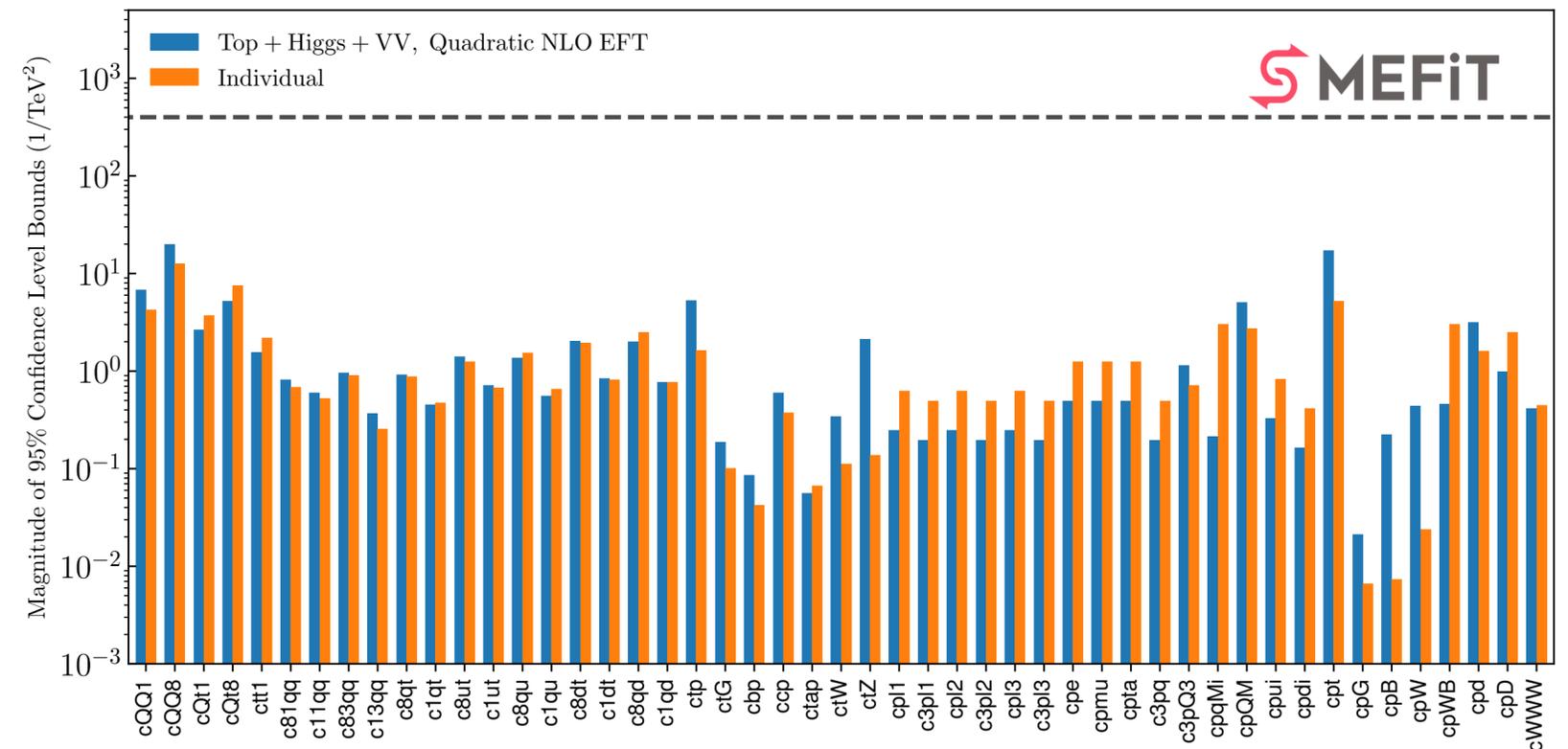
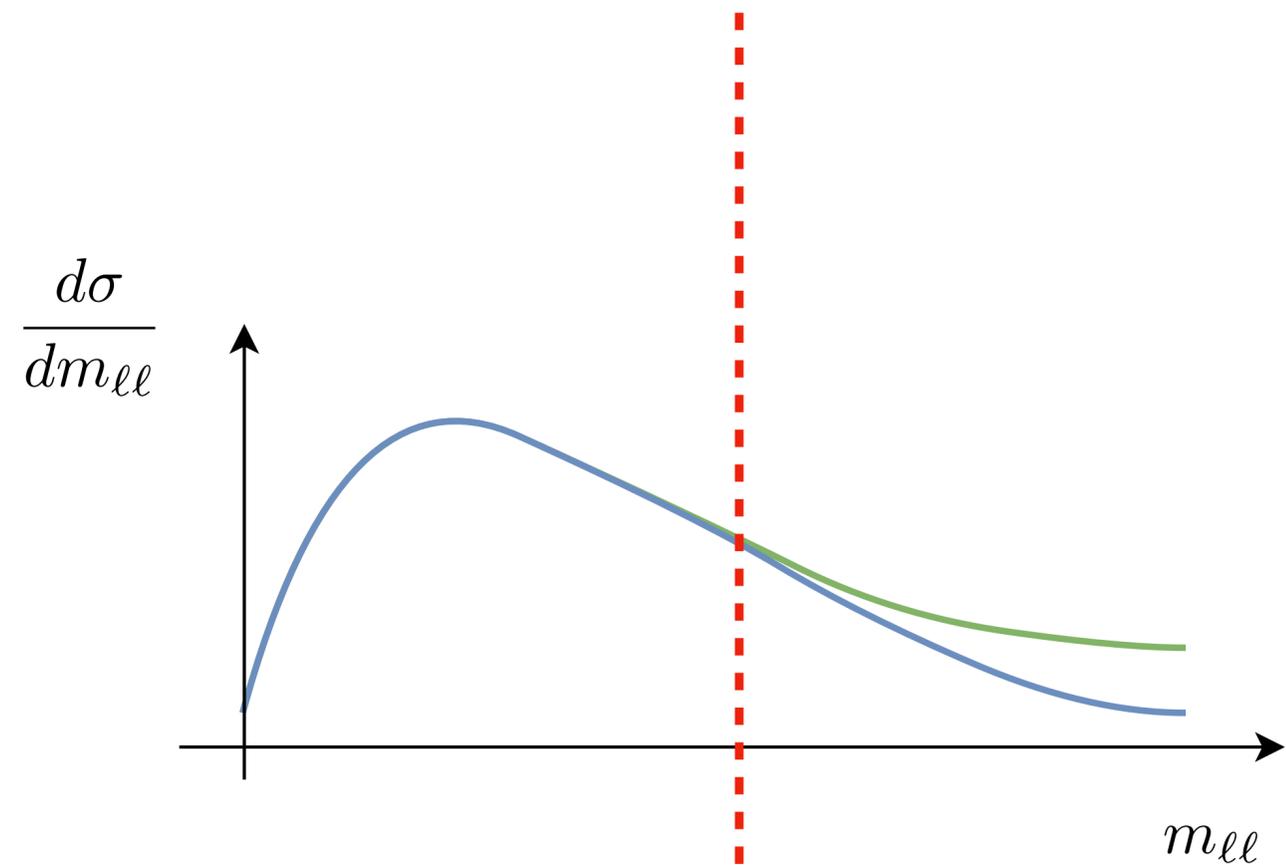


Conservative PDFs

Could we improve the SM PDF fits by removing the high-mass data from PDF fits?

- not in the spirit of global fits
- still have a theoretical inconsistency due to SM assumptions
- **but** much easier than doing a simultaneous PDF-SMEFT fit

See *J. Ethier et. al, 2105.00006* for the use of conservative PDFs in a global SMEFT fit

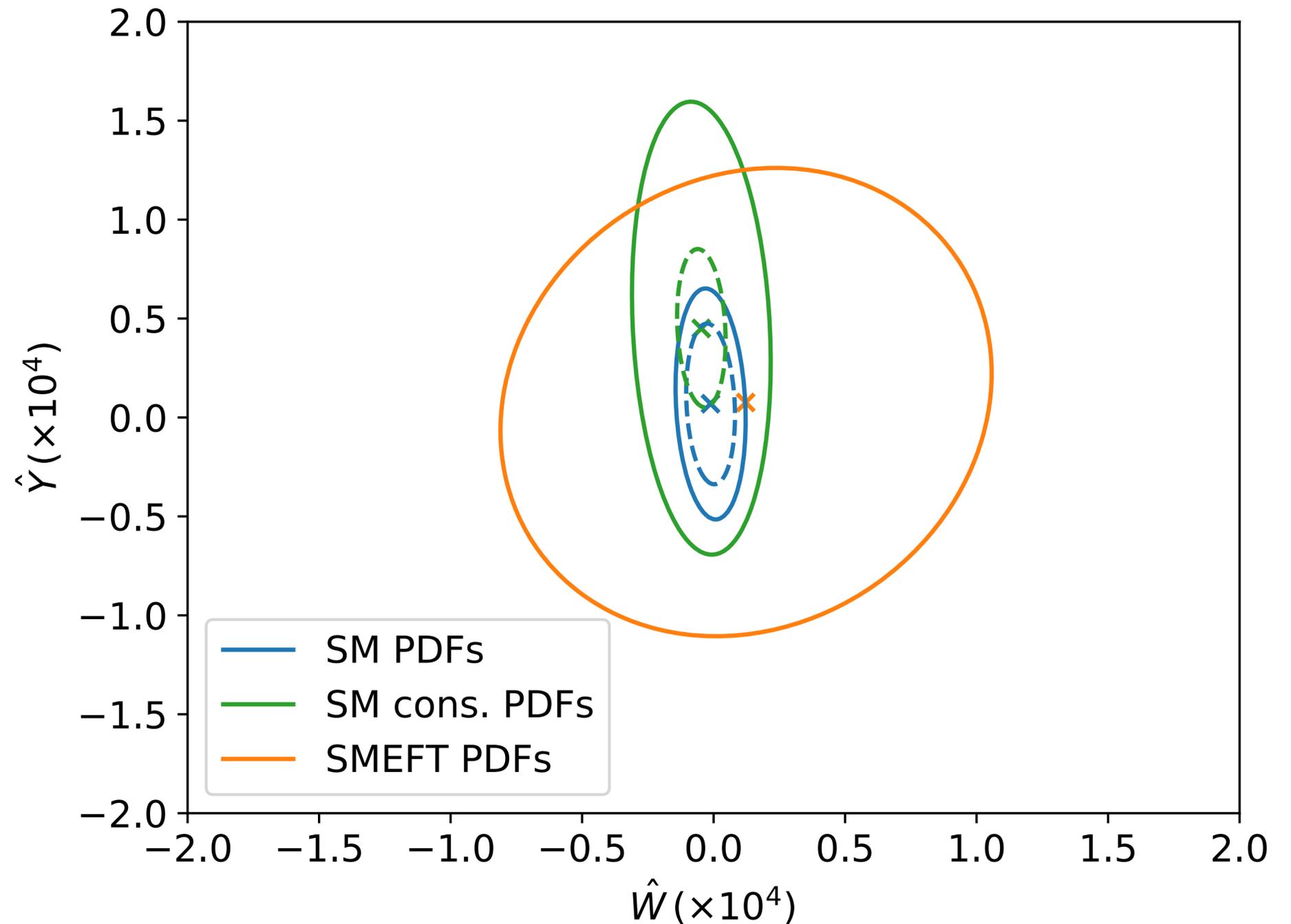


Conservative PDFs for high-mass Drell-Yan

Conservative PDFs:

- assume the SM
- are fit to data which does not receive large SMEFT corrections (i.e. no HL-LHC data, no high-mass DY data)

dashed: not including PDF uncertainties



Conservative PDFs for high-mass Drell-Yan

Conservative PDFs:

- assume the SM
- are fit to data which does not receive large SMEFT corrections (i.e. no HL-LHC data, no high-mass DY data)

Comparing green to orange:

- ▶ the constraints using SM conservative PDFs are closer to those using SMEFT PDFs
- ▶ still overestimating the constraints, especially in the \hat{W} direction

dashed: not including PDF uncertainties

