Combined SMEFT interpretation of Higgs, diboson and top quark data from the LHC

Higgs 2021, October 19th, 2021

Giacomo Magni

Nikhef Theory Group & VU Amsterdam

Motivation
Towards a SMEFT global fit

Present results based on:
SMEFiT Collaboration: *Combined SMEFT interpretation of Higgs, diboson, and top quark data from the LHC*
https://arxiv.org/abs/2105.00006

Goal: produce "*global*" SMEFT fit including Higgs, Top, VV data

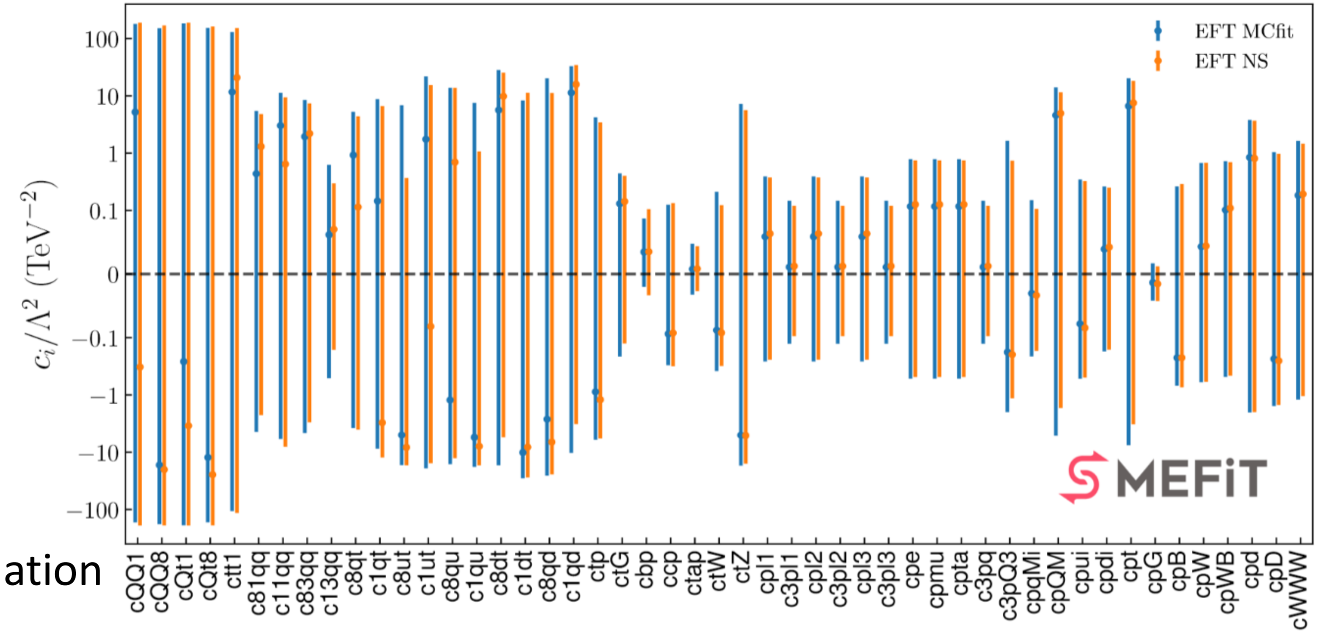
What do we fit:

- 317 experimental datapoints from LHC *(Run I, II and LEP)*
- Dim 6 EFT expanded including **Linear and Quadratic** contribution using *SMEFT@NLO*
- **36 independent dof** (14 related with EWPO) coming from the Warsaw basis

Motivation: a global SMEFT fit will tell us:

- If there is any deviation from the SM prediction in the LHC (i.e: where to look in the future)
- If the SMEFT is a correct interpretation of such deviations or we need a more complex EFT

Fitting Methodology
The SMEFiT framework

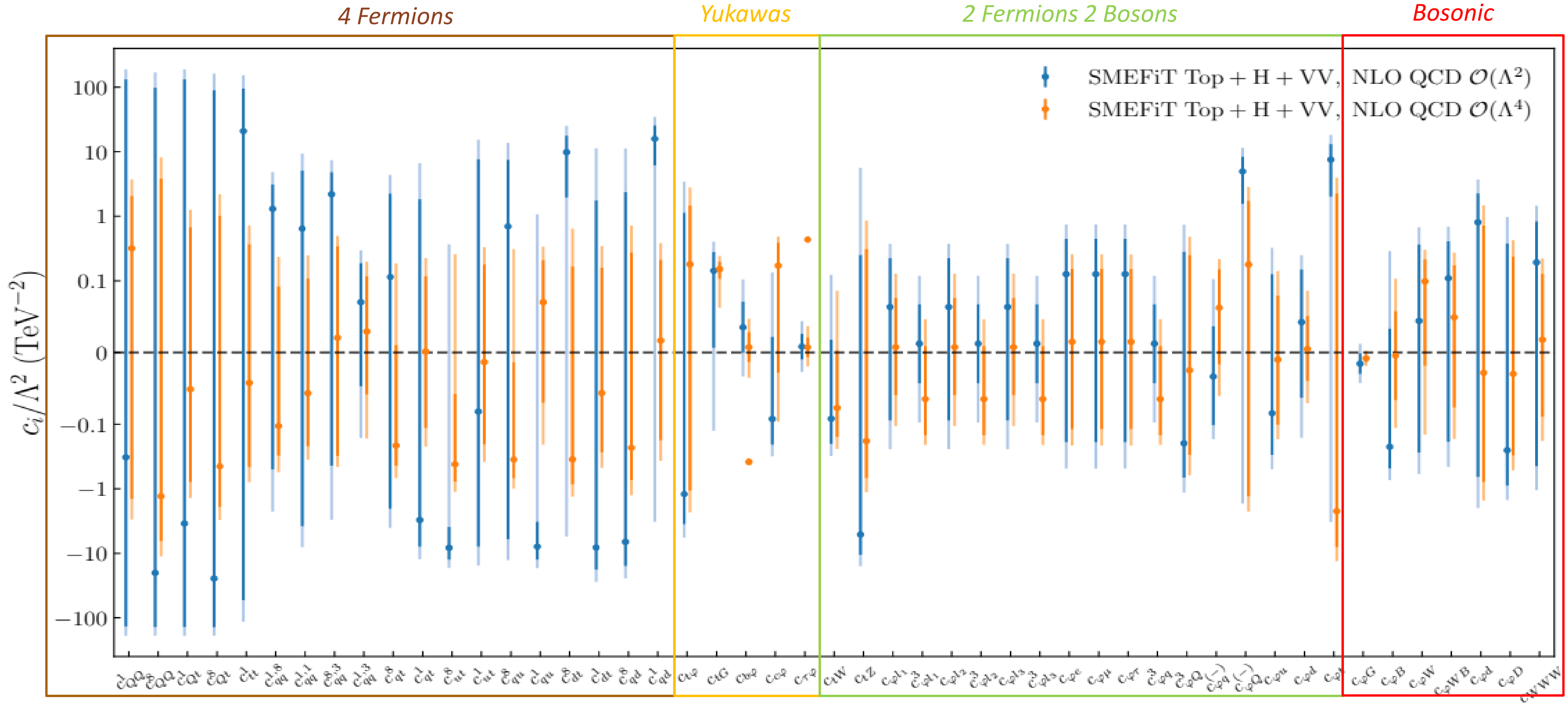


- Extensive **statistical toolbox** to validate results: information geometry, PCA, closure testing, …
- Two independent fitting methods: **McFit** *(frequentist approach)* and **Nested Sampling** *(Bayesian approach)* to cross-check each other
- Inclusion of the experimental **correlations** and **pdf uncertainties**
- Fitting also quadratic EFT contribution

↓

Full **posterior probabilities** for the EFT coefficients

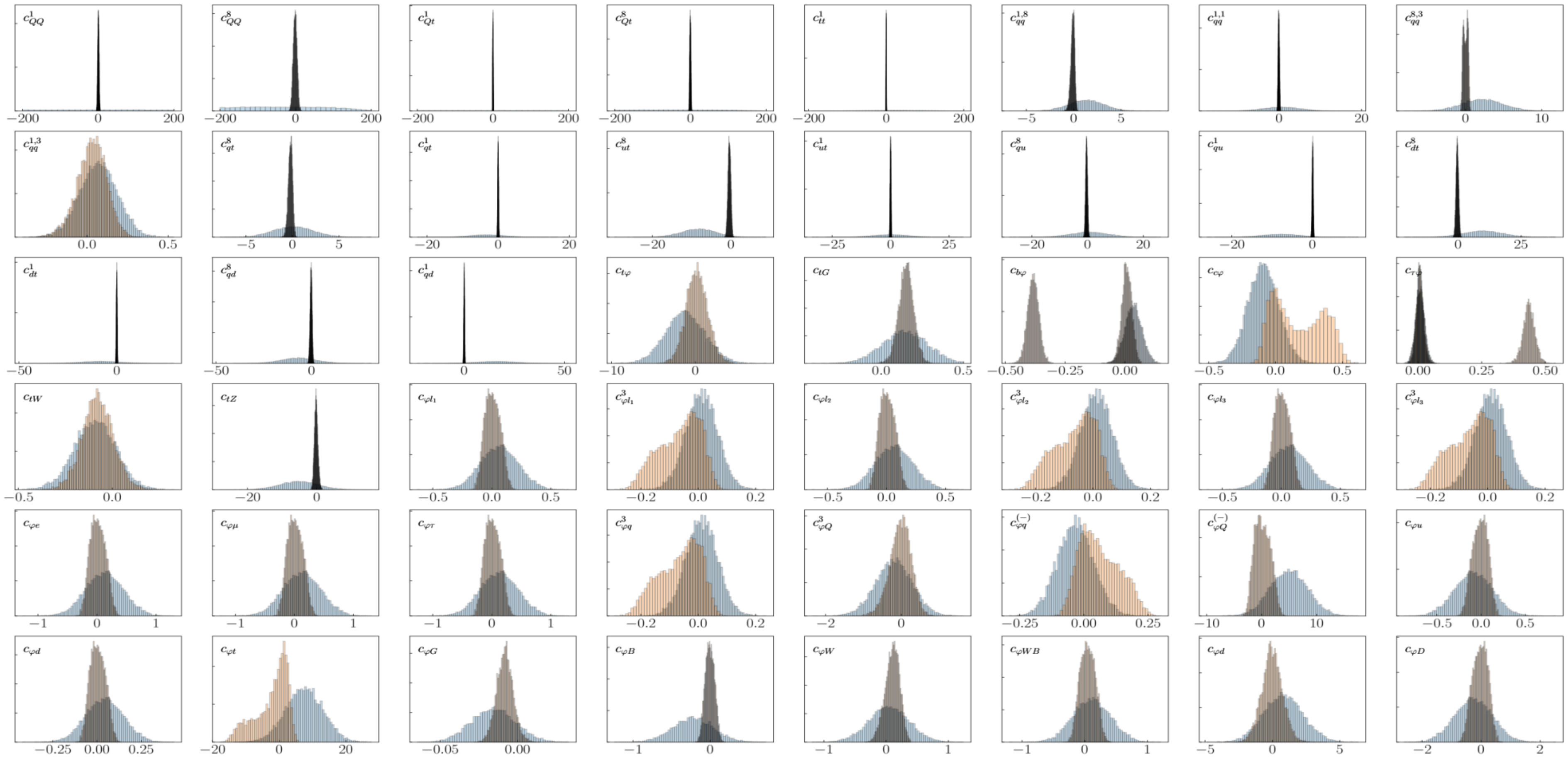
Linear vs Quadratic EFT fits

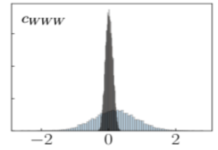


68% (95%) CL: Linear vs Quadratic

- All the op. are compatible with SM solution (95 % cl) except for c_{tG} *(quadratic)*
- Generally, quadratics give more stringent bounds bringing **new genuine information**

Linear vs Quadratic EFT fits





Posterior distributions: Linear vs Quadratic

- Posterior distribution are not more always Gaussian and double minima can appear

- Quadratic EFT **breaks flat directions** (see 4 Heavy Fermion)

Summary and Outlook

***SMEFiT** is a framework able to produce a comprehensive analysis of HEP data in the context of SMEFT*

Future plans:

- ✓ Keep expanding included datasets:
  - ✓ **new LHC observables** (DY, including flavour)
  - ✓ Also **non-LHC processes** (low-energy, neutrinos, EDMs)
- ✓ Keep stress-testing the **fit methodology** as it scales to a fit involving hundreds of coefficients
- ✓ Study of the **optimal observables** to maximise the EFT sensitivity
- ✓ **Benchmark comparisons** with other groups

- Not discussed here:
  - UV-motivated theory constraints and UV matching
  - Interplay with PDF fits *(here we used a PDF set without top data to avoid double counting)*
  - Treatment of theory uncertainties

Theory settings
SMEFT

Bottom-up approach: add higher dimension operator to the SM Lagrangian:

$$\mathcal{L}_{SMEFT} = \mathcal{L}_{SM} + \sum_{i=1}^{N_6} \frac{c_i}{\Lambda_i^2} O_i^{(6)} + \sum_{j=1}^{N_8} \frac{c_j}{\Lambda_j^2} O_j^{(8)}$$

dim 6 EFT from ***Warsaw Basis***

dim 8 EFT, not considered here

Observables gets modified as:

Best SM computed at N(N)LO QCD

Interference *terms and* ***EFT quadratics****,*
Computed with SMEFT@NLO

$$\sigma_{SMEFT} = \sigma_{SM} \times [\, 1 + \sum_{i=1}^{N_6} \frac{c_i}{\Lambda^2} \sigma_i^{(SM,6)} + \sum_{i,j=1}^{N_6} \frac{c_i c_j}{\Lambda^4} \sigma_{i,j}^{(6,6)}]$$

Theory settings

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 Warsaw operators** modifying Higgs, dibosons and top quark measurements: **36** (14) **independent** (dependent)
- Flavour assumption is **MFV**, with in quark sector with special treatment of role for top and bottom quark
- **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

317 experimental datapoints, with fiducial and differential cross sections, Top charge Asymmetries, STXS ….

from:

- LHC Run I
- LHC Run II
- LEP-2 (Diboson production)

Category	Processes	$n_{\rm dat}$
Top quark production	$t\bar{t}$ (inclusive)	94
	$t\bar{t}Z$, $t\bar{t}W$	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	40
	LHC	30
	Total	**70**
Baseline dataset	**Total**	**317**

Chi2 values after fitting: good reconstruction of the experimental data

Dataset	$n_{\rm dat}$	$\chi^2_{\rm SM}$	$\chi^2_{\rm EFT}$ $\mathcal{O}(\Lambda^{-2})$	$\chi^2_{\rm EFT}$ $\mathcal{O}(\Lambda^{-4})$
Total	**317**	**1.05**	**0.98**	**1.04**

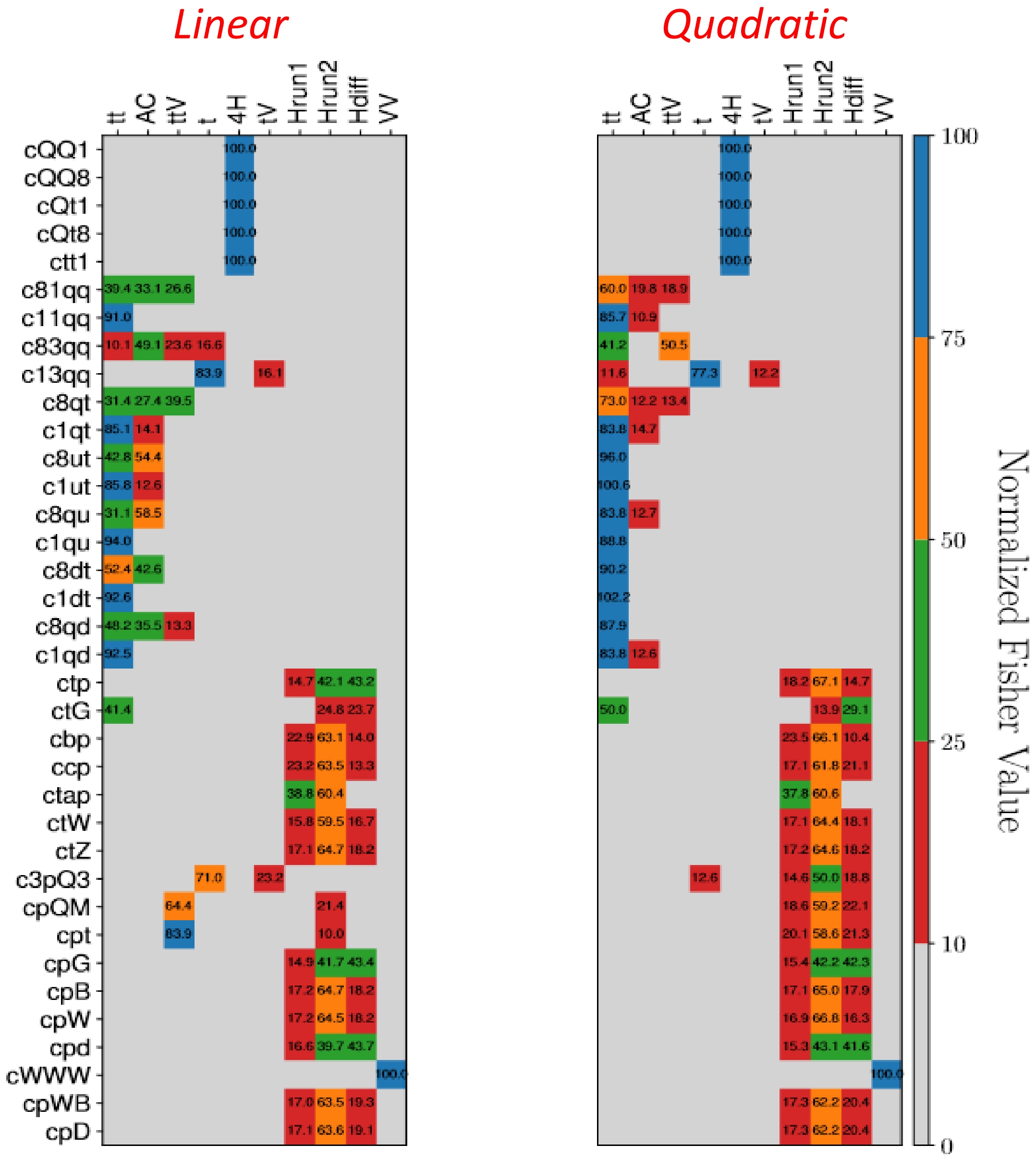
Fitting Methodology

Fisher Information and EFT sensitivity

- **Fisher information** eigenvalues quantifies how much each operator contributes to a process (experimental dataset)
- At linear level can be computed before fitting
- Can be evaluated with and without NLO EFT corrections

We observe:

- 4 Fermion operators are constrained mainly with top data
- 2 Fermion 2 Boson and purely bosonic by Higgs data



Fitting Methodology

The SMEFiT framework

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

- Define the figure of merit to minimize:

$$\chi^2(c_k) = \frac{1}{N_{data}} \sum (O_{exp,i} - O_{th,i})(cov^{-1})_{ij} O_{exp,j} - O_{th,j})$$

Covariance matrix including experimental and pdf uncertainties

Two complementary and equivalent fitting strategies:

- Bayesian reweighting (**Nested Sampling**):

$$p(c_k \,|\, data) = \frac{1}{Z} \mathscr{L}(data \,|\, c_k) \Pi(c_k)$$

Multi Gaussian likelihood *Flat prior*

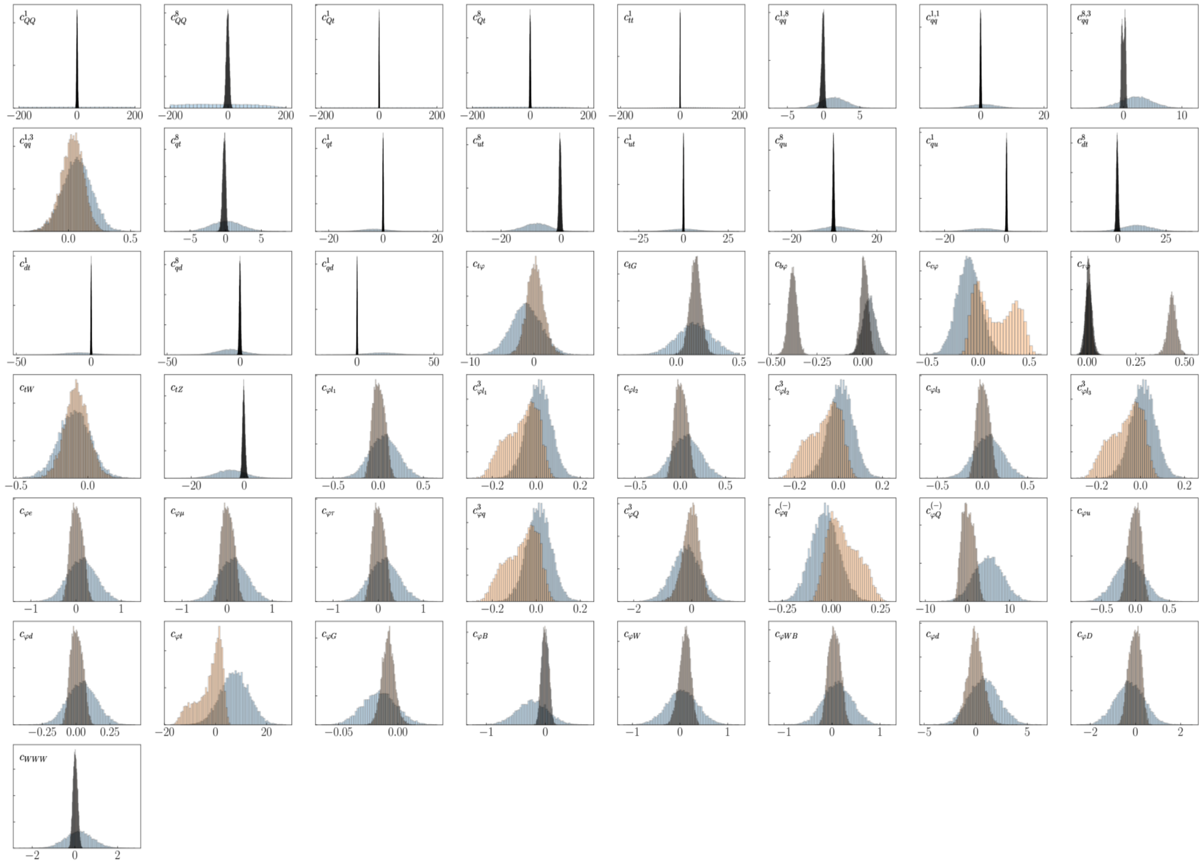
- **Monte Carlo replicas**: build many artificial replicas and determine the coefficients minimizing a cost function, replica by replica.

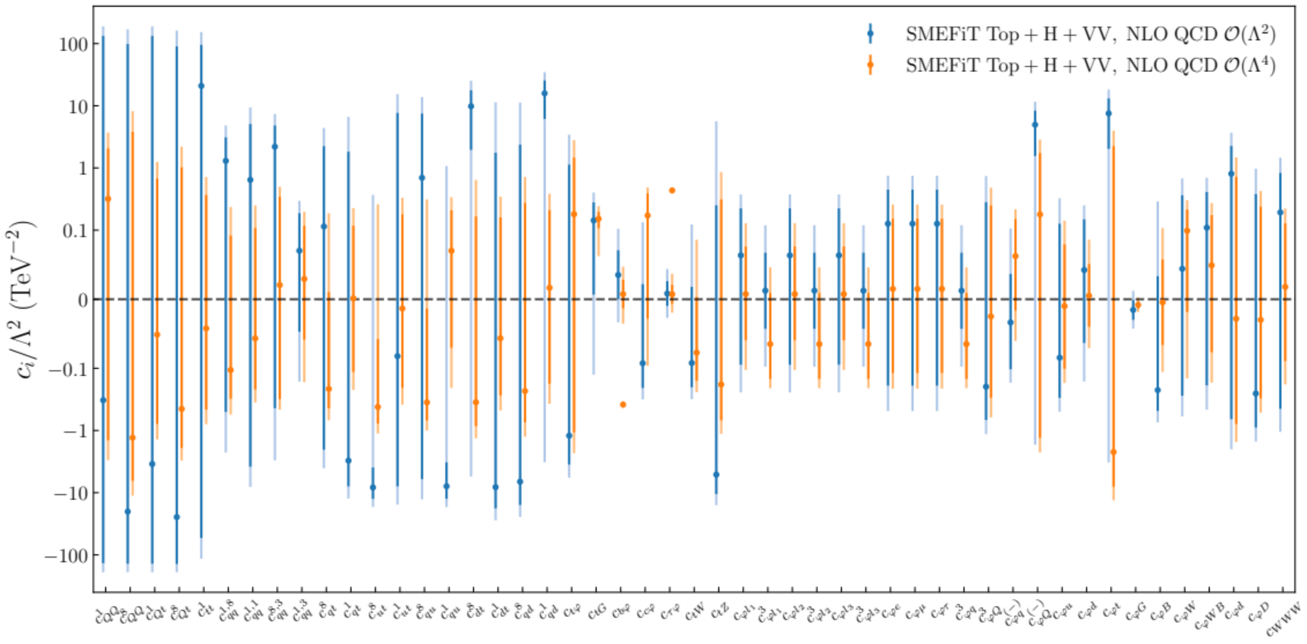
$$E(\{c_l^{(k)}\}) \equiv \frac{1}{N_{\rm dat}} \sum_{i,j=1}^{N_{\rm dat}} \left(\mathcal{O}_i^{(\rm th)}\left(\{c_n^{(k)}\}\right) - \mathcal{O}_i^{(\rm art)(k)} \right) ({\rm cov}^{-1})_{ij} \left(\mathcal{O}_j^{(\rm th)}\left(\{c_n^{(k)}\}\right) - \mathcal{O}_j^{(\rm art)(k)} \right)$$

Combined fit

Linear vs Quadratic EFT

Posterior distributions: Linear vs Quadratic





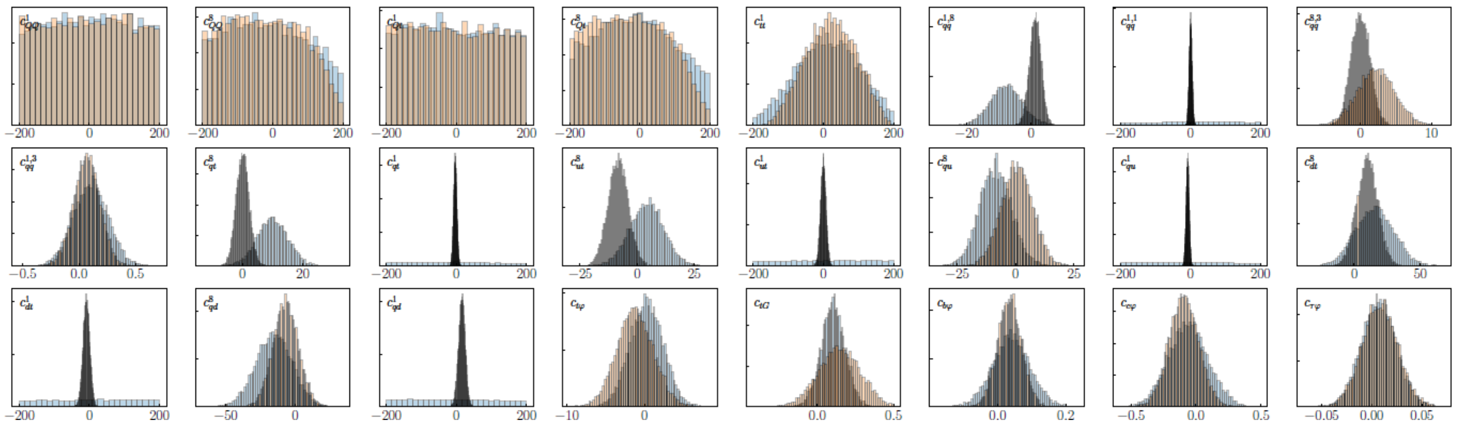
68% (95%) CL: Linear vs Quadratic

- All the op. are compatible with SM solution (95 % cl) except for c_{tG}
- Quadratic EFT **breaks flat directions** (see 4 Heavy Fermion)
- Generally, quadratics give more stringent bounds bringing **new genuine information**
- Posterior distribution are not more always Gaussian and double minima can appear

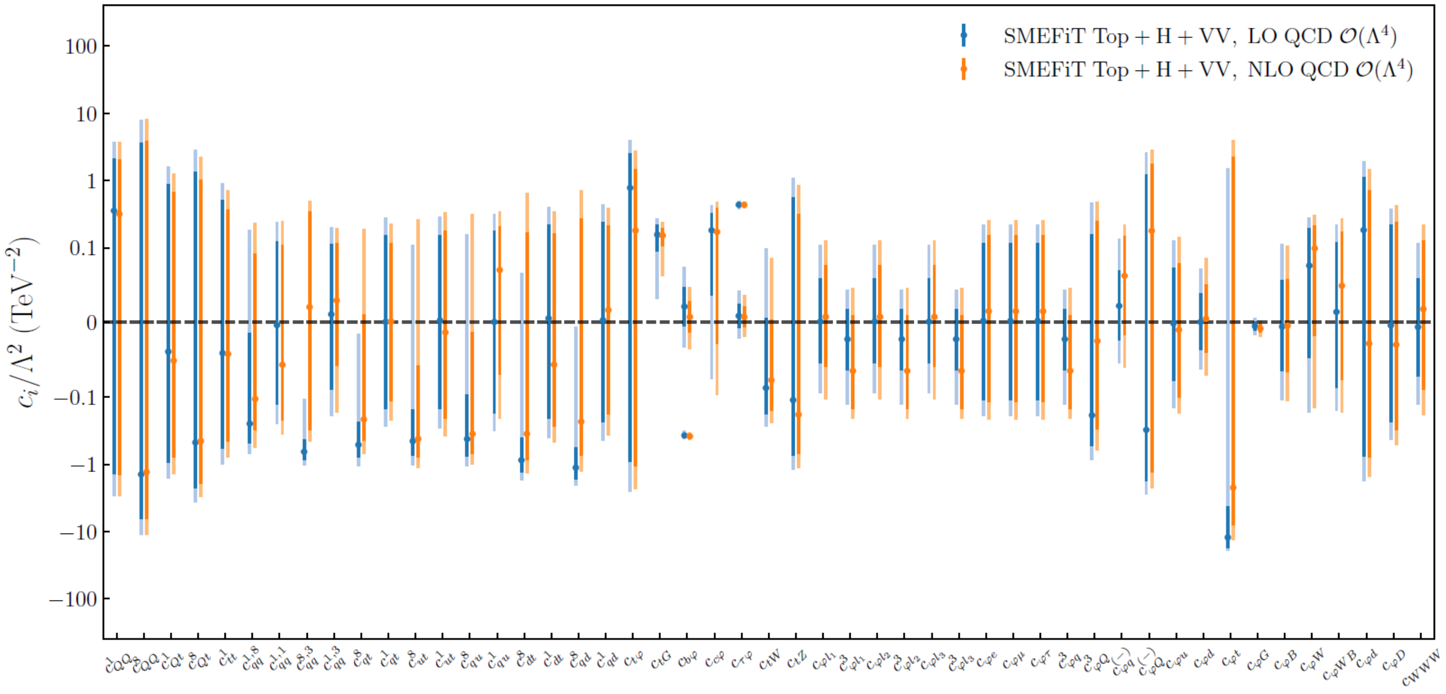
Combined fit

Effect of NLO corrections

Posterior distributions: Linear fit



68% (95%) CL: Quadratic fit

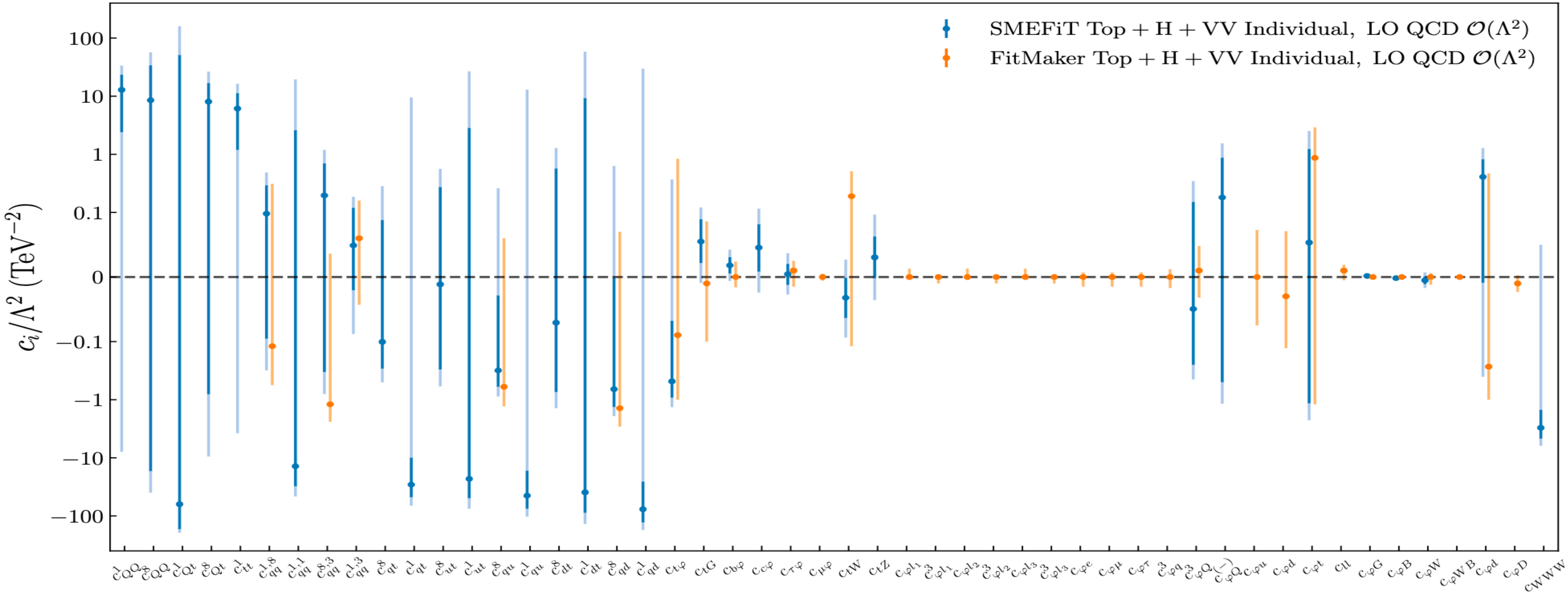


- NLO corrections break degeneracy on 4 Fermions op
- Effects are more evident for linear fit
- Nontrivial shifts both for central values and confidence levels
- NLO EFT corrections are needed: fits are more stable

Comparison with similar studies

Linear EFT

FitMaker: *J.Ellis et al.* *(see K. Mimasu talk)*

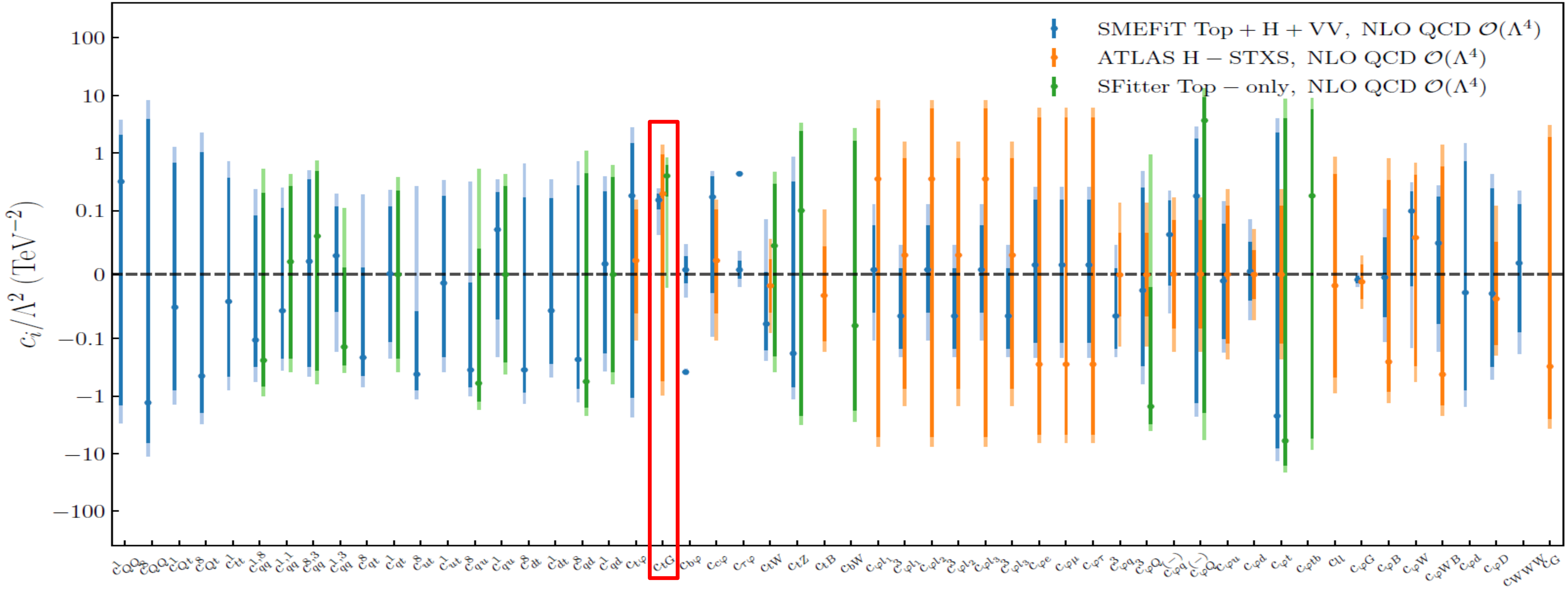


95 % Individual CL: reasonable agreement, but **not a fine tuned benchmark comparison**

Comparison with similar studies

Quadratic EFT

- ***SFitter***: I.Brivio et al.
- ***ATLAS – CONF note***: ATAS-CONF-2020-053
  (see , C. D. Burgard talk)



68% (95 %) CL: reasonable agreement, but **not fine tuned benchmark comparison!!**

Summary and Outlook

***SMEFiT** is a framework able to produce a comprehensive analysis of HEP data in the context of SMEFT*

Future plans:

- ✓ Keep expanding included datasets:
  - ✓ **new LHC observables** (DY, including flavour)
  - ✓ Also **non-LHC processes** (low-energy, neutrinos, EDMs)
- ✓ Keep stress-testing the **fit methodology** as it scales to a fit involving hundreds of coefficients
- ✓ Study of the **optimal observables** to maximise the EFT sensitivity *(see J.Ter Hoeve talk)*
- ✓ **Benchmark comparisons** with other groups *(see K. Mimasu, C. D. Burgard and E. Geoffray talks)*

- Not discussed here:
  - UV-motivated theory constraints and UV matching
  - Interplay with PDF fits *(here we used a PDF set without top data to avoid double counting)*
  - Treatment of theory uncertainties

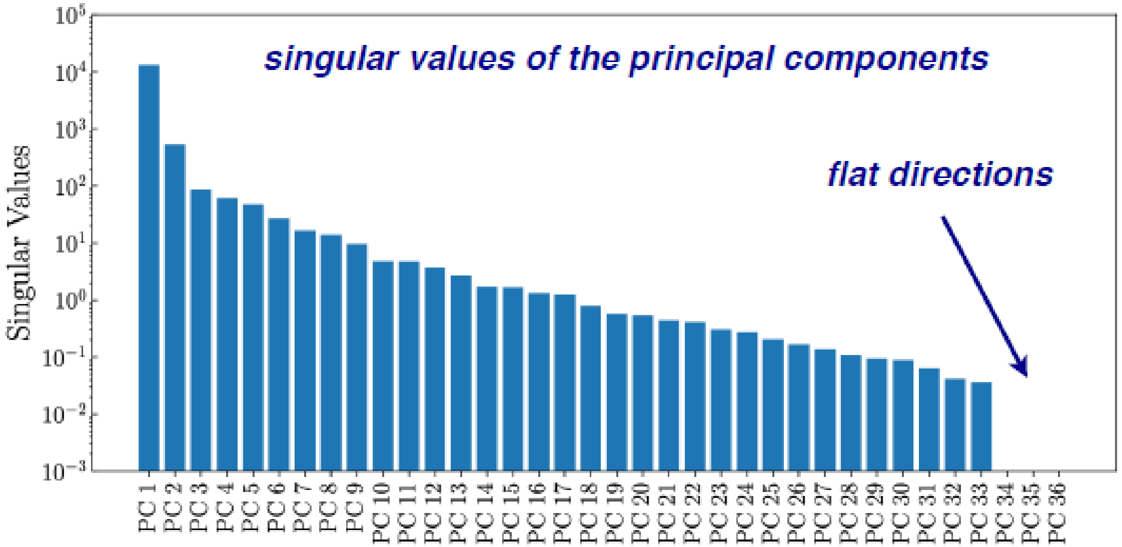
Thank you for the attention!

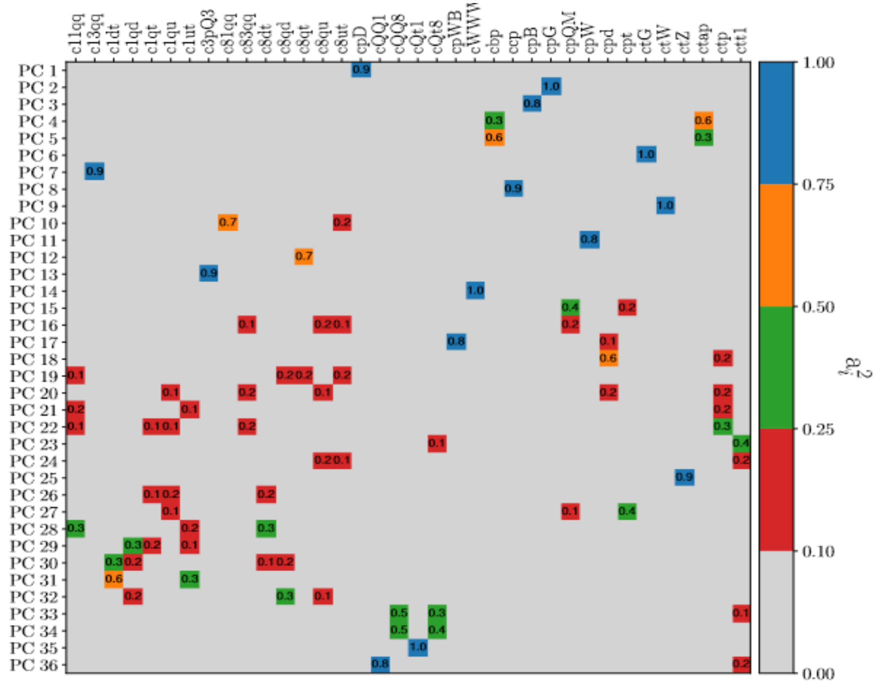
Fitting Methodology
Principal Component Analysis

- **PCA** is useful to identify and spot flat and “data driven” directions.
- Gives the best linear combination of coefficient to fit at linear level

We observe:

- 4 Heavy Fermions operators are basically unconstrained at linear level





Combined fit
Dataset Variations

- EWPO constrain are fixed in these fits.
- Global fit is stable upon dataset variations
- Top data improve bounds on bosonic operators, coupled with 3rd generation fermions.
- Dibosons data mainly relevant for c_{WWW}