A global SMEFT analysis of LHC data

Juan Rojo, VU Amsterdam &, Nikhef

Standard Model Physics at the TeV scale

PANIC 2021, 5th September 2021

based on **arXiv:2105.00006** by J. J. Ethier, G. Magni, F. Maltoni, L. Mantani, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, and C. Zhang

Theory calculations in the SMEFT

from Lagrangian …

$$\mathscr{L}_{\text{SMEFT}} = \mathscr{L}_{\text{SM}} + \sum_{m=1}^{N_6} \frac{c_m}{\Lambda^2} \mathscr{O}_i^{(6)} + \sum_{n=1}^{N_8} \frac{b_j}{\Lambda^4} \mathscr{O}_i^{(8)} + \ldots$$

SM — **EFT_{d6}** — **EFT_{d8}**

Theory calculations in the SMEFT

from Lagrangian …

$$\mathscr{L}_{\rm SMEFT} = \mathscr{L}_{\rm SM} + \sum_{m=1}^{N_6} \frac{c_m}{\Lambda^2} \mathcal{O}_i^{(6)} + \sum_{n=1}^{N_8} \frac{b_j}{\Lambda^4} \mathcal{O}_i^{(8)} + \ldots$$

SM **EFT_{d6}** **EFT_{d8}**

Linear EFT cross-sections:

interference SM-EFT_{d6}

Quadratic EFT cross-sections:

squares EFT_{d6}

to cross-sections ….

$$\sigma_{\rm SMEFT}\left(c, \Lambda\right) \simeq \sigma_{\rm SM} \times \left(1 + \sum_{m=1}^{N_6} \frac{c_m}{\Lambda^2} \sigma_m^{\rm (eft)} + \sum_{m,n=1}^{N_6} \frac{c_m c_n}{\Lambda^4} \sigma_{m,n}^{\rm (eft)}\right)$$

evaluate at (N)NLO QCD + NLO EW

evaluate at NLO QCD
with **SMEFT@NLO**

Theory calculations in the SMEFT

… to constraints on the EFT parameters

$$\chi^2(\boldsymbol{c}, \Lambda) = \frac{1}{n_{\rm dat}} \sum_{i,j=1}^{n_{\rm dat}} \left(\sigma_{\rm i,SMEFT}(\boldsymbol{c}, \Lambda) - \sigma_{\rm i,exp} \right) \left({\rm cov}^{-1} \right)_{ij} \left(\sigma_{\rm j,SMEFT}(\boldsymbol{c}, \Lambda) - \sigma_{\rm j,exp} \right)$$

log-likelihood minimisation

to cross-sections ….

Linear EFT cross-sections: interference SM-EFT$_{\rm d6}$

Quadratic EFT cross-sections: squares EFT$_{\rm d6}$

$$\sigma_{\rm SMEFT}\left(\boldsymbol{c}, \boldsymbol{\Lambda}\right) \simeq \sigma_{\rm SM} \times \left(1 + \sum_{m=1}^{N_6} \frac{c_m}{\Lambda^2} \sigma_m^{\rm (eft)} + \sum_{m,n=1}^{N_6} \frac{c_m c_n}{\Lambda^4} \sigma_{m,n}^{\rm (eft)} \right)$$

evaluate at (N)NLO QCD + NLO EW

evaluate at NLO QCD with **SMEFT@NLO**

The SMEFiT framework

SMEFiT

Search docs

OVERVIEW:

- Features
- Available Datasets

THEORY:

- SMEFT
- References

TALKS AND LECTURES:

- Talks and seminars

IMPLEMENTATION:

- Fitting strategies
- Nested Sampling
- MCFit

RESULTS:

- SMEFiT Top
- SMEFiT RW
- SMEFiT VBS
- SMEFiT2.0

SMEFiT

Welcome to the SMEFiT website!

SMEFiT is a Python package for global analyses of particle physics data in the framework of the Standard Model Effective Field Theory (SMEFT). The SMEFT represents a powerful model-independent framework to constrain, identify, and parametrise potential deviations with respect to the predictions of the Standard Model (SM). A particularly attractive feature of the SMEFT is its capability to systematically correlate deviations from the SM between different processes. The full exploitation of the SMEFT potential for indirect New Physics searches from precision measurements requires combining the information provided by the broadest possible dataset, namely carrying out extensive global analysis which is the main purpose of SMEFiT.

Project description

The SMEFiT framework has been used in the following **scientific publications**:

- *A Monte Carlo global analysis of the Standard Model Effective Field Theory: the top quark sector*, N. P. Hartland, F. Maltoni, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, C. Zhang [HMN+19].
- *Constraining the SMEFT with Bayesian reweighting*, S. van Beek, E. R. Nocera, J. Rojo, and E. Slade [vBNRS19].
- *SMEFT analysis of vector boson scattering and diboson data from the LHC Run II* , J. Ethier, R. Gomez-Ambrosio, G. Magni, J. Rojo [EGAMR21].
- *Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC*, J. Ethier, F. Maltoni, L. Mantani, E. R. Nocera, J. Rojo, E. Slade, E. Vryonidou, C. Zhang [EMM+21] **arXiv:2105.00006**

Results from these publications, including driver and analysis scripts, are available in the **Results** section.

Team description

The **SMEFiT collaboration** is currently composed by the following members:

- Jaco ter Hoeve, *VU Amsterdam and Nikhef Theory Group*
- Giacomo Magni, *VU Amsterdam and Nikhef Theory Group*
- Fabio Maltoni, *Centre for Cosmology, Particle Physics and Phenomenology Louvain and University of Bologna*
- Luca Mantani, *Centre for Cosmology, Particle Physics and Phenomenology Louvain*
- Emanuele Roberto Nocera, *Higgs Center for Theoretical Physics, University of Edinburgh*
- Juan Rojo, *VU Amsterdam and Nikhef Theory Group*
- Eleni Vryonidou, *University of Manchester*

https://lhcfitnikhef.github.io/SMEFT/

Quantifying EFT sensitivity

Quantify impact in fit using **information geometry** (Fisher discriminant)

linear

$$I_{ij} = \sum_{m=1}^{n_{\rm dat}} \frac{\sigma^{(\rm eft)}_{m,i}\sigma^{(\rm eft)}_{m,j}}{\delta^2_{{\rm exp},m}}$$

n.b. operator normalisation is arbitrary, thus absolute values of Fisher unphysical

normalise to the sum over a given operator: relative Fisher is physical

quadratic

$$I_{ij} = {\rm E}\left[\sum_{m=1}^{n_{\rm dat}} \frac{1}{\delta^2_{{\rm exp},m}} \left(\sigma_{m,ij}\left(\sigma^{(\rm th)}_m - \sigma^{(\rm exp)}_m\right) + \left(\sigma^{(\rm eft)}_{m,i} + \sum_{l=1}^{n_{\rm op}} c_l \sigma^{(\rm eft)}_{m,il}\right)\left(\sigma^{(\rm eft)}_{m,j} + \sum_{l'=1}^{n_{\rm op}} c_{l'}\sigma_{m,jl'}\right)\right)\right]$$

Determine **most sensitive directions** and identify possible flat directions using Principal Component Analysis (PVA) & Singular Value Decomposition (SVD)

$$\sigma^{(\rm th)}_m(\boldsymbol{c}) = \sigma^{(\rm sm)}_m + \sum_{i=1}^{n_{\rm op}} c_i \sigma^{(\rm eft)}_{m,i}$$

$$K = UWV^\dagger$$

singular value decomposition

$$K_{mi} = \sigma^{(\rm eft)}_{m,i}/\delta_{\rm exp,m;}$$

$${\rm PC}_k = \sum_{i=1}^{n_{\rm op}} a_{ki}c_i\,, \quad k = 1,\ldots,n_{\rm op}\,, \qquad \left(\sum_{i=1}^{n_{\rm op}} a^2_{ki} = 1 \;\; \forall k\right)$$

n.b. within our approach flat directions are not a problem, and can also be identified *a posteriori*

Operator basis and flavour assumptions

Class	$N_{\rm dof}$	Independent DOFs	DoF in EWPOs
four-quark (two-light-two-heavy)	14	$c_{Qq}^{1,8}$, $c_{Qq}^{1,1}$, $c_{Qq}^{3,8}$, $c_{Qq}^{3,1}$, c_{tq}^{8}, c_{tq}^{1}, c_{tu}^{8}, c_{tu}^{1}, c_{Qu}^{8}, c_{Qu}^{1}, c_{td}^{8}, c_{td}^{1}, c_{Qd}^{8}, c_{Qd}^{1}	
four-quark (four-heavy)	5	c_{QQ}^{1}, c_{QQ}^{8}, c_{Qt}^{1}, c_{Qt}^{8}, c_{tt}^{1}	
four-lepton	1		$c_{\ell\ell}$
two-fermion (+ bosonic fields)	23	$c_{t\varphi}$, c_{tG}, $c_{b\varphi}$, $c_{c\varphi}$, $c_{\tau\varphi}$, c_{tW}, c_{tZ}, $c_{\varphi Q}^{(3)}$, $c_{\varphi Q}^{(-)}$, $c_{\varphi t}$	$c_{\varphi\ell_1}^{(1)}$, $c_{\varphi\ell_1}^{(3)}$, $c_{\varphi\ell_2}^{(1)}$, $c_{\varphi\ell_2}^{(3)}$, $c_{\varphi\ell_3}^{(1)}$, $c_{\varphi\ell_3}^{(3)}$, $c_{\varphi e}$, $c_{\varphi\mu}$, $c_{\varphi\tau}$, $c_{\varphi q}^{(3)}$, $c_{\varphi q}^{(-)}$, $c_{\varphi u}$, $c_{\varphi d}$
Purely bosonic	7	$c_{\varphi G}$, $c_{\varphi B}$, $c_{\varphi W}$, $c_{\varphi d}$, c_{WWW}	$c_{\varphi WB}$, $c_{\varphi D}$
Total	50 (36 independent)	34	16 (2 independent)

- **Dim-6 SMEFT operators** modifying Higgs, dibosons, and top quark properties: **36** (14) **independent** (dependent) **DoFs**

- Flavour assumption is **MFV**, with $\mathrm{U}(2)_q \times \mathrm{U}(2)_u \times \mathrm{U}(3)_d$ in quark sector (special role for top quark) and $\left(\mathrm{U}(1)_\ell \times \mathrm{U}(1)_e\right)^3$ in lepton sector

- Constraints from **LEP EWPOs** imposed via restrictions in parameter space

$$\begin{pmatrix} c_{\varphi\ell_i}^{(3)} \\ c_{\varphi\ell_i}^{(1)} \\ c_{\varphi e/\mu/\tau} \\ c_{\varphi q}^{(-)} \\ c_{\varphi q}^{(3)} \\ c_{\varphi u} \\ c_{\varphi d} \\ c_{\ell\ell} \end{pmatrix} = \begin{pmatrix} -\frac{1}{t_W} & -\frac{1}{4t_W^2} \\ 0 & -\frac{1}{4} \\ 0 & -\frac{1}{2} \\ \frac{1}{t_W} & \frac{1}{4s_W^2}-\frac{1}{6} \\ -\frac{1}{t_W} & -\frac{1}{4t_W^2} \\ 0 & \frac{1}{3} \\ 0 & -\frac{1}{6} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_{\varphi WB} \\ c_{\varphi D} \end{pmatrix}$$

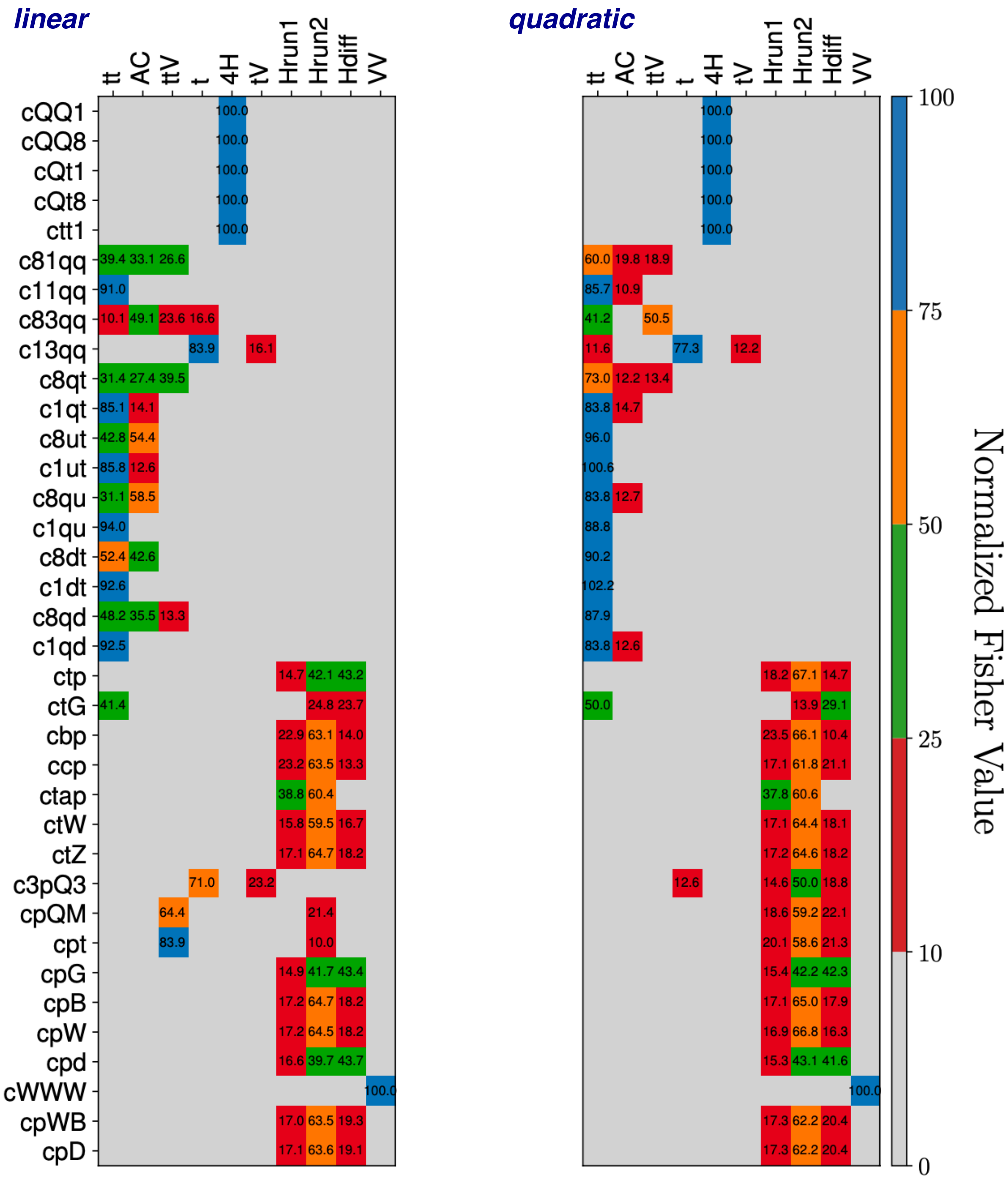
Experimental data

Category	Processes	$n_{\rm dat}$
Top quark production	$t\bar{t}$ (inclusive) **(incl LHC charge asy)**	94
	$t\bar{t}Z$, $t\bar{t}W$ **(incl ptZ in ttZ)**	14
	single top (inclusive)	27
	tZ, tW	9
	$t\bar{t}t\bar{t}$, $t\bar{t}b\bar{b}$	6
	Total	**150**
Higgs production and decay	Run I signal strengths	22
	Run II signal strengths	40
	Run II, differential distributions & STXS	35
	Total	**97**
Diboson production	LEP-2 **(WW)**	40
	LHC **(WW & WZ)**	30
	Total	**70**
Baseline dataset	**Total**	**317**

+ systematic assessment of fit results **wrt dataset variations**:

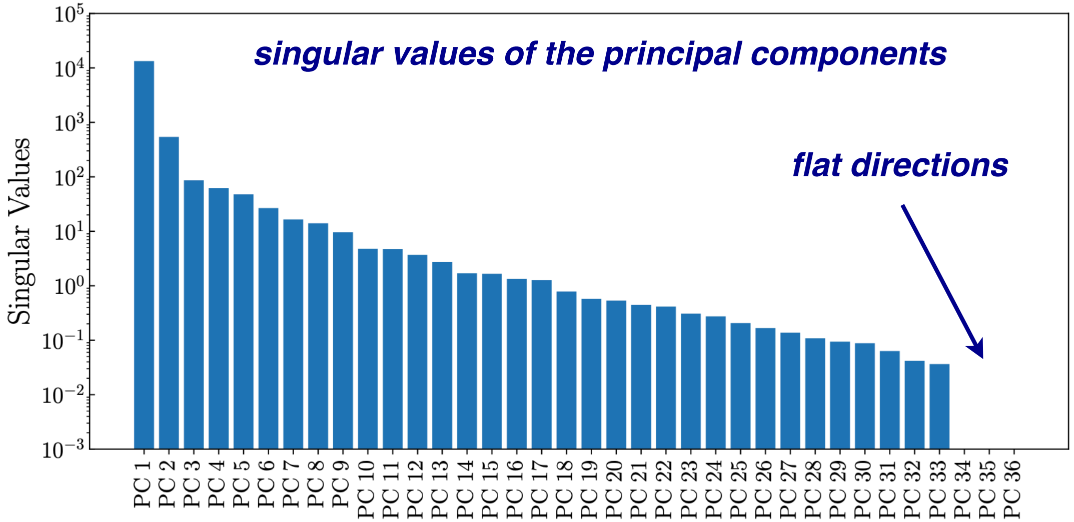
Higgs-only fit, top-only fit, no high-E data, no diboson data …

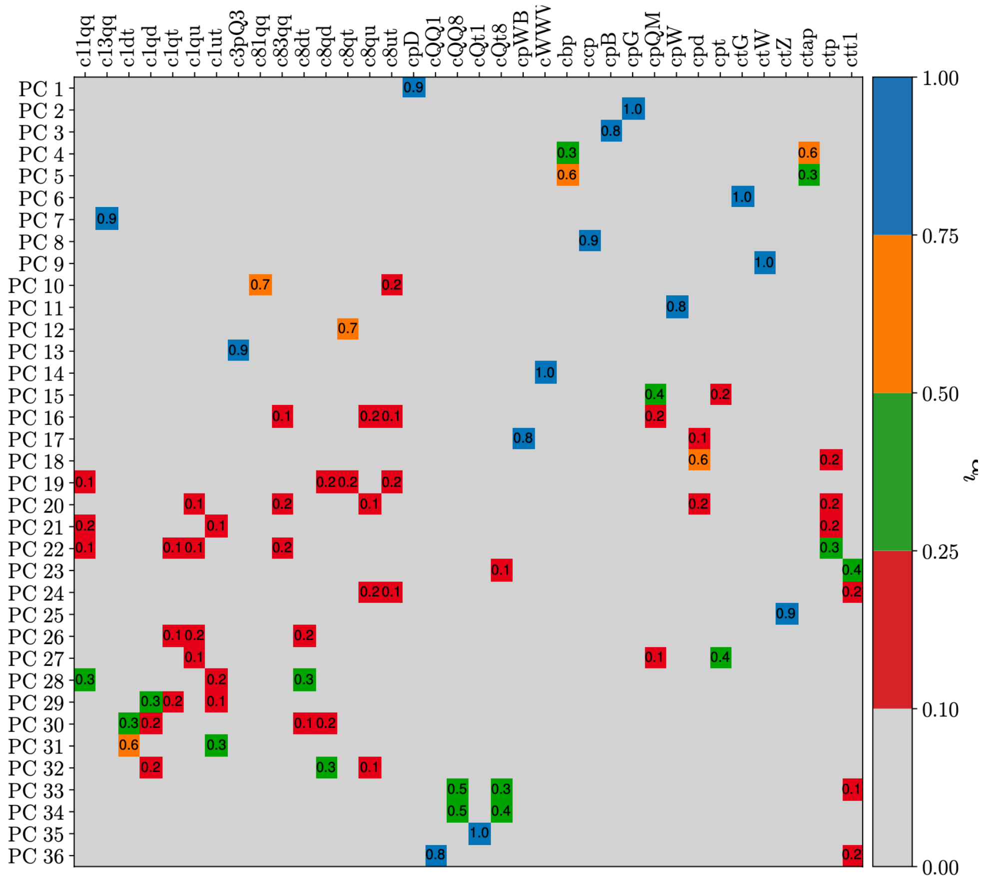
Quantifying EFT sensitivity



- Compare **relative impact of each process** on a given EFT coefficient
- Four-fermion operators constrained (mostly) by top data, two-fermion and purely bosonic (mostly) by Higgs
- **Sensitivity** depends on linear vs quadratic, but also LO vs NLO EFT
- Can be used at a finer level, *e.g.* identify **which differential distribution** of a given measurement carries more weight in the EFT fit

Quantifying EFT sensitivity





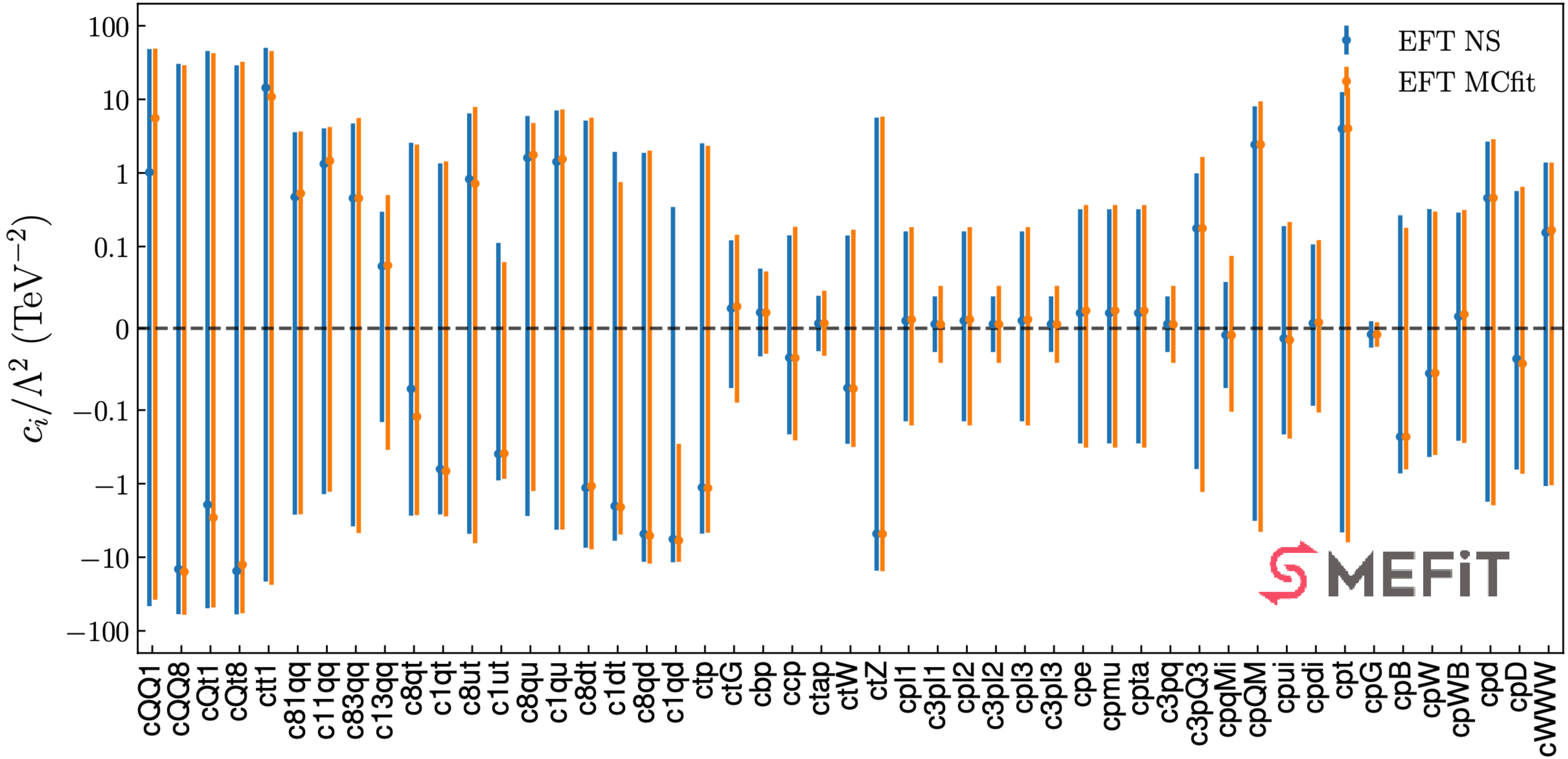
- Identify **flat directions** (in linear EFT fit) and which coefficient combinations have the higher variance

- Determine which coefficients are determined by one or a few processes, and which ones only enter at the level of linear combinations of many coefficients

- Some EFT parameters represent **``natural directions''**, other always appear in combination with several other coefficients

- Powerful tool to understand fit results, eventually could be used to **fit in the PCA basis** (though this is not required)

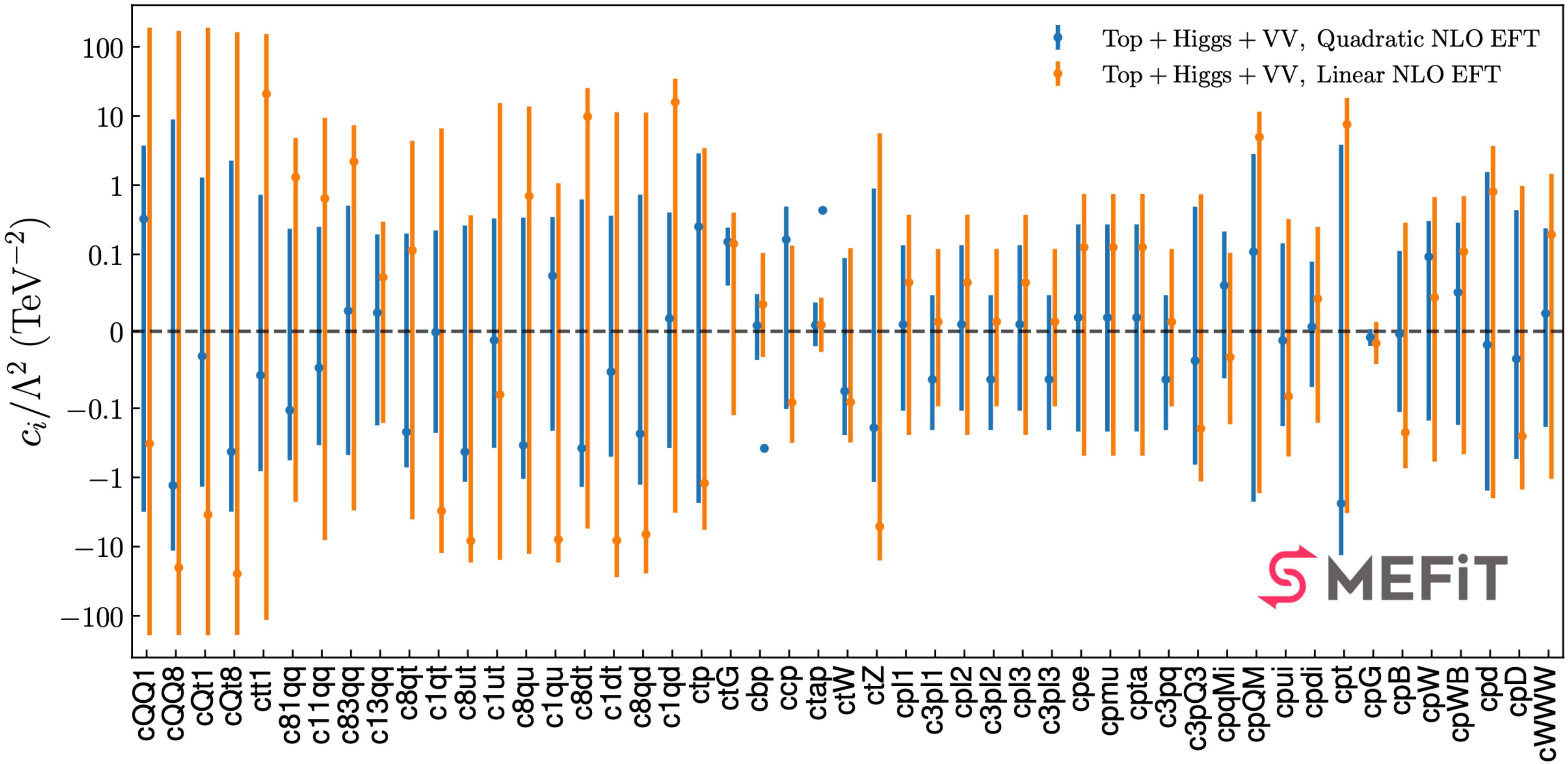
Fitting methodology



Median and 95% CL intervals for the **50 EFT parameters** considered in this analysis in linear fit

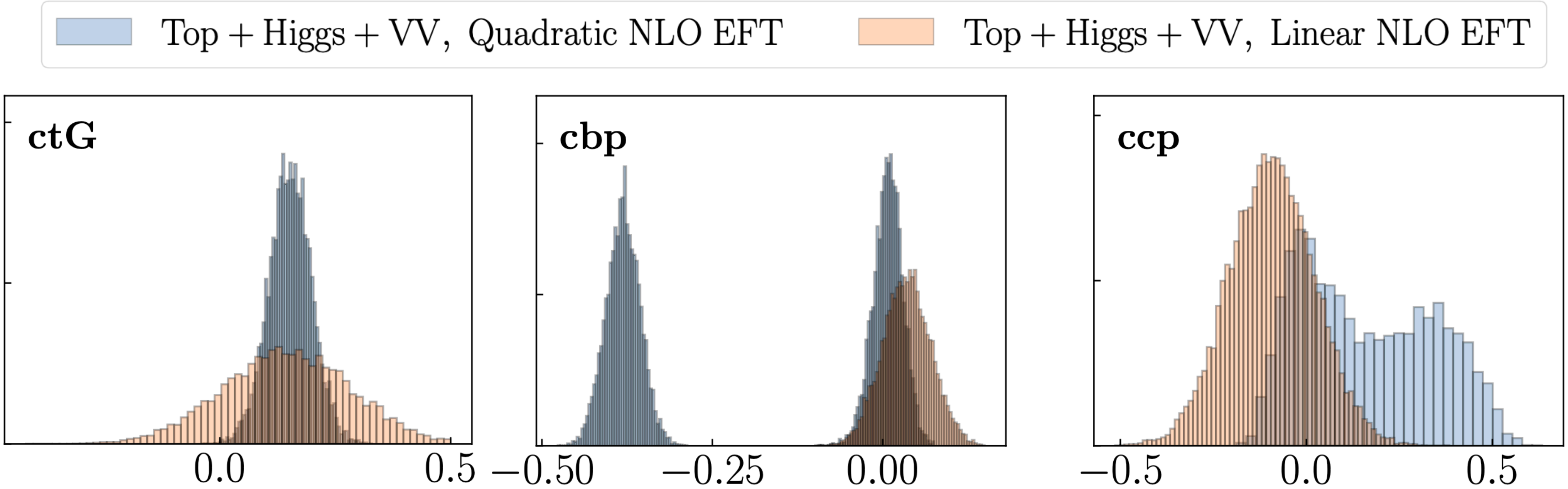
Equivalent results obtained with **MCfit** and **NS**: mutual validation of fit outcome

Results: global fit



- Agreement with SM at 95% CL for all EFT coefficients except for **ctG** in quadratic fit
- Quadratic corrections bring in sensitivity (more stringent bounds) *e.g.* for four-fermion operators
- Some DoFs exhibit a second ``BSM-like" solution in the quartic fit

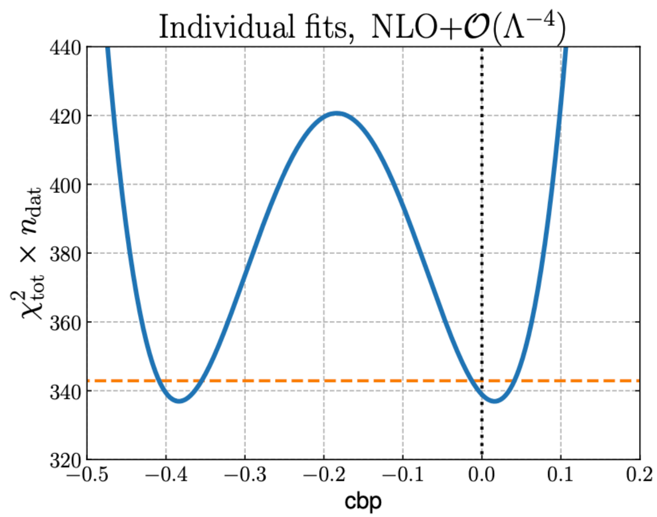
Results: global fit

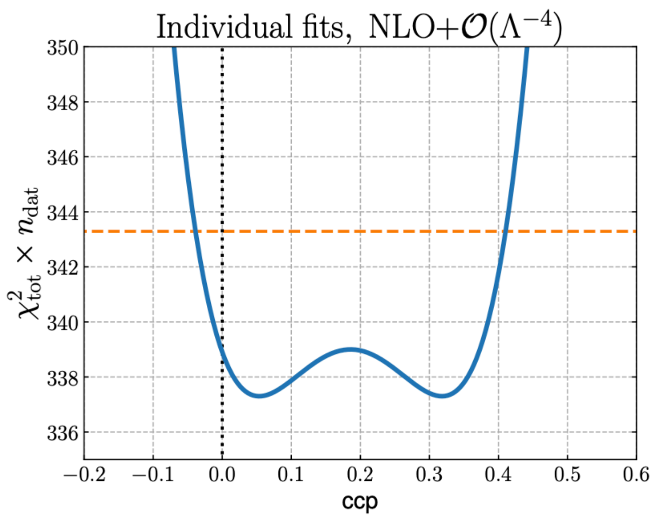


chromo-magnetic operator ***bottom Yukawa*** ***charm Yukawa***

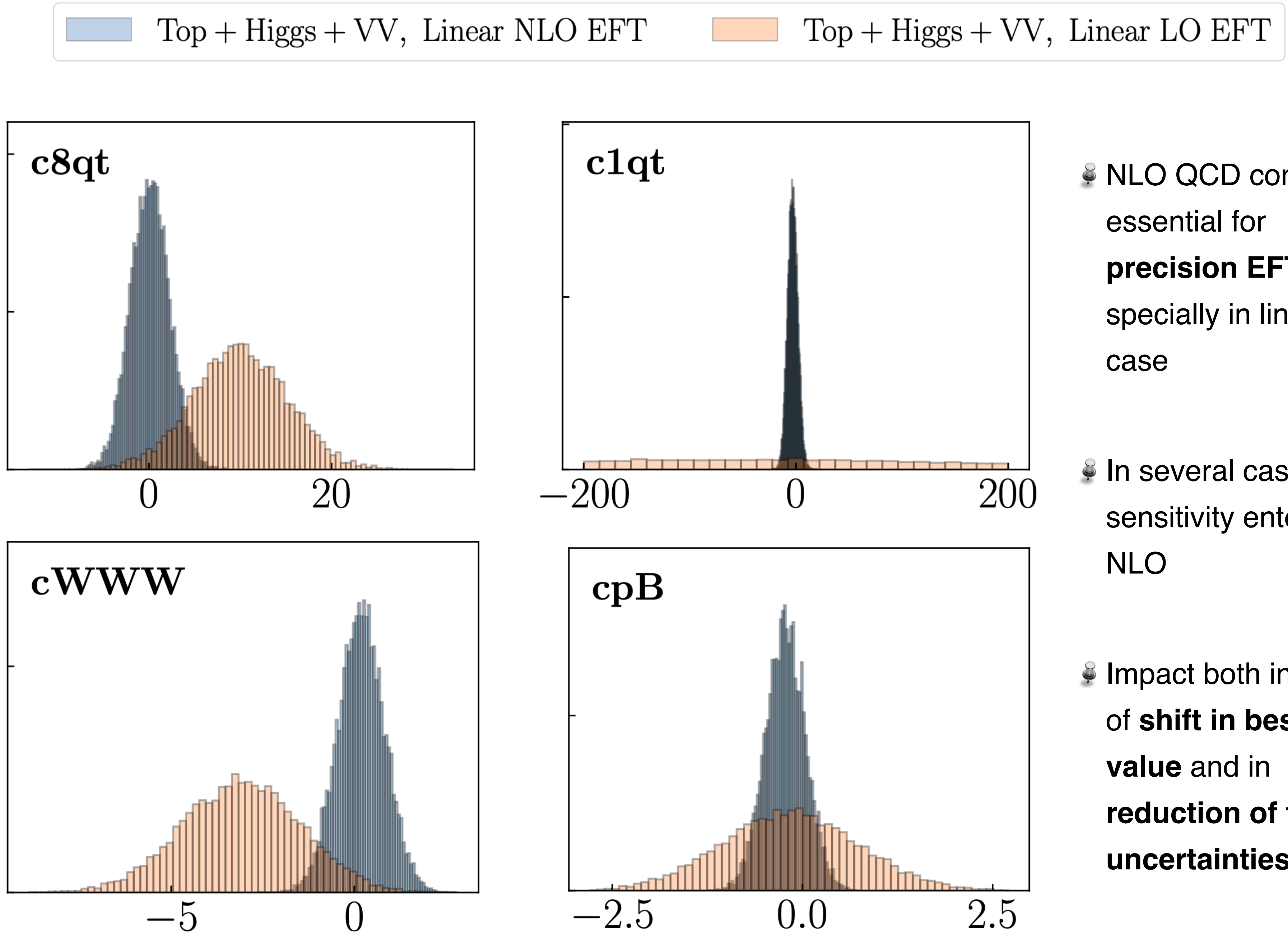
in general, sensitivity of fit results to inclusion of **quadratic EFT corrections**

1-parameter fits



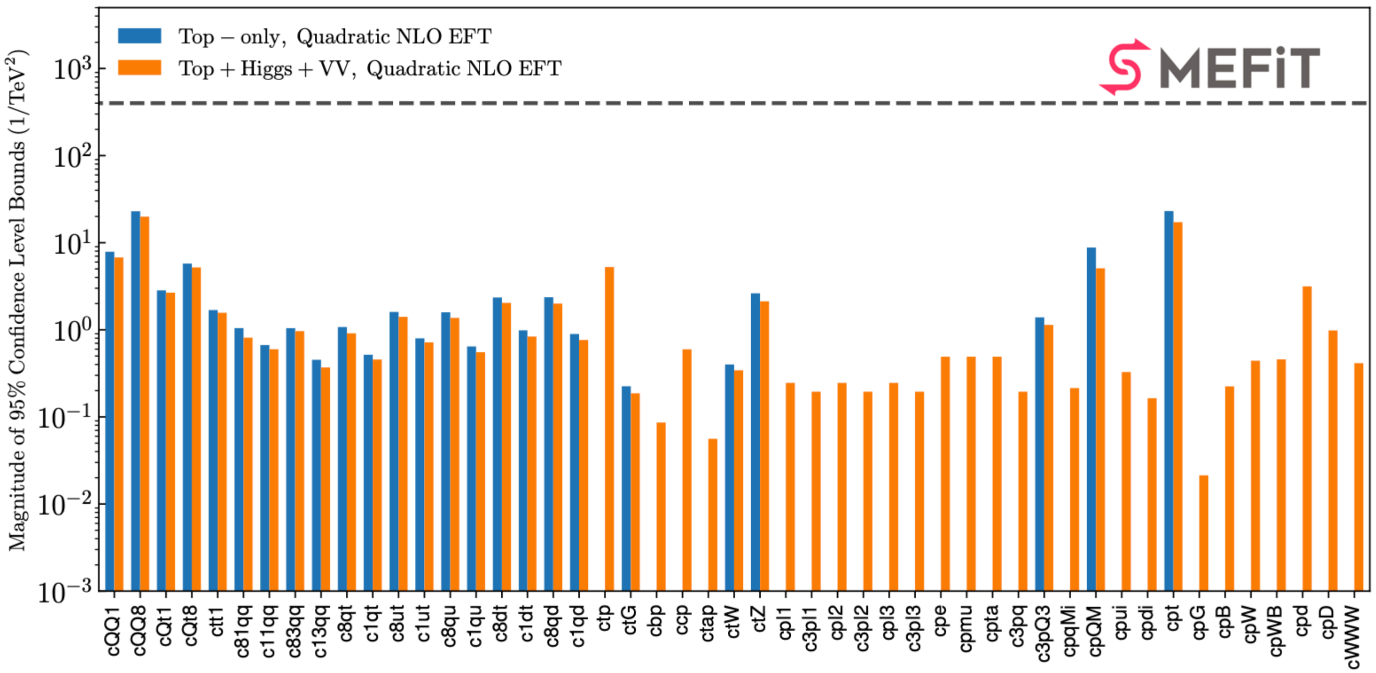


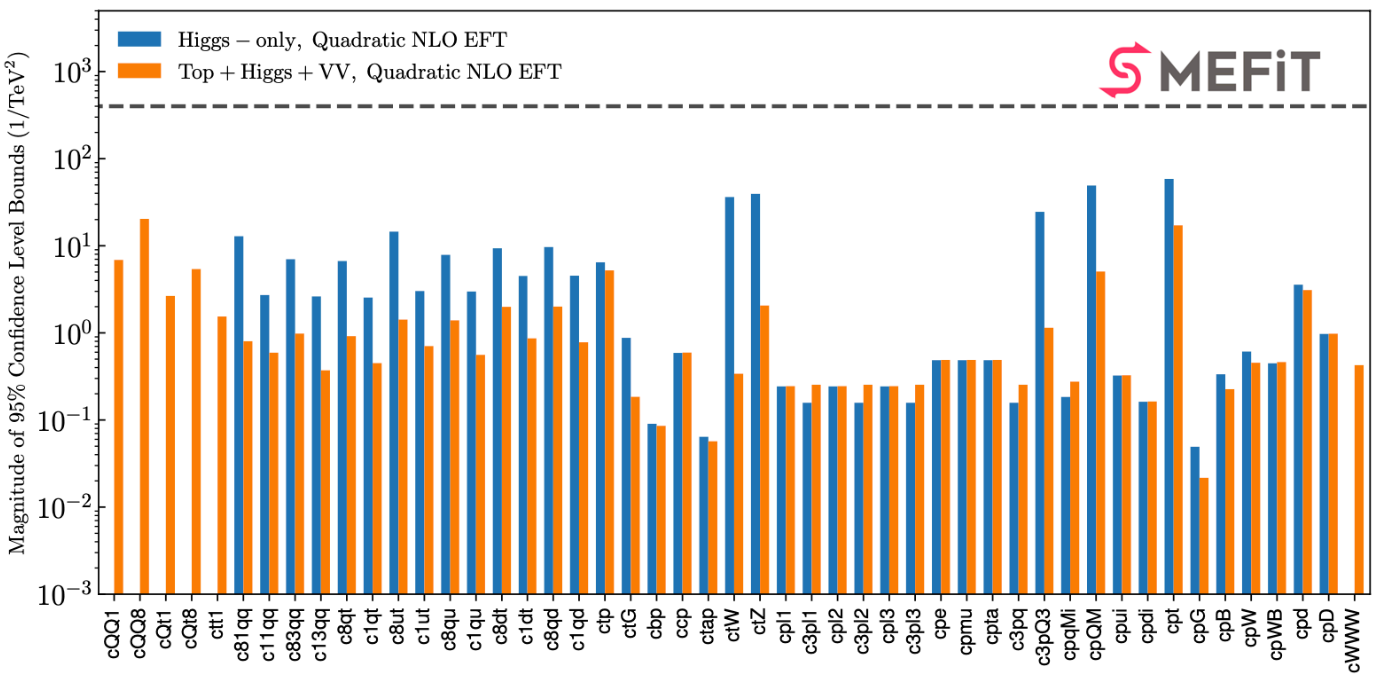
Results: impact of NLO corrections



- NLO QCD corrections essential for **precision EFT fits**, specially in linear case
- In several cases new sensitivity enters at NLO
- Impact both in terms of **shift in best-fit value** and in **reduction of fit uncertainties**

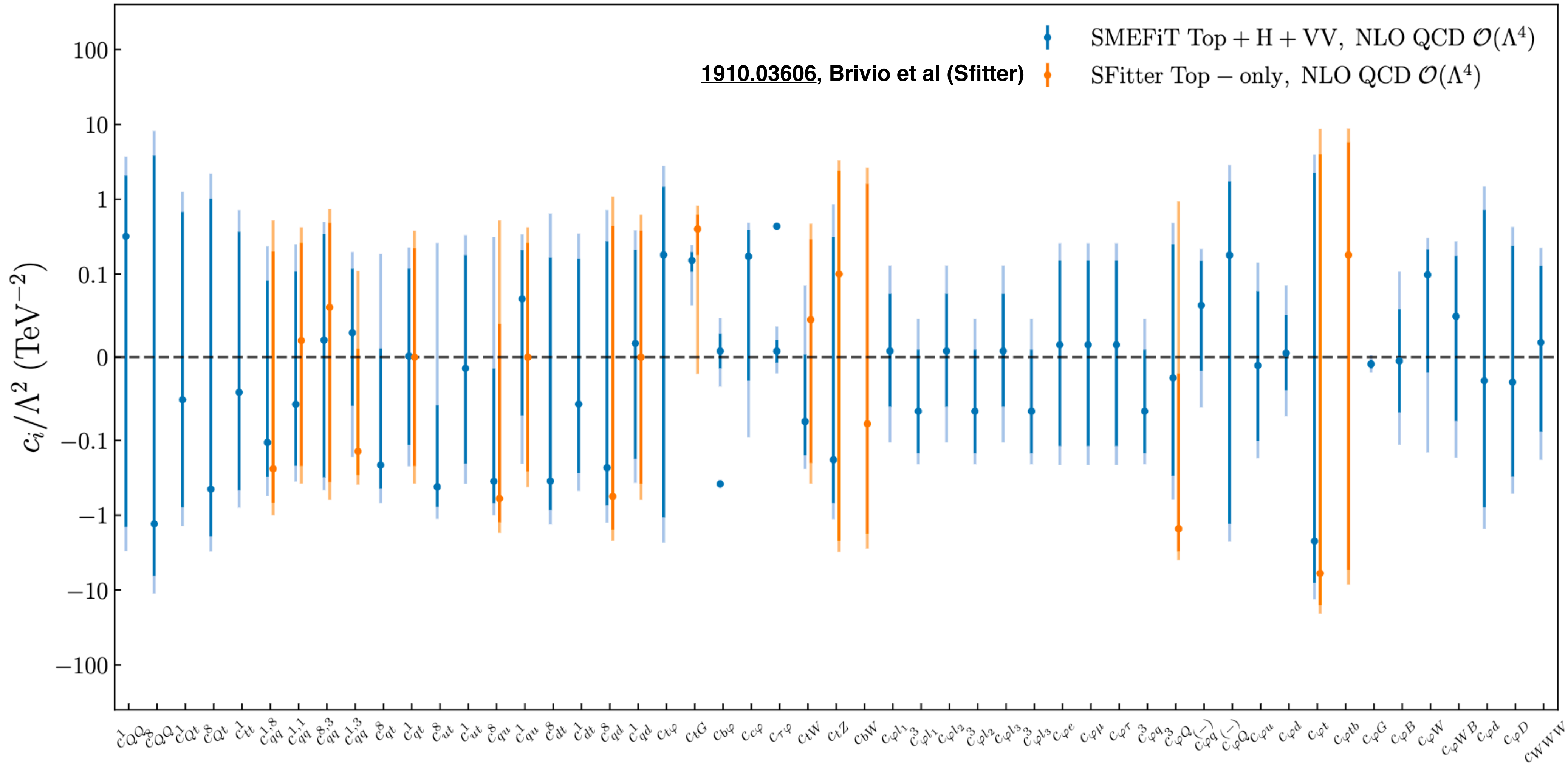
Results: dataset dependence





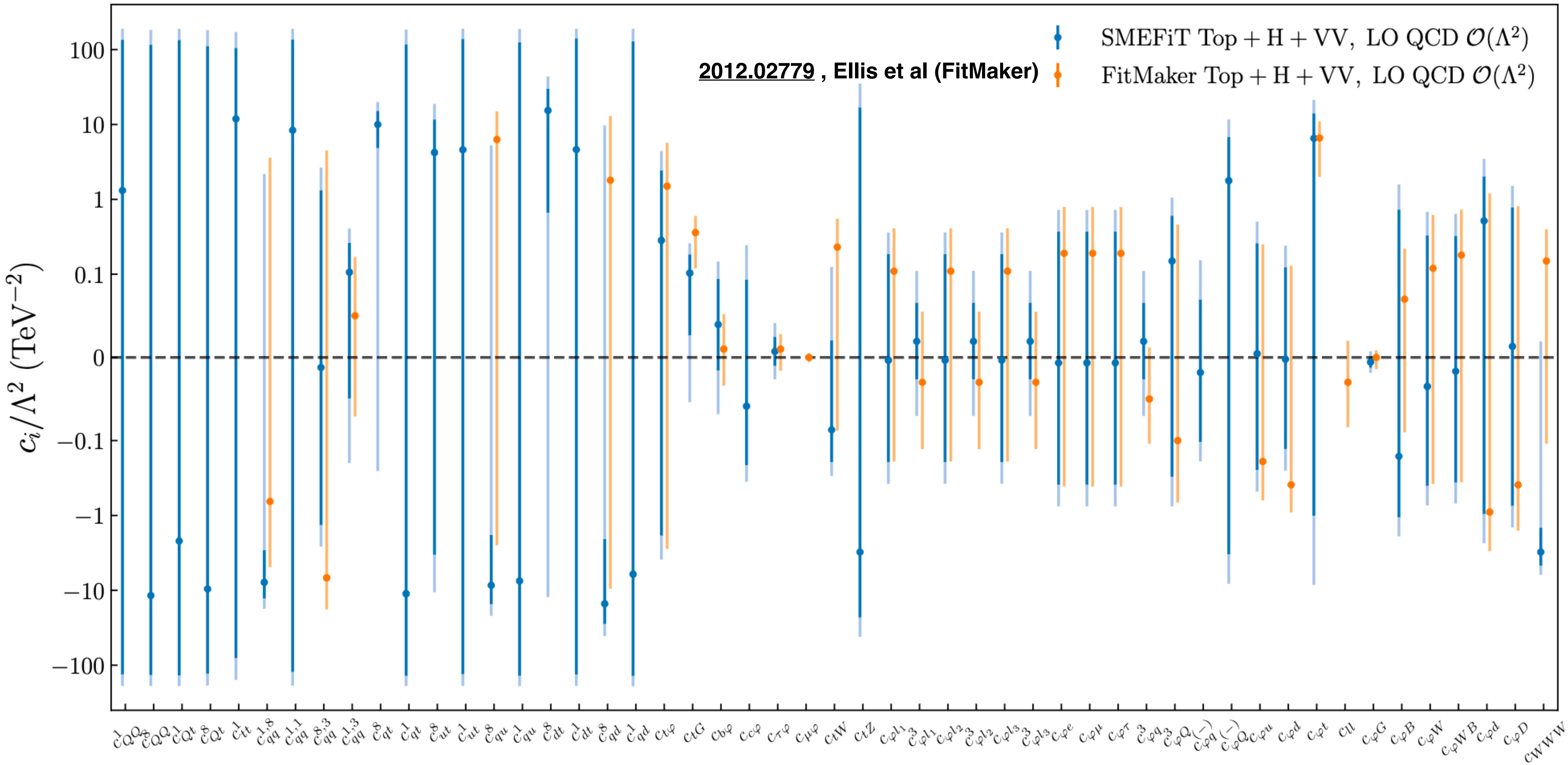
- Global fits consistent, but **more accurate**, with top-only or Higgs-only fit
- Top data **boosts the Higgs EFT fit** all across the board
- Diboson data only constraints **cWWW**
- Fit results stable upon **removal of high energy bins** (E > 1 TeV)

Comparison with SFitter (top-only)



global fit marginalised, 68% and 95% CL ranges (*not a tuned comparison*)

Comparison with FitMaker



Global (marginalised) fits, 68% and 95% CL ranges (*nb not a tuned comparison*)

Reasonable consistency but also noticeable differences: need **benchmark comparisons!**

ongoing efforts in LHC EFT WG

Summary and outlook

- **SMEFiT** is a novel framework to carry out global analyses of the SMEFT which exploits (but is independent from) ample expertise inherited from (NN)PDF fits

- Successfully deployed for various EFT interpretations, including a **global top+Higgs+diboson analysis,** EFT fits with Bayesian reweighting, and a dimension-six EFT analysis of VBS data

- Not discussed here: how to implement in the fit **UV-motivated theory constraints**, Bayesian inference for very fast EFT projections, **interplay with PDF fits**, treatment of theory uncertainties, matching to UV scenarios …

- Next steps in our program are the addition of **new LHC observables** (including flavour) and then that of **non-LHC processes** (low-energy, neutrinos, EDMs) as well as to keep improving the SM and EFT calculations used in the fit and ensuring a robust methodology that **scales to a fit involving hundreds of coefficients**

Why global SMEFT analyses?

- The **SMEFT is the new Standard Model,** once we assume that the SM is an effective description of Nature valid only up to some **cutoff energy *Λ***

- It provides a systematic, model-independent parametrisation of the low-energy deformations of a wide class of UV-complete BSM theories that reduce to the SM

- **Complete basis** at any given mass-dimension; **fully renormalizable**, full-fledged QFT: can compute higher orders in QCD and EW

- Exploits the full power of SM ``measurements'' for **model-independent BSM searches**

$$\mathscr{L}_{\rm SMEFT} = \mathscr{L}_{\rm SM} + \sum_{m=1}^{N_6} \frac{c_m}{\Lambda^2}\mathcal{O}_i^{(6)} + \sum_{n=1}^{N_8} \frac{b_j}{\Lambda^4}\mathcal{O}_i^{(8)} + \ldots$$

Fulfilling the potential of the SMEFT framework demands global analyses based on **a wide range of process** such that most (all?) **directions in the EFT parameter space** are covered

The SMEFiT framework

Theory

(N)NLO QCD + NLO EW for SM xsecs

NLO QCD, both linear and quadratic terms, **with SMEFT@NLO**

State-of-the-art **parton distributions** (avoid double counting)

Data

Higgs data (signal strengths, diff, STXS), diboson LEP and LHC, all available **top quark data** from Runs I+II, VBS, more in progress

Full experimental **correlations** included

SMEFiT

Validation

Extensive **statistical toolbox** to validate results: information geometry, PCA, closure testing, …

Full **posterior probabilities** in the EFT coefficients available, likelihoods WIP

Methodology

Two independent fitting methods, **MCfit** and **NestedSampling** (no reliance on linear approx) cross-check each other

Modular structure facilitates adding new datasets of better theory calculations

Fitting methodology

MCfit generate a large sample of **Monte Carlo replicas** to construct the **probability distribution** in the space of experimental data accounting for all uncertainties

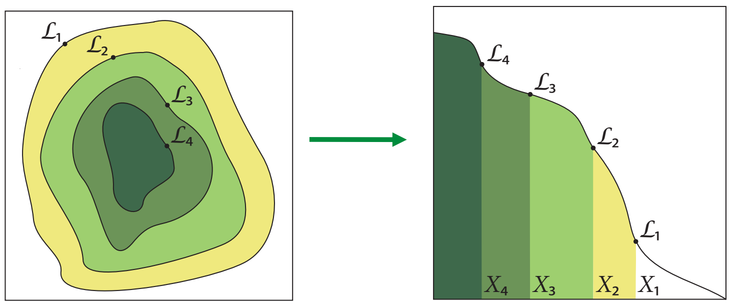
Determine the SMEFT coefficients **replica-by-replica** by minimising a cost function

$$E(\{c_l^{(k)}\}) \equiv \frac{1}{N_{\rm dat}} \sum_{i,j=1}^{N_{\rm dat}} \left(\mathcal{O}_i^{(\rm th)}\left(\{c_n^{(k)}\}\right) - \mathcal{O}_i^{(\rm art)(k)}\right)({\rm cov}^{-1})_{ij}\left(\mathcal{O}_j^{(\rm th)}\left(\{c_n^{(k)}\}\right) - \mathcal{O}_j^{(\rm art)(k)}\right)$$

where covariance matrix includes **all sources of experimental + theory errors**

Nested Sampling statistical mapping of the *N*-dimensional likelihood profile to 1D

$$Z = \int d^N c \mathscr{L}\left(\text{data} \,|\, \vec{c}\right) \pi(\vec{c}) = \int_0^1 dX \mathscr{L}(X)$$



- Samples directly from prior space to locate **regions of maximum likelihood**
- Main advantage: **no need for optimiser** (fitting)
- Exponential increase in runtime as prior volume increases